Research Article

Multiple Human Tracking Using Particle Filter with Gaussian Process Dynamical Model

Jing Wang, Yafeng Yin, and Hong Man

Department of Electrical and Computer Engineering, School of Engineering and Science, Stevens Institute of Technology, Hoboken, NJ 07030, USA

Correspondence should be addressed to Jing Wang, jwang@stevens.edu

We present a particle filter-based multitarget tracking method incorporating Gaussian process dynamical model (GPDM) to improve robustness in multitarget tracking. With the particle filter Gaussian process dynamical model (PFGPDM), a high-dimensional target trajectory dataset of the observation space is projected to a low-dimensional latent space in a nonlinear probabilistic manner, which will then be used to classify object trajectories, predict the next motion state, and provide Gaussian process dynamical samples for the particle filter. In addition, Histogram-Bhattacharyya, GMM Kullback-Leibler, and the rotation invariant appearance models are employed, respectively, and compared in the particle filter as complimentary features to coordinate data used in GPDM. The simulation results demonstrate that the approach can track more than four targets with reasonable runtime overhead and performance. In addition, it can successfully deal with occasional missing frames and temporary occlusion.

1. INTRODUCTION

Multitarget tracking is an important issue in security applications, which have attracted considerable attention and interest in recent years. Some classical approaches to multitarget tracking include the multiple hypothesis tracker (MHT) and the joint probabilistic data association filter (JPDAF) [1]. Particle filters have been recently used for multitarget tracking tasks, because they can deal with indeterministic motions, as well as nonlinear and non-Gaussian systems. However, joint particle filters can normally track up to three or four identical objects due to the exponential complexity [2]. A possible solution to this problem is to integrate the Gaussian process dynamical prediction function with the learning mechanism, which provides a particle filter with prior information to reduce sampling ambiguity, and improve particle efficiency. Furthermore, the high-dimensional learning datasets may increase classification and computation complexities. This can be alleviated by dimension reduction through nonlinear mapping, and incorporating Markov dynamics in the low-dimensional latent space for data prediction. Our major contribution in this work is a novel multitarget tracking algorithm that

incorporates particle filters with Gaussian process dynamical model to improve tracking accuracy and computational efficiency. Initial tests indicate that target objects (e.g., people) in a specific environment may have similar trajectory patterns, which make potentially efficient tracking algorithm possible. We use trajectory classification instead of pose and motion classification in motion tracking, therefore state sharing can be achieved on the latent space to take advantage of similar object trajectory properties. In addition, our research focuses on efficient multitarget trajectory tracking as well as handling missing frames and temporary occlusions to produce reliable tracking results for high-level analysis.

This article is organized as follows. Section 2 reviews previous work on tracking by using particle filter and Gaussian process dynamical model. The proposed particle filter Gaussian process dynamical model is described in Section 3. In Section 4, the experimental results are presented, and the article is summarized in Section 5.

2. PREVIOUS WORK

The previous work is summarized as follows. Khan et al. proposed a template-based particle filter system to track

interacting ants [1]. Compared with people, ants are more rotation invariant and have less contour changes, hence the learning system should be different. Okuma et al. studied multiple hockey players detecting and tracking by deploying particle filter incorporating Ada boost detection proposal generation algorithm [2]. The Kernel particle filter was developed to track multiple targets in image sequences by Chang et al. [3]. Zhou et al. proposed a particle filter-based tracking system with an appearance-adaptive model [4].

A trans-dimensional Markov Chain Monte Carlo (MCMC) particle filter was proposed for reliable tracking of indefinite number of interacting targets [5]. Multiple objects are formulated by a joint state-space model while efficient sampling is performed by deploying trans-dimensional MCMC on the subspace. It failed to track some targets due to the weakness of color models. Reference [6] employed particle filter to handle partial occlusion as a component of a proposed Hybrid Joint-Separable (HJS) filter framework in multibody tracking. A mean field Monte Carlo (MFMC), that is, particle filter modeled as a competition problem was used to address coalescence issue occurred in multitarget tracking [7]. Reference [8] employed particle filter incorporating with a multiblob likelihood function to track unknown and varied objects while assuming background modeling is effective given a static camera. A color particle filer embedded with a detection algorithm was proposed by [9] to track multiple targets deploying the same color description with internal initialization and cancelation functionality.

Gaussian process latent variable model described by Neil Lawrence handles probabilistic nonlinear dimensionality reduction problems to model the high-dimensional observation data and the corresponding projections onto the low-dimensional latent space [10]. Wang et al. incorporate Markov dynamics on latent variable state transitions lending Gaussian process latent variable model to handle time series data and robustly track human body motion and pose changes by classifying poses and motions [11]. Reference [12] used GPDM to track 3D human pose and motion. Raskin et al. proposed a Gaussian process annealing particle filter-based method to perform 3D target tracking by exploring color histogram features [13]. Our research is different in that multitarget trajectory tracking was performed, whilst the annealing particle filter GPDM framework proposed by Raskin et al. tracked 3D pose and motion of one target. In addition, our particle generation mechanism and classified elements are different.

Reference [14] described a framework combining the particle filter, GPDM, and discriminative learning approaches to avoid 3D human model in tracking 3D human motion. Image latent space mapping on joint angle latent space is performed by employing relevance vector machine (RVM) on small training sets. Reference [15] proposed a shared latent dynamical model derived from GPLVM and GPDM to diminish the dimensionality of the pose state space, hence facilitate the manipulation of tracking data. The latent space can be projected to both state space and observation space by learning approach with dynamic mechanism. SLDM is integrated with condensation framework to estimate positions in the latent space and reconstruct human poses. Reference [16] presented a full-3D edge tracker by using particle filter to track complex 3D objects of flexible motion and under self-occlusions with hidden line removal capability. Real-time rate is obtained by employing accelerating hardware implementation on hidden line removal and likelihood calculation.

3. PARTICLE FILTER GAUSSIAN DYNAMICAL PROCESS

3.1. Particle filter and GPDM

A particle filter is a Monte Carlo method for nonlinear, non-Gaussian models, which approximates continuous probability density function by using large number of samples. Hence, the accuracy of the approximation depends on high-dimensional state space which causes exponential increase of the number of particles. Given the time complexity constraint, the reduction of particles, and hence the computation power is a potential solution.

In GPDM, an observation space vector represents a pose configuration and motion trajectory captured by a sequence of poses. At the beginning of the learning procedure, the target data from observation is projected to a subspace-latent space by principal components analysis (PCA). During this projection, with an assumption of Gaussian prior distribution over the latent space, the projection will become a nonlinearization through Gaussian process, so it can be viewed as probability PCA (PPCA) [17]. Then, scaled conjugated gradient (SCG) is applied to optimize and smooth the initialized coordinates. Once a GPDM is created, sampling from the dynamical field provides meaningful prediction on the future motion changes. The latent space defines the temporal dependence between poses by employing Gaussian process integrated by Markov chain on the latent variable transitions. Since motion prediction, the temporal dependence, and sampling are performed on the latent space, potential computation benefits may be obtained.

3.2. Particle filter Gaussian process dynamical model

This research aims at developing a low-complexity and highly-efficient algorithm for tracking variable number of targets with competitive tracking performance in term of accuracy. With the general framework of GPDM, it can be extended to estimate pose and motion changes as proposed by Wang et al. Hence, if a target is suspected of malicious behavior, the system can trade performance off time complexity.

The basic procedure of the proposed particle filter Gaussian process is as follows.

(1) *Creating GPDM*. GPDM is created on the basis of the trajectory training data sets, that is, coordinate difference values, and the learning model parameters

$\Gamma = \{Y^T, X^T, \overline{\alpha}, \overline{\beta}, W\}$, where Y^T is the training observation dataset, X^T is the corresponding latent variable sets, $\overline{\alpha}$ and $\overline{\beta}$ are hyperparameters, and W is a scale parameter.

(2) *Initializing the model parameters and the particle filter.* The latent variable set of the training data and parameters $\{X^T, \overline{\alpha}, \overline{\beta}\}$ are obtained by minimizing the negative log posterior function $-\ln p(X^T, \overline{\alpha}, \overline{\beta}, W \mid Y^T)$ of the unknown parameters $\{X^T, \overline{\alpha}, \overline{\beta}, W\}$ with scaled conjugate gradient (SCG) on the training datasets.

The prior probability is derived on the basis of the created model. In this step, target templates are obtained from the previous frames as reference images for similarity calculation in the later stage.

(3) *Projecting from the observation space to latent space.* The test observation data is projected on the latent coordinate system by using probabilistic principal component analysis (PPCA). As a result, the dimensionality of the observed data is reduced.

(4) *Predicting and sampling.* Particles are generated by using GPDM in the latent space and the test data to infer the likely coordinate change value $(\Delta x_i, \Delta y_i)$.

(5) *Determining probabilistic mapping from latent space to observation space.* The log posterior probability of the coordinate difference values of the test data is maximized to find the best mapping in the training datasets of the observation space. In addition, the most likely coordinate change value $(\Delta x_i, \Delta y_i)$ is used for predicting the next motion.

(6) *Updating the weights.* In the next frame, the similarity between the template's corresponding appearance model and the cropped region centered on the particle is calculated to determine the weights w_i, and the most likely location $(\hat{x}_{t+1}, \hat{y}_{t+1})$ of the corresponding target, as well as to decide whether resampling is necessary or not.

(7) Repeat Steps 3–6.

3.2.1. Observation space

The targets of interest are detected and tracked for trajectory analysis. Instead of studying the coordinate values, the differences of the same target in two neighboring frames are calculated as the observed data. The location of the target can be obtained by adding the difference to the previous coordinate values. The 2D coordinate difference values of the head, centroid, and feet form a 6-dimensional vector for each object, given by $Y_k = (\Delta(x_1), \Delta(y_1), \Delta(x_2), \Delta(y_2), \Delta(x_3), \Delta(y_3))$, where Y_k is the observation value of the kth target, and $(x_k + \Delta(x_k), y_k + \Delta(y_k))$ is the coordinate value of the corresponding body part. With the 3 sets of coordinate values, the boundary, width, and height of an object can be determined. If there are 5 targets, the observation data has 30 dimensions.

3.2.2. Establishing trajectory learning model and obtaining appearance templates

GPDM is deployed to learn the trajectories of moving objects. The probability density function of latent variable X and the observation variable Y are defined by the following equations:

$$P(X_k \mid \alpha) = \frac{p(x_k)}{\sqrt{(2\pi)^{(N-1)d} \mid K_X \mid^d}} \exp\left(-\frac{1}{2}\mathrm{tr}(K_X^{-1} X_{2:N} X_{2:N}^T)\right),$$

(1)

where α is the hyperparameter of kernel, $p(x)$ can be assumed to have Gaussian prior, N is the length of latent vector, d is the dimension of latent space, and K_X is the kernel matrix

$$P(Y_k \mid X_k) = \frac{\mid W \mid^N}{\sqrt{\left(2\pi^{ND} \mid K_Y \mid^D\right)}} \exp\left(-\frac{1}{2}\mathrm{tr}(K_Y^{-1} Y W^2 Y^T)\right),$$

(2)

where k is the kth target, K_Y is the kernel function, and W is the hyperparameter.

In our study, RBF kernel given by the following is employed for GPDM model

$$k_Y(x, x') = \exp\left(-\frac{\gamma}{2}\|x - x'\|^2\right) + \beta^{-1}\delta_{X,X'},$$ (3)

where x and x' are any latent variables in the latent space, γ controls the width of the kernel, and β^{-1} is the variance of the noise.

Given a specific surveillance environment, certain patterns may be observed and worth exploring for future inferences. To initialize the latent coordinate, the d (dimensionality of the latent space) principal directions of the latent coordinates is determined by deploying probabilistic principal component analysis on the mean subtracted training dataset Y^T, that is, $Y^T - \overline{Y^T}$. Given Y^T, the learning parameters are estimated by minimizing the negative log posterior using scaled conjugate gradient (SCG) [18]. SCG was proposed to optimize the multiple parameters of large training sets by deploying Levenberg-Marquardt approach to avoid line search per learning iteration, which increases calculation complexity.

Besides position training datasets, the appearance database is created by obtaining the template images of human head, feet, and torso from the initial frames.

3.2.3. Latent space projecting, predicting and particle sampling

Since GPDM was constructed in the latent space, at the beginning of the test process, the target observation data has to be projected to the same 2-dimensional latent space in order to be compared to the trained GPDM. This projection is achieved by using probabilistic principal component analysis (PPCA), same as the first stage in GPDM learning. The feature vector of each frame contains three

pairs of coordinate change values for every target being tracked in that frame. For n targets, the feature vector will contain $3 \times n$ pairs of coordinate change values. The PPCA projection will reduce this $3 \times n \times 2$ dimensional feature vector to a 1×2 latent space vector to be used in particle filtering. The purpose of projecting the test data from the observation space to the latent space is to initialize the testing data in the latent space and obtain a compact representation of the similar motion patterns in the training dataset. With PPCA and trained GPDM, we can learn certain common motion patterns (e.g., velocities, directions, etc.) from multiple training targets, and then use the learned latent space motion behavior to predict multiple targets' future trajectories using particle filter with much improved efficiency. This is based on the presumption that many human trajectories possess similar properties in common video surveillance applications. It should be noted that the number of targets being tracked does not need to be identical to that in the training data. This is possible because that PPCA aggregates (or projects) multiple training objects as well as test objects onto the same low-dimensional space, and therefore the number of objects does not pose a constraint on the tracking process. If we can obtain the templates and the corresponding initial coordinates of n objects at the beginning of the test phase, the proposed framework can track these n targets regardless the number of training targets.

Particles are generated on the basis of the Gaussian process dynamical model in the latent space, taking the motion model property and unpredictable motion into consideration. The next possible position is predicted by determining the most similar trajectory pattern in the training database and using the corresponding position change value plus noise. The number of particles are reduced from over *one hundred* to about *twenty* by deriving the posterior distribution over functions, instead of parameters, and taking advantage of the learned knowledge. The simulation indicates that the decreasing number of particles does not compromise the tracking results, even in temporary occlusion cases (see Section 4). An example of the learned GPDM space is shown in Figure 1. Each point on this 2D latent space is a projection of a feature vector representing two training targets, that is, 6 pairs of coordinate change values. A total of 72 points in the figure correspond to feature vectors of these two targets over 73-image frames. The grayscale intensity represents the precision of mapping from the observation space to the latent space, and the lighter the pixel appears the higher the precision of mapping is.

3.2.4. Mapping from latent space to observation space

Thereafter, the latent variables are mapped in a probabilistic way to the location difference data in the observation space, defining the active region (i.e., distribution) of an observed target. However, the exact predicted coordinate values of the motion trajectory in the observation space need be calculated so that the importance weight for each particle in the observation space can be updated. Estimation maximization
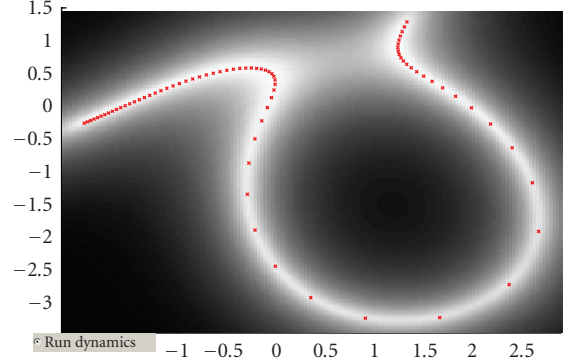

FIGURE 1: Latent space projections of a 2-target training vector sequence.
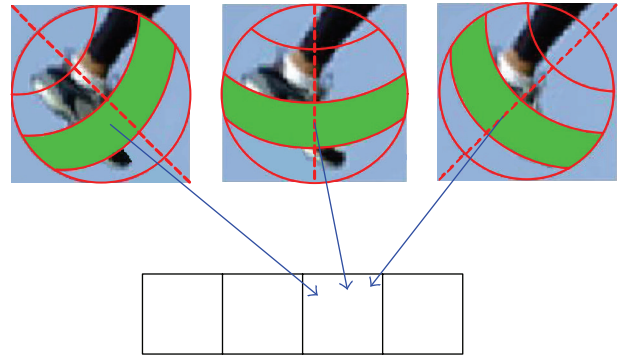

FIGURE 2: Construction of a rotation invariant appearance model for feet representation.

(EM) approach is employed to determine the most likely observation coordinates in the observation space after the distribution is derived.

The nondecreasing log posterior probability of the test data is given by

$$\log\left(P\left(Y_k \mid X^T, \bar{\beta}, W\right)\right)$$
$$= \log\left(\frac{|W|^N}{\sqrt{\left(2\pi^{ND}|K_Y|^D\right)}} \exp\left(-\frac{1}{2}\mathrm{tr}(K_Y^{-1}YW^2Y^T)\right)\right),$$
(4)

where W is the hyperparameter, N is the number of Y sequences, D is the data dimension of Y, and K_Y is a kernel matrix defined by a RBF kernel function given by (3). The log posterior probability is maximized to search for the most probable correspondence on the training datasets. The corresponding trajectory pattern is then selected for predicting the following motion. The simulation results show that it returns better prediction results than averaging the previous motion values. In addition, various targets can share the same database to deal with different future situations.
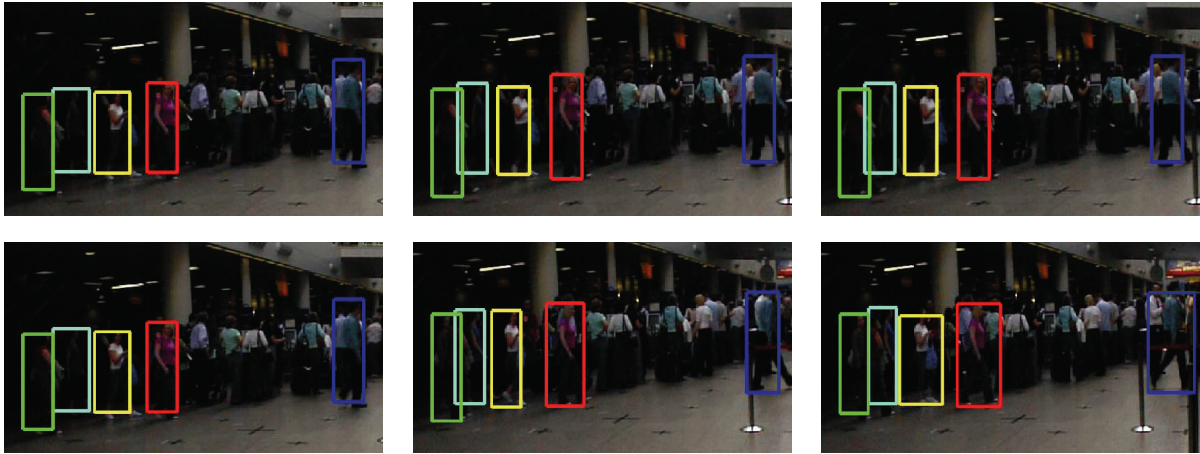
FIGURE 3: Sample results of tracking 5 targets using Histogram-Bhattacharya approach.

FIGURE 4: Sample results of tracking 5 targets using GMM-KL appearance model.
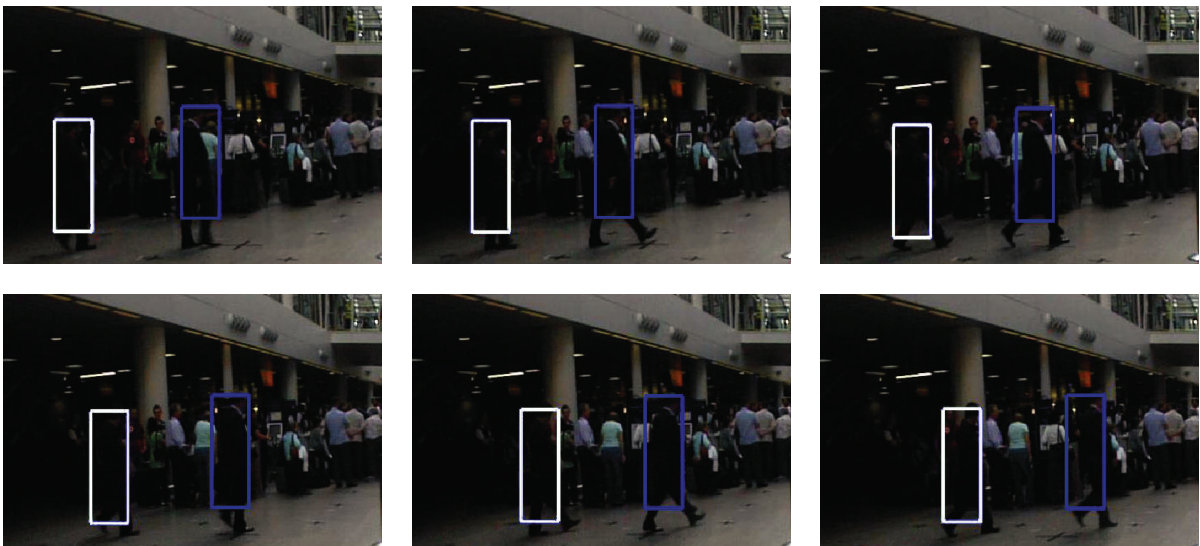

FIGURE 5: Sample results of tracking 2 Targets using rotation invariant appearance model.
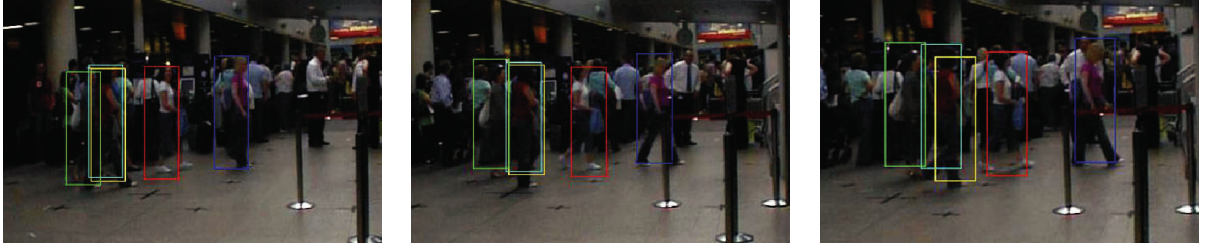
FIGURE 6: Sample results of tracking targets with temporary occlusion.

TABLE 1: Tracking performance of PFGPDM with three appearance models.

No. of frames	No. of targets	Error rate	Runtime
Histogram-Bharttacharyya			
30	2	3%	
30	5	6.7%	
40	5	0	120 sec
GMM-KL			
30	2	0%	
30	5	6.7%	
40	2	2.5%	
40	5	0	209 sec
Rotation invariant			
30	2	0%	
40	2	0%	
40	5	5%	196 sec

TABLE 2: Comparison of three methods on number of particles and error rates.

Algorithm	No. of targets	No. of particles	Error rate
AAMPF [4]	1	87~176	0
PFGPDM	1	20	0
TDMCPF [5]	4	300	0
PFGPDM	4	100	0

3.2.5. Importance weights update

The weights of the particles are updated in terms of the likelihood estimation based on the appearance model. The importance weight equation is given by

$$P(\hat{Y}_t \mid Z_t, k_t) = \frac{P(Z_t \mid k_t, \hat{Y}_t)P(k_t, \hat{Y}_t)}{P(Z_t)},$$

$$w_t \propto P(Z_t \mid k_t, \hat{Y}_t)P(k_t, \hat{Y}_t),$$

(5)

where $\hat{Y}_t$ is the estimation data, Z_t is the observation data, k_t is the identity of the target, and w_t is the weight of a particle. In our study, the likelihood function $P(Z_t \mid k_t, \hat{Y}_t)$ is defined to be dependent on the similarity between the appearance model distribution of the template and that of the test object. Therefore, the choice of appearance model is

important for updating the weights of particles. Edge feature is not used in this study due to its ambiguity in term of foreground and background, as well as the computation efficiency consideration. Histogram-Bhattacharya, GMM-KL appearance model, and rotation invariant model were tested to determine the resulting performance and time complexity.

3.2.6. Histogram-Bhattacharya and GMM-KL appearance model

Histogram-Bhattacharya was used for its simplicity and efficiency [19]. The RGB histogram of the template and the image region under consideration are obtained, respectively. The likelihood $P(Z_t \mid k_t, \hat{Y}_t)$ is defined to be proportional to the similarity between the histogram of the template and the candidate, that is, the region centered on the considered particle of the same size as the template. The above-mentioned similarity is measured by using Bhattacharya distance, since it provides complex nonlinear correlations between distributions.

GMM-KL frame is employed to measure the similarity between the image of template and the test object. GMM is a semiparametric multimodal density model consisting of a number of components to compactly represent pixels of image block in color space with illumination changes. Image can be represented as a set of homogeneous regions modeled by a mixture of Gaussian distributions in color feature space [20]. In comparison, Histogram-Bhattacharya framework presents an image without taking spatial factor into computation. The Kullback-Leibler distance is a measure of the distance between two-probability distributions given the metric of relative entropy [21]. Since the image approximated by Gaussian mixture model can be considered as independently identically distributed (iid) samples following Gaussian mixture distribution, comparison of the template image to that of the test image is formulated as measuring the distance between the two Gaussian mixture distributions. Symmetric version and nonsymmetric version are given by the following:

$$D(p_1, p_2) \cong \frac{1}{n} \sum_{t=1}^{n_1} \log \frac{p_1(x_{1t})}{p_2 x_{1t}} + \frac{1}{n_2} \sum_{t=1}^{n_2} \log \frac{p_2(x_{2t})}{p_1(x_{2t})},$$

$$D(p_1, p_2) \cong \frac{1}{n} \sum_{t=1}^{n} \log \frac{p_1(x_t)}{p_2(x_t)},$$

(6)

where p_1 and p_2 are Gaussian mixture distributions.

FIGURE 7: Sample results of tracking targets with 2 missing frames.

FIGURE 8: Sample results of tracking 1 target to be compared with [5].

The likelihood $P(Z_t \mid k_t, \hat{Y}_t)$ is defined to be proportional to the Kullback-Leibler distance between the associated Gaussian mixture distribution of the template and that of the test region. RGB intensity value is selected as the feature of the appearance model, since it provides reasonable computation complexity and tracking performance, given the efficiency and robustness requirements of the proposed tracking system.

3.2.7. Rotation invariant appearance model

In this work, feet are represented by rotation invariant appearance model, whilst heads are defined by Gaussian mixture model. Since movements of feet normally involves frequent angle changes, rotation invariant approach may render more robust and adaptive appearance model. In addition, the incorporation of spatial color information enables the model to be more discriminative.

In [22], the appearance model represented by multiple polar counterparts is claimed to be invariant to rotation and translation. The original algorithm was tailored to fit our computation-essential framework. First, a detected blob is fully surrounded by a reference circle. Along each of the three directions as shown in Figure 2, 4-control points are sampled uniformly within the reference circle. This forms a group of

4-concentric circles along the corresponding radii. Then the regions with the same control point in the three copies of the blob (shown as the shaded regions) are grouped into one of the 4 bins at the bottom of Figure 2, where all pixels in the corresponding bin are represented by a Gaussian color model with a mean μ and a variance σ. The similarity function given by the following is measured to determine the weights of particles

$$\Gamma = \frac{1}{2N} \sum_N \left\{ (\mu_B - \mu_A)^2 \left(\frac{1}{\sigma_A^2} + \frac{1}{\sigma_B^2} \right) + \frac{\sigma_B^2}{\sigma_A^2} + \frac{\sigma_A^2}{\sigma_B^2} \right\}, \quad (7)$$

where μ and σ are the mean and variance of the color feature given the current bin, and N is the total number of bins defined.

For head region, GMM-KL appearance model is sufficient for static and moving states. Theoretically, particles close to the true centroid in template image have similar probability distributions, and therefore deserve higher weights in the hope of performing more accurate prediction for the future frames. A threshold value is determined to select the particles accurately approximate the posterior probability of the target. When a particle has the weight below the threshold value, resampling is performed to adapt to motion changes.
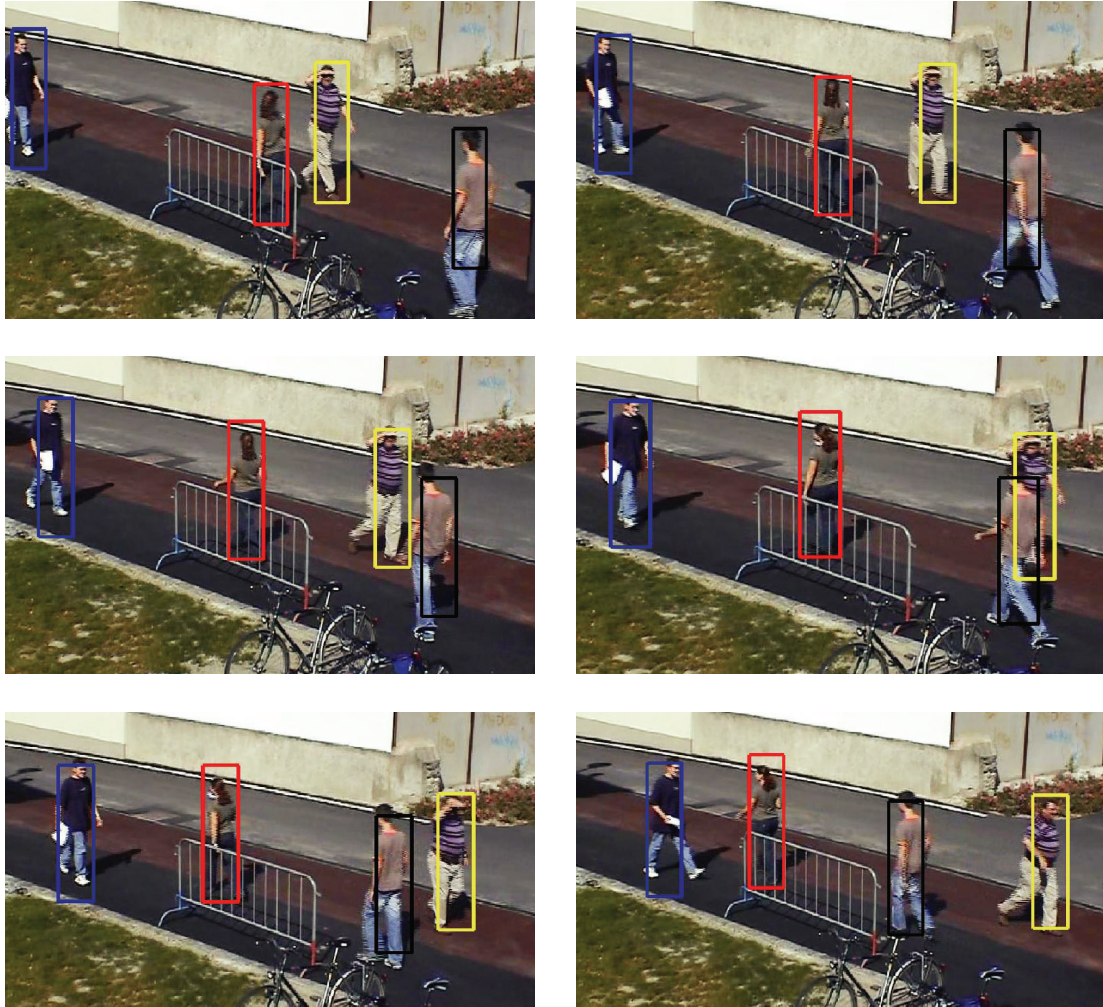
FIGURE 9: Sample results of tracking 4 targets on the IDIAP dataset [5].

4.   SIMULATION RESULTS AND DISCUSSION

The proposed PFGPDM was implemented by using MAT-LAB running on a desktop of 2.53 GHz Pentium 4 PC, with 1 GB memory and tested on the PETS 2007 datasets [23] and the IDIAP datasets used in [5]. Neil Lawrence's Gaussian process softwares provide the related GPDM functions for conducting simulations [24].

The experiments were designed to evaluate the performance of the proposed PFGPDM method under regular test conditions, as well as on sequences with occasional missing frames. The performance measures include sample image frames labeled with tracking results, error rate, runtime, and number of particles used. Error rate is defined as the percentage of frames that contain one or more miss-tracked target.

The training dataset consists of four sequences from the PETS dataset with a total of 276 frames. One target in each sequence is identified and tracked to build up a latent space trajectory database. The selected PETS test dataset includes one sequence of thirty frames with two walking people,

one sequence of thirty frames with five walking people, and one sequence of forty frames with five walking people. These targets have clearly different trajectory patterns, and the forty-frame sequence also contains temporary target occlusion.

Table 1 summarizes the experimental results in terms of error rate and run time. Samples of tracking results on 30-frame test sequences are shown in Figures 3, 4, and 5 for three different appearance models. Figures 3 and 4 shows the tracking results which use the Histogram-Bhattacharya approach and GMM-KL appearance model to track 5 targets, while Figure 5 utilizes the rotation invariant appearance model to track 2 targets. From these results one can see that, just using approximately 20 particles, the PFGPDM approach can effectively track multiple targets that are following trajectories similar to the trained database.

Simulation results also indicate that GMM-KL approach is more discriminative in terms of the background and the object, compared to Bharttacharyya distance on histograms, because the latter approach may not represent the image structure as robust as the GMM-KL method.

However, Bharttacharyya distance approach is simple to implement and efficient in terms of computation time. The rotation invariant model with 4-control points and $\pi/2$ polar representation showed promising tracking results on feet, which was as expected. In addition, this appearance model is sensitive to the number of control points, which leads to performance and time complexity tradeoff. In general, rotation invariant model and GMM-KL appearance model provided more adaptive tracking results than Histogram-Bharttacharyya model at the expense of computation resource.

Another observation is that the particles do not deviate from the target in dark regions or feet under considerable occlusion. This is a result of particle filtering integrated with the Gaussian process prediction, despite that the importance update function of the particle filter relies on the appearance model of the templates and the test regions. The constraint on the length difference between the head and feet prevents mis-association of the targets. Figure 6 shows that the temporary occlusion in the test sequence was successfully resolved by our proposed framework. The yellow bounding box represents the passage with the dark red clothes; the cyan bounding box denotes the passage with the blue clothes. The two passengers were separated in the left frame and overlapped in the middle frame, and finally they were correctly tracked when they appeared separately again in the right frame. Gaussian process can also help to predict the next movement in sequences with missing frames. Figure 7 shows the tracking results of a missing frame case, in which 2 consecutive frames were arbitrarily selected and discarded. In addition our method was tested using all three appearance models on all 30-frame test sequences under missing frame situations. We found that, with 2 consecutive missing frames, the tracking error rates were identical to what appear in Table 1. However, if more frames were missing, we saw a clear increase in tracking error rate. Both Figures 6 and 7 were based on the GMM-KL appearance model.

Two comparative studies were also conducted, in which our method was compared with two existing methods with excellent performance, namely, the adaptive appearance-model based particle filter (AAMPF) proposed by Zhou et al., [4] and the trans-Dimensional Monte Carlo Particle filter (TDMCPF) proposed by Smith et al. [5]. Our method and these two methods share a similar particle filter framework. They differ at feature selections and appearance models. However the AAMPF can only track one target, and the TDMCPF can track indefinite number (up to four) of targets. The results of these studies are summarized in Figures 8, 9, and Table 2. The tracking results of the AAMPF was obtained using the software provided by the authors of [4] and tested on a PETS sequence. The results of the TDMCPF can be found at the author's website (http://www.idiap.ch/~smith/). To compare with the TDMCPF results, our method was tested on the IDIAP dataset that was used in [5]. It should be noted that we still use the trained trajectory database based on the PETS dataset in the tests on the IDIAP dataset. From these results we can see clearly that our method can achieve comparable object tracking performance with much less numbers of particles. Also our trained trajectory database as well as our training method are robust enough to accommodate substantial motion variations. These results of our method were based on the GMM-KL appearance model.

5. CONCLUSION

An integrated Gaussian process dynamical model with particle filter framework is proposed to track multiple targets, and handle temporary occlusion as well as noncontinuous frames. The experimental results indicate that the proposed PFGPDM approach can reliably track multiple targets at very low error rates with much reduced computational complexity and the number of particles. Under temporary occlusion and missing frame cases, the impacted targets were correctly tracked due to the accurate predictions from Gaussian process.

It should be pointed out that, although the test sequences used in this paper only contain close to linear motion patterns, there is no inherent difficulty for the proposed method to handle more complex motions. This is because that the particle filter framework is generally not constrained to linear motion. However, tracking such complex motion patterns may comprise the computational efficiency introduced in this work. The exact capability of the proposed method in dealing with various complex motion patterns can be a very interesting topic for future study.

REFERENCES

[1] Z. Khan, T. Balch, and F. Dellaert, "An MCMC-based particle filter for tracking multiple interacting targets," in *Proceedings of the 8th European Conference on Computer Vision (ECCV '04)*, pp. 279–290, Prague, Czech Republic, May 2004.

[2] K. Okuma, A. Taleghani, N. de Freitas, J. J. Little, and D. G. Lowe, "A boosted particle filter: multitarget detection and tracking," in *Proceedings of the 8th European Conference on Computer Vision (ECCV '04)*, pp. 28–39, Prague, Czech Republic, May 2004.

[3] C. Chang, R. Ansari, and A. Khokhar, "Multiple object tracking with kernel particle filter," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 568–573, San Diego, Calif, USA, June 2005.

[4] S. K. Zhou, R. Chellappa, and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters," *IEEE Transactions on Image Processing*, vol. 13, no. 11, pp. 1491–1506, 2004.

[5] K. Smith, D. Gatica-Perez, and J.-M. Odobez, "Using particles to track varying numbers of interacting people," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 962–969, San Diego, Calif, USA, June 2005.

[6] O. Lanz, "Approximate Bayesian multibody tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1436–1449, 2006.

[7] T. Yu and Y. Wu, "Collaborative tracking of multiple targets," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '04)*, vol. 1, pp. 834–841, Washington, DC, USA, June-July 2004.

[8] M. Isard and J. MacCormick, "BraMBLe: a Bayesian multiple-blob tracker," in *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV '01)*, vol. 2, pp. 34–41, Vancouver, Canada, July 2001.

[9] J. Czyz, B. Ristic, and B. Macq, "A particle filter for joint detection and tracking of color objects," *Image and Vision Computing*, vol. 25, no. 8, pp. 1271–1281, 2007.

[10] N. Lawrence, "Probabilistic non-linear principal component analysis with Gaussian process latent variable models," *The Journal of Machine Learning Research*, vol. 6, pp. 1783–1816, 2005.

[11] J. Wang, D. Fleet, and A. Hertzmann, "Gaussian process dynamical models," in *Advances in Neural Information Processing Systems 18*, Y. Weiss, B. Schölkopf, and J. Platt, Eds., pp. 1441–1448, MIT Press, Cambridge, Mass, USA, 2006.

[12] R. Urtasun, D. J. Fleet, and P. Fua, "3D people tracking with Gaussian process dynamical models," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, vol. 1, pp. 238–245, New York, NY, USA, June 2006.

[13] L. Raskin, E. Rivlin, and M. Rudzsky, "Using Gaussian process annealing particle filter for 3D human tracking," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 592081, 13 pages, 2008.

[14] F. Guo and G. Qian, "3D human motion tracking using manifold learning," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 1, pp. 357–360, San Antonio, Tex, USA, September-October 2007.

[15] M. Tong and Y. Liu, "Shared latent dynamical model for human tracking from videos," in *Proceedings of the International Workshop on Multimedia Content Analysis and Mining (MCAM '07)*, pp. 102–111, Weihai, China, June-July 2007.

[16] G. Klein and D. Murray, "Full-3D edge tracking with a particle filter," in *Proceedings of the 17th British Machine Vision Conference (BMVC '06)*, vol. 3, pp. 1119–1128, Edinburgh, UK, September 2006.

[17] M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B*, vol. 61, no. 3, pp. 611–622, 1999.

[18] M. Riedmiller and H. Braun, "RPROP—a fast adaptive learning algorithm," in *Proceedings of the 7th International Symposium on Computer and Information Sciences (ISCIS '92)*, pp. 279–285, Antalya, Turkey, 1992.

[19] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '00)*, vol. 2, pp. 142–149, Hilton Head Island, SC, USA, June 2000.

[20] H. Greenspan, J. Goldberger, and L. Ridel, "A continuous probabilistic framework for image matching," *Computer Vision and Image Understanding*, vol. 84, no. 3, pp. 384–406, 2001.

[21] S. Kullback, *Learning Textures*, Dover, New York, NY, USA, 1968.

[22] J. Kang, K. Gajera, I. Cohen, and G. Medioni, "Detection and tracking of moving objects from overlapping EO and IR sensors," in *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW '04)*, vol. 8, p. 123, Washington, DC, USA, June 2004.

[23] PETS 2007 Benchmark Data, "Pets In Conjunction with 11th IEEE International Conference on Computer Vision," http://www.cvg.rdg.ac.uk/PETS2007/data.html.

[24] "Neil lawrence Gaussian process software," http://www.cs.man.ac.uk/~neill/software.html.

Special Issue on
Dependable Semantic Inference

Call for Papers

After many years of exciting research, the field of multimedia information retrieval (MIR) has become mature enough to enter a new development phase—the phase in which MIR technology is made ready to get adopted in practical solutions and realistic application scenarios. High users' expectations in such scenarios require high dependability of MIR systems. For example, in view of the paradigm "getting the content I like, anytime and anyplace" the service of consumer-oriented MIR solutions (e.g., a PVR, mobile video, music retrieval, web search) will need to be at least as dependable as turning a TV set on and off. Dependability plays even a more critical role in automated surveillance solutions relying on MIR technology to analyze recorded scenes and events and alert the authorities when necessary.

This special issue addresses the dependability of those critical parts of MIR systems dealing with semantic inference. Semantic inference stands for the theories and algorithms designed to relate multimedia data to semantic-level descriptors to allow content-based search, retrieval, and management of data. An increase in semantic inference dependability could be achieved in several ways. For instance, better understanding of the processes underlying semantic concept detection could help forecast, prevent, or correct possible semantic inference errors. Furthermore, the theory of using redundancy for building reliable structures from less reliable components could be applied to integrate "isolated" semantic inference algorithms into a network characterized by distributed and collaborative intelligence (e.g., a social/P2P network) and let them benefit from the processes taking place in such a network (e.g., tagging, collaborative filtering).

The goal of this special issue is to gather high-quality and original contributions that reach beyond conventional ideas and approaches and make substantial steps towards dependable, practically deployable semantic inference theories and algorithms.

Topics of interest include (but are not limited to):

- Theory and algorithms of robust, generic, and scalable semantic inference
- Self-learning and interactive learning for online adaptable semantic inference

- Exploration of applicability scope and theoretical performance limits of semantic inference algorithms
- Modeling of system confidence in its semantic inference performance
- Evaluation of semantic inference dependability using standard dependability criteria
- Matching user/context requirements to dependability criteria (e.g., mobile user, user at home, etc.)
- Modeling synergies between different semantic inference mechanisms (e.g., content analysis, indexing through user interaction, collaborative filtering)
- Synergetic integration of content analysis, user actions (e.g., tagging, interaction with content) and user/device collaboration (e.g., in social/P2P networks)

Authors should follow the EURASIP Journal on Image and Video Processing manuscript format described at http://www.hindawi.com/journals/ivp/. Prospective authors should submit an electronic copy of their complete manuscripts through the journal Manuscript Tracking System at http://mts.hindawi.com/, according to the following timetable:

Manuscript Due	December 1, 2009
First Round of Reviews	March 1, 2010
Publication Date	June 1, 2010

Guest Editors

Alan Hanjalic, Delft University of Technology, 2600 AA Delft, The Netherlands; a.hanjalic@tudelft.nl

Tat-Seng Chua, National University of Singapore, Singapore 119077; chuats@comp.nus.edu.sg

Edward Chang, Google Inc., China; University of California, Santa Barbara, CA 93106, USA; echang@ece.ucsb.edu

Ramesh Jain, University of California, Irvine, CA 92697, USA; jain@ics.uci.edu

Special Issue on
Advances in Multidimensional Synthetic Aperture Radar Signal Processing

Call for Papers

Synthetic Aperture Radar (SAR) represents an established and mature high-resolution remote-sensing technology that has been successfully employed to observe, study, and characterize the Earth's surface. Over the last two decades, one-dimensional SAR systems have evolved more and more toward multidimensional configurations, enabling quantitative remote sensing with SAR sensors. The relevance of the multidimensional SAR technology is primarily supported by the recent launch of the Japanese ALOS, the German TERRASAR-X, and the Canadian RADARSAT-2 orbital systems, and also by the development of airborne and ground-based SAR systems.

Multidimensional SAR data can be generated by different sources of diversity: baseline, polarization, frequency, time, as well as their different combinations. This degree of freedom makes multidimensional SAR data sensitive to a wide range of geophysical and biophysical features of the Earth's surface. Consequently, the definition of novel multidimensional signal processing techniques is essential to benefit from this information richness, especially when the objective is the quantitative retrieval of new parameters, and also necessary to extend already existing one-dimensional SAR data tools.

This special issue is seeking for original contributions in the definition of novel signal processing techniques, and also for works on the assessment of new physical or statistical models to improve the understanding of multidimensional SAR data and the extraction of information, considering the important challenges and limitations imposed by the physics governing the imaging process. This issue is also open to contributions oriented toward the exploitation of multidimensional SAR data for novel and exciting applications.

Topics of interest include, but are not limited to:

- Multibaseline interferometry and differential interferometry
- 3D reconstruction and multidimensional SAR focusing techniques
- Polarimetry and polarimetric interferometry
- Multitemporal and multifrequency SAR
- Novel multidimensional SAR system configurations
- Multidimensional SAR data classification and change detection
- Information extraction from multidimensional SAR data
- Multidimensional SAR data statistical modeling, filtering, and estimation
- Definition and assessment of electromagnetic models
- Extraction and estimation of geophysical and biophysical parameters
- Space-time adaptive processing (STAP)

Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://www.hindawi.com/journals/asp/guidelines.html. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at http://mts.hindawi.com/ according to the following timetable:

Manuscript Due	June 1, 2009
First Round of Reviews	September 1, 2009
Publication Date	December 1, 2009

Lead Guest Editor

Carlos López-Martínez, Department of Signal Theory and Communication, Technical University of Catalonia, c/Jordi Girona 1-3, 08034 Barcelona, Spain; carlos.lopez@tsc.upc.edu

Guest Editors

Laurent Ferro-Famil, Institute of Electronics and Telecommunications of Rennes, University of Rennes 1, 263 Avenue Gènèral Leclerc, 35042 Rennes Cedex, France; laurent.ferro-famil@univ-rennes1.fr

Andreas Reigber, Microwaves and Radar Institute, German Aerospace Center (DLR), P.O. Box 1116, 82230 Weßling, Germany; andreas.reigber@dlr.de