True Views from Depth-Exaggerated Images

Kokichi Sugihara[1]

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
Keywords: Depth perception, optical illusion, wide-angle lens, true appearance, viewpoint.

ABSTRACT

Wide-angle lenses are commonly used to show a large part of a scene such as a hotel room by a single image. However, images taken with such lenses exaggerate depth. We present a method for converting such images into images that look as if we really stand at the same position as the camera and see the scene directly. This method can remove depth exaggeration and thus give faithful information about the objects in the scene.

1 Introduction

Optical illusions can cause misunderstandings [1, 2]. It is thus important to understand the mechanism of optical illusions and develop methods to remove or decrease them. Here, we consider optical illusions in photographs used for advertisement.

For example, images on hotel reservation websites are often depth-exaggerated that thus mislead the viewer about room size. This effect is not intentional. Such images are taken using a wide-angle lens to show a large part of a room in a single image [3].

An image taken with a camera provides correct information about the appearance of a scene, because imaging is a purely scientific process based on optical physics. Nevertheless, the human brain misinterprets depth information because of an optical illusion. In this paper we construct a method for avoiding such illusion and thus giving faithful appearance of the scene.

The rest of this article is organized as follows. The source of the optical illusion in wide-angle images is described in Section 2. A method for removing the illusion is proposed in Section 3. The proposed method is validated using computational experiments in Section 4. Possible applications and extensions of our method are discussed in Section 5. The concluding remarks are given in Section 6.

2 Source of Illusion

One of the main sources of the illusion considered here is inappropriate viewpoint [3]. The viewpoint from which an image should be viewed is usually not specified. Images are thus viewed from an arbitrary viewpoint that is convenient to the viewer. However, what is perceived in an image depends on the viewpoint.

Fig. 1 shows how the depth illusion occurs. Suppose that we fix the lens center at P and take a photograph of object A. The object is projected onto imaging plane S. Because for a wide-angle lens the distance between P and S is short, the viewer sees the image from viewpoint Q, which is much farther from S than is P. The image does not contain explicit depth information and thus the absolute distance to the object cannot be determined from a single image. Light travels in straight lines and hence we see the object somewhere along the red lines. Suppose that we know the dimensions of object A. If we extend the image along the red lines to a point at which these dimensions are reached, we will perceive the object to be at that point (e.g., B in the figure). We thus interpret the object as being farther away and thicker than it actually is.

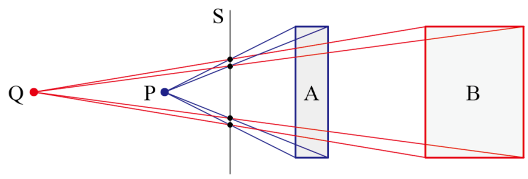


Fig. 1 Imaging of an object and perception by a viewer of the resulting image.

If we view the image from P, not Q, we will see the true appearance of the object (i.e., without illusion). This indicates that for objects to appear as they really are, the image must be viewed from the location of the lens center at which it was taken.

Based on this principle, we develop a method for removing the illusion and recovering the true appearance of objects in a single image.

3 True View Recovery

Consider image A’ in Fig. 2, which was taken using a wide-angle lens. Assume that we know the position of the lens center (point P).

Let A be the image frame that would be obtained if we took a photograph with a standard lens (i.e., a 35- to 85-mm lens for a 35-mm full-frame camera) at P facing the same direction. In this figure, we fixed the lens center P and adjusted the frame size according to the focal length of the lens. A standard lens has a longer focal length and hence frame A is smaller than frame A’.

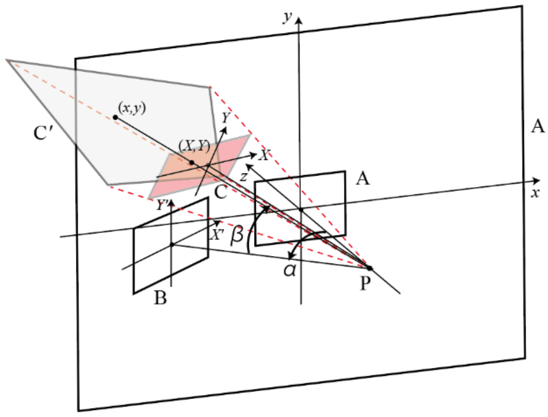


Fig. 2. Correspondence between images obtained using wide-angle and standard lenses.

An image corresponding to that taken using a standard lens can be considered as an approximation of the true appearance of the scene because the image is seen from a viewpoint that is close to the lens center. Therefore, if we do not change the view direction, image in frame A is an approximation of the true appearance of the scene.

Next, let us consider the case where we change the view direction. As shown in Fig. 2, we create an (x, y, z) Cartesian coordinate system and place the lens center P at the origin. The image plane is parallel to the xy plane. We use the (x, y) coordinate system to represent positions on frame A′. Let d be the distance from P to frame A′.

Suppose that at P we rotate the view direction around the vertical axis by angle *α* to get frame B, and then rotate around the horizontal axis by angle *β* to get frame C. Let (X', Y') be the coordinate system obtained from (x, y) by the first rotation, and let (X, Y) be the coordinate system obtained by the second rotation. Then, for each point (X, Y) on C, we get

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

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

and

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

$$y = \frac{d\,Y'}{(d + X'\tan\alpha)\cos\alpha}$$

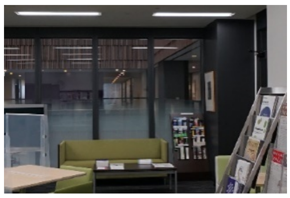
Using these equations, we can get point (x, y) on A’ that corresponds to point (X', Y') on C.

Given rotation angles *α* and *β*, we scan frame C and at each point (X, Y) we get the corresponding coordinates (x, y) and retrieve the color value from frame A’. Thus, we can construct the image in frame C. This image is an approximation of the true appearance that would be obtained if we had rotated our head horizontally by *α* and vertically by *β*.

4 Examples

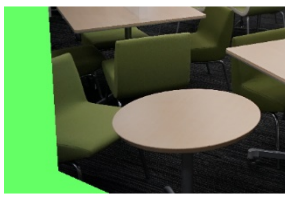
We now show some examples of our method. Fig. 3(a) shows an image of a lounge taken using a wide-angle lens with a focal length of 12 mm. This 12 mm is the distance between the image plane and the lens center for an image with the width of 36 mm. Therefore, the lens center of this image is at a distance of 1/3 the width, which is very short. When viewed from a natural distance, the image gives the impression that the room is much larger than it actually is.

(a) Original wide-angle image

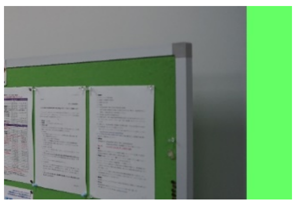


(b) Center

(c) Square table



(d) Circular table



(e) Sheets of paper

Fig. 3. Examples of true view recovery using proposed method

Fig. 3(b) to (e) show the output images of our method corresponding to a lens with focal length of 70 mm (i.e., a standard lens). They would be obtained if we viewed the same scene from point P.

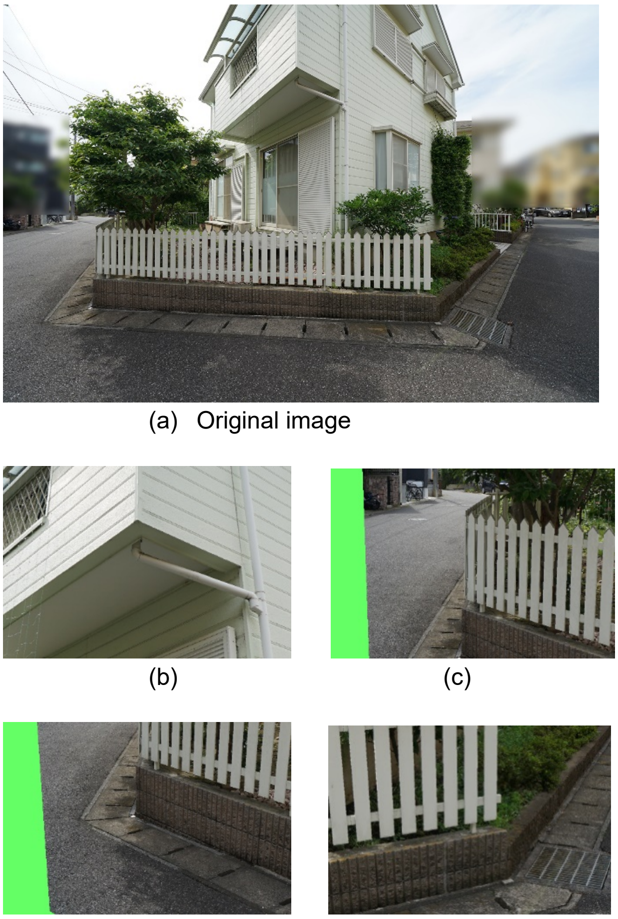
Fig. 3(b) shows the scene obtained by the camera with the same orientation. This image gives a more faithful impression of the depth of the room.

Fig. 3(c) shows the scene obtained with 20-degree rotation to the left and 15-degree rotation downward. We can understand that the table is square, although it looked non-square in the original image.

Fig. 3(d) shows the scene obtained with 40-degree rotation to the left and 20-degree rotation downward. We understand that the table is circular, although it looked elliptic in the original image.

Fig. 3(e) shows the scene obtained with 40-degree rotation to the right. We can understand that the sheets of paper are in portrait orientation, although they looked in landscape orientation in the original image.

Fig. 4(a) shows another example of an image taken by a wide-angle lens. Here is a house. From this image, we have an impression that the left and right walls of the center house meet in a sharp angle instead of the rectangle, and that the left road and the right road are almost parallel. In reality, on the other hand, the walls meet at right angle as a usual house and the left wall is parallel to the left road and the right wall is parallel to the right road. Consequently, the left and right roads meet at a right angle. The distortion we perceive from Fig. 4(a) is what we usually experience when we see the advertising images of real estates.



(a) Original image

(b) (c)

(d) (e)

Fig. 4. Image of a house taken by a wide-angle lens (a), and the recovered true appearances ((b) to (e)).

Fig. 4(b) to (e) show the images corresponding to a lens with focal length of 70 mm obtained by our method.

Fig. 4(b) shows the scene that would be obtained when we turn our head 30 degrees upward. From this image, we understand that the walls meet at right angles.

Fig. 4(c) shows the scene that would be obtained when we rotate our head 45 degrees to the left and 10 degrees downward. From this image we understand that the left and front white fences meet at about 45 degrees, although it looks almost 90 degrees in the original image.

Fig. 4(d) corresponds to the scene that would be obtained when we rotate our head 45 degrees to the left and 30 degrees downward. From this image, we again understand that the left road and the front fence meet at about 45 degrees, although the original image gives an impression that they meet at right angle.

Fig. 4(e) shows the scene that would be obtained when we rotate our heads 30 degrees to the right and 20 degrees downward. From this image we understand that the right road and the front fence meet at about 45 degrees, although it looks 90 degrees in the original image. From Figs. 4(c), (d) and (e), we can convince that the left and right roads meet at right angle.

5 Discussion

The proposed method can recover the true appearances of a scene from depth-exaggerated images. The fundamental observation is that the optical illusion created by a wide-angle lens is removed by matching the viewpoint and the lens center. There are two possible implementations of the method. One is to guide the user's viewpoint to the lens center at which the image was taken. The other is to convert the image in such a way that the associated lens center coincides with the user's viewpoint. In the experiments presented here, we converted an image into images taken with a standard lens, which can be considered as approximation of true appearance because they are naturally viewed from somewhere near the lens center.

Our method requires only a single image and the lens center at which the image was taken. However, images do not necessarily carry explicit information about the lens center (i.e., the focal length of the lens). In those cases, the focal length can be estimated from vanishing points [4, 5], from images of known objects [6], or using machine learning [7]. For ordinary images used at web sites, for example, the information about the lens center is lost. This fact might be a bottle neck when we want to apply our method to real situations.

However, we may expect that the owners of the hotel reservation websites, for example, are willing to provide the information about the focal length if it is necessary, because they want to show their commodities as faithfully as possible in order to avoid troubles with consumers. They use wide-angle images only because they want to show a wider range of the appearance by a single image. Indeed, one possible application of our method is that the hotel reservation websites provide the option by which we can see the true appearances of the rooms.

In our method, we can change the view direction arbitrarily; still, we can recover the true appearance of objects in the original images. There are many methods for recovering images by changing the viewpoint [8, 9]. However, they require or must estimate the three-dimensional shapes of objects. In case we have those data, it is possible to change the viewpoint and view direction arbitrarily. In contrast, our method cannot change the viewpoint because depth information about the scene is not available.

6 Concluding Remarks

We proposed a method for recovering the true appearance of a scene from a wide-angle image, which gives the wrong impression about the sizes and distances of objects in the scene. We can recover the appearance for the same orientation or change the view direction arbitrary. The true appearance can be recovered as long as it is given in the original image.

Wide-angle images are commonly used in advertisements, such as hotel reservation and rental property websites. The proposed method can recover a more faithful appearance and thus improve the reliability of such advertising images. So, it can be applied to various types of internet shopping sites.

Acknowledgments

The author would like to express his thanks to Mr. Michiro Otsubo of 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 JST A-STEP Tryout Grant (JPMJTM20MU) and a JSPS KAKENHI Grant (JP21H03530). Computational experiments were conducted on a MIMS SMP Computer at the MEXT Joint Usage/Research Center CMMA, Meiji University.

References

[1] R. L. Gregory, “The Intelligent Eye,” Weidenfeld & Nicolson, London, 1970.

[2] J. O. Robinson, “The Psychology of Visual Illusion,” Dover Publications, Inc., New York, 1998.

[3] K. Sugihara, “Mathematics of 3D Illusion,” Kyoritu-Shuppann Publisher, Tokyo, 2006.

[4] R. Cipolla, T. Drummond and D. Robertson, Camera calibration from vanishing points in images of architectural scenes, “Proceedings of the British Machine Vision Conference 1999,” pp. 382-391.

[5] R. Orghidan, J. Salvi, M. Gordan and B. Orza, Camera calibration using two or three vanishing points, “Proceedings of the Federated Conference on Computer Science and Information Systems,” 2012, pp. 123-130.

[6] A. Penate-Sanchez and J. Andrade-Cetto, Exhaustive linearization for robust camera pose and focal length estimation, “IEEE Transactions on Pattern Analysis and Machine Intelligence,” vol. 35 (2013), pp. 2387-2400.

[7] S. Workman, C. Greenwell, M. Zhai, R. Baltenberger and N. Jacobs, DEEPFOCAL: A method for direct focal length estimation, Proceedings of IEEE International conference on Image Processing 2015, pp. 1369-1373.

[8] G. Wolberg and S. Zokai, Image registration for perspective deformation recovery, “Proceedings of SPIE, Automatic Target Recognition X,” Orlando, April 2000. https://doi.org/10.1117/12.395570

[9] R. Hu, N. Ravi, A. Berg and D. Pathak, Worldsheet: Wrapping the world in a 3D sheet for view synthesis from a single image, arXiv.org, 2021. https://arxiv.org/pdf/2012.09854