Temporal Filtering: Framing Sonic Objects

Barry Moon

Department of Music, Brown University
email: Barry_Moon@Brown.edu

Abstract

Granular sampling techniques can be used to decontextualize sonic objects. Grains of sound can be recontextualized in continuums of amplitude or timbre by analysis.

1 Introduction

My DVD “Mince” is a study in framing. The camera was used to decontextualize objects and landscapes by framing their smaller parts, while the computer did the same with sound. I have been framing sonic objects (or sonic landscapes if you like) using techniques of granular sampling. It is the filtering, or assemblage of samples into something that is sonically, or musically interesting that is the challenge.

2 Temporal Filtering or Cataloging

I use the term “temporal filtering” to describe techniques where a sound source is sampled when it passes analysis criteria. Automated sample recording and playback based on amplitude analysis is a popular technique in real-time computer music. When I began working on the sound for my DVD, I sampled the radio based on amplitude analysis. A narrow window of extremely low amplitude values was set. The radio was sampled when its amplitude fit within the amplitude window, whereupon the amplitude window was incremented. The outcome was a sound that grew increasingly louder over a duration that I could predetermine. It was an extremely chaotic, noisy sound with a clear trajectory. The down side was that it took many hours to produce several seconds of sound.

When I started applying temporal filtering techniques to spectral analysis, and was having an even harder time matching source sounds to analysis criteria that I thought about the idea of cataloging. I realized that I could put each segment of sound in some form of sonic “catalog” based on analysis. In this case, the source sound was sampled into a position in the sample table, or catalog, according to a discrete value obtained by amplitude or spectral analysis.

For the granular sampling techniques discussed in this paper, I decided to go with a grain of 512 samples, or 11.609977 milliseconds at a sampling rate of 44.1khz. This grain size was chosen because it is long enough to be perceived as having some quality of sound, and short enough so that there is no perceivable change in the quality of sound over the duration of a grain.

2.1 Cataloging by Amplitude Analysis

It can be seen from Figure 1 that this is a fairly straightforward process. The "resolution" of the amplitudes used to catalog the sound is also going to determine the length of the catalog. The avg~ object outputs an average amplitude between 0. and 1., although, it should be noted that dc is the only input signal that can be measured as having an average amplitude of 1.. From testing, I have found ~0.6 to be a practical maximum amplitude. In this case the maximum length of the catalog will be 90000 * 0.6 * 11.609977 milliseconds = 626938.758 milliseconds, or roughly 10 minutes.

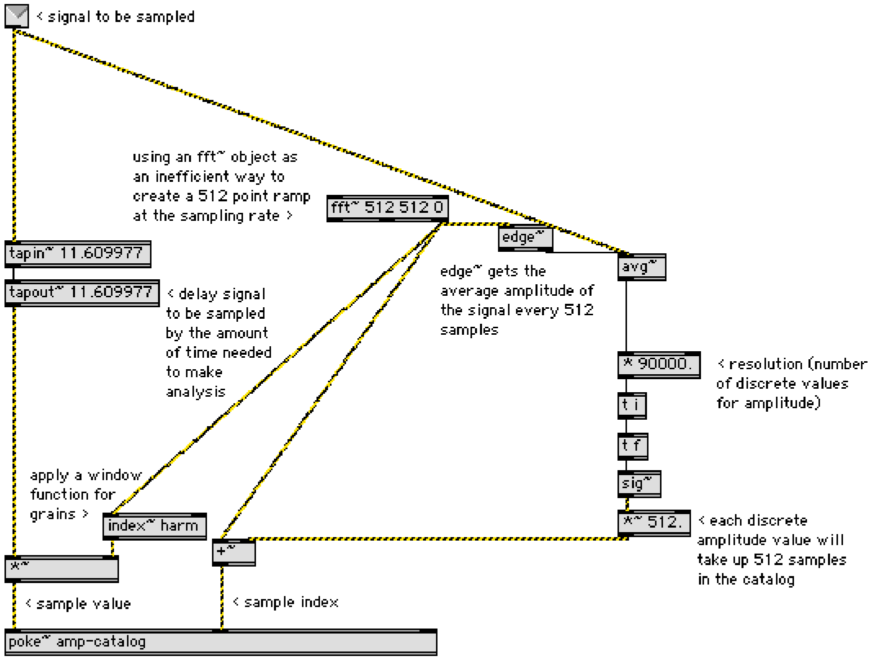

Figure 1. Cataloging by Amplitude

2.2 Cataloging by Zerocrossing Analysis

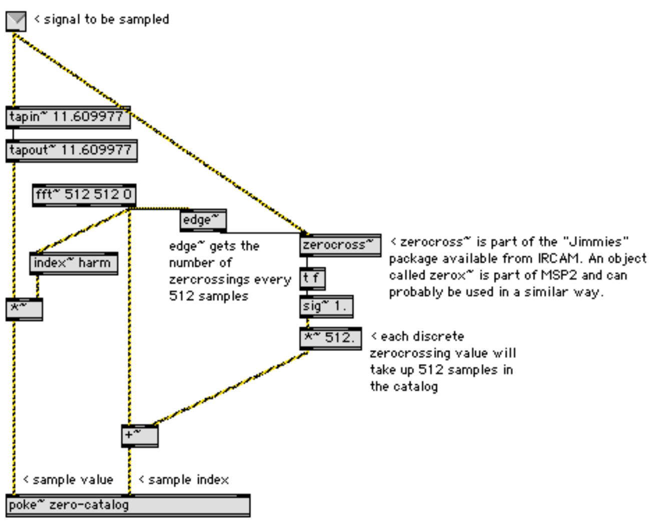

Figure 2a. Cataloging by Zerocrossing

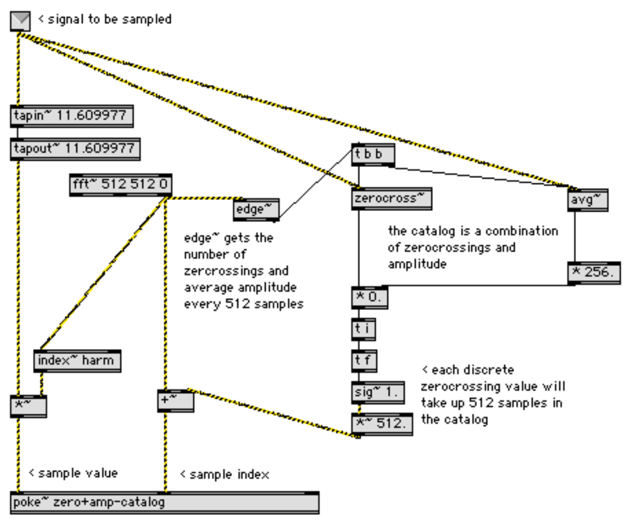

Figure 2b. Cataloging by Zerocrossing and Amplitude

Zerocrossing is another form of analysis that can produce meaningful linear relationships between sounds. However, as opposed to measuring amplitude, measuring zerocrossings can only produce a small range of values. The maximum number of zercrossings possible in an n-point sample would be n-1. And the only sound that could produce n-1 zerocrossings would be at the Nyquist frequency. From testing, I found that white noise is measured with approximately 280 zerocrossings. This can be considered a good maximum number.

Since zerocrossing and amplitude analysis are both somewhat linear, and amplitude analysis gives floating point values, the two can be combined using multiplication as in Figure 2b. I have found it productive to take the inverse of the zerocrossing value (not shown in Figure 2b), since for most signals, extremely low amplitudes tend to be noisy.

2.3 Cataloging by Spectral Analysis

As with cataloging by zerocrossing analysis, cataloging by spectral analysis provides transitions through continuums of timbres. I should point out that using the word "timbre" in this context could be considered problematic. Timbre analysis has traditionally been done using multidimensional models. For example, the model proposed by John Grey and his colleagues (Grey, 1977) relies on both the brightness of spectrum, and how that spectrum changes over time. Since temporal aspects of analysis are built into the granular sampling techniques, I am only concerned with the brightness of spectrum. The two types of spectral analysis I have been experimenting with are channel vocoder (Dudley, 1939), and fft based analysis.

For cataloging by spectral analysis, channel vocoder analysis has an advantage over fft analysis in that frequency bands for analysis can be chosen according to a logarithmic scale. The algorithm in Figure 3 will only detect energy in the lowest 16 bands of the fft analysis. This only extends to around 1400hz. The index output of the fft~ can be scaled to allow more bands to be included in the analysis. However, if it scaled by more than 0.5, the edge~ object can give misleading results because of multiple changes in logic per vector, or changes in logic spanning multiple vectors.

A 16 band resolution in the spectral domain, even if the analysis bands are logarithmically spaced, is fairly inaccurate. But it should be pointed out that the 65536 discrete values obtained by the analysis at this resolution still generate a catalog of around 12 minutes. I was originally thinking of a 31 band resolution, until I realized that the resulting catalog would be around 7 hours in duration (time to buy more RAM!).

Another issue of using the method shown in Figure 3 is that the threshold testing the fft magnitudes produces results that depend strongly on overall amplitude. If the threshold is set high, all quiet signals will be sampled at the minimum point in the catalog (zero), and if set low (as it is in Figure 3), all loud signals will be sampled at the maximum point in the catalog. Using a low threshold works well for my current project, as I am more interested in quiet sounds. If one wanted to make the spectral analysis independent of overall amplitude, the threshold could be set differently for each analysis window by first making an amplitude analysis.

One final word on the technique shown in Figure 3 is that the transition through timbres is in no way linear. Using powers of two, as I have, creates rhythmic changes in timbres. In cases where the catalog has many empty spaces, rhythms become more pronounced.

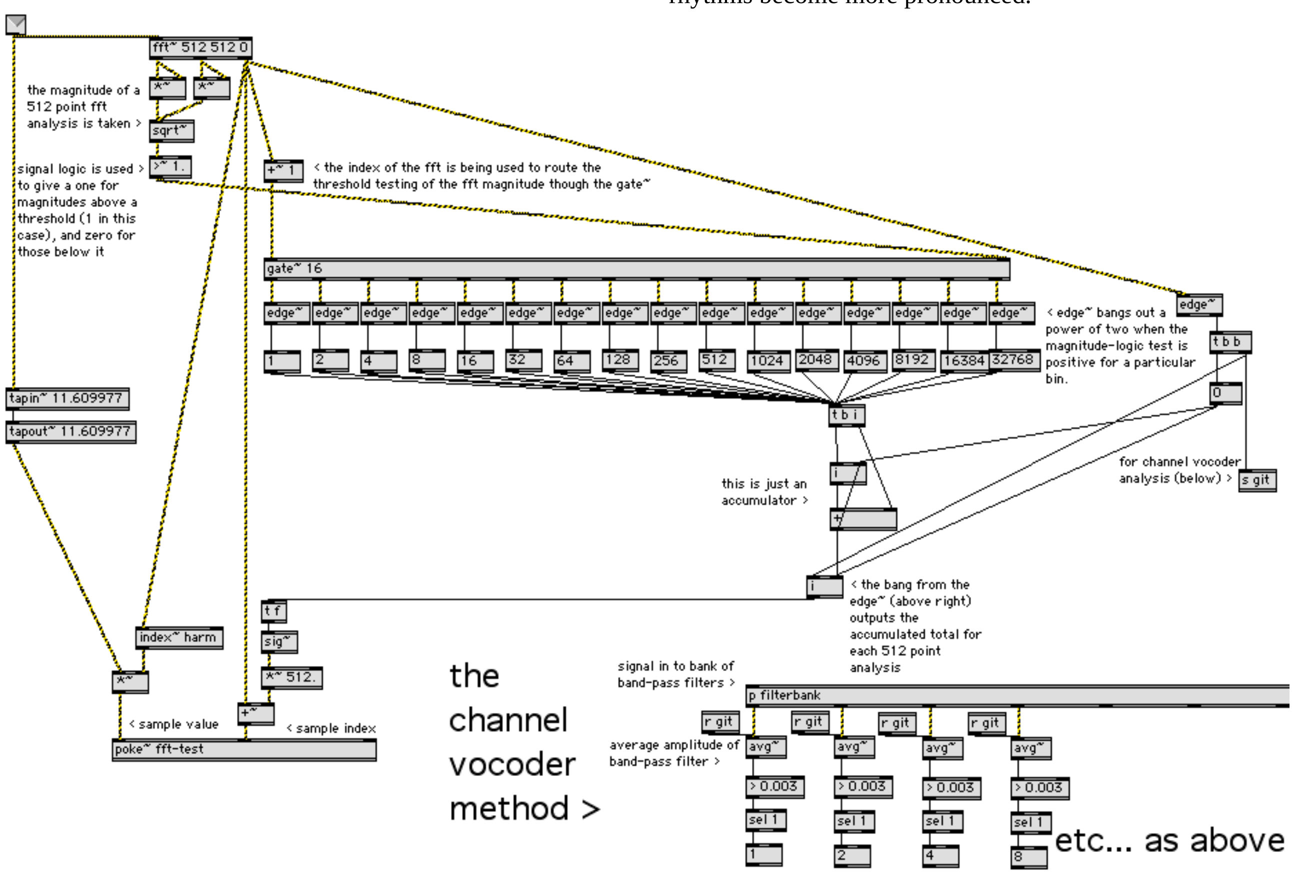

Figure 3. Cataloging by fft magnitude/channel vocoder

3 Results

Cataloging different sound sources using the methods discussed generates highly individual results. Conclusions about the nature of a sound source can be drawn from the sound of the resulting catalog. For example, the only low amplitude signals coming from an alternative rock station on the radio are noise based, whereas low amplitude signals from classical music are of a more traditionally musical nature. Also, using spectral cataloging, there is generally very little recognizably pitched information coming from an alternative rock station, it sounds more like filtered noise, whereas spectral cataloging of classical music reveals very clear pitches. When one makes several catalogs of one particular type of input source according to one parameter of analysis, the catalogs tend to sound very similar. This is useful for overlapping catalogs on a single channel to minimize the granularity of sound, and also for playing the catalogs over multiple channels to create the immersive quality that I desire.

There is still a large amount of redundancy inherent in the cataloging methods. Recording for periods of 12 hours or longer will not necessarily fill each point in a catalog. In this way it still fits my earlier definition of temporal filtering. There are also many points that will not be filled in a spectral catalog using the methods discussed. It may seem obvious, but there are certain combinations of spectral energy that will rarely, if ever be analyzed. For example, it is unlikely that a sound will have measurable energy in only high and low bands without at least the same amount of energy in several bands in between.

4 Conclusion

Temporal filtering, or cataloging techniques described in this paper grew out of techniques developed for real-time interactive music. Applications to acousmatic composition have yielded some interesting results. I am now thinking of ways of applying these techniques to the world of real-time. One interesting application that I have been experimenting with is using a catalog as a lookup table for a live input source. This is a kind of resynthesis using secondary sound sources, or mapping of one sound onto an assemblage of other sounds.

References

Dudley, H. 1939. "Remaking speech." *Journal of the Acoustical Society of America* 11:169-177.

Grey, J. M. 1977. "Multidimensional perceptive scaling of musical timbres." *Journal of the Acoustical Society of America* 61(5):1270-1277.