

# A Modal Distribution Approach to Piano Analysis and Synthesis<sup>1</sup>

Rowena Cristina L. Guevara and Gregory H. Wakefield

*Department of Electrical Engineering and Computer Science  
University of Michigan, Ann Arbor MI 48109  
gev@eecs.umich.edu ghw@eecs.umich.edu*

## Abstract

This paper introduces a method of analysis that points to additive synthesis techniques that accommodate different performance techniques including dynamics, articulation (legato, portato, staccato), and initial damper position. Previous attempts at such synthesis have been hampered by the resolution limits in time and frequency of the short-time spectral analysis used to extract the time-evolving partials. While such analysis sufficiently resolves the long-term decay of each partial of a piano sample, it substantially smooths their onset and thereby robs the synthesized sound of its percussive nature. We present results for piano analysis and synthesis using the modal distribution, which is an alternative representation based on time-frequency (t-f) distributions.

## 1 Introduction

The modal distribution has been designed specifically for signals that can be represented as a sum of isolated time-varying partials [Pielemer and Wakefield, 1996]. Unlike other distributions (e.g., the spectrogram, the Gabor transform, or linear transforms of the Wigner distribution), the modal distribution minimizes the cross-term ambiguities associated with bilinear t-f distributions while maintaining limited superposition. These two properties allow us to apply standard Hilbert techniques for each isolated partial to estimate the instantaneous amplitude and frequency.

The analysis suggests how different piano performance techniques can be incorporated into an additive synthesizer. We present analyses of piano sounds that vary in the above-mentioned performance techniques. Based on these analyses, piano sounds were synthesized. Psychophysical testing demonstrates that the proposed method yields perceptually accurate synthesized piano sounds.

## 2 Analysis

Using a digital audio tape, notes were recorded from three Steinway pianos. The sounds were transferred to a PC and analyzed as follows.

### 2.1 Methodology

The software implements the modal distribution which is a time-smoothed discrete pseudo-Wigner distribution. Time-smoothing is introduced by the cross-term filter which implements cross-term suppression. Partial appears as ridges along the time axis. The frequency support for each ridge defines the neighborhood of points over which instantaneous

power and frequency are estimated. Power is estimated as the sum of the distribution over the ridge neighborhood, while frequency is estimated as the centroid of the ridge neighborhood.

### 2.2 Results of Analysis

The physics of the piano lend certain characteristics to the partial structure of the piano sound [Fletcher and Rossing, 1991]. The soundboard favors certain frequency ranges, leading to weak fundamental for bass notes. The striking position of the hammer produces spectral nulls or attenuated modes. The frequency difference between partials increases with partial number as a consequence of string stiffness.

Aside from the above, the modal distribution of the piano sound reveals unexpected partials. These occur at frequencies that do not belong to the partial series of the note, even when inharmonicity is taken into account. Characterization of these 'rogue' partials for two-stringed notes leads to the conclusion that the slight mistuning of the string unison gives rise to two partial series, one for each string, each with comparable amplitudes and similar attack and decay characteristics. More surprisingly, rogue partials are present in single-stringed notes, but are characterized by lower amplitude and faster attack and decay. This is consistent with the motion of the massive damped bass string as seen by the bridge through which coupling between adjacent strings occurs.

Inclusion of rogue partials in piano synthesis means added computation and memory. The decision to include them was made after conducting an objective, two-interval forced choice discrimination experiment in which the subject's task was to discriminate between synthesized sounds with or without rogue partials. The results for two trained subjects was 100% discrimination; even the two untrained subjects were able to correctly discriminate on 75% of the trials.

<sup>1</sup> This research was supported by an ESEP Scholarship from the Philippine government to the first author and by funds from the Office of the President of the University of Michigan for the MusEn Project.

Some piano compositions call for releasing the damper for a given note only after the following note has been struck. In this case, the strings of the undamped note will vibrate sympathetically with those of the struck note. Usually, the notes are chosen such that there are common partial frequencies among the notes. The following are observed from our analysis of pertinent piano sounds:

- The shape of the amplitude evolution of the partial is unchanged with different initial damper position.
- The upper partials of the struck note are reinforced when strings of higher-pitched notes are undamped.

In our study, we consider three playing articulations: legato (notes are connected smoothly, with neither a perceptible break in the sound nor special emphasis), portato (each of several notes are separated slightly; a style of performance between legato and staccato) and staccato (notes are separated from its neighbors by a perceptible silence of articulation and given a certain emphasis; the opposite of legato); and two levels of dynamics: loud and soft. The following were observed and are consistent with previously published work [Askenfelt, 1990]:

- Across different articulation and dynamics, the shape of the amplitude evolution of a partial of a note is constant, except for the lower partials of staccato which exhibit nulls that may be attributed to the hammer bouncing on the string for this touch. These lower partials have periods comparable to the time that elapse between hammer rebounds.
- The relative amplitude of the partials changes with dynamics, with the higher partials attenuated markedly for soft notes.
- A note played staccato has more energy in the higher-frequency partials than when it is played either legato or portato.

These observations suggest a model for a single piano note that consists of partials with fixed (shape-wise) evolutions and variable initial values that depend on the articulation and dynamics. This model is consistent with additive synthesis.

### 3 Synthesis

The method of synthesis is dictated by the signal model assumed in the analysis. For the modal distribution, synthesis will be additive. Each partial is modeled as a sinusoidal generator with frequency updated at the same rate as the frequency estimates, and amplitude linearly interpolated between amplitude estimates. The impulsive nature of excitation of the piano causes wideband onsets in the modal distribution. Partial locations in the modal distribution are established by looking at time points that are beyond the onset, when the partials appear as salient ridges. Based on

these locations, the ridge neighborhood is chosen and the amplitude and frequency estimates are computed at every display point. The phase of each sinusoid is tracked during each update such that the transition will not introduce discontinuities in the synthesized sound.

The cross-term filter of the modal kernel may smooth the onsets of the amplitude estimates. If we let  $T$  be the length of the impulse response of the cross-term filter and model the initial decay of the partial as exponential, then the decay rate of the partial is preserved for times greater than  $T$  and we can estimate accurately the decay rate. Therefore, a first-order correction to the smoothed onsets of the piano partial extends the exponential backwards in time, up to the actual onset of the sound.

Notes synthesized as described above were presented to subjects in a psychophysical discrimination experiment. Initial performance across all subjects was approximately 60%; after training with feedback, subjects appear to perform no better than 70% suggesting that the synthesized signals are perceptually equivalent to the original signals.

### 4 Future Work

We believe that further refinement of the synthesis model can reduce this discrimination to chance. Comparison of the modal distribution of the original piano onset to that of the synthesized partials reveals the presence of weakly excited resonant frequencies in the original sound which are not included in the synthesis. We are currently exploring the possibility of modeling each partial as a cluster of poles with a dominant pole, where the dominant pole correspond to the partial defined by the amplitude and frequency estimate and the low-order approximation of the rest of the poles represent these weakly excited resonant frequencies. These poles could be considered as the characterization of the transient response of the piano to the impulsive excitation of striking a piano key. The objective is to find a configuration of pole clusters such that the perceptual discrimination between the synthesized and original piano sounds is minimized.

### References

- [Askenfelt, 1990] Anders Askenfelt (Ed.): *Five lectures on the acoustics of the piano*, Royal Swedish Academy of Music, Stockholm, Sweden. 1990.
- [Fletcher and Rossing, 1991] Fletcher, N.H., and Rossing, T.D. *The Physics of Musical Instruments*, Springer-Verlag, New York. Chapter 12.
- [Pielemeier and Wakefield, 1996] William J. Pielemeier and Gregory H. Wakefield. A high-resolution time-frequency representation for musical instrument signals. *J. Acoust. Soc. Am.*, **99** (4): pp. 2382-2396, 1996.