VISION - BASED HUMAN TRACKING WITH COLOR DETECTION FOR AUTONOMOUS MOBILE ROBOT

[1]THEINT THEINT HTWE, [2]WUT YI WIN, [3]LEI LEI TUN SHWE

[1,2,3] Mechatronic Engineering Department, Mandalay Technological University, The Republic of Union of Myanmar
E-mail: [1]theinttheintstar@gmail.com, [2]wutyiwin@gmail.com, [3]leilei89.mdy@gmail.com

Abstract - This paper proposes to build an autonomous human tracking mobile robot which can solve the occlusion problem during tracking. The robot can make human tracking efficiently by analysing the information obtained from a camera which is equipped on the top of the robot. The system performs human detection by using Histogram of Oriented Gradient (HOG) and Support Vector Machine (SVM) algorithms and then uses HSV colour detection system for detecting stranger. The color detection method for tracking humans in indoor environments providing the target color is distinctive from the environment. Kalman filter can estimate the target human when the human is occluded with walls or something. The experimental results are given to prove the effectiveness of the proposed method.

Keywords - color detection, HOG, HSV, Human tracking, Kalman filter, Occlusion problem, SVM

I. INTRODUCTION

For a mobile robot that performs autonomously, the communication between the human and the robot is the most important factor. The main difficulty in this kind of work is that the detection of the target is a sensitive thing to carry out. To perform this task accurately, robot needs a mechanism that enables it to visualize the human and act accordingly [1]. The robot must be intelligent enough to track the human in the environment. When the autonomous mobile robot moves, it depends on vision system to detect the objects. So vision system is one of the key technologies for intelligent robot.

Typically human tracking robots are equipped with several different sensors i.e. laser ranger finder (LFR), ultrasonic sensors, infrared (IR) sensing modules, thermal imaging sensors, camera, etc. for recognition and locating the target [2]. All the sensors and modules work in unison to detect and track the target. An intelligent decision is being made by the robot control unit based on the information obtained from the above sensors and modules.The robot moves within the environment by sensing and coming across its way by following the target human.

In this approach, information is taken from a vision sensor. Object detection is done by utilizing a mobile robot with vision sensor and is achieved by image processing techniques. In order to track the human, both methods for the initial detection of humans and the subsequent following need to be implemented. The detection system presented here is based on HOG and SVM which is trained to detect human in the video frame. After finishing the human detection process, the system performs stranger detection by using HSV colour detection algorithm. Then, the robot will track the stranger. During the process of target human tracking, the robot follows the stranger all the time not to be missed from its vision. The stranger can be occulded with something during tracking. To solve this problem, the system will track target human by using Kalman filter. The Kalman filter (KF) has commonly been used for estimation and prediction of the target human position in succeeding frames. This target human tracking system is activated after a successful detection. With this approach, the robot can even track a human who is hidden or occulded with something in the camera vision.

Most human tracking systems have been designed to operate in indoor and outdoor environments. When a human tracking algorithm is implemented in an outdoor environment, sensors will be exposed to more noise than indoor. In such unstructured environments, human tracking is a difficult task. Human tracking approaches have been developed with most based on visual tracking. So, this robot is intended to use in indoor environment. The main advantage of using this system is the increase in reliability and efficiency provided by target human detection algorithm and then using with Kalman filter. The paper is organized as follows. Section II depicts related works for human tracking system. Section III details importance functions of human tracking robot with vision sensor. Section IV presents the testing results with mobile robot. Finally, section V summarizes the contributions and discusses future extension

II. RELATED WORKED

Autonomous robots which work without any human interpolation are required in robotic fields. The robot has to move in any environment even in the one it has never seen before. This robot is essentially designed to move on a floor. The capability of a robot to track and follow a moving object can be used for several purposes such as helping humans, creating ease for people and using for defence purpose.

Laser sensor is used by Burgard in his tour guide robot for human tracking [3]. The Kinect sensor is also exploited for person following applications as in

[4]. The Kinect sensor measures the depth information by the use of infrared structured light. The field of view of Kinect sensor is narrower than that of other sensors which makes it less applicable for detection and tracking applications. Laser based approaches use scanning laser range finders to detect and track person [5], [6], [7]. This technique is simple to implement, requires little processing power, and allows for fast tracking. However, detection is limited to a single plane. Numerous objects such as trees and poles can be falsely detected, and following multiple people who cross paths is difficult. Depth imaging was used by SongminJia to carry out the detection. The model of a person was determined using the depth imaging [8]. The particular style of clothing was used by MehrezKristou. He used a multidirectional camera. LRF was also incorporated by him in the design [9]. A research was conducted by Wilhelm with the focus on the color of the particular person□s skin. Information from different sensors was also used by him in the research [10].

In this paper, the system is presented a method of a human tracking robot based on human feature identification and clothes' colours detection by using a camera. Intelligent tracking of specified target is carried out by the use of vision sensor.

III. THE PROPOSED VISION-BASED HUMAN TRACKING SYSTEM

The autonomous capability of mobile robots is a social and technological issue. Many potential applications add to the importance of human tracking technology. In this paper, the system proposes the detection and tracking schemes for human by the use of a camera. Mobile robot's camera captured the image sequences which are used as the input to the system. The robust target human tracking is achieved by exploiting human features. Human features are extracted by using histogram of oriented gradient (HOG). The detection and tracking schemes are realized by the combined application of support vector data description algorithm and HSV colour detection algorithm. Then, the robot performs the tracking. During human tracking, Kalman filter is used for the pose estimation of the target person. It will predict next state from the previous state. So, the robot can track the target human even when the target person is out of vision or occluded with something in the environment. The proposed scheme is successfully verified in real time environments. The system block diagram of human tracking system is shown in **Fig 1**.

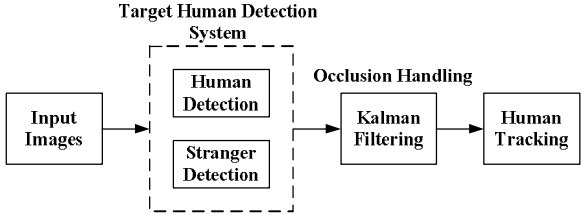


Fig.1. System block diagram

3.1. Target Human Detection System with Image Processing Techniques

For the image processing, a method which simultaneously uses to detect the human feature with HOG is the reliable and efficient method. In this section, the steps of image processing on a captured image are explained to detect the human feature and generate the target human to track.

Feature extraction is the main step of image understanding. Feature is a significant piece of information extracted from an image which provides more detail understanding of the image. Histogram of Oriented Gradients (HOG) is a technique to extract features from an image. Image processing algorithms for target human detection system is shown in **Fig. 2.**

The Support Vector Machines classifier is a binary classifier algorithm that classifies the detected object features as human or not. A set of training image is used determining with and without human, and described by their HOG, to learn a decision function. When the SVM is trained for human, it will classify the moving objects features as human or not.

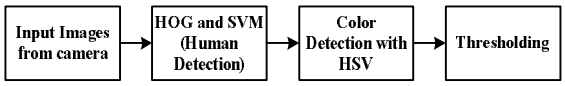


Fig.2. Image processing block diagram

To make human tracking more accurately, the robot has to make accurate human selection so that it can correctly track the human who is not wearing the school uniform. After detecting the detected objects with HOG and SVM, the system will make stranger detection by detecting colours of human's clothes. The mobile robot detects the stranger with HSV color detection algorithm. In this system, thresholding is used to filter the detected human's clothing color to be accurate at detection.

The system processing results for human detection is shown in **Fig.3**. In **Fig.3**, the mobile robot detects the entire human in the environment. The detection results of HOG and SVM are shown with green color box.

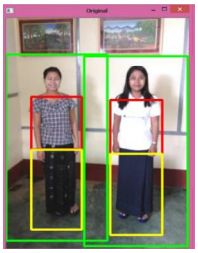


Fig.3. Processing result of Human detection

3.2. Color Detection System

Color detection technique can be robust against non stationary background. To make human tracking more accurately, the robot has to make accurate human detection so that it can correctly track the human who is not wearing the school uniform. After detecting the objects with HOG and SVM, the system will make stranger detection by detecting colors of human's clothes. The mobile robot detects the stranger with HSV color detection algorithm by separate color detection on the upper cloth color (white) and lower cloth color (dark blue) of the stranger so the robot can detect the human with less error to make human tracking. Color information is commonly represented in the widely used RGB coordinate system. This representation is suitable for acquisition or display devices but not particularly applicable in describing the perception of colors. On the other hand, the HSV (hue, saturation, value) color model corresponds more closely to the human perception of color. To be robust to the change in brightness, the color histogram of hue *h* and saturation *s* is made by using pixel values converted into HSV. Therefore it is applicable in the case of mobile robot.

These color detections are shown with upper cloth (red box) and lower cloth (yellow box).The system makes color detection by comparing the HSV color assign range with the detected human's upper and lower cloth color range.. The color detection results with HSV algorithm are shown in **Fig.4.**

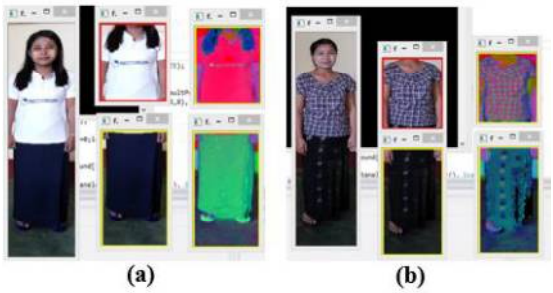


Fig.4. Processing result for stranger detection with HSV

3.3. Human tracking based on Kalman filter

In vision-based target human tracking, the system uses Kalman filter to be more accurate. There can be many condition errors because of the targeted human's movements. The targeted human can sometimes out of vision from the camera. Kalman filter can solve this problem by pose estimation of the targeted person all the time. Human tracking starts after *N* video frames have been used for the initialization of tracking based on Kalman filter. *N* is the number of the first *N* consecutive frames in which human detection was successful. In order to fulfill this task, the tracking system works under the assumption from the target human between consecutive frames and therefore it is able to correctly locate the selected human in successive frames. This holds as long as the humans in the scene can be clearly distinguished. However if the target human temporarily occludes with other human or something, target human can't be in the field of the camera's view any more, but the proposed Kalman filter predict target human's position continuously.

The Kalman filter addresses the general problem of trying to estimate the state of a discrete-time controlled process that is governed by the linear difference equation

$$\hat{X}_{k|k-1} = A\hat{X}_{k-1|k-1} + Bu_k \quad (1)$$

$$P_{k|k-1} = AP_{k-1|k-1}A^T + Q \quad (2)$$

$\hat{X}_{k|k-1}$ is the initial state or previous state estimate, is control input. A and B are state matrix which applies the effect of each system state parameter got from human detection system and control input matrix which applies the effect of each control input parameter. In practice, $P_{k|k-1}$ is the covariance matrix. The process noise covariance, and measurement noise covariance, *R* matrices might change with each time step or measurement, however here the value of R assumes constant. Y_k is the measurement value.

$$J_k = \frac{P_{k|k-1}C^T}{(CP_{k|k-1}C^T + R)^{-1}} \quad (3)$$

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + J_k(Y_k - C\hat{X}_{k|k-1}) \quad (4)$$

$$P_{k|k} = (I - J_kC)P_{k|k-1} \quad (5)$$

The matrix *A* in the difference equation (1) relates the state at the previous time step *k-1* to the state at the current step *k*, in the absence of either a driving function or process noise. Note that in practice *A* might change with each time step, but here it is assumed as constant. The matrix *B* relates the optional control input to the state x. In practice, measurement, *Y* might change with each time step or measurement, but here it is assumed as constant. The time update equations project the state and covariance estimates forward from time step *k-1* to step *k*. The first task during the measurement update is to compute the Kalmangain, J_k. The next step is to actually measure the process to obtain Y_k, and then to generate a posteriori state estimate by incorporating the measurement as in equation (4). The final step is to obtain an a posteriori error covariance estimate via

equation (5). After each time and measurement update pair, the process is repeated with the previous a posteriori estimates used to project or predict the new a priori estimates. Kalman filter makes practical implementations much more feasible. Flowchart of the Kalman filter used for human tracking system is shown in **Fig.5**.

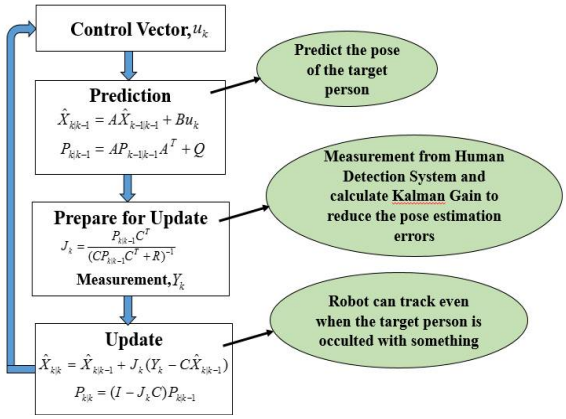
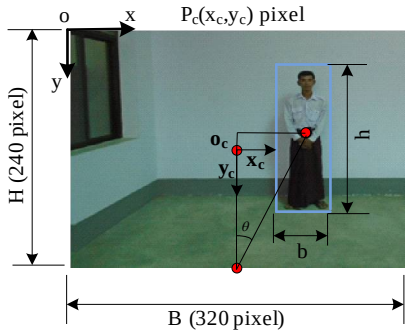


Fig.5. Flowchart of the Kalman filter used for human tracking system

3.4. Position of target human in coordinate frame fixed to a camera for tracking

The position of a human in the camera coordinate frame fixed to a camera is shown in **Fig. 6.** It is necessary to find a position of a human in the coordinate frame fixed to a camera in the case of a human tracking using the whole photo image captured by the camera.



o-xy: Global coordinate frame
oc-x_cy_c:Center of the camera coordinate frame
Pc-x_cy_c: position of the detected human center

Fig.6. Position of a human in the camera coordinate frame fixed to a camera

The o_c-x_c,y_c coordinate frame denotes the camera coordinate frame fixed to the camera. The Pc (x_c , y_c) [pixel] denotes the position of the detected human of center. Where *B* is breadth [pixel] and *H* is height [pixel] of a whole photo image and *h* is the high and *b* is the breadth of the detection window. It is possible to calculate the coordinates (x_c , y_c) [pixel] and can calculate the P_c-x_c,y_c,[pixel] because the whole photo image is two-dimensional image. From this pixel differences the angle (Θ) can be calculated by using triangulation method.

EXPERIMENTAL RESULTS

In order to validate the accuracy of the vision based target human tracking system, an experiment was planned in indoor environment. The true location of the target human with respect to the robot was then derived from the analysis of the image processing system. The mobile robot equipped with a camera which is used to recognize the target human and people around the robot. The information of the robot is used in real time by the Kalman filtering algorithm installed in the control system.

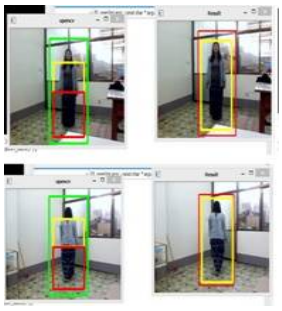


Fig. 7. The processing results for human detection system using Kalman



(a) (b)

Fig. 8. The processing results for detection system using Kalman filter (a)Stranger detection, (b)Occlusion condition

In target human tracking system, the system detects all moving objects to know whether it is human or not by detecting with HOG and SVM and then makes stranger detection with HSV color detection algorithm. The system processing results for detection system is shown in **Fig.7 and Fig.8**. The mobile robot detects the entire human in the environment. The detection results of HOG and SVM are shown with green colour box. Then the system detects the human upper and lower cloths colours by using HSV colour detection algorithm. These colour detections are shown with upper cloth (red box) and lower cloth (yellow box). The system makes colour detection by comparing the HSV colour assign range with the detected human's upper and lower cloth colour range. This system uses Kalman filter to be more efficient in tracking method. KF can perform well even when the target human occulded with something during its motion. The human tracking system using Kalman filter can give out more accurate results.



Fig. 9. The experimental results for human and stranger detection

The mobile robot is moving and detecting the objects in the environment. When the robot found the stranger, it will note that human as target human and make tracking. The experimental results for human and stranger detection are shown in **Fig.9**.



Fig.10. Target human Tracking

When the robot knows the target human, it will follow that human. It can still follow the target human who turns to one side or occluded with something by estimating the target human state with Kalman filter. So, it can't miss the target human during tracking. The results of target person following is shown in **Fig.10** and **Fig.11.**



Fig. 11. Target person following without missing in occluded condition

CONCLUSIONS

A method of human tracking with a mobile robot using a camera was presented in the paper. An effective implementation of human tracking robot is illustrated in this research. This robot does not only have the detection capability but also the tracking and following ability as well. And then, the robot can track the target human simultaneously whenever the target human is occluded with something. The tracking is basically performed on the human's upper and lower clothes colors and then tracks the human on the basis of that detection. This system showed that the use of color detection reduces the errors of the tracking system and can give more accurate results.

ACKNOWLEDGMENTS

The author would also like to thank all the teachers from the Department of Mechatronic Engineering, Mandalay Technological University (M.T.U). And she also would like to thank her family and her friends who have helped her throughout the year.

REFERENCES

[1] K. Morioka, J.-H. Lee, and H. Hashimoto, "Human-following mobile robot in a distributed intelligent sensor network," IEEE Trans. Ind. Electron., vol. 51, no. 1, pp. 229–237, Feb. 2004.

[2] H. Takemura, N. Zentaro, and H. Mizoguchi, "Development of vision based person following module for mobile robots in/out door environment," in 2009 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2009, pp. 1675–1680.

[3] W. Burgard, A. B. Cremers, D. Fox, D. Hähnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun, "The interactive museum tour-guide robot," in AAAI/IAAI, 1998, pp.

[4] http://www.ros.org/news/2010/12/kinect-based-personfollower.html

[5] AjoFod, Andrew Howard, and Maja J. Matari´c. A laser-based peopletracker. In IEEE International Conference on Robotics and Automation,pages 3024–3029, Washington DC, May 2002.

[6] WoojinChung,HoyeonKim,YoonkyuYoo,Chang-Bae Moon and Jooyoung Park, "The Detection and Following of Human Legs Through Inductive Approaches for a Mobile Robot With a Single Laser Range Finder,"IEEE Trans.Ind .Electron, Vol. 59, no. 8,pp. 3156-3165, Aug 2012.

[7] AnandPanangadan, Maja J. Matari´c, and Gaurav S. Sukhatme. Detecting anomalous human interactions using laser range-finders. In IEEE/RSJInternational Conference on Intelligent Robots and Systems, pages2136–2141. IEEE Press, Sep 2004.

[8] S. Jia, L. Zhao, and X. Li, "Robustness improvement of human detecting and tracking for mobile robot," in Mechatronics and Automation (ICMA), 2012 International Conference on, 2012, pp. 1904–1909

[9] M. Kristou, A. Ohya, and S. Yuta, "Target person identification and following based on omnidirectional camera and LRF data fusion," in RO-MAN, 2011 IEEE, 2011, pp. 419–424.

[10] T. Wilhelm, H. J. Böhme, and H. M. Gross, "Sensor fusion for vision and sonar based people tracking on a mobile service robot," in Proceedings of the International Workshop on Dynamic Perception, 2002, pp. 315–320.

★★★