Graph-based Particle Filter for Human Tracking with Stylistic Variations

Jesús Martínez del Rincón
http://cism.kingston.ac.uk/people/researcher/1214
Jean-Christophe Nebel
http://cism.kingston.ac.uk/people/academic/423
Dimitrios Makris
http://cism.kingston.ac.uk/people/academic/326

Digital Imaging Research Centre
Kingston University
UK

Articulated human tracking is one of the most active areas in computer vision due to its numerous applications. However, it still remains a major challenge due to the high complexity and dimensionality of the human pose space, which impact negatively on existing trackers by reducing their reliability at reconstructing human-like poses and making recovery from failure impossible.

Such difficulties have led to the development of approaches that address the size of the solution space, either using efficient search strategies such as annealing [1] and space partition [2] or by reducing its dimensionality [3,4,5,6]. Since computational cost increases with space dimensionality, many dimensionality reduction methods (DR) have been developed and explored as prior models for articulated human tracking. However, these processes may result in a loss of generality by compressing important information such as style, intra-activity variance and inter-subject variability.

In this paper we propose to address those natural limitations of model priors for articulated motion tracking. Tracking priors are learned by applying Generalised Laplacian Eigenmaps (GLE) whose capacity to produce general manifolds allows preserving both temporal continuity and stylistic variations among people and activities. In addition, we suggest a novel and integrated Graph-based tracking scheme which is consistent with the prior formation process.

Prior model learning: GLE. We present a new DR methodology, Generalised Laplacian Eigenmaps (GLE), which combines temporal and stylistic information as integral part of its objective function. Both the temporal structure and the style variance of the original data are expressed as complementary constraints, by building neighbourhood graphs G_T and G_S between the training samples. In this manner, local style neighbours as well as local temporal neighbours are placed nearby in the embedded space without the need of enforcing any artificial embedded geometry as in [6].

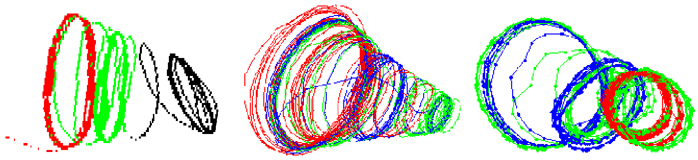

Fig. 1. Manifolds created using GLE. Left: Mocap data from 1 subject containing 3 variations of an activity (walking, fast walking and running). Centre: Mocap data from 3 subjects (red, green and blue) and with 3 variations of an activity per sequence. Right: Foreground masks, profile view, 3 subjects (red, green and blue), two variations (walking and fast walking)

Graph-based PF. This prior embedding supports a specialised particle filter in two ways. First it supplies a propagation model with both temporal (dynamic model) and stylistic constraints so that it can be applied to different scenarios and actors. Secondly it provides automatically a suitable process noise model in the manifold created from training data. This prevents divergence towards invalid poses in the low dimensional space by ensuring motion in the vicinity of the manifold; this also increases the probability of recovering after failure.

Since GLE connectivity graphs regulate the proximity and locality of the poses on the manifold, this connectivity is exploited to propagate and predict plausible hypotheses. Thus, we propose to replace the traditional deterministic propagation and prediction steps of particle filters by a stochastic propagation based on a triple resampling process. This graph-based resampling process associates each particle x_t^i to training points in the manifold m^k and projects particles in time and style based on the temporal G_T and stylistic G_S graphs. Conceptually, given a previous position of a particle x_{t-1}^i, the prediction x_t^i is:

$$p(x_t^i|x_{t-1}^i) \propto p(x_t^i|m^s) \cdot G_S(m^s, m^{k+1}) \cdot G_T(m^k, m^{k+1}) \cdot p(m^k|x_{t-1}^i)$$

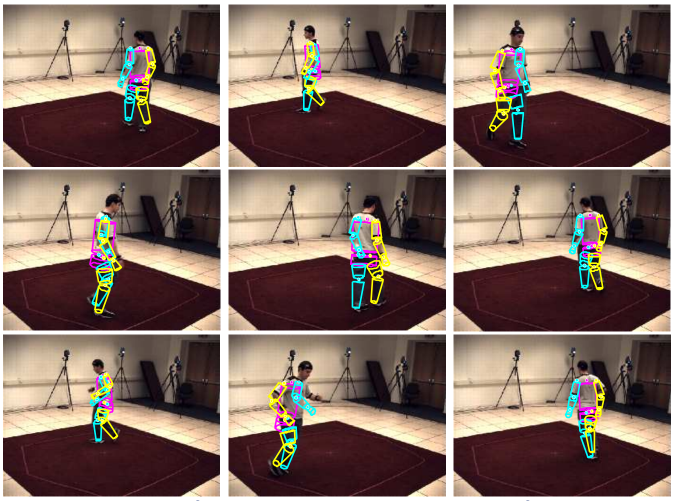

Figure 2: Results for graph-based Particle Filter for S4_Combo_1 (HumanEva II) sequence using bi-directional silhouettes as observation.

Results. The proposed algorithm is validated using a standard and well-known framework for articulated human pose estimation and tracking, HumanEVA. We demonstrate how graph-based propagation provides added robustness preventing divergence from the manifold. Such improvement, in combination with the capacity of GLE for representing stylistic variations, is able to cope with the running phase, the walking phase and their transitions, outperforming other state of the art methodologies.

Table 1: Performance comparison on HumanEVA II. Results are given for frames [1-437] (before divergence of all the methods). Two different observation functions were used, i.e. based on edges plus silhouettes (E+S) and based on bidirectional silhouettes (BS).

Error [cm] (std dev)	S4 _Combo_1
E+S PF	14.1 (6.5)
E+S APF [1]	14.5 (9)
E+S GPLVM-PF [5]	17.58 (10.1)
E+S GLE-PF	13.7 (9.5)
E+S GLE-GbPF	12.2 (7.5)
BS PF*	11.9 (8.1)
BS GLE-GbPF*	11.4 (7.8)

[1] J. Deutscher, A. Blake, and I. Reid, "Articulated Body Motion Capture by Annealed Particle Filtering," in *CVPR*, 2000.

[2] J. MacCormick and M. Isard, "Partitioned sampling, articulated objects, and interface-quality hand tracking," in *ECCV 2*, 2000, pp. 3-19.

[3] N. R. Howe, M. E. Leventon, and W. T. Freeman, "Bayesian reconstruction of 3D human motion from single-camera video," in *NIPS 12*, 2000, p. 820–826.

[4] H. Sidenbladh, M. J. Black, and L. Sigal, "Implicit probabilistic models of human motion for synthesis and tracking," in *ECCV 1*, 2002, p. 784–800.

[5] R. Urtasun, D. J. Fleet, and P. Fua, "Gaussian process dynamical models for 3D people tracking," in *CVPR*, 2006, p. 238–245.

[6] R. Urtasun, D. J. Fleet, and N. Lawrence, "Modeling human locomotion with topologically constrained latent variable models," in *HMUMCA Workshop*, 2007, pp. 104-118.

[7] M. Lewandowski, J. Martínez del Rincón, D. Makris, and J. C. Nebel, "Temporal extension of laplacian eigenmaps for unsupervised dimensionality reduction of time series," in *ICPR*, 2010, pp. 161-164.