Real-time Background Subtraction for Video Avatar

Hasup Lee, Yoshisuke Tateyama, and Tetsuro Ogi

***Abstract*—Background subtraction is widely used method of detecting moving objects in static background in computer vision fields. In this paper, we present real-time background subtraction method for video avatar system or video conferencing without blue screen. We observed the noise of digital camera and made a range matrix for background masking. Noise reductions were applied to both background and mask using median filter blurring and open operation. Our experiment was performed using notebook PC and built-in webcam. The results showed our approach is feasible to real-time indoor environments.**

***Index Terms*—Computer Vision, Background Subtraction, Video Avatar, Indoor.**

I. Introduction

Background subtraction is widely used method of detecting moving objects in static background in computer vision fields [1]. The objects are human, cars, ships and etc. and the environments are varied from indoors to outer space.

And it is also used for entertainment and broadcasting - produce the scenes of actors on computer graphic backgrounds or a caster on weather map. Chroma key or blue screen method is commonly used for this purpose.

Video avatar technology made users can share a virtual space between remote sites [7]. Users can point out the 3D data more precisely using video avatar with stereoscopic devices. Video avatar was applied to digital museum [8]. Chroma key was used for background subtraction in that system. But Chroma key needs space and blue or green curtain for making video stream.

Video avatar system does not need skeleton modeling of human. The background is indoor environment then no complex modeling of background is needed. The system sends and receives a video stream so real-time performance is required.

In this paper, we present real-time background subtraction method for video avatar system or video conferencing without blue screen. We observed the noise of digital camera and made a range matrix for background masking. Noise reductions were applied to both background and mask. Our approach is feasible to real-time indoor environments.

Manuscript received June 15, 2012; revised July 10, 2012.

This research is supported by Keio University Global COE (Center of education and research of symbiotic, safe and secure system design) Program.

The authors are with Visual Simulation Laboratory, Graduate School of System Design and Management, Keio University, Yokohama, Japan (e-mail: hasups@sdm.keio.ac.jp, tateyama@sdm.keio.ac.jp, ogi@sdm.keio.ac.jp).

II. Previous Works

There are many background subtraction approaches. These are well reviewed in [1]. Some methods made models for background and others made models for objects. There are various approaches from simple to complex one. A human body model using blob detection is made in pFinder [2]. A statistical background models for separated color and brightness are used in [3][4]. We first observed digital camera noise and made range matrix for determining background.

III. Digital Camera Noise

There are various noises caused by many reasons in digital camera images [5]. Inherent noise is well analyzed in [6] but there are other noise caused by external factors. For example, the fluorescent ramps are usually used in our experiment environment. These kinds of light have oscillation that humans cannot notice.

Fig. 1 shows plotting certain point of static background over 100 frames in RGB color space in each. The number of same RGB value is more, the point is darker. Fig. 1 shows there is no simple and fast model that represents background.

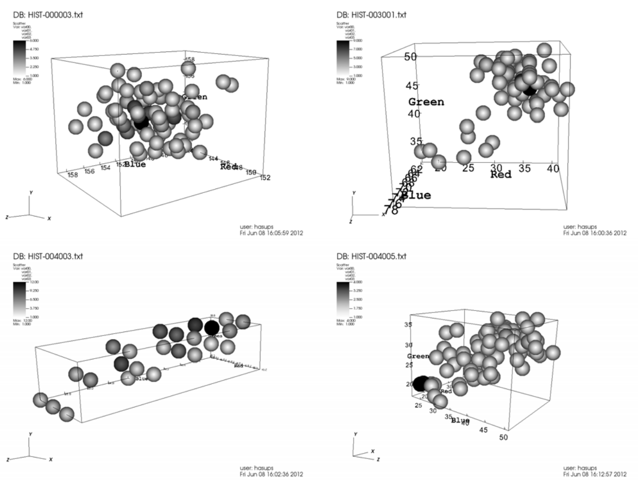

Fig. 1. Examples of plotting a point in RGB space over 100 frames.

Because the object of our approach was not analysis of camera noise, we developed simple range matrix as our background masking model.

IV. Preprocessing

Before background subtraction, the range matrix for masking was generated using static background frames. First predefined numbers of frame were stored in memory. Then noise was reduced in each frame and the range matrix was made. This range matrix was used for masking background from current frame during real-time video streaming.

DOI: 10.7763/IJFCC.2012.V2.117

A. Storing Background

First, we captured predefined *n* number of camera images in array. Images were static backgrounds without objects and stored to be sorted later.

B. Background Noise Reduction

To reduce the noise of background, median filter blurring was applied. Same sized blurring would be also applied to current camera image later.

C. Determining Background Range

Stored images were sorted in each RGB value at each point and let it be sorted array *S*. For example, sort {red value of (10, 10) in 1st image, red value of (10, 10) in 2nd image, …, red value of (10, 10) in *n*th image}. Then up value and down value of background range was stored in range matrix. The number of stored images is *n*, the number of error data is *e*, and tolerance is *t* then for each point, the up and down matrix *Up* and *Down* of background range are as (1). Their dimensions are the same as camera image.

$$Up_{x,y} \leftarrow \text{MIN}(\ (n\text{-}e)\text{th biggest value in } S_{x,y} + t,\ 255\ )$$
$$Down_{x,y} \leftarrow \text{MAX}(e\text{th smallest value in } S_{x,y} - t,\ 0\ ) \quad (1)$$

The values of *Up* and *Down* matrix mean the maximum and minimum values of bounding box in Fig. 1 except considering error *e* and tolerance *t*.

V. Background Subtraction

After making range matrix, background was subtracted from current camera image in real-time. A mask matrix was generated using the range matrices and applied by opening operation for noise reduction. Before mask matrix was generated, the current camera image is blurred using median filter of previous size.

A. Generating Mask Matrix

For each point in current camera image, if the red, green, and blue values are all between the values of *Up* and *Down* matrix, it was regarded as background. The mask matrix was generated using this determination.

B. Mask Noise Reduction

We applied opening operation to the mask matrix. The opening operation was used for noise reduction and consists of two steps; the erode operation and the dilate operation. The opening operation does not change the size of the area of the object but the blur operation does.

VI. Implementation

We used built-in webcam of notebook PC as input camera. The Motion Eye™ is built-in webcam of Sony Vaio™ notebook PC and its image's resolution is 640 X 480. The CPU is 2.8GHz dual core processor and memory is 4GB.

OpenCV 2.4 was used as computer vision software library. And develop environment was Microsoft Visual Studio™ 2010. The experiment was performed under indoor laboratory or office conditions.

VII. Results

The experimental environment was indoor laboratory or office like Fig. 2. The lights were fluorescent ramps. We captured and stored 100 frame images as background (n = 100). Each image was blurred using median filter of size 3 pixel. Fig. 2 is blurred image.

Fig. 2. Background image.

When determining range matrix, the number of error data was 20 (e = 20) and tolerance was 10 (t = 10). Using this range matrix, the mask matrix was generated as Fig. 3. The 7X7 filter for open morphology operation was used for noise reduction. Fig. 3 is the image after the open operation.

Fig. 3. Mask image.

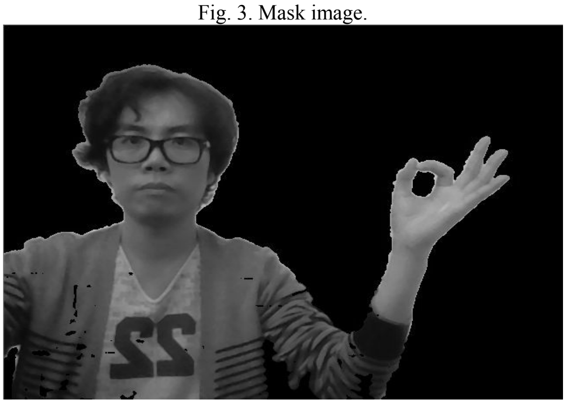

Fig. 4. Background subtraction by our approach.

The final result is shown as Fig. 4. There were books or stationery of various colors in background, so bottom part of object was not clearly subtracted.

But the result of Chroma key method as shown in Fig. 5

contains more noise in entire region. The value of thresholding is 10.

The parameters – error data, tolerance, Chroma key thresholding, open operation filter size, median filter size – can be controlled using track bars.

Fig. 5. Background subtraction by chroma key method.

VIII. Conclusion

The environment of video avatar or video conferencing is static and the lights are usually fluorescent ramps or other sources which have the oscillation that humans cannot notice. Various external factors made the distribution of camera image noise complex.

In this paper, background subtraction was performed using the range matrix. Noise was reduced using median filter blurring and open operation. Our approach ran in real-time and show feasible results.

References

[1] M. Piccardi, “Background subtraction techniques: A review,” *IEEE Conf. Syst, Man Cybern*, vol. 4, pp. 3099–3104, 2004.

[2] C. Wren, A. Azarhayejani, T. Darrell, and A. P. Pentland, “Pfinder: real-time tracking of the human body,” *IEEE Trans. on Pattern Anal. and Machine Intel*, vol. 19, no. 7, pp. 780-785, 1997.

[3] T. Horprasert, D. Harwood, and L. S. Davis, “A statistical approach for real-time robust background subtraction and shadow detection,” In *IEEE ICCV*, pp. 1-19, 1999.

[4] L. Jia and Y. Liu Y, "A novel thresholding approach to background subtraction," *IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 1-6, Jan. 2008.

[5] Cambridge in Colour, (June 2012), "Digital Camera Image Noise, Part 2," [Online]. Available: http://www.cambridgeincolour.com/tutorials/image-noise-2.htm

[6] R. Boie and I. Cox, “An analysis of camera noise,” *IEEE Trans. Pattern Anal. Mach. Intell*, vol. 14, no. 6, pp. 671–671, June 1992.

[7] K. Tamagawa, T. Yamada, T. Ogi, and M. Hirose, "Developing a 2.5-D Video Avatar," *IEEE Signal Processing Magazine*, vol. 18, no. 3, pp. 35-42, May 2001.

[8] H. Lee, Y. Tateyama, T. Ogi, T. Nishioka, T. Kayahara, and K. Shinoda, “Interactive Exhibition with Ambience using Video Avatar and Animation on Huge Screen,” *The 14th International Conference on Human-Computer Interaction (HCII 2011)*, LNCS 6774, pp. 253-259, July 2011.