Vision Research 48 (2008) 2143–2153

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Ambiguities and conventions in the perception of visual art

Pascal Mamassian *

CNRS & Université Paris Descartes, Laboratoire Psychologie de la Perception (UMR 8158), 45 rue des Saints-Pères, 75006 Paris, France

ARTICLE INFO

Article history:
Received 26 January 2008
Received in revised form 2 June 2008

Keywords:
Visual perception
Visual arts
Ambiguous perception
Visual constraints
Conventions

ABSTRACT

Vision perception is ambiguous and visual arts play with these ambiguities. While perceptual ambiguities are resolved with prior constraints, artistic ambiguities are resolved by conventions. Is there a relationship between priors and conventions? This review surveys recent work related to these ambiguities in composition, spatial scale, illumination and color, three-dimensional layout, shape, and movement. While most conventions seem to have their roots in perceptual constraints, those conventions that differ from priors may help us appreciate how visual arts differ from everyday perception.

I believe it is only by considering these psychological aspects of image making and image reading that we may come closer to an understanding of the central problem of the history of art [...], that is, why representation should have a history; why it should have taken mankind so long to arrive at a plausible rendering of visual effects that create the illusion of life-likeness; and why artists such as John Constable, who strove to be true to his vision, still had to admit that no art is ever free of convention.

Sir Ernst Gombrich (1960), p. 291
Art and Illusion

1. Introduction

Every vision scientist has an opinion about visual arts, apparently all of them different, and these opinions are frequently put forward in lectures and books. Most visual artists have a reciprocal interest in visual perception, although, in contrast, they rarely tour universities offering lessons on visual perception. Yet both vision scientists and artists have in common the passion to uncover the mysteries of visual perception, either to understand its mechanisms or to better communicate a message. While there are still relatively few scientific papers on the perception of visual arts, there is a growing number of books preaching that scientists and artists should look more often at each other's work (Gregory & Harris, 1995; Hecht, Schwartz, & Atherton, 2003; Livingstone, 2002; Maffei & Fiorentini, 1995; Zeki, 1999). Several of these books grew out of the enthusiasm that has followed the rapid advances in knowledge about how humans perceive, especially from the point of view of visual neuroscience. Even though some of this knowledge is likely to be necessary for a better understanding of the perception of visual arts, it is clearly not sufficient. The question of interest, in personal terms, is what is different when I look at an artist's painting compared to when I look at my kitchen appliances. Knowing that the same mechanisms are at work when I decipher the shape and color of objects in a Turner painting and in my microwave oven is not a satisfyingly complete answer.

One important difference between the perception of visual arts and what could be called "everyday perception" is the task of the observer. Vision scientists know very well the importance of specifying precisely the task of their participants, even when it is as simple as detecting or discriminating a small grey patch (Watson & Robson, 1981). In everyday perception, the task of the observer is well defined, often by the action that the perception supports. As we watch the incoming traffic before crossing the road, our perception of the traffic is oriented to the extraction of useful information such as the recognition of a car and the estimation of its speed, while at the same time disregarding irrelevant information such as the make or color of the car. Once the task is established, one can define the decisions necessary to perform it, and if one so wishes, the efficiency of the observer in this task can be computed by normalizing the performance to that of the ideal observer for this task (Barlow, 1962). Moreover, because these everyday perceptions are linked to a decision, we can describe the cost of making the wrong decision and thus study perception under risk (Landy, Goutcher, Trommershäuser, & Mamassian, 2007). It is more difficult to identify an appropriate task in the perception of visual arts, or indeed to agree that there is one. Without specifying a task, the question of

* Fax: +33 0 1 42 86 33 22.
E-mail address: pascal.mamassian@univ-paris5.fr

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.06.010

how good one is at looking at a painting becomes irrelevant, and the notion of risk associated to an alleged wrong perception becomes meaningless.

One way to identify a plausible task in visual arts perception is to return to the challenges of everyday perception. One major issue is that sensory information is inherently ambiguous, a single image being consistent with an infinite number of three-dimensional scenes. To resolve the ambiguities of the retinal images, one can approach visual perception as an inference problem (Gregory, 1980; Rock, 1983): what is the most likely scene in front of the observer given the retinal information? In its contemporary formulation, perceptual inferences are cast within a Bayesian framework, and the inference problem is solved thanks to additional knowledge referred to prior constraints (Kersten, Mamassian, & Yuille, 2004). For instance, the luminance gradient on an object provides ambiguous information about its shape but a prior knowledge that the light source is located above the observer will resolve the shape ambiguity.

Is visual arts perception subject to ambiguities similar to that faced in everyday perception? We shall argue that ambiguity reduction is indeed an important task of visual arts perception. The importance of ambiguity in the perception of art has been appreciated before, for instance by Semir Zeki: "Great art is that which corresponds to as many different concepts in as many different brains over as long a period of time as possible. Ambiguity is such a prized characteristic of all great art because it can correspond to many different concepts." (Zeki, 2002, p. 67). Everyday perception and visual arts perception seem to share the basic task of resolving ambiguities, but in ways that may be very different. While everyday perception uses prior knowledge, visual arts use conventions. These conventions may take their inspiration from visual prior knowledge but they can also be completely arbitrary. Take for instance the convention for the depiction of the human figure in ancient Egyptian art: the head, arms and legs are shown in a profile view, whereas the shoulders and torso face forward. In so doing, Egyptian artists violated the constraints of possible human contortions, maybe for the purpose of removing the ambiguity of the person's identity by displaying each body part in its most informative pose. Because artistic conventions are different across periods, across cultures, and across human development, the appreciation of art will necessarily vary from one person to the next.

The present review will emphasize the nature of the ambiguities common in everyday and art perception, and compare the conventions used in visual arts with the prior knowledge used in everyday perception. In the interest of presenting a reasonable size and coherent review, the scope will be restricted to topics directly linked to visual perception and the focus will be on research done in the last decade or so. To focus on perception, this review will not discuss more cognitive aspects of picture understanding (Solso, 1994; Willats, 1997) or the emotional responses to art (recently reviewed by Silvia, 2005). The review will also be restricted to behavioral studies of perception, and so will not address the emerging field of neuroesthetics that concerns the search for the neural substrates of aesthetics experience (Kawabata & Zeki, 2004; Solso, 2001). Finally, because they merit reviews on their own, the review will not cover two topics that could find their place here, that of face perception (see the book by Bruce and Young (1998) and the recent review by Rhodes (2006)) and that of eye-movements (recently discussed by Conway & Livingstone, 2007).

2. Composition

Composition refers to the way objects are arranged relative to each other and relative to the frame. Composition serves multiple purposes, from improving the harmony of the painting to helping understand the scene. Harmony is influenced by the way objects in the painting relate to each other, how similar repetitive patterns group together and how some figures segregate from the background. Which objects group together and which part of the image constitute the background are ambiguous relationships that Gestalt psychologists have helped disentangle with their principles of perceptual organization. For instance, the preference to see objects of similar orientations group together may be related to some statistical regularities of our environment (Elder & Goldberg, 2002; Geisler, Perry, Super, & Gallogly, 2001).

Are these visual prior expectations also conventions in visual arts? In a very influential book, Arnheim (1954) took these Gestalt principles and searched for correspondences in the visual arts. He argued that such principles were indeed present in paintings, and he emphasized several principles such as the power of balance where the depicted objects are distributed around the canvas, for instance by using symmetrical arrangements. Pinna (2007) is currently pursuing this Gestalt view of visual arts perception. However, other research groups have addressed the issue of whether compositional balance is really critical to distinguish a masterwork from other more regular paintings and they have found mixed results (Locher, 2003; Vartanian, Martindale, Podsiadlo, Overbay, & Borkum, 2005). In this kind of studies, there is always the danger that the observer already knows the painting, or at least recognizes the style of the painter, so that rating judgments can be related to familiarity rather than anything else. In an attempt to objectify and quantify the importance of composition, several experimenters have thus turned towards modern art. The advantage of using modern paintings such as the tri-color compositions of Piet Mondrian is that one can easily manipulate several dimensions independently. By changing the size, location and color of individual rectangles in a Mondrian-like stimulus, some combinations are systematically perceived more balanced (Locher, Overbeeke, & Stappers, 2005), but generic and robust principles of good composition are difficult to extract.

Another aspect of composition is to help the observer look for meaningful information in the canvas. The ambiguity here lies in the multiple locations in the paintings where attention can be caught. From a visual point of view, it appears that salient points in the image, defined primarily by low-level features such as high-contrast edges, will *a priori* be the best candidates for visual attention (Itti & Koch, 2001). These visual prior locations appear to be very different from the compositional conventions encountered in visual arts, where some locations in the canvas have a privileged status. One first example of a location with special status was described by Tyler (1998, 2007a) in a study of Western art portraits. Based on the analysis of numerous portraits covering over four centuries, Tyler noticed a high likelihood that one eye of the portrait fell near the vertical midline. Whether this predominance is merely the result of geometrical constraints and whether there is an aesthetic advantage for those portraits following this principle are issues still being debated (McManus & Thomas, 2007; Tyler, 2007b).

A second example of a location in the canvas that has a special status can be found in paintings that use one-point perspective. In these paintings, there is a single vanishing point on the canvas that corresponds to the far end of a set of parallel lines in space whose orientation is roughly orthogonal to the canvas. The location of the vanishing point appears to be the place to put an object of major importance. For instance, in "The School of Athens" by Raphael (1510–1511) that displays more than twenty famous Greek philosophers, the vanishing point lies intentionally in between the two main characters, on the book Timaeus that Plato on the left gives to his student Aristotle on the right (Fig. 1a). There might be a good technical reason to attract the attention of the observer to the van-

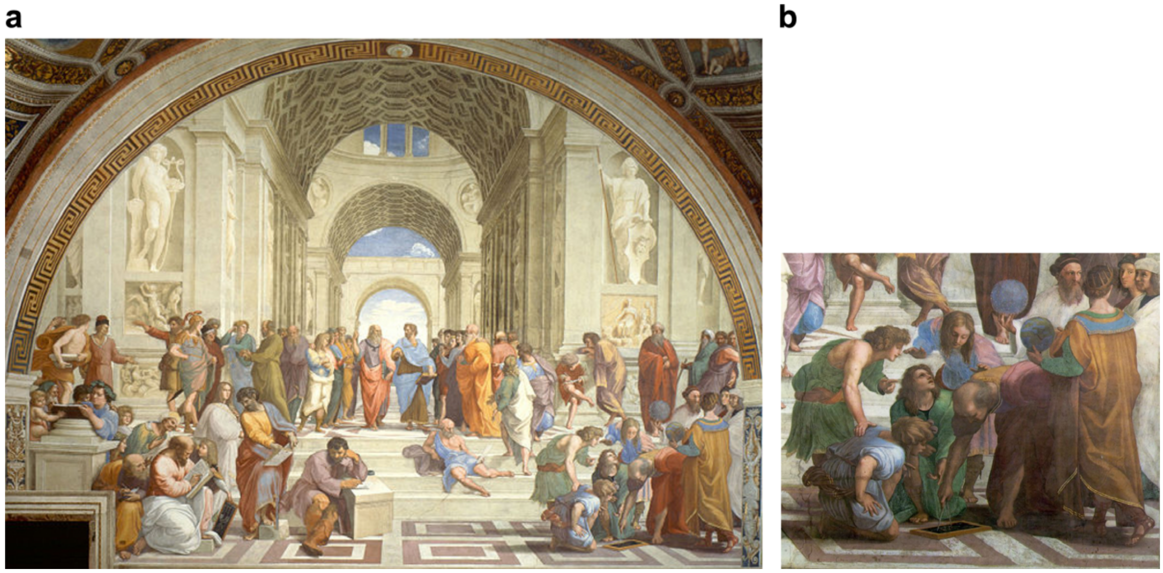

Fig. 1. (a) Illustration of one-point perspective. The two most important characters, Plato and Aristotle are placed next to the vanishing point. Raphael (1510–1511) "The School of Athens". Fresco, Palazzi Pontifici, Vatican. (b) Anomalous perspective for spheres. The whole scene is carefully following the laws of linear perspective, except for these two spheres that should have an oblong form in the image. Bottom right detail of the same fresco.

ishing point. Especially for large paintings or frescoes, the observer would then move in front of the vanishing point, and thus get closer to the center of projection where the scene can be perceived without any three-dimensional deformation (see section on *Three-dimensional layout* below). There are of course alternative explanations for placing objects of importance near the vanishing point. One alternative explanation is that, because of a well-known optical illusion, objects tend to appear bigger near the vanishing point, so this placement will artificially magnify the size of the objects. Another alternative explanation is that the vanishing point tends to be placed close to the center of the painting, so the power of the center might be sufficient to justify the placement of important objects at that location.

In summary on composition, the harmony of multiple objects in visual arts may have similar characteristics as the Gestalt principles of organization, but the placement of individual objects of importance in a painting appears to follow conventions that are quite different from everyday visual expectations.

3. Spatial scale

While composing a painting, the issue arises as to how best represent both large and small objects. This scaling problem is exacerbated when observers move closer or further away, thereby revealing or discarding different levels of details. The impact of the viewing distance is best characterized with the spatial frequency description of the painting. The high spatial frequency range represent the fine details of the image, whereas the low spatial frequencies are associated to the coarse information. Getting closer to a painting will therefore reveal high spatial frequencies that were before outside the range of frequencies the human visual system is sensitive to (typically represented by the contrast sensitivity function). Another action that will reveal high spatial frequencies is to make an eye movement towards a region of interest. Because spatial resolution is highest in the fovea and gradually decreases as one gets to the periphery, some fine details will only be visible when the observer is looking straight at that part of the image.

Because the painter does not know exactly where the observer will stand and where she will look, there is an ambiguity on the spatial frequencies of the painting that will actually be seen by the observer. In natural scenes, all spatial frequencies are available with a distribution that is independent of the observer's position. This distribution is well described as "1/f", that is the power spectral density is inversely proportional to the spatial frequency (Field, 1987). In other words, a natural image is largely dominated by low spatial frequencies (coarse blobs), and higher frequencies (finer details) are gradually less present. It is likely that the human visual system has adapted to these statistics of the natural environment and that the 1/f spatial frequency distribution is the one expected by the observer as she looks at an image.

Is the dominance for low spatial frequencies also a convention in visual arts? One could argue that Impressionists were possibly aware of the asymmetric role of low and high spatial frequencies, since by emphasizing large brush strokes, the interpretation of the painting was mostly carried out by the low spatial frequencies. However, it appears that some painters prefer to maintain a certain ambiguity in the interpretation. Maybe the most celebrated example of ambiguity in a painting is the enigmatic smile of the portrait of "Mona Lisa" painted by da Vinci (1503–1506). Livingstone (2000, 2002) has argued that the ambiguity of the facial emotion resulted from the superposition of two conflicting information in two different spatial frequency bands. A smile is clearly visible when only the low spatial frequencies are considered but a more neutral emotion is seen when all spatial frequencies are considered. Livingstone thus proposed that the ambiguity of Mona Lisa's smile will be most striking when an observer makes an eye movement looking first at the mouth and then at the eyes. When the mouth is seen in visual periphery, only the low spatial frequency information is available and the interpretation of the emotion is a smile. To quantify the fact that the change of emotion came mostly from the mouth region, other authors have added random luminance noise in the lower region of Mona Lisa's face and used the classification image technique to localize the relevant information (Kontsevich & Tyler, 2004).

The technique of adding luminance noise to a painting is interesting because it masks some information and preserves other information. This technique was modified in a study by Bonnar, Gosselin, and Schyns (2002) who presented an ambiguous painting masked by occluders of various sizes. The chosen painting was the "Slave Market with the Disappearing Bust of Voltaire" (Fig. 2a) by Salvador Dali

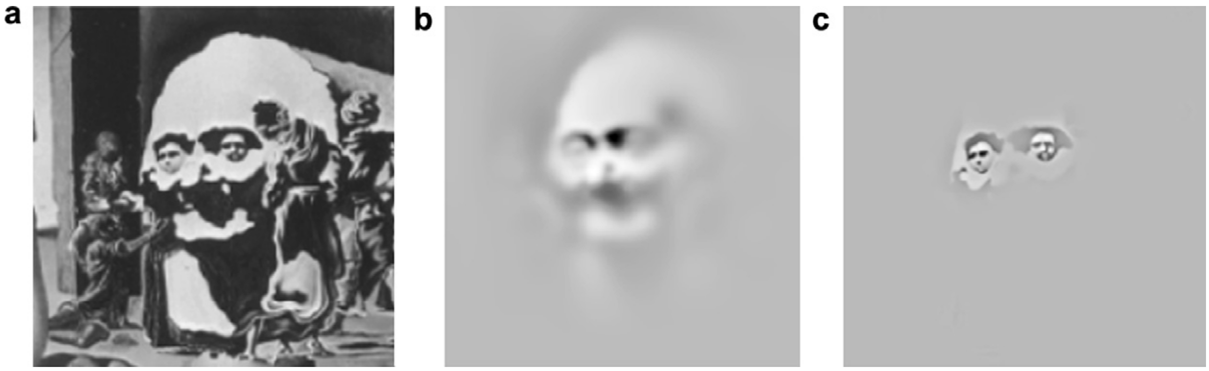

Fig. 2. (a) Ambiguous painting. The bust of Voltaire is perceived if one focuses on the coarse scale (or look from a far distance) and two nuns are perceived instead if one focuses on the fine scale (or look very close). Salvador Dali (1940) "Slave Market with the Disappearing Bust of Voltaire" (grey-scaled detail). Oil on canvas, 46.5 × 65.5 cm; Salvador Dali Museum, St. Petersburg, Florida. (b) The low spatial frequency (coarse) information used by observers when they reported seeing the bust of Voltaire. (c) The higher spatial frequency (fine) information used by observers when they reported seeing two nuns. Reproduced from Bonnar et al. (2002).

(1940) where a large face of the French philosopher Voltaire gives way to two nuns when one pays attention to the fine details of the painting. Observers were presented with a series of images each time masked by a different set of occluders and they had to report whether they saw the Voltaire or the nuns interpretation. Then, the authors reconstructed the visible parts of the image that led to the Voltaire interpretation (Fig. 2b) and the image that led to the nuns interpretation (Fig. 2c). As expected, the Voltaire interpretation involved mostly the low spatial frequencies of the image whereas the nuns interpretation involved higher spatial frequencies.

An alternative interpretation for both Mona Lisa's enigmatic smile and the disappearing bust of Voltaire comes from the ambiguous interpretation of the shadows in the paintings. Leonardo da Vinci took great troubles to depict highly saturated shadows around the mouth thanks to the delicate *sfumato* technique which consists in overlaying multiple translucent layers of paint (Elias & Cotte, 2008). The smile of Mona Lisa at low spatial frequencies comes from the large shadows of her cheekbones. If these shadows were properly interpreted, they could be discounted when the observer looks at the mouth, and thus the emotion of the mouth should not change when one considers different frequency bands. Similarly, the nuns in Dali's painting are defined from the shadows of the brow ridges and cheekbones of Voltaire, providing the hats and the dresses of the nuns. One sees Voltaire when the dark regions of the painting are interpreted as shadows and the two nuns when they are interpreted as dark surface material. By default, shadows tend to form large dark areas in a painting and as such contribute to the low spatial frequency information of the image. If these shadows are placed in specific areas (near the mouth in Mona Lisa and under the brow ridge in the disappearing bust of Voltaire), they can lose their role as shadows and offer an ambiguity to the interpretation of the painting.

In summary, natural scenes have a characteristic distribution of spatial frequencies and figurative art, to appear more natural, will necessarily match this prior distribution. However some artists have played cunningly with the spatial frequency contents of the painting to maintain some ambiguous interpretations. Other painters, such as abstract artists who freed themselves from representational depictions, have produced paintings that have a frequency distribution very different from that of natural images, and this convention on spatial frequency is a characteristic of their style.

4. Illumination and color

Rare are the paintings where the light source is depicted on the canvas. For all the other paintings, not knowing the location of the light source leads to ambiguities on the shape of objects, and not knowing the nature of the illuminant leads to ambiguities on the colors of objects. Naturally, these ambiguities are also present in everyday perception, and several prior constraints have been proposed to overcome these ambiguities. For instance, it is assumed that there is only one stationary light source in the scene (Mamassian, Knill, & Kersten, 1998) and that it is sufficiently diffuse so that color variations are due to changes in the surface materials rather than the illuminant (Brainard, 1997).

Do conventions in visual arts follow the perceptual priors for illumination and color? The placement of the illuminant has confused artists until the Renaissance, to the point that shadows were completely absent in paintings for centuries (Casati, 2004). Let us take the example of the "Adoration of the Magi" by Konrad Witz (Fig. 3). The artist was rightly appreciated for his mastering of shadows that presented fine geometrical details and nice penumbras (Stoichita, 1997). However a closer look at this painting shows several inconsistencies in the positions and forms of the shadows. For instance, the two statues at the corner of the building cast their shadow on the right and the left, something that is impossible unless there are two light sources in the scene.

Fig. 3. Unnoticed impossible shadows. There are multiple inconsistencies in this painting that are only visible if one pays close attention to the details. For instance, light is coming from the right but the statue in the left corner of the building cast a shadow on the right, and there are inconsistent shadows of the beams cast on the wall. Konrad Witz (1444) "Adoration of the Magi". Panel, 132 × 154 cm; Musée d'Art et d'Histoire, Geneva.

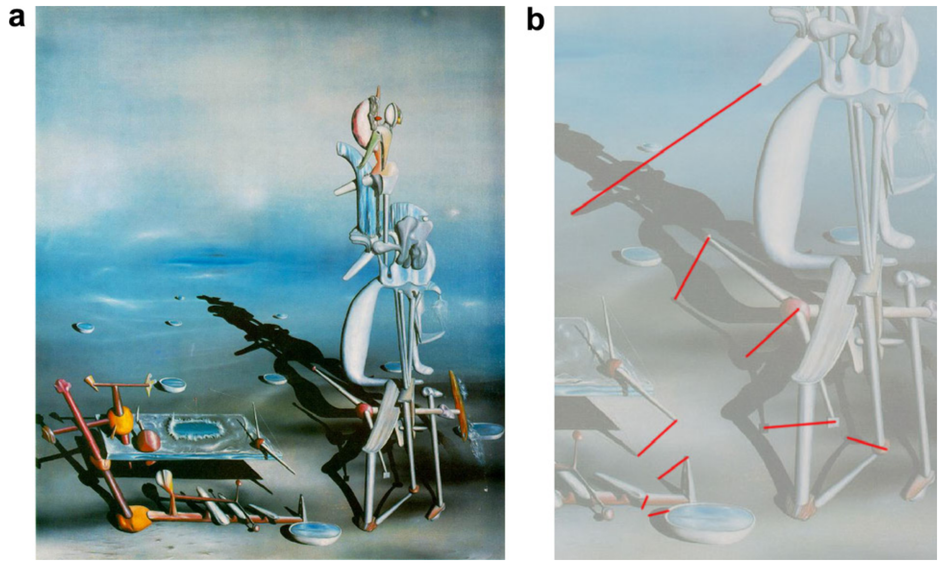

Fig. 4. Analyzing impossible shadows. (a) This surrealist painting offers shadows with fine detailed shapes. Yves Tanguy (1942) "Indefinite Divisibility". Oil on canvas, 102 × 89 cm; Albright-Knox Art Gallery, Buffalo, NY. (b) To determine the consistency of the light direction, one can trace constraint lines between identifiable object features and their cast shadow. Reprinted from Mamassian (2004).

Interestingly, these impossible shadows are not noticed, unless one voluntarily scrutinizes the painting (Fig. 4; Mamassian, 2004). Therefore, these inconsistencies do not take anything away from the quality of the painting, but allow the artist to emphasize the really important aspects of the painting. In other words, "the artist can take shortcuts, presenting cues more economically, and arranging surfaces and lights to suit the message of the piece rather than the requirements of the physical world." (Cavanagh, 2005). In contrast to the visual assumption that there is a single light source in the scene, the profusion of improbable or impossible shadows in western arts suggests that artists prefer to depict shadows as if there were multiple light sources, rather than having a shadow cross an object of interest.

The positions of shadows are constrained by the light source position. Human observers assume that light is coming from above their head, but also, and more surprisingly, slightly from the left (for a review, see Mamassian, 2004). Is this perceptual preference for a light on the left also a convention in the visual arts? There might be such a predominance for purely practical reasons. For instance, in his *Elements of Drawing*, John Ruskin wrote "Sit so that the light may come from the left, else the shadow of the pencil point interferes with your sight of your work." (Ruskin, 1857, p. 49). However, there might also be an aesthetic preference for a predominance of left light in paintings. By looking at some paintings in a mirror, Metzger (1936) has speculated that when the painting had the light coming from the left, it was more aesthetically pleasing than when light was from the right.

In an effort to quantify the predominance of paintings in the Western world where the light source is on the left, the paintings of the Louvre museum were surveyed by taking all the 659 paintings in their current catalogue (Mettais, 2002; see also McManus, Buckman, & Woolley, 2004). The paintings were sorted by periods, covering seven centuries from the 13th to the 19th century. They were also sorted in 194 portraits (Fig. 5a) and 465 non-portraits (Fig. 5b), the latter category including landscapes but also groups of people. Each painting was analyzed to see whether light was coming from the left, the right, or a general category called "non-specific" that included illumination from straight above, from the middle of the canvas (e.g. from a candle), or from inconsistent or multiple light sources. On average across all periods, the portraits

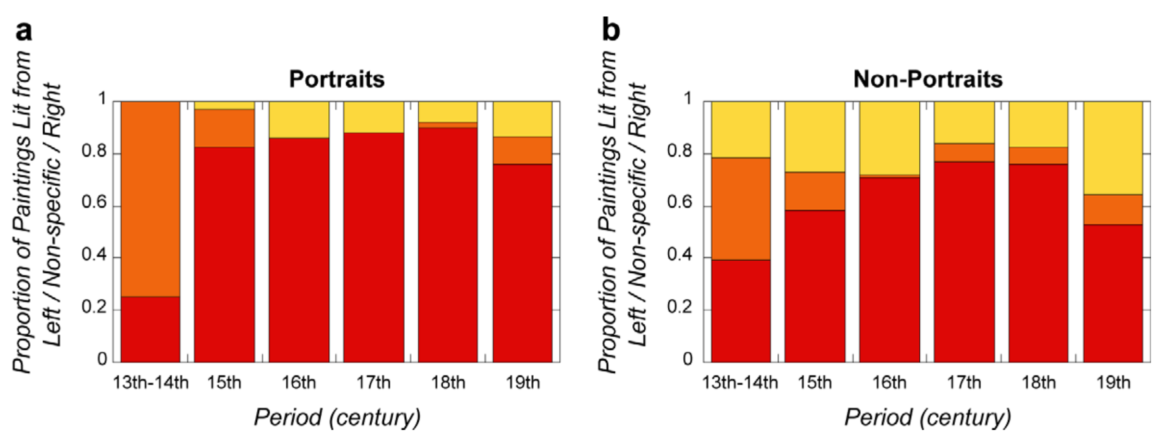

Fig. 5. Predominance of paintings lit from the left in the Louvre museum in Paris. (a) Portraits. This category corresponded to the painting of a single person posing for the artist and whose head or body occupied most of the canvas. (b) All paintings other than portraits. This category included landscapes, groups of people, and in general anything that was not categorized as a portrait. Color code: red is left lighting, yellow is right lighting, and orange corresponds to a "non-specific" category that included lights from straight above and other illumination conditions that were impossible to categorize as left or right light source positions.

were lit 84.0 percent of the time from the left and 9.8 percent from the right (6.2 percent from a non-specific direction), and the non-portraits were lit 67.3 percent of the time from the left and 23.0 percent from the right (9.7 percent non-specific). There was therefore an overwhelming predominance of paintings lit from left, and this bias was even more pronounced in portraits than in other paintings. The stronger bias for portraits is maybe not that surprising if we presume that portraits are in general more subject to conventions and studio constraints than outdoor scenes. In addition, it seems that this bias to the left for the light source position reached its peak between the 16th and the 18th century. Before that period, illumination conditions were not necessarily coherent in the scene, and after that period, artists started to free themselves from imposed conventions.

The choice of the light source position has an effect on the three-dimensional shape depicted. The problem of recovering shape-from-shading is under-constrained in the sense that the amount of light reflected by a surface is the result of the combination of the light intensity, the color of the surface and the orientation of the surface relative to the light source. From a perceptual point of view, shading remains a rather weak cue because of the ambiguity resulting from these multiple combinations (Mamassian & Kersten, 1996). In visual arts, the *chiaroscuro* technique from the Renaissance has helped provide a sense of shape from shading by the careful use of light contrasts, highlights and shadows. However, some ambiguities remain, in particular to differentiate the color of the surface and the intensity of the light source. For instance in engravings, artists were faced with hard choices to depict objects that varied both in color and in amount of reflected light (Zavagno & Massironi, 2006).

Given that a painting could be seen under various illumination conditions, from natural to artificial lights, how can the colors be reliably conveyed? One convention may be to over-emphasize the contrasts of the edges between objects. Chevreul (1839) is widely acknowledged for providing an early empirical description of simultaneous color contrast that has influenced both vision scientists and artists, in particular impressionists and pointillists. Opposing colors, especially red versus green and blue versus yellow, enhance each others' contrasts when they are placed side by side. The effect is most spectacular when the colors on the canvas are mostly from one pair of opposing colors, as in "The Night Café" by Vincent van Gogh (1888) (Fig. 6). Emphasizing color contrasts might be a convention that helps maintain color constancy irrespective of the illumination conditions.

The effects of color contrast take a new dimension when the depicted surfaces are small. The perceived color of a small patch is then very much affected by the background. However, recent research has shown that the shift in perceived color of the patch is larger when the background is itself composed of small regions rather than being uniform (Fig. 7a; Monnier & Shevell, 2003). This finding might provide a new perspective on pointillist paintings where small patches of color covering the whole canvas were preferred over large uniform color regions (Fig. 7b).

Within the contemporary view of the human visual system, the processing of color is split between luminance and chrominance pathways (e.g. Livingstone & Hubel, 1987). While natural objects are determined both by luminance and chrominance, some objects can therefore be defined exclusively from their luminance properties (i.e. different gray levels) and others from their chromatic properties (i.e. hue and saturation). In addition, the chrominance channel contributes weakly to the perception of three-dimensional position and movement. This view led Livingstone (2002) to propose that if an object is painted with a color that is equiluminant to (i.e. that has no luminance difference with) its background, that object should be difficult to maintain at a fixed position and its boundaries might be difficult to determine. Let us take the example of the "Coquelicots" by Claude Monet (1873) (Fig. 8a). The red poppies are near equiluminance to the green grass (Fig. 8b), maybe helping the impression that the poppies shiver in the wind.

In summary, it seems that conventions on the light source positions match well humans preferences for a placement above their head and slightly to the left. In contrast, the color of objects appear to follow conventions that emphasize the outlines of objects. While these conventions may be inspired from studies of color perception (Chevreul, 1839), they are not directly linked to specific prior constraints used by the visual system.

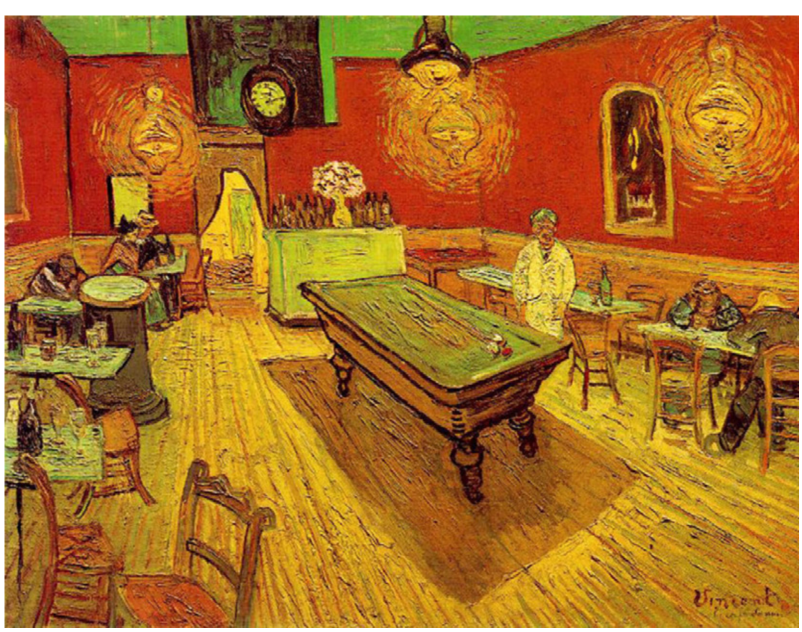

Fig. 6. Simultaneous color contrast. The green surfaces are adjacent to surfaces in red and orange tones. Vincent van Gogh (1888) "The Night Café". Oil on canvas, 72 × 92 cm; Yale University Art Gallery.

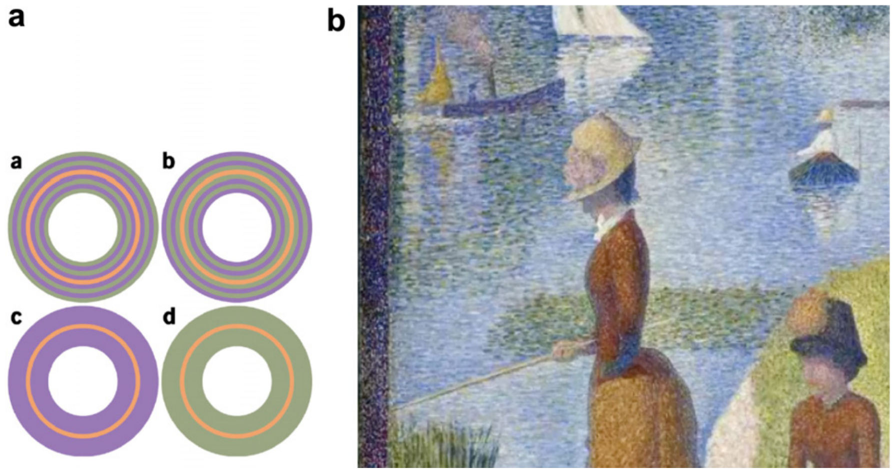

Fig. 7. (a) Apparent color shifts from textured background. The central ring in all four donut shapes is identical but has different color appearances depending on the background. The color shift is stronger when the background is itself composed of small rings (top figures) rather than uniform (bottom figures). Reprinted from Monnier and Shevell (2003). (b) Example of the use of small patches of color contrasts in pointillism. Red surfaces are placed next to green ones but a close examination of the painting shows that each surface is itself composed of numerous small patches. Georges Seurat (1884) "Un dimanche après-midi à l'Île de la Grande Jatte" (detail in top-left corner). Oil on canvas, 207 × 308 cm; Helen Birch Bartlett Memorial Collection, 1926.224; The Art Institute of Chicago (http://www.artic.edu/aic).

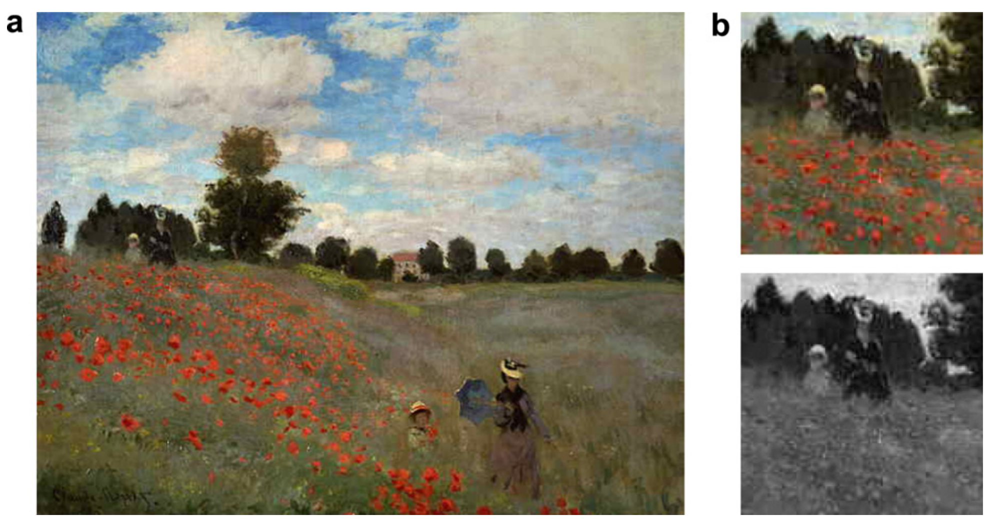

Fig. 8. Movement from equiluminance. (a) The red poppies are mostly defined from the color contrast with the green grass. Claude Monet (1873) "Coquelicots". Oil on canvas, 50 × 65 cm; Musée d'Orsay, Paris. (b) A detail of the same painting transformed to grey-scale suggests that the poppies are close to be equiluminant to the background.

5. Three-dimensional layout

Apart from composition and scaling effects, not knowing where the observer will stand in front of a painting can have dramatic effects on the interpretation of a depicted three-dimensional scene. Even if the painter has mastered the laws of linear perspective as they were derived at the beginning of the fifteenth century, all viewpoints except one should theoretically lead to geometrical deformations of the scene. The only viewpoint from where the painting is a faithful rendition of the scene is the center of projection, that is the location where the painter was standing if the scene was painted on a transparent canvas. Therefore, there is a fundamental ambiguity in the geometry of the depicted scene that depends on where the observer is standing.

The ambiguity on the geometry of the three-dimensional scene is an issue for museums but also for everyday perception as images seen from the wrong viewpoint abound in magazines, billboards or even movie theaters. Luckily, human observers rarely notice the deformations of the scene, a phenomenon known as La Gournerie's paradox. The resolution of this paradox is still a matter of debate. Some people argue that observers are just not sensitive to these deformations because of the nature of their internal representation of the three-dimensional world, namely a representation that is insensitive to the type of stretches that occur when one is not placed at the center of projection (Cutting, 1988). Other people argue that the deformations are not perceived because the observer has been able to calibrate herself, and then interpret the scene from the correct center of projection (Goldstein, 1987). To this end, the observer could use two kinds of information. First, there is some information in the frame and the canvas texture that helps determine the orientation of the canvas relative to the observer (Vishwanath, Girshick, & Banks, 2005). Second, there is some information in the depicted scene, such as the knowledge of the way the image of a cube is deformed when seen from different viewpoints (Kubovy, 1986). There is some evidence that human observers use such a rectangularity constraint, that is the belief that by default edges of objects are rectangular (Mamassian, 1998).

What are the conventions in visual arts to overcome the ambiguity of the three-dimensional layout associated to not knowing where the observer stands? One convention seems the prevalence of depicting geometrical objects that have a regular shape. The classical example of these objects are floor tiles that are overwhelmingly present in Renaissance paintings. The fact that tiles are almost always square, rather than an arbitrary aspect ratio, can help determine the proper center of projection for the painting. However, some objects resist the fore-shortening deformations they should undergo (Kubovy, 1986). For instance, in "The School of Athens" by Raphael (1510–1511) (see again Fig. 1b), the two spheres at the bottom right that represent the globe of the earth and the sphere of the fixed stars should project as oblong discs because of their eccentricity relative to the vanishing point. That the spheres are represented as circles rather than ellipses probably reflects the over-generalization of the fact that when we foveate a sphere, its retinal image is always a circular disc.

Other ambiguities of the three-dimensional layout can be found in the geometry of reflection. Mirrors and other specular surfaces are difficult to perceive because of confounding factors between the material properties, the objects surrounding the surface and the illumination conditions. Everyday perception seems to use knowledge of the statistics of illumination rather than contextual objects to make sense of glossy surfaces (Fleming, Dror, & Adelson, 2003). In contrast, artists seem to play carefully with the placement of contextual objects near the mirrors. In particular, one convention seems to represent the face of a depicted person in the mirror, even though, from an optical point of view, her face cannot be visible to both herself and the observer. This convention, sometimes called the "Venus effect" (Fig. 9), is effective because human observers have a naive understanding of physics (Bertamini, Latto, & Spooner, 2003). However, there are other examples related to reflections, for instance with respect to the size and the orientation of the reflected image, that suggest that artists used our misconception of optical reflections to depict a scene as they wanted it to be rather than veridically (Bertamini & Parks, 2005; Cavanagh, 2008).

In summary, it appears that to overcome the ambiguities on three-dimensional layout that are caused by the unknown viewpoint chosen by the observer, the convention to use regular shapes such as square tiles may be related to a generic perceptual prior to assume that objects have rectangular corners. In contrast, there are other conventions that appear unrelated to visual perception, such as the depiction of spheres as circles and the depiction of faces in mirrors.

6. Shape

Not knowing where the observer is will have effects on the perceived three-dimensional layout of the depicted scene, but also on the shape of the objects in the scene. Both two-dimensional (2 D) and three-dimensional (3 D) objects will be affected. In 2 D, there is an ambiguity on the aspect ratio of an object because an object seen from the side will undergo fore-shortening that will reduce its size along the slant direction. In 3 D, there is an ambiguity on the curvature of an object, because a concave object seen from below can produce the same image as a convex object seen from above. This latter ambiguity is well illustrated in a lithograph print from M. C. Escher (1955) called "Convex and Concave".

The ambiguity on the aspect ratio of 2 D objects may be resolved if we assume that objects have by default a unitary aspect ratio. In other words, an ellipse in an image is more likely to be the projection of a slanted circle (Geisler & Kersten, 2002). The ambiguity on the curvature of the 3 D object may be resolved if we assume that the viewpoint of the observer is located above the scene (Mamassian & Landy, 1998). When the shape is depicted as a line drawing, two intersecting lines are sufficient to represent an "egg" shape when the lines have the same curvature (both convex or both concave). In contrast, when the two line segments have opposite curvature, they generate the percept of a surface patch that is shaped as a saddle (Fig. 10a). The ambiguity inherent to the mapping between line curvature and surface shape can be lift by using the prior assumption that the object is seen from above (Mamassian & Landy, 1998).

The conventions on the representation of 2 D and 3 D shapes seem to match quite well the prior assumptions of the visual system. In 2 D, much effort was dedicated to promote the idea that the golden ratio (or golden section, i.e. 1.618:1) represented the perfect proportion of a rectangle (Green, 1995; McManus, 1980). How-

Fig. 9. The Venus effect. Most observers trust that Venus is seeing her reflection in this painting, even though this is not physically possible. Is this a license of the artist to show us Venus' face or is Venus looking at us? Diego Velázquez (1647–1651) "The Toilet of Venus ('The Rokeby Venus')". Oil on canvas, 122.5 × 177 cm; The National Gallery, London.

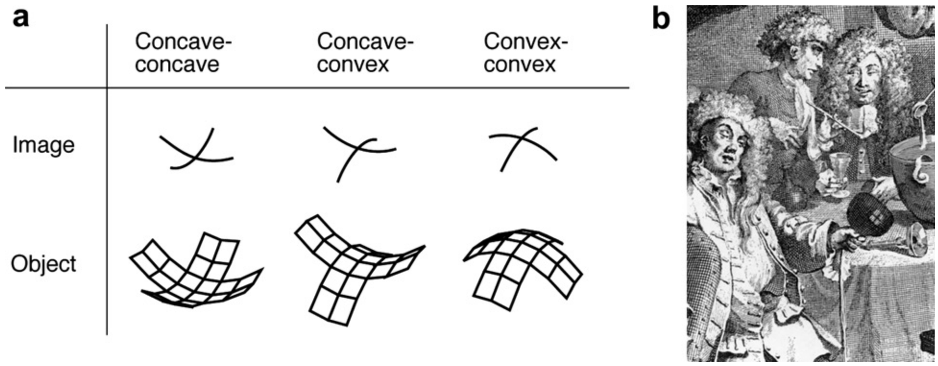

Fig. 10. (a) Correspondence between curved lines in an image and curved surface patches in 3 D space. Two curved line segments in an image are usually interpreted as a particular surface shape. Reprinted from Mamassian and Landy (1998). (b) William Hogarth (1732) "A Modern Midnight Conversation". Etching with engraving in black ink (detail). The character on the left is holding a bottle with a window reflection that helps determine the shape of the bottle.

ever, most recent studies found little evidence for a preference close to the golden ratio (e.g. Ohta, 1999). In contrast, there seems to be a robust preference for a unity ratio, that is for square figures (e.g. Davis, 2007). In 3 D, there is good evidence that artists used some conventions in line drawings that preceded our understanding of shape perception. Let us take for instance the etching "A Modern Midnight Conversation" by William Hogarth (1732) (Fig. 10b). One character is holding a bottle that shows the reflection of a four-pane window. It is plausible that this reflection is drawn more to help the observer determine the shape of the bottle than as a realistic representation of the scene. Indeed, the window would need to be tilted by 45 degrees to produce the oriented reflection, and the external light source at that time of the night (midnight) would need to be very powerful to produce such an intense reflection. This and other properties of line drawings are now being used in computer graphics to extract the sketch of an object (e.g. Girshick, Interrante, Haker, & Lemoine, 1999).

7. Movement

The absence of time in a static canvas is obviously a challenge for an artist who wants to depict a dynamic scene. Not only is there an ambiguity in the direction and speed of an object when it is frozen in a single frame, but there is also the more basic ambiguity of knowing whether the object is moving at all. In everyday perception, there is a prior for slowness that can explain numerous motion perception phenomena (Weiss, Simoncelli, & Adelson, 2002). Using a similar convention in visual arts would be counterproductive to display a dynamic scene, because that scene would tend to appear even more static. Artists have therefore struggled to propose conventions that exaggerate the motion to counteract the perceptual prior.

An old convention to depict movement in a static figure is "dynamic balance", where the impression of movement is brought by breaking the symmetry of the composition or by placing an objet in an unstable pose. One can argue that *contrapposto* is an example of dynamic balance, where a human figure stands with one leg supporting most of the body weight and the other leg is free and slightly flexed. The asymmetrical role of the legs can often be detected by noticing the twist of the shoulders relative to the hips (Fig. 11a).

Another convention used to depict movement is to present multiple, superimposed views of the object in motion. Such a representation corresponds to what one would see under a stroboscopic light (rapid sequence of bright flashes). Among the earliest scientific studies that used stroboscopic flashes are the representations of human movement by Etienne-Jules Marey (Fig. 11b). These studies influenced twentieth century artists such as Marcel Duchamp and his "Nu descendant l'escalier" (1912). Other ways to represent movement include photographic blur and direction vectors superimposed in the image (for a review, see Cutting, 2002).

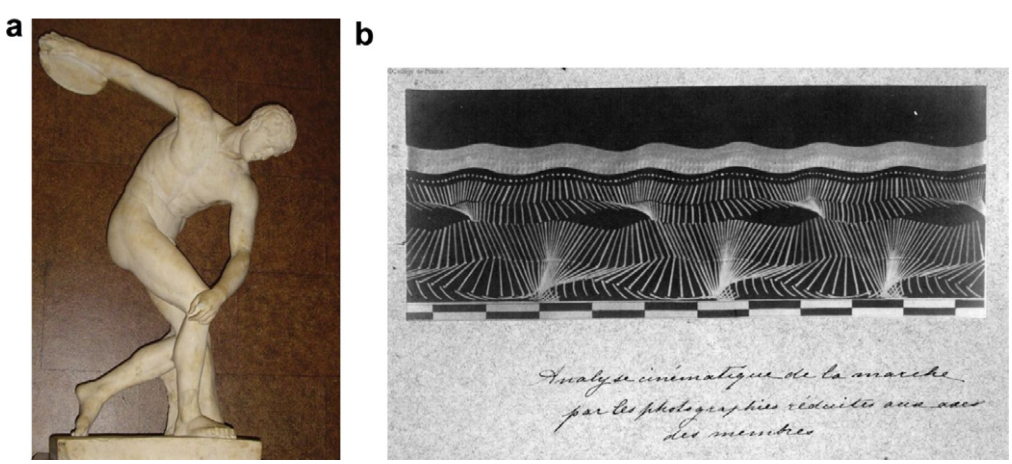

Fig. 11. (a) An example of *contrapposto*. The body stands on one leg and the other leg is flexed giving the impression that the discus thrower is about to release his throw. Copy of Myron (ca. 450 BC) "Discobolus". Marble Roman copy after a bronze original, 170 cm; The British Museum, London. (b) Example of movement under stroboscopic flashes. The movement of a person walking past the camera is fixed by flashing the scene at regular intervals. Reprinted from Marey (1884).

In summary, the conventions used in visual arts to depict movement are very different from the prior assumptions of the human observer. These conventions are often violating the structure of natural images, such as when multiple views are superimposed as if to simulate a stroboscopic lighting. Other artists have also used other tricks to produce an impression of motion. For instance, some Op art paintings, such as the closely packed parallel black and white lines of "Fall" by Bridget Riley (1963), create an intriguing movement perception that may be the result of small involuntary eye-movements (Zanker & Walker, 2004). More recently, the paintings of Patrick Hughes on canvases composed of truncated pyramids do generate a perception of non-rigidity because the linear perspective in the scene is in conflict with the slant of the canvas panes (Papathomas, 2002). These various attempts to produce an artistic perception of movement may one day reach the status of convention if they are effective and well accepted by the observers.

8. Discussion

Visual perception is ambiguous and visual arts play with these ambiguities. Ambiguities in visual perception are resolved thanks to prior constraints that are often derived from the knowledge of statistics of natural scenes. Ambiguities in visual arts are resolved thanks to conventions that found their inspirations from perceptual priors or, more interestingly, from other sources such as stylistic or arbitrary choices. Of course, not all ambiguities in visual perception are resolved, as witnessed by the multiple instances of bistable perception such as the one arising from looking at a Necker cube. Similarly, not all ambiguities in paintings are resolved, and artists probably strive to leave the right amount of ambiguities to let the observer contribute to his experience in a personal way.

The purpose of this review was to establish the relationship between visual prior constraints and artistic conventions. It is evident from this survey that most conventions have their roots in visual perception, such as the preference for above illumination that also presents a bias to the left. On the other hand, some conventions seem to differ from the priors of visual perception, such as the exaggerations of color contrasts or the depiction of figures in a mirror. Dividing the conventions that are perceptually-based from those that are not might be a fruitful exercise to make sense of what at first glance could appear as a bag of artistic tricks (Ramachandran & Hirstein, 1999). Visual art is different from visual perception, and one important difference might come from the conventions that differ from perceptual priors.

One can be sure that as vision scientists unravel visual ambiguities and constraints, and art historians report artistic conventions, there will always be an artist to violate these principles and promote a new style. However, these styles only make sense when one refers to the conventions they attack. What would be fauvism and the bewildering colors in Matisse paintings without centuries of hard work and conformism to reproduce colors as faithfully as possible? How could we understand cubism and the deconstructed figures of Picasso without referring to the constraint of a unitary viewpoint? Among these violations of conventions, "The Annunciation" by Crivelli (1486) is a playful painting full of amusing details, in particular the artist's decision to place completely irrelevant characters near the central vanishing point.

One practical application of the appreciation of ambiguities and conventions in visual arts is in product design. An attractive new product should be a compromise between its innovations and the ease with which it can be used. Understanding how to use a novel object presents some similarities with appreciating or not a painting. People working in product design often see the interaction with an object as a set of affordances (e.g, You & Chen, 2007), but some of these affordances can probably be redefined as conventions (Norman, 1999). As such, product design offers an interesting and practical forum to investigate the role of ambiguities and conventions in visual arts.

There are a lot of similarities in the work of artists and scientists (Kemp, 1990). But there are also differences. Nick Wade rightly notes that "Artists enhance and elaborate the effects, whereas scientists contract and constrain them." (Wade, 2003). More contact with visual arts can maybe help expand the horizon of some vision scientists.

Acknowledgments

This work benefited from grants by the France-Berkeley Funds (collaboration with Marty Banks) and a chaire d'excellence from the French Ministry of Research. I thank Patrick Cavanagh, Christopher Tyler and one insightful reviewer for discussions and comments on an earlier draft of the manuscript.

References

Arnheim, R. (1954). *Art and visual perception: A psychology of the creative eye*. Berkeley, CA: University of California Press.
Barlow, H. B. (1962). A method of determining the overall quantum efficiency of visual discriminations. *Journal of Physiology (London), 160*, 155–168.
Bertamini, M., Latto, R., & Spooner, A. (2003). The Venus effect: People's understanding of mirror reflections in paintings. *Perception, 32*, 593–599.
Bertamini, M., & Parks, T. E. (2005). On what people know about images on mirrors. *Cognition, 98*, 85–104.
Bonnar, L., Gosselin, F., & Schyns, P. G. (2002). Understanding Dali's Slave Market with the Disappearing Bust of Voltaire: A case study in the scale information driving perception. *Perception, 31*, 683–691.
Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *The Journal of the Optical Society of America A, 14*, 1393–1411.
Bruce, V., & Young, A. (1998). In *The eye of the beholder*. Oxford, UK: Oxford University Press.
Casati, R. (2004). Methodological issues in the study of the depiction of cast shadows: A case study in the relationships between art and cognition. *The Journal of Aesthetics and Art Criticism, 62*, 163–174.
Cavanagh, P. (2005). The artist as neuroscientist. *Nature, 434*, 301–307.
Cavanagh, P. (2008). Reflections in art. *Spatial Vision, 21*, 261–270.
Chevreul, M. E. (1839). *De la loi du contraste simultané des couleurs et de l'assortiment des objets colorés*. Paris: Pitois-Levrault et Cie.
Conway, B. R., & Livingstone, M. S. (2007). Perspectives on science and art. *Current Opinion in Neurobiology, 17*, 476–482.
Cutting, J. E. (1988). Affine distortions of pictorial space: Some predictions for Goldstein (1987) that LaGournerie (1859) might have made. *Journal of Experimental Psychology—Human Perception and Performance, 14*, 305–311.
Cutting, J. E. (2002). Representing motion in a static image: Constraints and parallels in art, science, and popular culture. *Perception, 31*, 1165–1193.
Davis, S. T. (2007). Aesthetic preferences for the unity ratio resist the influence of color illusions. *American Journal of Psychology, 120*, 47–71.
Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision, 2*, 324–353.
Elias, M., & Cotte, P. (2008). Multispectral camera and radiative transfer equation used to depict Leonardo's sfumato in Mona Lisa. *Applied Optics, 47*, 2146–2154.
Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *The Journal of the Optical Society of America, A, 4*, 2379–2394.
Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*, 347–368.
Geisler, W. S., & Kersten, D. (2002). Illusions, perception and Bayes. *Nature Neuroscience, 5*, 508–510.
Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurence in natural images predicts contour grouping performance. *Vision Research, 41*, 711–724.
Girshick, A., Interrante, V., Haker, S., & Lemoine, T. (1999). Line direction matters: An argument for the use of principal directions in 3D line drawings. *Proceedings SIGGRAPH 1999*.
Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. *Journal of Experimental Psychology—Human Perception and Performance, 13*, 256–266.
Gombrich, E. H. (1960). *Art and illusion*. Princeton, NJ: Princeton University Press.
Green, C. D. (1995). All that glitters—A review of psychological-research on the aesthetics of the golden section. *Perception, 24*, 937–968.
Gregory, R. L. (1980). Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London, B, 290*, 181–197.
Gregory, R. L., & Harris, S. (1995). *The artful eye*. Oxford, UK: Oxford University Press.
Kawabata, H., & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology, 91*, 1699–1705.

Hecht, H., Schwartz, R., & Atherton, M. (Eds.). (2003). *Looking into pictures: An interdisciplinary approach to pictorial space*. Cambridge, MA: MIT Press.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Kemp, M. (1990). *The science of art*. New Haven, CT: Yale University Press.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271–304.

Kontsevich, L. L., & Tyler, C. W. (2004). What makes Mona Lisa smile? *Vision Research, 44*, 1493–1498.

Kubovy, M. (1986). *The psychology of perspective and Renaissance art*. Cambridge, UK: Cambridge University Press.

Landy, M. S., Goutcher, R., Trommershäuser, J., & Mamassian, P. (2007). Visual estimation under risk. *Journal of Vision, 7:4*(6), 1–15.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *The Journal of Neuroscience, 7*, 3416–3468.

Livingstone, M. S. (2000). Is it warm? Is it real? Or just low spatial frequency? *Science, 290*, 1299.

Livingstone, M. S. (2002). *Vision and art: The biology of seeing*. New York: Harry N. Abrams.

Locher, P. J. (2003). An empirical investigation of the visual rightness theory of picture perception. *Acta Psychologica, 114*, 147–164.

Locher, P., Overbeeke, K., & Stappers, P. J. (2005). Spatial balance of color triads in the abstract art of Piet Mondrian. *Perception, 34*, 169–189.

Mamassian, P. (1998). The rectangularity constraint in linear perspective (abstract). *Perception, 27*(Suppl.), 34.

Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception, 33*, 1279–1290.

Mamassian, P., & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research, 36*, 2351–2367.

Mamassian, P., Knill, D. C., & Kersten, D. (1998). The perception of cast shadows. *Trends in Cognitive Sciences, 2*, 288–295.

Mamassian, P., & Landy, M. S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research, 38*, 2817–2832.

Maffei, L., & Fiorentini, A. (1995). *Arte e Cervello*. Bologna: Zanichelli Editore.

Marey, E.-J. (1884). Analyse cinématique de la marche. Comptes Rendus des Séances de l'Académie des Sciences, Paris, XCVIII.

McManus, I. C. (1980). The aesthetics of simple figures. *British Journal of Psychology, 71*, 505–524.

McManus, I. C., Buckman, J., & Woolley, E. (2004). Is light in pictures presumed to come from the left side? *Perception, 33*, 1421–1436.

McManus, I. C., & Thomas, P. (2007). Eye centring in portraits: A theoretical and empirical evaluation. *Perception, 36*, 167–182.

Mettais, V. (2002). *Louvre: 7 centuries of paintings*. Versailles: Art Lys.

Metzger, W. (1936). *Gesetze des Sehens*. Frankfurt am Main: Waldemar Kramer.

Monnier, P., & Shevell, S. K. (2003). Large shifts in color appearance from patterned chromatic backgrounds. *Nature Neuroscience, 6*, 801–802.

Norman, D. (1999). Affordance, conventions, and design. *Interactions, 6*, 38–42.

Ohta, H. (1999). Preferences in quadrangles reconsidered. *Perception, 28*, 505–517.

Papathomas, T. V. (2002). Experiments on the role of painted cues in Hughes's reverspectives. *Perception, 31*, 521–530.

Pinna, B. (2007). Art as a scientific object: Towards a visual science of art. *Spatial Vision, 20*, 493–508.

Ramachandran, V. S., & Hirstein, W. (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies, 6*, 15–51.

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology, 57*, 199–226.

Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.

Ruskin, J. (1857/1971). *The elements of drawing*. New York, NY: Dover.

Silvia, P. J. (2005). Emotional responses to art: From collation and arousal to cognition emotion. *Review of General Psychology, 9*, 342–357.

Solso, R. L. (1994). *Cognition and the visual arts*. Cambridge, MA: MIT Press.

Solso, R. L. (2001). Brain activities in a skilled versus a novice artist: An fMRI study. *Leonardo, 34*, 31–34.

Stoichita, V. I. (1997). *A short history of the shadow*. London, UK: Reaktion Books.

Tyler, C. W. (1998). Painters centre one eye in portraits. *Nature, 392*, 877–878.

Tyler, C. W. (2007a). Some principles of spatial organization in art. *Spatial Vision, 20*, 509–530.

Tyler, C. W. (2007b). Eye-centring in portraits: Reply to McManus and Thomas. *Perception, 36*, 183–188.

Vartanian, O., Martindale, C., Podsiadlo, J., Overbay, S., & Borkum, J. (2005). The link between composition and balance in masterworks v.s paintings of lower artistic quality. *British Journal of Psychology, 96*, 493–503.

Vishwanath, D., Girshick, A. R., & Banks, M. S. (2005). Why pictures look right when viewed from the wrong place. *Nature Neuroscience, 8*, 1401–1410.

Wade, N. (2003). Movements in art: From Rosso to Riley. *Perception, 32*, 1029–1036.

Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research, 21*, 1115–1122.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5*, 598–604.

Willats, J. (1997). *Art and representation*. Princeton, NJ: Princeton University Press.

You, H.-C., & Chen, K. (2007). Applications of affordance and semantics in product design. *Design Studies, 28*, 23–38.

Zanker, J. M., & Walker, R. (2004). A new look at Op art: towards a simple explanation of illusory motion. *Naturwissenschaften, 91*, 149–156.

Zavagno, D., & Massironi, M. (2006). Colours in black and white: The depiction of lightness and brightness in achromatic engravings before the invention of photography. *Perception, 35*, 91–100.

Zeki, S. (1999). *Inner vision: An exploration of art and the brain*. Oxford, UK: Oxford University Press.

Zeki, S. (2002). Neural concept formation and art: Dante, Michelangelo, Wagner. *Journal of Consciousness Studies, 9*, 53–76.