Robust Tracking of Spatial Related Components *

Thomas Mauthner, Michael Donoser and Horst Bischof
Institute for Computer Graphics and Vision
Graz University of Technology, Austria
{mauthner,donoser,bischof}@icg.tugraz.at

Abstract

This paper introduces a hierarchical approach for multi-component tracking, where the object-to-be-tracked is modeled as a group of spatial related parts. We propose to use a robust particle filtering framework for tracking the individual components and outline how the spatial coherency between the parts can be efficiently integrated by analyzing a two-level hierarchy of particle filters. Including spatial information allows to handle common tracking problems like occlusions, clutter or blur. Furthermore, the dynamic calculation of particle set uncertainties allows a dynamic adaption of stiffness values for the spatial model to e. g. force occluded parts to stay in spatial relation. The experimental section proves the robustness of the proposed tracker on challenging sequences of the VIVID-PETS database.

1 Introduction

Visual tracking is an important task in many computer vision applications like visual surveillance, human computer interaction, traffic monitoring or sports analysis. Many different methods have been proposed for visual tracking but recently especially probabilistic methods like particle filtering [1] were of high interest. Particle filtering has been widely applied to tracking problems where it is also known as Condensation algorithm. The particle filter can be interpreted as a probabilistic search algorithm where a set of particles, each representing one possible state, models the posterior probability representing the current knowledge about the object state.

Particle filters applied for tracking are typically based on color, contours or other appearance models as underlying features. Because for every hypothesized state all features have to be calculated the computational complexity strongly depends on the number of particles. Unfortunately the required number of particles increases exponentially with the dimensionality of the state space. Therefore, various authors describe ways to decrease the computational complexity by using hierarchical approaches. In [14] and [6] the complexity of features is increased in every level, while the number of particles is decreased. Efficient methods for partitioning the state space are shown in [13, 12], nevertheless a prior knowledge of possible states has to be given.

Many approaches also tried to simplify the problem by using part based modeling. For example Brandao et al. [3] introduced a subspace particle filter for hand tracking including a graph-based representation of the hand. The graph (of the hand) has to be given a-priori. A part-based method for tracking loose-limbed people in 3D over multiple views is presented in [11], which makes use of a bottom-up part detectors to estimate possible part locations in each frame. The recent work of Schindler and Dellaert [10] presents a method for efficiently tracking objects represented as constellation of parts. Rao-Blackwellization is used to integrate out continuous parameters. This allows to maintain multiple hypotheses for the object pose without the need to sample in the high dimensional space, but the constellation of parts has to stay constant.

We formulate the problem of visual tracking as tracking a group of spatial related components. These components can either be understood as parts of a single object or as several objects sharing a common behavior e. g. a group of cars. The main difference to the approaches mentioned above, is that the spatial model between components has not to be trained before, and is updated during the tracking process. We propose a two-level hierarchy of particle filters to efficiently model the state space as is illustrated in Figure 1. In the first level, independent appearance based particle filters are used per component. This can be performed very efficient using appearance model computation based on integral structures e.g. [9, 4]. Depending on the confidence value of the particle set, different numbers of particles are passed on. A second level combines the results of the individual components in a spring-mas model, considering their spatial arrangement. The stiffness between different components is set proportional to the uncertainty values of their

*This work was supported by the Austrian Science Fund (FWF P18600)

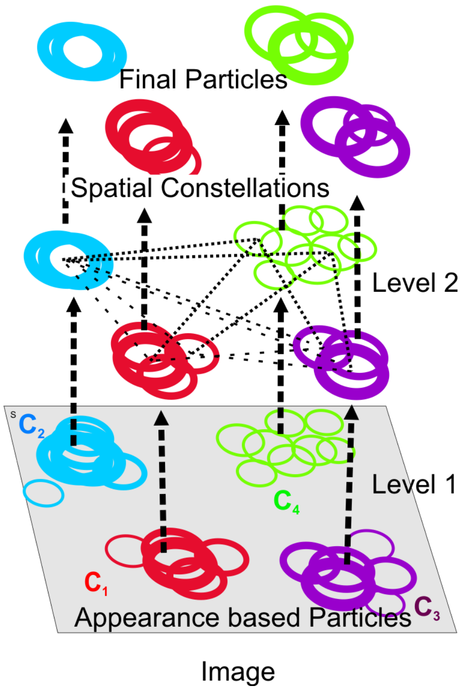

Figure 1. Overview of the proposed approach. Level 1: Appearance based particle filtering – circle size corresponds to particle weights. Level 2: After resampling a number of particles from level 1 is passed on to the second spatial coherency level. Connections to unstable particle sets are stiffer, illustrated by differently dashed lines.

particle sets. By using such a robust probabilistic tracking framework and integrating a spatial coherence model we are able to handle problems related to clutter, occlusions and dynamic object changes. For example it is possible to update non-occluded components while occluded parts are forced to stay in relation to the other parts.

The rest of the paper is organized as follows. Section 2 presents our tracking method in detail. Section 3 demonstrates the quality of the proposed tracker on challenging video sequences of VIVID-PETS 2005 database. Finally Section 4 draws some conclusions.

2 Tracking framework

In our tracking framework the object-to-be-tracked O is modeled as a constellation of N components C_n, or more generell a group of objects. The tracker has to be initialized by providing these individual components in the first frame. After initialization we propose to use a two level hierarchy of particle filters for tracking such constellation models through an image sequence. The first level models the posterior probability of each individual component state and the second level analyzes the spatial coherency of all components.

2.1 Review of particle filter

Particle filtering for tracking [1] provides a probabilistic framework, which maintains multiple hypotheses of the current object state and has proven to yield impressively robust tracking results. The probability distribution of the hidden target state $\mathbf{x}_t$ of the tracked object at time step t is estimated using a set of N_P weighted particles $S_t = \{x_t^i, w_t^i\}$ with $i = 1...N_P$ at time-step t, and associated measurements z_t^i. Each particle x_t^i simulates the real hidden state of the the object. Using the dynamic model $p(x_t^i|x_{t-1}^i)$ and the observation likelihood $p(z_t^i|x_t^i)$, the posterior distribution $p(\mathbf{x}_t|\mathbf{z}_t)$ is approximated by this finite set of particles.

2.2 Information exchange between levels

As shown in [6] several observers can be combined by using multiple stages of importance sampling. Every stage k uses the weighted particle set of its previous stage $S_{k-1,t} = \{x_{k-1,t}^i, w_{k-1,t}^i\}$ as a proposal distribution to sample a new set $S_{k,t} = \{x_{k,t}^i, 1/N_{P,k}\}$ by re-sampling from $S_{k-1,t}$. In our two-level hierarchy the first level observation likelihood $p(z_{1,t}|x_{1,t})$ is calculated by analyzing appearance information which is efficiently estimated based on integral images as e. g. shown in [9, 4]. Therefore, a high number of particles can be used and only important ones are passed to the second level, see Figure 1, which analyzes the spatial coherence as it is shown in Section 2.3.

In the case of background clutter, occlusion or ambiguities the distribution of the particles can become unsubstantial, which leads to drifting, inaccurate tracking results or lost objects. Therefore, it is an important task to measure the quality of a given weighted particle set $\{x_t^i, w_t^i\}$ to recognize when tracking of objects fail. As shown by [7, 2] it is possible to calculate a particle set uncertainty value U_C to measure tracking quality.

We use the particle set uncertainties to dynamically regularize the influence of the spatial coherence. The estimated uncertainty contributes to the final constellation likelihood and also to the spring-mass analysis by setting the stiffness value Ψ_{ij} of the model proportional to the uncertainty of the components. In such a way, that springs to a component with high uncertainty are set dynamically to a higher stiffness value. To enhance the influence of the uncertainty in the spatial model, a sigmoid function was used. This stabilizes the uncertainty behavior on the upper and lower bound and forces clearer decisions.

2.3 Incorporation of spatial coherency

The second particle filter level models the posterior probability of the state of the entire constellation model, analyzing the spatial arrangement of the individual components by a physically-based mass-spring model [5]. A mass-spring system consists of N mass points $\vec{p}_n$ and a set $\mathcal{S}$ of predefined springs $\vec{s}_{ij}$ each linking two mass points $\vec{p}_i$ and $\vec{p}_j$ having a fixed stiffness value Ψ_{ij}. Adapting the stiffness values allows to control the sensitivity of the model to changes of the spatial configuration. In our case, the mass points are the center points of the individual components of our object and we define a spring between all possible combinations of these points. Thus, for N components we have $N(N-1)/2$ springs. We use the deformation energy of the system to define the particle likelihood of the second particle filter level.

The described two-level hierarchy also ensures that a low number of particles is sufficient to sample the high-dimensional state-space. The first level provides a set of P hypothesized states $\{x_t^1, \ldots, x_t^P\}$ for every component, the corresponding observation likelihoods $p\left(z_t^i | x_t^i\right)$ and the particle set uncertainty U_C. We propose to use the likelihoods as importance density for randomly drawing the hypothesized first level states of each of the components to hypothesize entire constellation models, as depicted in Figure 1. For components having a low uncertainty, i. e. components where the corresponding first level feature analysis provides distinct results, only the highest likelihood particles should be used. In contrast, for components having high uncertainty the particles should be drawn almost independent of the first level likelihoods. Therefore, only a fixed, uncertainty dependent number of particles is considered in the constellation sampling process. This number is set reciprocally proportional to the uncertainty and always the particles with the highest likelihood are chosen. In such a way only potentially good constellations are considered in the second level sampling process.

3 Experiments

This section shows examples for tracking objects sharing common behavior as e. g. cars on a street. The shown video data is from the VIVID-PETS 2005[1] database one of the most challenging due to occlusions, camera motion, zooming and blur. In all experiments the appearance likelihoods of the particles in level 1 are calculated using covariance features based on a feature vector and update function as described in [8]. Each component uses 200 particles for the first level, whereas in the second spatial coherency level 1000 different constellations are analyzed. Please note, that the proposed method can also be used for non-rigid object tracking, as well as with different features in the first level.

Figure 2. Tracking multiple object sharing common behavior (Egtest04).

Figure 2 shows tracking results under several occlusions during the Egtest04 sequence, which is one of the most challenging sequences in VIVID-PETS. All three objects start with low uncertainty measurements until the first car gets occluded around frame 650. Successful handling of occlusion in several situations in addition with the adaption of the spatial model is illustrated by the sequence. Incorporation of spatial relations using our approach guides all objects through the occlusions. On reappearance, particles in the first level are resampled again on the object, leading to a small uncertainty. The corresponding uncertainty values U_{Cp} of the three objects are plotted in Figure 3. Please note that during several frames even two objects are occluded.

To compare our results to a state-of-the-art tracker we also applied a particle filter embedded covariance tracker as proposed by Porikli et al. [8] to the test sequences. Figure 4 compares our robust tracking results to the results of the covariance tracker, which fails to handle the blurring effects, because appearance changes too much. Figure 5 demonstrates that the dynamic adaption of the spatial model stiffness based on particle set uncertainties also allows to handle dynamic changes of the spatial model.

4 Conclusion and outlook

This paper introduces a two-level hierarchy of particle filters for tracking constellations of individual components. We show how the appearance based particle sets of the first level are fed to a second level which analyzes the spatial coherency by a mass-spring system. We further described that the two level design allows to reduce the number of re-

[1] http://www.vividevaluation.ri.cmu.edu/datasets/datasets.html

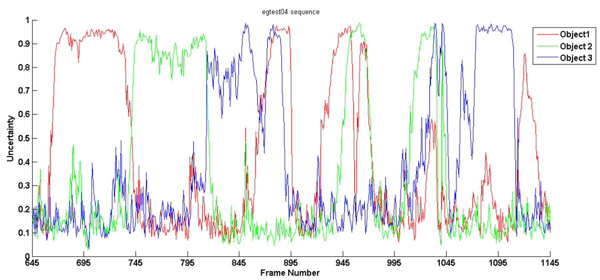

Figure 3. Uncertainty measurements used to dynamically set the spatial stiffness between objects during Egtest04 sequence.

Figure 4. Including spatial coherence information allows to prevent that the proposed tracker (red) drifts, while a state-of-the-art particle filter embedded covariance tracker (green) fails because of the appearance change by blurring.

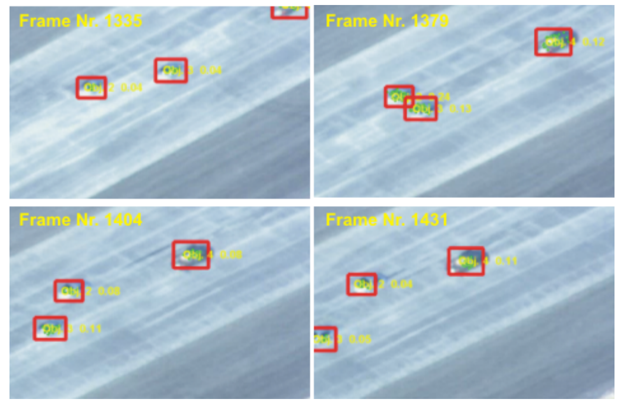

Figure 5. Dynamic adaption of spatial relations during tracking based on analyzing the uncertainty of each particle set.

quired particles to sample the high-dimensional state space. A dynamic adaption of the spatial model stiffness allows to handle occlusion and clutter problems by forcing e. g. occluded components to stay in relation to the other parts. Experimental evaluation proved that robust tracking results are achieved on challenging data sets.

References

[1] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.

[2] V. Badrinarayanan, P. Pérez, F. Le Clerc, and L. Oisel. Probabilistic color and adaptive multi-feature tracking with dynamically switched priority between cues. In *Proc. ICCV*, 2007.

[3] B. Brandao, J. Wainer, and S. Goldenstein. Subspace hierarchical particle filter. In *SIBGRAPI*, 2006.

[4] M. Donoser and H. Bischof. ROI-SEG: Unsupervised color segmentation by combining differently focused sub results. In *Proc. CVPR*, 2007.

[5] M. A. Fischler and R. A. Elschlager. The representation and matching of pictorial structures. *IEEE Transaction on Computers*, 22(1):67–92, 1973.

[6] Y. Li, H. Ai, T. Yamashita, S. Lao, and M. Kawade. Tracking in low frame rate video: A cascade particle filter with discriminative observers of different lifespans. In *Proc. CVPR*, 2007.

[7] E. Maggio, F. Smerladi, and A. Cavallaro. Combining colour and orientation for adaptive particle filter-based tracking. In *Proc. BMVC*, 2005.

[8] F. Porikli, O. Tuzel., and P. Meer. Covariance tracking using model update based on lie algebra. In *Proc. ECCV*, 2006.

[9] F. M. Porikli. Integral histogram: A fast way to extract histograms in cartesian spaces. In *Proc. CVPR*, 2005.

[10] G. Schindler and F. Dellaert. A rao-blackwellized parts-constellation tracker. In *Workshop on Dynamical Vision (ICCV)*, 2005.

[11] L. Sigal, S. Bhatia, S. Roth, M. J. Black, and M. Isard. Tracking loose-limbed people. In *Proc. CVPR*, 2004.

[12] B. Stenger, A. Thayananthan, P. H. S. Torr, and R. Cipolla. Model-based hand tracking using a hierarchical bayesian filter. *PAMI*, 28(9):1372–1384,, September 2006.

[13] V. Verma, S. Thrun, and R. G. Simmons. Variable resolution particle filter. In *IJCAI*, 2003.

[14] C. Yang, R. Duraiswami, and L. S. Davis. Fast multiple object tracking via a hierarchical particle filter. In *Proc. ICCV*, 2005.