International Journal of Electrical and Electronics Research ISSN 2348-6988 (online)
Vol. 3, Issue 1, pp: (296-301), Month: January - March 2015, Available at: www.researchpublish.com

Target Detection and Tracking in Real Time with Multi frame Marked Point Process Model

N. Mohanambigai[1], M. Arul pugazhendhi[2]

[1]B.E Student, [2]Assistant Professor, Electronics and Communication Engineering,
IFET college of Engineering, Villupuram.

Abstract: **In this paper, the target is considered as human, we propose a multiframe marked shape model based human detection and tracking. Geometric features are measured through principle component analysis features are reduced to half of its ratio and finally a realtime decision making system detects human and by representing region of interest through bounding box followed by tracking in order to speed up the process. Optical tracking method is used for tracking purpose. To shrink or expand the bounding box as required, the aforementioned sample rectangles generated during the tracking process and their associated confidence ratings which indicate the likelihood of a rectangle containing the region of interest Finally our proposed method, shape based target detection algorithm successfully tracking a human target walking across the view of the camera.**

Keywords: **Human Detection, Tracking, Realtime, Shape Characteristic Recognition.**

1. INTRODUCTION

Automatically detecting and tracking humans is of interest in many application areas, particularly in security and surveillance operations. Realtime imagery is particularly useful for this purpose as it can operate under any level of target illumination, and it delivers high contrast between humans and their surroundings due to temperature differences.

A number of tracking algorithms capable of following the path of a moving person have been presented in literature. Four state of theart, open source, tracking algorithms have been identified as being applicable to realtime data [14]. These trackers are known as MILTrack [1], Incremental Visual Tracking (IVT) [2], Tracking Learning Detection (TLD) [3], and RealTime Compressive Tracking) [4].

During the operation, MILTrack, IVT, and TLD all depend on the improvement of an appearance model of a target to account for appearance or lighting changes. Thus, if a target's shape is altered during the tracking process, each and every algorithm is able to learn the new form of the object of interest while operational.

Motion flow tracking incorporates a very sparse measurement matrix to efficiently extract the features for the updating of an appearance model. Compression arises during the application of a random matrix in its feature extraction process, which effectively reduces the dimensionality – and therefore the size – of the data extracted. It is shown in [4] that initializing trajectory generally outperforms other state of the art trackers – including MILTrack and TLD – when applied to a set of benchmark testing data.

While useful for single person tracking, the above algorithms were found to be susceptible to target loss when encountering occlusion and situations involving multiple humans in realtime video. Furthermore, an automated tracker initialization stage was desirable, as current trackers typically require an initial bounding box to be manually entered.

We have observed that certain fluctuations in the intensity of pixels within realtime image sequences can be indicative of the presence of humans and their associated movements. In this paper, we present a robust method for detecting and tracking individuals within realtime video sequences using both spatial and shape features of a targeted human. The proposed technique builds upon the motion flow tracking tracking algorithm, due to its demonstrated reliability. A standalone human detection algorithm based on Support Vector Machine (Matching Hypothesis) [5] classification of

International Journal of Electrical and Electronics Research ISSN 2348-6988 (online)
Vol. 3, Issue 1, pp: (296-301), Month: January - March 2015, Available at: www.researchpublish.com

shape features automates the initialisation procedure of the growing trajectory tracking algorithm, and modifications are made to this algorithm to accommodate simultaneous spatial and shape feature evaluation during tracking. While other methods of spatio shape tracking and detection have been investigated for use in realtime imagery [69], it is believed that the shape signatures used in this paper introduce a new concept.

A specially recorded dataset containing realtime videos of multiple humans was captured for testing purposes.

Results obtained show the successful automation of the tracker's initialization, and that the inclusion of shape feature recognition offers improvements in human tracking results, suggesting that such techniques may be viable in future tracking efforts.

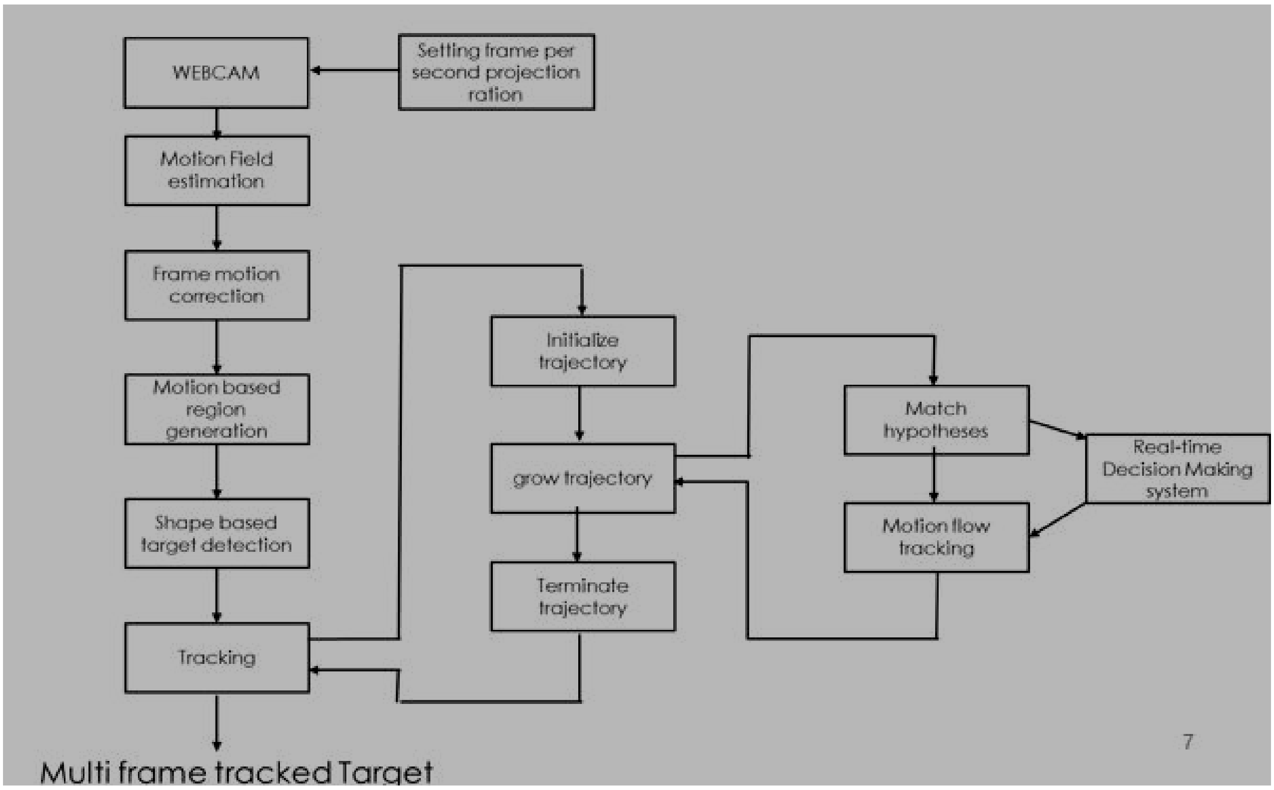

Fig 1. Real Time Compressive Tracking

2. DESCRIPTION OF DATA

The input data used for the testing of the existing and developed human tracking algorithm was in the form of 12, 16bit realtime videos; all had a resolution of 640 x512 pixels, and an effective frame rate of 25 frames per second. These videos were tailored to depict different scenarios in an attempt to emulate possible situations in the real world and to give images of humans at various ranges.

Simple two dimensional matrices sufficed to store individual frames of a video sequence, but for shape analysis, multiple frames were stored simultaneously in three dimensional matrices. In this case, the third dimension contained time varying realtime data and was of a size determined by the length of the sequence being analysed.

3. REALTIME COMPRESSIVE TRACKING

The detailed description of the operation of motion field estimation is provided in [4], which is a brief summary of the main components is relevant to the modifications described in this paper is provided here to minimize the need for cross referencing.

The initialisation stage of the frame motion correction algorithm requires a bounding box – which surrounds the object to be tracked in subsequent frames – to be defined by a preexisting text file. This file contains the location and size of the box to be used. Once initialized, the algorithm begins tracking the object within the specified box.

During tracking, the motion based region generation algorithm determines a number of sample rectangles defined around th e location of the current bounding box. When transitioning between frames, the sample rectangle containing features that most closely match previously observed characteristics of the object being tracked is chosen as the new bounding box. The tracking algorithm then updates the characteristics of the tracked object by analysing sample rectangles near to the current bounding box, and learns features of the surrounding background by recording the contents of rectangles placed fur illustrates the placement of sample rectangles around an original bounding box.

For each sample rectangle, a feature extraction template is applied and a number of subsamples obtained. This process is shown by where the smaller internal rectangles are representative of the subsamples taken. Following this, spatial data can be extracted from the subsamples within each rectangle and used to form a feature list that defines the spatial characteristics of the overall sample. Demonstrates this procedure. Newly acquired spatial samples are compared with previous positive (human) and negative (nonhuman) samples and allocated a confidence rating, which indicates how well they match the target of interest.

The shape based target detection algorithm successfully tracking a human target walking across the view of the camera. The algorithm's need for an initial bounding box requires prior knowledge of the target's location or a manual input stage. Thus, an algorithm to automatically detect the presence of humans and present the tracker with an initial bounding box is desirable. Furthermore, while the initialize trajectory algorithm worked well for simple human tracking applications, it was prone to failure when the target human was occluded by another in a scenario identified as a 'crossover'. In Section 5, we demonstrate that by including time varying realtime data in the list of features computed by the hypothesis tracker, this issue can be overcome.

4. HUMAN DETECTION ALGORITHM

When analyzing pixel intensities over time, pixels containing humans typically display greater variance than those containing the relatively static background, as demonstrates. It was therefore considered viable that a human detection algorithm could be developed that specifically identifies any time varying signatures similar to those of a human.

Principal Component Analysis (PCA) [10] is a mathematical data processing technique that can be employed to reduce the dimensionality – and therefore the size – of a data set prior to further processing.

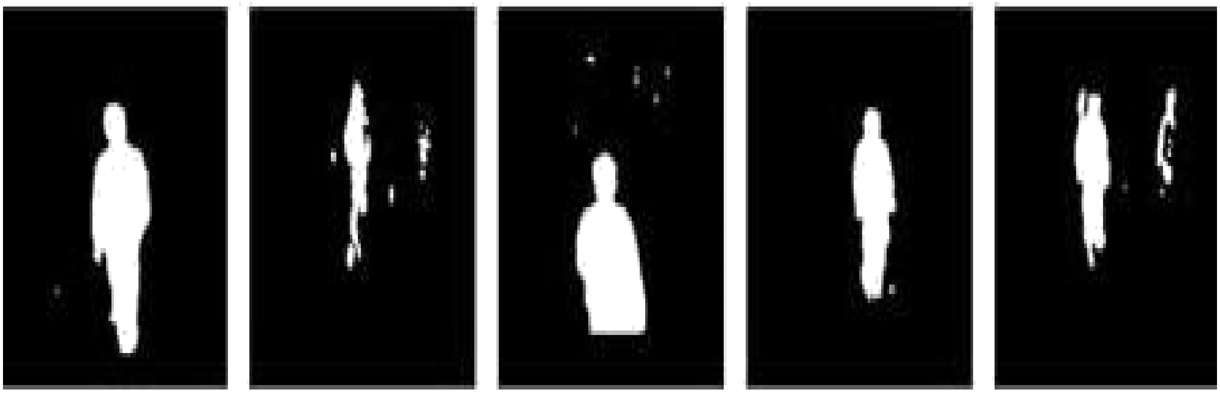

Fig 2. Optical flow Tracking.

Here, PCA can be used on time varying pixel intensities extracted from video sequences to reduce the dimensionality of the variance of the pixel intensities over time. This process enables the separation of dissimilar video features – such as humans and static background objects – based on their respective variances. The separation observed when plotting the first two principal components against each other following the application of PCA to the timevarying pixel intensities.

Matching hypothesis are a form of supervised learning binary classifier that can be trained to recognise the characteristics of positive and negative features – in this case, human and nonhuman traits – and subsequently be deployed on unseen data sets where they can detect the presence of these features [5]. Their initial training stage uses data of known class and chosen parameters to determine the orientation and placement of support vectors, which define the boundary that separates the two classes in the given spatial dimensions.

For human detection in realtime video, PCA can be used to reduce the dimensionality of datasets prior to their input to the Matching hypothesis for training or classification purposes. As the classification of each pixel within a frame is a time consuming process, steps must be taken to reduce the number of pixels to be classified.

One way to reduce the number of pixels to be processed is to apply motion detection to eliminate pixels exhibiting insufficient movement from the classification process. This results in a large portion of the background being ignored. For a sequence of frames, the mean intensities of each pixel can be calculated. These average values can then be subtracted from corresponding pixel intensities in the initial frame. Subsequent thresholding of the resulting image can then eliminate pixels with low variance.

Blob detection can also be used to limit the pixels to be analyzed to those contained within regions of similar intensity – it is assumed that any humans present in a realtime video will give rise to such regions. The Maximally Stable Extremal Regions (MSER) [11] method of blob detection was used for this application. MSER blob detection operates by incrementing through an image's intensity profile by a set threshold delta; changes in size of different regions within the image as this occurs are calculated and used to judge if each region is stable or not. Stable regions are then classified as blobs within the image. The maximum possible area of blobs can be restricted to the maximum expected size of humans within the image sequences used. An Matching hypothesis is then applied to the blobs in the image to determine if they are human in origin.

When applying this algorithm to a video sequence containing two humans – one running towards the camera and another far in the distance – the algorithm gave the result shown. It can be seen that the method has correctly placed red outlines around the locations of the two humans. For the initialization of the trajectory algorithm, the intensity or number of pixels contained within any outlined areas can be used to specify the characteristics of the human to be tracked. For example, if tracking of a nearby human is preferential, the area containing the largest number of pixels can be automatically selected by the algorithm as the initial bounding box for the tracker.

5. MODIFIED REALTIME COMPRESSIVE TRACKING

We propose in this paper a Modified algorithm which uses extracted shape and spatial features to track humans in realtime video. The extraction of these features is enabled by the modification of the existing feature extraction functionality of the Realtime decision making algorithm. Once extracted, both spatial and time varying realtime signatures can be used to form a feature list that defines the characteristics of a sample.

For comparison of shape features in Realtime decision making, additional functionality must be implemented. The technique of crosscorrelation [12] is considered viable for these purposes. Of particular interest is the cross correlation coefficient, which is a measure of the strength of relationship between two normalized datasets. Datasets that progress in an identical fashion incur a cross correlation coefficient of 1, while datasets that progress in an opposite fashion result in a value of 1. A cross correlation coefficient of zero implies that there is no relationship between the tested datasets.

Through determination of the cross correlation coefficients linking newly taken samples with previously recorded positive and negative samples, the direct comparison of time varying signatures are possible. Following cross correlation of shape features in successive frames, a confidence rating can again be associated with each potential bounding box to quantify its shape similarity to the object of interest.

With confidence ratings for each sample available from both spatial and shape classifiers, a simple cost function enables the generation of a new parameter. The peak value of this parameter identifies the sample rectangle that possesses the spatial and shape characteristics most similar to that of the desired target.

A decay function to weight previous shape characteristics reduces the overall impact that any erratic training samples have, while allowing a shape profile of the target to be built upon over time. Within this function, a learning rate parameter determines the speed at which the tracker adapts to changes in the object of interest's shape features.

The inability of the Realtime decision making algorithm– and therefore the motion flow tracking algorithm – to shrink an object's bounding box as it moves further away was thought to influence its operation at longer distances. While the bounding box may contain only the tracked person upon initialisation, a large portion of the box will contain the background as this person moves away from the camera.

To shrink or expand the bounding box as required, the aforementioned sample rectangles generated during the tracking process and their associated confidence ratings – which indicate the likelihood of a rectangle containing the object of interest – can be used. For example, if multiple rectangles carry a high confidence rating while being in close proximity, it is likely that a new bounding box can be generated using a combination of their areas.

6. EXPERIMENTAL RESULTS

Initial bounding boxes generated using the human detection algorithm of Section 4 were successfully used within the motion flow tracking and Realtime decision making system trackers. Demonstrates the acquisition of an initial bounding box through human detection, and show its recognition within the tracking algorithm and the beginning of the tracking process, respectively.

The scenario shown in shows the failure of the motion flow tracking algorithm to track the correct person following a crossover. It was hypothesised that this result could be improved upon through application of the proposed spatial and shape data fusion method for tracking. During testing of this hypothesis, the initial bounding box – which was automatically generated using the human detection algorithm of Section 4 – and parameters relevant to the spatial feature recognition were kept constant.

Early results – obtained when bounding box shrinking functionality was not present – indicated that in some crossovers, the inclusion of shape data improved correct target selection, but the problem of incorrect selection was still present.

The addition of functionality to shrink the bounding box improved results considerably for both the motion flow tracking and motion flow tracking algorithms; however, motion flow tracking consistently performed well, and in some cases better than the unmodified tracker. Demonstrates the ability of both trackers to track the same person in a sequence containing a large number of crossovers at various ranges. The individual frames to allow direct comparison of the trackers' performance at stages within this sequence. Both trackers achieve a similar level of accuracy for the majority of the sequence, but in the bottom frames to acquire that trajectory has outperformed growing trajectory; at this point, motion flow tracking has tracked the correct target while trajectory has failed.

The Centre Location Error (CLE) – as used in [4] – is defined as the distance between the central locations of the tracked target and the manually labelled ground truth for a sequence. Use of the CLE enabled a quantitative comparison of the accuracy of the system with the two algorithms achieving average CLEs of 11.6 and 25.1 pixels, respectively, for the sequence.

7. CONCLUSIONS

This paper presents the successful automation of a state oft heart tracking algorithm's initialisation stage through the application of a human detection algorithm based on temporal feature recognition. The automation of a tracker's initialisation stage in this way is of great importance if the tracker is to operate autonomously.

Furthermore, the modification of the motion flow tracking algorithm, known as Realtime decision making – which accommodates simultaneous spatial and shape feature evaluation during tracking – has given promising results that suggest there may be a case for future research in this area.

(a)

(b)

Fig: 3. (a), (b) Tracking of human

REFERENCES

[1] B. Babenko, M.H. Yang, and S. B elongie, "Robust Object Tracking with Online Multiple Instance Learning," IEEE Transactions on Pattern A nalysis and Machine Intelligence, vol. 33, no. 8, pp. 16191632, August 2011.

[2] D.A. Ross, J. Lim, R.S. Lin, and M.H. Yang, "Incremental Learning for Robust Visual Tracking," International Journal of Computer Vi ion, vol. 77, no. 13, pp. 125141, May 2008.

[3] Z. Kalal, K. Mikolajczyk, an d J. Matas, "TrackingLearningDetection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 7, pp. 14091422, July 2012.

[4] K. Zhang, L. Zhang, and M .H. Yang, "RealTime Compressive Tracking," in ECCV 2012, Florence, 2012, pp. 864877.

[5] N. Cristianini and J. ShaweT aylor, An Introduction to Support Vector Machines a nd Other Kernelbased Learning Methods, Cambridge University Press, 2000.

[6] J. Han and B. Bhanu, "Human Activity Recognition in Realtime Infrared Imagery," in Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, San Diego, 2005, pp. 1724.

[7] A. Leykin, Y. Ran, and R. Ham moud, "RealtimeVisible Video Fusion for Moving Target Tracking and Pedestrian Classification," in Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, Minneapolis, 2007, pp. 18.

[8] C. Bregler, "Learning and Rec ognizing Human Dynamics in Video Sequences," in Proc. IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition, San Juan, 1997, pp. 568574.

[9] C.N. Padole and L.A. Alexand re, "Wigner Distribution based Motion Tracking of Human Beings using Thermal Imaging," in Proc. IEE E Comp. Soc. Conf. on Computer Vision and Pattern Recognition, San Francisco, 2010, pp. 914.

[10] I. Jolliffe, Principal Component Analysis, John Wiley & Sons, Ltd, 2005.

[11] J. Matas, O. Chuma, and T. Pajdlaa, "Robust widebaseline stereo from maxima lly stable extremal regions," Image and Vision Co mputing, vol. 22, no. 10, pp. 761–767, September 2004

[12] A.G. Asueroa, A. Sayagoa, a nd A.G. González, "The Correlation Coefficient: An Overview," Critical Reviews in Analytical Chemistry, vol. 36, no. 1, pp. 4159, 2006.