Asian Journal of Information Technology 5 (10): 1034-1038, 2006

An HMM-Based Model for Moving Object Detection

A.R. Debilou and S. Aouragh
Department of Electronics, Faculty of Sciences and Engineering Sciences,
University of Mohamed Kheder, Biskra, Algeria

Abstract: A new probabilistic background-foreground model based on Hidden Markov Models (HHMs) is presented. The hidden states of the model enable discrimination between Foreground and Background. This method is composed of two phases. First, an ICE (Iterative Conditional Estimation) algorithm is introduced to learn the unknown HMM parameters. In the second stage, each pixel is classified with an MPM (Maximum Posterior Marginal) classification algorithm. The potential and efficiency of the method have been proven through simulations under Matlab.

Key words: Detection, hidden markov model, moving object, estimation, classification, stationary camera

INTRODUCTION

Detection of moving objects from an image sequence taken with a stationary camera is a basic task for several applications of computer vision and automatic surveillance, e.g., video monitoring systems, remote sensing, tracking and identification. It is important to achieve very high sensibility in the detection of moving objects, with the lowest possible false alarm rates. Background substraction is a method typically used to detect motion in the scene by comparing each new frame to a reference image. The simplest form of the reference image is a time averaged background one. So, a training period with the empty scene is necessary. A popular method for extracting moving objects is to maintain an adaptive model of the background and compare it to each incoming image frame. As result, a binary card is obtained, where every pixel is classified as either background or foreground. Background models used include those featuring per-pixel Kalman filters[1-4] or those featuring per-pixel Gaussian distributions[5-7]. Friedman and Russell modeled each pixel by an adaptive parametric mixture model of three Gaussian distributions[6]. While a non parametric approach is taken by Harwood[8], that increases detection sensitivity and reducing false positives.

A background model based on Hidden Markov Models (HMMs) was introduced by Blake[2] that modeled regions of background, foreground and shadow. The HMMs are a powerful tool that has shown to be of great use in signal and speech processing[3]. These models are increasingly being applied in vision domain, to problems such as image segmentation[7], handwritten symbols recognition[9], gesture interpretation and event understanding.

This study introduced a new segmentation method based on Hidden Markov Models. Moving object detection is divided on two phases, the learning phase and the segmentation phase. In the first one, model parameters are estimated with an ICE (Iterative Conditional Estimation) algorithm. Then in the second stage, each pixel is labeled as background or foreground, with the MPM (Maximum Posterior Marginal) classification algorithm.

HIDDEN MARKOV MODEL

Let I be the finite set corresponding to the N pixels of an image. We consider two random processes $X = (X_t)_{t\in T}$ and $Y = (Y_t)_{t\in T}$, where Y represents the background substraction image and X the unknown class image.

Each random variable X_t takes its values from the finite set of classes $\Omega = \{\omega_1, \omega_2,..., \omega_K\}$, whereas each Y_t is a real value. We denote realizations of X and Y by $x = (x_t)_{t\in T}$ and $y = (y_t)_{t\in T}$, respectively.

We suppose that the random variables $Y = (Y_t)_{t\in T}$ are independent conditionally on X and that the distribution of each Y_t conditional on X is equal only to its distribution conditional on X_t. Hence, all the distributions of Y conditional on X are determined by the K distributions of Y_t with respect to $X_t = \omega_i|_{1\le i\le K}$, which will be denoted f_1, f_2,..., f_K:

$$\begin{aligned} P(Y = y/X = x) &= \prod_{t\in T} P(Y_t = y_t/X_t = x_t) \\ &= \prod_{t\in T} f_{X_t}(Y_t = y_t) \end{aligned} \qquad (1)$$

Corresponding Author: A.R. Debilou, Department of Electronics, Faculty of Sciences and Engineering Sciences, University of Mohamed Kheder, Biskra, Algeria

We consider that the prior distribution of the process presenting the states of the pixels is a Markov Chain. We generally refer to it as Hidden (Hidden Markov model) because X is not directly observable. The classification or detection problem consists in estimating the realization x of the hidden states X from the only required information y; the realization of the observable process.

Before any HHM based processing on images, it is important to order pixels into a vector. A 2D image is then transformed into a 1D chain. Deferent scans can be utilized. We can adopt the line by line or column by column scan. Another alternative is to use a Hilbert-Peano scan[7].

Hidden markov chains: According to the definition, X is a Markov chain if

$$P(X_t / X_{t-1}, X_{t-2} ..., X_1) = P(X_t / X_{t-1}) \quad \text{for } 1 \le t \le T, \tag{2}$$

with T denote length of the markov chain[3].

The distribution of X will consequently be determined by the initial distribution of X_t denoted by π and the set of transition matrix elements denoted by $a_{ij} = P(X_{t+1} = \omega_j / X_t = \omega_i)$ each one[3,4,7]. In the following, we assume that the transition probabilities written as:

$$c_{ij} = P(X_t = \omega_i, X_{t+1} = \omega_j) \tag{3}$$

Our independent of t. The initial distribution then becomes:

$$\pi_i = P(X_1 = \omega_i) = \sum_{1 \le j \le K} c_{ij} \tag{4}$$

and the transition elements of matrix A are given by:

$$a_{ij} = P(X_t = \omega_j / X_{t-1} = \omega_i) = \frac{c_{ij}}{\sum_{1 \le j \le K} c_{ij}} \tag{5}$$

with

$$A = [a_{ij}]_{1 \le i, j \le K}$$

Hence, the a priori distribution of X is entirely determined by the K^2 parameters $(c_{ij})_{1 \le i \le K, 1 \le j \le K}$. So, we can write:

$$P(X) = P(X_1 = \omega_{X_1}, X_2 = \omega_{X_2}, ..., X_T = \omega_{X_T}) = \pi_{X_1} \cdot a_{X_1 X_2} \cdot a_{X_2 X_3} \cdot ... \cdot a_{X_{T-1} X_T} \tag{6}$$

The so-called forward and backward probabilities

$$\alpha_t(i) = P(X_t = \omega_i, Y_1 = y_1, Y_2 = y_2, ..., Y_t = y_t) \tag{7}$$

and

$$\beta_t(i) = P(Y_{t+1} = y_{t+1}, Y_{t+2} = y_{t+2}, ..., Y_T = y_T / X_t = \omega_i) \tag{8}$$

Can be calculated recursively? Unfortunately, the original forward-backward recursions derived from (12) and (13) are subject to serious numerical problems[7]. Devijver *et al* have proposed to replace the joint probabilities by a posteriori probabilities:

$$\alpha_t(i) \approx P(X_t = \omega_i / Y_1 = y_1, ..., Y_t = y_t) \tag{9}$$

$$\beta_t(i) \approx \frac{P(Y_{t+1} = y_{t+1}, ..., Y_T = y_T / X_t = \omega_i)}{P(Y_{t+1} = y_{t+1}, ..., Y_T = y_T / Y_1 = y_1, ..., Y_t = y_t)} \tag{10}$$

based on theses approximations, both forward and backward probabilities are calculated recursively as following:

Forward algorithm

Initialisation:

$$\alpha_1(i) = \frac{\pi_i \cdot f_i(y_1)}{\sum_{1 \le j \le K} \pi_j \cdot f_j(y_1)} \quad \text{for: } 1 \le i \le K$$

Induction:

$$\alpha_t(i) = \frac{f_i(y_t) \cdot \sum_{1 \le j \le K} \alpha_{t-1}(j) \cdot a_{ij}}{\sum_{1 \le l \le K} f_l(y_t) \cdot \sum_{1 \le j \le K} \alpha_{t-1}(j) \cdot a_{jl}} \quad \text{for: } 1 \le i \le K, \quad t = 2, ..., T \tag{11}$$

Backward algorithm

Initialisation:

$$\beta_T(i) = 1 \quad \text{for: } 1 \le i \le K \tag{12}$$

Induction:

$$\beta_t(i) = \frac{\sum_{1 \le j \le K} a_{ij} \cdot f_j(y_{t+1}) \cdot \beta_{t+1}(j)}{\sum_{1 \le l \le K} f_l(y_{t+1}) \cdot \sum_{1 \le j \le K} \alpha_t(j) \cdot a_{jl}} \quad \text{for: } 1 \le i \le K, \quad t = T-1, ..., 1 \tag{13}$$

The joint probability of two subsequent classes given all the observations

$$\psi_t(i,j) = P\left(X_t = \omega_i, X_{t+1} = \omega_j / Y = y\right) \quad (14)$$

can be written as a function of the forward-backward probabilities?

$$\psi_t(i,j) = \frac{\alpha_t(i)\cdot a_{ij}\cdot f_j(y_{t+1})\cdot\beta_{t+1}(j)}{\sum_{1\le m\le K}\sum_{1\le l\le K}\alpha_t(m)\cdot a_{ml}\cdot f_l(y_{t+1})\cdot\beta_{t+1}(l)} \quad (15)$$

The marginal a posteriori probability, i.e., the probability of having class ω_i in element number t given all the observations Y, can also be expressed in terms of the forward-backward probabilities:

$$\xi_t(i) = P\left(X_t = \omega_i / Y = y\right) = \frac{\alpha_t(i)\cdot\beta_t(i)}{\sum_{1\le l\le K}\alpha_t(l)\cdot\beta_t(l)} \quad (16)$$

Thesis' probabilities are needed in MPM algorithm computing.

It can be shown that the a posteriori distribution of X, i.e., P(X = x/Y = y), is that of a non stationary Markov chain, with transition matrix

$$t_{ij}^t = P\left(X_{t+1} = \omega_j / X_t = \omega_i, Y = y\right) \quad (17)$$

$$t_{ij}^t = \frac{a_{ij}\cdot f_j(y_{t+1})\cdot\beta_{t+1}(j)}{\sum_{1\le l\le K} a_{il} f_l(y_{t+1})\cdot\beta_{t+1}(l)} \quad (18)$$

We can therefore simulate a posteriori realisations of X directly, using Eq.18 recursively. The class of the first pixel $X_1 = \omega_i$ is drawn randomly according to the marginal a posteriori distribution $\xi_1(i)$ (16). Subsequently, for each new pixel, the transition probability's t_{ij}^t are computed Eq. 18, the class of the precedent pixel ω_i being fixed and the class ω_j of the next pixel is obtained by random sampling according to this distribution.

HMM PARAMETER ESTIMATION

We consider the set of model parameters that includes initial probabilities $(\pi_i)_{1\le i\le K}$, transition matrix $A = \lfloor a_{ij}\rfloor_{1\le i,j\le K}$ and observation distributions $(f_i)_{1\le i\le K}$. Generally, the model parameters are unknown. They must be estimated from the observation Y = y. We do not know the characteristics of the classes and we also do not know witch pixels are representative for each class. We suppose that each distribution f_i is Gaussian, with mean value μ_i and σ^2_i variance. Usually, the Baum-Welch algorithm is adopted, in this case. But in our approach, we have introduced the ICE (Iterative Conditional Estimation) algorithm witch is specially applied in image segmentation.

Algorithm: The ICE algorithm for an HMM, uses the forward and backward probabilities too. Each ICE iteration q is based on one forward-backward computation and includes the following steps:

- For every element t in the chain and for every possible class ω_i, we compute
- the forward probability's $\alpha_t^{(q)}(i)$ (11) based on $\theta^{(q-1)}$.
- the backward probabilities $\beta_t^{(q)}(i)$ (12) based on $\theta^{(q-1)}$.
- the marginal a posteriori probabilities $\xi_t^{(q)}(i)$ (15). This allows us to compute
- the new joint conditional probabilities $\Psi_t^{(q)}(i,j)$ (14),
- the new elements of the stationary transition matrix $a_{ij}^{(q)}$

$$a_{ij}^{(q)} = \frac{\sum_{1\le t\le T-1}\psi_t^{(q)}(i,j)}{\sum_{1\le t\le T-1}\xi_t^{(q)}(i)} \quad (19)$$

- and the new initial probabilities $\pi_i^{(q)}$

$$\pi_i^{(q)} = \sum_{1\le t\le T-1}\xi_t^{(q)}(i) \quad (20)$$

We compute a series of a posteriori realisations based on $\alpha_t^{(q)}(i)$ and $f_i^{(q-1)}$. For each realisation with index m ($1\le m\le M$, M total number of simulated realisations at iteration q), we estimate the class parameters $\theta^{(q),m}$, witch are averaged to obtain $\theta^{(q)}$.

$$\hat{\pi}_i^{(q)} = \frac{1}{M}\sum_{1\le m\le M}\hat{\pi}_i^{(q),m} \quad (21)$$

$$\hat{a}_{ij}^{(q)} = \frac{1}{M}\sum_{1\le m\le M}\hat{a}_{ij}^{(q),m} \quad (22)$$

$$\hat{\mu}_i^{(q)} = \frac{1}{M}\sum_{1\le m\le M}\hat{\mu}_i^{(q),m} \quad (23)$$

$$\left(\hat{\sigma}_i^{(q)}\right)^2 = \frac{1}{M}\sum_{1\le m\le M}\left(\hat{\sigma}_i^{(q),m}\right)^2 \quad (24)$$

STATE ESTIMATION

At this stage, the parameters of the model are known. Our objective is to determine the class of each pixel of the courant image, in a Bayesian framework. In other terms, choosing the realisation X = x which best explains the observation Y = y, according to a certain criteria. Given the observation sequence, several criteria for selecting an optimal state sequence may be used. One can adopt the MAP based estimator which maximises the global a posteriori probability P(X/Y). In the study of HMM, this estimator is called the Viterbi algorithm[3]. Posterior marginal distribution MPM maximizing the distribution $P(X_t/Y)$ for each pixel can also envisaged. The latest one is chosen in our study as it minimizes the rate of false classified pixels and do not use the total path.

Algorithm: The MPM algorithm for an HMM can be calculated directly for hidden Markov chains. It bases on forward-backward computation:

For every element t in the chain and for every possible class ω_i, we compute:

- the forward probabilities $\alpha_t(i)$ (11).
- the backward probabilities $\beta_t(i)$ (12).
- the marginal a posteriori probabilities $\xi_t(i)$ (15).

Then, each pixel t is attributed to the class that maximises $\xi_t(i)$:

$$x_t = \arg\max_{\omega_i} P\left(X_t = \omega_i / Y = y\right)$$

RESULTS

We have considered an ergodic HMM in which every can be reached in a single step from every other state. We use the gray-scale intensities as observations. We have assumed that the observation densities are Gaussian. First, image pixels are scanned with a Hilbert-piano path. Model parameters are estimated and updated by the ICE algorithm. We limit the number of ICE iterations to one and we compute only one a posteriori realisation per ICE iteration. Finally, the MPM-based state estimation process classifies each pixel as Background or moving object. The algorithm has been tested on a wide range of image sequences. Fig. 1 are shows detection results obtained in four deferent sequences. A number of two frames is sufficient to get a high detection quality. It is important to note that all the results shows the detection process without any a priori information on targets and initial models are chosen arbitrary.

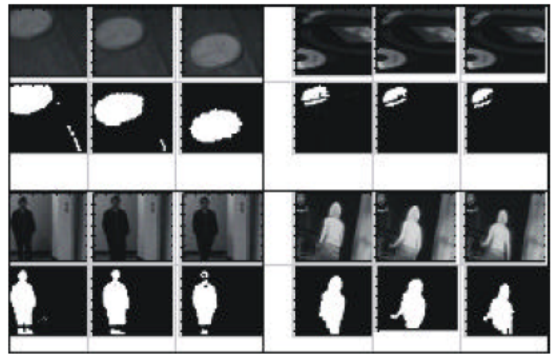

Fig. 1: Detection of moving object with ICE

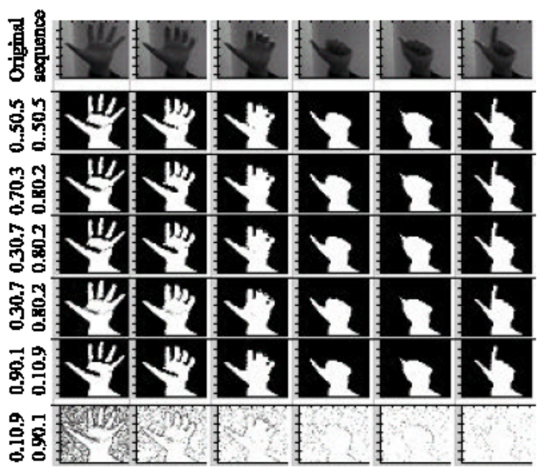

Fig. 2: Initial transition matrix effects on detection results

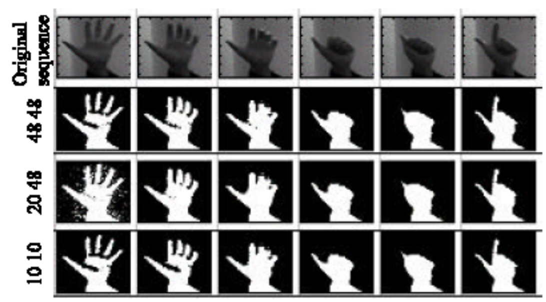

Fig. 3: Initial mean effects on detection results

To examine the performances more carefully, we have added some secondary analysis and comparisons. We have tested the effects of initial parameter setting as it is essential to guarantee an appropriate initial model. This point is critical in estimation-based processing. Fig. 2 are shows the effects of initial transition matrix. Those belong to Gaussian (mean and variance) distributions are illustrated by Fig. 3 and 4.

We have also compared our method to the EM-MPM algorithm and the EM-Viterbi one which is usually associated with HMMs, Fig. 5.

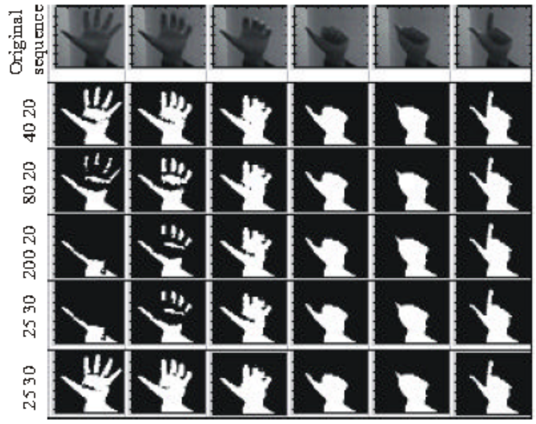

Fig. 4: Initial variance effects on detection results

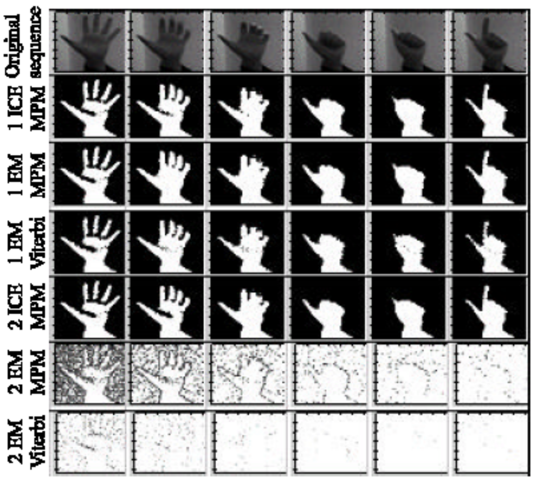

Fig. 5: Comparison of ICE+MPM algorithm vs EM+MPM algorithm vs EM+Viterbi algorithm with initial models: 1: $\pi = 0.5\ 0.5$, $A = 0.5\ 0.5, 0.5\ 0.5$, $\sigma = 48\ 48$, $\mu = 40\ 20$. 2: $\pi = 0.3\ 0.7$, $A = 0.1\ 0.9, 0.9\ 0.1$, $\sigma = 10\ 10$, $\mu = 40\ 20$

Our method detect moving object with efficiency. Unlike other approaches, it presents no sensibility to initial parameters unless they are extreme.

CONCLUSION

A new probabilistic study for extracting moving objects from stationary background has been described. It is an HMM-based segmentation method which is able to model Foreground as well Background regions. A considerable advantage is that no appropriate initial model is needed, unlike other approaches. It is also no longer necessary to provide specific data for training. Parameters are initially estimated from the first frame and then updated after every new frame. Another advantage is that the model can easily be extended. By adding a third state, it can detect shadows too.

To improve robustness, we can consider colour or other pixel characteristics instead of intensities. We can also take small blocs with equal sizes ($k \times k$ pixels) and filter them, instead of using gray-scale intensities directly in order to reduce noise. In specific applications such as traffic monitoring, is better to do some analysis in order to define the topology of the model and the form of the observation densities and then getting a suitable model.

Future study involves the implementation of the algorithm on a DSP kit (C6000 IDK from Texas Instruments), to get real-time data analysis.

REFERENCES

1. Karmann, K.P. and A. Von Brandt, 1990. Moving object recognition using an adaptative background memory, Time varying Image, Processing and Moving Object Recognition.
2. Rittscher, J., J. Katto, S. Joya and A. Blake, 2000. A probabilistic background model for tracking in European Conference on Computer Vision.
3. Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition, Proc. IEEE, 77: 257-286.
4. Ridder, C., O. Munkelt and H. Kirchner, 1995. Adaptive background estimation and foreground detection using Kalman filtering, Proc. ICAM.
5. Stauffer, C. and W. Grimson 1999. Adaptative background mixture model for", IEEE Conf. On Computer Vision and Pattern Recognition CVPR.
6. Friedman, N. and S. Russell, 1997. Image segmentation in Video Sequences: a Probabilistic Approach, in the Thirteen Conf on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers. Inc. San Francisco.
7. Wojciech Pieczynski and Btissam Benmiloud, 1995. Parameters estimation of hidden markov chains and image segmentation, Signal Analysis, (in french).
8. Ahmed Elgammal, David Harwood and Larry Davis, 2000. Non-parametric Model for Background Subtraction 6th European Conference on Computer Vision, Dublin.
9. Marisa Emika Morita, 2003. Automatic Recognition Of Handwritten Dates On Brazilian Bank Cheques, Doctoral thesis, Québec University, Montreal.