A Stochastic Statistical Approach for Tracking Human Activity

Md Alamgir Hossain¹ and Prof (Dr) Goutam Sanyal ²

¹Department of CSE, NIT, Durgapur, West Bengal, India
alamgir_sim@yahoo.com

²Department of CSE, NIT, Durgapur, West Bengal, India
nitgsanyal@gmail.com

Abstract

Modern research for tracking Human-Object by means of various statistical and mathematical shows a new method enriching the conventional methods. The use covariance considering the images in matrix form for collecting features is a proven approach for tracking object/image in preference to the usual histogram based object/image depiction replica frequently used in well-liked methodology. In this paper we propose few robust statistical approaches for tracking an object/image with the help of mathematical model. The improved mathematical model followed by covariance-chaser is capable to prove its superiority with an eye to generate a prominent algorithm leading to generate improved object/image tracking accurateness by decreasing the completing time. These mathematical models are evaluated to a covariance-chaser and the popular histogram-supportive tracker method. With the help of publicly available dataset a huge quantitative assessment is done pinpointing the effectiveness of the obtainable model. Our model is capable to achieve momentous speeding in human-object tracking dynamically in the better way and is capable to decrease the error for false-tracking compared to the earlier histogram-based and other approaches. It is proved that the accuracy rate based on mathematical-detection-model (MDM) is approximately 94.3% as compared to the conventional model with 89.1 percent.

Keywords

Mathematical-detection-model; Human-Object; covariance-specified-region; region-of-intensity

1. Introduction

The significance of Human-Object identification and their cataloguing is a very sophisticated phenomenon keeping our goal to give birth of a high-end detection-model based on mathematical and statistical approaches. A noticeable, strong and straightforward algorithm is very rare for feature selection depending on the tracked image. Human-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 and tracking tasks are very difficult if a good and robust model could not be established with the comparison of a diversity of interlinked tasks.

The gadget such as image-colour, image-noise, its texture, filtering algorithm applied in human-image with the help of unrefined pixel values are the conventional choice for in the field of computer-vision at the primary phase of tracking method. 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.

The resourcefulness of [2] histogram for tracking of non-rigid human-object was applied in various application. Moreover, a rapid universal corresponding histogram has been able to demonstrate its supremacy [3].

Besides the tracking based on histogram-application is found in texture portrayal [8, 9] and pronouncement their association [10] in the domain of image-processing.

The integral-image for receiving the essence of speediness in calculation process is noticed [11] in Haar-like-features withdrawal technique.

The superiority of cascaded classifiers have proved their dominancy by their routine algorithm for paving their way in image detection, but these algorithms demand a lot number of data to read out and categorize the human-object.

Juan Pavona, Antonio Fernandez-Caballero, Jorge Gomez-Sanz, Julian J. Valencia-Jimenez have anticipated [4] to give a novel systematic frame of surveillance system controlled by a numerous-agent technique. They have obligatory the idea of be appropriating an agent-orientated method, which is pondered with concrete situation.

Takeo Kanade, Robert T Collins and Alan J, Lipton have recommended [5] to enlarge up a willing, numerous-sensor video surveillance system that builds accessible with incessant exposure over battle field areas. Hironobu Fujiyoshi, David T, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt and Lambert Wixon, David Duggins, Yanghai T [6] have imagined a system for self-determining Video surveillance application. They have measured methodically and tried to apply it with video sensors for impetuous exposure of people and vehicles in an mixed up environment.

Barry Brumitt, Brian Meyers Kentaro Toyama, John Krumm have wished-for [12] to underline on background safeguarding.

C. E. Liedtk and R. Koch [13] have reachable a progress towards the modelling 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 wrought is produced..

Kim C. Ng, Hiroshi Ishiguro, Mohan Trivedi and Takushi Sogo have urbanized [5] 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. The number one is to track and silhouette human-object activity and the second one is to amalgamation the implicit interpretation of people for keeping an eye to them for their surrounding involvement.

Anthony R. Dick and Michael J. Brooks proposed [17] an programmed visual surveillance to get free the operative from the weigh down by putting in a nutshell the pertinent software in their tactic that can study surveillance activities regularly. Their thesis paper expose major advancement in the field of visual surveillance and solves the prime troubles in automatic video surveillance.
Liang Wang, Weiming Hu and Tieniu Tan have proposed [15] absolute revise on computer vision based on human being movement study. They have frizzled on three main issues unswervingly related to general people motion analysis scheme: people detection, people tracking and their activity perceptivity.

Sebastien Marcel, Christopher Mc Cool, Norman Poh, Chi Ho Chan, and all [16] have proposed an opinion of a person’s rareness facial video data and tried to verify it using their methodological approach. Their method dealt with a variety set of presumption, along with feature portrayal and pre-processing divergences. They have proposed the consequence of inauspicious circumstances to admittance information, for query selection, to build template for video-to-video face authentication.

2. COVARIANCE FOR HUMAN-OBJECT TRACKER

The tracking tactic is being represented in the subsequent steps:

Taking into consideration frame by frame, we have tried to construct an absolute human-image. While capturing an image-region we are to outline the image-based-covariance-matrix for extracting the features of the images. Now from our frame, we are isolating area with smallest covariance distances and assign them to their feasible positions. To adjust the divergences, we are creating and keeping a covariance-matrix set with a set of multiple samples image-set.

2.1. Mathematical Approach to Compute an Integral-Image

The regions-based on covariance of interest can be carried out depending on the Integral images leading to an intermediate point between two images accelerating speedy computational time. The intensity of an Integral-image can be depicted by the equation:

\[ (a', b') = \sum_{a < a', b < b'} \mu(a, b) \quad (i) \]

Using the above mentioned equation the outline of a specified rectangular section can be measured in fixed time schedule dynamically. The Integral images could be applied in extensive way for fast calculation if histogram-based concept is applied depending on based on region-of-intensity of interest. The (m, n)th module of the of the measurement matrix based on covariance as depicted above can be represented as

\[ p_r(x, y) = \frac{1}{q-1} \sum_{i=1}^{q} (v_r(x) - \rho(x))(v_r(y) - \rho(y)) \quad (ii) \]

By reshuffling it in a stretched way we can modify the equation (iii) as:

\[ p_r(x, y) = \frac{1}{y-1} \left[ \sum_{i=1}^{q} v_r(y)v_r(y) - \frac{1}{q} \sum_{i=1}^{q} u_r(x) \sum_{i=1}^{q} v_r(y) \right] \quad (iii) \]

To find the covariance encircled depending on the intensity-of-a-rectangular area R, we need to compute the every dimension of their features, u(m)m=1…c, and make summation preceded by multiplication with more than one dimensional features, say, \{v(x), v(y)\} x, y=1….q. However we can build w + w2 Integral-images for v(x) and multiplication of {v(x), v(y)}.
Let be an Integral images I and we get the adapted equation as:

\[ \text{IMG}(a, b', x) = \sum_{a < a', b < b'} \Pi(a, b, x) \quad x = 1 \ldots w \quad (iv) \]

and I can be modified with four parameters as below in the next order:

\[ \psi(a, b', x, y) = \sum_{a < a', b < b'} \Pi(a, b, x) \Pi(a, b, y) \quad x, y = 1 \ldots w \quad (v) \]

In our above notation, I(a, b) is a vector with proper dimension and \( \psi(x, y) \) is a matrix with \( w \times w \) dimension. The mentioned vector and matrix can be represented as:

\[
\begin{bmatrix}
I_{a, b} & I_{a, b} & \ldots & I_{a, b} \\
I_{a, b} & I_{a, b} & \ldots & I_{a, b} \\
\vdots & \vdots & \ddots & \vdots \\
I_{a, b} & I_{a, b} & \ldots & I_{a, b}
\end{bmatrix} = \begin{bmatrix}
\psi(a, b, 1, 1) & \psi(a, b, 1, w) \\
\psi(a, b, w, 1) & \psi(a, b, w, w)
\end{bmatrix} = \begin{bmatrix}
\psi(a, b, 1, 1) & \ldots & \psi(a, b, 1, w) \\
\psi(a, b, w, 1) & \ldots & \psi(a, b, w, w)
\end{bmatrix} \quad (vi)
\]

Now we get \( \psi(p, q) \) as a matrix and \( w + (w^2 + w)/2 \) repetition are required to calculate both I and \( \psi \). The timing complexity to construct the Integral images is \( \text{IMG} (w2 k l) \).

Let us consider that there is a rectangular shaped region \( R(a', b'; a'', b'') \), where \( (a', b') \) is the west-north and \( (a'', b'') \) south-east coordinate. Now the region of interest based on covariance is bounded by \((1, 1)\) and \((a', b')\) leading to the modified equation:

\[
p_f(1, 1; a', b') = \frac{1}{q - 1} \left[ \psi_{a', b'} - \frac{1}{q} \text{IMG}_a \text{IMG}_b \right] \quad (vii)
\]

If we can make small number modification we can achieve the equation below based the covariance-region-of-interest \( R(a', b'; a'', b'') \):

\[
p_f(a', b'; a'', b'') = \frac{1}{q - 1} \left[ \psi_{a', b'} + \psi_{a', b''} - \psi_{a'', b'} - \psi_{a'', b''} - \frac{1}{q} \left( \text{IMG}_{a', b'} + \text{IMG}_{a', b''} - \text{IMG}_{a'', b'} - \text{IMG}_{a'', b''} \right) \right] \quad (viii)
\]

Where \( q = (a'' - a') \quad (b'' - b') \). Now we have attained the timing complexity of \( O (w2) \) by rebuilding Integral images depending on the covariance considering rectangular shaped-region-region of interest.

### 2.2. Covariance Matrix for Distance Measurement

The measurement of distance based on Euclidean-space calculating method adjoined to it is anticipated and confirmed that covariance-matrices are not based on Euclidean-space measurement technique. The neighbouring vicinity events which have been used in the consequent part only demands a technique to measure a procedural distances between two point. The distance measuring tactic of the two covariance matrices can be achieved by the equation (ix).
\[
\chi(p_1, p_2) = \sqrt{\sum_{i=1}^{q} m q i \mu_i(p_1, p_2) - - - - (ix)}
\]

where \(\{ \mu_i(p_1, p_2) \}_{i=1}^{\ldots} \) are the random Eigen-values of x1 and x2 assessed from the equation(x):

\[
\psi, p_1 a_w - p_2 a_1 = 0 \quad x = 1 \ldots w \quad - - (x)
\]

And \(ax \neq 0\) are the random Eigen-vectors. The distance measured by \(\psi\) gratify the optimistic discrete symmetric matrices p1 and p2.

The clarification of the Eigen-value calculation with the aim at acquiring a degree of three mathematical exercises depending on statistical approach for distance computation. This method is superior than histogram based method.

2.3. Finding a Match with Appropriate Module

The covariance based region-intensity-of-interest of an image can be captured for measuring distances among the matrices. The covariance-matrices-space accompanied by intensity is dissimilar to vector space. It is not any easy task to compute the distances for region-of-interest between two matrices based on computation on arithmetic formula but also we require statistical computation, too. As the features of vector-space would not be Hence we have used the equation below as that we can remove the inconsistency amid covariance-matrices-based region:

\[
\chi(p_1, p_2) = \sqrt{\sum_{i=1}^{q} m q i \psi_i(p_1, p_2) - - - - (x)}
\]

Where \(\{ \psi_i(p_1, p_2) \}_{i=1}^{\ldots} \) are the random Eigen values allowing the parameter of the equation (xi).The most correct alike matching of covariance-region-of-interest categorizes the required human image in the present image-frame.

3. HUMAN-IMAGE DETECTION

Image detection phase deals with inputting image taking it as object so that it can accept image randomly. Here we consider the pixels position (p, q) with respect the color values (RGB). We have derivate more than on to acquire the intensity-based-region allowing for the pixels a and b as the parameter. Then it is transformed to the requisite intention dimension to acquire the final features in vector-form with the help of equation (xii).

\[
\chi(a, b) = \left[ \begin{array}{c} \frac{\partial \psi}{\partial a} R(a, b) \nabla (a, b) G(a, b) B(a, b) \nabla (a, b) \\
\frac{\partial \psi}{\partial b} R(a, b) \nabla (a, b) G(a, b) B(a, b) \nabla (a, b) \end{array} \right] \left[ \begin{array}{c} \frac{\partial \psi}{\partial a} (a, b) \\
\frac{\partial \psi}{\partial b} (a, b) \end{array} \right] \left[ \begin{array}{c} \frac{\partial \psi}{\partial a} (a, b) \\
\frac{\partial \psi}{\partial b} (a, b) \end{array} \right] \nabla (x)
\]

Where \(R=\text{Red}, G=\text{Green}, B=\text{Blue color and} \psi\) is the covariance based intensive region.
3.1. Covariance used for Human-Image Identification

It is obvious that the variance of pixel position (a, b) is alike for the whole regions of the equivalent size; they are at rest imperative since their correlation is calculated with other features with the help of the covariance-matrix. We have randomly generated covariance-matrices from a human image and then the image features are extracted encircled by human-image intensity region. We have started our calculation for finding the covariance of the entire region from the input human-image. Then we search for our expected image from the intensity of region of interest with a comparable covariance matrices.

We started our search based on the stochastic input images because we want to analyse the covariance of region of interest rapidly. If the size with respect to the intensity of the same regions alters then their variance as well as covariance should changes so we have taken the normalized values of those images.

Firstly, we have recognized alike locations with their scales. Secondly we have restated the same with the help of the covariance-matrices. We have used the following formula to identify the deviation of the source human-image with our object model-image.

\[ \chi(I,t) = \sum \left[ \chi(p_i, p_j) - \chi(p_i', p_j') \right] - (xiii) \]

where \( p_i \) and \( p_i' \) are the source and object-image covariance respectively. The covariance-region with the smallest dissimilarity is considered as the identical covariance-region.

3.2. Covariance-Tracker as a Feature Extractor

Extraction of features by covariance-tracker has given an improved effect in our application. This covariance-tracker produces output leading to achieve the target location and this is possible if we can scan the whole image-set. To reduce the computation time we should find out the histograms of the Integral-image and a modelled simulated simultaneous Image-tracker is gained. The attainment of an improved similar image to make a global search is easy for numerous synchronized functions and it cannot a enhanced scaled for larger images. Moreover, piercing the entire image may lead to imprecision that may be unfocused by disrupting from equivalent image within in the panorama-image set. An algorithm used to anticipate the position of the tracked-image from frame to frame changes gradually. To prevent this we use few optimized-techniques helping to achieve a strong and novel tracking technique based on extracted features with the help of covariance-of-region of interest.

4. Human-Object Contact in Tracking

A human being grasping an object-image in his/her hand will alter the image likelihood assessment as human-object- model is unable to elucidate all parts of the experimental silhouette. Hence if the object is implicitly signified, it can be helpful in the course tracking cycle. The pose of any object/image can be engaged in two dissimilar etiquette:

a) With elucidation non-human fore frontal areas in the likelihood assessment;
b) With restricting the human-body poses assessment in the phase of chronological update.

The principal hypothesis is that the grappled object-image is simple to track. As the object-image has lesser degrees of choice and most over and again is easier to model. Now we never supplementary consider the object-tracking problem in this concerned phase but rather focusing
on the phenomenon in what way the object pose become a helpful thing for human-object tracking.

4.1. Human-Object Likelihood

Encapsulating human-object tracking methodology we are on the verge of achieving a human-object tracker with each time footstep with respect to time. In this way a conversion covariance-matrix is achieved with the help of global coordinated system. These possibly help to achieve the global point of object-surface-position. The object-surface can thus be predictable onto the image plane depicting a silhouette together with the human-object-model. This improved silhouette approximation can be engaged in the human-object tracking likelihood method.

4.2. Human-Object Contact Constraints

We have considered human-image as an object as that we can use them in our application very easily. It is obvious that if we have to depend on the environmental circumferences for collecting an image then we have to bear some extra headache for that. We additionally permeated ourselves to one lengthened object-image. It is somewhat easy to enlarge the reasoning to more object-images. With one lengthened object-image, the probable human-object contact is

a) no object-image contact,
b) right hand-object-image contact,
c) left hand-object-image contact
4) both hands-object-image contact.

5. IMAGE CLASSIFICATION

At present, a huge number of sophisticated methods for object-image classification are available to point the centre point of it and then the object-image features are gained from it. The object-image feature is built from the gained result of the sifted the object-image features of the gained-pixels. For penetrating an object-image the usual tactic is to take as a means of the k-means-clustering system. With the help of covariance-matrix accompanied by with time series we are capable to identify the object-image. This method is computationally expensive as the object-images are needed to be kept in the large database and it needs an eminent huge space in this method.

5.1. Text Human-Object Classifier

We propose a robust tactic for image classification quandary without taking the whole object-image. We begin to take out several features from every pixel. Here we think about object-image-region-intensities and have taken many order derivatives with respect to the variables. So we gain the equation for feature classification as:

\[
\chi(a, b) = \left[ \psi(a, b) \begin{bmatrix} \frac{\partial \psi(a, b)}{\partial a} \\ \frac{\partial \psi(a, b)}{\partial b} \\ \frac{\partial^2 \psi(a, b)}{\partial a^2} \\ \frac{\partial^2 \psi(a, b)}{\partial b^2} \end{bmatrix} \right] - (x iv)
\]

The object-image is chosen arbitrarily from each image and the sizes will lie within 16×16 and 256×256. Integral object-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 aforesaid procedure is repetitif if required.
5.2. Experimental outcome on Human-Object-Image

We gained the result of our experiment based on Brodatz object-image dataset comparing 75 images and we have allowed the various images on the dataset. Each object-image is farther divided into sub-images so as to build it available as train image set for our test. Here we have achieved 91.43% correctness depending on our object-image-dataset.

Comparing our 50 images set with our train image set that are randomly selected based on covariance-matrix for every image we have gained the aforesaid result i.e. 91.431% It is superior with respect to quicker detection as well as reformation rate with our stochastic object-image(SOI) selection procedure.

Here we have measured some essential features namely-intensity-of-region-of-interest (IRI), intensity-gradient interest (IGI) and Laplacian-Method (LM) as that we can achieve better performance for classification and we reached to attain 85.31%. Thus we have gained the following tabular result:

<table>
<thead>
<tr>
<th></th>
<th>IRI</th>
<th>IGI</th>
<th>LM</th>
<th>SOI</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESULT</td>
<td>81.7</td>
<td>85.3</td>
<td>89.5</td>
<td>91.43</td>
</tr>
</tbody>
</table>

Table 1. Stochastic Enhanced Result

Fig: Stochastic Resultant Chart

REFERENCES


Authors Profile

Md. Alamgir Hossain has received his M.Tech degree from Bengal Engineering and Science University, Shibpur (BESUS), West Bengal. He has more than 15 years in the field of teaching, research and industry. He has published more than 12 papers in International Journals and Conferences. He is pursuing his PhD in Computer Science and Engineering from National Institute of Technology, Durgapur, India in the area of Image Processing. His areas of interest are Artificial Intelligence, Natural Language Processing and Image Processing. He has guided more than 46 PG (MCA) and 43 UG (B Tech, BCA) projects and thesis. He worked as a Lecturer in TIIT, Kolkata, West Bengal and as an Assistant Professor of Calcutta Institute of Technology under the Department of Computer Application (MCA), Uluberia, Howrah, West Bengal.
Prof (Dr) Gautam Sanyal is a member of the IEEE. He has received his B.E and M.Tech degree from National Institute of Technology (NIT), Durgapur, India. He has received Ph.D.(Engg.) from Jadavpur University, Kolkata, India, in the area of Robot Vision. He possesses an experience of more than 25 years in the field of teaching and research. He has published more than 123 papers in International and National Journals / Conferences. Three PhD (Engg) have already been awarded under his guidance. He is guiding PhD scholars in the field of steganography, Cellular Network, High Performance Computing and Computer Vision. He has guided over 10 PG and 130 UG theses. His research interests include Natural Language Processing, Stochastic modeling of network traffic, High Performance Computing, Computer Vision. He is presently working as a Professor in the department of Computer Science and Engineering and also holding the post of Dean (Students’ Welfare) at National Institute of Technology, Durgapur, India.