

## كبس ملفات الكلام وتشفيرها باستخدام الشبكة العصبية ذات الانتشار العكسي

نجلاء متي اسحق سفر<sup>(١)</sup>

### الملخص

يعد كبس الكلام حقلاً مهماً جداً في تطبيقات شبكات الانترنت، أو نقل المعلومات خلال شبكات الاتصالات الرقمية والشبكات الهاتفية، لذا فقد تم في هذا البحث دراسة أسلوب الكبس باعتماد الشبكات العصبية لتطبيقها في كبس بيانات الكلام، في حين شاع استخدام هذه الطريقة مؤخراً في كبس الملفات الصوتية. وقد تم التركيز على الملفات الصوتية التي تتضمن الكلام فقط أي: لا تحتوي أصواتاً أخرى كالموسيقى أو أصوات للسيارات أو للحيوانات وإلى آخره. في هذا البحث، استخدمت شبكة الانتشار العكسي (Back Propagation Neural Network BPNN) لعملية كبس إشارات الكلام. تم إجراء عدة تجارب تختلف فيما بينها من ناحية هيئة البيانات المقدمة للشبكة. وإن أعلى نسبة كبس تم التوصل إليها (1:10) من البيانات الأصلية. الملفات المكبوسة التي نحصل عليها تعد ملفات تشفير للكلام، إذ إنها تحوي بيانات ضئيلة نسبة إلى الملف الأصلي وكذلك ليست ذات فائدة عند تشغيلها في حالة تعرضها للسرقة خلال عملية النقل في الشبكات، إذ لا يستطيع فكها إلى أصلها إلا بامتلاك مصفوفة الأوزان الخاصة بها، لذلك فالفائدة المرجوة من هذا الكبس تكمن في تحقيق أهداف مزدوجة، إذ تُعد أمانة الكلام من الأهداف المهمة لمستخدمي أنظمة الاتصالات التي تستخدم الكلام. وللحصول على مستوى مرغوب من الأمانة فمن الضروري إضافة مرحلة تشفير لإشارة الكلام قبل إرسالها.

(١) كلية الزراعة والغابات / جامعة الموصل.

### Abstract

Speech compression is a very important field of internet application or to transfer through network and telecommunication. This research studies speech compression using neural network, which is commonly used with images files compression. It has been focused on the files that include speech only and do not contain other sounds like music or car sound, or animal sound etc. In this research it has been used Back Propagation Neural Network BPNN for speech signals compression. Several experiments have been performed which differ in configuration of data that entered to the network. The highest compression ratio obtained is (1:10) from the original data. The compression files obtained represent as encipher files because they have little data with compared to original file, and also they might be not useful if it has been stolen through transmitting operation on the network, they be able to be decompressed to the original without their weights matrix can not, so the promising benefit from this compression is to achieve double aims. Speech security is an important goal for users of many speech communication systems. To obtain a desired level of security an encryption scheme should be added for speech signal before transmission.

### المقدمة

الكلام فعل حي ديناميكي وهو الأساس في الحدث اللغوي (فرغلي، ١٩٨٩)، وعلى وفق تصريح بيل جيتس مدير شركة مايكروسوفت عن الكلام بأنه المفتاح التقني لتمكين المستخدمين من الاتصال فيما بينهم طبيعياً مع الماكنة (Cawley, et al., 2001)، فالكلام الوسيلة اللغوية الوحيدة المستخدمة عالمياً للاتصال بين أفراد الجنس البشري. وما هو إلا استحداث موجات صوتية بواسطة الحركة الإرادية للتركيب التشريحي في نظام توليد الكلام لدى الإنسان لتنقل المعلومات من المتكلم إلى السامع، (Abdul\_Majeed, 2003).

ينتج الصوت الإنساني وغير الإنساني نتيجة اهتزاز ذرات جسم ما بتأثير قوة ما. وتنقل هذه الاهتزازات أو الذبذبات من ذرات الجسم المهتز إلى ذرات الهواء أو ذرات سلك التلفون أو طبقات الجو الأثرية فتهتز بدورها. وتوصف الذبذبة بتحديد أمرين، الزمن الذي تستغرقه ويسمى (الفترة period) والبعد بين نقطة بدء الحركة ونقطة ارتدادها ويسمى اتساع الذبذبة amplitude وتوصف الموجة بعدد الذبذبات التي تكونها في الثانية ويسمى بدرجة الموجة frequency وبمتوسط اتساع ذبذباتها وهو يسمى اتساع الموجة (أيوب، ١٩٨٩).

يتم تمثيل إشارة الكلام عادة بصيغة رقمية تكون سلسلة من الخانات الثنائية، لأجل تطبيقات الخزن والنقل، ومن المستحب كبس إشارة بتمثيلها بخانات قليلة قدر الامكان، مع الحفاظ على نوعيتها بشكل محسوس ومدرك. في حالة كبس الحيز الضيق في إشارة الكلام الرقمي، يتم تعيين إشارة الكلام بنسبة ٨٠٠٠ عينة بالثانية. على نحو نموذجي، تمثل كل عينة بثمانية بتات. وهذا يتوافق مع معدل البتات في حالة ٦٤ كيلوبت بالثانية. ولتحقيق المزيد من الكبس سوف يكون ذلك على حساب النوعية (MathWorks,2004).

إن فهم الأنظمة الحديثة في الكلام يدمج مجالات تقنيات مختلفة من معالجة الإشارة وتمييز الأنماط واللغات الطبيعية واللغويات إلى نطاق إحصائي موحد. هذه الأنظمة لها تطبيقات على مدى واسع في مسائل معالجة الإشارة، تمثل ثورة في معالجة الإشارة الرقمية (Picone, 1996).

### كبس إشارة الكلام

إن كبس إشارة الكلام له عدة تطبيقات عملية. احد الأمثلة يتمثل في تقنية الخليوي الرقمي، عندما يتشارك عدة مستخدمين في حزمة مدى الترددات. فالكبس يتيح لعدة مستخدمين أن يتشاركوا في النظام خلافا لما هو مؤهل.

الهدف الرئيس لتقنية أي كبس هو تحقيق أقصى حد لتقليل حجم البيانات مع بقاء حفظ المميزات المهمة لتشكيل الإشارة عند تكوينها. كبس البيانات هي عملية اكتشاف وإلغاء الإسهاب الموجود في مجموعة البيانات، ويوجد الإسهاب في الإشارة الرقمية عندما تكون عينات الإشارة المتجاورة معتمدة إحصائياً و/أو مدى الإشارة المكتملة لا يحدث باحتمالية متساوية.

الخطوة الأولى نحو كبس بيانات الكلام هي اختيار اقل معدل أعتيان. بناءً على ذلك، يمكن أن يحقق كبساً إضافياً لإشارة الكلام من خلال استخدام الخصائص الإحصائية المعروفة للإشارة.

تسمح تقنية كبس الكلام بالعمل مع الأجهزة محدودة الذاكرة، أو ذات انخفاض في عرض الموجة (Weinschenk and Barker, 2000).

تحدد الخصائص الآتية لإجراء أية مقارنة بين تقنيات كبس إشارات الكلام:

١- تردد عينات الإشارة (Signal sampling frequency ( $f_0$ )) يستعمل في محول التناظري/الرقمي لتحويل إشارة الكلام إلى حالة رقمية، تردد العينات ويمكن أن يكون بترددات مختلفة على وفق الغرض (Kocyigit et al.,2002).

٢- عدد البتات في العينة الرقمية والتي تبين دقة الخزن وكثافته لبيانات إشارة الكلام ويمكن ان تكون ٨ أو ١٦ بت (Kocyigit et al.,2002).

٣- نسبة الكبس (Compression Ratio(CR)) نسبة الكبس هي مدى تقليل البيانات الذي يحصل عليه نتيجة لعملية الكبس، في نظام كبس إشارة الكلام باستخدام الشبكة العصبية ذات الانتشار العكسي، تعرف نسبة الكبس بنسبة البيانات التي تغذى إلى خلايا طبقة الإدخال إلى البيانات الخارجة من خلايا الطبقة المخفية. يمكن ان يعبر عن نسبة الكبس بالمعادلة (1) (قدو،٢٠٠٤):

$$CR = \frac{\text{The number of samples after compression}}{\text{The number of samples before compression}} \times 100\% \quad \dots\dots\dots (1)$$

ويمكن أن يعبر عنها كنسبة عدد العينات قبل الكبس إلى عدد العينات بعد الكبس، (Kocyigit et al.,2002).

كذلك يمكن أن تحسب نسبة الكبس (CR) بالمعادلة (Allaf,1999) :

$$CR = (1 - (N_h/N_i)) \times 100\% \quad \dots\dots\dots (2)$$

إذ أن  $N_i$  تمثل عدد الخلايا في طبقة الإدخال. و  $N_h$  هي عدد الخلايا في الطبقة المخفية.

٤- مؤشر الأداء (PRD, "Percent Root Mean Square Difference) Performance index

PRD : جذر فرق مربع المعدل المئوي للإشارة ويعد باراميتراً آخر مهماً لأية خوارزمية كبس وتبين القيمة الصغيرة للـ PRD نجاح الخوارزمية وتحسب كما في المعادلة (3) :

$$PRD = \sqrt{\frac{\sum_{i=1}^n [X_{org}(i) - X_{rec}(i)]^2}{\sum_{i=1}^n [X_{org}(i)]^2}} \cdot 100 \dots\dots\dots (3)$$

إذ إن  $X_{org}$  و  $X_{rec}$  هما عينات لسلاسل البيانات الأصلية والمعاد تكوينها .

### استخدام الشبكات العصبية في الكبس

لقد تزايد استخدام الشبكات العصبية الاصطناعية واختبارها والتي تأخذ بنظر الاعتبار حل العديد من المسائل والتطبيقات التي تحتاج إلى دقة حسابات عالية. والفائدة الأساسية للشبكات العصبية تكمن في قابليتها على المعالجة المتوازية، وقابليتها على استخلاص دالة التحويل المطلوبة من الأمثلة.

لقد تبين أن الإمكانيات الثلاث الآتية تعد مهمة في نماذج الشبكة العصبية لتوظيفها كمكينة تعلم وهذه الإمكانيات هي :

١. قابلية كبس البيانات لتخطيط مجموعة من بيانات أصلية إلى حيز المميزات لتقليل الأبعاد.

٢. قابلية العمومية لكبس بيانات غير متعلم عليها فضلاً عن البيانات المتعلم عليها. (Gary, et al, 1993)

٣. الشبكات العصبية لكبس الصور والإشارات يمكن أن تحدد من تأثيرات أخطاء القناة وأخطاء الدوائر. (Carrato, and Marsi, 1993)

### مسألة التشفير

يعد كبس الصور والإشارات نوعاً من مسائل التشفير (الترميز). إن الشبكات العصبية تعطي حلاً لمسألة تطابق المخططات من مجاميع إدخلات الإشارة. تُبنى الشبكة لانجاز هذا المخطط من خلال قناة ضيقة، لجعلها تطور تشفيراً كفواً في تلك القناة. يوجد مفهوم مهمان في ذلك: (أ) تقوم الشبكة بتطوير تمثيل مضغوط للبيئة (المحيط) و (ب) على الرغم من ان الخوارزميات المستخدمة طورت كنموذج تعلم بإشراف، هذه المسألة يمكن ان تؤخذ بنظر الاعتبار من دون إشراف وذلك لان الإخراج للشبكة هو نموذج الإدخال نفسه – نظام ذاتي- تم تنظيمه لتشفير البيئة. (Cottrel, 1989)

### التطبيق العملي

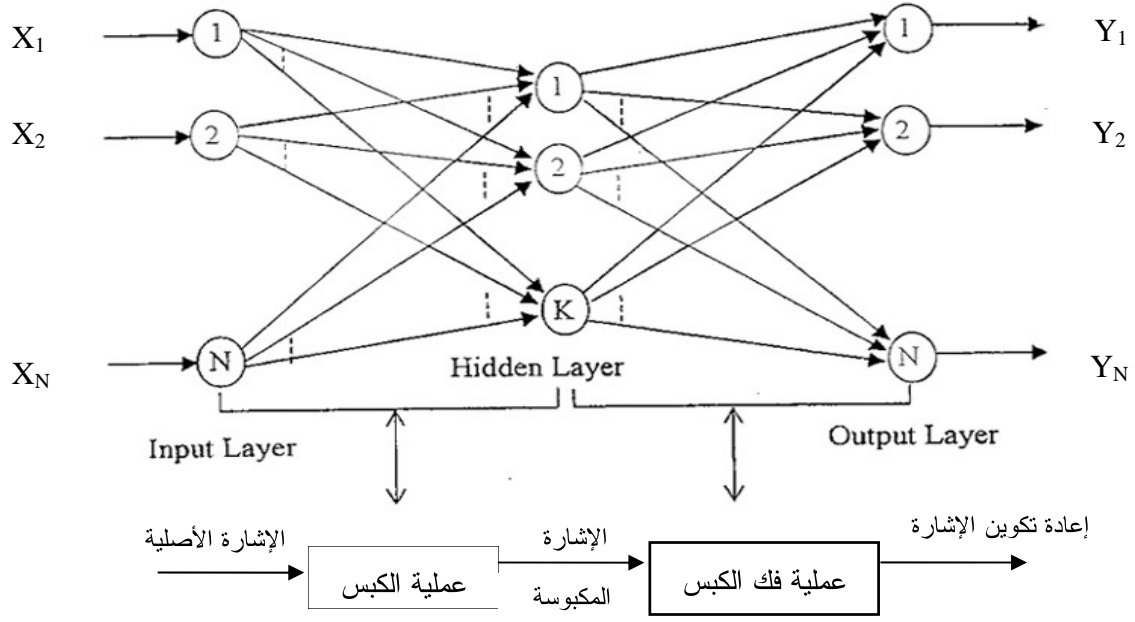
في هذا البحث، تم استخدام ثمانية ملفات كلام ذات أحجام مختلفة كل منها يحتوي على كلام لشخص ما، وصيغة التشكيل (format) التي تم التعامل معها هي (Windows)

Audio- Visual واختصاراً (WAV)، كونها من الصيغ المستخدمة بشكل واسع في بيئة النوافذ (Windows). (Kientzle, 1998). والجدول (١) يبين الخصائص للملفات المستخدمة في البحث.

الجدول (١) معلومات عن الملفات المستخدمة في البحث

| number of bits per sample | sample rate ( $f_0$ ) in Hertz | حجم الملف الأصلي | ت |
|---------------------------|--------------------------------|------------------|---|
| 8                         | 11025                          | 117 KB           | 1 |
| 8                         | 11025                          | 1.43 MB          | 2 |
| 8                         | 11025                          | 890 KB           | 3 |
| 8                         | 11025                          | 951 KB           | 4 |
| 16                        | 11025                          | 314 KB           | 5 |
| 16                        | 11025                          | 537 KB           | 6 |
| 16                        | 11025                          | 785 KB           | 7 |
| 16                        | 11025                          | 947 KB           | 8 |

استخدمت شبكة الانتشار العكسي (Back Propagation Neural Network BPNN) لعملية كبس إشارات الكلام. طورت معمارية BPNN باستخدام الفكرة في مسألة التشفير. تتألف الشبكة العصبية لكبس الكلام من ثلاث طبقات من الخلايا (إدخال ومخفية وإخراج) كما في الشكل (١).



شكل (١) شبكة الانتشار العكسي لنظام كبس الإشارات

تم إجراء عدة تجارب تختلف فيما بينها من ناحية هيئة البيانات المقدمة للشبكة، والغرض منها إيجاد أفضل أسلوب يحقق أعلى نسبة كبس واسترجاع للبيانات بشكل مقبول. ففي التجربة الأولى تم تقطيع إشارة الكلام إلى مقاطع غير متداخلة يحتوي كل منها على ٥٠٠ عينة يمثل كل منها الإدخال للشبكة العصبية، لذلك كان تنظيم الشبكة يحتوي على ٥٠٠ خلية في طبقة الإدخال، واحتوت الطبقة المخفية على ١٠٠ خلية، أما طبقة الإخراج فقد احتوت على ٥٠٠ خلية تمثل طبقة الإدخال لغرض الحصول على نسبة كبس أعلى ١:٥.

أما التجربة الثانية فقد تم تقسيم إشارة الكلام للملفات نفسها المستخدمة في التجربة الأولى إلى مقاطع غير متداخلة كل مقطع يحتوي على ٢٠٠ عينة. وكان تنظيم الشبكة كالآتي: ٢٠٠ خلية في طبقة الإدخال و ٢٠ خلية في الطبقة المخفية و ٢٠٠ خلية في طبقة الإخراج للحصول على نسبة كبس أعلى ١:١٠.

وفي التجربة الثالثة، تم تقسيم إشارة الكلام إلى مقاطع غير متداخلة كل مقطع يحتوي على ١٠٠ عينة وكان تنظيم الشبكة كالآتي: ١٠٠ خلية في طبقة الإدخال و ٥ خلايا في الطبقة

المخفية و ١٠٠ خلية في طبقة الإخراج للحصول على نسبة كبس أعلى ١:٢٠، لكن تعذر الوصول إلى تدريب للشبكة بسبب احتياجها إلى وقت طويل جدا في عملية التدريب وكذلك عدم التوصل إلى بارامترات التعلم المناسبة التي تزيد من كفاءة التدريب وحتى في بعض الحالات التي استطعنا فيها من الوصول إلى تدريب الشبكة فان الملفات المسترجعة بعد عملية الكبس لم تكن مقبولة.

التطبيق العملي لتنفيذ عملية كبس إشارة الكلام وتشفيرها لكل تجربة موضحة في الخطوات الآتية:

الخطوة ١: قراءة إشارة الكلام من الملف.

الخطوة ٢: تقطيع إشارة الكلام إلى مقاطع غير متداخلة يحتوي كل مقطع على ٥٠٠ عينة.

الخطوة ٣: تكوين ملفات التدريب من البيانات المقطعة، إذ يحتوي ملف التدريب على كل مقاطع الإشارة للملف.

الخطوة ٤: تدريب الشبكة العصبية ذات الانتشار العكسي، إذ يكون الإدخال مقاطع الإشارة (متجهات) من ملفات التدريب المهيأة للبيانات، ويحسب الإخراج لخلايا الطبقة المخفية بضرب متجه الإدخال بمصفوفة الأوزان بين طبقة الإدخال والطبقة المخفية.

الخطوة ٥: خزن الإخراجات للطبقة المخفية في ملف (الملف المكبوس الناتج).

تم إعادة خطوات عملية الكبس أعلاه نفسها للتجربة الثانية والثالثة عدا تغيير الخطوة الثانية و تقطيع إشارة الكلام إلى مقاطع غير متداخلة يحتوي كل مقطع على ٢٠٠ و ١٠٠ عينة على التوالي.

أما التطبيق العملي لتنفيذ عملية فك الكبس والتشفير لإشارة الكلام لكل تجربة فهو موضح في الخطوات الآتية:

الخطوة ١: فتح الملف المكبوس.

الخطوة ٢: قراءة متجه من الملف (إخراجات خلايا الطبقة المخفية: ١٠٠ في التجربة الاولى و ٢٠ في التجربة الثانية و ٥ في التجربة الثالثة).



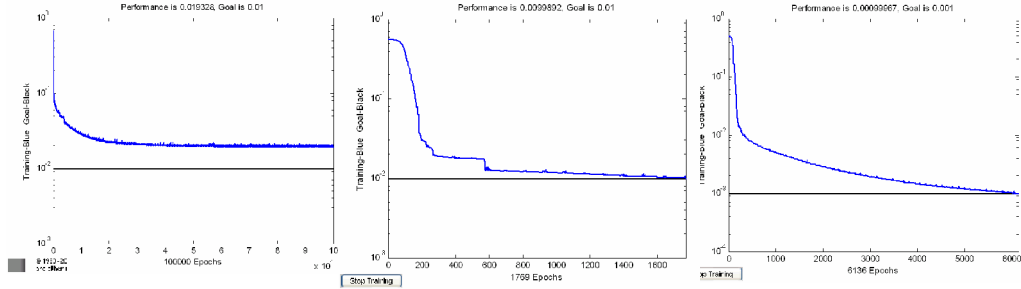
الخطوة ٣: حساب الإخراجات لخلايا طبقة الإخراج، بضرب الإخراجات لخلايا الطبقة المخفية بالأوزان التي بين الطبقة المخفية وطبقة الإخراج.

الخطوة ٤: إعادة ترتيب مقطع الإشارة الناتج من طبقة الإخراج للشبكة في موقعه المناسب لغرض إعادة بناء الإشارة.

الخطوة ٥: خزن الإشارة (الإخراجات لطبقة الإخراج) في ملف ذات تنسيق .Wav.

تم استخدام المترجم (Matlab V.7) في إعداد كل البرامج المتعلقة بالنظام.

في التجارب جميعها استخدمت دالة التعلم انحدار الميل مع العزم Gradient descent with momentum and adaptive في تعديل أوزان الشبكة. والشكل (٢) يوضح انحدار الخطأ في الشبكة أثناء التدريب والوصول إلى الهدف لأحد الملفات.



(ج)

(ب)

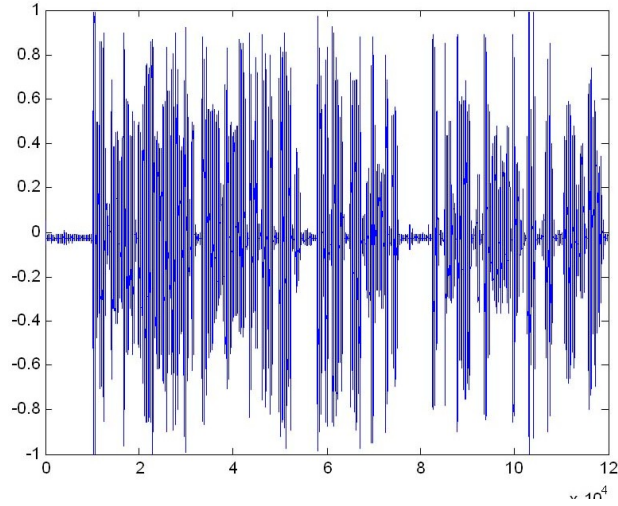
(أ)

الشكل (٢) يوضح انحدار الخطأ في الشبكة أثناء التدريب والوصول إلى الهدف لأحد الملفات

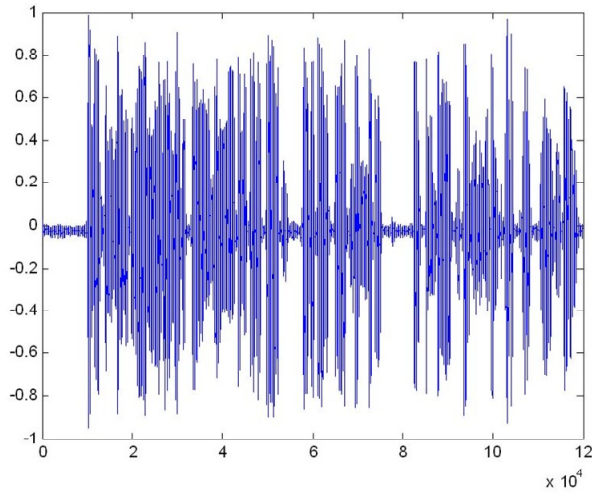
(أ) في التجربة (١) - (ب) في التجربة (٢) - (ج) في التجربة (٣)

مناقشة النتائج

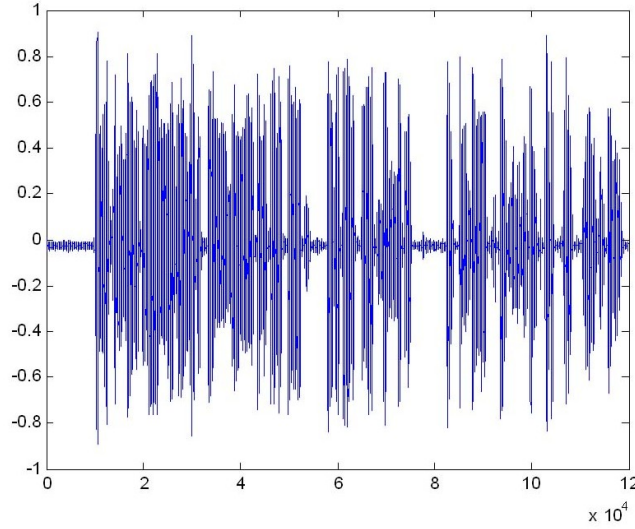
تم إجراء ثلاث تجارب كما ذكر في التطبيق العملي، في كل منها وتم استخدام الملفات ذاتها في عملية التدريب، إذ استخدمت عدة ملفات ذات أحجام مختلفة في عملية التدريب لغرض الكبس. شكل (٣ - أ - ب - ج) يوضح الإشارة الأصلية والمعاد تكوينها من التجربتين الأولى والثانية.



شكل (٣ - أ) الإشارة الأصلية قبل عملية الكبس



شكل (٣ - ب) الإشارة بعد عملية فك الكبس في التجربة (أ)



شكل (٣-ج) الإشارة بعد عملية فك الكبس في التجربة (٢)

في التجربة الأولى تم كبس كل مقطع من الإشارة يحتوي على ٥٠٠ عينة إلى ١٠٠ عينة وفيها تم الحصول على نسبة كبس CR ٥:١ كما موضح في الجدول (٢).

الجدول (٢) نسب الكبس الناتجة من تجربة (١)

| ت | حجم الملف الأصلي قبل عملية الكبس | عدد العينات قبل عملية الكبس | حجم الملف بعد عملية الكبس | عدد العينات بعد عملية الكبس | نسبة الكبس (CR) |
|---|----------------------------------|-----------------------------|---------------------------|-----------------------------|-----------------|
| 1 | 117 KB                           | 120000                      | 46.9 KB                   | 24000                       | 1:5             |
| 2 | 1.43 MB                          | 750000                      | 293 KB                    | 150000                      | 1:5             |
| 3 | 890 KB                           | 455500                      | 177 KB                    | 91100                       | 1:5             |
| 4 | 951 KB                           | 487000                      | 190 KB                    | 97400                       | 1:5             |
| 5 | 314 KB                           | 160500                      | 62.7 KB                   | 32100                       | 1:5             |
| 6 | 537 KB                           | 275000                      | 107 KB                    | 55000                       | 1:5             |
| 7 | 785 KB                           | 402000                      | 157 KB                    | 80400                       | 1:5             |
| 8 | 947 KB                           | 485000                      | 189 KB                    | 97000                       | 1:5             |

في التجربة الثانية تم كبس كل مقطع من الإشارة يحتوي على ٢٠٠ عينة إلى ٢٠ عينة وفيها تم الحصول على نسبة كبس CR ١٠:١ كما موضح في الجدول (٣).

الءءول (٣) نسب الكبس الناتءة من ءءربة (٢)

| ء | ءءم الملف الأصلى قبل عملىة الكبس | ءءم العىنات قبل عملىة الكبس | ءءم الملف بعء عملىة الكبس | ءءم العىنات بعء عملىة الكبس | نسبة الكبس ( CR ) |
|---|----------------------------------|-----------------------------|---------------------------|-----------------------------|-------------------|
| 1 | 117 KB                           | 120000                      | 23.4 KB                   | 12000                       | 1:10              |
| 2 | 1.43 MB                          | 750000                      | 146 KB                    | 75000                       | 1:10              |
| 3 | 890 KB                           | 455500                      | 89 KB                     | 45550                       | 1:10              |
| 4 | 951 KB                           | 487000                      | 95.1 KB                   | 48700                       | 1:10              |
| 5 | 314 KB                           | 160500                      | 31.3 KB                   | 16050                       | 1:10              |
| 6 | 537 KB                           | 275000                      | 53.7 KB                   | 27500                       | 1:10              |
| 7 | 785 KB                           | 402000                      | 78.5 KB                   | 40200                       | 1:10              |
| 8 | 947 KB                           | 485000                      | 94.7 KB                   | 48500                       | 1:10              |

إن عملىة إعاءة ءكوىن الإءارة للكلام المكبوس (المشفرف) فى ءءربة الأولى ءءم بءءهىز مءءه ذى ١٠٠ إءءال لءالا طبقة المءفىة للشبكة (لان ءءم لءالا للطبقة المءفىة ١٠٠ فى الكبس) ونءصل على ٥٠٠ قىمة من لءالا طبقة الإءراج، ءم نقوم بءمء القوالب مع بعضها لاسءرجاع إءارة الكلام  $x(n)$  ونءزن الإءارة المسءرجعة  $x(n)$  فى الملف لإعاءة عرضها.

وفى ءءربة ءانىة ىءم ءءهىز ٢٠ إءءالاً (ءءم لءالا للطبقة المءفىة ٢٠ فى الكبس) ونءصل على ٢٠٠ قىمة من لءالا طبقة الإءراج، ءم نقوم بءمء القوالب مع بعضها لاسءرجاع إءارة الكلام  $x(n)$  ونءزن الإءارة المسءرجعة  $x(n)$  فى ملف لإعاءة عرضها. ءنء ءساب مؤءر الأداء PRD لإءارة الكلام ءبىن انه فى ءالة ءنءظىم الذى ءم اسءءءامه فى ءءربة ءانىة أى ( ٢٠٠ : ٢٠ : ٢٠٠ ) كان PRD أعلى من ءنءظىم للءءربة الأولى ( ٥٠٠ : ١٠٠ : ٥٠٠ ) والءءىر بالذكر ان القىم الصءىرة للمؤءر PRD ءءل على نءاع الءوارزمية أى: إن النءائء المسءءصلة من ءءربة الأولى كانت أفضل من نءائء ءءربة ءانىة.. والءءول (٤) ىبىن قىم PRD ءى ءم ءءصول علىها من ءءربءىن.

الجدول (٤) يوضح مؤشر الأداء (PRD) لعملية الكبس في التجريبتين

| PRD (%)                                     |  | حجم الملف الأصلي |
|---|--|------------------|
| تجربة ٢<br>عدد الخلايا في الطبقة المخفية ٢٠ | تجربة ١<br>عدد الخلايا في الطبقة المخفية ١٠٠ |                  |
| مقياس PRD للإشارة بعد فك الكبس              | مقياس PRD للإشارة بعد فك الكبس               |                  |
| 14.86                                       | 4.18   | 117 KB           |
| 24.43                                       | 15.83  | 1.43 MB          |
| 18.85                                       | 7.03   | 890 KB           |
| 29.6  | 13.47  | 951 KB           |
| 21.97                                       | 14.33  | 314 KB           |
| 19.40                                       | 11.94  | 537 KB           |
| 18.33                                       | 11.60  | 785 KB           |
| 17.22                                       | 11.27  | 947 KB           |

كذلك من المحتمل حدوث بعض الأخطاء في أجهزة النقل لإشارات الكلام، لتقييم أداء عملية الكبس في هذه الحالة يمكن استخدام بعض معايير الدقة.

ويتم تقييم جودة الكلام المسترجع وقابلية الفهم المتبقية للكلام المكبوس (أو المشفر) بواسطة الاختبارات الشخصية، والموضوعية (Subjective & Objective tests). فالاختبارات الشخصية تستخدم الأذن البشرية في الإصغاء، و نتمكن من تقييم جودة الكلام المسترجع يجب الاستماع إلى إشارة الكلام الأصلية، والمسترجعة ثم مقارنتهما معاً. أما الاختبار الأوثق لتقييم الكلام المكبوس، والمسترجع فيتم بالاعتماد على طريقة الاختبارات الموضوعية، (Tashan, 2001).

الاختبارات الموضوعية طرق رياضية تستخدم لتقييم جودة الكلام المسترجع. توجد عدة أنواع من الاختبارات الشبئية، أهمها هو اختبار نسبة طيف مقطع الإشارة إلى الضوضاء (Segmental Spectral Signal –to- Noise Ratio: SSSNR) والذي يمكن حسابه من المعادلة (4) الآتية:

$$(SNR_r)_{db} = d(x, y) = 10 * \log_{10} \left[ \frac{\sum_{k=0}^{N-1} |X_r(k)|^2}{\sum_{k=0}^{N-1} (|X_r(k)| - |Y_r(k)|)^2} \right] \quad \text{-----} \quad (4)$$

إذ إن  $X_r(k)$  تمثل قيم تحويل فوريير السريع FFT لقلب من عينات إشارة الكلام الأصلية، و  $Y_r(k)$  تمثل قيم تحويل فوريير السريع FFT لقلب من عينات إشارة الكلام المكبوسة (المشفرة) / المسترجعة، و  $r$  تمثل تسلسل القلب، و  $N$  تمثل حجم القلب (Lee and Chou, 1984).

يسبب تعريف SSSNR مشكلة في حالة وجود فترة من السكون في أثناء نطق الكلام فإن أية كمية من الضوضاء تؤدي إلى رفع معامل الضوضاء، ولحل هذه المشكلة يتم إيجاد المقاطع الساكنة واستبعادها، ويمكن إجراء ذلك بحساب الطاقة (Energy) لكل مقطع من الإشارة الأصلية باستخدام المعادلة (5) ومن ثم حساب (SNR) للمقطع الذي طاقته أعلى من قيمة العتبة المحددة (Threshold)

$$Er = \sum_{k=1}^{N-1} x_r^2(k) \quad \text{-----} \quad (5)$$

من الجدول (5) ، نلاحظ أن نتائج SNR للتجربة الثانية أفضل من التجربة الأولى، أي إن الكلام المسترجع من التجربة الثانية كان أكثر جودة من الكلام المسترجع من التجربة الأولى. كما نلاحظ أن العامل الرئيس المسيطر على مستوى الكبس والتشفير هو طول القلب للإشارة . إذ إن تقليل عدد العينات داخل المقطع سوف يقلل من قابلية الفهم المتبقية للكلام المكبوس.

الجدول (٥) يوضح اختبار نسبة طيف مقطع الإشارة إلى الضوضاء لعملية الكبس في التجريبتين

| تجربة ٢ *           | تجربة ١ *           | حجم الملف الأصلي |
|---------------------|---------------------|------------------|
| (SNR) <sub>db</sub> | (SNR) <sub>db</sub> |                  |
| 25.0                | 15.5                | 117 KB           |
| 26.9                | 19.1                | 1.43 MB          |
| 32.0                | 19.4                | 890 KB           |
| 26.0                | 15.3                | 951 KB           |
| 24.5                | 14.6                | 314 KB           |
| 34.1                | 28.4                | 537 KB           |
| 24.3                | 19.4                | 785 KB           |
| 20.2                | 16.3                | 947 KB           |

\* db اختصار الديسيبل Decibel : مقياس للإحساس السمعي بالصوت يختلف عن المقياس الآلي لقوة الصوت ابتكره علماء الفيزياء [1].

كما اتضح من نتائج تطبيق التجارب على عدد من الأمثلة المختلفة أن الخوارزمية لا تتأثر بصفة المتكلم أو نوع الكلام وصيغته المعتمدة في التطبيقات العملية للخوارزمية، ولم يسجل على الخوارزمية أي اختلاف في نسبة الكبس أو نسبة التشوه مع اختلاف نوع الكلام أو صيغته وإنما أعطى نتائج متقاربة من بعضها البعض.

وقد قدم عدد من الباحثين في هذا المجال نتائج جيدة وكان أبرزها :

الباحثة (عبد القادر، ٢٠٠١)، في بحثها الموسوم (كبس الصوت عند الزمن الحقيقي) والذي استخدمت فيه خوارزمية (LZW) في عملية كبس الصوت قبل تخزينه على ذاكرة الحاسوب والعمل على فك الكبس عند محاولة استماع الصوت مجدداً، إذ وصلت نسبة كبس عندها إلى 1:5 من نسبة البيانات الأصلية.

أما الباحثة (قدو، ٢٠٠٤)، فقد قامت بقراءة الملف والبدء بعملية الكبس والتي مرت بثلاث مراحل تم فيها العمل على إزالة فترات الصمت، ومن ثم الاحتفاظ بمعدل عينات متناوبة من الإشارة لغرض معالجتها، بعدها قامت بعمل مواءمة لمنحني الإشارة الناتجة ليتم تخزينها في ملف جديد وتحت صيغة تخزين جديدة وحصلت على نسبة اختزال في حجم الملف الصوتي بحدود (16.283%) ونسبة SEGSR بحدود (25.195 dB).

أما في مجال استخدام الشبكات العصبية في كبس الإشارة، فقد قام الباحثون (Kocyigit et al.,2002). بكبس إشارات تخطيط القلب باستخدام شبكة عصبية مألوفة Ordinary type neural network ، (لم تذكر اسمها في متن البحث المذكور) وحصلوا على نسب كبس 1:20 و 1:40 وكانت نتائج مؤشر الأداء (PRD) المستحصلة لعملية الكبس بنسبة 1:20 أفضل من تلك التي حصلوا عليها بنسبة كبس 1:40 .



### المصادر

١. أيوب، عبدالرحمن، (١٩٨٩) تحليل عملية التكلم وبعض نتائجها التطبيقية، منشور في مجلة عالم الفكر، المجلد العشرون – العدد الثالث.
٢. عبد القادر، إسراء عبد السلام، (٢٠٠١)، "كبس الصوت عند الزمن الحقيقي"، بحث ماجستير، كلية علوم الحاسبات والرياضيات، جامعة الموصل.
٣. فرغلي، علي صبري، (١٩٨٩) عرض وتحليل لكتاب اللغة العربية والحاسوب تأليف نبيل علي منشور في مجلة عالم الفكر، المجلد العشرون – العدد الثالث.
٤. قدوس، سجي جاسم محمد، (٢٠٠٤)، "كبس إشارة الكلام بواسطة استخلاص الخواص"، بحث ماجستير، كلية علوم الحاسبات والرياضيات، جامعة الموصل.
5. Abdul\_Majeed, A. M. (2003), Speech encryption using scrambling, Thesis, Department of Computer Sciences, College of Computers and Mathematics Sciences, University of Mosul, Iraq.
6. Allaf, O. N., (1999), Backpropagation on neural network for image compression and enhancement, M.Sc. Thesis, Department of Computer Sciences, College of Computers and Mathematics Sciences, University of Mosul, Iraq.
7. Carrato, S. and Marsi, S. (1993) Adaptive Structure based on neural networks for subband-filtered image compression, neural network world, vol.1 , pp. 25-40.
8. Cawley., G., Bangham., J. A., Theobald., B., (2001), (PG). Funding: BBC [http://www.microsoft.com/billgates/speeches/2001/04\\_02chi.asp](http://www.microsoft.com/billgates/speeches/2001/04_02chi.asp).
9. Cottrel, G.W., Munro, P., and Zipser, D. (1989) Image compression by backpropagation: An example of extensional programming, Models of Cognition, shorkey (Ed.) Norwood: Ablex, pp.208-240.
10. Gary, D. K., Trevor, J.H. and Timothy, J. N. (1993) An investigation of the generalization performance of neural networks applied to lofargram classification, Neural Computing & applications 1(2):147-159.
11. Kientzle, T (1998) "A programmer's Guide to Sound" Addison-Wesley.
12. Kocyigit., Y., Korurek., M., Karlik., B., (2002), ECG compression by artificial neural networks. From internet.

13. Lee.,L.S., Chou., G.C., (1984), “ A New Time Domain Speech Scrambling System Which Does Not Require Frame Synchronization”, IEEE Journal on Selected Areas in Communications, Vol. SAC-2, No.3.
14. Picone J., (1996), FUNDAMETALS OF SPEECH RECOGNITION:A SHORT COURSE, INSTITUTE FOR SIGNAL AND INFORMATION PROCESSING , <http://isip.msstate.edu>
15. Signal Processing Blockset for Matlab V7.0 Copyright 1984-2004, The MathWorks, Inc.
16. Tashan, T.M.K., (2001), “Objective Test of Speech Signal” M.SC. Thesis, Al-Mustansiryah University.
17. Weinschenk.,S., Barker.,D.T., (2000), Designing effective speech interfaces, published by John Wiley & Sons, Inc. pp 137 .

This document was created with Win2PDF available at <http://www.daneprairie.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.