Detecting Occlusion and Camouflage during Visual Tracking

T. Chandesa
School of Computer Science
University of Nottingham
Malaysia Campus
Jalan Broga, Semenyih, 43500,
Selangor Darul Ehsan, Malaysia

T. Pridmore
School of Computer Science
University of Nottingham,
Jubilee Campus
Wollaton Road, Nottingham, UK

A. Bargiela
School of Computer Science
University of Nottingham
Malaysia Campus
Jalan Broga, Semenyih, 43500,
Selangor Darul Ehsan, Malaysia

Abstract **Visual tracking is an important scientific problem; the human visual system is capable of tracking moving objects in a wide variety of situations. It is also of considerable practical importance; many actual and potential applications of visual tracking algorithms exist in domains such as surveillance, medicine, robotics and the media. Although many effective tracking algorithms exist, occlusion and camouflage remain a common problem. These can cause a tracker to become dissociated from its target, so that the data it produces is unrelated to the target's behaviour. We focus on the detection of occlusion and camouflage during particle filter-based tracking. We use a Gaussian Mixture Model of particle distribution, extracted via the EM algorithm, to investigate the effects of occlusion and camouflage on the particle set representing a given target. The information gained from this investigation informs the design of process-behaviour chart which alerts the tracker of the occurrence of occlusion or camouflage. Data produced by the process-behaviour chart is also used to map out the boundary of the interfering object, providing valuable information about the viewed environment.**

I. INTRODUCTION

Visual tracking is a motion estimation process used when the target's velocity is too large to allow the use of derivatives. Features are extracted from neighbouring frames and matched between those frames to allow their motion to be recovered. Data is provided only at selected targets, but both the camera and/or target can be in motion. Visual tracking is an important scientific problem; the human visual system is capable of tracking moving objects in a wide variety of situations. It is also of considerable practical importance; many actual and potential applications of visual tracking algorithms exist in domains such as surveillance, medicine, robotics and the media.

Several powerful visual tracking techniques exist. Most are based upon model estimation concepts, such as Kalman filters or Sequential Monte Carlo methods, with particle filters receiving particular attention [1]. Though these have been shown to be successful in their own right, each has its strengths and weaknesses. For example, the Kalman filter is proven to converge, but is limited to linear motion and uni-modal priors. Particle filters can represent multi-modal distributions. These multi-modal distributions allow multiple hypotheses to be maintained and are represented by sets of discrete particles that are often shown as spots of different colour or intensity overlaid on image data.

While many effective tracking algorithms have been based on the particle filter, common problems such as image and motion noise, occlusion and camouflage remain. Occlusion arises when another object, usually with different features, falls between the camera and the target. Camouflage occurs when an object with similar features lies behind the target and makes the target invisible from the camera's point of view. Either of these events can cause the tracker to become dissociated from its target, so that the data it produces is unrelated to the target's behaviour.

Previous tracking techniques have sought to avoid these problems by keeping the tracker more tightly focused on the target, so that camouflage is simply not noticed and the tracker is more likely to reacquire the target following (short periods of) occlusion. Multiple motion models have been used, as have more detailed texture and colour cues that better model target appearance. Alternatively, areas of the environment in which the target is more likely to appear might be identified. While they improve tracker performance, these implicit methods do not entirely remove the problems of occlusion and camouflage. We take an alternative approach, arguing that the solution lies not in avoiding, but in detecting and explicitly reacting to these disruptive events.

In what follows, we focus on detecting occlusion and camouflage in a particle-filter based tracker by investigating the particle distributions that arise as particle-filter-tracked targets become occluded and/or camouflaged. The Expectation Maximization (EM) algorithm [2] is employed to cluster particles at each time step, providing vital information about the behaviour of particles whenever a target experiences occlusion and/or camouflage. This information is assessed using a process-behaviour chart which alerts the tracker to the occurrence of occlusion or camouflage. The information produced by the process-behaviour chart can then be used to map out the boundary of the interfering object, providing valuable information about the viewed environment. The ability to detect occlusion and camouflage raises the possibility of the tracker reacting appropriately to each event, though consideration of this is beyond the scope of the current paper. Though the method could be applied to any particle filter algorithm, Condensation [1] is used throughout.

The remainder of the paper is structured as follows. Section II reviews related work on occlusion and camouflage handling, the EM algorithm and particle clustering. The implementation of the particle clustering approach is discussed in section III. Section IV presents the process-behaviour charts used to detect occlusion and camouflage. Section V discusses the mapping approach used to identify the boundaries of interfering objects, and presents results obtained from real and artificial image sequences. Possible alternatives and extensions to the current methods are discussed in section VI before conclusions are drawn in section VII.

II. Related Work

A. *Occlusion and Camouflage Handling*

The occurrence of occlusion and camouflage is inevitable in visual tracking, and a variety of approaches have been adopted in response.

Some researchers [3] have tried to overcome occlusion by improving the representation used within their tracker or modifying the underlying search engine [4]. Others attempt to use depth or trajectory reasoning [5, 6] to overcome it, while [7] incorporates a binary visibility process into the observation model.

Camouflage has received less attention in the vision literature. In a substantial review, the only literature found that explicitly discusses camouflage is [8], where they used a stochastic process to handle the occurrence of camouflage whenever a new observation cannot be associated with an existing track. Their method does not, however, seek to detect these events as is done here.

Examination of the literature has shown only limited analysis of the changes taking place within a tracker during occlusion and camouflage: over the few frames it takes for the target to transition from normal visibility to full occlusion or camouflage. Closer examination of the measurement stage of a particle filter-based tracker provides useful information on the changes taking place within the particle set during occlusion and camouflage events. Measurement steps are often quite complex, powerful operations but can provide a lot of information about the target's local environment. In addition, particle filters are particularly good at providing this information, as they sample from both the target and its surroundings.

B. *Expectation Maximization and Particle Clustering*

The Expectation Maximization (EM) algorithm is an iterative maximum likelihood estimation process, mainly used when the data of interest are incomplete. The EM algorithm is widely applicable and has been used in numerous fields including biology, artificial intelligence and computer vision. Examination of the literature shows the EM algorithm to have been applied to a diverse range of problems in computer vision. Problems such as classification [9], motion estimation [10], optimization [11], segmentation and clustering [12, 13] have all utilized the EM algorithm successfully.

In visual tracking, the EM algorithm has been used to improve temporal appearance model parameters [14], motion and trajectory models [15, 16], and colour and spatial models [17, 18]. Meanwhile, in [19], the interpretation of scene content and camera positions was improved by using the EM algorithm. The work reported in [20] showed the EM algorithm to be capable of empirically estimating tracking errors.

The method presented here uses the Expectation Maximization algorithm to cluster particles at each time step of a particle filter-based tracking (Condensation [1]) process. The resulting descriptions of the particle set contain valuable information about the behaviour of the particles during occlusion and camouflage events. This information informs the design of a process-behaviour chart which alerts the tracker in the event of occlusion and camouflage.

Clustering of particles is not the norm in visual tracking, but has been done, as shown in [21] and [22]. In [22], the EM algorithm was used as a clustering tool to cluster together particles from one time frame. Cluster membership was then taken into account when particles were propagated into the next time frame.

III. Modelling Particle Distributions

Our approach differs from the work reported in [21] and [22] in a number of ways. Here, cluster membership is less important that the global properties (or parameters) of the clusters produced. Moreover, clusters are not propagated forward in time, but are recalculated at each time step. Particles do not stay within the same cluster throughout the tracking process, and clusters may vary dramatically in size. The number of clusters found is also expected to change as tracking proceeds

In our approach, an iterative EM algorithm is used to build a Gaussian mixture model to describe the particle set present at every time frame. The clusters' mean and deviation values are weighted using (1) and (2) respectively.

$$\mu_x = \frac{\sum_{i=1}^{n} P(\theta_i \mid x) x_i}{\sum_{i=1}^{n} P(\theta_i \mid x)} \tag{1}$$

$$\sum{}_x = \sqrt{\frac{\sum_{i=1}^{n} P(\theta_i \mid x)\left(\sqrt{(x_i - \mu_i)^2}\right)^2}{\sum_{i=1}^{n} P(\theta_i \mid x)}} \tag{2}$$

$P(\theta_i \mid x)$ is the probability of the particle being in a particular cluster and n is the total number of particles. Improved parameters for each cluster at time $t+1$ are computed by taking into consideration the probability of each particle being in each cluster. The convergence of the EM algorithm is based on the Euclidean distance between clusters at time $t+1$ and t.

When the EM algorithm has successfully converged, particles are assigned to mixture components based on the clusters having the highest probability value and the clusters' parameters are recomputed before feeding back the information to the tracker.

Fig. 1 and Fig. 2 show graphical representations of the result of applying the Condensation + EM algorithm to real videos. In each case EM reports a single particle cluster, with the mean particle position being shown as a pink dot and the deviation shown as a pink circle. Fig. 1 exhibits an occurrence of occlusion, as the person walking along the path steps behind the tree. Fig. 2 exhibits an occurrence of camouflage, as the tracked white football moves in front of the white-shirted players

Fig. 1. Gaussian Mixture Model describing Condensation particles during occlusion

Fig. 2. Gaussian Mixture Model describing Condensation particles during camouflage

Analysis of these (and similar) examples leads to the following hypotheses:

- When occlusion occurs, the particles stop tracking the target (woman) and become closely clustered together. The particles fail to re-attach themselves onto the target when the target reappears from being occlusion.
- When camouflage occurs, the particles get transferred onto the camouflage object and start to spread widely. The particles never again re-attach themselves onto the target (ball).

These effects are reflected in the parameters of the Gaussian Mixture Model produced by EM:

- During occlusion, the cluster deviation drops but the speed of the particles within the cluster increases. The reduction in deviation is due to the particles being clustered at the back of the target. The increase in speed occurs because during occlusion, a small number of particles will land on the occluding object, move quickly across it and then be destroyed as their weight drops to zero.
- During camouflage, the cluster deviation increases, as does the speed of the particles within the cluster. The increase in deviation and speed can be attributed to a general expansion of the cluster when the particles are transferred onto the similar, but larger, camouflaging object.

The ability to reliably detect these changes in the particle set would enable the tracker to automatically identify occlusion and camouflage events.

IV. Process-Behaviour Charts

Process-behaviour charts are a statistical graphical tool used to monitor the behaviour of fluctuation data. Literature has shown that process-behaviour charts are primarily a control engineering tool and have not previously been used in visual tracking. The Shewhart [23] and Exponentially Weighted Moving Average (EWMA) [24] control charts are the most commonly used. EWMA charts are used when small changes in data mean and variance must be detected, as shown in [25]. Shewhart charts are typically used to detect large changes in the mean and variation of some data value.

The Shewhart chart approach is employed here because particle fluctuations are quite large. Maximum and minimum control limits are calculated using (3) and (4). These control limits are used as a measurement threshold to determine whether the data in question is still under control or has gone out of control. The maximum and minimum control points are computed from the mean of the data.

$$MaximumControlLimit = \mu + 3\sigma \tag{3}$$

$$MinimumControlLimit = \mu - 3\sigma \tag{4}$$

In our method, the Shewhart chart parameters are computed over a moving window using the cluster results obtained from the EM algorithm. The moving window interval information is not gathered during the first 5 frames of the sequence as, during this stage, the particles are spreading within the boundary of the target object. The control limit is determined based on the Nelson rule. In the event of the mean value exceeding either of these control limits, the tracker notifies the user that the target in question is experiencing either occlusion or camouflage.

Fig. 3 and Fig. 4 show the Shewhart control chart analysis for occlusion applied to the image sequence illustrated in Fig. 1. Meanwhile, Fig. 5 and Fig. 6 show the Shewhart control chart analysis for camouflage applied to data extracted from the image sequence illustrated in Fig. 2. The control charts shown in Fig. 3 and Fig. 5 focus on the clustered particle deviation, while those shown in Fig. 4 and Fig. 6 focus on the clustered particle speed.

Analysis shows that the clustered particle deviation means provide a clearer indicator of the occurrence of occlusion and camouflage. As a result, the clustered particle deviation means are used as the primary tool to determine when occlusion and camouflage have occurred, with the clustered particle speed mean providing supplementary information.

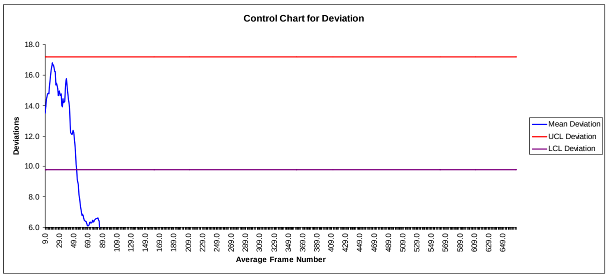


Fig. 3. Control chart for clustered particle deviation when tracking experiences occlusion (Fig. 1)

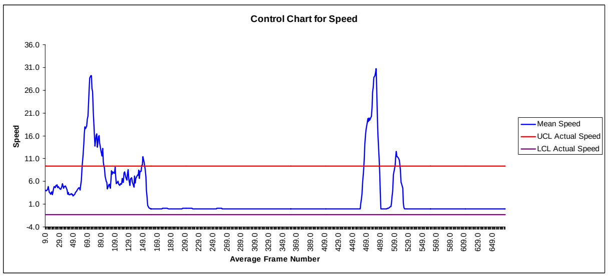


Fig. 4. Control chart for clustered particle speed when tracking experiences occlusion (Fig. 1.)

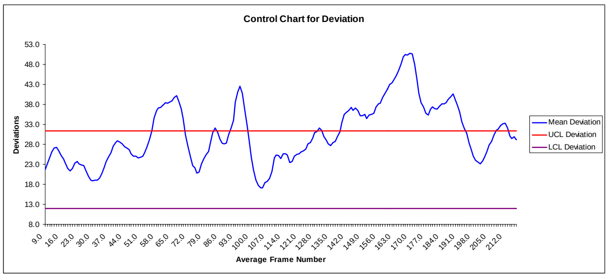


Fig. 5. Control chart for clustered particle deviation when tracking experiences camouflage (Fig. 2.)

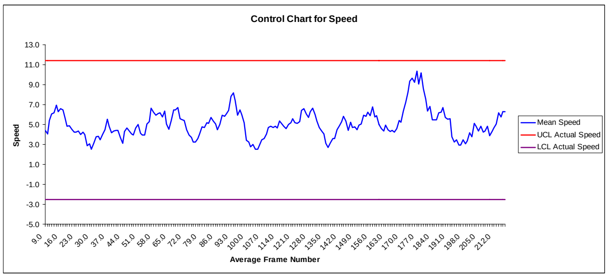
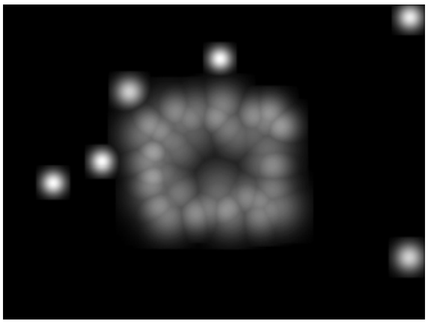


Fig. 6. Control chart for clustered particle speed when tracking experiences camouflage (Fig. 2.)

V. Results: Occlusion and Camouflage Mapping

When the process-behaviour charts fires, i.e. when the mean deviation value of a cluster moves outside the acceptable range, the position and spread of the cluster provide an indication of the image location at which occlusion or camouflage occurred. If the camera is fixed, results obtained by tracking multiple targets through the same environment can be combined to produce an occlusion or camouflage map of the background scene. Each time an event is detected, the mixture model component describing the relevant particle cluster is added to a second Gaussian Mixture Model, situated on the image plane and describing the occlusion or camouflage structure of the viewed environment. The resulting maps are both useful in themselves, and a good indicator of the performance of the event detection method.

Fig. 7 shows occlusion and camouflage maps created using simple artificial image sequences. Here, coloured circles are randomly placed at the boundaries of an image containing an identically (when considering camouflage) or differently (when considering occlusion) coloured rectangle. The circles are tracked until the process-behaviour chart fires, and then the Gaussian describing the particle cluster is added to a developing camouflage or occlusion map. The boundaries of the central rectangular region are clear in both, though it should be noted that the occlusion detector fires just outside, and the camouflage detector just inside, the interfering object.



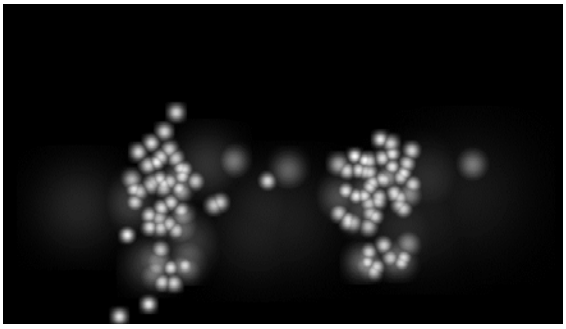
(a)

(b)

Fig. 7. (a) camouflage map, and (b) occlusion map obtained from simple artificial sequences in which the interfering object is a centrally placed rectangle.

Fig. 8 shows an occlusion map obtained from a fixed camera monitoring an outdoor walkway at the University of Nottingham Malaysia Campus. The walkway comprises a roof supported by a series of narrow vertical pillars which occlude those using it. Several tens of pedestrians were tracked using Condensation employing an RGB colour histogram target representation. Fig. 8a shows a sample image from the sequence, Fig 8b shows the resulting occlusion map. Note that although the occlusion map entries are gathered around the pillars, the pillar boundaries are less clear than in the artificial data of Fig. 7. This is because the targets tracked here are of variable size and, more importantly, most are wider than the occluding object.

(a)



(b)

Fig. 8. (a) Actual scene, (b) Occlusion map.

VI. Discussion

The EM algorithm-enhanced Condensation tracker implemented in this paper exploits only the position of the particles, the Condensation weights are not taken into consideration. The Condensation weights provide the tracker with information as to which particles are most likely to remain on the target and which ones are most likely to be propagated into the next frame. This extra information can be incorporated into the occlusion/camouflage detection method described above by simply extending the EM algorithm from two to three dimensions. However, comparing Fig. 9 with Fig. 11 and Fig. 10 with Fig. 12, demonstrates that the clustered particle deviation and speed undergoes no significant changes with the inclusion of the Condensation weights.

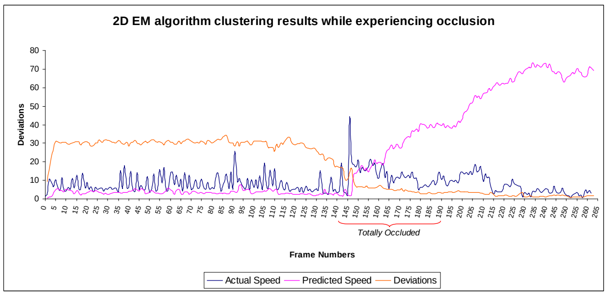


Fig. 9. 2D EM algorithm clustering while experiencing occlusion

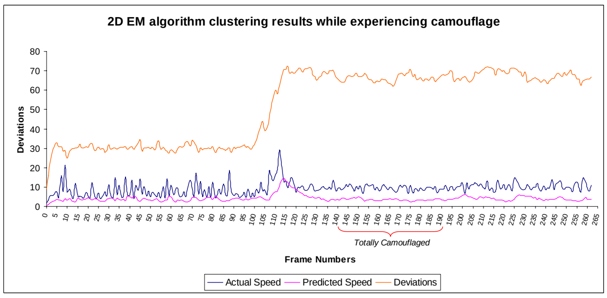


Fig. 10. 2D EM algorithm clustering while experiencing camouflage

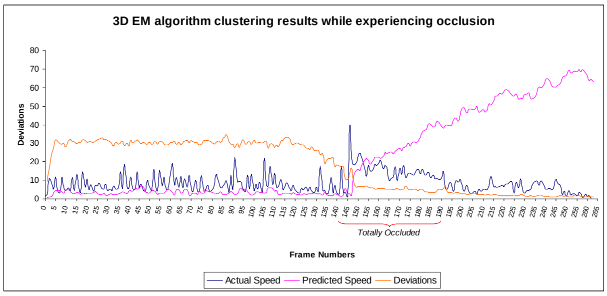


Fig. 11. 3D EM algorithm clustering while experiencing occlusion

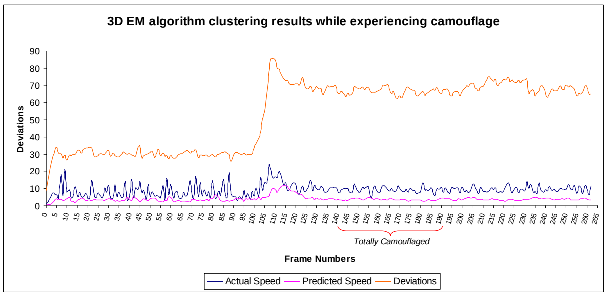


Fig. 12. 3D EM algorithm clustering while experiencing camouflage

In this initial study we have adopted the Shewhart control chart. Other control chart methods could be employed. A more sensitive EWMA could be implemented if smaller changes were required to be detected. The EWMA algorithm, however, relies on a user-specified parameter, lambda, the choice of which would be vital in making the chart sensitive to the changes sought. The value of lambda must be between 0 and 1, with values nearing 0 giving priority to older data and values near 1 will giving priority to new data. A lambda value of 1 will produce a Shewhart chart-like result. The choice of lambda will determine the distance of the maximum and minimum control points from the mean. Moreover, an initial lambda choice may not be applicable at detecting changes in mean and variance for all events, e.g. tracking normally, occlusion and/or camouflage.

The experiments conducted here considered only a single cluster at each time step. Future work will extend the approach to deal with multiple particle clusters, allowing the tracker to detect (and react appropriately to) clutter. Clutter disrupts tracking by camouflaging the target. However, camouflage, as the word is normally used, has only one effect on the particle set (to spread it out) and clutter has two (to spread it out and have it become multimodal).

VII. Conclusion

The novelty of the work reported here is three-fold. First, the EM algorithm is used within a particle-filtered (Condensation) tracker to analyse the effects of occlusion and camouflage and identify cues related to occlusion or camouflage. Secondly, by using a process-behaviour chart, the information gathered from that analysis is exploited to allow the tracker to determine when a target is experiencing occlusion and/or camouflage by monitoring the control points. Finally, the information produced by the process-behaviour chart is used by the tracker to build a Gaussian mixture model map of the boundaries of the interfering object(s), marking areas of the background environment in which occlusion and camouflage are likely to occur. This opens ups the possibility of the tracker both detecting and responding appropriately to inevitable, highly disruptive events.

References

[1] M. Isard, and A. Blake, “CONDENSATION—Conditional Density Propagation for Visual Tracking,” *International Journal of Computer Vision*, vol. 29, no. 1, pp. 5-28, 1998.

[2] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1-38, 1977.

[3] J. Jeyakar, R. V. Babu, and K. R. Ramakrishnan, “Robust object tracking with background-weighted local kernels,” *Computer Vision and Image Understanding*, vol. 112, no. 3, pp. 296-309, 2008.

[4] Y. Ma, Q. Yu, and I. Cohen, “Target tracking with incomplete detection,” *Computer Vision and Image Understanding*, vol. 113, no. 4, pp. 580-587, 2009.

[5] D. Greenhill, J. Renno, J. Orwell *et al.*, “Occlusion analysis: Learning and utilising depth maps in object tracking,” *Image and Vision Computing*, vol. 26, no. 3, pp. 430-441, 2008.

[6] R. Rosales, and S. Sclaroff, "Improved Tracking of Multiple Humans with Trajectory Prediction and Occlusion Modelling," *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Workshop on the Interpretation of Visual Motion*, Santa Barbara, 1998.

[7] P. Pérez, and J. Vermaak, "Bayesian tracking with auxiliary discrete processes. Application to detection and tracking of objects with occlusions," *In Proc. ICCV'05 Workshop on Dynamical Vision*, Beijing, China, October, 2005.

[8] P. KaewTrakulPong, and R. Bowden, “A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes,” *Image and Vision Computing*, vol. 21, no. 10, pp. 913-929, 2003.

[9] A. Lobay, and D. Forsyth, “Shape from Texture without Boundaries,” *International Journal of Computer Vision*, vol. 67, no. 1, pp. 71-91, 2006.

[10] W. Yu, G. Sommer, and K. Daniilidis, “Three dimensional orientation signatures with conic kernel filtering for multiple motion analysis,” *Image and Vision Computing*, vol. 21, no. 5, pp. 447-458, 2003.

[11] S. Ji, L. T. Watson, and L. Carin, “Semisupervised Learning of Hidden Markov Models via a Homotopy Method,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 275-287, 2009.

[12] J. F. Khan, R. R. Adhami, and S. M. A. Bhuiyan, “A customized Gabor filter for unsupervised color image segmentation,” *Image and Vision Computing*, vol. 27, no. 4, pp. 489-501, 2009.

[13] C. J. Veenman, M. J. T. Reinders, and E. Backer, “A Maximum Variance Cluster Algorithm,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 9, pp. 1273-1280, 2002.

[14] A. D. Jepson, D. J. Fleet, and T. R. El-Maraghi, "Robust online appearance models for visual tracking," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 10, pp. 1296-1311, 2003.

[15] P. Smith, T. Drummond, and R. Cipolla, “Layered motion segmentation and depth ordering by tracking edges,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 4, pp. 479-494, 2004.

[16] M. Gelgon, P. Bouthemy, and J.-P. Le Cadre, “Recovery of the trajectories of multiple moving objects in an image sequence with a PMHT approach,” *Image and Vision Computing*, vol. 23, no. 1, pp. 19-31, 2005.

[17] Y. Ren, and C. S. Chua, “Bilateral learning for color-based tracking,” *Image and Vision Computing*, vol. 26, no. 11, pp. 1530-1539, 2008.

[18] Y. Wu, and T. S. Huang, “Robust Visual Tracking by Integrating Multiple Cues Based on Co-Inference Learning,” *International Journal of Computer Vision*, vol. 58, no. 1, pp. 55-71, 2004.

[19] R. Stolkin, A. Greig, M. Hodgetts, and J. Gilby, “An EM/E-MRF algorithm for adaptive model based tracking in extremely poor visibility,” *Image and Vision Computing*, vol. 26, no. 4, pp. 480-495, 2008.

[20] H. Veeraraghavan, P. Schrater, and N. Papanikolopoulos, “Robust target detection and tracking through integration of motion, color, and geometry,” *Computer Vision and Image Understanding*, vol. 103, no. 2, pp. 121-138, 2006.

[21] A. Milstein, J. N. Sánchez, and E. T. Williamson, “Robust global localization using clustered particle filtering,” in *Eighteenth national conference on Artificial intelligence*, Edmonton, Alberta, Canada, pp. 581-586, 2002.

[22] J. Vermaak, A. Doucet, and P. Perez, "Maintaining multimodality through mixture tracking," *Proceedings. Ninth IEEE International Conference on Computer Vision, IEEE Computer* Society, pp. 1110-1116 vol. 2, October 13-16, 2002.

[23] W. A. Shewhart, *The ecomomic control of the quality of manufactured product*, New York: Macmillan, 1931.

[24] S. W. Roberts, “Control Chart Test based on Geometric Moving Averages,” *Technometrics*, vol. 1, pp. 239-250, 1959.

[25] J. Zhang, Z. Li, and Z. Wang, “Control chart based on likelihood ratio for monitoring linear profiles,” *Computational Statistics & Data Analysis*, vol. 53, no. 4, pp. 1440-1448, 2009.