Multiple Body Part Tracking Using a Probabilistic Data Association Filter

Harish Bhaskar*, Lyudmila Mihaylova* and Simon Maskell**
*Dept. of Communication Systems, Lancaster University, UK
**QinetiQ, Malvern, Worcestershire, UK
{*h.bhaskar,mila.mihaylova*}*@lancaster.ac.uk*, *srmaskell@qinetiq.com*

Abstract

This paper presents a framework for multiple body part tracking based on a probabilistic data association (DA) filter. The body parts are extracted using iterative cluster background subtraction and foreground modeling with pictorial structures. The background subtracted silhouette is cluttered and the body parts are subject to occlusions. The main novelty of the paper is in the effective solution for data association which involves tracking body parts based on the expected likelihood method. We also show the advantage of the expected likelihood DA over the standard Probabilistic Data Association Filter (PDAF). A number of experiments have been conducted on several synthetic and real-time data sets and encouraging results have been obtained.

1 Introduction and Previous Work

Multiple body part object tracking is a challenging area of research within computer vision and tracking communities due to the various challenges that it poses both from theoretical and practical point of view. The high degree of freedom of articulating regions with inter-dependencies between them requires efficient techniques able to cope effectively with the dynamic changes of the objects and background. Motion tracking consists of two distinct phases: *detection* and *tracking*. The detection process primarily aims at segmenting the human object from the frames of the video sequences. Tracking, on the other hand, involves spatially locating these detected regions in time. A number of techniques have been proposed for the detection and tracking phases [8, 10, 25]. In this paper, we present a probabilistic data association filter for multiple body parts tracking of a single articulating human object. These body parts are extracted using an iterative cluster level background subtraction mechanism and pictorial structure for foreground modeling.

Various techniques have been proposed for solving body parts object tracking problems that include, but are not restricted to features based methods, gradient techniques and probabilistic methods [24, 9, 25]. The presence of multiple parts of the body with similar feature characteristics leads to uncertainty in the origin of measurements. In this paper, we focus our attention on tracking human body parts based on multiple independent particle filters and the probabilistic data association algorithm which afford handling the uncertainty due to the measurement origin.

Data association (DA) refers to deciding which of the received multiple measurements to use to update the trajectory of the moving target [6, 26, 17]. Two important solutions to the DA problem are in the form of the strongest neighbour filter and the nearest neighbour filter [6, 15]. While the strongest neighbour filter chooses the measurement with the highest intensity specified within a bounding gate, the nearest neighbour filter prefers the closest measurement. These techniques are well suited for applications where the number of validated measurements is low and when the objects do not manoeuvre. In such situations a PDAF is useful. The general idea behind the probabilistic data association is, instead of using only one measurement among the multiple received ones and discarding the others, rather use all of the validated measurements with different weight (probabilities).

Data association is a problem of crucial importance for also multiple target tracking because of the necessity to relate each measurement to the correct object. Many methods have been proposed in the estimation and tracking literature [16, 6, 15, 20, 18], [19, 12, 7, 1, 23] both in the cases of known and unknown number of multiple targets. The observations are assumed to originate from different targets or from clutter. In some applications only one measurement is assumed available from each target, where in other applications several returns are available. This will of course reflect which data association method to use. The number of targets is estimated in other works [22, 21, 27, 14, 13], by using, e.g., the so called birth-death model.

2 Particle Filtering for Body Parts Tracking

In order to track the objects of interest, the posterior probability density function $p(\boldsymbol{x}_k|\boldsymbol{z}_k)$ of the state vector $\boldsymbol{x}_k$ has to be estimated at each time step k given the measurement vector $\boldsymbol{z}_k$. Within the recursive Bayesian framework this is performed based on the two step process of prediction and update. In the *prediction step* the probability of the current state based on its previous measurement is made:

$$p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1}) = \int p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})p(\boldsymbol{x}_{k-1}|\boldsymbol{z}_{k-1})d\boldsymbol{x}_{k-1}. \tag{1}$$

The *filtering step* is specified using:

$$p(\boldsymbol{x}_k|\boldsymbol{z}_k) \propto p(\boldsymbol{z}_k|\boldsymbol{x}_k)p(\boldsymbol{x}_k|\boldsymbol{z}_{k-1}). \tag{2}$$

The prediction step follows from marginalisation, and the new distribution is obtained by directly applying the Bayes' rule. The dynamics of the body parts is specified by the model $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$ and the feature likelihood is estimated in the light of the current observation $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$. The initial state probability density function $p(\boldsymbol{x}_0)$ is assumed to be known (supplied by a background subtraction technique [4, 5]). Once the sequence of the filtering distribution is known, point estimates of the state vector can be obtained using any of the appropriate function such as maximum-a-posteriori (MAP) estimates or minimum mean square estimates (MMSE).

A particle filter is implemented where the weighted set of samples $\left\{\boldsymbol{x}_k^{(i)}, \boldsymbol{w}_k^{(i)}\right\}_{i=1}^{N}$ approximates the state probability density function $p(\boldsymbol{x}_k|\boldsymbol{z}_k)$. New samples are generated from a suitable proposal distribution, which in turn depends on the previous state and new measurements, i.e., $\boldsymbol{x}_k^{(i)} \sim q_p\left(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{z}_k\right)$, $i = 1, 2, ..., N$. The new importance weights are set based on:

$$\boldsymbol{w}_{k+1}^{(i)} \propto \boldsymbol{w}_k^{(i)} \frac{p(\boldsymbol{z}_k|\boldsymbol{x}_k^{(i)})p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})}{q_p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1}^{(i)}, \boldsymbol{z}_k)} \tag{3}$$

where the sum of the weights is 1. The efficiency of the particle filter algorithm depends on the quality of the proposal distribution for the state and the likelihood function.

3 Motion and Measurement Models

The state space approach requires specifying the motion model $p(\boldsymbol{x}_k|\boldsymbol{x}_{k-1})$, and a measurement model $p(\boldsymbol{z}_k|\boldsymbol{x}_k)$ which links the state vector to the measurement vector. The state space of any object within a frame k can typically consist of different kinematic and region parameters. The state vector $\boldsymbol{x}_k = (x_k, \dot{x}_k, y_k, \dot{y}_k)'$ comprises the position coordinates x_k, y_k and the velocities $\dot{x}_k$, $\dot{y}_k$ of the centre of the target regions of the object in the image. The evolution of the state vector is described using the linear constant velocity model [2]:

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{G}\boldsymbol{v}_{1,k}, \tag{4}$$

where $\boldsymbol{F} = \mathrm{diag}(\boldsymbol{F}_1, \boldsymbol{F}_1)$, $\boldsymbol{F}_1 = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}$, $\boldsymbol{G} = \begin{pmatrix} T/2 & 1 & 0 & 0 \\ 0 & 0 & T/2 & 1 \end{pmatrix}'$ and $T = 1$ is the sampling interval. The noise $\boldsymbol{v}_{1,k}$ is assumed to be zero mean, white, Gaussian with covariance matrix $\boldsymbol{Q}$, i.e., $\boldsymbol{v}_{1,k} \sim \mathcal{N}(0, \boldsymbol{Q})$.

Cluster background subtraction [4, 5] is applied to a set of $\mathfrak{R}$ images and as a result the silhouette of the human is obtained in the presence of clutter. The colour cue and histogram are used as a measurement model. Due to this nonlinearity a particle filter is used to estimate the state vector for each body part. The PDAF approach assumes that all measurements are in a particular target extension gate and originated either from a target or from random clutter. We suggest solving our single object multiple parts tracking problem in video sequences with the assumption that the positional interference of different body parts are resolved spatially and each different body part is tracked using a separate particle filter where the data association ambiguity is resolved by the proposed Expected Likelihood Data Association Filter (described in the next section).

4 An Expected Likelihood Data Association Filter for Body Parts Motion Analysis

In this paper we propose a solution to the data association while tracking body parts in clutter based on the expected likelihood method [18]. Human body parts are subject to partial and full occlusions which creates problems to the algorithms for motion analysis. In the presence of a number of different body parts to be tracked, the standard approach [6, 15] to tracking in clutter is with the use of a Kalman filter or an extended Kalman filter and gating the measurement based on the predicted measurement covariance. The predicted state is then updated using this measurement covariance within probabilistic data association.

However, the predicted measurement covariance is not directly available in the particle filter and can be reconstructed approximately from the current cloud of particles. In this paper, the expected likelihood is calculated from the measurements and clutter statistics which is further used to validating the measurements against clutter. In subsection 4.1, we illustrate the standard mechanism of tracking using the PDAF and in subsection 4.2, we introduce the proposed expected likelihood probabilistic data association (ELPDA). We show the advantage of the ELPDA over the standard PDAF.

4.1 A Probabilistic Data Association Filter (PDAF)

Suppose the set of measurements at time k is $\boldsymbol{Z}_k = \{\boldsymbol{z}_{k,j}\}_{j=1}^{m_k}$ and m_k is the number of measurements falling within a validation gate. The cumulative set of validated measurements up to time k is denoted as $\boldsymbol{Z}^k = \{\boldsymbol{Z}_\ell\}_{\ell=1}^{k}$. At each scan, the validation gate, centered around the predicted measurement $\widetilde{\boldsymbol{z}}_{k|k-1}$ of the target, is setup to select the measurements associated probabilistically with the target. The validation region is

$$\left(\boldsymbol{z}_k - \widetilde{\boldsymbol{z}}_{k|k-1}\right)^{'} \boldsymbol{S}_k^{-1} \left(\boldsymbol{z}_k - \widetilde{\boldsymbol{z}}_{k|k-1}\right) < g^2, \tag{5}$$

where $\boldsymbol{S}_k$ is the covariance matrix of the innovation process $\boldsymbol{v}_k = \boldsymbol{z}_k - \widetilde{\boldsymbol{z}}_{k|k-1}$ corresponding to the true measurement and g is the gate size and $'$ is the transpose operation.

The standard PDAF for dealing with the measurement uncertainty incorporates the Kalman filter and the respective relationships are briefly summarised below. The prediction $\widetilde{\boldsymbol{x}}_{k|k-1}$ of the state vector and of the prediction $\widetilde{\boldsymbol{z}}_{k|k-1}$ of the measurement vector at time k are defined, respectively, as: $\widetilde{\boldsymbol{x}}_{k|k-1} = \boldsymbol{F}\widetilde{\boldsymbol{x}}_{k-1|k-1}$ and $\widetilde{\boldsymbol{z}}_{k|k-1} = \boldsymbol{H}\widetilde{\boldsymbol{x}}_{k-1|k-1}$ where $\boldsymbol{H}$ is the matrix of the linear measurement equation

$$\boldsymbol{z}_k = \boldsymbol{H}_k \boldsymbol{x}_k + \boldsymbol{v}_{2,k}, \tag{6}$$

where $\boldsymbol{v}_{2,k}$ is the measurement noise, assumed to be a white Gaussian process with covariance matrix $\boldsymbol{R}_v$. The state covariance matrix $\boldsymbol{P}$ is given by:

$$\boldsymbol{P}_k = \boldsymbol{F}\boldsymbol{P}_{k-1}\boldsymbol{F}^{'} + \boldsymbol{G}\boldsymbol{Q}\boldsymbol{G}^{'}. \tag{7}$$

The state update equation of the PDAF is represented using

$$\widetilde{\boldsymbol{x}}_k = \widetilde{\boldsymbol{x}}_{k|k-1} + \boldsymbol{W}_k \boldsymbol{v}_k, \tag{8}$$

where the Kalman gain $\boldsymbol{W}_k$ is computed as follows,

$$\boldsymbol{W}_k = \boldsymbol{P}_k \boldsymbol{H} \boldsymbol{S}_k^{-1}, \tag{9}$$

$$\boldsymbol{S}_k = \boldsymbol{H}\boldsymbol{P}_k\boldsymbol{H}^{'} + \boldsymbol{R}_v. \tag{10}$$

The combined innovation process is

$$\boldsymbol{v}_k = \Sigma_{j=1}^{m_k} \beta_{j,k} \boldsymbol{v}_{j,k}. \tag{11}$$

where the separate innovation vectors are

$$\boldsymbol{v}_{j,k} = \boldsymbol{z}_{j,k} - \widetilde{\boldsymbol{z}}_{k|k-1}. \tag{12}$$

The overall covariance $\boldsymbol{P}_k$ associated with the state vector is given in the form [17]

$$\boldsymbol{P}_k = \boldsymbol{P}_0 + \boldsymbol{W}_k \left\{ \beta_{0,k} \boldsymbol{S}_k + \Sigma_{j=1}^{m_k} \left[\beta_{j,k} \boldsymbol{v}_{j,k} \boldsymbol{v}_{j,k}^{'} \right] - \boldsymbol{v}_{j,k} \boldsymbol{v}_{j,k}^{'} \right\} \boldsymbol{W}_k^{'}, \tag{13}$$

$$\boldsymbol{P}_0 = \boldsymbol{P}_{k-1} - \boldsymbol{W}_k \boldsymbol{S}_k \boldsymbol{W}_k^{'}, \tag{14}$$

where $\beta_{j,k}$ denotes the probability that the jth measurement comes from the target in track at time k; $\beta_{0,k}$ denotes the probability that none of the measurements is originated from the target, or equivalently, the probability that the current measurement is a false alarm (or clutter). To evaluate the association probabilities, the conditioning is broken down into the past data $\boldsymbol{Z}^{k-1}$ and the latest data $\boldsymbol{Z}^k$. A probabilistic inference can be made on both the number of measurements in the validation region (from the clutter density, if know) and on their locations, expressed as [3]

$$\begin{aligned}\beta_{j,k} &= P(\theta_{j,k}|\boldsymbol{Z}^k) \\ &= \frac{1}{c_k} p\left[\boldsymbol{Z}^k|\theta_{j,k}, m_k, \boldsymbol{Z}^{k-1}\right] \boldsymbol{P}(\theta_{j,k}|m_k, \boldsymbol{Z}^{k-1}), \quad j = 0, 1, ..., m_k, \end{aligned} \tag{15}$$

where c_k is a normalisation constant, $\theta_{j,k}$ denotes the event that the measurement $\boldsymbol{z}_{j,k}$ is originated from the target, $\theta_{0,k}$ denotes the event that none of the measurements is originated from the target.

Using the parametric PDAF [15], the association probabilities are:

$$\beta_{0,k} = \frac{b}{b + \Sigma_{i=1}^{m_k} e_i}, \quad j = 0, \tag{16}$$

$$\beta_{j,k} = \frac{e_j}{b + \Sigma_{i=1}^{m_k} e_i}, \quad j = 1, 2, ...m_k, \tag{17}$$

where

$$b = \left[\frac{2\pi}{g}\right]^{m/2} \lambda V_k \frac{1 - P_D P_G}{P_D}, \tag{18}$$

$$e_j = \exp\left[-\boldsymbol{v}'_{j,k} \boldsymbol{S}_k^{-1} \boldsymbol{v}_{j,k}/2\right], \tag{19}$$

$$V_k = \frac{\pi^{m/2}}{\Gamma(m/2+1)} g^{m/2} \left|\boldsymbol{S}_k\right|^{1/2}, \tag{20}$$

where V_k is the volume of the ellipsoidal validation region, P_D is the probability of detection, P_G is the probability that the target measurement falls in the m-dimensional validation region, λ is the spatial density of false measurements and $\Gamma(.)$ is the Gamma distribution. The weights $\beta_{j,k}$ fulfill the constraint:

$$\Sigma_{j=0}^{m_k} \beta_{j,k} = 1, \; 0 \leq \beta_{j,k} \leq 1. \tag{21}$$

4.2 Expected Likelihood Probabilistic Data Association (ELPDA)

The ELPDA technique is used in the context of tracking in clutter, where it is necessary to identify and decide which of the measurements are correct and which are not. Generic particle filters are capable of modeling non-Gaussian probability density functions. However, when tracking an object with a particle filter, an analog of the predicted measurement covariance is not directly available. One possible solution then is the expected likelihood particle filter [18]. In order to account for the uncertainty in the measurement origin, the expected likelihood can be computed as the sum of the individual likelihoods of the measurements, with the weights provided by the PDA.

For the set of available measurements, we assume that one of the measurements originates from the target, and the rest are due to spurious clutter. In the case of body parts tracking the colour cue is used and the measurement equation is highly nonlinear. According to [18] the conditional probability $p(\theta_{j,k}|\boldsymbol{Z}^k)$ of the association event $\theta_{j,k}$ that the jth measurement in the gate is the correct measurement is given by

$$p(\theta_{j,k}|\boldsymbol{Z}^k) = p(\theta_{j,k}|\boldsymbol{z}_k, m_k, \boldsymbol{Z}^{k-1}) \tag{22}$$

for the set of m_k measurements $\boldsymbol{z}_k$ that fall within the validation gate.

From the Bayesian theorem, it follows that

$$p(\theta_{j,k}|\boldsymbol{z}_k, m_k, \boldsymbol{Z}^{k-1}) \propto p(\boldsymbol{z}_k|\theta_{j,k}, m_k, \boldsymbol{Z}^{k-1}) \times p(\theta_{j,k}|m_k, \boldsymbol{Z}^{k-1}). \tag{23}$$

If the incorrect measurements have a uniform probability density function within the gating volume V_k, and with the assumption of a normal measurement error for a correct measurement, we have

$$p(\boldsymbol{z}_k|\theta_{j,k}, m_k, \boldsymbol{Z}^{k-1}) = \begin{cases} V_k^{-(m_k-1)} P_G^{-1} \mathcal{N}(\check{\boldsymbol{z}}_{j,k}; 0; \boldsymbol{S}_k), & j = 1, \ldots, m_k; \\ V_k^{-m_k}, & j = 0 \end{cases} \tag{24}$$

where P_G is the probability of gating, $\boldsymbol{S}_k$ is the covariance matrix of the innovation vector $\check{\boldsymbol{z}}_{j,k}$.

Finally, it can be shown [18] that the probability $p(\theta_{j,k}|\boldsymbol{z}_k, m_k, \boldsymbol{Z}^k)$ of association events can be calculated as follows:

$$p(\theta_{j,k}|m_k, \boldsymbol{Z}^k) = \begin{cases} \frac{1}{C} \times P_G^{(-1)} \frac{(P_D P_G)\mu_F(m_k-1)}{m_k P_D P_G \mu_F(m_k-1) + (1-P_D P_G)\mu_F(m_k)} V_k^{-(m_k-1)} \times \\ \mathcal{N}(\boldsymbol{z}_k; 0; \boldsymbol{S}_k), \qquad j = 1, ..., m_k \ ; \\ \\ \frac{(1-P_D P_G)\mu_F(m_k)}{P_D P_G \mu_F(m_k-1) + (1-P_D P_G)\mu_F(m_k)}, j = 0, \end{cases}$$

where P_D is the probability of detection, μ_F is the probability mass function of the number of incorrect measurements, $P_D P_G$ refers to the probability that a target us detected and its measurements fall within the gate and $\boldsymbol{R}_k$ is the measurement error covariance matrix.

Substituting the expected likelihood into the relation for the particle weights, we have:

$$w_k^{(i)} \propto w_{k-1}^{(i)} \times \frac{p(\theta_{0,k}|m_k,\boldsymbol{z}_{k-1})V_k^{-(m_k)}+\sum_{j=1}^{m_k} p(\theta_{j,k}|m_k,\boldsymbol{z}_{k-1})V_k^{-(m_k-1)} P_G^{-1}\mathcal{N}(\boldsymbol{z}_k,\check{\boldsymbol{z}}_{j,k},\boldsymbol{R}_k)p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})}{q(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)},\boldsymbol{z}_k)}. \quad (25)$$

where $i = 1, \ldots, N$. We assume that the importance function is chosen to be the prior $p(\boldsymbol{x}_k^{(i)}|\boldsymbol{x}_{k-1}^{(i)})$ and then the weight update becomes the mixture of the PDA weighted measurement likelihoods for the correct and incorrect measurements.

5 Results

This section of the paper presents results comparing the tracking performance of PDAF and ELPDA approaches on video data from Carnegie Mellon University [11].

First, a cluster background subtraction technique [4, 5] is applied to extract the silhouette of the human from the first $\mathfrak{R} = 10$ video frames. Some of the BS images are shown in Figure 1 where we can see the occlusion between the body parts and the presence of clutter. A foreground modeling technique using pictorial structures and a genetic search algorithm [5] is applied to isolate each body part. Since the background subtracted silhouette is subject to clutter and occlusions, particle filters with the proposed ELPDAF is applied to track the body parts of the moving person.

Here, we give the graphs of the trajectories of 4 different body parts: head, torso, an arm and a leg across a number of different frames in a sequence. We also compare the results of ELPDA, PDAF for different body parts with the results of a foreground detection method and manually labeled ground truth.

The trajectories in Figure 2 represent the predicted spatial coordinates of the body parts using the proposed ELPDA technique, PDAF algorithm, the ground truth of the positions that are manually labeled and the results of the detection of foreground using the genetic algorithm search. Results from separate video frames are shown in Figure 3.

A number of important conclusions can be drawn from the above graphs. First, it is evident that the PDAF performs accurately and produces results similar to the ground truth data. Second, the foreground detection method produces less accurate results compared with the other methods particularly

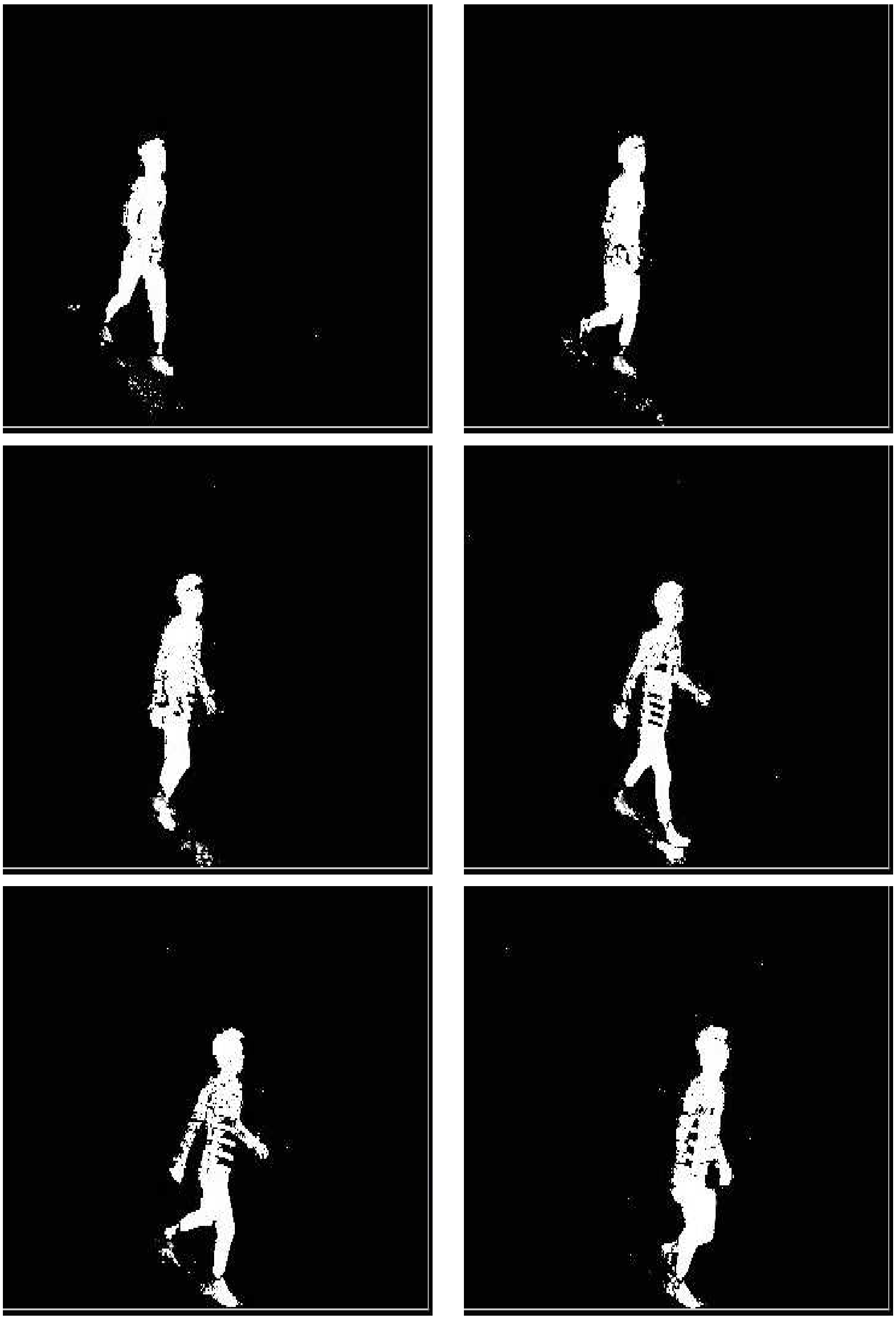

Figure 1: Background subtracted images

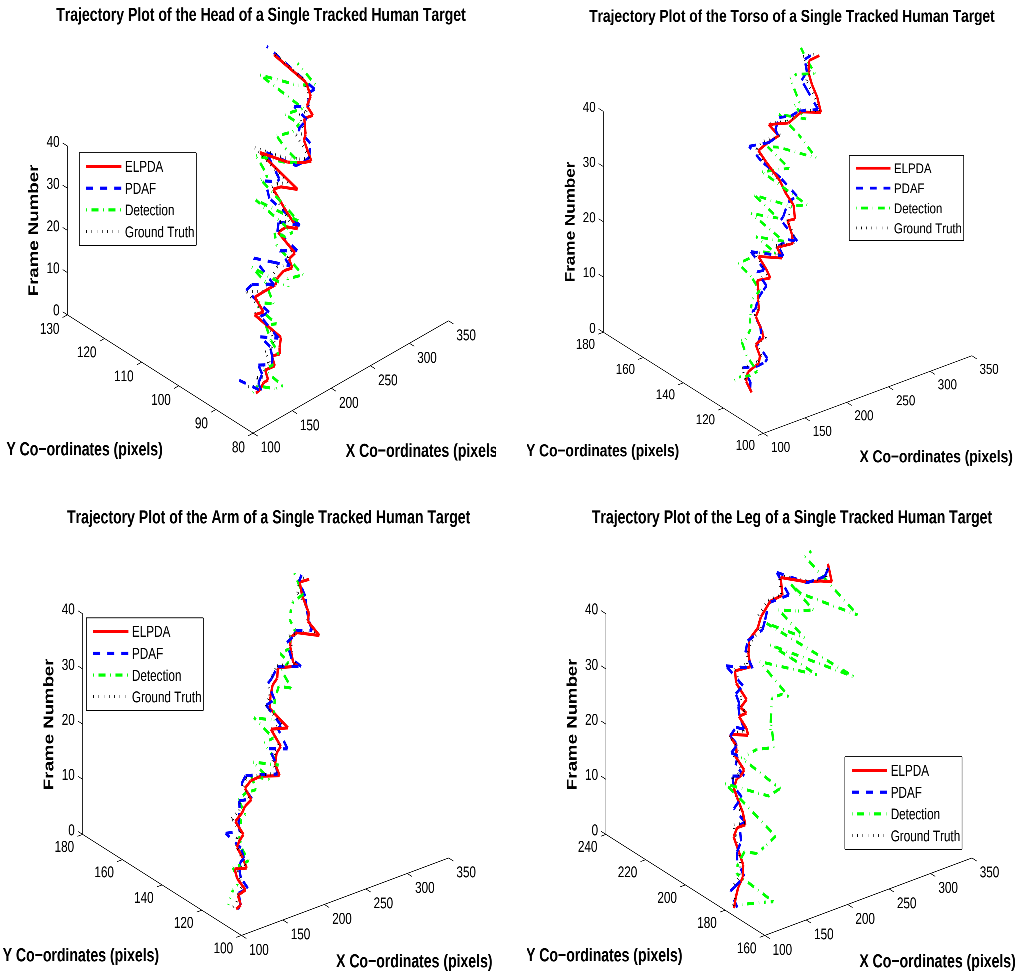

Figure 2: Position coordinates of the centre of the body parts estimated using the ELPDA and PDAF are compared with the ground truth data and foreground detection using a genetic search algorithm [5]

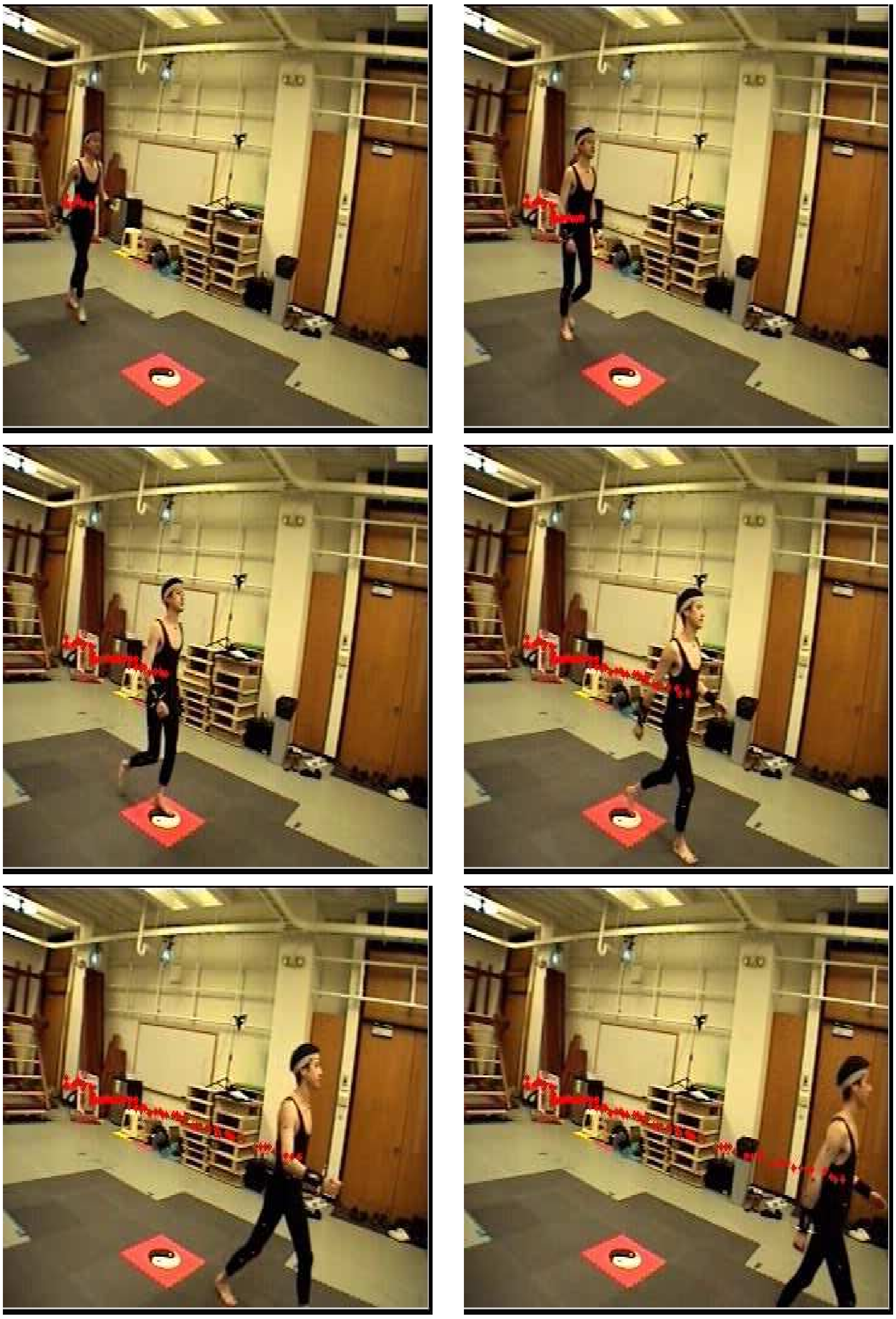

Figure 3: Results with the CMU video data for frames: 9, 21, 33, 45, 57 and 69. The dots in the video frames show the estimates for the centre of the torso.

because of the evolution search mechanism that fails to converge to a solution in efficient time. Next, it is obvious that the head and the torso are quite accurately tracked over the sequence as compared with the legs and the arms. The main reason attributed to this is because of the degree of freedom that these parts of the body have (i.e., rotational and spatial movements) as against the heard and torso regions. The only main drawback of the proposed scheme is that the tracking of parts works in serial which increases the computation demands of the algorithm. Also, in some areas of movements where one arm or leg region cross another during occlusion, there is a high degree of possibility that the measurements from them can be misinterpreted. However, these limitations can be overcome with the Joint PDAF approach.

6 Conclusions

This paper presents a framework for multiple body parts tracking in the presence of measurement uncertainty. Promising results are presented based on an expected likelihood PDAF and they are compared with the standard PDAF. Current work is focussed on multiple people body parts tracking, where all parts of a human will be considered as a whole, instead of using separate tracking algorithms for each body part.

Acknowledgements. The authors acknowledge the support from the UK MOD Data and Information Fusion Defence Technology Centre under the Tracking Cluster project DIFDTC/CSIPC1.

References

[1] Y. Bar-Shalom and W. Dale Blair (Eds.). *Multitarget-Multisensor Tracking: Applications and Advances, vol. 3*, volume II. Artech House, 2000.

[2] Y. Bar-Shalom and X.R. Li. *Estimation and Tracking: Principles, Techniques and Software*. Artech House, 1993.

[3] Y. Bar-Shalom and X.R. Li. *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[4] H. Bhaskar, L. Mihaylova, and S. Maskell. Automatic target detection based on background modeling using adaptive cluster density estimation. In *LNCS from the 3rd German Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, pages 130–134, Universität Bremen, Germany, 24-28 Sept., 2007.

[5] H. Bhaskar, L. Mihaylova, and S. Maskell. Automatic human body parts detection based on cluster background subtraction and foreground learning. *IEEE Transactions on Systems, Man and Cybernetics, submitted*, 2008.

[6] S. Blackman and R. Popoli. *Design and Analysis of Modern Tracking Systems*. Artech House Radar Library, 1999.

[7] M. Briers, S. Maskell, and M. Philpott. Two-dimensional assignment with merged measurements using Lagrangian relaxation. In *Proc. of SPIE Conference on Signal Processing of Small Targets*, pages 283–292, 2003.

[8] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(5):603–619, 2002.

[9] P. Felzenswalb. Learning models for object recognition. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2001.

[10] D. Forsyth, O. Arikan, L. Ikemoto, and D. Ramanan. *Computational Studies of Human Motion: Part 1, Tracking and Motion Synthesis. Foundations and Trends in Computer Graphics and Vision*. Hanover, Massachusetts. Now Publishers Inc., 2006.

[11] R. Gross and J. Shi. The CMU motion of body (MoBo) database (CMU-RI-TR-01-18). Technical report, Robotics Inst., Carnegie Mellon Univ. The data are available at: $http://mocap.cs.cmu.edu/$, 2001.

[12] P. Horridge and S. Maskell. Real-time tracking of hundreds of targets with efficient exact JPDAF implementation. In *Proceedings of International Conf. on Information Fusion*, 2006.

[13] C. Hue, J. Le Cadre, and R. Pérez. Sequential Monte Carlo methods for multiple target tracking and data fusion. *IEEE Trans. on Aerospace and Electronic Systems*, 38(3):791–812, 2002.

[14] C. Hue, J. Le Cadre, and R. Pérez. Tracking multiple objects with particle filtering. *IEEE Trans. on Aerospace and Electronic Systems*, 38(3):791–812, 2002.

[15] T. Kirubarajan and Y. Bar-Shalom. Probabilistic data association techniques for target tracking in clutter. *Proc. of the IEEE*, 92(3):536–557, 2004.

[16] X. R. Li. *Engineer's Guide to Variable-Structure Multiple-Model Estimation for Tracking. In Multitarget-Multisensor Tracking: Applications and Advances. Ch. 3. Eds. Y. Bar-Shalom and W. D. Blair, pp. 499–567*, volume 3. Artech House, 2000.

[17] Li Liangqun, Ji Hongbing, and Gao Xinbo. Maximum entropy fuzzy clustering with application to real-time target tracking. *Signal Processing*, 86(11):3432–3447, 2006.

[18] A. Marrs, S. Maskell, and Y. Bar-Shalom. Expected likelihood for tracking in clutter with particle filters. In *Proc. of the SPIE Conf., Vol. 4728*, Apr. 1-5 2002.

[19] S. Maskell, M. Briers, and R. Wright. Fast mutual exclusion. In *Proceedings of SPIE Conference on Signal Processing of Small Targets*, 2004.

[20] S. Maskell, M. Rollason, N. Gordon, and D. Salmond. Efficient particle filtering for multiple target tracking with application to tracking in structural images. *Image and Video Computing*, 21(10):931–939, 2003.

[21] W. Ng, J. Li, S. Godsill, and J. Vermaak. A hybrid approach for on-line joint detection and tracking for multiple targets. In *Proc. of the IEEE Aerospace Conf.*, 2005.

[22] W. Ng, J. Li, S. Godsill, and J. Vermaak. Multi-target tracking using a new soft-gating approach and sequential Monte Carlo methods. In *Proc. of the IEEE Intl. Conf. on Acoustic, Speech and Signal Processing*, 2005.

[23] L. Pao. Multisensor multitarget mixture reduction algorithms for target tracking. *AIAA Journal of Guidance, Control and Dynamics*, 17:1205–1211, 1994.

[24] A.P. Pentland. Recognition by parts. In *In IEEE International Conference on Computer Vision*, pages 612–620, 1987.

[25] D. Ramanan, D. A. Forsyth, and A. Zisserman. Tracking people by learning their appearance. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(1):65–81, 2007.

[26] D. Salmond, D. Fisher, and N. Gordon. Tracking and identification for closely spaced objects in clutter. In IEEE, editor, *Proc. of the European Control Conf.*, Brussels, Belgium, July 1997. IEEE.

[27] J. Vermaak, S. Godsill, and P. Pérez. Monte carlo filtering for multi-target tracking and data association. *IEEE Trans. on Aerospace and Electronic Systems*, 41(1):309–332, 2005.