



Room-size illusion and recovery of the true appearance

Kokichi Sugihara¹

Received: 14 January 2022 / Accepted: 10 April 2022 / Published online: 11 January 2023
© The JJIAM Publishing Committee and Springer Nature Japan KK, part of Springer Nature 2023

Abstract

It is widely known among photographers that photographing a room with a wide-angle lens exaggerates the size of the room; nevertheless, such images are commonly found on hotel-reservation web sites. The present paper points out that the size exaggeration is a kind of optical illusion caused by an inappropriate viewpoint from which the image is seen, and presents a method we developed for removing the illusion and thus recovering the true appearance of the room. This method requires only a single image together with the lens center at which the image is taken. From this information, we can generate images that would be obtained if we stand at the same point as the camera and pan around the original scene, changing the view direction. The validity of the method is shown by examples. Possible applications to size-exaggerated images posted on web sites are also discussed.

Keywords Depth exaggeration · Optical illusion · Wide-angle lens · Viewpoint · Room size · Standard lens

Mathematics Subject Classification 3MAH45000

1 Introduction

Optical illusion is visual phenomenon in which what we see is different from the physical reality. This is not a simple mistake, because the same phenomenon occurs even when we know the real shape. In this sense, an optical illusion stems from the intrinsic nature of the human visual system, and it cannot be removed by training a neural network [3, 13].

The purpose of seeing is to understand the surrounding environment in order to perform actions such as grasping objects and avoiding obstacles. Therefore, optical illusions may cause accidents and other inconveniences [4]. For example, the slope

✉ Kokichi Sugihara
kokichis@meiji.ac.jp

¹ Meiji Institute for Advanced Study of Mathematical Sciences, Meiji University, 4-21-1 Nakano, Nakano-ku, Tokyo 164-8525, Japan

illusion causes us to misinterpret downhill roads as uphill roads [1, 8, 18], and the Tarski illusion causes us to perceive road curves as more gentle than they are in reality [16]. These misperceptions might cause traffic accidents. Hence, it is practically important to understand the mechanisms of optical illusions, and develop methods to remove or decrease illusions as much as possible to provide safety and reliability in our daily lives.

An important class of optical illusion occurs when we see photographic images. Illusions belonging to this class include the leaning-tower illusion [7], stretching-tub illusion [7], and corridor illusion, to mention a few. In this paper we concentrate on a class of photographic images that cause misunderstanding during human communication.

Photographic images are convenient tools for visual communication. We see what we want to see even if they are located too far away to see directly. However, images are not necessarily faithful to the actual scene, and sometimes they give us wrong impression. When we visit a hotel-reservation web site, for example, we often come across an image of a room that makes the room look bigger than its actual size. Similar room-size exaggerations are found in images for various advertisements, such as for real estate properties, rental office spaces, and event halls.

It is widely known by photographers that room-size exaggeration occurs when images are taken using wide-angle lenses [15]. Photographers want to convey visual information of a room as much as possible by a single image. If they use a standard lens, they can cover only a small portion of the room. Instead, they use wide-angle lenses, which can cover a wider range of a room. However, the resulting images exaggerate the depth of the room.

These images might be regarded as false advertising, but this is often unintentional. The photographers just want to show more of the room in the advertisement. However, they also want to depict information about the room as accurately as possible. Based on this understanding, we try to find a method to show more of a room without size exaggeration.

From a physics point of view, taking an image is a purely optical process, and hence the precise appearance of a scene is recorded in the image. Indeed, if we see the image from the viewpoint corresponding to the lens center, we can see the true appearance of the scene. The room-size exaggeration occurs simply because we see the image from an inappropriate viewpoint. In this sense, the room-size exaggeration is a type of optical illusion. On the basis of this recognition, we propose a method to recover the true appearance of the same scene from images with an exaggerated room size. This paper is an extended study of the author's previously published short abstract on this topic [17].

The rest of the paper is organized as follows. The source of the optical illusion in wide-angle images is clarified in Sect. 2, and a method for removing the illusion and recovering the true appearance of the scene is constructed in Sect. 3. The proposed method is validated using computational experiments in Sect. 4, and supporting techniques and possible applications are discussed in Sect. 5. Concluding remarks are given in Sect. 6.

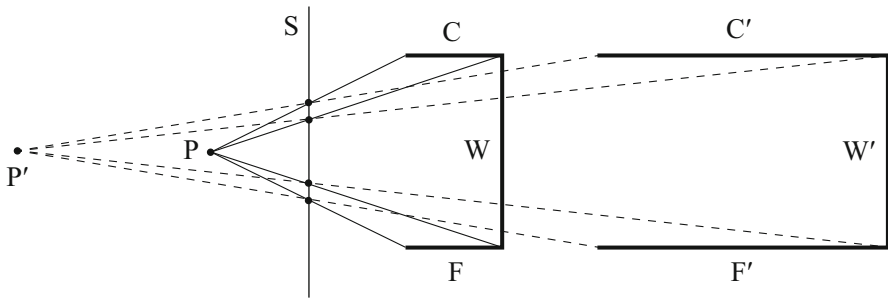


Fig. 1 Reason for the depth exaggeration

2 Room-size illusion

We consider why the room size is exaggerated in an image taken by a wide-angle lens [15]. As shown in Fig. 1, let us consider a room with floor F , wall W , and ceiling C , and take its image using a camera with the lens center at P and the image plane at S . The projected images of the corner vertices are shown by small dots on S . If we place our viewpoint at P , the image has the same appearance as the original room. In this sense the projected image conveys faithful information about the room.

However, it is difficult to see the image from P , because P is too close to the image plane S . A wide-angle lens has a short focal length, which means that the distance from P to S is short. Therefore, we usually see the image from a viewpoint that is much farther from S than P . Let this viewpoint be at P' . Because light travels in straight lines, we extend each vertex in the image along the half line starting at P' and passing through the vertex (broken lines in Fig. 1) until the object reaches its known size, and thus reconstruct the 3D structure. As shown in Fig. 1, the reconstructed floor, wall, and ceiling are F' , W' , and C' , respectively. We reconstruct the room in such a way that the distance between F' and C' is approximately the same as the known height of the ceiling of an ordinary room. As a result, we perceive the room as being farther than the actual distance, and the depth of the room appears much larger than the actual depth. This is the geometric mechanism by which the room-size illusion occurs.

An example of this visual effect is shown in Fig. 2 using a paper model of a cubic room. A general view is shown in Fig. 2a. The front wall is removed to take images of the inside, and the ceiling is removed to make the inside bright. Figure 2b, c shows two images of this room taken from the front with different lenses. Panel (b) was taken by a wide-angle lens with 14-mm focal length for a 35-mm full-frame camera, and panel (c) was taken by a standard lens with a 70-mm focal length for a 35-mm full-frame camera. When we see these two images from the same distance, we perceive the depth of the room to be much larger in (b) than in (c). For Fig. 2b, the distance from the image to the lens center is equal to 0.39 times the width of the image. For Fig. 2c, it is equal to 1.9 times the width of the image. If we see the image from the associated lens center, we will have the same visual angle between any two points in the scene because it is equivalent to see the room directly from the camera position, and consequently we can expect to perceive the depth of the room correctly.

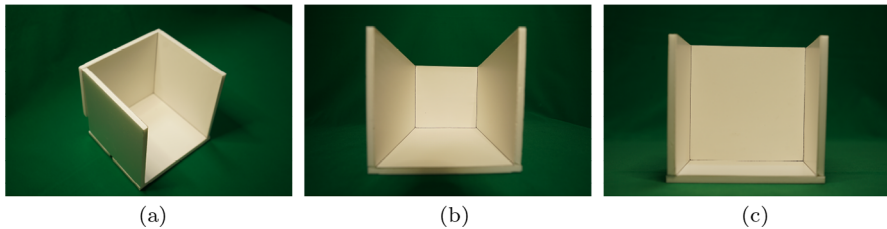


Fig. 2 Cubic room and images of it taken with different focal-length lenses: **a** general view, **b** 14-mm focal-length image, **c** 70-mm focal-length image

As we see in Fig. 1, P is the lens center at which the image on S is taken. Geometrically the image is the central projection of the scene with the center of the projection at P and the plane of the projection at S . The stored image may be displayed on a computer screen or sheet of paper at various sizes. We assume that the image will be expanded isotropically in every direction by a common scale factor, say s . This means that if we expand the image by s in the horizontal direction, we also expand it by s in the vertical direction. At the same time, we expand the surrounding 3D space by the same factor so that the distance between P and S is also multiplied by s . We call the resulting position of P the *associated lens center* with respect to the expanded image.

Suppose that two points Q_1 and Q_2 are seen from viewpoint P . The angle formed by two half lines starting at P and passing through either Q_1 or Q_2 is called the *visual angle* of Q_1 and Q_2 formed at P . The visual angle between any two points in the image formed at the associated lens center is invariant under scaling and is the same as the visual angle we obtain when we see the original scene directly from the camera position.

We can state our observation in the form of the following principle:

True-Appearance Principle. An image gives the same visual angle as the original scene if and only if it is seen from the associated lens center.

On the basis of this principle, we constructed a method for recovering the true appearances of the scene from a given image.

3 Recovery of true appearance

As shown in Fig. 3, let A' be the image taken by a camera with respect to the lens center at P . Because we use a wide-angle lens, the distance d from P to A' is relatively short.

Suppose that we want to take another image of the same scene at the same camera position using a standard lens. The standard lens has a focal length between 50 mm and 70 mm for a 35-mm full-frame camera, which is much larger than that of the wide-angle lens. So, as shown in Fig. 3, the resulting image A will be much smaller than A' if we adjust the scale in such a way that the distance between the image plane and the associated lens center is equal to d . Thus, cutting off the surrounding area in image A' gives us a standard-lens image in the same view direction.

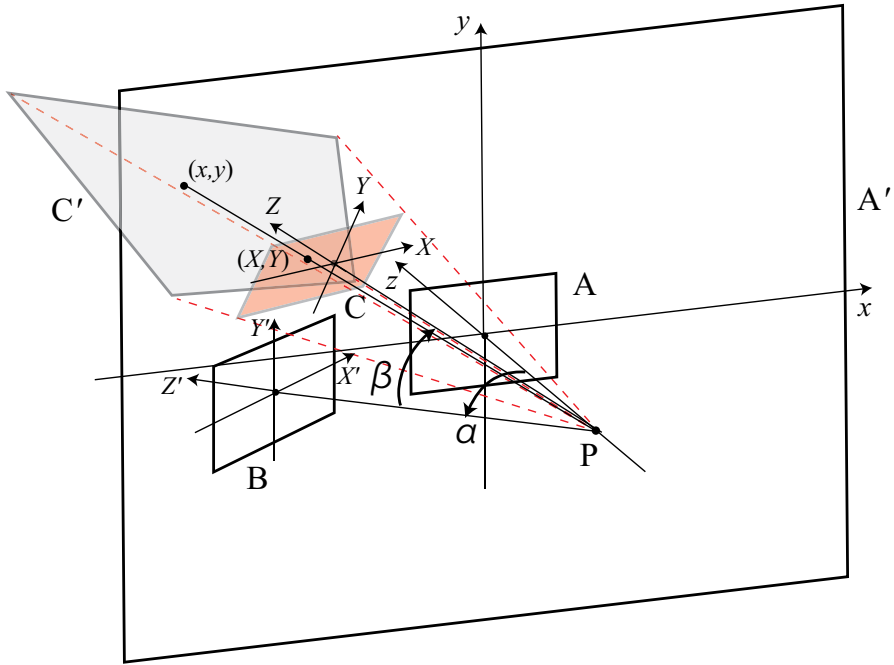


Fig. 3 Transformation from a wide-angle image to standard-lens images

Because the wide-angle image also conveys visual information outside A, we may recover the standard-lens images that would be obtained if we change the viewing direction. We next consider a method for achieving this.

We fix an (x, y, z) Cartesian coordinate system, place the lens center P at the origin, and place the original wide-angle image A' on the plane $z = d$. Because the image A' is parallel to the xy plane, we use symbol (x, y) to represent the 2D coordinate system on A' , as shown in Fig. 3.

Suppose that we change the view direction by rotating it around the y axis by angle α , and let B be the image frame corresponding to the standard lens. Let (X', Y', Z') be the coordinate system obtained from (x, y, z) by this rotation. Then, we have

$$X' = x \cos \alpha + z \sin \alpha, \tag{1}$$

$$Y' = y, \tag{2}$$

$$Z' = -x \sin \alpha + z \cos \alpha. \tag{3}$$

B is obtained from A by the same rotation. Hence, the plane containing B is represented by

$$z = \frac{d}{\cos \alpha} + x \tan \alpha. \tag{4}$$

Let $Q = (x, y, d)$ be an arbitrary point in image A. The line of sight from P to Q is represented by the vector $l(x, y, d)$, where l is a parameter. The value of l corresponding to the point of intersection of this line with the plane is obtained by substituting $l(x, y, d)$ into Eq. (4). Thus, we get

$$ld = \frac{d}{\cos \alpha} + lx \sin \alpha, \quad (5)$$

and hence

$$l = \frac{d}{(d - x \tan \alpha) \cos \alpha}. \quad (6)$$

Thus, the point of intersection is represented by

$$\frac{d}{(d - x \tan \alpha) \cos \alpha} (x, y, d). \quad (7)$$

This point is represented by the coordinate system (X', Y', Z') as

$$X' = \frac{d}{(d - x \tan \alpha) \cos \alpha} (x \cos \alpha + d \sin \alpha), \quad (8)$$

$$Y' = \frac{dy}{(d - x \tan \alpha) \cos \alpha}, \quad (9)$$

$$Z' = \frac{d}{(d - x \tan \alpha) \cos \alpha} (-x \sin \alpha + d \cos \alpha) = d. \quad (10)$$

Next, suppose that we change the view direction by rotating around the X' axis by angle β , and call the resulting image frame C. Let the resulting coordinate system be (X, Y, Z) . Then, we obtain

$$X = X', \quad (11)$$

$$Y = Y' \cos \beta + Z' \sin \beta, \quad (12)$$

$$Z = -Y' \sin \beta + Z' \cos \beta. \quad (13)$$

The plane containing C is represented by

$$Z' = \frac{d}{\cos \beta} + Y' \tan \beta. \quad (14)$$

The line of sight from P to (X', Y', d) is represented by $l'(X', Y', d)$, and the point of intersection of this line with the plane (14) satisfies

$$l' = \frac{d}{(Z' - Y' \tan \beta) \cos \beta}. \quad (15)$$

Consequently, the point of intersection is represented with respect to the (X, Y, Z) coordinate system as

$$X = \frac{dX'}{(d - Y' \tan \beta) \cos \beta}, \tag{16}$$

$$Y = \frac{dX'}{(d - Y' \tan \beta) \cos \beta} (Y' \cos \beta + d \sin \beta), \tag{17}$$

$$Z = d. \tag{18}$$

Equations (8), (9), (10) and (16), (17), (18) respectively represent the relation of the point (x, y, d) in the original image frame A' with the corresponding point (X, Y, d) in the rotated frame C.

To summarize the above arguments, we can construct the procedure to obtain the true appearance in the following way.

Suppose that we want to change the view direction by first rotating the viewpoint around the vertical axis by α and then by rotating the viewpoint around the horizontal axis by β , and then want to make the image corresponding to the frame C. For each point (X, Y) in C, we compute the corresponding point (X', Y') in B by the inverse transformation of Eqs. (16) and (17),

that is,

$$X' = \frac{dX}{(d + Y \tan \beta) \cos \beta}, \tag{19}$$

$$Y' = \frac{d}{(d + Y \tan \beta) \cos \beta} (Y \cos \beta - d \sin \beta), \tag{20}$$

and then compute the corresponding point (x, y) in A' by the inverse transformation of Eqs. (8) and (9), that is,

$$x = \frac{d}{(d + X' \tan \alpha) \cos \alpha} (X' \cos \alpha - d \sin \alpha), \tag{21}$$

$$y = \frac{dY'}{(d + X' \tan \alpha) \cos \alpha}. \tag{22}$$

We extract the color value of (x, y) from A' and embed it at point (X, Y) in C.

For actual computation, we cover frame C with a finite number of pixels. For each pixel (X, Y) , we compute the corresponding coordinates (x, y) , and determine its color by collecting the pixels near (x, y) and averaging their color values in A' .

4 Examples

This section presents the behavior of our method using wide-angle images. In those images, not only depths are exaggerated, but also the shapes of objects are distorted,

especially near the margins of the image frame. We show that both the depth exaggeration and the shape distortion are removed by our method.

In this paper we present the focal length of a lens with respect to a 35-mm full-frame camera. Thus, we hereafter omit the phrase “with respect to a 35-mm full-frame camera.”

The true appearance is recovered by changing the given image in such a way that the associated lens center coincides with the viewpoint from which we see the image. Therefore, strictly speaking, we need the location of the viewpoint. Instead of measuring the viewpoint, however, we assume that the image will be seen from the viewpoint corresponding to the lens center of the so-called “standard lens.” The standard lens is a lens with a focal length between 50 mm and 70 mm. It is widely accepted that the image taken by a standard lens gives a nearly faithful appearance of the original scene. This is probably because the focal length of a standard lens is close to the distance at which we usually see the image, or at least the viewpoint is not so far from the lens center, and hence our brain can correct the distortion. Thus, we convert the given wide-angle images to standard-lens images.

4.1 Example 1: lounge

Figure 4a shows the image of a lounge of the institute to which the author belongs taken by a wide-angle lens with a 12-mm focal length. Panels (b), (c), and (d) are recovered images assuming a focal length of 70 mm.

Figure 4b shows the recovered image, which assumes that the view direction is rotated 20 deg to the right and 30 deg upward. From this image, we can understand that the roof ventilation hatch is almost a square. In the original image (a), however, this hatch appears as a rectangle whose length in the depth direction is much larger than its width. This is a distortion effect of the wide-angle lens that can be removed by our method.

Figure 4c shows the image obtained if we rotate the view direction 40 deg to the right and 15 deg downward. From this image, we understand that the table is square and the chairs have a normal size for one person.

Figure 4d shows the image obtained when we rotate the viewpoint 40 deg to the left and 15 deg downward. From this image, we understand that the round tables are circular instead of elliptical, and the chairs have a normal size instead of being long seats like beds.

All the distortions we can see in (a) elongate the sizes of the objects in the depth direction. In this way, they create the impression that the depth is larger than the real size.

4.2 Example 2: utility room

Figure 5a shows an image of a utility room next to a bath taken by a wide-angle lens with 12-mm focal length. From this image, we have the impression that the center wall with the dark door meets the right wall at an obtuse angle much larger than 90 deg, and that the mirror on the left wall is larger in the horizontal direction than in



Fig. 4 Image of an institutional lounge: **a** wide-angle-lens image, **b** recovered image in the direction of 20 deg to the right and 30 deg upward, **c** recovered image in the direction of 40 deg to the right and 15 deg downward, **d** recovered image in the direction of 40 deg to the left and 15 deg downward

the vertical direction. From this image, we can recover images corresponding to those obtained with a 70-mm standard lens.

Figure 5b shows the scene obtained by 30-deg rotation to the right and 20-deg rotation downward. From this image we can understand that the center wall meets the right wall orthogonally.

Figure 5c–e provides the standard-lens images we obtain when we change the view direction 15 deg to the left and respectively 20, 10, and 0 deg upward. From these three images we understand that the width of the mirror is smaller than its height.

Figure 5f is a standard-lens image we get when we rotate our viewpoint 40 deg to the left and 10 deg downward. We can understand that the soap-bubble-toy holder is circular, although it seems elliptical in Fig. 5a.

The false impression about the wall angle, the mirror shape, and the toy-holder shape together contribute to an exaggerated room depth.

4.3 Example 3: house

Figure 6a represents an image of a house taken with a 12-mm wide-angle lens. From this image, we gain an impression that the left and right walls meet at an acute angle instead of a right angle, and that the left and right roads are almost parallel. Images with



Fig. 5 Image of a utility room: **a** wide-angle-lens image, **b** recovered image in the direction of 30 deg to the right and 20 deg downward, **c** recovered image in the direction of 15 deg to the left and 20 deg upward, **d** recovered image in the direction of 15 deg to the left and 10 deg upward, **e** recovered image in the direction of 15 deg to the left, **f** recovered image in the direction of 40 deg to the left and 10 deg downward

similar tendencies can be found in advertisements for real estate properties. However, these impressions are false. Figure 7 shows the plan view of the house and the roads, where the dot and the arrow represent the position and orientation, respectively, of the camera by which the image was taken. The walls meet at a right angle, and the left and right roads also meet at a right angle.

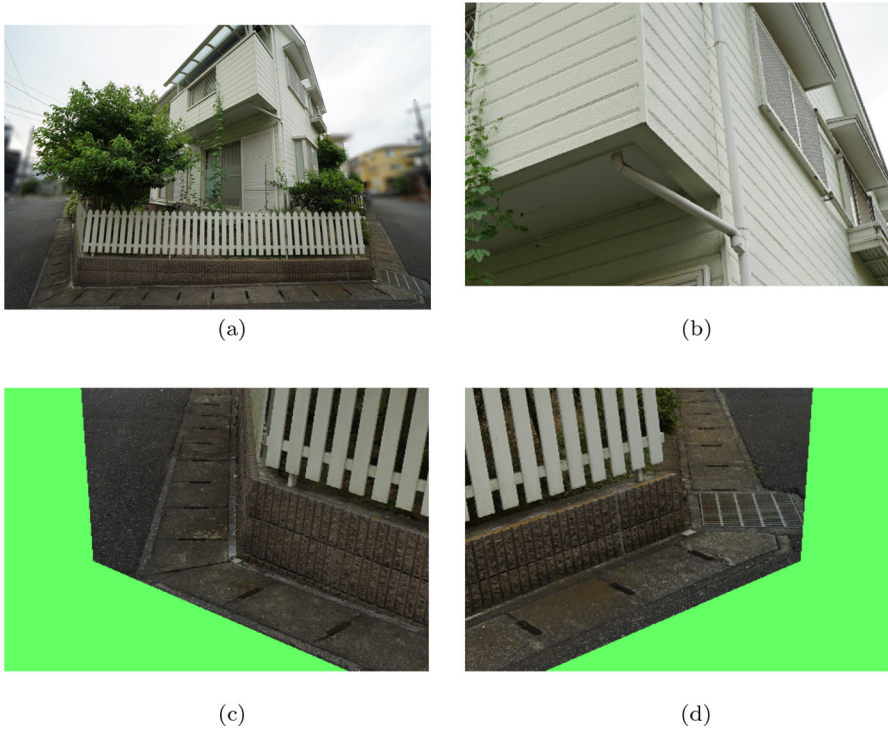


Fig. 6 Image of a house: **a** wide-angle-lens image, **b** recovered image in the direction of 15 deg to the right and 20 deg upward, **c** recovered image in the direction of 40 deg to the left and 25 deg downward, **d** recovered image in the direction of 40 deg to the right and 25 deg downward

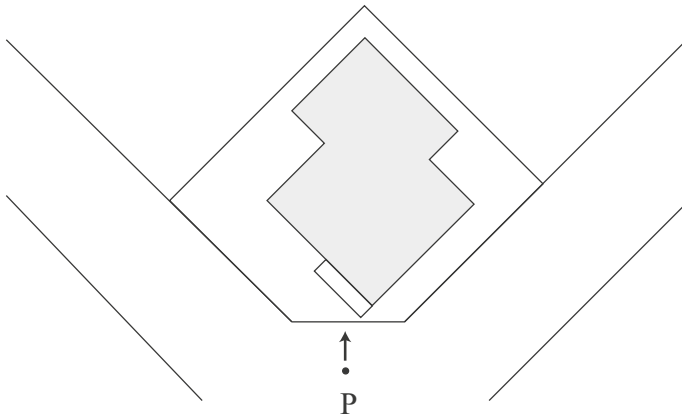


Fig. 7 Plan view of the house and roads in Fig. 6

Figure 6b is a 70-mm standard-lens image obtained by rotating the view direction 15 deg to the right and 20 deg upward. From this image we understand that the left and right walls meet at a right angle.

Figure 6c is a standard-lens image obtained by rotating the viewpoint 40 deg to the left and 25 deg downward. Figure 6d is the image obtained by a rotation of 40 deg to the right and 25 deg downward. From these images, we understand that the front road meets the left and right roads with obtuse angles much larger than 90 deg, and that the left and right roads are not parallel. We can also generate a video of the standard-lens appearance of the scene by continuously changing the view direction. When we see the continuous change from Fig. 6c, d, we understand that the roads meet at almost right angles.

5 Discussion

5.1 Lens-center information

Our method requires knowing the position of the lens center from which the image is taken. However, the lens-center information is usually not included when images are stored. The best way to retain that information seems to be to change our technology culture to record the lens-center information in the digital image file.

Depth-exaggerated images are widely used in advertisements, but the exaggeration is understood to not be intentional. Photographers just want to show a wide range of space in a single image, and they also want to show faithful appearances if possible. Because we are proposing a method for more faithful representation, we can expect that photographers will be willing to provide the lens-center information.

For example, a hotel-reservation web site company can record and attach the lens-center information to all images posted in the web site, and install our method so that users can see the true appearance of the room by changing the view direction. This is the ideal way to apply our method.

From the user's point of view, however, we want to recover the true appearance even if the lens-center information is not given. In that case, we need to estimate the lens center from the image itself, which is often possible.

Ordinary rooms are rectangular, and hence have three groups of parallel lines as the intersections of the walls, floor and ceiling, and they are orthogonal to each other. A group of parallel lines in 3D space, unless the lines are parallel to the image plane of the camera, forms a set of radial lines emanating from a common point in the image. This point is called the *vanishing point*. Images of a rectangular room are classified into three types of perspective projections called *one-point*, *two-point*, and *three-point perspective projections* according to how many groups of parallel lines form the vanishing points. For the three-point perspective projection, the lens center can be determined as the intersection of three spheres having diameters defined by pairs of vanishing points [2, 9, 15]. For the two-point perspective projection, the lens center can be determined as the intersection of the sphere having the diameter connecting the two vanishing points and the line passing through the center of the image perpendicular to the image plane [15]. For the one-point perspective projection, the lens center

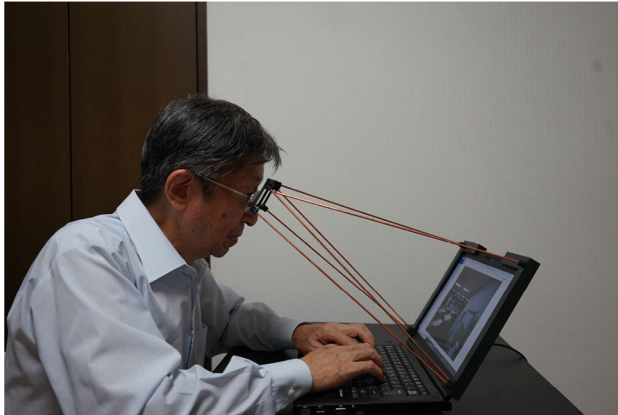


Fig. 8 Viewpoint guide

cannot be determined from the vanishing point, and hence we need to use some other information to estimate the lens center.

The lens center can also be estimated if we have knowledge about objects in the images [10]. Machine learning techniques can also be used to estimate the lens center [19].

It is interesting to note that the lens center can be easily determined from an image taken by a 360-deg camera, if the image is recorded as the cylindrical projection. For this kind of an image, we make a cylinder by gluing the left and the right edges of the image. Then, the lens center is at the center of the cylinder. This is true no matter in what size the image is printed.

5.2 Guiding the viewpoint

As we have discussed, standard-lens images can be approximately used to match the lens center and the user's viewpoint. We can give the true appearance more faithfully if the user's viewpoint matches the associated lens center exactly. One method for that purpose is to provide eyeglass-like equipment, such as that shown in Fig. 8, to guide the eyes. However, such equipment might be undesirable because it inconveniences the user.

A more flexible method might be to measure the user's viewpoint online and adjust the size and the position of the image on the display screen dynamically in a way such that the lens center matches the user's viewpoint. This method requires some equipment to detect the user's eye position.

There is a trade-off between the use of equipment to measure the user's eye position, and the use of fixed-focal-length images, such as standard-lens images. The former is more precise but constrains the users. The latter requires no action by the users but is less precise. It is one of our future research topics to investigate which option is more practical.



Fig. 9 Picture in Fig. 6c rotated 10 deg counterclockwise

A head-mounted display, if it is available, can also solve the viewpoint problem. If a user wears this equipment, the eyes are fixed with respect to the display screen. Therefore, there is no need to measure the eye position.

5.3 Direction of gravity

We are sensitive to the orientation of the image in interpreting it as 3D structures [5]. In particular, the direction of gravity seems important. For example, the house in Fig. 6 has a white fence composed of vertical pickets. In the recovered images in Fig. 6c, d, however, the fence looks tilted instead of vertical. This side effect occurs because the lens axis of the camera was not horizontal. The original image in Fig. 6a was taken by a camera whose axis was directed upward. If we take an image by a horizontally directing camera, the fence would be nearly vertical in the recovered image.

For example, if we rotate the image in Fig. 6c by 10 deg counterclockwise, we get the view in Fig. 9. The orientation of the scene seems more natural in the sense that the fence looks vertical. It would be more preferable if we could adjust the orientation of the recovered image in a way such that the vertical direction of the scene is kept vertical in the image. This is possible if we know the orientation of the axis of the camera with which the image was taken. It might be also possible to estimate the camera axis orientation from the image itself. This is one of our future research topics.

5.4 Rectangularity preference

An image comprises 2D information, while a scene includes depth to comprise 3D data; the depth of a scene is lost in its image. Hence, from a mathematical point of view, a single image can represent various 3D scenes [6, 14]. However, we usually perceive a unique 3D scene up to a scale when we see an image. This means that



Fig. 10 Wide-angle images of a house entrance: **a** one-point perspective projection, **b** two-point perspective projection

our brains implicitly assume a certain class of scenes. A typical assumption is the rectangularity assumption. It is well known that the human brains prefer rectangular structures when they interpret images [11, 12]. Because of this preference, we can perceive rectangular rooms from wide-angle images even though shapes of objects are distorted in the image.

One-point perspective projection has a remarkable feature in the sense that rectangular corners of the room are perceived as rectangular without any distortion. This class of image is obtained when one group of parallel lines is parallel to the lens axis, and the other two groups of parallel lines are parallel to the image plane of the camera. The vanishing point is located at the center of the image, and the lens center is on the line passing through the vanishing point and perpendicular to the image plane. However, the distance of the lens center from the image plane cannot be determined from the vanishing point information. No matter which point on the line we locate the center of projection, the image is a mathematically correct projection of some rectangular room. Indeed, we perceive a rectangular room more naturally when we see such an image than when we see a two-point or three-point perspective projection.

To show this visual effect, Fig. 10 shows two images of the same scene; (a) is a one-point perspective projection and (b) is a two-point perspective projection. Both of them were taken by a 12-mm wide-angle lens. We can see that rectangular corners look rectangular in (a), while they are distorted in (b). For example, the steps of the staircase and the gate to the next room look considerably distorted from rectangles in (b). For this reason, one-point perspective projections are widely used in the advertisement of real estate properties. Thus, the users must be careful.

6 Concluding remarks

This paper points out that the depth-exaggeration effect of images used in advertisements is an optical illusion caused by the difference of the viewpoint and the lens center. In response, we developed a method to remove the illusion and thus provide a faithful appearance of the scene. Our method requires only a image and the associated

lens center, and hence can be put to practical use as long as the lens center information is stored with the photograph or estimated from the image.

Depth-exaggerated images are widely used, but this situation is not desirable because consumers cannot obtain accurate information when deciding on a purchase. However, it may not be practical to prohibit exaggerated images by law. Indeed, this is related to the issue of “freedom of expression.” Photographers use wide-angle lenses to make special visual effects when they want to represent more of the actual space in a photograph. Hence, this problem is controversial.

However, providing faithful visual information about commercial goods is important and useful to both the advertisers and the consumers. Thus, it may not be difficult to establish a new culture in which images are accompanied with the associated lens center. Once this information is available routinely, our method can be used widely. Even if this lens-center information is not established, we can use various methods to estimate the lens center.

Remote shopping is becoming more and more popular, especially due to the COVID-19 pandemic. Hence, it is important to present faithful information of offered goods by images in order to raise market reliability. The proposed method may contribute to this goal.

Acknowledgements The author would like to express his thanks to Mr. Michiro Otsubo of the Ikuta Intellectual Property Division of Organization for the Strategic Coordination of Research and Intellectual Properties of Meiji University for his constructive discussions. This work is supported by a Japan Science and Technology Agency A-STEP Tryout Grant (JPMJTM20MU) and two Japan Society for the Promotion of Science KAKENHI Grants (21H03530 and 21K19801). Computational experiments were conducted on a MIMS SMP Computer at the Ministry of Education, Culture, Sports, Science and Technology, Joint Usage/Research Center CMMA, Meiji University.

Declarations

Conflict of interest The author declares that he has no conflict of interest.

References

1. Bressan, P., Garlaschelli, L., Barracano, M.: Antigravity hills are visual illusions. *Psychol. Sci.* **14**, 441–449 (2003)
2. Cipolla, R., Drummond, T., Robertson, D.: Camera calibration from vanishing points in images of architectural scenes. *Proc. Br. Mach. Vis. Conf.* **1999**, 382–391 (1999)
3. Gregory, R.L.: *The Intelligent Eye*. Weidenfeld & Nicolson, London (1970)
4. Iida, Y., Kitamura, R. (eds.): *Traffic Engineering* (in Japanese). Ohmsha, Tokyo (2008)
5. Mamassian, P., Landy, M.S.: Observer biases in the 3D interpretation of line drawings. *Vis. Res.* **38**, 2817–2832 (1998)
6. Marr, D.: *Vision*. W. H. Freeman and Company, New York (1982)
7. Martinez-Conde, S., Macknik, S.L.: *Champions of Illusion*. Scientific American/Farrar, Straus and Giroux, New York (2017)
8. Ninio, J.: *The Science of Illusions*. Cornell University Press, Ithaca (2001)
9. Orghidan, R., Salvi, J., Gordan, M., Orza, B.: Camera calibration using two or three vanishing points. In: *Proceedings of the Federated Conference on Computer Science and Information Systems*, pp. 123–130 (2012)
10. Penate-Sanchez, A., Andrade-Cetto, J.: Exhaustive linearization for robust camera pose and focal length estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 2387–2400 (2013)

11. Perkins, D.N.: Visual discrimination between rectangular and nonrectangular parallelepipeds. *Percept. Psychol.* **12**, 293–331 (1972)
12. Perkins, D.N.: Compensating for distortion in viewing picture obliquely. *Percept. Psychol.* **14**, 13–18 (1973)
13. Robinson, J.O.: *The Psychology of Visual Illusion*. Dover Publications Inc., Mineola (1998)
14. Sugihara, K.: *Machine Interpretation of Line Drawings*. MIT Press, Cambridge (1986)
15. Sugihara, K.: *Mathematics of 3D Illusion* (in Japanese). Kyoritu-Shuppan Publisher, Tokyo (2006)
16. Sugihara, K.: Misperception of road curvature due to slope change. *Jpn. J. Ind. Appl. Math.* **38**, 379–389 (2021)
17. Sugihara, K.: True views from depth-exaggerated images. *Proc. Int. Display Workshops* **28**, 1038–1041 (2021)
18. Tuinashi, S.: *Study on the Inclination Perception of Slopes* (in Japanese). Kazama Shobo, Tokyo (2013)
19. Workman, S., Greenwell, C., Zhai, M., Baltenberger, R., Jacobs, N.: DEEPFOCAL: a method for direct focal length estimation. *Proc. IEEE Int. Conf. Image Process.* **2015**, 1369–1373 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.