

Robust Visual Tracking Using Illumination Invariant Features in Adaptive Scale Model

Muhammad Muazzam Hussain*, Kashif Faheem
Department of Telecom Engineering, University of Engineering
and Technology, Peshawar, Pakistan
*Email: muazzam.riu [AT] gmail.com

Arslan Majid
Department of Electronics Engineering, Iqra University,
Islamabad, Pakistan

Abstract—When entering into the realm of Computer Vision, the first thing which comes in to mind is Visual tracking. Visual tracking by far comes into one of the most actively investigated research areas because of the fact that it has an extensive collection of applications in areas such as activity recognition, surveillance, motion analysis and as well as human computer interaction. Some serious challenges of this area which still create hindrance in achieving 100% accuracy are abrupt appearance and pose changes of an object along with its background blockage due to blockages called occlusion, illumination and lighting variances and changes in scale of target object in the frames. Moreover, diverse algorithms had been proposed for the resolution of said issue. Now in such cases, if we study the statistical analysis of correlation between two frames in a certain video, it can be efficiently utilized to get the most exact location of the targeted object. The algorithms in existence today do not completely exploit a strong spatio-temporal relationship that very often occurs between the two successive frames in a video sequence. Recent advances in correlation-based tracking systems have been proposed to address the problem in successive frames. In this thesis a very simple yet quite speedy and robust algorithm that in actual brings all the relevant information used for Visual Tracking. Two of the Models proposed are the “Locality Sensitive Histogram” and “Discriminative Scale Tracking Method”. These are robust enough to the variations which are based on appearance which are normally presented by blockage, pose, illumination and lighting variations alike. A scheme is proposed called scale adaptation which is very much clever to adapt variations of targeted scale in the most efficient manner. The Discriminative Scale Tracking Method is used for detection as well as scale change ultimately resulting in an effective tracking method in the end. Various different experiments with the best algorithms have demonstrated on challenging sequences that the suggested methodology attains promising results as far as robustness, accuracy, and speed is concerned.

Keywords- Visual Tracking, Illumination Invariant Features, Correlation Based Filtering, Adaptive Scale Space, Locality Sensitive Histogram

I. INTRODUCTION

Visual Tracking basically encompasses tracking and tracing of a person, object or thing moving or displacing in either a video or a set of images in a relevant background from frame to frame, etc. [1]. In this technique, a person or object is assigned with labels and points, by the help of which that particular subject is tracked in a scene or video. Now the tracking of objects that are moving or displacing from point to point and frame to frame is sometimes difficult because of the light, background, illumination and other environmental variations. Majority of algorithms pertinent to Tracking that are based on the assumption that background is constant and there is a very little or slight variation in a scene. Esp. abrupt changes are neglected [1]. For Visual Tracking, previous info. Including sizes, shapes, count in numbers of the objects are required.

This technique has various applications in Modern day technology like Surveillance, object detection, its recognition, augmented reality, etc. [2]. The major difficulties arise when the object comes into illumination and intensity variations, changes in object size, shape, pose and geometry [3]. It often happens that object labels lose their presence before tracker due to the abrupt changes in shapes, illumination, etc. this particular problem is still a chief challenge as far as the precision and effectiveness of the concerned algorithm is there while tracing.

The main assignment of visual tracking goes with the prediction of the target object while at the same time receiving the stream of framed data. There may be changes that an object may experience like changes in its some of its illumination along with its pose w.r.t the camera viewpoint, its shift, its rotation, change in scale, move speedily and abruptly and deforming by pulsating and articulating. Due to these changes, the object may lose its path and the tracker may lose its contact over the period of time. It is of vital importance that the prediction of the object must be carried out in the next data frame. The prediction of such objects can then be solved by adapting theory of probability or formulating the probabilistic problem where the chance of variation is calculated in the next data frame.

The process of prediction of such objects in computer vision, artificial intelligence and its branch known to be pattern recognition, is the formulation of learning and classification problem. Classification problem is solved by using two approaches, one is generative approach and the second is discriminative approach. The generative approach deals with the generative calculation method, which uses the joint

probability distribution $p(x,y)$ whereas the discriminative approach mainly deals with the factor of conditional probability distribution $p(y|x)$ of the provided data. The generative approach models the generation of the signal and the discriminative approach simply categorizes the signal. On the basis of these approaches, the visual tracking can be composed of two approaches i.e., generative approach and discriminative approach [4].

II. RESERCH METHODOLOGY

In this proposed methodology a system was developed to address the issues of illumination variation and change of scale of an object. The frame work was divided into two major blocks. The first block is the design of an illumination invariant feature extractor which can adjust if there are changes in intensities of light source or sources. And second an adaptive scale frame work was developed that can progressively adjust itself if the scale of the tracking object suddenly changes. The block diagram of the system being proposed is shown in figure 2.1. As it can be seen from the figure that design has two major portions. The first block will compute illumination invariant features and second block is the tracker based on discriminative scale space.

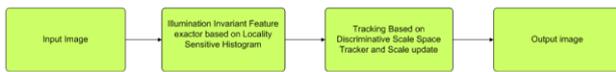


Figure 2.1: Block-Diagram of Method Proposed

The flow of proposed system as the first frame of the video sequence was taken the user that provided the initial coordinates which specify the boundaries of tracker box. After that the frames went through the first block i.e. calculation of illumination invariant features. In illumination invariant feature block the image histogram was calculated, as oppose to traditional histogram equalization we have used the concept of locality sensitive histogram calculation as specify in paper [5]. After the determination of histogram, the image passed through the illumination invariant feature filter.

A. LOCALITY SENSITIVE HISTOGRAM

The difference from traditional histogram in such a way that it considers the contribution of neighboring pixels by assigning weights as they move further away from the center of the image. Where as in normal histogram equalization computes the intensity values by spreading on an equal scale.

B. ILLUMINATION INVARIANT FEATURES

Most of the time, Images which are shot under distinct brilliance circumstances have extreme effects on target object's appearance. If we assume the illumination change

which is affine, images of the scene can be created as shown in Figure 2.2(a) with those that are taken under dissimilar lighting situations as revealed in Figure 2.2(b) and 2.2(c).



Figure 2.2: Illumination invariant features. (a)-(c) are those input images having dissimilar illuminations. (d)-(f) show corresponding outputs.

The intensity values (or color) are not effective enough by which we can match an image. We are hereby proposing a method to extract thick illumination invariant features which are built on an image transformation. For this purpose, Let “ I_p ” & “ I'_p ” be defined as the intensity values of pixel “ p ” which were (just before and a little) after illumination change which is affine.

We then have:

$$I'_p = A_p(I_p) = a_{1,p}I_p + a_{2,p}, \quad (1)$$

In the above equation, $a_{1,p}$ and $a_{2,p}$ are the two parameters related to the “Affine Transform” A_p at pixel “ p ”. Let us take H_p^S to be a specific histogram which is figured from a window S_p which gets centered at pixel p , and let b_p be the corresponding bin with value of intensity “ I_p ”. If we see the histogram definition, the pixels’ number in “ S_p ” among which intensity values fall within $[b_p - r_p, b_p + r_p]$ is shown by the formula as:

$$I_p = \sum_{b=b_p-r_p}^{b_p+r_p} H_p^S(b), \quad (2)$$

In the above formula, parameter r_p basically controls the integration interval at pixel p . Likewise, If “ r_p ” is scaled linearly with illumination in such a way that

$$r'_p = a_{1,p}r_p, \quad (3)$$

The integrated value I_0 p acquired from a dissimilar condition of illumination matches to the pixels' number with values of intensity which lie within

$$\begin{aligned} & [a^1_p b_p + a_{2,p} - a_{1,p} r_p, a^1_p b_p + a^2_p + a^1_p r_p] \\ & = [a^1_p (b_p - r_p) + a^2_p, a^1_p (b_p + r_p) + a^2_p] \\ & = [A_p (b_p - r_p), A_p (b_p + r_p)]. \end{aligned}$$

Now, if we overlook the error of quantization, I'_p gets equal to I_p . Therefore, " I_p " is affine illumination change which is independent and thus be utilized as an invariant feature prevailing under various different illumination conditions as long as Eq. 3 holds.

We can set the equation as:

$$r_p = \kappa |I_p - \bar{I}_p|, \quad (4)$$

In the above equation, $\kappa = 0.1$ is a constant $\bar{I}_p = \frac{1}{|S_p|} \sum_{q \in S_p} I_q$ Is basically the Mean Intensity Value of S_p window. Whereas, $|S_p|$ is the total number of pixels in S_p . With some further supposition that the Affine Illumination Change is made smooth locally in such a way that our transform gets the same for all of the pixels lying inside the window S_p , we have now,

$$\begin{aligned} r'_p & = \kappa |I'_p - \bar{I}'_p| \\ & = \kappa |a_{1,p} I_p + a_{2,p} - \frac{1}{|S_p|} \sum_{q \in S_p} (a_{1,p} I_q + a_{2,p})| \\ & = a_{1,p} \kappa |I_p - \bar{I}_p| \\ & = a_{1,p} r_p. \end{aligned} \quad (5)$$

As a consequence, Eq. 3 gets true when the interval " r_p " is acquired adaptively from Eq. 4.

The suggested illumination invariability lies on the hypothesis that the "Affine Illumination Change" is equal for all image pixels from that which our conventional histogram is being calculated. However, the ground reality says that it is incorrect to define a precise Local Window inside of which the assumed "Affine Illumination Transform" still rests unaffected. Therefore, it is needed to adapt such estimation in which we will be assigning greater weightage to pixels which are somewhat nearer to the central pixel & vice versa. Henceforth, we replace histogram H^S_p in Eq. 2 with our proposed and suggested locality sensitive histogram H^E_p , which in essence sees the contributions of all image pixels, adaptively. In addition to it, we make in use a "soft" interval which shrinks the Quantization Error, thus Eq. 2 becomes:

$$I_p = \sum_{b=1}^B \exp\left(-\frac{(b - b_p)^2}{2 \max(\kappa, r_p)^2}\right) \cdot H^E_p(b), \quad (6)$$

where $\bar{I}_p = \sum_{b=1}^B H^E_p(b)$ Since a^2_p is comparatively small, r_p can be swapped by $\kappa |I_p - \bar{I}_p|$. The never changing features calculated from Figure 2.2(a)-(c) are shown in Figure 2.2(d) (f). In contrast to the values of intensity, they still hold same even under the influence of strong variation in illumination. This altered image basically formulates one of the inputs for tracking algorithms being suggested in tracking formulation.

As it can be seen from figure 2.3 the change in light intensity does not affect the output of the filter. This demonstrates that changes in intensities of the light does not affect the quality of the picture.



Figure 2.3: Showing the change in light intensity does not effect the output of the filter.

The output of the previous block is fed to the next block i.e. Exhaustive Scale Space tracking to address scalability issue in the sequence.

A. C. Exhaustive Scale Space Tracking

In this searching technique a grey scale image is converted in to its subsequent patches f_1, \dots, f_t which will model the target appearance of training data. The expected outputs of correlation function g_1, \dots, g_t are compared with the training data at time " t " this can be evaluated by minimizing the sum of squared errors.

$$\epsilon = \sum_{j=1}^t \|h_t \star f_j - g_j\|^2 = \frac{1}{MN} \sum_{j=1}^t \|\bar{H}_t F_j - G_j\|^2. \quad (7)$$

Functions f_j, g_j and h_t are having dimensions of order $M \times N$. Circular correlation is denoted by \star . Parseval's identity theorem is followed for calculation of \bar{H}_t . The capital letters

of $\overline{H}_t F_j$ specify that their Desecrate Fourier Transforms (DFT) are calculated for the corresponding functions.

$$H_t = \frac{\sum_{j=1}^l \overline{G}_j F_j}{\sum_{j=1}^l \overline{F}_j F_j} \quad (8)$$

To determine the output of correlation function g_i can be built up as a Gaussian function where its peak likely hood will be at the center of the f_i . The terms of A_t i.e. is numerator and B_t which is the denominator of the equation 8 are separately updated with this new observation (f_t) of the target while taking weight average.

If given some image patch “z” of size $M \times N$ in a new frame, y being the correlation scores are computed as $y =$

$\mathcal{F}^{-1}\{\overline{H}_t Z\}$. Here, this \mathcal{F}^{-1} denotes the “Inverse DFT Operator”. Now, the new target location is projected to be at the maximum correlation score of y .

Furthermore, a joint translation-scale tracking method was devised which was primarily built on the learning of a correlation filter based on scale space in 3-dimensions. The size of the filter taken being fixed to $M \times N \times S$, where M denotes the height and n denotes the filter’s width, whereas “S” denotes the scales in numbers. We initially compute a pyramid of features within an area of rectangular around the target object to update the relevant filter. That pyramid is made in such a way that the size of the targeted subject gets $M \times N$ taken at its predictable scale. In this case, training sample which is “f” is then set to a feature pyramid of rectangular cuboid. The cuboid size is $M \times N \times S$ which is placed at the target object’s projected scale and location. We, afterwards use a Gaussian function which is 3-dimensional in nature as the wanted output of correlation “g”. At the end, our filter called scale space tracking is updated using equation (9).

After that, an $M \times N \times S$ rectangular cuboid “z” was extracted from this pyramid of features to try to trace the object in a new video frame as discussed above. The cuboid is centered at projected target scale and location. The scores of correlation denoted by “y” are now calculated using (7). The new target location with the scale are afterwards gained by looking for the maximum score in “y”.

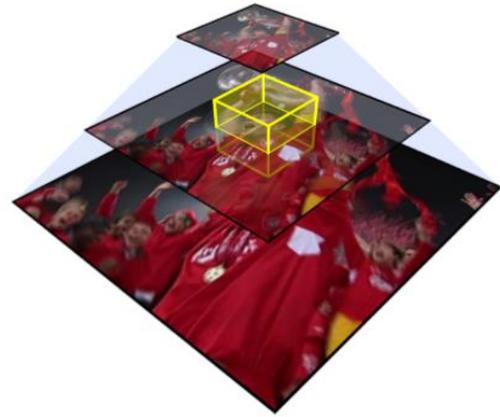


Figure 2.4: Above is the visualization of joint scale space filter approach. This training sample (yellow box) is extracted by scale pyramid constructed about the target center.

$$A_t^l = (1 - \eta)A_{t-1}^l + \eta \overline{G}_t F_t^l$$

$$B_t = (1 - \eta)B_{t-1} + \eta \sum_{k=1}^d \overline{F}_t^k F_t^k \quad (9)$$

Input:

- Image I_t .
- Previous target position p_{t-1} and scale s_{t-1} .
- Translation model $A_{t-1}^{trans}, B_{t-1}^{trans}$ and scale model $A_{t-1}^{scale}, B_{t-1}^{scale}$.

Output:

- Estimated target position p_t and scale s_t .
- Updated translation model A_t^{trans}, B_t^{trans} and scale model A_t^{scale}, B_t^{scale} .

Translation estimation:

- 1: Extract a translation sample z_{trans} from I_t at p_{t-1} and s_{t-1} .
- 2: Compute the translation correlation y_{trans} using $z_{trans}, A_{t-1}^{trans}$ and B_{t-1}^{trans} in (9).
- 3: Set p_t to the target position that maximizes y_{trans} .

Scale estimation:

- 4: Extract a scale sample z_{scale} from I_t at p_t and s_{t-1} .
- 5: Compute the scale correlation y_{scale} using $z_{scale}, A_{t-1}^{scale}$ and B_{t-1}^{scale} in (9).
- 6: Set s_t to the target scale that maximizes y_{scale} .

Model update:

- 7: Extract samples f_{trans} and f_{scale} from I_t at p_t and s_t .
- 8: Update the translation model A_t^{trans}, B_t^{trans} using (8).
- 9: Update the scale model A_t^{scale}, B_t^{scale} using (8).

The output of block two addressed the scalability issue in the sequential frames. Output of block two is the final output of the proposed tracker in which both illumination invariance features and scalability issues are addressed simultaneously.

III. RESULTS AND DISCUSSIONS

A. Tracking and Dataset Sequence

The proposed algorithm of Visual Tracking was tested on 11 sequences which were challenging in nature and are online available on visual tracking benchmark. Our proposed tracker was measured with the other two methods i.e Locality Sensitive Histogram Tracker (LSHT) [12] and Discriminative Scale Space Tracker (DSST) [13]. The parameters from the trackers source codes were used for best results for compared tracking results. Our proposed tracker is now implemented in MATLAB.

B. Experimental Setup

Initially, the context region's size is set by multiplying by 2 the targeted object. The parameter σ (sigma) initially set to

$$\sigma 1 = \frac{h+w}{2}$$

As in the above equation, "h" is related height and whereas "w" denotes initial tracking rectangle' width, respectively.

Other Parameter in code are set as following

Extra area surrounding the target (padding)= 1.0

Standard deviation for the desired scale filter output(σ)= 1/4

Number of scale levels (denoted "S")= 33

Scale increment factor (denoted "a")= 1.02

C. Experimental Results

The video frames are in gray scale, all of them and to quantitatively evaluate and asses the suggested algorithm with the other mentioned trackers, two metrics were used. The metrics being used are

i. Success Rate (SR)

ii. Central Location Error (CLE)

1) Success Rate (SR)

We estimate Rate of Success on the basis of the scores which are obtained. This is basically the ratio of correctly tracked frames by total frames in a sequence. For this very purpose, the Tracked frames along with the bounding box values are gauged with the manually labeled ground truth. Now the Ground truth values are those values which are expected to be achieved by the tracker. The obtained score if is greater than or equal to 0.5, then that frame is considered a success otherwise not. For this, the overlap rate certainly has to be computed between the bound boxes pertinent to the tracker along with its bounding box of the ground truth then. If

we see the Figure 3.2, the red bounding box shows the "Ground Truth Box" and another blue one represents the box which is calculated by the tracker. While, the Overlap rate is determined by the equation given in 5.1.

$$\text{Overlap Rate} = \frac{\text{area}(R_G \cap R_T)}{\text{area}(R_G \cup R_T)} \quad (10)$$

Whereas in the above equation, R_G and R_T are the regions of the box of the truth along the tracker respectively.



Figure 3.1: the above figure shows the Overlapping of a manually labeled ground truth and the tracker box[11].

The Overall Success rate of all the trackers which are compared with the proposed one is shown in figure 3.2 a and 3.2 b in the form of a bar chart to explain clearly. It can easily be observed that the suggested tracker beats the other two trackers which are used in the evaluation thereof.

Table3-1: Success Rate(SR)in eachframe.

Sequence	Total Frames	Proposed Method	DSST	LSHT
Bolt	350	350	350	298
Boy	602	602	602	303
Coke	291	275	261	212
David	770	770	700	632
David3	252	151	151	223
Deer	71	56	56	63
Doll	3872	3845	3845	1600
Tiger1	354	294	294	329
Tiger2	365	256	110	330
Trellis	569	569	569	469
Car	393	393	360	340
Successful Frames	7889	7561	7298	4799
Average SR		96%	93%	61%

2) Central Location Error (CLE)

This error comes basically when the calculation of the Euclidean distance which lie in the middle of the central locations of the object being tracked with the ground truth labeled manually. CLE can be calculated by the equation 11 which is shown as below,

$$center\ error = \sqrt{(x_g - x_t)^2 + (y_g - y_t)^2} \quad x_t \quad (11)$$

Whereas in the above formula x_g and y_g introduces initial coordinates which are the ground truth bounding box along x and y axes whereas x_t represents the initial coordinates along x and y axis of the tracker bounding box. Then again, the average CLE of each and every frame in a particular sequence is calculated to summarize an accumulative progress and performance of that particular sequence. Now, the below Table 3-2 enlists the average CLE of all the sequences which are used in the evaluation phase thereof.

Table 3-2: Center location error (CLE) in pixels.

Sequence	Proposed Method	DSST	LSHT
Bolt	4.79	4.79	2.9
Boy	2.17	2.17	5.4
Coke	9.7	13	10.1
David	1.78	2.4	4.5
David3	8.8	8.8	6.5
Deer	16.7	16.7	9.8
Doll	3.12	3.12	4.7
Tiger1	1.7	1.7	1.9
Tiger2	12.4	41.3	2.3
Trellis	2.71	2.71	9.2
Car	1.01	1.33	4.2
Average CLE	5.90	8.91	5.59

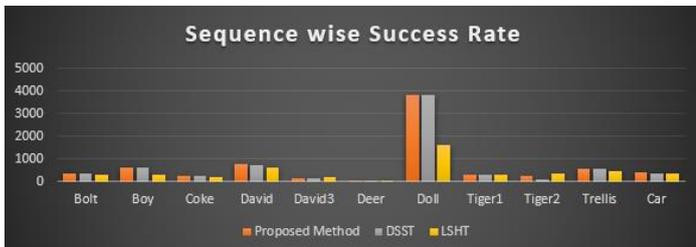


Figure 3.2: a) shows graphical comparison of proposed tracker with other two trackers for each sequences. b) shows the average success rate of the proposed tracker

Again, the bar graph of the average CLE of all the sequences is shown in figure 3.3 as shown.

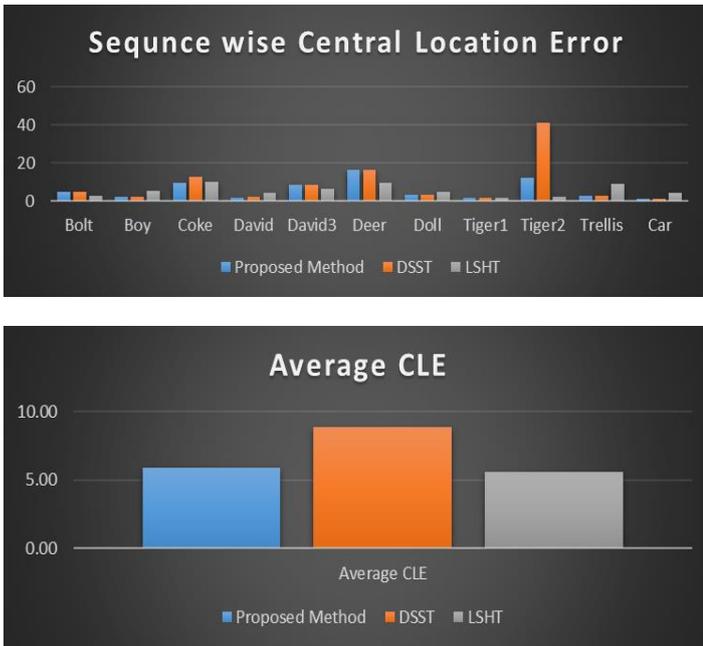


Figure 3.3: Central Location Error in terms of pixel

D. Qualitative Evaluation

Based on the different challenging of mentioned sequences, Qualitative evaluation was done. So, for qualitative evaluation illumination variation and scale space variations were evaluated.

i. Sequences Having Illumination Invariant Features

It was previously seen that in sequences where there was a Car, Trellis, Coke, David the objects' looks varies due to the forming shadows and the variations in lightening. Proposed methods really did well in this sequence. In the indoor sequences of Trellis and David, the targeted objects somehow experienced the gradual illumination change. In case of Indoor sequence of David, the object was initially completely in darker region. It was observed that the subset of image sequences was used by other mentioned trackers and not the whole image sequence. Whereas, while evaluating David and Trellis, some image frames from the evaluation phase were cut. While in the David indoor sequence, suggested method at initial stage drifted away but it recover and rejuvenated in latter stages and hence majority of tracking done by it was a good success. Whereas in the Tre sequence, the target object was initially placed in a bright region and at the end of the sequence, it abruptly goes under darker region. The suggested method in this respect perform well again. Main reason for this is that the proposed method update works reasonably well when the illumination invariant filtering process is deployed in it. The illumination Invariant filter empowers the tracker to grip gradual illumination changes very efficiently. This can be accredited to the

proposed feature mechanism which is pretty much insensitive to the gradual illumination variations. In Figure 3.5 some screen shots of Coke, Trellis, David.

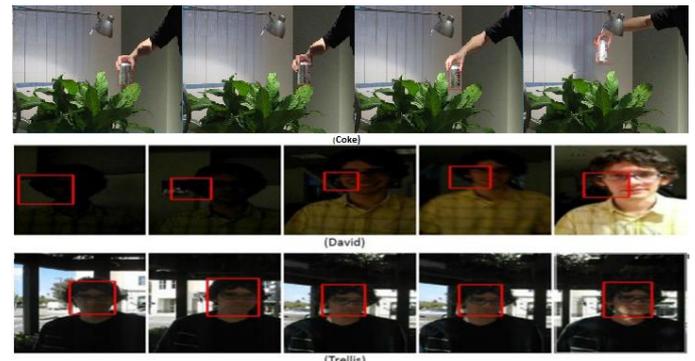


Figure 3.4: Screen shots of Coke, David and Trellis sequences are shown. These sequences had illumination variation and as well as pose changes. The coke and Trellis sequence performs well as expected due to the adaptive nature of the features used to handle illumination changes. Whereas In the David Sequence, the tracker is partially drifted but it still covers most of its targeted object.

ii. Sequence having Change in Scale

To test the changes in scale, we have taken the Doll and Bolt sequence in which is shown a continuous change in scale of object in the respective frames. The tracker did very well in the both sequences. The suggested method by us for scale changes that is Joint Translation Scale Tracking Method. The scale sequence of Doll undergone scale changes as it was observed in its specific sequence. For Bolt sequence, we can see that the runner which is running from one target to another, is continuously changing the scale while the tracker is updating its scale along with the movement of the runner.



Figure 3.5(a): Qualitative evaluation of doll sequence



Figure 3.5(b): Qualitative evaluation of Bolt

IV. CONCLUSION

Conclusively, we have offered a very simple yet quite speedy and robust algorithm that in actual brings all the relevant information used for Visual Tracking. Two of the Models proposed were the “Locality Sensitive Histogram” and “Discriminative Scale Tracking Method”. These are robust enough to the variations which are based on appearance which are normally presented by blockage, pose, illumination and lighting variations alike. A scheme was proposed called scale adaptation which is very much clever to adapt variations of targeted scale in the most efficient manner. The Discriminative Scale Tracking Method was used for detection as well as scale change ultimately resulting in an effective tracking method in the end. Various different experiments with the best algorithms have demonstrated on challenging sequences that the suggested methodology attains promising results as far as robustness, accuracy, and speed is concerned.

It was revealed that the suggested method of tracking was competent in comparison with other trackers, from the given test sequences. The algorithm which is exploiting some back ground evidence assists in decision making about prediction of the next incoming frame object, hence by comparing with the earlier frame, the confidence value can be figured out. This method typically aided in those cases where the bounding box drifts or sways away from the targeted object. It was revealed that such method is very much robust to abrupt illumination changes when they occurred. The success rate and the central location error actually shows that the accumulative efficiency got improved by 33 percent when it was compared with LSHT tracker, but with the improvement in efficiency its performance speed wise was very low.

REFERENCES

- [1] Maggio, Emilio, and Andrea Cavallaro. “Video tracking: theory and practice”, John Wiley & Sons, 2011.
- [2] Vapnik, Vladimir, “Statistical learning theory”, 1998, Wiley, New York.
- [3] A. Yilmaz, O. Javed, and M. Shah. “Object Tracking: A survey”, ACM Computing Surveys, 2006.
- [4] Y. Andrew and M. Jordan, “On Discriminative vs. Generative Classifiers: A comparison of logistic Regression and Naive Bayes”, Neural Information Processing Systems, 2001.
- [5] H. Shengfeng, Y. Qingxiong, L. Rynson, J. Wang and M. Yang, “Visual Tracking via Locality Sensitive Histograms”, 2013, Computer Vision and Pattern Recognition, June.
- [6] T. Dinh, N. Vo, and G. Medioni, “Context Tracker: Exploring supporters and distracters in unconstrained Environments”, 2011, European Conference on Computer Vision.
- [7] Yi Wu ,Jongwoo Lim and Ming-Hsuan Yang, “Online Object Tracking: A Benchmark ”, 2013, IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [8] M. Danelljan, G. Häger, F. Shahbaz Khan, and M. Felsberg. “Accurate scale estimation for robust visual tracking”, 2014, In Proceedings of the British Machine Vision Conference (BMVC).
- [9] M. Danelljan, A. Robinson, Fahad Khan and M. Felsberg, “Beyond Correlation Filters: Learning Continuous Convolution Operators for Visual Tracking”, 2016, In Proceedings of the European Conference on Computer Vision (ECCV).
- [10] V. Boddeti, T. Kanade, and B. Kumar, 2013, “Correlation filters for object alignment” , IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [11] H. Bolme, R. Beveridge, A. Draper and Yui Man Lui “Visual Object Tracking using Adaptive Correlation Filters”, 2011, IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [12] H. shengfeng, Y. Qing, L. Rynson and Yang Ming “Visual Tracking via Locality Sensitive Histograms”, 2013, IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
- [13] M. Danelljan, H. Gustav, Fahad Khan and M. Felsberg, “Discriminative Scale Space Tracking”, 2017 ,in Proceedings of the IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI).