A Mathematical Approach for Understanding Optical Illusions: The Miracle of the Human Vision System

Kokichi Sugihara

1. Introduction

There is a class of anomalous pictures called "pictures of impossible objects". They are popular in that Dutch artist M C Escher used them as inspiration for his art. When we see these anomalous pictures, we have impressions of 3D structures, but at the same time we feel that these structures are physically impossible. In other words, we have inconsistent impressions of the objects; 3D structures on the one hand, and physical impossibility on the other. This phenomenon belongs to optical illusion.

From a mathematical perspective, "impossible objects" are not necessarily impossible; indeed some of them can be realised as 3D objects. These objects cheat our perceptions in two manners. First, we feel they are impossible although we are looking at existing objects. Second, although we recognise that our perception is inconsistent, we cannot mentally adjust them to their true shapes. In addition to anomalous pictures, we encounter various types of optical illusions when we try to interpret 2D pictures as 3D structures. In this article we present a mathematical model for a computer vision system to understand 2D images as 3D objects, and compare it with the human vision system. Determining the differences between human and computer vision systems might then help us understand why certain types of optical illusions arise.

2. Illusions Generated by Impossible Objects

Let us start by discussing some typical examples of optical illusions related to 3D structures.

The first class of optical illusions is that generated by anomalous pictures. Figure 1 shows an example of an anomalous picture called here "Impossible Stairs". This drawing was presented by Penrose and Penrose in their scientific paper on psychology [3], and was used by M C Escher in his famous artwork "Ascending and Descending" (1960) [1]. In this drawing, there are walls surrounding a square yard, and stairs are located on top of the walls. However, if we follow the stairs in the ascending direction, we eventually reach the starting point, and hence there is no end to the stairs. This is physically impossible and hence it belongs in the pictures of impossible objects. This optical illusion can be explained by the fact that each subsection of the drawing corresponds to a physically realistic 3D structure.

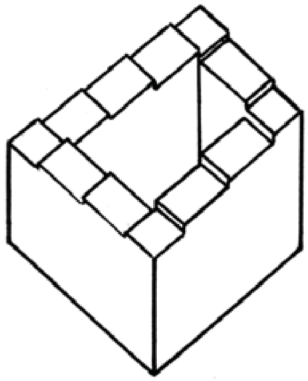
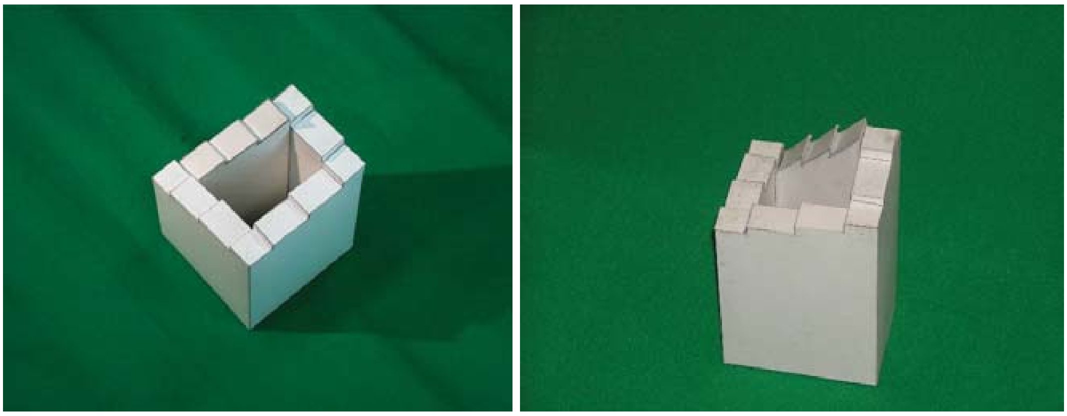


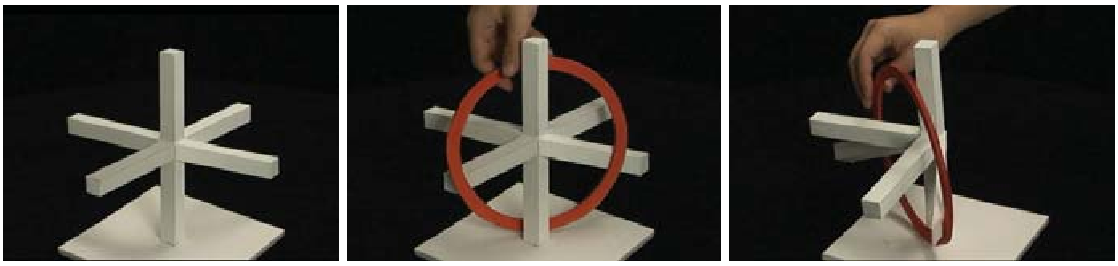
Fig. 1. Drawing of an impossible object "Impossible Stairs".

The second class of optical illusion is 3D objects that appear to be impossible. An example is shown in Fig. 2. Note that this is not a 2D drawing but a picture of an actual 3D object. When we see this object from a specific direction as shown in Fig. 2(a), it appears to be the same as the "Impossible Stairs". Thus, we feel that we are viewing an impossible object. The actual shape of this object is revealed when we view it from a different direction, as shown in Fig. 2(b); three of the sets of stairs have horizontal steps, but the fourth set consists of slanted steps, thus explaining the apparent difference in heights.



(a) (b)

Fig. 2. Model representation of the "Impossible Stairs", viewed from two different directions: (a) object seen from a special viewpoint; (b) another view of the same object.



(a) (b) (c)

Fig. 3. Impossible motion "Four Perches and a Ring".

We also call this kind of solid an impossible object, because it is a 3D realisation of what is represented in a picture of an impossible object. The impossible object appears to have an impossible structure when it is seen from a specific direction, but not from other directions.

The third example is a 3D object that looks familiar, but if it interacts with another object, their joint configuration appears to be physically impossible [7]. Figure 3 shows an example. In Fig. 3(a), the object itself appears to be a familiar solid composed of a vertical column and four horizontal perpendicular perches. However, a ring can be hung around the object as shown in Fig. 3(b). The ring passes behind the column but it is in front of all four perches, in a way which appears to be physically impossible. The actual shape of the object is revealed in Fig. 3(c); all four perches extend backwards, away from the ring, and hence the ring can be hung in front of the perches.

The perceived structure of the 3D object in Fig. 3(a) appears to be incompatible with the placement of the second object in Fig. 3(b). Therefore, we suspect something is incorrect regarding our perception of the structure. However, we cannot correct our perception of the structure to make it consistent with the placement of the ring. The interpretation of the object as having horizontal and mutually orthogonal perches is very stable. Indeed even after we rotate it as shown in Fig. 3(c) and understand the true shape, the horizontal and mutually orthogonal four perches are again evoked in our mind when we come back to the initial viewpoint. This example seems to suggest that our brain interprets the structure of an object automatically, ignoring our knowledge about the true shape of the object.

These are typical examples of optical illusions related to 3D objects. They have several common characteristics. First, even though we notice that our perception contains inconsistencies, we cannot correct our perception to make it physically consistent. Second, we remain predisposed towards certain object configurations even after we understand the true shape of an object. Third, we are predisposed to symmetric or rectangular objects although there are many other possibilities. In the next section we consider why these characteristics arise using a mathematical model for image interpretation.

3. Mathematical Model for Image Interpretation

Figure 4 shows a possible flow of information processing for interpreting a given image as a 3D object. It consists of three steps, and it is a typical model for a computer vision system [5, 6].

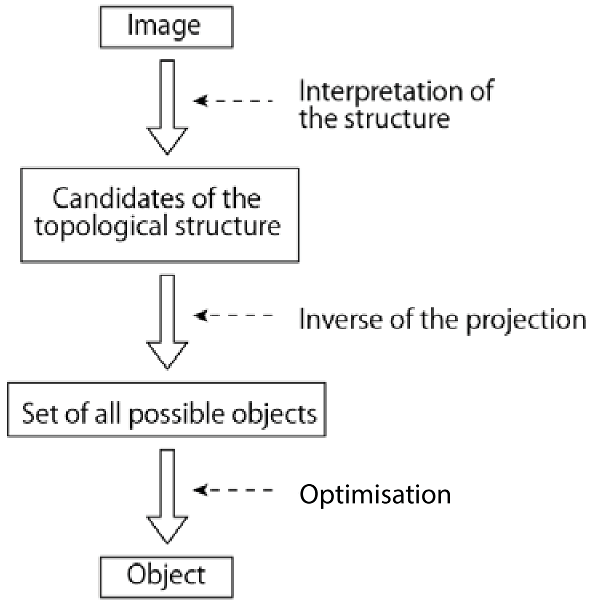


Fig. 4. Information processing model for extracting 3D structures from 2D images.

In the first step, possible candidates for the topological structure of the object are extracted. This task is typically accomplished by describing the nature of the object world with a vertex dictionary and drawing grammar, and by searching for interpretations that are consistent with the dictionary and grammar [2].

In the second step, we judge whether a given interpretation is realisable as a 3D solid object. For this objective we assign variables to represent the vertices and the faces of unknown objects, and represent what is known about the image with equations, and search for objects whose projection coincides with the given image by solving the equations [5]. We determine that the object is realisable if the equations have solutions, and unrealisable otherwise. For the case that the object is realisable, there are infinitely many solutions and hence the shape of the object cannot be determined uniquely [2].

In the third step, we choose from among all the solutions the one that is most likely. To this end, we use the psychological nature of human perception, such as the preference for highly symmetric structures and the preference for rectangularity, and reduce the problem to an optimisation problem.

For the first and the second steps we refer the reader to [2, 5, 6], and here we concentrate on the third step.

According to Gestalt psychology, humans are apt to group figure elements into simple and well-shaped objects, such as highly symmetric objects. Observing the responses of human vision systems to impossible objects and impossible configurations, we feel that they prefer objects with many rectangles [4]. In other words, the human brain selects the object that has the most rectangles. It seems that the human brain determines the most likely shape in a picture is correct, and ignores other possible shapes. This preference for rectangularity seems very strong.

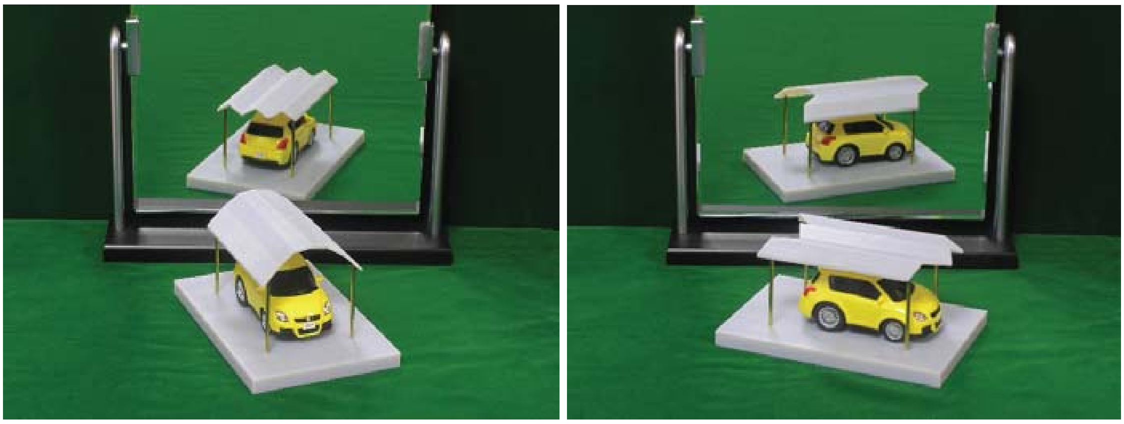
On the basis of this observation, we reduce the third step into an optimisation problem, in which we search for the object that has as many rectangles as physically possible. Then, we can explain the optical illusions shown in Figs. 1, 2 and 3.

The drawing in Fig. 1 is interpreted as stairs consisting of horizontal and vertical plates, although many other interpretations are mathematically possible, and consequently it is judged to be physically inconsistent. Similarly, the object in Fig. 2(a) is interpreted as stairs consisting of horizontal and vertical plates. Even though we know the real structure of the object from Fig. 2(b) and that the interpretation of horizontal steps is contradictory, we are predisposed towards the physically inconsistent interpretation. The object in Fig. 3(a) is also interpreted as a vertical column with perpendicular perches. Even though we feel this interpretation is not consistent with the motion of the ring, we usually do not perceive other shapes.

This preference for rectangles can explain many optical illusions related to 3D objects. Therefore, we have a strategy to cheat the human brain, that is, we employ angles other than right angles to construct 3D structures that look rectangular. Actually, the objects in Figs. 2 and 3 were designed by adopting this strategy.

4. Toward New Illusory 3D Objects

Once we accept the assumption of the brain's preference for right angles, we can design 3D objects that generate new types of optical illusions. Examples of newly developed 3D objects are now presented.



(a) (b)

Fig. 5. The Garage Roof Optical Illusion.

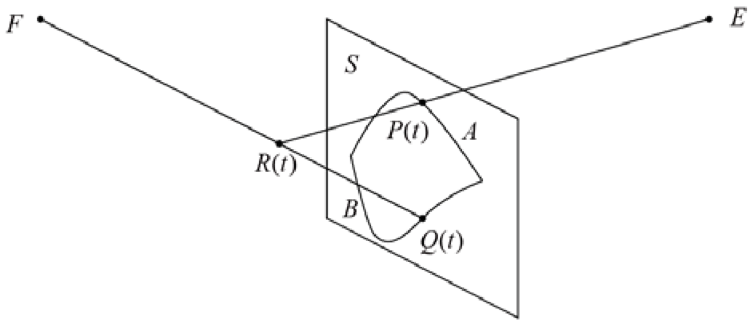


Fig. 6. Construction of a space curve that appears to be constructed from two separate curves when it is seen from two viewpoints.

Figure 5(a) shows a miniature carport and a vertical mirror positioned behind it. The mirror is an ordinary plane mirror. However, the shape of the roof appears to be round when viewed from the front and corrugated in the mirror image. The direct view and the reflected view are quite different; it is almost impossible to believe that they are the same object. The true shape of the roof is neither round nor corrugated; it has an irregular shape as shown in Fig. 5(b).

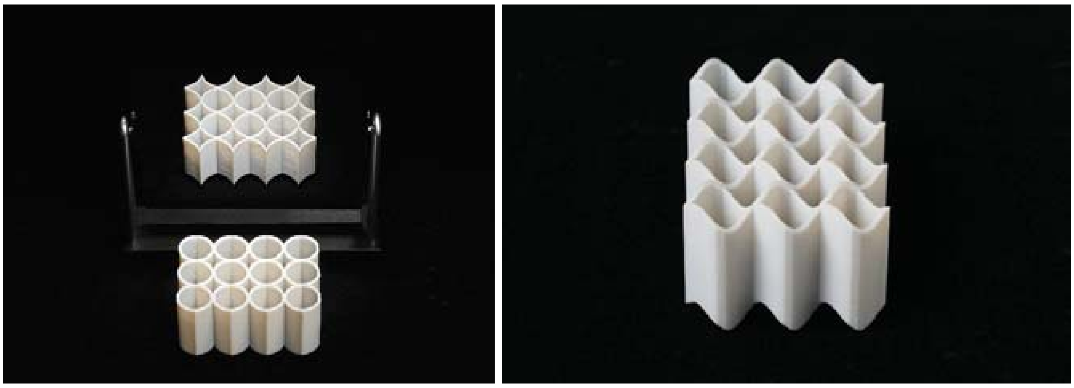
This roof is created as a surface swept by a line segment that moves in the 3D space without changing its orientation. This class of surface is called a cylindrical surface, and the direction parallel to the generating line segment is called the direction of the axis. Because of the preference for rectangles, we can expect that the edge curve of a cylindrical surface appears to be a planar curve obtained by cutting the surface with a plane perpendicular to the axis of the surface. This actually happens in our brains when we see Fig. 5(a). We call this type of surface an "ambiguous cylinder" because the perceived shape changes with the direction it is viewed from.

This illusion is strong in the following sense. First, although we logically know that the object cannot change its shape when it is reflected in a mirror, we cannot correct our perception, that is, we cannot resolve the inconsistency between the two appearances. Second, even after we know the true shape of the object, our brains return to the original perception when we come back to the original viewpoint.

An ambiguous cylinder can be generated with the following procedure. As shown in Fig. 6, let S be a vertical plane, and A and B be two curves on S such that they are horizontally monotone and their terminal points coincide with each other. Let E and F be two points outside S. We construct a space curve that coincides with A when it is seen from the viewpoint E and that coincides with B when it is seen from the viewpoint F. For that purpose, we consider the point $P(t)$ that moves along the curve A from one terminal point to the other terminal point for $0 \le t \le 1$. For each t, let T be the plane passing through $P(t)$, E, F, and let $Q(t)$ be the point of intersection of T and B. $Q(t)$ is unique because A and B are horizontally monotone and have the same terminal points. Let $R(t)$ be the point of intersection of the line passing through E and $P(t)$ and the line passing through F and $Q(t)$. $R(t)$, $0 \le t \le 1$, is the space curve

(a) (b)

Fig. 7. Ambiguous cylinder "A Full Moon and a Star".



(a) (b)

Fig. 8. Ambiguous cylinder "Capricious Pipes".

we want to construct. Finally we choose the line segment L that is perpendicular to the plane S, and move it in such a way that one terminal point of L travels along the space curve $R(t)$, $0 \leq t \leq 1$. The surface swept by L is the ambiguous cylinder.

Figures 7 and 8 show two more examples of ambiguous cylinders. In both figures, (a) shows the object and its mirror image, which appear to be very different from each other, and (b) shows the object seen from a general viewpoint.

5. Concluding Remarks

We have shown several classes of optical illusions generated by 3D objects, and a conceptual model of object perception that can explain these illusions. We can understand that these illusions arise firstly because 2D images lack depth information and secondly because the human brains prefer right angles when extracting 3D objects from 2D images.

When viewing real objects, we use both eyes. Hence, a triangle is formed between the two eyes and the target. This allows us to perceive the distance from the object. This is the principle of binocular stereo. On the other hand, when we see 2D images of objects we lose this depth perception, and hence we must make guesses about the shape of the object. The guesses are not necessarily correct and hence optical illusions arise.

These observations show us the importance of binocular stereo. Viewing 3D objects directly and viewing their 2D images are completely different. The former gives us additional information about the distance to the objects and hence we can make additional inferences about the structure. The latter, on the other hand, contains no direct information about the distance and hence there is no guarantee of the correctness of our perceptions. We thus have to be careful of the differences between 3D objects and their 2D images.

Acknowledgements

This work is supported in part by Grant-in-Aid for Basic Scientific Research No. 16H01728 and Challenging Exploratory Research No. 15K12067 of MEXT.

References

[1] B. Ernst, *The Magic Mirror of M. C. Escher* (Taschen America L.L.C., 1994).

[2] D. A. Huffman, Impossible objects as nonsense sentences, in *Machine Intelligence 6*, eds. B. Meltzer and D. Michie (Edinburgh University Press, Edinburgh, 1971), pp. 295–323.

[3] L. S. Penrose and R. Penrose, Impossible objects — A special type of visual illusion, *British J. Psychology*, **49** (1958) 31–33.

[4] D. N. Perkins, Visual discrimination between rectangular and nonrectangular parallelopipeds, *Perception & Psychophysics* **12** (1972) 293–331.

[5] K. Sugihara, *Machine Interpretation of Line Drawings* (The MIT Press, Cambridge, 1986).

[6] K. Sugihara, Design of solid for antigravity motion illusion, *Computational Geometry: Theory and Applications* **47** (2014) 675–682.

[7] K. Sugihara, Design of ambiguous cylinders, *Proceedings of the 10th Asian Forum on Graphic Science* (*AFGS2015*) (Bangkok, August 4–7, 2015) (to appear).

Translated from Sugaku Tushin Vol. 20-2

Kokichi Sugihara

Professor, Meiji University, Japan

Deputy Director, Meiji Institute for Advanced Study of Mathematical Sciences, Japan

kokichis@meiji.ac.jp

Kokichi Sugihara received Dr. Eng. in mathematical engineering from the University of Tokyo in 1980. He worked at Electrotechnical Laboratory of the Ministry of International Trade and Industry of Japan, Nagoya University, the University of Tokyo, and is now a professor of Meiji University. His research areas include computer vision, computational geometry, robust computation and computational illusion. He found a method for realising impossible objects in a 3D space without utilising traditional tricks, and extended the method to design new classes of 3D illusions such as impossible motion and ambiguous cylinders. His illusions won the first prize of the Best Illusion of the Year Contest in 2010 and 2013, and the second prize in 2015 and 2016.