2008-01-01

Audio Data Verification and Authentication using Frequency Modulation Based Watermarking

Jonathan Blackledge
Technological University Dublin, jonathan.blackledge@tudublin.ie

Omar Farooq
College of Engineering and Technology, Uttar Pradesh, India

Recommended Citation

Blackledge, J., Farooq, O.: Audio Data Verification and Authentication using Frequency Modulation based Watermarking. International Society for Advanced Science and Technology, Journal of Electronics and Signal Processing. Vol. 3 (ISSN 1797-2329), issue: No 2, pages: 51 - 63, 2008. doi:10.21427/D7G90N

Audio Data Verification and Authentication using Frequency Modulation based Watermarking

Jonathan M Blackledge, Fellow, IET and Omar Farooq, Member, IEEE

***Abstract*— An approach to watermarking digital signals using frequency modulation - 'Chirp Coding' - is considered. The principles underlying this approach are based on the use of a matched filter to reconstruct a 'chirp stream' code that is uniquely robust. The method is generic in the sense that it can, in principle, be used for a variety of different signal (the authentication of speech and biomedical signals, for example). Further, by generating a bit stream that is signal dependent, chirp coding provides a method of self-authentication, which has a wide range of applications including copyright protection and digital rights management. However, in this paper, we focus on the application of chirp coding for the verification, authentication and self-authentication of audio signals. We also consider the effect of using a multi-level chirp coding approach to increase the 'volume' of data that can be embedded into a host signal.**

The theoretical and computational aspects of the matched filter with regard to the properties of a chirp are briefly revisited to provide the essential background to the method. Coding and decoding methods are then addressed and the results of different 'attack strategies' considered including Objective Difference Grades that are evaluated using Perceptual Evaluation of Audio Quality .

***Index Terms*— Audio Signal Processing, Frequency Modulation, Chirp Coding, Digital Watermarking, Authentication, Wavelet Transform.**

I. INTRODUCTION

WITH the increase in computing power and high bandwidth internet connectivity, copying, editing and the illegal distribution of audio data has become relatively easy and common place. This is a problem that is an issue in the distribution of most data and digital signals and is a major concern of the multimedia industry, in general. For this reason, the demand for copyright protection and tamper proofing of audio data has significantly increased in recent years and a range of digital audio watermarking methods have been proposed for applications such as copyright, annotation, authentication, broadcast monitoring, and tamper-proofing [1] - [4].

Adding special codes in the file header or embedding digital watermarks in the data can prevent/detect illegal copying or tampering. The former technique, however, has a limitation which is that once the header has been analysed, it can be easily removed, unlike a watermark that remains with the host data and can only be removed at the cost of substantial loss of information from the data.

An important goal in any watermarking method is to make the watermark imperceptible so that the end user(s) of the data is unaware of its presence. This is especially important when the audio data is music where degradation in quality cannot be tolerated. Most of the watermarking algorithms developed for this purpose take advantage of Human Auditory System (HAS) limitations to embed a perceptually transparent watermark into the host signal.

A wide range of time domain embedding techniques such as Least Significant Bit (LSB) alteration [5], Echo Addition [6], [7], Quantization Index Modulation (QIM) [8] and Spread Spectrum [9], [10] methods have been attempted together with transform domain techniques such as Fourier [11], Cepstral [12] and Wavelet-based ([13], [14] and [15]) approaches. Most of the algorithms developed fall into the category of robust watermarking due to their high tolerance toward different attacks. However, there are some applications where there is a need for checking the authenticity and originality of audio data. Such applications occur in areas such as broadcasting and sound recordings that may be used as evidence, e.g. police interviews, telephone intercepts and verbal agreements that are taken to be legally binding. Thus, watermarks for such applications must be fragile, i.e. the watermark should break if any tampering is undertaken on the watermarked signal.

It is desirable to have a watermark extraction process that is 'blind', implying that the original 'host signal' is not required to extract the watermark. To achieve this, a fixed watermark sequence can be embedded. However, this decreases the security of the scheme since a unique watermark sequence should ideally be used for a given signal. A solution to this problem is to use a signal dependent watermark code that can be both regenerated and reconstructed from the watermarked signal (implying that the sequence extracted from the original and watermarked signals are same).

High robustness to attacks along with high data rate watermark embedding cannot be achieved simultaneously [16]. There have been various attempts to increase the payload capacity for robust watermarking [17] and, more recently, multi-level watermarking in different domains has been proposed [18]. In this paper we propose a robust multi-level audio watermarking scheme based on embedding different frequency modulated 'chirp' signals by exploiting the unique property of a chirp in that it can be recovered in very low Signal-to-Noise Ratio (SNR) environments. Further, the HAS limitation of having poor sensitivity to frequencies below 100Hz can be exploited. Together with the robustness of the method,

Manuscript received October 1, 2008. This work is supported by the Science Foundation Ireland and by BSkyB.

Professor Jonathan Blackledge (jonathan.blackledge@dit.ie) is the Stokes Professor of Digital Signal Processing, School of Electrical Engineering Systems, Faculty of Engineering, Dublin Institute of Technology (http://eleceng.dit.ie/blackledge). Dr Omar Farooq is a Reader in the Department of Electronics Engineering, College of Engineering and Technology, Uttar Pradesh, India.

this approach has the capability of providing multi-level self-authentication as well and thus, provides a multi-level robust watermarking procedure with tamper detection capability.

Developing robust watermarking techniques for digital audio signals is relatively difficult compared to watermarking digital images. This is due to the high sensitivity of the human ear over a large dynamic range in comparison to the human eye. In general, the characteristics of the HAS can be utilised to achieve an imperceptible performance in audio watermarking [19]-[24]. Most of the audio watermarking techniques use either a frequency domain [25], [26] or time domain [27], [6] masking property to embed a watermark. In [28], for example, a spread spectrum technique is proposed that considers the host signal in terms of a communications channel through which the watermark is transmitted. In all of these techniques, the host signal is required to extract the watermark, which, in many applications, may not be feasible and/or desirable (depending on the security protocols associated with a watermarking system). Thus, a blind extraction technique is usually preferable. In this paper, following an approach originally developed by Blackledge [29] and developed further by Blackledge and Farooq [30]-[33], we present a method of self-authenticating an audio signal and study the robustness of the technique subject to a range of attacks.

The International Federation of the Phonographic Industry (IFPI) has set the following requirements on audio watermarking [19], [16] and [20]: (i) The watermark should be imperceptible; (ii) the embedding algorithm should give more than 20 dB SNR with a 20 bps data payload; (iii) the watermark should resist most common audio processing operations and attacks, such as digital-to-analogue and analogue-to-digital conversions, temporal scaling, additive/multiplicative noise corruption and MP3 compression; (iv) the watermark should be able to prevent unauthorized detection, removal and embedding, unless the quality of audio becomes very poor.

Most audio watermarking techniques focus on embedding watermarks for data verification. The development of fragile watermarking schemes has not been researched to the same extent. The main application of fragile watermarking is in the assessment of authenticity (tamper detection) of the data. In this paper we propose a frequency modulation (chirp coding) watermarking scheme which is not only robust but also serves for the purpose of tamper-proofing through self-authentication. The advantages associated with the low human ear sensitivity at low frequencies are taken into consideration to embed low frequency chirps within the host signal. The watermarking sequence is derived from the host signal's sub-band energies obtained by wavelet decomposition. The sub-band energies are not changed after watermarking thereby giving the method the capability of implementing a blind watermark detection process. For self-authentication the extracted watermark sequence is compared to the sequence derived from the sub-band energies associated with the watermarked signal.

Section II of this paper gives the basic background of the proposed scheme and Sections III revisits the use of the match filter for the detection of frequency modulated signals, i.e. the deconvolution of frequency modulated signals by correlation. The use of frequency modulation for watermarking is explained in Section IV and the generation of watermark sequences based on sub-band energies obtained by wavelet decomposition is discussed in Section V. Application of the method for self-authentication is considered in Section VI and the results of different attacks based on the Perceptual Evaluation of Audio Quality (PEAQ ITU-R recommendation BS.1387) [21] measurements are presented in Section VII. Finally, Section VIII investigates the use of multi-level watermarking with regard robustness and self-authentication.

II. Background to the Proposed Scheme

Methods of watermarking digital data have applications in a wide range of areas. Digital watermarking of images has been researched for many years in order to achieve methods which provide both anti-counterfeiting and authentication facilities. A principal equation that underpins this technology is based on the fundamental model for defining a signal and image, in general, and is given by (e.g. [34], [35])

$$s = \hat{P}f + n$$

where f is the information content for the signal, $\hat{P}$ is some linear operator, n is the noise and s is the output signal. This equation is usually taken to describe a stationary process in which the operator $\hat{P}$ is invariant of time and the noise n is characterized by stationary statistics (i.e. the probability density function of n is invariant of time). The most typical operation associated with this model is the convolution operation, i.e.

$$s(t) = (p \otimes f)(t) + n(t)$$

where $\otimes$ denotes the convolution operation, $p(t)$ is the Impulse Response Function and t denotes time.

In the field of cryptology, the operation $\hat{P}f$ is referred to as the processes of 'diffusion' and the process of adding noise is referred to as the process of 'confusion'. In Cryptography and Steganography (the process of hiding secret information in images [20]) the principal 'art' is to develop methods in which the processes of diffusion and confusion are maximized, an important criterion being that the output should be dominated by the noise n which in turn should be characterized by a maximum entropy criterion (i.e. a uniform statistical distribution).

Digital watermarking and Steganography can be considered to form part of the same field of study, namely, covert communications. Being able to recover f from s provides a way of reconstructing the information content of the signal. If we consider f to be information that constitutes a 'watermark', and n is a host signal (an audio signal, for example, which is the focus of the work reported in this paper) in which the watermark is embedded, then our problem is to recover the watermark from the host signal. If, in addition, it is possible to determine that a copy of s has been made leading to some form of signal degradation and/or corruption that can be conveyed through an appropriate analysis of f, then a scheme can be developed that provides a check on: (i) the authenticity of the signal s; (ii) its fidelity [22] and [23].

Formally, the recovery of f from s is based on the inverse process

$$f = \hat{P}^{-1}(s - n)$$

where $\hat{P}^{-1}$ is the inverse operator. Clearly, this requires the signal n to be known *a priori*. If this signal has been generated by a pseudo random number generator, for example, then the 'seed' (representing a 'key' in cryptography) used to generate this signal must be known *a priori* in order to recover the signal f. In this case, the seed represents the private key required to recover f. However, in principle, n can be any signal that is considered appropriate for confusing the information $\hat{P}f$ including a pre-selected signal or image. Further, if the process of confusion is undertaken in which the SNR is set to be very low (i.e. $\|n\| >> \|\hat{P}f\|$ where $\|\bullet\|$ defines a norm, e.g. a Euclidean norm), then the watermark f can be hidden covertly in the signal n provided the inverse process $\hat{P}^{-1}$ is well defined and computationally stable. In this case, it is clear that the host signal or image n must be known in order to recover the watermark f leading to a private watermarking scheme in which the signal n represents a key. This data can of course be (lossless) compressed and encrypted as required. In addition, the operator $\hat{P}$ (and its inverse $\hat{P}^{-1}$) can be key dependent. The value of operator key dependency relies on the mathematical properties of the operator that is available and whether it is compounded in an algorithm that is required to be in the public domain.

Another approach is to consider the case in which the signal n is unknown and to solve the problem of extracting the watermark f in the absence of this signal. In this case, the reconstruction is based on the result

$$f = \hat{P}^{-1}s + m$$

where

$$m = -\hat{P}^{-1}n$$

Now, if a process $\hat{P}$ is available in which $\|\hat{P}^{-1}s\| >> \|m\|$, then an approximate (noisy) reconstruction of f can be obtained in which the (processed) noise m is determined by the original SNR of the signal s and hence, the level of covertness of the diffused watermark $\hat{P}f$. In this case, it may be possible to post-process the reconstruction (de-noising, for example) and recover a relatively high-fidelity version of the watermark, i.e.

$$f \sim \hat{P}^{-1}s.$$

This approach (if available) does not rely on a private key (assuming $\hat{P}$ is not key dependent). The ability to recover the watermark only requires knowledge of the operator $\hat{P}$ (and its inverse) and post-processing options as required. The problem here is to find an operator that is able to recover the watermark effectively in the presence of the signal *n*. Ideally, we require an operator $\hat{P}$ with properties such that $\|\hat{P}^{-1}n\| \to 0$.

In this paper, the operator is based on a chirp function. A 'chirp' is a signal which has a frequency sweep which increases or decreases with time. This increase/decrease in frequency may be linear, quadratic or logarithmic as illustrated by the spectrograms (time-frequency maps) given in Figure 1. One of the most common chirp signals is a linear

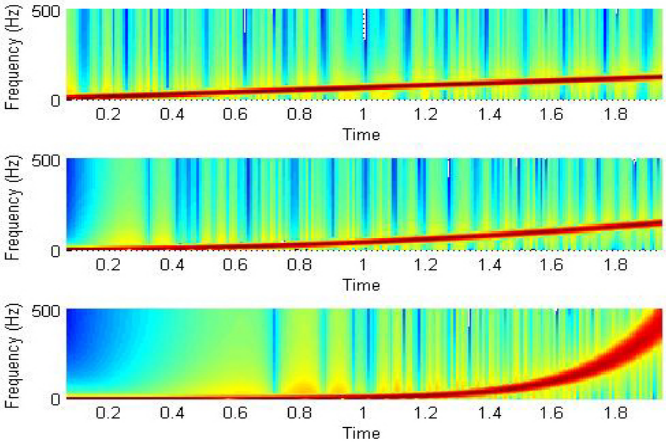

Fig. 1. Spectrograms (time-frequency maps) for three different types of chirp signal: linear (top), quadratic (centre) and logarithmic (bottom).

Frequency Modulated (FM) chirp which is of (complex) type $\exp\left(-i\alpha t^2\right)$ where α is the 'chirp parameter'. If this function is convolved with f, then the inverse process can be simply undertaken by correlating with the (complex) conjugate of the chirp $\exp(i\alpha t^2)$. This provides a reconstruction for f in the presence of noise n that is accurate and robust with very low SNRs [29]. If we consider a watermark based on some coding method in which the noise *n* is the host signal, then the watermark f becomes n-dependent. This allows an authentication scheme to be developed in which the watermark is generated from the signal in which it is to be embedded. Authentication of the watermarked data is then based on comparing the code generated from $s = \hat{P}f + n$ and that reconstructed from s when $\|\hat{P}f\| << \|n\|$. This is an example of a self-generated coding scheme which avoids the use, distribution and application of reference codes. It is the foundation for the method we call self-authentication described in Section VI in which code generation (in this case, the conversion of the host signal to a bit stream) is based on the application of a wavelet transform.

In practice, the diffusion of a bit stream (represented by the function f) by convolution with a chirp function, does not provide a reconstruction (through correlation with the complex conjugate of the chirp function) with an accuracy and fidelity that is viable. To overcome this, we consider a 'chirp stream' as illustrated in Figure 2. Here, the operation $\hat{P}f$ is replaced with a chirp stream which, with regard to Figure 2, can be written as

$$\hat{P}f = \cos(\alpha t^2) \otimes [-\delta(t) + \delta(t+T) + \delta(t+2T) - \delta(t+3T)$$
$$-\delta(t-4T) + \delta(t+5T) - \delta(t-6T)]$$

where $\cos(\alpha t^2) \quad \exists \quad \forall t \in [0, T]$. This solution provides a more accurate reconstruction but at the expense of increasing the size (the product of the chirp length with the total number of bits used to represent the watermark code) of the host signal required for watermarking.

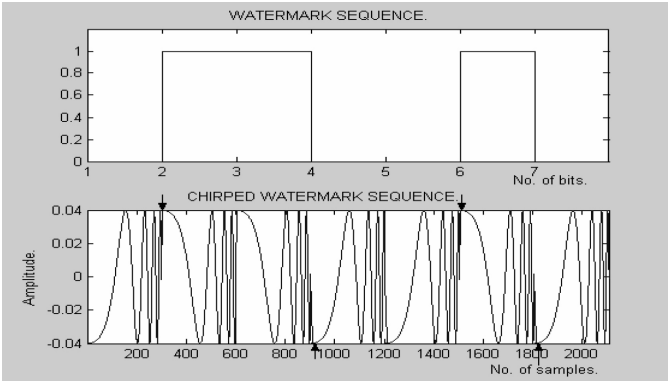

Fig. 2. Watermark sequence (0110010) and the corresponding chirp stream.

There are numerous applications of this technique in areas such as telecommunications and speech recognition where authentication is mandatory. The method can readily be applied to audio data with no detectable differences in the audio quality of the data. The watermark code is able to be recovered accurately and can be 'engineered' to be very robust. On the other hand, the method can be 'calibrated' to output relatively significantly changes if the data is distorted through cropping, filtering, noise or a compression system, for example, thus, providing a way of making a signal tamper proof.

III. Deconvolution of Frequency Modulated Signals

The matched filter is frequently used in systems that utilize linear frequency modulated (FM) pulses known as chirps or chirped pulses [36], [37] and [38]. Examples of where this particular type of pulse is used include real and synthetic aperture radar [39], [40], active sonar and some forms of seismic prospecting, for example. Interestingly, some mammals (including dolphins, whales and bats) use frequency modulation for communication and detection. The reason for this is the unique properties that chirps offer in terms of the quality of extracting information from signals with very low SNRs and the simplicity of the process that is required to do this (i.e. correlation) as discussed below.

The liner FM pulse is given (in complex form for unit amplitude) by

$$p\left(t\right)=\exp\left(-i\alpha t^{2}\right),\ \ |t|\leq\frac{T}{2}$$

where α is a constant and *T* is the length of the pulse. The phase of this pulse is αt^2 and the instantaneous frequency is given by

$$\frac{d}{dt}\left(\alpha t^{2}\right)=2\alpha t$$

which varies linearly with *t*. Hence, the frequency modulation is linear which is why the pulse is referred to as a linear FM pulse. If $p(t)$ is taken to be the Impulse Response Function, then the signal $s(t)$ that is recorded is given by (neglecting additive noise).

$$s\left(t\right)=\exp\left(-i\alpha t^{2}\right)\otimes f\left(t\right),\ \ |t|\leq\frac{T}{2}.$$

Matched filtering (i.e. correlating the signal $s(t)$ with the complex conjugate of $p(t)$ where the correlation operation is denoted by $\odot$) we have

$$\hat{f}(t)=\exp(i\alpha t^{2})\odot\exp\left(-i\alpha t^{2}\right)\otimes f(t)$$

$$=T\exp(i\alpha t^{2})\mathrm{sinc}\left(\alpha Tt\right)\otimes f\left(t\right),\quad|\,t\,|\leq\frac{T}{2}.$$

If we now consider the length of the linear FM pulse to be relatively long (i.e. $T>>$), then

$$\cos\left(\alpha t^{2}\right)\mathrm{sinc}\left(\alpha Tt\right)\sim\mathrm{sinc}\left(\alpha Tt\right)$$

and

$$\sin\left(\alpha t^{2}\right)\mathrm{sinc}\left(\alpha Tt\right)\sim0$$

so that

$$\hat{f}\left(t\right)\sim T\mathrm{sinc}\left(\alpha Tt\right)\otimes f\left(t\right).$$

In Fourier space, this last equation can be written as (ignoring scaling)

$$\hat{F}(\omega)=\begin{cases}F(\omega), & |\,\omega\,|\leq\alpha T;\\ 0, & |\,\omega\,|>\alpha T.\end{cases}$$

The estimate $\hat{f}$ is therefore a band limited estimate of f whose bandwidth is determined by the product of the chirping parameter α with the length of the pulse *T*.

Given that

$$s\left(t\right)=\exp\left(-i\alpha t^{2}\right)\otimes f\left(t\right)+n\left(t\right)$$

after match filtering we obtain the estimate

$$\hat{f}\left(t\right)\cong T\mathrm{sinc}\left(\alpha Tt\right)\otimes f\left(t\right)+\exp\left(i\alpha t^{2}\right)\odot n\left(t\right).$$

The correlation signal produced by correlating $\exp\left(i\alpha t^{2}\right)$ with $n\left(t\right)$ will, in general, be relatively low in amplitude under the assumption that $n\left(t\right)$ will not normally have features that correlate or 'match' those of a chirp. It is therefore reasonable to assume that

$$\|T\mathrm{sinc}\left(\alpha Tt\right)\otimes f\left(t\right)\|>>\|\exp\left(i\alpha t^{2}\right)\odot n\left(t\right)\|$$

and that, in practice, $\hat{f}$ is a band-limited reconstruction of f with high SNR. Thus, using chirps with matched filtering for the purpose of reconstructing an input in the presence of additive noise provides a relatively simple and computationally reliable method of 'diffusing' and reconstructing information encoded in the input function f. The ability for the matched filter to accurately recover information from linear FM type signals with very low SNRs leads naturally to consider its use for covert information embedding. This is the basis for the chirp coding method discussed in this paper - covertly watermarking digital signals for the purpose of signal authentication.

IV. Chirp Coding, Decoding and Watermarking

We now return to the issue of watermarking using chirp coding. The scheme of embedding and detecting a watermark based on frequency modulation is illustrated in Figure 3 and Figure 4.

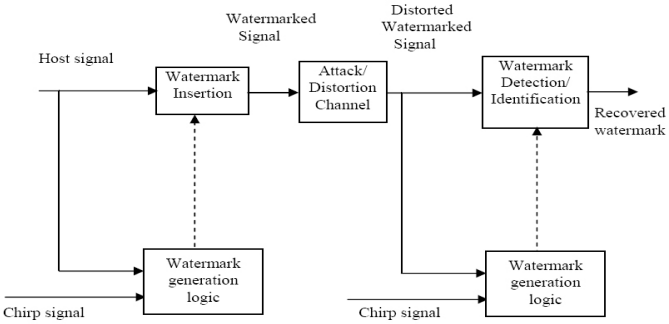

Fig. 3. Basic scheme for embedding and detecting a chirp coded watermark.

The model for the watermarked signal (which is real) is

$$s\left(t\right)=\text{chirp}\left(t\right)\otimes f\left(t\right)+n\left(t\right)$$

where[1]

$$\text{chirp}\left(t\right)=\cos\left(\alpha t^{2}\right),\quad \mid t\mid\leq\frac{T}{2}.$$

The function $f(t)$ is taken to be a sum of delta functions

$$f(t)=\sum_{k}b\delta(t+kT)$$

whose polarity $b=\pm1$ is used to differentiate between 0 and 1[2]

For the purpose of authenticating the signal $n\left(t\right)$, two basic criterion must be satisfied: (i) $f\left(t\right)$ must be a code that can be reconstructed accurately and robustly; (ii) the watermarking code should be very sensitive to any degradation in the signal $n\left(t\right)$. The degradation in the signal may be due to lossy compression, filtering operations, re-sampling etc. To satisfy the first condition it is reasonable to consider $f\left(t\right)$ to represent a bit stream, i.e. to consider a digitised version of $f\left(t\right)$ - the vector f_i - to be composed of a set of elements with values *0* or *1*. This binary code can be generated by using a key or a set of keys, which, when reconstructed, is compared to the key(s) for the purpose of authentication of the data. However, this requires the distribution of the keys (public and/or private). The code can also represent plaintext converted into binary form which is the basis for the case study considered in Section IX. However, we can also consider a scheme that involves the generation of a binary sequence using the spectral characteristic of the signal $n\left(t\right)$ itself. Once the binary sequence is generated, chirp coding can be applied as discussed below.

[1]The use of a sine function is equally valid.

[2]For a digital signal, as illustrated in Figure 2, for example, these delta functions are taken to be Kronecker delta functions - discrete form delta functions - of unit amplitude.

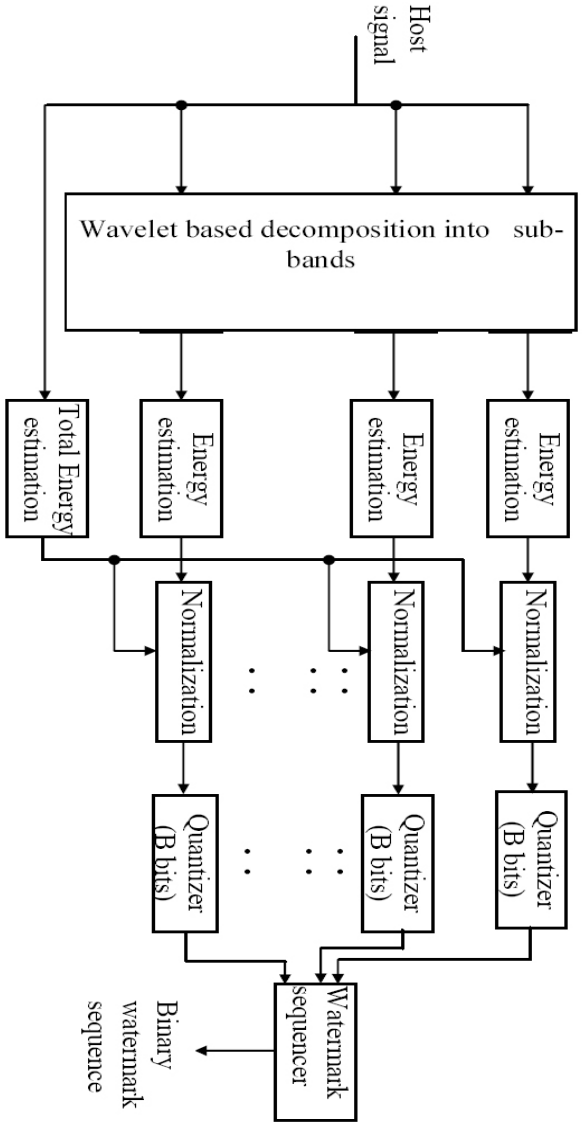

Fig. 4. Schematic logic for the generation of a watermark sequence based on the sub-band energies of the input signal.

A. Chirp coding

The purpose of this coding is to diffuse each bit over a range of compact support. In order to differentiate between *0* and *1* the polarity of the chirp is reversed for *0*. Thus, a binary sequence *101*, for example will be transformed into the signal $f(t)$ given by

$$f(t)=\begin{cases}+\text{chirp}(t), & t\in[0,T);\\ -\text{chirp}(t), & t\in[T,2T);\\ +\text{chirp}(t), & t\in[2T,3T).\end{cases}$$

where *T* is the period of chirp. The period over which the chirp is applied depends upon the length of the host signal and the length of the binary sequence. In order to avoid aliasing and interference with the high frequency components in the chirp stream, a logarithmic chirp is applied. The instantaneous frequency ω of a logarithmic chirp is given by

$$\omega(t) = \omega_0 + 10^{\beta t}$$

where

$$\beta = \log_{10}(\omega_1 - \omega_0)/t_1$$

Here, ω_0 is the initial frequency and ω_1 is the final frequency at time t_1, the final frequency being greater than the initial frequency. These parameters are used to generate a low frequency watermark that is below the frequency range of the HAS.

B. Decoding

Decoding or recovery of the binary sequence is based on correlating the watermarked signal with the original chirp function. This produces a correlation function that either *+1* or *-1* depending upon the original polarity of the chirp. If the value is positive then the sequence bit is assumed to be *1* and if a negative value is obtained then the bit is assumed to be *0*). In practice, however, the correlation function may not be exactly *1* or *-1* when reconstruction is undertaken and the binary sequence is effectively recovered by searching the correlation function for changes in sign. The chirp used to recover the watermark must have the same parameters (including the chirp length) as those used to generate the chirp stream. These parameters can be used to define part of a private key.

C. Watermarking

The watermarking process is based on adding the chirp stream f to the host signal n. However, we require that the magnitude of *f* is significantly less than *n* in order that *n* is only slightly perturbed by the chirp stream. We therefore consider the watermarking process to be described by the 'normalisation equation'

$$s(t) = \frac{a}{b}\left[\frac{cf(t)}{\|f(t)\|_\infty} + \frac{n(t)}{\|n(t)\|_\infty}\right]$$

where the coefficient $0 < c < 1$ determines the the SNR (the chirp stream-to-host signal ratio) of *s* respectively and where

$$a = \|n(t)\|_\infty$$

and

$$b = \|[cf(t)/\|f(t)\|_\infty] + n(t)/\|n(t)\|_{infty}]\|_\infty.$$

The coefficient *a* is required to provide a watermarked signal whose amplitude is compatible with the original signal *n*. The value of *c* is adjusted to provide an output that is acceptable in the application to be considered and to provide a robust reconstruction of the binary sequence by correlating $s(t)$ with $\mathrm{chirp}(\mathrm{t}), \mathrm{t} \in [0, T)$. To improve the robustness of the reconstruction, the value of *c* can be increased, but this has to be off-set with regard to the perceptual quality of the output and/or the tamper proofing required of the watermark.

D. Wavelet Decomposition

The wavelet transform is defined by (e.g. [41]-[44])

$$\hat{W}[f(t)] = F_L(t) = \int f(\tau)\, w_L(t,\tau)\, d\tau$$

where the wavelet function w_L is given by

$$w_L(t,\tau) = \frac{1}{\sqrt{|L|}} w\left(\frac{t-\tau}{L}\right).$$

This transform is essentially a convolution transform in which $w(t)$ is the convolution kernel but with a factor *L* introduced. The introduction of this factor introduces dilation and translation properties into the convolution integral (which is now a function of *L*) which gives it the ability to analyze signals in a multi-resolution role.

We consider a code generating method that is based on computing the energies of the wavelet transformation over *S* levels. Thus, the signal $f(t)$ is decomposed into wavelet space to yield the following set of functions:

$$F_{L_1}(t), F_{L_2}(t), ..., F_{L_S}(t).$$

The (percentage) energies of these functions are then computed as defined by

$$E_i = \frac{100}{E}\int |F_{L_i}(t)|^2\, dt$$

where

$$E = \sum_{i=1}^{S} E_i.$$

Concatenating the binary sequence representation of E_i generates the watermarking code. The watermark is then chirp coded as discussed in Section IV(A) with the computations being undertaken in digital form using a DWT (Discrete Wavelet Transform). In the study undertaken for this paper, the wavelets used are Daubechies wavelets.

E. Self-Authentication

By generating a bit stream that is signal dependent we can self-authenticate the signal. The reason for this is that the chirp stream can be added to the original signal with such a high signal-to-chirp stream ratio, that the computation of the energy values for the watermarked signal remains unaffected. Thus, once the signal has been watermarked with the chirp stream, we can apply two separate processes to it: (i) computation of the bit stream from the (watermarked) host signal; (ii) correlation of the watermarked signal with the appropriate chirp to recover the same watermark (bit stream). When both bit streams are found to be the same, the signal can be taken to be authentic.

Only a specified segment of the data need be extracted for watermarking which is equivalent to applying an off-set to the data. The segment can be user defined and if required, form the basis for a (private) key system. In principle, different wavelets can be used for the process of wavelet decomposition provide the decomposition provide results that are robust with regard to a specific criterion. The actual wavelet used then

provides another component that can form part of the private key needed to extract the watermark. The flow chart of the proposed watermarking scheme is shown in Figure 5 where the off-set is assumed to be zero.

The approach considered allows a code to be generated directly from the input signal and that same code used to watermark the signal. The code used to watermark the signal is therefore self-generating. Reconstruction of the code only requires a correlation process with the watermarked signal to be undertaken. In other words, the method can be seen as a way of authenticating data by extracting a code (the watermark) within a code (the signal) and is consistent with approaches that attempt to reconstruct information without the host data [45].

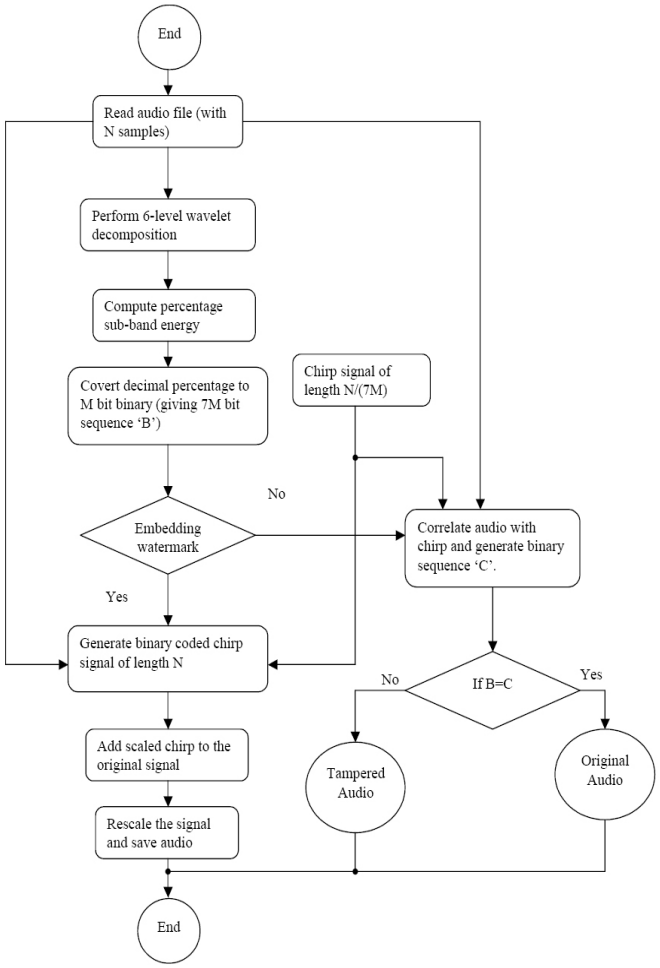

Fig. 5. Flow chart of the proposed watermarking scheme where the off-set is assumed to be zero.

V. Results of Attacks on Watermarked Audio Signals

Different audio files with sampling frequencies in the range of 11025Hz to 48kHz were chosen and a (logarithmic) chirp generated in the frequency band of 1-100Hz for watermarking. Since the human ear has low sensitivity in this low frequency band, the embedded watermark is not perceptible. A six-level wavelet decomposition using a 'Daubechies 6' wavelet was applied on the audio signal giving seven sub-bands. The energy in these sub-bands was normalized by the total energy of the signal to give percentage energies in each sub-band as discussed in Section IV(D). A total of 14 bits was used for quantization of the percentage energy features, thereby giving a total of 98 bits to be embedded into the audio signal. When the watermarking method is aimed at tamper proofing of the audio data, it is desirable to spread the watermark in which one bit is embedded per second. Depending upon the band and amplitude of the chirp, a Signal-to-Watermark Ratio (SWR) in excess of 40dB can be achieved. Figure 6 shows a segment of the original and the watermarked signal with SWR of 43.4dB (calculated for the entire duration of the original signal). For tamper proofing, the duration of the chirp can be increased to cover the entire music signal.

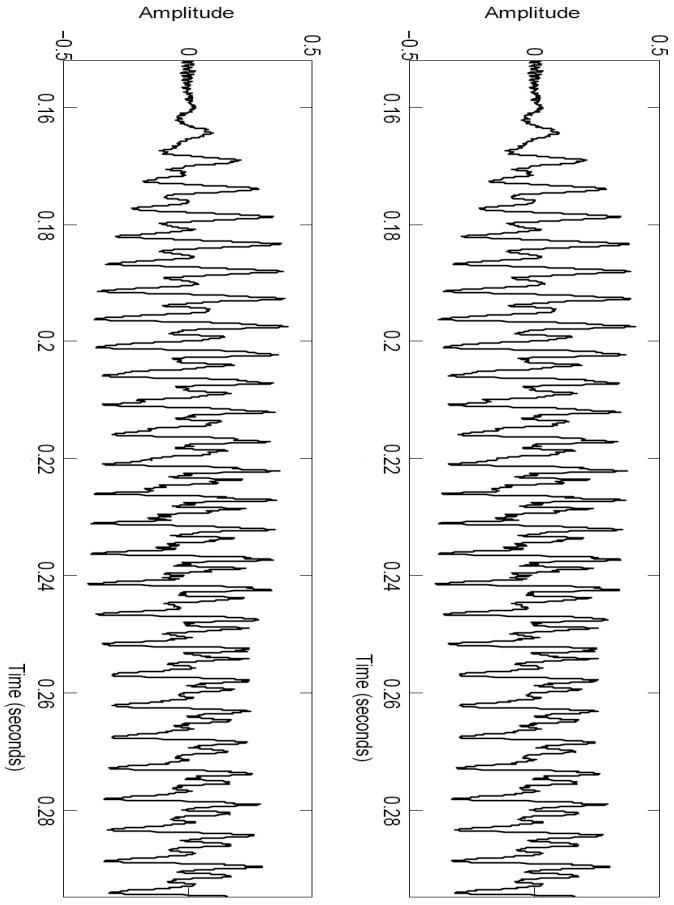

Fig. 6. Segment of an original audio signal (left) and a chirp-code watermarked signal for tamper proofing (right).

In order to detect the effectiveness of the proposed fragile watermarking scheme, various attacks were simulated and the watermark recovered. Since an attack changes the distribution of percentage sub-band energies that are originally present in the signal, the recovered watermark does not match with the percentage sub-band energies of the tampered signal. This results in the ability to detect any tampering in the signal. The robustness of the chirp-based watermark was also studied for its application to robust audio watermarking. This was

evaluated by comparing the original watermark bit sequence to that extracted from a tampered (attacked) audio signal. Since the most common attacks are based on filtering, compression, cropping and noise addition to the watermarked signal, these attacks were simulated.

A. Additive Noise

Different magnitudes of uncorrelated additive white Gaussian noise were injected into the watermarked signal giving various watermarked SNRs. The watermarking scheme was able to detect tampering (i.e. the addition of noise) even when the noise was 60dB below the watermarked signal. This attack was easily detected because additive noise changes the percentage sub-band energies (even with 'white noise' characterised by a flat power spectral density function). If the noise has uniform density, addition by a constant will change the percentage energy distribution feature vector causing it to be easily detected.

For evaluating robustness of the proposed scheme, the SNR was varied from 50dB to 5dB and the watermark extracted. The original watermark was detected without any error up to a SNR of 20dB. However, as the noise power was increased, error in the detection of the watermark was observed. This error is plotted in Figure 7 where it can be seen that for a SNR up to 10dB, the total error in detecting the watermark is less than 5%.

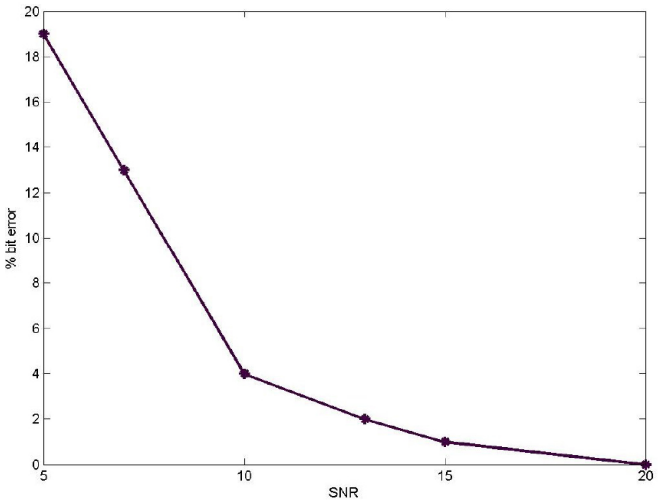

Fig. 7. Plot of percentage bit error detected in the chirp based watermarking scheme for different signal- to-noise ratios (SNR).

B. MPEG 1 Layer 3 (MP3) Compression

Watermarking was carried out on a wave file format with only one channel being watermarked. Experiments were conducted for a wide range of Constant Bit Rate (CBR) and Variable Bit Rate (VBR) MP3 compression (in the range of 128-320kbps). In all cases, tampering was detected by the proposed scheme. For the range of bit rates obtained by MPEG 1 Layer 3 compression, all the watermark bits were correctly recovered, showing the robustness of the technique toward compression attack.

C. Cropping

Arbitrary samples of the watermarked signal were removed and watermark detection applied. Since cropping reduces the energy in some of the sub-bands, tampering was easily detected when the signal was correlated with the matching chirp. However, depending upon the length of the crop and the portion of the signal over which the crop was taken, different results were observed for robust watermarking. If a crop is taken from the end of the audio file, then the watermark prior to this point is detectable. However, if a crop is taken over portions of the audio file toward the start or over the central portion the watermarked signal, then the watermark is not detectable after and beyond the cropped segment of the signal. This is due to the fact that the offset parameter changes after cropping and hence, the correlator cannot detect the watermark.

D. Filtering

Lowpass and highpass filtering was applied to the watermarked signal and the filtered signal correlated with the appropriate chirp. Tampering was easily detected in all cases. In the case of highpass filtering, the watermark is itself removed. Thus, correlating filtered signals with a chirp yields percentage sub-band energies, which do not match with the watermarked signal. Figure 8 shows the power spectral densities of the original, watermarked and the attacked signal using a bandpass filtering attack. The bandpass of the filter was chosen to be 0.01B to 0.99B where B was the signal bandwidth. Although there is a very small difference between the watermarked and the attacked signal, it is still easily detected by the proposed scheme.

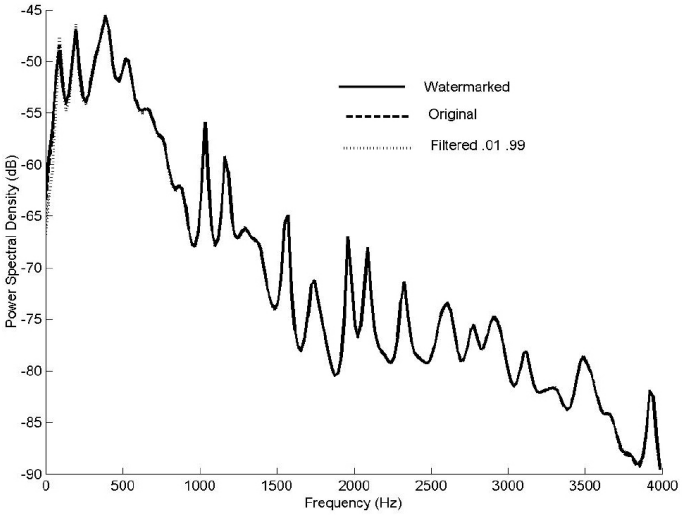

Fig. 8. Difference in the power spectral density of the original, watermarked and tampered signal. The tampering attack is done by using a bandpass filter with a normalised lower cut-off frequency of 0.01 and higher cut-off frequency 0.99.

For a lowpass filtered watermark signal, the detected bit error was less than 1% for a normalised higher cut-off frequency range of 0.01 to 0.99. This shows the effectiveness of the chirp based watermarking scheme. Even if some of the

frequency components of the chirp signal are not present in the watermarked signal, the watermark can still be detected. However, when using a highpass filter with different cutoff frequencies, errors are obtained in the detected watermark as shown in Figure 8. The reason for this error, is that during filtering, some portions of watermark may be removed while the signal remains, thereby resulting in erroneous detection. This problem is absent when lowpass filtering is applied because it removes the entire audio signal itself along with some portions of the watermark. Thus the lower frequency watermark that is left provides a matching 'signature' when correlated with the original chirp.

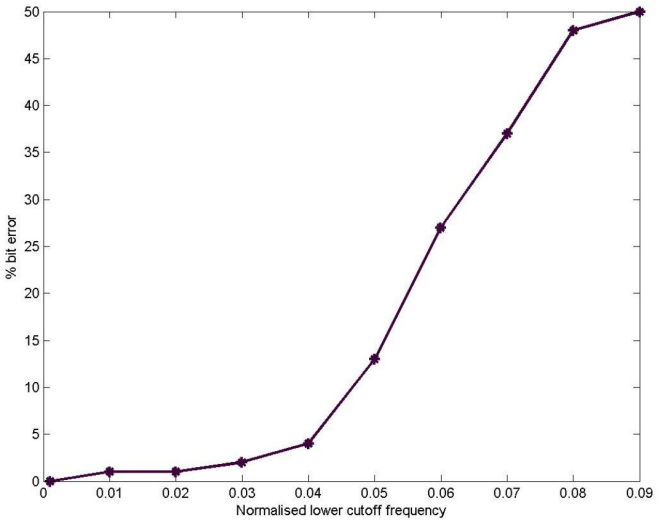

Fig. 9. Percentage bit error detected in the watermark for different lower cut-off frequencies of a highpass filter.

It is equally important for a watermarking scheme to be perceptually transparent along with its robustness and self-authentication capability. To evaluate this, the watermarking scheme was applied to different audio files taken from the Sound Quality Assessments Material (SQAM). These files included speech as well as music files. The subjective evaluation of the watermarked signal showed no perceptual difference from the original audio signal.

A Perceptual Evaluation of Audio Quality (PEAQ ITU-R recommendation BS.1387) for the coders has been proposed [21]. This evaluates a measure called Objective Difference Grade (ODG) on a scale of -4 to 0 where 0 corresponds to no perceptual difference. The evaluation can be carried out using two different models; a basic version and an advance version. In the basic version of PEAQ (Perceptual Assurance of Audio Quality), there are 11 model output variables that are combined with a neural network with 11 input nodes, 1 hidden layer with 3 nodes and a single output [21]. The average ODG using the basic version of PEAQ was evaluated and found to be -0.2567 for the speech, -1.318 for music, while the overall average was -0.721 taking all 16 SQAM (Speech Quality Assessment Material) audio. This shows that the watermarked speech has imperceptible difference as compared to the original speech. However it is perceptible for the musical recording used in SQAM. The main reason for this perception is due to the presence of large period of silence in the signal during which the SWR is low.

TABLE I
SNR AND ODG ACHIEVED FOR DIFFERENT EMBEDDING LEVELS.

Embedding Level	Chirp frequency range (Hz)	SNR (dB)	ODG
1	0-15	30.9050	-0.46556
2	0-30	28.1748	-0.77894
3	0-45	27.3172	-0.88256
4	0-60	26.7349	-0.95156

VI. MULTI-LEVEL WATERMARKING

To verify the multi-level watermarking scheme, audio files were selected from the Speech Quality Assessment Material (SQAM) which is usually used for assessing the quality of speech coders. A logarithmic chirp signal with a frequency sweep less than 100Hz was used to generate the chirp stream. This frequency range was chosen to keep the watermarks imperceptible, wavelet decomposition being carried out using a 'Daubechies 4' wavelet. The resultant sub-bands obtained for a 48kHz sampled signal are 0-3kHz, 3-6kHz, 6-9kHz, 9-12kHz, 12-15kHz, 15-18kHz and 18-24kHz. A binary watermark sequence was derived by representing the percentage sub-band energies using 12 bits and the total energy by 32 bits, thus giving a total of 116 bit of information. The coded chirp was scaled and embedded into the signal. The first level of watermark was embedded by using a logarithmic chirp with a frequency sweep of 0-15Hz and a total of four levels of embedding carried out with the chirp specification as given in Table I.

In order to analyze the worst case performance, the same sub-band based watermark sequence was embedded at all four levels. Further, the length of the chirp and the starting point for watermark embedding was kept the same at all levels. This generated overlapping chirp frequency ranges over a common time frame, thereby giving a maximum possibility for error during the detection process. Subjective assessments of the speech quality was carried out by calculating the SNR at each embedding level. The average SNR was evaluated for the different levels, the results obtained being given in Table I. This SNR can be improved if the length of the chirp is increased. However, it is important to take into consideration the human hearing curve to interpret the SNR achieved. Since the human ear has very poor sensitivity below 100 Hz, a lower SNR is still imperceptible upon listening.

To further verify the above results, tests based on Perceptual Assessment of Audio Quality (PEAQ, Basic) [21] were also carried out. The PEAQ algorithm is the ITUR recommendation (ITU-R BS.1387) for perceptual evaluation of wide-band audio codecs. This algorithm models fundamental properties of the auditory system along with physiological and psychoacoustic effects. It uses an original and test signal, and applies techniques to find difference between them. An Objective Difference Grade (ODG) is evaluated using a total of eleven Model Output Variables (MOV) of the basic version of PEAQ. The original signal and watermarked signal after different embedding levels was used to evaluate the ODG. The

results obtained are shown in Table I where the ODG values mimic the listening test ratings with values between -4.0 (very annoying) to 0 (imperceptible difference).

Although all the MOVs were calculated from the PEAQ (basic version), only those relevant to watermarking are reported here. The Noise-to-Mask Ratio (NMR) is an estimate in dB of the ratio between the actual distortion (caused due to the embedding watermark in this case) and the maximum inaudible distortion. The total NMR is the average of the NMR calculated over all the frames. Negative NMR values indicate inaudibility whereas values larger than 0dB indicate audible distortions caused by the watermark. This is an important test for checking the inaudibility of the embedded watermark at different levels. The result for speech signals and music signals were analysed separately and are plotted in Figure 10 along with the overall average.

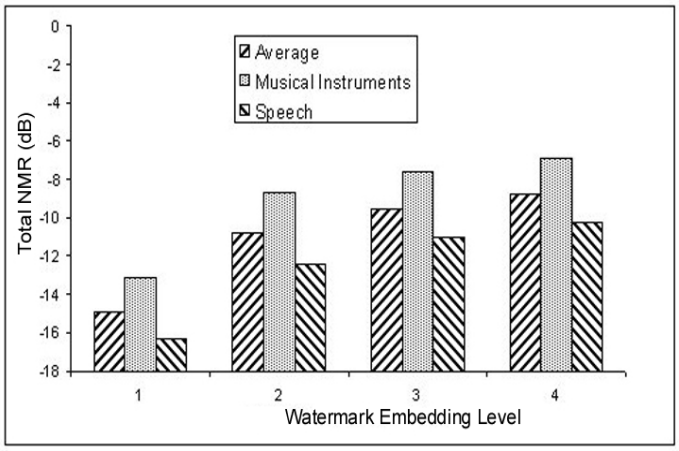

Fig. 10. Plot of total Noise-to-Mask Ratio (NMR) for different levels of watermark embedding.

As stated earlier, although the SNR values (as shown in Table I) indicate a high watermark level, it is still not audible because of the low chirp frequencies used. The noise loudness quantifies the partial loudness of distortions that is introduced during injection of watermark into the host signal. The root mean square value of noise loudness has a maximum limit of 14.8197. Figure 11 shows a normalised plot of the average RMS (Root Mean Square) noise loudness achieved on the SQAM data for all of the 4 levels of watermark embedding used.

It is possible for the total NMR to be below 0dB (implying inaudibility), but there may be a large number of frame with small positive values and a few frames with large negative values. This distribution can be observed by evaluating the number of disturbed frames. A relatively disturbed frame is one in which the maximum NMR exceeds 1.5 dB and is expresses a fraction of the total number of frames. The results in Figure 11 show that at four levels of watermark embedding, less than 3.5% of the frames have a NMR above 1.5dB. Thus the multi-level watermarking proposed is imperceptible. To evaluate the robustness and self-authentication capability of this multi-level approach, different attacks were simulated as discussed below.

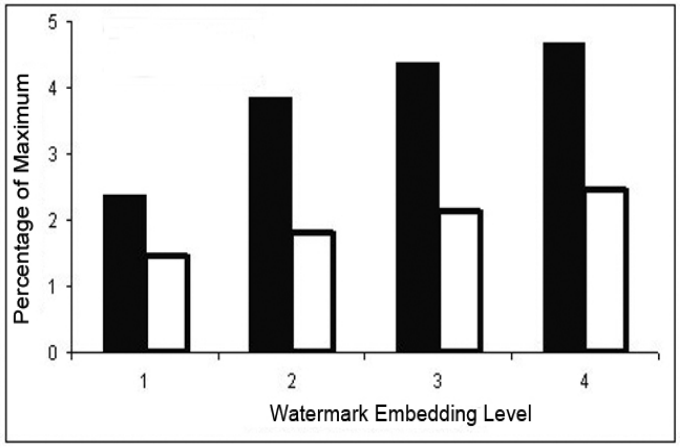

Fig. 11. Plot of the normalized values of the root mean square noise loudness (solid bar) and relatively disturbed frames for different watermark embedding levels.

A. Robustness

To evaluate the robustness of the multi-level watermarking method, correlations between the watermark sequence (obtained from the sub-band energies of the original signal) and the recovered watermark (obtained by correlating the watermarked signal with the appropriate chirp) were carried out. Different attacks such as the addition of additive white Gaussian noise, up-sampling, down-sampling, re-sampling, lowpass and highpass filtering were undertaken on multi-level watermarked audio data. The tests were undertaken on all the SQAM files and the average results are reported here. For an attack using additive noise, it was found that the watermark embedded using a low frequency sweep was more robust when compared to a high frequency sweep. The average SNR at which detection errors start for level-1 embedding was found -1.9060dB while for level-4 it was 12.14dB. The overall noise robustness was 13.0376dB which clearly shows that level-4 watermark (sweep from 0-60 Hz) is more sensitive to noise. Thus multi-level watermarking can help in embedding critical information using a lower frequency sweep making it more resistant to attack.

The audio sampling rate was varied from 10% to 200% of the sampling frequency and watermark recovered at all levels without any error. Thus, watermark recovery is not effected by alteration in the sampling rate. Amplitude scaling also has no effect on watermark recovery provided it is constant over the entire frequency band.

To simulate a filtering attack, the watermarked signal was passed through a finite impulse response high-pass filter of order ‘50 with cut-off frequency of f_c. This cut-off frequency was varied and extraction of watermark carried out until an error was obtained. It was observed that an error in recovering the watermark occurred at f_c=936Hz. Since a filter has a smooth transition from stopband to passband, the embedded chirp stream is not removed but severely attenuated. Since the chirp stream can be extracted from a high noise background it is therefore possible to extract the watermark at a high cut-off frequency. However, appreciable degradation of in

audio quality occurs using highpass filters with f_c=936Hz. The results obtained are shown in Figure 12. Higher cutoff frequencies can be achieved if the order of designed filter is lowered. The scheme is resistant to lowpass filter attack since the embedded watermark occupies a very low frequency band. Removing the watermark by lowpass filtering effectively removes the entire signal along with the watermark. To have intelligible audio quality, the bandwidth of the lowpass filter should be at least 4 kHz for which the watermark is fully recovered without any error.

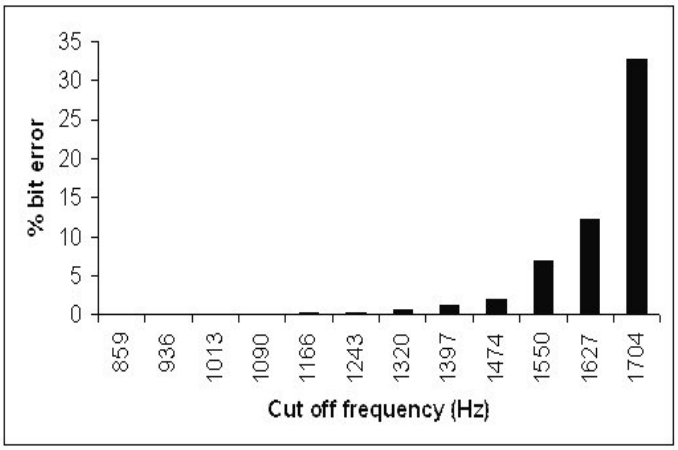

Fig. 12. Percentage bit error achieved for different cut-off frequencies of a highpass filter.

B. Self-Authentication

For appraisal of the self-authentication of an audio signal, two watermark sequences were extracted from the same watermarked signal, the original signal not being required. This provides the advantage of applying blind signal authentication. A signal is authenticated if the two watermark sequences extracted match perfectly. To evaluate this capability, white Gaussian noise was generated and added to the watermarked signal which was detectable at a level of about 45.0732dB SNR. This is because of the change in percentage energies of the sub-bands as a result of noise addition, even if the noise is perfectly 'white'. Further, the total energy also changes through noise addition. Thus the two extracted watermark sequences will not match. A lowpass filter with a cut-off frequency of 23996Hz (which is 99.98% of the full bandwidth) was designed and the watermarked signal passed through this filter. Although the signal undergoes imperceptible changes, because the filtering causes a change in the percentage energy distribution, the recovered watermark sequences did not perfectly match. A difference in 3.83% of bits was found when the two watermarks were compared. Similarly, a highpass filter with a cut-off frequency of 3.84Hz was used and 5.23% bits were found in error. Both filters had finite impulse response of order 50. Altering the sampling rate also changes percentage of sub-band energies and is thus easily detected. By re-sampling at 80% of the sampling frequency a bit error in 16.46% was observed, while for re-sampling at 120% of the sampling frequency an error of 14.33% was observed. While keeping the final rate unaltered, the signal was first up-sampled and then down-sampled. This attack was also detected, the results being shown in Table II.

TABLE II
SIGNAL AUTHENTICATION TEST

Attack type	% bit error	Authentication test
White noise (SNR=45.0732 dB)	0.59	Failed
Lowpass filter (bandwidth 0-23996Hz)	3.83	Failed
Highpass filter (bandwidth 3.84-24000Hz)	5.23	Failed
Sampling rate conversion (4*fs/5)	16.46	Failed
Sampling rate conversion (6*fs/5)	14.33	Failed
Up-sampling by 2+Down-sampling by 2	4.58	Failed
Down-sampling by 2+Up-sampling by 2	6.90	Failed

VII. CASE STUDY: AUDIO DATA VERIFICATION SYSTEM

The proposed scheme has been implemented using MATLAB to provide a simple tagging facility for (wav or MP3) audio files. The system designed for this purpose is available from http://eleceng.dit.ie/arg/downloads/ and is provided by the file Audio_Data_Verification.zip. It system has been designed with a view to tagging audio data prior to a download over the Internet. After installation of the software (which comes with two tagged MP3 files in the Folder *AudioFiles*), the executable file *AudioCode* provides the interface shown in Figure 13 which provides the user with the following options: *Name* (typically of the user of the data to be downloaded), *Password* (defined by the user of the data), *Author* (typically of the audio file, for example). In each case, the field length is limited to eight characters. The fields specifying the date and time are generated automatically. The I/O data fields and the accompanying operations *Tag*, *Compress*, *Show Tag* are specified as shown in Figure 13. There are two principal operations; those associated with tagging a .wav file and tagging a MP3 file as discussed below.

A. Tagging a .wav File

The .wav file is selected (through application of *Browse*) and the name of the output file specified (typically by *Browsing* and then editing the file name as required - the extension is not required). Clicking on the *Tag* button watermarks the file with the information entered by the user. The user has the option of additionally creating an MP3 file of the watermarked audio data by clicking on the *Compress* button.

B. Tagging a MP3 File

The .mp3 file is selected (through application of *Browse*) and the name of the output file specified (typically by *Browsing* and then editing the file name as required - the extension is not required). The system automatically converts the mp3 file to a .wav file for the purpose of watermarking the data. Clicking on the *Tag* button watermarks the file with the information entered by the user. The user then has the option of re-creating an MP3 file of the watermarked audio data by clicking on the *Compress* button.

C. Data Recovery

The watermarked file is selected through *Browse* as an input file (either a .wav or mp3 file). Clicking the *Show Tag* button executes recovery of the watermark which reconstructs the information entered by the user and is displayed in the appropriate fields - *Name, Password, Author, Date, Time*. If the input file is an MP3 file, the system automatically decompresses the file before extraction of the watermark as part of the *Show Tag* option. Assuming that the data has been watermarked, the information it conveys will be legible. However, if the data has not been watermarked or has undergone some form of degradation that is incompatible with the robustness parameter settings that have been hard-wired into the system, the information displayed will be illegible.

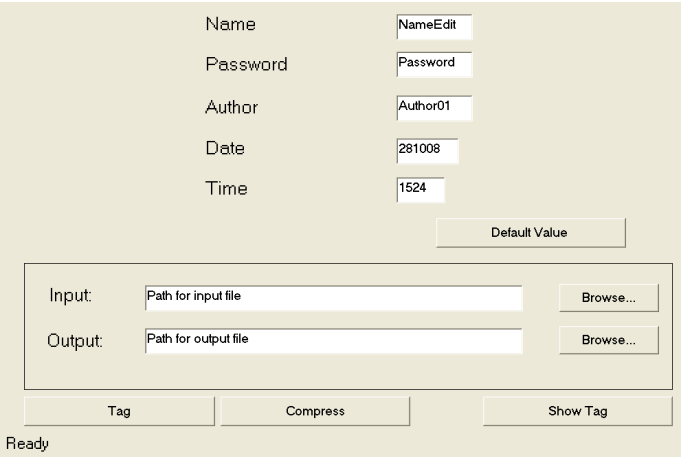

Fig. 13. Interface for audio data authentication system.

VIII. Conclusions

The frequency modulation based watermarking of audio data proposed in this paper is of specific value for the self-authentication of the data for which the method is unique. The proposed scheme has been simulated and tested for various attacks and has been shown to be robust to some of the attacks but with the capability to detect tampering of the signal. This is due to the embedding of a watermark sequence which is derived using the spectral properties of the signal. The Objective Difference Grade evaluated using the basic version of PEAQ with ten model output variables was -0.721 which is in the imperceptible range.

The multi-level watermarking scheme is found to be robust to various attacks as reported. Further, it is observed that a low frequency chirp sweep provides more robustness when compared to a high frequency sweep. Thus, different levels of robustness can be achieved. Since the watermark sequence is derived from the percentage sub-band energies, it is unique and signal dependent. Due to two different processes associated with information extraction, an additional advantage of self-authentication is achieved, thereby making this multi-level watermarking scheme simultaneously robust and fragile.

Acknowledgments

The authors are grateful for the support of the Audio Research Group, School of Electrical Engineering Systems, Dublin Institute of Technology, the Advanced Signal Processing Research Group, Loughborough University, the Science Foundation Ireland and BSkyB.

References

[1] W. Bender, D. Gruhl and N. Morimoto, *Techniques for Data Hiding*, Technical Report, MIT Media Lab, 1994.

[2] D. Gruhl, A. Lu, and W. Bender, *Echo hiding*, 'Information Hiding', 1996, 293-315.

[3] L. Boney, A. H. Tewfik and K. N. Hamdy, *Digital Watermarks for Audio Signals*, IEEE International Conference on Multimedia Computing and Systems, June 1996, 473-480.

[4] M. D. Swanson, B. Zhu, A. H. Tewk and L. Boney, *Robust Audio Watermarking using Perceptual Masking*, Signal Processing 66, 1998, 337-355.

[5] Y. Xiong and Z. X. Ming, *Covert Communication Audio Watermarking Algorithm Based on LSB*, International Conference on Communication Technology, 2006. ICCT 06, 1-4.

[6] B. S. Ko, R. Nishimura and Y. Suzuki, *Time-spread Echo Method for Digital Audio Watermarking*, IEEE Transactions on Multimedia, 7(2), 2005, 212-221.

[7] H. O. Oh, J. W. Seok, J. W. Hong and D. H. Youn, *New Echo Embedding Technique for Robust and Imperceptible Audio Watermarking*, Proc. ICASSP 2001, 1341-1344.

[8] Y. W. Liu and J. O. Smith, *Watermarking Sinusoidal Audio Representations by Quantization Index Modulation in Multiple Frequencies*, Proceedings of ICASSP 2004, Vol. 5, 373-376.

[9] D. Kirovski and H. S. Malvar, *Spread-Spectrum Watermarking of Audio Signals*, IEEE Transactions on Signal Processing, 51(4), 2003, 1020-1033.

[10] L. Lili, H. Jianling and F. Xiangzhong, *Spread-Spectrum Audio Watermark Robust Against Pitch-Scale Modification*, IEEE International Conference on Multimedia, 2007, 1770-1773.

[11] L. Xie, J. Zhang and H. He, *Robust Audio Watermarking Scheme Based on Nonuniform Discrete Fourier Transform*, IEEE International Conference on Engineering of Intelligent Systems, 2006, 1-5.

[12] S. K. Lee and Y. S. Ho, *Digital Audio Watermarking in the Cepstrum Domain*, IEEE Transactions on Consumer Electronics, 46(3), 2007, 744-750.

[13] R. Vieru, R. Tahboub, C. Constantinescu and V. Lazarescu, *New Results using Audio Watermarking Based on the Wavelet Transform*, International Symposium on Signals, Circuits and Systems, 2005. Vol. 2, 441-444.

[14] X. Quan and H. Zhang, *Audio Watermarking based on a Psychoacoustic Model and Adaptive Wavelet Packets*, Proceedings of the 7th International Conference on Signal Processing, 2004 Vol. 3, 2518-2521.

[15] X. Y. Wang and H. Zhao, *A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT*, IEEE Transactions on Signal Processing, 54(12), 2006, 4835-4840.

[16] I. Cox, M. Miller and J. Bloom, *Digital Watermarking*, Morgan-Kaufmann Publishers, 2003.

[17] J. Chou, K. Ramchandran and A. Ortega, *High Capacity Audio Data Hiding for Noisy Channels*, Proceedings of the International Conference on Information Technology: Coding and Computing, 2001, 108-111.

[18] N. Cvejic and T. Seppnen, *Fusing Digital Audio Watermarking and Authentication in Diverse Signal Domains*, Proceedings of the European Signal Processing Conference, 2005, 84-87.

[19] C. I. Podilchuk and E. J. Delp, *Digital watermarking: Algorithms and Applications*, IEEE Signal Processing Magazine, July, 2001, 18(4), 33-46.

[20] S. Katzenbeisser and F. A. P. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking* Artech House, Inc., 2000.

[21] P. Kabal, *An Examination and Interpretation of ITU-R BS.1387: Perceptual Evaluation of Audio Quality*, Technical Report, McGill University, version 2, 2003.

[22] R. J. Anderson and F. A. P. Petitcolas, *On the Limits of Steganography*, IEEE Journal of Selected Areas in Communication (Special issue on Copyright and Piracy Protection), 1998, 16(4), 474-481.

[23] F. A. P. Petitcolas, R. J. Anderson and M. G. Kuhn, *Information Hiding - A Survey*, Proc. IEEE, 1999, 87(7), 1062-1078.

[24] W. Shaoquan, H. Jiwu, D. Huang and Q. Y. Shi, *Efficiently Self-Synchronized Audio Watermarking for Assured Audio Data Transmission*, IEEE Transactions on Broadcasting, March, 2005, 51(1), 69-76.
[25] L. Boney, A. H. Tewfik and K. N. Hamdy, *Digital Watermarks for Audio Signals*, Proc. EUSIPCO, Sept. 1996, Vol. III, 1697-1700.
[26] M. D. Swanson, B. Zhu, A. H. Tewfik and L. Boney, *Robust Audio Watermarking using Perceptual Masking*, Elsevier Signal Processing, Special Issue on Copyright Protection and Access Control, 1998, 66(3), 337-355.
[27] A. N. Lemma, J. Aprea, W. Oomen and L. Van de Kerkhof, *A Temporal Domain Audio Watermarking Technique*, IEEE Transactions on Signal Processing, April 2003, 51(4), 1088-1097.
[28] I. J. Cox, J. Kilian, F. T. Leighton and T. Shamoon, *Secure Spread Spectrum Watermarking for Multimedia*, IEEE Transactions on Image Processing, December 1997, 6(12), 1673-1687.
[29] J. M. Blackledge, *Digital Signal Processing*, Horwood Publications, 2003.
[30] O. Farooq, S. Datta and J. M. Blackledge, *Digital watermarking of Audio Signals using Chirps*, Proceedings of the National Conference on Electronic Circuits and Communication Systems (ECCS-2004),Thapur Institute of Engineering and Technology, September 2004, 178-180.
[31] J. M. Blackledge, *Digital Watermarking and Self-Authentication using Chirp Coding*, ISAST, Transaction on Electronic and Signal Processing, 2007, 1(1), 61-71.
[32] O. Farooq, S. Datta and J. M. Blackledge, *Robust Watermarking of Audio with Blind Self-Authentication*, 7th WSEAS International Conference on Electronics, Hardware, Wireless and Optical Communications, Cambridge, UK, February, 2008, 35-39
[33] O. Farooq, S. Datta and J. M. Blackledge, *Blind Tamper Detection in Audio using Chirp based Robust Watermarking*, WSEAS Transactions on Signal Processing, 4(4), April 2008, 190-200.
[34] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice-Hall 2002.
[35] J. M. Blackledge, *Digital Image Processing*, Horwood Publishing, 2005.
[36] A. Jazinski, *Stochastic Processes and Filtering Theory*, Academic Press, 1970.
[37] A. Papoulis, *Signal Analysis*, McGraw-Hill, 1977.
[38] A. Bateman and W. Yates 1988, *Digital Signal Processing Design*, Pitman, 1988.
[39] A. W. Rihaczek, *Principles of High Resolution Radar*, McGraw-Hill, 1969.
[40] J. J. Kovaly, *Synthetic Aperture Radar*, Artech, 1976.
[41] D. Kundur and D. Hatzinakos, *A Robust Digital Image Watermarking Method using Wavelet Based Fusion*, Proceeding of the International Conference on Image Processing 1997, 544-547.
[42] D. Kundur and D. Hatzinakos, *Digital Watermarking using Multiresolution Wavelet Decomposition*, Proceedings of the International Conference on Acoustics, Speech and Signal Processing, ICASSP '98, 2969-2972.
[43] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.
[44] D. M. A. Lumini, *A Wavelet-based Image Watermarking Scheme*, Proceedings of the International Conference on Information Technology: Coding and Computing 2000, 122-127.
[45] J. J. Chae and B. Manjunath, *A Technique for Image Data Hiding and Reconstruction without a Host Image*, *Security and Watermarking of Multimedia Contents*, SPIE 3657, 1999, 386-396.

Jonathan Blackledge graduated in physics from Imperial College in 1980. He gained a PhD in theoretical physics from London University in 1984 and was then appointed a Research Fellow of Physics at Kings College, London, from 1984 to 1988, specializing in inverse problems in electromagnetism and acoustics. During this period, he worked on a number of industrial research contracts undertaking theoretical and computational research into the applications of inverse scattering theory for the analysis of signals and images. In 1988, he joined the Applied Mathematics and Computing Group at Cranfield University as Lecturer and later, as Senior Lecturer and Head of Group where he promoted postgraduate teaching and research in applied and engineering mathematics in areas which included computer aided engineering, digital signal processing and computer graphics. While at Cranfield, he co-founded Management and Personnel Services Limited through the Cranfield Business School which was originally established for the promotion of management consultancy working in partnership with the Chamber of Commerce. He managed the growth of the company from 1993 to 2007 to include the delivery of a range of National Vocational Qualifications, primarily through the City and Guilds London Institute, including engineering, ICT, business administration and management. In 1994, Jonathan Blackledge was appointed Professor of Applied Mathematics and Head of the Department of Mathematical Sciences at De Montfort University where he expanded the post-graduate and research portfolio of the Department and established the Institute of Simulation Sciences. In 2002 he was appointed Visiting Professor of Information and Communications Technology in the Advanced Signal Processing Research Group, Department of Electronics and Electrical Engineering at Loughborough University, England (a group which he co-founded in 2002 as part of his appointment). In 2004 he was appointed Professor Extraordinaire of Computer Science in the Department of Computer Science at the University of the Western Cape, South Africa. His principal roles at these institutes include the supervision of MSc and MPhil/PhD students and the delivery of specialist short courses for their Continuous Professional Development programmes. He currently holds the prestigious Stokes Professorship in Digital Signal Processing under the Science Foundation Ireland Programme based in the School of Electrical Engineering Systems at Dublin Institute of Technology, Ireland.

Omar Farooq obtained a BSc in Engineering and MSc Engineering from the Z.H. College of Engineering and Technology, AMU, in 1991 and 1993, respectively. He joined the Department of Electronics Engineering as a Lecturer in 1992 and promoted to reader in 2002. He obtained his PhD from Loughborough University, England, under a Commonwealth Scholarship in 2002, continuing his research at Loughborough University on a Post Doctoral Fellowship under the UKIERI scheme from 2006-2007. His area of research is in speech recognition and digital watermarking. He has authored and co-authored over 80 papers in refereed academic journals and national/international conference proceedings. He is a Member of the IEEE USA, ISTE India, ASI India, SSI India and a Fellow of the IETE, India.