Proc. 2006 IEEE Int. Conf. on Multisensor Fusion and Integration in Intelligent Systems (MFI2006), pp. 547-552, Heidelberg, Germany, Sep. 2006.

Visual Tracking of Multiple Persons in a Heavy Occluded Space Using Person Model and Joint Probabilistic Data Association

Maarten Liebens, Takuro Sakiyama and Jun Miura

Abstract— **This paper presents a new approach to the problem of image-based tracking of multiple persons in a heavy occluded space using a single camera. The presence of heavy occlusions results in uncertain measurement data. Examples of heavy occlusions are objects which impede the observation of a person, the overlap of multiple persons, etc. This measurement uncertainty can be partially by-passed if the process knows more about a person's expected size, i.e. person model. This way the observed measurement can be corrected using the introduced person model. Also the uncertainty of the measurements will be calculated with this person model. Subsequently, the corrected measurement is used to estimate the person's state (i.e. position and velocity) in the Kalman filter resulting in a more robust tracking. Next, tracking multiple persons jointly implies the need for a data association technique. This paper uses the Joint Probabilistic Data Association (JPDA) filter which calculates the a posteriori probabilities of the measurements having probably originated from the tracked persons. Finally, the approach has been implemented and tested on a single static camera bearing in mind that it will be applied on a mobile camera or robot. The approach presented here will verify whether the use of a single camera, wherefrom only 2D image-based data is gathered, delivers satisfactory tracking results using the Kalman and JPDA filter.**

I. INTRODUCTION

This paper presents a new approach to the problem of tracking multiple persons jointly using a single camera. Visual analysis of human motions has been widely studied [15]. In order to analyze human motions, we first need to detect and track humans reliably. Various visual features are used for detection and tracking such as optical flow, color, depth, and their combinations [11], [17]. Color information is especially useful when a target wears clothing with distinctive colors [10]. Recently, *mean shift* tracking has been proposed [4] which calculates the most probable target position using a spatially-smooth similarity function. Perez et al. [12] extended this idea to particle filter-based tracking to cope with similar-colored nearby objects or short-term occlusions. The Multiple Person Tracker (MPT) presented in this paper will perform Kalman filter-based tracking and to cope with similar-colored nearby objects/persons and short-term occlusions, it will use a person model and the Joint Probabilistic Data Association (JPDA) theory.

To sense a certain space, former approaches made use of sonar [6] or laser-range scans [7], [13], [14] wherefrom 3D data can be gathered but they are superseded, too expensive or take too much processing time. Other 2D laser sensors only provide range data on a scanning plane. Stereo vision can be used, but needs two or more cameras and extra processing. The MPT described in this paper performs image-based tracking using a single camera.

One of the key features of this paper is how measurement uncertainties are determined and how measurements can be corrected based upon the comparison between the extracted measurement and the expected size of the tracked person, i.e. a person model. Section II-B explains the computation of the measurement uncertainties and section III will expound how measurements are corrected using a person model.

Further, the MPT has to deal with additional uncertainty caused by the uncertainty of the origin of the measurements and the correctness of associations of measurements to a person. This measurement-to-person correspondence has extensively been studied for person tracking and the surveillance community. A number of statistical data association techniques are compared with each other by Cox [5]. The MPT applies the JPDA technique [1], [2], [3], [6] and is discussed in more detail in section IV. Former approaches to track multiple persons already investigated the use of the JPDA filter. However, most of these approaches use laser sensed data and particle filters [7], [14]. There are not many works on applying the JPDA theory to image-based tracking.

II. STATE ESTIMATION

A. Kalman Filtering

The MPT tracks a specific person by calculating its state using the centers of gravity of the extracted regions. However, due to occlusion and lack of information about a person's true size and motion over time, the MPT extracts often disfigured regions, i.e. measurements perturbed with noise. In spite of this noise, the MPT needs to estimate a person's state. One of the most well-known and often-used tools for tracking is the Kalman filter [2], [8], [16].

The dynamics of a person are modeled by the equation

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{u} \tag{1}$$

where $\mathbf{x}_k$ is the 4-dimensional state vector (i.e. position and velocity in two directions) at frame k, $\mathbf{A}$ is the known

This research was conducted while the first author was with Osaka University. His stay was supported by the DeMaMech student exchange program runned by the project leaders Prof.Dr. Tetsuo Tomiyama and Prof.Dr. Hiromasa Suzuki for the European Union and Japan respectively. The MPT was designed, implemented and tested in the Active Intelligent Systems Laboratory of Prof.Dr. Jun Miura.

M.Liebens is with the Department of Mechanical Engineering, Katholieke Universiteit Leuven, 3000 Leuven, Belgium. maarten.liebens@student.kuleuven.be

T.Sakiyama is with the Department of Electrical Engineering and Information Science, Niihama National College of Technology, Niihama, Ehime 792-0871, Japan. sakiyama@ele.niihama-nct.ac.jp

J. Miura is with the Department of Mechanical Engineering, Osaka University, Suita, Osaka 565-0871, Japan. jun@mech.eng.osaka-u.ac.jp

transition matrix which maps the previous state on the current state based on the previous velocity and $\mathbf{u}$ is the process noise, assumed to be normally distributed with zero mean and variance $\mathbf{Q}$.

The measurement system is modeled as follows. If the measurement originates from the person, then

$$\mathbf{z}_k = \mathbf{H}\,\mathbf{x}_k + \mathbf{v}_k \tag{2}$$

where H is the known (2×4) matrix which maps the true state on the observed state and $\mathbf{v}_k$ represents the measurement noise, also assumed to be normally distributed with zero mean and variance $\mathbf{R}_k$.

The MPT's estimation of a person's state $\mathbf{x}_k$ at time k, given data up to time i, is denoted $\hat{\mathbf{x}}_{k|i}$. The error in the state estimation is represented by its variance matrix $\mathbf{P}_{k|i}$. In the absence of measurement origin uncertainty (which will be introduced in section IV) the discrete-time Kalman filter calculates a person's state estimation and its variance by correcting their predictions as follows

$$\begin{aligned}\hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{W}_k\tilde{\mathbf{z}}_k \\ &= \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + \mathbf{W}_k\tilde{\mathbf{z}}_k \qquad (3)\\ \mathbf{P}_{k|k} &= \mathbf{P}_{k|k-1} - \mathbf{W}_k\mathbf{S}_k\mathbf{W}_k^T \\ &= \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^T + \mathbf{Q} - \mathbf{W}_k\mathbf{S}_k\mathbf{W}_k^T \qquad (4)\end{aligned}$$

where $\mathbf{W}_k$ is the Kalman gain, $\tilde{\mathbf{z}}_k$ is the innovation vector and defined by the difference between the observed measurement and the predicted person's state

$$\tilde{\mathbf{z}}_k = \mathbf{z}_k - \mathbf{H}\,\hat{\mathbf{x}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1} \tag{5}$$

and where $\mathbf{S}_k$ is the variance matrix of the innovation

$$\mathbf{S}_k = \mathbf{A}\mathbf{P}_{k|k-1}\mathbf{A}^T + \mathbf{R}_k \tag{6}$$

B. Uncertainty of Measurement Using Person Model

Former approaches [6], [7], [14] often apply a constant measurement noise variance $\mathbf{R}$ into the Kalman filter. However, the MPT extracts measurements having their own size, characterized by height h and width w. Compared with the expected height $E[h]$ and expected width $E[w]$ of a specific person, i.e. person model, the (2×2) variance of the measurement's noise can be calculated from frame to frame during tracking

$$R_{(1,1)} = \alpha(E[w] - w)^2 + \sigma_X^2 \tag{7}$$

$$R_{(2,2)} = \alpha(E[h] - h)^2 + \sigma_Y^2 \tag{8}$$

$$R_{(1,2)} = R_{(2,1)} = 0 \tag{9}$$

where $\boldsymbol{\sigma}$ is the deviation from the mean of the expected size of the unoccluded tracked person. This previous variance is added in order to avoid obtaining variance $\mathbf{0}$ if the expected size matches the current measured size. On the other hand when occlusion occurs and the measured region suddenly enlarges or shrinks, the measurement's noise variance should increase, for instance when an object starts to obstruct the presence of the tracked person. This event is presented in Fig. 1(a). The observed width of the extracted region is smaller compared to the person model. It is not immediately clear

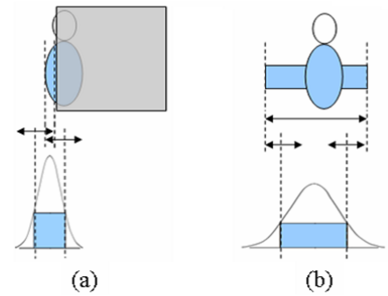


Fig. 1. Distributions of a person's state during ambiguous events. (a) Beginning of full occlusion. (b) Overlap of a person and a same colored object.

whether the left side corresponds to the real expected left side of a person or whether this left observed side is the boundary between the person and the occluding object. The same is considered for the right side. The two measures drawn in Fig. 1(a) depict the outer possible locations of the expected width. The position of the center of gravity becomes more uncertain, quantified by equation (7), and the MPT assumes that the probability of a person's position is uniformly distributed between the middles of the two outer positions of width. However, the measurement's noise is normally distributed. The coefficient α in (7) and (8) are meant to spread out the normal distribution of the measurement's noise in order to be in agreement with the uniform distribution like presented in Fig. 1(a).

On the other hand, a measurement can also be wider than the person model when an overlap occurs between a person and a same colored object. This event is presented in Fig. 1(b). The upper drawn measure represents the measured width of the tracked person. The center of gravity of this measurement is uncertain and this uncertainty is quantified by (7). The lowest two measures in Fig. 1 represent the two outer possible positions of the person's true width. The explanation about the distributions in the former case counts also here.

III. MEASUREMENT CORRECTION USING PERSON MODEL

A person walking in a heavy occluded space results often in disfigured and implausible measurements. Fortunately, measurements can be corrected in some cases by the MPT if there is *a priori* information about a person's expected height and width, i.e. a person model. The intention of this correction is to reduce a measurement's uncertainty.

The MPT can determine whether only the upper or lower part of a tracked person is visible or whether a same colored object in space suddenly enlarges the extracted region. The use of the person model for both cases will be explained together with Fig. 2. Fig. 2(a) represents the case when a person's lower part of the body is occluded. The height of the extracted measurement has changed gravely from the previous frames. When the MPT detects this change, it checks whether the upper pixel of the disfigured extracted region can be found within a threshold around the upper pixel

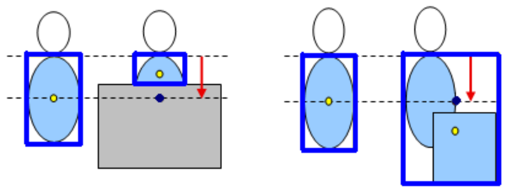


Fig. 2. Measurement correction. (a) Partial occlusion of the lower part of the tracked person. The bounding boxes represent the demarcation of the extracted measurements, the bright dots are the centers of gravity of the measurements and the dark dot (accompanied by the arrow) is the corrected measurement. (b) Overlap of a person and a same colored object.

of the region extracted in the frame before the occurrence of partial occlusion. The same is done for the lowest pixel to find out whether the upper part is occluded. When the MPT concludes that the lower part is occluded, the Y coordinate of the center of gravity of the measurement is corrected by positioning the center of gravity h_{exp} below the upper pixel.

According to this course of reasoning, the measurement, obtained from a frame where a tracked person and a same colored object are in overlap, can be corrected the same way and this correction is presented in Fig. 2(b).

IV. DATA ASSOCIATION

A. Problem Formulation

The appearance of multiple measurements, due to occlusion or simply due to the presence of multiple persons, complicates the process of tracking. The analysis for this appearance requires knowledge of measurement-to-person correspondence. That is, what is the possibility that the measurement currently originates from this person? Which measurements are used to correct the predicted state of a person in order to estimate a person's position? Any measurement might have originated from any person. In practice, some measurements are more likely to originate from one track than another. A distance measure is therefore needed that quantifies this likelihood. The smaller the distance between a measurement and its predicted value, the more likely it is to have originated from it. From the Kalman filter theory, it is known that the measurement is normally distributed about its predicted value [16]. In that case, a common distance measure is the Mahalanobis distance [9]. The MPT associates only the measurements to a person with distance $\sqrt{\tilde{\mathbf{z}}_k^T \mathbf{S}_k^{-1} \tilde{\mathbf{z}}_k} \leq \gamma$. The Mahalanobis distance quantifies the likelihood that a measurement originates from a specific person and can be considered to be a generalization of the Euclidean distance which also takes the relative uncertainties in innovations into account.

The m validated measurements extracted from the frame at time k are denoted as $\mathbf{z}_j$:

$$Z^k = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\} \bigcup Z^{k-1} \tag{10}$$

and the corresponding individual innovations at time k for person t are, for $j = 1, \ldots, m$:

$$\tilde{\mathbf{z}}_j^t \triangleq \mathbf{z}_j - \hat{\mathbf{z}}^t \tag{11}$$

where $\hat{\mathbf{z}}^t$ is the predicted measurement for person t. The time subscript k will be suppressed for convenience from all variables except from $\mathbf{P}$ and Z.

The main purpose of the JPDA filter is to combine all these measurements' innovations weighted with their *association probability* in order to correct the prediction of person t in the Kalman filter with a total innovation $\tilde{\mathbf{z}}^t$:

$$\tilde{\mathbf{z}}^t = \sum_{j=1}^{m} \beta_j^t \, \tilde{\mathbf{z}}_j{}^t \tag{12}$$

where β_j^t is the *association probability* that measurement j originated from person t.

B. Joint Event Probabilities

The key to the JPDA algorithm and to the calculation of the *association probabilities* β_j^t is the evaluation of the following feasible joint events

$$\chi = \bigcap_{j=1}^{m} \chi_{j\,t} \tag{13}$$

where $\chi_{j\,t}$ is the event that measurement j originated from person t. The feasible events are those joint events in which only one measurement originates from only one person and vice versa [1], [2], [6]. The probability of the joint event conditioned on the measurements up to time k is denoted as $P(\chi|Z^k)$. After obtaining these feasible joint events χ and their probabilities, the *association probability* β_j^t is the probability of the event where measurement j is associated to person t conditioned on all measurements up to the present time k:

$$\beta_j^t \triangleq P(\chi_{j\,t}|Z^k) \tag{14}$$

The extra computational burden resulting from the consideration of all the events (also those events with negligible probability or even probability zero), in order to calculate the *association probability* β_j^t, can be avoided with suitable logic limiting the probability calculations to only those events involving validated measurements. This logic, which has actually a negligible effect on the numerical results but not on the computational burden, is presented in [7].

The probability of a joint event conditioned on all measurements up to the present time k can be presented as:

$$\begin{aligned} P(\chi\,|\,Z^k) &= P(\chi\,|\,\mathbf{z}_1, \ldots, \mathbf{z}_m, Z^{k-1}) && (15)\\ &= P(\chi\,|\,\mathbf{z}_1, \ldots, \mathbf{z}_m, X^k) && (16)\\ &= \frac{1}{c}\, p(\mathbf{z}_1, \ldots, \mathbf{z}_m\,|\,\chi, X^k)\; P(\chi\,|\,X^k) && (17) \end{aligned}$$

The measurements Z^k in (15) can be split up in the current detected measurements $\mathbf{z}_1, \ldots, \mathbf{z}_m$ and the measurements Z^{k-1} obtained before time k. However, to consider the condition of all measurements obtained before time k is complex and laborious. To overcome this problem, the distributions of the predictions of the persons' states X^k are introduced in (16). All previous measurements are indirectly encapsulated in these predictions. On the other hand, the prediction and the estimation problem are assumed to be Markovian. The state of a person depends only on the measurements obtained at time k and the state at time $k-1$.

Using Bayes' rule, the probability of a joint event χ can be presented as in (17) where $P(\chi|X^k)$ is the prior probability conditioned on the distribution of the current predicted states X^k and the term $1/c$ is the prior probability of the measurements conditioned on the past data, and acts as a normalizing constant.

The probability of each measurement-to-person association is derived in [6] and defined as follows:

$$\begin{aligned} P(\chi|Z^k) &= \frac{1}{c}\prod_{j:\tau_j=1}\frac{1}{\sqrt{2\pi|\mathbf{S}^t|}}\exp\Big(-\frac{(\tilde{\mathbf{z}}_j^t)^T(\mathbf{S}^t)^{-1}(\tilde{\mathbf{z}}_j^t)}{2}\Big) \\ & \qquad \prod_{t:\delta_t=1}P_D\prod_{t:\delta_t=0}(1-P_D) \end{aligned} \tag{18}$$

where P_D is the probability of detecting a person in the frame and related to the observed space, δ_t is the target detection indicator and equals 1 if any measurement is associated with target t in event χ (i.e. whether target t is detected) and τ_j is the measurement association indicator and equals 1 if measurement j is associated with any target in event χ. The first product term in (18) is the probability density of a measurement $\mathbf{z}_j$ if it is associated to a person t. This probability is normally distributed about its predicted value $\hat{\mathbf{z}}^t$ with variance $\mathbf{S}^t$.

C. Association Probabilities

The probability β_j^t that measurement j, from the collection of all m validated measurements, belongs to person t may now be obtained over all feasible events χ for which this condition is true:

$$\beta_j^t = \sum_{\chi} P(\chi\,|\,Z^k)\omega_{j\,t}(\chi) \tag{19}$$

where $\omega_{j\,t}(\chi)$ equals unity if measurement j belongs to person t in event χ and zero otherwise. These probabilities are used to calculate the combined innovation (12) for every person.

V. MODIFICATION OF STATE ESTIMATION

The Kalman filter as described in section II needs to be modified to account for the possibility of incorrect measurement-to-person associations. Therefore, the estimator has to assimilate the *association probabilities*. This way, the estimator becomes nonlinear because the estimation and its variance are nonlinear functions of the measurements.

A. State estimation

The state estimation of a person is modified in the following way:

$$\hat{\mathbf{x}}_{k|k}^t = \sum_{j=1}^{m}\hat{\mathbf{x}}_{k|k,j}^t\,\beta_j^t \;=\; \hat{\mathbf{x}}_{k|k-1}^t + \mathbf{K}^t\,\tilde{\mathbf{y}}^t \tag{20}$$

where $\hat{\mathbf{x}}_{k|k,j}^t$ is the estimation using a single measurement out of m validated measurements and weighted by the *association probability* β_j^t. Also the Kalman gain $\mathbf{K}^t$ depends on the measurements because every measurement j has its own uncertainty $\mathbf{R}_j$. Subsequently, the total Kalman gain is, similar to the estimation of a person's state, composed out of partial Kalman gains originating from individual validated measurements and weighted by the *association probability* of the measurement:

$$\begin{aligned} \mathbf{K}^t &= \sum_{j=1}^{m}\mathbf{K}_j^t\,\beta_j^t && (21)\\ &= \sum_{j=1}^{m}\beta_j^t\,\mathbf{P}_{k|k-1}^t\,\mathbf{H}^T(\mathbf{H}\mathbf{P}_{k|k-1}^t\mathbf{H}^T+\mathbf{R}_j) && (22) \end{aligned}$$

B. Uncertainty of state estimation

The variance associated with the above estimation (20) is derived in [3] and can be represented as follows:

$$\begin{aligned} \mathbf{P}_{k|k} &= \int(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k})(\mathbf{x}_k-\hat{\mathbf{x}}_{k|k})^T \\ &\qquad p(\mathbf{x}_k|\mathbf{z}_1,\ldots,\mathbf{z}_m,\mathbf{X}_k)\,d\mathbf{x}_k && (23)\\ &= \beta_0^t\;\mathbf{P}_{k|k-1} + \sum_{j=1}^{m}\beta_j^t\;\mathbf{P}_{k|k,j} + \mathbf{P}_k && (24) \end{aligned}$$

The uncertainty consists of three parts. The first term $\beta_0^t\;\mathbf{P}_{k|k-1}$ covers the contribution of the possible event that no measurements are associated to the tracked person where $\beta_0^t = 1-\sum_{j=1}^{m}\beta_j^t$. The second term $\sum_{j=1}^{m}\beta_j^t\;\mathbf{P}_{k|k,j}$ represents the share of all validated measurements to the total uncertainty. The last term is an additional share reflecting how measurements are located from one another where the previous terms reflect how measurements are located from the predicted position of a tracked person. The precise formula for this last term is derived and given in [3], [6].

VI. EXPERIMENTAL RESULTS

The MPT is implemented and successfully applied to several image sequences. To quantify the computation time, in the case where there are four persons tracked, the average processing time per frame is about 0.12s when running on a 1.5GHz CPU. This shows that real-time tracking is possible with a rate of 8 frames per second. However, the processing time depends on the number of persons that are tracked, the number of extracted measurements, on the specifications of the CPU where the MPT is running on, etc.

A. Color-Based Person Detection

The MPT uses a color model from the upper part of a person's body as a feature in order to detect this person and to extract measurements. When a person enters the observed space and its presence isn't associated to an existing track, a new track is initiated and a corresponding color model is constructed based on the HSV color space. Charged with the color models from all tracked persons, the MPT will extract regions from every frame of the image sequence and after noise removal, regions are labeled with a number. Regions become measurements when they are labeled, surrounded by a bounding box and characterized with a center of gravity.

(a) (b)

Fig. 3. Comparison of *iterative Kalman filter* and *extended JPDAF Kalman filter*. (a) Use of *iterative Kalman filter*. The ellipses show the 99% probability area where the estimated position is located. (b) Use of *extended JPDA Kalman filter*. The ellipse is smaller. The estimated position is more certain when using the *extended JPDA Kalman filter*.

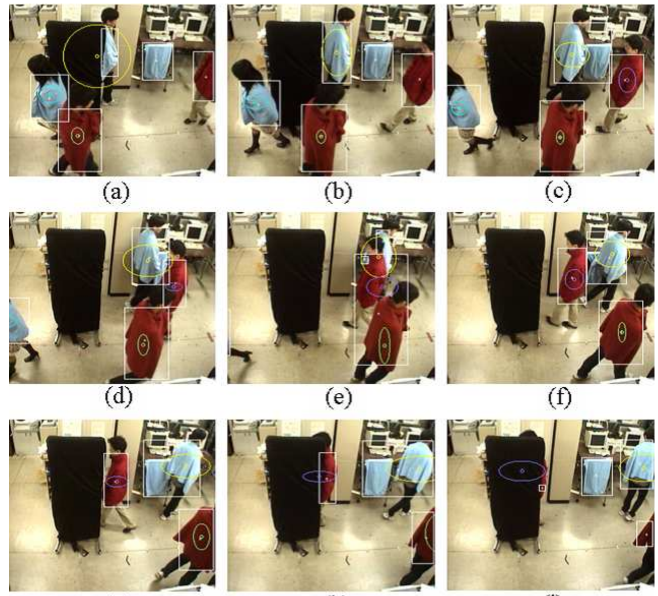
B. Effect of JPDA Filter

The MPT performs for every initialized person a correction of the prediction using a measurement. When multiple measurements are feasible, two options arise for this correction. The first option is the use of an *iterative Kalman filter*. This filter corrects the prediction with the first measurement and this correction will be corrected with the next measurement, etc. This way, all measurements will have the same share in correcting the prediction. This results in a more uncertain estimation when some measurements are more likely to originate from a current tracked person. The second option is the *extended JPDA Kalman filter*. Here the prediction is corrected only once using a combination of all validated measurements weighted with their share, i.e. their *association probability*.

The comparison of both options is presented in Fig. 3. In this example the hindmost person is occluded by the person in the front resulting in two extracted measurements. These two measurements are handled by an *iterative Kalman filter* in Fig. 3(a) and handled by the Kalman filter extended by the JPDA theory in Fig. 3(b). There is a neglected effect on the estimated position but a noticeable effect on its uncertainty which is smaller when the *extended JPDA Kalman filter* is used.

So it is clear that the MPT will solve the association problem with the JPDA theory. From every frame of the image sequence and for every person that is tracked, the MPT performs a number of loops and recursive algorithms in order to link measurements to different tracks and to calculate a measurement's share in the state estimation of the tracked person.

In Fig. 4, four persons are tracked. When two same colored persons approach each other, measurements originating from both persons will be used for the state estimation of one person. A measurement's share will be calculated by the MPT based on the innovation, the difference between measured size and the person model, etc. Also several occlusions and exceptional cases occur in Fig. 4. Fig. 4(a) shows a person appearing after full occlusion by an object in the observed space. During its occlusion no measurements from this person were gathered although the measurement from the same colored person who was walking in front of the object can be associated to this tracked person. However, the



(a) (b) (c)
(d) (e) (f)
(g) (h) (i)

Fig. 4. Representation of a sequence where four persons are tracked when walking at random and with several occlusions.

MPT calculates a much lower *association probability* for this latter measurement than the probability that no measurements are detected from this person. This none detection event is pushed through during the evaluation and results in a marginal correction of the prediction. According to (24) the first term is from considerably more importance during the calculation of the estimation's uncertainty. Subsequently, this uncertainty and its representation as an ellipse will increase during a full occlusion. From the moment when a representative measurement is extracted according to the person model and the prediction, the estimation's uncertainty will decrease again. However in Fig. 4(c) occurs an overlap with a same colored object. The measurement suddenly expands in its width and the estimation's uncertainty increases in this direction. In Fig. 4(e) the measurements of the two red colored persons are merged together. The height of this measurement differs significantly from both person models. However, this measurement can be corrected for the upper person as described in section III which results in a better state estimation and lower uncertainty. The measurement can not be corrected for the lower person because the measurement doesn't satisfy the requirements for using of the person model. After this crossing the two persons are tracked correctly.

C. Effect of Using Person Model

The person model of a tracked person contains information about its expected height and width. Assuming a person isn't occluded when he enters the observed space, a mean value and deviation of it's height and width is calculated. During tracking, the measurement's size is compared with the person model of the tracked person in order to detect occlusion and to calculate a measurement's variance $\mathbf{R}_k$.

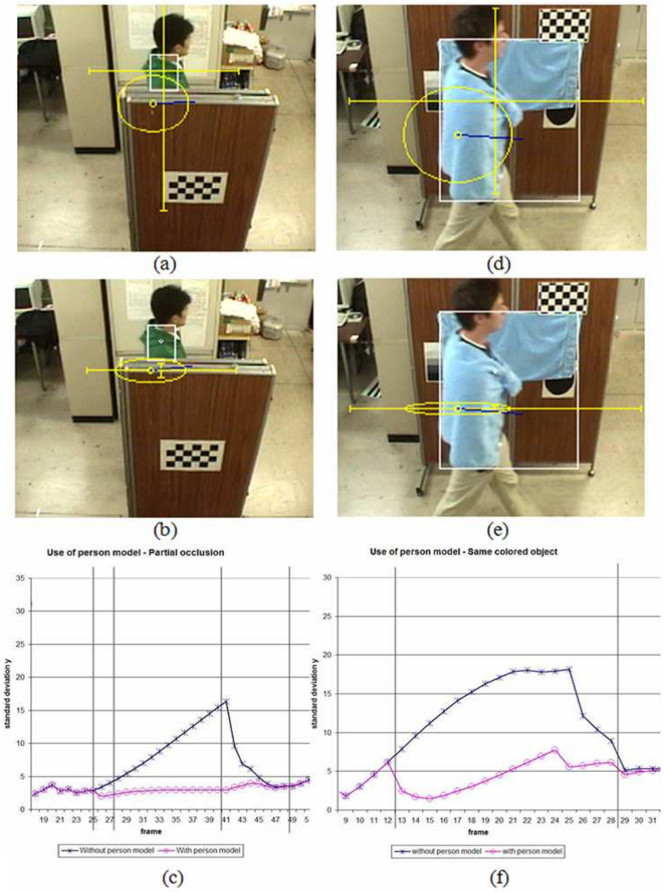


Fig. 5. Effect on measurement and estimation when using a person model. (a) Partial occlusion before correction of the measurement. The bounding box represents the observed measurement, the cross represents the measurement's uncertainty in both directions. (b) Partial occlusion after correction using the person model. (c) Quantitative representation of the correction. The X axis presents time and the Y axis the Y component of the measurement's noise deviation. The crossed line represents tracking without using the person model. The line marked with the circles represents the deviation when using the person model. The first vertical line shows when the tracked person starts to walk behind the scene. The last line shows when he is completely visible again for the camera. (d), (e) and (f) Analogue but in case of an overlap with a same colored object.

Fig. 5 represents the use of the person model to correct a measurement for the case of partially occluded persons and enlarged regions because of a same colored object in the background. Section III described how the MPT handles these two cases.

VII. CONCLUSION

The MPT is in a state where it can perform fail-safe tracking in a populated and heavy occluded space where every person is tracked with a Kalman filter. The Kalman filter has been shown to provide highly efficient state estimates for image-based tracking, although assuming unimodal Gaussian distributions. Furthermore, the Kalman filter is adjusted to image-based tracking using a person model for a more precise calculation of the uncertainties. But before estimating a person's state, the measurements are corrected by the MPT using also this person model in order to reduce the uncertainties. Next, the MPT decides which measurements are allowed to compete for correcting the estimation's prediction and in what extent by calculating its *association probability*. This extension of the Kalman filter is realized by the JPDA theory.

Continuing research in this topic includes tracking of non-continuous motion (currently possible for a not so sudden change of motion and depending on the estimation uncertainties), the use of more complex models for state estimation and the implementation of the MPT on a mobile camera/robot. The ultimate objective is the application of the MPT in a more complicated situation where many persons are distinguished and tracked. The computational cost for every state estimation will be a drawback in such a situation. Another consequence is that other features than only the color of the upper part of the body will be necessary for detection and association of measurements, for instance color of skin and hair, size, etc.

REFERENCES

[1] Y. Bar-Shalom, "Tracking Methods in a Multitarget Environment", *IEEE Transactions on Automatic Control*, vol. AC-23, No.4, August 1978, pp. 618-626.

[2] Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*, Mathematics in Science and Engineering, Academic Press, 1998.

[3] Y. Bar-Shalom and A. Jaffer, "Adaptive Nonlinear Filtering for Tracking with Measurements of Uncertain Origin", *Proc. of the 1972 IEEE Converence on Decision and Control*, December 1972, pp. 243-247.

[4] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-Based Object Tracking", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 5, 2003, pp. 594-607.

[5] I.J. Cox, "A review of Statistical Data Association Techniques for Motion Correspondence", *International Journal of Computer Vision*, 10:1, 1993, pp. 53-66.

[6] T.E. Fortmann, Y. Bar-Shalom and M. Scheffe, "Sonar Tracking of Multiple Targets Using Joint Probabilistic Data Association", *IEEE Journal of Oceanic Engineering*, vol. OE-8, No. 3, July 1983, pp. 173-184.

[7] O. Frank, *Multiple Target Tracking*, master thesis, Swiss Federal Institute of Technology Zurich (ETH), Switzerland, 2003.

[8] R.E. Kalman, "A New Approach to Linear Filtering and Prediction problems", *J. Basic Eng.*, vol. 82-D, 1960, pp. 35-45.

[9] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, North-Holland; 1996.

[10] Y. Ohno, J. Miura, and Y. Shirai, "Tracking Players and Estimation of the 3D Position of a Ball in Soccer Games", *Proc. 15th Int. Conf. on Pattern Recognition*, 2000, pp. 145-148.

[11] R. Okada, Y. Shirai, J. Miura, "Object Tracking Based on Optical Flow and Disparity", *Proc. IEEE/SICE/RSJ Int. Conf. on Multisensor Fusion and Integration in Intelligent Systems*, 1996, pp. 565-571.

[12] P. Perez, C. Hue, J. Vermaak, and M. Cangnet, "Color-Based Probabilistic Tracking", *Proceedings of 7th European Conf. on Computer Vision*, 2002, pp. 661-675.

[13] D. Schultz, W. Burgard, D. Fox, and A.B. Cremers, "Tracking multiple moving targets with a mobile robot using particle filters and statistical data association", *IEEE International Conference on Robotics and Automation*, 2001, pp. 1665-1670.

[14] D. Schulz, W. Burgard, D. Fox, and A.B. Cremers, "People Tracking with a Mobile Robot Using Sample-based Joint Probabilistic Data Association Filters", *International Journal of Robotics Research (IJRR)*, vol. 22, No. 2, 2003, pp. 99-116.

[15] L. Wang, W. Hu,and T. Tan, "Recent Developments in Human Motion Analysis", *Pattern Recognition*, Vol. 36, 2003, pp. 585-601.

[16] G. Welch and G. Bishop, *An Introduction to the Kalman Filter*, Technical Report TR 95-041, Computer Science, UNC Chapel Hill, 1995.

[17] T. Yamane, Y. Shirai, J. Miura, "Person Tracking Using Optical Flow and Uniform Brightness Regions", *Proc. IEEE Int. Conf. on Robotics and Automation*, 1998, pp. 3267-3273.