3D VOLUMETRIC OBJECT PERCEPTION IN BINOCULAR VISION

Qi Zhang[*], *Masanori Idesawa*[**], *and Ken Mogi*[*]

[*]Sony Computer Science Laboratories,
Takanawa Muse Bldg. 3-14-13 Higashi-gotanda, Shinagawa-ku, Tokyo, Japan 141-0022.
[**]Graduate School of Information Systems, The University of Electro-Communications,
1-5-1 Chofugaoka, Chofushi, Tokyo, Japan 182-8585.

ABSTRACT

3D object perception is one of the important issues in the study of human visual functions. It is a seemingly effortless process that requires no conscious thought for the human beings, but a difficult computational problem for machines. The object perception with binocular viewing utilizes the disparities between the two eyes to recover the 3D information of an object, and takes the advantages of the fact that the human beings have two eyes. Recently, a new visual effect named as mime effect was found, in which an illusory 3D volumetric object is perceived due to some stereoscopically displayed inducing objects (Zhang et al. 1998). Here we propose a processing model with both top-down and bottom-up processes, and discuss about the involvement of the early and higher-level visual cortical areas to this 3D volumetric object perception. It is hoped to provide new clues to the understanding of human 3D visual system.

1. INTRODUCTION

One of the greatest mysteries in cognitive science is the human ability to perceive visually presented objects with high accuracy and lightening speed. For more than 100 years, psychologists and neuroscientists have sought to understand the mechanisms underlying our innate ability to perceive and identify objects, using techniques ranging from psychophysical studies of human subjects to single neuron recording from non-human primates. Despite these efforts, the neural substrates of object perception remain elusive. Interest in how human object perception works is heightened by the fact that efforts to duplicate this ability in machines have not met with extraordinary success.

Some recent studies used some 2D images such as photos or line drawings of 3D objects to study the 3D object recognition [2,3]. The healthy human beings, however, have two eyes, and the disparities between the two eyes are very important cues for the perception of 3D objects. So the study on the 3D object perception with binocular viewing is very important for understanding the mechanism underlying the 3D object perception in human visual systems. In this paper, I will introduce a recently found new visual effect in binocular vision and discuss about its processing mechanisms in the brain.

2. MIME EFFECT IN 3D OBJECT PERCEPTION

When some inducing objects are suitably related and arranged in 3-D space without occlusion relationship and displayed stereoscopically, a volumetric illusory object is perceived with binocular viewing. These inducing objects sustain the volumetric object in the same way as the hands of mimers, and they are named as sustaining cues and this phenomenon as the "Mime Effect". A typical example is

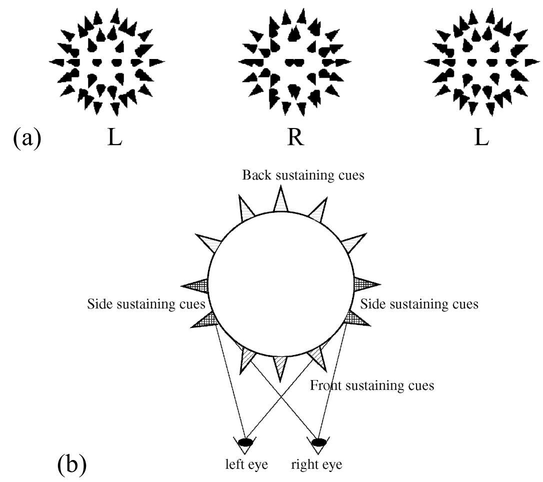

Fig. 1: Mime effect in 3-D object perception. (a) A complete transparent sphere is perceived in the presence of all the three types of sustaining cues. (b) The diagrammatic sketch of three types of sustaining cues. (L: left eye view, R: right eye view; left and center are for uncrossed fusion, while center and right are for crossed fusion.)

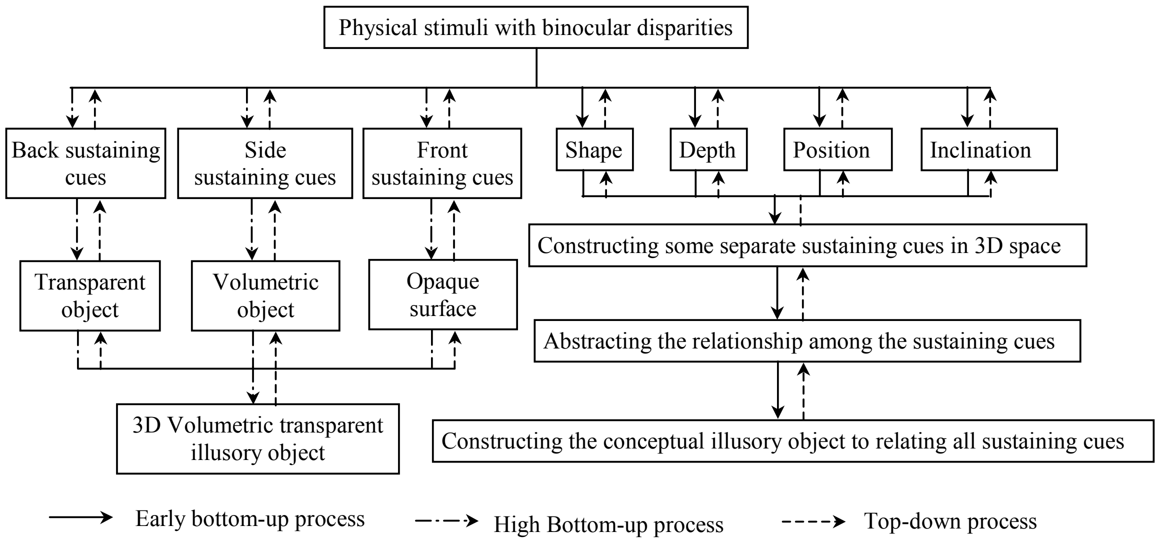

Fig. 2: Heterarchical model for perception in the mime effect. The attributes of the sustaining cues are first constructed from the physical stimuli with binocular disparities. Then, the relations among them are abstracted and an illusory object is computed based on these relations. These processes are bottom-up. Later the conceptual of the illusory object will be processed top-down in the early cortical areas to collect the necessary information for constructing a illusory object with specific properties. Then Bottom-up processes are conducted again to determine the properties of the illusory object, and three types of sustaining cues contribute differently to the 3D volumetric transparent illusory object perception in the mime effect.

shown in Fig. 1(a), where a number of cones are depicted with their centers of bases positioned on a spherical surface. A volumetric transparent sphere is perceived, as if inside brightness is different from that of the surrounding space and like as a transparent glass sphere; then the cones are perceived as standing on a transparent spherical object that, in fact, is not depicted at all. The sphere also appears to be "filled" with transparent medium and to occupy volumetric space, not simply a spherical surface.

To explore the mechanism of 3-D perception involving the mime effect, sustaining cues were studied by depicting only part of them. Depending on their positions on the illusory object, they play different actions to mime effect perception and were classified into three types: back, side and front sustaining cues (Fig. 1(b)). *Back sustaining cues* are in the back area of the illusory object, and are seen beyond the illusory object for both eyes, determining the transparency of the volumetric object. *Side sustaining cues* are on the side of the illusory object and in the binocular unpaired area, i.e., they exist in the front area for one eye and back area for the other. They evoke the outline of the illusory object and lead to the perception of its contour. Because they exist only on the side area and the effects for the two eyes are inconsistent, the perception of transparency induced by the side sustaining cues alone is rather unstable. *Front sustaining cues* are located in front of the illusory object for both eyes and only support the illusory surface, giving an impression of an opaque surface alone, but the perception itself is somewhat weaker than for the other two types.

When these three types of cues appear in different combinations, a different perception of the object is observed: when all the three types of sustaining cues are presented simultaneously, an entire transparent object will be perceived clearly; when back sustaining cues appear together with side sustaining cues, the illusory object will be perceived as transparent; an opaque illusory object is perceived if front sustaining cues exist either alone or together with side sustaining cues but without back sustaining cues.

3. PROCESSING MODEL FOR MIME EFFECT

Marr (1982) pointed out that the ability of human beings to generate a robust, three-dimensional percept of the world based on a pair of two-dimensional retinal images (stereopsis) requires the neurons in the visual cortex that can detect interocular image disparities, and proposed three relatively independent levels of processing for 3D vision, the low-level modules aimed to recover a meaningful description of the input intensity image, the middle-level modules to recover shape or location in space, and the high-level modules to construct the 3D representations[4]. Marr also advocated a 2.5D sketch is constructed first during the processing.

We think that the modules to processing 3D object perception are not independent, they will interact each other. In addition, a 3D concept is reconstructed directly without passing the step of building a 2.5D sketch.

3.1. A model with bottom-up and top-down processes

What is the mechanism underlying the 3D volumetric object perception in the mime effect? To answer this question, we considered a heterarchical processing model from the perspective of optic pathway.

In non-human primates, object perception is mediated by a network consisted of over 30 distinct cortical areas, which is arranged in a multilayered hierarchy [5]. Information moves through the system both bottom-up — from the eye to the top of the hierarchy — and top-down. A series of ventrally located cortical areas are thought to be primarily responsible for object perception by now.

The visual representation areas, however, are organized along both serial hierarchical and parallel distributed principle, with interconnections within and between levels. This arrangement, which may be described as an interconnected serial-parallel heterarchy, appears to be a general principle of neural organization. A heterarchy contains multiple parallel hierarchies, any one of which may dominate its local network domain at any given time by virtue of lateral interactions between them. The local connectivity of arrays of individual neurons forms the same type of heterarchy within the visual system from the retina to the cerebellum.

In the perception of the mime effect, the 3D volumetric illusory object is perceived due to the inducing objects of sustaining cues, for example, the cones in Fig. 1(a). The attributes of the cones, such as shapes, depths, positions, and inclinations, play very important roles in this perception. So before attaining the 3D illusory object perception, the cones should be perceived first via these bottom-up processes, and their attributes such as inclinations and positions are determined simultaneously in some parallel processes. Then the perceived cones will be compared each other, and their relationships are abstracted. Since the brain prefers to perceived objects in integration rather than as many separate objects, an illusory object contacting to all the cones is computed to relate the cones together based on the relationships among the cones. Consequently, the cones are perceived as a simple aggregate that is standing on the illusory object and contacting with it.

Then, the concept of the illusory object is feedback down from the higher-level cortical areas to the lower level areas. This top-down processes is to collect the information for determining the properties of the 3D illusory object.

Again, the collected necessary information is forwarded bottom-up again to construct the 3D illusory object with some specific properties. Thus, the cortical loop formed by these bottom-up and top-down processes would provide the neural basis for the illusory object perception under consideration.

In the last section, three types of sustaining cues were introduced to explain the mime effect. Their different contributions to the specific properties are mainly discriminated in this bottom-up processing step, depending on their different positions relative to the illusory object contribute. The back sustaining cues determine the transparency of the volumetric object, and the side sustaining cues silhouette the illusory object. In addition, since the side sustaining cues are located in the binocularly unpaired position, they also contribute to the volume perception, which is completed in order to remove the conflicts caused by the inconsistency between the binocularly unpaired information from the two eyes. Although the front sustaining cues only support the illusory surface and give an impression of the opaque surface alone, a volumetric object with the identical property is constructed when three types of sustaining cues appear simultaneously, and the transparency induced by back sustaining cues is extended to the entire object. The conceptual model is shown in Fig. 2.

3.2 Involved visual cortical areas in human brain

We have proposed a conceptual model of the processing for the 3D volumetric illusory object perception in the mime effect. Now, we would like to discuss about the involved visual cortical areas for this perceptual process.

In the human visual system, visual signals received by retina are first transferred to the lateral geniculate nucleus (LGN), where signals from the left and right eyes are shuffled, remaining in the same order as on the retina, then go to the occipital cortex to be processed in a number of differentiatable visual areas of V1, V2, V3, V4, and MT.

One property that many cortical cells have in common is that they receive visual input from both eyes. This is in marked contrast to LGN cells, which only respond to visual stimuli presented to one eye or the other. Cells that receive information from both eyes are called binocular neurons; V1 is the first level in the visual system where there is an excitatory combination of input from the eyes. The research with non-human primates found that the binocular cells in the early visual areas of V1 and V2 were activated by some specific visual perception, such as illusory contour [6], and binocular depth [7]. Cells involved in binocular disparity processing have also been found in other cortical areas, including V3 [8] and the medial superior temporal area (MST) [9]. Another area, MT (middle temporal or V5), has also been shown to contain neurons sensitive to binocular disparity [10].

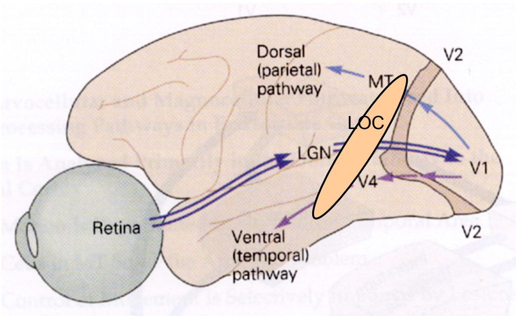

Fig. 3: Visual cortical areas involved in the mime effect located in both dorsal and ventral pathway in the human visual systems.

The human and macaque ventral streams appear to be structurally and functionally homologous up to the level of V4 [11]. In the mime effect, positions and inclinations of the sustaining cues are inferred from the depth information determined by the binocular disparities. So binocular cells in the above areas of dorsal stream (Fig. 3) all contribute to the 3D volumetric object perception.

On the other hand, visual input reaches the ventral stream via a thalamocortical projection from the LGN to V1, and V1 in turn projects to V2, V2 to V4. V4 is the area to process shape and color information. So it should also contribute to the perception of the sustaining cues' shape and color, and be necessary for the 3D volumetric object perception in the mime effect.

The above areas conduct the bottom-up processes to determine the properties of the sustaining cues. After this, the concept of the 3D illusory object is thought to be construct in higher-level areas.

Recent researches reported that several areas beyond V4 might participate in object perception. Potential higher visual areas in the human brain include the LOC [12], which is consisted of a dorsal subdivision (LO) [2] and a ventral-anterior subdivision located in the fusiform gyrus (pFs/LOa) area [13], and the parahippocampal place area [14]. These candidate areas are clustered together near the temporal-occipital junction, and all are active during object recognition. They include both dorsal and ventral areas, and are the first reasonable place to construct the 3D illusory object in the mime effect. There, the bottom-up processing results from both dorsal and ventral pathway mentioned above can be integrated together. These high-level areas then will interact with earlier visual cortical areas and conduct the top-down processes. After multiple repetitions of the top-down and bottom-up processes, the representation of the 3-D volumetric illusory object is constructed.

Besides of these visual areas, some other high level cortical areas in parietal and anterior lobes, which are usually considered to be in charge of some high level processing, such as somatosensory motor controlling and intelligent reasoning, may also be possible to participate in the process for the perception of 3D volumetric illusory object. This is because that the volume perception needs much intention and computation to remove the conflicts occurred due to the binocularly unpaired side sustaining cues. Anyway, LOC may be the first area to have the representation of the volume perception.

4. DISCUSSION

Visual function is typically studied by parametrically varying a single attribute of the visual scene, while keeping all other attributes constant. Correlations between attribute variations and neural activity provide information about areas or neural circuits involved in perception of the attribute of interest. In the case of luminance and color, which are easily quantified, this approach works well. Visual neuroscientists using this strategy have made substantial progress in identifying the specific brain structures and circuits responsible for relatively simple perceptual capacities, such as luminance and color perception and the ability to see simple patterns [15]. Unlike luminance, complex shapes and three-dimensional objects cannot easily be quantified and parameterized. As a result, neuroscientists have been less successful at understanding complex shape processing and object perception in visual areas beyond V4.

In this paper, we utilized a newly found visual effect of the mime effect to study the processing mechanism for the 3D volumetric object perception in human brain. This perception involves a large range of the visual cortical areas both in the dorsal and ventral pathways. In addition, some high-level visual areas, such as LOC, also participate in the processing. This visual effect is very effective to systematically study the mechanism of the 3D visual perception in human brain, including the high level cortices.

The proposed model and the consideration of the visual cortical areas are just a conjecture so far. Future work using fMRI measurements are necessary to provide evidences to support these inferences.

5. CONLUSION

In this paper, we introduced a newly found visual effect named as mime effect, in which a 3D volumetric illusory object is perceived from some interrelated inducing objects displayed stereoscopically. We proposed a processing model including both top-down and bottom-up processes for this perception. In addition, we discussed about the involved visual cortical areas to this perception in human brain. This study is hoped to open a new epoch of systematically studying on the 3D perception mechanism in the both early and high-level visual cortex.

6. REFERENCES

[1] Zhang Q., Idesawa M., and Sakaguchi Y., "Mime effect in the perception of volumetric transparent illusory objects with binocular viewing," *Jpn. J. Appl. Phys.*, 37, 3B, pp. L329-L332, 1998.

[2] Grill-Spector K., Kourtzi Z., and Kanwisher N., "The lateral occipital complex and its role in object recognition," Vision Research, **41**, pp. 1409–1422, 2001.

[3] Kanwisher N. G., Chun M. M., McDermott J., and Ledden P. J., "Functional imagining of human visual recognition," *Brain Res. Cogn. Brain Res.*, 5, pp. 55–67, 1996.

[4] Marr D., Vision. *W.H. Freeman and Company*, New York, 1982.

[5] Felleman DJ, and Van Essen DC, "Distributed hierarchical processing in the primate cerebral cortex," *Cerebral Cortex*, 1:1-47, 1991.

[6] Lee T.S., and Nguyen M., "Dynamics of subjective contour formation in the early visual cortex," Proc. Natl. Acad. Sci. USA, 98, pp. 1907-1911, 2001.

[7] Freeman R. D., "Stereoscopic vision: Which parts of the brain are involved?" Current Biology, 9: R610-R613, 1999.

[8] Felleman DJ, Van Essen DC: "Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex," *J Neurophysiol,* 57:889-920, 1987.

[9] Roy JP, Komatsu H, Wurtz RH: "Disparity sensitivity of neurons in monkey extrastriate area MST," *J Neurosci*, 12:2478-2492, 1992.

[10] DeAngelis GC, Newsome, WT: "Organization of disparity-selective neurons in macaque area MT," *J Neurosci*, 19:1398-1415, 1999.

[11] Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell, RBH: "Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging," *Science*, 268:889-893, 1995.

[12] Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady, TJ, Rosen BR, Tootell RB: "Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex," *Proc Natl Acad Sci USA*, 92:8135-8139, 1995.

[13] Kanwisher N, McDermott J, Chun MM: The fusiform face area: "a module in human extrastriate cortex specialized for face perception," *J Neurosci*, 17:4302-4311, 1997.

[14] Epstein R, Harris A, Stanley D, Kanwisher N: The parahippocampal place area: Recognition, navigation, or encoding?" *Neuron*, 23:115-125, 1999.

[15] Mazer J.A. and Gallant J.L., "Object recognition: Seeing us seeing shapes," Current Biology, 10:R668–R670, 2000.