Subgraph Iosmorphism Applied to Feature Correspondence in Timbre Morphing

Thomas Lysaght, David Vernon and Joseph Timoney
Department of Computer Science
National University of Ireland, Maynooth
Co. Kildare
Ireland

e-mail: Tom.Lysaght@may.ie, David.Vernon@may.ie, jtimoney@cs.may.ie
WWW: http://www.cs.may.ie

May 8, 2000

Abstract

This paper presents a subgraph isomorphism technique for correspondence matching in a Wigner Distribution timbre morphing procedure. The aim is to overcome the shortcomings found with previous STFT based methods. The results presented demonstrate that this technique produces a perceptually convincing sound.

keywords: Timbre morphing, wigner distribution, subgraph isomorphism.

1 Introduction

Timbre morphing is a technique for music sound synthesis that combines existing sounds (timbres) to form a new sound with intermediate timbre and duration [1]. Manipulation of the audio signals is normally done in the time-frequency plane using a suitable representation. Early work in this area was based on using the well-known Short-Time Fourier transform (STFT) or

Spectrogram [2] but this approach presented problems of both temporal and spectral smearing. More recently, current research has shifted towards using more optimal time-frequency representations such as the Wigner Distribution [3]. In particular, the Wigner distribution offers benefits of less spectral smearing and temporal smearing [4]. To blend the sounds, the Wigner-distribution based timbre morphing process requires techniques of linear interpolation and non-linear warping. However, the most significant problem is the identification of correspondences between features, such as peak of attack and vibrato cycles, in both sounds to ensure a perceptually smooth transition between them once warping is applied. Here, it is proposed to use subgraph isomorphism [5] [6] as a method to find these correspondences rather than a direct brute force matching process. Previous work contained an outline of the relevance of this technique [7] while this paper provides a detailed exposition of the implementation combined with results to demonstrate the accuracy and efficiency of the technique. Results are also presented of signals generated from Wigner Distributions using additive synthesis to demonstrate the overall effectiveness of the method.

2 Subgraph Isomorphism

The process of morphing requires the identification of certain important features in the time-frequency representations of the audio signals, establishing a correspondence between these features and then blending them together using interpolation and warping techniques. For both techniques, it is required that corresponding features (e.g., peak of attack, loudest point, vibrato cycles, etc.,) in each sound are aligned so that only one new feature results when the sounds are morphed. This is the correspondence problem and it is very difficult in that for n control points there are $n!$ possible configurations or correspondences. This represents an NP hard problem and a brute force approach requires significant computational effort $O(N!)$ for a graph with N nodes [8]. To overcome this, subgraph isomorphism [5] can be employed to identify the corresponding features in each sound. 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 subgraphs of a further graph G_β.

The enumeration algorithm works by generating all possible permutation matrices, each of which is used to permute the adjacency matrix of G_β. The alogorithm operates on graphs based on a connectivity analysis alone. For this a matrix M_0 is generated in accordance with:

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

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:

$$C = [c_{ij}] = M'(M'B)^T \tag{2}$$

where T denotes transposition. If it is true that:

$$(\forall i \forall j)(a_{ij} = 1) \Longrightarrow (c_{ij} = 1) \begin{cases} 1 \leq i < p_\alpha \\ 1 \leq j < p_\beta \end{cases} \tag{3}$$

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.

In applying subgraph isomorphism to features in sounds, critical points are first identified in their Wigner Distribution which then become nodes in the input graphs to the isomorphism procedure. For example, Figure 1 shows graphs for the 'Peaks of attack' feature, where each node represents a peak of attack identified in the sound and the arrows determine the ordering of nodes for finding the correspondences. A total of 15 permutation matrices were found for these directed graphs. Some of these are given below. The isomorphisms give by the first two matrices are given by $P_1 = \{m_{11}, m_{22}, m_{33}, m_{44}\}$ and $P_2 = \{m_{11}, m_{22}, m_{33}, m_{45}\}$.

$$P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

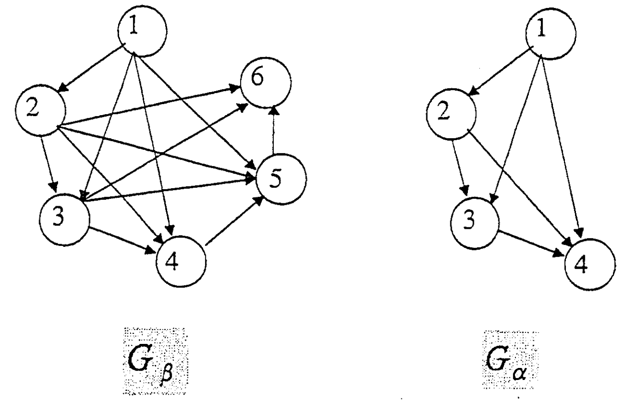


Figure 1: Example graphs for subgraph isomorphism. Graphs G_α and G_β representing different combinations of harmonic features in ascending order of harmonic number.

$$P_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \quad P_4 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$P_5 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

Following this, interpolation and warping are then performed. The corresponding peaks of attack identified through subgraph isomorphism in each sound form the control points for morphing. The control points for the new morphed sound (output image) are calculated as the midpoints between those in both input sounds so that both sets of fiducial points warp to the same location. Each input SPWD surface is warped separately using the same control points file for the resultant output image. The order of warping is second order which requires more than 9 control points for the overdetermined case. To warp each SPWD correctly, however, it was necessary to pin each surface with fiducial points along the edges and also along local areas on each harmonic.

Without this pinning, the points other than control points can be warped unpredictably to skewed or wildly varying positions in the SPWD surface. For example, in order to obtain stable results, extra pinning points must be included applying the procedure to a total of 74 control points in all by pinning each harmonic at the edges and close to the first feature point on each harmonic to be warped.

Trial and error is used to find the best approximation of warping to the specified control points.

Difficulties were also encountered when attempting to warp two points on neighbouring harmonics in opposite directions. This attempt to skew the surface gave very unsatisfactory and erronous results. A decision to leave one point unaltered and to map the other point to its position is necessary in these cases. For testing, initially, all harmonics from the reed-like sound are mapped onto the harmonics of the horn-like sound. The horn-like harmonic has one more harmonic present.

It may be added at this point that mappings are restricted to sequences of harmonics that have ascending harmonic number. This choice is arbitrary. Mapping ascending sequences to descending sequences may indeed give interesting and useful results when morphing.

3 Generation of synthesised signals

As the Wigner Distribution gives the decay envelope over time for each frequency, it is possible to reproduce the new sound by additive synthesis. This is achieved by generating a sinusoid for each discrete frequency, weighted it by the respective time decay envelopes (Wigner Distribution amplitudes) and then summing over all frequencies. To prevent the integrity of the signal being compromised by interference or cross terms and the frequency spreading property of the distribution, it is necessary to use a decision-based signal processing procedure to ensure that the frequencies of the synthesised sinusoids exhibit a harmonic relationship [7]. In applying 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. For computational efficiency and improved accuracy, the signals were Hilbert-transformed and then represented in the discrete time-frequency plane using the Smoothed Pseudo Wigner Distribution (SPWD). A graded morph between the two sounds was implemented using steps of 0.1.

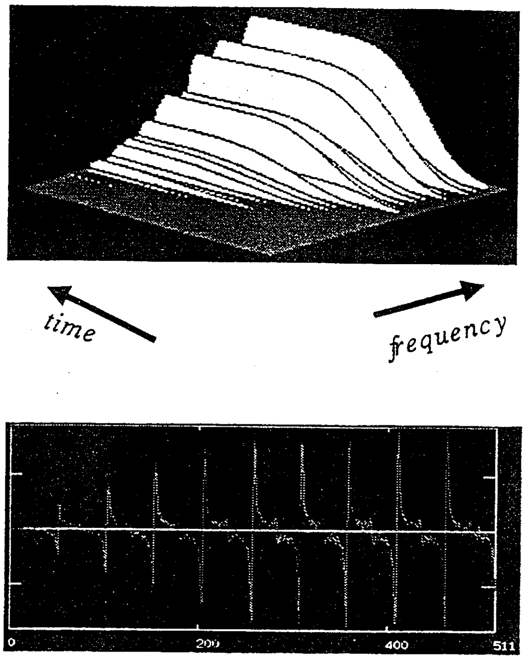
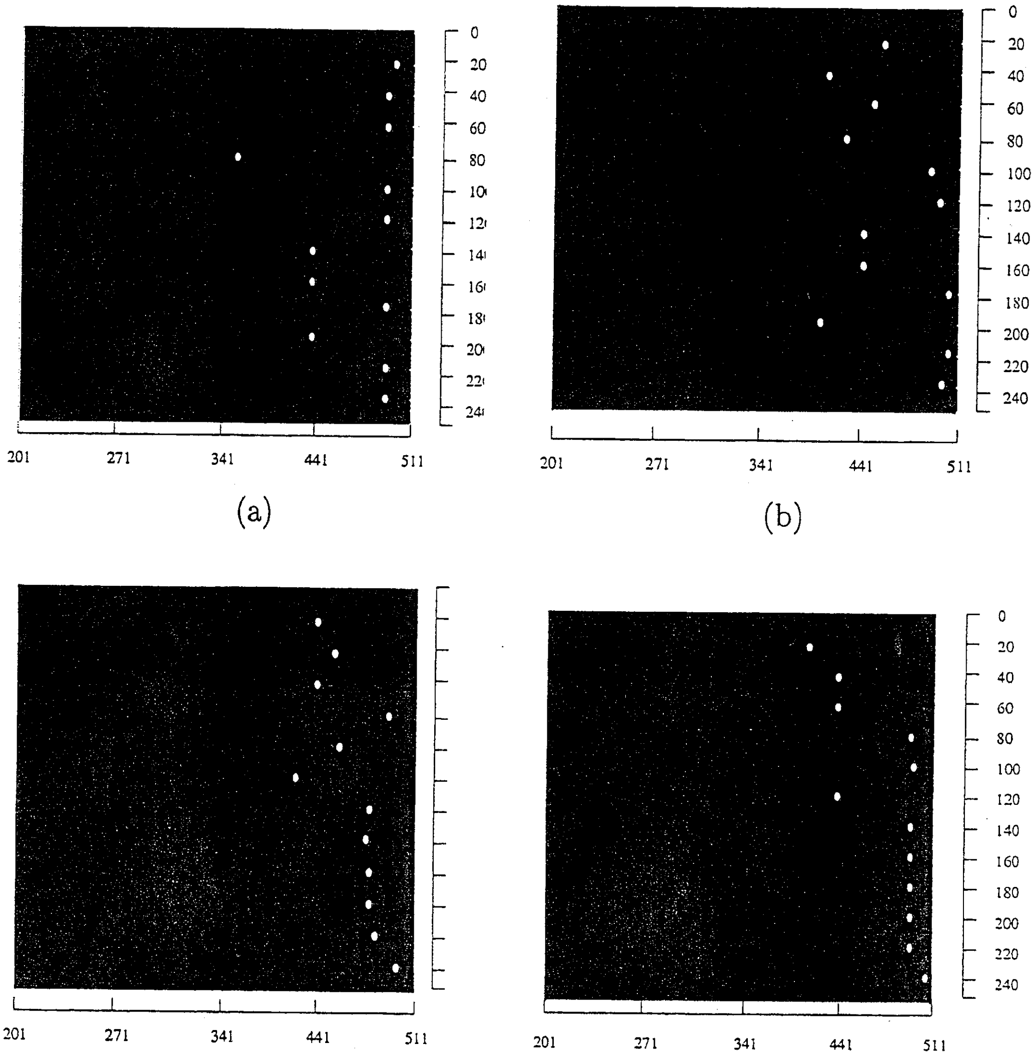


Figure 2: Wigner Distribution of morphed sound and synthesised result

A step of 0.1 involved a 0.1 contribution of first sound and 0.9 contribution of second sound to the morph. Figure 2 shows a resulting SPWD of the morphed sounds and Figure 3 shows the peak of attack features before and after warping. The new signal was then synthesised from its SPWD at each stage of the morph. Preliminary listening tests have found that signals morphed with a step of 0.5 produces a sound that is most easily distinguishable from both the original sounds



(a) (b)

(c) (d)

Figure 3: Peak of attack features: (a) and (d) peaks of attack on cl and tr SPWDs respectively. (b) and (c) peaks of attack on warped results of cl and tr SPWDs respectively.

References

[1] Telleman, Edwin. Lippold. Holloway, Brian. "Timbre Morphing of Sounds with Unequal Numbers of Features. ". J.Audio eng. Soc., Vol. 43, No. 9, Sept. 1991.

[2] Slaney M.,Covell M. and Lassiter B., "Automatic audio morphing." Proc. ICASSP, Atlanta, GA, May 1996.

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

[4] Hope C. and Furlong D., "Time-Frequency distributions for Timbre morphing: the Wigner distribution versus the STFT." Proc. SBCMIV, Brasilia, Brazil, August 1997.

[5] Ullmann, J. R., "An Algorithm for Subgraph Isomorphism". Journal of the Association for Computing Machinery, Vol. 23, No. 1, January 1976.

[6] Messmer, B. T. and Bunke, T. (1995). "Subgraph Isomorphism in Polynomial Time." Technical Report IAM-95-003, University of Bern.

[7] Lysaght T. and Vernon D., "Timbre morphing of synthesised transients using the Wigner Time-Frequency distribution." Proc. SBCMIII, Rio de Janeiro, Brazil, August 1999.

[8] 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.

T. Lysaght, D. Vernon, and J. Timoney, 'Subgraph Isomorphism Applied to Feature Correspondence in Timbre Morphing', Proceedings of Irish Signals and Systems Conference, 2000, University College Dublin, pp. 250-257, June 29-30, 2000. Ed. A. Fagan, O. Feely.