An Overview on Object Tracking in Motion Pictures

Amit[1] Kirti Bhatia[2] Rohini Sharma[3] Shalini Bhadola[4]
[1]Student [2,3,4]Assistant Professor
[1,2,4]Sat Kabir Institute of Technology and Management, Haryana, India
[3]Government College for Women, Rohtak, India

Abstract— To extract some useful information from an Image, Image processing method is used prominently. The image is converted into a digital form through applying few operations on it. Detection and tracking of object is a challenging task and it is useful in video observation and vehicle direction-finding. Observing a video is mainly used in dynamic situations like tracking of sportsperson in sports, public security, and organization of traffic. This article provides an overview on the numerous methods, challenges and characteristics of recognition and tracking of objects in a motion environment.

Keywords: Object Tracking, Image Processing

I. Introduction

Video tracking is a procedure of tracing a relocating object/ objects across regular time interval by the use of a camera. It is used in multiple applications like for interaction among human and computer, protection and observation, video transmission and condensing, improved reality, traffic management, health imaging [1] and video expurgation [2][3]. Object tracking in a video is a time spending procedure since the quantity of data that is available in the video is very vast. Moreover, it is difficult to select right recognition techniques for object tracking.

In video tracking, the target objects are associated in successive video frames. This linking may be extremely complex if the objects are changing their positions very fast as comparative to the frame rate. One more condition which enlarges the difficulty of the challenge is when the tracked object alters its orientation with time. Owing to these conditions, video tracking schemes commonly implement a motion prototype and this model defines the change in the image of the target for dissimilar probable movements of the object.

To accomplish the video tracking a procedure examines consecutive video frames and concludes the motion of targets among the frames. A lot of methods exist in the literature, each one with some power and some flaws. The selection of the procedure depends upon the condition of the environment. Mainly there are two factors of a visual tracking scheme: target illustration, localization, and filtering and data relationship.

Illustration and localization of target is a bottom-up method and provides a large number of tools for recognizing the movement of the object. Successful tracking of an object is subject to the selection of the algorithm. E.g., since the profile of a person is changed dynamically, blob tracking is helpful in recognizing human motion [4]. Normally the mathematical intricacy of these procedures is very low.

Object tracking in videos is also known as standard computer vision challenge. Not only has it discovered an object in a sight but also recognize it in all subsequent frames. It makes the object discriminate from other objects in static as well as dynamic manner. There are various tracking processes available in the literature, and few of them are now considered as excellent benchmarks fundamental architect for the employment of others.

The Lucas-Kanade-Tomasi structure [5] has proposed tracking of main feature points through decreasing a cost function which is based on the consecutive concentration maps of all frames as well as probable affine conversion of the feature set. Some frameworks have optical flow according to examination of gradient maps. Video tracking is universal and may exist in applications like video games, technical microscopic examinations, observation, criminal studies, record tracking, and driving support methods.

Although various efficient frameworks are available for tracking the objects, however, real sights and videos may cause severe complications while implementing these frameworks. Generally, the tracking methods should be regulated and twisted to attain lighting situation, features trait of the object traced and numerous sound and irregularities. Approximately for all circumstances of object tracking through videos a general framework elucidation should be formed.

II. Related work

The area of object tracking in video surveillance has attracted many researchers. The object tracking has many significant applications and it is a developing research field of image processing. Owing to main characteristics of video surveillance, it has a variation of applications such as human-computer communications, protection and observation, video transmission, traffic management, common areas like airstrips, subways, highways, etc. Tracking an object in a messy ground is also one of the difficult problems of video observation. A successive course of moving object recognition, its grouping, tracing and detecting the conduct concludes the computational framework of video observation.

The authors in [6] have done a thorough study of tracking procedures, their classification into various categories, emphases on imperative as well as beneficial tracking approaches. They have provided a concise synopsis of tracing approaches such as zone based, dynamic contour based, along with their advantages and disadvantages. They have mentioned diverse tracing procedures with thorough explanation. They have examined universal approaches in literature study on dissimilar procedures and lastly affirming the examination of probable research guidelines.

Object tracking is a popular challenge in the field of image processing. Although, the capability of tracking objects has upgraded extremely since last few years; but, nonetheless, it is contemplated as a difficult problem to resolve. The significance of object tracing is shown in the various applications like video observation, human-computer communication, and direction-finding in robots. The main goal of the authors [7] is to study, assess, and to build a brief overview of the maximum object tracking approaches. The authors have studied representation of object, traits of the object, methods for object detection, and methods for tracking the object over subsequent frames.

There are three fundamental phase in video investigation; object recognition, object tracing, and identification of object actions through track analysis [8]. Considering the tracing procedure there are numerous unique procedures which can be distributed into separate groups, or classes. A general method is to divide the procedures into three key classes; point tracing, kernel tracing, and silhouette tracing. Another ways of classifying the algorithm are region-based tracing, contour based tracing, and boundary-based tracing.

III. Object Tracking Procedures

There are three main methods of object tracking can be further classified into specific methods as shown in figure 1.

A. *Kernel-based Tracking [9]:*

It is a repetitive localization method based on the expansion of a similar measure. It is also known as mean shift algorithm. It is a non-parametric examination method for locating the maxima of density function. Let begin with an initial estimate x. Let a kernel function k is provided. This k governs the weight of adjacent points for re-assessment of the mean. The weighted mean of the density in the window defined by k is

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x) x_i}{\sum_{x_i \in N(x)} K(x_i - x)} \quad (1)$$

Where, N(x) is the neighborhood of x. Types of kernel based tracking is discussed as follows:

1) *Mean Shift Tracking:*

For visual tracking, it make a confidence map in the new image according to the color histogram of the object in the preceding image, and implement mean shift to discover the highest confidence map nearby the object's previous location. For the new image, the map is a probability density function which assigns a probability to every pixel of the new image. It is the probability of the pixel color appearing in the object of the preceding image. Figure 2 represents the processing of mean shift tracking.

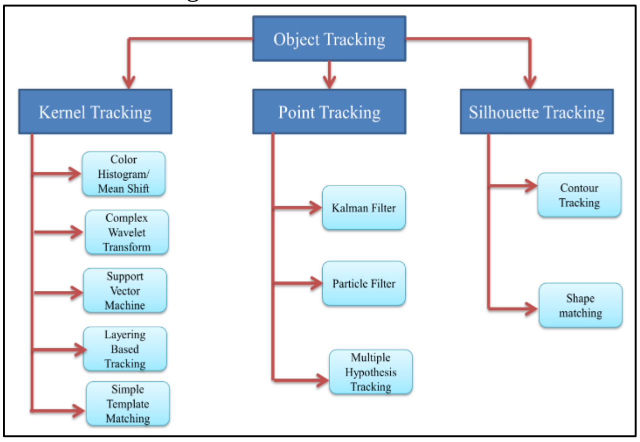

Fig. 1: Object Tracking Classification

2) *Complex Wavelet Transforms (CWT):*

It is a complex-valued expansion to the standard discrete wavelet transform. It is a 2-D wavelet transform that specifies multi-resolution, sparse illustration, and valuable representation of the configuration of an image. In the field of computer vision, by using the idea of visual frameworks, candidate regions can be rapidly focused. In this region objects of interest can be discovered, and then calculate extra characteristics by the CWT for those regions merely. These extra characteristics are valuable in precise recognition and identification of minor objects. Likewise, the CWT can be implemented to recognize the initiated voxels of cortex and in addition the temporal independent component analysis (TICA) can be employed to mine the fundamental autonomous resources [10].

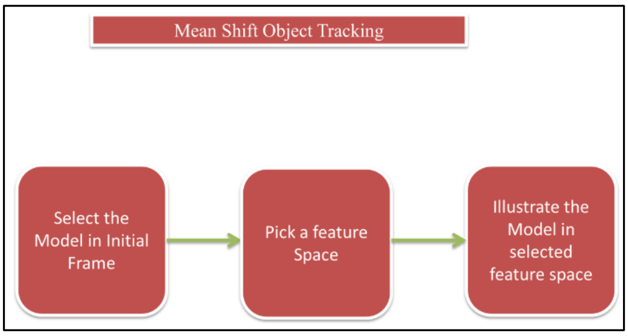

Fig. 2: Mean Shift Tracking

3) *Support vector Machine (SVM):*

It is a supervised learning paradigm which is used for classification and regression investigation. SVM are used to classify images. Simulation results illustrate that SVMs significantly provides great discovery precision than conventional query modification methods. Using a modified version SVM, image segmentation can also be performed [11]. Figure 3 shows the working of SVM for object tracking.

4) *Layered Based Tracking:*

Computerized object identification and tracking has been extensively employed in the video observation structures for public security and data synthesis in the remote detecting and airbase images. The most common applications are human motion investigation and the automobile discovery. The authors have implemented object identification and tracking in shape graphs of attracting objects integrating local background facts of the objects. The author has used layered approach as follows: At the top layer, figures and drawings gives a distinct measure to define the universal condition of the interesting objects. This type of knowledge is quite valuable to enhance the object tracking enactment for obstruction. The image can be formed as a graph or hyper graph by its native geometric characteristics. At the bottom layer, native geometric characteristics are utilized to catch native characteristics of objects and accomplish equivalent estimate of high-level figures. The native characteristic gives a direction to overcome imprecise object subdivision and abstraction [12].Figure 4 shows layered based tracking.

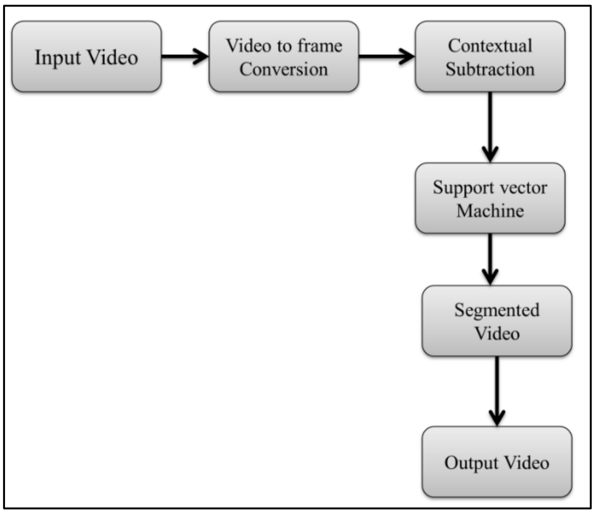

Fig. 3: Object Tracking using SVM

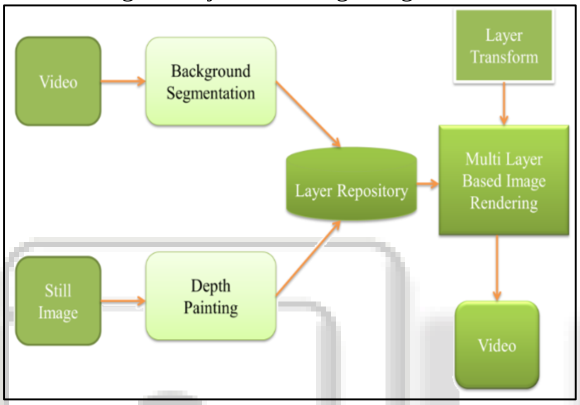

Fig. 4: Layered Based Tracking

5) *Template Based matching:*

Template matching is a method of digital image processing which is used to discover minor portions of a picture and match a template image. It may be utilized in developing a quality control, a method of detecting the direction of a mobile robot, or a method to find edges in a picture [13]. However, the key problems of template matching are: obstruction, identification of non-rigid conversions, lighting and contextual variations and disorder and scale variations. Figure 4 shows the processing of Template matching of an image.

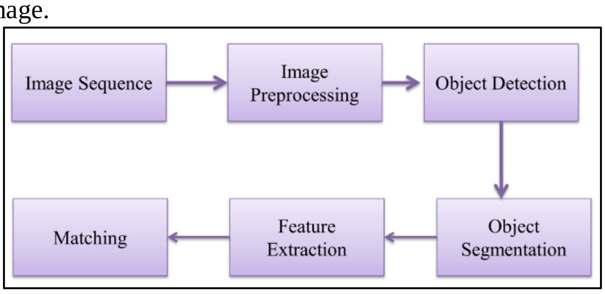

Fig. 5: Template Based Matching

B. *Point Tracking:*

We can track countless characteristics or patterns from an image. We can track one, two, or more features and it depends on the amount of data required and the amount of accurateness required in the result. There are four types of point tracking:

- One-point tracking - Track x and y axis of one feature position, with small or no variation in the image. We can use this data to transfer other components in the composite or can apply the reverse to steady the image.
- Two-point tracking - Track horizontal and vertical location for two facial appearances. The feature locations, related to each other, specify if the picture is revolving clockwise or counter-clockwise. In few events, two tracking points are adequate to compute the scaling of the facial appearances.
- Three-point tracking - Track horizontal and vertical location for three facial appearances. It gives complete profits of two-point tracking by an extra collection of tracking information for extra accurateness on z-rotation and scaling.
- Multi-point tracking - Entire benefits of less tracks by extra collection of tracking information. Generally, the three-point is adequate for maximum 2D tracking requirements; however multi-point prepares it to change and match-move additional component into the points of the facial appearances being tracked. Three variety of tracking are popular; kalman, Particle and multiple hypothesis tracking.

Various types of tracking are shown in figure 5[14].

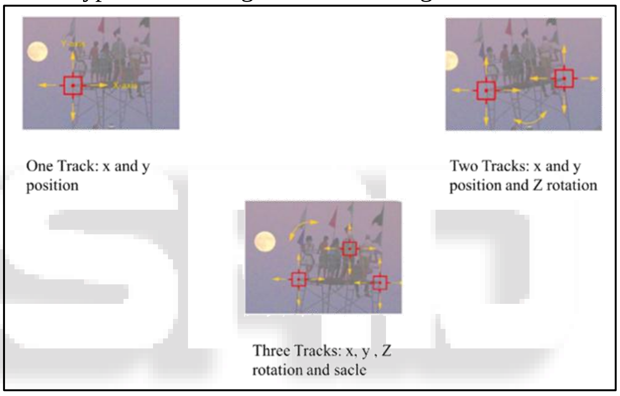

Fig. 6: Various types of point tracking

1) *Kalman Filter:*

It is an optimum iterative Bayesian filter for linear operations with respect to Gaussian noise. It is a method which makes use of a sequence of measurements perceived with time, comprising noise and other imprecisions, and creates approximations of anonymous variables which is likely to be more accurate than those related to a single measurement [15] Figure 6 shows the working of Kalman filter.

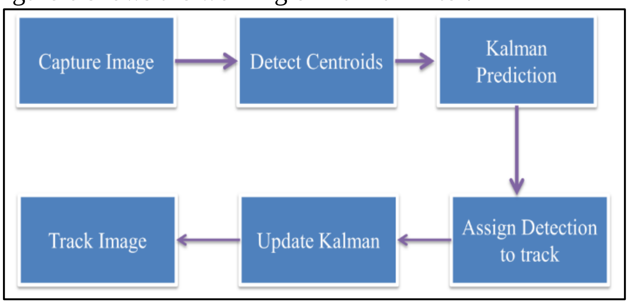

Fig. 7: Kalman Filter

2) *Particle Filter:*

It can track single and multiple objects simultaneously. It is a hypothesis tracker which estimates the filtered posterior allocation by a set of weighted particles. It weights particles on the basis of probability score and then transmits these particles according to a movement prototype. The probability function is used to diminish the number of particles surrounded near the object. Particle over the object has few RGB value while particle far from the object have zero RGB values. Particles with extra weight generate new particle in their vicinity and rest of the particles are moved from the object. A discrete collection of particles illustrates the object-state and progress with time through the survival of the fittest. It is beneficial for sampling the fundamental state-space allocation of nonlinear and non-Gaussian methods [16]. It is able to illustrate random densities. It uses temporal values for tracking. However, it is computationally difficult to implement [17]. Figure 7 shows movement of particles in Particle filter.

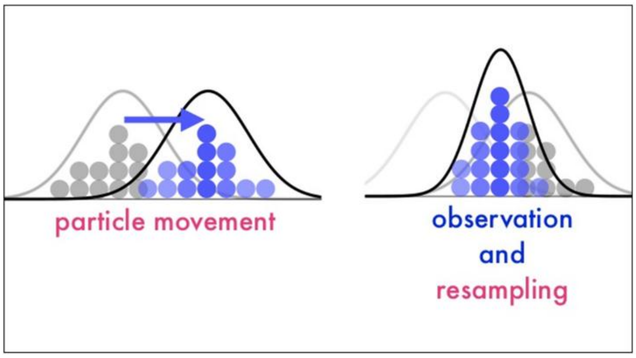

Fig. 8: Particle Filter

3) Multiple Hypothesis Tracking (MHT):

This tracking procedure decides the previous N scans. It makes tracks more consistent. The hypotheses are transmitted into the future in expectation that succeeding information may determine the ambiguity. The main standard of the MHT process is that problematic data alliance choices are postponed till further data are obtained. Processing of MHT is depicted in figure 8.

C. Silhouette Tracking:

The term Silhouette has been derived from the name of a French finance minister Étienne de Silhouette [18]. It is the low-cost method to record the appearance of a person. It is quite hard to identify the actions and process object tracking from silhouette images, particularly from a video. Researchers have suggested a new human action detection method as a grouping of numerous micro action series implemented by one or more body components of the people. This paradigm approximates the actions of dissimilar body portions for any specific time section to categorize movements. Some significant complexities are the existence of obstructions, disorder, communication between numerous objects, and varying illuminations. Atmosphere intricacy is another major problem is as it is affected by various environmental circumstances of the sight components; and procurement intricacy, which relies on video acquirement, that diverges regarding to view point, motion of the camera etc. Generally human movements are complex and hence a projected prototype has to manage all of these problems.

Academicians have selected the videos that comprise only human silhouettes, without reflecting silhouette extraction methods. To decrease the complication, the planned procedure measured videos that confined only one person object in every frame. The main modules of a standard recognition scheme through human silhouettes involve three major steps: foreground extraction/classification, and facet mining and movement categorization. The foreground is mined by removing backdrop setting of the video and consequently it assist in reduction of the searching zone of the present frame. The foreground arrangement decides if the foreground zone comprises a person or not. The examination of the human figure components motion is completed for the successive frames to decide the human movement in consecutive frames. Figure 9 depicts the silhouette and Figure 10 illustrates human motion detection from silhouette.

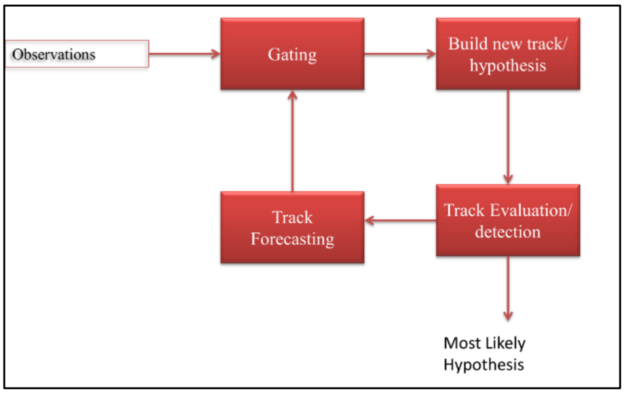

Fig. 9: Multiple hypothesis tracking

Contour tracking and shape matching are two popular methods of Silhouette tracking.

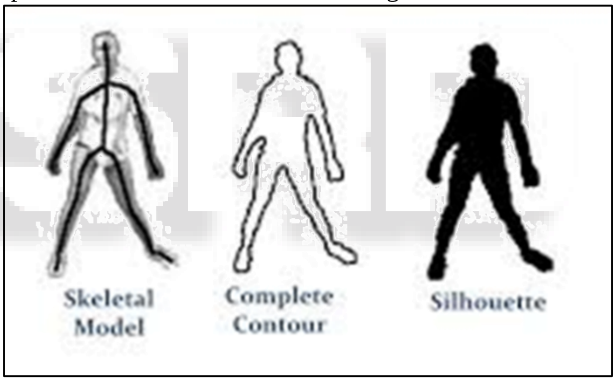

Fig. 10: Silhouette Image

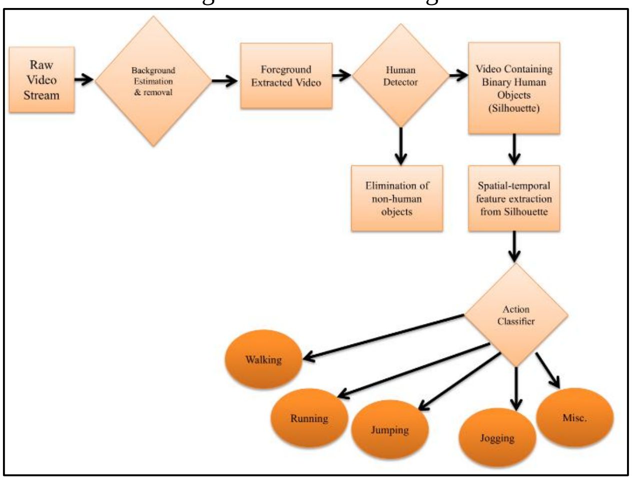

Fig. 11: Human Motion detection from Silhouette

1) Contour Tracking:

Recognition of object edge (dynamic contours or Compression procedure). Contour tracking procedures repetitively develop an original contour initialized from the preceding frame to its new location in the present frame. This method to contour tracking openly develops the contour by decreasing the contour power by gradient descent. Hand tracking through contour has been shown in [19] and by figure 11.

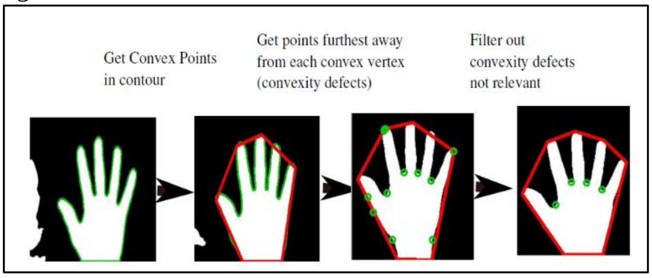

Fig. 12: Contour Tracking

2) Shape Matching:

This technique is used to track an object by using the following steps:

– Arbitrarily select a group of points that reside on the edges of an identified shape and one more set of points of an unidentified shape.
– Calculate the shape perspective of every point found in previous step.
– Relate all points from the identified shape to some point on the unidentified shape. To reduce the price of matching, initially select a transformation (affine or spline) that enclose the edges of the identified shape to the unidentified shape. Then pick the point on the unidentified shape that is very close to every warped point on the identified shape.
– Compute the distance among every pair of points on the two shapes. Use a weighted sum of the distance of shape framework.

IV. Conclusion

Tracking is the process of approximating the trajectory of an object in the picture plane as it travels over a sight. Filtering and data relationship is a top-down approach which includes integrating previous statistics regarding the sight or object, managing the object dimensions and assessment of dissimilar hypotheses. These approaches permit the tracking of composite objects along with additional composite object communication such as tracking objects movement in the presence of occlusions. In this article, we performed an extensive literature overview on object tracking methods and its various applications. All tracking procedures suppose that the movement of the object is easy and there is no occlusion. Object tracking procedures needs extra computation and storage constraints if there is an enhancement in the quantity of data confined in the video. While few tracking procedures can discover and track many objects simultaneously and can manage obstruction. However it requires extra computation and storage prerequisite. Various object tracking procedures can manage the changing lighting, contextual disorder, disguise, and obstruction distinctly.

References

[1] Peter Mountney, Danail Stoyanov & Guang-Zhong Yang, "Three-Dimensional Tissue Deformation Recovery and Tracking: Introducing techniques based on laparoscopic or endoscopic images", IEEE Signal Processing Magazine, 2010 July. IEEE Signal Processing Magazine. 27 (4): pp. 14–24.
[2] Lyudmila Mihaylova, Paul Brasnett, Nishan Canagarajan and David Bull, "Object Tracking by Particle Filtering Techniques in Video Sequences", In: Advances and Challenges in Multisensor Data and Information. NATO Security Through Science Series. Netherlands: IOS Press. pp. 260–268.
[3] Kato, Hirokazu & Mark Billinghurst, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System", IWAR '99 Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. IEEE Computer Society, Washington, DC, USA.
[4] S. Kang; J. Paik; A. Koschan; B. Abidi & M. A. Abidi, "Real-time video tracking using PTZ cameras". Proc. SPIE. 5132: pp.103–111.
[5] https://en.wikipedia.org/wiki/Lucas%E2%80%93Kanade_method
[6] Shipra Ojha , Sachin Sakhare, "Image processing techniques for object tracking in video surveillance- A survey", 2015 International Conference on Pervasive Computing (ICPC), Pune, India.
[7] Sanna Agren, "Object tracking methods and their areas of application: A meta-analysis", Master Thesis in Computing Science.
[8] Joshan Athanesious and P Suresh. Systematic survey on object tracking methods in video. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) October, pages 242–247, 2012.
[9] Comaniciu, D.; Ramesh, V.; Meer, P., "Real-time tracking of non-rigid objects using mean shift," Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, vol.2, no., pp.142,149 vol.2, 2000.
[10] Kingsbury, N G (May 2001). "Complex wavelets for shift invariant analysis and filtering of signals" (PDF). Journal of Applied and Computational Harmonic Analysis. 10 (3): 234–253. doi:10.1006/acha.2000.0343.
[11] Vapnik, Vladimir N.: Invited Speaker. IPMU Information Processing and Management 2014).
[12] Qiang He; Chee-Hung Henry Chu, Layer-based object detection and tracking with graph matching, Proceedings Volume 8020, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications VIII; 80200N (2011).
[13] Talmi, Itamar; Mechrez, Roey; Zelnik-Manor, Lihi (2016-12-07). "Template Matching with Deformable Diversity Similarity".
[14] learn.foundry.com/nuke/content/tutorials/written_tutorials/tutorial2/1_2_3_4_point_tracks.html
[15] M. Arulampalam; S. Maskell; N. Gordon & T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", IEEE Transactions on Signal Processing. 50 (2): 174.
[16] Emilio Maggio; Andrea Cavallaro, Video Tracking: Theory and Practice. 1. "Video Tracking provides a comprehensive treatment of the fundamental aspects of algorithm and application development for the task of estimating, over time", 2010.

[17] Karthik Chandrasekaran, Parametric & Non-parametric Background Subtraction Model with Object Tracking for VENUS. 1. "Background subtraction is the process by which we segment moving regions in image sequences", 2010.
[18] http://visagetechnologies.com/silhouette
[19] sa-cybernetics.github.io/blog/2013/08/12/hand-tracking-and-recognition-with-opencv/
[20] Black, James, Tim Ellis, and Paul Rosin (2003). "A Novel Method for Video Tracking Performance Evaluation". Joint IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance: pp:125–132.
[21] Gyro Stabilized Target Tracker for Off-shore Installation