Published in final edited form as:

Proc IEEE Annu Northeast Bioeng Conf. 2014 April 25; 2014: 1–3. doi:10.1109/NEBEC.2014.6972996.

A Novel Binocular Vision System for Wearable Devices

Haitian Zhai^{1,3}, Hui Li¹, Yicheng Bai³, Wenyan Jia², and Mingui Sun^{2,3,4}

Mingui Sun: drsun@pitt.edu

¹School of Electronics and Information, Northwestern Polytechnical University, Shaanxi, China

²Departments of Neurosurgery, University of Pittsburgh, Pittsburgh, PA, USA

³Departments of Electrical and Computer Engineering, University of Pittsburgh, Pittsburgh, PA, USA

⁴Departments of Bioengineering University of Pittsburgh, Pittsburgh, PA, USA

Abstract

We present a novel binocular imaging system for wearable devices incorporating the biology knowledge of the human eyes. Unlike the camera system in smartphones, two fish-eye lenses with a larger angle of view are used, the visual field of the new system is larger, and the central resolution of output images is higher. This design leads to more effective image acquisition, facilitating computer vision tasks such as target recognition, navigation and object tracking.

Keywords

binocular vision; wearable device; fish-eye lens

I. Introduction

The human visual perception system provides us with a unique ability to recognize objects, track targets of interests, and navigate in the real-world environment. These tasks are easy for people but not so easy for computers, especially for miniature wearable computers. A number of wearable vision systems have been studied in recent years [1]. However, none of them produces comparable images to that produced by the human visual system. We believe that the reason is that the visual field of the human eyes is much larger than that of the wearable cameras [2]. The field of each human eye is approximately 95 degree away from the nose, 60 degree toward the nose and, for binocular vision, the angle of view can be as large as 180 degree [3]. In order to overcome the shortcoming of the limited field, we have investigated a different system called a virtual motion visual system. Previously, the motion vision system was usually designed to control the viewing direction of the camera using motors. However, the movement of the camera is restricted by its mechanical functions and the system is difficult to be integrated into a miniature wearable device with a strict space limitation and a low power budget.

The main strategy utilized in our virtual motion visual system is to take advantage of two binocular fish-eye lenses to obtain images with wide views and then extract partial (retinal) images around the fixation points from the images taken by these lenses. Such a virtual

system can be easily implemented by software without any special hardware. It can also be utilized to detect moving objects within the wide angle of view by processing the images.

II. Methods

The structure of the proposed system is shown in Fig. 1. Image stitching process is implemented after correcting the distortion introduced by the fish-eye lenses. Then, the virtual motion vision is obtained from the stitching result.

A. Distortion correction

In order to obtain a large visual field, a fish-eye lens is used for each camera. We applied a perspective projection method which keeps straight lines in the real-world straight in the processed image to calibrate and correct the distorted images produced due to the use of the fish-eye lens [4]. It is assumed that the distorted image only has the radial distortion which means that the distortion model is the same in every direction along an arbitrary ray from the viewing center. A calibration checkerboard is used to calibrate the camera as is shown in Fig. 2(a), and the distortion curve after calibration is shown in Fig. 2(b). According to the distortion curve, real-world straight edges still appear straight in the corrected image.

B. Image stitching

Image stitching is necessary in order to obtain a single output image with a large visual field. The SIFT features between the two input images are extracted to establish point correspondences. Since the two cameras are installed in our wearable system are not parallel but with a small angle between them, the features extracted from the two images have different scales. However, this problem can be handled by our method since SIFT features are invariant under rotation and scale changes.

Once features are extracted from the two input images, they are matched by random sample consensus which is an algorithm to estimate parameters of a mathematical model from a set of observed data [6]. Due to the parallax effect caused by unwanted motion of the optical center and the misregistration error caused by mis-modelling of the lens distortion, multiband blending algorithm is used in the output image. After stitching and post-processing, the field of view of the stitched image becomes much larger than each of the raw input images.

C. Virtual motion vision

Human continuously directs their eyes toward interesting points of the scene to recognize an object. Such visual attention plays an important role in human vision. To imitate this property of the human eye, we set a fixed-size window in the stitched image as the retinal area, and the rest as the peripheral area. The retinal area virtually moves through the image with a certain freedom while the retinal area and the peripheral area interchange dynamically. The foveal vision is efficient in gist extraction because less data are needed to be loaded into the memory of the image processor, and the scanning process corresponds to adding or deleting a small amount of information from the visual space. Thus, this virtual

motion property provides us with an efficient way to scan the image space with less computational resource and power consumption.

III. Results

We have conducted experiments to test our system in both indoor and outdoor environments. Shown in Fig. 3 is one of the outdoor scenes. Fig. 3(a) is the image taking by a wide angle camera. It has 111° angle of view in the horizontal direction and 80° in the vertical direction. The radial perspective projection undistortion method was utilized. The output image after distortion correction is shown in Fig. 3(b). It can be seen that the distorted straight lines in the raw input image appears straight in the output image by using the perspective projection.

Since each of the lens has 111° visual field horizontally, the angle of view of the stitched image (Fig. 4a) from two input images is 151° which is similar to that of the human eye. The vertical angle of view is approximately 88° . Compared with the commonly used cameras in smartphones, which typically has angles of views of only 54.5° horizontally, 37.8° vertically and 63.5° diagonally, the new design has a better performance benefited from the large visual field. We set the retinal image the same size as that would be obtained by the smartphone camera as foveal images (red boxes in Fig. 4 (a)), so that our virtual motion system has a horizontal freedom of 96.5° . Shown in Fig. 4(b) are four virtual foveal images obtained from our virtual motion system.

IV. Conclusion

A novel binocular vision system for wearable device which incorporates the model of the human eye is presented. The system significantly enlarges the visual field without significant distortion effect in the output image. Our virtual motion visual system facilitates image processing tasks such as target recognition, navigation and tracking.

Acknowledgement

This work was supported in part by the National Institutes of Health grants R01CA165255 and R21CA172864 of the U.S. and the National Natural Science Foundation grant No. 61171155 and Natural Science Foundation of Shaan Xi Province Grant No. 2012JM8010 in China.

References

1. Mayol W, Cuevas, Tordoff BJ, Murray DW. On the choice and placement of wearable vision sensors. *IEEE Trans. on Systems Man and Cybernetics*. 2009; 39(2):414–425.
2. Ude A, Kyoto, Gaskett C, Cheng G. Foveated vision systems with two cameras per eye. *Proceedings of the IEEE International Conference on Robotics and Automation*. 2006
3. Henson, DB. *Visual Fields*. Oxford University Press; 1993.
4. Forstner, Wolfgang; Abraham, Steffen. Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2005; 59(5):278–288.
5. Larson, Adam M.; Loschky, Lester C. The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*. 2009; 9(10):6, 1–16. [PubMed: 19810787]
6. Brown, Matthew; Lowe, David G. Automatic Panoramic Image Stitching using Invariant Features. *International Journal of Computer Vision*. 2007; 74(1):59–73.

Figure 1.
Structure of the binocular visual system

(a)

(b)

Figure 2.

(a) Calibration board used for calibrating the camera, (b) Radial correspondences between distorted image and corrected image.

(a) *(b)*

Figure 3.
Original images (*a*). Images after distortion correction (*b*)

(a) *(b)*

Figure 4.
Our large view image (*a*) and 35mm camera image (*b*)