

تصميم ومحاكاة حاسوبية لرمز إشارة كلامية بالإعتماد على تقنية إثارة التنبؤ الخطي من نوع التحليل بالتركيب

مرسالة مقدمة إلى

مجلس كلية العلوم - جامعة بابل

وهي جزء من متطلبات نيل درجة ماجستير علوم

في علوم الحاسبات

من قبل

إيناس حمود محيسن السعدي



رمضان - ١٤٢٦

تشرين الأول - ٢٠٠٥ م

هـ

*Design and Computer Simulation
of Speech Coding Based on
Code Excited Linear Predictive
Technique of Type
Analysis by Synthesis*

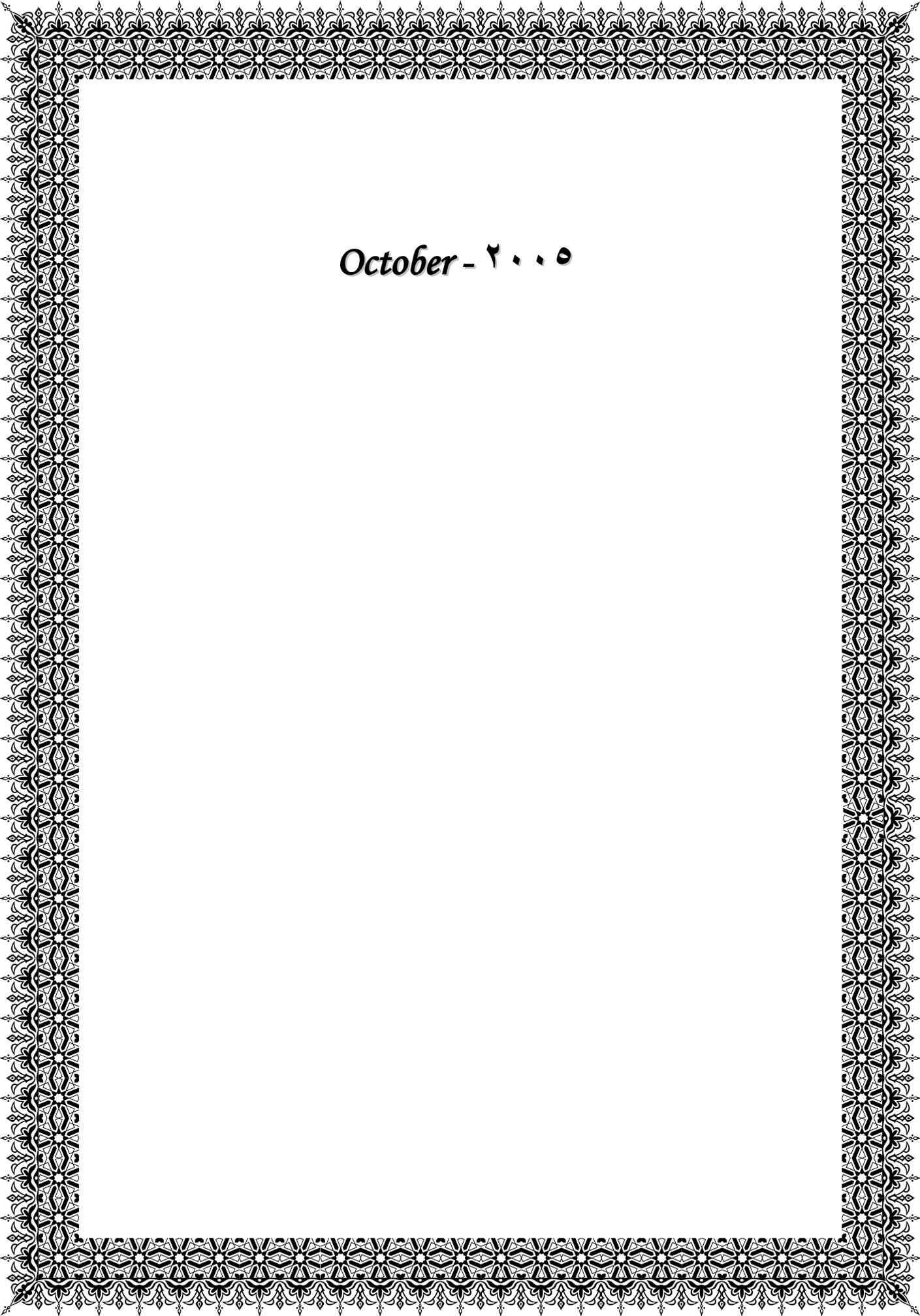
A Thesis

*Submitted to The Council of The Science College of
Babylon University
in Partial Fulfillment of The Requirements for The
Degree of Master of Science in Computer Science*

By

Enas Hamood Al-Saadi



A decorative border with a repeating geometric pattern of stars and floral motifs, framing the central text.

October - ۲۰۱۵

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

الَّذِينَ آمَنُوا وَتَطْمَئِنُّ قُلُوبُهُمْ بِذِكْرِ اللّٰهِ

أَلَّا يَذَّكَّرُ لَهُمُ اللّٰهُ تَطْمَئِنُّ الْقُلُوبُ ﴿٢٨﴾

الَّذِينَ آمَنُوا وَعَمِلُوا الصَّالِحَاتِ طُوبَىٰ لَهُمْ

وَحَسَنُ مَا بِئِهِمُ ﴿٢٩﴾

صَدَقَ اللهُ الْعَلِيُّ الْعَظِيمُ

سورة الرعد

شكر وتقدير

الحمد لله رب العالمين والصلاة والسلام على سيد المرسلين محمد (صلى الله عليه وآله وسلم) .

وبعد فأتقدم بالشكر الجزيل إلى كل من أستاذي المشرفين الأستاذ الدكتور نبيل هاشم كاغد والأستاذ الدكتور ستار بدر سدخان لما بذلاه من جهود في أثناء إشرافهما على الرسالة وتوجيهاتهما القيمة طوال مدة البحث .

ويسرني أن أتقدم بالشكر الجزيل إلى رئاسة ومنتسبي قسم علوم الحاسبات في كلية العلوم جامعة بابل .

كما وأتقدم بالشكر الجزيل إلى الآنسة سعاد عبد الزهرة لما أبدته من مساعدة وتشجيع .

ومن دواعي الوفاء أن أشكر عائلتي التي أعانتني كثيراً على مواصلة مسيرتي العلمية .

والله ولي التوفيق

قائمة المختصرات

المصطلح باللغة الإنكليزية	المختصر باللغة الإنكليزية
Analysis by Synthesis	AbS
Algebraic Code Excited Linear Predictive	ACELP
Absolute Category Rating	ACR
Adaptive Delta Modulation	ADM
Adaptive Differential Pulse Code Modulation	ADPCM
Adaptive Fixed-Code Excited Linear Predictive	AF-CELP
Adaptive Predictive Coding	APC
APC with Adaptive Bit Allocation	APC-AB
Adaptive Pulse Code Modulation	APCM
Adaptive Transform Coding	ATC
bit per second	bps
Code Excited Linear Predictive	CELP
Comparative Mean Opinion Score	CMOS
Coder – Decoder	Codec
Conjugate Structure Algebraic Code Excited Linear Predictive	CS-ACELP
Cutoff Frequency	fc
decibels	dB
Degradation Category Rating	DCR
Discrete Cosine Transform	DCT
Discrete Fourier Transform	DFT
Delta Modulation	DM
Degradation Mean Opinion Score	DMOS
Differential Pulse Code Modulation	DPCM
Diagnostic Rhyme Test	DRT
Enhanced Waveform Interpolative	EWI
Finite Impulse Response	FIR
Hertz	Hz
Integrated Service Original Networks	ISON
Java - Digital Signal Processing	J-DSP
kilo bit per second	kbps
Low Delay Code Excited Linear Predictive	LD-CELP
المصطلح باللغة الإنكليزية	المختصر باللغة الإنكليزية

Linear Predictive Coding	LPC
Multi Band Excitation	MBE
Mixed Excitation LPC	MELP
Minimum Mean Square Error	MMSE
Mean Opinion Score	MOS
Multi Pulse - Excited LPC	MPC
milli second	ms
Pulse Code Modulation	PCM
Power Spectral Density	PSD
Rapidly – Evolving Waveform	REW
Regular Pulse Excitation	RPE
Subband Coding	SBC
second	sec
Slowly – Evolving Waveform	SEW
Signal to Noise Ratio	SNR
Segmental Signal to Noise Ratio	SNR _{seg}
Short-Time Fourier Transform	STFT
Very Large Scale Integration	VLSI
Voiced coder decoder	Vocoder
Vector Sum Excited Linear Prediction	VSELP
Waveform Interpolative	WI
Warped Linear Prediction	WLP

Abstract

Speech coding is considered very important scientific issue, and it takes wide care in different researching sides, specially in current applications that occupy wide space that connect tightly with progressing of the satellite, mobile and cellular communications and other applications that to witnessed development with the information revolution.

The thesis aims to present analytical study for the researcher towards that specific to speech coding issue through the different literatures that deal with this issue. and design codec system of speech coding that depends on the Code Excited Linear Predictive (CELP) principle and to exploit analysis by synthesis technique of closed loop analysis type. The thesis reaches through the designed system and depending on the results of computer simulation in (Visual C++) version (٦), and (MATLAB) version (٦.٥) languages reaches (for the first time according to our thought) to suggestion of some improvements on the designed system to improve the quality of reconstructed speech signal through three approaches :

١. Suggestion of method for improvement of the structure of analysis by synthesis that used in the previous literatures in its known conventional manner .
٢. Suggestion of algorithm to improve loop of the fixed and adaptive codebooks by using suggestive equations to improve the gain values for codebooks that used in the conventional algorithm .

۲. Improved of the reconstructed speech signal in the post processing stage to achieve high quality similar to the quality of the original speech signal .

The evaluation of the algorithm achieved with its suggestions improvement by using objective and subjective evaluations, through the following standards :-

۱. Signal to Noise Ratio (**SNR**) .
۲. Segmental Signal to Noise Ratio (**SNR_{seg}**) .
۳. Mean Square Error (**MSE**) .
۴. In the subjective evaluations it is depended on Mean Opinion Score (**MOS**) .

The results were promising .

In this thesis (and for the first time according to our thought) suggestion changing of the window used in the Levenson Durbin algorithm through number of the suggestions was performed, based on which reaching to the best window. This window differs from the original window that used by Levenson Durbinin in the original algorithm . The compression between all suggestive cases was performed also coding achievement of each suggestive case was calculated.

الخلاصة

تُعَدُّ مسألة ترميز الإشارة الكلامية من المسائل العلمية المهمة ، وقد حظيت باهتمامٍ واسعٍ في جوانب بحثية متعددة ، لاسيما وأن تطبيقاتها الحديثة شغلت حيزاً واسعاً تضمن ارتباطاً وثيقاً مع تنامي اتصالات الأقمار الإصطناعية واتصالات الهاتف المحمول واتصالات الهاتف الخليوي وغيرها من التطبيقات المتعددة التي شهدت نمواً تحت مظلة الثورة المعلوماتية.

تهدف الرسالة إلى تقديم دراسة تحليلية للاتجاهات البحثية الخاصة بمسألة ترميز الإشارة الكلامية من خلال تناول الأدبيات المختلفة التي تعاملت مع هذه المسألة. وتصميم نظام مرمز - فاك ترميز للإشارة الكلامية يعتمد على مبدأ إثارة التنبؤ الخطي (CELP) ويستثمر تقنية التحليل بالتركيب ((Analysis by Synthesis (AbS)) ، من نوع تحليل الحلقة المغلقة. وقد توصلت الرسالة من خلال النظام المصمم وبالاعتماد على نتائج المحاكاة الحاسوبية بلغتي (Visual ++ الإصدار ٦) و (MATLAB) الإصدار (٦.٥) بالتوصل (وللمرة الأولى حسب اعتقادنا) إلى اقتراح بعض التحسينات على النظام المصمم لتحسين جودة إشارة الكلام المعاد بناؤه ، من خلال ثلاثة محاور ، هي :

١. اقتراح طريقة لتحسين هيكل عملية التحليل بالتركيب والمستخدم في الأدبيات السابقة بأسلوبه التقليدي المتعارف عليه.
 ٢. اقتراح خوارزمية لتحسين حلقة بحث كتابي الرموز المتكيف والثابت باستخدام معادلات مقترحة لتحسين قيم الربح لكتابي الرموز المستخدم في الخوارزمية التقليدية.
 ٣. تحسين إشارة الكلام المعاد بناؤه في مرحلة ما بعد المعالجة للحصول على جودة عالية مشابهة لجودة إشارة الكلام الأصلية.
- لقد تم تقويم الخوارزمية بمقترحاتها التحسينية باستخدام التقويم الرياضي والتقويم الحسي، من خلال المعايير الآتية :-

١. نسبة الإشارة إلى الضوضاء (Signal to Noise Ratio SNR) .
٢. نسبة الإشارة إلى الضوضاء المجزأة (Segmental Signal to Noise Ratio (SNR_{seg}) .
٣. معدل مربع الخطأ (Mean Square Error MSE) .

٤ . بالنسبة للمقاييس الحسية فقد تم اعتماد علامة معدل الرأي (Mean Opinion Score

MOS) .

وكانت النتائج واعدة .

لقد تم خلال هذه الرسالة (وللمرة الأولى حسب اعتقادنا) اقتراح تغيير النافذة المستخدمة في خوارزمية ليفنسن داربن من خلال عدد من الاقتراحات ، والتي تم على أساسها التوصل إلى أفضل نافذة . اختلفت عن النافذة الأصلية التي استخدمها ليفنسن داربن في الخوارزمية الأصلية. وقد تمت المقارنة بين جميع الحالات المقترحة ، وحساب إنجازية المرمز تحت كل حالة مقترحة.

قرار لجنة المناقشة

نحن أعضاء لجنة المناقشة نشهد إننا قد اطلعنا على الرسالة الموسومة
(تصميم ومحاكاة حاسوبية لمرمز إشارة كلامية بالإعتماد على تقنية إثارة التنبؤ
الخطي من نوع التحليل بالتركيب) ، وقد ناقشنا الطالبة ((إيناس حمود محسن)) في
محتوياتها وفيما له علاقة بها بتاريخ ٢٠/١٠/٢٠٠٥ ، ووجدنا إنها جديرة بالقبول بتقدير (امتياز) لنيل درجة الماجستير في علوم الحاسبات .

التوقيع :-

رئيس اللجنة :- د. توفيق عبد الخالق عباس
المرتبة العلمية :- أستاذ مساعد
العنوان :- جامعة بابل / كلية العلوم
التاريخ :- / /

التوقيع :-

عضو اللجنة :- د. عبد الحسن خلاوي
المرتبة العلمية :- أستاذ مساعد
العنوان :- الجامعة التكنولوجية / كلية الرشيد
التاريخ :- / /

عضو اللجنة :- د. حسين عطية
المرتبة العلمية :- مدرس
العنوان :- جامعة بابل / كلية العلوم للبنات
التاريخ :- / /

التوقيع :-

عضو اللجنة (مشرفاً) :- د. نبيل هاشم كاغد
المرتبة العلمية :- أستاذ
العنوان :- جامعة بابل / كلية العلوم
التاريخ :- / /

عضو اللجنة (مشرفاً) :- د. ستار بدر سدخان
المرتبة العلمية :- أستاذ
العنوان :- جامعة النهريين / كلية العلوم
التاريخ :- / /

مصادقة عمادة كلية العلوم

أصادق على ما جاء في قرار اللجنة أعلاه

التوقيع :-

الاسم :- د. عودة مزعل ياسر الزامل
المرتبة العلمية :- أستاذ
العنوان :- جامعة بابل / كلية العلوم
التاريخ :- / /

قائمة المحتويات

رقم الصفحة	الموضوع	العنوان
i	شكر وتقدير	
ii	الخلاصة	
iv	قائمة المختصرات	
vi	قائمة الرموز	
viii	المحتويات	
xi	قائمة الأشكال	
xiii	قائمة الجداول	
	المقدمة	الفصل الأول
١	مقدمة عامة	١.١
٣	البحوث السابقة حول رموز الإشارة الكلامية	٢.١
٩	أهداف الرسالة	٣.١
١٠	بنود الرسالة	٤.١
	أنظمة ترميز المصدر	الفصل الثاني
١١	المقدمة	١.٢
١٣	تقنيات ترميز الإشارة الكلامية	٢.٢
١٥	تقنيات ترميز الشكل الموجي	١.٢.٢
١٥	تقنيات الترميز الصوتي	٢.٢.٢
١٦	تقنيات الترميز الهجين	٣.٢.٢
١٧	أهمية ترميز إشارة الكلام	٣.٢
١٨	القابليات الحالية في ترميز إشارة الكلام	٤.٢
٢٠	تحليل - تركيب إشارة الكلام	٥.٢
٢١	تحليل الحلقة المفتوحة	١.٥.٢
٢٢	تحليل الحلقة المغلقة أو التحليل بالتركيب	٢.٥.٢
٢٣	رموزات التنبؤ الخطي للتحليل بالتركيب	١.٢.٥.٢
٢٤	تقويم كفاءة ترميز الإشارة الكلامية	٦.٢

رقم الصفحة	الموضوع	العنوان
٢٤	معدل نقل البيانات	١.٦.٢
٢٥	تأخير الاتصال	٢.٦.٢
٢٥	جودة الإشارة الكلامية	٣.٦.٢
٢٦	الطرائق الرياضية	١.٣.٦.٢
٢٧	الطرائق الحسية	٢.٣.٦.٢
٢٧	تثمين الصنف المطلق	أ.
٢٩	تثمين صنف التنشوة	ب.
٢٩	تثمين المقارنة	ت.

ث. مفهومية إشارة الكلام ٣٠

رقم الصفحة	الموضوع	العنوان
	تقنية ترميز إشارة التنبؤ الخطي	الفصل الثالث
٣١	المقدمة	١.٣
٣٤	التنبؤ الخطي	٢.٣
٣٦	طرائق تقنية ترميز التنبؤ الخطي	١.٢.٣
٣٧	المحاسن الأساسية لطريقة التنبؤ الخطي	٢.٢.٣
٣٨	طريقة الارتباط الذاتي	٣.٢.٣
٣٩	طريقة ليفنسن دارين التكرارية	٤.٢.٣
٤٠	بحث كتاب الرموز	٣.٣
٤١	بحث كتاب الرموز المتكيف	١.٣.٣
٤٢	بحث كتاب الرموز الثابت	٢.٣.٣
٤٣	الترشيح	٤.٣
٤٤	مرحلة احتساب معاملات الإشارة	٥.٣

رقم الصفحة	الموضوع	العنوان
	تصميم منظومة برمجية لمحاكاة المرمز المصدرى الهجين المعتمد على تقنية التحليل بالتركيب	الفصل الرابع
٤٨	مراحل تصميم المنظومة	١.٤
٥٢	الوصف العام للمرمز	٢.٤
٥٢	تسجيل النص المنطوق	١.٢.٤
٥٣	مرحلة المعالجة الابتدائية	٢.٢.٤

رقم الصفحة

الموضوع

العنوان

٥٤	مرحلة تحليل التنبؤ الخطي	٣. ٢. ٤
٦٢	مرحلة بناء كتاب الرموز	٤.٢.٤
٦٢	كتاب الرموز المتكيف	.١
٦٣	كتاب الرموز الثابت	.٢
٦٣	احتساب استجابة النبضة	٥.٢.٤
٦٧	احتساب إشارة الهدف	٦.٢.٤
٦٧	احتساب الريح	٧.٢.٤
٦٩	احتساب إشارة الإشارة	٨.٢.٤
٧٠	فاك الترميز	٣.٤
٧١	احتساب إشارة الكلام المعاد بنائها	١.٣.٤
٧٣	حساب وتوزين واختزال الخطأ	٢.٣.٤
٧٤	مرحلة ما بعد المعالجة	٣.٣.٤
٧٤	التقويم	٤.٤

النتائج ، والمناقشة والاستنتاجات، والأعمال المستقبلية

الفصل الخامس

٧٨	النتائج	١.٥
٧٨	المقاييس الرياضية	١.١.٥
٧٩	المقاييس الحسية	٢.١.٥
٧٩	الحالات الدراسية	٣.١.٥
٧٩	الحالة الدراسية الأولى	١.٣.١.٥
٩٥	الحالة الدراسية الثانية	٢.٣.١.٥
٩٩	الحالة الدراسية الثالثة	٣.٣.١.٥
١٠٢	الحالة الدراسية الرابعة	٤.٣.١.٥

١٠٦	المناقشة والاستنتاجات	٢.٥
١١٢	الأعمال المستقبلية	٣.٥
١١٣	المصادر	
١٢١	ملحق (A)	
١٣١	ملحق (B)	
١٤١	ملحق (C)	

الفصل الأول

المقدمة (Introduction)

١.١ مقدمة عامة (General Introduction)

يُعدّ الكلام طريقة أساسية لإيصال المعلومات من شخص إلى آخر لذا تم التأكيد على تصميم رموزات إشارة كلامية جديدة وكفؤه للاتصالات الصوتية ونقلها [١]. شهد القرن الماضي إزدهاراً بإتجاه تطبيقات رموزات الإشارة الكلامية عند معدلات نقل بيانات واطئة (Low - Bit Rate Speech Coders) للاتصالات العسكرية والمدنية وكذلك تطبيقات الصوت المرتبطة بالحاسوب (Computer - Related Voice Applications) [٢].

يُعدّ موضوع ترميز الإشارات والبيانات موضوعاً خصباً لأنواع مختلفة من البحوث ولعدة سنوات مضت فقد تزايدت الحاجة لترميز الإشارة الكلامية ونقلها بمدى محدود وخبزنها وتنقيتها من الضوضاء [٣].

تعددت في الوقت الحاضر تطبيقات رموزات الإشارة الكلامية . تتضمن هذه التطبيقات ترميز الوقت الحقيقي لإشارات الكلام لاستعمالها في اتصالات الأقمار الإصطناعية للهاتف المحمول والهاتف الخليوي ولأنظمة أجهزة الفيديو الصوتية ، كما تشمل خزن الإشارة الكلامية لتكوين الإشارة الكلامية أو نقل الصوت وأنظمة بريد الصوت [١].

إن أساس تقدم هذه البحوث هو تطوير رموزات إشارة كلامية جديدة قادرة على إنتاج إشارة كلامية ذات جودة عالية عند معدلات نقل بيانات واطئة ، حيث تقوم معظم هذه الرموزات بدمج تقنيات لتمثيل المميزات الطيفية لإشارة الكلام لغرض مطابقة شكل موجة الإشارة الكلامية (Waveform Speech) وتحسين كفاءة المرمز لملائمة أذن الإنسان . وبسبب مشكلة وجود أعداد كبيرة من المستخدمين لهذه الخدمات بنطاق ترددي محدود (Limited Allocated Bandwidth) وظهور الحاجة لتصنيع أجهزة اتصال تستخدم

مرمزات إشارة كلامية ذات معدلات نقل بيانات واطئة مع شبكة اتصالات ذات جودة جيدة ، فقد اعتمدت عدد من هذه المرمزات في تحديد القياسات العالمية للهواتف الخلوية الوطنية والعالمية . [٢] [٤]

تصنف تقنيات ترميز الإشارة الكلامية إلى ثلاثة تقنيات وهي تقنية ترميز الشكل الموجي (Waveform Coding Technique) وتقنية الترميز الصوتي (Voice Coding Technique) وتقنية الترميز الهجين (Hybrid Coding Technique) [١][٣]. تهدف تقنية ترميز الشكل الموجي إلى المحافظة على جودة الصوت المركب وبمعدل نقل بيانات عالي يتراوح بين (٨ kbps) إلى (٦٤ kbps) [٤] ، حيث إن تقنية ترميز الشكل الموجي الشائعة هي التضمين الترميزي النبضي (Pulse Code Modulation (PCM)) والتي تستعمل في أنظمة الاتصالات الهاتفية في الوقت الحاضر [١] ، بينما تعمل تقنية الترميز الصوتي عند معدلات نقل بيانات واطئة جداً تصل إلى (٠.٢ kbps - ٨ kbps) وتركز على إنتاج إشارة كلامية مفهومة من دون ضرورة مطابقة الشكل الموجي . يُعدّ ترميز التنبؤ الخطي (Linear Predictive Coding (LPC)) إحدى تقنيات المرمزات الصوتية شائعة الاستخدام [٥]. تستعمل تقنية الترميز الهجين تقنيات من كلا النوعين وهما تقنية ترميز الشكل الموجي وتقنية الترميز الصوتي وتعطي جودة إشارة كلامية جيدة عند معدلات نقل بيانات متوسطة تتراوح بين (٢ kbps) إلى (١٦ kbps). يُعدّ ترميز إثارة التنبؤ الخطي (Code Excited Linear Predictive (CELP)) أحد المرمزات الهجينة ذات الجودة العالية [١][٢] . هنالك ثلاثة معالم رئيسية لكفاءة تمثيل الصوت وهي معدل نقل البيانات (Bit Rate) وجودة الإشارة الكلامية (Speech Quality) وتأخير الاتصال (Communication Delay) [٦] .

لغرض تطوير ترميز الإشارة الكلامية تم تحليل ذلك الترميز باستعمال التحليل الرياضي والحسي (Subjective and Objective Analysis) حيث يتكون التحليل الحسي من الإصغاء إلى إشارة الكلام المسترجع وتطبيق الحكم على جودة تلك الإشارة الكلامية . وتعتمد جودة الإشارة الكلامية المسترجعة على رأي الشخص السامع للصوت حيث يتراوح رأي هذا الشخص بين صوت غير مفهوم أو صوت مفهوم أو صوت طبيعي . ونظراً لكون هذا التحليل يُعدّ تحليلاً غير عادلاً لجودة الإشارة الكلامية فقد أُستخدم نوعاً آخر من التحليل وهو التحليل الرياضي حيث يعتمد على تقويم جودة الإشارة الكلامية تقنياً وذلك لتقليل تحيز الشخص السامع للصوت حيث يتم تحليل العناصر باحتساب كل من نسبة الإشارة إلى الضوضاء ((SNR) (Signal to Noise Ratio) ونسبة الإشارة إلى الضوضاء المجزأة (Segmental Signal to Noise Ratio (SNR_{seg}) بين الإشارة الكلامية الأصلية والمسترجعة [٧] .

٢.١ البحوث السابقة حول مرمزات الإشارة الكلامية

بدأت بحوث ترميز الإشارة الكلامية منذ أكثر من سبعين عاماً حيث يعد (Homer Dudley) في عام (١٩٣٩) أول من قام بإجراء بحوثه في مختبرات هاتف بل (Bell Telephone Laboratories) . لقد كان الدافع على بحوث ترميز الإشارة الكلامية في ذلك الوقت هو تطوير أنظمة نقل الإشارة الكلامية عبر أسلاك التلغراف ذات عرض نطاق ترددي ضيق . أثبت (Dudley) عملياً وجود الفائضية في إشارة الكلام كما قام بتوفير أول طريقة تحليل - تركيب لترميز الإشارة الكلامية. لقد كانت الفكرة الأساسية التي انبثقت من

المرمز - فاك الترميز الصوتي (Dudley's Vocoder) هي تحليل الإشارة الكلامية بدلالة طبقة الصوت (Pitch) والطيف (Spectrum) وتركيبها من خلال إثارة مجموعة من عشرة مرشحات إمرار النطاق المتناظرة (تمثل قناة الصوت) مع إثارة دورية أو عشوائية (للأصوات المسموعة أو غير المسموعة بالترتيب) [٨] [٩].

لقد حظيت قناة المرمرز - فاك الترميز الصوتي باهتمام كبير خلال الحرب العالمية الثانية بسبب قوتها في النقل الكفوء للإشارة الكلامية المشفرة . خلال الخمسينيات والستينيات تمت دراسة الفورمانات [١٠] ومطابقة النماذج للمرمرز - فاك الترميز الصوتي مع تنفيذ تناظري محسّن لقناة المرمرز - فاك الترميز الصوتي [١١] . إن أفضل مطابقة لنماذج المرمرز - فاك الترميز الصوتي هي المطابقة بين طيف الزمن القصير (Short-Time Spectrum) للإشارة الكلامية ومجموعة محددة من نماذج استجابة التردد المخزون . إذ تنتج الإشارة الكلامية بإثارة مرشح القناة المترافقة مع النموذج المختار . على الرغم من أن تنفيذ المرمرز - فاك الترميز الصوتي يعتمد على تمثيل الإشارة الكلامية المتناظرة فإن التمثيل الرقمي أدى إلى أرباح كبيرة وسريعة بسبب نجاحه في التشفير والنقل عالي الكفاءة والخزن ، وعملياً فقد لاقى التضمين الترميزي النبضي (PCM) إهتماماً كبيراً في الأربعينيات . يُعدّ التضمين الترميزي النبضي أبسط أنواع تقنيات ترميز الشكل الموجي ولا يمتلك أي ميكانيكية لإزالة الفائضية [١٢].

مع المرونة المتوافرة باستخدام الحاسبات الرقمية هنالك ميل إلى التجربة مع تمثيل رقمي معقد للإشارة الكلامية (Sophisticated Digital Representation) [١٣] . ركزت الجهود الأولية على التنفيذ الرقمي للمرمرز - فاك الترميز الصوتي [١٤] . تم التركيز بشكل كبير على نموذج إنتاج نظام مصدر إشارة الكلام الخطي الذي طوره (Fant) في عام (١٩٦٠) . استخدم هذا النموذج لاحقاً في نظام إنتاج الإشارة الكلامية الشائع . يتكون هذا النموذج من نظام خطي ذو تغير زمني بطيء (لقناة الصوت ونموذج المزمار) ويثار بسلسلة نبضات الإثارة الدورية (للكلام المسموع (Voiced Speech)) والإثارة العشوائي (للكلام غير المسموع (Unvoiced Speech)) [١٥] .

أصبح نموذج نظام مصدر إشارة الكلام الخطي مرتبط مع طرائق سلاسل الزمن ذاتية الارتداد حيث يمكن الحصول على معاملاته بالتنبؤ الخطي ، وهي العملية التي يتم بها التنبؤ بنموذج الإشارة الكلامية بالترابط الخطي للعينات السابقة [١٦] . يُعدّ (Itakura) و (Saito) في عام (١٩٦٦) [١٧] و (Atal) و (Schroeder) في عام (١٩٦٧) [١٨] أول من طبق تقنيات التنبؤ الخطي ((Linear Prediction (LP)) للإشارة الكلامية . ثم قام كل من (Atal) و (Hanauer) في عام ١٩٧١ بتسجيل نظام التحليل - التركيب بالاعتماد على التنبؤ الخطي [١٩] . عنونت مشكلة التحليل الطيفي للإشارة الكلامية باستعمال التنبؤ الخطي من قبل (Makhoul) و (Wolf) في عام (١٩٧٢) [٢٠] بينما أختبرت الحقائق العلمية والنظرية لترميز التنبؤ الخطي (LPC) بوساطة (Gray و Markel) (١٩٧٦) [٢١] . لا يُعدّ التنبؤ الخطي الطريقة الوحيدة لتحليل أنظمة المصدر . إذ يمكن إستعمال التحليل المتجانس لفصل الإشارات التي أتحدت بالتلفيف (Convolution) والذي أستعمل أيضاً لتحليل إشارة الكلام ، و يُعدّ (Oppenheim) و (Schafer) أول من اقترحا هذه الطريقة في عام (١٩٦٨) [٢٢] حيث إن أول فوائد تحليل الإشارة الكلامية المتجانس هو القابلية على توفير معلومات طبقة الصوت من السبستروم (Cepstrum) [٢٣] .

يُعدّ نشوء تقنيات التكامل بمقياس كبير جداً (Very Large Scale Integration (VLSI)) مع تقدم نظرية معالجة الإشارة الرقمية خلال الستينيات والسبعينيات دافعاً للحصول على حلول جديدة ومحسنة لمشكلة ترميز الإشارة الكلامية . استعان كل من (Flanagan) و (Golden) في عام (١٩٦٦) بمحول فورير ذي الزمن القصير (Short-Time Fourier Transform (STFT)) في تحليل - تركيب إشارة الكلام [٢٤] ، كما وقاما (Schafer) و (Rabiner) في عام (١٩٧١) بتصميم ومحاكاة نظام التحليل - التركيب بالإعتماد على (STFT) [٢٥] ، وفي عام (١٩٧٦) قدم (Portnoff) قاعدة نظرية لتحليل الإشارة الكلامية في مجالي التردد والوقت باستخدام (STFT) [٢٦] .

شهدت أواسط السبعينيات نشاطاً مستمراً في التنبؤ الخطي [٢٧] والترميز التحويلي (Transform Coding) [٢٨] وترميز النطاق الجزئي (Sub-band Coding) [٢٩] ، إذ قام كل من (Tribolet) و (Crochier) في عام (١٩٧٩) باقتراح محول التحليل الأحادي ومرمzat النطاق الجزئي [٢٨] . كما شهدت السبعينيات جهوداً لتطبيق التنبؤ الخطي في الاتصالات السرية العسكرية [٣٠] . ركزت الجهود في الثمانينات والتسعينيات على تطوير مرمzat الإشارة الكلامية ذات معدل نقل بيانات واطى والقادرة على إنتاج إشارة كلامية ذات جودة عالية لتطبيقات الاتصالات .

أشتقت الكثير من الأعمال نتيجة الحاجة للنطاق الترددي الضيق والإرسال ذو الأمانة العالية في أنظمة الاتصالات الخلوية والاتصالات العسكرية [٢] . فقد اقترح (Atal) و (Schroeder) في عام (١٩٨٤) خوارزمية تنبؤ خطي مع إثارة المتجه العشوائي والذي يدعى ترميز إثارة التنبؤ الخطي (CELP) . إذ تحدد الإثارة العشوائية في ترميز إثارة التنبؤ الخطي باستعمال الحلقة المغلقة (التحليل بالتركيب) . كما تدعى مرمzat إثارة التنبؤ الخطي بالمرمzat الهجينة وذلك لأنها تجمع مميزات المرمzat الصوتية التقليدية ومميزات مطابقة شكل الموجة لمرمzat الشكل الموجي [٣١][٣٢] . برهنت البحوث على أن مرمzat ترميز إثارة التنبؤ الخطي قادرة على إنتاج معدلات نقل بيانات متوسطة وواطئة للإشارة الكلامية كافية لتطبيقات الاتصالات [٣٣] .

بعد ذلك تم اقتراح طريقة التحليل - التركيب الجيبية (Sinusoidal Analysis - Synthesis) لإشارة الكلام من قبل (McAulay) و (Quatieri) في عام (١٩٨٦) [٣٤] و مرمzat الإثارة متعددة النطاق (Multi-Band Excitation) من قبل (Griffin) و (Lim) في عام (١٩٨٨) [٣٥] .

أصبح تنفيذ الوقت الحقيقي للمرمzat الهجينة ملائماً مع تطورات هيكل كتب الرموز . مع استمرار التقدم في مرمzat الإشارة الكلامية (خاصة في نهاية الثمانينات) أصبح من الممكن اختيار خوارزميات المعدل الواطى للهاتف النقال . حيث تم اعتماد خوارزمية إثارة النبضة النظامية (Regular Pulse Excitation Algorithm) ذات معدل نقل بيانات (١٣ kpbs) في أوربا عام (١٩٨٨) [٣٦] .

في عام (١٩٩٠) اختير المرمز الهجين (CELP) بمعدل (٨ kbps) كمقياس لاتصالات الخلوي الرقمي في شمال أمريكا [٣٧] . كما واختير المرمز نفسه بمعدل (٦.٧ kbps) كمقياس لاتصالات الخلوي الرقمي الياباني في العام نفسه [٣٨] .

قدم (Chen) وجماعته في عام (١٩٩٠) نظام ترميز إثارة التنبؤ الخطي واطىء التأخير (LD-CELP) (Low - Delay Code Excited Linear Predictive) لاستخدامه في خدمات الشبكات الرقمية المتكاملة ((Integrated Service Digital Networks (ISDN)) حيث يتميز نظام الترميز هذا بزمن تأخيره القصير والذي يبلغ (٢ ms) إذ تم إنجاز ذلك باستخدام إطارات قصيرة الطول للإشارة الكلامية ، وتقنية التنبؤ الراجع بدلاً من التنبؤ الأمامي المستخدمة في الطرائق التقليدية [٣٩].

ثم قام العالمان (Gerson , Jasiuk) في السنة نفسها باقتراح مرمز (Vector Sum Excited Linear Prediction (VSELP)) وفيه يولد مصدر الإثارة بواسطة الترابط الخطي لعدة متجهات أساس ثابتة وهذا يحسن قوة المرمز ضد أخطاء الانتقال [٤٠].
أقترح (Kataoka) وجماعته في عام (١٩٩٣) مقياساً حديثاً وجديداً لأنظمة الترميز وهو مرمز ((Conjugate Structure Algebraic CELP (CS-ACELP)) والميزة الأساسية في نظام CS-ACELP هو هيكل كتاب رموزه المترابط في مولد إثارة المصدر فضلاً عن كونه أقصر نظام تأخير (وذلك لاعتماده على تقنية التنبؤ الراجع) مقارنة مع الطرائق التقليدية إذ اختزل الهيكل المترابط من متطلبات الذاكرة كما وحسن قوته تجاه أخطاء الانتقال [٤١].

أدى اكتشاف العالمان (Kleijn) و(Haagen) مرمز (Waveform Interpolation (WI)) في عام (١٩٩٤) [٤٢] ومرمز التنبؤ الخطي للإثارة المخلوطة (Mixed-Excitation LPC (MELP)) من قبل (Supplee) وجماعته في عام (١٩٩٧) [٤٣] إلى إنتاج إشارات إثارة وهي عبارة عن اتحاد المكونات الدورية والشبيهة بالضوضاء ، وقد أنتجت هذه المرمزات الحديثة جودة إشارة كلامية ممتازة مقارنة مع نظيراتها السابقة ، على الرغم من أنها بقيت أقل قوة من مرمزات الشكل الموجي ذات المعدل العالي للبيانات المنقولة حيث إنها أكثر تأثراً بالضوضاء ولا تستطيع ترميز الموسيقى بشكل جيد .

بعد ذلك تم اقتراح مرمز (Adaptive Fixed-Code Excited Linear Prediction (AF-CELP)) من العالمين (Kim) و(Lee) في عام (١٩٩٩) بواسطة تضمين مفهوم كتاب الرموز الثابت المتكيف ((Adaptive Fixed Codebook (AFC)) في مرمز CELP حيث يمكن الحصول على معالم إثارة AFC تسلسلياً مع بحث كتاب الرموز المتكيف وهذا يؤدي إلى زيادة قليلة في حجم الذاكرة والتعقيد الحسابي [٤٤].

في عام (٢٠٠٠) قدم (Capin) وجماعته استعراضاً حول (MPEG-٤) النسخة الثانية (٢ version) القياسية. حيث يتضمن مقياس (MPEG-٤) التمثيل والترميز لنماذج الإنسان التقريبي وتداخله مع أهداف (MPEG-٤) أخرى مع متطلبات نقل بيانات أقل من (١ kbps) . تقدم الطريقة ترميزاً كفوءاً وملائماً لتطبيقات اتصالات الوقت الحقيقي مثل تداخلات الفيديو والألعاب واستعمال الجات عبر العالم [٤٥].

شهدت السنوات الأخيرة تطوراً في الخدمات الرقمية كالاتصالات اللاسلكية الرقمية واتصالات الصوت عبر شبكات الإنترنت واتصالات الأقمار الاصطناعية وأنظمة خزن الصوت الرقمية مما أدى إلى خلق ثورة في مجال اكتشاف مرمزات الإشارة الكلامية والتي امتدت إلى الأعوام اللاحقة حيث شهد عام (٢٠٠١) اقتراح وتنفيذ مرمز - فاك ترميز ((Algebraic CELP (ACELP)) من قبل (Siddiqui) وجماعته بمعدلي نقل بيانات ذات

قيمة (٥.٣ kbps) و (٦.٣ kbps) حيث حُسِّنَ هذا المرمز لتمثيل الإشارة الكلامية بجودة عالية وبتعقيد أقل [٤٦].

ثم قام العالمان (Cottesman) و (Gersho) في عام (٢٠٠١) بإقتراح مرمز الإشارة الكلامية (Enhanced Waveform Interpolative (EWI) عالي الجودة عند معدل نقل بيانات واطئ ويتضمن هذا المرمز مميزات جديدة مثل تحسين شكل الموجة البطيئة (Slowly-Evolving Waveform (SEW)) للتوليد والتحليل بالتركيب لطور شكل الموجة البطيئة ، ومقياساً كفوءاً لشكل الموجة السريعة (Rapidly-Evolving Waveform (REW)) وتشير نتائج الاختبار الحسي بان الجودة لمرمز EWI جودة قريبة جداً من الجودة الخاصة بالتطبيق الشبكي على الأقل تحت ظروف الإشارة الكلامية النقية [٤٧].

في عام (٢٠٠٢) قام كل من العالمين (Ko) و (Yu) بإقتراح خوارزمية ترميز صوتي جديدة وهي تقنية التنبؤ الخطي المحرفة (Warped Linear Prediction (WLP)) لتحسين جودة الترميز الصوتي فرعي النطاق (Subband Audio Coding) مما أدى إلى خوارزمية ترميز صوتي فرعي النطاق مهجنة التنبؤ الخطي المحرفة (WLP) [٤٨].

تم تقديم نماذج محاكاة (J-DSP) وتمارين لتعليم تقنيات ترميز الإشارات الصوتية لطلبة الدراسات الأولية والدراسات العليا من قبل (Atti) و (Spanias) في عام (٢٠٠٣). إذ لا يعد برنامج محاكاة (J-DSP) مفيداً لطلبة الدراسات الأولية والعليا فقط وإنما للمهندسين في مجال الصناعات التي تحتاج إلى فهم وتطبيق خوارزميات ترميز الإشارة الكلامية. أستعمل المدربون التمارين المصممة على طريقة محاكاة (J-DSP) في حالة التعامل مع الإشارات الكلامية في الهواتف الخلوية الرقمية والتطبيقات الأخرى [٤٩].

وفي العام نفسه قام (Chu) و (Miki) بتقديم إجراء محسن يطبق على سلسلة النافذة المستخدمة لتحليل التنبؤ الخطي لمرمز (CS-ACELP) بإبدال النافذة الأصلية القياسية بوساطة النسخ المحسنة ، وقد تم الحصول على جودة مشابهة عند تقليل الكلفة المحسوبة أو تقليل تأخير الترميز [٥٠]. وفي عام (٢٠٠٤) قام كل من (Alku) و (Backstrom) بتطبيق طريقة التنبؤ الخطي لتحسين النموذج الطيفي للمكونات الواطنة للإشارة الكلامية مع رتب تنبؤ صغيرة ، حيث يُعدّ النموذج الطيفي الأساس في العديد من تطبيقات الإشارة الكلامية مثل ترميز الإشارة الكلامية والتركيب والتمييز [٥١].

٣.١ أهداف الرسالة

١. دراسة الأنواع الرئيسية الخاصة بمرمّات الإشارة الكلامية .

٢. تصميم وتنفيذ مرمز - فاك ترميز الإشارة الكلامية باستخدام مرمز إثارة التنبؤ الخطي (CELP) وبالاعتماد على تقنية التحليل بالتركيب ((Analysis by Synthesis (AbS) (تحليل الحلقة المغلقة) عند معدل نقل بيانات (٤ kbps) ومن ثم استرجاع إشارة الكلام المركب (المسترجع) مع إنجاز جودة كلام عالية وبدون اختلاف في الإشارة الكلامية الأصلية والمسترجعة أي المعاد بنائها .

٣. إقتراح تحسينات على النظام المصمم لتحسين جودة إشارة الكلام المعاد بناؤه والتي تتمثل بما يأتي :

- أ- تطبيق طريقة لتحسين هيكل التحليل بالتركيب المستخدم في رموزات - فاك ترميز إثارة التنبؤ الخطي .
 - ب- تحسين حلقة بحث كتابي الرموز المتكيف والثابت باستخدام معادلات مقترحة لتحسين قيم الربح لكتابي الرموز والذي يؤدي إلى تحسين إشارة الكلام المعاد بناؤه .
 - ت- تحسين إشارة الكلام المعاد بناؤه في مرحلة ما بعد المعالجة للحصول على جودة عالية مشابهة لجودة إشارة الكلام الأصلية .
٤. تقويم الخوارزمية باستخدام كلاً من التقويمين الرياضي والحسي .

٤.١ بنود الرسالة

يتكون الهيكل العام للرسالة من خمسة فصول يتضمن **الفصل الأول** مقدمة عامة عن موضوع رموزات الإشارة الكلامية ، ودراسة تحليلية للبحوث السابقة حول مسألة ترميز الإشارة الكلامية .

يتضمن **الفصل الثاني** تعريفاً لترميز إشارة الكلام وبيان هدفه وتوضيح تقنياته وأهميته فضلاً عن قابليته الحديثة ، ويتناول الفصل طرائق التحليل - التركيب وأخيراً إستعراض طرائق تقويم كفاءة ترميز الإشارة الكلامية .

يتضمن **الفصل الثالث** مقدمة عن رمز - فاك ترميز إثارة التنبؤ الخطي (CELP Codec) ويلقي نظرة حول التنبؤ الخطي وطرائقه ومحاسنه ، كما ويتناول الفصل بحث كتاب الرموز وكيفية احتساب معاملات الإثارة .

يتضمن **الفصل الرابع** الخطوات الأساسية للتصميم والبناء البرامجي لخوارزمية ترميز - فاك ترميز الإشارة الكلامية .

يتضمن **الفصل الخامس** النتائج التنفيذية ، والمناقشة والاستنتاجات ، والأعمال المستقبلية .

الفصل الثاني

أنظمة ترميز المصدر (Source Coding System)

١.٢ المقدمة (Introduction)

ترميز إشارة الكلام هو الحقل الذي يهتم بالحصول على تمثيل رقمي ملائم لإشارات الكلام لأغراض مختلفة مثل الخزن أو الضغط أو النقل الكفوء [٦] [٥٢].

ويُعدّ ترميز إشارة الكلام المسألة الأساسية في معالجة إشارة الكلام الرقمي إذ يعمل على تحويل إشارة الكلام إلى صيغة أكثر ترابطاً ، والتي يمكن فيما بعد نقلها عبر قناة اتصال بأصغر ذاكره وأصغر عرض نطاق ترددي [٧] ، حيث عملت خوارزميات الترميز على تقليل معدل نقل البيانات ، وبتمثيل رقمي للإشارة بدون فقدان لجودتها عند إعادة استرجاعها .

يمكن الحصول على جودة عالية وبمعدلات نقل بيانات واطئة بإستثمار فائضية الإشارة، وكذلك عن طريق معرفة أنواع معينة من تشوهات الترميز غير المدركه حسيّاً بسبب كونها مقنعة بالإشارة [٣] ، ويبرز الاحتياج إلى ترميز الإشارة الكلامية في منظومات اتصالات المسافات المحددة وخرن إشارة الكلام العالي الجودة وتشفير الرسائل . وعلى سبيل المثال ، في تقنية اتصال الخلوي الرقمي يحتاج العديد من المستخدمين إلى الاشتراك بحزم التردد نفسها ، حيث أتاحت عملية ترميز الإشارة الكلامية للعديد من المستخدمين الاشتراك بالنظام نفسه ، كذلك فإننا نحتاج ترميز الإشارة الكلامية في خزن الصوت الرقمي لكمية ثابتة من الذاكرة المتوافرة حيث جعل ترميز إشارة الكلام من الممكن خزن رسائل ذات حجم كبير [٧] [١].

يقسم معدل ترميز الإشارة الكلامية حسب مدى الترميز وكما مبين في الجدول (١-٢) الآتي [٥٢].

جدول (١-٢) معدلات ترميز الإشارة الكلامية

معدل الترميز	المدى (kbps)
متوسط	٨ إلى ١٦
منخفض	أقل من ٨ وأعلى من ٢.٤
منخفض جداً	أقل من ٢.٤

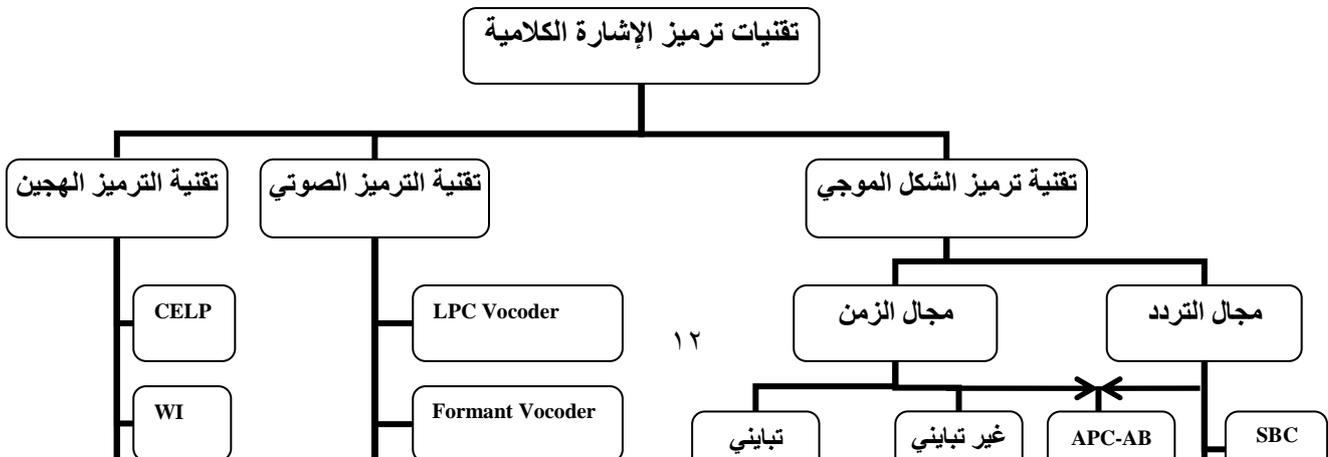
إن الهدف الرئيسي من ترميز الإشارة الكلامية هو تمثيل إشارة الكلام بعدد قليل من البيانات المنقولة مع المحافظة على جودة الإشارة الكلامية فضلاً عن إنجاز تحسيناً موضوعياً في الجودة عن طريق زيادة طول الإشارة الكلامية المسترجعة . وبصورة عامة توجد علاقة ايجابية بين كفاءة المرمز والتعقيد الخوارزمي المطلوب لإنجاز ذلك وهذا يعني إن الخوارزمية كثيرة التعقيد تأخذ معالجة وكلفة تنفيذ أكثر [٥] .

٢.٢ تقنيات ترميز الإشارة الكلامية (Speech Coding Techniques)

تقسم تقنيات ترميز الإشارة الكلامية إلى ثلاثة أصناف رئيسية [٣] وكما مبين في الشكل (١.٢) :-

١. تقنية ترميز الشكل الموجي (Waveform Coding Technique).
٢. تقنية الترميز الصوتي (Voice Coding Technique).
٣. تقنية الترميز الهجين (Hybrid Coding Technique).

ويبين الجدول (٢-٢) مقارنة بسيطة بين تقنيات ترميز الإشارة الكلامية.



شكل (١.٢) تقنيات ترميز الإشارة الكلامية
جدول (٢-٢) مقارنة بين تقنيات ترميز الإشارة الكلامية

الطريقة			
المميزات	تقنية ترميز الشكل الموجي	تقنية الترميز الهجين	تقنية الترميز الصوتي
معدل نقل البيانات	(٨ kbps) - (٦٤ kbps)	(٢ kbps) - (١٦ kbps)	(٠.٢ kbps) - (٨ kbps)
مادة الترميز	أي صوت	صوت متكلم مفرد	صوت متكلم مفرد
تقويم كفاءة الترميز	SNR	MOS و SNR	تشوه طيفي
المشاكل	صعوبة تقليل معدل البيانات المنقولة إلى مدى نطاق ضيق	تعقيد المعالجة	جودة إشارة الكلام المحدودة
أمثلة	في المجال الزمني : PCM APC,ADPCM,DM,ADM, في المجال الترددي : SBC, ATC في إتحاد المجالين: APC-AB	CELP, MPC, RPE	Channel Vocoder, Formant Vocoder, Phase Vocoder, LPC Vocoder

١.٢.٢ تقنيات ترميز الشكل الموجي (Waveform Coding Techniques)

تقنيات الشكل الموجي هي تقنيات غير كلامية قليلة التعقيد تعمل على إعادة بناء (إنتاج) شكل موجي صحيح في المجال الزمني قريب قدر المستطاع لإشارة الإدخال الأصلية وهذا يعني إنه يجب أن تكون الإشارة مستقلة وتعمل بشكل جيد مع الإشارات غير الكلامية (Non Speech Signals) [٥٣] ، وتكون قادرة على إنتاج إشارة كلامية ذات جودة عالية بمعدل نقل بيانات عالٍ بما فيه الكفاية (أعلى من ١٦ kbps) وعندما يكون معدل نقل البيانات أقل من هذا المستوى فإن جودة الإشارة الكلامية سوف تنخفض بشكل سريع [٤] .

تمتلك تقنيات الشكل الموجي عدة فوائد إذ تتحمل خصائص كلامية ذات مدى واسع ومقاومة لبيئة الضوضاء إلا إنها لا تخلو من بعض التعقيدات كما إنها ذات نقل اقتصادي لمعدل نقل البيانات .

أبسط صيغ تقنية ترميز الشكل الموجي هي التضمين الترميزي النبضي (PCM) [٣] فعند ترميز إشارة كلام الهاتف (Telephone Speech) بهذه الصيغة بالاعتماد على مميزات إختلاف السعة (Amplitude Variation Characteristics) فإننا نحصل على جودة عالية عند معدل نقل بيانات (٦٤ kbps) وعندما نستخدم مميزات الترابط والمميزات الطيفية للشكل الموجي فإن معدل نقل البيانات يمكن أن يقل إلى (٣٢ kbps) أو إلى (٢٤ kbps) [٥٣] .

تنفذ مرمزات الشكل الموجي أما في مجال الزمن (Time Domain) أو في مجال التردد (Frequency Domain) [١] .

٢.٢.٢ تقنيات الترميز الصوتي (Voice Coding Techniques)

تعتمد التقنيات الصوتية على نماذج إشارة الكلام وتحافظ على الصفات الطيفية لإشارة الكلام في الإشارة المرزمة وتركز على إنتاج إشارة كلام واضح ومفهوم عند معدلات نقل بيانات واطئة ، وواطئة جداً من دون ضرورة مطابقة الشكل الموجي ولكن مستوى جودة الإشارة الكلامية وبشكل طبيعي وموحد لمختلف المتكلمين يكون قليلاً ، حيث تكون تطبيقات تقنيات الترميز الصوتي محدودة لقنوات الإتصال الرقمية ذات معدل نقل البيانات الواطئ [٣] .

يتم في تقنيات الترميز الصوتي إرسال معاملات مرشح مصدر نموذج إشارة الكلام وليس عينات إشارة الكلام ويتم تمثيل تركيب مرشح المصدر (Source Filter) حسب نموذج إنتاج إشارة الكلام [٥٣] .

يطابق المرشح التركيبي ذو التغير الزمني (Time Varying Synthesis Filter) القناة الصوتية (Vocal Tract) أو الأنبوب السمعي (Acoustic Tube) أو حركة الشفاه (Lip Radiation) حيث يتم استعمال هذا المرشح طباقاً لمبادئ ترميز التنبؤ الخطي [٥٤] . تحتوي إشارة إثارة المزمار المقيسة بالربح على قطع دورية حيث تأتي هذه القطع الدورية من مولد النبضة (Impulse Generator) والذي يمثل إشارات الصوت المسموع أو الضوضاء مثل القطع التي تأتي من الصوت غير المسموع وبدلاً من التحيز إلى استخدام الإثارة الصوتية أو

اللاصوتية فقد تم اكتشاف نموذج موزون طيفياً (Spectrally Weighted) وهو عبارة عن مزيج من كلا النوعين [٤]. تستخدم تقنيات الترميز الصوتي النقية بشكل خاص للتطبيقات ذات معدل نقل البيانات الواطئ مثل الإثارة متعددة النطاق (MBE) [١].

٣.٢.٢ تقنيات الترميز الهجين (Hybrid Coding Techniques)

تُعَدُّ التقنيات الهجينة نوعاً وسطياً بين تقنيات الشكل الموجي و التقنيات الصوتية إذ إنها تجمع بين كفاءة الترميز لتقنيات الترميز الصوتي مع قوة الجودة العالية لتقنيات ترميز الشكل الموجي وتستخدم هذه التقنيات تحليل الحلقة المغلقة [٥٢] ، اشتقت هذه المرمزات لكي تستخدم كطريقه وسطية بين أنظمة تقنيات الترميز الصوتي باستخدام النبضات والضوضاء كمصادر صوتية وأنظمة تقنيات ترميز الشكل الموجي [٥٥].

تنتج تقنيات الترميز الهجينة جودة إشارة كلامية عالية عند معدلات نقل بيانات متوسطة وتطبق بشكل خاص في الهواتف الرقمية والخطوط السلكية والهاتف النقال وأنظمة الاتصال الأخرى وكما في تقنيات الترميز الصوتي فإنه يتم إرسال معاملات مرشح ترميز التنبؤ الخطي (LPC Filter) المركب كمعاملات للقناة [٥٤].

تعتمد أغلبية تقنيات ترميز إشارة الكلام الهجينة الحديثة على مبدأ (ترميز التنبؤ الخطي للتحليل بالتركيب) مثل مرمرز- فاك ترميز إشارة التنبؤ الخطي (CELP) وهي أحد الخوارزميات المفضلة لمعدل نقل بيانات بين (٤ kbps) و (١٦ kbps) [٥٥] وعلى الرغم من إن هذه الخوارزمية تضمن جودة إشارة كلامية جيدة عند معدل نقل بيانات واطئ باستخدام التحليل بالتركيب إلا إن تنفيذها ليس سهلاً بسبب التعقيد العالي لبحث الإثارة المثالي في عملية تركيب الإشارة الكلامية حيث إن متجه إشارة الإثارة المثالي يحدد بواسطة محددات الدورة المغلقة [٥٦].

٣.٢ أهمية ترميز إشارة الكلام (The Importance of Speech Coding)

يُعَدُّ ترميز إشارة الكلام ضرورة أساسية بالنسبة لتقنيات الاتصالات عبر المسافات الطويلة وعند خزن إشارات الكلام ذات الجودة العالية وتشفير الرسائل . وأصبح ترميز إشارة الكلام المفتاح الرئيسي للاتصالات على الرغم من وجود وسائل الإعلام المرئية غير المحدودة [٣].

إن الاستعمال المتزايد لوسائل الإعلام محدودة النطاق مثل الراديو والأقمار الاصطناعية وكذلك وسائل الإعلام ذات الخزن المحدود لمعدل نقل البيانات مثل الذاكرة المكتنزة (CD_Rom) وذاكرة السيلكون كما إن خزن كميات كبيرة من أصوات المتكلمين في الأرشيف أدى إلى تزايد الحاجة إلى ترميز إشارة الكلام حتى بوجود ذاكرة الحالة الصلبة والذاكرة الضوئية [٧].

تُعَدُّ تقنيّة ترميز الإشارة الكلامية ذات معدل نقل بيانات واطئ الأساس في الاتصالات الرقمية اللاسلكية وتركز البحوث الحديثة على تطوير تقنيات ترميز الإشارة الكلامية ذات الجودة العالية عند معدلات نقل بيانات واطئة وبكلفة منخفضة [٥٥].

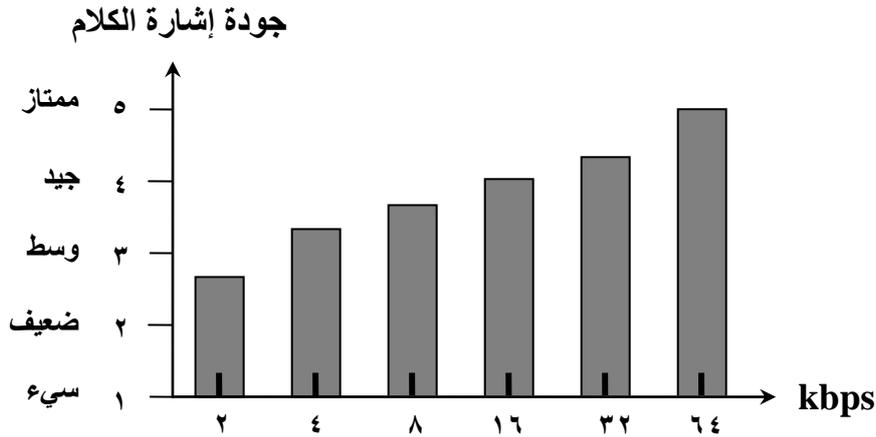
خلال السنوات العشر الماضية لم يكن بالإمكان إنتاج إشارة كلامية بجودة عالية وبمعدل نقل بيانات أقل من (٢٤ kbps) بينما تم في الوقت الحاضر إنتاج إشارة كلامية بجودة عالية وبمعدل نقل بيانات (٨ kbps). وباستعمال تقنيات جديدة لترميز القناة يمكننا أن ننقل إشارة كلامية ذات معدل نقل بيانات (٨ kbps) عبر قنوات الراديو المتنقلة (Mobile Radio Channel) على الرغم من ضوضاء القناة وخفوت الصوت لاسلكياً والتداخل بين الرموز [٥].

تتركز بحوث الإشارة الكلامية على إنتاج إشارة كلام رقمي يمكن أن ينتقل وبجودة عالية عبر شبكات الاتصال بوجود أكثر من (١٠ %) أخطاء القناة وبمعدل انتقال نهائي في المدى (٨-٤ kbps).

٤.٢ القابليات الحالية في ترميز إشارة الكلام

(Current Capabilities in Speech Coding)

يبين الشكل (٢-٢) جودة إشارة الكلام المعاد بناؤه والتي تنجز عند معدلات نقل بيانات متنوعة من (٢ kbps) - (٦٤ kbps) لإشارة كلام نطاق التليفون الضيق (Narrowband Telephone) وهو (٣٤٠٠ - ٣٠٠ Hz) حيث إن مفهومية إشارة الكلام المسترجع تكون متباينة عند هذه المعدلات لنقل البيانات ويعبر عن جودة إشارة الكلام بخمس نقاط لعلامة معدل الرأي (Mean Opinion Score (MOS)) على طول الإحداثي العمودي في الشكل (٢.٢) [٣].



شكل (٢.٢) علامة معن رأي جودة إشارة الكلام المعاد بناؤه لمعدلات نقل بيانات مختلفة

أبسط أنظمة الترميز هي التضمين الترميزي النبضي (PCM) وهي مكمئ بلا ذاكرة (memoryless quantizer) ويجهز ترميزاً واضحاً لإشارة كلام التليفون عند ٦٤ kbps [٤] ومع تطویر التضمين الترميزي النبضي تم إنتاج مُتَّبئئ مكيف وهو التضمين الترميزي النبضي التبايني المتكيف (Adaptive Differential Pulse Code Modulation (ADPCM)) والذي يجهز إشارة كلامية ذات جودة عالية عند معدل نقل

بيانات (32 kbps) حيث تكون جودة إشارة الكلام المنتج من (ADPCM) أقل من جودة إشارة الكلام للـ (PCM) عند معدل نقل بيانات (64 kbps) [1][8].

تتجزء المعدلات 16 kbps وأقل من ذلك جودة إشارة كلامية عالية بإستعمال التنبؤ ذي التكيف المعقد مثل ترميز التنبؤ الخطي (LPC) وتنبؤ طبقة الصوت (Pitch Prediction). أمثلة مهمة لمثل هذه المرمزات هي مرمزات الإثارة متعددة النبضة ومرمزات إثارة النبضة النظامية ومرمزات ترميز إثارة التنبؤ الخطي [3][5], وتجمع خوارزمية ترميز إثارة التنبؤ الخطي (CELP) قوة الجودة العالية لتقنية ترميز الشكل الموجي مع كفاءة الترميز لتقنية الترميز الصوتي، وترمز إشارة الكلام عند معدل نقل بيانات 16 kbps وأقل [5]. تجهز خوارزمية ترميز إثارة التنبؤ الخطي واطئ التأخير (LD-CELP) عند معدل نقل بيانات (16 kbps) جودة عالية قريبة من الجودة لإشارة الكلام التي يجهزها (PCM) فضلاً عن كونها قليلة التأخير ومقبولة كمقياس دولي لنقل إشارة الكلام عبر شبكات التلفون [3]. اختير معدل نقل البيانات (8 kbps) كمقياس للخلوي الرقمي في شمال أمريكا والذي يعطي جودة إشارة كلامية جيدة على الرغم من أنها أقل جودة من تلك الناتجة عند معدل نقل البيانات 64 kbps لرمز التضمين الترميزي النبضي [4]. إذ يركز كل من الجيل الأول الياباني والأمريكي الشمالي على تقنية (CELP) ويعتمد المقياس الخلوي الرقمي الأوربي الأول على خوارزمية إثارة النبضة النظامية (RPE) عند معدل نقل بيانات (13 kbps) [5].

يُعدّ المعدل (4.8 kbps) مهماً جداً حيث إنه يُمكننا من نقل الإشارة الكلامية عبر خطوط التلفون المحلية في الولايات المتحدة. أختيرت نسخة (CELP) التي تعمل عند معدل نقل بيانات (4.8 kbps) كمقياس في الولايات المتحدة للإتصالات الصوتية السرية. المعايير الأخرى تستخدم المرمز - فاك الترميز الصوتي لترميز التنبؤ الخطي عند معدل نقل بيانات (2.4 kbps) حيث ينتج المرمز - فاك الترميز الصوتي لترميز التنبؤ الخطي إشارة كلامية مفهومة ولكن جودة الإشارة الكلامية غير طبيعية [3].

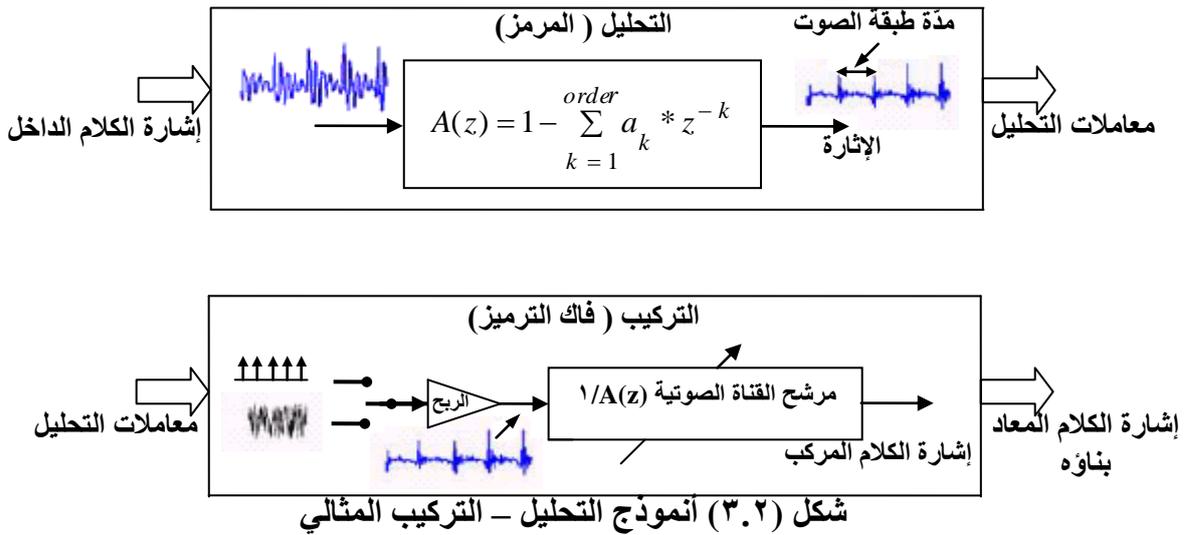
٥.٢ تحليل / تركيب إشارة الكلام (Speech Analysis / Synthesis)

تمثل إشارة الكلام مجموعة معاملات متكاملة تستعمل لتركيب إشارة الكلام. ويعرف التحليل - التركيب على أنه العملية التي يتم فيها إعادة إنتاج أو تركيب موجة إشارة الكلام (Speech Wave) بإستعمال مصدر صوت ومعاملات النطق [57]. تستخلص المعاملات بالاعتماد على دائرة مكافئة مفصولة خطياً لميكانيكية إنتاج إشارة الكلام وتصمم هذه المعاملات لأربعة أنواع من المعلومات:

- ١ . التميز بين الصوت المسموع (مصدر النبضات) والصوت غير المسموع (مصدر الضوضاء).
- ٢ . التردد للأصوات المسموعة .
- ٣ . سعة المصدر (Source Amplitude) .
- ٤ . خصائص المرشح الخطي (الرنين (Resonance)) .

تمثل أول ثلاث معلومات المصدر بينما يعطي المعامل الأخير معلومات الغلاف الطيفي (النطق) [٥].

يُعدّ التحري بعناية للإجراءات الأساسية الثلاثة لأنظمة تحليل وتركيب إشارة الكلام ضرورياً للتأكد من تحسن جودة إشارة الكلام المركب . حيث أن **الأجراء الأول** هو استخلاص هذه المعاملات والتي تنقل بدقة المعلومات السمعية المهمة فقط بإهمال المعلومات الفائضة المتضمنة في موجات إشارة الكلام . **الأجراء الثاني** هو حساب الإثارة للإشارة الكلامية بكفاءة. **الأجراء الثالث** هو إعادة إنتاج إشارة الكلام الأصلي بدقه وبوضوح وبشكل طبيعي قدر الإمكان باستخدام المعامل المميزة (Feature Parameters) (المعاملات والإثارة) [٥٨] . ويوضح الشكل (٣.٢) نموذج التحليل – التركيب المثالي [٥٠].



يوجد نوعان من تحليل - تركيب إشارة الكلام وهما :-

١.٥.٢ تحليل الحلقة المفتوحة (Open - Loop Analysis)

تستخلص المعاملات في تحليل الحلقة المفتوحة بدون الأخذ بنظر الاعتبار الاختلاف بين إشارة الكلام الأصلي وإشارة الكلام المعاد بناؤه لذلك لا يتضمن تحليل الحلقة المفتوحة مرحلة سيطرة التغذية المرتدة التكرارية (Iterative Feedback Control) لذا تدعى بعملية تحليل الحلقة المفتوحة [٥٧].

ينتج تحليل الحلقة المفتوحة إشارة كلامية وبجودة جيدة عند معدلات نقل بيانات أعلى من (١٦ kbps) وتكون قليلة التعقيد عادةً (أي تتطلب حسابات قليلة) إلا أنها أقل جودة من تحليل الحلقة المغلقة (التحليل بالتركيب) لأن معاملاتها لا تعدل كما في تحليل الحلقة المغلقة في كل مرة بمرحلة سيطرة التغذية المرتدة التكرارية . ومن أمثلة المرمزات التي تعتمد على مفهوم الحلقة

المفتوحة هي رمز التضمين الترميزي النبضي (PCM)، ورمز التضمين الترميزي النبضي التبايني المنكف (ADPCM) فضلاً عن العديد من الرموز الأخرى [٥٩].

٢.٥.٢ تحليل الحلقة المغلقة أو التحليل بالتركيب

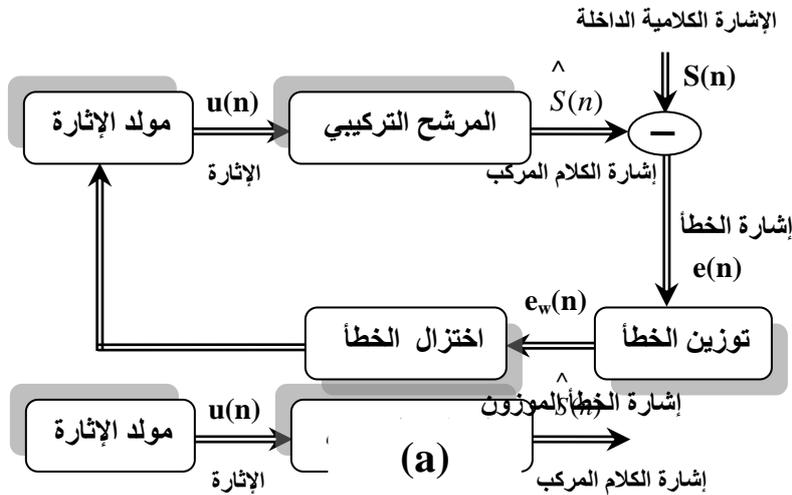
(Closed – Loop Analysis or Analysis by Synthesis)

هو عملية تحديد المعالم التي تميز النظام بالاعتماد على نموذج إنتاج الإشارة المفروضة [٦٠] وتستخلص وتعديل المعاملات في تحليل الحلقة المغلقة بمرحلة سيطرة التغذية المرتدة التكرارية بتقليل الاختلاف بين إشارة الكلام الأصلي والمعاد بناؤه. لذلك يتضمن تحليل الحلقة المغلقة تركيب، لذلك تدعى هذه العملية التحليل بالتركيب [٥٧].

ومن المهم في منظومة التحليل بالتركيب (Abs) اختيار نموذج الإنتاج المفترض وقيم المعاملات الابتدائية (Initial Parameters) وقياس تقويم الخطأ (Error Evaluation Measure) وخوارزمية الاختزال (Minimization Algorithm) ولقد تم استخدام معدل مربع الخطأ (Mean Square Error) بين إشارة الكلام الأصلي وإشارة الكلام المعاد بناؤه لنموذج إشارة الكلام كقياس لتقويم الخطأ [٥].

تظهر فائدة التحليل بالتركيب في استخلاص معامل الإشارة الكلامية كما تظهر في تطبيقات عديدة والتي يمكن أن يستخدم بها نموذج الإنتاج.

على الرغم من أن طريقة التحليل بالتركيب في مبادئها أفضل من أية طريقة أخرى إلا أنه هنالك مشكلة في الحسابات التي تتطلبها، خاصة وإنها تحتاج إلى عدد كبير من التكرارات لسيطرة التغذية المرتدة خلال تحليل الإشارة الكلامية الحقيقية بسبب التداخلات المتبادلة بين معاملات مختلفة تؤثر في إنتاج الغلاف الطيفي [٦٠]. تعود خوارزميات النوع CELP إلى هذا الصنف ويكون تحليل الحلقة المغلقة عادة عالي التعقيد، وتعد دمرزات التنبؤ الخطي للتحليل بالتركيب أحد التطبيقات المعتمدة على مفهوم الحلقة المغلقة. يوضح الشكل (٤.٢) النموذج العام لرمز - فاك ترميز التحليل بالتركيب (Abs) [١] [٥٧].



شكل (٤.٢) هيكل مرمز (b) فاك ترميز التحليل بالتركيب:

(a) المرز- فاك الترميز. (b) فاك الترميز.

١.٢.٥.٢ مرمزات التنبؤ الخطي للتحليل بالتركيب

(Analysis by Synthesis Linear Predictive Coders)

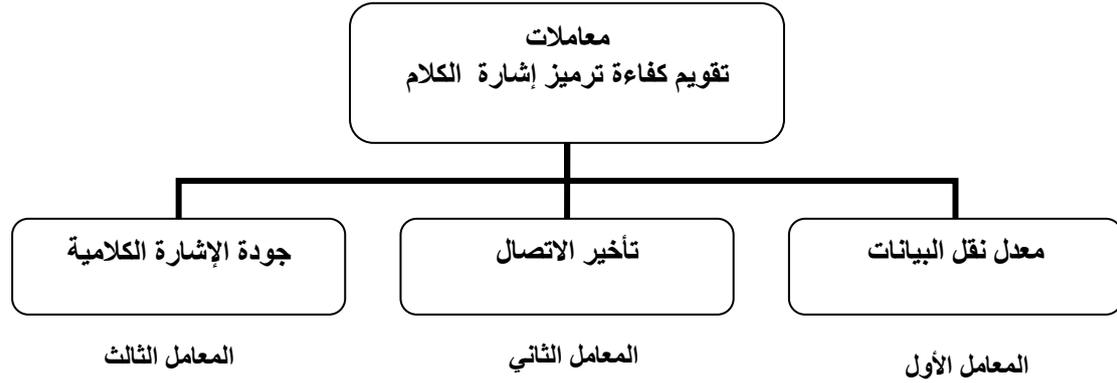
تتضمن مرزمزات الإشارة الكلامية للتحليل بالتركيب أصنافاً مهمة من المرزمزات مثل مرزمزات إثارة التنبؤ الخطي (CELP) والتي تقع بين مرزمزات الإشارة الكلامية الجديدة والأكثر فاعلية [٢]. تُعدّ مرزمزات التنبؤ الخطي للتحليل بالتركيب من المرزمزات الهجينة طالما تجمع بين صفات النماذج التي تعتمد على المرزمزات الصوتية بتمثيل الفورمانات (Formants) وهيكل طبقة الصوت للإشارة الكلامية وخصائص مرزمزات الشكل الموجي بتجهيز المطابقة للشكل الموجي لإشارة الكلام الداخل [٣٢].

وقد اعتمدت خوارزميات التحليل بالتركيب لترميز الإشارات الكلامية عند معدلات نقل بيانات متوسطة وأقل من ذلك. وتعتمد أنظمة الترميز المعيارية الحديثة على التنوع بهذا المبدأ (Abs). يعزى الإستقرار في مرشح التركيب $1/A(z)$ للمرزمز بسب امتلاكه لحلقة التغذية المرتدة (Feedback Loop) ويعني الإستقرار إن الخارج من النظام للإدخال المحدود هو نفسه ذو قيمة محدودة، فإذا حصلنا على معاملات التنبؤ الخطي من خلال طريقة الارتباط الذاتي لتحليل التنبؤ الخطي فنظرياً نؤكد الإستقرار لمرشح التركيب [١].

٦.٢ تقويم كفاءة ترميز الإشارة الكلامية

(Evaluation of The Speech Coding Performance)

تحاول مرزمزات إشارة الكلام تقليل معدل نقل البيانات للاستفادة من ذلك في نقل وخرن الإشارة بينما نحافظ على المستوى المطلوب لجودة إشارة الكلام، وتأخير الإتصال [٣] ويوضح الشكل (٥.٢) معاملات تقويم كفاءة ترميز الإشارة الكلامية.



شكل (٥.٢) معاملات تقويم كفاءة ترميز الإشارة الكلامية

١.٦.٢ معدل نقل البيانات (*Bit Rate*)

نعبر عن معدل كفاءة الترميز بالبيانات لكل ثانية (bps) [٣] وتتراوح معدلات نقل البيانات من (٦٤ kbps) لرمز التضمين ترميزي النبضي (PCM) إلى أقل من (١ kbps) ويمكن ترميز إشارة كلام الهاتف مع جودة جيدة عند معدل نقل بيانات (٤.٨ kbps) [٥٣].

٢.٦.٢ تأخير الاتصال (*Communication Delay*)

تقوم مرمزات الإشارة الكلامية بمعالجة إشارة الكلام بهيئة مجاميع (Blocks) وهذه الطريقة تؤدي إلى تأخير الإتصال [٥٣] وبالاعتماد على التطبيق يمكن أن يكون مجموع التأخير أقل من (١ sec) كما في شبكة التلفون أو أعلى من (٥ sec) كما في التلفون الفديوي [٣].

٣.٦.٢ جودة الإشارة الكلامية (*Speech Quality*)

تصنف جودة الإشارة الكلامية في الاتصالات الرقمية إلى أربعة أصناف وهي : واسع الانتشار (Broadcast)، الشبكي (Network)، الاتصالي (Communication) والتركيبي (Synthetic) [٥٧].

تشير إشارة الكلام واسعة النطاق واسعة الانتشار إلى جودة إشارة كلامية عالية والتي تنجز عند معدلات نقل بيانات أعلى من (٦٤ kbps). تشير الجودة الخاصة بالتطبيق الشبكي إلى جودة مشابهة لإشارة الكلام التناظري وتنجز عند معدلات نقل بيانات أعلى من (١٦ kbps). تتضمن الجودة الخاصة بالتطبيق الاتصالي جودة إشارة كلامية منخفضة نوعاً ما على الرغم من كونه طبيعي ومفهوم بشكل عالي ومكافئاً للاتصالات الهاتفية، وتنجز عند

معدلات نقل بيانات أعلى من (4.8 kbps). تكون إشارة الكلام التركيبي عادة مفهومة ولكنها غير طبيعية ومترافقة مع عدم القدرة على تميز المتكلم [5].

تعمل مرمزات إشارة الكلام حالياً بشكل جيد عند معدلات أقل من (4 kbps) لإنتاج إشارة كلامية ذات جودة خاصة بالتطبيق التركيبي حيث عندما تقل معدلات نقل البيانات فإن مفهومية إشارة الكلام تتأثر أيضاً [57]. يُعدّ تقويم جودة إشارة الكلام مهماً ولكنه بنفس الوقت عملية صعبة جداً ويقوم بالطرائق الرياضية والطرائق الحسية وتحاول المقاييس الرياضية أن تعطي رقماً إلى جودة إشارة الكلام، ولكي تكون هذه المقاييس مفيدة يجب أن تمتلك الخاصيتين التاليتين: الخاصية الأولى تتضمن وجوب تطابق الجودة الرياضية العالية والواطنة مع الجودة الحسية العالية والواطنة. الخاصية الثانية تتضمن وجوب إمكانية التحليل رياضياً وتنفيذها في بعض الخوارزميات. تعمل طرائق التقويم الحسية مع اختبارات الإصغاء وتدرج على مقاييس مختلفة [61].

١.٣.٦.٢ الطرائق الرياضية (Objective Methods)

طرائق المجال الزمني :- هي نسبة الإشارة إلى الضوضاء (SNR) وتُعدّ أحد المقاييس الرياضية الشائعة لتقويم كفاءة خوارزمية الترميز، حيث إنها مقياس لدقة إشارة الكلام المعاد بناؤه وتعرف بالمعادلة الآتية :-

$$SNR = 10 * \log_{10} \left\{ \frac{\sum_{n=0}^{L_0-1} S^2(n)}{\sum_{n=0}^{L_0-1} \left[S(n) - \hat{S}(n) \right]^2} \right\} \quad (\text{in dB}) \quad (1-2)$$

حيث إن $S(n)$ هي إشارة الكلام الأصلي (Original Speech Signal) و $\hat{S}(n)$ هي إشارة الكلام المعاد بناؤه (Reconstruction Speech Signal) و L_0 هي عدد العينات في الإشارة الكلامية [5]. إلا إن الـ (SNR) لا تتطابق بشكل جيد مع التقويم الحسي. وكمثال على ذلك في حالة التوقف وبوجود كميات صغيرة من الضوضاء فإن ذلك يؤدي إلى إنتاج قيم SNR سالبة كبيرة [1]. يمكن حل ذلك بقياس الـ (SNR) لإطارات قصيرة ثم يأخذ معدلها للحصول على مقياس جودة محسن ويدعى بنسبة الإشارة إلى الضوضاء المجزأة (SNR_{seg}) والمعرف بالمعادلة الآتية :-

$$SNR_{seg} = \frac{1}{M} \sum_{j=0}^{M-1} 10 * \log_{10} \left[\frac{\sum_{n=0}^{L-1} S^2(n)}{\left(S(n) - \hat{S}(n) \right)^2} \right] \quad (\text{in dB}) \quad (2-2)$$

حيث إن M هو عدد الإطارات ذات الحجم (L) لحساب نسبة الإشارة إلى الضوضاء المجزأة (SNR_{seg}) على $S(n)$ وهي إشارة الكلام الأصلي و $\hat{S}(n)$ هي إشارة الكلام المعاد بناؤه.

بالرغم من التحسينات لمقاييس التقويم فإنه نادراً ما تعتمد المقاييس على تشوهات الشكل الموجي إذ تكون هذه التحسينات غير كافية خاصة لتقنيات الترميز الهجينة ذات معدل نقل بيانات واطئ ، ويعزى ذلك لكون القيمة الواطنة للـ (SNR) غير ضرورية للدلالة على الجودة الواطنة حيث يمكن إستعمال مقاييس الـ (SNR) في بعض الحالات لأغراض التصميم كرقم تقريبي بينما تعد غير مناسبة للمقارنات بين مخططات الترميز المختلفة وكمقياس مطلق للجودة [٥٣][٥٨].

تعدّ الـ (SNR) مقياس جودة غير ملائم في تقويم مرمزات التحليل بالتركيب ، وذلك لأن شكل موجة الإشارة الكلامية يعاد إنتاجها تركيبياً باستعمال معالم تمثل معلومات مهمة مع عدم الأخذ بنظر الاعتبار شكل الموجة الحقيقي ولهذا فمن الأفضل تقويم جودة الإشارة الكلامية باستخدام الاختبارات الحسية [٥٨].

٢.٣.٦.٢ الطرائق الحسية (Subjective Methods)

هنالك طرائق مختلفة للاختبارات الحسية وهي :-

أ. تمييز الصنف المطلق (Absolute Category Rating (ACR))

وهو أحد المقاييس الحسية الشائعة و الذي يتم فيه تقويم جودة إشارة الكلام باستعمال تدرج خماسي النقاط (Five_Point Scale) [٣] ، عن طريق السماح لمجموعة من الأشخاص (١٢-٢٤ شخص) بالاستماع إلى إشارة الكلام لمدة (٦) إلى (١٠) ثواني ومن ثم تقويم الجملة بأخذ معدل تقاويم هؤلاء المستمعين والتي على أساسها يتم الحكم على المرمز بالرجوع إلى المستويات الخمسة المتفق عليها ، وهي : (ممتاز (Excellent) ، جيد (Good) ، وسط (Fair) ، ضعيف (Poor) ، سيء (Bad)) حيث تعطى هذه المستويات قيم رقمية من ٥ إلى ١ [٦٠] ، وكما مبين في الجدول (٢-٣) [٦] ، فإذا كانت نتيجة الإختبار تصل إلى (٣.٥) أو أعلى فتعدّ درجة مقبولة لتمييز الصوت وفهمه [٦٢].

يقوم الشخص الذي يجري الاختبار بتقويم مجموعة من الجمل متضمنة عدد من الجمل والتي تُعدّ كمصدر (جملة أصلية) . يجب القيام بتوزيع الجمل بشكل عشوائي لمجموعة من المصغين . تمثل النتيجة النهائية علامة معدل الرأي (MOS) والتي تحسب من نتائج المصغين . يمكن أن تؤثر مواد الاختبار (الجمل) في مقياس علامة معدل الرأي بمد أو ضغط الجمل ، لذا يجب إضافة جمل كمصدر (جملة أصلية) [٥].

جدول (٢-٣) مقياس MOS

مقياس MOS	جودة إشارة الكلام
-----------	-------------------

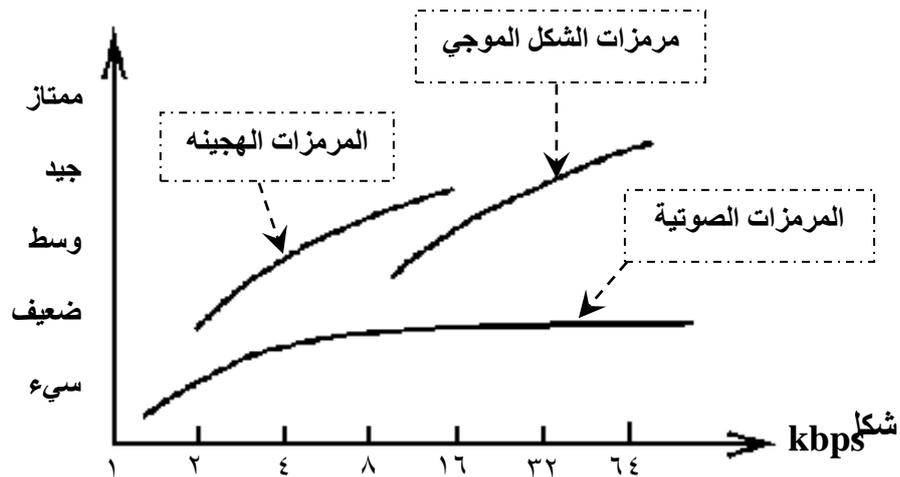
١	سيء
٢	ضعيف
٣	وسط
٤	جيد
٥	ممتاز

ويتعلق مدى علامة معدل الرأي (MOS) بجودة الإشارة الكلامية وكما يلي :-

- MOS ٢.٥ - ٣.٥ : جودة خاصة بالتطبيق التركيبي .
- MOS ٣.٥ - ٤.٥ : جودة خاصة بالتطبيق الإتصالي .
- MOS ٤.٥ - ٤.٥ : جودة خاصة بالتطبيق الشبكي .
- MOS ٤.٥ - ٥.٥ : جودة خاصة بالتطبيق واسع الانتشار .

يختلف تقدير علامة معدل الرأي (MOS) بشكل مهم من اختبار إلى آخر لذا فهو لا يعد مقياساً مطلقاً للمقارنة بين الرموز المختلفة ، ويوضح الشكل (٦.٢) جودة إشارة الكلام لأنواع الثلاثة الرئيسية من الرموز والتي تختلف حسب معدل نقل البيانات للرمز [٥][٥٧] .

جودة إشارة الكلام

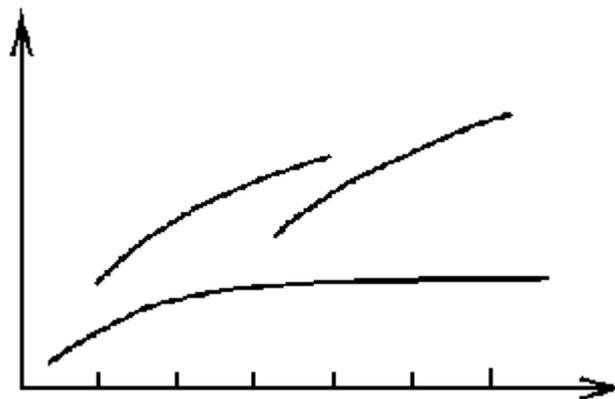


(Deg1)

معدل نقل البيانات (kbps)

ب. تميز

نارن مع جودة
سى جودة عالية
إلى المستويات
ع (لكنه غير
(Slightly A1
ة تميز صنف
بة . تقدم نتيجة



يتم إن
عالية أصليا
أصلية ثم يد
الخمسة ال
مزعج) (g
، مزعج (g
التشوه أو الأ

الاختبار كعلامة معدل الرأي التشوهي أو الأنخفاضي (Degradation Mean Opinion Score (DMOS) [٥٨][٦١].

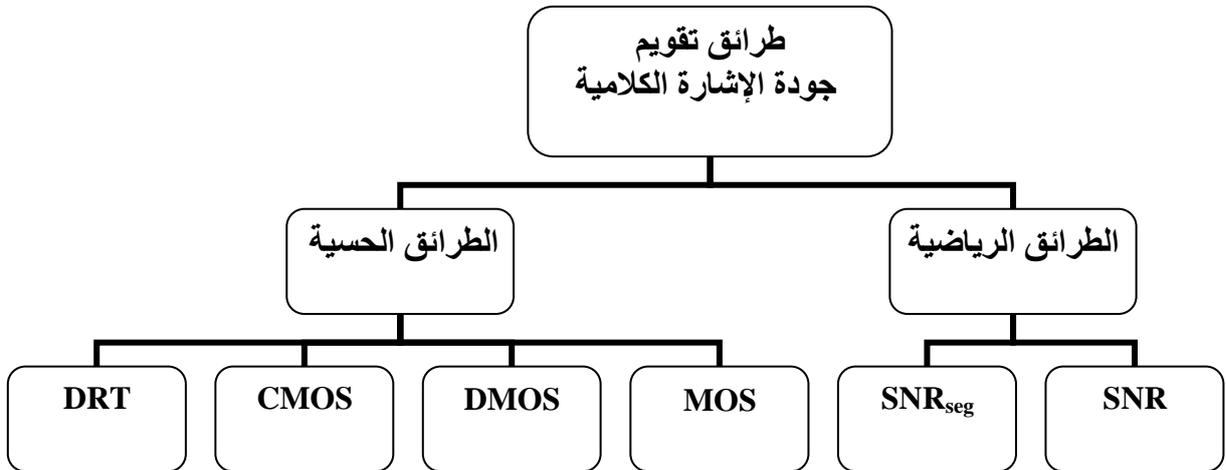
ت. تمييز المقارنة (Comparative Rating)

علامة معدل الرأي المقارن (Comparative MOS (CMOS)) هو اختبار إصغائي حيث يتم فيه الإستماع إلى نموذجين الأول يُعدّ مصدراً للمقارنة معه أما الثاني فيكون نموذجاً للاختبار حيث يرتب النموذجان بشكل عشوائي ثم ، يتمنان عن طريق السؤال الآتي : هل إن النموذج الأول مقارنةً مع النموذج الثاني ، أولاً: جيد بلا شك (Definitely Better) ، ثانياً : جيد بشكل طفيف (Slightly Better) ، ثالثاً : متساويان (Equal) ، رابعاً : سيء بشكل طفيف (Slightly Worse) ، خامساً : سيء بلا شك (Definitely Worse) تعطي هذه الطريقة مقياساً تقريبياً عن مدى أفضلية أو رداءة هذا الاختبار مقارنةً مع المصدر [٥٨][٦٢].

ث. مفهومية إشارة الكلام (Speech Intelligibility)

من الضروري القيام بقياس إمكانية فهم إشارة الكلام المفكوك ترميزها خاصة لرمز - فاك ترميز (Codec) الجودة الواطئة . يمكن القيام بقياس هذه الرموز ذات الجودة الواطئة بالإختبارات المفصلية (Articulation Tests) والتي تقيس النسبة المئوية للإشارات الكلامية والتي تستلم بشكل صحيح [٥] . الاختبار المفصلي هو اختبار القافية التشخيصي (Diagnostic Rhyme Test (DRT)) حيث يقوم المستمعون بسماع كلمات متكونة من مقطع واحد ومن السجع نفسه ومعروفة مسبقاً مثل (Meat And Beat) بعد ذلك يستمع المصغون إلى كلمة مسترجعة من هذه الكلمات ونختبر إمكانية التعرف عليها من ضمن الكلمات الأصلية . وطبيعياً يتراوح مدى علامة المفهومية بين ٧٠ إلى ٩٢ ، بينما تتراوح علامات نظام الهاتف عالي الجودة بين ٩٢ إلى ٩٦% [٦٣].

يوضح الشكل (٧.٢) طرائق تقويم جودة الإشارة الكلامية



شكل (٧.٢) طرائق تقويم جودة الإشارة الكلامية

الفصل الثالث

تقنية ترميز إثارة التنبؤ الخطي

(Code Excited Linear Predictive (CELP) Technique)

١.٣ المقدمة (Introduction)

تقنية ترميز إثارة التنبؤ الخطي (CELP) هي أحد طرائق ترميز إشارة الكلام الفعالة عند معدلات نقل (أو توليد) بيانات واطئة وقد اقترحت في منتصف الثمانينات من العالمين (Atal) و (Schroeder) [٦٤][٢] وتعني كلمة (CELP) ترميز إثارة التنبؤ الخطي. تعتمد تقنية ترميز إثارة التنبؤ الخطي على ثلاثة مفاهيم وهي :-

١. استخدام نموذج التنبؤ الخطي (LP) لنمذجة القناة الصوتية (Vocal Tract).
٢. استخدام مدخلات كتاب الرموز (المتكيف والثابت) كإدخال (الإثارة) لنموذج التنبؤ الخطي.
٣. يتم البحث بدورة مغلقة في المجال الموزون (Weighted Domain) [٥٧].

تُعدّ رموزات CELP رموزات قريبة من رموزات الشكل الموجي لأنها تنتج شكل موجة خارجة قريبة من شكل الموجة الأصلية. يوفر ترميز إثارة التنبؤ الخطي تجاوزاً لحواجز معظم المقاييس الحكومية لذا تم تطبيق بحثه في اتصالات الراديو النقال [١]، حيث إنه مجهزنا بترميز عند معدلات جيدة جداً وذات جودة ممتازة مقارنة مع رموزات إشارة الكلام عند معدلات متوسطة وذلك بربط الفجوة بين رموزات الشكل الموجي والرموزات الصوتية [٥٣]، ويعتمد على إجراءات بحث التحليل بالتركيب (تحليل الحلقة المغلقة) والتنبؤ الخطي. يمثل ترميز إثارة

التنبؤ الخطي كما مبين في الشكل (١.٣) طريقة تحليل وتركيب ضمن دورة مغلقة فعالة لأنظمة ترميز الإشارة الكلامية ذات نطاق ترددي ضيق ومتوسط وبمعدل نقل بيانات ضمن الحدود (٤- ١٦ kbps) [٦٤][٦٥].

يتضمن تحليل ترميز إثارة التنبؤ الخطي ثلاثة وظائف أساسية وهي :

١. التنبؤ الخطي Linear Prediction .

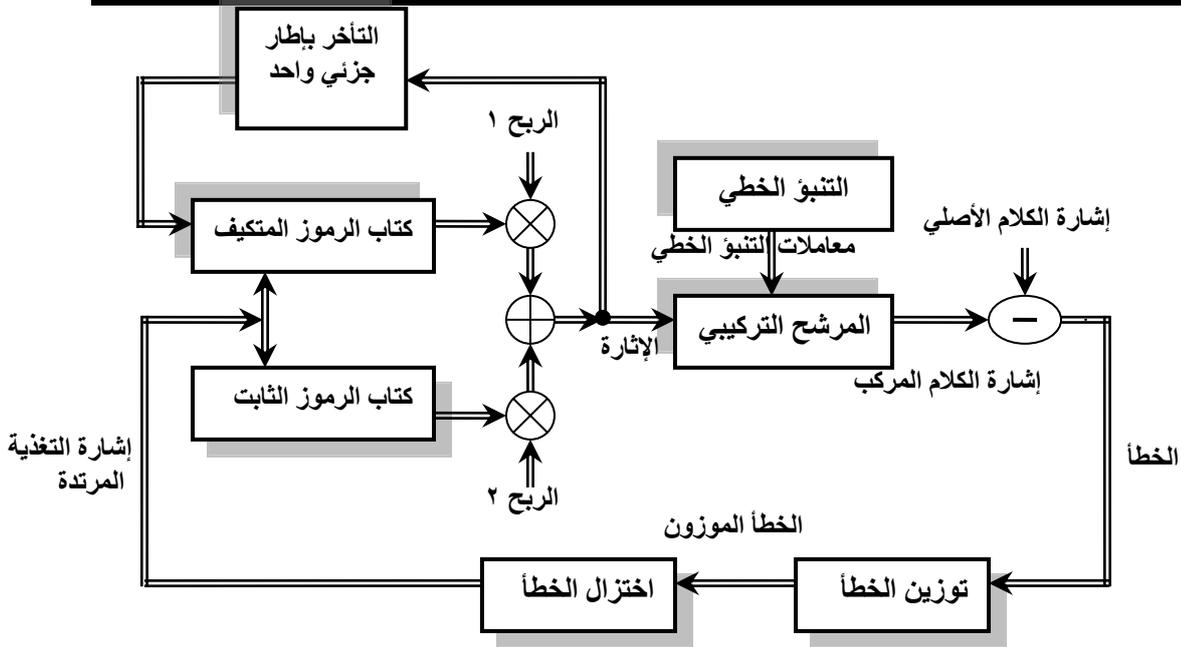
٢. بحث كتاب الرموز المتكيف Adaptive Codebook Search .

٣. بحث كتاب الرموز الثابت Fixed Codebook Search .

مع إضافة مثلى لوظيفة رابعة وتسمى مرشح التحسين لتحسين إشارة الكلام الخارج [٥٤] .

تقطع الإشارة الكلامية الداخلة إلى مجاميع من العينات تدعى الإطارات (Frames) (و المّة الزمنية النموذجية لكل إطار (ms ١٠) إلى (ms ٣٠)) وتعالج كوحدة واحدة . عندما تحلل الإشارة الكلامية ينجز أولاً تحليل التنبؤ الخطي عبر إطار الإشارة الكلامية الكلي لاستخلاص معاملات التنبؤ الخطي بعد ذلك يجزء كل إطار إلى عدد من الإطارات الجزئية (Subframes) (المّة الزمنية النموذجية لكل إطار جزئي هي (ms ٧.٥)) ومن ثم يطبق عليها بحث كتاب الرموز [٢][٦٥] ، حيث يتم اختيار متجهات الإثارة المقيسة المثالية (Optimal Scaled Excitation Vectors) من كتاب الرموز لتقليل المجموع الموزون للخطأ المربع بين موجة إشارة الكلام الأصلي وموجة إشارة الكلام المركب ، ويعتمد التعقيد وجودة إشارة الكلام المسترجع على حجم كتاب الرموز. إن عملية تركيب الإشارة الكلامية هي ببساطة عكس عملية تحليل الإشارة الكلامية [٥٤].

لا تتمكن مرمزات ترميز إثارة التنبؤ الخطي التي تعمل عند معدلات نقل بيانات منخفضة لإعادة إنتاج مكونات التردد العالية أو إزالتها في الإشارة الكلامية حيث تعاني الإشارة المعاد بناؤها من بعض الانخفاض والذي يستحق الأخذ بنظر الاعتبار للأصوات ذات الطبقة العالية مع نبضات إثارة المزمارة (Glottal Excitation Pulses) القوية . تعزى هذه المحددات (في بعض الأحيان) إلى حجم كتاب رموز الإثارة وكثافة محتوياته . (بصوره عامة) إن تحسن دقة إعادة الإنتاج يعزى إلى :- أولاً زيادة حجم كتاب رموز الإثارة [٢] ، لكن إذا كان كتاب الرموز ذا حجم كبير جداً فإن البحث سوف يستغرق وقتاً كبيراً . لذا حلت المشكلة باستخدام كتابي رموز صغيرة بدلاً من كتاب رموز واحد ذي حجم كبير إذ يتم تثبيت أحد كتابي الرموز من قبل المصمم وهو كتاب الرموز الثابت والثاني هو كتاب الرموز المتكيف ، حيث يبدأ فارغاً ويمتلئ خلال التشغيل مع نسخ من البقايا السابقة المتأخرة بكميات مختلفة [٦٥] ، وثانياً عندما تكون محتويات كتاب الرموز قريبة جداً من توزيع إثارة إشارة الكلام [٢].



شكل (١.٣) مرمز- فاك ترميز CELP للتحليل بالتركيب

٢.٣ التنبؤ الخطي (Linear Predictive (LP))

تعدّ طريقة ترميز التنبؤ الخطي إحدى تقنيات تحليل إشارة الكلام الأكثر قوة والكفاءة والمفيدة ، حيث أصبحت هذه الطريقة تقنية سائدة لتمثيل إشارة الكلام بمعدل خزن ونقل منخفض . وتوفر حسابات دقيقة وكفاءة لمعاملات الإشارة الكلامية [٦٦] . يُعدّ التنبؤ الخطي الأساس لعدد من تقنيات ترميز إشارة الكلام ومن ضمنها مرمزات إثارة التنبؤ الخطي [١] وفكرتها هي إمكانية تقريب عينة الإشارة الكلامية كتجميع خطي لعينات إشارة الكلام السابقة بتقليل مجموع مربع الخطأ (عبر مة زمنية محددة) بين عينات الإشارة الكلامية الحقيقية والمتنبئ بها خطأً .

وتأتي أهمية هذه الطريقة كونها طريقه قوية ويمكن الوثوق بها حيث إنها دقيقة لحساب المعاملات التي تميز نظام تغير الزمن الخطي وفي سرعتها في الحساب [٥٣].

تعتمد كفاءة طريقة تحليل الإشارة الكلامية على نموذج يعكس آلية إنتاج هذه الإشارة ويوضح الشكل (٢.٣) نموذج آلية الإنتاج من القناة الصوتية والتي تُعدّ مرشح تغير الزمن الخطي والذي يثار بمصدر نبضة متكرر أو بمصدر ضوضاء بيضاء . يمكن التعبير عن دالة انتقال $H(z)$ القناة الصوتية كالاتي [٥٣][٦٦] :-

$$H(z) = \frac{S(z)}{U(z)} = \frac{G}{1 - \sum_{k=1}^{order} \text{Alpha}_k * z^{-k}} \quad (1-3)$$

حيث يمثل كل من

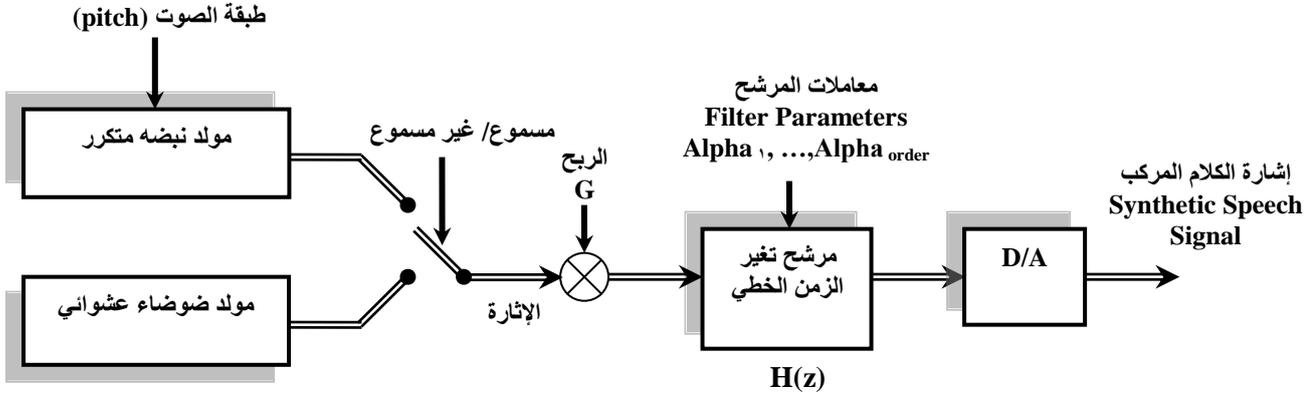
G : الربح (Gain)

$Alpha_k$: معاملات التنبؤ الخطي مستخرجة بخوارزمية ليفنسن- داربن التكرارية والموضحة

في الجزء (٤.٢.٣)

z : تحويل z (z - Transform)

$order$: رتبة المرشح



شكل (٢.٣) النموذج المبسط لآلية إنتاج الإشارة الكلامية

في الشكل (٢.٣) نترابط عينات الإشارة الكلامية $S(n)$ مع الإثارة $u(n)$ بالمعادلة الآتية :

$$S(n) = \sum_{k=1}^{order} Alpha_k * S(n-k) + G * u(n) \quad (٢-٣)$$

حيث إن

$S(n)$: إشارة الكلام الأصلية.

$u(n)$: مصدر الإثارة .

n : رقم العينة .

إذ يتم التنبؤ بالإشارة المسترجعة من التجميع الخطي لعينات إشارة الكلام السابقة وكما في المعادلة الآتية:-

$$\hat{S}(n) = \sum_{k=1}^{order} Alpha_k * S(n-k) \quad (٣-٣)$$

حيث إن

$\hat{S}(n)$: تمثل إشارة الكلام المسترجعة

يمكن حساب خطأ التنبؤ كالاتي:

$$e(n) = S(n) - \hat{S}(n) = S(n) - \sum_{k=1}^{order} Alpha_k * S(n-k) \quad (٤-٣)$$

إن هدف تحليل التنبؤ الخطي هو إيجاد معاملات التنبؤ الأفضل والتي تقلل قيمة طاقة الخطأ على طول الإطار (L) [٦٥].

$$E = \sum_{n=0}^{L-1} [e(n)]^2 = \sum_{n=0}^{L-1} \left[S(n) - \sum_{k=1}^{order} \text{Alpha}_k * S(n-k) \right]^2 \quad (٥-٣)$$

توجد ثلاث طرائق أساسية للحصول على معاملات التنبؤ الخطي وهي طريقة الارتباط الذاتي وطريقة التباين المشترك وطريقة الشبكية ، وسوف نقترن على ذكر طريقة الارتباط الذاتي لكونها الأفضل في مجال اهتمام موضوع بحث الرسالة . وبسبب طبيعة تغير المدة الزمنية للإشارة الكلامية فإن معاملات التنبؤ يجب أن تحسب مباشرة من إطارات الإشارة الكلامية [١] .

١.٢.٣ طرائق تقنية ترميز التنبؤ الخطي (LPC)

هنالك مجموعة من طرائق تقنية ترميز التنبؤ الخطي التي تستعمل للحصول على معاملات التنبؤ الخطي للإشارة الكلامية ، وينبع الاختلاف بين هذه الطرائق من فلسفة أو طريقة استعراض المشكلة وكذلك من الاختلاف في تفاصيل احتساب معاملات التنبؤ وهذه الطرائق هي : [٦٦]

١. طريقة الارتباط الذاتي (Autocorrelation Method)
٢. طريقة التباين المشترك (Covariance Method) .
٣. الطريقة الشبكية (Lattice Method) .
٤. صيغة المرشح المعكوس (Inverse Filter Formulation) .
٥. صيغة التخمين الطيفي (Spectral Estimation Formulation) .
٦. صيغة الإمكان الأعظم (Maximum Likelihood Formulation) .
٧. صيغة الضرب الداخلي (Inner Production Formulation) .

وسوف يتم إيجاز طرائق تحليل التنبؤ الخطي الثلاث الأولى وذلك لكونها طرائق أساسية للتحليل بينما تُعد الطرائق الباقية مماثلة لواحدة من هذه الطرائق الثلاث .

⊗ الطريقة الأولى :- طريقة الارتباط الذاتي : في هذه الطريقة يتكون كل إطار من عدد من العينات والذي يتم ترشيحه عن طريق ضربه بنافذة بعد ذلك يتم احتساب دالة الارتباط الذاتي. توضع المعاملات الناتجة في مصفوفة متناظرة والتي يكون فيها عناصر كل قطر من أقطارها متساوية ويطلق عليها مصفوفة توبليتز [١] .

⊗ الطريقة الثانية :- طريقة التباين المشترك : في هذه الطريقة لا تستخدم نافذة ويمكن حل المصفوفة الناتجة من هذه الطريقة باستخدام طريقة جولسكي إلا إن المعاملات الناتجة تكون غير مستقرة بشكل مؤكد ، أما في حالة وجود عدد كبير من العينات فإنها تكون غالباً مستقرة [٥٣] .

✪ الطريقة الثالثة :- الطريقة الشبكية : في هذه الطريقة يتم حساب معاملات الإنعكاس نسبة إلى أنبوب الصوت وفيها لا تستخدم دالة الارتباط الذاتي كما إنها لا تستخدم نافذة إلا إنها تحتاج إلى حسابات أكثر من الطريقتين السابقتين . يمكن استخدام تقنية المعالجة المتوازية في حل هذه الطريقة بسبب هيكلها المنتظم وتكون معاملات الإنعكاس الناتجة مستقرة بصورة أكيدة حتى في حالة استعمال الحسابات الرياضية الدقيقة المحددة [٦٦].

٢.٢.٣ المحاسن الأساسية لطريقة التنبؤ الخطي

- تمتاز طريقة التنبؤ الخطي بعدد من المحاسن التي يمكن إيجاز الأهم منها بالآتي [١][٥٣]:-
١. يعد نموذج ترميز التنبؤ الخطي نموذجاً تحليلياً قابلاً للتطبيق .
 ٢. دقيق رياضياً .
 ٣. بسيط وسهل الإنجاز .
 ٤. يعمل بصورة جيدة في تطبيقات الترميز .

٣.٢.٣ طريقة الارتباط الذاتي (Autocorrelation Method)

وهي إحدى الطرائق للحصول على معاملات التنبؤ الخطي بتقليل خطأ التنبؤ $e(n)$ لكل عينات إطار [٦٦]. حيث تفترض أن إطار الإشارة الكلامية $S_m(n)$ المكون من عدد من العينات تكون قيمته صفراً خارج الفترة $0 \leq n \leq L-1$ ويضرب بناقذة $W(n)$ بطول مُحدد والتي تكون أيضاً صفراً خارج الفترة $0 \leq n \leq L-1$ وكما في المعادلة الآتية [١]:-

$$(٦-٣) \bar{S}_m(n) = S_m(n) * W(n)$$

حيث m يمثل رقم الإطار

يمكن حساب معاملات الارتباط الذاتي $Aut(i)$ لكل إطار من الإشارة الكلامية كالاتي:-

$$Aut(i) = \sum_{n=i}^{L-1-i} S_m(n) * \bar{S}_m(n-i) \quad i = 0, \dots, order \quad (٧-٣)$$

لطريقة الارتباط الذاتي تكون معادلة المصفوفة لحل معاملات التنبؤ بالصيغة الآتية [٦٧]:-

$$\sum_{k=1}^{order} Aut(|i-k|) * Alpha_k = Aut(i) \quad i = 1, \dots, order \quad (٨-٣)$$

ونعبر عن المعادلات السابقة بنظام المصفوفات الموضح في المعادلة (٩-٣)

$$\begin{bmatrix} Aut(0) & Aut(1) & Aut(2) & \dots & Aut(order-1) \\ Aut(1) & Aut(0) & Aut(1) & \dots & Aut(order-2) \\ Aut(2) & Aut(1) & Aut(0) & \dots & Aut(order-3) \\ \vdots & \vdots & \vdots & \dots & \vdots \end{bmatrix} \begin{bmatrix} Alpha_1 \\ Alpha_2 \\ Alpha_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} Aut(1) \\ Aut(2) \\ Aut(3) \\ \vdots \end{bmatrix}$$

(٩-٣)

المصفوفة الناتجة لمعاملات الارتباط الذاتي سوف تكون من نوع مصفوفة توبليتز (Toeplitz) ، وتتميز مصفوفة معاملات الارتباط الذاتي التي تم الحصول عليها بالمعادلة (٩-٣) بالمواصفات الآتية [٥٣][٦٦]:-

١. إذا كانت الإشارة الكلامية دورية فإن دالة الارتباط الذاتي تكون دورية أيضاً .
٢. تصل قيمتها العظمى في حالة $i = 0$ أي $Aut(0)$ هي قيمة الطاقة للإشارة .
٣. تكون دالة الارتباط الذاتي دالة زوجية أي $Aut(i) = -Aut(i)$.

ويمكن استغلال هذه الصفات للحصول على معاملات التنبؤ الخطي باستخدام العديد من الطرائق [١].

٤.٢.٣ طريقة ليفنسن داربن التكرارية

(Levenson Durbin Recursion Method)

إن استثمار طبيعة التوبليتز لمصفوفة معاملات الارتباط الذاتي الناتجة عن المعادلة (٩-٣) أدى إلى ابتكار العديد من الطرائق التكرارية لحل هذه المعادلة ، أكثرها كفاءة هي طريقة ليفنسن داربن التكرارية لما تتميز به من اختزال للمتطلبات الحسابية واختزال الكلفة واختزال متطلبات الخزن مقارنة مع الطرائق التقليدية الأخرى . يمكن توضيح الخطوات الأساسية لهذه الطريقة كالآتي [١][٦٧]:-

$$E^{(0)} = Aut(0) \quad (١٠-٣)$$

$$Alpha_0^{(i-1)} = 1 \quad (١١-٣)$$

$$Re f_i = - \left[\sum_{j=0}^{i-1} Alpha_j^{(i-1)} * Aut(i-j) \right] / E^{(i-1)} \quad (١٢-٣)$$

$$Alpha_i^{(i)} = Re f_i \quad (١٣-٣)$$

$$Alpha_j^{(i)} = Alpha_j^{(i-1)} + Re f_i * Alpha_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1 \quad (١٤-٣)$$

$$E^{(i)} = (1 - \text{Re } f_i^2) * E^{(i-1)} \quad (10-3)$$

المعادلات (3-11)-(3-10) تحل بالتكرار (recursive) للقيم $i = 1, 2, \dots, \text{order}$ والحل الأخير لهذه المعادلات هو :-

$$Alpha_k = Alpha_j^{order} \quad 1 \leq j, k \leq \text{order} \quad (16-3)$$

حيث تمثل كل من

$Aut(i)$: معاملات الارتباط الذاتي .

$Alpha_k$: معاملات التنبؤ الخطي .

Ref_i : معاملات الإنعكاس .

من الملاحظ إن معاملات الارتباط الذاتي $Aut(i)$ يعوض عنها بمعاملات طبيعية $Auto(i)$ (Normalized) .

$$Auto(i) = Aut(i) / Aut(0) \quad (Auto(i) \leq 1) \quad (17-3)$$

٣.٣ بحث كتاب الرموز (*The Codebook Search*)

يستخدم هذا النوع من الترميز نوعين من كتب الترميز وهما :

١.٣.٣ بحث كتاب الرموز المتكيف (*Adaptive Codebook Search*)

بعد إيجاد معاملات ترميز التنبؤ الخطي ينشط بحث كتاب الرموز المتكيف لتقليل إضافي للإشارة المتبقية . مبدئياً إذا كان مرشح القناة الصوتية منمذجاً بشكل مضبوط بوساطة مرشح التنبؤ الخطي فإن الإشارة الناتجة تمثل إشارة الإثارة المزمارية تماماً . إشارة الإثارة المزمارية هي دورية بطبيعتها وتدعى فترتها فترة طبقة الصوت أو تأخير طبقة الصوت [٥٤] . في ترميز إثارة التنبؤ الخطي تختار إشارة إثارة المزمار المركبة المطابقة إلى فترة طبقة الصوت المثلى لكل إطار جزئي من كتاب الرموز المتكيف خلال مخطط حساب الحلقة المغلقة . هذا الإجراء مكيف لإشارة الإثارة السابقة والتي تجمع إخراج بحث كتاب الرموز الثابت مع بحث كتاب الرموز المتكيف بسبب الإطار الجزئي السابق شكل (١.٣) [٦٥][٦٨] .

يعبر عن متجه كتاب الرموز المتكيف المقيس بالربح في عملية بحث كتاب الرموز المتكيف بالمعادلة الآتية [٥٤]:-

$$a_i(n) = G_1 * (u(n) - \alpha) \quad n = 0, \dots, N-1 \quad (18-3)$$

٣)

حيث يمثل

i : فهرس متجه كتاب الترميز المتكيف .

N : طول الإطار الجزئي .

α : التأخير وتبحث خوارزمية CELP عن اختيار أفضل تأخير بالاعتماد على فترة طبقة الصوت [٦٥] .

يشتق الربح (G_1) لكتاب الرموز المتكيف لكل إطار جزئي كما سيوضح في الجزء (٥.٣) إذ أن (G_1) يحسن جودة إشارة الكلام .

٢.٣.٣ بحث كتاب الرموز الثابت (*Fixed Codebook Search*)

بعد تحليل التنبؤ الخطي وبحث كتاب الرموز المتكيف ينشط بحث كتاب الرموز الثابت لتوليد إشارة الكلام المعاد بناؤه والذي ينتج من إمرار الترابط لمتجه الإثارة المتكيف المثالي مع متجه الإثارة الثابت خلال المرشح التركيبي التنبؤ الخطي [٦٨] . ويمتلك متجه الترميز الثابت تأثيراً كبيراً في جودة إشارة الكلام المعاد بناؤه وكلفة الاحتساب وحجم خزن الذاكرة حيث يُعدّ متجه الترميز الثابت المفتاح لكفاءة خوارزمية CELP [٦٩] .

يمتلك كل متجه ترميز في كتاب الرموز الثابت فهرس (*Index*) حيث يمكن الحصول على أفضل جودة إشارة كلامية باستعمال كتاب الرموز وبمجاهات ترميز أكثر وبشكل مشابه إلى بحث كتاب الرموز المتكيف ينجز بحث كتاب الرموز الثابت لكل إطار جزئي بعملية الحلقة المغلقة وكذلك تبنيه مقياس أصغر مربع خطأ [٦٧] .

يعبر عن متجه الإثارة الثابت المقيس بالربح في بحث كتاب الرموز الثابت بالمعادلة الآتية [٥٤]:-

$$v_i(n) = G_2 * c_i(n) \quad n = 0, \dots, N-1 \quad (19-3)$$

حيث يمثل i الفهرس لمتجه كتاب الرموز الثابت .

يتم اشتقاق الربح (G_2) لكتاب الرموز الثابت لكل إطار جزئي كما سيوضح في الجزء (٥.٣) إذ أن (G_2) يحسن جودة إشارة الكلام .

٤.٣ الترشيح (Filtering)

يُعدّ الترشيح أحد عمليات معالجة الإشارة المعقدة الواسعة الاستعمال ويستعمل الترشيح لإمرار مكونات تردد معينة في الإشارة خلال النظام بدون أي تشوهات وحجز مكونات تردد أخرى حيث إن النظام الذي ينفذ هذه العملية يدعى المرشح [٧٠].

مدى الترددات الذي يسمح بالإمرار خلال المرشح يدعى نطاق المرور (Passband) ، ومدى الترددات الذي يحجز من قبل المرشح يدعى نطاق التوقف (Stopband) [٧١].

نستطيع تعريف أنواع عديدة من المرشحات بالاعتماد على طبيعة عملية الترشيح .

* يمرر مرشح الإمرار الواطئ (Lowpass Filter) مكونات التردد الواطئ التي أقل من تردد قطع معين fc ويدعى (تردد قطع) (Cutoff Frequency) ويحجز مكونات التردد العالي التي أعلى من تردد fc .

* يمرر مرشح الإمرار العالي (Highpass Filter) كل مكونات التردد العالي التي أعلى من تردد قطع معين fc ويحجز كل مكونات التردد الواطئ التي أقل من fc .

* يمرر مرشح إمرار النطاق (Passband Filter) كل مكونات التردد بين ترددي قطع fc_1 و fc_2 ويحجز كل مكونات التردد أقل من التردد fc_1 وأعلى من التردد fc_2 .

* يحجز مرشح نطاق التوقف (Bandstop Filter) كل مكونات التردد بين ترددي قطع fc_1 و fc_2 ويمرر كل مكونات التردد أقل من التردد fc_1 وأعلى من التردد fc_2 [٧٠][٧٢].

قد يحصل تشوه (Corrupted) للإشارة بإشارة التداخل (Interfering Signal) أو إشارة الضوضاء (Noise) . في مثل هذه الحالات يمكن معالجة الإشارة المرغوب بها من الإشارة المشوهة بالضوضاء بإمرار الأخيرة خلال مرشح الإمرار الواطئ مع تردد قطع fc [٦٩][٧١] وعلى هذا الأساس فقد استعمل في هذا البحث مرشح الإمرار الواطئ لاستجابة النبضة المتناهية (Lowpass FIR Filter) لتخليص إشارة الكلام المسترجعة من الضوضاء مع تردد قطع ($fc = ٣.٤$ kHz) وبذلك تكون إشارة الكلام المسترجعة قريبة قدر المستطاع من إشارة الكلام الأصلية .

٥.٣ مرحلة احتساب معاملات الإثارة

(Calculation of the Excitation Parameters Stage)

عادةً يكون المرشح التركيبي بالصيغة

$$H(z) = 1/A(z) \quad (٢٠-٣)$$

حيث

$$A(z) = 1 - \text{Alpha}_1 * z^{-1} - \text{Alpha}_2 * z^{-2} \dots \text{Alpha}_{order} * z^{-order} \quad (٢١-٣)$$

أما مرشح الخطأ الموزون فتكون صيغته

$$W(z) = A(z)/A(z/y) \quad (٢٢-٣)$$

حيث y هو ثابت تكون قيمته بين (٠-١) ($٠ < y < ١$) والقيمة النموذجية له ($y=٠.٩$) ويعتمد مرشح الخطأ الموزون على معاملات مرشح التنبؤ الخطي [٦٧] ، سلسلة المرشح التركيبي ومرشح الخطأ الموزون يكافئ استخدام مرشح التركيب الموزون بالصيغة $١/A(z/y)$. وبالاعتماد على مبادئ الحلقة المغلقة (طريقة التحليل بالتركيب) المستخدمة في خوارزمية (CELP) فإنه يجب اختزال الخطأ الموزون [٦٥].

$$e_w(n) = \hat{S}_w(n) - S_w(n) \\ = \hat{S}_w(n) - S_0(n) - G_1[(u(n) - \alpha) * h(n)] - G_2[c_i(n) * h(n)] \quad n = ٠, \dots, N-1 \quad (٢٣-٣)$$

حيث يمثل

$S_w(n)$: إشارة الكلام المدخل الموزون .

$$S_w(n) = S(n) + \sum_{k=1}^{order} \text{Alpha}_k * y * S(n-k) \quad n = ٠, \dots, N-1 \quad (٢٤-٣)$$

$\hat{S}_w(n)$: إشارة الكلام المركب الموزون .

$\hat{S}_0(n)$: استجابة الإدخال الصفرية للمرشح التركيبي الموزون نسبة إلى إدخاله في الإطار الجزئي السابق .

$h(n)$: استجابة النبضة للمرشح التركيبي الموزون [٦٦].

$$h(n) = \sum_{k=1}^{order} \text{Alpha}_k * h(n-k) + AE * \delta(n) \quad n = ٠, \dots, N-1 \quad (٢٥-٣)$$

حيث يمثل

$\delta(n)$: وحدة العينة (Sample Unite) أو وحدة النبضة (Impulse Unite) .

AE : سعة الإثارة (Amplitude of The Excitation) .

$$AE = \sqrt{Auto(0) - \sum_{k=1}^{order} Alpha_k * Auto(k)} \quad (26-3)$$

وتعمل طريقة بحث كتاب الرموز على إيجاد قيم الربح لكتبي الرموز المتكيف والثابت G_1, G_2 على التوالي الذي يختزل معدل مربع الخطأ E_w المأخوذ على طول الإطار الجزئي (N) ويكتب هذا الخطأ كالاتي [٦٥] :-

(٢٧-

$$E_w = \frac{1}{N} \left(\sum_{n=0}^{N-1} X^2(n) - T_{ai} \right) \quad (3)$$

حيث يمثل T_{ai} : حد عتبة خاضع للتعظيم .

$$T_{ai} = 2 * (G_1 * C_\alpha + G_2 * C_i - G_1 * G_2 * Y_{ai}) - G_1^2 * \zeta_\alpha - G_2^2 * \zeta_i \quad (28-3)$$

ζ_α : طاقة إشارة كتاب الرموز المتكيف

$$\zeta_\alpha = \sum_{n=0}^{N-1} [(u(n) - \alpha) * h(n)]^2 \quad (29-3)$$

C_α : الترابط بين إشارة كتاب الرموز المتكيف وإشارة الهدف $X(n)$

$$C_\alpha = \sum_{n=0}^{N-1} X(n) * [(u(n) - \alpha) * h(n)] \quad (30-3)$$

ζ_i : طاقة إشارة كتاب الرموز الثابت

$$\zeta_i = \sum_{n=0}^{N-1} [c_i(n) * h(n)]^2 \quad (31-3)$$

C_i : الترابط بين إشارة كتاب الرموز الثابت وإشارة الهدف $X(n)$

$$C_i = \sum_{n=0}^{N-1} X(n) * [c_i(n) * h(n)] \quad (32-3)$$

Y_{ai} : الترابط بين الإشارات من كتابي رموز .

$$Y_{ai} = \sum_{n=0}^{N-1} [(u(n) - \alpha) * h(n)] * [c_i(n) * h(n)] \quad (33-3)$$

و $X(n)$: الإشارة الهدف لبحث كتاب الرموز .

$$X(n) = S_w(n) - \hat{S}_o(n) \quad n = 0, \dots, N-1 \quad (34-3)$$

ومن أجل تحسين بحث الحلقة المغلقة لكتابي الرموز المتكيف والثابت فقد تم استخدام الطريقة الآتية لإيجاد معالم كتب الرموز [65][68] :-

١. جعل $(G_2 = 0)$ بالمعادلة رقم (3-28) .
٢. ومن ثم بإعطاء قيمة إلى التأخير α فان G_1 نستطيع إيجاده بضبط الاشتقاق الجزئي إلى T_{ai} بالترتيب مع G_1 إلى الصفر .
٣. ثم إيجاد قيمة T_{ai} لكل قيم α وإختيار تأخير كتاب الرموز المتكيف الذي يعظم T_{ai} .
٤. وبتثبيت معالم كتاب الرموز المتكيف بطريقه مشابهة يمكننا إيجاد معالم كتاب الرموز الثابت.

إن مساواة الاشتقاق الجزئي لـ T_{ai} بالترتيب مع G_1 و G_2 إلى الصفر يعطي زوج من المعادلات التي نستطيع حلها لتعطي قيماً مثالية ومحسنة للربح وهذه القيم هي [65][67] :-

$$G_1 = \frac{C_a * \zeta_i - C_i * Y_{ai}}{\zeta_a * \zeta_i - Y_{ai}^2} \quad (35-3)$$

$$G_2 = \frac{C_i * \zeta_a - C_a * Y_{ai}}{\zeta_a * \zeta_i - Y_{ai}^2} \quad (36-3)$$

تحسب طريقة العمل الحدود $(\zeta_a, \zeta_i, C_a, C_i, Y_{ai})$ والتي تستخدم في حساب الربح G_1 و G_2 . فيما بعد نستطيع تعويض هذه الأرباح في معادلة (3-28) لإعطاء T_{ai} حيث

الرمز يمتلك أعلى قيمة بالاختيار الصحيح إلى α [٦٥]. استخدام هيكل كتاب الرموز الجبري الثابت (Algebraic Fixed Codebook) يسمح بإيجاد \hat{c}_i و C_i بشكل كفوء وباستعمال سلسلة من أربع دورات متشابهة [٤٦] [٦٩].

الفصل الرابع

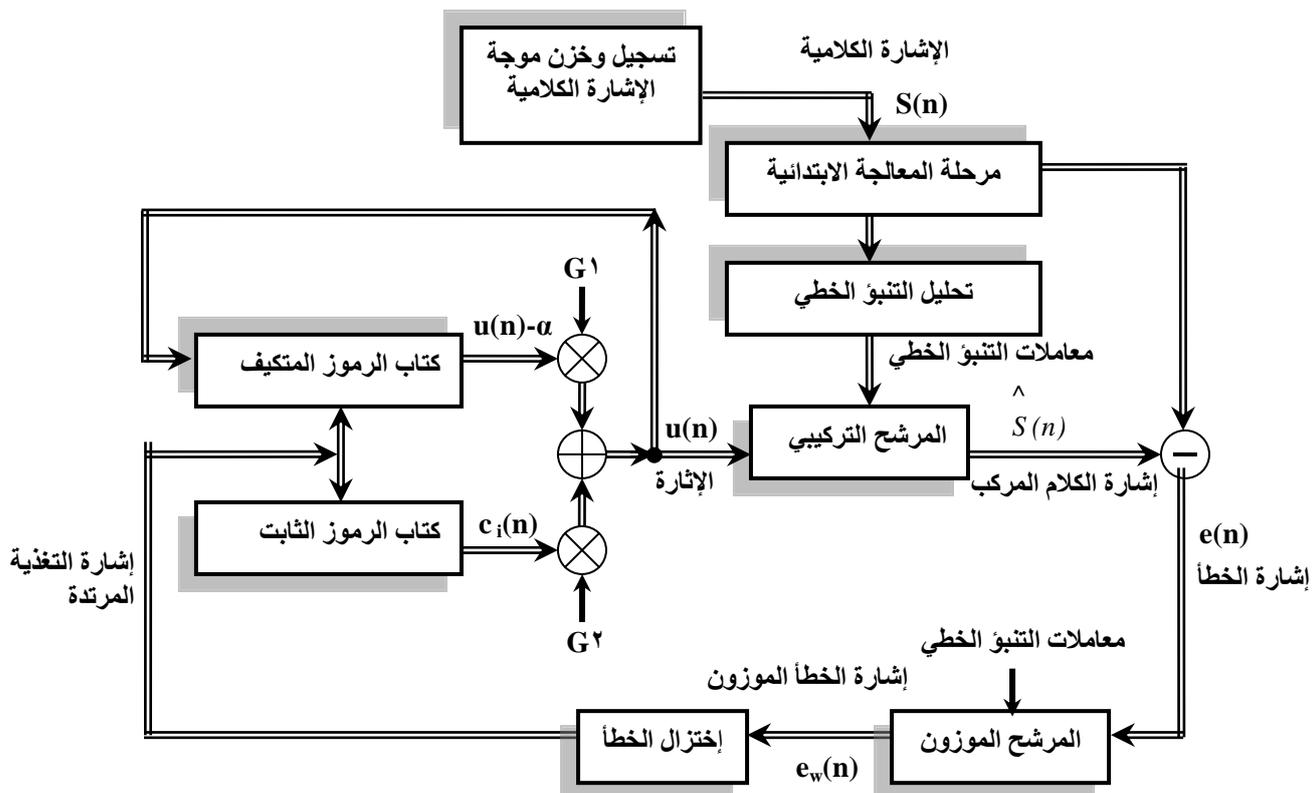
تصميم منظومة برمجية لمحاكاة المرز المصدرى

الهجين

المعتمد على تقنية التحليل بالتركيب

١.٤ مراحل تصميم المنظومة

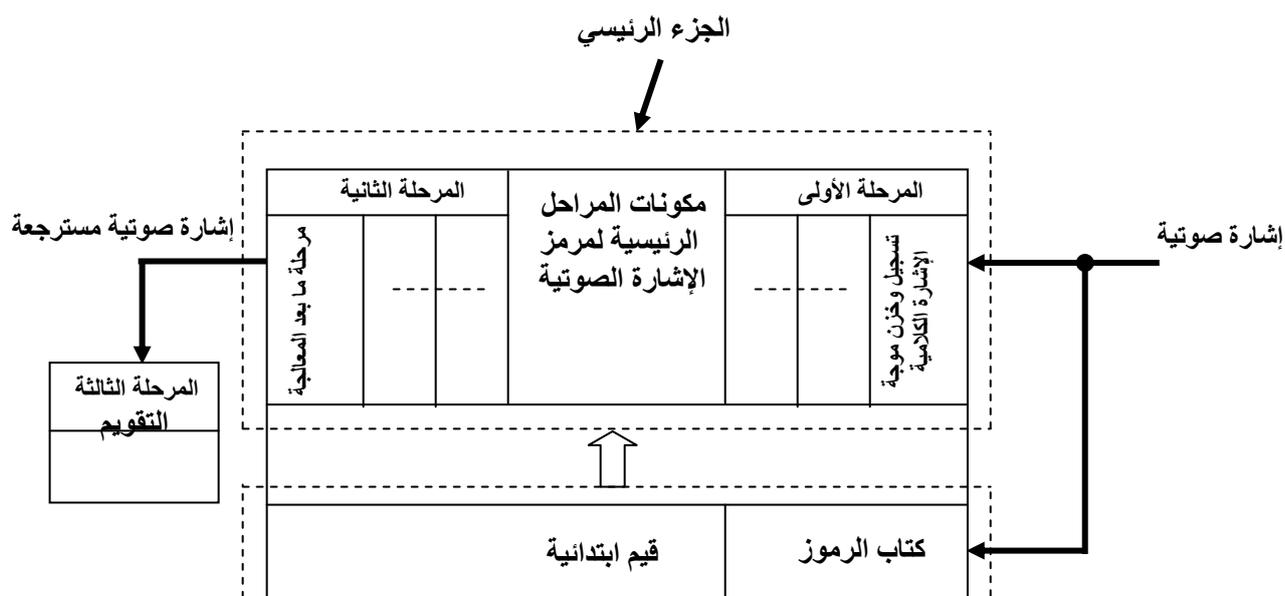
نقدم في هذا الفصل الخطوات الأساسية لتصميم مرز صوتى ذي معدل نقل بيانات واطى لتعزير كفاءة الإشارة الصوتية الناتجة بعد استرجاع الصوت حيث يشمل تصميم هذا المرز عدد من المراحل تبدأ بإدخال الإشارة الكلامية (الكلام المدخل (Input Speech)) (بعد تسجيل الإشارة الكلامية) وكيفية إجراء عمليات التجزئة والتحليل لهذه الإشارات ومن ثم بناء المرز المناسب وتنتهي باسترجاع إشارة الكلام المسترجع (الكلام المركب (Synthetic Speech)) مع إنجاز جودة كلام عالية عند معدل نقل بيانات (٤ kbps) وبدون اختلاف في الإشارة الكلامية الأصلية والمسترجعة أي المعاد بنائها وكما موضح بالمخطط التالي شكل رقم (١.٤) :-



شكل (١.٤) المرز - فاك ترميز إثارة التنبؤ الخطي (CELP)

يمكن وصف الشكل العام لأي مرز صوتي من نوع (AbS) بالشكل رقم (٢.٤) والذي يتكون هيكل مكوناته الرئيسية من جزئين :

الجزء الرئيسي يتعلق بالمرحلتين الرئيسيتين من مكونات المرز الصوتي ، والجزء الساند الذي يتكون من مرحلة إعداد معاملات كتاب الرموز والقيم الابتدائية التي يتحدد على أساسها سرعة الصوت المرز .



شكل (٢.٤) هيكل مكونات المرز الصوتي

يتكون النظام المقترح من ثلاثة مراحل أساسية وهي :

١. مرحلة الترميز (Encoding) ٢. مرحلة فك الترميز (Decoding) ٣. مرحلة التقويم (Evaluation) .

١. مرحلة الترميز: تتألف مرحلة الترميز بدورها من عدة مراحل فرعية ، وهي :-

- أ- مرحلة تسجيل وخرن موجة إشارة الكلام بملف صوتي .
- ب- مرحلة المعالجة الابتدائية وذلك من أجل تهيئة الملف الصوتي للمعالجة .
- ت- مرحلة تحليل التنبؤ الخطي والتي يتم فيها تجزئة الإشارة الكلامية المدخلة إلى عدد من الإطارات (Frames) ، ثم يمرر كل إطار من هذه الإطارات خلال تحليل التنبؤ الخطي لحساب معاملات التنبؤ الخطي .
- ث- مرحلة بناء كتابي رموز وهما كتاب الرموز المتكيف وكتاب الرموز الثابت وحساب معالمهما بصورة تؤدي إلى تقليل الخطأ الموزون بين إشارة الكلام الأصلية والمعاد بناؤها وقد تم تطبيق تقنية مثالية لتحسين حلقة بحث كتابي الرموز بينما نحاول الحفاظ على مستوى مقبول من التعقيد .
- ج- مرحلة حساب معالم الإثارة لكل إطار جزئي بعد تجزئة كل إطار إلى أربعة إطارات جزئية ، والذي تم شرحه في البند (٥.٣) .
- ح- مرحلة حساب إشارة الإثارة التي تستخدم لاحقاً لإعادة بناء إشارة الكلام .

٢. مرحلة فك الترميز وتتضمن المراحل الآتية :

- أ- مرحلة إعادة بناء إشارة الكلام الخارج (الكلام المركب) باستخدام المعاملات وإشارة الإثارة التي تم الحصول عليهم من مرحلة الترميز .
- ب- مرحلة حساب وتوزين واختزال الخطأ .
- ت- مرحلة ما بعد المعالجة لتحسين جودة الإشارة الكلامية المسترجعة .

٣. مرحلة التقويم وتتضمن المراحل الآتية :

- أ- تقويم إشارة الكلام الخارج باستخدام التقويم الرياضي .
- ب- تقويم إشارة الكلام الخارج باستخدام التقويم الحسي .

سوف نأخذ كل مرحلة والمراحل الجزئية التي تتكون منها بشيء من التفصيل وكالاتي:-

٢.٤ الوصف العام للمرمز

(The General Description of The Coder)

المرمز - فاك الترميز المقترح هو ترميز إثارة التنبؤ الخطي (CELP). والذي يرمز عينات إشارة الكلام بطريقة التحليل بالتركيب (تحليل الحلقة المغلقة)، ويتكون من ثلاثة أجزاء أساسية وهي التنبؤ الخطي، وبحث كتاب الرموز المتكيف، وبحث كتاب الرموز الثابت.

يُقطع سبل الإشارة الكلامية المدخلة إلى مرز إثارة التنبؤ الخطي إلى إطارات ذات مدة زمنية (30 ms) ثم يُقطع كل إطار إلى أربعة إطارات جزئية طول كل منها (7.5 ms) وعندما يحل الكلام يُنجز أولاً تحليل التنبؤ الخطي عبر إطار إشارة الكلام الكلي لاستخلاص معاملات التنبؤ الخطي، بعد ذلك يتم حساب معاملات الإثارة لكل إطار جزئي، وبعد استخلاص المعالم لنموذج ترميز إثارة التنبؤ الخطي وهي (معاملات مرشح التنبؤ الخطي ومؤشرات وإرباح كتاب الرموز الثابت والمتكيف) تستخدم هذه المعالم عند فاك الترميز لاسترجاع إشارة الكلام المعاد بناؤه بترشيح الإثارة خلال المرشح التركيبي (Synthesis Filter) ويعتمد المرشح التركيبي على مرشح التنبؤ الخطي ذي الرتبة (10) حيث تم اختيار هذه الرتبة بعد سلسلة من التجارب. بعد احتساب إشارة الكلام المعاد بناؤه يحسن إضافياً بمرحلة ما بعد المعالجة (Post Processing Stage).

١.٢.٤ تسجيل النص المنطوق (*Spoken Utterance Recording*)

١. لقد تم التسجيل باستخدام النظام البرمجي (Cool Edit ٢٠٠٠).

٢. تم اختيار ترددي نمذجة وهما ٤٤١٠٠ و ٨٠٠٠ هرتز والذي يُعدّ المعدل المناسب مع إشارة الكلام المنقولة بالتراسل الصوتي في منظومات الاتصالات ويُعدّ معدل العينات هذا ذا كلفة احتساب وخرن معقولة حيث إن تقليل هذا المعدل يؤدي إلى خسارة كبيرة في المعلومات التي تخص الإشارة أما في حالة اختيار معدل عينات كبير فهذا يعني الوقوع في مشاكل الحسابات المعقدة وتوفير مساحات الخزن الكبيرة.

٣. تم تسجيل مجموعة من إشارات الاختبار ولمختلف المتكلمين (رجالاً ونساءً) وفي أوقات مختلفة.

٤. تم قراءة ومعالجة الإشارة بالكامل ومن دون إهمال أي مقطع منها علماً إن أطوال الإشارات كانت مختلفة .
٥. كما تم تخزين إشارة الكلام الناتجة من مرحلة التسجيل في ملف صوتي (Wave File) مقسم إلى جزئين أساسيين الجزء الأول يسمى مقدمة الملف (Header File) وحجمه (٤٤) بايت بينما تخزن البيانات (عينات إشارة الكلام) في الجزء الآخر .
٦. تم التسجيل في غرفة مناسبة للتسجيل وباستخدام اللاقطة نفسها في كل عملية وتحت ظروف اعتيادية.

٢.٢.٤ مرحلة المعالجة الابتدائية (Preprocessing Stage)

يجب تهيئة الملف الصوتي لعملية الترميز بعد مرحلة التسجيل مباشرةً لذلك يجب البدء في مرحلة المعالجة الابتدائية والتي تُعدّ أحد المراحل المهمة في معالجة إشارة الكلام وتتضمن هذه المرحلة خطوتين أساسيتين وهما : -

١. إزالة السكون (Silence) يدوياً من بداية ونهاية إشارة الكلام بالنظام البرمجي (Cool Edit ٢٠٠٠) .
٢. يتم تحويل الملف الصوتي (Wave File) ذي الامتداد (wav) إلى ملف بيانات (Data File) ذي امتداد (dat) باستخدام المحاكاة بالحقيبة البرمجية MATLAB وذلك لدقة وكفاءة الـ MATLAB في الاحتساب وعدم فقدان أي بيانات من ملف الصوت .

٣.٢.٤ مرحلة تحليل التنبؤ الخطي (The Linear Predictive Analysis Stage)

في هذه المرحلة يتم تطبيق التقنيات الملائمة لإستخلاص المعاملات الأساسية لإشارة الكلام وكما موضح بالمخطط الإنسيابي في الشكل (٣.٤).

حيث تم إستخلاص معاملات المرشح التركيبي بتحليل التنبؤ الخطي وبصيغة الارتباط الذاتي مع نافذة متناظرة ذات مّة زمنية (٣٠ ms) وتستخدم خوارزمية ليفنسن دارين لتحويل معاملات الارتباط الذاتي إلى معاملات التنبؤ الخطي وحسب الخطوات الآتية :-

١. قسمة بيانات ملف البيانات إلى عدد من الإطارات المّة الزمنية لكل إطار هي (٣٠ ms).
٢. ضرب كل إطار بنافذة وتتضمن نافذة تحليل التنبؤ الخطي المتناظرة نافذة هجينة ذات إتجاهين :-

الاتجاه الأساسي : نافذة هجينة مكونة من جزئين ، الجزء الأول يعتمد على دالة الهامنك ويتعامل مع بداية ثلاثة أرباع إطار الإشارة الكلامية ، أما الجزء الثاني من النافذة فيتضمن استخدام الربع الأول من دالة الجيب تمام والذي يتعامل مع الربع الأخير من إطار الإشارة الكلامية ، كما موضح في الشكل (٤.٤) لذلك تعرف النافذة W_{lp} رياضياً كالاتي :-

$$W_{lp}(n) = \begin{cases} 0.54 - 0.46 * \cos(2 * \pi * n / 199) & \text{if } n = 0, \dots, 199 \\ \cos((2 * \pi * (n-200)) / 159) & \text{if } n = 200, \dots, 239 \end{cases} \quad (١-٤)$$

الاتجاه المقترح : نافذة هجينة مكونة من جزئين ، الجزء الأول يعتمد على الربع الثاني والربع الثالث من دالة الجيب تمام وتتعامل مع بداية نصف إطار الإشارة الكلامية ، ويستخدم الجزء الثاني من النافذة المقترحة دالة الهامنك ، حيث تعامل النصف الأخير من إطار الإشارة الكلامية ، كما موضح في الشكل (٥.٤) . لذلك تعرف النافذة الهجينة المقترحة رياضياً كالاتي :-

$$W_{lp}(n) = \begin{cases} \cos(X + \pi/2) & \text{if } n = 120, \dots, 239 \\ 0.54 - 0.46 * \cos(2 * \pi * n / 119) & \text{if } n = 0, \dots, 119 \end{cases} \quad (٢-٤)$$

حيث

$$X = ((2 * \pi * (n-0)) / 240) \quad n=0, \dots, 119 \quad (٣-٤)$$

٤)

وتعرف إشارة الكلام المضروبة بنافاذة كالاتي :-

$$\bar{S}(n) = W_{lp}(n) * S(n) \quad n=0, \dots, 239 \quad (٤-٤)$$

٤)

٣. حساب معاملات الارتباط الذاتي (Auto-Correlation Coefficients) وفقاً للترتبة (order = ١٠) وبالاعتماد على المعادلة (٧-٣).

٤. تعديل معاملات الارتباط الذاتي وذلك بضرب معاملات الارتباط الذاتي المحسوبة أعلاه بالعامل الآتي:

$$W_{lag}(k) = \exp \left[-\frac{1}{2} \left(\frac{2 * \pi * f_0 * k}{f_s} \right)^2 \right] \quad k=1, \dots, 10 \quad (٥-٤)$$

حيث يمثل

f_0 : وسع النطاق الترددي ويساوي (٦٠ Hz)

f_s : تردد النمذجة

كما ويضرب $Aut(0)$ بمعامل ترابط الضوضاء البيضاء الذي يساوي (1.0001) ولذلك تعطى معاملات الارتباط المعدلة كالآتي :

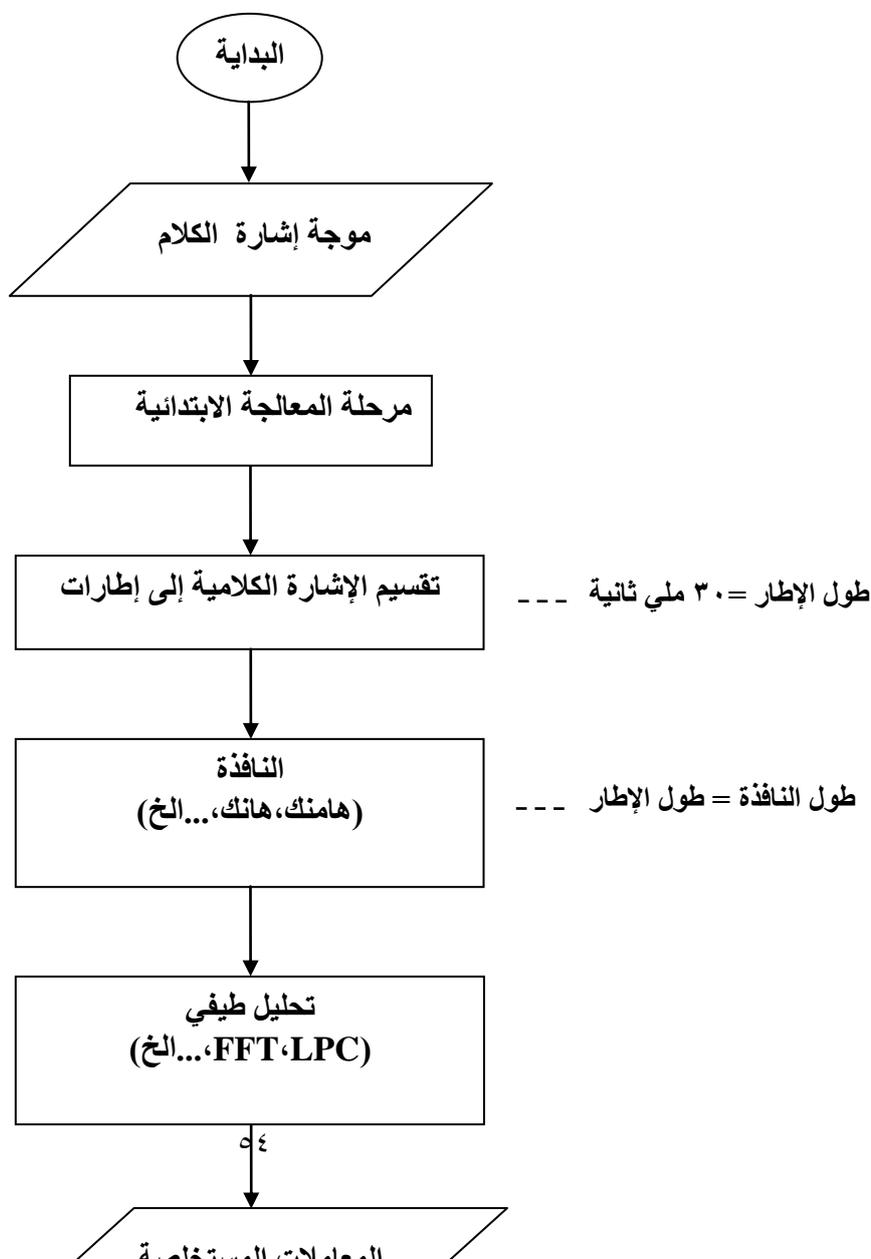
$$Aut_{apt}(0) = 1.0001 * Aut(0)$$

$$Aut_{apt}(k) = W_{lag}(k) * Aut(k) \quad k=1, \dots, 10 \quad (6-$$

٤)

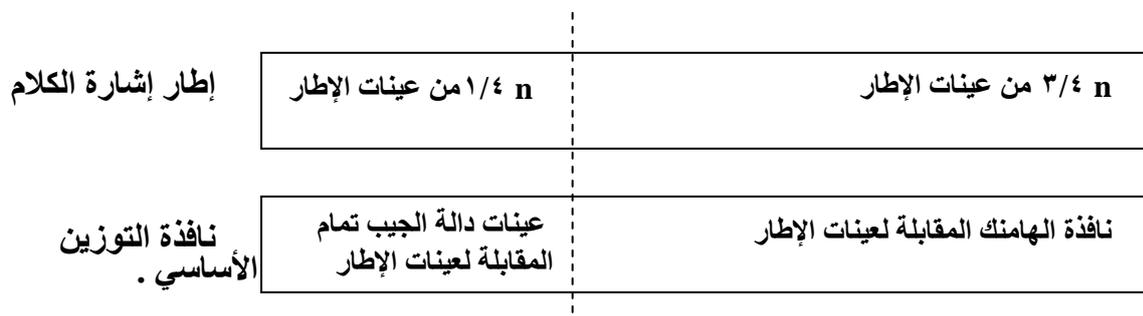
٥. التطبيع لقيم معاملات الارتباط الناتجة بتقسيمها على أول معامل من معاملات الارتباط الذاتي وبالإستناد إلى المعادلة رقم (٣-١٧) وبحسب ما موضح في المخطط الإنسيابي رقم (٦.٤) حيث أن $(Auto_k)$ تشير إلى معاملات الارتباط الذاتي بعد التطبيع والتي أنجزت للتخلص من القيم العالية جداً مما يسهل من عملية الحساب كثيراً خاصة أن قيم معاملات الارتباط الذاتي أساسية للاستمرار في عملية التحليل .

٦. الحصول على معاملات التنبؤ الخطي بتطبيق خوارزمية ليفنسن داربن التكرارية على المعاملات الناتجة في الخطوة (٥) وباستخدام المعادلات من (٣-١٠) إلى (٣-١٦) وكما موضح في الشكل (٧.٤) .

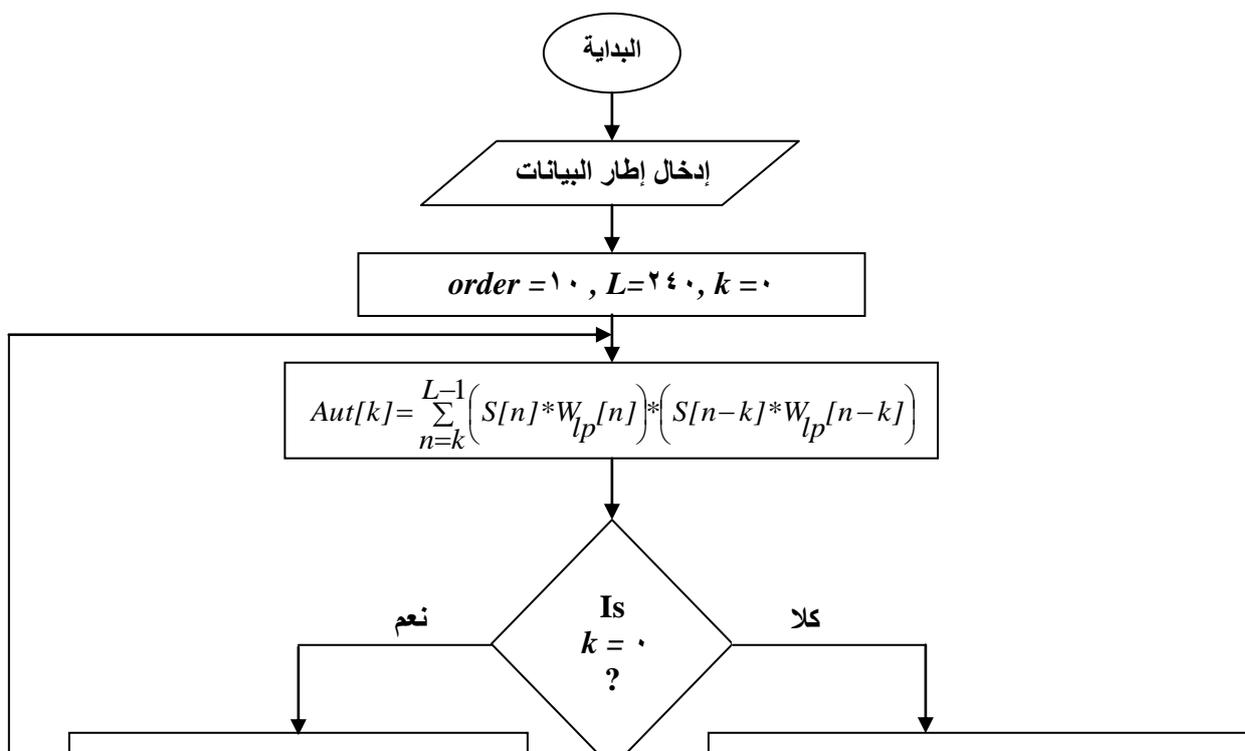
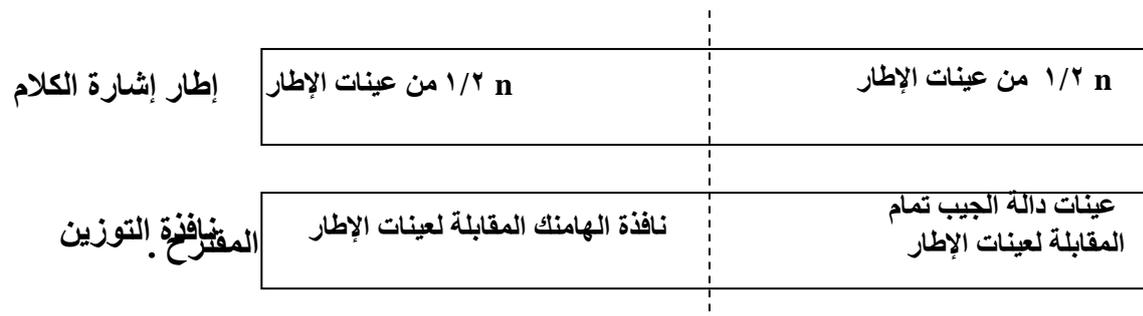


شكل (٣.٤) المخطط الانسيابي لإجراء تحليل إشارة الكلام النموذجي .

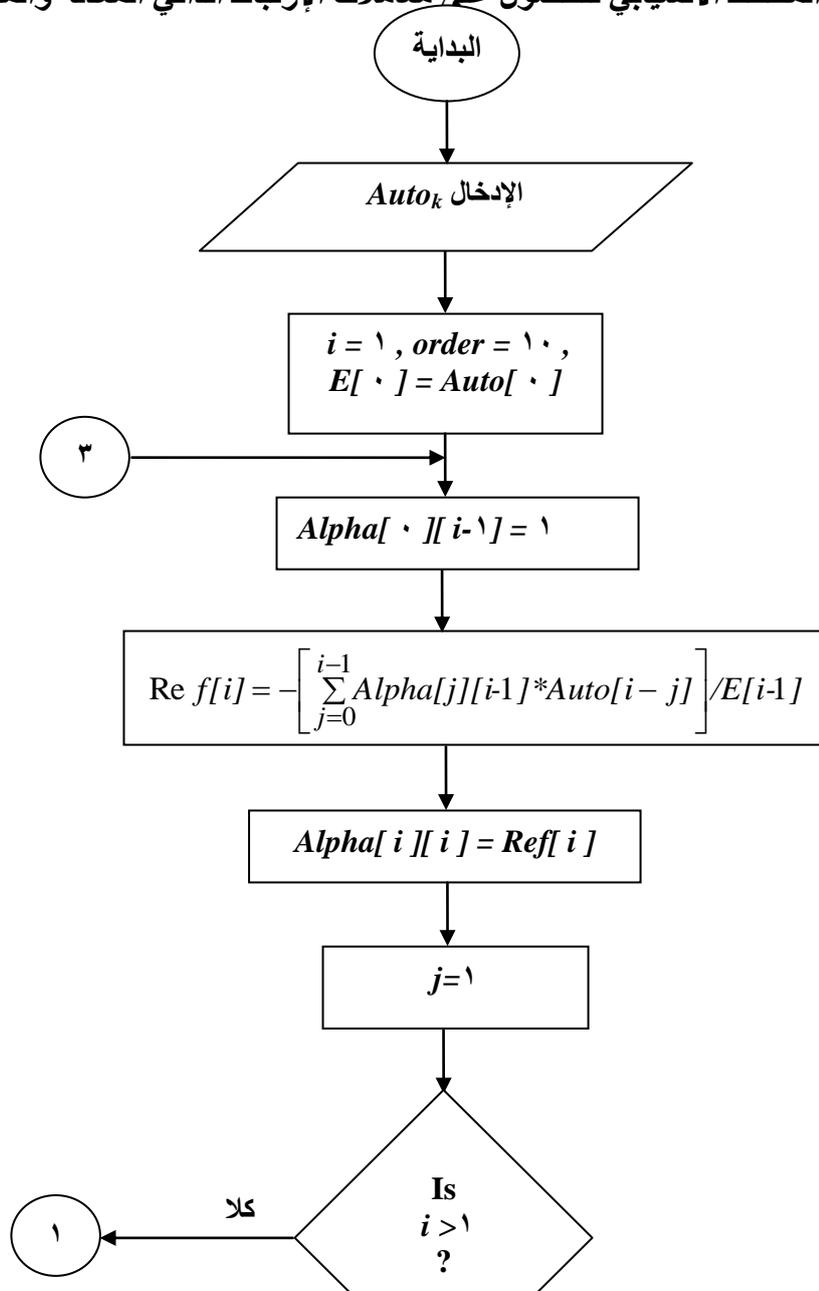
الحد الفاصل بين الدالة الأولى والدالة الثانية

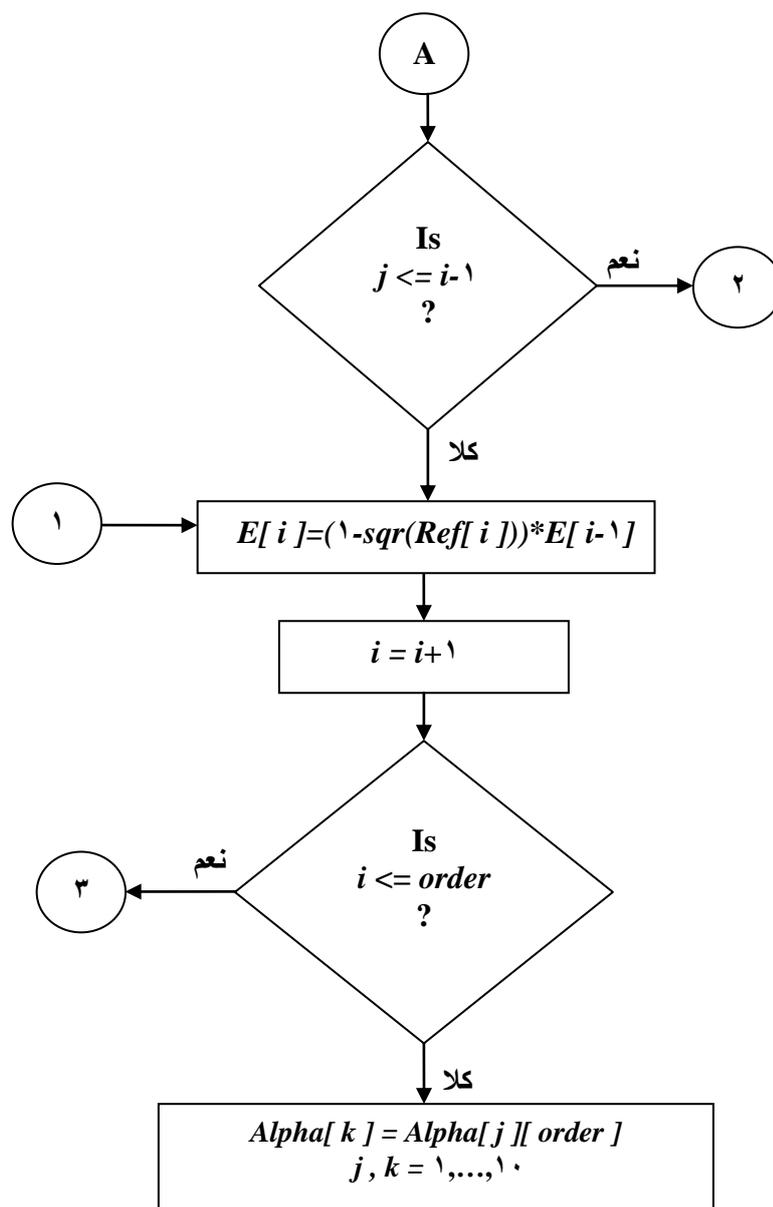


الحد الفاصل بين الدالة الأولى والدالة الثانية



شكل (٦.٤) المخطط الانسيابي للحصول على معاملات الارتباط الذاتي المعدلة والمتطبعة.

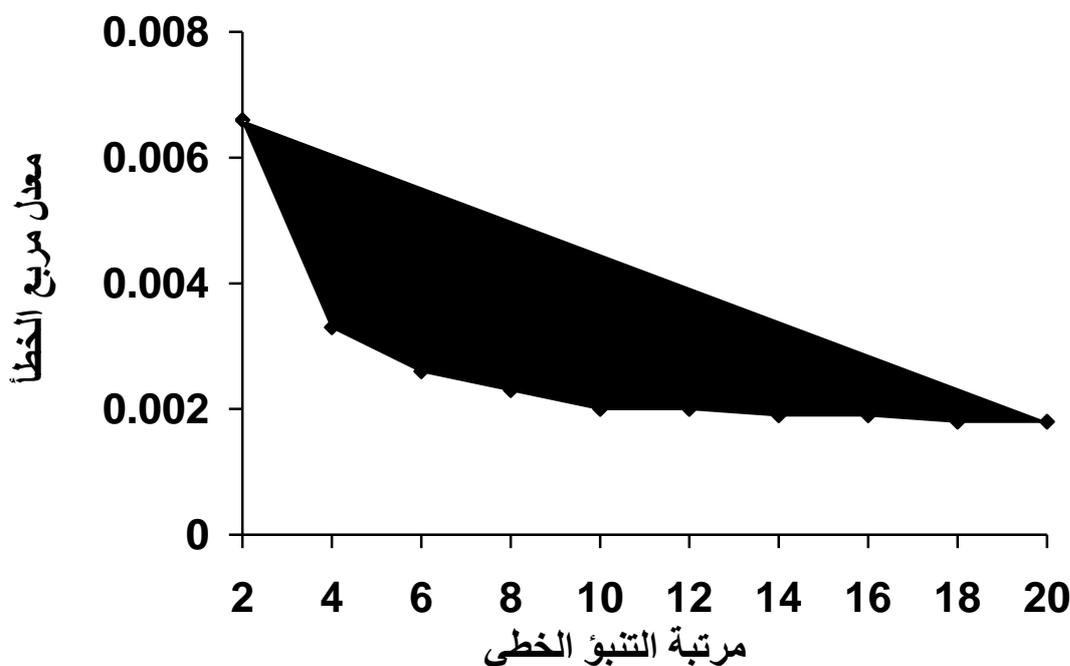




شكل (٧.٤) المخطط الانسيابي لخوارزمية ليفنسن داربن التكرارية .

الإخراج $Alpha_k$

تم إجراء اختبار لغرض اختيار عدد من معاملات التنبؤ الخطي (الرتبة ($order$)) حيث يبين المنحني في الشكل (٨.٤) العلاقة بين معدل مربع الخطأ (Mean Square Error) وعدد معاملات التنبؤ الخطي ($order$) إذ يتبين **النهاية** مع الخطأ سوف يثبت تقريباً بحدود (١٠ فما فوق) وعلى هذا الأساس فقد تم اختيار $order=$ **النهاية** (إن اختيار $order$ قليلة تنتج معاملات أقل وحسابات أقل ولكن بدقة قليلة وباختيار $order$ كبيرة ينتج معاملات أكثر وحسابات أكثر ولكن بدقة أكثر ومن هنا يتم التنسيق بين عدد المعاملات والدقة باختيار الرتبة $order=10$).



شكل (٨.٤) معدل مربع الخطأ بطريقة ترميز التنبؤ الخطي بمراتب مختلفة

٤.٢.٤ مرحلة بناء كتاب الرموز (Codebook Construction Stage)

تعد هذه المرحلة من أهم مراحل ترميز إشارة الكلام بطريقه كفؤة للحصول على جودة إشارة كلامية عالية بدون زيادة كثيرة في معدل البيانات المنقولة ، إذ يُصمم كتاب الرموز من قبل مصمم النظام.

تتضمن إشارة الإثارة في النظام المقترح متجه ترميز متكيف مقاس و متجه ترميز ثابت مقاس لذا فقد تم بناء كتابي رموز وهما كتاب الرموز المتكيف وكتاب الرموز الثابت وقد تم استخدام كتاب الرموز الجبري لإثارة كتاب الرموز الثابت .

وجد من الممكن إجراء تحسين مهم في كفاءة المرز بحساب قيم مثالية ومحسنة لبرحي كتابي الرموز ، وهينة كتاب الرموز وبسبب طبيعة التكيف الراجع للمرز – فاك الترميز كان من الضروري استعمال التدريب على كتاب رموز مغلق الحلقة .

١. كتاب الرموز المتكيف (Adaptive Codebook)

هو جدول للإثارة السابقة $u(n)-\alpha$ والذي يبني من قبل المصمم خلال بحث الحلقة المغلقة للتحليل بالتركيب حيث يكون فارغاً في البدء ثم يمتلئ خلال التشغيل مع النسخ من الإثارة السابقة المتأخرة بإطار جزئي واحد .

يتم بحث كتاب الرموز المتكيف لكل إطار جزئي ذي مدة زمنية (7.5 ms) ، يحسب الربح (G_1) لمتجه كتاب الرموز المتكيف بالاعتماد على المعادلة (3-35) وبصورة تؤدي إلى تقليل الخطأ الموزون بين إشارة الكلام الأصلي والمعاد بناؤه .
يعبر عن متجه كتاب الرموز المتكيف كما في المعادلة (3-18) .

٢. كتاب الرموز الثابت (Fixed Codebook)

يستخدم النظام المقترح كتاب رموز موقع العينة لغرض إثارة كتاب الرموز الثابت وتتضمن كتب رموز موقع العينة من مواقع البيانات ، ففي الطريقة المقترحة تتغير المواقع من إطار جزئي إلى إطار جزئي بالاعتماد على إشارة الكلام المدخلة ويحتوي كل متجه كتاب رموز ثابت على أربعة قنوات وتتضمن مواقعها كل العينات في الإطار الجزئي ، حيث ينجز البحث في أربعة دورات متشابهة طبقاً إلى كل مواقع العينة ، وطالما إن مواقع العينة مهمة لجودة إشارة الكلام لذا يجب تقيداً بطريقة كفاءة والتي يتم بها تقيد مواقع العينة إلى مواقع العينات الزوجية كما مبين في الجدول (4-1).

جدول (4-1) : هيكل متجه كتاب الرموز الثابت

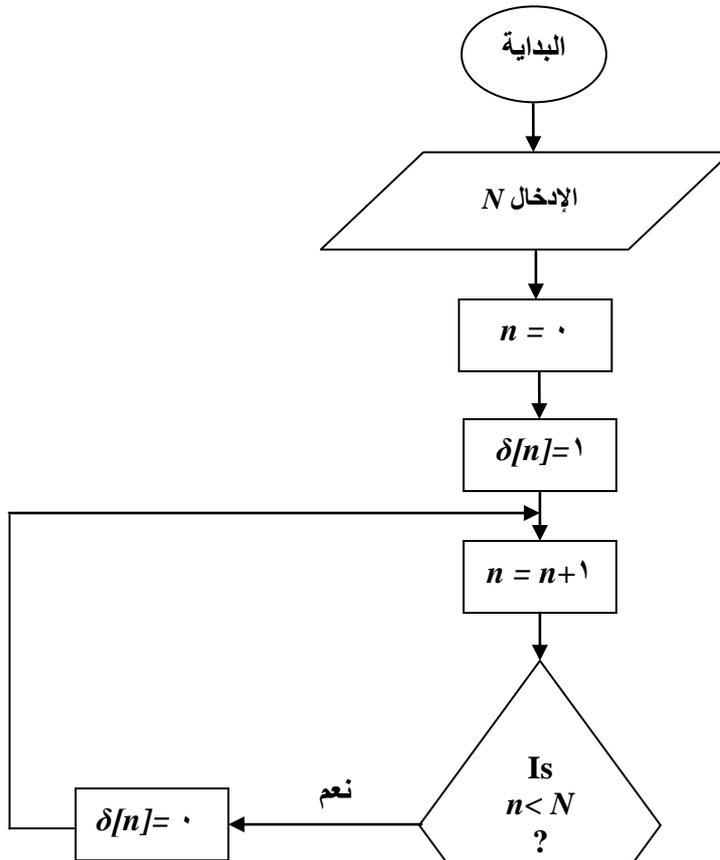
Channel Number	Positions
٠	٠, ٨, ١٦, ٢٤, ٣٢, ٤٠, ٤٨,
١	٢, ١٠, ١٨, ٢٦, ٣٤, ٤٢, ٥٠,
٢	٤, ١٢, ٢٠, ٢٨, ٣٦, ٤٤, ٥٢, (٦٠)
٣	٦, ١٤, ٢٢, ٣٠, ٣٨, ٤٦, ٥٤, (٦٢)

حيث يبين الجدول (4-1) كيفية توزيع مواقع العينة لأربعة قنوات لمتجه كتاب الرموز الثابت. ويتم بحث كتاب الرموز الثابت لكل إطار جزئي ذي مدة زمنية (7.5 ms) ، يحسب الربح (G_2) لمتجه كتاب الرموز الثابت بالاعتماد على المعادلة (3-36) وبصورة تؤدي إلى تقليل الخطأ الموزون بين إشارة الكلام الأصلي والمعاد بناؤه .
يعبر عن متجه الترميز الثابت كما في المعادلة (3-19) .

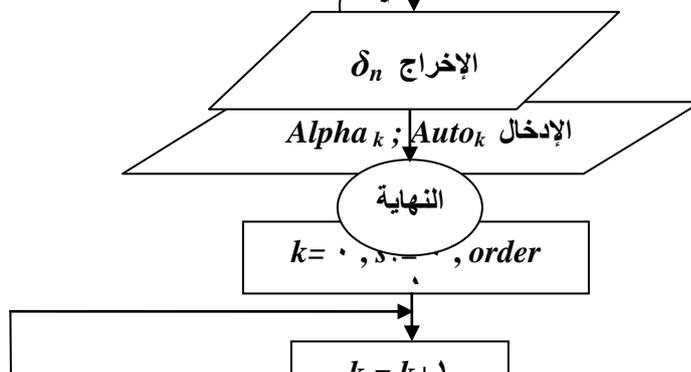
٥.٢.٤ احتساب استجابة النبضة

(Computation of The Impulse Response)

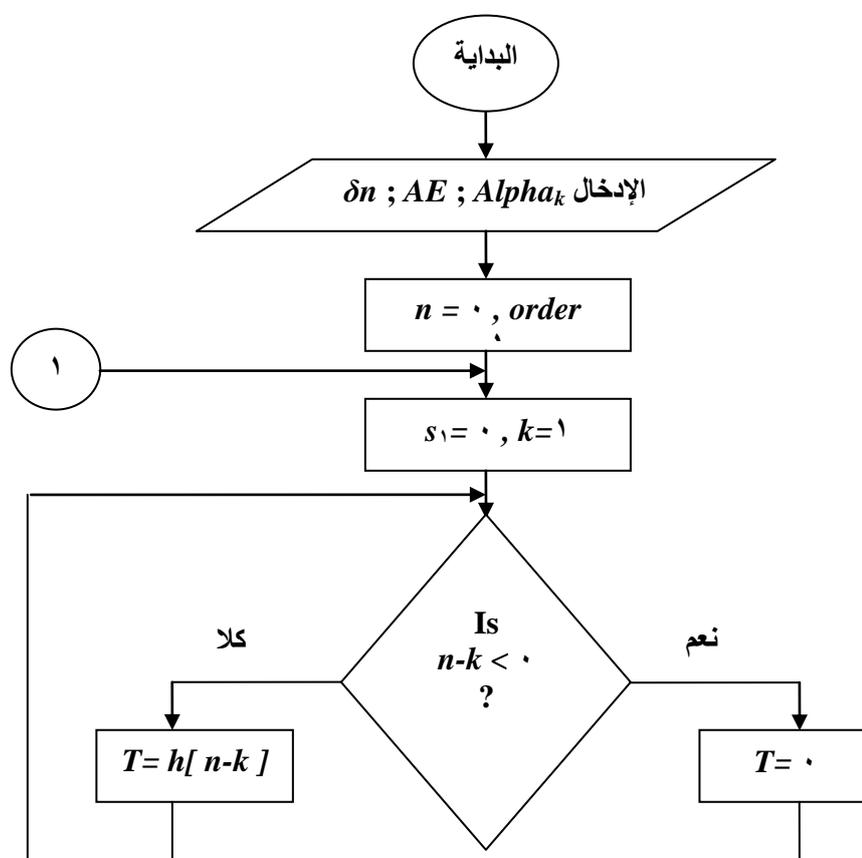
١. تحسب استجابة النبضة $h(n)$ لكل إطار جزئي (N) وحسب الخطوات الآتية :-
 حساب وحدة العينة أو وحدة النبضة $\delta(n)$ وبحسب ما موضح في المخطط الانسيابي رقم (٩.٤).
٢. حساب سعة الإثارة AE وبلاستناد إلى المعادلة رقم (٣-٢٦) ، وكما موضح بالمخطط الانسيابي رقم (١٠.٤).
٣. وبالاتماد على الخطوتين السابقتين تحسب استجابة النبضة $h(n)$ وبالاتماد على المعادلة (٣-٢٥) ، وكما موضح بالمخطط الانسيابي رقم (١١.٤).

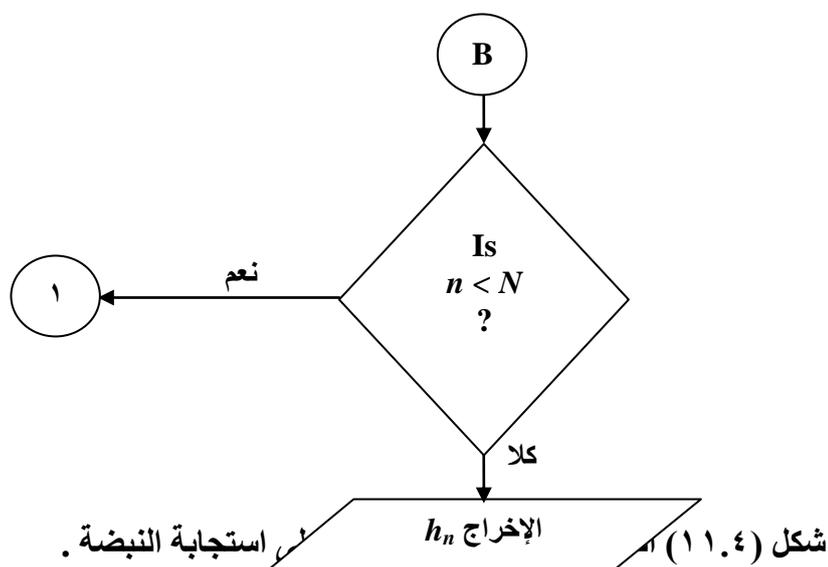


شكل (٩.٤) المخطط انسيابي للحكم على وحدة العينة أو وحدة النبضة.



شكل (١٠.٤) المخطط الانسيابي للحصول على سعة الإثارة .



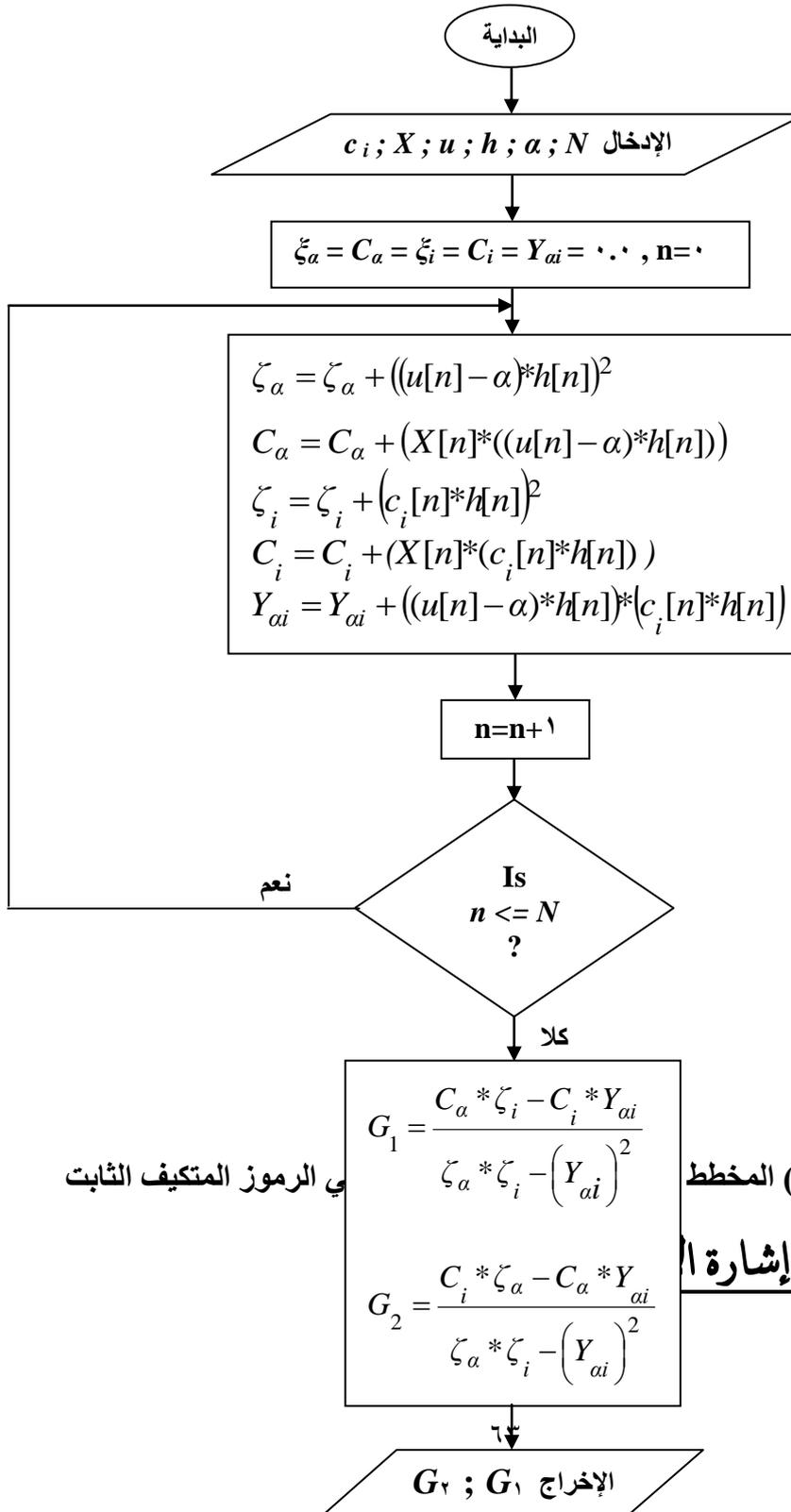


٦.٢.٤ احتساب إشارة الهدف (Computation of The Target)

تحتسب إشارة الهدف $X(n)$ لبحث كتابي الرموز المتكيف والثابت عادةً لكل إطار جزئي (N) بطرح استجابة الإدخال الصفرية لمرشح التركيب الموزون من إشارة الكلام الموزون $S_w(n)$ وبالاعتماد على المعادلة (٣-٣٤)

٧.٢.٤ حساب الربح (Computation of The Gain)

يحسب الربح لمتجهي كتابي الرموز المتكيف و الثابت كما موضح بالمخطط الانسيابي رقم (١٢.٤):



ي الرموز المتكيف الثابت

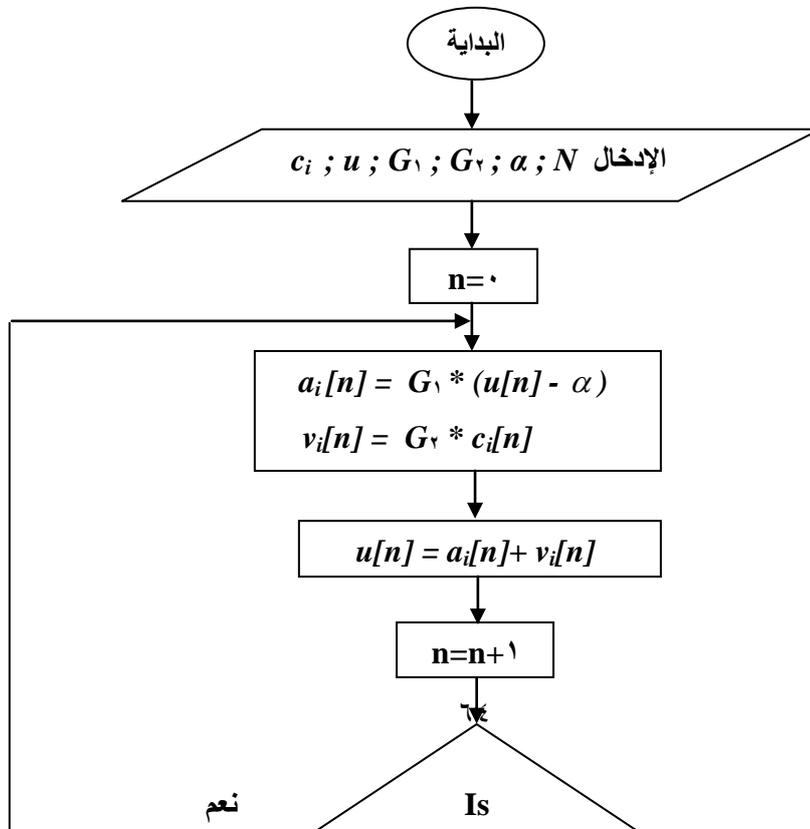
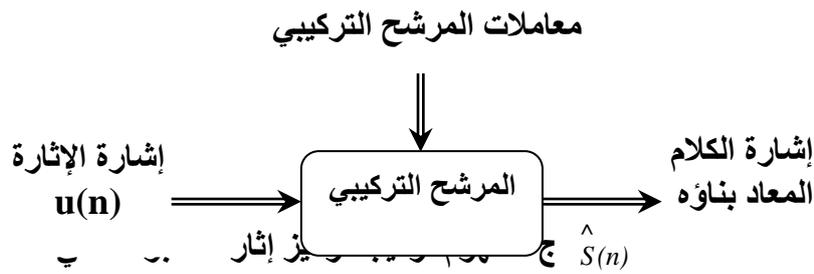
شكل (١٢.٤) المخطط

٨.٢.٤ احتساب إشارة

(Computation of The Excitation Signal)

الإثارة النهائية $u(n)$ هي مجموع إشارة كتاب الرموز المتكيف المقيسة مع الإشارة المقيسة من كتاب الرموز الثابت.

تستعمل الإثارة $u(n)$ لقيادة المرشح التركيبي لنمذجة التأثيرات للقناة الصوتية ، وعند فاك الترميز ترشح إشارة الإثارة خلال المرشح التركيبي للتنبؤ الخطي لإنتاج إشارة الكلام المعاد بناؤه $\hat{S}(n)$ والتي تكون مشابهة قدر الإمكان لإشارة الكلام الأصلية كما مبين في الشكل (١٣.٤) ، ويوضح المخطط الانسيابي (١٤.٤) احتساب إشارة الإثارة .

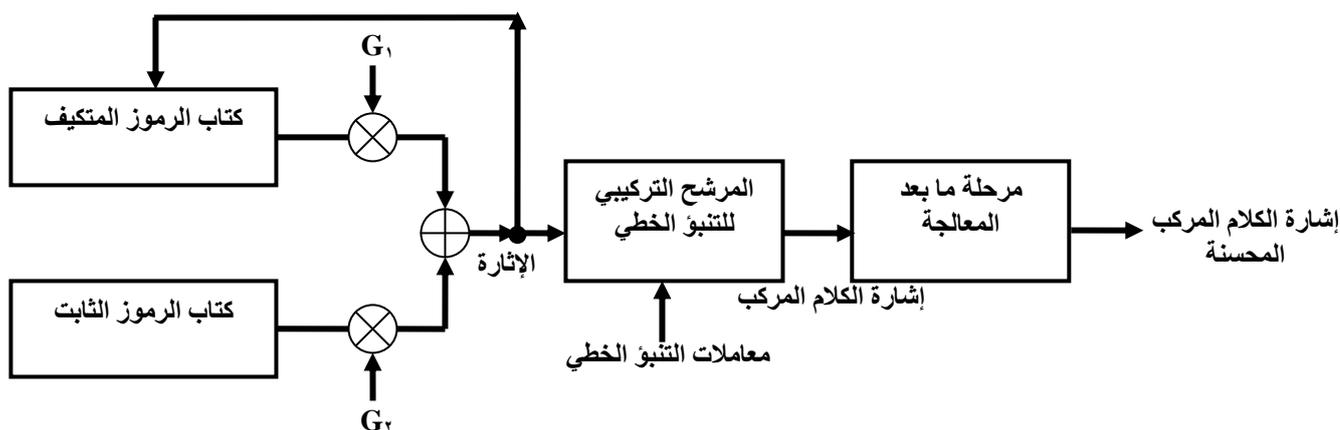


شكل (١٤.٤) المخطط الانسيابي لاحتساب إشارة الإثارة

٣.٤ فك الترميز (Decoder)

يبين الشكل (١٥.٤) مبادئ فك الترميز. بعد الحصول على معالم المرز وهي معاملات مرشح التنبؤ الخطي ومتجهي كتاب الرموز الثابت والمتكيف ومجموعتين من الريح لكتابي الرموز الثابت والمتكيف فإن هذه المعالم تستخدم لاحتساب إشارة الكلام المعاد بناؤه كما سيوصف في الجزء (١.٣.٤) وتحسن هذه الإشارة المعاد بناؤها بعمليات ما بعد المعالجة وكما سوف يُبين في الجزء (٣.٣.٤).
في مرحلة فك الترميز يتم إجراء الخطوات الآتية :-

١. يعاد تركيب الإثارة بإضافة متجهات كتاب الرموز الثابت والمتكيف المقاسة بأرباحها.
٢. يعاد تركيب إشارة الكلام بترشيح الإثارة خلال المرشح التركيبي للتنبؤ الخطي .
٣. تمرر إشارة الكلام المعاد بناؤه خلال مرحلة ما بعد المعالجة .

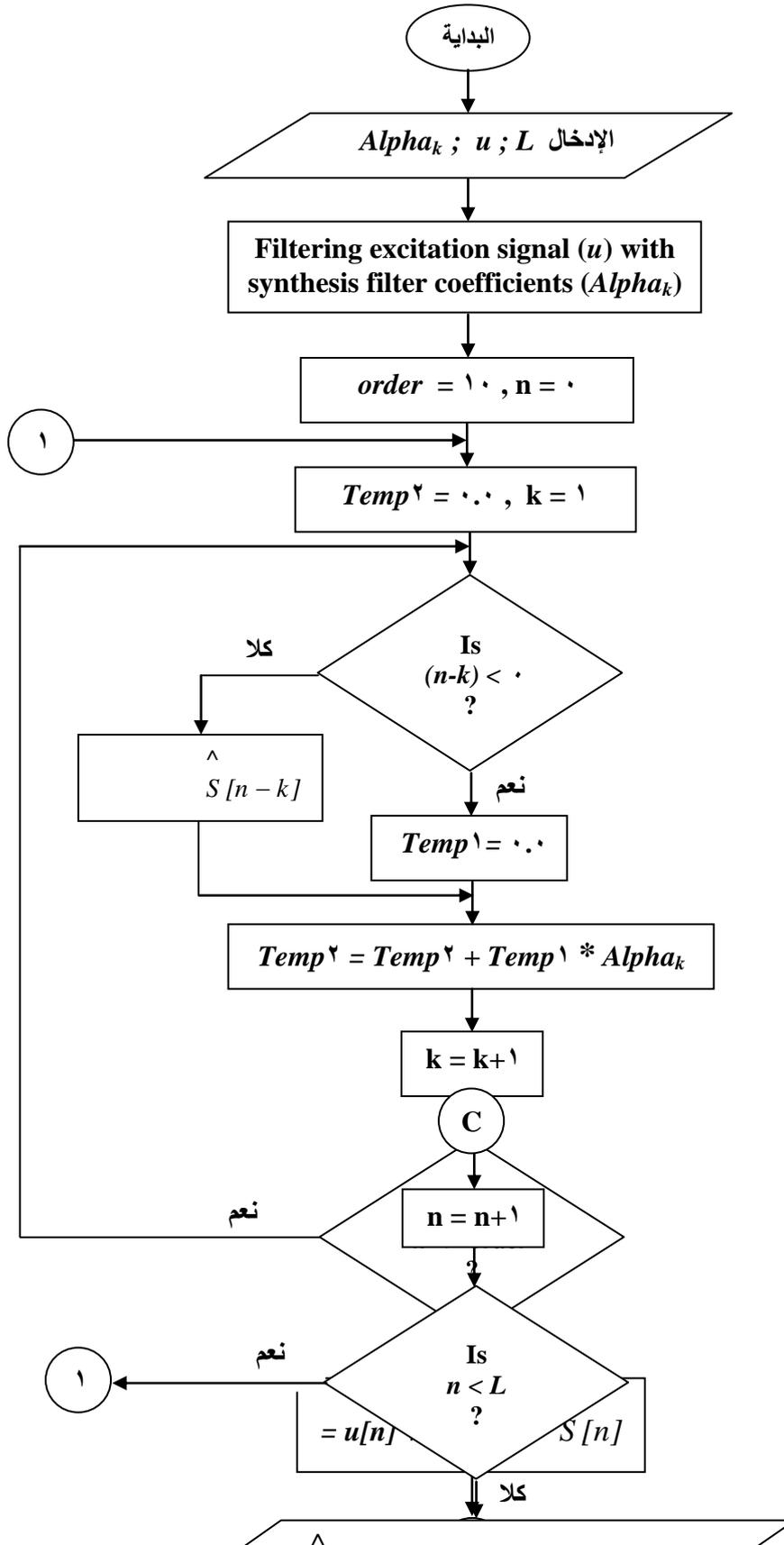


شكل (١٥.٤) فك الترميز لمرز إشارة التنبؤ الخطي (CELP)

١.٣.٤ احتساب إشارة الكلام المعاد بناؤها

(Computation of The Reconstructed Speech Signal)

تدخل إشارة الإثارة $u(n)$ الناتجة من المخطط الانسيابي رقم (١٤.٤) ومعاملات التنبؤ الخطي إلى المرشح التركيبي للتنبؤ للحصول على إشارة الكلام المعاد بناؤه وكما موضح في المخطط الانسيابي رقم (١٦.٤) ، بعد ذلك تعالج الإشارة المعاد بناؤها $\hat{S}(n)$ بواسطة مرحلة ما بعد المعالجة الموصوفة بالجزء (٣.٣.٤) لتحسين جودة إشارة الكلام المعاد بناؤه.



شكل (١٦.٤) المخطط الانسيابي لاحتساب إشارة الكلام المعاد بناؤه

٢.٣.٤ حساب وتوزيع واختزال الخطأ

يحسب الخطأ بين إشارة الكلام الأصلية والمسترجعة لكل إطار ومن ثم يوزن بمرحلة التوزيع بالاعتماد على المعادلة (٣-٢٣) ، وتستخدم كتلة توزيع الخطأ لتشكيل طيف إشارة الخطأ وذلك لتقليل الإرتفاع غير الموضوعي لهذا الخطأ وهذا ممكن لأن إشارة الخطأ في مناطق التردد التي تمتلك فيها الإشارة الكلامية طاقة عالية ستكون مقنعة جزئياً بهذه الإشارة الكلامية . يقوي مرشح التوزيع الضوضاء في مناطق التردد ، حيث تكون مكونات الإشارة الكلامية واطئة . لذا فإن إختزال الخطأ الموزون سيعمل على تقليل الارتفاع غير الموضوعي لإشارة الخطأ . لقد وجد التوزيع والإختزال لإنتاج تحسين مفيد في جودة إشارة الكلام المعاد بناؤه لرمزات - فاك ترميز التحليل بالتركيب (AbS Codex) .

٣.٣.٤ مرحلة ما بعد المعالجة (Post-Processing)

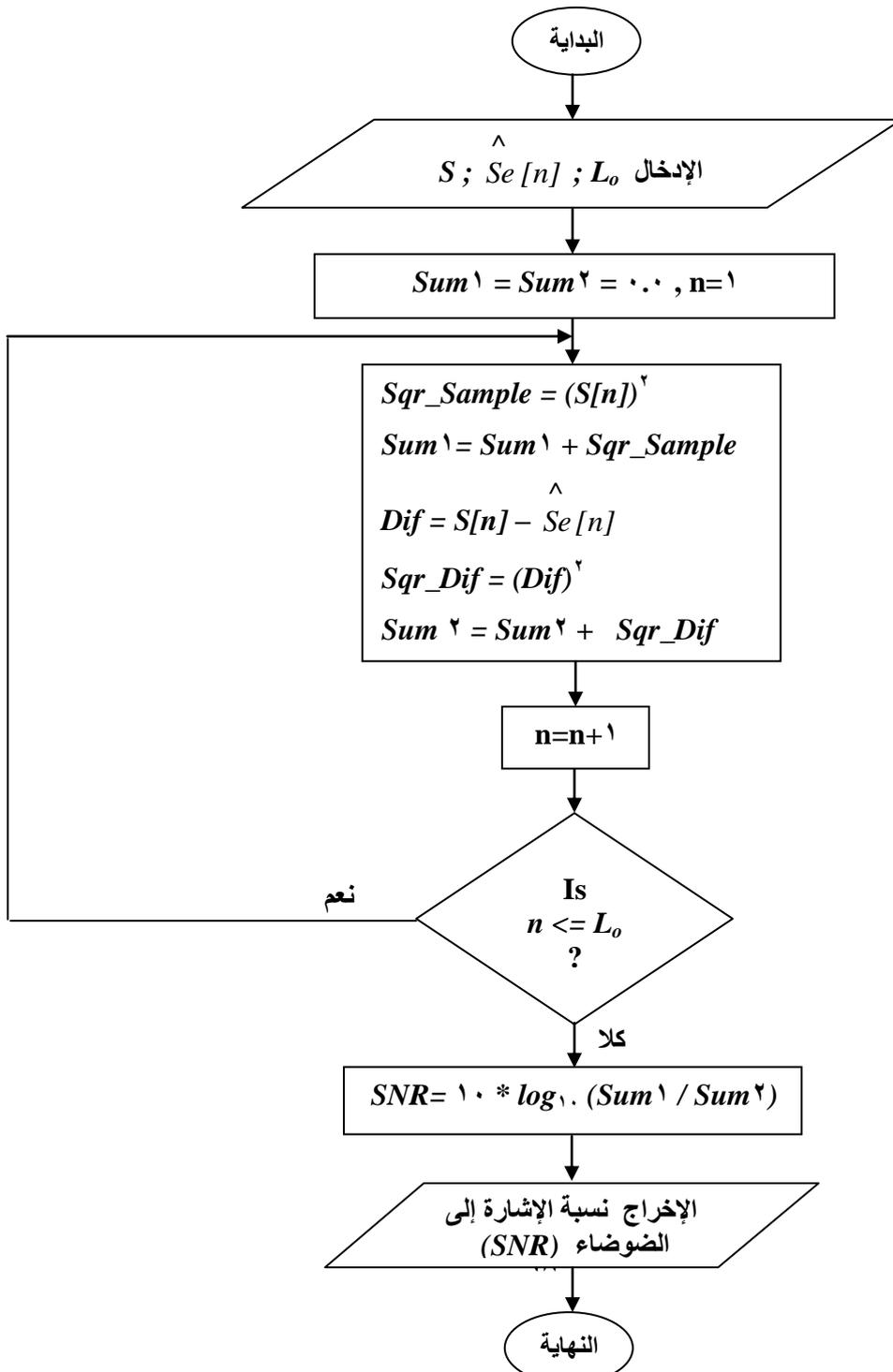
تطبق هذه المرحلة من المرز - فاك الترميز (Codec) على فاك الترميز فقط لتحسين جودة إشارة الكلام المعاد بناؤه للحصول على جودة إشارة كلام عالية مشابهة لجودة إشارة الكلام الأصلية .

تقسم مرحلة ما بعد المعالجة إلى جزئين أساسيين استعملا حسب الترتيب التالي :-

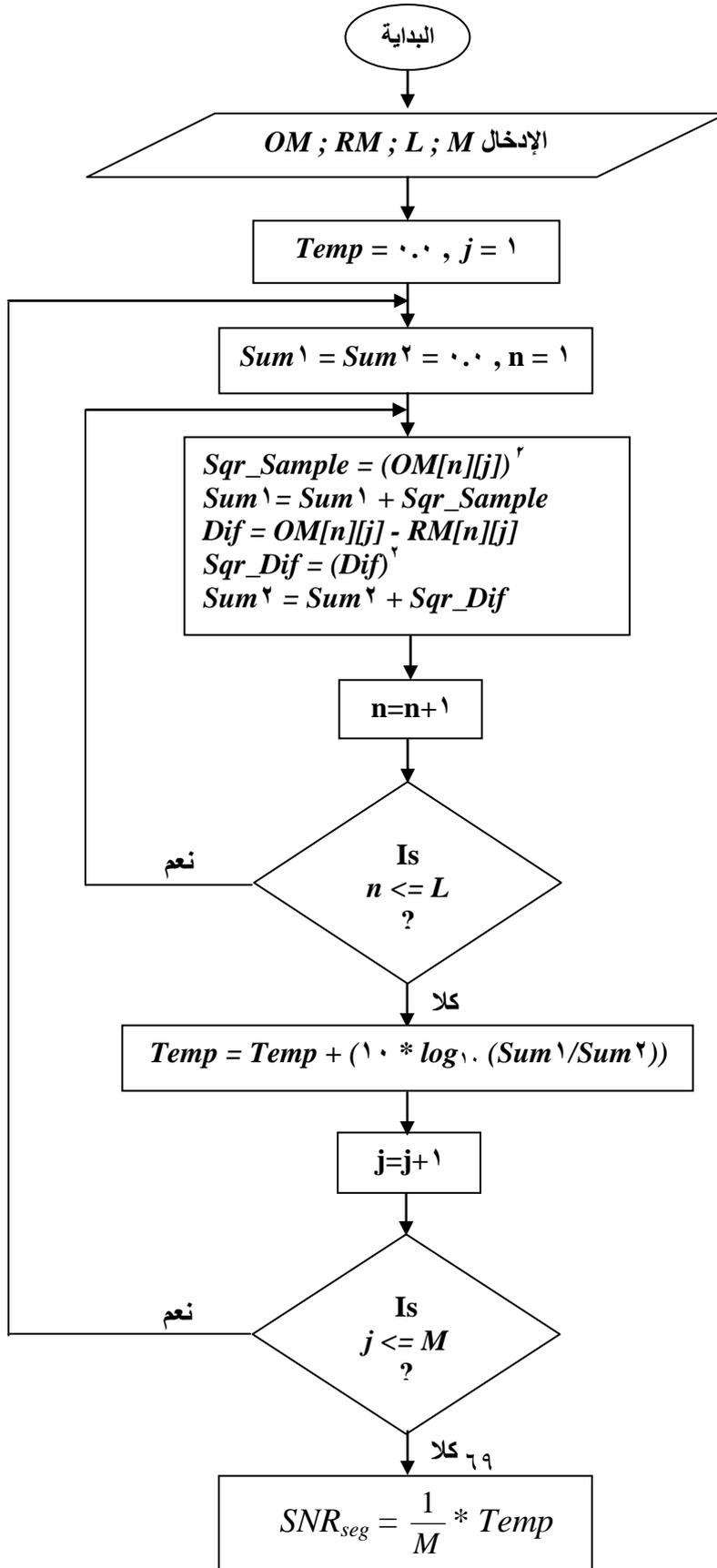
١. بعد الحصول على إشارة الكلام المعاد بناؤه تحسن جودته بزيادة طول إشارة الكلام.
٢. استخدام مرشح الإمرار الواطئ لاستجابة النبضة المتناهية (FIR) لتحسين جودة الصوت الناتج من عملية التركيب.

٤.٤ التقييم (Evaluation)

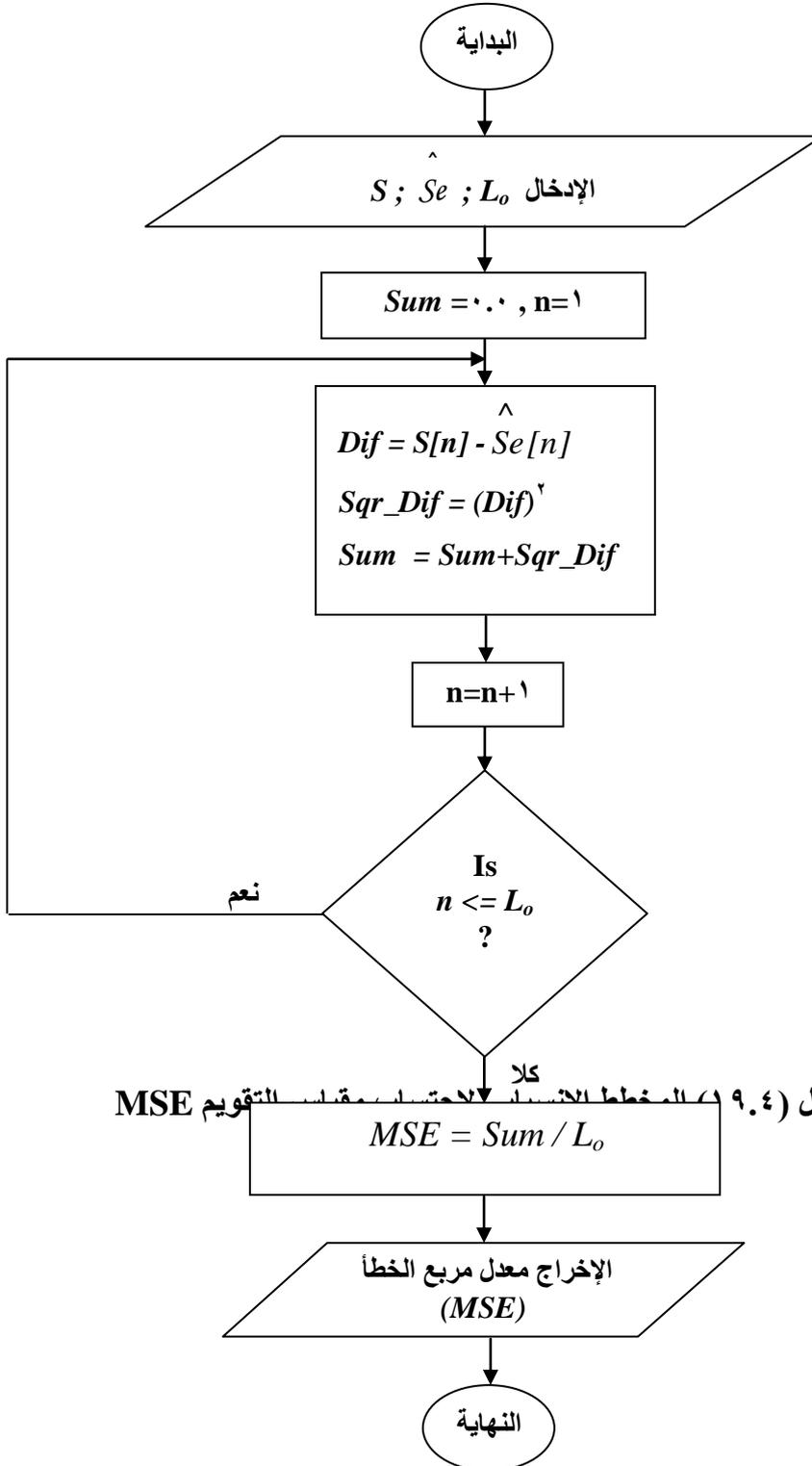
تطبق هذه المرحلة بعد مرحلتي الترميز وفك الترميز لتقويم كفاءة النظام المقترح وباستخدام المقاييس الرياضية (SNR, SNR_{seg}, MSE) على التوالي وكما موضح بالمخططات الانسيابية [(١٧.٤), (١٨.٤), (١٩.٤)] الآتية:



شكل (١٧.٤) المخطط الانسيابي لاحتساب مقياس التقويم SNR



شكل (١٨.٤) المخطط الانسيابي
لاحتساب مقياس التقويم SNR_{seg}



شكل (١٩.٤) المخطط الانسيابي للاحتساب مقياس التقويم MSE

الفصل الخامس

النتائج، المناقشة والاستنتاجات، والأعمال المستقبلية

١.٥ النتائج

اختبار أداء النظام المقترح , فقد تم تقويمه من خلال استخدام إشارات الكلام الأصلية (المدخلة) وإشارات الكلام المعاد بناؤها (المسترجعة) ولعدد من المتكلمين (رجالاً ، ونساءً) عند ترددي نمذجة (٨٠٠٠ Hz) و(٤٤١٠٠ Hz) كأتملة تجريبية لبيان سلوكية كفاءة النظام المقترح ، وقد تمت المقارنة بالاعتماد على نوعين من المقاييس وهما المقاييس الرياضية والمقاييس الحسية .

١.١.٥ المقاييس الرياضية (Objective Measures)

يمكن استعمال العديد من المقاييس الرياضية لتقويم أداء مرمز إشارة الكلام المعتمد على التحليل بالتركيب (AbS) لجودة الإشارة المعاد بناؤها بعد فك الترميز . حيث تم حساب المعايير الآتية :

- ◀ نسبة الإشارة إلى الضوضاء (Signal to Noise Ratio SNR) بين إشارة الكلام الأصلية والنتيجة (المسترجعة) .
- ◀ نسبة الإشارة إلى الضوضاء المجرأة (Segmental Signal to Noise Ratio) بين إشارة الكلام الأصلية والنتيجة (المسترجعة) .
- ◀ معدل مربع الخطأ (Mean Square Error MSE) بين إشارة الكلام الأصلية والنتيجة (المسترجعة) .

٢.١.٥ المقاييس الحسية (Subjective Measures)

هنالك طرائق مختلفة للاختبارات الحسية , إلا أنه سيكون الاعتماد على علامة معدل الرأي (Mean Opinion Score (MOS)) فقط لكونها الأفضل في مجال تطبيق ترميز إشارة الكلام

ولأن الاختبارات الأخرى مشابهة لعلامة معدل الرأي ولا تقدم أية فائدة إضافية . علماً إنه قد تم استخدام الأسلوب الآتي لحساب المقياس الحسي (MOS) :

- ◀ السماح لعشرين (٢٠) شخص بالاستماع إلى إشارة الاختبار .
- ◀ يُقوم كل مستمع إشارة الاختبار عن طريق إعطاء درجة لجودة ومفهومية إشارة الاختبار ، حيث تتراوح هذه الدرجة بين الدرجة (١) المقابلة لمستوى سيء - إلى الدرجة (٥) المقابلة لمستوى ممتاز .
- ◀ تقويم إشارة الاختبار بأخذ معدل تقاويم هؤلاء المستمعين والتي على أساسها يتم الحكم على المرمز بالرجوع إلى المستويات الخمسة المتفق عليها ، وهي :
(Excellent) ممتاز ، (Good) جيد ، وسط (Fair) ، ضعيف (Poor) ، سيء (Bad) حيث تعطى هذه المستويات قيم رقمية من ٥ إلى ١ .

وقد تم تطبيق المقياس الحسي (MOS) على الحالات الدراسية الأربع التالية وسنورد نتائج المقياس الحسي (MOS) للحالة الدراسية الأولى فقط لكونها الأفضل .

٣.١.٥ الحالات الدراسية

١.٣.١.٥ الحالة الدراسية الأولى

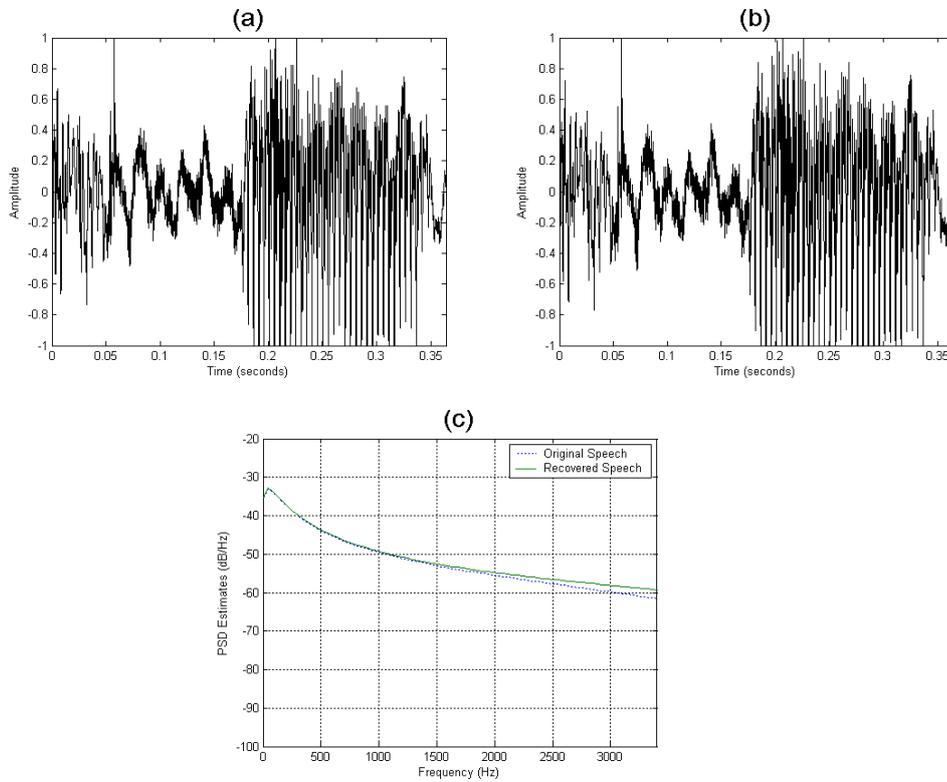
تم تنفيذ النظام المقترح بالاعتماد على المفهوم الذي يُعدّ النظام عبارة عن حلقة مغلقة (Closed Loop) . في هذا المفهوم تتضمن عملية تحليل الإشارة الكلامية ، تحليل التنبؤ الخطي عبر إطار إشارة الكلام الكلي بعد تقطيع موجة إشارة الكلام الداخلة إلى إطارات ، وذلك لاستخلاص معاملات التنبؤ الخطي لكل إطار ويعاد تركيب الإشارة الكلامية لكل إطار وباستخدام مرشح التحسين (FIR Filter) بعد تركيب موجة الإشارة الكلامية المسترجعة لتحسينها .

وتوضح الأشكال (١.٥) و(٢.٥) الإشارة الكلامية الأصلية (Original Speech) ، والمسترجعة (Recovered Speech) ، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٤٤١٠٠ Hz) .

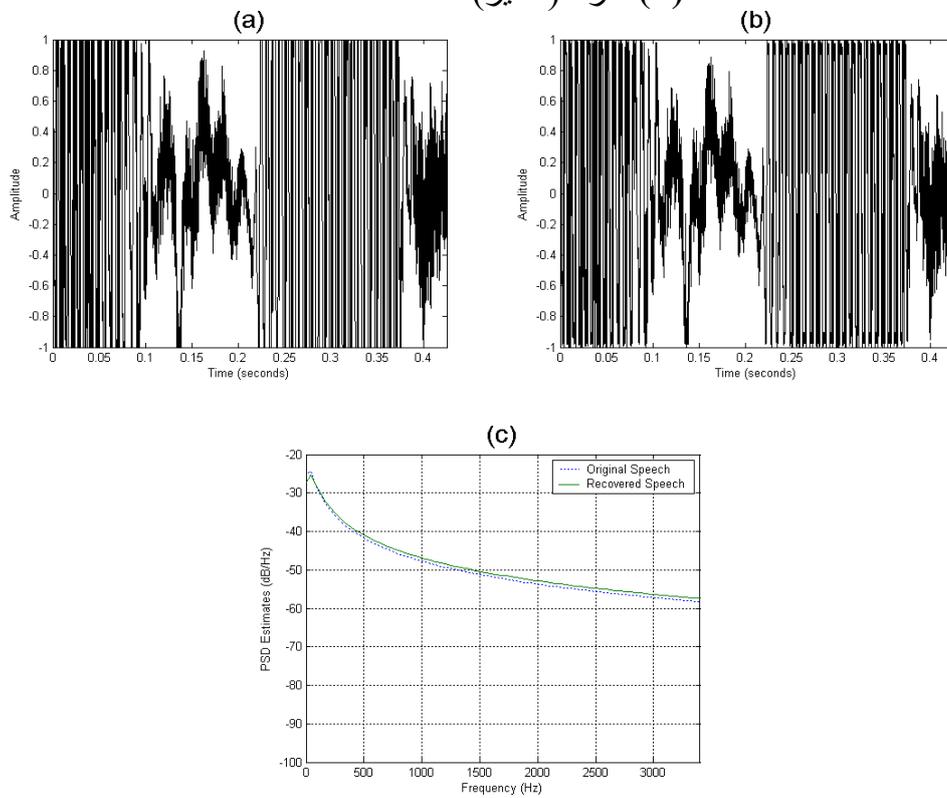
علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين .

ملاحظة : إن الأشكال من (١-٦) ستكون العناوين الخاصة بأجزائها الثلاثة كالتالي:

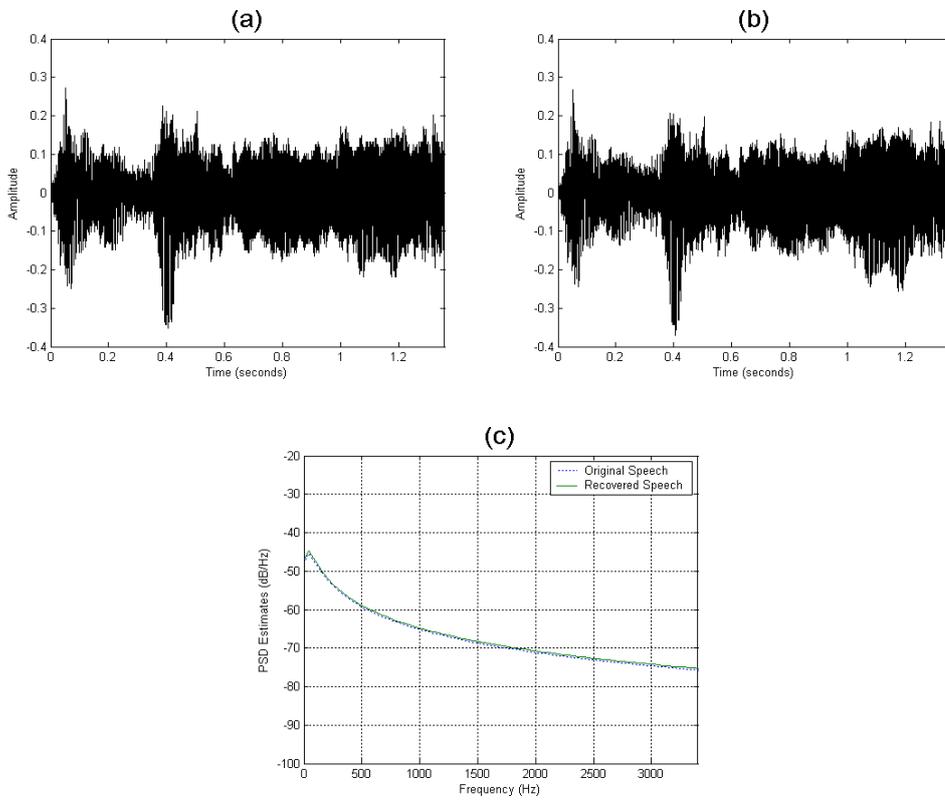
- a. الإشارة الكلامية الأصلية .
- b. الإشارة الكلامية المسترجعة .
- c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة .



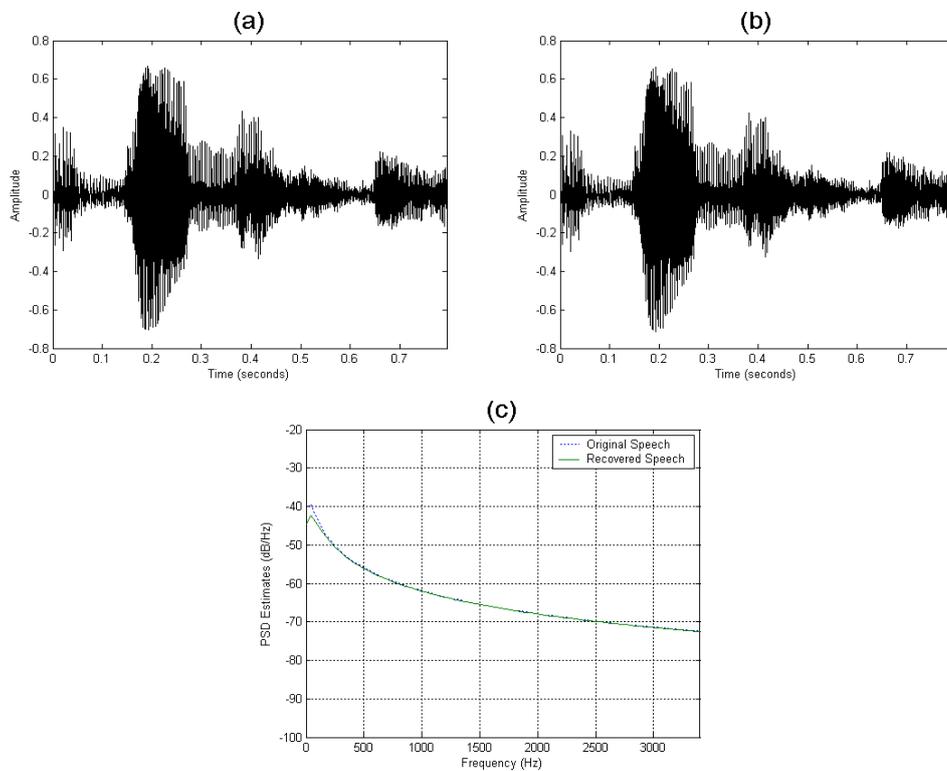
(١) امرأة (الخير)



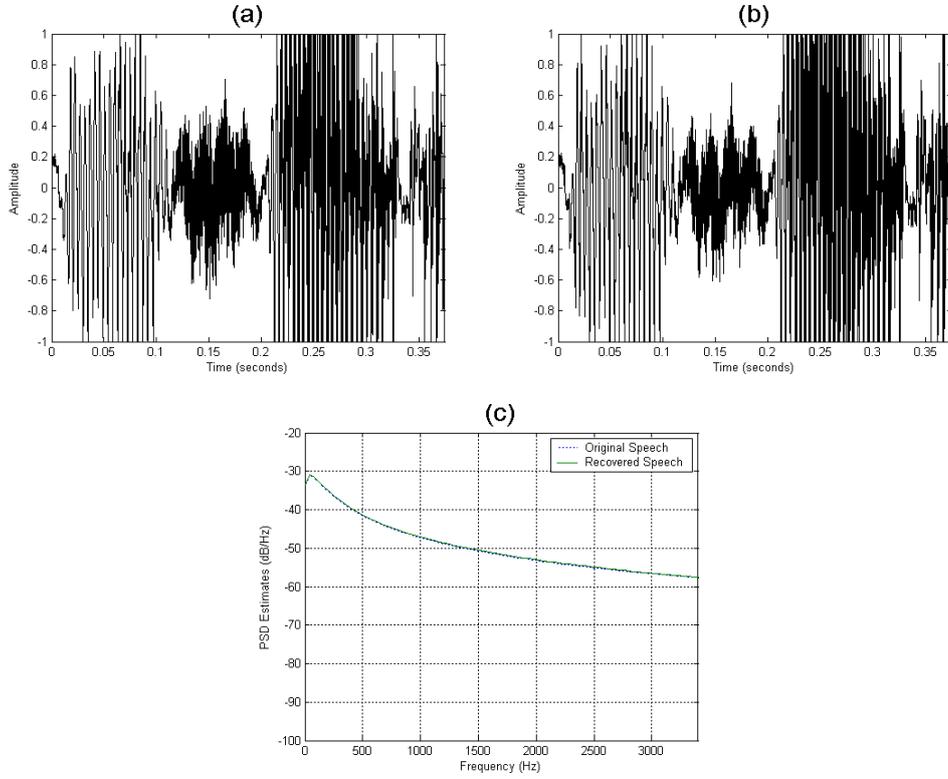
(٢) رجل (الخير)



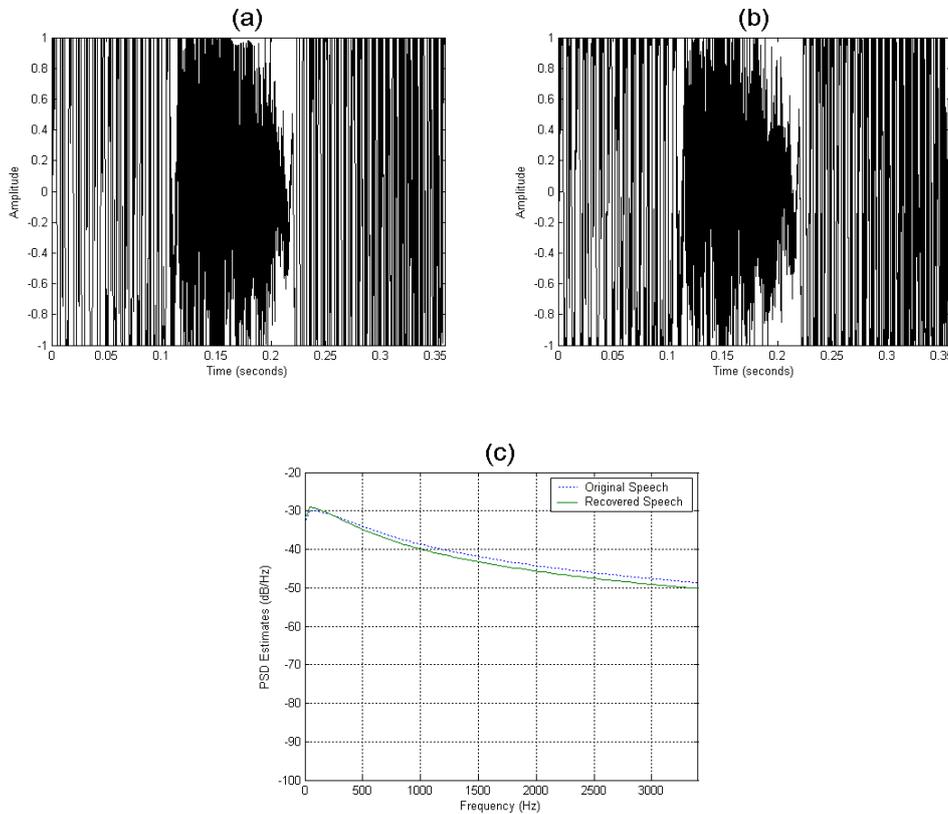
(٣) رجل (الحمد لله)



(٤) رجل (الله أكبر)



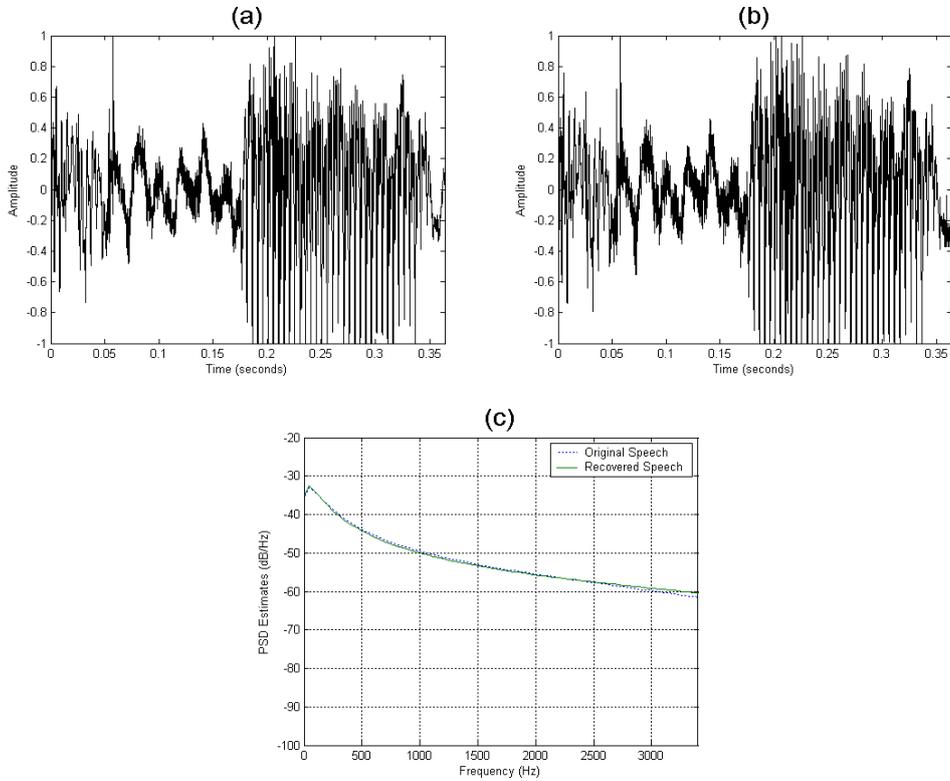
(٥) امرأة (مساء)



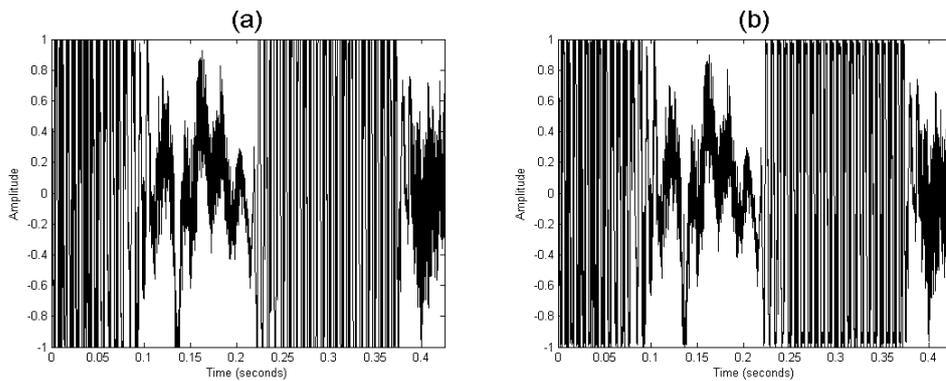
(٦) رجل (مساء)

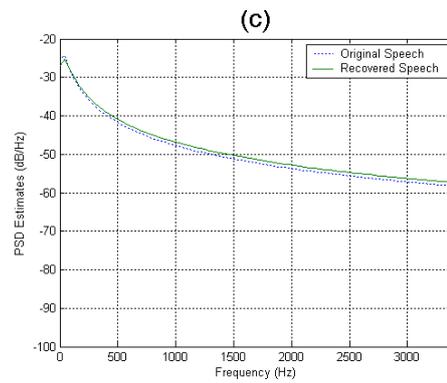
شكل (١.٥) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة : (٤٤١٠٠ Hz)

- a. الإشارة الكلامية الأصلية .
- b. الإشارة الكلامية المسترجعة .
- c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة .

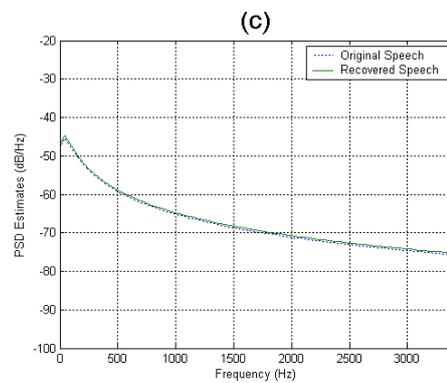
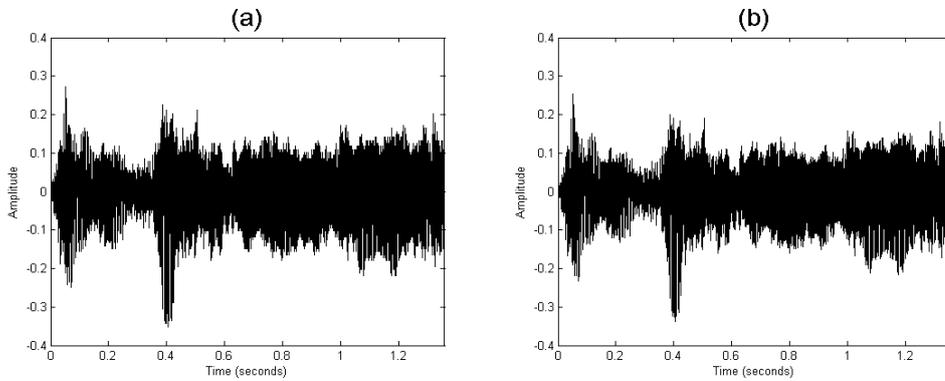


(١) امرأة (الخير)

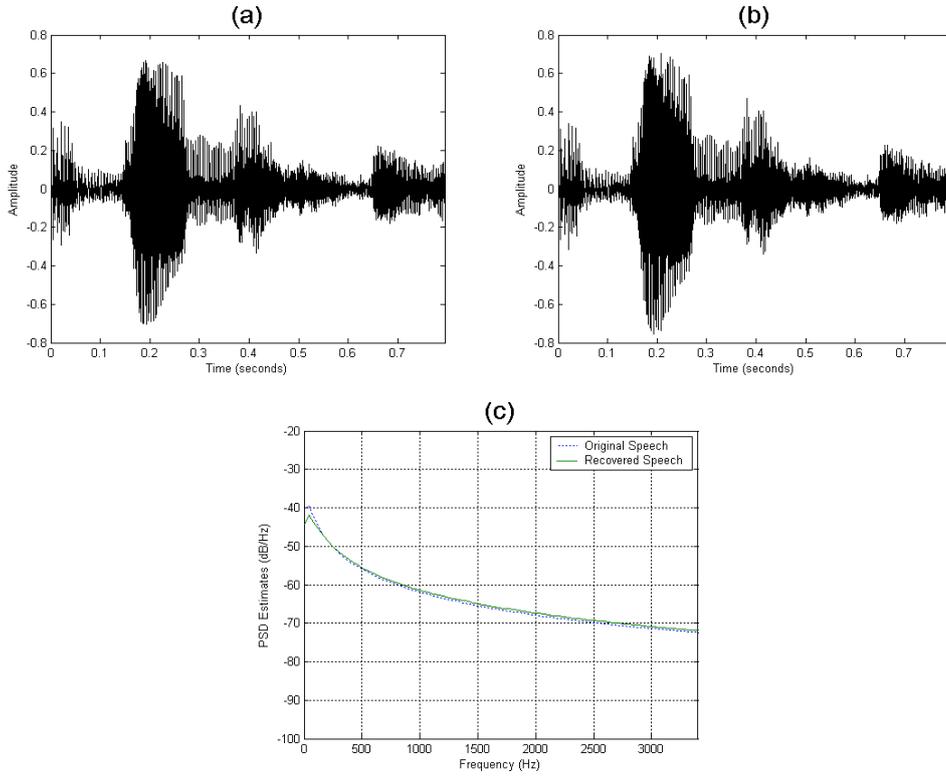




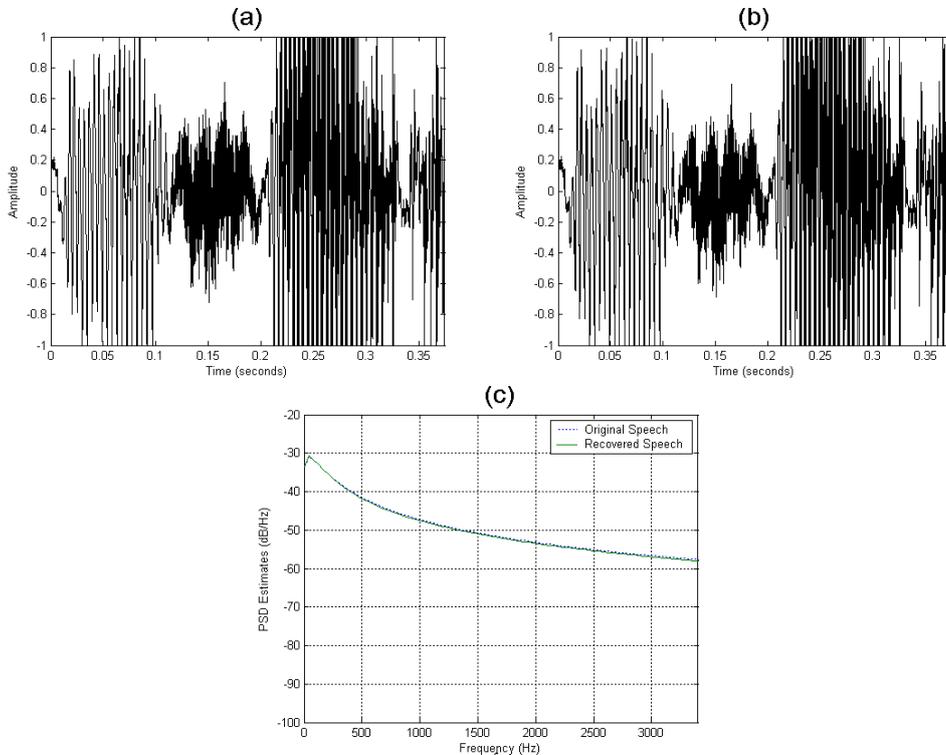
(٢) رجل الخير



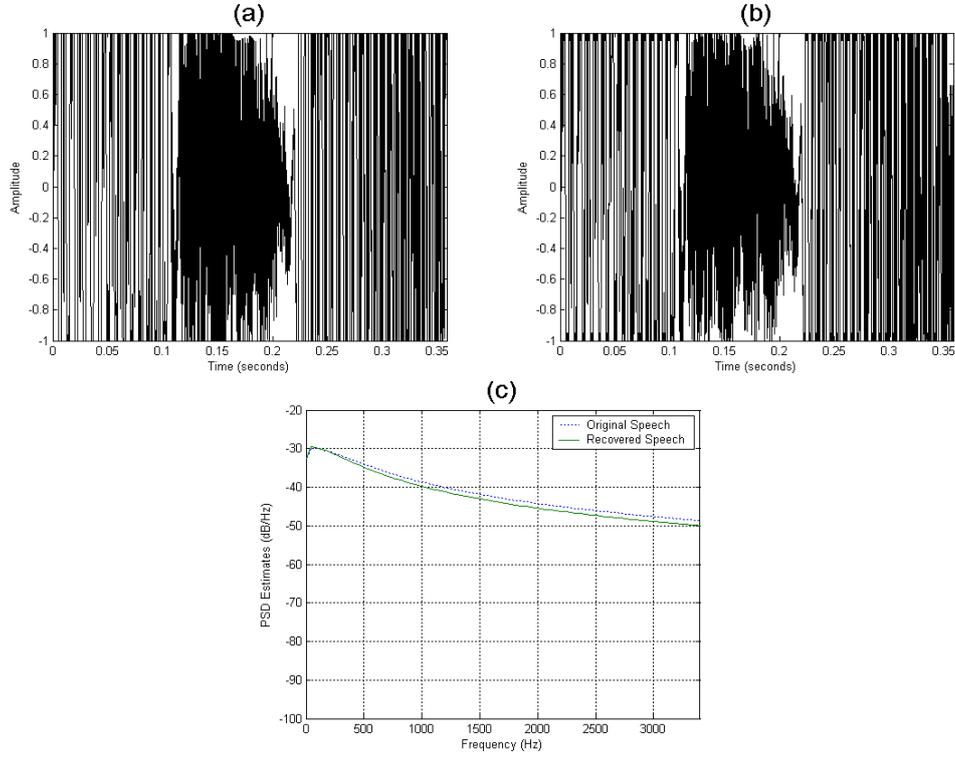
(٣) رجل (الحمد لله)



(٤) رجل (الله أكبر)



(٥) امرأة (مساء)



يتضح من الأشكال (a) و (b) التطابق بين الإشارات الأصلية والإشارات المسترجعة .

أما الأشكال (c) فيتضح منها وجود فرق قليل أو عدم وجوده بين الإشارتين الأصلية والمسترجعة وهذا الفرق غير ثابت إذ يزداد بزيادة التردد حيث إن قيمته قليلة ولا تؤثر في مدى التطابق بين الإشارة الكلامية الأصلية والمسترجعة .

تم تقويم كفاءة الحالة الدراسية الأولى باستخدام المقاييس الرياضية والمقياس الحسي MOS ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (١-٥) و(٢-٥) على التوالي عند تردد نمذجة (٤٤١٠٠ Hz) .

جدول (١-٥) نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE	MOS
امراة (الخير)	١٠.٨٦٠٠	١٠.١٠٣٧	٠.٠١٠٠	٤.١٤
رجل (الخير)	١٠.٩٦١٧	٩.٤٦٨٨	٠.٠٤٠٨	٤.٢٣

رجل (الحمد لله)	٨.٧٥٠٥	٩.٥٧٨٦	٠.٠٠٠٦	٣.٩١
رجل (الله أكبر)	٧.٣٧٤٧	٧.٨٠٤٦	٠.٠٠٣١	٣.٦٥
امرأة (مساء)	٧.٦٢٨٩	٩.١٢٨٨	٠.٠٣٥٠	٣.٥٢
رجل (مساء)	٨.٨٠٤٧	٩.٢٣٠٠	٠.٠٨٣٥	٤

جدول (٢-٥) نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالاعتماد على النافذة الهجينة من الاتجاه المقترح

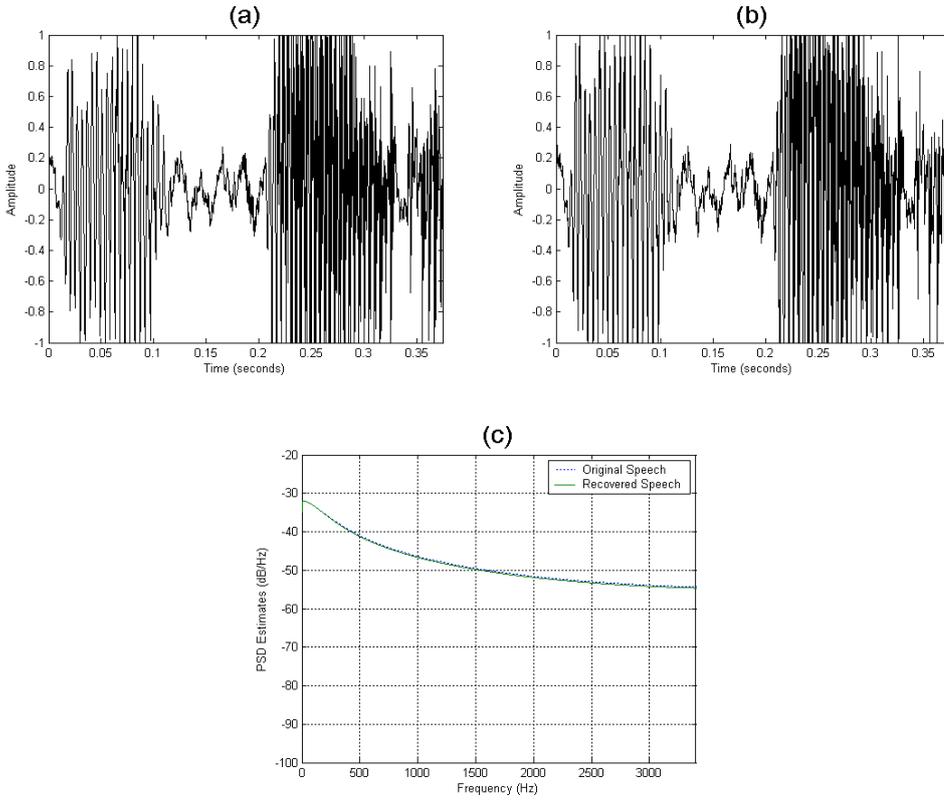
مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE	MOS
امرأة (الخير)	١١.١٤٦٣	١٠.٢٣٤٠	٠.٠٠٩٤	٤.٨٤
رجل (الخير)	١١.٠٣٦٦	٩.٥٢٩٥	٠.٠٤٠١	٤.٧٩
رجل (الحمد لله)	٩.٣٧٤٥	١٠.١١٤٩	٠.٠٠٠٥	٤.٧
رجل (الله أكبر)	٨.٨٣٥٨	٨.٨٠٦٠	٠.٠٠٢٢	٤.٥٥
امرأة (مساء)	٧.٩١٠٠	٩.٢٣٥٤	٠.٠٣٢٨	٤.٧٨
رجل (مساء)	٩.٠٧٣٣	٩.٦٢٦٩	٠.٠٧٨٥	٤.٦٩

وتوضح الأشكال (٣.٥) و (٤.٥) الإشارة الكلامية الأصلية ، والمسترجعة ، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٨٠٠٠ Hz) .

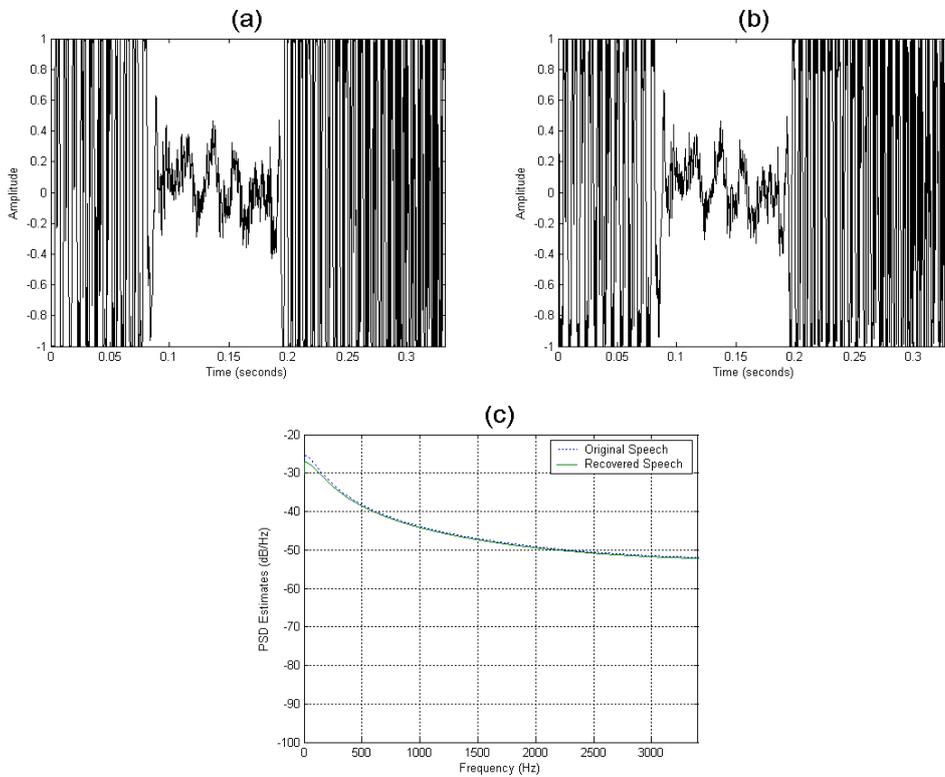
علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين .

ملاحظة: إن الأشكال من (٤-١) ستكون العناوين الخاصة بأجزائها الثلاثة كالاتي:

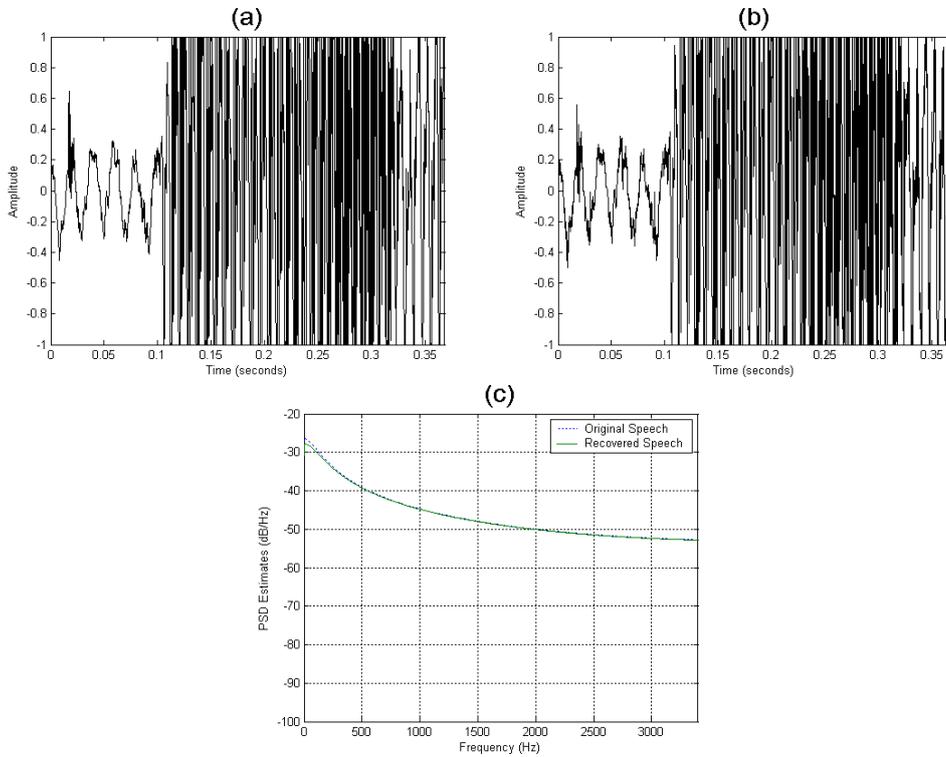
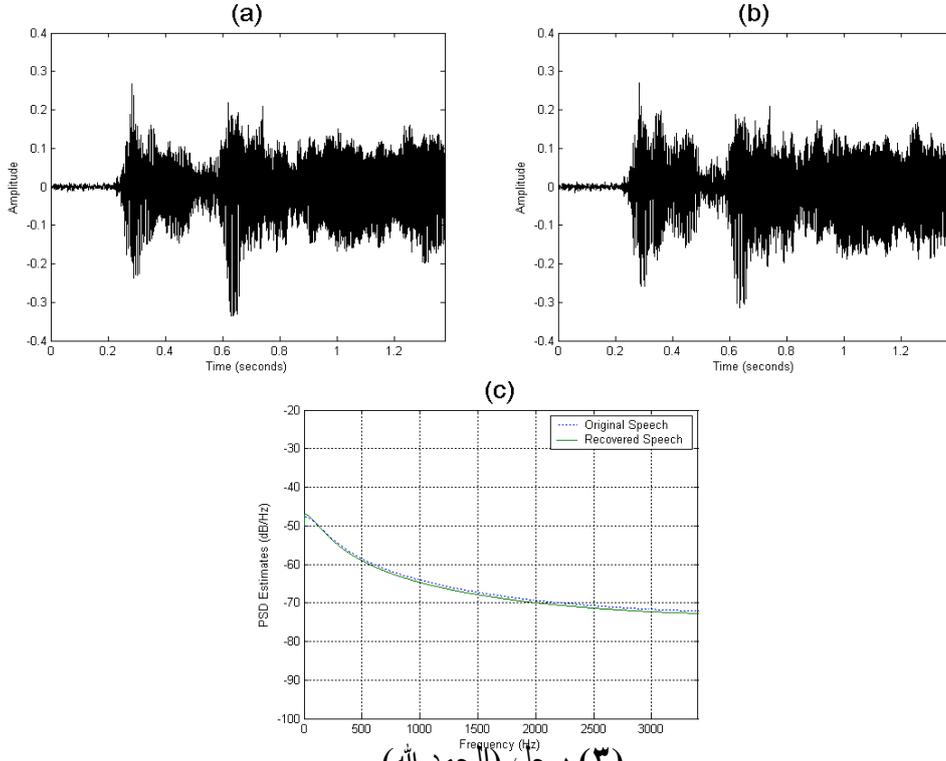
- الإشارة الكلامية الأصلية .
- الإشارة الكلامية المسترجعة .
- مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة .



(١) امرأة (مساء)



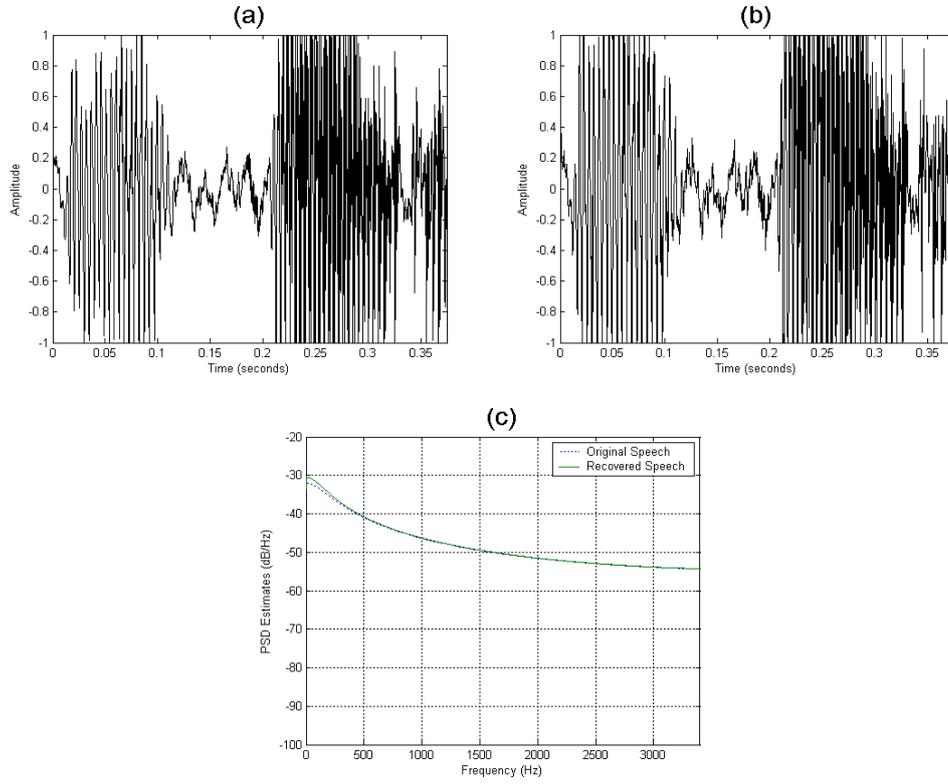
(٢) رجل (مساء)



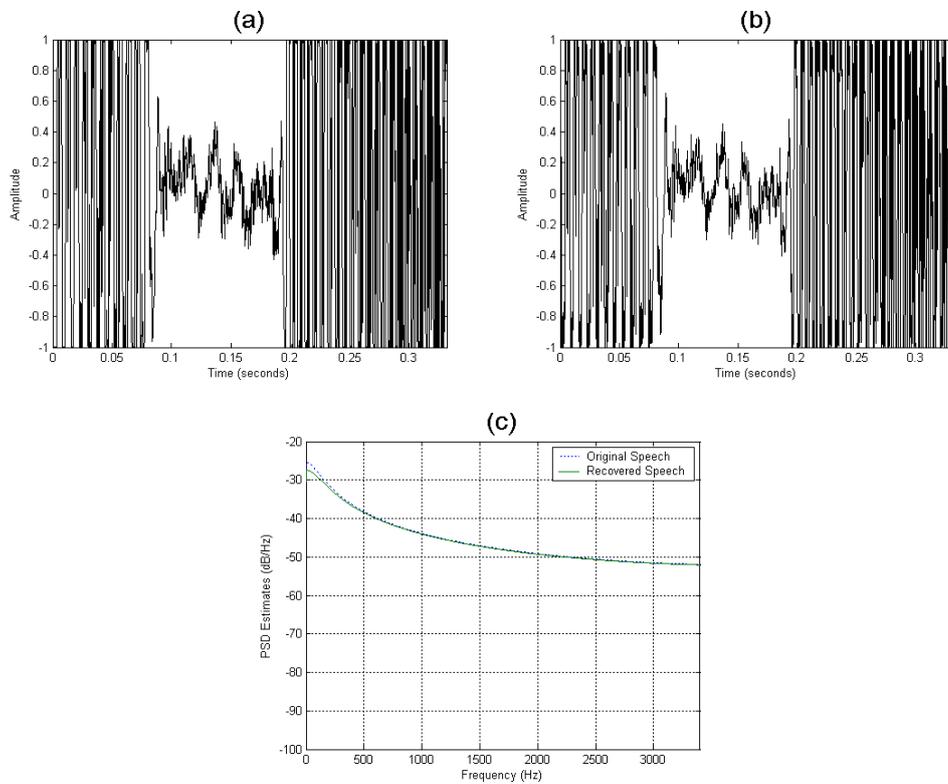
شكل (3.5) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (8000 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

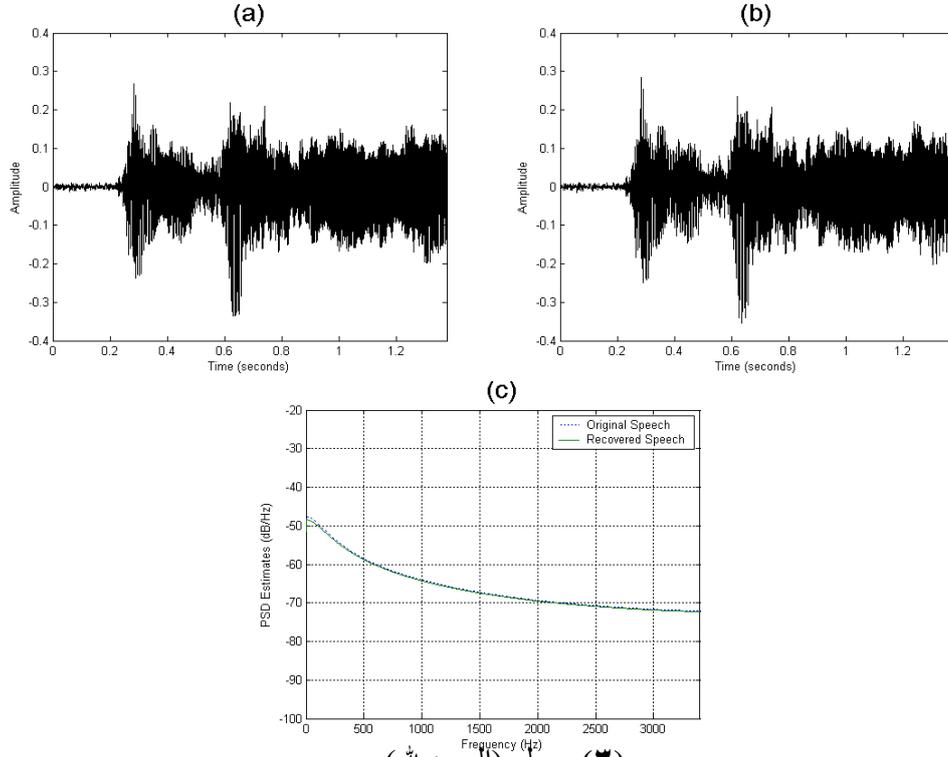
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.



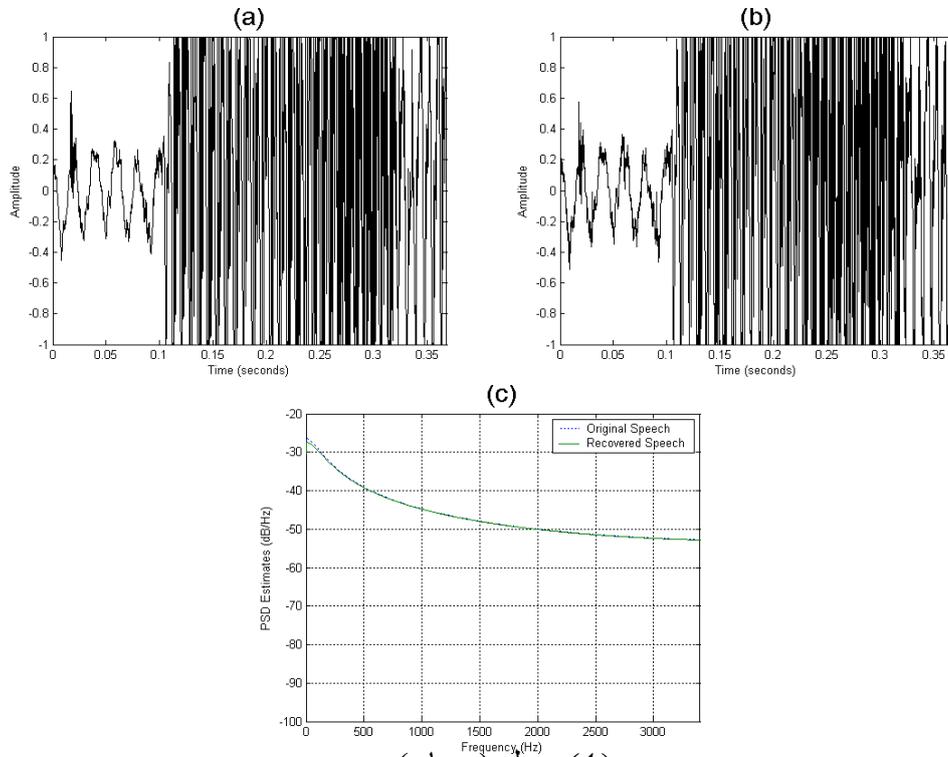
(١) امرأة (مساء)



(٢) رجل (مساء)



(٣) رجل (الحمد لله)



(٤) رجل (صباح)

شكل (٤.٥) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (8000 Hz) :

ينتضح من الأشكال (a) و (b) التطابق بين الإشارات الأصلية والإشارات المسترجعة. a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.

أما الأشكال (c) فيتضح منها وجود فرق قليل أو عدم وجوده بين الإشارتين الأصلية والمسترجعة وهذا الفرق غير ثابت إذ يزداد بزيادة التردد حيث إن قيمته قليلة ولا تؤثر في مدى التطابق بين الإشارة الكلامية الأصلية والمسترجعة .

تم تقويم كفاءة الحالة الدراسية الأولى باستخدام المقاييس الرياضية والمقياس الحسي MOS ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (٣-٥) و(٤-٥) على التوالي عند تردد نمذجة (٨٠٠٠ Hz) .

جدول (٣-٥) نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE	MOS
امرأة (مساء)	٣.٢٦٢٨	٤.٠١٣٢	٠.٠٨٨٢	٣.٦٥
رجل (مساء)	٦.٣٩٦٢	٦.٠٠٣٣	٠.١٢١٢	٤.٢٥
رجل (الحمد لله)	٢.١٤٩٨	٢.٤٨٥٤	٠.٠٠٢٤	٣.٤٨
رجل (صباح)	٤.٥٦٧٠	٥.٦٨٦٤	٠.١٤٨٥	٤.١٩

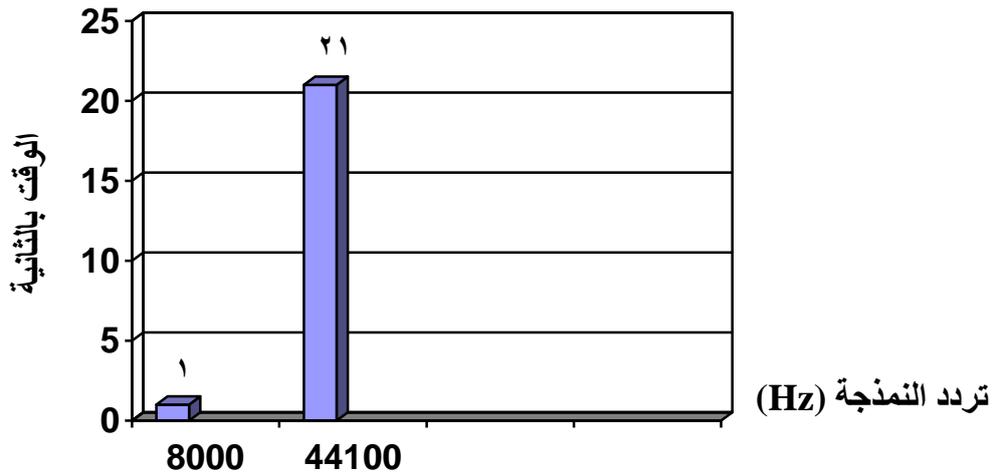
جدول (٤-٥) نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE	MOS
امرأة (مساء)	٣.٣٨٣٦	٤.٢٦٤٩	٠.٠٨٥٨	٤.٧٢
رجل (مساء)	٦.٤٢٩٥	٦.٠١٤٢	٠.١٢٠٢	٤.٩٢

رجل (الحمد لله)	٢.٥٣٩٦	٢.٩٠٦٦	٠.٠٠٢٢	٤.١٧
رجل (صباح)	٤.٥٩٢٣	٥.٧١٣٠	٠.١٤٧٦	٤.٨٥

توضح الجداول (١-٥) إلى (٤-٥) نتائج الحالة الدراسية الأولى للترميز عند ترددي نمذجة (٤٤١٠٠ Hz) و (٨٠٠٠ Hz). حيث يتضح من الجدول بان أفضل النتائج للمقاييس الرياضية (MSE, SNR_{seg}, SNR) والمقياس الحسي (MOS) لجميع حالات إشارات الاختبار المأخوذة هي التي تم الحصول عليها من النافذة الهجينة من الاتجاه المقترح والتي يكون فيها الجزء الأول هو الربع الثاني والربع الثالث من دالة الجيب تمام المضروبة بنصف عينات الإطار بينما يتكون الجزء الثاني من دالة الهامنك المضروبة بالنصف المتبقي من عينات الإطار لذا يفضل اعتماد هذه النافذة في الحالة الدراسية الأولى لعملية الترميز.

ويوضح الشكل (٥.٥) الوقت المستغرق لكلمة مساء منطوقة من قبل رجل وبالاعتماد على نافذة الاتجاه المقترح عند ترددي نمذجة (٨٠٠٠ Hz) و (٤٤١٠٠ Hz).



شكل (٥.٥) الوقت المستغرق لكلمة مساء

٢.٣.١.٥ الحالة الدراسية الثانية

تم تنفيذ النظام المقترح بالاعتماد على المفهوم الذي يُعدّ النظام عبارة عن حلقة مغلقة. في هذا المفهوم تتضمن عملية تحليل الإشارة الكلامية، تحليل التنبؤ الخطي عبر إطار إشارة الكلام الجزئي بعد تقطيع موجة إشارة الكلام الداخلة إلى إطارات ومن ثم تجزئة كل إطار إلى أربعة إطارات جزئية، وذلك لاستخلاص معاملات التنبؤ الخطي لكل إطار جزئي وبعاد تركيب الإشارة الكلامية لكل إطار جزئي وباستخدام مرشح التحسين (FIR Filter) بعد تركيب موجة الإشارة الكلامية المسترجعة لتحسينها.

وتوضح الأشكال (1.A) و(2.A) في الملحق (A) الإشارة الكلامية الأصلية، والمسترجعة، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (3.2.4) عند تردد نمذجة (44100 Hz).

علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD)، مقاساً للإشارتين.

تم تقويم كفاءة الحالة الدراسية الثانية باستخدام المقاييس الرياضية السابقة، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (3.2.4)، والنتائج موضحة في الجداول (5-5) و(6-5) على التوالي عند تردد نمذجة (44100 Hz).

جدول (5-5) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امراة (الخير)	8.9927	8.2310	0.0154
رجل (الخير)	7.6282	6.5383	0.0880
رجل (الحمد لله)	7.9919	8.0490	0.0008
رجل (الله أكبر)	6.5018	6.7060	0.0038
امراة (مساء)	6.7885	6.8188	0.0425
رجل (مساء)	7.6902	7.4825	0.1079

جدول (6-5) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (الخير)	٩.٠١٨٨	٨.٣٨٦٥	٠.٠١٥٣
رجل (الخير)	٧.٨٦٢٣	٦.٦٤١٠	٠.٠٨٣٤
رجل (الحمد لله)	٨.١٨١٩	٨.٢٩٦٦	٠.٠٠٠٧٨٢٩ ٥
رجل (الله أكبر)	٧.١٦٧١	٧.٢٤٥٤	٠.٠٠٣٢
امرأة (مساء)	٦.٩٥٠٧	٧.٠٦٥١	٠.٠٤٠٩
رجل (مساء)	٧.٨٣٨٨	٧.٧٣٣٥	٠.١٠٤٣

وتوضح الأشكال (٣.A) و (٤.A) في الملحق (A) الإشارة الكلامية الأصلية ، والمسترجعة، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٨٠٠٠ Hz) .

علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين .

تم تقويم كفاءة الحالة الدراسية الثانية باستخدام المقاييس الرياضية السابقة ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (٧-٥) و (٨-٥) على التوالي عند تردد نمذجة (٨٠٠٠ Hz) .

جدول (٧-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	٢.٨١٥٤	٣.٤٠٢٠	٠.٠٩٧٨
رجل (مساء)	٥.٦١٨٥	٥.٣٠٠٥	٠.١٤٤٩
رجل (الحمد لله)	٢.٤٨٣٤	٢.٧٥٦٩	٠.٠٠٢٣

رجل (صباح)	٣.٦٦٠٣	٤.٠١٥٧	٠.١٨٢٩
------------	--------	--------	--------

جدول (٨-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	٣.١٣٠٨	٣.٧٨٧٣	٠.٠٩٠٩
رجل (مساء)	٥.٧٦٣٨	٥.٣٩٤٦	٠.١٤٠٢
رجل (الحمد لله)	٢.٤٩٤٧	٢.٨٠٣٤	٠.٠٠٢٣
رجل (صباح)	٣.٧٤٥٢	٤.١١٣٣	٠.١٧٩٤

توضح الجداول (٥-٥) إلى (٨-٥) نتائج الحالة الدراسية الثانية للترميز عند ترددي نمذجة (٤٤١٠٠ Hz) و (٨٠٠٠ Hz) حيث يتضح من الجدول بان أفضل النتائج للمقاييس الرياضية (MSE , SNR_{seg} , SNR) لجميع حالات إشارات الاختبار المأخوذة هي التي تم الحصول عليها من النافذة الهجينة من الاتجاه المقترح والتي يكون فيها الجزء الأول هو الربع الثاني والربع الثالث من دالة الجيب تمام المضروبة بنصف عينات الإطار الجزئي بينما يتكون الجزء الثاني من دالة الهامتك المضروبة بالنصف المتبقي من عينات الإطار الجزئي . لذا يفضل اعتماد هذه النافذة في الحالة الدراسية الثانية لعملية الترميز .

٣.٣.١.٥ الحالة الدراسية الثالثة

تم تنفيذ النظام المقترح بالاعتماد على المفهوم نفسه الذي أعتمد في الحالة الدراسية الأولى إلا إنه لم يتم استخدام مرشح التحسين (FIR Filter) بعد تركيب موجة الإشارة الكلامية المسترجعة لتحسينها في هذه الحالة .

وتوضح الأشكال (١.B) و (٢.B) في الملحق (B) الإشارة الكلامية الأصلية ، والمسترجعة ، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٤٤١٠٠ Hz) .

علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين .

تم تقويم كفاءة الحالة الدراسية الثالثة باستخدام المقاييس الرياضية السابقة، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤)، والنتائج موضحة في الجداول (٩-٥) و (١٠-٥) على التوالي عند تردد نمذجة (٤٤١٠٠ Hz).

جدول (٩-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امراة (الخير)	٢.٨٥٣١	٢.٦٤٣١	٠.٠٦٣٣
رجل (الخير)	٢.٨٢٢٩	٢.٥٦٠٤	٠.٢٦٦٠
رجل (الحمد لله)	٢.٧٠٥٧	٢.٥٩٧٢	٠.٠٠٢٨
رجل (الله أكبر)	٢.٧٥٨٨	٢.٧٧٤٤	٠.٠٠٨٩
امراة (مساء)	٢.٥٦٥٧	٢.٢٧١٣	٠.١١٢٣
رجل (مساء)	٢.٥٤٤٠	٢.٢٤٥٠	٠.٣٥٣٠

جدول (١٠-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امراة (الخير)	٢.٨٣٦٦	٢.٥٧٨٦	٠.٠٦٣٥
رجل (الخير)	٢.٨٢٧٢	٢.٥٥٨٩	٠.٢٦٥٧
رجل (الحمد لله)	٢.٧٩٣٣	٢.٧٢٧٤	٠.٠٠٢٧
رجل (الله أكبر)	٢.٨٠٦٩	٢.٧٩٣٩	٠.٠٠٨٨
امراة (مساء)	٢.٦٢٣٨	٢.٣٠٩١	٠.١١٠٨

رجل (مساء)	٢.٥٦٨٦	٢.٢٦٦٢	٠.٣٥١٠
------------	--------	--------	--------

وتوضح الأشكال (٣.B) و (٤.B) في الملحق (B) الإشارة الكلامية الأصلية ، والمسترجعة، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٨٠٠٠ Hz) .

علمًا إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين .

تم تقويم كفاءة الحالة الدراسية الثالثة باستخدام المقاييس الرياضية السابقة ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (١١-٥) و (١٢-٥) على التوالي عند تردد نمذجة (٨٠٠٠ Hz) .

جدول (١١-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	١.٦٩٧٩	١.٨٧١٧	٠.١٢٦٥
رجل (مساء)	٢.٣٥١٨	٢.٢٥٧٤	٠.٣٠٧٥
رجل (الحمد لله)	٠.٨٢٤١	٠.٩٢٥٩	٠.٠٠٣٣
رجل (صباح)	٢.٠١٦٧	٢.١٩٥٦	٠.٢٦٧١

جدول (١٢-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	١.٦٧٤٢	١.٧٩٩٧	٠.١٢٧١

رجل (مساء)	٢.٣٦٦٣	٢.٢٦٦٦	٠.٣٠٦٥
رجل (الحمد لله)	١.٣٨٤٠	١.٤٨٥٥	٠.٠٠٢٩
رجل (صباح)	٢.٠١٦١	٢.١٩٥٩	٠.٢٦٧١

توضح الجداول (٩-٥) إلى (١٢-٥) نتائج الحالة الدراسية الثالثة للترميز عند ترددي نمذجة (٤٤١٠٠ Hz) و (٨٠٠٠ Hz).

٤.٣.١.٥ الحالة الدراسية الرابعة

تم تنفيذ النظام المقترح بالاعتماد على المفهوم نفسه الذي أعتمد في الحالة الدراسية الثانية إلا إنه لم يتم استخدام مرشح التحسين (FIR Filter) بعد تركيب موجة الإشارة الكلامية المسترجعة لتحسينها في هذه الحالة .

وتوضح الأشكال (١.C) و (٢.C) في الملحق (C) الإشارة الكلامية الأصلية ، والمسترجعة ، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة ، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٤٤١٠٠ Hz) .

علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD) ، مقاساً للإشارتين تم تقويم كفاءة الحالة الدراسية الرابعة باستخدام المقاييس الرياضية السابقة ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (١٣-٥) و (١٤-٥) على التوالي عند تردد نمذجة (٤٤١٠٠ Hz) .

جدول (١٣-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امراة (الخير)	٢.٥٩٠٣	٢.٣٥٧١	٠.٠٦٧٢
رجل (الخير)	٢.٦٠٣٦	٢.٣٤٥٣	٠.٢٧٩٨
رجل (الحمد لله)	٢.٥٤٩٥	٢.٥٠٠٧	٠.٠٠٢٩
رجل (الله أكبر)	٢.٥٩٨٣	٢.٥٧٤٦	٠.٠٠٩٣

امرأة (مساء)	٢.٣٨٢٠	٢.١٤٤٢	٠.١١٧١
رجل (مساء)	٢.٣٦٣٥	٢.٠٩٩٣	٠.٣٦٨٠

جدول (١٤-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (الخير)	٢.٦٢٥٩	٢.٤٤١٠	٠.٠٦٦٧
رجل (الخير)	٢.٥٤٧١	٢.١٤٢٢	٠.٢٨٣٤
رجل (الحمد لله)	٢.٥٧٤٦	٢.٥٢١٤	٠.٠٠٢٨
رجل (الله أكبر)	٢.٥٨٥٨	٢.٥٧٣١	٠.٠٠٩٣
امرأة (مساء)	٢.٤١٢١	٢.١٥٦٣	٠.١١٦٣
رجل (مساء)	٢.٣٠٨٨	١.٩٥٣٢	٠.٣٧٢٧

وتوضح الأشكال (٣.C) و (٤.C) في الملحق (C) الإشارة الكلامية الأصلية، والمسترجعة، ومدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة، على التوالي وبالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) عند تردد نمذجة (٨٠٠٠ Hz).

علماً إن معلم التقويم لمدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة هو كثافة طيف القدرة (PSD)، مقاساً للإشارتين.

تم تقويم كفاءة الحالة الدراسية الرابعة باستخدام المقاييس الرياضية السابقة ، بالاعتماد على اتجاهين من النوافذ الهجينة المشار إليها في البند (٣.٢.٤) ، والنتائج موضحة في الجداول (١٥-٥) و (١٦-٥) على التوالي عند تردد نمذجة (٨٠٠٠ Hz) .

جدول (١٥-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه الأساسي

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	١.٤٤٥٦	١.٦٠٧٢	٠.١٣٤٠
رجل (مساء)	٢.١٧٩٥	٢.١٠٦٣	٠.٣١٩٩
رجل (الحمد لله)	١.٥١٩٢	١.٥٩٥٦	٠.٠٠٢٨
رجل (صباح)	١.٨٩١٢	١.٩٩٨٩	٠.٢٧٤٩

جدول (١٦-٥) نتائج استخدام المقاييس الرياضية بالاعتماد على النافذة الهجينة من الاتجاه المقترح

مقاييس التقويم إشارة الاختبار	SNR (dB)	SNR _{seg} (dB)	MSE
امرأة (مساء)	١.٥٠٥٩	١.٥٩٩١	٠.١٣٢٢
رجل (مساء)	٢.١٨٣٣	٢.١٠٨٠	٠.٣١٩٦
رجل (الحمد لله)	١.٤٣٧٣	١.٥٢٦٤	٠.٠٠٢٩
رجل (صباح)	١.٨٥٠٢	١.٩٧٤٠	٠.٢٧٧٥

توضح الجداول (١٣-٥) إلى (١٦-٥) نتائج الحالة الدراسية الرابعة للترميز عند ترددي نمذجة (٤٤١٠٠ Hz) و (٨٠٠٠ Hz) .

٢.٥ المناقشة والاستنتاجات

١. توجد تغيرات سريعة جداً في الإشارة الكلامية وتأخذ حيزاً زمنياً قليلاً، إن سعة الإشارة الكلامية عند التغير ربما لا يمكن ملاحظته في الطرائق التقليدية لتحليل الإشارة الكلامية وعليه فإن عملية التحليل تجرى على قطع قصيرة من إشارات الكلام والتي تسمى الإطارات تتراوح بين (٣٠-٥٠) إطار لكل ثانية والتي تمثل كلاماً مفهوماً مع ترميز جيد، وعلى هذا الأساس فقد تم تقطيع سيل الإشارة الكلامية المدخلة إلى مرز إثارة التنبؤ الخطي إلى إطارات كل إطار ذي مدة زمنية (٣٠ ms).

٢. من خلال تجربة العديد من القيم لرتبة مرشح التنبؤ الخطي فقد وجد بأن الاختيار الملائم في مرحلة التحليل لقيمة الرتبة هي (١٠ = order) وذلك من خلال إجراء اختبار لغرض اختيار الرتبة شكل (٨.٤) إذ لوحظ ثبوت معدل مربع الخطأ تقريباً ابتداءً من قيمة الرتبة (١٠) فما فوق، حيث إن الاختيار الملائم للرتبة يؤدي إلى تقليل معدل مربع الخطأ الكلي وبالتالي الحصول على إشارة كلامية مشابهة للإشارة الكلامية المدخلة الأصلية (أما اختيار رتبة قليلة فينتج معاملات أقل وحسابات أقل ولكن بدقة قليلة، وباختيار الرتبة كبيرة تنتج معاملات أكثر وحسابات أكثر ولكن بدقة أكثر. ومن هنا يتم التنسيق بين الرتبة والدقة باختيار الرتبة (١٠)).

٣. إن تعديل معاملات الارتباط الذاتي أفضل من عدم تعديلها وذلك لتجنب مشاكل الإشارات الداخلة منخفضة المستوى وبالتالي الحصول على مرشح تركيبى مستقر.

٤. إن تقسيم معاملات الارتباط الذاتي على أول معامل من معاملاتنا يؤدي إلى التخلص من القيم العالية جداً مما يسهل من عملية الحساب حيث إن معاملات الارتباط الذاتي أساسية للاستمرار في عملية التحليل.

٥. إن وجود المرشح يُعد حالة إيجابية نسبة إلى عدم وجوده لأنه من خلاله يمكن للنظام أن يتحكم بعدد من مواصفات الإشارة الناتجة من المعالجة سواء كان ربح هذه الإشارة أو تردد قطعها، إذ يفضل في أنظمة معالجة الإشارات الاعتماد على استخدام المرشحات بمختلف أنواعها. إذن يُعد المرشح مكون أساسي من مكونات معالجة الإشارة الكلامية ضمن هذا النظام، هذه المسألة ترتبط أيضاً بطريقة تصميم المرشح ونوعية المرشح المستخدم، لذا فقد تم استخدام مرشح الإمرار الواطئ لاستجابة النبضة المتناهية (FIR) لتحسين جودة الصوت الناتج من عملية التركيب.

٦. أثبت قيمة الرتبة (٤) لمرشح الإمرار الواطئ لاستجابة النبضة المتناهية (FIR) بأنها اختيار ملائم بعد سلسلة من التجارب والاختبارات العلمية (علماً أنه قد تم تجربة العديد من القيم الأخرى لقيمة الرتبة كانت أكثر واقل من هذه القيمة ولكن وجد بأن اختيار قيمة للرتبة أقل أدى إلى تقليل وضوح وشدة الإشارة الكلامية المعاد بناؤها وباختيار قيمة أعلى للرتبة أدى إلى الحصول على النتائج نفسها التي حصل عليها من قيمة الرتبة المختارة وعلى هذا الأساس تم الاختيار).

٧. في المنظومات التي تتعامل مع الإشارة الكلامية في مسألة التحليل بالتركيب غالباً ما تكون الإشارة الكلامية مرشحة بمرشح إمرار واطئ ذي تردد قطع (٣.٤ kHz) ومن المعروف جيداً

إن الطاقة الموجودة بالإشارة الكلامية طاقة قليلة جداً بعد هذا التردد لذلك تكون غير مهمة ولذا فعند مناقشة الأشكال (c) والتي تمثل (مدى التطابق) (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة) يجب أن ننتبه إن النطاق الترددي المهم بالنسبة لنا هو من صفر إلى حد (3.4 kHz) ، لذلك من خلال قياس الفرق بين الإشارة الكلامية الأصلية والمسترجعة لكل صوت منطوق نكتشف إن الفرق يختلف من كلمة إلى أخرى حيث يكون هذا الفرق أحياناً صفرأ في بعض الكلمات وأحياناً الفرق موجب وأحياناً أخرى الفرق سالب أي يعني عدم وجود سياق موحد لطبيعة الفرق الناتج .

٨. هنالك نقاط على محور التردد يكون فيها الفرق صفر أو قيمة موجبة أو سالبة بين الإشارة الأصلية والإشارة المسترجعة ، مسألة نوعية الفرق كاختلاف من موجب إلى سالب أو صفر توجد بوضوح ضمن المدى الترددي الخاضع للاهتمام ، بينما هذه المسألة لا تظهر بنفس الوضوح التي تظهر به في حالة (8000 Hz) عندما يكون تردد النمذجة (44100 Hz) ، مثل هكذا مرمرات يمكن أن يستخدم مرشح بتردد قطع (2.4 kHz) حتى نقلل الفرق الكلي بين الإشارتين الأصلية والمسترجعة نسبة مما لو أخذ مرشح بتردد قطع (3.4 kHz) وهذا ما يتم استخدامه في الأجهزة العسكرية .على الرغم من إن هذه المسألة تؤدي إلى تقليل مفهومية الكلام .

٩. مقاييس مقارنة الكفاءة للمرمز – فاك الترميز الذي يعتمد على مبدأ علامة معدل الرأي (MOS) لا يعتمد على استخدام شخص واحد للتقويم ، إنما باعتماد عدد من المقومين بحيث يؤخذ المعدل لتقاويمهم وعلى أساسه يتم الحكم على المرمز حيث كلما كان عدد المقومين كبيراً كلما كانت دقة التقويم أفضل بالاعتماد على مبدأ المعدل للتقاويم.

١٠. إن أفضل نتائج التقويم هي التي تم الحصول عليها من تطبيق النافذة الهجينة من الاتجاه المقترح لذا يفضل اعتماد هذه النافذة في الحالات الدراسية الأولى والثانية لعملية الترميز . حيث تكون مقاييس التقويم SNR و SNR_{seg} و MOS للنافذة المقترحة أعلى من SNR و MOS للنافذة من الاتجاه الأساسي بينما يكون مقياس التقويم MSE للنافذة المقترحة أقل من MSE للنافذة من الاتجاه الأساسي ، ومن هنا نستنتج بأن نافذة الاتجاه المقترحة هي الأفضل ولكل إشارات الاختبار المطبقة لذا نوصي باعتمادها.

١١. لم توضح خوارزمية ليفنسن – داربن سبب استخدام النافذة الموضحة في الأدبيات ، ولذلك يعد المقترح الذي قدم في هذه الرسالة مدخلاً جديداً ومهماً في إيجاد أفضل ترتيب لنافذة هجينة تستخدم مع خوارزمية ليفنسن – داربن المستخدمة في مرمرات إشارة الكلام من نوع CELP .

وقد تم اقتراح تركيبات مختلفة للنافذة الهجينة بإبدال دالة الهامنك ، ومعرفة مقدار تأثير تبديل هذه الدالة على فاعلية الخوارزمية ، فعند تصميم نوافذ لتوزيع إطار الإشارة الكلامية المأخوذ في الاعتبار في موضوع الترميز وفك ترميز الإشارة الكلامية يفضل أن تكون النوافذ متغيرة الأشكال والأحجام حيث يمكن أن تؤدي إلى إيجاد تشكيله هجينة ذات مواصفات تقويمية توفر أفضل النتائج . لذلك تم اقتراح تركيب نافذة هجينة نموذجية بعد سلسلة من التجارب والاختبارات العملية .

١٢. من خلال التجارب الاختبارية السابقة نلاحظ إن القيم التي تم الحصول عليها من الحالة الدراسية الأولى لعملية الترميز هي قيم جيدة جداً وذات مفهومية عالية وجودة ممتازة وذلك يعزى إلى سببين الأول هو لإجراء عملية التحليل والتركيب للإشارة الكلامية لكل إطار. إن التحليل وتركيب الإشارة الكلامية باستخدام إطار كلي للإشارة الخاضعة للاختبار يُعدّ أفضل مما لو اتخذنا مبدأ التحليل وتركيب الإشارة الكلامية باستخدام إطار جزئي من الإشارة الخاضعة ، أي أن الإطار الكلي يخضع إلى عملية تجزئة إلى أربعة إطارات جزئية، ويخضع كل منها إلى عملية تحليل وتركيب كاملة . ولكن أوضحت النتائج عدم أفضلية هذه الطريقة الأخيرة نسبة للطريقة الأولى ، وذلك لأن عدد عناصر الإشارة الخاضعة للاختبار في حالة الإطار الجزئي أصبح أقل بكثير من عدد العناصر في الإطار الكلي ، مما يسبب في أحيان كثيرة غلبة بعض الصفات السيئة على عملية التحليل والتركيب، ولذلك تتأثر معاملات القياس بهذه الحقيقة ، لأن عينة النماذج المأخوذة من الإشارة (الخاضعة للاختبار) قد اختزلت بدرجة كبيرة .

أما السبب الثاني هو استخدام مرشح التحسين (مرشح FIR) لتحسين جودة إشارة الكلام المسترجع والحصول على إشارة كلامية مسترجعة تشبه تماماً الإشارة الكلامية المدخلة.

بينما نلاحظ من خلال التجارب الاختبارية السابقة إن القيم التي تم الحصول عليها من الحالة الدراسية الثانية لعملية الترميز هي قيم جيدة وذات جودة جيدة جداً وأفضل من نتائج الحالات الدراسية الثالثة والرابعة وذلك بسبب استخدام مرشح التحسين (مرشح FIR) والذي يمكننا من الحصول على إشارة مسترجعة تشبه إشارة الكلام الداخلة ، إلا إن الإشارة المسترجعة لا تشبه تماماً الإشارة الكلامية الداخلة وذلك بسبب إجراء عملية التحليل والتركيب للإشارة الكلامية لكل إطار جزئي.

أما بالنسبة للحالة الدراسية الثالثة لعملية الترميز ومن خلال التجارب الاختبارية السابقة نلاحظ إن القيم التي تم الحصول عليها هي قيم أفضل من نتائج الحالة الدراسية الرابعة وذات جودة واضحة وذلك بسبب إجراء عملية التحليل والتركيب للإشارة الكلامية لكل إطار، ولكنها أسوء من نتائج الحالة الدراسية الثانية بسبب عدم استخدام مرشح التحسين .

وأخيراً من خلال التجارب الاختبارية السابقة يلاحظ إن القيم التي تم الحصول عليها من الحالة الدراسية الرابعة لعملية الترميز تكون مقبولة إلا إنها تعد نتائج قليلة الجودة والمفهومية وذات جودة ضعيفة وذلك يعزى إلى سببين الأول هو لإجراء عملية التحليل والتركيب للإشارة الكلامية لكل إطار جزئي ، أما السبب الثاني فيعود إلى إن عدم استخدام مرشح التحسين في هذه الحالة يؤدي إلى استرجاع إشارة كلام لا تشبه إلى حد ما إشارة الكلام الداخلة .

١٣. أظهرت نتائج التطبيق بان أفضل حالة دراسية كانت الحالة الدراسية الأولى والتي تم تنفيذ النظام المقترح فيها بالاعتماد على المفهوم الذي يُعدّ النظام عبارة عن حلقة مغلقة . في هذا المفهوم تتضمن عملية تحليل الإشارة الكلامية ، تحليل التنبؤ الخطي عبر إطار إشارة الكلام الكلي بعد تقطيع موجة إشارة الكلام الداخلة إلى إطارات ، وذلك لاستخلاص معاملات التنبؤ الخطي لكل إطار ويعاد تركيب الإشارة الكلامية لكل إطار وباستخدام مرشح التحسين (مرشح FIR) بعد تركيب موجة الإشارة الكلامية المسترجعة لتحسينها .

حيث وجد بأن نتائج الحالة الدراسية هذه أفضل من الحالات الدراسية الأخرى وذلك لأن عملية التحليل والتركيب أجريت لكل إطار ، إذ أثبتت التجارب العلمية بأن عملية التحليل والتركيب

للإشارة الكلامية لكل إطار كانت نتائجها أفضل مما لكل إطار جزئي ، بالإضافة إلى استخدام مرشح التحسين (مرشح FIR) لتحسين جودة إشارة الكلام المسترجع والحصول على إشارة كلامية مسترجعة تشبه تماماً الإشارة الكلامية المدخلة.

١٤. في أي مسألة من المسائل التي تتعامل في تحويل الإشارة من الصورة التماثلية إلى الصورة الرقمية وبالعكس يكون استخدام تردد نمذجة عالي أفضل بكثير من استخدام تردد نمذجة قليل (مع الالتزام بشرط Niquist) ، ولذلك في نتائج العمل كان استخدام تردد النمذجة (٤٤١٠٠ Hz) يعطي نتائج أكثر دقة من استخدام تردد النمذجة (٨٠٠٠ Hz) .

١٥. من النتائج التقويمية لعلاقة الـ (PSD) مع التردد بين الإشارة الأصلية والمسترجعة نلاحظ من الأشكال (c) التي تم الحصول عليها إن مدى التطابق (الاختلاف) بين إشارة الكلام الأصلية وإشارة الكلام المسترجعة يعتمد على معالم ثلاثة هي :

- إشارة الكلام المدخلة .
- النافذة المستخدمة .
- تردد النمذجة .

بالنسبة إلى المعلم الأول والذي يمثل الإشارة الكلامية المدخلة إلى النظام باعتقادنا إن هذه المسألة يمكن أن تعزى إلى خاصية مهمة من خواص الإشارات الصوتية وهي خاصية المدى الديناميكي للإشارة يكون مختلفاً من إشارة إلى أخرى وهذا الاختلاف يعني اختلاف الإشارات الصوتية في المحتوى المعلوماتي بالنسبة للتردد وبالتالي تعتمد النتيجة على الإشارة الداخلة أيضاً.

١٦. لقد اثبت مرمز - فاك ترميز إثارة التنبؤ الخطي الذي يعتمد على مبادئ التحليل بالتركيب ويحاول تقليل إشارة الخطأ الموزون كفاءة عالية جداً في الترميز حيث أنتج إشارة كلامية مسترجعة ذات جودة عالية عند معدل نقل بيانات (٤ kbps) برغم تغير إشارات الكلام مع الوقت وبرغم إجراء الاختبار على تسجيلات في أوقات مختلفة .

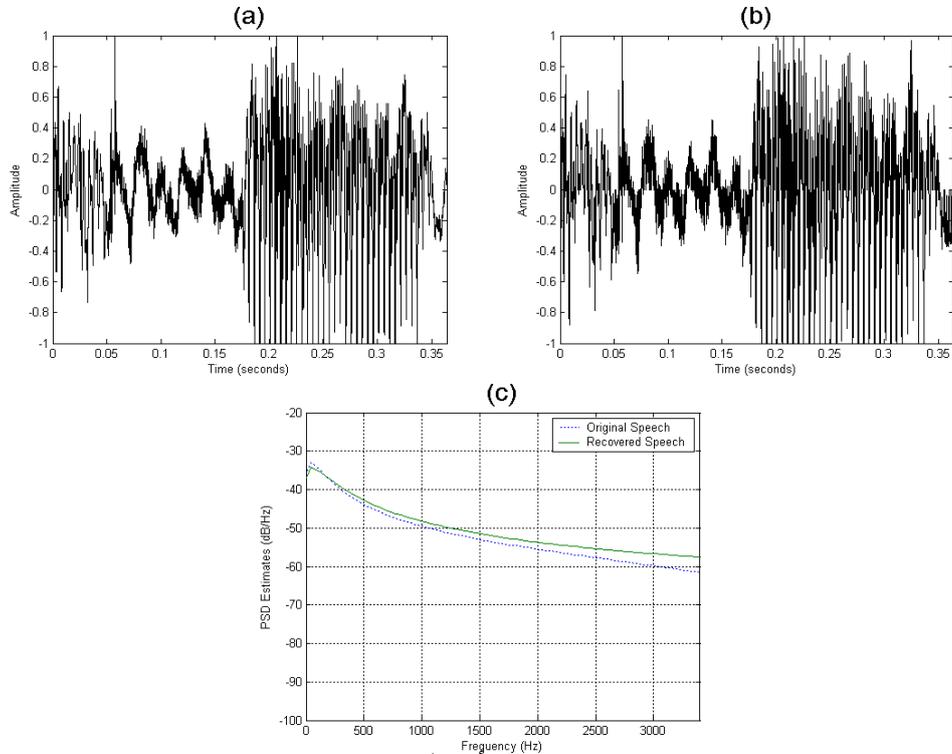
١٧. في مثل هذه الجوانب البحثية يفضل عدم الاعتماد على بنية برمجية واحدة لجميع العمليات المشمولة ضمن المسألة البحثية . واعتماداً على ما تم التوقف عنده من أدبيات سابقة وما تم من إجراءات بحثية خاصة بالباحث ، توصلنا إلى قناعة إن هذه المسألة يمكن أن تعالج باستخدام البنات البرمجية الآتية :

- بنية الـ (MATLAB) لإنجاز الجوانب الخاصة بالتحويل من الملف الصوتي ذي الامتداد (wav) إلى ملف صوتي ذي امتداد (dat) ، وذلك للدقة العالية التي توفرها هذه البنية البرمجية هنا .
- استخدام الـ (Visual C++) في عمليات المعالجة الخاصة بترميز الإشارة الكلامية .
- استخدام بنية الـ (MATLAB) في عمليات الاسترجاع والتقويم .

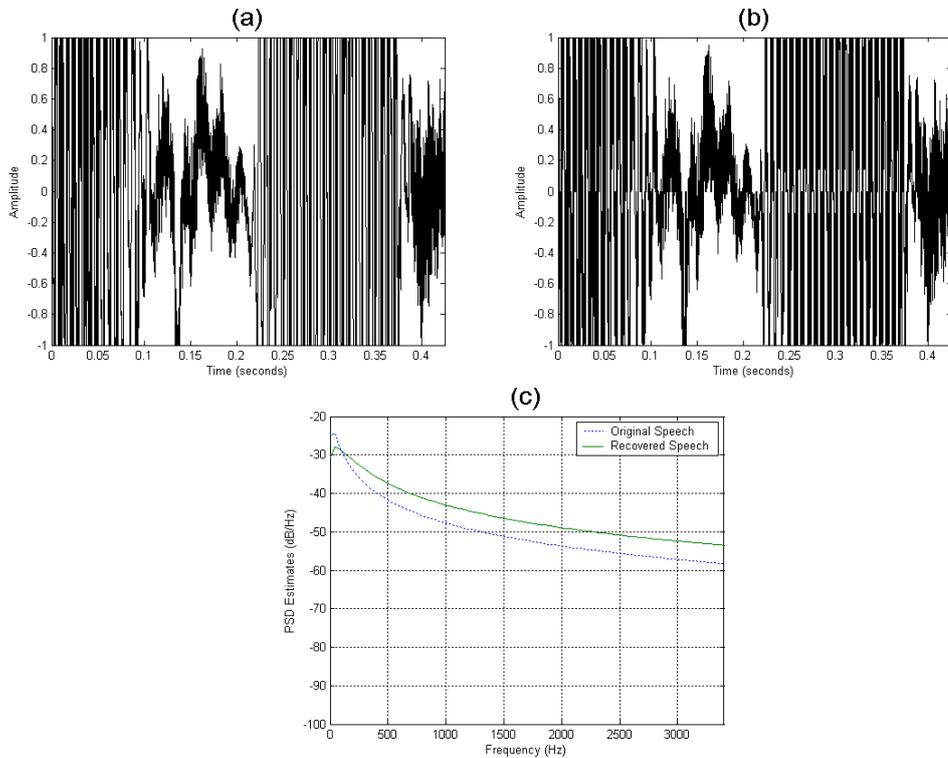
٣.٥ الأعمال المستقبلية

١. استخدام المنطق المضرب لتحسين المقاييس الحسية وتحويلها إلى تقاويم عديدة بدلاً من تقاويم حسية .
٢. استخدام الشبكات العصبية كتقنية لتمييز أو تحديد الأنواع المختلفة من الرموز بالاعتماد على درجة تقويمها .
٣. إن مسألة اختيار النافذة يمكن أن تخضع إلى بحث منفصل من خلال اتجاهين :
الاتجاه الأول : استخدام الخوارزميات الجينية لاختيار أفضل نافذة .
الاتجاه الثاني : استخدام الشبكات العصبية لمقطع معين من المقاطع لاختيار أفضل نافذة.
٤. يمكن تحسين كفاءة الترميز وتوسيعها باتجاه تطبيقات الإشارات غير الكلامية بواسطة دمج تقنيات مجال التردد خاصة لترميز الإشارات واسعة النطاق .

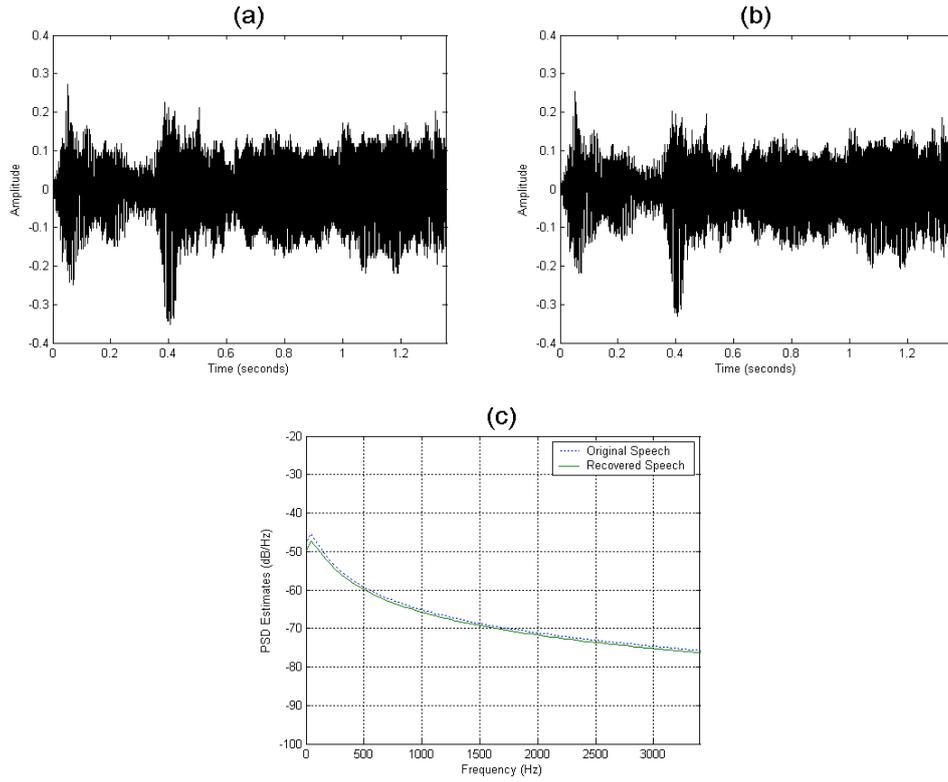
ملحق (A)



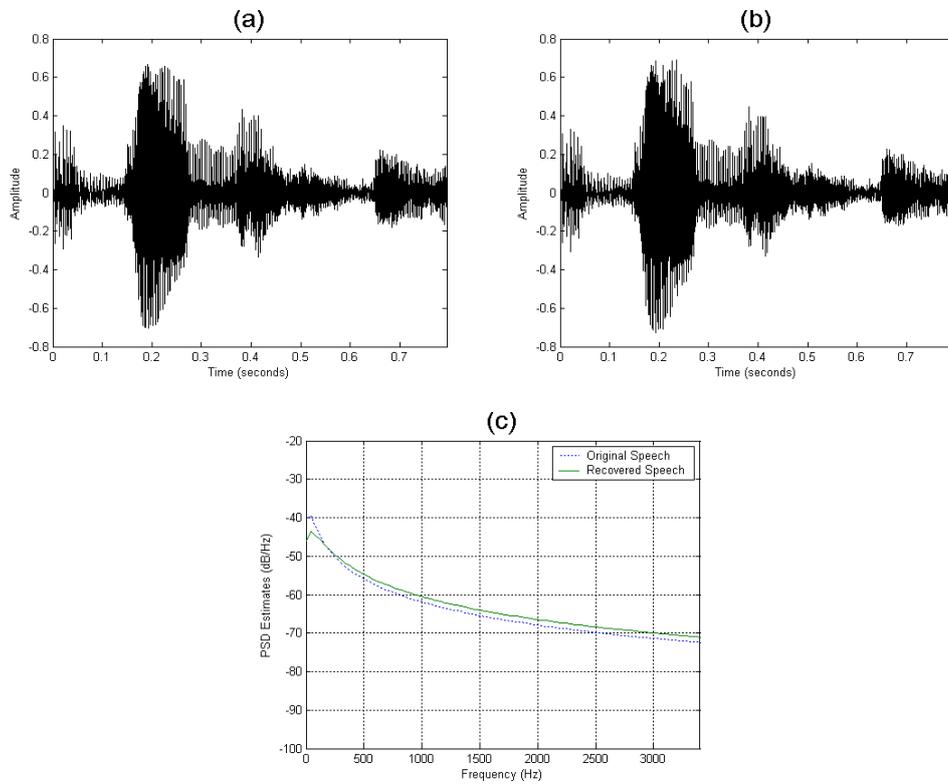
(١) امرأة (الخير)



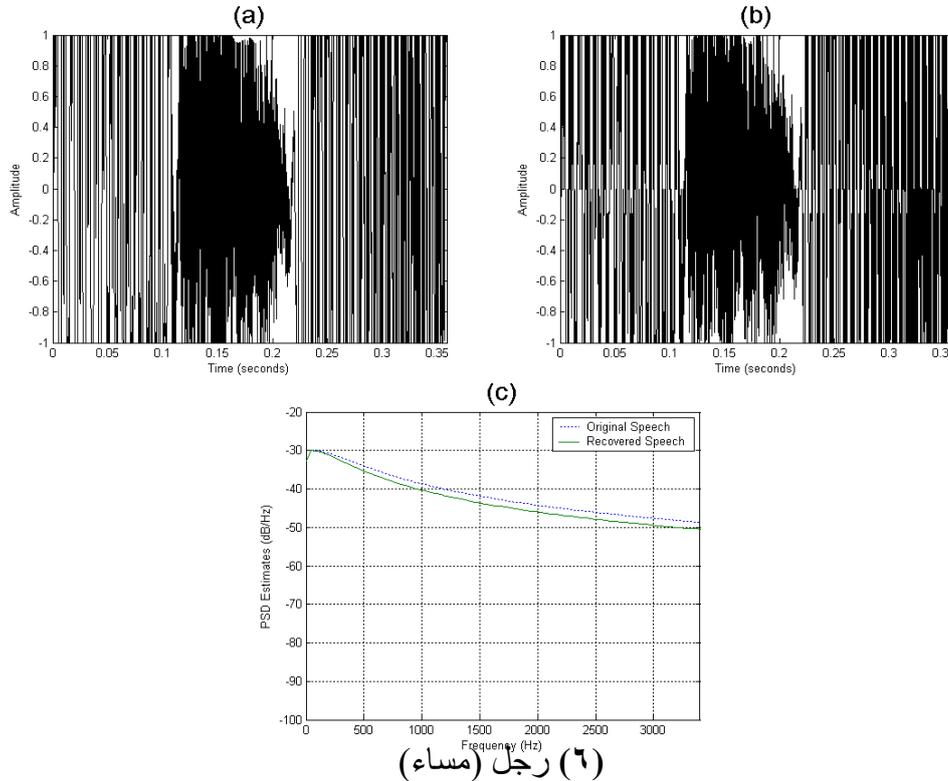
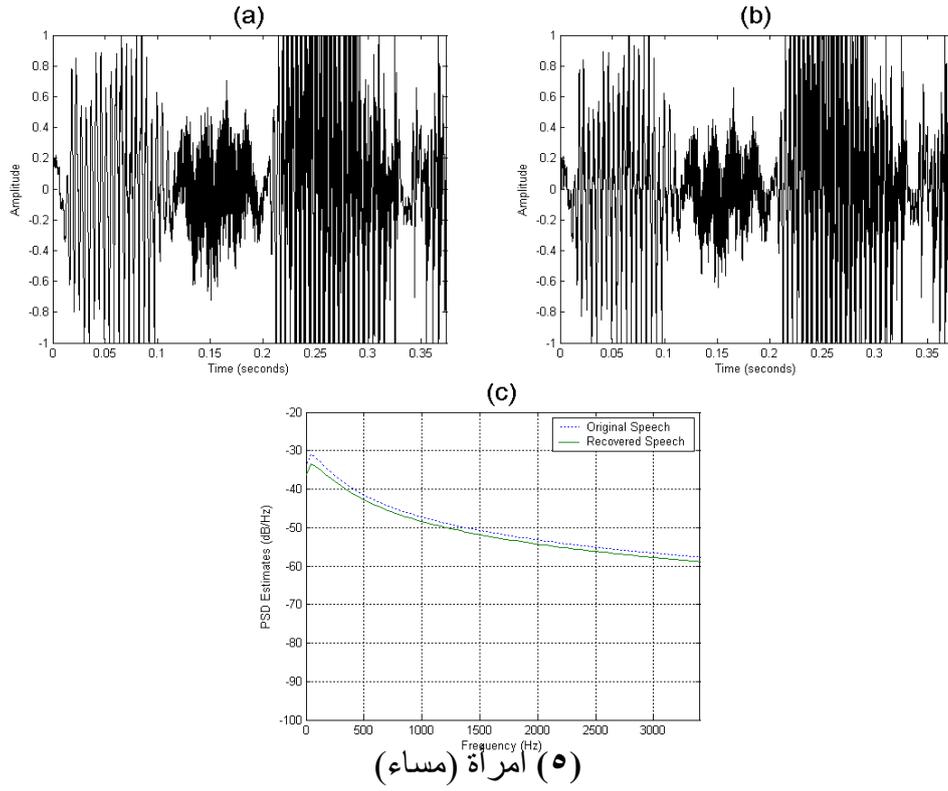
(٢) رجل (الخير)



(٣) رجل (الحمد لله)



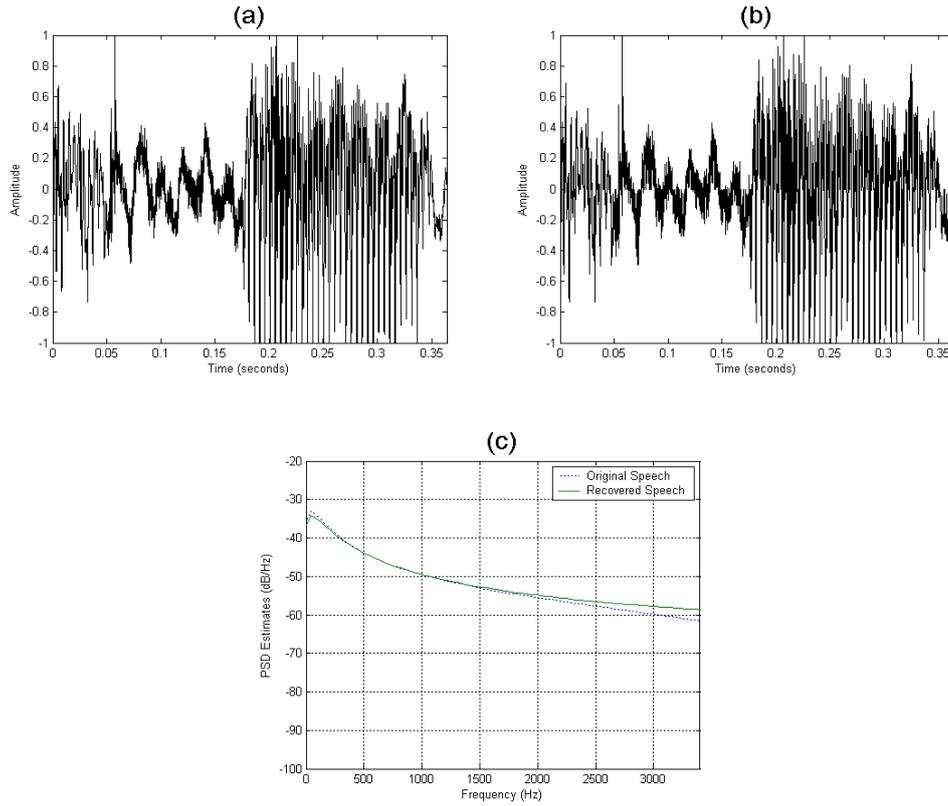
(٤) رجل (الله أكبر)



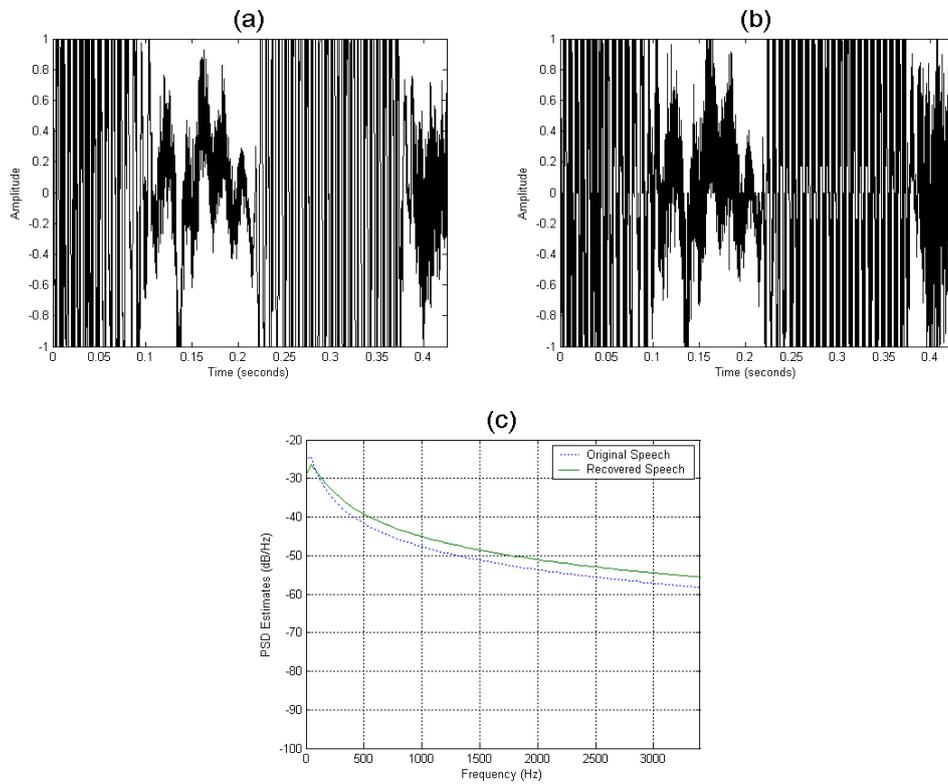
شكل (1.A) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (44100 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

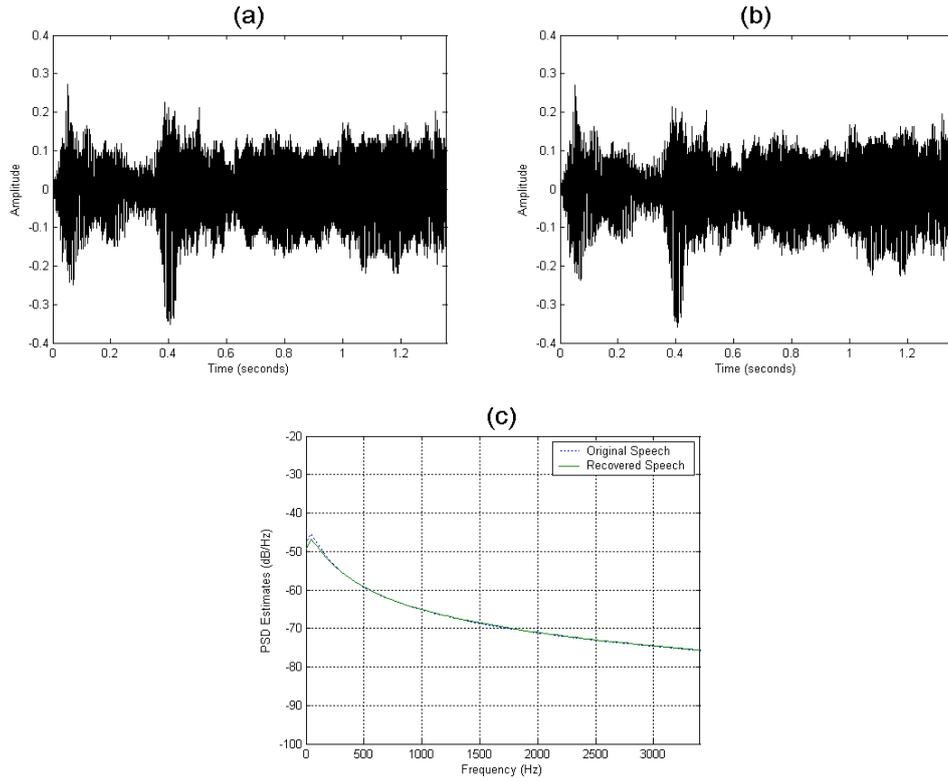
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.



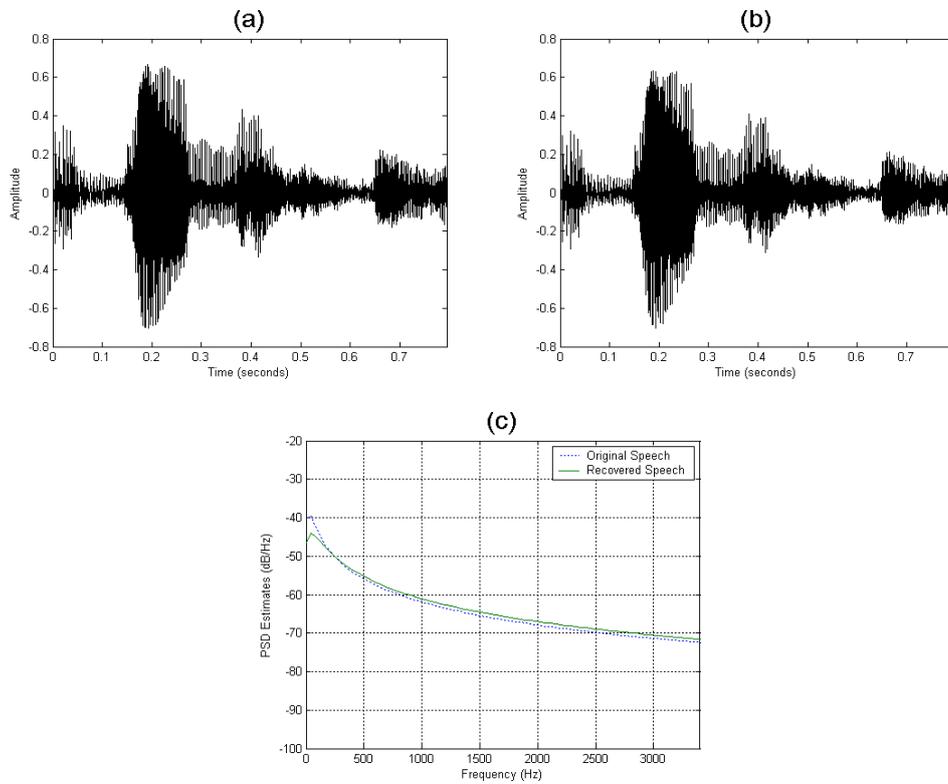
(١) امراة (الخير)



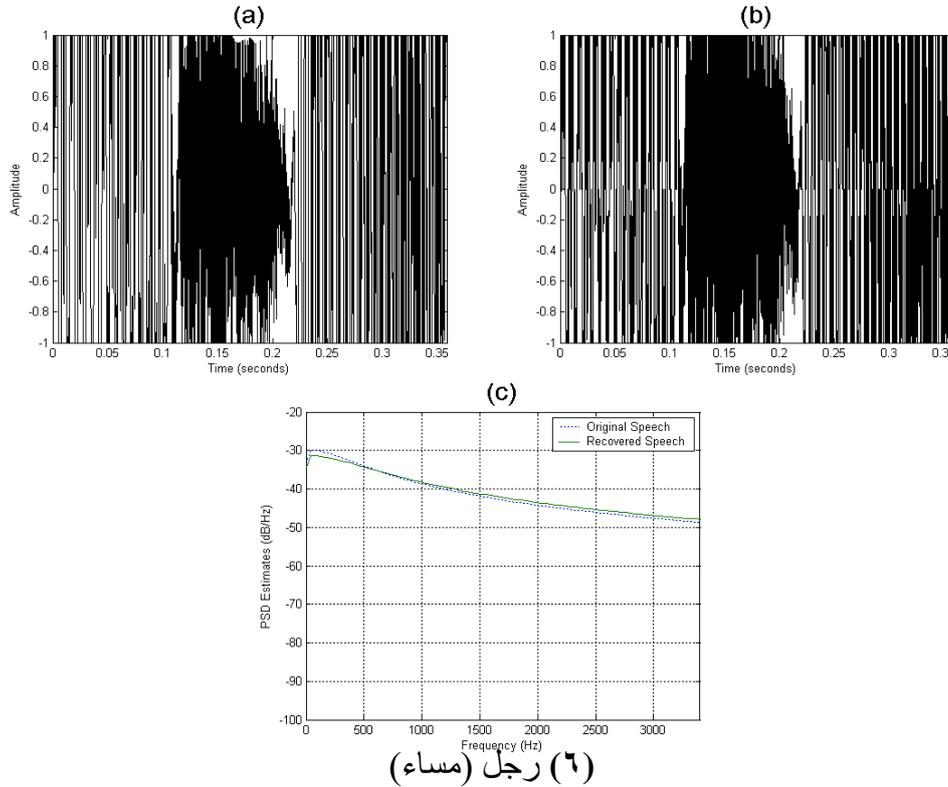
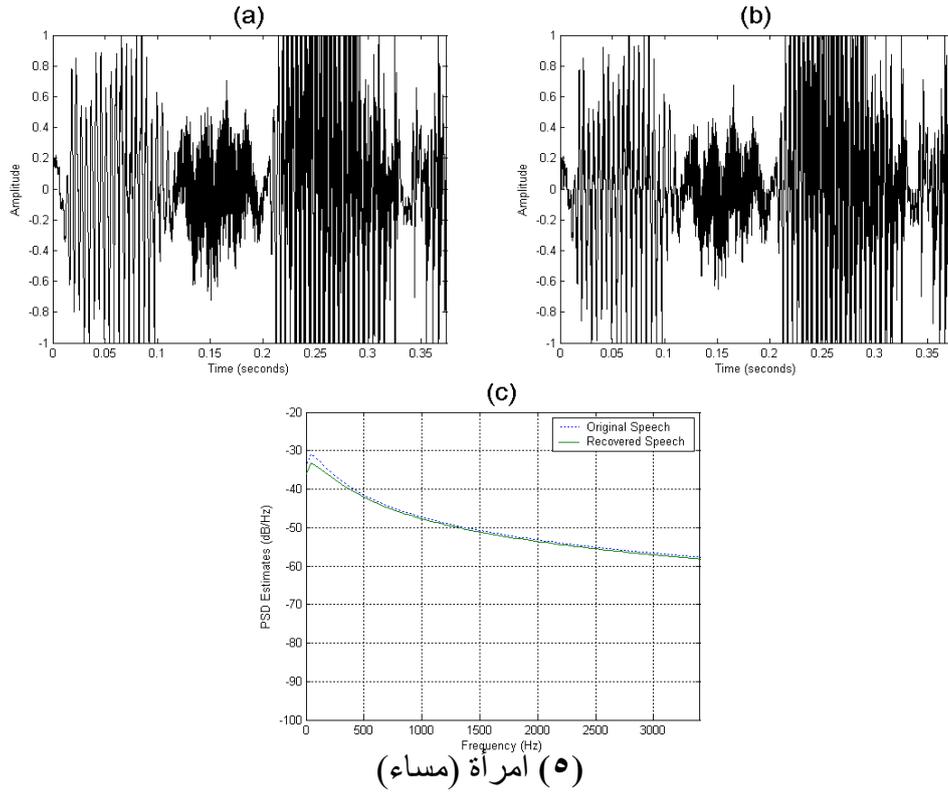
(٢) رجل (الخير)



(٣) رجل (الحمد لله)

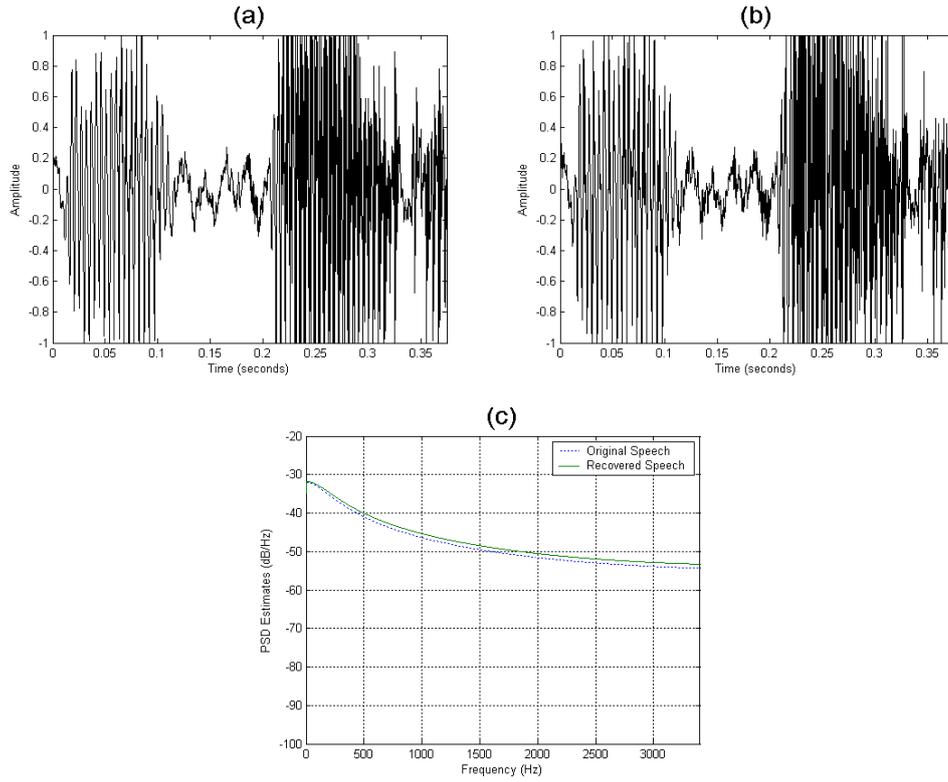


(٤) رجل (الله أكبر)

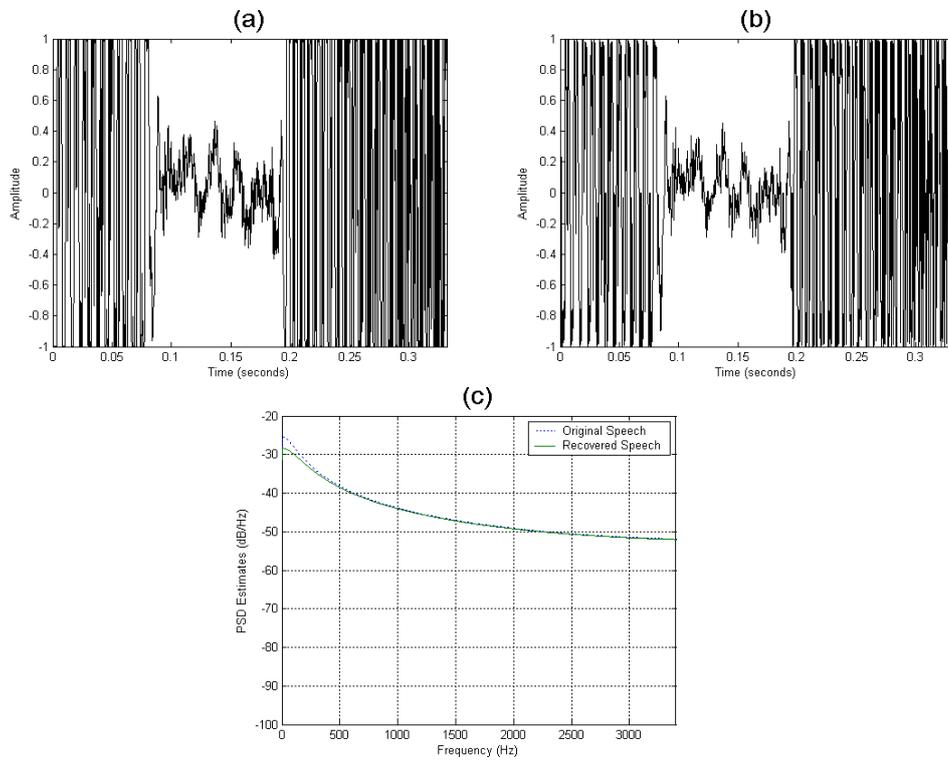


شكل (٢.A) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (44100 Hz) :

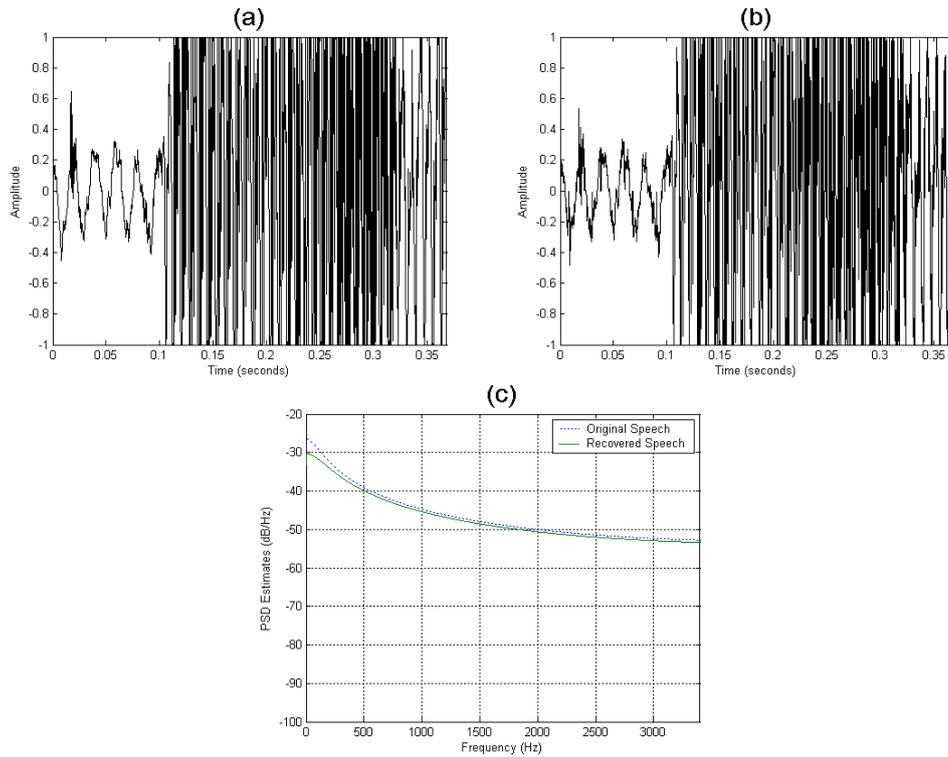
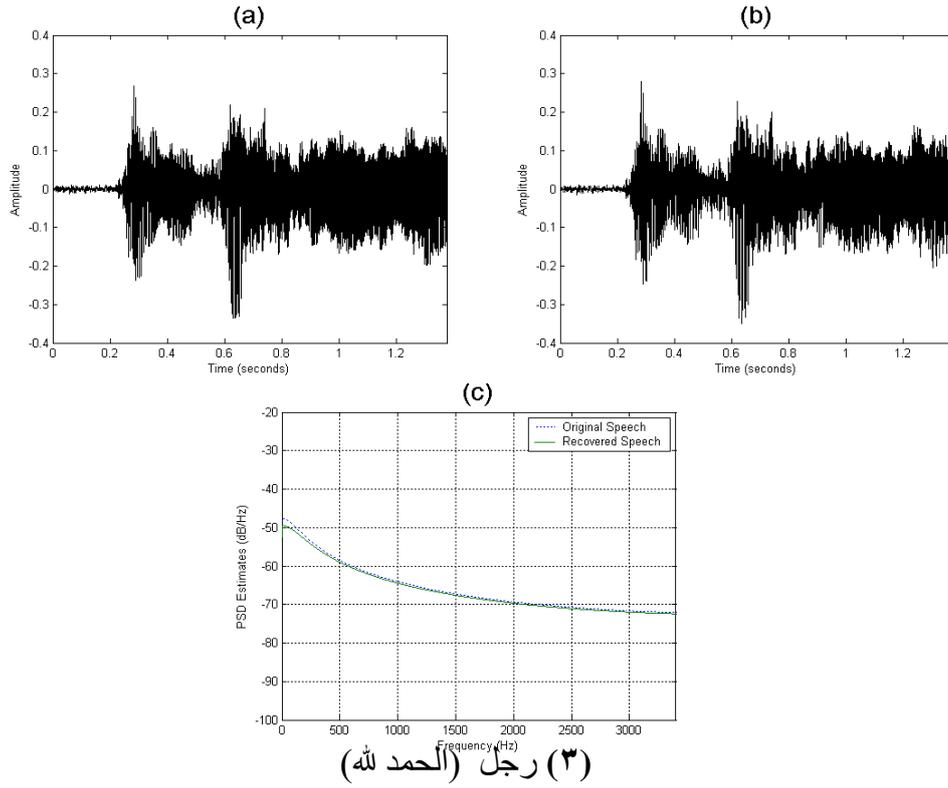
a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.
 c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.
 والمسترجعة.



(١) امرأة (مساء)

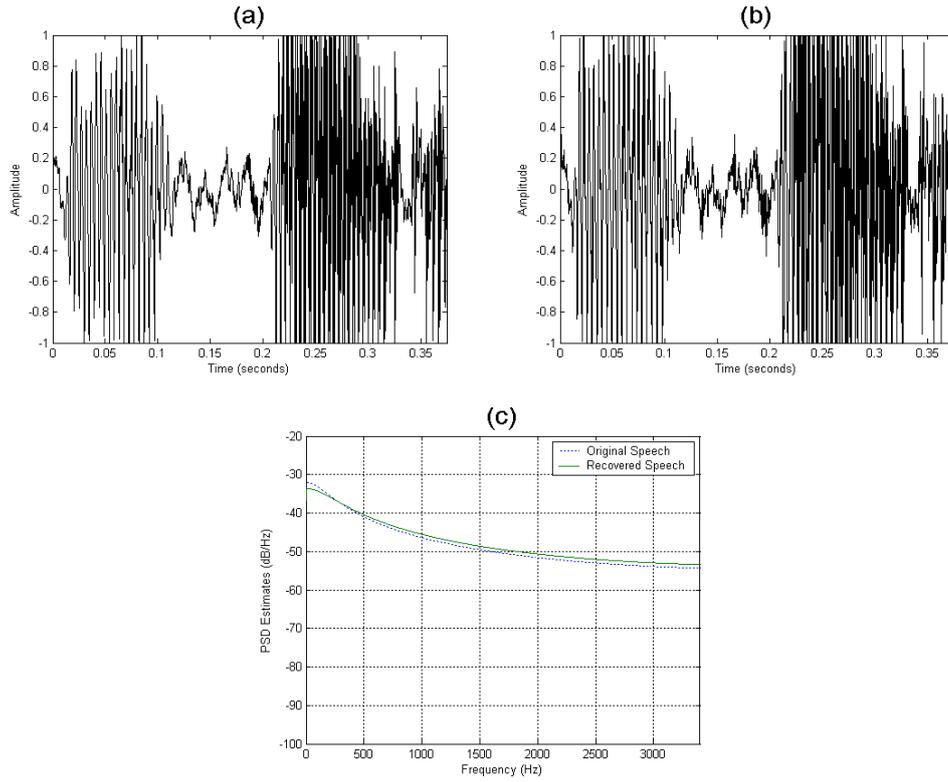


(٢) رجل (مساء)

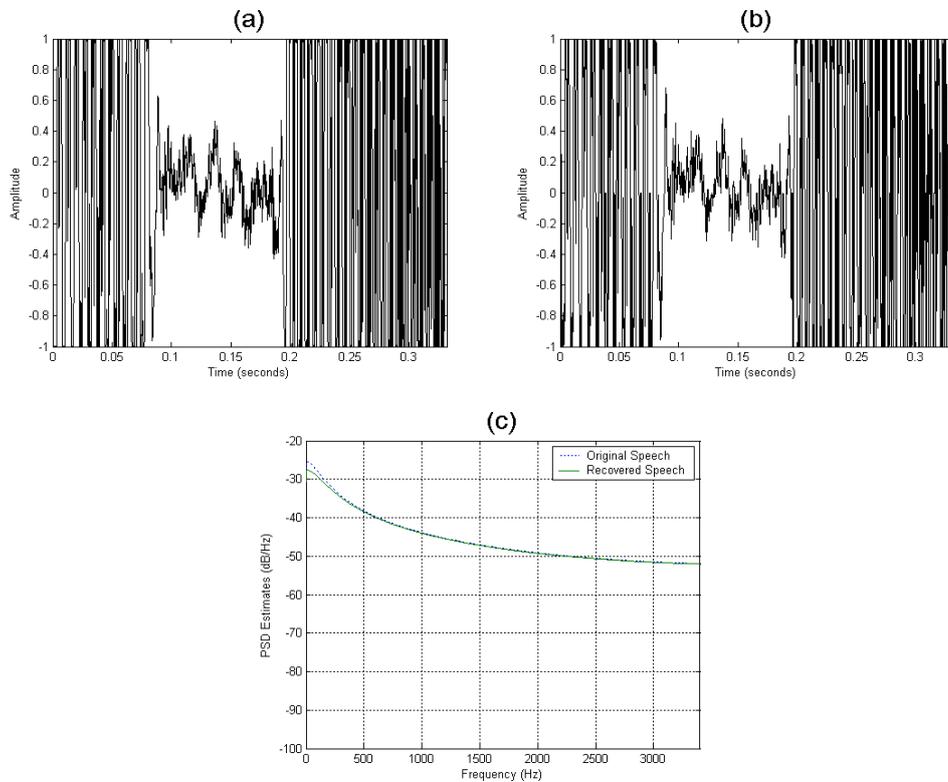


شكل (٣.A) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (8000 Hz) :

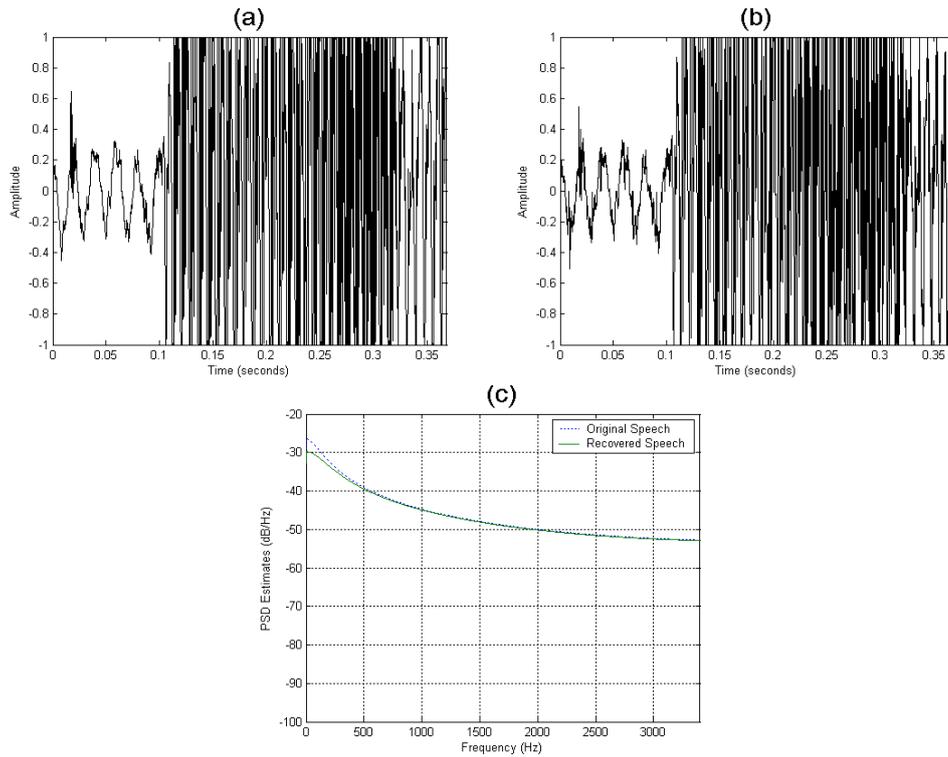
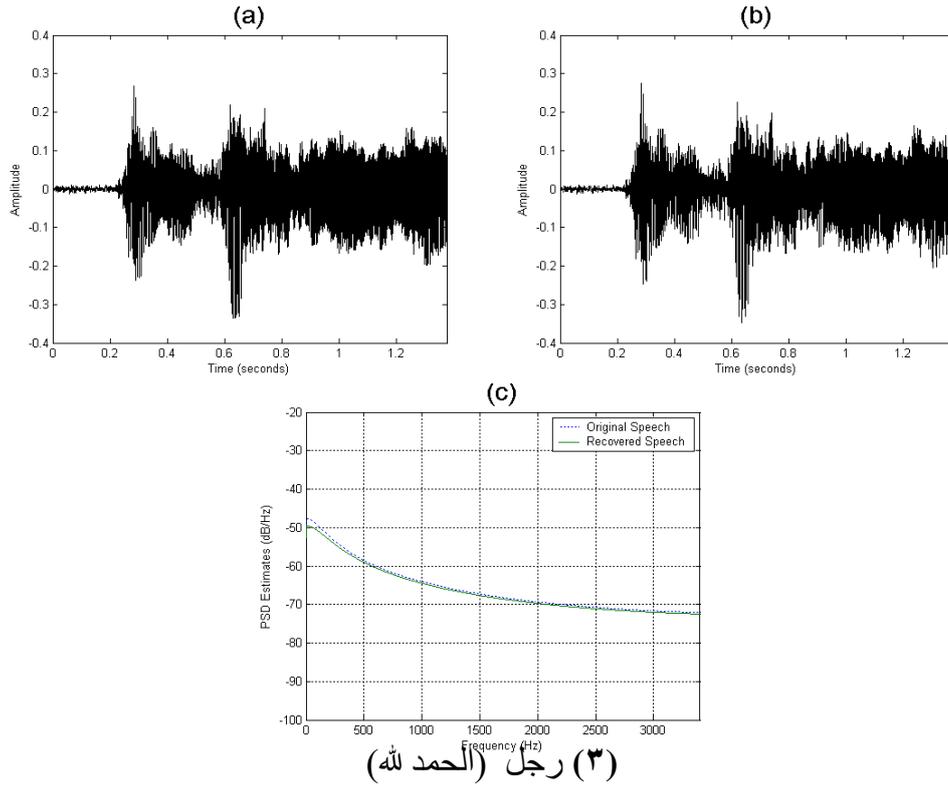
- a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة. c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.



(١) امرأة (مساء)



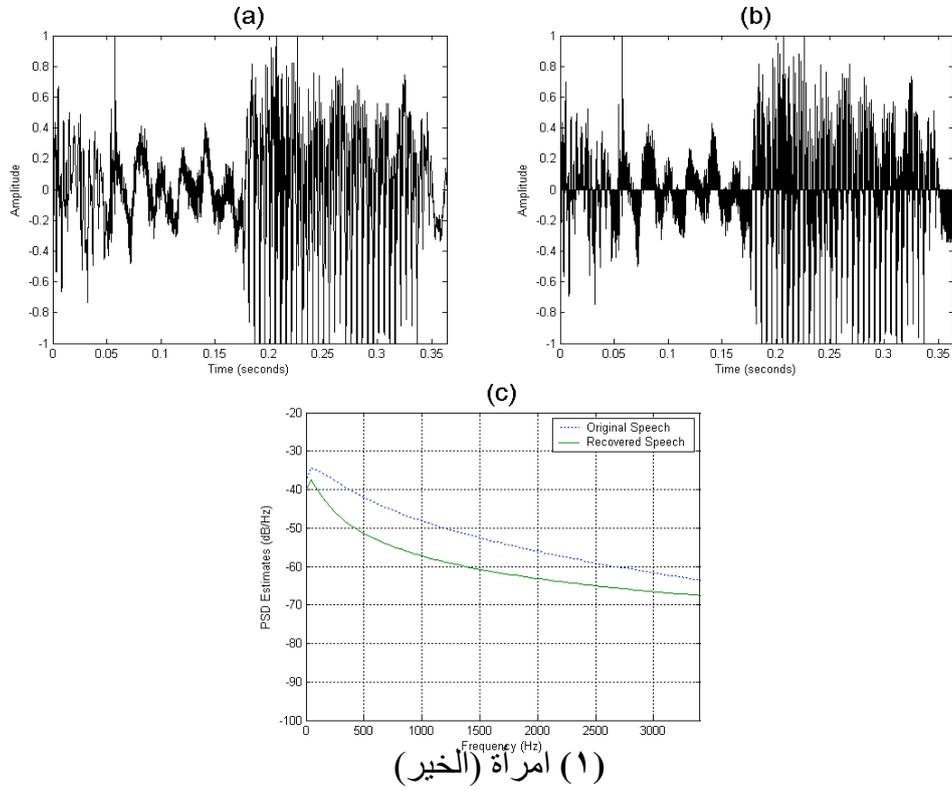
(٢) رجل (مساء)



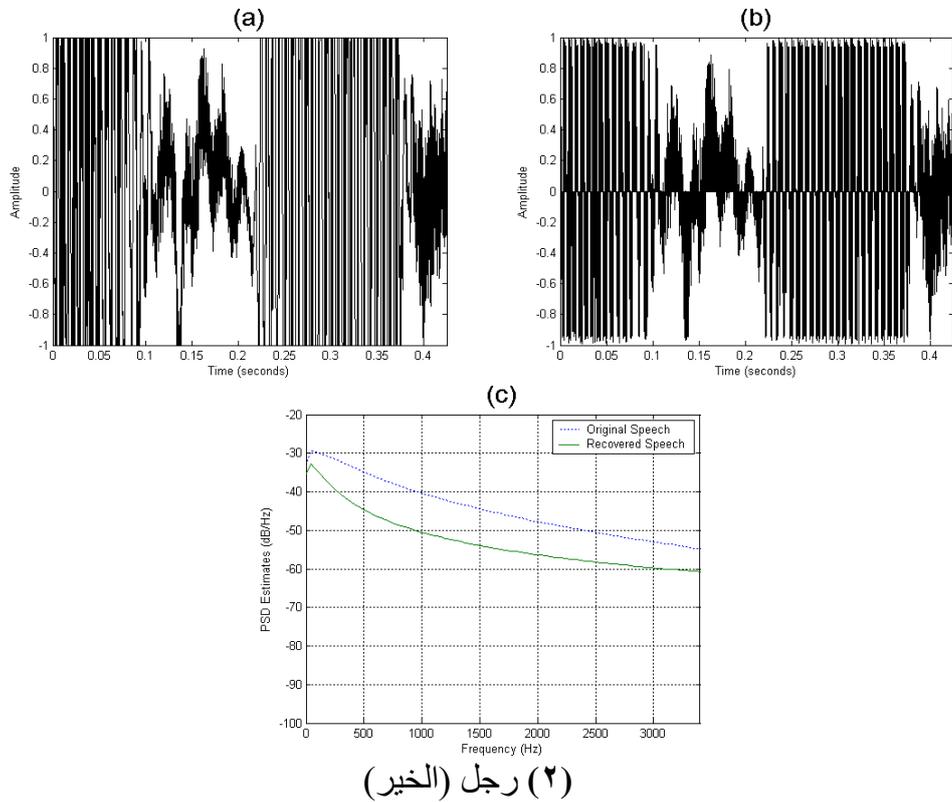
شكل (٤.A) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (8000 Hz) :

- a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.

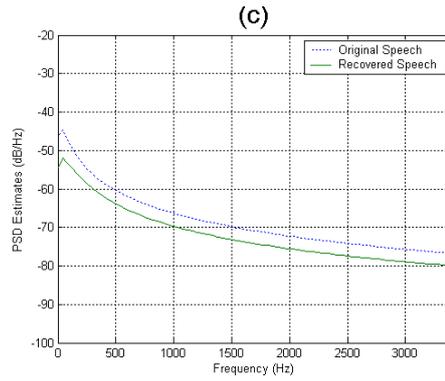
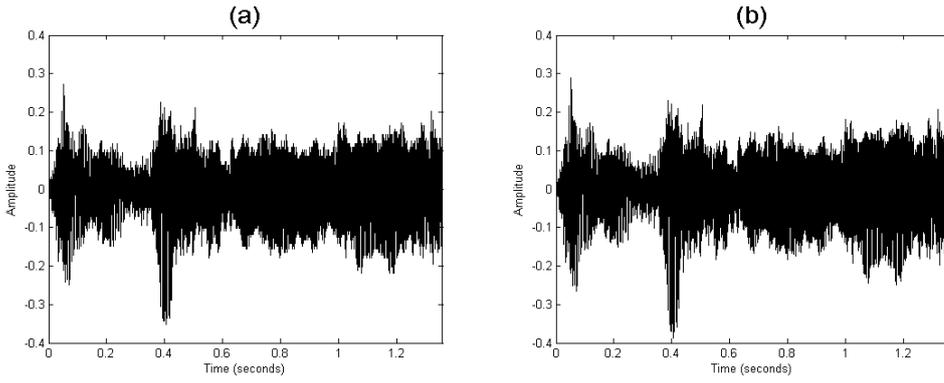
ملحق (B)



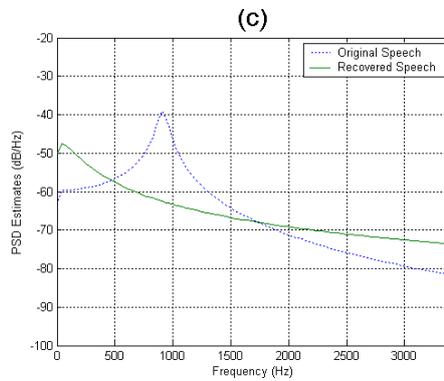
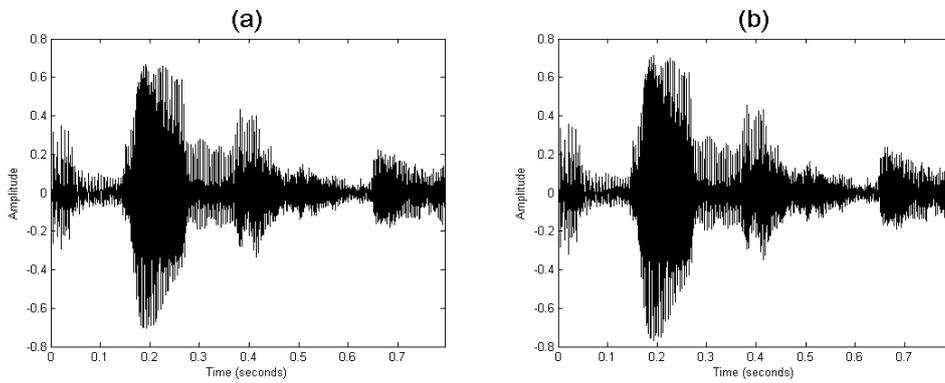
(1) امرأة (الخير)



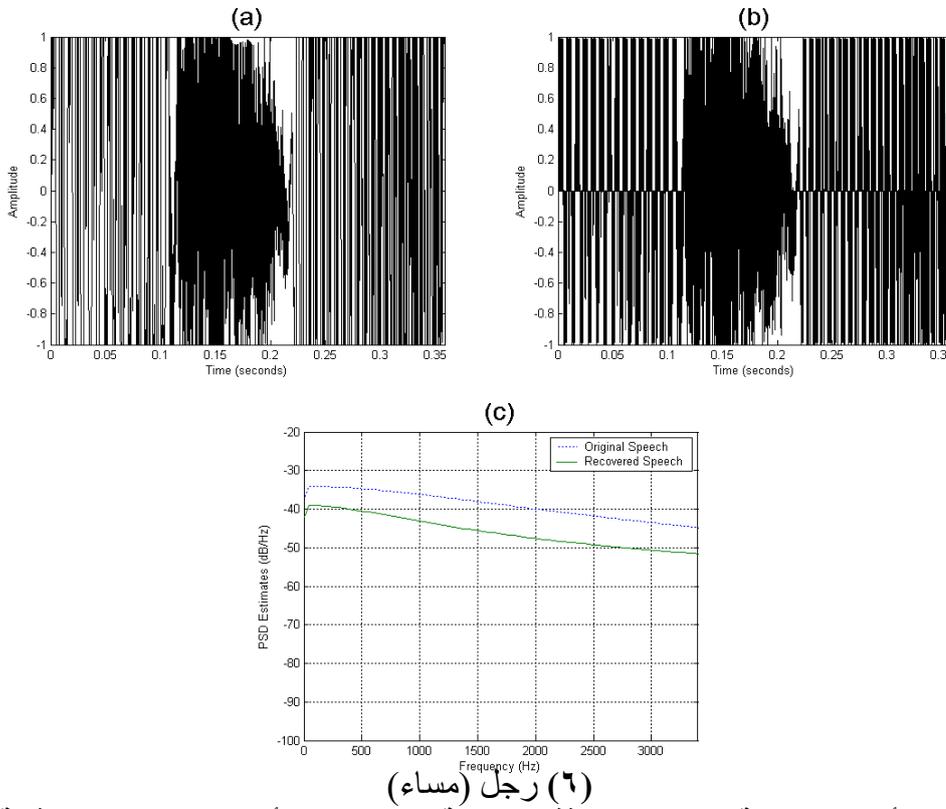
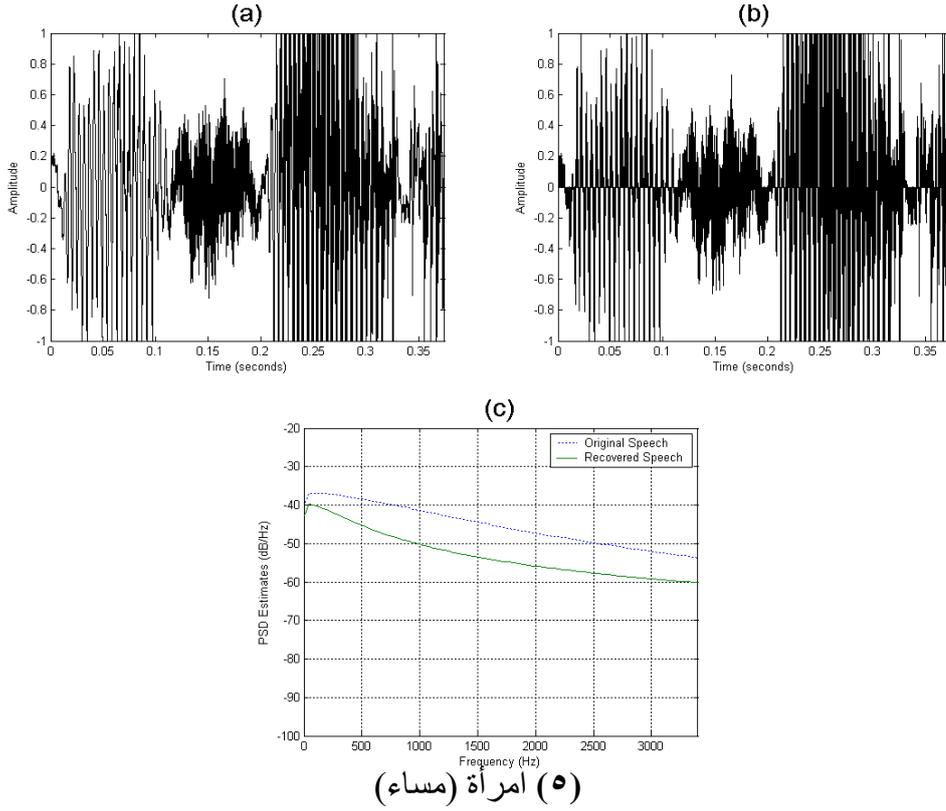
(2) رجل (الخير)



(٣) رجل (الحمد لله)

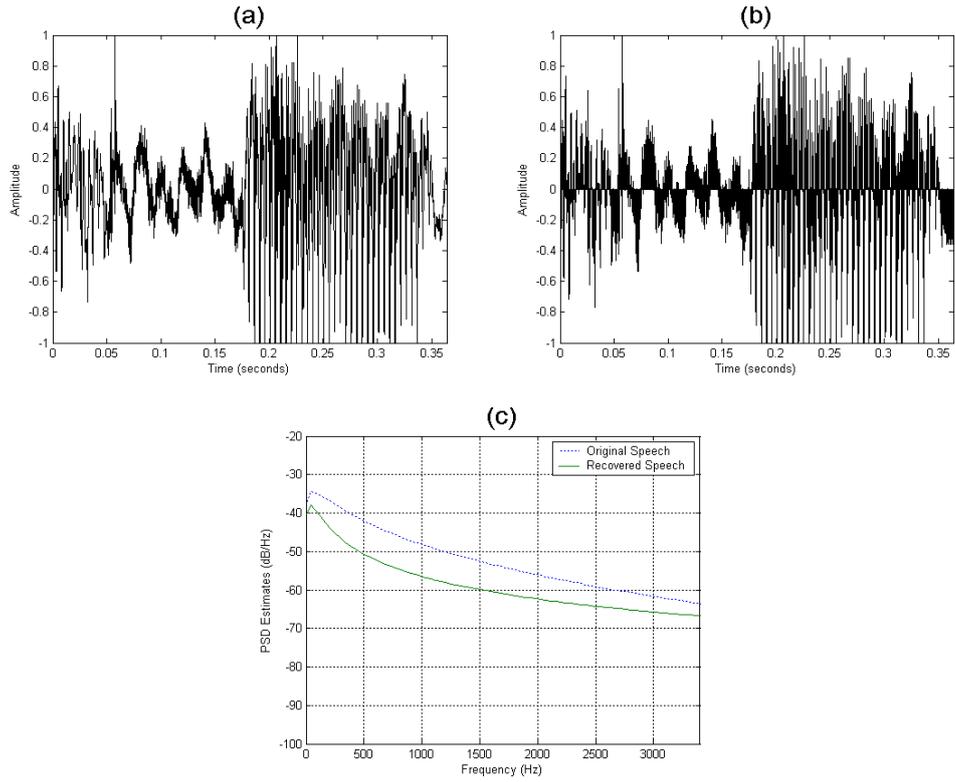


(٤) رجل (الله أكبر)

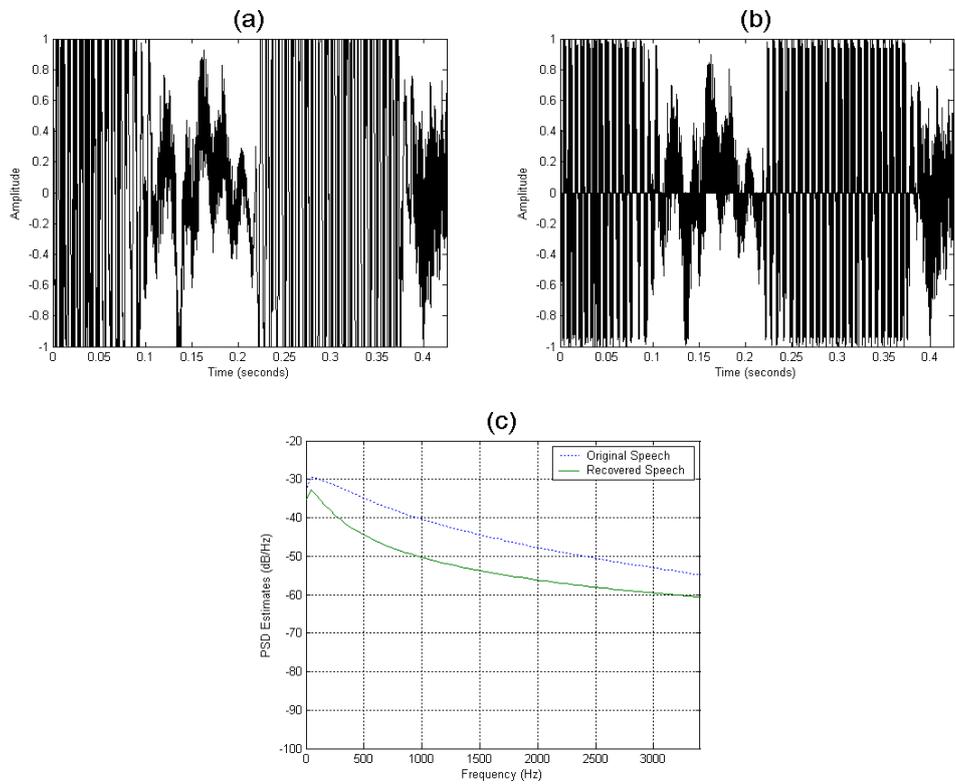


شكل (١.B) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (44100 Hz) :

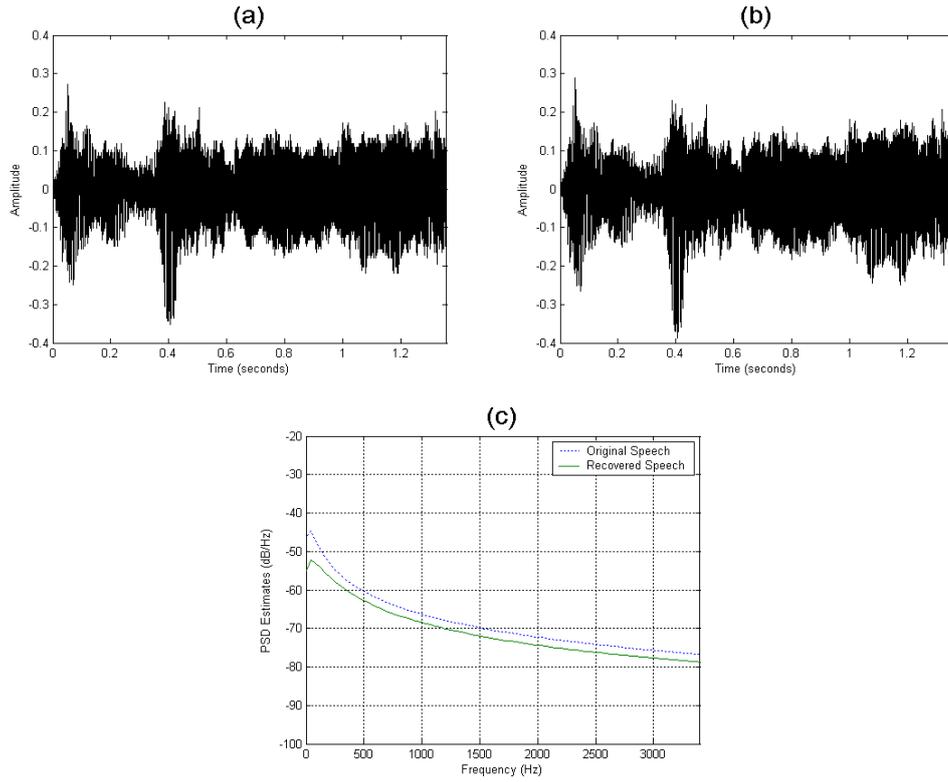
- a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.
- c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.



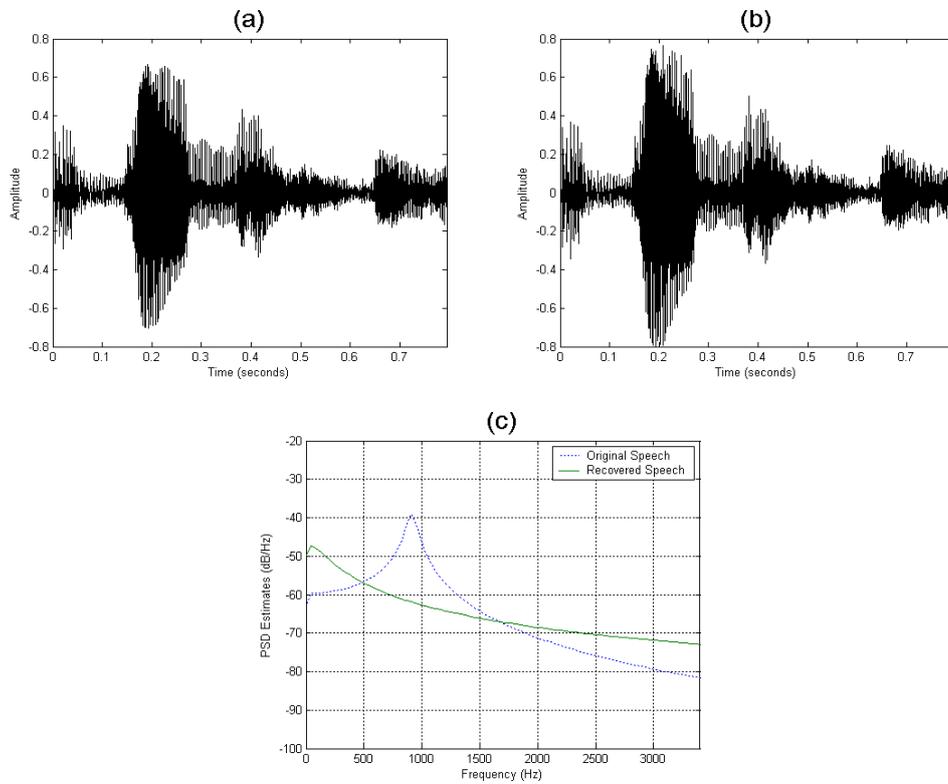
(١) امرأة (الخير)



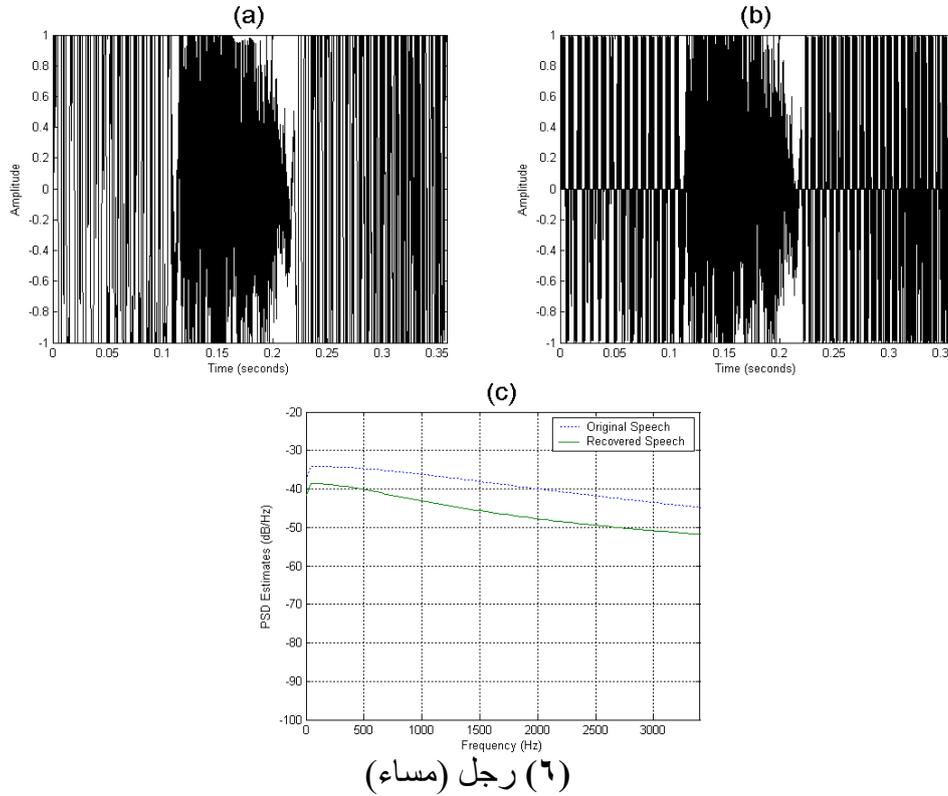
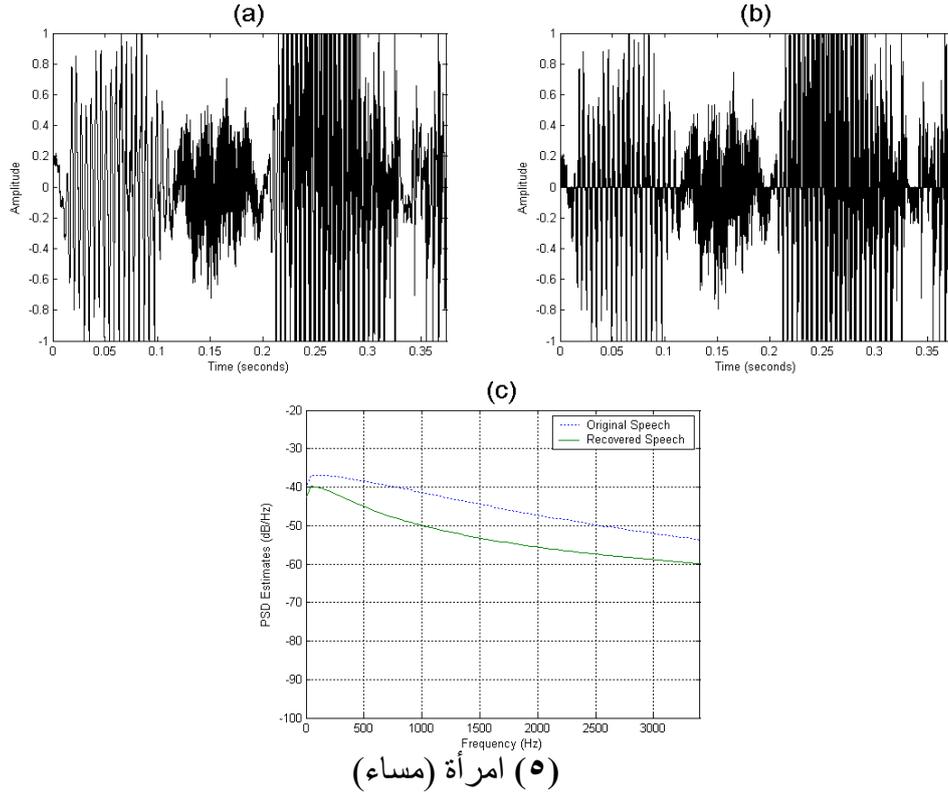
(٢) رجل (الخير)



(٣) رجل (الحمد لله)



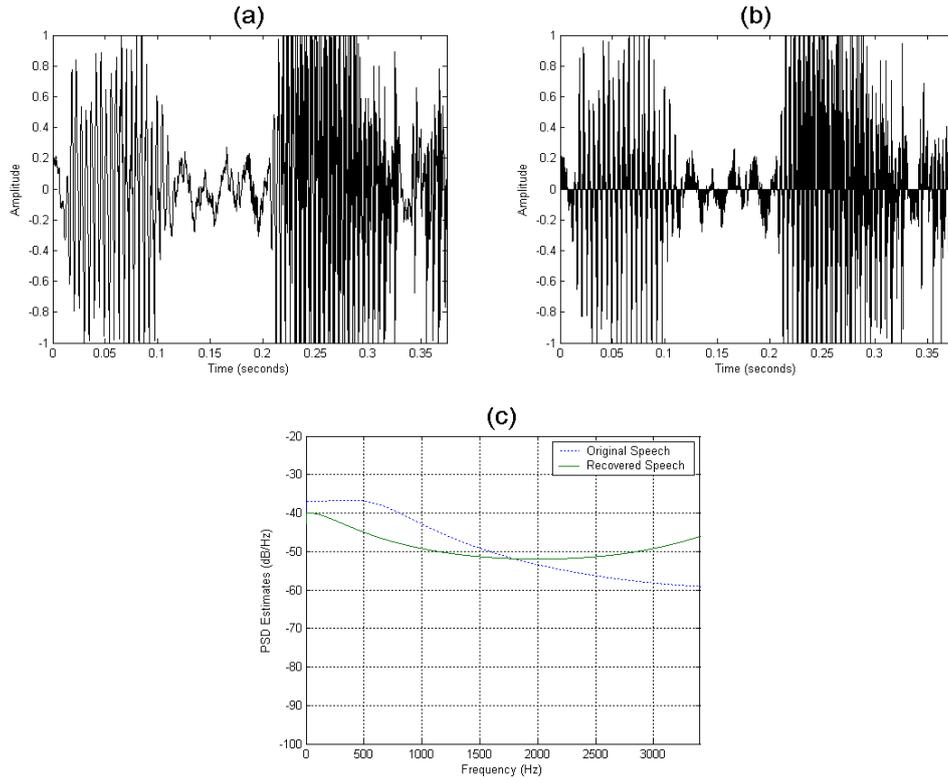
(٤) رجل (الله أكبر)



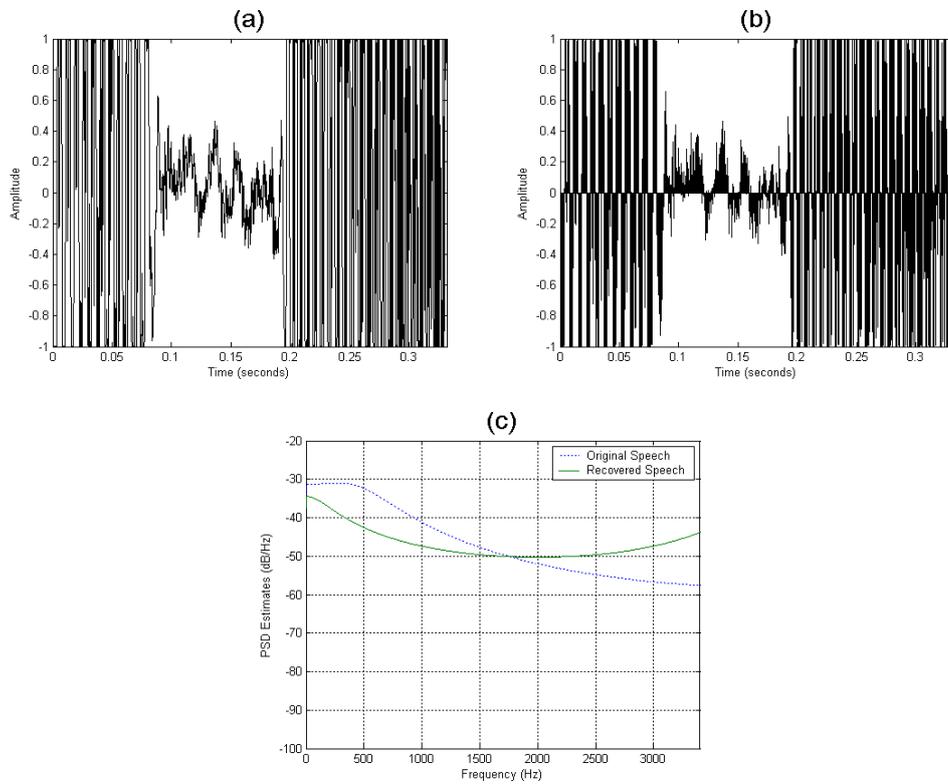
شكل (2.B) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (44100 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

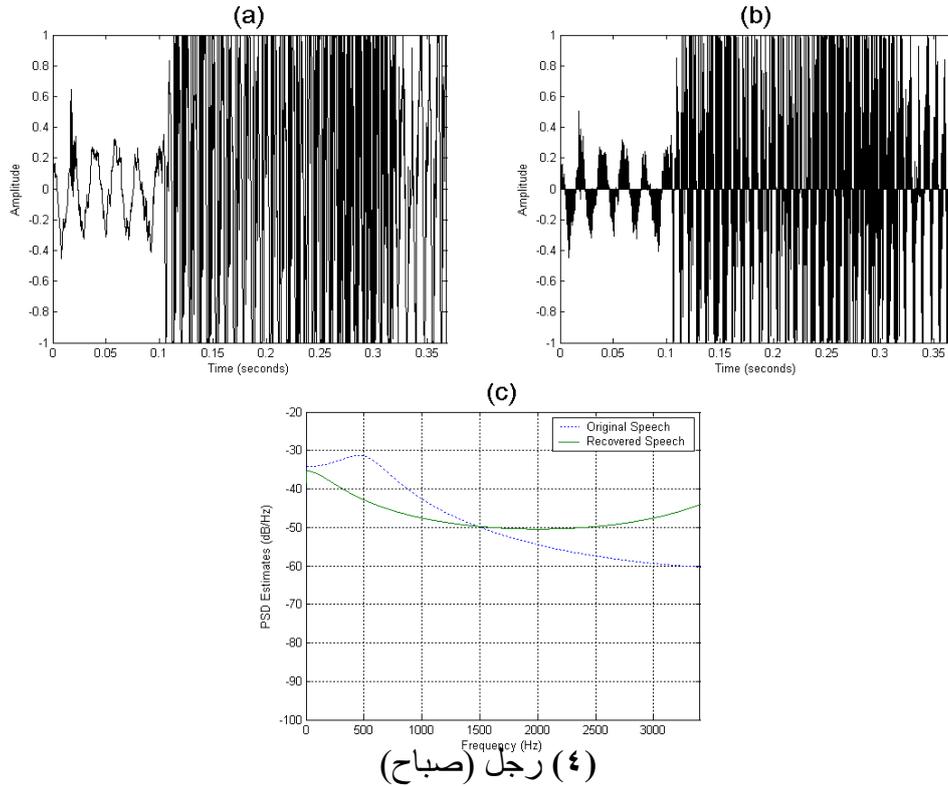
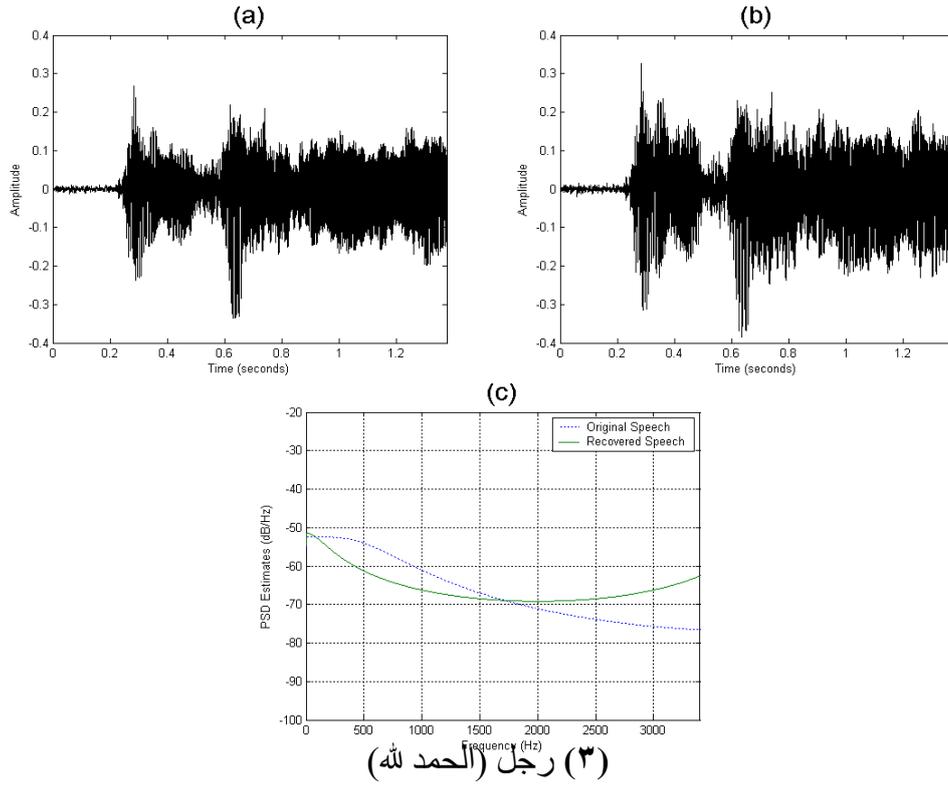
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.



(١) امرأة (مساء)



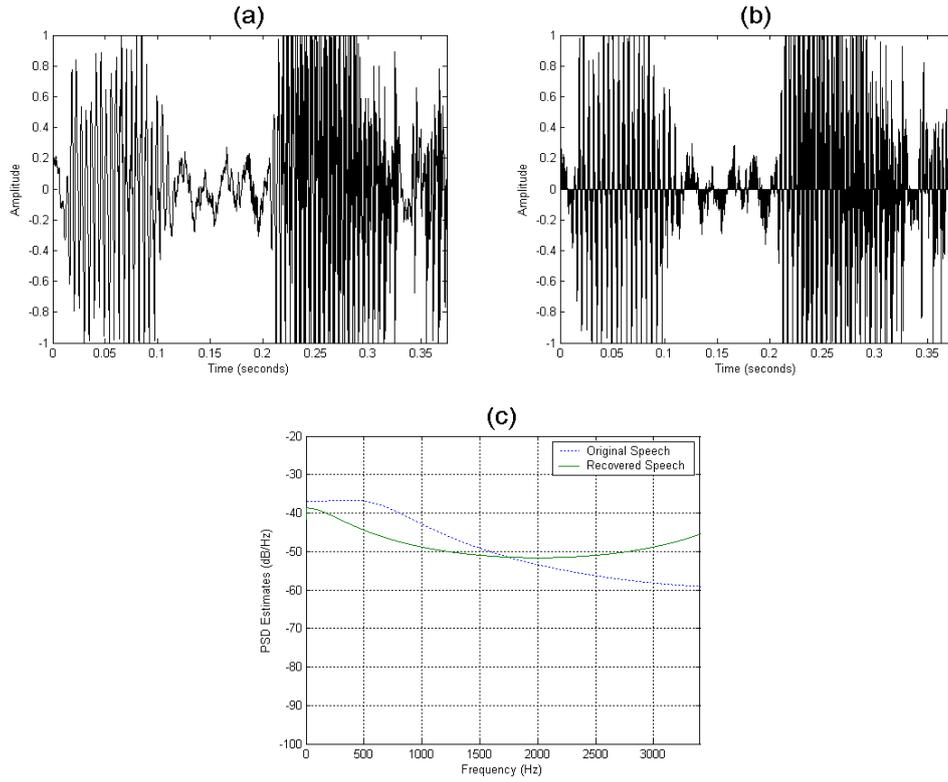
(٢) رجل (مساء)



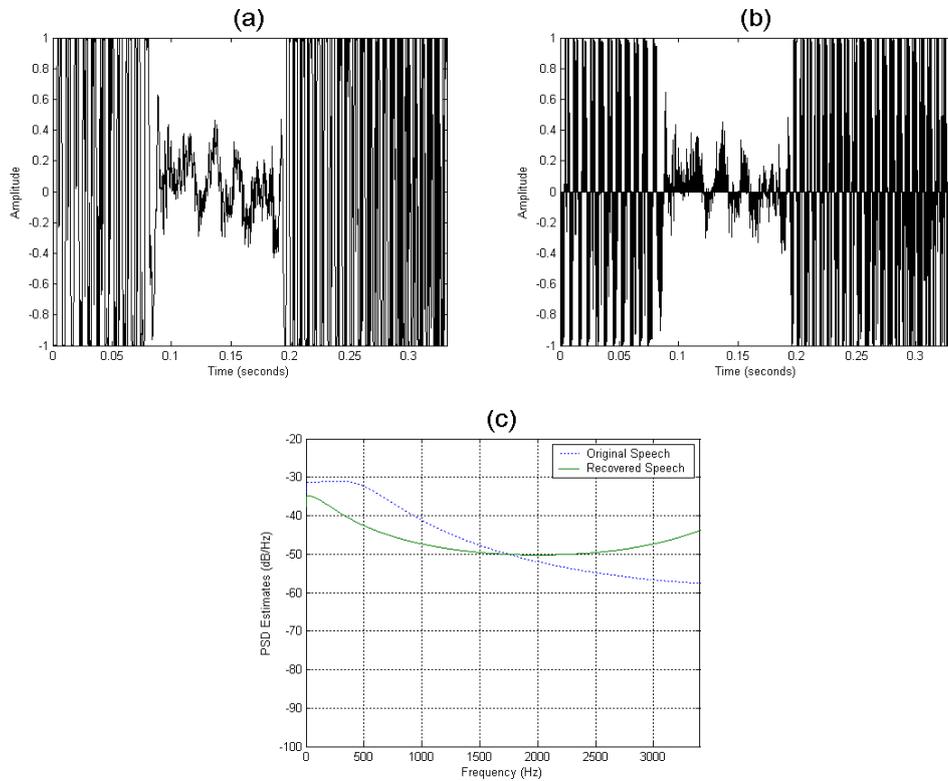
شكل (٣.B) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (8000 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

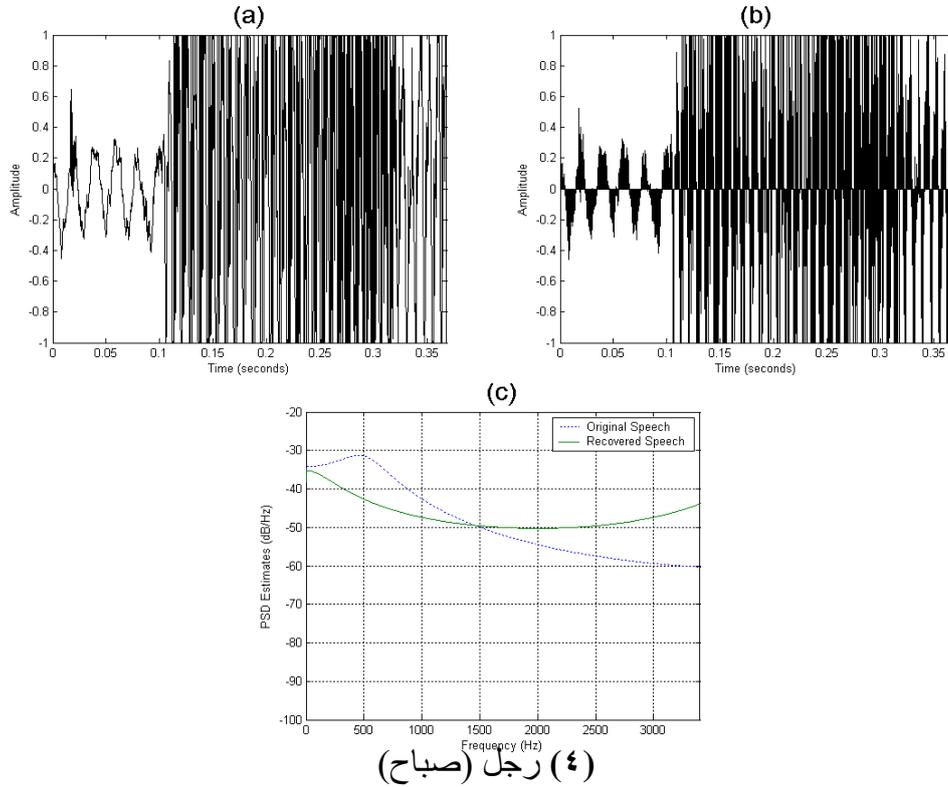
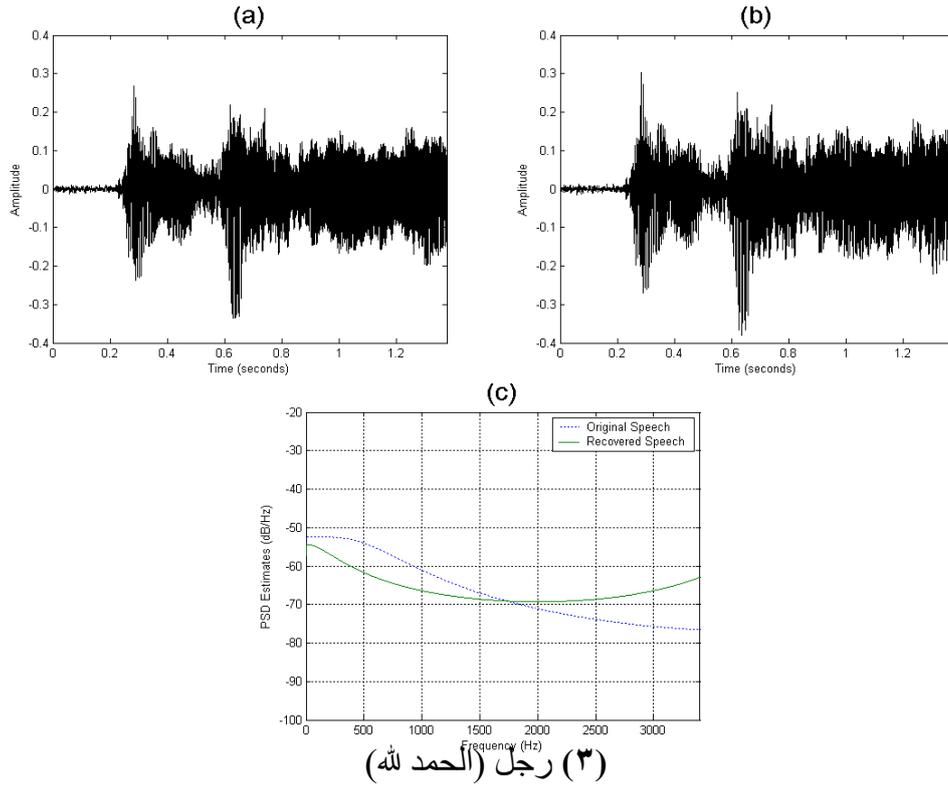
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة.



(١) امرأة (مساء)



(٢) رجل (مساء)

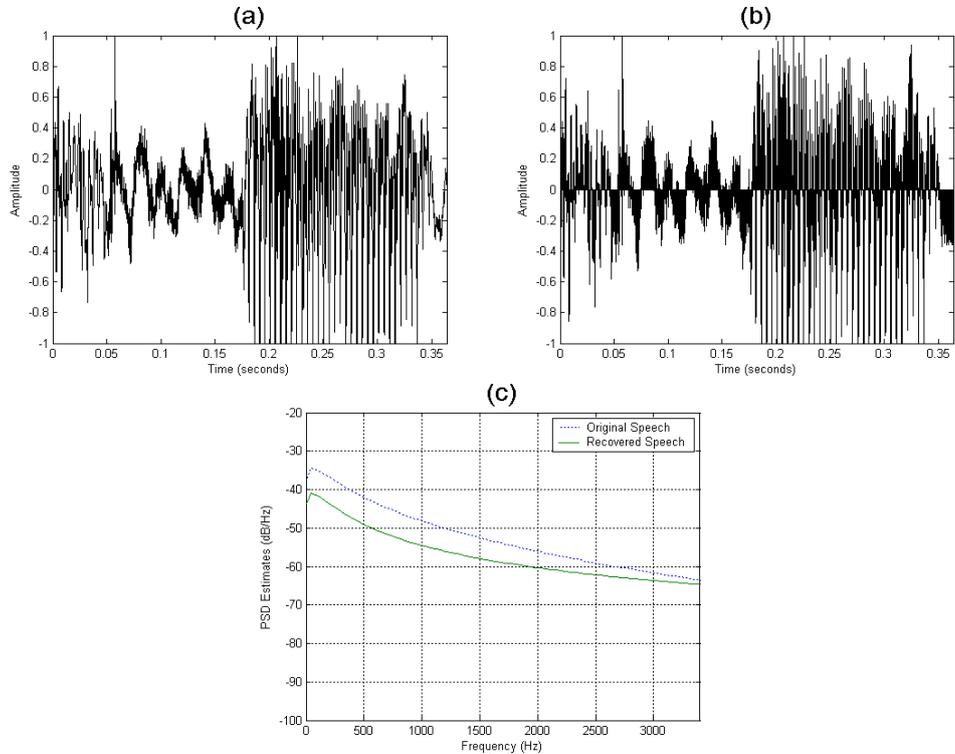


شكل (٤.B) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (8000 Hz) :

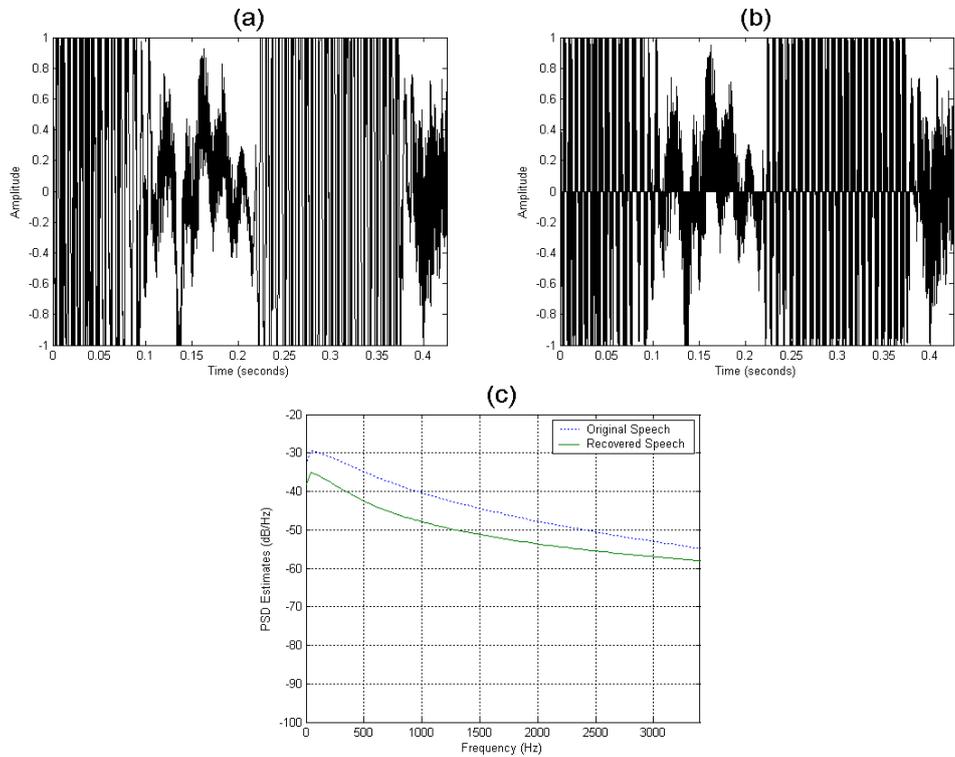
a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.

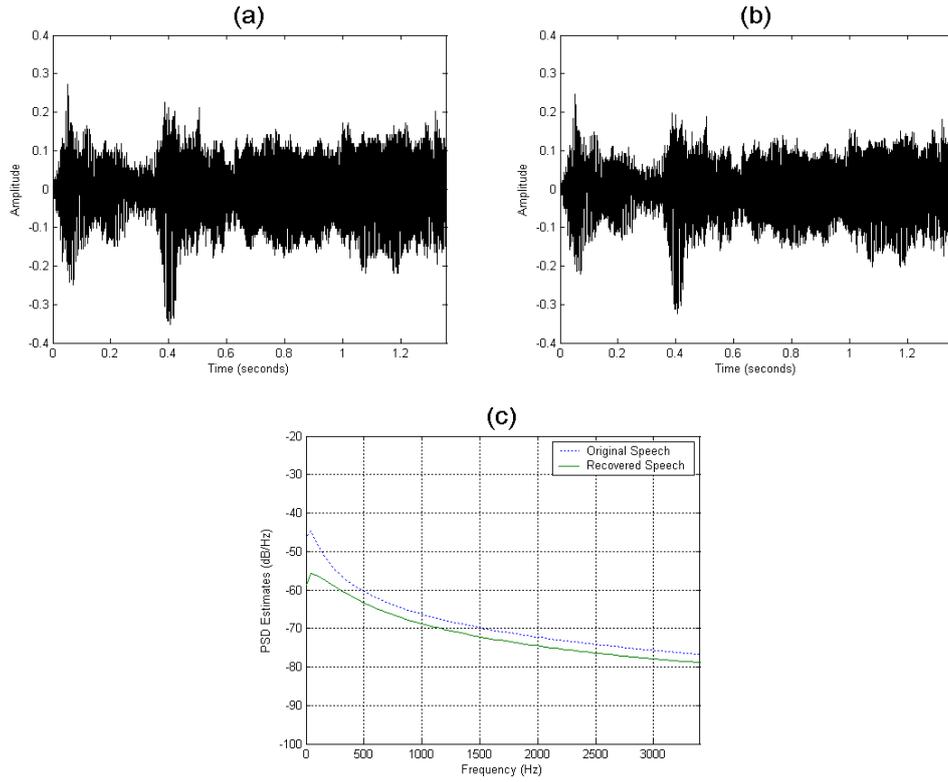
ملحق (C)



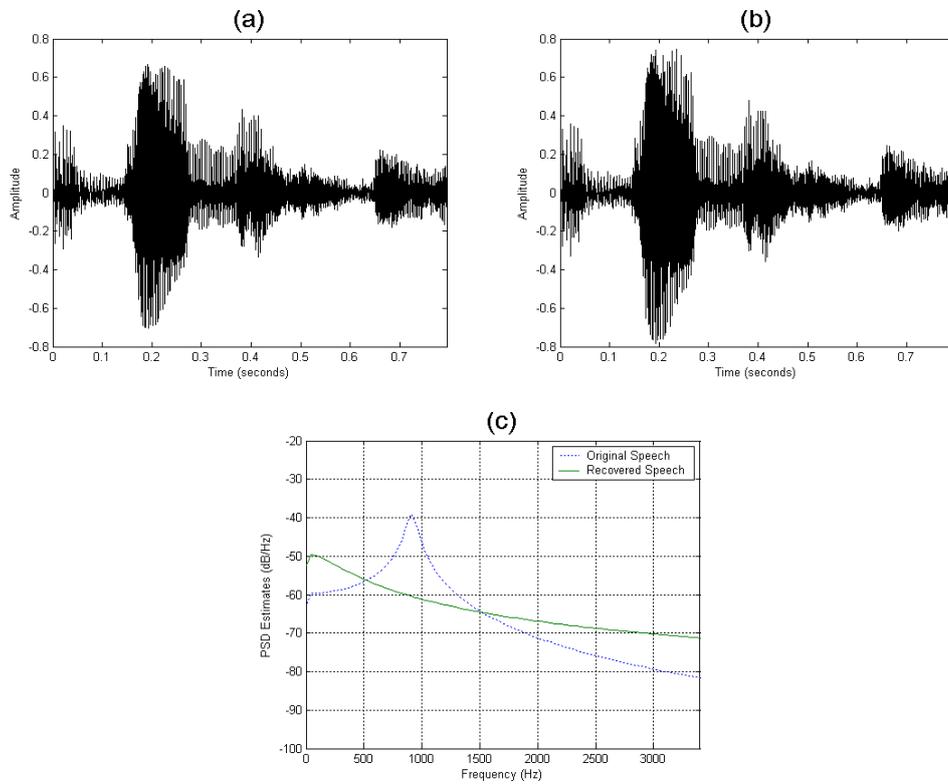
(١) امراة (الخير)



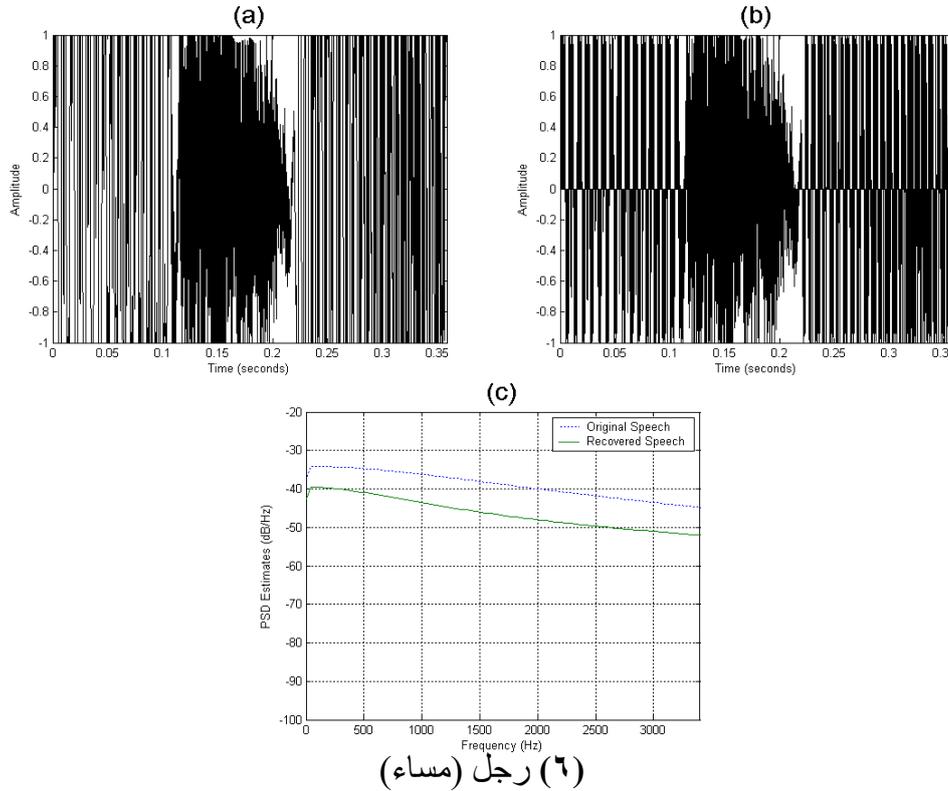
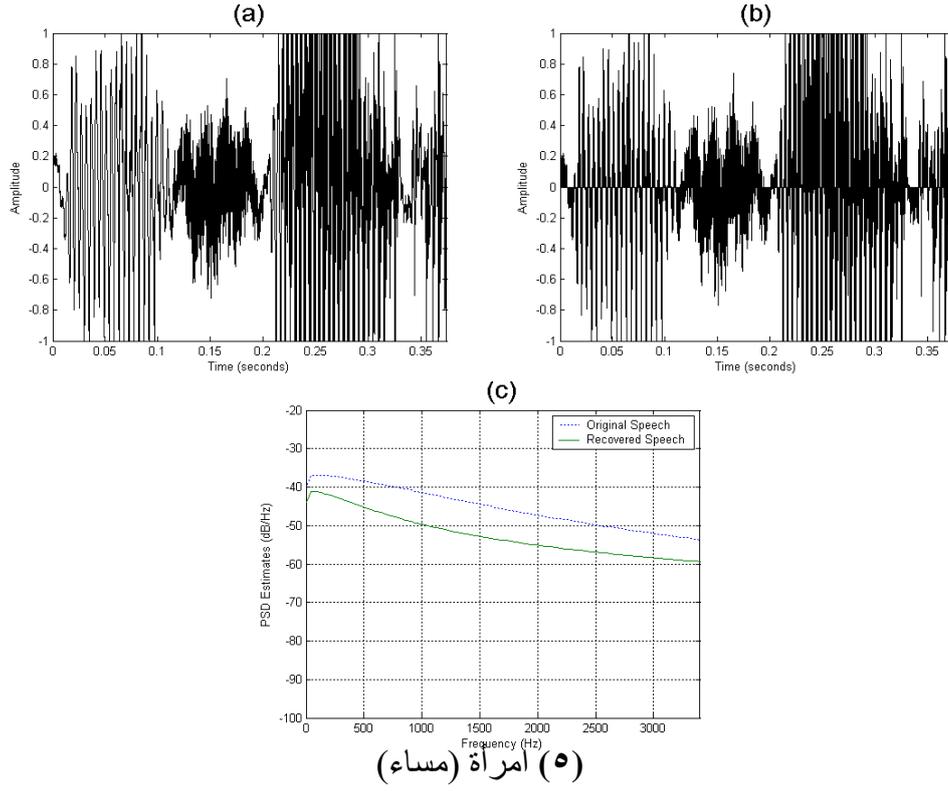
(٢) رجل (الخير)



(٣) رجل (الحمد لله)

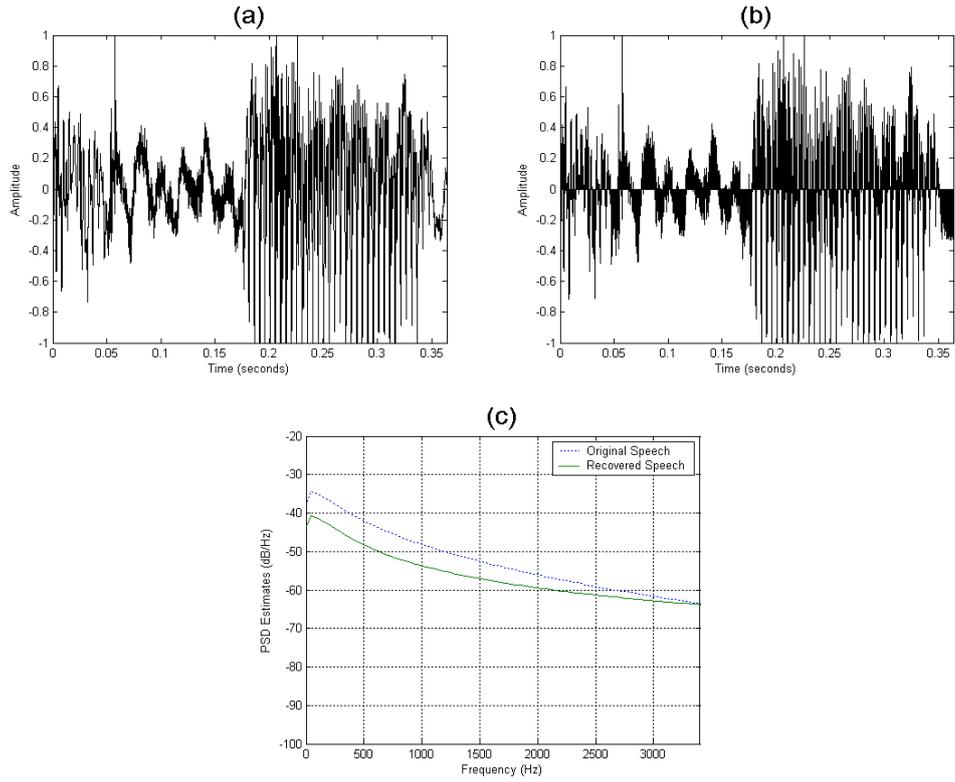


(٤) رجل (الله أكبر)

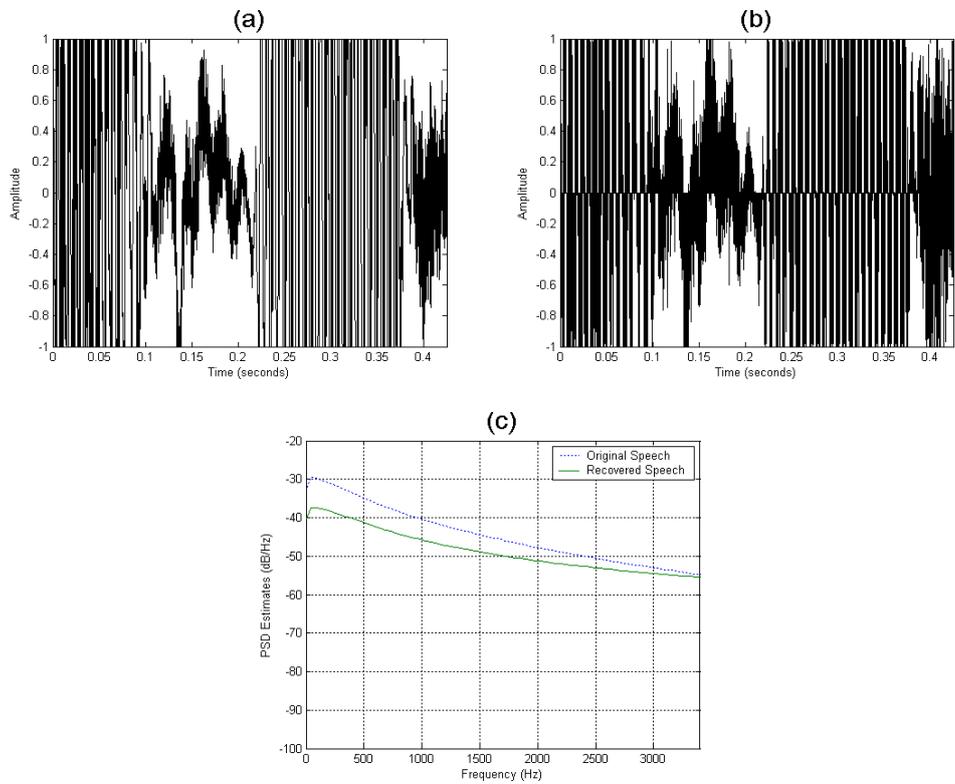


شكل (1.C) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (44100 Hz):

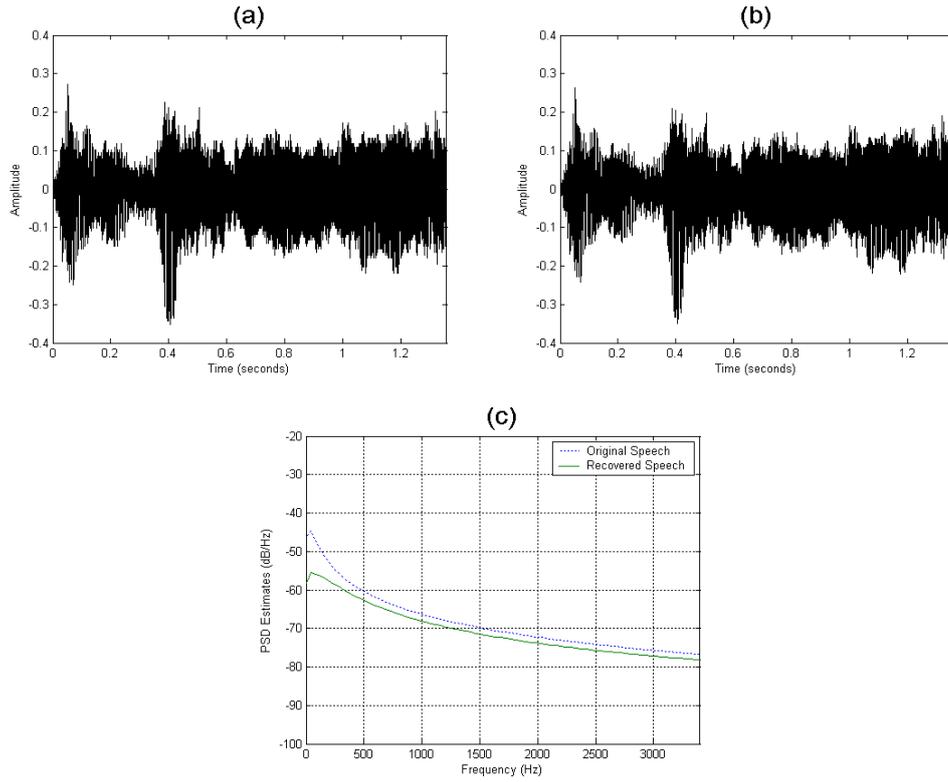
- a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.
- c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة والمسترجعة.



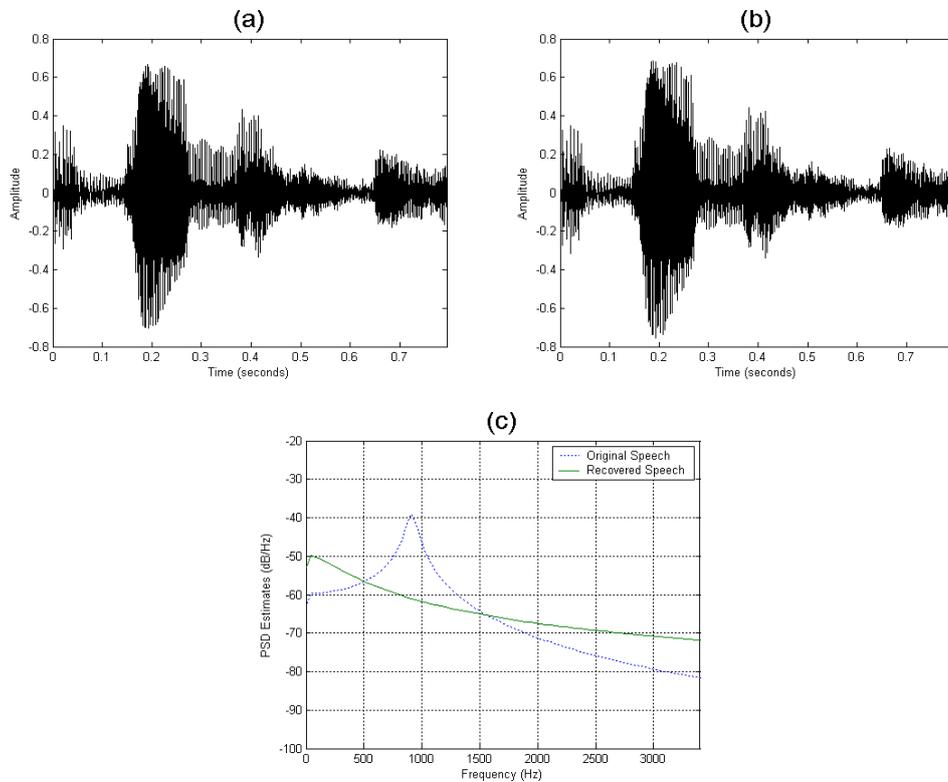
(١) امرأة (الخير)



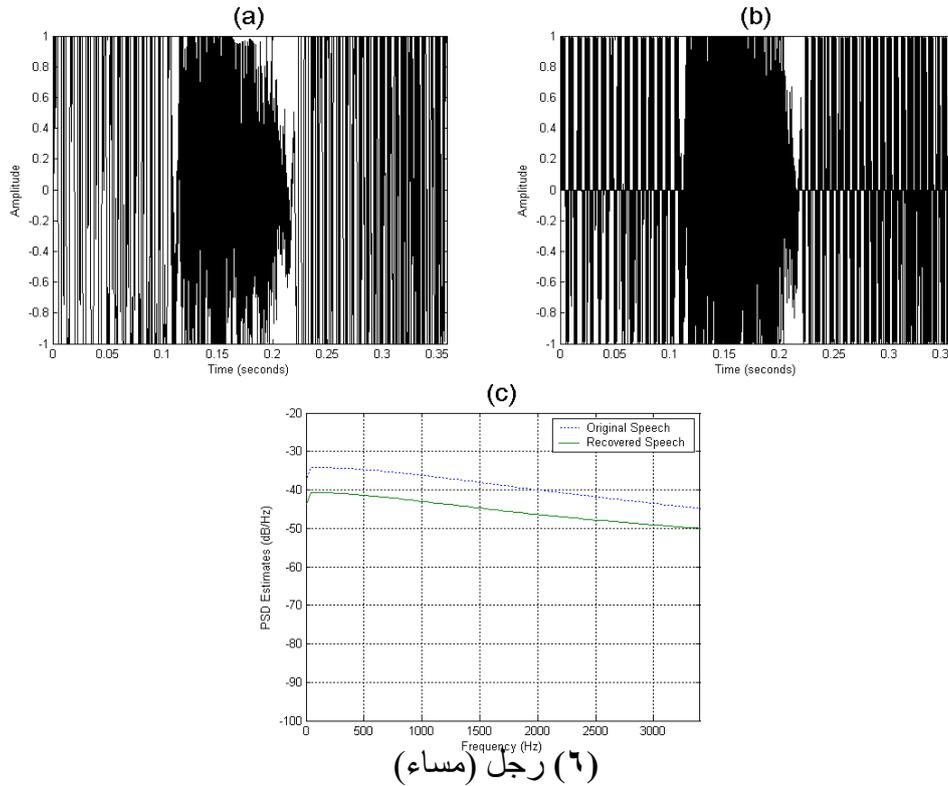
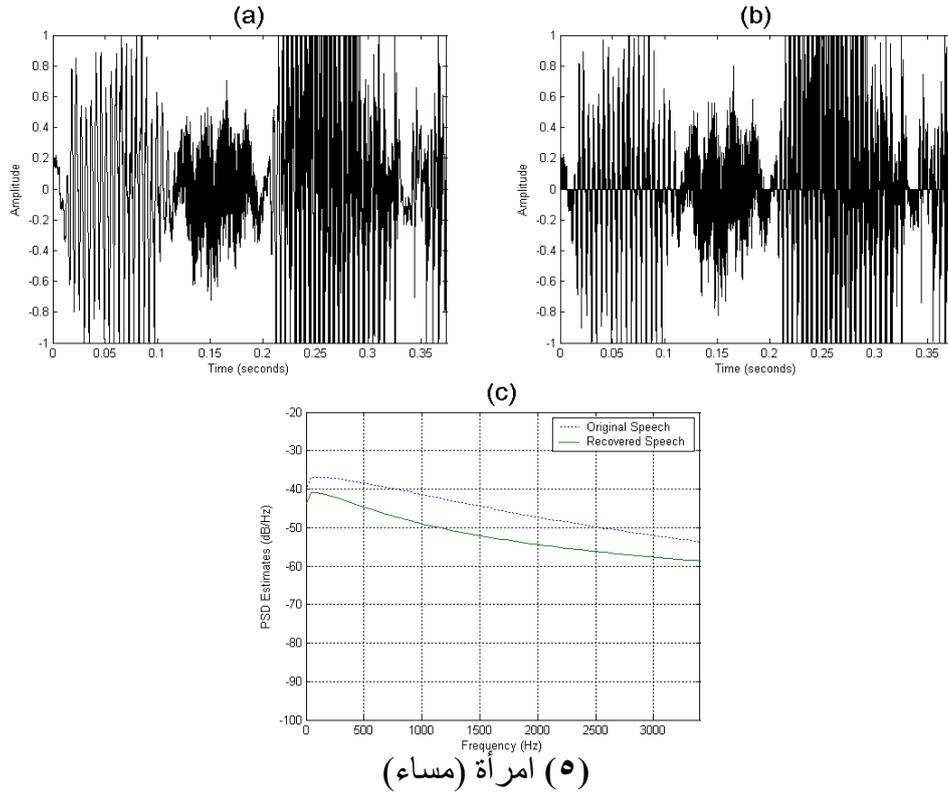
(٢) رجل (الخير)



(٣) رجل (الحمد لله)



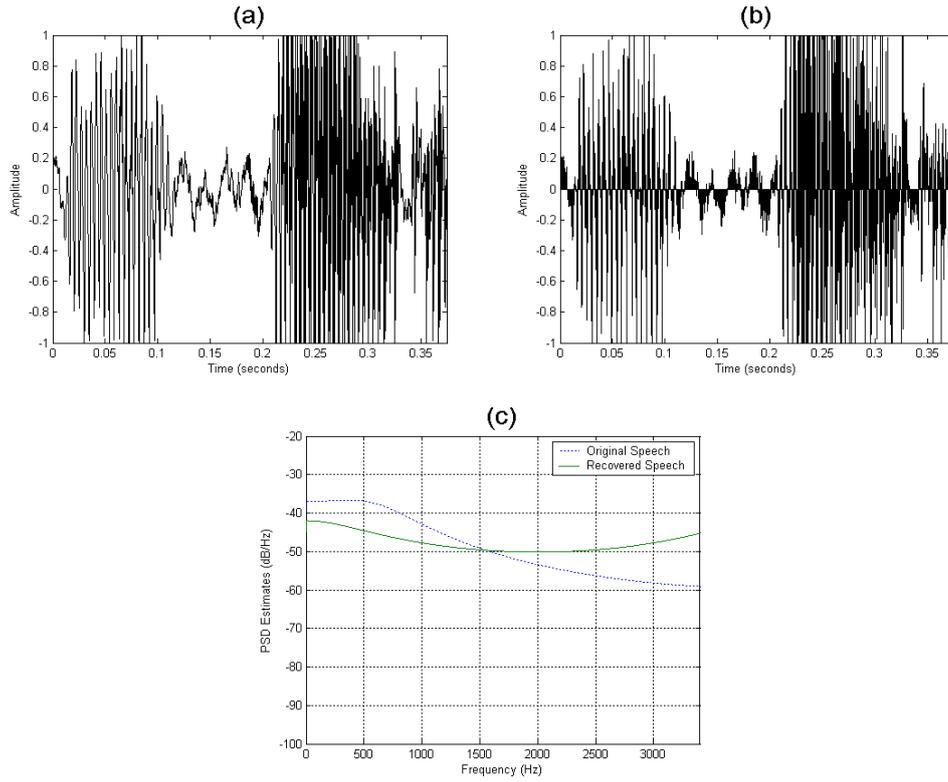
(٤) رجل (الله أكبر)



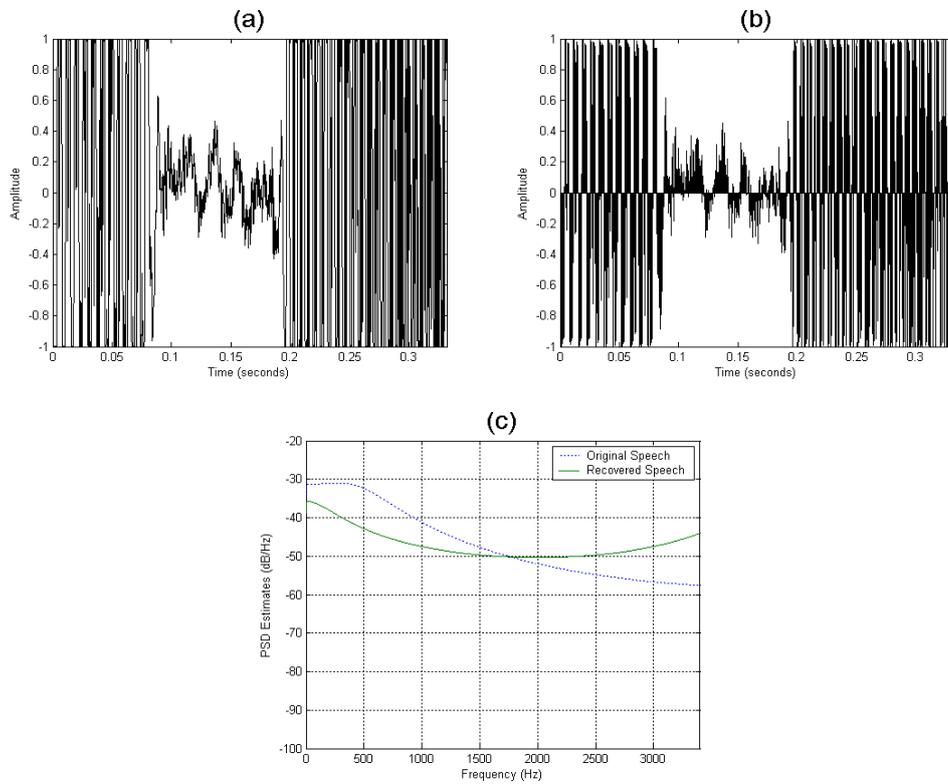
شكل (٢.C) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (44100 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

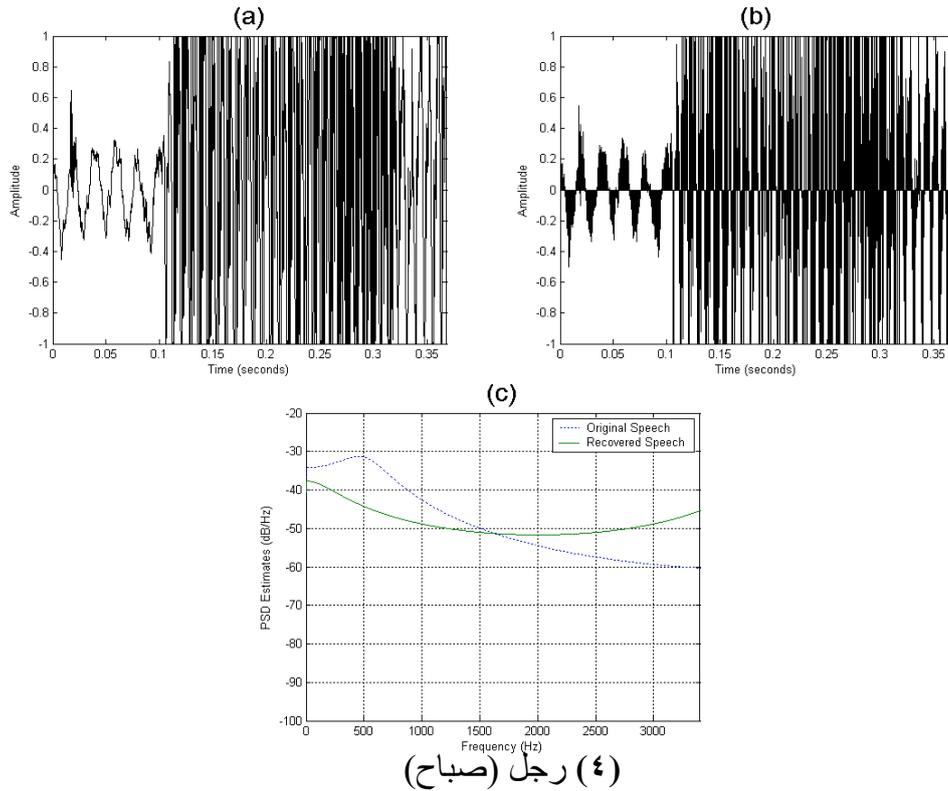
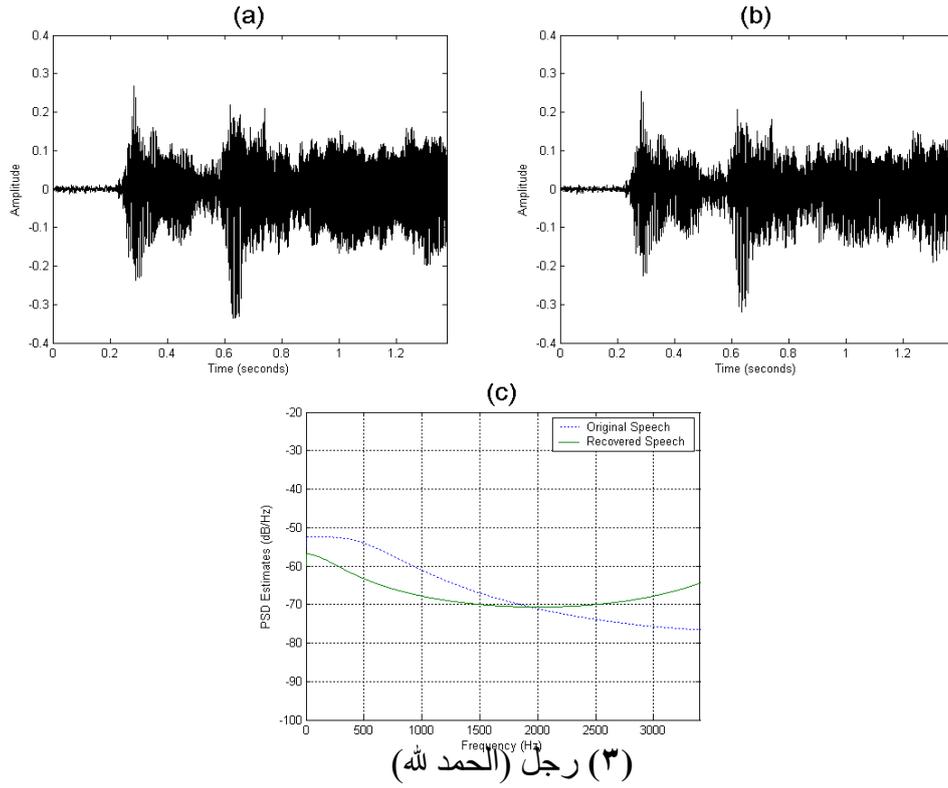
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة والمسترجعة.



(١) امرأة (مساء)



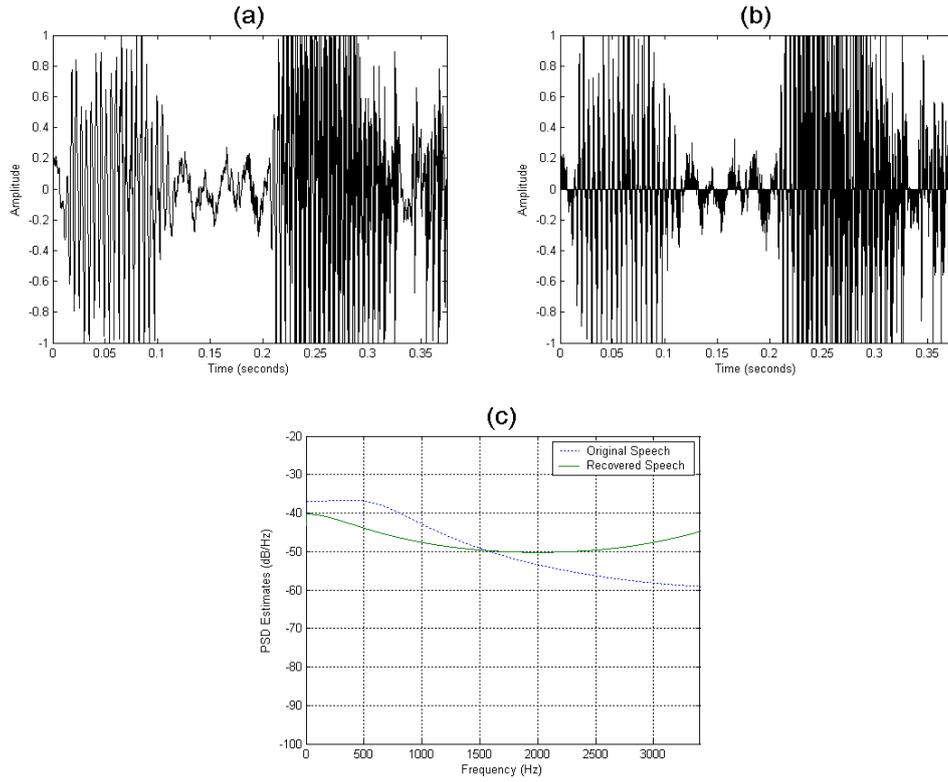
(٢) رجل (مساء)



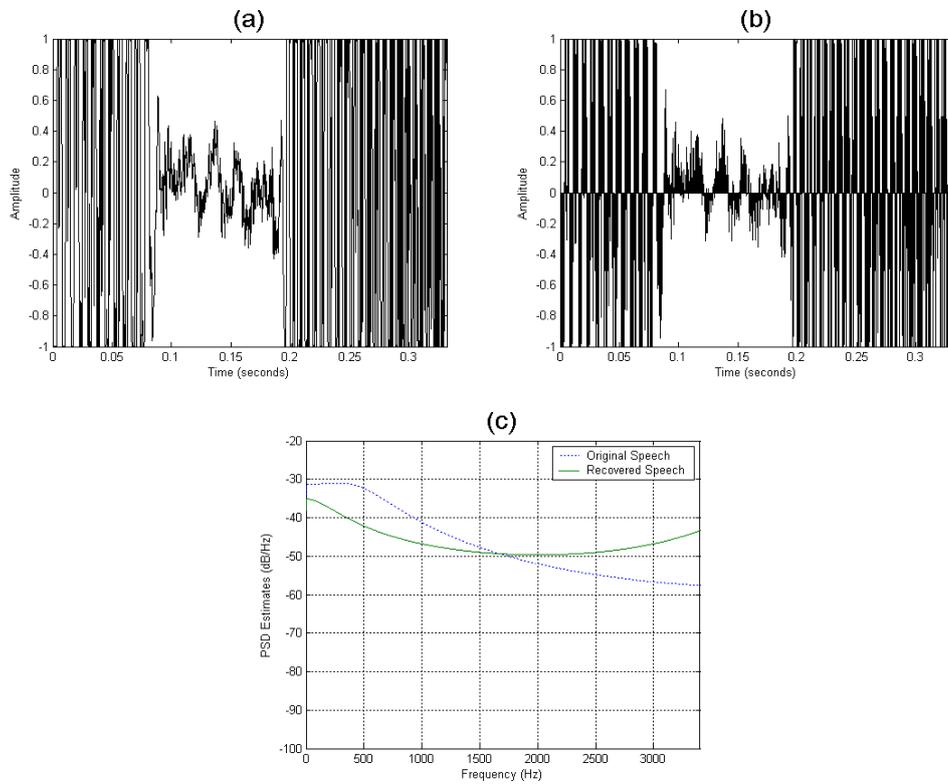
شكل (٣.C) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه الأساسي عند تردد نمذجة (٨٠٠٠ Hz):

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

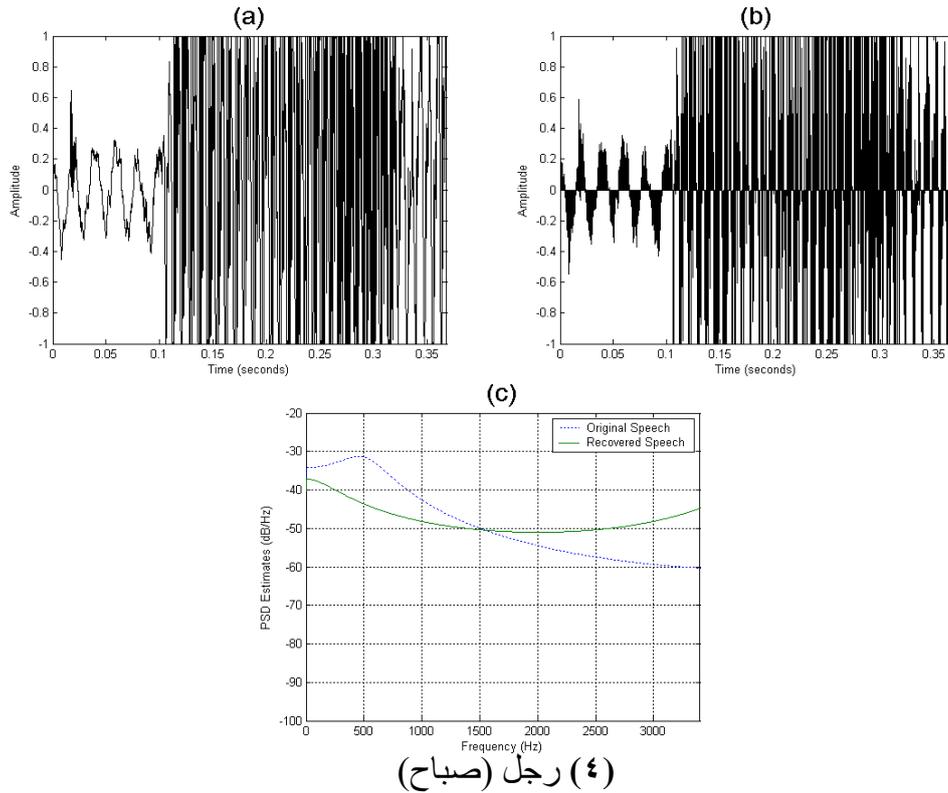
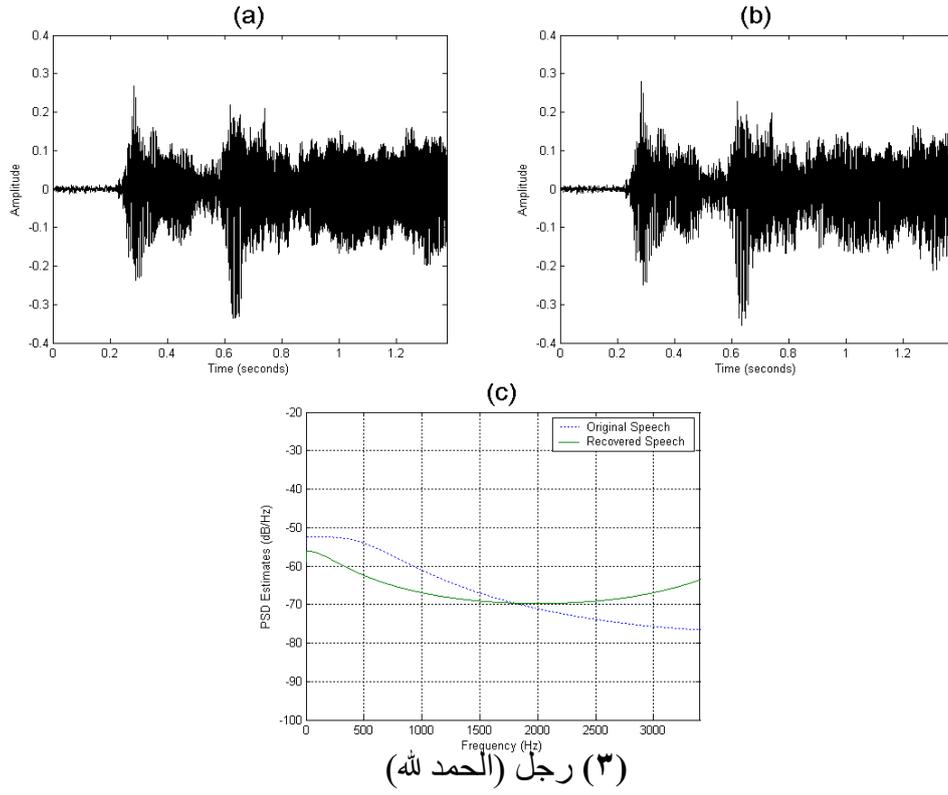
c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.



(١) امرأة (مساء)



(٢) رجل (مساء)



شكل (٤.C) الأشكال الناتجة باستخدام النافذة الهجينة من الاتجاه المقترح عند تردد نمذجة (8000 Hz) :

a. الإشارة الكلامية الأصلية. b. الإشارة الكلامية المسترجعة.

c. مدى التطابق (الاختلاف) بين الإشارة الكلامية الأصلية والمسترجعة. والمسترجعة.

قائمة الأشكال

رقم الصفحة

عنوان الشكل

رقم الشكل

١٣	تقنيات ترميز الإشارة الكلامية	(١.٢)
١٨	علامة معدل الرأي لجودة إشارة الكلام المعاد بناؤه لمعدلات نقل بيانات مختلفة	(٢.٢)
٢١	أ نموذج التحليل – التركيب المثالي	(٣.٢)
٢٣	هيكل مرمرز – فاك ترميز التحليل بالتركيب	(٤.٢)
٢٤	معاملات تقويم كفاءة ترميز الإشارة الكلامية	(٥.٢)
٢٩	جودة إشارة الكلام مقابل معدل نقل البيانات لأصناف الترميز العامة	(٦.٢)
٣٠	طرائق تقويم جودة إشارة الكلام	(٧.٢)
٣٣	مرمرز- فاك ترميز CELP للتحليل بالتركيب	(١.٣)
٣٥	النموذج المبسط لآلية إنتاج الإشارة الكلامية	(٢.٣)
٤٩	المرمرز – فاك ترميز إثارة التنبؤ الخطي (CELP)	(١.٤)
٥٠	هيكل مكونات المرمرز الصوتي	(٢.٤)
٥٦	المخطط الانسيابي لأجراء تحليل إشارة الكلام النموذجي	(٣.٤)
٥٧	إجراء الضرب بنافاذة تحليل التنبؤ الخطي لنافاذة الإتجاه الأساسي	(٤.٤)
٥٧	إجراء الضرب بنافاذة تحليل التنبؤ الخطي لنافاذة الإتجاه المقترح	(٥.٤)
٥٨	المخطط الانسيابي للحصول على معاملات الارتباط الذاتي المعدلة والمتطبعة	(٦.٤)
٦٠	المخطط الانسيابي لحوارزمية ليفنسن داربن التكرارية	(٧.٤)
٦١	معدل مربع الخطأ بطريقة ترميز التنبؤ الخطي بمراتب مختلفة	(٨.٤)
٦٤	المخطط انسيابي للحصول على وحدة العينة أو وحدة النبضة	(٩.٤)
٦٥	المخطط الانسيابي للحصول على سعة الإثارة	(١٠.٤)
٦٧	المخطط الانسيابي للحصول على إستجابة النبضة	(١١.٤)
٦٨	المخطط الانسيابي لاحتساب ربحي كتابي الرموز المتكفي الثابت	(١٢.٤)
٦٩	نموذج لمفهوم تركيب ترميز إثارة التنبؤ الخطي	(١٣.٤)
٧٠	المخطط الانسيابي لاحتساب إشارة الإثارة	(١٤.٤)
٧١	فاك الترميز لمرمرز إثارة التنبؤ الخطي (CELP)	(١٥.٤)
٧٣	المخطط الانسيابي لاحتساب إشارة الكلام المعاد بناؤه	(١٦.٤)
٧٥	المخطط الانسيابي لاحتساب مقياس التقويم SNR	(١٧.٤)
٧٦	المخطط الانسيابي لاحتساب مقياس التقويم SNR _{seg}	(١٨.٤)

رقم الصفحة

عنوان الشكل

رقم الشكل

٧٧	المخطط الانسيابي لاحتساب مقياس التقويم MSE	(١٩.٤)
٨٣	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٤٤١٠٠ Hz)	(١.٥)
٨٦	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٤٤١٠٠ Hz)	(٢.٥)
٩٠	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٨٠٠٠ Hz)	(٣.٥)
٩٢	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٨٠٠٠ Hz)	(٤.٥)
٩٥	الوقت المستغرق لكلمة مساء	(٥.٥)
١٢٣	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٤٤١٠٠ Hz)	(١.A)
١٢٦	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٤٤١٠٠ Hz)	(٢.A)
١٢٨	الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٨٠٠٠ Hz)	(٣.A)

- ١٣٠ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٨٠٠٠ Hz) (٤.A)
- ١٣٣ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٤٤١٠٠ Hz) (١.B)
- ١٣٦ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٤٤١٠٠ Hz) (٢.B)
- ١٣٨ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٨٠٠٠ Hz) (٣.B)
- ١٤٠ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٨٠٠٠ Hz) (٤.B)
- ١٤٣ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٤٤١٠٠ Hz) (١.C)
- ١٤٦ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٤٤١٠٠ Hz) (٢.C)
- ١٤٨ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه الأساسي عند تردد نمذجة (٨٠٠٠ Hz) (٣.C)
- ١٥٠ الأشكال الناتجة باستخدام النافذة الهجينة من الإتجاه المقترح عند تردد نمذجة (٨٠٠٠ Hz) (٤.C)

قائمة الجداول

رقم الصفحة	عنوان الجدول	رقم الجدول
١٢	معدلات ترميز الأشاره الكلامية	(١-٢)
١٤	مقارنة بين تقنيات ترميز الإشارة الكلامية	(٢-٢)
٢٨	مقياس MOS	(٣-٢)
٦٣	هيكل متجه كتاب الرموز الثابت	(١-٤)
٨٧	نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(١-٥)
٨٨	نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(٢-٥)
٩٣	نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(٣-٥)
٩٤	نتائج استخدام المقاييس الرياضية والمقياس الحسي MOS بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(٤-٥)
٩٦	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(٥-٥)
٩٧	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(٦-٥)
٩٨	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(٧-٥)
٩٨	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(٨-٥)
١٠٠	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(٩-٥)
١٠٠	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(١٠-٥)
١٠١	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(١١-٥)

رقم الصفحة	عنوان الجدول	رقم الجدول
١٠٢	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(١٢-٥)
١٠٣	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(١٣-٥)
١٠٤	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه المقترح	(١٤-٥)
١٠٥	نتائج استخدام المقاييس الرياضية بالإعتماد على النافذة الهجينة من الإتجاه الأساسي	(١٥-٥)

قائمة الرموز

الرمز	معناه
AE	سعة الإثارة
a_i	متجه كتاب الرموز المتكيف المقيس بالريح
$Alpha_k$	معاملات التنبؤ الخطي
Aut_k	معاملات الارتباط الذاتي
Aut_{apt_k}	معاملات الارتباط الذاتي المعدلة
$Auto_k$	معاملات الارتباط الذاتي المعدلة والمتطبعة
c_i	متجه كتاب الرموز الثابت
C_i	الترابط بين إشارة كتاب الرموز الثابت وإشارة الهدف $X(n)$
C_a	الترابط بين إشارة كتاب الرموز المتكيف وإشارة الهدف $X(n)$
E	طاقة الخطأ لكل إطار
$e(n)$	خطأ التنبؤ
E_w	معدل مربع الخطأ الموزون على طول الإطار الجزئي
$e_w(n)$	الخطأ الموزون
f_o	وسع النطاق الترددي
f_s	تردد النمذجة
G	الريح
G_1	ريح كتاب الرموز المتكيف
G_2	ريح كتاب الرموز الثابت
$h(n)$	إستجابة النبضة
L	طول إطار الإشارة الكلامية
L_o	طول الإشارة الكلامية
M	عدد الأطارات
m	رقم الإطار
N	طول الإطار الجزئي للإشارة الكلامية
الرمز	معناه
n	رقم العينة
OM	مصفوفة إطارات الإشارة الكلامية الأصلية
$order$	الرتبة
Ref_k	معاملات الإنعكاس
RM	مصفوفة إطارات الإشارة الكلامية المسترجعة
$S(n)$	إشارة الكلام الأصلية
$S_m(n)$	شريحة كلامية للإطار m
$S_w(n)$	إشارة الكلام الأصلية الموزونة
T_{ai}	حد عتبة خاضع للتعظيم

مصدر الإشارة	$u(n)$
متجة كتاب الرموز الثابت المقيس بالربح	v_i
نافذة معينة مطبقة على العينة n	$W(n)$
نافذة الارتباط التأخيري	W_{lag}
نافذة تحليل التنبؤ الخطي مطبقة على العينة n	$W_{lp}(n)$
إشارة الهدف	$X(n)$
معامل التوزين لمرشح التوزين	y
الترابط بين الإشارات من كتابي رموز	Y_{ai}
تحويل z	z
حاصل ضرب إشارة الكلام في نافذة تحليل التنبؤ الخطي	$\bar{S}(n)$
إشارة الكلام المسترجعة	$\hat{S}(n)$
إشارة الكلام المسترجعة المحسنة	$\hat{S}_e(n)$
إشارة الكلام المسترجعة الموزونة	$\hat{S}_w(n)$
إستجابة الإدخال الصفرية	$\hat{S}_0(n)$
التأخير	α
وحدة العينة أو وحدة النبضة	$\delta(n)$
طاقة إشارة كتاب الرموز الثابت	ζ_i
طاقة إشارة كتاب الرموز المتكيف	ζ_α

المصادر (References)

- [١] Z. Yuan , " The Weighted Sum of The Line Spectrum Pair For Noisy Speech " , M.Sc. Thesis , Department of Electrical and Communications Engineering , Helsinki University of Technology , May ٢٠٠٣.
- [٢] W. Li , A. Sridhar , and T. Teng , " Compression of Speech Coding Algorithms : ADPCM ,CELP and VSELP " , Project presented to University of Texas at Dallas , pp. ١-١٢, ١٩٩٩ .
- [٣] B. S. Atal , and N. S. Jayant , " Transmission and Storage " , AT & T Bell Laboratories , Murray Hill , New Jersey , USA , <http://cslu.cse.ogi.edu/HLTsurvey/ch١٠node٤.html> .
- [٤] M. H. Johnson , and A. Alwan , " Speech Coding: Fundamentals and Applications " , Encyclopedia of Telecommunications , Wiley, Dec. ٢٠٠٢.
- [٥] A. S. Spanias , "Speech Coding : A Tutorial Review " , Department of Electrical Engineering , Arizona State University, <http://www.fulton.asu.edu/~spanias/E٦٠٧S٠٥/SpeechCoding.pdf> ,pp. ١- ٩٥ , Jan. ٢٠٠٢ .
- [٦] R. V. Cox , and P. Kroon , " Low Bit - Rate Speech Coders For Multimedia Communication " , IEEE Communications Magazine, Vol. ٣٤ , No. ١٢ , pp. ٣٤ - ٤١, Dec. ١٩٩٦.
- [٧] O. Ozun , P. Steurer , and D. Thell , " Wideband Speech Coding With Linear Predictive Coding (LPC) " , Department of Electrical Engineering , University of California at Los Angeles , pp. ١-١٨ , Winter ٢٠٠٢.
- [٨] H. Dudley , " The Vocoder " , Bell Labs Record , Vol. ١٨ , No. ٤ , pp. ١٢٢-١٢٦ , ١٩٣٩.

- [٩] H. Dudley, " Remaking Speech " , J. Acoust. Soc. Am. , Vol. ١١ , pp. ١٦٩-١٧١ , ١٩٣٩.
- [١٠] W. Munson , and H. Montgomery , " A Speech Analyzer and Synthesizer " , J. Acoust. Soc. Am. , Vol. ٢٢ , pp. ٦٧٨-٦٨١ , ١٩٥٠.
- [١١] H. Dudley , " Phonetic Pattern Recognition Vocoder For Narrow-Band Speech Transmission " , J. Acoust. Soc. Am., Vol. ٣٠ , pp. ٧٣٣-٧٣٩ , ١٩٥٨.
- [١٢] N. S. Jayant , and P. Noll , " Digital Coding of Waveforms " , Prentice-Hall , INC. Englewood Cliffs , ١٩٨٤.
- [١٣] R. Schafer , and L. Rabiner , " Digital Representations of Speech Signals " , Proc. IEEE , Vol. ٦٣ , pp. ٦٦٢-٦٧٧ , Apr. ١٩٧٥.
- [١٤] R. Golden , " Digital Computer Simulation of A Sampled-Datavoice -Excited Vocoder " , J. Acoust. Soc. Am. , Vol. ٣٥ , pp. ١٣٥٨-١٣٦٦ , ١٩٦٣.
- [١٥] G. Fant , " Acoustic Theory of Speech Production " , Mouton and Co. , Gravenhage , The Netherlands, ١٩٦٠.
- [١٦] J. Makhoul , " Linear Prediction : A Tutorial Review " , Proc. IEEE , Vol. ٦٣ , No. ٤ , pp. ٥٦١-٥٨٠ , Apr. ١٩٧٥.
- [١٧] F. Itakura , and S. Saito , " Analysis Synthesis Telephony Based on The Maximum Likelihood Method " , Proc. ٦th International Congress of Acoustics , pp. C-١٧-C-٢٠ , Aug. ١٩٦٨.
- [١٨] B. Atal , and M. R. Schroeder, " Adaptive Predictive Coding of Speech Signals " , Conf. Comm. and Proc. , pp. ٣٦٠-٣٦١ , ١٩٦٧.
- [١٩] B. Atal , and S. Hanauer, " Speech Analysis and Synthesis by Linear Prediction of The Speech Waves " , J. Acoust. Soc. Am. , Vol. ٥٠ , pp. ٦٣٧-٦٣٩ , ١٩٧١.

- [٢٠] J. Makhoul , and J. Wolf , " Linear Prediction and Spectral Analysis of Speech " , BBN Inc. , Rep. ٢٣٠٤ , ١٩٧٢.
- [٢١] J. Markel , and A. Gray, Jr. , " Linear Prediction of Speech " , Springer-Verlag , New York , ١٩٧٦.
- [٢٢] A. Oppenheim , and R. Schafer , " Homomorphic Analysis of Speech " , IEEE Trans. , Vol. AU-١٦ , pp. ٢٢١-٢٢٦ , Jun. ١٩٦٨.
- [٢٣] D. G. Childers , " The Cepstrum " , Proc. IEEE , No. ١٠ , pp. ١٤٢٨-١٤٤٣ , ١٩٧٧.
- [٢٤] J. Flanagan , and Golden , " Phase Vocoder " , Bell Tech. J., Vol. ٤٥ , pp. ١٤٩٣-١٤٩٧ , Nov. ١٩٦٦.
- [٢٥] R. Schafer , and L. Rabiner , " Design and Simulation of A Speech Analysis-Synthesis System Based on Short-Time Fourier Analysis " , IEEE Trans. , Vol. AU-٢١ , No. ٣ , pp. ١٦٥-١٦٨ , Jun. ١٩٧٣.
- [٢٦] M. R. Portnoff , "Short-Time Fourier Analysis of Sampled Speech" , IEEE Trans. , Vol. ASSP-٢٩ , No .٣ , pp. ٣٦٤-٣٧٣ , Jun. ١٩٨١.
- [٢٧] V. Viswanathan , " Speech-Quality Optimization of ١٦ kb/s Adaptive Predictive Coders " , Proc. , Denver , Vol. ICASSP-٨٠ , pp. ٥٢٠-٥٢٤ , ١٩٨٠ .
- [٢٨] J. Tribolet , and R. Crochiere , " Frequency Domain Coding of Speech " , IEEE Trans. , Vol. ASSP-٢٧ , No. ٥ , pp. ٥١٢-٥١٦ , Oct. ١٩٧٩.
- [٢٩] R. Crochiere , S. Webber , and J. Flanagan , " Digital Coding of Speech in Sub-bands " , Bell Tech. J., Vol. ٥٥ , No. ٨ , pp. ١٠٦٩-١٠٧٢ , Oct. ١٩٧٦.
- [٣٠] G. Kang , and L. Fransen , " High Quality ٨٠٠ b/s Voice Processing Algorithm " , NRL Report ٩٣٠١ , Feb. ١٩٩١.

- [٣١] M. R. Schroeder , and B. Atal , " Code-Excited Linear Prediction (CELP) : High Quality Speech at Very Low Bit Rates " , Proc. , Tampa , Vol. ICASSP-٨٥, pp. ٩٣٧-٩٣٩ , Apr. ١٩٨٥.
- [٣٢] B. Atal , and M. R. Schroeder , " Stochastic Coding of Speech Signals at Very Low Bit Rates " , Proc. Int. Conf. Comm., pp. ١٦١٠-١٦١٣ , May ١٩٨٤.
- [٣٣] J. Chen , and R. Cox , " A Fixed-Point LD-CELP Algorithm and Its Real-Time Implementation " , Proc. , Toronto, Vol. ICASSP-٩١, pp. ٢١-٢٤ , May ١٩٩١.
- [٣٤] R. McAulay , and T. Quatieri , " Multirate Sinusoidal Transform Coding at Rates From ٢.٤ kbits/s to ٨ kbits/s " , Proc. , Dallas, Vol. ICASSP-٨٧, Apr. ١٩٨٧.
- [٣٥] D. Griffin , and J. Lim , " Multiband Excitation Vocoder " , IEEE Trans. , Vol. ASSP-٣٦ , No. ٨ , pp. ١٢٢٣-١٢٢٦ , Aug. ١٩٨٨.
- [٣٦] P. Vary , "Speech Codec For The European Mobile Radio System " Proc. , Vol. ICASSP-٨٨ ,pp. ٢٢٧-٢٣١, Apr. ١٩٨٨.
- [٣٧] I. Gerson , and M. Jasiuk , " Vector Sum Excited Linear Prediction (VSELP) Speech Coding at ٨ kbits/s " , Proc. , New Mexico , Vol. ICASSP-٩٠ , pp. ٤٦١-٤٦٤ , Apr. ١٩٩٠.
- [٣٨] I. Gerson , M. Jasiuk , M. McLaughlin , and E. Winter , "Combined Speech and Channel Coding at ١١.٢ kbits/s" , Signal Processing V: Theories and Applications , Elsevier , pp. ١٣٣٩-١٣٤٢ , ١٩٩٠ .
- [٣٩] J. H. Chen , M. J. Melchner , R. V. Cox , and D. O. Bowker , "Real – Time Implementation and Performance of A ١٦ kb/s Low – Delay CELP Speech Coder " , Proc. IEEE Int. Conf. Acoust. , Speech , Signal Processing , pp. ١٨١-١٨٤ , ١٩٩٠ .
- [٤٠] I. A. Gerson , and M. A. Jasiuk , " Vector Sum Excited Linear Prediction (VSELP) Speech Coding at ٨ kbs " , Proc. IEEE Int. Conf. Acoust. , Speech , Signal Processing , pp. ٤٦١- ٤٦٤ , ١٩٩٠ .

- [٤١] A. Kataoka , T. Moriya , and S. Hayashi , " An \wedge -kbit/s Speech Coder Based on Conjugate Structure CELP " , Proc. IEEE Int. Conf. Acoust. , Speech , Signal Processing , pp. ٥٩٢-٥٩٥ , ١٩٩٣ .
- [٤٢] W. B. Kleijn , and J. Haagen , " Transformation and Decomposition of The Speech Signal For Coding " , IEEE Signal Processing Lett. ,Vol. ١ , No. ٩ , PP. ١٣٦-١٣٨ , ١٩٩٤.
- [٤٣] L. Supplee, R. Cohn , J. Collura , and A. McCree , " MELP : The New Federal Standard at ٢٤٠٠ bps " , Proc. IEEE Int. Conf. Acoust. , Speech , Signal Processing , pp. ١٥٩١-١٥٩٤ , ١٩٩٧ .
- [٤٤] H. K. Kim , and H. S. Lee , " A Low – Bit – Rate Extension Algorithm to The \wedge kbit/s CS-ASELP Based on Adaptive Fixed Codebook Modeling " , IEICE Trans. Inf. & Syst. ,Vol. E٨٢-D , No.٧ , pp. ١٠٨٧-١٠٩٢ , Jul. ١٩٩٩.
- [٤٥] T. K. Capin , E. Petajan , and J. Ostermann , " Very Low BitRate Coding of Virtual Human Animation in MPEG-٤ " , <http://www.research.att.com:٩٠٠٠/projects/tts/papers/٢٠٠٠/ICME/coding> , ٢٠٠٠.
- [٤٦] K. I. Siddiqui , N. M. Sheikh , H. Raza , M. A. Farooq , and Baha uddin , " Real –Time Implementation of ITU-T's G.٧٢٣.١ Dual Rate Speech Coder For Multimedia Communications Transmitting at ٥.٣ and ٦.٣ kbits/s on Trimedia's TM-١٠٠٠ VLIW DSP CPU " , Proc. of IEEE International Multitopics Conference, Lahore , Pakistan , Dec. ٢٠٠١.
- [٤٧] O. Gottesman , and A. Gersho , " Enhanced Waveform Interpolative Coding at Low Bit-Rate " , IEEE Transactions on Speech and Audio Processing , Vol. ٩ , No. ٨ , pp. ٧٨٦-٧٩٧ , Nov. ٢٠٠١.
- [٤٨] R. Yu , and C. C. Ko , " A Warped Linear - Prediction – Based Subband Audio Coding Algorithm " , IEEE Transactions on

- Speech and Audio Processing , Vol. ١٠ , No. ١ , pp. ١-٨ , Jan. ٢٠٠٢.
- [٤٩] V. Atti , and A. Spanias , " On – Line Simulation Modules For Teaching Speech and Audio Compression Techniques " , ٣٣rd ASEE/IEEE Frontiers in Education Conference , pp. T٤E-١٧–T٤E-٢٢ , Nov. ٢٠٠٣ .
- [٥٠] W. C. Chu, and T. Miki , " Optimization of Window and LSF Interpolation Factor For The ITU-T G.٧٢٩ Speech Coding Standard " , Eurospeech , http://www.isca-speech.org/archive/papers/eurospeech_٢٠٠٣/e٠٣_١٠٦١.pdf , pp. ١٠٦١-١٠٦٤ , ٢٠٠٣.
- [٥١] P. Alku , and T. Backstrom , " Linear Predictive Method For Improved Spectral Modeling of Lower Frequencies of Speech With Small Prediction Orders " , IEEE Transactions on Speech and Audio Processing , ٢٠٠٤ .
- [٥٢] A. Spanias, "Communication Systems: Digital Communications Introduction to Source Coding" , Lecture ١٢ , Department of Electrical Engineering , Arizona State University , <http://www.eas.asu.edu/~spanias> , pp.١-٢١ , Apr. ٢٠٠٢.
- [٥٣] W. Pereira , " Modifying LPC Parameter Dynamics to Improve Speech Coder Efficiency " , M.Sc. Thesis , Department of Electrical & Computer Engineering , Faculty of Graduate Studies and Research , McGill University, Sept. ٢٠٠١.
- [٥٤] C. K. Lin , " Low-Complexity CodeBook Searching Algorithms For FS١٠١٦ " , M.Sc. Thesis , Department of Communications Engineering , National Chiao Tung University , Jun. ٢٠٠٢ .
- [٥٥] S. S. Kapadia , " Protocol Design For Congestion Management in Narrowband Integrated Networks " , M.Sc. Thesis , Department of Electrical Engineering and Computer Science , Faculty of The Graduate School , Kansas University, ١٩٩٢ .
- [٥٦] J. McDonough , C. Chienchung , P. Hantak , and C. Sakamaki , " A Single Chip QCELP Vocoder For CDMA Digital Cellular " , CICC '٩٤ , pp. ٢١١-٢١٤ , May ١٩٩٤.

- [٥٧] A. Spanias , " Communication Systems : LPC Vocoders in Cellular Systems " , Lecture ١٧ , Department of Electrical Engineering , Arizona State University , <http://www.eas.asu.edu/~spanias> , pp.١-٢٩, ٢٠٠٢.
- [٥٨] A. Perkis , " Multimedia Signal Processing Application in Audiovisual Communications " , Partly published in "Satellite Communication ; Fixed and Mobile Services" , Kluwer Academic Publishers , New York , pp. ١-٣٦ , Jun. ١٩٩٣ .
- [٥٩] A. Kiviluoto , " Speech Coding Standards " , <http://mia.ece.uic.edu/~papers/WWW/MultimediaStandards/chapter٣.pdf>.
- [٦٠] C. G. Bell , H. Fujisaki , J. M. Heinz , K. N. Stevens , and A. S. House , " Reduction of Speech Spectra by Analysis - by - Synthesis Techniques " , J. Acoust. Soc. Amer. , Vol. ٣٣ , No.١٢ , pp. ١٧٢٥-١٧٣٦ , ١٩٦١.
- [٦١] R. M. Gray , A. Buzo , A. H. Gray, Jr. , and Y. Mastuyama , " Distortion Measures For Speech Processing " , IEEE Trans. Acoustics , Speech and Signal Processing , Vol. ASSP-٢٨ , No. ٤ , pp. ٣٦٧-٣٧٦ , Aug. ١٩٨٠ .
- [٦٢] R. Kubichek , " Standards and Technology Issues in Objective Voice Quality Assessment " , Digital Signal Processing : A Review Journal , DSP ١ , pp. ٣٨-٤٤ , Apr. ١٩٩١.
- [٦٣] W. D. Voiers , " Evaluating Processed Speech Using The Diagnostic Rhyme Test " , Speech Techn. , Jan. ١٩٨٣.
- [٦٤] B. S. Atal , and M. R. Schroeder, "Code-Excited Linear Prediction(CELP) High-Quality Speech at Very Low Bit Rates", Int. Conf. Acoust. , Speech and Signal Processing , Vol. ICASSP-٢٥, pp. ٩٣٧-٩٤٠, Mar. ١٩٨٥.
- [٦٥] J. P. Woodard , and L. Hanzo , " Improvements to The Analysis -by- Synthesis Loop in CELP Codecs " , Southampton University , <http://www.ecs.soton.ac.uk/publications/rj/١٩٩٥-١٩٩٦/comms/jpw٩١r.pdf> , Jun. ١٩٩٦ .

- [٦٦] L. R. Rabiner , and R. W. Schafer , " Digital Processing of Speech Signals " , Prentice-Hall , Inc. , Englewood Cliffs , New Jersey , ١٩٧٨ .
- [٦٧] ITU-T Recommendation G.٧٢٩ , " Coding of Speech at \wedge kbit/s Using Conjugate - Structure Algebraic – Codec - Excited Linear-Prediction (CS-ACELP) " , ١٩٩٦.
- [٦٨] H. Kumar , and K. Sundaresan , " Implementation of The Code Excited Linear Predictive (CELP) Codec For VOIP " , Department of Electrical Engineering , State University Of New York at Buffalo , ١٩٩٦ .
- [٦٩] ITU-T Recommendation G.٧٢٣.١ , " Dual Rate Speech Coder For Multimedia Communications Transmitting at ٥.٣ & ٦.٣ kbit/s " , Mar. ١٩٩٦ .
- [٧٠] S. K. Mitra , " Digital Signal Processing (A Computer – Based Approach) " , The McGraw-Hill Companies , Inc. , United States of America , ١٩٩٨.
- [٧١] T. H. Andersen , " Filtering " , DIKU , Spring ٢٠٠٣ .
- [٧٢] R. W. Hamming , " Digital Filters " , Prentice-Hall , Inc. , Englewood Cliffs , New Jersey , ١٩٧٧.