CHALMERS

Chalmers Publication Library

Grassmann Manifold Online Learning and Partial Occlusion Handling for Visual Object Tracking under Bayesian Formulation

This document has been downloaded from Chalmers Publication Library (CPL). It is the author´s version of a work that was accepted for publication in:

21st International Conference on Pattern Recognition (ICPR 2012) (ISSN: 1051-4651)

Citation for the published paper:
Gu, I. ; Khan, Z. (2012) "Grassmann Manifold Online Learning and Partial Occlusion Handling for Visual Object Tracking under Bayesian Formulation". 21st International Conference on Pattern Recognition (ICPR 2012) pp. 1463-1466.

Downloaded from: http://publications.lib.chalmers.se/publication/160433

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

21st International Conference on Pattern Recognition (ICPR 2012)
November 11-15, 2012. Tsukuba, Japan

Grassmann Manifold Online Learning and Partial Occlusion Handling for Visual Object Tracking under Bayesian Formulation

Irene Y.H. Gu Zulfiqar H. Khan
Dept. of Signals and Systems, Chalmers Univ. of Technology, Gothenburg, 41296, Sweden

Abstract

This paper addresses issues of online learning and occlusion handling in video object tracking. Although manifold tracking is promising, large pose changes and long-term partial occlusions of video objects remain challenging. We propose a novel manifold tracking scheme that tackles such problems, with the following main novelties: (a) Online estimation of object appearances on Grassmann manifolds; (b) Optimal criterion-based occlusion handling during online learning; (c) Nonlinear dynamic model for appearance basis matrix and its velocity; (b) Bayesian formulations separately for the tracking and the online learning process. Two particle filters are employed: one is on the manifold for generating appearance particles and another on the linear space for generating affine box particles. Tracking and online updating are performed in alternative fashion to mitigate the tracking drift. Experiments on videos have shown robust tracking performance especially when objects contain significant pose changes accompanied with long-term partial occlusions. Evaluations and comparisons with two existing methods provide further support to the proposed method.

1 Introduction

Visual tracking has drawn increasing interest in recent years. Many promising results have been obtained by, e.g. trackers using mean shift, local point feature and particle filters [1, 2, 3, 4]. Online learning is essential for the robustness of video object tracking since video objects are dynamic with deformable shape, pose changes and various other changes. Early work of online learning includes, e.g., incremental subspace learning [5], however, tracking drift or failure remains for video objects in complex scenes, e.g. significant pose changes, occlusions and intersections. For planar video objects with significant pose changes, manifold tracking is more suitable since a dynamic object with continuous pose changes is better described by a set of subspaces, or points on a smoothing manifolds. Manifold-based video object tracking has drawn much interest lately. [6] proposes piecewise geodesics on complex Grassmann manifolds using projection matrices for synthetic array sensor signals. [7] proposes visual tracking by using a Kalman filter to velocity vectors in the tangent planes of Grassmann manifold, that only works for objects with small/moderate pose changes. Several covariance tracking methods on Riemannian manifolds have also been proposed [8, 9]. In these methods online learning is designed for learning object changes, where object occlusion scenarios are not considered. Despite reasonably good results from manifold tracking, challenges remain for tracking objects with significant pose changes especially when this is accompanied with long-term object partial occlusions or object intersections. The main reasons could be the lack of robust online learning methods, the lack of online learning with simultaneously occlusion handling, and also the lack of robust dynamic models on manifolds.

Motivated by the above, we propose a novel tracking method that tackles these issues. The proposed tracking scheme is a Grassmann manifold-based Bayesian tracker, where the main novelties are the online appearance learning combining with occlusion handling. Comparing with our previous work on Riemannian manifolds in [9], this paper deals with a different type of manifolds, also an occlusion handling strategy is introduced on top of the online learning. To the best of our knowledge, for Grassmann manifold tracking scenarios, no successful criterion of object online learning with occlusions handling has so far been reported.

2 Grassmann Manifolds: Review

A Grassmann manifold $\mathcal{G}_{n,k}$ is defined as a set of all k-dimensional subspaces in $\mathbb{R}^n$. Let $\mathbf{p} \in \mathcal{G}_{n,k}$ and $\mathbf{U}$ be the $n \times k$ orthonormal bases of $\mathbf{p}$. To form a full bases for $\mathbb{R}^n$, an $n \times n$ orthonormal $\mathbf{Q}$ is defined as $\mathbf{Q} \triangleq [\mathbf{U} \mid \mathbf{U}_\perp]$, where $\mathbf{U}^T\mathbf{U}_\perp = 0$. A Grassmann manifold can be equivalently described by the basis matrix $\mathbf{U}$ in $[\mathbf{U}]$ for achieving computational efficiency [10]. Let $\mathbf{Y} \in \mathbb{R}^{n,k}$ be an observation matrix, $\mathbf{m}_Y = \frac{1}{k}\sum_{i=1}^{k}\mathbf{Y}_i$ be the mean, then $\mathbf{U} \in \mathbb{R}^{n,k}$ is computed by a compact singular value decomposition (SVD) of mean subtracted $\mathbf{Y}$, $\mathbf{UDV} = \text{SVD}(\mathbf{Y} - \mathbf{m}_Y)$. More general descriptions of manifold theories can be found in [10].

Mapping functions: Two important mapping functions are essential for manifold tracking and online learning. One is the exponential map ($\mathcal{T} \to \mathcal{G}_{n,k}$)). Given $\mathbf{p}, \mathbf{q} \in \mathcal{G}_{n,k}$, the exponential mapping function maps a tangent vector $\boldsymbol{\Delta}$ to a manifold point $\mathbf{q}$ at t=1, starting from $\mathbf{p}$ along the geodesic. The exponential mapping function can be described by using basis matrices for computational efficiency [11]:

$$\exp_{\mathbf{p}}(\Delta) = \mathbf{W} = \mathbf{U}\mathbf{V}\cos(\boldsymbol{\Sigma})\mathbf{V}^T + \mathbf{R}\sin(\boldsymbol{\Sigma})\mathbf{V}^T$$

where $\mathbf{U}$ and $\mathbf{W}$ are $n \times k$ basis matrices for $\mathbf{p}$ and $\mathbf{q}$. Another is the logarithmic map ($\mathcal{G}_{n,k} \to \mathcal{T}$). Given $\mathbf{p}, \mathbf{q} \in \mathcal{G}_{n,k}$, the logarithmic mapping function maps $\mathbf{p}$ to $\mathbf{q}$ on $\mathcal{G}_{n,k}$ along the geodesic that results in a tangent vector $\boldsymbol{\Delta}$ in $\mathcal{T}$. Using basis matrices, this can be efficiently computed by $\boldsymbol{\Delta} = \log_{\mathbf{p}}(\mathbf{q}) = \mathbf{S}\sin^{-1}(\boldsymbol{\Sigma})\mathbf{V}^T$, where $\mathbf{R}\boldsymbol{\Sigma}\mathbf{D}^T = \mathbf{W} - \mathbf{U}\mathbf{U}^T\mathbf{W}$, $\mathbf{V}\mathbf{C}\mathbf{D}^T = \mathbf{U}^T\mathbf{W}$ is the generalized SVD, $\mathbf{C}^T\mathbf{C} + \boldsymbol{\Sigma}^T\boldsymbol{\Sigma} = I$, and $\sin^{-1}(\cdot)$ acts element-by-element along the diagonal of $\boldsymbol{\Sigma}$.

3 General Description: Proposed Scheme

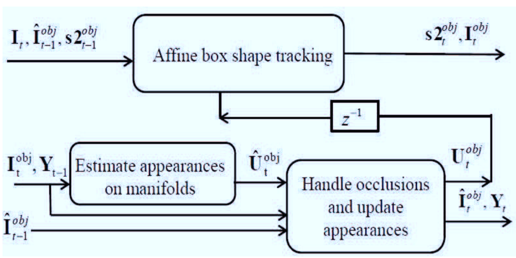

Figure 1. Block diagram of the proposed integrated scheme. $\hat{\mathbf{I}}_{t-1}^{obj}$ is the reference object image at t-1, $\hat{\mathbf{I}}_{t}^{obj}$ and $\mathbf{s2}_{t-1}^{obj}$ are the tracked object image and its box parameters at t-1; $\mathbf{I}_{t-1}^{obj}$ is tracked object image used as the new observation image at t; $\hat{\mathbf{U}}_t$ is the estimated manifold appearance, and $\mathbf{U}_t^{obj}$ is the final updated appearance after occlusion handling; ($\mathbf{Y}_{t-1}$, $\mathbf{Y}_t$) are the observation matrix with a sliding window size L at t-1 and t; $\mathbf{I}_t$ is the current video frame; and $z^{-1}(\mathbf{U}_t^{obj}) = \mathbf{U}_{t-1}^{obj}$ is the reference object appearance at t-1 used for the tracking process at t.

Fig.1 shows the block diagram of the proposed scheme, which can be split into two parts: a Bayesian object tracking process (block-1 on the top), and a process of Bayesian manifold online appearance estimation, and manifold updating with criterion-based occlusion handling (block-2 on bottom left, and block-3 on bottom right), respectively. In the tracking process, object bounding box affine parameters are tracked by a PF (Particle Filter). This is different from the conventional PF tracking such that the embedded visual object appearance is on a Grassmann manifold rather than in a vector space. In the online updating process, the appearance subspace is first estimated on the manifold by another PF. The PF uses a nonlinear dynamic model, the exponential and logarithmic mapping functions between the tangent planes and the manifold. The likelihood in this PF is computed by the subspace angles between the current observation and predicted manifold particles. The online learned object appearance is then obtained as the posteriori manifold point. A criterion is applied to estimate the occlusion. If no occlusion is detected, updating the basis matrix of reference object appearance is then performed. These two parts, tracking and updating, are performed in an alternation fashion as an integrated tracking scheme.

4 Dynamic Model, Bayesian Manifold Appearance and Occlusion Handling

4.1. Nonlinear Dynamic State Space Model

Let the object appearance at t be described by a point on a Grassmann manifold by the basis matrix $\mathbf{U}_t$, and the change of appearance (or speed) be $\boldsymbol{\Delta}_\mathbf{t}$. Define the state vector as $\mathbf{s}_t = [\mathbf{U}_t \quad \boldsymbol{\Delta}_t]^T$. Let the state dynamics be described by the nonlinear dynamic model,

$$\begin{aligned} \mathbf{U}_t &= h(\mathbf{U}_{t-1}, \boldsymbol{\Delta}_t) = \exp_{\mathbf{U}_{t-1}}(\boldsymbol{\Delta}_t) \\ \boldsymbol{\Delta}_t &= \boldsymbol{\Delta}_{t-1} + \mathbf{V}_1 \end{aligned} \tag{1}$$

where $\mathbf{V}_1$ (including the acceleration and model noise) is assumed to be zero-mean white distributed, $\boldsymbol{\Delta}_{t-1}$ is constant in each sample interval $T = t_k - t_{k-1}$, and $T = 1$ for mathematical convenience, $h(\cdot)$ is nonlinear. We refer (1) as the dynamic model that deals variables in two different spaces: The first equation models the dynamic appearances on the manifold where two manifold points at successive time instants are related by $\boldsymbol{\Delta}_t$ in the tangent plane; the second equation is a constant velocity model in the tangent plane whose acceleration is considered as white noise. The above dynamic model can be considered as a 2nd-order discrete white noise acceleration model for Grassmann manifold points.

4.2. Online Bayesian Appearance Estimation

The aim here is to perform online estimation of $\mathbf{U}_t$ given a new object appearance at t. We first assume that no object occlusion occurs (See Section 4.3 for occlusion handling). This is realized by a PF on the manifold. Let the current observation at t be $\mathbf{Z}_t = \tilde{\mathbf{U}}_t^{obj}$ (provided by the tracking process, see Section 5). $\tilde{\mathbf{U}}_t^{obj}$ is the basis matrix of tracked object at t obtained by first stacking the tracked object appearance $\mathbf{I}_t^{obj}$ in a sliding windows of size L (noting, $\mathbf{Y}_t$ below corresponds to non-occlusion cases): $\mathbf{Y}_t = [\mathbf{I}_{t-L+1}^{obj} \cdots \mathbf{I}_t^{obj}]$, and then calculating the basis matrix. Let $\mathbf{U}_t$ be the Bayesian estimate through:

$$p(\mathbf{U}_t|\mathbf{Z}_{0:t}) \propto p(\mathbf{Z}_t|\mathbf{U}_t) \int p(\mathbf{U}_t|\mathbf{U}_{t-1}, \boldsymbol{\Delta}_t) p(\mathbf{U}_{t-1}|\mathbf{Z}_{t-1}) d\mathbf{U}_{t-1}$$

where $\mathbf{U}_t$ is object manifold appearance, $\mathbf{Z}_{0:t}$ is the observations up to t. The posterior pdf estimate is approximated by $p(\mathbf{U}_t|\mathbf{Z}_{0:t}) \approx \sum_{j=1}^{N_1} w_t^j \delta(\mathbf{U}_t - \mathbf{U}_t^j)$, where $\mathbf{U}_t^j$ is the jth particle, w_t^j is the normalized weight, and N_1 is the total number of particles. Since this PF is performed on the manifold where the dynamic model describes state variables in two inter-connected spaces, realization of this PF requires the interaction between the manifold points and their tangent planes. This estimation process is subdivided into the following steps:

Prediction: Let $\mathbf{U}^j_{t-1}$ be a manifold particle point at $t-1$ and $\boldsymbol{\Delta}^j_{t-1}$ be the velocity particle that connects ($\mathbf{U}^j_{t-2}, \mathbf{U}^j_{t-1}$), where $\mathbf{U}^j_{t-1}$ is the end point of the geodesic starting from $\mathbf{U}^j_{t-2}$. First, a set of velocity particles $\boldsymbol{\Delta}^j_t$ (originated from $\mathbf{U}^j_{t-1}$) is generated in tangent planes using the previous velocity particles $\boldsymbol{\Delta}^j_{t-1}$ according to $\boldsymbol{\Delta}^j_t = \boldsymbol{\Delta}^j_{t-1} + \mathbf{V}_1$ (using (1)), $j = 1, \cdots, N_1$. Then, a set of new manifold particles $\mathbf{U}^j_t$ is obtained from $\boldsymbol{\Delta}^j_t$ through the exponential mapping, according to $\mathbf{U}^j_t = exp_{\mathbf{U}^j_{t-1}}(\boldsymbol{\Delta}^j_t)$ (using (1)). $\mathbf{U}^j_t$ are predicted manifold points at t.

Appearance Likelihood and Particle Weights: The likelihood is computed from Gaussian distributed principal angles between the observation bases $\tilde{\mathbf{U}}^{obj}_t$ (computed from $\mathbf{Y}_t$) and predicted manifold point $\mathbf{U}^j_t$: $p(\tilde{\mathbf{U}}^{obj}_t|\mathbf{U}^j_t) = \exp\left\{-\frac{d(\tilde{\mathbf{U}}^{obj}_t,\mathbf{U}^j_t)}{\sigma^2_l}\right\}$, where σ^2_l is the measurement noise (σ^2_l=0.1 in our tests), and $d(\tilde{\mathbf{U}}^{obj}_t, \mathbf{U}^j_t)$ is defined according to principal angle [10]. The weight is then updated by $w^j_t \propto w^j_{t-1} p(\tilde{\mathbf{U}}^{obj}_t|\mathbf{U}^j_t)$ and subsequently normalized. Resampling is applied if $\hat{N}_{eff} = 1/\sum_{j=1}^{N_1}(w^j_t)^2 < N1_{th}$, to prevent the degeneracy [4].

Posterior Estimation of Manifold Point: MMSE estimate of $\hat{\mathbf{U}}^{obj}_t$ is obtained as the expected value of weighted predicted particles on the manifold by:

$$\hat{\mathbf{U}}^{obj}_t = \exp_{\tilde{\mathbf{U}}^{obj}_t}\left(\frac{1}{N_1}\sum_{j=1}^{N_1}\mathbf{w}^j_t \log_{\tilde{\mathbf{U}}^{obj}_t}\mathbf{U}^j_t\right) \quad (2)$$

4.3. Partial Occlusion Handling

It is important that an online updating method *only* updates the reference object appearance for changes caused by the dynamic object itself (e.g. pose, deformation). If online updating is applied when manifold appearance changes is due to partial occlusions, it could lead to tracking drift. In video object tracking, there is an ambiguity between changes due to object and due to occlusions. Our aim here is to introduce a criterion that gives a rough estimation on whether changes are due to object dynamics, or occluding objects/background. If latter case appear, the updating process would be frozen to prevent absorbing wrong information to the object.

The occlusion handling strategy is based on the observation that relatively large differences may occur when an object experiences occlusions as compared with object pose changes. We use the Bhattacharyya coefficient between a tracked and the reference object as the distance measure between two subspaces, i.e.: $\rho_t = \sum_u \sqrt{p^t_u\, q^{t-1}_u}$, where p^t_u and q^{t-1}_u are uth histogram bin of spatial kernel-weighted intensity histograms for a tracked and the reference object region at t. The criterion for estimating occlusion is done by comparing ρ_t with an empirical threshold ρ_{th}:

$$\rho_t < \rho_{th} \quad (3)$$

If (3) is satisfied, then the object is considered as occluded and no updating is performed. The rationale behind the choice is that the subspace change introduced by occluding object/background is usually larger than that from pose/appearance change from an object itself. Since a large ρ_t indicates that two subspaces are closer, a small ρ_t value is as an indication of object experiencing occlusions and deviating from its own appearance. If (3) is satisfied, the tracked object region $\mathbf{I}^{obj}_t$ would not be added to the sliding window observation matrix $\mathbf{Y}_t$, and no updating is performed for (4). Otherwise, changes are considered as caused by object itself, and hence $\mathbf{I}^{obj}_t$ is added to $\mathbf{Y}_t$, and the basis matrix from the observation matrix $\mathbf{Y}_t$ as well as the reference image object $\hat{I}^{obj}_t$ are then updated by:

$$\begin{aligned} &\text{update } \tilde{\mathbf{U}}^{obj}_t: \quad \tilde{\mathbf{U}}^{obj}_t\mathbf{D}^{obj}_t\mathbf{V}^{obj}_t = \text{SVD}(\mathbf{Y}_t - \mathbf{m}_Y) \\ &\text{update } \hat{I}^{obj}_t: \quad \hat{I}^{obj}_t = \kappa\mathbf{m}_Y + (1-\kappa)\hat{I}^{obj}_{t-1} \\ &\text{update } \mathbf{U}^{obj}_t: \quad \mathbf{U}^{obj}_t = \hat{\mathbf{U}}^{obj}_t \end{aligned} \quad (4)$$

where κ is a constant controlling the learning rate and $\mathbf{m}_Y$ is the mean of $\mathbf{Y}_t$.

5 Bayesian Object Tracking

The aim in this part is to estimate the posterior pdf of affine object bounding box, while taking into account of the manifold object appearance within the box. This is realized by utilizing another PF, PF-2, where the manifold object appearance is embedded. Let the state vector $\mathbf{s}2_t = [y^1_t\ y^2_t\ \beta_t\ \gamma_t\ \alpha_t\ \phi_t]^T$ at t be the affine bounding box parameters (2D box center, scale, rotation, aspect ratio and skew). Given a set of particles at $t-1$, new particles $\{\mathbf{s}2^i_t\}_{i=1}^{N_2}$ are generated by PF-2 according to the state equation $\mathbf{s}2_t = \mathbf{s}2_{t-1} + \mathbf{v}2$, $\mathbf{v}2 \sim \mathcal{N}(0, \boldsymbol{\Omega})$. The likelihood is moedelled as Gaussian distributed dynamic prediction error *on the manifold,*

$$p(\mathbf{z}2_t|\mathbf{s}2^i_t) = \exp\left\{-\frac{||\mathbf{d}_{I^i_t} - \mathbf{U}^{obj}_{t-1}\mathbf{U}^{obj\,T}_{t-1}\mathbf{d}_{I^i_t}||}{\sigma^2_{w2}}\right\}, \quad \text{where}$$

$\mathbf{I}(\mathbf{s}2^i_t) = \mathbf{I}^i_t$ describes the candidate object appearance within the bounding box, $\mathbf{d}_{\mathbf{I}^i_t} = \mathbf{I}^i_t - \hat{\mathbf{I}}^{obj}_{t-1}$, $\mathbf{U}^{obj}_{t-1}$ is the manifold bases of the reference object at $(t-1)$, $\mathbf{z}2_t$ is the current observation (image frame), and σ^2_{w2} is the variance of measurement noise. The particle weights $w2^i_t$ are updated by $w2^i_t \propto w2^i_{t-1}p(\mathbf{z}2_t|\mathbf{s}2^i_t)$ followed by the normalization. Further, resampling is applied if $\hat{N}_{eff} = 1/\sum_{i=1}^{N_2}(w2^i_t)^2 < N2_{th}$ [4]. Finally, the MAP estimate of bounding box $\mathbf{s}2_t$ is computed.

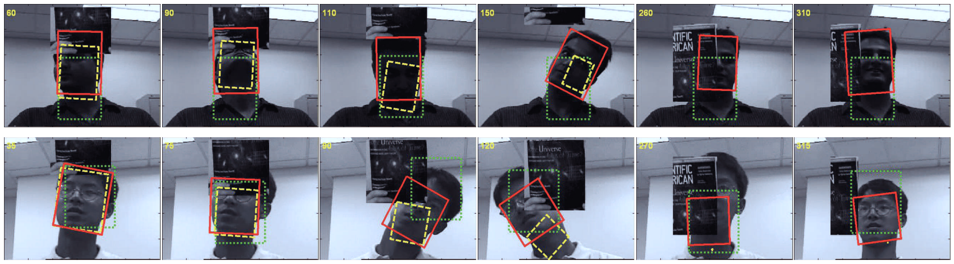

Figure 2. Tracking results on "Danni" and "Behzad" with added occlusions (by superposition of a real occluding (book) image on face images). Red box: from the proposed tracker; Green: from *Tracker-1*; Yellow: from *Tracker-2*.

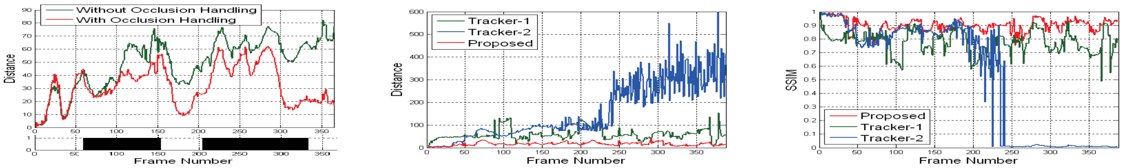

Figure 3. Performance evaluation and comparisons. Left: Proposed method with/without occlusion handling strategy: Euclidean distance on video "Danni + occlusion". The black bar indicates the frames with occlusion; Middle (Euclidian distances between bounding box corners of tracked and ground truth object); Right (SSIM between images of tracked and ground truth object (the larger value the better): Proposed tracker and *tracker-1, tracker-2* on video "Chai' without occlusions.

6 Experiments and Results

To test and evaluate the proposed scheme, several videos containing deformable objects with significant object pose changes, captured by a moving/static camera, are used. For all videos, initial object bounding boxes are manually selected. Each box is further normalized to 32×32 pixels. Parameters used for PF-1 are $N_1 = 400$, $\sigma^2_{v_1} = 0.01$, $\sigma^2_l = 0.1$, $\kappa = 0.1$, $N1_{th}$=50; for PF-2, $N_2 = 600$, $\sigma^2_{w2} = 0.25$, and $N2_{th}$=75. Two existing trackers: Tracker-1 (covariance-based tracking in [8]), and Tracker-2 (subspace tracking on Grassmann manifold [7]) are used for comparisons.

Tests are performed on several videos that contain objects with *both large pose changes and partial occlusions*. Fig.2 shows the tracking results on several videos with partial occlusions. Fig.3 (left plot) shows the Euclidean distances (of 4 corners of tracked box and ground truth box) with and without occlusion handling. Observing Fig.2 and Fig.3, the proposed tracker has clearly shown a better performance in these tests.

Tests and comparisons are also performed on videos where objects contain large pose changes but without occlusions. Fig.3(Middle and Right plots) shows the Euclidian distances and SSIM (Structural Similarity) measure as a function of video frames, from the proposed tracker and Tracker-1,2. Comparing the results, the proposed scheme has shown clear improvement.

7 Conclusion

Tests on the proposed tracking scheme, consisting of visual tracking on the manifold and online manifold basis updating, has shown very robust tracking performance for objects containing moderate to large pose changes. The online updating of basis matrices by exploiting the nonlinear dynamic model and two state variables enables effective posterior estimates of Grassmann manifold points. A method to detect partial occlusion is shown to be effective. The online tracking by integrating dynamic appearance and shape on the manifold and its tangent plane in single particle filter is efficient. Comparisons with two existing and most relevant manifold tracking methods have provided further support to the robustness of the proposed scheme.

References

[1] Z.H.Khan; I.Y.H. Gu, "Joint Feature Correspondences and Appearance Similarity for Robust Visual Object Tracking", IEEE Trans IFS, vol. 3 pp.591-606, 2010.

[2] S.Haner, I.Y.H. Gu, "Combining foreground/background feature points and anisotropic mean shift for enhanced visual object tracking", ICPR 2010, pp.3488-3491, 2010.

[3] Z.H.Khan, I.Y.H. Gu, A.Backhouse, "Robust visual object tracking using multi-mode anisotropic mean shift and particle filters", IEEE Trans CSVT, vol.1, pp.74-87, 2011.

[4] A.Dore, M.Soto, C.Regazzoni, "Bayesian tracking for video analytics", IEEE Trans IP, 27(5), 46-55, 2010.

[5] D.Ross, J.Lim, et al, " Incremental learning for robust visual tracking", *Int.J. Comput. Vis.* , 77(1), pp.125-141, 2008.

[6] A.Srivastava,et al, "Bayesian and geometric subspace tracking", *Adv. Appl. Prob, (SGSA)*, vol.36, pp.43-56, 2004.

[7] T.Wang, A.Backhouse, I.Y.H. Gu, "Online subspace learning on Grassmann manifold for moving object tracking in video", ICASSP, 2008.

[8] F.Porikli, O.Tuzel, P.Meer, "Covariance tracking using model update based on Lie algebra", IEEE CVPR, 2006.

[9] Z.H.Khan, I.Y.H.Gu, "Bayesian online learning on the Riemannian manifold using a dual model with applications to video object tracking", 1st IEEE Workshop ITCVPR,2011.

[10] A.Edelman et al, "The geometry of algorithms with orthogonality constraints",*SIAM J.Matrix Anal. Appl*,20(2),1998.

[11] R.Subbarao, P.Meer, "Nonlinear mean shift over Riemannian manifolds", *Int. J. Comput. Vis.*, 84(1), pp.1-20, 2009.

[12] A.Tyagi, J.W.Davis, "A recursive filter for linear systems on Riemannian manifolds", IEEE CVPR, pp.1-8, 2008.