Motion Processing Using
Variable Harmonic Components

Jonas Gomes[1] Luiz Velho[1] Fernando Wagner da Silva[1,2] Siome Klein Goldenstein[3]

[1] VISGRAF Laboratory, IMPA, Rio de Janeiro, Brazil *

[2]LCG/UFRJ – COPPE Sistemas, Rio de Janeiro, Brazil *

[3]VAST Laboratory - University of Pennsylvania, USA †

Abstract

This paper discusses the problem of motion processing and proposes the use of a mathematical model, which describes a motion signal as a path with variable harmonic components. We show that this model is quite suited to representing a motion path as a periodic part with additional motion texture (noise), as described in [18]. This provides a good mathematical formulation to the idea of "motion content". We also describe some mathematical tools that can be used with the model in order to construct different motion paths filters. As an aplication, we describe an algorithm to change automatically the time duration of a motion. We also point to interesting research directions on motion processing, using the model here introduced.

CR Categories: I.3.7[Computer Graphics]: Three-dimensional graphics–Animation; I.4.8[Image Processing and Computer Vision]: Scene analysis–Time varying images.

Additional Keywords: Motion processing, Motion editing, Windowed Fourier Transform, Gabor wavelets, Lapped Cosine Transform, Cyclification, Reparametrization, Harmonic components.

1 Introduction

Motion processing uses operations over motion paths in order to modify an existing animation. With the widespread use of motion capture devices, the area of motion processing has turned out to assume an important role in computer animation systems.

*{jonas|lvelho|nando}@visgraf.impa.br

†siome@graphics.cis.upenn.edu

In fact, the development of robust motion processing techniques allows us to store captured motion, creating motion databases, which could be further modified to achieve some desired animation goals [13], [9].

These techniques should allow us to envision for animation the same paradigm of image based rendering and modeling. In fact, by capturing some specific motions from a family of characters, we should be able to create a "motion panorama" which would allow us to reconstruct any motion related with the family.

The image based rendering paradigm has produced fruitful results so far mainly because of the well established mathematical foundations in the area of image processing and computer vision. The existence of both functional and stochastic models for images, along with techniques for inverse problems allow us to develop a plethora of linear, projective and more complex non-linear filters. These filters allow us to analyze, process and sinthetize families of images, which entails the whole paradigm of image based rendering and modeling.

2 Previous Work

A *motion signal* is a vector function of one variable $f\colon \mathbb{R} \to \mathbb{R}^n$. This signal represents the positional and rotational values of the nodes of a character hierarchy (see Figure 1). We have $f(t) = (f_1(t), \dots, f_n(t))$, where each component $f_i(t)$ is called a *motion path*.

Motion processing consists in the use of operators on a space of motion signals. Motion processing techniques are used in many animation production environments and some processing techniques are found in commercial software currently available. Computer games such as FIFA Soccer [2] and virtual reality applications such as ActiveWorlds [1] employ motion databases with hun-

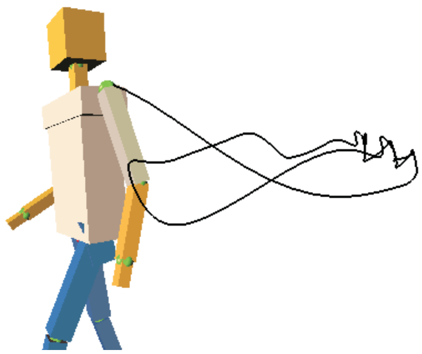


Figure 1: Motion signal.

dreds of small pieces of movements. According to user interaction, these motion paths are combined in real-time in order to achieve the desired goal.

There are two basic approaches to construct these operators:

- Using spacetime constraints;
- Constructing operators directly on motion signal spaces.

Spacetime constraint methods reduce the problem to the use of optimization techniques to minimize some energy functionals obtained by imposing different spatial constraints on the motion signal (i.e. the configuration space of the motion) (see [24]). Examples of this approach can be found in [25], [19] [10], [4], [14].

The direct approach, also called the *signal processing approach*, mimics on the idea of traditional image processing. A classical example occurs in keyframe animation systems, where sampling, reconstruction and resampling techniques are used to create motion paths from the keyframes. Also these techniques enable us to change some of the motion parameters (e.g. motion speed and acceleration) [6].

The signal processing approach to motion processing has been well addressed in [3]. In [22] a motion processing approach is used to create "emotional filters" for a character. In [4] operators are defined to change the time duration of a motion (cyclification).

In order to further develop the techniques of direct motion path operator construction, a point of fundamental importance consists in devising good mathematical models for motion path representation. Currently, there are essentially two basic approaches to representing a motion path:

- Spline-based representation;
- Periodic representation.

In the first model the motion signal is considered as a curve in the Euclidean n-dimensional space $\mathbb{R}^n$, and splines of different nature are used to represent and process the motion path. All of the "spline carpentry" is used to devise operators over motion signals. As an example, when a motion is captured spline-based interpolation techniques are used to reconstruct the motion signal from the samples [21].

In the periodic representation the motion path is assumed to be a periodic path. In this case the motion is completely characterized by its amplitude, frequency and phase. This representation allows us to use Fourier based techniques in order to construct motion filters. An interesting example of this approach appeared in [22], where "emotional filters" are developed representing the motion signal by periodic paths. These filters are computed on the Fourier domain by modifying the phase, amplitude and frequency of the motion signal.

The analysis and detection of motion cycles is an easy task for perfectly periodic motions, where the beginning and end of the curves match precisely. However, due to the nature of human locomotion, it is very unlikely that a perfectly periodic motion will occur. Small variations in phase and amplitude components of a "potentially periodic" human motion signal are caused by a series of factors, including oscillations of torque in muscles, uneven terrain and other external factors. These biomechanic and external factors introduce an important noise component in the signal, which is a fundamental aspect of natural-looking motion. We will call a motion with these properties as *near–periodic*.

We conclude that the use of periodic paths to represent a motion path is a crude approximation. Figure 2 shows the motion path of the elbow corresponding to two cycles.

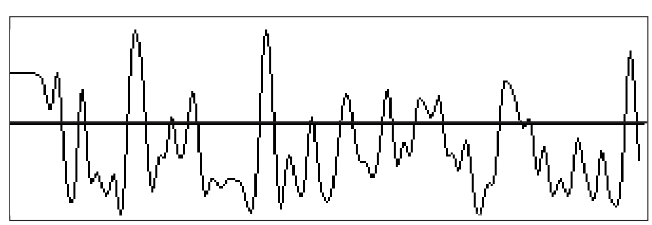


Figure 2: Motion path of the elbow motion.

Ken Perlin, [18], has pointed out that the presence of noise, (Perlin noise), is an essential component to obtain expressive motion of characters: The noise enhances the character gestures and allow us to obtain more pleasing gestures when animating the character. This established the fact that a motion signal is composed of periodic components plus noise, which is commonly referred as "motion texture".

3 Our contribution

The basic idea of the direct computation of motion path filters consists in extending the classical technique of image processing to motion. The success of image processing comes from the existence of very good mathematical tools to process images: good image models (functional and stochastic), and good mathematical tools to work with, arising from the theory of function spaces and stochastic processing.

Therefore, it is clear that the mathematical representation for a motion signal needs further investigations in order to design better motion filtering techniques. This is the main contribution of this work.

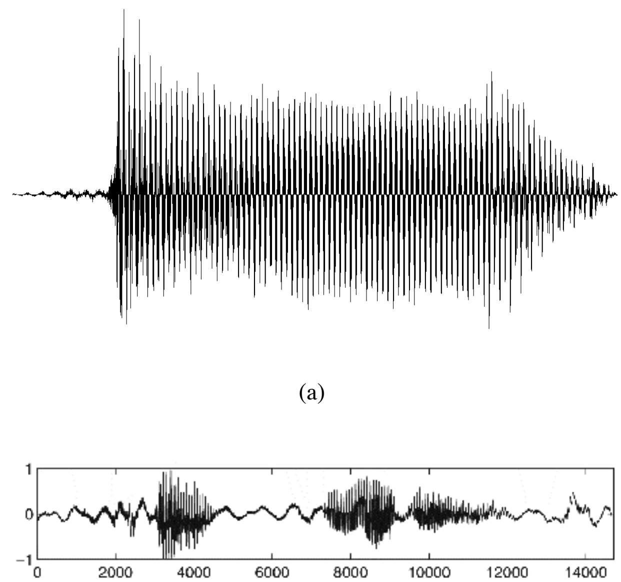
We propose the use of a new model for a motion signal based on the idea of using variable harmonic components. This idea is borrowed from the study of music and speech signals. We describe some mathematical tools to create motion filters using the model, and we describe an automatic algorithm to change the duration of a motion without modifying the "motion contents".

Also, we point to new mathematical tools, based on wavelets, to compute the harmonic components of a motion signal in order to create a whole new family of filters.

4 Variable harmonic contents

Periodic signals posses distinguished harmonic components characterized by three parameters: amplitude, frequency and phase. The study of these signals, in particular their description, representation and processing, is attained with the use of Fourier theory. Unfortunately, there are important signals that are not periodic, and the use of Fourier theory to study these signals has several limitations.

Figure 3 shows the plot of the amplitude variation along time of a music signal (a) and a speech signal (b). These two signals have some resemblance, although the speech signal has noticeable noise. This comes from the fact that both of them have harmonic components.



(a)

(b)

Figure 3: Musical signal (a), speech signal (b).

Since the signals are not periodic, these harmonic components have a variable nature, and an important issue consists in characterizing mathematically the nature of these components. This characterization should enable us to devise new tools to describe, represent and define operators on these signals.

4.1 The LCT decomposition

A well known approach to study harmonic contents of audio signals consists in subdividing the time domain of the signal into a number of "harmonic components packets". This is attained using a block basis transform. Essentially, we obtain a partition of the time interval, and for every interval of the partition we measure the harmonic components using an interval based cosine basis at different scales. An example of this approach is the classical Discrete Cosine Transform (DCT). Nevertheless, because of the periodicity of the cosine, the DCT basis has the disadvantage of having discontinuities at the boundaries.

A more recent approach was devised by H. Malvar, [16, 17], in the discrete case and, R.R. Coiffmann, Y. Meyer, [5], for continuous signals: Lapped block transforms. In this transform the cosine basis are windowed in such a way to avoid boundary discontinuities, and the intervals of the time partition overlap so as to maintain the orthogonality of the DCT.

The most used block basis transform is the Lapped Cosine Transform (LCT) which can be considered as a smoothed version of the DCT. Figure 4 illustrates the time-frequency decomposition of a signal using the LCT.

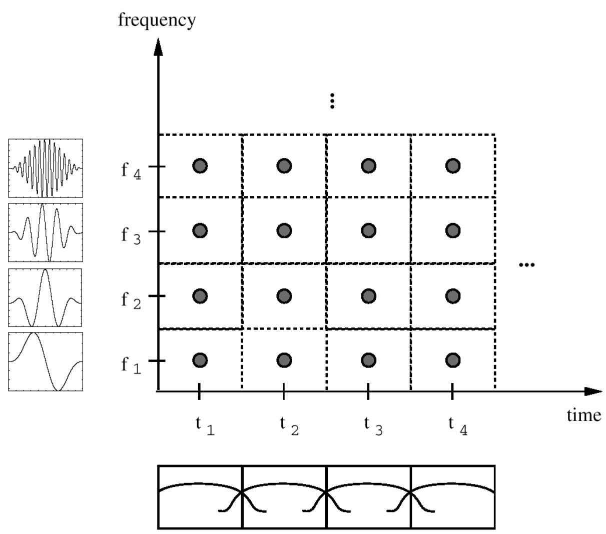


Figure 4: Time-frequency representation.

The LCT time×frequency representation computes, for each time interval, the harmonic contents of the signal at different scales. Therefore the transform is able to measure the variable nature of the signal's harmonic components.

4.2 The harmonic decomposition

Another approach to study the harmonic components of a signal consists in decomposing the signal using sums of sinusoidal partials ([20]):

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) \cos \phi_k(t). \qquad (1)$$

This is called the *harmonic decomposition* of the signal.

In [12] speech signals are studied by using a decomposition of the type

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} a_k(t) \cos \phi_k(t) + Noise. \qquad (2)$$

This is in accordance with our previous remark that speech signals have harmonic contents, as those of music sound, with some additional noise.

The decomposition in equation (1) resembles a Fourier series decomposition. Nevertheless both the amplitude a_k and the "phase function" ϕ_k change along the time. The derivative $\phi_k'(t)$ is called the *instantaneous frequency* of the signal. Note that in the case of periodic signals the amplitude is constant and $\phi_k(t) = bt+c$. Therefore, the instantaneous frequency coincides with the frequency of the signal.

The characterization of the harmonic components of music and speech signals using instantaneous frequency is not new. J. Ville, [23], used the so called Wigner-Ville transform

$$P_V f(u,\omega) = \int_{-\infty}^{+\infty} f\left(u + \frac{\tau}{2}\right) f^* \left(u - \frac{\tau}{2}\right) e^{-i\tau\omega} d\tau,$$

in order to compute the instantaneous frequency of a signal. In fact, theorem below was proved by him:

Theorem (Ville) *If* $f(t) = a(t)e^{i\phi(t)}$ *then*

$$\phi'(u) = \frac{\int_{-\infty}^{+\infty} \omega P_V f_a(u,\omega) d\omega}{\int_{-\infty}^{+\infty} P_V f_a(u,\omega) d\omega}$$

Nevertheless, it is well known that the Wigner-Ville transform has computational problems caused by the interference of crossed terms of the quadratic factor.

Recently, other time-frequency transforms have been used to compute the harmonic decomposition of a signal. In [8] a method is described to compute the instantaneous frequency and the variable amplitude using the Windowed Fourier Transform. On the other hand, S. Mallat, [15], has used Gabor wavelets, in order to obtain the harmonic sum decomposition of a music signal.

5 Harmonic content of motion signals

We have seen in Section 2 that motion signals have a near–periodic nature, and also posses noise. This fact lead us to study these signals by characterizing them by having variable harmonic components as in the case of music and speech signals. By comparing the elbow motion signal from Figure 2, with those of Figure 3, it is clear the resemblance between them. A major distinction is that the frequency of motion signals are very low compared with those of speech or musical sound. This is in fact a computational advantage.

The above remark leads us to propose the use of the mathematical models described in the previous section in order to study motion signals. Therefore, we could either use the harmonic decomposition in (1) to obtain the amplitude and instantaneous frequency of the signal, or we could use the Lapped Cosine Transform to obtain a decomposition of the harmonic contents.

5.1 Motion retiming

A common problem when reusing a motion path is that of changing the time duration. When a virtual player is pursuing the ball in a soccer game, a small piece of motion (e.g. a captured run motion) needs to be repeated several times to produce the effect of a running character. In another example, captured facial motions usually need some time-scaling processing in order to be synchronized with previously recorded audio.

Motion retiming is usually attained with two distinct techniques: motion reparametrization or motion cyclification. The first technique uses sampling and reconstruction methods techniques and it is quite suitable when a motion path is described using splines (see [6]). *Motion cyclification* performs a motion cycle analysis and detection in order to retime the motion (see [7]). Therefore, it is strongly based on the periodic nature of the motion signal.

5.1.1 Retiming using harmonic decomposition

Consider a motion path f decomposed into harmonic components as in equation (1). For a given constant $\alpha \in \mathbb{R}$, define the signal $g_\alpha(t)$ by

$$g_\alpha(t) = \sum_{k=1}^{n} a_k(\alpha t) \cos\left(\frac{\phi_k(\alpha t)}{\alpha}\right) \tag{3}$$

Note that the signal f has been rescaled in time in order to obtain g_α. Moreover, from the definition of g_α it follows that the amplitude of f at t_0 is equal to the amplitude of g at αt_0. Also, the instantaneous frequency of f at t_0 equals the instantaneous frequency of g at $t = \alpha t_0$. We conclude that the operator $f \mapsto g_\alpha$ changes the duration period of motion without changing its harmonic components (amplitude and instantaneous frequency).

5.1.2 Retiming using LCT decomposition

In this section we use the LCT decomposition of a motion signal in order to present an automatic method for changing the duration of a motion without altering the "motion content". A complete description of this technique has been published in [7].

The idea of the retiming algorithm is to use the LCT decomposition to obtain the "Harmonic content packets" of the signal. From this representation we scale the time content of the packets, replicating them, without changing the frequency content of each packet.

The retiming operation (W) in the frequency domain uses an affine dilation on the time axis of the time $\times$ frequency representation (which is equivalent to scaling the image on the time axis). This results in a replication of the atom elements of the representation, as shown in Figure 5.

The process of time dilation is as follows: we transform the 1D signal, $T(f)$; apply the time scaling (dilation or compression), $W(T(f))$; and reconstruct the curve in the time domain using the inverse transform, $T^{-1}(W(T(f)))$. Since regions of the image represent the presence of certain frequency components in the time segment limited by its boundaries, its stretching is responsible for a replication of the oscillations (prolonging the phenomenon in the expansion case). A very good illustration of the technique for speech signal is shown in the dilation Figure 6 (from [11]).

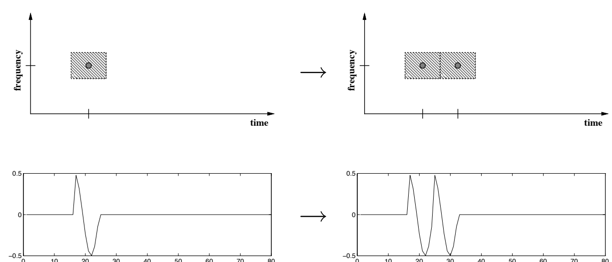


Figure 5: Time dilation in time $\times$ frequency domain.

The window computation

An essential step in the process of retiming of a motion path is the computation of the window size of the LCT. We compute this size based on the lowest frequency present in the motion signal. This frequency is called the *fundamental cycle* of the motion path.

Our method employs a circular autocorrelation function which measures the similarity between translated versions of a signal, as shown in Figure 7 (a). The fundamental cycle is given by the distance between consecutive maximum points of the correlated signal (Figure 7,

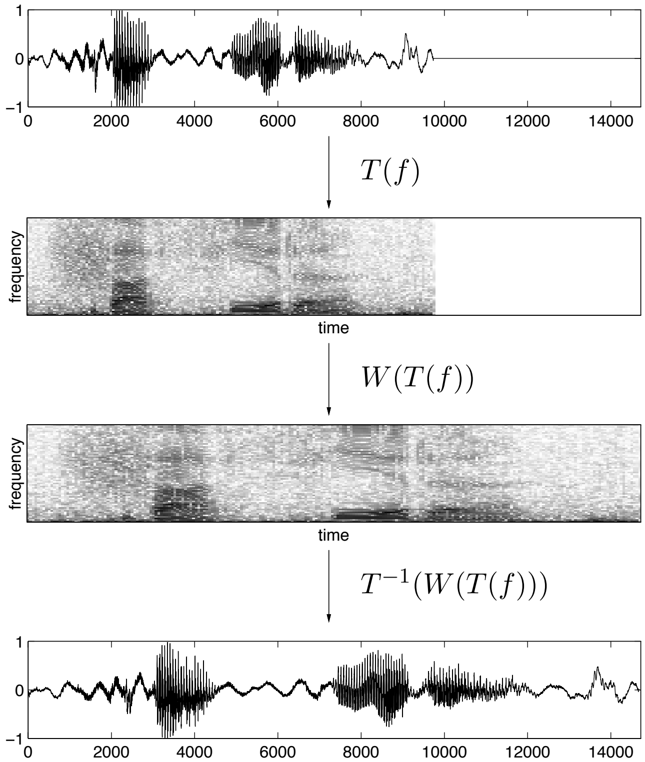
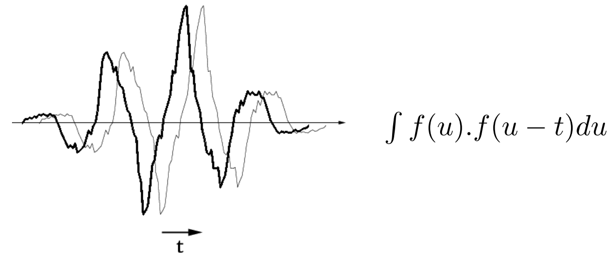


Figure 6: Time × frequency representation and time dilation of a signal f (top). (from [11])

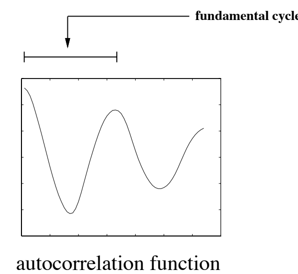
b). In our method the signal borders are smoothed using a windowing process, thus giving more weight to the central part of the data. This is extremely important because in periodic or near–periodic functions, there will exist at least one maximum at each multiple of the fundamental cycle. However, there is no guarantee that other (lower) local minima will also exist. Without the windowing, all maxima concerning the multiples of the fundamental cycle will have amplitude similar to the signal energy, which makes difficult the task of choosing the fundamental cycle.

5.1.3 Examples

In the following examples our method was applied to several individual motion curves. In all cases, the fundamental cycle detected by the algorithm is represented as a gray rectangle over the original signal, which is placed at the top. A dilation factor of two has been used in all examples.



(a)



(b)

Figure 7: Autocorrelation of a near–periodic signal (a) and its autocorrelation function (b).

Periodic motion

Figure 8 (top) shows a $sine$ function with fixed period. Note that although this function is periodic, in this example the beginning and end of the signal doesn't match, and therefore a simple concatenation would not generate good results. Figure 8 (bottom) shows the result of the cyclification applying our algorithm. Note that are no discontinuities in the boundaries of the cycle.

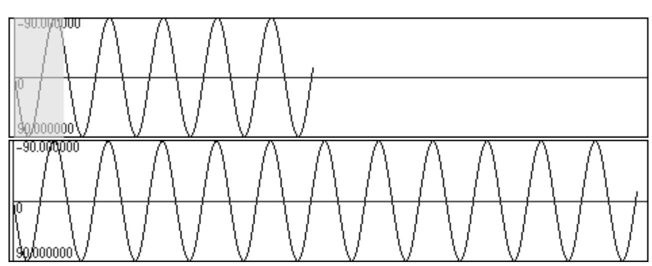


Figure 8: Time duplication of a periodic motion with a dilation factor of 2.

Pendulum with friction

Figure 10 (top) shows the motion curve resulting from a kinematic simulation of a pendulum presented in Figure 9.

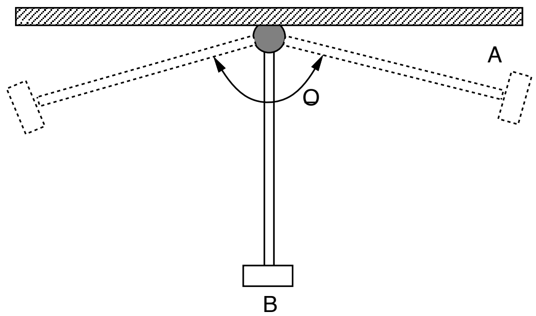


Figure 9: Kinematic simulation of a pendulum.

It is important to notice that in this case there is only a basic frequency which is repeated along the curve, but its amplitude decreases quadratically with time due to a simulated friction coefficient imposed to the system. The resulting signal (Figure 10, bottom) shows a replication of the frequency component, while preserving the quadratic decaying of its amplitude.

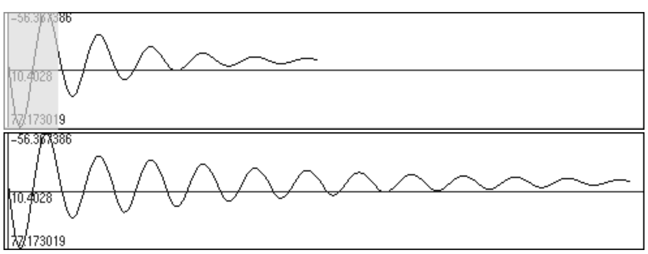


Figure 10: Time duplication of a pendulum motion curve using a dilation factor of 2.

Left uparm joint curve

In this example, a motion curve (left uparm joint) from a captured walking sequence was used as input to our algorithm. Note that the resulting signal is a perfect cyclification of the original one, with no discontinuities during the motion loops.

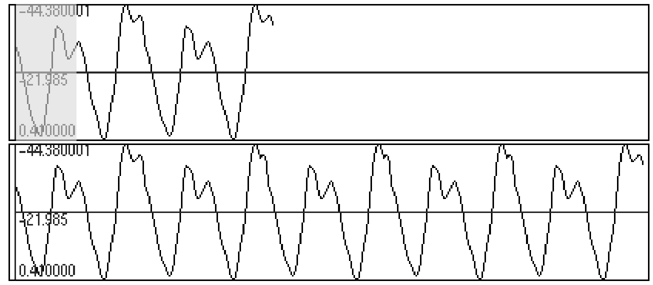


Figure 11: Time duplication of a left uparm motion curve. The dilation factor is 2.0.

6 Conclusions and current work

We have presented a novel mathematical framework to study motion paths, based on the characterization of the motion signal by the variable nature of its harmonic contents. We have exploited this motion signal representation using the LCT transform with very promising results, as is the case with the retiming algorithm here described.

Currently, we are implementing algorithms to compute the harmonic decomposition of a motion signal, using the Windowed Fourier Transform and Gabor wavelets. It is our purpose to use this decomposition in order to experiment with them in the definition of a new family of motion path operations. In particular, we are implementing an algorithm of motion retiming based on harmonic decompositions, using equation (3).

We are also investigating the application of our cyclification method to articulated figure motion by using captured human motion paths as input. In this case, an important aspect that must be considered in the cyclification process is that of synchronism between segments in periodic or near–periodic human motions.

We have detected coupling patterns during the movement of joints or groups of joints in specific sets of near–periodic human motions (e.g walking). These joints may have what we call strong and weak dependencies on their phases. A strong dependence within a group of joints means that their motion curves have a common periodic behavior, with phases that are multiples of a predominant fundamental cycle. In a weak dependence, the motion curves of joints are being influenced by the movement of other joints or groups of joints. As an example, in a walk movement (Figure 12), the motion of knees, feet, elbows and hands is strongly influenced

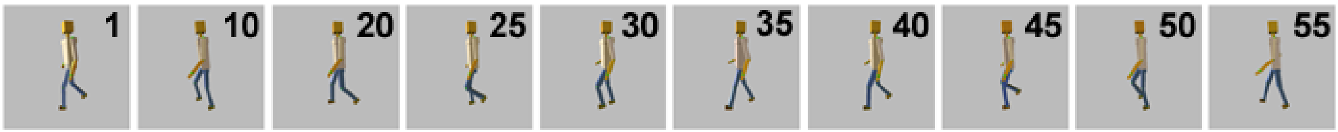


Figure 12: Selected frames from a walk sequence.

by the motion of upper arm and upper leg joints. This happens due to the structural relationship existing between these joints and also due to the nature of the walk motion. Events such as heel-strike and toe-touch are interpreted and processed by the human locomotor system, triggering actions that will control the basic aspects of a human gait. Moreover, there is a weak dependence between the joints of the arms and legs. This happens due to the necessity of a balance control that is achieved by a cross synchronization of arms and legs motions.

By detecting the predominant cycle associated to groups of joints and using it as starting point to the cyclification process, we believe that our method will preserve and correctly replicate all the fundamental cycles of the motion path.

Acknowledgments

The authors would like to thank Viewpoint Datalabs, Inc. and Biovision, Inc. for access to motion capture data, and to the Brazilian Council for Scientific and Technological development (CNPq) for the financial support. This research has been developed in the VISGRAF laboratory at IMPA, as part of the Ph.D. programme of the first author. VISGRAF laboratory is sponsored by CNPq, FAPERJ, FINEP and IBM Brasil.

References

[1] Activeworlds inc. http://www.activeworlds.com.

[2] Fifa99 website. http://www.fifa99.com/.

[3] Armin Bruderlin and Lance Williams. Motion signal processing. In *Proceedings of Computer Graphics (SIGGRAPH'95)*, pages 97–104, Los Angeles, USA, August 1995.

[4] Bobby Bodenheimer Charles Rose, Brian Guenter and Michael F. Cohen. Efficient generation of motion transitions usign spacetime constraints. In *Proceedings of Computer Graphics (SIGGRAPH'96)*, pages 147–154, New Orleans, USA, August 1996.

[5] R.R. Coifman and Y. Meyer. Remarques sur l'analyse de fourier a fenêtre. *C.R. Acad. Sci.*, pages 259–261, 1991.

[6] Fernando Wagner da Silva, Luiz Velho, and Jonas Gomes. Motion reparametrization. In *EUROGRAPHICS (short-papers issue)*, pages 1.5.1–1.5.4, Lisbon, Portugal, September 1998. Eurographics, Springer-Verlag.

[7] Fernando Wagner da Silva, Luiz Velho, Jonas Gomes, and Siome Goldenstein. Motion cyclification by time x frequency warping. In *SIBGRAPI'99 - International Symposium on Computer Graphics and Image Processing*, pages 49–58, Campinas, Sao Paulo, october 1999. IEEE Computer Society.

[8] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani. Asymptotic wavelet and gabor analysis: extraction of instantaneous frequency. *IEEE Trans. Info. Theory.*, 38(2):644–664, March 1992.

[9] Michael Gleicher. Motion editing with spacetime constraints. In *Proceedings of the 1997 Symposium of Interactive 3D Graphics*, 1997.

[10] Michael Gleicher and Peter Litwinowicz. Constraint-based motion adaptation. Report 96-153, Apple Computer, June 1996.

[11] Siome Goldenstein and Jonas Gomes. Time warping of audio signals. In *Proceedings of Computer Graphics International - CGI'99*, July 1999.

[12] Y. Stylianos J. Laroche and E. Moulines. Hns: speech modification based on a harmonic plus noise model. In *Proceedings of the IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, Minneapolis, Minnesota, 1993.

[13] Armin Bruderlin Kenji Amaya and Tom Calvert. Emotion from motion. In *Proceedings of Computer Animation (CA'96)*, 1996.

[14] J. Lee and S. Y. Shin. A hierarchical approach to interactive motion editing for human-like figures. In *SIGGRAPH 99 Conference Proceedings*, Annual Conference Series, pages 39–48. ACM SIGGRAPH, Addison Wesley, August 1999.

[15] S. Mallat. A wavelet tour of signal processing. Academic Press, New York, 1998.

[16] H. S. Malvar. Signal processing with laped transforms. Artech House, Norwood, MA.

[17] H. S. Malvar and D. H. Staelin. The lot: transform coding without blocking effects. *IEEE Trans. Acoust., Speech and Signal Proc.*, 37(4):553–559, June 1989.

[18] Ken Perlin. Real time responsive animation with personality. *IEEE Transactions on Visualization and Computer Graphics*, 1(1):5–15, March 1995.

[19] Z. Popoviĉ and A. Witkin. Phisically based motion transformation. In *SIGGRAPH 99 Conference Proceedings*, Annual Conference Series, pages 11–20. ACM SIGGRAPH, Addison Wesley, August 1999.

[20] X. Rodet and P. Depalle. A new additive synthesis method using inverse fourier transform and spectral envelopes. In *Proceedings of ICMC*, pages 265–272, May 1992.

[21] S. Sudarsky and D. House. Motion capture data manipulation and reuse via b-splines. In *Proceedings of CAPTECH'98*, 1998.

[22] Munetoshi Unuma, Ken Anjyo, and Ryozo Takaeuchi. Fourier principles for emotion-based human figure animation. In *Proceedings of Computer Graphics (SIGGRAPH'96)*, pages 75–84, 1996.

[23] J. Ville. Theorie et applications de la notion de signal analytique. *Cables et Transm.*, 2A(1):61–74, 1948.

[24] Andrew Witkin and Michael Kass. Spacetime constraints. In John Dill, editor, *Proceedings of Computer Graphics (SIGGRAPH'88)*, volume 22, pages 159–168, August 1988.

[25] Andrew Witkin and Zoran Popović. Motion warping. In *Proceedings of Computer Graphics (SIGGRAPH'95)*, pages 105–108, Los Angeles, USA, August 1995.