TIMBRE MORPHING USING THE WIGNER TIME-FREQUENCY DISTRIBUTION

Thomas Lysaght, Diarmuid O'Donoghue, David Vernon
Department of Computer Science
National University of Ireland
Maynooth
Ireland

June 30, 1997

Abstract

Timbre morphing is a technique of music sound synthesis which combines existing sounds (timbres) to form a new sound. Musical timbre is defined as the temporal development of harmonics within a sound . We implement the Wigner time-frequency distribution as our representation of the temporal variation of harmonics. Similarity between input signals provides a basis for generating the Wigner distribution of the new timbre. In this paper we develop signal processing techniques for multi-signal timbre morphing using linear region-based interpolation and geometric warping. A cost function is used for the correspondence problem encountered in matching timbral features (peak of attack etc.) in each sound. Recovery of the new timbre is achieved through additive synthesis. Due to interference terms, the Wigner distribution needs to be perturbed for the correct recovery of any signal. Results of synthesis are presented for synthetic signals. These provide a means to test and ensure a robust synthesis.

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 [1]. We apply timbre morphing to sampled instrumental sounds of equal pitch. We us the Wigner time-frequency distribution as our representation of timbre to show how harmonics develop over time within a sound. The principal method used for time frequency analysis has been the short-time Fourier transform (STFT), or spectrogram, developed in response to the analysis of human speech in the 1940's [2]. The spectrogram, however, contains an inherent tradeoff between time and frequency resolution [3], whereas, the Wigner distribution gives good localisation in both time and frequency. The synthesis of a new musical timbre is enabled by identifying an analogy [4,5,6] between the input signals and using their similarity as a basis for constructing the new "synthesised" timbre. To facilitate this process, we represent all input signals by their respective Wigner distributions (see Figure 2). The Wigner distributions represent the variations in signal strength over the frequency and time domains, with the vertical axis representing loudness. Features in the input spaces are identified and used as a basis for identifying the inter-signal similarity, with the most significant features including peak of attack, loudest point, quietest point, and start of decay. Each feature is identified in the input spaces, and then mapped to each corresponding feature in the other signals to ensure the generation of a suitable output signal. This mapping and transfer process in guided by a number of constraints which serve to ensure that the output signal is a Wigner distribution of the required timbre. Such constraints may, for example, influence the role played between the base frequency and its harmonics or the treatment of cross-terms generated during the mapping process. It is these constraints which ensure that the generated signal contains the required timbre features of the input signals.

Indeed, the structural diagram (Figure 1) detailing our approach to musical synthesis highlights the similarity between our approach and that of conceptual blends [7]. While some blend principles such as integration and good reason have parallels in our model, other principles find no place within our model. The unpacking and web principles for instance, do not apply to our model as there is no need to manipulate the generated conceptual space as an entity, while maintaining the association between the generated space and the input spaces.

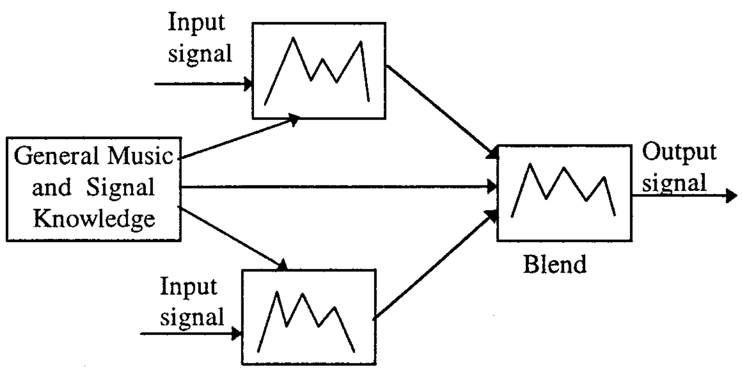


Figure 1. Timbre morphing as a blend.

We develop signal processing techniques for morphing consisting of linear interpolation and non-linear warping. The identification of a correspondence between features in both sounds is recognised as an NP-hard problem. We, therefore, limit the number of features used in morphing to reduce the complexity of correspondence matching.

1.1 Timbre

Musical timbre is an important ingredient in the composition, performance and appreciation of music both vocal and instrumental. The word timbre, also termed 'sound quality' and 'die klangfabre' is used to describe that characteristic quality which distinguishes sounds played on one instrument from those of another or to distinguish between different sounds played on the same instrument. We can immediately recognise the difference between a note played on, e.g., a trombone, from one played on a violin. Much work has been done through the physical [8,9,10] and the psychological [11,12] approaches in musical acoustics. Helmholtz [8] lists timbre as one of the three distinguishing attributes of musical tones, i.e. force (intensity or loudness), pitch (frequency) and quality (timbre). We often associate particular timbres with words such as 'smooth', 'sharp', 'hollow' etc., with colours and feelings as a means of comparing our perceptions of sounds. 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. The importance of timbre can be appreciated when we distinguish between 'good' or 'poor' quality of tone in an instrument. Various explanations have been proposed for the superior quality of tone of the Stradivarius string instruments and attempts have been made to reproduce this quality in modern instruments [13].

All musical tones are comprised of a series of pure (sinusoidal) tones. The lowest and generally the loudest of these 'partial' tones is termed the fundamental, or prime, tone. This, with the other partial tones, the upper partials or harmonics, forms an harmonic series as propounded by J.B. Fourier (1768-1830) and the series of upper partials is the same for all musical tones. The frequencies of the partials are in the ratio 1:2:3:4, etc. and are termed *fundamental*, *first upper partial* (second harmonic), *second upper partial* etc. The *fundamental* is generally the loudest harmonic and its frequency gives the frequency of the compound tone. We distinguish between musical sounds and noise (unwanted information in a signal [14]), and are interested, primarily, in sounds within the aural range (20Hz - 20kHz). Furthermore, the importance of the transient part of a musical sound in the identification of a particular instrument is well established. An attack transient is characteristic of all acoustic musical instruments and is the most important identifying element of the sound [9]. For timbre morphing, then, we choose sounds with short impulse responses(~0.4secs), where onset attack is followed almost immediately by the decay transient. We sample sounds at a chosen pitch (e.g. A = 440Hz, i.e., A above middle C) over the instruments' ranges and use these as the sounds to be morphed.

2 Representations of sound.

In order to analyse the timbre of a sound we need to know which frequencies are present in the signal and also when they are present. For the first requirement, namely, *which* frequencies are present, a suitable prediction can be made as all musical sounds contain the same relative makeup of harmonics and their relative strengths are known for different instruments. Fourier analysis provides the mathematical tool for the frequency domain representation of a signal. By expanding the signal as a set of functions it reflects the compound nature of the signal. The energy density spectrum (the absolute square of the Fourier transform) contains this information [2]. The second requirement, that is, *when* the frequencies are present, requires a frequency-time distribution to show how the spectral content is changing with time. For timbre morphing we use the Wigner Distribution[1] [15] which belongs to the Cohen General Class of time-frequency distributions. The General Class can be used to derive all time-frequency distributions by choosing a suitable, two dimensional, kernel function. The kernel function for the Wigner Distribution is unity. Other time-frequency distributions belonging to the General Class are due to Rihaczek, Page and Margenau-Hill. The Wigner Distribution as given by Cohen [2] is as follows:

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

The Wigner distribution in terms of the signal, $s(t)$ or its spectrum, $S(\omega)$ are equivalent. We have used the former of the two definitions. The Wigner distribution, however, contains cross terms and negative values which are difficult to interpret [2]. For computational purposes the Wigner distribution is windowed giving the Pseudo Wigner Distribution (PWD) [16]. 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:

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

where $p(k) = w(k)w^*(-k)$ 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.

The transition from the Wigner distribution $W(t,\omega)$ of a continuous time-signal to that of a discrete time-signal $SPWD(n,\theta)$ is non trivial requiring sampling of the signal at twice the Nyquist rate, i.e., $f_c \le 4f_s$ where f_s is the sampling frequency and f_c is the

[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.

highest frequency present in the signal [16]. Equation (2) may be viewed as a discrete Fourier Transform with respect to variable k of

$$\left[p(k) \sum_{l=-M+1}^{M-1} z(l) g(n,k) \right]$$

for window size $2L-1$. However, the Wigner Distribution is periodic in π compared with a period of 2π for the DFT [16]. The difference is caused by the factor 2 in the exponent of (2). We, therefore, need a sampling frequency which is twice as high as for the Fourier Transform to avoid aliasing in the Wigner Distribution. Alternatively we can compute the analytic signal [2,16] where the frequency spectrum vanishes for negative frequencies and then use the Nyquist sampling rate. Nuttall's Fourier domain implementation produces an alias-free representation using just the Nyquist sampling rate [17].

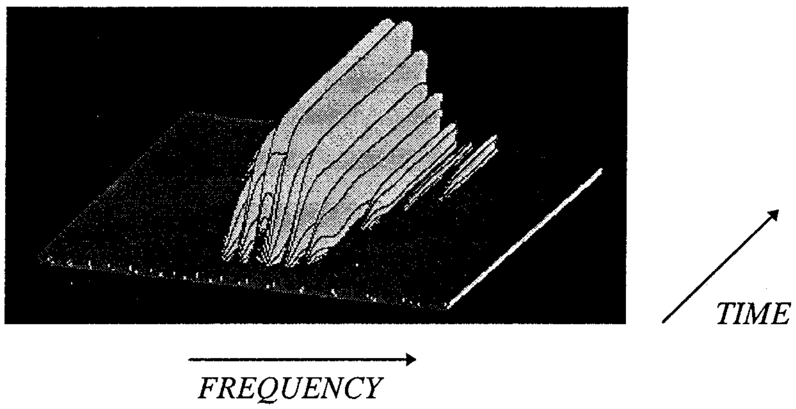


Figure 2. Wigner Distribution of synthetic signal for fundamental frequency and harmonics 3,5,7,9,12,15,and 17.

3 Interpolation and Warping - Morphing.

3.1 Interpolation

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. (See Figure 3). We give a general formulation for interpolation in terms of the signature value at a point, the neighbourhood around the point and its position in the time-frequency plane:

$$S(f,t) = I_s\{N_s(f,t), s(f,t), f, t\} \tag{3}$$

where:

$S(f,t)$ is the new interpolated time-frequency surface representing the new timbre.

I_s is the interpolation function where subscript s denotes the series of signals used in the interpolation. $s = s_1, s_2, \ldots s_n$, where n is the number of signals. e.g., $I_{c,v}$ could represent interpolation between a clarinet sound and a violin sound.

$N_s(f,t)$ is a series of neighbourhood functions for the series s defined above. I_s is therefore a function of n neighbourhood functions, e.g.

$$S(f,t) = I_{c,v}\{N_c(f,t), N_v(f,t), c(f,t), v(f,t), f, t\} \qquad (4)$$

could be the interpolation function for a clarinet (subscript c) and a violin (subscript v) sound.

$s(f,t)$ is the series of signal signature values at frequency f and time t.

f, t are the frequency and time co-ordinates of the time-frequency representation.

The interpolation function is also, therefore, a function of position.

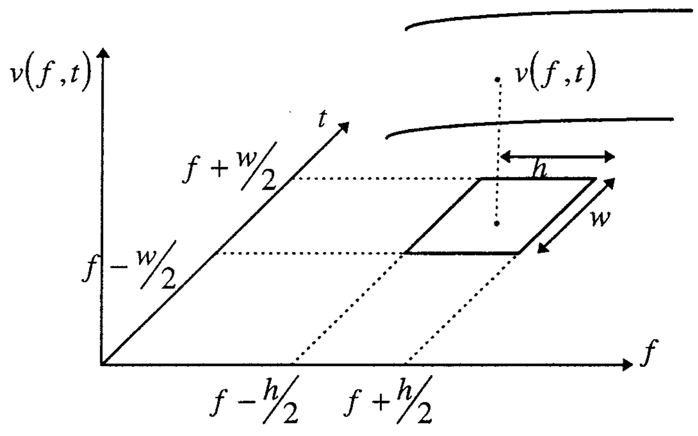


Figure 3. Rectangle (region) of interpolation $(w \times h)$ for violin surface.

Some linear interpolation methods can now be defined:

i. *linear mean value interpolation* for two sounds (e.g., clarinet and violin).

$$S(f,t) = \tfrac{1}{2}\{N_v(f,t) + N_c(f,t)\} \qquad (5)$$

where

$$N_c(f,t) = \frac{1}{w \times h} \sum_{i=f-h/2}^{f+h/2} \sum_{j=t+w/2}^{t+w/2} c(i,j)$$

i.e. the local average of the frequency-time signature values on, e.g., a clarinet surface, where $w \times h$ is the dimension of the morph window (region size). We require that the points $c(i,j)$ lie on harmonics and we can specify which harmonics are used in the interpolation. Similarly for $N_v(f,t)$.

ii. *Variable gradient-weighted interpolation.* The interpolation function is given by:

$$s(f,t) = \frac{(1-k+k*N_v(f,t))*v(f,t) + (1-k+k*N_c(f,t))*c(f,t)}{(1-k+k*N_v(f,t)) + (1-k+k*N_c(f,t))} \tag{6}$$

where the gradient(for the violin surface) is given as:

$$N_v(f,t) = \sqrt{(v(f+1,t)-v(f,t))^2 + (v(f,t+1)-v(f,t))^2} \tag{7}$$

Similarly for $N_c(f,t)$. The variable k determines the relative sharing between N_v and v or N_c and c. k determines how the gradient at a point weights the signature value at that point. The gradient here is the normalised gradient. Varying k in increments of *0.1*, $0 \le k \le 1$, will result in a progression from a simple mean of the surface signatures (*k=0*) to a maximum weighting by the gradient *(k=1)*.

Interpolation as a function of position (f,t) may occur when the points for interpolation are in the neighbourhood of a specified timbral feature (e.g. peak of attack).

3.2 Warping

Warping is a geometric operation which distorts the original image by specifying control points in both the original and required image [14]. While the interpolation function is defined as non-geometric (i.e. time-frequency or position invariant), the warping function is geometric (Figure 5.). Timbre morphing requires that certain corresponding features in each sound are aligned so that one new feature results when the sounds are morphed. The peak of attack, for example, in each sound will need to be aligned giving one attack peak in the new sound. Other features such as loudest point, quietest point, start of decay, or repeatable features such as vibrato peaks will also be aligned. The position of the harmonic in the harmonic series (low or high), its magnitude and its maximum spread will be taken as features for morphing. These will then constitute the control points for warping. We identify features initially by finding extrema in each surface. For this we use the simple mathematical formula:

$$\Delta = AC - B^2 \tag{8}$$

where

$$AC - B^2 = f_{xx} f_{yy} - f_{xy}^2 = \begin{vmatrix} f_{xx} f_{xy} \\ f_{yx} f_{yy} \end{vmatrix} \text{ is the discriminant of } f. \tag{9}$$

$A = f_{xx}(a,b)$, $B = f_{xy}(a,b)$ and $C = f_{yy}(a,b)$ and f_{xx} and f_{yy} are the second partial derivatives of f in the x and y directions respectively and f_{xy} is the mixed second partial derivative of f. For a local minimum at (a,b) we have $\Delta > 0$ and $A > 0$. For a local maximum at $(a,b), \Delta > 0$ and $A < 0$. For a saddle point at $(a,b), \Delta < 0$ and $\Delta = 0$ implies an inconclusive test.

Correspondence between features (see Figure 4), then, needs to be established before morphing. The number and choice of features will be kept to a minimum due to the complexity involved in correspondence matching. For n control points there are $n!$ possible configurations or correspondences. This represents an NP hard problem. However, for small n, the problem is manageable. To find which correspondence is best we implement a cost function which uses all possible correspondences and we then minimise this cost function. An example is the sum of distances between control points (Figure 4). In the case of two sounds this cost function is given by:

$$C = \sum \left|\left(p_i p_j\right)\right| \tag{10}$$

where:

C is the cost function.
P_i is a control point in the first image (timbre).
P_j is a control point in the second image (timbre).
$\left.\begin{matrix} i = 1..n. \\ j = 1..n \end{matrix}\right\}$ where n is the number of extrema in each sound.

and

$$\left|\left(p_i p_j\right)\right| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{11}$$

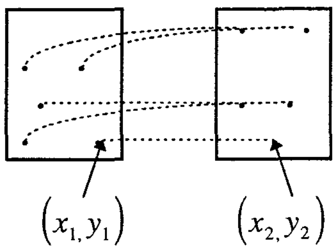
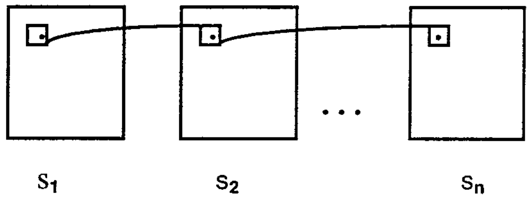
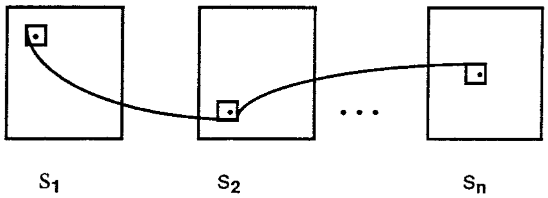


Figure 4. Correspondence between features for two sounds.

We can have a number, n say, of cost functions C_i, $i=1,2,3,..n$, each as a measure of the correspondence between control points. Finally, we specify the order or the warping function (e.g., third order). Interpolation and warping are summarised in Table1 and Table2.



Interpolation (non-geometric)



Warping (geometric)

Figure 5. Interpolation and Warping

TABLE 1.

Interpolation		
Window (region) size	Linear	Gradient Weighted
1×1	Point wise linear	Variable Gradient weighted
3×3, 5×5, 7×7, …	Region based linear	

TABLE 2.

Warping (order specified, e.g. third)		
Choice of Extrema	Features (control points) Extrema	Optimality Criteria For Correspondence Match (Cost Function)
Start of attack, Peak of attack Loudest point, Quietest point Vibrato peaks, Start of Decay Harmonics (Low, High, Random, Magnitude Maximum Spread)	1,2,3,…n	C_1 C_2 C_3 … C_n

4 Synthesis of new signals from the Wigner Distribution.

New signals will by synthesised from the Wigner Distribution of the new timbre by 'additive' synthesis. The transformation is from the 2-D time-frequency representation to the 1-D signal (Eq. 12).

$$S(f,t) \Rightarrow s(t) \tag{12}$$

Since the Wigner Distribution gives us the decay envelope over time for each frequency component 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 (the Wigner Distribution amplitudes) and then we sum over all frequencies (Eq. 13).

$$\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} A_{i,j} f_{i,j} \tag{13}$$

where:

M is the number of frequency sampled points.
N is the number of time samples in the sound.
$A_{i,j}$ is the amplitude in the Wigner Distribution at point (i,j) .
$f_{i,j}$ is the amplitude of a sinusoid with frequency $\frac{i\pi}{M}$ at time j.

The discrete frequency components are given by $\frac{m\pi}{M}$, where $m = 0 \ldots M-1$ is the length of the frequency window chosen.

As already stated, the Wigner Distribution will contain interference or cross terms. These provide frequency contributions which are not present in the original signal. Therefore, the summation over all frequencies in Eq. 13 will contain unwanted components. The amplitude of these cross terms will determine their contributory effect on the resultant synthesised signal. 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. Smoothing in the SPWD has the effect of reducing the effect of the cross terms. Another factor is the effect of a spreading in frequency of the SPWD which gives rise to extra frequency contributions for each discrete frequency component of a signal. The synthesised signal will, therefore, deviate from a 'correct' synthesis(i.e. a synthesis which returns the original signal from its Wigner Distribution) by the contribution of these interference terms. We choose to ignore the energy contributions due to the spread in frequency and use only the energy of frequencies which are harmonics of a fundamental for synthesis. We can now redefine our polynomial to sum only the fundamental and its harmonics as follows:

$$s'(i) = \sum_{i=0}^{N} \sum_{\substack{h=0 \\ h+=f_0}}^{M} \sin\left(\frac{h}{M} \times i \times \pi\right) \times W(h,i) \tag{14}$$

where:

$\frac{h\pi}{M}$ are the discrete frequency components for harmonics in the Wigner Distribution.

h denotes the discrete frequency of each harmonic in the Wigner Distribution. The value of h is incremented at each iteration by the offset of the fundamental f_0 from the origin on the frequency axis of the Wigner Distribution.

$W(h,i)$ is the Wigner Distribution signature at frequency point h at time t.

$s'(i)$ is the synthesised signal.

N is the number of time samples on the Wigner Distribution.

M is the number of frequency samples on the Wigner Distribution.

There still may be some deviation due to cross terms and we therefore need a method which retrieves the correct signal from the Wigner Distribution. We can measure the size of the deviation by taking the RMS error of the original and synthesised signals. The RMS is given by:

$$rms = \left| \frac{1}{N} \sum_{k=0}^{N-1} \left| s(k) - s'(k) \right|^2 \right|^{1/2} \tag{15}$$

where:

$s(k)$ is the original signal.

$s'(k)$ is the synthesised signal.

Since the Wigner distribution may need to be perturbed in order to recover a correct signal, we can modify Eq.(14) by including an error $e(h,i)$:

$$s'(i) = \sum_{i=0}^{N} \sum_{h=0}^{M} \sin\left(\frac{h}{M} \times i \times \pi\right) \times \left(W(h,i) + e(h,i)\right) \tag{16}$$

A test suite of synthesised 'music' signals will be used initially to test the synthesis routine (see Figures 7,8,9,10). A simple sinusoid (pure tone) will be processed and subsequently more complex (multi-component) signals (compound tones). A fundamental with just a few harmonics (e.g. second and third) will be taken as a compound tone. The amplitudes of the harmonics relative to the fundamental will be chosen so as to simulate different instrumental timbres.

4.1 Multi-Note Synthesis

A series of notes (tones) can be synthesised from the Wigner distribution of a single tone by specifying the new fundamental frequency. For example, the note middle C = 264Hz has the following harmonic structure:

$$C(t) = f_{264}(t) + f_{528}(t) + f_{792}(t) + \ldots \ = \sum_{n=1}^{N} n \times f_{264}(t) \quad , t = 0 \ldots T-1 \qquad (17)$$

where N is the number of harmonics specified and T is the time extent of the signal and $1 \times f_{264} = f_{264}, 2 \times f_{264} = f_{528}$, etc.

A complete scale (e.g. middle C,C#,D,D#,E,F,F#,G,G#,A,A#,B,C' [2]) of notes (tones) could be synthesised from one Wigner Distribution. Other scales could likewise be synthesised from notes sampled from lower or higher registers in each instrument. The timbre of an instrument varies from low to high registers and therefore sampling of new tones from one tone (we will call it the source tone) should be restricted to within an octave or two of the source tone. A whole keyboard of synthesised tones of new timbres is thus possible. Also, by simply adding combinations of different notes (e.g. middle C + E(330Hz) + G(396(Hz))[3], chords can be produced, making possible the production of harmonised melodies played using multiple new timbres.

5 Conclusion

We have presented signal processing techniques for timbre morphing and have identified an analogy between the two inputs signals using their similarity as a basis for constructing the new "synthesised" timbre. We have defined the varying frequency-temporal nature of timbre and have implemented the Wigner Distribution as its representation. The interpretative difficulties of the Wigner distribution due to cross terms have been highlighted. We have given the theoretical foundations for non-linear interpolation and have implemented geometric warping as a means of aligning corresponding features in each sound. This constrains the mapping and transfer process of timbre morphing which serves to ensure that the output signal is a Wigner distribution of the required timbre. We have highlighted the correspondence problem in matching features and suggested a cost function to compute the mapping. We use the RMS as a correcting measure in perturbing the Wigner distribution to produce a 'correct' synthesis of the original timbre. This has been verified by using synthetic signal data. This still requires some refinement but present results are quite accurate and suggest only slight alterations. Finally, we have suggested creative methods of synthesising a whole keyboard of new timbres from the Wigner distribution of morphed sounds using our synthesis technique.

[2] This is known as a Chromatic Scale as it contains all the semitones (half-tones) between the first note (middle C) and its octave (C'). The # sign means a sharp in musical notation and signifies the semitone directly above the note.

[3] This combination CEG is known as the root chord of the key of C and is denoted $\binom{5}{3}$,E being a third above middle C and G a fifth above middle C.

References

[1]. Telleman, Edwin. Haken, Lippold. Holloway, Brian. "Timbre Morphing of Sounds with Unequal Numbers of Features." J. Audio Eng. Soc., Vol. 43, No. 9, 1995 September.

[2]. Cohen, Leon. "Time-Frequency Analysis". Prentice Hall PTR 1995.

[3]. Cohen, Leon. "Time-Frequency Distributions-A Review". Proceedings of the IEEE, Vol 77, No.7, July 1989.

[4]. Gentner, D. *"Structure-Mapping: A Theoretical Framework for Analogy"*, Cognitive Science, 7, 155-170, 1983.

[5]. Keane, M. T. Ledgeway, T., Duff, S. *"Constraints on Analogical Mapping: A comparison of Three Models"*, Cognitive Science, 18, 387-438, 1994.

[6]. Veale, R. A. O'Donoghue, D. Keane, M. T. *"Computability as the Ultimate Cognitive Constraint"*, Annual Conference of the International Cognitive Linguistics Association, Albuquerque: NM, 1995.

[7]. Turner, M. Fauconnier, G. *"Conceptual Integration and Formal Expression."*, Metaphor and Symbolic Activity, 10:3, 183-203. 1995

[8]. Helmholtz, Hermann. "On the Sensations of Tone". Dover Publications 1954.

[9]. Woods, Alexander. "The Physics of Music". Chapman and Hall Ltd., 1975.

[10]. Winckwl, Fritz. "Music, Sound and Sensation. A Modern Exposition". 1967 Dover Publications, Inc.

[11]. Melka, A., "Methodological approaches to the investigation of musical timbre". Journal de Physique IV, may 1994.

[12]. Seashore, Carl. E. "Psychology of Music". Dover Publications, Inc., 1967, New York.

[13]. Krupowicz, Stanislaw and Skalmierski, Bodgan. "Towards an explanation of the shape of the violin".

[14]. Vernon, David. "Machine Vision, Automated Inspection and Robot Vision", Prentice Hall International (UK) Ltd., 1991.

[15]. Wigner, e. P., "On the Quantum Correction for Thermo-dynamic Equilibrium". Physic Review 40, pp.749-759 (1932).

[16]. Janse, Cornelis, P., and Kaizer, Arie J., "Time-Frequency Distributions of Loudspeakers: The Application of the Wigner Distribution". J. Audio Eng. Soc., Vol 31, No. 4. 1983 April.

[17]. Nuttall, Albert,H. "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.

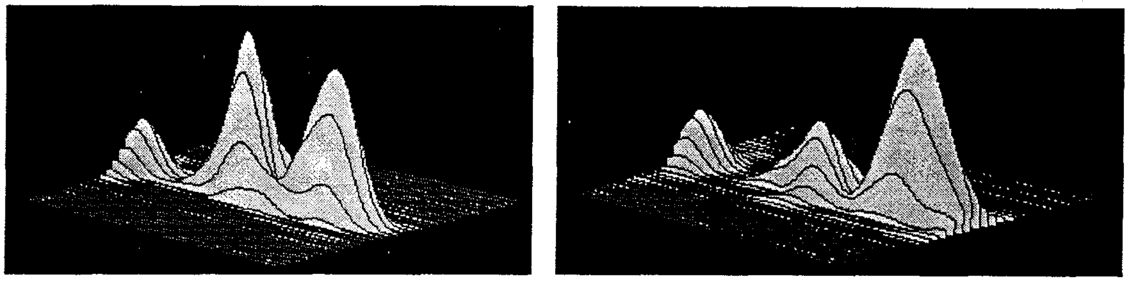


Figure 6. Peaks in the Wigner distribution as features for morphing

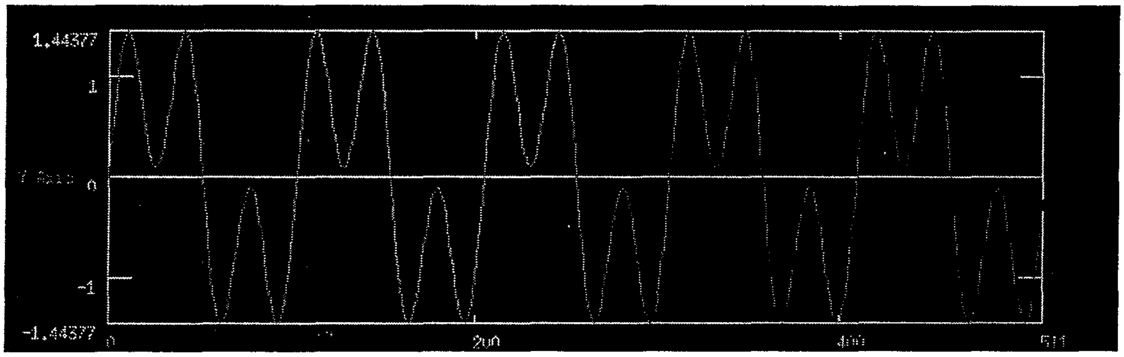


Figure 7. Synthetic signal. Fundamental and third harmonic.

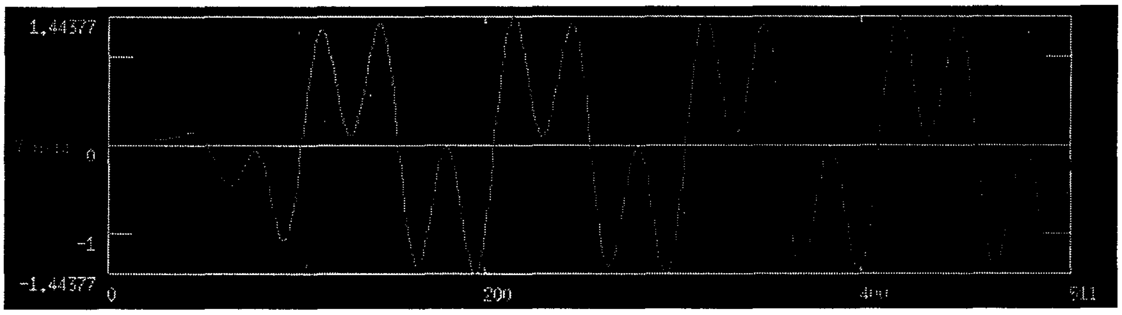


Figure 8. Signal in Figure 7 synthesised from its Wigner distribution.

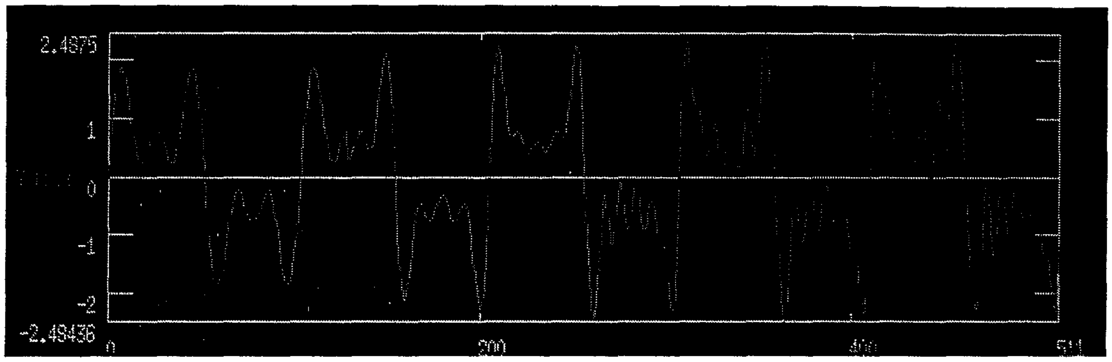


Figure 9. Synthetic timbre (fundamental and 7 harmonics)

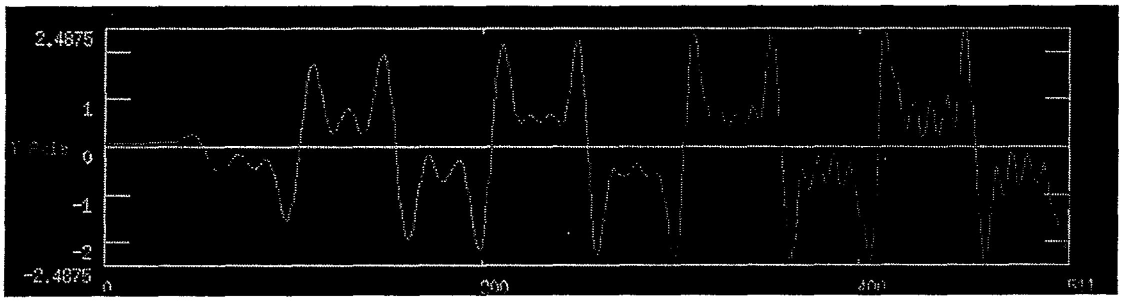


Figure 10. Signal 9 synthesised from its Wigner distribution.