

The Phase Vocoder: A Tutorial

Mark Dolson

Computer Audio Research Laboratory
Center for Music Experiment, Q-037
University of California, San Diego
La Jolla, California 92093

The phase vocoder is a digital signal processing technique of potentially great musical significance. It can be used to perform very high fidelity time scaling, pitch transposition, and myriad other modifications of recorded sounds. In this tutorial, I attempt to explain the operation of the phase vocoder in terms that musicians can understand.

Introduction

For composers interested in the modification of natural sounds, the phase vocoder is a digital signal processing technique of potentially great significance. By itself, the phase vocoder can perform very high fidelity time-scale modification or pitch transposition of a wide range of sounds. In conjunction with a standard software synthesis program, the phase vocoder can provide the composer with arbitrary control of individual harmonics. But use of the phase vocoder to date has been limited primarily to experts in digital signal processing. Consequently, its musical potential has remained largely untapped. In this article, I attempt to explain the operation of the phase vocoder in terms that musicians can understand. I rely heavily on the familiar concepts of sine waves, filters, and additive synthesis, and I employ a minimum of mathematics. My hope is that this tutorial will lay the groundwork for widespread use of the phase vocoder, both as a tool for sound analysis and modification, and as a catalyst for continued musical exploration.

Overview Historically, the phase vocoder comes from a long line of voice coding techniques which were developed primarily for the electronic processing of speech. Indeed, the word ``vocoder'' is simply a contraction of the term ``voice coder.'' There are many different types of vocoders. The *phase* vocoder was first described in 1966 in an article by Flanagan and Golden. However, it is only in the past ten years that this technique has really become popular and well understood. The phase vocoder is one of a number of digital signal processing algorithms which can be categorized as analysis-synthesis techniques. Mathematically, these techniques are sophisticated algorithms which take an input signal and produce an output signal which is either identical to the input signal or a modified version of it. The underlying assumption is that the input signal can be well represented by a model whose parameters are varying with time. The analysis is devoted to determining the values of these parameters for the signal in question, and the synthesis is simply the output of the model itself. For example, in linear prediction the signal is modeled as the output of a time-varying filter whose input and frequency-response are determined by the analysis. The benefits of analysis-synthesis formulations are considerable. Since the synthesis is based on an analysis of a specific signal, the synthesized output can be virtually identical to the original input; this can occur even when the signal in question bears little relation to the assumed model. Furthermore, the parameter values which are derived from the analysis can be modified to synthesize any

number of useful modifications of the original signal. In this case, however, the perceptual significance and musical utility of the result depends critically on the degree to which the assumed model matches the signal to be modified. In the phase vocoder, the signal is modeled as a sum of sine waves, and the parameters to be determined by analysis are the time-varying amplitudes and frequencies of each sine wave. Since these sine waves are not required to be harmonically related, this model is appropriate for a wide variety of musical signals. As a result, the phase vocoder can serve as the basis for a variety of very high fidelity modifications. It is this feature which makes the phase vocoder so useful in computer music. In the sections which follow, I show in detail how the phase vocoder analysis-synthesis is actually performed. In particular, I show that there are two complementary (but mathematically equivalent) viewpoints which may be adopted. I refer to these as the *Filter Bank* interpretation and the *Fourier Transform* interpretation, and I discuss each in turn. Lastly, I show how the results of the phase vocoder analysis can be used musically to effect useful modifications of recorded sounds.

The Filter Bank Interpretation

The simplest view of the phase vocoder analysis is that it consists of a fixed bank of bandpass filters with the output of each filter expressed as a time-varying amplitude and a time-varying frequency (see Figure 1). The synthesis is then literally a sum of sine waves with the time-varying amplitude and frequency of each sine wave being obtained directly from the corresponding bandpass filter. If the center frequencies of the individual bandpass filters happen to align with the harmonics of a musical signal, then the outputs of the phase vocoder analysis are essentially the timevarying amplitudes and frequencies of each harmonic. But even when this situation does not obtain, the phase vocoder analysis is still surprisingly useful.

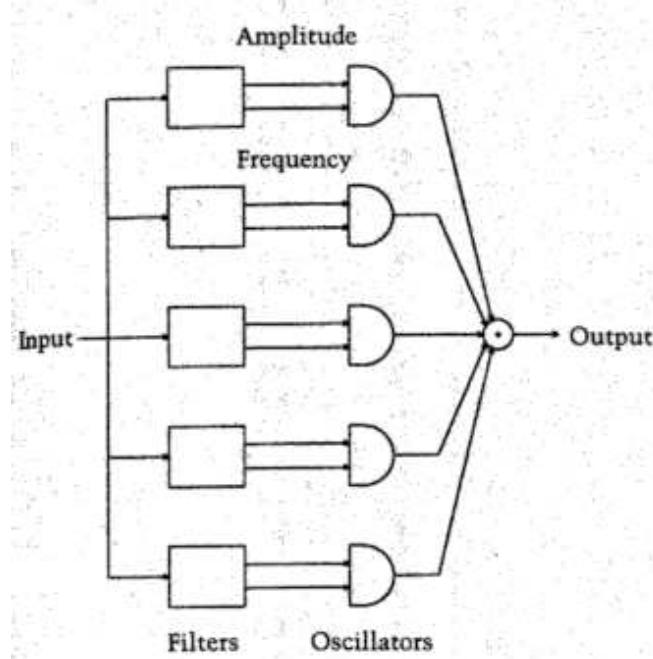


Figure 1. The Filter Bank Interpretation

The filter bank itself has only three constraints. First, the frequency response characteristics of the individual bandpass filters are identical except that each filter has its passband centered at a different frequency. Second, these center frequencies are equally spaced across the entire spectrum from 0 Hz to half the sampling rate. Third, the individual bandpass frequency response is such that the combined frequency response of all the filters in parallel is essentially flat across the entire spectrum. This ensures that no

frequency component is given disproportionate weight in the analysis, and that the phase vocoder is in fact an analysis-synthesis identity. As a consequence of these constraints, the only issues in the design of the filter bank are the number of filters and the individual bandpass frequency response. The number of filters must be sufficiently large so that there is never more than one partial within the passband of any single filter. For harmonic sounds, this amounts to saying that the number of filters must be greater than the sampling rate divided by the pitch. For inharmonic and polyphonic sounds, the number of filters may need to be much greater. If this condition is not satisfied, then the phase vocoder will not function as intended because the partials within a single filter will constructively and destructively interfere with each other, and the information about their individual frequencies will be coded as an unintended temporal variation in a single composite signal. The design of the representative bandpass filter is dominated by a single consideration: the sharper the filter frequency response cuts off at the band edges (i.e., the less overlap between adjacent bandpass filters), the longer its impulse response will be (i.e., the longer the filter will ``ring''). Thus, to get sharp cut-offs with minimal overlap, one must use filters whose time response is very sluggish. In the phase vocoder, this tradeoff is ever-present, and the best solution is generally discovered experimentally by simply trying different filter settings for the sound in question.

Filter Bank The above paragraphs provide an adequate description of the phase vocoder from the standpoint of the user, but they leave unanswered the question of how it actually works. In this section, I show in detail how the output of a single bandpass filter is expressed as a time-varying amplitude and a time-varying frequency. The actual operation of a single phasevocoder bandpass filter is shown in Figure 2. This diagram may appear complicated, but it can easily be broken down into a series of fairly simple mathematical steps.

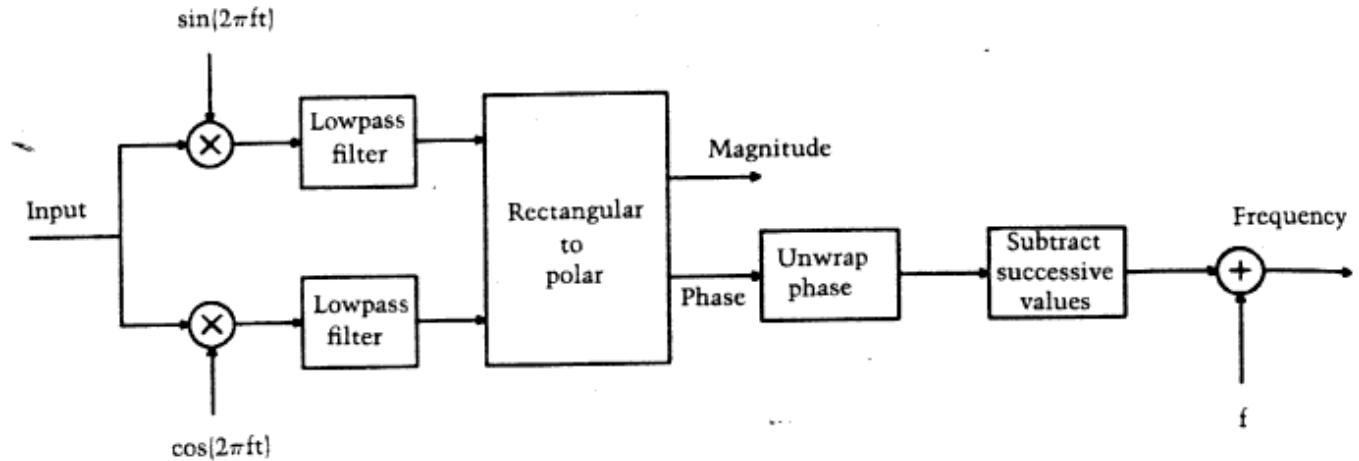


Figure 2. An Individual Bandpass Filter

In the first step, the incoming signal is routed into two parallel paths. In one path, the signal is multiplied by a sine wave with an amplitude of 1.0 and a frequency equal to the center frequency of the bandpass filter; in the other path, the signal is multiplied by a cosine wave of the same amplitude and frequency. Thus, the two parallel paths are identical except for the phase of the multiplying waveform. Then, in each path, the result of the multiplication is fed into a lowpass filter. The multiplication operation itself should be familiar to musicians as simple ring modulation. Multiplying any signal by a sine (or cosine) wave of constant frequency has the effect of simultaneously shifting all the frequency components in the original signal by both plus and minus the frequency of the sine wave. An example of this is shown in Figure 3 in which a 100 Hz sine wave multiplies an input signal of 101 Hz. The result is a sine wave at 1 Hz (i.e., 101 Hz - 100 Hz) and a sine wave at 201 Hz (i.e., 101 Hz + 100 Hz). Furthermore, if this result is now passed

through an appropriate lowpass filter, only the 1 Hz sine wave will remain. This sequence of operations (i.e., multiplying by a sine wave of frequency f and then lowpass filtering) is useful in a variety of signal processing applications and is known as *heterodyning*. Any input frequency components in the vicinity of frequency f are shifted down to the vicinity of 0 Hz and allowed to pass; input frequency components not in the vicinity of frequency f are similarly shifted but not by enough to get through the lowpass filter. The result is a type of bandpass filtering in which the passband is frequency-shifted down to very low frequencies.

$$\sin(A) \sin(B) = \sin(A-B) + \sin(A+B)$$

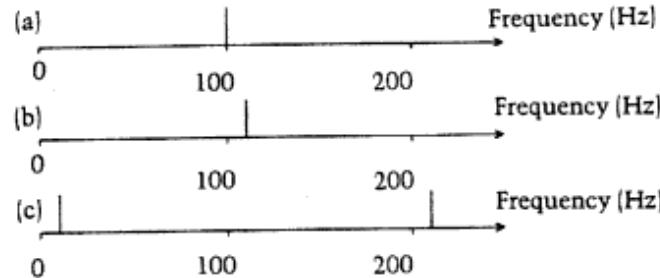
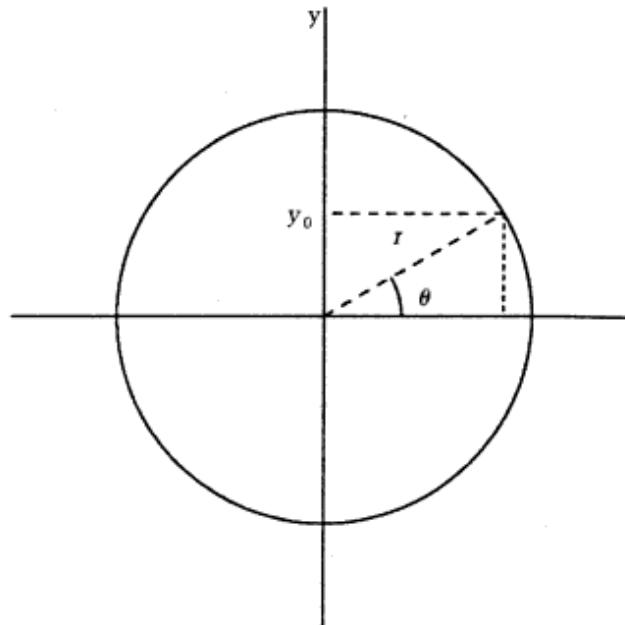


Figure 3. Multiplying Two Sine Waves

In the phase vocoder, heterodyning is performed in each of the two parallel paths. But since one path heterodynes with a sine wave while the other path uses a cosine wave, the resulting heterodyned signals in the two paths are out of phase by 90 degrees. Thus, in the above example, both paths will produce a 1 Hz sinusoidal wave at the outputs of their respective lowpass filters, but the two sinusoids will be 90 degrees out of phase with respect to each other. To understand what the phase vocoder does next with these signals, we can consider the rotating wheel illustrated in Figure 4.



$$r = \sqrt{x_0^2 + y_0^2}$$

$$\theta = \arctan\left(\frac{y_0}{x_0}\right)$$

Figure 4. Rectangular and Polar Coordinates

Suppose that we wish to plot the position of some point on the wheel as a function of time. We have a choice of using "rectangular" coordinates (e.g., horizontal position and vertical position) or "polar" coordinates (e.g., radial position and angular position—also known as magnitude and phase). With rectangular coordinates we find that both the horizontal position and the vertical position are varying sinusoidally, but the maximum vertical displacement occurs one quarter cycle later than the maximum horizontal displacement. With polar coordinates we simply have a linearly increasing angular position and a constant radius. Clearly, the latter description is simpler. The situation within the phase vocoder is very much analogous. The two heterodyned signals can be viewed as the horizontal and vertical signals of the rectangular representation, whereas the desired representation is in terms of a time-varying amplitude (i.e., radius) and a time-varying frequency (i.e., rate of angular rotation). Happily, the translation between the two different representations is easily accomplished. As shown in Figure 4, the amplitude at each point in time is simply the square root of the sum of the squares of the two rectangular coordinates. The frequency cannot be calculated directly, but it can be very well approximated by taking the difference in successive values of angular position and then dividing by the time between these successive values. To see this, we can note that the difference between two successive values of angular position is some fraction of an entire cycle (i.e., a complete revolution), and that "frequency" is simply the number of cycles which occur during some unit time interval. As a result, we need only worry about how to calculate the angular position. Figure 4 also gives a formula for the angular position, but it produces answers only in the range of 0 to 360 degrees. Thus, if we examine successive values of angular position, we may find a sequence such as 180, 225, 270, 315, 0, 45, 90. This suggests that the instantaneous frequency (i.e., rate of angular rotation) is given by the sequence: $(225 - 180)/T = 45/T$, $(270 - 225)/T = 45/T$, $(315 - 270)/T = 45/T$, $(0 - 315)/T = -315/T$, $(45 - 0)/T = 45/T$, $(90 - 45)/T = 45/T$, where T is time between successive values. But the $-315/T$ element is clearly not quite right. What has actually happened is that we have gone through more than a single cycle. Therefore, if we want our frequency calculation to work properly, we should really write the sequence as 180, 225, 270, 315, 360, 405, 450. Now the result of the frequency calculation is, as it should be, a sequence of $(45/T)$'s. This process of adding in 360 degrees whenever a full cycle has been completed is known as *phase unwrapping* (see Figure 5). It is the final necessary step in the sequence of operations which makes the phase vocoder work.

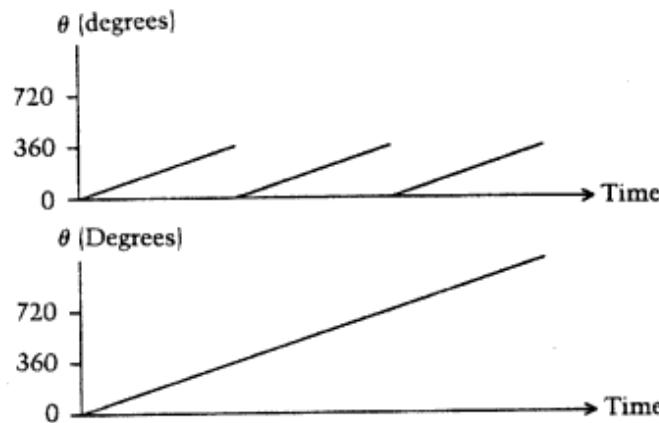


Figure 5. Phase Unwrapping

Thus, the internal operation of a single phase vocoder bandpass filter consists of (1) heterodyning the input with both a sine wave and a cosine wave in parallel, (2) lowpass filtering each result, (3) converting the two parallel lowpass filtered signals to radius and

angular-position signals, (4) unwrapping the angular-position values, and (5) subtracting successive unwrapped angular-position values and dividing by the time to obtain a rate-of-angular-rotation signal. But it should be noted that this rate-of-rotation signal (i.e., the instantaneous frequency) actually refers only to the difference frequency between the heterodyning sinusoid (i.e., the filter center frequency) and the input signal. Therefore the final step is simply to add the filter center frequency back in.

The Fourier Transform Interpretation

A complementary (and equally correct) view of the phasevocoder analysis is that it consists of a succession of overlapping Fourier transforms taken over finite-duration windows in time. It is interesting to compare this perspective to that of the Filter Bank interpretation. In the latter, the emphasis is on the temporal succession of magnitude and phase values in a single filter band. In contrast, the Fourier Transform interpretation focuses attention on the magnitude and phase values for all of the different filter bands or *frequency bins* at a single point in time (see Figure 6).

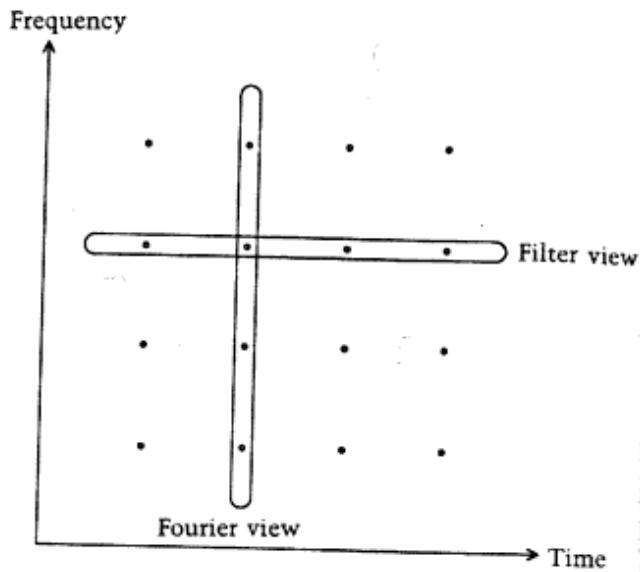


Figure 6. Filter Bank Interpretation vs. Fourier Transform Interpretation

These two differing views of the phase-vocoder analysis suggest two equally divergent interpretations of the resynthesis. In the Filter Bank interpretation (as noted above), the resynthesis can be viewed as a classic example of additive synthesis with timevarying amplitude and frequency controls for each oscillator. In the Fourier view, the synthesis is accomplished by converting back to real-and-imaginary form and overlap-adding the successive inverse Fourier transforms. This is a first indication that the phase vocoder representation may actually be more generally applicable than would be expected of an additive-synthesis technique. In the Fourier interpretation, the number of filters bands in the phase vocoder is simply the number of points in the Fourier transform. Similarly, the equal spacing in frequency of the individual filters can be recognized as a fundamental feature of the Fourier transform. On the other hand, the shape of the filter passbands (e.g., the steepness of the cutoff at the band edges) is determined by the shape of the window function which is applied prior to calculating the transform. For a particular characteristic shape (e.g., a Hamming window), the steepness of the filter cutoff increases in direct proportion to the duration of the window. Thus, again, we see the fundamental tradeoff between rapid time response and narrow frequency response. It is important to understand that the two different interpretations of the phase vocoder analysis apply only to the implementation of the bank of bandpass filters. The operation

(described in the previous section) by which the outputs of these filters are expressed as time-varying amplitudes and frequencies is the same for each. However, a particular advantage of the Fourier interpretation is that it leads to the implementation of the filter bank via the more efficient Fast Fourier Transform (FFT) technique. The FFT produces an output value for each of N filters with (on the order of) $N \log 2N$ multiplications, while the direct implementation of the filter bank requires N^2 multiplications. Thus, the Fourier interpretation can lead to a substantial increase in computational efficiency when the number of filters is large (e.g., $N = 1024$). The Fourier interpretation has also been the key to much of the recent progress in phase-vocoder-like techniques. Mathematically, these techniques are described as Short-Time Fourier-Transform techniques (Rabiner & Schafer, 1978; Crochiere, 1980; Portnoff, 1980; Portnoff, 1981a,b; Griffin & Lim, 1984). Such algorithms may also be referred to as Multirate Digital Signal Processing techniques (for reasons which will be made clear below) (Crochiere & Rabiner, 1983).

Sample-Rate Considerations

The input and output signals to and from the phase vocoder are always assumed to be digital signals with a sampling rate of at least twice the highest frequency in the associated analog signal (e.g., a speech signal with a highest frequency of 5 KHz might be digitized—at least in principle—at 10 KHz and fed into the phase vocoder). However, the sample rates within the individual filter bands of the phase vocoder do not need to be nearly so high. This is most easily understood via the Filter Bank interpretation. Within any given filter band, the result of the heterodyning and lowpass filtering operation is a signal whose highest frequency is equal to the cutoff frequency of the lowpass filter. For instance in the above example, the lowpass filter may only pass frequencies up to 50 Hz. Thus, although the input to the filter was a speech signal sampled at 10 KHz, the output of the filter can be sampled (at least in the ideal case) at as little as 100 Hz without any aliasing error. This is true for each of the bandpass filters, because each filter operates by heterodyning a certain frequency region down to the 0 - 50 Hz region. In practice, the lowpass filter can never have an infinitely steep cutoff. Therefore to really avoid aliasing error, it is advisable to sample the output of the filter at four times the cutoff frequency (e.g., 200 Hz) as opposed to two. Still, this represents an enormous savings in computation (e.g., the filter output is calculated 200 times per second instead of 10,000 times). A detail worth noting here is that this savings is only possible because the filter is a finite impulse response (FIR) filter, (i.e., the present output is calculated entirely on the basis of present and past inputs). If we now seek to resynthesize the original input from the phase vocoder analysis signals, we face a minor problem. The analysis signals (which in the Filter Bank interpretation are thought of as providing the instantaneous amplitude and frequency values for a bank of sinewave oscillators) are no longer at the same sample rate as the desired output signal. Thus, an additional interpolation operation is required to convert the analysis signals back up to the original sample rate. Even so, this is a lot more computationally efficient than avoiding the sample-rate reduction in the first place. In the Fourier Transform interpretation the details of these multiple sample rates within the phase vocoder are less apparent. In the above example, where the internal sample rate is only 2% ($200/10000$) of the external sample rate, we simply skip $10000/200 = 50$ samples between successive FFT's. As a result, the FFT values are computed only $10000/50 = 200$ times per second. In this interpretation, the interpolation operation is automatically incorporated in the overlap-addition of the inverse FFT's. Lastly, it should be noted that we have so far considered the bandwidth of the output of the lowpass filter without any mention of the conversion from rectangular to polar coordinates. This conversion involves highly nonlinear operations which (at least in principle) can significantly increase the bandwidth of the signals to which they are applied. Fortunately, this effect is usually small enough in practice that it can generally be ignored.

Applications

The basic goal of the phase vocoder is to separate (as much as possible) temporal information from spectral information. The operative strategy is to divide the signal into a number of spectral bands, and to characterize the time-varying signal in each band. This strategy succeeds to the extent that this bandpass signal is itself slowly varying. It fails when there is more than a single partial in a given band, or when the time-varying amplitude or frequency of the bandpass signal changes too rapidly. ``Too rapidly'' means that the amplitude and frequency are not relatively constant over the duration of the FFT. This is equivalent to saying that the amplitude or frequency changes considerably over durations which are small compared to the inverse of the lowpass filter bandwidth. To the extent that the phase vocoder does succeed in separating temporal and spectral information, it provides the basis for an impressive array of musical applications.

Historically, the first of these to be explored was that of analyzing instrumental tones to determine the time-varying amplitudes and frequencies of individual partials. This application was pioneered by Moorer and Grey at Stanford in the mid '70's in a landmark series of investigations of the perception of timbre (Grey & Moorer, 1977; Grey, 1977; Grey & Gordon, 1978; Moorer, 1978). (The ``heterodyne filter'' technique developed by Moorer is essentially a special case of the phase vocoder.) More recently, interest in the phase vocoder has focused more on its ability to modify and transform recorded sound materials in musically useful ways. The possibilities in this realm are myriad. However, two basic operations stand out as particularly significant. These operations are *time scaling* and *pitch transposition*.

Time Scaling

It is always possible to slow down a recorded sound simply by playing it back at a lower sample rate; this is analogous to playing a tape recording at a lower playback speed. But this kind of simplistic time expansion simultaneously lowers the pitch by the same factor as the time expansion. Slowing down the temporal evolution of a sound without altering its pitch requires an explicit separation of temporal and spectral information. As noted above, this is precisely what the phase vocoder attempts to do. To understand the use of the phase vocoder for time scaling, it is helpful to once again consider the two basic interpretations described above. In the Filter Bank interpretation, the operation is simplicity itself. The time-varying amplitude and frequency signals for each oscillator are control signals which (hopefully) carry only temporal information. Stretching out these control signals (via interpolation) does not change the frequency of the individual oscillators at all, but it does slow down the temporal evolution of the composite sound. The result is a time-expanded sound with the original pitch. The Fourier transform view of time scaling is more complicated, but it is no less instructive. The basic idea is that in order to time-expand a sound, the inverse FFT's can simply be spaced further apart than the analysis FFT's. As a result, spectral changes occur more slowly in the synthesized sound than in the original. But this overlooks the details of the magnitude and phase signals in the middle. Consider a single bin within the FFT for which successive phase values are incremented by 45 degrees. This implies that the signal within that filter band is increasing in phase at a rate of 1/8 cycle (i.e., 45 degrees) per time interval, where the time interval in question is the time between successive FFT's. Spacing the inverse FFT's further apart means that the 45 degree increase now occurs over a longer time interval. Hence, the frequency of the signal has been inadvertently altered. The solution is to rescale the phase by precisely the same factor by which the sound is being time-expanded. This ensures that the signal in any given filter band has the same frequency variation in the resynthesis as in the original (though it occurs more slowly). The reason that the problem of rescaling the phase does not appear in the Filter Bank interpretation is that the interpolation there is assumed to be performed on the frequency control signal as opposed to the phase. This is perfectly correct conceptually, but the actual implementation generally conforms more closely to the Fourier interpretation. Also, by emphasizing that the time expansion amounts to spacing out successive ``snapshots'' of

the evolving spectrum, the Fourier view makes it easier to understand how the phase vocoder can perform equally well with non-harmonic material. Of course, the phase vocoder is not the only technique which can be employed for this kind of time scaling. Indeed, from the standpoint of computational efficiency, it is probably the very least attractive. But from the standpoint of fidelity (i.e., the relative absence of objectionable artifacts), it is decidedly the most desirable.

Pitch Transposition

Since the phase vocoder can be used to change the temporal evolution of a sound without changing its pitch, it should also be possible to do the reverse (i.e., change the pitch without changing the duration). In fact, this operation is trivially accomplished. The trick is simply to time scale by the desired pitch-change factor, and then to play the resulting sound back at the wrong sample rate. For example, to raise the pitch by an octave, the sound is first time-expanded by a factor of two, and the time-expansion is then played at twice the original sample rate. This shrinks the sound back to its original duration while simultaneously doubling all frequencies. In practice, however, there are also some additional concerns. First, instead of changing the clock rate on the playback digital-to-analog converters, it is more convenient to simply do a sample-rate conversion on the time-scaled sound via software. Thus, in the above example, we would simply designate a higher sample rate for the time-expanded sound, and then sample-rate convert it down by a factor of two so that it could be played at the normal sample rate. It is possible to embed this sample-rate conversion within the phase vocoder itself, but this proves to be of only marginal utility and will not be further discussed. Second, upon closer examination it can be seen that only time-scale factors which are ratios of integers are actually allowed. This is clearest in the Fourier view because the expansion factor is simply the ratio of the number of samples between successive analysis FFT's to the number of samples between successive synthesis FFT's. However, it is equally true of the Filter Bank interpretation because it turns out that the control signals can only be interpolated by factors which are ratios of two integers. Of course, this has little significance for time scaling because, while it may be impossible to find two suitable integers with precisely the desired ratio, the error is perceptually negligible. However, when time scaling is performed as a prelude to pitch transposition, the perceptual consequences of such errors are greatly magnified (by virtue of the ear's sensitivity to small pitch differences), and considerable care may be required in the selection of two appropriate integers. An additional complication arises when modifying the pitch of speech signals because the transposition process changes not only the pitch, but also the frequency of the vocal tract resonances (i.e., the formants). For shifts of an octave or more, this considerably reduces the intelligibility of the speech. (This same phenomena occurs in the pitch transposition of non-speech sounds as well, but for these sounds intelligibility is not an issue.) To correct for this, an additional operation may be inserted into the phase vocoder algorithm as shown in Figure 7. For each FFT, this additional operation determines the spectral envelope (i.e., the shape traced out by the peaks of the harmonics as a function of frequency), and then distorts this envelope in such a way that the subsequent sample-rate conversion brings it back precisely to its original shape.

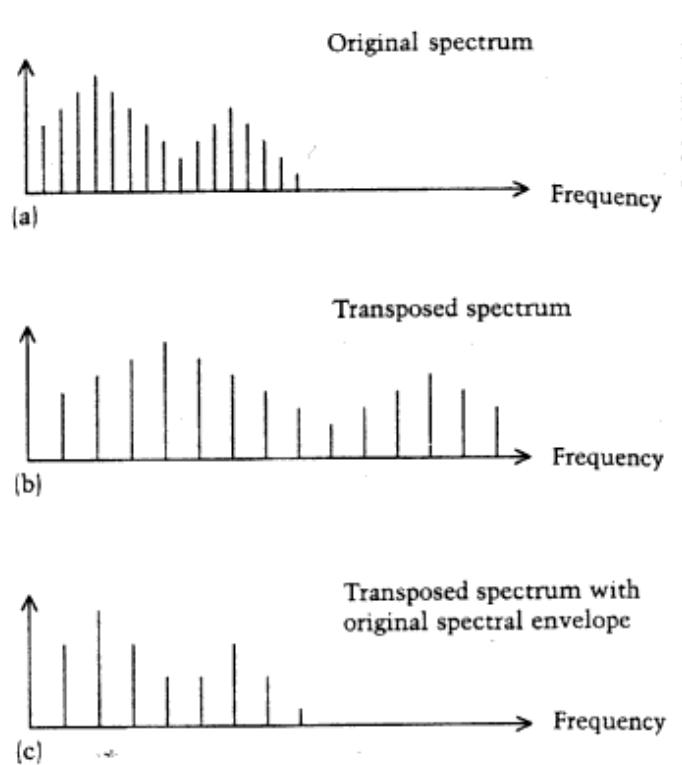


Figure 7. Spectral Envelope Correction

Summary

The above descriptions address only the most elementary possibilities of the phase vocoder technique. In addition to simple time scaling and pitch transposition, it is also possible to perform time-varying time scaling and pitch transposition, time-varying filtering (e.g., cross synthesis), and nonlinear filtering (e.g., noise reduction), all with very high fidelity. A detailed description of these applications, and of their musical implications, is the subject of a separate paper.

References

- Crochiere, R. E. (1980). *A weighted overlap-add method of Fourier analysis-synthesis*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-28(1), 55-69.
- Crochiere, R. E. & Rabiner, L. R. (1983). *Multirate Digital Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall.
- Flanagan, J. L. & Golden, R. M. *Phase vocoder*. *Bell System Technical Journal*, 45, 1493-1509.
- Grey, J. M., & Moorer, J. A. (1977). *Perceptual evaluations of synthesized musical instrument tones*. *Journal of the Acoustical Society of America*, 62, 454-462.
- Grey, J. M. (1977). *Multidimensional perceptual scaling of musical timbres*. *Journal of the Acoustical Society of America*, 61, 1270-1277.
- Grey, J. M., & Gordon, J. W. (1978). *Perceptual effects of spectral modifications on musical timbres*. *Journal of the Acoustical Society of America*, 63, 1493-1500.

Griffin, D. W. & Lim, J. S. (1984). *Signal estimation from modified short-time Fourier transform*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-28(2), 236-242.

Moorer, J. A. (1978) *The use of the phase vocoder in computer music applications*. *Journal of the Audio Engineering Society*, 24(9), 717-727.

Portnoff, M. R. (1980). *Time-frequency representation of digital signals and systems based on short-time Fourier analysis*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-28(1), 55-69.

Portnoff, M. R. (1981a). *Short-time Fourier analysis of sampled speech*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3), 364-373.

Portnoff, M. R. (1981b). *Time-scale modification of speech based on short-time Fourier analysis*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3) 374-390.

Page updated Apr 23, 2014 at 04:22 • Email: jens@panix.com

No unauthorized duplication, copying, mirroring, archival, or redistribution/retransmission allowed! Any offensively categorical statements passed off as facts herein should only be construed as my very opinionated opinions.