Timbre Morphing of Synthesised Transients Using the Wigner Time-Frequency Distribution

Thomas Lysaght and David Vernon

Department of Computer Science
National Univeristy of Ireland, Maynooth
Co. Kildare, Ireland

Abstract

In this paper we develop signal-processing techniques for Timbre Morphing consisting of linear interpolation and non-linear warping. We use the Wigner Time-Frequency Distribution as our representation of timbre. We use a simple mathematical formula for extrema to identify features in each Wigner Distrubution sufrace. Subgraph isomorphism is proposed as a method of identifying corresponding features in each sound. Signals are synthesised from Wigner Distributions using additive synthesis. We apply timbre morphing to short duration synthesised transients of equal pitch.

1. INTRODUCTION

The process of timbre morphing is a technique of music sound synthesis which combines existing sounds (timbres) to form a new sound with intermediate timbre and duration (Telleman, Lippold and Holloway 1991). We apply timbre morphing to short duration synthesised sounds of equal pitch. We use the Wigner time-frequency distribution as our representation of timbre to show the development of harmonics over time. The Wigner distribution gives good localisation in both time and frequency (Cohen 1995). We have implemented the discrete smoothed version, the Smoothed Pseudo Wigner Distribution (SPWD). We develop signal-processing techniques for timbre morphing consisting of linear interpolation and non-linear warping. Identifying a correspondence between features in both sounds is necessary for morphing. We propose using subgraph isomorphism as a method of finding correspondences. This represents an NP hard problem. We synthesise signals from Wigner Distributions using additive synthesis. The importance of the transient part of a musical sound in the identification of a particular instrument is well established (Woods 1975). We, therefore, choose only short duration transients for timbre morphing to reduce the cost of computing the SPWD.

2. The Wigner Distribution

We define timbre, formally, as the temporal development of harmonics within a sound i.e., the time-dependent variation in amplitude of each frequency comprising the sound. To represent this temporal variation we use the Wigner Distribution[1] (Wigner 1932). The Wigner Distribution as given by Cohen (Cohen 1995) is as

$$W(t,\omega) = \frac{1}{2\pi}\int s^*\left(t-\frac{1}{2}\tau\right)s\left(t+\frac{1}{2}\tau\right)e^{-j\tau\omega}d\tau$$
$$= \frac{1}{2\pi}\int S^*\left(\omega+\frac{1}{2}\theta\right)S\left(\omega-\frac{1}{2}\theta\right)e^{-jt\theta}d\theta$$

follows:

The Wigner distribution in terms of the signal, or its spectrum, are equivalent. The Wigner distribution, however, contains cross terms and negative values which are difficult to interpret (Cohen 1989). For computational purposes the Wigner distribution is windowed giving the Pseudo Wigner Distribution (PWD) (Janse and Kaizer 1983). Further smoothing of the PWD produces the Smoothed Pseudo Wigner Distribution (SPWD) which helps reduce the influence of cross terms. We use the SPWD defined as follows:

---

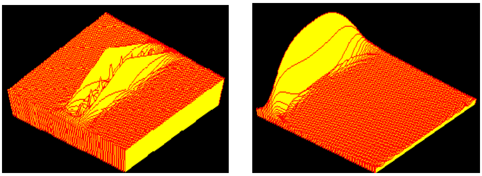
[1] Wigner developed his distribution in the field of quantum mechanics in order to calculate the quantum correction to the second virial coefficient of a gas, indicating how it deviates from the ideal gas law. This required a joint distribution of position and momentum.

$$SPWD(n,\theta) = 2\sum_{k=-L+1}^{L-1} e^{-jk2\theta}\, p(k) \sum_{l=-M+1}^{M-1} z(l)\, g(n,k)$$

where:

$$p(k) = \omega(k)\omega^*(-k) \text{ and } g(n,k) = s(n+k)s^*(n-k)$$

w(k) is the window function, *s(n)* is the signal and *z(l)* is the smoothing window.



(a) (b)

Figure 1 – (a) SPWD (Real) of sinusoid. The cross terms are visible between negative and positive frequencies. (b) SPWD (Analytic) of sinusoid.

Furthermore, the Wigner Distribution requires sampling of the signal at twice the Nyquist rate, i.e.,

$$f_c \le 4 f_s$$

where f_s is the sampling frequency and f_c is the highest frequency present in the signal (Janse and Kaizer 1983). We, therefore, need a sampling frequency which is twice as high as for the Fourier Transform to avoid aliasing. We compute the analytic signal (Cohen 1995, Janse and Kaizer 1983, Nuttall 1990) and then use the Nyquist sampling rate.

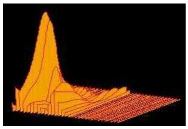
3. INTERPOLATION AND WARPING – MORPHING

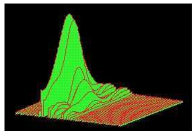
We develop a general formula for interpolation between *n* instrumental timbres (2-D Wigner Distribution surfaces). Interpolation is implemented as region-based linear interpolation, a region size of 1×1 giving point wise linear interpolation. We give a general formulation for interpolation in terms of the signature value at a point *s(f,t)*, the neighbourhood around the point $N_s(f,t)$ and its

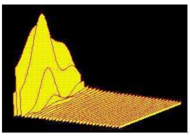
$$S(f,t) = I_s\{N_s(f,t), s(f,t), f, t\}$$

position in the time- frequency plane *f, t*:

Initially we use a simple linear mean value interpolation for two sounds. We also define a variable gradient- weighted interpolation where there is a relative sharing between the gradient at a point and the signature value at that point.







SPWD1 SPWD2 SPWD3

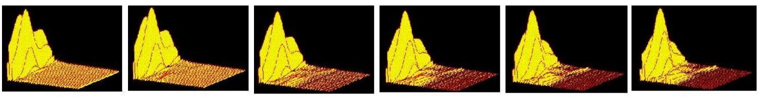


FIG0 FIG1 FIG2 FIG3 FIG4

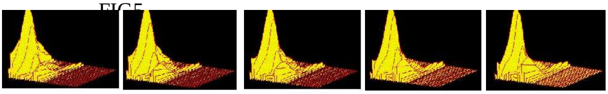


FIG6 FIG7 FIG8 FIG9

Figure 2 - SPWD 1, SPWD 2 and SPWD 3 are three Wigner Distributions. The peaks in SPWD 3 have been warped to align with peaks in SPWD 1. SPWD 2 is the warped surface. Figures FIG0 to FIG1 show a gradual morph between SPWD 1 and SPWD 2 using steps of 0.1 for each morph. At the halfway point, FIG0.5, SPWD 1 and SPWD 2 contribute equally to the morph.

Warping is a geometric operation, which distorts the original image by specifying control points in both the original and required image (Vernon 1991). Timbre morphing requires that corresponding features (e.g., peak of attack, loudest point, vibrato cycles, etc.,) in each sound are aligned so that one new feature results when the sounds are morphed. These will then constitute the control points for warping. We use a simple mathematical formula for extrema to identify features in each surface. Other features such as quietest point, start of decay, the position of the harmonic in the harmonic series, its magnitude and its maximum spread will also be taken as features for morphing.

In establishing a correspondence between features we note that for *n* control points there are *n!* possible configurations or correspondences. This represents an NP hard problem and a brute force approach requires effort *O(N!)* for a graph with *N* nodes (Depiero, Trivedi and Serbin 1996). We propose using subgraph isomorphism as a method of identifying corresponding features in each sound. We have implemented an algorithm by J. R. Ullmann (Ullmann 1976). Here subgraph isomorphism is determined by means of a simple enumeration procedure with backtracking, designed to find all isomorphisms between a given graph G_α and subgraps of a further graph G_β The algorithm operates on graphs based on a connectivity

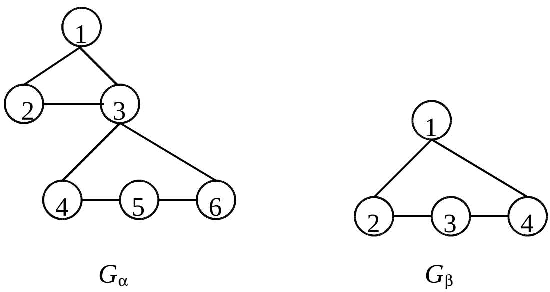


Figure 3 – Example graphs for subgraph isomorphism. Graph $G_\alpha$is isomorphic to a subgraph of graph G_β

analysis alone (Depiero, Trivedi and Serbin 1996). For this, a matrix M^0 is generated in accordance with:

$$m_{ij}^0 = \begin{cases} 1 & \text{if the degree of the } jth \text{ point of } G_\beta \geq \text{the degree} \\ & \text{of the } ith \text{ point of } G_\alpha, \\ 0 & \text{otherwise} \end{cases}$$

The enumeration algorithm works by generating all possible matrices M', each of which is used to permute the adjacency matrix B of G_β Each matrix M' is generated by systematically changing to 0 all but one of the 1's in each of the rows of M^0, such that no column of M', contains more than one 1. A further matrix, C, is defined as follows:

$$C = \left[c_{ij}\right] = M'\left(M'B\right)^T$$

where T denotes transposition. If it is true that:

$$\left(\forall i \forall j\right)\left(a_{ij} = 1\right) \Rightarrow \left(c_{ij} = 1\right) \begin{cases} 1 \leq i \leq p_\alpha \\ 1 \leq j \leq p_\beta \end{cases}$$

where p_α and p_β are the number of points in G_α and G_β respectively, then M' specifies and isomorphism between G_α and a subgraph of G_β. If $m_{ij} = 1$, then the *jth* point in G_β corresponds to the *ith* point in G_α in this isomorphism. Messmer and Bunke (Messmer and Bunke 1995) give a recursive procedure for Ullman's algorithm. We have also implemented this procedure. Figure 3 shows two example graphs G_α and G_β The input permutation matrix P for Messmer and Bunke's procedure is shown below and two distinct matrices M' for which subgraph isomorphism was found between G_α and a subgraph of G_β

$$P = \begin{pmatrix} 000001 \\ 100000 \\ 010000 \\ 001000 \\ 000100 \\ 000010 \end{pmatrix} \quad M' = \begin{pmatrix} 001000 \\ 000100 \\ 000010 \\ 000001 \end{pmatrix} \quad M' = \begin{pmatrix} 001000 \\ 000001 \\ 000010 \\ 000100 \end{pmatrix}$$

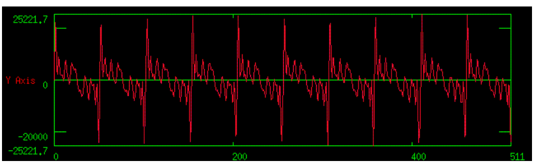
In applying subgraph isomorphism to features in sounds, we identify critical points in the Wigner Distribution of sounds. These points are the nodes in the input graphs to the isomorphism procedure.

4. SYNTHESIS OF NEW SIGNALS FROM THE WIGNER DISTRIBUTION

The Wigner Distribution gives us the decay envelope over time for each frequency and so we can reproduce the new sound by additive synthesis. The amplitude information at each discrete frequency component in the Wigner Distribution is the decay rate for a sinusoid of that frequency. We use a polynomial to model this synthesis. Sinusoids of each discrete frequency are weighted by their respective time decay envelopes (Wigner Distribution amplitudes) and then we sum over all frequencies. The Wigner Distribution will contain interference or cross terms and there is also a spreading in frequency of the SPWD. Therefore, our summation over all frequencies will contain unwanted components. In the case of music signals this could result in non- harmonic components or extra harmonic components in the synthesised signal with significant magnitudes which are not contained in the original signal. We choose the energy of frequencies, which are harmonics of a fundamental for synthesis. We define our polynomial to sum only the fundamental and its harmonics as follows:

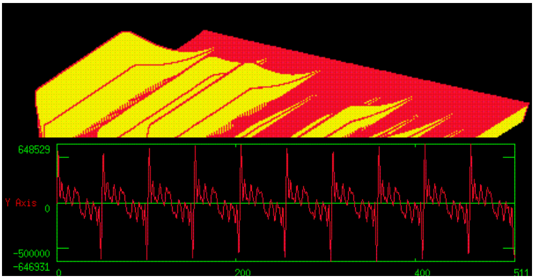
$$S'(i) = \sum_{i=0}^{N} \sum_{\substack{h=0 \\ h+=0}}^{M} \sin\left(\frac{h}{M} \times i \times \pi\right) \times W(h,i)$$

where $h\pi/M$ are the discrete frequency components for harmonics, $W(h,i)$ is the SPWD signature at frequency point h at time t and $s'(i)$ is the synthesised signal. (see Figure 4).



(a)

(b)



(c)

Figure 4 – (a) synthetic signal 10/512 Hz with 11 harmonics. (b) Wigner distribution of signal in (a). (c) signal in (a) synthesised from its Wigner distribution

We have applied timbre morphing to short duration synthesised sounds. Two sounds with different harmonic spectra were chosen, one with mostly even harmonics and another with mostly odd harmonics. The SPWD was run on the analytic version of the signals and a graded morph between the two sounds was implemented using steps of 0.1. A step of 0.1 involved a 0.1 contribution of first sound and 0.9 contribution of second sound to the morph. We have used simple mean value interpolation to create the new timbre. The new signal was then synthesised from its SPWD at each stage of the morph. This produced a good result on the synthetic sounds with a step of 0.5 producing the sound which was most easily distinguishable from both the original sounds.

Current work focuses on the application of timbre morphing to synthesised transients using subgraph isomorphism to identify correspondences between features. Future work will include the implementation of the Wigner Distirbution on a high performance computing facility for musical sounds sampled at high frequencies.

5. REFERENCES

Telleman, Edwin. Lippold. Holloway, Brian. (1991) *Timbre Morphing of Sounds with Unequal Numbers of Features*. J. Audio eng. Soc., Vol. 43, No. 9, Sept. 1991.

Cohen, Leon (1995). *Time-Frequency Analysis*. Prentice Hall PTR.

Cohen, Leon (1989). *Time-Frequency Distributions-A Review*. Proceedings of the

IEEE, Vol. 77, No.7, July 1989.
Woods, Alexander (1975). *The Physics of Music*. Chapman and Hall Ltd.
Vernon, David (1991) *Machine Vision, Automated Inspection and Robot Vision*. Prentice Hall International (UK) Ltd.
Wigner, E. P. (1932). *On the Quantum Correction for Thermo-dynamic Equilibrium*. Phys. Rev. 40, pp.749-75.
Janse, Cornelis, P., and Kaizer, Arie J. (1983). *Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution*. J. Audio Eng. Soc., Vol. 31, No. 4. 1983 April.
Nuttall, Albert, H. (1990). *Alias-free smoothed wigner distribution function for discrete-time samples*. Naval Underwater Systems Centre, Newport, Rhode Island, New London, Connicticut. Technical Report 8785, Oct. 1990.
Ullmann, J. R. (1976). *An Algorithm for Subgraph Isomorphism*. Journal of the Association for Computing Machinery, Vol. 23, No. 1, January 1976.
Depiero, Fred., Trivedi, Mohan., and Serbin, Steven. (1996). *Graph Matching Using a Direct Classification of Node Attendance*. Pattern Recognition, Vol. 29, No. 6, pp. 1031-1048.
Messmer, B. T. and Bunke, T. (1995*). Subgraph Isomorphism in Polynomial Time*. Technical Report IAM-95-003, University of Bern.