Journal of Information Technology and Engineering
Vol. 3 No. 1 (June, 2018)

Shape Constancy Computation Based on Visual Perceptual Theory

Bing LI, Xi LI, De XU and Song-He FENG
Institute of Computer Science, Beijing Jiaotong UniversityBeijing, 100044, China
E-mail: binggege@people.com.cn

Received: 13th March 2017 Revised: 24th April 2017 Accepted: 10th October 2017

Abstract: *Shape constancy is one of the perceptual constancy in visual psychology. This paper presents a computable model for shape constancy based on visual perception theory. In this approach, we use the depth information in 2-D image and the computable model for size constancy to get the perceptual size of each part of the object; and then the object's shape, which has transformed as the result of depth in 2-D image, can be recovered. And in this paper, visual perception theory is used to solve computer vision problems. This method may be a novel way to solve some difficult problems in this field in the future. The experiments show that our algorithm is effective.*

Key words: *Shape Constancy, Visual Perception, Perceptual Constancy.*

1. INTRODUCTION

According to geometrical optics theory, the image on human retina changes with the relative position between the observer and the object. But however it changes; we perceive a fixed object that has constant size, constant shape, constant color and constant brightness. Human beings don't perceive objects only according to the images on his retinas. This phenomenon is called perceptual constancy in visual psychology theory [1].

Our perception of objects is far more constant or stable than our retinal images. Retinal images change with the movement of the eyes, the head and our position, together with changing light. If we relied only on retinal images for visual perception we would always be conscious of people growing physically bigger when they came closer, objects changing their shapes whenever we moved, and colors changing with every shift in lighting conditions. Counteracting the chaos of constant change in retinal images, the visual properties of objects tend to remain constant in consciousness. We are not usually conscious of people appearing to get bigger as they approach us or of things appearing to change shape according to the angles from which we view them.

Human vision uses perceptual operations to achieve many goals in complex environments. Visual perceptual psychology has been identified as the major aspect in these operations. It has been considered as the key contributor to the efficiency of the vision system. In recent years, visual perception based computing models has received more and more attention in computer vision. The perceptual organization theory is most frequently applied in computer vision and image process [2,3]. But only a few researches are based on perceptual constancy. And most of these researches are mainly based on color constancy and brightness constancy, like optical flow estimation which is based on image brightness constancy [4,5]. Fewer researches are on the size constancy and shape constancy. I do my utmost to search some papers about shape constancy used in computer vision, I only find that Mr. Qigang Gao has studied shape constancy [6,7]. But Mr. Qigang Gao just used shape constancy principle to propose other computer vision problem. So almost no one did any researches on shape constancy principle itself. The shape constancy principle is so important in human vision that it is necessary to study it and to use it in computer vision. The traditional researches in computer vision mostly are based on the physical model, which less considers the human perception. So it maybe gets some good unimaginable results to introduce the visual perception theory to computer vision. In this paper, we do some researches on shape constancy principle and propose a computable model for shape constancy. We use this model to recover the object's shape, which changed in image caused by depth. This paper is organized as follows. In Section 2, the principle of perceptual shape constancy is described. In Section 3, we propose the computable model for shape constancy. The experiment results using this model are presented in Section 4. Section 5 concludes this paper.

2. PERCEPTUAL SHAPE CONSTANCY

The stability of our perceptual world in spite of the variations in physical stimulation is called perceptual constancy [6]. Psychologists classified the perceptual constancy into four categories: color constancy, brightness constancy, size constancy and shape constancy [6]. The shape perception is playing a key role in human vision in recognizing objects. So we ought to consider human perceptual factor in the research of computer vision.

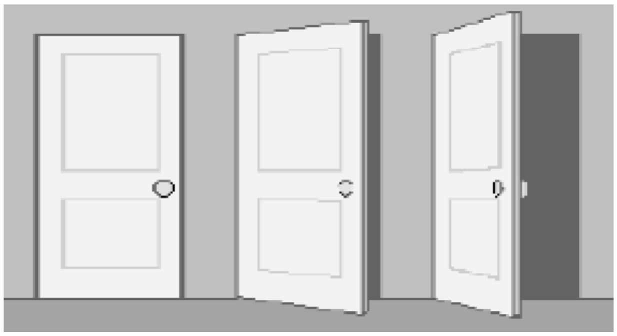


Figure 1: Perceptual shape constancy: The three doors have different shapes because of the view angles. But the shapes of them that human perceive are all rectangle.

Shape constancy is the tendency of an object to appear as the same shape, even when the view angle changes. Fig. 1 is an illustration of the perception of shape constancy. As one look at a door from different angle, the projected shapes of the door go through series transformations. Nevertheless, by our perception the door retains the same shape. Explanation of such phenomenon is relevant to the perception of depth and orientation [1]. So the depth of the image is the key factor for shape constancy. In the next section, we will consider the factors to recover the original shapes of the objects.

3. SHAPE CONSTANCY COMPUTATION

Shape constancy is not as easy as size constancy; the computation it involved bases on the depth of all parts of the object. Suppose observer can perceive the depth of all parts of the object accurately, the size of each part can be perceived accurately, consequently the shape of the object can also be perceived correctly. From the analyses above we can draw a conclusion that shape constancy can be derived from size constancy [1].

3.1 Perceptual Size Constancy

Size constancy is also one of the perceptual constancy properties. Although the size of object in the image on the human retina changes continuously, the size the observer perceives is fixed. Set trees as an example, there are many trees stand along the road with same heights. When you look down the road, you will find that the trees far from you look shorter than these near to you. But you can perceive that all these trees almost have the same heights. Psychologists called this phenomenon size constancy. It is the basic for shape constancy.

Psychologists have discovered computation theory of size constancy. They have got an expression to compute the perceptual size as

$$S = k \times A \times D \tag{1}$$

where S is the object's perceptual size, A is the angle of view, D is the perceptual depth of object, that is the depth that human beings perceived from camera or his eyes to the object, k is the zoom coefficient of human eyes or camera, it keeps invariable in a certain imaging process. The angle of view A can be represented by the size of object in the image.

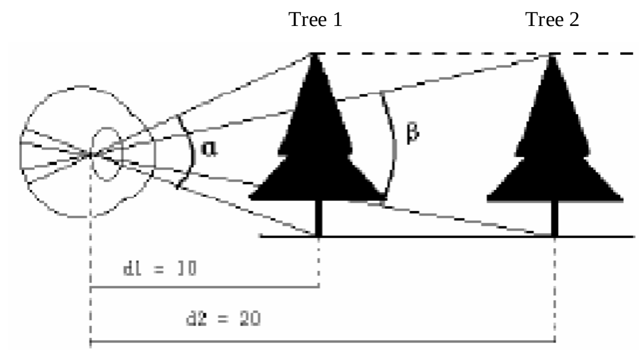


Figure 2: Example of size constancy computation: The images of two trees on human retina have different heights. Tree 1's image is higher than Tree 2's on the retina, but the heights of them, which human perceive, are same.

Fig. 2 shows an example of the computation theory of the size constancy. The distance between tree1 and the observer is 10 units, and the tree2 is 20-units-long from the observer, that is $d_1 = 10$, $d_2 = 20$. The view angles of the two trees are α and β. The view angle can be represented by the size of object in the image. We define the sizes of the trees in the image as S_1 and S_2. In addition, according to the pinhole imaging model, the size of the object in image is in inverse proportion to the distance between the object and the observer. Consequently, we can get equation (2) as follow.

$$\frac{\alpha}{\beta} = \frac{S_1}{S_2} = \frac{d_1}{d_2} = \frac{2}{1} \tag{2}$$

From equation (1) and equation (2), we can compute the perceptual size of the two trees, we define the perceptual size of the two trees as PS_1 and PS_2, the relationship between PS_1 and PS_2 is as follow

$$\frac{PS_1}{PS_2} = \frac{k \times \alpha \times d_1}{k \times \beta \times d_2} = \frac{\alpha \times d_1}{\beta \times d_2} = 1 \tag{3}$$

From the computation of the size constancy, we find the perceptual size of Tree1 is equal to the perceptual size of Tree2. Although they have different size images on human retina, the sizes human perceive are the same. From the equations above, the distance between object and observer is the key variable in the computation. This distance is just the perceptual depth in the 2-dimision image. We will discuss how to get the perceptual depth in 2-dimision in the next section.

3.2 Perceptual depth Computation

The perceptual depth of image is the most important factor in size and shape constancy computation. There have been many methods to estimate depth of image. According to the difference of camera number and image number, these methods can be classified into 3 categories. First, multicameras and multi images method, that is depth estimation from stereo based binocular matching [8]. Second, single camera and multi images method, that is depth estimation from defocus based single image [9, 10]. Third, single camera and single image method [11, 12]. All these methods based on the imaging model of camera. In this paper, just for validating the computation of shape constancy, we will use a simple single camera and single image model to estimate the image depth [13]. The model is generally described as follow.

According to the pinhole imaging model as Fig. 3, the line *OO'* is the optical axis. So the point *O'* is the center point of this image. Point *P* and *U* are on the ground, which is a reference plane. We can find the point *U'*, which is the image of *U*, is the highest point of image. That is, points between *U* and *E* can't be shown in this image. From the imaging geometry in Fig.3, the Triangle POE is similar to the Triangle OP'O', that is

$$\Delta POE \sim \Delta OP'O' \tag{4}$$

We define the distance from *O* to the image plane, that is distance between *O* and *O'*, as focus distance and note it as *f*. Note the distance between *O* and *E* as *h*. The distance from point *P* to *E* is just the perceptual depth of *P*, and we note it as D_p. Thus we have

$$D_p = h \cdot f / |P'O'| \tag{5}$$

Because in a certain imaging process, *h* and *f* are fixed, we can use ε to represent $h\,f\,i$. $|P'O'|$ is the distance from point P' to the center O' in the image. Suppose the width of image is m and the height is n. we note the coordination of P' as (p'_x, p'_y) and the coordination of *O'* as (O'_x, O'_y). If we express in pixel unit, we have

$$|P'O'| = |P'_y - O'_y| = |P'_y - \frac{n}{2}| \tag{6}$$

From equitation (5) and (6), we can compute D_p as

$$D_p = \varepsilon / |P'_y - \frac{n}{2}| \tag{7}$$

This depth estimation method is very simple, but there must be reference ground in the image.

3.3 Shape Constancy Computation

The main reason of shape transformation in image is the size change of some of all parts of object. In order to recover the transformed shape in image, we must recover the size of all parts of object. So using the algorithm in 3.1 and 3.2, we

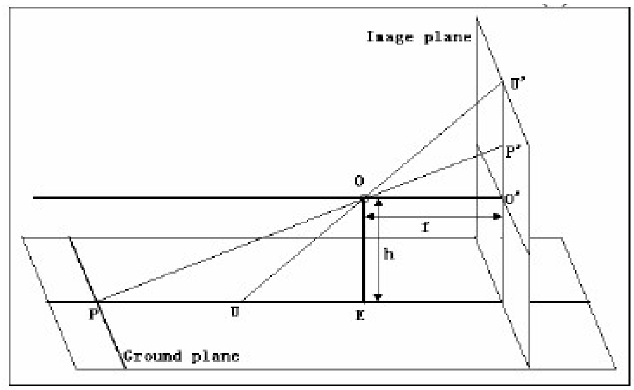


Figure 3: The pinhole imaging model with real ground: The points P, U and E are on the ground plane. O is the pinhole, and points O', P' and U' are on the image plane.

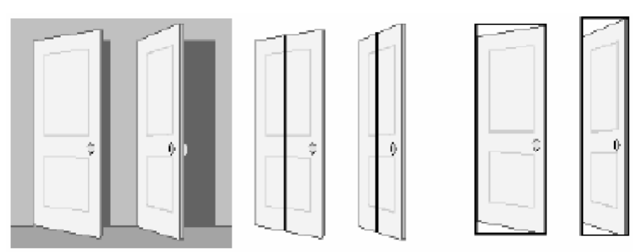
can compute the perceptual size of each part of object; consequently, we can also recover the shape of object. We define size of a certain object in image as a referenced size and note it as Sr, then compute other part's size, which we note it as S, according to the rate of perceptual size. From equation (1) and (7) ,we can get S as follow:

$$S = \frac{PS}{PS_r} \times S_r = \frac{k \times SI \times \varepsilon / |P_{sy} - n/2|}{k \times SI \times \varepsilon / |P_{ry} - n/2|} \times S_T = \frac{SI \times |P_{ry} - n/2|}{|P_{sy} - n/2|} \tag{8}$$

where S is the perceptual size referring to referenced object; *PS* and PS_r are the perceptual size, *SI* is the size of the object that waits for computation in image, P_{sy} is its Y coordinate and P_{ry} is the *Y*-coordinate of the referenced object.

4. EXPERIMENT RESULTS

Because our aim is just to validate our computation model, we suppose that the contours of objects have already been got. We use the two open doors in Fig.1 to experiment. The door ought to be rectangle, but it becomes trapeziums in Fig. 1 by reason of depth. We use our algorithm to get experiment result as Fig.4.



(a) (b) (c)

Figure 4: Experiment of Doors: The Wide Contour in the Picture (c) is the Perceptual Shape the two Doors.

In Fig.4, Picture (a) is original image of the two doors; Picture (b) is the contours of the doors. The wide contours in Picture (c) are the shapes of doors after recovery. We can find that the two doors recover to rectangle from trapeziums. We define the right border of each door as its referenced size, and then get some representative lines along the direction of depth change, for example the wide lines in Picture (b). Then compute these lines' perceptual size, according to these sizes we can get a proximal shape of object, just like shapes in Picture (c). Fig.5 shows another 2 experiments of library ground and road. The top 3 pictures show the experiment of library ground and the bottom 3 ones are the experiment results of road. The wide contour in Picture (c) of Fig.5 is the recovered shape of the ground and wide contour in Picture (f) is the recovered shape of the road. We have the experiments' detail data in Table 1.

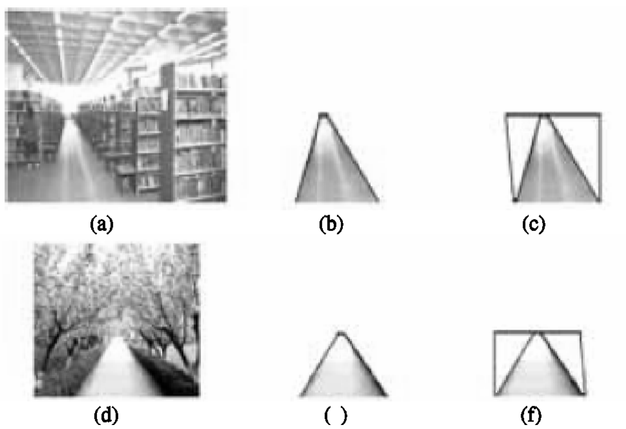


Figure 5: Experiments of the library ground and the road: The top 3 pictures show the experiment of library ground, The bottom 3 pictures are the experiment results of experiment of road.

CONCLUSIONS

The shape of object is important in human visual perception. So we introduce it to computer vision. In this paper, we propose a shape constancy computation method based on visual perception theory. Experiment results show that this method is effective. It can recover the shape transformation in a certain degree. Above all, we apply the perceptual constancy theory in computer vision. It is maybe another way to solve some difficult problems in this field.

Table 1
The Detail Date of the Experiments. Note: SI=size in Image; ISR = Rate of Sizes in the Image; D is the value of |P'y – n/2| in the Equation (6); PS=perceptual size; PSR = Rate of the Perceptual Sizes.

Exp	*Object*	*Border*	*SI*	*ISR*	*D*	*PS*	*PSR*
1st	Left	Left	140	0.87	69	162	1.01
	Door	Right	160	5	80	160	4
	Right	Left	142	0.86	69	168	1.00
	Door	Right	167	2	82	167	6
2nd	Lib	Top	13	0.11	10	114	1.04
	Ground	Bottom	109	9	88	109	6
3rd	Road	Top	48	0.13	11	336	0.96
		Bottom	349	8	77	349	3

REFERENCES

[1] I. Rock, Perception, Scientific American Books, Inc., 1984.

[2] S. Sarkar, "An Introduction of Perceptual Organization", in *IEEE Conf. on KIMAS*, 2003, pp. 330-335.

[3] S. Sarkar, and K. L. Boyer, "Perceptual Organization in Computer Vision: a Review and a Proposal for a Classificatory Structure", *IEEE Trans. on System, Man, Cybernetics*, 1993, pp. 382-399.

[4] Q. Zhou, D. Chelberg, Y. J. Zeng, and J. Starzyk, "Robust Optical Flow Estimation Using Invariant Feature", Southeastern Symposium on System Theory, 2001, pp. 263-267.

[5] M. J. Black, and P. Anandan, "A Framework for a Robust Estimation of Optical Flow", in *4th International Conf. on Computer Vision*, 1993, pp. 12-20.

[6] Q. G. Gao, "A Computation Model for Understanding Three-Dimensional View Space", in *IEEE Conf. on ICSMC*, 1996, pp. 941-946.

[7] Q. G. Gao, A. K. C. Wong, and S. H. Wang, "Estimating Face-pose Consistency Based on Synthetic View Space", *IEEE Trans. on System, Man, Cybernetics*, 1998, pp. 1191-1199.

[8] S. D. Ma, and Z. Y. Zhang, *Computer Vision*, Beijing: Science Press, 1998, pp. 52-93.

[9] P. Favaro, and S. Soatto, "A Geometric Approach to Shape from Defocus", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005, pp. 406-417.

[10] A. N. Rajagopalan, and S. Chaudhuri, "Depth Estimation and Image Restoration Using Defocused Stereo Pairs", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2004, pp. 1521-1525.

[11] S. H. Lai, C. W. Fu, and S. Chang, "A Generalized Depth Estimation Algorithm With a Single Image", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1992, pp. 405–411.

[12] A. Torralba, and A. Oliva, "Depth Estimation from Image Structure", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, pp. 1226-1238.

[13] A. M. Wu, and D. XU, "Generalized Perceived Depth Estimation from a Single 2D Image without any Priori Information", in *ICSCA' 2006*, 2006.