

Segmentation of Arabic Word into Letters and Recognition

Faten Basher Abd Alahad

Enaam Ghanim Saeed

Enaamghanim@uomosul.edu.iq

College of Computer Sciences and Mathematics

University of Mosul, Iraq

Received on: 06/04/2005

Accepted on: 09/08/2005

ABSTRACT

In this research a system for Arabic word segmentation into letters and recognition has been designed by dividing it into five groups (plosives, fricatives, nasals, glide and semi-vowel sounds) based on articulatory phonetics. This system include four main stages:

Stage one: Endpoint algorithm has been used to identify the beginning and the end of word.

Stage two: A new segmentation algorithm has been suggested and implemented depending upon time domain features and Arabic phonology rules.

Stage three: This stage includes letters feature extraction depending on linear predictive coding and system data base constructing which include vectors of features for the segmented letters from words and regarding letter recurring in different positions in the word.

Stage four: Includes Arabic letters recognition according to articulation which entails using Dynamic Time Warping (DTW) method that uses dynamic programming basics to obtain the matching path for the least distance accumulated value, where the word used in segmentation and recognition belongs to the four persons who create the data base and the results were in consistence which ranged from (75- 80)%

Keywords: segmentation, letters, word

تقطيع الكلمة العربية إلى أحرف وتمييزها

إنعام غانم سعيد

فاتن بشير عبد الأحد

كلية علوم الحاسبات والرياضيات، جامعة الموصل

تاريخ قبول البحث: 2005/08/09

تاريخ استلام البحث: 2005/04/06

الملخص

في هذا البحث تم تصميم نظام لتقطيع الكلمات العربية إلى أحرف وتمييزها من خلال تقسيمها إلى خمس مجاميع وهي (الأصوات الانفجارية ، الأصوات الاحتكاكية ، الأصوات الأنفية ، الأصوات الانزلاقية والأصوات شبه الحركية) اعتماداً على علم الأصوات النطقي. يتضمن النظام أربع مراحل رئيسية وهي كما يأتي:

المرحلة الأولى : يتم فيها تطبيق خوارزمية لتحديد بداية نطق الكلمة ونهايته.

المرحلة الثانية : تم اقتراح وتنفيذ خوارزمية تقطيع جديدة تعتمد على خصائص المنطلق الزمني وقواعد علم الفونولوجي في اللغة العربية.

المرحلة الثالثة: تتضمن استخلاص الصفات للحروف بالاعتماد على طريقة ترميز التنبؤ الخطي LPC وبناء قاعدة البيانات للنظام والتي تشمل متجهات الصفات للأحرف المستقطعة من الكلمات مع الأخذ بنظر الاعتبار ورود الحرف في مواضع مختلفة من الكلمة.

المرحلة الرابعة: تشتمل على تمييز الحروف العربية حسب كيفية نطقها فقد تم استخدام طريقة انحراف الوقت الديناميكي (DTW [dynamic time warping]) التي تستخدم أساسيات البرمجة الديناميكية لإيجاد مسار التطابق الذي يمتلك أقل قيمة تراكمية للمسافة إذ أن الكلمات المستخدمة في التقطيع والتمييز تعود للأشخاص الأربعة المكونين لقاعدة البيانات وكانت النتائج جيدة فقد تراوحت بين (75-80)%.

الكلمات المفتاحية: تقطيع , حروف , كلمة.

1. المقدمة:

الكلام هو إحدى الوسائل المستخدمة عالمياً للاتصال بين أفراد الجنس البشري، وهو هبة الله سبحانه وتعالى للبشر الذين ميزهم به عن سائر المخلوقات الأخرى، والكلام هو استحداث موجات صوتية بواسطة الحركة الإرادية للتركيب التشريحي في نظام توليد الكلام لدى الإنسان لتنتقل المعلومات من المتكلم إلى السامع. هذا التركيب يتألف من ربط قنوات عديدة لتشكل بمجموعها جهاز النطق وأعضائه (Vocal tract and speech organs) ، وشكلها مختلف من شخص إلى آخر ولهذا تتغير طريقة إنتاج الكلام [2] [3].

فالكلام هو سلسلة صوتية متصلة ببعضها البعض اتصالاً وثيقاً ، يمكن وصف الكلام بصيغة إشارة تحمل معلومات معينة ، أي تمثيلها بشكل موجة سمعية (Wave Form).

في كل أنظمة تمييز الكلام تتم معالجة الإشارة من البداية إلى النهاية (Front – end) processing ، مثل (المعالجة بالنافذة ، الترشيح ، التعيان ، التكميم) . حسب نوع التطبيق فان إشارة الكلام يمكن ان تتمثل بالاعتماد على معالم (Parameters) متنوعة للكلام مثل التمثيل باستخدام معلمات المجال الزمني (Time domain parameters) للإشارة أو المعلمات الطيفية (Spectral parameters) ليتم استخراج خصائص الكلام المطلوبة من قبل النظام المصمم ، إن اختيار نوع التمثيل للكلام يؤدي دوراً كبيراً ومهماً في إظهار خصائص الكلام [7][9].

2. تقنيات تمييز الكلام:

إن التقنيات المستخدمة لتمييز الكلام يمكن تصنيفها كالاتي :

1. تمييز الكلمة المنفصلة (Isolated word recognition)

أنظمة تمييز الكلام المنفصل تميز الكلمات المنطوقة بصورة منفصلة (Separately) ، وهنا يجب على المستخدم ان يتوقف بين كلمة وأخرى.

2. تمييز الكلمة المتصلة (Connected word recognition)

تعدد كحالة وسطية بين تمييز الكلمة المنفصلة (Isolated word recognition) و تمييز الكلام المستمر (Continuous speech recognition) تسمح للمستخدم بان يتكلم بعدد من الكلمات في وقت واحد ودون توقف بين كلمة وأخرى.

3. تمييز الكلام المستمر (Continuous – speech recognition)

أنظمة تمييز الكلام المستمر اكثر صعوبة من أنظمة تمييز الكلام المنفصل ، النظام يدرّب على تمييز وحدات الكلام مثل (الفونيمات Phonemes ، المقاطع Syllables) . وان التمييز في الأنظمة المستمرة يتطلب تسلسلاً سمعياً (Acoustic continuum) ليتم تقطيعه إلى وحدات صوتية اصغر من الكلمة مثلا إلى مقاطع (Syllables) أو إلى أنصاف المقاطع (Semi – syllables) أو إلى فونيمات (Phonemes) وبعدها يتم تمييز هذه المقاطع أو الفونيمات. هذه النوعية من التمييز تستخدم الفونيمات وحدة أساسية للتمييز وتحتاج فضلاً عن الخصائص المميزة للفونيمات إلى معرفة بإشارات اللغة وقواعدها ، وهذه المعرفة ضرورية لحل بعض المشاكل ، في هذا النوع من التمييز حجم المصادر يكون كبيراً وهذا يعتمد بصورة مباشرة على اللغة المستخدمة [5][6][8][10].

3. الفونيمات :

تعد الفونيمات الوحدة الرئيسية التي تنقل المعنى اللغوي للكلام، فالفونيم هو اصغر وحدة صوتية محددة بعدد من الصفات الواضحة والمناسبة التي تميزه من غيره من الفونيمات لتلك اللغة. إن دراسة وتصنيف هذه الأصوات الكلامية نفسها تعرف بعلم الأصوات (Phonetics). تتألف اللغة العربية الفصحى (Standard Arabic language) من 29 فونيم من الفونيمات الصامتة (Consonants) و 6 فونيمات من فونيمات العلة (Vowels) ، وبذلك فإن الفونيمات يمكن أن تصنف إلى صنفين العلة (Vowels) والصامتة (Consonants) [4][8] . ويمكن تصنيف الصوامت على النحو الآتي :

1. الأصوات الانفجارية (The plosive sounds)

عند النطق بهذه الأصوات يسمع انفجار خفيف ناتج عن اندفاع الهواء المنحبس خلف مخرج الصوت ، وهناك مصطلح آخر للانفجارات هو (" الوقفات Stops ") إذ أن مصطلح الوقفات يشير إلى قفل مجرى الصوت قفلاً كاملاً. الأصوات الانفجارية في اللغة العربية هي : ب، ض، د، ك، ت، ء، ق، ط.

2. الأصوات الاحتكاكية (The fricative sounds)

الأصوات الاحتكاكية عند نطقها لا يسمع معها مثل هذا الانفجار وذلك لان الهواء لا ينحبس عند مخرج الصوت بل ينساب في يسر وهواده وعليه فإن الأصوات الاحتكاكية في اللغة العربية هي: غ، ع، ج، ز، ظ، ذ، هـ، ح، خ، ص، ش، س، ث، ف.

3. الأصوات الأنفية (The nasals sounds)

هي الأصوات التي تشترك القناة الصوتية (Vocal tract) والقناة الأنفية في إنتاجها ويوجد في اللغة العربية حرفان أنفیان هما : م ، ن.

4. الأصوات الإنزلاقية (The glides sounds)

أهم ما يميز الأصوات الإنزلاقية إن الهواء المنحبس عند النطق بها يخرج من جانب من جوانب اللسان أو من كليهما واللغة العربية تمتلك ثلاثة أصوات إنزلاقية هي: ل المرققة ، ل المضخمة ، ر .

5. أصوات أشباه الحركات (Semi _ Vowels sounds)

تمتلك اللغة العربية صوتين من أشباه الحركات هما : و ، ي [1][8].

4. تسجيل الكلام:

المتكلمون : هم أربعة أشخاص اثنان من الذكور واثنان من الإناث تتراوح أعمارهم بين (20-26) عاما ، قام كل واحد منهم بتسجيل ما يقارب الـ(35) كلمة مختلفة. عند إصدار الكلمة أمام لاقط الصوت تتغير الموجات الصوتية إلى موجات تناظرية ثم تقوم بطاقة الصوت (Sound card) بتحويلها إلى عينات. في هذا البحث تم تسجيل الكلمة بمعدل عينات يساوي (11025 HZ) وتمثيل كل عينة بـ (64 bit) واستخدام الأسلوب الأحادي (Mono) في التسجيل. خزنت هذه العينات الصوتية داخل ملفات بصيغة (Wav) وهي من الصيغ المستخدمة بشكل واسع في بيئة النوافذ (Windows).

5.مرحلة إزالة الصمت

من المشاكل التي تواجه عملية معالجة إشارة الكلام هي تحديد بداية النطق الحقيقي ونهايته وذلك لتقليل حجم الملفات الصوتية التي تؤدي إلى تقليل المعالجة المطلوبة لإشارة الكلام والحصول على كفاءة أكبر في التمييز . وتتخصص إزالة الصمت في هذا البحث بخطوتين أساسيتين هما:

1. حساب معدل السعة لإشارة الكلمة
2. إزالة السكوت بتطبيق خوارزمية تحديد بداية الكلام ونهايته (Endpoint algorithm)[11]

5-1 حساب معدل السعة لإشارة الكلمة

إن السعة لإشارة الكلام تتغير مع الزمن ، وتمثيل إشارة الكلمة بالاعتماد على معدل السعة لوقت قصير يعطي تمثيلاً ملائماً لإشارة الكلمة لأنه يعكس هذه التغييرات بشكل واضح وجيد، ويتم إنجاز هذه العملية من خلال الخطوات الآتية:

1. معالجة إشارة الكلمة من البداية إلى النهاية نافذة هامينك (Hamming window)

لتنعيم الإشارة وتثقيتها من الضوضاء وكالاتي

$$S^{\wedge}(n) = S(n) * W(n) \quad n = 1,2,\dots,\text{length}(s) \quad (1)$$

W(n) : نافذة هامينك ، S(n): بيانات إشارة الكلمة الداخلة .

2. تقطيع الإشارة S^{\wedge}(n) من البداية إلى النهاية إلى اطارات متداخلة ومتساوية في الحجم كل إطار يتكون من (N = 512) من العينات ويتداخل مقداره (M=256) من العينات، وكل إطار تتم معالجته بشكل منفصل بأخذ نافذة هامينك ثم حساب معدل السعة لذلك الإطار .

5-2 استخدام خوارزمية تحديد بداية الكلمة ونهايتها:

يوضح الشكل (1) المخطط الانسيابي لخوارزمية إزالة السكوت من بداية الإشارة للكلمة

* لتكن A تمثل مصفوفة حساب معدل السعة للإشارة الداخلة فنأخذ أعلى قيمة واقل قيمة من قيم معدل السعة التي يمكن تمثيلها بـ Amax و Amin على الترتيب.

* نحسب قيمة العتبة (Threshold) للكلام وكالاتي [11]:

$$I_1 = 0.03 * (A_{max} - A_{min}) + A_{min} \quad (2a)$$

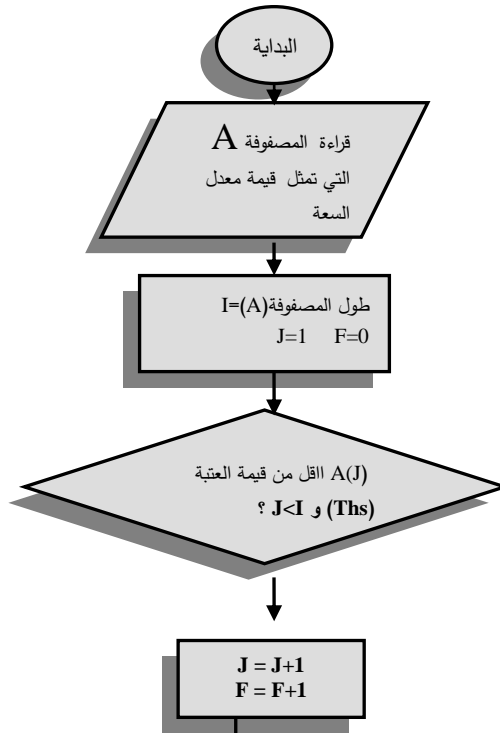
$$I_2 = 4 * A_{min} \quad (2b)$$

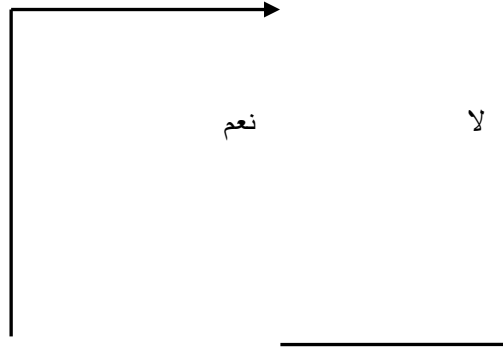
$$I_3 = \text{MIN} (I_1, I_2) + 0.4 \quad (2c)$$

* يتم حساب الإطارات التي ستتم إزالتها من بداية الإشارة $S(n)$ بالاعتماد على قيمة العتبة ، الإطارات المحذوفة من قيم معدل السعة للإشارة تمثل الإطارات التي معدل سعتها اقل من قيمة العتبة فإذا فرضنا أن F تشير إلى عدد الإطارات التي ستتم إزالتها من بداية الإشارة ومن ثم يمكن حساب عدد العينات التي تتم إزالتها على النحو الآتي:

$$\text{عدد العينات المزالة} = ((F-1)*M) + N$$

* وفي حالة حذف الصمت من نهاية الكلمة تطبق نفس الخطوات السابقة ، لإزالة الصمت من بداية الإشارة مع الأخذ بنظر الاعتبار ان $I=1$ و J طول المصفوفة (A) .
وعليه نستمر بحذف الإطارات مادامت قيمة السعة اقل من العتبة وان $I < J$ ثم تتم إزالة العينات التي تمثل الصمت من نهاية الإشارة بنفس الطريقة السابقة.





الشكل (1) المخطط الانسيابي لإزالة الصمت من بداية الإشارة

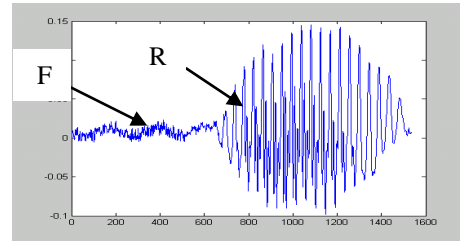
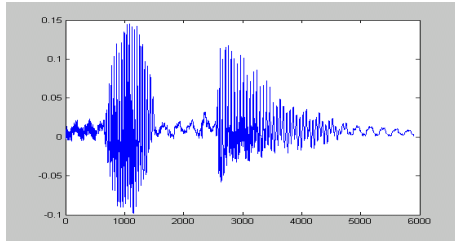
6. تقطيع الكلام أوتوماتيكياً إلى أحرف:

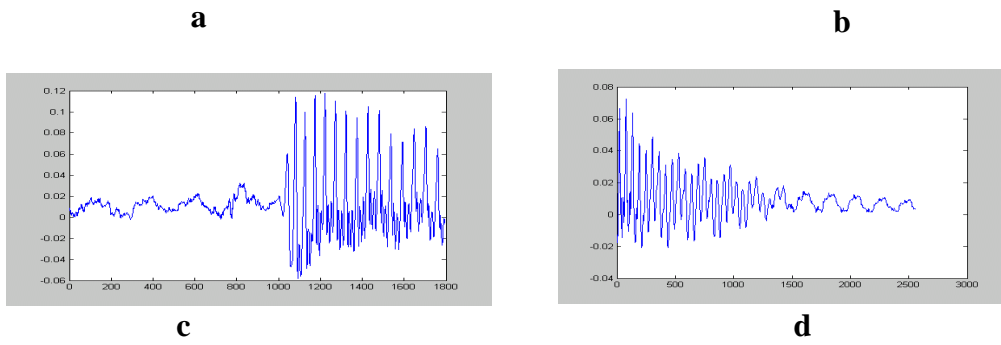
يعد تقطيع الكلام أوتوماتيكياً مرحلة مهمة في تطوير عملية تمييز الكلام لأنه سيؤدي إلى تمييز عدد أكبر من الكلمات من خلال بناء مصادر (References) للنظام تعتمد في حجمها على عدد أحرف لغة المتكلم بدلاً من بناء مصادر محددة بعدد معين من الكلمات ، وإن عملية تقطيع الكلام يدوياً هي عملية عرضة للخطأ وفيها هدر للوقت لهذه الأسباب كلها يفضل تقطيع الكلام أوتوماتيكياً .

في معظم الأحيان نستطيع من خلال النظر إلى الشكل الموجي المميز (Formant) لإشارة الكلمة أن نميز حرفاً من آخر ، فالشكل الموجي بصورة عامة يرتفع (rises) عند الانتقال من الفونيم الصامت إلى فونيم الحركة وعند الانتقال من فونيم الحركة إلى الصامت يحدث هبوط (Falls) بالشكل الموجي كما موضح في الشكل 2(a) .

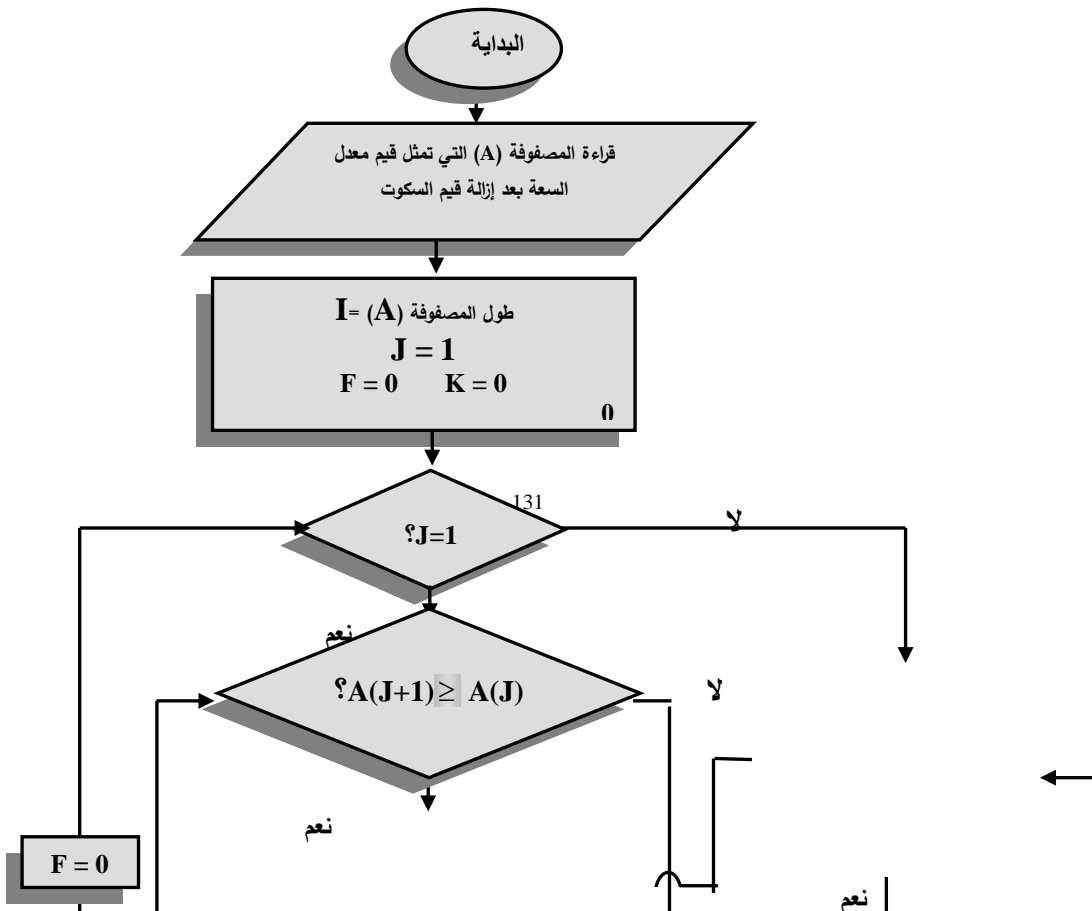
وهذا ما يعكسه بصورة واضحة معدل السعة للإشارة فبالاعتماد على معدل السعة الذي تم حسابه في الخطوات السابقة وبعد إزالة الصمت من بداية إشارة الكلمة ونهايتها نطبق خوارزمية التقطيع والشكل (3) يوضح المخطط الانسيابي لعملية التقطيع إذ نأخذ قيم معدل السعة للحرف الأول (صامت + حركة) تكون قيمتها في حالة تصاعد إلى أن يأتي نطق الصامت الثاني للكلمة فنلاحظ بدء هبوط قيم معدل السعة ونستمر بمتابعة قيم معدل السعة إلى أن تبدأ بالارتفاع ومن ثم بالهبوط لتشير إلى الصامت الآخر وهكذا مع مراعاة كون طول أي حرف لا يقل عن أربعة إطارات التي تمثل اقصر مدة زمنية لنطق أي حرف في اللغة العربية والتي تم تحديدها من خلال التجارب العملية التي قمنا بها .وبهذه الطريقة يمكن تقطيع الكلمة إلى حروفها ، كل مقطع يمثل حرف يتكون من (F) من الإطارات وكل إطار يمثل (M) من العينات لإشارة الكلمة $S(n)$ لذلك يعوض عن كل مقطع بما يقابله من إشارة الكلمة ($S(n)$) .

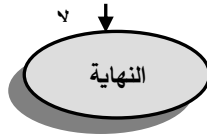
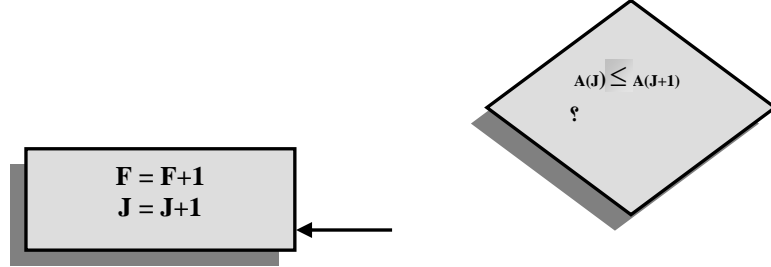
والشكل (2) يوضح إشارة كل حرف من أحرف كلمة "حسن" المقطعة





الشكل (2) a: إشارة كلمة "حسن", b: إشارة الحرف "ح", c: إشارة الحرف "س"
 F: falls R: raises d: إشارة الحرف "ن"





الشكل (3) المخطط الانسيابي لعملية التقطيع

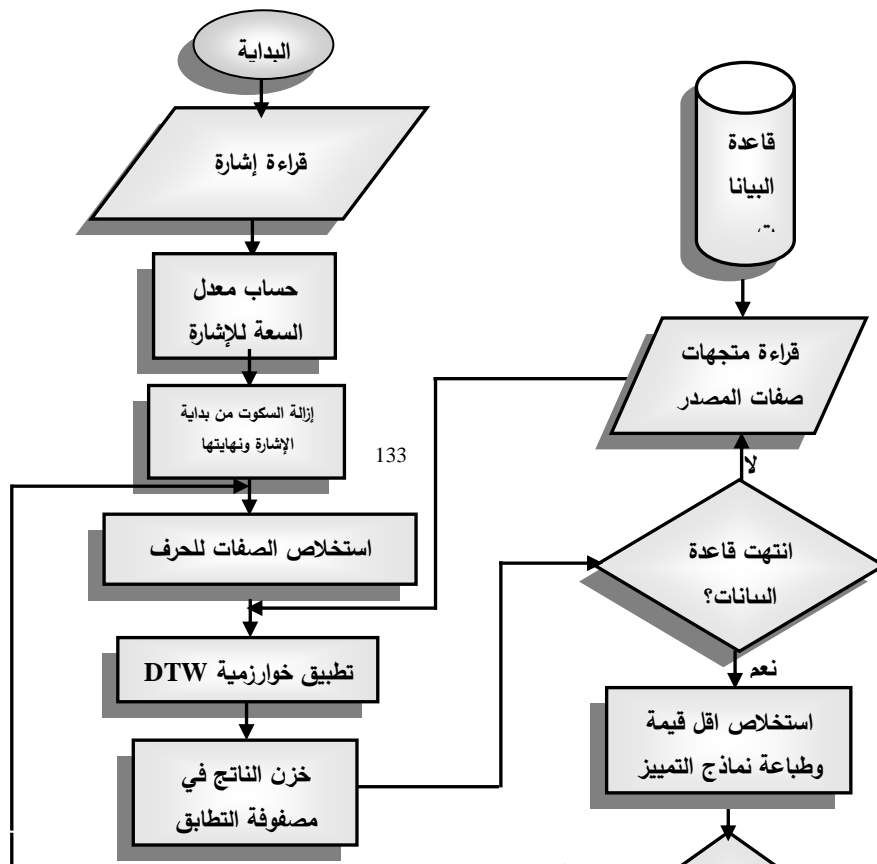
7. تحليل الإشارة وبناء قاعدة البيانات:

تحتوي الإشارة الصوتية الرقمية على عدد كبير من المعلومات غير المهمة في عملية التمييز فضلاً عن أنها تحتاج إلى مساحة خزن كبيرة، فلتسهيل خطوات المعالجة وتقليلها نستخلص الصفات الأساسية من هذه الإشارة مما يؤدي إلى تقليص حجم البيانات. في هذا البحث يتم استخلاص الصفات للإشارة الصوتية للأحرف العربية المستقطعة باستخدام طريقة التحليل (LPC) كل على انفراد وبناء قاعدة البيانات من الأحرف المحللة التي تعد المرجع الرئيس لتمييز أي حرف، في هذا البند يتم أولاً تحديد المتكلم وقراءة الملفات الخاصة بذلك المتكلم ثم يحلل كل ملف بصورة منفصلة لاستخلاص متجهات الصفات لغرض تخزينها في قاعدة

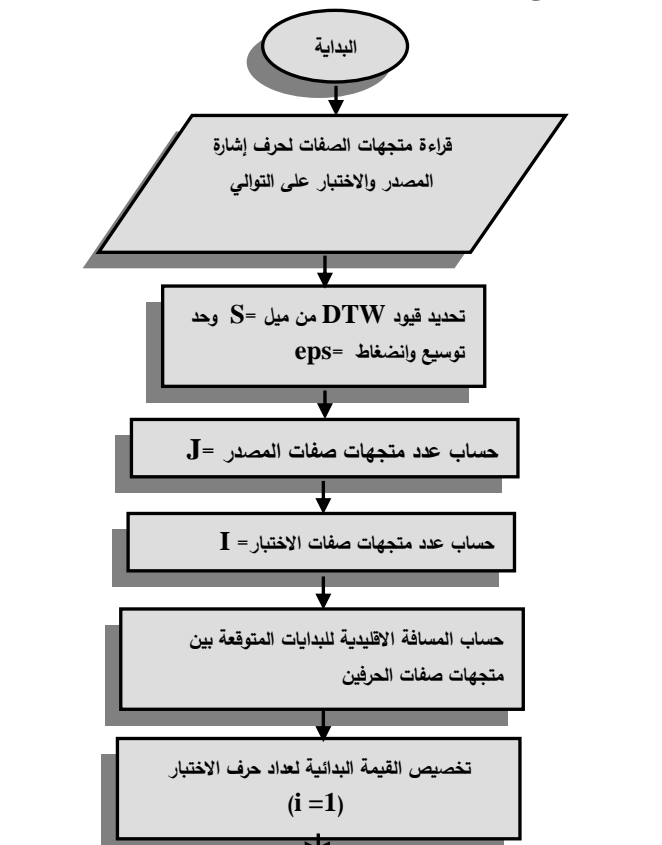
البيانات لذلك المتكلم . وتكرر هذه العملية مع جميع الأحرف لغرض استكمال قاعدة البيانات ، يتم استخلاص الصفات لكل حرف بتطبيق خطوات خوارزمية ترميز التنبؤ الخطي المذكورة آنفاً . علماً أن عدد معاملات المتجه الواحد لكل إطار من إطارات الإشارة يساوي ($P = 10$) بالاعتماد على النتائج التي حصلنا عليها خلال التجربة والنتائج موضحة في الجدول (2) .

8. مرحلة التمييز:

تتم عملية التمييز كما مبين في الشكل (4) الذي يوضح المخطط الانسيابي لعملية التمييز بعد أن يتم استخلاص الصفات لإشارات المصادر وبناء قاعدة البيانات للنظام يستلم النظام إشارة الاختبار التي سوف تمر بنفس المراحل السابقة (إزالة السكوت من بداية ونهاية الإشارة ، تقطيع إشارة الكلمة إلى أحرف ، واخيراً استخلاص متجهات صفات كل حرف من حروف الكلمة) . يتم بعدها حساب اقل قيمة تراكمية بين متجهات صفات إشارة حرف الاختبار ومتجهات صفات إشارة حرف المصدر باستخدام خوارزمية (DTW) التي ستوضح لاحقاً ، ثم تخزن النتيجة في مصفوفة التطابق المحددة ، وتكرر هذه العملية مع كل إشارات المصادر يستخرج في النهاية اقل قيمة من مصفوفة التطابق والحرف الذي يعود إلى تلك القيمة يعد ناتج التمييز .



يتم تطبيق خوارزمية انحراف الوقت الديناميكي (DTW) الموضحة في الشكل (5) حيث يوضع المخطط الانسيابي بخوارزمية (DTW) . في البداية تعرف المتغيرات وفضاء البحث الذي سوف يستخدم لتطبيق الخوارزمية ، فنستخدم سلسلة متجهات الصفات (LPC) لإشارة المصدر وإشارة الاختبار الناتجة من عملية التحليل ونطابقها وان نتيجة التطابق تمثل اقل قيمة للكلفة وهي الكلفة التراكمية لمسار التطابق . نحتاج إلى مصفوفتين لخزن إشارة المصدر وإشارة الاختبار اللتين يشار إليهما بـ Ref و Test على الترتيب .

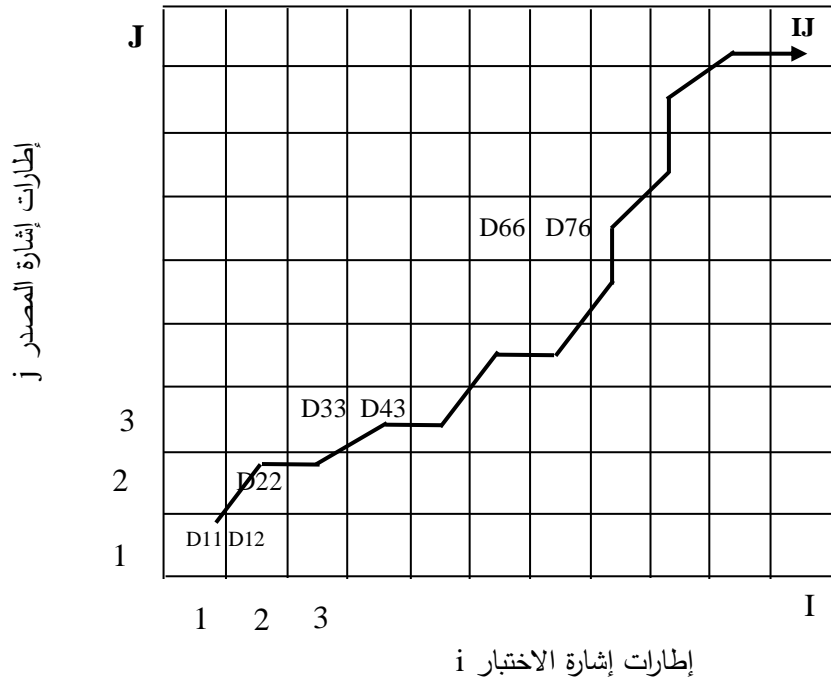




في الشكل (6) يلاحظ ان كل عمود يعكس إطار إشارة الاختبار ، وكل صف يعكس إطار إشارة المصدر، يتم الحصول على قيمة العقدة D من قياس المسافة الاقليدية ، كل عقدة تمثل المسافة بين إطار i من إشارة الاختبار والإطار j من إشارة المصدر وكالاتي :

$$D_{ij} = 1/P \sqrt{\sum_{K=1}^P (\text{ref}(k,j) - \text{test}(k,i))^2} \quad (3)$$

P: عدد المعاملات لكل إطار



الشكل (6) انحراف الوقت الديناميكي

إن هناك قيودا خاصة لخوارزمية (DTW) لتحقيق الانحراف الزمني بشكل صحيح ، في هذا البحث تم تطبيق القيود الشاملة على مسار البحث العام لحصره ضمن منطقة قانونية تقع حدودها ما بين اكبر واقل ميل (S) معين يشار إليه بـ Smax و Smin على التوالي والموضحة بالشكل (7) فضلاً عن تقييد حد الانضغاط والتمدد في إشارات الحرف لمدى معين من الوقت ويشار إليه بـ (eps) ويتم تحقيق هذه القيود الشاملة من خلال المعادلتين الآتيتين:

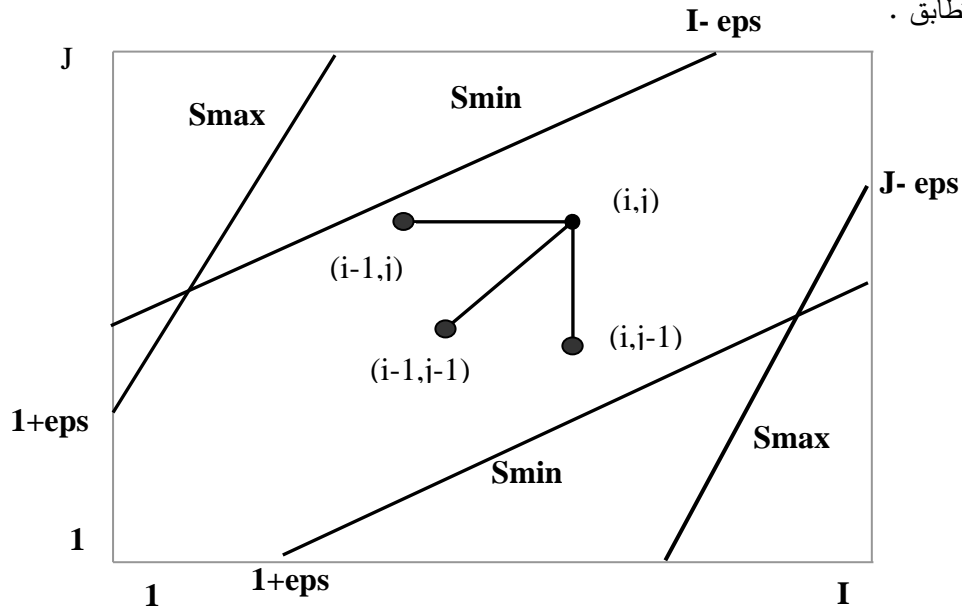
$$\text{Max } J = \text{Min} \left[\begin{array}{l} S_{\max} (i-1) + \text{eps} + 1 \\ S_{\min} (i-1) + J - S_{\min} (I-\text{eps}-1) \end{array} \right] \quad (4)$$

$$\left[\begin{array}{l} S_{\min}(i-1-\text{eps})+1 \\ \dots \end{array} \right]$$

$$\text{Min } J = \text{Max } S_{\text{max}} * i - (S_{\text{max}} * I - J) - \text{eps} \quad (5)$$

علماً إن $S_{\text{min}} = 1/S_{\text{max}}$

أما القيود المحلية التي تم استخدامها في هذا البحث فهي قيد المسار المحلي الموضح فضلاً عن قيدي الاستمرارية وثبوت الوتيرة ، تحت هذه القيود تجمع كل المسافات ذات الكلفة المتراكمة الأقل للمسارات القادمة من كل العقد السابقة الممكنة ، وتخزن النتيجة في مصفوفة التتابع .



الشكل (7) القيود المستخدمة

9. مناقشة النتائج والاسـ

1.9 عملية التقطيع

لغرض اختبار الخوارزمية المستخدمة في تقطيع الكلمة إلى أحرف قمنا باستخدام ثلاثة أطوال مختلفة للإطارات لغرض معرفة الأكفأ بينها .

فإذا فرضنا ان N يمثل حجم الإطار الواحد، و M حجم التداخل بين الإطارات فانه في الحالة الأولى ($M=128$, $N=256$) لاحظنا انه تمت تجزئة الحرف الواحد إلى عدة أجزاء . في الحالة الثانية ($M=256$, $N= 512$) كانت النتائج جيدة وكان التقطيع صحيحاً بشكل عام. اما في الحالة الثالثة ($M=512$, $N=1024$) فقد أظهرت النتائج تداخل الحرف الواحد مع الحرف الذي يسبقه أو الذي يليه .

وعليه فشل التطبيق في الحالتين الأولى والثالثة .
والنتائج التي حصلنا عليها من الحالة الثانية موضحة بالجدول (1)

الجدول (1) نتائج التقطيع

| النسبة المئوية % | عدد الكلمات المقطعة خاطئاً | عدد الكلمات المقطعة بصورة صحيحة | الأشخاص |
|------------------|----------------------------|---------------------------------|----------|
| 88.5 | 4 | 31 | Male 1 |
| 85.7 | 5 | 30 | Male 2 |
| 82.8 | 6 | 29 | Female 1 |
| 82.8 | 6 | 29 | Female 2 |

9.2 عملية التمييز

أما طريقة التمييز فقد تم اختبارها بوساطة قيم مختلفة للمحددات إذ أن P تمثل عدد معاملات المتجه الواحد لكل قالب من قوالب متجهات الصفات للإشارة ، ϵ هي حد الانضغاط والتمدد في الوقت ، و S_{max} و S_{min} تمثل القيمتين العظمى والصغرى للميل (S) على التوالي- عند اختبار كل حالة من حالات الجدول (2) يلاحظ جودة النتائج في الحالتين الأولى والثانية وان إنقاص P من 10 إلى 8 لا يؤثر في نتيجة التمييز، أما الحالات الأخرى فان إنقاص قيم الميل والـ ϵ أدى إلى انخفاض دقة التمييز كذلك فإن ارتفاع الـ ϵ إلى 5 كان له تأثير في دقة نتائج التمييز ، لهذا فان اختيار قيم المحددات الشاملة له تأثير كبير على دقة النتائج لدورها بتحديد منطقة البحث القانونية أي استبعاد فضاء البحث غير المعقول بحيث تتضمن هذه المنطقة اكثر احتماليات المسار الممكنة.

الجدول (2) حالات التمييز

| الحالات | P | S_{max} | S_{min} | ϵ |
|----------------|----|-----------|-----------|------------|
| الحالة الأولى | 10 | 3 | 0.333 | 3 |
| الحالة الثانية | 8 | 3 | 0.333 | 3 |
| الحالة الثالثة | 10 | 3 | 0.333 | 2 |
| الحالة الرابعة | 10 | 2 | 0.25 | 3 |
| الحالة الخامسة | 10 | 3 | 0.333 | 0 |

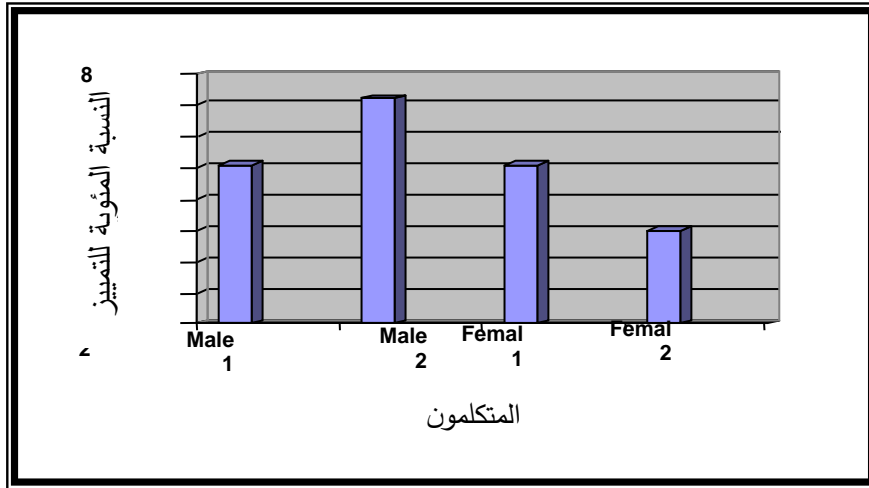
| | | | | |
|---|-------|---|----|----------------|
| 5 | 0.333 | 3 | 10 | الحالة السادسة |
|---|-------|---|----|----------------|

والجدول (2) يوضح حالات التمييز

والجدول (3) يظهر نتائج تمييز الأحرف بالاعتماد على قيم المحددات للحالة الأولى والتي تم استخدامها في هذا البحث والموضحة بالشكل (8)

الجدول (3) نتائج تمييز الأحرف

| النسبة المئوية % | عدد الأحرف الخطأ | عدد الأحرف الصحيحة | الحالات |
|------------------|------------------|--------------------|----------|
| 77.1 | 22 | 74 | Male 1 |
| 80 | 19 | 76 | Male 2 |
| 77.1 | 22 | 74 | Female 1 |
| 75 | 24 | 72 | Female 2 |



الشكل (8) دقة نتائج التمييز للمتكلمين

3-9 الاستنتاجات

تم من خلال هذا البحث التوصل إلى الاستنتاجات الآتية :

1. لإزالة الصمت (Silence) من بداية الإشارة ونهايتها تأثير كبير في دقة النتائج.
2. إن اختيار نسبة التعيان لها تأثير كبير في عملية التمييز ، فزيادة نسبة التعيان تؤثر تأثيراً فعالاً في زيادة دقة التمييز والعكس بالعكس .

3. ان معالجة الإشارة باستخدام طريقة التحليل (LPC) فضلاً عن سرعتها ودقة نتائجها تزودنا بنموذج جيد لإشارة الكلام وذلك بإعطائها افضل تمثيل ممكن للإشارة وهذا يقود إلى دقة أعلى في أداء التمييز .
4. إن انحراف الوقت الديناميكي (DTW) هي طريقة رياضية مرنة وتعطي نتائج ذات دقة عالية في التمييز .
5. بالرغم من وجود عدة طرائق لحساب المسافة لكن لاحظنا ان المسافة الاقليدية المستخدمة في الحسابات أثبتت فعاليتها في هذا المجال
6. استخدام قيود تطبيع الزمن (Time normalization constraints) يضمن التزامن بشكل صحيح بين إشارة الاختبار وإشارة المصدر المختلفتين في الطول.

المصادر

- (1) بدري ، كمال إبراهيم ، (1988) ، " علم اللغة المبرمج ، الأصوات والنظام الصوتي مطبقا على اللغة العربية"، معهد اللغة العربية ، جامعة الملك محمد بن سعود الإسلامية.
- (2) حلیم ، علياء موفق عبد المجید ، (2003) "تشفير إشارة الكلام بطريقة البعثة" ، بحث ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل .
- (3) الدليمي ، شهلة عبد الوهاب ، (2002) ، "تمييز أصوات الحروف العربية" بحث ماجستير ، علوم حاسبات ، كلية علوم الحاسبات والرياضيات ، جامعة الموصل.
- [4] AL-Ani, S.H., (1970), **Arabic Phonology an Acoustical and Physiological Investigation**, Indiana University, Mouton and Co.N.V. Publishers, The Hague.
- [5] Coust, A.P., (2002), "**Connected Word Recognition**", [http:// ska.ai-depot.com/paper/nodeq.html](http://ska.ai-depot.com/paper/nodeq.html).
- [6] Land, J., (2000), "**Isolated Word Speech Recognition of English Digits**", <http://www.gerc.eng.ufl.edu/default.htm>
- [7] Riadh, M.H., (2001), "**Information Retrieval using Spoken Language Interface. (Isolated)**", Msc.Thesis computer science, National computer center/ Institute of Higher studies in Computer and information.
- [8] AL – Sayegh, S.W., (2002), "**Arabic Phoneme Recognizer Based on Hybrid Neural Network**", Ph.D. Thesis, information institute for postgraduate studies at in computer science Iraq Commission for computer and Information.
- [9] Sofia, F.B., (2003), "**Mathematical model for ARABIC Speech Recognition**", Ph.D.Thesis in Mathematics, The College of Computer Sciences and Mathematics, University of Mosul.
- [10] (2003), "**Speaker Dependent /Speaker Independent**", [http://www.Imagesco.com/articles/hm_2007/Speech Recognition Tutoriat02.htm](http://www.Imagesco.com/articles/hm_2007/Speech%20Recognition%20Tutoriat02.htm) Speech Recognition circuit.
- [11] Rabiner, L.R ;Sambur,M.R.,(1975), " **An algorithm for determining the end point of isolated utterances**",*American telephone and telegraph company* .