Proceedings of the International Symposium on AI-Inspired Biology, Jackie Chappell, Susannah Thorpe, Nick Hawes and Aaron Sloman (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

Generation of Biological Motion Patterns and Their Emergent Transitions with a Trainable Neural Network Oscillator

Emilian Lalev [1]

Abstract. Biological Motion perception is the ability to perceive gestalts of moving human or animal figures from two-dimensional projections of their joints. These projections are very salient stimuli. A neural network system for generation of such sequences and transitions between them is described. The network can produce several sequential motion patterns learned from human motion capture data. Realistic transitions between them can then be emergently generated while switching between modes of oscillation (e.g., from one type of gait to another). This system might be further explored as a cognitive model for human motion generation showing that motions can be stored in exactly the same structure.

1 Introduction

Biological Motion (BM) perception was first investigated by [8, 9]. It refers to the ability to perceive moving human or animal figures only given the projections of their joints on a 2D plane (the so-called point-light displays) [8]. Human movements contain many degrees of freedom and include both rigid and nonrigid motion which makes them very complex. BM carries rich information about the perceived object: First, it works for identification of anumals and other people from their surrounding [14]. The observed body kinematic structure is revealed and the identity of the living being is classified [2]. BM also reveals the particular type of action being performed, such as walking, running or dancing. And finally, perception of BM can even help define qualitative characteristics of the agent such as gender, weight, mood, etc. [4].

The phenomena related to BM perception and production are approached by several computational and statistical methods. The so called ”morphable models” [6, 10] have been applied for the purposes of motion editing [1], re-targeting motion from one character to another [7], tracking a human figure from video data [11], or recognizing type and style of gait patterns [6]. Fourier techniques [13, 3] are based on analyses of sinusoidal oscillations and are applied for periodic motions, such as locomotion patterns.

In addition to the former approaches, an artificial neural network system was designed to process BM [12]. In the system, periodic point-light displays are approximated as combinations of complex sinusoidal functions. The advantage of such a system relative to similar existing ones is that generalization properties of neural networks seem able to emergently solve particular problems of motion generation and transitions, thus reducing computational demands and efforts for the user.

[1] Central and Eastern European Center for Cogntitive Science, New Bulgarian University, Sofia Bulgaria, 1618 Montevideo Str., email: elalev@Cogs.nbu.bg

2 System description

An Elman recurrent network [5] and one feed-forward network are the parts of the system. The recurrent network (BM Generator) is used as a trainable oscillator. One single network generates each of several learned point-light displays and transitions. It is optimized with the feed-forward network (Transition Filter, representing the physical restrictions of the moving body) to ensure the smoothness of transitions. The work of the Transition Filter is to try to associate only known body postures from novel and noisy frames initially obtained from the Generator during transitions.

2.1 Networks

The BM Generator (Figure 1) has 300 hidden units. The Transition Filter (also included in Figure 1) has 150 hidden units. Both networks have tahn-sigmoid transfer functions.

2.2 Inputs and Targets

As training data for the system, human motion capture data of several motions is used. 15 markers are chosen/computed from the original recorded markers [13]. The coordinate system that is used is object-centered. The number of frames for a particular motion depends on the speed of motion execution. The temporal resolution is 24 frames per second.

From every motion capture record, one locomotion cycle is extracted such that the dynamic pattern can be repeated infinitely.

The input training data for the Generator is composed of the classification vector (or the label) of a given motion repeated constantly in time, quite similar to bias connections attached to the input nodes (see Figure 1).

The training inputs and targets for the Transition Filter are identical as it learns to auto-associate each known frame of the trained motions. During the work of the system, the Transition Filter takes as input the output from the Generator (see Figure 1).

2.3 Training

An algorithm for adaptive training goal changes the goal according to the current performance. The networks are trained to fit 2 different walking patterns, and one running pattern as the mean squared error does not exceed 0.001.

Proceedings of the International Symposium on AI-Inspired Biology, Jackie Chappell, Susannah Thorpe, Nick Hawes and Aaron Sloman (Eds.), at the AISB 2010 convention, 29 March – 1 April 2010, De Montfort University, Leicester, UK

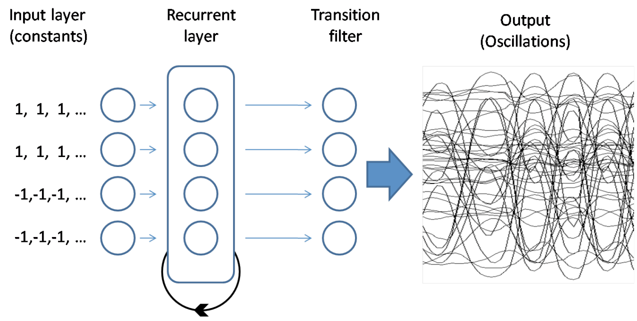

Figure 1. Architecture of the BM Generator and its inputs and targets with the transitions filter.

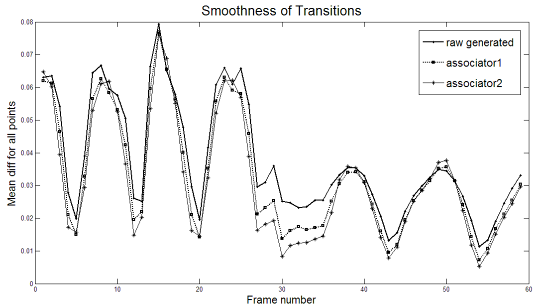

Figure 2. Differentiated coordinates' trajectories as a measure for the smoothness of transitions from the BM Generator - raw or processed once or twice by the transition filter.

3 Results from the BM Generator

1. It learns to oscillate infinitely each of several dynamical patterns.
2. The Generator can perform switching between different BM sequences when the input is changed accordingly. Depending on the exact frame that switching happens, different emergent transitions may be observed.
3. Despite the fact that the BM generator is not trained to any transitions, it generates them quite plausibly in most cases. One such example is accelerating during switching from walking to running. It is shown in the short animation $transitions1.gif$ at http://sites.google.com/site/emilianzaharievlalev/upload/
4. Transitions are additionally smoothed and restricted to realistic transformations by the Transiton Filter network (Figure 1) which reproduces the coordinates of each motion frame. Filtered transitions are qualitatively smoother and more realistic than the ones that come directly from the network's output. A short animation $transitionsFilteredAndRaw.gif$ showing this is available at http://sites.google.com/site/emilianzaharievlalev/upload/

Smoothness of transitions can also be quantitatively measured by the mean absolute value of differentiated oscillations (Figure 2). In the figure, a transition is initiated at frame 21 and ends about frame 40. Obviously, the usage of the auto-associator (associator1) smoothes the sharp changes in this period. Applying it once again (associator2), smoothing of the border between the motions becomes even better.

4 Conclusion

I showed implementation of recurrent neural networks in processing of real human motion capture data. Recurrent networks can be utilized as trainable oscillators of different body motions. Switching between oscillations is smooth and very realistic. Such smooth transitions are due to the recurrent activations in the network remaining from the previous oscillating regime. They work as initial context for the new oscillation.

Networks learn dynamics in BM as approximated functions. They develop several dynamical attractors, resulting in oscillations representing separate locmotion patterns. Maybe this system has common grounds with the biological mechanisms for oscillatory motions like human and animal gaits. The basic inference would be that a particular and exactly the same brain structure may be responsible for several different command tasks.

ACKNOWLEDGEMENTS

I would like to acknowledge Prof. Dr. Nikolaus Troje, head of Biomotion Lab, Queen's University, Kingston, Ontario, Canada for providing me with human motion capture data.

REFERENCES

[1] A. Bruderlin and L. Williams, 'Motion signal processing', in *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles*, pp. 97–100, (1995).
[2] J. E. Cutting and L. T. Kozlowski, 'Recognizing friends by their walk: Gait perception without familiarity cues', *Bulletin of the Psychonomic Society*, **9(5)**, 353–356, (1977).
[3] J. Davis, A. Bobick, and W. Richards, 'Categorical representation and recognition of oscillatory motion patterns', in *IEEE Conference on Computer Vision and Pattern Recognition, Hilton Head Island, SC*, pp. 628–635, (2000).
[4] W. H. Dittrich, T. Troscianko, S. E. G. Lea, and D. Morgan, 'Perception of emotion from dynamic point-light displays represented in dance', *Perception*, **25(6)**, 727–738, (1996).
[5] J.L. Elman, 'Finding structure in time', *Cognitive Science*, **14**, 179–211, (1990).
[6] M. A. Giese and T. Poggio, 'Morphable modelsfor the analysis and synthesis of complex motion patterns', *International Journal of Computer Vision*, **38**, 59–73, (2000).
[7] M Gleicher, 'Retargetting motion to new characters', in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, Orlando*, pp. 32–42, (1998).
[8] G. Johansson, 'Visual perception of biological motion and a model for its analysis', *Perception and Psychophysics*, **14**, 201–211, (1973).
[9] G. Johansson, 'Spatio-temporal differentiation and integration on visual motion perception', *Psychological Research*, **38**, 379–393, (1976).
[10] M. J. Jones and T. Poggio, 'Multidimensional morphable models: A framework for representing and matching object classes', *International Journal of Computer Vision*, **29**, 107–131, (1999).
[11] S. X. Ju, M. J. Black, and Y. Yacoob, 'Cardboard people: A parameterized model of articulated image motion', in *Proceedings of the International Conference on Face and Gesture, Vermont*, (1996).
[12] E. Lalev, 'Connectionist framework for processing of biological motion sequences and their transitions', in *Proceedings of the 20th Irish Conference on Artificial Intelligence and Cognitive Science AICS2009*, eds., John Dunnion Lorcan Coyle and Jill Freyne, pp. 198–209, (2009).
[13] N. F. Troje, 'Decomposing biological motion: A framework for analysis and synthesis of human gait patterns', *Journal of Vision*, **2**, 371–387, (2002).
[14] N. F. Troje and C. Westhoff, 'The inversion eff ect in biological motion perception: Evidence for a 'life detector'?', *Current Biology*, **16(8)**, 821–824, (2006).