

Self organized character animation based on learned synergies from full-body motion capture data

Aee-Ni Park, Albert Mukovskiy, Lars Omlor, Martin A. Giese

Abstract—We present a learning-based method for the real-time synthesis of trajectories for character animation. The method is based on the approximation of complex movements by a small set of spatial movement primitives, or synergies. Applying kernel methods, such learned primitives are associated with nonlinear dynamical systems that, similar to central pattern generators in biological systems, produce highly complex trajectories in real-time by self-organization of behaviour. It is demonstrated that the novel method is suitable for the simulation of highly realistic human movements in interactive character animation. Different simple applications are discussed, including crowd-animation, navigation combined with the dynamic variation of movements styles, and the animation of a folk dance.

I. INTRODUCTION

Generating realistic complex body movements is a core problem in computer graphics and robotics. This problem is particularly difficult for systems with many degrees of freedom, like for full-body animation and in humanoid robots. In addition, many applications in robotics and computer games require methods that are real-time capable and permit an online synthesis of character behaviour. In the computer graphics, motion capture has become the standard approach for the modeling of naturalistic movements. Usually, recorded trajectories are retargeted to kinematic or physical models [1], typically requiring tedious editing of recorded data in order to adapt it to the constraints of the animation. Different approaches have been developed to automatically select segments of motion capture data from large data bases and to concatenate them to longer sequences that fulfil specific constraints that are specified by the animator [2], [3], [4]. Other approaches, based on physical or dynamical models [5], [6], focused on the simulation of scenes with many interactive agents or crowds, navigating autonomously or showing collective behaviours, where the underlying character models are often simplified, resulting in a manageable complexity of the system dynamics and dynamics simulation, but lacking subtle details of realistic human body movements. Some approaches have managed to simulate highly realistic human behaviour using dynamic models combined with sophisticated hierarchical control architectures [7]. However, the design of such systems is complex and the adjustment of their parameters requires much expertise of the animator. In this context, it seems

an interesting challenge to develop simple dynamical architectures for generating complex human movements that integrate the information learned from motion capture data, and which avoid a detailed simulation of all details of the human body dynamics, however resulting in real-time animation with a high degree of realism.

The proposed novel method combines a compact parameterization of movements recorded by motion capture with a dynamical systems approach for the self-organization of interactive behaviour. Using unsupervised learning methods, we approximate classes of full-body movements by combinations of a very small number of hidden source signals. The proposed method provides significantly more compact representations than standard approaches for dimension reduction, like PCA or ICA. In a second step, applying supervised learning, these source signals are associated with the stable solutions of low-dimensional nonlinear dynamical systems whose stability properties can be systematically designed. By coupling of these dynamical systems the coordination between the degrees of freedom within individual characters is maintained, and coordinated behaviour between multiple characters can be simulated.

The proposed approach is biologically inspired and exploits the concept of synergies, which plays a central role in the field of motor control. Synergies specify lower-dimensional control units that typically encompass only a subset of the available degrees of freedom. A decomposition in such low-dimensional sub-units has been proposed as way to solve the 'degrees-of-freedom problem', which arises in the synthesis and control of movements in effector systems with many degrees of freedom [8], [9]. Recent studies in motor control have successfully applied unsupervised learning methods to extract low-dimensional spatio-temporal components from trajectories or EMG signals, which have been interpreted as correlates of synergies [10], [11].

However, for an accurate resynthesis of the trajectories standard methods for dimension reduction, like PCA or ICA, require typically 8-12 components for an accurate approximations of human full-body movements [12]. It is shown in the following that by choice of an appropriate mixture model much more compact representations of human movements can be learned. These representations can be associated with relatively simple dynamical systems for real-time animation with well-defined stability properties. Like 'central pattern generators' in biological systems, these dynamical systems consist of simple basic units that are dynamically coupled and that generate complex coordinated behaviour by self-organization.

All authors are with the Laboratory for Action Representation and Learning, Dept. for Cognitive Neurology, Hertie Institute for Clinical Brain Research, Tübingen, Germany. M.A. Giese is also with the School of Psychology, University of Bangor, UK. {aee-ni.park,albert.mukovskiy,lars.omlor}@medizin.uni-tuebingen.de,martin.giese@uni-tuebingen.de

The contributions of this paper are as follows: (1) Novel method for the learning of highly compact models for the accurate approximations of human full-body movements; (2) method for the mapping of the learned components (synergies) onto dynamical systems with controlled stability properties that are easy to design; (3) demonstration of the potential of this method for a few selected problems from character animation.

In the following we first describe the novel method for the learning of trajectory models (Section II). We then discuss the mapping onto the stable solutions of nonlinear dynamical systems (Section III), and how the model can be augmented by a dynamic model for navigation (Section IV). Some exemplary applications of the method are presented in Section V. Implications of the framework and future extensions are discussed in Section VI.

II. SYNERGY-BASED TRAJECTORY MODEL

A. Motion Capture

Motion data were recorded using a VICON 612 Motion Capture System with 8 cameras and 41 reflecting markers with a sample frequency of 120 Hz. The animations presented in this paper are based on a data set of gaits that includes neutral and emotional straight walking (sad, happy, angry, and fearful), and neutral walking in a circle (with a rotation of 45 deg left and right per double step). The recorded position trajectories were fitted with a hierarchical kinematical model (skeleton) with 17 joints. The rotations between adjacent segments were parameterized as quaternions. The angles given in an axis angle representation served as basis for the modeling by unsupervised learning.

B. Extraction of learned Synergies

Joint angle trajectories, after subtracting the means, were approximated by a weighted mixture of source signals. As shown elsewhere [13], an especially compact model for gait trajectories can be obtained by fitting an anechoic mixture model given by the equation:

$$x_i(t) = \sum_j w_{ij} s_j(t - \tau_{ij}) \quad (1)$$

The functions s_j denote hidden source signals and the parameters w_{ij} are the mixing weights. Opposed to common blind source separation techniques, e.g., PCA or ICA, this mixing model allows for time shifts τ_{ij} of the sources before the linear superposition. Time shifts (delays), source signals and mixing weights are determined by a blind source separation algorithm described in the following.

Classical applications in acoustics for anechoic demixing assume frequently an underdetermined mixing model of the form (1). These are models where the sources outnumber the signals. A new algorithm for the solution of overdetermined problems, i.e., if the number of sources is smaller than the number of signals, can be derived based on the Wigner-Ville transform [14]. The Wigner-Ville spectrum (WVS) of a

random process x is defined by the partial Fourier transform of the symmetric autocorrelation function of x :

$$W_x(t, \omega) := \int E \left\{ x\left(t + \frac{\tau}{2}\right) \overline{x\left(t - \frac{\tau}{2}\right)} \right\} e^{-2\pi i \omega \tau} d\tau \quad (2)$$

Applying this integral transform to equation (1) results in the equation

$$W_x(t, \omega) := \sum_j w_{ij}^2 W_{s_j}(t - \tau_{ij}, \omega) \quad (3)$$

under the assumption that the sources are statistically independent. As two dimensional representation of one dimensional signals, this equation is redundant and can be solved by computing a set of projections onto lower dimensional spaces that specify the same information as the original problem. In short, all parameters in the model (1) can be estimated using the following two steps:

1) Solve:

$$|\tilde{x}_i(\omega)|^2 = \sum_j |w_{ij}|^2 |\tilde{s}_j(\omega)|^2 \quad (4)$$

where \tilde{x}_i and \tilde{s}_j specify the normal Fourier transform of x_i and s_j . This equation can be solved using nonnegative matrix factorization (NMF) or a positive ICA algorithm to obtain the power spectra of the sources.

2) Use the results from the previous step to solve the following equation numerically in order to obtain the phase spectra and delays of the sources:

$$\begin{aligned} |\tilde{x}_i(\omega)|^2 \frac{\partial}{\partial \omega} \arg \{ \tilde{x}_i(\omega) \} = \\ \sum_n |w_{in}|^2 |\tilde{s}_n(\omega)|^2 \left[\frac{\partial}{\partial \omega} \arg \{ \tilde{s}_n(\omega) \} + \tau_{in} \right] \end{aligned} \quad (5)$$

Detailed comparisons for periodic and non-periodic trajectory data show that the model (1) results in more compact trajectory approximations for human movement data, requiring less source terms than models based on standard PCA and ICA for the same level of accuracy. This is illustrated in Fig. that shows the approximation quality as function of the number of sources. Approximation quality was defined by the expression $1 - \frac{\|D-F\|}{\|D\|}$, where D signifies the original data matrix, F its approximation by the source model, where the norm is the Frobenius norm. For the approximation of the gait data basis the new method requires only three source signals to achieve an approximation quality of 97 % while while PCA and ICA require more than 6 sources to achieve the same level of accuracy. See [15] for further details.

This time-frequency approach allows a very accurate estimation of the delays in Fourier domain without the need for discretization. A detailed analysis shows additionally that the mixing weights vary with the emotional style of the movement, while source functions and delays are very similar between different emotions. This makes it possible to morph between different emotions by interpolation of the mixing weights.

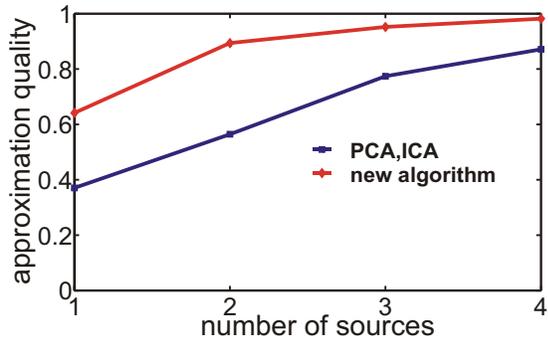


Fig. 1. Comparison of different blind source separation algorithms. Approximation quality is shown as a function of the number of sources for traditional blind source separation algorithms (PCA/ICA) and our new algorithm.

III. DYNAMICAL SYSTEM FOR TRAJECTORY GENERATION

The previous model for the compact approximation of trajectories based on synergies is not suitable for real-time animation, e.g. for computer games, since the trajectory has to be synthesized by superposition and delaying of the complete source signals. A real-time capable algorithm can be devised by specifying a dynamical system that produces the same trajectories as solution. We design such a dynamical system, again exploiting the fact that the movement can be approximated by a superposition of a few basic components. For this purpose, we establish a mapping between the solutions of simple dynamical systems and the sinusoidal source signals of the trajectory representation. The complete trajectory is then generated by a set of such simple dynamical systems, which are coupled to ensure temporal coordination between the different sources. In an abstract sense, the resulting system is similar to a set of coupled 'central pattern generators' in a biological system. Such architectures have been proposed as model for the generation of gaits and other motor patterns [16], [17]. In the following, we first introduce the attractor dynamics and discuss how the mapping between its solutions and the source signals of the trajectory representation is learned. Finally, we demonstrate how dynamic couplings can be introduced that stabilize the coordination within single characters, and which are suitable for the simulation of coordinated behaviour of multiple avatars.

1) *Attractor Dynamics*: Often, behaviour can be mapped onto stable solutions of dynamical systems [18]. Various applications in the field of robotics (e.g. [16], [17]) successfully demonstrate this procedure. In the experiments of this paper we simulate only periodic gaits, while our approach generalizes for non-periodic movements. Limit cycle oscillators seem to be a natural choice as dynamics for the generation of such periodic behaviours (e.g. [16], [17]). As basis elements of our architecture we use a van der Pol oscillator. This oscillator can be understood as a harmonic oscillator with an amplitude-dependent damping.

Its dynamics is given by the differential equation:

$$\ddot{y}(t) + \lambda \left(y(t)^2 - k \right) \dot{y}(t) + \omega_0^2 y(t) = 0 \quad (6)$$

The parameter ω_0 determines the eigenfrequency of the oscillator, and the parameter k the amplitude of the stable limit cycle. After perturbations the state will thus return to this attractor. In addition, the structure of the dynamics does not fundamentally change for moderate couplings with other system components (structural stability). For more details see [19].

2) *Mapping of phase space onto source signals*: In order to generate the source signals in real-time, we construct nonlinear mappings between the attractor solutions of the van der Pol oscillators and the values of the source signals. These mappings are learned by Support Vector Regression (SVR) [20]. For each of the three sinusoidal source signals of the model (1) a separate oscillator is introduced. In order to model the delays for joint angles (17 joints \times 3 angles resulting in 51 joint angles) we introduce new auxiliary signals for the individual sources with different delays according to the relationship:

$$\tilde{s}_{ij}(t) = s_j(t - \tau_{ij}) \quad (7)$$

For each of these modified source signals we construct a separate mapping from the phase space of the van der Pol oscillator that corresponds to the source s_j , defined by the variable pair $\mathbf{y}_j(t) = [y_j, \dot{y}_j(t)]$ and the values of the source function. The mapping is defined by the nonlinear functions:

$$\tilde{s}_{ij}(t) = f_{ij}(\mathbf{y}_j(t)) \quad (8)$$

The nonlinear functions f_{ij} are learned by SVR from T training data pairs $\{\mathbf{y}_j(t_k), s_{ij}(t_k)\}_{1 \leq k \leq T}$ that are derived by sampling one cycle of the stationary solution of the oscillator and the modified source signal equidistantly over time. After the functions f_{ij} have been learned, the dynamics corresponds to three limit cycle oscillators with nonlinear instantaneous observers that map the state of each oscillator onto the corresponding set of delayed source signals. These signals then are linearly combined according to (9). The complete reconstruction of the joint angle trajectories requires the addition of the average joint angles m_i , which have subtracted from the data before the blind source separation (Figure 2):

$$x_i(t) = m_i + \sum_j w_{ij} \tilde{s}_{ij}(t) \quad (9)$$

3) *Dynamic couplings*: In order to stabilize the timing-relationship between the different sources ('synergies') we introduced dynamic couplings between the three oscillators that drive the source signals of the same avatar. It has been shown applying contraction theory that stable behaviour of such oscillator networks can be accomplished by introduction velocity couplings [21].

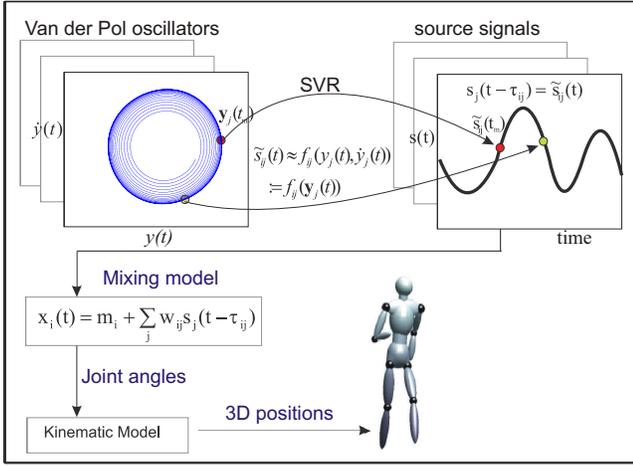


Fig. 2. Illustration of the dynamics for real-time animation. The phase space of the limit cycle oscillator is mapped onto the time-shifted source signals using Support Vector Regression. Joint angles are synthesized by combining the signals linearly according to the learned mixture model (1). A kinematic model converts the joint angles into 3-D positions for animation.

The couplings within the same avatar were given in the form (cf. Figure 3):

$$\begin{aligned}
 \ddot{y}_1 + \lambda (y_1^2 - k) \dot{y}_1 + \omega_0^2 y_1 &= \\
 &\alpha (\dot{y}_2 - \dot{y}_1) + \alpha (\dot{y}_3 - \dot{y}_1) \\
 \ddot{y}_2 + \lambda (y_2^2 - k) \dot{y}_2 + \omega_0^2 y_2 &= \\
 &\alpha (\dot{y}_1 - \dot{y}_2) + \alpha (\dot{y}_3 - \dot{y}_2) \\
 \ddot{y}_3 + \lambda (y_3^2 - k) \dot{y}_3 + \omega_0^2 y_3 &= \\
 &\alpha (\dot{y}_1 - \dot{y}_3) + \alpha (\dot{y}_2 - \dot{y}_3)
 \end{aligned} \tag{10}$$

Such dynamic couplings can also be exploited for simulating coordinated behaviour of multiple characters and crowds, for example, to enforce that multiple avatars walk in synchrony (e.g. soldiers in lock-step). To implement such couplings we only connected the oscillators assigned to the source with the lowest frequencies (oc1). By introducing unidirectional couplings it is possible to make multiple characters following one, who acts as a leader [21].

4) *Dynamic variation of movement style*: The proposed model for the real-time generation of trajectories permits style morphing in a straight-forward way. To interpolate between two movement styles (a) and (b), e.g. neutral and emotional walking, the mixing weights and the mean joint angles are linearly interpolated. For the weights w_{ij} this results in the equation:

$$w_{ij}(t) = \lambda(t) w_{ij}^a + (1 - \lambda(t)) w_{ij}^b \tag{11}$$

The time-dependent morphing parameter $\lambda(t)$ specifies the movement style. Additionally, the gait speed is adjusted by interpolating the eigenfrequencies of the oscillators:

$$\omega_0(t) = \lambda(t) \omega_0^a + (1 - \lambda(t)) \omega_0^b \tag{12}$$

The same type of morphing was also applied to implement direction changes of the avatars for navigation, morphing

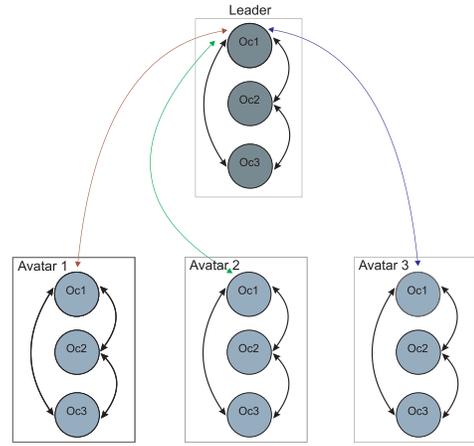


Fig. 3. Coupling of multiple avatars, each of them comprising three coupled oscillators.

between straight and curved gait steps. The change of the morphing parameter can be made dependent on the behaviour of other avatars in the scene, e.g. to influence the emotional style dependent on the distance d of the avatar from another 'dangerous' colleague, introducing a distance-dependent morphing weight $\lambda(t) = g(d(t))$. Similarly, the eigenfrequency parameter ω_0 can be made dependent on the distance from other agents. In this way, the walking speed of the characters can be made dependent on the behaviour of the others.

IV. NAVIGATION

For the simulation of realistic interactive character behaviour, navigation and obstacle avoidance have to be included in the system for real-time animation. From the generated articulated motion of the avatars, first, the translatory motion and the rotation of the hip in the horizontal plane is computed by enforcing the floor contact constraints of the feet. In this way, we generate the propagation and heading direction of the avatars indirectly from the joint angle movements. By interpolating between straight and curved walking we were able to simulate walking with different curved walking paths. The addition of a navigation dynamic makes it possible to simulate avatars that move towards specific targets in space, or to avoid obstacles and other avatars.

A. Computation of translation and heading

The pelvis forms the root of the kinematic chain of our avatar model. The core of the algorithm for computing the avatar's propagation is to adjust the horizontal pelvis translation and rotation in order to fulfil the constraint that the feet that make contact to the ground do not translate. For this purpose, we first detect whether the feet make ground contact using a simple threshold criterion for the vertical position of the foot centers. This criterion works well when the original motion capture data did not contain foot slipping. We add a differential translation and horizontal rotation to the pelvis coordinate system in order to minimize foot slipping.

In addition, the vertical position of the pelvis is adjusted by assuming that the foot forms the lowest point of the figure.

B. Navigation dynamics

We use a simplified version of a dynamic navigation model which has been applied successfully in robotics before [22], [23]. The temporal change of the heading direction φ_i of avatar i is given by the differential equation:

$$d\varphi_i/dt = h^{goal}(\varphi_i, \mathbf{p}_i, \mathbf{p}_i^{goal}) + \sum_j h^{avoid}(\varphi_i, \mathbf{p}_i, \mathbf{p}_j) + \sum_j h^{pcoll}(\varphi_i, \varphi_j, \mathbf{p}_i, \mathbf{p}_j) \quad (13)$$

The right-hand side of this equation has three different components, where \mathbf{p}_i denotes the position of character i . The first term

$$h^{goal}(\varphi_i, \mathbf{p}_i, \mathbf{p}_i^{goal}) = \sin\left(\arctan\left(\frac{(\mathbf{p}_i - \mathbf{p}_i^{goal})_2}{(\mathbf{p}_i - \mathbf{p}_i^{goal})_1} - \varphi_i\right)\right) \quad (14)$$

models a tendency to move in the direction of the goal, where the goal of avatar i is specified by the 2D position vector \mathbf{p}_i^{goal} .

The second term implements obstacle avoidance, where the obstacles are defined by the other avatars. This term is given by:

$$h^{avoid}(\varphi_i, \mathbf{p}_i, \mathbf{p}_j) = \sin(\Delta\varphi_{ij}) \cdot \exp\left(-\frac{\Delta\varphi_{ij}^2}{2\sigma_\varphi^2}\right) \cdot \exp\left(-\frac{d_{ij}^2}{2\sigma_d^2}\right) \quad (15)$$

with $\Delta\varphi_{ij} = \varphi_i - \arctan((\mathbf{p}_i - \mathbf{p}_j)_2/(\mathbf{p}_i - \mathbf{p}_j)_1)$ and $d_{ij} = |\mathbf{p}_i - \mathbf{p}_j|$.

Much more realistic collision avoidance is accomplished by inclusion of a third term in the dynamics that controls the heading direction. This term has the same form as the previous term.

$$h^{pcoll}(\varphi_i, \varphi_j, \mathbf{p}_i, \mathbf{p}_j) = \sin(\Delta\varphi_{ij}) \cdot \exp\left(-\frac{(\Delta\varphi_{ij}^{pc})^2}{2\sigma_\varphi^2}\right) \cdot \exp\left(-\frac{(d_{ij}^{pc})^2}{2\sigma_d^2}\right) \quad (16)$$

However, d_{ij}^{pc} and $\Delta\varphi_{ij}^{pc}$ are computed like d_{ij} and $\Delta\varphi_{ij}$ replacing the point \mathbf{p}_j by the predicted collision point of the avatar i with avatar j .

The walking direction of the characters is changed by interpolation between straight walking and walking along curved paths to the left for $d\varphi_i > 0$, and to the right for $d\varphi_i \leq 0$ using (11). The morph parameter $\lambda(t)$ was taken proportional to $|d\varphi_i|$, normalizing in a way that ensures $\lambda = 1$ for the maximum possible value of this derivative. The heading direction $\varphi_i(t)$ generated by the navigation dynamics was low-pass filtered with a time constant equal to one step cycle to improve the smoothness of the navigation behaviour.

V. RESULTS

The proposed technique for motion capture-based self-organized simulation of character behaviour has a broad application spectrum, which can only be partially explored in this paper. In the following, we present a number of example applications that highlight possibilities of the approach.

A. Approximation quality

In general, the blind source separation method discussed in Section II provides accurate approximations of the joint trajectories by using less components than other standard methods for dimension reduction. Beyond the quantitative analysis in Fig. 1, the influence of approximation quality on animations is illustrated in by an attached *Demo*¹.

As further evaluation of the method, we compared the reconstruction errors of the model for different conditions, including on-line and off-line models. For this purpose, we extracted 3 shift invariant source signals from the 7 motion-captured emotional walking gaits (including left and right turning walks). We compared the reconstructed angles trajectories of neutral straight walking of synthesized gaits with the original motion-captured trajectories of the same style (corresponding to avatar 1). The second case was the off-line generation of the trajectory by the blind source separation model (1), used to animate avatar 2. The movements of avatar 3 were generated by the simplest possible on-line model, mapping the phase spaces of 3 coupled oscillators by Support Vector Regression onto the source signals. The propagation speed of this avatar deviates slightly (a few percent) from the avatar 1 due to approximation errors.

Finally, we created a more complex online situation where another avatar (avatar 4) follows avatar 1 implementing a distance-to-frequency coupling in order to keep the distance between the avatars constant (cf. Section III-4). The distance-frequency coupling was implemented by a sigmoidal function:

$$\lambda(t) = g(d(t)) = \frac{1}{1 + \exp(-\gamma d(t))} \quad (17)$$

with the positive constant γ . The parameters ω_0^b and ω_0^b in (12) were appropriately adjusted.

The normalized errors (unexplained variance) between the trajectories in joint angle space of avatars 2, 3 and 4 compared with avatar 1 (original trajectory) were: 6.89% (2-1), 8.62% (3-1), and 7.51% (4-1). This implies that the major error is introduced by the source approximation, while the support vector regression has a much smaller influence.

B. Coordination of a crowd

A first example is illustrated in Figure 4. A crowd of characters is first locomoting individually without fixed phase relationship. An instruction to move synchronously (in lock-step) is introduced, modeled by introducing couplings between the oscillator triples of the individual avatars, taking

¹www.uni-tuebingen.de/uni/knv/arl/avi/compareN.avi
www.uni-tuebingen.de/uni/knv/arl/avi/compareH.avi

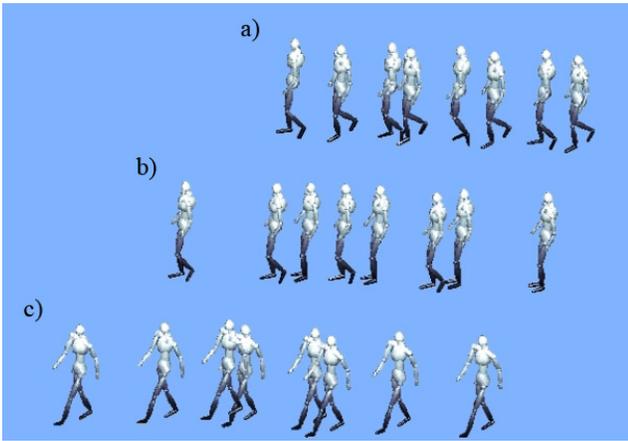


Fig. 4. Synchronization of the gaits within a crowd. (a) Avatars start with self-paced walks that are out of phase. After a transitory period (b), the gaits of the characters become completely synchronized (c).

one oscillator as a leader (Section 3). For appropriate coupling, the transition between uncoordinated and coordinated crowd behaviour is quite short (less than three steps) and looks quite natural. After the transition, characters adapt walking behaviour that looks highly realistic [Demo²].

The same example is useful to illustrate the benefit of more compact generative trajectory models for on-line animation. We compared the synchronisation behaviour of two avatars (see [Demo³]). One of them (avatar 1) was driven by three oscillators corresponding to the source signals extracted by the discussed blind source separation algorithm. The other avatar (avatar 2) was driven by 5 oscillators that were mapped onto PCA components of the trajectories, in the stationary state resulting in a similar approximation quality. Both avatars were coupled to a third avatar (leader) with the same coupling strength. Starting the avatars 1 and 2 with equal initial phases which differ from the phase of avatar 3, it shows that the model with three oscillators synchronizes faster and produces more natural-looking behaviour.

The same example also nicely demonstrates that the model, even though it has been trained in the attractor of the dynamics, generalizes in a meaningful way to transient states. This is illustrated by the phase diagram in Fig. 5 that illustrates the trajectories of the first leading oscillators of the two avatars. First, this figure shows that the oscillators during the synchronization period are not near the attractor, requiring the system to generalize to trajectories that were not used during the raining of the system. Second, it is obvious that the system with less oscillators (red trajectory) returns faster to the attractor than the system with 5 oscillators (green trajectory). More detailed investigations testing dependence on coupling strength and coupling structure are underway. Since the linear weights of the PCA components are not sparse the mixing of the many oscillator inputs results in jerky motion in the starting phase of the synchronization period.

²www.uni-tuebingen.de/uni/knv/ar/avi/synchronization.avi

³www.uni-tuebingen.de/uni/knv/ar/avi/transient.avi

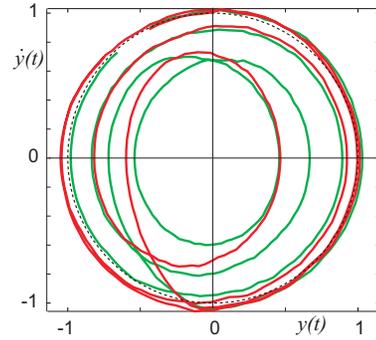


Fig. 5. Phase plot of the first (principal) oscillators of two avatars that are coupled to a leader. The diagram shows the state space of the oscillator in the $[y(t), \dot{y}(t)]$ plane. The green line corresponds to the avatar driven by 5 PCA components, and the red line to the avatar animated with 3 sources extracted with our algorithm. The coupling is switched on after short starting period.

C. Navigation and style change

The second example [Demo⁴] is illustrated in Figure 6. A group of avatars that meets in the center of the scene changes their affect upon the contact with the other characters. This behaviour was implemented by making the affect of each avatar dependent on the distance from the others. In addition, the avatars avoid each other, due to the navigation dynamics described in section IV. In this simulation, navigation and changes of emotional styles were combined, based on only three prototypical gaits: neutral walking with rotation right or left, and emotional walking. In order to produce the emotional gait for blending with left and right neutral paces, we first created an intermediate balanced mixture by interpolating the mixing weights. Corresponding with section IV, gait morphing was based on a piece-wise linear interpolation dependent on the sign of the change of the heading direction $d\varphi/dt$.

D. Animation of a 'folk dance'

The next example is a self-organized scenario, where eight pairs of avatars execute a type of 'folk dance' that requires them to walk in synchrony to the music along a straight line, forming a corridor. Once the couples have reached the end of the corridor, they are required to walk quickly to the other end and re-enter the corridor from the other side. At the point of reentrance the individual avatars need to synchronize with their partner and the other avatars within the corridor. In spite of this relatively complex scenario quite natural behaviour of the whole group of 16 avatars can be self-organized [Demo⁵].

To simulate this complex behaviour, we divided this scenario into four main sectors that are described separately in the following (see Figure 7):

- 1) At the entrance of the corridor the characters gather and wait for the corresponding partner to move synchronously. To simulate the synchronized walking of

⁴www.uni-tuebingen.de/uni/knv/ar/avi/avoidance.avi

⁵www.uni-tuebingen.de/uni/knv/ar/avi/dance.wmv

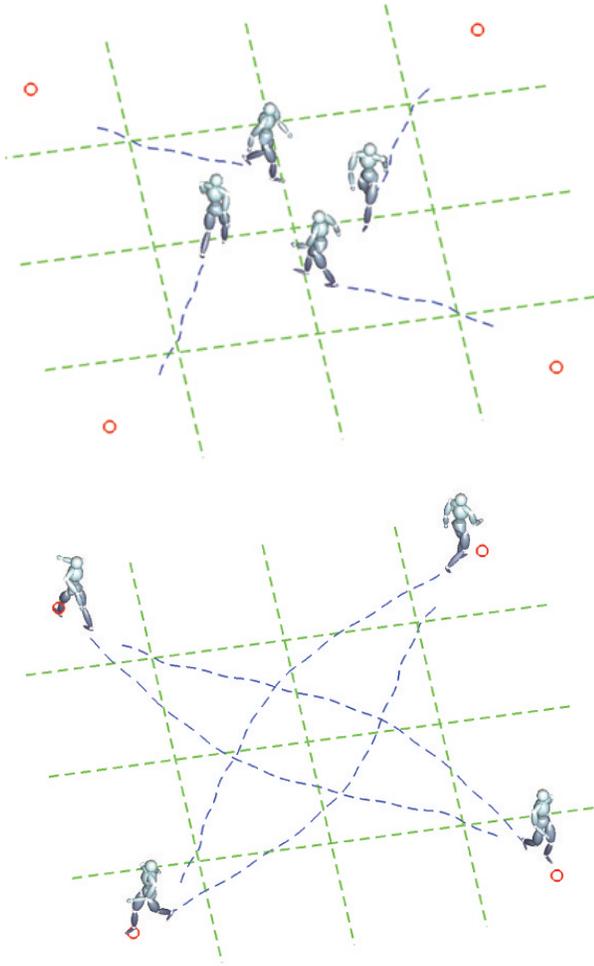


Fig. 6. Avoidance behaviour and change of emotional style. a) At the meeting point their emotional styles change from sad to happy. In addition, the characters avoid each other. b) Avatars heading towards their goals (red circles) with a happy emotion.

all couples within the corridor the corresponding oscillators are coupled similar to equation (10), introducing couplings not only between the avatars of one couple but also between the subsequent couples within the corridor.

- 2) At the end of the corridor reaching the second zone the two avatars of each couple separate and the coupling between their oscillators is removed. Their movements from this point are asynchronous and determined by the the navigation dynamics according to Section IV. In addition, within this zone the emotional walking style of the characters changes from happy to neutral. The curved walking paths are generated by defining appropriate goal points.
- 3) Along the straight paths outside the corridor the avatars accelerate to catch up with their partner at the beginning of the corridor in time, simulated by a temporary increase of the eigenfrequency of the corresponding oscillators.
- 4) In the last zone the characters decelerate, modeled by

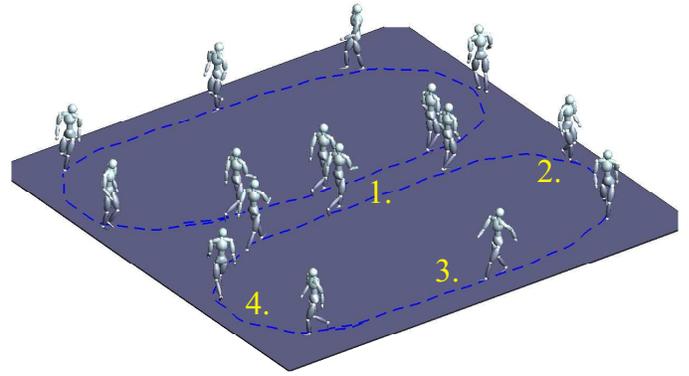


Fig. 7. Simulation of a 'folk dance' behaviour is organized within four zones: 1) Avatars within the corridor move in synchrony, simulated by coupling the corresponding oscillators. 2) Leaving the corridor the avatars of the individual dancing couple become separated and start to move along curved paths. Their emotional style changes from happy to neutral. 3) Avatars walk asynchronously and 'hurry up' to reach the entrance of the corridor, simulated by increasing the eigenfrequency of the oscillators. 4) When the avatars approach the entrance of the corridor their walking speed decreases, and they need to get in synchrony again with the appropriate leg. Both can be simulated by appropriate adjustment of the oscillator frequencies.

decreasing the frequency of the oscillators. A difficult problem is the re-synchronization with the correct foot at the entrance of the corridor. This is accomplished by slightly adjusting the oscillator frequencies to ensure re-synchronization with the appropriate leg. Again, the curved paths are generated by defining appropriate goal point exploiting the navigation dynamics (13).

VI. CONCLUSION AND FUTURE WORK

We presented a new framework that links the synthesis of realistic human movements based on motion capture data with the self-organization of behaviour using nonlinear dynamical systems. The proposed method exploits biologically inspired concepts and is based on the learning of highly compact models for human movement trajectories by a superposition of learned 'synergies', which are controlled by nonlinear attractors dynamics. By introducing of appropriate dynamic couplings complex realistic looking behaviour, reproducing the fine structure of human movements, could be self-organized, even for complex scenarios. Presently, we extend this approach also for non-periodic movements. In addition, we work on scenarios with contact between multiple avatars inducing additional kinematic constraints.

Future work will try to exploit more the synergy concept, learning sparse components that encompass only limited sets of degrees of freedom. This will potentially result in more flexible control of motion styles. In addition, such components offer the possibility to make individual synergies reactive and dependent on external constraints, potentially providing a basis for a much more fine-grained adaptation of the generated behaviour to external constraints. In this context, possible applications in robotics will be considered.

VII. ACKNOWLEDGMENTS

Supported by DFG Forschergruppe 'Perceptual Graphics' and EC FP7 project COBOL. We thank C. Roether for help with the acquisition of the motion capture data. Additional support was provided by the Hermann und Lilly-Schilling-Stiftung. We are grateful to T. Flash and J.J. Slotine for many interesting discussions and W. Strasser for his support.

REFERENCES

- [1] M. Gleicher, "Retargetting Motion to New Characters", *Proc. ACM SIGGRAPH, Int. Conf. on Comp. Graph. and Interactive Techniques*, 1998, pp. 33-42.
- [2] M. Gleicher, H.J. Shin, L. Kovar and A. Jepsen, "Snap-together motion: Assembling run-time animation", *ACM Trans. on Graph. SIGGRAPH*, 2003, pp. 702-702.
- [3] O. Arikan, D.A. Forsyth and J.F. O'Brien, "Motion synthesis from annotations", *ACM Trans. on Graph. SIGGRAPH*, 2003, pp. 402-408.
- [4] J. Chai and J.K. Hodgins, "Performance Animation from Low-dimensional Control Signals", *ACM Trans. on Graph. SIGGRAPH*, 2003, pp. 702-702.
- [5] R. Grzeszczuk, D. Terzopoulos and G. Hinton, "NeuroAnimator: Fast neural network emulation and control of physics based models", *Proc. ACM SIGGRAPH, Int. Conf. on Comp. Graph. and Interactive Techniques*, 1998, pp. 9-20.
- [6] W. Shao and D. Terzopoulos, "Autonomous pedestrians", *Proc. ACM SIGGRAPH*, 2005, pp. 19-28.
- [7] J.K. Hodgins, W.L. Wooten, D.C. Brogan and J.F. O'Brien, "Animating human athletics", *Proc. ACM SIGGRAPH*, 1995, pp. 71-78.
- [8] N.A. Bernstein, *The coordination and regulation of movements*, Pergamon Press, Oxford, 1967.
- [9] T. Flash and B. Hochner, "Motor primitives in vertebrates and invertebrates", *Curr. Opin. Neurobiol.*, vol.15, no.6, 2005, pp. 660-666.
- [10] A. d'Avella and E. Bizzi, "Shared and specific Muscle synergies in neural motor behaviours", *Proc. Natl Acad Sci USA*, vol.102, no.8 2005, pp. 3076-3081.
- [11] Y. Ivanenko, R. Poppele and F. Lacquaniti, "Five basic muscle activation patterns account for muscle activity during human locomotion", *J Physiol*, vol.556, 2004, pp. 267-282.
- [12] A. Safanova and J.K. Hodgins and N.S. Pollard, "Synthesizing physically realistic human motion in low-dimensional, behaviour-specific spaces", *ACM Trans. on Graphics, Proc. SIGGRAPH*, 2004, pp. 514-521.
- [13] L. Omlor and M. Giese, "Extraction of spatio-temporal primitives of emotional body expressions", *Neurocomputing*, 2007, pp. 10-12.
- [14] G. Matzand and F. Hlawatsch, "Wigner distributions (nearly) everywhere: Timefrequency analysis of signals, systems, random processes, signal spaces, and frames", *Signal Processing*, vol.83, no.7, 2003, pp. 1355-1378.
- [15] L. Omlor and M.A. Giese, "Blind source separation for over-determined delayed mixtures", *Advances in Neural Information Processing Systems*, 2006, pp. 1049-1056.
- [16] A. Ijspeert, J. Nakanishi and S. Schaal, "Learning Attractor Landscapes for Learning Motor Primitives", *NIPS*, 2002, pp. 1547-1554.
- [17] S. Schaal, A. Ijspeert and A. Billard, "Computational approaches to motor learning by imitation", *Philos Trans R Soc Lond B Biol Sci.*, vol.358, no.1431, 2003, pp. 537-547.
- [18] J.A.S. Kelso, *Dynamic Patterns: The Self-Organization of Brain and Behaviour*, MIT Press, Cambridge, 1995.
- [19] A.A. Andronov, A.A. Vitt and S.E. Khaikin, *Theory of oscillators*, Dover Publ. Inc., New York, 1987.
- [20] V. Vapnik, *Statistical Learning Theory*, Wiley-Interscience, New York, 1998.
- [21] W. Wang and J.J.E. Slotine, "On Partial Contraction Analysis for Coupled Nonlinear Oscillators", *Biological Cybernetics*, vol.92, no.1, 2005, pp. 38-53.
- [22] G. Schöner, M. Dose and C. Engels, "Dynamics of behaviour: Theory and applications for autonomous robot architectures", *Robotics and Autonomous Systems*, vol.16, no.2-4, 1995, pp. 213-245.
- [23] W.H. Warren, "The Dynamics of Perception and Action", *Psychological Rev.*, vol.113, no.2, 2006, pp. 213-245.