ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

Improved Hand Tracking System Based Robot Using MEMS

M.Ramamohan Reddy
P.G Scholar,
Department of Electronics and communication Engineering, Malla Reddy College of engineering.

I.V.Prakash, Ph.D
Associate Professor,
Department of Electronics and communication Engineering, Malla Reddy College of engineering.

ABSTRACT:

This paper presents an improved hand tracking system using pixel-based hierarchical-feature Ada Boosting (PBHFA), skin color segmentation, and codebook (CB) background cancelation. The proposed PBH feature significantly reduces the training time by a factor of at least 1440 compared to the traditional Haar-like feature. Moreover, lower computation and high tracking accuracy are also provided simultaneously. Yet, one of the disadvantages of the PBHFA is the false positive which is the consequence of the appearance of complex background in positive samples.

To effectively reduce the false positive rate, the skin color segmentation and the foreground detection by applying the CB model are catered for rejecting all of the candidates which are not hand targets. As documented in the experimental results, the proposed system can achieve promising results, and thus it can be considered as an effective candidate in handling practical applications which require hand postures.

I. INTRODUCTION:

HAND POSTURES are powerful means for communication among humans. They facilitate a recipient to understand what a speaker is communicating. In daily life, some common postures are also used, such as “Stop,” “OK,” “Yes,” and “No.” Many applications are designed by using the motion of hand, such as human–computer interfaces, robot control, and communication with the deaf. Yet, to use the motion of hand for supporting visual aspects of interaction, it is necessary to track the hand in real time. In general, the main challenge of building a real-time system is to maintain a good balance between the accuracy and the processing time. The objective of these applications requires fast detection to process information carried by the hand postures while maintaining a good detection rate.

In particular, the hand tracking is rather difficult because most of the backgrounds change across frames, and the hand shape is a complex object to detect Kolsch and Turk [1] proposed a new method to detect a hand using image cues obtained from the optical flow and a color probability distribution. The pyramid-based Kanade–Lucas– Tomasi (KLT) feature tracking was employed as their first modality because it shows excellent performance on quickly moving rigid objects, and can be processed very efficiently [2], [3].

The method can deal with dynamic backgrounds and light condition changing. Moreover, it can follow a rapid hand movement despite arbitrary finger configuration changes. However, the color of a hand must be learned with a normalized-RGB histogram, which is contrasted to the background color in the image around the hand. It means that there is no other part with skin color from the same person appearing in the reference background area. Thus, this approach cannot classify wooden objects that are not within the reference background area during learning.

As a result, the wooden objects will be considered as foreground. Zhu et al. [4] introduced a novel statistical approach to hand segmentation based on Bayes decision theory. The Gaussian mixture model with the restricted expectation and maximization algorithm [5] were used to build the color models of a hand and background for a given image.

These models can classify each pixel in an image as either a hand pixel or a background pixel. The advantage of this approach is able to segment hand region from various backgrounds and light conditions with unknown color of a hand. Unfortunately, the method proposed in [4] cannot distinguish multitargets in an image and the error rate of 11.5% is a bit too high. Moreover, it takes about 0.1 s for each image on a personal computer platform, which does not support real-time scenarios. Binh and Ejima [6] applied skin color tracking for face detection to extract the hand region by separating a hand from a face. An image is classified into regions according to the human skin color by threshold with a proper threshold value. Subsequently, the face region is detected and removed to reserve the hand region. By doing this, the hand regions of each input image can be extracted in real time.

Other skin models such as [7] can be embedded into the above systems to achieve good skin segmentation. Nevertheless, these approaches have the same problem as those methods using human skin detection; they cannot distinguish between skin regions or some objects with similar color or shape as the hand. Thus, this is not a reliable modality, which requires users to wear long-sleeve shirt to avoid the problem of misclassifying the arm region as a part of the hand.

Other approaches were proposed to use a learning-based method such as AdaBoost, Gaussian mixture model (GMM), and SVM [8]. AdaBoost is a well-known algorithm to maintain a good classification capability. With this approach, choosing feature for classification is highly important. Local binary pattern (LBP) [9] is one of the powerful features with low computation. Its grayscale invariant characteristic has coped with many of practical computer vision problems.

Another feature used in Chen et al.'s work [10], called Haar-like feature [11], was adopted for hand detection. They collected 1712 positive samples (including four different postures) with different scales, and 500 negative samples (only backgrounds) for experiment. Compared with other methods which operate on multiple scales, this approach has high accuracy, which reduces the processing time by using the integral image.

However, as described in their paper [10], the positive samples used for training do not contain any complex background, and the results show that the detection is simply on uniform background. Thus, the background subtraction is required before detecting, and the Gaussian filter is employed to reduce noise.

Thus, the overall performance of the system highly relies on the subtraction. Just et al. 12] also employed AdaBoost for both of the hand classification and recognition.

An additional feature, called modified census transform (MCT), was used to transform the original image into another feature space using 3×3 kernels and then comparing to its neighborhood. Then, the two-class strong classifier of the AdaBoost is trained to distinguish the posture and nonposture. The training and testing sets include images of size 30×30 for ten postures. However, in the experiment, images are cropped and the posture is put perfectly centered in each image, and this limitation does not guarantee a good detection for different scales when scanning image in real time.

Donoser and Bischof [13] applied appearance-based approach to combine a state-ofthe- art interest point tracker with efficiently calculated color likelihood maps. First, the color distribution of the hand is built by using the Gaussian model. Second, the color distribution of the hand is combined with maximally stable external region (MSER) detector. The MSER is one of the good interest region detectors in computer vision proposed by Mikolajczyk– Schmid [14]. The color likelihood map is used to compute ordered set for each pixel. Based on the pixel ordering, MSERs connect regions which can be detected in any image and the hand region can be segmented with high accuracy. After detection phase, the system tracks the hand based on the detected targets in frame. However, the system cannot deal with the non-hand target that is similar to the color model.

II.PROPOSED HAND DETECTION SYSTEM:

One of the differences between hand detection and face detection is that most of the positive samples officially provided in hand detection are with complex background behind the objects of interest. Yet, most of the official positive samples of face contain only face area; most of the positive samples for training contain various backgrounds as indicated with the green squares and these areas are considered as a part of the object. And thus reduces the tracking rate. Nonetheless, the appearance of background in the positive training samples can benefit the proposed method to gain the advantage of detecting hand against complex background. This is the main difference to other methods when most of the former approaches tried to separate the pure hand region from background, which made their systems not to solve the complexity background problem effectively. Thus, the proposed hand tracking system employs the combination of pixel-based performance.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

III. DESIGN OF MEMS ROBOT :

The design of fabricated MEMS micro robot is shown in Figure 1. The number of the legs of the micro robot was 6. The structure and the step pattern of the robot was emulated those of insects. The micro robot consisted of frame parts, rotary type actuators and link mechanisms. The rotary type actuator generated the locomotion of the robot by supplying the electrical current to the artificial muscle wires. The size of the micro robot fabricated by the MEMS technology was designed as 4.0, 2.7, 2.5 mm, width, length, height.

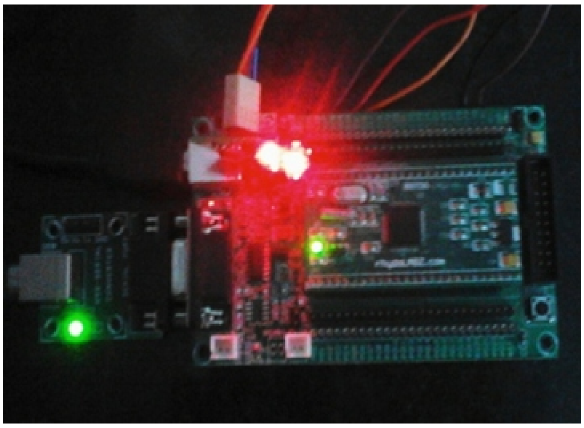

Figure 1: MEMS Robot

IV. HSV COLOR SPACE:

In the practical usage stage, a detected hand by the Ada Boost algorithm is further filtered by the skin color tore duce the false positives. The range of the skin colors heavily rely on the lighting conditions, thus the HSV color model which is robust to lighting change is adopted for skin color localization. One of the most important advantages of this color model in skin color segmentation is that it allows users to intuitively specify the boundary of the skin color class in terms of the hue and saturation. In this paper, the ranges of the hue and saturation are set in between 0° and 5° and 0.23 to 0.68, respectively, as specified in the segmentation results.

V. FOREGROUND DETECTION:

The first level of tracking is the detection of targets of interest. A successful approach toward this detection is the foreground–background segregation.

Unfortunately, most of the background and foreground in the video sequence are non stationary in practice, such as waving trees, rippling water, and light conditions. One of the popular methods in foreground detection is the mixture of Gaussian. The main idea of the GMM is the construction of the background distributions. These distributions are used to classify the pixels to foreground and background. Yet, the MoG still has some disadvantages: the low learning rate makes it difficult to cope the sudden change problem.

Conversely, slowly moving objects will be considered as background. Moreover, it normally causes high false positive rate. Proposed the CB model for foreground detection. The advantages of the CB include fast processing speed, capability to handle scenes with dynamic backgrounds, and robust to handle illumination variations. Also, by using the CB, the background can be classified into many layers based on the applications. The concept of the CB is to train background pixel wise over a period of time. Sample values at each pixel are clustered as a set of code words. The combination of multiple code words can model the mixed background.

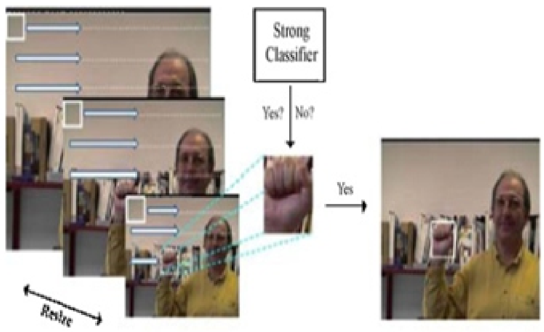

Figure 2: Example of detecting hands of various sizes.

VI. HAND TRACKING METHODOLOGY:

Hand tracking is an application in field of object tracking. In the proposed system, the hand tracking phase is the second step after hand detection. Many algorithms for objects tracking were presented in the literature. These methods can be divided into three categories point tracking (Kalman filter), kernel tracking (mean shift and KLT tracker) and silhouette tracking (variation method and condensation algorithm).

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

In this paper, the Euclidean distance is chosen to track the hand because of its simplicity and effective result in practice. The center of each hand is denoted as and the hand area is limited within a square with the parameter Eucliddis where the variable is denoted on the bottom-right corner of The green line in indicates the trace of the hand in previous frames.

VII. HAND TRACKING:

The main objective of this paper is to track a hand. In this implementation, the Logitech webcam is used for testing in real time, and the captured images are of size 320 × 240 under 15 f/s with natural lighting conditions. The HP laptop with Core 2 Duo 2.4GB and 4GB RAM is employed as the platform. Also, the tracking methodology guarantee that only the hand of interest is tracked even there are many targets in the frame.

Shows another scenario in which there are two targets in frame 1514, and the "small hand" is the object we are interested. In frame 1546, the "small hand" disappears from the scene, and it simply leaves the "big hand" as a target. Nonetheless, the system does not track with the "big hand" instead because it is not the tracked object in the previous frame. In frames 1584 and 1619, when the "small hand" appears again, the system recognizes it as the object of interest, and keeps tracking.

Because the distance between the centers of a tracked object of consecutive frames is often very short, the system mostly can track the correct hand among many interfering targets. The advantages and disadvantages among various methods in the field of hand tracking. Complex background is the key issue of the hand detection, and most of the former researchers tried to solve this problem. For this, the proposed system can be considered as a good candidate for providing both efficient processing time and detection accuracy.

VIII. EXPERIMENTAL RESULTS:

In this paper, the public Sebastian Marcel's hand posture database is employed in our experiment, and some of the samples are shown in Fig. 11. In this database, six different postures are involved, including "A," "B," "C," "Point," "Five," and "V," and the numbers of each posture in the database are organized in Table I. and the numbers of each posture in the database are organized in Table I.

However, the samples in both "Point" and "Five" sets have a wide variety of shapes, and some of them are even in unreasonable shapes. Thus, the two datasets are excluded in our experiments. In the reduced database, 80% of them are adopted for the training, and in total 200 PBH-based weak classifiers are employed for the following experiments. Moreover, the remained database (20%) is adopted for the following simulation.

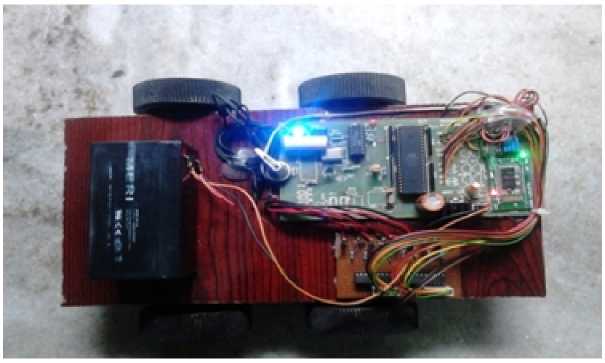

Figure 3: Robot Movement using MEMS.

IX. CONCLUSION:

In this paper, a hand tracking system was proposed by using the PBHFA, skin color segmentation, and CB model. The goal of this paper was to use PBH feature to reduce the required training time and further reduce the required computation in tracking phase. According to the experimental results, the above tasks were achieved; meanwhile the tracking accuracy was still maintained in high level as that of the Haarlike feature. The superiorities of the PBH also induced some benefits for the applications.

1) Short training time:

This feature makes the proposed features that can be applied for self-learning hand tracking system, since the proposed features only require little time for adapting with unusual positive hands.

2) Low computation complexity:

It indicates that the proposed features algorithm can also be embedded on lower price systems. On the other hand, the combination of PBH, skin color segmentation, and CB model was employed to achieve some additional advantages.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Monthly Peer Reviewed Open Access International e-Journal

The task of the skin color segmentation was to separate hand region from background, which can remove most of the background. Yet, there were some backgrounds which had color in skin tone, and it was the main reason of applying the CB model. The advantage of the CB model was to be able to subtract moving foreground and background with high accuracy. The background was trained in every frame and therefore it detects the moving hand in front of a camera. As presented in experimental results, the system showed promising performance.

References:

[1] M. Kolsch and M. Turk, "Fast 2D hand tracking with flocks of features and multi-cue integration," in Proc. CVPRW, vol. 10. Jun. 2004, p. 158.

[2] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in Proc. Imag. Understand. Workshop, 1981, pp. 121–130.

[3] J. Shi and C. Tomasi, "Good features to track," in Proc. IEEE Conf. Comput. Vision Patt. Recog., Jun. 1994, pp. 593–600.

[4] X. Zhu, J. Yang, and A. Waibel, "Segmenting hands of arbitrary color," in Proc. IEEE Int. Conf. Automat. Face Gesture Recog., Mar. 2000, pp. 446–453.

[5] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. Royal Statist. Soc., vol. 39, no. 1, pp. l–38, 1977.

[6] N. D. Binh and T. Ejima, "Hand gesture recognition using fuzzy neural network," in Proc. Int. Conf. Graphics, Vision Image Process., Dec. 2005, pp. 362–368.

[7] A. Cheddad, J. Condell, K. Curran, and P. M. Kevitt, "A skin tone detection algorithm for an adaptive approach to steganography," Signal Process. J., vol. 89, pp. 2465–2478, Dec. 2009.

[8] C. Bergamini, L. S. Oliveira, A. L. Koerich, and R. Sabourin, "Combining different biometric traits with one-class classification," Signal Process. J., vol. 89, pp. 2117–2127, Nov. 2009.

[9] T. Ojala, M. Pietikaine, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Patt. Recog., vol. 29, no. 1, pp. 51–59, Jan. 1996.

[10] Q. Chen, N. D. Georganas, and E. M. Petriu, "Hand gesture recognition using Haar-like features and a stochastic context-free grammar," IEEE Trans. Instrument. Measurement, vol. 57, no. 8, pp. 1562–1571, Aug. 2008.

[11] P. Viola and M. J. Jones, "Robust real-time face detection," Int. J. Comput. Vision, vol. 57, no. 2, pp. 137–154, 2004.