كبس الصوت باستخدام تحويلات المويجة

خليل إبراهيم السيف

كلية علوم الحاسبات والرياضيات

جامعة الموصل

تاريخ استلام البحث:٤/١٠/٢٠٠٥ تاريخ قبول البحث:٥/٣/٢٠٠٧

ABSTRACT

In this research wavelet technique was used for speech compression. It was translated and scaled to groups which represent mother wavelet function.

Applying the adopted algorithm with MATLAB software using Daubechies filters family (with multi-scale for compression level) to compress speech signal for (N=1..9). When the compression level and the distortion level, a best compression level seen at level (5), with different type of wavelet and fixed compression ratio at level (5) with different type of Daubechies filters family.

Minimum signal distortion was seen when the adopted algorithm was applied. And with multi type of wavelet it was seen that the scale had direct proportional relation ship with the compression ratio, also speech compression does not need smoothing for the speech signal as compared to other compression methods

الملخص

تم في هذا البحث استخدام تقنية كبس لإشارات الصوت باستخدام تحويلات المويجة، إذ يتم تمثيلها إلى مجموعة نسخ مقاسه ومترجمة تتكون من دالة المويجة ألام(*mother wavelet*)، إذ يتم اشتقاق المرشح الذي هو مجموعة من المعاملات لإشارة المويجة التي تقوم بعملية كبس الصوت.

تم تطبيق الخوارزمية المعتمدة في البحث باستخدام برامجيات (*MATLAB*) واستخدام مرشحات مويجة دوبيجز ، وعند استخدام قياسات متعددة لمستوى الكبس تم كبس إشارة صوت لعدة مستويات مع اعتماد(*N= 1…9*). عند حساب مستوى الكبس ومقدار التشوه (*SNR*) تمت ملاحظة أحسن مستوى للكبس عند المستوى (*5*) ، مع استخدام أنواع متعددة من المويجات مع بقاء مستوى الكبس مساويا إلى (*5*) لأنواع متعددة من عائلة مويجة ديوبجز (*Daubechies*)

تم الحصول على اقل ما يمكن من تشوهات الإشارة عند تطبيق الخوارزمية. أما عند استخدام مويجات متعددة من هذه المويجة لوحظ انه كلما كان القياس اكبر يكون الكبس أفضل. وكذلك عند عملية كبس الصوت باستخدام تحويل المويجة لا تحتاج إلى تنعيم لإشارة الصوت مقارنة مع طرائق الكبس الأخرى.

1. المقدمة:

يعتبر الصوت هو الوسيلة الأساسية لتبادل المعلومات من شخص لأخر وهو إحدى الوسائل المهمة لتمييز الأشياء خاصة تمييز الأشخاص من أصواتهم. والصوت مجموعة من الترددات يحدث نتيجة جسم مهتز ينتقل عبر وسط ناقل فمثلا الكلام يحدث نتيجة حركة الأوتار الصوتية للحنجرة وكذلك الموسيقى تحدث نتيجة حركة الأوتار للآلة الموسيقية وهكذا ، وصوت الإنسان بالإمكان تغطيته بحزمه ذات عرض4KHZ تحمل معلومات كافيه عن الصوت، وقد أصبح كبس الصوت وتشفيره ضروريا في العديد مِنْ التطبيقاتِ التي تستخدم الزمن الحقيقيَ(real time) مثل الإشارة المستعملة للاستعمال في اتصالات القمر الصناعي النقالة للإرسال الهاتفي الخلوي (mobile) والتي يسمح لمجموعة من المستخدمين بتقاسم الوقت، وكذلك يستخدم في التسجيل الصوتي لـــ videophones أَو عقد المؤتمرات من خلال وسائل الاتصالات teleconferencing وفي تطبيقات أخرى .

إما الكبس: فهو عملية تقليل من حجم البيانات المرسلة عبر وحدات النقل. هناك نوعان من تقنيات الكبس، الأول الكبس بفقدان (Lossy compression) والآخر الكبس من دون فقدان (Lossless compression) ومن طرائق التقنية الأولى طريقة التشفير بطول التنفيذ (Run length Code) التي تسمى ببعض الأحيان التشفير المتكرر والتي تستند إلى حساب الرموز المتكررة ووضع إشارة محلها (flag)، وكذلك طريقة هوفمان (Huffman) التي تستند إلى الاحتمالية الحاصلة في الرموز، وطريقة(LZW) التي تضع سلسلة من الرموز مقابل شفرة معينة، أما من طرائق التقنية الأخرى طريقة التحويل ألموجيwavelet التي هي محور هذا البحث، وطريقة المكمم الإتجاهي (vector..uantization) التي تقوم بعملية ترميز وفك الرموز لكتل معينه توضع في دليل الكتل الصورية(codebook)، وكذلك طريقة الشفرة التنبؤية(Leaner Predicative code) التي تعتمد على عملية الترابط بين النقاط المتجاورة والقيمة المتوقعة (المرجوة) من هذه النقاط أساسها مجموع القيم لتلك النقاط .

إن عملية تمثيل الصوت بالحاسبة أوما يسمى الصوت الرقمي هو إشارة تناظرية مستمرة (continuous analog signal) أو هو عدد غير منته من الترددات(العينات) التي

تكون بمعدل (8000sam/ sec) وكل عينة تتكون من (8bit) وتمتلك معلومات عن الطور والسعة ، وقد جرى تقليديا الاعتماد على تصنيف مشفرات الصوت إلى صنفين:

الأول: مشفرات شكل الموجه, وتتضمن نوعين من المشفرات(coders):

- مشفرات شكل المـــويجة (waveform coders)
- مشفرات تحليل التوليف vocoders

إذ تقوم مشفرات شكل الموجة باستنساخ الشكل الحقيقي للاشارة الصادرة من لاقطة الصوت, وقد أصبحت شائعة بالتوليف النبضي المرمز(PCM Pulse-code-Modulation) في تشفير شكل الموجه .

الثاني: اعتمد على معاملات الإشارة واستخدم بشكل مبكر, من اغلب التقنيات الشائعة لهذا النوع هي التشفير التنبؤي الخطـــي (LPC: Leaner-Predicative-code).

في هذا البحث يتم استخدام تقنية كبس لإشارات الصوت باستخدام تحويلات المويجة، إذ يتم تمثيلها إلى مجموعة نسخ مقاسه ومترجمة تتكون من الدالة ألام للمويجة (mother wavelet) إذ يتم اشتقاق المرشح الذي هو مجموعة من المعاملات لإشارة المويجة التي تقوم بعملية كبس الصوت. بما إن بيانات الصوت غير ثابتة وعشوائية بسبب التغيير الذي يطرأ على عملية إنتاج الصوت من إشارات غير ثابتة ومميزه ذات اتجاهات وتغيرات غير متوقعة. وبما أن المويجة تتميز بمعاملات وترددات ثابتة, عليه تصبح تقنية تحويلات المويجة مناسبة لتشفير إشارات الصوت .

2. تحويلات المويجة(wavelet)

الفكرة الأساسية وراء تحويلات المويجة إمكانية تحليلها طبقا لقياس معـــين, إذ يمكـــن تمثيل دالة المويجة بضرب داخلي(inner product) يؤدي الى تمثيل المويجه من خلال دالـــه رياضيه لفترة زمنيه محدده ذات معدل مساو إلى الصفر**[7],[5]**, والذي يكون ذا فائدة لغـــرض تمثيل البيانات. إذ يمكن تمثيل آية إشارة من خلال مجموعات التغيير والتحويل لمجموعة مـــن دوال القاعدة (base) التي تمثـــل تلك الإشارة**[3]**، والدالة الرئيسية للمويجة التي تسمى بدالـــة الام (mother function) هي:

حيث إن :

$$y_{a,b}(x) = a^{-1/2} y\left(\frac{x-b}{a}\right) \quad \text{.....(1)}$$

a : معامل التوسع

b : معامل النقل

إما دالة المويجة فتتكون من دالتين رئيسيتين:

الأولى: تسمى بدالة المويجة(wavelet function) ويرمز لها بالرمز (Ψ).

الثانية: دالة التعيير(scaling function) ويرمز لها بالرمز (Φ).

هناك نوعان من دالة المويجة الأولى تسمى دالة المويجة المستمرة (continues wavelet transform) التي معادلتها الأساسية هي[3]:

$$C_{(s,u)} = \int_{-\infty}^{\infty} f(t)\sqrt{s} * y\left[\frac{(x-u)}{a}\right] \quad \text{.....} \quad \text{.....} \quad \text{...(2)}$$

حيث أن :

S : معامل التوسع

u : معامل النقل

X : الإشارة الداخلة

علية يمكن ملاحظة : إن قياس مدى الإشارة يتغير مع الزمن فعندما يكون التغيير عاليا تتمثل المويجة بالمعلومات العامة(معلومات التردد الواطئ) حول الإشارة ،أما عندما يكون التغير واطئا فان المويجه تمثل المعلومات التفصيلية (معلومات التردد العالي) حول الإشارة ويمكن ملاحظة العلاقة مع التردد والزمن على أنها علاقة عكسية كما موضحة بالشكل (1).

إذ تتم ملاحظة إن مساحة النافذة ثابتة ولكن أبعادها متغيرة فعند الترددات العالية تكون النوافذ ذات عرض ضيق (يعود إلى زمن وضوحية جيد) وارتفاع عال(يعود إلى تردد ووضوحية ضعيفين)على العكس بالترددات الواطئة[7]

الشكل(1) شكل المويجة

أما النوع الآخر من تحويلات المويجة فأنها تقسيم المويجة المستمرة إلى مجموعة من الكتل أو العينات(block or sample) التي تسمى تحويلات المويجة المتقطعة (discrete wavelet transform) والتي تستخدم لتحليل الصوت أو الصور ويعبر عنها بالدالة الآتية:

حيث إن :

$$f(x)=\sum_{j=1}^{L}\sum_{k=-\infty}^{\infty}d(j,k)y\,(2^{-j}t-k)+\sum_{k=-\infty}^{\infty}a(L,k)f\,(2^{-L}t-k).....(3)$$

$a(L,K)$ تمثل المعلومات التقريبية

$d(j,K)$ تمثل المعلومات التفصيلية

ومن المعادلة أعلاه يتم اشتقاق المعادلتين الآتيتين:

$$a(L,k)=\frac{1}{\sqrt{2^{L}}}\int_{-\infty}^{\infty}f\,(t)f\,(2^{-L}t-k)dt.....(4)$$

أما دالة المعلومات التفصيلية والتي تمثل مرشح التردد العالي فهي :

$$f\,2(t)=\sum_{k=-\infty}^{\infty}d(l,k)f\,(2^{-l}t-k).....(5)$$

إذ يلاحظ من المعادلتين أعلاه إن دالة المويجة تمثل مرشح التمرير العالي ودالة التعيير تمثل مرشح التمرير الواطئ .

وهناك عدة عوائل من المويجات التي تستخدم في مجالات مختلفة ومن هذه العوائل [5]:

1 – مويجة كوفلت(coflict wavelet)

٢- مويجة ديوبجز(Daubechies wavelet)

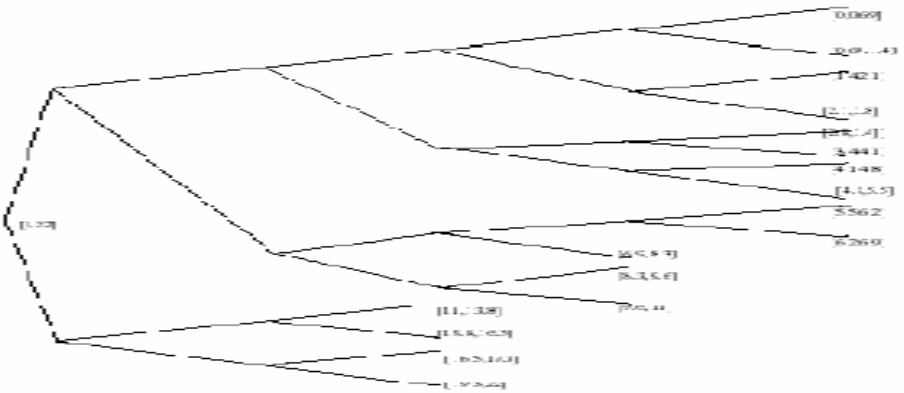
٣- مويجة هار(Haar wavelet)

٤- مويجة ساملت (synmelet wavelet)

وكل من هذه العوائل مرشح معين وكل مرشح له معاملات خاصة به وكذلك يمكن أن تصنف هذه العوائل على أساس العزوم المتلاشية التي تستند إلى مجموعة من الإضافات على العلاقات الرياضية لمعاملات هذه المرشحات وفي اغلب الأحيان تكون محدده إذ أن درجه اندماج إشارات العزوم تكون عالية. ففي هذا البحث تم استخدام مرشح ديوبجز من ضمن هذه العوائل في كبس الصوت الذي يمتلك عددا من المعاملات التي يمكن تكرارها واستخدامها في قياس معين[6].

٣ -كبس الصوت [4] [9][8]

تنطلق الفكرة من وراء كبس الصوت باستخدام المويجة من معلومات الصوت التي هي (طاقة ونغمه) إذ إن عمل المويجات هو تفكيك الإشارة إلى مجموعة من الحزم بترددات مختلفةِ، وتنفذ هذه العملية حسب اختيار وظيفةِ المويجة ، إذ يتم حساب دالة المويجة المتقطعة (DWT: Discrete Wave Transformation) التي تقوم بالكبس بالاستناد إلى عدد صغير من معاملات التقريب(التي يكون لها مستوى مناسب لتفكك المويجة تعتمد على نوع الإشارة آو مستوى الطاقة إذ يتم اختيار المشفر الذي يَستعملُ في كبس الصوت باستخدام تحويلات المويجة المتقطعة (DWT)، كما في الشكل (2)



الشكل (2) الشجرة المهيكلة لتحويلات المويجة حسب ترددات معينة

إذ تقوم المويجة بتحويل الإشارة إلى حزم منفصلة وتَعطي حرية القرار في اختيار الترددِ طبقاً لحاجاتِ التطبيقِ المحدّدةِ. يلاحظ من الشكل(2) إن المعلومات التفصيلية للصوت تتمثل بالترددات العالية كما في حالة الشجرة الثنائية(Binary Tree).

عند استخدام معاملات المويجة المتقطعة التي تستخدم مويجة دوبيجز تستخدم في عملية إمرار المرشحات التردد العالي والوا طئ على إشارة الصوت التي يتم تفككها إلى حزم ، ويمكن أن نعبر عن دالة هذه المويجة رياضيا بالمعادلات الآتية :

$$\Phi(t) = \sum_{k} c(k)\Phi(2t-k) \ldots\ldots\ldots\ldots\ldots (6)$$

إن المعادلة (6) تسمى بدالة التعيير التي تمثل المرشح الذي يقوم بعملية التنعيم للإشارة والحفاظ على الترددات الواطئة لها ، أما المعادلة (7) فتمثل دالة المويجة في قياس دالة التعيير ، في حين أن المعادلة (8) هي التي تحقق خاصية التعامد بين دالتي القياس والتحويل.

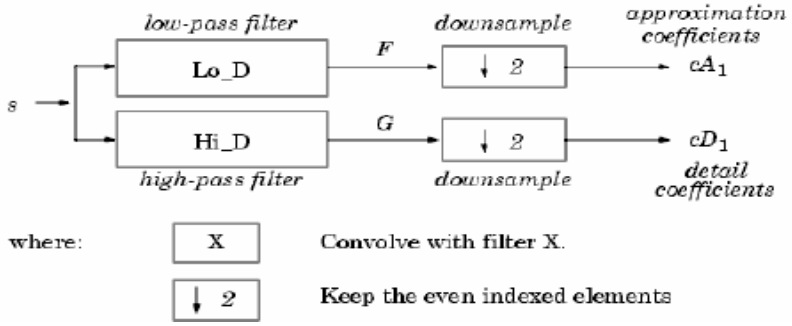
$$j(t)=\sum_{k}(-1)^{k}c(1-k)f(2t-K).....(7)$$

$$\sum_{k} c_k c_{k-2n} = 2d_{0,m} ....(8)$$

الملاحظ من المعادلتين (7 و 8) إن كلا من (c(k),c2, c(n-1)) هي التي تمثل الومضات (impulse) المطلوبة في مرشح التمرير الواطئ التي يكون طوله (2N). أما مرشح التمرير العالي فتحسب معاملاته مناظرة لمعاملات مرشح التمرير الواطئ من المعادلة(9).

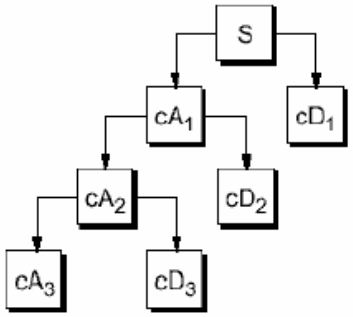
$$g_k = (-1)^k c(1-k)..... ....(9)$$

للحصول على صوت مكبوس يتم إدخال أشارات الصوت (S) وعند استخدام خوارزمية المويجة المتقطعة يتم تفكك الإشارة باستخدام الحالة الأولى أو مستوى واحد من الكبس نحصل على مجموعتين من المعاملات هي (cA1) التي تمثل إشارة معاملات التردد الواطئ والأخرى معاملات (cD1) التي تمثل معاملات التردد العالي ، إذ تتم معالجة إشارة الصوت مع المرشحين باستخدام اللف الرياضي(convolution)للحصول على مستوى واحد من الكبس ، كما نلاحظ بالشكل الاتي:



الشكل (3)

ومن الممكن إمرار نفس المرشحين على الإشارة الناتجة من إمرار مرشح التمرير ألواطئ مرة أخرى للحصول على مستويين آخرين و بالإمكان التوصل لأكثر من مستوى ممكن من خلال تكرار العملية أكثر من مرة كما مبين بالشكل (4)

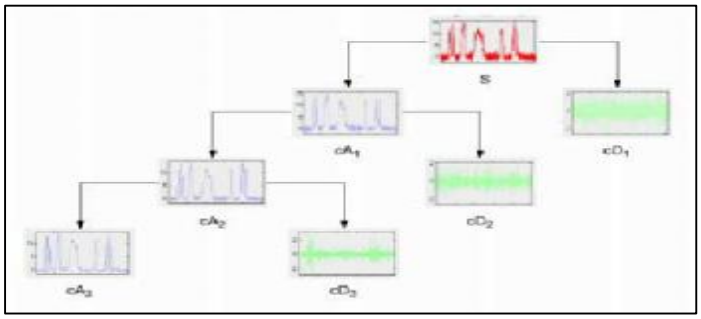


الشكل (4) الكبس لأكثر من مستوى

ممكن حساب عدد معاملات كل من (cA1,cD1) بالصيغة الاتية :

$$floor\ (\frac{n-1}{2}) + n \ldots\ldots a \ldots\ldots\ldots \ldots (10)$$

الشكل (5) يمثل المخطط الانسيابي الذي يوضح شكل الإشارة عند كل مستوى[2]



الشكل (5) يوضح شكل الموجة بعد الكبس

4- الخوارزمية المطبقة في البحث:

أدناه الخطوات المتبعة في تنفيذ الخوارزمية المقترحة :

1) اختيار نوع المويجة المستخدمة ومستوى الكبس.
2) إدخال إشارة الصوت المطلوب كبسها ويكون نوع الإشارة (OD) أو (wav).
3) حساب العتبة التي تسمح بمستوى الكبس. تحسب حسب القانون(المعادلة 11)

$$n = \frac{m}{(j+2-i)^a} \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

حيث أن α هي معاملات الكبس في هذا البحث تم الاعتماد على (α =1.5 لكونه حقق أفضل نتائج)

j : مستوى الكبس

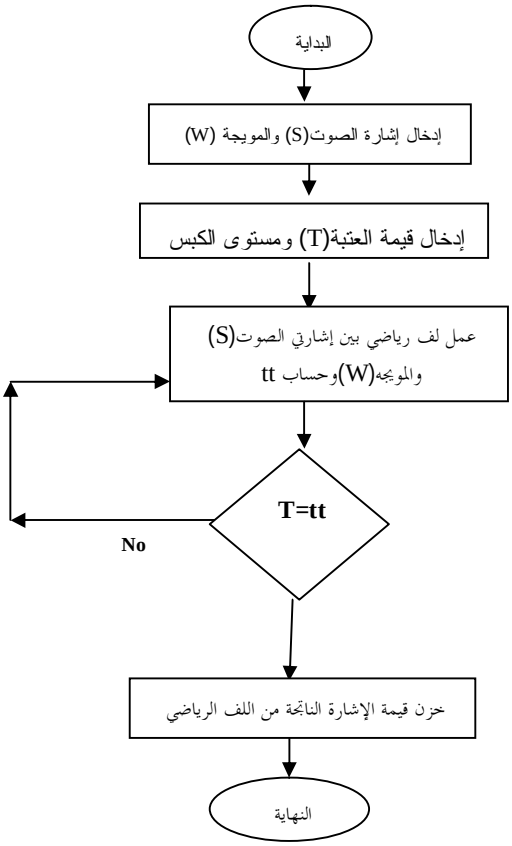
i: عدد المعاملات التي تحفظ عند ذلك المستوى وتتزايد بالفترة (J---1)

أما قيمة التوزيع(m) فتحسب بالقانون (L*1.5).

حيث أن (L) يمثل طول المعاملات التقريبية للمستويات العالية للإشارة.

4) عمل لف رياضي بين مرشحي المويجة والصوت وتستمر عملية اللف على آن لا تتجاوز قيمتها المحسوبة بالخطوة (3).

5) خزن نتيجة اللف الرياضي الذي يمثل الصوت المكبوس وحساب نسبة التشوة.

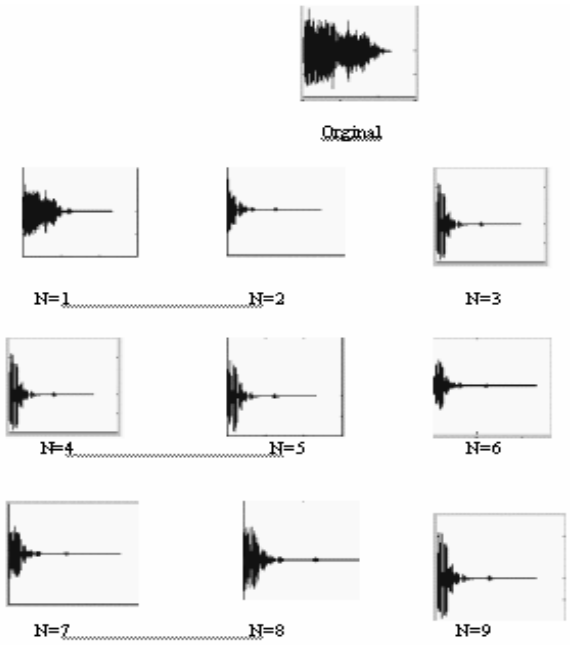


الشكل (6)

المخطط الانسيابى لخوارزمية كبس الصوت باستخدام تحويلات المويجة

5- النتائج

تم تطبيق الخوارزمية المذكورة انفا باستخدام نظام(matlab) واستخدام مرشحات مويجة دويبجز ، وعند استخدام قياسات متعددة لمستوى الكبس تم كبس إشارة صوت لعدة مستويات مع الاعتماد على(N= 1...9) للحصول على إشارات لصوت مقطع موسيقي ويمكن ملاحظة مستويات الكبس بالشكل(7).



الشكل (7)

المخطط الانسيابي لخوارزمية كبس الصوت باستخدام تحويلات المويجة

أما الجدول (1) فيبين مستوى الكبس ومقدار التشوة (Signal to Noise Ratio) الذي تم الحصول عليه.

الجدول (١)قيمة الكبس لمستويات متعددة

N(compression level)	Compression rate	SNR(signalto Noise Ratio)
1	**0. 93**	**20.0176**
2	**0. 86**	**21.5743**
3	**0. 78**	**21.3181**
4	**0. 71**	**22.1721**
5	***0. 63***	***23.4351***
6	**0. 56**	**23.2999**
7	**0. 49**	**22.4267**
8	**0. 42**	**21.7135**
9	**0. 35**	**21.5912**

من الجدول (1) تمت ملاحظة إن أحسن مستوى للكبس يكون عند المستوى (5). يمثل الجدول(2) مستوى الكبس لحالات متعددة عند استخدام أنواع مختلفة من المويجات مع بقاء مستوى الكبس مساويا إلى (5) لأنواع متعددة من عائلة مويجة ديوبجز (Daubechies).

الجدول (2) يمثل الكبس لأنواع مختلفة من المرشح

Dubachies type	Compression rate	SNR(db)
Db2	0. 97	19.9698
Db4	0. 94	19.9698
Db6	0. 89	20.02822
Db8	0. 85	24.3818
Db10	***0.79***	***25.316***
Db12	0. 76	23.886
Db14	0. 72	22.5985
Db16	0. 68	21.1755
Db24	0, 51	20.8704

مما سبق تبين من الجدولين أعلاه انه عندما يتم كبس الصوت باستخدام تحويلة المويجة فأنة يتم الحصول على أفضل كبس للصوت باستخدام قياسات متعددة لمويجة ديوبجز كما يلاحظ انه يتم الحصول على أفضل مستوى للكبس عندما يكون القياس (5) ،إذ ظهر اقل ما يمكن من تشوهات على الإشارة، وعند استخدام مويجات متعددة من هذه المويجة (يلاحظ من الجدول (2)) انه كلما كان القياس اكبر (أي عدد المعاملات أكثر) يكون الكبس أفضل. وعند حساب مقدار التشوه بالكبس (SNR) من الملاحظ الحصول على كبس مثالي عندما تكون عملية الكبس مساوية ل (10db) وهو أفضل مستوى للكبس إذ أن التشوه اقل ما يكون. كذلك عند كبس الصوت

بالمويجة فإنها لا تحتاج إلى عملية تنعيم لإشارة الصوت مقارنة مع الطرائق الأخرى للكبس التي تحتاج إلى مرشحات لتحسين الإشارة بعد كبسها مما يؤدي إلى إضافة خوارزميات أخرى لتحسين النتائج، أما عند الاعتماد على المويجة فان دالة التعيير هي التي تقوم بعملية تحسين الإشارة .

المصادر

[1] Agbinya, J.I. 'Discrete Wavelet Transform Techniques in Speech Processing'.*IEEE Tencon Digital Signal Processing Applications Proceedings*, IEEE, New York, NY, 1996, pp 514-519.

[2] Blahut, R.E. '*Digital Transmission of Information*', Addison-Wesley, NewYork, 1990.

[3] Chen, A . et. al. 'Discrete Wavelet Transform for Audio Compression',http://is.rice.edu/~welsh/elec431/wavelets.html, (current July., 16, 2001).

[4] N.R. Chong, I.S. Burnett, J.F. Chicharo and M.M. Thomson, ihUse of the Pitch Synchronous Wavelet Transform As a NewDecomposition Method for WI,lo *Proc. Int'l Conf. Acoustics, Speech and Signal Processing*, Vol 1, IEEE, New York, NY, 1998, pp. 513-516.

[5] Fgee, E.B. et. al.,'Comparing Audio Compression using Wavelets with other Audio Compression Schemes', *IEEE Canadian Conference on Electrical and Computer Engineering*, IEEE, Edmonton, Canada, 1999, pp.698-701.

[6] Gersho, A., Speech Coding,l. *Digital Speech Processing*, A.N. Ince, ed., Kluwer Academic Publishers, Boston, 1992, pp. 73-100.

[7] Graps, A., 'An Introduction to Wavelets' *IEEE Computational Sciences&Engineering*,http://www.amara.com/IEEEwave/IEEEwavelet.html *(current Mar.15, 2001).*

[8] Haykin, S. '*Communication Systems*', John Wiley & Sons, NewYork,

[9] Holmes, J.N., '*Speech Synthesis and Recognition*', Chapman Hall, London, 1988

[10] Viswanathan, V. W. et. al. 'Real-Time Implementation of a Wavelet-Based Audio Coder on the T1 TMS320C31 DSP Chip,.*5th International Conference on Signal Processing Applications & Technology (ICSPAT)*, Dallas, TX, Oct. 1994.