

دراسة تأثير معاملات خوارزمية الطرح الطيفي المعدلة وطول النافذة الزمنية في تحسين الإشارات الصوتية

الدكتور فائق عراج*

عادل ديوب**

(تاريخ الإيداع 7 / 6 / 2015. قُبل للنشر في 10 / 12 / 2015)

□ ملخص □

إن إلغاء الضجيج من الكلام الملوث هو حقل هندسي يدرس التقنيات المستخدمة لاستعادة الإشارة الأصلية من الإشارة الملوثة بأنواع مختلفة من الضجيج مثل الضجيج عريض الحزمة وضيق الحزمة وأنواع أخرى موجودة بالبيئة وتعتبر تقنية الطرح الطيفي الأبرز في هذا المجال. في هذا البحث سوف نناقش تأثير معاملات خوارزمية الطرح الطيفي المعدلة وطول النافذة الزمنية في تحسين الكلام الملوث بالضجيج عريض الحزمة. تم دراسة وتحديد قيم المعاملات المثالية وطول النافذة المثالي عند قيم مختلفة لنسبة الإشارة إلى الضجيج SNR للإشارة الملوثة وهي **0, 5, 10, 15 dB** وتم مناقشة 18 حالة عند كل قيمة. أجريت المحاكاة باستخدام برنامج الماتلاب (matlab) وتمت مقارنة النتائج سماعياً وبالاعتماد على تحسين قيمة SNR لكل حالة.

الكلمات المفتاحية: الطرح الطيفي ، تحسين الكلام ، نسبة الإشارة إلى الضجيج ، النافذة الزمنية ، معامل الطرح، معامل قاعدة الطيف .

* أستاذ مساعد-قسم هندسة الاتصالات-كلية الهندسة الميكانيكية والكهربائية-جامعة تشرين-اللاذقية-سورية.
**طالب دراسات عليا(ماجستير)-كلية الهندسة الميكانيكية والكهربائية-جامعة تشرين-اللاذقية-سورية.

Studying the effect of modified spectral subtraction algorithm parameters and time window length in speech signals enhancement

Dr. Faek Arraj*
Adel Dayoub**

(Received 7 / 6 / 2015. Accepted 10 / 12 / 2015)

□ ABSTRACT □

Speech denoising is a field of engineering that studies techniques used to recover the original signal from the noisy signal corrupted with different types of noise, such as broadband noise and narrowband noise, and other types present in environment, but the spectral subtraction technique consider the most prominent in this area .

In this search we will discuss the parameters impact of the modified spectral subtraction algorithm and the time window length in the enhancement of speech that corrupted with broadband noise. We done the study and determine the ideal parameters values and the ideal window length with different values for the signal -to-noise ratio SNR for noisy speech **0db, 5db, 10db, 15db** and we discuss 18 case for each value. We done the simulation using MATLAB software and the results were compared based on improving the value of SNR for each case .

Keywords: spectral subtraction , speech enhancement , Signal-to-noise ratio, time window , subtraction factor , spectral floor factor .

*Associate Professor , Department of Communications Engineering , Faculty of Mechanical and Electrical Engineering , Tishreen University , Lattakia , Syria.

**Postgraduate Student , Department of Communications Engineering , Faculty of Mechanical and Electrical Engineering , Tishreen University , Lattakia , Syria.

مقدمة:

يمكن أن يفسد الضجيج الموجود في إشارات الكلام جودة الإشارة وينقص الوضوح ويزيد إرهاق المستمع وفي الواقع يوجد عدة أنواع من الضجيج التي تؤثر على إشارة الكلام وتعد مسألة إنقاص الضجيج ضرورية في عالم الاتصالات .

تطورت الأنواع المختلفة لخوارزميات تقليل الضجيج بالاعتماد على تقنيات مجال التحويل والترشيح المنكف ومن بين تقنيات تحسين الكلام، تقنية الطرح الطيفي والتي تعد الأبرز في هذا المجال وأيضاً الطرح الطيفي المعدل والطرح الطيفي التكراري ومرشح وينر ومرشح كالمن وعتبة الضجيج وغيرها من التقنيات. بالرغم من أن خوارزمية الطرح الطيفي لها تعقيد حسابي منخفض جداً، فإنها تقلل الضجيج الخلفي بفعالية وتبين النتائج التجريبية انه يوجد بعض الضجيج المتبقي في الإشارة المعالجة والذي يؤثر في السمع.

كانت إزالة الضجيج من إشارات الكلام خلال العقود الماضية في موقع اهتمام الباحثين أثناء معالجة الكلام وارتكز العديد منهم إلى خوارزمية الطرح الطيفي وأجرى دراسة عليها وفق منحى معين وهناك أيضاً بعض الباحثين درسوا الطرق الأخرى وأجروا تعديلات عليها .

سنركز في بحثنا إلى خوارزمية الطرح الطيفي المعدلة وسندرس تأثير معاملات وطول النافذة الزمنية المستخدمة من أجل التوصل إلى أفضل القيم لها بحيث يتم إنقاص الضجيج الخلفي وزيادة تعريف إشارة الكلام . يتطلب القيام بهذه الدراسة تصميم خوارزمية الطرح الطيفي المعدلة والتي هي محور بحثنا باستخدام برنامج المحاكاة المشهور ماتلاب عن طريق اختيار البلوكات الخاصة بكل مرحلة من مراحل الخوارزمية وربطها معاً ومن ثم سنقوم باختيار عدة قيم لمعاملات هذه الخوارزمية وطول النافذة والتي سيتم التعرف عليها تفصيلاً ضمن البحث ونقارن النتائج التي سنحصل عليها بواسطة مفهوم نسبة الإشارة إلى الضجيج للإشارة المحسنة SNR وسماعياً. تظهر الدراسة أيضاً منحنيات كل من الإشارة الصوتية الأصلية و الإشارة الملوثة والإشارة المحسنة في خرج الخوارزمية.

أهمية البحث وأهدافه:

يتضمن هذا البحث دراسة لبعض الخوارزميات المستخدمة في تخفيض الضجيج مثل تقنية الطرح الطيفي والطرح الطيفي المعدل وعتبة الضجيج وستكون خوارزمية الطرح الطيفي المعدلة محور البحث. من الممكن تطبيق خوارزمية الطرح الطيفي المعدلة التي سيتم التوصل إليها بشكلها الأمثل في تطبيقات الاتصالات الصوتية وفي التخلص من الضجيج من التسجيلات الصوتية القديمة وهنا سنطبقها على إشارة صوت مسجلة واضحة وسيتم إضافة ضجيج عريض الحزمة إليها ضمن برنامج ماتلاب وسنجري مقارنة سماعية بين الصوت الملوث بالضجيج والصوت المحسن أمام عدد من المستمعين وأيضاً مقارنة رقمية لنبيين الفرق الواضح بينهما . يمكننا النجاح في هذه الدراسة من تحسين الإشارات الصوتية الملوثة بأكثر قدر ممكن وبالتالي من الممكن استثمار هذه الخوارزمية بشكلها الأمثل في تطبيقات الاتصالات الصوتية وأجهزة تقوية السمع وتطبيقات نقل الصوت عبر الانترنت (VOIP:voice over internet protocol) ونظم التعرف على الكلام وفي التخلص من الضجيج من التسجيلات الصوتية القديمة أو الملوثة بالضجيج وغيرها من التطبيقات الأخرى.

منهجية البحث:

سيتم إجراء البحث وفق المنهجية التالية:

1. تصميم خوارزمية الطرح الطيفي المعدلة باستخدام برنامج المحاكاة ماتلاب.
2. تحميل مقطع صوتي مسجل واضح وصافي إلى داخل برنامج المحاكاة وإضافة ضجيج عريض الحزمة له داخل البرنامج.
3. دراسة عدة قيم لمعاملات الخوارزمية وطول النافذة لبيان تأثيرها.
4. تحليل النتائج والاستنتاجات.

خوارزميات تخفيض الضجيج في الإشارات الصوتية:

1- خوارزمية الطرح الطيفي Spectral Subtraction Algorithm:

يمكن تصنيف نظم تحسين الكلام بطرق مختلفة مثل: تحسين الكلام مفرد القناة وتحسين الكلام متعدد القناة، وفي هذا الإطار فإن خوارزمية الطرح الطيفي تعتبر تقنية لتحسين الكلام مفردة القناة. تفترض تقنية تحسين الكلام مفردة القناة أن الضجيج ثابت (مستقر) وهي تعتبر رخيصة وسهلة البناء والتصميم، أما بالنسبة لتقنية تحسين الكلام متعددة القناة فأنها تتعامل مع الضجيج غير الثابت حيث يوجد قناة مرجعية تساعد في حساب الضجيج بطريقة فعالة وهي تعتبر تقنية باهظة الثمن ومعقدة التصميم [1].

إن الطرح الطيفي يعتمد على مفهوم أن الإشارة المحسنة يمكن استخراجها بطرح مكونات الطيف المحسوب للضجيج من طيف إشارة الدخل الملوثة. بافتراض أن الضجيج $d(n)$ يضاف إلى إشارة الكلام $x(n)$ فإن الكلام الملوث $y(n)$ يمكن كتابته بالشكل [2]:

$$y(n) = x(n) + d(n), \text{for } 0 \leq n \leq N - 1 \quad (1)$$

حيث إن n دليل الزمن، N عدد العينات.

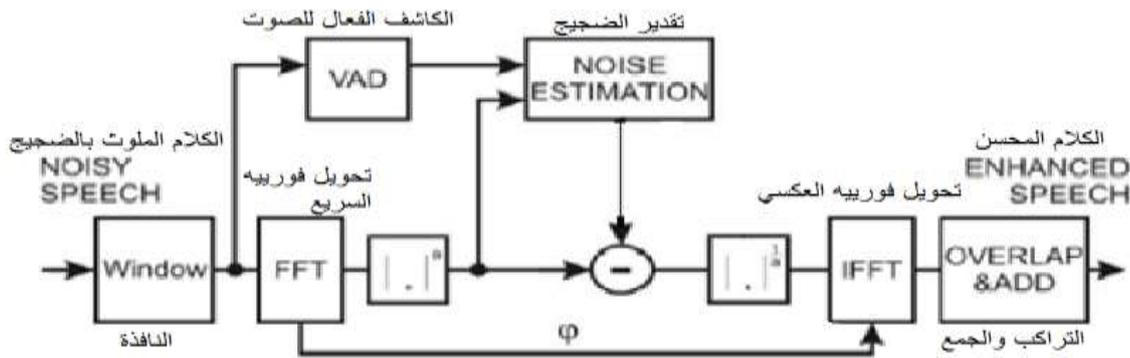
إن الغاية من تحسين الكلام هو إيجاد الإشارة المحسنة $x(n)$ من الإشارة $y(n)$ وذلك بافتراض أن $d(n)$ غير مرتبطة مع $x(n)$ إشارة الدخل $y(n)$ سوف تقطع على K مقطع بنفس الطول. إشارات المجال الزمني يمكن أن تحول إلى المجال الترددي كما يلي [2]:

$$Y_k(\omega) = X_k(\omega) + D_k(\omega), \text{for } 0 \leq k \leq K - 1 \quad (2)$$

حيث أن k هو دليل المقطع $Y_k(\omega)$, $X_k(\omega)$, $D_k(\omega)$ يدل على مقادير تحويل فورييه المنقطع DFT بالزمن القصير بأخذ $y(n)$, $x(n)$, $d(n)$ على التوالي، و $Y_k(\omega)$ مرفوع إلى القوة $a = 1$ يشير إلى الطرح الطيفي للمطالات، $a = 2$ يشير إلى الطرح الطيفي للاستطاعة). إذا كان من الممكن إيجاد طيف الضجيج \bar{D}_k فإن تقريب الكلام \bar{X}_k يمكن إيجاده من Y_k [2]:

$$\bar{X}_k(\omega) = Y_k(\omega) - \bar{D}_k(\omega) \quad (3)$$

لا يمكن حساب طيف الضجيج بدقة لكن يمكن حسابه خلال الفترة التي لا يوجد فيها كلام في إشارة الدخل. تستخدم معظم طرق الطرح الطيفي وحيدة القناة كاشف فعال للصوت (VAD: Voice Activity Detector) لتحديد فترة الصمت من أجل حساب الضجيج بدقة. يفترض أن يكون الضجيج ثابتاً لفترة قصيرة ولذلك يمكن أن يستخدم الضجيج من قطاعات الصمت لإزالة الضجيج من قطاعات الكلام. يوضح الشكل (1) المخطط الصندوقي لخوارزمية الطرح الطيفي.



الشكل (1) المخطط الصندوقي لخوارزمية الطرح الطيفي.

بعد الطرح فإن مقدار الطيف يمكن أن يكون سالباً أي أن قيمة المتوسط المطالي لطيف الضجيج تكون أكبر من مطال طيف الكلام الملوث بالضجيج. هناك بعض الإمكانيات بإزالة المكونات السالبة وعلى سبيل المثال نستخدم تقويم نصف الموجة (ضبط الأجزاء السالبة على الصفر) أو تقويم الموجة الكاملة (قيمة مطلقة). يستخدم تقويم نصف الموجة بشكل عام لكنه يقدم أصوات موسيقية في الإشارة المعالجة. إن تقويم الموجة الكاملة يتجنب إصدار ضجيج موسيقي لكنه أقل فعالية في إنقاص الضجيج وتكون معادلة الطرح الطيفي [2]:

$$|\hat{X}(\omega)|^2 = \begin{cases} |Y(\omega)|^2 - |\hat{D}(\omega)|^2, & \text{if } |Y(\omega)|^2 > |\hat{D}(\omega)|^2 \\ 0, & \text{Otherwise} \end{cases} \quad (4)$$

بعد الطرح فإن $|\hat{X}(\omega)|^{\frac{1}{2}}$ يستخرج ليؤمن مكونات مطال فورييه المرتبطة. يستخدم تحويل فورييه العكسي مع أخذ مكونات الإزاحة بشكل مباشر من وحدة تحويل فورييه، وكذلك يستخدم التراكب والجمع (overlap&add) من أجل إعادة تركيب الكلام في المجال الزمني.

إن هذه الخوارزمية لها عدة عيوب تتلخص فيما يلي:

❖ عدم إمكانية استخدام الضجيج الغير ثابت:

إن حساب طيف الضجيج مستخرج من مناطق كلام غير محلية من الإشارة الكلامية الملوثة. هذا الافتراض صحيح في حال الضجيج السابق والذي فيه طيف الضجيج لا يتغير كثيراً عبر الزمن. تراقب الكواشف التقليدية إطارات الضجيج للإشارة الملوثة لتحديث حساب الضجيج. لكن تحديث حساب الضجيج في كلا الطريقتين محدد بالإطارات التي يغيب فيها الكلام وهذا غير كافٍ في حالة الضجيج غير الثابت والذي فيه طيف طاقة الضجيج يتغير حتى خلال فعالية الكلام.

❖ الاعتماد على دقة الكاشف:

إن أداء الطرح الطيفي محدد بدقة حساب الضجيج، والذي هو بالإضافة مُحدد بأداء كواشف التوقف-الكلام. إن أداء خوارزمية إنقاص الضجيج عبر الطرح الطيفي بالإضافة إلى أداء VAD ينخفض بشكل ملحوظ عند نسب SNR المنخفضة. قد تتسبب المعالجة بإزالة بعض الأجزاء من إشارة الكلام خلال القطاعات الأضعف وذلك في حالة انقطاع قطاعات الصمت الهامة.

ذكرت بعض الأبحاث أنه يمكن استخدام الحساب المرتكز على الحجز (quantile_based estimation) لطيف الضجيج من طيف الإشارة الملوثة وذلك من أجل إنقاص الضجيج بطريقة الطرح الطيفي [3]. تركز هذه الطريقة

على ملاحظة أن طاقة الإشارة في عينة ترددية محددة تكون منخفضة في أغلب الإطارات وعالية فقط في 10-20% من الإطارات المرتبطة بقطاعات الكلام الصوتية.

❖ الضجيج الموسيقي:

بالرغم من أن طريقة الطرح الطيفي تؤمن تحسين في شروط تخميد الضجيج، فإنها تنتج غالباً نوع ضجيج متموج عشوائي جديد، مشار إليه بالضجيج الموسيقي بسبب طيفه الضيق الحزمة ووجود خصائص تشابه الرنين. هذه الظاهرة يمكن شرحها عن طريق أخطاء حساب الضجيج التي تقود إلى ذرى خاطئة في الطيف المعالج. عندما يعاد تركيب الإشارة المحسنة في المجال الزمني، فإن هذه الذرى تنتج في منحنيات جيبية صغيرة، والتي تردداتها تتغير من إطار إلى آخر وبالرغم من أن الضجيج الموسيقي مختلف جداً عن الضجيج الأصلي فإنه أحياناً يكون مزعج جداً. إن طريقة الطرح الطيفي التي سببت الضجيج الموسيقي، يمكن أن تنتج أحياناً في الإشارة جودة محسوسة ومحتوى معلومات أخفض عن الإشارة الملوثة الأصلية. معظم الأبحاث في الوقت الحالي تتركز حول الطرق الممكنة من أجل إزالة الضجيج الموسيقي.

2- خوارزمية عتبة (بوابة) الضجيج Noise Threshold (Gate) Algorithm:

تسمح بوابة الضجيج للإشارة بالمرور عبرها فقط عندما تكون هذه الإشارة فوق قيمة العتبة. إذا كانت الإشارة تحت العتبة فلا يسمح للإشارة بالمرور وتغلق البوابة. تستخدم بوابة الضجيج عندما يكون مستوى الإشارة فوق مستوى الضجيج تضبط العتبة فوق مستوى الضجيج ولذلك عندما لا يكون هناك إشارة فإن البوابة سوف تغلق [4]. تضبط العتبة (عادةً يدوياً) عالية كفاية، لذلك عندما تغيب الإشارة المطلوبة (على سبيل المثال عندما يوجد توقف بين الجمل أو الرسائل) فإنه لا توجد هسهسة خشنة في إشارة الخرج. إن العتبة لا يجب أن تضبط عالية كثيراً لأن الإشارة المطلوبة تتأثر عندها. إذا كانت الإشارة المستقبلية (الدخل) تحت العتبة، فإنه من المفترض أن تحتوي على الضجيج فقط، ومستوى إشارة الخرج ينقص أو "gated" (يعني أنه غير مسموح للإشارة بالمرور عبرها). هذه العملية يمكن أن تجعل صوت الإشارة المستقبلية بعض الشيء أقل ضجيجاً لأن الهسهسة تختفي خلال التوقف بين الكلمات أو الجمل، لكنها ليست فعالة عملياً.

إن هذه الخوارزمية لها عدة عيوب تتلخص فيما يلي [5]:

- 1- ربما من غير الممكن ضمان أنه لا توجد أية إشارة تحت العتبة، فإذا كانت موجودة سوف يحدث قطع عندما تغلق البوابة.
- 2- لا تستطيع دارة الحساسية والمفتاح التصرف بشكل آني، لذلك فإن الإشارة التي تتسبب بفتح البوابة سوف تكون دائماً مقصوفة لبعض المدى وهذا قد يكون حاد جداً إذا كانت استجابة البوابة بطيئة.
- 3- لا تزال هذه الخوارزمية الضجيج في حال وجود الإشارة المطلوبة (خلال الفجوات فقط).
- 4- إذا تجاوزت الإشارة الملوثة العتبة فإن البوابة مفتوحة ويسمح للإشارة بالمرور عبرها. لذلك فإن البوابة ربما تفتح إذا حدث طفحان مفاجئ للضجيج أو بعض الأصوات العالية والتي تجعل مستوى الإشارة يتجاوز العتبة. في تلك الحالة تكون جودة إشارة الخرج جيدة فقط إذا كانت الإشارة قوية بما فيه الكفاية لتقتنع بالضجيج. لذلك السبب هذه الطريقة تعمل فقط في حالة النسب العالية للإشارة إلى الضجيج.

3- خوارزمية الطرح الطيفي المعدلة:

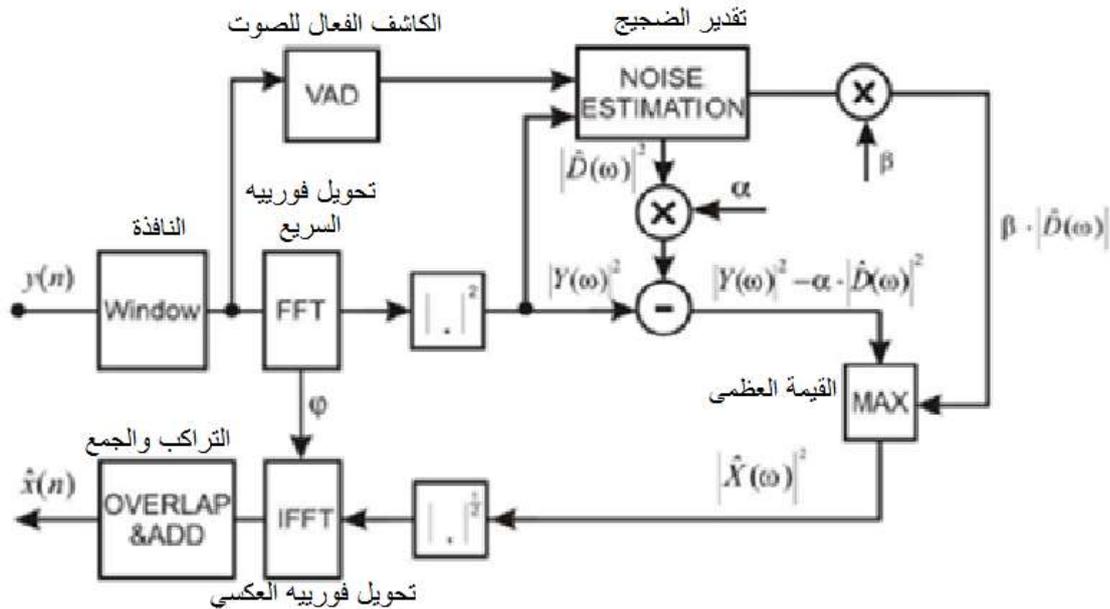
يتم إنقاص الضجيج بطريقة الطرح الطيفي ولكن يبقى بعض الضجيج عريض الحزمة (الضجيج الموسيقي) في الكلام المعالج والذي يجب أخذه بعين الاعتبار حيث أن طيف الضجيج المتبقي يتألف من قمم (peaks) وهبوطات بأحداث عشوائية. لحل هذه المشكلة قدمت خوارزمية الطرح الطيفي المعدلة بارامترين إضافيين إلى خوارزمية الطرح الطيفي وهما عامل الطرح α وعامل قاعدة الطيف β . عند اختيار عامل طرح α أكبر من الواحد فإنه يُنقص القمم عريضة الحزمة في الضجيج المتبقي ولكنه يمكن أن يؤدي إلى هبوطات عميقة مسبباً الضجيج الموسيقي الذي يؤثر سلباً على جودة الصوت. يمكن تقنيّة الضجيج الموسيقي عن طريق عامل قاعدة الطيف β ولذلك فهذان العاملان يقدمان مرونة أكبر في الخوارزمية [6].

طريقة الطرح الطيفي باستخدام عامل الطرح وقاعدة الطيف تحاول إنقاص انحراف الطيف من أجل تحسين

جودة الكلام. هذه التقنية المقترحة بالشكل (2) يمكن التعبير عنها كما يلي [2]:

$$|\hat{X}(\omega)|^2 = \begin{cases} |Y(\omega)|^2 - \alpha \cdot |\hat{D}(\omega)|^2, & \text{if } |Y(\omega)|^2 - \alpha \cdot |\hat{D}(\omega)|^2 > \beta \cdot |\hat{D}(\omega)|^2 \\ \beta \cdot |\hat{D}(\omega)|^2, & \text{Otherwise} \end{cases} \quad (5)$$

حيث أن $0 \leq \alpha$ و $\beta \gg 1$. إن قسوة الطرح يمكن أن تتغير بتغيير قيم عامل الطرح، عند قيم α الأعلى من 1 ينتج مستوى SNR عالي من أجل الإشارة التي لا تحتوي ضجيج، لكن القيم المرتفعة جداً ربما تسبب تشويهاً في جودة الكلام ولذلك فإن قيمة α يجب اختيارها بدقة لتجنب الضجيج الموسيقي وتشويه الإشارة. إن تقديم قاعدة الطيف β يمنع مكونات الطيف للكلام المحسن من النزول تحت الحزمة الدنيا وبذلك يتم إملاء الهبوطات العميقة حول القمم الضيقة (من الطيف المحسن). إن إنقاص انحراف الطيف لقمم الضجيج (عندما يتم ضبط المكونات السالبة إلى الصفر) سوف ينقص كمية الضجيج الموسيقي.

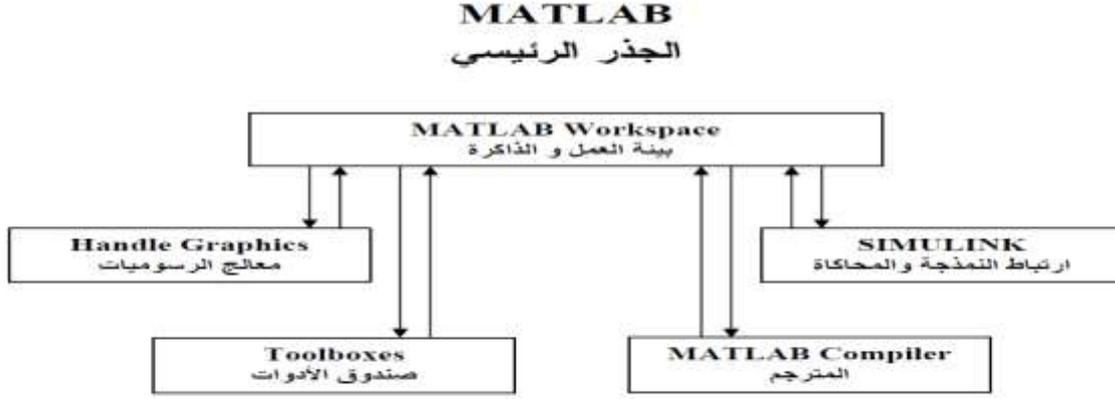


الشكل (2) المخطط الصندوقي للطرح الطيفي باستخدام عامل الطرح α وقاعدة الطيف β .

النتائج والمناقشة :

1- برنامج المحاكاة ماتلاب (MATLAB):

يعد ماتلاب (MATLAB:Matrix-Laboratory) برنامجاً رائداً في التطبيقات الهندسية والرياضية من إنتاج شركة ماثوروكس (mathworks) [7]. إن الارتباطات والتفرعات المهمة لبرنامج ماتلاب مبينة بالشكل (3):



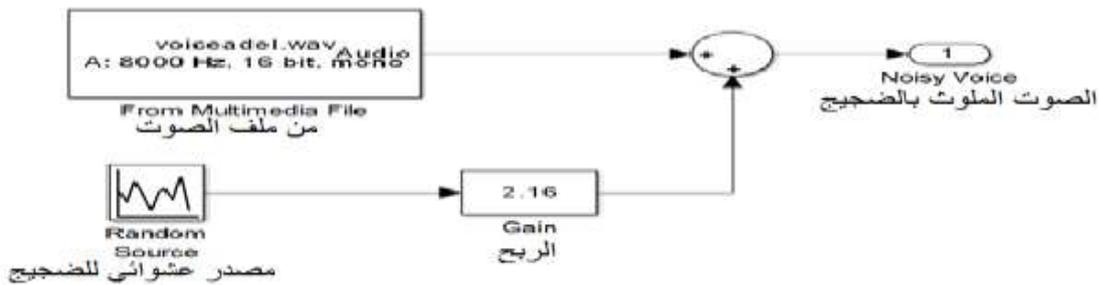
الشكل (3) يبين الارتباطات المهمة لبرنامج ماتلاب.

يعد هذا البرنامج بيئة تطوير برمجية مخصصة للمهام الحسابية، حيث تتوفر فيه الكثير من الوظائف والدوال الرياضية المبنية داخليا والتي تسهل حل مختلف أنواع المعادلات الرياضية، وذلك بالإضافة للعديد من المميزات الأخرى به ونذكر من استعمالات البرنامج على سبيل المثال لا الحصر ما يلي من مجالات: الرياضيات و الحساب، تحليل واستكشاف وتصوير البيانات، الرسوم الهندسية والبيانية وإنشاء واجهات المستخدم الرسومية والتواصل مع البرامج المكتوبة بلغات أخرى بما في ذلك C, C++ ، جافا [8].

2- بناء نموذج (model) خوارزمية الطرح الطيفي المعدلة باستخدام الماتلاب:

سيتم إجراء المناقشة والحصول على النتائج باستخدام برنامج MATLAB R2014a حيث سنبنّي نموذج (model) محاكاة لخوارزمية الطرح الطيفي المعدلة المبينة في الشكل (2) باستخدام البلوكات المناسبة كما هو مبين فيما يلي:

في البداية لدينا مقطع صوتي بطول زمني معين وبتردد 8000Hz مأخوذ بقناة أحادية حيث نضيف إليه ضجيج وليكن ضجيج عشوائي بتوزيع منتظم فنحصل على إشارة صوتية ملوثة بالضجيج (noisy voice) ستكون دخل لبرنامجنا كما هو مبين في الشكل (4).



الشكل (4) يبين كيفية توليد الإشارة الصوتية الملوثة بالضجيج.

نقوم بعد ذلك بتخزين عينات الصوت في مخزن معطيات (buffer) بطول معين، ومن ثم نمرر هذه العينات على نافذة زمنية معينة وهنا لابد من الإشارة إلى أن تابع النافذة هو تابع رياضي يكون له قيمة الصفر خارج فاصل معين يتم اختياره وتستخدم النوافذ بشكل عام في تصميم المرشحات وفي أنظمة معالجة الكلام وفي التحليل الطيفي للإشارات [9]. وكل نافذة تكون موصوفة بعرض الفص الرئيسي ومستوى الفص الجانبي حيث يفضل أن يكون الفص الرئيسي ضيق ومستوى الفصوص الجانبية منخفض [10].

إن طول النافذة يتم قياسه بالنقاط (1 point=1 sample) ومدة النافذة تعتمد على عدد العينات (طول النافذة) وكذلك معدل التقطيع للإشارة وذلك وفق العلاقة :

$$\text{window duration} = \frac{\text{window length}}{\text{sampling frequency}} \quad (6)$$

وتبين النتائج التجريبية أن الخيار المثالي لمدة نافذة التحليل هي 15 – 35 ms عند إعادة بناء إشارة الكلام [11].

بالنسبة لدراستنا سنستخدم نافذة هامينغ (hamming) والتي لها الشكل الرياضي [12]:

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N}\right) \quad (7)$$

N : تمثل عدد العينات في القطاع ، n : تمثل دليل العينة.

وستتم الدراسة على قيمتين لطول النافذة الأولى بطول 128 نقطة والثانية بطول 256 نقطة وبالنسبة لخيارنا الأول 128 نقطة ستكون مدة النافذة كما يلي:

$$\text{window duration} = \frac{\text{window length}}{\text{sampling frequency}} = \frac{128}{8000} = 0.016 \text{sec} = 16 \text{ms}$$

وهو يحقق المجال السابق أكبر من 15ms وأصغر من 35ms .

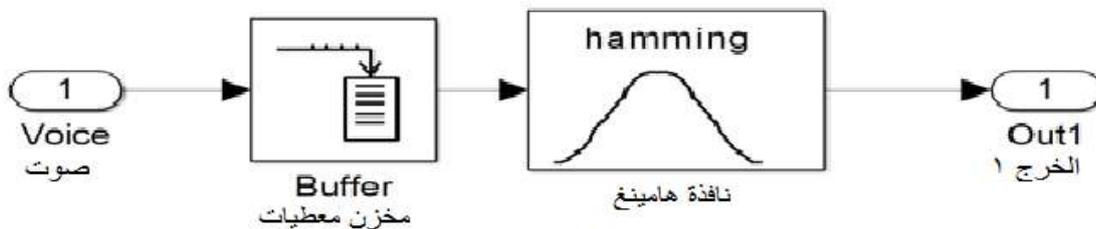
وبالنسبة لخيارنا الثاني 256 نقطة ستكون مدة النافذة كما يلي:

$$\text{window duration} = \frac{\text{window length}}{\text{sampling frequency}} = \frac{256}{8000} = 0.032 = 32 \text{ms}$$

وهو يحقق أيضاً المجال السابق أكبر من 15ms وأصغر من 35ms .

وبالتالي ستتم دراسة كل من الحالتين والمقارنة بينهما وفي النهاية نعود ونقوم بتجميع العينات جميعها سويةً

ويوضح الشكل (5) كيفية تطبيق نافذة هامينغ في الخوارزمية .



الشكل (5) يبين عملية تخزين العينات وتطبيق نافذة هامينغ.

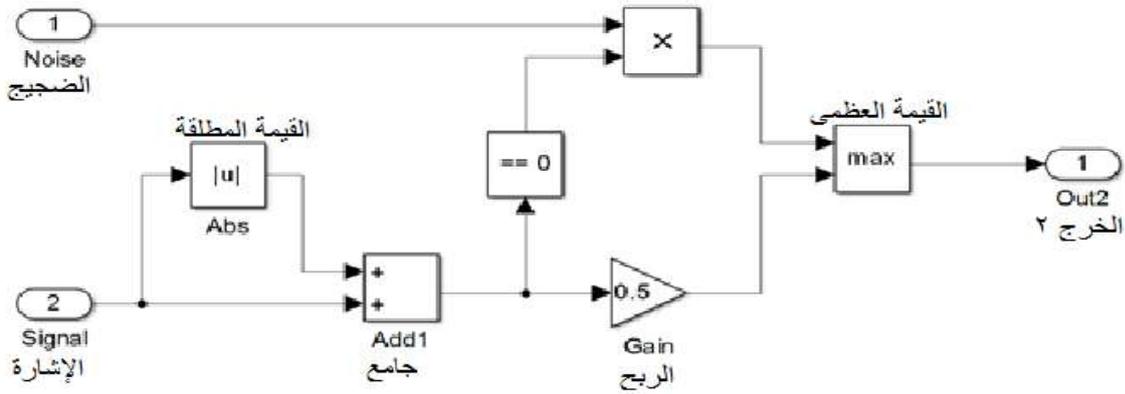
ثم نقوم بتمرير إشارة الخرج الناتجة على كتلة تسمى الكاشف الفعال للصوت (VAD: Voice Activity Detector) وذلك لتحديد فترات الصمت وفترات الكلام في الإشارة ويكون خرج الكاشف 1 في حالة الصمت و 0

في حالة وجود كلام ويفيد هذا الكاشف في حساب القيمة المتوسطة للإشارة الملوثة بالضجيج في لحظات الصمت والتي تمثل في الحقيقة القيمة المتوسطة للضجيج .

نقوم الآن بأخذ تحويل فورييه السريع لإشارة الخرج الناتجة في الشكل (5) وبعد حساب القيمة المتوسطة للضجيج نقوم بعملية الطرح الطيفي المعدلة أي تطبيق للمعادلة (5) التالية :

$$|\hat{X}(\omega)|^2 = \begin{cases} |Y(\omega)|^2 - \alpha \cdot |\hat{D}(\omega)|^2, & \text{if } |Y(\omega)|^2 - \alpha \cdot |\hat{D}(\omega)|^2 > \beta \cdot |\hat{D}(\omega)|^2 \\ \beta \cdot |\hat{D}(\omega)|^2, & \text{Otherwise} \end{cases}$$

قد ينتج لدينا بعد عملية الطرح الطيفي عينات في الإشارة لها طولية سالبة لذلك نحتاج لبناء كتلة تسمى تقويم نصف الموجة كما هو موضح بالشكل (6) وذلك للحفاظ على الطويلة الموجبة وضبط الطويلة السالبة إلى الصفر حيث تمثل هذه اللحظات التي تكون فيها الطويلة سالبة ترددات معينة يكون فيها الكلام له طولية أقل من الضجيج [13].



الشكل (6) يبين عملية تقويم نصف الموجة.

نقوم بعد ذلك بعملية تحويل فورييه السريع العكسي للخرج الأخير ومن ثم نقوم ببناء المقطع الصوتي من جديد من خلال كتلة تسمى التراكم والجمع (Overlap And Add) وذلك لمنع وجود انقطاعات في الإشارة ونحصل في النهاية على الإشارة الصوتية المحسنة (enhanced voice) التي يمكننا الاستماع إليها .

الآن يكون لدينا النموذج الكامل لخوارزمية الطرح الطيفي المعدلة جاهز ويمكننا إجراء الدراسة عليه.

3- دراسة تحسين أداء خوارزمية الطرح الطيفي المعدلة:

تلعب معاملات خوارزمية الطرح الطيفي المعدلة والتي تم التعرف إليها سابقاً وكذلك طول النافذة المستخدمة دوراً أساسياً في مقدار التحسين الذي يمكن أن تقدمه هذه الخوارزمية من حيث تخفيض الضجيج في الإشارة الملوثة بالضجيج وسوف نحري فيما يلي دراسة تفصيلية ومعقدة لتأثير معاملات عامل الطرح α وقاعدة الطيف β و طول النافذة في تحسين أداء خوارزمية الطرح الطيفي المعدلة ويظهر الشكل (2) المخطط الصندوقي لهذه الخوارزمية وتعبير المعادلة (5) عن هذه الخوارزمية رياضياً .

بعد بناء نموذج خوارزمية الطرح الطيفي المعدلة وفق الخطوات المذكورة سابقاً سنقوم الآن بتقييم أداء الخوارزمية عند قيم مختلفة لنسبة الإشارة إلى الضجيج (مقدرة بالديسبل) في الدخل SNR_{INPUT} والتي يمكن حسابها رياضياً كما يلي:

$$SNR_{INPUT} = 10 \log_{10} \frac{|x(n)|^2}{|d(n)|^2} \quad (8)$$

وبالتالي فإن نسبة الإشارة إلى الضجيج (مقدرة بالديسبل) للإشارة المحسنة في الخرج SNR_{OUTPUT} تكون [14]:

$$SNR_{OUTPUT} = 10 \log_{10} \frac{|x(n)|^2}{|\bar{d}(n)|^2} \quad (9)$$

$\bar{d}(n)$: يمثل كمية الضجيج المتبقي في الخرج ويحسب كما يلي:

$$\bar{d}(n) = \bar{x}(n) - x(n) \quad (10)$$

$\bar{x}(n)$: تمثل الإشارة المحسنة في الخرج النهائي .

وبالتالي فإن قيمة التحسين (مقدرة بالديسبل) في نسبة الإشارة إلى الضجيج $SNR_{IMPROVEMENT}$ تكون:

$$SNR_{IMPROVEMENT} = SNR_{OUTPUT} - SNR_{INPUT} \quad (11)$$

وهي تمثل الفرق في الديسبل بين الخرج والدخل.

سوف نقيم أداء الخوارزمية عند أربع حالات تمثل أربعة مستويات لنسبة الإشارة إلى الضجيج في الدخل وهي

0db, 5db, 10db, 15db وفي كل حالة سوف ندرس القيم التالية للبارامترات:

❖ طول النافذة، سوف نناقش قيمتين لها وهي: 128 عينة و 256 عينة .

❖ عامل المعايرة α ، سوف نناقش ثلاثة قيم له وهي: 3 , 5 , 7 .

❖ قاعدة الطيف β ، سوف نناقش ثلاثة قيم له وهي: 0.01 , 0.03 , 0.07 .

أي أنه من أجل كل حالة سيكون لدينا 18 قيمة في الخرج .

A- الحالة الأولى: عندما $SNR_{INPUT} = 0db$ سوف تنتج لدينا 18 قيمة سيجري مقارنتها وفق الجدول (1)

التالي :

الجدول (1) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عندما $SNR_{INPUT} = 0db$.

window length	α	β	$SNR_{INPUT}(db)$	$SNR_{OUTPUT}(db)$	$SNR_{IMPROVEMENT}(db)$
128	3	0.01000	-0.0017722561	6.5127397	6.5145121
128	3	0.03000	-0.018369164	6.2273917	6.2457609
128	3	0.07000	0.015829895	5.7571979	5.7413678
128	5	0.01000	-0.0040178071	7.5441375	7.5481553
128	5	0.03000	-0.017999003	7.2477527	7.2657518
128	5	0.07000	-0.0093891462	6.6838999	6.6932888
128	7	0.01000	0.0018364800	7.5804968	7.5786605
128	7	0.03000	-0.023873080	7.3586755	7.3825483
128	7	0.07000	-0.026628044	6.5412798	6.5679078
256	3	0.01000	0.0042721890	7.0570455	7.0527735
256	3	0.03000	0.0069401748	6.8394508	6.8325105
256	3	0.07000	-0.016513469	6.2758818	6.2923951

256	5	0.01000	-0.0042494480	8.4270124	8.4312620
256	5	0.03000	-0.0084856804	8.2016268	8.2101126
256	5	0.07000	-0.0041211890	7.5720577	7.5761790
256	7	0.01000	-0.012901556	8.3421068	8.3550081
256	7	0.03000	-0.015305614	8.0586700	8.0739756
256	7	0.07000	-0.0064444314	7.4414926	7.4479370

بالمقارنة بين هذه النتائج التي حصلنا عليها نجد أن أفضل تحسين لنسبة الإشارة إلى الضجيج هو

$SNR_{IMPROVEMENT} = 8.431 \text{ db}$ وكانت قيم البارامترات كما يلي:

$window \text{ length} = 256 \text{ points}$, $\alpha = 5$, $\beta = 0.01$

B- الحالة الثانية: عندما $SNR_{INPUT} = 5 \text{ db}$ سوف تنتج لدينا 18 قيمة سيجري مقارنتها وفق الجدول (2)

التالي :

الجدول (2) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عندما $SNR_{INPUT} = 5 \text{ db}$.

window length	α	β	$SNR_{INPUT}(\text{db})$	$SNR_{OUTPUT}(\text{db})$	$SNR_{IMPROVEMENT}(\text{db})$
128	3	0.01000	5.0259519	10.449844	5.4238925
128	3	0.03000	5.0170469	10.070830	5.0537834
128	3	0.07000	5.0298624	9.4864931	4.4566307
128	5	0.01000	5.0232706	11.179710	6.1564398
128	5	0.03000	5.0241952	10.758657	5.7344623
128	5	0.07000	5.0254679	10.024745	4.9992771
128	7	0.01000	5.0443959	10.719137	5.6747413
128	7	0.03000	5.0194206	10.738620	5.7191992
128	7	0.07000	5.0380158	10.045777	5.0077615
256	3	0.01000	5.0426655	10.907331	5.8646650
256	3	0.03000	5.0237117	10.527061	5.5033498
256	3	0.07000	5.0437317	9.9782562	4.9345245
256	5	0.01000	5.0107851	11.786480	6.7756948
256	5	0.03000	4.9969907	11.448062	6.4510713
256	5	0.07000	5.0215893	10.855380	5.8337908
256	7	0.01000	5.0139394	11.569929	6.5559897
256	7	0.03000	4.9988213	11.356795	6.3579741
256	7	0.07000	5.0440707	10.808846	5.7647748

بالمقارنة بين هذه النتائج التي حصلنا عليها نجد أن أفضل تحسين لنسبة الإشارة إلى الضجيج هو
 $SNR_{IMPROVEMENT} = 6.775 \text{ db}$ وكانت قيم البارامترات كما يلي:

$$window \text{ length} = 256 \text{ points} , \quad \alpha = 5 , \quad \beta = 0.01$$

C- الحالة الثالثة: عندما $SNR_{INPUT} = 10 \text{ db}$ سوف تنتج لدينا 18 قيمة سيجري مقارنتها وفق
 الجدول (3) التالي :

الجدول (3) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عندما $SNR_{INPUT} = 10 \text{ db}$.

window length	α	β	$SNR_{INPUT} \text{ (db)}$	$SNR_{OUTPUT} \text{ (db)}$	$SNR_{IMPROVEMENT} \text{ (db)}$
128	3	0.01000	9.9709721	14.154403	4.1834307
128	3	0.03000	9.9706707	13.682008	3.7113371
128	3	0.07000	9.9601383	12.969029	3.0088911
128	5	0.01000	9.9766140	14.598582	4.6219683
128	5	0.03000	9.9584856	14.222486	4.2639999
128	5	0.07000	9.9757128	13.394598	3.4188852
128	7	0.01000	9.9219160	14.396235	4.4743195
128	7	0.03000	9.9885769	14.120447	4.1318703
128	7	0.07000	9.9749460	13.313554	3.3386078
256	3	0.01000	10.050386	14.472695	4.4223089
256	3	0.03000	10.040423	14.114988	4.0745649
256	3	0.07000	10.039216	13.491342	3.4521255
256	5	0.01000	10.048695	15.055617	5.0069227
256	5	0.03000	10.044240	14.684235	4.6399946
256	5	0.07000	10.044227	14.025397	3.9811707
256	7	0.01000	10.023160	14.859884	4.8367243
256	7	0.03000	10.015090	14.587923	4.5728331
256	7	0.07000	10.005301	14.036354	4.0310526

بالمقارنة بين هذه النتائج التي حصلنا عليها نجد أن أفضل تحسين لنسبة الإشارة إلى الضجيج هو
 $SNR_{IMPROVEMENT} = 5.006 \text{ db}$ وكانت قيم البارامترات كما يلي:

$$window\ length = 256\ points , \quad \alpha = 5 , \quad \beta = 0.01$$

D-الحالة الرابعة: عندما $SNR_{INPUT} = 15db$ سوف تنتج لدينا 18 قيمة سيجري مقارنتها وفق الجدول (4) التالي :

الجدول (4) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عندما $SNR_{INPUT} = 15db$.

window length	α	β	SNR_{INPUT} (db)	SNR_{OUTPUT} (db)	$SNR_{IMPROVEMENT}$ (db)
128	3	0.01000	15.051100	17.492256	2.4411564
128	3	0.03000	15.072814	17.020876	1.9480619
128	3	0.07000	15.070881	16.155840	1.0849590
128	5	0.01000	15.061081	17.777945	2.7168636
128	5	0.03000	15.071916	17.362122	2.2902060
128	5	0.07000	15.084104	16.431677	1.3475733
128	7	0.01000	15.038366	17.670353	2.6319866
128	7	0.03000	15.055021	17.225491	2.1704693
128	7	0.07000	15.064513	16.443083	1.3785696
256	3	0.01000	14.862667	17.451820	2.5891533
256	3	0.03000	14.867938	17.101263	2.2333250
256	3	0.07000	14.873667	16.412672	1.5390053
256	5	0.01000	14.871689	17.899401	3.0277119
256	5	0.03000	14.844844	17.490868	2.6460238
256	5	0.07000	14.847187	16.742594	1.8954067
256	7	0.01000	14.877882	17.625149	2.7472668
256	7	0.03000	14.851386	17.486532	2.6351461
256	7	0.07000	14.869222	16.766045	1.8968229

بالمقارنة بين هذه النتائج التي حصلنا عليها نجد أن أفضل تحسين لنسبة الإشارة إلى الضجيج هو $SNR_{IMPROVEMENT} = 3.027\ db$ وكانت قيم البارامترات كما يلي:

$$window\ length = 256\ points , \quad \alpha = 5 , \quad \beta = 0.01$$

وبالتالي نستنتج من الحالات الأربعة السابقة التي تم مناقشتها:

1. من أجل تحسين أداء خوارزمية الطرح الطيفي المعدلة يفضل أن يكون طول النافذة 256 نقطة $window\ length = 256\ points$ وقد لاحظنا أن أداء الخوارزمية يقل عندما يكون طول النافذة 128 نقطة $window\ length = 128\ points$.
2. إن أفضل قيمة لعامل المعايرة هي $\alpha = 5$ حيث تقدم هذه القيمة أكبر تحسن في أداء الخوارزمية ويفضل أن لا تزيد هذه القيمة عن 7 وأن لا تقل عن 5 حيث يقل أداء الخوارزمية عند ذلك.
3. تعطي الخوارزمية أفضل أداء عندما تكون قيمة معامل قاعدة الطيف هي $\beta = 0.01$ ويفضل أن لا تزيد قيمة β عن 0.01 حيث يهبط أداء الخوارزمية بشكل ملحوظ .
ولتحديد مجال أدق لقيمة β سندرس قيم أصغر من 0.01 وذلك بعد معرفتنا المسبقة أن أفضل قيمة لطول النافذة هي $window\ length = 256\ points$ وأن أفضل قيمة لعامل المعايرة هي $\alpha = 5$ كما هو موضح بالجدول (5) التالي:

الجدول (5) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عند قيم مختلفة لـ β .

window length	α	β	$SNR_{INPUT}(db)$	$SNR_{OUTPUT}(db)$	$SNR_{IMPROVEMENT}(db)$
256	5	0	0.023606373	0.80173075	0.77812439
256	5	0.01	-0.033060756	8.4805679	8.5136290
256	5	0.0001	-0.036533110	8.4478493	8.4843826
256	5	0.0000001	-0.010039301	8.4247026	8.4347420

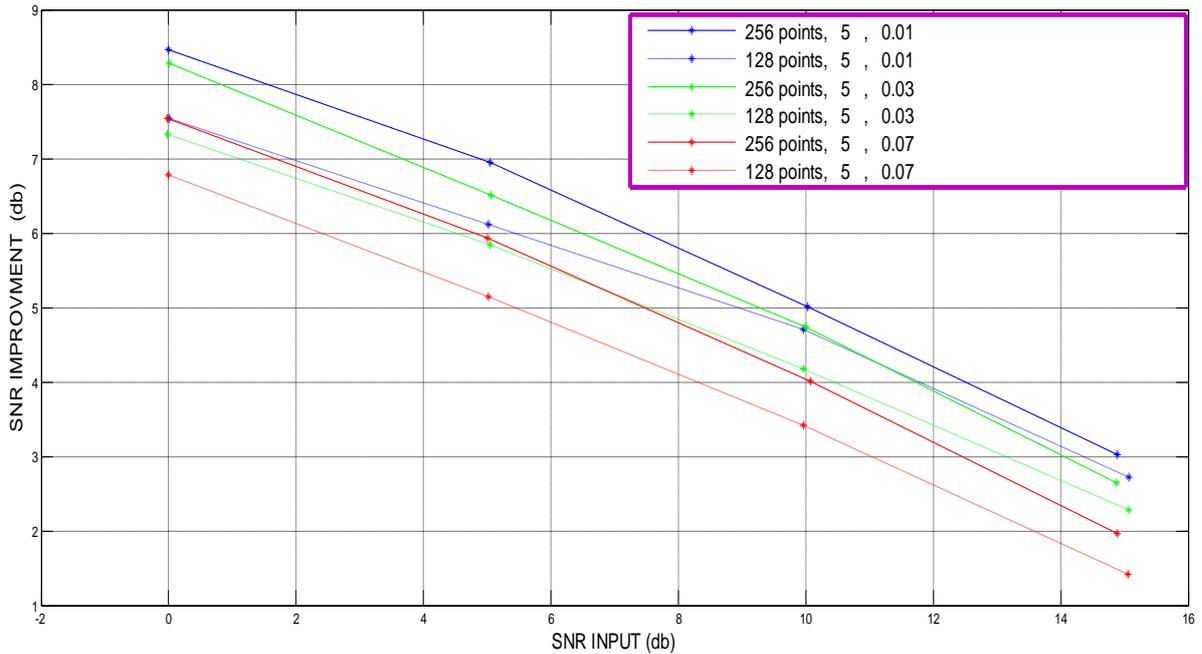
نستنتج من هذا الجدول أن الخوارزمية تعطي أداء متقارب من أجل جميع قيم β الأصغر من 0.01 ولكن لا يجوز أن تكون قيمة معامل قاعدة الطيف مساوية تماما للصفر حيث يهبط أداء الخوارزمية كلياً. سندرس الآن أداء الخوارزمية عند القيم العالية لنسبة الإشارة إلى الضجيج في الدخل فمثلاً $SNR_{INPUT} = 20\ db, 21\ db$ وتكون النتائج كما هو موضح بالجدول (6):

الجدول (6) يبين مقدار تحسين نسبة الإشارة إلى الضجيج عندما $SNR_{INPUT} = 20, 21\ db$.

window length	α	β	$SNR_{INPUT}(db)$	$SNR_{OUTPUT}(db)$	$SNR_{IMPROVEMENT}(db)$
256	5	0.01	20.23	20.17	-0.05
256	5	0.01	21.13	20.47	-0.66

نستنتج من هذا الجدول أن الخوارزمية تعطي أداء جيد حتى نسبة إشارة إلى ضجيج في الدخل $SNR_{INPUT} = 20\ db$ أي أن الإشارة تكون حينها أكبر من الضجيج بعشر أضعاف وبعدها يسوء أدائها وهذا منطقي حيث أنه لا داعي حينها لاستخدام خوارزميات تخفيض ضجيج حيث يقوم الصوت الأصلي بتقنيع الضجيج .
الآن سنجري المقارنة سماعياً حيث سيتم جلب ثلاثة مستمعين ذكور وثلاثة مستمعين إناث لتقييم أداء الخوارزمية عند القيم السابقة التي توصلنا إليها $SNR_{INPUT} = 20, 21\ db$, $\alpha = 5$, $\beta = 0.01$, $window\ length = 256\ points$ وقد لوحظ الفرق الواضح بين الإشارة الملوثة والإشارة المحسنة.

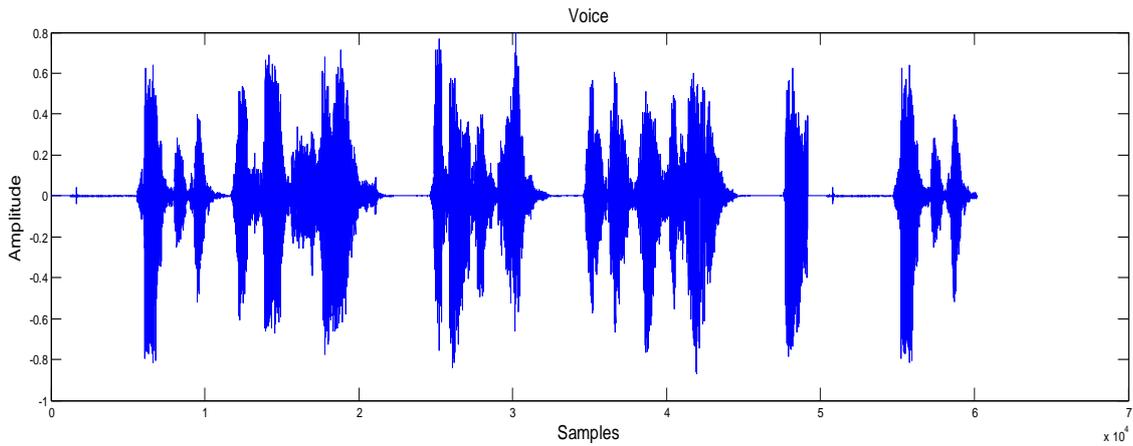
- بعد ذلك سنقوم بأجراء مقارنة بواسطة المنحنيات المبينة بالشكل (7) التي تمثل التحسين في نسبة الإشارة إلى الضجيج $SNR_{IMPROVEMENT}$ بالمقارنة مع نسبة الإشارة إلى الضجيج في الدخل عند القيم التالية
- $SNR_{INPUT} = 0db, 5db, 10db, 15db$ حيث سندرس ستة حالات هي :
1. $window\ length = 256\ points$, $\alpha = 5$, $\beta = 0.01$
 2. $window\ length = 128\ points$, $\alpha = 5$, $\beta = 0.01$
- حيث يمثل الخط الأزرق المستمر الحالة الأولى ويمثل الخط الأزرق المنقط الحالة الثانية .
3. $window\ length = 256\ points$, $\alpha = 5$, $\beta = 0.03$
 4. $window\ length = 128\ points$, $\alpha = 5$, $\beta = 0.03$
- حيث يمثل الخط الأخضر المستمر الحالة الثالثة ويمثل الخط الأخضر المنقط الحالة الرابعة .
5. $window\ length = 256\ points$, $\alpha = 5$, $\beta = 0.07$
 6. $window\ length = 128\ points$, $\alpha = 5$, $\beta = 0.07$
- حيث يمثل الخط الأحمر المستمر الحالة الخامسة ويمثل الخط الأحمر المنقط الحالة السادسة.
- ونلاحظ بوضوح أن أفضل تحسين في أداء الخوارزمية تم في الحالة الأولى
- $window\ length = 256\ points$, $\alpha = 5$, $\beta = 0.01$ والمبينة باللون الأزرق المستمر .



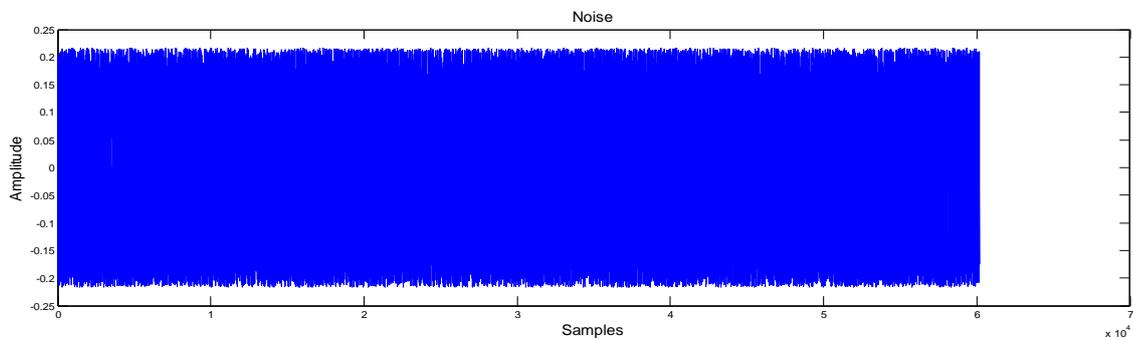
الشكل (7) يبين منحنيات تحسين نسبة الإشارة إلى الضجيج.

الآن ومن أجل القيم التي تحقق أفضل أداء للخوارزمية

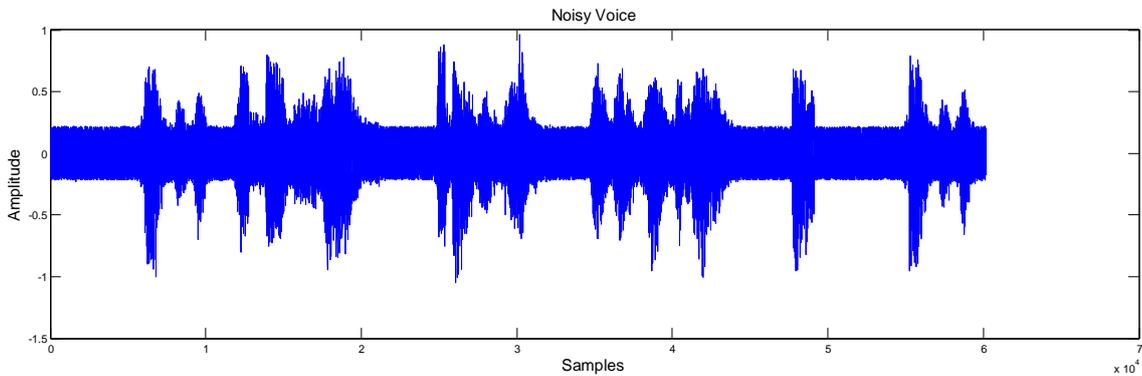
- $window\ length = 256\ points$, $\alpha = 5$, $\beta = 0.01$ سوف نقوم برسم المنحنيات التي تمثل مطالات كل من إشارة الصوت الأصلية المبينة بالشكل (8) والضجيج المضاف لها في الشكل (9) ثم الإشارة الملوثة بالضجيج في الشكل (10) والضجيج المتبقي في الشكل (11) والإشارة المحسنة في الشكل (12) .



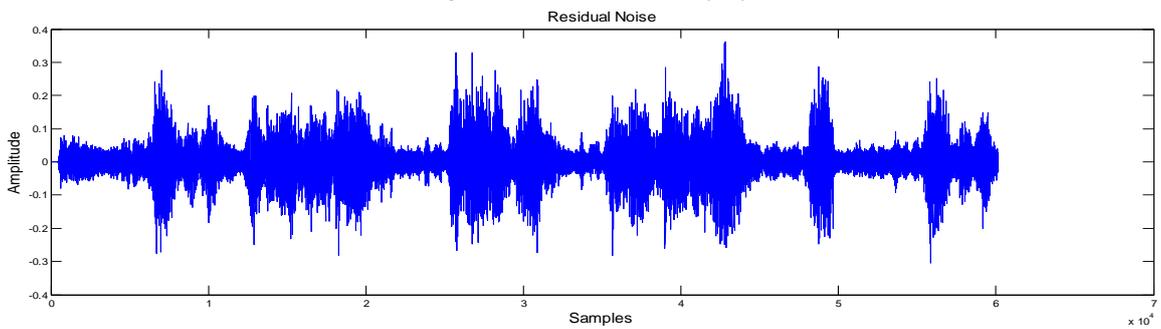
الشكل (8) يبين الإشارة الأصلية.



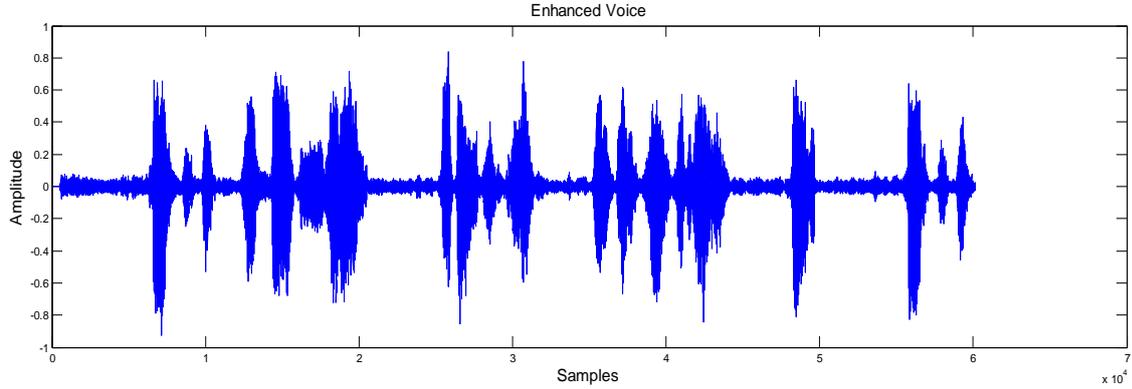
الشكل (9) يبين الضجيج المضاف.



الشكل (10) يبين الإشارة الملوثة بالضجيج.



الشكل (11) يبين الضجيج المتبقي .



الشكل (12) يبين الإشارة المحسنة.

الاستنتاجات و التوصيات :

1 الاستنتاجات:

1. تخفض خوارزمية الطرح الطيفي المعدلة الضجيج الخلفي بشكل جيد ويمكن ملاحظة ذلك من خلال مقارنة الإشارة الملوثة مع الإشارة المحسنة سماعياً وكذلك بمقارنة قيمة SNR .
2. تقدم خوارزمية الطرح الطيفي المعدلة أفضل أداء عندما يكون طول النافذة 256 نقطة $window\ length = 256\ points$ وإن أفضل قيمة لعامل المعايرة هي $\alpha = 5$ ويفضل أن لا تزيد هذه القيمة عن 7 وأن لا تقل عن 5 وأفضل قيمة لمعامل قاعدة الطيف هي $\beta = 0.01$ ويفضل أن لا تزيد قيمة β عن 0.01 وتغطي الخوارزمية أداء متقارب من أجل جميع قيم β الأصغر من 0.01 ولكن لا يجوز أن تكون قيمته مساوية تماماً للصفر حيث يهبط أداء الخوارزمية كلياً.
3. تقوم الخوارزمية المقترحة بتخفيض الضجيج الخلفي بشكل فعال وملحوظ ولكن مع ذلك لوحظ وجود بقايا تشويه للكلام قليلة جداً تكاد لا تذكر.
4. تؤدي الخوارزمية المقترحة عند مستويات SNR_{INPUT} منخفضة لإشارة الدخل الملوثة مثل $0db$ إلى نتائج عالية لتحسين SNR_{OUTPUT} في الخرج يصل إلى $8.42db$.
5. يقل أداء الخوارزمية المقترحة عندما تكون نسبة الإشارة إلى الضجيج في الدخل $SNR_{INPUT} > 20\ db$ وهذا منطقي حيث أنه لا داعي حينها لاستخدام خوارزميات تخفيض ضجيج حيث يقوم الصوت الأصلي بتقنيع الضجيج .

التوصيات:

1. إعادة الدراسة بشكل كامل على نفس الخوارزمية ولكن باستخدام نافذة زمنية أخرى غير نافذة هامينغ وكذلك دراسة إمكانية تطبيق تحويل ويفلت بدلا من تحويل فورييه السريع.
2. دراسة إمكانية تحسين دقة الكاشف الفعال للصوت المستخدم في الخوارزمية ومحاولة استخدام الحساب المرتكز على الحجز (quantile_based estimation) لطيف الضجيج من طيف الإشارة الملوثة والمقارنة بين الطريقتين.

3. إعادة الدراسة بشكل كامل على خوارزمية الطرح الطيفي المعدلة باستخدام نافذة هامينغ ولكن عند أطوال أخرى مثل $window\ length = 64,512,1024\ points$ ودراسة تأثير قيم مختلفة لمعدل التراكب (overlap) مثل النصف أو الربع أو الثلث.
4. دراسة إمكانية تخفيض الضجيج الخلفي بالاستناد إلى ترشيح كالمن أو وينر أو مربع متوسط الخطأ الأصغري (MMSE) ومقارنة النتائج.
5. تضمين معايير أخرى غير SNR لمقارنة الأداء في تخفيض الضجيج الخلفي مثل مقارنة النتائج سماعياً أو رسم المخططات الطيفية أو غير ذلك من المعايير.

المراجع:

- [1]Kaladharan,N.*Speech Enhancement by Spectral Subtraction Method*.International Journal of Computer Applications (0975 – 8887) Volume 96– No.13, June 2014 .
- [2]Verteletskaya,E;Simak,B.*Noise Reduction Based on Modified Spectral Subtraction Method*.IAENG International Journal of Computer Science,38:1, IJCS_38_1_10, February ,2011.
- [3]Tiwari,N;Pandey,P.*Speech Enhancement Using Noise Estimation Based on Dynamic Quantile Tracking for Hearing Impaired Listeners*.Proc. 21th National Conference on Communications 2015 (NCC 2015), Mumbai, Feb. 27 - Mar. 1, 2015.
- [4]Sinclair,I.*Audio And Hi-Fi Handbook*.British Library Cataloguing in Publication Data, 1998.
- [5]Kosek,D.*Audio Spectral Noise Reduction Method and Apparatus*.United States Patent,11/073-820,March,2005.
- [6]Shah,M.S;Geevarghese,G.*Spectral Subtraction for Enhancement of Speech Quality*.Special Issue of International Journal of Electronics, Communication & Soft Computing Science and Engineering, ISSN: 2277-9477,2015.
- [7]ماتلاب , 23 May.2015. < <http://ar.wikipedia.org/wiki/ماتلاب> > .
- [8]Kiusalaas,J.*Numerical Methods in Engineering with MATLAB*.Cambridge University Press,2005 .
- [9]Window_function,20Mar.2015.< http://en.wikipedia.org/wiki/Window_function >.
- [10]MuralidharP.V; Nsastry,L.D; Nayak,S.K.*Interpretation of Dirichlet,Bartlett,Hanning and Hamming windows using Fractional Fourier Transform*.International Journal of Scientific & Engineering Research,Volume 4,Issue 6, June 2013.
- [11]Paliwal,K.K; Lyons,J.G; Wojcicki,K.K.*Preference for 20-40 ms window duration in speech analysis*. Signal Processing Laboratory Griffith University, Nathan, QLD 4111, Australia,2010.
- [12]hammingwindow,2 Dec.2014.< <http://www.mathworks.com/help/signal/ref/hamming.html> >.
- [13]Karam,M;Khazaal,H.F;Aglan,H;Cole,C.*Noise Removal in Speech Processing Using Spectral Subtraction*. Journal of Signal and Information Processing,Vol.5,2014,32-41.
- [14]snr-of-noisy-signal,21May.2015 < <http://dsp.stackexchange.com/questions/4889/how-do-i-calculate-snr-of-noisy-signal> >.