Human Motion Analysis Using Hidden Markov Model

Uttam Kumar Kar[1], Md. Shafiul Alam Chowdhury[2]

[1, 2]Uttara University, Department of Computer Science and Engineering,
House-4, 5 & 6, Road-15, Sector-06, Uttara Model Town, Dhaka-1230, Bangladesh

Abstract: *The Hidden Markov Model (HMM) is a popular statistical tool for modeling a wide range of time series data. In this paper, we have used a feature based bottom up approach with HMMs that is characterized by its learning capability and time-scale invariability. To apply HMM to our aim, one set of time-sequential image is transformed into an image feature vector sequence, and the sequence is converted into a symbol sequence by vector quantization. In learning human action categories, the parameters of the HMMs, one per category were optimized so as to best describe the training sequences from the category. To recognize an observed sequence, the HMMs best match sequence is chosen. The reorganization rate can be improved by increasing the number of people used to generate the training data.*

Keywords: Hidden Markov Model, Feature Vector, Binary Image, Symbol Sequence, Weizmann database

1. Introduction

Human motion analysis has many applications in many areas, such as computer vision field, modeling. The main scientific challenges in human motion analysis are to detect, track and identify people and to recognize the human activity from observations. HMM is very rich mathematical structure for this application. Our purpose is to recognize human action from time-sequential images, not obtaining geometric representations of human bodies. Since we focus on recognition, we avoid reconstruction because the representation obtained by geometric reconstruction is not essential. The model parameters are obtained from time sequence of image. Our objective is to apply HMM model for Human Motion Analysis and lay down a comparison between HMM results and actuality.

1.1 Motivation

The world is undergoing its tremendous and enormous fascinations day-after-day. Initially the motivation is to introduce with Hidden Markov Model (HMM) and then to recognize the human action by the help of HMM. Human action recognition and synthesis are of major interest in a variety of disciplines: sport analysis, dancer training and choreography, scientist simulation, 3D animation, medical rehabilitation, virtual reality and entertainment. There are many processes for human action recognition. But we use HMM because by using HMM process the way to recognized human action from time-sequential images is more conformable and agreeable than any other process. So our motivation is to recognize human action from time-sequential images by applying HMM.

2. Human Action

Human action is change of location or position of human body or any parts of the body with respect to time. Change in motion is the result of an applied force. Human action is typically described in terms of velocity, displacement, and time. Human action is always observed and measured relative to a frame of reference. So, the term Human action signifies any spatial and/or temporal change in a physical system. Human motion analysis is receiving increasing attention from computer vision researchers. This interest is motivated by a wide spectrum of applications, such as athletic performance analysis, surveillance, man-machine interfaces, content-based image storage and retrieval, and video conferencing. Different types of human actions are walk, jump, run, bend, one-hand wave, two-hand wave etc.

3. Hidden Markov Model (HMM)

A Markov model is a probabilistic model of symbol sequences in which the probability of the current event is conditioned only the previous event. An HMM is a non-deterministic Markov model that is, one where acknowledge of the emitted symbol does not determine the state-transition. This means that in order to determine the probability of a given string, we must take more than one path through the states into account. A HMM is characterized by three matrices: State transition probability matrix A, symbol output probability matrix B, initial state probability matrix π. To apply HMM to time sequential image, the image must be transformed into symbol sequence. With a hidden Markov model, we usually model a temporal process whose output we can observe but do not know the actual understanding mathematical or physical model.

HMMs make it possible to deal with time-sequential data and can provide time-scale invariability in recognition. HMMs are characterized by their learning ability which is achieved by presenting time-sequential data to a HMM and automatically optimizing the model with the data. A HMM consists of a number of states each of which is assigned a probability of transition from one state to another state. States at any time depends only on the state at the preceding time. One symbol is yielded from one of the HMM states according to the probabilities assigned to the states. HMM states are not directly observable, and can be observed only through a sequence of observed symbols. HMM has many applications in many areas, such as weather prediction,

speech & handwriting recognition, signature & finger print identification & so on.

3.1 Elements of a HMM

N=number of states in model
M=number of distinct observation symbols
Q={1,2,… ,N} : set of hidden states
V={1,2,… ,M} : set of observation symbols
A: state transition probability matrix
  ✓ $a_{ij} = P(q_{t+1}=j|q_t=i)$
B: observation symbol probability
  ✓ $b_j(k) = P(o_t=k|q_t=j)$
π: initial state distribution
  ✓ $\pi_i = P(q_1=i)$
λ: the entire model λ=(A,B,π)

4. Extraction

Visual recognition of human actions in video clips has been an active field of research in recent years. We will operate not on entire video sequences, but on very short sub-sequences. The initial stage of human motion analysis problem is the extraction of moving targets from a video stream. There will be spurious pixels detected, holes in moving features, "interlacing" effects from video digitization processes, and other anomalies. Foreground regions are initially filtered for size to remove spurious features, and then the remaining targets are pre-processed before motion analysis is performed. To apply HMM to our aim, one set of time-sequential image is transformed into an image feature vector sequence, and the sequence is converted into a symbol sequence by vector quantization. Model parameters are obtained from the symbol sequence.
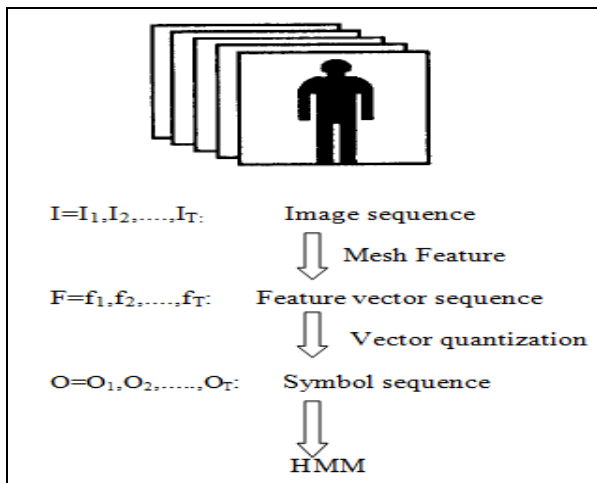

Figure 1: Processing flow

4.1 Steps of symbol sequence calculation:
1) First we need to take two image frame at t=0 sec and t=1 sec from Weizmann database [12] for each action.
2) Second we need to find out the difference image. For this purpose the difference image must be same size.
3) Third the binary image obtained from the difference image by thresholding. We take threshold value 30 for this calculation.

4) Fourth we need to subdivide the binary image at different block. We take sixteen blocks here.
5) Fifth we find the total black/white pixels of whole image and total black/white pixels at each block. Divided the black/white pixels at each block by total black/white pixels of the whole image.
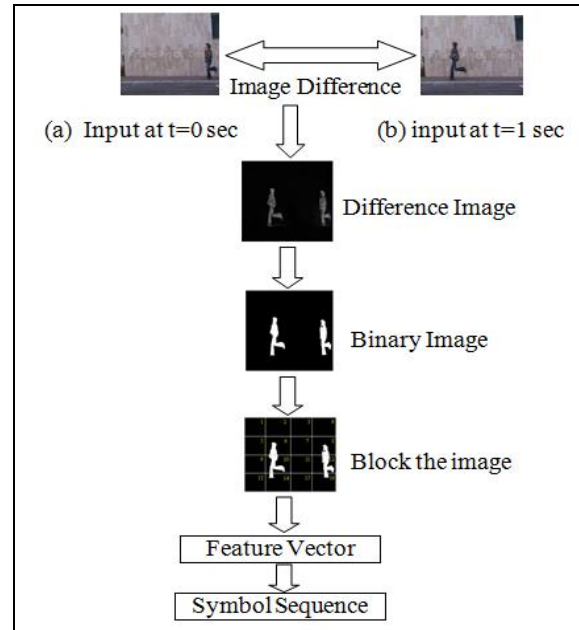6) Sixth by vector quantization we calculate the output symbol sequences.

Figure 2: Steps of Symbol Sequence calculation

4.2 Symbol Sequence Calculation

We take four persons for each action (Bend, Jump, Run & Walk) and take two time sequential image, one (a) at t=0 sec and another (b) at t=1 sec for each person which is shown in figure-3 to figure-6 respectively. Then we calculate four symbol sequences for each action which is shown in table-1 to table-4 respectively.
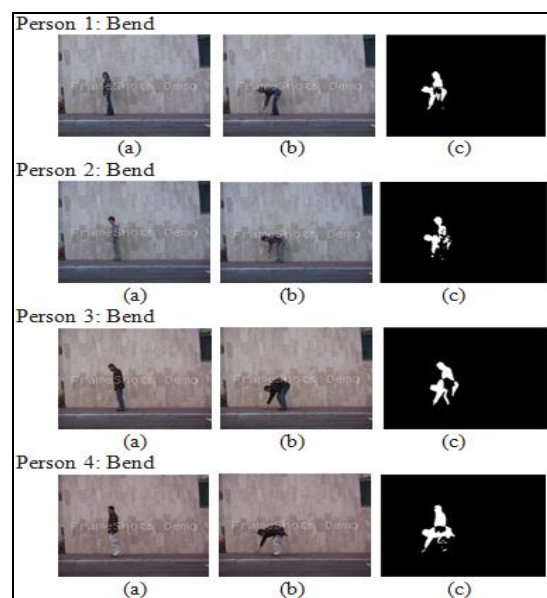

Figure 3: Bend ((a) at t=0s, (b) at t=1s, (c) Binary Image)

Table 1: Output symbol sequences for Bend

Person1:	1	1	1	1	1	9	1	1	3	12	1	1	1	1	1	1
Person2:	1	1	1	1	1	8	1	1	1	15	1	1	1	1	1	1
Person3:	1	1	1	1	1	9	1	1	1	14	1	1	1	1	1	1
Person4:	1	1	1	1	1	10	1	1	1	13	1	1	1	1	1	1

Table 3: Output symbol sequences for Run

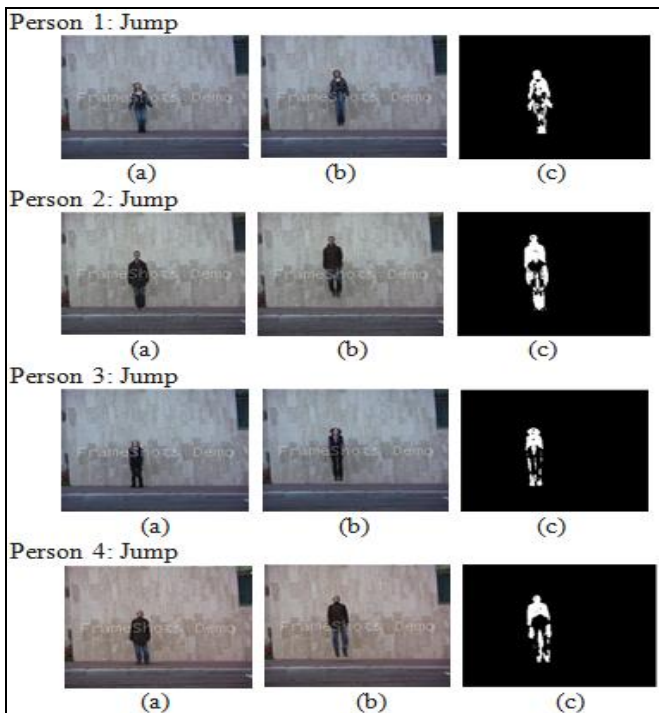| Person1:	1	1	1	1	1	5	1	4	1	8	1	8	1	2	1	2
Person2:	1	1	1	1	1	5	1	3	1	7	2	8	1	2	1	1
Person3:	1	1	1	1	1	7	1	6	1	6	1	7	1	1	1	1
Person4:	1	1	1	1	1	7	1	5	1	6	1	7	1	1	1	1

Figure 4: Jump ((a) at t=0s, (b) at t=1s, (c) Binary Image)
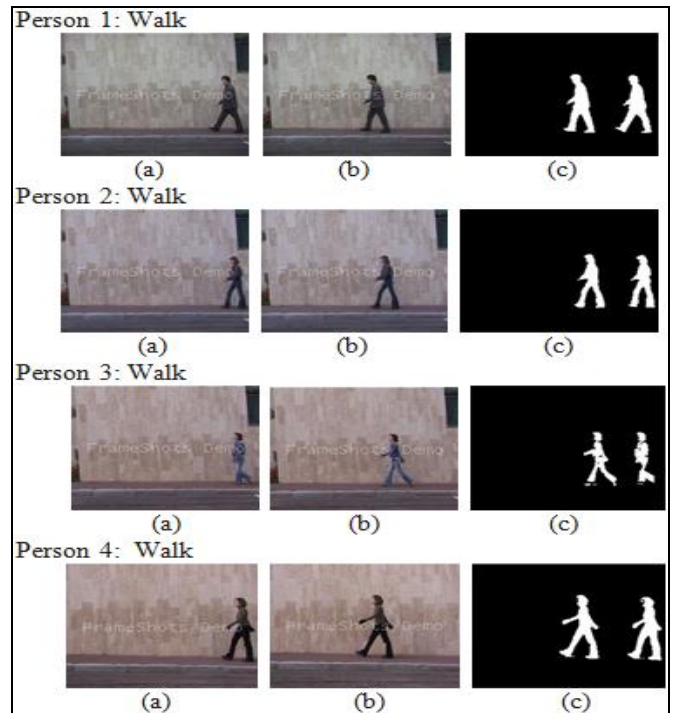

Figure 6: Walk ((a) at t=0s, (b) at t=1s, (c) Binary Image)

Table 2: Output symbol sequences for Jump

Person1:	1	1	1	1	1	15	1	1	1	7	1	1	1	3	1	1
Person2:	1	2	1	1	1	14	1	1	1	8	1	1	1	2	1	1
Person3:	1	1	1	1	1	17	1	1	1	4	1	1	1	3	1	1
Person4:	1	1	1	1	1	15	1	1	1	7	1	1	1	3	1	1

Table 4: Output symbol sequences for Walk

Person1:	1	1	1	1	1	1	4	4	1	1	8	8	1	1	3	3
Person2:	1	1	1	1	1	1	4	4	1	1	10	8	1	1	2	2
Person3:	1	1	1	1	1	1	5	5	1	1	7	8	1	1	1	2
Person4:	1	1	1	1	1	1	4	4	1	1	8	8	1	1	3	2

5. HMMs Calculation

We recognized the HMMs from symbol sequence of four persons. Who perform Jump, walk, bend and run action respectively. We get three different models for each person for each action. Then we calculate log-probability from observation sequence from other four different persons by using Viterbi Algorithm & HMMs. We compare log-probability & maximum log-probability represents the correct action. Our expectation will be maximum log-probability for same action. For Jump, walk, bent and run we calculate total 16 symbol sequences. Each symbol is used as a test & training sequence.

We calculate the HMMs for Jump, walk, bend and run for one person and calculate log-probability by using other 15 test symbol sequence for Bend, Walk, Run and Jump. Similarly we calculate the HMMs for each symbol sequence and also calculate log-probability by using other 15 test symbol sequence for Bend, Walk, Run and Jump. We get three HMMs for one symbol sequence.

Figure 5: Run ((a) at t=0s, (b) at t=1s, (c) Binary Image)

HMM 1 for Jump (Person 1):

```
M= 16, N= 3

A:
0.459610  0.540951  0.001439
0.001948  0.451226  0.548826
0.999256  0.001740  0.001003

B:
0.999694 0.001000 0.001124 0.001000 0.001000 0.001000 0.001120 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001061 0.001000

0.999613 0.001000 0.001146 0.001000 0.001000 0.001000 0.001132 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001109 0.001000

0.001382 0.001000 0.333786 0.001000 0.001000 0.001000 0.333824 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.334009 0.001000

pi:
0.999735  0.001263  0.001002
```

HMM 2 for Jump (Person 1):

```
M= 16, N= 3

A:
0.450536  0.001407  0.550057
0.999212  0.001003  0.001785
0.001959  0.540274  0.459767

B:
0.999718 0.001000 0.001116 0.001000 0.001000 0.001000 0.001112 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001055 0.001000

0.001383 0.001000 0.333785 0.001000 0.001000 0.001000 0.333824 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.334007 0.001000

0.999588 0.001000 0.001155 0.001000 0.001000 0.001000 0.001141 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001115 0.001000

pi:
0.999707  0.001002  0.001291
```

HMM 3 for Jump (Person 1):

```
M= 16, N= 3

A:
0.458770  0.541794  0.001436
0.001946  0.452048  0.548006
0.999252  0.001744  0.001003

B:
0.999697 0.001000 0.001124 0.001000 0.001000 0.001000 0.001119 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001060 0.001000

0.999611 0.001000 0.001147 0.001000 0.001000 0.001000 0.001133 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.001109 0.001000

0.001382 0.001000 0.333786 0.001000 0.001000 0.001000 0.333824 0.001000
0.001000 0.001000 0.001000 0.001000 0.001000 0.001000 0.334009 0.001000

pi:
0.999732  0.001265  0.001002
```

Similarly we get three HMMs for each symbol sequence.

6. Recognition Process

We calculate different HMMs & pass test sequence through these HMMs & obtained different log-probability. Highest rank represents the recognize action. The recognition process is shown in below-
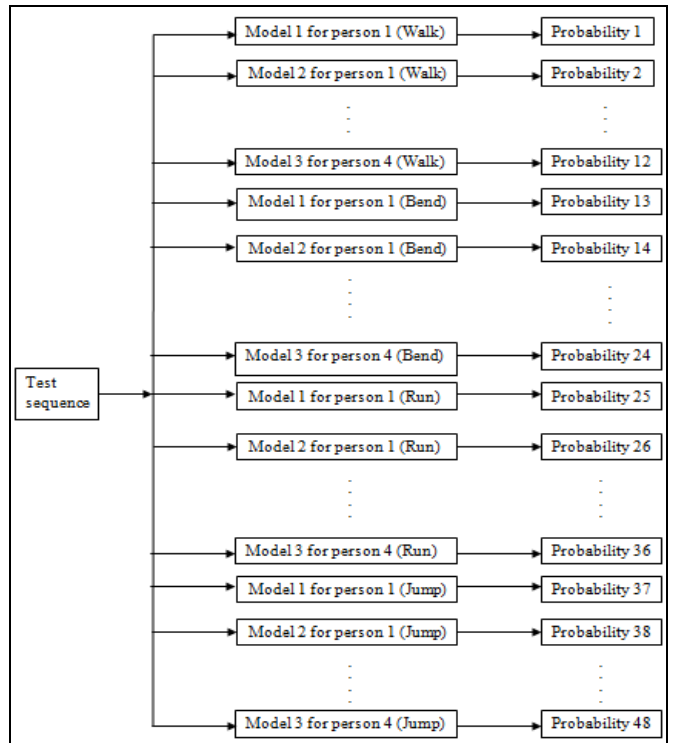

Figure 7: Recognition Process

6.1 Log- probability calculation for different sequences

We have total 16 symbol sequences. We get three HMMs for one symbol sequence and pass another 15 symbol sequences through these HMMs and calculate log-probability.

Table 5: Log-Probability for Bend

Training Sequence	Test Sequence		First Model P(O|λ)	Second Model P(O|λ)	Third Model P(O|λ)
Bend Person-1	Bend	Sequence-2	-16.05	-16.05	-16.05
		Sequence-3	-11.77	-11.77	-11.77
		Sequence-4	-16.05	-16.05	-16.05
	Jump	Sequence-1	-22.88	-22.88	-22.88
		Sequence-2	-29.71	-29.71	-29.71
		Sequence-3	-22.88	-22.88	-22.88
		Sequence-4	-22.88	-22.88	-22.88
	Run	Sequence-1	-43.38	-43.38	-43.38
		Sequence-2	-43.38	-43.38	-43.38
		Sequence-3	-29.71	-29.71	-29.71
		Sequence-4	-36.55	-36.55	-36.55
	Walk	Sequence-1	-38.87	-38.87	-38.87
		Sequence-2	-43.38	-43.38	-43.38
		Sequence-3	-36.55	-36.55	-36.55
		Sequence-4	-39.58	-38.87	-39.58
Bend Person-2	Bend	Sequence-1	-27.25	-22.76	-27.25
		Sequence-3	-20.34	-15.85	-20.35
		Sequence-4	-20.34	-15.85	-20.35
	Jump	Sequence-1	-19.60	-17.70	-19.60
		Sequence-2	-26.12	-24.20	-26.12
		Sequence-3	-24.25	-22.46	-24.65
		Sequence-4	-19.59	-17.60	-19.60
	Run	Sequence-1	-34.12	-32.25	-34.13
		Sequence-2	-39.56	-37.72	-39.56
		Sequence-3	-31.18	-29.07	-31.19
		Sequence-4	-38.09	-35.10	-38.10
	Walk	Sequence-1	-40.31	-38.32	-40.32
		Sequence-2	-40.31	-38.32	-40.32

		Sequence-3	-33.40	-31.41	-33.41
		Sequence-4	-40.31	-38.32	-40.32
Bend Person-3	Bend	Sequence-1	-21.94	-21.28	-21.94
		Sequence-2	-20.57	-19.41	-20.58
		Sequence-4	-20.57	-19.41	-20.58
	Jump	Sequence-1	-27.71	-25.96	-27.69
		Sequence-2	-28.98	-27.51	-28.97
		Sequence-3	-27.71	-25.96	-27.69
		Sequence-4	-27.71	-25.96	-27.69
	Run	Sequence-1	-48.43	-45.61	-48.41
		Sequence-2	-48.43	-45.97	-48.41
		Sequence-3	-34.61	-32.51	-34.60
		Sequence-4	-41.52	-39.42	-41.52
	Walk	Sequence-1	-48.43	-46.69	-48.41
		Sequence-2	-48.43	-46.69	-48.41
		Sequence-3	-41.52	-39.78	-41.51
		Sequence-4	-48.43	-46.69	-48.41
Bend Person-4	Bend	Sequence-1	-27.43	-27.75	-27.45
		Sequence-2	-20.53	-20.84	-20.54
		Sequence-3	-20.53	-20.84	-20.54
	Jump	Sequence-1	-22.38	-27.75	-22.39
		Sequence-2	-29.00	-34.66	-29.00
		Sequence-3	-22.38	-27.75	-22.39
		Sequence-4	-22.38	-27.75	-22.39
	Run	Sequence-1	-41.22	-48.47	-42.23
		Sequence-2	-42.52	-48.47	-42.52
		Sequence-3	-28.99	-34.66	-29.00
		Sequence-4	-35.90	-41.56	-35.91
	Walk	Sequence-1	-43.10	-48.47	-43.11
		Sequence-2	-43.10	-42.71	-43.11
		Sequence-3	-36.20	-41.56	-36.21
		Sequence-4	-43.10	-48.47	-43.11

Similarly, we calculate log-probability for Bend, Jump, Run and Walk. For bend (person 1) we get 3 HMMs & pass other 15 test sequence through these. Then we obtained different log-probability. For bend (person 1) our required result was Bend. Our output highest log- probability for bend which indicate our expectation and output results are same. Similarly based on log-probability (shown above) we recognize different action.

7. Conclusion

We have used Hidden Markov Model based approach for human action recognition from a set of time-sequential images. In our algorithm, a mesh feature vector sequence extracted from time sequential images is converted to a sequence of symbols. In learning, symbol sequences obtained from training image sequence data are used to optimize HMMs for action categories. In recognition, a symbol sequences from an observed image sequence is processed by HMMs, and the recognition result is determined as the category which best matches the observed sequence. We performed four different types of human action- bend, jump, run and walk. We obtained high recognition rate for jump and bend but low recognition rate for run and walk. We need to increase the recognition rate. If we increase the training data then increase the recognition rate.

8. Acknowledgement

References

[1] B.Horowits and A. Pentland. "Recovery of Non-Rigid Motion and structure", In Proceeding of CVPR, pp. 325-330, 1991.

[2] Yang He and Amlan Kundu. "Planar Shape Classification using Hidden Markov Model", In Proc. CVPR, pp. 10-15, 1991.

[3] David Hogg."Model –based vision:a program to see a walking person", Image and vision computing, Vol. 1, No. 1, pp. 5-20, Feb 1983.

[4] B.W. Hwang and S.Takaba. "Real-time Measurement of Pedestrian Flow Using Processing of ITV Images", Trans. IEICE, Vol. J66, No. 8, pp. 917-924, 1983. (in Japanese).

[5] J.Aloimonos. "Purposive and Qualitive Active Vision", In Proceeding of ICPR, pp. 346- 360, 1990.

[6] A.Kunda, Y.He, and P.Bahl. "Handwritten Word Recognition: a Hidden Markov Model Based Approach", Pattern Recognition, pp.283-297, May 1989.

[7] J.O'Rourke and N.Badler. "Model BasedImage Analysis of Human Motion Using Propagation", IEEE Trans.PAMI, Vol. PAMI-2, No. 6, pp. 522-536, Nov 1980.

[8] E.Oscarsson. "TV-Camera Detecting Pedestrians for Traffic Light Control", In Technological and Methodological Advances Measurement (ACTA IMEKO), Volume 3, pp.275-282, 1982.

[9] L.R. Rabiner and B.H. Juang. "An Introduction to Hidden Markov Models", IEEE ASSPMAGAZINE, pp. 4-16, Jan 1986.

[10] T.Aong and M.Ishikawa. "Sensor Fusion Using Stochastic Process", In 2nd Symposium of Autonomous Distributed Systems, pp. 115-118, 1991.

[11] Weizmann DB, http://www.wisdom.weizmann.ac.il.

Author Profile

Uttam Kumar Kar received the *B.Sc. degree in Electronics & Communication Engineering* from *Khulna University of Engineering & Technology*, *M.Sc. degree in Computer Science & Engineering* from *Uttara University* in 2010 and 2017 respectively. He served as a Site Engineer in Crystal Engineers & Builders Limited in Bangladesh during 2010-2011. He has College level teaching experience in Dhaka Medical Institute in Bangladesh during 2011-2013. Now he is working as a System Engineer in LM Ericsson Bangladesh Ltd since 2013.

Md. Shafiul Alam Chowdhury received the *Bachelor of Computer Applications* degree from *University of Madras (India)*, *M.Sc. in Computer Science* degree from *Bharathidasan University (India)*, *M.Sc in Interactive Systems Engineering* degree from *Royal Institute of Technology-KTH (Sweden)* in 2001, 2003 and 2007 respectively. He stayed in Expertise Centre for

Digital Media & Hasselt University in Belgium under a research project during 2007-2008. He served Information-Technology Manager Post in Rahman Export-Import-Trading (S) Ltd Company in Singapore during 2009-2011. He has College and University level of teaching experience in Cantt-Public School & College - Dinajpur, Uttara University, and City University in Bangladesh respectively. He now is working as an Associate Professor and Chairman (Head of the dept.) of Computer Science & Engineering department at Uttara University, Bangladesh. He actively participated in number of national/ international conferences, seminars, workshops and journals as a paper presenter, session chair, editorial board-member etcetera. He has research interest in the area of Automatic Speech Recognition System, Ubiquitous Computing, Pervasive Computing, Human-Computer-Interaction (HCI), IT-Management, IT-Strategy, E-government, CRM etcetera.