2011 11th International Conference on Control, Automation and Systems
Oct. 26-29, 2011 in KINTEX, Gyeonggi-do, Korea

Human Position Estimation in Intelligent Space for an Active Information Display

Jae-Hyun Ahn[1], Jeong-Eom Lee[2], Joo-Hyung Kim[2], Sang-Jun Kim[1], Joo-Ho Lee[3] and Gwi-Tae Park[2]

[1] Department of Mechatronics, Korea University, Seoul, Korea
[2] School of Electrical Engineering, Korea University, Seoul, Korea
(Tel : +82-2-3290-3673; E-mail: gtpark@korea.ac.kr)
[3] College of Information Science and Engineering, Ritsumeikan University, Shiga, Japan

Abstract: In this paper, we introduce a new active information display system which is based on Intelligent Space and a projector mounted mobile robot. This system is able to afford a human with relevant information by projecting it on where the human is looking such as a wall, a floor, and a door. For this system, Intelligent Space should recognize location of human. To estimate position of human in Intelligent Space using multiple camera system, we propose a human detection method based on the Histogram of Oriented Gradients(HOG) feature descriptors and the boosting technique in each camera and a position estimation method using multiple view geometry. The performance of proposed method for human position is verified by real-time experiments.

Keywords: Human Position Estimation, Intelligent Space, Ubiquitous Display, Active Information Display System, Human Computer Interaction.

1. INTRODUCTION

The main goal of this research is to realize an active information display system which is able to afford a human centered information transfer method. In our living environment, the most famous and representative ways of transferring information are passive information display methods such as road signs, guide maps, notice boards, etc. In these kinds of information display methods, constant information is written on fixed media in advance and those media are located at some specified place. It is the most common way of transferring information from past to now since the birth of human being. However, there are lots of problems in such methods and they cannot be human centered. Thus, to cope with these problems, we proposed a new active information display method based on a projector mounted mobile robot[1,2,3]. The proposed system is able to afford a user with relevant information by projecting it on where the user is looking, so that the user does not need to move for seeking information.

To realize this system, position of human should be recognized by Intelligent Space. In our system, Intelligent Space is equipped with multiple cameras to detect position of human. To estimate position of human in Intelligent Space using multiple camera system, we propose a human detection method based on the Histogram of Oriented Gradients(HOG) feature descriptors and the boosting technique in each camera and a position estimation method using multiple view geometry.

2. SYSTEM OVERVIEW

Fig. 1 shows comprehensive outline of newly designed active information display system. When a human seeks new information, the system obtains the location and gaze of user using cameras, and determines the projected area to display visual information, and calculates the goal position of the robot. After that, the mobile robot moves to the goal position and provides visual information without occlusion to the user by using a projector. To reduce distortion and to provide intact visual information the warped image is projected on the direction where a user looking.

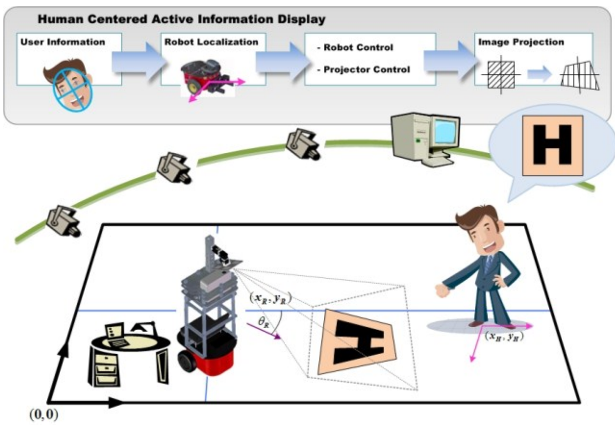


Fig. 1 System Overview

In order to realize this process, our system consists of Intelligent Space and Ubiquitous Display (i.e., a mobile robot with a projector).

Intelligent Space is an environmental system, which supports people in it, informationally and physically. The space can detect state-changes by using network sensors or control equipments connected via network. In Intelligent Space, network devices and data obtained from them are shareable. Thus, Intelligent Space can reduce some of the equipments installed on a physical agent such as a service robot. In general, a stand-alone robot has many sensors for perceiving circumstances in an environment, and builds a map using data obtained around the space, and estimates the robot's location by itself. On the other hand, a robot in Intelligent Space can perceive not only information about the robot, but also other information about circumstances in the space. So, Intelligent Space is a good platform for a network robot[4,5,6].

As a physical agent in Intelligent Space, Ubiquitous Display(UD) executes tasks in the physical domain to support people in the space. UD consists of five parts; a projector, a pan-tilt mechanism, a power supply, a mobile robot and a notebook. The main role of UD moves to a user, and project visual information on where the user looks at. The user can interact with Intelligent Space through UD.

In our system, we use multiple cameras to detect state-changes. Some cameras are used to find the location of human and robot. Others recognize the human's face and the human's gaze. Each camera is connected to a desktop through Giga Ethernet, and the desktops are connected via network, so that the information that the desktops process are shared. And we developed UD-m which is a mobile robot with built-in projector similar to the UD. The difference between UD and UD-m is a pan-tilt mechanism to rotate a mirror instead of a projector. Because, a mirror is much smaller and lighter than a projector .

3. Human Position Estimation using Mulitple Cameras in Intelligent Space

In the active information display system, the information of human positions in 3D space should be obtained by Intelligent Space. To find position of human in Intelligent Space using multiple camera system, firstly, the position of human on image coordinates should be detected in each camera. Multiple cameras, which are connected each other via a network, are able to detect 2D points of the objects. Secondly, a process of estimating 3D position from the positions, which are detected in each camera, is needed. We approach this process from multiple view geometry. In multiple view geometry, a relation between 3D point and image point is interpreted by pin-hole camera model. 3D positions of human can be estimated by camera model. Using each 2D points of the human from cameras, human positions in 3D space can be estimated.

3.1 Human Detection

Our human detection algorithm consists of three major steps (see Figure 2), as listed below.

1. Build template-based HOG feature set: We setup a set of templates that summarize the response of each template patch using one HOG[7], which represents marginal distribution of the patch. To speed up the HOG computation, we use “integral image” proposed by Viola and Jones[8] from one-dimensional to d-dimensional integral image, called “Integral Histogram Image”(IHI).
2. Design classifiers for the template HOG features: For each template patch, the classifiers can well distinguish between human and non-human. Also, these must be simple algorithm to reduce operation for an exhaustive template set. We consider three classifiers consists of principal components analysis (PCA), fisher linear discriminant (FLD) and support vector machines (SVM).
3. Choose classifiers by AdaBoost learning to construct human detector: The best classifiers for template HOG features able to distinguish human and non-human samples are chosen by AdaBoost learning. Finally, to construct human detector, we utilize a classifier with a ensemble combination of the chosen classifiers by AdaBoost.

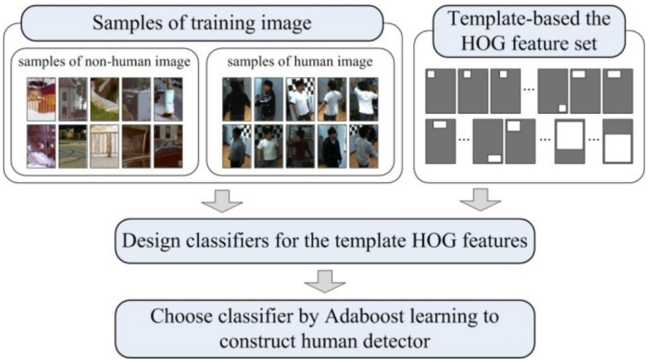


Fig. 2 Framework to construct human detector

3.2 Estimating Position of Human

To find 3D position of human in intelligent space using multiple camera system, firstly, the position of human on image coordinates should be detected in each camera using the human detection method described in previous section. Secondly, a process of estimating 3D position from the positions, which are detected in each camera, is needed. We approach this process from multiple view geometry [9]. In multiple view geometry, a relation between 3D point and image point is interpreted by pin-hole camera model, generally. Using pin-hole camera model, the rays from the focus of camera to human position detected on image in each camera can be obtained. And, 3D position is estimated by the rays like Fig. 3.

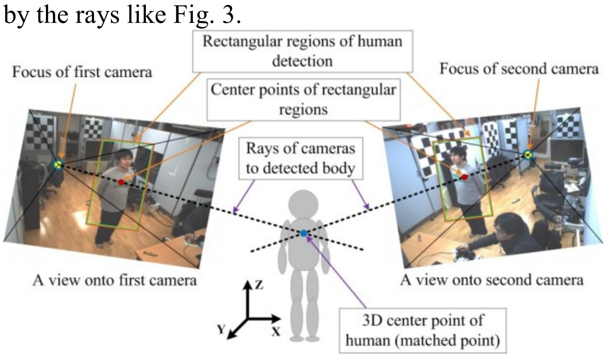
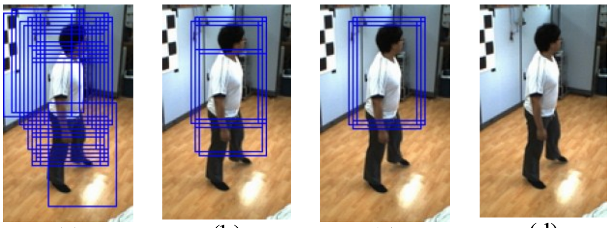


Fig. 3 The conceptual process of estimating the 3D position

4. EXPERIMENTS

4.1 Human Detection

Fig. 4 shows human detections using cascade classifiers with various hit rates. In this figure, there is a relative trade-off between the false positive rate and hit rate. The cascade classifier with high hit rate has just not best performance.



(a) (b) (c) (d)

Fig. 4 Examples of human detections using cascade classifiers with various hit rates.

(a) with 0.992^{14} hit rate (b) with 0.93^{14} hit rate

(c) with 0.89^{14} hit rate (d) with 0.80^{14} hit rate

4.2 Estimating Position of Human

Fig 5 is a diagram of the demonstration room setup for the our Intelligent Space. The area of the room is $5.6 \times 2.9\text{m}^2$. Four of cameras were mounted at approximately 2.1m from the floor for estimating position of human in the entire room.

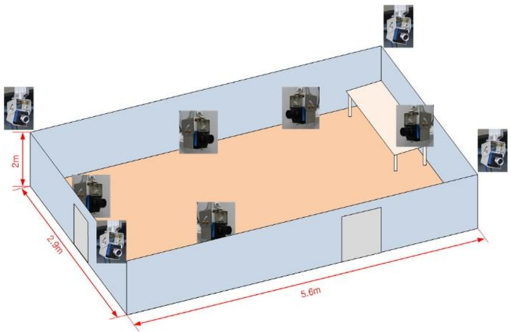
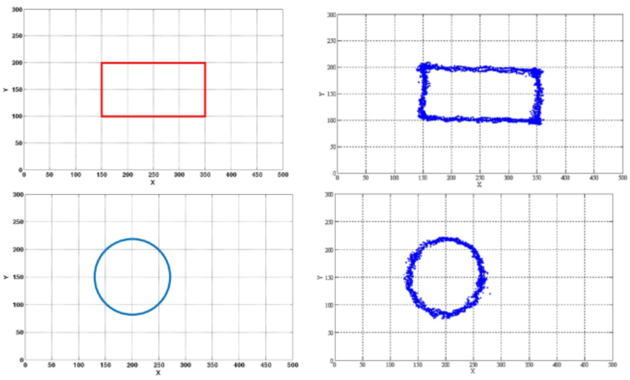


Fig. 5 Setup of our Intelligent Space.

For experiments of estimating human position in our Intelligent Space, two types of walking paths are setup as shown in Fig. 6(a) and Fig. 6(b) shows the results of estimated positions. In table 1, we summarize the result of average error between walking paths and estimated human positions. Fig. 7 shows the examples of estimated positions.



(a) two types of walking paths (b) the results

Fig. 6 Walking paths and the experimental results.

Table 1 The result of average error.

Type of walking path	Average error
Type-1 (rectangle)	3.1 *cm*
Type-2 (circle)	21.7 *cm*

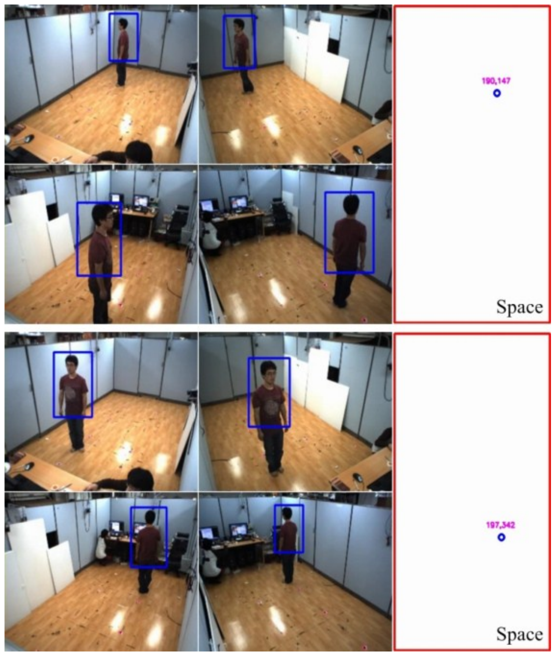


Fig.7 The examples of estimated positions.

5. CONCLUSIONS

In this paper, we described a human position estimation method using multiple camera system in Intelligent Space. Firstly, the position of human on image coordinates was detected in each camera. Multiple cameras, which are connected each other via a network, were able to detect 2D points of the objects. Secondly, a process of estimating 3D position from the positions, which are detected in each camera, was suggested using multiple view geometry. In experimental results, we presented that our method have produced the results of the human positions with reasonable distributed range. But, further studies are required to utilize the proposed methods.

REFERENCES

[1] Joo-Ho Lee, "Human Centered Ubiquitous Display in Intelligent Space", *The 33rd Annual Conf. of the IEEE Industrial Electronics Society(IECON)*, 2007.

[2] S. Miyashita, Joo-Ho Lee, "Ubiquitous Display for Human Centered Interface -Fixed Shape Projection and Parameter Optimization-", *Proceedings of the 17th IFAC World Congress*, 2008.

[3] Jeong-Eom Lee, S. Miyashita, K. Azuma, Joo-Ho Lee, and Gwi-Tae Park, "Anamorphosis Projection by Ubiquitous Display in Intelligent Space", *HCI International 2009*, 2009.

[4] Joo-Ho Lee, Kazuyuki Morioka, Noriaki Ando, Hideki Hashimoto, “Cooperation of Distributed Intelligent Sensor Network”, *IEEE Trans. on Mechatronics*, vol. 9, no. 3, Sep. 2004.

[5] Joo-ho Lee and Hideki Hashimoto, “Controlling Mobile Robots in Distributed Intelligent Sensor Network”, *IEEE Trans. Industrial Electronics*, vol. 50, no. 5, pp.890-902, Oct. 2003.

[6] Joo-Ho Lee, and Hideki Hashimoto, "Intelligent Space - Concept and contents", *Advanced Robotics*, Vol. 16, No. 3, pp. 265-280, 2002.

[7] N. Dalal, B. Triggs, “Histograms of oriented gradients for human detection”, Proc. Computer Vision and Pattern Recognition, 2005.

[8] P. Viola, M. Jones, “Robust real-time object detection”, IEEE ICCV workshop on statistical and computational theories of vision, 2001.

[9] R. Hartley, A. Zisserman, “Multiple View Geometry in Computer Vision”, Cambridge University Press, 2003.