SEMI-REVERSIBLE QUANTIZATION BASED DATA HIDING USING MISSING SAMPLES RECOVERY TECHNIQUE

Ali Ameri, Mohammad J. Saberian, Mohammad A. Akhaee, and Farokh Marvasti
Advanced **C**ommunications **R**esearch **I**nstitute (ACRI)
Electrical Engineering Dept, Sharif University of Technology, Tehran, Iran
E-mail: a_ameri83@std.sbu.ac.ir, ess.saberian@gmail.com akhaee@ee.sharif.edu marvasti@sharif.edu

ABSTRACT

A blind semi-invertible quantization based data hiding scheme, which reconstructs the original signal with high precision has been proposed. In order to produce correlated quantization yielding reversibility of the quantization based approach, a new transform domain has been introduced. In decoder, by compensating the quantization error and using the iterative technique, the original signal is recovered; then the watermarked signal is compared with the reconstructed original signal and hidden data is retrieved. Simulation results show that the proposed method in comparison with other reversible methods imposes less distortion and thus a higher Signal to Noise Ratio (SNR) is achieved.

Index Terms *— Semi-reversible, Data Hiding, Quantization based embedding, Iterative method*

1. INTRODUCTION

In some practical issues, any distortion caused by watermarking is intolerable. In these cases, it is necessary to recover the original signal after data extracting without any loss. For instance, in medical diagnosis and law enforcement, it is vital to reconstruct the original signal when the embedding message is retrieved. Also, high resolution aerial images, in expense of high cost, are obtained for some other post processing. Thus, any distortion that reduces this precision should be removed completely. The watermarking algorithm satisfying these qualifications are called reversible, lossless or distortion-free techniques [1]

Although, several watermarking algorithms have been already developed, they rarely were reversible. For example, the Least Significant Bit plane (LSB) scheme proposed in [2] is not invertible due to bit substitution by data. Another well-known watermarking approach called Spread Spectrum (SS) [3] is lossy because of truncation and round-off error. Furthermore, the Quantization Index Modulation (QIM) which is the base of several watermarking techniques [4] is not distortion-free owing to the error of quantization.

Several reversible watermarking algorithms have been proposed recently [5] using modulo 256 addition for data hiding in the eight-bit gray scale image. Macq *et al*. [6] used a lossless multi resolution transform domain, considering the patchwork algorithm, and modulo 256 addition, to develop a robust reversible marking approach. The same idea for distortion free data hiding has been investigated by using integer discrete cosine transform (DCT) [7]. Celik *et al*. developed a generalized reversible version of LSB technique by employing a prediction based conditional entropy coder [8]. Furthermore, by using integer transform and modulo N addition the conventional SS technique has been changed to distortion-free algorithm in [9]. Moreover, in [10], an invertible watermarking scheme based on integer to integer wavelet transform has been proposed.

In this paper, a novel semi-reversible quantization based data hiding method has been proposed. Quantization is performed in a transform domain called Point to Point Graph (PPG), in such a way that some pair samples of the cover media have the correlated quantization error [11]. In the detection process, using this correlation, suitable subtraction, the periodic non-uniform miss sampled version of the original signal is achieved. For original signal recovery the iterative method [12] has been employed and the embedded data is retrieved. Although this algorithm is suitable for any kind of signal, we apply the proposed scheme to audio signals for better comparison with existing schemes.

2. PROPOSED METHOD

2.1. PPG Transformation

In this transform domain, the time samples of the host signal in each frame are put together in pairs in such a way that they make the set $\{(s_1, s_{k+1}),(s_2,s_{k+2}),\ldots\}$ where k is an arbitrarily index and called index of PPG. This process is done on the $2k$ first samples and after that it is repeated on the other samples.

The most important advantage of using this transform domain is that the quantization noise of paired samples (which are not necessarily adjacent) is related to each other. This relation will be helpful to extract the original signal and the embedded data from watermarked one [11].

2.2. Data Embedding Process

First, the host signal S is separated into two parts (by a filter bank): S_{LP} which has the low frequency components of S and S_{HP} which has the high frequency ones of that. Embedding in low frequency components causes more robustness while keeps the watermark transparent. The PPG transformation is applied to the frames of signal S_{LP} with the index of k which should be selected in such a way to achieve both more correlation among the samples and more transparency of embedded data in the time domain shape.

The basis of data embedding is on the quantization of PPG points to the concentric circles which is called quantization levels. In this approach, each point is stretched in direction of $tan^{-1}(10/9)$ to intersect the two adjacent quantization levels. This value is attained in such a way that the algorithm to be robust against lowpass filtering attack while keeping the differential filter in (3) invertible. In fact this value is hand-optimized to compromise the existing trade off. For embedding message one, the point with bigger radius and for embedding message zero, the one with smaller radius is selected.

In order to determine the intersection points depending on the embedded message, only a quadratic equation must be solved as follows:

$$\begin{cases} x = x_0 + d\cos(\theta) \\ y = y_0 + d\sin(\theta) \end{cases}$$

where $\theta = \tan^{-1}(10/9)$

$$x_0^2 + y_0^2 = R_0^2$$

$$x^2 + y^2 = R_q^2 \Rightarrow d^2 + 2d(x_0\cos(\theta) + y_0\sin(\theta)) + R_0^2 - R_q^0 = 0 \quad (1)$$

which (x_0,y_0) is a typical PPG point of the signal, (x,y) is the corresponding point of (x_0,y_0) after data hiding process and d represents the magnitude of the displacement vector between them.

The data embedding procedure can be described as follows:

Step1: By using a filter bank, the host signal S is separated into two parts: S_{LP} and S_{HP}.

Step2: The PPG transform is applied to each frame of S_{LP}.

Step3: Depending to the embedded message, for all points of a frame, the intersection equation (1) must be solved to achieve the quantized point in the watermarked signal.

Step4: After information embedding in PPG points of the host signal, by applying the inverse of PPG transformation (PPG^{-1}), the signal S_{WLP} can be obtained.

Step5: The signals S_{WLP} and S_{HP} are summed together to form the watermarked signal S_w.

2.3. Data Extraction Process

At the receiver, first, by using the same filter bank, the low pass part of the watermarked signal, S_w, is separated (S_{WLP}). Next, the out of band quantization noise will be compensated nearly by a re-quantization process. This process maps each point of the signal S_{WLP} to the nearest level. The resulted signal is shown by S_{RQ}. Then, we try to obtain a well approximation of the low pass part of the original signal. Afterward, by comparing the low pass part of the received signal and that of the reconstructed original one, the embedded data can be retrieved. The extraction method of the low pass part of the host signal from the received one is as following:

The effect of the quantization noise between each sample is 9/10 of the k^{th} next sample (this is because of the mapping in direction with slope 10/9). Thus, a kind of correlation is existed among the quantization noise. In the first step, by exploiting this kind of correlation, the original signal can be attained. To achieve this aim, the host signal is subtracted from 9/10 of its k^{th} shifted version. The resulted signals forms of two types of samples: one is the differential version samples of the signal (without any noise), and the other is the noisy signal samples that the variance of the noise is twice of the received one. As these noisy samples are unworthy, we force them to be zero. These zeros are periodically non-uniform.

By using this technique, the achieved signal is a missed sampled version of the original signal passed through the differential filter which can be written as follow:

$$y[n] = x[n] - 0.9x[n-k]$$

$$F_{diff}(e^{jw}) = 1 - 0.9e^{jkT_s w} \quad (2)$$

which Ts is the sampling period time , k is the index of the PPG transformation. An important issue about the differential filter $F_{diff}(e^{jw})$ is that it is invertible. The $F^{-1}_{diff}(e^{jw})$ can be written as:

$$F_{diff}^{-1}(e^{jw}) = \frac{1}{1 - 0.9e^{jkT_s w}} \quad (3)$$

This is the main reason of the kind of quantization proposed in the last part. It should be notice that if we quantized the host signal in direction of the line $y=x$, then the resulted differential filter is not be invertible.

The reconstruction methods using nonuniform samples are investigated in the next sub-section. After applying the reconstruction methods, the distortion caused by differential process should be compensated. Hence, $F^{-1}_{diff}(e^{jw})$ is applied.

The achieved signal is a suitable approximation for the low pass part of the original one. In the dewatermarking process, this signal is used in data extracting. If the message one is embedded in the host signal, its points will be quantized to the next levels (with larger radii). Thus, the sum of the radii of the points will be larger than that of the host signal and vice versa for embedding zero. Therefore, by comparing this characteristic of the received signal with the obtained original one, the embedded bit can be estimated. Since the low frequency part of the host signal is well reconstructed, by adding the high frequency components of the host signal which is kept unchanged, the original signal is recovered completely.

The steps of de-watermarking algorithm are explained as follows:

Step1: Using the same filter bank, the low pass part of the watermarked signal, S_w, is separated (S_{WLP}).

Step2: The pre-quantize process explained above is performed on the S_{wLP}. The obtained signal is shown by S_{PQ}.

Step3: The filter $F_{diff}(e^{jw})$ defined in relation (2), is applied to S_{PQ} to achieve the signal S_{diff}.

Step4: Using the iterative technique (4), the filtered version of the S_{LP} is attained.
Step5: The inverse of the differential filter ($F_{diff}^{-1}(e^{jw})$) is applied to the signal resulted in step 4.
Step6: The radii summation of the PPG points of the recovered signals and the lowpass component of the received signal (S_{WLP}) are compared to estimate the embedded message.

2.4. The Iterative reconstructing Method

Reconstruction of the signal using the non-uniform samples is a problem which has been investigated and developed by [12]. According to the theory of non-uniform sampling, if the rate of remained samples is more than the Nyquist rate, the reconstruction is possible. The standard iterative method is given by [13]:

$$x_{n+1} = \lambda G(x(t)) + (I - \lambda G)x_n(t) \qquad (4)$$

where λ is the relaxation parameter which indicates the convergence rate of the method.

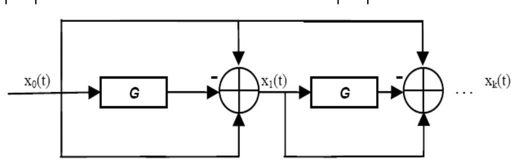
The system G used in the iterative method is formed of two parts: one is a zero putting function and another is a low pass filter. The block diagram of G and iterative method is illustrated in Fig. 1 It should be noticed that in order to satisfy the Nyquist rate after miss sampling leading to perfect reconstruction, the bandwidth of the low pass filter used in this method must be at least more than twice the bandwidth of S_{LP}. Furthermore, the proof of convergence of this method and the value of optimum relaxation parameter which results in the best convergence rate are investigated in the next section.

3. Performance Analysis

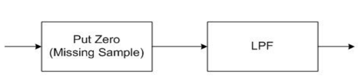
3.1. Capacity Calculation

Assume a typical PPG point (P_i) which is mapped in the data hiding process to another point Q_i. It can be written that $\vec{Q}_i = \vec{P}_i + \vec{D}_i$ which $\vec{D}_i$ represents the exerted distortion vector. This distortion can be mapped on the horizontal and vertical axes which are shown by d_i and d_{i+k} respectively. It is worth to notice that the d_i and d_{i+k} represent the inserted distortions on the pair S_i and S_{i+k} samples of the host signal. Thus, it can be written that:

$$d_i = \left|\vec{D}_i\right|.Cos\,(tan^{-1}(10/9)) \quad , \quad d_{i+k} = \left|\vec{D}_i\right|.Sin\,(tan^{-1}(10/9))$$



a



b

Fig. 1. a) The reconstruction block diagram using standard iterative method and **b)** The block G in the standard iterative method

In order to calculate the distortion power, the distribution of $|D_i|$ is supposed to be uniformed over [0, Δ], which Δ is the distance between quantization levels. P_D, the power of the exerted distortion can be derived as follows:

$$P_D = E\{\sum_i (d_i^2 + d_{i+k}^2)\} = \sum_i E\{D_i^2\} = \frac{1}{12}\Delta^2 \qquad (5)$$

which N is the frame length of the host signal. The distortion power can be controlled by changing Δ and N to achieve acceptable quality of the watermarked signal. However, choosing smaller Δ, causes more fragility against noise. Thus, a trade-off is existed in choosing Δ which represents the physical implementation of the process.

The derivation of Distortion to Noise Ratio (DNR) and the capacity are as following [4]:

$$C = \frac{1}{2}\log_2(1 + DNR) \qquad (6)$$

where $DNR = \frac{P_D}{\sigma_n^2}$ and $\sigma^2{}_n$ is the variance of the AWGN.

3.2. Iterative Convergence

The applied system in the used iterative method includes of a zero-putting block and a low pass filter (Fig. 1). This system can be assumed as a function G: R^N - R^N where N is the frame length. According to the [12], to prove the convergence of the iterative method, it is enough to show that $\|I - \nabla G\| < 1$ which ∇ indicates the gradient operator.

If an m order zero putting function (a function which puts m zero in the signal with period 2m) is applied to the input signal shown by $x[n]$, the output can be written as follows:

$$y[n] = \begin{cases} x[n] & n = im, \ldots (i+1)m - 1 \qquad i = 0, 2, \ldots. \\ 0 & otherwise \end{cases}$$

which $x[n]$ is the input signal and $y[n]$ is the output one. The frequency spectrum of the output signal can be derived as follows:

$$Y\left(e^{jw}\right) = \sum_k C_k X\left(e^{j(w - k\frac{\pi}{m})}\right)$$

$$\text{Where } C_k = \frac{1}{2m}\frac{\sin\left(m \times \frac{w}{2}\right)}{\sin\left(\frac{w}{2}\right)}\Bigg|_{w = k\frac{\pi}{m}} \qquad k = 0, 1, \ldots, 2m - 1$$

Since the low frequency part of the host signal has been used, according to Fig. 2, it can be assumed that in the above process, the main frequency spectrum is only overlapped with the its two adjacent shifted spectrums. Using the N dimensional DFT and assuming the bandwidth of the input signal $x[n]$ is L, it can be written that:

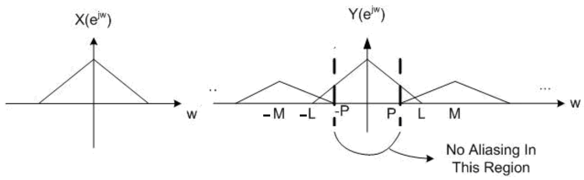


Fig. 2. The spectrum of the input and output signals.

$$Y[k]=\begin{cases} X[k] & |k|<P \\ X[k]+\alpha X[-(M-k)] & P\le |k|\le L \\ 0 & otherwise \end{cases} \qquad (7)$$

where $M=\dfrac{2m}{N}$

This relation indicates that the spectrum of the output signal is pure in the first P points while in the other points it some interference exist. In the matrix notation, the DFT points of the output can be written as follows:

$$Y=G(X)=WX \qquad (8)$$

where $W=\begin{bmatrix} 0 & . & & & & . & & & \\ & 0 & & & & & & & \\ & & 1 & . & . & \alpha & 0 & & \\ & & 0 & 1 & . & . & \alpha & & \\ & & & & . & . & & & \\ & & \alpha & . & . & 1 & . & & \\ & & 0 & \alpha & . & . & 1 & & \\ & & & & . & & & 0 & \\ & & & & . & & & & 0 \end{bmatrix}$

and $\alpha=C_1=C_{-1}=\dfrac{1}{2m}\dfrac{1}{\sin\left(\pi/2m\right)}$

Because this system is linear for any input signal, *x[n]*, hence, $\nabla G=W$.

Investigating the convergence of the iterative method, $\|I-\nabla G\|$ must be calculated.

$$I-W=\begin{bmatrix} 1 & . & & & & . & & & \\ & 1 & & & & & & & \\ & & 0 & . & . & \alpha & 0 & & \\ & & 0 & 0 & . & . & \alpha & & \\ & & & & . & . & & & \\ & & \alpha & . & . & 0 & . & & \\ & & 0 & \alpha & . & . & 0 & & \\ & & & & . & & & 1 & \\ & & & & . & & & & 1 \end{bmatrix}$$

It is necessary to notice that in the applied iterative technique, the used input signals are inherently low pass; i.e. $X[k]=0$ for $|k|>L$. Thus, the one elements on the main diagonal of the matrix *I-W* can be ignored since they are no effect in the norm of resulted signal. In other words, in the multiplication of the *(I-W)* matrix into *X*, they are no effective in the norm of the result. As a consequence, in the space of low pass signals, *I-W* can be written as:

$$W'=\begin{bmatrix} 0 & . & . & \alpha & 0 \\ 0 & 0 & . & . & \alpha \\ & & . & . & \\ \alpha & . & . & 0 & . \\ 0 & \alpha & . & . & 0 \end{bmatrix}$$

Finally, the $\|W'\|$ must be calculated.

$$\|W'\|=Sup\frac{\|W'x\|}{\|x\|}=Sup\frac{\alpha\sqrt{\sum_{P}^{L}x[i]^2+\sum_{-L}^{-P}x[i]^2}}{\sqrt{\sum_{-N/2}^{-N/2}x[i]^2}}<\alpha<1 \qquad (9)$$

where $X=[x_{-N/2},x_{-(N/2)+1},\ldots,x_{N/2}]$.Using (9), the convergence of the iterative method is apparent.

3.3. Calculating the Optimum λ

A good approach in the iterative method to achieve a better convergence rate is using a proportion of the system G ($\lambda.G$) instead of G itself. In this part, the optimum relaxation parameter (λ_{opt}) that causes the best convergence rate is derived. In relation (4), if the system $\lambda.G$ is used instead of G, the matrix of W' can be rewritten as follows:

$$W'=\begin{bmatrix} 1-\lambda & . & . & \lambda\alpha & 0 \\ 0 & 1-\lambda & . & . & \lambda\alpha \\ & & . & . & \\ \lambda\alpha & . & . & 1-\lambda & . \\ 0 & \lambda\alpha & . & . & 1-\lambda \end{bmatrix}$$

Similar to relations (9), it can be written that:

$$\|W'x\|=\sup\frac{\|W'x\|}{\|x\|}=\sup\frac{\sqrt{\sum_{-N/2}^{N/2}|(1-\lambda)x[k]+\lambda\alpha x[k+P]|^2}}{\sqrt{\sum_{-N/2}^{N/2}|x[i]|^2}} \qquad (10)$$

By derivation, the optimum λ can be derived as follows:

$$\lambda_{opt}=\frac{\sum_{-N/2}^{-N/2}(|x[k]|^2-|x[k+P]|^2)}{\sum_{-N/2}^{-N/2}(|x[k]|^2+|x[k+P]|^2-2\,\mathrm{Re}\{x[k]x[k+p]^*\})} \qquad (11)$$

4. SIMULATION RESULTS AND DISCUSSION

The proposed algorithm can be applied to any kind of media such as audio, image, or other type of data, depending on the application. Here, as an example, we simulate the proposed method over three (march, medium, and light) audio signals sampled at 44100 kHz and quantized with 16-bit. The cut off frequency of the analyzed filter is set to 4-kHz. The index of PPG transformation, *k*, is fixed to four in all frames. The number of quantization levels is 30; however, it can be adjusted to control the distortion imposed to the host signal.

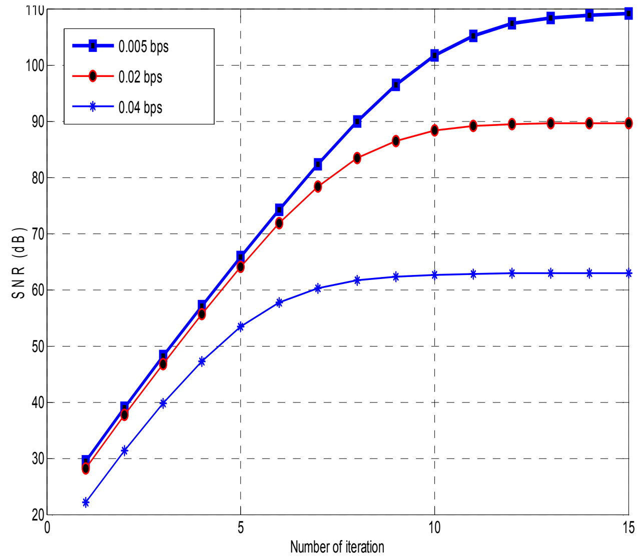


Fig. 3. SNR improvement of the reconstructed signal for various embedding capacity (medium audio)

Fig. 3 shows the SNR improvement of the reconstructed signal for various embedding capacity versus the iteration number. The optimum value of λ is set to .95 according to (11). From this figure it can be deduced that about 40-80 dB SNR improvement is achieved which illustrates the original signal is well reconstructed.

Table I depicts the amount of distortion (signal quality) for the proposed method in comparison with generalized LSB [8], the integer DCT-based scheme [7], and integer to integer wavelet-based method [10]. However, since these recent investigations are based on image signals, for fair comparison, we simulate these techniques over audio signals.

In spite of other available methods, in the proposed scheme, all samples of cover media in each frame participate in data hiding process. Therefore, the quality of the host signal slightly differs with increasing the payload. From Table 1 it is obvious that the proposed scheme achieves higher payload with less distortion in the host signal. Thus, it outperforms other invertible techniques reported in literatures. Moreover, it should be noted that by decreasing the embedding capacity and keeping the distortion value, depending on the applications, desirable robustness can be attained.

Table. 1 Comparison of embedding capacity versus distortion (SNR dB)

Capacity (bps)	Proposed method	G-LSB [8]	Int. DCT-based [7]	Int. to Int. Wavelet-based [10]
0.5	**32.5**	20.4	22.7	24.5
0.25	**32.0**	23.7	26	28.7
0.16	**31.3**	28.7	29.6	**31.3**
0.125	**33.3**	30	30.8	**33.4**
0.0625	32.7	34.2	35.2	**42.1**

5. CONCLUSION

The main achievement of this paper is indicating how quantization based algorithms can be semi-reversible. Thus, this technique can be applied to a wide category of methods using QIM for data hiding purposes. Experimental results show that the original signal can be reconstructed up to SNR about 100 dB. This precision is enough in many applications such as secure communication in which the rendering of original signal is required. Moreover, the distortion of the proposed method can easily controlled by quantization levels and the SNR of the marked signal slightly differs with payload. This is another advantage of the proposed method in comparison with common approach reported so far.

REFERENCES

[1] Z. Ni, Y. Q. Shi, N. Ansari, and W. Su, "Reversible Data Hiding," *IEEE Trans. Circuits Syst.: Video Technol.*., vol. 16, no. 3, pp. 354-362, MAR 2006.

[2] J. Irvine and D. Harle, *Data Communications and Networks: An Engineering Approach*. New York: Wiley, 2002.

[3] J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process*., vol. 6, no.12, pp. 1673–1687, Dec. 1997.

[4] B. Chen and G. W. Wornell, "Quantization index modulation: a class of provably good methods for digital watermarking and information embedding," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.

[5] C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel, "Lossless Recovery of an Original Image Containing Embedded Data," *U.S. Patent 6 278 791 B1*, Aug. 21, 2001.

[6] B. Macq and F. Deweyand, "Trusted headers for medical images," *presented at the DFG VIII-D II Watermarking Workshop*, Erlangen, Germany, Oct. 1999.

[7] B. Yang, M. Schmucker, W. Funk, C. Busch, and S. Sun, "Integer DCT-based reversible watermarking for images using compandingtechnique," *in Proc. SPIE, Security, Steganography, and Watermarkingof Multimedia Contents*, San Jose, CA, Jan. 2004, pp. 405–415.

[8] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Reversible data hiding," *in Proc. IEEE ICIP*, Rochester, NY, vol. 2, pp. 157–160, Sep. 2002.

[9] X. Tong, C. Y. Shao, X. G. Xu, and X. M. Niu, "REvesible Data-Hiding Scheme for 2-D Vector Maps Based on Difference Expansion," *IEEE Trans. .Inform. Foren. and Secur.*, vol. 2, no. 3, pp. 321-330, Sept 2007.

[10] S. Lee, C. D. Yoo, and T. Kalker, "Reversible Image Watermarking Based on Integer-to-Integer Wavelet Transform", *IEEE Trans. .Inform. Foren. and Secur.*, vol. 2, no. 3, pp. 321-330, Sept 2007.

[11] M. A. Akhaee, M. J. Saberian, S. Feizi, and F. Marvasti, "Robust Audio Data Hiding Using Correlated Quantization with Histogram Based Detector" *IEEE Trans. MultiMedia*, to be appeard.

[12] F. Marvasti, *Nonuniform Sampling: Theory and Practice*, Kluwer Academic/Plenum Publishers, 2001.

[13] F. Marvasti, "An iterative method to compensate for the interpolation distortion," *IEEE Trans. ASSP*, vol. 3, no.1, pp 1617-1621, May 1989