January 2014

Object Tracking by Normalized Cross Correlation and PCA Based Template Updating: Comparative Analysis

M.H. Sidram

Department of E & EE, Sri Jayachamarajendra College of Engineering , Mysore, Karnataka, India.,
mhsidram@gmail.com

N. U. Bhajantri

Department of Computer Science Engineering Government College of Engineering Chamarajnagar,
Mysore District, Karnataka, INDIA, bhajan3nu@gmail.com

Follow this and additional works at: <https://www.interscience.in/ijcsi>

Part of the [Computer Engineering Commons](#), [Information Security Commons](#), and the [Systems and Communications Commons](#)

Recommended Citation

Sidram, M.H. and Bhajantri, N. U. (2014) "Object Tracking by Normalized Cross Correlation and PCA Based Template Updating: Comparative Analysis," *International Journal of Computer Science and Informatics*: Vol. 3 : Iss. 3 , Article 8.

DOI: 10.47893/IJCSI.2014.1143

Available at: <https://www.interscience.in/ijcsi/vol3/iss3/8>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in International Journal of Computer Science and Informatics by an authorized editor of Interscience Research Network. For more information, please contact sritampatnaik@gmail.com.

Object Tracking by Normalized Cross Correlation and PCA Based Template Updating: Comparative Analysis

¹M.H.Sidram, ²Nagappa.U.Bhajantri

Department of E & EE, Sri Jayachamarajendra College of Engineering , Mysore, Karnataka, India.

²Department of CS&E, Government Engineering College, Chamarajanagar, Karnataka, India.

mhsidram@gmail.com, bhajan3nu@gmail.com

Abstract: The principle behind to detect and track non-stationary object via a sequence of frames is addressed. The proposed strategy pushed the Normalized Cross-Correlation (NCCR) to track object by matching the template and updating the template is encouraged through Principal Component Analysis (PCA). This work remarked with exhaustive experiment and witnessed with comparative analysis over dataset related to outdoor environment. The system kernel reveals the capability to track the object and outcome is fairly acceptable to great extent under different light conditions.

Keywords: *Object tracking, Frame difference, Template matching, Normalized Cross-Correlation, PCA*

1. RELEVANT FACTS

An attention drawing mechanism, accountable to establish the correspondence among the objects in set of frames is object tracking. Perhaps it unearths many applications but important among them are in video surveillance, monitoring the traffic and as a vision to the robot. There is no dearth of relevant literature in tracking object emerged in a moderate scene. It could be possible through spatial or appearance based model. Secondly several processes are evolved from frequency sphere. Further too hybrid approaches are celebrating effective performance. There are several approaches for tracking object in a scene that are Point tracking, Kernel tracking and silhouette tracking. Template matching is sub-class of Kernel tracking [8].

Some of the factors make object tracking complex due to change in color and illumination, noise in the images, abrupt motion of the objects and computational aspects for real-time processing [8]. Today the prime research in computer vision algorithm is detection and tracking of object. One such application is analysis of traffic scene. Thus vehicle detection is important for civilians as well as military

usage especially in aerial and usual traffic scene since vehicles are vital part of human life.

This paper attempts to propose a system which tracks the object vigorously with the correlation between object and template. However it takes care of updating the template with the help of different approaches like work [12] NCCR (Normalized Cross-Correlation) and novelty exhibited by freshly proposed PCA [7] for updating the template. In order to emphasize the proposed process with the help of three building blocks, such as correlating the template and image is aspired which is motivated based on the correlation score. Secondly the frame differencing algorithm is employed to produce the motion regions. Finally, sub-images are cropped and stored via frames which correspond to motion regions. In the sequel, existing template will be correlated with sub-images and the best match will be replaced using NCCR in the first experiment and updating method is changed in the second experiment using PCA. In case of PCA based updating, the existing template is compared with the sub-images and hence the best match is declared as new template based on the Euclidean distances values.

It gives high value when the template (test image) has a best match with sub-images (training images) as a new template. This process is repeated for different interval of frames and is called the updating frequency f . An experiment has been conducted exhaustively employing the benchmark datasets such as PETS 2001 [3] and VISOR [11] and their details are tabulated in the TABLE- 1.

Table-I Different Dataset

Dataset	PETS 2001 (1)	PETS 2001 (2)	PETS 2001 (3)	VISOR
# of frames	2353	2240	2688	1495
Contents	Human, Cars and People	Human, Cars and People	Human, Cars and People	Human, Cars
Camera	Side fixed, Moving tree	Top-Down fixed	Side fixed	Side fixed

This paper stuffing is set as follows. The related work has been dealt in Section 2. The proposed method is emphasized in Section 3. The experiment and results are discussed in section 4. Conclusion and future work presented in section 5.

2. COMPREHENSION OF ESSENTIAL LITERATURE

In the work of [4] J.P. Lewis et al. encourages the potentiality of normalized cross-correlation based template matching in the spatial domain.

The author [1] Alan .J. Lipton et al. attempted to employ the combination of frame differencing and template matching to highlight the object in a scene. The template matching is guided by temporal differencing and image based correlation to make tracking process robust. Further the Impulse Response filter (IIR) is used to update the template, in other words it is known as adaptive template matching method. Researcher Hieu T. Nguyen et al. [2] tried to comprehend the tracking process for a rigid object through Kalman filter and consequently updating the template to adapt changing illumination and orientation of the object is achieved via an adaptive Kalman filter.

In the work of [6] Longin Jan Latecki et al. proposed strategy which is based on selective hypothesis tracking algorithm. It includes the motion regions, image alignment and minimum cost estimation to update the template dynamically. In other words minimum cost matching is established through association between the motion region and the aligned template. Thus motion vector is updated

Dynamic template matching and controlling the field of view of camera by PTZ was remarked by [5] Karan Gupta et al. using frame difference approach and choosing the proper threshold. This strategy basically tries to consider the instant updating the template although limited to a single object in a scene.

In the work of Xue mei et al. [9] used the probabilistic algorithm for tracking, which included template matching and incremental subspace update. The templates are modeled using mixed probabilities and updated based on considerably changes of the object appearance. The augmentation of the Kernel Gream matrix with a row and column yields the updating.

Wenhui Liao et al. [10] introduced a new method called Case Based Reasoning (CBR) to maintain accurate template of object automatically. In other words algorithm dynamically updates the case base (template). With this, real time face tracking is built to track the face robustly under different orientations and conditions.

M.H.Sidram et al. [12] evolved a novel strategy in which the object is tracked by matching the template and updating the template through Normalized Cross-Correlation.

The literature surveyed till this point has encouraged us to propose a system based on NCCR for template matching to track the object and update the template by conducting experiments using NCCR method and PCA method separately and hence compare the results.

3. PROCESS ERECTED

This para dedicated to present a proposed work and aims to track the object and update the template by two approaches. Compare the tracking efficiency of both methods. The simplified block diagram of a general system is shown in Figure.1

Figure 1: Block Diagram of Proposed system

The efficient object tracking involves two steps i.e. firstly template matching and secondly template updating. In this attempt template matching is done by NCCR. But updating is done by proposed PCA approach. In other words the proposed method is the extension of our previous work [12].

(A) The computation of normalized cross-correlation involves through the following mathematical expression [12] displayed. The location where the maximum value of correlation score occurs and corresponding location is the best match. Thus it gives the evidence to put the bounding box over the object.

$$C(u, v) = \sum_{x,y} f(x, y) t(x - u, y - v) \quad (1)$$

It is used as a measuring unit of similarity between the image and the template. The difficulties are noticed such as image energy which causes correlation score minimum, sorting of $C(u, v)$ depends on template size,

change in illuminations not affecting the equation (1) are eliminated through a process of normalization. Therefore the normalized cross-correlation (γ) expressed through equation 2 as follows.

$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2\}^{0.5}} \quad (2)$$

Where $\bar{f}_{u,v}$ and $\bar{t}$ are means of image and template respectively.

Further the necessity of template updating as we discussed and same is achieved through the equation 3. In order to obtain the absolute value of moving object by frame differencing below equation is exploited.

$$D = |f_m - f_{m-1}|$$

$$P(i, j) = \{f(i, j) \mid D \geq T\}$$

$$\{0 \mid D < T\} \quad (3)$$

(B) PCA is also known as karhunen Loeve transform

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \text{ n*1 vector} \quad (4)$$

$$m_x = 1/m \sum_{k=1}^m x_k \text{ mean} \quad (5)$$

$$c_x = 1/m \sum_{k=1}^m x_k x_k^T - m_x m_x^T$$

Co-variance matrix (6)

The goal of PCA is to find a set of μ_i for $i=1, 2, \dots, n$, which have largest projection on to each $x_k - m$, which is maximization of quantity

$$\lambda_i = 1/m \sum_{k=1}^m (\mu_i^T [x_k - m])$$

Using Rayleigh's principle it can be proved that the solution of this equation is given by eigen values λ_i and eigen

vectors of co-variance matrix c_x

$$Y = \psi(x - m_x)$$

$$c_y = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Eigen values of c_x and c_y are identical

In the first case of experiment close the switch to the position-1 of the algorithm in figure-1 and conduct an experiment on various bunch of frames of different dataset at various updating frequencies f and in the second case close the switch to the position-2 and repeat the experiment. Compared the results and tabulated through the tables II to V.

In view of obtaining the binary form from the difference image by selecting suitable threshold and post-process it in the later stage using the morphological operations. Then, connected component helps to label the moving objects. In the sequel the centroids of moving objects are estimated.

The proposed strategy encourages to mound the cropped sub-images with the help of centroid followed by process of computation of cross-correlation score between the template and sub-images. Therefore the best match will be the new template and process of updating is repeated for k - interval of frames. As it is empirically observed by the proposed experiment, the value of the k reflects with the dataset. This entire process is illustrated in subsequent section through two phases. First algorithm is predominantly exhibit object tracking task and second one dedicates to update the template in turn which supports and provide enhanced knowledge to track the object.

Algorithm-I

1. Unfurl video into frames.
2. Median filter is employed to obtain noise free frame.
3. Initialize with template.
4. Read the r^{th} frame and the template, compute the correlation score. Put the bounding box over the object for the best match.
5. Generating and updating the template after every fixed interval of frames using Algorithm-II.
6. Step 4 and 5 are repeated for n frames

Algorithm-II

Generate and update the template after every fixed interval

1. Initialization of count through k - interval of frames.
2. Get absolute value by subtracting m^{th} frame from $(m-1)^{th}$ frame.
3. Using threshold, the difference image is converted to binary form
4. The moving objects are labeled using connected component analysis.
5. Determine the Centroids of moving objects.
6. Cropped sub-images corresponding to centroids are stored.
7. Declare a new template using PCA between the template and sub-images

4. EXPERIMENTS AND RESULTS

The experiment conducted to corroborate the performance efficacy of the NCCR approach. The computational aspects of the evolved methods turn out to be polynomial and its order is $O(n^6)$ for NCCR and $O(n^5)$ for PCA approach. The same is tested over the available machine Pentium(R) Dual-core CPU, T4200 @ 2.00 GHz and 2.83 GB of RAM of 1.20 GHz.

The proposed approach capable to digest single object as small as 50 pixels is tracked efficiently. Template update is done empirically for every k frames which yields better performance. In the experiment environment k predominantly represents template updating at interval of frames and also known as updating frequency f . Here we have noticed some of the interesting observations which made us keen upon further exploration in the future work.

The critical observation done for updating template at every alternate frame becomes computationally expensive. On the other hand updating after many frames will fail the tracking. Hence it is empirically chosen a suitable update frequency as k for each dataset because of the stability. It is also further noticed by experimentation, that the tracking performance is directly proportional to the size of template. In other words larger the template, tracking is better. The proposed system has robustly performed over the different datasets and varied illumination conditions.

Dataset used are PETS2001 (Video clips 1, 2 and 3 clips) and VISOR (Video for traffic surveillance clip). The experiment is conducted on the different bunch of

frames of each dataset which includes different objects. Hence the results of the Tracking (T) and Mis-tracking (T) at different updating frequencies f have been tabulated. But for the sake clarity, one of the tables on each dataset has been portrayed through the TABLES II to V. Tracked frames are shown through the Figures 2 to 5.

Individual objects are tracked for PETS 2001(1) dataset using respective templates at fixed updating frequency 5 and few of them are selected to experiment are tabulated in the TABLE VI and this reveals that the Mis-tracking (MT) rate is minimal in case of PCA based updating method.

From the TABLE-II it is clear that for NCCR based updating, for $f=7$, frame 1380th will get wrongly selected (updated) from sub-images which leads to the Mis-tracking and therefore number of tracking (T) is 23 and number of Mis-tracking is 7. On the other hand PCA method selects (updates) the template exactly and leads to the proper tracking thereby resulting in number of tracking (T) 30 and Mis-tracking as a nil is a remarkable experimental observation.

TABLE III reveals that for NCCR based method, for $f=8$, frame 783rd will get wrongly updated and therefore number of tracking (T) is 22 and number of Mis-tracking (MT) is 8. Whereas PCA method updates the template exactly and number of tracking (T) 30 and Mis-tracking (MT) is nil.

TABLE-IV also unearths that for NCCR based method, for $f=4$, frame 621st will get wrongly updated and number of tracking (T) is 14 and number of Mis-tracking (MT) is 16. Whereas PCA method updates the template exactly and resulted in number of tracking (T) 26 and Mis-tracking (MT) is 4.

Similarly TABLE-V reveals that for NCCR based updating, for $f=10$, frame 941st will get wrongly updated and therefore number of tracking (T) is 10 and number of Mis-tracking (MT) is 20. Whereas PCA method updates the template exactly and leads to the proper tracking thereby resulting in number of tracking (T) 30 and Mis-tracking (MT) is nil. The same effect may also be observed at other updating frequencies in the respective tables.

The experimental observations divulge that the PCA based updating is efficient compared to NCCR approach and hence this reveals the worthiness of novelty of proposed system.

Table- II. *Pets 2001(1)_Dataset_2353_Frames*

f	# of frames	NCCR		PCA	
		T	MT	T	MT
1	30	27	03	29	01
2	30	26	04	28	02
3	30	24	06	30	00
5	30	29	01	30	00
7	30	23	07	30	00
8	30	30	00	30	00
12	30	18	12	27	03
15	30	30	00	30	00
20	30	30	00	30	00
25	30	26	04	29	01

Frequency -Track and Mis-Track : People (1360-1390) # of frames 30.

Table- III. *Pets 2001(2)_Dataset_2240_Frames*

f	# of frames	NCCR		PCA	
		T	MT	T	MT
1	30	07	23	30	00
2	30	28	02	30	00
3	30	16	14	30	00
4	30	18	12	30	00
6	30	22	08	30	00
7	30	19	11	30	00
8	30	22	08	30	00
9	30	25	05	30	00

15	30	30	00	30	00
25	30	23	07	30	00

Frequency -Track and Mis-Track : People_Car (760-790)

Table- IV. *Pets 2001(3)_Dataset_2688_Frames*

<i>f</i>	# of frames	NCCR		PCA	
		T	MT	T	MT
1	30	16	14	18	12
2	30	16	14	20	10
3	30	16	14	19	11
4	30	14	16	26	04
8	30	14	16	22	08
10	30	18	12	28	02
12	30	10	20	22	08
15	30	14	16	28	02
20	30	18	12	18	12
25	30	23	07	23	07

Frequency -Track and Mis-Track : Human_Car (610-640)

TABLE- V. *VISOR_DATASET_1495_FRAMES*

<i>f</i>	# of frames	NCCR		PCA	
		T	MT	T	MT
1	30	20	10	30	00
2	30	24	06	30	00
3	30	18	12	30	00
6	30	18	12	30	00
7	30	30	00	30	00
10	30	10	20	30	00
12	30	22	08	30	00
15	30	30	00	30	00

20	30	30	00	30	00
25	30	30	00	30	00

Frequency -Track and Mis-Track : Human_Car_Car (930-960)

TABLE- VI. *PETS 2001(1)_DATASET_2353_FRAMES at f=5*

Type of Object	# of frames	NCCR		PCA	
		T	MT	T	MT
Car (1090-1240)	150	43	108	149	01
Cycled Human (625-662)	38	28	10	37	01
Human (480-541)	62	24	38	31	30
People (1360-1454)	95	89	06	94	00

Tracking and Mis-Tracking using NCCR and PCA approaches.

5. INFERENCE AND FUTURE VIEW

In this work it is established through normalized cross-correlation feature to track multiple objects. This procedure being able to track object as small 50 pixels and update frequency is empirically decided as k frames. It is observed that larger the template, tracking is better on the contrary poor tracking. Experimental results on PETS 2001 and VISOR video dataset reveal that the approach is capable of spotting and tracking the object correctly. This experiment reveals and helps to draw inference that template updating via PCA is quite efficient comparatively NCCR based template updating. Future work can be focused to track the object for different set of videos and handle the partial and full occlusions. Hence many future avenues can be thought of based on the success reported in this proposed research.

6. ACKNOWLEDGMENT

Thanks are due to the “JSS Research Foundation” for constant support in carrying out the research work.

Figure 2(a)

Figure 3(a)

Figure 2(b)

Figure 3(b)

Figure 2(c)

Figure 3(c)

Figure 2(c)

Figure 3(d)

Figure 2- PETS 2001 (1)

Figure 3- PETS 2001 (2)

Human_Car-610

Human_Car-638

Figure 4(a)

Human_Carblack-839

Human_Carblack-857

Figure 5(a)

Human_Car_Human-950

Human_Car_Human-980

Figure 4(b)

Human_Cargreen-873

Human_Cargreen-897

Figure 5(b)

Human-2050

Human-2080

Figure 4(c)

Human_Car_Car-930

Human_Car_Car-957

Figure 5(c)

Car_Car-2575

Car_Car-2605

Figure 4(d)

Human-1220

Human-1250

Figure 5(d)

Figure 4- PETS 2001 (3)

Figure 5 -VISOR (video for traffic surveillance)

REFERENCES

- [1]. [1] Alan J. Lipton, Horonobo Fujiyoshi and Raju S. Patil, "Moving Object Classification and Tracking from Real-Time Video" submitted to IEEE WACV 98, 1998. Fourth edition, 2005.
- [2]. Hieu T. Nguyen, Marcel Worring and Rein van den Boomgaard. "Occlusion Robust Adaptive Template Tracking", Intelligent Sensory Information Systems, University of Amsterdam, Faculty of Science, Kruislaan 403, NL-1098 SJ, Amsterdam, The Netherlands..
- [3]. <http://www.cvg.cs.rdg.ac.uk/PETS2001/pets2001-dataset.html>
- [4]. Lewis.J.P et.al. "Fast Normalized Cross-Correlation". Vision Interface, P.120-123, 1995.
- [5]. Karan Gupta and Anjali V.Kulkarni. "Implementation of an Automated Single Camera Object Tracking System Using Frame Differencing and Dynamic Template Matching"
- [6]. Longin Jan Latecki and Roland Mieziako. "Object Tracking with Dynamic Template Update and Occlusion Detection".
- [7]. Original version by Amir Hossein Omidvarnia, October 2007: <http://cswww.essex.ac.uk/mv/allfaces/faces94.zi>
- [8]. Yilmaz, A., Javed, O., and Shah, M., "Object tracking:A survey", ACM Computing Surveys, 38, 4, Article 13, Dec. 2006
- [9]. Xue Mei, Shaohua Kevin Zhouy and Fatih Porikliz "Probabilistic Visual Tracking via Robust Template Matching and Incremental Subspace Update", Mitsubishi electric research laboratories.
- [10]. Wenhui Liao, Yan Tong, Zhiwei Zhu, and Qiang Ji "Robust Object Tracking with a Case-base Updating Strategy" IJCAI-07
- [11]. [http://imagelab.ing.unimore.it/visor/video_videosInCategory - video for traffic surveillance.](http://imagelab.ing.unimore.it/visor/video_videosInCategory-video-for-traffic-surveillance)
- [12]. M.H.Sidram and Nagappa.U.Bhajantri "Normalized Cross- Correlation for Tracking Object and Updating the Template: Exploration with Extensive Dataset" International Journal of Computer Science and Information Security, Vol. 10 No. 02, February 2012, pp 102-112.