A Novel Approach for Detection and Tracking of Vessels on Maritime Sequences

Disha Patel
U and P U. Patel Department of Computer Engineering, CSPIT, CHARUSAT, Changa, Gujarat, India
Email:pateldisha1408@gmail.com
Dippal Israni
U and P U. Patel Department of Computer Engineering, CSPIT, CHARUSAT, Changa, Gujarat, India
Email : dippalisrani.ce@charusat.ac.in
Mrugedrasinh Rahevar
U and P U. Patel Department of Computer Engineering, CSPIT, CHARUSAT, Changa, Gujarat, India
Email :mrugendrarahevar.ce@charusat.ac.in

---**ABSTRACT**--

Object detection, classification and tracking are prime components in all computer vision application. The research problem here is to detect, classify and track small objects (such as ships, boats etc.) on maritime scenario. Main purpose of object detection in maritime is to secure the country from various rocket launchers, sea side firearms. According to today's scenario, security is very important in maritime applications. This paper showcases experiment results for object detection using Speeded Up Robust Features (SURF), Binary Robust Invariant Scalable Key points (BRISK), Lukas-kanade on standard dataset IPATCH PETS 2016. However, the state of the art methods limits in handle camouflage scenario. This paper proposes an adaptive Lucas-kanade approach that handles such scenarios. The proposed approach utilizes interaction of arithmetic mean and histogram equalization with optical flow (Lucas-Kanade) approach to resolve camouflage. Finally, the proposed approach is evaluated using standard parameters such as recall, precision and f1 score. The performance measures depict that the proposed approach outperforms state of the art trackers.

1. INTRODUCTION

The start of 21st century has been a resurgence of the sea theft [1]. In today's scenario where the security is very important to maritime, the applications of image processing helps to detect sea skimming missiles located at long ranges. It can be beneficial for the Infrared Search and Track (IRST) technology to secure the country from sea skimming rockets, seaside firearms, and various rockets launchers. It is utilized in several areas such as aerial defense, accurate guidance, remote sensing of satellite images and many more [2].

For littoral nations, maritime surveillance has always been a crucial segment for law enforcement and also for environment protection [3].Object detection, classification and tracking are prime components of image processing. The problem of detecting small target is of high research interest, as it has complex background that includes low contrast, noise, small size and obscure shape. Hence, it has been a challenge to focus on maritime problems for many military experts in recent year. A few occasions are helpful pointers for an early warning of a sea risk [1].

Object detection is an integral segment of video surveillance scenarios. It aids as the essential enabler for crucial jobs such as detection of moving object, tracking objects and classification of object [4]. Object detection techniques are primarily classified in three categories (i) temporal (ii) spatial and (iii) integration of spatial and temporal. To detect moving object in a visual frame, the most reliable cue is motion details of an object [5].

Object detection techniques are mostly utilized in remote sensing, especially for on board satellites that provides a repetitive and consistent view of the earth [6]. This is very important for monitoring the earth system and the impact of human activity on the earth. Long distance video capturing is used in Surveillance [7] that plays a key role in monitoring the behavior, activities and other changing information of objects in order to influence, direct and protect them. For surveillance cameras, visual object tracking is an important part of video processing [8].

This paper describes various methods for detecting, classifying and tracking objects. The related techniques are depicted in section 2. The novel approach is explained in section 3. In section 4, simulation results and are showcased. In section 5 the results are analyzed using standard parameters. Finally, section 6 represent conclusion.

2. RELATED WORK

As the beginning of the 21st century has seen a new interest in sea-related piracy, surveillance in maritime has become subject of extreme importance. In recent years, several studies have discussed small infrared target detection techniques. This portion surveys the state-of-the-art approaches regarding their use of information. For example, background information, target data, decision information and visual setting [9].

An object is always characterized as a point of interest. It includes vehicle on road, fish inside sea, boats on an ocean, planes in the sky, human walking on a road, and players in sports [10]. These are important for tracking in their own specific domain.

In maritime surveillance, there are many challenges that include wide area detection and tracking, Night sequences, weather or other environmental issues, differing object size [11].

In order to provide robustness, Bloisi et al [12] have used Haar-like classifier to detect boat amongst totally still boats anchored off the coast. They have represented an automated sea-related secretly recording system. The system can provide the user worldwide view adding a visual dimension to AIS data.

Panchal payal et al [13] has discussed a basic method that includes background subtraction, frame differencing and shaped based methods. Object tracking and detection is most trending research area and it is used to detect the motion of various objects of a video sequence.

Stefan Leutenegger et al [14] proposed a new method to detect, describe and match key points. A comprehensive benchmark evaluation reveals the adaptive BRISK method. It has high performance like in state of the art algorithms, although at a very low cost (in cases order of magnitude faster than SURF). To increase the computational speed, a new FAST based detector with a bit string descriptor assembled from intensity comparisons is applied. This is obtained through specific samples of each key point area.

Himaniparekh et al suggested background subtraction method [15]. This approach is categorized into approximate median and Gaussian of mixture. This approach does not require sub sampling of frame to create an adaptive background model and it also requires less memory. Other suggested approach is frame Differencing method. The suggested approach is less complex and efficient for static background. However, this approach requires a background without moving object. Background subtraction techniques are mainly used in many real-time video surveillance applications in the field of video surveillance research [16].

Remya Ramachandran et al. suggested SURF (Speeded Up Robust Features) Features for object detection [17]. This approach is serval times faster than traditional SIFT. However, this approach is not suitable to rotation and illumination changes. Dhara Patel et al suggested Optical Flow (Lucas-Kanade) method that has high computational speed and accurate time derivatives [18]. However, in this approach boundaries of moving object are not precisely estimated.

Speeded up robust features (SURF) is a standard algorithm that is utilized mostly in computer vision [19]. SURF was initially presented by Herbert at al. This algorithm is primarily categorized in three different parts: point detection of interest, description of local neighborhood and matching. The SURF algorithm is used for object recognition tasks [20].

2.1 Frame extraction from input video
In this step the temporal sequences are extracted from the video sequences for detection of objects.

2.2 Interest point detection
This paper uses SURF for extracting features from a frame. SURF uses square-shaped filters as a Gaussian smoothing approximation. Filtering the image with a square is much faster when the integral image is used as shown in eq. 1 and eq. 2.. In SURF, blob detector based on the Hessian matrix is used to extract points of interest [19].
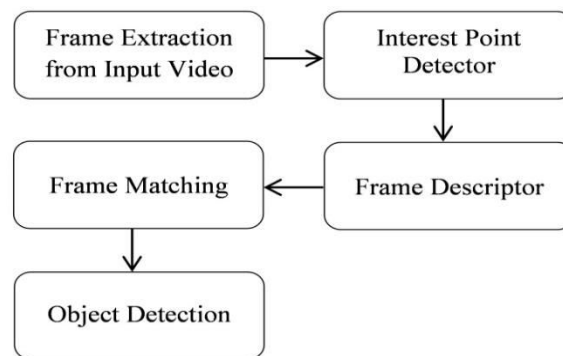

Fig 1. SURF algorithm

Equation (1) depicts summation of the original image's ROI that can be quickly assessed utilizing the integral image for evaluation at the ROI's four corners [19].

$$S(a,b) = \sum_{i=0}^{a} \sum_{j=0}^{b} I(i,j) \qquad (1)$$

Given a point A = (a, b) in an image I, the Hessian Matrix H (A, σ) in A at scale σ is defined as follows:

$$H(A,\sigma) = \begin{bmatrix} L_{aa}(A,\sigma) & L_{ab}(A,\sigma) \\ L_{ab}(A,\sigma) & L_{bb}(A,\sigma) \end{bmatrix} \qquad (2)$$

Here, Laa (A, σ) represents the Gaussian Second order derivative convolution with the image I, the same is followed for Lab(A, σ) and Lbb (A, σ) [21].

2.3 Feature descriptor
A feature descriptor's goal is to estimate a unique and robust description of an image feature. The descriptor's dimensionality consumes effect on computational complexity as well as robustness accuracy for matching points. The first phase involves of setting a reproducible orientation b from a circular area within the interest point. After that a square segment is formed that is attached to the selected orientation [19].

2.4 Feature Matching
Matching pairs are estimated using the comparison of descriptors obtained from various images [19].

2.5 Object detection

Based on the comparison of extracted features an object is detected. This process continues in an iterative manner till all the objects are detected in an image.

Motion based on the optical flow method are detected using flow vector features. These functions detect motion objects in a visual frame over time that detect moving pixels. Meyer et al. 1998 [22] [23] proposes active rays to extract the articulated rays, which is a contour - based tracking algorithm for computing the field of displacement vectors. These vectors of displacement are intended for gait analysis and even in the camera motion scenarios moving objects can be distinguished with the use of optical flow based technology separately. All flow computing methods are generally computationally challenging and subtle to errors as well. These methods are not used in real - time simulation on visual frames. Barron's 1994 [24] proposed another variant of optical flow.

3. PROPOSED APPROACH

The previous section discussed serval approaches for object detection and tracking. This paper proposes an adaptive Lucas-Kanade approach for object detection and tracking on maritime visual scenarios. The flow of proposed approach is as discussed in fig 2.
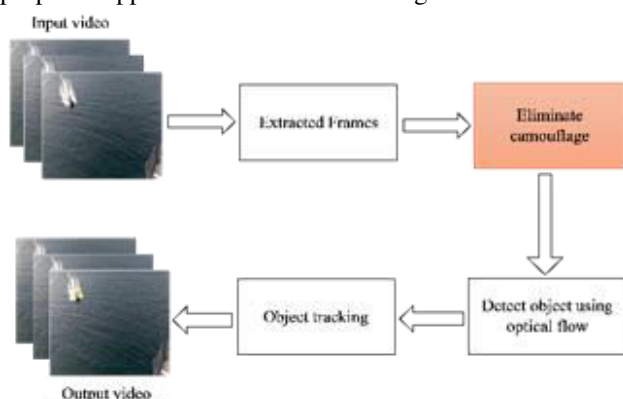

Fig 2. Proposed object detection approach

3.1 Input Video and Frame Extracted

In the preliminaries stage of object detection is to extract frames from the video sequence for processing. Input video is first converted into frames during the process. The input video frame is partitioned off into little blocks. This video sequence capture the by long rang object is detected in different camera view. Standard dataset IPATCH PETCH 2016 is used here to test the proposed approach. The description of input video is showcased in Table 2. The number of frames in video depends on the size of the video. These frames occupy large memory space. Hence, frames are extracted from video sequence as and when required. These frames are utilized as an input for object detection.

3.2 Eliminate camouflage

There are scenarios where objects are hardly different from its background and surroundings. This makes the tracking and detection process challenging. This scenario

is known as camouflage. The proposed mitigation process is showcased in fig 3.

- **Arithmetic Mean**

The arithmetic mean is one of several different types of average descriptor used to describe a data sets center or representative value.

- **Histogram stabilizes**

Equalization of histograms is a technique of contrast enhancement in image processing in a spatial domain using image histogram. Histogram equation usually increases global contrast of processing image [25]. This approach is helpful for noise background.
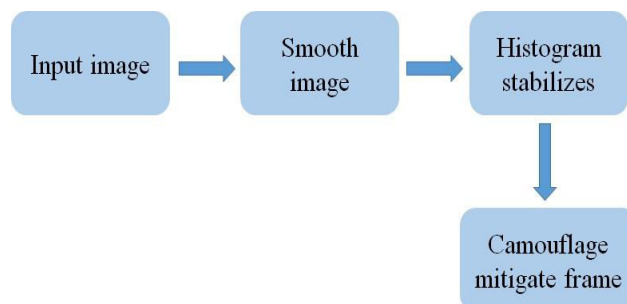

Fig 3. Camouflage elimination approach

3.3 Detect object using Optical flow

For optical flow detection computing the motion of every pixel in image sequences is very challenging. Hence, optical flow based approach is used to resolve this issue in proposed approach. Bruce D. Lucas and Takeo Kanade suggested a differential technique for estimating optical flow. This technique utilized least square criterion, for all the pixels in a neighborhood to solve the equating of basic optical flow [26] [27]. Hence, in the proposed approach the adaptive Lucas Kanade is proposed. This approach integrate camouflage estimation and optical flow based Lucas-Kanade method to detect the object. The advantage of the proposed approach is it detects the object even in noisy environments. In maritime environments sea wave is also detected as an object. However, the proposed approach eliminates this noise and precisely detects only the real motion.

3.4 Object Tracking

Optical flow technique is utilized as a feature descriptor for tracking objects. Object trajectories are signified by edges and objects are traced in each frame based on centroid of object. In tracking multiple objects, the overlap of edge details or the construction of bounding box is estimated to highlight the movement

3.5 Output video

Once the objects are detected in the frame, they are annotated with the bounding box. The result of the proposed approach on every frame sequence are combined and the output video sequence is generated.

4. EXPERIMENT RESULTS

This section showcases the evaluation the proposed approach on standard dataset. This chapter shows

simulation results of object detection of several scenarios of boats on multiple video sequences.

4.1 Dataset Description

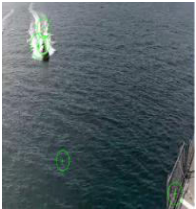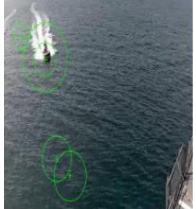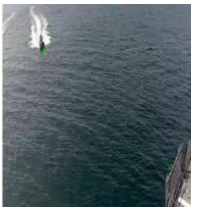The dataset that is used for evaluation in this paper is PETS 2016 (Performance Evaluation Tracking and Surveillance) [28]. The proposed algorithm's performance is estimated on aIPATCH PETS 2016public dataset. The dataset covers six sequences with several challengessuch asvariation of illumination, shadows and different moving direction. Also in this dataset the subjects often change direction, accelerate and loiter. For every scenario there are 2 video sequences.
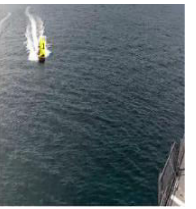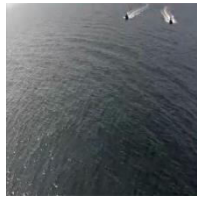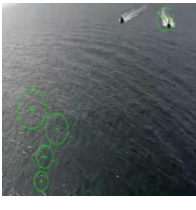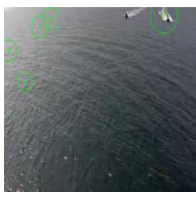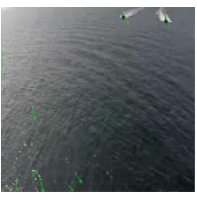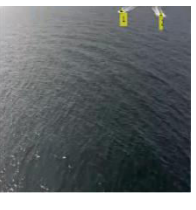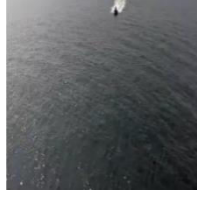
Table 1. Dataset Description

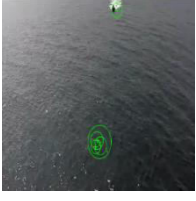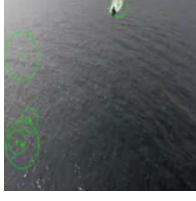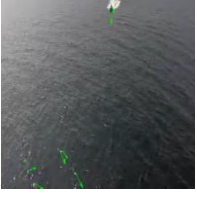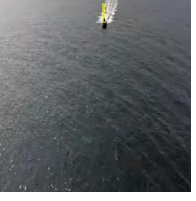| Dataset name	Sc3_Tk1		Sc3_Tk3
Video	Sequence 2	Sequence 4	Sequence 4
Dataset Provider	PETS 2016	PETS 2016	PETS 2016
Dimension (Height*Width)	303*424	305*427	295*421
No of Frame present	4761	3746	5164
No of frame use for processing	620	491	411
Frame rate (frames/second)	30	30	30

4.2 Experiments and Simulation

Table 2. Sc3_Tk1 and Sc3_Tk3 of the IPATCH PETS dataset using different algorithm implementation

Dataset Name	Video	Input Image	SURF	BRISK	Lucas-Kanade	Proposed Approach
Sc3_Tk3	Sequence 2					
	Sequence 4					
Sc3_Tk3	Sequence 4					

The first column represents standard dataset boat loitering. Second column represents the sequence of the dataset utilized for simulation. The third column showcases the input images and the fourth, fifth and sixth column is describing the resultof state-of-the-art approaches. The last column represents the result of proposed approach. The sequence utilized for simulation results covers different scenarios of object position. In sequence 2 of Sc3_Tk1,

object is too near to the camera. In sequence 4 of same dataset, multiple objects are present. Sequence 4 of Sc3_Tk3 covers object that is too far of the camera. Hence, the section of sequence is based on covering all different aspects of object position. The test results of the proposed tracker on the PETS 2016 maritime dataset are presented in this section. A common maritime data comparing their detection performance in maritime scenes was evaluated for the adaptive approach proposed.

5. PERFORMANCE EVALUATION

This paper utilizes standard parameters used for evaluation of the proposed approach. Precision and recall are important element for performance evaluation. To compute precision and recall, some parameters such as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are required.

True Positive represent an outcome which detectobject if the object is existing in ground truth.

True Negative represent an outcome which not detects object if the object is not existing in ground truth.

False Positive represent an outcome which detects the object if the object is not existing in ground truth.

False Negative represent an outcome which not detects the object if the object is existing in ground truth.

5.1 Precision

Precision can bemeasured as the ratio of currently positive predicted observations to the total positive predicted observation [29].

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

It is a relation between TP and the total number of TP and FP.

5.2 Recall

Recall can be measured as correctly anticipated observation ratio to actual observations [29].

$$Recall = \frac{TP}{TP+FN} \quad (4)$$

It describes the relationship between true positive(TP) and the total number of false negative(FN) and true positive(TP).

5.3 F1 Score

The F1 score can be estimated as the weighted average of recall and precision [30].

$$F1score = 2*(Recall*precision) /Recall + precision \quad (5)$$

Table 3. Parameters results

Dataset Name	Video	Algorithm	Precision	Recall	F1 score
Sc3_Tk1	Sequence 2	SURF	0.1129	0.4084	0.1779
		BRISK	0.0558	0.1610	0.0829
		Lucas-kanade	0.2509	**0.8565**	**0.3881**
		Proposed approach	**0.7831**	0.2485	0.3773
	Sequence 4	SURF	0.2090	0.4573	0.2869
		BRISK	0.1485	0.2744	0.1927
		Lucas-kanade	0.2606	**0.9764**	0.4113
		Proposed approach	**1.7190**	0.3710	**0.6103**
Sc3_Tk3	Sequence 4	SURF	0.1821	0.5000	0.2669
		BRISK	0.2443	0.7185	0.3647
		Lucas-kanade	0.1083	**0.9019**	0.1934
		Proposed approach	**0.7726**	0.7926	**0.7824**

As per the outcomes showcased in Table 3, the proposed approach shows efficient results. In this table results of precision, recall and f1 score are showcased. The measures depict that the precision of proposed approach is high in all sequences. However, recall measure of proposed approach is less. In such circumstances the third parameter f1 score depict the efficiency of the proposed approach.

6. CONCLUSION

Object detection, classification and tracking techniques are used to detect, track and classify missiles and vessels in maritime video and image sequences. Various methods are defined and described in this paper. Amongst them the main focus is on automatic maritime surveillance and visual target detection as it is normally used by researchers for defense purpose. Moreover, the data set used for

experimentation is also mentioned. Here, an adaptive method is proposed to remove noise and small moving multiple objects detected. This paper showcases simulation results of state of the art trackers such as SURF, BRISK, Lucas-Kanade and proposed adaptive Lucas-Kanade approach. The evaluation standard parameter prove that proposed approach is more efficient then state of the art methods.

ACKNOWLEDGEMENTS

REFERENCES

[1] L. Patino and J. Ferryman, Loitering Behaviour Detection of Boats at Sea, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, United Kingdom, 2017, 44-50.

[2] G. Liu, S. Liu, F. Wang, and J. Ma, Infrared Aerial Small Target Detection with NSCT and Two-Dimensional Property Histogram, *International Journal of Pattern Recognition and Artificial Intelligence*, *32(10)*, 2018, pp. 1-17. "doi.org/10.1142/S0218001418500313".

[3] D. Prasad, D.Rajan, L.Rachmawati, E.Rajaballyand C.Quek,Video processing from electro-optical sensors for object detection and tracking in a maritime environment: a survey, *IEEE Transactions on Intelligent Transportation Systems*, *18(8)*, 2017, 1993-2016.

[4] B.Kalantar, S. B.Mansor, A. A.Halin, H.Zulhaidi, M. Shafri and Mohsen Zand, Multiple Moving Object Detection from UAV Videos Using Trajectories of Matched Regional Adjacency Graphs, *IEEE Transactions on Geoscience and Remote Sensing*, *55(9)*, 2017, 5198-5213.

[5] H. Shen, S. Li,C. Zhu, H. Chang, J. Zhang, Moving object detection in aerial video based on spatiotemporal saliency, *Chinese Journal of Aeronautics*, *26(5)*, 2013, no. 5, pp. 1211-1217.

[6] L.M.G. Fonseca, L. M.Namikawa, and E. F.Castejon, Digital image processing in remote sensing, *IEEE Tutorials of the XXII Brazilian Symposium on Computer Graphics and Image Processing*, Rio de Janeiro, Brazil, 2009, 59-71. "doi.org/10.1109/SIBGRAPI-Tutorials.2009.13".

[7] D. Israni and H. Mewada, Identity Retention of Multiple Objects under Extreme Occlusion Scenarios using Feature Descriptors, *Journal of Communications Software and Systems (JCOMSS), 14(4)*, 2018, 290-301.

[8] R.S. Shirbhate, N. D. Mishra, R. P. Pande, Video Surveillance System Using Motion Detection - A Survey, *International Journal of Advanced Networking and Applications*, *03(05)*, 2012, 19-22.

[9] S. Kim and J. Lee, Small infrared target detection by region- adaptive clutter rejection for sea-based infrared search and track, *Sensors*, *14(7)*, 2014, 13210-13242, "doi.org/10.3390/s140713210"

[10] D.Israni and H. Mewada, Feature Descriptor Based Identity Retention and Tracking of Players under Intense Occlusion in Soccer Videos, *International Journal of Intelligent Engineering and Systems*, *11(4)*,2018, 31-41. "doi.org/10.22266/ijies2018.0831.04".

[11] D.Bloisi, L.Iocchi, M.Fiorini, G.Graziano, Camera Based Target Recognition for Maritime Awareness, In *15th IEEE International Conference on Information Fusion*, 2012, 1982-1987.

[12] D. Bloisi, L.Iocchi, M.Fiorini, and G.Graziano, Automatic maritime surveillance with visual target detection, In *Proceedings of the International Defense and Homeland Security Simulation Workshop (DHSS)*, Rome, 2011, 141-145.

[13] P. Panchal Payal, G.Prajapati, S. Patel, H. Shah, and J. Nasriwala, A review on object detection and tracking methods, *International Journal for Research in Emerging Science and Technology*, *2(1)*, 2015, 7- 12.

[14] S. Leutenegger,M. Chli andR. Siegwart, BRISK: Binary robust invariant scalable keypoints. In *IEEEInternational Conference on Computer Vision (ICCV)*, 2011, 2548-2555.

[15] H.S. Parekh, D. G. Thakore, and U. K. Jaliya, "A survey on object detection and tracking method", *International Journal of Innovative Research in Computer and Communication Engineering*, *2(2)*, 2014, 2970- 2978.

[16] K. Patel,D. Israni,D. Garg,A Roadmap to Mitigation Techniques: Bedrock for Atmospheric Turbulence, *International Journal of Advanced Networking and Applications*, *10(5)*, 2019, 4007-4012.

[17] R.Ramachandran, 'A Survey on Image Matching Techniques', *International Journal for Research in Applied Science and Engineering Technology*, *2(5)*, 2014, 275-280.

[18] D. Patel, and S. Upadhyay, Optical flow measurement using lucaskanade method,*International Journal of Computer Applications*, *61(10)*, 2013, 6-10.

[19] M.K. Bhure, and M.J.D. Dhande, Object Detection using SURF features, *International Research Journal of Engineering and Technology (IRJET)*, *4(7)*, 2017, no. 7, 1524-1526.

[20] B. BhosaleSwapnali, S. Kayastha Vijay and K. HarpaleVarsha, Feature extraction using surf algorithm for object recognition", *International Journal of Technical Research and Applications*,*2(4)*, 2014, 197-199.

[21] H. Bay, A.Ess, T. Tuytelaars, L. Van Gool, Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding*, *110(3)*,2008, 346-359, "doi.org/10.1016/j.cviu.2007.09.014".

[22] D. Meyer, J. Denzler, and H. Niemann, Model based extraction of articulated objects in image sequences for gait analysis, in *Proceedings of International Conference on Image Processing*, 1998, 78–81.

[23] D. Meyer, J. Posl, and H. Niemann, Gait classification with HMM's for trajectories of body partsextracted by

mixture densities, in *Proceedings of British Machine Vision Conference*, 1998, 459–468.

[24] J. Barron, D. Fleet, and S. Beauchemin, Performance of optical flow techniques, *International Journal of Computer Vision*, *12(1)*, 1994, 42–77.

[25] O.Patel, Y. P.Maravi, and S. Sharma, A comparative study of histogram equalization based image enhancement techniques for brightness preservation and contrast enhancement,*Signal & Image Processing: An International Journal (SIPIJ)*,*4(5)*, 2013, 11-25.

[26] Lucas, Bruce D., and Takeo Kanade, 'An iterative image registration technique with an application to stereo vision, in *Proceedings of the 7th International Joint Conference on Artificial Intelligence,3*, 1981, 674-679.

[27] B. D. Lucas, Generalized image matching by the method of differences, doctoral diss., Pittsburgh, PA, 1984.

[28] L. Patino, J. Ferryman, Semantic modelling for behaviourcharacterisation and threat detection, in*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops,* 2016: 43-49.

[29] C. Goutte, and E. Gaussier, A probabilistic interpretation of precision, recall and F-score, with implication for evaluation, in *European Conference on Information Retrieval*, Berlin, Heidelberg, 2005, 345-359.

[30] Y. Sasaki, The truth of the F-measure *Teach Tutor mater, 1(5)*, 2007, pp. 1-5.

AUTHOR'S PROFILE

Disha Patel received M. Tech in Computer Engineering from CHARUSAT University. She has received her B.E. in Computer Engineering from IIET, Dharmaj. Her major research interest includes Image processing, Computer Vision, Machine learning.

Dippal Israni received his M. Tech in Computer Engineering from Dharmsinh Desai University He is Assistant Professor in CSPIT, CHARUSAT and Assistant Scientist in Charusat Space Research and Technology Centre, CHARUSAT. His major research interests include Computer Vision, Image Processing and Machine Learning.

Mrugendrasinh Rahevarreceived his M. Tech in Computer Science and Engineering from GEC, Modasa. He is Assistant Professor in, CSPIT, CHARUSAT. His major research interests include Compute**r** Vision, and Deep learning