Morph: Timbre Hybridization Tools Based on Frequency Domain Processing

Giovanni Cospito (1)
Raffaele de Tintis (2)
(1) LaSDIM – Laboratorio per la Sperimentazione e la Didattica dell' Informatica Musicale
Sezione Musica Contemporanea, Civica Scuola di Musica di Milano
Corso di Porta Vigentina 15, I-20122 Milano, Italia. Fax +39 2 58314433
gcospito@micronet.it
(2) LIM – Laboratorio di Informatica Musicale. Dipartimento di Scienze dell'Informazione
Università degli Studi di Milano. Via Comelico 39, I-20135 Milano. Italia. Fax +39 2 5500637
rdt@lalim.lim.dsi.unimi.it

Abstract

This paper presents a hybridization method using Morph, a sound timbre hybridization and modeling software running on the Windows 95 platform. Elaboration occurs in Fourier space and is based on analysis/re-synthesis operations through FFT and FFT-1. The hybridization method works through the segmentation of two starting spectra in zones centered on the energy peaks. Each identified spectral zone, is numbered in increasing order so that a matching can be established between pairs of zones with the same value during re-synthesis.

A full description of the Morph elaboration performance includes cross-synthesis, timbre hybridization, harmonic/inharmonic component separation and filtering.

An initial catalogue of hybrid sounds has been compiled using this system, in order to tackle hybridization problems in an organic and systematic manner. The criteria used for the drafting and organization of the catalogue are discussed in the latter part of this paper.

1 The method

Having passed into frequency-domain, the initial sound analysis continues by segmenting each of the two spectra into different zones. As in the spectra modeling techniques following the Serra and Smith (4) method, we proceed by researching spectrum peaks in order to identify the stable component of the sounds being elaborated.

However, our aim here is different, as the two components - stochastic and deterministic - are not processed separately at the end of the analysis, but together.

Once recognized, the spectrum peaks are numbered and used to segment the spectra into different zones, each centered on a peak. During the synthesis of the hybrid result, the computation occurs between those pairs of zones which, in both sounds, are centered around peaks of the same value.

For each pair of zones, elaboration starts from the centers (the identifying peaks) and continues with the lateral components.

Using the two spectrums in *Figure 1* and *Figure 2*, we should match on a logarithmic scale, components 3 and 2 respectively from the first and second spectrum, then the second with the first, and so on.

The procedure can be schematized as:

Mod[t] = exp(C1 * log(Mod1[t]) + (1-C1) * log(Mod2[t]))
Ph[t] = Ph[t-1]+exp(G * log((Ph1[t]-Ph1[t-1])) + (1-G) * log((Ph2[t] - Ph2[t-1])))
FB[t] = exp(C3 * log(FB1[t]) + (1 - C3) * log(FB2[t]))

Mod[t],Ph[t],FB[t] = Moduli, Phases and Frequency-Bins of the new Hybrid Component
Mod1[t], Ph1[t], FB1[t], Mod2[t], Ph2[t], FB2[t] = Moduli, Phases and Frequency-Bins of the Initial Sounds
C1, C2, C3 = Hybridization Coefficients of the Moduli, Phases and Frequency-Bins
t = Analysis Frame Index

With the exception of particular cases, the identified zones will be of different size, and it is therefore possible that a component of one of the two spectra will have no match in the other sound.

In this case, the matching process will be carried out using a frequency and energy component equal to zero. Once the two components forming the hybrid sound are identified, three variables will be assigned: the new phase, the energy and the frequency-bin where the result will be placed. Maintaining two separate variables for phase and frequency-bin, rather than frequency only, allows better flexibility during the elaboration phase.

In fact, if our goal is to generate hybrids, we will go from the two values to the frequency value before proceeding with the elaboration. To produce the new hybrid component, an interpolation on a logarithmic scale will be applied both to moduli and frequencies.

Furthermore, it is also possible to redraw the spectrum by placing the resonance areas characteristic of one sound at the positions (frequency-bins) taken by those of another sound. Being centered on energy peaks, the zones into which the spectra are segmented will identify the resonance areas if those peaks are kept sufficiently low in number. By recomputation alone, and without interpolating energy and frequency values, it is possible to move spectrum areas identified in a sound to the positions taken by those of another sound. A stretching or compressing of the distances between zones of the spectral envelope takes place, creating in one sound the resonances which occur in another.

Fig. 1

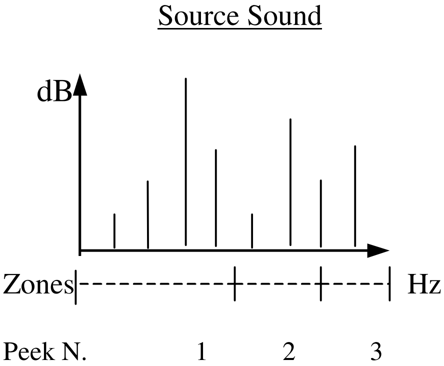


Fig. 2

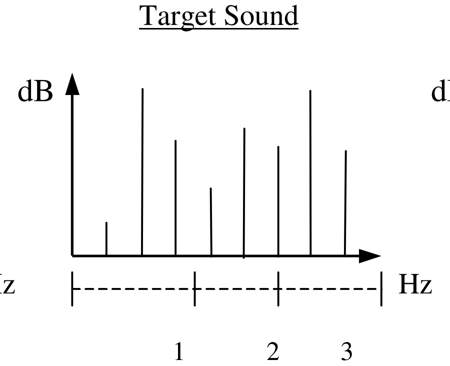
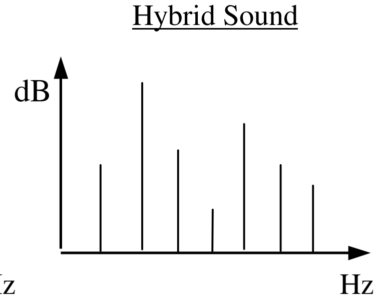


Fig. 3



2 Spectrum Segmentation

Spectrum segmentation is influenced by values occurring in two variables: Selectivity and Harmonicity. After a local maximum within the spectral envelope has been found, the selectivity parameter establishes a threshold below which the next local minimum will be found, so that the maximum may be classified as a peak.
This parameter will, therefore, proportionally influence the number of recognized peaks, and accordingly the areas into which the spectrum will be divided.

The lower this parameter is kept, the wider and better defined the identified areas will be.

Before the hybridization process begins, it is good to choose a selectivity value, for each of the two starting sounds, in order to identify an equal or very similar number of spectral zones in each one, and to be able to match them one-to-one.

The harmonicity parameter has been introduced to support a double function.

Since no *a priori* hypothesis is made on the pitch of the two starting sounds, and it is not possible to predict the positions of the identified spectral peaks, during the first tests, while hybridizing two sounds with a predominantly harmonic structure, the result was often distinctly inharmonic.

Peak positions in the spectrum of the hybrid result became virtually random and the harmonic structure of the starting sounds was lost.

Vice-versa, inharmonic elements introduced in the final result gave a completely new perception compared to the sources, thus nullifying the possibility of considering it a hybrid from a perceptual point of view.

The harmonicity parameter, therefore, serves to contain - and often solve- this problem.

When selecting the peaks, its function is to reduce their number only to those belonging to the series of multiples best represented within the spectrum, with an error inversely proportional to its value.

Keeping this parameter low, the margin for error will be high, and therefore all the identified peaks will be kept. If its value is set high, all peaks away from the harmonic structure will be discarded.

The harmonicity coefficient's second function is to limit the identified spectral zones in order to include only the lateral components nearer to the peaks, i.e. to their center. Starting from these and moving outwards towards the zones' edge, we find more and more inharmonic and noisy partials, which can be excluded in phase of re-synthesis by assigning high values to the harmonicity coefficient.

3 Harmonic / Inharmonic Component Separation

As seen before, in Morph the harmonic/ inharmonic components are re-synthesized together through a single inverse FFT. However, it is possible to isolate one of the components, and to separate both through two re-synthesis operations. Morph classifies spectral peaks as harmonic structure, while components found while moving towards each zone's edge are considered more and more inharmonic.

By using the two selectivity and harmonicity coefficients, it is possible to separate both of the signal's components. By assigning a very low value to the selectivity coefficient, the recognized peaks will be few and well spaced. By further assigning a high value to the harmonicity coefficient, the only peaks that will be kept are those in multiple spectral positions relative to the fundamental.

With the latter parameter kept high, re-synthesis will take place considering only those components very near the peaks, discarding all others. By assigning the maximum value to the harmonicity coefficient, re-synthesis will apply to the peaks alone.

Using a flag in the main panel menu it is then possible to invert the procedure and consider, in a re-synthesis phase, only those components away from the peaks and placed on the edges of each zone which make up the most inharmonic and noisy part of the two starting sounds.

4 An experimental catalogue of hybrid sounds

Morph proposes a general timbre hybridization technique, not tied to the nature of the sound sources but with enough control modes to allow it to adapt to different types of timbres. It was therefore necessary to identify a method which, by circumscribing the vast array of timbres through clear organizational criteria, would allow systematic experimental applications.

We therefore chose timbres exclusively produced by musical instruments, and we then added to the instrument groups the timbre classes, as defined through various methods by organology and western culture literature on orchestration. Organology does not specifically concern itself with the classification of timbres, but it is nevertheless possible to relate the classes of musical instruments to the classes of timbres, as evidently happens during the practice of orchestration.

This choice has the advantage of referring to an accurate and systematic selection of the already available instruments. Another incentive to use this class systemization has been the verification of a remarkable overlap of the classification criteria for musical instruments in organology and the primary parameters in timbre perception taken into account by "Ecological Psychology" (1).

For instance, we can refer to the instrumental classification methods born out of the work of A. Schaeffer in his *Classification des Instruments de Musique* of 1936, in which identification of the two fundamental classes is based on the state of vibrating material (solid and air) and the following subclasses on the physical qualities of the instruments' body (degree of flexibility, elasticity and tension). We immediately notice that these categories for analysis bear a striking similarity to the definition of primary perceptive parameters taken into account by "Ecological Psychology."

The systematic application and experimental method consists of starting with groups of instruments considered as a class of timbres and building hybrid sounds between pairs of instruments belonging to different classes. The organization into classes or groups of instruments/timbres that we considered derives from the classification methods of Curt Sachs and Moritz von Hornbostel in *Systematik der Musikinstrumente* of 1914, those of A. Schaeffer previously mentioned, and the most widely known academic manuals on orchestration.

With the techniques offered by Morph, once the optimal number of peaks is defined, we then have three independent parameters whose values define a hybrid (these parameters or indexes, as previously mentioned, are tied to the module, the phase and the frequency-bin of the spectral components of the two source sounds). The index-adjusting criteria used to produce a hybrid sound are perceptive recognition of both the original sounds' contribution and the possibility to establish an order of approximation of the hybrid towards one of the same sounds.

Having fixed a reasonable order for the hybrids, starting from the first sound towards the second, we obtain a family of functions which identifies the behavior of indexes tied to the construction of the ordered group. These functions have revealed similar behavior in hybridization indexes between instruments belonging to two of the same class. In this phase, experimentation occurs between isolated sounds. The group of sounds from the first instrument is hybridized with the symmetrical group of sounds from the second, i.e. the sounds from both groups are chosen of equal length and height. In this way all problems related to the different behaviors of instruments during the transition between heights and the performance of musical phrases (legato, portamento, articulation, vibrato, etc.) are avoided. The catalogue could also serve as a tool for the experimental identification of continuous timbric spaces.

Bibliography

[1] S. McAdams and E.Bigand (ed.). *Thinking in Sound: the cognitive psychology of human audition*, Oxford University Press, 1993.

[2] S. McAdams, Winsberg, S. Donnadieu, G. De Soete, J. Krimphoff. "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes", *Psychol Res* (1995) 58: 177-192, Springer-Verlag.

[3] CERL SoundGroup. "Timbre Morphing of Sounds with Unequal Numbers of Features", *J. Audio Eng. Soc.*, Vol. 43, N. 9, 1995, September.

[4] X. Serra, J. Smith III. "Spectral Modeling Synthesis: A Sound Analysis/ Synthesis System Based on a Deterministic plus Stochastic Decomposition", *Computer Music Journal*, Vol. 14, No. 4, Winter 1990.

[5] R. de Tintis. "A System Based on Fourier Analysis/ Synthesis for the Hybridization of Sound Timbres", *Proc. of the XII Colloquium on Musical Informatics*, Gorizia, Italy, 1998.

[6] R. de Tintis. "Morph: A System for Timbre Hybridisation", in *Computer Music for the Electronic Musician*, E. Reck Miranda Editor, Focal Press, 1998.

[7] M. H. Serra, D. Rubine, R. Dannenberg. "Analysis and Synthesis of Tones by Spectral Interpolation", *J. Audio Eng. Soc.,* Vol. 38, N.3, 1990, March.

[8] J. Strawn "Analysis and Synthesis of Musical Transitions Using the Discrete Short-Time Fourier Transform", *J. Audio Eng. Soc.*, Vol.35, N. 1/2, 1987, January/February.

[9] D. Arfib. "Analysis, Transformation, and Resynthesis of Musical Sounds with the Help of a Time-Frequency Representation", in *Representation of Musical Signals*, MIT Press, 1991.

[10] D. Arfib, N. Delprat. "Musical Transformation Using the Modification of Time-Frequency Images", *Computer Music Journal*, Vol.17, No.2, Summer, 1993.

[11] "Rapid Measurement of Digital Instantaneous Frequency", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-23, N.2, April 1975.

[12] "A Comparison of Computational Methods for Instantaneous Frequency and Group Delay of Discrete-Time Signals", *J. Audio Eng. Soc.*, Vol.46, N.3, 1998 March.