MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11-13, 2002, Nara-ken New Public Hall, Nara, Japan

3—28 Detecting moving objects with an omnidirectional camera and subtraction whose background image is renewed.

Kazumasa Yamazawa[1]
Graduate School of Information Science
Nara Institute of Science and Technology

Naokazu Yokoya[2]
Graduate School of Information Science
Nara Institute of Science and Technology

Abstract

When we use a PTZ camera and background subtraction to detect moving objects, we must synchronize the PTZ of the camera and the PTZ of the background image. This paper describes a method a new method of detecting moving objects with an omnidirectional camera. The proposed method uses a fixed omnidirectional camera which can acquire images surrounding the camera, and background subtraction based on updating background image. When we subtract the background image from the input image, we consider oscillation of intensity. Therefore, the proposed method can robustly detect moving objects under the condition that there were fluorescent lights and windows which outside light came into. In the experiment, the system can detect moving objects and show the images of the objects in the room.

1 Introduction

When we detect moving objects with a common video camera, the field of view is limited. There are some methods as solve this problem. They are methods which use a PTZ (panning, tilting and zooming) camera[1], methods which use multiple cameras[2], methods which use an omnidirectional camera[3], and so on. However, if we use the background subtraction as a major method for detecting moving objects, and use a PTZ camera, we must synchronize the PTZ of the camera and the PTZ of the background image. Moreover, when we use a common camera, the camera has a limited field of view and detects limited moving objects. When we use multiple cameras, we requires multiple cameras and computers, and so the system becomes complex. On the other hand, the omnidirectional camera, which can look omnidirectional view at a time, is suitable for detecting moving objects in the surrounding environment. This paper describes a new method of detecting moving objects with an omnidirectional camera and subtraction whose background image is updated.

This paper is structured as follows. First, we describe the method of detecting moving objects with a fixed omnidirectional video camera. Next, we present experiments to detect and show moving objects. Finally, we summarize the present work.

2 Detecting moving objects

2.1 Outline of the method

The proposed method uses an omnidirectional camera and background subtraction. The fixed omnidirectional camera acquires the omnidirectional images surrounding the camera. At first, we estimate the background image and amplitude image from the first several frames of the input omnidirectional images. After the background image and amplitude image are estimated, we detect moving objects from the input images, the background image and the amplitude image, updating the background image and the amplitude image. Next, we calculate the fan-shaped bounding-boxes which enclose the detected moving objects, generate the common perspective images of the moving objects, and show the user the common perspective images.

This section is structured as follows. First, we describe the omnidirectional camera: HyperOmni Vision. Next, we describe the method of detecting moving objects from the omnidirectional input images. At last, we describe the method of calculating the fan-shaped bounding-boxes which enclose the detected moving objects.

2.2 Omnidirectional camera: HyperOmni Vision

We use HyperOmni Vision[4] as an omnidirectional camera in the present work. HyperOmni Vision is composed of a hyperboloidal mirror and a common video camera as illustrated in Fig.1. The camera acquires an omnidirectional scene reflected by the hyperboloidal mirror. The hyperboloidal mirror is constructed of a hyperboloid of two sheets of revolution, which has two focal points (O_M and O_C). The camera lens center is fixed at the focal point O_C. Given a world coordinate (X,Y,Z) and an image coordinate (x,y) as shown in the Fig.1(b), the shape of hyperboloidal mirror and the two focal points are represented as follows.

$$\begin{aligned}&\text{Mirror Shape :} && \frac{X^2+Y^2}{a^2}-\frac{Z^2}{b^2}=-1,\\ &\text{Inner focal point } O_M: && (0,0,+c), \quad (1)\\ &\text{Outer focal point } O_C \text{ (Camera lens center)}: && (0,0,-c),\\ &&& \text{where } c=\sqrt{a^2+b^2}.\end{aligned}$$

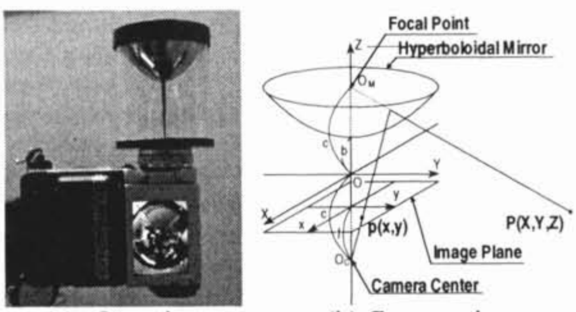
A ray going from the point $P(X,Y,Z)$ in 3D toward

[1] Address: 8916-5 Takayama-cho, Ikoma, Nara 630-0101 Japan. E-mail: yamazawa@is.aist-nara.ac.jp
[2] Address: 8916-5 Takayama-cho, Ikoma, Nara 630-0101 Japan. E-mail: yokoya@is.aist-nara.ac.jp

the inner focal point O_M is reflected by the mirror and passes through the outer focal point O_C intersecting the image plane at the point $p(x, y)$. Therefore, the projection of HyperOmni Vision is not common planar perspective, but satisfies the single viewpoint constraint. The relationship between $P(X, Y, Z)$ and $p(x, y)$ is given by:

$$x = \frac{f(b^2 - c^2)X}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}},$$
$$y = \frac{f(b^2 - c^2)Y}{(b^2 + c^2)(Z - c) - 2bc\sqrt{X^2 + Y^2 + (Z - c)^2}}. \quad (2)$$

By using the above equation, an omnidirectional input image can be converted to a common perspective image at the viewpoint O_M [5] (see Fig.2). In this study, we show the user the common perspective images of the moving objects.



(a) Overview (b) Construction

Fig.1 HyperOmni Vision.

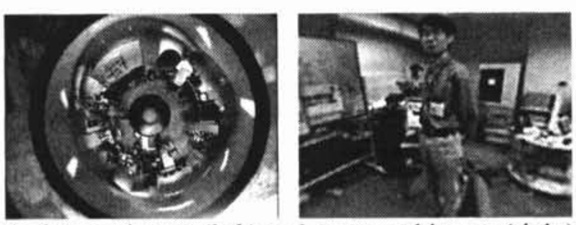


Fig.2 Input image (left) and generated image (right).

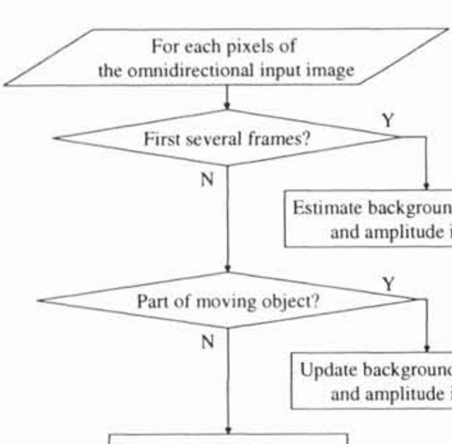


Fig.3 Flow chart of detecting moving objects area.

2.3 Detecting moving objects area from omnidirectional input images

The proposed method which detects moving objects is based on a background subtraction method. In the background subtraction process, we subtract at each pixel, but we don't consider the neighborhood. Figure 3 shows the flow chart of detecting moving objects.

We made the model of intensity of background pixel as follows.

I = Iave + σsin(2π*ω*t) + k*noise (3)
I : intensity of background pixel
Iave : average of I
σ: amplitude of oscillation of intensity
ω : frequency of oscillation of intensity
t : time
k : -1, 0, or +1
noise : noise which is dependent only on the video camera

The term σsin(2π*ω*t) means a flicker of a fluorescent light, CRT and etc. And, the term k*noise means a noise which is dependent only on the video camera due to gain-up and etc. Therefore, the intensity has the range of following.

Iave-σ-noise ≦ I ≦ Iave+σ+noise (4)

If the intensity I of the pixel is outside the range, we estimate that the pixel is a part of a moving object.

First, we predetermine parameter noise which is dependent only on the video camera setup.

Next, we consider the slow change of the background. When the pixel is estimated at the background, the parameter Iave and σ are updated by following equation.

I'ave = (n-1)/n*Iave + 1/n*I
σ' = (n-1)/n*σ + 1/n*2*(I-Iave)2 (5)
n : parameter of the update speed
(When n increases, updating speed decreases.)

We calculate the equation (5) about the pixel which is estimated at the background on each frame. We use I'ave and σ' as Iave and σ of next frame.

When the pixel is estimated at a part of moving object, we don't calculate the equation (5). But, we consider cases that an object is put or taken, and cases of the rapid change of the background. And so, when the pixel is estimated at a part of moving objects, the parameter σ is updated by following equation.

σ' = (m-1)/m*σ + 1/m*2*(I-Iave)2 (6)
m: parameter of the speed that the detected object fade out
(When m increases, the speed of fading out decreases.)

Normally, we determine m larger than n. Therefore, if the detected moving object stays at a position, σ increases, and then the detected moving object fade out.

2.4 Generating the common perspective image of the moving objects

The above method can detect pixels of moving objects with updating the background image continuously. Next, the common perspective images of the moving objects are shown to the user. It is similar to the method of the reference [3]. The method is as follows.

1. Compute a histogram with respect to the longitude direction θ of the detected moving objects (see Fig.4).
2. Detect blobs whose values are more than 0 in the histogram as the ranges of the longitude of the moving objects $((\theta_{\alpha 1}, \theta_{\alpha 2})$ & $(\theta_{\alpha 3}, \theta_{\alpha 4}))$.
3. Compute the maximum and minimum latitude of the detected moving objects for the each ranges of the longitude.
4. Generate the common perspective images based on the computed ranges of the longitude and the latitude, and show to the user according to the presentation policy.

When we generate the common perspective images, we use the method of the reference [5].

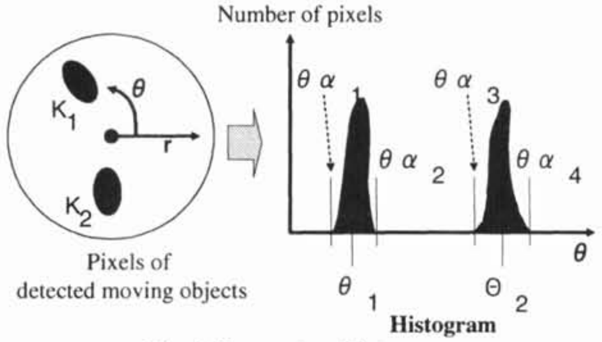


Fig.4 Compute a histogram

3 Experiments

We implemented the proposed method on a Linux PC. The spec of the PC is as follows.

OS: Redhat 7.1
CPU: Pentium4 2GHz
Memory: 512MB

Figure 5 shows the screen of the system which we applied the above method to. The screen is divided into 4 images. The 3 images show the 3 common perspective images of the largest moving objects. The other image shows an omnidirectional image, histogram, and etc. The magenta line means the histogram of the pixels of the moving objects in the longitude direction.

We set the parameters as follows.

$0 \leq I \leq 255$
noise = 4
n=360, m=1080

We set the parameters heuristically. But, the parameters can be applied to common indoor environments: lobby, council rooms, and so on.

In this experiment, there were 2 walkers and a putted doll under the condition that there were fluorescents and windows which outside light came into (see Fig.6). Figure 7 shows the stream of screen images. The system can detect moving objects robustly. The system updates images at intervals of 0.2 sec.

4 Conclusions

We have proposed a new method of detecting moving objects with an omnidirectional video camera. The proposed method uses a fixed omnidirectional camera which can acquire images surrounding the camera, and background subtraction based on updating background image. When we subtract the background image from the input image, we consider oscillation of intensity. Therefore, the proposed method can robustly detect moving objects under the condition that there were fluorescent lights and windows which outside light came into. In the experiment, the system can detect moving objects and show the images of the objects in the room.

The future work includes the improvement of updating the area of the detected moving objects and the tracking of detected moving objects between each frame.

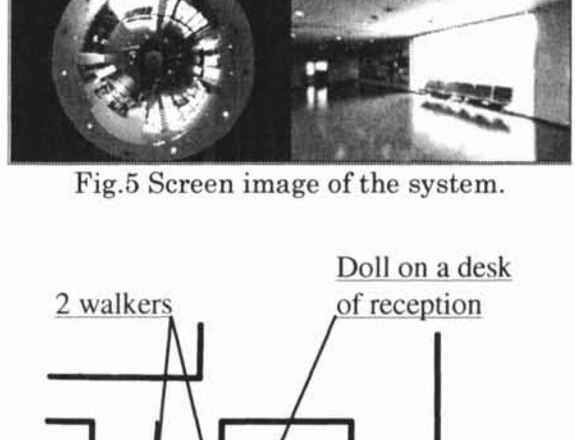


Fig.5 Screen image of the system.

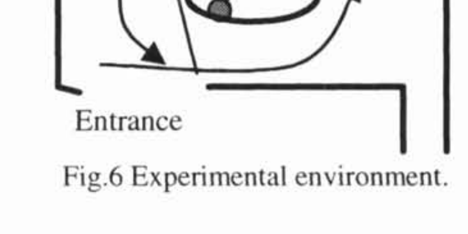


Fig.6 Experimental environment.

References

[1] T.Wada and T.Matsuyama: "Appearance sphere: Background model for pan-tilt-zoom camera", Proc. ICPR, vol.A, pp.718-722, 1996.

[2] Mikic, S.Santini, and R.Jain: "Tracking Objects in 3D using Multiple Camera Views", Proc. ACCV2000, pp.234-239, 2000.

[3] R.Miki, N.Yokoya, K.Yamazawa, and H.Takemura: "A Real-time Surveillance and Monitoring System Using Multiple Omnidirectional Video Cameras"

[4] K. Yamazawa, Y. Yagi, and M. Yachida. Obstacle Detection with Omnidirectional Image Sensor HyperOmni Vision. In Proc. Int. Conf. on Robotics and Automation, pages 1062-1067, 1995.

[5] Y. Onoe, K. Yamazawa, H. Takemura, and N. Yokoya. Telepresence by Real-Time View-Dependent Image Generation from Omnidirectional Video Streams. Computer Vision and Image Understanding, Vol. 71, No.2, pages 154-165, 1998.

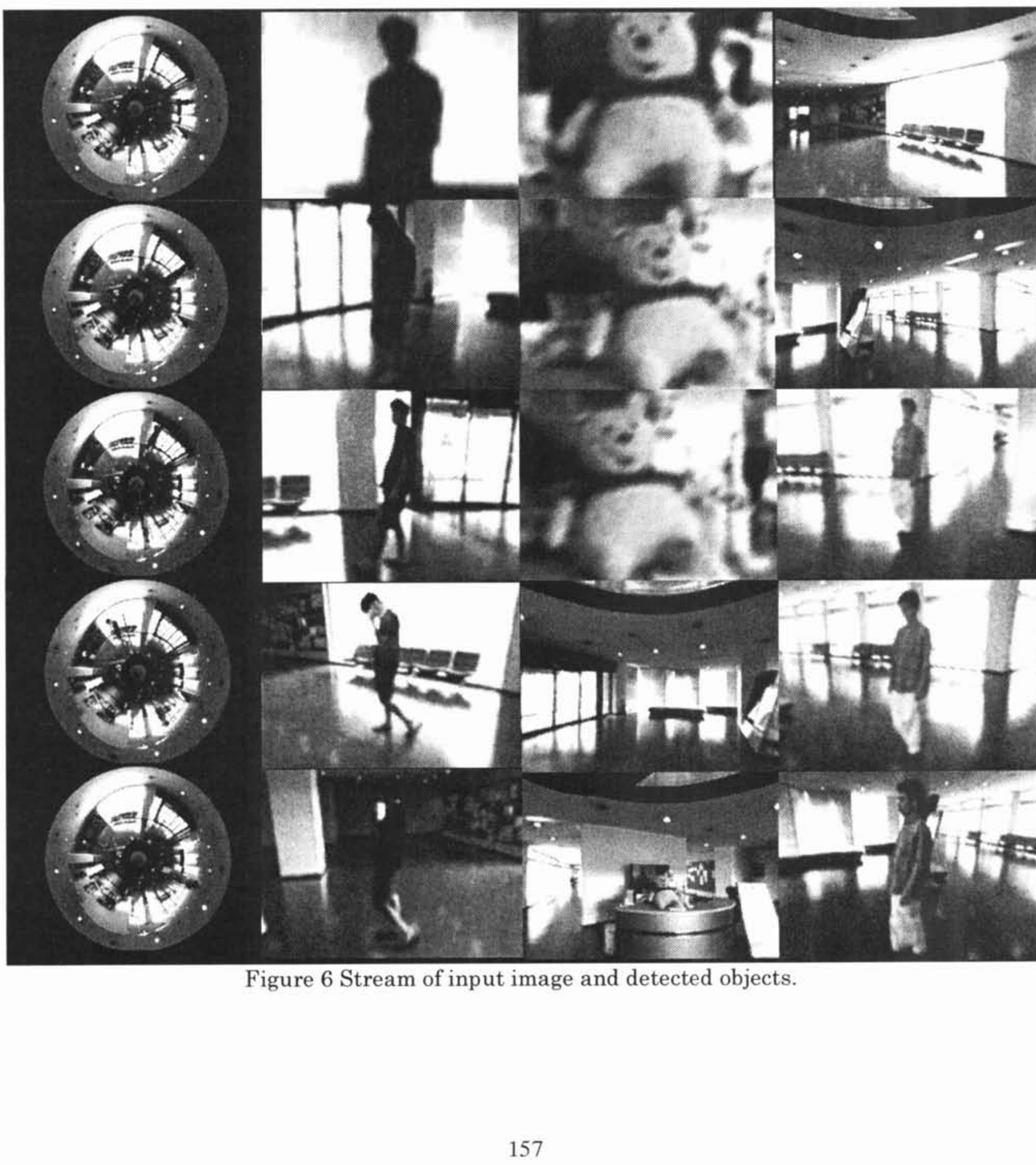


Figure 6 Stream of input image and detected objects.