

DYNAMIC SURFACE COMPLETION:
THE JOINT FORMATION OF COLOR,
TEXTURE, AND SHAPE

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität
zu Kiel

vorgelegt von

Daniel Wollschläger

Kiel, Februar 2006

Referent: Prof. Dr. R. Mausfeld

Korreferent: PD Dr. Johannes Andres

Tag der Mündlichen Prüfung: 26.01.2006

Zum Druck genehmigt: Kiel, den 26.01.2006

Prof. Dr. J. Grotemeyer, Dekan

“To be is to be the value of a bound variable”

– WILLARD VAN ORMAN QUINE

Contents

1	Introduction	9
I	Theoretical background	13
2	Input fragmentation	15
2.1	Optical sampling	16
2.2	Occlusion	17
2.3	Camouflage	20
2.4	Physical transparency and translucency	22
2.5	Motion	22
3	Visual completion across space and time	25
3.1	Structuring completion phenomena	26
3.2	Static completion	28
3.2.1	Modal completion	29
3.2.2	Amodal completion	54
3.2.3	Phenomenal interactions involving statically completed features	63
3.3	Dynamic completion	69
3.3.1	Boundaries and surfaces from spatiotemporal discontinuities	71
3.3.2	Spatiotemporal relatability theory	80
3.4	Preliminary synopsis	84
II	Experiments	87
4	Dynamic color spreading	89
4.1	Aperture viewing	90
4.2	Color from motion	92
4.3	Flank transparency	94
4.3.1	Stimulus construction and phenomenology	94
4.3.2	Experiment 1: Velocity of moving lines	97
4.3.3	Experiment 2: Modulating gap size	102
4.3.4	Experiment 3: Spatial range of the color-spreading effect	105
4.3.5	Additional observations	108
4.3.6	Discussion	109

5	Dynamic texture spreading	113
5.1	Stimulus construction and phenomenology	113
5.2	Visual persistence	115
5.3	Experiment 4: Distance-speed tradeoff	120
5.4	The timecourse of visual completion	126
5.4.1	General visual processing	127
5.4.2	Completion times	129
5.5	Methodological issues and the functional role of illusory percepts	130
5.5.1	Qualitative demonstrations	132
5.5.2	Quantitative methods	135
5.5.3	Limits and problems of quantitative methods	137
5.6	Experiment 5: Discrimination performance	138
5.7	Experiment 6: Gabor-frequency matching	147
5.8	Discussion	151
6	General Discussion	155
6.1	Perceiving completed surfaces, objects, and events	155
6.2	Summary	159
6.3	Conclusions and outlook	165
6.4	Acknowledgements	167
	References	169
	Appendices	185
A	Stimuli and adaptive algorithm used in experiment 5	185
A.1	Algorithm development	185
A.2	Stimulus construction	197
B	Tables	201
B.1	Experiments 1, 2 and 3	201
B.2	Experiments 4a and b	202
B.3	Experiment 5	204
B.4	Experiment 6	211
C	Abstract	213
D	Zusammenfassung	215

Chapter 1

Introduction

Visual perception relies on the physico-geometric characteristics of the light that impinges on the photoreceptor arrays of the retinae. Light that is reflected or emitted from objects in the environment passes through different media and interacts with many objects in the scene before it finally reaches our eyes. As a consequence, the wave fronts coming from a particular object have undergone severe alterations once they arrive on the retinae. One of the most drastic changes happens when the light array hits an opaque obstacle which is positioned closer to the eyes, resulting in a partial loss of the information carried by the light about the originating object.

Consider, e.g., the scene depicted in figure 1.1. Different parts of the geometric objects are visible from different vantage points, but regions corresponding to a particular solid are always interrupted by regions from occluding obstacles. It is because of this process of occlusion that in ordinary environments, physically continuous objects are often projected onto disconnected regions on the retinae, interrupted and altered by light coming from different parts of the scene. Hence, spatially separated image fragments that distally correspond to the same object must not be interpreted as isolated regions. Instead, these fragments should be grouped to establish their belongingness. Then again, there are of course many disconnected image regions which must not be grouped because they really do originate from separate surfaces.

The characteristics of the visual input get even more complicated when we take into account the temporal aspects of our surrounding: Human observers are mobile agents, actively exploring the environment with their perceptual systems. The sampling of the incoming light array is thus dynamic and takes place from constantly changing vantage points: We incessantly move our eyes, head, and body, thus causing the scene's projection to vary considerably over space and time. Since many objects also tend to change location over time independently from ourselves, the spatial relations between the observer and the scene, and between different objects in the scene can change substantially from one moment to the next. Most importantly, the combination of motion and occlusion yields an always changing pattern of what is visible and what is hidden in the retinal projection.

Our conscious perceptual world, however, reflects little of this—we are seldom aware of any loss of information about the external world, although the surfaces of many distal

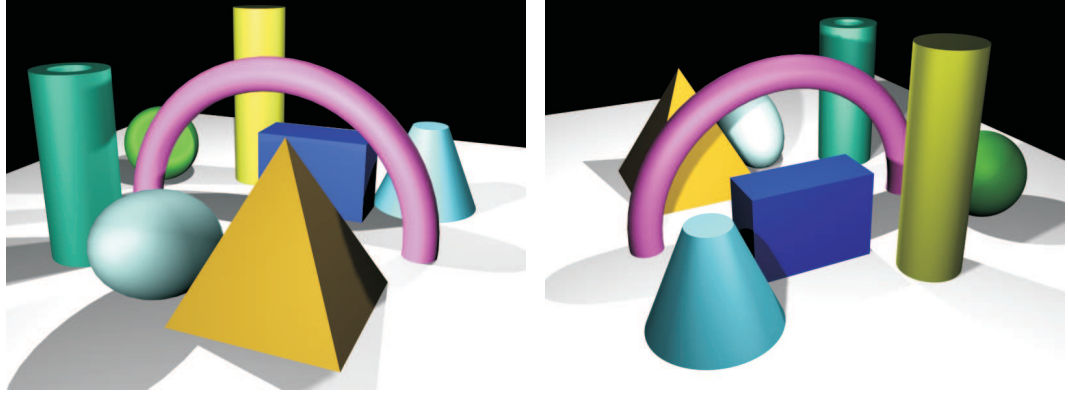


Figure 1.1: Two views of the same scene. Different parts of an object are visible to the eye depending on the vantage point, and no object can be seen in its entirety.

objects have no correspondence to areas in the retinal image. Furthermore, we typically have no trouble in perceiving, identifying and manipulating objects, swiftly recognizing their contiguous three-dimensional shape and maintaining their unity and identity over time. Our perception is characterized by a subjective completeness that pervades on many different levels of analysis: The perceptual world seems to be complete, stable over time, populated by continuous objects, and fully detailed or “sharp” in all directions. In other words, while the input to the visual system is limited in scope, of high temporal variation, and fragmented with respect to regions corresponding to light emitted from the same object, our perception seems to be complete, stable and continuous.

Since our interaction with the world normally is a successful one, it is obvious that the visual system does not deal with incomplete input information in random or arbitrary ways. Instead, the visual system succeeds in obtaining functionally relevant scene, object, and surface structure from physically disconnected image regions. This process of perceiving fragmented image parts as originating from the same object, of identifying corresponding stimulus regions, grouping and perceptually connecting them, is commonly referred to as visual completion or filling-in. From the perspective of vision science, visual completion has proven quite difficult to understand. The present work theoretically and empirically addresses some characteristics of the purported mechanisms underlying the perceptual completion of a spatiotemporally fragmented visual input.

Unfortunately, just as the visual input is highly fragmented, so does the scientific account of visual completion suffer from a high degree of partitioning: Owing to a predominantly neurophysiological perspective on perception, many facets of visual processing are commonly treated separately with little cross-talk between the insular disciplines,

among them motion analysis, color vision, shape perception, or low-level recognition of boundaries. It is evident that the research on spatiotemporal integration touches many of these aspects simultaneously, thus requiring an integrative approach. However, each of the aforementioned aspects is associated with its own historically-developed theoretical approaches, more or less idiosyncratic terminology, experimental paradigms, and literature domains. Part of the difficulty in studying perceptual completion thus is that it draws from several different fields, implying the need to accumulate and reconcile evidence that was collected within different research contexts. Among the relevant perceptual aspects, boundary integration has by far received the most scrutiny ever since completion was on the agenda of vision scientists. More recently however, aspects of motion processing and surface representation have gained in attention. I will try to shed light on completion from different angles, discussing, among others, boundary completion as a necessary basis for the understanding of dynamic surface formation, which is at the heart of this work.

Partly because the problem of understanding visual completion is separately tackled from many different sides, no established coherent picture or widely agreed-upon basic models have surfaced so far. This makes it difficult to single out a well-defined sub-problem, place it within an accepted theoretical frame, and straightforwardly treat it in due course. Owing to this situation, the following work itself will not follow a perfectly straight path from presenting the relevant theoretical background to deriving a set of clear-cut research questions to their experimental treatment and, finally, the answer to the question posed. Rather, it will meander through the landscape that defines the broad problem of stimulus integration: I will first give an integrative review of the general theoretical background to then address the more concrete theoretical questions in close proximity to the corresponding experiments. The presented work should thus not be conceived of as a systematically-crafted puzzle piece in a concerted effort to unravel visual completion, but rather as the attempt to create order in the intermediary picture that is formed by ongoing efforts in the study of filling-in.

To this end, I will, in chapter 2, give a more structured account of the causes for fragmentation in the input to the visual system before turning to aspects of perceptual completion phenomena themselves in chapter 3. The treatment of completion progresses from instances apparent in static stimuli to those found in the perception of dynamic scenes. While it is at first useful to order completion phenomena according to what feature is actually completed (e.g., boundaries vs. surface qualities like color or texture), it will become apparent that complex interactions between different features are more the rule than an exception in most cases of completion. Accordingly, special emphasis will be placed on these multi-faceted interactions in section 3.2.3, but also throughout the treatise in general. Different theoretical approaches to perceptual completion will be encountered throughout, among them image-based geometrical accounts, models of quasi-cognitive perceptual inference, and also ideas about the neural correlates of perceptual filling-in.

With respect to empirical work, chapter 4 will be concerned with establishing the

Chapter 1 Introduction

general characteristics of a novel multi-aperture viewing phenomenon called flank transparency, and to place it within the larger context of already-known effects. Results from the respective experiments will then motivate a discussion about perceptual mechanisms possibly responsible for them in chapter 5, focusing especially on the role of visual persistence as outlined in section 5.2. Sometimes dubbed “retinal painting”, sensory persistence has played a pivotal role in explaining dynamic completion phenomena and will serve as the starting point for the experiments presented here. The affirmative results regarding its relevance in turn prompt, in section 5.4, for the closer examination of another quantitative aspect of filling-in—the timecourse of perceptual completion’s microgenesis. This examination will reveal important methodological issues associated with the attempt to study visual completion in objective ways, a topic that section 5.5 is devoted to. The two subsequent experiments try to incorporate some of the emerging conclusions in the design of their methods, and generally address the question how it can be shown that visual completion provides functionally relevant information about partially invisible stimuli.

In the concluding discussion in chapter 6, I then seek to shortly place the presented completion phenomena within a larger framework of object and event perception. The chapter will end with some brief remarks on philosophical as well as conceptual issues in vision science as spurred by the examination of perceptual completion.

Part I

Theoretical background

Chapter 2

Input fragmentation

As has been alluded to in the introduction, the light array stemming from an object in a natural scene undergoes several substantial changes before it reaches the eye. Processes such as occlusion often block out light coming from a particular region of the object and thus prevent it from reaching the eye, thus leading to disconnected areas in the retinal projection that all belong together as they originated from the same object.

In order to illustrate the problem of stimulus fragmentation under natural viewing conditions, figure 2.1 shows three frames of a short animation, representing momentary views of a moving observer. The frames make visible the high degree of instability and fragmentation in space and time which governs the input available at the eyes. At each fixed point in time as well as over different moments, image regions belonging to the same object are actually separated, often multiple times, by regions belonging to several different objects. Therefore, interrupted pattern fragments not only have to be linked to the correct fragments in the same frame, but also to regions in neighboring time-intervals stemming from the same object: A fragment of the car in frame 2 may correspond to another fragment seen through an aperture in frame 2, may have been hidden in frame 1 and reappear in frame 3. The several apertures that render visible parts of the car are themselves in motion, making it hard to predict the pattern of covering and uncovering of the bypassing car.

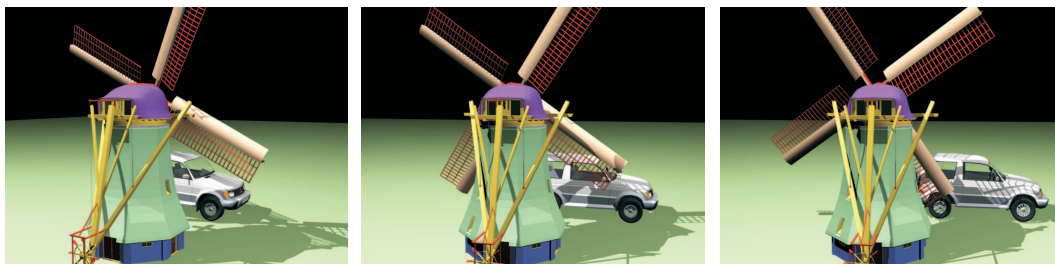


Figure 2.1: Three frames from an animation showing the spatiotemporal fragmentation of objects under natural viewing conditions.

To disentangle the different sources of stimulus fragmentation in dynamic scene perception, a systematic outline of the relevant psychophysical influences operative in the input generation process is presented below.

2.1 Optical sampling

Before turning to sources of regional discontinuities in the stimulus that result from the physical characteristics of the distal scene, it should be remembered that the optical sampling done by the eye is itself a source of fragmentation: In each moment, light reaches our eyes only from a limited solid angle in the environment, thus excluding much of the potentially available visual information. The photoreceptor array which transduces the light's energy into a neural code is discrete, thus imposing a theoretical limit on the input's level of detail that can be resolved. The result is spatial frequency aliasing, which means that information present in some spatial frequencies is tampered with while other information is ultimately lost. Furthermore, the retinae are inhomogeneous with respect to the density and arrangement of the different photoreceptor types (Hofer, Carroll, Neitz, Neitz and Williams, 2005; Roorda and Williams, 1999), thus creating input variations that are not lawfully related to characteristics of the distal scene.

The blind spot represents an extreme case of the inhomogeneity of the photoreceptor distribution on the retina. Characterized by the complete absence of rods and cones, the blind spot does not transduce any light energy at all, resulting in a complete loss of the information available in that area of the retinal image. The clinical condition of a scotoma is similar to the situation at the blind spot in that portions of the retinal field are non-functional with respect to signal transduction. As with the case of the blind spot, the missing active neural signaling in scotomas also implies that no absence of information from the affected locations can be explicitly coded. This raises the question if blind spots and scotomas really create any stimulus discontinuity at all, and, equivalently, if the blind spot is represented as an area of missing information in further processing. As an analogy, one might think of the blind spot as the space between our fingers when touching a surface (Pessoa, Thompson and Noë, 1998). Even though the surface information coming from touch is clearly fragmented, one does not have the impression of touching isolated islands of material.

Two of the most palpable causes for the *temporal* loss of information are eye blinks and saccades. While eye blinks can be performed consciously, they are usually automatic and go unnoticed. We blink about eight times a minute, each time creating a temporal gap of about 50 ms (although blink duration varies widely depending on fatigue and other factors). During saccadic eye movements, conscious visual processing is actively suppressed, making us effectively blind for stimulus changes during a time period of 20–120 ms, depending on the size of the saccade (Rodieck, 1998). Taking eye blinks and saccadic suppression together, temporal gaps sum up to about one to two

hours of our daily wake time. Contrary to this conclusion, our subjective perceptual world is not constituted of discrete perceptual epochs but seems to be temporally continuous (van Rullen and Koch, 2003). Obviously, there have to be mechanisms which lead to a smooth temporal event perception based on a highly disruptive input.

2.2 Occlusion

In accordance with the depth relations within a scene, occlusion determines what is visible to and what is hidden from an observer. Put simply, occlusion has as its consequence that visible boundaries and shapes in the retinal image do not fit the distal object boundaries and shapes. As a separation between regions in the 2-D projection of a scene is often not diagnostic for a separation of their originating scene elements, grouping processes are of the utmost importance to vision, and in particular play a key role in visual completion. Consequently, already the Gestaltists tried hard to understand the mechanisms involved, and many current approaches may be seen as formalizations of their intuitions (cf. section 3.2.2). The fragmentation of the 2-D image has another far-reaching consequence for visual processing: It dramatically limits the use of 2-D representations, and implies the need to cast visual processing in 3-D space early on, because only here can object fragmentation be overcome. Recent work indeed supports the idea that a volumetric representation is deeply rooted in visual processing (cf. section 3.2.2).

It is important to distinguish early on some qualitatively different occlusion situations because, as will become evident later, they lead to different regularities in the projected image. The first and most widely studied case of occlusion is the partial interposition of two opaque objects along the line of sight (figure 2.2a). Here, a depth discontinuity in the scene corresponds to the projected boundary that separates the two objects. The same, however, is not true for the case of two interpenetrating objects (figure 2.2b), where there is only a curvature discontinuity at the contour of penetration in the scene, but no depth discontinuity.¹ Likewise, in situations of conformation or wrapping (figure 2.2c), no abrupt depth change between occluding and occluded object is extant in the scene.

An important geometric regularity of visual mapping concerns the image borders which separate two surfaces in the occlusion cases of interposition and penetration. Given that the corresponding surfaces in the scene do not, by chance, perfectly align, edges represent the boundary of only one of the two abutting sides, and in that sense can only belong to one surface. More specifically, the border must be intrinsic to the figure, i.e., the occluder (Rubin, 1915). Also called belongingness (Koffka, 1935), or “Gesetz der Innenseite” (Metzger, 1975), this border ownership (Nakayama and

¹Strictly speaking, for every image resulting from penetration, an interposition scene can be constructed which has the same projection. Said interposition situation would, however, require a non-generic viewpoint (see below) to be equivalent in this respect.

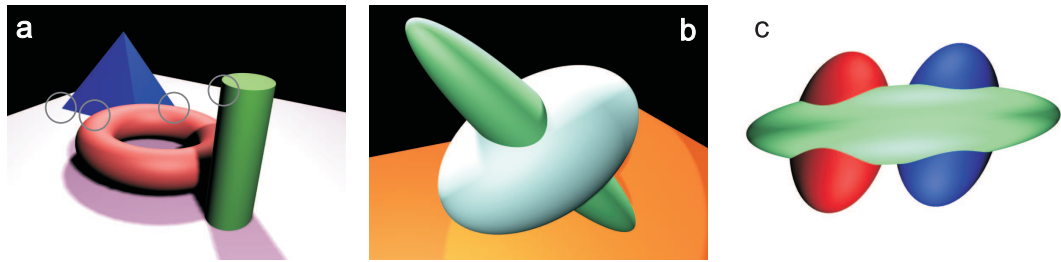


Figure 2.2: Different types of occlusion. a) Interposition of objects along the line of sight, causing L- and T-junctions (circled). b) Penetration. c) Conformation or wrapping of one object around the other. Note that no depth discontinuities in the scene are caused by the latter two cases.

Shimojo, 1990b), thus specifies which surface actually ends at said border and which one is not limited by it—and may thus continue behind the other one (Michotte, Thinès and Crabbé, 1964; Sampaio, 1943). Knowledge of border ownership now disambiguates the occlusion configuration as it must be the occluding one of the two abutting surfaces which owns the boundary, except for conformation situations (Tse, 1999).

From the geometric relations described above, it follows that cues to border ownership are an important source of information to the scene’s spatial layout. This conclusion is backed by a number of bi-stable figures, whose interpretation can switch drastically with respect to the represented perceptual units, their shape, and figure-ground assignment. Importantly, these reversals always entail a change in border ownership (Metzger, 1975).

For the large class of interposition-type occlusion, it can be shown that image-defined regularities concerning border-ownership can be derived from natural geometry (Hoffman and Richards, 1984; Koenderink, 1976): Instances of this occlusion class produce, when viewed from almost all vantage points, a special type of discontinuity in the projected contour where the edges from occluding and occluded object meet (figure 2.2a). These discontinuities represent a point of the continuous contour where the first-order derivative or tangent is not defined. For this reason, it is often called “first-order” or “tangent” discontinuity. Depending on the form of the contour, the projected junctions between the borders of occluding and occluded objects are called, e.g., T- or L-junctions, and occlusion junctions in general. Coming back to the concept of border ownership in interposition occlusion, it can be shown that it is the location of the tangent discontinuity which divides the contour into segments with different ownership, therefore, they are also called contour terminators. Even more specific information can be derived from the junction structure of an image: For example, in the case of T-junctions, it is most often the nearer, occluding surface which owns the “cap” of the T and the occluded surface that owns the stem.

Since the junction structure of an image is usually directly evident in the input,² it represents a potential low-level cue to scene properties like occlusion and depth stratification which are by themselves not immediately derivable from image properties. Numerous empirical studies indeed testify to the importance of junctions in disambiguating the visual input with respect to the spatial scene layout (see e.g., Anderson, 1997; Rubin, 2001). It is, however, important to remember that the regularities described above only hold if interposition occlusion and a so-called non-accidental viewpoint are given. The former limitation will play a major role when we consider competing models of occlusion detection and processing discussed in sections 3.2.1 and 3.2.2. Without going into the formal details of the latter assumption (see Albert, 2001; Albert and Hoffman, 1995; Hoffman, 1998 for an overview), its essence may be stated in the following way: For a given stimulus, there is often a small set of viewpoints which lead to projections of it that are topologically different from the projections from all other viewpoints (figure 2.3). If these special vantage points are, in a here unspecified way, “rare”, they are called accidental, as opposed to the “generic” other viewpoints. The special qualitative structure of projections from accidental viewpoints is immediately destroyed once small random perturbations in it occur, and give way to an image structure of a generic type. Linked to the distinction between stable generic and unstable accidental viewpoints is the “generic viewpoint assumption” which postulates a special importance of considering generic viewpoints in visual processing (Albert and Hoffman, 2000; Nakayama, He and Shimojo, 1995; Nakayama and Shimojo, 1990b).

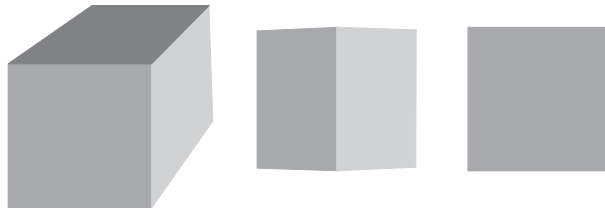


Figure 2.3: Three views of a cube. From a generic viewpoint, three sides are always visible. Accidental viewpoints change the topological characteristics of the projection by reducing the number of visible sides.

While occlusion is a major cause for fragmentation and loss of information in the incoming light array, its interaction with the depth ordering of a scene also introduces an asymmetry in the generation of the visual input which is both simple and rich in implications: Objects nearer to an observer can occlude objects that are farther away, but not vice versa. Since the depth structure within a scene must itself be determined indirectly on the basis of the retinal projections, identified occlusion relations can be exploited as an important qualitative source of information for pinpointing the three-

²A reliable distinction between first-order discontinuities and localized regions of very high curvature of course assumes infinite resolution in image sampling. For this reason, it can be expected that a perceptual concept of tangent discontinuity is continuous with a zone of transition between regions of high curvature and actual tangent discontinuities.

dimensional scene layout. More specifically, occlusion asserts that fragmented regions belonging to the same object must be located farther away from an observer than the occluding object.

Exactly what portion of the space behind an obstacle is hidden slightly differs between the two eyes because of their horizontal separation. Hence, there are image regions in the left eye that are—necessarily on their right side—bounded by regions coming from the occluding object, and have no correspondence in the right eye. As for the right eye, there are regions that are—now on their left side—bounded by areas coming from the occluding object, that are invisible to the left eye. This lawful visibility pattern called half occlusion or da Vinci stereopsis was found to be a depth cue in its own right (Nakyama and Shimojo, 1990a).

Finally, it should be noted that, although seldom consciously perceived, occlusion is not an exception but the rule in the projection of real world scenes. In any naturally complex scene, there are many objects along most lines of sight in different depth planes. Moreover, because of self-occlusion, in fact *no* object can ever be visible to the eyes in its entirety. Taking into account the regularities associated with occlusion formation is therefore of eminent importance in object recognition and scene perception.

2.3 Camouflage

Camouflage happens when there is no physical discontinuity between the retinal projection of a nearer surface and that of the background. One way for this situation to occur is when color and texture characteristics of the foreground object's surface are sufficiently similar to those of the background, thus destroying the delineating border between their respective projected regions. The particular importance of contours for shape identification is exploited in camouflage techniques commonly used by animals and humans alike: In addition to changing the surface properties themselves, they often focus on the deliberate alteration of an object's silhouette. These techniques often aim at multiply disrupting the usually smooth and convex silhouette with irregular protruding objects to create a zone of transition in which the object imperceptibly blends in with the background (figure 2.4).

Strictly speaking, camouflage is not responsible for breaking up any image regions stemming from the same object, but instead for exactly the opposite effect—the linking of stimulus areas which belong to different objects. Camouflage eliminates the formation of image contours between separate surfaces and thus illustrates the other side of the coin of stimulus fragmentation: On the one hand the need to group and ultimately join disconnected but corresponding regions, and on the other hand to also segregate those image regions not belonging together, even in the absence of physical discontinuities between them. In a general sense, camouflage therefore also contributes to the disparate relation between regions in the image and the layout of scene surfaces.



Figure 2.4: A so-called Ghillie suit used to camouflage a hunter in the woods. The suit mainly seeks to destroy the person’s characteristic silhouette.

Perfect camouflage in the environment may be rare, but several circumstances lead to an approximation of its consequence by considerably lowering the luminance contrast between foreground and background. This effect of reduced boundary salience may, e.g., be induced by darkness, fog, or rain (Parks, 1984).

Compared to occlusion, camouflage entails the reverse depth relation between hiding and hidden object: The structure that is rendered invisible is nearer to the observer, the hiding part of the scene more distant—camouflage means blending in with the *background*. With respect to the color and texture conditions that can give rise to camouflage, there is another important difference to occlusion: The environment characteristics that can lead to camouflage are much more restricted than those for occlusion. Occlusion relies on scene geometry only and is independent from the color composition of the near and far object, whereas camouflage crucially depends on it (Anderson, Singh and Fleming, 2002).

2.4 Physical transparency and translucency

Physical transparency and translucency describe situations where light that was reflected from an object interacts with partially light-transmittant media before it reaches the eyes. This interaction causes the light to change in ways that strongly depend on the physical configuration of the medium it passes through. If the medium contains no embedded particles that could scatter the light, its only effect is the attenuation of the incident energy-distribution over the spectrum. The pattern of attenuation is called the light-absorption spectrum, which, theoretically, can vary within very large bounds. Even though, it seems that the characteristics of many realistically occurring media fall in a restricted subset of all possible absorption spectra (Faul and Ekroll, 2002). The case of pure energy attenuation by the medium is called physical transparency, and it is by far less common than translucency which also incorporates light scattering from embedded particles (Koenderink, 2003). In addition to attenuating and scattering light that passes through them, translucent media also reflect light on their own, which then mixes additively with the light they transmit.

Transparency and translucency can cause image fragmentation because objects may be simultaneously visible in free view, and behind the light-transmittant filter (figure 2.5). The resulting image from that object can be divided into different areas by sharp borders. The stimulus characteristics of the areas resulting from filtered light can differ greatly from those resulting from an unobstructed view of the same object. It is thus unobvious which differing image regions contain portions of light reflected from the same object and thus belong together, and which differing image regions really result from different objects. A successful object identification in this case requires a correct grouping of image fragments: It entails finding corresponding image regions to determine where the same object is visible both under the transparent layer and in free view.

While transparency and translucency as discussed so far are properties defined for physical objects, there is also the perceptual quality of seeing transparent objects. The relationship between physical and perceptual transparency is intricate, and it turns out that both are largely decoupled: Physical transparency is neither necessary nor sufficient for perceptual transparency, which instead seems to rely more on global stimulus patterns (D’Zmura, Colantoni, Knoblauch and Laget, 1997; Faul, 1997). Section 3.2.1 will later pick up on the role of transparency perception in visual completion phenomena.

2.5 Motion

The sources of input fragmentation presented in the preceding sections were restricted to those relevant to purely static scenes. Under natural viewing conditions, however, motion introduces temporal aspects to fragmentation which act in conjunction with

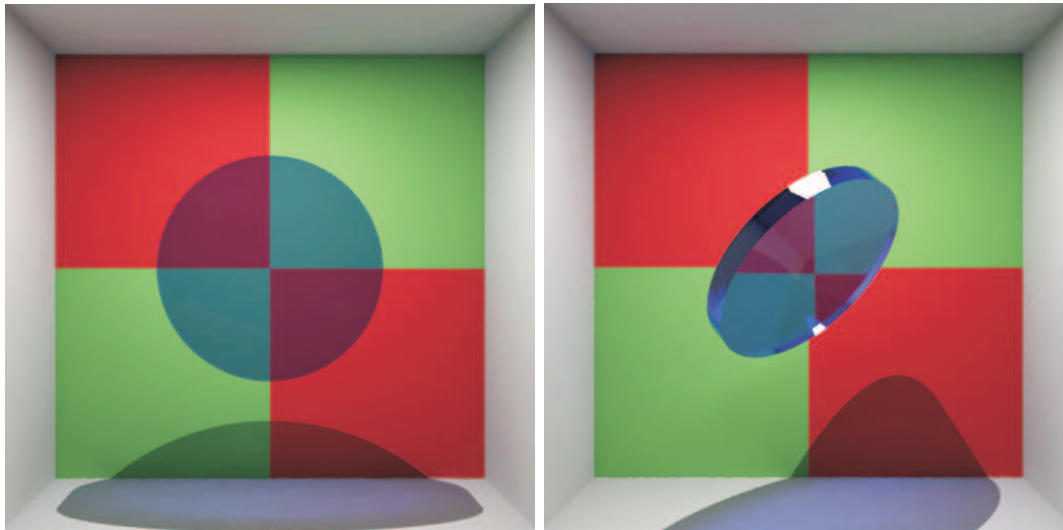


Figure 2.5: Two simulated examples of color transparency. The projection of the background is fragmented as a result of the transparent disk: While some of its area is freely visible, other parts are seen through the layer. Renderings courtesy of Franz Faul.

the influences considered so far, and which must therefore be taken into account in the analysis of the stimulus generation process.

With respect to the inhomogeneities in the optical sampling process, motion leads to a temporal variation in the level of detail with which input regions can be represented. For example, at one point in time, the projection of part A of a surface could lie in the fovea and therefore have fine-grained resolution while part B lies in the peripheral visual field with low spatial resolution. In the next moment, A might be placed in the periphery and B in the fovea. A high-detail representation of A in the first moment must now be temporally linked to a low-detail stimulus region. Since low levels of detail could introduce ambiguities regarding possible correspondence partners, this causes a difficulty in maintaining object rigidity and identity. In this context, it is interesting to note that there obviously are compensation mechanisms in the control of body, head and eye movements which succeed in keeping the projection of a moving target surprisingly stable with respect to its location on the retina (Rodieck, 1998). This prevents large temporal resolution changes of stimulus regions belonging to the same object.

Regarding occlusion, motion changes the belongingness pattern of separated input regions because it leads to a process of covering and uncovering of different scene elements in the vicinity of the occluding object. This dynamic accretion and deletion of image regions again introduces the problem of temporal correspondence between spatiotemporally disconnected regions. But it also can, at the same time, be used to

gather depth information: The temporally continuous regions must be closer to the observer than the changing ones. However, the possibility of extracting depth information from the pattern of accretion and deletion is complicated by possible interactions between several sources of motion, i.e., self motion and independent motion of different objects in the scene.

With rigid object motion, all object parts and surface texture elements share a common motion component, a characteristic known as “common fate” (Uttal, Spillmann, Stürzel and Sekuler, 2000; Wertheimer, 1923). Due to this fact, motion discontinuities are likely to be present between objects, but rarely occur within one. An exception to this rule is the movement of shadows which needs to be separated from the surface it occurs on, and instead needs to be linked to the shadow casting object. The special status of shadow motion was indeed illustrated in experiments which show how the perception of a dynamic three-dimensional scene layout can be drastically influenced by the manipulation of shadow motion alone (Kersten, Knill, Mamassian and Bühlhoff, 1996; Kersten, Mamassian and Knill, 1997).

In line with the natural constraints of common fate, a shared motion component of stimulus elements seems to be a very strong cue for grouping input fragments and determining the unity of separated stimulus regions. This becomes especially obvious when common fate leads to the perception of objects in the complete absence of static shape information, as demonstrated in random dot kinematograms (Gibson, Kaplan, Reynolds and Wheeler, 1969; Kaplan, 1969). A natural instance of this process can be observed when motion interacts with camouflage which often instantly breaks down once the camouflaged object starts to move (Cicerone and Hoffman, 1997; Metzger, 1975). As a countermeasure, some animal species like the cuttlefish seem to have developed dynamic camouflage techniques which can compensate for the common motion of their surface texture, thus tampering with the motion analysis mechanisms of beasts of prey, and misleading them in the perceived motion direction (Anderson, Baddeley, Osorio, Shashar, Tyler, Ramachandran, Crook and Hanlon, 2003).

Chapter 3

Visual completion across space and time

We are usually oblivious to the high degree of spatial fragmentation of the visual input—our perceptual world is populated by complete objects with seemingly no parts intrinsically missing, even if they are temporally hidden from our view. In ordinary life, this rarely strikes us as curious. We commonly accept our perception simply as a mostly faithful representation of some relevant properties of the physical world. Thus we are not surprised that just like the physical world is characterized by constant properties like temporal stability and spatial contiguity, so is our perceptual reflection of it.

Contrary to this view, vision science regards the described relation between visual input and perceptual representation as in dire need of explanation. The question more or less implicitly asked in most accounts of the problem may be posed as how the integrity and permanence of objects can be preserved in spite of an input that is fragmented in space and time. The terms “perceptual completion” and “filling-in” are commonly used as labels for both the phenomenological aspects of perceiving complete scenes on the basis of incomplete stimuli, and the visual processing involved. Although I will adopt this established terminology, I do so under reserve: Talking about “perceptual completion” suggests a theoretical perspective on the issue which too easily falls into the trap of understanding perception simply as the result of multiple transformations of the input. Following this slippery slope, visual processing quickly becomes “filling in the gaps in the stimulus”. This line of thinking, however, may be inadequate and may actually prove to be a hindrance for a better understanding of object perception in the face of stimulus fragmentation. The deeper problem of theorizing about visual completion raised here will be addressed later in section 6.1.

Before turning to the conceptual intricacies associated with the topic, section 3.1 will illustrate some of the difficulties one encounters when trying to structure the field of perceptual completion. It will also outline the particular strategy to do so that is employed here. Section 3.2 will address the phenomenology and empirical evidence relating to static completion; the more general case of dynamic displays will be covered in section 3.3. Both sections will make apparent that visual completion involves the interaction of many aspects of our perceptual world which are commonly analyzed separately from one another. Section 3.2.3 therefore specifically highlights the interactions between aspects of perceptually completed features like boundaries, color, depth,

and transparency. Since perceptually completed entities cannot be based on information that is different from that underlying the non-completed entities, one might wonder whether there is any use or functional role for perceptual completion. Section 5.5 will argue that subjectively completed objects are often functionally equivalent to non-completed ones in visual processing, and will give an overview over the relevant methodology used to tackle this question. One particular aspect that bears special relevance to the models for, and philosophical status of completion phenomena is its time course. The question whether completion is an active process spurred several inquiries in how much time it takes the visual system to develop a complete representation of a fragmented stimulus. The experiments and results associated with this question will be covered in section 5.4.

3.1 Structuring completion phenomena

The study of visual-completion phenomena is an inquiry into the question how the visual system deals with the problem of generating a functionally complete perceptual environment when faced with input fragmentation. Specifically, vision science tries to understand how disconnected stimulus regions are processed to the end of determining unity, shape, position, surface attributes and identity of 3-D scene elements. One postulated early mechanism serving unit formation is the clustering of corresponding elements, a process traditionally called grouping. In order to achieve this goal, information from multiple sources must be combined, an attempt which has to deal with incomplete and even conflicting information. Even when perceptual units are formed, their shape may still be indeterminate (Koffka, 1935). In fact, in many cases a precise shape representation of a unified object is never generated (cf. section 6.1). But when it is, it constitutes the result of combining information from fragments deemed stemming from one object. The aspect of information integration is not unique to perceptual completion but also plays an important role, e.g., in motion perception (Hildreth and Koch, 1987), the combination of different cues to one perceptual aspect such as depth or color, and more generally in the spatiotemporal binding of different features into coherent objects or events (cf. section 6.1).

On a more general level, visual completion may also be seen as an especially obvious case of perceptual inference in response to the fact that the input is vastly underdetermined with respect to the originating scene: As a result of the many-to-one mapping from physical scene to proximal stimulus, the same local image properties may usually have been caused by an infinitely large number of distal stimuli (Hoffman, 1998). The precise quantitative aspects of the disparate relation between distal and proximal stimulus have so far been analyzed for a variety of aspects of visual processing, e.g., for the case of perceiving lightness on the basis of luminance (Adelson, 2000), or for perceiving stable surface colors despite changing illuminants (MacLeod and Golz, 2003; Mausfeld, 1998). Applied to the case of input fragmentation, the ambiguity of the stimulus manifests itself as an infinite number of possibilities to replace the lost information that was

3.1 Structuring completion phenomena

carried by those parts of the incoming light array which were blocked from reaching the eyes. Put bluntly, since all surface structures “look” the same when they are not visible, an occluded area can, in principle, correspond to anything in the scene.

As was famously argued for by Helmholtz (1894) and, before him, Alhazen (Sabra, 1978; 1989), the visual system cannot arrive at a solution to the general form of this problem without adding information itself. From the external perspective of scientific analysis, this may be circumscribed as perceptual assumptions or processing heuristics. Arguably though, this terminology merely reflects the attempt to attach a meaningful label to the fixed structure of the visual system which “adds information” only implicitly by the virtue of its processing design. Thus the inference is not thought to be cognitive, but to be an automatic perceptual process governed by idiosyncratic rules, i.e., following an *intrinsic* logic. What is called perceptual assumption from an information processing perspective may establish itself in a specific situation of input fragmentation like this: It is the way in which the available information constrains the replacement of missing information in the process of integrating image fragments. It is thus the goal of the present work to identify characteristics of the rules that govern processes and effects of visual completion.

When trying to give a systematic account of visual completion, one faces the difficulty of providing a coherent structure for a very rich and diverse phenomenology arising from several different situations which all promote some sort of filling-in. Ideally, such an overview should be theory driven, grouping those instances and aspects together that share common processing components. Unfortunately, a theory which may provide that sort of structure is not yet in place, and has only begun to emerge piecemeal (Pessoa et al., 1998). For this reason, I will largely follow a mixed approach which takes into account theoretical as well as phenomenological aspects. I will therefore first point out what perceptual components can be involved in visual completion, and will then separately consider each one in turn. This account prepares for the presentation of empirical findings on complex interactions of the components participating in the formation of complete perceptual objects and scenes. Throughout, I will let the following questions guide my treatment of filling-in:

- Is the completion purely spatial (as in static stimulus material), or spatiotemporal (as with dynamic stimuli)?
- Are the completed entities represented with explicit visual qualities or not (modal vs. amodal completion)?
- Does the completion pertain to shapes, filled surfaces, or three-dimensional volumes?
- What visual feature is completed, i.e., contours, color, texture, material quality, depth, or motion?

Especially the last of the abovementioned taxonomic criteria should be regarded as highly provisional since it is all but clear-cut what the basic features of visual processing

actually are. It will in fact become apparent that many of the experimental results reviewed here strongly argue for a different set of visual primitives, as will be discussed further in section 3.2.3.

3.2 Static completion

Even though our ultimate goal is to understand perceptual completion as it occurs in dynamic environments, it is useful to first separately consider the case of shape and object formation in static scenes. Historically, this was the main starting point for inquiries into visual completion, and so far has received much more scrutiny than dynamic interpolation. Partly for these reasons, theories of spatiotemporal object formation are heavily influenced by models and ideas gained in the examination of the static case, as will become especially apparent in section 3.3.2.

Instances of visual completion are often divided into categories according to the question whether the perceptually completed portion of the scene is explicitly represented or not. Owing to Helmholtz (1894), the standard terminology is such that if the completed part possesses the same visual qualities as the rest of the scene, completion is said to be “modal”. Otherwise, i.e., if the completed part is deprived of sensory qualities, completion is labeled “amodal” (Michotte and Burke, 1951; Michotte et al., 1964).

Here, the question instantly arises in what sense visual completion without any phenomenological qualities can be sensibly termed completion at all. After all, what could be completed if not the visual qualities? A plausible reason to assume that amodal completion is a truly visual phenomenon will be brought forward in sections 3.2.2 and 5.5. On the basis of several experimental results, we will argue that amodally completed scene elements share many functional properties with elements that are not completed.¹ Therefore, amodally completed parts are just as “real” perceptually as elements that carry an explicit phenomenological status.

As with the distinction between modal and amodal completion, I will adhere to another influential separation in the following account of completion phenomena: The segregation of contour versus surface feature completion that was argued for, e.g., by Grossberg and Mingolla (1985). This distinction seems to be warranted, since several demonstrations show that a dissociation between completed contours and surface aspects like color or texture is possible: Just as the presence of subjective contours does not always coincide with phenomenal surface changes (figure 3.1), there are instances of completed surface features without subjective boundaries (figures 3.13 and 3.14). Other such cases can also be observed with variants of the dynamic stimuli that were used in experiments underlying this work (cf. section 4.2).

¹Conceptual issues concerning the terminology of the dichotomy of “completed” and “real” perceptual elements will be discussed in section 6.1

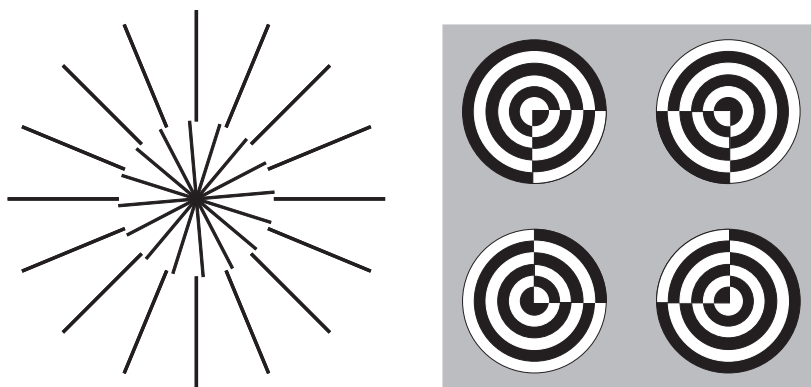


Figure 3.1: Examples of modally completed contours in the absence of surface changes.

3.2.1 Modal completion

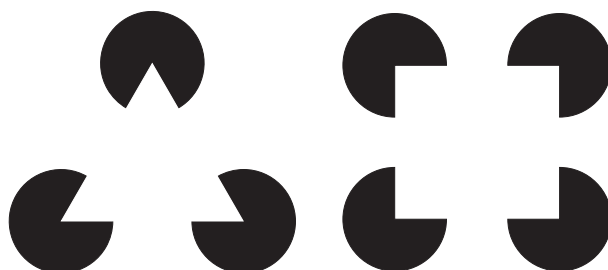


Figure 3.2: Modally completed boundaries together with an illusory lightness enhancement in the Kanizsa square and triangle. Boundary completion is here accompanied by perceived surface changes.

When do we modally complete? In the treatment of modal completion, it is instructive to first bring to mind that modal completion is not limited to laboratory-devised stimuli of which the Kanizsa triangle and square (figure 3.2) are probably the most widely circulated. Rather, there is a great diversity of circumstances that all promote an explicit representation of visual qualities not defined in the input. Among these are, first for the case of healthy observers, the following naturally-occurring and artificial ones:

- The blind spot in the eyes is not represented as an area where visual information is absent—instead, it is modally filled in with the same visual qualities that characterize other peripheral regions in the visual field (Ramachandran, 1992a; 1992b; 1993).

- Using retinal stabilization techniques Gerrits, de Haan and Vendrik (1966; see also Yarbus, 1967) showed that areas enclosed by retinally stabilized borders are subjectively completed with the surface quality of the surrounding region (see figure 3.3).
- For Ganzfeld situations, there exist reports claiming that the representation of a surface with a corresponding color is not held up, but rather filled in entirely with a visual quality known as “Eigengrau” (Knau and Spillmann, 1997). This, however, is a contested observation.
- With prolonged fixation, items in the peripheral visual field can fade and be filled in with the quality of the surrounding area, a process termed “Troxler fading” (Troxler, 1804).
- In situations mimicking the visible occlusion of elements by otherwise perfectly camouflaged objects, modally completed contours can be observed, most famously mimicked in the case of the Kanizsa triangle (figure 3.2; Kanizsa, 1979).
- Modally completed contours can also be obtained from random dot stereograms (Julesz, 1971), random dot kinematograms (Gibson et al., 1969; Kaplan, 1969), and half occlusion, also known as da Vinci stereopsis (Nakayama and Shimojo, 1990a). Here, there is no static local shape information at all which indicates the course of the boundaries (cf. section 3.2.3).
- Even in seemingly inconspicuous situations like typical center-variegated surround displays (Mausfeld, 1998) used in the study of color constancy, the central patch is arguably completed with a uniform surface quality: Due to different color inductions from the surrounding regions, it should be expected that the center obtains an inhomogeneous appearance. Nevertheless, this is usually not noticed when attention is spread globally on the display. But it can become apparent in color matching procedures that require focused attention on the center (Whittle, 1998).

Concerning clinical conditions, there are several pathological developments which entail the modal completion of visual features: Hemianopia, a reduction in vision in one half of the visual field that can, in its extreme form, lead to blindness in the affected half. It can be caused by damage to the optic pathways due to brain injuries following a stroke, tumor, or trauma (Fuchs, 1921; Koffka, 1935; Poppelreuter, 1917). Some severe methodological difficulties notwithstanding, it seems that the affected part of the visual field is in some instances modally completed without the awareness of the patient (Walker and Mattingley, 1998). Patients with a visual neglect (the behavioral ignorance of one particular side of elements in the environment) report the perception of whole objects even though they behaviorally respond only to one half of them. Arguably, they are in a sense blind to the other half. Finally, degenerations of the retinal periphery known as scotomas can eliminate neural transduction of visual input in parts of the visual field without having the patient consciously experience a loss of

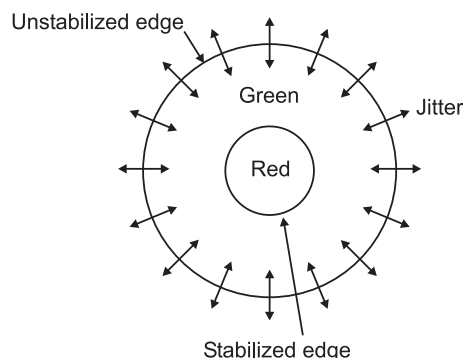


Figure 3.3: Retinal stabilization. The inner contour is retinally stabilized whereas the outer contour is not, and thus constantly jiggles over the retina. As a result, the surround color perceptually fills in the central patch.

information. Instead, surface qualities of regions surrounding the scotoma fill in the affected area (Ramachandran, 1991).

When talking about modal perceptual completion, one evidently deals with a wide variety of phenomena arising in very different circumstances. What unites them, at least superficially, is the conscious representation of visual qualities that have no local counterpart in the retinal image and stand in no obvious relation to either the physical input or the neural transduction thereof. They all highlight the discrepancy between the sparse input to the visual system and the rich perceptual structure that is typical for our visual world. Simply put, the aforementioned examples are instances of seeing things that are not “really” there or would—from a pre-theoretical point of view—be at least expected to look different.²

The question remains whether the abovementioned phenomena share some common perceptual processes that are responsible for the grouping and unification of fragmented image regions. While this question is still largely up for debate (see Pessoa et al., 1998 for an overview), one common aspect seems to be that neither is the majority of perceptual completion cases susceptible to cognitive control, nor does it require familiarity with the stimulus material. This indicates that the mechanisms involved work, with notable exceptions mentioned in due course, largely automatic and do not require the involvement of reasoning.

In the following sections, I will review different instances and aspects of boundary as well as surface completion, tacitly assuming that all of them shed light on the perceptual processes governing visual filling-in. The treatment necessarily concentrates

²Insofar as visual perception in general might be characterized as being grossly underdetermined by the input to the visual system, this simple description however quickly becomes less clear than it may seem at first glance. Again, this train of thought leads to conceptual issues which will be discussed further in section 6.1.

on only a few of the completion phenomena listed above, as they are not equally suited for empirical examination, leading to the situation that some cases are much better studied and analyzed than others. It is by no means intended to provide an exhaustive account of the different phenomena, but I will nevertheless try to incorporate the most important facets that appear relevant for the empirical work presented in chapters 4 and 5.

Boundary completion

Modally completed boundaries (figures 3.1 and 3.2) are often called “subjective” or “illusory” to emphasize their special status compared to contours derived from extant stimulus discontinuities (Gregory, 1972; Schumann, 1904). Even though this terminology is dubitable if taken literally, it is adopted here to match the existing literature, and both terms are used synonymously. It should be noted up front that even though this section will concentrate on the perception of modally interpolated contours, nearly all examples given below may equally well figure in the treatment of amodal completion, which often (but not always, see Pinna, Ehrenstein and Spillmann, 2004) complements the modal one in camouflaged occlusion displays. In the Kanizsa triangle for example, the “pacmen” look like discs with one section occluded by a camouflaged triangle. The nature of the close relationship between modal and amodal completion hinted at here will be discussed in more detail when amodal completion is considered in-depth in section 3.2.2.

In the study of perceptual filling-in, boundary completion has received the major share of attention (see, e.g., Parks, 1984; Petry and Meyer, 1987), owing to the widely accepted view that boundaries act as the most basic building blocks in object recognition. Since objects typically begin and end at boundaries, object recognition as a bottom-up process is thought to involve the finding and classification of contours as the first processing steps (Kellman, Guttman and Wickens, 2001; for a region-based approach, see Palmer, 1992). Consistent with this emphasis on the importance of boundaries is the ability of artists to provoke sophisticated interpretations even with simple line drawings: In such drawings, one can often perceive a three-dimensional scene layout, object curvature and volume, and even facial expressions of emotion and mood. The identification of object boundaries often implies a grouping of enclosed image fragments into a unit with defined shape, and boundaries are finally assigned to a constructed surface which owns them. (Kellman, 2003). In the traditional, largely hierarchical approach, once boundaries are determined, they define surfaces whose qualities are apparent within the defined contours. Surfaces in turn then link up to form three-dimensional volumes (see below for findings challenging this view). This conventional role of contours notwithstanding, illusory boundaries can also arise without belonging to any explicit surface at all, as figure 3.1 illustrates. Moreover, the reverse constellation of illusory surfaces without clear boundaries is also possible (figures 3.13 and

3.14). This will be discussed later as a potential indication of a general dissociation between contour and surface mechanisms.

As figure 3.4 illustrates, subjective contours do not represent a dichotomous all-or-none phenomenon. Rather, interpolated boundaries can differ greatly along a continuum of strength or distinctiveness. Models trying to provide an explanation for boundary completion thus need not only predict under which circumstances contour filling-in occurs at all, but also must account for the continuous variability of the percept. As will become apparent in the following sections, illusory contours are not unique with respect to the continuous nature of their salience and clarity. Many surface characteristics like color or texture share this seemingly continuous variability along a given perceptual dimension.



Figure 3.4: The strength of completed boundaries can vary gradually.

While commonly being the case, the sources of edge information that provide input for contour completion need not be luminance based. Instead, edges from color, texture and disparity discontinuities can all be functionally equivalent with respect to their ability of being linked together. For example, the Kanizsa triangle is also visible in random dot stereograms when the boundaries are defined by disparity alone (Nakayama, Shimojo and Silverman, 1989). Section 3.3, will also further extend this result to dynamic edge cues that may merely be abstract spatiotemporal discontinuities, sometimes not even consistent with possible physical scene configurations. Taken together, this indicates that the origin of shape information seems to be largely irrelevant, as shape formation mechanisms appear to equally accept input from many different channels, not just those blatantly associated with it.

The study of subjective contours has mostly been concerned with the isolation of geometrical relations in the input that are necessary and sufficient for the formation of illusory boundaries. On a general level of analysis, subjective boundaries seem to form as the result of integrating different sources of edge information in the stimulus into a unified contour. Areas enclosed by the stimulus-defined edges that initiate the formation of illusory contours are called “inducers”, and in line with the idea that the combination of different edge information sources is crucial for obtaining subjective boundaries, no instances of illusory contours have been reported in displays with just one inducer.

Regarding more concrete requirements for the geometrical relations between existing edge fragments in the stimulus to being integrated, several propositions have been made: Most models assume that for projected fragmented edges to be potentially linkable, they need to end abruptly as indicated by a tangent discontinuity (cf. section 2.2). As is obvious in figure 3.5, there are more constraints on the spatial relation between two edges for being linkable than just ending abruptly in tangent discontinuities. This is likely the case because not all tangent discontinuities result from occlusion, they may just as well be the consequence of the object-inherent shape configuration. Therefore, other criteria for potential connectedness between contours across gaps are required: The well-known Gestalt condition of “good continuation” (Wertheimer, 1923) tries to capture a certain type of contour smoothness, as does the influential concept of “relatability” (Kellman and Shipley, 1991), which may be taken as a more rigorous formulation in the same spirit (but see Kellman, Garrigan, Kalar and Shipley, 2003). It basically states that for two edges to be relatable, they must end in a tangent discontinuity, be connectable by a smooth curve with at most one inflection (Singh and Hoffman, 1999), and their extensions must intersect at an angle between 90° and 270° (figure 3.5). The presence of relatable pairs of edge fragments that end in tangent discontinuities is then conjectured to be the initiating condition for the process of contour integration. The precise form that the interpolated edges are then actually seen in are, however, not specified by the relatability concept. It seems that perceived contours are always smooth, but their specific shape has so far proved to be more difficult to model than the triggering conditions for contour completion (Kellman, Garrigan and Shipley, 2005).

One rationale behind the definition of relatability is that projective geometry implies that smooth curves in 3-D will generically project to smooth curves in 2-D, and, more specifically, that straight edges in 3-D lead to straight edges in 2-D. This implies that boundaries that are smoothly connected in 3-D will preserve this quality also after projection to two dimensions, and will therefore have relatable edge segments if the edge is disjointed by occlusion. Apart from highlighting the conditions that relatability does use, it should also be noted what potentially relevant geometrical properties do not count in its definition: Tangent discontinuities do not have to be luminance defined, but may also arise from texture differences or sharp transitions in disparity. This liberal attitude towards possible input channels will be expanded to purely temporal sources of information in section 3.3.2. Even though much work in shape perception is devoted to the special importance of concave curvature minima for segmenting objects (Hoffman and Richards, 1984; Singh and Hoffman, 2001), tangent discontinuities do not have to be concave³ (figure 3.6; Kellman and Shipley, 1991). In addition, sources of edge information do not even need to have the same orientation with respect to the shape, i.e., some inducers may specify the object from the outside, and some from the inside (figure 3.6). In line with the theoretical irrelevance of these factors for relatability, several demonstrations indeed show that illusory contours are obtained without these

³The concept of concavity can be generalized to boundary points that are not first-order continuous (Hoffman and Richards, 1984).

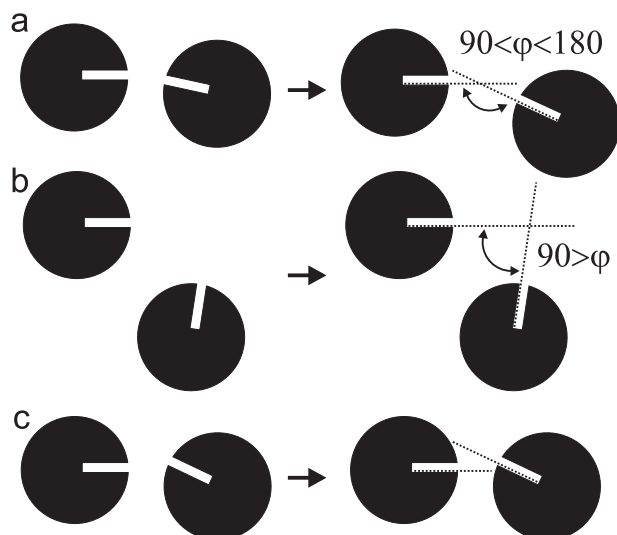


Figure 3.5: The idea of relatability. a) Illusory contours are perceived between edges whose tangent extensions form an angle between 90 and 180° . b) No contours are completed between edges whose extensions intersect at an angle less than 90° . c) No contours are completed between edges whose extensions do not intersect.

conditions met by the stimulus.



Figure 3.6: Modally completed contours between reliable edges that define the object from the inside (convex fragment boundary) as well as from the outside (concave fragment boundary).

Fitting the phenomenology of illusory contours, relatability is thought to be of variable magnitude, continuously degrading towards the limits of its boundary conditions. Many studies have shown that relatability is able to predict the occurrence as well as the strength of illusory contours in a variety of stimulus classes to a good degree (Kellman et al., 2001). One important characteristic of the relatability concept is that it is entirely input driven, i.e., testing whether two edges are reliable does not require specific knowledge about the stimulus, e.g., from prior experience. As such, it fits

well with findings indicating that knowledge of the fragmented shape normally has no influence on its completion (Kellman et al., 2001).

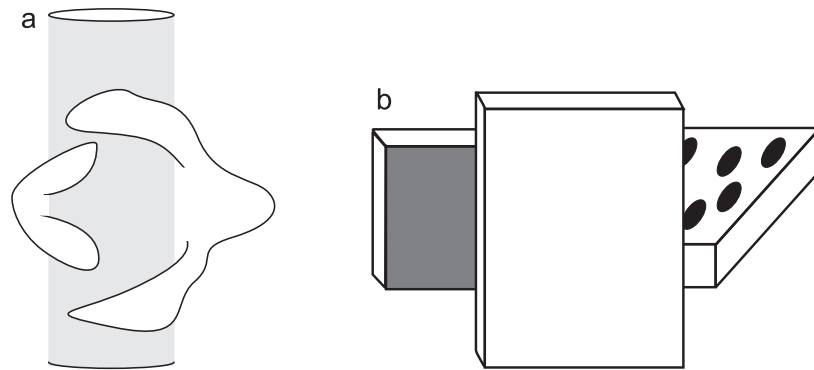


Figure 3.7: a) Amodal completion in the absence of relatable edges. b) No completion despite the presence of relatable edges. Figure adapted from Tse (1998a; 1999)

Recently, however, some evidence accumulated that the principle of relatability does not capture the geometrical requirements for boundary formation as well as initially thought. This is suggested by demonstrations of contour integration between non-relatable edges (figure 3.7a) and the lack of completed contours between relatable edge fragments (figure 3.28b; Tse, 1998; 1999). Furthermore, lines and dots can elicit strong subjective boundaries in appropriate configurations, and can also alter the perceived illusory shape in displays containing tangent discontinuities (figure 3.8; Bruno, 2001; Leshner and Mingolla, 1993; Shipley and Kellman, 2003). However, it is unobvious how to integrate the induction of illusory contours by thin lines and dots into the concept of relatability. Neither do lines or dots contain tangent discontinuities, nor do they have a defined orientation at their end as required for constructing imaginary extensions and checking their intersection angle (Kellman and Shipley, 1991; Shipley and Kellman, 2003). Furthermore, extended boundaries that do not end in any sort of tangent discontinuity can also sometimes trigger a filling-in process. As will be discussed below in the context of amodal interpolation, it seems that these results call into question the general idea of illusory contours exclusively being the result of a purely stimulus-driven process, but rather suggest that boundary interpolation has to be considered within the more general context of perceptual formation of surfaces in depth.

Surface feature interpolation

Although boundaries undeniably constitute a highly important aspect of visual processing, recent evidence suggests shifting away from regarding them as the basic building blocks of our visual experience, and rather highlights the fundamental role of surfaces as

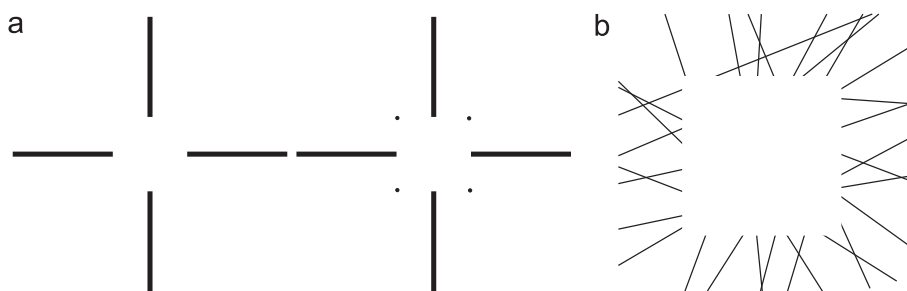


Figure 3.8: Dots and thin lines can influence interpolated shapes. a) In the Ehrenstein cross, an illusory disk is perceived in the central area. Four dots arranged as a square lead to the perception of an illusory square instead, in an otherwise identical configuration. b) Abruptly ending thin lines can induce the perception of a modally completed square.

perceptual primitives and functional units of visual processing. Many convincing reasons for this development were brought forward by Nakayama et al. (1995; Nakayama and Shimojo, 1990b; 1992) who strongly argued for the existence of a surface concept with special status in visual representations. Some of their arguments and experimental work will be presented in section 3.2.3.

Already Katz (1911) stressed the relevance of distinguishing between different modes of color appearance, and specifically commented on the opaque surface quality some colors exhibit in contrast to so-called film colors which lack it (see also Koffka, 1935; Metzger, 1975). Surfaces are thus characterized by their perceived solidity and the impression of representing a cohesive unit. They can be segmented as figures on a background, and may be defined along several perceptual dimensions: They exhibit features such as a position in depth, coloration and brightness, texture, and degree of transparency. In the following paragraphs, it will become evident that perceptually completed objects are often defined along the very same dimensions as uncompleted objects are, with all of the aforementioned features visually filled in.

Color and brightness spreading. As a generic term, color spreading subsumes a number of different visual phenomena. They all demonstrate how certain spatial configurations can give rise to the perception of surface qualities, such as color or brightness, attributed to stimulus regions that are themselves devoid of corresponding physical information about such features. Figure 3.9 shows a recently introduced variant of this effect (Wollschläger, Rodriguez and Hoffman, 2001; 2002). Here, color is not only seen in the physically-colored parts of the display, but also in the regions between them, giving rise to a lightly-colored surface. The induced color spreading is thereby bounded by modal subjective contours (Day, 1983; Kanizsa, 1979; Schumann, 1904). As will be discussed in chapters 4 and 5, cases of color spreading are not confined to static stimuli

but may be elicited by appropriately crafted dynamic scenes, as well.

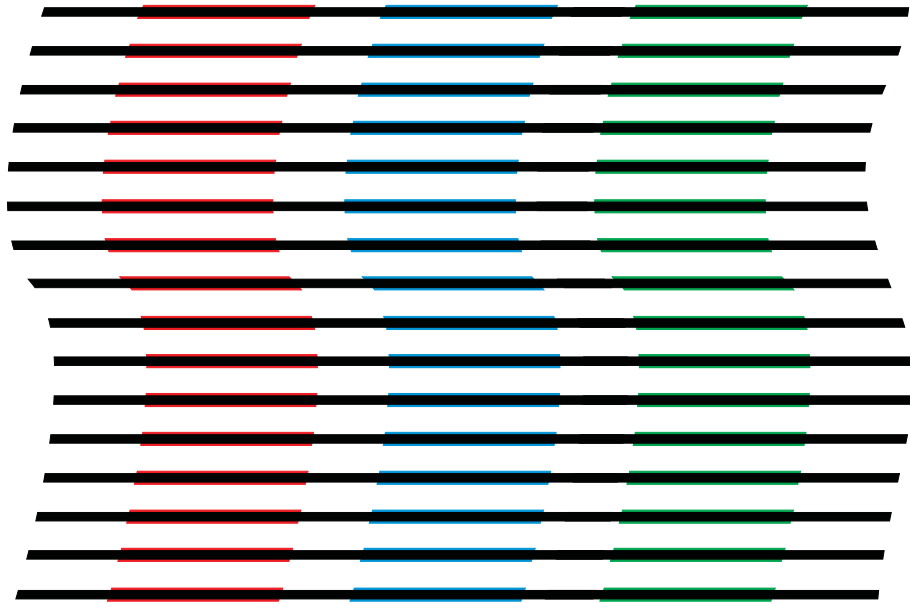


Figure 3.9: Color spreading induced by adding thin colored flanks to black lines.

While the phenomenon of color spreading was probably known and systematically exploited for practical use as early as in the 16th and 17th century, the first scientific account was given by Wallach (1935). The phenomenon was later independently re-discovered by Varin (figure 3.10b and c; Varin, 1971) and vanTuijl (figure 3.10a; van Tuijl, 1975), and only then received more attention from vision science.

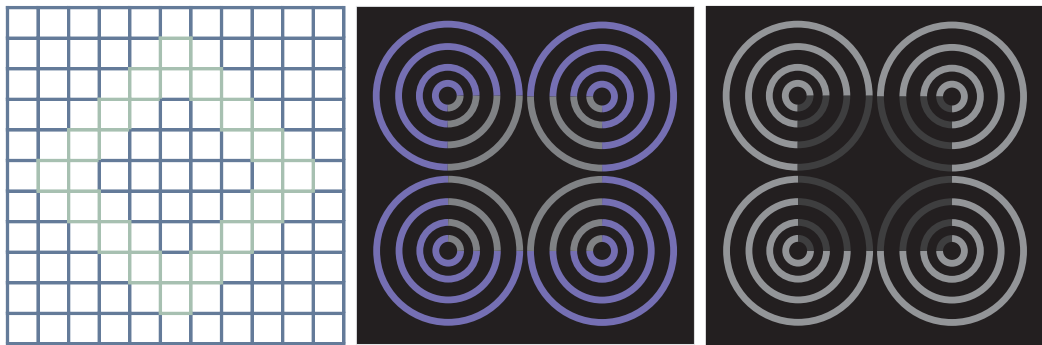


Figure 3.10: Neon color spreading. a) In the van Tuijl figure. b) Complementary color spreading in the Varin figure. c) Darkness spreading.

Depending on the particular stimulus configuration employed, the subjective impression of color spreading can vary considerably in quality and strength, indicating

that the occurrence of color spreading, similar to the one of modal subjective contours, should not be considered as a dichotomous phenomenon. Finding out about regularities exhibited by the waxing and waning appearance of the filled-in color will indeed be one objective pursued in the empirical part of this work (cf. chapter 4). Apart from the subjective color itself, color spreading conceptually integrates a number of effects that may vary continuously in influence, and are of potentially different origin. This implies that their connection mechanism-wise cannot be assumed *a priori* but needs to be demonstrated by additional means. The following paragraphs try to convey an idea of the sizable gamut of effects that are all color-spreading related, trying to pinpoint important commonalities as well as differences.

Neon Color Spreading. Neon color spreading (figures 3.10 and 3.11a; for a review, see Bressan, Mingolla, Spillmann and Watanabe, 1997) refers to the perceptual phenomenon that color seems to disperse from image elements into their surround, thereby creating a subtle, neon-like veil. The observed coloration overcomes real figure boundaries and typically covers an area confined by subjective contours (Takeichi, Shimojo and Watanabe, 1992). The filled-in color resembles a less saturated variant of the inducer's color, similar to a highly diluted watercolor, but its impression is usually quite vague, or ethereal in quality. If the inducers are colored and the embedded segments achromatic, the spreading appears tinted in the color complementary to that of the inducing segments (figure 3.10b). In purely achromatic displays which are structurally identical to neon-color-spreading stimuli, one can observe an analogous spreading of brightness, or, depending on the luminance relation between the relevant stimulus elements, also darkness onto the homogeneous background (figure 3.10c; van Tuijl, 1975; van Tuijl and de Weert, 1979).

Many neon-color-spreading displays share a common construction principle: Neon color spreading can typically be observed in line drawings where the color of some segments is abruptly changed to a different one, i.e., where differently-colored segments are embedded into an ordered arrangement of lines. The embedded segments are placed such that they are compatible with a simple geometrical object that is only visible in the areas of the lines (which serve as viewing apertures to a farther depth-level) but not in the homogeneously-colored area enclosing the apertures (figure 3.12). The colored object that is thus suggested to be partially visible can take on a modal, fully colored appearance within its subjectively completed contours. In this case, the change in color associated with neon color spreading often coincides with a change in the material appearance of the filled-in surface as it turns into a slightly transparent layer (see below). An amodal interpretation (cf. section 3.2.2) of the stimulus is also possible (figure 3.12).

Whereas neon color spreading is usually associated with the presence of illusory contours, color and brightness spreading can also arise without them, as is evident in figure 3.13, which shows brightness enhancement in a vaguely defined, borderless region, and the so-called “neon flanks” phenomenon (Redies and Spillmann, 1981; Redies,

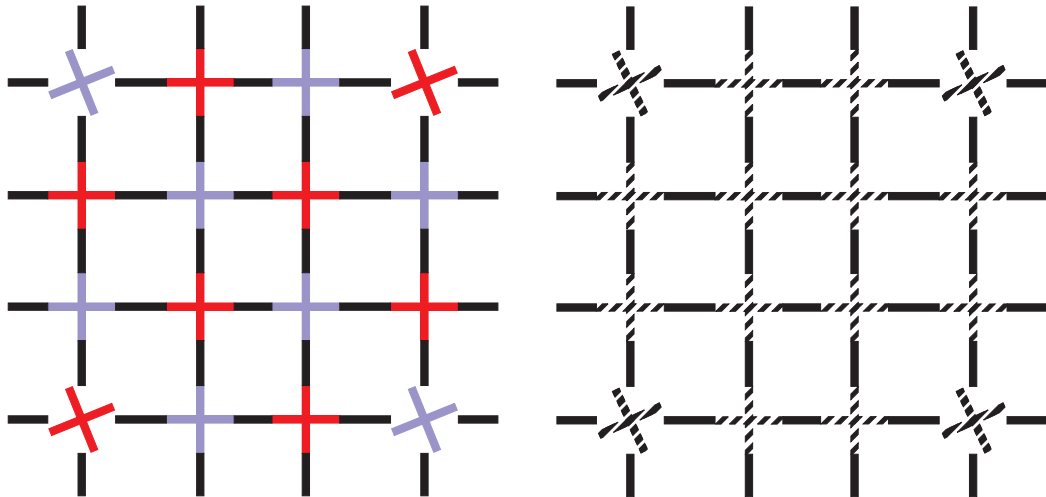


Figure 3.11: Neon color spreading and texture spreading in an array of modified Ehrenstein figures. The texture spreading can be best appreciated in the periphery of the field of view.

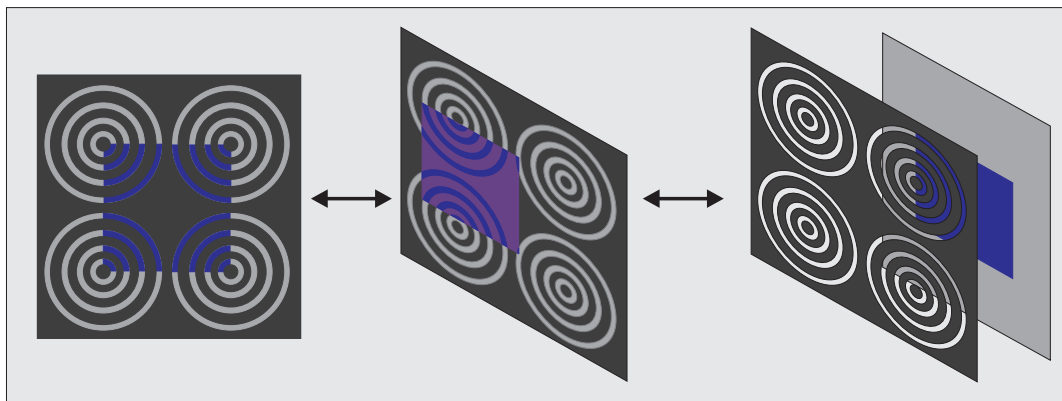


Figure 3.12: Interpretation of a typical neon-color-spreading stimulus in terms of a 3-D scene. Modal completion corresponds to the perception of a transparent filter floating in front of a surface with rings painted on it. Amodal completion corresponds to the perception of a square painted on a background, both visible through apertures cut into a front mask. Note that both interpretations differ with respect to the type of objects they involve.

Spillmann and Kunz, 1984). In a similar vein, Watanabe and Sato (1989) observed unbounded, diffusely fading color spreading in the modified Ehrenstein figure at isoluminant color conditions. Taken together with illustrations showing that illusory contours also need not enclose visible surfaces (figure 3.1), these demonstrations possibly entail

a dissociation between contour and surface interpolation mechanisms, implying that boundary and surface formation are more independent structurally than hierarchical processing schemes often imply.

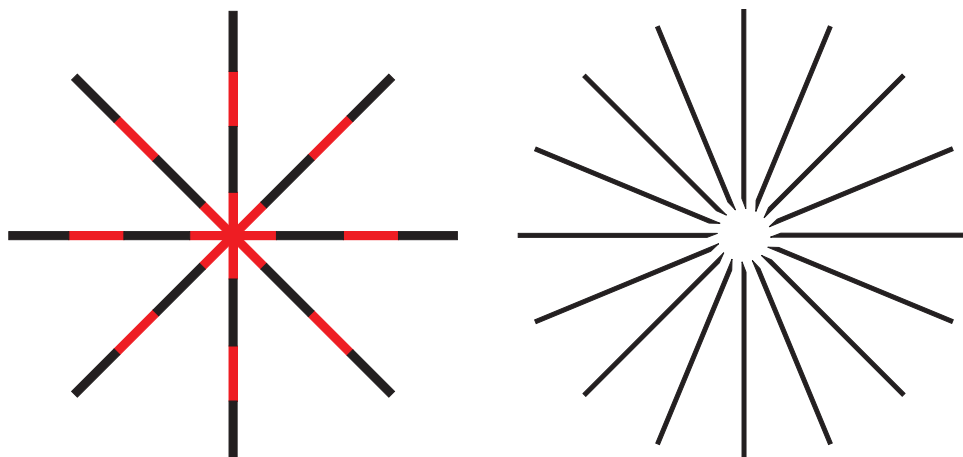


Figure 3.13: Flank-induced fuzzy color spreading that is not contained by illusory contours, and blurred brightness enhancement that is also unbounded.

Neon-flank stimuli also demonstrate how in some cases color disperses only in the immediate vicinity of colored figure elements, yielding the perception of narrow tinted flanks alongside them. The color spreading fades gradually away from the inducing segment without forming a clear surface area enclosed by subjective contours. Redies and Spillmann (1981) propose that the neon flanks phenomenon can be taken as a basic local effect which in accumulation serves to create global neon spreading as a lateral extension (see also Watanabe and Takeichi, 1990). If one simply extends the height of the inducers in the neon flanks effect, one obtains another display exhibiting a fuzzy brightness spreading called “visual phantoms” that is not bounded by illusory contours (figure 3.14; Rosenbach, 1902; Tynan and Sekuler, 1975; Weisstein, Maguire and Berbaum, 1977). Apart from their constructional similarities, neon color spreading in the guise of neon flanks and visual phantoms exhibit a largely overlapping set of characteristics with respect to the effect of stimulus changes and their effect on color spreading perception, strongly arguing for their theoretical relatedness. One aspect, however, sets visual phantoms apart from neon color spreading: Despite the completion being modal in quality, the completed elements are perceived to be located behind an occluder. This is in contrast to the coupling between type of completion and depth ordering normally observed in color spreading (cf. section 3.2.3).

Edge-induced color spreading. In accordance with the examples presented so far, already van Tuijl (1975) noted that, probably counterintuitively, larger colored areas do not contribute to color spreading as much as thin lines can do. It has indeed



Figure 3.14: Dark visual phantoms induced by an apparently occluded grating that is modally completed behind the occluding rectangle. Surface features are completed in the absence of illusory boundaries.

be shown that considerable filling-in of surface color can be induced by changing the color only of thin lines of a figure—even more than in the neon-flanks phenomenon. With his watercolor-effect displays (figure 3.15; Pinna, 1987; Pinna, Brelstaff and Spillmann, 2001), Pinna made an especially instructive case for this claim. Three other demonstrations highlighting the importance of information changes at edges for the perception of homogeneous surface colors are “flank transparency” (Wollschläger et al., 2002), the Craik-O’Brien-Cornsweet effect (Wachtler and Wehrhahn, 1997), and the perception of center-surround displays with a retinally stabilized border between them (Yarbus, 1967).

Flank transparency is the perception of a colored transparent filter evoked by stimuli containing as few as two colors (figure 3.9, Wollschläger et al., 2001; 2002). Displays of flank transparency contain an array of line segments placed on a uniform background. Small flanks are added to the line segments if the segments fall in the interior of a virtual shape, such as a virtual disk. This leads to the perception of a colored transparent object with well-defined boundaries lying over the array of lines. Although flank induced and not looking self-luminous, the color filling in of flank transparency shares some aspects with neon color spreading, most notably a transparency impression and its short range of operation. Static flank transparency displays can easily be turned into dynamic ones, yielding even more pronounced subjective boundaries, as well as color and texture spreading (cf. chapters 4 and 5).

In watercolor-effect displays with a white background (figure 3.15), the coloration is entirely induced by colored flanks snugly fitted to one side of a figure’s black contour.

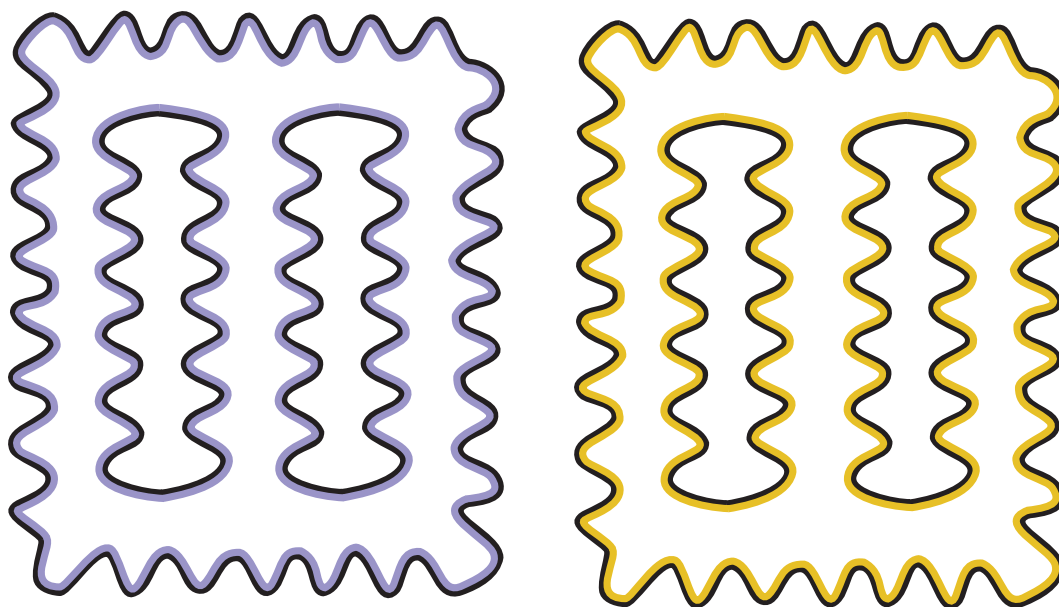


Figure 3.15: The watercolor effect. Matte illusory coloration induced by continuous thin edges is perceived over a large visual area.

The color then spreads on the side of the two-line contour that has the lower contrast to the background. The flank-induced coloration resembles pastel surface colors. The area covered by the “diffused watercolor” can encompass as much as 40 deg, a much larger region than what is usually bridged by neon color spreading which typically spreads only over a few minutes of angle. In this respect, the watercolor effect is more similar to the Craik-O’Brien-Cornsweet effect (see below) that also influences the perceived coloration over several degrees of visual angle. In contrast to the aforementioned neon flanks phenomenon, watercolor spreading may originate at thin lines but is in no ways confined to their vicinity. Pinna et al. (2001) further stress that as an additional difference to most classical neon-color-spreading examples, the color spreading in watercolor displays has neither a self-luminous nor a transparent quality associated with it. As such, it does not elicit a separate surface impression at all like neon color spreading does, but simply seems to be attached to the background surrounding the contour. Using the term “edge colors”, Broerse, Vladusich and O’Shea (1999) demonstrate fringe-induced color spreading in the modified McCollough effect (figure 3.16) which also lacks luminosity, transparency and objecthood.

The Craik-O’Brien-Cornsweet effect (figure 3.17) may also be taken as a spreading of brightness or color (Wachtler and Wehrhahn, 1997) which is fundamentally edge induced. Worth noting in this case is the construction of the display which differs from traditional neon-color-spreading stimuli as well as from the edge-induced ones. Instead of containing two operative color discontinuities—one from the outer elements to the

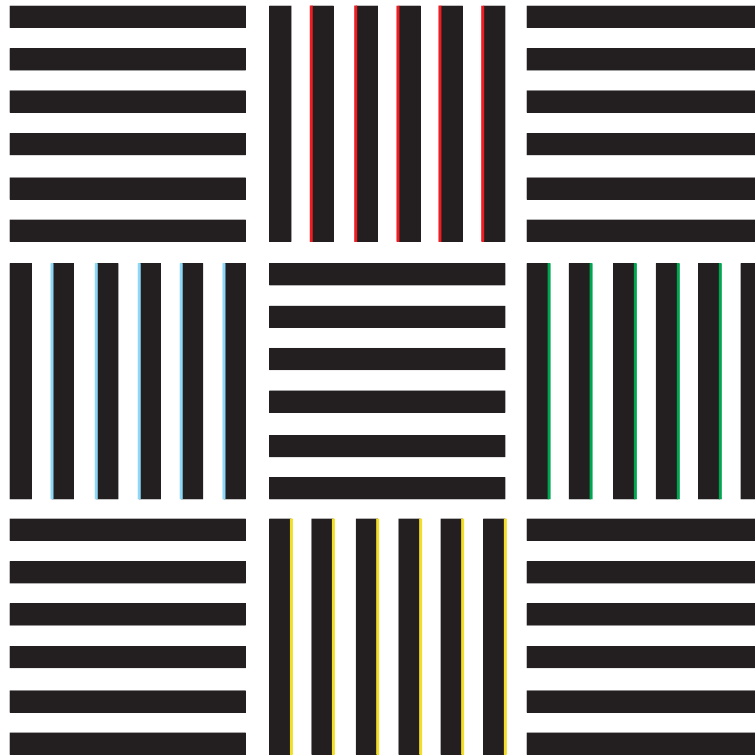


Figure 3.16: Flank-induced color spreading in the McCollough effect. Figure adapted from Broerse, Vladusich and O’Shea (1999).

colored ones, and one from the colored elements to the background— Craik-O’Brien-Cornsweet-effect stimuli have only one, separating the two visible areas. The inducing color thus stems from the gradient band which connects areas of constant color. As mentioned above, the range of operation is quite large, similar to that of the watercolor effect.

Findings regarding the perception of retinally stabilized center-surround images (Gerrits et al., 1966; Yarbus, 1967) also support the idea that edge information can be decisive for the perceived coloration of larger homogeneous areas: If the border separating the central patch from the immediately surrounding colored annulus is retinally stabilized, the central patch slowly acquires the color of its immediate surround (figure 3.3). The retinal stabilization of the contour compensates the eye’s natural tremor and in turn causes a constant input to those cells whose receptive field includes the border’s projection—thus supposedly fatiguing them and removing the change signal. The phenomenon may therefore suggest that there is a tendency in the visual system to maintain a given assignment of some surface attribute until explicit change information coming from a border indicates otherwise.

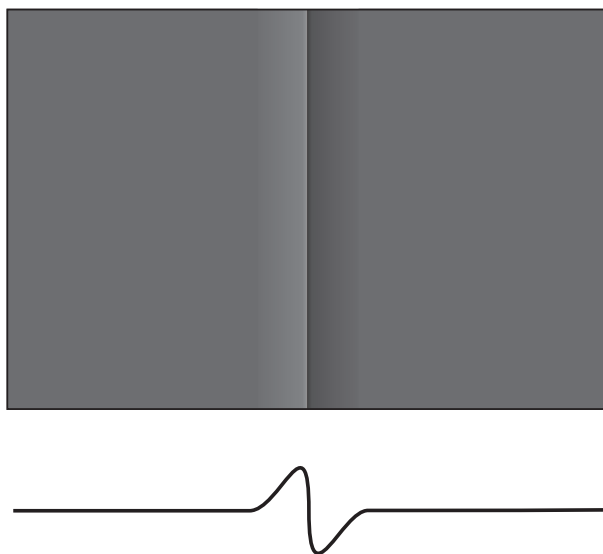


Figure 3.17: Craik-O'Brien-Cornsweet effect and the display's intensity profile. The left and right areas have identical physical intensity, but the inner gradient induces a perceived brightness difference.

In an attempt to pinpoint the origin of color spreading in neural processing, several studies addressed the question whether color spreading also persists in displays which present different structural stimulus elements to different eyes, thus requiring binocular fusion to perceive a complete color-spreading display. For the modified Ehrenstein cross (figure 3.11a), Redies and Spillmann (1981) concluded that no color spreading is seen if the embedded colored segments are separated from the achromatic grid, and projected to a different eye. Takeichi et al. (1992) confirmed this result, but also showed that subjective contours as well as color spreading obtain if the stimulus is segregated into the horizontal and vertical bar together with their respective embedded element. Regarding edge-induced color spreading in the watercolor effect, Pinna et al. (2001) reported that a binocular presentation of inducing edge and outer contour does not impede color filling-in, speaking against purely retinal processing mechanisms.

Texture spreading. Some neon-color-spreading displays can be altered to elicit texture spreading (figure 3.11b) when the uniformly-colored embedded line segments are replaced with, e.g., segments of a stripe texture (Watanabe and Cavanagh, 1991). In these stimuli, a *non-uniform* color pattern perceptually fills in a homogeneous background region, bounded by subjective contours. When the extant texture fragments are put in motion, the completed texture seems to move as a whole. Although undisputed, texture- and the associated motion spreading are weak effects and appear only in the peripheral visual field—both break down instantly when the inducing elements are foveated (Watanabe and Cavanagh, 1991). This suggests that the tolerance for con-

flicting positive information is lower for texture spreading than for subjective contours or color spreading.

In support of the idea that a lack of conflicting information is necessary for texture spreading to occur, texture spreading is also abundant in the part of the visual field corresponding to the blind spot (Kawabata, 1984; Koffka, 1935; Ramachandran, 1992a). There, no information at all can veto against the filled-in features. The same is true for the clinical condition of scotomas, where neural tissue is substantially damaged, and texture spreading has been repeatedly reported (Ramachandran, 1991). Furthermore, peripheral texture filling-in was reported for cases of prolonged fixation (Kawabata, 1984; Ramachandran and Gregory, 1991), a technique which supposedly eliminates change information at image contours due to fatigue (Troxler fading), similar to the case of retinally stabilized images (see above).

A perceived change of texture information of completely different origin is illustrated in figure 3.18: Here, texture completion in the sense of an increase in texture density is evoked as an aftereffect (Durgin, 1998). As there is a subjective impression of a denser texture, texture elements can be said to be filled in, even though no “added” individual elements can be pinpointed. Interestingly, this case of filling-in occurs not only in the absence of subjective contours but without any notion of a subjective surface, whose attribute the completed texture could be.

Material changes: Perceptual transparency. Typically, the region encompassed by neon color spreading appears—albeit sometimes faintly so—as a transparent layer (Bressan et al., 1997; Ekroll and Faul, 2002; Nakayama, Shimojo and Ramachandran, 1990; Varin, 1971). Apart from its connection to color spreading, perceptual transparency is also interesting in its own right as it involves the decomposition of *one* color designator at a particular retinal image location into distinct causal layers which are explicitly represented in the same direction of view (D’Zmura et al., 1997; Faul, 1997; Metelli, 1974). Since the segmentation of one local input value into several contributing factors is inevitably ambiguous (MacLeod and Golz, 2003; Mausfeld, 1998; Mausfeld and Niederée, 1993), the perception of transparency is by necessity based on global input relations.

At first glance, perceptual transparency may appear to be a somewhat curious percept—after all, instances of perceptual transparency in real-world scenes do appear to arise only rarely. However, closer inspection reveals that transparency can be encountered quite often in natural scenes, e.g., in the form of fog or haze. Reflections, e.g., on wet surfaces, also constitute a case where two or more objects are represented simultaneously at the same image location: The reflecting surface itself and the reflected scene.

Perceptual transparency is also related to the representation of a material quality like gloss or roughness which opens up a new perceptual dimension. This dimension is based on, but goes far beyond the more widely studied attributes like color or texture. As

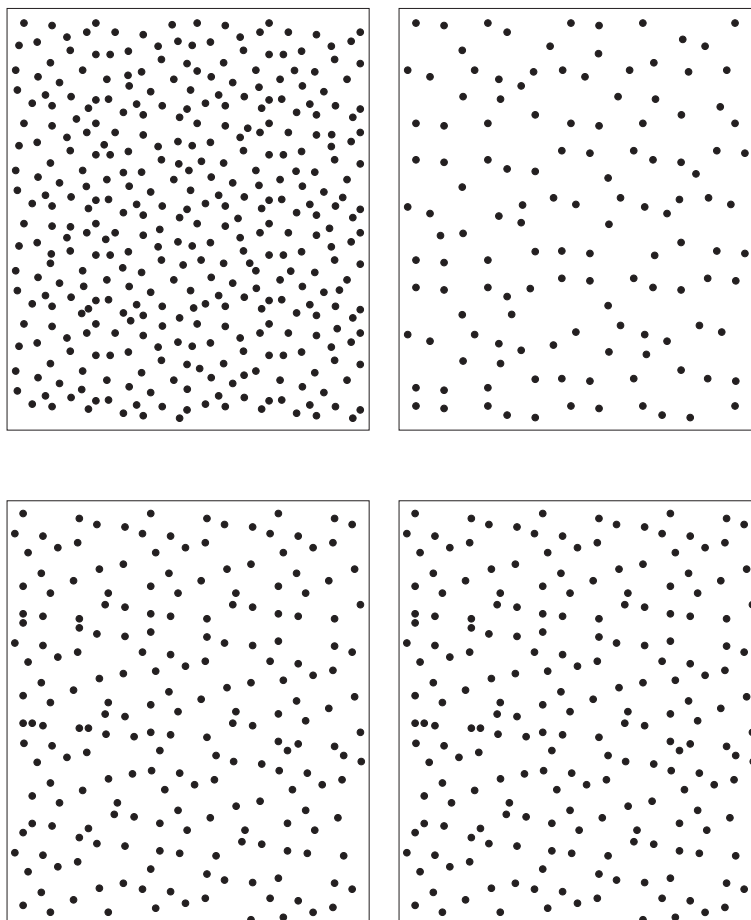


Figure 3.18: Texture density aftereffect. Stare at the top squares for a minute, then look at the bottom squares. The texture density of the left square seems to be noticeably lower than that of the right one. Yet, no additional *individual* elements are perceived.

such, it illustrates how certain global stimulus patterns are exploited to generate a new perceptual quality which is liberated from the physical process of stimulus generation: Although it seems at first obvious to link perceptual transparency to the existence of physically light-transmitting media, it is in fact neither necessary nor sufficient to have these media in the scene to also perceive transparency. Moreover, stimuli can evoke perceptual transparency that are ecologically invalid in the sense that no filter exists that could have this effect. Bozzi (1975), and later Koenderink (2003) provided examples of perceptual transparency resulting from decreased line width (figure 3.19) or increased sharpness in a circumscribed region representing the hypothesized filter. Owing to the dissociation between physical and perceptual transparency, experiments serving in the study of perceptual transparency almost never employ light-transmitting

media in the physical stimulus generation—just as, e.g., 3-D perception is commonly studied with computer-generated flat stimuli.

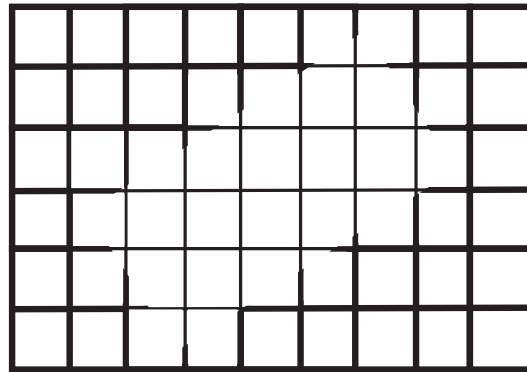


Figure 3.19: Perceptual transparency in response to decreased line width in a two-tone stimulus.

In the study of perceptual transparency, two stimulus aspects that are usually distinguished are the figural and the color conditions for the emergence of perceptual transparency. The color relations between surfaces supposedly seen behind a transparent filter and those supposedly under free view apparently have to adhere to a specific pattern which is being increasingly better understood (figure 3.20c; Faul and Ekroll, 2002). Just as well, there are requirements pertaining to the *spatial* stimulus layout needed to evoke perceptual transparency (figure 3.20a and b). As hinted at in section 2.4, perceptual transparency itself represents a case of grouping image fragments to the end of determining shape and unity of scene objects, a task which is redolent of the fundamental challenge of functionally successful visual completion. Differently-colored regions of the stimulus are combined to form one transparent layer lying over a bipartite background. This amounts to extracting three causal surfaces from an input with four separate regions, a process which is structurally similar to the grouping of several disconnected image fragments into one occluded object. It may thus be concluded that perceptual transparency carries important information about the spatial scene layout: It implies information about the depth configuration of the participating objects, as well as about their surface quality and color. As such, it is not surprising that perceptual transparency figures prominently in visual completion which fundamentally revolves around determining the proper element grouping, and finding a stable spatial configuration compatible with all image aspects.

As has been noted several times, perceptual transparency accompanies some completion phenomena like neon color spreading or flank transparency. Perceptual transparency in these cases is a quality that applies to a completed surface as a whole, and is not evoked by the physically specified fragments themselves when seen in isolation. Color spreading in contrast is clearly traceable to the colored input fragments, although the precise input-output relation is poorly understood. Hence, there is no simple rela-

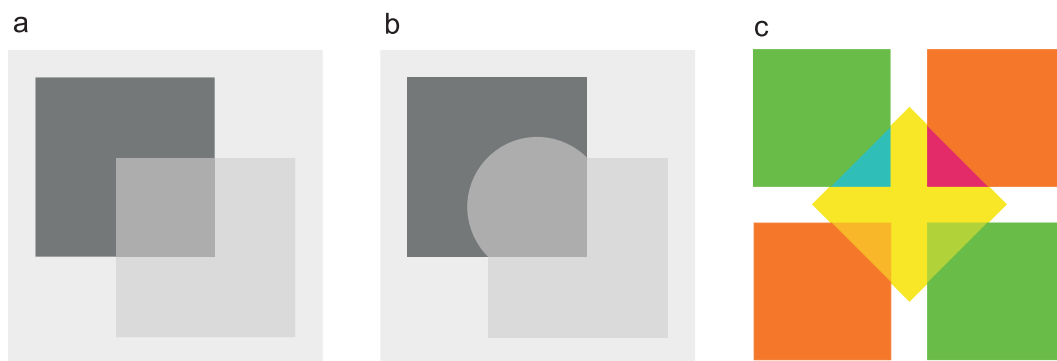


Figure 3.20: Perceptual transparency only obtains if figural as well as chromatic conditions are met. a) and b) Identical coloration only leads to perceptual transparency if no figural cues veto the scission. c) Identical shape information only leads to perceptual transparency if the coloration supports scission.

tion between the image elements and perceptual transparency in completion displays, but perceptual transparency arises cooperatively with other completed features like depth and illusory contours. The relations between perceptual transparency and other features are complex as they can mutually interact without exhibiting a clear causal lineage, a fact which will be taken up again in section 3.2.3.

While transparency serves as a defining attribute of neon color spreading, its intrinsic relationship to the color-spreading component remains a debated issue. Several authors have gathered evidence that stimulus conditions favoring perceptual transparency are necessary for obtaining neon color spreading (Bressan, 1993a; 1993b; Nakayama et al., 1990). Furthermore, the perceptual quality of perceptual transparency arising in neon-color-spreading displays was indeed shown to be compliant with the same quantitative models which were developed to describe perceptual transparency in non-spreading displays (Ekroll and Faul, 2002). More specifically, perceptual transparency in neon-color-spreading displays was optimal under exactly those color conditions that were predicted by a quantitative model previously shown to give a good account of occurrence and quality of perceptual transparency in conventional cases (D’Zmura et al., 1997; Faul, 1997).

Despite the aforementioned results supporting an intimate relationship between color spreading and perceptual transparency in visual processing, there is also evidence hinting at a less tight coupling between these two aspects: The existing instances of color spreading that do not look transparent (see above), for one argue against a general link between all forms of color spreading and perceptual transparency. A second indication along this line is a result which suggests that neon color spreading itself is most pronounced under color conditions which are different from those required for an optimal impression of perceptual transparency: Given the task of picking a color arrangement

in order to create a display eliciting the most distinctive neon color spreading, subject's choices differ systematically from those supporting the most convincing impression of perceptual transparency (Wollschläger, 2002). The two phenomenal aspects of neon color spreading and perceptual transparency thus may pertain to the same completed surface, but may be segregated in visual processing. This of course leaves untouched the interesting finding that perceptual transparency in neon-color-spreading displays is governed by the same rules also applying to perceptual transparency in non-spreading situations.

Depth interpolation. While illusory contours as well as color and brightness spreading often are the most salient features in demonstrations of visual completion, it is important to note that the assignment of explicit depth information to a scene element can also be based on filling-in. A demonstration for this effect, which—despite its simplicity—already illustrates some complexities of the completion process is given by the stereogram in figure 3.21 (Nakayama et al., 1990). Here, the vertical boundaries of the horizontal bar have crossed disparity, but no source of vertical disparity is available in the stimulus. As a result, the horizontal contours of the horizontal bar do not carry any depth information. As pointed out by Nakayama et al. (1990), this configuration is highly ambiguous with respect to the spatial scene layout that might have caused it, leaving room for many different ways of how to interpolate the depth information compatible with the stimulus.

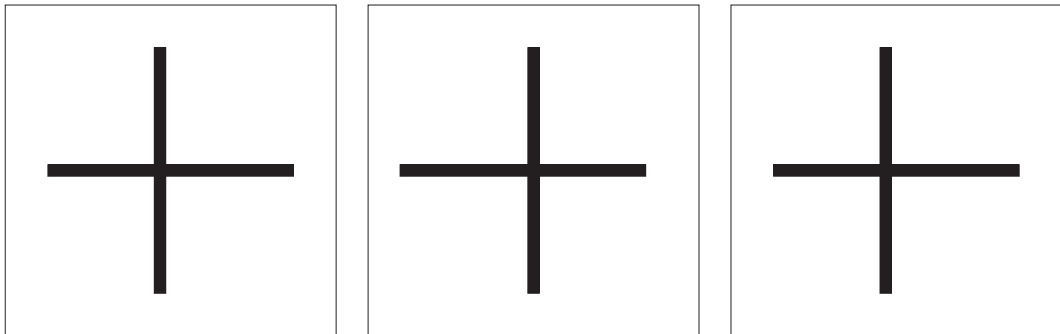


Figure 3.21: Stereoscopic version of the Ehrenstein cross. Only the vertical edges of the horizontal bar carry disparity. Yet, when near disparity is applied to these edges, the whole horizontal bar is perceived to be lying in front. Cross-fusers should use the left pair of figures, diverge-fusers the right one.

Interestingly, the preferred interpretation of the display is seeing a continuous horizontal bar in front of the vertical one, indicating the completion of a constant depth throughout the bar. This means that depth information is not simply interpolated linearly between the available vertical disparity signals, which would have led to the perception of a V-shaped configuration. One reason for the rejection of this arrangement may be the instability of its projected disparity values with respect to small

changes in the vantage point of the observer, i.e., it would require the assumption of a non-generic view (cf. section 2.2).

Another case of constant depth filling-in comes from Bradley and Petry (1977), who devised a subjective Necker cube (figure 3.22) in a stimulus similar to the Kanizsa triangle, that is equivalent to the normal version in its bi-stability of completion: It fluctuates between being modally completed in front of the circles, and being amodally completed as if seen through the circles which work as apertures to the space behind them.

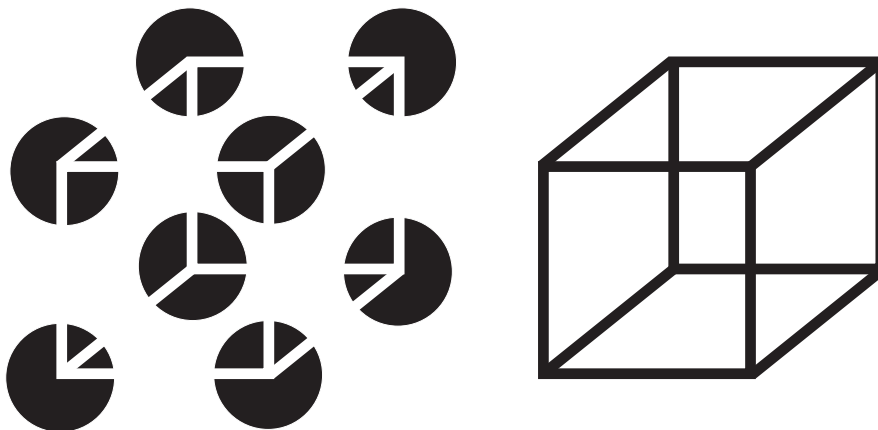


Figure 3.22: Necker cube completed from fragments. The cube may be modally or amodally completed, together with a coupled change in depth position. The completed cube can flip in orientation just as a non-completed one.

Depth filling-in may also take the form of depth capture, as was described by Julesz (1971) for a texture stereogram with black and white dots on a grey background: The black dots make out a total of 5% of all dots. Only some of them have crossed disparity while the remaining ones have zero disparity just like the white dots. Instead of seeing some black dots in front of a black- and white-dotted background, one however sees *all* black dots as lying in front, thereby forming a continuous surface. Note that in this case no ambiguity regarding the depth of the zero-disparity black dots has to be overcome, yet the positively available depth information is apparently overruled by the surface generation process demanding constant depth for the surface.

Nearly all of the considered completion phenomena involve some kind of depth stratification (for exceptions, see Pinna et al., 2001 and Pinna et al., 2004), but so far, each filled-in object either was itself, or belonged to a constant-depth 2-D surface. However, there is no restriction to illusory surfaces of needing to be flat and fronto-parallel. Several demonstrations have shown that subjective surfaces can also bend through different depth planes and thus encompass smoothly varying depth values: Carman and

Welsh (1992) first presented stereoscopic stimuli similar to the Kanizsa square that, when fused, triggered the perception of curved, brightness-enhanced illusory surfaces bounded by subjective contours. Extending this effect to the domain of color spreading, Kojo, Liinasuo and Rovamo (1995) used a similar adaptation of the Varin (1971) neon-color stimulus to induce colored, slightly transparent looking surfaces curving in space.

Also creating curved illusory surfaces on the basis of sparsely distributed disparity signals, Ishikawa and Geiger (2006) more recently confirmed that the perceptually preferred interpolation of surface structure cannot be explained by simple linear interpolation. Extending the conclusions reached by Nakayama and Shimojo (1990b), they also showed that no schemes that assume the generation of minimal surfaces with optimized tension properties (like, e.g., soap films on a wire) are able to account for the observed interpolation preferences. Contrary to the observable depth interpolation in the cross, the assumption of a generic view here does not provide any insights into the preferred perceptual depth interpretation.

A common qualitative aspect of the illusory surfaces considered here is that they all perceptually behave indistinguishably from completely defined surfaces with respect to their view stability: They exhibit shape invariance under changing observation angles (Carman and Welsh, 1992; Nakayama and Shimojo, 1992). This seemingly tangential observation is a sign of the general equivalence of perceptually completed entities compared to normal ones: I will argue in section 5.5 that it allows for important conclusions regarding the question what kind of information constitutes a possible input to various visual processing modules.

Apart from demonstrating that the feature of depth itself can be visually completed in non-trivial ways on the basis of little positive information about it, i.e., from few available sampling points, the presented illustrations also hint at an important generalization of completion as considered so far: It seems that the visual system can test for, and smoothly connect linkable stimulus edges not only in the image plane, but also in depth. With respect to ideas such as good continuation or relatability, this calls into question their 2-D image-based nature, and may require to re-cast them in 3-D terms, i.e., to assume that they accept input not directly from the stimulus domain but from the domain of 3-D perceptual entities. Additional evidence pointing in the same direction will be considered below in the treatment of illusory volumes and amodal completion, together with references to the ongoing efforts along this way (Kellman, Garrigan and Shipley, 2005; Kellman, Garrigan, Shipley, Yin and Machado, 2005; Kellman and Shipley, 1991).

Bounded volumes

The completion phenomena considered so far all have in common that they apply to two-dimensional perceptual entities: Subjective contours normally bound closed sur-

faces which may be qualified by attributes such as depth, color, texture, transparency. Interestingly, perceptually completed elements can also be fundamentally volumetric in nature, as figure 3.23 illustrates. Despite the fact that only the volume-bounding surface can be modally represented, there is a clear sense of inside-outside relationship attached to it which is conveyed by the completed scene. Stimuli eliciting the completion of solids usually are images of a camouflaged, hence invisible volume partly occluding objects that are positioned farther away—configurations closely resembling those that lead to subjective contours bounding two-dimensional surfaces (figure 3.2). Contrary to the interposition-type occlusion depicted most often in boundary completion stimuli, illusory volumes mainly result from conformation configurations (cf. section 2.2; Tse, 1998) where the occluded object wraps around the camouflaged occluding solid. Here, the visual system evidently derives information about the three-dimensional shape of the invisible object through the curvature changes along occlusion edges of the enveloping element. As figure 3.23a shows, such modal illusory volumes need not be accompanied by strong subjective contours, providing one more example for the perceptual dissociation between contours, surfaces, and, in this case, volumes.



Figure 3.23: Examples of volume completion. a) The pole as well as the blob seem to be 3-dimensional volumes, not flat surfaces. Note that no tangent discontinuities are in the display. Adapted from Tse and Albert (1998). b) The figure depicts a Dalmatian that is camouflaged by the background. Once modally completed, the Dalmatian also appears volumetric.

The tangent discontinuities that were found to play a major role in the formation of illusory contours are often absent in conformation displays (figure 2.2c and 3.23a) leading to modal volume completion. Information about the solid's three-dimensional

shape here seems to come from second-order, i.e., curvature discontinuities in the contours of those elements that envelope the object, as discussed in the following section.

3.2.2 Amodal completion

Originally, amodal completion was a label applied to the perception of objects which are *temporally* hidden from view by an occluder, but which do, nevertheless, evoke a subjective impression of perceptual presence, even without having any explicit visual quality associated with them (Michotte et al., 1964). In the context of discussing static completion phenomena, amodal completion is said to take place when *spatially* interrupted regions of the retinal image are seen as parts of the same object that extends behind an occluder, and are connected behind a partially occluding obstacle. Plus, there is no visual quality associated with the occluded object parts.

As a consequence of the missing visual qualities, and unlike the case of modally completed objects, one cannot easily sort amodal completion phenomena according to the visual feature that is completed. In addition, clear evidence needs to be provided that amodal completion really is a phenomenon of visual perception proper, since there is no obvious way of distinguishing it *a priori* from purely cognitive reasoning. However, as will also be further discussed in section 5.5, numerous experiments clearly indicate that amodally completed objects are in many instances as perceptually “real” and functionally relevant as are objects conventionally regarded as completely defined. A counterintuitive implication of the perceptual reality of amodally completed entities therefore is that phenomenal visibility is not an essential attribute of visual object representations. Instead, the intrinsic object concept may have something analogous to a flag parameter which signifies whether the object is visible or not, but which does not otherwise concern the object’s existential status. The important concept of object permanence is most often thought of as applying to the dimension of time, but we see that it is equally well fit to pertain to the dimension of space.

A simple demonstration (figure 3.24) already hints at this fact by showing that when faced with an occlusion situation, we have strongly biased expectations about how the object would look like behind the occluder. As a result, some object shapes that are perfectly consistent with the image data in the occlusion situation surprise us more when revealed under the occluder than others. Still, one might argue that this effect is cognitive in nature. Speaking against this hypothesis are experiments which demonstrate that already young infants are equally surprised as adults to see some previously unseen, in principle consistent, shapes appear behind an occluder while others do not arouse their suspicion, indicating that they were somehow expected from the occlusion situation (for a review, see Arteberry, 2001).

Amodal completion may thus be seen as the representation of object unity, permanence and even shape in the absence of positive information about it. In everyday life, amodal completion is much more frequent than we normally realize, as partial occlusion

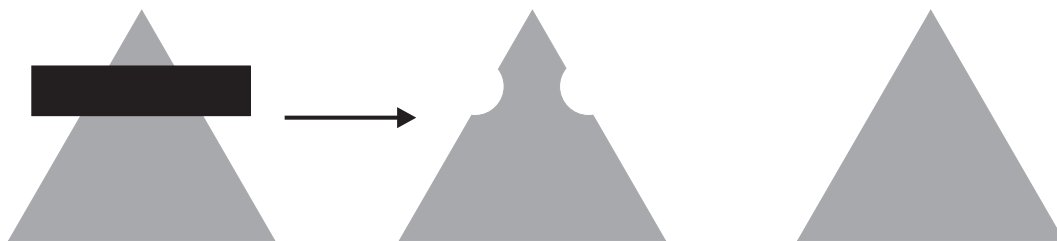


Figure 3.24: Demonstration of amodal completion. After seeing the occlusion configuration, we would be surprised to learn that the shape under the bar was the middle one. Instead, we would expect a normal triangle to be the occluded object.

leading to stimulus fragmentation is abundant (cf. chapter 2). Yet, our conscious perception of the scene is not populated by equally fragmented elements. Indeed, it seems quite hard to even render oneself conscious about what the disconnected fragments are, as introspection and anecdotal evidence from painting techniques indicate: If one tries to realistically draw the image of a three-dimensional scene, it is sometimes much easier to paint the scene with complete objects from back to front, simply overpainting those that are farther behind in the same direction of view, than trying to draw the fragments themselves as they result from the projection.

According to the influential “identity” hypothesis (Kellman, Yin and Shipley, 1998; Shipley and Kellman, 1992), amodal completion of objects is a product of the same rules of boundary integration as modal completion is. Specifically, the identity hypothesis assumes that image tangent discontinuities are the initiating condition for amodal contour interpolation. The type of tangent discontinuity offers cues to border ownership (cf. section 2.2) which in turn indicates which surface fragment neighboring the border is located in front (the fragment owning the border), and which behind and thus lacks boundedness. The fragment located behind the other one is occluded and therefore may not end at said contour, but instead may be linkable to other fragments (Nakayama and Shimojo, 1990b). The concept of relatability (cf. section 3.2.1) is then assumed to provide the conditions which determines whether two abruptly ending edges that belong to unbounded surfaces enter in an interpolation process. As was the case for modal completion, the concept of relatability was successful in a number of cases in correctly predicting the occurrence of amodal completion (Kellman et al., 2001). Just as for the case of modal contour completion, image borders need not be defined by luminance differences to provide an input to the relatability analysis. This is in line with empirical observations, illustrating that assignments of border ownership, and amodal completion can be triggered in random dot stereograms on the basis of disparity differences alone (Nakayama et al., 1989).

According to the identity hypothesis, the integration of boundary fragments into completed contours is a stimulus-driven, bottom-up mechanism. It is conceived of as

a necessary early processing step which leads to the abstract specification of unity and shape of surfaces. The boundary integration and surface completion stages are thereby thought to be independent from the type of completion which later applies to the surface: Unit formation and depth stratification are separable processes (Kellman and Shipley, 1991). In this view, the distinction between modal and amodal appearance then is just a superficial one, representing two modes of appearance of the same visual filling-in mechanisms. This means that given the same image fragments, modal and amodal completion would lead to identical completed contours and perceptual units, differing only in their immediate visibility as a result of perceived depth placement (figure 3.25). Since the identity hypothesis tries to explain findings from different phenomena with the same underlying processing principles, it has the advantage of providing a parsimonious explanation for a wide variety of phenomena.

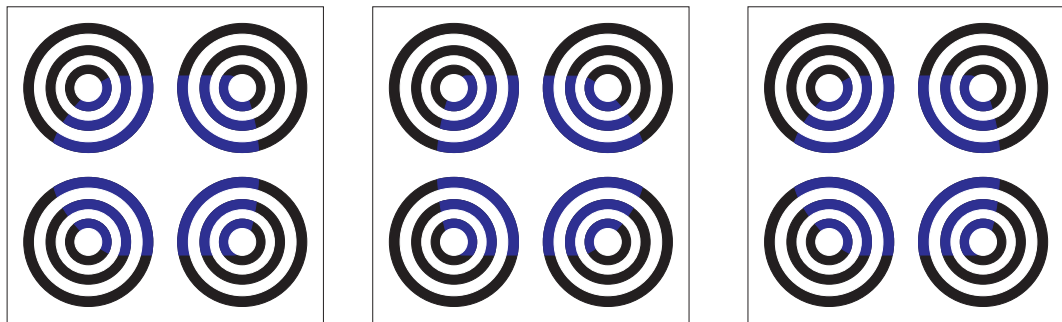


Figure 3.25: Stereoscopic version of the Varin figure. A curved surface can be perceived when fusing an image pair. Fusing different image pairs leads to a different depth positioning of the surface. When the surface is perceived to be in front, it is modally completed. When it is perceived to be behind, as if seen through circular apertures, it is amodally completed. The surface's shape however stays the same.

As it turns out, this supposition is testable because stimuli with an identical spatial structure exist that are either modally or amodally completed, depending on their coloration or binocular depth assignments to fragments (cf. section 3.2.3). In some cases, switching between amodal and amodal completion is possible even within exactly the same bi-stable stimulus, as indicated by so-called spontaneously splitting objects (figure 3.26). It was generally found to be the case that occluded contours arise under the same spatial conditions as modally completed ones do. More specifically, amodal completion often obtained only if abruptly ending image contours forming T- or L-junctions conformed to the relatability criterion discussed in the context of subjective contours. The resulting amodal contours were judged to be equal to the corresponding modally completed ones, and the same number of equally shaped objects resulted from either case—the same perceptual units were formed irrespective of the type of completion (Kellman et al., 2001). In addition, modally and amodally completed

contours can interact to simultaneously define one object by being perceptually linked together, as demonstrated by quasi-modal displays (figure 3.27b).



Figure 3.26: A spontaneously splitting object. A horizontal bar is segmented from a triangle. Sometimes the triangle seems to be in front, and sometimes the bar. Structurally identical configurations can lead to modal completion and perceptual transparency.

According to proponents of the identity hypothesis, these results strongly speak in favor of a shared process of boundary integration and perceptual unit formation (Kellman, Garrigan and Shipley, 2005; Kellman and Shipley, 1991). They argue that contour interpolation logically must precede depth assignment in, e.g., spontaneously splitting objects, and must therefore be independent from later assignments of modality or occludedness. In addition, it might seem cumbersome to hypothesize that two different interpolation processes are invoked in modal and amodal completion which give exactly the same output and can also interact with each other. The identity hypothesis thus indeed seems to be in good accordance with empirical observations.

More recently, however, a growing body of evidence points to shortcomings of the identity hypothesis as well as the principle of relatability itself in fully explaining relevant completion phenomena (Singh, 2004; Singh and Anderson, 2002; Tse, 1998; 1999). As already mentioned in section 3.2.1, some configurations involving only lines and points can create convincing subjective contours (figure 3.8), but are not accessible to the concept of relatability as they have no defined orientation at their endpoints and as such cannot end in tangent discontinuities. As a solely geometric account, relatability theory is essentially ignorant to variations in the color and luminance structure of a stimulus, as long as the pattern of tangent discontinuities remains unchanged: “The unit formation process that we have described as color blind, in that similarity of surface color plays no role in determining whether spatially separated projections will be perceived as a unit.” (Kellman and Shipley, 1991, p.190). Contrary to this notion, Spehar (2000; 2002; see also He and Ooi, 1998) showed that the strength of interpolated edges in the Kanizsa square is significantly reduced if the inducers are inhomogeneous in luminance or color.

With respect to amodal completion, there exist several cases of it in displays with abruptly ending image contours that are not relatable: In some of these situations,

reliability only fails because too much of the conjectured object is occluded, and no edge can be found to pair with one that ends in a tangent discontinuity. As illustrated in figure 3.27a, the smaller rectangle seems to extend a little behind the larger one, even though the reliability condition is not met. This observation was confirmed in experiments conducted by Kanizsa (1979), who found that the size of the smaller rectangle is consistently misperceived by subjects who generally overestimate it.

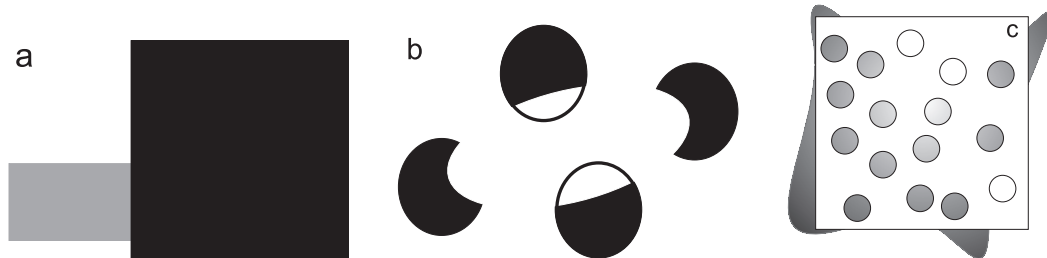


Figure 3.27: a) A rectangle seems to be occluded by a larger square despite the absence of reliable edges. Subjects reliably overestimate the horizontal size of the smaller rectangle. b) Quasi-modal completion where modally and amodally completed contours join to form the complete boundary. c) Surface spreading that generates fuzzy shape information.

In order to reconcile amodal completion in these displays with reliability theory, Yin, Kellman and Shipley (1997; 2000; Kellman and Shipley, 1991) proposed that an independent mechanism of surface-feature spreading complements boundary completion. Surface filling in thereby operates within the extrapolated boundaries of unpaired discontinuous edges. This means that there is supposed to be a general attribute of surface quality which spreads within the linear tangent extensions of unpaired tangent discontinuities, and which can capture patches whose surface features are sufficiently similar in texture and color to the unbounded surface.

In contrast to the amodal completion of boundaries, surface filling in does not generate any definite information about shape—it can only create unity, and in some cases constrain possible shapes a little (figure 3.27c; Kellman, 2003). The sense of unity induced by surface spreading, however, appears to be weaker than that caused by interpolation of reliable edges. The fact that occluded regions are sometimes specified only up to some level of uncertainty about its shape is interesting in its own right since it points to a general capability of the visual system to have fuzzy representations. Moreover, it shows that not only scene attributes themselves can be represented implicitly or explicitly, but also associated “meta-data” which specifies the degree of uncertainty associated with a representation (cf. section 6.1).

The problem of accounting for point- and line-displays eliciting subjective contours as well as the need to conjecture a surface spreading mechanism which complements boundary integration both show that image tangent discontinuities and the concept

of relatability cannot tell the whole story about the visual integration of static image fragments. Still, the two aforementioned cases do not directly contradict the approach in general. More substantial issues, however, were raised by Tse (1998, 1999) that point to deeper problems in it: There are stimuli showing failures of amodal completion despite relatable image edges (figure 3.7b). Furthermore, amodal completion occurs in projections arising from conformation- and penetration-type occlusion configurations (figures 3.23a and 3.28a), where there are no first-order discontinuities in the participating image contours at all, only second-order curvature discontinuities. In some cases, such as the elliptical projection of a smoothly-shaded sphere which volumetrically completes behind the rim,⁴ the boundary is even first-order as well as second-order continuous (Michotte et al., 1964).

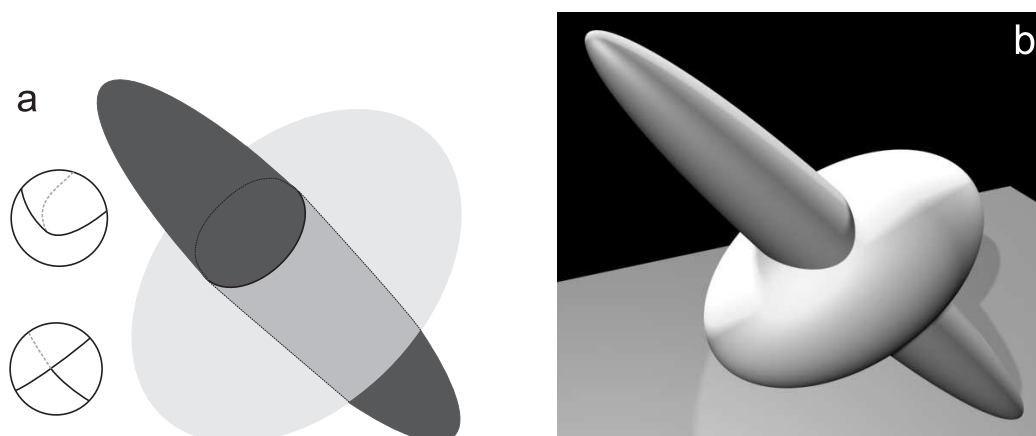


Figure 3.28: Penetration occlusion does not involve a depth discontinuity at the contour of penetration. Generically, the contour of penetration’s projection has no tangent discontinuities.

Occlusion configurations of conformation and penetration are especially interesting because they do not involve a depth discontinuity in the scene between the two participating objects, just a surface orientation discontinuity (Hoffman and Richards, 1984). These occlusion situations are reliably perceived as such, but as a consequence of the missing depth discontinuity, their generic projections deviate systematically from projections of interposition-type occlusion: In the case of conformation, tangent discontinuities only arise from self-occlusion of the bounding volume, often rendering the discontinuities not relatable.⁵ Situations in which occlusion results from penetration generically project to boundaries that have no tangent discontinuities where the contours of the two objects meet. Instead, there is only a second-order, i.e., curvature

⁴The rim is the projection of the set of points where the line of sight is orthogonal to the surface curvature. The rim separates the visible volume from self-occluded space (Hoffman and Richards, 1984; Richards, Koenderink and Hoffman, 1987)

⁵As opposed to often relatable edges belonging to the occluding volume (cf. section 3.2.1).

discontinuity where border ownership changes between them (figure 3.28a). Curvature discontinuities in image boundaries, however, do not arise just from occlusion, but from many different causes, most notably from part boundaries (Hoffman and Richards, 1984; Singh and Hoffman, 2001). This leaves us with the fact that interposition situations of occlusion generically cause image tangent discontinuities and only accidentally create projections without them, whilst a lack of image first-order discontinuities is generic for conformation- and penetration-occlusion (Tse, 1998; 1999). It follows that neither first-order nor second-order discontinuities of image boundaries theoretically are, by themselves, reliably diagnostic cues to the existence of occlusion in the scene. Moreover, empirical evidence suggests that the visual system somehow takes this result into account: It does not have these discontinuities as automatic triggers for occlusion perception and object completion, as supported by examples of failing completion in spite of tangent discontinuities, and of existing completion in the their absence.

Contrary to the influential view that discontinuous image junctions such as T- or L-junctions are direct input-driven cues to occlusion, it thus seems that these are, just like curvature discontinuities, context-dependent cues which are evaluated depending on other perceived scene characteristics. Figure 3.29 depicts an especially intriguing illustration for this line of reasoning as it leads to a case of amodal completion which pre-supposes the generation of perceptual volumes: When taken to be an image of just two-dimensional entities, the stimulus contains no indication of occlusion because, as a consequence of the lacking junctions, there is no cue to an occluder. This indicates that a volumetric representation of the depicted fragments has to precede the image analysis with respect to linkable elements and their hidden contours, suggesting that the idea of relatable scene fragments is probably better placed at a level of surfaces and volumes than at the level of boundaries. Conditions then have to be specified that formalize the analogue concepts of “good continuation” in three-dimensional space, i.e., surface relatability and volume mergeability (Kellman, Garrigan and Shipley, 2005; Kellman, Garrigan, Yin, Shipley and Machado, 2005; Tse, 1998; 1999; 2002).

Since the surfaces or volumes that are bounded by contours are inferred perceptual entities and not image elements, it follows that contour interpolation itself cannot be generally input-driven. It already needs to incorporate scene information resulting from a higher level of analysis, such as surfaces, volumes, and depth stratification. Likewise, the classification of *some* first- or second-order discontinuities as actual occlusion junctions cannot be determined purely locally (Tse, 1998), but rather seems to critically depend on the scene configuration at an intermediate level of representation as evaluated in the context of a number of different cues. The conclusion that occlusion perception cannot be the result of deterministic image cues is also backed by the phenomenon of figure-ground reversible volumes (figure 3.30). Each switch in the assignment of figure and ground implies a change in the occlusion relationship between the two objects. In unison with this reversal, it also entails different object shapes and a different depth configuration of the elements. Furthermore, border-ownership at the shared contours also switches with figure-ground reversals, meaning that it is no

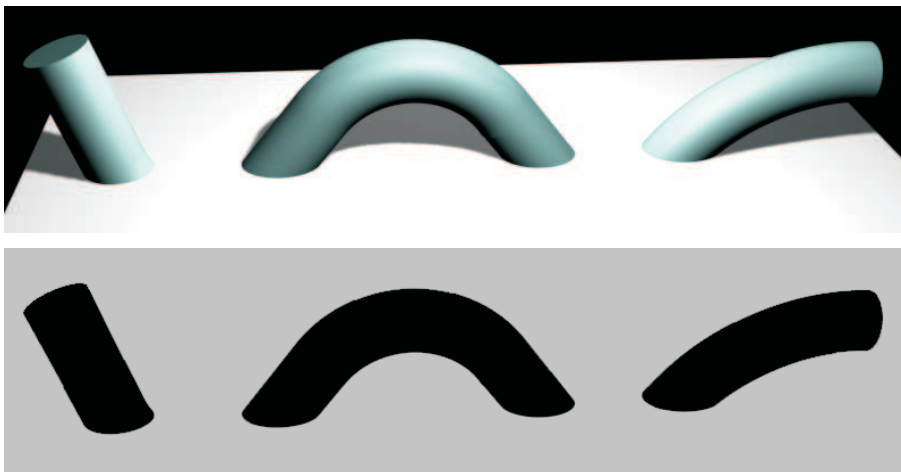


Figure 3.29: Amodal completion in the complete absence of junctions or any first-order discontinuity, and therefore of any two-dimensional cues to occlusion. Only in a volumetric representation, cues to occlusion exist.

direct cue to occlusion relations by itself, as it is often presumed to be (Tse, 1999). Since figure-ground reversal and all implied scene interpretation changes occur in response to the same stimulus, it is evident that occlusion perception is not entirely image-determined, and also not completely pre-attentive or encapsulated from cognitive influences.

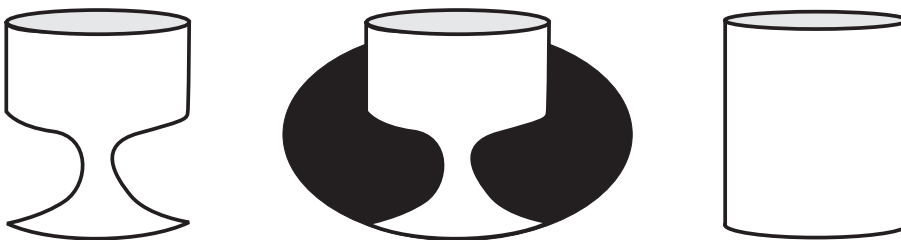


Figure 3.30: Reversible volumetric completion. The central figure is bi-stable and can flip between either being seen as a wine-glass and a conforming part of a torus, or as a cylinder occluded by a wrapping ribbon. Note that the wine-glass is the figure, whereas the cylinder is ground. Figure adapted from Tse (1998).

The fact that relatability systematically fails to capture one class of amodal completion situations points to a possibly deeper-rooted problem of the concept—its limitation to stimulus-driven processing and consequent disregard of the influences of higher-order representations, such as surfaces or volumes. A second problem pertains to the claim of the identity hypothesis that both modal and amodal completion are based on the

same rules of boundary integration and perceptual unit formation. As Anderson et al. (2002) pointed out, there is a theoretical asymmetry between scenes leading to modal and amodal completion: Modally completed objects result from cases of camouflaged elements placed in front of other elements and thus occluding them, whereas amodally completed object parts are perceived to be occluded. Thus, they are further away than the occluding elements. While camouflage critically depends on the color characteristics of the scene, occlusion does not. In line with this distinctive difference, and contrary to the claim by Shipley and Kellman (1992), edge interpolation for modally completed objects was shown to depend on color relations in the image (Anderson et al., 2002; see also He and Ooi, 1998; Spehar, 2000; 2002), and is thus constrained by mechanisms which do not affect amodal completion. In addition, Anderson et al. (2002; Singh, 2004) were able to show that, depending on the type of completion, interpolated boundaries were the same, but different perceptual units resulted from structurally identical stimuli (figure 3.12). Finally, Anderson et al. (2002) and Singh (2004) devised stimuli which even lead to different completed boundaries, according to whether completion is modal or amodal.

The interpretation of the empirical results given above sparked a debate about their theoretical relevance for the concept of relatability and the identity hypothesis. Perhaps not surprisingly, the proponents of relatability claim that they constitute only very weak evidence against their line of reasoning, and that relatability theory in combination with surface spreading can readily account for the vast majority of demonstrations given by Tse, Singh, and Anderson (Kellman et al., 2001; Kellman, Garrigan and Shipley, 2005). So far however, they have failed to provide a more detailed argument or analysis which could support this claim. Consequently, it seems that the empirical results presented here still suggest that, although the concept of relatability fits well with many perceptual completion phenomena, significant deviations from its predictions can be found. Meanwhile, Kellman, Garrigan and Shipley (2005) recently tried to reformulate the concept of relatability and place it in 3-D space, without, however, elucidating the obvious connection to Tse's (1998; 1999; 2002) ideas, or commenting upon the drastic switch from being an image-based criterion to a percept-based one.

The two main ingredients in the identity hypothesis, namely a shared boundary completion as well as unit formation process, have both been challenged with counterexamples. The latter result hints at the possibility that the identity hypothesis fails because it is negligent to the regularities associated with the projective geometry of occlusion and its implications for the depth ordering of scenes. Anderson et al. (2002) infer that relative depth provides strong constraints on what features participate with what results in the process of object completion (cf. section 3.2.3). They further conclude that substantial changes of perceived surface structure, e.g., concerning the number and form of objects, can occur without being influenced by contour interpolation mechanisms. Together with the findings from volume completion briefly hinted at above, this represents additional evidence that contours may prove incapable to provide the basic building blocks in modeling the perception of completed objects. The traditional view of surface interpolation assumes that surface features like color or tex-

ture merely spread within interpolated contours, and thus complement the more basic boundary completion (Yin et al., 1997). It now rather seems that intermediate-level perceptual concepts such as surfaces and volumes themselves are the main input into object completion, possibly owing to the fact that a representation of depth structure plays a major role in visual processing.

One reason why limitations of the concepts of contour relatability, of the identity hypothesis, and of the emphasis on contours in general have gone unnoticed for so long, may be grounded in the long-standing preference for simple two-dimensional-looking stimulus material. Although this minimal-stimulus approach was undoubtedly able to uncover a great many phenomena as well as important characteristics of perceptual completion, it may also have clouded the view regarding more involved interdependencies in the processing mechanisms. It thus seems appropriate to further develop the use of more complex stimulus classes that provoke unified percepts of interpolated boundaries, filled-in surface features and depth assignments to study the intrinsic intertwined rules of completion mechanisms. This approach will be presented in more detail in the following section.

3.2.3 Phenomenal interactions involving statically completed features

So far, the point was repeatedly made that many illusory percepts are gradual in nature, as they can appear in different levels of strength or conspicuousness. In contrast to this continuous variability, the following account will demonstrate that, depending on how other scene attributes are set, completed features can also change abruptly, exhibiting qualitative changes in perceptual appearance rather than smooth transitions from one state to another.

Since visually completed features—be it contours, color, or depth—are in a sense illusory, one might hypothesize that they are merely unimportant epi-phenomena, detached from ordinary, functionally relevant visual processing. Quite to the contrary, several observations indicate that illusory features are not isolated at all, but are completely tied into the normal perceptual mechanisms: They interact in non-arbitrary ways with each other as well as with features that seem to be more directly derived from the retinal input. In particular, visually completed scene attributes may provide input to other processing mechanisms. While the following account highlights phenomenal interactions involving completed features, section 5.5 later addresses also measurable effects of completion in performance tasks.

The upcoming presentation of the complex pattern of lawful co-dependencies involving illusory features revolves around the question about what constitutes the intrinsic logic of the visual system regarding the ways in which perceptual primitives can interact and be coupled in the formation of objects. This touches the general issue of cue coupling (Yuille and Bülthoff, 1996) but will also raise the deeper question what might be considered a basic visual feature, or perceptual primitive in the first

place. To anticipate, many reasons will argue against the global usefulness of the classical, physiologically-inspired separation between attributes of, e.g., color, motion, and depth. Instead, it will prove fruitful to consider, e.g., surfaces located in depth as a fundamental visual unit. After relevant insights from the study of dynamic completion will have been presented, this question will also be addressed more directly in section 6.1.

Depth and contour completion. One of the most robust effects of coupled features in visual interpolation exists between assignments of depth and the type of perceptual completion, i.e., whether it is modal or amodal: As discussed above, the perception of interpolated contours as either being modally or amodally completed entails a corresponding depth attribution. Modally completed contours are lying in front, amodally completed ones are perceived to be in the back of the scene. The strong correlation between type of completion and depth layer can be observed within one fixed bistable stimulus, where perceived depth and type of completion always switch in unison, e.g., in spontaneously splitting objects (figure 3.26). Upfront, it should be emphasized that completed distinct surfaces always have a special depth position, probably meaning that the intrinsic surface concept has this attribute as a mandatory parameter.

If the depth assignment can be manipulated directly, e.g., through changes in binocular disparity, reversals between modal and occluded perception of contours happen in dependence on disparity settings. Examples for this effect were given by, e.g., Nakayama et al. (1990) for the stereoscopic Kanizsa square and the modified Ehrenstein figure (see also Kellman and Shipley, 1991). Taking the Kanizsa square and giving the vertical contours a crossed disparity results in modal completion of the square and amodal completion of the inducers, whereas an uncrossed disparity leads to the perception of an occluded square seen through modally completed holes.

The explicit assignment of depth through disparity can be a potent cue that is apparently able to overrule other sources of information pertaining to the depth layout of the scene. As was already discussed in sections 2.2 and 3.2.1, the junction structure of a retinal image provides cues to border ownership, which in turn can specify what side of a contour should be placed in front of the other one. The T-junction is a prime example for this, as the cap generically owns the border and thus makes the two sides of the stem unbounded and placed behind. If, however, binocular disparity is set such that it conflicts with this interpretation, i.e., if the cap carries an uncrossed disparity, it overrides the junction structure. This turns the stem sides into bounded surfaces that cannot link up with relatable edges, and accordingly, no completion between them occurs (Nakayama and Shimojo, 1992; Nakayama et al., 1989). Depth thus vetoes interpolation processes that would normally be triggered had the display been viewed monocularly. In addition, this example shows that border ownership may not always be straightforwardly determined from monocular image cues, a result that already appeared in section 3.2.2.

While the coupling of depth and type of completion supports the notion that interpolated features stand in lawful relations to each other, it does not by itself attest to the claim that completed features can serve specific purposes in functional visual processing. Many arguments in favor of this idea will be put forth in section 5.5, but two illustrative cases shall already be considered in this context: In the aforementioned study of Nakayama et al. (1989), performance effects of completion—or the absence thereof—were tested in a face recognition task. Face fragments were presented with occluders between them that promoted amodal completion of the face if placed in front, and prevented it if placed in the background. Face recognition was significantly better in the completion condition, suggesting that a unified face percept aided in the recognition process. Davis and Driver (1997) studied attention spreading on fragmented surfaces. This method can be used as a way of probing the functional unity of image fragments because cued attention spreads within same-object regions but does not spill over to other objects. Davis and Driver (1997) manipulated the type of completion of the surface fragments through depth-placement, and found that attention only spreads on near, modally completed surfaces, but not on far, occluded ones. Both of these studies show that the same available input information can be used by the visual system in functionally very distinct ways, depending only on perceptual changes regarding aspects of completion.

Depth, color spreading, and transparency. If completion concerns not only boundary, but also surface aspects, these are often also affected when depth assignment is manipulated. Just as boundaries, color spreading and transparency, too, have a close relationship with the three-dimensional scene interpretation: In the modified Ehrenstein figure or in the Varin square, color spreading and transparency are absent if the figure is placed behind and thus amodally completed, whereas both aspects abound if the figure is lying in front and has modally completed contours (Nakayama et al., 1990). Watanabe and Takeichi (1990) also observed that transparency, color spreading, and type of contour completion all change in unison depending on disparity settings in variants of the modified Ehrenstein figure. In a similar vein, Meyer and Dougherty (1987) observed that for flicker-induced, non-stereoscopic depth, transparency, color spreading and type of completion always switch in unison with depth assignment.

In some configurations, the functional relevance of color spreading and transparency appears to be powerful enough to prevent other common percepts from occurring. If a modified Ehrenstein figure is presented binocularly with a colored cross embedded into only one of the two stimuli, the two images do not correspond, and one expects binocular rivalry to occur. According to Nakayama et al. (1990), a rather different percept obtains: The absence of disparity notwithstanding, a modally completed colored transparent filter is sometimes seen as lying in front of the larger cross. This percept fluctuates with seeing the embedded cross as being part of the background layer, where no color spreading, transparency or modal illusory contours are seen. These percepts are thus again correlated and switch in unison, notably without any changes to the

stimulus, i.e., in response only to purely perceptual changes.

Despite the aforementioned results seeming analogous to the tight correlation between depth and contours, the precise relationship between depth, transparency, color spreading, and contours is by far not as straightforward as has been described so far. This conclusion will be supported by the following cursory presentation of relevant findings. It should, e.g., be noted that in the binocular modified Ehrenstein figure, only the horizontal bar is assigned a crossed or uncrossed disparity, whereas the vertical bar is not shifted and thus has zero disparity. If, however, the horizontal bar has crossed disparity, a transparent layer is formed that captures the depth of the vertical bar which forms part of the transparent surface and as such is seen as lying in front, too (Nakayama and Shimojo, 1992).

While transparency can itself capture the depth of other scene parts, it can also block depth capture that occurs without it (Nakayama et al., 1990). Transparency can therefore be said to determine the depth positioning of other scene elements: If the element is fused with the transparent surface, its depth is set to that of the near surface, if it is segmented from the transparent layer, then depth capture that would raise the element from the background is vetoed. In other cases of stimuli with equal disparity signals, transparency, as manipulated by changing global figural or color cues, can also actively determine perceived depth of scene elements, not just by the promotion or suppression of depth capture (Nakayama et al., 1990). It follows that the occurrence of transparency is not simply passively determined by depth, it rather seems that the exactly reversed dependence relationship is also possible.

Similar to the prevention of depth capture, transparency can also suppress the formation of illusory contours, as is again apparent in the modified Ehrenstein cross. If the embedded cross is binocularly seen without the larger cross, and crossed disparity is assigned to the horizontal bar, this bar is seen in front, clearly delineated from the vertical bar by modal illusory contours. If, however, the larger cross is added again, the aforementioned percept of a transparent layer is seen, i.e., the horizontal bar is fused with the vertical bar, and both are part of the transparent surface. Even though both stimuli are identical with respect to the embedded cross, no contours separate its horizontal and vertical bar in the transparency condition, whereas subjective contours are present in the isolated condition (Nakayama et al., 1990).

For the general relation between depth, color spreading, transparency, and illusory contours, this leaves us with a somewhat murky picture. Not only do the cited studies indicate that, on the one hand, depth assignment can determine color spreading and transparency, and that, on the other hand, transparency can determine perceived depth and contours, it also appears that all these perceptual features can be dissociated completely: As was already discussed in section 3.2.1, color spreading can occur without illusory contours (Redies and Spillmann, 1981; Redies et al., 1984; Watanabe and Sato, 1989) as well as without transparency. And while the transparency perception in neon color spreading displays was successfully accounted for by models created for conventional perceptual transparency (Ekroll and Faul, 2002), the optimal color conditions

for neon spreading do seem to deviate systematically from those for a most convincing transparency impression (Wollschläger, 2002). Perceptual transparency itself does not rely on color spreading or illusory contours anyway. One of the few reliable constraints seems to be that color spreading only appears transparent when it is part of a modal surface, i.e., when it is also bounded by subjective contours. This condition is violated in the watercolor effect (Pinna et al., 2001), as well as by the diffuse color spreading cases, both of which elicit no perceptual transparency.

Color spreading and figure-ground segmentation. Color spreading has so far appeared as a feature that is detached from functional processing, and merely stands out for its ephemeral phenomenal quality. As it turns out, color spreading can be an important factor in determining the grouping of image fragments, and thus the figure-ground assignment in a scene. Pinna et al. (2001), Pinna, Werner and Spillmann (2003) found that fragments linked through watercolor spreading tend to be grouped together to forming the figure of a scene, even if powerful Gestalt influences such as proximity or symmetry would indicate otherwise.

Wollschläger et al. (2002) suggested that, taken together with the considerable spatial extent of the color spreading, this key role in figure-ground organization offers some insight into coloring techniques used by Renaissance map makers during the late 16th and the 17th century. In some maps from this epoch, regions are painted in outlines only with flanking colored lines alongside them (figure 3.31). This “outline-color technique” is prominently featured, e.g., in Joan Blaeu’s maps (Bagrow and Skelton, 1985; Shirley, 1983). The result is a layout that looks strikingly similar to watercolor effect stimuli. Map makers evidently employed color-spreading effects in a systematic way to ensure that countries are easily silhouetted against each other without having to be drawn in solid colors. Although the induced color spreading must overcome frequent “barriers” in the form of location names—and is therefore somewhat weak—the effectiveness of the technique is evident in original maps made this way.

Occlusion relations and motion perception. When motion interacts with occlusion, the visual system not only faces the challenge of having to establish the correspondence between objects that are stable in shape and fully visible. It then also has to classify different fragments that are visible at different moments in time as belonging to the same object, and thus as representing a potential indicator for object motion (this is discussed in more detail in the subsequent section 3.3). Since having the attribute of being occluded makes an image fragment unbounded and potentially linkable to other fragments in space or time, it is predestined to be involved in motion perception of partially visible objects. The depth configuration of a scene as well as figural stimulus aspects can both give cues about occlusion, and therefore are prime candidates for being a determinant of motion perception. Nakayama et al. (1989) as well as Shimojo and Nakayama (1990) indeed provided convincing evidence that motion perception can change dramatically dependent on the evident incompleteness of an object resulting



Figure 3.31: Map drawn by Joan Blaeu (1663) using the outline-color technique. By permission of The British Library: Maps.C.5.b.1.

from occlusion. More specifically, they were able to influence what object in a scene is perceived as stationary, and what as moving, purely by a change in depth assignment through binocular disparity. This result, however, has an important qualification: If figural cues conflicted with an occlusion relation, disparity changes had no effect.

Conclusions. The experimental findings on perceptual interactions briefly reviewed here all suggest, their diversity and multi-facetedness notwithstanding, that there are substantial interactions between perceptual codes for scene attributes that are conventionally treated separately. Small changes to one perceptual attribute often result in big qualitative changes in others, pointing to discontinuous cross-feature processing in visual perception. This suggests that it may not be adequate to consider aspects such as depth, shape, and color as distinct perceptual primitives, as they seem to be too closely intertwined. This of course raises the deeper question what then might be considered a basic visual feature, or perceptual primitive instead. Hinting at a possible solution to this question, the presented interactions do not seem to be arbitrary or accidental couplings, but are very systematic, especially when considered in the context

of surface perception. These surfaces always come with a distinct depth assignment that is lawfully tied to other of their attributes. Therefore, surfaces underline the fundamental importance of perceptual representations of three-dimensional entities—even in what may be considered early aspects of visual processing (cf. section 5.5).

As a second important conclusion, there seem to be virtually no restrictions in the internal architecture of the visual system regarding what kind of feature representation can supply an input to other processing mechanisms. Illusory or completed features may equally well provide the informational basis for determining other scene aspects as image-based features do, just as the completion of visual qualities itself rests upon cues which *are* extant in the stimulus. Where stimuli lead to a bistable phenomenology, pure perceptual changes can initiate the switch of other visual qualities in the absence of any image input to do so, thus being functionally equivalent to them (cf. section 5.5).

Essentially the same interactions that were considered here will be met again in dynamic displays throughout the following sections. Naturally, special emphasis will be placed on the involvement of motion in the spatiochromatic interactions between color, depth, and subjective contours.

3.3 Dynamic completion

While the preceding section already covered a large variety of perceptual completion phenomena along with associated empirical results and theoretical models, it was limited to stimulus situations associated with static scenes. However, since one motivation for the study of perceptual completion is derived from observations about the highly fragmented nature of the visual input in *natural* viewing environments, the temporal dimension that is inevitably operative in a realistic stimulus generation process ultimately needs to be taken into account as well. Occlusion, one of the most important causes for image fragmentation, directly interacts with observer as well as with object motion (cf. section 2.5). This implies that the ability of the visual system to successfully retrieve object information from the incoming stream of information must rest on mechanisms that go beyond a mere spatial analysis of the input. The importance of these dynamic mechanisms for interacting with our environment is emphasized by experimental results indicating that they are already operative in infants of 2 months age (see, e.g., Arteberry, 2001; Johnson, 2002).

Much like occlusion is a source of stimulus fragmentation, but can, once determined, also impose important constraints on possible spatial scene layouts, motion, too, may be seen both as a challenge and an opportunity for the visual system. Knowledge about object motion can help in gathering object information from disconnected input regions: Certain parts of an object may be blocked from view at one point in time, but be visible moments later, depending on the motion trajectories of itself, other objects and the observer. If the observer in principle possesses the ability to accumulate

information pieces that are available only sequentially over time, the rules governing motion might provide the constraints required to obtain information about the objects' trajectories. These in turn could be used to spatiotemporally group the input flow into same-object chunks. Another important motion-related source of information is that its flow pattern is ordinarily smooth within contiguous objects, but is likely to be discontinuous across different, autonomously moving scene elements. Common motion components and motion-defined borders are thus theoretically good cues for determining the connectivity between spatially separated regions or to test for necessary segmentation within otherwise homogeneous areas.

The previous remarks underscore the need to direct our attention to the dynamic aspects of interpolation processes for gaining a better understanding of perceptual completion, and consider motion as a genuine influence on it. Research about motion processing itself is indeed one of the main roots in the study of dynamic object completion (Morgan, 1980; Morgan, Findlay and Watt, 1982). When concerned with the spatiotemporal input characteristics that allow the visual system to collect and integrate object information sprinkled across space and time, one may place different emphasis on either dimension: Research perspectives originally grounded in the perception of incomplete static stimuli may regard time as an influence in the stimulus generation process that impedes the process of identifying temporally invariant relative geometric properties of the scene. While occlusion *spatially* separates chunks of information about a contiguous object, motion draws corresponding input events *temporally* apart from one another. Motion therefore needs to be discounted in order to get at the stable geometric layout of a scene as it unfolds in time, just as it would be apparent in an unobstructed static snapshot (Feldman 1985; Kellman et al., 2001; Palmer, 2003). This train of thought forms the basis of spatiotemporal relatability theory discussed in section 3.3.2.

If coming from a background of motion or event perception, one may instead consider time as a distinct quality itself. Time imposes its own input regularities, e.g., for binding spatially unrelated fragments into one event by temporal correlation, or for causing segmentations according to the purely temporal relations between spatially contiguous input fragments. The importance of timing for the grouping of discrete input fragments into coherent, moving objects can already be appreciated in φ -motion.

As aspects of boundary formation and perception of surface features are often intertwined (for a review, see Bruno, 2001), no strict division exists between the empirical study of dynamic boundary interpolation and dynamic surface-feature completion. The following review of dynamic completion will still use this aspect to provide a useful ordering structure, because many earlier studies placed special emphasis on boundary aspects and were mostly negligent to surface characteristics. These in turn received more attention later on in experiments that used—with a different focus—a very similar stimulus class as research on contour interpolation. The following exposé of dynamic completion therefore starts with general findings and theoretical ideas gathered from

experiments that are mostly concerned with aspects of motion-induced boundary formation. This account then provides the basis for an in-depth treatment of dynamic surface-feature spreading in the following chapters 4 and 5. The account of dynamic surface interpolation, together with aspects surfacing at various points in this chapter, will directly lead to the theoretical questions empirically addressed also in chapters 4 and 5.

3.3.1 Boundaries and surfaces from spatiotemporal discontinuities

From early on, questions concerned with the integration of spatiotemporally disconnected stimuli have mainly revolved around the possibility to identify shapes despite a high degree of fragmentation (Hecht, 1924; Parks, 1965; Rock, 1981; Rothschild, 1922; Wallach, 1935; Zöllner, 1862). Possibly owing in part to technical limitations of the stimulus-generation process, this interest was accompanied with a restriction to achromatic, two-dimensional line drawings as stimulus material and an almost exclusive focus on boundary integration. Three experimental paradigms emerged as the most influential over the years: Situations of dynamic occlusion typically simulate the translatory or rotational motion of a camouflaged, hence invisible object in front of background elements of variable size (figure 3.33a and b). Originally coming from the early development of cinematography (Kalkofen, 2000; Plateau, 1836), another stimulus class simulates the motion of an object behind narrow apertures which uncover only very small portions of it at each point in time. Finally, a third stimulus category employs abstract spatiotemporal discontinuities to delineate otherwise invisible borders, a situation which does not necessarily conform to any possible scene configuration in natural environments.

Dynamic occlusion. Situations of dynamic occlusion may elicit percepts of both modal and amodal completion. As reviewed by Michotte et al. (1964), one of the earliest demonstrations was the so-called tunnel effect which involved the amodally completed identity of an object that seemingly moves behind a rectangular occluder and re-appears at the opposite end of its entrance (figure 3.32a). The progressively disappearing object at one end and the appearing object at the other end of the occluder are perceived to be identical despite belonging to spatiotemporally separated elements. Moreover, the object seems to maintain its shape, despite the deformations of its contour as it is successively covered and uncovered by the occluder. While the object is invisible, a clear sense of its unity and shape is preserved, together with information about its current position on the motion trajectory (Burke, 1952; Michotte et al., 1964). This is also true for non-linear trajectories, i.e., when the connecting line between entrance and exit points needs to be curved if smoothness is stipulated. A smooth motion path is indeed perceived, even when entrance and exit point are not relatable and inflections in the connecting line are required. The conclusion that the visual system represents information about the trajectory of a temporarily invisible

object was later confirmed several times using a number of different non-introspective methods (Demkiw and Michaels, 1976; Scholl and Pylyshyn, 1999; Shiori and Cavanagh, 1992; Shiori, Cavanagh, Miyamoto and Yaguchi, 2000). As will be discussed further below (cf. section 3.3.2), this bears importance for current models of spatio-temporal boundary formation which require the visual system to have this kind of knowledge (Palmer, 2003).

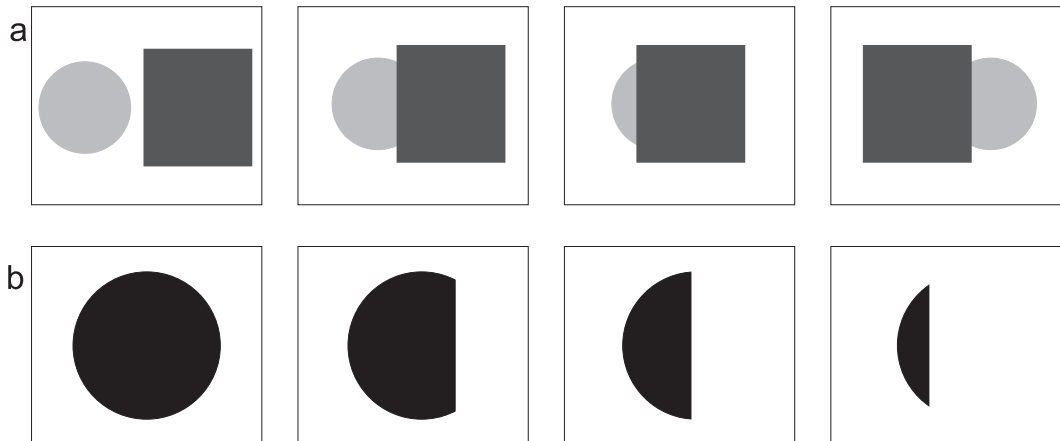


Figure 3.32: a) The tunnel effect. Frames from an animation which simulates the dynamic occlusion of a moving disk by a stationary square. b) Frames from the screen-effect animation which simulates the occlusion of a moving disk by an invisible square.

The amodal-completion effect described above also persists if the occluding tunnel is itself camouflaged by the background and thus invisible (figure 3.32b; Sampaio, 1943). This “screen-effect” configuration promotes an additional effect of subjective contours—the edges corresponding to the entrance and exit points of the occluded object are modally perceived as slits in the background. The simultaneous presence of modal and amodal completion is reminiscent of static demonstrations like the Kanizsa triangle, where the inducing pacmen are amodally completed to circles while the triangle is perceived modally. In addition to modal and amodal completion of subjective contours, the tunnel as well as the screen effect induce a perceived depth stratification that is also found in the more complex demonstrations described below: The amodally completed surface is always perceived as an entity that is detached from the background and is assigned a depth layer behind the occluder.

Similar to the first versions of dynamic occlusion, later demonstrations also used stimuli that simulated the gradual occlusion of bigger homogeneous objects by an invisible occluder: Kellman and Cohen (1984) for example animated the Kanizsa figure such that the invisible triangle rotated over a set of black shapes (figure 3.33a). Similar to the tunnel effect, the animated setup induces the simultaneous modal completion of the triangle and amodal completion of the disks. While some edge segments of the

occluder are physically specified at one time or other in this configuration, the occluder also contains segments that never have a visible effect on the background shapes—these segments, too, are modally interpolated if the display is animated. Importantly, no illusory contours are perceived in each still view, even though some edge information is contained in a static image as at least one disk is partially occluded at each point in time.

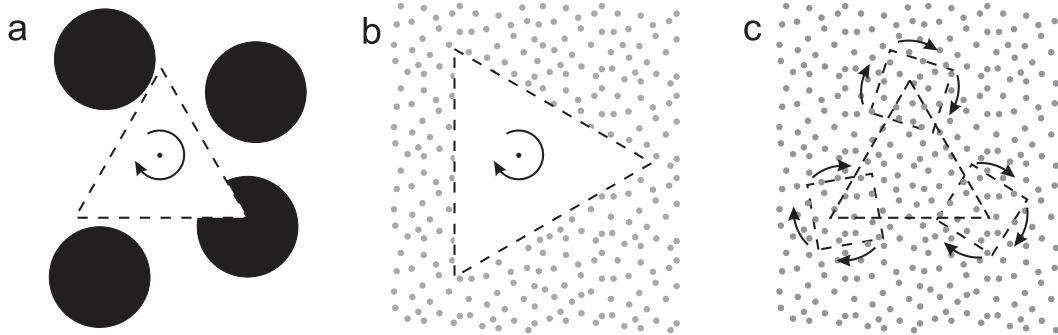


Figure 3.33: a) Dynamic occlusion of disks by a camouflaged, hence invisible rotating triangle. b) Dynamic occlusion of a background texture by an invisible rotating triangle. c) Kinetic induction of illusory contours through dynamically-defined inducers.

The absence of illusory contours in the static condition sets Kellman and Cohen's (1984) stimulus apart from earlier demonstrations of moving virtual contours with a Kanizsa triangle version that used inducing lines and points (cf. section 3.2.1; Bradley and Lee, 1982): Here, each stationary view also elicits the perception of subjective contours, albeit less compellingly than when put in motion. Using thin lines instead of solid background shapes as inducers, Bruno and Bertamini (1990; Bruno and Gerbino; 1991) later provided a straightforward extension of Bradley and Lee's (1982) dynamic-occlusion stimulus where no static view alone is sufficient to trigger the perception of subjective contours.

The dynamic occlusion stimuli presented by Kellman and Cohen (1984) and Bruno and Bertamini (1990) may require motion to elicit visible contour interpolation, but they still contain static edge cues. One may therefore object that the visual completion in these cases is not a matter of response to purely dynamic occlusion. Several other displays inducing dynamically completed contours do indeed get by without such static cues by using only randomly-placed small texture elements in the background that are covered or uncovered as a whole (figure 3.33b). As a consequence, there are also no moving edges—at least on a fine spatial scale—providing motion information. This is in contradistinction to the stimulus setups described above, where there are moving edges caused by the partial occlusion of large background objects or lines. Picking up results from static filling-in, the input to dynamic completion mechanisms need not be luminance based: In a dynamic-occlusion variant schematically displayed in figure

3.33c, the imaginary triangle figure rests at its place, but is dynamically specified by visually completed squares, which are themselves only defined by relative motion to the background (Kellman and Loukides, 1987).

If the background texture is sparse enough, the shape of the occluder seen when moving cannot be interpolated precisely in still view (Andersen and Cortese, 1989; Bruno and Bertamini, 1990; Hine, 1987; Shipley and Kellman, 1994). Likewise, if the occluder is covered with the same texture as the background—e.g., in random dot kinematograms (Gibson et al., 1969; Kaplan, 1969), there is no static shape information at all: Neither does a static texture density gradient between occluded and unoccluded region exist, nor do the two areas differ in average luminance or color. The dynamic information coming from the common motion of the occluder texture, as well as the pattern of accretion and deletion of background texture elements alone suffice to give cues to the occluder's shape. This is illustrated by the instant break-down of camouflage when a hidden object starts to rigidly move (Metzger, 1975). Finally, if an occluder with a small area is sandwiched in between two layers with a random dot texture, one closer to the observer and another farther away, it dynamically occludes only texture elements of the farther layer (Cunningham, Shipley and Kellman, 1998a; Stappers, 1989). If such a display is seen at rest, no occluder is perceived between occluded and unobstructed background area because, by raising the overall texture density, the nearer layer prevents the texture gradient from becoming, according to Weber's law, big enough to be noticeable.

In spite of their differing construction principles, all of the stimuli described above were used to demonstrate that articulate subjective contours are perceived, and surprisingly precise shape perception can be retrieved from spatiotemporally separated occlusion events alone. This result convincingly shows that boundary information can not only be interpolated over spatial gaps but also be accumulated over time. The cited studies either used tasks of shape discrimination with a set of basic, differently curved geometric shapes (Andersen and Cortese, 1989; Hine, 1987; Prazdny, 1986; Shipley and Kellman, 1994; Stappers, 1989), or simply employed an introspective method and asked subjects if they perceived an object with definite shape (Kellman and Cohen, 1984; Michotte et al., 1964). Those studies that varied the parameters of texture density and velocity all agreed that shape perception improves with the number of occlusion events, either achieved through higher element densities or higher velocities (Andersen and Cortese, 1989; Bruno and Bertamini, 1990; Bruno and Gerbino, 1991; Shipley and Kellman, 1994). The observed increase in accuracy may be the result of more articulate subjective contours, whose clarity and salience was found to vary continuously with texture density, again indicating that the perception of subjective contours is not an all-or-none phenomenon (cf. section 3.2.1). Interestingly, when the animation of dynamic occlusion is stopped, subjects often report that as the contours vanish, they do so in a continuous manner, as if they were slowly faded out (Hine, 1987). The same observation will resurface in the discussion of dynamic surface spreading (cf. section 4.3).

Some of the cited studies (Bruno and Bertamini, 1990; Bruno and Gerbino, 1991; Kellman and Cohen, 1984; Prazdny, 1986) tested for and commented on a possible asymmetry in shape perception with respect to the type of relative motion between the invisible occluder (the figure) and the occluded texture (the ground): Bruno and Bertamini (1990) found that in cases of rotational figure motion, the figure's shape was not clearly delineated by subjective contours, and was also not perceived to be stable. Instead, it appeared to deform over time, similar to the breathing square illusion (Bruno, 2001; Meyer and Dougherty, 1990). In contrast, figure vs. ground motion did not produce different shape perceptions with translatory motion. Prazdny (1986) already stated that only weak subjective contours were reported in the rotational figure motion case, but Kellman and Cohen (1984) merely found a small asymmetry in reported shape perception between rotational figure and ground motion conditions. Evidence for a dissociation between these two conditions is interesting because the relative motion signals in both cases are identical. This result suggests that shape perception in dynamic occlusion may not be entirely governed by low-level image-based information, but may rather be influenced by figure-ground assignments resulting from global cue integration.

In a similar vein, Prazdny (1986) found that explicit occlusion events from accretion and deletion of texture elements were essential to an accurate shape perception, and a mere difference in relative motion between figure and ground was not sufficient to elicit subjective contours in the animated Kanizsa triangle: In one condition, the rotationally moving inducers were covered by a static random dot texture, whose elements retained their position throughout the animation. The inducer rotation only made visible different portions of this texture depending on the rotation position, as if acting as an aperture to the texture behind it. The background was covered by a randomly moving dynamic white noise, thus producing motion discontinuities at the inducer edges. In this case, no shape perception obtained. This is in contrast to the condition with reversed assignments of static random dot texture and dynamic white noise, which featured the same motion discontinuities, but also produced occlusion events with the static background texture elements being occluded by dynamic white noise of the rotating inducers. Kellman and Loukides (1987) later confirmed that kinematically specifying the inducers of the Kanizsa triangle in a random dot pattern can be functionally equivalent to luminance-defined pacmen in inducing the modal completion of the triangle (figure 3.33c).

Bruno and Bertamini (1990) found that translatory motion overall led to a higher level of accuracy in shape perception than rotational motion and speculated about a possible connection to the difficulties of performing ocular tracking on rotationally moving objects. This difficulty arises because the eyes cannot rotate fully around the axis perpendicular to the sagittal plane. As a consequence, eye movements do not, in this case, provide any information about the moving object's position as opposed to the situation with linear motion trajectories. This line of reasoning will be continued with a more detailed account on the possible role of eye movements in dynamic perceptual completion in section 5.2.

Although it was only rarely remarked in the original studies (for an exception, see Bruno, 2001; Kellman and Cohen, 1984), many of the traditional dynamic occlusion stimuli are consistent with a different depth stratification and assignment of figure and ground than the conventional one. Normally, the displays are interpreted as having an occluding figure in front of occluded background shapes. But what appears to be an occluded inducer may also represent a hole, or, in the case of inducing lines, a slit in a surface that is formed by what is normally perceived as the background. The occluding figure then becomes a shape that is seen through the apertures and is itself occluded by the surface (figure 4.1). This physically possible arrangement is indeed sometimes perceived, together with a simultaneous switch from modal to amodal completion of the figure, and a reversal of the conventional depth assignment: The formerly occluding figure is now located behind the surface, whereas the formerly occluded elements are closer to the observer.

The bi-stable character of this kind of display is also typical for other examples of static and dynamic completion, as was already described in sections 3.2 and 3.2.3. The different interpretations of dynamic occlusion displays reveal that it simply is a special case of so-called aperture viewing displays that will form the basis of the empirical investigations in this work, and will be discussed in depth in chapters 4 and 5.

General spatiotemporal discontinuities. While a dynamic occlusion display with a textured background usually has a straightforward interpretation in terms of a corresponding natural scene that projects to the respective stimulus, it can also be analyzed without any reference to physically possible stimulus generation processes. The element transformations caused by a simulated occlusion event, i.e., sudden or progressive disappearance, may then be regarded as just one particular spatiotemporal element transformation among many theoretically conceivable ones: As an imaginary moving object passes over a stimulus element, that element may, among others change color, position, orientation, or shape upon entrance into, and exit from the imaginary object's bounded area. These changes neither need to correspond to physically possible transformations nor need they be the same for all elements. Figure 3.34 illustrates this principle with randomly placed black and white background elements on a grey background. Whenever the imaginary circle passes over a white element, its color changes to black, and to white again when the element leaves the area of the disk. If the element is black, it changes to white within the circle and back to black outside of it. Likewise, each element may change its orientation or position randomly once inside the passing object and return to its previous state once outside of it again.

Using the general type of textured dynamic transformation display outlined above, Shipley and Kellman (1993a; see also Cicerone and Hoffman, 1992; Cicerone, Hoffman, Gowdy and Kim, 1995; Kellman et al., 2001; Shipley and Kellman, 1994) extensively studied the question whether abstractly specified information fragments provided in these displays can be spatially and temporally integrated into percepts of object features. In analogy to geometric discontinuities found in static displays (cf. sections

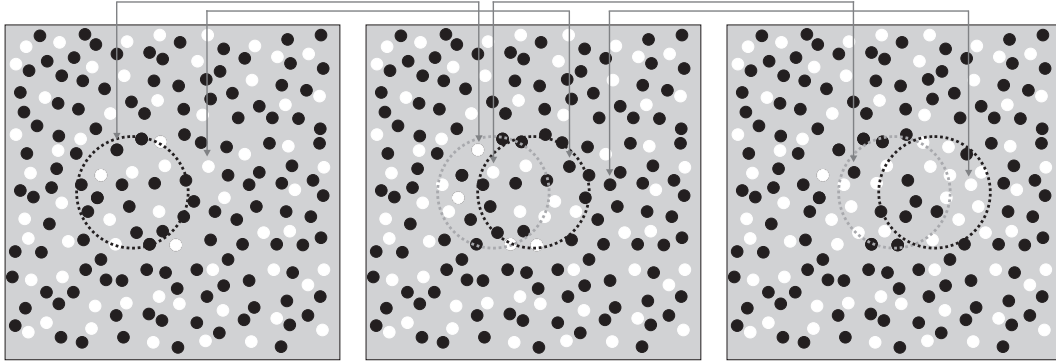


Figure 3.34: Frames from an animation involving bi-directional spatiotemporal discontinuities. Once inside a moving virtual disk, stationary white dots change their color to black, and black ones to white. This change is reversed once the dots are again outside the virtual disk. Arrows indicate examples of color change from one frame to the next.

3.2.1 and 3.2.2), they termed the element changes “spatiotemporal discontinuities” if they suddenly disrupt the projective identity of the elements (Kellman and Shipley, 1991). Such a discontinuity is thus said to occur if no continuous scene transformation of the element could cause the change in its displayed projection. Simply put, all stimulus changes that cannot result from rigid scene motion are considered a spatiotemporal discontinuity. As mentioned above, textured dynamic occlusion displays are one example of these discontinuous transformations, in this case disappearance and reappearance. Since all elements in dynamic occlusion are subject to the same change, it constitutes an example of so-called uni-directional spatiotemporal discontinuities, whereas the display given in figure 3.34 represents a bi-directional change.

Using shape discrimination tasks, Shipley and Kellman studied a variety of spatiotemporal-discontinuity displays that incorporated uni- as well as bi-directional changes in color, orientation, and position. They generally found that subjects can accumulate fragmented shape as well as surface information and readily identify the moving imaginary object in most spatiotemporal-discontinuity displays. This object usually appears to be bounded by sharp subjective contours—a phenomenon, which they called “spatiotemporal boundary formation”. Even though a bi-directional spatiotemporal-discontinuity displays is an ecologically invalid stimulus, it also succeeds in inducing some object qualities like boundaries and global rigid motion. This effect is reminiscent of Bozzi’s (1975; Koenderink, 2003) observation that impressions of transparency can result from displays that feature regional stimulus changes which could not possibly be the result of a filter application (cf. section 3.2.1). Replicating standard results for dynamic occlusion, it was generally found that higher signal element densities improve discrimination accuracy in uni- and bi-directional spatiotemporal-boundary-formation displays (Cunningham, Shipley and Kellman, 1998a; Shipley and Kellman,

1993a; 1994).

Observers also report the presence of a special surface appearance corresponding to the completed object for a smaller set of spatiotemporal-discontinuity stimuli. More specifically, it appears that all of the studied uni-directional *color* change displays can evoke crisp subjective contours as well as a surface appearance of the imaginary object. For some uni-directional displays, the surface impression includes even more sophisticated percepts: If, e.g., on a white background all elements within the imaginary object change their color from black to red, the moving surface is perceived to be homogeneously colored. It also has a transparent quality (Cicerone et al., 1995), similar to that found in neon color spreading (cf. section 3.2.1). Chapters 4 and 5 will focus on these surface quality changes in dynamic completion displays. It is instructive to also consider those stimulus manipulations that Shipley and Kellman (1993) reported as unfit to initiate contour and surface integration: Small positional displacements of the texture elements for example did result in the perception of a moving object, supposedly because the changes were considered within-surface deformations, not between-surface discontinuities.

In contrast to uni-directional displays, the bi-directional color change displays failed to induce the perception of a rigid surface and only elicited illusory contours of weaker clarity than the uni-directional displays (Shipley and Kellman, 1994). As indicated by a substantially inferior level of shape discrimination performance, the object transformations in orientation and position induced weaker subjective contours than changes in color. No performance difference was found between uni- and bi-directional displays. The last results may be interpreted as possibly constituting a hint at the relevance of ecological validity of the stimulus: Uni-directional displays that are mostly consistent with physical situations of occluding or transparent media seem to promote the perception of object features more convincingly than bi-directional displays that are incompatible with physics.

When changing between being set in motion and being stationary, spatiotemporal-discontinuity displays undergo several qualitative changes with respect to the induced phenomenology. Given the right stimulus parameters, local motion signals are instantly grouped and integrated into the global motion of a newly-created, phenomenally present object. This object is perceived to be of definite shape, and also possesses an explicit depth assignment. Halting the display, all these interpretations fade out, and only single static texture elements are perceived.

Just as it is necessary to analyze what kinds of perceptual impression can be induced by dynamic spatiotemporal-discontinuity displays, it is also instructive to bring to mind what possible scene interpretations do *not* arise: Notwithstanding the fact that the stimuli are animated through individual changes of single texture elements—be it that they disappear, change color or orientation—there is usually no local motion perceived (Shipley and Kellman, 1997). Instead of registering all change events separately, they are grouped and ascribed to one causal influence, i.e., the moving object. This is the same result as was found for dynamic completion displays, where coherent

boundary motion is perceived in favor of local element motion (Bruno and Gerbino, 1991). Consistent with the implied importance of grouping the different motion vectors, shape discrimination performance degrades with the injection of spurious local motion components independent from the occluding object (Shipley and Kellman, 1997). This finding may be interpreted as another indication that shape perception can strongly rely on motion analysis.

When spatiotemporal-discontinuity parameters do not favor the formation of a coherent object, observers do have conscious access to local motion signals. This suggests that in the case of global motion, they are actively prevented from becoming phenomenally available (Shipley and Kellman, 1993a; 1994; cf. section 5.5). In the sense that the object is now simultaneously responsible for all changes, just one global percept exists, and informational access to individual elements is preempted.

Note that for uni-directional change displays of the described type, there exist static cues to the location of the imaginary object because of a difference in mean luminance, or because of a different texture inside and outside of that object. In bi-directional displays, there is none such static difference. Bi-directional change displays are also set apart from random dot kinematograms simulating the accretion and deletion of texture resulting from kinetic occlusion, because in that case, all elements within the moving surface share motion components, i.e., have the same common fate (cf. section 2.5). As such, the area enclosed by the imaginary object in bi-directional spatiotemporal-discontinuity displays in no way differs from the surrounding regions, and local element changes at the imaginary object's borders alone specify its shape. The lack of consistent features that may differentiate the enclosed area of the moving object from the background may be seen as an obvious candidate for explaining the failure to form a surface representation in that case. This topic will be given more consideration in the following section 3.3.2 together with other theoretical ideas developed in response to spatiotemporal-discontinuity displays.

While spatiotemporal-discontinuity displays have so far been regarded from the perspective of shape and surface completion, the fact that observers perceive smooth motion of the imaginary shape also deserves attention. This is because, strictly speaking, there are no moving elements—neither in uni-, nor in bi-directional displays. However, because of the static luminance differences in uni-directional color change displays, a spatial frequency low-pass version of the display would exhibit a stimulus border between the area inside the moving shape and outside. An input into luminance-based motion detection systems would thus be provided at the borders, as it generates a Fourier-motion component in the direction of the moving imaginary object (Shipley and Kellman, 1997). In contrast to this extant first-order motion signal (Sperling and Lu, 1998), bi-directional color change displays do not contain any purely luminance-based motion information (Shipley and Kellman, 1997)—in the terminology of Chubb and Sperling (1988), they constitute drift-balanced stimuli. Nevertheless, local motion signals at the borders may still be picked up by second-order motion detectors (Sperling and Lu, 1998). But only if these local signals are properly integrated can they explain

the percept of global object motion which is observed also in bi-directional change displays. This “correct” integration of distributed motion signals is all but trivial and not given *a priori* in the visual system, as illustrated by the breathing illusions discussed in the preceding paragraph on dynamic occlusion. Therefore, motion integration and shape perception are inseparably tied together with a need to be cooperatively specified as parameters of the same object.

3.3.2 Spatiotemporal relatability theory

Consistent with the findings about static boundary completion discussed in section 3.2.1, the aforementioned results about spatiotemporal-discontinuity displays demonstrate that edge information need not be luminance based in order to enter the object completion process as a source of boundary information. In addition to information coming from disparity, texture, or color, it rather seems that edges can also be defined based on information from purely spatiotemporal stimulus discontinuities, with static spatial discontinuities completely missing. In spatiotemporal relatability theory, the sudden breakage of projective identity of an element now plays a role equivalent to a tangent discontinuity in static displays: It serves as the initiating condition for contour interpolation and feeds information into the same interpolation mechanisms as static discontinuities do. Empirical evidence coming from experiments with general spatiotemporal-discontinuity displays seems to lend support to this idea. Small texture-element displacements as well as gradual element transformations were found to be inadequate to trigger the perception of clear illusory contours (Shipley and Kellman, 1993a).

The definition of a spatiotemporal discontinuity and its presumed role in triggering the contour completion process immediately leads to the question how this dynamic interpolation is actually performed by the visual system. A natural attempt in trying to figure out boundary interpolation processing for spatiotemporal-discontinuity displays is to try to connect it to mechanisms of static contour integration that may already be better understood. Static situations can, after all, be conceived of as the limiting case in the continuous velocity range of dynamic scenes. Thus it seems unreasonable to assume two separate sets of perceptual mechanisms which redundantly serve the same goal of shape perception given a highly fragmented visual input. Trying to reduce the problem of dynamic contour interpolation based on general spatiotemporal discontinuities to the problem of static contour interpolation based on spatial discontinuities indeed represents the path that spatiotemporal relatability theory has chosen to attack the problem (Kellman et al., 2001; Palmer, 2003).

Spatiotemporal relatability theory was explicitly formulated as an extension of relatability theory for static boundary integration (cf. section 3.2.1) to dynamic stimuli, and thus focuses on the temporally stable geometric properties of dynamic scenes. However, for the established criterion of relatable edges to be of use in dynamic interpolation of image fragments, several obstacles have to be cleared. Simply searching

for, and then connecting reliable edges simultaneously present within each frame is not always possible and can lead to spurious fragment connections: As too much of the object can sometimes be occluded, what is visible as a fragment in a given frame might not have a reliable paired fragment within that frame, but only in the following one (figure 3.35a)—this is essentially the situation of dynamic occlusion displays. Moreover, visible fragments within one frame might not actually belong to the same objects, so they should not be blindly linked even if they are reliable.

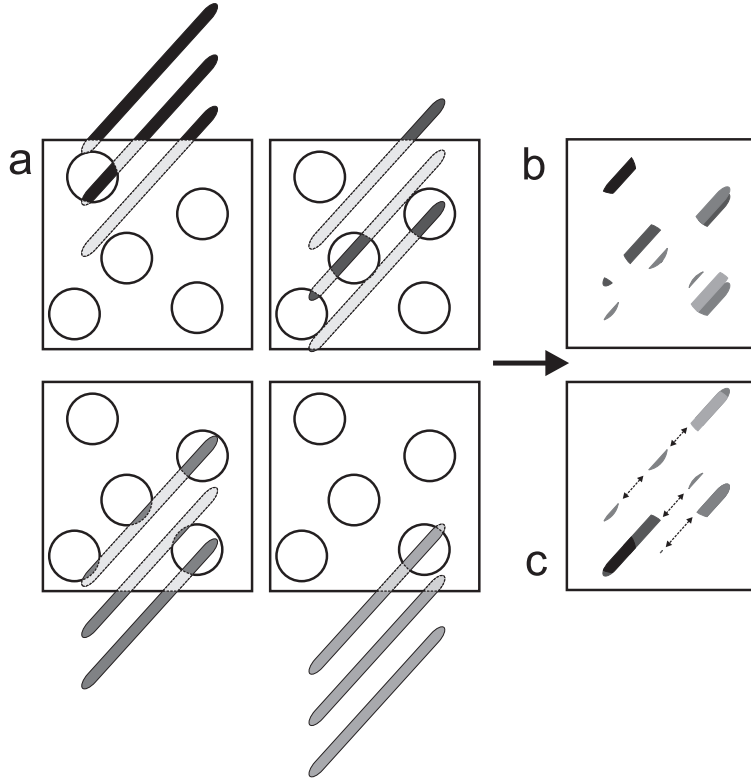


Figure 3.35: Illustration of spatiotemporal reliability. a) Four frames simulating the dynamic occlusion of three tilted bars by a mask with circular apertures. b) Simple temporal summation of all frames does not lead to a recovery of the bars. c) If the bars are tracked perfectly by the eyes, temporal integration leads to reliable fragments whose completion can recover the bars. This is equivalent to the persistent representation of visible fragments within a distal frame of reference. Figure adapted from Palmer (2003).

An idea to overcome the problem that linkable object edges are not specified simultaneously but need to be registered sequentially might be to apply reliability analysis to a temporally integrated version of the dynamic scene. But even if visual information somehow persists in accessible form for some time, multiple successive frames cannot

simply be integrated by temporally adding them into one frame (figure 3.35b). This would only give a relatively accurate representation of the scene geometry and allow for a relatability computation if a stationary observer looks at static objects, and the moving occluder uncovers the object consecutively over time. Or, what would be equivalent with respect to the proximal stimulus, the static observer might also perfectly track the moving object behind a stationary occluder, thus keeping the retinal image of the object stable (figure 3.35c). If neither of both conditions are satisfied, however, mere temporal summation will not lead to a reconstruction of the scene that is a useful input to relatability analysis.

With the goal of transforming the dynamic signal into a stable static representation of its relative geometry, it is thus necessary to precisely single out those aspects of dynamic displays that need to be discounted before a static relatability computation can be sensibly performed. Not surprisingly, the crucial element is motion itself. In the two examples of successful temporal summation, object motion is precisely cancelled, either because it is absent in the scene from the very beginning, or because it is nullified in the proximal stimulus through pursuit eye movements. Knowledge of object velocity in both cases is the key to obtaining information about the relative geometry of the scene. If one hypothesizes that information about the object's motion trajectory is somehow available to the visual system also in situations different from the two cases considered above, the same effect of removing motion can be achieved through motion extrapolation: At each moment, the position of stored invisible object fragments are updated to the location where they should currently be according to the estimated motion velocity and direction. This way, all fragments belonging to one object keep their relative position because the individual fragment's location becomes independent from the position it was actually registered at. Partial dynamic information about object fragments at varying locations (in absolute measures) is thus transferred to stable relative position information within a static reference system that is integrated over time (figure 3.35c). This is what was called a "stable features frame" before (Feldman, 1985). The reference frame is thus assumed to hold two sets of fragments from the dynamically occluded object: Those portions that are momentarily visible plus those that were registered before and are still accessible through some kind of visual buffer. Importantly, the stored fragments enter the reference frame with their positions updated according to motion extrapolation, to the end of keeping relative position between same-object areas constant. This processing scheme thus incorporates temporal summation, but only after motion extrapolation and subsequent position updating.

According to spatiotemporal relatability theory, a static analysis for relatable edges, however originally defined, is then thought to take place within the distal representation containing the accumulated information. Still, there may be object parts that are, even over time, never fully specified, as is the case for some dynamic occlusion displays (Kellman and Cohen, 1984; Palmer, 2003). However, assuming that the static reference frame contains enough accumulated edge information, the interpolation of missing boundary segments can be performed exactly like in the static case, i.e., between

relatable edges that abruptly end in a discontinuity, be it spatial or spatiotemporal in nature. In summary, as Palmer (2003) puts it, spatiotemporal relatability theory thus essentially combines informational persistence and position updating based on motion extrapolation with boundary integration between relatable edges as defined for static stimuli.

As spatiotemporal relatability theory requires correct motion assessment for extrapolating the current position of previously registered and stored invisible fragments, it needs to assume that local motion signals can be reliably integrated to yield the global object motion velocity and direction. Local motion signals are thereby thought to be the result of pairing two spatiotemporal-discontinuity events, thus building up a motion vector field. How this vector field is exactly integrated into coherent global motion is of no concern to spatiotemporal relatability theory proper. It entails, e.g., solving the aperture problem (Shimojo and Richards, 1986), and the possibility to disambiguate the uncertainty met when trying to trace back a complex motion field to a rigid object motion. In spite of a large body of research (Bruno, 2001; Hildreth and Koch, 1987) and undisputable advances, mechanisms of motion signal integration are all but well understood (Bruno and Gerbino, 1991; Shipley and Kellman, 1994; 1997).

Even though motion perception resulting from the integration of distributed motion signals in some cases systematically deviates from the global motion used to generate the stimulus, it remains an empirical observation that humans can reliably estimate motion paths in many real-world situations: The ability to play, e.g., Baseball, Tennis, or Racquetball strongly attests to this prowess. In addition, experimental evidence suggests that observers have access to motion information about temporally invisible objects that are still represented modally or amodally (Burke, 1952; Demkiw and Michaels, 1976; Michotte et al., 1964; Scholl and Pylyshyn, 1999; Shiori and Cavanagh, 1992; Shiori et al., 2000).

Just as spatiotemporal relatability theory requires the extraction of global motion, it presumes that stimulus information is stored in some form that makes it accessible at a later time. Only this informational persistence allows the cumulative integration of edge information into the proposed stable reference frame. So far, spatiotemporal relatability theory itself is mute on how exactly this visual short-term storage is achieved, although this issue is currently being worked on (Palmer, 2003). In any case, numerous empirical results provide clear evidence that access to past visual information is possible in many tasks for a short amount of time, even if conclusions as to the characteristics and implementation of the storage are not unequivocal (Coltheart, 1980). One mechanism that is often discussed in this context is simple sensory persistence, also termed retinal painting. Retinal painting together with perfect smooth pursuit eye movements would, in theory, eliminate the problem of visual storage in many instances of stimulus integration. The possible influence of retinal painting represents one of the starting points for the experimental work presented here, and will be discussed in depth in section 5.2.

Limitations. As a straightforward extension of static relatability theory to dynamic stimuli, spatiotemporal relatability keeps many of the former's limitations, the main one possibly being that it is still rooted in the 2-D image domain of the visual input. Consequently, its applicability to the dynamic interpolation of 3-D objects is as of yet unclear, and a large body of well-known demonstrations involving the dynamic interpolation of volumes escapes this approach: Structure from motion, the kinetic depth effect, biological motion perception in point-light walkers, stereokinetic figures, and the rotating Necker cube seen through apertures (Bruno, 2001) are just some examples demonstrating that the representation of rigid objects in 3-D space can be triggered by a suitable combination of very reduced stimuli containing only highly fragmented shape and motion cues. This again implies that the spatial three-dimensionality figures early on in visual processing, probably being an inseparable aspect of surface and object representations. Being reminiscent of Tse's (1998; 1999) analysis of the static relatability concept with respect to capture his demonstrations of volume completion (cf. section 3.2.2), this means that spatiotemporal relatability theory probably needs to be re-cast in perceptual entities, not stimulus ones.

Furthermore, spatiotemporal relatability formally does not handle the induction of illusory contours and surfaces from the dynamic interruption of line drawings, as in the animated Kanizsa figure of Bradley and Lee (1982). This is because the concept of sudden breakage of projective identity does as such not apply to thin lines. Every thin line in the image is a possible projection of the same straight line in 3-D, shortening or lengthening of the projected line may therefore be the result of smooth scene transformations (Bruno, 2001).

Another result reported in the dynamic occlusion literature is also not handled by spatiotemporal relatability: Despite the same relative motion, and therefore the same spatiotemporal discontinuities, figure and ground motion do not seem to have identical effects (Bruno and Bertamini, 1990; Prazdny, 1986; cf. section 3.3.1).

3.4 Preliminary synopsis

As we have seen, phenomena of perceptual completion are legion. They encompass a multitude of static and dynamic stimulus situations, and concern many different perceptual aspects. Their study has unearthed evidence of rich interactions between commonly separated features of visual perception, and thus offers intriguing insights into the internal processing architecture of vision. However, the nature of the intrinsic relationship between different kinds of completion remains elusive. There is hitherto no common theoretical framework which could tie together all the fragmented bits of evidence and integrate the empirical findings from the various partaking research domains.

To single out only a few of the loose ends in need of being tied up, one might mention that the chromatic constraints for color spreading, be it in neon color spreading or in

flank-induced ones have as of yet not been precisely analyzed. Evidence arguing for the importance of certain contrast patterns is often only anecdotal. Furthermore, when systematic experimental findings exist, their results tend to be idiosyncratically tied to a specific stimulus, and thus difficult to generalize to other situations.

While the study of color relations can still be embedded into a formal frame of analysis, no such established formalism exists for structuring the influence of figural aspects, at least beyond junction classification. Still, figural aspects are apparently of utmost importance to surface integration. Their interactions with chromatic conditions of object specification, however, are only beginning to be studied (see, e.g., Nakayama et al., 1990). The same is true for aspects of motion processing. So far, it is unclear how motion and color work together in creating dynamically-derived surface percepts. Only contrast-reversing bi-directional dynamic displays have received considerable attention. A more fine-grained analysis why bi-directional displays do not afford surface integration but uni-directional ones do, is still missing.

Another big shortcoming of the current state of affairs is the lack of convincing methodology to assess completion phenomena without being susceptible to response biases and to other uncertainties when dealing with data gained in purely introspective tasks. Put bluntly, nailing down subtle effects like neon color spreading has proven to be a non-trivial goal. In addition, aspects of boundary and surface completion are often confounded in widely-used tasks like shape identification.

Evidently, there are still several open questions awaiting answers in the study of dynamic completion. The following empirical sections will try to pick up some of the problems that have surfaced during the preceding review of visual completion in space and time. In particular, it will focus on aspects of surface-feature completion, dynamic spatiochromatic interactions, and the development of an adequate methodology.

Part II

Experiments

Chapter 4

Dynamic color spreading

While the preceding part of this treatise provided the general theoretical background of dynamic completion, more concrete theoretical aspects and relevant empirical findings need to be discussed in order to prepare for the experiments carried out. To this end, the immediately-relevant theoretical aspects are considered in direct conjunction with the particular experiments they apply to. One main focus of the forthcoming sections is the question of adequate methodology with which to approach completion phenomena. Although this aspect plays a role throughout the forthcoming chapters, section 5.5 and experiment 5 (cf. section 5.6) are especially concerned with it.

But first, sections 4.1, 4.2, and 4.3 will describe in detail the development of a new kind of flexible stimulus type belonging to the class of dynamic-color-spreading displays. This stimulus will, in different variations, be used in all presented experiments. The experimental sections 4.3.2, 4.3.3, and 4.3.4 serve to demonstrate the particular robustness of induced filling-in effect, which makes this stimulus material a useful tool in quantitative studies of visual completion. I will report a first set of studies aimed at, so to speak, putting the phenomenon under consideration on stable empirical grounds. As such, the experiments are intended to determine the basic psychophysical characteristics of the stimulus type, and shall show its general accessibility to quantitative methods.

Section 5.1 will extend the experimental display to the induction of non-homogeneous surface completion in the form of dynamic texture spreading. Building upon the empirical results from experiments 1 to 3 and additional theoretical considerations, section 5.2 will then be mainly concerned with the discussion of simple perceptual mechanisms potentially responsible for the observed surface filling-in. Experiments 4a and 4b were designed to test some straightforward predictions of these theoretical ideas, and will be reported in section 5.3. Their results will prompt for a deeper analysis of some temporal aspects of visual processing of normal and completed stimuli in section 5.4, and for a second look at common experimental methods with their underlying assumptions in section 5.5. The experiments brought forward in sections 5.6 and 5.7 try to incorporate some of the lessons learned from examining several methodologies, and the results will argue for a measurable, functional performance benefit of dynamic surface completion. With section 5.8, the empirical part will conclude by giving an integrative recapitulation of the results obtained in experiments on dynamic texture spreading.

4.1 Aperture viewing

As was pointed out in the previous sections, dynamic-occlusion stimuli as well as a subset of spatiotemporal-discontinuity displays may all be regarded as instances of so-called multi-aperture viewing, a generalization of the classical anorthoscopic perception situation. Strangely, articles on dynamic occlusion and spatiotemporal boundary formation tend to comment on this connection at best in passing (Bruno and Bertamini, 1990; Bruno and Gerbino, 1991; Kellman and Shipley, 1991), if they mention it at all. As a consequence, overlap between kinetic occlusion literature and that on anorthoscopic perception is still surprisingly small. This is unfortunate because studies on anorthoscopic perception have a comparatively long tradition, and their results bear direct relevance to the perception of dynamic occlusion and spatiotemporal discontinuities.

Anorthoscopic perception (see Kalkofen, 2000, for a historical review and explanation of the term) was introduced to vision science by Plateau (1836; 1850), and was later studied more intensely by Zöllner (1862). Zöllner in turn inspired many more modern studies (Anstis, 2005; Fendrich, 2005; Hecht, 1924; Nishida, 2004; Rock, 1981; Rothschild, 1922). Anorthoscopic perception is said to occur in situations where a moving object is almost completely hidden by an occluder—typically a sheet of paper, the only visible portion being what can be seen through a narrow slit that is cut into the paper mask (figures 4.1 on page 91 and 5.2). Even though the visible fragment alone carries by far not enough information to render the object discernible, the object seems to be fully visible and can be readily identified when moving back and forth behind the slit. As the slit itself can be of very narrow proportions, the fact that the objects appear to be extended way beyond the aperture’s dimensions deserves special emphasis (Rock, 1981; Sohmiya, 1994).

While most of the earlier studies on anorthoscopic perception only used one slit, later ones generalized this situation to evenly-spaced multiple apertures (Fahle and Poggio, 1981; Mateeff, Popov and Hohnsbein, 1993; Nishida, 2004). This multi-slit viewing can now be further generalized to apertures of arbitrary shape, in which case it mimics real-world configurations where objects are seen behind, e.g., dense foliage or a picket fence. Dynamic occlusion and uni-directional spatiotemporal-discontinuity displays are equivalent to this kind of general multi-aperture display: The texture elements can be regarded as invisible apertures through which portions of the background are seen. The background is then partly occluded by a moving object located between it and the occluding layer (figure 4.1). If the moving object has the same color as the occluder, it camouflages the apertures, which is equivalent to occluding them. In cases of invisible slits, aperture displays are also reminiscent of the screen effect (Michotte et al., 1964; Sampaio, 1943; cf. section 3.3.1), which may be said to incorporate two very large slits adjoining one occluding stripe.

The straightforward interpretation of multi-slit displays in terms of possible physical arrangements which they simulate means that they form a subset of general spatiotemp-

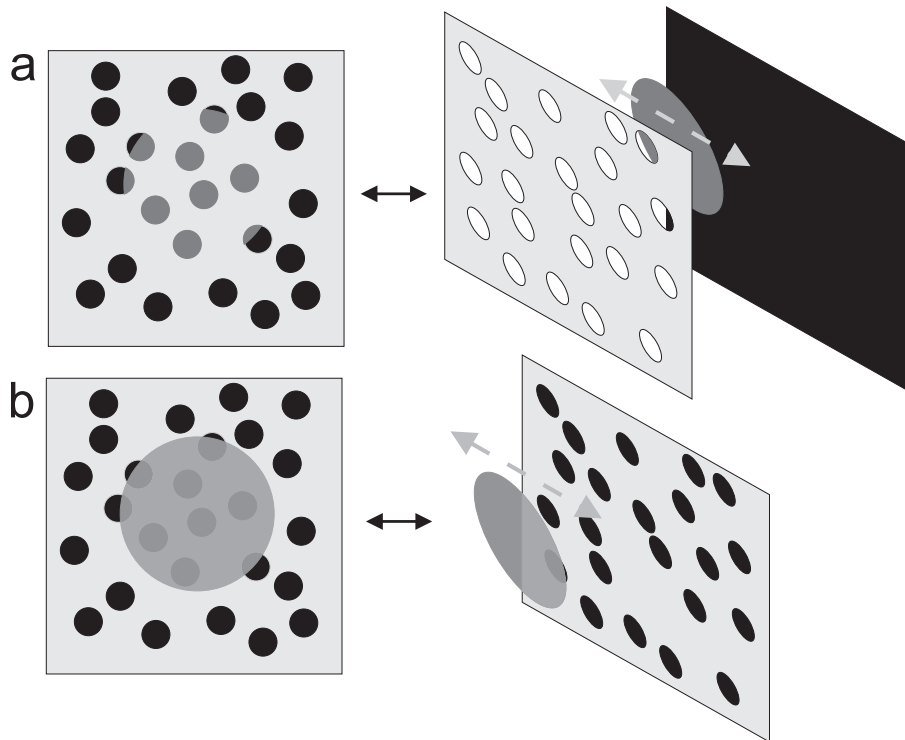


Figure 4.1: Possible percepts in dynamic multi-aperture viewing and their corresponding 3-D scene organization. a) Amodal completion represents the presence of three layers: A frontal mask with numerous apertures in it, through which an object is seen as moving over a background. b) Modal completion corresponds to the perception of a transparent layer, moving over a background with dynamically occluded texture elements painted on it.

oral-discontinuity stimuli, the bi-directional displays being excluded by the restriction to physical plausibility. Naturally, this provokes the question whether the already-discussed perceptual differences between uni- and bi-directional stimuli (cf. section 3.3.1) are the consequence of some perceptual implementation of physical constraints. Alternatively, other differences between the two display classes may be responsible. This question will be taken up in the general discussion in section 6, when all experimental results have been presented.

Most research on multi-aperture viewing stresses aspects of contour integration and form perception, customarily using methods like shape discrimination. This emphasis on boundary aspects parallels the research focus of dynamic occlusion and general spatiotemporal-discontinuity experiments as discussed in section 3.3.1. Since apertures play a major role in anorthoscopic perception, their precise influence on the perceived shapes has been widely studied. It was for instance known from early on that the

perceived shape of the object is often distorted, depending on the object's speed (Zöllner, 1862), the aperture shape (Hecht, 1924; Rothschild, 1922), and the observer's eye movements (Anstis and Atkinson, 1967). The important role of aperture characteristics is apparent in motion perception in general, which must deal with the aperture problem. As a further indication of the aperture's importance, it was found that perceived motion direction can change with the orientation of elongated apertures, with a preference for motion orthogonal to the longer side of the aperture. In addition to the results already discussed in section 3.3.1, more findings from experiments using multi-aperture displays will be reported in section 5.2 in the context of possible explanations for the phenomenon of anorthoscopic perception in general.

Various kinds of multi-aperture displays will be used as stimuli in the experiments presented in the subsequent sections, but in contrast to the majority of studies treated so far, we will here emphasize aspects of surface quality that are attached to the dynamically completed objects. An important phenomenon of spatiochromatic interaction that prepares the ground for this is the so-called color-from-motion effect, described in the following section.

4.2 Color from motion

The color-from-motion displays introduced by Cicerone and Hoffman (1992; Cicerone et al., 1995; Hoffman, 2003) are very similar to the uni-directional kind of spatiotemporal-discontinuity displays: Their only difference being that the moving object between the background layer and the occluding mask with holes punched into it is of a different color than the front sheet of paper (figure 4.1). Thus, whenever the object passes behind a circular aperture, it changes the color of what is visible through the hole from the background color to the color of itself. It should again be noted that no stimulus elements actually move, they just switch coloration according to the position of a moving virtual disk (figure 4.2).

The induced perception in these displays can be of amodal quality (cf. section 3.2.2), in which case the scene interpretation and depth stratification is the one described above. However, a second percept is also common when viewing color-from-motion displays: The moving object can then be modally perceived as a transparent filter, delineated by crisp subjective contours, and completely filled with color that looks like a diffused variant of the moving object's coloration. Compatible with this impression, the object is seen as being located in front of the sheet of paper. The simultaneous switch of coupled perceptual attributes bears strong resemblance to neon color spreading (cf. sections 3.2.1 and 3.2.3). But while the latter effect is often vague and perceptually unstable, the motion-induced coloration and impression of perceptual transparency here are quite strong and unambiguous.

Color from motion was shown to be strongly affected by the amount of perceived motion in the display, more precisely, apparent motion was shown to considerably

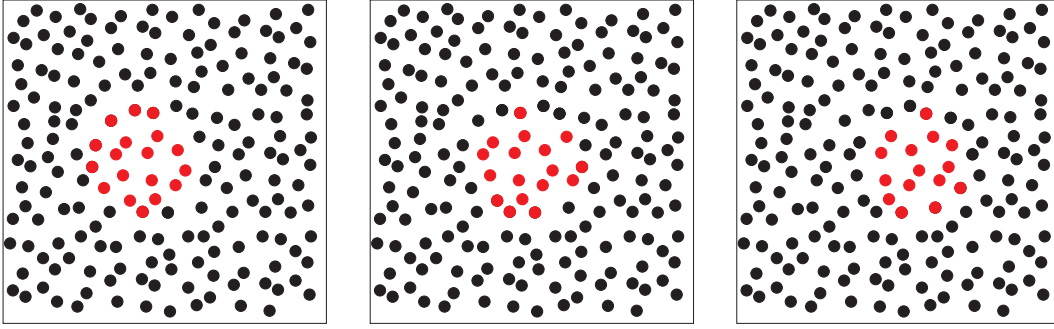


Figure 4.2: Three frames from a color-from-motion display. The dots are stationary and change color according to the position of a moving virtual disk. Whenever a dot falls within the disk's area, it is colored red, otherwise black.

enhance the color-spreading effect (Cicerone et al., 1995). Paralleling results about object recognition in other animated sparse displays, signal element density was also shown to positively influence the distinctiveness of the object's appearance. Fidopiastis, Hoffman and Prophet (2000) finally demonstrated that the strength of color from motion is modulated by the particular layout of the dot, i.e., the aperture array which provides signal elements for form and color information. Specifically, it was established that the color-spreading effect increases with a denser and more regular dot array. Note that these results all represent gradual variations in illusory object perception, whereas the switch between amodal and modal completion and its associated changes in object quality is a discontinuous one.

Pertaining to the question whether color from motion may be of retinal origin, Cicerone and Hoffman (1997) reported that a dichoptic presentation of the stimulus elements does not disrupt the perception of a completed transparent disk. More specifically, the authors presented every odd-numbered frame of the normal animation sequence to the left eye, and every even-numbered frame of the sequence to the right eye. While one eye was seeing a frame of the full stimulus, the other eye was simultaneously shown a frame of the stimulus where all dots within the moving virtual disk were eliminated, and the background color was shown instead. Viewed monocularly with either eye, this stimulus does not give rise to the perception of a moving transparent disk. A moving transparent disk was however what subjects reported seeing when having the stereoscopic input, meaning that information from each eye was combined into one percept. The authors concluded that binocular neurons at cortical processing stages must be responsible for the color-spreading effect.

Later, experiments carried out by Shipley and Cunningham (2001) led to a qualification of these results, by indicating that monocular patterns of change do matter for color from motion: Whereas in Cicerone and Hoffman's (1997) study, the pattern of element changes was identical for each eye, Shipley and Cunningham (2001) devised

a binocular display where, on identical black backgrounds, one eye saw dots switching from white to red once within the virtual disk, and the other eye saw identically positioned dots switching from red to white. Thus, the display was in effect the binocular version of a bi-directional spatiotemporal-discontinuity display. Contrary to the stimulus chosen by Cicerone and Hoffman (1997), the display induced color-spreading monocularly in each eye. Dot contrasts to the background were identical in each eye, hence, when viewed binocularly, the single displays were easily fusible. Nevertheless, no surface perception obtained, indicating that binocular fusion does not work with an abstract surface description as its input. This would have been the same for each eye, meaning that color from motion does take specific monocular qualities into account.

Informal observations made with variations of the color-from-motion display show that it affords a dissociation between boundary and surface integration also found in static displays (cf. section 3.2.1). One manipulation to achieve amoeba-like color spreading not bounded by subjective contours is the following one: Suppose that the coloration of the dots within the virtual disk is not constant but can be subject to random changes. For example, only half of the dots within the disk get colored red, while the remaining half keeps its black color. Now, every fixed interval, the assignment of which dots within the disk are colored and which are not, changes randomly. As a result, small fuzzy islands of color spreading are generated in the vicinity of haphazardly high concentrations of colored dots, and these murky blobs waft around as the coloration changes. Importantly, the unbounded color spreading is not observed if the overall density of dots within the virtual disk is simply cut in half, but the color assignment within the disk is homogeneous and stable. A similar effect was reported by Cunningham, Shipley and Kellman (1998b), but obtained with different means: In a dynamic spatiotemporal-discontinuity display, they inserted static texture elements inside a moving form and found that the contours perceptually fluctuated. Miyahara and Cicerone (1997) also reported fuzzy color spreading in color-from-motion displays if the dots inside and outside of the imaginary disk were of equal luminance—a result which reminds of Watanabe and Sato's (1989) observation of unbounded neon color spreading in an isoluminant modified Ehrenstein cross.

4.3 Flank transparency

4.3.1 Stimulus construction and phenomenology

Color from motion may be regarded as a dynamic variant of static neon color spreading: Both effects share important characteristics with respect to the phenomenology of the illusory coloration and its delineating subjective boundaries, as well as in the functional coupling of depth placement and mode of completion. With flank transparency (Wollschläger et al., 2001; 2002), static flank-induced color spreading like the water-color effect (Pinna et al., 2001) similarly finds an analogy in the dynamic domain.

In a sense, flank transparency arises from combining color from motion with the watercolor effect: The display contains a virtual disk that appears to move relative to an array of solid line segments distributed on a uniformly-colored background. The line segments are arranged such that their center points form a square grid, and their orientations are pseudorandom. All line segments have the same color that is different from the uniformly-colored background. A stationary virtual disk is then defined which is horizontally and vertically centered on the screen. The line segments are rigidly translated by a slight amount in a horizontal direction in each frame, and narrow colored flanks are added parallel to these segments where they lie in the interior of the virtual disk (figure 4.3).

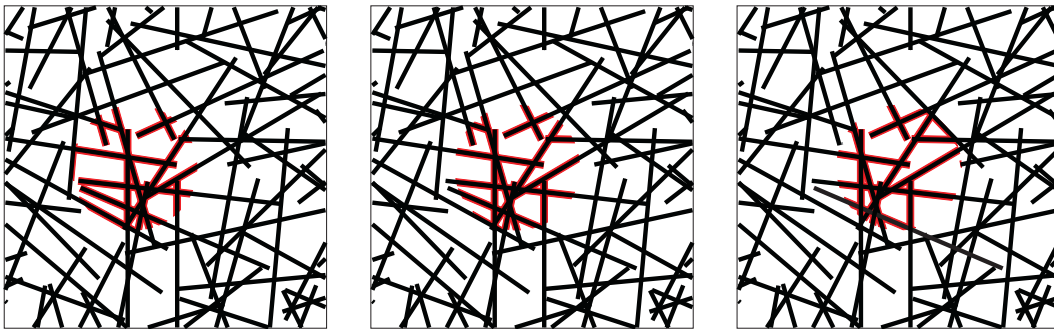


Figure 4.3: Three frames from a dynamic flank-transparency display. The lines are stationary, and thin flanks are added alongside them according to the position of a moving virtual disk. Where a line falls within the disk’s area, a colored flank is attached to it.

Instead of arranging randomly-oriented lines on a grid, the lines may as well be lined up on a horizontal strip, each one being vertically oriented (figure 4.5b on page 98). This configuration makes it more obvious that flank-transparency displays closely resemble standard multi-slit stimuli. However, the described setup differs from standard multi-aperture viewing stimuli in that the latter may have visible or invisible apertures, but they do not incorporate visible lines *between* them. This seemingly subtle distinction has important consequences. Flank transparency can also be observed in static displays (figure 3.9), and it can evoke transparent color spreading even in two-colors displays (figure 4.4)—an observation that is currently not handled by models of perceptual transparency (Wollschläger, 2001).

While a single frame of the flank-transparency animation alone already produces a relatively weak impression of illusory contours and color spreading, the animated display yields compelling color spreading and the clear perception of a moving transparent filter with well-defined boundaries. Flank transparency does not look self-luminous, but it shares other characteristics with neon spreading in addition to perceptual transparency: The possibility of both modal and amodal completion goes hand-in-hand with a concomitant depth stratification of objects, leading to the perception of the

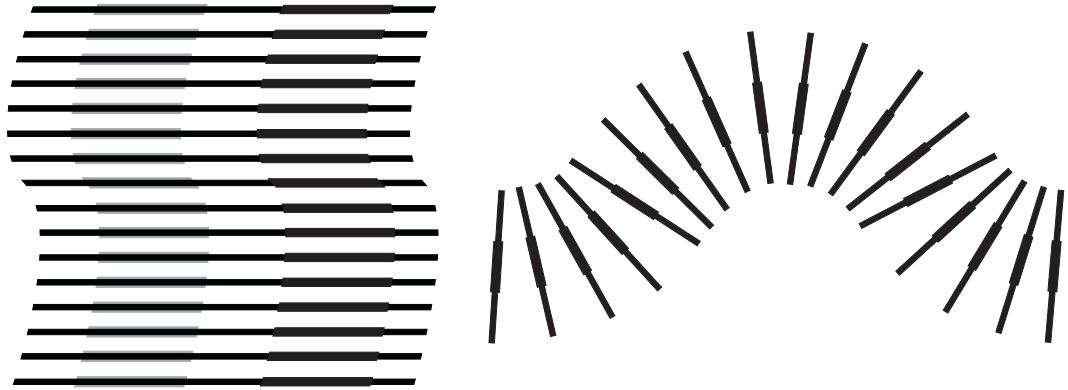


Figure 4.4: Static flank transparency in an achromatic display. Darkness spreading and the perception of a transparent layer can be observed even in two-color versions of the display.

virtual shape as being in front of or behind the background layer. When the flank-transparency animation is halted, the interpolated boundaries as well as the transparent color spreading do not disappear instantaneously, but fade away in a noticeable time-span—similar to what Hine (1987) reported for illusory boundaries resulting from dynamic occlusion.

We found that the color-spreading effect evoked by a flank-transparency display is sufficiently strong to warrant the use of standard quantitative psychophysical tasks like matching or adjustment procedures. This is in contrast to many static neon-spreading stimuli, which often have a slightly elusive quality and are perceptually unstable, thus only allowing for subjective ratings of stimulus-appearance aspects. In the following three experiments, the robustness of flank transparency was exploited using standard asymmetric color matching. In preliminary experiments, we also tested a boundary-strength matching task, but found that it yielded equivalent results as a boundary rating task, which we later employed. Note that while color matching still is a subjective task in the sense that there is no objectively correct answer, it represents a well-established psychophysical procedure whose application was so far restricted to non-illusory stimuli.

The experiments reported here were carried out to be able to place flank transparency in the context of other completion phenomena by determining its basic characteristics and comparing them to related effects. Since flank transparency may be regarded as a crossbreed of the watercolor effect and color from motion, it seems important to establish commonalities and differences to both of them, in particular with respect to its range of operation: Whereas color from motion and neon color spreading are typically confined to a few degrees of visual angle, the watercolor effect can operate over up to 40 deg. In general, the experiments stress the *continuous* variability of the associated percept depending on some of the main stimulus parameters.

Three experiments were performed in order to examine the effect of apparent motion, of introducing small spatial discontinuities (gaps) between lines and flanks, and the spatial extent of the color spreading. Experiment 1 studied the influence of apparent motion on brightness and saturation of the virtual disk, as well as on perceived strength of its boundaries. We observed that increased apparent motion resulted in stronger subjective boundaries and a darker, more saturated illusory disk. The results are compared to those by Cicerone et al. (1995), obtained in a similar task. In experiment 2, we explored the relationship between the gap size between lines and flanks and observers' color matches and boundary-strength ratings. We showed that with the introduction of gaps, boundary strength decreases, and the illusory disk appears lighter and desaturated. The results are related to findings in comparable experiments investigating color spreading in the Ehrenstein figure (Redies and Spillmann, 1981) and to characteristic attributes of the watercolor effect (Pinna et al., 2001). In experiment 3, we estimated the spatial extent over which color spreading can operate in our display. We obtained a measure for spatial drop-off and compared it to that of other neon-spreading displays. Our results indicate that the color spreading in dynamic flank-transparency displays extends over distances much smaller than those described by Pinna et al. (2001) for the watercolor effect.

4.3.2 Experiment 1: Velocity of moving lines

Participants. Five observers with normal color vision and normal or corrected-to-normal visual acuity participated in the experiment, including a collaborating scientist (TR) and myself (DW). The remaining subjects were naïve as to the purpose of the experiment.

Apparatus and stimuli. The stimuli were displayed on an Apple Studio 17in color monitor with a vertical refresh rate of 75 Hz using the OpenGL graphics library. Color resolution for each pixel was 8 bits for each channel R, G, B . Following a procedure described by Brainard (1995), the monitor was calibrated and gamma-corrected using a Spectracolorimeter (Photo Research PR-650). Observers viewed the display in a darkened room at a distance of 80 cm.

The display's background was white with CIE 1931 (Wyszecki and Stiles, 1982) coordinates $x = 0.29$, $y = 0.31$, $L = 70 \text{ cd} \cdot \text{m}^{-2}$, and subtended a visual angle of 18 by 11 deg. The display was bipartite: The upper half contained the dynamic-color-spreading stimulus, and the lower half the figures used for a match and a rating task described below.

The dynamic stimulus contained randomly-oriented black lines (CIE $x = 0.29$, $y = 0.31$, $L < 0.2 \text{ cd} \cdot \text{m}^{-2}$) of width 13.6 min of arc and length 2.27 deg.¹ Red (CIE $x =$

¹Throughout the experimental section, distances are reported in terms of their respective visual angle, measured in degree or min of arc.

0.48, $y = 0.33$, $L = 14 \text{ cd} \cdot \text{m}^{-2}$), and, in a separate session, green (CIE $x = 0.29$, $y = 0.45$, $L = 14 \text{ cd} \cdot \text{m}^{-2}$) flanks of 2.73 min of arc width were added parallel to the lines on both of their sides in the interior of a virtual half disk with a radius of 1.6 deg (figure 4.5). A fixation cross was placed on the horizontal center line at an eccentricity of 9.55 min of arc to the disk boundary, because our previous pilot studies had indicated that color spreading was strongest under parafoveal viewing conditions.

In the first task, a solid half disk for doing a color match was displayed in the lower half of the screen on the same white background as the upper half (figure 4.5). The two half disks were separated from each other by 1 deg in order to avoid perceptual interactions between the illusory and the real half disk. Otherwise, we reasoned that the real half disk would adversely affect color spreading in the motion-defined disk when placed immediately adjacent to it. In the second task, a number between 0 and 10 was displayed, corresponding to an observer's rating entered on the keyboard.

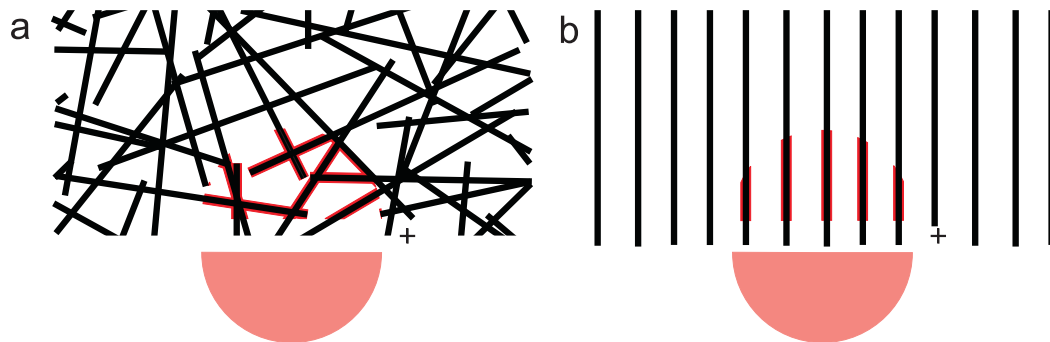


Figure 4.5: Flank-transparency stimulus as used for the color-matching procedure. The bottom half displays an adjustable half disk used for the matching. a) In experiments 1 and 2, line orientation was random. b) In experiment 3 it was vertical.

The line array rigidly moved to the right in unison, whereas the virtual half disk remained stationary at the center of the screen. The trailing side of the display was continuously replenished with lines. The velocity of the moving lines was modified purely through timed delays between changes of their position, translating into speed of apparent motion. The length of the discrete “jumps” of the lines from one frame to the next, as well as the physical refresh rate of the monitor, was constant over different speeds.

Procedure. The independent variable was the velocity of the moving line array. A flank-transparency display was presented using one of three different speeds, either 0.5, 1.5, or 2.5 $\text{deg} \cdot \text{s}^{-1}$. Subjects performed 10 repetitions for each of the three speeds in

pseudo-random order. The experiment was conducted twice, once for red and once for green flanks.

The observer's task in a given trial was twofold. First, the observer matched the color of the virtual disk in the upper half of the screen by adjusting the color of a solid half disk in the lower half with the computer keyboard's arrow keys. We took the adjusted color as a measure for the strength of the spreading effect: Darker, more saturated colors indicated a strong effect, whereas lighter, less saturated colors similar to the background color indicated an effect that was only weak.

Pilot studies had indicated that while perceived saturation and brightness of the illusory disk were not affected by the experimental manipulations, hue did not seem to vary. Consequently, we restricted the subjects' settings to a plane of constant dominant wavelength, corresponding perceptually to a constant hue. Subjects thus controlled luminance and saturation. The hue was fixed to be identical to that of the flanks. This restriction helped reduce the complexity of the color-matching procedure and increase reliability. Before each new trial, the luminance and saturation of this half disk were preset to a random value in the chosen plane of constant hue. Prior to the experiments, all subjects completed as many training trials as were needed to become sufficiently familiar with the task of adjusting a color in the chosen two-dimensional plane in color space.

In the second task subjects were asked to give a rating of perceived boundary strength of the illusory disk on a scale ranging from 0 to 10, using the keyboard's arrow keys to increase or decrease the displayed number on the screen. The endpoints of the scale were anchored verbally, with "0" corresponding to no perception of illusory contours, and "10" corresponding to most highly apparent illusory contours. Examples of different illusory boundary strengths were provided to illustrate the scale's possible range, using examples of apparent illusory contours such as the Kanizsa triangle and Albert and Hoffman's (1995) magic square (figure 3.8b, page 37).

Subjects were instructed to maintain fixation on the cross to reduce possible blurring effects of eye movements. There was no time limit for either task—subjects pressed a key to confirm their settings, and continued with the next trial.

Results. Figure 4.6 shows the means of the subjects' ratings in the first experiment for red and green flanks. Ratings for perceived boundary strength increased monotonically for each subject with increasing speed of the line array.

Figure 4.7 presents data obtained from the color-matching task in experiment 1. It shows the settings for red and green flanks for two of the five observers in an excitation purity (Wyszecki and Stiles, 1982) by luminance color space (see table B.1 for the complete data). The axes correspond to the two dimensions available to the subjects for the color matches. Results indicate that with increasing speed, the illusory disk appeared darker and more saturated than at slower speeds. This pattern was observed

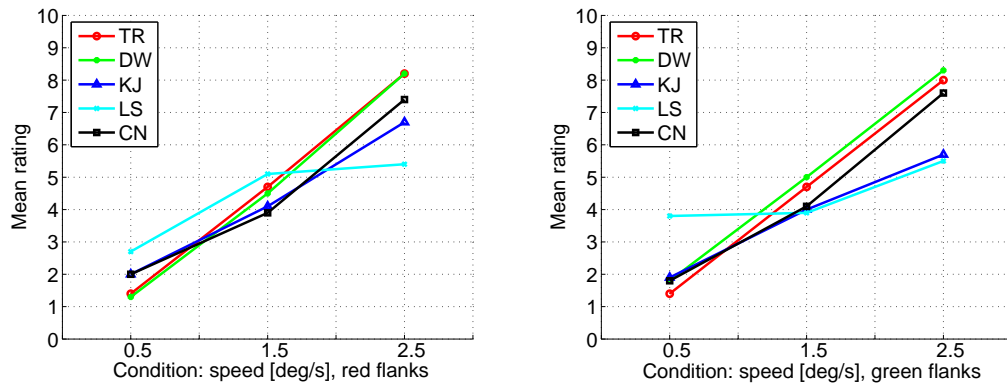


Figure 4.6: Mean boundary-strength ratings in experiment 1 for slow, medium, and fast speed, respectively. a) For red flanks. b) For green flanks.

in four of the five subjects, with considerable inter-subject variability in the absolute location of the matches. One subject’s (KJ) settings did not substantially differ between conditions. Except for subject KJ, the centroids for each subjects’ matches in the three conditions were approximately equally spaced, i.e., no two conditions grouped together.

With green flanks, the color-spreading effect was apparently a lot weaker than with red flanks, indicated by the low saturation settings in the matchings. This is probably the reason why these data were generally less systematic than with red flanks. Increasing saturation resulting from higher velocities was observed for three (TR, KJ, CN) of the five subjects, decreasing luminance for two subjects (TR, CN) (table B.1). The remainder of the observers did not exhibit a systematic pattern of change between conditions.

Discussion. Experiment 1 provides good support for the conclusion of Cicerone et al. (1995) that “color spreading is yoked to the perception of apparent motion” (p. 763). Thus, despite not being a novelty, experiment 1 shows convincingly that perceptual codes for color and motion can show strong interactions. Although constructed in a different way, our display behaves like color-from-motion stimuli employed by Cicerone et al., who also found that ratings for boundary strength are higher for faster-moving stimuli. In addition, we quantified the effects of increasing speed on subjects’ settings in a color-matching task. Color spreading increased with higher velocities in the sense that the illusory disk acquired a darker and more saturated color.

Speaking in terms of perceptual transparency, the observed reduction in color spreading with slower speeds corresponds to the impression of a filter that becomes more and more transparent and thus less visible itself. The color matches approach to the color

4.3 Flank transparency

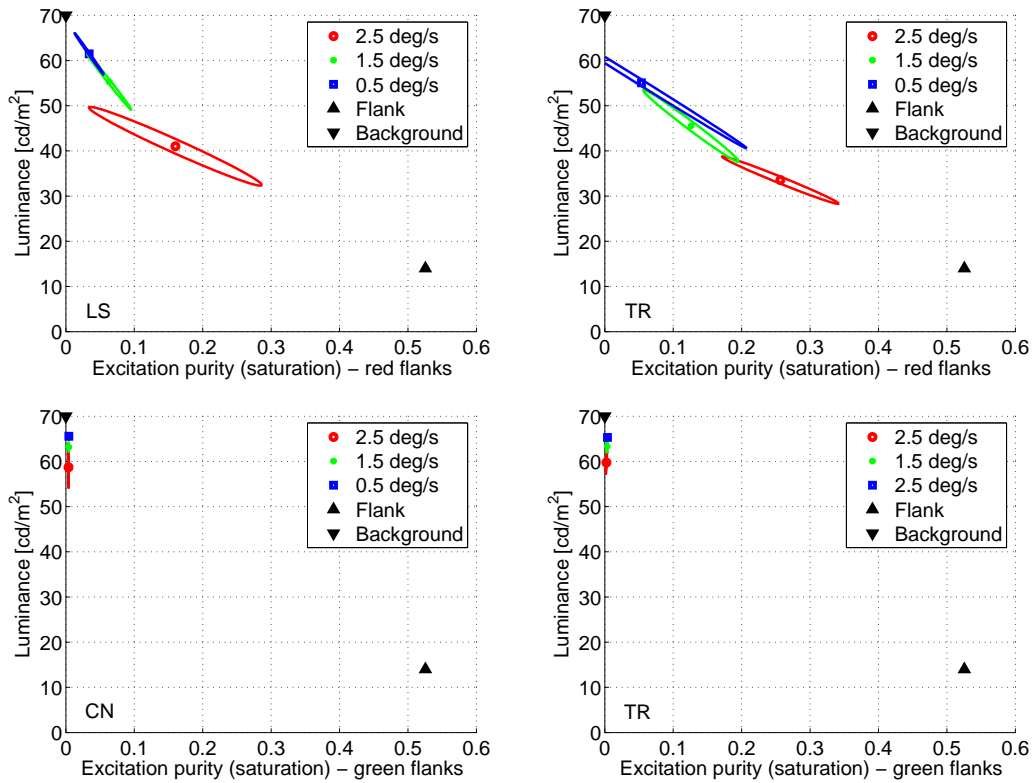


Figure 4.7: Color matches and corresponding 95% confidence ellipses in experiment 1 for red and green flanks, observers LS, TR, CN, and TR. See table B.1 for the complete data.

and luminance of the background can be interpreted within additive transparency models, for instance, the episcotister model (D’Zmura et al., 1997; Metelli, 1974). Here, the same effect results if more and more light reflected from the white background is mixed with proportionally less light reflected from the filter.

The increase in color spreading with increasing speed is paralleled by an increase in ratings for perceived boundary strength of the illusory disk, indicating that the perceived layer became more visible as a separate object through better specification of its contours and its surface color.

However, the pattern of increasing saturation and decreasing luminance with higher speeds was not evident in the data of all subjects. We will consider two reasons for this inter-subject variability.

First, one subject’s (KJ) matches were all very close to the monitor gamut, possibly indicating that only unsatisfactory matches were possible due to monitor restrictions. Second, one might argue that the nature of the matching task is fairly abstract due to

the different phenomenological qualities of the regions that are to be matched: whereas the illusory disk appears transparent, the adjusted comparison patch appears as a matte opaque surface color. As a result, the two regions never look truly identical (see Ekroll, Faul and Niederée, 2004, for a discussion of non-matchability problems in asymmetric color matching). Thus, the subjects' task effectively was to maximize the similarity between target and matching region. It required a concentration on particular aspects of the display, like saturation and lightness, while other aspects like opacity had to be discounted. This apparently led to a reduced accuracy of the matches and to greater dissimilarities between data of different subjects. Nevertheless, as we have just seen, our data can be readily interpreted.

4.3.3 Experiment 2: Modulating gap size

Participants. The same five subjects as in experiment 1 participated in experiment 2.

Apparatus and stimuli. The apparatus for displaying and viewing the stimuli was the same as in experiment 1.

The stimuli were the same as in experiment 1 with the following difference: Gaps of varying size were introduced between the colored flanks and the line segments (figure 4.8). Over time, this manipulation did not affect the averaged area of flanks in the motion-defined disk, but merely changed the spatial location of the flanks. The speed of the moving line segments was kept constant at $2.5 \text{ deg} \cdot \text{s}^{-1}$. Among the tested speeds in experiment 1, this was the setting that created the strongest color-spreading effect. In this and the following experiment, only red (CIE $x = 0.48$, $y = 0.33$, $L = 14 \text{ cd} \cdot \text{m}^{-2}$) flanks were used because of their more compelling color spreading effect compared with the green stimuli.

Procedure. The independent variable was the size of the gap between the colored flanks and the line segments. The three gap sizes employed were no gap, 1.36, and 2.73 min of arc. Subjects performed the same two tasks as in experiment 1 in 10 repetitions for each of the three conditions in pseudo-random order.

The subjects performed the same color-matching and boundary-strength rating task, and received identical instructions as in experiment 1.

Results. Figure 4.9 shows the means for the subjects' ratings in the second experiment. Ratings for perceived boundary strength decreased monotonically for each subject with increasing gap size.

Regarding the color matches, the effect of increasing gaps between lines and flanks was similar to that of decreasing speeds in the first experiment: Color spreading was

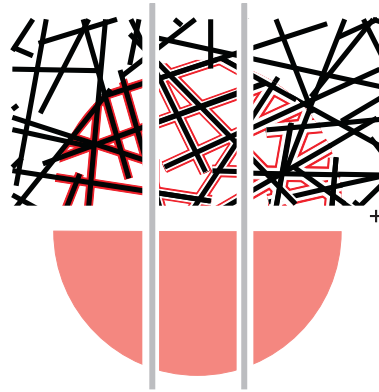


Figure 4.8: Schematic depiction of the stimulus used in experiment 2. There were three conditions of varying gap size.

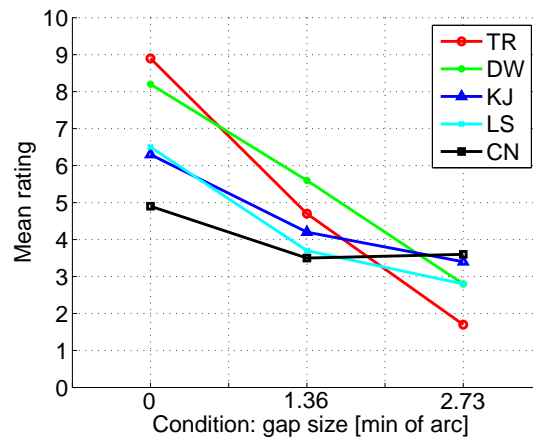


Figure 4.9: Mean boundary-strength ratings in experiment 2 for no, medium and large gaps between lines and flanks.

impaired in four of the five subjects with the introduction of gaps (table B.2). In these cases the disk appeared less saturated and lighter, as can be seen in figure 4.10, which shows the data of three subjects. For one subject (CN, cf table B.2), the matches did not substantially differ between conditions: they were all highly desaturated and close to background luminance. As for the other subjects, the matches resulting from the small-gap condition and the matches from the large-gap condition tended to lie close together with highly desaturated settings near background luminance. In contrast, the matches resulting from the no-gap condition were separated considerably from this group in having more saturated and darker settings. The no-gap condition replicates the fastest condition ($2.5 \text{ deg} \cdot \text{s}^{-1}$) of experiment 1, and each subject's settings and

ratings showed good reliability when compared between the two experiments.

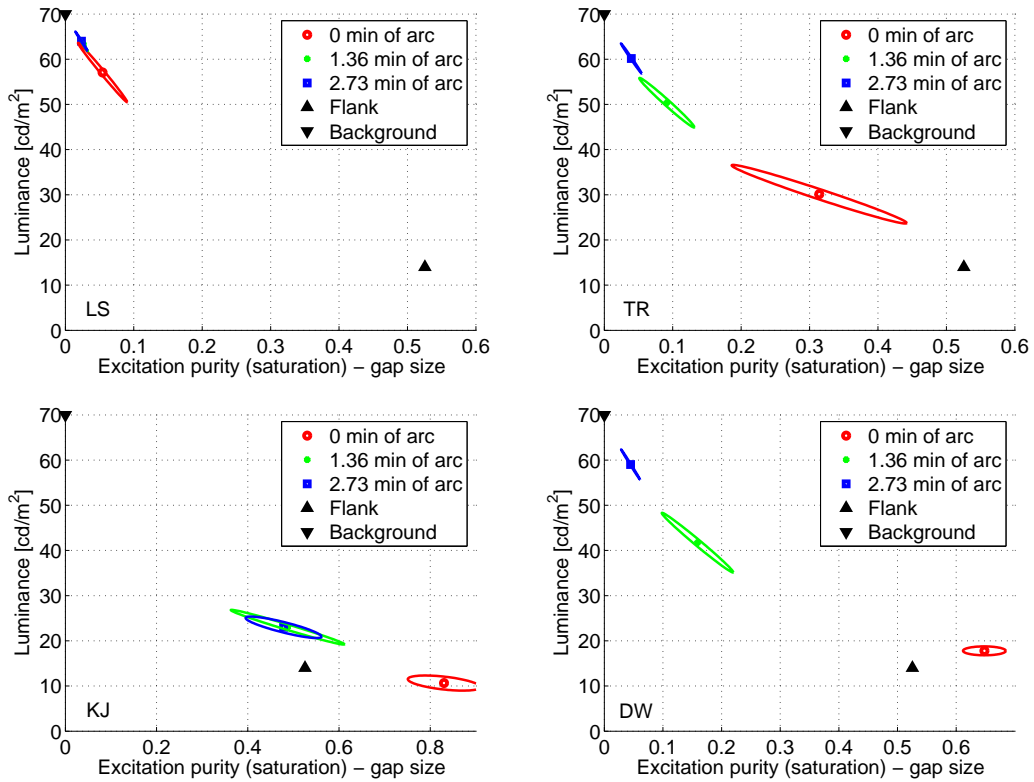


Figure 4.10: Color matches and corresponding 95% confidence ellipses in experiment 2 for observers LS, TR, KJ, and DW. See table B.2 for the complete data.

Discussion. The results are similar to data obtained by Redies and Spillmann (1981) who showed that spatial discontinuities between the outer and inner cross in the modified Ehrenstein figure resulted in a loss of color spreading. However, violations of physical contiguity seemed to have a more detrimental effect in neon color spreading than here, possibly owing to the generally more robust nature of the color-spreading effect in flank transparency. For static displays of the watercolor illusion, Pinna et al. (2001) demonstrated that color spreading vanishes quickly if flanks and lines are not immediately adjacent. Similarly, in our dynamic displays, color spreading persisted with larger gaps, but was greatly reduced in strength. This experiment thus underscores the importance of figural determinants of color spreading, showing that spatial contiguity of line segments and flanks plays a critical role in flank-transparency displays. Still, the reduction in color spreading was a continuous one as indicated by the remaining faint spreading in the weaker conditions. This is opposed to the often qualitative influence of figural aspects in static configurations. As in experiment 1, the

changes in the degree of color spreading covary with a reduction in perceived boundary strength, showing that the contours and surface color of the perceived layer together became less visible.

On the one hand, the gradually degrading color spreading with increasing gap size is compatible with explanations in terms of simple spatiotemporal averaging processes (cf. section 5.2). But, on the other hand, it is also implausible to assume these processes as the only explanation of flank transparency: The introduction of gaps left the number and summed length of red flanks within the imaginary circle largely unaffected, i.e., an averaging that ignores figural aspects should not be influenced by the experimental manipulations.

4.3.4 Experiment 3: Spatial range of the color-spreading effect

Participants. The same five subjects as in the previous experiments also participated in experiment 3.

Apparatus and stimuli. The apparatus for displaying and viewing the stimuli was the same as in experiments 1 and 2.

The stimulus was similar to the one in experiment 1, with the following differences: The line array was now rearranged such that all lines were vertical with a uniform distance to each other at all points (figure 4.5b). This was done in order to control for the exact distance to be bridged by the color spreading. The spacing between line segments varied from 0.89 to 1.84 deg between conditions. The size of the virtual disk was adapted accordingly so as to keep the mean number of lines and flanks in the virtual disk constant over different spacings (figure 4.11). The speed of the moving line segments was kept fixed at $2.5 \text{ deg} \cdot \text{s}^{-1}$, the speed we had found optimal in experiment 1 for creating a strong color-spreading effect.

Procedure. The independent variable in experiment 3 was the distance between facing pairs of colored flanks in the interior of the virtual circle.² The three levels chosen for the distance were 0.89, 1.43, and 1.84 deg. In order to keep the number of flanks in the interior of the disk constant over levels of separation, the radius of the disk was adjusted to 1.27, 2.11, and 2.73 deg for the three distances, respectively. As in experiments 1 and 2, observers performed 10 repetitions of the same tasks for each of the three conditions in pseudo-random order.

²Note that this manipulation automatically affects the proportion of physically-specified area relative to the total area of the disk. Line distance and proportion of colored area both reflect the signal element density of the stimulus and are related in a strictly-monotonic increasing way. Since line distance additionally specifies *how* this element density is achieved, it is the independent variable reported here. This reasoning also applies to experiments 4 (cf. section 5.3) and 6 (cf. section 5.7).

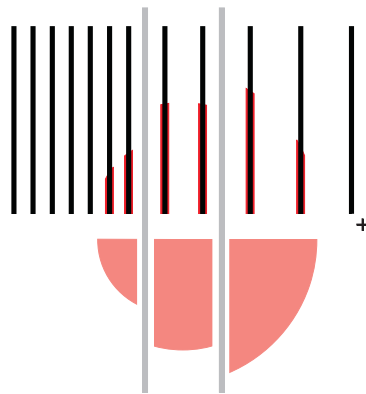


Figure 4.11: Schematic depiction of the stimulus used in experiment 3. There were three conditions of varying line distance.

The subjects performed the same color-matching and boundary-strength rating task, and received identical instructions as in experiments 1 and 2.

Results. Figure 4.12 shows the means for the subjects' ratings in experiment 3. Ratings for perceived boundary strength decreased monotonically for each subject with increasing distance between flanks attached to neighboring lines.

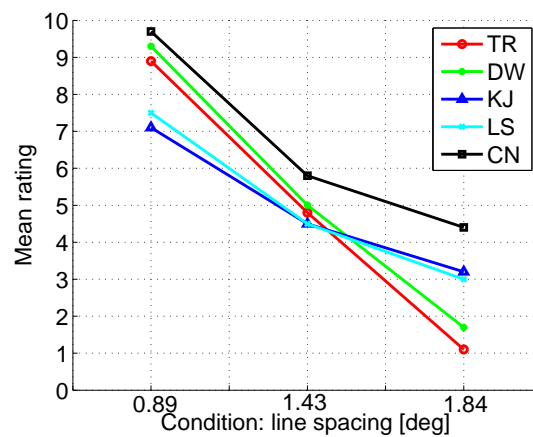


Figure 4.12: Mean boundary-strength ratings in experiment 3 for narrow, medium, and wide spacing of the lines.

The matching data from four of the five subjects showed the same pattern as obtained in experiments 1 and 2 with different manipulations: Increasing the distance between line segments and thereby between pairs of opposing flanks lead to a reduction in color

spreading in the sense that the illusory disk desaturates and became lighter (table B.3, figure 4.13). Only for subject KJ did the illusory disk appear more saturated and slightly darker when seen in displays with greater separation between lines. Also evident from the data is the great dissimilarity between the settings for the nearest condition as opposed to the settings in the two other conditions which for all observers lied close together. In particular, the settings for the two larger distances were close to the background white for four of the five subjects.

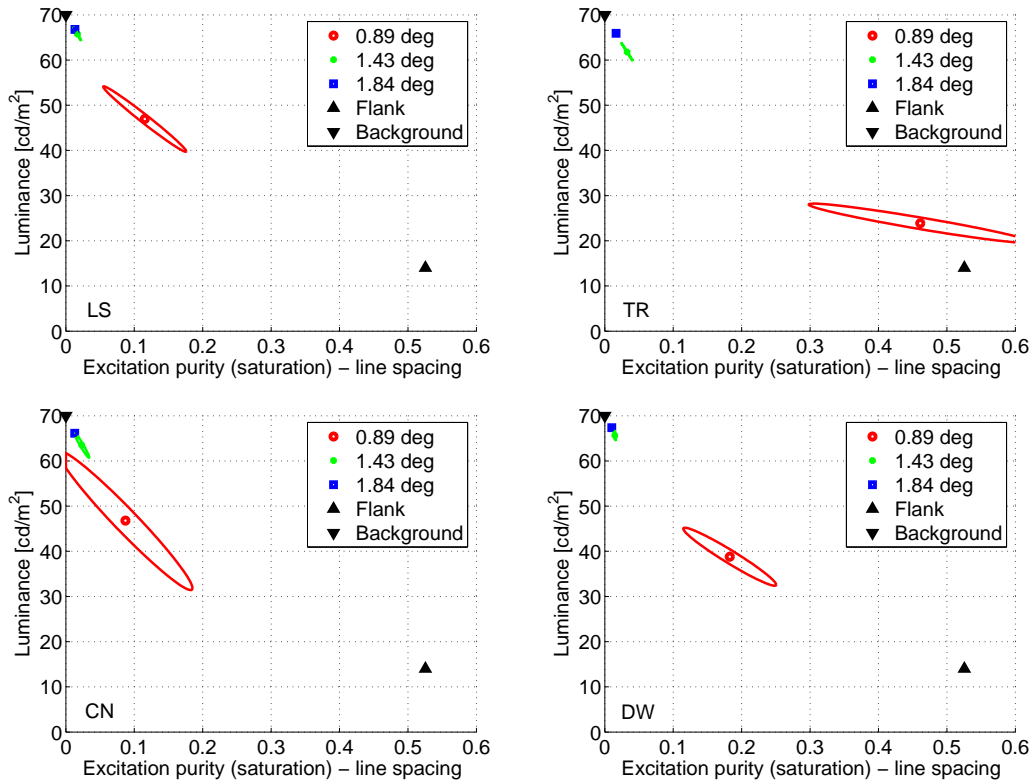


Figure 4.13: Color matches and corresponding 95% confidence ellipses in experiment 3 for observers LS, TR, CN, and DW. See table B.3 for the complete data.

Discussion. Experiment 3 shows that the strength of the color spreading quickly degenerates as the spatial extent is increased over which the spreading must bridge. At larger distances, the illusory disk’s color is nearly identical to that of the background. As in the experiments before, the reduction in color spreading is expressed by increased luminance and reduced saturation in the color matches of the virtual disk. Except for subject KJ, a reduction in color spreading is correlated with weaker illusory contours defining the disk. Whereas Pinna et al. (2001) reported that the watercolor effect could effectively operate over 40 deg, our display—in this respect—resembles more traditional

displays of neon spreading which are greatly limited in the spatial extent covered by the spreading color.

4.3.5 Additional observations

While we made several arbitrary design choices in the construction of our display, e.g., concerning the shape of background elements and of the virtual figure, or the direction of motion, the spreading effect does not depend much on this particular setup. Adding flanks to moving elements in a circumscribed region leads to color spreading much like other uni-directional element manipulations, such as changes in color, do (cf. section 3.3.1). As such, the structural layout of the flank transparency display is easily generalized to other choices of inducing elements and virtual shapes, thus permitting studies on interpolation strategies of the visual system when facing arbitrary incomplete contours. For instance, the moving lines spread across the background can be replaced with small dots furnished with colored fringes when the dots are in the interior of the virtual shape. Likewise, substituting the virtual disk with a rectangular or concave shape is possible without altering the effect. With respect to the restrictions on color choices, flank transparency even abounds if the same color is chosen for lines and flanks, making the display a purely two-tone stimulus (Wollschläger et al., 2001).

We concentrated on presenting results from displays that evoke a clear perception of color spreading and transparency. However, it should be noted that it is possible to generate structurally identical displays that fail to elicit such effects when colors for background, lines, and flanks are appropriately chosen. This occurs, for instance, with a greenish background (CIE $x = 0.35, y = 0.56, L = 10$), light-gray lines (CIE $x = 0.29, y = 0.31, L = 15$), and blue flanks of the same luminance as the background (CIE $x = 0.17, y = 0.15, L = 10$). As is typically found in other displays, the perception of motion is strongly impaired near isoluminance of flanks and background. The flanks seem to move in discrete jumps, and no illusory boundaries are perceived. Pinna et al. (2001) also reported degraded color spreading at isoluminance for watercolor-effect stimuli. The impeded motion analysis resulting from equiluminance emphasizes the involvement of motion processing for dynamic completion: As motion can be a powerful cue to fragment grouping (cf. section 3.3.1), its absence is a likely candidate to explain the failure of color spreading. It also follows that the color-spreading and transparency effects each depend on luminance conditions as well as on figural determinants.

A figural manipulation which prevents color spreading from taking place is to not juxtapose the flanks to the lines, but to rotate them instead by 90° . Even though the flanks' rotation does not change the absolute amount of surface information within the virtual disk, the lacking contiguity with the black lines seems to constitute a hindrance to a segmentation into different layers. As a consequence, no grouping of fragments and surface integration obtains. This result is similar to the one found for introducing gaps between flanks and lines. However, unlike the latter result, it seems to be qualitative in nature, i.e., already small rotations immediately abolish the color spreading.

Cicerone and Hoffman (1997) suggested that a functional advantage of color from motion might be a facilitation to detect moving camouflaged objects (cf. section 2.3). They reported that objects which were invisible in still view could be perceived as moving behind a partially occluding screen in the apparent-motion condition. A very similar phenomenon can be observed in the flank-transparency display: Given equally colored flanks and lines, a sufficient reduction of the width of the flanks in the interior of the virtual disk (e.g., to 1 px) can render the flanks virtually invisible in still view—only an array of uniform lines randomly spread over the background is perceived. Using apparent motion, the disk surface defined by the presence of the narrow flanks immediately pops out. A similar salience enhancement due to static color spreading was reported by Redies et al. (1984).

4.3.6 Discussion

As has already become evident in sections 3.2 and 3.3, there is much diversity in the phenomenology of color spreading and associated degrees of perceptual transparency. Their phenomenological variety naturally prompts for the question whether to each type of color spreading corresponds a qualitatively different perceptual mechanisms or whether they are due to a combination of common underlying processes. Of particular interest is the question if there exists some shared rudimentary form of color spreading which is at the core of all phenomena, but which manifests itself in the guise of different perceptual interpretations depending on the stimulus class employed.

In the following section, I will first give a brief overview of selected models for transparent color spreading together with a discussion of their applicability to our stimulus class and of their compatibility with our data (for a broader review, see, e.g., Bressan, 1997; Bressan et al., 1997). I will then try to systematize the color-spreading phenomenology in terms of accordances and differences with respect to key qualities like transparency, spatial extent and dependence on perceived depth order. Our goal is to evaluate the proposal of a common basal color-spreading effect whose final perceptual appearance may be modified through a limited set of structural stimulus characteristics.

Models

While some authors reason that qualitatively different mechanisms are responsible for the various color-spreading effects (Broerse et al., 1999; Pinna et al., 2001), others have attempted to conceptually integrate many of the phenomena in a single parsimonious theory. These theories mainly focus on the color determinants of spreading effects.

Bressan (1993b) proposes that “neon spreading and assimilation processes share the same low-level mechanism, and that the special sensory conditions for the occurrence of neon spreading are in fact the sensory conditions for the occurrence of perceptual

transparency” (p. 55). In this account, the diffusion component of neon spreading is entirely due to von Bezold-type assimilation. The remaining phenomenal difference between pure assimilation stimuli and neon-spreading displays is then attributed to a layer decomposition of the assimilation color into a transparent filter and a background component in the case of neon spreading. The assimilation-based color filling-in is thus understood as a first processing stage. It is presumably followed by a layer decomposition if the stimulus structure present after the filling-in process meets the criteria for perceptual transparency. “Bezold-type assimilation turns into the neon spreading effect whenever (i.e., if and only if) it takes place within a surface that is also seen as transparent.” (p. 55). This approach denies extra theoretical relevance of the neon spreading-phenomenon beyond that of assimilation processes and transparency.

In contrast, Pinna et al. (2001) object that their fringe-induced color spreading cannot be explained by assimilation effects since the latter are effective only over much smaller spatial extents than the illusory coloration observed in the watercolor effect. If one follows this argument and assumes a different underlying cause, then it seems possible to also attribute the spreading component of neon color spreading to factors other than assimilation.

Anderson (1997) suggests a different mechanism for the illusory change in brightness that is observed in various displays, including achromatic versions of neon spreading. He argues that a “phenomenal scission of homogeneous luminance into multiple contributions” (p. 419) gives rise to a change in brightness in order to reflect the inferred properties of the different contributing layers. This laminar segmentation (Mausfeld, 1998) is presumably triggered by the presence of contour junctions between areas with different amounts of luminance contrast of the same polarity (see also Todorović, 1997). While Bressan assumes that illusory spreading results from assimilation, and the neon effect from transparency, Anderson claims that changes in brightness as well as transparency simultaneously stem from perceptual scission into separate layers of the colored embedded elements. In a strict sense, however, Anderson’s reasoning seems to apply to the embedded elements only, and therefore does not explicitly provide a mechanism for the characteristic brightness changes *in between* the embedded elements in neon-spreading displays. Also note that watercolor-effect stimuli lack junctions completely and thus are not easily explained by Anderson’s approach.

In a general account of color and form perception, Grossberg and Mingolla (1985) discuss applications of their neuro-computational model to stimuli that elicit illusory contours and neon spreading (see, e.g., Grossberg, 1992; 1997 for further development of this idea into the FACADE model). They propose that a Boundary Contour System (BCS) and a complementary Feature Contour System (FCS) work as independent modes on extracting image properties. While the BCS serves to generate perceptual boundaries, the FCS creates information about the features (e.g., color or texture) that then fill in the regions defined by the BCS. Grossberg and Mingolla conjecture that in the case of neon-spreading displays, a local inhibition of the BCS output in the

colored region can lead to a perceptual dispersion of color into the surround until it is “stopped” by the next extracted contours.

Pinna et al. (2001) demur that FACADE largely depends on the presence of line-end terminators which are not found in watercolor displays.

While the classical color-spreading demonstrations considered by all of the models require at least three colors to evoke the effect, our results on perceptual transparency in two-tone displays suggest that this is not an inherent characteristic of color spreading in general. The necessity to use three colors might instead be due to limitations imposed by the particular structural layout (insertion of colored elements in line drawings) shared by most known neon-color-spreading displays. Note that the Ehrenstein figure as well as most other static color-spreading demonstrations cannot separately specify figural stimulus composition on the one hand, and the colors of inducing pattern and colored shape on the other hand. For instance, identical colors of the inducing pattern and the inner cross in the modified Ehrenstein figure turn the image into one of a uniform cross—thereby crucially changing the object’s composition of parts.

It is a common feature of the abovementioned models of color spreading that—their differences notwithstanding—they all assume different relative contrasts between inducing pattern and background on the one hand, and between inserted colored elements and the background on the other hand as the initial trigger of the effect, i.e., assimilation, scission, or contour weakening. This approach necessarily fails to account for flank-transparency displays in which the same colors are used for both fringes and line segments—a setup that can induce a distinct impression of color spreading and transparency (Wollschläger et al., 2001). Figure 4.4 (page 96) illustrates this situation with a static display that leads to a slightly weaker effect than the dynamic version. Transparency in two-color stimuli was apparently first observed jointly by Kanizsa and Bozzi in 1960/61 who later published an example quite similar to it (figure 3.19, page 48; Bozzi, 1975).

If one tries to explain color spreading in two-color displays with assimilation processes, one faces the problem to specify the color conditions under which assimilation should not take place. Interpreted physically, this limiting case would correspond to a transparent layer which does not attenuate the energy spectrum of the incident light at all. As this layer would then be invisible, the situation would be equivalent to the absence of such a layer. This is a problem because, with respect to their color composition, two-tone color-spreading displays are identical to the vast majority of two-color displays where no color spreading is observed. This is probably why current color models of perceptual transparency cannot account for the transparency impression reported here. With respect to transparency, the chromatic information present in a two-tone flank-transparency display is no different from the usual case of a two-color display that does not elicit transparency. Consequently, scission-based models of neon color spreading should make identical predictions for both cases, thus failing to explain the different phenomenology.

It is therefore clear that none of the models can be regarded as being sufficiently complete in specifying *necessary* conditions for the occurrence of color spreading and transparency. But it should be noted that at least Anderson (1997) explicitly refrains from making this claim. The deficits of color-based models that become most apparent in the two-color case reveal the importance of other image characteristics in a cooperative cue-integration process leading to the final percept of color spreading.

The flank transparency effect provides an important link between two different instantiations of color spreading: On the one hand the fringe-induced color spreading as presented by Broerse et al. (1999) and Pinna et al. (2001) with its pastel appearance and long-range operation, and on the other hand traditional, neon-like color spreading that appears transparent. I was able to show in the presented experiments that stimuli similar to watercolor-effect displays behave like neon-spreading stimuli once apparent motion is introduced to provide figural cues for the presence of an object. Instead of looking matte and opaque, the color spreading then appears transparent and is bounded by subjective contours—another typical feature of classical neon spreading that is absent in the watercolor effect. As shown in the present experiments, flank transparency shares with color-from-motion displays devised by Cicerone and Hoffman (1992) its dependency on apparent motion, as well as its connection between perceived boundary strength and amount of color spreading. It is also similar to classical neon spreading in its characteristic short-range operation demonstrated in experiment 3.

Chapter 5

Dynamic texture spreading

5.1 Stimulus construction and phenomenology

The phenomenon of dynamic texture spreading (figure 5.1) builds upon the possibilities of the flank-transparency stimulus. It is a spatiochromatic interaction that not only comprises the formation of strong subjective contours but also the vivid illusory perception of surface qualities like a colored texture as well as depth stratification. Unlike static texture spreading, the effect manifests itself reliably, and is of high strength in general. As such, dynamic texture spreading extends flank transparency to the completion of *non-uniform* surface features. Static neon color spreading and texture spreading thus both have counterparts in the domain of dynamic stimuli, i.e., flank transparency and dynamic texture spreading—each one qualitatively similar but more robust compared to their static analogue.

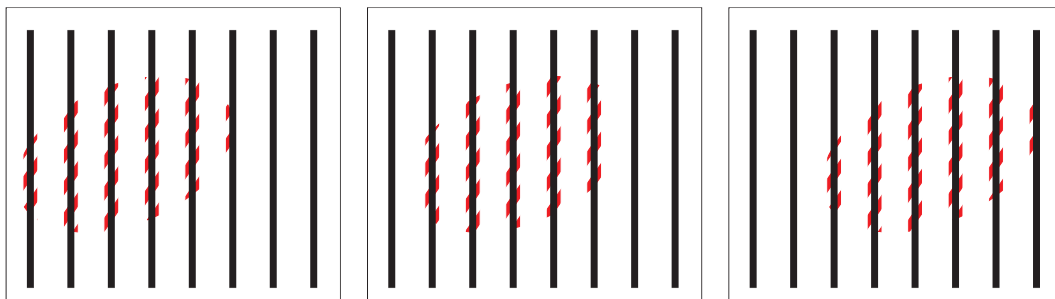


Figure 5.1: Three frames from a dynamic-texture-spreading stimulus as the virtual disk moves to the right. Note that within each flank, the motion direction is a vertical one like in the Barberpole illusion. The flank size displayed here is slightly exaggerated for illustration purposes.

Dynamic texture spreading can be observed with the same variant of multi-slit view displays as flank transparency (cf. section 4.3; Fahle and Poggio, 1981; Mateeff et al., 1993; Nishida, 2004). These are generalizations of Zöllner's (1862; Rock, 1981) anorthoscopic view apparatus from one to multiple apertures: Suppose you put a disk

covered with a stripe pattern on a white background. In front of it, you place a white sheet of paper with black vertical lines painted on it. Seamlessly alongside the lines, you cut narrow rectangular slits, making visible small strips of the background and the textured disk behind it (figure 5.2). The animation procedure may be such that the front sheet of paper is translated gradually on a horizontal path, accompanied by a corresponding change in what is visible of the stationary textured shape through the apertures flanking the lines. Alternatively, one may hold the front sheet of paper in a fixed position and horizontally move the striped disk back and forth behind the stationary apertures while tracking the disk with the eyes. At a suitable speed and line density, both animations lead to perceptual integration of uncontiguous image regions; one has a vivid illusory impression of colored texture in an area defined by crisp subjective contours tracing the outline of the disk.

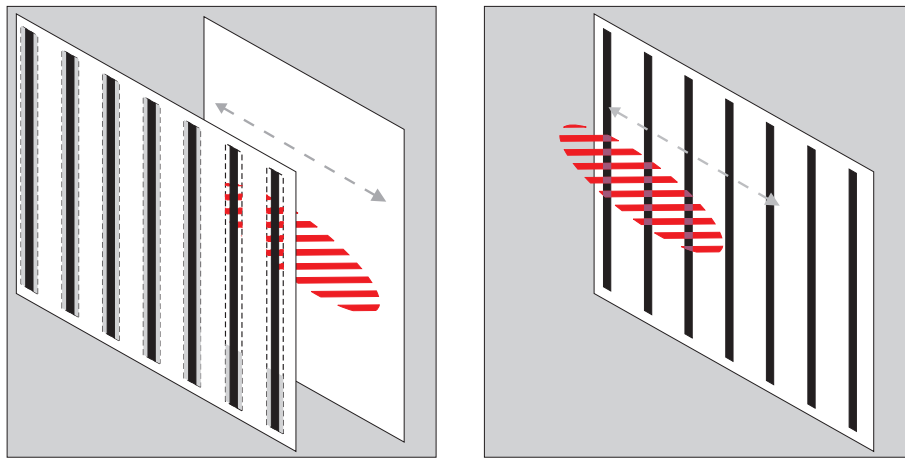


Figure 5.2: Possible 3-D interpretations of dynamic texture spreading: A frontal mask has several invisible slits in it, through which a textured object is seen as moving over a background. Black lines are painted between every second pair of slits. Alternatively, transparent colored stripes may be seen as moving in front of a background covered with lines.

The disk is perceived to move on a horizontal path, together with its attached texture. This is noteworthy because the motion-vectors generated by the texture fragments are equivalent to those found in the Barberpole illusion (Wallach, 1935). Hence, one could expect not to see coherent horizontal motion of all texture fragments, but individual vertical motion of each dotted line that appears through a single slit. That this is not the case shows that motion vectors are grouped and assigned to a unified surface. We may interpret this finding as another indication of the far-reaching influence of representations of a cohesive surface.

When the dynamic-texture-spreading animation is stopped, the completed boundaries as well as the filled-in texture do not disappear instantaneously, but fade away

gradually—analogous to flank transparency (cf. section 4.3) and dynamic occlusion (Hine, 1987).

5.2 Visual persistence

One important result that spanned all reported experiments on flank transparency was the smooth change in its appearance depending on gradual variations in spatial and temporal display parameters: Speed, the distance the color spreading needs to bridge, and figural contiguity all influenced the strength of surface appearance in a continuous manner. Notably the complementary influences of speed and distance deserve a closer inspection: To recapitulate, while increasing speed augmented perceived saturation of the illusory disk and the strength of its boundaries, increasing the distance had the converse effect.

If one leaves aside all figural aspects of completing a highly fragmented stimulus, and muses about explanations for the two result patterns mentioned above, it is not difficult to combine and understand them in terms of signal element density in space-time. Just as higher velocities lead to an increase in the number of fragments supporting an additional surface in a given time-window and retinal area, widening the space between disk fragments decreases the number of elements that pass over a portion of the retina during that time. What is still needed to account for the perception of an extended object based on a number of discrete input fragments is the notion that each fragment's influence on perception lasts longer than its actual imprint on the retina. Only then may the space-time gaps in the stimulus be bridged by the continuing effect of one fragment at a fixed location until the next one arrives. This idea has in fact been empirically supported in many studies on visual persistence (for a review see, e.g., Coltheart, 1980). To single out only one result, it was shown that an intense light flash visually persists for about 120 ms (Di Lollo, 1977), a temporal estimate that was often confirmed later on (cf. section 5.4).

As was forcefully argued for in Coltheart's (1980) seminal review, different semantic aspects of persistence need to be isolated early on in the discussion. These aspects refer to different levels of analysis and therefore are—to some degree—independent from one another. For one, the stimulus may continue to be present phenomenally after its physical disappearance, in which case we talk about *visual* persistence. Typically, the visible remnants of the stimulus fade out quickly. Second, there is an informational sense in which stimuli may persist: If one does not, in a strict sense, see the stimulus after its offset, but can still retrieve encoded visual information, one has to assume the stimulus persists in some form of iconic memory or *sensory* persistence. Finally, neural persistence refers to the neural activity resulting from a visual stimulus, but which lasts beyond the latter's physical duration. As these kinds of persistence are easily confused in different experimental paradigms, it is important to first identify the type that bears relevance to the question of dynamic completion, and then precisely

tailor the experimental methods to allow for its analysis. Since color and texture spreading by definition are phenomenological effects, it is visual persistence that is of interest here. As an explanation for visual persistence, the continuing access to visual information appears to be an obvious candidate, and sensory persistence is therefore predominantly considered in this context.

Sensory persistence asserts that sensory information remains in some form of retinal or post-retinal storage and could thus be integrated by a simple local mechanism over a certain time-frame. If, e.g., in the case of aperture viewing an observer would be able to perfectly track an object that is moving behind a stationary slit with his eyes, stimulus fragments would be presented such that the *complete* stimulus is subsequently spread over the retina over time. In that case, such temporal integration might be able to automatically recover the extended stimulus just by means of sensory persistence. No complex form of shape interpolation would then be required in addition to this form of “retinal painting”, as it is often called (figure 5.3). Indeed, if retinal painting was able to *fully* explain the observations made in multi-slit viewing situations, the observed visual completion would not constitute an interesting effect at all: It would then simply be an automatic side-effect of peripheral visual processing characteristics, and not the result of a specialized computational analysis.

Taken less literally, retinal painting may also be identified with some form of unspecified representation that is strictly retinotopic. This stance is preferred by some authors (Palmer, 2003; Shipley and Cunningham, 2001) because it does not restrict itself to actual retinally persistent information, and more explicitly stresses its contrasting qualities compared to visual representations in distal terms, such as object-centered ones.

Note again that this sketch of a possible explanation is completely negligent to questions asking for a reason why the visual system would at all interpret areas of background color as gaps in a depicted object, and not simply as, indeed, the background. This train of thought leads back to fundamental issues about what the cues are that signal object fragmentation, and may thus trigger mechanisms of perceptual completion. Candidates for this class of cues, e.g., tangent discontinuities or spatiotemporal discontinuities, were discussed in sections 3.2.1, 3.2.2, and 3.3.2.

Many of the available studies on dynamic boundary completion tested whether simple sensory persistence is a possible explanation for the observed discrepancy between the fragmented stimulus material and the perception of extended contours. Some authors found indeed strong support for a causal role of retinal painting in the perception of outline shapes under aperture-viewing conditions.

With respect to the perceived shape, retinal painting makes some straightforward predictions that are easily testable because the temporally summed retinal projection of a moving object seen through a slit depends on eye movements: If, e.g., an observer does not track a moving object perfectly but moves his eyes faster than perfect pursuit speed, the retinal image is elongated. Analogously, the retinal shape is compressed when eye

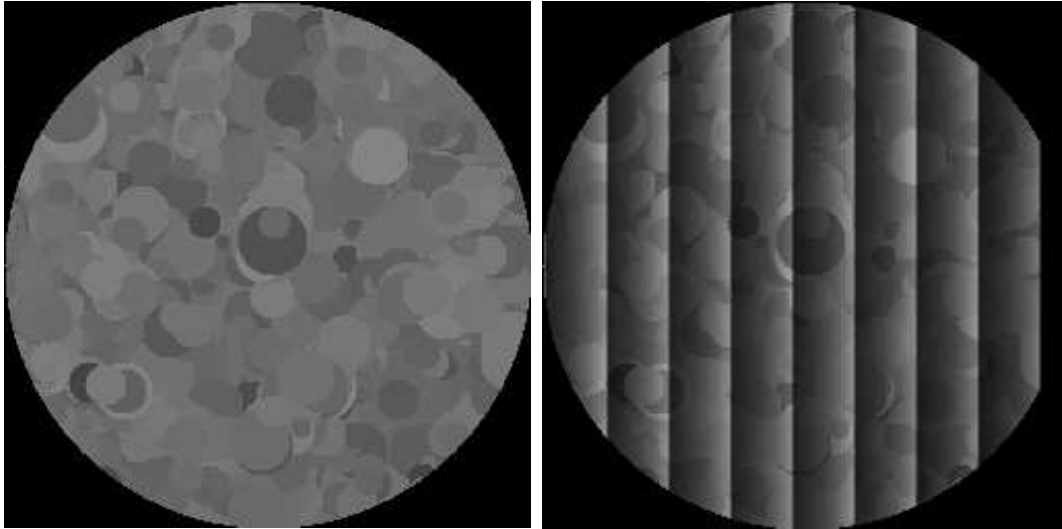


Figure 5.3: Simulation of the effect of retinal painting when viewing a randomly textured disk moving behind a set of slit apertures.

movements are slower than perfect pursuit velocity (Anstis and Atkinson, 1967). If eye movements are exactly the opposite of pursuit, the integrated retinal image is a mirror-reversed version of the normal one. It is clear that in general, projected retinal form does not necessarily correspond directly to perceived shape. Under anorthoscopic viewing conditions however, the integrated retinal shape may still be of high relevance. Anstis and Atkinson (1967) as well as many others (for a review, see Morgan et al., 1982) indeed demonstrated that the shape distortions of compression, elongation and mirror reversal are observed in aperture viewing, and follow closely the predictions made by retinal painting.

However, other authors claimed that retinal painting is neither necessary nor sufficient to explain the observed phenomena. As for *empirical* arguments against the importance of sensory persistence, a number of experiments tried with several means to prevent eye tracking of the moving object and thus retinal painting: Already Zöllner (1862), von Vierordt (1868), and Rothschild (1922) tried to show that anorthoscopic perception can occur even when a steady fixation on the stationary slit prevents the image of the object from being spread over the retina over time. With modern eye-tracking equipment, Fendrich et al. (2005) recently extended that work and found that boundary integration in anorthoscopic perception can still take place with retinally stabilized images. Anstis (2005) reported that anorthoscopic perception in multi-slit viewing also occurs with rotational movement which is hard to track with the eyes due to the inability to fully rotate them in the sagittal plane. Continuous object perception measured through object-tracking performance in a task from Scholl and Pylyshyn (1999) could not be achieved through eye movements, as four differently-moving disks

had to be tracked in parallel.

Another type of empirical evidence against retinal painting comes from experiments on motion-direction specific noise masking. Drawing inferences from his experimental results gained with a reverse-correlation method, Nishida (2004) argued that high-level motion analysis is involved in the perception of patterns viewed through multiple slits. He concludes that simple sensory persistence does not constitute a possible underlying mechanism.

In addition to the empirical arguments presented so far, several theoretical issues with sensory persistence were also pointed out: Shimojo and Richards (1986) showed that retinal painting can only work if the aperture problem was previously solved and the motion direction already determined by the visual system. While this still allows for cues from eye movements, Palmer (2003) pointed out that eye movements can, in principle, not compensate for simultaneous object motion in different depth planes, due to motion parallax. Another more general theoretical issue was put forward by Dixon and Di Lollo (1994): On the one hand, it is necessary for the visual system to integrate stimuli over time to guarantee perceptual continuity, but on the other hand, it must also detect rapid changes and must therefore segregate those stimulus elements that do not belong to one spatiotemporal object. This asks for a selective mechanism whose action is controlled by another processing instance. Retinal painting by itself does not provide a way to distinguish coextensive fragments in the stream of information from disjoint ones, i.e., for deciding when to integrate and when to segregate. This problem results from the fact that persistence itself is necessarily blind to any semantic structure of the scene in terms of objects and events. If active indiscriminately, persistence thus would operate not only within objects but also across object boundaries. This characteristic would blur the demarcation of objects, a highly undesirable consequence from the point of view of object recognition.

It appears reasonable to conclude that sensory persistence may be only one among several, and possibly more important, mechanisms involved in boundary integration. Yet, this conclusion does not automatically extend to the integration of *surface* attributes. Surface filling-in may constitute a different perceptual process that may therefore be mediated by different computational mechanisms. Whether surface completion involves or even completely relies on sensory persistence has not yet been explored. Considering that visual completion phenomena encompass a multitude of complex aspects regarding the reconstruction of boundary and surface information, it seems worthwhile to try to separate them and individually test which mechanisms might be responsible for each one. Retinal painting seems to be a very simple and evident candidate to explain some characteristics of dynamic texture spreading.

In section 3.3.1, we already came across some hints arguing for the involvement of retinal painting in surface quality perception: In bi-directional spatiotemporal-discontinuity displays, surface characteristics were absent, whereas in uni-directional displays, such as color from motion (cf. section 4.2), they are clearly apparent. An important difference between the two display types is that in bi-directional displays, no

coherent difference between the illusory surface area and surround exists that might be summed up over time. Therefore, we take sensory persistence as a first starting point in our investigation of dynamic texture spreading. More specifically, one motivation for our experiments was the question whether retinal painting provides a possible explanation for the interpolation of a surface feature like texture.

Retinal painting applied to dynamic color and texture spreading. Visual persistence allows for the possibility that temporal fragment frequency and spatial element density have equivalent effects within certain limits, and may thus compensate for each other. It asserts that a complete surface representation can be sustained by discrete, scattered signal elements barring that each point of the represented surface is frequently enough supported by new input signals consistent with that surface (figure 5.4). Applied to our stimulus setup, this means that a complete textured object would be perceived as long as the information signaling the presence of a piece of texture is refreshed often enough at each fixed location, regardless of how this refresh frequency is achieved. As the refresh frequency depends on the density and speed of the moving rectangular apertures through which object information becomes available, these two stimulus parameters should be important determinants of the texture-spreading effect.

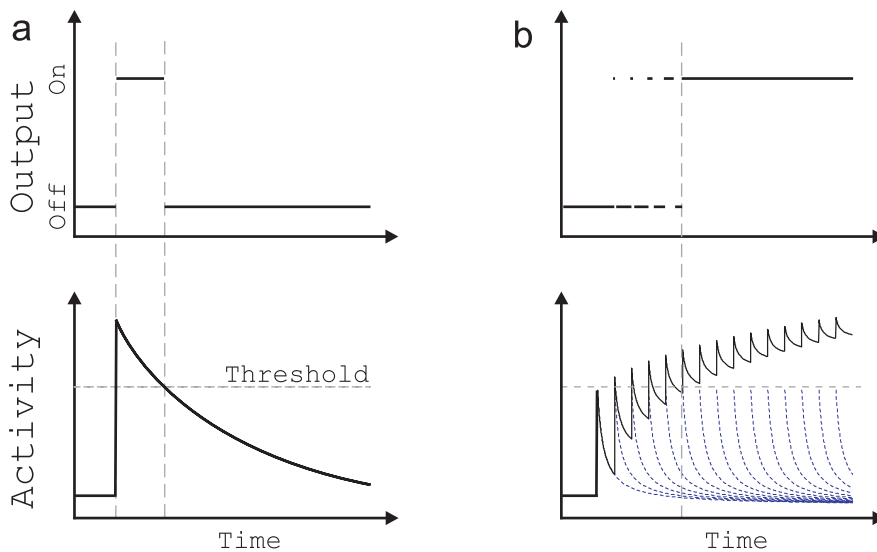


Figure 5.4: Schematic model of sensory persistence, read from bottom to top. a) A processing unit is assumed to have constant On output when activity is above some threshold. Activity is assumed to decay exponentially with time. b) When the processing unit gets externally stimulated repeatedly, temporal summation prevents the activity from falling below threshold level. As a result, the unit has a prolonged On output despite a discontinuous input.

This consideration is backed by empirical results about the relevance of signal element density for object recognition in animated sparse displays. In research on spatio-temporal boundary formation, Shipley and Kellman (1993a; Cunningham et al., 1998) found that shape identification improved with more signal elements in their displays, a result which was also reported by Mateeff et al. (1993) and Bull, Hunter and Finlay (2003). Fidopiastis et al. (2000) showed that the strength of color from motion is modulated by the particular layout of the dot array which provides signal elements for form and color information. Specifically, it was established that the color-spreading effect increases with a denser and more regular dot array. For color from motion (Cicerone et al., 1995) and for flank transparency (Wollschläger et al., 2002), a close dependence on the amount of motion has been observed.

It therefore seems natural to analyze dynamic texture spreading by asking about the specific interplay between the density of the line array and its speed in determining the filling-in effect. This will be the starting point for experiments 4a and 4b. Dynamic texture spreading particularly lends itself as a tool to conduct studies on surface completion since the contiguity of a texture or lack thereof is immediately visible.

5.3 Experiment 4: Distance-speed tradeoff

Experiment 4 consisted of two parts with complementary designs. Both addressed the question to what extent quantitative properties of texture spreading are compatible with the mechanism of sensory integration as delineated in section 5.2. This is done using a dynamic-texture-spreading stimulus to study the interplay between the density of the line array and its speed in determining the filling-in effect. Although experiments 4a and 4b do not directly test how much *surface* information is used to complete the task, results from experiments 5 and 6 will later attest that the dynamic-texture-spreading stimulus indeed provides relevant information about the surface layout. Since experiments 4a and 4b employ the same stimulus class, there is no reason to assume otherwise for their case.

Subjects. Ten subjects (five female, five male) participated in experiments 4a and 4b, among them nine students from the University of Kiel who were naïve as to the purpose of the experiment, and myself (DW). Subjects had normal color vision and normal or corrected-to-normal visual acuity.

Apparatus and Stimuli. The stimuli were displayed on a Sony Multiscan 500 PS 21 in color monitor with a vertical refresh rate of 75 Hz using the OpenGL graphics library. The timing of the stimulus frames was synchronized with the vertical refresh of the monitor. Color resolution for each pixel was 8 bits for each channel R, G, B . Following a standard procedure (Brainard, 1995), the monitor was calibrated and gamma-corrected

5.3 Experiment 4: Distance-speed tradeoff

using a colorimeter (LMT C1210). Observers placed their heads on a chin rest and binocularly viewed the display in a darkened room at a distance of 80 cm.

The background of the display was white (CIE $x = 0.30$, $y = 0.31$, $L = 55 \text{ cd} \cdot \text{m}^{-2}$) and subtended a visual angle of 26.7 by 21.9 deg. An array of vertically-oriented black lines (CIE $x = 0.30$, $y = 0.31$, $L < 0.2 \text{ cd} \cdot \text{m}^{-2}$), 6.75 min of arc wide and 3.37 deg long, was placed on the background.

In the interior of a virtual ellipse, red (CIE $x = 0.63$, $y = 0.34$, $L = 12.6 \text{ cd} \cdot \text{m}^{-2}$) striped flanks of 1.27 min of arc width and 2.78 cpd stripes were added parallel to the lines on either side. The virtual ellipse was centered on the screen and had main radii of 1.69 and 1.46 deg (figure 5.2). From trial to trial, its orientation was varied randomly in steps of 45 deg.

A fixation cross was placed on the horizontal center line at an eccentricity of 0.23 min of arc to the disk boundary, as informal observations had indicated that texture spreading was strongest under parafoveal viewing conditions, similar to what has been reported for anorthoscopic perception (Hecht, 1924), and static texture spreading (Watanabe and Cavanagh, 1991). The line array rigidly moved to the right, whereas the virtual disk remained stationary at the center of the screen. The velocity of the moving lines was varied by changing the length of the discrete “jumps” of the lines from one frame to the next. The trailing side of the display was continuously replenished with lines.

Procedure. In both experiments, we used a method very similar to that employed by Haber and Nathanson (1968), and asked subjects to find the speed or line-density threshold for perceiving a complete textured ellipse using an interactive adjustment procedure. Subjects were given two criteria for the perceptual presence of a *complete* ellipse. First, they should be able to discern the striped texture, and second, they should be able to identify the ellipse’s orientation, asserting that they saw the complete boundary. The suggested strategy was to first explore the perceptual outcome of the setting over a wide range of the values of the dependent variable, and then to close in on threshold level.

In experiment 4a, the speed of the moving line array served as the independent variable. It was varied in five steps: For the first group of four subjects, these five steps were 1.69, 3.37, 5.06, 6.75 and 8.44 $\text{deg} \cdot \text{s}^{-1}$. To a second group of six subjects, we later administered the same experiment with a slightly larger range of the independent variable which varied in the five steps of 3.37, 5.06, 6.75, 8.44, and 10.12 $\text{deg} \cdot \text{s}^{-1}$. The initial value of the line spacing was set to 0.67 deg. Subjects interactively controlled the spacing of the vertical lines in steps of 1 pixel using the arrow keys of the computer keyboard until the illusory ellipse was just noticeable.

In experiment 4b, the spacing of the vertical lines was varied in five steps. For the first group of subjects, these were 0.33, 0.49, 0.65, 0.81, 0.97 deg, corresponding to line

densities of 85, 58, 44, 35, and 29 lines visible. For the second, later tested group of six subjects, these were 0.27, 0.54, 0.81, 1.08, and 1.35 deg. The initial speed was set to $6.75 \text{ deg} \cdot \text{s}^{-1}$. Subjects interactively controlled the speed of the line array in steps of one pixel displacement per frame until they just perceived the illusory ellipse.

The chosen ranges for the two independent variables were derived from results in preliminary studies with the goal to obtain easily comparable functions in the two experiments relating line density and speed of the line array. That is, the values were selected such that the values for the independent variable in experiment 4b would be similar to the dependent-variable settings in experiment 4a, and vice versa.

Experiments 4a and 4b were carried out separately with the order of the independent variables counterbalanced across subjects. In each experiment, subjects performed 10 repetitions for each of the 5 conditions in pseudo-random order, resulting in a total of 50 trials.

Prior to the experiments, all subjects familiarized themselves with the task of adjusting the speed and spacing of the line array, respectively. There was no time limit for either task—subjects pressed a key to confirm their setting, and continued with the next trial.

Results. Figure 5.5 displays the combined results for the two experiments for two individual subjects (one from each of the two subject groups) as well as averaged across subjects within each subject group (see table B.4 for individual data from all subjects). Line spacing is plotted against the speed of the moving line array. In both experiments, these two variables seemed to be linked by a positive, approximately linear relationship. In the following account of the results, I will refer to the data in terms of line density, which is inversely related to line spacing. It is evident that for both types of threshold determination, the density of the moving line array and its speed were in a compensatory relationship. This means that in experiment 4a, line density could be reduced with higher speeds while maintaining the illusory percept, or, conversely, line density had to be increased with slower speeds to obtain the illusory percept. For experiment 4b, results indicate that with lower line density, speed had to be increased to reach threshold, and with denser line spacing, speed could be reduced.

As intended, the functions relating speed and line spacing in the two experiments had mostly overlapping domains and ranges, and they can thus be easily compared. The functions were separated by a small horizontal offset, but were quite similar in shape. This suggests that the results of the two experiments indeed describe the same relationship, and can be combined into one data set. Yet, the two functions did not coincide exactly. The observed difference between them may be attributed to small shifts of the subjective criterion between experimental sessions.

As is apparent in figure 5.5, there was some inter-subject variability in the absolute settings of the dependent variables. The narrow error bars indicating the SEM however

5.3 Experiment 4: Distance-speed tradeoff

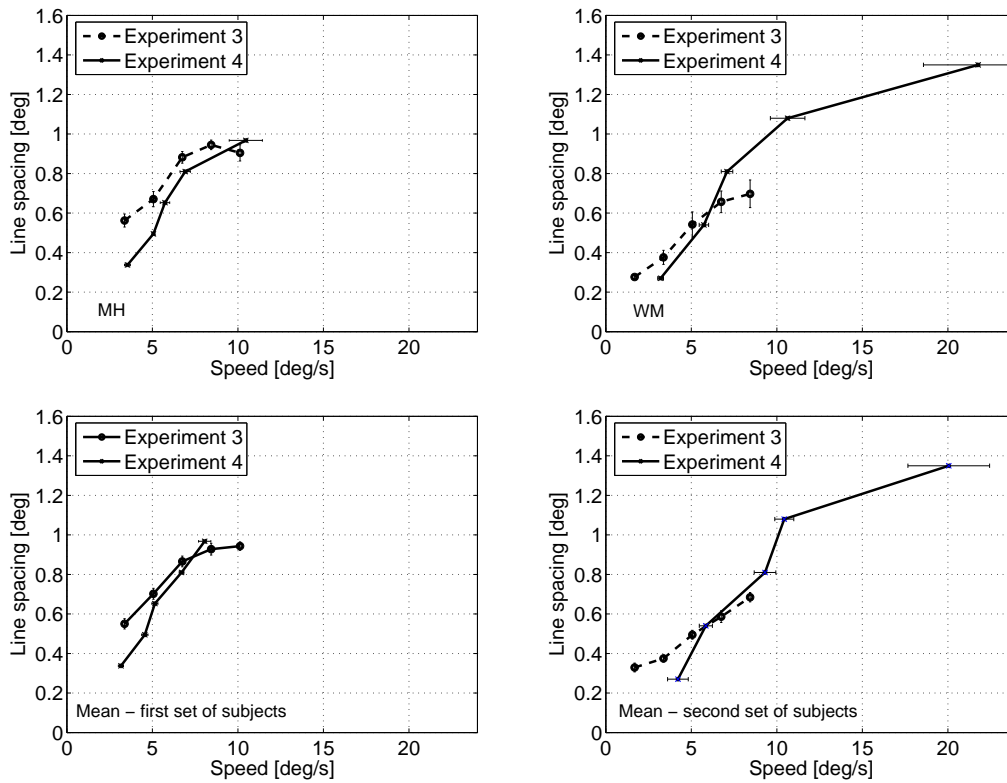


Figure 5.5: Results from experiments 4a and 4b. Distance of the moving lines is plotted against the speed of the moving lines. Error bars indicate ± 1 SEM (see text).

show that subjects' settings were in general quite consistent within conditions.

Based on these data, it is possible to calculate the delay between two consecutive appearances of a flank at a given fixed location within the virtual ellipse at threshold level. Note that in the calculation of the tolerated delay, we make several simplifying assumptions: First, we ignore the contribution of physical persistence from the monitor phosphors as most phosphors have a fast initial decay. Furthermore, dynamic texture spreading can be observed equally well and with similar parameter settings when the stimulus is generated by a video projector. Second, we only take into account the delay that results from the separation between different lines, not that resulting from the separation of the two colored flanks along one single black line. With large line distances, this influence is negligible, with very short distances, however, it could add noticeably to the data. The reported delay data should therefore be taken as approximate. The delay between two refreshes of texture information at a stationary point is plotted in figure 5.6 against the speed of the moving line array for both experiments. The pattern of results slightly differed between subjects, but in each case the data

points varied within a close range from 100 to 150 ms. The tolerated delay increased non-linearly with very low speeds in experiment 4a, indicating that the line spacing setting never fell below a fixed threshold value. This is to be expected from the fact that a small distance between lines can be bridged by color or texture spreading even in static displays.

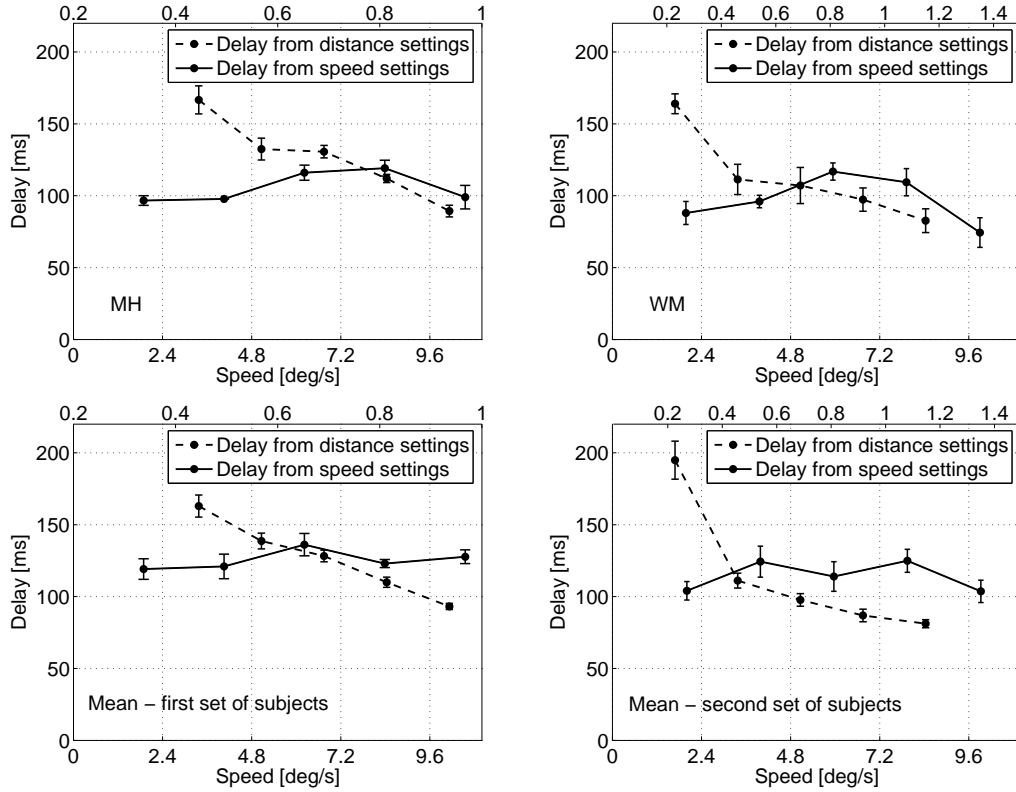


Figure 5.6: Results from experiments 4a and 4b, re-analyzed in terms of the maximally tolerated delay between two consecutive appearances of a textured flank at threshold level. Error bars indicate ± 1 SEM (see text).

Discussion. The results from both experiments show a fair degree of consistency within, as well as across subjects. This seems worth noting, considering the peculiar perceptual nature of most illusory percepts on the one hand, and the employed adjustment method of threshold determination on the other hand. Any vagueness in the percept would have necessarily resulted in highly variable threshold settings. Since the individual settings actually stayed in a reasonably close range, one may infer that the percept itself is stable—a prerequisite for any further data analysis with regard to possible perceptual mechanisms. Furthermore, the results are consistent with those obtained in experiment 1 regarding the enhancing influence of motion on the perceived

amount of spreading.

The very existence of dynamic texture spreading strongly calls into question the explanation attempt of homogeneous color spreading through color assimilation or the “diffusion” of neural activity (cf. section 4.3.6). These mechanisms would predict a diffuse spreading of color in all directions, and thus do not seem to be compatible with the observed non-homogeneous texture filling-in which is visible as a clearly demarcated pattern.

In line with results presented by Morgan (1980; Morgan et al., 1982), the idea of simple temporal integration through sensory persistence (cf. section 5.2) was able to capture the data from both experiments surprisingly well: The density of the line array and its speed seem to act together in a compensatory manner in determining the strength of the illusory percept. Within the studied parameter range, this means that a change in the strength of the illusory percept due to, e.g., a reduction in line density could be compensated by a corresponding increase in the speed of the line array. We may thus further conclude that according to our data, retinal painting may qualify to act as a mechanism subserving position updating in spatiotemporal relatability theory (cf. section 3.3.2).

That sensory persistence accounts well for our data warranted a further analysis in terms of one of its important parameters: The maximum delay between two input signal refreshes that is tolerated by the visual system while still yielding a complete percept. This delay was not exactly constant over experimental conditions as would be expected from retinal painting, but it stayed in a narrow range from around 100 to 150 ms (figure 5.6), indicating the maximum time window for integration processes. This estimate is remarkably similar to results obtained in a number of other experiments on object recognition, as will be outlined in the following section 5.4.

Despite the compatibility of our results with retinal painting, figural influences on dynamic texture spreading make it clear that persistence alone cannot tell the whole story about the dynamic completion mechanisms: By simply removing the black lines from the display such that only the invisible slits remain in the frontal masking layer (figure 5.2), no information that would be relevant for the integration process itself is removed from the stimulus. Yet, informal observations indicate that no texture spreading occurs with this setup. Apparently, without the black lines, no cues for a segmentation into different perceptual layers, as well as for a subsequent grouping and surface completion remain in the display. A second informal demonstration also points towards the importance of visual processing steps that go beyond mere peripheral transformations of the spatiotemporal input: One can not only construct the apertures which make visible the textured surface in the form of elongated slits, but can also implement them as small rectangular “windows” which open and close randomly from one frame to the next. Given a suitable choice for the aperture size and the number of “open” apertures at each point in time, the relative proportion of visible to invisible surface area can be made the same as in the multi-slit case. Therefore, from the point of view of spatiotemporal integration, visual completion of the textured surface should

obtain equally well with both stimulus types. This, however, is not what is actually observed with random apertures. Here, the shape of the surface can be easily discerned, but no surface filling-in obtains. Instead, local element motion is seen, much like it is with dynamic white noise.

The aforementioned examples are equivalent to multi-slit viewing in terms of the available amount of surface information that can be spatiotemporally integrated, e.g., through a mechanism of sensory persistence. Nevertheless, the perceptual outcome is quite different from texture spreading. Together with observations indicating that homogeneous surface filling-in in the form of flank transparency is subject to figural influences (cf. section 4.3.5), these examples of failing surface completion thus demonstrate that surface filling-in is indeed an interesting effect that cannot be completely explained away with a trivial mechanism.

5.4 The timecourse of visual completion

As was pointed out in the discussion of the previous experiments, it looks like integration of stimulus elements occurred within a time window of about 100–200 ms. To interpret this conclusion, it is necessary to embed it into the relevant research context of processing times in vision. We will thus get slightly side-tracked here to briefly delve into the discussion about the general timecourse of visual perception, and the role that temporal integration plays in it. While this discussion is too broad to be adequately covered here in its entirety, it needs to be at least sketched out because it bears high relevance to the question to what degree visual persistence is responsible for spatiotemporal completion (cf. section 5.2).

Subjectively, visual perception is almost always instantaneous. This means that we perceive external events to take place at the same time as we internally gain consciousness about them. In other words, our perceptual world does not seem to “lag behind” the events of the physical world in any way—when we open our eyes, the visual world is instantly at our disposal, and does not gradually build up. This is in contrast to the rare occasions where it takes us noticeable time to develop a clear visual representation of an object, e.g., under adverse viewing conditions such as at dusk.

However, in physiology it was discovered early that neural activation spreads only at a finite speed. Nowadays, the measurement of the velocity of neural propagation is a standard procedure, and leads to a range of about $10\text{--}120\text{ m}\cdot\text{s}^{-1}$ for mammalian motor neurons, depending on the thickness of their myelin sheath. As a simple consequence, this implies that cortical processing of the visual input necessarily lags behind the actual physical event some tenths of milliseconds. How can this result be reconciled with the subjective immediacy of vision? The following paragraphs are intended to provide the background for this question. First, we will try to make the theoretical limits on the speed of processing a little more precise, and will then present empirical findings on the timecourse of forming complete visual percepts.

5.4.1 General visual processing

Physiological limits in the nervous system cause that after a brief light flash, the information flow from the retina to V1 takes about 40–100 ms (Berry, Brivanlou, Jordan and Meister, 1999). This time is partially due to sluggish signal transduction in the photoreceptors and the limited speed of neural activity propagation. Cortical processing itself requires roughly 110–130 ms (Rodieck, 1998), and another 40 ms are then spent while motor signals travel to the efferent motor system and the muscles are activated. Adding up the different latencies, this means that up to 200 ms pass before a motor reaction can occur in response to an event. This figure is consistent with reaction-time experiments showing that the majority of subjects can execute a motor response within 180–200 ms after a stimulus onset (Rodieck, 1998). Of course, the specific timing may vary considerably between situations, as different tasks may require differential processing.

In many instances, however, such a reaction would occur far too late because motion may change an object's position considerably during that time as is evident when considering the classical example of Baseball: A good pitcher can throw a fastball at 90 mph $\approx 40 \text{ m}\cdot\text{s}^{-1}$. The pitching mound's center is located 60.5 ft $\approx 18.4 \text{ m}$ away from the home plate, meaning that when the batter sees the ball 200 ms after it had left the pitcher's hand, the ball has already travelled 8 m, nearly half distance. This leaves the batter only about 260 ms to completely execute a swing. Note that even this extremely short timeframe only allows the prediction of the ball's motion trajectory to be based on the glimpse of the ball leaving the pitcher's hand. For a batter to know 260 ms in advance where to hit the ball, visual perception needs to be a fundamentally predictive system. Apparently, evidence for this can already be found in the retina (Berry et al., 1999). By foretelling the state of the environment for a few hundred milliseconds, the visual system serves the goal of seeing things as they presumably are at present, and not as they were moments ago. Predicting an object's future state requires the integration of past information about it, which is why the idea of temporal averaging of sensory information here plays an important role, especially in the discussion of motion perception. In motion perception, already the basic phenomenon of φ -motion makes evident that the perception of smooth motion paths involves some mechanism of spatiotemporal integration.

Connected to the subject how our phenomenological sense of simultaneity and event perception is tied to the chronology in the physical world is the question whether visual perception is inherently continuous or discrete. Again, subjectively, our experience of the world is a continuous and seamless one. Nevertheless, there may be a reason to believe that perception occurs in short processing epochs, or snapshots, in which perception is assumed to be constant (VanRullen and Koch, 2003). Events occurring within these snapshots would be integrated and thus be indistinguishable from each other with respect to their timing. While the merits of this idea are hotly debated, the undisputed concept of temporal integration itself is independent from it, because it

need not be confined to working only within one epoch, but may also be continuously active.

If no discrete processing is assumed, temporal integration can be understood as a continuous convolution of the input with a smoothing kernel, i.e., as a temporal low-pass filter. This filter blurs sudden state changes by letting past events influence the perceptual state that results in response to current ones, e.g., in the form of simply taking the moving average as the current state. Sensory persistence would indeed have just this effect, and thus represents a possible mechanism for it.

Since it probably represents the hard limit of temporal resolution, several studies tried to determine the precise interstimulus interval between two events that prevents them from being integrated into one percept. Estimates ranged from 20–60 ms between two successive visual stimulations to render them distinguishable from each other (Paradiso and Nakayama, 1991; VanRullen and Koch, 2003). These results are in line with Bloch's law which states that approximately within 100 ms, the duration of a stimulus and its intensity have equivalent effects and may compensate each other. Since no difference between two stimuli of different presentation duration but appropriately adjusted intensity may be detected, normal temporal resolution cannot be much better than 100 ms.

Different approaches have been taken to figure out precisely how integration processes work, some of which pertain to the perception of fully visible stimuli, and some to visual completion proper. The former ones shall be considered first: Measuring perceived elongation of moving objects, Burr (1980) found that for moving stimuli which are presented long enough, motion smear is much less than what the commonly reported integration-time estimate of 120 ms may suggest. Only for stimuli presented very shortly, his subjects observed elongations consistent with a persistence of 120 ms. Burr concluded that sensory persistence is normally active, but can be suppressed when motion processing is triggered. In EEG studies on the timecourse of neural signals in object recognition, Johnson and Olshausen (2003) found recognition-correlated ERP components between 150–300 ms. Similarly, Thorpe, Fize and Marlot (1996) reported an EEG component at 150 ms linked to the detection of objects in complex natural scenes. Studying temporally weighted spatial averaging in the flash-lag effect, Eagleman and Sejnowski (2000) came to results suggesting that integration takes place within an 80 ms window.

When one assumes that the complete space-time development of an event is given, general temporal smoothing would also allow for future stimuli to take effect on the perception of past events. This would make it equivalent to, e.g., the perception of motion direction by smoothing the motion vector field (Hildreth and Koch, 1987). Theoretically, time-symmetrical smoothing would often lead to better interpolation of motion paths than mere prediction. Empirical evidence points to the possibility that the visual system may actually employ this strategy with a symmetrical integration window of 90–112 ms around an event (Rao and Sejnowski, 2001). However, such

behavior constitutes a case of postdiction which cannot be implemented by sensory persistence alone.

5.4.2 Completion times

The aforementioned results on the general nature of time and event perception, in addition with hints at the importance of temporal integration—possibly through sensory persistence—encourage further inquiries into the timecourse of visual completion in particular. Only very few filling-in phenomena take noticeable time, among them those happening with stabilized images (Gerrits et al., 1966) or in scotomas (De Weerd, Desimone and Ungerleider, 1998; Ramachandran and Gregory, 1991). In these cases completion is a slow process taking 2–3 s. In contrast, almost all examples of visual completion, be they modal or amodal, appear to be instantaneous, i.e., object parts that are blocked from view are represented very rapidly. As Tse (1998) argues, this rules out top-down processing as a major influence in filling-in, simply because it would be too slow. Still, the results presented below all make it clear that completion happens at finite speed, and has indeed a measurable timecourse. By comparing the different studies and their conclusions, it will become apparent that time to completion is probably not constant, but depends both on the chosen stimulus and task.

Studying the integration of local element changes in the perception of subjective surfaces in spatiotemporal boundary formation, Shipley and Kellman (1993b; 1994) came to the conclusion that signal integration takes place within 150–165 ms, a range that is similar to the value of 120 ms reported by Guttman and Kellman (2004) for contour completion. This duration is also consistent with Mather's (1988) results that apparent motion of subjective contours becomes visible 85–100% of the time if the stimulus lasts longer than 160 ms. Supposedly, by this time object fragments were joined to form a complete percept.

Johnson and Olshausen (2005) recently conducted an EEG study with a recognition task using incomplete objects in the presence and absence of their occluders. Their conclusion was that, depending on the particular task, recognition-related potentials could be measured between 130 and 180 ms after stimulus onset. Using the performance in a shape discrimination procedure as a measure of completion with highly occluded stimuli, Ringach and Shapley (1996) estimated that 120–170 ms is required to build up a functional stimulus representation. Using stimuli that are typically subjected to amodal completion due to available occlusion cues, Rauschenberger and Yantis (2001) showed in a visual search task that access to pre-amodal completion information is still available for up to 100 ms after stimulus onset. Thereafter however, stimuli behaved equivalently to fully specified ones, indicating that the completion process had already finished. Only when larger areas were occluded, completion took longer, i.e., up to 250 ms. Note that all of these time windows are fairly consistent with the one calculated from the results of experiments 4a and 4b, reported in section 5.3.

Evidently, the presented studies come to similar, but not identical conclusions regarding the temporal development of completed percepts. Importantly, estimates of stimulus integration time for the perception of fully visible objects do not differ substantially from those obtained with fragmented stimuli. Based on the observation that dealing with highly incomplete input information indeed represents the normal viewing conditions for the visual system (cf. chapter 2), such a conclusion is not surprising. Still, comparing the results obtained in different experiments should only be done cautiously, since it cannot be assumed *a priori* that they are all comparable with respect to the type of completion they analyze. This caveat notwithstanding, the situation that a time window of about 100–200 ms consistently shows up as the required completion duration represents converging evidence that completion is not instantaneous, arguing for an actively progressing process that takes measurable time. The diversity of the results then implies that completion should not be regarded as a constant in visual processing, but rather seems to be variable, depending on the stimulus and task that the visual system is facing.

One single factor that was repeatedly shown to influence completion time is the size of the area that needs to be filled in, with less occlusion resulting in faster completion (Rauschenberger and Yantis, 2001; Shore and Enns, 1997). Sekuler, Gold, Murray and Bennett (2000) showed that time to completion varies from 100–200 ms depending on degree of occlusion. This range that was replicated by Guttman et al. (2003) in a primed matching study with an estimate of 75–200 ms.

5.5 Methodological issues and the functional role of illusory percepts

In the previous section, we were repeatedly confronted with experimental methods seeking to turn the timecourse of visual completion into a measurable entity. Often, the approach taken was to devise a task whose successful processing ideally requires a completed stimulus representation. If one looks closer at the employed methods however, this path looks much less straightforward than what it may seem at first. In addition, the choice of a well-suited methodology is inextricably connected to several assumptions about the general nature of illusory percepts. These are of high theoretical interest by themselves, but only rarely spelled out explicitly. This insight motivates a second, more in-depth look at some of the experiments already reported, with a focus on general methodological questions related to the study of visual completion. At the same time, the following paragraphs will also provide the basis for considerations that led to the design of experiments 5 and 6, reported on in sections 5.6 and 5.7.

If one wants to measure the timecourse of visual completion, only the beginning of the involved mechanisms is easy to figure out as it simply coincides with the onset of the stimulus. In contrast, knowing *if* a completed percept is eventually formed is much more difficult, and determining *when* exactly the relevant visual processing finishes

even more so. Regarding the mere presence of a completed feature, introspection may in many cases be unambiguous and may lead to uncontroversial results: Few people will contest that strong illusory contours can be seen in the Kanizsa triangle. Examples of amodal completion (cf. section 3.2.2), however, are already more difficult to judge. Of course there may be a strong feeling that an object is incomplete and that it continues behind an occluder, but how reliable is this impression? And, perhaps more importantly, to what extent does it make an assertion about visual perception proper, and not about cognitive reasoning?

In some instances, it is obvious that reasoning is involved in our assessment of incomplete objects: When we see a familiar person being occluded by a table from its mid-section downwards, we might very well say that we “see” this person in the sense of knowing its appearance. But this statement then just means that the shape of the missing bottom half is *suggested* by memory traces from previous exposure. In the same vein, some fragmented figures may *remind* us of certain shapes without them being actually visible in the strong sense. Kellman et al. (2001) argue that in other cases, cognitive reasoning is harder to tell apart from genuine perceptual processes, e.g., in the perception of occluded, possibly symmetric or familiar objects (figure 5.7). In section 3.2.1, we also came across some effects like static texture spreading which are very faint in general. Simply telling subjects to describe or rate their appearance is prone to be unreliable, yields inconsistent results across subjects, and asks for instruction-induced biases in the response.



Figure 5.7: Amodal completion of a possibly symmetric, and a familiar object. Statements about the objects’ perceived shape may be inadvertently influenced by cognitive reasoning.

Clearly, introspection may sometimes not be the best task to select when studying research questions in visual completion. This is especially so when amodal completion is concerned or the effects are barely visible. For this reason, studies of visual completion have in recent years increasingly employed novel methods that are designed to be less susceptible to cognitive influences. Some of them even seek to be objective, i.e., to use tasks that objectively have a right answer. For many of these tasks, the general idea behind their construction is to assess the effect of the completed percept in question on some measurable entity, and to compare it to the effect of an uncompleted percept: If a

filled-in percept has the same perceptual consequences as a normal one with respect to some relevant measurable function, perception is assumed to be complete as indicated by functional equivalence.

The aforementioned methodological approach rests on several assumptions which are worthwhile to be made explicit and analyzed in their own right. For one, no principal distinction is assumed to exist between “real” and “illusory” percepts. Given the fundamentally inferential nature of visual perception, such a distinction would be hard to argue for anyway. Supporting this assessment, several studies have found that for many cortical processes, visually completed stimuli are indistinguishable from non-completed ones. There are, e.g., neurons in V1 that respond to subjective contours, i.e., their receptive field is not directly stimulated, but would have been, had the contour been physically closed (Kellman et al., 2001). Many other cases of neural equivalence of filled-in percepts are now known (Kellman et al., 2001). In this context more important however, is the large set of behavioral observations, both qualitative and quantitative, that indicate the same.

5.5.1 Qualitative demonstrations

The following paragraphs will give a brief survey of qualitative demonstrations of illusory percepts. They suggest that completed objects can act just like non-completed percepts in many different cases of visual processing: For example, with the double-neon-color illusion, van Lier (2002) presented an animation of two illusory neon color stripes, induced by the Varin figure (figure 3.10a; cf. section 3.2.1). By letting the visually-completed portions of these transparent-looking stripes partially slide over each other, this stimulus leads to the illusory perception of a somewhat darker, less transmittant part just where the two stripes seem to be stacked. This effect thus mimics the consequence of using real transparent films lying over each other and shows that, together with visual completion to a transparent surface, the illusory percept automatically has attributes and causal consequences tied to it that are the same as for real ones.

Many other phenomena imply that once a perceptual category is established for a stimulus, it can have the same causal consequences in visual processing, irrespective of its real or illusory origin. Some of these studies concern different kinds of aftereffects while others show that illusory entities either have the same additional qualities as real ones do, or can feed information into other systems as effectively as those. Interpreted more generally, this has important implications for questions regarding the kind of input that perceptual mechanisms operate upon, and hence the visual system’s architecture in general (cf. section 6.1). All of the forthcoming examples illustrate that visual completion does not simply generate illusory percepts that are completely detached from normal visual processing, but are an integral part of it. Since completed percepts are intimately tied into functionally relevant perceptual mechanisms, it will also become

clear that they are indeed a genuine visual phenomenon, and not only appear to be so as a result of cognitive reasoning:

- Shimojo, Kamitani and Nishida (2001) showed that illusory surface color generated by neon color spreading causes the same kind of chromatic afterimages as real coloration does.
- Similarly, Weisstein et al. (1977) demonstrated that when put in motion, visual phantoms (cf. section 3.2.1) generate motion aftereffects that can affect non-illusory percepts.
- Measuring the consequence of color filling-in that takes place in response to retinal stabilization, Piantanida (1985) concluded that filled-in color reduces subsequent temporal modulation sensitivity as effectively as physically-based color percepts of the same kind.
- Berkley, DeBruyn and Orban (1994) and Paradiso, Shimojo and Nakayama (1989) conducted studies indicating that the tilt aftereffect crosses between real and illusory lines. This means that just as the orientation of real lines will modify perceived orientation of subsequently-seen illusory lines, the tilt of illusory lines can alter perceived orientation of real lines.
- Michotte et al. (1964) reported that amodally completed objects can initiate φ -motion.
- Carman and Welsh (1992; see also Nakayama and Shimojo, 1992) stressed that modally completed 3-D surfaces have attributes of view stability and shape invariance under changes of viewpoint.
- In an intriguing study, Chen and Cicerone (2002) used the color-from-motion effect (cf. section 4.2) to induce “illusory” binocular disparity exclusively from completed modal boundaries, and showed that this causes a corresponding depth perception with high reliability.
- Nakayama et al. (1989) created a face recognition task using a strongly fragmented face. When an occlusion interpretation was favored by giving the blocking fragments near disparity, face recognition was easy, indicating face fragments were grouped together in the course of amodal completion. Performance was much worse, however, in an otherwise identical display where the face fragments carried near disparity and did not appear occluded. A purely perceptual change thus can have important functional consequences. More recently, Johnson and Olshausen (2005) came to similar conclusions after obtaining data in recognition tasks involving incomplete objects either in the presence or in the absence of their occluders.
- Yantis (1995) studied the perception of group vs. element motion in the Ternus display which depends on the interstimulus interval. He found that by “occluding” the middle element in an interleaved frame with a perceptually filled-in Kanizsa

square, he could influence an observer's perception towards a strong bias for element motion. Moreover, this causal role of the illusory occluder was tied to it being in front, consistent with an equivalent physical setup.

- A reverse correlation technique was used by Sekuler et al. (2000) in combination with a shape discrimination task involving stimuli of curved Kanizsa squares with added noise. The authors established that observers used for their decisions spatial locations of the stimulus where no border was physically present, but where interpolated boundaries were seen. This again implies the use of information that is perceptually generated without direct physical foundation.

Information blocking. Whereas completed objects can carry information that is adds to that which is present in the fragments themselves, there are also experiments which indicate that a completed stimulus representation actively preempts information about features of the non-completed stimulus. This means that once an object has been visually completed, information about aspects of the stimulus *fragments* that gave rise to it may become inaccessible.

Yin et al. (2000) showed that sensitivity to binocular disparity is impaired when surface feature integration indicates a depth ordering that is in conflict with the disparity of single fragments. Thus, surface integration is a process that interacts with other cues, and is not simply to be thought of as a uni-directional, bottom-up integration of information sources. Not only can surface features compete with disparity, according to studies carried out by He and Nakayama (1992; 1994), the formation of completed surfaces preempts lower-level features such as luminance-based contours. This was shown in visual search and texture segmentation tasks where surface completion can override information in the proximal stimulus that would have been beneficial to the efficient handling of the task at hand. The loss of functional access to low-level image information is thereby not instantaneous, but is effective only after about 100 ms (Rauschenberger and Yantis, 2001; cf. section 5.4). Supporting evidence for an active elimination process of information existing prior to completion comes from Paradiso and Nakayama (1991): These authors described experiments where they were able to block the progression of brightness filling-in with a mask whose stimulus onset asynchrony ranged from 50–100 ms. Apparently, the brightness filling-in started at the target-disk's border and moved inwards. When stopped by a mask, only a ring-shaped area was reported as being filled-in.

Based on experiments with spatiotemporal-discontinuity displays (cf. section 3.3.1), Shipley and Kellman (1997) pointed out that local element motion is not seen when the percept is one of global object motion. Yet, this would be a viable interpretation of the stimulus: This becomes apparent when figural cues prevent local motion signals from being integrated into global object motion, the individual element motion signals are then indeed perceived. The same observations were made concerning boundary

motion vs. element motion in dynamic occlusion displays by Bruno and Gerbino (1991; cf. section 3.3.1).

5.5.2 Quantitative methods

The aforementioned effects qualitatively demonstrate that in many situations and for many processing demands that completed percepts can behave indistinguishably from normal ones. However, as we have seen throughout, visual completion often is a gradual phenomenon: Some completed percepts appear more distinct than others, and even for the same type of filling-in, different levels of completion strength are possible, as was concluded from experiments 1, 2, and 3. A further challenge is to make this continuous range of appearance somehow measurable. With color matching, the first three experiments already incorporated one method to quantitatively grasp the different qualities of a completed percept, but it was a subjective one. Next, we shall give a short overview over other possible experimental approaches to reveal qualities of completed percepts in quantitative tasks—many of them standard psychophysical paradigms adapted to the particular stimulus material.

Shape discrimination. Shape discrimination tasks are the most common quantitative methods found in studies on visual completion. A popular way to administer it is to devise a set of artificial, meaningless figures, present fragments of them, and have subjects decide which one they saw (e.g., Anstis, 2005; Kellman and Shipley, 1991). Variations of this methodology include one employed by Murray, Sekuler and Bennett (2001), who used partially visible rectangles as stimuli that were presented for different durations. Subjects had to decide whether the horizontal or vertical side was longer, and sensitivity could be assessed with respect to the physical ratio of these sides. The authors found that discrimination sensitivity was much higher when the fragments appeared to be occluded and were supposedly perceptually grouped than when no such interpretation was favored. In contrast to the similar task from Nakayama et al. (1989; see above), this method allows for more fine-grained assessment of degree of completion because sensitivity varied continuously with presentation time. A perceptual classification task was also used by Ringach and Shapley (1996). They had subjects discriminate between curved Kanizsa squares looking convex (fat) or concave (thin), depending on the four pacmen's angle of rotation. Similar accuracy was found for modal and amodal completion conditions, both leading to much better performance than a fragmented control condition.

As a variation of the fat vs. thin task described above, Guttman and Kellman (2004; see also Guttman, Sekuler and Kellman, 2003) asked subjects to localize small dots as being either inside or outside the curved Kanizsa square. This method was intended to make the interpolated contours more precisely traceable, and confirmed the conclusion that completion conditions allowed for better accuracy than a fragmented control condition. Nishida (2004) added a reverse correlation technique to a shape

discrimination task in multi-slit viewing (cf. section 4.1). Using letters as the to-be-identified objects, he showed that observers used spatial frequencies for successful discriminations that could only have been recovered by spatiotemporal integration.

Davis and Driver (1997) studied the spreading of attention on fragmented surfaces in conjunction with a discrimination task. This method can be used as a way of probing the functional unity of image fragments because cued attention spreads within same-object regions but does not spill over to other objects. Davis and Driver (1997) constructed stimuli containing objects that looked like they were interrupted in their middle but still belonged together. If one then probes the subject by presenting a target within a fragment on one side of the breach, and asks him to discriminate it, reactions will be faster if the fragmented object had been cued beforehand on the other, disconnected fragment. Supposedly, this means that after completion, attention could subsequently spread on the surface linking the two fragments. Davis and Driver (1997) manipulated the type of completion of the surface fragments through depth-placement, and found that attention spreads on near, modally completed surfaces just like on real ones. This, however, was not the case on far, occluded ones. Note that the difference between modal and amodal completion found here does not only concern the phenomenological level, but also has functionally relevant consequences (cf. Singh, 2004; Singh and Anderson, 2002).

Visual search. Visual search is also a widely-employed paradigm. It was used by Davis and Driver (1998; see also Gegenfurtner, Brown and Rieger, 1997) to show that completed shapes may behave equivalently to fully visible ones. In addition, they established that visual processing was parallel for modally, but not for amodally completed objects—a result that again stresses the important asymmetry between modal and amodal completion with respect to their functional roles (cf. section 3.2.2). Visual search was also the method chosen by Rauschenberger and Yantis (2001) in their quest to uncover the timecourse of completion (cf. section 5.4), allowing for a gradual measurement of identification precision.

Priming. By using fragmented stimuli as primes for fully visible ones, Guttman et al. (2003; see also Sekuler and Palmer, 1992) checked whether visually completed percepts can enhance subsequent performance in a same vs. different shape comparison task as effectively as fully specified ones. In line with the conclusions drawn from a similar experiment by Sekuler et al. (2000), their results indicated that occluded circles can act equivalently to complete ones only if stimulus conditions induce amodal completion, but not else.

Apparent motion. Nakayama et al. (1995; see also He and Nakayama; 1994) devised an apparent-motion task to demonstrate that amodally completed surfaces can have equivalent perceptual effects as fully visible objects. The stimulus was a φ -motion

animation with two apparently occluded rectangles whose motion direction was inherently ambiguous (figure 5.8): Subjects alternate in their perceptions between horizontal and vertical motion. The preference for a particular direction thereby depends on the horizontal and vertical distance between the rectangles: Motion is dominantly seen between the closer rectangle locations. Amodal completion now effectively reduces the vertical distance, thus shifting preference towards the perception of vertical motion. The stimuli allows for an assessment of the gradual dependence of a preference for horizontal motion on the ratio of horizontal to vertical distance between subsequently shown squares. By comparing this dependence between an occluded condition (as depicted in figure 5.8) and a condition with the large middle rectangle removed from the display, it is possible to calculate exactly how much vertical distance between squares could be compensated for by amodal completion.



Figure 5.8: a) Two frames of a bi-stable φ -motion animation. b) Either horizontal or vertical motion can be perceived, depending on the horizontal and vertical distance between the squares' two positions. Amodal completion effectively reduces vertical distance, thus shifting preference to the perception of vertical motion.

5.5.3 Limits and problems of quantitative methods

With the use of priming, visual search, and rapid perceptual classification methods, the aforementioned experiments all make clear that seemingly early tasks like parallel visual search can have as input completed surfaces located in depth, which themselves are complex representations. This again stresses the important role of mid-level visual concepts (cf. section 3.2.3; Nakayama et al., 1995) and illustrates that the entry level into many perceptual processing modules can be quite high in terms of the representations involved.

Notwithstanding the fact that the introduction of performance measures and objective tasks to the study of visual completion has brought a significant advance in methodological rigor, it should not be overlooked that their predominant use and hasty interpretation can be problematic as well. This is because the objectivity comes at a prize: While these experiments can provide an objective measure of some functional influence of a subjective percept, inferences to the qualities of that percept always stay

indirect ones. In a sense, they are sometimes farther away from the phenomenon under consideration than plain introspective methods. These have a quality of immediacy and should therefore not be neglected—they remain useful tools in the study of visual completion (Nakyama et al., 1995).

In addition, objective tasks like shape recognition may sometimes be processed successfully based on local image characteristics without having been intended or noticed by the experimenter. In these cases, even though global perceptual aspects, e.g., surface qualities, are sought to be studied, they may actually not contribute to a subject's performance at all (Bruno, 2001). As an example, consider the recognition of letters, as used by Nishida (2004): Letters as such were specially designed to be optimally recognizable even under adverse reading conditions. They therefore contain many redundant local cues to recognition, keeping them still legible when portions are missing or are only vaguely visible. A similar argument pertains to the stimulus class of meaningless line drawings, widely used to study visual completion. These, too, may inadvertently be identifiable from only small portions due to idiosyncrasies in their shape. The adequacy of a chosen task for the desired research question thus needs to be established every time, and no stock methodology like shape discrimination or visual search may be used uncritically.

When explicitly studying surface aspects of visual completion, one faces the challenge to devise a task which does not confound local aspects, e.g., concerning the boundary, and global surface qualities. The development of adequate stimulus material thus remains a crucial prerequisite for the study of surface interpolation. Ideally, one needs a task which requires that a fully completed surface representation is at the observer's disposal in order to show a successful reaction. Preferably, such a task should also be a quantitative and objective one, allowing for a gradual assessment of the percept, without being susceptible to response biases.

Unfortunately, of the methods presented above, many do in fact confound global surface and local contour aspects—and of those that do not, some are qualitative, like, e.g., the ones used by Yantis (1995). This implies the need to develop new, more suitable methods, that are specifically tailored to assess surface spreading. Attempts at doing so will be presented in the following sections 5.6 and 5.7, together with the experiments employing them.

5.6 Experiment 5: Discrimination performance

Using a phenomenological task, results obtained in experiments 4a and 4b already let us derive quantitative information about the texture interpolation process. The pattern of results also allowed us to relate the observed phenomena to the concept of sensory persistence. With experiments 5 and 6, we now intend to show that texture spreading is a functionally relevant effect that involves more complex processing and provides more information than contour interpolation. To this end, we used highly fragmented

5.6 Experiment 5: Discrimination performance

stimuli in performance tasks whose successful completion required not only boundary information but—ideally—a complete surface representation.

The task was designed to also counter some possible objections against our interpretation of experiments 4a and 4b. After all, we could not ascertain directly that subjects really did have the experience of clearly accentuated texture spreading, and did use this visual impression as a criterion for the threshold settings. They may also have used mere shape information, and alternatively, they may have experienced surface completion, but only in a diffuse homogeneous form. Experiments 5 and 6 address these concerns and will make plausible retrospectively that it was texture spreading which we measured in experiments 4a and 4b.

In experiment 5, we used a two-alternative forced-choice discrimination task to objectively measure the degree to which texture spreading can help to successfully recover the texture composition of a fragmented surface. The stimuli were designed not to contain local features that would be diagnostic for their discrimination. Instead, we wanted to be able to infer from a successful task completion the full perceptual presence of the surface layout. We therefore tried to construct stimuli whose identification rests on features that become apparent only in full view. This, however, proved to be no trivial goal. In particular, two commonly used stimuli, meaningless line drawings and letters, turned out to be inappropriate. The former lacks any surface qualities at all, and the latter contains—by the very virtue of its construction—many local cues which aid in its discrimination. To come up with a suitable stimulus class and an adequate experimental procedure, several preliminary experiments and numerical simulations were carried out—these are described extensively in appendices A.2 and A.1.

Subjects. Five students from the University of Kiel (two female, three male) participated in experiment 5, all were naïve as to the purpose of the experiment. We carried out a separate control experiment with five subjects, among them one (JJ) who also participated in the experimental condition, and myself (DW). All subjects had normal color vision and normal or corrected-to-normal visual acuity.

Apparatus and Stimuli. The same technical setup was used as in experiments 4a and 4b, with the exception of the monitor’s vertical refresh rate, which was set to 100 Hz.

The background of the display was grey (CIE $x = 0.30$, $y = 0.31$, $L = 38 \text{ cd} \cdot \text{m}^{-2}$) and subtended a visual angle of 17.4 by 13.3 deg. An array of vertically-oriented black lines (CIE $x = 0.30$, $y = 0.31$, $L < 0.2 \text{ cd} \cdot \text{m}^{-2}$), 6.75 min of arc wide and 10 deg long, was placed on the background.

The target stimulus was placed in the interior of a virtual disk with a radius of 2.9 deg. Throughout the experiment, only thin stripes (1.27 min of arc wide) of the target stimulus became visible as flanks alongside the black vertical lines.

The target stimuli were non-meaningful images (figure 5.9) which were constructed in the following way: First, the disk was completely covered with randomly placed circles whose varying radii were chosen from a Gaussian distribution with a mean of 13.5 min of arc and a standard deviation of 2.70 min of arc. Next, we defined three invisible, non-intersecting Bézier curves with a thickness of 0.25 deg which started and ended at the disk’s boundary. As a final step, we colored those circles whose center lay outside the curves green (CIE $u = 0.16$, $v = 0.47$, $L = 12$) and then those whose center fell within one of the curves red (CIE $u = 0.24$, $v = 0.45$, $L = 12$). Gaussian red-green chromaticity noise as well as Gaussian luminance noise was added to each circle’s color to make the curves’ boundaries less detectable by simple local edge operators (see appendix A.2 for a more detailed account of the stimulus choice and construction).

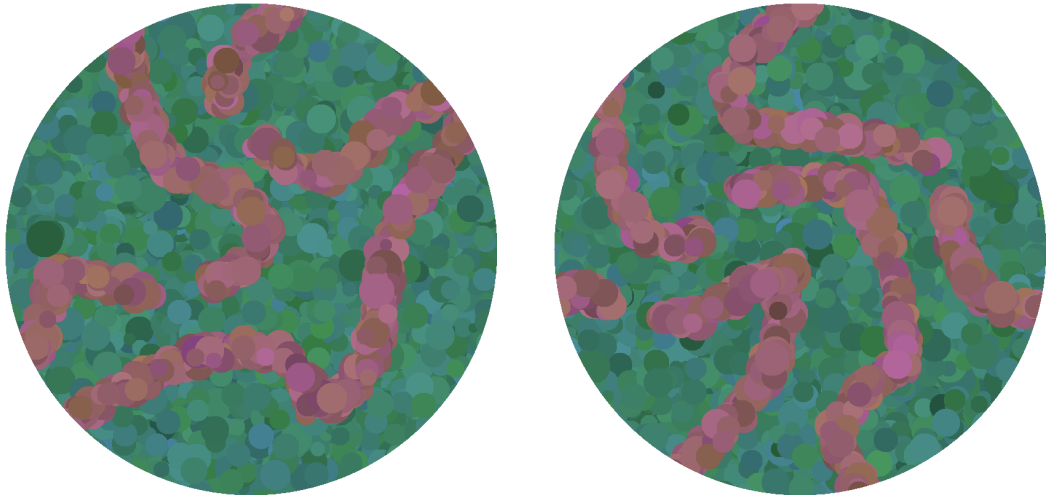


Figure 5.9: Examples of the two stimulus classes used in experiment 5. Circles are colored red if their center falls within one of three imaginary Bézier curves, else green. The target stimulus either contains a continuous line, or all lines are broken.

We designed five individual stimuli each for two different stimulus classes. These classes differed with respect to the presence of a Bézier curve which continued uninterrupted between the disk’s periphery. As will be described below, it was the subjects’ task to discriminate between these two classes, i.e., to decide whether a continuous curve was present or not. In order to reduce the chance of recognizing a particular exemplar of a class, the target stimulus was randomly rotated on each trial.

In the motion conditions, we used four velocities and two types of motion: In the first motion condition (subjects JJ, DWI), the line array rigidly moved to the right at one of the four speeds, whereas the virtual target stimulus remained stationary at the center of the screen. In the second one (subjects TG, MH, KR), the line array was stationary, and the virtual target stimulus moved horizontally back and forth at one of

5.6 Experiment 5: Discrimination performance

the four speeds. Subjects in the second condition were instructed to track the target which made the condition theoretically equivalent to the first one with respect to the resulting retinal image (see section 5.2). In the control experiment, neither lines nor the virtual disk moved, providing a static reference case which indicated the achievable performance without temporal integration.

The four velocities were 5.32, 2.66, 1.77, and 1.33 $\text{deg} \cdot \text{s}^{-1}$. The different speeds were achieved by changing the interval at which the screen was vertically synchronized, which is why the highest velocity is exactly twice as fast as the next slower one, three times as fast as the third, and four times as fast as the slowest. This approach was chosen to guarantee that in all four conditions, the motion path of the to-be-identified object would be equally specified by spatiotemporal fragments. Had we instead increased the length of the discrete jumps of the object between frames, less of the actual motion path of the object would have been displayed, thus reducing the available information. Changing the vertical refresh rate has the same effect as the introduction of timed delays between drawing routines—as was done in experiment 1, but is more precise.

Procedure. In addition to manipulating the velocity, we varied the spacing of the vertical lines from 0.25 to 1.1 deg, corresponding to line densities of 80 to 16 lines visible. The dependent variable was the correctness of the subject’s answer to the question whether the target stimulus contained a continuous line or not. Each target stimulus class had an equal probability to be chosen on a given trial, as was the case for all five exemplars within one such class.

The spacing of the vertical lines was varied according to a custom adaptive method. We used it in order to minimize the number of trials necessary to reliably reconstruct the function relating independent and dependent variable for a given subject. The combination of the independent variables speed and line spacing in a particular trial was chosen in a pseudo-random way. Details of the employed algorithm are given in appendix A.1. Unlike stock adaptive methods like, e.g., QUEST (Watson and Pelli, 1983), we did not assume any characteristics of the function other than monotonicity, and wanted to reconstruct the function over the whole range of the independent variable, as opposed to only estimating the parameters of a fixed function. In addition, the usefulness of, e.g., QUEST for saving trials depends on the accuracy of the initial threshold estimate (figure A.8 on page 196)—a value which we thought may vary considerably between subjects.

Each animation lasted for 8.4 s, after which subjects pressed one of two cursor keys to indicate their classification of the target stimulus. This keystroke automatically started the next trial. Subjects were encouraged to make pauses at their own disposal. Prior to the experiments, all subjects familiarized themselves with the task in 20 training trials which were not included in the analysis. The experiment was split up in sessions of 150 trials, roughly corresponding to a duration of 75 minutes, and it took subjects from 1192 (JJ) to 1625 (MH) total trials to complete the described procedure. Broken

down to the separate velocities, between 220 (TG) and 440 trials were required to estimate the performance curve. The average number of trials thus was a bit lower than what was expected from the simulations (cf. section A.1), which led to an estimate of about 450–500 trials (figure A.6), but was clearly in the same order of magnitude. This provides some supporting evidence that the simulations used to devise the experimental procedure were generally adequate.

Results. Figure 5.10a displays individual data from all subjects in all motion conditions, and for subject JJ also the results from the static control (see also table B.6). The two motion conditions are theoretically equivalent with respect to the retinal image and—as expected—did not produce systematically different results. Therefore, no difference will henceforth be made between these two conditions. The data of individual subjects were also in good general agreement, which warranted their pooling in each of the four velocity conditions and the static control, respectively. Figure 5.10 thus also shows the averaged data. For plotting the data, we divided the full parameter range of the independent variable into intervals. The data points were located at the interval’s mid-point and represent the mean percent correct of the subject’s classification in these intervals. They are guaranteed to be based on at least 40 trials (cf. appendix A).

For all subjects, the discrimination performance was inversely linked to line distance in the motion as well as in the control condition. It continuously varied from near perfect discrimination at low distances to about chance performance at high distances. The performance degradation seems to roughly follow the shape of a logistic psychometric function, as is especially apparent in the averaged data. In the static control condition, the drop in performance with wider line spacing happened faster than in the motion conditions, the difference becoming more pronounced with line spacings wider than 0.56 deg. At a line distance of, e.g., 0.67 deg, performance of JJ in the $1.33 \text{ deg}\cdot\text{s}^{-1}$ motion condition was still at nearly 90% correct, whereas in the static control condition, it had already dropped to about 65%. For the averaged data, mean performance in the static condition at a line distance of, e.g., 0.72 deg was 61% correct, while in the $1.33 \text{ deg}\cdot\text{s}^{-1}$ motion condition it was still 79%.

The pattern of results regarding the influence of velocity is not that clear cut. Surprisingly, discrimination performance did not seem to benefit from an increase in the magnitude of the object’s speed. If anything, rather the reverse seems to be the case. In all subjects except for JJ, performance for the fastest speed dropped before performance in the other velocities. In the averaged data, the static control condition was nearly on par with the $5.32 \text{ deg}\cdot\text{s}^{-1}$ condition, except for very wide line spacings. In contrast, the slowest speed was for most subjects the one where discrimination performance could be held up best with increasing line distance. Only for subject JJ did all motion conditions yield relatively equal performance which was better than that in the static control throughout the parameter range of line spacings.

To abstract from the noisy empirical data, the averaged results from the $1.33 \text{ deg}\cdot\text{s}^{-1}$

5.6 Experiment 5: Discrimination performance

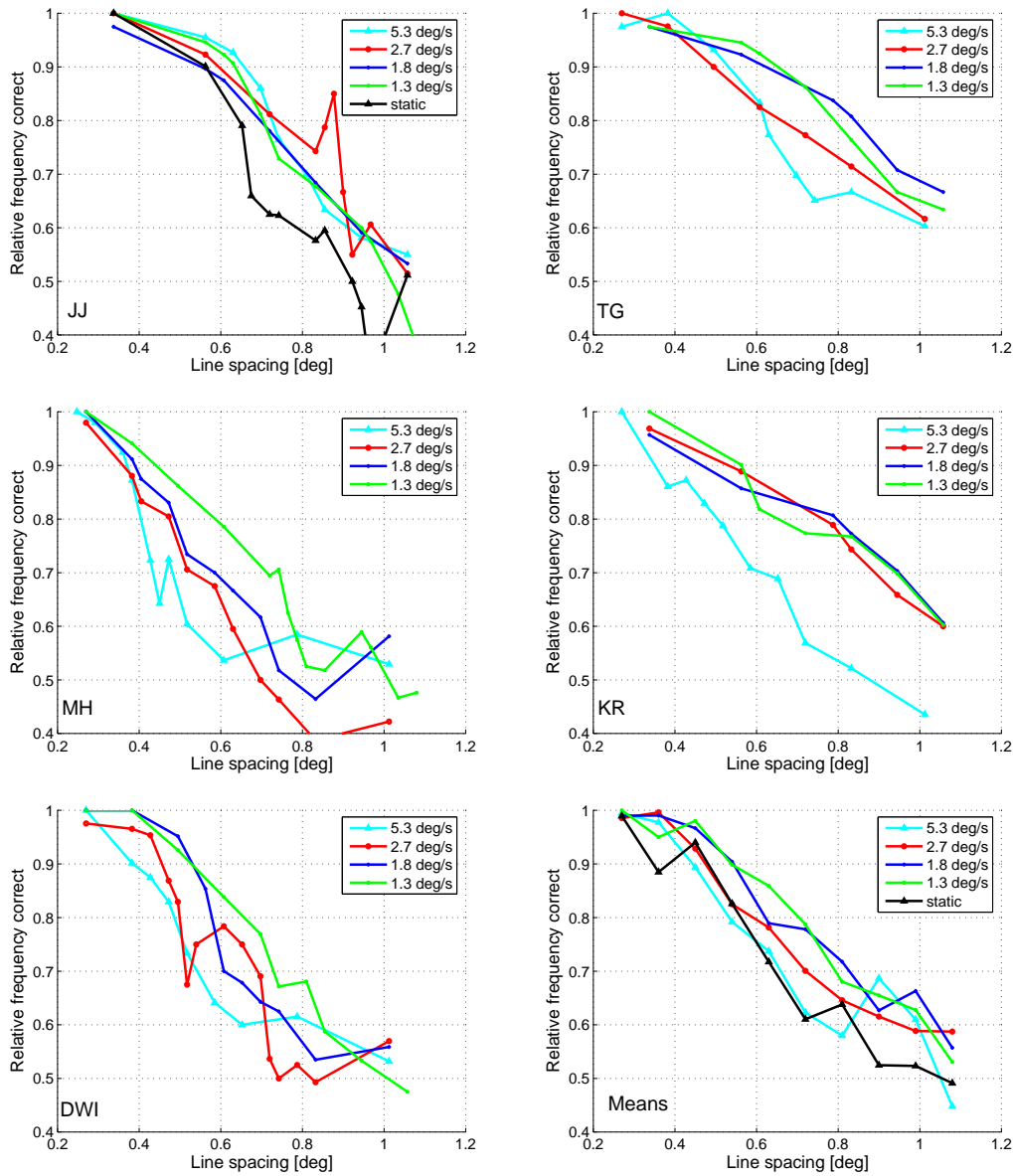


Figure 5.10: Results from experiment 5. Individual data from all subjects in the motion conditions, for subject JJ also in the static control. The figure also shows the combined results averaged over subjects for each velocity as well as in the control condition.

motion condition, and the static control were each visually fitted with logistic psychometric functions. In the motion condition, it was $f(x) = 0.5/(1 + \exp(-6 + 8 \cdot x)) + 0.5$, and in the static condition $f(x) = 0.5/(1 + \exp(-6 + 10 \cdot x)) + 0.5$. The relative perfor-

mance benefit of the motion condition over the static control is plotted in figure 5.11 for the averaged empirical data as well as for the fitted psychometric functions. Mean discrimination was up to 30% better with a speed of $1.33 \text{ deg}\cdot\text{s}^{-1}$ than in the static control.

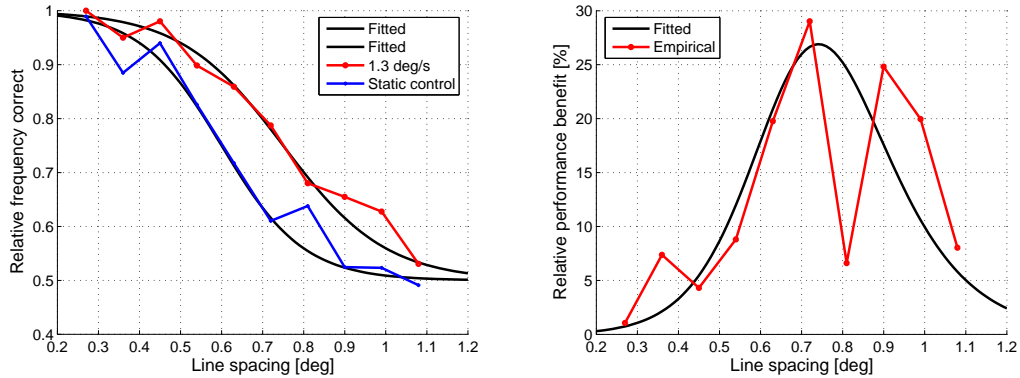


Figure 5.11: Logistic psychometric functions visually fitted to the $1.33 \text{ deg}\cdot\text{s}^{-1}$ motion condition and the static control. The mean relative performance benefit of the motion over the static control condition was computed for the averaged empirical data, as well as for the fitted logistic psychometric functions.

Discussion. The data show the computational relevance of dynamic surface completion: Discrimination performance for fragmented texture patterns can objectively benefit from spatiotemporal integration. The performance difference to a static control condition may be seen as a measure for the functionally relevant amount of surface feature information that is gained through integration. Note that this result allows for no conclusion how exactly the integration was achieved, e.g., whether it was through retinal painting or higher-order motion analysis—it is compatible with both mechanisms.

The results also imply that the integration of a complex texture does not simply lead to its global spatiotemporal averaging or blending into a uniform color, thus setting dynamic texture spreading apart from homogeneous color spreading. Nevertheless, integration does occur, but only *within* the texture pattern. This result supports the relevance of the conflict noted by Dixon and Di Lollo (1994): The visual system must selectively integrate corresponding fragments in the input stream, but it must at the same time also segregate those that do not belong to each other (cf. section 5.2). Integration can therefore not be unconstrained, but must operate on a pre-defined domain, i.e., only within the object boundaries. Finding out exactly how these boundaries are established, and how the regions to be filled in are determined by the visual system remains as the deeper, probably crucial problem.

5.6 Experiment 5: Discrimination performance

Additional informal observations with a slightly different stimulus also point in the same direction. The presented display was constructed as follows: In a color-from-motion stimulus (cf. section 4.2), not one but two imaginary disks of different color were defined, moving along the same horizontal path, but in different directions. When the two circles overlapped, the dots in the area of intersection were colored such that one random half of the dots was assigned to the first disk, and the other half to the second one. Thus, if the first disk was green, one half of the dots in the intersection were also colored green, and the other half, say, red. The percept induced by this setup was not one of two transparent layers that, when sliding over each other, produced a darker virtual layer filled-in with a mixed color—as in the double-neon demonstration by van Lier (2002; cf. section 5.5). Instead, color spreading in the area of the disks was only perceived when the disks were separated. When they intersected, the perception of a homogeneous layers was instantly abolished, and single colored dots were seen. This shows again that there is no simple averaging in the form of integrating everything within a certain spatiotemporal frame. Rather, constraints seem to be imposed on the integration process to work only within regions belonging to each other.

In experiment 5, we noted that the amount of motion did not per se have a beneficial effect on texture spreading, that is to say, this was the case at least with respect to the functionally relevant consequences. Even if this result does not squarely contradict the results from experiments 1 with flank transparency, and from experiment 4 with dynamic texture spreading, it may still be said to be somewhat at odds with them. In a pilot study with the same stimuli, we had already gotten the same result of motion failing to augment the discrimination performance. Said experiment strongly resembled experiment 5 with a moving disk, but with line spacing being manipulated in four fixed steps. Line spacing was fully crossed with object speed which also varied in four steps. Each of the resulting 16 condition was tested 20 times. Results from three of the four subjects in this preliminary experiment are shown together with the averaged data in figure 5.12.

Being the result of using only four line spacings, the interpolated data curves are obviously a lot coarser than the ones resulting from experiment 5. The inability of finer reconstruction within an acceptable number of trials was exactly one of the main reasons for developing the adaptive algorithm. The data show no performance advantage for higher speeds, and thus may be considered consistent with the data yielded in experiment 5. Also note that the data resulting from the highest speed of $5.32 \text{ deg} \cdot \text{s}^{-1}$ represents the worst performance in the averaged results, and is somewhat separated from the other curves who group together.

The pattern of results just described is reminiscent of the one obtained in experiment 5 (figure 5.10), where mean performance at the highest speed was also the worst, being almost identical to the one in the static control condition. In experiment 5, this result was most pronounced for subjects TG, MH, and KR—those who had been issued a display in which the target disk moved, and not the lines. As noted above, this was also the type of motion setting used in the pilot study. As a possible cause for the results,

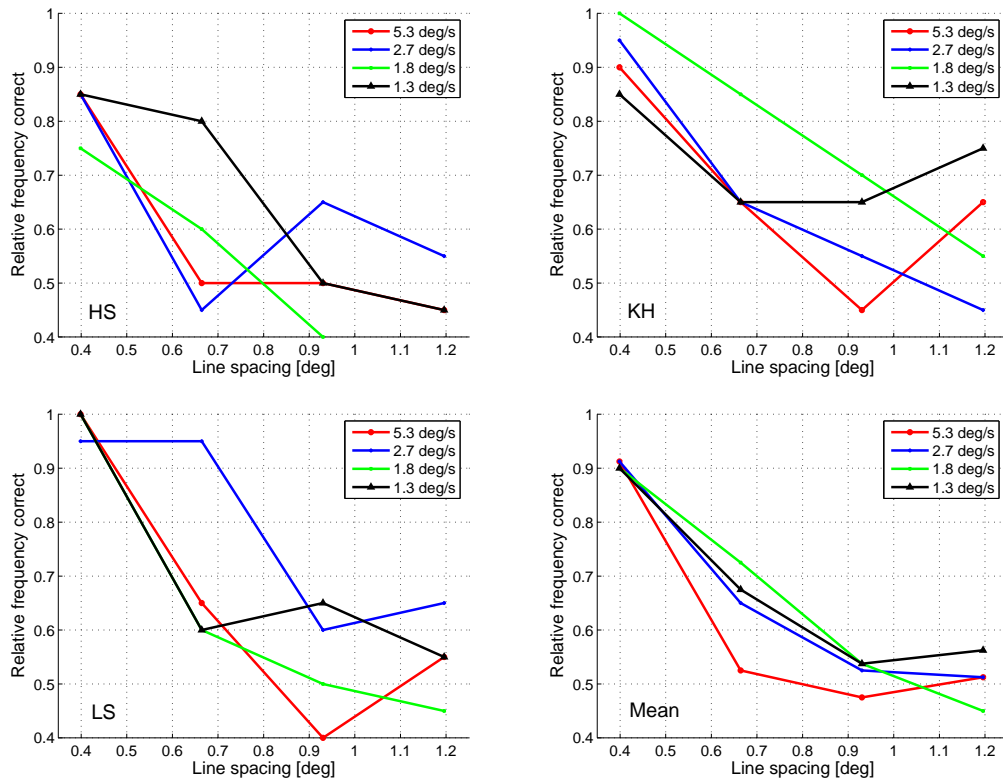


Figure 5.12: Results from a pilot study which strongly resembled experiment 5, but had line spacing manipulated in four fixed steps. Individual data from three of the four subjects together with the averaged results are shown.

we may thus speculate that the observers were incapable to smoothly track the moving target disk at a speed of $5.32 \text{ deg} \cdot \text{s}^{-1}$, and thus unable to exploit persistence. Consistent with the conclusions drawn from results in experiment 4, this interpretation would again highlight the importance of eye movements for spatiotemporal integration, possibly mediated by sensory persistence. Under ideal conditions however, human observers can smoothly track objects that move faster than $100 \text{ deg} \cdot \text{s}^{-1}$ (Ilg, 1997). While this general capability does not automatically imply that subjects could reliably execute smooth pursuit eye movements with our viewing conditions and stimuli, a possible role of smooth eye movements in our result can also not be presumed without further empirical evidence. Other factors which might have caused a detrimental effect of high speeds are motion smear and aliasing effects, i.e., the introduction of spurious spatial frequency information as a result of a spatiotemporally discrete sampling process. Again, these considerations are speculative without further evidence from empirical inquiries. Furthermore, they should not overshadow the situation that while experiment 5 soundly established a measurable motion benefit for shape recognition, there remains

5.7 Experiment 6: Gabor-frequency matching

a certain tension between our results from experiment 5 and those from experiments 1 and 4.

Taking together, on the one hand, the results from experiments 1 and 4, and, on the other hand, from the pilot study and experiments 5, we might infer that there could be a dissociation between phenomenological aspects of surface-feature spreading and functionally relevant ones: Whereas motion clearly leads to a subjective enhancement of the surface spreading in a continuous fashion (experiments 1 and 4), no equally continuous functional benefit seems to be associated with higher speeds regarding discrimination performance. Thus surface integration for functional purposes seems to be qualitative in nature: Either it is possible to exert spatiotemporal integration to obtain relevant surface structure—regardless of speed, or it is not possible due to the absence of motion in the display or the inability to smoothly follow the target with the eyes.

The stimuli were constructed such that the necessary information for a successful discrimination could not come from simple contour integration alone, as was the case with most discrimination tasks reported in the literature which used line drawings or letters. We were thus able to show that spatiotemporal integration of a surface feature with non-local properties, such as a complex texture, can indeed have computational relevance in a given task.

Our goal was to design a task that, to be successfully accomplished, requires a *complete* surface representation available to the visual system. However, as is evident from the static control condition, information relevant to the discrimination task was, up to a certain degree, also available with incomplete stimuli. One explanation for this finding is that with a line distance smaller than the gaps between Bézier curves, detection of gaps becomes trivial. Gap information is then available in spatial frequencies low enough to be fully recovered with the available spatial sampling frequency without the need for temporal integration. Another explanation may be required to account for the relatively smooth performance drop with increasing line spacing in the static control. It is possible that diagnostic local features were still present in the target stimuli to a certain extent. This illustrates the difficulty in designing appropriate stimuli for the objective assessment of the degree to which a complete object representation has formed. Nevertheless, the fact that spatiotemporally integrating the stimulus leads to a measurable performance gain once line distance becomes larger than curve gaps, means that non-local surface feature information were relevant and could be recovered in the integration process.

5.7 Experiment 6: Gabor-frequency matching

In experiment 6, a matching procedure is used to obtain another objective measure of the subjects' ability to perceptually integrate a fragmented stimulus. To this end, subjects had to match the frequency of a Gabor patch to that of a vertically oriented test Gabor which was shown in a multi-slit view setup with visible slits. A Gabor patch

is well suited to be used in this task as it constitutes a textured surface without clear boundaries, thereby reducing the diagnostic value of information coming from simple boundary integration mechanisms.

Subjects. Five male subjects participated in experiment 6, among them three students from the University of Kiel who were naïve as to the purpose of the experiment, one collaborating scientist (FF), and myself (DW). Subjects had normal color vision and normal or corrected-to-normal visual acuity.

Apparatus and Stimuli. The same technical setup was used as in experiment 5.

The stimuli were constructed such that in the top half of the display, subjects were presented a multi-slit view setup which made partially visible a Gabor patch within the region covered by the slits. The background color was black, the slits were 1.35 min of arc wide, 5.7 deg long and grey ($CIE\ x = 0.30, y = 0.31, L = 12\text{ cd} \cdot \text{m}^{-2}$). The vertically oriented Gabor patch (figure 5.13) of radius 2.9 deg was horizontally centered on the screen. It was nominally equiluminant and varied in its chromaticity according to the Gaussian-windowed cosine-wave on the axis defined by the colors green ($CIE\ u = 0.16, v = 0.47, L = 12\text{ cd} \cdot \text{m}^{-2}$) and red ($CIE\ u = 0.24, v = 0.45, L = 12\text{ cd} \cdot \text{m}^{-2}$).



Figure 5.13: Example Gabor patch used as a target stimulus in experiment 6. Please note that the colors may be incorrectly reproduced in this print.

The slit array varied in its density. It rigidly moved to the right at a speed of $4.5\text{ deg} \cdot \text{s}^{-1}$, whereas the virtual Gabor patch remained stationary.

5.7 Experiment 6: Gabor-frequency matching

In the bottom half of the display, subjects were presented a fully visible Gabor patch of the same size as the test Gabor. Its chromaticity pattern was identical to that of the test Gabor, but its luminance was reduced to $8 \text{ cd} \cdot \text{m}^{-2}$. Its luminance was also windowed with a two-dimensional Gaussian to make it blend in with the black background. The match Gabor was adjustable with respect to the frequency of the underlying cosine-wave.

The vertical orientation of the Gabor together with the narrow slit width ensured that, in a given frame, no frequency information would be available within each slit. As such, the highest frequency that could be reliably recovered from a static version of the display is given by the Nyquist limit: It is equal to one half of the slit frequency.

Procedure. The frequency of the test Gabor was set to 0.69, 1.05, 1.39, 1.74, 2.09, 2.79, 3.48 or 4.18 full cycles $\cdot \text{deg}^{-1}$, corresponding to 4, 6, 8, 10, 12, 16, 20 and 24 amplitudes in the area of the Gabor. The slit frequency was one of 2.44, 3.14, 3.83 or 4.53 slits $\cdot \text{deg}^{-1}$, corresponding to 14, 18, 22 and 26 slits in the area of the Gabor. The two independent variables were combined in a fully crossed 8×4 design.

Subjects adjusted the frequency of the match Gabor with the keyboard's cursor keys. Ten repetitions for each setting were carried out, resulting in a total of $8 \times 4 \times 10 = 320$ trials. If a setting could not be made because the Gabor was not seen and its frequency therefore indeterminable, subjects could flag this trial as invalid using the keyboard's space bar.

A separate control condition with a fully visible test Gabor patch was used to obtain individual assessments of the adjustment method's inherent inaccuracy. All aforementioned frequencies had to be matched by each subject in the control condition, resulting in $8 \times 10 = 80$ additional trials. The results provided a reference case for the adjustments in the experimental conditions.

Results. Trials that were flagged as invalid, a total of 193, equaling 12% of the trials, were excluded from the following analysis. Since the subjects did not exhibit systematic differences in their frequency settings, their data were pooled. The resulting data are shown in figure 5.14 (see table B.16 for individual data from all subjects). The graph depicts the mean relative deviations of subjects' settings from the respective frequency of the test Gabor. The relative deviation was measured in percentage of the test frequency. The graph also contains the respective Nyquist limit for each slit frequency. In addition to the experimental conditions, the graph provides the range of relative deviations in the control condition together with their standard errors.

The settings were in general rather precise, the mean relative deviation of the subjects' settings typically ranged from only 10 to 15% of the test frequency. Most importantly however, mean deviations in the experimental conditions were about as large as, or only slightly larger than in the control condition—with a clear exception of the

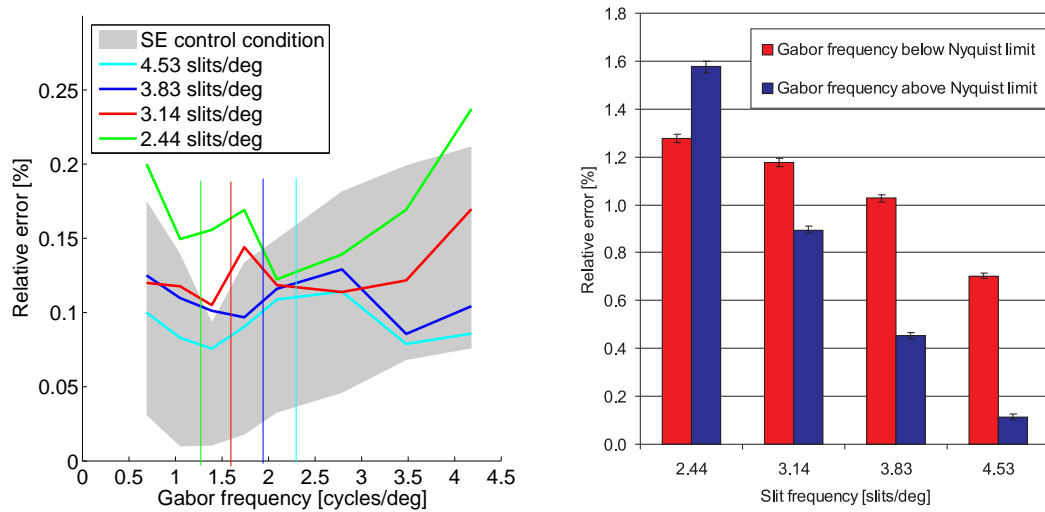


Figure 5.14: Results from experiment 6. a) Mean relative deviation of the subject's settings is plotted against the frequency of the test Gabor. Each line represents the settings for one slit frequency. The grey background area shows the range of the subject's settings in the control condition, i.e., the mean relative deviation ± 1 SE. Vertical lines indicate the respective Nyquist limits induced by the corresponding slit frequencies. b) Mean relative deviation of the frequency settings for all tested slit frequencies below and above their respective Nyquist limit. Error bars indicate ± 1 SEM.

lowest line density. A general trend for all four slit frequencies was that mean relative deviation was higher for the lowest and highest Gabor frequencies. This also means that relative deviations did not simply increase with Gabor frequency in a monotonic way. With regard to slit frequency, it is apparent that lower slit frequencies were associated with higher mean relative deviations of Gabor frequencies.

Figure 5.14b depicts the data in a way that makes explicit the relationship between mean relative deviation in cases where the Gabor frequency was below the Nyquist limit imposed by the slit frequency, compared to cases where the slit frequency was above that limit. Contrary to intuition, mean relative deviation was not higher beyond the Nyquist limit, in fact it was lower, with a decreasing difference between them towards the lowest slit frequency. Only here subjects showed worse performance beyond the Nyquist limit. The figure also shows more directly that adjustment inaccuracies generally decreased with higher slit frequencies.

Discussion. Results from experiment 6 again show that spatiotemporal integration of a texture does not lead to its blending into a uniform color. Instead, the surface structure is preserved as indicated by the fact that the fragmented Gabor pattern's frequency could be reliably recovered well beyond the Nyquist limit theoretically imposed by the frequency of the slits. In fact, for the two highest slit frequencies, subject's performance was just as good over nearly all tested frequencies as in the reference case with an unobstructed test Gabor. This means that due to the possibility of temporal integration, surface feature perception was almost perfect in these two cases.

Based on these results, one may ask whether the employed task is at all sensitive to the amount of surface information that is available in the display in each frame. This however follows from the pattern of results which indicates that lower slit frequencies indeed led to higher mean inaccuracies. The analysis of the performance differences with respect to test frequencies below and above the Nyquist limit (figure 5.14b) also suggests a similar conclusion: For the lowest slit frequency, the Nyquist limit had the expected effect of separating better from worse performance. For all higher slit frequency, the Nyquist limit lost this theoretical property, hinting at the fact that there is no resolution limit when spatiotemporal integration is possible (Nishida, 2004). Still, it remains elusive as to why the performance above the Nyquist limit was actually better for the three highest slit frequencies.

5.8 Discussion

The results obtained in all four experiments were systematic, reliable, and of sufficient interindividual consistency. This warrants the use of dynamic texture spreading as a tool to investigate visual completion processes with phenomenological (experiments 4a and 4b) as well as with more rigorous quantitative methods (experiments 5 and 6). We believe that this possibility constitutes a significant advantage over static color or texture spreading, which are often perceptually unstable and therefore prove elusive to grasp with quantitative methods.

Experiments 4a and 4b tried to relate the phenomenon under consideration to sensory persistence, a simple perceptual mechanism whose involvement in spatiotemporal integration is frequently discussed (cf. section 5.2). The data from these two experiments were in good agreement with a retinal painting account of dynamic texture spreading. The results of experiments 5 and 6 may also be interpreted as being compatible with retinal painting. The performance drop associated with a progressive reduction of diagnostic stimulus fragments was gradual rather than abrupt. In this respect, the results are similar to those from experiments 1 to 4 which showed that the phenomenal variability of dynamic surface completion can be of a continuous nature given suitable manipulations of spatial and temporal stimulus parameters.

Experiments 5 and 6 served to illustrate that texture spreading allows observers to recover significantly more information about the surface layout of complex and highly

fragmented stimuli than just their boundaries. This was done using performance tasks whose successful completion required that a rich perceptual surface pattern representation was at the observer's disposal. Texture spreading brought with it a measurable performance benefit and thus proved to be a functionally relevant effect.

So far, the reported results do not by themselves show that dynamic surface completion is an interesting effect in its own right which cannot be readily explained as a mere by-product of low-level peripheral transformations of the input. However, qualitative observations indicate that several aspects of the phenomenon are not accounted for by sensory persistence and likely require the assumption of a computational analysis in visual processing. First, since sensory persistence and the corresponding integration mechanism are strictly local, the effect should not be sensitive to global changes in the stimulus layout, as long as these changes do not affect the refresh frequency of texture information. Nevertheless, it can be observed that texture spreading does not occur if the black lines—which do not directly contribute to the sensory persistence of texture cues—are removed from the display. Likewise, a stimulus with randomly opening apertures does not elicit dynamic surface completion, even though the amount of spatiotemporal surface information is the same as in multi-slit displays (cf. section 5.3).

While this influence of figural cues is not readily explained by retinal painting, it can be naturally integrated into a high-level perspective which emphasizes the global stimulus layout and scene organization: The Bregman-Kanizsa display (Bregman, 1981; Kanizsa, 1979; Nakayama et al., 1989) for example illustrates that fragmented letters are immediately recognized only if the occluding material is also visible. Applying this reasoning to dynamic texture spreading, we may interpret the lines as occluders, signaled by the T-junctions they form with the flanks. Their presence seems to constitute a cue to selectively integrate those fragments in the stream of information that are classified as coextensive, and to determine which ones are disjoint and thus need to be segregated (Dixon and Di Lollo, 1994).

Equally important as the problem of dealing with figural influences, is that retinal painting also fails to account for the perception of depth in dynamic-completion displays. Not only do dynamically-integrated surfaces regularly have a depth assignment attached to them, there are also reports of inducing the perception of three-dimensional objects by anorthoscopic perception: Day (1989) and Fujita (1990) showed the projection of a rotating wireframe cube, a typical kinetic-depth-effect stimulus, behind a narrow slit. The animated display had subjects indeed report the perception of a rotating 3-D object, even though no single static view was enough to do so. As one can easily imagine, spatiotemporal integration of the moving edges through retinal painting does not at all lead to a combined pattern that would be useful for perceiving the correct 3-D shape (Fujita, 1990). Of course, after discussing the methodological problems with introspective report tasks in section cf. 5.5, the aforementioned results require cautious interpretation. It remains unclear to what degree the reported impression of

a rotating wireframe was a truly perceptual one: An alternative would be that it was merely reminded of by the display.

As a third issue, it is unobvious how to apply the temporal integration scheme to static neon-spreading displays. In this case, it is not plausible to assume that the gaps between corresponding stimulus elements are temporally joined with eye movements which recurrently spread the complete stimulus onto the retina. However, it should be noted that in general, the color and texture spreading is a lot weaker in static displays compared to their dynamic counterparts. Finally, one of the main objections against a major involvement of sensory persistence was that anorthoscopic contour perception can still occur when experimental conditions preclude the possibility that a complete image is temporally spread over the retina. This argues against a critical role of retinal painting in boundary integration. However, apart from the possibility that a retinotopic and a distal representation may co-exist (Shipley and Cunningham, 2001), it does not touch the question whether other perceptual attributes, such as surface filling-in, trace back to sensory persistence. Taken together, this suggests that, unlike certain spatial and figural stimulus aspects, sensory persistence by itself may not be a *necessary* prerequisite for all spreading effects. Nevertheless, when stimulus conditions favor its occurrence, it seems to be able to provide performance-relevant information about surface attributes and to determine crucial characteristics of the percept. In particular, it may be responsible for the particular vividness of the spreading effect in dynamic stimuli.

Chapter 6

General Discussion

6.1 Perceiving completed surfaces, objects, and events

Illusory percepts. At the core of this treatise is the examination of spatiotemporally-completed object percepts. Hence, it might be rightfully expected to first learn what a *completed* feature is, i.e., what distinguishes it from plain regular ones. Throughout the previous chapters, I have indeed often contrasted “completed” features of a percept with “fully-specified”, “stimulus-based” or even “real” ones. This parlance was adopted for reasons of convenience, to match the habitual language use and free my account of cumbersome terminology as much as possible. However, it goes without saying that the fundamentally constructive nature of visual perception does not, in a strict sense, allow for drawing a clear line between percepts that merely are veridical reflections of the physical environment, and ones that are illusory (Mausfeld, 2002). What we then mean when we designate a particular percept as being due to visual completion is that the physical basis of a specific perceptual outcome is *obviously* hard to pinpoint. This is in contrast to simpler percepts like that of a luminance edge, which seem to be more easily traceable to the physical properties of the stimulus. Yet, there is a twofold problem with this distinction: So far, nobody knows how even normal vision works. When we consider all kinds of naturally occurring percepts, e.g., those of objects, biologically relevant categories, or psychological states, the rules that govern the mapping or the triggering relation between physics and visual perception are not yet established. Furthermore, all interesting visual representations are grossly underdetermined by the physical input, such that speaking of fully specified percepts is a problem by itself. Making a serious ontological distinction between real and subjectively completed percepts thus seems unwarranted at this time, if not in principle.

Furthermore, when trying to understand the mechanisms that allow us to dynamically form complete object percepts despite the many obstacles that lie in the visual system’s path, we inevitably create the false impression that these phenomena are more complex or in greater need of explanation than conventional vision. As the above discussion has shown, studying visual completion is in no ways different from studying regular vision since no sensible distinction can be made between these two. Visual perception *per se* can be regarded as a fundamentally inferential process which is guided

by biologically-rooted, but nonetheless idiosyncratic rules of the species whose survival it serves (Hoffman, 1998; Mausfeld, 2002). What sets phenomena of visual completion apart from regular vision is that the former are blatantly obvious cases of inference in the absence of a good data basis. To appreciate that the very same problems apply to all of vision just requires us to ponder over the basic constraints of visual perception a little longer.

The latter conclusion should not come as a surprise since it constitutes one fundamental justification to use so-called visual illusions to study the general characteristics of normal vision. Nevertheless, in this context, it seems unwieldy to adopt a new terminology and dispense with the distinction between completed and regular percepts as the latter allows us to easily connect to the established way of reasoning about the topic.

Irrespective of these conceptual reflections, we have also seen much empirical evidence which argues for a functional equivalence between fully specified and completed objects on a computational level of visual processing. This indicates that visual completion is not an arbitrary epiphenomenon completely detached from normal visual processing, but is tightly integrated into the inner workings of perception. We have seen that perceptual mechanisms may well operate upon input that was generated from completion. Regarding the architecture of the visual system, this may be interpreted to mean that once the barrier to the formation of a specific kind of perceptual representation has been crossed, no further distinctions will be made with respect to the origin of said representation. In addition, these representations seem to carry with them a fixed set of parameters like depth, motion, and color in the case of surfaces, which will be set to a fixed value in a coordinated process. Interestingly, amodal completion seems to make a strong case for the assumption that being phenomenally visible is not among the compulsory settings for a visual surface. Instead, the reviewed evidence argues that perceptual entities can be genuinely visual representations without being visible at all.

We shall, however, not keep quiet about the points put forth by some authors arguing that functional differences between regular and completed percepts do exist. This issue concerns the characteristics of mental imagery and visualization, and their relationship to visual processing (see, e.g., Kosslyn, 2003; Pylyshyn, 2003a; 2003b). The ongoing debate reminds us that all claims pertaining to the functional equivalence between illusory and normal percepts need to rest on empirical evidence, and each case needs to be established separately.

Multiple representations. As an important qualification to the account just given, it needs to be said that in no ways do we mean to imply that real and illusory percepts are always indistinguishable from each other. When looking at stimuli depicting some visual illusions, we are often well aware of the curious nature of the elicited percept. On the one hand, we can perceive contours and colors in between the physically specified

image fragments, but on the other hand, we know distinctly that these stand on less stable empirical grounds than the remainder of our percept. Due to the cognitive impenetrability of most aspects of vision, we cannot, of course, force us not to see the features that we regard as illusory.

Similarly, when we appreciate paintings, we can have simultaneous visual representations of the depicted image space with its depth structure, but at the same time *see* that the picture is flat in the reference frame we rest in. The phenomena outlined above may be taken to show that perceptual entities can figure within multiple reference frames, and can have attributes attached to them that carry information about their credibility. This credibility can be gradual in nature: Perceptual scene elements may be said to carry “meta-data” which specifies the degree of uncertainty associated with their representation. In many cases of perceptual completion, we find that despite the strong conviction that some object is completed behind an occluder, a precise shape of the unified object beyond that of a bounding box is never generated. This was noted by many authors (Michotte et al. 1964; Tse and Albert, 1998, among others), and we have also discussed some relevant examples here, e.g., for static surface spreading (figure 3.27c on page 58) and volume completion (figure 3.23a on page 53). In these cases, the occluded form stays indeterminate and appears somehow amorphous. While the completed objects are characterized by a sense of unity and cohesiveness in distal terms, they can still be perceived to be incomplete in proximal terms—a distinction which reminds of the multiple reference frames in picture perception.

For some cases of visual illusions, it is evident that they are phenomenally indistinguishable from common percepts. Many of the widely-circulated size and color illusions certainly fall in this category, as does the filling-in of the blind spot. What theoretically sets these cases apart from instances of noticeably completed percepts is still unclear. In any case, it seems that perceptual objects are probably fairly abstract entities—they are allowed to be half-real, may be visual yet invisible, and need not have a definite shape. Naturally, these remarks raise the question what a perceptual object then is. This, however, is a topic that all by itself is equally important as it is broad. We shall therefore only touch it briefly in the following paragraphs.

Objects and events.

“Surprisingly, a significant obstacle in the path of understanding object recognition is that we lack a precise definition of what constitutes an object. Without such a definition, how can we possibly know where we are headed? Furthermore, any computational theory of object recognition becomes impossible, for what is to be computed?”¹

– WHITMAN RICHARDS

¹Richards (1988, p. 17).

“Our view is that higher functions require, as an input, a data format that explicitly represents the scene as a set of surfaces.”²

– KEN NAKAYAMA, ZIJIANG HE AND SHINSUKE SHIMOJO

When wondering about what kind of perceptual features can be subjected to filling-in, we quickly come to a more general question: What are the specific parameters that illusory as well as ordinary visual objects as theoretical entities both may be specified through? It then becomes evident that to advance the understanding of visual filling-in, we need a far better idea about the abstract entities that are to be completed. Otherwise, completion would remain a rather diffuse act, for we would keep talking about completion as a *process*, but would at the same time remain mute about the *structures* that these processes operate upon. In some cases, it may be less obvious that the crucial question is overshadowed of what perceptual entities we tacitly assume in our discourse: For instance, an important aspect of completion concerns the grouping of image fragments, and the assignment of figure and ground to elements in a scene. However, it is seldom made explicit how rich a perceptual entity must be to play the role of figure, e.g., whether it can be any closed region, or whether it must be a surface in depth, a bounded volume, or a full-fledged spatiotemporal object.

If we were able to name a set of perceptual kinds, and list for each one what the free parameters are that must be set to a specific value on the basis of the available input data, we were in a position to pose the questions concerning visual completion in a much more specific way. These could be of the following kind: What is the process that assigns a value to the free parameter x of structure S , and what attributes of what other structures does it take as input? Currently, the study of the data structures of visual perception and their specific roles in different computational mechanisms is still in its early infancy. But what has emerged so far points towards the importance of the concepts of objects and events. It seems that much processing revolves around these categories, not just in perception proper, but also in cognition, attention, and language.

One of the main characteristics of an object is that it spatiotemporally links contingent information fragments from many different channels into one cohesive unit. The same may be said about events, which also constitute the binding of multiple input sources into a single entity. That objects and events play closely together is argued for, e.g., by demonstrations of cross-modal perceptual interactions, such as the number of heard beeps influencing the number of seen flashes (Shams, Kamitani and Shimojo, 2000; 2002). The study of object formation thus is by necessity an inquiry into the binding of dissociated information fragments, quite similar to that of visual completion. Object and event perception require the binding of information across the borders of different perceptual codes, and possibly also across those of different computational systems, such as the ones for cognition, attention, and perception. Exactly what kind of parameters objects and events can have in what processing contexts is as of yet

²Nakayama et al. (1995, p. 6).

not understood. Yet it seems plausible that once they are uncovered a bit more, they could make for an apt guide on how to put order into the wide variety of completion phenomena.

In the study of visual completion, just as in much of vision science in general, the question how to properly structure the domain under consideration is so far answered only by naming intuitively separable aspects, e.g., color, depth, texture, or shape. These are then usually considered in isolation from one another, an approach that is also reflected in the way that standard text-books on vision are structured. For analytic purposes, it is of course inevitable to slice the rich visual phenomenology into smaller pieces that promise to be more easily understood when considered just by themselves. The standard segmentation of visual processing, however, may be misleading as it may not reflect the intrinsic conceptual structure of vision (Mausfeld, 2002; 2003). We have already seen throughout this treatise that the superficial aspects of color, shape, and motion seem to provide only an unsatisfactory structure for the relevant phenomena. Too much intimate interaction between them exists to let them appear as a plausible set of independent dimensions for analysis.

In other words, to overcome the current state of affairs in the study on filling-in, we may ultimately need to be able to consider visual completion within the larger context of perceiving objects and events. From this context, it should ideally be possible to derive more precise research questions, and to have a frame of reference to analyze the bits of empirical evidence in.

However, delving further into the topic of object perception would be well beyond the scope of this work. Despite still lacking a coherent picture, a wealth of evidence has already accumulated on the role of objects in perception and cognition, calling for an account of their details that devotes to them the space and care they deserve.

6.2 Summary

In the following paragraphs, I shall give a short retrospective summary of the chief theoretical aspects of dynamic visual completion, together with an integrated evaluation of the empirical findings reported both here and in the existing literature. I will do so by tracing the proceedings of this treatise, highlighting the main issues that were addressed along the way.

The journey into perceptual completion started off in chapter 2 by considering sources of spatiotemporal input fragmentation that the visual system is facing in natural environments. Fragmentation can be the result of the physical scene composition and the laws governing the geometry of perspective projection. The signal-transduction characteristics of our sensory organs further add to this effect. Of the many ways in which the reflected wave fronts from an object are altered before reaching the eye, we especially concentrated on occlusion and its interaction with motion. Stated simply,

the interception of light by objects nearer to the observer has as its consequence that boundaries and contiguous areas in the retinal image do not match the distal object boundaries and shapes. Grouping only those fragments stemming from one distal object while maintaining a separation between fragments belonging to different objects thus is a major obstacle that needs to be overcome for object recognition. The regularities following from the geometry of projection dictate that depth perception needs to be inextricably intertwined with grouping: Since only objects closer to an observer can occlude objects that are farther away, fragments should only be linked together if classified as lying *behind* an occluder. Accordingly, it cannot surprise much to encounter the many results which indicate that depth plays a pivotal role in object completion.

Just as the perception of depth is theoretically and phenomenally linked to occlusion, motion, too, plays an important role in the processes of both stimulus generation and perceptual processing with respect to occlusion. The pattern of dynamic occlusion and revelation of background elements as well as the structure of a scene's motion vector field are in fact powerful grouping factors used by the visual system to link corresponding image fragments. How exactly the visual system uses these cues in the parsing of disconnected 2-D image regions into coherent surfaces and volumes later emerged as the main question behind the different instances of dynamic visual completion.

Chapter 3 prepared the ground for approaching the experimental examination of a more concrete research question. In order to provide the necessary theoretical background, it gave a broad review of relevant phenomena and empirical findings, shedding light on visual completion from different angles. Since the heterogeneous category of visual completion subsumes a multitude of different effects, special emphasis was placed on providing a structured account of completion. To do so, it followed the rough taxonomic structure of static vs. dynamic, as well as modal vs. amodal kinds of completion.

For a long time, the main focus in research on static completion has been the perceptual integration of boundary fragments into closed contours. The influential concept of relatability tries to formulate the necessary and sufficient conditions the input image has to meet in order to trigger a completion process between two disconnected edges. It thereby treats amodal and modal completion in an identical fashion, their only difference supposedly being a superficial phenomenological one. Although attempts are underway to generalize the notion of relatability to 3-D space, so far it is a purely geometric condition rooted in the 2-D domain of the stimulus. A fair number of examples of modal and amodal completion were indeed found to be in good accordance with it.

However, instances of boundary completion being triggered by stimuli in violation of the rules of relatability accumulated, and raised doubts concerning its general adequacy in explaining boundary completion. One of the main arguments calling into question the usefulness of relatability comes from cases of volume completion in response to stimuli depicting situations of conformation. Many instances of volume completion are

elicited despite the absence of relatable edges. In these stimuli, cues to occlusion only exist once image entities are interpreted volumetrically. Another set of evidence that is unaccounted for by relatability theory comes from situations in which modal and amodal completion yield very different scene interpretations in terms of the perceptual units involved. Both classes of phenomena suggest that relatability has problems explaining effects that crucially involve the role of depth. This points to the important question whether 2-D image based information, or 3-D volumetric representations enter the visual completion process as input.

When closely examining the rich functional interactions between the representations of color, depth, and motion in visual completion, more arguments turn up that point towards the importance of mid-level representations like surfaces which include aspects of depth. This is because the perceptual codes for a completed object's color, placement in depth, and transparency appear to be strongly coupled in the sense that combinations of them cannot be of arbitrary nature. They rather obey the rules imposed by a coherent surface concept: If a perceptual entity is perceived to lie behind an object, it does not have a modal quality. If, however, the object is perceived to be in front, its surface is modally represented including aspects of color, texture, and transparency. Experimentally toggling the state of one of these features will cause qualitative changes in the others, too. The lawful constraints on their interactions suggest that even though aspects of, e.g., color, shape, and motion are often considered to be modular perceptual primitives, surfaces may be a much more promising candidate for such a concept.

Many examples of boundary and surface completion elicited by dynamic stimuli indicate that motion is used extensively by the visual system to bind together dissociated fragments in the input stream. The results are coherent spatiotemporal chunks corresponding to single objects and events. Amazingly, these grouping processes can be initiated not only by stimuli which are in principle compatible with a physical scene layout, but also by ones which implement spatiotemporal discontinuities only in an abstract way. Still, in contrast to dynamic occlusion, these bi-directional displays only lead to boundary completion and fail to elicit the creation of a distinct surface percept with its associated qualities of color and definite location in depth. As a relatively recent development, the theory of spatiotemporal relatability tries to straightforwardly generalize the concept of static relatability to dynamic stimuli. While this appears to be a sensible approach, it also inherits from of static relatability theory all of the latter's abovementioned limitations, some of them due to their restriction to 2-D image geometry. As such, it has difficulties to account for the general involvement of depth in filled-in percepts, and also for more complex phenomena like the dynamic completion of 3-D objects.

After exploring the breadth of visual-completion phenomena and the many areas of unfinished work in their study, chapters 4 and 5 served to narrow down the topic towards experimentally treatable research questions. In total six experiments on dynamic color and texture spreading were presented that addressed two main issues: To

what extent can low-level processes like local spatiotemporal integration through visual persistence account for the effects? And what is the adequate methodology to approach the dynamic visual completion of surface qualities, given the often peculiar nature of their appearance? In the process of dealing with these questions, we concentrated on the role that motion plays in the completion of uniform and non-uniform surface percepts. To this end, we started by developing a new kind of multi-aperture-viewing stimulus eliciting a particularly strong impression of transparent dynamic color spreading that we termed flank transparency. The first set of three experiments reported in chapter 4 served to, so to speak, put the phenomenon under consideration on stable empirical grounds. As such, the experiments were intended to determine the basic psychophysical characteristics of the stimulus type, and to show its general accessibility to quantitative methods.

In experiment 1, we quantified the effects of increasing speed on subjects' settings in a color-matching task. We found that color spreading increased with higher velocities, indicated by the illusory disk acquiring a darker and more saturated color. This provided good support for the idea that the appearance of color spreading in dynamic displays strongly depends on apparent motion. Thus, experiment 1 convincingly showed that, consistent with the theoretical importance of motion as a cue to fragment grouping, perceptual codes for color and motion are assigned in a tightly coupled manner.

In experiment 2, we manipulated the contiguity between the inducing elements of a flank-transparency display as a figural influence on color spreading. The results indicated that violations of physical contiguity have a detrimental effect on surface completion, but less so than in corresponding static displays. We interpret this finding as being due to the very robust nature of the color-spreading effect in flank transparency. The experiment underlined the importance of figural determinants of color spreading, an aspect which has so far proven difficult to integrate into isolated research perspectives which focus only on chromatic stimulus conditions.

In order to place flank transparency within the context of other color-spreading phenomena, we determined in experiment 3 its spatial extent. We found that the strength of the color spreading quickly degenerates as the distance is increased which the spreading must bridge. At larger inducer spacings, the illusory object's color was nearly identical to that of the background, whereas under optimal conditions, it was of a highly saturated red. While the watercolor effect effectively operates over up to 40 deg, flank transparency resembles more traditional displays of neon spreading which are greatly limited in the spatial extent covered by the spreading color.

The results from all three experiments showed a fair degree of consistency within, as well as across subjects in the task of color matching. This emphasizes the stability of the dynamic effect under consideration and encourages the further use of quantitative psychophysical methods. As a side note, flank transparency turned out to be a challenge to current models of chromatic conditions for perceptual transparency. Perceptual transparency in flank-transparency displays obtains even if these are composed

of only two colors, whereas current models all require the presence of at least three colors to be applicable. For theoretical reasons, this number is not further reducible for purely chromatic models. It thus appears possible that motion as a powerful grouping cue may relax the chromatic requirements for perceptual transparency when compared to static stimuli.

Exploiting the flexibility of the multi-aperture stimulus class, we extended flank transparency to the perceptual completion of non-uniform surface features in chapter 5: Dynamic texture spreading is a novel spatiochromatic interaction that not only comprises the formation of strong subjective contours but also the vivid illusory perception of surface qualities like a colored texture along with depth stratification. Unlike static texture spreading, the effect manifests itself reliably, and is of high distinctiveness in general.

Building upon the empirical results from experiments 1–3 and additional theoretical considerations, we discussed whether the simple mechanisms of sensory persistence are potentially responsible for the observed surface filling-in. Even though often rightfully criticized, sensory persistence has played a pivotal role in explaining many aspects of shape perception in multi-aperture displays. Experiments 4a and 4b were designed to apply sensory persistence to the study of surface-feature integration by testing some straightforward predictions with a method of threshold-adjustments.

As was the case in the first three experiments, results showed good reliability and consistency over subjects. We inferred that the percept of dynamic texture spreading is indeed a stable one—a necessary prerequisite for any further data analysis with regard to possible perceptual mechanisms. In line with earlier evidence gathered from studies of dynamic boundary completion, the idea of simple temporal integration through sensory persistence was able to capture the data from both experiments surprisingly well: Two manipulations which affected the spatiotemporal amount of fragmented surface data in opposite ways acted together in a compensatory manner in determining the strength of the illusory percept. According to our data, the percept was thus mainly governed by the gross amount of information that could be integrated within one space-time window, no matter whether it was achieved by temporal or spatial factors.

That sensory persistence accounts well for our data warranted a further analysis in terms of one of its important parameters: The maximum delay between two input signal refreshes that is tolerated by the visual system while still yielding a complete percept. This delay was not exactly constant over experimental conditions as would be expected from retinal painting, but it stayed in a narrow range from around 100 to 150 ms, indicating the maximum time window for integration processes. This estimate turned out to be consistent with the general timecourse of visual processing, and was remarkably similar to results obtained in a number of other experiments reported in the literature on object completion.

The results from experiments 4a and 4b prompted for a deeper analysis of some temporal aspects of visual processing of normal and completed stimuli. They also spurred

a critical look at experimental methods commonly used in the study of visual completion, together with an evaluation of their underlying assumptions. To experimentally assess mechanisms of filling-in beyond relying on introspective data, it is necessary to turn visual completion into a reliably measurable entity that is not sensitive to observer biases. A common approach is to try to devise a performance task whose successful processing completely rests on the presence of a completed stimulus representation. The performance yielded by a subject in such task is then taken as the degree to which the percept was completed and thus functionally operative in perceptual processing. While in theory this sounds good, actually coming up with an adequate methodology is a different thing altogether.

As with all research questions, having good methodological tools at hand can be of decisive importance for advancing one's understanding of the studied domain. Many of the considered methods currently used in studies on visual completion proved to be inadequate to measure the phenomenon of dynamic texture spreading. For this reason, great care was taken to incorporate some of the lessons learned from examining the several methods used so far. Consequently, we developed a new stimulus class in combination with an objective performance task to isolate aspects of surface completion from local boundary integration: The stimuli were constructed such that the necessary information for a successful discrimination could not come from simple contour integration alone—as was the case with most discrimination tasks reported in the literature which used line drawings or letters. After fine-tuning its final implementation in several pilot studies, we used this task in experiment 5.

The obtained data show that discrimination performance for a fragmented texture pattern objectively benefits from spatiotemporal integration. The performance difference to a static control condition was taken as a measure for the functionally relevant amount of surface feature information that is gained through integration. The results also implied that the integration of a complex texture does not simply lead to its global spatiotemporal averaging or blending into a uniform color, thus setting dynamic texture spreading apart from homogeneous color spreading. Integration does occur, but constraints seem to be imposed on the integration process to work only within regions belonging to each other. In our case, these were the objects depicted by the texture pattern.

Experiment 6 was devised to support the abovementioned conclusions with a different methodology, an objective matching task also using dynamic-texture-spreading stimuli. Results from experiment 6 again showed that spatiotemporal integration of a texture preserves the surface structure of the underlying object as indicated by the fact that the fragmented pattern could be reliably recovered. That temporal integration was indeed used was indicated by a performance well above the limit theoretically imposed by the spatial layout of the stimulus. Compared with a condition using an unobstructed control stimulus, surface feature perception was almost perfect, supposedly due to spatiotemporal integration.

6.3 Conclusions and outlook

As was announced in the introduction, it would be a vast overstatement to claim that a coherent picture of visual completion emerged after engaging in the not so stringent review of relevant theoretical ideas together with the associated old as well as new empirical findings. This being said, several conclusions still seem to be warranted from all things considered so far.

For one, there are aspects of genuine surface completion, which appear to be irreducible to the mere interpolation of boundaries. This means that visual surfaces are defined along many more dimensions than just their two-dimensional projected shape. Surfaces also have a definite position in depth, and can be in a state of coherent motion. This surface motion is able to group, and thus to override the differing local motion signals of fragments which define the surface. Completed surfaces behave, with respect to many processing aspects, functionally equivalent to surfaces which are considered to be more directly based on the physical input to the visual system. Interpolated features are fed as input into the computation required for seemingly early tasks like visual search or the generation of disparity-based depth percepts. Interpolated surfaces are thus tightly integrated into the normal processing flow of visual perception.

This conclusion is supported by results indicating that visual completion can lead to measurable performance benefits in object-recognition tasks. Far from being a functionally irrelevant by-product of visual processing with only phenomenally pleasing qualities, filling-in seems to contribute to the superordinate objective of perception: To establish a stable connection between the biological organism and its physical surrounding in the service of the organism's survival.

With respect to possibly underlying mechanisms which contribute to the aforementioned feat, our empirical results indicate that the simple mechanism of visual persistence can account for some observed data patterns. Despite being justifiably criticized on theoretical as well as on empirical grounds as unfit to explain a number of results on boundary integration, persistence thus remains a candidate to figure in some aspects of surface-feature completion. One of the arguments in favor of this mechanism is a set of observations indicating that completed features exhibit a continuous variability along several perceptual dimensions, calling for a mechanism with an equally gradual output. However, precisely for this reason, persistence fails to cover qualitative perceptual state switches, e.g., in response to figural influences which seem subtle from a perspective of pure spatiotemporal averaging.

An analogous problem exists in the experimental ways in which completion is currently approached: While introspective tasks are indispensable for getting an immediate access to what kind of perception it actually is that is in need of explanation, they leave much to be desired when it comes to reliability and proneness to observer biases. Performance tasks promise to excel in this regard and to allow for objective

assessments of the visual processing under question, but necessarily dislodge themselves from the phenomenon proper. By often reducing the response possibilities of observers for the sake of reliability and objectivity to a highly restricted gamut, they risk overlooking important subtle characteristics of the induced percepts. And while performance tasks are quite apt to capture the continuous aspects of perceptual variability, they equally fail to grasp the discrete qualitative changes that often result from figural manipulations.

In line with these considerations, the new adaptive shape discrimination task proved to be a useful tool in the study dynamic surface completion as it affords the reliable measurement of associated functional effects. The observed performance benefit for spatiotemporal integration itself does not imply a particular underlying mechanism for said integration, e.g., sensory persistence. It is, however, in a limited sense compatible with such a notion, i.e., if one is willing to acknowledge that persistence is only allowed to be operative within regions that have been grouped as belonging to one object *beforehand*. This qualification is necessary because persistence alone cannot provide a mechanism to distinguish between those cases where the sudden onset of a new object requires integration not to take place, and those cases where corresponding fragments indeed need to be subjected to spatiotemporal summation.

Filling-in itself may thus be a relatively simple process, possibly implemented in part by persistence, but understanding how the visual system determines wherein or whereupon a perceptual feature should spread is much harder. This problem results from the fact that persistence itself is necessarily blind to any semantic structure of the scene in terms of objects or events. If active indiscriminately, persistence thus would operate not only within objects but also across object boundaries. This characteristic would blur the demarcation of objects, a highly undesirable consequence from the point of view of object recognition. Integration can therefore not be unconstrained, but must operate on a pre-defined domain, i.e., only within the object boundaries. Finding out exactly how these boundaries are established, and how the regions to be filled in are determined by the visual system remains as the deeper, probably crucial problem. In any case, integration cannot simply operate on the level of the retinal image, but already presupposes a semantic segmentation of the input.

Reaching a deeper understanding of completion thus depends on accruing more knowledge about the perceptual system's conceptual architecture, and in particular about what role visual objects play in it. This takes us back to the question of what may rightfully be regarded a set of perceptual primitives, i.e., along which dimensions the visual phenomenology should be sliced into intrinsically coherent parts. So far, this question is unresolved, although the common strategy of separating aspects like color, shape, motion seems to be increasingly inadequate.

A number of more concrete open questions in the study of dynamic completion also remain: Reconciling the aspects of, on the one hand, gradual parameter variations and simple spatiotemporal averaging processes with, on the other hand, qualitatively

acting figural constraints is one such challenge. Furthermore, not only is a more precise formulation of the chromatic constraints for static and dynamic color spreading missing, there is not even an established formalism for the structured analysis of the important interplay of, say, figural and chromatic aspects. More generally, there is hitherto no theoretical framework which could tie together all the fragmented bits of evidence and integrate the empirical findings into a common picture of the matter. The various partaking research domains all have their own approaches together with an often idiosyncratic terminology and a set of preferred standard models. Crossing these borders seems to be one of the most imminent requirements to further advances in explaining dynamic completion, an undertaking which is still in its infancy.

6.4 Acknowledgements

A treatise of equal proportions to the present one would be in order to express my gratitude to the many who fostered my studies. Here, I can only single out some of those who provided me with outstanding support over the years: I am indebted to my advisors Rainer Mausfeld, Don Hoffman, and Johannes Andres for the encompassing education they bestowed on me, as well as for the time and energy they invested in my training.

I heartily thank Franz Faul and Vebjørn Ekroll for freely sharing their intellectual insights, for helping me out countless times, and for a guidance that stretches well beyond intellectual matters. Heike Jores, Oliver Essing, Fritz Geiger, and Vincent van Houten selflessly let me draw the energy from them that I needed to pursue my work and that allowed me to still keep a semi-sane mind.

This dissertation was supported by a Fulbright fellowship and a doctoral scholarship of the Studienstiftung des deutschen Volkes. I am extremely grateful to both organizations for the financial aid they provided, and even more so for giving me the opportunity to engage in numerous inspiring discussions with scholars of many different fields.

References

- Adelson, E. H. (2000). Lightness perception and lightness illusions. In M. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 339–351). Cambridge, MA: MIT Press.
- Albert, M. K. (2001). Surface perception and the generic view principle. *Trends in Cognitive Sciences*, *5*, 197–203.
- Albert, M. K., & Hoffman, D. D. (1995). Genericity in spatial vision. In R. D. Luce, M. D'Zmura, D. D. Hoffman, G. Iverson, & A. K. Romney (Eds.), *Geometric representations of perceptual phenomena: Papers in honor of Tarow Indow on his 70th birthday* (pp. 95–112). Mahwah, NJ: Lawrence Erlbaum Associates.
- Albert, M. K., & Hoffman, D. D. (2000). The generic-viewpoint assumption and illusory contours. *Perception*, *29*, 303–312.
- Andersen, G. J., & Cortese, J. M. (1989). 2-D contour perception resulting from kinetic occlusion. *Perception & Psychophysics*, *46*(1), 49–55.
- Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception*, *26*(4), 419–452.
- Anderson, B. L., Singh, M., & Fleming, R. W. (2002). The interpolation of object and surface structure. *Cognitive Psychology*, *44*, 148–190.
- Anderson, J. C., Baddeley, R. J., Osorio, D., Shashar, N., Tyler, C. W., Ramachandran, V. S., et al. (2003). Modular organization of adaptive colouration in flounder and cuttlefish revealed by independent component analysis. *Network Computation in Neural Systems*, *14*, 321–333.
- Anstis, S. M. (2005). Local and global segmentation of rotating shapes viewed through multiple slits. *Journal of Vision*, *13*(3), 194–201.
- Anstis, S. M., & Atkinson, J. (1967). Distortions in moving figures viewed through a stationary slit. *American Journal of Psychology*, *80*, 572–585.
- Arteberry, M. E. (2001). Perceptual unit formation in infancy. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Grouping and segmentation in vision* (pp. 37–69). New York, NY: Elsevier Science.
- Bagrow, L., & Skelton, R. A. (1985). *History of cartography*. Chicago, IL: Precedent Publishing.
- Berkley, M. A., De Bruyn, B., & Orban, G. A. (1994). Illusory, motion, and luminance defined contours interact in the human visual system. *Vision Research*, *34*, 209–216.
- Berry, M. J., Brivanlou, I. H., Jordan, T. A., & Meister, M. (1999). Anticipation of moving stimuli by the retina. *Nature*, *398*, 334–338.

References

- Bozzi, P. (1975). Osservazioni su alcuni casi di trasparenza fenomenica realizzabili con figure a tratto. In G. B. Flores d'Arcais (Ed.), *Studies in perception: Festschrift for Fabio Metelli* (pp. 88–110). Milano: Martello-Giunti.
- Bradley, D. R., & Lee, K. (1982). Animated subjective contours. *Perception & Psychophysics*, *32*, 393–395.
- Bradley, D. R., & Petry, H. M. (1977). Organizational determinants of subjective contour: The subjective Necker cube. *American Journal of Psychology*, *90*, 253–262.
- Brainard, D. H. (1995). Colorimetry. In M. Bass (Ed.), *OSA handbook of optics* (pp. 26.1–26.54). New York, NY: McGraw-Hill.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bregman, A. S. (1981). Asking the “what for” question in auditory perception. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 99–118). Hillsdale, NJ: Erlbaum.
- Bressan, P. (1993a). Neon color spreading with and without its figural prerequisites. *Perception*, *22*(3), 353–361.
- Bressan, P. (1993b). Revisitation of the luminance conditions for the occurrence of the achromatic neon color spreading illusion. *Perception & Psychophysics*, *54*(1), 55–64.
- Bressan, P. (1997). Guest editorial: Contextual effects of colour appearance: Lightness and colour induction, transparency and illumination. *Perception*, *26*(4), 375–377.
- Bressan, P., Mingolla, E., Spillmann, L., & Watanabe, T. (1997). Neon color spreading: A review. *Perception*, *26*(11), 1353–1366.
- Broerse, J., Vladusich, T., & O'Shea, R. P. (1999). Colour at edges and colour spreading in McCollough effects. *Vision Research*, *39*, 1305–1320.
- Bruno, N. (2001). Breathing illusions and boundary formation in space-time. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping in vision* (pp. 402–427). New York, NY: Elsevier Science.
- Bruno, N., & Bertamini, M. (1990). Identifying contours from occlusion events. *Perception & Psychophysics*, *48*(4), 331–342.
- Bruno, N., & Gerbino, W. (1991). Illusory figures based on local kinematics. *Perception*, *20*, 259–274.
- Bull, N. J., Hunter, M., & Finlay, D. C. (2003). Cue gradient and cue density interact in the detection and recognition of objects defined by motion, contrast, or texture. *Perception*, *32*(1), 29–39.
- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, *4*, 121–138.
- Burr, D. C. (1980). Motion smear. *Nature*, *284*, 164–165.
- Carman, G. J., & Welsh, L. (1992). Three-dimensional illusory contours and surfaces. *Nature*, *360*, 585–587.
- Chen, V. J., & Cicerone, C. M. (2002). Depth from subjective color and apparent motion. *Vision Research*, *42*, 2131–2135.
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis

- for studying non-Fourier motion perception. *Journal of the Optical Society of America. A*, 5, 1986–2007.
- Cicerone, C. M., & Hoffman, D. D. (1992). Dynamic neon colors: Perceptual evidence for parallel visual pathways. *Advances in Color Vision Technical Digest (Optical Society of America)*, 4, 66–68.
- Cicerone, C. M., & Hoffman, D. D. (1997). Color from motion: Dichoptic activation and a possible role in breaking camouflage. *Perception*, 26, 1367–1380.
- Cicerone, C. M., Hoffman, D. D., Gowdy, P. D., & Kim, J. S. (1995). The perception of color from motion. *Perception & Psychophysics*, 57(6), 761–777.
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27(3), 183–228.
- Cunningham, D. W., Shipley, T. F., & Kellman, P. J. (1998a). The dynamic specification of surfaces and boundaries. *Perception*, 27(4), 403–415.
- Cunningham, D. W., Shipley, T. F., & Kellman, P. J. (1998b). Interactions between spatial and spatiotemporal information in spatiotemporal boundary formation. *Perception & Psychophysics*, 60(5), 839–851.
- Davis, G., & Driver, J. (1997). A functional role for illusory colour spreading in the control of focused visual attention. *Perception*, 26, 1397–1411.
- Davis, G., & Driver, J. (1998). Kanizsa subjective figures can act as occluding surfaces in parallel stages of visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1), 169–184.
- Day, R. H. (1983). Neon color spreading, partially delineated borders, and the formation of illusory contours. *Perception & Psychophysics*, 34(5), 488–490.
- Day, R. H. (1989). Apparent depth from progressive exposure of moving shadows: The kinetic depth effect in a narrow aperture. *Bulletin of the Psychonomic Society*, 27(4), 320–322.
- De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Perceptual filling-in: A parametric study. *Vision Research*, 38(18), 2721–2731.
- Demkiw, P., & Michaelis, C. F. (1976). Motion information in iconic memory. *Acta Psychologica*, 40(4), 257–264.
- Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature*, 267(5608), 241–243.
- Dixon, P., & Di Lollo, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology*, 26, 33–63.
- Durgin, F. H. (1998). Quasi-modal encounters of the third kind: The filling-in of visual detail. *Behavioral and Brain Sciences*, 21(6), 756–757.
- D’Zmura, M., Colantoni, P., Knoblauch, K., & Laget, B. (1997). Color transparency. *Perception*, 26(4), 471–492.
- Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science*, 287, 2036–2038.
- Ekroll, V., & Faul, F. (2002). Perceptual transparency in neon color spreading displays. *Perception & Psychophysics*, 64(6), 945–955.

References

- Ekroll, V., Faul, F., & Niederée, R. (2004). The peculiar nature of simultaneous colour contrast in uniform surrounds. *Vision Research*, *44*(15), 1765–1786.
- Fahle, M., & Poggio, T. (1981). Visual hyperacuity: Spatiotemporal interpolation in human vision. *Proceedings of the Royal Society London B*, *213*, 451–477.
- Faul, F. (1997). *Theoretische und experimentelle Untersuchung chromatischer Determinanten perzeptueller Transparenz*. PhD thesis, Christian-Albrechts-Universität, Kiel, Germany.
- Faul, F., & Ekroll, V. (2002). Psychophysical model of chromatic perceptual transparency based on subtractive color mixture. *Journal of the Optical Society of America. A*, *19*(6), 1084–1095.
- Feldman, J. A. (1985). Four frames suffice: A provisional model of vision and space. *The Behavioral and Brain Sciences*, *8*(2), 265–289.
- Fendrich, R., Rieger, J. W., & Heinze, H.-J. (2005). The effect of retinal stabilization on anorthoscopic percepts under free-viewing conditions. *Vision Research*, *45*, 567–582.
- Fidopiastis, C., Hoffman, D. D., & Prophet, W. D. (2000). Constructing surfaces and contours in displays of color from motion: The role of nearest neighbors and maximal disks. *Perception*, *29*(5), 567–580.
- Fuchs, W. (1921). Untersuchungen über das Sehen der Hemianopiker und Hemiamblyopiker. II: Die totalisierende Gestaltauffassung. *Zeitschrift für Psychologie*, *86*, 1–143.
- Fujita, N. (1990). Three-dimensional anorthoscopic perception. *Perception*, *19*(6), 767–771.
- Gegenfurtner, K., Brown, J. E., & Rieger, J. (1997). Interpolation processes in the prediction of real and illusory contours. *Perception*, *26*, 1445–1458.
- Gerrits, H. J. M., Haan, B. de, & Vendrik, A. J. H. (1966). Experiments with retinal stabilized images. Relation between the observations and neural data. *Vision Research*, *6*, 427–440.
- Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception & Psychophysics*, *5*(2), 113–116.
- Gregory, R. L. (1972). Cognitive contours. *Nature*, *238*(5358), 51–52.
- Grossberg, S. (1992). Neural FACADEs: Visual representations of static and moving Form-and-Color-and-Depth. In G. W. Humphreys (Ed.), *Understanding vision* (pp. 232–271). Oxford, UK: Blackwell Publishers.
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, *104*(3), 618–658.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*(2), 173–211.
- Guttman, S. E., & Kellman, A. B. S. P. J. (2003). Temporal variations in visual completion: A reflection of spatial limits? *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 1211–1227.

- Guttman, S. E., & Kellman, P. J. (2004). Contour interpolation revealed by a dot localisation paradigm. *Vision Research*, *44*, 1799–1815.
- Haber, R. N., & Nathanson, L. S. (1968). Post-retinal storage? Some further observations on Parks' camel as seen through the eye of a needle. *Perception & Psychophysics*, *3*(5A), 349–355.
- He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature*, *359*(6392), 231–233.
- He, Z. J., & Nakayama, K. (1994). Apparent motion determined by surface layout not by disparity or three-dimensional distance. *Nature*, *367*, 173–175.
- He, Z. J., & Ooi, T. L. (1998). Illusory-contour formation affected by luminance contrast polarity. *Perception*, *27*, 313–335.
- Hecht, H. (1924). Neue Untersuchungen über die Zöllnerschen anorthoskopischen Zerrbilder. *Zeitschrift für Psychologie*, *94*, 153–194. (Herausgegeben von F. Schumann: I. Die simultane Erfassung der Figuren)
- Hildreth, E. C., & Koch, C. (1987). The analysis of visual motion: From computational theory to neuronal mechanisms. *Annual Review of Neuroscience*, *10*, 477–533.
- Hine, T. (1987). Subjective contours produced purely by dynamic occlusion of sparse-points array. *Bulletin of the Psychonomic Society*, *25*(3), 182–184.
- Hofer, H., Carroll, J., Neitz, J., Neitz, M., , & Williams, D. R. (2005). Organization of the human trichromatic cone mosaic. *Journal of Neuroscience*, *25*, 9669–9679.
- Hoffman, D. D. (1998). *Visual intelligence. How we create what we see*. New York, NY: W. W. Norton & Company.
- Hoffman, D. D. (2003). The interaction of colour and motion. In R. J. Mausfeld & D. Heyer (Eds.), *Colour – Mind and the physical world* (pp. 361–378). Oxford, UK: Oxford University Press.
- Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*, 65–96.
- Ilg, U. (1997). Slow eye movements. *Progress in Neurobiology*, *53*, 293–329.
- Ishikawa, H., & Geiger, D. (2005). Illusory volumes in human stereo perception. *Vision Research*, *46*, 171–178.
- Johnson, J. S., & Olshausen, B. A. (2003). Timecourse of neural signatures of object recognition. *Journal of Vision*, *3*(7), 499–512.
- Johnson, J. S., & Olshausen, B. A. (2005). The recognition of partially visible natural objects in the presence and absence of their occluders. *Vision Research*, *45*, 3262–3276.
- Johnson, S. P., & Mason, U. (2002). Perception of kinetic illusory contours by two-month-old infants. *Child Development*, *73*(1), 22–34.
- Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.
- Kalkofen, H. (2000). Anorthoskop (Plateau 1836) und anorthoskopische Erscheinungen. In *42. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Jena*.
- Kanizsa, G. (1979). *Organization in vision*. New York, NY: Praeger.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics*, *6*, 193–198.

References

- Katz, D. (1911). *Die Erscheinungsweisen der Farben und ihre Beeinflussung durch die individuelle Erfahrung*. Leipzig: Barth.
- Kawabata, N. (1984). Perception at the blind spot and similarity grouping. *Perception & Psychophysics*, *36*(2), 151–158.
- Kellman, P. J. (2003). Interpolation processes in the visual perception of objects. *Neural Networks*, *16*, 915–923.
- Kellman, P. J., & Cohen, M. H. (1984). Kinetic subjective contours. *Perception & Psychophysics*, *35*(3), 237–244.
- Kellman, P. J., Garrigan, P. B., Kalar, D., & Shipley, T. F. (2003). Good continuation and relatability: Related but distinct principles [Abstract]. *Journal of Vision*, *3*(9), 120a.
- Kellman, P. J., Garrigan, P. B., & Shipley, T. F. (2005). Object interpolation in three dimensions. *Psychological Review*, *112*(3), 586–609.
- Kellman, P. J., Garrigan, P. B., Yin, C., Shipley, T. F., & Machado, L. (2005). 3-D interpolation in object perception: Evidence from an objective performance paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 558–583.
- Kellman, P. J., Guttman, S. E., & Wickens, T. D. (2001). Geometric and neural models of object perception. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Segmentation and grouping in vision*. Oxford, UK: Elsevier Science.
- Kellman, P. J., & Loukides, M. G. (1987). An object perception approach to static and kinetic subjective contours. In S. Petry & G. E. Meyer (Eds.), *The perception of illusory contours* (pp. 151–164). New York, NY: Springer.
- Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141–221.
- Kellman, P. J., Yin, C., & Shipley, T. F. (1998). A common mechanism for illusory and occluded object completion. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 859–869.
- Kersten, D., Knill, D. C., Mamassian, P., & Bühlhoff, H. H. (1996). Illusory motion from shadows. *Nature*, *279*, 31.
- Kersten, D., Mamassian, P., & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, *26*(2), 171–192.
- Knau, H., & Spillmann, L. (1997). Brightness fading during Ganzfeld adaptation. *Journal of the Optical Society of America. A*, *14*, 1213–1222.
- Koenderink, J. J. (2003). Guest editorial: Another Babel. *Perception*, *32*, 391–394.
- Koenderink, J. J., & van Doorn, A. J. (1976). The singularities of the visual mapping. *Biological Cybernetics*, *24*(1), 51–59.
- Koffka, K. (1935). *Principles of Gestalt Psychology*. New York, NY: Harcourt, Brace, and World.
- Kojo, I. V., Liinauso, M. E., & Rovamo, J. M. (1995). Neon colour spreading in three-dimensional illusory objects [Abstract]. *Investigative Ophthalmology & Visual Science*, *36*(4), S665.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2003). Mental imagery: Against the nihilistic hypothesis. *Trends in Cognitive Science*, *7*(3), 109–111.

- Leshner, G. W., & Mingolla, E. (1993). The role of edges and line-ends in illusory contour formation. *Vision Research*, *33*, 2253–2270.
- MacLeod, D. I. A., & Golz, J. (2003). A computational analysis of colour constancy. In R. J. Mausfeld & D. Heyer (Eds.), *Colour – Mind and the physical world* (pp. 205–242). Oxford, UK: Oxford University Press.
- Mateeff, S., Popov, D., & Hohnsbein, J. (1993). Multi-aperture viewing: perception of figures through very small apertures. *Vision Research*, *33*(17), 2563–2567.
- Mather, G. (1988). Temporal properties of apparent motion in subjective figures. *Perception*, *17*, 729–736.
- Mausfeld, R. J. (1998). Color perception: From Grassmann codes to a dual code for object and illumination colours. In W. G. K. Backhaus, R. Kliegl, & J. S. Werner (Eds.), *Color perception: Perspectives from different disciplines* (pp. 219–250). Berlin: Walther de Gruyter.
- Mausfeld, R. J. (2001). What's within? Can the internal structure of perception be derived from regularities of the external world? *Behavioral and Brain Sciences*, *24*, 689–690.
- Mausfeld, R. J. (2002). The physicalistic trap in perception theory. In R. J. Mausfeld & D. Heyer (Eds.), *Perception and the physical world* (pp. 75–112). Chichester, UK: Wiley.
- Mausfeld, R. J. (2003). 'Colour' as part of the format of different perceptual primitives: The dual coding of colour. In R. J. Mausfeld & D. Heyer (Eds.), *Colour – Mind and the physical world* (pp. 381–429). Oxford, UK: Oxford University Press.
- Mausfeld, R. J., & Niederée, R. (1993). An inquiry into relational concepts of colour, based on incremental principles of colour coding for minimal relational stimuli. *Perception*, *22*, 427–462.
- Metelli, F. (1974). The perception of transparency. *Scientific American*, *230*(4), 90–98.
- Metzger, W. (1975). *Gesetze des Sehens* (3rd ed.). Frankfurt a.M.: Verlag von W. Kramer & Co.
- Meyer, G. E., & Dougherty, T. J. (1987). Effects of flicker-induced depth on chromatic subjective contours. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 353–360.
- Meyer, G. E., & Dougherty, T. J. (1990). Ambiguous fluidity and rigidity and diamonds that ooze. *Perception*, *19*, 491–496.
- Michotte, A., & Burke, L. (1951). Une nouvelle énigme de la psychologie de la perception: Le 'donnée amodal' dans l'expérience sensorielle [A new enigma in the psychology of perception: Amodal data in sensory experience]. In *Actes du 13ème Congrès Internationale de Psychologie, Stockholm* (pp. 179–180).
- Michotte, A., Thinès, G., & Crabbé, G. (1964). Amodal completion of perceptual structures. In G. Thinès, A. Costall, & G. Butterworth (Eds.), *Michotte's experimental phenomenology of perception* (pp. 140–167). Hillsdale, NJ: Erlbaum.
- Miyahara, E., & Cicerone, C. M. (1997). Color from motion: Separate contributions of chromaticity and luminance. *Perception*, *26*, 1381–1396.
- Morgan, M. J. (1980). Analogue models of motion perception. *Philosophical Transactions of the Royal Society of London B*, *290*, 117–135.

References

- Morgan, M. J., Findlay, J. M., & Watt, R. J. (1982). Aperture viewing: A review and a synthesis. *Quarterly Journal of Experimental Psychology*, *34A*, 211–233.
- Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2001). Time course of amodal completion revealed by a shape discrimination task. *Psychonomic Bulletin & Review*, *8*, 713–720.
- Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition* (pp. 1–70). Cambridge, MA: MIT Press.
- Nakayama, K., & Shimojo, S. (1990a). DaVinci stereopsis: Depth and subjective contours from unpaired monocular points. *Vision Research*, *30*, 1811–1825.
- Nakayama, K., & Shimojo, S. (1990b). Toward a neural understanding of visual surface perception. In T. J. Sejnowski, E. R. Kandel, C. F. Stevens, & J. D. Watson (Eds.), *The Brain. cold Spring Harbor symposia on quantitative biology. Volume 55* (pp. 911–924). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science*, *257*(4), 1357–1363.
- Nakayama, K., Shimojo, S., & Ramachandran, V. S. (1990). Transparency: Relation to depth, subjective contours, luminance, and neon color spreading. *Perception*, *19*(4), 497–513.
- Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping and the recognition of occluded objects. *Perception*, *18*, 55–68.
- Nishida, S. (2004). Motion-based analysis of spatial patterns by the human visual system. *Current Biology*, *14*(10), 830–839.
- Palmer, E. M. (2003). *Spatiotemporal reliability in the perception of dynamically occluded objects*. PhD thesis, University of California Los Angeles, Los Angeles, CA.
- Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, *24*(3), 436–447.
- Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, *31*, 1221–1236.
- Paradiso, M. A., Shimojo, S., & Nakayama, K. (1989). Subjective contours, tilt-aftereffects, and visual cortical organization. *Vision Research*, *29*, 1205–1213.
- Parks, T. E. (1965). Post-retinal visual storage. *The American Journal of Psychology*, *78*, 145–147.
- Parks, T. E. (1984). Illusory contours: A (mostly) atheoretical review. *Psychological Bulletin*, *95*, 282–300.
- Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: A guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences*, *21*(6), 723–802.
- Petry, S., & Meyer, G. E. (1987). *The perception of illusory contours*. New York, NY: Springer.
- Piantanida, T. P. (1985). Temporal modulation sensitivity of the blue mechanism:

- Measurements made with extraretinal chromatic adaptation. *Vision Research*, 25, 1439–1444.
- Pinna, B. (1987). Un effetto di colorazione. In V. Majer, R. Maeran, & M. Santinello (Eds.), *Il laboratorio e la città. XXI Congresso degli Psicologi Italiani* (p. 158).
- Pinna, B., Brelstaff, G., & Spillmann, L. (2001). Surface color from boundaries: A new ‘watercolor’ illusion. *Vision Research*, 41(20), 2669–2676.
- Pinna, B., Ehrenstein, W. H., & Spillmann, L. (2004). Illusory contours and surfaces without amodal completion and depth stratification. *Vision Research*, 44, 1851–1855.
- Pinna, B., Werner, J. S., & Spillmann, L. (2003). The watercolor effect: A new principle of grouping and figure-ground organisation. *Vision Research*, 43, 43–52.
- Plateau, J. A. F. (1836). Notice sur l’Anorthoscope. *Bulletin de l’Académie des Sciences et Belles Lettres de Bruxelles*, 3, 7–10; 364.
- Plateau, J. A. F. (1850). Zweite Notiz über neue sonderbare Anwendungen des Verweilens der Eindrücke auf der Netzhaut. *Annalen der Physik und Chemie*, 79, 269–290.
- Poppelreuter, W. (1917). *Die psychologischen Schädigungen durch Kopfschuss im Kriege 1914–1916*. Leipzig: Voss.
- Prazdny, K. (1986). Illusory contours from inducers defined solely by spatiotemporal correlation. *Perception & Psychophysics*, 39(3), 175–178.
- Pylyshyn, Z. (2003a). Return of the mental image: Are there really pictures in the brain? *Trends in Cognitive Science*, 7(3), 113–118.
- Pylyshyn, Z. (2003b). *Seeing and visualizing: It’s not what you think*. Cambridge, MA: MIT Press.
- Ramachandran, V. S. (1992a). Blind spots. *Scientific American*, 266, 85–91.
- Ramachandran, V. S. (1992b). Filling in gaps in perception: I. *Current Directions in Psychological Science*, 1, 199–205.
- Ramachandran, V. S. (1993). Filling in gaps in perception: II. *Current Directions in Psychological Science*, 2, 56–65.
- Ramachandran, V. S., & Gregory, R. L. (1991). Perceptual filling-in of artificially induced scotomas in human vision. *Nature*, 350, 699–702.
- Rao, R. P. N., Eagleman, D. M., & Sejnowski, T. J. (2001). Optimal smoothing in visual motion perception. *Neural Computation*, 13, 1243–1253.
- Rauschenberger, R., & Yantis, S. (2001). Masking unveils pre-amodal completion representation in visual search. *Nature*, 410, 369–372.
- Redies, C., & Spillmann, L. (1981). The neon color effect in the Ehrenstein illusion. *Perception*, 10, 667–681.
- Redies, C., Spillmann, L., & Kunz, K. (1984). Colored neon flanks and line gap enhancement. *Vision Research*, 24, 1301–1309.
- Richards, W. A. (1988). Image interpretation: Information at contours. In W. A. Richards (Ed.), *Natural computation* (pp. 17–35). Cambridge, MA: MIT Press.
- Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America. A*, 4(7), 1168–1175.

References

- Ringach, D. L., & Shapley, R. M. (1996). Spatial and temporal properties of illusory contours and amodal boundary completion. *Vision Research*, *36*, 3037–3050.
- Rock, I. (1981). Anorthoscopic perception. *Scientific American*, *244*(3), 145–153.
- Rodieck, R. W. (1998). *The first steps in seeing*. Sunderland, MA: Sinauer.
- Roorda, A., & Williams, D. R. (1999). The arrangement of the three cone classes in the living human eye. *Nature*, *397*(6719), 520–522.
- Rosenbach, O. (1902). Zur Lehre von den Urteilstäuschungen. *Zeitschrift für Psychologie*, *29*, 434–448.
- Rothschild, H. (1922). Untersuchungen über die sog. Zöllnerschen anorthoskopischen Zerrbilder. *Zeitschrift für Psychologie*, *90*, 137–166.
- Rubin, E. (1915). *Synsoplevede Figurer. Studier i psykologisk Analyse*. Copenhagen: Gyldendalske Boghandel Nordisk Forlag.
- Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception*, *30*(3), 339–366.
- Sabra, A. I. (1978). Sensation and inference in Alhazen's theory of visual perception. In P. K. Machamer & R. G. Turnbull (Eds.), *Studies in perception: Interrelations in the history of philosophy and science*. Columbus, OH: Ohio State University Press.
- Sabra, A. I. (1989). *The optics of ibn al-Haytham: Books I-III: On direct vision*. London, UK: Warburg Institute.
- Sampaio, A. C. (1943). *Etude sur la translation des objets comme facteur de leur permanence phénoménale. L'effet écran [The translation of objects as a factor in their phenomenal permanence]*. Louvain-la-Neuve: Warny.
- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, *38*(2), 259–290.
- Schumann, F. (1904). Beiträge zur Analyse der Gesichtswahrnehmungen. Vierte Abhandlung. Zur Schätzung der Richtung. *Zeitschrift für Psychologie und Physiologie der Sinnesorgane*, *36*, 161–185.
- Sekuler, A. B., Gold, J. M., Murray, R. F., & Bennett, P. J. (2000). Visual completion of partly occluded objects: Insights from behavioral studies. *Neuro-Ophthalmology*, *23*, 165–168.
- Sekuler, A. B., & Palmer, S. E. (1992). Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology: General*, *121*(1), 95–111.
- Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature*, *408*, 788.
- Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research*, *14*, 147–152.
- Shimojo, S., Kamitani, Y., & Nishida, S. (2001). Afterimage of perceptually filled-in surface. *Science*, *293*, 1677–1680.
- Shimojo, S., & Nakayama, K. (1990). Amodal representation of occluded surfaces: Role of invisible stimuli in apparent motion correspondence. *Perception*, *19*, 285–299.
- Shimojo, S., & Richards, W. A. (1986). “Seeing” shapes that are almost totally

- occluded: A new look at Park's camel. *Perception & Psychophysics*, *39*(6), 418–426.
- Shiori, S., & Cavanagh, P. (1992). Visual persistence of figures defined by relative motion. *Vision Research*, *32*(5), 943–951.
- Shiori, S., & Cavanagh, P. (2000). Nonlinearity in color space measured by relative motion. *Perception & Psychophysics*, *62*(6), 1182–1190.
- Shipley, T. F., & Cunningham, D. W. (2001). Perception of occluding and occluded objects over time: Spatiotemporal segmentation and unit formation. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Grouping and segmentation in vision* (pp. 557–585). New York, NY: Elsevier Science.
- Shipley, T. F., & Kellman, P. J. (1992). Perception of partly occluded objects and illusory figures: Evidence for an identity hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 106–120.
- Shipley, T. F., & Kellman, P. J. (1993a). Optical tearing in spatiotemporal boundary formation: When do local element motions produce boundaries, form, and global motion? *Spatial Vision*, *7*, 323–339.
- Shipley, T. F., & Kellman, P. J. (1993b). Spatiotemporal boundary formation: Temporal integration is confined to a 150 msec window. *Investigative Ophthalmology and Visual Science Supplement*, *34*(4), 1082.
- Shipley, T. F., & Kellman, P. J. (1994). Spatiotemporal boundary formation: Boundary, form, and motion perception from transformations of surface elements. *Journal of Experimental Psychology: General*, *123*, 3–20.
- Shipley, T. F., & Kellman, P. J. (1997). Spatiotemporal boundary formation: The role of local motion signals in boundary perception. *Vision Research*, *37*(10), 1281–1293.
- Shipley, T. F., & Kellman, P. J. (2003). Boundary completion in illusory contours: Interpolation or extrapolation? *Perception*, *32*(8), 985–999.
- Shirley, R. W. (1983). *The mapping of the world: Early printed world maps, 1472–1700*. London, UK: Holland Press.
- Shore, D. I., & Enns, J. T. (1997). Primitive shape completion time depends on the size of the area to be completed. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 980–998.
- Singh, M. (2004). Modal and amodal completion generate different shapes. *Psychological Science*, *15*(7), 454–459.
- Singh, M., & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review*, *109*, 492–519.
- Singh, M., & Hoffman, D. D. (1999). Completing visual contours: The relationship between relatability and minimizing inflections. *Perception & Psychophysics*, *61*(5), 943–951.
- Singh, M., & Hoffman, D. D. (2001). Part-based representations of visual shape and implications for visual cognition. In T. F. Shipley & P. J. Kellman (Eds.), *From fragments to objects: Grouping and segmentation in vision* (pp. 401–459). New York, NY: Elsevier Science.

References

- Sohmiya, T., & Sohmiya, K. (1994). What is a crucial determinant in anorthoscopic perception. *Perceptual and Motor Skills*, *78*(3), 987–998.
- Spehar, B. (2000). Degraded illusory contour formation with non-uniform inducers in Kanizsa configurations: The role of contrast polarity. *Vision Research*, *40*, 2653–2659.
- Spehar, B. (2002). The role of contrast polarity in perceptual closure. *Vision Research*, *42*, 343–350.
- Sperling, G., & Lu, Z. (1998). A systems analysis of visual motion perception. In T. Watanabe (Ed.), *High-level motion processing* (pp. 153–183). Cambridge, MA: MIT Press.
- Stappers, P. J. (1989). Forms can be recognized from dynamic occlusion alone. *Perceptual and Motor Skills*, *68*, 243–251.
- Takeichi, H., Shimojo, S., & Watanabe, T. (1992). Neon flank and illusory contour: Interaction between the two processes leads to color filling-in. *Perception*, *21*(3), 313–324.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381*, 520–522.
- Todorović, D. (1997). Lightness and junctions. *Perception*, *26*, 379–394.
- Troxler, D. (1804). Ueber das Verschwinden gegebener Gegenstände innerhalb unseres Gesichtskreises. In K. Himly & J. A. Schmidt (Eds.), *Ophthalmologische Bibliothek II* (pp. 1–119). Jena: F. Frommann.
- Tse, P. U. (1998). Illusory volumes from conformation. *Perception*, *27*, 977–992.
- Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, *39*, 37–68.
- Tse, P. U. (2002). A contour propagation approach to surface filling-in and volume formation. *Psychological Review*, *109*(1), 91–115.
- Tse, P. U., & Albert, M. K. (1998). Amodal completion in the absence of image tangent discontinuities. *Perception*, *27*, 455–464.
- Tynan, P., & Sekuler, R. (1975). Moving visual phantoms: A new contour completion effect. *Science*, *188*, 951–952.
- Uttal, W. R., Spillmann, L., Stürzel, F., & Sekuler, A. B. (2000). Motion and shape in common fate. *Vision Research*, *40*, 301–310.
- van Lier, R. (2002). A double neon colour illusion. *Perception*, *31*, 31–38.
- van Tuijl, H. F. J. M. (1975). A new visual illusion: Neonlike color spreading and complementary color induction between subjective contours. *Acta Psychologica*, *39*, 441–445.
- van Tuijl, H. F. J. M., & de Weert, C. M. M. (1979). Sensory conditions for the occurrence of the neon spreading illusion. *Perception*, *8*, 211–215.
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, *7*(5), 207–213.
- Varin, D. (1971). Fenomeni di contrasto e diffusione cromatica nell'organizzazione spaziale del campo percettivo. *Rivista di Psicologia*, *65*, 101–128.
- von Helmholtz, H. (1894). *Handbuch der physiologischen Optik* (2nd ed.). Hamburg: Voss.

- von Vierordt, K. (1868). *Der Zeitsinn nach Versuchen; §27. Scheinbare Verzerrung bewegter Gegenstände*. Tübingen: Laupp.
- Wachtler, T., & Wehrhahn, C. (1997). The Craik-O'Brien-Cornsweet illusion in colour: Quantitative characterisation and comparison with luminance. *Perception*, *26*(11), 1423–1430.
- Walker, R., & Mattingley, J. B. (1998). Pathological completion: The blind leading the mind? *Behavioral and Brain Sciences*, *21*(6), 778–779.
- Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *208*, 325–380.
- Watanabe, T., & Cavanagh, P. (1991). Texture and motion spreading, transparency and the aperture problem. *Perception & Psychophysics*, *50*, 459–464.
- Watanabe, T., & Sato, T. (1989). Effects of luminance contrast on color spreading and illusory contour in the neon color spreading effect. *Perception & Psychophysics*, *45*(4), 427–430.
- Watanabe, T., & Takeichi, H. (1990). The relation between color spreading and illusory contours. *Perception & Psychophysics*, *47*(5), 457–467.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120.
- Weisstein, N., Maguire, W., & Berbaum, K. (1977). A phantom-motion aftereffect. *Science*, *198*, 955–958.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, *4*, 301–350.
- Whittle, P. (1998). Filling-in does require a mechanism, and some persistent doubts. *Behavioral and Brain Sciences*, *21*(6), 779–780.
- Wollschläger, D. (2002). *Zum Zusammenhang von Perzeptueller Transparenz und Neon-Color-Spreading*. Masters Thesis, Christian-Albrechts-Universität, Kiel, Germany.
- Wollschläger, D., Rodriguez, A. M., & Hoffman, D. D. (2001). Flank transparency: Transparent filters seen in dynamic two-color displays. *Perception*, *30*(12), 1423–1426.
- Wollschläger, D., Rodriguez, A. M., & Hoffman, D. D. (2002). Flank transparency: The effects of gaps, line spacing, and apparent motion. *Perception*, *31*(9), 1073–1092.
- Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae* (2nd ed.). New York, NY: John Wiley and Sons.
- Yantis, S. (1995). Perceived continuity of occluded visual objects. *Psychological Science*, *6*(3), 182–186.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.
- Yin, C., Kellman, P. J., & Shipley, T. F. (1997). Surface completion complements boundary interpolation in the visual integration of partly occluded objects. *Perception*, *26*, 1459–1479.
- Yin, C., Kellman, P. J., & Shipley, T. F. (2000). Surface integration influences depth discrimination. *Vision Research*, *40*, 1969–1978.
- Yuille, A. L., & Bülthoff, H. H. (1996). Bayesian decision theory and psychophysics.

References

- In D. C. Knill & W. A. Richards (Eds.), *Perception as Bayesian inference* (pp. 123–161). Cambridge, UK: Cambridge University Press.
- Zöllner, F. (1862). Ueber eine neue Art anorthoskopischer Zerrbilder. *Annalen der Physik und Chemie*, 117, 477–484.

Appendices

Appendix A

Stimuli and adaptive algorithm used in experiment 5

A.1 Algorithm development

In the following paragraphs, I will first present in detail the sequence of steps that defines the adaptive algorithm employed in experiment 5 (cf. section 5.6). This description will be followed by a section that provides the motivation behind the chosen procedure. It also explains how we carried out several simulations to identify the convergence characteristics of the algorithm, to determine suitable choices for its free parameters, and to compare its properties with an already-established adaptive procedure.

Procedure. The algorithm we used in Experiment 5 to determine the tested line distances first divided the complete parameter range into two intervals (10–30 px and 30–50 px).¹ It then randomly drew and tested values from each interval until each interval would contain at least 40 data points. Our simulations (cf. section 5.6) had indicated that averaging over 40 trials would yield sufficiently stable probability estimates. It then calculated the mean percentages correct for each interval (thereby including its boundaries) and compared the results between intervals. If the percentage correct differed by more than 0.1, they were subdivided into four intervals (10–20 px, 20–30 px, 30–40 px and 40–50 px). Values from each new interval were then randomly drawn and tested for the subject’s response until the intervals all contained at least 40 data points. Mean percentages correct for all existing intervals were (re-)calculated and again compared between adjacent intervals. If the performance between adjacent intervals would differ by more than 0.1 (cf. section 5.6), the process of recursive subdivision and testing would be repeated for these intervals. This process continued either until all differences between the mean percentages correct of adjacent intervals were less than 0.1, or until the minimum interval width of 1 px was reached. If the subdivided intervals

¹In this section, pixels will henceforth be used instead of degrees of visual angle, since the former were the unit that the algorithm operated on. For reference, 50 px \approx 1.33 deg with the employed viewing distance of 80 cm.

had boundaries at pixel fractions, the boundaries were set to the nearest integer pixel value.

The described method may be seen as an implementation of the standard breadth-first-search algorithm that searches each branch at one level of the search tree before descending into the next-deeper level. The nodes of the search tree here correspond to pairs of tested intervals, and the depth of the search tree represents the resolution, i.e., the size of tested intervals (figure A.1). The decision criterion used at each node to decide whether to descend into a deeper level of the tree corresponds to the question whether the performance difference between adjacent intervals was greater than 0.1. We favored a breadth-first search over a depth-first approach because the latter searches each branch of the search tree down to the deepest level before taking on the next branch. For the experiment, this would have had the consequence of presenting only line spacings within the same interval for a long period of time before moving to the next branch with other line distances. No randomization of this variable would have been possible. In contrast, the breadth-first algorithm allows for picking distances from a comparatively wide parameter range, thus making randomization possible.

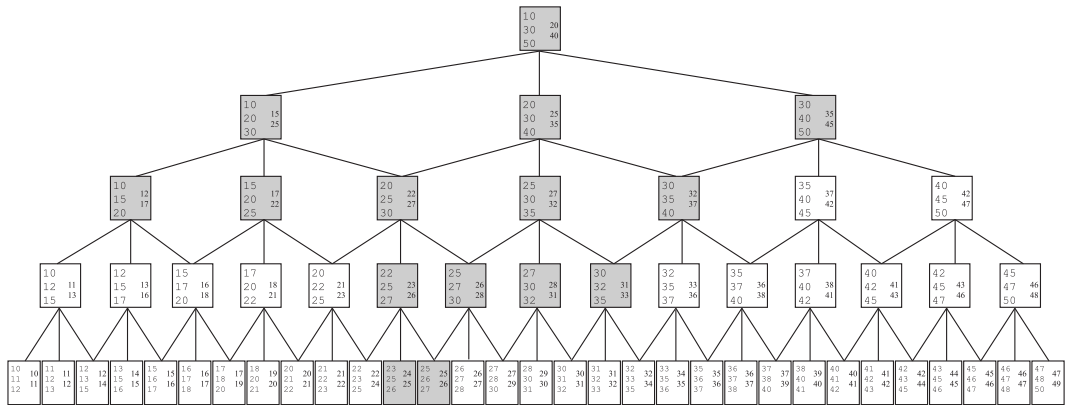


Figure A.1: Search tree used for the adaptive algorithm in experiment 5. The numbers in each box designate the interval boundaries as well as the interval mid-points used to plot the performance assessment. Colored boxes indicate which nodes were visited for subject DWI with a speed of $1.77 \text{ deg} \cdot \text{s}^{-1}$.

Note that the number of trials within an already tested interval further up the search tree could still increase in a progressing experiment as finer-grained intervals, corresponding to nodes deeper down the search tree, were filled up with trials. As a consequence, mean percentages correct for already tested intervals could change later on. This, however, was no problem because each newly-opened interval was guaranteed to be tested until it contained at least 40 trials. In addition, after each trial, performance differences between adjacent intervals would be re-calculated, and the search tree would be re-entered at the top. Therefore, if there was no difference between two adjacent intervals after trial c , but later after trial $c + 50$, the difference

would be detected, the intervals subdivided and tested. Conversely, if there was a difference between two intervals after trial c , but no more so after trial $c + 50$, the subdivision and testing had already happened, resulting only in a finer sampling than what would have been actually necessary.

When averaging data over subjects, the chosen approach leads to the difficulty that different subjects are likely to have different patterns of tested intervals. No set of intervals can thus be chosen that precisely reflects that intervals that were actually tested in the experiment with all the subjects. As a compromise, we chose a fixed interval width when pooling the data. This was done after assessing the effect of different widths with respect to the shape of the resulting curve (figure A.2). If the interval width is chosen too small, there may be not enough data points within all intervals to make the averaging stable, if it is chosen to be very large, the resulting curve becomes coarse. For the data of experiment 5, the resulting curves were very similar in shape with interval widths from 2 to 5 px (figure A.2). An interval width of 4 px was chosen as a result. We also checked whether having different offsets before the first interval would have an effect on the curve's shapes. Theoretically, this might be the case as different intervals with different number of trials in it result. Nevertheless, we found it not to be the case, and thus chose an offset of 0.

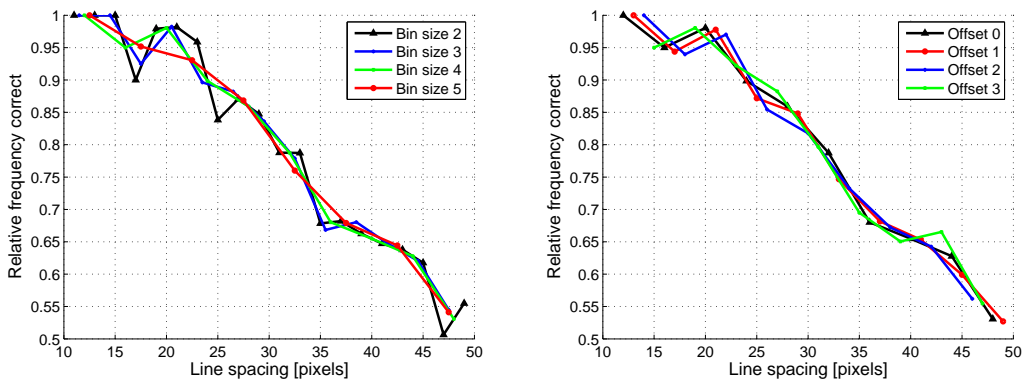


Figure A.2: Using different interval widths (left) and offsets (right) when averaging data over multiple subjects. Figure shows exemplary effects for the data in experiment 5 with a speed of $1.33 \text{ deg}\cdot\text{s}^{-1}$. When varying the interval size, the offset was set to 0, when varying the offset, interval size was set to 4.

Algorithm development. Preliminary experiments where line distance was an independent variable with fixed steps that was fully crossed with object velocity indicated two things: First, with line spacings from 10–50 px, recognition performance varied between chance performance and perfect discrimination. We thus chose this interval for any further pilot and final experiments. Second, it became clear that with a fully

crossed design, the experiment would take far too long if a sufficient sampling of the respective underlying curve was desired. Therefore, the rationale behind the described method was to be able to adjust the sampling rate depending on the underlying function's rate of change. Having a constantly high sampling rate not only increases the total number of trials, it is also moot to spend much time sampling in parameter intervals where the curve does not change much. Conversely, in order to precisely recover the shape of the curve in areas of high volatility, a fine sampling is needed.

As an additional difficulty, we neither wanted to assume a specific shape of psychometric the function, nor did it seem wise to assume *a priori* that all observers had highly similar curves. The former point ruled out choosing a stock adaptive method like QUEST (Watson and Pelli, 1983), because they typically assume a fixed function, e.g., of Weibull type, and only estimate its free parameters. In contrast, we wanted to reconstruct a function of unknown type over the whole range of the independent variable. As a consequence the algorithm needed to detect areas of high rate of change automatically, depending on an observer's available responses. As a measure of rate of change, we chose the difference between two adjacent intervals in terms of their respective average values. If this difference exceeded a limit of 0.1, intervals were made smaller, corresponding to a more fine-grained sampling. By limiting the maximally tolerated difference between adjacent intervals, the taken approach promises to lead to an approximately constant smoothness of the reconstructed curve because each adjacent data point differs by no more than 0.1 (provided that the resolution limit was not reached before). Furthermore, the lowered sampling rate in parameter ranges of slow performance change helped save trials and thus reduced the experiment's duration.

As an unwanted side effect, nearly all adaptive methods have the disadvantage of risking subjects' frustration when carrying out the experiment, and our method is no exception to this rule. This is because most time is spent over parameter ranges with performances of about 0.8–0.6 where the subject is never quite sure if he perceives a signal- or noise-only stimulus. By keeping the subject in constant limbo with respect to the certainty of his response, his patience may get strained a little. We accepted this disadvantage as a fair trade for the reduced number of trials the adaptive method promised.

The employed algorithm has two main parameters which needed to be set to a fixed value for the experiment, namely the minimum number of trials per interval required for averaging, and the maximally tolerated performance difference between two sampling points. Preliminary experiments and several simulations were carried out to determine suitable values for these two parameters. The goal was to find a combination which allowed for a sufficiently fine sampling of the curve to reconstruct it without too much error, but also to simultaneously keep the number of necessary trials as low as possible.

As mentioned above, a preliminary experiment was carried out where line distance was an independent variable with four steps, with 80 trials for each distance. The obtained data was then analyzed within a signal-detection framework: The presence of a

contiguous snake in the stimulus (figure 5.9 on page 140) was treated as the signal-, its absence as the noise-only condition. Perceived line distance (assumed to be equivalent to physical line distance) represented the parameter continuum. Assuming normal distributions with equal standard deviation of 1 for a subject’s response intensity to both signal- and noise-stimuli, we calculated the signal detection parameters d' , $\mu(\text{signal})$, $\mu(\text{noise})$, and the criterion likelihood as the ratio of densities at the criterion point. Since no pay-off was associated with correct or false classifications, and both stimulus classes appeared at an equal probability of 0.5, we expected the likelihood to be about 1, as was indeed the case (except for the closest line spacing; cf. table A.1).

Parameter	Values			
Line distance [px]	45	35	25	15
$p(\text{hit})$	0.400	0.725	0.850	0.990
$p(\text{false alarm})$	0.525	0.300	0.175	0.025
$\mu(\text{noise})$	-0.063	0.524	0.935	1.960
$\mu(\text{signal})$	0.253	-0.598	-1.036	-2.326
d'	-0.316	1.122	1.971	4.286
likelihood	1.031	1.042	1.106	2.193

Table A.1: Signal-detection parameter values calculated from data in a preliminary experiment.

In a simulation, we then generalized the behavior of the subject to the full range of line-distance settings by assuming that his behavior conformed to the signal-detection model. The likelihood for choosing the criterion point was assumed to be constant for all parameter values and was set to the average of calculated likelihoods. The parameters necessary to calculate the probability of having a correct response for each line distance were $\mu(\text{noise})$ and d' . We interpolated each parameter by visually fitting a logistic psychometric function to its four empirical values, which was possible to a sufficiently good degree (figure A.3): $d' = 5/(1 + \exp(-3.5 + 0.15 \cdot x))$, $\mu(\text{noise}) = 2.5/(1 + \exp(-3.5 + 0.15 \cdot x))$. This observer will be henceforth called the signal-detection observer.

The aforementioned steps allowed us to simulate the behavior of the observer over the whole range of line distances by randomly sampling responses according to the interpolated signal and noise distributions, and the criterion given by the likelihood. This was done in order to assess over how many trials we would need to calculate the relative frequency of correct answers to get a stable estimate. Figure A.4 displays the result of these simulations in terms of percentage correct answers for a given line distance. Each line distance (with a granularity of 1 px), was drawn 10, 50, 100, or 500 times. The figure also shows the expected probability of a correct answer as exactly calculated from the interpolated signal-detection parameters. The according psychometric function can be fit to a very good degree with the following Weibull

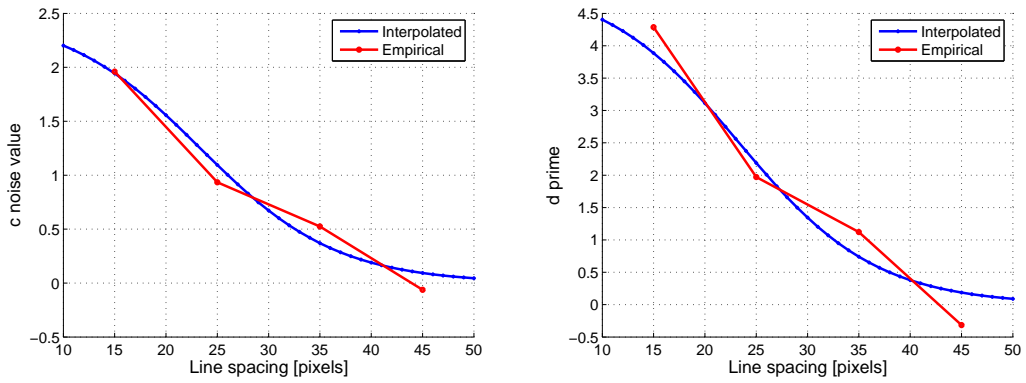


Figure A.3: Empirical $\mu(\text{noise})$ and d' values as well as their respective interpolation with visually-fitted logistic psychometric functions.

function: $f(x) = 0.01 \cdot 0.5 + (1 - 0.01) \cdot (1 - (1 - 0.5) \cdot \exp(-10^{-0.06 \cdot (x-27)}))$.²

Results indicate that, not surprisingly, the exact probability is increasingly better approached with more repetitions for each line distance. As is also apparent, the higher variance with probabilities approaching 0.5 leads to an increase in deviations between simulated percentages correct and the exact probabilities in the upper range of line distances.

Figure A.4 shows that to adequately reconstruct an observer's performance curve, averaging over 10 trials for each line distance is by far not enough to get stable percentages correct that are similar to the exact probabilities. Averaging over 50 trials seems to already give a good estimate. However, this number would result in a total number of trials that is unfeasible to carry out in an actual experiment if many distances are to be tested. And even then, larger errors in the upper parameter range could probably not be avoided. As a consequence, the approach of having line distance as an independent variable with fixed steps was dropped in favor of the adaptive procedure described above.

The adaptive algorithm was now subjected to several simulations to test whether it actually works in the sense that it gives results that eventually converge against the true underlying function, given enough trials. To this end, we simulated the algorithm for different settings of trials per interval and of the maximally-tolerated performance difference between intervals. In the simulations, we used two different theoretical observers—for one, the signal-detection observer just described, and second, an observer with a linear psychometric function ranging from 1.0–0.5 performance over the parameter range. The latter one will be referred to as the linear observer. We took

²Meaning that the threshold, defined as the point with a performance of 0.82, was reached at a line spacing of 27 px. The steepness was -0.06 , chance performance was 0.5, and the observer makes input errors at a rate of 0.1.

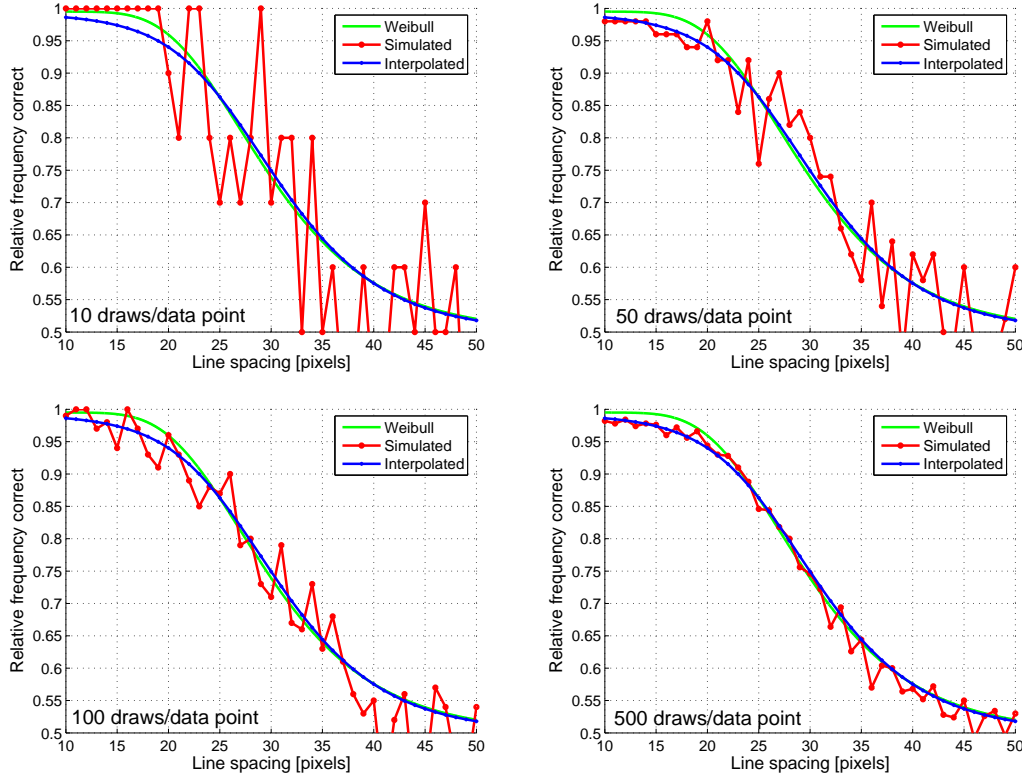


Figure A.4: Simulations of a signal-detection observer based on data from a preliminary experiment. Each line distance (with a granularity of 1 px), was drawn 10, 50, 100, and 500 times. Figures also depict exact percentages correct and the fit with a Weibull function.

the respective psychometric functions of these observers as the to-be-recovered true function

For each simulation, we calculated the required total number of trials until the algorithm would stop, and the square-root of mean squared deviations (RMS) between exact probabilities and simulated relative frequencies of correct answers. This was done using the data at the interval mid-points, i.e., the output of the adaptive algorithm. We also calculated the absolute difference between real thresholds, defined as the line distance where performance was 0.82, and thresholds estimated from the simulated drawings. These were calculated by fitting an approximating cubic spline through the data points that the algorithm yielded. It was ensured that the resulting spline was monotonic, such that the threshold was unambiguously defined. To reduce artifacts from random deviations, each simulation was repeated 20 times, and the results averaged.

Appendix A Stimuli and adaptive algorithm used in experiment 5

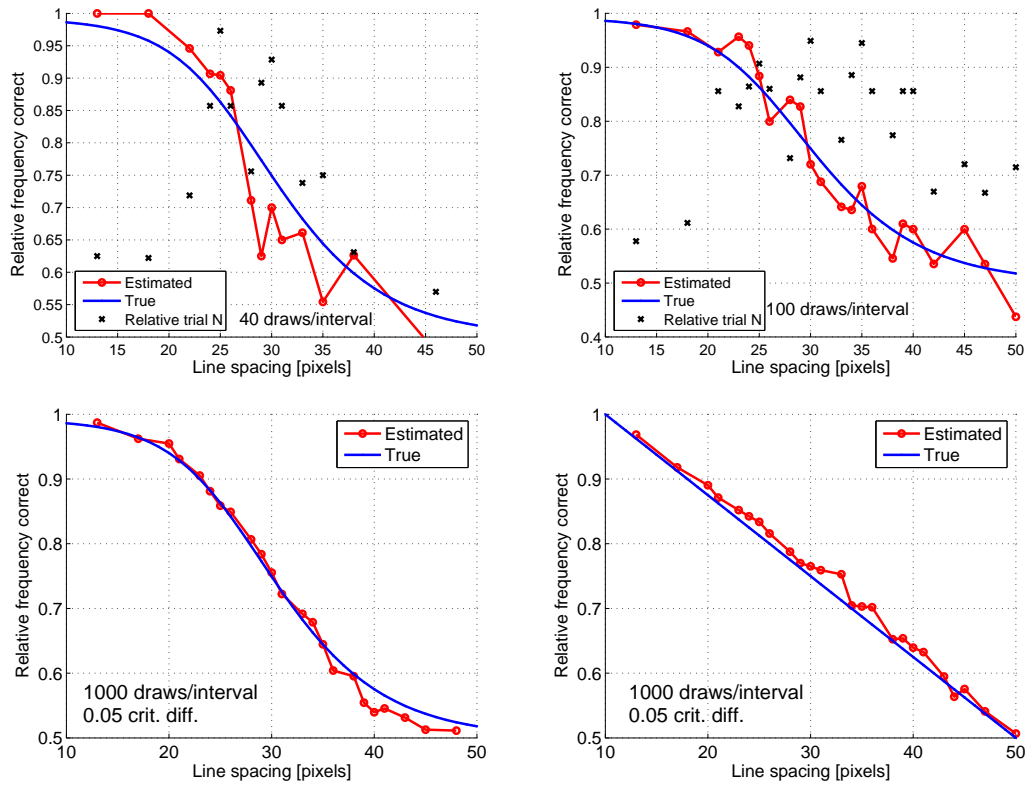


Figure A.5: Results from simulating the adaptive algorithm with the different observers. The figures show the estimated relative frequencies of correct answers depending on settings for the minimum number of draws from each interval. The critical difference is kept constant at 0.1 (top) or 0.05 (bottom). The top figures also show in relative terms how many trials were spent over sections of the parameter range.

Figure A.5 takes a look at some detailed results for individual simulations with the signal-detection observer before the full set of simulations will be commented on below. The critical performance difference was set to 0.1, and the number of trials per interval varied. Among other things, the figure shows the relative number of trials spent over each section of the parameter range (as does figure A.10b). This allows for checking whether the algorithm worked as intended, i.e., whether it saved trials in relatively constant sections and increased the number of trials in sections of quicker performance changes. To a fair degree, this seems to be the case. More importantly, the algorithm did succeed in reconstructing the curve given enough trials per interval. With many number of trials per interval and a low tolerance for performance differences between adjacent intervals, the algorithm seems to converge against the true function.

To assess these characteristics more formally and over a wider parameter range,

figure A.6 shows the complete data. The figure depicts how the total number of trials and the RMS develop when increasing the number of drawings from each interval, and when changing the maximum performance difference between adjacent intervals. When varying the trials per interval, the critical difference was kept constant at 0.1. When varying the critical difference, trials per interval were set to 40. Note that the calculation of the RMS was done not to learn its precise values, but rather to get a qualitative assessment of its development, i.e., to see if it would abruptly drop with some parameter value.

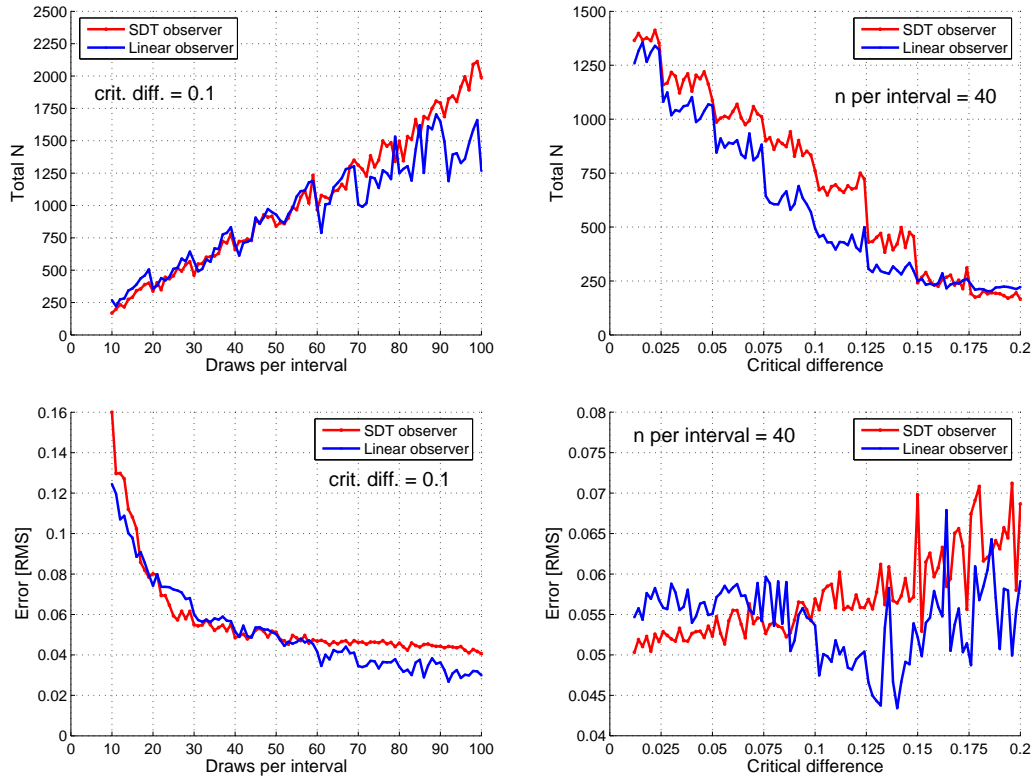


Figure A.6: Simulations for a signal-detection, and a linear observer for various choices of draws per interval and maximum performance difference between adjacent intervals. The RMS error between randomly drawn subject's settings and exact probabilities was calculated with the interval mid-points. The total number of trials is also shown.

With respect to the total number of trials, we can observe an approximately linear increase with more trials per interval, and with a near linear decrease with larger allowed performance differences between intervals. The algorithm seems to perform a bit better with the linear observer, but not much. The RMS falls roughly exponentially with more trials per interval, with a slightly steeper descent for the linear observer.

Regarding the critical difference, its increase leads to a slow, linear increase in RMS, but only for the signal-detection observer. For the linear observer, the tolerated performance difference does not seem to be systematically detrimental to RMS. Indeed, the logic of the algorithm would predict this behavior, since the underlying linear function could theoretically be approximated without error by just two data points and their interpolation. The absolute error of threshold estimates first seems to fall linearly with the number of draws per interval and reaches a stable plateau of 1 px at about 40 draws for the signal-detection, and at about 50 draws for the linear observer. Increasing the critical difference again leads to an increase in estimation errors, but only after a value of about 0.1.

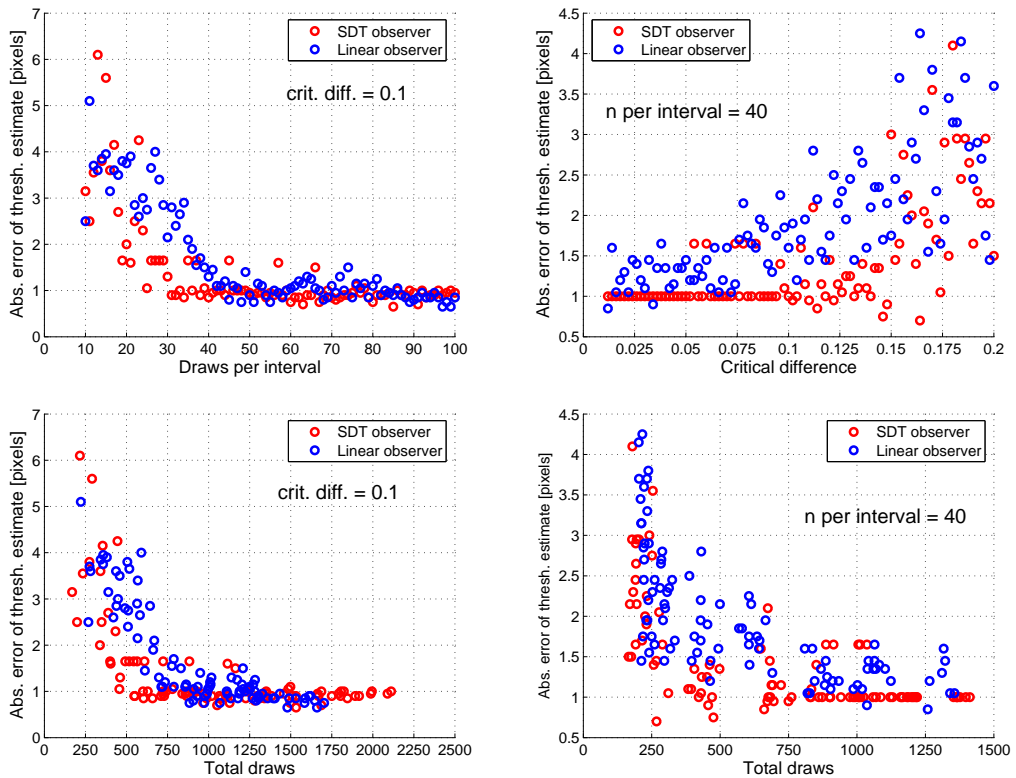


Figure A.7: Simulations for a signal-detection, and a linear observer. Figures show how the absolute error of threshold estimates depends on the settings for draws per interval, for the critical difference between intervals, and on the total number of trials. In one condition, the number of draws per interval was varied, in the other condition, the critical difference between adjacent intervals, while keeping the respective other parameter value constant.

Figure A.7 plots the the error of threshold estimate, depending on the settings for draws per interval, for the critical difference between intervals, and on the total number

of trials. For both observers, we can conclude that with an increasing number of total trials, the threshold estimate approaches the real threshold with suitable parameter choices. This pattern can be observed when the number of trials per interval is varied and the critical performance difference held constant, and vice-versa. However, in both conditions the threshold estimate would not come closer than 1 px to the real threshold. This seems curious but may simply be due to the granularity with which the parameter range can be manipulated. Note that for the linear observer, threshold estimates were worse than for the signal-detection observer with low number of draws from each interval. A likely cause for this result is that spline interpolation naturally fits linear functions not as well as curving ones when there are only a few knots.

By mainly observing the exponential drop of both the RMS and the threshold estimate's error with an increasing number of trials per interval, we see that with a setting of 40, we can expect a sufficient reconstruction of the curve. Further increasing this number does not seem to buy much advantage in terms of precision, but would result in a linear increase in total number of trials. Based on this compromise, trials per interval were set to 40 and the critical performance difference was set to 0.1 for the actual experiment reported in section 5.6. Given these two values, we expected that about 700 trials would be necessary for reconstructing a single curve. This estimate was largely borne out in the experiment in terms of order of magnitude—actual numbers of required trials were about half that estimate.

Importantly, the simulations also showed that the algorithm was not very sensitive to the kind of underlying function and its parameters. At least for the two different monotonic functions used, results were highly similar. This means that its efficiency in reducing the number of trials would not be much affected by being confronted with different real subjects and their possibly differing performance levels or psychometric functions. This was the reason for not simply choosing an established algorithm like QUEST (Watson and Pelli, 1983). In order to see how QUEST performs when its assumptions are not met, some further simulations were carried out.

When trying to compare the performance of QUEST against the new adaptive procedure, it is important to bear in mind that the two have very different objectives. Our own algorithm seeks to recover a complete function, irrespective of its type. QUEST, however, assumes a Weibull function whose steepness and horizontal offset (equivalent to the threshold) need to be initially guessed. Keeping the steepness estimate fixed, QUEST then tries to determine the real threshold and chooses parameter values accordingly. The output of QUEST is thus just a threshold estimate (with an associated confidence), much less information than what is generated by our own procedure. The functioning of QUEST lets one assume that it might be sensitive to the accuracy of the initial parameter estimates, and in particular of the threshold. To test this intuition, simulations were carried out, on the one hand, with the signal-detection observer, whose psychometric function can be fit quite well with a Weibull function (figure A.4). In the simulation, we set all QUEST parameters to that of the fitting Weibull function, except for the initial threshold guess, which was varied. We also tested the linear

observer, in which case we left the parameters equal to those for the signal-detection observer. The number of trials was held constant at 400, the simulation was carried out 20 times, and results averaged.

Figure A.8 plots the error of the threshold estimate that was the output of QUEST as implemented by Denis Pelli for the Matlab Psychophysics Toolbox (Brainard, 1997), depending on the deviation of the initial guess from the real threshold. Clearly, the accuracy of the initial threshold guess has a strong effect on the correctness of QUEST's threshold estimate. The error increases linearly with the offset of the guess from the real value once this offset surpasses 2 px. Two aspects of the results seem especially worth mentioning: The linear function relating the error of the initial guess and the error of the final estimate has a slope that is close to 1. Furthermore, the results are almost identical whether QUEST is given 400 or 800 trials to find the real threshold. The two results taken together probably indicate that QUEST starts off at the initial guess and then tries to find the real threshold within an environment of the guess. However, this environment seems to be constant, such that QUEST does not even try to look for the real threshold outside of, in our case, 1 px to the left and to the right of the initial guess. 400 trials seem to be enough to fully search within this environment, and can thus compensate for a deviation of 2 px.

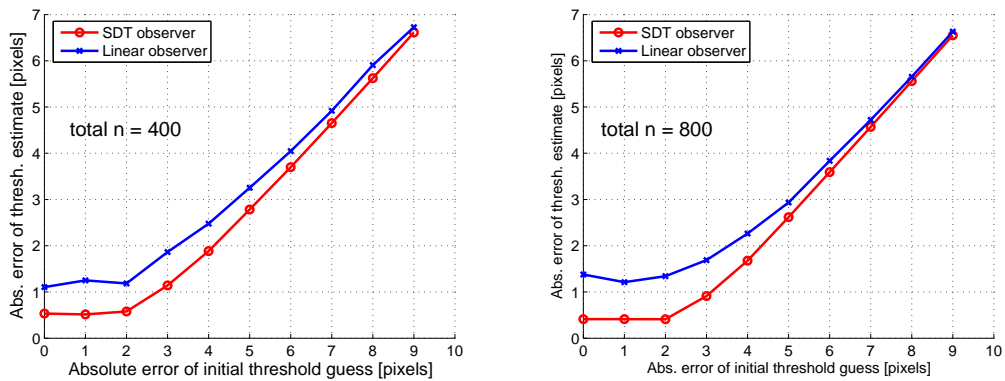


Figure A.8: Error of QUEST's final threshold estimate after 400 and after 800 trials depending on the deviation of the initial threshold guess compared to the real threshold. All other Weibull parameters were set to match the ones of the true function.

The efficiency and usefulness of QUEST thus crucially depend on how much accurate information about an observer one has before conducting an experiment. Given that in experiment 5, we could not very well guess this information, QUEST would not have been a suitable choice. In addition, because of its reliance on the initial threshold guess, ending QUEST prematurely carries the danger of leading one far astray with respect to the final estimate. As might be expected, QUEST's performance is a little better for the signal-detection observer, whose psychometric function is very close to the shape of

the supplied Weibull function. However, even though the function types do not match, the best estimate is also close to perfect for the linear observer, indicating that QUEST is probably not very sensitive to the accuracy of its other initial parameters.

A.2 Stimulus construction

The stimulus class chosen for experiment 5 (cf. section 5.6) consisted of non-meaningful images (figure 5.9, page 140) which were constructed in the following way: First, a disk was completely covered with randomly placed circles whose varying radii were chosen from a Gaussian distribution with a mean of 13.5 min of arc and a standard deviation of 2.70 min of arc. Next, we defined three invisible, non-intersecting Bézier curves with a thickness of 0.25 deg which started and ended at the disk's boundary. As a final step, we colored those circles whose center lay outside the curves green and then those whose center fell within one of the curves red.

Subjects had to discriminate between two classes of the described stimulus. In one class, all three Bézier curves were interrupted once within the disk such that a gap would separate two parts of the curve. In the other class, only two curves were interrupted such that there remained one continuous curve without a gap. The rationale behind the stimulus construction was that in order to verify the presence of a continuous line, the observer must have a complete (i.e., gapless) perceptual representation of that line. The two classes should not be distinguishable based on local features. Since both classes contained continuous line segments as well as gaps between segments, the detection of a continuous segment or of a gap did not suffice to make a reliable discrimination. In addition, the number of circles within the curves did not vary systematically between classes, such that the total amount of redness in a stimulus was not diagnostic for a correct classification.

This stimulus may be varied along several dimensions, the main ones being the size of the circles that the texture is composed of, and their coloration inside as well as outside of the Bézier curves. The coloration itself was chosen to be isoluminant, differing only in the chromaticities. But $2\text{cd}\cdot\text{m}^{-2}$ luminance noise was added to compensate for inter-individual differences or calibration imprecisions with regard to luminance. The colors were randomly chosen from two uni-dimensional Gaussian distributions with red (CIE $u = 0.16$, $v = 0.47$, $L = 12$) and green means (CIE $u = 0.24$, $v = 0.45$, $L = 12$) and standard deviations of 0.02 along the axis which connected the two means in CIE u, v, L chromaticity space with the achromatic point.

In preliminary experiments, we tested how more similar colorations inside and outside of the curves as well as different settings for chromatic variance affected discrimination performance. To that end, we conducted two kinds of experiments: In the first one, we simply flashed stimuli of the described kind for 150 ms and then masked it with a chromatic noise mask consisting of dots the same kind of as the target. We recorded discrimination performance depending on settings for chromatic similarity, determined

by the absolute distance from each other CIE u, v, L chromaticity space, and chromatic variance. First, we set the mean chromaticities of the dots inside and outside of the shape such that they were maximally saturated while being equidistant to the achromatic point. As an experimental manipulation, we now varied the distance between them in the following way: The tested mean chromaticity was a convex combination of the respective maximally distant chromaticity and the achromatic point. To scale the convex combination, we used factors of 0.9, 0.8, 0.7, and 0.6 (figure A.9). This means that a factor of 1 would have resulted in the maximally saturated chromaticities (preserving equal distance to the achromatic point), while a factor of 0.5 would have resulted in both chromaticities being the same. As a second, fully crossed factor, we set the chromatic standard deviation of the two Gaussians to be 0, 0.01, and 0.02.

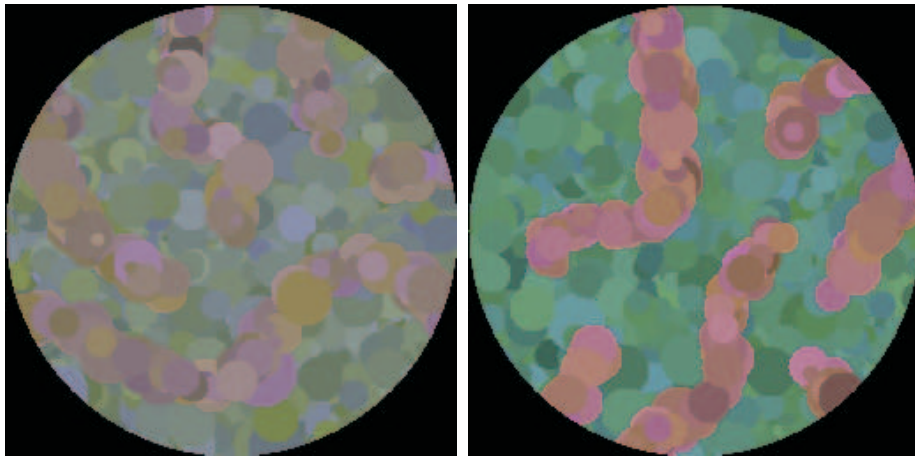


Figure A.9: Examples for stimuli with mean chromatic distances of 0.6 and 0.8 between points inside and outside of the curves (see text). Please note that the reproduction here can differ significantly from the intended coloration.

Figure A.10a shows the results from an individual subject in the described experiment. As is apparent, increasing the chromaticity noise only had a small effect on discrimination performance. This result was unexpected since chromaticity noise renders the areas inside and outside of the curve less homogeneous. Chromatic similarity between the two areas, however, did have a substantial effect on discrimination performance, with the expected pattern of more contrast between the two areas allowing for better discrimination.

As chromatic noise did not seem to have much of an effect, we only varied color similarity in a second experiment, but with a lot more steps than in the first one. Specifically, we employed the adaptive algorithm described above in appendix A.1 but had chromatic similarity as the parameter which was adaptively chosen. Figure A.10b shows the data from an individual subject. The results confirm those of the first pilot study, clearly indicating that discrimination performance increased with chromatic

distance. The figure also shows the relative number of trials that the algorithm had to spent over each section of the parameter range, again indicating that most trials were indeed spent over sections of high volatility of the performance curve.

In order to set boundary conditions for experiment 5 that theoretically allowed for near perfect discrimination, we chose a relative chromatic distance of 0.9 and a chromatic standard deviation of the Gaussians of 0.02.

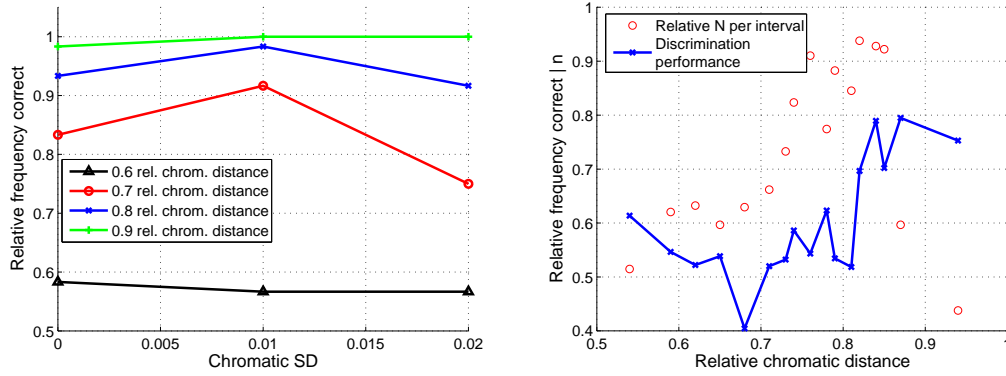


Figure A.10: Results from an individual subject in two preliminary experiments carried out to determine suitable values for the stimulus parameters of chromatic standard deviation and relative chromatic distance between circles inside and outside of the shape.

Appendix B

Tables

B.1 Experiments 1, 2 and 3

Flank	Speed [deg·s ⁻¹]	TR	DW	KJ	LS	CN
Red	0.5	.31 .31 55.1	.46 .33 23.8	.52 .33 18.1	.30 .31 61.5	.30 .31 63.3
	1.5	.34 .32 45.6	.48 .33 22.1	.52 .33 18.9	.31 .31 55.3	.30 .31 60.8
	2.5	.38 .32 33.5	.50 .33 21.0	.52 .33 18.3	.35 .32 41.0	.32 .31 53.7
Green	0.5	.29 .33 65.3	.29 .53 12.5	.29 .48 8.9	.29 .34 65.0	.29 .32 65.6
	1.5	.29 .35 63.3	.29 .55 11.1	.29 .48 9.6	.29 .33 65.2	.29 .32 63.2
	2.5	.29 .38 59.8	.29 .55 13.9	.29 .50 9.4	.29 .34 64.8	.29 .33 58.7

Table B.1: Mean color matches (CIE xyL) in experiment 1 for red and green flanks.

Gap size [min of arc]	TR	DW	KJ	LS	CN
0	.40 .32 30.1	.52 .34 17.8	.59 .34 10.7	.31 .31 57.1	.31 .31 59.7
1.36	.32 .31 50.4	.35 .32 41.7	.47 .33 23.0	.30 .31 63.3	.30 .31 63.4
2.73	.30 .31 60.2	.31 .31 59.0	.46 .33 23.0	.30 .31 64.0	.30 .31 60.8

Table B.2: Mean color matches (CIE xyL) in experiment 2.

Spacing [deg]	TR	DW	KJ	LS	CN
0.89	.46 .33 23.9	.36 .32 38.8	.50 .33 16.4	.33 .31 46.9	.32 .31 46.8
1.43	.30 .31 61.8	.30 .31 65.6	.60 .34 8.8	.30 .31 65.8	.30 .31 63.5
1.84	.30 .31 65.9	.29 .31 67.4	.60 .34 7.8	.29 .31 66.8	.29 .31 66.1

Table B.3: Mean color matches (CIE xyL) in experiment 3.

B.2 Experiments 4a and b

Subject	Speed [deg · s ⁻¹]	Distance [deg]	Delay [ms]	SE distance [deg]	SE delay [ms]
TP	1.69	0.31	183	0.032	18.97
	3.37	0.31	93	0.016	4.71
	5.06	0.37	73	0.021	4.10
	6.75	0.41	61	0.041	6.11
	8.44	0.61	72	0.034	4.05
TH	1.69	0.22	129	0.005	2.85
	3.37	0.30	89	0.008	2.23
	5.06	0.42	84	0.022	4.39
	6.75	0.57	84	0.029	4.27
	8.44	0.71	84	0.037	4.36
WM	1.69	0.28	164	0.012	6.90
	3.37	0.38	111	0.035	10.52
	5.06	0.54	107	0.064	12.56
	6.75	0.66	97	0.055	8.13
	8.44	0.70	823	0.070	8.26
RH	1.69	0.20	120	0.005	2.81
	3.37	0.25	73	0.008	2.32
	5.06	0.36	71	0.031	6.14
	6.75	0.40	59	0.030	4.37
	8.44	0.47	56	0.040	4.81
IG	1.69	0.67	400	0.037	21.95
	3.37	0.59	176	0.041	12.10
	5.06	0.59	117	0.026	5.05
	6.75	0.61	90	0.043	6.36
	8.44	0.76	90	0.022	2.61
CA	1.69	0.29	173	0.006	3.44
	3.37	0.42	125	0.014	4.19
	5.06	0.68	134	0.060	11.93
	6.75	0.88	130	0.094	13.90
	8.44	0.87	103	0.048	5.64
TG	3.37	0.62	183	0.028	8.21
	5.06	0.82	162	0.068	13.40
	6.75	1.02	151	0.066	9.73
	8.44	1.07	127	0.069	8.14
	10.12	1.06	105	0.039	3.81
KR	3.37	0.69	205	0.038	11.37
	5.06	0.79	156	0.030	5.99
	6.75	0.84	124	0.040	5.92
	8.44	0.94	112	0.037	4.41

	10.12	0.93	92	0.026	2.61
MH	3.37	0.56	167	0.033	9.74
	5.06	0.67	132	0.038	7.58
	6.75	0.88	131	0.029	4.35
	8.44	0.94	112	0.024	2.87
	10.12	0.90	89	0.041	4.05
DW	3.37	0.33	97	0.007	2.04
	5.06	0.53	104	0.017	3.33
	6.75	0.72	107	0.031	4.61
	8.44	0.75	89	0.055	6.49
	10.12	0.88	87	0.042	4.17

Table B.4: Mean adjusted line distances for the individual subjects in experiment 4a.

Subject	Speed [deg · s ⁻¹]	Distance [deg]	Delay [ms]	SE speed [deg]	SE delay [ms]
TP	2.19	0.27	141	0.36	12.60
	5.74	0.54	99	0.37	8.00
	10.12	0.81	83	0.62	5.47
	13.67	1.08	83	1.05	5.57
	21.09	1.35	68	1.65	6.32
TH	3.38	0.27	80	0.00	0.00
	8.10	0.54	67	0.22	2.13
	9.28	0.81	88	0.38	3.25
	11.64	1.08	94	0.39	3.17
	13.16	1.35	104	0.49	3.46
WM	3.20	0.27	88	0.17	8.00
	5.74	0.54	96	0.28	4.35
	7.09	0.81	117	0.34	5.99
	10.63	1.08	109	1.01	9.47
	21.77	1.35	74	3.20	10.29
RH	4.56	0.27	68	0.67	6.51
	7.42	0.54	75	0.52	3.62
	8.10	0.81	102	0.34	4.31
	7.93	1.08	142	0.51	10.19
	9.28	1.35	150	0.52	8.24
IG	1.69	0.27	160	0.000	0.00
	2.70	0.54	273	0.720	31.45
	4.56	0.81	228	0.619	44.24
	6.07	1.08	199	0.675	23.10
	9.62	1.35	143	0.439	7.61
CA	10.29	0.27	87	2.873	24.28
	5.40	0.54	136	1.665	14.03

Appendix B Tables

	16.71	0.81	66	2.256	15.79
	12.66	1.08	122	2.108	27.76
	45.39	1.35	83	10.438	37.45
TG	3.37	0.34	100	0.000	0.00
	4.39	0.49	120	0.373	9.25
	4.89	0.65	142	0.394	11.96
	6.41	0.81	130	0.337	7.04
	7.09	0.97	141	0.421	9.32
KR	1.86	0.34	190	0.169	10.00
	3.54	0.49	171	0.467	27.51
	4.39	0.65	168	0.373	25.78
	6.75	0.81	122	0.252	4.89
	7.26	0.97	139	0.439	9.69
MH	3.54	0.34	97	0.169	3.33
	5.06	0.49	98	0.000	0.00
	5.74	0.65	116	0.276	5.26
	6.92	0.81	119	0.303	5.52
	10.46	0.97	99	0.968	8.18
DW	3.88	0.34	90	0.258	5.09
	5.23	0.49	95	0.169	2.44
	5.57	0.65	119	0.258	4.92
	6.75	0.81	122	0.252	4.89
	7.42	0.97	132	0.276	4.68

Table B.5: Mean adjusted line speeds for the individual subjects in experiment 4b.

B.3 Experiment 5

Subject TG	Speed [deg · s ⁻¹]	Line spacing [deg]	Trials in interval	Relative frequency correct
	5.32	0.32	40	0.98
		0.45	45	1.00
		0.58	59	0.93
		0.72	90	0.83
		0.74	53	0.77
		0.82	43	0.70
		0.88	43	0.65
		0.98	42	0.67
		1.20	63	0.60
	2.66	0.32	40	1.00
		0.45	41	0.98
		0.58	40	0.90

B.3 Experiment 5

		0.72	40	0.83
		0.85	44	0.77
		0.98	49	0.71
		1.20	60	0.62
	1.77	0.40	40	0.98
		0.66	52	0.92
		0.93	80	0.84
		0.98	52	0.81
		1.12	65	0.71
		1.25	42	0.67
	1.33	0.40	40	0.98
		0.66	55	0.95
		0.72	40	0.93
		0.85	51	0.86
		0.98	72	0.76
		1.12	45	0.67
		1.25	41	0.63

Table B.6: Mean results from experiment 5 for subject TG.

Subject MH	Speed [deg · s ⁻¹]	Line spacing [deg]	Trials in interval	Relative frequency correct
	5.32	0.29	40	1.00
		0.35	49	0.98
		0.43	40	0.93
		0.45	47	0.87
		0.50	65	0.72
		0.53	42	0.64
		0.56	40	0.73
		0.61	48	0.60
		0.72	41	0.54
		0.93	65	0.58
		1.20	51	0.53
	2.66	0.32	49	0.98
		0.45	67	0.88
		0.48	42	0.83
		0.56	41	0.80
		0.61	51	0.71
		0.69	40	0.68
		0.74	42	0.60
		0.82	40	0.50
		0.88	41	0.46
		0.98	80	0.39

Appendix B Tables

		1.20	45	0.42
	1.77	0.32	70	1.00
		0.45	68	0.91
		0.48	48	0.88
		0.56	59	0.83
		0.61	64	0.73
		0.69	40	0.70
		0.74	51	0.67
		0.82	47	0.62
		0.88	56	0.52
		0.98	84	0.46
		1.20	43	0.58
	1.33	0.32	48	1.00
		0.45	51	0.94
		0.58	58	0.86
		0.72	70	0.79
		0.85	72	0.69
		0.88	51	0.71
		0.90	40	0.63
		0.93	47	0.57
		0.96	40	0.53
		1.01	56	0.52
		1.12	56	0.59
		1.14	41	0.56
		1.22	45	0.47
		1.28	42	0.48

Table B.7: Mean results from experiment 5 for subject MH.

Subject KR	Speed [deg · s ⁻¹]	Line spacing [deg]	Trials in interval	Relative frequency correct
	5.32	0.32	79	1.00
		0.45	75	0.86
		0.50	47	0.87
		0.56	41	0.83
		0.61	66	0.79
		0.69	48	0.71
		0.77	45	0.69
		0.85	58	0.57
		0.98	46	0.52
		1.20	62	0.44
	2.66	0.40	64	0.97
		0.66	63	0.89

B.3 Experiment 5

		0.93	109	0.79
		0.98	74	0.74
		1.12	41	0.66
		1.25	40	0.60
	1.77	0.40	70	0.96
		0.66	49	0.86
		0.93	57	0.81
		0.98	44	0.77
		1.12	64	0.70
		1.25	94	0.61
	1.33	0.40	76	1.00
		0.66	81	0.90
		0.72	44	0.82
		0.85	53	0.77
		0.98	43	0.77
		1.12	53	0.70
		1.25	78	0.60

Table B.8: Mean results from experiment 5 for subject KR.

Subject JJ	Speed [deg · s ⁻¹]	Line spacing [deg]	Trials in interval	Relative frequency correct
	5.32	0.40	40	1.00
		0.66	67	0.96
		0.74	41	0.93
		0.82	43	0.86
		0.88	51	0.76
		0.96	40	0.70
		1.01	41	0.63
		1.12	43	0.58
		1.25	40	0.55
	2.66	0.40	40	1.00
		0.66	65	0.92
		0.85	69	0.81
		0.98	70	0.74
		1.01	47	0.79
		1.04	40	0.85
		1.06	42	0.67
		1.09	40	0.55
		1.14	99	0.61
		1.25	68	0.51
	1.77	0.40	40	0.98
		0.66	58	0.90

Appendix B Tables

		0.72	40	0.88
		0.85	41	0.78
		0.98	57	0.68
		1.12	71	0.59
		1.25	45	0.53
	1.33	0.40	40	1.00
		0.66	74	0.95
		0.72	52	0.92
		0.74	43	0.91
		0.82	48	0.81
		0.88	48	0.73
		0.98	62	0.68
		1.12	60	0.60
		1.14	40	0.58
		1.22	42	0.48
		1.28	42	0.38

Table B.9: Mean results from experiment 5 for subject JJ.

Subject DWI	Speed [deg · s ⁻¹]	Line spacing [deg]	Trials in interval	Relative frequency correct
	5.32	0.32	46	1.00
		0.45	61	0.90
		0.50	40	0.88
		0.56	41	0.83
		0.61	45	0.73
		0.69	53	0.64
		0.77	45	0.60
		0.93	52	0.62
		1.20	47	0.53
	2.66	0.32	41	0.98
		0.45	58	0.97
		0.50	43	0.95
		0.56	61	0.87
		0.58	41	0.83
		0.61	40	0.68
		0.64	40	0.75
		0.72	74	0.78
		0.77	40	0.75
		0.82	42	0.69
		0.85	41	0.54
		0.88	40	0.50
		0.93	40	0.53

		0.98	71	0.49
		1.20	79	0.57
	1.77	0.32	40	1.00
		0.45	41	1.00
		0.58	62	0.95
		0.66	41	0.85
		0.72	40	0.70
		0.77	56	0.68
		0.82	42	0.64
		0.88	40	0.63
		0.98	43	0.53
		1.20	68	0.56
	1.33	0.32	46	1.00
		0.45	40	1.00
		0.58	54	0.93
		0.72	68	0.84
		0.82	52	0.77
		0.88	70	0.67
		0.96	47	0.68
		1.01	63	0.59
		1.12	45	0.53
		1.25	40	0.48

Table B.10: Mean results from experiment 5 for subject DWI.

Subject JJ	Line spacing [deg]	Trials in interval	Relative frequency correct
	0.40	76	1.00
	0.66	91	0.90
	0.77	43	0.79
	0.80	47	0.66
	0.85	40	0.63
	0.88	61	0.62
	0.98	59	0.58
	1.01	42	0.60
	1.09	60	0.50
	1.12	42	0.45
	1.14	40	0.33
	1.17	51	0.37
	1.25	43	0.51

Table B.11: Mean results from static control condition of experiment 5 for subject JJ.

Appendix B Tables

Subject TP	Line spacing [deg]	Trials in interval	Relative frequency correct
	0.32	40	0.93
	0.45	42	0.95
	0.58	58	0.88
	0.61	40	0.83
	0.69	45	0.76
	0.74	64	0.66
	0.85	60	0.60
	0.88	43	0.56
	0.96	42	0.50
	1.01	41	0.49
	1.20	62	0.453

Table B.12: Mean results from static control condition of experiment 5 for subject TP.

Subject DWO	Line spacing [deg]	Trials in interval	Relative frequency correct
	0.32	40	0.98
	0.45	63	0.94
	0.58	117	0.84
	0.61	70	0.76
	0.69	40	0.73
	0.74	46	0.65
	0.77	57	0.49
	0.85	66	0.52
	0.98	52	0.60
	1.20	52	0.46

Table B.13: Mean results from static control condition of experiment 5 for subject DWO.

Subject WM	Line spacing [deg]	Trials in interval	Relative frequency correct
	0.32	40	0.98
	0.45	49	0.96
	0.58	73	0.86
	0.61	53	0.83
	0.64	40	0.80
	0.66	49	0.65
	0.69	40	0.75
	0.74	52	0.79

	0.85	47	0.68
	0.98	44	0.66
	1.12	64	0.58
	1.14	47	0.55
	1.17	42	0.57
	1.20	44	0.48
	1.22	40	0.38
	1.28	45	0.38

Table B.14: Mean results from static control condition of experiment 5 for subject WM.

Subject MP	Line spacing [deg]	Trials in interval	Relative frequency correct
	0.32	70	0.99
	0.45	56	0.89
	0.48	40	0.88
	0.56	42	0.79
	0.61	45	0.69
	0.72	55	0.71
	0.85	75	0.61
	0.98	53	0.53
	1.20	42	0.55

Table B.15: Mean results from static control condition of experiment 5 for subject MP.

B.4 Experiment 6

Subject	Slits frequency [slits · deg ⁻¹]	Gabor frequency [cycles · deg ⁻¹]							
		0.69	1.05	1.39	1.74	2.09	2.79	3.48	4.18
DW	4.53	0.11	0.06	0.08	0.12	0.15	0.12	0.04	0.10
	3.83	0.13	0.06	0.06	0.09	0.10	0.09	0.08	0.14
	3.14	0.12	0.09	0.08	0.10	0.12	0.08	0.11	0.15
	2.44	0.20	0.13	0.10	0.13	0.08	0.23	0.11	0.21
VE	4.53	0.24	0.15	0.10	0.11	0.09	0.07	0.07	0.05
	3.83	0.29	0.20	0.17	0.10	0.14	0.11	0.11	0.10
	3.14	0.25	0.15	0.19	0.25	0.13	0.07	0.06	0.11
	2.44	0.22	0.23	0.17	0.10	0.15	0.09	0.08	0.09
FF	4.53	0.09	0.04	0.03	0.03	0.04	0.06	0.12	0.09
	3.83	0.10	0.04	0.03	0.03	0.06	0.11	0.07	0.05
	3.14	0.06	0.07	0.04	0.03	0.11	0.06	0.06	0.09

Appendix B Tables

	2.44	0.05	0.06	0.09	0.07	0.07	0.10	0.15	0.15
GW	4.53	0.14	0.06	0.08	0.09	0.11	0.14	0.08	0.08
	3.83	0.15	0.13	0.10	0.13	0.12	0.16	0.08	0.14
	3.14	0.23	0.19	0.12	0.17	0.06	0.10	0.19	0.32
	2.44	0.21	0.16	0.18	0.14	0.17	0.13	0.34	0.50
TP	4.53	0.15	0.12	0.09	0.10	0.16	0.18	0.09	0.11
	3.83	0.12	0.12	0.14	0.14	0.17	0.19	0.08	0.10
	3.14	0.11	0.09	0.11	0.18	0.17	0.25	0.19	
	2.44	0.17	0.17	0.25	0.41	0.14			

Table B.16: Mean relative differences between subjects' settings and test Gabor frequency in experiment 6. Empty table cells indicate that all trials in this condition were flagged as invalid.

Subject	Slit frequency [slits · deg ⁻¹]	Gabor frequency [cycles · deg ⁻¹]							
		0.69	1.05	1.39	1.74	2.09	2.79	3.48	4.18
DW	4.53	0.82	0.43	0.44	0.96	1.28	1.68	1.52	1.65
VE	4.53	1.00	0.78	0.34	0.65	0.83	1.08	1.51	1.35
FF	4.53	0.25	0.33	0.16	0.17	0.14	0.29	0.43	0.49
GW	4.53	1.52	0.45	0.56	0.34	0.43	0.53	1.02	1.18
TP	4.53	0.80	1.28	0.93	1.19	1.19	1.31	1.13	1.65

Table B.17: Mean relative differences between subjects' settings and test Gabor frequency in the static control condition of experiment 6.

Appendix C

Abstract

The main subject of my dissertation is the dynamic completion of homogeneous as well as non-uniform surface qualities, i.e., color and texture. Dynamic surface completion can be seen as a grouping phenomenon incorporating the interaction of many different perceptual codes: Among them are color, motion, depth, and shape.

To place dynamic surface completion within the appropriate research context, the first part of my thesis is concerned with an analysis of the general problem of spatio-temporal input fragmentation. A review of relevant completion phenomena follows, including static and dynamic examples, as well as instances of boundary and surface feature completion.

Since the heterogeneous category of visual completion subsumes a multitude of different effects, special emphasis was placed on providing a structured account of them. Moreover, the review pays attention to functional interactions between the coupled representations of color, depth, and shape. It finally bears on the question how the phenomena may be tied together in a common theoretical picture by analyzing their shared underlying processing aspects. It ends by pointing out loose ends in the current state of completion research, some of which are picked up in the empirical part.

The empirical part with six experiments on dynamic color and texture spreading addresses two main open questions: To what extent can low-level processes like local spatiotemporal integration through visual persistence account for the effects? What is the adequate methodology to approach visual completion in general, given the often peculiar nature of their appearance? To this end, I developed a new flexible framework based on multi-aperture viewing in order to generate dynamic stimuli which trigger strong surface completion, including color and texture spreading. In addition, I devised a new objective performance task designed to isolate aspects of surface completion from local boundary integration, which is at the heart of most current completion research.

My results indicate that visual persistence can account for some patterns indicating a continuous variability of completed percepts, but fail to cover qualitative perceptual changes and subtle figural influences. The new adaptive shape discrimination task proved to be useful tool in the study of reliably measure the functional effects of surface completion.

Appendix D

Zusammenfassung

Ausgangspunkt dieser Arbeit sind zwei sich entgegenstehende Beobachtungen: Einerseits ist das von Objekten reflektierte Licht auf seinem Weg zum Auge zahlreichen raumzeitlichen Einflüssen ausgesetzt, die dazu führen, daß physikalisch abgegrenzte Regionen des retinalen Bildes nicht mit der Objektstruktur der Außenwelt korrespondieren: Insbesondere sorgt die zeitlich variable Verdeckung von Objekten dafür, daß das zu *einem* Objekt gehörende Lichtsignal auf mehrere, raumzeitlich unzusammenhängende Regionen des retinalen Bildes verteilt sein kann. Andererseits ist unsere phänomenale Wahrnehmung der Außenwelt keineswegs dadurch charakterisiert, daß Objekte unvollständig und zerrissen erscheinen. Vielmehr nehmen wir Objekte auch dann als vollständig und zusammenhängend wahr, wenn große Teile von ihnen verdeckt und deswegen unsichtbar sind.

Das visuelle System überwindet also die starke Fragmentierung des Inputs mit dem Ergebnis eines durch zusammenhängende Oberflächen und Objekte charakterisierten Perzepts einer Szene. Dies beinhaltet, daß eigentlich zusammengehörige Fragmente des Inputs offenbar trotz ihrer räumlichen und zeitlichen Dissoziation einander zugeordnet und gruppiert werden können. Weiterhin wird bei näherer Betrachtung einer Reihe von Befunden deutlich, daß die etwa durch Verdeckung unwiederbringlich verlorene Information über ein Objekt in gewisser Hinsicht vom Wahrnehmungssystem aktiv und phänomenal sichtbar vervollständigt wird. Der Fragestellung meiner Arbeit liegen nun im wesentlichen folgende Fragen zugrunde:

- Welche Objektqualitäten außer der Form eines Objekts können von einer solchen perzeptuellen Vervollständigung betroffen sein? Insbesondere betrifft dies Oberflächenaspekte wie Farbe, Textur und Tiefenzuordnung eines Objekts.
- Ist die perzeptuelle Vervollständigung von Objekten rein phänomenologischer Natur, d.h. ein von von funktionalen Aspekten der visuellen Verarbeitung abgekoppeltes Epiphänomen? Oder läßt sich zeigen, daß die Vervollständigung von Objekten einen Beitrag zu leistungsbezogenen Wahrnehmungsaspekten liefert?
- Inwieweit lassen sich Phänomene der Vervollständigung von Oberflächeneigenschaften durch einen vergleichsweise einfachen Mechanismus raumzeitlicher Summation mittels sensorischer Persistenz erklären?

- Welche methodischen Zugänge sind geeignet, um sich der perzeptuellen Vervollständigung verschiedener Objekteigenschaften jenseits einer rein introspektiven und subjektiven Herangehensweise zu nähern?

Im theoretischen Teil konzentriert sich die Arbeit zunächst darauf, den notwendigen Kontext zur Einordnung der untersuchten Sachverhalte zu stellen. Zu Beginn steht daher ein Überblick der vielfältigen Einflüsse, die zur Fragmentierung des visuellen Inputs führen. Dabei werden räumliche wie auch zeitlich wirkende Aspekte berücksichtigt. Der Schwerpunkt des theoretischen Teils liegt in einer integrativen, auf Strukturierung bedachten Darstellung der breitgefächerten Phänomene visueller Vervollständigung. Die Darstellung wird geleitet durch die wiederholte kritische Auseinandersetzung mit theoretischen Modellen der perzeptuellen Vervollständigung, die auf der zweidimensionalen Geometrie des retinalen Bildes basieren.

Die bisherige experimentelle Untersuchung visueller Vervollständigung hat sich inhaltlich überwiegend auf den Aspekt des Umrisses eines Objekts und damit seiner Form konzentriert. Dabei wurden meist nur statische Reizkonstellationen berücksichtigt. Abweichend davon liegt in meiner Arbeit der Schwerpunkt auf der Vervollständigung von *Oberflächeneigenschaften* in *bewegten* Reizen und berücksichtigt dabei vor allem die Aspekte Farbe und Textur.

Neben der inhaltlichen Ausrichtung auf die gemeinsame Vervollständigung von Oberflächenattributen bildet die Entwicklung einer angemessenen Methodik zur Untersuchung von Vervollständigungsphänomenen einen zweiten Schwerpunkt der Arbeit. Ergebnis eines Überblicks über etablierter experimenteller Herangehensweisen ist die Feststellung, daß die bisher meist verwendeten introspektiven und auf subjektive Schätzurteile abzielenden Methoden eine hohe Anfälligkeit für Verzerrungen besitzen.

Ausgehend vom letztgenannten Ergebnis wird in der Arbeit eine neue, dynamische Variante eines bekannten Untersuchungsparadigmas vorgestellt, die zusammen mit einer neuen Reizklasse entwickelt wurde. Ziel des Verfahrens ist es insbesondere, die funktionale Relevanz der Vervollständigung von Oberflächenattributen in unverzerrter Form messbar zu machen. Pilotstudien und numerische Simulationen zur genauen Festlegung und Validierung des Verfahrens sind dabei seinem experimentellen Einsatz vorausgegangen.

Ergebnis der insgesamt sechs vorgestellten Experimente ist zunächst, daß das Phänomen der subjektiven Vervollständigung von Oberflächeneigenschaften in der Tat einer vergleichsweise strengen Methodik zugänglich ist. Nachdem so die Angemessenheit des gewählten methodischen Zugangs festgestellt wurde, wurde als weiteres Ergebnis erzielt, daß einige Aspekte des zu untersuchenden Phänomens mit der Annahme einfacher sensorischer Persistenz als kausalem Mechanismus vereinbar sind. Andererseits stellt sich jedoch gleichzeitig heraus, daß dieser Mechanismus andere, qualitative Ergebnisse nicht zu erklären vermag. Schließlich ist es gelungen, funktional relevante Vorteile raumzeitlicher Vervollständigung in einer Weise messbar zu machen, die über subjektive Selbstauskunft hinausgeht.

Erklärung

Hiermit erkläre ich, Daniel Wollschläger, dass die Dissertation “Dynamic surface completion: The joint formation of color, texture, and shape” abgesehen von der Beratung meines Betreuers, Herrn Prof. Dr. R. Mausfeld, in Inhalt und Form meine eigene Arbeit ist. Diese Arbeit war bisher weder teilweise noch ganz Gegenstand eines Prüfungsverfahrens. Teile der in dieser Arbeit berichteten Daten wurden bereits veröffentlicht oder sind zur Veröffentlichung angenommen. Planung und Durchführung der zugrundeliegenden Experimente lagen ausschließlich, die Abfassung der zugehörigen Artikel weitestgehend in meiner Verantwortung:

Wollschläger, Rodriguez und Hoffman. 2002. *Flank transparency: The effects of gaps, line spacing, and apparent motion*. **Perception** 31(9), 1073-1092.

Wollschläger und Faul. Im Druck. *Dynamic texture spreading: Probing the mechanisms of surface interpolation*. **Spatial Vision**.

(Daniel Wollschläger)

Lebenslauf (Curriculum Vitae)

Persönliche Daten

- Name: Daniel Wollschläger
- Geburtsdatum: 19. Juni 1976
- Geburtsort: Göttingen, Niedersachsen
- Familienstand: Ledig
- Nationalität: Deutsch
- Adresse: Telgter Str. 38, 33619 Bielefeld

Schulbildung

- 1982–1986: Grundschule Dornberg, Bielefeld
- 1986–1995: Ratsgymnasium Bielefeld
- Juni 1995: Erlangung der allgemeinen Hochschulreife (Abitur)

Hochschulstudium

- 1996–2002: Christian-Albrechts-Universität zu Kiel, Diplomstudiengang Psychologie (12 Studiensemester)
- Oktober 1998: Vordiplom
- 2000–2001: University of California at Irvine, USA, Graduate Studies of Cognitive Sciences (2 Studiensemester)
- September 2002: Diplom
- Oktober 2002–Dezember 2005: Promotionsstudium, betreut von Prof. Dr. R. Mausfeld, Institut für Psychologie, Christian-Albrechts-Universität zu Kiel (7 Studiensemester)