

## مقارنة الطرق الإحصائية والنماذج السمعية لتخفيض الصفير في المعينات السمعية الثنائية لمرضى نقص السمع الشخي

د. م. فطمة الطراب<sup>(1)</sup>

### الملخص

تستخدم المعينات السمعية الثنائية بشكل واسع في تدبير حالات نقص السمع الثنائي الجانب، ومن بينها نقص السمع الشخي ثنائي الجانب، فتقوم بتضخيم الأصوات بغض النظر عن مصدرها، وبما أن حالات نقص السمع الشخي تتطلب زيادة الريح مع تراجع السمع التدريجي، فهذا يؤدي إلى احتمالية أكبر لحدوث الصفير مما ينعكس سلباً على أداء المعينات السمعية. وتعد مشكلة الصفير عائقاً كبيراً عند تصميم المعينة السمعية لأنها تزداد بزيادة الريح. وقد قام العديد من الباحثين بمعالجتها بطرق مختلفة باستخدام المرشحات الموائمة والنماذج السمعية والطرق الإحصائية، وأظهرت هذه الطرق أداءً مقبولاً في تخفيض الصفير بالنسبة للمتحدثين باللغة الإنكليزية والفارسية [1-5,12,23].

نقدم في هذا البحث مقارنة عملية بين الطرق الإحصائية والنماذج السمعية لتخفيض الصفير في المعينات السمعية الثنائية لمرضى نقص السمع الشخي باستخدام بيانات كلام باللغتين العربية والإنكليزية، ومن كلا الجنسين، ويقوم ريح تمتد من 75 ديسبل حتى 90 ديسبل الخاصة بمرضى نقص السمع الشخي الشديد والعميق. بهدف تحديد الطرق الأكثر فعالية في تخفيض الصفير في البلاد التي تتحدث باللغة العربية. أظهرت النتائج أن أداء طريقة فصل المنبع الأعمى الإحصائية (BSS+GMM) أفضل من طريقة النماذج السمعية (CASA) من خلال التحسن في القيم المتوسطة لكل من نسبة الإشارة الهدف للضجيج (SNR)، وكذلك نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR)، وذلك لكل من اللغة العربية والإنكليزية معاً، كما وتسمح هذه الطريقة للمستخدم بزيادة ربح المعينة السمعية دون أن تتأثر جودة الصوت المسموع.

الكلمات المفتاحية: المعينات السمعية الثنائية - النماذج السمعية - الصفير - الطرق الإحصائية.

<sup>(1)</sup> مدرسة أصيلة في قسم الأذنية - كلية الطب البشري - جامعة دمشق.

## Comparison of Statistical Methods and Auditory Models for Acoustic Feedback Suppression in Binaural Hearing Aids For Presbycusis's patients

Dr.Fatmah Tarrab<sup>(2)</sup>

### Abstract

Binaural hearing aids are widely used in managing binaural hearing loss, including Presbycusis, they amplifies sounds regardless of their source, Presbycusis require increased gain along with a gradual decline in hearing, this leads to a greater probability of occurrence of wheezing, which is affecting the performance of hearing aids.

The acoustic feedback is a major problem when designing a hearing aid because it increases with increasing gain. Many researchers have reduced it in different ways using Adaptive Filters, Auditory Models, and Statistical Methods. These methods show an acceptable performance in feedback reduction for English and Persian speakers [1-5,12,23].

In this paper, we present a comparison between statistical methods and auditory models for Acoustic Feedback Suppression in binaural hearing aids for Presbycusis's Patients using Arabic and English speech data of both sexes and with gain from (75 to 90 dB) for patients with severe and profound Presbycusis.

The aim of this paper is identifying the most Effective method for Acoustic Feedback Suppression in Arabic-Speaking Countries.

The results showed that the performance of the statistically blind source separation method (BSS + GMM) is better than the performance of auditory models (CASA) by improving the average values for both the signal to noise ratio (SNR) and the signal-to-feedback ratio (SFR) for Arabic and English speech data. Moreover, (BSS+GMM) allowed the user to increase the gain of the hearing aid without affecting the sound quality.

**Keywords:** Bilateral Hearing Aids, Auditory Models, Acoustic Feedback, Statistical Methods.

---

(2) Lecture - ENT Department -Faculty of Medicine-Damascus University

**1- المقدمة:**

يعدُّ نقص السمع الشيخي (presbycusis) أكثر أنواع نقص السمع شيوعاً عند كبار السن، ويكون ثنائي الجانب، ويحدث عند التقدم بالعمر بسبب تنكُّس الخلايا الشعرية أو العصب السمعي، وتتأثر بشكل رئيس بالتواترات العالية ومن ثم يتراجع السمع تدريجياً ببطء، ولا يوجد علاج شاف لنقص السمع الشيخي، وينصح باستخدام معينات سمعية ثنائية للحفاظ على الانتقائية الاتجاهية للنظام السمعي، ولتعويض توازن الأذنين، مما يحسن نسبة الكلام إلى الضجيج بين الأذنين [7,8,35]. إلا أنه يمكن أن ينخفض أداء المعينات السمعية عندما يزداد ربح المعينة السمعية كما في حالات نقص السمع الشيخي الشديد والعميق، فيحدث الصفير، حيث تتسبب مشكلة الصفير في تشويه الصوت المسموع، وقد تؤدي إلى عدم القدرة على السمع مع زيادة ربح المعينة السمعية لتعويض نقص السمع، وتنعكس هذه المشكلة سلباً على المريض، فينزعج من المعينة السمعية، ويتخلّى عنها [1-6,34].

تم عبر السنوات الماضية معالجة ظاهرة الصفير في أنظمة المعينات السمعية بشكل عام باستخدام تقنية المرشحات الموائمة، وعلى الرغم من الاستخدام الواسع لتقنية المرشحات الموائمة في المعينات السمعية، إلا أنها لا تزال تعاني من مشكلة الانحياز، مما يؤثر في أداء نظام إلغاء الصفير، ومن ثم تشويه خرج المعينة السمعية [5,6,22]. كما قامت أبحاث أخرى بمعالجة مشكلة الصفير في أنظمة المعينات السمعية الثنائية باستخدام تقنيات فصل المصدر الأعمى مع المرشحات الموائمة بمفهوم تقنية (blind source separation: BSS)، وتحليل النظام تبين أنه لا يزال يعاني من مشكلة الانحياز في حل وينر [9]. وقام باحثون آخرون بمعالجة الصفير في المعينات السمعية الثنائية باستخدام تقنية فصل المنبع

الأعمى مع خوارزميات التعليم غير الموجه، وهي طريقة إحصائية تعتمد على الاحتمالات في عملية الفصل، وحصلت هذه الطريقة على نتائج مقبولة في تخفيض الصفير، ونجحت في تحسين نسبة الإشارة إلى الضجيج [4,10]. وفي الأعوام الأخيرة انتشرت تقنية حديثة لتخفيض الصفير تعتمد على النماذج السمعية للإنسان بشكل مشابه لعمل النظام السمعي للإنسان، وهي تقنية التحليل الحسي السمعي الحسابي (CASA) [11,14].

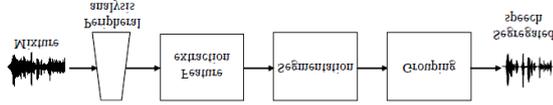
يقدم هذا البحث مقارنة عملية بين الطرق الإحصائية والنماذج السمعية في تخفيض الصفير في المعينات السمعية الثنائية لمرضى نقص السمع الشيخي باستخدام بيانات كلام باللغتين العربية والإنكليزية، ومن كلا الجنسين، وقيم ربح تمتد من 75 ديسبل حتى 90 ديسبل، بهدف تحديد الطرق الأكثر فعالية في تخفيض الصفير لاستخدامها في المعينات السمعية الثنائية في البلاد التي تتحدث اللغة العربية.

**2- منهج البحث:****2-1 الطرق الإحصائية في تخفيض الصفير:**

تعد تقنية فصل المنبع الأعمى تقنية إحصائية غير مستوحاة من النظام السمعي، ومع هذا تُستخدم هذه التقنية لفصل مزيج من الإشارات الصوتية، وقد حققت نجاحاً كبيراً في عمليات فصل إشارات الضجيج عن الإشارة الهدف في العديد من التطبيقات الصوتية [13,17,25]. ويبين الشكل (1) المخطط العام لتقنية فصل المنبع الأعمى. وقد سميت بالأعمى كون مزيج المصادر الصوتية وعملية المزج غير معروفين.

التحليل الحسي السمعي عند الإنسان بطرق حسابية فأطلقوا عليها تقنية التحليل الحسي السمعي الحسابي (CASA) [14,21].

ويظهر الشكل (2) مخططاً عاماً لمراحل العمل في نظام (CASA)، والتي تحاول تقليد عمل النظام السمعي عند الإنسان عندما يستقبل الأصوات من مصادر مختلفة، ويتكون هذا النظام من أربعة مراحل رئيسية: وهي مرحلة التحليل المحيطي (peripheral analysis)، ومرحلة استخراج السمات (Feature extraction)، ومرحلة التقطيع (Segmentation)، ومرحلة التجميع (Grouping) [11].



الشكل (2) مخطط مراحل العمل في نظام CASA [11]

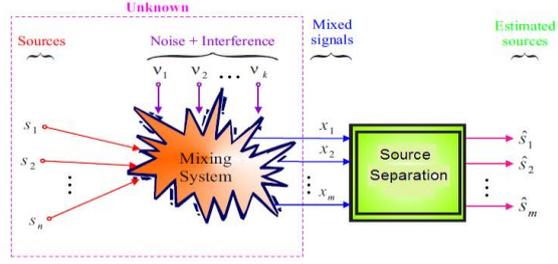
### 3- تصميم الخوارزميات المستخدمة في

#### البحث:

تم في هذا البحث تصميم خوارزميتين، الأولى: لمحاكاة طريقة فصل المنبع الأعمى، والثانية: لمحاكاة طريقة التحليل الحسي السمعي الحسابي عبر برنامج ماتلاب، وذلك عبر ميكروفونين لفصل الإشارة الصوتية المسببة للصفير عن إشارة الكلام حسب الشكلين (1) و (2) على الترتيب، وهذا البرنامج سيمكننا من اختبار أداء الطريقتين في تخفيض الصفير في المعينات السمعية الثنائية لمعرفة، أي الطريقتين أفضل أداءً في البلاد التي تتحدث اللغة العربية. وقد تم إعداد برنامج المحاكاة كما يأتي:

#### 3-1 بيانات الدخل:

يتكون دخل النظام الصوتي من مجموعة من إشارات الكلام تحوي مزيج الإشارة الهدف والإشارة الصوتية المسببة للصفير بتردد تقطيع (16 كيلوهرتز)، حيث إن إشارات الكلام عبارة عن تسجيلات كلام متنوعة باللغتين العربية والإنكليزية ولهجات مختلفة ومسجلة بمواقع مختلفة



الشكل (1) مخطط توضيحي لتقنية فصل

#### المنبع الأعمى [13]

ووفقاً للعديد من الأبحاث العلمية تعدُّ تقنية فصل المصدر الأعمى الطريقة الواعدة لتخفيض الصفير مقارنة مع الطرق السابقة في فصل المنابع بسبب قوتها في عملية الفصل وقلة الافتراضات عن المصادر الصوتية، ولا تتطلب معلومات سابقة (مثل خصائص إشارات المصدر، أو مصفوفة المزج، أو ترتيبات الميكروفونات). كما وتمكننا هذه التقنية من استخدام ميكروفونين لنحاكي عمل النظام السمعي للإنسان، وهي تمثل تطبيقات المعينات السمعية الثنائية، كما وتعمل هذه التقنية على استعادة إشارات المصادر الأساسية المكونة للمزيج المراقب عبر الميكروفونات فقط، لذا لا تحتاج إلى كاشف صوتي (voice-activity detector) [13,10,15].

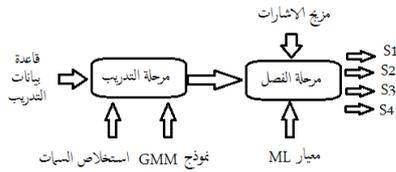
فيمكننا تطبيق تقنية (BSS) على تشكيلة واسعة من الأنظمة الصوتية ومن بينها تخفيض الصفير في المعينات السمعية الثنائية.

#### 2-2 النماذج السمعية في تخفيض الصفير:

يمتاز النظام السمعي عند الإنسان بقدرته على تمييز الأصوات ضمن غرفة تحوي مزيج مختلف من الأصوات. فعندما يصل الصوت إلى أذني الإنسان فإنه يخضع لتحليل حسي سمعي (ASA: auditory scene analysis) عبر مرحلتين التقطيع والتجميع إلى أن يتم إدراك الصوت في القشرة السمعية [14,20]. وقامت أبحاث عديدة بدراسة هذا

حيث يدل الرمز (R) على الأذن اليمين والرمز (L) على الأذن اليسار. أي تصبح إشارة كل ميكرفون عبارة عن مزيج من إشارة الكلام  $S(n)$  والإشارة الصوتية المسببة للصفير  $F(n)$ .

وتتم عملية الفصل بتقنية فصل المنبع الإحصائية في هذه المرحلة باستعادة الإشارة الهدف والإشارة الصوتية المسببة للصفير من مزيجها عند كل ميكرفون. وستتم عملية الفصل وفق مرحلتين: مرحلة التدريب، ومرحلة الفصل كما هو مبين في الشكل (3)، والذي يظهر مخطط مراحل العمل في تقنية فصل المنبع الأعمى الإحصائية.



الشكل (3): مخطط العمل بتقنية فصل المنبع الأعمى الإحصائية

#### 1- مرحلة التدريب:

سيتم التدريب في هذه المرحلة على مجموعتي تدريب باللغة العربية واللغة الإنكليزية، وتتضمن (16) ملفاً صوتياً، كلٌّ منها عبارة عن مزيج من الإشارة الهدف والإشارة الصوتية المسببة للصفير من ضمن مجموعة الملفات الصوتية المأخوذة من قاعدة البيانات، وتتضمن مرحلة التدريب الخطوات الآتية:

#### أ- استخراج السمات:

يتم في هذه المرحلة تحليل طيفي للإشارات الصوتية المأخوذة من قواعد بيانات التدريب. واستخراج بارامترات المجال الترددي من إشارة الكلام، ومن ثم تحويل كل إطار إلى شعاع خصائص واحد. باستخدام تحويل فورييه قصير المدى STFT كما في المعادلة الآتية:

$$X_i(k, L) = \sum_{T=-1/2}^{1/2} \text{win}(T) x_i(T + kT_0) e^{-jL\omega_0 T} \quad (3)$$

وبأصوات من كلا الجنسين. وهذه القاعدة الصوتية مأخوذة من قاعدة بيانات (TIMIT)، وهي مصممة خصيصاً من أجل اختبار وتقييم أنظمة تعرف الصوت واستحصال الإشارات في أبحاث معالجة الصوت [29-33].

وتتم الحصول على الإشارة الصوتية المسببة للصفير بنرشح الإشارة الهدف (الكلام) لكل ميكرفون وتضخيمها بقيمة ربح مختلفة تمتد من (75 ديسبل) إلى (90 ديسبل) بخطوة مقدارها (5 ديسبل) لمحاكاة ربح المعينة السمعية في تضخيم الإشارات لمرضى نقص السمع الشيخي الشديد والعميق.

#### 2-3 بارامترات الخوارزميتين المستخدمتين في هذا

#### البحث:

#### أ- خوارزمية فصل المنبع الأعمى الإحصائية:

تم اعداد الخوارزمية الأولى في هذا البحث لمعالجة مشكلة الصفير باستخدام تقنية فصل المصدر الأعمى واستخدام نموذج غوص الإحصائي. وحسب مفهوم تقنية (BSS) سيكون لدينا في كل معينة سمعية مزيج واحد مراقب بمصدرين للتقدير، وبما أنه لدينا في حالتنا ميكرفونين سيكون لدينا مزيجين مراقبين أحدهما للأذن اليمين والآخر للأذن اليسار بأربعة مصادر إشارة للتقدير، وهي الإشارات الواصلة لكل من الأذن اليمين واليسار، بالإضافة إلى الإشارات الصوتية المسببة للصفير. فيمكن كتابة الدخل الصوتي لكل أذن على حدة في المجال الترددي كما يأتي:

$$X_L(n) = \sum_{p=0}^{p-1} h_{ji}(p) s_i(n - p_L) + f_L(n) \quad (1)$$

$$= S_L(n) + F_L(n)$$

$$X_R(n) = \sum_{p=0}^{p-1} h_{ji}(p) s_i(n - p_R) + f_R(n) \quad (2)$$

$$= S_R(n) + F_R(n)$$

**1. E-step** : يتم في هذه المرحلة حساب الاحتمالات المتوقعة لكل نقاط البيانات الصوتية من أجل جميع الصفوف الصوتية وفق العلاقة :

$$P(\omega_i \setminus x) = \frac{P(x \setminus \omega_i) \cdot P(\omega_i)}{P(X)} \quad (4)$$

حيث:

$P(\omega_i)$  : تعطينا احتمال ظهور كل منبع صوتي قبل قياس أي سمة .

$P(x \setminus \omega_i)$ : يمثل احتمال الصف الصوتي الذي يعطي شعاع السمة المشاهد. و  $P(\omega_i \setminus x)$ : الاحتمالات السابقة. طبعاً نحن لا نعرف  $P(\omega_i \setminus x)$ ، إلا أنه يمكننا تقدير  $P(x \setminus \omega_i)$  من بيانات التدريب، لذلك يعدّ نموذج غوص أحد الطرق المناسبة لحساب  $P(x \setminus \omega_i)$ . حيث يتم تمثيل كلّ صنف صوتي بتوزع غوص، ومن ثم حساب الاحتمال بقياس قيمة التابع. ويتم من خلال بيانات التدريب تدريب كل من  $P(x \setminus \omega_i)$  و  $P(\omega_i)$ . ومن ثم يتم استخدام هذه المعلومات لفصل البيانات الجديدة في المرحلة الآتية. اما  $P(X)$  فهو نفسه لجميع الصفوف الصوتية ويحسب من العلاقة الآتية [27]:

$$P(X) = \sum_{i=1}^K P(x \setminus \omega_i) \cdot P(\omega_i) \quad (5)$$

**2. M-step**: يتم في هذه المرحلة إعادة تقدير البارامترات باستخدام قيم المتغيرات المخفية الحالية كما في العلاقتين (6) و (7) [27]:

$$\mu_i(t+1) = \frac{\sum_k P(w_i | x_k, \lambda_i) x_k}{\sum_k P(w_i | x_k, \lambda_i)} \quad (6)$$

$$\Sigma_i(t+1) = \frac{\sum_k P(w_i | x_k, \lambda_i) [x_k - \mu_i(t+1)][x_k - \mu_i(t+1)]^T}{\sum_k P(w_i | x_k, \lambda_i)} \quad (7)$$

وفي نهاية مرحلة التدريب نحصل على بارامترات نموذج مزيج غوص  $P(x \setminus \omega_i)$  والاحتمالات الآلية  $P(\omega_i)$ ,

حيث يدل  $(i)$  على عدد الميكروفونات و  $K = \{0, \dots, K\}$  دليل الإطار الزمني،  $L = \{0, \dots, L-1\}$  دليل القناة الترددية،  $\omega_0$  تابع النوفذة و  $T_0$  بارامترات دقة شبكة التمثيل الترددي - الزمني T-F. وعندها نحصل على السمات الطيفية للإشارة الصوتية.

### ب- النموذج الإحصائي:

في هذه المرحلة نقوم بنمذجة تابع الكثافة الاحتمالية المشترك لطيف إشارات الميكروفونين اليمين واليسار بنموذج مزيج غوص لترجمة شعاع السمات الصوتية  $(x)$  المعبر عن الإشارة الصوتية. ويعدّ نموذج مزيج غوص (GMM) Gaussian Mixture Model أحد الطرق الفعالة من أجل نمذجة البيانات وتصنيفها، ويمتاز من غيره من النماذج الإحصائية أنه لا يتطلب أن يكون عدد المكونات الغوصية مساوٍ لعدد المصادر فيمكن تطبيقه في حالة المعينات السمعية الثنائية، كما ويؤخذ بالحسبان التركيب المتنوع للأصوات. مما يسمح لنا بفصل الإشارة الصوتية المسببة للصفير لأن الإشارة التي سيتم معالجتها في بحثنا (الكلام) هي إشارة عشوائية وغير ثابتة [26].

ويكون في هذه المرحلة عدد السمات المحسوبة من المكونات الغوصية مساوياً لثلاثة أضعاف عدد المكونات الغوصية. بحيث يكون طول شعاع السمة لـ  $(Q)$  مكون غوصي هو  $(Q+1) \cdot 3$ . وسنستخدم الخوارزمية التكرارية المعروفة بالتوقع الأعظمي (Expectation-Maximization) EM للحصول على مجموعة البارامترات الأمثل التي توافق نموذج مزيج غوص الممثل للإشارات الصوتية لكل صنف صوتي من طيف الكلام. ويتم الحساب في خوارزمية EM وفق خطوتين بشكل تكراري [27]:

## ب- خوارزمية تقنية التحليل الحسي السمعي الحسابي (CASA):

تم إعداد هذه الخوارزمية في برنامج ماتلاب، حيث يتم أولاً تحويل إشارات المزيج المراقبة عبر الميكروفونات عند معدل تقطيع (6 كيلوهرتز) إلى المجال الترددي-الزمني باستخدام مرشح غاما تون ثنائي ذو الترتيب (4) بمجال ترددات مركزية يمتد من (50 هرتز) إلى (8000 هرتز)، ومن ثم يتم حساب الكوليغرام لكل مزيج عند كل ميكروفون لاستخراج السمات، ليتم في المرحلة الأخيرة الحصول على الإشارة الثنائية الصافية. وفيما يأتي نبين مراحل هذه العملية بالتفصيل:

### مرحلة التحليل المحيطي:

تمثل هذه المرحلة وظيفة الغشاء القاعدي في الحلزون، سنستخدم في هذا البحث تمثيلات الكوليغرام (Cochleagram) والتي تم تقديمها من قبل الباحثين (brown and wang) عام (2005) واستخدمت في العديد من أنظمة (CASA) [28]. وهذا التمثيل في الكوليغرام يعطي توازناً جيداً بين الدقة والفعالية الحسابية، فهو يحتوي العديد من خصائص استجابات العصب السمعي، ويمتاز كونه قريباً من التمثيلات السمعية عند الإنسان، ويعطي دقة ترددية عند الترددات المنخفضة أدق من الترددات العالية، كما ويمكننا بناء وإعادة تشكيل الإشارة من الكوليغرام من أجل الاستفادة منها في مرحلة التقييم. لذلك سنقوم في مرحلة التحليل الترددي باستخدام بنك من مرشحات غاما تون (gamma tone filter bank) لتمثيل الغشاء القاعدي في كل من الأذنين اليمين واليسار، وهي عبارة عن مرشحات تمرير حزمة تم اشتقاقها من الخواص الفيزيائية الصوتية للجهاز السمعي المحيطي للإنسان. وتعد هذه المرشحات نموذجاً قياسياً للترشيح الحلزوني. ويعتمد عرض حزمة هذه المرشحات على التردد، وتعطى

والتي سيتم استخدامها في المرحلة الآتية لفصل البيانات الجديدة.

### 2- مرحلة الفصل:

سنستخدم في مرحلة الفصل إحدى تقنيات التقدير المعروفة (maximum likelihood estimation (MLE)). لتفكيك مكونات مزيج إشارات الميكروفونين اليمين واليسار  $X_L, X_R$  بالمفهوم الاحصائي. وذلك بتحليل مكونات

المزيج الغوصي للنماذج والتي توافق معيار الاحتمال الأعظمي للبارامترات من أجل الحصول على القيمة الأعظمية لاحتمال نموذج المزيج الصوتي  $P(X/\omega_i)$ . فمن أجل أي مزيج جديد يحوي إشارة صوتية مسببة للصغير يتم نسبها للمصدر الصوتي الموافق الذي يملك القيمة الأعظمية عبر حساب الاحتمال لبيانات التدريب لكل صنف صوتي  $\omega_i$ . كما هو مبين بالعلاقة الآتية [27]:

$$\omega_i = \arg \max [P((x_i/\omega_i)P(\omega_i))] \quad (8)$$

حيث:  $X_i$ : الإشارة المشاهدة عبر الميكروفون (i).

### 3- إعادة تشكيل إشارات المصدر:

بعد الحصول على الصفوف الصوتية لكل من الإشارة الهدف والإشارة الصوتية المسببة للصغير سنقوم بإعادة تركيب إشارات الكلام المقدر في المجال الزمني من أجل استخدامها في مرحلة التقييم. ونعبر عن ذلك بالعلاقة الآتية:

$$M_n(t, f) = p(x \mid \mu, \Sigma) \quad (9)$$

حيث تمثل ( $\mu$  و  $\Sigma$ ) كلاً من المتوسط ومصفوفة التباين على الترتيب لنموذج المزيج الغوصي ذو  $p$  مكون .

وبعدها يتم حساب مقدر الصورة الخاص  $\widehat{S}_{mn}$  (n) إشارة مستقبلية عند (m) ميكروفون بتطبيق  $M_n(t, f)$  على عدد المشاهدات (m) كما يأتي :

$$\widehat{S}_{mn}(t, f) = M_n(t, f) \cdot X_m(t, f) \quad (10)$$

حيث:  $n=1, \dots, N$

(CASA) [14,21,20]. ويتم في هذه المرحلة حساب النماذج الثنائية (ILD, ITD) عبر كل حزمة ترددية ناتجة في مخطط الكوكليغرام السابق باستخدام تابع (cross-correlogram) كما هو مبين في العلاقة (14):

$$R_i(t, \tau) = \frac{\sum_{k=0}^{k-1} \bar{l}_i(t.k/2-k) \bar{r}_i(t.k/2-k-\tau)}{\sqrt{\sum_{k=0}^{k-1} \bar{l}_i^2(t.k/2-k)} \sqrt{\sum_{k=0}^{k-1} \bar{r}_i^2(t.k/2-k-\tau)}}$$

حيث تمثل  $(\bar{l}_i)$  و  $(\bar{r}_i)$  القيم المتوسطة لإشارات الأذن اليمين واليسار على الترتيب، والتي يتم تقديرها لكامل عدد الإطارات (n) و (W): تابع النافذة. وقيمه العظمى تستجيب لـ (ITD) المقدر (بالمعينات) وتعطى بالعلاقة (15):

$$\hat{\tau}_i(t) = \arg \max_{\tau} C_i(t, \tau). \quad (15)$$

ويتم حساب الاختلاف في مستوى الشدة بين المزيجين الصوتيين اليمين واليسار للمحافظة على اتجاه الصوت الهدف بمقارنة الطاقة المكاملة عبر فترات زمنية محددة بين الأذن اليمين واليسار. ويتم التعبير عن نموذج (ILD) في قناة غاما تون بالديسبل وحسب بالعلاقة الآتية:

$$ILD_i(t) = 20 \log_{10} \left\{ \frac{\sum_{k=0}^{k-1} r_i^2(t.k/2-k)}{\sum_{k=0}^{k-1} l_i^2(t.k/2-k)} \right\} \quad (16)$$

## 2- مرحلة التقطيع (الفصل):

بعد أن قمنا في المرحلة السابقة بتوليد النماذج الثنائية (ILD, ITD) لنظام (CASA)، سنقوم في هذه المرحلة بدمجها كسمات ونمذجتها لحساب تابع الكثافة لوحداث التمثيل الزمني الترددي في الكوكليغرام، وسنقوم في هذه المرحلة بفصل مصدرَي الصوت (الكلام الهدف والمسببة للصفير) بتطبيق معيار الاحتمال الأعظمي (ML)

استجابتها النبضية في المجال الزمني عند التردد (f) المركزي بالعلاقة الآتية:

$$g(f, t) = \begin{cases} t^{a-1} e^{-2\pi b t} \cos(2\pi f t), & t \geq 0. \\ 0, & \text{else,} \end{cases} \quad (11)$$

حيث: يمثل (a) ترتيب المرشح، و (t) الزمن، ويمثل (f) التردد المركزي للمرشح مقدراً بالهرتز. ويحدد التابع (b) عرض حزمة مرشح غاما تون. وسنعمد في بحثنا على الترددات الممتدة من (50 Hz) إلى (8KHz) حسب تردد تقطيع الكلام (16 كيلو هرتز)، وهي الترددات المهمة بالنسبة للمعينات السمعية.

ففي المعينات السمعية الثنائية لدينا ميكروفونين، أحدهما للأذن اليمين، والثاني للأذن اليسار. وبآتي سيتشكل لدينا مزيجين. بفرض لدينا مزيج إشارة دخل الأذن اليمين  $(X_R(t))$  ومزيج إشارة دخل الأذن اليسار  $(X_L(t))$  والتردد المركزي  $(f_c)$  فإن عملية تطبيق بنك الفلاتر على القناة (c) ينتج:

$$X_L(c, t) = X_L(t) * g_L(f_c, t) \quad (12)$$

$$X_R(c, t) = X_R(t) * g_R(f_c, t) \quad (13)$$

## 2- مرحلة استخراج السمات:

سنقوم في هذه المرحلة باستخراج النماذج الثنائية للتعبير عن اتجاهات المصادر الصوتية في كل مزيج على حدة والمراد فصلها. (ITD) وهو يمثل اختلاف زمن الوصول بين الإشارات السمعية الثنائية في القناة (i). و (ILD) وهو يمثل الاختلاف في مستوى الشدة بين المزيجين الصوتيين اليمين واليسار. حيث تمتاز هذه النماذج بكونها مستقلة عن تركيب الإشارة، ويمكن استخدامها عبر جميع المكونات الصوتية على عكس نماذج التجميع الأخرى المستخدمة في العديد من أنظمة

**السمعية في تخفيض الصفير:**

ولمقارنة أداء طريقة فصل المنبع الأعمى الإحصائية وطريقة التحليل الحسي السمعي الحسابي في فصل الإشارات الصوتية قمنا بإنشاء 64 مزيجًا يتألف من إشارات الكلام والإشارات المسببة للصفير. وتم اختيار مصادر الكلام الحقيقية بشكل عشوائي من قاعدة بيانات TIMIT.

وتم تقييم أدائهما باستخدام حاسب شخصي نوع (HP Pavilion Entertainment) ، بمعالج (Intel core 2 Duo) ، وسرعة (2GHz) ، وذاكرة (RAM) بسعة (2GB) باستخدام برنامج ماتلاب، وتم قياس جودة الإشارات المنفصلة من حيث نسبة الإشارة إلى الضجيج (SNR: Signal-to-noise ratio) التي يتم حسابها باستخدام المعادلة (20) وحساب التحسن في نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (Signal-to-feedback ratio: SFR) [16,18,19,24].

$$SNR = 10 \log_{10} \frac{\|s_{target}\|^2}{\|s_{target}-s_{estimate}\|^2} \quad (20)$$

وتم حساب التحسن في نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) قبل تطبيق الخوارزميتين وبعده كما في العلاقتين (21) و (22)، ومن ثم حسبنا تحسن نسبة SFR كما في العلاقة (23).

$$SFR\text{-before} = 10 \log_{10} \frac{E[S]^2}{E[F]^2} \quad (21)$$

$$SFR\text{-after} = 10 \log_{10} \frac{E[S]^2}{E[F-\bar{F}]^2} \quad (22)$$

$$SFR\text{ improv} = SFR(A) - SFR(B) \quad (23)$$

**5-النتائج ومناقشتها:**

تم في هذا البحث تقييم أداء الطريقتين باستخدام (64) عينة من مزيج الصفير بقيم ربح توافق نقص السمع الشخي الشديد والعميق من (75) ديسبل حتى (90) ديسبل)

بحساب القيمة العظمى لاحتمال المشاهدة الحالية لمصدر الصوت ( $\lambda$ ) التي تكون عندها السمات الثنائية للمصدر الهدف هي المسيطرة في كل وحدة من وحدات الكوكليغرام كما يأتي:

$$\hat{s}(x_i) = \arg \max [P(x|\lambda_i) P(x_i)] \quad (17)$$

حيث:

$X_i$ : شعاع السمات الثنائية للإشارة الصوتية المشاهدة عبر الميكروفون (i). ويتم تنفيذ هذه الخطوة لكل إطار زمني على طول وحدات التردد - الزمن في الكوكليغرام حتى نحصل على الوحدات التي تكون فيها النماذج الثنائية المرتبطة بالمصدر الهدف مسيطرة.

**4- مرحلة التجميع:**

في مرحلة التجميع يتم استخدام تمثيل الكثافة للنماذج الثنائية (ITD,ILD) بين الإشارة الهدف والإشارة المسببة للصفير لاستخراج الإشارة الهدف وإعادة تشكيلها وسماعها كإشارة ثنائية من أجل استخدامها في مرحلة التقييم. لذلك سنقوم أولاً بحساب قناع تمثيلات الزمن الترددية لكل مصدر صوتي تم فصله بشكل مستقل عبر (i) ميكروفون عند الزمن (t) والتردد (f) بالشكل الآتي:

$$M_i(t,f) = p(x \setminus \lambda) \quad (18)$$

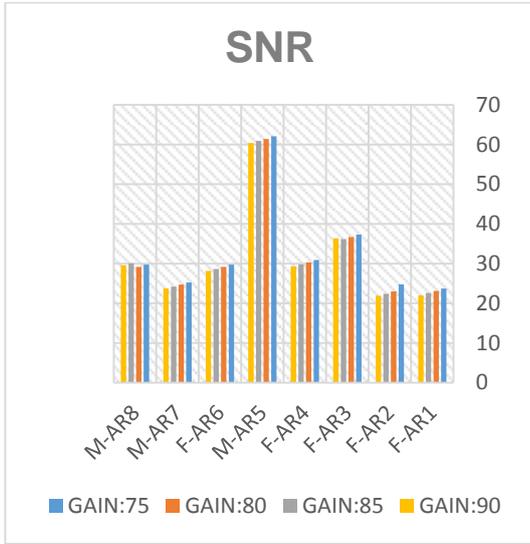
وبعدها سنحسب مقدر الصورة الخاص  $\widehat{s}_{mn}$  (spatial image estimate) ل (N) إشارة مستقبلية عند (m) ميكروفون بتطبيق  $M_i(t,f)$  على عدد المشاهدات (m) كما يأتي :

$$s_{mn}(t,f) = \widehat{M}_i(t,f) \cdot X_m(t,f) \quad (19)$$

حيث:  $n=4, i=2, m=2$ .

وفي المرحلة الأخيرة سنعيد تركيب صور إشارات الكلام الهدف المقدر في المجال الزمني للحصول على  $\widehat{s}_{mn}(t)$  باستخدام التحويل العكسي لتحويل فورييه قصير المدى (ISTFT).

**4-تقييم أداء الطريقتين الإحصائية والنماذج**



الشكل (5) نتائج اللغة العربية نسبة إشارة الكلام إلى الضجيج (SNR) بطريقة (BSS+GMM)

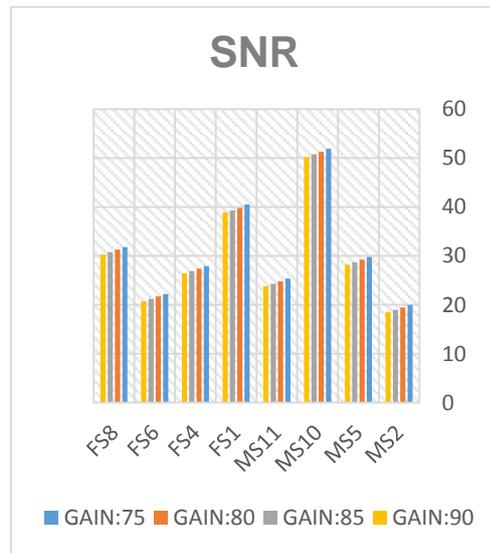
كما يبين الشكلان (6) و(7) جودة إشارة الكلام الهدف المنفصلة من حيث متوسط تحسن نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) لثماني إشارات صوتية عند قيم الربح الأربعة المختلفة بطريقة (BSS+GMM) باللغة الإنكليزية والعربية على الترتيب.



الشكل (6) متوسط تحسن SFR بطريقة (BSS+GMM) ملفات اللغة الانكليزية.

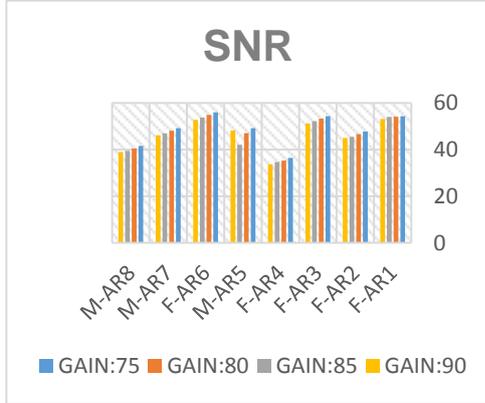
بخطوة مقدارها (5) ديسبل وتم إظهار النتائج باستخدام الواجهات الرسومية عبر برنامج ماتلاب.

يظهر الشكلان (4) و(5) جودة إشارة الكلام الهدف المنفصل من حيث نسبة الإشارة إلى الضجيج (SNR) لثمانية إشارات صوتية عند قيم الربح الأربعة المختلفة بطريقة (BSS+GMM) باللغة الإنكليزية واللغة العربية على الترتيب.



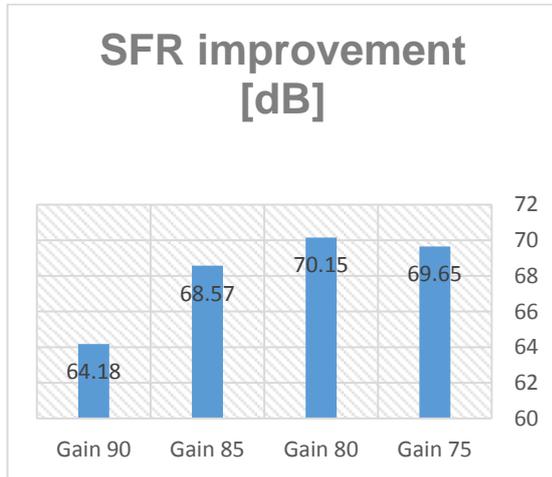
الشكل (4) نتائج اللغة الإنكليزية نسبة إشارة الكلام إلى الضجيج (SNR) بطريقة (BSS+GMM)

الشكل (8) نتائج اللغة الإنكليزية نسبة إشارة الكلام إلى الضجيج (SNR) بطريقة (CASA)



الشكل (9) نتائج اللغة الإنكليزية نسبة إشارة الكلام إلى الضجيج (SNR) بطريقة (CASA)

ويبين الشكلان (10) و(11) جودة إشارة الكلام الهدف المنفصلة من حيث متوسط تحسن نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) لثمانى إشارات صوتية عند قيم الريح الأربع المختلفة بطريقة (CASA) باللغة الانكليزية واللغة العربية على الترتيب.

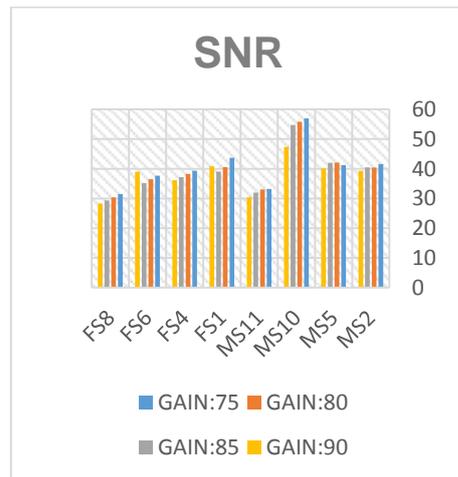


الشكل (10) متوسط تحسن SFR بطريقة (CASA) ملفات اللغة الإنكليزية.



الشكل (7) متوسط تحسن SFR بطريقة (BSS+GMM) ملفات اللغة العربية.

ويظهر الشكلان (8) و(9) جودة إشارة الكلام الهدف المنفصل من حيث نسبة الإشارة إلى الضجيج (SNR) لثمانى إشارات صوتية عند قيم الريح الأربع المختلفة بطريقة (CASA) باللغة الإنكليزية والعربية على الترتيب.



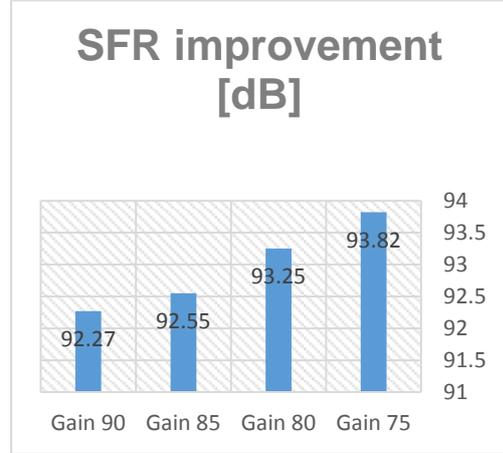
الهدف والإشارة المسببة للصفير من خلال معلومات الأذنين ، والتي كانت مستقلة عن بنية الإشارة.

نستج مما سبق أن طريقة فصل المنبع الأعمى الإحصائية (BSS+GMM) أفضل من طريقة النماذج السمعية (CASA) من حيث تحسن نسبة الإشارة الهدف للضجيج (SNR)، وكذلك تحسين نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) بالنسبة لكل من اللغة العربية والإنكليزية معاً، وهذا ما يسمح لنا باستخدامها كطريقة فعالة في تخفيض الصفير في المعينات السمعية الثنائية لتدبير حالات نقص السمع الشخي الشديد والعميق في البلاد التي تتحدث اللغة العربية أيضاً. وللتأكد من النتائج تم إجراء الاختبار الإحصائي (T) للعينات المزدوجة، ويظهر الجدول (1) وجود اختلافات ذات دلالة إحصائية بين جودة الصوت في تحسين نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) بين الطريقتين لملفات اللغة الإنكليزية حيث بلغت قيمة  $(p < 0.05)$ . بينما الجدول (2) يظهر عدم وجود اختلافات ذات دلالة إحصائية بين جودة الصوت في تحسين نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) بين الطريقتين لملفات اللغة العربية حيث بلغت قيمة  $(p > 0.05)$ .

الجدول (1) نتائج اختبار (Paired-Samples T-test) للمقارنة بين الطريقتين في تحسين (SFR) لملفات اللغة الإنكليزية

Gain [dB]	75	80	85	90	
Mean	SFR – BSS [dB]	97.79	97.81	97.82	97.83
	SFR – CASA[d B]	69.65	70.15	68.57	64.18
Variance	SFR – BSS [dB]	10.66	10.61	10.52	10.49
	SFR – CASA[d B]	844.29	1131.68	10208.64	1389.94
P-value	0.01	0.02	0.02	0.02	

p:probability



الشكل (11) متوسط تحسن SFR بطريقة (CASA) لملفات اللغة العربية.

أظهرت طريقة فصل المنبع الإحصائية كما نرى من الشكلين (4) و(5) أداءً متشابهاً في تحسين نسبة الإشارة الهدف للضجيج (SNR) بالنسبة لجميع ملفات الكلام الصوتية باللغة العربية والإنكليزية وجميع قيم الريح المختلفة.

في حين اختلف أداء طريقة النماذج السمعية (CASA) في نسبة الإشارة الهدف للضجيج (SNR) بالنسبة لملفات الكلام باللغة الإنكليزية كما نرى في الشكل (8) وكذلك بالنسبة لملفات اللغة العربية في الشكل (9) حيث اختلفت نسبة (SNR) بشكل واضح باختلاف قيم الريح المختلفة.

وحققت أيضاً طريقة فصل المنبع الإحصائية مستوى أعلى من SFR بالنسبة لجميع ملفات الكلام الصوتية باللغة الإنكليزية والعربية مع زيادة الريح .

في حين كان أداء طريقة النماذج السمعية (CASA) في تحسين نسبة (SFR) أفضل بالنسبة لملفات الكلام باللغة الإنكليزية عند قيم الريح (75 و80 ديسبل )، في حين انخفض أداؤها تدريجياً مع زيادة الريح عند قيم الريح (85, 90) ديسبل كما نرى في الشكل (10)، وكذلك بالنسبة لملفات اللغة العربية كما في الشكل (11) حيث اختلفت نسبة (SFR) بشكل واضح مع زيادة الريح، وكان هذا الاختلاف أكبر بالنسبة لملفات اللغة العربية. إلا أنها كانت قادرة على التمييز بين إشارة الكلام

## المراجع References

- [1] Guo, M. (2020, May). An Empirical Study on Acoustic Feedback Path Across Hearing Aid Users. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 441-445). IEEE.
- [2] Lee HW, Jeon MY. A combined feedback and noise cancellation algorithm for binaural hearing aids. *Advances in Electrical and Computer Engineering*. 2011;11(3):35-40.
- [3] Jorge I. M. , (August , 2012 )," Robust Binaural Noise -Reduction Strategies With Binaural-Hearing-Aid Constraints: Design, Analysis And Practical Consideration" , the School of Electrical and Computer Engineering ,Ph.D. Thesis, Georgia Institute of Technology,PP.148 .
- [4] Jafari I, Haque S, Togneri R, Nordholm S. Evaluations on underdetermined blind source separation in adverse environments using time-frequency masking. *EURASIP J Adv Signal Process*. 2013;2013(1):162.
- [5] Hamacher V., Fischer E., Kornagel U. and Puder H. (2011). "Applications of Adaptive Signal Processing Methods in High-End Hearing Aids", *Topics in Acoustic Echo and Noise Control*, Springer, Berlin, Heidelberg, New York.
- [6] van Waterschoot T., Moonen M., (February 2011),"Fifty Years of Acoustic Feedback Control: State of the Art and Future Challenges", *Proceedings of the IEEE* , Vol.99, No. 2.
- [7] Avan P, Giraudet F, Büki B. Importance of binaural hearing. *Audiol Neurootol*. 2015;20 Suppl 1:3-6.
- [8] Dillon H. *Hearing aids*. 2nd ed. London: Thieme Medical Pub; 2001
- [9] Lombard A., Reindl K., and Kellermann W. , (2009) , "Combination of Adaptive Feedback Cancellation and Binaural Adaptive Filtering in Hearing Aids", *EURASIP Journal on Advances in Signal Processing*.
- [10] Tarrab F., Marmar Z. and Alamine I. ,(2016), "Howling Reduction in Binaural Hearing Aids Using Blind Source Separation Technique" , *Damascus University journal for the engineering science*, Damascus University, Damascus.
- [11] Hu G., and Wang D., (2011). "An Auditory Scene Analysis Approach to

الجدول (2) نتائج اختبار (Paired-Samples T-) للمقارنة بين الطريقتين في تحسين (SFR) ملفات اللغة العربية

Gain [dB]	75	80	85	90	
Mean	SFR – BSS [dB]	95.7	95.7	95.7	95.7
	SFR – CASA [dB]	93.82	93.25	92.55	92.27
Variance	SFR – BSS [dB]	13.04	13.04	13.36	12.87
	SFR – CASA [dB]	41.75	49.19	56.01	60.12
P-value	0.28	0.23	0.19	0.18	

p:probability

## 6-الخاتمة:

تم في هذا البحث تصميم خوارزميتين لمقارنة الطرق الإحصائية ممثلة بطريقة فصل المنبع الأعمى مع نموذج غوص الإحصائي (BSS+GMM)، وكذلك النماذج السمعية ممثلة بطريقة التحليل الحسي السمعي الحسابي (CASA) لاختبار أيهما أفضل في التعامل مع حالات نقص السمع الشخي الشديد والعميق للاستخدام في المعينات السمعية الثنائية، وتبين لنا من نتائج الاختبار فعالية أداء طريقة فصل المنبع الإحصائية في تحسين نسبة الإشارة الهدف للضجيج (SNR)، وكذلك تحسين نسبة الإشارة الهدف إلى الإشارة المسببة للصفير (SFR) بالنسبة لكل من اللغة العربية والإنكليزية معاً، كما وتسمح هذه الطريقة للمستخدم بزيادة ربح المعينة السمعية دون أن تتأثر جودة الصوت المسموع.

- [21] Shao Y, Srinivasan S, Jin Z, Wang D. A computational auditory scene analysis system for speech segregation and robust speech recognition. *Comput Speech Lang.* 2010;24(1):77-93.
- [22] Sprieta A., Rombouts G., Moonena M., Wouters J., (2006), "Adaptive feedback cancellation in hearing aids", The Franklin Institute. Published by Elsevier Ltd.
- [23] Matsuo K. , Kawano S., Nakashima† H., Chisaki Y., (2005), "A simple howling control for binaural hearing aid system", 12th International Congress on Sound and Vibration.
- [24] Meng G., Thomas Bo E., Søren H. J. and Jesper J. , (2013) , "Evaluation of State-of-the-Art Acoustic Feedback Cancellation Systems for Hearing Aids " , J. Audio Eng. Soc.
- [25] Mukai R., Sawada H., Araki S., and Makino S. (2004), " Frequency Domain Blind Source Separation for Many Speech Signals " , Springer-Verlag Berlin Heidelberg .
- [26] Hirasawa Y, Yasuraoka N, Takahashi T, Ogata T, and Okuno G. H, (2012), "A GMM Sound Source Model for Blind Speech Separation in Under-determined Conditions", Springer-Verlag Berlin Heidelberg.
- [27] Jeff A. Bilmes , (April 1998), "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models". International Computer Science Institute Berkeley CA, 94704 and Computer Science Division Department of Electrical Engineering and Computer Science U.C. Berkeley.
- [28] Brown G.J., Wang D.L.,(2005)," Separation of speech by computational auditory scene analysis", *Speech Enhancement*, J. Benesty, S. Makino and J. Chen (Eds.), Springer, New York,
- [29] Walker, Kevin, et al.( 2014), "GALE Phase 2 Arabic Broadcast News Speech Part 1 ",LDC2014S07. DVD. Philadelphia: Linguistic Data Consortium.
- [30] Walker, Kevin, et al.(2016), "GALE Phase 3 Arabic Broadcast Conversation Speech Part 2 " ,LDC2016S01. Web Download. Philadelphia: Linguistic Data Consortium.
- [31] Mohamed M., Graff D., and Cieri Ch.,(2006)," Arabic Broadcast News Speech LDC2006S46". DVD. Philadelphia: Linguistic Data Consortium.
- Monaural Speech Segregation", *Topics in Acoustic Echo and Noise Control*, Springer, Berlin, Heidelberg, New York.
- [12] Ji X., Ruiyu L., Qingwu L., Z. Xuewu,(April,2013)," A Novel Acoustic Feedback Suppression Method based on Linear Prediction and Parallel Gradient Projection", *Advances in information Sciences and Service Sciences(AISS)* Vol.5, Number7,pp.60-66.
- [13] Cichocki A, Amari S.I ,(2003), "adaptive Blind signal and image Processing, learning lgorithms and applications " , JOHN WILEY&SONS.LTD.
- [14] Rabiee A., Setayeshi S., and Lee S.,(2012), "CASA: Biologically Inspired Approaches for Auditory Scene Analysis " , *Natural Intelligence: the INNS Magazine* 50 Vol. 1, No.2.
- [15] Emura, S., Sawada, H., Araki, S., & Harada, N. (2020, May). A Frequency-Domain BSS Method Based on  $\ell_1$  Norm, Unitary Constraint, and Cayley Transform. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 111-115). IEEE.
- [16] Xiaomei, C., Guodong, L., & Bo, Z. (2019, November). Objective assessment of hearing aid performance based on Gaussian process. In *2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)* (pp. 1606-1613). IEEE.
- [17] Ganesh R. N., Wang W.,( 2014), "Blind Source Separation Advances in Theory, Algorithms and Applications " , Springer-Verlag Berlin Heidelberg .
- [18] Guo M, Jensen SH, Jesper J. Evaluation of state-of-the-art acoustic feedback cancellation systems for hearing aids. *J Audio Eng Soc.* 2013;61(3):125-37.
- [19] Samuel M., (Mars , 2006) , " Subjectively preferred sound representation for different listening situations " , Department of Speech, Music and Hearing Royal Institute of Technology.
- [20] Lynne A. Werner , Richard R. Fay and Arthur N. Popper,(2012),"Human Auditory Development", Springer Science+Business Media.

- [32] Alsulaiman M., et al. (2014), "King Saud University Arabic Speech Database LDC2014S02". Hard Drive. Philadelphia: Linguistic Data Consortium.
- [33] Walker, Kevin, et al. (2015), "GALE Phase 2 Arabic Broadcast News Speech Part 2 LDC2015S01". DVD. Philadelphia: Linguistic Data Consortium.
- [34] Ma G.,(2006), "Source Separation for Feedback Cancellation in Hearing Aids", Master thesis, at Oersted, DTU.
- [35] Gelfand, S. A. (2017). Essentials of audiology.

Received	2020/6/22	إيداع البحث
Accepted for Publ.	2020/9/2	قبول البحث للنشر