17th International Conference
ENGINEERING MECHANICS 2011
Svratka, Czech Republic, 9 – 12May 2011

AN ENHANCED APPROACH IN MOTION DETECTION OF HUMAN MOVEMENTS

A. Guran*, A. Vahidi-Shams*

Abstract: *In this paper, we will present a robust and fast motion detection approach, which can localize and keeps track an object in motion under hard real-time constraints. To eliminate the undesirable problems such as lighting and Shadows we used some techniques. The enhanced digital image is very closer to the scene perceived by the human visual system. Furthermore, we could decrease the effect of little and Periodic movement in back ground such as the leaves move in the wind, Moving shadow or Sea wave in Presence or absence of the movement objects. Experiment results to certify this efficiency showed desirable process time and performed with high accuracy. Introduced Model might be advantageous in various tasks of object tracking such as Analysis of Fundamental Human Movements.*

Keywords: ***Computer vision, human detection, motion detection, human movements.***

1. Introduction

Currently, motion detection and human interaction detection is one of the most active research topics in machine vision (Azad et al., 2007, Bandouch and Beetz, 2009), which used in wide spectrum of applications in many areas such as human-robot interaction, smart surveillance, virtual reality and perceptual-interface (Krügeretal, 2007; Kulic et al., 2008). An extensive literature exists regarding motion analysis in video streams using fixed camera. There are various numbers of approaches involving statistical analysis of motion suggested over the years. Some publications discuss the effect of camera, light-setting and the color spaces have in the process (Blostein and Huang, 1991, Deutscher and Reid, 2005).Various techniques for moving object detection have been proposed. Commonly three approaches are background subtraction, temporal differencing and optical flow. Background subtraction is the most commonly used approach in presence of still cameras. The principle of this approach is to use a background model and compare the current frame with a reference. The foreground objects present in the scene are detected in this way. The statistical model based on the background subtraction is flexible and fast, but in this method the background scene and the camera are required to be stationary well (Schuldtetal, 2004). Temporal differencing is based on frames difference that attempts to detect moving regions by making use of the difference of consecutive two frames in a video sequence. This approach is very adaptive to dynamic environments, but in general performs a poor job of extracting the complete forms of certain types of moving objects (Bobick and Davis, 2001). The other approach is the optical flow which is an approximation of the local image movement and specifies how much each frame pixel moves between adjacent images. According to the smoothness constraint, the corresponding points in the two successive frames should not move more than a few pixels. For an uncertain environment, this means that the camera motion or background changing should be relatively small. The optical flow based methods are complex, but it can detect the motion accurately even without knowing the background (Yun et al., 2005). Edwin Land coined “Retinex” word for his model of human color vision, combining the retina of the eye and the cerebral cortex of the brain. The retinex Image Enhancement Algorithm is an automatic image enhancement method that enhances a digital image in terms of dynamic range compression, color independence from the spectral distribution of the scene illuminant, and color/lightness rendition. The digital image enhanced by the retinex ImageEnhancement Algorithm is much closer to the scene perceived by the human visual system, under all kinds and levels of lighting variations, than the digital image enhanced by any other method. Although retinex models are still widely used in computer

* prof. Dr. Ardeshir Guran, Ph.D. and Ing. Afra Vahidi Shams, M.Sc: Institute of Structronics, 275 Slater Street, 9th Floor, Ottawa Canada; e-mails: ardeshir.guran@mail.mcgill.ca, afra.vahidishams@gmail.com

vision, they have been shown not to accurately model human color perception (Land and Mccann, 1971). Periodic motion is ubiquitous in the natural world, with instances ranging from the simple harmonic movement, which could appear in the background of our scene. Also shadows' movement beside motion of targets can decrease accuracy of MDS (motion detector systems). Furthermore, because of our main aim to implement it in embedded DSP(digital signal processor), we needed more robust algorithm that is fairly simple to compute, has the potential to run in real-time, and is relatively immune to noise. Here, we present the modified and developed to eliminate the effects of these undesirable problems. The principal idea in this study is to integrate the benefits of these three above approaches and a robust Model is represented. The principal goal of this algorithmic rule is to separate the background interference and foreground information effectively and discover the moving object accurately.

2. Overview of the Method

System presented here was done according to following principles:

- The aim of the system is robust and fast motion detection and tracking for surveillance systems.
- Images are received from a stable or non stable video camera.
- System is localizing movement object on the bases of motion picture analysis in camera view.
- System is designed to work in real time (now on standard PC class computer).

2.1. Algorithm Flow

The proposed algorithm consists of seven stages for image acquisition, preprocessing, characteristic extraction and motion detection. Fig. 1 shows the process flow of the proposed robust motion detection algorithm. Each of these stages will be described in detail in the next Section (Shams and Guran, 2011).

3. Detailed Algorithm Description

Step1. Each image frame is taken from signal source. Since we are analyzing color pictures we got three brightness difference values for each base color (RGB).at this stage, input image is enhanced based on retinex method.

Step2. At this stage image is pre-processed. The main aim of it is to make image properties better for generating objects data. Removing shadow improves the accuracy, According to this fact and to do this, we first convert pixel information from the RGB (Red/Green/Blue) color space to the HSV (Hue-Saturation-Value) color space. HSV color space corresponds closely to the human perception of color (Hurlbert, and Wolf, 2002) and it has revealed more accuracy to distinguish shadows. Then, it tries to estimate how the occlusion due to shadow changes the value of H, S and V. The rationale is that cast shadow's occlusion darkens the background pixel and saturate its color.

Step3. In this step we needs binary image for entry the next layer, to achieve this goal, all pixels with values below the threshold are set to 0, all those above are set to 1. This is a simple way to remove the effects of any non-linearity in the contrast in the frame. However, the threshold value must be defined. It could be achieved by calculating the threshold based on a histogram of a given differential image. And again filtration is performed, this time using logic or median filters. Threshold can be described by the following equation:

$$Th = \left(\frac{\left(\sum_{x=1}^{n}\sum_{y=1}^{m}(x,y)\right)}{n\times m} - \sum_{i=1}^{n\times m} P_i \cdot (img_i - \mu)^2\right)/2, \mu = \sum_{i=1}^{n\times m} P_i \cdot img_i \text{ where n and m are size of image (img)} \quad (1)$$

Step4. Differential image is calculated as difference between current frame and current background (in our case last frame). For the first read frame the calculation is not performed, since at this moment the background bitmap is not available.

Step5. The aim of this phase to find association between objects form following levels (objects from two following frames of motion picture) and label the object as a movement one, not identified or as matched to some object form the previous level. At first, similar pixels are rejected. Next, for each pixel on current frame list the different pixel from the previous list is selected with the highest value.

Step6. Last layer is the background update. The background is then used in this step and updated recurrently. The idea of background generating is based on integration of information inserted by every new frame with background. It is suitable for inhibiting the little periodic movement effects. We repeated for 4 times each as one layer and then repeated for each layer.

Step7. At last, Comparison between two layers which shown in Fig. 1 leads to detection of movement objects and area of motion is shown with colorful rectangles. Steps are depicted in Fig. 2.

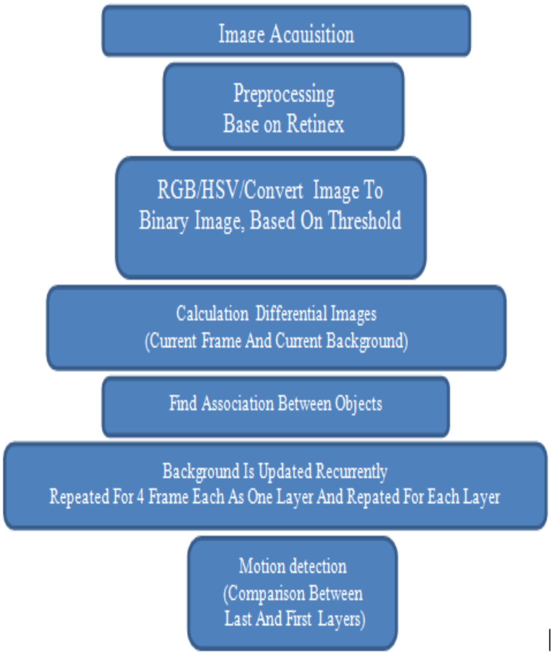

Fig. 1: The flow of algorithm.

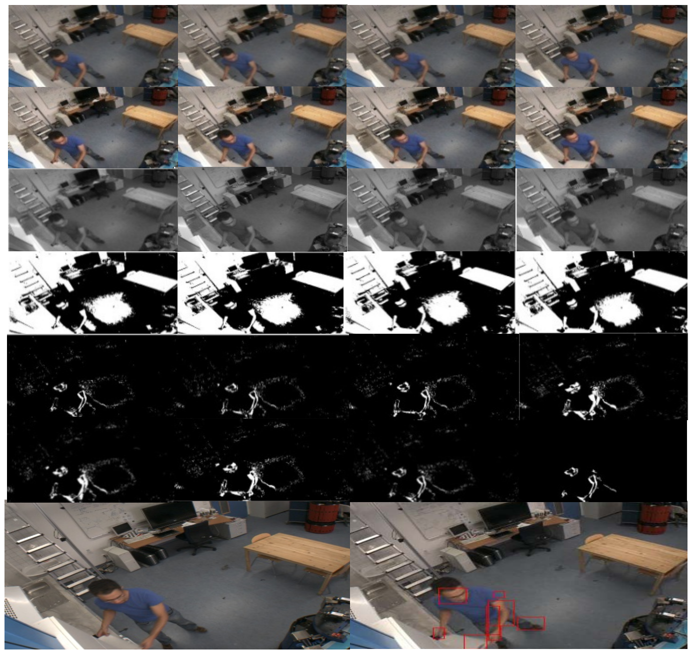

Fig. 2: Four frames and results of steps are depicted from top: a) frame 1 to 4 of each cycle b) enhanced image (Based on Retinex) c) gray value images d) binary images based on threshold e) differential between frames(L1: 1,2 & L2: 2,3 & L3: 3,4 & L4: L1,L2 & L5: L2,L3 & L6: L4,L5).

4. Experimental Setup and Results

4.1. Experiment 1: human motion detection

We design first experiment to detect human movement. To evaluate our system, we performed several experiments on the TUM kitchen data. The algorithm is implemented in Matlab2010a program. The size of the input video image is 320×240 pixels and the sample rate is about 25 frames per second. In our experiment each cycle just takes about 120 ms time (each 4 frames).

4.2. Dataset

The TUM Kitchen Data Set has been recorded in the TUM kitchen (Tenorthet Al, 2009).The recorded (synchronized) data consists of calibrated videos, motion capture data and recordings from the sensor network, and is well-suited for evaluating approaches for motion tracking as well as for activity recognition. In particular, the following modalities are provided: Video data (25 Hz) from four static overhead cameras (Fig. 1) provided as 384x288 pixel RGB color image sequences (JPEG) or compressed video files (AVI). Additionally, full resolution (780x582 pixels) raw Bayer pattern video files are available.

4.3. Experiment 2: motion detection under hard real-time constraints

In experiment 1, we saw that the method yielded the best results, which is completely consistent with our expected results. We employed this method to solve motion detection under hard real-time constraints problem. To be able to compare the robustness of this method to solve the little and \periodic movement problem, we repeated the first experiment but with different set of data (Several video frequencies in sea surface as experimental video frequencies). In this experiment we used a video which one Jockey was riding a horse at the beach. Also the robustness to outdoor lighting changes was investigated. The results are relative ideal. Detection of moving targets in different sea surfaces is shown in Fig. 3.

Fig. 3: Human motion detection (top), motion detection under hard real-time constraints (bottom).

5. Conclusion

This paper introduces a robust and comparatively accurate method for detecting moving targets, especially for human motion detection, under hard real-time constraints. The particular advantage of the proposed method is that it can be applied to complex scenes, when does not rely on camera stabilization. The results show that it can detect moving targets correctly and be robust even for the noise of shadows and ocean wave. Experiments results show that the method detecting moving targets high efficiency and high accuracy. Also when two or more moving objects are superimposed together, it can detect them as different moving object. Second, because of our aim and with this proposed method this is easily usable for real-time detection. Third, in the condition of complex background, it can detect small targets as well as common targets. However, there are still something need to improve. We believe these problems are important issues in motion detection and more comprehensive study is required to investigate. In the future work, we are going to investigate these complications with more details and also ways of minimizing their effects should be considered in future.

References

Azad, P., Ude, A., Asfour, T. & Dillman, R. (2007) Toward an Unified Representation for Imitation of Human Motion on Humanoids, In IEEE International Conference on Robotics and Automation (ICRA). Rome, Italien, April, pp. 2558-2563.

Bandouch, J. & Beetz, M. (2009) Tracking humans interacting with the environment using efficient hierarchical sampling and layered observation models, IEEE Int. Workshop on Human-Computer Interaction.In conjunction with ICCV2009. Sept. 27 Oct. 4, pp. 2040–2047.

Bobick, A. F. & Davisj, W. (2001) The recognition of human movement using temporal templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(3): pp. 257–267.

Blostein, S & Huang, T. (1991) Detecting small, moving objects in image sequences using sequential hypothesis testing. IEEE Transactions on Signal Processing, 39(7):pp. 1611–1629.

Deutscher, J. & Reid, I. (2005) Articulated body motion capture by stochastic search. International Journal of Computer Vision (IJCV),61(2):pp. 185–205.

Hurlbert, A. & Wolf, K. (2002) The contribution of local and global cone-contrasts to colour appearance: a Retinex-like model. In B.E. Rogowitz and T.N. Pappas Human Vision and Electronic Imaging VII. Proc. of SPIE 4662, pp. 286-297.

Kulic, D., Takano, W. & Nakamura, Y. (2008) Incremental learning, clustering and hierarchy formation of whole body motion patterns using adaptive hidden markov chains, International Journal of Robotics Research, 27(7):pp. 761–784.

Krüger, V., Kragic, D., Ude, A. & Geib, C. (2007) The meaning of action:a review on action recognition and mapping. Advanced Robotics, 21(13):pp. 1473–1501.

Land, H. E. & McCann, J. J. (1971) "Lightness and Retinex Theory," J. Opt. Soc. Am. 61, pp.1-11.

Schuldt, C., Laptev, I. & Caputo, B. (2004) Recognizing human actions:a local svm approach. In 17th International Conference on Pattern Recognition (ICPR 2004), volume 3, pp. 32–36.

Tenorth, M., Bandouch, J. & Beetz, M. (2009) The TUM kitchen data set of everyday manipulation activities for motion tracking and action recognition, in IEEE Int. Workshop on Tracking Humans for the Evaluation of their Motion in Image Sequences (THEMIS) at ICCV, pp. 1089 - 1096 .

Vahidi-Shams, A. & Guran, A (2011) A Robust And Fast Motion Detection And Tracking Method For Video Surveillance Systems Accepted in The Eighth IASTED International Conference on Biomedical Engineering (Biomed 2011) February.16-18, Innsbruck, Austria.

Yun, J., Yan, D. & Ze-ping, L. (2005) Study on improved arithmetic of image segmentation, Optical technique, pp. 155-157.