Markerless Articulated Human Body Tracking for Gait Analysis and Recognition

Tomasz Krzeszowski[1,2], Bogdan Kwolek[1,2], Konrad Wojciechowski[1,3], Henryk Josinski[1,3]
[1]*Polish-Japanese Institute of Information Technology, Aleja Legionów 2, 41-902 Bytom, Poland*
[2]*Rzeszow University of Technology, Al. Powstańców Warszawy 12, 35-959 Rzeszów, Poland*
[3]*Silesian University of Technology, Akademicka 16, 41-100 Gliwice, Poland*

June 18, 2011

Abstract. We present a particle swarm optimization (PSO) based system for markerless full body motion tracking. The fitness function is smoothed in an annealing scheme and then quantized. In this manner we extract a pool of candidate best particles. The swarm of particles selects a global best from such a pool of the particles to force the PSO jump out of stagnation. Experiments on 4-camera datasets demonstrate accuracy of our method on image sequences with walking persons. The system was evaluated using ground-truth data from marker-based motion capture system by Vicon. We compared the joint motions and the distances between ankles, which were extracted by both systems. Thanks to high precision of the markerless motion estimation the curves illustrating the distances between ankles overlap considerably in almost all frames of the image sequences.

Key words: Markerless motion tracking, particle swarm optimization, gait analysis and recognition

1. Introduction

In 1995, Dyer et al. in Motion Capture White Paper that was released by SGI stated that "motion capture is one of the hottest topics in computer graphics today". According to paper mentioned above [3], motion capture involves measuring an object's position and orientation in physical space, then recording that information in a computer-usable form. Today there is a great interest in motion capture and the number of papers related to this research area grows exponentially. In the last years many commercial motion capture (also known as MoCap) systems [13][16] have been developed for gaming and animation [19], rehabilitation and medical treatment [12][18], sport [17], and military applications. There are several methods of motion capture, including optical, electromagnetic and mechanical ones. Mechanical systems require performers to wear exoskeletons. Magnetic systems detect the positions and orientation using a magnetic field and they are sensitive to the presence of metallic objects in the environment. In optical MoCap systems either passive reflective markers or active markers are attached to a performer, and a system of fixed cameras records the position of these markers. In a passive system each camera has a ring of LEDs around the lens that emit light towards the subject. The light then bounces off the reflective markers. Afterwards, the information captured from the

reflective markers is triangulated.

There are intrinsic problems in using markers, mainly the inconvenience of attaching them to the body, failures in tracking them, and the requirement for special lighting conditions. In consequence, such a technology can only be used in highly controlled laboratory conditions. During capturing of the motion the surface markers can move relative to underlying bone, resulting in soft-tissue artifacts. Thus, many researchers believe that the ultimate solution to human MoCap is markerless tracking [14]. The research in markerless tracking is motivated mainly by unreliable anatomical landmark identification and the usability of such a technology in controlled laboratory environments only. Markerless systems utilize computer vision techniques to estimate motion parameters directly from video footage without the use of special markers. These approaches are less accurate than optical systems but they are more affordable and portable.

The experiments with moving light displays [1] demonstrated that people can recognize human activities on the basis of motion of a small set of points attached to the human body. The experimental results stimulated discussion whether the recognition takes place on the basis of 2D motion patterns, or rather it is achieved through 3D reconstructions from the motion of patterns. In [5] the authors argue that in activity analysis, motion cues are more important than spatio-temporal representations. In general, articulated models in action analysis make easier the occlusion handling. The model-based approaches are view independent and less dependent on the training data.

Human gait is a very complex activity to analyze because it is a complex set of coordinated sub-activities [15]. The lengths of limbs, body mass and muscularity, stride length, and several other factors have influence on how a person walks. Walking is a periodical activity and thus frequency analysis of spatio-temporal signals characterizing the walking style can be a very attractive approach [11][24]. Current approaches to determining the identity of individuals in image sequences by the way they walk can be classified into two broad categories, namely appearance-based ones that deal directly with image statistics and model-based ones that analyze the variation of the model parameters that has been fitted in advance to the image data. The majority of the approaches proposed for gait recognition are based on analyzing image sequences acquired by a single camera. The major drawback of such approaches is that they are typically designed only for a specific viewpoint, usually fronto-parallel. Furthermore, achieving resistance against illumination variations and clothing changes is not straightforward. To cope with difficulties mentioned above much effort was expanded to this end, for example by using image-based reconstruction to achieve view-independent motion classification [20]. However, the inability to take into account the self-intersection constraints as well as to encode joint angle limits makes 2D models unsuitable for tracking of real human movements. As it was previously mentioned, the methods that are based on holistic space-time features or space-time shapes depend more on the training data in comparison to the model-based approaches.

For natural scenarios, a system that operates in unconstrained environment where, maybe, there are varying illumination conditions, and where a subject moves freely is needed. In real-world environments, 2D analysis based gait recognition can lead to poor identification performance due to varying viewpoints, occlusion and appearance variations, clothing and illumination changes, and in consequence it cannot provide sufficiently accurate results. In conventional gait analysis, the trajectories of points on the human body that correspond to anatomical landmarks are employed [8]. Urtasun and Fua [10] proposed an approach that relies on matching 3D motion models to images, and then tracking and restoring the motion parameters. The tests were performed on datasets with four people, i.e. 2 men and 2 women walking at 9 different speeds ranging from 3 to 7 km/h by increments of 0.5 km/h. A Vicon motion capture system was used to build motion models. In [22] Sigal et al. demonstrated that on the basis of 3D articulated pose estimates it is possible to infer subtle physical attributes of human, like gender and weight, and even some aspects of mental state, *e.g.*, happiness or sadness. In this context it is worth noting that authors of the work mentioned above point that with 3D-model based methods the inferring attributes of unfamiliar people does not presuppose that test subjects exist in the training data. Moreover, by the use of 3D articulated tracking we do not need view-based models. One of the benefits of model-based trackers is that they permit a comprehensive exploration of the space of possible poses. In [21] a markerless system for motion tracking in surveillance videos has been proposed. In this work it is demonstrated that recovering of the human pose is possible at a distance even in case of temporal occlusions, imperfect camera calibration and synchronization.

In this paper, video sequences captured by four synchronized cameras are used as input. The motion of walking person is inferred with the help of a 3D human model. The 3D human pose is reconstructed through matching the projection of the human body with the image observations. The human silhouette is extracted via background subtraction and then the edges are located within the extracted silhouette. The objective function takes into account the normalized distance between the model's projected edges and the closest edges in the image. Estimating poses from frame to frame is done assuming temporal coherence over time and it is achieved through a global search around the pose estimated in the previous frame. In the search for the best match we employ particle swarm optimization [4]. We demonstrate that a modified particle swarm optimization algorithm [23] allows us to obtain far better tracking accuracy in comparison to ordinary particle swarm optimization as well as the particle filter [6][7]. The investigated searching schemes were compared by analyses carried out both through qualitative visual evaluations as well as quantitatively using data captured by Vicon system as ground truth. We demonstrate visually the velocities of the ankles that were determined by our markerless system and marker-based system from Vicon. To show the potential of our approach for gait analysis and recognition we show the plots with distances between ankles for both motion capture systems.

2. Searching Schemes for Human Motion Tracking

The Particle Swarm Optimization algorithm (PSO) is a population-based stochastic search algorithm to the complex non-linear optimization problems. The PSO algorithm was first introduced by Kennedy and Eberhart in 1995 [4] and its central idea was inspired by imitating of the social behavior of animals such as bird flocking, fish schooling, etc. Its searching power stems from the communication mechanisms among individuals to share both individual and global knowledge when a population of birds or insects search food or migrate in a searching space, although not all the individuals know where the best position is. In particular, if any population member can find out a promising position or a promising path, the rest of the swarm will follow such an individual quickly. More specifically, PSO does not use the gradient of fitness function of the problem being optimized, which means that PSO does not require the optimization function to be differentiable as is required by classic optimization methods such as gradient descent. PSO can therefore be utilized to solve optimization problems that are non-linear, multimodal, noisy, change over time and so on.

In the PSO algorithm each particle represents a potential solution to the optimization problem. Much of the success of PSO algorithms comes from the fact that individual particles have tendency to diverge from the best known position in any given iteration, enabling them to ignore local optima, while the swarm as a whole gravitates towards the global extremum. Starting with randomly initialized locations and moving in randomly chosen directions, each particle flies through the multidimensional search space with a certain velocity and remembers its best previous positions. Each swarm member communicates good positions to each other as well as dynamically adjusts its own position and velocity derived from the best positions of all particles. All particles are subsequently scored by a fitness function $f : R^n \rightarrow R$. In consequence, all particles tend to fly towards better and better positions over the searching process until the swarm as a whole does not converge to the position with the best fitness value.

Since the object tracking is a kind of dynamic optimization, the tracking can be attained through incorporating the temporal continuity information into the ordinary PSO. Consequently, the tracking can be obtained by a sequence of static PSO-based optimizations, followed by re-diversification of the particles to cover the potential poses that can arise in the next time step. The re-diversification of the particle i can be achieved on the basis of normal distribution, which is concentrated around the state estimate $\hat{x}_{t-1}$ in time $t-1$: $x_t^i \leftarrow \mathcal{N}(\hat{x}_{t-1}, \Sigma)$, where Σ denotes the covariance matrix of the Gaussian distribution, whose diagonal elements are proportional to the predicted velocity $v_t = \hat{x}_{t-1} - \hat{x}_{t-2}$.

The original PSO is not guaranteed to converge on a local extremum. Clerc and Kennedy [9], in their study on stability and convergence of particle swarm optimization have demonstrated that the utilization of a constriction factor may be necessary to ensure

the convergence of the algorithm. Their work indicated that the inclusion of properly defined constriction factor leads to better convergence. Moreover, such a coefficient can prevent explosion and induce particles to converge on local optima. In the discussed study, they introduced a constriction factor ω and the following form of the formula expressing the i-th particle's velocity:

$$v^{i,k+1} = \omega[v^{i,k} + c_1 r_1(p^i - x^{i,k}) + c_2 r_2(g - x^{i,k})] \tag{1}$$

where constants c_1 and c_2 are responsible for balancing the individual's self-knowledge and group's social knowledge, respectively, r_1 and r_2 denote uniformly distributed random numbers, x^i stands for position of the i-th particle, p^i is the local best position of particle, g is the global best position, whereas ω is a function of c_1 and c_2, which is given by the following equation:

$$\omega = \frac{\kappa}{|2 - \phi - \sqrt{\phi^2 - 4\phi}|} \tag{2}$$

where $\kappa = 2$, $\phi = c_1 + c_2$, and $\phi > 4$ to guarantee computational stability. As ϕ increases, the constriction factor ω decreases and diversification is reduced. Typically, ϕ is set to 4.1 (i.e. $c_1, c_2 = 2.05$) and the ω is equal to 0.729.

Building on the successful applications of PSO with constriction factor we employ a linearly decreasing constriction factor, instead of a constant one. In our approach the value of ω depends on annealing factor α in the following manner:

$$\omega = \omega_{min} + (\omega_{max} - \omega_{min})(1.1 - \alpha) \tag{3}$$

where $\omega_{min} = 0.52$, $\omega_{max} = 1.32$, $\alpha = 0.1 + \frac{k}{K+1}$, $k = 0, 1, \ldots, K$, and K is the number of iterations. The values of ω_{min} and ω_{max} were determined experimentally, whereas the maximal velocity was limited to the dynamic range of the state variable in each dimension.

The annealing factor α is also employed to smooth the objective function. The larger the iteration number is, the smaller is the smoothing. In consequence, in the last iteration the algorithm utilizes the non-smoothed function. The algorithm termed as annealed PSO (APSO) [23] can be expressed as follows:

1. For each particle i
2. initialize $v_t^{i,0}$
3. $x_t^{i,0} \sim \mathcal{N}(g_{t-1}, \Sigma_0)$
4. $p_t^i = x_t^{i,0}$, $f_t^i = f(x_t^{i,0})$
5. $u_t^i = f_t^i$, $\tilde{u}_t^i = (u_t^i)^{\alpha_0}$
6. $i^* = \arg\min_i \tilde{u}_t^i$, $g_t = p_t^{i^*}$, $w_t = u_t^{i^*}$
7. For $k = 0, 1, \ldots, K$
8. update ω_α on the basis of (3)

9. $G = \arg\min_i round(num\_bins \cdot \tilde{u}_t^i)$
10. For each particle i
11. Select a particle from $\{G \cup g_t\}$ and assign it to g_t^i
12. $v_t^{i,k+1} = \omega_\alpha[v_t^{i,k} + c_1 r_1(p_t^i - x_t^{i,k}) + c_2 r_2(g_t^i - x_t^{i,k})]$
13. $x_t^{i,k+1} = x_t^{i,k} + v_t^{i,k+1}$
14. $f_t^i = f(x_t^{i,k+1})$
15. if $f_t^i < u_t^i$ then $p_t^i = x_t^{i,k+1},\ u_t^i = f_t^i,\ \tilde{u}_t^i = (u_t^i)^{\alpha_k}$
16. if $f_t^i < w_t$ then $g_t = x_t^{i,k+1},\ w_t = f_t^i$

The smoothed objective functions are quantized, see 9th line in the pseudo-code. Owing to this the similar function values are clustered into the same range of values. In each iteration the algorithm determines the set G of the particles, which after the quantization of the smoothed fitness function from the previous iteration, assumed the smallest values (best fitness scores). For each particle i the algorithm selects the global best particle g_t^i from $\{G \cup g_t\}$, where g_t determines the current global best particle of the swarm. By means of this operation the swarm selects the global best location from a pool of candidate best locations to force the PSO jump out of stagnation. We found that this operation contributes considerably toward better tracking, particularly in case of noisy observations. It is worth noting that in the literature devoted to dynamic optimization the problem of optimization of noisy objective functions is considered very rarely.

The fitness score for i-th camera's view is calculated on the basis of following expression: $f^i(x) = 1 - ((f_1^i(x))^{w_1} \cdot (f_2^i(x))^{w_2})$, where w denotes weighting coefficients that were determined experimentally. The function $f_1(x)$ reflects the degree of overlap between the extracted body and the projected 3D model into 2D image corresponding to camera i. The function $f_2^i(x)$ reflects the edge distance-based fitness [23] in the image from the camera i. The objective function for all cameras is determined according to the following expression: $f(x) = \frac{1}{4}\sum_{i=1}^{4} f^i(x)$.

3. Experiments

The markerless motion tracking algorithm was evaluated on three image sequences acquired by four synchronized and calibrated cameras. Each pair of the cameras is approximately perpendicular to the other camera pair. The cameras acquire color images of size 1920×1080 with 25 frames per second. The experiments were done on images subsampled by factor of 2 both horizontally and vertically. Ground truth data were obtained by a commercial motion capture system from Vicon Nexus. The system employs reflective markers and sixteen cameras to recover the 3D location of such markers at rate of 100 Hz. The motion capture system is capable of differentiating overlapping markers from each camera's view. The synchronization between the MoCap and multi-camera

system was done using hardware from Vicon Giganet Lab. The location of the cameras and our laboratory layout is depicted in Fig. 1.

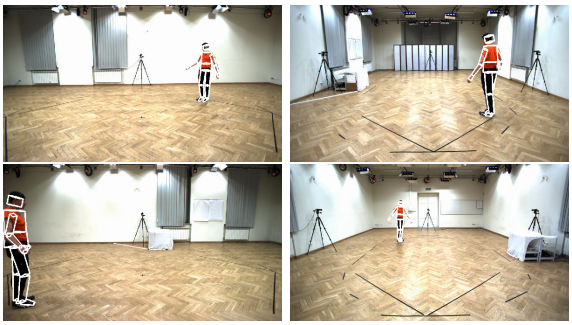

Fig. 1. Layout of the laboratory for human motion tracking. The images illustrate the initial model configuration, overlaid on the image in first frame and seen in view 1 and 2 (upper row), and in view 3 and 4 (bottom row).

The accuracy of our algorithm for human motion tracking was evaluated experimentally in scenarios with a walking person. Although our focus was on tracking of torso and legs, we also estimated the head's pose as well as the pose of both arms. The human body model consists of eleven rigid body parts, including pelvis, upper torso, two upper arms, two lower arms, two thighs, two lower legs and a head, each of which was represented by a truncated cone. The body configuration is parameterized by position and orientation of the pelvis in the global coordinate system as well as relative angles between the connected limbs. Although the hands and feet appear in Fig. 1, we do not consider them in the model parameterization. Figure 2 depicts some tracking results, which were obtained in a sequence #1 with a person following a line joining two non-

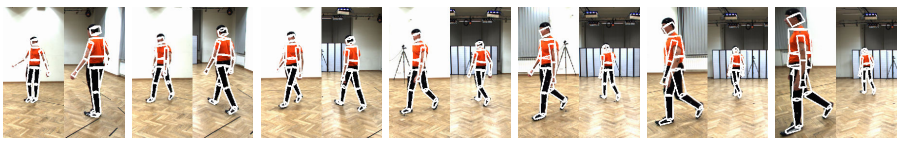

Fig. 2. Articulated 3D human body tracking in four camera setup. Shown are results in frames #0, 20, 40, 60, 80, 100, 120. The left sub-images are seen from view 1, whereas the right ones are seen from view 2.

consecutive laboratory corners. The overlap of the projected 3D model on the subject

undergoing tracking can be utilized to illustrate the tracking accuracy.

In Fig. 3 are shown some tracking results that were obtained in a sequence #2. The results were obtained using frontal (back) and side views. As one can observe, the projected model matches the person silhouette reasonably well.

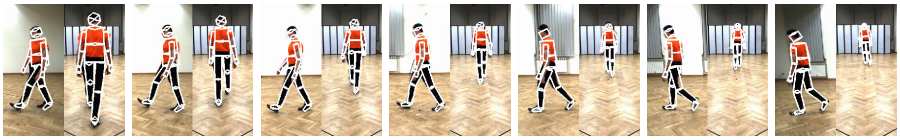

Fig. 3. Articulated 3D human body tracking in four camera setup using frontal (back) and side views. Shown are results in frames #20, 40, 60, 80, 100, 120, 140. The left sub-images are seen from view 1, whereas the right ones are seen from view 2.

The plots shown in Fig. 4 illustrate the accuracy of motion estimation for some joints, which were obtained by APSO algorithm. As we can observe the tracking error of both

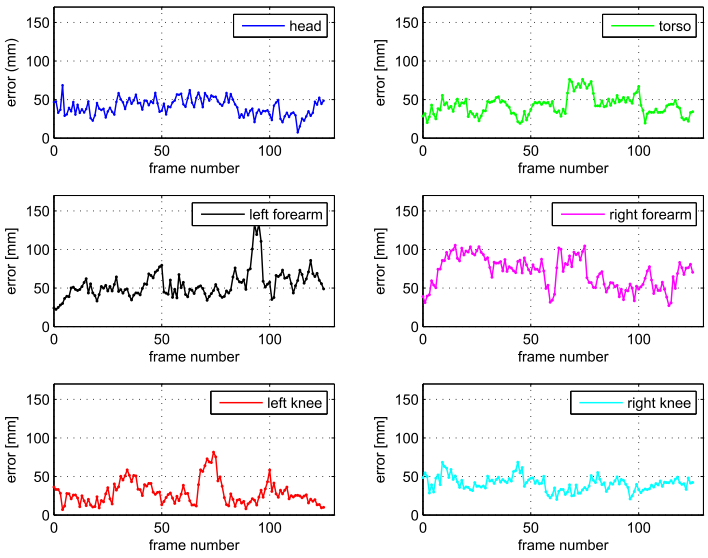

Fig. 4. Tracking errors [mm] versus frame number for APSO.

forearms is somewhat larger in comparison to the remaining limbs. The average error of

both knees is about 50 mm, whereas the maximal errors do not exceed 80 mm for the left knee and 70 mm for the right one. Gait is mainly a motion of lower limbs, although motion of upper limbs can also be considered in gait analysis. Therefore, a bit larger errors of both limbs do not have crucial influence on the performance of gait analysis. The discussed results were obtained by APSO in 20 iterations and using 300 particles.

In order to compare our APSO-based motion tracker with trackers built on ordinary PSO and particle filter (PF) we conducted experiments in sequences of images with a walking actor. The experimental results depicted in Fig. 5 demonstrate that APSO algorithm outperforms both PSO and PF in terms of tracking accuracy. The experimental results with APSO and PSO were obtained in 20 iterations and using 300 particles, whereas in experiments with PF we employed 6000 particles. As we can see, the tracking accuracy of PF-based algorithm is much worse in comparison to algorithms built on particle swarm optimization.

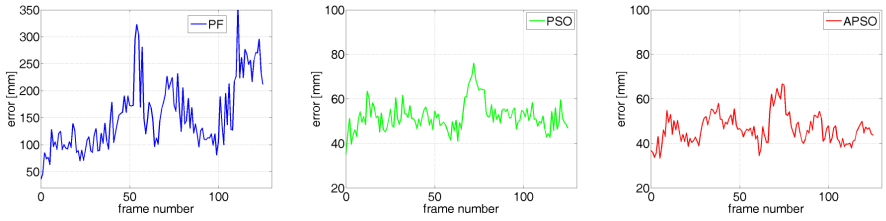

Fig. 5. Tracking errors of PF, PSO and APSO (from left to right).

We evaluated the tracking accuracy of APSO, PSO and PF algorithms for motion tracking on three image sequences with a walking person. Some images from sequence #1 are depicted in Fig. 2, whereas some images from sequence #2 are shown in Fig. 3. In Tab. 1 are depicted some quantitative results, which are averages over ten runs of the motion tracker with unlike initializations. In the PF we employed the classical likelihood function $L(z|x) \propto e{-}\frac{f(x)^2}{\sigma^2}$, where z stands for the observation, x denotes state, and σ determines how fast the likelihood will decrease when the fitness function is large (bad particles).

The discussed results for the full body tracking were obtained for $M = 39$ markers. From the above set of markers, 4 markers were placed on the head, 7 markers on each arm, 12 on the legs, 5 on the torso and 4 markers were attached to the pelvis. Given such a placement of the markers on the human body and the estimated human pose, which has been calculated by our algorithm, the corresponding positions of virtual markers were determined and then utilized in calculating the average Euclidean distance between corresponding markers. The average Euclidean distance $\overline{d}_i$ for each marker i

Tab. 1. Average errors for $M = 39$ markers in three image sequences. The images from Seq. 1 are shown on Fig. 2, whereas the images from Seq. 2 are depicted on Fig. 3.

			Seq. 1	Seq. 2	Seq. 3
	#particles	it.	error [mm]	error [mm]	error [mm]
PF	1000		230.2±121.9	241.7±135.5	158.8±96.7
	2000		190.4±106.6	195.1±112.0	142.1±88.0
	3000		187.6±110.5	173.2±100.9	137.8±86.8
	6000		163.3±96.1	163.1±99.4	121.6±76.8
	15000		147.9±84.4	143.6±89.8	112.9±71.3
PSO	100	10	59.5±28.9	60.1±27.1	69.4±34.8
	100	20	53.6±23.3	57.0±26.2	61.1±30.1
	300	10	52.4±23.0	54.5±24.9	61.0±30.4
	300	20	49.1±21.6	53.3±22.8	59.0±28.0
APSO	100	10	55.3±24.8	56.1±24.7	62.9±30.6
	100	20	50.4±22.1	50.2±22.1	53.7±23.7
	300	10	48.3±19.8	51.0±22.7	55.0±25.8
	300	20	46.0±18.0	48.6±20.8	54.5±26.4

was calculated using real world locations $m_i \in R^3$ on the basis of the following equation:

$$\overline{d}_i = \frac{1}{T} \sum_{t=1}^{T} ||m_i(\hat{x}_t) - m_i(x_t)|| \tag{4}$$

where $m_i(\hat{x})$ stands for marker's position that was calculated using the estimated pose, $m_i(x)$ denotes the position, which has been determined using ground-truth information, whereas T stands for the number of frames. The errors reported in columns 4-6 of Tab. 1 indicate the distance errors, which are averages for $M = 39$ markers. For each marker i the standard deviation σ_i was calculated as follows:

$$\sigma_i = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} \left(||m_i(\hat{x}_t) - m_i(x_t)|| - \overline{d}_i \right)^2} \tag{5}$$

The standard deviation $\overline{\sigma}$ shown in Tab. 1 is the average over all markers. As we can observe, the APSO algorithm outperforms both PSO and PF based motion trackers.

Figure 6 depicts the plots of the velocities of the ankles over time for both our markerless motion capture system and Vicon MoCap system. The cone is calculated

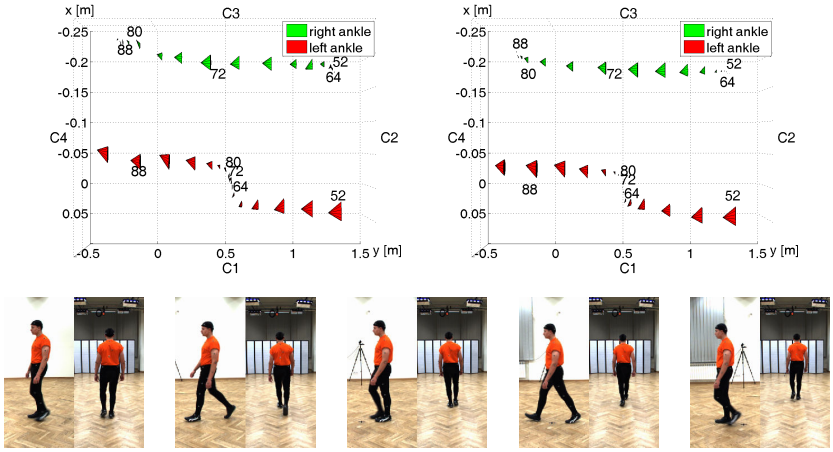

Fig. 6. Temporal velocities of the ankles in a single step estimated by APSO (left) and Vicon MoCap (right) and the person's poses, acquired by the camera C1 and C2 in frames #52, 64, 72, 80 and 88. The bigger the height of the cone is, the larger is the ankle velocity.

for every second frame. Below the plots, some corresponding images are shown. The location of the cone reflects the location of the ankle on xy plane in particular frame, see also frame numbers near the cones. The bigger the height of the cone is, the larger is the ankle velocity. For the visualization purposes the sizes of the cones were downscaled by factor 36. As one can observe, the average velocity of the performer is about 1 m/s, whereas the maximal temporal velocity of the right ankle is in frames 68-70 and takes the value about 3 m/s. For walking there exists a phase where both feet are in contact with the ground. Such a phase can be detected easily from the depicted plots. Moreover, as we can observe, the difference between the velocities estimated by our system and the velocities obtained by marker-based system is not too large.

In Fig. 7 are shown plots illustrating the distance between the ankles of the walking performer. The distances were estimated using data from our markerless algorithm and marker-based system from Vicon. As one can observe, the curves that were determined by two different motion capture systems considerably overlap in almost all frames of the considered sequences. High overlap between both curves formulates a rationale for the usage of the markerless motion tracking to achieve view-independent gait recognition. Therefore, the distances between ankles together with the joint angles were employed in experiments consisting in identification of the person on the basis of the well known

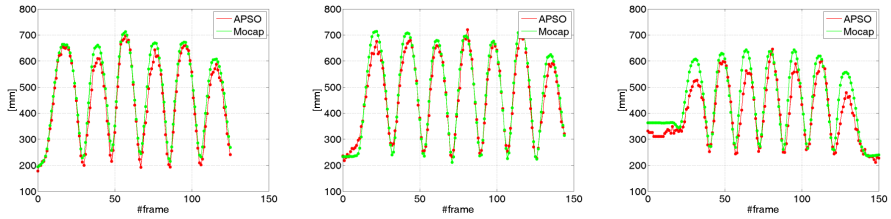

Fig. 7. Distance between ankles for markerless APSO-based system and marker-based system from Vicon. Sequences Seq. 1, Seq. 2 and Seq. 3 (from left to right).

Dynamic Time Warping (DTW). The DTW algorithm [2] calculates the distance between each possible pair of points of two sequences (e.g. time series). It uses these distances to calculate a cumulative distance matrix and finds the least expensive path through this matrix. Such a path represents the best warp for which the feature distances between their synchronized points are minimized. In consequence, similarities in motion patterns are detected, even if in one video the individual is walking slowly and if in another one he or she is walking more quickly, or even if accelerations and decelerations during the walking take place. We used DTW to calculate the distance between a test sequence and a reference sequences. The experiments were carried out on ten sequences with walking performers. Overall, our initial results indicate that we could achieve a classification accuracy of over 95%.

4. Conclusions

We have presented a vision system for markerless articulated human body tracking. The tracking is done using a modified particle swarm optimization algorithm. The objective function is smoothed in an annealing scheme and then quantized. This allows us to extract a pool of candidate best particles. The algorithm selects a global best from such a pool to force the PSO jump out of stagnation. Experiments on 4-camera datasets demonstrate accuracy of our approach. The joint velocities and distances between ankles were compared with the features extracted by marker-based commercial MoCap system from Vicon.

Acknowledgment

This work has been supported by the National Science Centre (NCN) within the research project N N516 483240 and the National Centre for Research and Development (NCBiR) within the project OR00002111.

References

1973

[1] Johansson G.: Visual perception of biological motion and a model for its analysis, Perception and Psychophysics, vol. 14, 201-211, 1973.

1978

[2] Sakoe H., Chiba S.: Dynamic programming algorithm optimization for spoken word recognition, IEEE Trans. on Acoustics, Speech and Signal Processing, 26(1), pp.4349, 1978.

1995

[3] Dyer S., Martin J., Zulauf J.: Motion Capture White Paper, SGI, on-line at `ftp://ftp.sgi.com/sgi/A%7CW/jam/mocap/MoCapWP_v2.0.html`.

[4] Kennedy J., Eberhart R. C.: Particle swarm optimization, Proc. of IEEE Int. Conf. on Neural Networks, Piscataway, NJ, pp. 1942-1948, 1995.

[5] Cedras C., Shah M.: Motion-based recognition a survey, Image and Vision Computing, vol. 13, 129-155, 1995.

1996

[6] Isard M., Blake A: Contour tracking by stochastic propagation of conditional density, In Proc. of the 4th European Conference on Computer Vision, Springer-Verlag, Cambridge, London, UK, 343-356, 1996.

2000

[7] Doucet A., Godsill S., Andrieu Ch.: On sequential Monte Carlo sampling methods for Bayesian filtering, Statistics and Computing, vol. 10, no. 3, 197-208, 2000.

2001

[8] Tanawongsuwan R., Bobick A.: Gait recognition from time-normalized joint-angle trajectories in the walking plane, Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2001, vol.2, pp. II-726-731, 2001.

2002

[9] Clerc M., Kennedy J.: The particle swarm - explosion, stability, and convergence in a multidimensional complex space, IEEE Trans. on Evolutionary Computation 6(1), 58-73, 2002.

2004

[10] Urtasun R., Fua P.: 3D tracking for gait characterization and recognition. In Proc. of the Sixth IEEE Int. Conf. on Automatic Face and Gesture Recognition, IEEE Computer Society, Washington, DC, USA, 17-22, 2004.

2005

[11] Boulgouris N. V., Hatzinakos D., Plataniotis, K. N.: Gait recognition: a challenging signal processing technology for biometric identification, Signal Processing Magazine, IEEE, vol.22, no.6, pp. 78-90, Nov. 2005.

[12] Saboune J., Charpillet F.: Markerless human motion capture for gait analysis, Clinical Orthopaedics and Related Research, Springer, 2005.

2006

[13] Moeslund T. B., Hilton A., Krüger, V.: A survey of advances in vision-based human motion capture and analysis, Comput. Vis. Image Underst., vol. 104, no. 2-3, 90-126, Nov. 2006.

[14] Mündermann L., Corazza S., Andriacchi T. P.: The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications, J. Neuroengineering Rehabil., vol. 3, no. 6, 2006

[15] Nixon M. S., Carter J.N.: Automatic Recognition by Gait, Proc. of the IEEE, vol. 94, no. 11, pp. 2013-2024, 2006.

2007

[16] Poppe R.: Vision-based human motion analysis: An overview, Comput. Vis. Image Underst., vol. 108, no. 1-2, 4-18, 2007.

2008

[17] Barris S., Button C.: A review of vision-based motion analysis in sport, Sports Medicine, vol. 38, no. 12, 1025-1043, 2008.

[18] Zhou H., Hu H.: Human motion tracking for rehabilitation - A survey, Biomedical Signal Proc. and Control (3):118, 2008.

2009

[19] Quah C. K., Koh M., Ong A., Seah H. S., Gagalowicz A.: Video-based Motion Capture for Measuring Human Movement, [in:] Digital Sports for Performance Enhancement and Competitive Evolution: Intelligent Gaming Technologies, Information Science Publishing, 209-228, 2009.

[20] Bodor R., Drenner A., Febr D., Masoud O., Papanikolopoulos N.: View-independent human motion classification using image-based reconstruction, Image Vision Comput., vol. 27, no. 8, 1194-1206, 2009.

2010

[21] Kwolek B.: Multiple views based human motion tracking in surveillance videos, In Proc. of 8th IEEE Int. Conf. on Advanced Video and Signal-Based Surveillance, pp. 492-497, 2011.

[22] Sigal L., Fleet D. J, Troje N. F., Livne M.: Human attributes from 3D pose tracking, In Proc. of the 11th European Conf. on Computer Vision, vol. III, Springer-Verlag, Berlin, Heidelberg, 243-257, 2010.

2011

[23] Kwolek B., Krzeszowski T., Wojciechowski K.: Swarm intelligence based searching schemes for articulated 3D body motion tracking. Int. Conf. on Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science, vol. 6915, Springer-Verlag, pp. 115-126, 2011.

[24] Świtoński A., Polański A., Wojciechowski K.: Human Identification Based on Gait Paths, Int. Conf. on Advanced Concepts for Intelligent Vision Systems, Lecture Notes in Computer Science, vol. 6915, Springer-Verlag, pp. 531-542, 2011.