



Factored Interval Particle Filtering for Gait Analysis

Jamal Saboune, Cédric Rose, François Charpillet

► To cite this version:

Jamal Saboune, Cédric Rose, François Charpillet. Factored Interval Particle Filtering for Gait Analysis. 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - IEEE EMBC 2007, Aug 2007, Lyon, France. 4 p. inria-00170996

HAL Id: inria-00170996

<https://inria.hal.science/inria-00170996>

Submitted on 11 Sep 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Factored Interval Particle Filtering for Gait Analysis

Jamal Saboune, Cédric Rose, and François Charpillet

Abstract—Commercial gait analysis systems rely on wearable sensors. The goal of this study is to develop a low cost marker less human motion capture tool. Our method is based on the estimation of 3d movements using video streams and the projection of a 3d human body model. Dynamic parameters only depend on human body movement constraints. No trained gait model is used which makes this approach generic. The 3d model is characterized by the angular positions of its articulations. The kinematic chain structure allows to factor the state vector representing the configuration of the model. We use a dynamic bayesian network and a modified particle filtering algorithm to estimate the most likely state configuration given an observation sequence. The modified algorithm takes advantage of the factorization of the state vector for efficiently weighting and resampling the particles.

I. HUMAN MOTION CAPTURE

Gait analysis is becoming a very useful tool for many applications in medicine. It is used for biometric identification [1], rehabilitation programs and dynamic equilibrium evaluation among others. Current commercial gait analysis systems are essentially based on wearable sensors and instrumented walkways. Electromyography signals, extracted using electrodes that are fixed on leg muscles, can provide useful information on the muscular activity while walking and thus provide a diagnosis of gait disorders [2]. Using footswitch equipped soles [3] and FSR (Force Sensing Resistor) sensors attached the foot, different phases of movement (heel strike, stance phase, heel and toe off) can be detected while normal walking. The GAITRite system [4] uses a pressure sensors carpet to measure many gait parameters in order to determine dynamic balance and predict fall risk. Electro-goniometers are widely used to evaluate the amplitude of flexion and extension movements of body articulations (ankles, knees etc.); this information can then be useful for estimating the ability of a person to maintain its balance while moving [5]. Tri-axial accelerometers and gyroscopes placed on the body provide information on daily activity [6][7] and individual acceleration profiles [8]. Using these acceleration signals Menz. et al. [9] extract spatio-temporal gait parameters. All these techniques use expensive wearable sensors that can be restraining and provide each a diagnosis of a single aspect of gait. Knowing this, gait analysis using video feeds captured by commercial conventional cameras seems to be a passive and low cost solution.

An in-depth inclusive gait analysis requires the estimation of the 3d positions of several key points of the body while walking. These positions would then be used to evaluate

all the spatio-temporal parameters, accelerations and body articulations movements. In order to obtain this information, a 3d human motion capture system has to be developed. Marker-based systems [10] have been widely used for years with applications found in biometrics. In typical systems, a number of reflective markers are attached to several key points of the patient's body and then captured by infrared cameras fixed at known positions in the footage environment. The markers positions are then transformed into 3d positions using triangulation from the several cameras feeds, making it impossible to track a point's motion when it is not visible by two or more cameras. Moreover, using markers could be considered obtrusive, implicates the use of expensive specialized equipment and requires a footage taken in a specially arranged environment. Using video feeds from conventional cameras and without the use of special hardware, implicates the development of a marker less body motion capture system. Research in this domain is generally based on the articulated-models approach. Haritaoglu et al. [11] present an efficient system capable of tracking 2d body motion using a single camera. However this system is unable to provide 3d positions, restricting the information we can extract from the feeds. Bregler et al. [12] used gradient descent search with frame-to-frame region based matching and applied this method on multi camera sequences. This method proved to be unable to track agile motions with cluttered backgrounds. Combining 2d tracker and learned 3d configuration models, Howe et al. [13] were able to produce 3d body pose from short single camera feeds. Gavrilu and Davis [14] use an explicit hierarchical search to sequentially locate parts of the body's kinematic chain, reducing the search complexity significantly. In real world situations, it seems to be hard to specify each body part in the image independently without using labels or color cues. Sidenbladh et al. [15] use the Condensation algorithm [16] with learned stochastic models and a generative model of image formation to track full body motion. However, using dynamic trained models would restrict the generality of the approach and prevent the system from tracking gait abnormalities. Deutscher et al. [17] use weak dynamical modeling, multi-layers search and a modified Condensation algorithm to track the 3d movement from multiple camera feeds. Applying this type of layered search augments the risks of falling into local minima, especially in the case of a low frequency capture.

In fact we need to extract the exact 3D position of several points of the human body in order to detect gait abnormalities, using conventional digital camera feeds (25 frames/s) only. Respecting these conditions requires conceiving a new

approach. The method we present here is based on a modified version of the Condensation algorithm, dynamic bayesian networks and an articulated 3d model of the human body.

II. AN OVERVIEW OF THE CONDENSATION ALGORITHM

Condensation algorithm introduced by Isard et al. [16] is considered as a robust Bayesian state estimator. In this context, the system we observe is described by a state model and a state vector (containing the variables we want to estimate). In addition to the state model, we define an observation Y , through which the likelihood of a state vector at time t is evaluated; This likelihood is called weight. The normalized weight of a state's configuration m , would represent the observation probability $w^t(m) = P(Y_t|X = m)$. The system is thus characterized by a probability density function called the observation probability $P(Y_t|X)$ and the posterior probability density $P(X|Y_{0:t})$, where $Y_{0:t}$ is the observations history at t . The Condensation algorithm, also known as particle filtering, proved to be able to handle non-Gaussian, multi-modal probability densities. In fact it can model uncertainty by transmitting less fitting state configurations at t , to later time steps, and thus giving them a chance to be chosen. For each particle a weight is assigned at each time step. The posterior density is represented by the set of particles and their weights. The particle filtering can be viewed as a search for the best particle in a well defined particles set at each time step. In order to have a good state vector estimation, a minimum number of particles is necessary. This number becomes relatively big for large state vectors; Thus, the complexity of the algorithm would make it practically inapplicable in this case.

III. THE ARTICULATED BODY MODEL AND THE LIKELIHOOD FUNCTION

The 3d articulated human body model we use simulates the human movement through the configuration of 31 degrees of freedom. These degrees of freedom represent rotations of 19 joints of the body (head, elbows, knees etc..) and thus have constrained values. These constraints can easily be modified depending on the nature of the movement to be tracked. The body parts form an open loop kinematic chain composed of four branches starting from the torso. This 3d model is also characterized by a specific point which we call body origin (the sacrum point). The position of this point in our 3d referential (camera referential) defines the positions of all other points of the 3d model. No dynamic nor trained walking models were used, which makes our approach simple and generic.

In addition to defining a model, we define a function to evaluate its configuration (31 parameters values) likelihood to the real image; This function is called weight in a particle filtering context where each particle represents a configuration of the model. A silhouette image of the tracked body is constructed by subtracting the background from the current image and then applying a threshold filter. This image will then be compared to the synthetic image representing the model's configuration (2d projection of the 3d body's

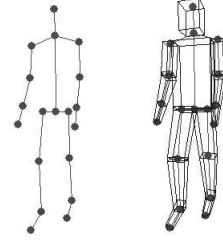


Fig. 1. 3d articulated human body model composed of 19 joints each represented by a 3d point.

model) to which we want to assign a weight; This weight will be greater when the interleaving zone of the two images is wider.

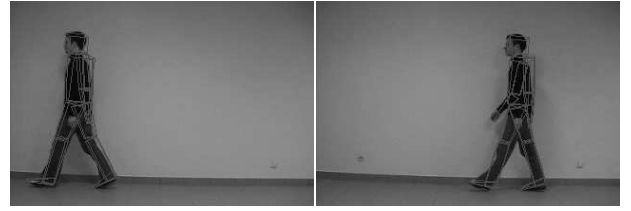


Fig. 2. Projection of articulated 3D model on sample images

IV. INTERVAL PARTICLE FILTERING

In our approach we consider the problem of the human motion tracking as a state estimation problem. The 3d model's parameters configuration forms the state vector X and the real image would be the observation Y . Condensation algorithm was used due to its capacity to handle the multi modal and non Gaussian observation probability. In parallel, we modified the basic algorithm by introducing the Interval Particle Filtering that tends to reconfigure the particles search space in an optimal way. These modifications preserve the advantages a particle filter offers and tend to reduce the complexity of the basic algorithm. Neither dynamic modeling nor an evolution model was used. Details and results of this algorithm can be found in [18].

The Interval Particle Filtering creates a fitter and more reliable search space than that created by the Condensation algorithm and present satisfactory results. In order to have a more precise tracking, we need to discretize the intervals of particles we create, with more particles. However, using more particles introduces more complexity; Our idea is to reduce the number of these particles created in accordance to the weights at time $t - 1$, by eliminating those that would be less fit at time t . This improvement is accomplished using the Dynamic Bayesian Networks formalism.

V. FACTORED REPRESENTATION OF A MARKOV PROCESS

Motion capture can be viewed as the monitoring of a Markov process. As mentioned before, the state of the system is the configuration of the articulated model. Observation is given by the video feeds.

Denoting as X_t the state of the system at time t and Y_t the corresponding observation, a Markov process is characterized by a transition model $P(X_t|X_{0:t-1}) = P(X_t|X_{t-1})$ and an observation function $P(Y_t|X_{0:t}) = P(Y_t|X_t)$.

Bayesian networks (BN) are a marriage between probability theory and graph theory. They provide a natural tool for dealing with uncertainty and complexity. A BN gives a graphical representation of the joint probability distribution of a set of variables. Each variable is associated to a local probability law and the joint probability distribution is the product of the local laws. When working with a time-evolving set of variable X_t we call the BN a Dynamic Bayesian Network (DBN)[19]. A DBN is a temporal probabilistic model which is based on a factored representation of the probability distributions $P(X_t|X_{t-1})$ and $P(Y_t|X_t)$ where X_t and Y_t are decomposable into several variables: $X_t = (X_t^1, \dots, X_t^N)$ and $Y_t = (Y_t^1, \dots, Y_t^M)$. The factored representation of $P(X_t, Y_t|X_{t-1})$ is given by a directed acyclic graph in which each node is associated to a variable X_t^i or Y_t^j .

$$P(X_t, Y_t|X_{t-1}) = \prod_i P(X_t^i|Pa(X_t^i)) \times \prod_j P(Y_t^j|Pa(Y_t^j))$$

where $Pa(X_t^i) \in (X_{t-1}, X_t), Pa(Y_t^j) \in (X_t)$ are the parents of X_t^i, Y_t^j in the directed acyclic graph.

Given such a representation, the goal of inference algorithms is to efficiently compute the *a posteriori* probability $P(X_t|Y_{0:t})$. Particle filtering can be used to perform approximate inference in DBNs with any kind of conditional probability distributions.

VI. FACTORIZING FROM THE KINEMATIC CHAIN

A kinematic chain is defined by a set of joints between segments $\{S_i\}_{i=1..N}$ allowing relative motion of the neighboring segments. In this work the value of the observation function depend on the absolute position $\{X_i\}_{i=1..N}$ of each element of the 3D model which is calculated from the relative angles between the body parts. It is possible to take advantage of the structure of the kinematic chain by using a BN to factor the representation of the state vector. We simply need to build a directed acyclic graph with the variables X_i by adding a link between two variables wherever a joint exists between the corresponding segments. One, or more particular segments, in our case the torso of the articulated 3d body model, has to be arbitrarily taken as root for choosing the direction of the links. The probability distribution $P(X)$ is now factored as $P(X) = \prod_{i=1}^N (P(X_i|Pa(X_i)))$ where $Pa(X_i)$ are the parents of X_i in the directed graph.

Due to occlusion problems, the evaluation of the position of the different body parts cannot be done independently. Nevertheless, it can be factored. The idea is that we can evaluate the matching of the segments following a chain process by successively applying

$$(w_n, Img_n) = score_n(X_i, Img_{n-1})$$

where Img_{n-1} is the background image used for the projection of the i 'th segment, w is the resulting score and

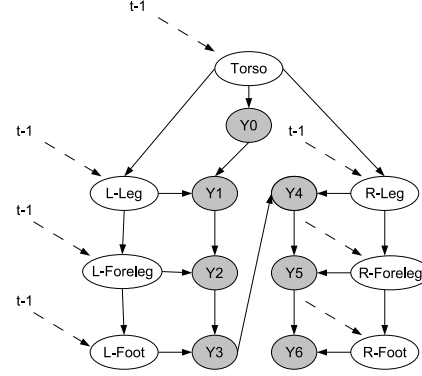


Fig. 3. A (partial) DBN for gait analysis, based on the kinematic chain of a 3d articulated body model. Grayed nodes (Y_i) correspond to observation functions. The parts of the kinematic chain corresponding the arms are missing on this representation for simplification reasons.

Img_n the new background image. We denote the observation probability given the position of the i 'th segment as $P(y_n|X_i, Y_{n-1})$.

We now have a factored representation of the state vector as well as of the observation function.

$$P(x, y) = \prod_i P(x_i|Pa(X_i)) \times \prod_j P(y_j|Pa(Y_j))$$

Since we do not use a model for gait analysis, we make the hypothesis that the position of the i 'th segment at time t only depends on the position of the i 'th segment at time $t-1$: $P(X_t^i|X_{t-1}) = P(X_t^i|X_{t-1}^i)$. Adding the corresponding edges to the Bayesian Network, we get Dynamic Bayesian Network which we will use to enhance the Interval Particle Filtering algorithm.

VII. HIERARCHICAL WEIGHTING AND RESAMPLING

As mentioned before, a critical issue in high dimensional particle filtering is that a majority of the particles get "killed off" during resampling step. The main idea behind this work is to use the factorization of the process for hierarchically resampling the components of the state vector. Our proposal is inspired from the likelihood weighting procedure [20][21] in which the elements of the state vector are sampled in a topological order. Being able to take observations as soon as possible into account, we propose to resample the particles whenever it is necessary.

- 1) The Diffusion Step : for each particle z_{t-1}^i we build a set of diffused particles z_t^{ij} .

We search the graph in a topological order.

- If the node corresponds to a state variable X_t^n , each particle z_t^{ij} gives birth to a new subset of particles which are completed with samples from $P(X_t^n|Pa(X_t^n) = u_j)$ where u_j is the value of $Pa(X_t^n)$ in the particle z_t^{ij} . Thus, the set of diffused particles is extended each time a state node is encountered.

- If the node corresponds to an observation function. The weight $w_t^{i,j}$ of the particles $z_t^{i,j}$ is updated by the rule :

$$w_t^{i,j} = w_t^{i,j} * P(Y^k = y_t^k | Pa(Y_t^k) = u_i)$$

where u_j is the value of $Pa(Y_t^k)$ in particle $z_t^{i,j}$. If necessary the particles $z_t^{i,j}$ are resampled according to their weight $w_t^{i,j}$. Thus, the set of diffused particles is reduced each time an observed node is encountered.

- 2) The Drift Step : the diffused particles $z_t^{i,j}$ are already weighted. Select the ones that fit the best the observation (ie: with the highest weight) for the next diffusion step.

VIII. RESULTS

The number of particles evolving in the state space can be chosen according to the application constraints regarding movement tracking accuracy. When applied to gait analysis, the variables of interest are the ones related to the leg and torso movements. They represent 5 degrees of freedom out of 31. These variables are explored extensively by IPF[18]. Choosing an exploration step of 5° for these variables with a total of 6000 particles, each frame requires 4.5 seconds of processing time on a P4 3GHz PC. Occlusion problems that may lead for example to confusion between the left leg and the right leg were solved using two cameras. The motion capture given by IPF was compared to the tracking performed by a 6 camera Vicon system. Sample curves are presented on figure VIII.

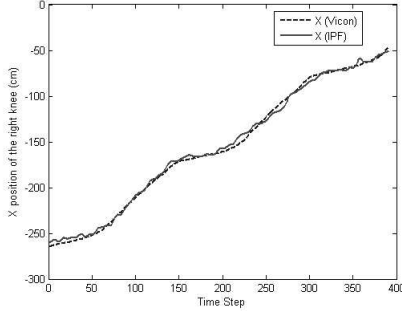


Fig. 4. Comparison of the longitudinal displacement x of the right knee estimated by IPF (in plain) and by Vicon (in dotted). These x values are calculated in the camera referential.

The analysis of the video sequences was then performed using Factored IPF. Due to factorization, we were able to increase the accuracy of tracking by increasing the number of interesting variables and reducing the exploration step for these variables to 2° . The computation time was reduced to 1.5 sec for each frame which represents a diminution by a factor 3.

IX. CONCLUSION

In this paper we described a marker less motion capture system for gait analysis. The system is based on the analysis

of conventional video streams. Viewing this application as a Markov monitoring problem we proposed to use a Dynamic Bayesian Network to factor the state vector and the observation function. We defined a new particle filtering method taking advantage of the factorization for hierarchically weighting and resampling the particles. The factored approach is a way of overcoming complexity in this high dimensional monitoring problem.

REFERENCES

- [1] L. Lee and W. Grimson. Gait analysis for recognition and classification, 2002.
- [2] S.G Brauer, Y.R Burns, and Galley P. A prospective study of laboratory and clinical measures of postural stability to predict community-dwelling fallers. *Journal of Gerontology*, 55:469–76, 2000.
- [3] Noraxon®, <http://www.noraxon.com>.
- [4] Gaitrite®, <http://www.gaitrite.com>.
- [5] Jennifer C Nitz and Nancy Low Choy. The relationship between ankle dorsiflexion range, falls and activity level in women aged 40 to 80 years. *New Zealand Journal of Physiotherapy*, 32(3):121–125, 2004.
- [6] N. Noury. *Un capteur intelligent pour la détection des chutes - Fusion multi-capteurs et décisions à base de règles*, volume 2. Hermès-Elsevier, 2004.
- [7] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. Robert. An ambulatory system for physical activity monitoring in elderly. In *Microtechnologies in Medicine and Biology, 1st Annual International Conference On*. 2000, pages 562–566, Lyon, France, 2000.
- [8] R. Moe-Nilssen. A new method for evaluating motor control in gait under real-life environmental conditions. *Journal of Clinical Biomechanics*, 13:320–335, 1998.
- [9] Hylton B. Menz, Stephen R. Lord, and Richard C. Fitzpatrick. Age-related differences in walking stability. *Age and Ageing*, 32:137–142, 2003.
- [10] Vicon®, <http://www.vicon.com>.
- [11] I. Haritaoglu, D. Harwood, and L. Davis. Who, when, where, what: A real time system for detecting and tracking people in 2.5d. In *5th European Conf. Computer Vision*, pages 877–892, Freiburg Germany, 1998.
- [12] C. Bregler and J. Malik. Tracking people with twists and exponential maps. In *Computer Vision and Pattern Recognition 1998*, 1998.
- [13] Nicholas R. Howe, Michael E. Leventon, and William T. Freeman. Bayesian reconstruction of 3d human motion from single-camera video. In *Advances in Neural Information Processing Systems MIT Press 2000*, pages 820–826, 2000.
- [14] D.M. Gavrilu and L. Davis. 3d model based tracking of humans in action: A multi view approach. In *Computer Vision and Pattern Recognition 1996*, pages 73–80, 1996.
- [15] Hedvig Sidenbladh, Michael J. Black, and David J. Fleet. Stochastic tracking of 3d human figures using 2d image motion. In *6th European Conference on Computer Vision*, pages 702–718, Dublin, 2000.
- [16] Michael Isard and Andrew Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1):5–28, 1998.
- [17] J. Deutscher, A. Blake, and I. Reid. Articulated body motion capture by annealed particle filtering. In *Computer Vision and Pattern Recognition*, volume 2, pages 126–133, Hilton Head Island, SC, USA, 2000.
- [18] J. Saboune and F. Charpillat. Using interval particle filtering for marker less 3d human motion capture. In *17th IEEE International Conference on Tools with Artificial Intelligence*, pages 621–627, Hong Kong, 2005.
- [19] Kevin Patrick Murphy. *Dynamic Bayesian Networks : Representation, Inference and Learning*. PhD thesis, University of California Berkeley, 2002.
- [20] Robert M. Fung and Kuo-Chu Chang. Weighing and integrating evidence for stochastic simulation in bayesian networks. In *UAI '89: Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence*, pages 209–220, Amsterdam, The Netherlands, The Netherlands, 1990. North-Holland Publishing Co.
- [21] Ross D. Shachter and Mark A. Peot. Simulation approaches to general probabilistic inference on belief networks. In *UAI '89: Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence*, pages 221–234, Amsterdam, The Netherlands, The Netherlands, 1990. North-Holland Publishing Co.