OBJECT TRACKING USING DOMINANT SUB BANDS IN STEERABLE PYRAMID DOMAIN

Rikin J. Nayak, Jitendra P. Chaudhari

V. T. Patel Department of Electronics & Communication Engineering, CHARUSAT Space Research and Technology Center, Chandubhai S Patel Institute of Technology, Charotar University of Science and Technology, Changa, Gujarat
e-mails: rikinnayak.ec@charusat.ac.in, jitendrachaudhari.ec@charusat.ac.in
India

Abstract: Video Surveillance system is ubiquitous today for which Object tracking is a challenging task which has many applications in the computer vision area. Sparse representation has been applied in the field of object tracking comprehensively. In the proposed method we have used a steerable pyramid for object tracking using dominant sub bands only as features for reducing computation complexity and to increase efficiency. For steerable, the extraction of coefficient at any orientation can be done using a linear combination of basis filters. Features are characterized according to dimensionality and direction. Experimental results show that the proposed method gives favorable results. For the given videos using the proposed method we have a success rate for tracking is more than 90 percentages.

Key words: object tracking, steerable pyramid, Dominant Sub-bands.

1. INTRODUCTION

Object tracking is the method of estimating an object´s position throughout the video. Tracking is important for surveillance systems like transport systems, military areas, public places and industrial places. The object tracking system consists of three parts [22], Appearance based, motion based and searching based. The tracking method consists of object detection, object representation and object tracking. Numerous object detection, representation and tracking methods have been discussed in [1]. Although momentous progress in the field of tracking still due to challenges like complex shape, illumination change, environment effect, etc. it is a complicated problem [3]. Tracking object in the video sequence is the process of finding and marking a moving object with the time information using a camera. It is a significant task in many applications like human computer interaction, surveillance system, traffic system, etc. [2]. This paper is focused on searching the target based on its sparse information as features. Steerable pyramid transform can decompose image into multiple scales, multiple orientations, it is widely used in image processing applications. An image pyramid is the hierarchical image representation at different resolutions. The idea is to generate several images at different scales and different orientations that characterize the response of a number of filters. Object tracking using steerable uses the energy of steerable coefficients as a feature template. Steerable pyramid transform is rotation invariant and translation invariant so that it can represent of position or orientation of image more sparsely. The goal of feature selection and extraction is to obtain features which maximize the similarity of interested objects and dissimilarity of non-interested objects. It helps in understanding the matching between extracted reference features and target features. It also results in good efficiency in terms of time and cost [7]. In this algorithm searching is based on feature matching with fixed distance however one can introduce prediction using Kalman or such other methods to increase the speed of the algorithm. The paper is organized as follows: In section II we present the related work and concept of steerable pyramid transform. In section III we have described the proposed Method, in section IV we have described experimental results for tracking and in section V we conclude the paper.

2. RELATED WORK

Recently many tracking methods have been proposed in different literature. Different tracking strategies have been discussed in [1]. However Most of these work well for fixed scenarios or conditions. The object tracking algorithms can be generally divided into four groups: region based [25], contour based [26], model based [27] and feature based [28] algorithms. Some of the popular methods for object tracking are based on Kalman filter [29, 30], Particle filter [31], and Mean Shift [32] trackers. Transform based tracking algorithm proposed by [6, 8, 21, 24]. Wavelet and other such frequency transform based object tracking methods are available but they have shift and rotation sensitivity issues, and poor in directionality [8]. In [6, 8], they have used Curvelet transform for object tracking by selecting the energy of coefficients as features and predict the location of the object in current frame with help of previous three frames and basic equations of straight line motion. They have used the energy as a matching parameter for object tracking. Tracking algorithm tracks the object in the next frame based on to the value of the current velocity measured from the previous three frames. In [23] we have used dominant curvelet sub bands for object tracking. Tracking can also be used in surveillance [35] as well as to track animal, vehicle or player in game [34]. However, to track object which changes its shape or appearance is very challenging. Application of tracking can be extended to smart surveillance where the behavior of tracked object can be detected and further action could be taken. Such a system can apply to detect abandoned object [33]. Deep Learning and Convolution neural network (CNN) indeed improve the tracking capability however it also adds complexity in computational requirements.

Steerable transform has the advantage of being aliasing free [9] because it uses a low pass filter at each decomposition stage and can be designed to produce any number of orientation bands. Oriented filters are used in many applications like object reorganization [10], enhancement [11], Palm print reorganization [12], denoising [13], orientation estimation [14], Image fusion [20], Brain disease diagnosis [15], Texture classification [18], Texture image retrieval [19] etc. It has an advantage of overcoming the limitations of orthogonal separable wavelet decompositions [16].

3. STEERABL PYRAMID TRANSFORM

In a steerable pyramid, Image is first decomposing in two sub bands: a low pass and high pass using filters. This low pass band continues divided in different band pass coefficients at different orientation using band pass filters, followed by a low pass band sub sampled by a factor of two. This is as shown in fig 1.

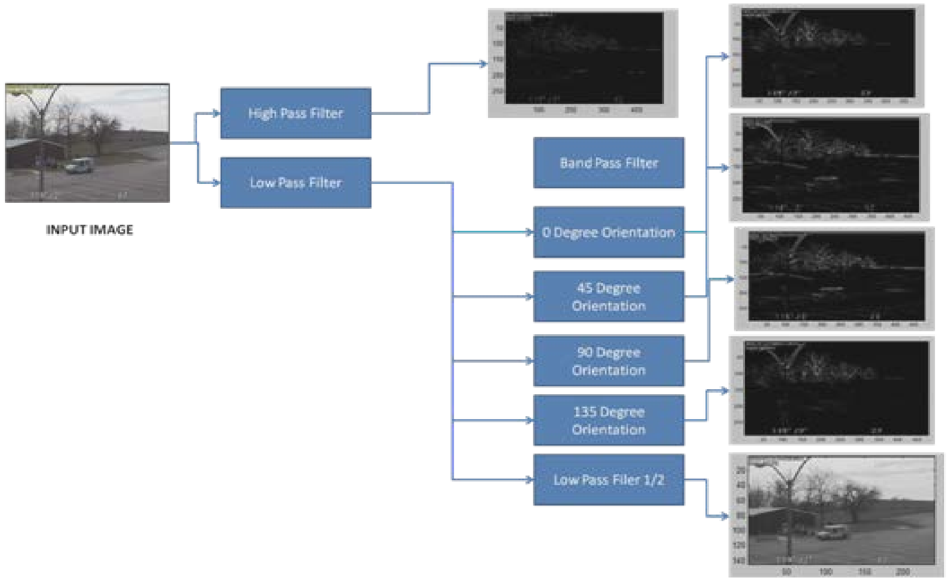

Fig. 1. Image decomposition for Steerable pyramid transform with four band pass

Steerability means wavelet image can be rotated at any orientation by using a linear combination of the primary set of components [9, 17]. One drawback of the steerable pyramid is its computation efficiency: the resulting transform is considerably over complete by a factor $\frac{4K}{3}$. Where K is the number of orientation bands. For the understanding case of steerable pyramid Consider a two dimensional symmetric filter g with Gaussian impulse response is [9],

$$G(x,y) = e^{-(x^2+y^2)} \tag{1}$$

Now the first x derivative of a Gaussian $G_1^{0^0}$ is

$$G_1^{0^0} = \frac{d}{dx}G(x,y) = -2xe^{-(x^2+y^2)} \tag{2}$$

and same function rotated 90^0 is

$$G_1^{90^0} = \frac{d}{dy}G(x,y) = -2ye^{-(x^2+y^2)} \tag{3}$$

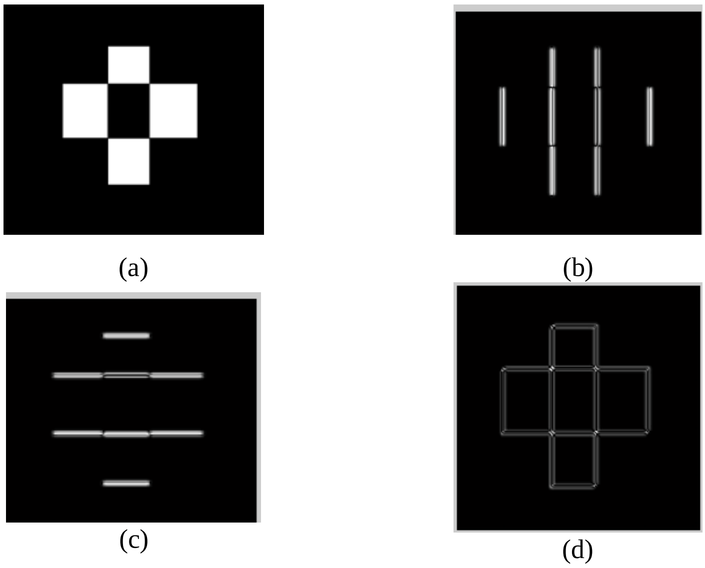

Fig. 2. Steerable pyramid Coefficients (b), (c) and (d) for original image (a)

Now using these two primary set of basis filter we can obtain coefficient for any orientation using linear combination of these basis filters.

Filter $G_1^{\Theta^0}$ for particular orientation band at Θ can be synthesized by taking linear combination of $G_1^{\Theta^0}$ at 0 and 90:

$$G_1^{\Theta} = \cos(\Theta)G_1^{0^0} + \sin(\Theta)G_1^{90^0} \tag{4}$$

This is shown in fig 2. Here figure (b) and (c) are steerable coefficients for orientation 0 and 90 degree. Fig (d) shows steerable coefficients for orientation θ = 45 degree.

$$\mathrm{Im}(d) = \frac{1}{\sqrt{2}}\mathrm{Im}(b) + \frac{1}{\sqrt{2}}\mathrm{Im}(c) \tag{5}$$

4. TRACKING ALGORITHM

The Steps for the algorithm are as below. The algorithm is based on, described by the author in [21].

Steps of the algorithm:

Step 1: Object selected by the user

Step 2: Extract the features of the selected object using coefficients of the steerable filter.

Step 3: Search the object using a fixed window by matching the features with the features extracted in step 2

Step 4: Identify the minimum difference in features with primary extracted features from step 2

Step 5: Create a bounding box around the identified location where the difference is minimum and search the object in next frame using step 3

Program Realization:

For Frame = 1:

Apply Contrast Enhancing on Frame

Target template is selected by the user in the first frame

Apply the Steerable Pyramid transform and select dominant coefficients

Calculate the energy for the first frame using

$E = \sum_{(i,j) of bounding box} |w(i,j)|^2$

For Frame = 2: n:

Apply Contrast Enhancing on Frame

Using a fixed window divide the frame into n block

Apply the Steerable Pyramid transform on each block; Select the same dominant coefficients as in first frame

Calculate the energy using $Enew = \sum_{(i,j) of bounding box} |w(i,j)|^2$

Match the template using abs(E-Enew) between two frames.

Create a box where abs(E - Enew) is minimum.

W(i, j) is steerable pyramid sub bands at (i, j) point. For steerable any orientation band can be synthesized by using a linear combination of primary set basic sub bands. Then the object's next position is estimated by measuring the similarity between two vectors of energy. Here we have considered dominant coefficients instead of all steerable coefficients. In proposed method we have presumed that object is not tilted beyond a certain limit however with change in orientation we can use a dynamic selection of dominants coefficients (i.e. Maximum energy). For tracking an object in video sequences we have considered a video as a combination of thousands of different frames.

5. EXPERIMENTAL RESULTS

5.1 Case Study 1

Fixed camera: Here in the video, the man is walking on the street (Fig.3). In this video object as well as background is moving. Here a frame size of the video is 160x120**.** We have applied our algorithm to 80 frames. The results are as shown in fig 3. For measuring the accuracy we have used centroid parameter compare with actual centroid measured manually.

Frame 1

Frame 4

Frame 5

Frame 6

Frame 9

Frame 10

Frame 51

Frame 52

Frame 53

Frame 54

Fig. 3. Tracked Object

Fig 4 shows a graph for the comparison of the centroid of a tracked object with the actual centroid. From the graph it is shown that the method gives very close results with actual.

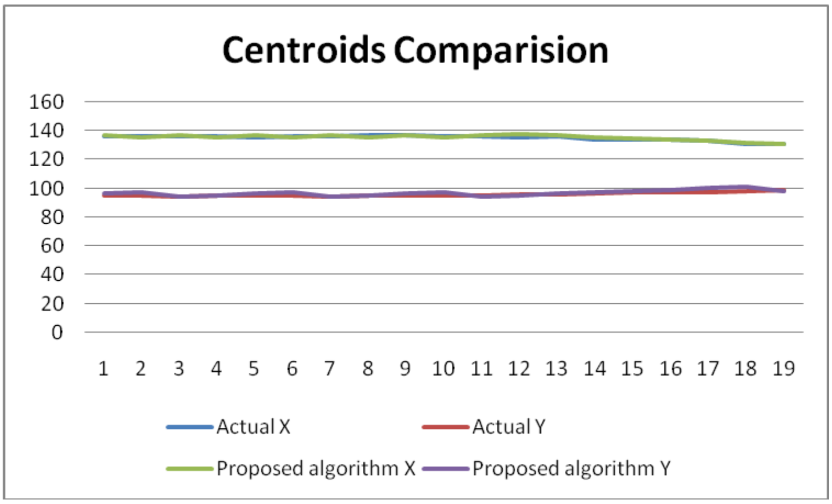

Fig. 4. Center location for man video

Table 1 shows actual centroid and centroid with the steerable based method. Here actual centroid computed manually for selected frame and centroid with the proposed method is centroid from the tracked frame. Success rate gives closeness in percentage.

Table 1. Center location of man object video

Frame No	Actual		Proposed algorithm		Success rate (%)	
	X	Y	X	Y	X	Y
1	136	95	136.5	96	99.63235	98.94737
2	136	95	135.5	97	99.63235	97.89474
3	136	94.5	136.5	94	99.63235	99.4709
4	136	95	135.5	95	99.63235	100
5	135	95	136.5	96	98.88889	98.94737
6	136	95	135.5	97	99.63235	97.89474
7	136	94.5	136.5	94	99.63235	99.4709
8	136.5	95	135.5	95	99.2674	100
9	136.5	95.5	136.5	96	100	99.47644
10	136	95	135.5	97	99.63235	97.89474
11	136	95.5	136.5	94	99.63235	98.42932
12	135	96	137.5	95	98.14815	98.95833
13	135.5	96	136.5	96	99.26199	100
14	134	96.5	135.5	97	98.8806	99.48187
15	134	97.5	134.5	98	99.62687	99.48718
16	133.5	97	133.5	99	100	97.93814
17	133	97.5	132.5	100	99.62406	97.4359
18	131	98	131.5	101	99.61832	96.93878
19	131	98.5	130.5	98	99.61832	99.49239
20	130	97	129.5	99	99.61538	97.93814

For measuring the performance we have used “success rate” as a parameter. It can be defined as

$$\text{Success rate (in \%)} = \frac{Centroid_{actual} - |Centroid_{actual} - Centroid_{proposedmethod}|}{Centroid_{actual}} * 100\%$$

5.2 Case Study II

Moving camera: In the second case we have used a video in which camera is moving. In video we have tracked a player in football match. The size of the frame is 640 x 360. We have tracked the object in 55 frames. Fig 5 shows tracked results.

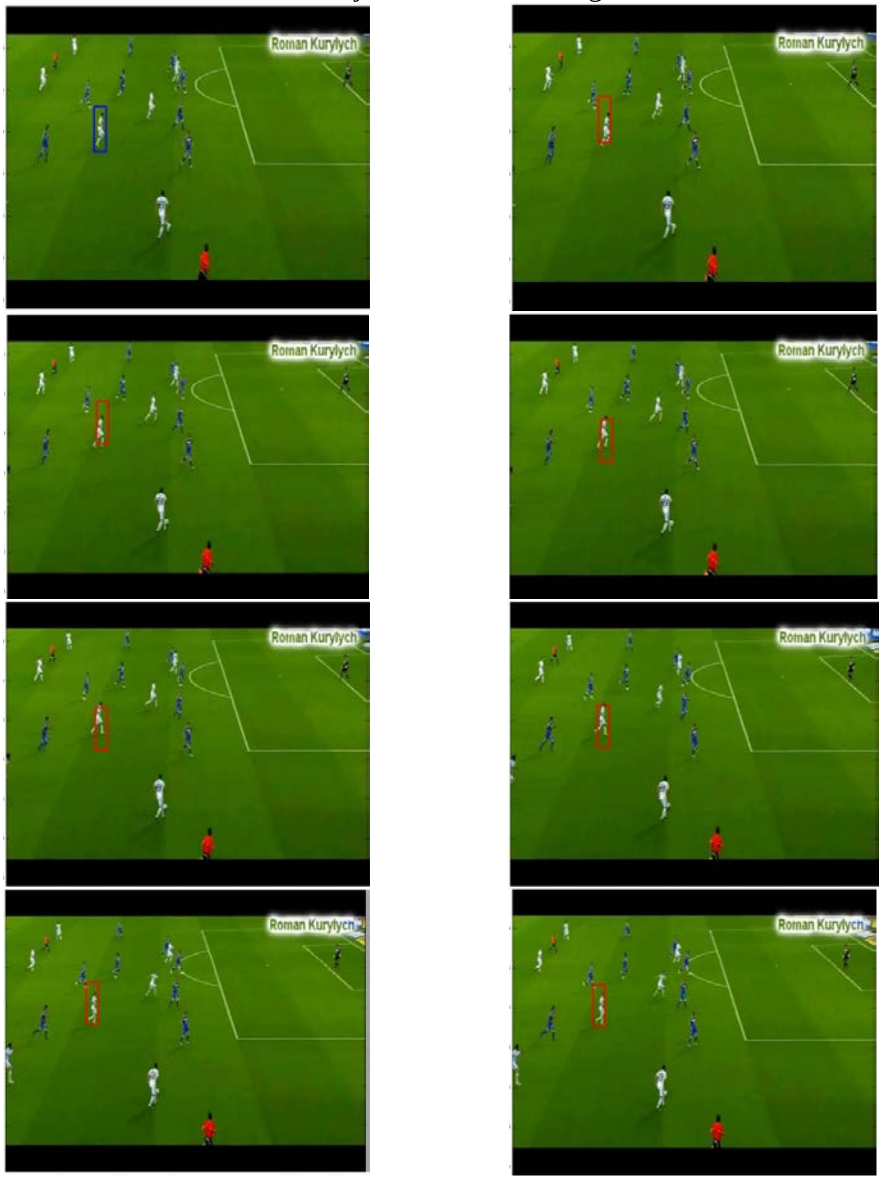

Fig. 5. Player track Frame No 1 to 10, Blue box shows selected object, In Red it shows tracked object

5.3 Case Study III

Moving camera: Moving truck is tracked for this scenario. The result is presented in Fig. 6.

Fig. 6. Truck Track

We have also compared the proposed algorithm in which we have used only dominant coefficients with all bandpass coefficients of the steerable pyramid. Here the algorithm applied for both cases for frame no 51 to 85. In the case of all bandpass it missed object after frame no 71. In the case of the proposed method due to sparse representation of image with dominant sub-bands, it works well for tracking compare to all sub-bands. We have compared the simulation results with curvelet and steerable pyramid with all bands at an appropriate scale. Table 2 shows the comparison.

Table 2 Comparison of Steerable with one BP, all BP and Curvelet

	Curvelet	Steerable with all bandpass	Steerable with dominant bandpass
No of frames	35	35	35
Tracked frames	2	21	35
Scale	4	1	1

6. CONCLUSION

We have developed and demonstrate a steerable pyramid transform-based method for tracking an object in the video sequence. For tracking we have used dominant sub-bands for extracting templates. For the given dataset we are able to track the object with more than 90 percent success rate. Results show that our method works well for different case study videos since dominant sub-bands give sparse representation for the object. We have also compared the proposed method with all band pass coefficients of steerable pyramid at one scale. In the case of all bands pass it missed object totally after 21st frame. The method has been tested in various cases. Here for the tracking we used simple template matching based searching methods however one can add prediction with changeable searching distance to reduce the timing of the algorithm.

REFERENCES

[1]. Yilmaz, Alper, Omar Javed, and Mubarak Shah. Object tracking: A survey. *ACM computing surveys*, 38.4, 2006, 13 p..

[2]. Li, X., Hu, W., Shen, C., Zhang, Z., Dick, A. and Hengel, A.V.D. A survey of appearance models in visual object tracking. *ACM transactions on Intelligent Systems and Technology (TIST)*, 4(4), 2013, p.58

[3]. J. Gao, T. Zhang, X. Yang and C. Xu. Deep Relative Tracking, *IEEE Transactions on Image Processing*, vol. 26, no. 4, April 2017, pp. 1845-1858, doi: 10.1109/TIP.2017.2656628

[4]. Luo, W., Xing, J., Zhang, X., Zhao, X. and Kim, T.K. Multiple object tracking: A literature review. *ArXiv preprint* arXiv: 1409.7618, 2014.

[5]. L. Wei and L. Xingwei, Multiple object tracking based on modified algorithm of GMMCP tracker, *2016 IEEE International Conference on Signal and Image Processing (ICSIP)*, Beijing, 2016, pp. 11-15. doi: 10.1109/SIPROCESS.2016.7888214

[6]. S. Nigam and A. Khare. Curvelet transform-based technique for tracking of moving objects, *IET Computer Vision*, vol. 6, no. 3, pp. 231-251, May 2012. doi: 10.1049/iet-cvi.2011.0023.

[7]. Khaire, Utkarsh Mahadeo, and R. Dhanalakshmi. Stability of Feature Selection Algorithm: A Review. *Journal of King Saud University-Computer and Information Sciences*, 2019.

[8]. Binh N.T., Khare A. Object Tracking Of Video Sequences In Curvelet Domain, *International Journal Of Image And Graphics*, World Scientific Publication, vol. 11, No 1 , 2011, pp. 1-20.

[9]. W. T. Freeman and E. H. Adelson. The design and use of steerable filters, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, 1991, pp. 891--906.

[10]. Xiantong Zhen and Ling Shao, "Spatio-temporal steerable pyramid for human action recognition," *10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*, Shanghai, 2013, pp. 1-6.

[11]. F. Denis, F. Davignon and A. Baskurt. Steerable pyramid for contrast enhancement and directional structures detection, *11th European Signal Processing Conference*, Toulouse, 2002, pp. 1-4.

[12]. D. Tamrakar and P. Khanna. Multispectral palmprint recognition using steerable filter, *9th International Conference on Industrial and Information Systems (ICIIS)*, Gwalior, 2014, pp. 1-5.

[13]. Ehsaeyan, E. An Improvement of Steerable Pyramid Denoising Method. *Iranian Journal of Electrical and Electronic Engineering*, 12.1, 2016, pp. 35-41.

[14]. Nestares, Oscar. *Automatic dominant orientation estimation in text images based on steerable filters*. U.S. Patent No. 8,139,894. 20 Mar. 2012.

[15]. Kale, Vandana V., Satish T. Hamde, and Raghunath S. Holambe. Brain disease diagnosis using local binary pattern and steerable pyramid. *International Journal of Multimedia Information Retrieval*, 2019, pp. 1-11.

[16]. H. Kaur and K. Kaur. Image Forgery Detection using Steerable Pyramid Transform and Lab Color Space, *International Journal of Advanced Research in Computer Science and Software Engineering*, vol.-5, No 8, August 2015.

[17]. M. Unser, N. Chenouard and D. Van De Ville. Steerable Pyramids and Tight Wavelet Frames in $L_{2}({BBR}^{d})$, *IEEE Transactions on Image Processing*, vol. 20, no. 10, pp. 2705-2721, October 2011. doi: 10.1109/TIP.2011.2138147

[18]. Dash, Sonali, and Uma Ranjan Jena. Texture classification using steerable pyramid based Laws' masks, *Journal of Electrical Systems and Information Technology*, 4.1, 2017, pp.185-197.

[19]. Pandiselvi, G., et al. Steerable Pyramid Decomposition-Rotation & Scale invariant texture image retrieval. *International Journal of Computer Science and Network Security* (IJCSNS), 16.3, 2016, p. 114.

[20]. Xu, Shengnan, and Zhenhua Li. Remote sensing image fusion using intensity-hue-saturation transform and steerable pyramid transform, *Chinese Automation Congress (CAC)*. IEEE, 2017.

[21]. A. Khare and U. S. Tiwary. Daubechies Complex Wavelet Transform Based Moving Object Tracking, *2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing*, Honolulu, HI, 2007, pp. 36-40.

[22]. L. Wen, Z. Cai, Z. Lei, D. Yi and S. Z. Li. Robust Online Learned Spatio-Temporal Context Model for Visual Tracking, *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 785-796, 2014. doi: 10.1109/TIP.2013.2293430.

[23]. Nayak, R., Bhavsar, J., Chaudhari, J., Mitra, S. K. Object tracking in curvelet domain with dominant curvelet subbands. *Energy*, 16(775.86), 2012, p=32

[24]. Prakash, Om, and Ashish Khare. Tracking of moving object using energy of biorthogonal wavelet transform. *Chiang Mai J. Sci.* 42.3, 2015, pp. 783-795.

[25]. Lee, Hyemin, and Daijin Kim. Salient region-based online object tracking. *IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2018.

[26]. Kanagamalliga, S., and S. Vasuki. Contour-based object tracking in video scenes through optical flow and gabor features. *Optik*, 157, 2018, pp. 787-797.

[27]. Zhou S. K., Chellappa R. and Moghaddam B. Visual tracking and recognition using appearance-adaptive models in particle filters, *IEEE Transactions on Image Processing*, 13(11), 2004, pp. 1491–1506.

[28]. Kang, Kyuchang, et al. Invariant feature based object tracking using discrete dynamic swarm optimization. *Etri Journal*, 39.2, 2017, pp. 151-162.

[29]. Ali A and K Terada 2009 A Framework for Human Tracking using Kalman Filter and Fast Mean Shift Algorithms, *IEEE Int'l* Conf. *on Computer Vision Workshops (ICCV Workshops)*, Kyoto, 2009.

[30]. Zhao J, W Qiao and G-Z Men. An approach based on mean shift and KALMAN filter for target tracking under occlusion, *Int. Conf. on Machine Learning and Cybernetics*, Volume 4, Baoding, 2009.

[31]. Tang D and Y-J Zhang. Combining Mean-Shift and Particle Filter for Object Tracking, *Sixth International Conference on Image and Graphics (ICIG)*, Hefei, Anhui, 2011.

[32]. Liu, Jinhang, and Xian Zhong. An object tracking method based on Mean Shift algorithm with HSV color space and texture features. *Cluster Computing* 22.3, 2019, pp. 6079-6090.

[33]. Luna, Elena, et al. Abandoned object detection in video-surveillance: survey and comparison. *Sensors*, 18.12, 2018, p. 4290.

[34]. D. Israni, H. Mewada. Feature Descriptor Based Identity Retention and Tracking of Players Under Intense Occlusion in Soccer Videos, *International Journal of Intelligent Engineering and Systems*, vol. 11, no. 4, 2018, pp. 31-41.

[35]. D. Israni, H. Mewada. Identity Retention of Multiple Objects under Extreme Occlusion Scenarios using Feature Descriptors, *Journal of Communications Software and Systems*, vol. 14, no. 4, 2018, pp. 290-301.

Information about the authors:

Rikin J Nayak – M.Tech in Communication System Engineering, Assistant Professor at Charotar University of Science and Technology, Changa, Anand, CSRTC, V.T.Patel Department of Electronics & Communication Engineering, Chandubhai S Patel Institute of Technology. His area of interest is Computer Vision and Deep Learning.

Dr. Jitendra P Chaudhari – Ph.D. in Electronics & Communication Engineering, Associate Professor at Charotar University of Science and Technology, Changa, Anand, CSRTC, V.T.Patel Department of Electronics & Communication Engineering, Chandubhai S Patel Institute of Technology. His area of interest is Image Processing and Power Electronics.

Manuscript received on 09 January 2020