A Novel Stochastic Tracking Approach on Human Movement Analysis

Mohammad Alamgir Hossain
Department of Computer Science,
Jazan University,
Kingdom of Saudi Arabia

Saleh M AL-Turki, Ph.D
College of Computer Science & Information Systems,
Jazan university,
Kingdom of Saudi Arabia

Goutam Sanyal, Ph.D
Department of Computer Science & Engineering,
National Institute of Technology, Durgapur, West Bengal, India

ABSTRACT
Modern research demands more mathematical, statistical information and proves on the research topic. Moving Object tracking with mathematical and statistical approaches paves a new modus operandi stirring over the conventional methods. The use of covariance as a detector of object-based image to extract features is a proven approach. An object can be tracked by using the conventional histogram-based depiction model and it is a well known as well as a popular approach. Our proposed methodological model is build up with numerical, statistical and mathematical formulas. To implement them the standard images database collected images are considered. This mathematical model has been enriched in pursuance with covariance-chaser and in subsequently is capable enough to establish its supremacy with an eye to spawn an important algorithm leading to generate improved object-image-tracker (OIMT) method correctly by taking minimal finishing time. With the help of publicly available dataset enormous quantitative estimation is done pinpointing the efficacy of the reachable model. Our model is capable to achieve momentous speeding in human-object-tracking dynamically in an enhanced way and this method is capable to decrease the error for false-tracking in comparison with the traditional histogram-based and other approaches. It is proved that the accuracy rate based on statistical-mathematical-detection-model (SMDM) is approximately 94.97% as compared to the MDM with 94.3% and the conventional model with 89.1%.

General Terms
Image-set, region-intensity, detection-model, human-object-model, filtering algorithm, schedule algorithm.

Keywords
Human-Object, mathematical-detection-model, SMDM, OIMT, HOM, covariance-specified-region, region-of-intensity.

1. INTRODUCTION
The identification and cataloging of Human-Image based on object is a very sophisticated phenomenon aiming at providing a high-end detection-model depending on mathematical and statistical approaches. A noticeable, strong as well as straight forward algorithm is very rare for feature selection based on tracked-images. Image-object based features would be treated a good one, if; it is evident, strong and simple to work with it. Superior enriched-features must be assessed and filtered out accompanied by dexterous algorithms and rationally correct technique. However, detection as well as tracking task is very difficult as it leads to establish a better and robust model with the comparison of a diversity of interlinked tasks.

The elements such as image-color, image-noise, texture, filtering algorithm applied in human-image with the help of unrefined pixel values are very conventional choice in the field of computer-vision. However, it has been provided with the evidence that features-extraction is vigorous if the enlightenment is ever changing as compared to the images that are escorted to motions. The unprocessed pixel values are being represented based on the histogram when it is applicable on a covariance-specified-region of a human-image if the region is recognized with the help distribution-method. Tracking based on histogram-application is found in texture depiction [1, 3] and assertion their union [4] in the domain of computer vision and image processing.

The ingenuity of [5] histogram to identify still picture is applied in a variety of application. Moreover, a rapid universal corresponding histogram has been able to demonstrate its supremacy [7] in many applications.

The integral-image for getting the soul of speediness in estimation process is observed [8] in Haar-like-features extraction system.

The dominance of cascaded classifiers have proved their superiority by their schedule algorithm for paving their way to detect image, but these algorithms requires huge dataset as test data as that it can be categorized in accordance with the requirements.

Juan Pavona, Antonio Fernandez-Caballero, Jorge Gomez-Sanz and Julian J. Valencia-Jimenez projected [9] to give a unique methodical frame over surveillance system controlled by a numerous-agent method.

Takeo Kanade, Robert T Collins and Alan J, Lipton have suggested [5] to widen up a courteous, numerous-sensor video surveillance system that makes obtainable with incessantexposure over battle field areas. Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt and Lambert Wixson [6] have predictable a system for self-determining video surveillance and scrutinizing application. Their methodical measures are applied with video sensors for impulsive exposure of people and vehicles in an implicated environment.

Kim C. Ng, Hiroshi Ishiguro, Mohan Trivedi and Takushi Sogo have urbanized [10] an incorporated classification with the goal to supervise a surveillance work. Their scheme was build up with manifold Omni-directional vision sensors and was emphasized to covenant with two unequivocal surveillance tasks.

C. E. Liedtk and R. Koch [12] have reachable a progress towards the modeling of versatile 3D views like outer surface street views from a sequence of stereoscopic-pictures. Initially put in a nutshell with a standard stereoscopic-image associated with 3D model-picture in 3D geometrical shaped is produced.

Barry Brumitt, Brian Meyers Kentaro Toyama, John Krumm have wished for [15] underlining a background safeguard.

Liang Wang, Weiming Hu and Tieniu Tan have proposed [16] absolute revise on computer vision based on human movement. They have drained on three main issues reliably related to general people motion analysis scheme: people detection, people tracking and their activity perceptivity.

Anthony R. Dick and Michael J. Brooks proposed [17] a programmed visual surveillance system. Their thesis paper rendering major improvement in the field of visual surveillance and explain the prime dilemma in automatic video surveillance.

Sebastien Marcel, Christopher Mc Cool, Norman Poh, Chi Ho Chan, Albert Ali Salah and Nicholas Costen [16] have proposed an opinion of a person's exclusivity authentication using facial video data. Their method dealt with a variety set of presumption, along with feature pre-processing deviation. They have proposed the outcome of inauspicious position to access information, for query selection, to build template for video-to-video face verification.

2. COVARIANCE AS TRACKER ON HUMAN-IMAGE

The tracking tactic is being represented in the subsequent steps:

With the help of a set of frames, an absolute human-image is constructed. While capturing an image-region there was an outlining for extracting the features of the images based on covariance-matrix. Now from the extracted frames, it was isolated some selected area with smallest covariance distances and assigned them to their feasible positions. To adjust the divergences, it was proposed to create and keep a covariance-matrix set with a set of multiple samples image-set.

2.1 Mathematical Computation on Integral-Image

The region based intensity of covariance matrix can be carried out depending on the Integral images. This leads to find a midway point of two images. The intensity of an Integral-image can depict as:

$$(x', y') = \sum_{x<x', y<y'} u(x, y) \quad ------(1)$$

Based on equation (1) the outline of a specified rectangular segment can be calculated in fixed time schedule stochastically. The Integral images could be applied in wide way for fast calculation if histogram-based concept is applied based on region-of-intensity of interest. The $(\boldsymbol{p}, \boldsymbol{q})\boldsymbol{th}$ module of the of the dimension matrix based on covariance as illustrated above can stand for as

$$P_{r(a,b)} = \frac{1}{n-1} \sum_{i=0}^{n} \left(v_{i\,(a)} - P_{(a)}\right) - \left(v_{i\,(b)} - P_{(b)}\right) ----(2)$$

Rearranging it the equation (3) can be adopted as:

$$P_r\,(a, b) = \frac{1}{b-1}\left[\sum_{i=1}^{n} v_i(b) v_i(b) - \frac{1}{n}\sum_{i=1}^{n} u_i(a) \sum_{i=1}^{n} v_i(b)\right] --(3)$$

The covariance surrounded on the intensity-of-a-rectangular area R, it is required to calculate the all elements of their features,µ(p) where p=1…i, and make summation preceded by multiplication with more than one dimensional features, say, $\{\boldsymbol{v}(\boldsymbol{a}), \boldsymbol{v}(\boldsymbol{b})\}\ \boldsymbol{a}, \boldsymbol{b} = \boldsymbol{1} \ldots . \boldsymbol{n}$. However, it can build Integral-image for v(a) and $\{\boldsymbol{v}(\boldsymbol{a}), \boldsymbol{v}(\boldsymbol{b})\}$.
Let be an Integral images M and the adapted equation as:

$$M(x', y', a) = \sum_{x<x', y<y'} \int (x, y, a) \qquad a = 1 \ldots i \quad --- (4)$$

and M can be customized with four parameters as underneath in the next order.

$$\theta(x', y', a, b) = \sum_{x<x', y<y'} \int (x, y, a) \int (x, y, b) \quad a, b = 1 \ldots i - (5)$$

In the above notation, $M(x, y)$ is a vector with proper dimension and $Q(, a, b)$ is a matrix with i × i dimension

Let there be a rectangular shaped region $(x', y', x", y")$, where (x', y') is in the west-north and $(x", y")$ is on the south-east coordinate. Now the covariance based region of interest is bounded by (1, 1) and (x', y'). It leads to the customized equation:

$$P_T(1{,}1; x', y') = \frac{1}{i-1}\left[\theta_{x',y'} - \frac{1}{n} M_{x',y'}\, M^t_{x',y'}\right] --- (6)$$

With a little alteration, the following equation of covariance-region-of-interest $R(x', y', x", y")$ can be achieved.

$$\begin{aligned} P_T(x', y'; x", y") = \frac{1}{n-1}\Big[&\theta_{x',y"} + \theta_{x',y'} - \theta_{x",y'} \\ &- \theta_{x',y'} \frac{1}{n}\big(M_{x',y"} + M_{x',y'} - M_{x",y'} \\ &- M\theta_{x',y"}\big)\big(M_{x',y"} + M_{x',y'} - M_{x",y'} \\ &- M\theta_{x',y"}\big)^t\Big] ----(7) \end{aligned}$$

Where $n = (x" - x')(y" - y')$.

2.2 Obtaining Covariance Matrix with Image-Space

The Euclidean-space measurement scheme based on covariance-matrices may not always be applicable. The bordering vicinity events which have been used in the subsequent part only stress a technique to measure a routine distances between two point. The distance computation tactic of the two covariance matrices can be gained as:

$$X(m_1, m_2) = \left(\sum_{a=1}^{n} Pn^2\, u_a\,(m_1, m_2)\right)^{\frac{1}{2}} ----(8)$$

Where $\{u_a\,(m_1, m_2)\}_{a=1\ldots\ldots i}$, be the random Eigen-values of a_1 *and* a_2 evaluate from the equation below.

$$\theta_a m_1 x_p - m_2 x_a = 0 \quad a = 1 \ldots . i \quad --- (9)$$

and $x_a \neq 0$ be the random Eigen-vectors. Here the distance measured by gratifies the optimistic discrete symmetric matrices m_1 and m_2.

The amplification of the Eigen-value computation aiming at acquiring a degree of three arithmetic approach for to measure distance. This method is better than histogram-base method.

2.3 An Appropriate Alike Method

With the help of covariance the region-intensity-of-interest of an image can be gained for measuring distances amongst the matrices. The covariance-matrices-space escorted by intensity is dissimilar to vector space. Only arithmetic operation is incomplete to evaluate the distances for region-of-interest between two matrices. As the features of vector-space would not be sufficient enough to do the same so it is necessary to use the equation below as that discrepancy can be removed:

$$Q(m_a, m_b) = (\sum_{a=1}^{n} Pn^2 \, \lambda_a(m_a, m_b))^{1_2} ----(10)$$

Where the random Eigen values is used as a parameter of the equation (11). The most correct alike finding of covariance-region-of-interest classifies the needed image in the current image-frame.

3. IMAGE RECOGNITION

Image finding phase starts with the acceptance of input images. The image is considered as an object. So it was proposed to take images stochastically. Here it is proposed to consider the pixels position $(m, n)$with respect the color values (RGB). It is proposed to take more parameters to obtain the intensity-based-region of x and y. The requisite features can be acquired from the equation (11).

$$\lambda(x, y) = \left[xyR(x,y)G(x,y)B(x,y) \left|\frac{\delta\lambda(x,y)}{\delta x}\right| \left|\frac{\delta\lambda(x,y)}{\delta y}\right| \left|\frac{\delta^2\lambda(x,y)}{\delta x^2}\right| \left|\frac{\delta^2\lambda(x,y)}{\delta y^2}\right|\right]^t --- (11)$$

Where R=Red, G=Green, B=Blue color and λ is the covariance-intensive-region.

3.1 Covariance Matrix to Detect Image

The variance of pixel position (x, y) is identical for the entire regions of the same size; they are at rest vital since their correlation is calculated with other features based on covariance-matrix. The covariance-matrices are generated at random from the input images and then the image features are extracted based on intensity-region. It is proposed a huge calculation for finding the covariance of the entire region from the input image. With the consideration of stochastic input images the covariance-region of interest can be analyzed very fast. If the size with respect alters then their variance as well as covariance will be changing so we took the normalized values of those images.

Firstly, it is accepted alike spots with their scales. Secondly the same is restated with the help of the covariance-matrices. Here the formula (13) is used to know the difference of the source human-image with our object-model-image.

$$\lambda(Im, t) = \frac{min}{b} \left[\sum_a^s \lambda(m_a^{Im}, m_a^t) - \lambda(m_b^{Im}, m_b^t) \right] --(12)$$

where m_a^{Im} and m_a^t *be* the source and object-image covariance respectively. The covariance-region with the smallest difference accepted as 'alike' covariance-region.

3.2 Covariance as Feature Extractor

Feature extraction by covariance-tracker has accelerated in the proposed application. This tracking output leads to attain the target location. This is possible if a scan the whole image-set is done. To reduce the calculation time the histograms of the Integral-image and a modeled simulated image-tracker is proposed. Hence piercing the whole image may lead to imprecision that may be unfocused by troublesome from equivalent image from the panorama-images. An algorithm anticipated the spot of the tracked-image from frame to frame changes step by step. To stop it an optimized-techniques to attain a strong and novel tracking method is proposed to extract features based covariance-of-region of interest.

4. CONTACT BASED OBJECT TRACKING

A person acquaints an object-image in his/her application can help to amend the image-likelihood on his human-object-model (HOM). It can't explain all parts of the experimental silhouette. Hence if the object is implicitly signified, it can be helpful in the subsequent course tracking cycle. The pose of any object-image can be occupied in two distinct manners as below.

a) With exposition non-human fore frontal areas in the likelihood assessment;

b) With restricting the human-body poses measurement in the phase of chronological update.

The primary assumption is that the grappled object-image is simple to track. As the object-image has minor degrees of choice so it is easier to represent. It rarely consider the object-tracking dilemma in this concerned phase but rather stressing on the phenomenon in what way the object pose helps to the tracking system.

4.1 Human-Image-Object Likelihood

In a nutshell, the proposed scheme is on the verge of achieving a human-object image tracking system. A conversion covariance-matrix is gained with the help of global coordinated system. This possibly helps to attain the global point of object-surface-position. The object-surface can thus be predictable onto the image plane portraying a silhouette along with the human-object-model. This enriched silhouette estimation can be occupied in the human-object tracking likelihood method.

4.2 Human-Object Contact Constraints

Human-images are taken as an object where all possible points on the human image should be presented as an object. Additionally, it is per metered itself to one extended object-image. It is somewhat easy to enlarge to more object-images. With one lengthened object-image, the probable human-object contacts are:

1. No object-image contact,

2. Right hand-object-image contact,

3. Left hand-object-image contact,

4. Both hands-object-image contacts.

5. IMAGE CLASSIFICATION

Numerous sophisticated techniques are available for image classification to point the centre-point of image and then from it the features are gained. The image features is built from the gained result of the sifted image-pixels values. For piercing an image the standard tactic is the k-means-clustering scheme. The covariance-matrix accompanied by with time series helps us to identify the images dynamically. This technique is expensive as the images are to be in the large database with huge space capacity.

5.1 Object Classifier

A robust tactic for image classification quandary without taking the whole object-image is proposed here. It is necessary to take out several features from every pixel. Here it is necessary to find out image based on region-intensities and as much as order of derivatives possible with respect to the variables. The following equation is proposed for feature classification:

$$Q(x,y) = \left[\lambda(m,n)\left|\frac{\delta\lambda(x,y)}{\delta x}\right|\left|\frac{\delta\lambda(x,y)}{\delta y}\right|\left|\frac{\delta^2\lambda(x,y)}{\delta x^2}\right|\left|\frac{\delta^2\lambda(x,y)}{\delta y^2}\right|\right] - - - - - (13)$$

The images chosen randomly from image-sets and the sizes will be 16×16 and 256×256. Images are taken to calculate the every region along with covariance-matrix. Each basis image is then treated as the covariance-matrices and we propose a train image from every image cluster. The above mentioned process is recurring if needed.

5.2 Experimental Result

The experimental is based on Brodatz image dataset computing more than 81 images and many other images taken by Nikon P 500 camera as a the dataset. Each image is farther separated into sub-images so as to construct an available trained image set for our future test image set. With the help of the proposed methodology the resultant accuracy is 94.97% based on our image-dataset**.**

Comparing 55 images set with the train image set that are randomly selected based on covariance-matrix for every image gained result is 94.97%. It is better compared to MDM and stochastic object-image (SOI) selection procedure. Here some measured for essential features namely-intensity-of-region-of-interest (IRI), intensity-gradient interest (IGI) and Laplacian-Method (LM) is used to achieve better performance. Thus we have gained the following tabular result:

Table 1. Experimental Results in the following table.

	IRI	IGI	LM	SOI	MDM	SMDM
RESULT	81.73	85.31	89.51	91.43	94.30	94.97

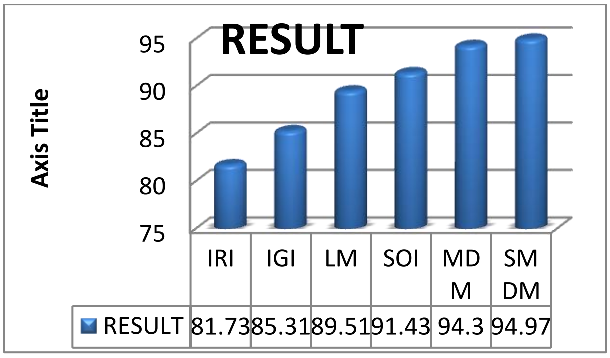

Fig: Resultant Chart

6. CONCLUSION

Numerous calculation and computation based on image dataset is done. It is our expectation to enrich the proposed methods with the consideration dynamic circumstances and multidimensional view of image-sets. Some sophisticated noise detection algorithm is necessary to embed in the subsequent research in the future.

7. REFERENCES

[1] Comaniciu, D., Ramesh, V., Meer, P.: Real-time tracking of non-rigid objects using mean shift. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC. Volume 1. (2000) 142–149.

[2] Viola, P., Jones, M.: "Rapid object detection using a boosted cascade of simple features". In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, HI. Volume 1. (2001) 511–518.

[3] Comaniciu, D., Ramesh, V., Meer, P.: Real-time tracking of non-rigid objects using mean shift. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC. Volume 1. (2000) 142–149.

[4] Viola, P., Jones, M.: "Rapid object detection using a boosted cascade of simple features". In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Kauai, HI. Volume 1. (2001) 511–518.

[5] Comaniciu, D., Ramesh, V., Meer, P.: Real-time tracking of non-rigid objects using mean shift. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Hilton Head, SC. Volume 1. (2000) 142–149.

[6] Porikli, F.: Integral histogram: A fast way to extract histograms in cartesian spaces.In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Diego,CA. Volume 1. (2005) 829 – 836.

[7] Takeo Kanade, Robert T Collins ,Alan J, Lipton, "Advances in Cooperative Multi-Sensor Video Surveillance", Peter Burt, Lambert Wixson - Proceedings of DARPA Image Understanding Workshop , 1998.

[8] N Leung, T., Malik, J.: Representing and recognizing the visual appearance of materials using three-dimensional textons. Intl. J. of Comp. Vision 43 (2001) 29–44. 2001 Kluwer Academic Publishers. Manufactured in The Netherlands.

[9] Robert T. Collins, Alan J. Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt and Lambert Wixson, "A System for Video Surveillance and Monitoring" CMU-RI-TR-00-12

citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.73... pdf.

[10] Varma, M., Zisserman, A.: Statistical approaches to material classification. In: Proc. European Conf. on Computer Vision, Copehagen, Denmark. (2002), IEEE PAMI, VOL. X, NO. Y, APRIL 2006.

[11] Juan Pavona, Jorge Gomez-Sanz, Antonio Fernandez-Caballero, Julian J. Valencia-Jimenez "Development of intelligent multisensor surveillance systems with agents " Elsever,Robotics and Autonomous Systems 55 (2007) 892–903 .

[12] Kim C. Ng, Hiroshi Ishiguro, Mohan Trivedi, Takushi Sogo "An Integrated Surveillance System—Human Tracking and View Synthesis using Multiple Omni-Directional Vision Sensors ",Image and Vision Computing Journal, June 2002.

[13] Georgescu, B., Meer, P.: Point matching under large image deformations and illumination changes. IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 26, NO. 6, JUNE 2004-Pgs: 674–688.

[14] C. E. Liedtk and R. Koch, "MODELLING OF 3D SCENES FROM THE ANALYSIS OF STEREOSCOPIC IMAGE SEQUENCES "citeseerx.ist.psu.edu/viewdoc/download.

[15] Ismail Oner Sebe, Jinhui Hu, Suya You , Ulrich Neumann, "3D Video Surveillance with Augmented Virtual Environments ",ACM SIGMM 2003 Workshop on Video Surveillance, pp 107-112, Berkeley, California, November 2003, Berkeley,CA, pp.107-112.

[16] Kentaro Toyama, John Krumm, Barry Brumitt, Brian Meyers," Wallflower: Principles and Practice of Background Maintenance ",Computer Vision, 1999. The Proceedings of the Seventh IEEE,Vol: 1 Dig Obj Iden: 10.1109/ICCV.1999.791228 ,1999 , Pg: 255 – 261.

[17] Liang Wang, Weiming Hu, Tieniu Tan, " Recent Developments in Human Motion Analysis ",National Laboratory of Pattern Recognition Institute of Automation, Chinese Academy of Sciences, Beijing, P. R. China, 100080.

[18] Anthony R. Dick, Michael J. Brooks ," Issues in Automated Visual Surveillance", New Scientist, 12 July 2003, page 4.

[19] Norman Poh, Chi Ho Chan, Josef Kittler, S´ebastien Marcel, Christopher Mc Cool, Albert Ali Salah,Nicholas Costen," Face Video Competition ",M. Tistarelli and M.S. Nixon (Eds.): ICB 2009, LNCS 5558, pp. 715–724, 2009. Springer-Verlag Berlin Heidelberg 2009.

[20] Norman Poh, Chi Ho Chan, Josef Kittler, Sébastien Marcel, Christopher Mc Cool, Enrique Argones Rúa, Jose Luis Alba Castro, Mauricio Villegas, Roberto Paredes,Vitomir Struc, Nikola Pavesic, Albert Ali Salah, Hui Fang, Nicholas Costen, "An Evaluation of Video-to-Video Face Verification ",IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 5, NO. 4, DECEMBER 2010.