STUDY ON SMART CAMERA

[1] Dr. B.Mukunthan, Ph.D, [2]T.Saravanan M.Sc.,M.A.,M.Phil.,(Ph.D), [3]R.Neethimathi M.C.A,

[1] Research Advisor, [2] Assistant Professor, [3] Research Scholar,
Jairam's Arts and Science College, Karur, Tamilnadu, India
[1]Research Department of Computer Science,
[1] Jairam's Arts and Science College,Karur,Tamilnadu,India

Abstract : ***Once the cameras are well placed and calibrated in a surveillance zone for a specific task, control flow among the cameras is a crucial stage toward development of a smart MCN based surveillance system. This chapter proposes a smart MCN model which uses the architecture of MCN but avoids its complexities and overheads, by letting single camera to track the subject at any instant of time. The control flow from one camera to another is governed by an occlusion determination algorithm that determines the chances of occlusion, so that with prior knowledge of occlusion, control can be forwarded to such camera that does not encounter any occlusion. This way multiple track of the subject can be avoided (an overhead in MCN based surveillance), at the same time uninterrupted track of the subject (a limitation in single camera based surveillance) can be achieved. The discussed multi-camera model for visual surveillance works on a single camera which is a part of multi-camera system and instead of handling occlusion; it pre-determines occlusion and avoids its occurrence. The proposed approach analyses the change in the dimension of the subject in camera coordinates as it moves in the 3D world coordinate. It analyses the data to decide the direction of motion and apparent speed of the subject and further determines chances of occlusion and its time and location of occlusion in the camera plane. Based on this, further decision towards avoiding occlusion can be made. The analytical framework contains.***

IndexTerms - Motion analysis, LUT,GPS, RAM,CASIA Database

I. INTRODUCTION

Background subtraction is a reliable method for localization of a foreground with respect to a fixed background. The results of background subtraction are used in this approach for analysis. A few reasonable assumptions are made in our approach while considering motion of subjects, such as eight possible directions of motion and three levels of speed for any subject on move have been considered. This assumption discritizes the approach at both the levels of direction as well as speed. The proposed system can be well described by sub-dividing into following three steps:

(i) Motion analysis
  (a) Direction of motion determination
  (b) Apparent speed determination
(ii) Occlusion determination
(iii) Mitigation of occlusion
(iv)

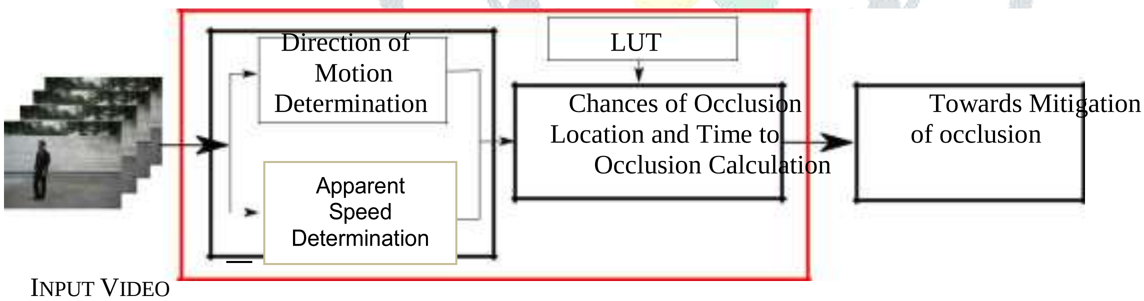


Figure 1: proposed camera control model governed by occlusion determination algorithm

The above three steps are depicted in the form of block diagram in Figure 1. Motion analysis of a subject is performed to determine the direction of motion as well as speed of the subject as observed in the image plane. Variation of height and width with respect to frame number of a subject's motion is been exploited to determine the direction of motion.

Kulkarni et al. have proposed an approach for efficient use of multiple cameras by devising multi-tier camera network called SensEye [1, 2]. This approach is energy efficient although it has a complex hardware architecture and diverse software requirement. Automation of the localization process started gaining importance to ascertain accuracy and real-time localization. One of the primitive automated solutions to localization has been through GPS [3]. However, it has failed mostly due to the poor resolution. Efforts have also been made towards developing localization algorithms on single processor after collecting images from all the networked cameras in a single room [4, 5]. But in practical scenario, large number of cameras producing high volume of images and video data makes the analysis time-consuming on single processor. The subsequent attempts of developing localization algorithms deploy more than one processor concurrently to achieve real-time localization. These approaches differ in variety of coverage areas, assumptions made on deployment of the nodes, and the way sensors work [6]. Early automated localization techniques for static sensors, viz. non-camera equipped networks have used ultra-sound, radio, or acoustic signals [7].

II. DATABASE USED

Experiments are performed over both publicly available as well as self acquired databases. Initially experiments are conducted over a set of self acquired database. In order to verify the algorithm over a globally available database, CASIA Dataset A [8] gait database has been used. Since the data is intended for angle invariant gait pattern based subject identification, they are not sufficient for the proposed experiments. For the sake of direction determination, the database is modified and mirror imaged. Our experiments of determining discrete direction, the same database has been modified and mirror image data have been generated to have enough video in all eight directions as needed for this experiment.

The modified database contains 20 subjects making 2 walks in each direction, thereby accumulating 40 videos in each direction, hence a total of 320 videos. These videos have to be classified into 8 directions based on the analysis of pattern of dimensional change in the bounding box of the subject.

Background subtraction method has been used for extracting foreground in self acquired sample videos for system testing, however the investigation results are generated on both self acquired and the CASIA database which is already background separated and algorithm to remove unwanted blobs are subsequently applied to them.

III. MOTION ANALYSIS

Motion analysis of the subject gives information about the direction and apparent speed of the subject's movement. This section analyses the motion of the subject to determine direction and speed and presents experimental steps, inferences and results.

3.1 Determination of Direction of Motion

A subject in a plane is free to move in all the directions. Calculation of exact direction of movement is neither a perceptive solution for real time execution nor it is feasible with low resolution video footage. Since the prime motive of direction calculation is to contribute towards occlusion mitigation, the possible direction of subject's motion has been reduced to eight discrete directions as encountered by the camera. Figure 3.2 shows the direction vectors labelled from 1 to 8 with respect to camera. These direction vectors are equidistant and hence distinguishable. However,

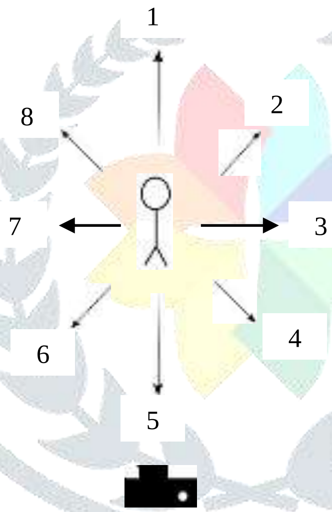


Figure 2: Discrete directions of motion with respect to camera

increasing the number of direction vectors decreases the estimation result on low resolution videos and decreasing the number of direction vectors, affects further processing , hence the discretization of eight direction vector is quite justified. Since the direction of motion has to be utilized for approximating the chances of occlusion, discrete direction can be applied. The discrete direction will also provide faster computation which is needed for real time processing. To realize the direction of motion of a subject, change in the width, height, and location co-ordinates of the bounding box of the subject is studied. The pattern change in the subsequent frames of sample video during the motion in the perspective view of camera is shown in Fig. 3. The change in the dimension as well as location in the camera frame of a subject together make a unique pattern for each of the direction. Variations of height and width with respect to frame number for four different direction of motion of a subject are plotted in Figures 3.4 and 3.5. This gives the pattern based on which the direction of motion can be explained.

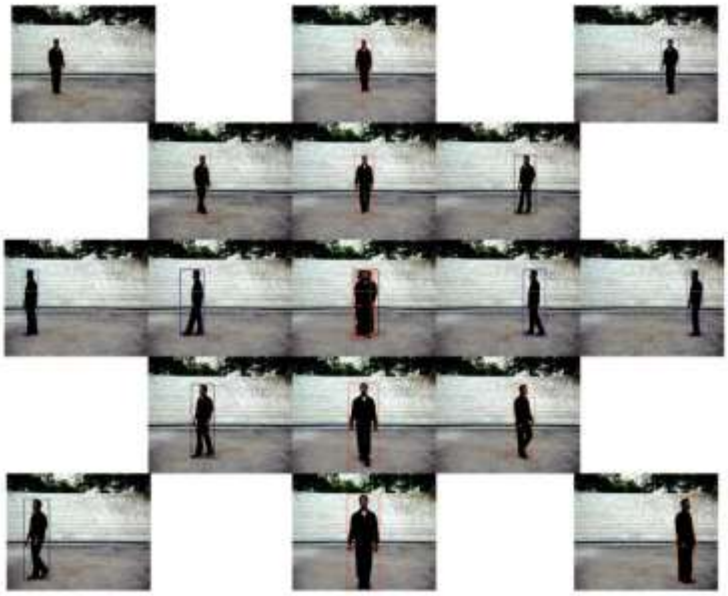


Figure 3: Pattern change in the dimension of subject

Study of Frames and Inferences

Fig. 4 and 5 show change of pattern in height and width of the subjects as it moves along different directions. When the subject is moving orthogonal to the view axis (i.e. along direction vector 3 or 6), the cyclic pattern is visible in regular interval in the width graph, however height remains constant as can be seen in the height graph. When the subject is moving along the view axis (i.e. along direction vector 1 or 5), then the subject appears to be growing in width and height from vanishing point and vice versa for opposite motion. Hence it is very obvious that the width and height has some pattern distinguishing them from another and a smart mechanism is needed to identify them. Next section of this chapter presents various steps taken towards direction estimation of moving subject and its further implementation.

Experiment

The aim of the experiment is to generate a phase or phase band of the graph that represents the direction of motion of the subject.

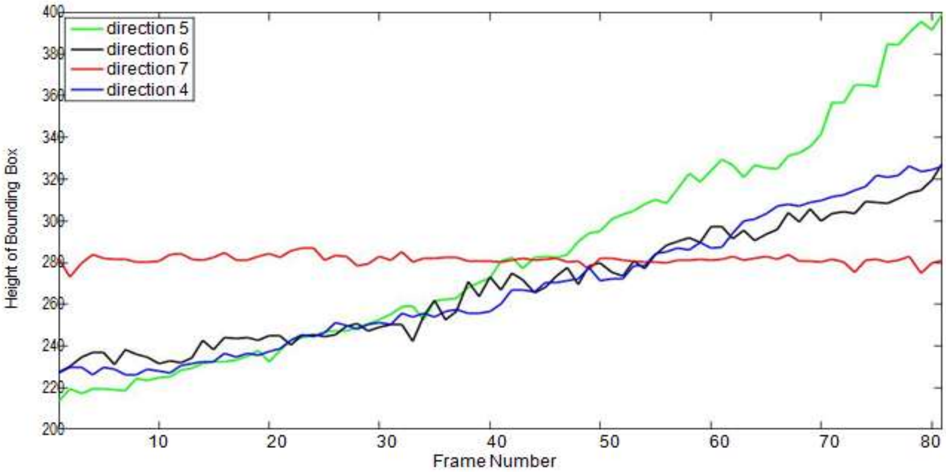


Figure.4: Variation of height with respect to frame number for four different direction of subject's motion

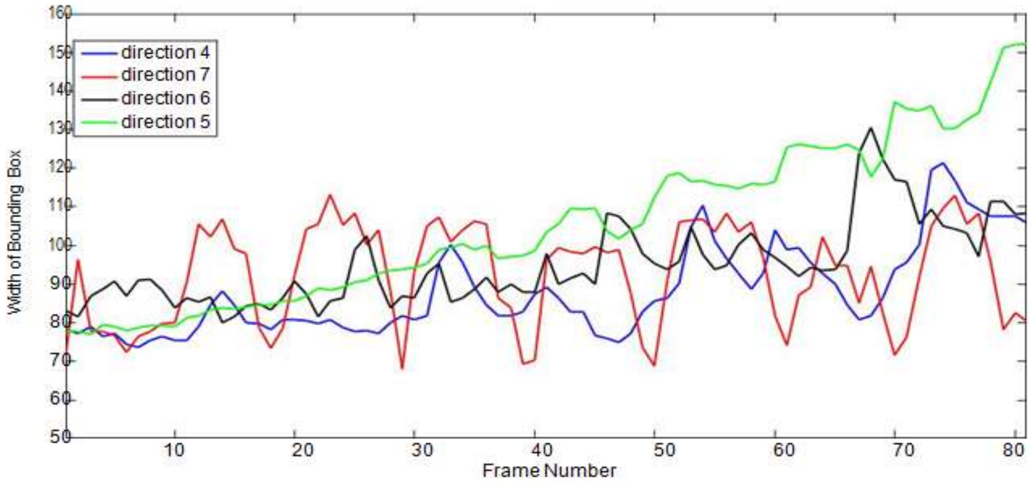


Figure 5: Plot of width for different direction

Challenges In order to generate a unique phase or phase band for motion in a particular direction, a system is required to be robust towards many issues that has been listed here:
Low resolution database.
Static occlusion causing foreground to be unavailable for few frames.

- Improper segmentation.
- Variable speed during the walk.
- Direction of motions being approximate rather than exact.
- Different distances of subject with respect to camera.
- Different amount of distances covered or different number of gait cycles of walk available for analysis.

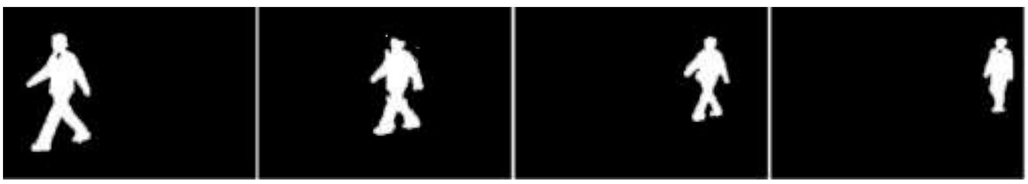
In order to meet above mentioned challenges, following stages are performed towards achieving a unique phase band for each direction. The following paragraphs elucidates each of these stages in sequence.

1) Frame Rectification and Removal of Undesired Blobs The frame sequence provided in the database are segmented and are converted to binary image sequences. Frame contains unwanted blobs due to improper segmentation, presence of noise and partial occlusion. Unwanted foregrounds are deleted by selecting of largest connected component. Optical flow based methods are applied for rectification of improper segmentation. Figure 6 shows, few frames from two video sequences and their rectified forms after unwanted blob removal.

2) Morphological Operations and Tracking the Subject On the rectified frame sequence, moving subject is identified and rectangular bounding box is fitted over them to get the track of the moving subject. Further, the subjects are tracked where the pattern of change in the dimensions of the subject are recorded for further reference. In the perspective view, the dimensions of the subject are varying, and this fact has been exploited to differentiate between the subjects that are moving in different directions. The pattern in the temporal change of width as well as height are plotted, and different plots for height and width are obtained which have been utilised for two way analysis for estimating the direction of motion of the subject under study. In the Fig. 7, a frame sequence with direction vector 6 is shown after noise removal and morphological operations.

(c) Extrema Detection and Putting Envelop Over the Plot After getting the plots of the frame sequence, next objective is to distinctively identify the plots such that plots of same direction of motion should come under same identifier. Subjects may be observed nearer or farther from the camera and hence the graph in both the cases may look different although there may be similarity in the pattern of the graph. Thus to achieve distance invariance, and to overcome a few of the improper segmentation, graphs have been proposed to be presented in terms of linear regression of the extrema boundaries of the graph. This step has been elaborated and divided into following sub-steps for better understanding:

(a) Frame sequence 1 with noise



(b) Frame sequence 1 after noise removal

(c) Frame sequence 2 with noise

(d) Frame sequence 2 after noise removal
Figure 6: Frame rectification and unwanted blob removal

(a) Frame sequence with bounding box
Figure 7: Morphological operations and tracking of subject

Covering envelope over the plots The prime objective of finding the envelope is to process the envelope further to boil down the graph into a line of the form y = mx + c,

However, the span of the envelope i.e. the difference between m1 and m2 (as shown in Fig. 8) and the area covered by envelop, can be used to study the distance of the subject from camera in particular cases as well. Higher the area covered, closer the subject is from camera. Upper envelope has been made from a set of local maxima points representing the maximum width or height of the subject recorded in a gait cycle while capturing subject's movement.

Soft Extrema Elimination The envelope has been made over local maxima and minima points in the graph, but improper segmentation of videos has resulted into some trough regions formed at upper fragment of the graph and certain crests are formed at lower fragment of the graph. Due to this, a few minima points exist in the maxima region and vice versa. This can be seen

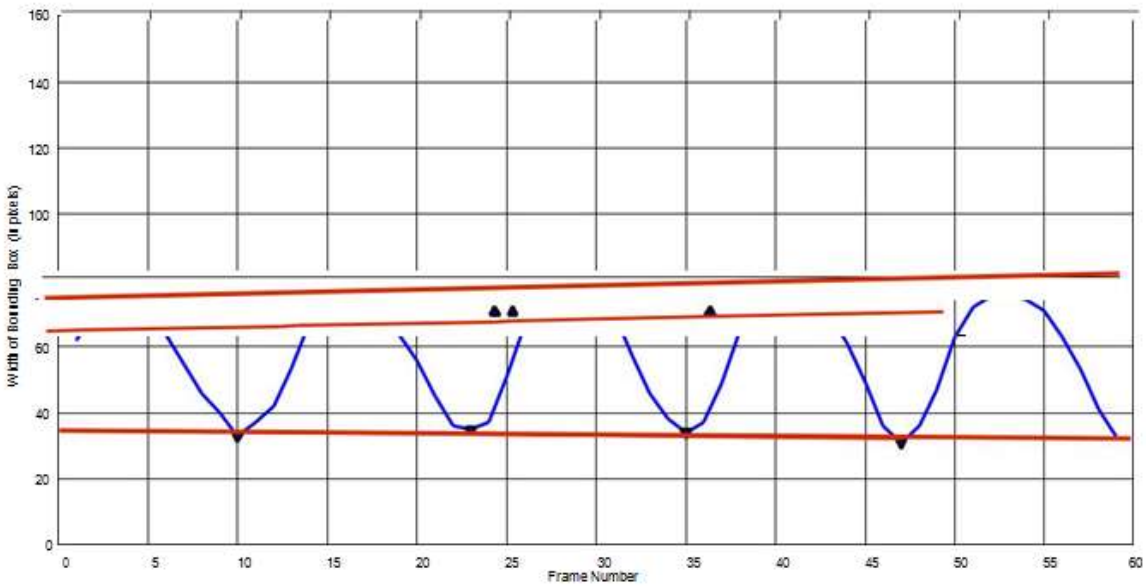
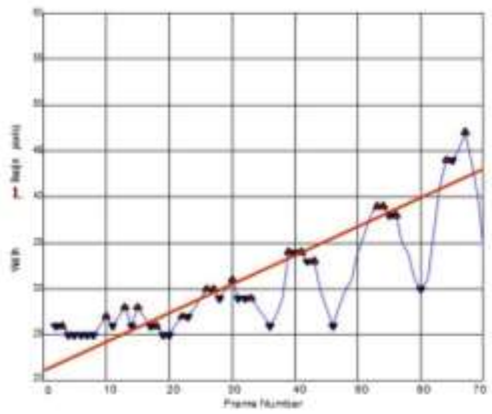
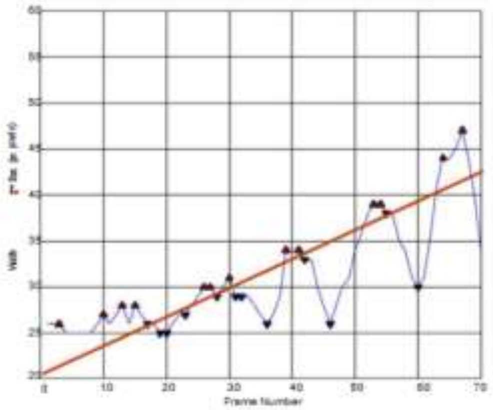


Figure 8: Envelop over the plot

in Fig. 9, which is a graph of a particular subject in CASIA Dataset A) walking at an angle of 45 with respect to camera. Such misplaced extremas are eliminated before calculation for putting envelop. A linear fitting is done over the curve to decide whether an extrema is correctly placed or not.





a) Extrema points including soft extremas (b) Extrema points after excluding soft extremas

Figure 9: Detection and removal of soft extremas

Linear regression is a statistical analysis for association between two variables. It is used to find the relation between them. In the context of the proposed work, our objective is to eliminate erroneous data that could contribute in constructing envelop over the graph. To identify such points, a linear regression function has been cast-off, where 2-tuple variable point's co-ordinates are, (x, y) = (dimension, frame number) where dimension: length or width in different graphs

The linear regression relationship between x and y is given in the slope-intercept straight line equation form as:

$$y = mx + c$$

where,

$$m = \frac{n\sum y - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$$

and,

$$c = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

and, n = number of variable pairs to be regressed, in this case they are the number of readings for each walk i.e. the number of frames in the video under study The resulted regression line y = mx + c is plotted on both the graph types i.e. width vs frame number and height vs frame number. Maxima points are always expected to lie above this line and minima are below this line. Those points that do not satisfy these criteria are named soft extremas in this context and they have to be eliminated. After elimination of such points, the envelop generated are represented by these two equations

$$y = m_1x + c_1$$

$$y = m_2x + c_2$$

Finally, a line as average of these two lines are plotted, which is represented as y = mx + c

Where, $m = \frac{(m1 + m2)}{2}$

and, $c = \frac{(c1+c2)}{2}$

Fig. 3.10(a), shows envelop drawn by the regression lines of the extremas after removal of soft extremas. Later, the average of the two lines has been taken are drawn as phase and shown in Fig. 10(a)

(c) Study of Phase of the Line The line represented as y = mx + c carries the information of phase i.e. direction of motion of the subject and its distance from the camera. Thus we have distance invariant, phase information of the motion of the subject with respect to the studies of width and height of the subject.

Thus two different equations obtained from different graphs are

$y = m_h + c_h$ from the study of change in height $y = m_w + c_w$ from the study of change in width Where

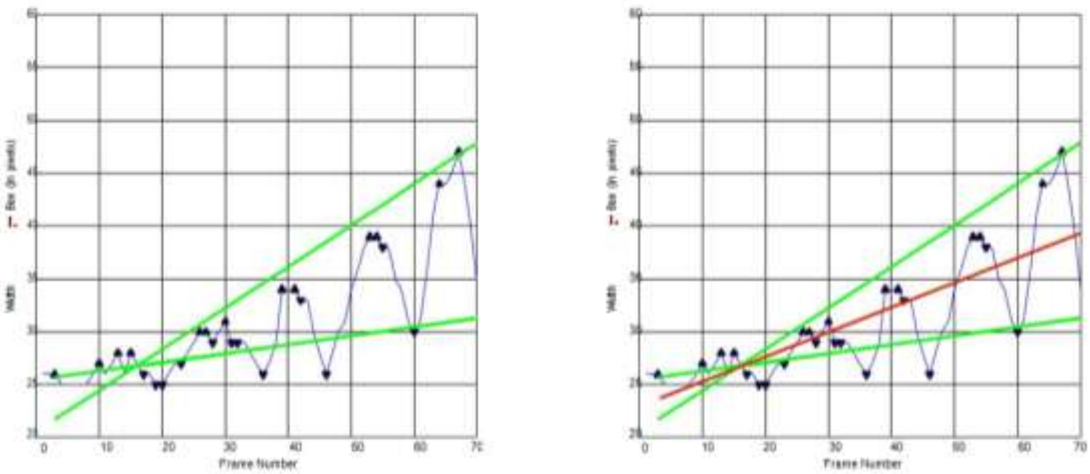
m_h : Phase representing subjects direction of motion with respect to height

m_w : Phase representing subjects direction of motion with respect to width

c_w : Representing subjects distance from camera with respect to width

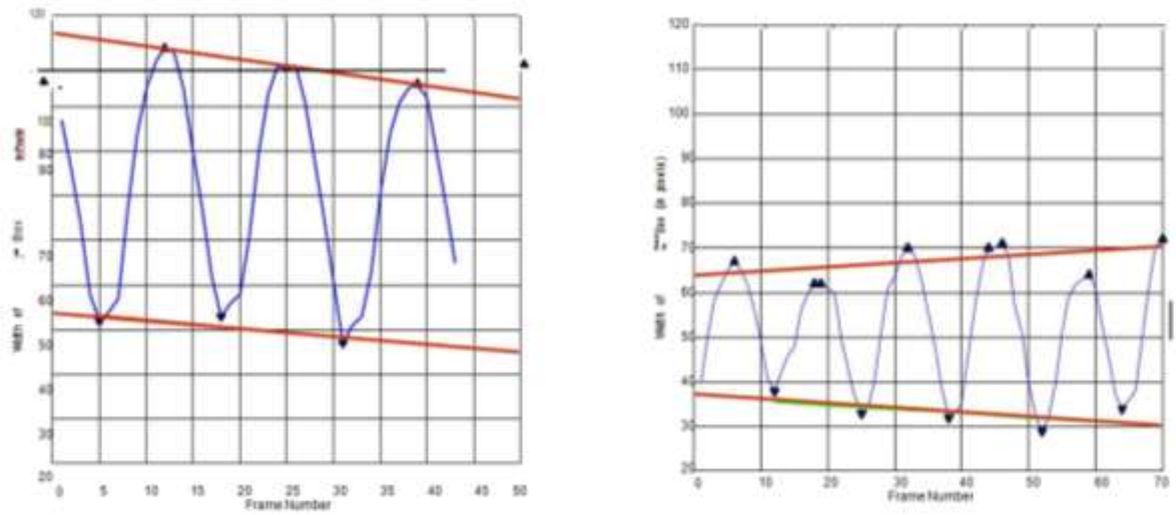
Subjects may be moving nearer to or farther from the camera; however the phase of the subject does not alter with the distance of the subject. Fig. 11(a) and 11(b) represent subject moving in the direction 3 but at different distances from the camera. Distance from the camera affects the span of the graph generated, however the phase of the subject in both the cases are nearly same thus bringing distance invariance in the system.

(e) Plotting of Phase and Determination of Direction of Motion For each sample video, we have its phase information with respect to height and width. All these values are plotted and two separate graphs are obtained having all m_h and m_w information. The m_h and m_w of the subjects moving in the same direction (i.e. the in-phase videos) are shown as connected points in separate graphs in the next section. Thus the phase information has been exploited to estimate the direction of



(a) Regression line on extremas as envelop to (b) Average line depicting phase of the graph the graph

Figure 10: Envelop and average line drawn based on envelop motion of the subject



(a) Subject moving closer to the camera (b) Subject moving farther to the camera

Figure 11: Distance invariant direction of motion estimation through phase

IV CONCLUSION

The direction of the subject has been estimated based on the height and width information of the moving subject. Fig. 10 and 11 shows the results in graphical representation where connected points in the graph represent the phase of the subject moving in the same direction. Out of 8 discrete directions, direction 2 & 8, and 4 & 6 overlap, since the videos are mirror image and mod of slopes are plotted in the graph. They are presented with different markers attached with them, as shown in corresponding legend chart. Further best fitting linear separator is applied on the curves to minimize misclassification and achieving best accuracy. The system has overall good estimation accuracy as follows:

- height based accuracy is 93.75%
- width based accuracy is 83.75%

Also, the algorithm copes well with diverse situations like presence of noise and occlusion, variable speed, inexact direction of motion, variable distances of subject and low resolution videos. However, with better segmentation the results can be further improved.

Further, that algorithm has been run on 320 different videos for 400 times on a system for estimating its suitability in terms of time consumption. The algorithm has been tested over following simulation platform: It has Intel Xeon processor with 4 parallel processing core of frequency 2.4 GHz each and 8 threads. It has 8 GB of RAM and 12 MB of cache memory and runs on 64 bit instruction set. The average time taken for direction estimation is 1.04 sec. with a maximum and minimum time consumptions of 1.97 sec and 0.62 sec.

REFERENCES

[1] P. Kulkarni, D. Ganesan, and P. Shenoy. The case for multi–tier camera sensor networks. In *Proceedings of the international workshop on Network and operating systems support for digital audio and video*, NOSSDAV, pages 141–146, 2005.
[2] P. Kulkarni, D. Ganesan, P. Shenoy, and Q. Lu. Senseye: a multi-tier camera sensor network. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 229–238, 2005.
[3] Hartley R. and Zisserman A. *Multiple view geometry in computer vision*. Cambridge Univ. Press, 2 edition, 2004.
[4] Davis L., Borovikov E., Cutler R., Harwood D., and Horprasert T. Multi-perspective analysis of human action. In *Proceedings of 3rd International Workshop on Cooperative Distributed Vision*, Kyoto, Japan, 1999.
[5] Kanade T., Rander P., and Narayanan P.J. Virtualized reality: Constructing virtual worlds from real scenes. 4(1):34–47, 1997.
[6] Piovan G., Shames I., Fidan B., Bullo F., and Anderson B.D.O. On frame and orientation localization for relative sensing networks. In *47th IEEE Conference on Decision and Control*, CDC08, pages 2326–2331. IEEE, 2008.
[7] Taylor C., Rahimi A., Bachrach J., Shrobe H., and Grue A . Simultaneous localization, calibration, and tracking in an ad hoc sensor network. In *5th International Conference on Information Processing in Sensor Networks*, IPSN. ACM, 2006.
[8] CASIA Database. http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp.