行政院國家科學委員會補助專題研究計畫成果報告

※※※※※※※※※※※※※※※※※※※※※※※※※※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※　　　　多媒體數位浮水印技術(Ⅲ)　　　　※
※　　　　　　　　　　　　　　　　　　　　　　　　※
※※※※※※※※※※※※※※※※※※※※※※※※※※

計畫類別：□個別型計畫　　□整合型計畫
計畫編號：NSC　90－2213－E－002－094
執行期間：90 年 08 月 01 日 至 91 年 07 月 31 日止

計畫主持人：貝 蘇 章

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
☑出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

執行單位：國立台灣大學電機工程學研究所

中 華 民 國 91 年 08 月 31 日

多媒體訊號處理（III）子計畫一：
多媒體數位浮水印技術（III）
Multimedia Digital Watermarking Techniques（III）

計畫編號：NSC-90-2213-E-002-094
執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：貝蘇章　　台灣大學電機系教授

中文摘要

本研究針對數位聲音訊號提出一展頻浮水印方法。浮水印在加入到主音樂（host audio）之前，先與一數列作相關運算（correlation），以打亂浮水印。把打亂的浮水印直接疊加到主音樂，就形成了藏有浮水印的音樂。

要解出浮水印，需要原始的主音樂。把藏有浮水印的音樂減去原始主音樂，應得到被打亂的浮水印，再把已打亂的浮水印與上述數列作一次相關運算，則可以還原浮水印。

用來作相關運算的數列，具有自相關函數為 delta function 的特性，因此可以提高浮水印的效率與安全性。

此方法可以抵抗 MPEG-1 Layer III（MP3）聲音壓縮攻擊。另外，本研究中也採用了一種定量的方法來量測音樂之間的相似度，以改進現行 MOS 方法的主觀性，進而客觀評判浮水印技術的好壞。

關鍵詞：浮水印、相關運算、MP3。

ABSTRACT

A robust audio watermarking technique using spread spectrum approach is proposed in this paper. The watermark is scrambled by a spread spectrum sequence before it is inserted into the host audio. In the extraction process, the watermark must be restored by the same sequence. The spread spectrum approach can increase both the efficiency and security of the watermarking method. Two kinds of sequences, perfect sequences and uniformly redundant arrays (URA's), are tested. This technique proves to be robust against MP3 attack in most cases except when speech clips are used together with perfect sequences.

Keywords: Watermarking, Spread spectrum, MP3.

1. INTRODUCTION

It has been very common to distribute exact copies of data electronically nowadays. This makes data keeping far more convenient, but other problems may occur. It is possible that illegal copies may be made. To deal with pirating and to protect the intellectual property, the author can place some information (usually called watermark or digital signature) into his/her audio productions without being perceived. Only those with the right key can successfully extract the watermark.

Various techniques have been proposed to cope with pirating problems in audio data. As Tilki et al. mentioned in [1], watermark can be inserted by replacing the Fourier transform coefficients over the middle frequency bands by spectral components of the watermark. In [2], watermark is embedded by modifying phase values of Fourier transform coefficients. Another technique is echo hiding, which employs multiple decaying echoes to place a peak in the cepstrum at a known location [2]. Also, watermark embedding making use of perceptual masking has been proposed by Swanson et al. [3]

In the proposed technique, the watermark is scrambled before it is added into the host audio. This is done by correlating the watermark with a perfect sequence or a uniformly redundant array (URA), whose autocorrelation is a delta function, so that the scrambled watermark can be totally restored by correlating with the sequence again [4,5].

2. SEQUENCES WITH AUTOCORRELATION OF ZERO SIDELOBE

A sequence with autocorrelation of zero sidelobe is equivalent to the direct sequence in spread spectrum communications. After the signal is correlated with the sequence, it becomes a random signal, just like a white noise. If this noise is correlated with the sequence again, the original signal is restored.

Reconstruction using sequences with autocorrelation functions of low sidelobe is mentioned in [4], in which URA (Uniformly Redundant Arrays) was introduced. Perfect sequences were further developed in [5].

2.1 Properties of Perfect Sequences

Assume a perfect sequence $s(n)$ with length N, and its periodic sequence $s_p(n)$ with period N. Some properties of perfect sequences are shown as follows: [5]

2.1.1 Correlation properties

The autocorrelation function, or the PACF (Periodic Repeated Autocorrelation Function) of $s_p(n)$ is

$$\varphi(m)=\sum_{n=0}^{N-1}s_p(n)s_p(n+m) \tag{1}$$

Then

$$\varphi(m)=\begin{cases}E, & m=0\\ 0, & m\neq 0\end{cases} \tag{2}$$

where the energy E of the sequence is given by

$$E=\sum_{n=0}^{N-1}s^2(n) \tag{3}$$

The magnitude of the spectrum of a perfect sequence is always the constant $\sqrt{E}$.

2.1.2 Product theorem of perfect sequences

Consider two periodic perfect sequences $s_1(n)$ and $s_2(n)$ whose periods are N_1 and N_2, with N_1 and N_2 relatively prime, and energy efficiencies η_1 and η_2. Then their product is also a perfect sequence with period $N_1 \cdot N_2$. Also, the energy efficiency of the product sequence is the product of the energy efficiencies of the two original sequences, i.e.

$$\eta = \eta_1 \bullet \eta_2 \tag{4}$$

This can be proven by the definitions of perfect sequences [5].

2.1.3 Synthesis of perfect sequences

From 2.1.2, each perfect sequence $s_p(n)$ possesses a DFT $S_p(k)$ of constant discrete magnitude. This property is used in perfect sequence synthesis. Combining a constant amplitude spectrum any odd-symmetrical phase spectrum

$$\psi(N-k) = -\psi(k), \quad \textit{in the region } 0 \le k < N \tag{5}$$

can always give a real, perfect sequence by inverse DFT.

2.2 Properties of Uniformly Redundant Arrays

Uniformly redundant arrays (URA's), as introduced in [4], are binary matrices with autocorrelation of zero sidelobe. They are firstly developed to enhance the performance of coded aperture array image processing. A complete URA set consists of a pair of matrices, A and G, where A is the key used in the embedding process, and G in the extraction process.

2.2.1 Synthesis of URA's

Given a URA of dimension r by s, it must be satisfied that r and s are both prime numbers and r-s = 2. The elements in the matrix are denoted as $A(i,j)$, where i = 0~r-1, and j = 0~s-1. URA's are generated as follows: [4]

$$\begin{aligned} A(i,j) &= 0 \quad \textit{if } i = 0 \\ &= 1 \quad \textit{if } j = 0,\ i \neq 0 \\ &= 1 \quad \textit{if } C_r(i)C_s(j) = 1 \\ &= 0 \quad \textit{otherwise} \end{aligned} \tag{6}$$

where $C_r(i) = 1$ if there exists an integer x, $1 \le x < r$ such that $i = x^2 \bmod r$

$= -1$ otherwise

The extraction key G is generated by assigning

$$\begin{aligned} G(i,j) &= 1 \quad \textit{if } A(i,j) = 1 \\ &= -1 \quad \textit{if } A(i,j) = 0 \end{aligned} \tag{10}$$

2.2.2 Correlation Property

The circular correlation function of A and G is a 2-D delta function with the element in the intersection of the first column and the first row proportional to the number of 1's in A, which is the value $(rs+1)/2$, and the rest all zeros.

3. PROPOSED TECHNIQUE

3.1 Watermark Embedding

The proposed technique makes use of spread spectrum approach and repeated insertion. The watermark W is first correlated with a sequence P, resulting in a noise-like signal I. This signal is scaled by a factor k and added into the host audio A, producing the stego audio S.

$$S = A + kI = A + k(W \otimes P) \tag{7}$$

where $\otimes$ stands for correlation.

3.2 Watermark Extraction

In the extraction process, it is necessary to refer to the original host audio A. The received stego audio S is subtracted by A, obtaining the noise-like signal I, which is then correlated with the sequence P to restore the watermark W.

$$\begin{aligned} W' &= (S-A) \otimes P \\ &= kI \otimes P \\ &= k(W \otimes P) \otimes P \\ &= kW \otimes (P \otimes P) \\ &= kW \otimes \delta \\ &= kW \end{aligned} \tag{8}$$

The block diagrams of watermark embedding and extraction are shown in Figs. 1 and 2.

3.3 Repeated Insertion

When I is added into the host audio A, repeated insertion is adopted. That is, each sample in I is added into several samples in A. This concept is illustrated in Fig. 3 [6].

4. AUDIO SIMILARITY MEASURE

This technique proves to be robust against MPEG I—Layer III (MP3) compression attack. To grade the watermark quality objectively and quantitatively, the approach in [7,8] is incorporated.

In the quality evaluation process, Measuring Normalizing Block (MNB) technique is developed. There are several steps in the measurement process:

(1) Two signals are normalized by removing the mean values and normalized to a common RMS level.

(2) Each signal is broken into overlapping frames. Each frame is multiplied by Hamming window and transformed into frequency domain. Only the samples of DC to Nyquist are retained.

(3) Select frames with energy above a given threshold. Transform the frequency domain samples into dB scale by taking logarithm.

(4) Apply the Frequency MNB (FMNB).

(5) Apply either Time MNB (TMNB) structure 1 or structure 2.

(6) Apply linear combination and logistic function to obtain Acoustic Distance (AD) and Logistic Function of AD (L(AD)).

The range of AD is from 0 to infinity, and the range of L(AD) is from 1 to 0. For AD closer to 0 and L(AD) closer to 1, the two audio signals are of higher similarity perceptually. For two identical signals, AD is 0 and L(AD) is 0.9909 [7,8].

5. EXPERIMENTAL RESULTS

5.1 Data Profile

Two types of audio clips are tested in this research, including both music and speech. The data profiles are listed in Tables 1 and 2. Two clips are used as host audio

while four are used as watermark audio clips. All the clips are 16-bit PCM stereo WAV format except that the host ones are of sample rate 44100 Hz and the watermark ones are of sample rate 44100/6 = 7350 Hz. Both perfect sequences and URA's are tested, and the experimental results are shown.

5.2 Using Perfect Sequences

The perfect sequence is used to scramble the watermark before watermark embedding. The extracted watermark after the stego audio has gone through MP3 compression and decompression is compared with the original watermark and the similarity results corresponding to different repeating block sizes are measured, as shown in Figs. 4 and 5. This value stands for the robustness of the watermarking technique against MP3 attack. The MP3 encoding specification tested here is the standard bitrate of 128kb/s.

5.3 Using URA's

The similarities of extracted and original watermarks using URA's under MP3 attack with different repeating block sizes are also plotted in Figs. 4 and 5.

6. DISCUSSION

If the extracted watermark qualities under two combinations are carefully investigated, it is clear that URA outperforms the perfect sequences no matter what types of audio clips are under consideration.

In both cases, the host and watermark combinations can be classified into two categories: "music in music" and "speech in speech", with the former combination more robust against MP3 attack. Also, the repeated insertion does not work much in audio watermarking. Given the same host audio, watermark audio, and scrambling sequence, larger repeating block sizes do not necessarily yield better results in similarity measurement. Therefore, the major factor that counts in robustness improvement is not the repeating block size, but the efficiency of the scrambling sequence, in the URA case, the number of 1's in the matrix.

7. CONCLUSION

An audio watermarking technique based on the spread spectrum approach is proposed in this paper. Sequences with autocorrelation function of zero sidelobe are introduced, investigated, and tested in the experiments. Also, their results under MP3 compression attack are presented in a new objective and quantitative audio similarity measure.

The experimental results show that URA provides better robustness for watermark embedding than the perfect sequence.

Repeated insertion is adopted but proved not very promising in robustness improvement. The major part that improves the robustness is the sequence correlation.

The main contribution of this research is that audio clips are used as watermark other than the commonly used binary sequences. The employment of audio clips as watermark introduces much more challenges than binary signals, but has also pointed out another way on digital watermarking. Watermarks can be larger and more meaningful signals other than binary sequences, carrying more information about the author, owner, or the creation. It will be of wide applications in the future multimedia and internet oriented environments.

8. REFERENCE

[1] J.F. Tilki, A. A. Beex, "Encoding a Hidden Digital Signature Onto an Audio Signal Using Psychoacoustic Masking", Proc. 1996 7th Int. Conf. On Sig. Proc. Apps. And Tech., Boston, MA 1996.

[2] W. Bender, D. Gruhl, N. Morimoto, "Techniques for Data Hiding", Tech. Rep., MIT Media Lab, 1994

[3] M. D. Swanson, B. Zhu, A. H. Tewfik, L. Boney, "Robust Audio Watermarking Using Perceptual Masking", Signal Processing 66, 1998

[4] E. E. Fenimore, T. M. Cannon, "Coded Aperture Imaging With Uniform Redundant Arrays", Applied Optics Vol. 17 No.3, 1 Feb 1978

[5] H. D. Luke, "Sequences and Arrays With Perfect Periodic Correlation", IEEE Trans. Aerospace and Electronic Systems, May 1998

[6] C. H. Lee, Y. K. Lee, "An Adaptive Digital Image Watermarking Technique For Copyright Protection", IEEE Trans. Consumer Electronics, Vol. 45, No. 4, Nov 1999

[7] S. Voran, "Objective Estimation of Perceived Speech Quality – Part I: Development of the Measuring Normalizing Block Technique", IEEE Trans. Speech and Audio Processing, Vol. 7, No. 4, Jul 1999

[8] S. Voran, "Objective Estimation of Perceived Speech Quality – Part II: Evaluation of the Measuring Normalizing Block Technique", IEEE Trans. Speech and Audio Processing, Vol. 7, No. 4, Jul 1999

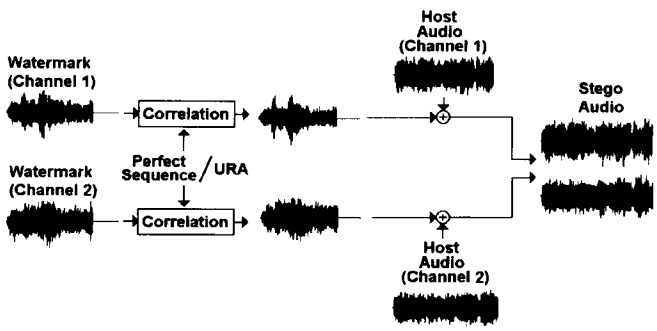


Fig. 1 Illustration of watermark embedding

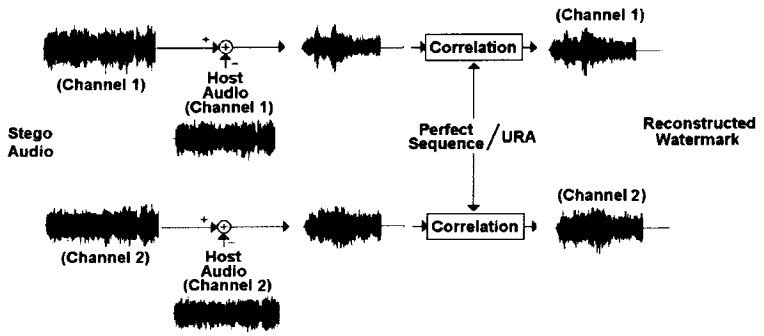


Fig. 2 Illustration of watermark extraction

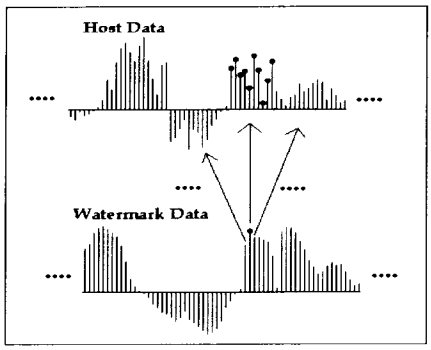


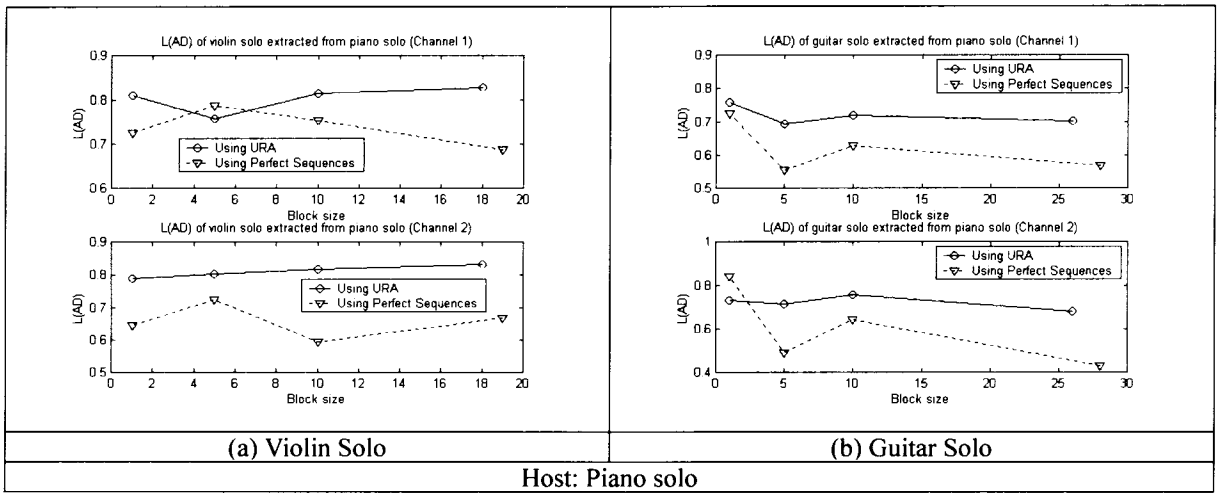
Fig. 3 Illustration of repeated insertion

Table 1: Audio clips used as host audio (sample rate = 44100Hz)

1	PIANO	Piano Solo: "Etude Op.25 No.12, Chopin"	13.531 seconds
2	SPFLE	English Speech: "Time Magazine", Female	17.580 seconds

Table 2: Audio clips used as watermark (sample rate = 7350Hz)

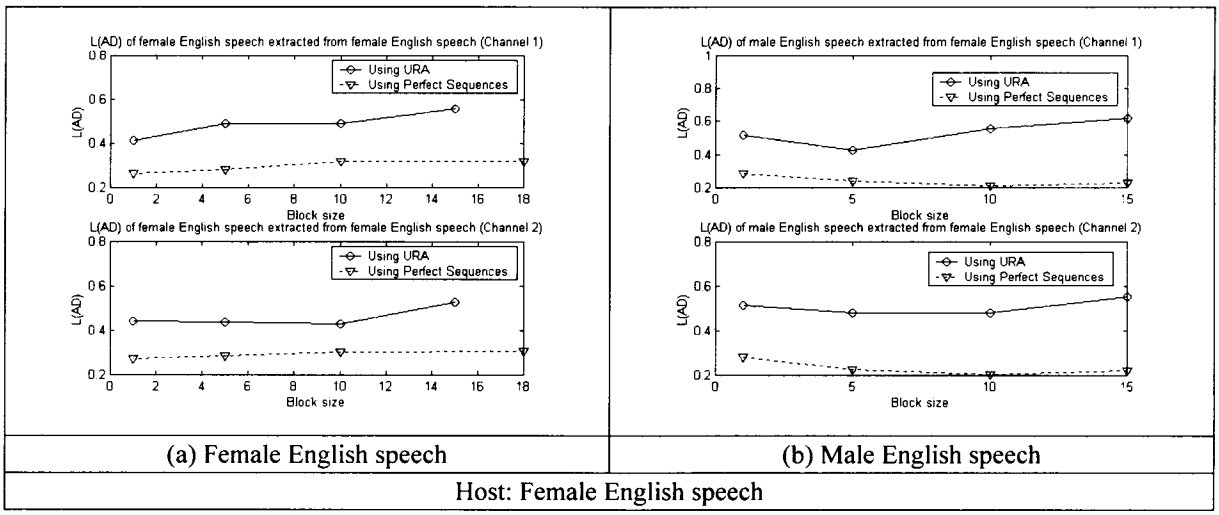
1	VOILN	Violin Solo: "Hungarian Dances No.1, Brahms"	4.232 seconds
2	GUITR	Guitar Solo: "Petenera Para Guitarra"	2.900 seconds
3	SPFSE	English Speech: "from Time Magazine", Female	5.538 seconds
4	SPMSE	English Speech: "from Time Magazine", Male	6.531 seconds



(a) Violin Solo | (b) Guitar Solo

Host: Piano solo

Fig. 4 Similarity values of music clips extracted from piano solo



(a) Female English speech | (b) Male English speech

Host: Female English speech

Fig. 5 Similarity values of English speech clips extracted from English speech clips

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

91 年 06 月 10 日

報告人姓名	貝蘇章	服務機構及職稱	台大電機系 教授
會議時間 會議地點	91/05/26~91/05/29 美國亞歷桑那州,鳳凰城	本會核定補助文號	國科會專題計劃補助
會議名稱	(中文)2002 國際電路與系統會議 (英文)2002 IEEE Int'l Symposium on Circuits and Systems		
發表論文題目	(中文)分數希伯特轉換器之設計 (英文) Maximally Flat Allpass Fractional Hilbert Transformers		

報告內容:

一、 出席 2002 國際電路與系統會議

二、 參加會議經過及與會心得

三、 大會的重點及發展趨勢

四、 攜回資料名稱

五、 結論

出席國際會議報告

『2002 國際電路與系統會議』(2002 IEEE Int'l Symposium on Circuits and Systems)假美國亞歷桑那州,鳳凰城市盛大舉行，筆者應邀出席該會，發表論文一篇，茲將參加會議的心得分述於下：

一、本會共有一仟伍佰多位來自世界各地四十多個國家的學者、專家與會，約有 1100 篇論文在大會發表，內容著重在電路、VLSI、CAD、數位訊號處理，影像壓縮與視訊、電力及系統、多媒體等是個內容非常豐富及具學術權威地位的重要會議。

二、此次大會的主辦單位為亞歷桑那州立大學電機系及國際電機電子工程師學會、電路與系統分會協辦(IEEE Circuits Systems Society)，其目的

想藉此會交換各國最新「電路與系統」科技，並促進學術交流。

三、鳳凰城市為亞歷桑那州的中心，會議地點為近郊的史考特(Scottsdale, Phoenix)渡假避暑勝地，風景優美，設備優良，是個理想的會議場所。

四、會議的頭一天共有 13 個短期課程班(Tutorials)，讓與會者有再接受新科技之機會。大會的學術論文發表方式分成 15 個場地同時進行，另外有個較大的場地以海報方式舉行，並有書展及電路系統及產品展示，海報方式能與原作者親自相互討論及交換心得，效果十分良好。

五、大會在開幕式邀請二位專家，發表學術專題演講，內容精采，受益良多。此次大會本人受邀擔任「多媒體系統」及「數位濾波器」之 Session Chair，主持論文發表。

六、大會並在星期二晚上舉行晚宴，並安排音樂演奏及印地安舞蹈表演，以宴與會嘉賓。

七、本人此次承蒙「國科會專題計劃」補助旅費，得以順利成行，在此深表謝意，並攜回大會論文集之光碟一片，引進及吸收不少新知及科技發展，供國內參考，並達到學術交流的目的。

台大電機系教授

貝蘇章　敬上　91/06/10

MAXIMALLY FLAT ALLPASS FRACTIONAL HILBERT TRANSFORMERS

Soo-Chang Pei and Peng-Hua Wang

Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.
Email address: pei@cc.ee.ntu.edu.tw

ABSTRACT

Recently, a generalization of the Hilbert transformer, the fractional Hilbert transformer, was defined and developed. In this paper, we propose a design of the allpass filter to realize the fractional Hilbert transformer based on the maximally flat approximation to the desired phase response. The coefficients are solved analytically for the traditional Hilbert transformer which is a special case of the fractional Hilbert transformer. Based on the closed-form coefficients, we show that the maximally flat allpass Hilbert transformers are stable. Design examples indicate that the proposed filters exhibit good approximation to the desired frequency response.

1. INTRODUCTION

Hilbert transform (HT) is a basic and important tool in signal processing. In a communication system, it is used for single side-band modulation and then reduces the bandwidth needed for transmission. HT is also used for edge detection [1]. In [2], the HT is generalized into the fractional Hilbert transform (FHT) and two alternative definitions are given. One is a modification of the definition of HT and the frequency response is expressed by

$$H_{fht}(\omega) = \begin{cases} e^{-j\alpha\pi/2} & \text{for } \omega > 0 \\ e^{j\alpha\pi/2} & \text{for } \omega < 0 \end{cases} \tag{1}$$

Note that this FHT can be regarded as an allpass (AP) filter with a suitable specification of phase response. Another is implemented by using the above definition as well as the fractional Fourier transform to achieve a two-parameter FHT system. The discrete version of the FHT is proposed in [3].

Both FIR and IIR filters are investigated to realize the HT in extensive literature. In [4] and [5], the impulse responses of the FIR HTs are analytically solved in the maximally flat (MF) sense. A realization scheme based on decomposing the transfer function of HT into allpass subfilters is proposed in [6].

In this paper, the FHTs with MF phase response will be designed. The filter coefficients are obtained by solving linear equations. For the HT, a special case of the FHT, the coefficients are solved analytically. We show that the MF AP HTs are always stable by applying the Eneström-Kakeya theorem to the closed-form coefficients. The stability problem for the general FHTs is illustrated by numerically showing that the poles lie in the unit circle $|z| < 1$ for a certain α.

2. MAXIMALLY FLAT DESIGN OF AP FHT

The transfer function $H(z)$ of an Nth-order AP filter can be represented by

$$H(z) = z^{-N}\frac{\sum_{n=0}^{N} a_n z^n}{\sum_{n=0}^{N} a_n z^{-n}} = z^{-N}\frac{A(z^{-1})}{A(z)} \tag{2}$$

where the coefficients a_n's are real. Without loss of generality, we let $a_0 = 1$ to prevent from the null solution for a_n. The phase response $\theta(\omega)$ of $H(e^{j\omega})$ can be expressed by

$$\theta(\omega) = -N\omega + 2\tan^{-1}\frac{\sum_{n=0}^{N} a_n \sin n\omega}{\sum_{n=0}^{N} a_n \cos n\omega}. \tag{3}$$

Given the desired frequency response $\theta_d(\omega)$, we want to find a set of coefficients a_n's so that the phase error $\theta_e(\omega) = \theta_d(\omega) - \theta(\omega)$ is minimized. By expressing the desired phase response as $\theta_d(\omega) = -N\omega + \tilde{\theta}(\omega)$, the phase error $\theta_e(\omega)$ can be represented by

$$\theta_e(\omega) = \tilde{\theta}(\omega) - 2\tan^{-1}\frac{\sum_{n=0}^{N} a_n \sin n\omega}{\sum_{n=0}^{N} a_n \cos n\omega}. \tag{4}$$

However, it is difficult to minimize the above error function due to its nonlinearity. Therefore, we minimize an equivalent error function $e(\omega)$ which is defined by

$$e(\omega) = \tan\frac{1}{2}\tilde{\theta}(\omega) - \frac{\sum_{n=0}^{N} a_n \sin n\omega}{\sum_{n=0}^{N} a_n \cos n\omega}. \tag{5}$$

It is easy to show that $e(\omega) = n(\omega)/d(\omega)$ where

$$n(\omega) = \sum_{n=0}^{N} a_n \sin\left[\frac{1}{2}\tilde{\theta}(\omega) - n\omega\right] \tag{6}$$

$$d(\omega) = \cos\frac{1}{2}\tilde{\theta}(\omega)\sum_{n=0}^{N} a_n \cos n\omega. \tag{7}$$

Finally, the numerator function $n(\omega)$ is the target error function to be minimized.

Let $\tilde{\theta}(\omega)$ in Eq.(6) be the ideal phase response expressed as

$$\tilde{\theta}(\omega) = \begin{cases} -\frac{1}{2}\alpha\pi, & \text{for } 0 < \omega < \pi; \\ \frac{1}{2}\alpha\pi, & \text{for } -\pi < \omega < 0 \end{cases} \tag{8}$$

where α is the desired fractional parameter. Putting the phase response $\tilde{\theta}(\omega)$ into Eq.(6), we can obtain the following equations

$$n(\omega) = -\sum_{n=0}^{N} a_n \sin\left(n\omega + \frac{1}{4}\alpha\pi\right), \text{ for } \omega > 0; \tag{9}$$

$$n(\omega) = \sum_{n=0}^{N} a_n \sin\left(-n\omega + \frac{1}{4}\alpha\pi\right), \text{ for } \omega < 0. \tag{10}$$

If $n(\omega)$ is minimized at $\omega = \omega_0$ in the MF sense, it has to satisfy the MF conditions of $d^k n(\omega)/d\omega^k = 0$ at $\omega = \pm\omega_0$. Because of the odd symmetry between Eq.(9) and (10), we can drop the case of $\omega_0 > 0$. Let $d^k n(\omega)/d\omega^k = 0$ at $\omega = \omega_0$ where $n(\omega)$ is expressed by Eq.(9). We have

$$\sum_{n=1}^{N} a_n n^k \sin\left(n\omega_0 + \frac{1}{2}k\pi + \frac{1}{4}\alpha\pi\right) = -\delta_k \sin\frac{1}{4}\alpha\pi \tag{11}$$

where δ_k is defined by

$$\delta_k = \begin{cases} 1, & \text{for } k = 0; \\ 0, & \text{otherwise.} \end{cases}$$

The coefficients a_n's can be solved by the N linear equations expressed in Eq.(11) for $k = 0, 1, \ldots, N-1$. In this paper, we choose the frequency $\omega_0 = \frac{1}{2}\pi$ to achieve the best approximation on the middle frequency range.

3. EXPLICIT SOLUTION FOR HILBERT TRANSFORMER

In the previous section, we show that the coefficients of the AP FHTs can be obtained by solving Eq.(11). The resulting FHTs exhibit the best approximation at $\omega = \frac{1}{2}\pi$ in the MF sense. However, these linear equations are ill-conditioned. For example, the reciprocal of the condition number is approximated to 1.281817×10^{-26} for the equations of Eq.(11) with $N = 20$ and $\alpha = \frac{1}{2}$ solved by MATLAB or similar software. That is, we can not solve Eq.(11) reliably for large N. The problem of numerical instability may be avoided by solving the equations analytically.

In this section, we will analytically solve the coefficients of MF AP FHTs for $\alpha = 1$ in Eq.(11). That is, we will find the closed-form solution of the AP MF HTs. Letting $\alpha = 1$ in Eq.(11), we can express the equations as

$$\sum_{n=1}^{N} a_n n^k \left(\cos\frac{1}{2}n\pi + \sin\frac{1}{2}n\pi\right) = -\delta_k, \text{ for even } k,$$

$$\sum_{n=1}^{N} a_n n^k \left(\cos\frac{1}{2}n\pi - \sin\frac{1}{2}n\pi\right) = 0, \text{ for odd } k. \tag{12}$$

The above equation can be solved by the ratio of two Vandermonde's determinants. If N is even, after some algebraic manipulations, we obtain the closed form of a_n's of

$$a_{2m} = \frac{\left(\frac{1}{2}\right)_m}{\left(M+\frac{1}{2}\right)_m}\binom{M}{m} \tag{13}$$

for $m = 0, 1, \ldots, M$, and

$$a_{2m+1} = -\frac{(M-m)\left(\frac{1}{2}\right)_m}{\left(M+m+\frac{1}{2}\right)\left(M+\frac{1}{2}\right)_m}\binom{M}{m} \tag{14}$$

for $m = 0, 1, \ldots, M-1$ where $M = N/2$, $\binom{M}{m}$ is the binomial coefficient, and $(x)_n$ is the Pochhammer's symbol defined by $(x)_0 = 1$ and $(x)_n = x \times (x+1) \times \cdots \times (x+n-1)$. On the other hand, if $N = 2M+1$, we have

$$a_{2m} = -a_{2m+1} = \frac{\left(\frac{1}{2}\right)_m}{\left(M+\frac{3}{2}\right)_m}\binom{M}{m} \tag{15}$$

for $m = 0, 1, \ldots, M$. Based on Eq.(15), it is obvious that $1 - z^{-1}$ is a factor of the denominators $A(z)$ of the odd-order MF AP HTs. Accordingly, we have the following property:

Property 1 *If N is an odd number and $M = (N-1)/2$, the denominator $A(z)$ of an Nth-order MF AP HT can be factored as $A(z) = (1 - z^{-1})\hat{A}(z^2)$ where*

$$\hat{A}(z) = \sum_{m=0}^{M} \hat{a}_m z^{-m}, \text{ and } \hat{a}_m = \frac{\left(\frac{1}{2}\right)_m}{\left(M+\frac{3}{2}\right)_m}\binom{M}{m}. \tag{16}$$

Remark. In[7] the authors analytically solved the coefficients of the MF AP orthonormal symmetric wavelet filters. These coefficients can be expressed by

$$b_m = (-1)^m \frac{(-M+K/4)_m}{(1+K/4)_m}\binom{M}{m} \tag{17}$$

where K must be odd to satisfy the orthonormal condition. The odd-order MF AP HTs with coefficients expressed in Property 1 can be related with the MF AP orthonormal symmetric wavelet filters by $\hat{a}_m = (-1)^m b_m$ for $K = 4M+2$.

Based on Eq.(1), one may synthesize the Nth-order FHTs by

$$H_\alpha(z) = \left(\cos\frac{1}{2}\alpha\pi\right) z^{-N} + \left(\sin\frac{1}{2}\alpha\pi\right) H_1(z) \tag{18}$$

where $H_1(z)$ is an Nth-order HT. That is, high order FHTs with flat phase response can be synthesized by the MF AP HTs without solving the ill-conditioned Eq.(11). We will prove that these FHTs are always stable by showing that the MF AP HTs are stable. However, the FHTs synthesized by AP HTs are not allpass filters. It is obvious that the FHT synthesized by Eq.(18) has the magnitude of $\cos\frac{1}{2}\alpha\pi + \sin\frac{1}{2}\alpha\pi$ at $\omega = 0$ if $H_1(z)$ is the MF AP HT. This DC value reaches its largest value of $\sqrt{2} > 1$ when $\alpha = 1/2$. Hence, a scaling factor is necessary to reduce this peak magnitude. A modified FHTs is proposed by $H'_\alpha(z) = kH_\alpha(z)$ where $k = (\cos\frac{1}{2}\alpha\pi + \sin\frac{1}{2}\alpha\pi)^{-1/2} = (1+\sin\alpha\pi)^{-1/4}$.

4. STABILITY PROBLEM

To test the stability, one can apply the Schur-Cohn criterion or the more efficient Jury-Marden criterion [9] on the proposed AP filters. Thiran show that the MF AP fractional delay filters were stable by applying the Schur-Cohn criterion[8]. However, there exists other stability criteria suitable for the proposed filters without evaluating the Schur-Cohen determinants or establishing the Jury-Marden arrays. In this paper, we will apply the Eneström-Kakeya theorem [10] which is stated as

Theorem 1 *Let $p(x) = \sum_{n=0}^{N} a_n x^{N-n}$, $N \geq 1$, be a polynomial with $a_n > 0$ for $0 \leq n \leq N$. Let $r_n = a_{n+1}/a_n$ for $0 \leq n < N$. Then all the zeros of $p(x)$ are contained in the annulus*

$$\min_n r_n \leq |x| \leq \max_n r_n.$$

Based on the Eneström-Kakeya theorem, we have the following properties about the stability of the MF AP HTs.

Property 2 *All the poles of the MF AP HTs of even order are contained in the unit circle $|z| < 1$.*

Proof: Let $B(z) = A(-z)$ where $A(z)$ is the denominator of an even-order MF AP HT with coefficients expressed in Eqs.(13) and (14). Then all the coefficients of $B(z)$ satisfy

$$\begin{aligned} r_{2m} &= -\frac{a_{2m+1}}{a_{2m}} = \frac{2M-2m}{2M+2m+1} < 1 \\ r_{2m-1} &= -\frac{a_{2m}}{a_{2m-1}} = \frac{2m-1}{2m} < 1. \end{aligned}$$

According to the Eneström-Kakeya theorem, we conclude that all the zeros of $B(z)$ are contained in the unit circle $|z| < 1$ since the ratios of successive coefficients are less than unity. Then all the zeros of $A(z)$ are also in $|z| < 1$.

Property 3 *The MF AP HT of odd order has a pole at $z = 1$. All the other poles are contained in the unit circle $|z| < 1$.*

Proof: By applying the Eneström-Kakeya theorem to $\hat{A}(z)$ in Property 1, we have this property.

We can not conclude that the general MF AP FHTs are stable. However, by numerically computing the poles of the AP FHTs, the largest moduli of these poles are less than unity within a range of interest. Fig. 1 shows the plot of largest moduli for $2 \leq N \leq 16$ and $0 < \alpha < 1$.

5. DESIGN RESULTS

Fig. 2 and 3 show the design results of the MF AP FHTs . Fig. 2 is the plot of the phase responses of the 10th- and 11-order MF AP FHTs for $\alpha = 0.1, 0.3, 0.5, 0.7$ and 0.9. The phase responses are normalized by $2\{\arg[H(\omega)] + N\omega\}/\pi$. The purpose of the normalization is to find out the approximation to the desired α. There are bumps around $\omega = 0$ for odd-order filters. Figs. 3 and 4 show the magnitude and normalized phase responses for the FHTs synthesized by the 30th-order MF AP HT. We can not obtain these FHTs by solving Eq.(11) due to numerical instability. The phase responses shown in Fig. 3 exhibit good approximation to the desired phase responses within the middle frequency band. However, the magnitude responses can not remain unity over the whole band. Fig. 4 shows the magnitude responses which are scaled according the discussion in Section 3.

6. CONCLUSIONS

The MF AP FHTs are proposed in this paper. The coefficients of the FHTs are obtained by solving a set of linear equations. For the special cases of the HTs, the coefficients are solved analytically. Based on the closed-form expressions, we prove that the MF AP HTs are stable according to the Eneström-Kakeya theorem. The stability of the general MF AP FHTs is investigated by numerically calculating their poles for a certain range of α and N. The largest pole in modulus is less than unity for $N \leq 16$. However, since the general FHTs can be synthesized by the HTs, we show that IIR FHTs with flat phase response can be implemented by the MF AP HTs.

7. REFERENCES

[1] K. Kohlmann, "Corner detection in natural images based on the 2-D Hilbert transform," *Signal Processing*, vol. 48, pp. 225-234, 1996.

[2] A. W. Lohmann, D. Mendlovic and Z. Zalevsky, "Fractional Hilbert transform," *Optics Letters*, vol. 21, no. 4, pp. 281-283, February 1996.

[3] S.-C. Pei and M.-H Yeh, "Discrete fractional Hilbert transform," *Proc. IEEE Int. Symposium on Circuits and Systems*, Monterey, California, May 1998.

[4] T. Cooklev and A. Nishihara, "Maximally flat FIR Hilbert transformer," *Int. J. of Circuit Theory Applicat.*, vol. 21, pp. 353-570, 1993.

[5] J. Le Bihan, "Coefficients of FIR digital differentiators and Hilbert transformers for midband frequencies," *IEEE Trans. Circuits and Systems-II*, vol. 43, no. 3, pp. 272-274, March 1996.

[6] H. Johansson and L. Wanhammar, "Digital Hilbert transformers composed of identical allpass subfilters," *Proc. IEEE Int. Symposium on Circuits and Systems*, Monterey, California, May 1998.

[7] X. Zhang, T. Muguruma, and T. Yoshikawa, "Design of orthonormal symmetric wavelet filters using real allpass filters," *Signal Processing*, vol. 80, pp. 1551–1559, 2000.

[8] Jean-Pierre Thiran, "Recursive digital filters with maximally flat group delay," *IEEE Trans. Circuit Theory*, vol. CT-18, no. 6, pp. 659–664, November 1971.

[9] A. Antoniou, *Digital Filters: Analysis, Design, and Applications*, 2nd ed., McGraw-Hill, 1993.

[10] N. Anderson, E. B. Saff, and R. S. Varga, "On the Eneström-Kakeya theorem and its sharpness," *Linear Algebra and its Applications*, vol. 28, pp 5–16, 1979.

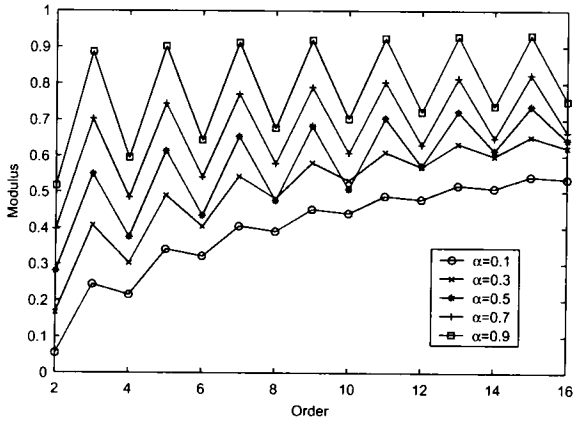


Fig. 1. The plot of the poles of largest modulus for $0 < \alpha < 1$ and $2 \leq N \leq 16$. These moduli are less than unity indicate the corresponding MF AP FHTs are stable.

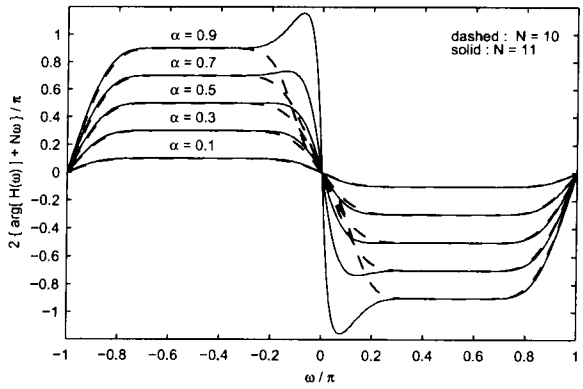


Fig. 2. The plot of normalized phase responses for $N = 10$ and 11.

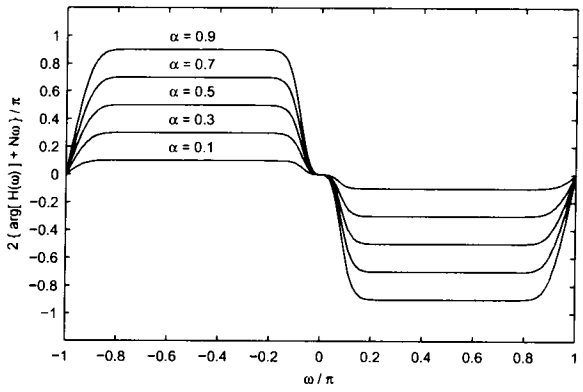


Fig. 3. The plot of normalized phase responses of 30th-order FHTs synthesized by the 30th-order MF AP HT.

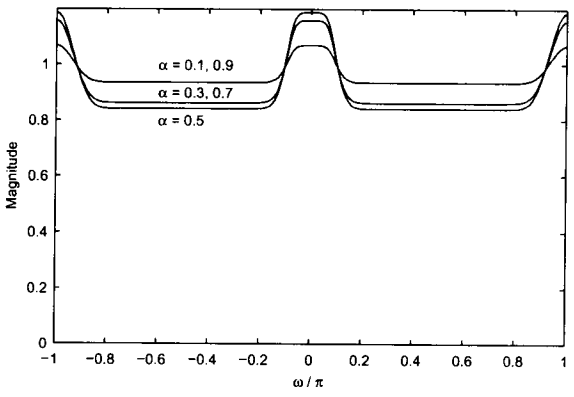


Fig. 4. The plot of magnitude responses of 30th-order FHTs synthesized by the 30th-order MF AP HT.