Human Visual Mechanism

***Shin'ya Nishida*†**

Abstract

What we see is not exactly what is there. In this paper, I review three visual illusions lurking in our daily experiences: slit view, surface reflectance perception, and color-motion asynchrony. I discuss the implications of these illusions for information processing by our visual system and for visual telecommunications technology for humans.

1. Introduction

The main task of our sense of vision is to determine the properties of objects and events in the external world from the image information projected on the retinas of our eyes. This is not as easy as you might expect, which is why the abilities of current machine vision systems are far inferior to those of humans and why robots that can see as we do will probably not appear in the next few decades.

Since visual processing is so complex, even the extraordinary human visual system sometimes makes errors, which result in various visual illusions. Visual illusions are not necessarily caused by the limitations of our visual system. In situations where the correct solution is impossible, the brain's effort to make the best guess results in illusions. Visual illusions are very important to researchers because they afford useful clues to the algorithms used by the visual system. By analyzing error patterns, one can see how the visual system accomplishes its miraculous feats.

Common visual illusions are mainly distortions of lengths, orientations, or shapes in line-drawing figures. However, these geometrical illusions make up just one limited class of visual illusion. Here, I review three illusions hidden in our daily experiences, address their scientific implications for the processing performed by our visual system, and mention how these implications may play a role in future visual telecommunications technology.

2. Slit view

Have you ever seen one of those electric signboards where the words scroll rapidly from right to left in front of some scene behind the signboard? Take a close look, and you will notice that the display consists only of an array of widely separated columns of LEDs (light emitting diodes), with the background being visible through the gaps. According to a maker of such signs, the separation between LEDs in the horizontal direction is ten times the vertical separation. This horizontal separation is too wide for us to recognize stationary letters, but the rapid scrolling enables us to see clear letter images (Fig. 1). This

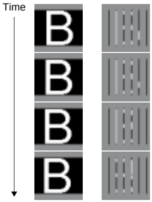

Fig. 1. Slit view. The letter 'B' is shown through slits. Letter recognition is difficult for stationary images, but easy when the letter moves behind the slits.

† NTT Communication Science Laboratories
Atsugi-shi, 243-0198 Japan
E-mail: nishida@brl.ntt.co.jp

means that by using temporal information, this display system successfully reduces the required number of pixels to 1/10 without losing apparent spatial resolution.

This is an example of an application of a classical visual illusion known as the slit view, which was originally noticed through the perception of perfect objects moving behind narrow slits. A classical account of the slit view is retinal painting. When slits are moving in front of a stationary pattern, the whole pattern is painted on the retina from one end to the other over time. The visible persistence of the retinal input (afterimage) causes it to appear to the observer that the whole pattern is presented at once. Even when a moving pattern is seen through stationary slits, as in the multi-slit display described above, similar retinal painting is expected to occur if the observer's eyes move while viewing the display to track the pattern movement.

It is true that the tracking eye movement makes it easy to read letters in a multi-slit display. The results of our experiments, however, indicate that the pattern movement improves pattern recognition even when the observer's eyes do not move. This finding argues against the idea that retinal painting can solely account for the multi-slit view. Our visual system has an ability to temporally integrate fragmented image inputs into a perceived coherent pattern by accurately taking into account the speed of the pattern movement behind slits [1].

The principle of animation is, as everyone knows, the sequential presentation of a series of stationary images. However, one should not take this fact to indicate that the perception of animation is just the sum of the perception of each stationary image. The visual system always has to analyze dynamic input changes in three-dimensional space-time coordinates. In an elegant fashion, the visual system merges and reconstructs space and time to obtain a valid perception of spatiotemporal events. The slit view is a valuable visual illusion that imparts this important fact.

3. Surface reflectance perception

Our eyes catch a pattern of lights, but our sense of surface brightness does not reflect the physical intensity of the incoming light. For instance, white paper looks white, and black coal looks black, regardless of whether they are placed under dark or bright illumination. It is not too much to say that our sense of brightness correlates with the rate of light reflection at the viewed surface, automatically canceling the effects of illumination on the surface. To avoid confusion, vision scientists sometimes use the special term 'lightness' to refer to the sense of surface reflection rate.

The visual system is said to compute lightness in comparison with surrounding surfaces. Since strong illumination would also elevate the brightness of adjacent surfaces, by judging whether a given surface is brighter than the surrounds, the visual system can to some extent infer the light reflection rate of the surface. However, this is not the only strategy that the visual system uses for illumination cancellation. In Fig. 2, the two regions indicated by arrows are the same gray and surrounded by grays of similar average brightness. Nevertheless, the top appears darker than the wall. Based on the nature of the real world, the visual system assumes that the top is probably illuminated more than the side, so it cancels the influence of illumination when computing surface lightness [2].

Like lightness perception, our sensation of surface color correlates with the chromatic properties of surface reflectance, rather than with the chromaticity of the reflected light. The visual system attempts to see the inherent color of a given surface by taking into account the color of illumination.

In addition, we can read complex surface reflectance properties, such as gross, matte, rough, and fuzzy. This ability is the basis for our perception of surface material (*e.g.*, metal, plastic, skin) and its state (*e.g.*, polished, wet, dusty). The perception of surface reflectance properties is a remarkable ability of our visual system.

Theoretically speaking, however, there must be

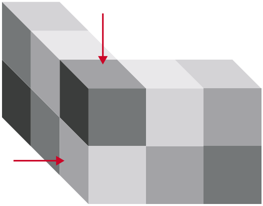

Fig. 2. Brightness illusion. The two regions indicated by arrows are painted the same gray, but the top looks darker, because it appears to be illuminated by stronger light.

some limitations for surface reflectance perception. On the one hand, the illumination pattern often becomes very complex due to spatial and temporal variations in intensity and spectrum, and inter-reflections between adjacent surfaces. On the other hand, surface reflectance is in general a complex function of incident and reflection angles, and also varies from one point to another even in apparently uniform surfaces. Even though correct estimation of surface reflectance requires accurate knowledge about the shape of the surface, it is known that the visual system cannot recover correct three-dimensional shapes from two-dimensional retinal images. It is therefore impossible to perfectly recover reflectance properties by just glancing at the surface.

So how well can we estimate surface reflectance? One limitation was suggested by an experiment conducted at NTT Communication Science Laboratories [3]. We required subjects to vary the reflectance (the intensity of matte reflection and the sharpness of specular reflection) of one surface to make it apparently identical to that of another surface with a different shape. Each surface had a smooth bumpy shape, with the height and scale of the bumps being systematically changed between conditions. To provide enough information about illumination and shape, we showed subjects an animation of the surface rotating about a horizontal axis. If the visual system can correctly estimate the surface reflectance properties, it should be able to match the reflectance between differently shaped surfaces. The results, however, indicated that the reflectance matching was very difficult. Instead, our subjects equated the intensity distribution of the surface image between the two surfaces.

As a result of recent advances in computer graphics, we often encounter rendered images that are apparently indistinguishable from real images. The point of the current technique of photo-realistic image rendering lies in how accurately and exclusively the rendering algorithm can simulate complex optical events in the real world. However, it is human observers who see reality in the rendered images, and the human visual system cannot perfectly check the physical consistency within the image. This means that perfect physical simulation is a sufficient but not a necessary condition for rendering images that look realistic to human observers. If we can understand how the visual system works, we will be able to develop a simple and efficient algorithm for generating realistic images.

4. Perceptual asynchrony of color and motion

In the perceptual asynchrony of color and motion (Fig. 3(a)), simultaneously occurring events are not seen as such [4]. A visual display demonstrating this

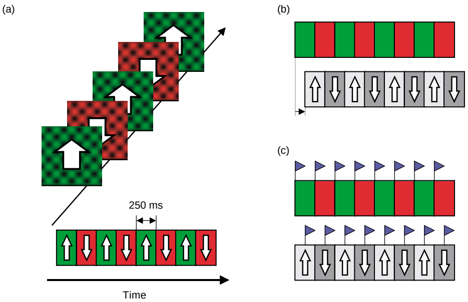

Fig. 3. Perceptual asynchrony of color and motion. (a) The oscillations in color and motion direction are physically synchronized, but are seen as asynchronous by observers. (b) According to the processing time hypothesis, this illusion reflects putative extra processing time for motion. (c) According to the time marker theory, the illusion is produced by erroneous matching of the time of color change with the time of motion (position change).

effect consists of a plaid pattern moving up and down. The direction reverses every 250 ms, so the rate of oscillation is 2 Hz. In synchrony with the motion reversal, the color of the stimulus changes. It is green during the upward phase and red during the downward phase. Subjects who view these moving images soon notice how difficult it is to judge the relationship between motion direction and color. They might sometimes see motion-color combinations that never actually occur (such as red moving upward). This is not simply due to the difficulty of binding attributes together at this alternation rate, because when the color change is delayed by about 100 ms relative to the direction change, an observer can bind color and motion much more easily. That is, the physically synchronous events appear not to be synchronous, while the physically asynchronous events seem to be synchronous.

A conventional account of this illusion is the processing time hypothesis (Fig. 3(b))[4]. It takes a certain amount of time from the occurrence of a physical event until its perception by the observer. At least 50 ms is needed for neural signals to reach the entrance of the visual cortex. Further time is required for the propagation of neural signals to other visual areas and for the perception of the event somewhere in the brain. From examinations of the cortical areas related to visual processing, it has been suggested that different attributes like color and motion are processed separately in different areas. Thus, the processing of a motion reversal might not finish at the same time as the processing of a color change even when the two events occur simultaneously in the physical world. Color-motion asynchrony could be a result of such a processing time difference. According to this processing time hypothesis, in order to account for the finding that color changes should be delayed to obtain apparent synchrony, the processing time for motion must be much longer than that for color.

Through a systematic investigation of the perceptual asynchrony of color and motion in NTT Communication Science Laboratories we obtained evidence against the processing time hypothesis [5]. First, the perceptual asynchrony occurred under limited temporal conditions. It was evident for rapid alternation of color and motion, but gradually disappeared as the interval between stimulus changes increased. Subjects could accurately judge the temporal order between a single color change and a single motion reversal. The apparent asynchrony should also be found regardless of the temporal condition, if a putative motion processing delay does in fact cause the illusion.

Second, the perceptual asynchrony was not accompanied by a corresponding difference in reaction time. We required subjects to press a button as soon as they saw a specific color or a specific motion direction in a stimulus sequence, and measured the reaction time from the stimulus onset until the button was pressed. We found no significant difference in the reaction times for color and motion.

An alternative explanation that we are proposing is the time marker theory (Fig. 3(c)), which states that human judgment of the temporal relationship is based on a comparison of salient temporal features in the stimulus (time marker). In the experiments, what had to be judged was the relationship between the timing of color change and the timing of motion direction change. Direction change, however, was not a salient feature for subjects because it is a higher-order temporal change in the sense that it is a (direction) change of (position) change. In the case of rapid alternation, the visual system could not detect direction reversals. Instead, the period of a given direction of motion became a salient feature (time marker) as indicated by the flags in Fig. 3(c). Since a wrong time marker was matched with the marker assigned to a color change, color and direction were no longer seen as synchronous. When the color change was delayed by about 100 ms, the time markers were nearly aligned, which led to an apparent synchrony.

The time marker theory does not require processing time differences. The illusion is expected to occur even if there is no significant processing time difference between color and motion, which agrees with the reaction time results. In addition, the theory predicts that the difference in the temporal structure of the stimulus change, not the difference in the stimulus attribute, is the critical factor for generating the perceptual asynchrony. This prediction is directly supported by the occurrence of the opposite effect (*i.e.*, apparent delay of color) when the temporal structure was swapped between color and motion.

We believe that the time marker theory not only explains the color-motion asynchrony illusion, but also represents an idea that can be developed into a general theory of human time perception. We are now trying to extend our theory to cross-modal temporal judgment, an important human factor for multimedia technology.

5. Conclusion

We can learn a lot from visual illusions. We learn from the slit view that we do not see a collection of snapshots; from surface reflectance perception, that we do not see incoming light; and from the perceptual asynchrony, that subjective time does not reflect the processing time in our brain.

References

[1] S. Nishida, "Direction-selective mechanism mediates identification of spatial patterns moving behind narrow slits," Proceedings of the 2nd Annual Meeting of Vision ScienceS Society, pp. 137, 2002.

[2] E. H. Adelson, "Perceptual organization and the judgement of brightness," Science, Vol. 262, No. 5142, pp. 2042-2044, 1993.

[3] S. Nishida and M. Shinya, "Use of image-based information in judgments of surface reflectance properties," Journal of the Optical Society of America, A, Vol. 15, No. 12, pp. 2951-2965, 1998.

[4] K. Moutoussis and S. Zeki, "A direct demonstration of perceptual asynchrony in vision." Proc R Soc Lond B Biol Sci Vol. 264, No. 1380, pp. 393-399, 1997.

[5] S. Nishida and A. Johnston, "Marker correspondence, not processing latency, determines temporal binding of visual attributes," Current Biology, Vol. 12, No. 3, pp. 359-368, 2002.

Shin'ya Nishida

Senior Research Scientist, Supervisor, Human and Information Science Laboratory, NTT Communication Science Laboratories.

He received the B.S., M.S., and Ph.D. degrees in psychology from Kyoto University, in 1985, 1987, and 1996, respectively. After a two-year stay at ATR Auditory and Visual Perception Laboratories in Kyoto as a research associate, he joined NTT Basic Research Laboratories in Tokyo (now Communication Science Laboratories) in 1992. Dr. Nishida is a member of the Vision Society of Japan, the Japanese Psychological Association, the Japanese Psychonomic Society, the Association for Research in Vision and Ophthalmology, and the Optical Society of America.