

A Comprehensive Survey of Algorithms for Face Tracking in different Background Video Sequence

Ranganatha S
Asst. Professor

Dept. of Computer Science and Engineering
Government Engineering College
Hassan, Karnataka, India

Y. P. Gowramma
Professor

Dept. of Computer Science and Engineering
Kalpataru Institute of Technology
Tiptur, Karnataka, India

ABSTRACT

Video processing is an interesting research zone in image processing. Face tracking is part of video processing, where the face regions need to be detected and tracked. In this paper, we present a survey of some of the familiar algorithms that are used for tracking the face(s) in different background challenging video sequences. Mean-Shift is an important algorithm that is based on the displacement of points. Improvisation of Mean-Shift lead to the development of CAMSHIFT; the latter is one of the robust chromatic tracking approach developed till date. KLT is an efficient point tracking algorithm. This paper also includes a survey of different motion estimation algorithms, which are classified as either pixel based or feature based. At the end, recent developments help in knowing the relevant works that are being carried out now a days.

Keywords

Face Tracking, Survey, Algorithms, Different Background, Video Sequence, Mean-Shift, CAMSHIFT, KLT, Motion Estimation, Recent Developments.

1. INTRODUCTION

Face tracking relies on monitoring the required facial region(s) based on the previous detection method, and the features extracted in the previous steps, that take place before tracking by having a bounding box or any other marker around the required facial region.

Face tracking algorithms face various technical challenges [1], some of them are as listed next. 1) Illumination: Face look changes when there are variations in the illumination. Reflectance properties of skin, the type of equipment and the way it is used for capturing the video, are the major sources of variations. 2) Pose: This creates a problem while detecting the face in the input image. Most of the existing algorithms are capable of tracking only the frontal posed faces. But, we expect a tracker to identify and trace the different poses of the face(s). 3) Expression: Robust face tracking becomes nightmare when there are variations in the human facial expressions. Most of the existing algorithms work robust except for the extreme expressions like scream. 4) Occlusion: Tracking becomes difficult when the face in the input image is not visible properly, may be due beard, moustache, glasses and headscarf. 5) Different backgrounds: Sometimes, the backgrounds of the videos are complex and cluttered. Especially, we say that a video sequence comes under different background, if it is recorded under the following conditions: a) Moving faces, but unmoving camera, b) Unmoving faces, but moving camera, and c) Moving faces, with moving camera.

Surveillance is one of the major area where the tracking of face(s) is intelligently used. Along with this, other areas such as indexing, compression and video structuring also need the assistance of tracking algorithms. Various algorithms are available for tracking; some of them are novel (base), and the remaining are the modified versions of the base algorithms to overcome the drawbacks/specific-drawback of those algorithms.

The upcoming sections of our paper are prepared as stated next. Section 2 houses the face tracking algorithms. Section 3 includes the recent developments. Lastly, the paper halts with conclusion in section 4.

2. FACE TRACKING ALGORITHMS

Let us discuss some of the major tracking algorithms in this section.

2.1 Mean-Shift Algorithm

Mean-Shift [2, 3] is a non-parametric, feature-space analysis technique that is used for locating the maxima of a density function, which is often referred to as mode-seeking algorithm. It is based on the set of data points at initial phase; and locating their initial mean in further iterations, based on the displacement of the points as shown if Figure 1.

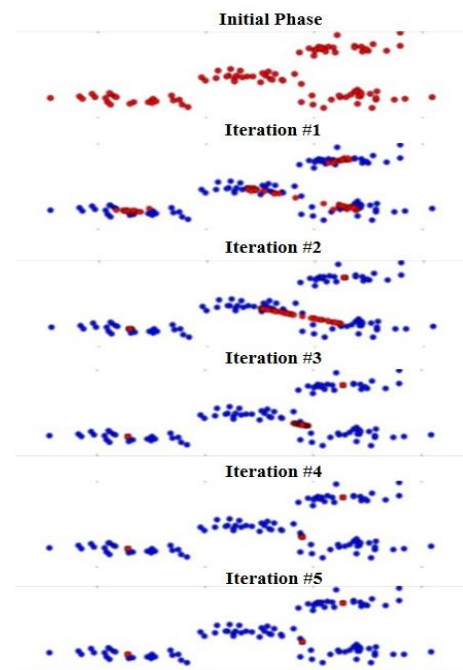


Fig 1: Illustration of the operation of Mean-Shift algorithm

The Mean-shift algorithm for a set of data points D :

1. For each data point $d \in D$, find the neighbouring points $E(d)$ of d .
2. For each data point $d \in D$, calculate the mean shift $s(d)$ using the equation:

$$s(d) = \frac{\sum_{d_i \in E(d)} P(d_i - d) d_i}{\sum_{d_i \in E(d)} P(d_i - d)} \quad (1)$$

where $s(d)$ is the neighbourhood of d , a set of points for which $P(d_i) \neq 0$.

3. For each data point $d \in D$, update $d \leftarrow s(d)$.
4. Repeat the first step for n iterations or until the points are almost not moving or not moving.

Mean-Shift can be applied for clustering, tracking, and smoothing purposes. The major drawback of this algorithm is that the window size is trivial.

2.2 Kanade-Lucas-Tomasi (KLT)

Algorithm

It is a method to feature mining. KLT uses spatial intensity statistics to hunt for the location that produces the correct match. It is quicker during the examination of less potential image matchings.

KLT feature tracer is centered on 2 papers. In the 1st paper, Lucas and Kanade [4] formulated an idea of local examination by using the approximation of gradient weights. In the 2nd paper, Tomasi and Kanade [5] adopted the similar approach for image record-keeping due to the transformation; but, improvised the method by adopting the features that are better suited for tracking algorithms. Let us look into the details of first paper. Consider m as the shift between 2 images $A(x)$ and $B(x) = A(x + m)$. Then the computation is made by,

$$A' \approx \frac{A(x+m) - A(x)}{m} = \frac{B(x) - A(x)}{m} \quad (2)$$

Which results in,

$$m \approx \frac{B(x) - A(x)}{A'(x)} \quad (3)$$

This computation to the gradient of an image is accurate only if the shift region between the 2 images is not excessively huge. So, computation of m depends on x . Hence, at many values of x , m will be as mentioned in the equation below.

$$m \approx \frac{\sum_x \frac{B(x) - A(x)}{A'(x)}}{\sum_x 1} \quad (4)$$

The average can be additionally enhanced by weighing the impact of every term that is inversely proportionate to an assessment of $|A''|$ as,

$$A''(x) \approx \frac{B'(x) - A'(x)}{m} \quad (5)$$

Hence, the weighting function would be, 1/|A''(x)| i.e.

$$W(x) = \frac{1}{|B'(x) - A'(x)|} \quad (6)$$

Then the average with weighting can be rewritten as,

$$m = \frac{\sum_x \frac{W(x) [B(x) - A(x)]}{A'(x)}}{\sum_x W(x)} \quad (7)$$

Using the Newton Raphson iteration, we can estimate the best value of m . That is,

$$m = \begin{cases} m_0 = 0 \\ m_{k+1} = m_k + \frac{\sum_x \frac{W(x) [B(x) - A(x)]}{A'(x)}}{\sum_x W(x)} \end{cases} \quad (8)$$

In the second paper, the recommended features are chosen, if the two eigenvalues of the gradient matrix are greater than certain threshold. It can be formulated as,

$$\nabla g = e \quad (9)$$

in which ∇ is the gradient. The formula is similar to m that is discussed above. Local path is a better feature to track, if the two eigenvalues (λ_1 and λ_2) of ∇ are greater than the threshold value. The tracking technique that is centred on these papers is known as KLT.

Further, Shi and Tomasi [6] have suggested a verifying stage, which verifies whether the features are tracked correctly or not. An affine conversion is apt among the picture of the presently traced feature and its picture from a not following earlier frame. If the affine benefited picture is excessively unlike, then the feature is discarded. The reason is that, among the successive frames, translation is an adequate prototype for tracking; however, due to the difficult movement, viewpoint issues etc. a more difficult prototype is necessary when the frames are away from each other. By means of an analogous basis as for the KLT, Shi and Tomasi exhibited that the quest can be accomplished with the formula,

$$Sx = b \quad (10)$$

where S is a gradients matrix, x is an affine coefficients vector and b is a fault vector.

2.3 Motion Estimation Algorithm

Motion estimation is an approach for deciding the motion vectors which express the renovation from nearby frames in a video sequence. It is an ill posed problem as motion is 3D scene and frames/images are 3D scenes projected on the 2D

plane. The motion vectors can act on whole frame or specific region or a pixel. The methods of finding motion vectors [7-9] can be classified as:

1. Pixel based/Direct method, and
2. Feature based/Indirect method.

2.3.1 Pixel based method

This category includes various algorithms such as:

1. Block matching algorithm,
2. Phase correlation and frequency domain, and
3. Optical flow.

2.3.1.1 Block matching algorithm

Block matching [10] is a method of finding the corresponding macro-blocks in the frames of a video. This algorithm divides the current frame of a video into macro-blocks; and compares every macro-block with a matching block in the neighbouring frames of the video. A vector is generated which models the motion of a macro-block from one position to another. This motion, computed for all the macro-blocks containing a frame, set up the motion expected in a frame. Figure 2 illustrates the discussion made till now.

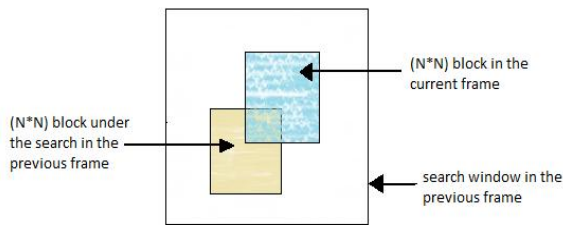


Fig 2: Illustration of block search procedure.

The hunt region for a correct macro-block match is determined by the ‘search factor’, s , where s is the count of pixels on all four flanks of the matching macro-block in the former frame. The search factor is a measure of movement. The higher the value of s , higher is the possible movement and the probability for discovering a noble match. A complete hunt of all the possible blocks is however a computationally costly job. Block matching includes various well known matching algorithms for finding the best matching block.

2.3.1.2 Phase correlation and frequency domain

Phase correlation [11] is a method to measure the offset among two analogous images (image correspondence) or other data groups. It is normally adopted in image registering, it depends on a frequency domain depiction of the data, and generally computed by Fast Fourier Transform. The word is applicable mainly to a group of cross-association methods that separate the phase statistics from the Fourier space illustration of the cross-correlogram.

Consider two input images i_p and i_q . The Fourier transform of these images can be calculated as follows:

$$I_p = T\{i_p\}, I_q = T\{i_q\}. \quad (11)$$

Further, compute the cross power spectrum by means of,

$$G = \frac{I_p \circ I_q^*}{|I_p \circ I_q^*|}. \quad (12)$$

Then to attain the normalized cross association, we put on inverse Fourier transform as follows:

$$g = T^{-1}\{G\}. \quad (13)$$

At the end, determine the highest value in g using,

$$(\Delta a, \Delta b) = \arg \max_{(a,b)} \{g\}. \quad (14)$$

2.3.1.3 Optical flow

Optical flow is an arrangement of possible motion of the entities, planes, and boundaries in a visual arena instigated by the associative kinetics among a viewer and the arena [12, 13]. The idea of optical flow was presented by a psychologist, J. J. Gibson, to define the visual stimulus given to the moving animals [14]. Gibson highlighted the significance of optic flow for affordance observation, the capability to determine the chances of the motion within the scenario. Supporters of Gibson as well as his ecological method to psychology have additionally explained the importance of optical flow stimulus for the realization of motion by the viewer in the realm, and the outlines (e.g. shapes) [15]. Figure 3 illustrates the rotation of observer, 3D optic flow and its 2D representation.

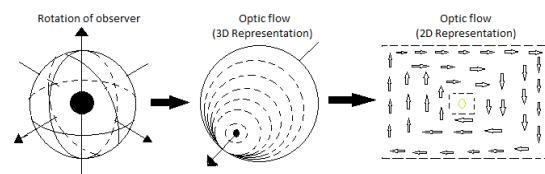


Fig 3: Illustration of the optical flow algorithm.

2.3.2 Feature Based method

Feature based approach extract the features, e.g. ‘‘corner points’’ in the video frames. Statistical procedure is enforced over a global or native region, which assists the function of eliminating the matches that does not resemble the actual motion. An effectively used statistical procedure is RANSAC (Random Sample Consensus). Let us look into the approaches that are based on features in the following sub-subsections.

2.3.2.1 Corner Detection

This approach is used to extract specific type of features and deduce the subjects of a frame, or just an image in the video sequence. A corner can be observed as the connection of two edges. There are many algorithms that are based on corner detection, namely, Moravec, Förstner, Harris & Stephens, multi-scale Harris, Gaussian, level curve curvature, scale-space interest points, Wang and Brady, affine-adapted interest points, SUSAN, AST-based, Trajkovic & Hedley, spatio-temporal interest points, automatic synthesis, and many more. Let us look into Moravec and Harris & Stephens corner detection algorithms in the succeeding part of this section.

Moravec algorithm [16] describes a corner as a location having small self-resemblance. It exams every element of the image to check whether any corner is existing, taking into account how analogous a spot centred on the element is to the close, hugely coinciding spots. The correspondence is gauged by doing the addition of squared differences (SSD) among the conforming elements of two spots. A small numeral designates higher resemblance. When the element is in an area of constant intensity, then the nearest spots will appear analogous. When the element is on an edge, then the nearest spots in a path vertical to the edge will appear somewhat

dissimilar; but, the nearest spots in a path parallel to the edge will end in a minor modification. When the element is on a feature with deviation in all the ways, then no nearest spot will appear similar.

Harris & Stephens [17] enhanced the Moravec approach by taking the differential of the corner total with regard to direction as an alternative for the shifted locations. Let us consider an image (frame) as F , the frame patch area as (c, d) that is shifted by (m, n) and the weighted sum of squared differences (SSD) among the two patches as D . Then,

$$D(m, n) = \sum_c \sum_d w(c, d) (F(c+m, d+n) - F(c, d))^2. \quad (15)$$

$F(c+m, d+n)$ can be extended using Taylor extension, nearly as:

$$F(c+m, d+n) \approx F(c, d) + F_m(c, d)m + F_n(c, d)n. \quad (16)$$

Hence,

$$D(m, n) \approx \sum_c \sum_d w(c, d) (F_m(c, d)m + F_n(c, d)n)^2. \quad (17)$$

It can be written in matrix form as,

$$D(m, n) \approx (mn)S \begin{pmatrix} m \\ n \end{pmatrix}. \quad (18)$$

Where S is a structure element that can be written as follows:

$$S = \begin{bmatrix} \langle F_m^2 \rangle & \langle F_m F_n \rangle \\ \langle F_m F_n \rangle & \langle F_n^2 \rangle \end{bmatrix}. \quad (19)$$

A corner is categorized by a huge deviation of D in each direction of the vector (mn) . After examining the eigenvalues of S , this categorization can be conveyed as stated next: S must have 2 “huge” eigenvalues for the concerned point. Then using the sizes of the eigenvalues, we can make the following inferences.

- If $\delta_1 \approx 0$ and $\delta_2 \approx 0$ then pixel (m, n) has no concerned feature point.
- If $\delta_1 \approx 0$ and δ_2 has certain huge positive value, then we get an edge.
- If δ_1 and δ_2 have huge values, then we get a corner.

Harris and Stephens found that precise calculation of the eigenvalues is costly, as it involves the calculation of a square root, then as an alternative recommended the below mentioned equation R_c , in which t is an adjustable responsive factor.

$$R_c = \delta_1 \delta_2 - t (\delta_1 + \delta_2)^2 = \det(S) - t \text{Trace}^2(S). \quad (20)$$

2.3.2.2 RANSAC

It is an iterative process [18] of computing the parameter in the group of the monitored data which comprises outliers. It is not a deterministic procedure i.e. it yields a realistic outcome merely with a definite likelihood. RANSAC can be adopted to answer the Location Determination Problem (LDP), where the

aim is to identify the elements of the region. An example for this is, fixing a contour in 2 dimensions to a group of observations. Figures 4 and 5 illustrates the process of RANSAC.

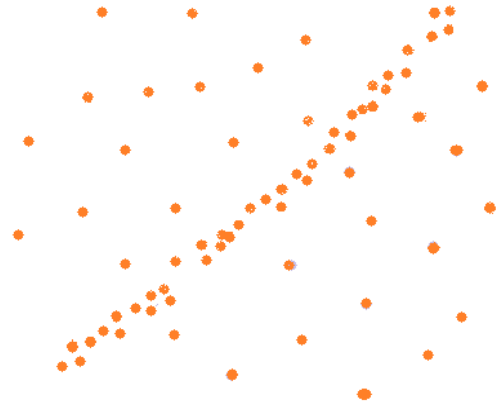


Fig 4: A data group with various outliers for which a line (contour) needs to be fixed.

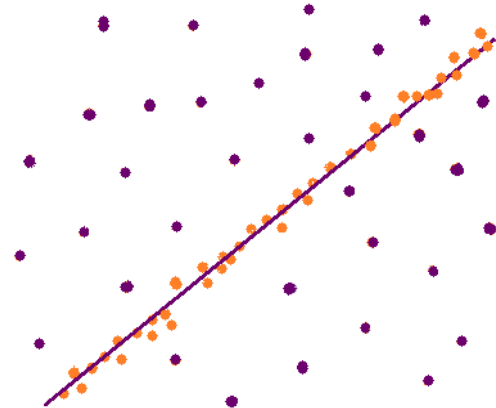


Fig 5: Fixed line (contour) with RANSAC in which the outliers have no impact on the result.

3. RECENT DEVELOPMENTS

In this section, we look into the recent developments that are happened, either using or improving the algorithms being discussed in the previous section of this paper.

CAMSHIFT [19] is based on the Mean-Shift algorithm, which applies Mean-Shift to every single frame of the video sequence. An improved CAMSHIFT algorithm proposed by Emami et al. [20] is based indirectly on the Mean-Shift algorithm. Ranganatha et al. [21] have developed an efficient tracking algorithm by joining the Mean-Shift algorithm and Kalman filter [22]. They have addressed the problem of illumination; and improved the speed of tracking.

KLT algorithm is a core for the feature tracking algorithms as many of the feature trackers are based on the point values. Ranganatha et al. [23-27] have proposed several algorithms that are based on KLT. These algorithms are robust enough to track both single and multiple face in different background video sequence. The algorithms addressed many questions which were left unanswered previously such as, 1) Tracking

multiple face [24, 25], 2) Selective face tracking [26] i.e. user specified number of face(s) tracking, 3) Selected single face tracking [27], 4) Occlusion (e.g. spectacle, beard, headscarf etc.) [25, 26], 5) Low resolution [23-27], 6) Face skin colored regions [26], 7) Abrupt motions [23-27], 8) Variations in pose, expression and illumination [23-27] etc. to name a few. Some of the primary features used by these algorithms are centroid [23], mean [24, 25] and color [26]. These algorithms have also adopted the secondary features such as SURF [21], Haar [21, 23], BRISK [24], FAST [25, 27], and HOG [27] etc.

The work in literature [23] efficiently used the corner detection method to compute the corner and edge points. Further, centroid was computed using the corner and edge points. This work fails in multiple face tracking. But, the effective fusion of KLT with corner detection method nullified the shortcomings of both KLT and corner detection methods.

The algorithms being discussed in the previous paragraphs, adopted either existing approaches or self-established techniques for detecting the face(s) in the videos during the first phase. Viola-Jones [28, 29] algorithm is an existing popular face detection approach, which detects only the frontal posed face(s); and for detecting the face(s) having different poses, it needs further training. Instead of this, images can be trained [25, 30], which help during the first phase to detect the face(s).

Figures below include few of the snap shots, which are outcomes of some of the major algorithms discussed in this section. In each Figure, the frame numbers that are displayed indicate the frames from left to right respectively. Figures show the frames of both indoor and outdoor videos, which belong to different datasets.

The video sequences being considered for testing as shown in Figures 6, 7, 11 and 13 are taken from Celebrities dataset [31]. The video sequence that is selected for testing as shown in Figure 8 is from Choke Point dataset [32]. The video sequence that is considered for testing as shown in Figure 9 belong to HOHA dataset [33]. The video sequence being considered for testing as shown in Figure 10 is part of Math-Works videos [34]. Lastly, the video sequence that took part in testing as shown in Figure 12 was obtained from VidTIMIT dataset [35].

Video sequences which are considered as shown in the frames of Figures 6, 7, 8, 10 and 12 comes under *Moving faces, but unmoving camera* category. The video sequence that is considered as shown in the frames of Figure 13 comes under *Unmoving faces, but moving camera* category. Lastly, the video sequences that are considered as shown in the frames of Figures 9 and 11 comes under *Moving faces, with moving camera* category.



Fig 6: Outcome of the work in literature [21]; showing the tracking result of single face, 641 frames video sequence 0741_01_001_harry_connick.avi with frames #36 and #357 displayed.



Fig 7: Outcome of the work in literature [23]; showing the result (face detection, feature selection and face tracking) of single face, 186 frames video sequence 0286_01_016_angelina_jolie.avi with frames #01, #01 and #45 displayed.



Fig 8: Outcome of the work in literature [24]; showing the tracking result of multiple face, 26 frames video sequence P2LS5C1.mp4 with frames #47, #55 and #67 displayed.



Fig 9: Outcome of the work in literature [24]; showing the tracking result of multiple face, 233 frames video sequence Dead_Poets_Society_00068.avi with frames #07, #40 and #126 displayed.



Fig 10: Outcome of the work in literature [25]; showing the tracking result of single face, 413 frames video sequence tilted_face.avi with frames #72 and #198 displayed.



Fig 11: Outcome of the work in literature [26]; showing the tracking result of single face, 23 frames video sequence 0193_01_004_alanis_morissette.avi with frames #01, #10, #19 and #22 displayed.



Fig 12: Outcome of the work in literature [27]; showing the result (face selection and tracking) of single face, 363 frames video sequence mcm0_head.mpg with frames #01, #38, #196 and #123 displayed.



Fig 13: Outcome of the work in literature [30]; showing the tracking result of single face, 69 frames video sequence 0673_01_004_gloria_estefan.avi with frames #04, #37, #49 and #58 displayed.

4. CONCLUSION

Video and video types are changing day by day, due to which, face tracking is an active research area in video processing. Face tracking is a thought-provoking problem because of a number of real-life challenging conditions, such as the shape, changes in the expression and pose, illumination, abrupt motions, occlusion, different backgrounds etc. to mention a few. In this paper, we have discussed some of the familiar algorithms which are used for human face tracking in different background challenging videos. Apart from the literature being included and discussed, new algorithms are also getting developed; because, face tracking has several issues which are still open for research.

5. REFERENCES

- [1] Ranganatha S and Dr. Y P Gowramma, "Face Recognition Techniques: A Survey", International Journal for Research in Applied Science and Engineering Technology (IJRASET), ISSN: 2321-9653, Vol.3, No.4, pp.630-635, April 2015.
- [2] K. Fukunaga and L. D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", in IEEE Trans. on Information Theory, Vol.21, No.1, pp.32-40, January 1975. DOI: 10.1109/TIT.1975.1055330
- [3] Yizong Cheng, "Mean Shift, Mode Seeking, and Clustering", in IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol.17, No.8, pp.790-799, August 1995. DOI: 10.1109/34.400568
- [4] Bruce D. Lucas and Takeo Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision", in Proc. of International Joint Conference on Artificial Intelligence, Vol.2, pp.674-679, August 1981.
- [5] Carlo Tomasi and Takeo Kanade, "Detection and Tracking of Point Features", Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
- [6] Jianbo Shi and Carlo Tomasi, "Good Features to Track", in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.593-600, June 1994. DOI: 10.1109/CVPR.1994.323794
- [7] Philip H.S. Torr and Andrew Zisserman, "Feature Based Methods for Structure and Motion Estimation", ICCV Workshop on Vision Algorithms, pp.278-294, September 1999.
- [8] Michal Irani and P. Anandan, "About Direct Methods", ICCV Workshop on Vision Algorithms, pp.267-277, September 1999.
- [9] Rui Xu, David Taubman, and Aous Thabit Naman, "Motion Estimation Based on Mutual Information and Adaptive Multi-scale Thresholding", in IEEE Trans. on Image Processing, Vol.25, No.3, pp.1095-1108, March 2016. DOI: 10.1109/TIP.2016.2514488
- [10] C. Je and H-M. Park, "Optimized Hierarchical Block Matching for Fast and Accurate Image Registration", Signal Processing: Image Communication, Vol.28, No.7, pp.779-791, August 2013.
- [11] M. H. G. Peeters, "Implementation of the Phase Correlation Algorithm: Motion Estimation in the Frequency Domain", Practical Training Report, June 2003.
- [12] Andrew Burton and John Radford, "Thinking in Perspective: Critical Essays in the Study of Thought Processes", Routledge, ISBN: 0-416-85840-6, 1978.
- [13] David H. Warren and Edward R. Strelow, "Electronic Spatial Sensing for the Blind: Contributions from Perception", Springer, ISBN: 90-247-2689-1, 1985.
- [14] James J. Gibson, "The Perception of the Visual World", Houghton Mifflin, 1950.
- [15] C.S. Royden and K.D. Moore, "Use of Speed Cues in the Detection of Moving Objects by Moving Observers", Vision Research, pp.17-24, 2012. DOI: 10.1016/j.visres.2012.02.006
- [16] Hans P. Moravec, "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover", Robotics Institute, Carnegie Mellon University Technical Report CMU-RI-TR3, September 1980.
- [17] C. Harris and M. Stephens, "A Combined Corner and Edge Detector", in Proc. of 4th Alvey Vision Conference, Manchester, UK, pp.147-151, 1988.
- [18] Martin A. Fischler and Robert C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", Commun. ACM, Vol.24, No.6, pp.381-395, June 1981. DOI: 10.1145/358669.358692
- [19] G. Bradski, "Computer Vision Face Tracking for Use in a Perceptual User Interface", Intel Technology Journal, pp.12-21, 1998.
- [20] E. Emami and M. Fathy, "Object Tracking Using Improved CAMShift Algorithm Combined with Motion Segmentation", in Proc. of 7th Iranian Machine Vision and Image Processing (MVIP), pp.1-4, 2011.
- [21] Ranganatha S and Y P Gowramma, "An Integrated Robust Approach for Fast Face Tracking in Noisy Real-World Videos with Visual Constraints", in Proc. of IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.772-776, September 2017. DOI: 10.1109/ICACCI.2017.8125935
- [22] Aziz Makandar and Daneshwari Mulimani, "Analysis of Multiple Object Detection Using Kalman Filter in Sports Video", IJCA Proc. of National Conference on Computer Science and Information Technology (NCCSIT), 2017(1):13-15, September 2018.
- [23] Ranganatha S and Y P Gowramma, "A Novel Fused Algorithm for Human Face Tracking in Video Sequences", in Proc. of IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), pp.1-6, October 2016. DOI: 10.1109/CSITSS.2016.7779430

- [24] Ranganatha S and Y P Gowramma, “Development of Robust Multiple Face Tracking Algorithm and Novel Performance Evaluation Metrics for Different Background Video Sequences”, *International Journal of Intelligent Systems and Applications (IJISA)*, Vol.10, No.8, pp.19-35, August 2018. DOI: 10.5815/ijisa.2018.08.03
- [25] Ranganatha S and Y P Gowramma, “Image Training, Corner and FAST Features based Algorithm for Face Tracking in Low Resolution Different Background Challenging Video Sequences”, *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, Vol.10, No.8, pp.39-53, August 2018. DOI: 10.5815/ijigsp.2018.08.05
- [26] Ranganatha S and Y P Gowramma, “Color Based New Algorithm for Detection and Single/Multiple Person Face Tracking in Different Background Video Sequence”, *International Journal of Information Technology and Computer Science (IJITCS)*, Vol.10, No.11, pp.39-48, November 2018. DOI: 10.5815/ijitcs.2018.11.04
- [27] Ranganatha S and Y P Gowramma, “Selected Single Face Tracking in Technically Challenging Different Background Video Sequences Using Combined Features”, *ICTACT Journal on Image and Video Processing (JIVP)*, Vol.9, No.2, November 2018.
- [28] P. Viola and M. Jones, “Robust Real-Time Face Detection”, *International Journal of Computer Vision (IJCV)*, Vol.57, pp.137-154, 2004.
- [29] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features”, in *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Kauai, USA, Vol.1, pp.511-518, December 2001. DOI: 10.1109/CVPR.2001.990517
- [30] Ranganatha S and Y P Gowramma, “Image Training and LBPH Based Algorithm for Face Tracking in Different Background Video Sequence”, *International Journal of Computer Sciences and Engineering (IJCSE)*, Vol.6, No.9, pp.349-354, September 2018. CrossRef-DOI: <https://doi.org/10.26438/ijcse/v6i9.349354>
- [31] M. Kim, S. Kumar, V. Pavlovic, and H. Rowley, “Face Tracking and Recognition with Visual Constraints in Real-World Videos”, in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-8, June 2008. DOI: 10.1109/CVPR.2008.4587572
- [32] Yongkang Wong, Shaokang Chen, Sandra Mau, Conrad Sanderson, and Brian C. Lovell, “Patch-Based Probabilistic Image Quality Assessment for Face Selection and Improved Video-Based Face Recognition”, in *proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp.74-81, June 2011. DOI: 10.1109/CVPRW.2011.5981881
- [33] Ivan Laptev, Marcin Marszalek, Cordelia Schmid, and Benjamin Rozenfeld, “Learning Realistic Human Actions from Movies”, in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1-8, June 2008. DOI: 10.1109/CVPR.2008.4587756
- [34] <https://in.mathworks.com/downloads/R2018a/toolbox/vision/visiondata>.
- [35] C. Sanderson and B.C. Lovell, “Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference”, *Lecture Notes in Computer Science (LNCS)*, Vol.5558, pp.199-208, 2009.