DID using stereovision

Dong-Jin Seo , Sang-il Na†*, Ye-Eun Lim*††* and Dong-Seok Jeong*†††

Department of Electronic Engineering, College of Engineering, Inha University, Incheon, Korea

Summary

DID (Digital information display) has recently shown a sharp increase in number according to its various uses. Most of them have been used for transmitting information, but, in recent years, products equipped with input units have been increasing. This paper designs a case to apply Stereovision to DID and deals with application methods thereto. Moreover, it adjusts an angle of a camera to be used and simplifies a part of pre-treatment for high-speed treatment. Through these processes, it separates a user from a background. Accordingly, it is possible to perform various motion detections and provide user-made advertisement services.

Key words:
DID, Stereovision, camera, sensor, display.

1. Introduction

Recently, DID is an advertising media that has been recently noticed. This paper suggests a stereo vision[1] to be used for motion detection. It deals with a method for separating a background and a person at high speed by taking notice of the point that traditional stereo vision methods have a great number of operations and limits in distance and space. For this, two cameras to be separated from each other as far as possible are arranged on both sides of a billboard, and set to perform high-speed processing by removing functions with a great number of operations. Accordingly, it is possible to provide a user-made service using this function.

2. DID (Digital Information Display)

DID stands for Digital information display and means a stand-type billboard including a large-sized LCD panel. In recent years, according as the price of the panel has gone down, and communication units such as Internet have been developed, the number of products introduced has been rapidly increased. Such products have been used for transmitting simple information or advertisement unilaterally. However, various input units have gradually been required according to requests of markets.

Because old input units such as a keyboard and a mouse needed fixed appliances, they had limits of time and space. According to that, a type inputted though contact of body (hands) such as a touch sensor has been preferred.

Fig. 1 Example of DID

A touch screen as this input unit has been mostly used. The touch screen is classified into types of an infrared sensor, an infrared camera, ultrasonic wave and capacitance, and the infrared sensor type touch screen among them has been most widely used. The infrared sensor type indicates a method for irradiating infrared rays from one direction and receiving relevant ray to the other side. The infrared camera type indicates mounting a camera detecting infrared rays on one edge, attaching an infrared reflecting filter on other edges, then irradiating infrared rays from edged on the opposite side. At this time, it is judged whether touch is reacted by sensing the reflected ray. The ultrasonic wave type indicates irradiating ultrasonic waves on a touch screen, reading reflected waves and sensing whether or not to be contacted. Moreover, the capacitance type becomes recently famous for being used for iPhone and iPad and indicates detecting capacitance of body.

Fig. 2 Touch sensor

Manuscript received April 5, 2010
Manuscript revised April 20, 2010

However, these touch sensors have conditions that a user has to come very close to DID. According to that, many studies on methods for acknowledging spaces have been recently made. These are studied on types of ultrasonic wave, Gyro-sensor, video detection, etc. The ultrasonic wave type has a problem of larger limits of space and a defect of difficulty in detecting complicated motions. The Gyro-sensor type is convenient for use, but has a problem that motion can be detected by a user who has to take a sensor. According to that, the video detection type is more preferable, but has defects of complicated configuration and installment of systems.

3. Stereovision

3.1 Stereovision

Stereovision indicates a technology for receiving videos from two cameras and acquiring space information. On the basis of human's method for acquiring visual information, it is applied to various fields such as 3-dimensional reconstruction and recognition by abstracting distance information from the videos and using the information. Even though Stereo-vision can acquire and express such information, the video data inputted become doubled. Moreover, as the distance information is calculated from two videos, it has a defect, that is, the number of operation becomes larger. Nevertheless, it has great advantages that it is possible to acquire distance information in comparison with wide areas and acknowledge various motions.

3.2 Basic algorithm of Stereovision

Stereo vision is largely classified into three stages. These indicate stages for camera calibration, disparity map calculation and distance calculation. The first stage, the camera calibration, is a process for compensating for internal parameters of two cameras that are different.[2] Without this process, it is difficult to measure distance accurately. The second stage is disparity map calculation. One object is shown in different positions of two videos. When the two coordinates are acknowledged and how many pixels they are separated from each other is calculated, then they are expressed numerically, it indicates relative distance information on each object. Last, as shown in Fig. 3, it is possible to calculate the actual distance easily by using information on the focal distance and distance between cameras.

If z value is estimated with reference to Fig. 3, it is possible to abstract distance information on a relevant object. If positions of an object are set as x, y and z, and the lens of the left camera is set as the starting point, the left camera can be indicated in a proportional expression as equation (1).

$$\frac{a}{f} = \frac{x}{z} \tag{1}$$

Here, 'a' indicates the number of pixels that are positioned from the center on a screen and 'f' indicates the focal distance. If the right camera is expressed in the same method, it is indicated as equation (2).

$$\frac{b}{f} = \frac{d-x}{z} \tag{2}$$

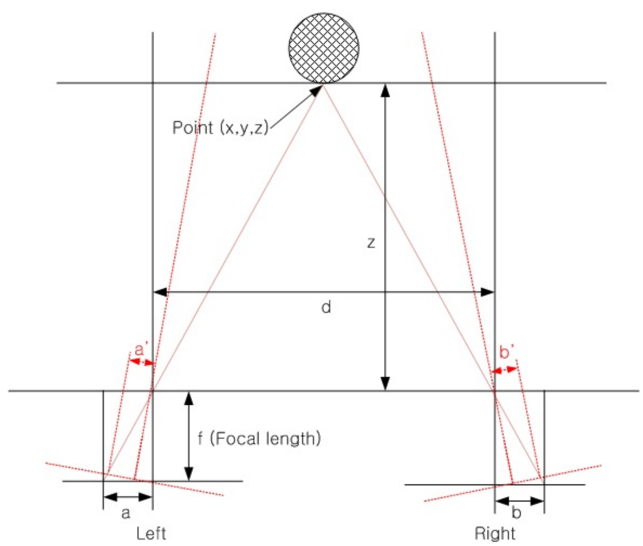


Fig. 3 Distance measurement method using Stereovision

If the focal distance of two cameras is the same, a desired value of z can be calculated.

$$z = \frac{fd}{(a+b)} \tag{3}$$

In other words, if you know the focal distance and the distance between two cameras, the value z can be calculated, the distance from the object.

3.3 Algorithm for Disparity map calculation

Disparity map indicates the actual distance information. Of course, the distance must be accurately calculated using these values, but it is possible to classify whether one coordinate is ahead of the other coordinate or not with such information. Traditional and easier methods for abstracting Disparity map are SSD and SAD methods.[3][4] The demo program for Bumblebee(POINT GAEY INC.), Stereo vision camera, also uses is the advanced SSD method. The SSD stands for Sum of squared distance. This method uses the difference in brightness between pixels. So, similarity of a specific area is calculated from the two pictures using this. According to the SSD method, the similarity gets higher, so the value gets lower, because of using the difference in brightness between the pixels.

$$SSD(x, y, d) = \sum_w [R_{x,y} - L_{x+d,y}]^2 \quad (4)$$

SAD stands for Sum of absolute distance. Also, the SAD method uses the difference in brightness between pixels, like the SSD method, however, the point of using absolute value for the difference in brightness between the pixels is different from the SSD method.

$$SAD(x, y, d) = \sum_w |R_{x,y} - L_{x+d,y}| \quad (5)$$

In this paper, the basic stereo vision algorithm is performed using the SSD method.

4. System

4.1 System configuration map

Generally, stereo matching includes Camera calibration and disparity map calculation. This process is required for so many calculations. A general method for applying camera calibration is the use of transformation matrix, but this process is also required for so many calculations. However, without the process, it is difficult to measure distance accurately. The reason is that information on the cameras is different.

This system is focusing on classifying person in front of DID and abstracting their motions. According to that, the speed of calculation is increased by removing the process for camera calibration. Because, if the camera with similar conditions is used, it is suggested that camera calibration is completed. The process is removed not to need a value of distance that is accurately measured but to perform real-time processing.

DID case is manufactured that is formed long laterally. The cameras are attached on both ends thereof, at this time, d', the distance between the cameras is 1250mm. D values of general stereovision cameras are indicated between 100 and 200mm. The reason is that it was designed to be similar to human as much as possible. Moreover, if it is installed in a robot, there is space limit for installation, so it is manufactures like this. The system can use the whole of lateral surface of DID, so it has an advantage of great increase in the value d. If the value d is large, it is possible to detect objects in a farther distance, so it has great advantages in comparison with general stereovisions. However, if the value d is farther due to a camera's an angle of view, a problem rises that it is operated from a distance.

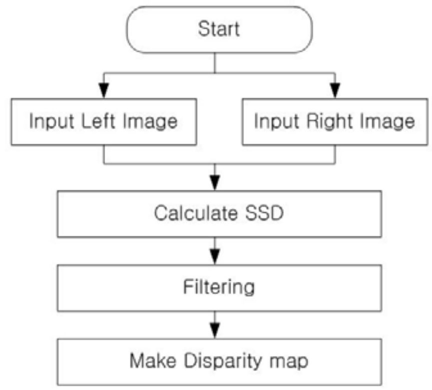


Fig. 4 System configuration

According as the value d is larger than the traditional system, the camera is installed to be turned inward at 5 degrees each. In this case, it makes the increase in the volume of calculation for the value z. Because it should calculate the central and actual value z using z values of right and left sides. However, the system is focusing not on accurate measurement of distance but on abstraction of objects and real-time separation of backgrounds, so it operates calculation with the same as the general formula.

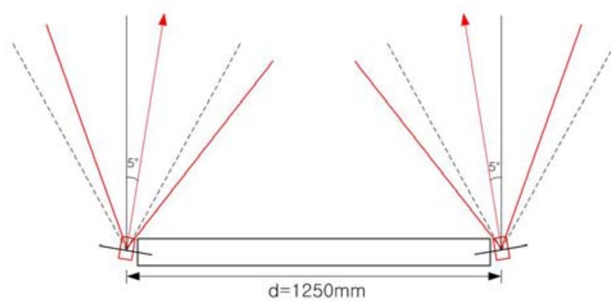
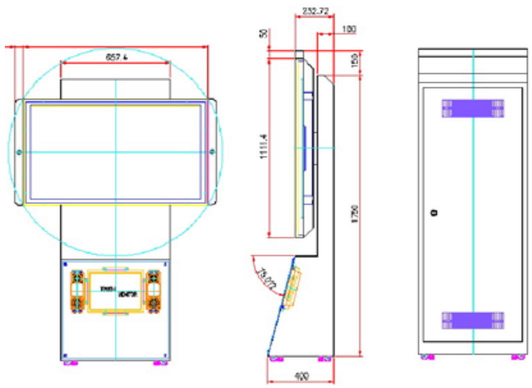


Fig. 5 Angle change of 5 degrees in case of camera installation

4.2 System explanation (explanation on pictures and contents)

In the present system 46-inch LCD panel is used, and DID case to be rotatable is manufactured for experiments. It is equipped with spaces for fixing and installing cameras on both ends of the case and it is possible to acquire information on a desired area through adjustment of the camera's angle.



(a) DID drawing

(b) Actual case manufactured

Fig. 6 Exterior of DID

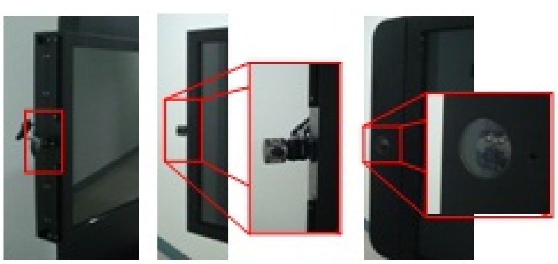
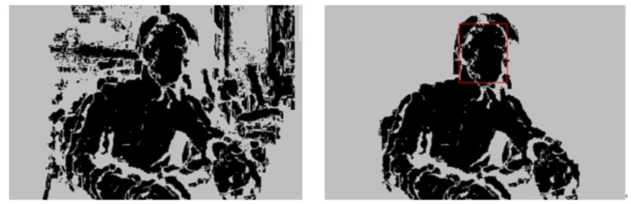


Fig. 7 Feature of camera installation

5.Result

5.1 Result



(a) (b)

Fig. 8 result images (a) Disparity map, standardized as the left camera, and (b) Separation of background and detection of face area

5.2 Result analysis

The present system is focusing on separating a user from background by application of stereovision to DID, and providing various services by such applications. According to application of SSD and calculation of disparity map, it is possible to get a result such as (a) in Fig. 8, and if the background and the user are separated from them, it is possible to get a result such as (b). The detection speed was measured within 13ms till the separation of the background and the user, with the standard of Intel E8600, DDR2 2G.

(b) in Fig. 8 is an example indicating application of face detection algorithm[5] to a relevant area after separation of front and rear videos, while setting the user as a standard among the distance information calculated.

With the use of the present system, it is possible to introduce user information to advertising on the basis of the distance from the user and then provide user-made services to them. Later, functions for tracking and forecasting motion information, based on the present system, will be added.

References

[1] David A.Forsyth and Jean Ponce, "Computer vision", *Prentice Hall*, pp.234-250, 2002

[2] Jean-Philippe Tarel and Jean-Marc Vezien, "A complete software Solution for Camera Calibration", *INRIA*, pp.5-10, 1996

[3] A.Luo and H.Burkhardt, "An intensity-based cooperative bidirectional stereo matching with simultaneous detection of discontinuities and occlusions", *International Journal of Computer Vision*, vol 15, 171-188, 1995

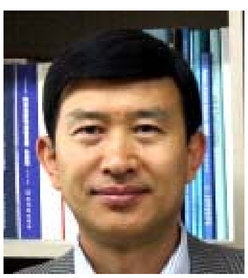
[4] Takeo Kanade and Masatoshi Okutomi, "A stereo matching algorithm with an adaptive window", *IEEE Transactions on pattern anaylsis and machine intelligence*, Vol 16, No 9, 1994

[5] Castillo, C.D. and Jacobs, D.W.,"Using Stereo Matching with General Epipolar Geometry for 2D Face Recognition across Pose", *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol 31, 2009

Dong-Jin Seo received the B.S. and M.S. degree in Electronic Engineering at the School of Electronic Engineering from INHA University, Incheon, Korea, in 2002 and 2004. He has been received the Ph.D. course in Electronic Engineering at the School of Electronic Engineering from INHA University, Incheon, Korea from 2006. His main research interests include stereo vision, machine vision and image/video coding.

Sang-il Na received the B.S. and M.S. degree in Electronic Engineering at the School of Electronic Engineering from INHA University, Incheon, Korea, in 2002 and 2004. He has been received the Ph.D. course in Electronic Engineering at the School of Electronic Engineering from INHA University, Incheon, Korea from 2006. His main research interests include video motion estimation and image/video coding.

Ye-Eun Lim received the B.S. degree in Electronic Engineering from INHA University, Incheon, Korea, in 2008. She has been received the M.S. course in Electronic Engineering at the School of Electronic Engineering from INHA University, Incheon, Korea from 2009. Her main research interests are in the areas of robot vision, object tracking system.



Dong-Seok Jeong received the B.S. degree from Seoul National University, Korea, in 1977. And he worked as research engineer in Agency for Defense Development. He received the M.S. degree and Ph.D degree from Virginia Tech, USA, in 1985 and 1988, respectively. Currently he is an associate professor of the Department of Electronic Engineering at Inha University. He is a senior member of IEEE. His research interests include machine vision, image/video coding, MPEG-7 and DRM.