التوصيـة ITU-R BS.1387-2

(2023/05)

السلسلة BS: الخدمة الإذاعية (الصوتية)

طريقة القياسات الموضوعية لجودة الصوت المدركة



**السلسلة SA**

**التطبيقات الفضائية والأرصاد الجوية**

**تمهيـد**

يضطلع قطاع الاتصالات الراديوية بدور يتمثل في تأمين الترشيد والإنصاف والفعالية والاقتصاد في استعمال طيف الترددات الراديوية في جميع خدمات الاتصالات الراديوية، بما فيها الخدمات الساتلية، وإجراء دراسات دون تحديد لمدى الترددات، تكون أساساً لإعداد التوصيات واعتمادها.

ويؤدي قطاع الاتصالات الراديوية وظائفه التنظيمية والسياساتية من خلال المؤتمرات العالمية والإقليمية للاتصالات الراديوية وجمعيات الاتصالات الراديوية بمساعدة لجان الدراسات.

سياسة قطاع الاتصالات الراديوية بشأن حقوق الملكية الفكرية (IPR)

يرد وصف للسياسة التي يتبعها قطاع الاتصالات الراديوية فيما يتعلق بحقوق الملكية الفكرية في سياسة البراءات المشتركة بين قطاع تقييس الاتصالات وقطاع الاتصالات الراديوية والمنظمة الدولية للتوحيد القياسي واللجنة الكهرتقنية الدولية (ITU‑T/ITU‑R/ISO/IEC) والمشار إليها في القرار ITU-R 1.   
وترد الاستمارات التي ينبغي لحاملي البراءات استعمالها لتقديم بيان عن البراءات أو للتصريح عن منح رخص في الموقع الإلكتروني <http://www.itu.int/ITU-R/go/patents/en> حيث يمكن أيضاً الاطلاع على المبادئ التوجيهية الخاصة بتطبيق سياسة البراءات المشتركة وعلى قاعدة بيانات قطاع الاتصالات الراديوية التي تتضمن معلومات عن البراءات.

|  |  |
| --- | --- |
| **سلاسل توصيات قطاع الاتصالات الراديوية**  (يمكن الاطلاع عليها أيضاً في الموقع الإلكتروني <https://www.itu.int/publ/R-REC/en>) | |
| **السلسلة** | **العنـوان** |
| **BO** البث الساتلي | |
| **BR** التسجيل من أجل الإنتاج والأرشفة والعرض؛ الأفلام التلفزيونية | |
| **BS الخدمة الإذاعية (الصوتية)** | |
| **BT** الخدمة الإذاعية (التلفزيونية) | |
| **F** الخدمة الثابتة | |
| **M** الخدمة المتنقلة وخدمة التحديد الراديوي للموقع وخدمة الهواة والخدمات الساتلية ذات الصلة | |
| **P** انتشار الموجات الراديوية | |
| **RA** علم الفلك الراديوي | |
| **RS** أنظمة الاستشعار عن بُعد | |
| **S** الخدمة الثابتة الساتلية | |
| **SA** التطبيقات الفضائية والأرصاد الجوية | |
| **SF** تقاسم الترددات والتنسيق بين أنظمة الخدمة الثابتة الساتلية والخدمة الثابتة | |
| **SM** إدارة الطيف | |
| **SNG** التجميع الساتلي للأخبار | |
| **TF** إرسالات الترددات المعيارية وإشارات التوقيت | |
| **V** المفردات والمواضيع ذات الصلة | |

|  |
| --- |
| ***ملاحظة****: تمت الموافقة على النسخة الإنكليزية لهذه التوصية الصادرة عن قطاع الاتصالات الراديوية بموجب الإجراء الموضح في القرار ITU-R 1.* |

*النشر الإلكتروني*جنيف، 2024

© ITU 2024

جميع حقوق النشر محفوظة. لا يمكن استنساخ أي جزء من هذه المنشورة بأي شكل كان ولا بأي وسيلة إلا بإذن خطي من الاتحاد الدولي للاتصالات (ITU).

التوصيـة ITU-R BS.1387-2

طريقة القياسات الموضوعية لجودة الصوت المدركة

(2023-2001-1998)

مجال التطبيق

توصف هذه التوصية طريقة للقياس الموضوعي لجودة الصوت المدركة.

مصطلحات أساسية

جودة الصوت المدركة، القياس الموضوعي، تشفير بمعدل بتات منخفض

إن جمعية الاتصالات الراديوية للاتحاد الدولي للاتصالات،

إذ تضع في اعتبارها

*أ )* أن الطرائق الموضوعية التقليدية (مثل قياس النسبة إشارة إلى ضوضاء والتشوه) لم تعد كافية لقياس جودة الصوت المدركة للأنظمة التي تستخدم مخططات تشفير بمعدل بتات منخفض أو التي تستخدم معالجة الإشارات التماثلية أو الرقمية؛

*ب)* أن مخططات التشفير بمعدل بتات منخفض آخذة في الانتشار بسرعة؛

*ج)* أن عمليات التنفيذ التي تمتثل لمواصفة ما أو لمعيار ما لا تضمن جميعها أفضل جودة يمكن تحقيقها في إطار تلك المواصفة أو ذلك المعيار؛

*د )* أن طرائق التقييم الشخصي الرسمية لا تصلح للمراقبة المستمرة لجودة الصوت، في ظل الظروف التشغيلية مثلاً؛

*ھ )* أن طرائق القياس الموضوعي لجودة الصوت المدركة قد تكمِّل في نهاية الأمر أو تحل محل طرائق الاختبار الموضوعي التقليدية في جميع مجالات القياسات؛

و ) أن القياس الموضوعي لجودة الصوت المدركة يمكن أن يكمِّل بصورة مفيدة طرائق التقييم الشخصي؛

*ز )* أنه بالنسبة لبعض التطبيقات، من الضروري وجود طريقة يمكن تنفيذها في الوقت الفعلي،

توصي

1 بأن تستخدم الطريقة الواردة في الملحق 2 للقياس الموضوعي لجودة الصوت المدركة، لكل تطبيق مدرج في الملحق 1.

تمهيد

توصف هذه التوصية طريقة للقياس الموضوعي لجودة الصوت المدركة لجهاز قيد الاختبار، على سبيل المثال، كوديك ذو معدل بت منخفض. وهي مقسمة إلى ملحقين. يقدم الملحق 1 للمستعمل نظرة عامة على الطريقة ويتضمن أربعة مرفقات. حيث يصف المرفق 1 التطبيقات وإشارات الاختبار. ويسرد المرفق 2 متغيرات مخرجات النموذج ويناقش قيود الاستخدام والدقة. ويقدم المرفق 3 الخطوط العريضة للنموذج بينما يصف المرفق 4 مبادئ وخصائص طرائق قياس جودة الصوت الإدراكية الموضوعية بشكل عام.

يوفر الملحق 2 للمنفذ وصفاً تفصيلياً للطريقة باستخدام نسختين من النموذج النفسي الصوتي الذي تم تطويره خلال مرحلة التكامل حيث تم دمج ستة نماذج. ويرد في المرفق 1 بالملحق 2 وصف لعملية التحقق من صحة طريقة القياس الموضوعي. ويقدم المرفق 2 بالملحق 2 نظرة عامة على جميع قواعد البيانات التي تم استخدامها في تطوير الطريقة والتحقق من صحتها.

جدول المحتويات

*الصفحة*

[سياسة قطاع الاتصالات الراديوية بشأن حقوق الملكية الفكرية (IPR) ii](#_Toc165969814)

[الملحق 1 - نظرة عامة 7](#_Toc165969815)

[1 مقدمة 7](#_Toc165969816)

[2 التطبيقات 7](#_Toc165969817)

[3 الصيغ 8](#_Toc165969818)

[4 الميدان الشخصي 8](#_Toc165969819)

[5 الاستبانة والدقة 10](#_Toc165969820)

[6 المتطلبات والقيود 10](#_Toc165969821)

[المرفق 1 بالملحق 1 - التطبيقات 10](#_Toc165969822)

[1 عام 10](#_Toc165969823)

[2 التطبيقات الرئيسية 11](#_Toc165969824)

[1.2 تقييم عمليات التنفيذ 11](#_Toc165969825)

[2.2 ترتيب الجودة الإدراكية 11](#_Toc165969826)

[3.2 المراقبة عبر الإنترنت 11](#_Toc165969827)

[4.2 حالة المعدات أو التوصيلة 11](#_Toc165969828)

[5.2 تحديد هوية الكوديك 12](#_Toc165969829)

[6.2 تطوير الكوديكات 12](#_Toc165969830)

[7.2 تخطيط الشبكة 12](#_Toc165969831)

[8.2 المساعدة في التقييم الشخصي 12](#_Toc165969832)

[9.2 ملخص التطبيقات 13](#_Toc165969833)

[3 إشارات الاختبار 13](#_Toc165969834)

[1.3 اختيار إشارات الاختبار الطبيعية 13](#_Toc165969835)

[2.3 مدة الإشارة 15](#_Toc165969836)

[4 التزامن 15](#_Toc165969837)

[5 مسائل حقوق التأليف والنشر 15](#_Toc165969838)

[المرفق 2 بالملحق 1 - متغيرات الخرج 15](#_Toc165969839)

[1 مقدمة 15](#_Toc165969840)

[2 متغيرات خرج النموذج 16](#_Toc165969841)

[3 جودة الصوت الأساسية 16](#_Toc165969842)

[4 هامش التشفير 17](#_Toc165969843)

[5 متطلبات المستعمل 17](#_Toc165969844)

[المرفق 3 بالملحق 1 - الخطوط العريضة للنموذج 17](#_Toc165969845)

[1 معالجة الصوت 18](#_Toc165969846)

[1.1 الإعدادات المحددة من قبل المستعمل 18](#_Toc165969847)

[2.1 النموذج الصوتي النفسي 18](#_Toc165969848)

[3.1 النموذج الإدراكي 19](#_Toc165969849)

[المرفق 4 بالملحق 1 - مبادئ وخصائص طرائق قياس جودة الصوت الإدراكية الموضوعية 19](#_Toc165969850)

[1 مقدمة والتسلسل التاريخي 19](#_Toc165969851)

[2 البنية العامة لطرائق قياس جودة الصوت الإدراكية الموضوعية 20](#_Toc165969852)

[3 الأسس النفسية والصوتية والإدراكية 21](#_Toc165969853)

[1.3 خاصية النقل للأذن الخارجية والوسطى 21](#_Toc165969854)

[2.3 سلالم التردد الإدراكية 21](#_Toc165969855)

[3.3 الإثارة 22](#_Toc165969856)

[4.3 الكشف 23](#_Toc165969857)

[5.3 التقنيع 23](#_Toc165969858)

[6.3 الجهارة والتقنيع الجزئي 24](#_Toc165969859)

[7.3 الحدة 24](#_Toc165969860)

[8.3 المعالجة الإدراكية 24](#_Toc165969861)

[4 النماذج المدمجة 25](#_Toc165969862)

[1.4 مؤشر التشويش (DIX) 25](#_Toc165969863)

[2.4 نسبة الضوضاء إلى الإشارة المقنعة (NMR) 26](#_Toc165969864)

[3.4 التقييم الموضوعي للإشارة الصوتية 26](#_Toc165969865)

[4.4 قياس جودة الصوت الإدراكية (PAQM) 27](#_Toc165969866)

[5.4 التقييم الإدراكي 27](#_Toc165969867)

[6.4 القياس الموضوعي الإدراكي 28](#_Toc165969868)

[7.4 نهج مجموعة الأدوات 28](#_Toc165969869)

[الملحق 2 - وصف النموذج 30](#_Toc165969870)

[1 الخطوط العريضة 30](#_Toc165969871)

[1.1 الصيغة الأساسية 31](#_Toc165969872)

[2.1 الصيغة المتقدمة 31](#_Toc165969873)

[2 نموذج الأذن المحيطية 31](#_Toc165969874)

[1.2 نموذج الأذن القائم على التحويل FFT 31](#_Toc165969875)

[2.2 نموذج الأذن القائم على بنك المرشحات 45](#_Toc165969876)

[3 المعالجة المسبقة لأنماط الإثارة 53](#_Toc165969877)

[1.3 تكييف المستوى والنمط 53](#_Toc165969878)

[2.3 التشكيل 55](#_Toc165969879)

[3.3 الجهارة 55](#_Toc165969880)

[4.3 حساب إشارة الخطأ 56](#_Toc165969881)

[4 حساب متغيرات خرج النموذج 56](#_Toc165969882)

[1.4 نظرة عامة 56](#_Toc165969883)

[2.4 فرق التشكيل 57](#_Toc165969884)

[3.4 جهارة الضوضاء 58](#_Toc165969885)

[4.4 عرض النطاق 60](#_Toc165969886)

[5.4 نسبة الضوضاء إلى القناع 61](#_Toc165969887)

[6.4 Relative Disturbed FramesB  61](#_Toc165969888)

[7.4 احتمال الكشف 61](#_Toc165969889)

[8.4 البنية التوافقية للخطأ 63](#_Toc165969890)

[5 التوسيط 64](#_Toc165969891)

[1.5 التوسيط الطيفي 64](#_Toc165969892)

[2.5 التوسيط الزمني 64](#_Toc165969893)

[3.5 التوسيط عبر القنوات الصوتية 66](#_Toc165969894)

[6 تقدير جودة الصوت الأساسية المدركة 66](#_Toc165969895)

[1.6 الشبكة العصبية الاصطناعية 66](#_Toc165969896)

[2.6 الصيغة الأساسية 67](#_Toc165969897)

[3.6 الصيغة المتقدمة 69](#_Toc165969898)

[7 مطابقة عمليات التنفيذ 70](#_Toc165969899)

[1.7 عام 70](#_Toc165969900)

[2.7 الاختيار 70](#_Toc165969901)

[3.7 إعدادات اختبار المطابقة 70](#_Toc165969902)

[4.7 فاصل التسامح المقبول 71](#_Toc165969903)

[5.7 عناصر الاختبار 71](#_Toc165969904)

[المرفق 1 بالملحق 2 - عملية التحقق 72](#_Toc165969905)

[1 عام 72](#_Toc165969906)

[2 المرحلة التنافسية 73](#_Toc165969907)

[3 المرحلة التعاونية 74](#_Toc165969908)

[4 التحقق 74](#_Toc165969909)

[1.4 المقارنة بين الدرجات SDG وODG 75](#_Toc165969910)

[2.4 الترابط 75](#_Toc165969911)

[3.4 درجة الخطأ المطلق (AES) 77](#_Toc165969912)

[4.4 مقارنة الدرجات ODG مقابل فاصل الثقة 79](#_Toc165969913)

[5.4 مقارنة الدرجات ODG مقابل فاصل التسامح 83](#_Toc165969914)

[5 اختيار صيغ النموذج المثلى 84](#_Toc165969915)

[1.5 معايير الاختيار المسبق استناداً إلى الترابط 84](#_Toc165969916)

[2.5 تحليل عدد القيم الشاذة 85](#_Toc165969917)

[3.5 تحليل شدة القيم الشاذة 86](#_Toc165969918)

[6 الخلاصة 87](#_Toc165969919)

[المرفق 2 بالملحق 2 - أوصاف قواعد البيانات المرجعية 87](#_Toc165969920)

[1 مقدمة 87](#_Toc165969921)

[2 العناصر لكل قاعدة بيانات 89](#_Toc165969922)

[3 الحالات التجريبية 89](#_Toc165969923)

[1.3 MPEG90 90](#_Toc165969924)

[2.3 MPEG91 90](#_Toc165969925)

[3.3 ITU92DI 90](#_Toc165969926)

[4.3 ITU92CO 90](#_Toc165969927)

[5.3 ITU93 90](#_Toc165969928)

[6.3 MPEG95 91](#_Toc165969929)

[7.3 EIA95 91](#_Toc165969930)

[8.3 DB2 91](#_Toc165969931)

[9.3 DB3 91](#_Toc165969932)

[10.3 CRC97 92](#_Toc165969933)

[4 العناصر لكل حالة بالنسبة لقاعدتي البيانات DB2 وDB3 92](#_Toc165969934)

[1.4 DB2 92](#_Toc165969935)

[2.4 قاعدة البيانات DB3 95](#_Toc165969936)

[المراجع 98](#_Toc165969937)

[بيبليوغرافيا 99](#_Toc165969938)

الملحق 1  
  
نظرة عامة

# 1 مقدمة

تعد جودة الصوت أحد العوامل الأساسية عند تصميم نظام رقمي للإذاعة. وقد أدى الإدخال السريع لمختلف مخططات خفض معدل البتات إلى بذل جهود كبيرة في إنشاء وتحسين إجراءات التقييمات الشخصية، وذلك ببساطة لأن اختبارات الاستماع الرسمية كانت الطريقة الوحيدة ذات الصلة للحكم على جودة الصوت. وكانت الخبرة المكتسبة هي الأساس الذي قامت عليه التوصية ITU‑R BS.1116، التي أصبحت فيما بعد الأساس لمعظم اختبارات الاستماع من هذا النوع.

وبما أن التقييمات الشخصية للجودة تستغرق وقتاً طويلاً وباهظة الثمن، فمن المستحسن تطوير طريقة قياس موضوعي من أجل التوصل إلى تقدير لجودة الصوت. ولم يثبُت أبداً أن طرق القياس الموضوعي التقليدية، مثل النسبة إشارة إلى ضوضاء (S/N) أو التشوه الكلي الناجم عن التوافقيات (THD) ترتبط بشكل موثوق بجودة الصوت المدركة. وقد أصبحت المشكلات أكثر وضوحاً عند تطبيق الطرائق على الكوديكات الحديثة غير الخطية وغير الثابتة، على السواء.

لقد تم تطوير عدد من الطرائق لإجراء قياسات إدراكية موضوعية لجودة الصوت المدركة خلال العقد الماضي. ولكن لم يتم التحقق من صحة أي من هذه الطرائق بشكل كامل، وبالتالي لم يتم تقييسها أو قبولها على نطاق واسع. وفي عام 1994، حدد قطاع الاتصالات الراديوية الحاجة الملحة إلى وضع معيار في هذا المجال وبدأ العمل في ذلك. وتم إصدار دعوة مفتوحة لتقديم المقترحات وتم تلقي الطرائق الست التالية كطرائق مرشحة للقياس: مؤشر الاضطراب (DIX)، ونسبة الضوضاء إلى القناع (NMR)، وقياس جودة الصوت الإدراكي (PAQM)، والتقييم الإدراكي (PERCEVAL)، والقياس الموضوعي الإدراكي (POM) ونهج الأدوات. ويرد وصف هذه الطرائق في المرفق 4 بالملحق 1.

وطريقة القياس الواردة في هذه التوصية هي نتيجة عملية تمت فيها دراسة أداء كل من الطرائق الست المذكورة أعلاه، وتم استخراج أفضل الأدوات الواعدة ودمجها في طريقة واحدة. ولقد تم التحقق من صحة الطريقة الموصى بها بعناية في عدد من مواقع الاختبار. وقد أثبتت أنها تنتج معلومات موثوقة ومفيدة للعديد من التطبيقات. ومع ذلك، لا بد من الأخذ في الاعتبار أن طريقة القياس الموضوعي الواردة في هذه التوصية ليست عموماً بديلاً عن إجراء اختبار استماع رسمي.

# 2 التطبيقات

يُوضح في الشكل 1 أدناه المفهوم الأساسي لإجراء قياسات موضوعية باستخدام الطريقة الموصى بها.

الشكل 1

المفهوم الأساسي لإجراء قياسات موضوعية



تنطبق طريقة القياس الواردة في هذه التوصية على معظم أنواع تجهيزات معالجة الإشارات الصوتية، الرقمية منها والتماثلية. ومع ذلك، من المتوقع أن تركز العديد من التطبيقات على الكوديكات الصوتية.

وقد تم تحديد الفئات الثمانية التالية من التطبيقات:

الجدول 1

التطبيقات

|  |  |  |  |
| --- | --- | --- | --- |
|  | التطبيق | وصف مختصر | الصيغة |
| 1 | تقييم عمليات التنفيذ | إجراء لتحديد خصائص عمليات التنفيذ المختلفة لمعدات معالجة الصوت، وفي كثير من الحالات الكوديكات الصوتية | أساسية/متقدمة |
| 2 | ترتيب الجودة الإدراكية | إجراء سريع يتم إجراؤه قبل إدخال أي جزء من المعدات أو دارة في الخدمة | أساسية |
| 3 | المراقبة عبر الإنترنت | عملية مستمرة لمراقبة الإرسال الصوتي في الخدمة | أساسية |
| 4 | حالة المعدات أو التوصيل | تحليل مفصل لجزء من المعدات أو دارة | متقدمة |
| 5 | تحديد الكوديك | إجراء لتحديد نوع كوديك معين وعملية تنفيذه | متقدمة |
| 6 | تطوير الكوديك | إجراء يحدد خصائص أداء الكوديك بأكبر قدر ممكن من التفصيل | أساسية/متقدمة |
| 7 | تخطيط الشبكة | إجراء لاستمثال تكلفة وأداء شبكة النقل في ظل قيود معينة | أساسية/متقدمة |
| 8 | المساعدة في التقييم الذاتي | أداة لفرز المواد الهامة لإدراجها في اختبار الاستماع | أساسية/متقدمة |

# 3 الصيغ

من أجل تحقيق التوافق الأمثل مع مختلف متطلبات التكلفة والأداء، فإن طريقة القياس الموضوعي الموصى بها في هذه التوصية لها صيغتان. وقد تم تصميم الصيغة الأساسية للسماح بتنفيذ فعال من حيث التكلفة في الوقت الفعلي، في حين تركز الصيغة المتقدمة على تحقيق أعلى دقة ممكنة. واعتماداً على التنفيذ، تزيد هذه الدقة الإضافية من التعقيد بمقدار أربعة أضعاف تقريباً مقارنة بالصيغة الأساسية.

ويقدم الجدول 1 بعض الإرشادات حول الصيغة التي يجب الذي تطبيقها لكل تطبيق من التطبيقات.

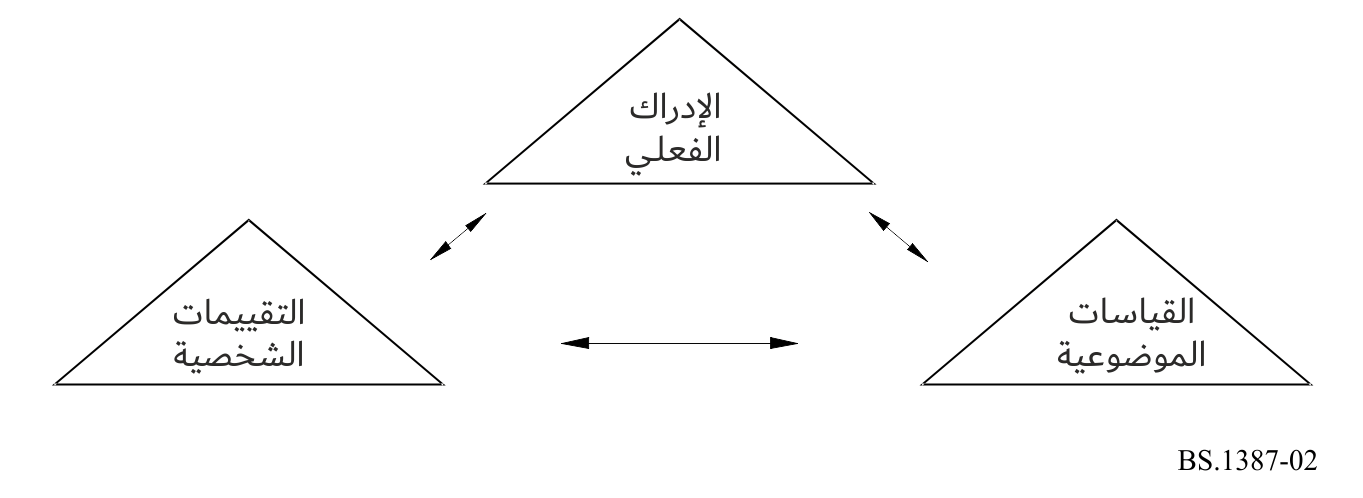
# 4 الميدان الشخصي

اختبارات الاستماع الشخصية الرسمية، على سبيل المثال، تلك المستندة إلى التوصية ITU-R BS.1116، مصممة بعناية لتقترب قدر الإمكان من تقدير موثوق للحكم على جودة الصوت. ومع ذلك، لا يمكن للمرء أن يتوقع أن تعكس نتيجة اختبار الاستماع الشخصي الإدراك الفعلي بشكل كامل. ويوضح الشكل 2 العيوب الضمنية في كل من الميدان الشخصي والموضوعي.

ومن الواضح أنه لن يتسنى التحقق من صحة طريقة موضوعية مباشرة. وبدلاً من ذلك، يتم التحقق من صحة طرائق القياس الموضوعي مقابل اختبارات الاستماع الشخصية.

الشكل 2

مفاهيم التحقق من الصلاحية



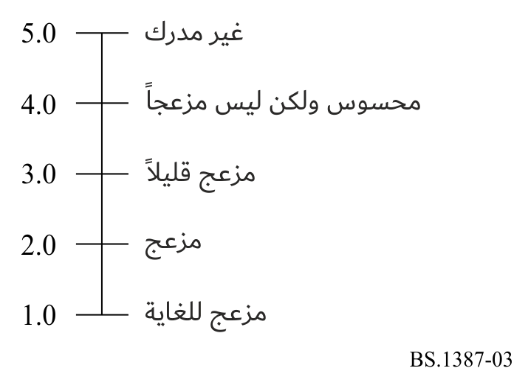
تركز طريقة القياس الموضوعي الواردة في هذه التوصية على التطبيقات التي يتم تقييمها عادة في الميدان الشخصي من خلال تطبيق التوصية ITU R BS.1116. ويمكن وصف المبدأ الأساسي لطريقة الاختبار المحددة هذه بإيجاز على النحو التالي: يمكن للمستمع الاختيار بين ثلاثة مصادر ("A" و"B" و"C"). والإشارة المرجعية المعروفة متاحة دائماً كالمصدر "A". وتتوفر الإشارة المرجعية المخفية والإشارة قيد الاختبار في وقت واحد ولكن يتم تخصيصهما "عشوائيًا" للمصدرين "B" و"C"، اعتماداً على التجربة.

ويُطلب من المستمع تقييم مظاهر الانحطاط في المصدر "B" مقارنة بالانحطاط في المصدر "A" وفي المصدر "C" مقارنة بـالانحطاط في "A" وفقاً لسلم الانحطاط المستمر خماسي الدرجات. وينبغي أن يتعذر تمييز أحد المصدرين "B" أو "C" عن المصدر "A". ويمكن أن يكشف المصدر الآخر عن مظاهر الانحطاط. ويجب تفسير أي اختلافات مُدرَكة بين المصدر المرجعي والمصدر الآخر باعتبارها انحطاطاً. وتُستخدم عادةً سمة واحدة فقط، وهي "جودة الصوت الأساسية". وتُعرف على أنها سمة شاملة تتضمن أي وجميع الاختلافات المكتشفة بين الإشارة المرجعية والإشارة قيد الاختبار.

ويتم التعامل مع سلم الدرجات باعتباره مستمراً مع وجود "نقاط ارتكاز" مستمدة من سلم الانحطاط خماسي الدرجات لقطاع الاتصالات الراديوية الوارد في التوصية ITU‑R BS.1284 والمبين أدناه.

الشكل 3

سلم الانحطاط خماسي الدرجات لقطاع الاتصالات الراديوية



يعتمد تحليل نتائج اختبار الاستماع الشخصي بشكل عام على درجة الاختلاف الشخصية (SDG) المحددة على النحو التالي:

*SDG*  *GradeSignal Under Test* – *GradeReference Signal*

وينبغي أن تتراوح قيم الدرجة SDG بشكل مثالي بين 0 و4–، حيث تقابل القيمة 0 انحطاطاً غير محسوس و4– انحطاطاً يُعتبر مزعجاً للغاية.

# 5 الاستبانة والدقة

درجة الاختلاف الموضوعية (ODG) هي متغير خرج من طريقة القياس الموضوعي وتقابل الدرجة SDG في الميدان الشخصي. وتقيد استبانة الدرجة ODG برقم عشري واحد. ومع ذلك، ينبغي للمرء أن يكون حذراً وألا يتوقع بشكل عام أن يكون الاختلاف بين أي زوج من الدرجات ODG بمقدار عُشر الدرجة كبيراً. وتسري نفس الملاحظة عند النظر في نتائج اختبار الاستماع الشخصي.

ولا يوجد رقم واحد يصف بشكل كامل دقة طريقة القياس الموضوعي. وبدلاً من ذلك، يتعين على المرء أن يأخذ في الاعتبار عدداً من أرقام الجدارة المختلفة. وأحد هذه الأرقام هو العلاقة بين الدرجات SDG وODG. ومن المهم أن نفهم أنه لا يوجد ضمان بأن الارتباط سيتجاوز قيمة محددة سلفاً. ومن المرجح إلى حد كبير أن يختلف أداء طريقة القياس، على سبيل المثال، باختلاف نوع ومستوى التدهور الـمُدخل.

وهناك رقم آخر يستحق الاهتمام وهو عدد القيم المتطرفة. وتُعرف القيمة المتطرفة على أنها قيمة مقاسة لا تلبي مخطط التفاوت المحدد مسبقاً. ووفقاً لمتطلبات المستخدم، ينبغي أن توفر طريقة القياس أعلى دقة ممكنة للطرف العلوي من سلم الدرجات (أي جودة صوت عالية). وبالتالي، يُسمح للدقة المتحصل أن تكون أقل في النطاق الأوسط والأدنى من سلم الدرجات.

وعلى الرغم من أن الارتباط عادة ما يعطي تقديراً جيداً لدقة طريقة القياس الموضوعي، فمن المهم أن نأخذ في الاعتبار أنه حتى رقم الارتباط المرتفع نسبياً يمكن أن يخفي أداء غير مقبول (من منظور القيم المتطرفة) لطريقة القياس.

ورقم الجدارة الثالث الذي تم استخدامه أثناء عملية التحقق من الصحة هو درجة الخطأ المطلق (AES)، والتي تعكس متوسط العلاقة بين حجم فاصل الثقة بالنسبة للدرجة SDG والمسافة بين الدرجتين SDG وODG.

ويمكن الاطلاع على مزيد من التفاصيل حول الأداء المتوقع لطريقة القياس وكذلك الأداء أثناء عملية التحقق من الصحة في المرفق 1 بالملحق 2.

# 6 المتطلبات والقيود

يجب أن تكون الإشارة الصادرة عن الجهاز قيد الاختبار والإشارة المرجعية متراصفتين زمنياً بدقة 24 عينة خلال فترة القياس الكاملة. ولا تشكل آلية التزامن جزءاً من هذه التوصية ومن المتوقع أن تختلف من تنفيذ إلى آخر.

المرفق 1  
بالملحق 1  
  
التطبيقات

# 1 عام

يقدم هذا المرفق التعاريف والمتطلبات المحددة للتطبيقات الرئيسية التي تستهدفها طريقة القياس الموضوعي الموصى بها لجودة الصوت المدركة.

وتتطلب بعض التطبيقات تنفيذاً في الوقت الفعلي لطريقة القياس الموضوعي، بينما يكون القياس في غير الوقت الفعلي كافياً بالنسبة لتطبيقات أخرى. وبالنسبة للتنفيذ في الوقت الفعلي، يوصى بألا يتجاوز الحد الأقصى للتأخير عبر معدات القياس 200 ms وألا يُقبل التأخير إذا كان أكثر من ثانية واحدة.

وعلاوة على ذلك، لا بد من التمييز بين القياسات عبر الانترنت وخارجها. ففي القياسات خارج الإنترنت، يتمتع إجراء القياس بإمكانية النفاذ الكامل إلى الجهاز أو التوصيل به بينما يشير القياس عبر الإنترنت إلى أن البرنامج قيد التشغيل، وهو ما يجب عدم قطع تشغيله بواسطة القياس.

# 2 التطبيقات الرئيسية

## 1.2 تقييم عمليات التنفيذ

تحتاج الهيئات الإذاعية ومشغلو الشبكات وغيرهم إلى تقييم عمليات التنفيذ المختلفة للمعدات، ولا سيما الكوديكات الصوتية، وذلك عند اختيار هذه المعدات لشرائها أو عند إجراء اختبارات القبول.

وبالنسبة لهذا النوع من التطبيقات، يلزم توفر دقة عالية خاصة لتقييم الانحطاطات الصغيرة وتصنيف التطبيقات المختلفة بشكل صحيح. وفيما يتعلق بمتغيرات الخرج، يعد الخرج البسيط مثل الدرجة ODG كافياً للمستعملين، ولكن يمكن لمطوري الكوديكات الصوتية إجراء تحليل أكثر شمولاً باستخدام مجموعة مناسبة من متغيرات خرج النماذج (MOV).

يمكن استخدام صيغتي النموذج كلتيهما، ولكن يوصى بالصيغة المتقدمة.

## 2.2 ترتيب الجودة الإدراكية

هذا إجراء سريع يتم قبل إدخال جزء من المعدات أو الدارة في الخدمة. والهدف هو التحقق من الأداء الوظيفي والجودة. ويتولى موظفو التشغيل التعامل مع معدات القياس. وقد يكون هناك أي نوع من التشوه.

ويلزم إجراء القياس في الوقت الفعلي. ويمكن استخدام إشارات الاختبار أو الإشارات الصوتية المحددة مسبقاً. وينبغي عرض الدرجات ODG بشكل صحيح وينبغي تقديمها مرتين على الأقل في الثانية، أو مباشرة بعد نهاية إشارة الاختبار في حالة استخدام إشارة اختبار خاصة.

ويكفي استخدام الصيغة الأساسية.

## 3.2 المراقبة عبر الإنترنت

هذه عملية مستمرة تحدث أثناء البث الصوتي المستمر. ويجب ألا يتم قطع تشغيل البرنامج بإجراء القياس. ومن ثم، يجب استخدام إشارة البرنامج نفسها أو جزء صوتي محدد مسبقاً لمن أجل القياس. وقد يكون الأخير إشارة محطة إذاعية أو ترنيمة. ويتولى موظفو التشغيل التعامل مع معدات القياس.

ويلزم إجراء القياس في الوقت الفعلي. ويجب أن يتم عرض الدرجات ODG بشكل صحيح ويجب تقديمها مرتين على الأقل في الثانية الواحدة أو مباشرة بعد نهاية الإشارة المحددة مسبقاً. ولا يُحبذ عرض الدرجات MOV.

ويكفي استخدام الصيغة الأساسية.

## 4.2 حالة المعدات أو التوصيلة

لضمان عمل التوصيلات الصوتية أو المعدات، يلزم إجراء فحص مكثف للجودة من وقت لآخر. وعلى النقيض من المراقبة عبر الإنترنت أو الترتيب الإدراكي، يتطلب هذا التطبيق التحقق من العديد من المعلمات التقنية.

وينبغي لنظام القياس أن يقدم معلومات مفصلة عن تأثير حالة المعدات أو التوصيلة على جودة الصوت المدركة من خلال عرض المجموعة الكاملة من المتغيرات MOV بالإضافة إلى الدرجات ODG. ولا يلزم إجراء القياس في الوقت الفعلي.

ويوصى باستخدام الصيغة المتقدمة.

## 5.2 تحديد هوية الكوديك

لتحديد هوية الكوديكات (خوارزميات مختلفة أو عمليات تنفيذ مختلفة لنفس الخوارزمية)، يجب أن يكون نظام القياس قادراً على تخزين واسترجاع ومقارنة أنماط الخصائص. ويمكن اعتبار التشابه بين الأنماط بمثابة مقياس للتشابه بين عمليات التنفيذ المختلفة للكوديكات. ويتم استخدام مثل هذا الإجراء لتحديد نوع أي كوديك معين وتنفيذه.

ويجب أن يسجل نظام القياس أكبر قدر ممكن من المعلومات عن الأنماط. والنظر في الدرجات ODG فقط قد لا يوفر معلومات كافية.

ويعد استخدام الصيغة الأساسية كافياً، على الرغم من عدم الحاجة إلى القياس في الوقت الفعلي.

**ملاحظة** - لا يوجد سوى القليل من الخبرة في استخدام الطريقة الموصى بها. علاوة على ذلك، لم يتم بعد تحديد مقياس واحد للتشابه بين الأنماط.

## 6.2 تطوير الكوديكات

بالنسبة لهذا التطبيق، يجب أن تحدد طريقة القياس خصائص أداء الكوديك قيد الاختبار بأكبر قدر ممكن من الدقة وبأكبر قدر ممكن من التفصيل، ولا سيما بالنسبة للانحطاطات الصغيرة.

وتتطلب اختبارات المراقبة المستمرة معالجة في الوقت الفعلي وهو ما لا تدعمه الصيغة المتقدمة بالضرورة. ومع ذلك، سوف تتطلب حالات التدهور الصغيرة والمعلومات التفصيلية الصيغة المتقدمة. ويجب أن يكون نظام القياس قادراً على عرض المخرجات بنفس المعدل الذي تم حسابها به. ومن المرغوب فيه النفاذ المباشر إلى تاريخ المخرجات خلال فترة مدتها 4 ثوانٍ.

ويُوصى باستخدام الصيغة المتقدمة. ومع ذلك، فإنه للقياس في الوقت الفعلي، تكفي الصيغة الأساسية. ويلزم إجراء تحليل في الوقت الفعلي وفي غير الوقت الفعلي للرتل تلو الآخر. ولا بد من الإشارة إلى أي تشوه شديد، على سبيل المثال، بواسطة عرض الذروة. ومن المرغوب فيه النفاذ إلى المجموعة الكاملة من المتغيرات MOV.

## 7.2 تخطيط الشبكة

يتطلب تخطيط الشبكات تقييم الجودة المتوقعة في نقاط مختلفة أثناء عملية التخطيط. ويمكن استخدام محاكاة برمجية لمكونات الشبكة، والتي تسمح بدمج مراحل مختلفة لمعالجة الصوت، لفحص التشكيلات المختلفة من أجل تحسين جودة الصوت. وفي مرحلة لاحقة، يمكن اختبار مكونات معالجة الصوت الفعلية في التشكيلة المختارة.

ويتولى مهندسو النظام تخطيط الشبكة حيث ينبغي لهم استرجاع معلومات مفصلة عن تأثير خصائص الشبكة على جودة الصوت. يجب أن يعتمد تصنيف تشكيلات الشبكة المحتملة المختلفة على مجموعة مناسبة من المتغيرات MOV حسب التطبيق المحدد للشبكة. ومن ثم فإن عرض الدرجات ODG فقط لا يكفي. ولا يلزم إجراء القياس في الوقت الفعلي للتقييم في هذا التطبيق.

يمكن استخدام صيغتي النموذج كلتيهما، ولكن يوصى بالصيغة المتقدمة.

## 8.2 المساعدة في التقييم الشخصي

توفر طريقة القياس الموضوعي أداة لفرز المواد الصوتية المهمة المقرر استخدامها في اختبارات الاستماع الشخصية. ويمكن استخدام المجموعة الكاملة من المتغيرات MOV لتصنيف المواد الهامة.

ويلزم تحقيق أعلى دقة ممكنة ويوصى باستخدام الصيغة المتقدمة. ومع ذلك، يُحبذ القياس في الوقت الفعلي لتقليل الوقت اللازم لاختيار المواد الحرجة.

## 9.2 ملخص التطبيقات

يلخص الجدول 2 المتطلبات فيما يتعلق بطريقة القياس للتطبيقات الرئيسية.

الجدول 2

المتطلبات فيما يتعلق بطريقة القياس

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | التطبيق | الفئة | الوقت الفعلي | الحد الأدنى لمعدل قيم الخرج،  ROV (1) (Hz) | عبر/خارج الإنترنت | صيغة النموذج |
| 1 | تقييم عمليات التنفيذ | تشخيصي | لا | – | خارج الإنترنت | الصيغتان |
| 2 | ترتيب الجودة الإدراكية | تشغيلي | نعم/لا | 2 | خارج الإنترنت | الأساسية |
| 3 | المراقبة عبر الإنترنت | تشغيلي | نعم | 2 | عبر الإنترنت | الأساسية |
| 4 | حالة المعدات أو التوصيلة | تشخيصي | نعم/لا | – | عبر/خارج الإنترنت | المتقدمة |
| 5 | تحديد هوية الكوديك | تشخيصي | لا | – | خارج الإنترنت | الصيغتان |
| 6 | تطوير الكوديكات | تطويري | نعم/لا | – | خارج الإنترنت | الصيغتان |
| 7 | تخطيط الشبكة | تطويري | نعم/لا | – | خارج الإنترنت | الصيغتان |
| 8 | المساعدة في التقييم الشخصي | تطويري | نعم/لا | – | خارج الإنترنت | المتقدمة |
| (1) معدل قيم الخرج (في الثانية) | | | | | | |

# 3 إشارات الاختبار

تُقسم إشارات الاختبار إلى مجموعتين: طبيعية واصطناعية. وتتألف قائمة إشارات الاختبار الطبيعية المقدمة هنا من تتابعات صوتية حرجة مستخدمة بالفعل في اختبارات الاستماع التي أجراها قطاع الاتصالات الراديوية ومنظمات أخرى لتقييم جودة الصوت. ويجب أن تكون الإشارات متاحة في موقع الإرسال وفي موقع القياس. وبالتالي، لا بد من وجود ذاكرة في جهاز القياس.

يتم تعريف الإشارات الاصطناعية رياضياً ويمكن تغييرها بطريقة يمكن التحكم فيها. ويمكن توليد هذه الإشارات في موقعي الإرسال والقياس. ولا توجد حاجة إلى ذاكرة إضافية في جهاز القياس. ونظراً لطبيعة هذه الإشارات، فمن الصعب، إن لم يكن من المستحيل، استخلاص درجات تقييم شخصية لها. ولذلك، لم يتم التحقق من صحة طريقة القياس مقابل النتائج الشخصية لهذه الإشارات.

## 1.3 اختيار إشارات الاختبار الطبيعية

يقدم الجدول 3 قائمة بمجموعة فرعية من إشارات الاختبار التي استخدمت أثناء إجراء التحقق الذي أنتج هذه التوصية. ويشار أيضاً إلى نوع الأخطاء البشرية، التي تكشف عنها هذه الإشارات عادةً بسبب التشفير ذي معدل البتات المنخفض.

الجدول 3

قائمة بمجموعة فرعية من إشارات الاختبار

| الرقم | العنصر | اسم الملف | ملاحظات |
| --- | --- | --- | --- |
| 1 | صنجات | cas | (1) |
| 2 | الكلارينيت | cla | (2) |
| 3 | العصا | clv | (1) |
| 4 | الفلوت | flu | (2) |
| 5 | آلة الاكسليفون | glo | (1) و(2) و(5) |
| 6 | البيان القيثاري | hrp | (1) و(2) و(4) |
| 7 | طبلة الغلاية | ket | (1) |
| 8 | الماريمبا | mar | (1) |
| 9 | بيانو شوبرت | pia | (2) |
| 10 | أنبوب النغمات المرجعية | pip | (4) |
| 11 | راي كودر | ryc | (2) و(4) |
| 12 | الساكسفون | sax | (2) |
| 13 | موسيقى القرب | sb1 | (2) و(4) و(5) |
| 14 | كلام لأنثى باللغة الإنكليزية | sfe | (3) |
| 15 | كلام لذكر باللغة الإنكليزية | sme | (3) |
| 16 | كلام لذكر باللغة الألمانية | smg | (3) |
| 17 | الطبول الجانبية | sna | (1) |
| 18 | سوبرانو لموزارت | sop | (4) |
| 19 | دف صغير (رق) | tam | (1) |
| 20 | بوق | tpt | (2) |
| 21 | آلة المثلث الإيقاعية | tri | (1) و(2) و(5) |
| 22 | آلة توبا | tub | (2) |
| 23 | سوزان فيجا | veg | (3) و(4) |
| 24 | إكسيليفون | xyl | (1) و(2) |
| (1) إشارات عابرة: حساسة لما قبل الصدى، وإثارة الضوضاء في الميدان الزمني.  (2) بنية نغمية: حساسة للضوضاء، خشونة.  (3) كلام طبيعي: (مزيج حرج من الأجزاء النغمية والأصوات الحادة): حساسة للتشوهات وإثارة الصوات الحادة.  (4) صوت مركب: يضغط على الجهاز قيد الاختبار.  (5) عرض نطاق كبير: يضغط على الجهاز قيد الاختبار، فقدان الترددات العالية، ضوضاء ذات تردد عال مشكلة بالبرنامج. | | | |

## 2.3 مدة الإشارة

ينبغي أن تكون مدة إشارة الاختبار الطبيعية هي نفس مدة إشارة الاختبار الطبيعية تقريباً كما لو كانت ستُستخدم في اختبار استماع. وتتراوح المدة عادةً من 10 إلى 20 ثانية. ومن المحتمل جداً أن يقتصر الجزء الحرج من إشارة الاختبار، الذي يكشف عن معظم الأخطاء البشرية، على جزء قصير فقط من المدة.

وينبغي أن تكون مدة إشارات الاختبار الاصطناعية طويلة بما يكفي للضغط على الكوديك قيد الاختبار، والذي قد يحتوي على دارئ للإشارة الصوتية المشفرة. وبالنظر إلى أطوال الدارئ هذه والثوابت الزمنية الموجودة في طريقة القياس، يجب أن تكون مدة كل عنصر اختبار واحد في التتابع أكثر من ms 500. ويمكن أن تقتصر المدة على هذه القيمة القصيرة لأنه ليس من المتوقع استخدام هذه الإشارات في اختبارات الاستماع الشخصية.

# 4 التزامن

بالنسبة لإجراء القياس، يجب أن تتم مزامنة الإشارة قيد الاختبار والإشارة المرجعية من حيث الزمن. وينطبق ذلك على إشارات الاختبار الطبيعية والاصطناعية.

# 5 مسائل حقوق التأليف والنشر

لا يمكن استعمال إشارات الاختبار الواردة في الجدول 3 دون حقوق التأليف والنشر إلا لأغراض القياس إلى جانب طريقة القياسات الموضوعية الموصوفة في الملحق 2 بهذه التوصية.

**ملاحظة** - يجب الحصول على إعفاء من حقوق التأليف والنشر لجميع التتابعات، بشكل رئيسي من اتحاد الإذاعات الأوروبية (EBU) (قرص EBU SQAM).

المرفق 2  
بالملحق 1  
  
متغيرات الخرج

# 1 مقدمة

تقيس طريقة القياس الموضوعي الموصوفة في هذه التوصية جودة الصوت وتخرج قيمة تهدف إلى أن تتوافق مع جودة الصوت المدركة. وتحدد طريقة القياس الخصائص الأساسية للنظام الصوتي. وتحدد عدة مراحل وسيطة التأثيرات الصوتية الفسيولوجية والنفسية.

ويمكن استخدام هذه المخرجات الوسيطة لتحديد خصائص الأخطاء البشرية. وتسمى المعلمات متغيرات خرج النموذج (MOV). وتجمع المرحلة النهائية من نموذج القياس بين قيم المتغيرات MOV لإنتاج قيمة خرج واحدة تقابل بشكل مباشر النتيجة المتوقعة من تقييم شخصي للجودة.

# 2 متغيرات خرج النموذج

يحتوي الجدول 4 على وصف لـلمتغيرات MOV المستخدمة للتنبؤ بدرجات الاختلاف الموضوعي. يتم اشتقاق الرموز SubscriptsA من جزء بنك التصفية في النموذج، بينما يتم اشتقاق الرموز SubscriptsB من جزء المحول FFT في النموذج. ويمكن التنبؤ بدرجات الاختلاف الموضوعي إما من جزء المحول FFT فقط (الصيغة الأساسية) أو من مزيج من جزئي المحول FFT وبنك التصفية (الصيغة المتقدمة). يتم إجراء التوسيط دائماً مع مرور الوقت.

# 3 جودة الصوت الأساسية

المعلمة الأكثر شهرة في اختبارات الاستماع الشخصية هي جودة الصوت الأساسية (BAQ). وتُقاس الجودة BAQ كدرجة اختلاف شخصية (SDG) يتم حسابها على أنها الدرجة المعطاة للمرجع مطروحة من الدرجة المعطاة للإشارة قيد الاختبار في اختبار شخصي[[1]](#footnote-1). عادة ما تكون للدرجة SDG قيمة سالبة. وتسمى معلمة الخرج المقابلة من النموذج بدرجة الاختلاف الموضوعي (ODG). ويعتمد التقابل بين المتغيرات MOV والدرجة ODG على عدد كبير من عناصر الاختبار الموثوقة، انظر المرفق 2 بالملحق 2.

الجدول 4

وصف معلمات خرج النموذج

|  |  |
| --- | --- |
| معلمة خرج النموذج | الوصف |
| *WinModDiffB* | *متوسط الاختلاف المقسم إلى نوافذ في التشكيل (الأغلفة) بين الإشارة المرجعية والإشارة قيد الاختبار* |
| *AvgModDiff1B* | *متوسط الاختلاف في التشكيل* |
| *AvgModDiff2B* | *متوسط الاختلاف في التشكيل مع التركيز على التشكيلات المقدمة وتغييرات التشكيل حيث يحتوي المرجع على تعديلات قليلة أو معدومة* |
| *RmsModDiffA* | *قيمة جذر متوسط التربيع للاختلاف في التشكيل* |
| *RmsMissingComponentsA* | *قيمة جذر متوسط التربيع لجهارة الضوضاء لمكونات التردد المفقودة (المستخدمة في المتغير RmsNoiseLoudAsymA)* |
| *RmsNoiseLoudB* | *قيمة جذر متوسط التربيع لمتوسط جهارة الضوضاء مع التركيز على المكونات المقدمة* |
| *RmsNoiseLoudAsymA* | *المتغير RmsNoiseLoudA + 0,5 المتغير RmsMissingComponentsA* |
| *AvgLinDistA* | *مقياس لمتوسط التشوهات الخطية* |
| *BandwidthRefB* | *عرض نطاق الإشارة المرجعية* |
| *BandwidthTestB* | *عرض نطاق إشارة الخرج للجهاز قيد الاختبار* |
| *TotNMRB* | *لوغاريتم متوسط إجمالي نسبة الضوضاء إلى القناع* |
| *RelDistFramesB* | *الجزء النسبي من الأرتال التي يحتوي نطاق تردد واحد على الأقل فيها على مكون ضوضاء مهم* |
| *AvgSegmNMRB* | *اللوغاريتم المتوسط القطاعي لنسبة الضوضاء إلى القناع* |
| *MFPDB* | *الحد الأقصى لاحتمال الكشف بعد ترشيح التمرير المنخفض* |
| *ADBB* | *متوسط المجموعة المشوهة (=الرتل)، يتم اعتباره لوغاريتم نسبة التشوه الإجمالي إلى إجمالي عدد الأرتال شديدة التشوه* |
| *EHSB* | *البنية التوافقية للخطأ مع مرور الوقت* |

الدرجة ODG هي المعلمة المقاسة بشكل موضوعي والتي تقابل الجودة المدركة شخصياً. ونظراً لأن مهمة المستمع في اختبار الاستماع هي تقييم الجودة BAQ لعنصر الاختبار، فإن الدرجة ODG تعد أيضاً مقياساً للجودة BAQ.

# 4 هامش التشفير

هناك معلمة أخرى قد تكون ذات قيمة في المستقبل وهي هامش التشفير (CM)، وهي طريقة لوصف الأخطاء البشرية غير المسموعة. ويمكن تقييم هامش التشفير الشخصي (SCM) عن طريق تكبير الأخطاء البشرية حتى تصبح مسموعة للشخص القائم بالاختبار. ويصف الهامش SCM الارتفاع إلى عتبة إمكانية سماع الأخطاء البشرية.

ولإيجاد العتبة، يجب تكبير الأخطاء البشرية أو توهينها أثناء اختبار الاستماع. والطريقة المناسبة هي طريقة الاختلاف. ويتم تكبير إشارة الفرق للإشارة الأصلية والمشفرة المتزامنة مع الوقت وإضافتها إلى الإشارة الأصلية. ومن الأفضل إجراء اكتشاف عتبة السمعية باستخدام طريقة الاختيار القسري. ويتم الحصول على هامش SCM عن طريق حساب متوسط قيم العتبة للتكبير أو التوهين التي تم الحصول عليها من الأشخاص القائمين بالاختبار. تمثل قيم الهامش CM السالبة أخطاء مسموعة بينما تمثل قيم الهامش CM الموجبة أخطاء غير مسموعة. وعلى عكس الجودة BAQ، يعد هامش التشفير مقياساً للوقت الذي تصبح فيه الخطاء البشرية (على أي مستوى) مسموعة وليس مدى إزعاج هذه الأخطاء. ويرد وصف التعريف والتحقق من صحة طريقة قياس الهامش SCM في [Feiten، 1997].

وهامش التشفير الموضوعي (OCM) مشتق أيضاً من المتغيرات MOV. وفي الوقت الحاضر، لم يتم تقييم سوى عدد قليل من بنود الاختبار لهامش التشفير الشخصي. ولم تتم بعد دراسة التقابل إلى الهامش OCM من النموذج الوارد في هذه التوصية.

# 5 متطلبات المستعمل

تختلف متطلبات المستعمل فيما يتعلق بمتغيرات الخرج من طريقة القياس باختلاف التطبيق. وبالنسبة لبعض التطبيقات، فإن الرقمين 2 و3، على سبيل المثال (انظر المرفق 1 بالملحق 1)، يكون القياس جزءاً من إجراء تشغيلي. وفي هاتين الحالتين، من المهم جداً أن تكون نتائج الطريقة سهلة القراءة وسهلة الفهم للأشخاص الذين ليس لديهم معرفة متعمقة بتقنية القياس. ويتحقق ذلك على أفضل وجه إذا قامت الطريقة بإخراج **قيمة واحدة** فقط تتوافق مع جودة الصوت المدركة.

وقد ينطبق الشيء نفسه أيضاً على التطبيقات الأخرى، على سبيل المثال، التطبيقان 1 و4. ومع ذلك، بالنسبة لهذين التطبيقين، وكذلك للتطبيقات 5-8، قد تكون متغيرات الخرج الأكثر تعقيداً مفيدة للمستعملين الذين لديهم معرفة أعمق عن آليات طريقة القياس.

المرفق 3  
بالملحق 1  
  
الخطوط العريضة للنموذج

وفقاً للتوصية ITU-R BS.1116، يتم الحصول على الدرجة SDG لعنصر اختبار صوتي في اختبار استماع، ويمثل متوسط الدرجات SDG على عدد من المستمعين الجودة الشخصية للعنصر. وقد يحتوي العنصر على أنواع مختلفة من التشوهات الصوتية، لذلك يتم دمج التغايرات في الجودة بمرور الوقت. لذلك، يتطلب التنبؤ بالدرجات SDG بناءً على القياسات المادية نموذجاً دقيقاً للنظام السمعي المحيطي بالإضافة إلى الجوانب الإدراكية لأحكام جودة الصوت.

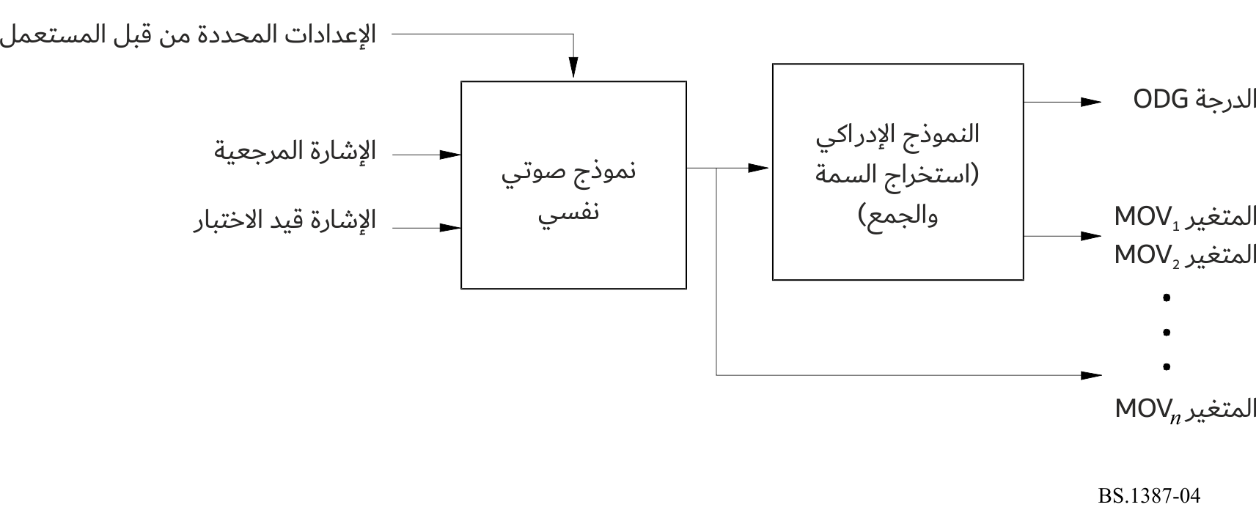
وينتج النموذج الموصى به للقياس الموضوعي عدداً من متغيرات خرج النموذج (MOV) بناءً على مقارنات بين الإشارة المرجعية والإشارة قيد الاختبار. ويُجرى تقابل بين المتغيرات MOV هذه والدرجة ODG باستخدام تقنية استمثال ي تقلل إلى أدنى حد من الفرق التربيعي بين توزيع الدرجات ODG والتوزيع المقابل لمتوسط الدرجات SDG لمجموعة بيانات كبيرة بما فيه الكفاية.

تم وصف نوعين مختلفين من النموذج - صيغة قائمة على تحويل فورييه المتمايز (DFT) يمكن استخدامها للمراقبة في الوقت الفعلي، وصيغة أخرى، تعتمد على كل من بنك التصفية والتحويل DFT، والتي كان من المتوقع أن تعطي نتائج أكثر دقة. يُطلق على الصيغة المستندة إلى التحويل DFT اسم الصيغة الأساسية، بينما يُطلق على الصيغة المركبة اسم الصيغة المتقدمة.

ويعرض الشكل 4 البنية رفيعة المستوى لكل من الصيغة الأساسية والصيغة المتقدمة.

الشكل 4

مراحل المعالجة المنفذة في النموذج



# 1 معالجة الصوت

كما هو الحال في اختبارات الاستماع الشخصية، يتم الحكم على جودة إشارة الاختبار بالنسبة للإشارة المرجعية. يتم تحويل كل من الإشارة المرجعية والإشارة قيد الاختبار (الإشارات غير المجسمة أو المجسمة) إلى تمثيل صوتي نفسي. وتُجرى مقارنة بين هذه التمثيلات من أجل استخلاص الدرجة ODG. ويتم تنفيذ هذه العمليات من خلال مراحل المعالجة الموضحة في الشكل 4.

## 1.1 الإعدادات المحددة من قبل المستعمل

تتطلب طريقة القياس مستوى الاستماع المفترض كمعلمة. ولذلك، يتعين على المستعمل توفير مستوى ضغط الصوت بوحدة dB SPL الناتج عن موجة جيبية كاملة النطاق بتردد Hz 1 019,5. وفي حالة عدم معرفة مستوى الاستماع الدقيق، يوصى بافتراض مستوى استماع يبلغ dB SPL 92.

## 2.1 النموذج الصوتي النفسي

يحول النموذج النفسي الصوتي الأرتال المتعاقبة لإشارة المجال الزمني إلى تمثيل غشاء قاعدي. وتبدأ هذه العملية باستخدام كل من التحويل DFT وبنك التصفية. ويقوم التحويل DFT بتحويل البيانات إلى مجال التردد، وتتم مقابلة النتيجة من مقياس التردد إلى مقياس درجة الصوت، وهو المكافئ النفسي الصوتي للتردد. وفي جزء بنك التصفية من النموذج، يتم أخذ التقابل بين التردد ودرجة الصوت في الاعتبار مباشرةً من خلال عروض النطاق والمباعدة بين مرشحات تمرير النطاق.

يُستخدم مفهومان مختلفان لتحقيق التقنيع المتزامن. وتُحسب بعض المتغيرات MOV باستخدام *مفهوم العتبة المقنعة*، بينما يعتمد البعض الآخر على *مقارنة التمثيلات الداخلية*. ويحسب المفهوم الأول العتبة المقنعة مباشرة باستخدام وظائف التقنيع النفسي والمادي. وتعتمد متغيرات خرج النموذج على المسافة بين إشارة الخطأ المادية والعتبة المقنعة تلك. وفي مفهوم مقارنة التمثيلات الداخلية، يتم نشر طاقة كل من الإشارة قيد الاختبار (SUT) والإشارة المرجعية إلى مناطق الصوت المجاورة من أجل الحصول على أنماط الإثارة. وتعتمد متغيرات خرج النموذج على إجراء مقارنة بين أنماط الإثارة تلك. ويُنفد التقنيع غير المتزامن عن طريق تلطيخ تمثيلات الإشارة بمرور الوقت.

تُنمذج العتبة المطلقة جزئياً من خلال تطبيق وظيفة للترجيح المعتمد على التردد وجزئياً عن طريق إضافة إزاحة تعتمد على التردد إلى أنماط الإثارة. وهذه العتبة عبارة عن تقدير تقريبي للحد الأدنى من الضغط المسموع [ISO 389-7، الصوتيات - الصفر المرجعي لمعايرة أجهزة قياس الصوت - الجزء 7: العتبة المرجعية للاستماع في ظل ظروف الاستماع للمجال الحر ومجال الانتثار، 1996].

والمخرجات الرئيسية للنموذج الصوتي النفسي هي الإثارة والعتبة المقنعة كدالة في الزمن والتردد. ويُتاح خرج النموذج على عدة مستويات لمزيد من المعالجة.

## 3.1 النموذج الإدراكي

يقوم النموذج الإدراكي بتكثيف المعلومات المستقاة من تتابع من الأرتال التي ينتجها النموذج الصوتي النفسي. وأهم مصادر المعلومات لإجراء قياسات الجودة هي الاختلافات بين الإشارة المرجعية والإشارة قيد الاختبار في كل من مجال التردد ومجال الصوت. ففي مجال التردد، يتم قياس عرض النطاق الطيفي لكلا الإشارتين، وكذلك البنية التوافقية في الخطأ. وعند قياس الخطأ في مجال الصوت، تُشتق مقاييس الخطأ من كل من تشكيل غلاف الإثارة ومقدار الإثارة.

وتُوزن الميزات المحسوبة، بحيث يؤدي تجميعها إلى درجة ODG قريبة بالقدر الكافي من الدرجة SDG للتشويه الصوتي الخاص محل الاهتمام. وتستخدم الصيغة الأساسية 11 سمة لإنتاج درجة ODG واحدة، بينما تستخدم الصيغة المتقدمة خمس سمات. وأُجري الاستمثال باستخدام خوارزمية تعلم الشبكة العصبية ذات الانتشار الخلفي (انظر الملحق 2، الفقرة 6). وتتألف بيانات التدريب من قاعدتي البيانات 1 و2 بالكامل وجزء من قاعدة البيانات 3. وتم الحصول على بيانات اختبار التعميم من بقية قاعدة البيانات 3 وقاعدة البيانات CRC97 بالكامل (انظر المرفق 2 بالملحق 2).

المرفق 4  
بالملحق 1  
  
مبادئ وخصائص طرائق قياس جودة الصوت الإدراكية الموضوعية

# 1 مقدمة والتسلسل التاريخي

يعتمد الإرسال الرقمي للإشارات الصوتية وتخزينها بشكل متزايد على خوارزميات خفض البيانات، والتي تتكيف مع خواص الجهاز السمعي البشري وتعتمد بشكل خاص على تأثيرات التقنيع. ولا تهدف هذه الخوارزميات بشكل أساسي إلى تدنية التشوهات، بل تحاول التعامل مع هذه التشوهات بطريقة بحيث لا تُدرك إلا بأقل قدر ممكن. ولم يعد ممكناً تقييم جودة هذه المشفرات الإدراكية بطرائق القياس التقليدية، التي تحدد عادة القيمة الإجمالية للتشوه. ومن الأمثلة التي كثيراً ما تُذكر لتوضيح هذه القيود ما يسمى بمعجزة dB 13: الضوضاء المتراكبة ذات البنية الطيفية المتكيفة مع بنية الإشارة الصوتية والتي تكاد تكون غير مسموعة حتى لو انخفضت نسبة الإشارة إلى الضوضاء *(S/N)* غير الموزونة الناتجة إلى dB 13.

ولهذا السبب، تتطلب تقييمات الكوديكات الإدراكية اختبارات استماع من أجل تقييم جودة الصوت. وتتطلب الموثوقية والتكرار الكافيان لاختبارات الاستماع إنفاقاً كبيراً للوقت والعمل.

ويمكن أن تساعد مخططات القياس الموضوعي التي تتضمن خواص الجهاز السمعي البشري في التغلب على هذه المشكلات. وتم نشر هذه الفكرة لأول مرة بواسطة [Schroeder وآخرون، 1979]. وفي هذه الورقة، التي تدور بشكل أساسي حول تشفير الكلام، يتم وصف مخطط قياس "جهارة الضوضاء (NL)".

وفي هذه الورقة، يتم تقدير جهارة الصوت المدركة لإشارة الضوضاء الخاصة بكوديك الكلام، والتي هي الفرق بين إشارة الدخل والخرج له، لكل رتل زمني يبلغ حوالي ms 20. فإذا كانت إشارة الضوضاء مقنعة تماماً، فإن جهارة الصوت المدركة تكون صفراً. ويعمل التقنيع الجزئي على خفض جهارة إشارة الضوضاء غير المقنعة. وتُستمثل العتبة المقنعة المستخدمة لضوضاء تقنيع النغمات ويتم حساب تدهور الكلام النهائي لكل رتل. ولا يتم حساب أي ملخص للجودة الإجمالية لعينة الكلام.

وفي عام 1985 نشر كارجالاينن مخطط القياس "الفرق الطيفي السمعي (ASD)" [Karjalainen، 1985]. ولقد بدأ بعدة أفكار من شرودر وأتال وهال، لكنه استعاض عن التحليل القائم على الأرتال ببنك تصفية به مرشحات متراكبة، وغير طريقة تضمين العتبة المطلقة وأضاف نموذجاً للتقنيع الزمني. وتُعالج كل من إشاراتي الدخل إلى نظام القياس بنفس الطريقة تماماً، مما ينتج عنه نوع من التمثيل الداخلي. وتُقارن هذه التمثيلات الداخلية مع بعضها البعض لشرح الاختلافات المدركة بين إشارتي الدخل والخرج لمخطط تشفير الكلام. ولا يتم حساب أي ملخص للجودة الإجمالية لعينة الكلام. وتُكيف الدقة الزمنية لـلفرق ASD بشكل أفضل مع خواص النظام السمعي البشري ولكنها تزيد من تعقيد الخوارزمية.

وفي عام 1987، نشر براندنبورغ مخطط القياس "نسبة الضوضاء إلى القناع (NMR) [Brandenburg، 1987]، والذي كان من المفترض استخدامه كأداة لتطوير مخططات التشفير الصوتي. تم تقليل تعقيد المخطط مقارنة بـالمخطط NL عن طريق حساب الانتشار على النطاقات الإدراكية باستخدام دالة الانتشار التي تم تصميمها على أنها منحنى الحالة الأسوأ. وتم استمثال العتبة المقنعة المستخدمة لنغمة تقنيع الضوضاء. وأُضيف مخطط بسيط لنمذجة ما بعد التقنيع والعديد من الطرائق لتقييم الجودة المدركة لمقتطفات الصوت الأطول. وكان هذا المخطط هو الأول الذي تم تنفيذه في عتاد الوقت الفعلي.

وفي عام 1989 قدم مور وجلاسبيرغ [Moore، 1989] نموذجاً إدراكياً ولكنهما لم يقدما طريقة للحكم على الجودة المدركة للإشارات الصوتية الضعيفة.

# 2 البنية العامة لطرائق قياس جودة الصوت الإدراكية الموضوعية

تعمل جميع مخططات القياس الإدراكي مع إشارتي دخل: إحداهما تسمى الإشارة المرجعية (REF)، والأخرى تسمى الإشارة قيد الاختبار (SUT). وفي الحالات التي لا يمكن فيها إرسال المرجع إلى جهاز القياس، ولكن الإشارة تكون معروفة جيداً، يمكن أن تكون الإشارة المرجعية مرجعاً داخلياً مخزناً في جهاز القياس نفسه. ومن الضروري أن تكون إشاراتي الدخل متراصفتين زمنياً.

ويمكن دمج النماذج الصوتية النفسية في مخططات القياس بطريقتين مختلفتين. الطريقة الأولى تشبه إلى حد كبير بنية مخططات التشفير الصوتي: تُستخدم الإشارة المرجعية لحساب تقدير العتبة المقنعة الفعلية (انظر أدناه). تتم مقارنة الفرق بين الإشارة قيد الاختبار والإشارة المرجعية بهذه العتبة المقنعة. وتسمى هذه الطريقة "مفهوم العتبة المقنعة" وتستخدم في جهارة الضوضاء والنسبة NMR. ويمكن حساب الفرق بين إشاراتي الدخل إما في المجال الزمني أو كالفرق بين أطياف الطاقة قصيرة الأجل. وتوفر الطريقة الثانية متانة أفضل ضد أخطاء التراصف الزمني ولكنها تقلل من الدقة الزمنية. وعادة ما يكون الفرق في المجال الزمني حساساً جداً لتشوهات الطور، وبالتالي لم يعد يستخدم.

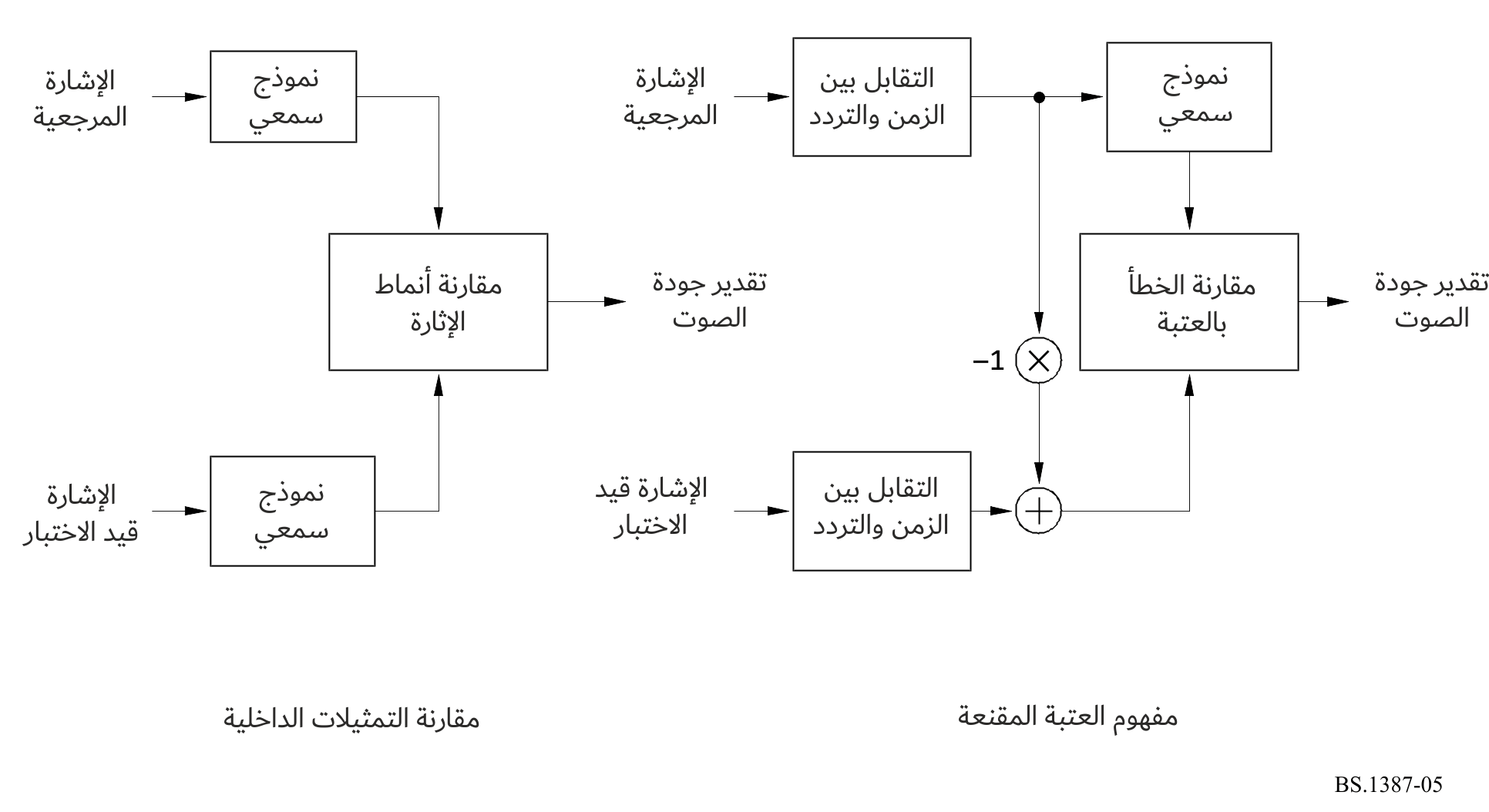
والطريقة الثانية أقرب إلى العمليات الفسيولوجية في الجهاز السمعي البشري: يتم حساب ما يسمى بالتمثيل الداخلي لكل من الإشارة المرجعية والإشارة قيد الاختبار. وهذا التمثيل الداخلي هو تقدير للمعلومات المتوفرة للدماغ البشري لمقارنة الإشارات. وتسمى هذه الطريقة "مقارنة التمثيلات الداخلية" وتستخدم في المخطط ASD.

# 3 الأسس النفسية والصوتية والإدراكية

يناقش هذا القسم خواص الجهاز السمعي البشري التي تعتبر الأبرز في تقييم الجودة المدركة للإشارات الصوتية. وينصب التركيز الرئيسي على كيفية نمذجة هذه الخواص.

الشكل 5

المفاهيم الصوتية النفسية المستخدمة في النهج المختلفة لمخططات القياس الإدراكية



## 1.3 خاصية النقل للأذن الخارجية والوسطى

بوجه عام، يجب أن تمر الإشارات الصوتية عبر الأذن الخارجية والوسطى حتى تصل إلى الأذن الداخلية حيث تُجرى عمليات كشف الصوت وتحليله. وتقوم الأذن الخارجية والوسطى بإجراء عملية مرشح تمرير النطاق على إشارة الدخل. وتُضاف الضوضاء الموجودة في العصب السمعي، إلى جانب الضوضاء الناتجة عن تدفق الدم، إلى إشارة الدخل. ويزداد اتساع هذه الضوضاء مع الترددات المنخفضة. وتحد وظيفة النقل للأذن الخارجية والوسطى إلى جانب الضوضاء الداخلية من القدرة على اكتشاف الإشارات الصوتية الصغيرة، ويكون لها التأثير الأكبر على عتبة السمع المطلقة.

## 2.3 سلالم التردد الإدراكية

مستقبلات ضغط الصوت في الأذن البشرية هي الخلايا الشعرية. وهي تقع في الأذن الداخلية، وبشكل أكثر دقة في القوقعة. وفي القوقعة، يتم إجراء تحويل التردد إلى موضع. ويعتمد موضع الإثارة القصوى على تردد إشارة الدخل. وكل خلية شعرية في موضع معين في القوقعة مسؤولة عن مدى متراكب على سلم التردد. ويرتبط الانطباع الإدراكي للصوت بمسافة ثابتة من الخلايا الشعرية.

واعتماداً على التجربة الصوتية النفسية المستخدمة، يتم العثور على وظائف تحويل مختلفة من التردد إلى الصوت:

وفي [Zwicker وFeldtkeller، 1967] يرد جدول يقسم سلم التردد بوحدات Hz إلى 24 نطاقاً غير متراكب، ما يسمى بالنطاقات الحرجة. وترد في الجدول 5 ترددات القطع العليا لهذه النطاقات. ويحتوي الجدول أيضاً على تعريف لسلم بارك (Bark): Bark 1 يقابل Hz 100، وBark 24 يقابل Hz 15 500.

الجدول 5

سلم النطاقات الحرجة كما عرفه زويكر

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| النطاق الحرج | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| تردد القطع العلوي (Hz) | 100 | 200 | 300 | 400 | 510 | 630 | 770 | 920 | 1 080 | 1 270 | 1 480 | 1 720 |
|  | | | | | | | | | | | | |
| النطاق الحرج | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| تردد القطع العلوي (Hz) | 2 000 | 2 320 | 2 700 | 3 150 | 3 700 | 4 400 | 5 300 | 6 400 | 7 700 | 9 500 | 12 000 | 15 500 |

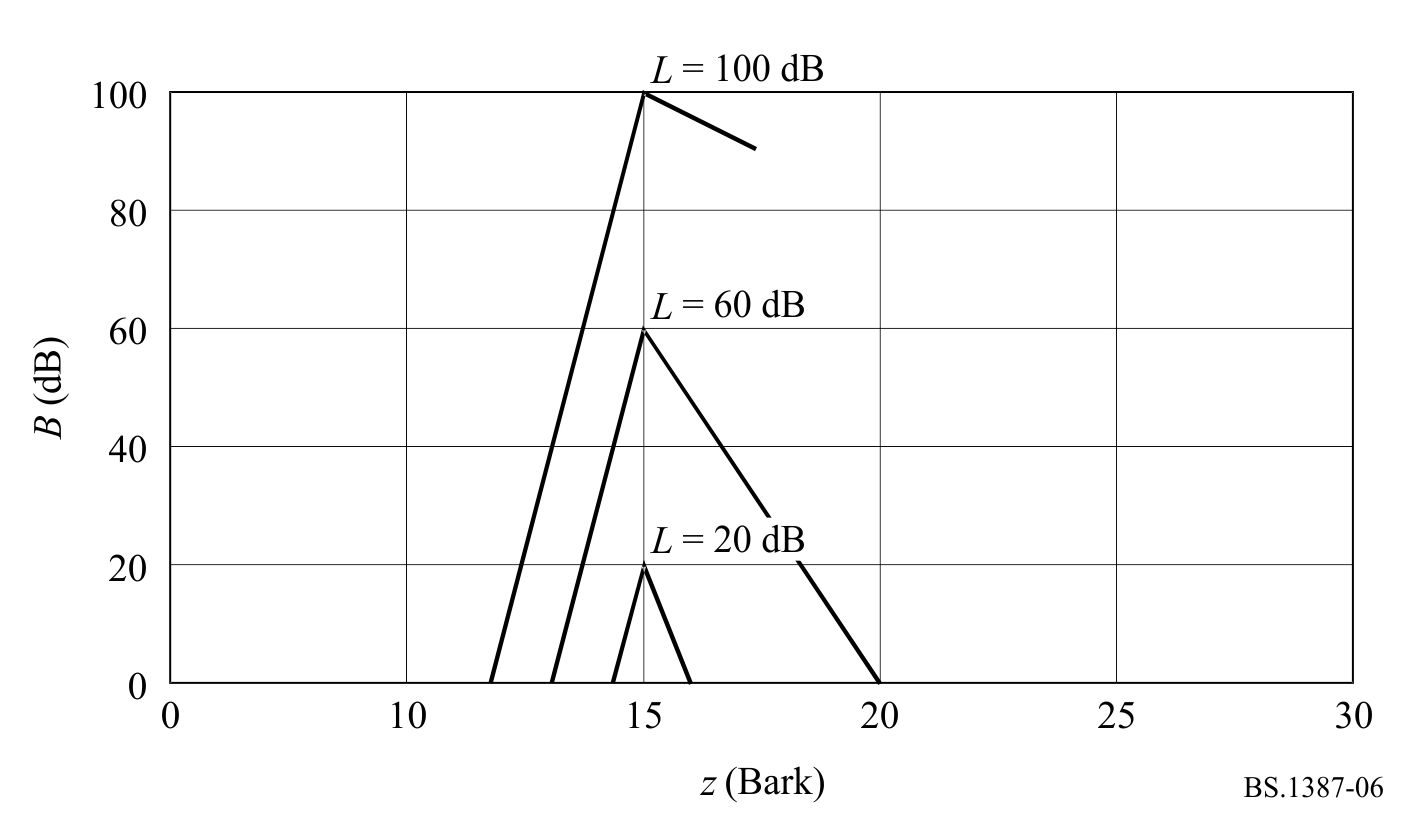
تم في الماضي تحديد عدة تقديرات تقريبية لسلم بارك. ويمكن الاطلاع على مناقشة تفصيلية للسلالم المختلفة في [Cohen وFielder، 1992]. وفي سياق القياس الموضوعي لجودة الصوت المدركة، تم تحقيق أفضل النتائج باستخدام سلم بارك.

## 3.3 الإثارة

تتفاعل كل خلية شعرية مع مجموعة من الترددات التي يمكن وصفها بخاصية المرشح. ويمكن التعبير عن ميل المرشحات بشكل أفضل على سلم إدراكي كما هو موضح أعلاه. ويكون شكل المرشحات على هذا السلم مستقلاً تقريباً عن التردد المركزي. والميل السفلي للإثارة مستقل عن المستوى L لإشارة الدخل (حوالي dB/Bark 27). ويكون الميل العلوي أكثر انحداراً بالنسبة للمستويات الأدنى منه بالنسبة للمستويات الأعلى لإشارة الدخل (من 5– إلى 30– dB/Bark). وتنجم هذه الخاصية المنحدرة عن آلية ردود الفعل بين نوعين مختلفين من الخلايا الشعرية وتحتاج إلى بعض الوقت حتى تستقر. ولذلك يتم تحقيق أفضل دقة للتردد السمعي للإشارات الثابتة بعد عدة ميلي ثانية من بداية الإشارة. وتُضاف أنماط الإثارة للإشارات المكونة من عدة مكونات بطريقة غير خطية.

الشكل 6

تبعية مستوى الإثارة وفقاً لترهاردت [1979]



بعد التعرض للإشارة، تحتاج الخلايا الشعرية والمعالجة العصبية إلى بعض الوقت للتعافي حتى يتم الوصول إلى الحساسية الكاملة مرة أخرى. وتعتمد مدة عملية التعافي على مستوى الإشارة ومدتها ويمكن أن تستمر حتى عدة مئات من الملي ثانية. وتُعالج الإشارات عالية المستوى بشكل أسرع من الإشارات منخفضة المستوى في الطريق بين الخلية الشعرية والدماغ. ولذلك، فإن بداية الإشارة العالية يمكن أن تقنّع الإشارة الأضعف السابقة.

وهناك طريقة أخرى لنمذجة الإثارة تعتمد على سلم ERB [Moore، 1986]. وتستخدم هذه الطريقة ما يسمى بمرشحات ROEX [Moore، 1986]. وفي سياق القياس الموضوعي لجودة الصوت المدركة، تحققت نتائج أفضل باستخدام النماذج المستندة إلى [Zwicker وFeldtkeller، 1967] و[Terhardt، 1979].

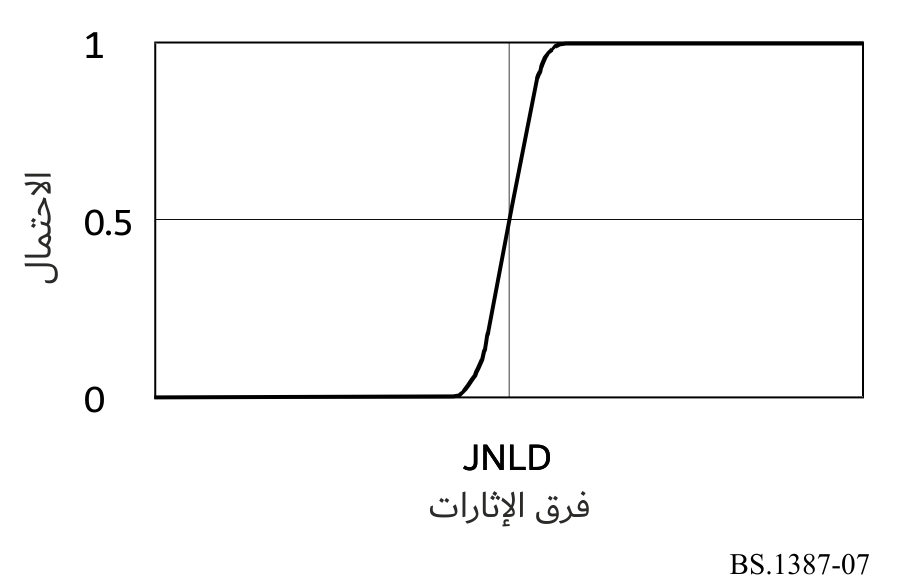
## 4.3 الكشف

تُنقل إثارة الإشارات الصوتية المختلفة إلى الدماغ البشري. وهناك ثلاثة أنواع مختلفة من الذاكرة تختلف حسب درجة التفاصيل ومدة وجود المعلومات: الذاكرة طويلة الأجل، والذاكرة قصيرة الأجل، والذاكرة قصيرة الأجل للغاية. وفي سياق اختبارات الاستماع، تلعب الذاكرة قصيرة الأجل للغاية الدور الأبرز. ويتم الاحتفاظ بمعظم تفاصيل الإشارة إذا كانت مدة المقتطف الصوتي أقل من خمس إلى ثماني ثوانٍ اعتماداً على المستمع والمقتطف الصوتي. ويؤخذ ذلك في الاعتبار في إجراء التقييم المحدد في التوصية ITU-R BS.1116 حيث يُسمح للأشخاص القائمين بالاختبار باختيار أجزاء قصيرة جداً من مقتطف صوتي للاستماع إليها عن كثب. وعند عتبة الكشف يكون احتمال الكشف %50. وحول العتبة، يزداد احتمال كشف الفروق بسلاسة من %0 إلى %100.

إن فرق المستوى الملحوظ بالكاد (JNLD) هو عتبة كشف فروق المستوى. ويتأثر الفرق JNLD بمستوى إشارات الدخل. وبالنسبة للإشارات الصغيرة، يلزم وجود فروق كبيرة للكشف (المستوى: 20 dBSPL، JNLD: 0,75 dB). وبالنسبة للإشارات العالية، تكون الحساسية للفروق الصغيرة أعلى بكثير (المستوى: dBSPL 80، JNLD: 0,2 dB). وتعتمد هذه الأرقام على تجارب التشكيل بالاتساع.

الشكل 7

المبدأ الخاص باحتمال الكشف



## 5.3 التقنيع

الإشارة التي تكون مسموعة بوضوح إذا تم تقديمها بمفردها يمكن أن تكون غير مسموعة تماماً في وجود إشارة أخرى، وهي الإشارة القائمة بالتقنيع. ويُسمى هذا التأثير بالتقنيع وتسمى الإشارة المقنعة بالمقنَع. ويجب التمييز بين حالتين:

*– التقنيع المتآون*

في هذه الحالة، يتم عرض المقنِع والمقنَع في نفس الوقت ويكونان شبه ثابتين. فإذا كان للمقنِع عرض نطاق منفصل، فسيتم رفع عتبة السمع حتى بالنسبة للترددات الأقل أو الأعلى من المقنِع. وتعتمد كمية التقنيع على بنيتي المقنِع والمقنَع. وفي الحالة التي تقوم فيها إشارة شبيهة بالضوضاء بتقنيع إشارة نغمية، يكون مقدار التقنيع مستقلاً عن التردد تقريباً. وإذا كان مستوى ضغط الصوت للمقنَع أقل بحوالي dB 5 عن مستوى المقنِع، يصبح غير مسموع. وفي الحالة التي تقوم فيها الإشارة النغمية بتقنيع إشارة شبيهة بالضوضاء، فإن مقدار التقنيع يعتمد على تردد المقنِع. ويمكن تقدير هذا المقدار بالمعادلة ، حيث *z* هي معدل النطاق الحرج للمقنِع. بالإضافة إلى ذلك، عند مستويات الإشارة العالية، تؤدي التأثيرات غير الخطية إلى خفض العتبة المقنعة بالقرب من المقنِع. وتحدث تأثيرات مماثلة مع الاتجاه نغمة تقنيع نغمة. وتُضاف العتبات المقنعة لعدة إشارات بطريقة غير خطية. وبشكل عام، تكون العتبة المقنعة الناتجة أعلى من العتبة المقنعة التي تنتجها كل إشارة فردية.

*– التقنيع المؤقت*

في هذه الحالة، يتم عرض المقنِع والمقنَع في أوقات مختلفة. وبعد وقت قصير من اضمحلال المقنِع، تكون العتبة المقنَّعة أقرب إلى التقنيع المتآون لهذا المقنِع أكثر من العتبة المطلقة. وتبعاً لمدة المقنِع، يمكن أن يتراوح وقت اضمحلال العتبة بين ms 5 (المقنِع: نبضة غوسية مدتها حوالي ms 0,05) وأكثر من ms 150 (القناع: ضوضاء وردية مدتها ثانية واحدة). والإشارات الضعيفة تكون قبل تقنيع الإشارات الأعلى جهارة مباشرة. وتبلغ مدة تأثير التقنيع العكسي هذا حوالي ms 5. وإذا كان المقنَع أعلى بقليل من العتبة، فلن يُدرك قبل المقنِع ولكن كتغيير للمقنِع. ويُظهر التقنيع العكسي انحرافات كبيرة من مستمع إلى آخر.

## 6.3 الجهارة والتقنيع الجزئي

تعتمد جهارة الصوت المدركة للإشارات الصوتية على ترددها ومدتها ومستوى ضغط الصوت. وبسبب التقنيع التلقائي، تكون جهارة الإشارة المركبة أقل من مجموع جهارة الصوت لجميع مكوناتها. وفي سياق قياس جودة الصوت، يتم تقليل جهارة التشوه غير المرغوب فيه المضاف إلى الإشارة المرجعية، أي جهارة الضوضاء، عن طريق التقنيع الجزئي الناتج عن الإشارة المرجعية.

## 7.3 الحدة

الحدة، إحدى القيم الأساسية للإحساس، ترتبط بطابع الصوت. ويُنظر إلى الصوت على أنه حاد إذا كان يحتوي بشكل أساسي على مكونات عالية التردد. فعلى سبيل المثال، يقال على نغمة جيبية أو ضوضاء محدودة بالنطاق عند الترددات العالية، أو ضوضاء تمرير مرتفع مع تردد قطع في مدى التردد فوق kHz 3 تقريباً أنها حادة. ومع ذلك، فإن بنية التردد التفصيلية للإشارة الصوتية ليس لها تأثير كبير على الحدة. وأُجري البحث الأساسي المتعلق بالحدة بواسطة فون بسمارك [1974].

تم إجراء أبحاث إضافية بخصوص الحدة بواسطة [Aures، 1984]. وكانت نتيجة هذه الأبحاث دالة ترجيح معدلة قليلاً مقارنة بدالة الترجيح التي حددها بسمارك. وهي تساهم بشكل أقل في إسناد الحدة عند معدلات النطاق الحرج المنخفضة جداً والمرتفعة جداً، وأكثر عند معدلات النطاق الحرج بين 14 و20 Bark. بالإضافة إلى ذلك، أظهرت هذه الأبحاث أن حدة الإشارات الصوتية ذات التباين الكبير في مستوى ضغط صوت الإشارة الصوتية والمحتويات القوية عالية التردد لا يمكن أن تعتمد فقط على جهارة الصوت الإجمالية، بل على دالة الترجيح، التي تعتمد على جهارة الصوت الإجمالية.

## 8.3 المعالجة الإدراكية

من الواضح أن جودة الصوت المدركة تتأثر بشدة بالتأثيرات الإدراكية. ويمكن إثبات ذلك من خلال تجربة بسيطة.

تُعالج إشارة مرجعية ذات ضوضاء أساسية مسموعة بوضوح بواسطة بعض المعدات الصوتية غير القادرة على نقل الضوضاء الأساسية هذه. وبما أن الضوضاء عبارة عن تشوه غير مرغوب فيه، فسيتم تصنيف الإشارة المرجعية بشكل أسوأ من الإشارة المعالجة في اختبار الاستماع. ومن ناحية أخرى، فإن نفس الإشارة المعالجة ستكون أسوأ في الدرجة إذا كان الجزء الأكثر أهمية في الإشارة المرجعية هو ضوضاء أساسية رخيمة.

إن إدراج جميع التأثيرات الإدراكية المحتملة يقع خارج نطاق هذا المرفق، ولكن فيما يلي بعض الأمثلة:

مثال 1: فصل التشوهات الخطية عن التشوهات غير الخطية

التشوهات الخطية أقل اعتراضاً من التشوهات غير الخطية. ويمكن فصل التشوهات الخطية عن التشوهات غير الخطية بسهولة إلى حد ما باستخدام الترشيح العكسي التكيفي لإشارة الخرج. وتستخدم الطريقة المحددة في هذه التوصية فصل التشوهات الخطية عنالتشوهات غير الخطية.

مثال 2: تحليل المشهد الصوتي

تحليل المشهد الصوتي [Bregman، 1990] هو عملية إدراكية تسمح للمستمعين بفصل الأحداث الصوتية المختلفة وتجميعها في كائنات مختلفة. وكان النهج العملي، كما هو مذكور في [Beerends وStemerdink، 1994]، مفيداً في التحديد الكمي لتأثير تحليل المشهد الصوتي. فإذا لم يُشفر مكون التردد الزمني بواسطة كوديك، فإن الإشارة المتبقية ستظل تشكل مشهداً صوتياً متماسكاً، في حين يؤدي إدخال مكون تردد زمني جديد غير ذي صلة إلى تصورين مختلفين. وبسبب الانقسام إلى تصورين مختلفين، سيكون التشوه أكثر اعتراضاً مما قد يتوقعه المرء على أساس جهارة عنصر التشوه الذي تم إدخاله حديثاً. ويؤدي هذا إلى عدم تناظر بينالاضطراب المتصور للتشوه الناجم عن عدم تشفير مكون التردد الزمني مقابل الاضطراب الناجم عن إدخال مكون تردد زمني جديد.

مثال 3: التقنيع المعلوماتي

يمكن نمذجة التقنيع المعلوماتي من خلال تحديد مقياس التعقيد الطيفي الزمني الذي يشبه الإنتروبيا. ويعتمد التأثير على الأرجح على مقدار التدريب الذي يتعرض له الأشخاص القائمين بالاختبار قبل إجراء التقييم الشخصي. وترد المحاولة الأولى لنمذجة هذا التأثير في [Beerends *وآخرون،* 1996] حيث تم حساب تقدير التعقيد المحلي خلال نافذة زمنية تبلغ حوالي ms 100. فإذا كان هذا التعقيد المحلي مرتفعاً، فسيكون سماع التشوهات خلال هذه النافذة الزمنية أكثر صعوبة مما لو كان التعقيد المحلي منخفضاً. ويمكن أن يؤدي التدريب إلى تقليل العتبة المقنعة بمقدار عدة عشرات dB [Leek وWatson، 1984].

المثال 4: الترجيح الطيفي الزمني

تحمل بعض المناطق الطيفية الزمنية في الإشارة الصوتية معلومات أكثر، وبالتالي قد تكون أكثر أهمية من غيرها. وقد ثبتت أهمية الترجيح الطيفي الزمني في تقديرات الجودة على كوديكات الكلام. ففي الكلام، من الواضح أن بعض المكونات الطيفية الزمنية، مثل التكوينات، تحمل معلومات أكثر من غيرها [Beerends وStemerdink، 1994]. أما في الموسيقى، فإن جميع المكونات الطيفية الزمنية في الإشارة، حتى فترات الصمت، قد تحمل معلومات.

# 4 النماذج المدمجة

## 1.4 مؤشر التشويش (DIX)

تعتمد طريقة القياس الإدراكي DIX (مؤشر التشويش) [Thiede وKabot، 1996] على بنك المرشحات الصوتية الذي ينتج دقة زمنية عالية وبالتالي يسمح (مقارنة بالنهج المعتمدة على محول FFT) بنمذجة أكثر دقة للتأثيرات الزمنية مثل ما قبل وبعد التقنيع. ويتم الحفاظ على البنية الدقيقة الزمنية للأغلفة عند كل مرشح صوتي ويتم استخدامها للحصول على معلومات إضافية حول الإشارات والتشوهات المدخلة.

تُوزع الترددات المركزية للمرشحات الفردية بالتساوي على سلم درجة الصوت الإدراكي. والجزء العلوي من شكل المرشح مستدير قليلاً لضمان أن العدد المختار من المرشحات يغطي نمدى التردد الكامل دون تموجات في استجابة التردد الإجمالية. ومن أجل نمذجة العتبات المقنعة، تنخفض ميول المرشحات بشكل كبير على سلم Bark. ويعتمد انحدار ميول المرشحات على مستوى إشارات الدخل. وتمت تغطية نطاق التردد المسموع بواسطة 80 مرشحاً في الإصدار الأول من النموذج DIX وتم تخفيضه لاحقاً إلى 40 مرشحاً، أي أن دقة التردد تقابل حوالي Bark 0,6. وتعد خوارزمية بنك المرشحات سريعة إلى حد ما مقارنة ببنوك المرشحات الأخرى ذات المرشحات الفردية، ولكنها لا تزال تستهلك وقتاً أطول بكثير من التحويلات القائمة على المجموعات مثل تحويل FFT وتحويلات حزمة المويجات.

ويقوم النموذج DIX بتكييف المستويات والأطياف دينامياً بين الإشارة قيد الاختبار والإشارة المرجعية من أجل فصل التشوهات الخطية عن التشوهات غير الخطية. ويقوم بتقييم بنية الأغلفة الزمنية عند مخرجات المرشحات من أجل نمذجة الكمية المتزايدة من التقنيع الناتج عن المقنِعات المعدلة والشبيهة بالضوضاء مقارنة بالنغمات النقية.

وبمقارنة التمثيلات الداخلية للإشارة قيد الاختبار والإشارة المرجعية، يتم حساب العديد من معلمات الخرج، بما في ذلك جهارة الصوت الجزئية للتشوهات غير الخطية، ومؤشرات مقدار التشوهات الخطية، ومقاييس التأثيرات الزمنية والخاصة بالأذنين. ومع ذلك، يمكن تحقيق تقدير جيد لجودة الصوت الأساسية باستخدام اثنين فقط من معلمات الخرج: تُقابل جهارة الصوت الجزئية للتشوهات غير الخطية مع أحد المؤشرات الخاصة بكمية التشوهات الخطية لتقدير جودة الصوت الأساسية المتوقعة للإشارة قيد الاختبار.

## 2.4 نسبة الضوضاء إلى الإشارة المقنعة (NMR)

يقوم مخطط القياس NMR (نسبة الضوضاء إلى الإشارة المقنعة) [Brandenburg، 1987] بتقييم فرق المستوى بين العتبة المقنعة وإشارة الضوضاء. ويُستخدم التحويل DFT مع نافذة Hann مقدارها ms 20 لتحليل محتوى تردد الإشارة. وتُدمج معاملات التحويل في نطاقات وفقاً لسلم Bark. وتُقدر العتبة المقنعة لكل نطاق. ويتم اشتقاق ميل العتبة المقنعة باستخدام نهج الحالة الأسوأ مع الأخذ في الاعتبار أن الميول تكون أكثر انحداراً بالنسبة للإشارات الضعيفة ولكنها تصل إلى العتبة المطلقة عند المستويات الأعلى. وتُكيف العتبة المطلقة مع دقة إشارة الدخل (عادةً 16 بتة)، ولكن ليس مع المتطلبات الصوتية النفسية. وبسبب هذه الحقائق يعتبر المخطط NMR قوياً في مواجهة التغيرات في مستوى الاستعادة. وتبلغ دقة سلم درجة الصوت حوالي Bark 1. ونظراً لأن القدرة الحاسوبية المطلوبة منخفضة، فقد كان من الممكن تنفيذ المخطط NMR كنظام في الوقت الفعلي في مرحلة مبكرة من تطويره.

وتم استخدام النموذج منذ عام 1987 وأثبت اعتماديته الأساسية.

وأهم قيم خرج المخطط NMR هي معدل علم التقنيع، الذي يعطي النسبة المئوية للأرتال ذات التشوهات المسموعة، بالإضافة إلى إجمالي ومتوسط نسبة الضوضاء إلى الإشارة المقنعة وهي طرق مختلفة لحساب متوسط المسافة بين طاقة الخطأ والعتبة المقنعة.

## 3.4 التقييم الموضوعي للإشارة الصوتية

يستخدم مخطط القياس OASE (التقييم الموضوعي للإشارة الصوتية) [Sporer، 1997] بنك مرشحات يحتوي على 241 مرشحاً لتحليل إشارات الدخل. والترددات المركزية متباعدة بشكل متساوٍ على سلم بارك بمسافة Bark 0,1. وتتراكب المرشحات مع بعضها البعض. ويُكيف كل مرشح مع استجابة التردد لنقطة على الغشاء القاعدي. ويُضمن اعتماد مستوى الميول من خلال نهج الحالة الأسوأ كما هو الحال في المخطط NMR. وتتطلب المرشحات ذات الترددات المركزية المنخفضة حساباً بمعدل أخذ العينات الكامل بينما يمكن حساب المرشحات ذات الترددات المركزية الأعلى بمعدل أخذ عينات منخفض. وبعد المرشحات، يتم حساب نموذج للتأثيرات الزمنية للجهاز السمعي البشري، كما هو الحال في ASD. وباتباع هذه الخطوة، من الممكن تقليل معدل أخذ العينات في جميع نطاقات المرشحات. ويؤدي هذا إلى دقة زمنية لبنك المرشحات تبلغ ms 0,66 بمعدل أخذ عينات يبلغ kHz 48. ويُقارن خرج مرشحات المواءمة للإشارة المرجعية والإشارة قيد الاختبار مع احتمال دالة الكشف. وتستخدم هذه الدالة جهارة صوت إشارات الدخل كدخل لحساب الفرق JNLD. ويُشتق الاحتمال الإجمالي للكشف من احتمال الكشف لكل نطاق. وتُجرى هذه العملية لكل من قنوات الدخل وأيضاً لما يسمى بالقناة المركزية. إن احتمال الكشف في القناة المركزية لكل نطاق هو أسوأ حالة لاحتمال الكشف للقناة اليسرى واليمنى. ولكل رتل مدته ms 0,66 يتم حساب مجموع الخطوات فوق العتبة أيضاً.

وتُستخدم عدة أساليب لحساب المتوسط الزمني لاحتمال الكشف والخطوات فوق العتبة:

– المتوسط الزمني لاحتمال الكشف؛

– تردد الأرتال مع احتمال كشف فوق 0,5؛

– الحد الأقصى لاحتمال الكشف مع ترشيح بتمرير منخفض؛

– الحد الأقصى لاحتمال الكشف مع ترشيح بتمرير منخفض مع النسيان؛

– متوسط عدد الخطوات فوق العتبة للأرتال التي يزيد احتمال اكشفها عن 0,5؛

– متوسط عدد الخطوات فوق العتبة؛

– الحد الأقصى لعدد الخطوات فوق العتبة؛

– متوسط عدد الخطوات فوق العتبة لنسبة %10 من أسوأ الأرتال.

## 4.4 قياس جودة الصوت الإدراكية (PAQM)

المبدأ الأساسي للمخطط PAQM [Beerends وStemerdink، 1992] هو طرح التمثيلات الداخلية (التمثيلات داخل رأس الشخص القائم بالاختبار) للإشارة المرجعية والإشارة المتدهورة ورسم خريطة للفرق بإجراء تقابل إدراكي لجودة الصوت المدركة شخصياً. ويتم التحول من المجال المادي الخارجي إلى المجال النفسي الجسدي الداخلي عن طريق أربع عمليات:

– إجراء تقابل للتردد الزمني الذي يتم عبر المحول DFT مع نافذة Hann مدتها حوالي ms 40؛

– تعديل التردد باستخدام سلم بارك؛

– نشر التردد الزمني (تحوير غير خطي)؛

– تعديل الشدة (الانضغاط).

يتيح الجمع بين التحوير والانضغاط نمذجة سلوك التقنيع للجهاز السمعي البشري عند العتبة المقنعة وفوقها. ويُستمثل الانضغاط باستخدام النتائج الشخصية لتقييم كوديك الصوت MPEG الأول [تقرير اختبارISO/IEC JTC 1/SC 2/WG 11 MPEG/Audio، الوثيقة MPEG90/N0030، أكتوبر 1990] [تقرير اختبار ISO/IEC JTC 1 /SC 2/WG 11 MPEG/Audio، الوثيقة MPEG91/N0010، يونيو 1991]. يتم التعبير عن الفرق في التمثيل الداخلي باضطراب الضوضاء. وفي أحدث إصدارات PAQM، كما قدمت إلى قطاع الاتصالات الراديوية، تم تضمين تأثيرين إدراكيين في التقابل بدءاً من اضطراب الضوضاء وحتى الجودة الشخصية والتدفق الإدراكي [Beerends وStemerdink، 1994] والتقنيع المعلوماتي [Beerendsوآخرون،1996].

تم تطوير إصدار مبسط من PAQM، [Beerends وStemerdink، 1994] باستخدام نموذج إدراكي كما هو معروض في [Beerends وStemerdink، 1994] ولكن تم تمديده مع ترجيح الفواصل الصامتة. وأثناء تطوير PSQM، اتضح أنه عند الحكم على جودة الكلام في سياق الاتصال الهاتفي، فإن الضوضاء التي تحدث أثناء الفواصل الصامتة تكون أقل أهمية من الضوضاء التي تحدث أثناء فترات الكلام النشطة. وفي معيار مرجعي أجراه قطاع تقييس الاتصالات، أظهر مقترح المخطط PSQM أعلى ارتباط بين الجودة الموضوعية والجودة الشخصية (لجنة الدراسات 12 لطاع تقييس الاتصالات، COM 12 74 – مراجعة اختبارات التحقق من صحة المقاييس الموضوعية لجودة الكلام). وتوصف التوصية ITU-T P.862 التقييم الإدراكي لجودة الكلام (PESQ) حيث تحتوي على خوارزمية محسنة لتقييم جودة الكلام.

## 5.4 التقييم الإدراكي

ينمذج PERCEVAL (التقييم الإدراكي) [Paillardوآخرون،1992] خصائص النقل للأذن الوسطى والداخلية لتشكيل تمثيل داخلي للإشارة. تتحلل إشارة الدخل إلى تمثيل تردد زمني باستخدام محول DFT. وعادةً، يتم تطبيق نافذة هان مقدارها ms 40 على بيانات الدخل، مع تراكب بنسبة %50 بين النوافذ المتعاقبة. ويتم ضرب طيف الطاقة في دالة تعتمد على التردد تنمذج تأثير قناة الأذن والأذن الوسطى. وتُقابل قيم الطاقة الطيفية الموهنة من سلم التردد إلى سلم درجة الصوت الأكثر خطية فيما يتعلق بكل من الخواص الفيزيائية للأذن الداخلية والتأثيرات الجسدية النفسية المرصودة. وتُحور بعد ذلك مكونات الطاقة المحولة بدالة انتشار لمحاكاة تشتت الطاقة على طول الغشاء القاعدي. وأخيراً، تُضاف طاقة متأصلة تعتمد على التردد إلى كل مكون من مكونات درجة الصوت لمراعاة العتبة المطلقة للسمع. ويؤدي تحويل الطاقة إلى ديسيبل إلى تمثيل الغشاء القاعدي للإشارة.

عند محاكاة تجارب التقنيع السمعي، يتم تشكيل تمثيل غشائي قاعدي لكل محفز، والفرق بين التمثيلات هو المعلومات المتاحة لأداء المهمة. وأحد التمثيلين هو للمقنِع وحده، والآخر هو للمقنِع وإشارة الاختبار مجتمعين. يمثل الفرق بينهما مكون الإشارة غير المقنَع. ويحسب المخطط PERCEVAL احتمال كشف هذا الفرق. ويُقدر احتمال عدم كشف الفرق لكل كاشف على طول الغشاء القاعدي المحاكى باستخدام دالة الاحتمال السينية. ومع افتراض أن الكاشفات مستقلة إحصائياً، يُحسب احتمال الكشف العام لمجموعة الكاشفات بأكملها باعتبارها مكملاً لناتج احتمالات عدم الكشف الفردية. وتمت محاكاة العديد من تجارب التقنيع بنجاح باستخدام هذا النهج، وتم استخدام النموذج لتقييم جدوى نمذجة المستمعين الفرديين [Treurniet، 1996].

وكأداة لتقدير جودة الصوت، يحسب المخطط PERCEVAL الفرق بين تمثيلي الإشارة المرجعية والإشارة قيد الاختبار. ومن خلال تطبيق افتراضات معقولة حول العمليات الإدراكية والمعرفية ذات المستوى الأعلى، يتم حساب عدد من المتغيرات ذات الصلة بالإدراك الحسي ومقابلتها لقياس الجودة الموضوعية للإشارة قيد الاختبار. وتم استمثال التقابل من خلال تدنية الفرق بين توزيع الجودة الموضوعية والتوزيع المقابل لمتوسط تقييمات الجودة الشخصية لمجموعة البيانات المتاحة.

## 6.4 القياس الموضوعي الإدراكي

الغرض من القياس الموضوعي الإدراكي (POM) [Colomesوآخرون،1995] هو تحديد مقدار معين من التدهور الذي قد يحدث بين الإشارة المرجعية ونسختها "المتدهورة". ويتحقق ذلك من خلال مقارنة التمثيل القاعدي الداخلي لكلتا الإشارتين، مهما كان التدهور الناتج عنه. وينمذج التمثيل القاعدي العمليات المختلفة التي تخضع لها الإشارة الصوتية عند انتقالها عبر الأذن البشرية. لذلك، فإن المرحلة الأولى من المخطط POM هي حساب التمثيل الداخلي للإشارة الصوتية. وتم اختيار نمط الإثارة (المعطى بالديسيبل)، المنتشر عبر الغشاء القاعدي، لنمذجة معدل الإطلاق في الخلايا العصبية على طول الغشاء القاعدي.

تسمى عملية حساب نمط الإثارة بالأذن الاصطناعية. وبعد ذلك، بمجرد وجود تمثيلين داخليين للإشارتين المراد مقارنتهما ببعضهما البعض، يجب على المخطط POM أن يشير إلى ما إذا كان الفرق بين تمثيلهما الداخلي مسموعاً أم لا، وإذا كان الأمر كذلك فبأي طريقة. وهذا ما يسمى عملية الكشف.

يستخدم المخطط POM محول DFT بنافذة هان قدرها ms 40 تقريباً (مع تراكب بنسبة %50 بين نافذتي هان). وعدد القنوات القاعدية للتحليل هو 620. الأجزاء المتبقية من النموذج السمعي تكاد تكون مطابقة لتلك المستخدمة في كل من المخططين PAQM وPERCEVAL.

ويتم وصف دالة الانتشار بدقة تامة من خلال تقريب أكثر دقة يأخذ في الاعتبار كلا من تبعية المستوى وفقاً لمؤلف [Terhardt، 1979] والشكل المستدير وفقاً لمؤلف [Schroederوآخرون،1979].

وخرج هذا النموذج هو احتمال كشف التشوه بين الإشارتين المقارنتين، بالإضافة إلى ما يسمى بالمسافة القاعدية التي تمثل الفجوة الإدراكية بين الإثارتين المقارنتين.

## 7.4 نهج مجموعة الأدوات

يستخدم مخطط مجموعة الأدوات (Toolbox) نهجاً من ثلاث خطوات لقياس المسافة المدركة في جودة الصوت لإشارة اختبار الصوت فيما يتعلق بالإشارة المرجعية الصوتية، مما يعطي مؤشراً لمستوى جودة الصوت الشخصية الإجمالية لإشارة الاختبار. وتعتمد الطريقة على نماذج إدراكية معروفة جيداً تُستخدم لوصف التمثيل الإدراكي للفروق بين الإشارتين الصوتيتين. علاوة على ذلك، فهي تتضمن إجراء ترجيح لجودة الصوت المدركة لإشارة اختبار مجسمة، مع الأخذ في الاعتبار نتائج كل من القناتين اليسرى واليمنى. ولا يلزم وجود ارتباط صارم على أساس كل عينة على حدة للإشارة المرجعية والإشارة الصوتية قيد الاختبار.

تعتمد الوظيفة الرئيسية لمخطط مجموعة الأدوات، الخطوة 1، على حساب جهارة الصوت المحددة، المحسوبة وفقاً لمؤلف [Zwicker وFeldtkeller، 1967]، باستخدام تحويل FFT قدره 2 048 نقطة، محاط بنافذة Hann، تقابل مدة قدرها حوالي ms 40. وتُزاح النافذة بأكملها بزيادات قدرها ms 10. بالإضافة إلى ذلك، يتم تطبيق تأثيرات التقنيع الزمني، مثل التقنيع المسبق واللاحق، وفقاً لزويكر (Zwicker). ومن هذه القيم الأساسية للإحساس، تُحسب المعلمات الإدراكية الأخرى، مثل جهارة الصوت المدمجة، وجهارة الصوت المقنعة جزئياً، والحدة، وفقاً لـمفهوم [von Bismarck، 1974] و[Aures، 1984]، وكمية الأصداء المسبقة كنتيجة لـمرحلة ما قبل المعالجة للخطوات اللاحقة.

وتتضمن الخطوة الثانية من مخطط مجموعة الأدوات إجراءات الترجيح التي تعتمد بشكل أساسي على مقدار الفرق المدرك في جهارة الصوت وتغير جهارة الصوت بمرور الوقت.

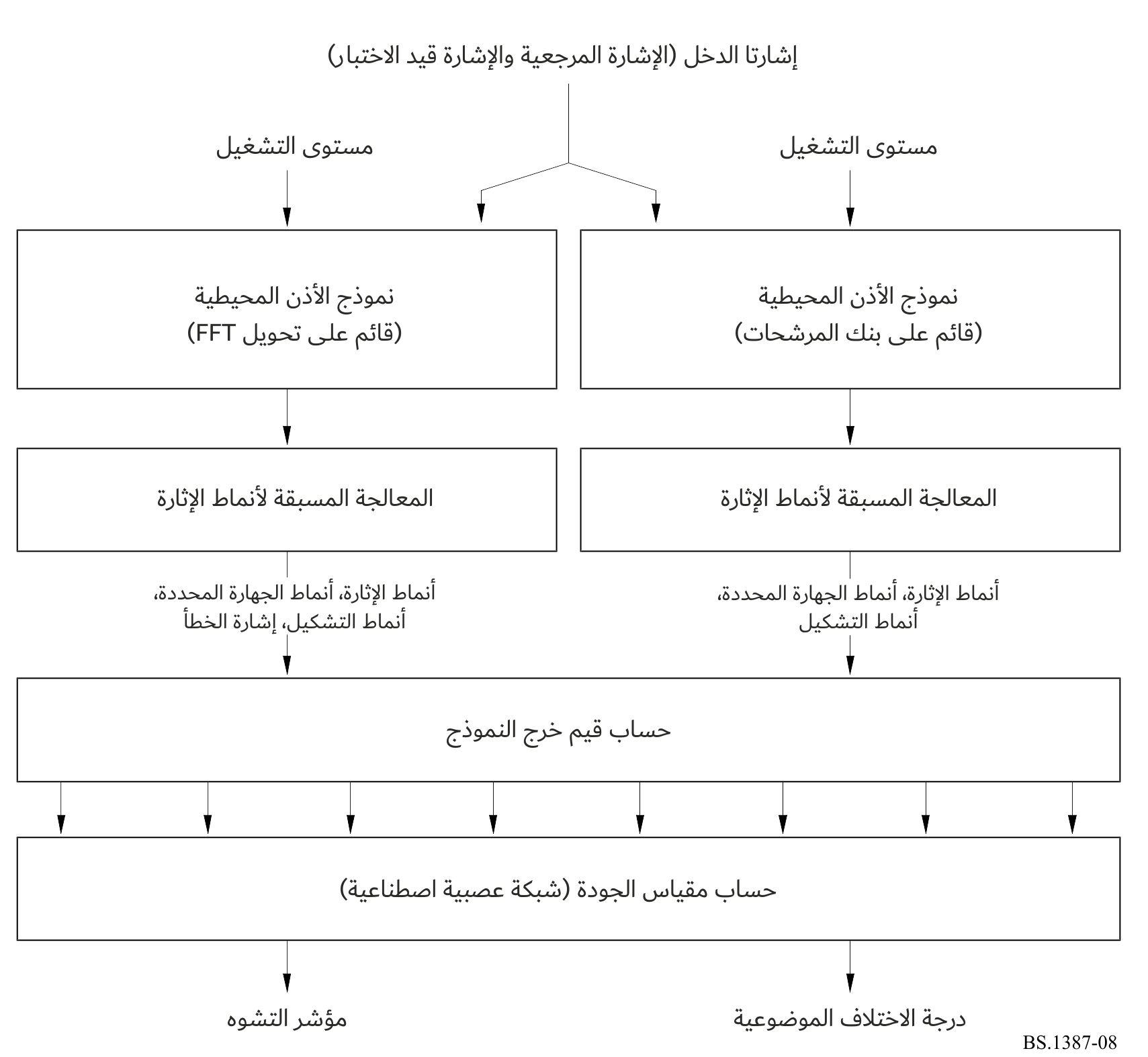
وتتضمن الخطوة الثالثة من مخطط مجموعة الأدوات إنشاء مجموعة من القيم المتوسطة لمجموعة الأدوات تعتمد على تحليل إحصائي للقيم التي تم الحصول عليها في الخطوتين 1 و2. ويتضمن خرج هذا التحليل الإحصائي القيم المتوسطة والقصوى وجذر متوسط التربيع وكذلك الانحراف المعياري للقيم المتوسطة. ويستخدم مجموع مرجح لقيم خرج مجموعة الأدوات المتوسطة هذه للضبط النهائي للمسافة الإدراكية بين الإشارة قيد الاختبار والإشارة المرجعية. وإذا لزم الأمر، يمكن مطابقة قيمة الخرج الفردية هذه مع درجة الاختلاف الشخصية، والتي يتم الحصول عليها عادةً في اختبارات الاستماع الشخصية عن طريق ضبط بيانات الخرج بشكل تسلسلي لكل زيادة في الوقت باستخدام إما دالة متعددة الحدود خطية أو ذات ترتيب أعلى.

الملحق 2  
  
وصف النموذج[[2]](#footnote-2)

# 1 الخطوط العريضة

الشكل 8

مخطط وظيفي عام لنظام القياس



تتكون *طريقة القياس الموضوعي لجودة الصوت المدركة* المقترحة من *نموذج الأذن المحيطية*، والعديد من الخطوات المتوسطة (يشار إليها هنا باسم " *المعالجة المسبقة لأنماط الإثارة*")، وحساب (في الغالب) *متغيرات خرج النموذج* ("MOV") المستندة إلى الصوت النفسي وإجراء التقابل من مجموعة من متغيرات خرج النموذج إلى قيمة واحدة تمثل *جودة الصوت الأساسية* للإشارة قيد الاختبار. وتتضمن نموذجين للأذن المحيطية، أحدهما قائم على التحويل FFT والآخر قائم على بنك المرشحات. وباستثناء حساب إشارة الخطأ (التي تُستخدم فقط مع الجزء القائم على التحويل FFT من نموذج الأذن)، فإن البنية العامة هي نفسها لكلا نموذجي الأذن المحيطية.

وفيما يلي قيم الدخل اللازمة لحساب المتغيرات MOV:

*–* أنماط الإثارة لكل من إشارة الاختبار والإشارة المرجعية.

– أنماط الإثارة المكيفة طيفياً لكل من إشارة الاختبار والإشارة المرجعية.

– أنماط جهارة الصوت المحددة لكل من إشارة الاختبار والإشارة المرجعية.

– أنماط التشكيل لكل من إشارة الاختبار والإشارة المرجعية.

– تُحسب إشارة الخطأ على أنها الفرق الطيفي بين إشارة الاختبار والإشارة المرجعية (لنموذج الأذن المعتمد على التحويل FFT فقط).

إذا لم يُذكر خلاف ذلك، ففي حالة الإشارات المجسمة، تُجرى جميع الحسابات بشكل مستقل وبنفس الطريقة للقناتين اليسرى واليمنى.

ويحدد الوصف إعدادين، أحدهما يسمى " *الصيغة الأساسية*" والآخر يسمى " *الصيغة المتقدمة*".

وفي جميع المعادلات المعطاة، المؤشر "Ref." يشير إلى جميع الأنماط المحسوبة من الإشارة المرجعية، والمؤشر "Test" إلى جميع الأنماط المحسوبة من الإشارة قيد الاختبار. ويشير المؤشر "*k*" إلى متغير التردد المنفصل (أي نطاق التردد) والمؤشر "*n*" إلى متغير الوقت المنفصل (أي إما عداد الأرتال أو عداد العينات). وإذا لم يتم تعريف قيم *k* أو *n* بشكل صريح، فيجب إجراء الحسابات لجميع القيم الممكنة للمؤشرين *k* و*n*. يتم شرح جميع الاختصارات الأخرى في مكان ورودها.

وبالنسبة لأسماء متغيرات خرج النموذج، يشير المؤشر "*A*" إلى جميع المتغيرات المحسوبة باستخدام الجزء القائم على بنك المرشحات من نموذج الأذن ويرمز المؤشر "*B*" إلى جميع المتغيرات المحسوبة باستخدام الجزء القائم على التحويل FFT من نموذج الأذن.

## 1.1 الصيغة الأساسية

لا تتضمن *الصيغة الأساسية* إلا المتغيرات MOV التي يتم حسابها من نموذج الأذن القائم على التحويل FFT. لا يتم استخدام الجزء القائم على بنك المرشحات من النموذج. وتستخدم *الصيغة الأساسية* إجمالي 11 متغيراً MOV للتنبؤ *بجودة الصوت الأساسية* المدركة.

## 2.1 الصيغة المتقدمة

تتضمن *الصيغة المتقدمة* المتغيرات MOV التي يتم حسابها من نموذج الأذن القائم على بنك المرشحات بالإضافة إلى المتغيرات MOV التي يتم حسابها من نموذج الأذن القائم على التحويل FFT. يتم حساب أنماط الإثارة المكيفة طيفياً وأنماط التشكيل من الجزء القائم على بنك المرشحات من النموذج فقط. وتستخدم *الصيغة المتقدمة* إجمالي 5 متغيرات MOV للتنبؤ *بجودة الصوت الأساسية* المدركة.

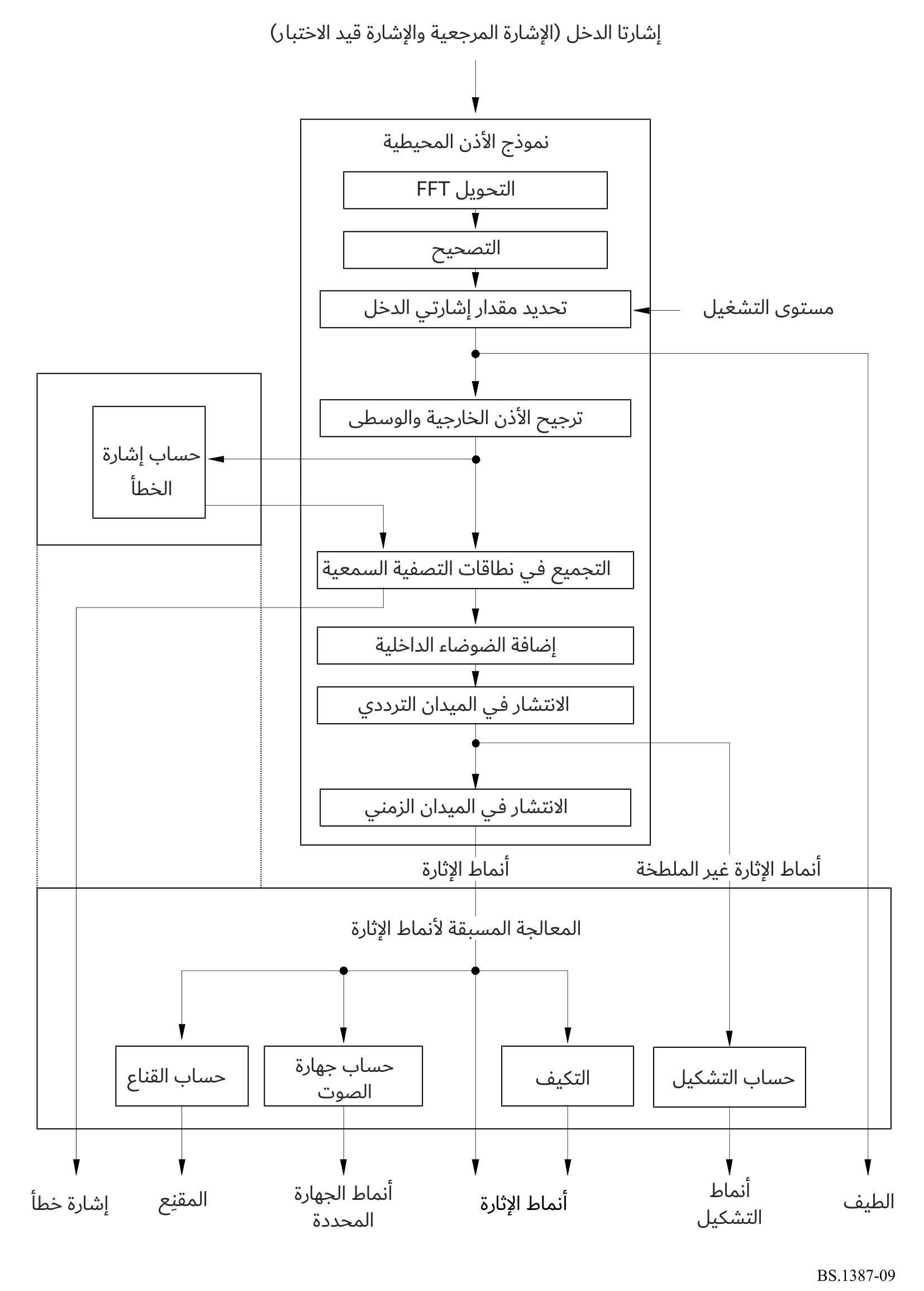
# 2 نموذج الأذن المحيطية

## 1.2 نموذج الأذن القائم على التحويل FFT

### 1.1.2 نظرة عامة

الشكل 9

نموذج الأذن المحيطية والمعالجة المسبقة لأنماط الإثارة لجزء النموذج القائم على التحويل FFT



يتم قطع دخل نموذج الأذن القائم على التحويل FFT، الإشارة المرجعية وإشارة الاختبار المتراصفتين زمنياً بعينات kHz 48، إلى أرتال تبلغ حوالي s 0,042 مع تراكب بنسبة %50. ويحول كل رتل الميدان الترددي باستخدام نافذة Hann ومحول FFT قصير المدى، ويتم تقييسه إلى مستوى التشغيل. وتُطبق دالة الترجيح على المعاملات الطيفية، والتي تنمذج استجابة تردد الأذن الخارجية والوسطى. ويتم التحويل إلى تمثيل درجة الصوت عن طريق تجميع المعاملات الطيفية المرجحة في نطاقات حرجة. وتتم إضافة إزاحة تعتمد على التردد لمحاكاة الضوضاء الداخلية في الجهاز السمعي. وتُستخدم دالة انتشار قائمة على المستوى لنمذجة المرشحات الصوتية الطيفية في الميدان الترددي. وهي تتبع انتشاراً في الميدان الزمني يراعي تأثيرات التقنيع الأمامي.

تُستخدم *أنماط الإثارة* التي تم الحصول عليها الآن لحساب *أنماط جهارة الصوت المحددة* و*أنماط التقنيع*. وتُستخدم الأنماط قبل النشر النهائي في الميدان الزمني ("*أنماط الإثارة غير الملطخة*") لحساب أنماط التشكيل.

ولنمذجة إشارة الخطأ، يتم دمج أنماط الإشارة المرجعية وإشارة الاختبار لخرج مرشح الأذن الخارجية والوسطى وإجراء تقابل لها مع سلم درجة الصوت من خلال التجميع إلى نطاقات حرجة.

وتُستخدم هذه المخرجات مع أنماط الإثارة لحساب قيم متغيرات خرج النموذج.

### 2.1.2 المعالجة الزمنية

يتم قطع دخل نموذج الأذن القائم على التحويل FFT، إشارة الاختبار والإشارة المرجعية، إلى أرتال من 2 048 عينة مع تراكب لعدد 1 024 عينة:

 (1)

حيث *n* عدد الأرتال الزمنية و*kt*العداد الزمني داخل الرتل.

### 3.1.2 تحويل فورييه السريع (FFT)

يُجرى التقابل من الميدان الزمني للميدان الترددي باستخدام نافذة هان (Hann):

 (2)

 (3)

يتبعها تحويل فورييه قصير الأمد:

 (4)

يُحسب عامل المقايسة للتحويل FFT من مستوى ضغط الصوت المفترض، *Lp*، لموجة جيبية كاملة كالتالي:

 (5)

 (6)

حيث يتم حساب عامل المعايرة *Norm* بأخذ موجة جيبية قدرها Hz 1 019,5 وdB 0 بمقياس كامل كإشارة دخل وحساب القيمة المطلقة القصوى للمعاملات الطيفية على مدى 10 أرتال.

وإذا كان مستوى ضغط الصوت غير معروف، يُوصى بضبط *Lp* على dBSPL 92.

### 4.1.2 الأذن الخارجية والوسطى

تُنمذج الاستجابة الترددية للأذن الخارجية والوسطى بدالة ترجيح قائمة على التردد:

 (7)

حيث:

 (8)

التمثيل الترددي عند الخط k المطبق على خرج المحول FFT (المعادلة (9)).

 (9)

*Fe*[*kf*] يُشار إليها على أنها "*خرج المحول FFT المرجح للأذن الخارجية*".

### 5.1.2 التجميع في نطاقات حرجة

يُحسب سلم درجة الصوت من تقريب يُعطى بمفهوم [Schroederوآخرون، 1979].

 (10)

يُطلق على وحدات درجة الصوت اسم *Bark* (على الرغم من أن هذا السلم لا يمثل بالضبط سلم Bark كما هو محدد بواسطة [Zwicker وFeldtkeller، 1967]).

تتراوح حدود تردد المرشحات من Hz 80 إلى Hz 18 000. يتوافق عرض نطاقات المرشحات وتباعدها مع دقة قدرها *res*=*0,25* Bark للصيغة الأساسية و*res*=*0,5* Bark للصيغة المتقدمة.

ويؤدي هذا إلى عدد نطاقات التردد *Z* = 109 للصيغة الأساسية و*Z* = 55 للصيغة المتقدمة.

الجدول 6

نطاقات تردد نموذج الأذن القائم على التحويل FFT في الصيغة الأساسية

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *k* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 0 | 80 | 91,708 | 103,445 | 23,445 |
| 1 | 103,445 | 115,216 | 127,023 | 23,577 |
| 2 | 127,023 | 138,87 | 150,762 | 23,739 |
| 3 | 150,762 | 162,702 | 174,694 | 23,932 |
| 4 | 174,694 | 186,742 | 198,849 | 24,155 |
| 5 | 198,849 | 211,019 | 223,257 | 24,408 |
| 6 | 223,257 | 235,566 | 247,95 | 24,693 |

الجدول 6 (*تابع*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *K* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 7 | 247,95 | 260,413 | 272,959 | 25,009 |
| 8 | 272,959 | 285,593 | 298,317 | 25,358 |
| 9 | 298,317 | 311,136 | 324,055 | 25,738 |
| 10 | 324,055 | 337,077 | 350,207 | 26,151 |
| 11 | 350,207 | 363,448 | 376,805 | 26,598 |
| 12 | 376,805 | 390,282 | 403,884 | 27,079 |
| 13 | 403,884 | 417,614 | 431,478 | 27,594 |
| 14 | 431,478 | 445,479 | 459,622 | 28,145 |
| 15 | 459,622 | 473,912 | 488,353 | 28,731 |
| 16 | 488,353 | 502,95 | 517,707 | 29,354 |
| 17 | 517,707 | 532,629 | 547,721 | 30,014 |
| 18 | 547,721 | 562,988 | 578,434 | 30,713 |
| 19 | 578,434 | 594,065 | 609,885 | 31,451 |
| 20 | 609,885 | 625,899 | 642,114 | 32,229 |
| 21 | 642,114 | 658,533 | 675,161 | 33,048 |
| 22 | 675,161 | 692,006 | 709,071 | 33,909 |
| 23 | 709,071 | 726,362 | 743,884 | 34,814 |
| 24 | 743,884 | 761,644 | 779,647 | 35,763 |
| 25 | 779,647 | 797,898 | 816,404 | 36,757 |
| 26 | 816,404 | 835,17 | 854,203 | 37,799 |
| 27 | 854,203 | 873,508 | 893,091 | 38,888 |
| 28 | 893,091 | 912,959 | 933,119 | 40,028 |
| 29 | 933,119 | 953,576 | 974,336 | 41,218 |
| 30 | 974,336 | 995,408 | 1016,797 | 42,461 |
| 31 | 1016,797 | 1038,511 | 1060,555 | 43,758 |
| 32 | 1060,555 | 1082,938 | 1105,666 | 45,111 |
| 33 | 1105,666 | 1128,746 | 1152,187 | 46,521 |
| 34 | 1152,187 | 1175,995 | 1200,178 | 47,991 |
| 35 | 1200,178 | 1224,744 | 1249,7 | 49,522 |

الجدول 6 (*تابع*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *K* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 36 | 1249,7 | 1275,055 | 1300,816 | 51,116 |
| 37 | 1300,816 | 1326,992 | 1353,592 | 52,776 |
| 38 | 1353,592 | 1380,623 | 1408,094 | 54,502 |
| 39 | 1408,094 | 1436,014 | 1464,392 | 56,298 |
| 40 | 1464,392 | 1493,237 | 1522,559 | 58,167 |
| 41 | 1522,559 | 1552,366 | 1582,668 | 60,109 |
| 42 | 1582,668 | 1613,474 | 1644,795 | 62,128 |
| 43 | 1644,795 | 1676,641 | 1709,021 | 64,226 |
| 44 | 1709,021 | 1741,946 | 1775,427 | 66,406 |
| 45 | 1775,427 | 1809,474 | 1844,098 | 68,671 |
| 46 | 1844,098 | 1879,31 | 1915,121 | 71,023 |
| 47 | 1915,121 | 1951,543 | 1988,587 | 73,466 |
| 48 | 1988,587 | 2026,266 | 2064,59 | 76,003 |
| 49 | 2064,59 | 2103,573 | 2143,227 | 78,637 |
| 50 | 2143,227 | 2183,564 | 2224,597 | 81,371 |
| 51 | 2224,597 | 2266,34 | 2308,806 | 84,208 |
| 52 | 2308,806 | 2352,008 | 2395,959 | 87,154 |
| 53 | 2395,959 | 2440,675 | 2486,169 | 90,21 |
| 54 | 2486,169 | 2532,456 | 2579,551 | 93,382 |
| 55 | 2579,551 | 2627,468 | 2676,223 | 96,672 |
| 56 | 2676,223 | 2725,832 | 2776,309 | 100,086 |
| 57 | 2776,309 | 2827,672 | 2879,937 | 103,627 |
| 58 | 2879,937 | 2933,12 | 2987,238 | 107,302 |
| 59 | 2987,238 | 3042,309 | 3098,35 | 111,112 |
| 60 | 3098,35 | 3155,379 | 3213,415 | 115,065 |
| 61 | 3213,415 | 3272,475 | 3332,579 | 119,164 |
| 62 | 3332,579 | 3393,745 | 3455,993 | 123,415 |
| 63 | 3455,993 | 3519,344 | 3583,817 | 127,823 |
| 64 | 3583,817 | 3649,432 | 3716,212 | 132,395 |

الجدول 6 (*تابع*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *K* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 65 | 3716,212 | 3784,176 | 3853,348 | 137,136 |
| 66 | 3853,348 | 3923,748 | 3995,399 | 142,051 |
| 67 | 3995,399 | 4068,324 | 4142,547 | 147,148 |
| 68 | 4142,547 | 4218,09 | 4294,979 | 152,432 |
| 69 | 4294,979 | 4373,237 | 4452,89 | 157,911 |
| 70 | 4452,89 | 4533,963 | 4616,482 | 163,592 |
| 71 | 4616,482 | 4700,473 | 4785,962 | 169,48 |
| 72 | 4785,962 | 4872,978 | 4961,548 | 175,585 |
| 73 | 4961,548 | 5051,7 | 5143,463 | 181,915 |
| 74 | 5143,463 | 5236,866 | 5331,939 | 188,476 |
| 75 | 5331,939 | 5428,712 | 5527,217 | 195,278 |
| 76 | 5527,217 | 5627,484 | 5729,545 | 202,329 |
| 77 | 5729,545 | 5833,434 | 5939,183 | 209,637 |
| 78 | 5939,183 | 6046,825 | 6156,396 | 217,214 |
| 79 | 6156,396 | 6267,931 | 6381,463 | 225,067 |
| 80 | 6381,463 | 6497,031 | 6614,671 | 233,208 |
| 81 | 6614,671 | 6734,42 | 6856,316 | 241,646 |
| 82 | 6856,316 | 6980,399 | 7106,708 | 250,392 |
| 83 | 7106,708 | 7235,284 | 7366,166 | 259,458 |
| 84 | 7366,166 | 7499,397 | 7635,02 | 268,854 |
| 85 | 7635,02 | 7773,077 | 7913,614 | 278,594 |
| 86 | 7913,614 | 8056,673 | 8202,302 | 288,688 |
| 87 | 8202,302 | 8350,547 | 8501,454 | 299,152 |
| 88 | 8501,454 | 8655,072 | 8811,45 | 309,996 |
| 89 | 8811,45 | 8970,639 | 9132,688 | 321,237 |
| 90 | 9132,688 | 9297,648 | 9465,574 | 332,887 |
| 91 | 9465,574 | 9636,52 | 9810,536 | 344,962 |
| 92 | 9810,536 | 9987,683 | 10168,013 | 357,477 |
| 93 | 10168,013 | 10351,586 | 10538,46 | 370,447 |

الجدول 6 (*تتمة*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *K* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 94 | 10538,46 | 10728,695 | 10922,351 | 383,891 |
| 95 | 10922,351 | 11119,49 | 11320,175 | 397,824 |
| 96 | 11320,175 | 11524,47 | 11732,438 | 412,264 |
| 97 | 11732,438 | 11944,149 | 12159,67 | 427,231 |
| 98 | 12159,67 | 12379,066 | 12602,412 | 442,742 |
| 99 | 12602,412 | 12829,775 | 13061,229 | 458,817 |
| 100 | 13061,229 | 13296,85 | 13536,71 | 475,48 |
| 101 | 13536,71 | 13780,887 | 14029,458 | 492,748 |
| 102 | 14029,458 | 14282,503 | 14540,103 | 510,645 |
| 103 | 14540,103 | 14802,338 | 15069,295 | 529,192 |
| 104 | 15069,295 | 15341,057 | 15617,71 | 548,415 |
| 105 | 15617,71 | 15899,345 | 16186,049 | 568,339 |
| 106 | 16186,049 | 16477,914 | 16775,035 | 588,986 |
| 107 | 16775,035 | 17077,504 | 17385,42 | 610,385 |
| 108 | 17385,42 | 17690,045 | 18000 | 614,58 |

الجدول 7

نطاقات تردد نموذج الأذن القائم على التحويل FFT في الصيغة المتقدمة

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *k* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 0 | 80 | 103,445 | 127,023 | 47,023 |
| 1 | 127,023 | 150,762 | 174,694 | 47,671 |
| 2 | 174,694 | 198,849 | 223,257 | 48,563 |
| 3 | 223,257 | 247,95 | 272,959 | 49,702 |
| 4 | 272,959 | 298,317 | 324,055 | 51,096 |
| 5 | 324,055 | 350,207 | 376,805 | 52,75 |
| 6 | 376,805 | 403,884 | 431,478 | 54,673 |
| 7 | 431,478 | 459,622 | 488,353 | 56,875 |
| 8 | 488,353 | 517,707 | 547,721 | 59,368 |

الجدول 7 (*تابع*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *k* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 9 | 547,721 | 578,434 | 609,885 | 62,164 |
| 10 | 609,885 | 642,114 | 675,161 | 65,277 |
| 11 | 675,161 | 709,071 | 743,884 | 68,723 |
| 12 | 743,884 | 779,647 | 816,404 | 72,52 |
| 13 | 816,404 | 854,203 | 893,091 | 76,687 |
| 14 | 893,091 | 933,119 | 974,336 | 81,245 |
| 15 | 974,336 | 1016,797 | 1060,555 | 86,219 |
| 16 | 1060,555 | 1105,666 | 1152,187 | 91,632 |
| 17 | 1152,187 | 1200,178 | 1249,7 | 97,513 |
| 18 | 1249,7 | 1300,816 | 1353,592 | 103,892 |
| 19 | 1353,592 | 1408,094 | 1464,392 | 110,801 |
| 20 | 1464,392 | 1522,559 | 1582,668 | 118,275 |
| 21 | 1582,668 | 1644,795 | 1709,021 | 126,354 |
| 22 | 1709,021 | 1775,427 | 1844,098 | 135,077 |
| 23 | 1844,098 | 1915,121 | 1988,587 | 144,489 |
| 24 | 1988,587 | 2064,59 | 2143,227 | 154,64 |
| 25 | 2143,227 | 2224,597 | 2308,806 | 165,579 |
| 26 | 2308,806 | 2395,959 | 2486,169 | 177,364 |
| 27 | 2486,169 | 2579,551 | 2676,223 | 190,054 |
| 28 | 2676,223 | 2776,309 | 2879,937 | 203,713 |
| 29 | 2879,937 | 2987,238 | 3098,35 | 218,414 |
| 30 | 3098,35 | 3213,415 | 3332,579 | 234,229 |
| 31 | 3332,579 | 3455,993 | 3583,817 | 251,238 |
| 32 | 3583,817 | 3716,212 | 3853,348 | 269,531 |
| 33 | 3853,348 | 3995,399 | 4142,547 | 289,199 |
| 34 | 4142,547 | 4294,979 | 4452,89 | 310,343 |
| 35 | 4452,89 | 4616,482 | 4785,962 | 333,072 |
| 36 | 4785,962 | 4961,548 | 5143,463 | 357,5 |
| 37 | 5143,463 | 5331,939 | 5527,217 | 383,754 |

الجدول 7 (*تتمة*)

| المجموعة | التردد الأدنى/Hz | التردد المركزي/Hz | التردد الأعلى/Hz | عرض التردد/Hz |
| --- | --- | --- | --- | --- |
| *k* | *fl[k]* | *fc[k]* | *fu[k]* | *fw[k]* |
| 38 | 5527,217 | 5729,545 | 5939,183 | 411,966 |
| 39 | 5939,183 | 6156,396 | 6381,463 | 442,281 |
| 40 | 6381,463 | 6614,671 | 6856,316 | 474,853 |
| 41 | 6856,316 | 7106,708 | 7366,166 | 509,85 |
| 42 | 7366,166 | 7635,02 | 7913,614 | 547,448 |
| 43 | 7913,614 | 8202,302 | 8501,454 | 587,84 |
| 44 | 8501,454 | 8811,45 | 9132,688 | 631,233 |
| 45 | 9132,688 | 9465,574 | 9810,536 | 677,849 |
| 46 | 9810,536 | 10168,013 | 10538,46 | 727,924 |
| 47 | 10538,46 | 10922,351 | 11320,175 | 781,715 |
| 48 | 11320,175 | 11732,438 | 12159,67 | 839,495 |
| 49 | 12159,67 | 12602,412 | 13061,229 | 901,56 |
| 50 | 13061,229 | 13536,71 | 14029,458 | 968,229 |
| 51 | 14029,458 | 14540,103 | 15069,295 | 1039,837 |
| 52 | 15069,295 | 15617,71 | 16186,049 | 1116,754 |
| 53 | 16186,049 | 16775,035 | 17385,42 | 1199,371 |
| 54 | 17385,42 | 17690,045 | 18000 | 614,58 |

يُجرى تقابل التردد إلى درجة الصوت بواسطة الخوارزمية الموضحة في القسم الفرعي التالي، حيث *Fsp[ kf ]* هو تمثيل طاقة "*مخرجات التحويل FFT المرجحة للأذن الخارجية":*

 (11)

أو تمثيل طاقة إشارة الخطأ:

 (12)

على التوالي. أنظر الفقرة 4.3 لحساب إشارة الخطأ.

وخرج هذه المرحلة من المعالجة هو طاقات المجموعات الترددية، *Pe*[ *k, n*].

#### 1.5.1.2 شبه شفرة

|  |  |
| --- | --- |
| /\* inputs \*/ |  |
| Fsp[ ] | طاقات الدخل |
| /\* outputs \*/ |  |
| Pe[ ] | : طاقات جرى تقابلها مع درجة الصوت |
| /\* intermediate values \*/ |  |
| i | : مؤشر للمجموعات الترددية |
| k | : مؤشر\ لخطط المحول fft |
| Z | عدد المجموعات الترددية:  109 للصيغة الأساسية  55 للصيغة المتقدمة |
| fl[] | : التردد الأدنى للمجموعة الترددية |
| fu[] | : التردد الأعلى للمجموعة الترددية |
| Fres | : ثابت للدقة الترددية |

Fres = 48000/2048;

for(i=0; i<Z; i++ )

{

Pe[i]=0;

for(k=0;k<1024;k++)

{

/\* line inside frequency group \*/

if( (( k-0.5)\*Fres >= fl[i]) && ((k+0.5)\*Fres <= fu[i]))

{

Pe[i] += Fsp[k];

}

/\* frequency group inside\*/

else if( (( k-0.5)\*Fres < fl[i]) && ((k+0.5)\*Fres > fu[i]))

{

Pe[i] += Fsp[k]\*(fu[i]-fl[i])/Fres;

}

/\* left border \*/

else if( ((k-0.5)\*Fres < fl[i]) && ((k+0.5)\*Fres > fl[i]))

{

Pe[i] += Fsp[k]\*( (k+0.5)\*Fres - fl[i])/Fres;

}

/\* right border

else if( ((k-0.5)\*Fres < fu[i]) && ((k+0.5)\*Fres > fu[i]);

{

Pe[i] += Fsp[k]\*(fu[i]- (k-0.5)\*Fres)/Fres;

}

/\* line outside frequency group \*/

else

{

Pe[i] += 0;

}

}

/\* limit result \*/

Pe[i]=max(Pe[i],0.000000000001);

}

### 6.1.2 إضافة الضوضاء الداخلية

تُضاف إزاحة قائمة على التردد، *PThres،* للطاقات في كل مجموعة ترددية:

 (13)

 (14)

وخرج هذه المرحلة من المعالجة، *Pp[* *k, n]*، يُشار إليه باسم " *أنماط درجة الصوت*".

### 7.1.2 الانتشار

يتم تلطيخ أنماط درجة الصوت *Pp[ k, n]* عبر التردد باستخدام دالة انتشار تعتمد على المستوى. ودالة الانتشار هي دالة أسية ذات وجهين. ويبلغ الميل الأدنى دائماً dB/Bark 27 ويعتمد الميل الأعلى على التردد والطاقة.

وتُحسب قيم الميل وفقا لما يلي:

 (15)

 (16)

مع:



ويُنفذ الانتشار بشكل مستقل لكل مجموعة من مجموعات الترددات *k*:

 (17)

حيث *Eline* تُشتق من:

 (18)

*NormSP*[*k*] تُحسب وفقاً لما يلي:

 (19)

مع:

 (20)

و*res* هي دقة سلم درجة الصوت بوحدات Bark (0,25 للصيغة الأساسية و0,5 للصيغة المتقدمة).

وتُستخدم الأنماط في هذه المرحلة من المعالجة، *E2[k, n]*، لاحقاً لحساب أنماط التشكيل ويشار إليها باسم "أنماط الإثارة غير الملطخة".

### 8.1.2 الانتشار في الميدان الزمني

من أجل نمذجة التقنيع الأمامي، يتم تلطيخ الطاقات الموجودة في كل مجموعة تردد بمرور الوقت بواسطة مرشحات التمرير المنخفض من الدرجة الأولى. وتعتمد الثوابت الزمنية على التردد المركزي لكل مجموعة (كما هو مُعطى في المعادلة (10) والجدول 6) ويتم حسابها وفقاً لما يلي:

 (21)

وتُحسب مرشحات التمرير المنخفض من الدرجة الأولى وفقاً لما يلي:

 (22)

 (23)

حيث يُحسب *a* من الثوابت الزمنية أعلاه بالمعادلة:

 (24)

*n* رقم الرتل الفعلي، و*k* مؤشر المجموعة و*Ef* [*k,* 0] = 0.

ويُشار إلى الأنماط في هذه المرحلة من المعالجة، *E*[*k, n*]، باسم "*أنماط الإثارة*".

### 9.1.2 عتبة التقنيع

تصف عملية التقنيع التأثير الذي تصبح به الإشارة الخافتة ولكن المسموعة بوضوح غير مسموعة عند حدوث إشارة أعلى جهارة. وتُحسب هذه العتبة عن طريق ترجيح أنماط الإثارة بدالة الترجيح *m[k]*.

 (25)

 (26)

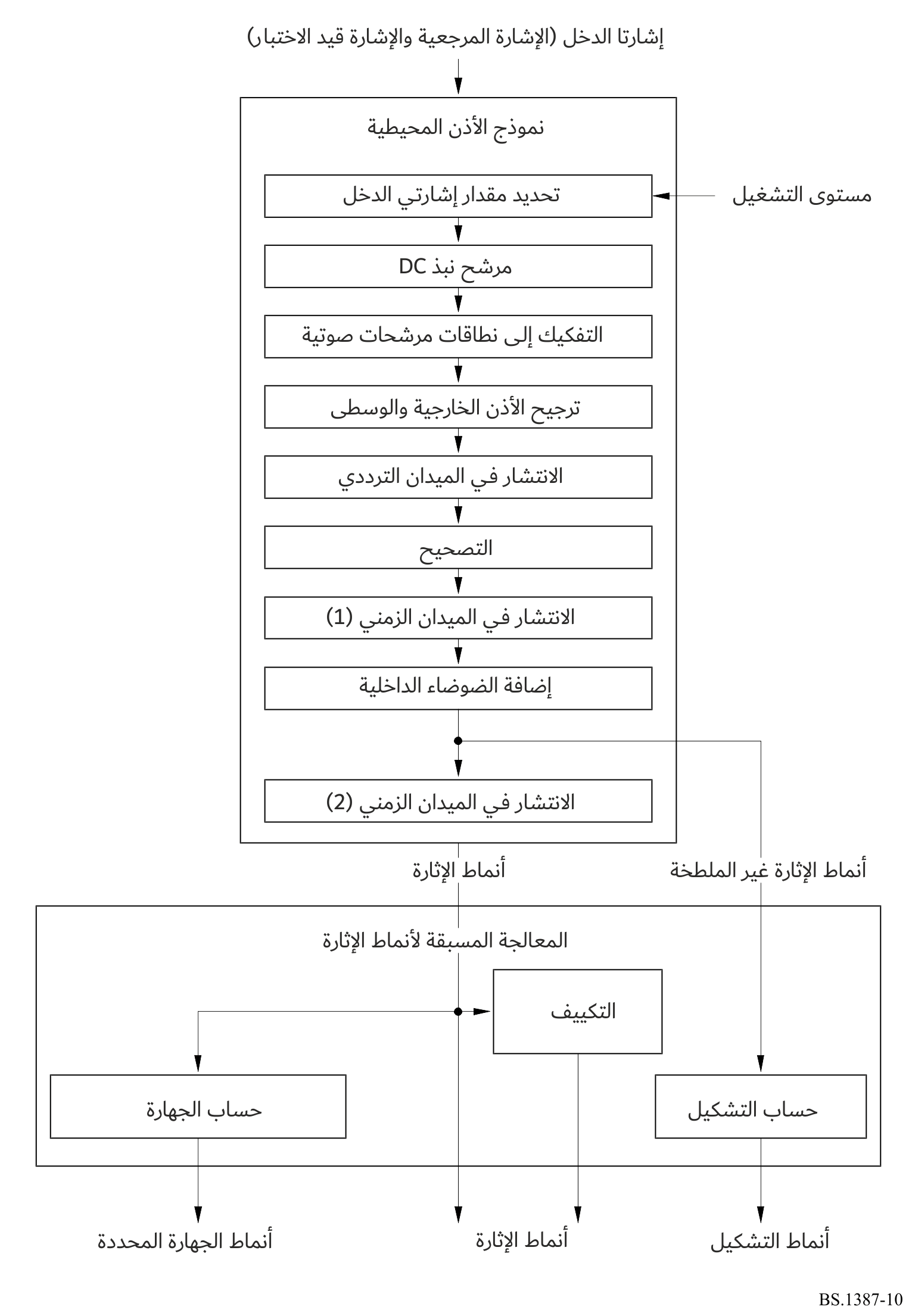
ويُشار إلى الأنماط في هذه المرحلة من المعالجة، *M[k, n]*، باسم "*أنماط القناع*".

## 2.2 نموذج الأذن القائم على بنك المرشحات

### 1.2.2 نظرة عامة

الشكل 10

نموذج الأذن المحيطية والمعالجة المسبقة لأنماط الإثارة لجزء النموذج القائم على بنك المرشحات



عند دخل نموذج الأذن القائم على بنك المرشحات، يتم ضبط الإشارة قيد الاختبار والإشارة المرجعية على مستوى التشغيل المفترض وتُرسل عبر مرشح تمرير عال من أجل إزالة المكونات DC والمكونات دون الصوتية للإشارات. ويتم بعد ذلك تفكيك الإشارات إلى إشارات تمرير النطاق بواسطة مرشحات الطور الخطي التي يتم توزيعها بالتساوي على سلم درجة الصوت الإدراكي. ويُطبق ترجيح يعتمد على التردد على إشارات تمرير النطاق من أجل نمذجة الخصائص الطيفية للأذن الخارجية والوسطى. وتُنمذج الدقة الطيفية المعتمدة على المستوى للمرشحات الصوتية من خلال تحوير الميدان الترددي للمخرجات باستخدام دالة انتشار تعتمد على المستوى.

تُحسب أغلفة الإشارات باستخدام تحويل هيلبرت لإشارات تمرير النطاق ("التصحيح") ويتم تطبيق تحوير الميدان الزمني مع وظيفة نافذة من أجل نمذجة التقنيع العكسي. وبعد ذلك، تُضاف إزاحة تعتمد على التردد تأخذ في الاعتبار الضوضاء الداخلية في الجهاز السمعي وتنمذج العتبة في وضع اللانشاط. وأخيراً، تُنفذ عملية تحوير ثانية للميدان الزمني باستخدام دالة الانتشار الأسية التي تراعي التقنيع الأمامي.

تُستخدم *أنماط الإثارة* التي تم الحصول عليها الآن لحساب *أنماط الجهارة المحددة*، وتُستخدم الأنماط قبل انتشار الميدان الزمني النهائي (" *أنماط الإثارة غير الملطخة*") لحساب *أنماط التشكيل*. وهذه، بالإضافة إلى أنماط الإثارة نفسها، هي الأساس الذي يتم على أساسه حساب قيم النموذج. ومن أجل فصل تأثير استجابة تردد الحالة الثابتة للجهاز قيد الاختبار عن التشوهات الأخرى، يتم أيضاً تكييف أنماط الإثارة للإشارة قيد الاختبار والإشارة المرجعية طيفياً مع بعضها البعض ("التكييف"). ويتم حساب أنماط التشكيل وأنماط الجهارة المحددة من أنماط الإثارة المكيَّفة وغير المكيَّفة.

### 2.2.2 الاعتيان الفرعي

عند خرج بنك المرشحات، يُخفض اعتيان الإشارات بعامل 32 وبعد انتشار الميدان الزمني الأول، يُخفض اعتيان الإشارات بعامل 6 (انظر الشكل 11).

### 3.2.2 ضبط مستوى التشغيل

يُحسب عامل المقايسة للدخل من مستوى التشغيل المفترض لإشارة دخل كاملة النطاق بواسطة:

 (27)

في حالة معرفة نستوى التشغيل الدقيق، يوصى بضبط *Lmax*على dBSPL 92.

### 4.2.2 مرشاح نبذ DC

نظراً لأن بنك المرشحات حساس للإشارات دون الصوتية في إشارات الدخل، يتم تطبيق مرشح نبذ DC على إشارات الدخل. ويستخدم مرشح التمرير العالي من الدرجة الرابعة لبتروورث بتردد قطع قدره Hz 20. ويُنفذ المرشح كتشلشل من مرشحين IIR من الدرجة الثانية.

*yn**xn – 2xn–1**xn–2**b1 yn–1**b2 yn–2* (28)

حيث معاملات المجموعة الأولى هي:

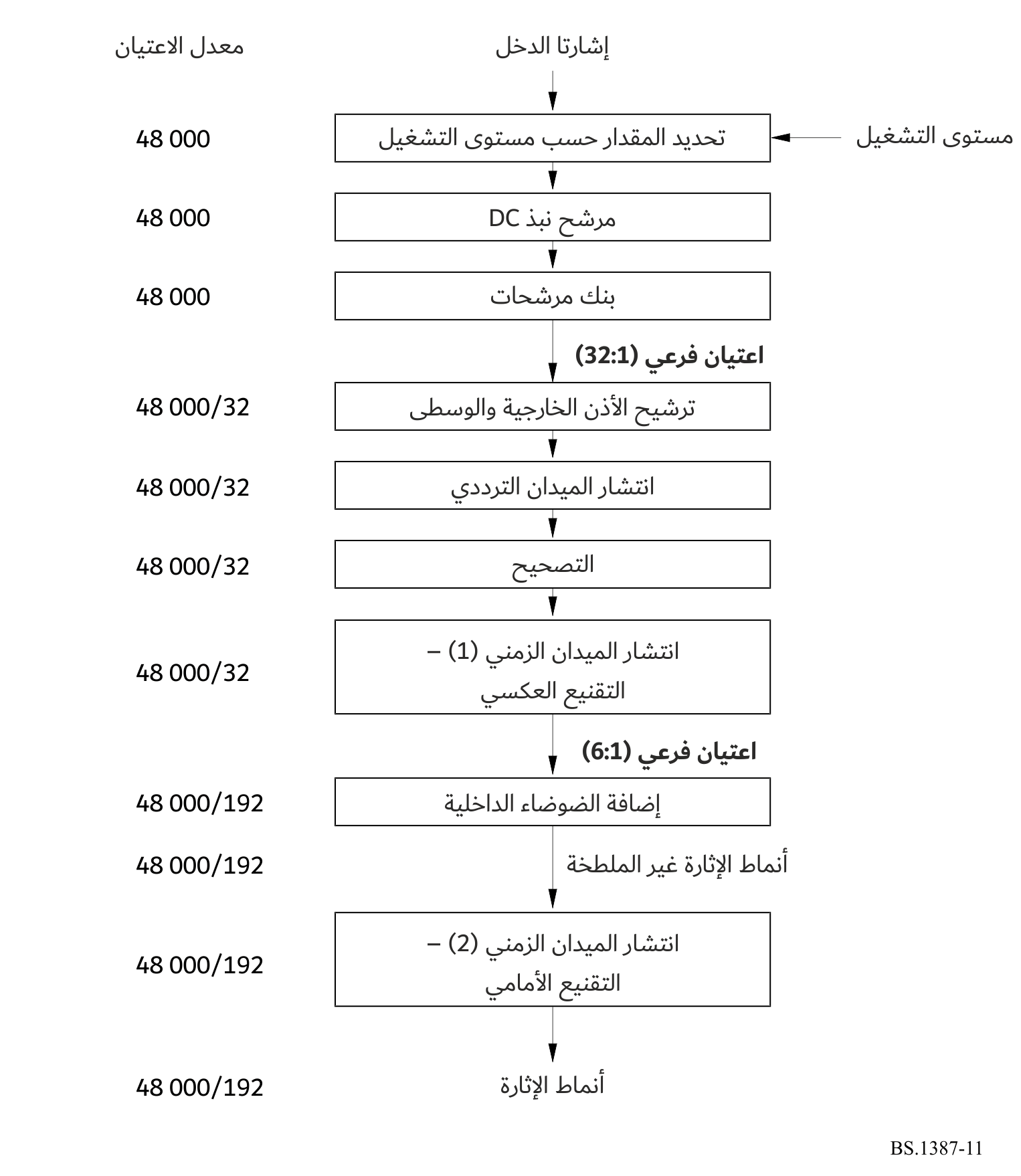
*b1, 2  1.99517, –0.995174*

ومعاملات المجموعة الثانية هي:

*b1, 2  1.99799, –0.997998*

الشكل 11

الاعتيان الفرعي في نموذج الأذن المحيطية القائم على بنك المرشحات



### 5.2.2 بنك المرشحات

يتألف بنك المرشحات من 40 زوجاً من المرشحات لكل قناة من قنوات الإشارة قيد الاختبار والإشارة المرجعية. والمرشحات متباعدة بشكل متساوٍ ولها عرض نطاق مطلق ثابت عندما يتعلق الأمر بسلم درجة الصوت السمعي. ويتكون كل زوج من المرشحات من مرشحين لهما استجابة تردد متساوية ولكن بفارق 90 درجة في استجابة الطور. وبالتالي، فإن خرج المرشح الثاني يمثل تحويل هيلبرت لخرج المرشح الأول (أو الجزء التخيلي، إذا افترض أن المرشح الأول يمثل الجزء الحقيقي من إشارة مركبة). وأغلفة استجاباتها النبضية لها شكل cos2. وتُحدد المرشحات في الجدول 8 والمعادلة (29) (حيث *k* هو مؤشر المرشح، و*n* هو مؤشر العينة الزمنية و*T* هو الوقت بين عينتين: *T = 1/48 000*). ويمكن تحقيقها كمرشحات FIR باستخدام القيم *hre(k, n)* و*him(k, n)* كمعاملات. وعندما تكون إشارات الدخل محدودة الوقت، يمكن أيضاً حساب مخرجات المرشحات بواسطة خوارزميات متكررة سريعة جداً.

 (29)

الجدول 8

التردد المركزي وطول الاستجابة النبضية والتأخير الإضافي لكل مرشح

| مؤشر المرشح | التردد المركزي/Hz | طول الاستجابة النبضية/العينات | التأخير الإضافي/العينات |
| --- | --- | --- | --- |
| *k* | *fc[k]* | *N[k]* | *D[k]* |
| 0 | 50,00 | 1456 | 1 |
| 1 | 116,19 | 1438 | 10 |
| 2 | 183,57 | 1406 | 26 |
| 3 | 252,82 | 1362 | 48 |
| 4 | 324,64 | 1308 | 75 |
| 5 | 399,79 | 1244 | 107 |
| 6 | 479,01 | 1176 | 141 |
| 7 | 563,11 | 1104 | 177 |
| 8 | 652,97 | 1030 | 214 |
| 9 | 749,48 | 956 | 251 |
| 10 | 853,65 | 884 | 287 |
| 11 | 966,52 | 814 | 322 |
| 12 | 1089,25 | 748 | 355 |
| 13 | 1223,10 | 686 | 386 |
| 14 | 1369,43 | 626 | 416 |
| 15 | 1529,73 | 570 | 444 |
| 16 | 1705,64 | 520 | 469 |
| 17 | 1898,95 | 472 | 493 |
| 18 | 2111,64 | 430 | 514 |
| 19 | 2345,88 | 390 | 534 |
| 20 | 2604,05 | 354 | 552 |
| 21 | 2888,79 | 320 | 569 |

الجدول 8 (*تتمة*)

| مؤشر المرشح | التردد المركزي/Hz | طول الاستجابة النبضية/العينات | التأخير الإضافي/العينات |
| --- | --- | --- | --- |
| *k* | *fc[k]* | *N[k]* | *D[k]* |
| 22 | 3203,01 | 290 | 584 |
| 23 | 3549,90 | 262 | 598 |
| 24 | 3933,02 | 238 | 610 |
| 25 | 4356,27 | 214 | 622 |
| 26 | 4823,97 | 194 | 632 |
| 27 | 5340,88 | 176 | 641 |
| 28 | 5912,30 | 158 | 650 |
| 29 | 6544,03 | 144 | 657 |
| 30 | 7242,54 | 130 | 664 |
| 31 | 8014,95 | 118 | 670 |
| 32 | 8869,13 | 106 | 676 |
| 33 | 9813,82 | 96 | 681 |
| 34 | 10858,63 | 86 | 686 |
| 35 | 12014,24 | 78 | 690 |
| 36 | 13292,44 | 70 | 694 |
| 37 | 14706,26 | 64 | 697 |
| 38 | 16270,13 | 58 | 700 |
| 39 | 18000,02 | 52 | 703 |

تتراوح الترددات المركزية من Hz 50 إلى Hz 18 000. ويُحسب سلم درجة الصوت السمعي من تقريب مقدم من [Schroederوآخرون، 1979]:

 (30)

يطلق على وحدات درجة الصوت اسم Bark (على الرغم من أن هذا السلم لا يمثل بالضبط سلم Bark كما حدده [Zwicker وFeldtkeller، 1967]).

من أجل الحصول على تأخيرات متساوية لجميع المرشحات، يتم تأخير دخل كل مرشح بعدد D من العينات، حيث D هو نصف الفرق بين طول الاستجابة النبضية للمرشح وطول الاستجابة النبضية للمرشح ذو الاستجابة النبضية الأطول[[3]](#footnote-3):

 (31)

ويُجرى اعتيان فرعي لخرج المرشحات بعامل مقداره 32، أي أن قيم الخرج تُحسب عند عينة الدخل رقم 32 لجميع المرشحات[[4]](#footnote-4).

### 6.2.2 ترشيح الأذن الخارجية والوسطى

تُنمذج الاستجابة الترددية للأذن الخارجية والوسطى بواسطة دالة ترجيح تعتمد على التردد تُطبق على خرج المرشحات (المعادلة (32)).

 (32)

#### 1.6.2.2 شبه شفرة

|  |  |
| --- | --- |
| /\* inputs \*/ |  |
| out\_re,out\_im | : خرج بنك المرشحات |
| W | : دالة ترجيح (أنظر المعادلة (32)) |
| /\* outputs \*/ |  |
| out\_re,out\_im | : خرج بنك المرشحات |
| /\* intermediate values \*/ |  |
| k | : مؤشر المرشح |
| Wt | : عامل الترجيح |

*/\* outer and middle ear filtering \*/*

for(k=0..39)

{

Wt = pow(10,W[k]/20)

out\_re[k] \*= Wt;

out\_im[k] \*= Wt;

}

### 7.2.2 انتشار الميدان الترددي

تُلطخ قيم خرج بنك المرشحات عبر التردد باستخدام دالة انتشار قائمة على المستوى. ودالة الانتشار هي دالة أسية ذات وجهين. يكون الميل الأدنى دائماً dB/Bark 31 ويتراوح الميل الأعلى بين 24- و4- dB/Bark.

ويُحسب الميل الأعلى، *s[k]*، وفقاً لما يلي:

 (33)

يُحسب المستوى *L[k]* بشكل مستقل لكل قناة مرشح عن طريق أخذ القيمة المطلقة المربعة لخرج المرشح وتحويلها إلى سلم dB. والترددات المركزية، *fc [k]*، مأخوذة من الجدول 8. ويتم تشذيب التمثيلات الخطية للميول مع مرور الوقت بواسطة مرشح تمرير منخفض من الدرجة الأولى له ثابت زمني قدره ms 100.

ويُنفذ الانتشار بشكل مستقل بالنسبة للمرشحات التي تمثل الجزء الحقيقي من الإشارات والمرشحات التي تمثل الأجزاء التخيلية من الإشارات (المعادلة (29)). ويُنفذ الانتشار الترددي أولاً للميل الأعلى (القائم على المستوى) وبعد ذلك للميل الأدنى باستخدام خوارزمية مرشح IIR من الدرجة الأولى.

#### 1.7.2.2 شبه شفرة

|  |  |
| --- | --- |
| /\* inputs \*/ |  |
| out\_re,out\_im | : خرج بنك المرشحات (الأجزاء الحقيقية والتخيلية) |
| z[ ] | : معدلات النطاقات الحرجة للترددات المركزية لنطاقات المرشحات بوحدات Bark (طبقاً للجدول 8 والمعادلة (30)) |
| /\* outputs \*/ |  |
| A\_re,A\_im | : أنماط الخرج |
| /\* intermediate values \*/ |  |
| j,k | : مؤشر المرشح |
| a,b | : معاملات للتشذيب الزمني |
| dist | : مقدار ثابت لحساب اللغط |
| L[ ] | : المستوى عند خرج كل مرشح |
| s[ ] | : الميل المحلي للانتشار لأعلى |
| d1,d2 | : دارئان |
| /\* static \*/ | (قيم محجوزة من الرتل السابق؛ قيم تُستهل بالصفر عند بدء القياس) |
| cl, cu[ ] | : جزء انتشار الإشارة |

*/\* level dependent upward spreading \*/*

dist = pow(0.1,(z[39]-z[0])/(39\*20));

*/\* (z[39]-z[0])/39 is the distance in Bark between two adjacent filter bands \*/*

a = exp(-32/(48000\*0.1));

b = 1 - a;

for(k=0..39)

{

A\_re[k] = out\_re[k];

A\_im[k] = out\_im[k];

}

for(k=0..39)

{

*/\* calculate level dependent slope \*/*

L[k] = 10\*log10(out\_re[k]\* out\_re[k] + out\_im[k]\* out\_im[k];

s[k] = max(4,(24 + 230/fcentre[k] - 0.2\*L[k]));

*/\* calculate spread fraction and smooth it over time\*/*

cu[k] = a\*pow(dist,s[k])+b\*cu[k];

*/\* spreading of band k \*/*

d1 = out\_re[k]

d2 = out\_im[k]

for(j=k+1..39)

{

d1 \*= cu[k];

d2 \*= cu[k];

A\_re[j] += d1;

A\_im[j] += d2;

}

}

*/\* downward spreading \*/*

cl = pow(dist,31);

d1 = 0;

d2 = 0;

for(k=39..0)

{

/\* spreading of band k \*/

d1 = d1 \* cl + A\_re[k];

d2 = d2 \* cl + A\_im[k];

A\_re[k] = d1;

A\_im[k] = d2;

}

### 8.2.2 التصحيح

تُحسب الطاقات عند خرج المرشحات عن طريق إضافة القيم التربيعية للمرشح الذي يمثل الجزء الحقيقي من الإشارة وذلك الذي يمثل الجزء التخيلي من الإشارة.

 (34)

وتُنفذ جميع العمليات التالية على هذه الطاقات.

### 9.2.2 تلطيخ الميدان الزمني (1) – التقنيع العكسي

من أجل نمذجة التقنيع العكسي، تُلطخ الطاقات عند خرج المرشحات بمرور الوقت بواسطة مرشح FIR مع استجابة نبضية على شكل cos2 مع 12 تفريعة (والتي تقابل استجابة مرشح مكونة من 384 عينة عند معدل أخذ عينات دخل بنك المرشحات). وبعد التلطيخ الزمني، يُخفض اعتيان الخرج بعامل مقداره 6. ويتم ضرب القيم الناتجة بعامل معايرة *cal1 = 0,9761* للحصول على مستويات الخرج المناسبة لمستوى التشغيل المحدد.

 (35)

### 10.2.2 إضافة الضوضاء الداخلية

بعد الانتشار الأول للميدان الزمني، تتم إضافة إزاحة *EThres* قائمة على التردد إلى الطاقات في قناة كل مرشح.

 (36)

 (37)

تُستخدم الأنماط في هذه المرحلة من المعالجة، *E2[k, n]*، لاحقاً لحساب أنماط التشكيل ويشار إليها باسم " *أنماط الإثارة غير الملطخة*".

### 11.2.2 تلطيخ الميدان الزمني (2) – التقنيع الأمامي

من أجل نمذجة التقنيع الأمامي، تُلطخ الطاقات الموجودة في قناة كل مرشح بمرور الوقت بواسطة مرشحات التمرير المنخفض من الدرجة الأولى. وتعتمد ثوابت الوقت على التردد المركزي لكل مرشح (كما هو موضح في الجدول 5) ويتم حسابها وفقاً لما يلي:

 (38)

وتُحسب مرشحات التمرير المنخفض من الدرجة الأولى وفقاً لما يلي:

 (39)

حيث يُحسب *a* من الثوابت الزمنية أعلاه بواسطة:

 (40)

ويُشار إلى الأنماط في هذه المرحلة من المعالجة، *E*[*k, n*]، باسم "*أنماط الإثارة*".

# 3 المعالجة المسبقة لأنماط الإثارة

تُستخدم معظم الحسابات الموضحة في هذا القسم مع نموذج الأذن القائم على بنك المرشحات وكذلك مع نموذج الأذن القائم على التحويل FFT. ونظراً لاختلاف عامل الاعتيان الفرعي وعدد نطاقات التردد بين نموذجي الأذن، يتم وصف الثوابت القائمة على هذا العامل باستخدام المتغيرين القائمين على نموذج الأذن *StepSize* و*Z*. فبالنسبة لنموذج الأذن القائم على التحويل FFT، تكون قيمة *StepSize* هي *1* *024* و*Z* إما *55* (الصيغة المتقدمة) أو *109* (الصيغة الأساسية). وبالنسبة لنموذج الأذن القائم على بنك المرشحات، تكون قيمة *StepSize* هي *192* و*Z* هي *40*. وإذا لم يُذكر خلاف ذلك، تتم ضبط كافة المتغيرات والمرشحات التكرارية على الصفر.

## 1.3 تكييف المستوى والنمط

من أجل التعويض عن فروق المستوى والتشوهات الخطية بين إشارة الاختبار والإشارة المرجعية، يتم تكييف متوسط مستويي إشارة الاختبار والإشارة المرجعية مع بعضهما البعض.

وفي الخطوة الأولى، يتم تشذيب الطاقات الموجودة في قناة كل مرشح بواسطة مرشحات التمرير المنخفض من الدرجة الأولى. وتعتمد الثوابت الزمنية على الترددات المركزية للمرشحات ويتم اختيارها على النحو التالي:

 (41)

وتُحسب مرشحات التمرير المنخفض من الدرجة الأولى وفقاً لما يلي:

 (42)

 (43)

حيث *ETest* و*ERef* هما أنماط الإثارة التي سيتم تكييفها مع بعضها البعض ويتم حساب *a* من الثوابت الزمنية بواسطة:

 (44)

### 1.1.3 تكييف المستوى

من أنماط الإدخال ذات التمرير المنخفض ** و**، يُحسب عامل التصحيح اللحظي *LevCorr* بواسطة:

 (45)

فإذا كان عامل التصحيح أكبر من واحد، تتم قسمة الإشارة المرجعية على عامل التصحيح، وإلا يتم ضرب إشارة الاختبار في عامل التصحيح.

 (46)

 (47)

### 2.1.3 تكييف النمط

تُحسب عوامل التصحيح لكل قناة من خلال مقارنة الأغلفة الزمنية لخرج المرشحات لإشارة الاختبار والإشارة المرجعية:

 (48)

تُحسب قيم *a* على النحو الوارد أعلاه (المعادلة (44)) من الثوابت الزمنية المعطاة في المعادلة (41). وإذا كانت *R[k, n]* أكبر من الواحد، يضبط عامل التصحيح لإشارة الاختبار على *R*[*k, n*]*–1* ويضبط عامل التصحيح للإشارة المرجعية على واحد. وفي الحالة المعاكسة، يضبط عامل التصحيح للإشارة المرجعية على *R[k, n]* ويضبط عامل التصحيح لإشارة الاختبار على واحد.

 (49)

إذا كان مقام المعادلة (48) صفراً (وبالتالي ستكون *R[ k, n]* قيمة غير محددة) وكان البسط أكبر من الصفر، فإن *RTest[ k, n]* تُضبط على الصفر وتُضبط *RRef* [*k, n*] على الواحد . وعندما يكون بسط المعادلة (48) صفراً أيضاً، يتم نسخ النسب *RTest*[*k, n*] و *RRef* [*k, n*] من نطاق التردد أدناه. وإذا لم يكن هناك نطاق تردد أدناه (أي *k* = 0) فإن النسب *RTest*[*k, n*] و *RRef* [*k, n*] تُضبط على واحد.

ويُحسب متوسط عوامل التصحيح على عدد *M* من قنوات المرشحات وتُشذب زمنياً (المعادلة (50)) باستخدام نفس الثوابت الزمنية المذكورة أعلاه (المعادلات من (41) إلى (44)). وعرض نافذة التردد *M* هو 3 لنموذج الأذن القائم على بنك المرشحات. وبالنسبة لنموذج الأذن القائم على التحويل FFT، فهو 4 (الصيغة المتقدمة) أو 8 (الصيغة الأساسية) على التوالي.

 (50)



وعند حدود سلم التردد حيث تتجاوز نافذة التردد مدى نطاقات المرشحات، يُخفض عرض نافذة التردد وفقاً لذلك:

*M*1  min(*M*1, *k*), *M*2  min(*M*2, *z*– *k*– 1), *M*  *M*1  *M*2  1 (51)

وتُرجح أنماط الدخل المكيفة المستوى مع عوامل التصحيح المقابلة *PattCorrTest/Ref* [*k, n*]من أجل الحصول على الأنماط المكيفة طيفياً.

 (52)

 (53)

## 2.3 التشكيل

من أنماط الإثارة غير الملطخة، *E2*[*k, n*]، تُحسب جهارة الصوت المبسطة عن طريق رفع الإثارة إلى أس قدره 0,3. وتُلطخ هذه القيمة والقيمة المطلقة لاشتقاقها الزمني زمنياً.

 (54)

 (55)

وتُحسب قيم *a* كما في المعادلة (44) من الثوابت الزمنية المعطاة بواسطة:

 (56)

ومن القيم الناتجة،*و*،يُحسب مقياس تشكيل الغلاف عند خرج كل مرشح:

 (57)

وتُستخدم قيم * فيما بعد لحساب فرق التشكيل.*

## 3.3 الجهارة

تُحسب أنماط الجهارة المحددة للإشارة قيد الاختبار والإشارة المرجعية طبقاً للمعادلة:

 (58)

كما ورد في [Zwicker وFeldtkeller، 1967]، تُحسب جهارة الصوت الإجمالية للإشارة قيد الاختبار والإشارة المرجعية كمجموع عبر جميع قنوات المرشحات لجميع قيم جهارة الصوت المحددة فوق الصفر.

 (59)

يتم اختيار ثابت القياس على أنه *const* = 1.07664 لنموذج الأذن المحيطية القائم على التحويل FFT و*const*= 1.26539 لنموذج الأذن المحيطية القائم على بنك المرشاحات من أجل إعطاء جهارة صوت إجمالية لصوت واحد لنغمة جيبية قدرها dBSPL 40 عند kHz 1. ويُحسب مؤشر العتبة s والإثارة عند العتبة *EThres* وفقاً لما يلي:

 (60)

و

 (61)

على التوالي.

**ملاحظة** - نظراً لاختلاف نماذج الأذن المحيطية، فإن جهارة الصوت المحسوبة هنا لا تتطابق مع جهارة الصوت المحددة في المعيار ISO 532 (الصوتيات - طريقة حساب مستويات جهارة الصوت 1975).

## 4.3 حساب إشارة الخطأ

تُحسب إشارة الخطأ فقط في النموذج القائم على التحويل FFT. ويُحسب في الميدان الترددي بأخذ الفرق بين أطياف القدر المرشّح في الأذن الخارجية والوسطى للإشارة المرجعية وإشارة الاختبار (انظر الفقرة 4.1.2).

 (62)

*Fnoise* تُقابل مع ميدان درجة الصوت باستخدام الخوارزمية الموصوفة في الفقرة 5.1.2.

ويُشار إلى خرج هذه الخوارزمية، *Pnoise*[*n, k*]، باسم " *أنماط الضوضاء*".

# 4 حساب متغيرات خرج النموذج

## 1.4 نظرة عامة

الجدول 9

نظرة عامة على متغيرات خرج النموذج المستخدمة في التنبؤ بجودة الصوت الأساسية

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| متغير خرج النموذج (MOV) | محسوب في نموذج الأذن ... | | يُستخدم في الصيغة ... | |
|  | التحويل FFT | بنك المرشحات | الأساسية | المتقدمة |
| WinModDiff1B | نعم | لا | نعم | لا |
| AvgModDiff1B | نعم | لا | نعم | لا |
| AvgModDiff2B | نعم | لا | نعم | لا |
| RmsModDiffA | لا | نعم | لا | نعم |
| RmsNoiseLoudB | نعم | لا | نعم | لا |
| RmsNoiseLoudAsymA | لا | نعم | لا | نعم |
| AvgLinDistA | لا | نعم | لا | نعم |

الجدول 9 *(تتمة)*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| متغير خرج النموذج (MOV) | محسوب في نموذج الأذن ... | | يُستخدم في الصيغة ... | |
|  | التحويل FFT | بنك المرشحات | الأساسية | المتقدمة |
| BandwidthRefB | نعم | لا | نعم | لا |
| BandwidthTestB | نعم | لا | نعم | لا |
| Total NMRB | نعم | لا | نعم | لا |
| RelDistFramesB | نعم | لا | نعم | لا |
| Segmental NMRB | نعم | لا | لا | نعم |
| MFPDB | نعم | لا | نعم | لا |
| ADBB | نعم | لا | نعم | لا |
| EHSB | نعم | لا | نعم | نعم |

## 2.4 فرق التشكيل

تُقاس الفروق في تشكيل الأغلفة الزمنية للإشارة قيد الاختبار والإشارة المرجعية عن طريق حساب فرق التشكيل المحلي لكل قناة مرشح (المعادلة (63))، حيث يتم اشتقاق *Modtest* و*ModRef* من تطبيق المعادلة (57) على إشارة *Rtest* المرجعية.

 (63)

ويحسب فرق التشكيل اللحظي باعتباره متوسط فروق التشكيل المحلية عبر جميع قنوات المرشحات (المعادلة (64)).

 (64)

تؤخذ عتبة اللانشاط في الاعتبار بواسطة عامل ترجيح قائم على المستوى (المعادلة (65)) محسوب من أنماط الإثارة المعدلة للإشارة المرجعية على النحو الوارد في المعادلة (55) ودالة الضوضاء الداخلية على النحو المحدد في المعادلة (36) لـنموذج الأذن القائم على بنك المرشحات والمعادلة (13) لنموذج الأذن القائم على الترشيح FFT.

 (65)

ويرد وصف التوسيط الزمني لفروق التشكيل اللحظي *ModDiff[n]* باستخدام عوامل الترجيح *TempWt[n]* في الفقرة 2.5. وترد في الجدول 10 قيم الثوابت *negWt* و*offset* و*levWt*.

الجدول 10

متغيرات خرج النموذج التي تقدر فرق التشكيل الإجمالي

| MOV (Xxx=Win/Avg/Rms) | negWt | offset | levWt |
| --- | --- | --- | --- |
| *XxxModDiff1B* | 1 | 1 | 100 |
| *XxxModDiff2B* | 0,1 | 0,01 | 100 |
| *XxxModDiffA* | 1 | 1 | 1 |

### 1.2.4 RmsModDiffA

متغير خرج النموذج *RmsModDiffA* هو المتوسط المربع لفرق التشكيل المحسوب من نموذج الأذن القائم على بنك المرشحات. انظر الفقرة 2.2.5 للاطلاع على التوسيط الزمني والجدول 10 للاطلاع على الثوابت.

### 2.2.4 WinModDiff1B

متغير خرج النموذج *WinModDiff1B* هو متوسط مقسم إلى نوافذ لفرق التشكيل المحسوب من نموذج الأذن القائم على التشكيل FFT. انظر الفقرة 3.2.5 للاطلاع على التوسيط الزمني والجدول 10 للاطلاع على الثوابت. ولا يطبق عامل الترجيح الزمني الوارد في المعادلة (65) على هذا المتغير.

### 3.2.4 AvgModDiff1B وAvgModDiff2B

متغيراً خرج النموذج *AvgModDiff1B وAvgModDiff2B* هما المتوسط الخطي لفرق التشكيل المحسوب من نموذج الأذن القائم على التحويل FFT. والفرق بين المتغيرين *AvgModDiff1B وAvgModDiff2B* هو أن الثوابت يتم اختيارها بشكل مختلف. انظر الفقرة 1.2.5 للاطلاع على التوسيط الزمني والجدول 10 للاطلاع على الثوابت.

## 3.4 جهارة الضوضاء

تقوم متغيرات خرج النموذج هذه بتقدير جهارة الصوت الجزئية للتشوهات الإضافية في وجود الإشارة المرجعية المقنِعة. وصُممت معادلة جهارة الصوت الجزئية (المعادلة (66)) للحصول على جهارة الصوت المحددة للضوضاء وفقاً لـمؤلف [Zwicker وFeldtkeller، 1967] في حالة عدم وجود مقنِع، ولإنتاج شيء مثل النسبة بين الضوضاء والقناع إذا كانت الضوضاء صغيرة جداً مقارنة بالمقنِع.

وتُحسب جهارة الضوضاء الجزئية وفقاً لما يلي:

 (66)

حيث تكون *E*0 دائمًا 1 و *EThres* هي دالة الضوضاء الداخلية *EThres*[*k*] كما هي محددة في المعادلة (36) وتُحسب *s* وفقاً لما يلي:

 (67)

وما لم توصف بشكل مختلف، تُستخدم *أنماط الإثارة المكيفة طيفياً* (انظر الفقرة 1.3) كمدخلات: *ETest*=*EP,Test*[*k, n*] و*ERef*=*EP,Ref* [*k, n*]. ويُحسب المعامل β، الذي يحدد مقدار التقنيع، بواسطة:

 (68)

لا تؤخذ قيم جهارة الضوضاء اللحظية في الاعتبار إلا بعد ms 50 من تجاوز جهارة الصوت الإجمالية للقناة السمعية اليسرى أو اليمنى مرة واحدة قيمة *NT NThres*=*0,1* *sone* لكل من إشارة الاختبار والإشارة المرجعية (انظر الفقرة 2.4.2.5).

وفي التوسيط الطيفي، تُعاير القيم اللحظية بعدد نطاقات المرشحات لكل نطاق حرج بدلاً من العدد الإجمالي لنطاقات المرشحات، أي يتم ضرب نتيجة التوسيط الطيفي في عامل قيمته 24.

وإذا كان جهارة الضوضاء اللحظية أقل من قيمة العتبة *NLmin*، يتم ضبطها على الصفر.

الجدول 11

متغيرات خرج النموذج التي تقدر جهارة الضوضاء الإجمالية

| MOV (Xxx=Win/Avg/Rms) | α | ThresFac0 | S0 | NLmin |
| --- | --- | --- | --- | --- |
| *XxxMissingComponentsB* | 1,5 | 0,15 | 1 | 0 |
| *XxxNoiseLoudB* | 1,5 | 0,15 | 0,5 | 0 |
| *XxxMissingComponentsA* | 1,5 | 0,15 | 1 | 0 |
| *XxxNoiseLoudA* | 2,5 | 0,3 | 1 | 0,1 |
| *XxxLinDistA* | 1,5 | 0,15 | 1 | 0 |

### 1.3.4 RmsNoiseLoudA

متغير خرج النموذج *RmsNoiseLoudA* هو المتوسط المربع لجهارة الضوضاء المحسوب من نموذج الأذن القائم على بنك المرشحات. انظر الفقرة 2.2.5 للاطلاع على التوسيط الزمني والجدول 11 للاطلاع على الثوابت.

### 2.3.4 RmsMissingComponentsA

متغير خرج النموذج *RmsMissingComponentsA* هو المتوسط المربع لجهارة الضوضاء المحسوب من نموذج الأذن القائم على بنك المرشحات. ويتم حسابه باستخدام أنماط الإثارة المكيفة طيفياً لإشارة الاختبار والإشارة المرجعية المتبادلة من أجل الحصول على جهارة المكونات في الإشارة المرجعية المفقودة في إشارة الاختبار. انظر الفقرة 2.2.5 للاطلاع على التوسيط الزمني والجدول 11 للاطلاع على الثوابت.

### 3.3.4 RmsNoiseLoudAsymA

متغير خرج النموذج *RmsNoiseLoudAsymA* هو المجموع المرجح للمتوسطات المربعة لجهارة الضوضاء (انظر الفقرة 1.3.4) وجهارة مكونات الإشارة المفقودة (انظر الفقرة 2.3.4)، وكلاهما محسوب من نموذج الأذن القائم على بنك المرشحات.

 (69)

### 4.3.4 AvgLinDistA

يقيس متغير خرج النموذج *AvgLinDistA* جهارة مكونات الإشارة المفقودة أثناء التكيف الطيفي للإشارة قيد الاختبار والإشارة المرجعية. ويستخدم الإثارة المكيفة طيفياً للإشارة المرجعية كمرجع والإثارة غير المكيفة للمرجع كإشارة اختبار. ويُحسب هذا المتغير من نموذج الأذن القائم على بنك المرشحات. انظر الفقرة 1.2.5 للاطلاع على التوسيط الزمني والجدول 11 للاطلاع على الثوابت.

### 5.3.4 RmsNoiseLoudB

متغير خرج النموذج *RmsNoiseLoudB* هو المتوسط المربع لجهارة الضوضاء المحسوب من نموذج الأذن القائم على التحويل FFT. انظر الفقرة 2.2.5 للاطلاع على التوسيط الزمني والجدول 11 للاطلاع على الثوابت.

## 4.4 عرض النطاق

تقدر قيم خرج النموذج هذه متوسط عرض النطاق للإشارة قيد الاختبار والإشارة المرجعية في خطوط التحويل FFT.

وبالنسبة لكل رتل، يُحسب عرض النطاق المحلي *BwRef*[*n*] و*BwTest*[*n*] وفقاً لشبه الشفرة أدناه.

### 1.4.4 شبه شفرة

|  |  |
| --- | --- |
| /\* inputs \*/ |  |
| FLevRef[], FlevelTest[] | : مستوى خرج التحويل FFT بوحدات dB |
| /\* outputs \*/ |  |
| BwRef, BwTest | : أنماط الخرج |
| /\* intermediate values \*/ |  |
| k | : مؤشر خطوط التحويل FFT |
| ZeroThreshold | : عتبة عرض النطاق |

ZeroThreshold = FLevelTst(921);

BwRef = BwTst = 0.0;

for(k=921;k<1024;k++)

{

ZeroThreshold=max(ZeroThreshold,FLevelTst(k));

}

for (k = 920; k>=0; k--)

{

if (FLevelRef[k] >= 10.0+ZeroThreshold)

{

BwRef = k+1;

break;

}

}

for (k = BwRef-1; k>=0; k--)

{

if(FLeveltest[k] >= 5.0+ZeroThreshold)

{

BwTest=k+1;

break;

}

}

### 2.4.4 BandwidthRefB وBandwidthTestB

المتغير BandwidthRefB هو المتوسط الخطي لـعرض النطاق BwRef، والمتغير BandwidthTestB هو المتوسط الخطي لـعرض النطاق BwTest. بالنسبة للعملية التوسيط، يتم أخذ الأرتال ذات BwRef > 346 فقط بعين الاعتبار. ويتم تجاهل الأرتال ذات الطاقة المنخفضة في بداية ونهاية العناصر (انظر الفقرة 4.4.2.5). انظر الفقرة 1.2.5 للاطلاع على التوسيط الزمني.

## 5.4 نسبة الضوضاء إلى القناع

تُحسب القيم التالية للنموذج من قيم الضوضاء والتقنيع.

النسبة *NMR* المحلية للرتل الحالي *n* تُعرف كالتالي:

 (70)

### 1.5.4 Total NMRB

متغير خرج النموذج *Total NMRB هو المتوسط الخطي للنسبة ضوضاء إلى قناع باستخدام*

 (71)

ويتم تجاهل الأرتال ذات الطاقة المنخفضة في بداية ونهاية العناصر (انظر الفقرة 4.4.2.5).

### 2.5.4 Segmental NMRB

متغير خرج النموذج *Segmental NMRB هو المتوسط الخطي للنسبة ضوضاء إلى قناع المحلية.* انظر الفقرة 1.2.5 للاطلاع على التوسيط الزمني.

ويتم تجاهل الأرتال ذات الطاقة المنخفضة في بداية ونهاية العناصر (انظر الفقرة 4.4.2.5).

## 6.4 Relative Disturbed FramesB

متغير خرج النموذج *Relative* *Disturbed FramesB (الاسم المختصر:* RelDistFramesB) يمثل عدد الأرتال ذات:



بالنسبة لعدد الأرتال الإجمالي في العنصر.

ويتم تجاهل الأرتال ذات الطاقة المنخفضة في بداية ونهاية العناصر (انظر الفقرة 4.4.2.5).

## 7.4 احتمال الكشف

تعتمد المتغيرات MOV المعرفة في هذا القسم على الدالة  (النطاق k، الرتل n)، وهي أنماط الإثارة *E*[*k, n*] المعبر عنها بوحدات dB:

 (72)

بالنسبة لكل رتل *n*:

تُنفذ الخطوات التالية بشكل مستقل لكل قناة *c* (تكون قيم *c* يسار ويمين). و*أنماط الإثارة اللوغاريتمية* هي  للإشارة المرجعية و للإشارة قيد الاختبار على التوالي.

ولكل نطاق *k*:

– حساب متوسط الإثارة غير المتماثلة.

 (73)

– حساب حجم خطوة الكشف الفعال s. المعادلة التالية هي تقريب لفرق المستوى الملحوظ فقط كما تم قياسه بواسطة مفهوم [Zwicker وFastl، 1990].

إذا كانت *L*[*k, n*] > 0:

*s*[*k, n*]    5.95072  ((6.39468)/*L* [*k, n*])1.71332  9.01033  10–11  *L* [*k, n*]4  5.05622   
10–6  *L* [*k, n*]3 – 0.00102438  *L* [*k, n*]2 0.0550197  *L*[*k, n*] –0.198719

وخلاف ذلك

*s* [*k, n*]  1.0  1030 (74)

– حساب الخطأ الجبري *e*:

 (75)

– إذا كانت ، يتم ضبط انحدار الميل *b* على 4,0 وخلاف ذلك ييتم ضبطه على 6,0. وهذا ينمذج التأثير المتمثل في أن زيادة طاقة الإشارة للإشارة قيد الاختبار مقارنة بالإشارة المرجعية تكون أكثر إثارة للإنتباه من النقصان.

– حساب عامل القياس *a*:

 (76)

– حساب احتمال الكشف. وتحدد المعادلة (76) عامل القياس a كما لو كانت *e*[*k, n*] تساوي *s*[*k, n*]، *pc*[*k, n*] تصبح *0,5*:

 (77)

– حساب العدد الإجمالي للخطوات فوق العتبة:

 (78)

– ويكون احتمال الكشف بكلتا الأذنين كالتالي:

*pbin* [ *k, n*]  max( *pleft* [ *k, n*], *pright* [*k, n*]) (79)

– ويكون عدد الخطوات فوق العتبة لقناة الأذنين كالتالي:

*qbin* [ *k, n*]  max(*qleft* [ *k, n*], *qright* [*k, n*]) (80)

ويكون الاحتمال الإجمالي لكشف القناة *c* للرتل *n* كالتالي:

 (81)

حيث قد تأخذ *c* القيمة *left* أو *right* أو *bin. ويكون العدد الإجمالي للخطوات فوق العتبة للقناة c للرتل n كالتالي:*

 (82)

### 1.7.4 الحد الأقصى لاحتمال الكشف المصفى (MFPDB)

تُحسب الصيغة المشذبة لاحتمال الكشف لكل قناة *c*:

 (83)

حيث *Pc*[*–1*] =*0*. ويعتمد المقدار الثابت *c*0 على *حجم الخطوة*:

*c*0  0.9*StepSize*/1 024 (84)

*ويقلل الثابت c*0 من الحساسية للتشوهات بالغة القصر.

ويُحسب الحد الأقصى لاحتمال الكشف المصفى (الاسم المختصر: MFPD):

 (85)

حيث *PMc*[*–1*] تساوي الصفر. ويعتمد المقدار الثابت *c*1 على *حجم الخطوة*:

*c*1  0.99*StepSize*/1 024 (86)

ينمذج الثابت *c*1 التأثير المتمثل في أن التشوهات في بداية مقتطف من الصوت تكون أقل حدة مما تكون عليه في نهاية المقتطف بسبب النسيان. ويُلاحظ أن هذا الثابت مفيد في نمذجة اختبارات الاستماع حيث لا يُسمح للأشخاص القائمين بالاختبار باختيار أجزاء أقصر من المقتطف. وبالنسبة للنموذج الحالي، الذي تتم معايرته باستخدام بيانات من اختبارات الاستماع وفقاً للتوصية ITU‑R BS.1116، ينبغي أن تكون قيمة الثابت *c*1 تساوي 1,0.

ومتغير خرج النموذج MFPD هو *PMbin*[*n*] للرتل الأخير.

### 2.7.4 متوسط المجموعة المشوهة [[5]](#footnote-5) (ADBB)

يُحسب عدد الأرتال الصالحة التي لديها احتمال كشف قناة الأذنين *Pbin*[*n*] فوق 0,5 (*ndistorted*).

وبالنسبة لجميع الأرتال الصالحة، يُحسب العدد الإجمالي للخطوات فوق عتبة قناة الأذنين *Qbin*[*n*]:



ويُحسب تشوه متوسط المجموعة المشوهة ADB:

– if *ndistorted* is zero then *ADB* = 0 (no distortion audible);

– if *ndistorted* > 0 and *Qsum* > 0 then *ADB* = log10 ((*Qsum*) / *ndistorted*));

– if *ndistorted*> 0 and *Qsum* is zero then *ADB* = –0.5.

## 8.4 البنية التوافقية للخطأ

الإشارة المرجعية التي تحتوي على توافقيات قوية (مثل الكلارينيت، هاربسكورد) لها طيف يتميز بعدد من القمم المتباعدة بانتظام والتي تفصلها قيعان عميقة. وفي بعض الظروف، قد ترث إشارة الخطأ تلك البنية. فعلى سبيل المثال، من المرجح أن تظل الضوضاء الممزوجة بمثل هذه الإشارة غير مقنعة حيث تكون الإشارة منخفضة في القيعان الطيفية. وسيحتوي طيف الخطأ الناتج بعد ذلك على بنية مشابهة للطيف الأصلي ولكن بإزاحة في التردد ليتوافق مع مواضع القيعان. وقد تؤدي هذه البنية إلى تشويه بالصفات النغمية التي يمكن أن تزيد من بروز الخطأ.

ويُعرف الخطأ على أنه الفرق في الأطياف اللوغاريتمية للإشارة المرجعية والإشارة المعالجة، كل منها مرجح بالاستجابة الترددية للأذن الخارجية والوسطى (انظر الفقرة 4.1.2، المعادلة (7)). ولا يُستخدم نمط الإثارة من النموذج النفسي الصوتي هنا لأن تحويل التردد غير الخطي إلى Bark من شأنه تلطيخ البنية التوافقية.

### 1.8.4 EHSB

يُتحصل على مقدار البنية التوافقية من خلال تحديد وقياس أكبر قمة في طيف دالة الترابط الآلي. يتم حساب كل ترابط على أنه جيب تمام الزاوية بين متجهين وفقاً للصيغة التالية، حيث هو متجه الخطأ وهو نفس المتجه بتأخير بمقدار معين. وطول الترابط هو نفس الحد الأقصى للتأخر (أي 256 في المثال أدناه).

 (87)

الحد الأقصى للتأخر في الحصول على دالة الترابط الآلي هو القوة الأكبر للرقم 2 ويكون أصغر من نصف رقم مكون تردد التحويل FFT المقابل للتردد kHz 18.

فعلى سبيل المثال، عند معدل اعتيان قدره kHz 48 وحجم نافذة FFT يبلغ 2 048 عينة، فإن مكون FFT المقابل للتردد kHz 18 هو (18/24) × 1 024 = 768. وبالتالي، سيكون الحد الأقصى للتأخر 384. وسيكون العدد الفعلي للتأخيرات 256، وهي أكبر قوة للرقم 2 أقل من 384. وسيتم الحصول على القيمة الأولى لدالة الترابط عن طريق تراصف *Ft*[0] مع *F*0[0] والقيمة الأخيرة عن طريق تراصف *Ft*[0] مع *F*0[255].

يوضع متجه الترابطات الناتج في نافذة هان معايرة، وبعد إزالة المكون DC عن طريق طرح القيمة المتوسطة، يتم حساب طيف القدرة باستخدام تحويل FFT. وتُحدد الذروة القصوى في الطيف بعد القاع الأول التردد السائد في دالة الترابط الآلي. ومتوسط قيمة هذا الحد الأقصى عبر الأرتال مضروباً في 1 000.0 هو متغير البنية التوافقية للخطأ (EHS).

# 5 التوسيط

## 1.5 التوسيط الطيفي

ما لم يُذكر خلاف ذلك في أوصاف المتغيرات MOV (انظر الفقرة 4)، تُستخدم الخوارزمية التالية عند توسيط القيم المحلية عبر نطاقات التردد.

### 1.1.5 المتوسط الخطي

تُحسب قيمة المتوسط الخطي كالتالي:

 (88)

حيث *S* ترمز إلى اسم متغير خرج النموذج و*Z* عدد نطاقات التردد.

## 2.5 التوسيط الزمني

ما لم يُذكر خلاف ذلك في أوصاف المتغيرات MOV (انظر الفقرة 4)، تُستخدم واحدة أو أكثر من الخوارزميات التالية عند توسيط القيم اللحظية مع الزمن. ويُشار إلى عامل الترجيح الزمني (في حالة تطبيقه) بالرمز *W*، و*Z* هو عدد نطاقات التردد.

### 1.2.5 المتوسط الخطي

تُحسب قيمة المتوسط الخطي (السابقة “*Avg*”) كالتالي:

 (89)

حيث يشير *X* إلى اسم المتغير MOV و*N* هو عدد العينات الزمنية التي تم حساب القيم اللحظية للمتغير *X* لها.

وفي حالة تطبيق الترجيح الزمني (انظر الفقرة 2.4)، يُحسب المتوسط الخطي وفقاً لما يلي:

 (90)

بدلاً من ذلك.

### 2.2.5 متوسط التربيع

تُحسب قيمة متوسط التربيع (السابقة“*Rms*”) كالتالي:

 (91)

حيث يشير *X* إلى اسم المتغير MOV و*N* هو عدد العينات الزمنية التي تم حساب القيم اللحظية للمتغير *X* لها.

وعند تطبيق الترجيح الزمني (انظر الفقرة 2.4)، يُحسب متوسط التربيع وفقاً لما يلي:

 (92)

بدلاً من ذلك.

### 3.2.5 المتوسط المقسم إلى نوافذ

تُحسب قيمة المتوسط المقسم إلى نوافذ (السابقة “*Win*”) كالتالي:

 (93)

حيث يشير *X* إلى اسم المتغير MOV و*N* هو عدد العينات الزمنية التي تم حساب القيم اللحظية للمتغير *X* لها وL طول النافذة الزمنية المتدرجة في العينات الزمنية. ويبلغ طول النافذة ms 100 تقريباً، أي أن L تساوي 4 لنموذج الأذن القائم على التحويل FFT و25 لنموذج الأذن القائم على بنك المرشحات.

### 4.2.5 اختيار الأرتال

#### 1.4.2.5 التوسيط المتأخر

بالنسبة إلى المتغيرات MOV التي تستخدم هذا المعيار، لا تؤخذ القيم المحسوبة خلال أول s 0,5 من القياس في الاعتبار في التوسيط الزمني. ويُستخدم *التوسيط المتأخر* لمتغيرات خرج النموذج التالية:

WinModDiff1، AvgModDiff1، AvgModDiff2، RmsNoiseLoudness، RmsNoiseLoudAsym، RmsModDiff، AvgLinDist.

#### 2.4.2.5 عتبة الجهارة

بالنسبة إلى المتغيرات MOV التي تستخدم هذا المعيار، فإن جميع القيم اللحظية المحسوبة حتى ms 50 بعد أن تصل الجهارة الإجمالية لإحدى القنوات الصوتية المقابلة إلى قيمة *NThres* sone لكل من إشارة الاختبار والإشارة المرجعية لا تؤخذ في الاعتبار في التوسيط الزمني. وتُستخدم *عتبة الجهارة* فقط من أجل المتغيرات MOV الموصوفة في الفقرة 3.4.

#### 3.4.2.5 عتبة الطاقة

عندما تكون طاقة النصف الأحدث من الرتل المؤلف من 2 048 عينة أقل من [[6]](#footnote-6)\*8 000، في القناة الأحادية، أو في كل من القناتين اليسرى واليمنى لكل من البيانات المرجعية وبيانات الاختبار، يتم تجاهل الرتل. وتتسم الأرتال بتراكب بنسبة %50 ويتم تقييم نصف الأرتال التي تحتوي على بيانات جديدة فقط. ويحول تطبيق هذا المعيار دون معالجة الأرتال ذات الطاقة القليلة جداً.

ويستخدم هذا المعيار فقط من أجل المتغير MOV الموصوف في الفقرة 8.4.

#### 4.4.2.5 حدود البيانات

إذا كان الملف المعالج يحتوي على ضوضاء قبل أو بعد بيانات الملف المرجعي الشرعي، فقد يكون الخطأ النسبي كبيراً جداً نظراً لأن المستوى المرجعي يقترب من -∞. وعندما يعتبر هذا الخطأ نقيصة، فقد يتم تجاهله من خلال تطبيق معيار رفض حدود البيانات.

وعند فتح الملفات لأول مرة، يتم تحديد مواقع بداية ونهاية البيانات الفعلية في الملف المرجعي. وتُحدد بداية البيانات أو نهايتها على أنها الموقع الأول، الذي يتم المسح فيه من بداية الملف أو نهايته، حيث يتجاوز مجموع القيم المطلقة على مدى خمس عينات متتالية 200، في إحدى القنوات الصوتية المقابلة. ومن ثم تُغفل تجاهل الأرتال التي تقع خارج هذا النطاق بالكامل.

ويُستخدم هذا المعيار لـلمتغيرات MOV.

## 3.5 التوسيط عبر القنوات الصوتية

ما لم يُذكر خلاف ذلك، في حالة الإشارات المجسمة، يتم توسيط المتغيرات MOV للقناة اليسرى واليمنى خطياً بعد التوسيط الزمني.

# 6 تقدير جودة الصوت الأساسية المدركة

تُقدر *جودة الصوت الأساسية المدركة* من خلال إجراء تقابل للعديد من المتغيرات MOV مع رقم واحد باستخدام بنية شبكة عصبية اصطناعية ذات طبقة مخفية واحدة.

## 1.6 الشبكة العصبية الاصطناعية

وظيفة التنشيط للشبكة العصبية هي دالة سينية غير متماثلة:

 (94)

تستخدم الشبكة مدخلات *I* وعقد *J* في الطبقة المخفية. ويُحدد التقابل من خلال مجموعة من عوامل قياس المدخلات *amin*[*i*] و*amax*[*i*] ومجموعة من أوزان المدخلات *wx*[*i*] ومجموعة من أوزان المخرجات *wy*[*j*] وزوج من عوامل قياس المخرجات *bmin* و*bmax*. ويُجرى تقابل للمدخلات مع *مؤشر التشوه*

 (95)

والذي يرتبط مباشرة *بجودة الصوت الأساسية المدركة* المقدرة من حيث درجة الاختلاف الموضوعي (ODG). ويُعبر عن العلاقة بين مؤشر التشوه ودرجة الاختلاف الموضوعي من خلال:

 (96)

## 2.6 الصيغة الأساسية

تستخدم الصيغة الأساسية نموذج الأذن القائم على التحويل FFT فقط. وتستخدم المتغيرات MOV التالية: BandwidthRefB وBandwidthTestB وTotal NMRB وWinModDiff1B وADBB وEHSB وAvgModDiff1B وAvgModDiff2B وRmsNoiseLoudB وmfpdb وRelDistFramesB. ويُجرى تقابل لهذه المتغيرات MOV البالغ عددها 11 متغيراً مع مؤشر جودة وحيد يستخدم شبكة عصبية على النحو الموصوف في الفقرة 1.6 (*الشبكة العصبية الاصطناعية*) التي تضم ثلاث عقد في الطبقة المخفية. وترد في الجداول من 12 إلى 16 معلمات التقابل.

الجدول 12

المتغيرات MOV المستخدمة في الصيغة الأساسية

| المتغير MOV | الغرض |
| --- | --- |
| WinModDiff1B |  |
| AvgModDiff1B | تغييرات في التشكيل (تتعلق بالاضطراب) |
| AvgModDiff2B |  |
| RmsNoiseLoudB | جهارة التشوه |
| BandwidthRefB | التشوهات الخطية (استجابة ترددية، إلخ) |
| BandwidthTestB |  |
| RelDistFramesB | تردد التشوهات المسموعة |
| Total NMRB | النسبة ضوضاء إلى قناع |
| MFPDB | احتمال الكشف |
| ADBB |  |
| EHSB | البنية التوافقية للخطأ |

الجدول 13

عوامل القياس لمدخلات الصيغة الأساسية

| index (*i*) | MOV (x[i]) | amin[i] | amax[i] |
| --- | --- | --- | --- |
| 0 | BandwidthRefB | 393,916656 | 921 |
| 1 | BandwidthTestB | 361,965332 | 881,131226 |
| 2 | Total NMRB | -24,045116 | 16,212030 |
| 3 | WinModDiff1B | 1,110661 | 107,137772 |
| 4 | ADBB | -0,206623 | 2,886017 |
| 5 | EHSB | 0,074318 | 13,933351 |
| 6 | AvgModDiff1B | 1,113683 | 63,257874 |
| 7 | AvgModDiff2B | 0,950345 | 1145,018555 |
| 8 | RmsNoiseLoudB | 0,029985 | 14,819740 |
| 9 | MFPDB | 0,000101 | 1 |
| 10 | RelDistFramesB | 0 | 1 |

الجدول 14

أوزان عقد الدخل للصيغة الأساسية

| index (*i*) | MOV (x[i]) | node 1 (wx[i,0]) | node 2 (wx[i,1]) | node 3 (wx[i,2]) |
| --- | --- | --- | --- | --- |
| 0 | BandwidthRefB | -0,502657 | 0,436333 | 1,219602 |
| 1 | BandwidthTestB | 4,307481 | 3,246017 | 1,123743 |
| 2 | Total NMRB | 4,984241 | -2,211189 | -0,192096 |
| 3 | WinModDiff1B | 0,051056 | -1,762424 | 4,331315 |
| 4 | ADBB | 2,321580 | 1,789971 | -0,754560 |
| 5 | EHSB | -5,303901 | -3,452257 | -10,814982 |
| 6 | AvgModDiff1B | 2,730991 | -6,111805 | 1,519223 |
| 7 | AvgModDiff2B | 0,624950 | -1,331523 | -5,955151 |
| 8 | RmsNoiseLoudB | 3,102889 | 0,871260 | -5,922878 |
| 9 | MFPDB | -1,051468 | -0,939882 | -0,142913 |
| 10 | RelDistFramesB | -1,804679 | -0,503610 | -0,620456 |
| 11 | bias | -2,518254 | 0,654841 | -2,207228 |

الجدول 15

أوزان عقد الخرج للصيغة الأساسية

| node 1 (wy[0]) | node 2 (wy[1]) | node 3 (wy[2]) | bias (wy[3]) |
| --- | --- | --- | --- |
| -3,817048 | 4,107138 | 4,629582 | -0,307594 |

الجدول 16

عوامل قياس خرج الصيغة الأساسية

|  |  |  |
| --- | --- | --- |
|  | bmin | bmax |
| ODG | 3,98– | 0,22 |

## 3.6 الصيغة المتقدمة

تستخدم *الصيغة المتقدمة* كل من نموذج الأذن القائم على بنك المرشحات ونموذج الأذن القائم على التحويل FFT. وتستخدم المتغيرات MOV: RmsModDiffA *و*RmsNoiseLoudAsymA *و*AvgLinDistA *و*Segmental NMRB *و*EHSB. ويُجرى تقابل لهذه المتغيرات MOV الخمسة مع مؤشر جودة وحيد يستخدم شبكة عصبية على النحو الموصوف في الفقرة 1.6 التي تضم خمس عقد في الطبقة المخفية. وترد في الجداول من 17 إلى 21 معلمات التقابل.

الجدول 17

المتغيرات MOV المستخدمة في الصيغة المتقدمة

| المتغير MOV | الغرض |
| --- | --- |
| RmsNoiseLoudAsymA | جهارة التشوه |
| RmsModDiffA | تغييرات في التشكيل (تتعلق بالاضطراب) |
| AvgLinDistA | التشوهات الخطية (الاستجابة الترددية، إلخ) |
| Segmental NMRB | النسبة ضوضاء إلى قناع |
| EHSB | البنية التوافقية للخطأ |

الجدول 18

عوامل القياس لعقد دخل الصيغة المتقدمة

| index (i) | MOV (x[i]) | amin[i] | amax[i] |
| --- | --- | --- | --- |
| 0 | RmsModDiffA | 13,298751 | 2166,5 |
| 1 | RmsNoiseLoudAsymA | 0,041073 | 13,24326 |
| 2 | Segmental NMRB | -25,018791 | 13,46708 |
| 3 | EHSB | 0,061560 | 10,226771 |
| 4 | AvgLinDistA | 0,024523 | 14,224874 |

الجدول 19

أوزان مدخلات الصيغة المتقدمة

| index (*i*) | MOV (x[i]) | node 1 (wx[i,0]) | node 2 (wx[i,1]) | node 3 (wx[i,2]) | node 4 (wx[i,3]) | node 5 (w4[i,4]) |
| --- | --- | --- | --- | --- | --- | --- |
| 0 | RmsModDiffA | 21,211773 | -39,913052 | -1,382553 | -14,545348 | -0,320899 |
| 1 | RmsNoiseLoudAsymA | -8,981803 | 19,956049 | 0,935389 | -1,686586 | -3,238586 |
| 2 | Segmental NMRB | 1,633830 | -2,877505 | -7,442935 | 5,606502 | -1,783120 |
| 3 | EHSB | 6,103821 | 19,587435 | -0,240284 | 1,088213 | -0,511314 |
| 4 | AvgLinDistA | 11,556344 | 3,892028 | 9,720441 | -3,287205 | -11,031250 |
| 5 | bias | 1,330890 | 2,686103 | 2,096598 | -1,327851 | 3,087055 |

الجدول 20

أوزان عقد خرج الصيغة المتقدمة

| node 1 (wx[i,0]) | node 2 (wx[i,1]) | node 3 (wx[i,2]) | node 4 (wx[i,3]) | node 5 (w4[i,4]) | bias (wy[5]) |
| --- | --- | --- | --- | --- | --- |
| -4,696996 | -3,289959 | 7,004782 | 6,651897 | 4,009144 | -1,360308 |

الجدول 21

عوامل قياس خرج الصيغة المتقدمة

|  |  |  |
| --- | --- | --- |
|  | bmin | bmax |
| ODG | -3,98 | 0,22 |

# 7 مطابقة عمليات التنفيذ

## 1.7 عام

يعرض هذا القسم مجموعة من عناصر الاختبار للتحقق من التنفيذ الأمثل للطريقة.

## 2.7 الاختيار

اُختيرت عناصر الاختبار من قاعدة البيانات 3 (DB3)، والتي اُستخدمت في التحقق من صحة النماذج. ولتبسيط الاختبار، اُختيرت مجموعة فرعية من عناصر قاعدة البيانات DB3 البالغ عددها 84 عنصراً. وتتكون هذه المجموعة الفرعية من 16 عنصراً. وكان المعيار الرئيسي للاختيار هو أن تغطي قيم المتغيرات MOV وDI (*مؤشر التشويه*) الناتجة نطاقاً واسعاً.

## 3.7 إعدادات اختبار المطابقة

عناصر الاختبار متاحة للتنزيل كملفات WAV (نسق Microsoft RIFF). تم اعتيان من جميع العناصر عند kHz 48 والتشكيل PCM بمعدل 16 بتة. والإشارة المرجعية وإشارة الاختبار التي قدمها الاتحاد الدولي للاتصالات هي بالفعل مكيفة من حيث الزمن والمستوى ببعضهما البعض، بحيث لا تكون هناك حاجة إلى كسب إضافي أو تعويض للتأخير. ويجب تعديل خوارزمية القياس إلى مستوى استماع يبلغ dB SPL 92.

## 4.7 فاصل التسامح المقبول

من أجل المطابقة مع التوصية، يجب أن تستنسخ قيم DI المحسوبة القيم الواردة في الجدولين 22 و23، مع تسامح أقل من 0,02± لجميع عناصر الاختبار[[7]](#footnote-7). وإذا لم يسفر التنفيذ عن نتائج ضمن هذا التسامح فإنها لا تتطابق مع هذه التوصية.

## 5.7 عناصر الاختبار

يعرض الجدولان التاليان أسماء العناصر المرجعية وعناصر الاختبار[[8]](#footnote-8)، وقيم DI الناتجة. ويرتبط الجدول 22 بالصيغة الأساسية، ويحتوي الجدول 23 على قيم الصيغة المتقدمة.

الجدول 22

عناصر الاختبار وقيم DI الناتجة للصيغة الأساسية

| العنصر | DI | ODG |
| --- | --- | --- |
| acodsna.wav | 1,304 | -0,676 |
| bcodtri.wav | 1,949 | -0,304 |
| ccodsax.wav | 0,048 | -1,829 |
| ecodsmg.wav | 1,731 | -0,412 |
| fcodsb1.wav | 0,677 | -1,195 |
| fcodtr1.wav | 1,419 | -0,598 |
| fcodtr2.wav | -0,045 | -1,927 |
| fcodtr3.wav | -0,715 | -2,601 |
| gcodcla.wav | 1,781 | -0,386 |
| icodsna.wav | -3,029 | -3,786 |
| kcodsme.wav | 3,093 | 0,038 |
| lcodhrp.wav | 1,041 | -0,876 |
| lcodpip.wav | 1,973 | -0,293 |
| mcodcla.wav | -0,436 | -2,331 |
| ncodsfe.wav | 3,135 | 0,045 |
| scodclv.wav | 1,689 | -0,435 |

الجدول 23

عناصر الاختبار وقيم DI الناتجة للصيغة المتقدمة

|  |  |  |
| --- | --- | --- |
| العنصر | DI | ODG |
| acodsna.wav | 1,632 | -0,467 |
| bcodtri.wav | 2,000 | -0,281 |
| ccodsax.wav | 0,567 | -1,300 |
| ecodsmg.wav | 1,594 | -0,489 |
| fcodsb1.wav | 1,039 | -0,877 |
| fcodtr1.wav | 1,555 | -0,512 |
| fcodtr2.wav | 0,162 | -1,711 |
| fcodtr3.wav | -0,783 | -2,662 |
| gcodcla.wav | 1,457 | -0,573 |
| icodsna.wav | -2,510 | -3,664 |
| kcodsme.wav | 2,765 | -0,029 |
| lcodhrp.wav | 1,538 | -0,523 |
| lcodpip.wav | 2,149 | -0,219 |
| mcodcla.wav | 0,430 | -1,435 |
| ncodsfe.wav | 3,163 | 0,050 |
| scodclv.wav | 1,972 | -0,293 |

المرفق 1  
بالملحق 2  
  
عملية التحقق

# 1 عام

في عام 1994 اعتمد قطاع الاتصالات الراديوية المسألة ITU-R 210/10 – طرائق التقييم الموضوعية لجودة الإدراكية، وتم إنشاء فريق مهام. وكان من بين الإجراءات الأولى إطلاق دعوة مفتوحة لتقديم المقترحات، وتم تلقي ردود من ستة من مقدمي النماذج.

تم بذل الكثير من الجهد لتحديد إجراءات عملية التحقق. وتبين أن من المفيد تجميع قاعدة بيانات أولى، يشار إليها بـاسم DB1، تتكون من مواد من اختبارات الاستماع التي تم إجراؤها بالفعل. وكان التركيز الرئيسي على الجودة الصوتية المتوسطة والعالية وبالتالي لم تؤخذ في الاعتبار إلا النتائج المستقاة من اختبارات الاستماع المتطابقة مع التوصية ITU-R BS.1116. وتمثل المواد المستقاة من هذه الاختبارات مواد بث مهمة للكوديكات ذات معدلات البتات المنخفضة مثل MPEG1 Layer II، وMPEG1 Layer III، وDolby AC2، وMini Disc، وNICAM وغيرها. وتم إنشاء قاعدة البيانات 1 لتزويد مقدمي النماذج بمنصة مشتركة تضم المواد التي تغطي نطاقاً كبيراً من الانحطاطات ومجموعة متنوعة من الكوديكات والتدهور الناتج عن الكوديكات المتشلشلة. يمكن الاطلاع على وصف تفصيلي للاختبارات المجمعة في قاعدة البيانات 1 في المرفق 2 بالملحق 2.

ومن الواضح أن طريقة القياس الموضوعي لجودة الصوت المدركة التي تحاكي السلوك البشري لا يمكن التحقق من صحتها إلا على قاعدة بيانات تحتوي على نتائج مستقاة من الاختبارات الشخصية. ويتطلب التحقق المناسب قاعدة بيانات تعتمد على مواد غير معروفة. ولهذا السبب كان من الضروري إجراء اختبارات استماع جديدة. وبما أن طريقة القياس ينبغي أن تستهدف بشكل مثالي أي نوع من النقائص التي يمكن أن تظهر في تطبيقات البث، فلا ينبغي إدراج نقائص التشفير فحسب، بل يجب أيضاً إدراج النقائص التقليدية مثل التشوه والضوضاء. وقد تم إنشاء قاعدة البيانات 2 وقاعدة البيانات 3 في عامي 1996 و1997 على التوالي لتلبية هذه المتطلبات. وإلى جانب الكوديكات المضمنة بالفعل في قاعدة البيانات DB1، تم أيضاً إدراج الكوديكين Dolby AC-3 وAAC. ويمكن الاطلاع على تفاصيل إضافية في المرفق 2 بالملحق 2.

وينبغي أن يُراعى في التحقق أوجه عدم اليقين، التي غالباً ما تُقدم على شكل فواصل ثقة، متأصلة في اختبارات الاستماع الشخصية. ويعتمد حجم فاصل الثقة على عدد من العوامل. وأهمها خبرة الأشخاص القائمين بالاختبار وإجراءات التدريب والسياق الذي تُعرض فيه عناصر الاختبار، فضلا عن عدد الأشخاص القائمين بالاختبار.

ويستند تكييف الطريقة الموضوعية الواردة في هذه التوصية والتحقق منها إلى "مستمع ذي خبرة متوسطة". وتُستخدم القيم المتوسطة من تقييمات الجودة الشخصية مع فواصل الثقة %95 لتحديد خصائص "المستمع ذي الخبرة المتوسطة ".

تعتبر اختبارات الاستماع الشخصية حساسة جداً للعوامل المختلفة التي تؤثر على النتائج وأُنتجت الدرجات SDG لكل من قاعدة البيانات 2 وقاعدة البيانات 3 في ثلاثة مواقع اختبار مختلفة وتحقق عدد من الدراسات من إمكانية دمج البيانات بالفعل. وعلى الرغم من عدم توصل جميع الدراسات إلى استنتاجات متطابقة، فقد وجد أن من المعقول دمج البيانات وشكلت قاعدة البيانات المدمجة هذه الأساس لعملية التحقق.

وقُسمت عملية التحقق إلى ثلاث مراحل:

– المرحلة 1: المرحلة التنافسية

– المرحلة 2: المرحلة التعاونية

– المرحلة 3: الاختيار النهائي

وسيرد وصف هذه المراحل بالتفصيل في الأقسام التالية.

# 2 المرحلة التنافسية

اُقترحت ست طرائق (DIX، NMR، PAQM، PERCEVAL، POM، TTA) للقياس الموضوعي لجودة الصوت المدركة وتقرر مقارنة أداء هذه الطرائق باستخدام قاعدة البيانات 2 ومجموعة فرعية من قاعدة البيانات 1. وتم إنشاء قاعدة البيانات 2 في بداية عام 1996. وكان اختيار مواد الاختبار النهائي بمثابة جهد مشترك بين SR (السويد) وهيئة الإذاعة البريطانية (المملكة المتحدة). وأُجريت اختبارات الاستماع في NRK في النرويج، وDR في الدنمارك، وNHK في اليابان. وأعد تحليل إحصائي للبيانات المستقاة من الاختبارات من قبل شركة Deutsche Telekom (ألمانيا) وTeracom (السويد). وخلال المرحلة 1، تم إنشاء البيانات الموضوعية في موقع محايد (Swisscom، سويسرا). بعد ذلك، تلقى مقدمو النماذج النصف الأول من قاعدة البيانات 2 للتكييف النهائي للطرائق (المرحلة 2). وأخيراً، تم إنشاء درجات الاختلاف الموضوعي الجديدة في Swisscom.

تم إجراء تحليل أداء للطرائق من قبل شركة Teracom (السويد) وكذلك من قبل مقدمي النماذج أنفسهم. وعلى الرغم من أن نتائج بعض الطرائق المقترحة أظهرت ترابطاً كبيراً بالدرجات SDG، فقد كان هناك إجماع على أن أياً من الطرائق لم تلب متطلبات المستعملين. وأظهرت دراسة منفصلة أن أياً من الطرائق المقترحة لم تكن أفضل بكثير من الطرائق الأخرى. ولذلك تقرر تطوير طريقة قياس محسنة كجهد مشترك بين جميع مقدمي النماذج الحاليين. وينبغي مقارنة أداء الطريقة الجديدة بإحدى الطرائق القائمة بالفعل والمشار إليها باسم النموذج B3.

# 3 المرحلة التعاونية

ارتكزت المرحلة التعاونية على فكرة الجمع بين أفضل عناصر الطرائق المختلفة في طريقة واحدة جديدة. ولتلبية احتياجات المستعملين على أفضل وجه، تقرر تطوير صيغتين من الطريقة. واحدة مناسبة لعمليات التنفيذ في الوقت الفعلي، والتي قد تتطلب قدرة حاسوبية أكبر لتحقيق دقة أعلى.

وصُمم إجراء التحقق للطرائق الجديدة بطريقة مماثلة لتلك الخاصة بالمرحلة التنافسية. وكان لا بد من إنشاء قاعدة بيانات جديدة (DB3). وحُددت العناصر والشروط في نهاية المطاف في ربيع عام 1997 وتم تجميعها في SR وSwisscom وBBC. ويمكن الاطلاع على وصف كامل لقاعدة البيانات في المرفق 2 بالملحق 2. وأجريت اختبارات الاستماع الشخصية في ثلاثة مواقع اختبار، وهي Deutsche Telekom وNHK وSR. وطبقت جميع المواقع "طريقة الحجب المزدوج الثلاثي الحوافز مع المرجع المحجوب" الموصوفة في التوصية ITU-R BS.1116. وتم جمع النتائج المستقاة من اختبارات الاستماع في السويد. وأُجري تحليل إحصائي شامل لنتائج اختبار الاستماع في شركة Teracom وكذلك من قبل أطراف أخرى. ونتيجة لهذا التحليل، تم استبعاد بعض المستمعين من مواصلة التقييم. وجُمعت النتائج من مواقع الاختبار لتكوين قاعدة البيانات 3.

وفي خريف عام 1997، تم إصدار 52 عنصراً من قاعدة البيانات إلى مقدمي النماذج. كُيفت الطرائق الجديدة مع البيانات الجديدة. ونظراً لوجود العديد من قيم ضبط المعلمات التي أعطت نتائج مماثلة، فقد تم تأجيل قرار الاختيار النهائي لأطول فترة ممكنة. وفي النهاية، اُستخدمت في سويسرا العناصر المتبقية البالغ عددها 32 عنصراً في التحقق من الطرائق الجديدة على مجموعة بيانات "غير معروفة".

وبالإضافة إلى ذلك، تم استخدام نتائج اختبار استماع جديد أجرته CRC (كندا) في التحقق من الطرائق الجديدة إزاء المواد "غير المعروفة". ويرد وصف عملية الاختيار والتحقق في الأقسام التالية.

# 4 التحقق

أُجريت اختبارات موسعة للصيغ الثماني عشرة المحددة لطريقة القياس الموضوعي. ويرد في هذا القسم وصف معايير الاختيار، وتُعرض نتائج مقارنة الدرجات SDG مع النتائج التي تم الحصول عليها من الصيغ الثماني عشرة لطريقة القياس. وكان الهدف هو اختيار الصيغ المثلى التي سيوصى بها الاتحاد الدولي للاتصالات والتحقق منها.

معايير الاختيار

العلاقة بين النتائج الشخصية والموضوعية هي المعيار الأكثر وضوحاً للتحقق من الطريقة الموضوعية. وبالإضافة إلى ذلك، تم طرح معيارين آخرين يأخذان في الاعتبار موثوقية القيم المتوسطة للتحقق – المعيار AES ومخطط التسامح.

ويُعرف المعيار AES، الذي طُرح لربط دقة النموذج بدقة اختبار الاستماع، وفقاً للمعادلة التالية:

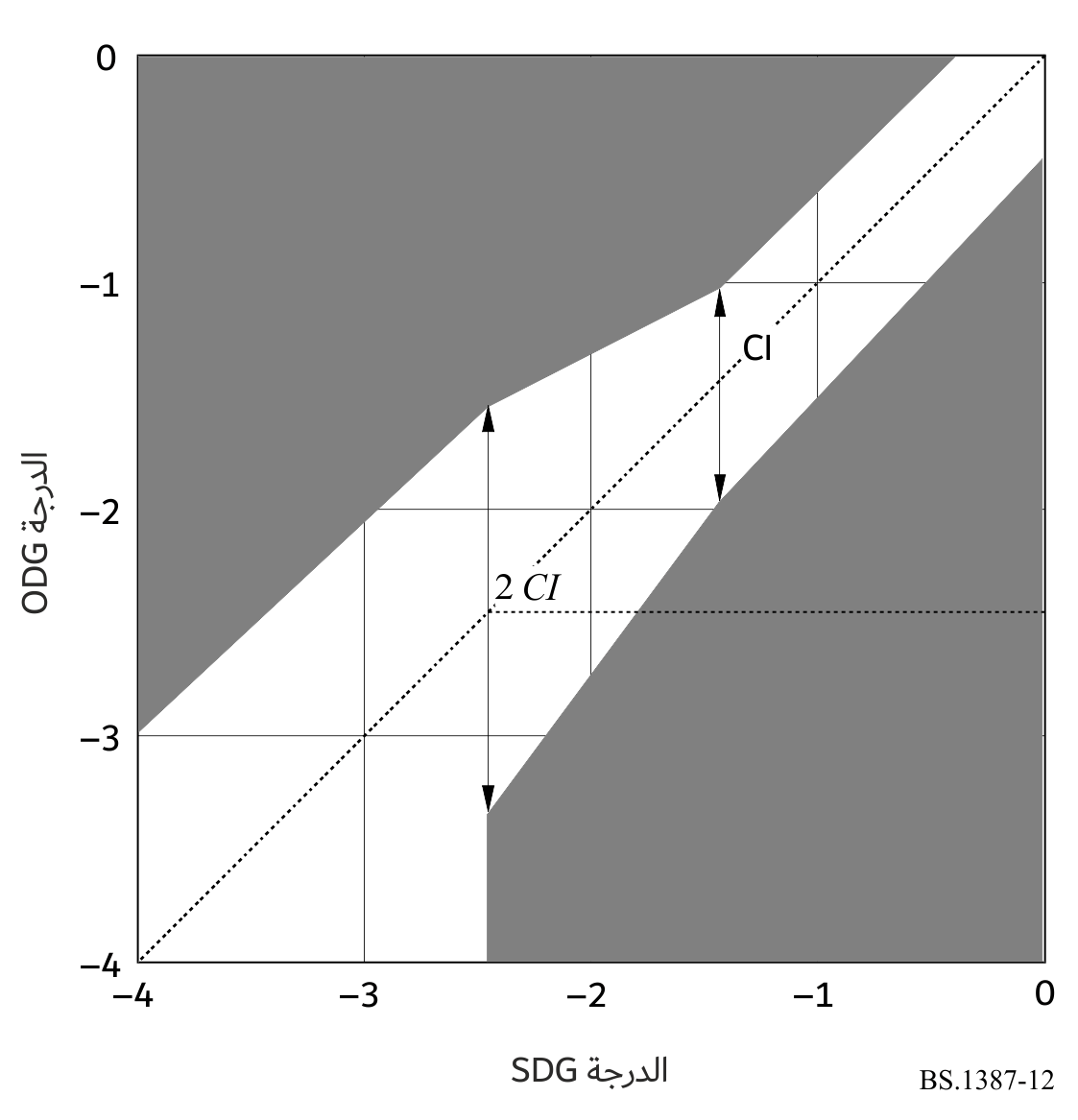
 if *CI*  0.25 then *CI*  0.25

حيث *CI فاصل الثقة.*

وصُمم مخطط التسامح للسماح بانحرافات مختلفة للدرجات ODG عن الدرجات SDG في الطرف العلوي والسفلي من سلم الانحطاط. ويرتبط نطاق التسامح بفواصل الثقة في اختبارات الاستماع. ويقتصر هذا النطاق على قيمة لا تقل عن 0,25 من الدرجات. واُستخدمت مسافة الدرجات ODG خارج مخطط التسامح إلى مخطط التسامح لتقييم جودة طريقة القياس.

الشكل 12

مخطط التسامح، فاصل الثقة *CI* ≥ 0,25



## 1.4 المقارنة بين الدرجات SDG وODG

قُسمت القياسات الموضوعية إلى ثلاث مراحل مختلفة. ففي المرحلة 1، كانت جميع عناصر الاختبار الـبالغ عددها 84 غير معروفة للجميع باستثناء فريق الاختيار. وفي المرحلة 2، تم إصدار معلومات عن 52 عنصراً. وتضمنت المعلومات كلاً من قيم الدرجات SDG والمقتطفات الصوتية الفعلية. وفي المرحلة الثالثة، اُستخدمت هذه المعلومات لتحسين أداء صيغ الطريقة. ويرجى العلم أنه تم اختبار أربع صيغ إضافية خلال المرحلة 3، مقارنة بالمرحلة 1. وحُسبت قيم الدرجات SDG المقدمة من البيانات التي تم إصدارها بواسطة 75 شخصاً مؤهلاً للاختبار.

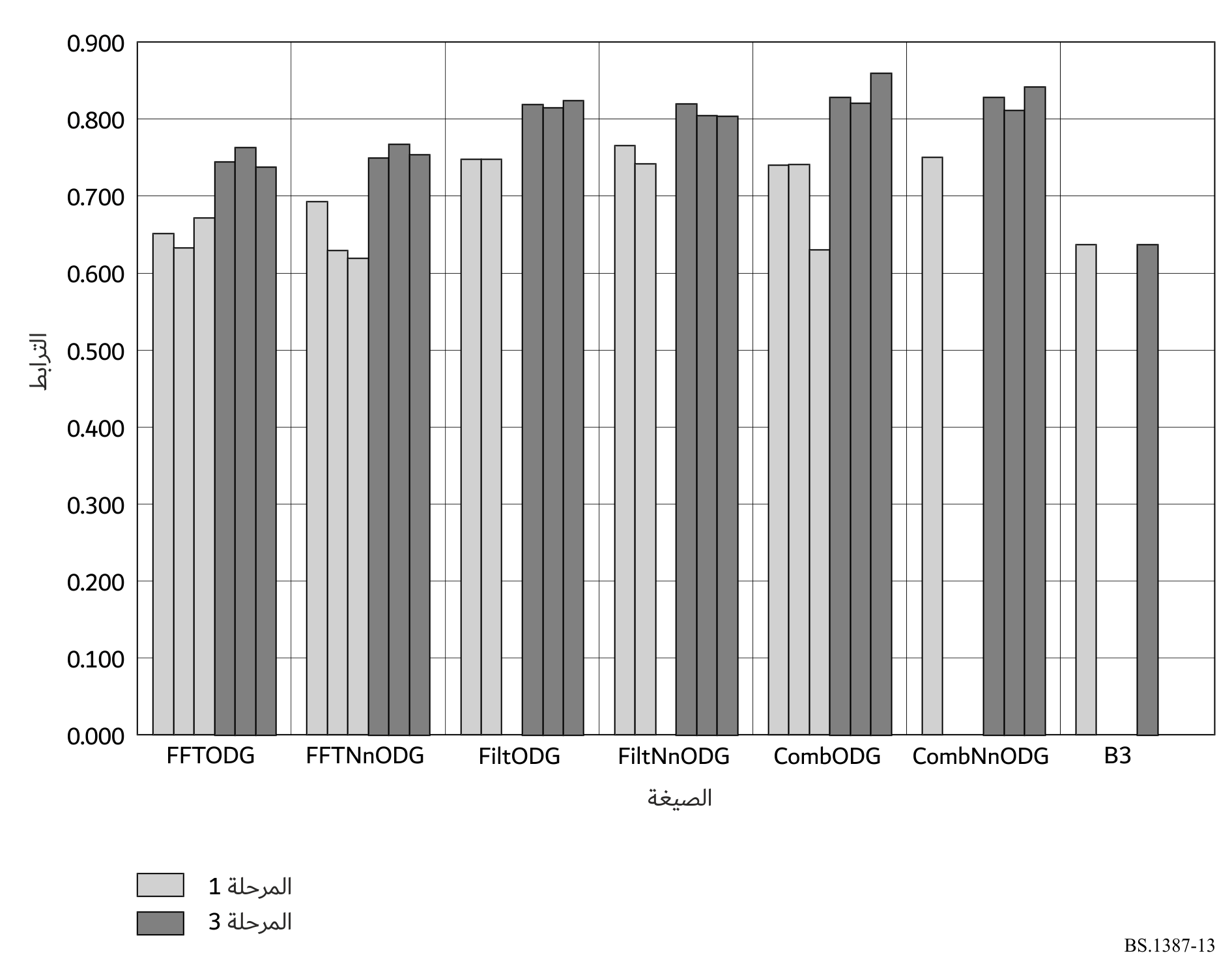
وهناك العديد من الأساليب المختلفة لتقييم مدى الدقة التي تبرز بها الدرجات ODG الدرجات SDG. ولسوء الحظ، لا توجد قيمة واحدة توضح الأداء الكامل بحق. وبدلاً من ذلك، يتعين على المرء أن يبحث في عدد من وجهات النظر. وترد الترابطات في الفقرة 2.4 وتُعرض درجات الخطأ المطلق (AES) في الفقرة 3.4. والنموذج B3 هو أحد النماذج التي اختبرها قطاع الاتصالات الراديوية في عام 1996 وتقرر مقارنة مختلف الصيغ الجديدة مع هذه الصيغة الأقدم.

## 2.4 الترابط

رُسمت مخططات الترابط من المرحلتين 1 و3 في الشكل 13 (84 عنصراً) والشكل 14 (32 عنصراً).

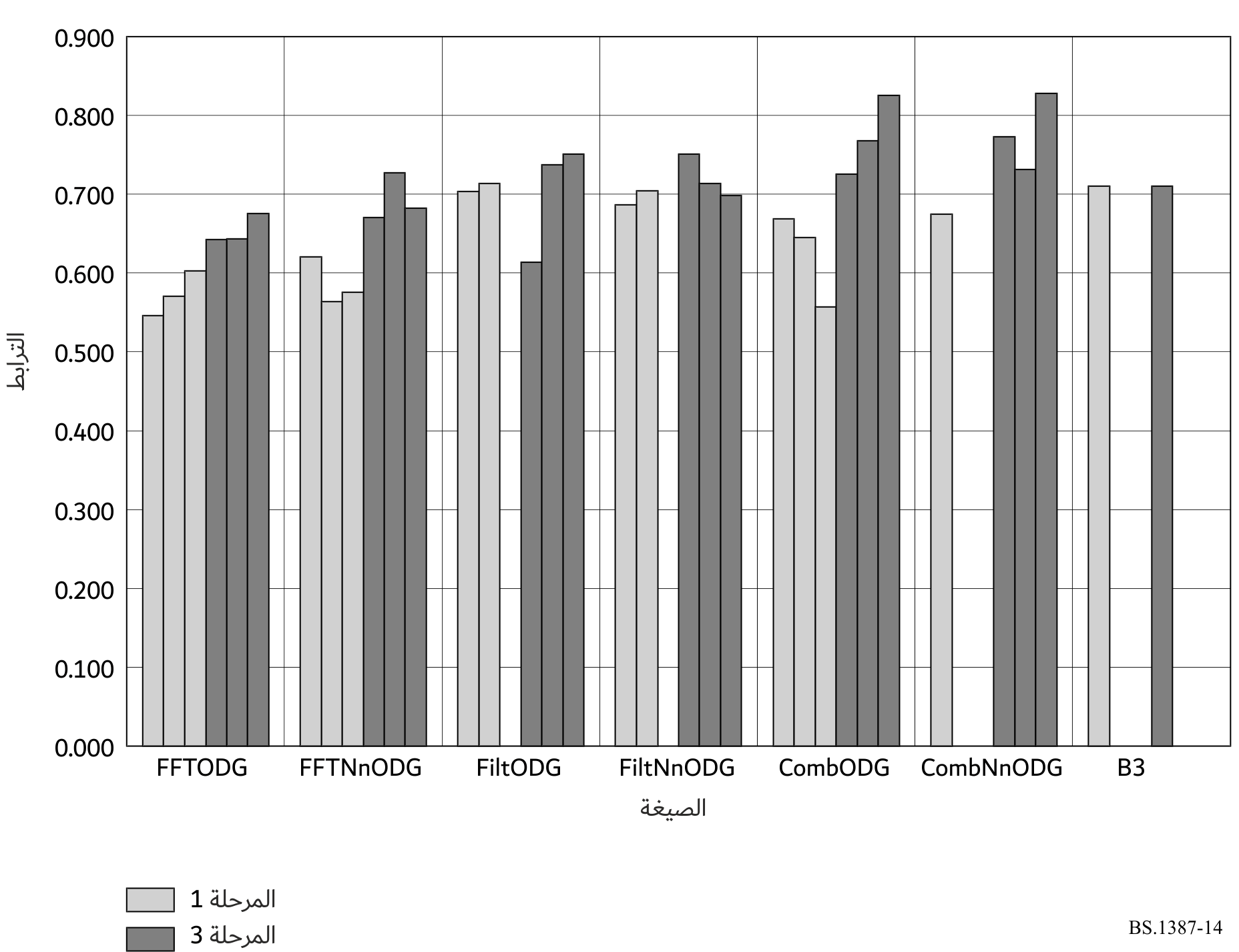
الشكل 13

الترابط بين الدرجات SDG وod  
تم إدراج جميع العناصر البالغ عددها 84



الشكل 14

الترابط بين الدرجات SDG وODG  
تم إدراج العناصر التي لم تصدر البالغ عددها 32

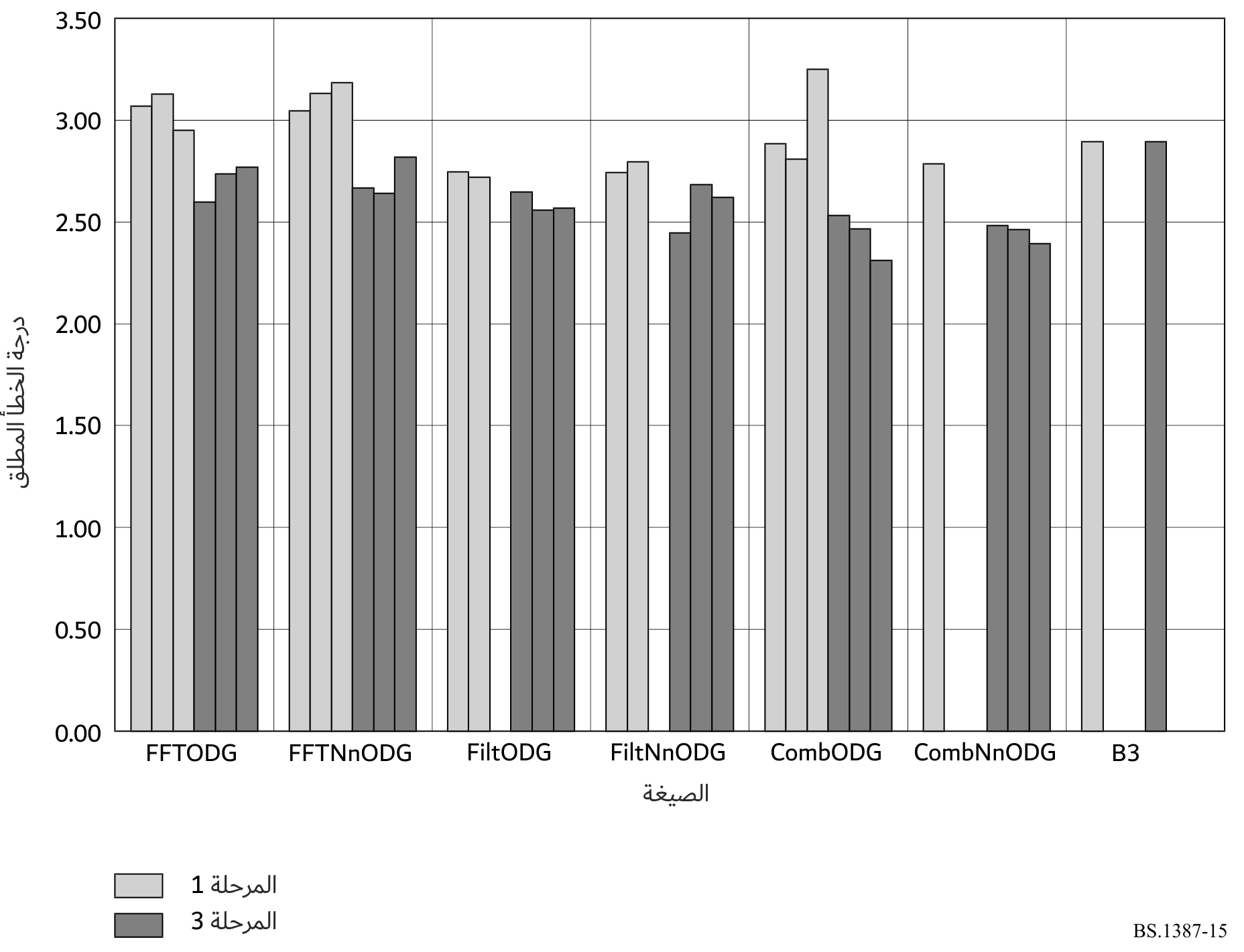


## 3.4 درجة الخطأ المطلق (AES)

أي نموذج ينتج، في المتوسط، قيم للدرجات ODG ضمن فاصل الثقة للدرجات SDG سيحصل على قيمة لدرجة الخطأ المطلق (AES) تقترب من 2. وترد نظرة عامة على قيم الدرجة AES في الشكلين 15 و16.

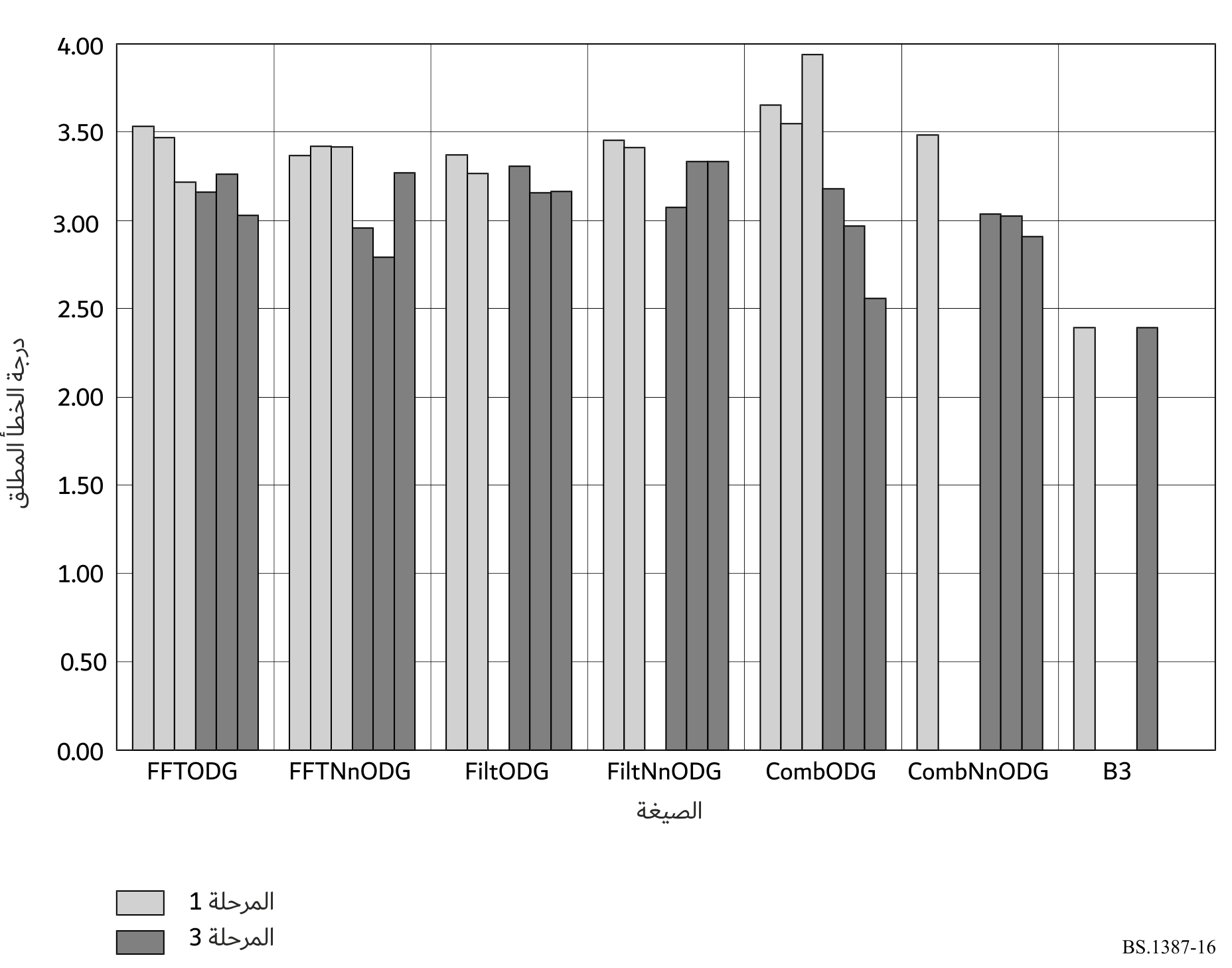
الشكل 15

درجة الخطأ المطلق للصيغ المختلفة  
تم إدراج جميع العناصر البالغ عددها 84



الشكل 16

درجة الخطأ المطلق للصيغ المختلفة  
تم إدراج العناصر التي لم تصدر البالغ عددها 32



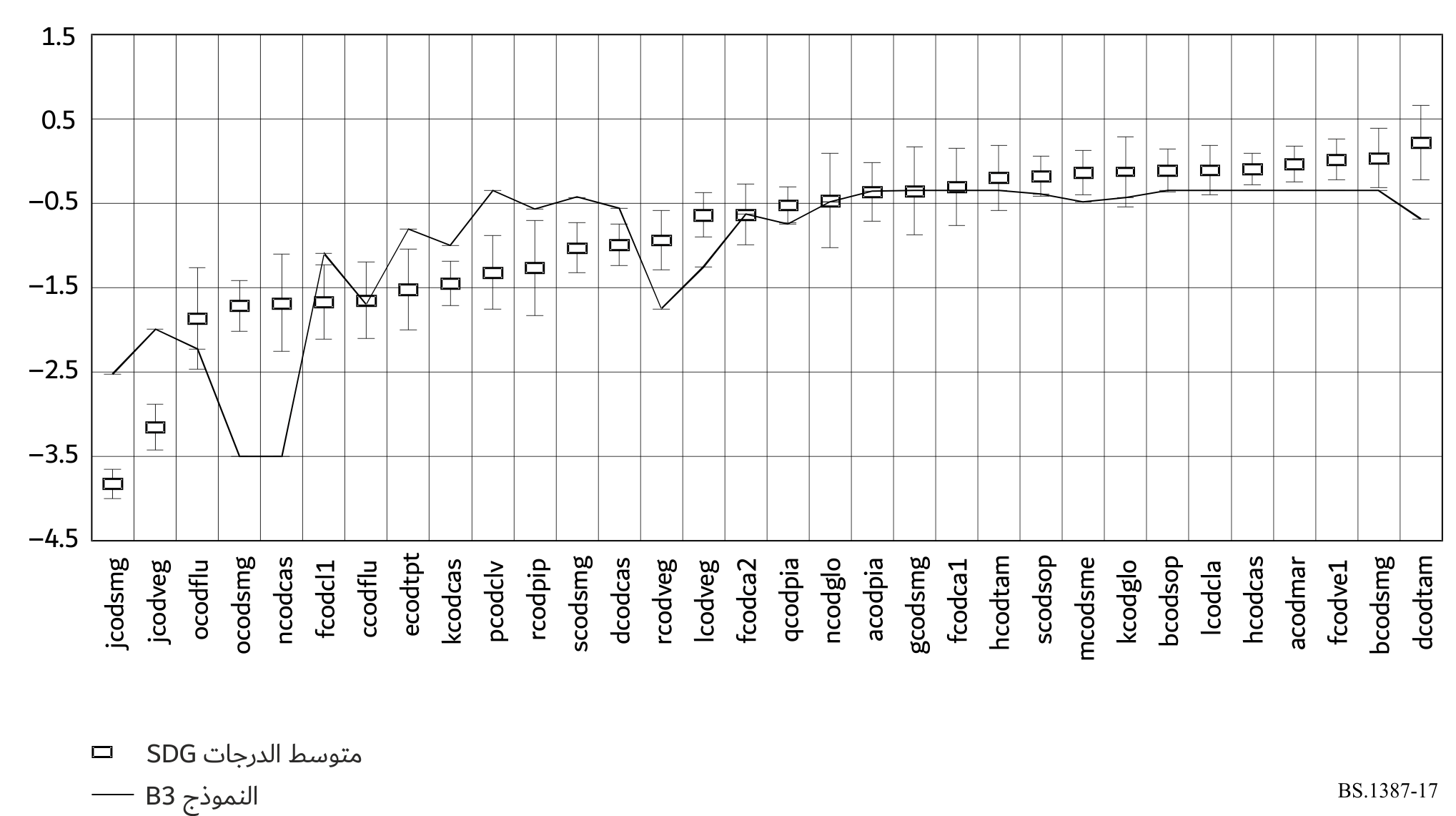
## 4.4 مقارنة الدرجات ODG مقابل فاصل الثقة

كان أداء معظم الصيغ متشابهاً، كما كان متوقعاً على الأرجح. وعُرض الكثير من المخططات في الاجتماع ولكن في هذه التوصية فإن مستودع المخططات محدود. ولمزيد من التفاصيل، يرجى الرجوع إلى تقرير اختبار التحقق الكامل.

وفي الأشكال من 17 إلى 22، تم رسم متوسط الدرجات SDG وفاصل الثقة والدرجات ODG للعناصر التي لم تصدر البالغ عددها 32 للنموذج B3 وصيغتي النموذج FFTNnODG1 وCombNnODG3.

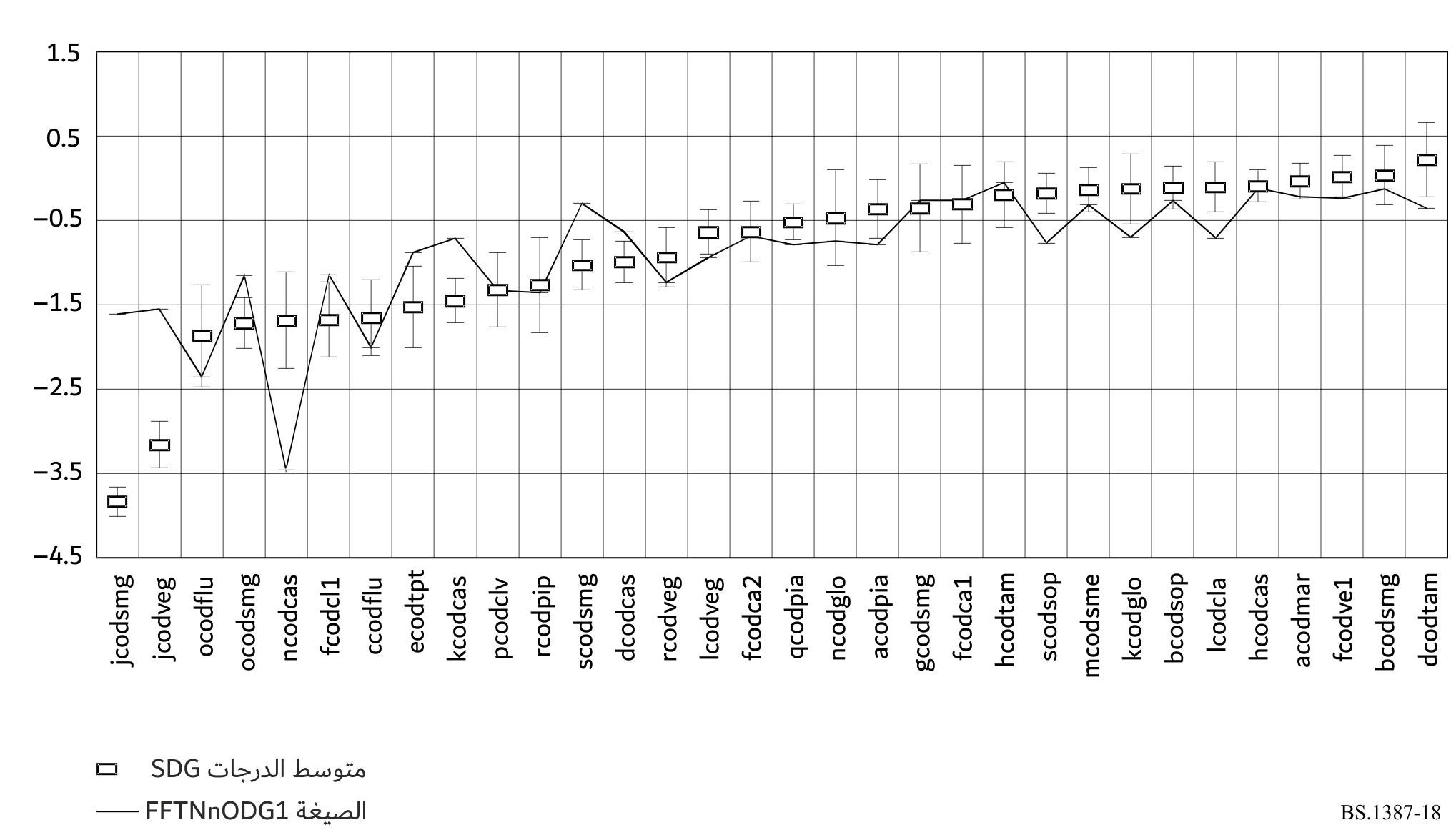
الشكل 17

النموذج B3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG للعناصر التي لم تصدر البالغ عددها 32



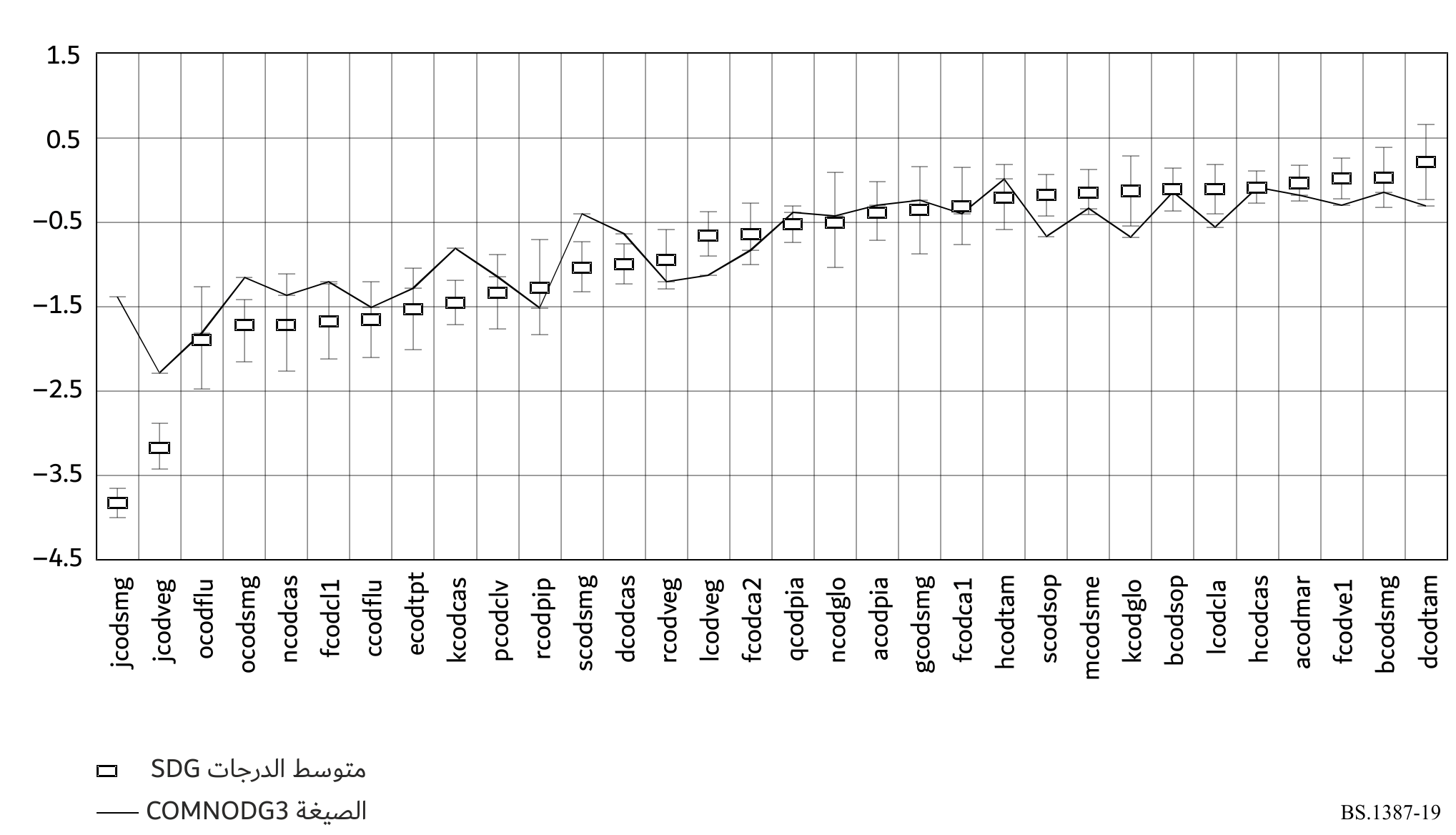
الشكل 18

الصيغة FFTNnODG1: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG بعد المرحلة الثالثة للعناصر التي لم تصدر البالغ عددها 32



الشكل 19

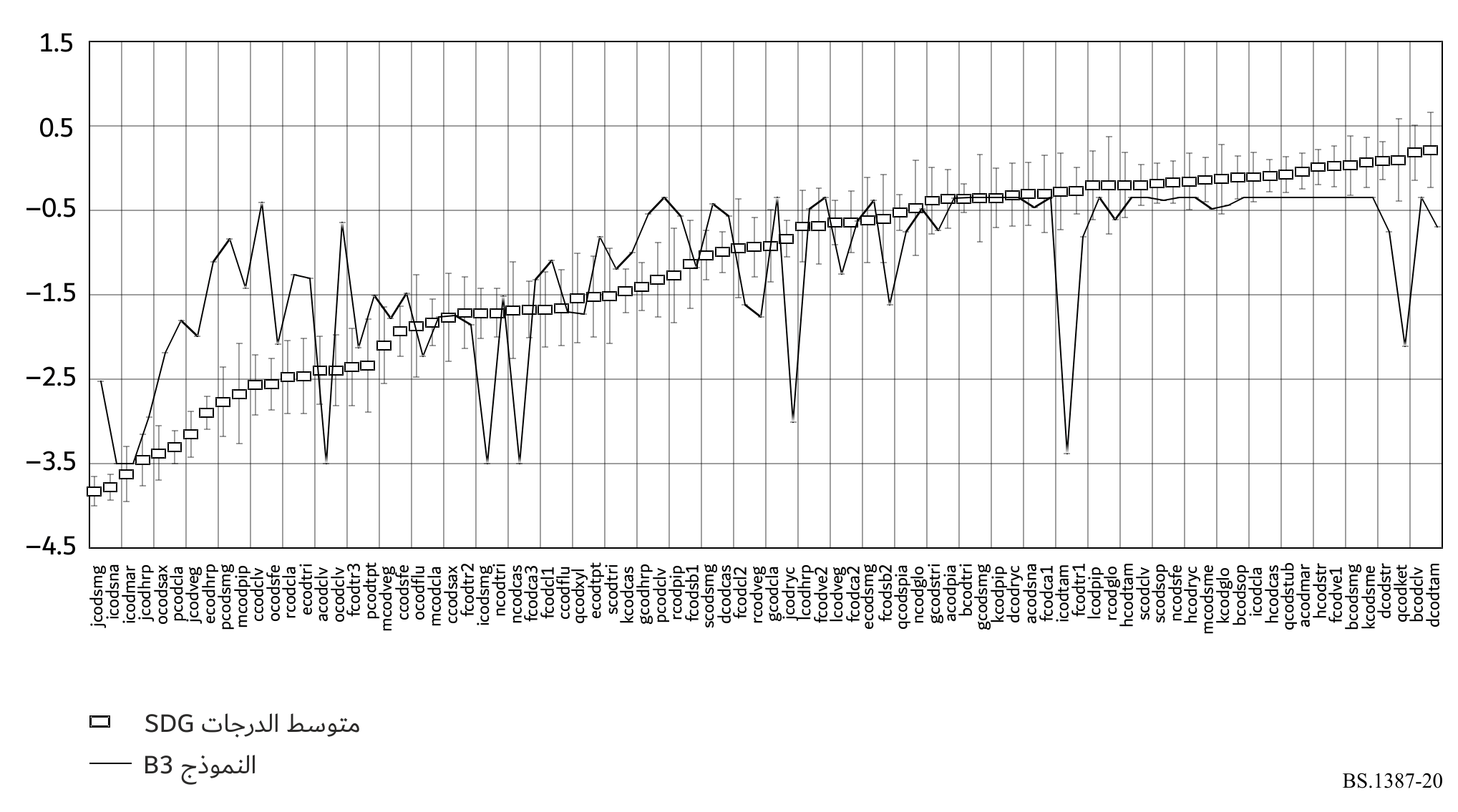
الصيغة CombNnODG3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG بعد المرحلة الثالثة للعناصر التي لم تصدر البالغ عددها 32



وترد في الشكل 20 والشكل 21 مخططات مماثلة، ولكن في هذه الحالة لجميع العناصر الـبالغ عددها 84 خلال المرحلة 3. وبالإضافة إلى ذلك، يوضح الشكل 22 أداء الصيغة CombNnODG3.

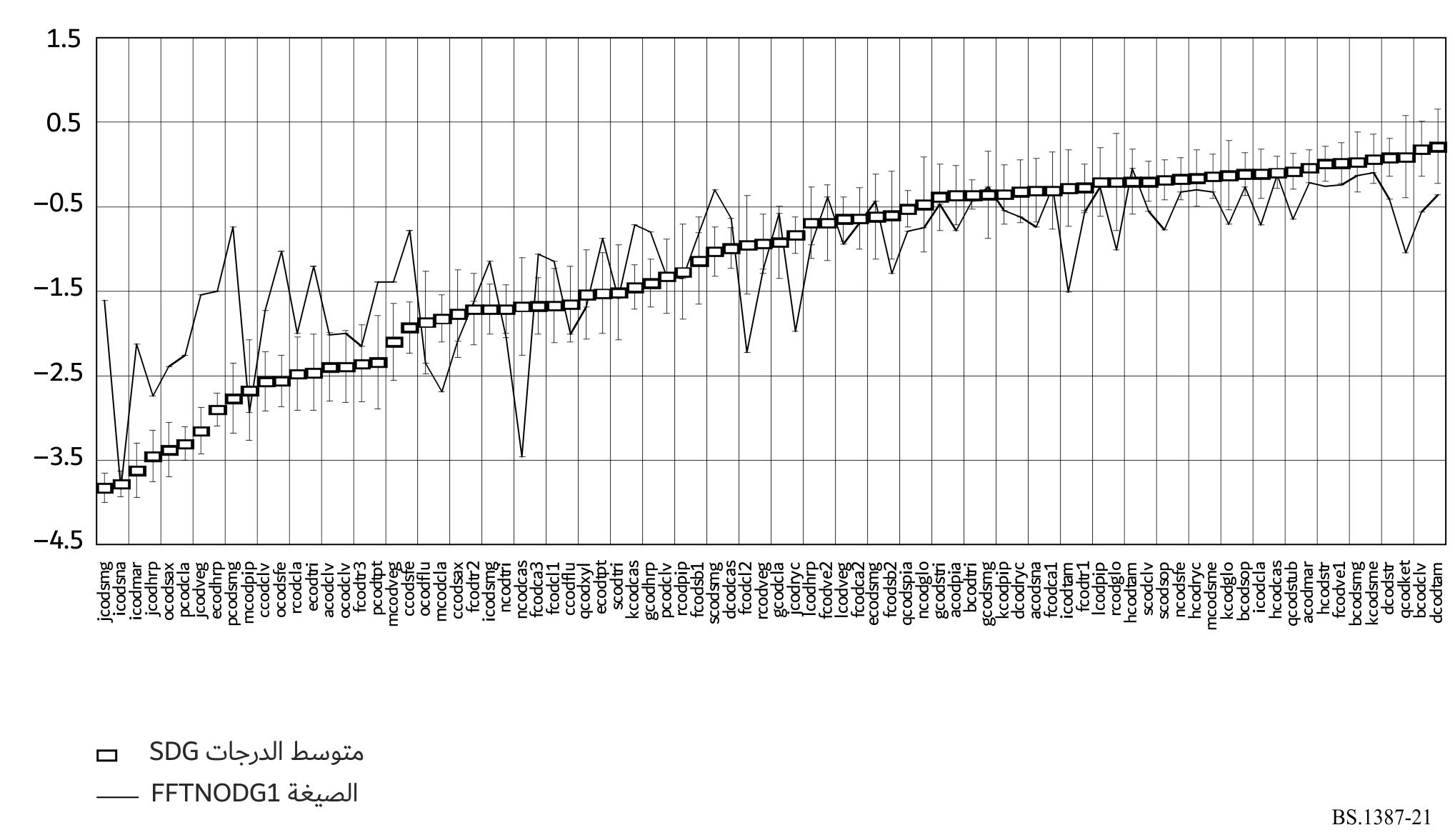
الشكل 20

النموذج B3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG لجميع العناصر البالغ عددها 84



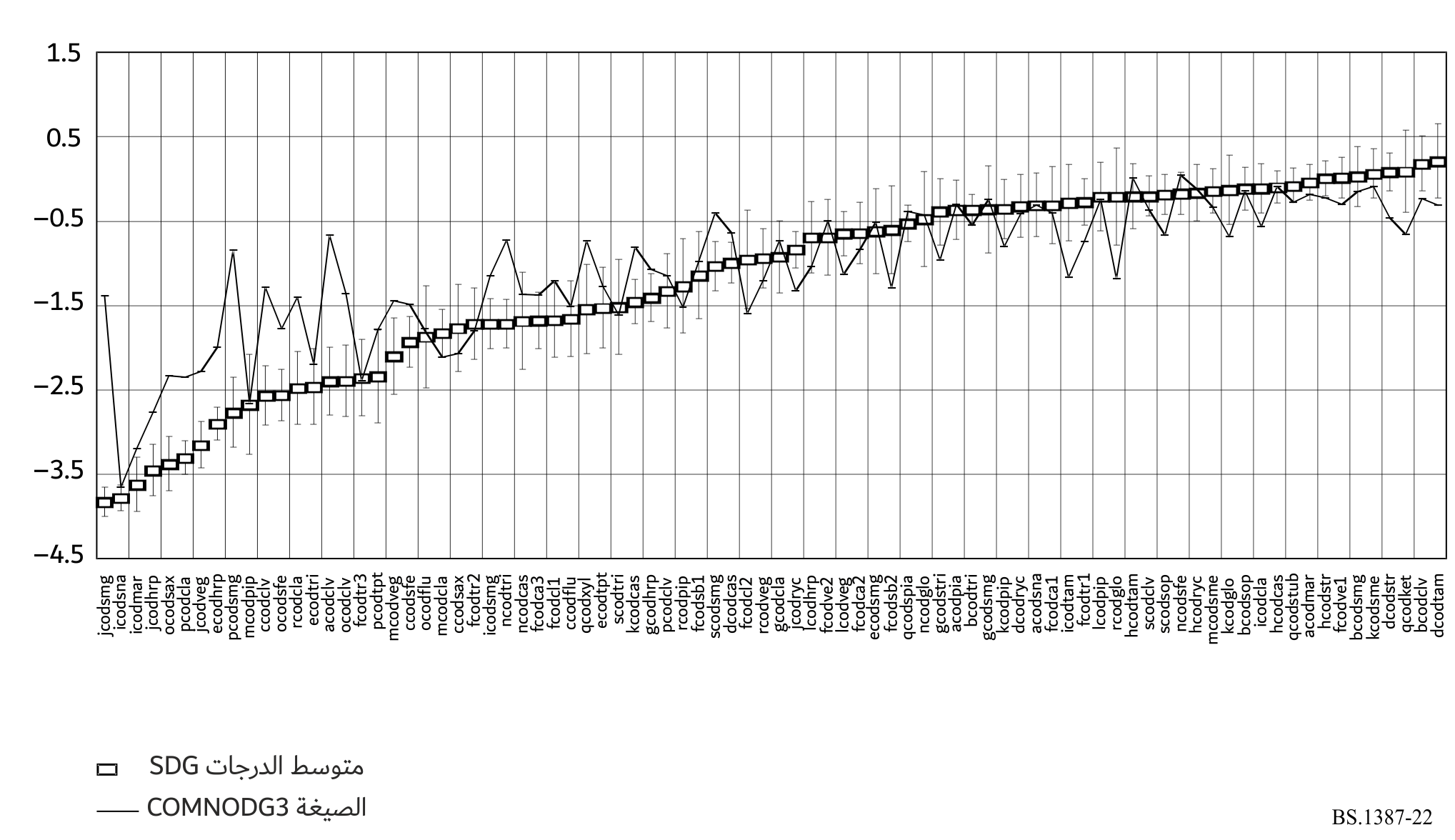
الشكل 21

الصيغة FFTNnODG1: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG أثناء المرحلة الثالثة لجميع العناصر البالغ عددها 84



الشكل 22

الصيغة CombNnODG3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG أثناء المرحلة الثالثة لجميع العناصر البالغ عددها 84

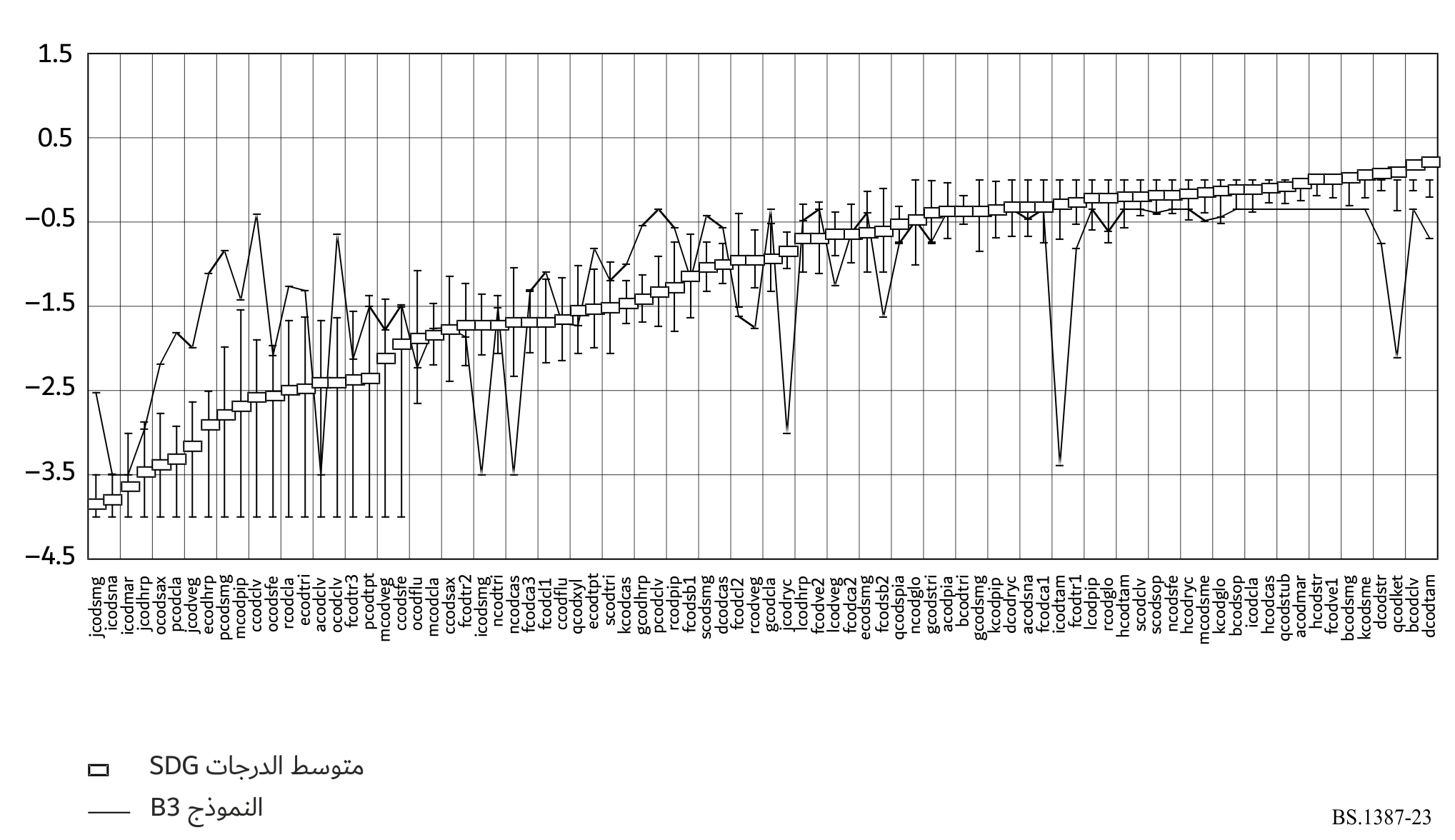


## 5.4 مقارنة الدرجات ODG مقابل فاصل التسامح

حدد قطاع الاتصالات الراديوية متطلباً مستهدفاً للمستعمل يمكن تقابله مع فاصل التسامح. وتكون المتطلبات المستهدفة أكثر صرامة بالنسبة للمستويات الأعلى من جودة الصوت وأقل صرامة بالنسبة لجودة الصوت المنخفضة. وتوضح الأشكال أدناه الأداء في هذا البعد للنموذج B3 وصيغتي النموذج FFTNnODG1 وCombNnODG3 لجميع العناصر البالغ عددها 84 خلال المرحلة 3.

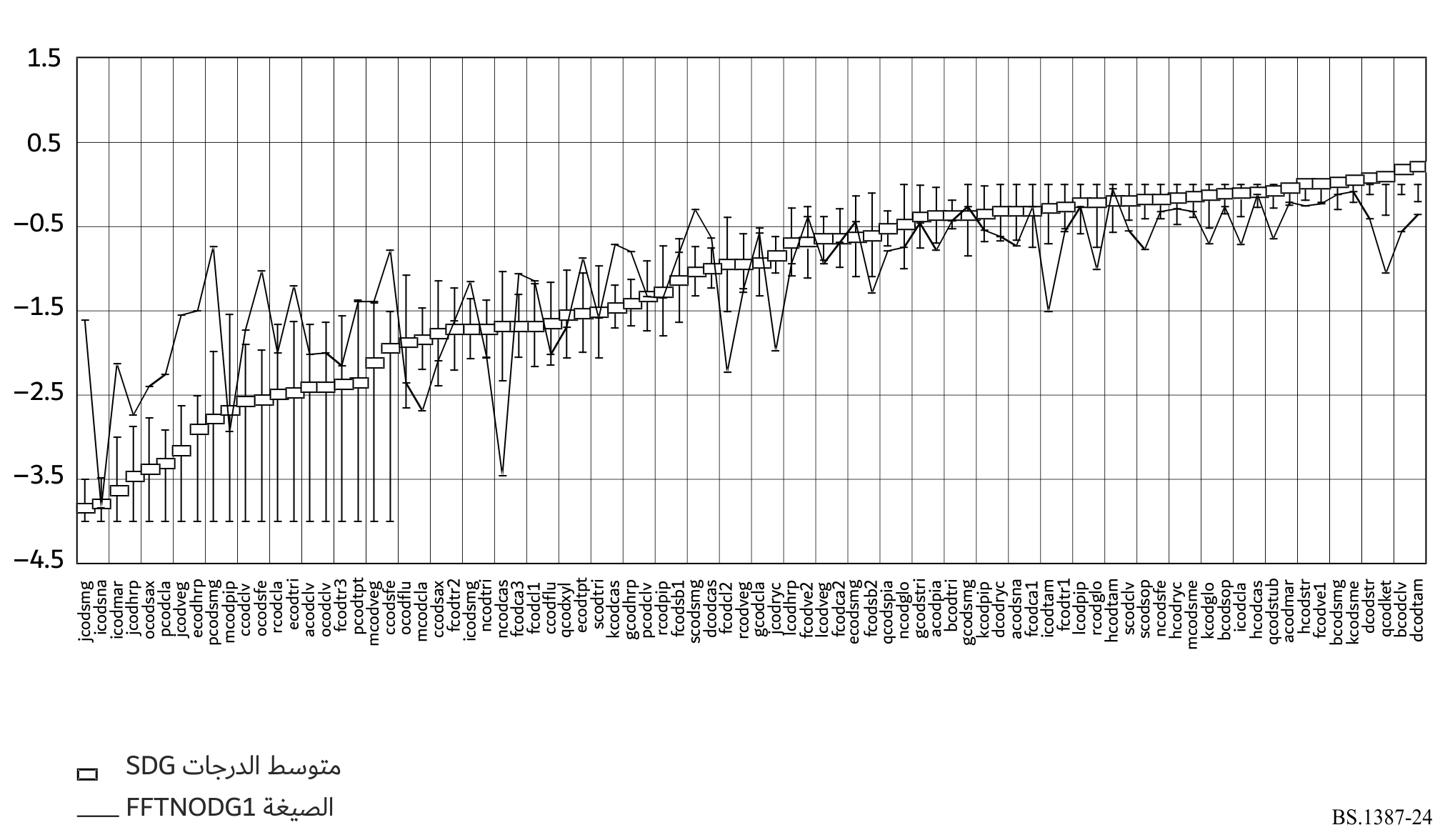
الشكل 23

النموذج B3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG أثناء المرحلة 3 لجميع العناصر البالغ عددها 84



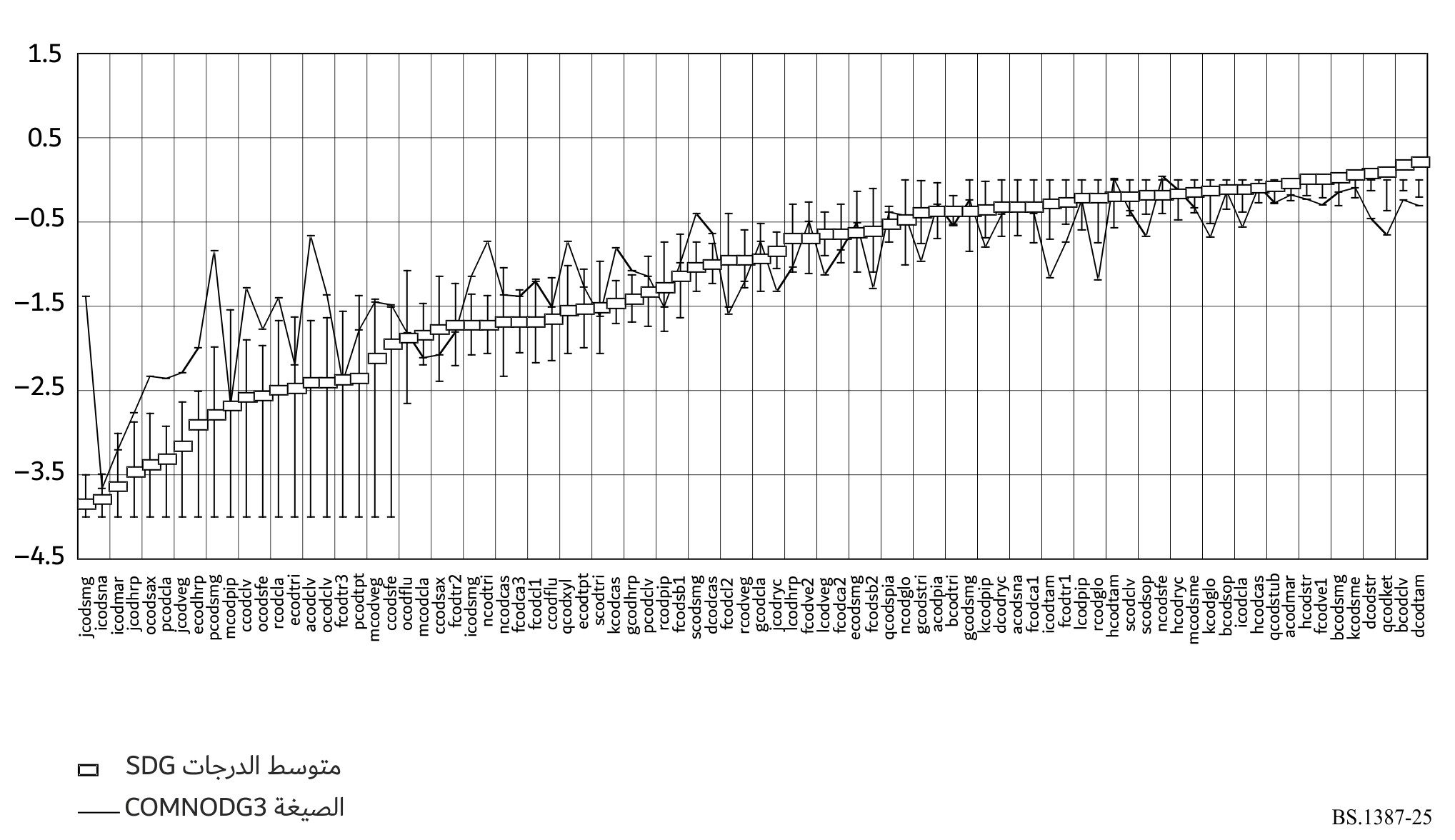
الشكل 24

الصيغة FFTNnODG1: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG أثناء المرحلة 3 لجميع العناصر البالغ عددها 84



الشكل 25

الصيغة CombNnODG3: مخطط متوسط الدرجات SDG وفاصل الثقة والدرجات ODG أثناء المرحلة 3 لجميع العناصر البالغ عددها 84



# 5 اختيار صيغ النموذج المثلى

قُدمت ثماني عشرة صيغة مختلفة للنموذج إلى موقع الاختبار الموضوعي: 6 صيغ قائمة على التحويل FFT، وست صيغ تستخدم بنك مرشحات مُكيَّف إدراكياً وست صيغ تستخدم مزيجاً من كل من التحويل FFT وبنك المرشحات. وتهدف مجموعة التحويل FFT إلى استخدامها في مخطط قياس قادر على التنفيذ في الوقت الفعلي، بينما تحتاج المجموعتان المتبقيتان إلى تعقيد حاسوبي أكبر ومن المتوقع أن تحققان دقة أعلى. ويمكن تقسيم صيغ النموذج الست في كل مجموعة إلى مجموعتين فرعيتين: إحداهما تستخدم الشبكات العصبية والأخرى **لا** تستخدم الشبكات العصبية. وأُدرج أيضاً أداء النموذج المرجعي من أبريل 1996 (B3) في قاعدة البيانات 3 لأغراض المقارنة.

## 1.5 معايير الاختيار المسبق استناداً إلى الترابط

– قُيم أداء 18 صيغة مختلفة للنموذج والنموذج المرجعي B3 باستخدام مجموعتي بيانات لم يتم استخدامهما في الغالب لتدريب النماذج (لم يتم استخدام الجزء الثاني من المرحلتين 1 و3 وقاعدة البيانات CRC للتدريب).

– أُنشئت قاعدة البيانات 3 خصيصاً لتقييم النماذج الإدراكية. واُستخدم حوالي نصف قاعدة البيانات هذه لتدريب النماذج. واُستخدم الترابط بين النتائج الشخصية والنتائج الموضوعية على العناصر المتبقية (DB3\_2nd) لتقييم النماذج الموضوعية.

– أنشأت CRC قاعدة بيانات CRC لتقييم مخططات التشفير الصوتي الإدراكي. ولم تُستخدم قاعدة البيانات هذه لتدريب النماذج الإدراكية. واُستخدم الترابط بين النتائج الشخصية والنتائج الموضوعية على جميع العناصر لتقييم النماذج الموضوعية.

وترد نتائج قاعدة البيانات 3 في الجدولين 24 و25 (الترابطات ودرجة الخطأ AES).

وفي مرحلة الاختيار المسبق، تم أخذ جميع المدخلات التي تحدد أداء جميع صيغ النموذج كمياً بعين الاعتبار مع إعطاء أهمية كبيرة لـقاعدة البيانات DB3. وبناءً على المقارنة الشاملة، تقرر إجراء مقارنة إضافية بين صيغتي النموذج في كل مجموعة والتي يبدو أنها الأفضل. ويبين الجدول 24 الترابط بين هذه الصيغ الست (ثلاث مرات لصيغتين من كل نموذج).

الجدول 24

الترابط بين الدرجات SDG وODG

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | FFTNnODG1 | FFTNnODG2 | FiltODG2 | FiltODG3 | CombNnODG3 | CombODG3 | B3 |
| DB3\_2nd | 0,671 | 0,728 | 0,738 | 0,751 | 0,828 | 0,826 | 0,710 |
| CRC | 0,837 | 0,779 | 0,862 | 0,839 | 0,851 | 0,777 | 0,656 |

الجدول 25

درجة الخطأ المطلق

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | FFTNnODG1 | FFTNnODG2 | FFTODG2 | FFTODG3 | CombNnODG3 | CombODG3 | B3 |
| DB3\_2nd | 2,96 | 2,79 | 3,16 | 3,16 | 2,91 | 2,56 | 2,39 |
| CRC | 1,55 | 1,85 | 1,61 | 1,67 | 1,61 | 1,90 | 2,78 |

تُظهر صيغتا النموذج القائمتان على التحويل FFT أداءً مختلفاً في قاعدتي البيانات. ويحتاج قرار اختيار إحدى الصيغتين إلى معايير إضافية. ومع الأخذ في الاعتبار أن قاعدة البيانات CRC كانت غير معروفة تماماً، يبدو أن أداء الصيغة FFTNnODG1 أفضل قليلاً.

وتعطي صيغ النموذج المدمجة نتائج أفضل على قاعدتي البيانات مقارنة بالصيغ القائمة على التحويل FFT. ومع ذلك، فإنها تظهر أداءً أسوأ على قاعدة البيانات CRC مقارنة بصيغ بنك المرشحات. ويمكن اعتبار صيغ بنك المرشحات حالة خاصة من صيغ النموذج المدمجة حيث يكون ترجيح متغيرات خرج الصيغة القائمة على التحويل FFT صفراً. ولذلك تُفضل صيغ النموذج المدمجة.

## 2.5 تحليل عدد القيم الشاذة

يُعبر عن دقة البيانات الشخصية الواردة من اختبارات الاستماع عبر فاصل قدره %95 حول متوسط عدد من المستمعين. وقُيم أداء صيغ النموذج أيضاً من خلال النظر في القيم الشاذة. ويعتبر العنصر بمثابة قيمة شاذة عندما يكون الفرق بين البيانات الشخصية والبيانات الموضوعية أكبر من ضعف فاصل الثقة.

ويوضح الجدول 26 عدد القيم الشاذة لصيغ النموذج الست لقاعدة البيانات DB3. و"حساس" يعني أن صيغة النموذج تشير إلى جودة صوت أقل من التقييم الشخصي، بينما "غير حساس" هو العكس.

الجدول 26

القيم الشاذة

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | FFTNnODG1 | FFTNnODG2 | FiltODG2 | FiltODG3 | CombNnODG3 | CombODG3 |
| حساس | 10 | 4 | 4 | 4 | 3 | 5 |
| غير حساس | 13 | 13 | 11 | 13 | 12 | 14 |
| الإجمالي | 23 | 17 | 15 | 17 | 15 | 19 |

وبالنظر إلى عدد القيم الشاذة، يبدو أن صيغة النموذج FFTNnODG2 تعمل بشكل أفضل من صيغة النموذج FFTNnODG1. وتُظهر صيغة النموذج CombNnODG3 أفضل أداء بين الصيغ الأكثر دقة.

## 3.5 تحليل شدة القيم الشاذة

الجدول 27

العناصر ذات الانحرافات التي تزيد عن 0,1 من درجات الفرق بين التنبؤ والدرجات SDG

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| عدد القيم الشاذة | 14 | 4 | 4 | 12 | 12 | 9 | 6 |
| صيغة النموذج | B3 | CombODG3 | CombNnODG3 | FFTNnODG1 | FFTNnODG2 | FiltODG2 | FiltODG3 |
| العنصر | jcodsmg | jcodsmg | jcodsmg | jcodsmg | jcodsmg | jcodsmg | jcodsmg |
|  | qcodket | qcodket | pcodsmg | qcodket | qcodket | rcodpip | rcodpip |
|  | pcodsmg | pcodsmg | ccodclv | pcodsmg | pcodsmg | qcodket | qcodket |
|  | pcodcla | icodmar | acodclv | ocodsfe | pcodcla | pcodsmg | pcodsmg |
|  | ocodclv |  |  | ncodcas | ocodsfe | ocodsax | ocodsax |
|  | ncodcas |  |  | mcodcla | ncodglo | kcodpip | fcodsb2 |
|  | mcodpip |  |  | jcodveg | kcodcas | jcodveg |  |
|  | jcodryc |  |  | jcodryc | jcodveg | fcodsb2 |  |
|  | icodtam |  |  | icodtam | jcodryc | fcodcl2 |  |
|  | icodsmg |  |  | icodmar | icodtam |  |  |
|  | fcodsb2 |  |  | fcodcl2 | ecodhrp |  |  |
|  | ecodhrp |  |  | ecodhrp | ccodclv |  |  |
|  | ccodclv |  |  |  |  |  |  |
|  | acodclv |  |  |  |  |  |  |

الجدول 28

العناصر ذات الانحرافات التي تزيد عن 1,5 من درجات الفرق بين التنبؤ والدرجات SDG

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| عدد القيم الشاذة | 8 | 2 | 2 | 3 | 2 | 2 | 2 |
| صيغة النموذج | B3 | CombODG3 | CombNnODG3 | FFTNnODG1 | FFTNnODG2 | FiltODG2 | FiltODG3 |
| العنصر |  | jcodsmg | jcodsmg | jcodsmg | jcodsmg | jcodsmg | jcodsmg |
|  | qcodket |  | pcodsmg |  |  |  |  |
|  | pcodsmg | pcodsmg |  | pcodsmg | pcodsmg |  |  |
|  |  |  |  |  |  | pcodsmg | pcodsmg |
|  |  |  |  | ncodcas |  |  |  |
|  | ncodcas |  |  |  |  |  |  |
|  | jcodryc |  |  |  |  |  |  |
|  | icodtam |  |  |  |  |  |  |
|  | icodsmg |  |  |  |  |  |  |
|  | ccodclv |  |  |  |  |  |  |
|  | acodclv |  |  |  |  |  |  |

# 6 الخلاصة

بعد القرارات المتخذة أثناء التطوير، تم اختيار صيغتين للنموذج من بين الصيغ المقترحة البالغ عددها 18 من أجل تلبية متطلبات التطبيقات المحددة لطرائق القياس الموضوعية. وحُددت صيغة قليلة التعقيد، مصممة للسماح بالتنفيذ في الوقت الفعلي بكفاءة من حيث التكلفة، وصيغة أعلى دقة، لا تعمل بالضرورة في الوقت الفعلي. وبالنسبة لعملية الاختيار، تم تطبيق وتحليل المعايير المذكورة أعلاه.

وباعتبارها صيغة الوقت الفعلي، اُختير نموذج قائم على التحويل FFT، يشار إليه باسم "FFTNnODG1" للأسباب التالية:

لم تُظهر أي من الصيغ القائمة على التحويل FFT ميزة كبيرة على الصيغ الأخرى فيما يتعلق بأحد المعايير المذكورة أعلاه. وفيما يتعلق بمعاملات الترابط بين الدرجات SDG وODG، فإن عدد وشدة القيم الشاذة لكل صيغة من الصيغ التي تم التحقق منها لها مزاياها وعيوبها. ومع ذلك، جدير بالذكر أن النموذج "FFTNnODG1" حقق أفضل ترابط في قاعدة بيانات CRC'97 التي كانت غير معروفة تماماً (0,837).

وبالنسبة للصيغة ذات الدقة الأعلى، تم إعطاء الأفضلية للصيغة المدمجة، التي تشتمل على التحويل FFT وبنك المرشحات، لأن هذا النهج يشتمل أيضاً على مجموعة فرعية من نموذج بنك المرشحات الخالص وبالتالي يجب أن يكون له إمكانات أداء أفضل. وكانت الصيغ الست المختلفة للنموذج المدمج متاحة لعملية الاختيار. وأظهرت الصيغة المختارة "CombNnODG3" قيماً شاذة أقل بالإضافة إلى ترابط أعلى (r=0.851 for CRC’97) مقارنة بالصيغ الأخرى. وكان ترابط هذه الصيغة بالنسبة لقاعدة البيانات 3 الكاملة بنفس الترتيب من حيث الحجم الموجود الخاص بالصيغة "CombODG3"، ولكنها أظهرت ترابطاً أعلى مقارنة بالصيغ الأخرى.

المرفق 2  
بالملحق 2  
  
أوصاف قواعد البيانات المرجعية

# 1 مقدمة

أثناء تطوير طريقة القياس الموضوعي لجودة الصوت المدركة، تم استخدام عدد من قواعد البيانات للتدريب والتحقق.

وتحتوي بعض قواعد البيانات المدرجة على بيانات سماعات الرأس ومكبرات الصوت، وبعضها يحتوي على بيانات سماعات الرأس فقط. وبالنسبة لقواعد البيانات التي تحتوي على مجموعات منفصلة من البيانات المتاحة لعرض مكبرات الصوت وسماعات الرأس، اُستخدمت بيانات سماعات الرأس فقط.

ويُعرف العنصر على أنه جزء صوتي يستخدم في التقييم الشخصي. وتشير الحالة إلى حالة تدهور واحدة واُستخدمت جميع العناصر لجميع الحالات التجريبية باستثناء دراسات قاعدتي البيانات DB2 وDB3. واُستخدمت قاعدة البيانات DB3 جزئياً للتدريب وجزئياً للتحقق (تم استخدام 52 عنصراً من أصل 84 عنصراً للتدريب في المرحلة الثانية من التحقق).

التدريب

– MPEG90

– يغطي متوسط الدرجات SDG لكل عنصر بشكل منتظم النطاق من 0,0 إلى 4,0–.

– تقرير اختبار MPEG/Audio للفريق ISO/IEC JTC 1/SC 2/WG 11، الوثيقة MPEG90/N0030، أكتوبر 1990.

– MPEG91

– كان ما لا يقل عن 88 في المائة من متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,1 إلى 3,8–.

– تقرير اختبار MPEG/Audio للفريق ISO/IEC JTC 1/SC 2/WG 11، الوثيقة MPEG91/N0010، يونيو 1991.

– ITU92DI

– كان ما لا يقل عن 80 في المائة من متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,1 إلى 3,4–.

– ITU92CO

– كان ما لا يقل عن 96 في المائة من متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,2 إلى 2,4–.

– ITU93

– كان معظم متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,1– إلى 2,3–. ولم يكن هناك فرق كبير بين البيانات من المختبرين.

– Grusecوآخرون[1997].

– MPEG95

– كان ما لا يقل عن 63 في المائة من متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,2– إلى 3,8–.

– Meares وKim [1995]

– EIA95

– كان ما لا يقل عن 93 في المائة من متوسط الدرجات SDG لكل عنصر أعلى من 2,0–، وكان النطاق من 0,1 إلى 3,7–.

– Grusecوآخرون[1997].

– DB2

– لم تُستخدم جميع العناصر لجميع الحالات.

التحقق

– DB3

– لم تُستخدم جميع العناصر لجميع الحالات.

– CRC97

– يغطي متوسط الدرجات SDG لكل عنصر بشكل منتظم النطاق من 0,1 إلى 3,6–.

– Soulodreوآخرون[1998].

وتصف الأقسام التالية العناصر التي تم إدراجها في قواعد البيانات المختلفة والحالات التي تم تطبيقها.

# 2 العناصر لكل قاعدة بيانات



# 3 الحالات التجريبية

بالنسبة لجميع معدلات البتات مع بيان بالإشارة المجسمة kbit/s، يُحدد معدل البتات الإجمالي، على سبيل المثال، صوت مجسم kbit/s 256 يعني أنه تم تخصيص ما مجموعه kbit/s 256 لكلتا قناتي إشارة الصوت المجسم. وإذا لم تتم الإشارة إلى أي شيء آخر، يشير الصوت المجسم إلى تشفير قناة مستقلة.

## 1.3 MPEG90

ثلاثة معدلات للبتات: kbit/s 64 للقناة الأحادية، وkbit/s 192 وkbit/s 256 للصوت المجسم، وجميع المواد غير متاحة لقاعدة البيانات هذه.

– Musicam.

– SB-ADPCM.

## 2.3 MPEG91

ثلاثة معدلات للبتات: kbit/s 64 للقناة الأحادية، وkbit/s 192 وkbit/s 256 للصوت المجسم.

– MPEG1 Layer I.

– MPEG1 Layer II.

– MPEG1 Layer III.

– MUSICAM.

– ASPEC.

– NICAM.

## 3.3 ITU92DI

خمسة كوديكات توزيع: kbit/s 240 للصوت المجسم.

خضع كل عنصر للمعالجة بواسطة الكوديك نفسه ثلاث مرات جنباً إلى جنب مع انخفاض في المستوى بمقدار dB 0,1 قبل كل مرة.

– MPEG1 Layer II.

– MPEG1 Layer III.

– Dolby AC-2.

– Aware.

– NHK.

## 4.3 ITU92CO

ستة كوديكات مساهمة: kbit/s 360 للصوت المجسم. خضع كل عنصر للمعالجة بواسطة الكوديك نفسه ثلاث مرات جنباً إلى جنب مع انخفاض في المستوى بمقدار dB 0,1 قبل كل مرة.

– MPEG1 Layer II.

– MPEG1 Layer III.

– Dolby AC-2.

– Dolby Low-Delay.

– Aware.

## 5.3 ITU93

تشكيلات MPEG1 Layer II مع الكوديكات الترادفية

– كوديك بث فقط مع صوت مجسم kbit/s 256.

– كوديك بث فقط مع صوت مجسم kbit/s 192 (تشفير مجسم مشترك).

– ثمانية كوديكات مساهمة بمعدل kbit/s 360 متبوعة بكوديك بث واحد بمعدل kbit/s 256، كلها بصوت مجسم.

– ثمانية كوديكات مساهمة بمعدل kbit/s 360 متبوعة بكوديك بث واحد بمعدل kbit/s 192، كلها بصوت مجسم.

– خمسة كوديكات مساهمة بمعدل kbit/s 360 متبوعة بثلاثة كوديكات توزيع بمعدل kbit/s 240 وكوديك بث واحد بمعدل kbit/s 256، كلها بصوت مجسم.

– خمسة كوديكات مساهمة بمعدل kbit/s 360 متبوعة بثلاثة كوديكات توزيع بمعدل kbit/s 240 وكوديك بث واحد بمعدل kbit/s 192، كلها بصوت مجسم.

## 6.3 MPEG95

عمليات تنفيذ الكوديكات (kbit/s 64)

– تم اختيار اثنين وعشرين شكلاً مختلفاً للتشفير من مجموعة أكبر من طرائق التشفير المتاحة من 6 كوديكات تنفذ مجموعة فرعية مكونة من 4 نماذج ذات دقة منخفضة و17 نموذجاً للزمن/التردد عالية الدقة.

– كانت المؤسسات المشاركة هي AT&T، وFraunhofer، وSony، وGCL، وRAI/Alcatel، وPhilips.

– كانت جميع العناصر عبارة عن تسجيلات أحادية الأذن تم عرضها ثنائية الأذنين.

## 7.3 EIA95

– Eureka 147/MPEG1 Layer II #1 224 kbit/s stereo (تشفير مجسم مشترك)

– Eureka 147/MPEG1 Layer II #2 192 kbit/s stereo (تشفير مجسم مشترك)

– AT&T/Lucent 160 kbit/s stereo

– AT&T/Lucent/Amati #1 128 kbit/s stereo

– AT&T/Lucent/Amati #2 160 kbit/s stereo

– VOA/JPL 160 kbit/s stereo

– USADR-FM #1 128-256 kbit/s stereo (معدل بتات متغير)

– USADR-FM #2 128-256 kbit/s stereo (معدل بتات متغير)

– USADR-AM 96 kbit/s stereo

## 8.3 DB2

– MPEG1 Layer II 256 kbit/s stereo, 1, 3, 5, 7, and 9 stages

– Dolby AC2 256 kbit/s stereo, 1, 3, 5, 7, and 9 stages

– MPEG1 Layer II 192 kbit/s stereo (joint stereo coding)

– MPEG1 Layer II 64 kbit/s mono

– MPEG2 Layer II 64 kbit/s mono

– MPEG1 Layer II 384 kbit/s stereo

– MPEG1 Layer III 128, 160, 192 kbit/s, all stereo.

– APT-X 256 and 384 kbit/s both stereo.

– التقدير الكمي للتشوه

– تسجيل تماثلي 1 و2 و3 مراحل

– تقليم البيانات

## 9.3 DB3

– NICAM

– MiniDisc and MiniDisc + Layer II 192 kbit/s, stereo (joint stereo coding)

– Dolby AC2 256 kbit/s stereo, 1, 3, 5, 7, and 9 stages

– MPEG1 Layer II selection from Swisscom database, >192 kbit/s stereo

– MPEG1 Layer III 128 and 160 kbit/s both stereo (joint stereo coding)

– MPEG AAC 128 kbit/s stereo (joint stereo coding)

– MPEG Layer III 128 + Layer II, 384 + Layer II, 224 kbit/s, all stereo

– Dolby AC3 256 kbit/s stereo

– Dolby AC3 256 + MPEG Layer II, 224 kbit/s, both stereo

– التقدير الكمي للتشوه

– التشوه الكلي الناجم عن التوافقيات

– ضوضاء

## 10.3 CRC97

– AT&T PAC 64, 96, 128, and 160 kbit/s, all stereo

– Dolby AC3 128, 160, and 192 kbit/s, all stereo

– MPEG1 Layer II software 128, 160, and 192 kbit/s, all stereo

– MPEG1 Layer II hardware (ITIS) 96, 128, 160, 192 kbit/s, all stereo

– MPEG4 AAC 96 and 128 kbit/s, both stereo

– MPEG1 Layer III 128 kbit/s stereo

# 4 العناصر لكل حالة بالنسبة لقاعدتي البيانات DB2 وDB3

## 1.4 DB2

رقم الحالة العناصر

موقع الاختبار I، NHK، اليابان

Layer II, 256 kbit/s 1 stage CO13 CLA,RYC,SB1,STR

3 stages CO11 CLA,RYC,SB1,STR

5 stages CO19 CLA,RYC,SB1,STR

7 stages CO18 CLA,RYC,SB1,STR

9 stages CO15 CLA,RYC,SB1,STR

NBC (Dolby AC2) 1 stage CO1A CAS,RYC,STR,WIN

3 stages CO12 CAS,RYC,STR,WIN

5 stages CO17 CAS,RYC,STR,WIN

7 stages CO16 CAS,RYC,STR,WIN

9 stages CO14 CAS,RYC,STR,WIN

رقم الحالة العناصر

موقع الاختبار II، DR، الدانمارك

Layer II, 256 kbit/s 1 stage CO2B CLA,RYC,SB1,STR

Layer II, 192 kbit/s js CO25 CLA,RYC,SB1,STR

Layer II, 64 kbit/s mono CO27 MLA,MPE,MTR,MYC

NBC (Dolby AC2) 5 stages CO29 CAS,RYC,STR,WIN

MPEG2/L2 LSF CO22 MLA,MPE,MTR,MYC

Analogue 1 CO23 PER

Analogue 2 CO2A PER

Analogue 3 CO28 PER

Errors 1 CO24 GLO,HRN,TRI

Errors 2 CO21 GLO,HRN,TRI

Clipping CO26 BAS,CL2,TUB

موقع الاختبار III، NRK، النرويج

Layer II, 384 kbit/s CO34 CLA,RYC,SB1,STR

Layer II, 256 kbit/s 1 stage CO31 CLA,RYC,SB1,STR

NBC (Dolby AC2) 5 stages CO3B CAS,RYC,STR,WIN

Layer III (ASPEC3), 192 kbit/s CO32 CLA,STR,TAM,VEG

Layer III (ASPEC3), 128 kbit/s CO39 CLA,STR,TAM,VEG

Layer III (ASPEC3), 160 kbit/s CO3A  CLA,STR,TAM,VEG

APT-X, 256 kbit/s CO33 HAR,SB2,STR,TPT

APT-X, 384 kbit/s CO36 HAR,SB2,STR,TPT

Quantizing dist. 1 CO35 DRU

Quantizing dist. 2 CO37 DRU

Quantizing dist. 3 CO38 DRU

عناصر الاختبار

STR الموسيقى الشعبية السويدية، تسجيل SR، تم استخدامه سابقاً

SB1 مزمار القربة، تسجيل SR

SB2 مزمار القربة، تسجيل SR

CLA الكلارينيت، SQUAM 16/2

TAM الدف الصغير، تسجيل SR، تم استخدامه سابقاً

WIN سترافينسكي، مجموعة الرياح، تم استخدامه سابقاً

TPT البوق، SQUAM 21/2

HAR مجموعة المهرج، تسجيل بي بي سي G 49/17

VEG سوزان فيجا، المعلم القديم، تم استخدامه سابقاً

CAS صنجات، SQUAM 27

SPE كلام بالألمانية، SQUAM 54

RYC راي كودر، الأسطوانة CD: JAZZ tr 11 (0.25 – 0.47)

PER قرع طبول، باس ماريمبا الياباني، الأسطوانة CD: Sony/CBS 32DC 5027

HRN النفير، SQUAM 54

GLO الإكسليفون، SQUAM 35/1، تم استخدامه سابقاً

TRI المثلث الرنان، SQUAM 32/2

DRU الطبول، SQUAM 28

CL2 الكلارينيت، SQUAM 16/2

BAS باس كلارنيت، SQUAM 17

TUB توبا، SQUAM 24

MPE مزيج أحادي القناة من SPE

MTR مزيج أحادي القناة من STR

MLA مزيج أحادي القناة من CLA

MYC مزيج أحادي القناة من RYC



## 2.4 قاعدة البيانات DB3



مسرد المصطلحات

درجة الخطأ المطلق (AES)

تُشتق درجة الخطأ المطلق من معادلة وُضعت خصيصاً لتقييم جودة النتائج المتحصل عليها من طريقة قياس إدراكية موضوعية. وهي تأخذ في الاعتبار فواصل الثقة للقيم المتوسطة لاختبارات الاستماع الشخصية.

جودة الصوت الأساسية

تُعرف جودة الصوت الأساسية بأنها خاصية شخصية عالمية تتضمن أي من وجميع الفروق المكتشفة بين الإشارة المرجعية والصيغة المعالجة منها.

هامش التشفير

هامش التشفير هو معلمة جودة تقيس المساحة العالية لنواقص التشفير غير المسموعة حتى العتبة عندما تصبح هذه النواقص مسموعة.

متغيرات خرج النموذج (MOV)

متغيرات خرج النموذج هي قيم الخرج المتوسطة لطريقة القياس الإدراكي. وتستند هذه المتغيرات إلى النتائج النفسية الصوتية الأساسية، وبالتالي يمكن استخدامها لتحديد خصائص نواقص التشفير بشكل أكبر.

درجة الاختلاف الموضوعية (ODG)

درجة الاختلاف الموضوعية هي معلمة الخرج الرئيسية لطريقة القياس الإدراكي. وهي تقابل الدرجات SDG وهي معلمة القياس تعطي جودة الصوت الأساسية العالمية. ويتراوح نطاق الدرجات ODG بين 0 و4.

القياس خارج الخط

إجراء القياس الذي لا يتفاعل مع إرسال البرنامج الجاري.

القياس على الخط

إجراء القياس الذي يعتمد على إرسال البرنامج الجاري، أو أجزاء منه.

درجة الاختلاف الشخصية (SDG)

في اختبار الاستماع وفقاً للتوصية ITU-R BS.1116، يتم تصنيف الجودة الصوتية الأساسية للمرجع المخفي والنسخة المعالجة من المرجع على سلم انحطاط من خمس درجات. وتُعرف درجة الاختلاف بأنها الدرجة المعطاة للإشارة قيد الاختبار مطروحاً منها الدرجة المعطاة للإشارة المرجعية. وينبغي لنطاق نطاق الدرجات SDG المثالي أن يتراوح بين 0 و4-. إذا لم يُحدد المرجع بشكل صحيح، تكون الدرجة SDG موجبة.

المختصرات

ADB متوسط المجموعة المشوهة *(Average Distorted Block)*

AES درجة الخطأ المطلق *(Absolute Error Score)*

ASD الفرق الطيفي السمعي *(Auditory Spectral Difference)*

Avg متوسط (خطي) *(Average (linear))*

BAQ جودة الصوت الأساسية *(Basic Audio Quality)*

Bw عرض النطاق *(Bandwidth)*

CI فاصل الثقة *(Confidence Interval)*

CM هامش التشفير *(Coding Margin)*

DBn قاعدة البيانات رقم *n* (1 أو 2 أو 3) *(Database n (1, 2 or 3))*

DC تيار مستمر *(Direct Current)*

DFT تحويل فورييه المتمايز *(Discrete Fourier Transform)*

DIX مؤشر الاضطراب *(Disturbance Index)*

DUT الجهاز قيد الاختبار *(Device Under Test)*

EHS البنية التوافقية للخطأ *(Error Harmonic Structure)*

ERB عرض النطاق المستطيل المكافئ *(Equivalent Rectangular Bandwidth)*

fac عامل *(factor)*

FFT تحويل فورييه السريع *(Fast Fourier Transform)*

FIR الاستجابة النبضية المحدودة *(Finite Impulse Response)*

IIR استجابة نبضية غير محدودة *(Infinite Impulse Response)*

ISO المنظمة الدولية للتوحيد القياسي *(International Standards Organization)*

JNLD فرق المستوى الملحوظ بالكاد *(Just Noticeable Level Difference)*

MFPD الحد الأقصى لاحتمال الكشف المصفى *(Maximum Filtered Probability of Detection)*

MOV متغير خرج النظام *(Model Output Variable)*

MPEG فريق خبراء الصور المتحركة *(Moving Picture Expert Group)*

NL جهارة الضوضاء *(Noise Loudness)*

NMR النسبة ضوضاء إلى قناع *(Noise-To-Mask Ratio)*

OASE التقييم الموضوعي للإشارة الصوتية *(Objective Audio Signal Evaluation)*

OCM هامش التشفير الموضوعي *(Objective Coding Margin)*

ODG درجة الاختلاف الموضوعية *(Objective Difference Grade)*

PAQM قياس جودة الصوت الإدراكية *(Perceptual Audio Quality Measure)*

PEAQ القياسات الموضوعية لجودة الصوت المدركة *(Objective Measurements of Perceived Audio Quality)*

PERCEVAL التقييم الإدراكي *(Perceptual Evaluation)*

POM القياس الموضوعي الإدراكي *(Perceptual Objective Measure)*

REF الإشارة المرجعية *(Reference Signal)*

res الدقة *(Resolution)*

r.m.s. جذر متوسط التربيع *(Root Mean Squared)*

ROEX أسي مقرب *(Rounded Exponential)*

ROV معدل قيم الخرج *(Rate of Output Values)*

SCM هامش التشفير الشخصي *(Subjective Coding Margin)*

SDG درجة الاختلاف الشخصية *(Subjective Difference Grade)*

SNR النسبة إشارة إلى ضوضاء *(Signal-to-Noise Ratio)*

SPL مستوى ضغط الصوت *(Sound Pressure Level)*

SUT الإشارة قيد الاختبار *(Signal under Test)*

THD التشوه الكلي الناجم عن التوافقيات *(Total Harmonic Distortion)*

Win متوسط مقسم إلى نوافذ *(Windowed Average)*

المراجع

AURAS, W. [September 1984] *Berechnungsverfahren für den Wohlklang beliebiger Schallsignale, ein Beitrag zur gehörbezogenen Schallanalyse*. Dissertation an der Fakultät für Elektrotechnik der Technischen Universität München, Federal Republic of Germany.

BEERENDS, J. G. and STEMERDINK, J. A. [December, 1992] A perceptual audio quality measure based on a psychoacoustic sound representation. *J.* *Audio Eng.* *Soc*., Vol. 40, p. 963-978.

BEERENDS, J. G. and STEMERDINK, J. A. [February, 1994] Modeling a cognitive aspect in the measurement of the quality of music codecs. Contribution to the 96th AES Convention, preprint 3800. Amsterdam, Netherlands.

BEERENDS, J. G. and STEMERDINK, J. A. [March, 1994] A perceptual speech quality measure based on a psycho-acoustic sound representation. *J.* *Audio Eng.* *Soc*., Vol. 42, p. 115-123.

BEERENDS, J. G., van den BRINK, W. A. C. and RODGER, B. [May, 1996] The role of informational masking and perceptual streaming in the measurement of music codec quality. Contribution to the 100th AES Convention, preprint 4176. Copenhagen, Denmark.

BRANDENBURG, K. [1987] Evaluation of quality for audio encoding at low bit rates. Contribution to the 82nd AES Convention, preprint 2433. London, United Kingdom.

BREGMAN, A. S. [1990] *Auditory Scene Analysis: The Perceptual Organisation of Sound.* MIT Press, Cambridge MA, United States of America.

COHEN, E. A. and FIELDER, L. D. [May, 1992] Determining noise criteria for recording environments. *J.**Audio Eng.* *Soc*., Vol. 40, p. 384-402.

COLOMES, C., LEVER, M., RAULT, J. B. and DEHERY, Y. F. [April, 1995] A perceptual model applied to audio bit-rate reduction. *J.Audio Eng.* *Soc*., Vol. 43, p. 233-240.

FEITEN, B. [March, 1997] Measuring the Coding Margin of Perceptual Codecs with the Difference Signal. 102nd AES-Convention, preprint 4417. Munich, Federal Republic of Germany.

GRUSEC, T., THIBAULT, L. and SOULODRE, G. [September, 1997] EIA/NRSC DAR systems subjective tests. Part 1: Audio codec quality. *IEEE Transactions on Broadcasting*, Vol. 43, **3**.

KARJALAINEN, J. [March, 1985] A new auditory model for the evaluation of sound quality of audio system. Proceedings of the ICASSP, p. 608-611. Tampa, Florida, United States of America.

LEEK, M. R. and WATSON, C. S. [1984] Learning to detect auditory pattern components. *J.**Acoust.* *Soc*. Am., Vol. 76, p. 1037‑1044.

MEARES, D. J. and KIM, S. W. [July, 1995] “NBC time/frequency module subjective tests: overall results”, ISO/IEC JTC 1/SC 29/WG 11 N0973 MPEG95/208.

MOORE, B. C. [1986] *Frequency Selectivity in Hearing*.Academic Press, London, United Kingdom.

MOORE, B. C. [1989] *An introduction to the psychology of hearing*. Academic Press, London, United Kingdom.

PAILLARD, B., MABILLEAU, P., MORISETTE, S. and SOUMAGNE, J. [1992] Perceval: Perceptual evaluation of the quality of audio signals. *J.* *Audio Eng.* *Soc*., Vol. 40, p. 21-31.

SCHROEDER, M. R., ATAL, B. S. and HALL, J. L. [December 1979] Optimizing digital speech coders by exploiting masking properties of the human ear. *J.* *Acoust.* *Soc.* *Am*., Vol. 66, p. 1647‑1652.

SOULODRE, G., GRUSEC, T., LAVOIE, M. and THIBAULT, L. [March 1998] Subjective evaluation of state‑of-the-art 2-channel audio codecs. *J.* *Audio Eng.* *Society*.

SPORER, T. [October 1997] Objective audio signal evaluation*–*applied psychoacoustics for modeling the perceived quality of digital audio. 103rd AES-Convention, preprint 4512. New York, United States of America.

TERHARDT, E. [1979] Calculating Virtual Pitch, Hearing Research. Vol. 1, p. 155-182.

THIEDE, T. and KABOT, E. [1996] A New Perceptual Quality Measure for Bit Rate Reduced Audio. Contribution to the 100th AES Convention, preprint 4280. Copenhagen, Denmark.

TREURNIET, W. C. [1996] Simulation of individual listeners with an auditory model. Proceedings of the Audio Engineering Society, Reprint Number 4154. Copenhagen, Denmark.

von BISMARCK, G. [1974] Sharpness as an attribute of the timbre of steady sounds. *Acustica*, 30, p. 159-172.

ZWICKER, E. and FASTL, H. [1990] *Psycho-acoustics, Facts and Models*. Berlin; Heidelberg: Springer Verlag, Federal Republic of Germany.

ZWICKER, E. and FELDTKELLER, R. [1967] *Das Ohr als Nachrichtenempfänger*. Stuttgart: Hirzel Verlag, Federal Republic of Germany.

بيبليوغرافيا

GRUSEC T., THIBAULT L. and SOULODRE, G. [1995] Subjective evaluation of high quality audio coding systems: methods and results in the two-channel case. Preprint 4065 (F-5), Proceedings of the AES. New York, United States of America.

ــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــــ

1. انظر التوصية ITU-R BS.1116. [↑](#footnote-ref-1)
2. قدم مؤيدو التكنولوجيا الموصوفة في هذه التوصية بيانات براءات اختراع تتوافق مع الملحق 1 بالقرار ITU-R 1. والتكنولوجيا الموصوفة في هذه التوصية محمية ببراءات اختراع دولية، كما هو الحال مع جميع توصيات الاتحاد الدولي للاتصالات الخاضعة لحقوق التأليف والنشر. وتعتبر الموافقة المسبقة للمالكين في شكل ترخيص إلزامية لاستغلال هذه التكنولوجيا. وللحصول على مزيد من المعلومات بشأن ترخيص هذه التكنولوجيا، يرجى الرجوع إلى قاعدة بيانات براءات الاختراع الخاصة بقطاع الاتصالات الراديوية أو إلى أمانة مكتب الاتصالات الراديوية. [↑](#footnote-ref-2)
3. التأخير الإضافي لعينة واحدة ليس ضرورياً للتنفيذ. ويُلاحظ أن التنفيذ المرجعي المستخدم لاختبار الامتثال يتضمن هذا التأخير الإضافي. [↑](#footnote-ref-3)
4. في الواقع، فإن أغلفة المرشحات في نطاقات التردد العليا لا تحقق بالضرورة نظرية الاعتيان. على الرغم من أن الاسترداف لن يحدث إلا في ظل ظروف خاصة جداً (أي مكونات عالية التردد مشكلة بترددات أكبر من kHz 1,5) وأن المشاكل المتعلقة بهذه التأثيرات لم تحدث أبداً داخل قواعد البيانات المعروفة، إلا أنه ينبغي الإشارة إلى أن مشاكل الاسترداف قد تحدث، خاصة عند استخدام إشارات اختبار اصطناعية. [↑](#footnote-ref-4)
5. مصطلح "مجموعة" يعادل "رتل" في هذا السياق. [↑](#footnote-ref-5)
6. \* يشير هذا الرقم إلى بيانات الدخل ذات نسق العدد الصحيح الجبري المكون من 16 بتة بمدى يتراوح بين 32 768– و32 767 على النحو المستخدم في القرص المدمج. [↑](#footnote-ref-6)
7. لتحقيق هذه الدقة ينبغي استخدام حساب النقطة الطليقة لمعهد مهندسي الكهرباء والإلكترونيات. [↑](#footnote-ref-7)
8. تُشتق أسماء العناصر المرجعية المقابلة بالاستعاضة عن السلسلة الفرعية "cod" في أسماء عناصر الاختبار بـالسلسلة "ref"، على سبيل المثال، العنصر المرجعي "bcodtri.wav" إلى "breftri.wav". [↑](#footnote-ref-8)