MOVING OBJECTS TRACKING ALGORITHM DESIGN AND TESTING FOR MOBILE ROBOTS

Alexander Bekiarski

Technical University – Sofia, Bulgaria
Kliment Ohridski, 8
Tel.: +359 965 3300; E-mail: aabbv@tu-sofia.bg

Abstract

Mobile robot object tracking is a popular method for testing the level of intelligence of a mobile robot. The robot visual system is the means for a mobile robot to interact intelligently and effortlessly with a moving object or a human inhabited an environment. The goal here is to design and testing a mobile robot tracking algorithm capable of interfacing intelligently with an object or human in the robot environment or area of observation.

1. INTRODUCTION

Currently one of the most active application in robotic is to observing some moving objects or people in the area of the robot. [1] The proposed algorithm starts with creating a certain background model, subtracting this model from each frame of tracking sequence to find the potential regions of movements [2]. These regions are treated as possible parts of the objects being tracked. Each of these regions is matching with the other regions to define the real place of the object or human tracking from the robot.

2. MOVING OBJECTS TRACKING

The objects in the robot area of observation are usually moving persons in applications like video surveillance [3]. The tracking algorithm of moving person consists of three main parts: motion detection, human body separation, person tracking.

2.1. Motion Detection Algorithm

The first step – motion detection algorithm is presented in the Fig.1.

The current input image frame is IFr(n). One Frame Time delay block is necessary to store the previous image frame IFr(n-1) in a frame memory. Then is calculated Image Absolute Difference IAD between of current and previous frame as:

$$IAD = \left| IFr(n) - IFr(n-1) \right| \qquad (1)$$

The space and temporal person movements are in the range of low space and temporal frequencies [2], so the next step in the algorithm presented on the Fig.1 is mark as Low Pass Filter. It is designed as a sliding local MxN windows with central element i,j moving in the Image Absolute Difference frame IAD with N_x and N_y - the horizontal and vertical image size. The goal is to calculate a value of absolute difference VAD(I,j):

$$VAD(i,j) = \sum_{k=1}^{M} \sum_{l=1}^{N} IAD(i-k, j-l) \qquad (2)$$

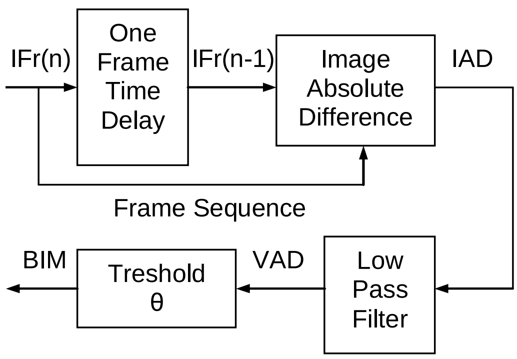

Fig.1. Algorithm for motion detection

The calculated values are compared with a chosen Threshold θ:

$$BIM(i,j) = \begin{cases} 0 \text{ } if \text{ } VAD \langle \theta \\ 1 \text{ } if \text{ } VAD \geq \theta \end{cases} \qquad (3)$$

The result is a binary image mask BIM containing values “0” for pixels belonging to the static image regions and values “1” for pixels belonging to the moving image regions.

2.2. Image Frame Pre-filtering

The algorithm on Fig.1 work well if there is not the image intercity changes caused from the fast illumination drifts, noise etc. These changes can lead to some errors of false moving object detection. To overcome these negative effects it is proposed to perform a pre-filtering of each frame IFr(n) using a filter structure presented on the Fig.2.

The goal of the proposed pre-filtration in the Fig.2 is to avoid false motion detection by fast illumination existing in each image frame IFr(n) as a composition of an illumination i and reflectance r:

$$IFr(n) = iIFr(n) + rIFr(n). \qquad (4)$$

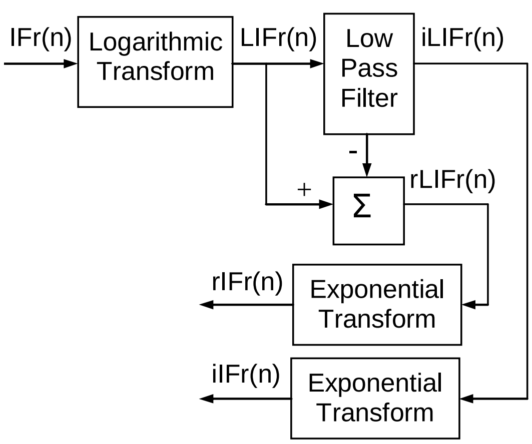

Fig. 2. Frame pre-filtering

An Logarithmic Transform to (4):

$$LIFr(n) = \log IFr(n) = \log(iIFr(n) + rIFr(n)) =$$
$$= iLIFr(n) + rLIFr(n). \qquad (5)$$

show that components $iLIFr(n)$ for illumination and $rLIFr(n)$ for reflection are mixed. They can be separated with a Low Pass Filter and a block of addition Σ, using the property of the luminance component $iLIFr(n)$ to contain only the low frequency spectral components:

$$rLIFr(n) = LIFr(n) - iLIFr(n). \qquad (6)$$

The separated illumination $iLIFr(n)$ and reflection $rLIFr(n)$ are put under inverse Exponential Transform to output them as two separate components iIFr(n) and rIFr(n).

Only the reflected part rIFr(n) carry the information for the moving objects or persons in the images. Only this part rIFr(n) is applied as input of the motion detection algorithm on the Fig.1.

2.2. Object or Human Body Separation

The binary image mask BIM, determined as the output in Fig.1. is used in the separation of moving person in area of robot observation. The binary mask in this step is processed with Freeman chain code to describe the object boundary as a set of points with their co-ordinates x(p) and y(p), combined as a complex number:

$$z(p) = x(p) + jy(p), \qquad (7)$$

for $p = 0,1,...,P$.

Each closed boundary is with perimeter P.

The discrete Fourier transform of z(p) is given by:

$$a(k) = \frac{1}{P}\sum_{n=0}^{P-1} z(p).e^{-j2\pi kp/P} \qquad (8)$$

for $k = 0,1,...,P-1$.

The complex coefficients $a(k)$ represent the Fourier description of the corresponding boundary. They can be processed to make them invariant of object position and site. The invariant properties are considered as translation, rotation and scaling. The independence of translation and rotation can be achieved by removing DC-components $a(0)$ and by using only magnitude of each $a(k)$ spectral coefficient. The scale invariance is realized by dividing all $a(k)$ by the magnitude of $a(1)$. Also it is made a subtractions of the phase $e^{j\varphi_1}$ of coefficient $a(1)$, weighted with "k". This removes the starting point variance. Finally the set of coefficients representing a corresponding boundary with removed variances are:

$$ai(k) = \frac{|a(k)|}{|a(1)|}.e^{-j\varphi_1 k} \qquad (9)$$

for $k = 2,3,...,P-1$.

There is an other way to boundary representation as the distance d(p) of the boundary points x(p), y(p) to the objects or person centre of gravity (xc, yc):

$$d(p) = \left((x(p) - x_c) + (y(p) - y_c)^2\right)^{1/2}. \qquad (10)$$

Using Fourier transform for the distance d(p) gives:

$$b(k) = \frac{1}{P}\sum_{p=0}^{P-1} d(p) e^{-j2\pi kp/P} \qquad (11)$$

for $k = 0,1,...,P-1$.

The advantage of second method is the real value of d(p) in equation (10), which gives the possibility to calculate and use only half of description b(k) from equation (11). Also the subtraction of the centre of gravity represent the position of the shape, i.e. the distance d(p) is invariant to translation. The other invariance properties are achieved in a similar way (equation (9)) to the first method:

$$bi(k) = \frac{|b(k)|}{|b(0)|} \qquad (12)$$

for $k = 0,1,...,P-1$.

For the moving objects or persons detection and tracking it is sufficiently to use only low frequency components from (9) or (12), for example first 10 coefficients in the object or person classification task.

2.3. Classification of Moving Object

For a precise separation only of moving human body from all moving objects it is performed an object classification with an appropriate feed-forward neural network, which is proposed and tested as a four layers type for the neural network.

The network consists of four layers:

- input layer with number of chosen fetures;
- two hidden layers with number of neurons chosen in the time of testing as seven neurons;
- one output layer with one neuron per class of the objects.

The activation function is chosen as a sigmoid and the training algorithm is well known back-propagation method.

3. EXPERIMENTAL RESULTS AND CONCLUSION

The proposed algorithm of moving object or human, tracking with a mobile robot is simulated and tested with some chosen images with objects like cars, persons, and landscape, shown in Fig.3.

The experiments are carried out with the described algorithm of moving objects detection and their classification with the four layer neural network. The network is trained with the set of objects type: human, cars and landscapes, shown in Fig.3, which are divided as training and testing images. The simulation gives the correct classification. Each objects, as it is shown in fig.3, is separated with a rectangle to show the correct classifications for the chosen types of moving objects human and cars. The landscape type is rejected, which is also an example of the ability of the proposed algorithm to recognize false moving objects.

Fig.3. The separated human, vehicle and rejected false moving objects

The results in the Table 1 are summarized and shown as a compartment of two methods for invariant properties (equations 9 and 12).

The great percentage of correct moving object classification is a guaranty of the efficiency of using and future improvements of the developed and tested algorithm.

Table 1

Method	Equation (9)		Equation (12)	
Target Class	human	vehicle	human	vehicle
Number of objects	777	463	773	456
Correct Classification	745	452	672	436
Miss Classification	32	11	101	20
Correct %	96	98	87	96

ACKNOWLEDGMENT

This work was supported by National Ministry of Science and Education of Bulgaria under Contract BY-I-302/2007: "Audio-video information and communi-cation system for active surveillance co-operating with a Mobile Security Robot".

References

[1] J. Fritisch, M. Kleinehagenbrock, S. Lang "Audiovisual person tracking with a mobile robot", in Proc. Int. Conf. on Intelligent Autonomous System, pp. 898-906, IOS Press, 2004.

[2] T. Aach, A. Kaup, and R. Mester. Statistical model-based change detectionin moving video. Signal Processing, 31(2):165-180, 2008.

[3] Taneva L., Development system with microcontroller MSP430F149 for education, FMNS, 2005, Blagoevgrad, Vol. 1, pp. 175 – 180.