التقليل من الضوضاء المتداخلة مع الكلام باستخدام المرشح التكيفي

ايمان فتحي احمد
قسم الحاسبات/كلية التربية

بسام علي مصطفى
قسم الحاسبات/كلية علوم الحاسبات والرياضيات

جامعة الموصل

تاريخ الاستلام
2006/6/27

تاريخ القبول
2006/11/14

Abstract

Suggested this project a filter for speech signal enhancement. While steps can be listed below : First recording speech signal from microphone. Second input signal should be divided and attached into overlapped frames. Third these frames should be identified and recognized at (Voiced Sound)frame signals and (Unvoiced Sound). By using two adaptive filter algorithms: The Least Mean Squares(LMS) algorithm and Recursive Least Squares (RLS) algorithm. Then the objective tests have been done by measuring the LMS errors between input and output signals. When the suggested filter applied to several of applied samples, the results showed the filter order has agreed influence upon filtration time, and on the square of the error signal. The tests revealed that the LMS algorithm gives good results of reducing the square of the error signal.

الملخص

اقترح هذا العمل مرشحا يعمل على تحسين إشارة الكلام الداخلة اليه ، اما مراحل العمل فيمكن إدراجها فيما يأتي: أولا اكتساب إشارة الكلام من لاقط الصوت، ثانياً تقسم الإشارة الداخلة على مقاطع متداخلة، ثالثاً تمييز هذه المقاطع إلى مقطع من نوع صوت الكلام المسموع (*Voiced Sound*) ومقطع من نوع صوت الكلام غير المسموع *Unvoiced* (*Sound*). استخدمت خوارزميتان من خوارزميات المرشح التكيفي (*Adaptive Filter*)هما:خوارزمية النهايات الصغرى للمربعات الوسطية LMS وخوارزمية النهايات الصغرى للمربعات التكرارية RLS تلا ذلك إجراء الاختبارات الشيئية وذلك عن طريق قياس الحد الأدنى لمربع الخطأ بين الإشارة الداخلة والإشارة الخارجة.وعند تطبيق المرشح المقترح على عدد من الأمثلة التطبيقية أظهرت النتائج ان درجة ترتيب المرشح لها تأثير كبير في زمن تنفيذ المرشح ، وكذلك على مربع إشارة الخطأ.أظهرت الاختبارات ان خوارزمية LMS تعطي نتائج أفضل من خوارزمية RLS في التقليل من مربع إشارة الخطأ.

1-1 المقدمة (*Introduction*)

شهد القرن الحالي توسعاً سريعاً في مجال البحث العلمي وتقدماً مناسباً فـي توثيـق حقائق العلوم والبحث والتطوير وهذا بدوره يعطي فهماً جيداً ودقيقاً فـي اختبـار الطرائـق القياسية المستخدمة[13] ولقد كان علم معالجة الكلام واحداً من هذه التطورات التي شهدها هذا القرن ، اذ ان الكلام هو الوسيلة اللغوية الوحيدة المستخدمة عالمياً للاتصال بين أفراد الجنس البشري، وما الكلام إلا عملية استحداث موجات صوتية بواساطة الحركة الاراديـة للتركيـب التشريحي في نظام توليد الكلام لدى الانسان لنقل المعلومات من المتكلم الى السامع[3]. ومن خصائص الكلام تردد حدة الصوت (*Pitch Frequency*) وتمثل تردداً لاهتـزاز الأوتـار الصوتية، وتختلف القيمة الوسطية لهذا التردد من فرد الى اخر ولكل متكلم مدى اعلـى مـن القيمة الوسطية او اقل. كذلك تحتوي إشارة الكلام على معلومات مهمـة تتعلـق بشخـصية المتكلم. والمقصود بشخصية المتكلم هي تلك الاختلافات القليلة في تردد طبقة الصوت[2].

2-1 الكلام (*Speech*)

الكلام هو ديناميكية الإشارة التي تحتوي على المعلومات الأساسية للكلام التي تـدعى بالموجة الصوتية. هذه الموجة التي تنتج خلال توليد الضغط داخل الفم للمتكلم وهـي نتيجـة لسلسلة حركات متناسقة ومتتالية لتراكيب النظام اللفظي لدى الإنسان[14].

3-1 مفهوم الديسيبل (*Decibel Scale dB*)

هو المقياس الذي يستخدم لقياس مستوى كثافة الصوت وهو واسع الاستعمال في مجال الإلكترونيات، الإشارة والاتصالات.البيل *bel* هو وحدة قياس منسوب القـدرة (تـساوي 10 ديسيبل) واكتشفت من قبل العالم(*Alexander Graham Bell*)الذي نال درجة الشرف على هذا الاكتشاف[7].والعامل المهم الذي نواجهه عندما نتعامل مـع الأصـوات هـو مفهـوم (الديسيبل dB) وليس القدرة ، لأننا نتعامل مع مستويات مختلفة ومتعددة مـن قـيم اتـساع الإشارة(*Signal Amplitude*) ، والسبب في ذلك يعود الى التعامل مـع مـديات واسـعة ومختلفة من مستويات الأصوات فمن الصعب التعامل مع الأرقام. لذلك اسـتخدام المقيـاس اللوغـاريتمي الـذي يـسمى مـستوى ضـغط الـصوت *decibel (Sound Pressure Level(SPL))* لقياس كثافة موجة الصوت ، فعند مـستوى صـفر ديـسيبل (0db SPL) تساوي موجة الصوت (10^{-16}watt/cm^2) وهو اضعف مـستوى للـصوت تـستطيع الأذن البشرية التقاطه، ويكون الكلام الاعتيادي عند مستوى (60 ديسيبل)، اما الصوت المؤلم الذي يـســبب الـــضرر بـــالأذن البـــشرية ويكـــون عنـــد درجة (140 ديسيبل)[1].

$$Y= log_{10} X \qquad (1)$$

4-1 تحليل الكلام (*Speech Analysis*)

ان الهدف الرئيس لتحليل الكلام هو تمييز الإشارة الصوتية وكذلك التقليل من عرض حزمة الـ(*bandwidth*) وتمييز الإشارة الكلامية بعلامات قليلة. وتتم هذه العملية في حيـز الزمن(*Time domain*) الذي يشمل إطارات مقاطع البيانات (*data windowing*) لاختبار مقاطع الكلام القصيرة (*Frame*) للبيانات التي يفترض ان تكون ثابتة خلال المدة الزمنيـة. فضلاً عن تحليل الكلام الذي أتاح للبشر التعامل مع الماكنة ، إذ استطاع الإنسان بناء دوائـر إلكترونية دقيقة ذات مساحة صغيرة تمكنت من ترجمة أوامر البشر وفهمها[6].

5-1 تقنيـات تمـيز الكـلام (*Speech Recognition Technique*)

ان تمييز الكلام هو عملية تحويل الإشارة الصوتية التـي يلتقطهـا مكبـر الـصوت او الهاتف الى مجموعة من الكلمات، وتمييز تلـك الكلمـات بوصـفها ناتجـا نهائيـا فـي عدد من التطبيقات مثل الأوامـر(*Command*)،والـسيطرة (*Control*)،وإدخـال البيانـات (*Data entry*)[15]. تعد معالجة إشارة الكلام واستخلاص الخواص المرحلة الأولية لنظـام تمييز الكلام وخلال هذه المكونات يعرض النظام الحاسبي اشارة الكلام نفسها، وتشير معالجة إشارة الكلام الى العمليات التي تجري على الإشارة الكلامية مثل استخدام المرشحات ،وعملية تحويل الإشارة المستمرة الى قيم رقمية(*Digitization*)، والتحليل الطيفي(*Spectrum*) والى أخره [17].

6-1 المرشحات (*Filters*)

تعـد المرشـحات مـن المكونـات الأساسـية لمعالجـة الإشـارة (*Signal processing*) وأنظمة الاتصالات والوظيفة الأساسية للمرشح هي واحدة من ضمن الوظائف الآتية[4]:

1-تحليل الإشارة الى اثنتين أو اكثر من حزم الإشارة الثانوية لغرض معالجتها.

2- تغيير طيف الترددات الإشارة مثل موازنة مخطط السمع .

3- التقليص من حجم الضوضاء.

4- المعالجة الفيديوية ومعالجة الإشارة الطبية، تحليل البيانات الاقتصادية والتمويلية.

5- فصل الإشارة عن الإشارات الأخرى المتداخلة مع الإشارة الأصـيلة مثـل إشـارة الضوضاء.

6- إعادة او تجديد الإشارة التي تكون مشوهة او منحرفة بسبب عدد من الطرائق.

المرشـح الرقمـي جـزء مهـم فـي معالجـة الإشـارة الرقميـة (*DSP*) وعندما تحول الإشارة التناظرية (*Analog*) الى سلسلة من العينات الرقمية يجـري عمل المرشح الرقمي على الإشارة من خلال عمليـة اللافـوف الرياضـي (*Cnvolution*)

لعينات الإشارة الداخلة مع استجابة ترددات المرشح(*Frequency Response*) وهي عمليـة تجري على البيانات وتتضمن عمليتي (الضرب والجمع)[8].

7-1 مرشح استجابة النبضات المحدود(*Finite Impulse Response(FIR)*)

مرشح *FIR* هو أحد النوعين الرئيسين من انواع المرشحات الرقمية المستخدمة في تطبيقات معالجة الإشارة الرقمية (*DSP*) كما موضح في المعادلة (2) والنوع الآخـر هو مرشح استجابة النبضات غير المحدود (*Infinite Impulse Response(IIR)*)[10] .

$$y(m) = \sum_{k=0}^{M} b_k \; x(m-k) \qquad (2)$$

إذ b_k هي معامل المرشح (*Filter Coefficients*)، M هي درجة ترتيب المرشـح و x(m)، y(m) هي قيم إشارة الإدخال و الإخراج للمرشح (*FIR*) وإذا نفذت معادلة المرشح (2) مع قيم نبضات الإدخال *Impulse input* للمرشح (*FIR*)، فان قيم استجابة نبـضات المرشـح (*FIR*) تكون مماثلة لسلسة معاملات المرشح b_k .

$$h(k) = \begin{cases} b_k & 0 \subseteq k \subseteq M \\ 0 & Otherwise \end{cases} \qquad (3)$$

ولهذا فان المعادلة (4) هي عملية اللافوف الرياضـي (*Convolution*) لسلـسلة اسـتجابة نبضات (*Impulse response*) للمرشح h(k) مع إشارة الإدخال x(m-k) [10].

$$y(m) = \sum_{k=0}^{M} h(k)\, x(m-k) \qquad (4)$$

8-1 المرشح التكيفي (*Adaptive Filters(AF)*)

المرشحات التكيفية هي نوع من المرشحات الرقمية لها القدرة على الـتحكم الـذاتي وتغيير المعاملات. ويستطيع هذا النوع من المرشحات التغير لمطابقة الإشارة الداخـة الـى المرشح.وستخدم المرشح التكيفي في التطبيقات التي تتطلـب مختلـف مواصـفات المرشـح استجابة لتغيرات الإشارة الآنية. وهو يتطلب نوعين من إشارات الإدخال[12]

1. إشارة الإدخال *Input signal x(n)*.
2. الإشارة المرجعية *Reference d(n)* .

ويمتلك المرشح التكيفي القابلية على تغيير معاملاته وهي المعاملات الجديدة المرسلة الى المرشح من مصدر مولـد المعـاملات (*Coefficient generator*). ذلـك ان مولـد المعاملات هو عبارة عن خوارزمية تقوم بتغيير المعاملات (*Coefficient*) استجابة للإشارة القادمة. ويعد المرشح الرقمي نموذجياً نوعا خاصا من المرشحات الرقمية المعروفـة *FIR* و *IIR* او أي نوع من المرشحات الرقمية الأخرى. ويستخدم المرشح التكيفي في العديـد مـن

التطبيقات مثل تقليص كمية الضوضاء وتحسين الإشارة والتنبؤ الخطي[5]. ويتكون المرشـح التكيفي من:

1- نظام الإدخال الأولي (*Primary input*) والذي ينمذج بواسطة المرشح *FIR* مـع القابلية على تعديل قيم المعاملات.

2- نظام لنموذج المرشح *FIR*.

3- إشارة الإدخال *x(n)* ،لتي تكون مدخلا الى كلا النوعين نظـام(*Primary input*) ، ونظام *FIR* .

4- درجة إشارة الإخراج *y(n)* تمثل قيم إخراج نظام نموذج المرشح *FIR*.

5- درجـة إشـارة الإخـراج*d(n)* تمثـل قـيم إخـراج نظـام الإدخـال الأولـي (*Primary input*).

6- درجة إشارة الخطأ *e(n)* تنتج من خلال مقارنة درجة إشارة الإخراج *d(n)* للنظـام الإدخال الأولي (*Primary input*) بدرجة إشارة الاخراج y(n) للنظام المرشح *FIR*.

9-1 خوارزمية (*Least Mean Squares (LMS)*)

خوارزمية *LMS* هي واحدة من الخوارزميات المستخدمة في مجال معالجة الإشـارة الرقمية. تبدأ خوارزمية *LMS* العمل بجعل قيم أوزان المعاملات جميعا تساوي صـفرا عنـد الزمن صفر[9]، ويعتمد أداء هذه الخوارزمية وعى نحو كبير على درجة ترتيـب المرشـح فضلا عن العامل الثاني وهو العامل (μ) الذي يساعد في عملية تـسريع تقـارب عوامـل المرشح(*Cefficient*) من درجة الإخراج المطلوب.وبعد ان يتم التقارب بين عوامل المرشح مع درجة الإخراج المطلوبة تصبح درجة هذا العامل صغيرة.اي تجعل درجة عامل التقـارب (μ) ذات درجة كبيرة في بداية العمل لتسريع خوارزمية التقارب ومع مرور الـزمن يـصبح هذا العامل تدريجيا ذات درجة صغيرة وعند ذاك يمتلك المرشح التكيفي اصغر درجة لمربـع الخطأ[5].

10-1 خوارزمـية(*Request Least Squares(RLS)*)

اذا كانــت خوارزميــة *LMS* تعــد خوارزميــة بــسيطة وســهلة التمثيــل، فان*RLS* تعد خوارزمية اكثر تعقيداً وتحتاج الى عمليات حسابية اكثر وزمن تنفيذ اطول و دقة اكبر،تعمل هذه الخوارزمية على تقليل نسبة الخطأ بين قيم الإشـارة الداخلـة وقـيم إشـارة الإخراج[8]. تشبه معمارية هذه الخوارزمية وتركيبها معمارية الخوارزمية *LMS*. وعملهـا تغيير عدد من المعاملات الأولية للخوارزمية، مثل استخدام المعامل (λ) بدل المعامـل (μ). ويتيح عامل التقارب (λ) للخوارزمية *RLS* التقليل من عامل الخطأ للعينـات الـسابقة عـن طريق ضرب قيم بيانات الخطأ السابقة من قيم العامل(λ). وتكون قيم هذا العامل محـصورة

بين (0 ، 1) ، فعندما تكون قيم هذا العامل تساوي واحدا، توجد قيم الخطأ بصورة كاملـة،أي توجد قيم الخطأ للعينات السابقة فضلا عن العينات الحالية،وعندما تقترب قيم هذا العامل مـن الصفر فان قيم الخطأ تحسب لقيم العينة الحالية فقط[14].

11-1 خوارزميـة تمييـز إشـارة الكـلام (*Speech Signal Recognition Algorithm*)

ميز مقطع إشارة الكلام المعطى بوصفه صوت كلام مـسموع (*Voiced sound*) او صوت كلام غير مسموع (*Unvoiced sound*) خلال الخطوات الاتية:

- مررت إشارة الكلام على مرشح من نوع مرشح الترددات العالية *HPF* لغـرض إزالة الترددات الواطئة من إشارة الكلام .
- عولجت إشارة الكلام لغرض عمل تمثيـل جديـد للإشـارة فـي مجـال تمييـز الكلام.ان موجة إشارة الكلام نفسها لاتعطي تمثيلا جيدا للإشارة في مجال تمييـز الكلام ذلك لأنها لاتعرض على نحو مباشر الخواص اللغوية للجملة. وتتغير هيئة الموجة (*Waveform*) على نحو أسـرع مـن حركـة الأعـضاء المتحركـة (*Articulaters*) للمتكلم. ويعد تحليل المقاطع القصيرة من المفاهيم الرئيسة فـي مجال معالجة إشارة الكلام الرقمية وهي عادة ما تطلب في معظم طرائق التحليـل التي تكون فيها الإشارة ثابتة خلال الزمن. لهذا فان إشـارة الكـلام ليـست ذات خواص ثابتة خلال الزمن اذ ان الإشارة تتغير دوريا وببطء خلال الـزمن، وقـد قادت هذه الافتراضية إلى المفاهيم الأساسية للتحليل خلال زمـن قـصير علـى مقاطع قصيرة من الإشارة الكلامية[2].اذ قطعت إشارة الكلام الداخلة *X(n)* الـى عدد من القوالب المتداخلة(*Overlap Frames*) المتساوية في الحجم اعتماداً على حجم القالب(*N*) المحدد ليحتوي كل قالب على N من العينات.

$$X_1(n) = \left[X_r(1) + X_r(2),........X_r(N)\right]$$

اذ تمثل r تسلسل القالب ، وتمثل N حجم القالب

ثم قام النظام ببناء خوارزمية تحتوي على عدد من الدوال تقوم بإيجاد المعاملات الاتية لكـل مقطع من مقاطع الإشارة:

1. معاملات التنبؤ الخطي *LP coefficients*.
2. طاقة *energy* كل مقطع .
3. التنبؤ بعامل الخطأ *Predication error* .
4. معامل الانعكاس الأول *First –order refliection coefficient*.

اذ كانت الإدخالات لهذه الدوال كما ياتي :

1. قيم إشارة الإدخال *X(n)*.
2. عدد مقاطع الإشارة *No. of frames*.
3. طول المقطع *Frame length*.
4. طول مقطع التداخل *Overlap*.

أ. جعل العداد يساوي صفرا .

ب. زيادة العداد بواحد .

ت. إيجاد معامل الانعكاس الأول (*First-order refliction coefficient*) للمقطع الأول الذي يقوم بحساب مقدار كثافة الصوت وسرعته لمنطقة القناة الصوتية *Vocal tract* من خلال المعادلة الاتية .

$$\hat{a}_j^i = \frac{-\frac{1}{N}\sum_{n=0}^{N-1} X(n)X(n-1)}{\frac{1}{N}\sum_{n=0}^{N-1} X(n-1)X(n-1)} = \frac{\hat{R}xx(1)}{\hat{R}xx(0)} \qquad (5)$$

اذ تمثل j a_j^i درجة المرشح .

و *i* عدد معاملات المرشح.

وتمثل *X(n)* قيم الإشارة لعينات المقطع .

و *X(n-1)* قيم العينات السابقة .

ث. نستطيع الآن الحصول على معامل التنبـؤ بعامـل الخطـأ *error predication* من خلال حساب ترميز التنبؤ الخطي *LPC* من المعادلة الاتية:

$$S(n) = \sum_{k=1}^{p} a_1 s(n-k) \qquad (6)$$

اذ *a* هو معامل التنبؤ الخطي الذي تم الحصول عليه من المعادلة (5) و*S* هي إشارة الكـلام الحقيقية

بينما معامل الخطأ e(n) الذي تم الحصول عليه من خلال المعادلة الاتية :

$$e(n) = S(n) - \bar{S}(n) = S(n) - \sum_{k=1}^{p} a_k S(n-k) \qquad (7)$$

ح. حساب الطاقة لوقتن قصير للمقطع الأول من خلال المعادلة الاتية:

$$E_n = 10\log(\varepsilon + \frac{1}{N}\sum_{m=n-N-1}^{n} X^2(m) \qquad (8)$$

اذ ε عـدد موجـب لمنـع اخـذ لوغـاريتم للـصفر وN عـدد المقـاطع *Frames* و *X(n)* هي قيم الإشارة .

خ. الخطوة التالية من الخوارزمية هي إيجاد قيمة معاملات التنبؤ الخطي الثاني والثالث والى اخره $\hat{a}_1^p \quad , \hat{a}_2^p \; ,......... .., \hat{a}_p^p$ من المعادلة (5)

ج. تحليل بقية المقاطع من خلال الاستفادة من ناتج تحليل المقطع الأول .

د. زيادة العداد بمقدار واحد.

ذ. هل حللت جميع مقاطع الإشارة جميعا ؟ إذا كان الجواب بنعم يتم الخروج من الـدارة وإلا فيتم الرجوع الى الخطوة (ب) .

بعد الحصول على مقدار الطاقة لكل مقطع ومعامل الـ *LPC Coefficient* وكـذلك نسبة الخطأ للـ *LPC error*. قورن ناتج (جمع قيم المقطع الحالي مع المقطع الـذي يليـه) بقيمة العتبة(*Thershold value*) فإذا كان ناتج المقارنة اكبر من قيمة العتبة فإن الناتج يدخل في مقارنة أخرى إذا كان مجموع (طاقة المقطع الحالي مع المقطع الذي يليه) اكبر من واحـد فان المقطع يعد من نوع صوت الكلام المسموع(*Voiced sound*) وإذا كان الناتج اقل مـن قيمة العتبة او اقل من واحد فإن المقطع يعد مـن نـوع صـوت الكـلام غيـر المـسموع (*Unvoiced sound*). اعطى ناتج التحليل النهائي قيم (واحد و صفر) اذ مثل المقطـع مـن نوع صوت الكلام المسموع (*Voiced sound*) بقيمة 1 (واحد) و تمثيل المقطع من نـوع صوت الكلام غير المسموع(*Unvoiced sound*) بقيمة 0(صفر).

12-1 خوارزمية التحسين (*Enhancement Algorithm*)

إشارة الضوضاء المتداخلة مع الإشارة الأصيلة عند عملية التسجيل فظروف الغرفـة الاعتيادية مثل صوت المروحة او صوت مكيف الهواء يعمل على إضعاف من وضـوح الكلام ويؤثر على خاصية السمع لدى الشخص المستمع لذلك كان لابـد مـن اسـتخدام خوارزميات التحسين لغرض التقليل من الضوضاء المضافة مع الإشارة الأصيلة[16].

13-1 الترددات الأساسية لنبرة الصوت (*Pitch Fundamental Frequency(F_0)*)

وتعرف على أنها تلك الترددات التي تحدث عندما يحدث تذبذب للحبال الصوتية عندما صدر صوت الكلام المسموع Voiced sound، وتعد F_0 تعتبر من العوامل المهمـة لتقـدير طيف الإشارة على نحو متقدم اذ تستخدم في أنظمة التمييز[15].

$$
\begin{aligned}
&F(n) = \log_{10}(f_0(n)) \\
&50\,Hz \subset f_0 \subseteq 500\,Hz \quad for Voiced \\
&for\ Unvoiced \quad f_0 \ \ is undefine\ \ d
\end{aligned}
\qquad (9)
$$

فنبرة الصوت تحسب بواساطة أربعة عوامل رئيسة والتي تتضمن (الطـول والتـوتر والجهد وحجم لحبال الصوتية) .تحمل تغيرات نبرة الصوت معظم التغيرات مـن انخفـاض وارتفاع في طبقة الصوت، فضلا عن أسلوب المتكلم وحالته الصحية ولهجته ومدى انفعالـه ، ويساوي معدل تردد صوت الرجل 128 دورة في الثانية ، أما معدل تردد صـوت النـساء فيساوي 256 دورة في الثانية[2].

- قام النظام بإيجاد النبرة *Pitch period* لكل مقطع من مقاطع الإشـارة الداخلـة عن طريق :

1. ضرب قيمة المقطع من نوع صوت الكـلام المـسموع *Voice sound* بنافـذة هاننيـك (*Digital Filter Window*) التي تتم في حيز الزمن ويقصد بالنافذة (*Window*) مجموعة من القيم التي تستخدم لإنتاج التأثير الواضح في الإشارة [15]، ان إشارة الكلام *X(n)* تضرب في نافذة محددة المدة *W(n)* للحصول على عينات كلام تعتمد على شكل النافذة ،وتعد واحـدة من الطرائق التي تستخدم لتجنب قطع نهايات المقاطع الكلامية عن طريق تناقص الإشارة الى الصفر اوالى قريب من الصفر وتقليل التداخل على نحو غير ملائم عـن طريـق اسـتخدام النوافذ،ويتـــــضمن اختيـــــار حجـــــم النافـــــذة عـــــددا مـــــن الشروط هي :

1. يجب ان تكون النافذة ذات طول مناسب بما لا يؤثر في خصائص الإشارة أي لا تغير من مغزى الإشارة داخل النافذة.
2. يجب ان تكون النافذة ذات طول مناسب لإنتـاج عينـات كافيـة لحـساب العوامـل المطلوبة.
3. يجب ان لا تكون النافذة الجيدة قصيرة لتجنب قطع إشارة الكلام فـي عمليـة تحليـل الإشارة التي تتكرر على نحو دوري[2].

$$y(n) = W(n) * X(n) \qquad (10)$$

X(n) إشارة الكلام وتضرب في قيم النافذة *W(n)* وينتج عن هذه العملية مجموعة عينات من إشارة الكلام على هيئة النافذة ، ومن أنواع النوافـذ المـشهورة *Bartlett* ،*Rectangular*، *Hanning* ، *Hamming*، *Blackman* .

- ثـم قـام النظـام بتـصميم مرشـح الـذي يطلـق عليـه المرشـح التكيفـي (*Adaptive Filter*) وهو أحد أنواع المرشحات من نوع *FIR* .
- ومن ثم استخدامت الخوارزمية *LMS* والخوارزمية *RLS* في عمليـة الترشـيح وهما من أنواع الخوارزميات المستخدمة من المرشح (*AF*) ،وحـددت العوامـل الرئيسة للخوارزمية وهي :

1. تحديد طول المرشح او درجته(*length* مابين (10-100).
2. تحديد قيمة μ المحصورة بين (0 ،1) للخوارزمية *LMS* وλ المحصورة قيمتها بـين (0 ،1) للخوارزمية *RLS* .
3. تحديد عدد دورات التنفيذ *Iteration* .
4. تحديد القيمة البدائية لعامل المرشح *Coefficient(W(n))* .
5. تحديد محدد مرور الترددات *Cutoff- frequency* .
6. تحديد زمن التأخير z^{-1} .
7. تحديد المحدد *bassband* و *Stopband*.

ان القاعدة الأساسية في عملية تحسين إشارة الكلام تتم من خلال مراقبة موجة صوت الكلام المسموع *(Voiced)* الذي يتذبذب على نحو دوري ومنـتظم إشـارة الـى التـرددات الأساسية*(Fundemental frequency)*،في حين تتاخر إشـارة الـضوضاء *(Unvoiced)* بدورة او اكثر*(Pitch period (T))* عن إشارة صوت الكلام المسموع. وتتمثل قاعـدة عمـل المرشح *(AF)* في التقليل من طاقة *(energy)* موجة إشارة الـضوضاء ،ويعمـل اسـتخدام المرشح لهذه المعرفة على إمرار الترددات التي تتذبذب عل نحو دوري ومنتظم فـضلا عـن حجز الترددات غير المنتظمة[11].

14-1 جودة الكلام المحسن

تم تقييم جـودة الكـلام المحـسن بوسـاطة الاختبـارات الشخـصية ، والـشيئية (*Subjective and Objective test*)، اما الاختبارات الـشيئية وهـي طـرق رياضـية تستخدم لتقييم جودة الكلام بعد التحسين. توجد عدة أنواع من الاختبارات الشيئية أهمها:

أولاً: إيجاد اقل قيمة لمربع الخطأ(*Minimum squared error*) بين إشارة الإدخال وبين إشارة الإخراج للمرشح كما موضح في المعادلة الآتية[13]:

$$MSE = \frac{1}{N}\sum_{n=0}^{N}\left[x(n)-\hat{x}(n)\right]^2 \qquad (11)$$

حيث تمثل *x(n)* إشارة الكلام الأصلية.

و $\hat{x}(n)$ إشارة الكلام المحسنة

وتمثل N عدد العينات إشارة الكلام.

ثانيـ ــأ: اختبـ ــار نـ ـــسبة طيـ ــف مقطـ ــع الإشـ ــارة الـ ــى الـ ــضوضاء (*Segment Spectral Signal-to- Noise Ratio(SSSNR)*) الذي يتم في حيـز التـردد قالب بعد قالب.

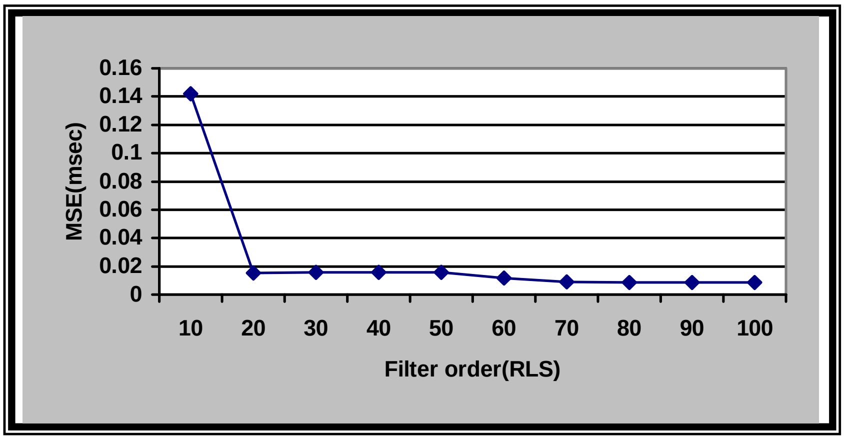
$$(SSSNR_r)_{db} = d(x,y) = 10*\log_{10}\left[\frac{\sum_{k=0}^{N-1}\left|X_r(k)\right|^2}{\sum_{k=0}^{N-1}\left(\left|X_r(k)\right| - \left|Y_r(k)\right|\right)^2}\right] \qquad (12)$$

اذ ان $X_r(k)$ تمثل قيم تحويل فورير السريع *FFT* لقالب من عينات إشارة الكلام الأصـلية و $Y_r(k)$ تمثل تمثل قيم تحويل فورير السريع *FFT* لقالب من عينات إشارة الكلام المحسنة و r تمثل تسلسل القالب و N تمثل حجم القالب[2].

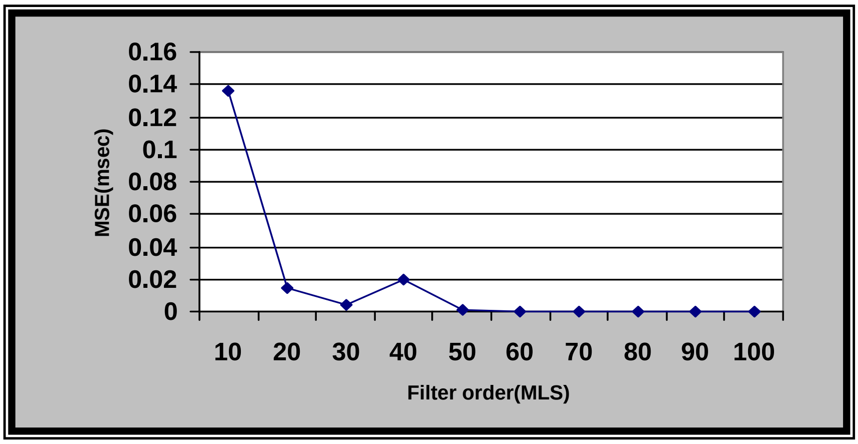
1-15 مناقشة النتائج

1- قياس الحد الأدنى لمربع الخطأ(MSE)

تم تطبيق المعادلة (11) على عدد من العينات الكلامية لقياس الحد الأدنى مربع الخطأ على لإشارة الداخلة والإشارة المحسنة من خلال إيجاد الفرق بين قيم إشارة الإدخال الأصلية وبين قيم الشارة الإخراج كما موضح في الأشكال الآتية:



الشكل(1): نتائج قياس الحد الأدنى لمربع الخطأ للعينة الكلامية الأولى باستخدام خوارزميةRLS

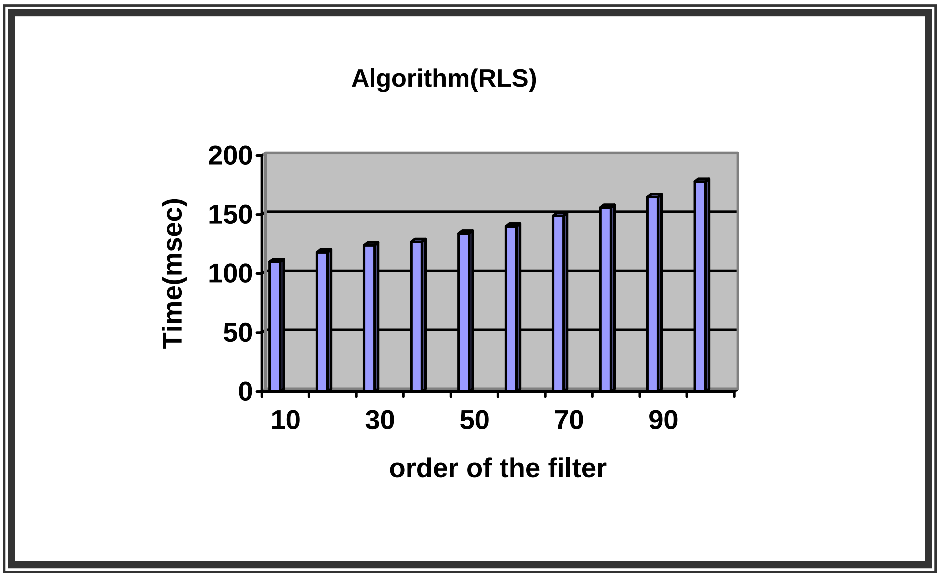


الشكل(2): قياس الحد الأدنى لمربع الخطأ للعينة الكلامية الأولى باستخدام الخوارزمية LMS

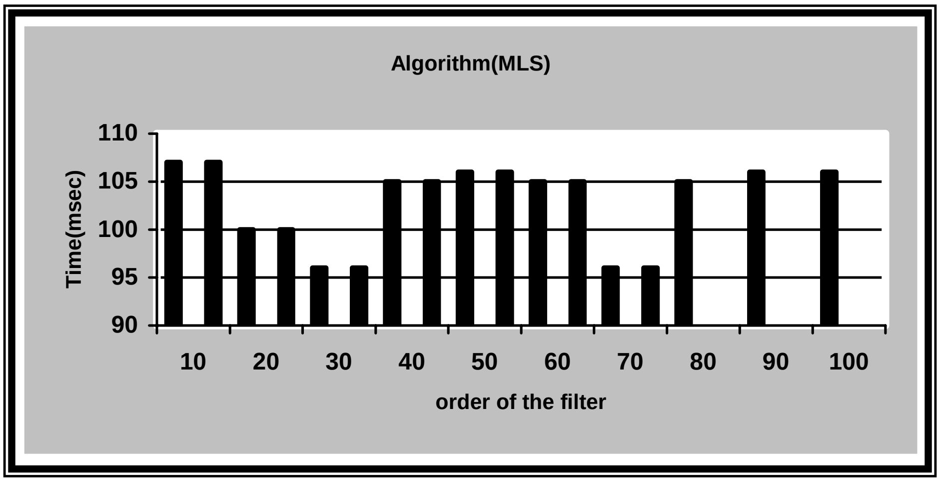
نلاحظ من الأشكال السابقة ، إن العامل الرئيسي المسيطر على مستوى قياس الحد الأدنى لمربع الخطأ هو درجة ترتيب المرشح ، آي ان افضل نتيجة للتقليل من قيمة الخطأ كانت عند ترتيب المرشح(100) للخوارزميةRLS والخوارزميةLMS .

2- قياس زمن التنفيذ

تم مقارنة درجة ترتيب المرشح مع زمن التنفيذ لكل من الخوارزمية RLS والخوارزمية LMS كما في الأشكال (4,3) زمن التنفيذ يزداد طردياً مع زيادة درجة ترتيب المرشح هذا بالنسبة لعمل الخوارزمية RLS . اما بالنسبة للخوارزمية LMS كما موضح في الشكل(4) زمن التنفيذ يتذبذب بين الزيادة والنقصان مع زيادة درجة ترتيب المرشح.



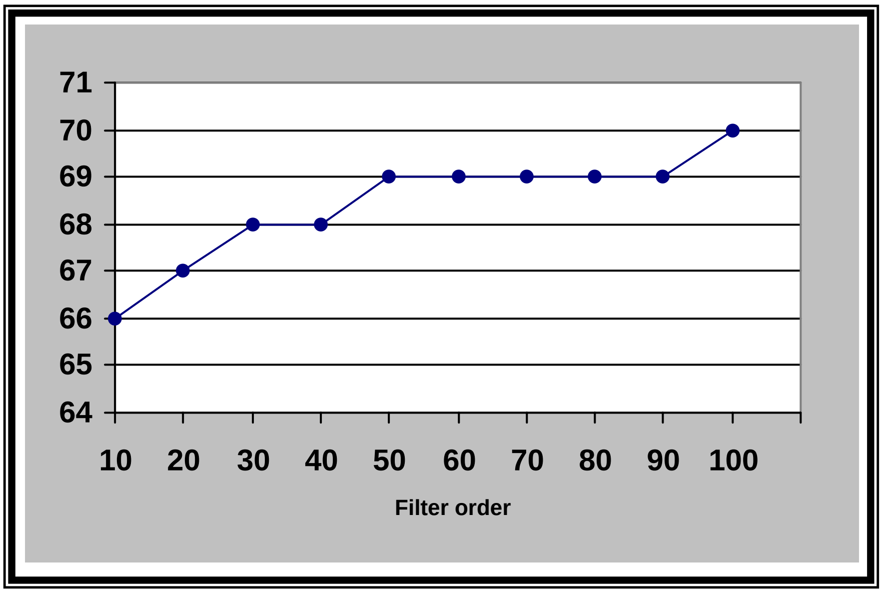
الشكل(3) مقارنة زمن التنفيذ مع درجة ترتيب المرشح باستخدام خوارزمية RLS



الشكل (4) مقارنة زمن التنفيذ مع درجة ترتيب المرشح باستخدام خوارزمية LMS

3- قياس جودة إشارة الكلام المحسنة

طبقت المعادلة (12) على إشارة الكلام الأصيلة و إشارة الكلام المحسنة لقياس مـدى التحسين على الإشارة الكلامية. وطبقت علـى الإشـارة الكلاميـة لاطـوال مختلفـة مـن القوالب(512،250،1024) كما موضح في الشكل الآتي:



الشكل(5): قياس درجة التحسين على الإشارة الكلامية ذات طول قالب(512)

يتبين من الشكال السابق ان درجة تحسين الإشارة الكلامية تعطي ناتجا افضل عند زيادة درجة ترتيب المرشح وكانت افضل نتيجة عند طوال القالب(512).

1-16 الاستنتاجات

1. من خلال دراسة خواص كل من الخوارزمية RLS و الخوارزمية LMS تبين وجود ثلاث خواص تؤثر في عمل هذه الخوارزميات وهي أولا زمن التنفيذ ، ثانياً درجة تسلسل المرشح ، ثالثاً عامل الحد الأدنى لمربع الخطأ MSE
2. أظهرت الدراسة ان درجة ترتيب المرشح تعد العامل الرئيس المؤثر من زمن تنفيذ المرشح وكذلك من التقليل من الحد الأدنى لمربع الخطأMSE.
3. تبين من النتائج ان زمن تنفيذ الخوارزمية RLS يزداد طردياً مع زيادة درجة ترتيب المرشح.
4. ان لعامل تقارب القيم (λ) تأثيرا كبيرا من زمن تنفيذ الخوارزمية، اذ يكون تأثيره كبيرا من زمن التأخير لاتمام عملية التحسين .
5. ان العامل (μ) يحافظ على استقرارية عمل المرشح طالما كانت قيمة هذا العامل اكبر من الصفر .

17-1 المصادر

[1]. ثابت، نادية معن محمد، (2000)، " محرر فيديو ال AVI " ، بحث ماجستير ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل.

[2]. حليم ، علياء موفق عبد المجيد ،(2003)، " تشفير إشارة الكلام بطريق البعثرة "، بحث ماجستير ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل.

[3]. Abdul-Majeed , E. O. ,(2000), “Voice Recognition Using Neural Net Work” M.SC. thesis , computer Science ,College of computer and Mathematics ,Mosul University.

[4]. Douglas L. Jones ,(2004), “Application Filter Based on the LMS-Algorithm”.
http:// www. pactcorp.com/FTP/technology/appl-notes/PACT-pp 006 .pdf.

[5]. Peter S. ,(2000), “ Band-Reject Filter”.
http:// www. infodotinc.com/content/nnets/1418/css/1418-51. html.

[6]. Childers,D. G. ,(2000), “ Speech Processing and Synthesis Toolboxes ”, John Wiley & Sons ,Inc.

[7]. Dunn, Bob,(2003),” Speech Signal Processing and Speech Recognition” .
http:// www.caip.rutgers.edu/~rabinkin/Dsp-no-audio.pdf.

[8]. Erik ,(2003), “IIR Filter & FIR Filters”.
http:// zone.ni.com/devzone/conceptd.nsf/webmain/A43….html.

[9]. Lau, Y.; Zahir,M.H.,(2003), “Perfomance of Adaptive Filtering Algorithms Comparative Study”, School of Electrical and Computer Engineering ,RMIT ,University Melbourne, Victoria Australia.
http:// www. dspquru.com/info/faqs/firfaq.htm

[10]. Ludeman, C.,L. ,(1986), “Fundamentals of Digital Signal Processing” ,ISBNO-471-60363-5

[11]. Lim, J.;Alan,V.;Louis,D. ,(1978), “ Evaluation of an Adaptive Comb Filtering Method for Enhancing Speech Degraded by white

Noise addition", IEEE Transaction on Acoustics ,Speech, and signal processing, Vol.Assp.26.NO.4.

[12]. Math , D.,(2000), " MATLAB The Language of Technical Computing version 6.5 September 22,2000.

[13]. Mustafa, B.,A.,(2003),"Data-Base system for Speaker identification" ,Ph.D Thesis, college of Computer Sciences & Mathematics ,University of Mosul.

[14]. Philips ,C.,(2001), " A multi-Band Spectral Subtraction Method for Speech Enhancement".
http:// www. Utdallas.edu/~Liozou/suni-ms.pdf.

[15]. Sofia,F.,B,(2003),"Mathematical Model for Arabic Speech Recognition", Computer Sciences & Mathematics ,University of Mosul Ph.D.

[16]. Ward, D.,(2002), " Project –Speech Enhancement".
http:// www. Ee.icac.uk/dward/labcourse/labscripts/project.pdf

[17]. Yonghe ,L.,(2000), " Voice and Unvoiced Decision".
http://www.owlent.rice.edu/~elec532/project500
/vocode/uv/uvdet.html