Iraqi Journal of Statistical Science (20) 2011
The Fourth Scientific Conference of the College of Computer Science & Mathematics
pp [309-320]

Speech compression based on wavelet and contourlet Transformation

كبس الكلام باستخدام التحويل ألموجي والتحويل الكنتوري

Dr.Khalil I. Alsaif* Hadia Salih Albadrani**

الخلاصة

كبس الكلام مهم جداً في كثير من التطبيقات. في هذا البحث تم كبس الكلام باستخدام التحويل الموجي والتحويل الكنتوري ،وذلك بأخذ إشارة الكلام وتحويلها إلى مصفوفة ذات بعدين (لكي تتلاءم مع التحويل الكنتوري) وإجراء التحويل الموجي ،ثم اخذ المعاملات لهذا التحويل وإجراء التحويل الكنتوري عليها . بعد نقل الكلام أو خزنه وعند استعماله يعاد فك الكبس باستخدام معاكس لهذه التحويلات . تم استخدام مقياس الكفاءة لقياس كفاءة النتائج ،وكانت النتائج جيدة .

Abstract

Speech compression is very important in many applications. In this research we applied compression on speech by using wavelet transform and contourlet transform. By taking speech (which is often single dimension) into a two-dimensional array (to be suitable for transferring it to contourlet transformation), and then we applied wavelet transform .Then we make contourlet transform on coefficients of high wavelet. After transforming or storing speech and when using the decompression is applied by using inverse way of those transformations. Applied measuring (SNR, PSNR, NRMSE and Corr.) to test performance of the results, the results are very good.

1_Introduction

Although the Wavelet Transform (WT) is known to be a powerful in many signal and image processing applications such as compression, noise removal, image edge enhancement, and extraction, wavelets are not optimal in capturing the two-dimensional singularities. Found in images and often required in many segmentation and compression applications [1]. In particular, natural images consist of edges that are smooth curves which cannot be captured efficiently by the wavelet transform. Therefore, several new transforms are required for image signals.

- Asst. Prof. \ College of Computers Sciences and Math.\ University of Mosul

** Assist.Lecturer\ College of Computers Sciences and Math.\ University of Mosul

Received:1/10 /2011 ____________________Accepted: 21 /12 / 2011

The Contourlet Transform (CT) is one of the new geometrical image transforms, which can efficiently represent images containing contours and textures. This transform uses a structure similar to that of curvelets, that is, a stage of sub band decomposition followed by a directional transform. In the contourlet transform, a Laplacian Pyramid (LP) is employed for the first stage, while Directional Filter Banks (DFBs) (as in Figure (1 - a) and figure (1 - b), are used in the angular decomposition stage. A comparison between the wavelet scheme and the contourlet shows the improved edge contours of the later. First applying a multiscale transform, followed by a local directional transform to gather the nearby basis functions at the same scale into linear structures. Consider the wavelet transform of a 2-D piecewise functions with a smooth discontinuity curve in addition to contourlet transform (as in figure 2(a)) both wavelet and contourlet transformation However, they are blind to the smoothness of this curve and it is easy to see that there are $O(2j)$ significant wavelet coefficients at the scale 2^{-j} . Comparing the wavelet scheme with the contourlet (as in Figure 2(b)), we see that the improvement of the new scheme can be attributed to the grouping of nearby wavelet coefficients, since they are locally correlated due to the smoothness of the contours. Therefore, we can obtain a sparse expansion [2].

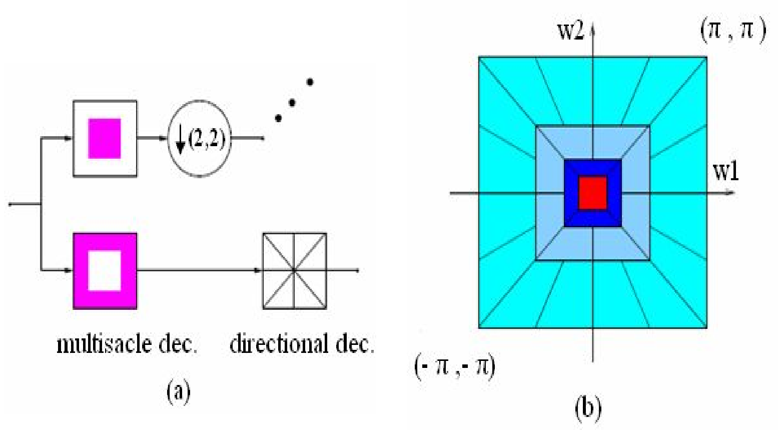


Figure (1) the original contourlet transform. (a) Block diagram. (b) Resulting frequency division

In essence, a wavelet-like transform for edge detection, and then a local directional transform for contour segment detection are applied. The overall result is an image expansion using basic elements like contour segments, and thus are named contourlets and the process is called the contourlet transform (CT) [3].

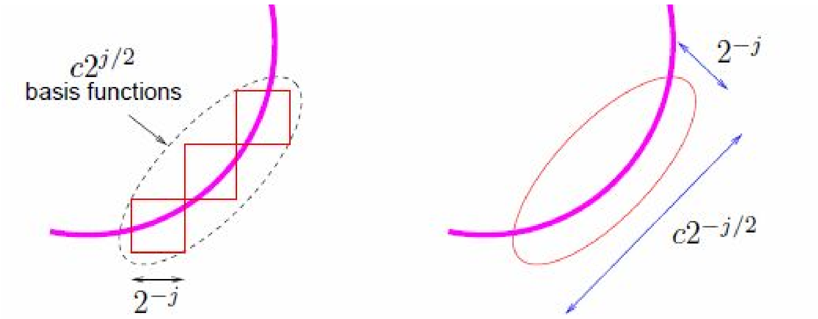


Figure (2) Illustration showing how capture point in wavelet and contourlet transformations.

In this research wavelet transform is used, it has obtained the traditional three high pass bands corresponding to the LH, HLand HH bands. We apply Directional Filter Banks (DFB) with the same number of directions to each band in a given level (*j*).Level of the wavelet transform *J*, we decrease the number of directions at every other dyadic scale when we proceed through the coarser levels (j < *J)*. In this way we could achieve it. Figure (3) illustrates a schematic plot of the wavelet based contourlet transform using 3 wavelet levels and 3 *d L* = directional levels.

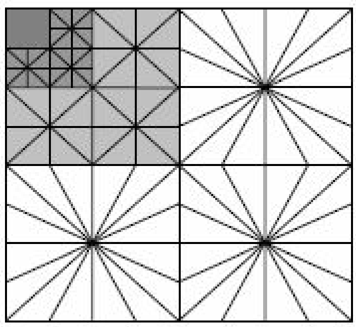


Figure (3) plot of the WBCT using 3 wavelet levels. The directional decomposition is overlaid the wavelet sub bands.

2- Wavelet transform

The fundamental idea behind wavelets is to analyze according to scale. The wavelet analysis procedure is to adopt a wavelet prototype function called an *analyzing* wavelet or *mother* wavelet. Any signal can then be represented by translated and scaled versions of the mother wavelet [4]. The wavelets can be translated about time in addition to being compressed and widened [5]. Speech compression is the technology of converting human speech into an efficiently encoded representation that can later be decoded to produce a close approximation of the original signal. The compressed speech signals using Discrete Wavelet Transform (DWT) techniques. Wavelet analysis is the breaking up of a signal into a set of scaled and translated versions of an original (or mother) wavelet. Taking the wavelet transform of a signal decomposes the original signal into wavelets coefficients at different scales and positions. These coefficients represent the signal in the wavelet domain and all data operations can be performed using just the corresponding wavelet coefficients. Wavelet transforms are broadly divided into three classes: continuous, discrete and multiresolution based [4].

2-1- Discrete wavelet transforms

It is computationally impossible to analyze a signal using all wavelet coefficients, so one may wonder if it is sufficient to pick a discrete subset of the upper half plane to be able to reconstruct a signal from the corresponding wavelet coefficients. One such system is the affine system for some real parameters *a>1*, *b>0.* The corresponding discrete subset of the half plane consists of all the points $(a^m, n\,a^m b)$ with integers $m, n \in \mathbb{Z}$. The corresponding *baby wavelets* are now given as [6, 8]

$$\psi_{m,n}(t) = a^{-m/2}\psi(a^{-m}t - nb).$$

A sufficient condition for the reconstruction of any signal *x* of finite energy by the formula

$$x(t) = \sum_{m\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}\langle x,\, \psi_{m,n}\rangle \cdot \psi_{m,n}(t)$$

is that the functions $\{\psi_{m,n} : m, n \in \mathbb{Z}\}$ form a tight frame of $L^2(\mathbb{R})$ [6, 7].

2-2-Multiresolution discrete wavelet transforms

In any discredited wavelet transform, there are only a finite number of wavelet coefficients for each bounded rectangular region in the upper half plane. Still, each coefficient requires the evaluation of an integral. To avoid this numerical complexity, one needs one auxiliary function, the father wavelet $\phi \in L^2(\mathbb{R})$. Further, one has to restrict *a* to be an integer. A typical choice is *a=2* and *b=1* [6, 9]. From the mother and father wavelets one constructs the subspaces

$$V_m = \operatorname{span}(\phi_{m,n} : n \in \mathbb{Z}), \text{ where } \varphi_{m,n}(t) = 2^{-m/2}\varphi(2^{-m}t - n)$$

And

$W_m = \mathrm{span}(\psi_{m,n} : n \in \mathbb{Z})$, where $\psi_{m,n}(t) = 2^{-m/2}\psi(2^{-m}t - n)$. [6, 7, 8, and 9].

From these one requires that the sequence

$$\{0\} \subset \cdots \subset V_1 \subset V_0 \subset V_{-1} \subset \cdots \subset L^2(\mathbb{R})$$

forms a multiresolution analysis of $L^2(\mathbb{R})$ and that the subspaces $\ldots, W_1, W_0, W_{-1}, \ldots$ are the orthogonal "differences" of the above sequence, that is, W_m is the orthogonal complement of V_m inside the subspace V_{m-1}. In analogy to the sampling theorem one may conclude that the space V_m with sampling distance 2^m more or less covers the frequency base band from 0 to 2^{-m-1}. As orthogonal complement, W_m roughly covers the band $[2^{-m-1}, 2^{-m}]$ [8, 9, 10, and 11].

From those inclusions and orthogonally relations follows the existence of sequences $h = \{h_n\}_{n\in\mathbb{Z}}$ and $g = \{g_n\}_{n\in\mathbb{Z}}$ that satisfy the identities

$$h_n = \langle \phi_{0,0}, \phi_{-1,n} \rangle \text{ And } \phi(t) = \sqrt{2}\sum_{n\in\mathbb{Z}} h_n \phi(2t - n)$$

And

$$g_n = \langle \psi_{0,0}, \phi_{-1,n} \rangle \text{ And } \psi(t) = \sqrt{2}\sum_{n\in\mathbb{Z}} g_n \phi(2t - n).$$

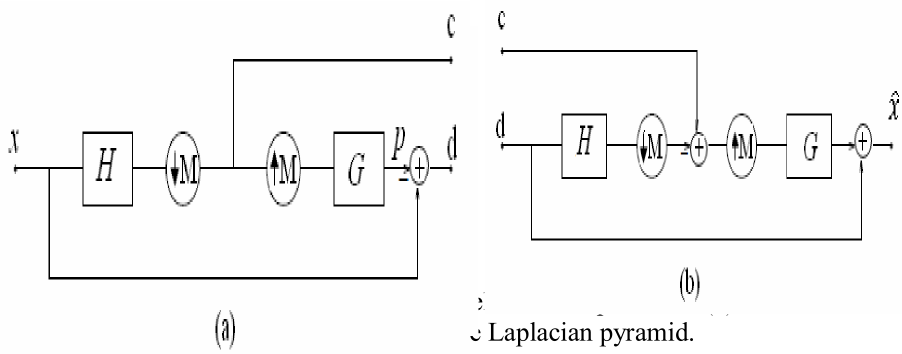
The second identity of the first pair is a refinement equation for the father wavelet φ. Both pairs of identities form the basis for the algorithm of the fast wavelet transform [6, 7, 8, 9, and 10].

3- Contourlet transform

The contourlet transform consists of two major stages: the subband decomposition and the directional transform [3].

3-1- Pyramid frames

One way to obtain a multiscale decomposition is to use the Laplacian pyramid (LP). The LP decomposition at each level generates a down sampled low pass version of the original and the difference between the original and the prediction, resulting in a band pass image. (As in Figure (4)) depicts this decomposition process, where H and G are called (low pass) analysis and synthesis filters, respectively, and M is the sampling matrix. The process can be iterated on the course (sampled down) signal (Low frequencies). Note that in multidimensional filter banks, sampling is represented by sampling matrices; for example, down sampling x[n] by M yields xd[n] = x[Mn], where M is an integer matrix [12].



…struction scheme for … Laplacian pyramid.

3-2- Iterated directional filter banks

Bamberger and Smith constructed a 2-D Directional Filter Bank (DFB) that can be maximally decimated while achieving perfect reconstruction. The DFB is efficiently implemented via an l-level binary tree decomposition that leads to 2^l sub bands with wedge-shaped frequency partitioning as in Figure (5). The original construction of the DFB involves modulating the input image and using quincunx filter banks with diamond-shaped filters [13].A proposed new construction for the DFB that avoids modulating the input image and has a simpler rule for expanding the decomposition tree [3, 12].

His simplified DFB is intuitively constructed from two building blocks. The first building block is a two-channel quincunx filter bank with fan filters (as in Figure (6)) that divides a 2-D spectrum into two directions: horizontal and vertical. The second building block of the DFB is a *shearing* operator. Figure (10) shows an application of a shearing operator where a −45◦ direction edge becomes a vertical edge. By adding a pair of shearing operator and its inverse (“unshearing”) before and after respectively, a two channel filter bank in Figure (9), obtain a different directional frequency partition while maintaining perfect reconstruction. Thus, the key in the DFB is to use an appropriate combination of shearing operators together with two-direction partition of quincunx filter banks at each node in a binary tree-structured filter bank, to obtain the desired 2-D spectrum division as shown in Figure (7).

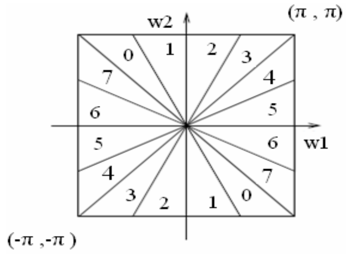


Figure (5) Directional filter bank. Frequency partitioning where l = 3 and there are 23 = 8 real wedge-shaped frequency bands. Sub-bands 0–3 correspond to the *mostly horizontal* directions, while sub bands 4–7 correspond to the *mostly vertical* directions.

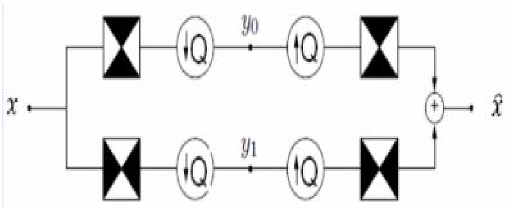


Figure (6). Two-dimensional spectrum partition using quincunx filter banks with fan filters. The black regions represent the ideal frequency supports of each filter. Q is a quincunx sampling matrix.

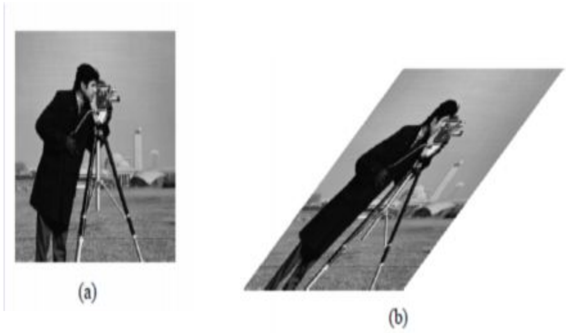


Figure (7) Example of shearing operation that is used like a rotation operation for DFB decomposition. (a) The “cameraman” image. (b) The “cameraman” image after a shearing operation.

In the left part of Figure (8), we see the block diagram of the contourlet transform with two levels of multiscale decomposition, followed by angular decomposition. By using the multirate identities, filter bank redraw into its equivalent parallel form, as shown above in the right part of figure (4) [3].

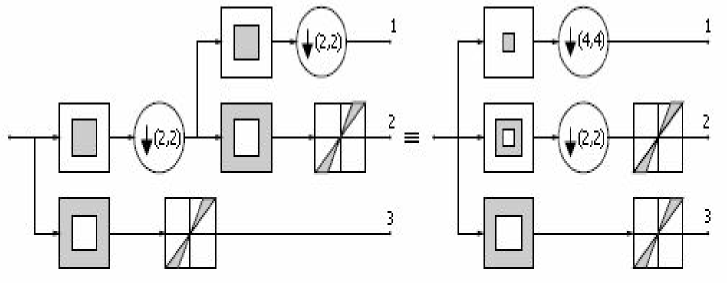


Figure (8). Block diagram of the contourlet transform with two levels of multiscale decomposition. Gray regions represent the ideal pass band support of the component filters. Left: The iterated form. Right: The equivalent parallel form.

4- Performance Measures

A number of quantitative parameters can be used to evaluate the performance of the wavelet based speech coder, in both reconstructed signal quality after decoding and compression .The following parameters are compared [10]

1-Signal to Noise Ratio (SNR) the following formula:

$$SNR = 10log10\left[\frac{\sigma_x^2}{\sigma_z^2}\right]$$

σ_x^2 Is the mean square of the speech signal σ_z^2 is the mean square difference between the original and reconstructed signal.

2-Peak Signal to Noise Ratio**:**

$$pSNR = 10log10\frac{NX^2}{\|x-r\|^2}$$

N is the length of reconstructed signal*, X* is the maximum absolute square value of the signal X and $\|x-r\|^2$ is the energy of the difference between the original and reconstructed signal.

3-Normalized Root Mean Square Error:

$$NRMSE = \sqrt{\frac{(x(n)-r(n))^2}{(x(n)-\mu_{x(n)})^2}}$$

x(n) is the speech signal , r(n) is the reconstructed signal ,and $\mu_{x(n)}$ the mean of the speech signal .

4- Correlation between original signals to the compressed signal.

5-Compression ratio: It is the ratio of the original signal to the compressed signal.

$$CR = \frac{bitpersample(original)}{bitpersample(compressed)}$$

5- Proposed algorithm

The proposed algorithm as shown in fig. (9) we used wavelet transform on speech signal after convert it from 1-D to 2-D. we applied contourlet transform on high coefficients of wavelet, and then make zeroing the high coefficients of contourlet for getting compression. To decompress this signal used inverse way and then listen. Measuring the correlated factor between the original signal and the retrieved ones show that they are closed to each other up to (-+0.90) as shown in table 1, in addition to evaluating the SNR plus the compression ratio gave a good result as seen in Table (1). The proposed algorithm can be used to build a small model to apply it practically to give result.

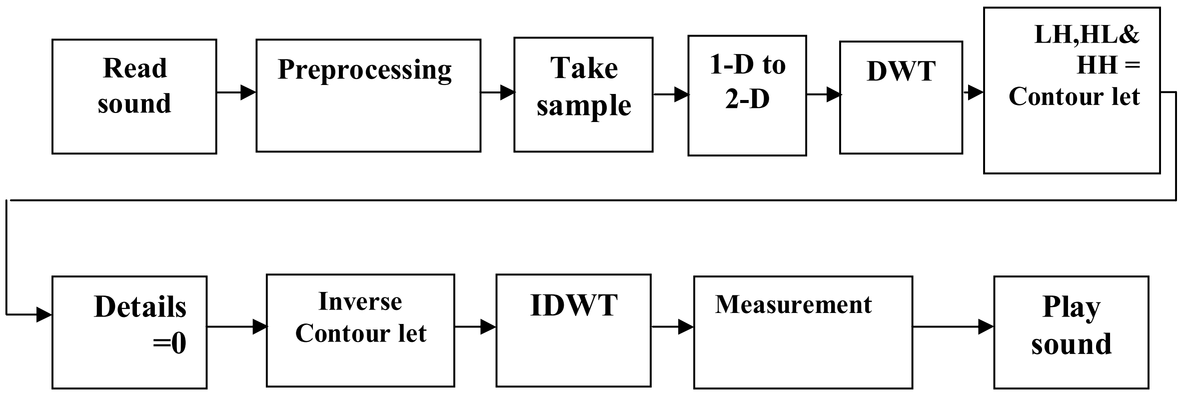


Figure (9) work scheme

6-Practical application

The proposed algorithm is applied with different speech signals and the performance was measured depending on evaluating some factors. The result of correlation factor, Normalized Root Mean Square Error, Peak Signal to Noise Ratio and Signal to Noise Ratio (SNR) between original signal and retriever signal after compression are very good when making HL LH HH contourlet transform and the coefficients of contourlet are set to zero, as we see in Table (1) and figure (10) Compression ratio equals **1: 2.2857** between original and reconstruct signal as we see in figure (11).

Table 1: Perfomace of test signal

	SNR	PSNR	NRMSE	Corr.
speech1	7.4285	16.5202	0.4267	0.9043
speech 2	7.5035	22.4607	0.4227	0.9062
speech 3	11.1338	18.4260	0.2781	0.9605
speech 4	12.0826	15.4380	0.2492	0.9684

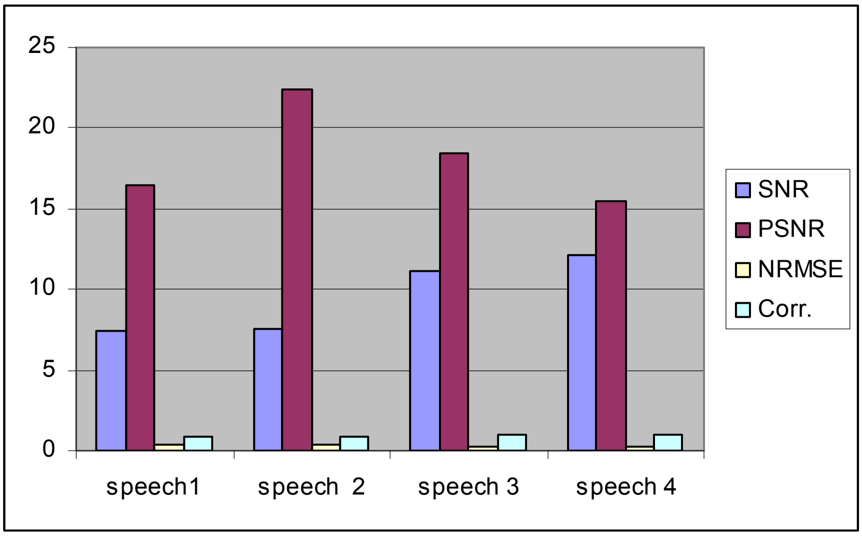


Figure (10): Performance of test signal

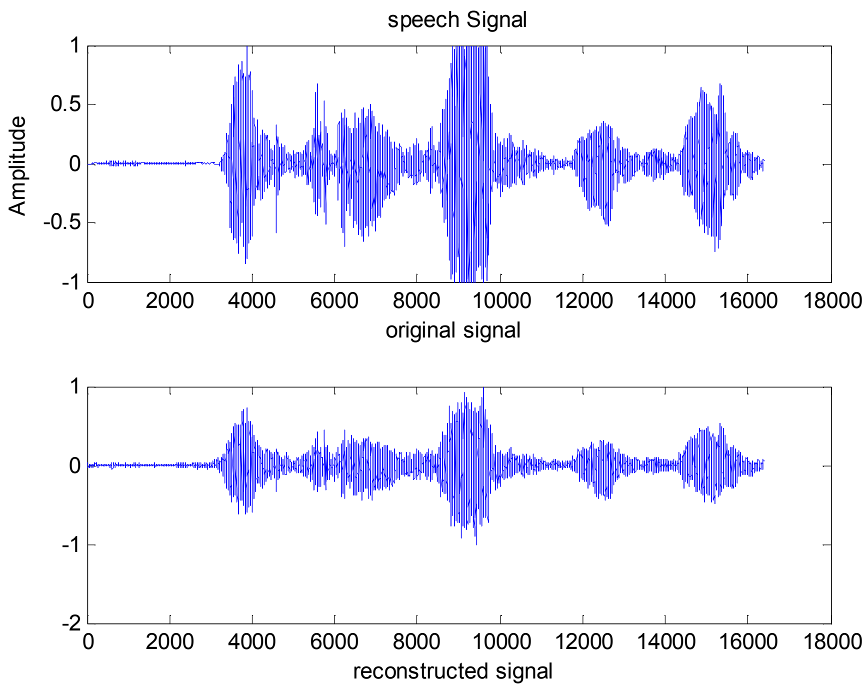


Figure (11): for original and reconstruct signal

7-REFRENCES

[1] Magboub H.M. , Ali N., Osman M.A. , Alfandi **S.A.,** "Multimedia speech compression techniques", Fac. of Eng., Comput. Eng. Dept., Al-Fateh Univ., Tripoli, Libya, **2010**.

[2] Silja M S, Soman K P," A Watermarking Algorithm Based on Contourlet Transform and Nonnegative Matrix Factorization ", International Conference on Advances in Recent Technologies in Communication and Computing **2009**.

[3] Dr Abdul-Jabbar. J. M. And Fathee H. N.," DESIGN AND REALIZATION OF CIRCULAR CONTOURLET TRANSFORM ",Dept. of Computer Eng., College of Eng., University of Mosul, Iraq ,**2010** .

[4] Rao , N. ,2001,"Speech Compression Using Wavelets **",**Thesis, Submitted for the degree of Bachelor of Engineering (Elec Hons) **,**School of Information Technology and Electrical Engineering**,** University of Queensland.

[5] Germano T., 1999," Wavelets in Multiresolution Analysis".

[6] Burrus,C.S.R.A.Gopinath and H.Guo, "Introduction to Wavelet and wavelet Transform", Prentice Hall, 1998.

[7] Daubechies, I.,M.Antoniui,M.Barland and P.Mathieu,"Image Coding using Wavelet Transform", *IEEE* Trns. On Image processing,Vol.1,No.2,PP.205-220, April1992.

[8] Levicky, D.,E.Matus and P.kral,"Wavelet transforms analysis-synthesis algorithms ",J.Elect.Eng. ,Vol.7,No.11-12,PP.281-286,1996.

[9] Mallat,S.G.,"Multifrequency channel decomposition of images and wavelet models",", *IEEE* Trns. On Acoustics,speech,and signal processing,37,2091-2110,1989a.

[10] Mohammed Z.,2007," WAVELET-BAISED COMPRESSION SYSTEM USING LINEAR PREDICTION", master thesis, UNIVERSITY OF BASRAH, College OF ENGINEERING.

[11] Hakim H.,"Encryption using wavelet coded image data",Athesis presented to the University of Basrah in partial fulfillment of the

requirements for the degree of master of science in computer Engineering,2004.

[12] D. D.-Y. Po and M. N. Do, “Directional multiscale statistical modeling of images,” Proc. Of SPIE Conference on Wavelet Applications in Signal and Image Processing X, San Diego, USA, August **2003**.

[13] "Two Dimensional Wavelets", ECE 802, Spring **2010**.