Optik 124 (2013) 2097–2100

Contents lists available at SciVerse ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Segmentation-based adaptive vergence control for parallel multiview stereoscopic images

Jianjun Lei*, Hailong Zhang, Chunping Hou, Liyuan Lin

School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China

ARTICLE INFO

Article history:
Received 20 January 2012
Accepted 17 June 2012

Keywords:
Multiview
Autostereoscopic display
Vergence control

ABSTRACT

The vergence control is important for multiview stereoscopic images captured by a camera array to acquire natural and clear stereoscopic effects. In this paper, we derive the basic characteristics of vergence control for parallel multiview stereoscopic images. Then we segment the stereoscopic images based on Mean Shift, and shift the images by the disparity of the object in the central region to achieve the adaptive vergence control. We have evaluated our method on Middlebury data sets and artificially synthesized images. Experimental results show that the proposed method is effective, and the generated stereoscopic images could reproduce the 3D scenes vividly in a multiview autostereoscopic display.

1. Introduction

Stereoscopic display technology is receiving increasing attention because it enables us to obtain the depth information of the objects and improve the perception of the distribution of the objects in the scene. There are many types of stereoscopic display technology including classic stereo systems that require glasses to more sophisticated multiview autostereoscopic displays that do not need glasses [1–3]. Multiview autostereoscopic displays provide highly realistic stereoscopic images and free-view navigation to viewers by generating various viewpoints of the scene [4–6].

There are two means to get the stereoscopic images for multiview autostereoscopic displays: toed-in camera array and parallel camera array [7,8]. Stereoscopic images captured directly usually have fusion problems which will lead to blurriness. It is essential to acquire natural and clear stereoscopic effects through vergence control. In the toed-in method, vergence control could be achieved by rotating all optical axes to converge on the key object. However, even with the addition of vergence control, inherent vertical disparity and keystone distortion still exists to result in visual fatigue. For the parallel method, vergence control is achieved by moving the CCD horizontally to alter the relative position with lens or shifting the captured images. Kwon et al. [9] implemented vergence control via disparity information for binocular stereoscopic images, but did not consider the multiview stereoscopic images. Deng et al. [10] shifted the images captured by a tri-view parallel camera to implement the vergence control. The processed stereoscopic images have both positive and negative disparity and are free from vertical disparities and keystone distortions. Nevertheless, the method only suits for fixed depth scenes because of its inflexibility of shifting manually or setting the shifting range in the camera.

In this paper, we derived the characteristics of vergence control in parallel multiview stereoscopic images. After the disparity of the object in the central region is obtained by means of Mean Shift, the images can be shifted by the disparity to realize the vergence control.

2. Characteristics of vergence control

Vergence control is indispensable to get natural and clear stereoscopic images using a parallel camera array. Taking an example for eight views, the principle of vergence control for multiview stereoscopic images is derived as follows. First, we establish the coordinate system based on view 1. Set $V_1, V_2, \ldots, V_8$ as the horizontal coordinate of view1 to view8, $V'_1, V'_2, \ldots, V'_8$ as the horizontal coordinate of view1 to view 8 after vergence control, $d_1, d_2, \ldots, d_7$ as the horizontal disparity of adjacent views, $d'_1, d'_2, \ldots, d'_7$ as the horizontal disparity of adjacent views after vergence control. Then, view 2 to view 8 are shifted by $\Delta x_1, \Delta x_2, \ldots, \Delta x_7$ to reduce the disparity as:

$$V'_{i+1} = V_{i+1} - \Delta x_i, \quad i = 1, 2, \ldots, 7 \tag{1}$$

$$d'_i = V'_{i+1} - V'_i = \begin{cases} (V_2 - \Delta x_1) - V_1 = d_1 - \Delta x_1, & i = 1 \\ (V_{i+1} - \Delta x_i) - (V_i - \Delta x_{i-1}) = d_i - \Delta x_i + \Delta x_{i-1}, & i = 2, 3, \ldots, 7 \end{cases} \tag{2}$$

* Corresponding author.
E-mail address: jjleitg@gmail.com (J. Lei).

0030-4026/$ – see front matter © 2012 Elsevier GmbH. All rights reserved.
http://dx.doi.org/10.1016/j.ijleo.2012.06.021

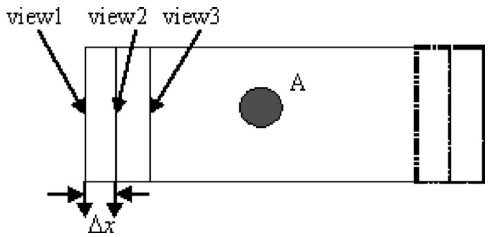

Fig. 1. Illustration of characteristics of vergence control.

It has been proved that the most comfortable and ideal stereoscopic effect occurs when $d_1 = d_2 = \cdots = d_7 = d$. Fig. 1 shows the desirable situation that the multiview stereoscopic images meet the condition $d_1 = d_2 = \cdots = d_7 = d$. We set zone A as the zero disparity object in view 1 and shift view 2, view 3 ... view 8 by $\Delta x, 2 \times \Delta x, \ldots, 7 \times \Delta x$ as the following equations:

$$V'_{i+1} = V_{i+1} - \Delta x_i = V_{i+1} - i \times \Delta x, \quad i = 1, 2, \ldots, 7 \tag{3}$$

$$\begin{aligned} d'_i &= V'_{i+1} - V'_i = (V_{i+1} - i \times \Delta x) - (V_i - (i-1) \times \Delta x) \\ &= d - \Delta x, \quad i = 1, 2, \ldots, 7 \end{aligned} \tag{4}$$

The zero disparity, negative disparity, and positive disparity zone will exist in the images with sufficient shifting distance.

3. Disparity estimation based on Mean Shift

The adaptive and robust image segmentation based on Mean Shift is insensitive to smooth and texture region and similar to analysis mechanism of human eye. In this paper, we segment the images based on Mean Shift, and then use the segmentation results to generate the smoothed dense disparity map.

An image is represented by spatial and color information. The eigenvector for spatial-color Mean Shift filters is defined as $X = (x^s, x^r)^T$. Because space and color is mutually independent, the multivariate kernel is defined as the product of two radially symmetric kernels [11].

$$K_{h_s,h_r}(x) = \frac{B}{h_s^2 h_r^3} k\left(\left\|\frac{x^s}{h_s}\right\|^2\right) k\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \tag{5}$$

where x^s is spatial information, x^r is the color feature, $k(x)$ the common kernel profile used in both two domains, h_s and h_r the employed kernel bandwidths, and B the corresponding normalization constant.

Every point is finally converged using the Mean Shift algorithm iteratively based on multivariate kernels. The initial points which converge to the same point are classified to the same region and then labeled.

After the segmentation, we get the initial disparity map through a window-based local method. We introduce the gradient information in light of [12] and combine it with BT (Birchfield and Tomasi) [13], so the cost function is finally designed as:

$$C(p, \bar{p}_d) = w * C_{BT}(p, \bar{p}_d) + (1 - w) * C_{GRAD}(p, \bar{p}_d) \tag{6}$$

where

$$C_{BT}(p, \bar{p}_d) = \sum_{c \in \{r,g,b\}} \min\{\bar{C}(p, \bar{p}_d), \bar{C}(\bar{p}_d, p)\} \tag{7}$$

$$\begin{aligned} C_{GRAD}(p, \bar{p}_d) = &\sum_{c \in \{r,g,b\}} |\nabla_x I_c(p) - \nabla_x I'_c(\bar{p}_d)|^2 \\ &+ \sum_{c \in \{r,g,b\}} |\nabla_y I_c(p) - \nabla_y I'_c(\bar{p}_d)|^2 \end{aligned} \tag{8}$$

$$\bar{C}(p, \bar{p}_d) = \max\{0, I_c(p) - I'_{c,\max}, I'_{c,\min} - I_c(p)\} \tag{9}$$

$$\bar{C}(\bar{p}_d, p) = \max\{0, I'_c(\bar{p}_d) - I_{c,\max}, I_{c,\min} - I'_c(\bar{p}_d)\} \tag{10}$$

where I, I' are the stereo pair, $p, \bar{p}_d$ are the two corresponding points in the stereo pair, ∇_x and ∇_y represent the horizontal and vertical gradient, w represents the weighting factor between 0 and 1. $I'_{c,\max}, I'_{c,\min}$ are respectively the maximum and minimum of $I'(\bar{p}_d)$, I'_+ and I'_- (see [13]).

Then we utilize cross-check to locate the error matching points in the initial disparity map and smooth the disparity map considering the result of image segmentation [14]. Specifically, suppose that the points lying in the same segment region possess the same disparity, we compute the mean disparity of correct matching points in the segment region and assign it to the region.

4. Experiments

3DSMAX simulation experiments and multiview autostereoscopic display experiments using stereoscopic image sets are conducted to evaluate our algorithms. Mean Shift segmentation is implemented in L^*u^*v color space, and we use 16, 16, 20 for h_s, h_r and the minimum region size, respectively.

4.1. 3DSMAX simulation experiments

We use 3DSMAX to establish a stereoscopic scene with four cubes. The largest cube is 30 cm long, 120 cm wide, and 106 cm high and the other three cubes are 10 cm long, 10 cm wide, and 10 cm high. The distance between the small cube in the middle and the largest cube is 80 cm. The other two small cubes which are 15 cm apart located 110 cm in front of the largest cube. Three cameras which are 5 cm apart respectively are placed 50 cm ahead of the two small cubes. The focal length and the horizontal view angle are 43 mm and 45°. The cameras are placed by the toed-in array method and the parallel array method to capture the stereoscopic images from the frontal view.

Fig. 2 shows the images captured by the toed-in camera array. The principal axes of cameras converge on the small cube in the middle. The largest cube has positive disparity and its positions from left boundaries in the three images are rightward one by one from left image to right image. The middle cube is zero disparity and located in the same positions in the three images. Two small symmetrical cubes have negative disparities and their positions from left boundaries are leftward one by one from left image to right image. However, vertical disparities and keystone distortions exist in the images by the edge of the largest cube.

Fig. 3 shows the images captured by the parallel camera array. The whole object's positions from left boundaries in the three

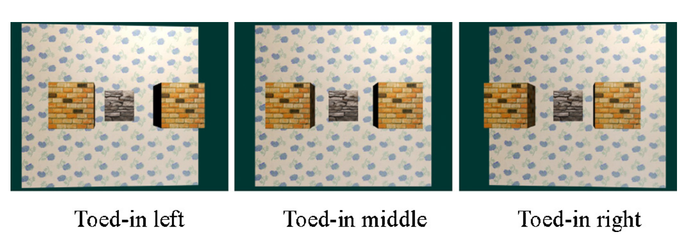

Fig. 2. Images captured by toed-in camera array.