

Review On Object Tracking Based On Image Subtraction

Rashi Jain, Versha Kunwar, Purna Priya

Abstract- The goal of this article is to review the state-of-the-art tracking methods, classify them into different categories, and identify new trends. In this paper, we discuss about the various techniques and method used for object tracking i.e. prediction of path of any object based on subtraction of two consecutive frames which have been generated through visualization of movement of object.

Index Terms- High-order Markov chain, graphical models, H.264/AVC spatial-temporal graph, 3D model acquisition, Augmented reality, object tracking, motion-blur, efficient second-order minimization, computed tomography (CT), motion learning, Markov random field (MRF), maximum a posterior probability (MAP), features from accelerated segment test (FAST).

1 INTRODUCTION

Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers, the availability of high quality and inexpensive video cameras, and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. There are three key steps in video analysis:

- detection of interesting moving objects,
- tracking of such objects from frame to frame, and
- analysis of object tracks to recognize their behavior.

Visual target tracking is an important task for many industrial applications such as intelligent video surveillance systems, human-machine interfaces, and remote sensing and defense systems. Despite the ubiquitous applications, visual target tracking still remains a challenging problem. This is due to the fact that in addition to the traditional tracking challenges (such as complex target motions and background clutter), the targets in visual data are always subject to deformations, occlusions, camera ego-motions, changes in scale, and various illuminations.

One can simplify tracking by imposing constraints on the motion and/or appearance of objects. For example, almost all tracking algorithms assume that the object motion is smooth with no abrupt changes. One can further constrain the object motion to be of constant velocity or constant acceleration based on a priori information. Prior knowledge about the number and the size of objects, or the object appearance and shape, can also be used to simplify the problem.

In a tracking scenario, an object can be defined as anything that is of interest for further analysis. For instance, boats on the sea, fish inside an aquarium, vehicles on a road, planes in the air, people walking on a road, or bubbles in the water

are a set of objects that may be important to track in a specific domain. Objects can be represented by their shapes and appearances.

Our survey is focussed on compiling some of the various methodologies and trackers tailored for specific objects at one place to provide readers with the information about which techniques have been used for different types of objects alongwith their advantages and limitations and therefore providing readers with the possibility to experiment with the techniques learnt.

2 TECHNIQUES AND METHODOLOGIES

Numerous approaches for object tracking have been proposed. These primarily differ from each other based on the way they approach the following questions:

- Which object representation is suitable for tracking?
- Which image features should be used?
- How should the motion, appearance, and shape of the object be modeled?

A large number of tracking methods have been proposed which attempt to answer these questions for a variety of scenarios. Some of the various algorithms used for object tracking are described as follows:

2.1 Visual Tracking Using High-Order Particle Filtering

It extended the first-order markov chain Model commonly used in visual tracking and presented a novel Framework of visual tracking using high-order Monte Carlo Markov chain. By using graphical models to obtain conditional Independence properties, a general expression for the Posterior density functions of an mth-order hidden markov model is derived. Subsequently, Sequential Importance Sampling (SIS) is used to Estimate the posterior density and obtain the high-order particle algorithm for visual object

tracking. Experimental results demonstrate that the performance of this algorithm is Superior to traditional first-order particle filtering (i.e., particle Filtering derived based on first-order markov chain).

Graphical models provide a simple way to visualize and capture the structure of the probability model. Moreover, the conditional dependence properties can be obtained by inspection of the graph. First-order hidden Markov model (HMM) is the model adopted in visual tracking, where the current state only depends on the previous state.

The algorithm has two stages, first is pre-processing of given image and after that Segmentation and then morphological operations. Steps of algorithm are as following:-

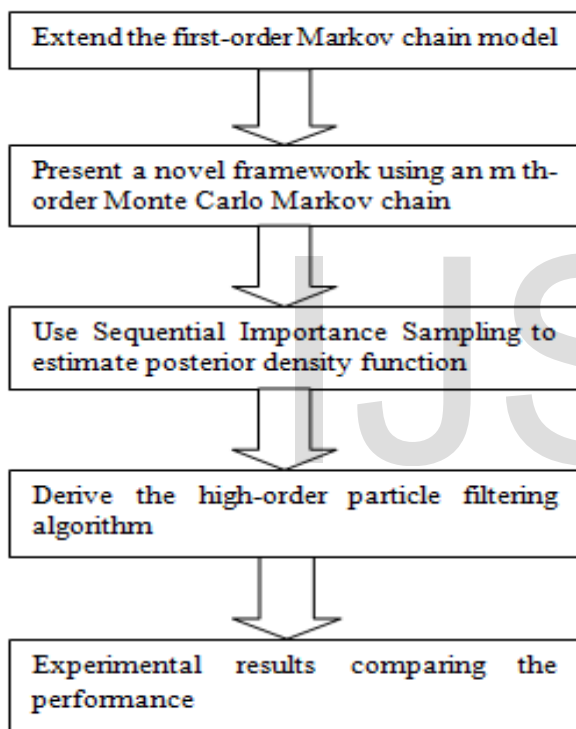


Fig. 1. Flowchart for Visual Tracking Using High-Order Particle Filtering

The extension of the first-order HMM model used in particle filtering to a high-order HMM model helps to achieve more robust and accurate visual object tracking. First assumption of first-order Markov model can simplify the expression of the posterior density estimate as well as the implementation of the visual tracking algorithm, but its disadvantage is the first-order model is not generally true and cannot accurately characterize the dynamics of moving objects. Secondly, particle filtering based on a first-order Markov model is very sensitive to loss of particle information from the previous time instant, when particles are lost or delayed this situation would arise in applications where particle updates must be transmitted over a

communication network.

2.2 Moving Object Detection And Tracking Using A Spatio-Temporal Graph In H.264/Avc Bit Streams For Video Surveillance

It is a spatio-temporal graph-based method of detecting and tracking moving objects by treating the encoded blocks with non-zero motion vectors and/or non-zero residues as potential parts of objects in H.264/AVC bit streams. A spatio-temporal graph is constructed by first clustering the encoded blocks of potential object parts into block groups, each of which is defined as an attributed sub graph where the attributes of the vertices represent the positions, motion vectors and residues of the blocks. In order to remove false-positive blocks and to track the real objects, temporal connections between sub graphs in two consecutive frames are constructed and the similarities between sub graphs are computed, which constitutes a spatio-temporal graph. The experimental results show that the proposed spatio-temporal graph-based representation of potential object blocks enables effective detection for the small-sized objects and the objects with small motion vectors and residues, and allows for reliable tracking of the detected objects even under occlusion. The identification of the detected moving objects is determined as rectangular regions of interest (ROIs) for which the ROI sizes and positions are adaptively adjusted to give the best approximation of the real shapes and positions of the objects. Method used are:

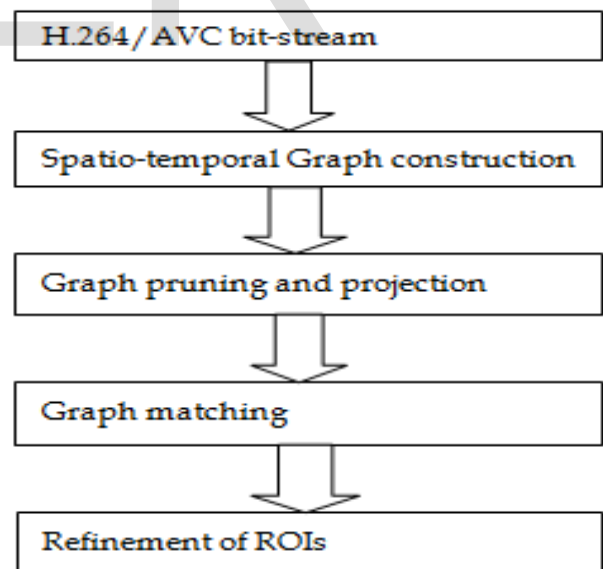


Fig. 2. Flowchart for Moving Object Detection And Tracking Using A Spatio-Temporal Graph In H.264/Avc Bit Streams For Video Surveillance

- 1) Successfully detect small objects by observing the block partitions in the smallest unit.
- 2) Remove false-positively detected block groups by spatio-

temporal graph pruning.

3) Reliably identify the moving objects frame to Frame as well as during occlusion using graph matching.

4) Adaptively adjust the ROIs to approximate the real object shapes despite of fluctuated numbers of detected blocks in block groups.

The experiment results also show that the proposed scheme can be effectively used to detect and track the objects with different sizes in various scenery conditions with accurate ROI refinement.

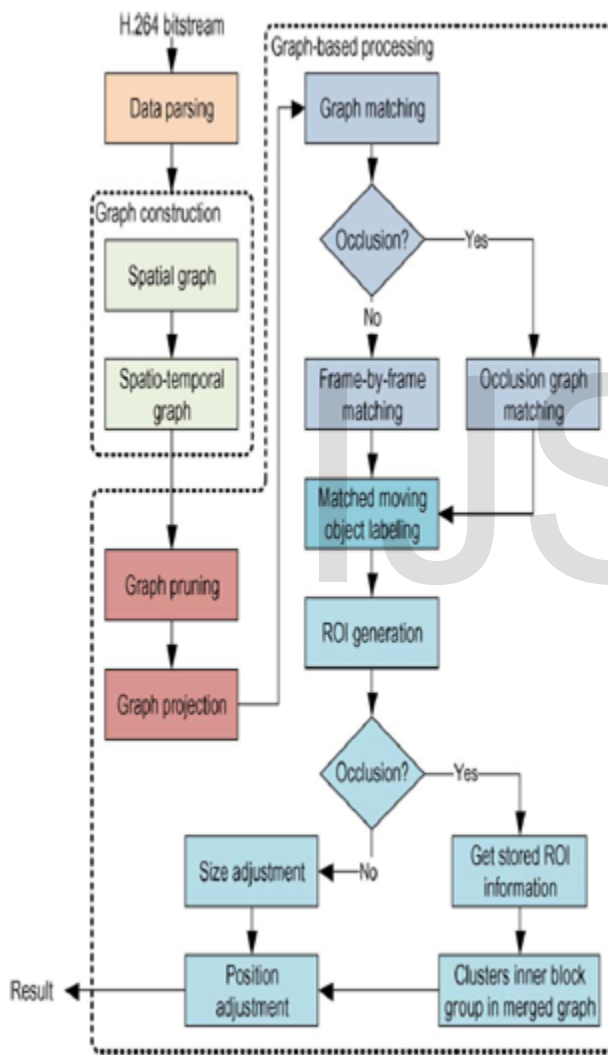


Fig. 3. can be described in two folds: 1) construction of a graph structure and 2) graph based processing. The detection and tracking of moving objects in this paper is focused only for static camera environment of surveillance with limited number of objects. However, the detection and tracking problems are dealt under more realistic environment with various objects of different sizes and movement directions, occlusions, as well as with various kinds of indoor or outdoor scenery. Although there are some works for moving camera [3]–[5], they only deal with object detection, not addressing tracking and Identification with and without occlusions.

Monitoring systems that can automatically perform such monitoring and annotating processes will increase the advantage of surveillance systems especially when humans cannot be present. Thanks to its capabilities in providing high quality compressed video with small bitrates, H.264/AVC compression standard has been widely utilized in many applications including surveillance video systems. Some real-field video surveillance systems are able to detect and even to track object movements, but not specifically identify the different objects. Thanks to its capabilities in providing high quality compressed. The number of tracked parts of the human body is limited. That is, only up to three regions of interest (ROIs) can be tracked at the processing speed of 26 frames per second.

2.3 Interactive 3D Model Acquisition and Tracking of Building Block Structures

It presents a prototype system for interactive construction and modification of 3D physical models using building blocks. This system uses a depth sensing camera and a novel algorithm for acquiring and tracking the physical models. The algorithm, Lattice-First, is based on the fact that building block structures can be arranged in a 3D point lattice where the smallest block unit is a basis in which to derive all the pieces of the model. The algorithm also makes it possible for users to interact naturally with the physical model as it is acquired, using their bare hands to add and remove pieces. The algorithm has two proof-of-concept applications: a collaborative guided assembly system where one user is interactively guided to build a structure based on another user’s design, and a game where the player must build a structure that matches an on-screen silhouette.

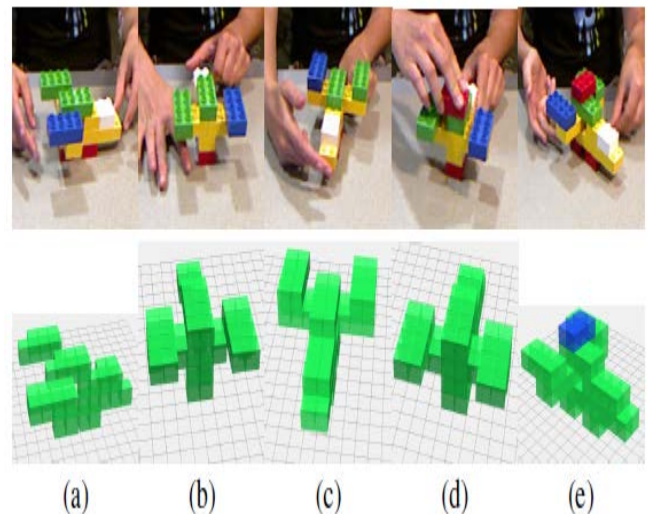


Fig. 4. Example of user interaction with a model acquisition system: (a) user places the structure in front of the sensor, and the virtual model is partially initialized; (b) (c) user rotates the structure to fill complete the model; (d) user adds a piece and then rotates the structure as (e) the model is updated with additional piece (rendered in blue).

It uses a novel algorithm, Lattice-First, that takes advantage of the orthogonal and grid-like properties of building block structures in order to achieve robustness to interference and occlusions from the user's hands, and to support a dynamic model in which pieces can be added and removed. The algorithms are fast and effective, providing users with the ability to incrementally construct a block based physical model while the system maintains the model's virtual representation. Highly interactive applications using this technology for virtual and augmented reality are still limited.

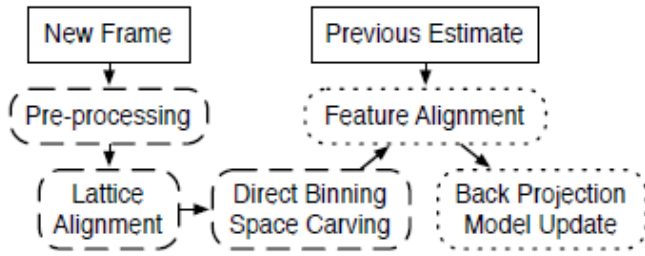


Fig. 5. High level flowchart of our proposed system. The processing stages up to Feature Alignment are computed directly for each frame, avoiding potential feedback loops with the dynamically updated model. Note that solid lines indicate data, dashed lines indicate processes computed directly for each frame, and dotted lines indicate processes that use both the current frame and the previous estimate.

2.4 Moving-Target Tracking in Single-Channel Wide-Beam SAR

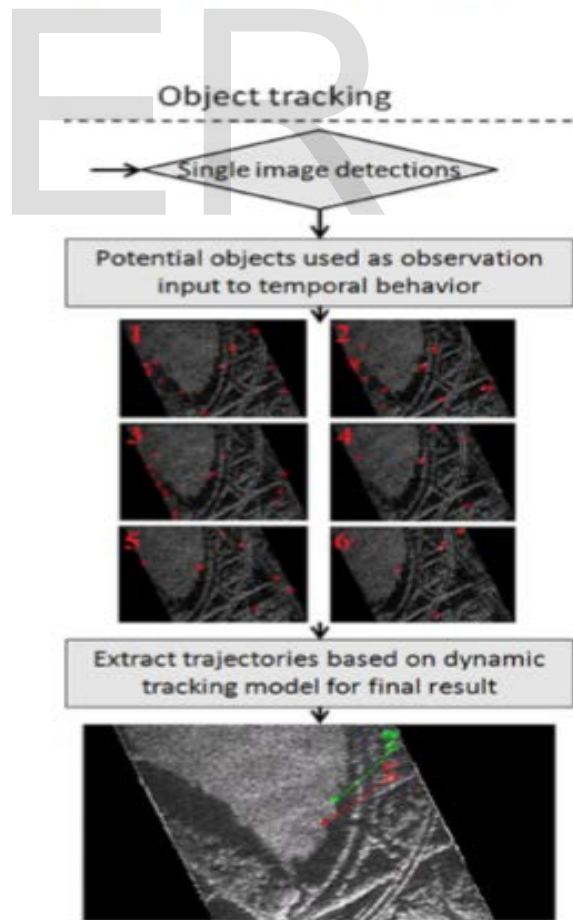
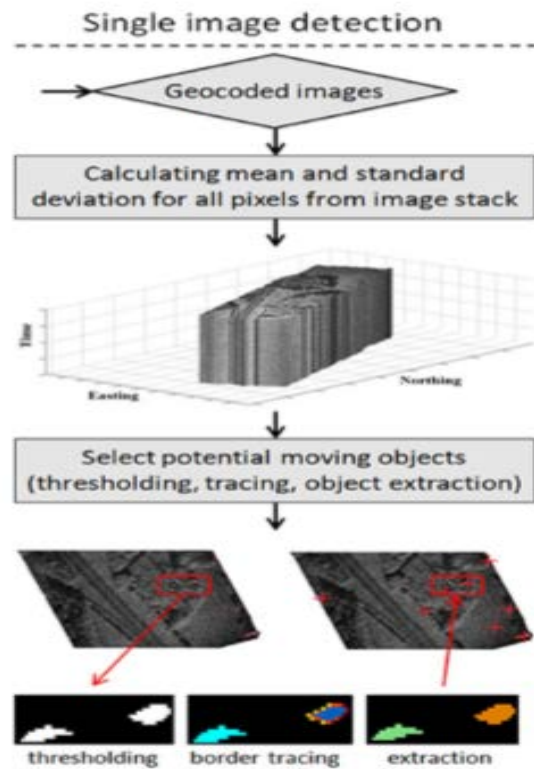
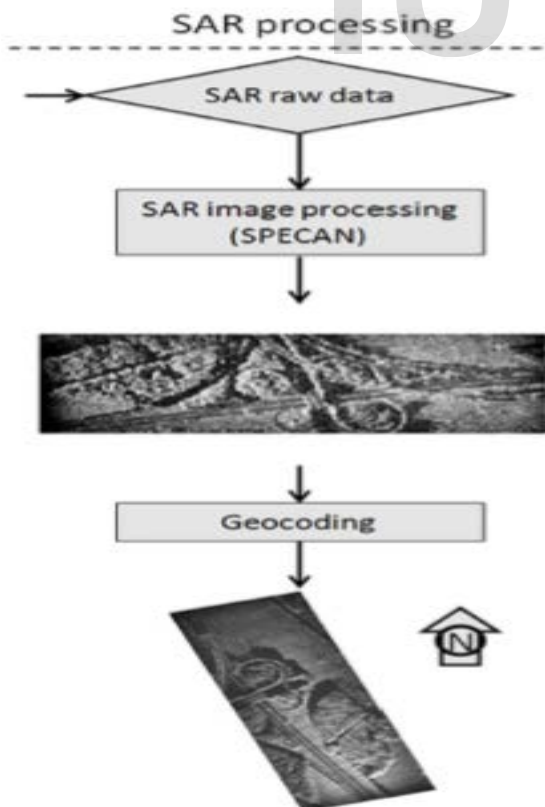


Fig. 6. Flowchart for Moving-Target Tracking in Single-Channel Wide-Beam SAR

A novel method for moving target tracking using single channel synthetic aperture radar (SAR) with a large antenna beam width is introduced and evaluated using a field experiment and real SAR data. The approach is based on sub aperture SAR processing, image statistics, and multitarget unscented Kalman filtering. The method is capable of robustly detecting and tracking moving objects over time, providing information not only about the existence of moving targets but also about their trajectories in the image space while illuminated by the radar beam. The method has been successfully applied on an experimental data set using miniature SAR to accurately characterize the movement of vehicles on a highway section in the radar image space.

This algorithm is not only used to identify moving targets but also estimate their trajectories in the radar image space. It can be practically applied to any single-channel wide beam SAR and shows promising results. The speed estimation was within an error margin of 1 m/s. A more complete validation is impaired by the lack of ground truth data.

2.5 Automatic Bootstrapping and Tracking of Object Contours

A new fully automatic object tracking and segmentation framework is proposed. The framework consists of a motion-based bootstrapping algorithm concurrent to a shape-based active contour. The shape-based active contour uses finite shape memory that is automatically and continuously built from both the bootstrap process and the active-contour object tracker. A scheme is proposed to ensure that the finite shape memory is continuously updated but forgets unnecessary information. Two new ways of automatically extracting shape information from image data given a region of interest are also proposed. Results demonstrate that the bootstrapping stage provides important motion and shape information to the object tracker. This information is found to be essential for good (fully automatic) initialization of the active contour. Further results also demonstrate convergence properties of the content of the finite shape memory and similar object tracking performance in comparison with an object tracker with unlimited shape memory. Tests with an active contour using a fixed shape prior also demonstrate superior performance for the proposed bootstrapped finite-shape-memory framework and similar performance when compared with a recently proposed active contour that uses an alternative online learning model.

This algorithm is not only applicable to object tracking in video data but also to static image object segmentation. System is capable of automatically detecting and tracking moving objects in video data without any manual intervention. Manual parameter setting may be necessary for other data. Parameters and various other configurations

can severely affect the performance of object segmentation and tracking systems.

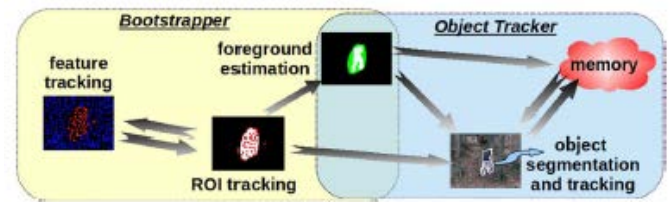


Fig. 7. Overview of the proposed system. The two major parts of the system are a bootstrapper and an object tracker. The bootstrapper is motion based and automatically extracts shape information for a subset of the frames. The object tracker continues the shape extraction process using a shape-based active contour. Both the bootstrapper and the object tracker subsystems perform shape extraction and tracking, with the object tracker performing the final role in this part of the process.

2.6 Prediction Based Collaborative Trackers (PCT): A Robust and Accurate Approach Toward 3D Medical Object Tracking

3D echocardiography and cardiac computed tomography (CT) are emerging diagnostic tools among modern imaging modalities for visualizing cardiac structure and diagnosing cardiovascular diseases. Cardiac CT can provide detailed anatomic information about the heart chambers, large vessels, and coronary arteries. Compared with other imaging modalities such as ultrasound and MR. Robust and fast 3D tracking of deformable objects, such as heart, is a challenging task because of the relatively low image contrast and speed requirement. A novel one-step forward prediction is introduced to generate the motion prior using motion manifold learning. The 3D tracking performance is limited due to dramatically increased data size, landmarks ambiguity, signal drop-out or complex nonrigid deformation.

A robust, fast, and accurate 3D tracking algorithm: prediction based collaborative trackers (PCT) is presented here. Collaborative trackers are introduced to achieve both temporal consistency and failure recovery.

We select two collaborative trackers: the detection tracker and the template tracker, which can mutually benefit each other. The detection tracker can discriminate the 3D target from background in low image quality and noisy environment. The template tracker respects the local image information and preserves the temporal consistence between adjacent frames. PCT provides the best results.

The new tracking algorithm is completely automatic and computationally efficient. It requires less than 1.5s to process a 3D volume which contains millions of voxels. It presents a robust, fast and accurate PCT and tested in both 3D ultrasound and CT. PCT can process a 3D volume in less than 1.5s and provides accuracy. It is demonstrated that PCT increases the tracking accuracy and especially speed

dramatically. The final accurate results are achieved by introducing the one step forward prediction motion prior.

The generality of PCT is already proven by adverse and extensive experiments using three challenging heart tracking applications. Currently, work is going on extending PCT to other tracking applications, such as lung tumor tracking in continuously acquired fluoroscopic video sequences.

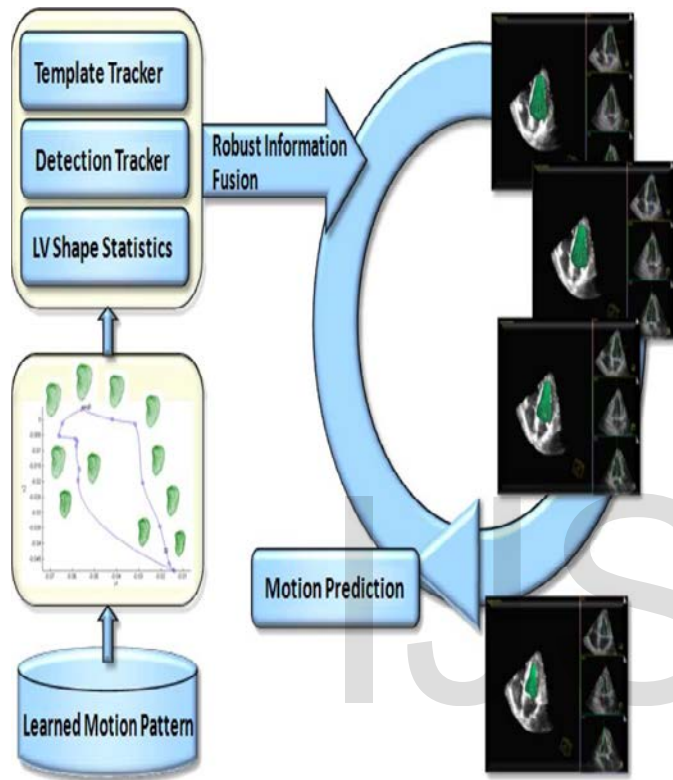


Fig. 8. The flow chart of the PCTs using one step forward prediction, marginal space learning, motion manifold learning, and detection/template trackers.

2.7 A Change Information Based Fast Algorithm for Video Object Detection and Tracking

The proposed algorithm includes two schemes: one for spatio-temporal spatial segmentation and the other for temporal segmentation. A combination of these schemes is used to identify moving objects and to track them. A compound Markov random field (MRF) model is used as the prior image attribute model, which takes care of the spatial distribution of color, temporal color coherence and edge map in the temporal frames to obtain a spatio-temporal spatial segmentation. In this scheme, segmentation is considered as a pixel labeling problem and is solved using the maximum a posterior probability (MAP) estimation technique. The MRF-MAP framework is computation intensive due to random initialization. To reduce this burden, change information based heuristic initialization technique is proposed. The scheme requires an initially segmented frame. For initial frame segmentation, compound MRF model is used to model attributes and

MAP estimate is obtained by a hybrid algorithm [combination of both simulated annealing (SA) and iterative conditional mode (ICM)] that converges fast.

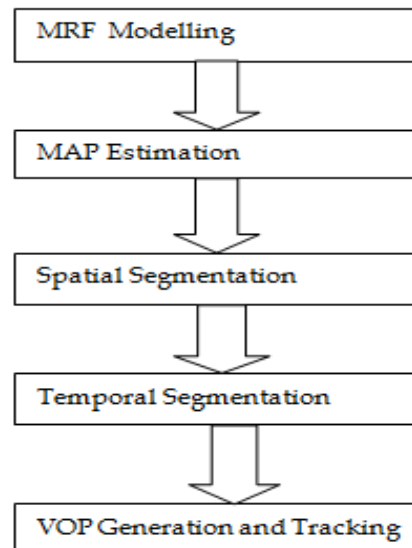


Fig. 9. Flowchart for A Change Information Based Fast Algorithm for Video Object Detection and Tracking

It produces better segmentation results compared to those of edgeless and JSEG segmentation schemes and comparable results with edge based approach. The proposed scheme gives better accuracy and is approximately 13 times faster compared to the considered MRF based segmentation schemes for a number of video sequences.

2.8 Adaptive Sampling for Feature Detection, Tracking, and Recognition on Mobile Platforms

It proposes a novel approach to speed up feature detectors and to inform feature tracking that speeds up the recognition process by using the concept of adaptive sampling. The algorithms used here are based on considering the result of the computation on a reduced or previous sample to decide whether to continue the sampling and computation steps or to stop. Therefore, these algorithms adapt based on the image content. The adaptive sampling term is used to refer to this property.

The feature detection is done using a modified version of features from accelerated segment test (FAST) called as M-FAST algorithm which results in speed up of up to 50%. It also proposes using adaptive sampling to combine tracking and recognition. It proposes a dynamic length descriptor based on BRIEF to make more efficient use of different performances within the tracking pipeline. This method has a lower memory footprint and was faster than using two different types of descriptors one for tracking and one for recognition. Since in tracking, the spatial distance between the matched feature pairs should be smaller than a given

threshold, when searching for a match, first the spatial distances should be computed. To speed up the computation of distance between pairs, spatial binning is used.

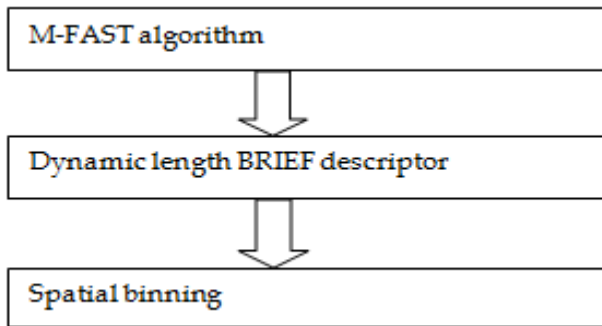


Fig. 10. Flowchart for Adaptive Sampling for Feature Detection, Tracking, and Recognition on Mobile Platforms

It shows how processing time and memory footprint can benefit by this approach with little impact on overall output quality.

2.9 Handling Motion-Blur in 3D Tracking and Rendering for Augmented Reality

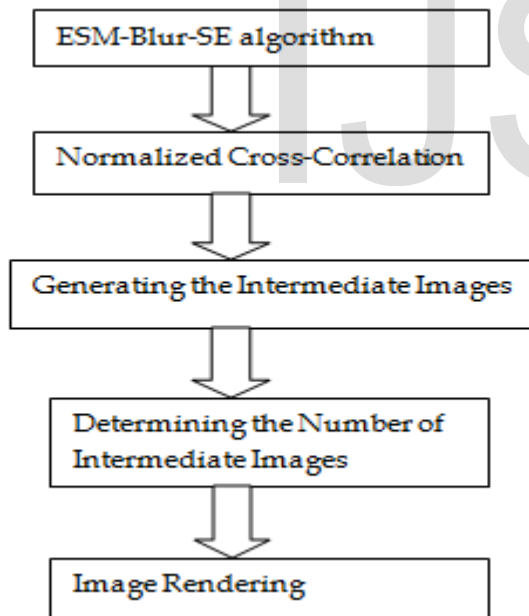


Fig. 11. Flowchart for Handling Motion-Blur in 3D Tracking and Rendering for Augmented Reality

Efficient Second-order Minimization (ESM) is a powerful algorithm for template matching-based tracking, but it can fail under motion blur. This paper introduces an image formation model that explicitly consider the possibility of blur, and shows its results in a generalization of the original ESM algorithm. This allows to converge faster, more accurately and more robustly even under large amount of blur. Template matching based approach is used here

because it does not rely on feature extraction, which can be dangerous when motion blur erases most of the image features, and because it usually yields very accurate results. This paper also presents an efficient method for rendering the virtual objects under the estimated motion blur. It renders two images of the object under 3D perspective, and warps them to create many intermediate images. By fusing these images a final image for the virtual objects is obtained blurred consistently with the captured image. Because warping is much faster than 3D rendering, it can create realistically blurred images at a very low computational cost.

2.10 Fusing Multiple Independent Estimates via Spectral Clustering for Robust Visual Tracking

One fundamental problem of object tracking is the convergence of estimates to local maxima not corresponding to target objects. To mitigate this problem, constructing a good posterior distribution of the target state is important. In this paper, a robust tracking approach by building a new posterior distribution model from multiple independent estimates of a target state is proposed. For this purpose, initial target estimates by running multiple independent estimators (object trackers or detectors with different visual cues) are obtained. Specifically a color (red and green)-histogram-based tracker, a SURF-based tracker, and an object detector trained with HoG features and a support vector machine (SVM) classifier is used. In addition, the target estimate based on a velocity model to ensure motion consistency of the target object is obtained. Given the multiple independent estimates of target states, a confidence score is assigned to each estimate by examining its consistency with others and photometric similarities with target models using a spectral clustering method. Then a new hypotheses to find the target in the subsequent frames on the basis of the confidences is generated.

The proposed approach is robust to severe or long-term occlusion, and hence is able to recover from significant tracking errors. VTD and particle-filter-based trackers commonly use adaptive appearance models but they could not overcome the local maxima problem.

3 CONCLUSION

In this article, we present a survey of object tracking methods and also give a brief review of related topics. We divide the tracking methods based on the names of the research papers from which they have been included. We believe that, this article can give valuable insight into this important research topic and encourage new research.

REFERENCE

[1] Pan Pan and Dan Schonfeld, "Visual Tracking Using High-Order Particle Filtering," *IEEE Signal Processing Letters*, vol. 18, no. 1, pp.

- 51-54, January 2011, doi: 10.1109/LSP.2010.2091406
- [2] Houari Sabirin and Munchurl Kim, "Moving Object Detection and Tracking Using a Spatio-Temporal Graph in H.264/AVC Bit streams for Video Surveillance," *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 657-668, June 2012, doi: 10.1109/TMM.2012.2187777
- [3] Andrew Miller, Brandyn White, Emiko Charbonneau, Zach Kanzler and Joseph J. LaViola Jr., "Interactive 3D Model Acquisition and Tracking of Building Block Structures," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 4, pp. 651-659, April 2012
- [4] Youngmin Park, Vincent Lepetit and Woontack Woo, "Handling Motion-Blur in 3D Tracking and Rendering for Augmented Reality," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 9, pp. 1449-1459, September 2012, doi: 10.1109/TVCG.2011.158
- [5] Jung-ho Kim, Jihong Min, In So Kweon and Zhe Lin, "Fusing Multiple Independent Estimates via Spectral Clustering for Robust Visual Tracking," *IEEE Signal Processing Letters*, vol. 19, no. 8, pp. 527-530, August 2012, doi: 10.1109/LSP.2012.2205916
- [6] Mosalam Ebrahimi and Walterio W. Mayol-Cuevas, "Adaptive Sampling for Feature Detection, Tracking, and Recognition on Mobile Platforms," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 10, pp. 1467-1475, October 2011, doi: 10.1109/TCSVT.2011.2163450
- [7] Badri Narayan Subudhi, Pradipta Kumar Nanda and Ashish Ghosh, "A Change Information Based Fast Algorithm for Video Object Detection and Tracking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 7, pp. 993-1004, July 2011, doi: 10.1109/TCSVT.2011.2133870
- [8] Lin Yang, Bogdan Georgescu, Yefeng Zheng, Yang Wang, Peter Meer and Dorin Comaniciu, "Prediction Based Collaborative Trackers (PCT): A Robust and Accurate Approach toward 3D Medical Object Tracking," *IEEE Transactions on Medical Imaging*, vol. 30, no. 11, pp. 1921-1932, November 2011, doi: 10.1109/TMI.2011.2158440
- [9] John Chiverton, Xianghua Xie and Majid Mirmehdi, "Automatic Bootstrapping and Tracking of Object Contours," *IEEE Transactions on Image Processing*, vol. 21, no. 3, pp. 1231-1245, March 2012, doi: 10.1109/TIP.2011.2167343
- [10] Daniel Henke, Christophe Magnard, Max Frioud, David Small, Erich Meier and Michael E. Schaepman, "Moving-Target Tracking in Single-Channel Wide-Beam SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4735-4747, November 2012, doi: 10.1109/TGRS.2012.2191561
- [11] Yan Zhai, Mark B. Yeary, Samuel Cheng, and Nasser Kehtarnavaz, "An Object-Tracking Algorithm Based on Multiple-Model Particle Filtering With State Partitioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 5, pp. 1797-1809, May 2009, doi: 10.1109/TIM.2009.2014511
- [12] Alper Yilmaz, Omar Javed and Mubarak Shah, "Object Tracking: A Survey," *ACM Computing Surveys*, Vol. 38, No. 4, Article 13, Publication date: December 2006, doi: 10.1145/1177352.1177355