International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 6 (2017) pp. 866-871

Khlui-Phiang-Aw Sound Synthesis Using A Warped FIR Filter

Korakoch Saengrattanakul
Faculty of Engineering, Khon Kaen University
Khon Kaen-40002, Thailand.
ORCID: 0000-0001-8620-8782

Kittipitch Meesawat*
Faculty of Engineering, Khon Kaen University
Khon Kaen-40002, Thailand.
*(*Corresponding author)*
ORCID: 0000-0002-1438-1735

Abstract

This article presents an implementation of the warped finite impulse response filter on the khlui-phiang-aw sound synthesis using subtractive method. The khlui-phiang-aw is an air-reed instrument in local Thai music. In general, the subtractive synthesis consists of a signal generator, and a filter. The generator generates a rich spectrum signal at the desired pitch, then the filter shapes the spectrum to the desired frequency profile. This filter can be implemented using a finite impulse response filter, but with a very high filter order. This is because the conventional filter design implies a uniform frequency resolution, while the human auditory frequency resolution is non-uniform. The fine resolution at low frequency results in excessive resolution of the high frequency. Frequency warping is a method that transforms the uniform resolution to a non-uniform one which can be adjusted to conform to the human auditory frequency resolution. The warped finite impulse response filter is a filter that is designed on the warped frequency axis thus reduces the filter order. The filter order can be reduced further by applying the signal decimation prior to the warped finite impulse response filter design. The quality of the synthesized sound is assessed by comparing the spectrum of the original khlui-phiang-aw sound to that of the synthesized sound using the cross-correlation coefficient in the frequency domain. It was found that the synthesis that used the finite impulse response filter yielded a better result. However, the warped finite impulse response filter implementation required a lot less filter order with a subtle sound quality degraded.

Keywords: subtractive synthesis, frequency warping, FIR filter

INTRODUCTION

Khlui – phiang – aw (KPA) is a Thai air-reed woodwind instrument. This kind of instrument is commonly found in southeast Asia with different names and minor characters, such as khlui in Laos, khloy in Cambodia, and suling in Indonesia. A diagram of the KPA is shown in Figure 1.

To play it, the musician blows into the instrument at the air-reed end to generate sound, and control the pitches by using fingers to close or open the tone holes. Its pitches range approximately from A4 to E6, called the middle range in Thai music [1]. The KPA sound synthesis would be beneficial in Thai music composition, teaching, and practicing.

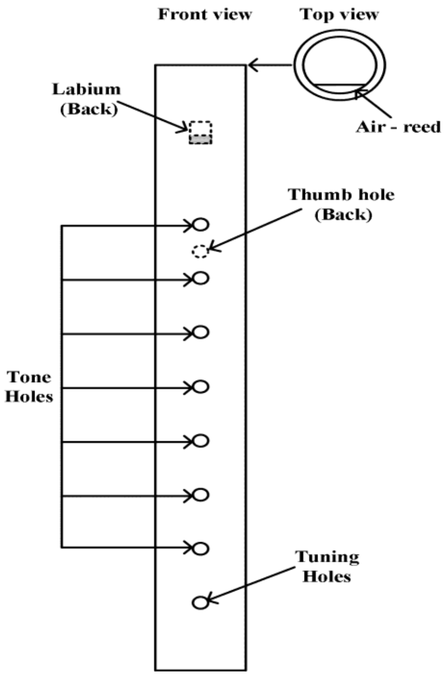

Figure 1: Diagram of the KPA.

The KPA sound synthesis in this article exploits the subtractive technique. The synthesis system consists of a sawtooth waveform generator and a resonance filter. The frequency of the waveform and the frequency response of the filter are adjusted to accommodate the desired pitch.

The sawtooth waveform is chosen because of its richness in harmonics and its resemblance to the physical phenomenon of the air pressure in the KPA resonator duct. To generate the alias-suppressed sawtooth waveform directly in the discrete domain, the differentiate parabolic wave (DPW) method [2], [3] was used in this work.

Once the formant of the KPA sound is found, it is straightforward to design a finite impulse response (FIR) filter for the resonator at any pitch. However, the filter order is drastically high, thus real-time implementation is not practical [4]. The filter order can be reduced using the decimation technique [5] and the frequency warping technique [6], [7]. This article presents an implementation of the decimation technique and the frequency warping technique for the subtractive synthesis of the KPA sound. The quality of the synthesized KPA sound using the proposed method was assessed.

SUBTRACTIVE SYNTHESIS

Subtractive synthesis [8] is a synthesis method that uses a filter to shape the frequency spectrum of the synthesized sound. That is the filter subtracts the excessive spectrum from a spectral rich input signal. The conceptual diagram of the subtractive synthesis is shown in Figure 2.

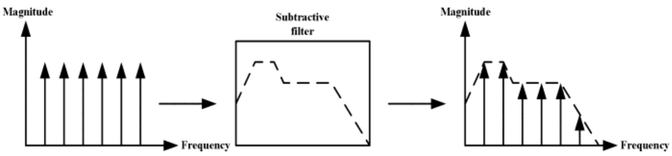

Figure 2: The subtractive synthesis.

The subtractive filter can be implemented using an FIR filter designed by the frequency sampling method [9]. The desired frequency response for each pitch is obtained from the original KPA sound spectrum at that pitch. The simplest realization structure of the FIR filter is shown in Figure 3.

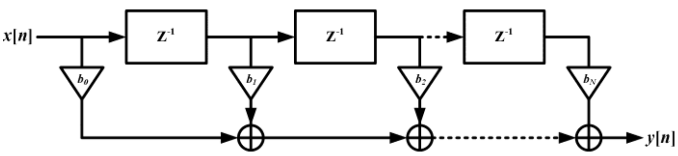

Figure 3: Realization structure of the FIR filter.

SIGNAL DECIMATION

At professional audio standard quality sampling rate, i.e. 48 kHz, an FIR filter requires a high order to be usable in the KPA sound synthesis. Reducing the sampling rate can lower the filter order by the factor of the reduction ratio. The sampling rate can be reduced by the signal decimation [5]. This is possible if the signal does not contain much of the significant information above the half of the reduced sampling rate.

The decimation process is shown in Figure 4. The lowpass filter gets rid of the frequency content of the signal above the haft of the target sampling rate. Then the downsampling removes the excessive data points, hence the desired frequency response shall have a lower data point and therefore a lower filter order.

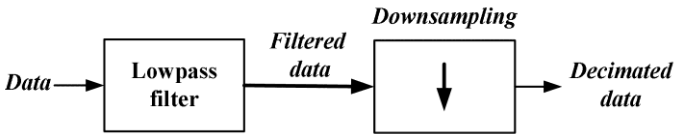

Figure 4: The decimation process.

FREQUENCY WARPING AND WARPED FIR FILTER

The frequency resolution of human hearing is non-uniform. It is high resolution at low frequency, and low resolution at high frequency. In contrast, the frequency resolution of the conventional FIR filter design, such as the frequency sampling method, is uniform. Coping the high resolution at low frequency requires a high order FIR filter, which its resolution is redundant in the high frequency range.

Frequency warping is a process that maps the uniform frequency axis to non-uniform frequency axis [6]. The frequency mapping from the uniform scale to a non-uniform scale is defined by Equation (2) [10]

$$\omega' = \omega + 2\tan^{-1}\left(\frac{\lambda \sin\omega}{1-\lambda\cos\omega}\right), \tag{2}$$

where ω and ω' are the frequency in the linear scale and the non-linear scale, respectively. The λ denotes the warping parameter. Equation (3) determines the appropriate warping parameter for the human hearing called the Bark scale [11] as a function of the system sampling rate (f_s),

$$\lambda = 1.0674\left[\frac{2}{\pi}\arctan(0.06583\frac{f_s}{1000})\right]^{\frac{1}{2}} - 0.1916. \tag{3}$$

The implementation of the frequency warping concept in an FIR filter is usually done as follows: An FIR filter is designed in the warped frequency domain which is defined by the Equation (2). The concept of the warped signal spectrum is depicted in Figure 5. Then the filter is realized by replacing the unit delays with allpass filters that their phase response is the inversed mapping of the Equation (2). This is equivalent to de-warping the designed frequency response. Figure 6 illustrates the de-warping frequency response of the filter. These show that the low frequency details of the prototype frequency response are handled evenly though the frequency axis.

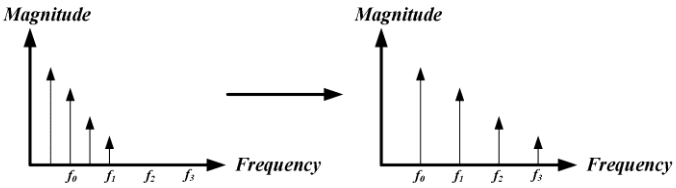

Figure 5: The prototype signal spectrum is warped so the low frequency details are spread through the frequency axis.

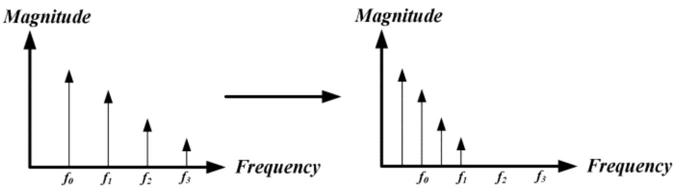

Figure 6: The frequency response of the filter is de-warped back to its original scale by replacing the unit delays in the filter structure with the allpass filters.

The transfer function of the allpass filter for the reverse mapping is given by

$$A(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}}. \tag{4}$$

The structure of the warped FIR filter is shown in Figure 7. Literally, this is not an FIR filter, but rather an IIR filter because its structure contains recursive loops. However, this type of filter is known as the warped FIR (WFIR) filter [12].

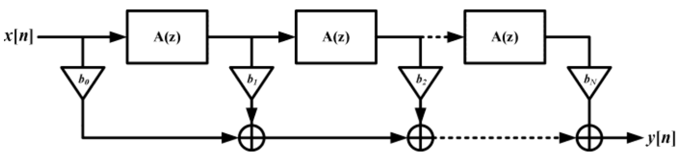

Figure 7: Realization structure of the WFIR filter.

PITCH DETERMINATION

The original KPA sounds were measured for their pitches. The measurement uses a 1.6 mV/Pa sensitivity, 40-15,000 Hz bandwidth dynamic microphone. The KPA sounds were captured by a sound card with a built-in amplifier. The sampling rate was set to 48,000 Hz. The 32,768 data points, which is 0.68 seconds long, each sound sample were measured. All measurements were controlled by a computer.

The pitch of each KPA sound was determined from the frequency at the first peak of the power spectrum density of the sound calculated by the Welch's Method [9]. In this experiment, the 8,192 points FFT and the Blackman-Harris window with a 50 percentages overlap were used. The pitch frequencies of the KPA sounds are shown in TABLE 1.

Table 1. The pitch frequencies of the KPA.

Pitch No.	P1	P2	P3	P4	P5	P6	P7	P8
Frequency (Hz)	486	545	615	662	732	809	938	967

The pitch frequencies of the KPA were used for the sawtooth signal generation. This sawtooth waveform is the source signal to be fed into the subtractive filter in the synthesis process. The differentiated parabolic waveform (DPW) algorithm [2], [3] was used to generate the sawtooth waveform to reduce the aliasing problem while having a low computational cost. The block diagram of the DPW algorithm is shown in Figure 8.

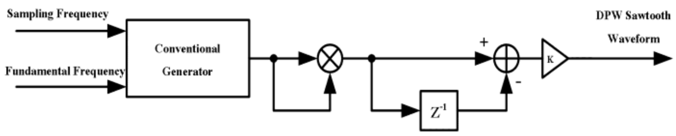

Figure 8: Realization of the DPW algorithm.

SUBTRACTIVE SYNTHESIS USING FIR FILTER

The procedures of the KPA sound synthesis using the FIR filter based subtractive method are described as follows:

The aliases suppressed sawtooth waveform was generated using the DPW algorithm. The generated sawtooth waveform was then fed into an FIR filter.

The frequency response of the FIR filter shall look very close to the original KPA sound spectrum with a small difference. That is because the uncompensated filter requires the input signal to have a flat or a comb-liked spectrum, while the sawtooth waveform spectrum rolls-off as the frequency increases. Therefore, the frequency response of the filter is compensated by the inverse spectrum of the first order autoregressive estimation of the sawtooth waveform. The inverse spectrum of the first order autoregressive estimation of the sawtooth waveform, the frequency response of the uncompensated filter and the compensated filter are shown in Figure 9.

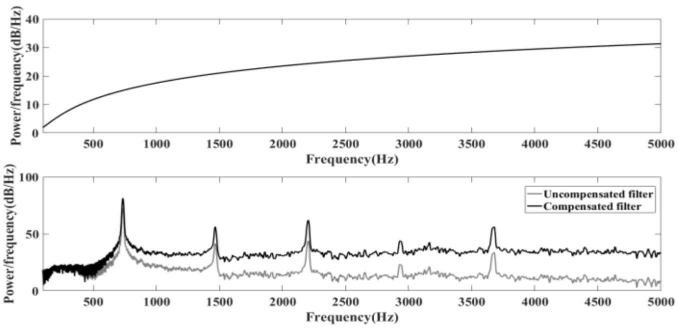

Figure 9: The inverse spectrum of the first order autoregressive estimation of the sawtooth waveform (top), and the frequency response of the 8,192nd order of the uncompensated and the compensated FIR filters (bottom).

After that, the white noise of the same level as the noise floor of the original KPA sound is added to the filtered sawtooth waveform, producing the synthesized KPA sound. Five filter orders, namely, 512nd, 1024th, 2048th, 4096th, and 8192nd orders were tested.

SYNTHESIS USING DECIMATED WFIR FILTER

The KPA sound synthesis using the decimated WFIR filter is similar to that of the FIR filter based synthesis, except that an FIR filter was replaced by a WFIR filter, and the WFIR filter design was done in the decimated domain. The original sound was first decimated. Then the prototype spectrum was determined and was warped. The FIR filter design was taking place in this warped frequency domain. Figure 10 presents the decimated WFIR filter design process.

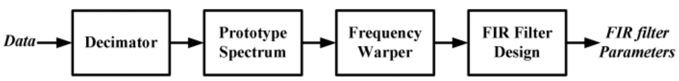

Figure 10: The concept of the WFIR filter design process.

Because the highest significant harmonic of the KPA sound under the range of consideration is less than 6 kHz [13], the operating sampling rate could be reduced to 12 kHz. This was done by using a decimation process with a decimation factor of 4. The spectrum of the decimated KPA sound was determined and was used as the target frequency response for the WFIR filter design.

The determined frequency response was then warped using Equation (2). By using Equation (3), the warping parameter for the sampling rate of 12 kHz was found to be 0.5048. The frequency sampling method was used in the decimated warped frequency domain. After the filter coefficients were obtained, the WFIR filter was implemented by replacing the unit delays with the allpass filters which are defined by Equation (4). Five WFIR filters with the order of 2nd, 4th, 8th, 16th, and 32nd, were tested. The filtered signal was resampled to its original sample rate, i.e. 48 kHz before it was sent to the output.

ASSESSMENT OF THE KPA SIGNAL QUALITY

It is necessary to assess the quality of the synthesized sound. An aspect of the synthesized sound that is related to its quality is its similarity to the original sound in the frequency domain. The cross-correlation coefficient [14] between the spectrum of the original KPA sound and that of the synthesized sound was chosen. The signal spectra were offset by the estimated means of their noise floor levels to avoid the level biases. It works in this particular case because the KPA sound spectrum exhibits the filtered sawtooth waveform spectrum, i.e. there is a dominant peak at the pitch frequency, and other peaks at approximately harmonic frequencies. A signal spectrum with the comparable peaks at the same frequencies yields a high value of the cross-correlation coefficient.

RESULTS AND DISCUSSIONS

The power spectrum density of the original KPA sound, the synthesized sound using the 1024th order FIR filter, and the synthesized sound using the 16th order decimated WFIR filter are shown in Figure 11. They were separated by an 80 dB offset. The spectra show their first peaks with tantamount levels at the same frequency. They also exhibit the similar pattern among their overtones. That is, the spectra of the synthesized sounds are close to that of the original sound with minor differences.

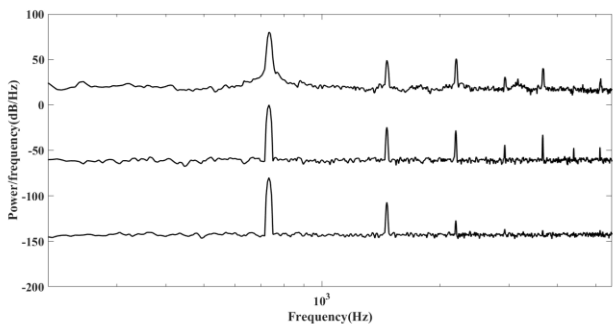

Figure 11: The frequency spectrum of the original (top), the frequency spectrum of the FIR based method (middle), and frequency spectrum of the proposed method (bottom).

The cross-correlation coefficients between the synthesized sound spectra and that of the original sound are shown in

TABLE 2. The label 0 in the table means that the synthesized sound was the unfiltered sawtooth waveform. This was used as the baseline for the worst performance of the synthesizer. The cross-correlation coefficient between the spectrum of the sawtooth waveform, and that of the original KPA sound is 0.4896.

The trend shows that the performance of the synthesizer is better as the filter order increases. Nevertheless, the performance of the FIR based method does not relatively improve much as the filter order goes beyond the 1024^{th} order. In the same manner, increasing the order of the decimated WFIR filter beyond the 16^{th} order does not improve the performance much neither. Therefore, these orders can be considered as the appropriate values for each technique.

Table 2. Cross-correlation coefficients between the spectrum of the synthesized and that of the original sound using an FIR filter and the decimated warped FIR filter at different filter orders.

FIR		WFIR	
Order	Correlation Coefficient	Order	Correlation Coefficient
0	0.4896	0	0.4896
512	0.5456	2	0.6132
1024	*0.7805*	4	0.6471
2048	0.8325	8	0.6954
4096	0.8312	*16*	*0.7132*
8192	0.8192	32	0.7158

It can also be seen that the use of an FIR filter gives the better synthesized sound quality than that of the decimated WFIR filter. Even though at the appropriate order, the cross-correlation coefficient has dropped from 0.7805 to 0.7132, or 8.62%, the filter order is reduced from 1024 to 16, or 98.43%. This is a worth trade-off between the quality of the synthesized sound and the cost of computation.

CONCLUSIONS

This article presents an implementation of the decimated warped FIR filter in the subtractive synthesis used to synthesize the sound of khlui-phiang-aw (KPA), a local Thai air-reed instrument. The quality of the synthesized sounds were indirectly measured using the cross-correlation coefficients between the original KPA sound spectrum and that of the synthesized sounds. It was found that the FIR filter based subtractive synthesis gives a better quality of the synthesized KPA sound, but with a cost of a high filter order. In contrast, the decimated warped FIR filter based synthesizer delivers a slightly lower sound quality, while it requires a notably lower filter order. Justly, this does not impress as it seems since it is a comparison between the FIR filter and the IIR filter. Nevertheless, the lower filter order is preferable in the real-time implementation. The proposed method using the decimated WFIR filter is a rational alternative to the FIR filter in the KPA sound synthesis using the subtractive method.

ACKNOWLEDGMENTS

This work was financially supported by Khon Kaen University, Thailand, under the Incubation Research Project.

REFERENCES

[1] Attakitmongcol, K., Chinvetkitvanit, R., and Sujitjorn, S., 2004, "Characterization of Traditional Thai Musical Scale," WSEAS Trans. On Acoustics and Music, 1(2), pp. 90-95.

[2] Valimaki, V., and Huovilainen A., 2007, "Antialiasing oscillators in subtractive synthesis," IEEE Signal Process, 24(2), pp. 116-125.

[3] Valimaki, V., 2005, "Discrete-time synthesis of the sawtooth waveform with reduced aliasing," IEEE Signal Process. Lett., 12(3), pp. 214-217.

[4] Türckheim, F., Smit, T., and Mores, R., 2010, "String Instrument Body Modeling Using Fir Filter Design and Autoregressive Parameter Estimation," Proc. of the 13th Int. Conference on Digital Audio Effects, Graz, Austria, pp. 1-6.

[5] Ifeachor, E. C., and Jervis, B. W., 2002, Digital Signal Processing A Practical Approach, Pearson Education Asia, India.

[6] Trzos, M., 2011, "Determining the Prediction Order of Warped Linear Prediction for Frequency Warping," Elektrorevue journal, 2(4), pp. 51-57.

[7] Asavathiratham, C., Beckmann, P. E., and Oppenheim, A. V., 1999, "Frequency Warping in the Design and Implementation of Fixed-Point Audio Equalizers," Proc. of 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, pp. 55-58.

[8] Cook, P. R., 2002, Real Sound Synthesis for Interactive Applications, CRC Press, USA.

[9] Proakis, J. G., and Manolakis, D. G., Digital Signal Processing Principles, Algorithms, and Applications,

Prentice Hall, USA.

[10] Lansky, P., and Steiglitz, K., 1981, "Synthesis of Timbral Families by Warped Linear Prediction," Computer Music Journal, 5(3), pp. 45-49.

[11] Smith, III, J., O., and Abel, J., S., 1999, "Bark and ERB Bilinear Transforms," IEEE Transactions on Speech and Audio Processing, 7(6), pp.697-708.

[12] Karjalainen, M., Harma, A., Laine U. K., and Huopaniemi J., 1997, "Warped filters and their audio applications," Proc. of 1997 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp.1-4.

[13] Saengrattanakun, K., Kanjanapadit, J., and Meesawat, K., 2009, "Significance of Harmonics of Klui-Piang-Aw Sounds," Thaksin Journal, 12(3), pp. 89-98.

[14] Taylor, F. J., 1994, Principles of signals and systems, McGraw- Hill.