International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 14 (2015) pp 33981-33986

Parameter Based Change Detection In Human Crowd Movement Using Hidden Markov Model

Nitish Ojha
Department of Information Technology
Indian Institute Of Information Technology Allahabad, India
nitishkumarojha@gmail.com

Abhishek Vaish

Department of Information Technology
Indian Institute Of Information Technology Allahabad, India
abhishek@iiita.ac.in

Abstract
The focus of security research in computer vision is getting shifted towards in intelligent automated videos surveillance in recent years. The process of anomaly detection involves object classification, detection, tracking and recognition. In classification process we use local features of moving object and group them on basis of their attributes. In this paper, we have used time based dynamic warping with Hidden Markov Model for clustering of anomalous human movement through images collected through optical flow based segmentation.

Keywords - Crowd Movement, HMM, Video Surveillance, Optical Flow.

Introduction

Action recognition is a challenging area of research which is attracting many researchers of computer vision and intelligent systems. Action recognition is getting more and more attention when it is being discussed from crowd movement perspective rather than individual human action movement.

Enormous number of researchers have been carried-out in recent years and very relevant results have been observed yet this field is still open consider different challenges. When we consider crowd behavior in human action recognition research the number of research outcome being reported is relatively small. HMM is highly used algorithm to approach this type of streamlined research area. Although HMM is a mainstream method for dynamic event outcome prediction yet less research are reported specially focused on crowd behavior prediction using Hidden Markov Model based on optical flow based analysis.

Related Work

Due to arising situation of new challenges of surveillance areas, high end security cameras are the need of the time which facilitate the investigation process of any crisis by presenting better form of evidences in digital format.

Security companies are putting their better efforts to make these surveillance cameras more advanced using different set of technology so that scenario based recognition can be there yet more research is required for appropriate understanding of event happening in a scene. Several studies have been done in last decade based on classical methods of background based activity recognition, frame differencing, background subtraction, spatio-temporal analysis etc.

In this paper, we present a model which detects multiple movement of crowd in a specified area. The system detects the different pattern of movement of crowd in a directional fashion.

Most of the work done in area of computer vision focusing action recognition using hidden markov model is somewhere limited to individual entity based approach where the state of the movement is being recognized and predicted based on atomic movement of individual human.[1]

In this study researchers took three subjects to predict six different swings of tennis balls, but still there was a lot of scope in this considering crowd domain.

Solving this issue with HMM methodology, several challenges lies in coping with the issue of shadows, noise, distortion, occlusions and number of good scientific studies have been performed but still it remains a good area attraction due to its thirst.

Irani et al. proposed a model for human action recognition in video based on hidden markov model without using segmentation, yet it was very fruitful model to detect the discontinuity based action recognition but it also poses the space complexity high.[2]

Silhouettes based action recognition with help of motion history images were also used by Martinez et al in 2D images which was different from others as they have used Kohonen SOM technique in 2D to learn. [3] [4].

In summary we can say that the result of accurate recognition of human action in a streamlined movement of crowd depends on the performance of the technique which is being used.

Some classical techniques have been proposed in last five years such as Zhong et al. developed a new technique based on clustering occurrence of information after segmentation of complete video. [5]

Some changes were also observed in these classical approaches such as distance measurement based techniques have been proposed for clustering of information along with some un-supervised non-parametric model.[6] Hautamaki et al.

In last 5 years, few researchers proposed hierarchical based clustering of information using HMM by making a model as outcome of learning the likelihood pattern of trajectory and combining new trajectory with developed model by comparing the immediate trajectory.

Some research hers carried away this idea and worked on the concept of grouping of similar motion movement and proposed a model based on HMM clustering side by side they also developed a new approach as extension of HMM to detect the deviation based outlier.

Methodology

Model Choice - When we consider the first step towards selection of model choice the primary objective is "Optimization".

The selected model should be enough mature that the class wise randomness in crowd movement can be interpreted easily. A generalized action recognition algorithm based model will detect and differentiate the variation in the movement and differentiate it in respect to other classes consisting the same type of action with slight variation. [7]

This model is based on the images collected through video surveillance sequences where it will select the movement based sequence with particular time stamp and gesture recognition is performed on the pixel value of the image in terms of feature vector.

Considering the scenario where variation in the movement generate increased amount of data which results as a good objective for recognition using efficient learning and classification algorithm and for this purpose hidden Markov model is one of main stream model choice which is used in these of problems.

(B) Feature Selection: While selecting the feature to represent the crowd-movement is posture based atomic level variation and the most important factor is representation of dependencies and independencies among all variable such as geometric features, parametric features, silhouettes based features etc. Among all available methodologies, Optical flow methodology efficiently represents velocity distribution over pixel components as it extract the motion as a distribution of velocity over an image. [8]

Optical Flow based Motion Extraction: Here we have used histogram based optical flow approach to compute the motion descriptor frame by frame.

The fundamental problem lying with optical flow is relative motion as optical flow is being used for motion vector extraction based depth perception when object to be surveillance are moving at uniform velocity and camera is static but when camera is also moving including the situation that en-surveilled objects are also moving at different level of speed within one frame i.e. in a fame when a person is running, the movement at upper part of the body is quietly different with movement of lower part of the body (legs). [9]

This issue considering under over-fitting was solved by decreasing the frame sampling rate to low level (in our study, it was 0.3 frames per second) so that more meaningful sequences can be grabbed for proper learning and classification. [10]

Assumptions for Optical flow-

(i) Camera Movement: In our study camera was static so that the optical flow obtained through it will be only be outcome of moving objects in-front of camera.

(ii) Uniform Movement: The objects were moving at similar speed with a specified range.

(iii) Occlusions: In the complete study, occlusions were relatively small and mathematically negligible.

(iv) Context: The movement was in well-versed format of context -aware system [11], all the three entities of the systems were as a part -

(A) Places: Rooms, Buildings, Lanes etc

(B) People: Individual, Groups etc

(C) Peripherals: Objects, Things, Components etc

In this research work, the crowd movement has inherent temporality as being a time-variant process which unfolds with time, , the state of movement at time "t" is directly influenced by a state at t-1 or it can be said that the prediction of future instance of motion depends solely on its present state [12] and in terms of mathematical notions, it is well known as "Markov Process" which states that in a series of action (Movements) the next state depends conditionally on the present state of the system and its future and past are independent.

The crowd movement process similarly follows the memory less pattern which can be stated as -

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with a filtration $(\mathcal{F}_t,\ t \in T)$ for some (totally ordered) index set T and let $(S, \mathcal{S})$ be a metric space. An S-valued stochastic process $X = (X_t,\ t \in T)$ adapted to the filtration is said to possess the Markov property with respect to the $\{\mathcal{F}_t\}$ if, for each $A \in \mathcal{S}$ and each $s, t \in T$ with s < t, $\mathbb{P}(X_t \in A|\mathcal{F}_s) = \mathbb{P}(X_t \in A|X_s).$

Here transitional probabilities are involved and events have associated transitional probabilities so they form a complete Markov Network and in nature these action based events are totally causal. [13]

In later part of the research, Hidden Markov process is formed while modeling the dynamic process of human action where there is strong mutual connection between different states (as in each class of training data of video

set , we select one or two parameter of HMM model to describe the training sequence of video in classes so finally we get 10 models at the time of testing (recognizing) if ideally we compute, there will be 10 likelihood scores and the model with highest score will result the more likelihood sequence), here we have used HMM in a unsupervised manner [14] which present next sequence as predicted output in similar footprint of clustering. The reason behind using HMM as unsupervised algorithm model, lies in the beauty of HMM which results clustered data instance homogeneously on time axis. Similar type of studies have been performed "Allain et al." [15]

Description of HMM -
HMM is characterized by the following parameters -

(1) N: The number of states in HMM (the number of hidden states).we denote N states as $S = \{S_1, S_2, S_3, ...S_N\}$ and hidden state at time t as $q_t \in S = \{ S_1, S_2, S_3, ...S_T \}$

(2) The number of observation symbol in the sequence. W denote the observation symbol sequences as $O = \{O_1, O_2, O_3, ...O_N\}$

(3) A: State transition Matrix $A = (a_{ij})_{*NXN}$ where $a_{ij} = p(q_{t+1}=S_J|q_t=S_i)$ $(1,\leq i,\ j\leq N)$ a_{ij} is the probability of reaching state S_j at time t+1 from state S_i at time t.

(4) B: Observation symbol probability distribution, B= $\{b_i\ (O_t)\}$, where $b_i\ (O_t) = p(O_t|S_i)$ $1<i\leq N$, where $b_i\ (O_t)$ is probability of generating observation symbol O_t from state S_i at time t.

(5) π: The initial state distribution $\pi = \{ \pi_1, \pi_2, ... \pi_n\}$

where π_i is the probability of initial state S_i. We denote HMM as $\lambda = \{A.B, \pi\}$ using the above parameter, where A stands for Markov chain and B for relationship between states.

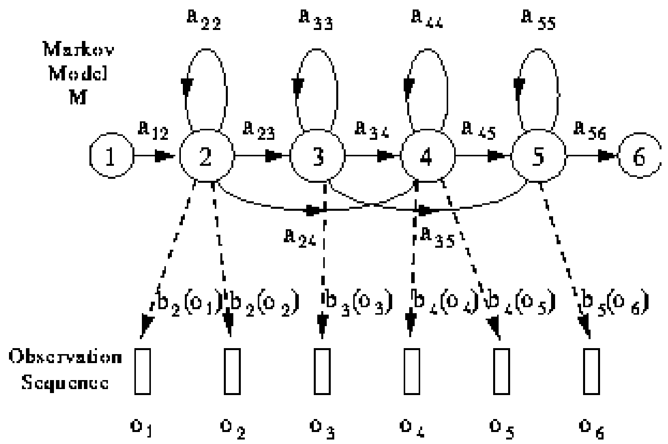

Fig. 1 Graphical Modeling of HMM having both hidden n visible observations

Experimental Process

Recognition: We followed HMM and observations were $O=\{O_1, O_2, O_3, O_T\}$ and $\lambda_1, \lambda_2, \lambda_3$, are the training samples for different action set so likelihood of the parameter of HMM-Backward-Forward algorithm -
$X(O|\lambda_1)$, $X(O|\lambda_2)$, ... $X(O|\lambda_3)$ and the best recognition based action will be the one which is having Maximum-Likelihood condition - Action Number = $\operatorname{argmax}(X|\lambda^h)$)
In this study, while training the algorithm, mix feature based samples were used in HMM and if we choose very small initial level of parameter for our algorithm the issue of localization arises at absorbing state of HMM [16], so we have used K-means algorithm for parameter updating between data and mean along with Chernoff bound.

$$\rho = \int \sqrt{(X|w_1)\ (X|w_2)}\ dx$$

as because the true expected error can be bounded well by Chernoff and computationally simple manner by Bhattacharya bound when test pattern has missing or distorted features.
where $P(error) \leq \sqrt{P(w1)\ P(w2)} \int \sqrt{P(X|w_1)P\ (X|w_2)}\ dx$

$$= \sqrt{P(w_1)\ P\ (w_2)}\ e^{-k(1/2)}$$

where $k(1/2)= 1/8\ (\mu_2- \mu_1)^t\ [\Sigma_1+ \Sigma_2/2]^{-1}\ (\mu_2- \mu_1) + 1/2\ \ln |\Sigma_1+ \Sigma_2/2| / \sqrt{|\Sigma_1||\Sigma_2|}$.

The video data set used in this study was "AGORA" data set. [17]. This dataset contains different action classes -

(i) Pedestrian are walking in a free space.
(ii) Pedestrian are walking, an obstacle is in the middle.
(iii) Pedestrian are walking in a scene with obstacle.
(iv) Escape situation, Pedestrian are crossing a door.
(v) Dispersion of crowd from a given point.
(vi) Four group of people are meeting in the middle of the scene.

Here we have first segment in which we first calculate the dominant flow in the moving crowd.[18]

Fast Algorithm For Detection Of Strongest Flow -

(i) Perform segmentation on image to break into finite set of cells.
(ii) Determine the connecting cells among all available cells.
(iii) Find out the homogenous directional flow and heterogeneous directional flow.
(iv) Compute and link all the flows in similar direction.
(v) If no similar motion flow present in the same cell then choose the closest cell having same flow from other cell and so on.
(vi) Symbolically present the flow in form of vectors.
(vii) Flow having the highest width should be computed.

Fast Algorithm for Flow based Clustering -

(i) From all available list of flows, find out the widest flow in the scene, remove clutters, dummy shadow pixels and extract the feature vector for template at certain time t.
(ii) Apply morphological operation to remove noise.
(iii) Apply vector quantization method.
(iv) Convert these features into symbols using non-metric technique.

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 10, Number 14 (2015) pp 33981-33986

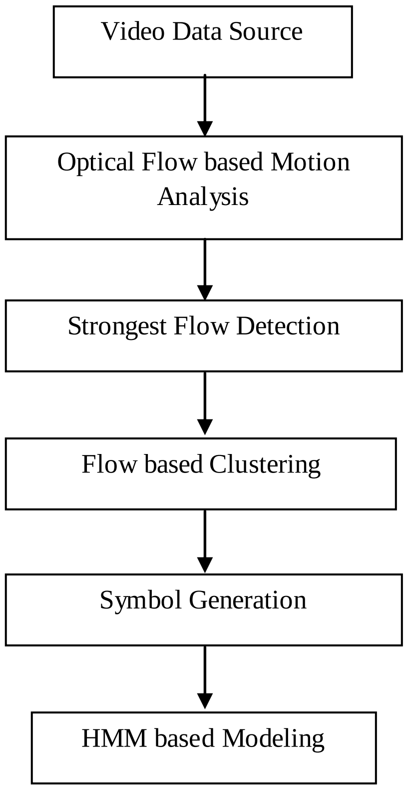

Fig. 2 Framework based representation of experimental process

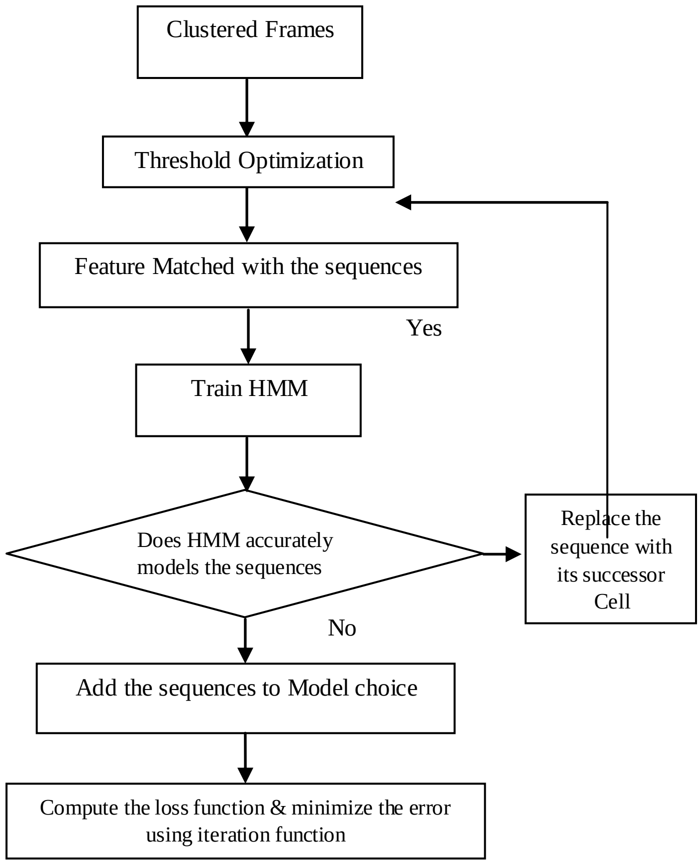

Fig. 3 Process flow of Recognition process using HMM

In OpenCV :

```
for (int h = θ; h < cluster container (); h++)
{
crpoint pt1, pt2
CvMat box[100]
pt1.x = (int) points -> data.f1 [h*2];
pt2.x = (int) points -> data.f1 [h*2+1];
CV circle (frame1, pt1, 2θ, CV_RGB, 4);
```

And for global orientation detection -

```
double calcglobalorientation (Input Array orientation, InputArray Mask, InputArray MHI, doubleduration);
Void
Opticalflow:calc(InputArray10, InputArray I1, InputOutputArray flow);
```

Process Flow For HMM based Anomaly Detection -

Result And Discussion

Out of all video sequences based action present in the data set, we selected 6 activities and segmented those into different frames at 0.3 second per frame.

We obtained 300 motion segments out of all and we manually labeled each of activity into different categories.

We performed 2 set of experiments, in first set of experiments, we applied our proposed algorithm of optical flow generation along with clustering operation for first 60 motion-sequences video, and in second set of experiment, we applied the HMM model for learning and classification.

As a classical model approach, it was a plain hypothesized concept to use HMM in such type of network but contrary to our belief, we obtained better results from our 1200 sequences per cluster.

Conclusion

In this paper we introduced a novel method of crowd movement analysis to differentiate normal movement from anomalous movement using amalgamation of two algorithm based on optical flow and HMM based cluster modeling.

We compared our result with other methodologies i.e. Ahmad et al (2012). [19], Ravi et al. (2005)[20], Du-Tran et al (2008)[21], and we found that our selected concept performed better in terms of time complexity and computational load in comparison to other models.[22]

The results also revealed that better recognition rates can be achieved if choose certain better parameters such as over segmentation may be handled in better way using extrapolation and subtraction (ES) rather than treating impurity in different class, secondly this methodology can be improved if we use grid based segmentation while calculating the strongest flow.

Certain limitation were also there such as labeling in crowd-movement poses the risk of high-computational load if it is case of high level of variability such as dispersion.

TABLE 1. Category wise results representation obtained through clusters

Activity		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
	Walking in a free space	93	58	16	20	61	0
	Obstacle is in the middle	0	30	48	78	20	8
	Walking in a scene with obstacle	0	10	18	0	0	63

TABLE: 2 Recognition results comparison among different features based HMM approaches.

Feature selected for Recognition	Spatio-Temporal Feature based	Motion History Images based State Model	Silhouettes based	Mixed Feature based Model	Optical Flow Grouping based
Accuracy	90.00%	88.00 %	90.00%	100	95.00%

References

[1] J. Yamato and A.F. Bohnick, Learning Visual Behavior for Gesture Analysis, to appear in Proc. of IEEE Symposium - Florida 1992.

[2] Schechman, E. Irani, Space-time behavior based correlation, IEEE Transction on Pattern Anaysis & Machine Inteligence-29(2009)2045-2056.

[3] Martinez Contire ras F, Orrite Vrunela, C Hero-Jaraba, Recogniio human action using silhouettes based.

[4] D. Ramnan, D.A. Farsynth, Tracking people by learning their appearance, IEEE Transaction on Pattern Recognition and Machines, Intelligent 2007-6581.

[5] Zhang, Y., Hu, S. and Martin, R. R. 2008. Shrink ability maps for content-aware video resizing. Computer Graphics Forum 27(7), pp. 1797-1804.

[6] Ouivirach, K.; Dailey, M.N., "Clustering human behaviors with dynamic time warping and hidden Markov models for a video surveillance system," Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010 International Conference on , vol., no., pp.884,888, 19-21 May 2010

[7] Hidden Markov Model for Modeling and Recognizing Gesture Under Variation - Andrew Wilson and Bobick.

[8] S. Baker, S. Roth, D. Scharstein, M.J. Black, J.P. Lewis, and R. Szeliski. A database and evaluation methodology for optical flow. Technical report, Microsoft Research, 2007. I. Lieberwerth. Presenting two time slice based optical flow algorithms. Internal Report:Vicar Vision, 2007.

[9] Li X, HMM based action recognition using oriented histograms of optical flow fields, Electronic Letters.

[10] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” Proceedings of the IEEE , pp. 257–286, 1989.

[11] A Survey on Context aware systems: Matthias Baldauf & Scharam Dustdar - International Journal of Ad. Hoc and Ubiquitous Computing. 2007.

[12] www. Britannica.com/EBChecked/365797/markov-process.

[13] B. N. Subudhi, P. K. Nanda & A. Ghosh, “A change information based fast algorithm for moving Object Detection and Tracking,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 7, pp. 993-1004, 2011.

[14] Ofli, F.; Erzin, E.; Yemez, Y.; Tekalp, A.M.; Erdem, C.E.; Erdem, A.T.; Abaci, T.; Ozkan, M.K., "Unsupervised dance figure analysis from video for dancing Avatar animation," Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on , vol., no., pp.1484,1487, 12-15 Oct. 2008.

[15] Pierre Allain, Nicolas Courty, Thomas Corpetti. AGORASET: a dataset for crowd video analysis. 1st ICPR International Workshop on Pattern Recognition and Crowd Analysis, Nov 2012, Tsukuba, Japan. pp.1-6. <hal-00904216>.

[16] Ouivirach, K.; Dailey, M.N., "Clustering human behaviors with dynamic time warping and hidden Markov models for a video surveillance system," Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), 2010 International Conference on , vol., no., pp.884,888, 19-21 May 2010.

[17] Nicolas Courty, Pierre Allain, Clement Creusot, Thomas Corpetti, Using the Agoraset dataset: Assessing for the quality of crowd video analysis

methods, Pattern Recognition Letters, Volume 44, Pages 161-170, ISSN 0167-8655, July 2014.

[18] Ozturk, O.; Yamasaki, T.; Aizawa, K., "Detecting Dominant Motion Flows in Unstructured/Structured Crowd Scenes," Pattern Recognition (ICPR), 2010 20th International Conference on , vol., no., pp.3533,3536, 23-26 Aug. 2010.

[19] Faizan Ahmad, Aaima Najam and Zeeshan Ahmed, Image-based Face Detection and Recognition:“State of the Art”, IJCSI International Journal of Computer Science Issues , Vol. 9, Issue 6, No 1. 2012.

[20] Ravi, Nishkam, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman. "Activity recognition from accelerometer data." In AAAI, vol. 5, pp. 1541-1546. 2005

[21] Du Tran, Junsong Yuan, and David Forsyth, Video Event Detection: from Subvolume Localization to Spatio-Temporal Path Search, IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2014.

[22] Tran, Du, and Alexander Sorokin. "Human activity recognition with metric learning." In Computer Vision–ECCV 2008, pp. 548-561. Springer Berlin Heidelberg, 2008.