International Journal of Systems Signal Control and Engineering Application 8 (1-4): 11-18, 2015
ISSN: 1997-5422

A Heuristics Based Adaptive Audio Watermarking Scheme Through SVD-KFDA

[1]Koyi Lakshmi Prasad and [2]T.C.H. Malleswara Rao
[1]Department of CSE, Bharath University, Chennai, India
[2]Department of CSE, VBIT, Hyderabad, India

Abstract: The heuristics based adaptive audio watermarking technique using SVD-KFDA is proposed in this study. The audio signal is segmented by sub-frames and synchronization code embedded with watermarked signal. On other side, the down-sampling and modulation process using energy relationship is anticipated in order ensure reliable and robust fidelity on watermarked signal. The heuristics process identifies the native energy relationship among sub-frames which is hidden in the watermarked signal through SVD-KFDA. The heuristics based adaptive model is more resilient and can propose extraordinary robustness on usual acoustic data processing and outbreaks associated with common watermarking methods. This system exposes feasibility on implementation and derives less evaluation complexity.

Key words: Audio watermarking, Kernel Fisher Discriminant Analysis (KFDA), Singular Value Decomposition (SVD), energy relationship modulation, synchronization code

INTRODUCTION

In the modern developments World Wide Web and digital multimedia technology consume broadcast and delivery of digital multimedia (such as image, audio and video) over long distances. However, this accessibility permits unlawful replication and spreading of multimedia data. To stretch a copyright fortification of digital data becomes a significant issue. Digital watermarking (Cox *et al*., 2002) technology conventional on prolonged deal of consideration to resolve this issue. The digital watermarking technique is a progression of embedding watermark facts into the transmitted audio signal. This entrenched information can later be perceived or extracted from digital audio signal for numerous applications. Moreover, several applications using an audio watermark which containing copyright protection, replica protection, content verification, biometric and broadcast observing.

A promising methodology to elucidate this issue is digital audio watermarking technique. This is developed through digital audio watermarking and patent information is implanted into multimedia facts without any noticeable variances linked to unique data this is more significant that the evidence embedded cannot be eradicated (Katzenbeisser and Petitcolas, 2000; Potdar *et al*., 2005). In most of the cases, the acoustic signal watermarking system has numerous performance criteria's. It highly concern with signal processing, reliability, feasibility and volume of watermark are basic significant metrics. Meanwhile entrenching watermark into an audio data fetches variations to the digital audio data this can consequence in deprivation of audio acoustic feature. Consequently, it is anticipated that the ruin is undetectable for Human Auditory System (HAS). In most of the cases speaking with high fidelity can simply fulfilled because the embedded watermark gesture isn't apparent by HAS providing PSNR value of the audio signal watermarking facts is beyond 45 dB (Cvejic and Seppanen, 2007). Nevertheless, the watermarked audio data might suffer with some noise, the lossy compression and mutual audio signal handling which are able to result in devastation of the implanted processed data. Henceforth, a watermarking system must retain resilient robustness against noise and the lossy compression with public audio signal processing or outbreaks. However, the conformity and forcefulness are characteristically contradicting between each other. Consequently, the balance among fidelity and strength is essential.

The watermark volume prerequisite by a watermarking structure usually relies on its application level objectives and its volume can be observed as a static constraint as long as the active watermarking scheme can deliver adequate embedding area. In digital audio watermarking scheme, there are huge number of investigation outcomes have been identified and possibly been segregated into two extensive categories: The Temporal domain technique's and the transformed domain technique's (i.e., DWT, DFT, DCT) (Lin *et al*., 2010).

In most of the cases, the temporal domain techniques deliver modest and operative schemes for implanting the

Corresponding Author: Koyi Lakshmi Prasad, Department of CSE, Bharath University, Chennai, India

watermark data into a digital audio, however, they are not forceful to battle common outbreaks. The existing outcomes have exposed that transform domain techniques can be additional to forcefully defend noise, the lossy compression and mutual digital audio processing likened with the temporal domain schemes.

Literature review: In the last few years, numerous audio watermarking approaches were progressed for copyright fortification. One of few DCT based methodologies are thought to be real and operative as like the watermark embedding is prepared in the transform domain. The DCT based method is beneficial as it attempts to gather the native energy in some of the coefficient quantities and likewise it diminishes the association among contiguous sections with in a digital audio. Suresh entrenches a DCT-SVD (Singular value decomposition) based method for embedding then prolonged it to DWT and SVD. Researcher associated with two basic approaches and privileged DCT based technique to be additional forceful. Guo also stretches based on DCT structure which it signifies to be reliable particularly against analog data to digital data vise-verse conversion for air based channel broadcast applications.

The SVD is a commanding tool in the field of image processing domain (Liu and Tan, 2002). Since, the SVD can familiarize to several transform fields, it is being broadly functional in digital audio signal watermarking process (Megias *et al*., 2010). For example, El-Samie (2009) employed a dual strategy to implant the watermark data. Afterwards relating the leading SVD to 2D matrix molded by the final audio signal, researcher slightly merged the projected watermark with diagonal matrix having singular data and formerly accomplished the subsequent SVD on the altered matrix. In addition, the design having matrices comprising left and right (L-to-R) singular vectors should be preserved to extract the final watermark data. Subsequently, Lei *et al.* (2011) extended further the twin strategy and delivered by means of audio data transferred over network systems through segment by segment basis.

Swanson *et al.* (1998) suggested a direct way of embedding watermarking pattern which inserts watermark bits by altering the audio trials. Chen *et al.* (2010a) familiarized a technique in which cluster scales are reformed to attain high forcefulness. Nevertheless, both approaches may falls under low data payload. In another research, Chen *et al.* (2010b) offered an adaptive system which consuming wavelet related entropy, however, robustness and resampling on low-pass filtering attacks are not sufficiently described. Chen *et al.* (2013) anticipated a procedure that implants watermark data through energy proportion scheme. Nevertheless, the SNR outcomes of this algorithmic procedure are not up to the mark. In another research, Chen *et al.* (2013) presented a watermarking method based on optimization scheme which embeds watermark data in very low frequency coefficients of DWT. Still the individual assessment of watermarked audio data has not clearly directed in this arrangement.

In this research study, a novel heuristics digital audio watermarking pattern in the Discrete Wavelet Transform domain using SVD and Kernel Fisher Discriminant Analysis (KFDA) is anticipated (Peng *et al.*, 2013; Bhat *et al.*, 2010). The Native energy relationship among two similar sub-frames of each audio frame depends QIM of the Singular Values (SVs) is embed the watermark in wavelet approximation (Chen and Wornell, 2001).

In order to assure decent fidelity, there are two techniques has implemented further. Down-Sampling Method (DS). The Energy Relationship Modulation Method (ERM). Initially the down-sampling technique can confirm native energies of the dual audio sub-frames to be nearly equal conferring to slow time fluctuating feature of digital audio signal. Consequently, the ERM process which magnitudes the quantity of alteration can meritoriously decrease the humiliation of audio hearing quality (Saadi *et al.*, 2007; Liang and Shi, 2004). This replicates the idea of dynamic watermarking. The watermark indicator will be fabricated by heuristics native energy relationship among two audio sub-frames which is essentially non-linear because the representation from estimating the sub-band coefficients to watermark indication is non-linear. The detection capacity of watermark detector would essentially contain both heuristics ability and generalization ability.

MATERIALS AND METHODS

The KFDA be predominant Kernel-based Adaptive Learning algorithms. It falls on non-linear discriminant analysis via the kernel trick. The Singular Value Decomposition Transformation is an active numerical examination tool used to analyze matrices of values. In the process of SVD each real matrix is disintegrated to three real matrices. The fundamental idea of SVD-KFDA is to resolve the issue of linear Fisher discriminant analysis into inherent feature space V. Still, it is challenging to do so straightly since the dimension of the feature space V could be subjectively huge or even unbounded. In the practical implementation, the inherent feature vector in V cannot be computed explicitly however, it is just prepared by calculating the inner product of two separate vectors in V with kernel process.

From the digital audio signal, we have taken two classes of segments Cx_1 and Cx_2. Each class contains S_i (i.e., i = 1, 2) samples $\{a_1^1, a_2^1, a_3^1 ... a_s^1\} \subseteq Cx_1$ and $\{a_1^2, a_2^2, a_3^2, ... a_s^2\} \subseteq Cx_2$. $S_1+S_2 = S$ and 1& -1 are the set labels for every individual classes. Assume that let φ be the non-linear mapping in which it maps facts as of the origin space G into large dimensional space V.φ: G→V. The prime kernel discriminant vector can be originated by exploiting the Fisher criterion function on high dimensional feature space V. The kernel trick with LDA-function:

$$J(w) = \frac{w^T Cs_B^{\varphi} w}{w^T Cs_w^{\varphi} w} \tag{1}$$

where, w∈V, Cs_B^{φ} and Cs_w^{φ} are the scattered class matrix and inside class scatter matrix correspondingly:

$$Cs_w^{\varphi} = \frac{1}{S}\sum_{i=1}^{2}\sum_{j=1}^{S_i}(\varphi(a_j^i)-q_i^{\delta})(\varphi(a_j^i)-q_i^{\varphi})^T \tag{2}$$

$$Cs_B^{\varphi} = \sum_{i=1}^{2}\frac{S_i}{S}(q_i^{\varphi}-q_0^{\varphi})(q_i^{\varphi}-q_0^{\varphi})^T \tag{3}$$

where, q_i^{φ} is the mean vector of the plotted prepared training samples in the ith class and the mean vector of entire mapped training samples q_0^{φ}:

$$q_i^{\varphi} = \frac{1}{S_i}\sum_{j=1}^{s}\varphi(a_j^i) \tag{4}$$

$$q_0^{\varphi} = \frac{1}{S}\sum_{j=1}^{s}\varphi(a_j) \tag{5}$$

By computing directly φ(a) is not possible at all the time, so by hosting kernel function is quite practicable:

$$\delta = (a, a') = (\varphi(a)\times\varphi(a')) \tag{6}$$

This equation helps to compute the dot product in space V through that we compute φ directly. The desirable selection of **δ** comprises the Gaussian and polynomial kernels. For the kernel matrices of entire samples $\tilde{\delta}$ each class $\tilde{\delta}_1$ and $\tilde{\delta}_2$ be illustrates as follows:

$$\delta = (\delta_{ij} = \delta(a_i, a_j))_{i,\, j=1}^{s}$$

$$\delta_i = (\delta_{ij}^i = \delta(a_j, a_1^i))_{j,\,1=1}^{j=s,\,1=s}, i = 1, 2$$

The theory behind in reproducing kernel function specifies w can be drawn as an extension in the form of:

$$w = \sum_{i=1}^{s}\alpha_i\varphi(a_i) \tag{7}$$

This above mentioned equation also disseminate the object function and also be described as the data α∈A however it appears only in the form of:

$$J(\alpha) = \frac{\alpha^T\delta_B\alpha}{\alpha^T\delta_w\alpha} \tag{8}$$

where $\alpha = \{\alpha_1, \alpha_2, \alpha_3, ..., \alpha_n\}^T$, U_i represented the column vector comprising S_i components with a mutual value of S_i^{-1} and $\boldsymbol{\delta}_B = (\mu_1-\mu_2)(\mu_1-\mu_2)^T$, $\mu_i = \tilde{\delta}_i U_i$:

$$\delta_w = \sum_{i=1}^{2}\tilde{\delta}_i(I_i-U_i)\tilde{\delta}_i^T \tag{9}$$

The I_i is the $S_i{\times}S_i$ identity matrix. The U_i is a $S_i{\times}S_i$ matrix including all elements equals to n-1 Eq. 9 is the KFDA function. In order to get optimal KFDA vector α by exploiting Eq. 9. This is equal to resolving the generalized feature equation:

$$\delta_B\alpha = \lambda\delta_w\alpha \tag{10}$$

Consider that $\boldsymbol{\delta}_w$ is probable to be singular or not properly organized the regularized solution is attained by replacing:

$$\delta_w^v = \delta_w + v\times 1 \tag{11}$$

where, v is denoted as regularization constant. Let $\hat{\alpha}$ be the Eigen vector correspondent to the highest eigen value gained from Eq. 11. Thus, the final version of kernel Fisher discriminant classifier is specified by the succeeding form:

$$f(x) = \text{sgn}\left(\sum_{i=1}^{s}\hat{\alpha}_i\delta(a, a_i) + b\right) \tag{12}$$

where bias b can be calculated by:

$$b = \frac{-\hat{\alpha}\times(s_1\mu_1 + s_2\mu_2)}{s}$$

Synchronization code embedding

Step 1: Originally the digital audio signal has separated into 2 fold segments so as to insert the proposed synchronization code. The major part A_1 of the digital audio signal, further A fragmented into P audio fragments by which every audio segment $SA_1(a)$ consuming N samples. Stretched as:

$$SA_1(a) = \{SA_1(a)(\delta) = A_1(\delta + (a\text{-}1)\times N), 1 \le \delta \le N, 1 \le a \le P\} \tag{13}$$

Step 2: After the expansion, the mean rate of expanded segment ($SA_1(a)$) is denoted as:

$$\overline{SA_1(a)} = \frac{1}{N}\sum_{\delta=1}^{N} SA_1(a)(\delta), (1 \le a \le P) \tag{14}$$

Step 3: Respectively each bit of the anticipated synchronization code will be inserted onto each individual $SA_1(a)$ as:

$$SA_1^{'}(a)(\delta) = SA_1(a)(\delta)\text{-}\overline{SA_1(a)}\text{-}\left|\overline{SA_1(a)}\right|), \text{ if } v(i) = 1 \tag{15}$$

$$SA_1^{'}(a)(\delta) = SA_1(a)(\delta)\text{-}\overline{SA_1(a)} + \left|\overline{SA_1(a)}\right|), \text{ if } v(i) = 0 \tag{16}$$

Subsequently, on other side is to accomplish DWT-SVD of respective audio frame of A_1, it is further decomposed into two levels of its double sub-frames correspondingly. Therefore, we acquire their three similar sub-bands (α_1, α_2, β_2) and select their estimated sub-bands β_2 as embedding points of watermark process. Occasionally, we signify the approximation sub-band of audio sub-frame of A_δ^x as follows:

$$S_\delta^x = \{s_\delta^x(a) | a = 1, 2, ..., \beta/4\}$$

where, $s_\delta^x(a)$, a = 1, 2, ..., $\beta/4$ are represented as the coefficients in approximation sub-band β_2 of A_δ^x depend on S_δ^x. On other hand, the native energy of audio sub-frame A_δ^x is denoted as:

$$U_\delta^x = \sum_{\alpha=1}^{\beta/4} \left|s_\delta^x(a)\right|^2 \tag{17}$$

where, x = 1, 2; δ = 1, 2, ..., E_M. Since, the approximation sub-band of corresponding audio sub-frame focuses high level of energy so, that we conclude $U_\delta^1 \approx U_\delta^2$, δ = 1, 2, ..., E_M the represented strategy of modulation is used to accomplish watermark embedding:

- In the process of embedding every audio frame, there is only one watermark bit will be embedded (1 or 0)
- The designated bit 1 or 0 is embedded and achieved by modulating entire coefficients of S_δ^1 and S_δ^2 such that $U_\delta^1 > U_\delta^2$ or $U_\delta^1 \le U_\delta^2$. This process is mentioned as energy relationship modulation in this study. Based on above mentioned items, the relationship is referred as:

$$v(x)_\delta = 1 \equiv U_\delta^1 > U_\delta^2,\ v(x)_\delta = 0 \equiv U_\delta^1 \le U_\delta^2$$

S_δ^2 from Eq. 15 and 16, it is clear that two energy associations is hidden in audio frames through watermark embedding. For this persistence to modify entire coefficients in S_δ^1 and S_δ^2 such that $U_\delta^1 > U_\delta^2$ or $U_\delta^1 \le U_\delta^2$. Consequently, the succeeding principle is engaged: if $v(x)_\delta$ = 1, we must upturn entire coefficients in S_δ^1 and diminish entire coefficients in S_δ^2 such that $U_\delta^1 > U_\delta^2$; if $v(x)_\delta$ = 0, we would diminish entire coefficients in S_δ^1 and upturn all coefficients in S_δ^2 such that $U_\delta^1 > U_\delta^2$. The values of coefficients in S_δ^x (x = 1, 2) are different in nature. In the meantime to guarantee decent acoustic feature, the coefficient modification scheme which magnitudes the amount of modification, relatively the upper coefficient carries on the bigger modification on other hand the lower coefficient proceeds on minor modification. The advantage of consuming the coefficient modification scheme is to confirm decent acoustic feature of the proposed watermarked audio signal data.

The energy relationship modulation is emphasized through above described from. Let $\Delta U_\delta = \left|U_\delta^1\text{-}U_\delta^2\right| + \eta$, η is a constant variable in turn it is used to manage smallest variation quantity. Correspondingly energy relationship modulation constraints of S_δ^1 and S_δ^2 is defined, respectively:

$$\begin{aligned} U'^1_\delta &= U_\delta^1 + \Delta U_\delta/2,\ U'^2_\delta = U_\delta^2\text{-}\Delta U_\delta/2, \text{ if } v(x)_\delta = 1 \\ U'^1_\delta &= U_\delta^1\text{-}\Delta U_\delta/2,\ U'^2_\delta = U_\delta^2 + \Delta U_\delta/2, \text{ if } v(x)_\delta = 0 \end{aligned} \tag{18}$$

However, the factor η in Eq. 18 having the role of embedding potential. In addition, the native energies of diverse digital audio frames could be variable in nature; obviously their total energy modulations might be also dissimilar. Thus, the above formation is clearly signifies that the total energy modulation is dynamic and adaptive watermarking technique, i.e., the amount energy of the embedded watermark signal is changeable with native audio signal characteristics. In order to accomplish energy relationship modulation through Eq. 18, the subsequent coefficient adaptation formation can be adopted:

$$\begin{aligned} &S'^x_\delta(a) = \text{sgn}(S_\delta^x(a)) \times \sqrt{\frac{(S_\delta^x(a))^2 \times U'^x_\delta}{(S_\delta^x(1))^2 + (S_\delta^x(2))^2 + ... + (S_\delta^x(\beta/4))^2}} \\ &a = 1, 2, ..., \beta/4, x = 1, 2;\ \delta = 1, 2, ..., E_M \end{aligned} \tag{19}$$

The sgn(.): sign function, $S'^x_\delta(a)$ (a = 1, 2, ..., $\beta/4$) are the real coefficients, $S_\delta^x(a)$ (a = 1, 2, ..., $\beta/4$) are post modification coefficients and U'^x_δ is native energies of sub-bands after applying modulation, x = 1, 2, δ = 1, 2, ..., E_M.

In this study, the embedding watermark technique consists of dual components: the reference data with length n represented with the letter T, the proprietor signature S_i with the length of $l_1 \times l_2$. Henceforth, $W = Ts_i = t_1 t_2 \ldots t_n s_1 s_2 \ldots s_{l1 \times l2} = w_1 \ldots w_n w_{n+1} \ldots w_{n+l}$ such that $t_i, s_i \in \{0, 1\}$, $l = l_1 \times l_2$. The proprietor signature S is acquired through a binary input figure with the minimum size of $l_1 \times l_2$ by applying permutation and reshapes process into line format. The letter R is represented as binary sequence produced through Gaussian distribution with zero mean value and slight unit variance.

In this research, the main objective of applying T is to define the energy relationship in dual audio sub-frames. Each audio frame of embedding stage is considered and then it progress into construction of data set to discover the energy relationship in final extraction stage.

Step 1: At the first step, partition the audio signal. The given audio signal A_1 is separated into a series of sub-audio frames, A_δ (δ = 1, 2, .., E_M). Subsequently, their two audio sub-frames, A_δ^1 and A_δ^2 are attained through to Eq. 14.

Step 2: Select the required audio frames to embed watermark signal. In order to attain the security feature, we randomly choose (n+l) audio frames from the (E_M) audio frames based on the secret key δ_1. The identified audio frames have of two basic parts: reference frames. i.e., used to entrench the reference data T. Watermark frames, i.e., to embed the proprietor signature S_i.

Step 3: Manipulating the native energy. The two audio sub-frames of currently selected audio frame be converted into two levels through DWT decomposition correspondingly and their approximation sub-bands S_δ^1 and S_δ^2 is observed. Based on Eq. 18 their native energies U_δ^1 and U_δ^2 are calculated, respectively.

Step 4: Finally, the watermark signal is embedded. The watermarks with reference data T and proprietor signature S_i are embedded onto the elected audio frames by integrating the native energy relationship modulation method. At last for each selected audio frame A_δ then we compute total energies of two audio sub-frames after the energy modulation process U_δ^1 and U_δ^2, respectively through Eq. 18. Then all the coefficients of S_δ^1 and S_δ^2 are altered through Eq. 19:

Step 5: At the receiver side to reconstruct actual audio signal, the audio sub-frame is undergone with inverse DWT transform. Subsequently all audio frames are collected into ultimate watermarked digital audio signal A_1.

Watermark extraction: Generally the noise, process of lossy compression tend the audio signal into low quality data since the modification on after watermarking process may reflect as corrupted data. This will include so many varieties of noise types. Henceforth, the vision on signal processing and the watermarking process having too many endangerous on watermark extracting. In this study, we proposed an adaptive watermark system for digital audio watermarking. Subsequently a watermark system having high detection capability, i.e., high reliability on audio signal processing. It is implemented via machine learning process, nevertheless, the existing static watermarking schemes are failed to guarantee superior simplification ability in theoretical. In this study, the anticipated heuristics-based watermark system is implemented through adaptive native energy relationship in sub-frames of every audio frame. Still, above mentioned native energy relationship fundamentally reproduces a non-linear based relationship using approximation of coefficients to watermark bit:

$$w = f(S_\delta^1(1), S_\delta^1(2), \ldots, S_\delta^1(L/2), S_\delta^2(1), S_\delta^2(2), \ldots, S_\delta^2(L/2) \quad (20)$$

Hence, the KFDA is made for our heuristics based adaptive watermark system since this signifies the powerful non-linear discovering ability. On the other side, the watermark system should also includes fine characterization ability. In addition, the watermark system has enough feasibility on implementation and having minimal computational complexity.

Step 1: Segmenting the received audio signal $\hat{A}$ and convert into a series of sub-audio frames, $\hat{A}_\delta$, δ = 1, 2, ..., E_M. After the segmentation the two sub-frames, $\hat{A}_\delta^1$ and $\hat{A}_\delta^2$ are accomplished by Eq. 14.

Step 2: Choose the sub-audio frames; we select (n+l) audio frames from the given watermarked audio signal using the same secret key δ_1 as it is embedded above. Then the constructed training set G using initial n audio frames with the reference data T is integrated. The training set G is prepared using the coefficients of S_δ^1 and S_δ^2 along with the reference data bit t_δ in T. For practical implementation firstly t_δ = 0 and t_δ = -1 is denoted. KFDA be practiced through the training set G. At final stage the well trained KFDA is used for watermark extraction.

Step 3: The KFDA will be trained using n reference sub-frames; hence the training set G can be constructed as:

$$G = \{(\theta_\delta, t_\delta)|\delta = 1, 2, ..., n\} = \{((S_\delta^1(1), ..., S_\delta^1(L/2),..., S_\delta^2(1)..., S_\delta^1(L/2), t_\delta)|\delta = 1, 2, ..., n\} \quad (21)$$

Such that $\hat{s}_\delta^1(a) \in S_\delta^1$, $\hat{s}_\delta^2(a) \in S_\delta^2$, a = 1, ..., L/2, δ = 1, 2, ..., n. In this presented work, the Gaussian based kernel K_G $(\theta, \theta_i) = \exp(-||\theta-\theta_i||^2/\gamma)^2$ is preferred as kernel process where γ is represented as width parameter. Where by resolving a comprehensive eigen vector issue in Eq. 11, we preserve to attain its optimal discriminant vector $\hat{\alpha}$, through calculating b_i. Therefore, the well trained KFDA is accomplished by Eq. 12.

Step 4: The final process of watermark extraction is performed using sub-audio frames in which proprietor signature is embedded:

$$G' = \{\hat{\theta}_\delta = (\hat{s}_\delta^1(1), ..., \hat{s}_\delta^1(L/2), ..., \hat{s}_\delta^2(1)...\hat{s}_\delta^1(L/2))\ \ \delta = 1, ..., 1 \quad (22)$$

Ultimately, the well-trained KFDA through Eq. 12, we can determine the subsequent outcomes, represented as $\{\hat{t}_\delta|d = 1, ..., 1\}$:

$$\begin{aligned}\hat{t}_\delta &= f(\hat{\theta}_\delta), \delta = 1, ..., 1\\ \hat{\theta}_\delta &= ((\hat{s}_\delta^1(1), ..., \hat{s}_\delta^1(L/2), ..., \hat{s}_\delta^2(1)...\hat{s}_\delta^1(L/2)) \in G'\end{aligned} \quad (23)$$

In conclusion, embedded proprietor signature is extracted by the subsequent constraints:

$$\hat{s}_\delta = \begin{Bmatrix} 1, \text{if } \hat{t}_\delta = 1; \\ 0, \text{if } \hat{t}_\delta = -1; \\ d = 1, 2, ..., l_1 \times l_2 \end{Bmatrix}$$

Step 5: The accomplished watermark image is described as one-dimensional sequences $s_1 s_2 ... s_{l1\times l2}$ of proprietor signature for clear data it again transformed into 2D watermark image.

RESULTS AND DISCUSSION

In our evaluation process, we have considered four different music with variable styles (i.e., Blues, Jazz, classical music and rock and roll) are utilized as basic input audio signals. These music having dissimilar regularity characteristics in turn it is used to assess the performance of our algorithm.

The main evaluation on fidelity and reliability against numerous disputes. The basic setup on all digital audio signals would be 20 sec length Mono format with 16 b/s and 44.1 kHz sample rate. Figure 1 depicts the binary image with size of 64×64 and used as a watermarking image. Figure 2 describes the proprietor monogram S is produced using the watermarking image by permuting and

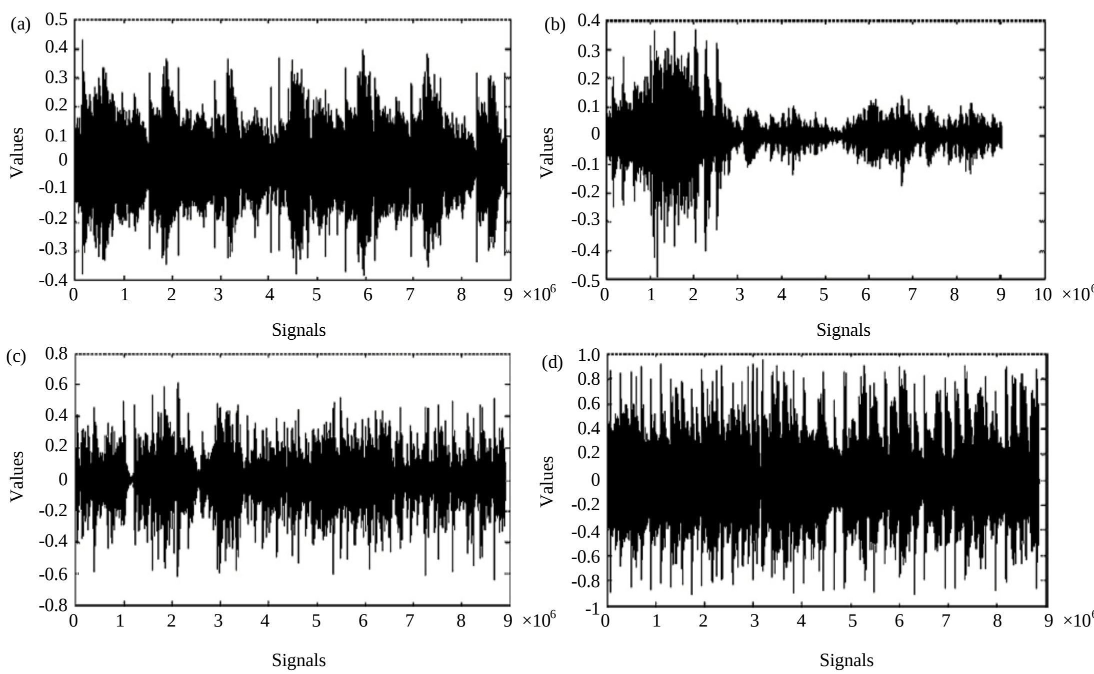

Fig. 1: a) Jazz original audio signal; b) classical original audio signal; c) blues original audio signal and d) rock and roll original audio signal

Fig. 2: The binary watermark image

restructuring through line ordering. Consequently, the constant m = 4098. For the evaluation results, we have setup many parameters as follows:

- Initially the reference data ($T = t_1, t_2, ..., t_n$) is utilized the pseudo random code which is formatted using Gaussian distribution function with zero mean and variable units. We also set the value for n = 400
- The span of each audio sub-frame is described as L = 66. For the energy relationship modulation process we concluded the empiric value $\nabla = 0.0028$.
- The KFDA kernel function setup a constant width parameter $\sigma = 260$ and penalty parameter C = 400

Therefore, the PSNR value is used to assess the variation among basic audio signal A and final outcome audio watermark signal A. Higher PSNR indication signifies that the processed watermarked audio signal is highly suitable with original data. In case of PSNR $\geq$45 dB which signifies that fare fidelity and reliability. In the meantime, standardized cross-correlation (NC) parameters are used to calculate the attack against reliability.

It is a normal justification that human observation might not substantiate well with the PSNR measure. Nevertheless, biased quality assessment of a watermarking method can be directed to deliver an improved test of imperceptibility depends on human sensitivity. Henceforth, the mean opinion score with five-rate point totals is defining as the particular test in our experiments. Table 1 delivers the typical MOS grade scores of the 20 spectators on diverse audio signals. The measureable and individual outcomes of reliability designate that the projected watermarking structure has decent reliability (Fig. 3).

To assess the strength of the anticipated watermarking method there are so common attacks have deployed which comprising, noise intervention, re-sampling low pass filtering and re-quantizing.

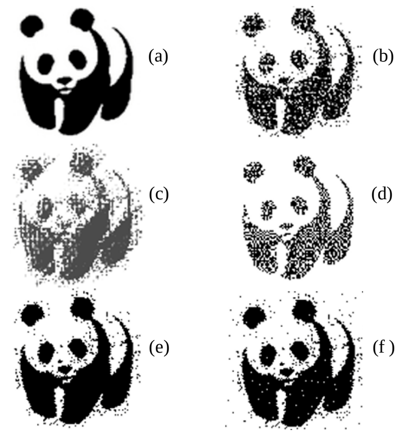

Fig. 3: a-f) The extracted images from watermark audio signals with multiple variations

Table 1: No attach-the performance report

Test audio signals	PSNR	MOS	NC	Capacity (bits)
Jazz	55.46	4.79	1	6976
Classical	53.39	4.78	1	7056
Blues	54.72	4.75	1	6960
Rock and roll	56.13	4.81	1	6928

Table 2: Normalized cross correlation (NC) in favor of low-pass filtering

Test audio signals	SVD-KFDA	Wang's Method	Wang's Method	Xu's Method	Hong's Method
Jazz	0.9869	0.8916	0.9238	0.9317	0.9569
Classical	0.9865	0.8831	0.9176	0.9195	0.9465
Blues	0.9734	0.8905	0.9214	0.9287	0.9534
Rock and roll	0.9602	0.8938	0.9276	0.9335	0.9502

Table 3: Normalized cross correlation (NC) in favor of low-pass filtering re-sampling

Test audio signals	SVD-KFDA	Wang's Method	Wang's Method	Xu's Method	Hong's Method
Jazz	0.9511	0.8412	0.9054	0.9176	0.9311
Classical	0.9649	0.8332	0.9011	0.9103	0.9249
Blues	0.9595	0.8374	0.9036	0.9164	0.9295
Rock-and-roll	0.9425	0.8426	0.9097	0.9218	0.9325

Table 4: Normalized cross correlation (NC) in favor of low-pass filtering re-quantizing

Test audio signals	SVD-KFDA	Wang's Method	Wang's Method	Xu's Method	Hong's Method
Jazz	1	0.8456	0.9095	1	1
Classical	1	0.8359	0.9053	1	1
Blues	1	0.8442	0.9078	1	1
Rock and roll	1	0.8476	0.9125	1	1

The anticipated scheme is compared with other relative methods (Xu *et al*., 2007; Wang *et al*., 2005; Bhat *et al*., 2010) (Fig. 3 and Table 2-4). The other methods utilized the middle coefficients and methods uses various techniques such ANN, SVR, etc., the outcome of watermarked audio datails again undergone with down-sampling and up sampling instantly (22.05-44.1 kHz). For the process of re-quantizing, the incoming audio signal with 16 bit watermarked into 8 bit and then immediately back to 16 bit. Table 3-4 dictate the experimental results of re-sampling and re-quantizing, correspondingly.

CONCLUSION

In this study, we proposed a heuristics-based adaptive audio watermarking scheme through DWT-Singular Value Decomposition and Kernel Fisher Discriminant Analysis (SVD-KFDA). In this anticipated model, the fair reliability is derived by modulating three basic formations, i.e., synchronization code embedding, down sampling technique and native energy relationship technique. The synchronization code is extracted separately and the initial audio data signal is segmented into two sub-frames and each of which holding same energy levels. With this fundamental idea, the native energy relationship modulation method utilized the rule of proportioning the total quantity of alteration to implant watermark bits. Sebsquently, this also encrypt the relative native energy into sub-audio frames. However, this will reduce audio hearing quality drastically. The heuristics based adaptive audio watermark process can learn the native energy relationships using training sets. However, the robustness on battling de-synchronization outbreaks is still a problematic issue in digital audio watermark method so far. Henceforth, our subsequent extension is to suggest an digital audio watermarking method against de-synchronization outbreaks by merging other relative schemes.

REFERENCES

Bhat, V., I. Sengupta and A. Das, 2010. An adaptive audio watermarking based on the singular value decomposition in the wavelet domain. Digital Signal Process., 20: 1547-1558.

Chen, B. and G.W. Wornell, 2001. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. IEEE Trans. Inform. Theory, 47: 1423-1443.

Chen, S.T., H.N. Huang, C.J. Chen and G.D. Wu, 2010a. Energy-proportion based scheme for audio watermarking. IET Signal Process., 4: 576-587.

Chen, S.T., G.D. Wu and H.N. Huang, 2010b. Wavelet-domain audio watermarking scheme using optimisation-based quantisation. IET Signal Process., 4: 720-727.

Chen, S.T., H.N. Huang, C.J. Chen, K.K. Tseng and S.Y. Tu, 2013. Adaptive audio watermarking via the optimization point of view on the wavelet-based entropy. Digital Signal Process., 23: 971-980.

Cox, I.J., M.L. Miller and J.A. Bloom, 2002. Digital Watermarking. Morgan Kaufmann Publishers, San Francisco, USA.

Cvejic, N. and T. Seppanen, 2007. Digital Audio Watermarking Techniques and Technologies Applications and Benchmarks. IGI Global, USA. UK., ISBN-10: 1599045133, pp: 328.

El-Samie, F.E.A., 2009. An efficient singular value decomposition algorithm for digital audio watermarking. Int. J. Speech Technol., 12: 27-45.

Katzenbeisser, S. and F.A.P. Petitcolas, 2000. Information hiding techniques for steganography and digital watermarking. EDP Audit Control Security Newslett., 28: 1-2.

Lei, B.Y., Y. Soon and Z. Li, 2011. Blind and robust audio watermarking scheme based on SVD-DCT. Signal Process., 91: 1973-1984.

Liang, Z. and P. Shi, 2004. An efficient and effective method to solve kernel fisher discriminant analysis. Neurocomputing, 61: 485-493.

Lin, S.D., C.C. Huang and J.H. Lin, 2010. A hybrid audio watermarking technique in cepstrum domain. ICIC Express Lett., 4: 1597-1602.

Liu, R. and T. Tan, 2002. An SVD-Based watermarking scheme for protecting rightful ownership. IEEE Trans. Multimedia, 4: 121-128.

Megias, D., J. Serra-Ruiz and M. Fallahpour, 2010. Efficient self-synchronised blind audio watermarking system based on time domain and FFT amplitude modification. Signal Process., 90: 3078-3092.

Peng, H., B. Li, X. Luo, J. Wang and Z. Zhang, 2013. A learning-based audio watermarking scheme using kernel fisher discriminant analysis. Digital Signal Process., 23: 382-389.

Potdar, V.M., S. Han and E. Chang, 2005. A survey of digital image watermarking techniques. Proceedings of the 3rd International Conference on Industrial Informatics, August 10-12, 2005, IEEE, pp: 709-716.

Saadi, K., N.L. Talbot and G.C. Cawley, 2007. Optimally regularised kernel fisher discriminant classification. Neural Networks, 20: 832-841.

Swanson, M.D., B. Zhu, A.H. Tewfik and L. Boney, 1998. Robust audio watermarking using perceptual masking. Signal Process., 66: 337-355.

Wang, C., X. Ma, X. Cong and F. Yin, 2005. An Audio Watermarking Scheme with Neural Network. In: Advances in Neural Networks-ISNN 2005. Wang, J., X. Liao and Z. Yi (Eds.). Springer, Berlin, Heidelberg, Germany, ISBN: 978-3-540-25913-8, pp: 795-800.

Xu, X., H. Peng and C. He, 2007. DWT-Based Audio Watermarking using Support Vector Regression and Subsampling. In: Applications of Fuzzy Sets Theory. Masulli, F., S. Mitra and G. Pasi (Eds.). Springer, Berlin, Heidelberg , Germany, ISBN: 978-3-540-73399-7, pp: 136-144.