Authentication and Quality Monitoring Based on Audio Watermark for Analog AM Shortwave Broadcasting *

Ming Li, Yun Lei, Xiang Zhang, Jian Liu and Yonghong Yan
ThinkIT Lab, Institute of Acoustics
Chinese Academy of Sciences
{ming.li,ylei,xzhang,jliu,yyan}@hccl.ioa.ac.cn

Abstract

A method based on audio watermarking techniques for authentication and monitoring broadcasting quality of existing analog amplitude modulation (AM) shortwave radio is presented. The content and number of extracted messages can be useful to authenticate the received radio, trace the transmitting route, and represent the quality of broadcast transmitting environment, thus can be used to improve the quality of broadcast. In this paper, a robust self-synchronizing audio watermark algorithm for shortwave narrow band channel is proposed. Experimental results show that the method is robust against shortwave channel's noises and is effective to authenticate and monitor audio quality in realtime shortwave channel.

1 Introduction

Shortwave is used by radio services intended to be heard at great distances from the transmitting station. The long range of shortwave broadcasts comes at the expense of lower audio fidelity [2]. AM shortwave radio traveling by sky-wave has close relations with the ionosphere. Different traveling routes and times represent various conditions of the ionosphere, which cause diverse broadcasting qualities [7]. To improve the broadcasting qualities, the statistical relations between the broadcasting qualities and transmitting surroundings are usually used as a guide of allocating transmitters and frequencies [7]. Audio watermarking, as a powerful technique for copyright protection for audio works, has been widely exploited [3, 1, 6]. A digital watermarking is imperceptibly and robustly embedded into the host data such that it cannot be removed [3]. The audio watermark is imperceptible to humans and is embedded into the cover audio media. In this paper, a method, which is based on audio watermarking techniques using spread spectrum (SS) [4] to evaluate the broadcasting quality and track the traveling route, is presented. The information of the radio station, transmitter and aerial is repeatedly embedded into the radio-cast prior to transmission, and continually retrieved after reception. The extracted information is used to monitor the quality of reception and trace the origin of the signal.

This paper is organized as following. Shortwave channel is analyzed in section 2. In section 3 embedding strategy algorithms are explained. The decoding and synchronization strategies are proposed in section 4. Experimental results are given in section 5 and the concluding remarks are presented in section 6.

2 Channel Analysis

The shortwave channel in the 3-30M frequency band (HF) using the ionosphere as a transmission medium is a time-variant channel with frequency and time dispersive fading. The time-variant impulse responses of this channel is a consequence of the constantly changing physical characteristics of the transmission media. The time variations appear to be unpredictable to the user of the channel, therefore, it is reasonable to characterize the time-variant multipath channel statistically. Based on the analysis in [10], there are multiple propagation paths, and each path is associated with a propagation delay and an attenuation factor as a result of changes in the structure of the medium. Let $x(t)$ denotes the transmitted signal, then the received signal $y(t)$ is expressed in the integral form:

$$x(t) = \int_0^\infty x(t-\tau)h(t,\tau)d\tau \qquad (1)$$

$$H(f,t) = \int_{-\infty}^\infty h(t,\tau)e^{-j2\pi ft}d\tau \qquad (2)$$

where $h(t,\tau)$ denotes the attenuation of the signal components at delay τ and at time instant t and $H(f,t)$ is the corresponding transfer function. Thus, the amplitude variations in the received signal, termed signal fading, can be expressed as:

$$y(t) = \int_{-\infty}^\infty X(f)H(f,t)e^{j2\pi ft}df \qquad (3)$$

The audio watermarking method applied to AM channel, especially sky-wave (short-wave), faces many practical problems. First, in this type of channel, the radio signal is often severely interfered with noise. Second, the signal might be attenuated or enhanced randomly and capriciously. Third, the channel bandwidth

*This work is supported by Chinese 973 program (2004CB318106), National Natural Science Foundation of China (10574140, 60535030)

is narrow, only 4 kHz generally. And Fourth, received signal is always amplitude overflowed. Therefore, the covert messages, including two aspects: synchronization and watermarking message must be robust enough, and the embedded bandwidth should be narrower than 4 kHz. On the other hand, the watermarking method used here requires less transmitting capacity than other applications. Thus the main challenge is coming from time-variant frequency selective fading and severe additive noise.

3 Embedding Strategy

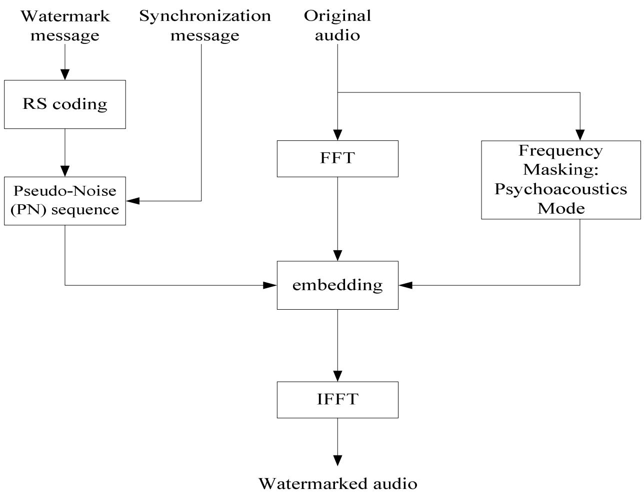

Figure 1. Overview of watermark embedding

Figure1 shows the overview of watermarking embedding strategy. Firstly, the N_{ori} bit watermarking message is encoded by Reed-Solomon (RS) code [12] which results in a N_w bit watermarking sequence M_w, secondly, the synchronization sequence M_s with the length of $N_s$is attached to the front of the watermarking message. Therefore, the length of the total embedding sequence M_e is extended to $N_e = N_w + N_s$ bits eventually. $x\{n\}$ is denoted as the original signal frame to be watermarked. It represents a block of samples from the analog broadcast. The frame $x\{n\}$ is filtered by hanning windows. $X\{f\}$ is the corresponding frequency magnitude components of $x\{n\}$. Each shift step of frames is the length of 1/4 frame to make the watermark more perceptually indistinguishable [15]. Each watermark bit $M_e(i)$ of a value $\{\pm 1\}$ from the sequence M_e is embedded into the audio signal. The corresponding watermarked frame is generated by (4):

$$\begin{aligned} Y\{f\} &= X\{f\} \cdot 10^{\alpha_j w\{f\} S_j M_e(i)} \qquad (4) \\ S_j &= (-1)^{\lfloor \frac{(N_{bp}/2)}{j} \rfloor} \end{aligned}$$

where S denote a bi-phase code $\{\pm 1\}$; α is the masking threshold based on the human auditory system (HAS) [1, 6] that guarantees the imperceptibility of the embedded watermark; $w\{n\}$ is Pseudo-Noise (PN) sequence generated by balanced Gold Sequences. A bit $M_e(i)$ is embedded in continuous N_{bp} frames as a whole. If $M_e(i) = 1$, then $S = 1$ in the front $N_{bp}/2$ frames and $S = -1$ in the next $N_{bp}/2$ frames. If $M_e(i) = -1$, then $S = -1$ in the front $N_{bp}/2$ frames and $S = 1$ in the next $N_{bp}/2$ frames. With no direct current component, the bi-phase code is beneficial to enhancing the robustness [4].

4 Decoding Strategy

4.1 Data Decoding

Let $C(R, P)$ denote the normalized correlation between real sequences R(k) and PN-sequence P(k):

$$\mathrm{C}(R, P) = \frac{\sum_{i=1}^{L} R(i) \cdot P(i)}{\left[\sum_{j=1}^{L} R(j)^2 \cdot \sum_{k=1}^{L} P(k)^2 \right]^{1/2}} \qquad (5)$$

where L is the length of the two sequences. Assuming that $Y_j(k)$ denotes the spectrum of the jth frame signal by sampling analog waveform on the receiver, the correlation C_j between $w\{k\}$ and $Y_j(k)$ is calculated according to:

$$\begin{aligned} C_j &= C(w, log_{10}(Y_j)) \\ &= C(w, log_{10}(X_j)) + \alpha_j \cdot S_j \cdot M_e(i) \cdot C(w, w) \\ &= \alpha_j \cdot M_e(i) \cdot S_j + C(w, log_{10}(X_j)) \qquad (6) \end{aligned}$$

As the second part should be near zero based on the property of PN sequences. The watermark bit $M_e(i)$ is determined by the correlations of the continual N_{bp} frames as well as frames in the embedding process according to:

$$I_i = \frac{1}{N_{bp}} \sum_{j=0}^{N_{bp}-1} C_j \cdot S_j \qquad (7)$$

If $I_i > 0$, the watermark bit is "1"; if $I_i < 0$, the watermark bit is "-1". Let x, y and $\overline{\alpha}$ denote $log_{10}(X)$, $log_{10}(Y)$ and $\frac{1}{N_{bp}} \sum_{j=0}^{N_{bp}-1} \alpha_j$ respectively, and assuming the elements x_i are independent identically distributed (i.i.d.) Gaussian random variables with standard deviation σ, i,e, $x_i \sim \mathcal{N}(0, \sigma_x)$. Thus, I_i can be rewritten as (8):

$$\begin{aligned} I_i &= \frac{1}{N_{bp}} \sum_{j=0}^{N_{bp}-1} (\alpha_j \cdot M_e(i) + S_j C(w, x)) \qquad (8) \\ &= \overline{\alpha} \cdot M_e(i) + \mathcal{N}\left(0, \frac{1}{\sqrt{L \cdot N_{bp}}}\right) \qquad (9) \end{aligned}$$

Therefore, in the proposed method the likelihood of a bit misdectection P_{MDB} once a watermark bit is detected equals (10) with $\tau = 0$:

$$\begin{aligned} P_{MDB} &= Pr[I_i < \tau] \qquad (10) \\ &= Pr[\mathcal{N}\left(0, \frac{1}{\sqrt{L \cdot N_{bp}}}\right) > \overline{\alpha} - \tau] \qquad (11) \\ &< \frac{1}{2} erfc\left(\frac{(\overline{\alpha} - \tau) \cdot \sqrt{N_{bp} \cdot L}}{\sqrt{2}} \right) \qquad (12) \end{aligned}$$

4.2 Synchronization

Up to now, several synchronization strategies are adopted to resist the synchronization attack in robust audio watermark algorithms, and the comparison of each strategy is performed in [13]. Kirovski *et al.*[4] utilize the combination of spread spectrum and spread spectrum code to effectively detect the HAS shaped watermark embedded into modulated complex lapped transform (MCLT) coefficients; and Wu *et al.*[14] adopt barker code as the synchronization code which is embedded into temporal domain with the watermark code together to achieve the ability of self-synchronization. In proposed application, both two strategies are used to achieve robust synchronization in realtime experiments. Due to the severe noise and distortion in the channel of AM shortwave broadcasting, only using the self-relative property of synchronization code has high false positive rate. In proposed method, if the extracted sequence equals to the synchronization code and the sum of their squared normalized correlations value with PN-sequence $\sum_{k=1}^{N_s} I_k^2$ is larger than a threshold I_0, synchronization finishes. Sliding correlation, which is computationally complex but robust way, is applied to search the threshold of synchronization sequence. Fortunately in synchronization the shift step is not necessary to be one point but the length of 1/2 frame or less [4]: Assuming that the length of a frame in the embedding process is L, such as 1024 points, thus, if only the bias between embedding region and detection region is less than $L/2$, the extracted bit, based on the sign of I, is the same as the case of no bias regardless of differences and overlaps among frames. In our application, the search shift step is $L/16$.

4.3 Modification in real experiments

In watermark embedding, the masking threshold based on the human auditory system (HAS) guarantees the imperceptibility of the embedded watermark. In the proposed algorithm in AM shortwave broadcasting channel, the number of sub-bands in HAS model is extended from 32 to 512, which results in higher masking threshold of psychoacoustic model without introducing audible distortion. To get the more reliable watermarking message, which is very important for monitoring, three consecutive watermarking messages are extracted after synchronization. The final output of extracted watermarking message is the voting result of above three watermarking messages and permutation of watermark bits is employed to resist the time-variant fading. In this way, the error bits caused by burst noise and attenuation could be corrected. The confidence of the output message is the number of the same bits in the three messages, only high confidence results are outputted. At last, the sequence is then decoded by RS code to finally generate the watermarking message.

5 Experimental Results

In our experiment, balance gold sequence is adopted as the PN sequence, N_{ori}, N_w and N_s is 24, 40, and 8 respectively. The sample rate is 32K, the length of one frame is 1024 points and the shift step in detection is 64 points, 1/16 of one frame. The threshold of synchronization I_0 is set between $[0.3, 1.3]$, which is adaptively modified by the output confidence value of the watermark decoder.

First we evaluate our algorithm in the test condition of DA/AD conversion and 32 kbps MP3 compression which is used to demonstrate the robustness against common manipulations. BAR, MAR, and WDR denote the percentage of bit accuracy rate, message byte accuracy rate and watermark detection rate respectively [9]. Second, the experiments in real time AM shortwave broadcasting channel are performed. The radio station A is chosen to transmit broadcast into which the watermarking message of 3 bytes length including the information of the radio station is embedded online. The broadcast is received in station A and station B which is about 2500 km away from station A respectively using the JRC NRD-545 DSP Receiver. The channel condition of received sinal in station A and station B is denoted as $A \rightarrow A$ and $A \rightarrow B$ respectively. The shortwave frequency in this experiment is 17880 khz and 11970 Khz respectively, and power of the transmitter is 100 kW. Real- time watermarking messages are continually extracted online.

Table 1. Experiments with different channels

channel condition	I_0	WDR (%)	BAR (%)	transmit frequency(Khz)
DA/AD	1	99.98	99.40	
32Kbps MP	0.8	99.88	99.42	
$A \rightarrow A$	1	68.02	99.61	17880
$A \rightarrow A$	0.9	76.74	99.49	17880
$A \rightarrow A$	0.8	81.98	99.63	17880
$A \rightarrow A$	0.7	87.02	99.34	17880
$A \rightarrow B$	0.9	13.37	97.41	17880
$A \rightarrow B$	0.8	25.00	97.15	17880
$A \rightarrow B$	0.7	42.64	96.93	17880
$A \rightarrow B$	0.9	17.64	98.26	11970
$A \rightarrow B$	0.8	25.39	97.93	11970
$A \rightarrow B$	0.7	34.69	97.09	11970

The experiments for each transmit frequency lasted for 45 hours which were spread into 15 days evenly. The broadcast on 8:00 - 11:00 a.m. of each day was used in the experiment and the content of the broadcast was mostly the news, music and talks.The experimental results are listed in table1 and table3. 45 hours are divided by the unit of half hour. Subjective evaluation of quality was conducted by 15 persons with the criteria listed in table2. Because level 5 is usually impossible in the AM shortwave broadcasting and level 1 did not appear in our experiment, only three levels are shown in table3.

For the aim of broadcast authentication and digital right protection in same frequency interference condition, the experiments of 2 station (A and B) transmitting different broadcast content ($C1$ and $C2$) with the same transmit frequency is performed. Table4 and Table5 show the experiments results of baseline system (A0), repeat coding (A1), data permutation (A2), and confidence judgement(A3) in this same frequency noise channel.

From table1, it can be seen that this algorithm is robust to AM shortwave channel. $A \rightarrow A$ performs better than $A \rightarrow B$ is due to no ionosphere interference, and the higher the threshold I_0, the higher the BAR with the cost of lower WDR. Table4 and Table5

Table 2. Characters of Subject Evaluation Scores

SEC	Characters
5	No interference
4	A little interference but the content is legible
3	Medium interference but the content could be comprehended
2	Severe interference and the great mass of the content could not be comprehended
1	Only the existence of the broadcasting could be judged

Table 3. quality monitoring experiments

SEC	WDR (%)	MAR (%)	Statistical Time
4	63.2	82	9 hours
3	49	75	25 hours
2	29.8	18	11 hours

demonstrate that the contents, which include the information of the radio station, transmitter and aerial, could reliably represent the origin of broadcast in the extracted messages even when two programs transmitted in the same frequency. The proposed modifications of our base algorithm further increase the accuracy of extracted message. Thus, the traveling route can be traced by the contents to monitor the quality of shortwave broadcast channel. Table3 shows that the system of evaluating the AM broadcast quality objectively could be established by the amount of extracted messages. The less the broadcast is interfered, the more the amount of extracted messages is, within each half hour.

6 Conclusion

This article presents a robust self-synchronizing audio watermark algorithm for shortwave narrow band channel. This algorithm has the following merits: Firstly, while the watermark information being hideaway, synchronized signal is inserted, which enables the watermark to have both the clock and data self-synchronizing abilities. Secondly the watermark signal can resist disturbances during shortwave transmission. Thirdly the watermark can be transmitted through narrow band channel, which is suitable for the shortwave broadcast. The experiments indicated that, this algorithm can be applied to monitor the AM broadcasting quality and authentication, and it showed satisfactory robustness for real-world application.

Table 4. Experiments for modification (Sec4.3) with same frequency interference

methods	A0	A0+A1	A0+A1+A2	A0+A1+A2+A3
BAR(%)	88.14	92.71	97.42	98.02
WDR(%)	56.78	50.19	50.19	40.89

Table 5. Experiments for same frequency interference

transmit power(Kw)		BAR (%) with A0		BAR (%) with A0+A1+A2+A3	
$C1$	$C2$	$C1$	$C2$	$C1$	$C2$
100	50	88.14	77.55	99.11	99.65
100	100	90.45	83.49	98.93	97.98
50	100	86.59	91.93	98.96	98.89
100	100	88.02	90.63	99.38	98.66

References

[1] L. Boney, A. Tewfik, and K. Hamdy. Digital watermarks for audio signals. *Proceedings of ICMCS*, Hiroshima, Japan 1996.

[2] K. Davies. *Nbs monograph 80*. U.S. Government Pringing Office, Washington. DC, 1965.

[3] F. Hartung and M. Kutter. Multimedia data-embedding and watermarking technologies. *Proceedings of the IEEE*, 87(7), January 1999.

[4] D. Kirovski and M. H. Spread-spectrum watermarking of audio signals. *IEEE Transactions on Signal Processing*, 51, January 2003.

[5] B. Ko, R. Nishimura, and Y. Suzuki. Time-spread echo method for digital audio watermarking using pn sequences. *ICASSP*, 2:2001–2004, 2002.

[6] G. L. Cano, P. Gmez, E. Bonnet, and M. Batlle. Audio watermarking and fingerprinting: For which applications? *Journal of New Music Research*, 32(1), 2003.

[7] G. Lane. Required signal-to-interference ratios for shortwave broadcasting. *Radio Science*, 32:2091–2098, January 1997.

[8] M. Li, Y. Lei, J. Liu, and Y. Yan. A novel audio watermarking in wavelet domain. *Proceedings of IIH-MSP*, California, USA 2006.

[9] H. Malik, A. Khokhar, and R. Ansari. robust audio watermarking using frequency selective spread spectrum theory. *Proceedings of ICASSP*, Montreal, Canada 2004.

[10] J. Proakis. *Digital Communications*. McGraw-Hill Book Company, 1983.

[11] S. Sedghi, M. Khademi, and N. Cvejic. Channel capacity analysis of spread spectrum audio watermarking for noisy environments. *Proceedings of IIH-MSP*, California, USA 2006.

[12] B. Stephen and K. Vijay. *Reed-solomon codes and their applications*. John Wiley Sons, Inc, 1999.

[13] X. Wang and H. Zhao. A novel synchronization invariant audio watermarking scheme based on dwt and dct. *IEEE Transactions on Signal Processing*, 54(12), January 2006.

[14] S. Wu, J. Huang, D. Huang, and Y. Shi. Efficiently self-synchronized audio watermarking for assured audio data transmission. *IEEE Transactions on broadcasting*, 51(1), 2005.

[15] H. Xing and M. Scordilis. Improved spread spectrum digital audio watermarking based on a modified perceptual entropy psychoacoustic model. *Proceedings of the IEEE*, 2005.