الاتحاد الدولي للاتصالات
قطاع الاتصالات الراديوية

منشورات ITU
التوصيات

التوصية ITU-R BT.500-15

(2023/05)

السلسلة BT: الخدمة الإذاعية (التلفزيونية)

منهجيات التقدير الشخصاني لنوعية الصور التلفزيونية

تمهيد

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	العنوان
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).

التوصية ITU-R BT.500-15*

منهجيات التقدير الشخصاني لنوعية الصور التلفزيونية[1]

(المسألة ITU-R 102-4/6)

(2023-2019-2012-2009-2002-2000-1998-1998-1995-1994-1992-1990-1986-1982-1978-1974)

مجال التطبيق

تقدم هذه التوصية منهجيات لتقدير نوعية الصورة، بما فيها طرائق الاختبار العامة وسلالم إسناد الدرجات المستعملة أثناء التقديرات وشروط المشاهدة الموصى بها لإجراء التقديرات. وتتكون التوصية من ثلاثة أجزاء.

– يصف الجزء 1 المتطلبات العامة لإجراء تقدير شخصاني للصور التلفزيونية وإرشادات بشأن ظروف استعمال منهجيات معينة.

– يصف الجزء 2 مختلف منهجيات التقدير الموصى بها التي يمكن استعمالها عند إجراء تقديرات شخصانية لنوعية الصورة.

– يصف الجزء 3 منهجيات خاصة بأنساق الصورة وتطبيقاتها استناداً إلى المواصفات الواردة في الجزأين 1 و2.

مصطلحات أساسية

التقدير الشخصاني، تقدير الصورة

إن جمعية الاتصالات الراديوية للاتحاد الدولي للاتصالات،

إذ تضع في اعتبارها

أ) أن كمية كبيرة من المعلومات قد جمعت حول الطرائق المستعملة في مختلف المختبرات لتقدير نوعية الصورة؛

ب) أن تفحص هذه الطرائق يظهر أن ثمة درجة كبيرة من التوافق بين مختلف المختبرات بشأن عدد من جوانب منهجيات الاختبار الشخصانية؛

ج) أن تبنّي منهجيات التقدير المقيسة مهم في تبادل المعلومات بين مختلف المختبرات؛

د) أن تقديرات نوعية الصورة الرتيبة أو التشغيلية و/أو الانحطاطات التي تستعمل سلم نوعية وانحطاط خماسي الدرجات والتي يقوم بها بعض المهندسين المشرفين في أثناء عمليات رتيبة أو عمليات خاصة، قد تستعمل أيضاً بعض جوانب المنهجيات الموصى بها للتقديرات في المختبر؛

هـ) أن الإدخال المستمر لإشارات تلفزيونية جديدة وأنماط جديدة لمعالجة الإشارات أو خدمات تلفزيونية جديدة أو معززة قد يتطلب منهجيات مختلفة لإجراء التقديرات الشخصانية للصورة؛

و) أن إدخال هذه الأنماط من المعالجة والإشارات والخدمات سوف يزيد من احتمالات أن يصبح أداء كل قسم من سلسلة الإشارة أكثر اعتماداً على العمليات المحققة في الأجزاء السابقة من السلسلة،

توصي

1 أن تُستعمل منهجيات الاختبار العامة وسلالم إسناد الدرجات وشروط المشاهدة لتقدير نوعية الصورة التي يصفها الجزء 1، في تجارب المختبرات وفي التقديرات أثناء التشغيل كلما أمكن ذلك؛

---

* أدخلت لجنة الدراسات 6 للاتصالات الراديوية تعديلات صياغية على هذه التوصية في سبتمبر 2023 وفي مايو 2024 وفقاً للقرار ITU-R 1.

[1] تنبغي إحاطة لجنة الدراسات 12 لقطاع تقييس الاتصالات علماً بهذه التوصية.

2 أن تُستعمل، على الرغم من وجود منهجيات بديلة ووضع منهجيات جديدة، المنهجيات التي يصفها الجزء 2 كلما كان مناسباً؛

3 أن تُستعمل المنهجيات العامة للاختبار وسلالم إسناد الدرجات وشروط المشاهدة لتقدير نوعية الصورة في النظام أو التطبيق المعين للصورة الذي يصفه الجزء 3 للتجارب في المختبر وفي تقديرات التشغيل كلما أمكن ذلك؛

4 أن تتبع متطلبات منهجية الاختبار المختارة التي يصفها الجزء 2 من أجل تسهيل تبادل المعلومات بين المختبرات المختلفة؛

5 أن تُعالَج البيانات المجمعة وفقاً للتقنيات الإحصائية المفصلة في الملحق 2 بالجزء 1، من أجل تسهيل تبادل المعلومات بين المختبرات المختلفة؛

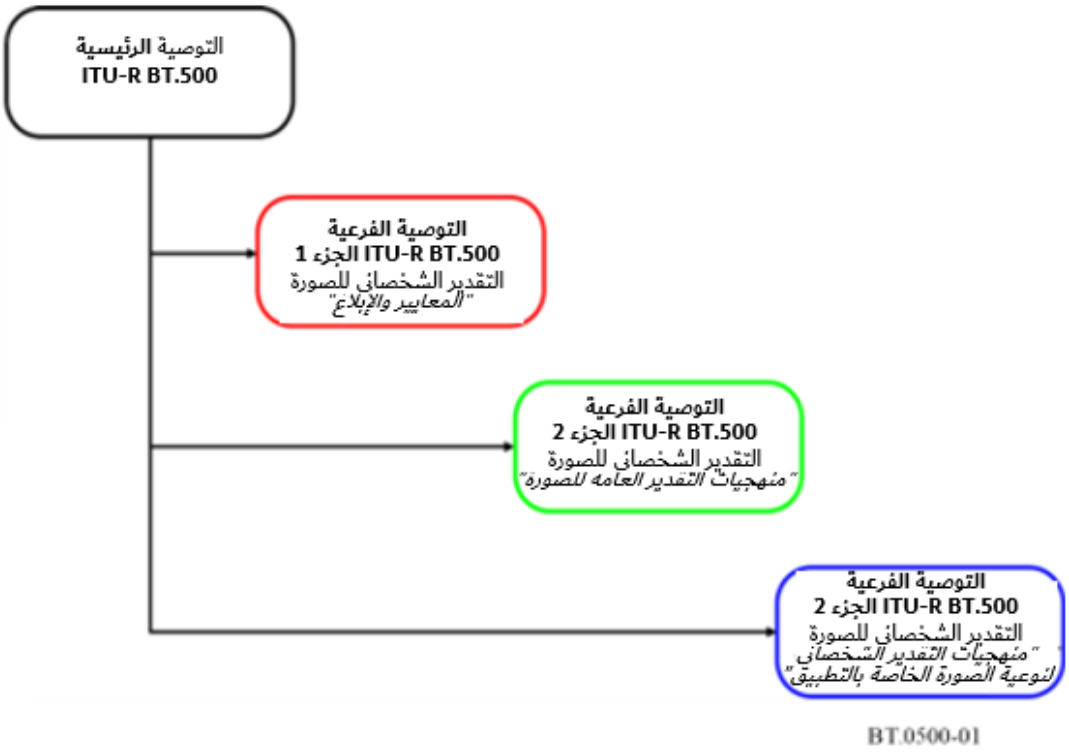
6 أن يتوفر في كل تقارير الاختبارات أشمل وصف ممكن لتشكيلات الاختبارات وموادها وللمراقبين والطرائق، نظراً إلى أهمية تحديد قاعدة التقديرات الشخصانية للصورة.

ملاحظات بشأن هيكل هذه التوصية واستعمالها (إعلامية)

تتألف التوصية ITU-R BT.500 من ثلاثة أجزاء شبه مستقلة تحت هذه التوصية الرئيسية على النحو الموضح في الشكل 1.

الشكل 1

هيكل التوصية ITU-R BT.500



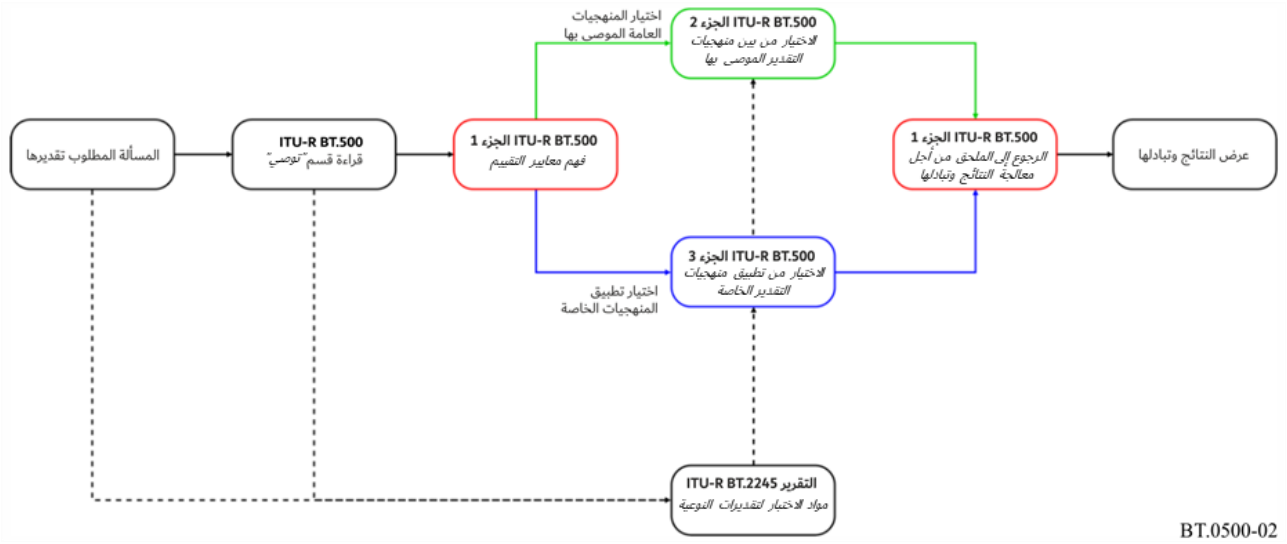
وتُنصح المختبرات التي ترغب في إجراء تقديرات شخصانية للصورة أن ترجع إلى القسم "*توصي*" أعلاه ثم أن تستعمل المعايير الواردة بالتفصيل في الجزء 1 من أجل فهم أنسب منهجية لإجراءات التقدير الخاصة بها. ويعرض الجزء 2 نظرةً عامةً على عدة منهجيات شخصانية لتقدير الصورة يمكن استعمالها بينما يعرض الجزء 3 معلومات عن تطبيق بعض المنهجيات الخاصة الإضافية التي يمكن أن تساعد في إعداد إجراءات التقدير الشخصاني للصورة ذات الصلة.

مشورة بشأن استعمال التوصية ITU-R BT.500

يبين الشكل 2 تدفق عمل محتمل لاستعمال التوصية ITU-R BT.500.

الشكل 2

استعمال التوصية ITU-R BT.500



الأساس المنطقي

تتيح الأجزاء المتعلقة بالهيكل من هذه النسخة من التوصية ITU-R BT.500 إضافة منهجيات جديدة للتقدير الشخصاني للصورة ومراجعة المنهجيات القائمة بدون الحاجة إلى إضافة توصيات جديدة تكرر المعلومات عبر وثائق متعددة أو إصدار مراجعات للأجزاء التي لا تتطلب تغييرات.

التوصيات الأخرى المتعلقة بتقدير الصورة

تتعلق التوصيات التالية بالقياسات الموضوعية لنوعية الصورة التي يمكن أن تتيح منهجيات أخرى لتطبيق تقديرات خاصة للصورة تستعمل بعض معايير التقدير الواردة في التوصية ITU-BT.500.

التوصية ITU-R BT.1683 تقنيات القياس الموضوعي للنوعية الفيديوية المرئية في تلفزيون رقمي عادي الوضوح بوجود المرجع الكامل

التوصية ITU-R BT.1866 تقنيات القياس الموضوعي لجودة الفيديو الحسية من أجل تطبيقات الإذاعة التي تستعمل تلفزيون منخفض الوضوح في وجود مرجع كامل

التوصية ITU-R BT.1867 تقنيات القياس الموضوعي للجودة المرئية الحسية من أجل تطبيقات الإذاعة باستعمال تلفزيون منخفض الوضوح في وجود عرض نطاق مرجعي مخفض

التوصية ITU-R BT.1885 تقنيات القياس الموضوعي للنوعية الفيديوية الحسية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض

التوصية ITU-R BT.1907 تقنيات القياس الموضوعي للجودة الفيديوية المدركة للتطبيقات الإذاعية التي تستعمل التلفزيون عالي الوضوح في وجود إشارة مرجعية كاملة

التوصية ITU-R BT.1908 تقنيات القياس الموضوعي للجودة الفيديوية المدركة للتطبيقات الإذاعية التي تستعمل التلفزيون عالي الوضوح في وجود إشارة مرجعية كاملة

الجزء 1

نظرة عامة على متطلبات التقدير الشخصاني للصورة

الصفحة

الجزء 1 4

نظرة عامة على متطلبات التقدير الشخصاني للصورة 4

1 مقدمة 6

2 الخصائص المشتركة للتقدير 6

1.2 الشروط العامة للمشاهدة 6

1.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المختبرية 6

2.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المنزلية 7

3.1.2 مسافة المشاهدة 7

4.1.2 زاوية المراقبة 9

5.1.2 مجموعة ألوان بيئة القاعة 9

6.1.2 شاشة العرض 9

2.2 إشارات المصدر 10

3.2 انتقاء مواد الاختبار 11

1.3.2 تتابعات اختبارات قطاع الاتصالات الراديوية 11

4.2 مدى الشروط وكيفية إرسائها 12

5.2 المراقبون 12

1.5.2 عدد المراقبين 12

2.5.2 فرز المراقبين 12

3.5.2 تعليمات خاصة بالتقدير 12

6.2 جلسة الاختبار 13

7.2 تقديم النتائج 13

3 انتقاء طرائق الاختبار 14

الملحق 1 بالجزء 1 – تحليل النتائج وتقديمها 14

1-A1 مقدمة 14

2-A1 طرائق التحليل المشتركة 14

1.2-A1 حساب العلامات المتوسطة 14

2.2-A1 حساب فترة الثقة 15

3.2-A1 الفرز اللاحق للمراقبين 16

4.2-A1 حساب متوسط العلامات وفترات الثقة في ظل ظروف الاختبار الصعبة 21

3-A1 العملية التي تسمح بتعيين علاقة بين العلامة المتوسطة والقياس الموضوعي لتشوه الصورة..... 22
1.3-A1 التقريب بواسطة دالة منطقية رياضية متناظرة 22
2.3-A1 التقريب بواسطة دالة لا تناظرية 23
3.3-A1 تصحيح الانحطاط/التحسين المتبقي وآثار حدود سلم التقدير 24
4.3-A1 إدماج جانب الاعتمادية في الصور البيانية..... 25
4-A1 الاستنتاجات..... 26
المرفق 1 بالملحق 1 – التنفيذ المرجعي للطريقة المأخوذة من الفقرة 4.2-A1..... 26
الملحق 2 بالجزء 1 – وصف نسق مشترك لتبادل الملفات..... 31
الملحق 3 (إعلامي) بالجزء 1 – خصائص العطل في محتوى الصورة 33
1-A3 مقدمة..... 33
2-A3 تحديد خاصية العطل..... 33
3-A3 استعمال خاصية العطل 35
الملحق 4 (إعلامي) بالجزء 1 – طريقة تحديد خاصية عطل مركبة بالنسبة إلى محتوى البرنامج وشروط الإرسال 35
1-A4 مقدمة..... 35
2-A4 تحليل محتوى البرنامج..... 36
3-A4 تحليل قناة الإرسال..... 36
4-A4 تحديد خصائص العطل المركب 37
الملحق 5 (إعلامي) بالجزء 1 – الآثار السياقية..... 37
الملحق 6 (إعلامي) بالجزء 1 – مقياسا المعلومات المكانية والزمانية 38
الملحق 7 (إعلامي) بالجزء 1 – مصطلحات وتعاريف 39

1 مقدمة

تُستعمل طرائق التقدير الشخصاني للصورة لتحديد أداء الأنظمة التلفزيونية بواسطة قياسات تأخذ مباشرةً في الحسبان ردات فعل الذين يشاهدون الأنظمة الخاضعة للاختبار. ومن المعروف في هذا الموضوع أن الوسائل الموضوعية قد لا تسمح بتمييز أداء النظام بصورة شاملة؛ ومن ثم يجب إضافة قياسات شخصانية إلى القياسات الموضوعية.

ويوجد عادةً صنفان من التقديرات الشخصانية: أولاً، تقديرات تحدد أداء الأنظمة في الشروط المثلى؛ وتسمى تقديرات النوعية. وثانياً، التقديرات التي تحدد قدرة الأنظمة على المحافظة على النوعية في شروط غير الشروط المثلى للإرسال أو البث؛ وتسمى تقديرات الانحطاط.

ويجب من أجل إجراء أنسب التقديرات الشخصانية، أن تنتقى، أولاً المنهجية الأفضل لشروط معينة وأهداف تقدير الصورة المطلوبة من بين مختلف الخيارات المتاحة.

وللمساعدة في هذا الاختيار، ينبغي مراعاة الخصائص العامة الموضحة في الفقرة 2 من أجل فهم أنسب الخيارات المتعلقة بالمشكلة أو العملية التي يجري تقديرها.

وبعد فهم هذه الخيارات، توفر الفقرة 3 من الجزء 1 نظرةً عامةً على منهجيات تقدير الصورة الموصى بها التي يمكن استعمالها لاختيار أنسب منهجية للمشكلة أو العملية التي يجري تقديرها، مع وضع في الاعتبار نمط أداة التقدير المستعملة وشروط بيئة التقدير.

غير أن اختيار المنهجية الأنسب يتعلق بأهداف الخدمة التي يسعى إلى تحقيقها النظام الخاضع للاختبار. ولهذا تعرض في الجزء 2 وتوصيات أخرى لقطاع الاتصالات الراديوية الإجراءات الكاملة لتقدير تطبيقات محددة.

2 الخصائص المشتركة للتقدير

ترد هنا الشروط العامة للمشاهدة لأغراض التقدير الشخصاني. أما شروط المشاهدة الخاصة للتقديرات الشخصانية لأنظمة محددة فترد في المنهجيات ذات الصلة.

ملاحظة – عند إجراء تقدير شخصاني لصور ذات مدى دينامي واسع، من المستصوب الرجوع إلى الوثائق الأخرى المحال إليها، حيثما وجدت، في القسم ذي الصلة.[2]

1.2 الشروط العامة للمشاهدة

تفترض بيئة المشاهدة المختبرية شروطاً حرجةً للتحقق من الأنظمة. وترد الشروط العامة للمشاهدة للتقديرات الشخصانية في البيئة المختبرية في الفقرة 1.1.2.

ويفترض أن توفر بيئة المشاهدة في المنزل وسائل لتقييم النوعية من جهة مستعمل القنوات التلفزيونية. وترد الشروط العامة للمشاهدة في البيئة المنزلية وسائل لتقييم النوعية من جهة مستعمل القنوات التلفزيونية. وترد الشروط العامة للمشاهدة في البيئة المنزلية في الفقرة 2.1.2. وقد اختيرت هذه المعلمات لتحديد بيئة أكثر حساسيةً بدرجة بسيطة من الحالات العادية للمشاهدة المنزلية.

1.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المختبرية

ينبغي أن تُرتب شروط مشاهدة المقيّمين على النحو التالي:

أ) إضاءة القاعة: ضعيفة

ب) لونية الخلفية: D_{65}

[2] ستجري مراجعة هذه التوصية لإدراج إرشادات إضافية بعد مزيد من العمل بشأن المدى الدينامي الواسع واكتساب خبرة في هذا المجال.

ج)	ذروة النصوع[3]:	250-70 cd/m^2 (انظر الفقرة 5.6.1.2)
د)	نسبة التباين في شاشة العرض:	≥ 0,02 (انظر الفقرة 4.6.1.2)
هـ)	نسبة نصوع الخلفية وراء شاشة عرض الصورة إلى ذروة نصوع الصورة:	≈ 0,15

2.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المنزلية

أ)	شدة إضاءة المكان على الشاشة (ينبغي قياس الضوء الوارد من المكان على الشاشة عمودياً نسبةً إلى الشاشة):	200 lux
ب)	ذروة النصوع:	500-70 cd/m^2 (انظر الفقرة 4.6.1.2)
ج)	نسبة نصوع الشاشة غير النشطة إلى نسبة التباين في ذروة النصوع لشاشة العرض:	≥ 0,02 (انظر الفقرة 4.6.1.2)

3.1.2 مسافة المشاهدة

تستند مسافة المشاهدة إلى حجم الشاشة ويمكن اختيارها وفقاً لمعايرين محددين: مسافة المشاهدة المفضلة (PVD) ومسافة المشاهدة المستهدفة (DVD). وسيعتمد اختيار أحد المعيارين على غرض الدراسة.

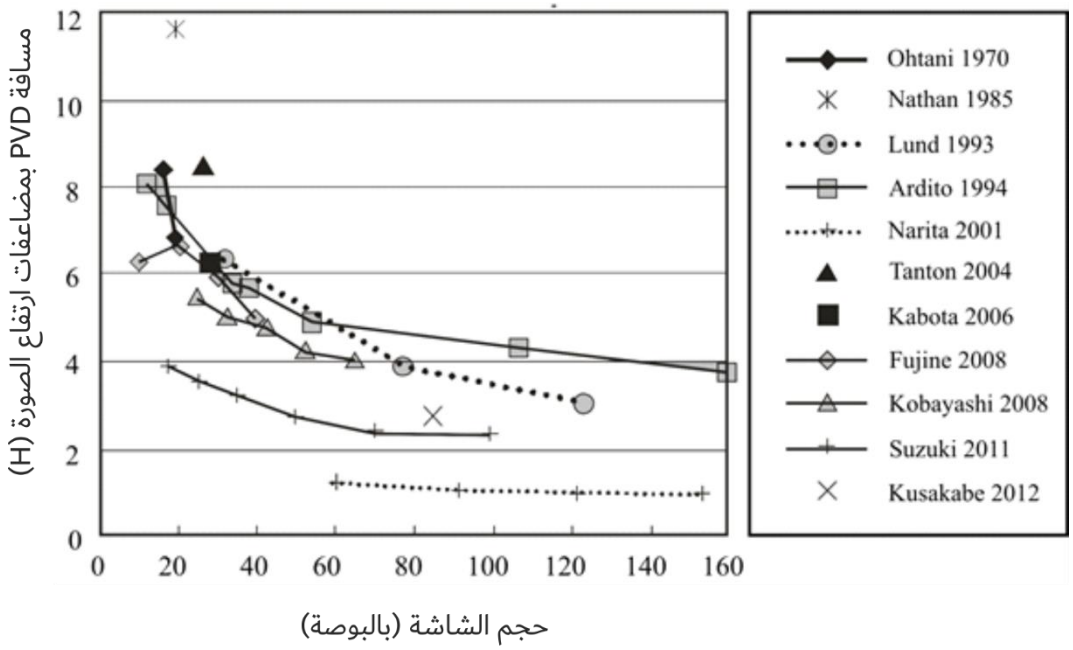
1.3.1.2 مسافة المشاهدة المفضلة

تستند مسافة المشاهدة المفضلة إلى تفضيلات المشاهدين التي تم تحديدها تجريبياً. ويرد في الشكل 1-1 المسافة PVD (بدلالة حجم الشاشة)، حيث يحتوي على عدد من مجموعات البيانات المجمعة من المصادر المتاحة. ويمكن الرجوع إلى هذه المعلومات لتصميم اختبار تقدير شخصاني.

[3] ينبغي ضبط ذروة النصوع وفقاً لإضاءة القاعة.

الشكل 1-1

مسافة المشاهدة المفضلة بدلالة أحجام الشاشات



BT.0500-01-1

2.3.1.2 مسافة المشاهدة المستهدفة

تمثل مسافة المشاهدة المستهدفة (DVD) أو مسافة المشاهدة المثلى فيما يتعلق بنظام رقمي، المسافة التي يقابل عندها بِكْسِلان متجاوران زاوية قدرها 1 قوس–دقيقة عند عين المشاهد؛ وزاوية المشاهدة الأفقية المثلى هي الزاوية التي تُرى الصورة تحتها على مسافة مشاهدتها المثلى.

ويقدم الجدول 1-1 مسافات المشاهدة المثلى (وزوايا المشاهدة الأفقية المثلى) من أجل أنظمة عديدة لاستبانة الصورة يُعبّر عنها بمضاعفات ارتفاع الصورة.

الجدول 1-1

زاوية المشاهدة الأفقية المثلى ومسافة المشاهدة المثلى بمضاعفات ارتفاع الصورة (H)

أنظمة الصورة	المرجع	النسبة الباعية	النسبة الباعية للبكسل	زاوية المشاهدة الأفقية المثلى	مسافة المشاهدة المثلى
720 × 483	التوصية ITU-R BT.601	4:3	0,89	°11	7 *H*
640 × 480	VGA	4:3	1	°11	7 *H*
720 × 576	التوصية ITU-R BT.601	4:3	1,07	°13	6 *H*
1 024 × 768	XGA	4:3	1	°17	4,5 *H*
1 280 × 720	التوصيان ITU-R BT.1543 وITU-R BT.1874	16:9	1	°21	4,8 *H*
1 400 × 1 050	SXGA+	4:3	1	°23	3,3 *H*
1 920 × 1 080	التوصية ITU-R BT.709	16:9	1	°31	3,2 *H*
3 840 × 2 160	التوصية ITU-R BT.2020	16:9	1	°58	1,6 *H*
7 680 × 4 320	التوصية ITU-R BT.2020	16:9	1	°96	0,8 *H*

ملاحظة: عندما ينطوي تقييم الصورة على الاستبانة، ينبغي استعمال القيمة الأقل لمسافة المشاهدة للنسقين 320 4 × 680 7 و160 2 × 840 3. وفي حالة عدم تقييم الاستبانة، يمكن استعمال أي مسافة مشاهدة في المدى (بالنسبة إلى النسق 160 2 × 840 3: ارتفاع صورة قدره 1,6 إلى 3,2؛ وبالنسبة إلى النسق 320 4 × 680 7، ارتفاع صورة قدره 0,8 إلى 2,3).

4.1.2 زاوية المراقبة

ينبغي أن تكون زاوية المراقبة القصوى بالنسبة إلى الزاوية العادية مقيدةً بحيث لا تكون الانحرافات في الألوان المستنسخة على الشاشة مرئيةً للمراقب. وينبغي أيضاً مراعاة زاوية المشاهدة الأفقية المثلى لنظام الصورة قيد الاختبار لتحديد زاوية المراقبة. انظر الفقرة 8.1 من التقرير ITU-R BT.2129 للاطلاع على مزيد من التفاصيل.

5.1.2 مجموعة ألوان بيئة القاعة

ينبغي أن تكون خلفية العرض بنفس لون النقطة البيضاء المرجعية؛ وينبغي استعمال لون داكن غير لامع في سائر أسطح القاعة. والهدف من ذلك هو التقليل إلى أدنى حد من شرود الضوء الواقع على شاشة العرض.

6.1.2 شاشة العرض

يؤدي استعمال شاشة عرض ذات خصائص مختلفة إلى تقديرات شخصانية مختلفة لنوعية الصورة. وبالتالي، يُوصى بشدة التحقق من خصائص شاشات العرض قبل الاستعمال. ويمكن الرجوع إلى التوصية ITU-R BT.1886 – وظيفة التحويل الكهربصري المرجعي من أجل شاشات العرض المسطحة المستعملة في الإنتاج في استوديو التلفزيون عالي الوضوح والتقرير ITU-R BT.2129 – متطلبات المستعمل من أجل شاشة عرض مسطحة لتكون بمثابة مرقاب رئيسي في بيئة إنتاج برنامج تلفزيوني عالي الوضوح، عند استعمال شاشات العرض المهنية ذات الشاشات المسطحة من أجل التقدير الشخصاني.

ويوفر التقرير ITU-R BT.2390 معلومات عن بيئات العرض والمشاهدة المختبرية والمنزلية لتقدير الصور ذات المدى الدينامي الواسع (HDR).

1.6.1.2 معالجة شاشة العرض

ينبغي أن تتم معالجة شاشات العرض كتدرج الصور وتحويل معدل الأرتال ومحسّنات الصورة، في حال تنفيذها، بطريقة تتفادى إدخال الآثار السلبية المصاحبة للصورة. وينبغي أن تكون معالجة المدى الدينامي الواسع مناسبةً لنظام المدى الدينامي الواسع الذي يجري تقديره أو الذي يجري استعماله أثناء التقدير. وبالنسبة إلى تقديرات بيئة المستهلكين أو التوزيع، يمكن أن يشمل ذلك استعمال بيانات شرحية سكونية أو دينامية مناسبة. وينبغي إدراج التفاصيل الكاملة لهذه البيانات الشرحية في ملاحظات التقديرات ليكون بوسع المختبرات الأخرى تكرار التقديرات بدقة.

وعند استعمال شاشات عرض المستهلكين لإجراء تقديرات شخصانية للصورة، من المهم تعطيل جميع خيارات معالجة الصور ما لم يكن أثر معالجة الصورة هذا هو موضع التقدير (التقديرات).

وعند النفاذ إلى الصور المشذرة، ينبغي أن يبيّن تقرير الاختبار ما إذا كان مزيل التشذير مستعملاً أم لا. ويُفضل عدم استعمال مزيل التشذير إذا تسنّى عرض الإشارات المشذرة بدونه.

2.6.1.2 استبانة شاشة العرض

تمتثل استبانة شاشة العرض المتخصصة عادةً للمعايير المطلوبة في التقدير الشخصاني ضمن مدى تشغيل نصوعها.

ويمكن التحقق من الاستبانة القصوى والدنيا (مركز الشاشة وزواياها) وتسجيلها للقيمة المستعملة للنصوع.

وفي حال استعمال شاشات العرض التلفزيونية المسطحة للتقديرات الشخصانية، يوصى بشدة التحقق من الاستبانتين القصوى والدنيا (مركز الشاشة وزواياها) وتسجيلهما بالقيمة المستعملة للنصوع.

والنظام الأكثر عمليةً والمتوفر حالياً لأداء التقدير الشخصاني من أجل التحقق من استبانات شاشات العرض أو التلفزيونات التجارية هو استعمال نموذج اختبار مسموح مولَّد كهربائياً.

3.6.1.2 إعداد شاشات العرض

ينبغي إعداد نصوع وتباين شاشة العرض طبقاً لإضاءة بيئة الاختبار باستعمال أشكال الموجات PLUGE وفقاً للتوصية ITU-R BT.814.

وبالنسبة إلى تقديرات الصورة ذات المدى الدينامي العادي (SDR)، ينبغي قياس نسبة تباين شاشة العرض وفقاً للتوصية ITU-R BT.815. وعند تقدير الصور ذات المدى الدينامي الواسع، ينبغي الرجوع إلى التقرير ITU-R BT.2390.

4.6.1.2 التباين في شاشة العرض

قد تؤثر شدة إضاءة الوسط تأثيراً كبيراً على التباين.

ونادراً ما استعملت شاشات العرض المتخصصة تقنيات لتحسين تباينها في وسط شديد الإضاءة، ولذا فمن المحتمل ألاّ تمتثل لمعايير التباين المطلوبة إن استعملت في وسط شديد الإضاءة.

وتستعمل شاشات عرض المستهلكين عادةً تقنيات للحصول على تباين أفضل في وسط شديد الإضاءة.

5.6.1.2 سطوع شاشة العرض

عند ضبط سطوع شاشة LCD، يفضل استعمال التحكم في شدة الإضاءة الخلفية بدلاً من استعمال تدرج سوية الإشارة للحفاظ على دقة البتة. وفي حالة تكنولوجيات العرض الأخرى التي لا تستعمل الإضاءة الخلفية، ينبغي ضبط سوية اللون الأبيض بوسائل أخرى غير تدرج سوية الإشارة. وجدير بالملاحظة أن جهاز العرض PDP يتحكم في السطوع من خلال عدد إشعاعات الإضاءة وفي حال الضبط على سطوع أقل، يتعرض إنتاج الضوء إلى التدهور.

6.6.1.2 الآثار السلبية الناجمة عن حركة شاشة العرض

ينبغي ألا ينشأ عن شاشة العرض أي آثار سلبية للحركة التي ترتبط ببعض تكنولوجيات أجهزة العرض. ومن جهة أخرى، ينبغي أن تُمثَّل على شاشات عرض المستهلكين آثار الحركة المتضمنة في إشارة الدخل. وعند استعمال شاشات العرض الخاصة بالمستهلكين، من الأهمية بمكان تعطيل **جميع** خيارات معالجة الحركة.

7.6.1.2 المساحات الآمنة لشاشات عرض ذات نسق عريض بنسبة باعية 16:9

ترد في التوصية ITU-R BT.1848 المساحات الآمنة لشاشات عرض ذات نسق عريض بنسبة باعية 16:9.

2.2 إشارات المصدر

توفر إشارة المصدر الصورة المرجعية مباشرةً، ودخل الإشارة للنظام الخاضع للاختبار. ويجب أن تكون نوعيتها مثلى بالنسبة إلى المعيار التلفزيوني المستعمل. ويعتبر غياب العيوب من الجزء المرجعي في زوج الصور المقدم أساسياً للحصول على نتائج مستقرة.

وتشكل الصور الثابتة والتتابعات الفيدوية المخزنة رقمياً الإشارات التي يمكن استعادتها بصورة أفضل، ولهذا تعتبر المصدر المفضل للإشارات. ويمكن تبادلها فيما بين المختبرات، من أجل الحصول على مقارنات أكثر دلالة فيما بين الأنظمة.

ويطلب غالباً أن تراعى طريقة تأثر أداء النظام الخاضع للاختبار، بأية معالجة قد أجريت في مرحلة سابقة من تاريخ الإشارة. ويستحسن لذلك، عندما يصار إلى اختبار على أجزاء السلسلة التي يمكن أن تدخل تشوهات في المعالجة، حتى لو كانت غير مرئية، أن تُسجَّل الإشارة الناتجة تسجيلاً شفافاً، ثم تتيسر لاختبارات أخرى في اتجاه التيار عندما يرغب في التحقق من كيفية التراكم على طول السلسلة للانحطاطات العائدة إلى المعالجة الترادفية. ويجب أن يُحافَظ على هذه التسجيلات في مكتبة مواد الاختبار لاستعمال مستقبلي وفقاً للحاجة، وأن يضاف إليها تقرير مفصل عن تاريخ الإشارة المسجلة. وإذا تطلب الأمر، يمكن أن تكون ماسحات الشرائح التي يبلغ حجمها 35 mm مصدراً للصور الثابتة. وتكون الاستبانة المتاحة مناسبةً لتقدير التلفزيون التقليدي. وقد يعطي قياس اللون ومعلمات الفيلم الأخرى مظهراً ذاتياً مختلفاً لصور آلات التصوير في الاستوديو. وإذا كانت هذه المعلمات

تؤثر في النتائج، فيجب استعمال مصادر الاستوديو المباشرة، مع أن ذلك يبدو غالباً أقل ملاءمةً. ويجب، بصورة عامة، أن تضبط ماسحات الشرائح لكل صورة على حدة، من أجل الحصول على أفضل تقدير شخصاني ممكن، لأن الوضع يكون على هذا النحو في الواقع العملي.

وتُجرى عادةً تقديرات سعة المعالجة اللاحقة بواسطة تقنية كمد اللون. ويكون هذا الأخير حساساً جداً لإضاءة الاستوديو، في حالة التصوير في الاستوديو. ولهذا يفضل أن تتم التقديرات بواسطة زوج شرائح خاصة بكمد اللون تعطي نتائج عالية النوعية. ويمكن إدخال الحركة في الشريحة الأمامية، عند الحاجة.

3.2 انتقاء مواد الاختبار

استُعملت عدة طرائق لتحديد أنواع مواد الاختبار المطلوبة في تقديرات التلفزيون. غير أنه ينبغي، في التطبيق العملي، أن تُستعمل أنماط خاصة من المواد لمعالجة مشاكل التقدير الخاصة. ويقدم الجدول 1-2 عرضاً لمشاكل التقدير النمطية ومواد الاختبار المستعملة لمعالجة هذه المشاكل.

الجدول 1-2

انتقاء مواد الاختبار*

مشكلة التقدير	المواد المستعملة
الأداء الإجمالي مع مواد متوسطة	عامة، "حرجة دون إفراط"
السعة، التطبيقات الحرجة (مثل المساهمة والمعالجة اللاحقة، إلخ.)	مدى واسع، بما في ذلك المواد الحرجة جداً للتطبيق الخاضع للاختبار
أداء الأنظمة "التكييفية"	مواد حرجة جداً للمخطط "التكييفي" المستعمل
التعرف إلى نقاط الضعف والتحسينات الممكنة	مواد حرجة مميزة للنعت المعني
التعرف إلى العوامل التي تميز الأنظمة المختلفة	مدى واسع من المواد المعقدة جداً
تحويل المعايير المختلفة	حرجة بالنسبة إلى الفروقات (مثل تردد المجال)

* من الواضح أن كل مواد الاختبار قد تكون جزءاً من محتوى البرامج التلفزيونية. انظر الملحقين 3 و4 من أجل معلومات أخرى عن انتقاء مواد الاختبار.

يمكن أن تؤدي بعض المعلمات إلى انحطاطات مماثلة لمعظم الصور أو التتابعات. ويمكن، في هذه الحالة، أن توفر النتائج التي يحصل عليها من خلال عدد صغير جداً من الصور أو التتابعات (اثنين مثلاً)، تقديراً دلالياً.

غير أن غالباً ما يتعلق تأثير الأنظمة الجديدة بمحتوى المشهد أو التتابع تعلقاً شديداً. ويظهر، في هذه الحالة، في أثناء مجمل ساعات البرنامج توزيع إحصائي لاحتمالات الانحطاط ولمحتوى الصور أو التتابعات. ونظراً إلى أن شكل هذا التوزيع غير معروف في العادة، يجب أن يتم انتقاء مواد الاختبار وتأويل النتائج بحذر شديد.

من الهام، بصورة عامة، أن تدرج مواد حرجة، نظراً لإمكانية أخذ ذلك في الاعتبار عند تأويل النتائج، لكن من غير الممكن استخلاص النتائج انطلاقاً من مواد غير حرجة. ويجب في الحالة التي يؤثر فيها محتوى المشهد أو التتابع في النتائج، أن يتم انتقاء المواد على نحو تكون فيه "حرجة دون إفراط" بالنسبة إلى النظام الخاضع للاختبار. وتعني عبارة "دون إفراط" أن الصور يمكن أن تشكل مبدئياً جزءاً من ساعات البرنامج العادي. ويجب أن تُستعمل، في هذه الحالات، أربعة بنود على الأقل: فيكون نصفها مثلاً حرجاً فعلاً، والنصف الآخر حرجاً بعض الشيء.

1.3.2 تتابعات اختبارات قطاع الاتصالات الراديوية

وضع عدد من المنظمات صوراً وتتابعات ثابتة للاختبار. ويعرض التقرير ITU-R BT.2245 – مواد اختبار التلفزيون عالي الوضوح وفائق الوضوح بما في ذلك التلفزيون ذو المدى الدينامي العالي لتقييم جودة الصورة – معلومات مفصلة عن مواد اختبار التلفزيون عالي الوضوح وفائق الوضوح التي يمكن استعمالها للتقدير الشخصاني. ويقدم الملحقان 1 و2 بالجزء 1 من هذه التوصية معلومات أخرى عن انتقاء مواد الاختبار.

4.2 مدى الشروط وكيفية إرسائها

نظراً إلى أن أغلب طرائق التقدير حساسة للتغيرات في مدى شروط المشاهدة وتوزيعها، يجب أن تتضمن جلسات التقدير الشخصاني المديات الكاملة للعوامل المتغيرة. غير أنه يمكن تقريب ذلك بواسطة مدى أكثر تقييداً، فتقدم أيضاً بعض الشروط التي قد تقع عند أطراف السلالم. ويمكن تمثيلها على أنها أمثلة والتعرف إليها كحالات قصوى (إرساء مباشر) أو توزيعها على كل الدورة وعدم التعرف إليها كحالات قصوى (إرساء غير مباشر).

5.2 المراقبون

يمكن أن يكون المراقبون خبراء أو غير خبراء وفقاً لأهداف التقييم. فالمراقب الخبير هو من لديه خبرة في مصنوعات الصور التي يمكن أن يطالعه بها النظام تحت الاختبار. أما المراقب غير الخبير ("الغر") فهو من ليس لديه هذه الخبرة. على أي حال، ينبغي ألا يكون المراقبون معنيين مباشرةً حاضراً أو فيما مضى، أي بما يكفي لاكتساب معرفة محددة ومفصلة في تطوير نظام قيد الدراسة.

1.5.2 عدد المراقبين

ما لم تنص المنهجية المختارة على خلاف ذلك، ينبغي استعمال 15 مراقباً على الأقل. ويتعلق العدد اللازم من المقيّمين بحساسية إجراء الاختبار واعتماديته وبالحجم المتوقع للتأثير المقدر. وفي الدراسات محدودة النطاق ذات الطبيعة الاستكشافية مثلاً، يمكن استعمال أقل من 15 مراقباً. وفي هذه الحالة، ينبغي تحديد هذه الدراسة على أنها "غير رسمية". وينبغي الإبلاغ عن مستوى خبرة المراقبين المشاركين في تقييم جودة الصورة التلفزيونية.

2.5.2 فرز المراقبين

بشكل عام، قبل بدء الجلسة، يختار المراقبون وفقاً لحدة بصرهم العادية (أو المحولة إلى العادية) بواسطة مخطط اختبار "سنلين" (Snellen) أو "لندولت" (Landolt)، ووفقاً لرؤيتهم العادية للألوان بواسطة مخططات تنتقى خصيصاً لهذا الغرض (مخطط إشيهارا (Ishihara)، مثلاً).

يوضح القسمان 3.2-A1 و4.2-A1 تفاصيل سيناريوهات فرز المراقبين المختلفة التي يمكن تطبيقها على منهجيات الاختبار المختلفة. عندما يتم إجراء اختبارات معملية أو أقل رسمية كجزء من برنامج اختبار متعدد المواقع أو لمنظمة، من المهم تبادل التفاصيل الكاملة لطريقة ومعايير فرز المراقبين وإدراجها كجزء من النتائج المنشورة.

وبشكل عام، أكبر قدر ممكن من التفاصيل حول خصائص لجان التقدير الخاصة بهم والتي يمكن أن تشمل فئة المهنة (على سبيل المثال، موظف في الهيئة الإذاعية، طالب جامعي، عامل مكتب، ...) والجنس والفئة العمرية.

ملاحظة – أظهرت دراسة أجريت للتحقق من اتساق النتائج التي حصلت عليها مختبرات مختلفة احتمال وجود تباين منهجي بين هذه النتائج. ويصبح هذا التباين شديد للغاية إذا ما جمعت عدة نتائج من عدة مختبرات مختلفة بهدف تحسين حساسية واعتمادية تجربة ما.

ويمكن عزو هذا التباين بين المختبرات إلى اختلاف مستوى مهارة مجموعات المقيّمين. ولا بد من إجراء مزيد من البحث للتأكد من صلاحية هذه الفرضية، وإن ثبت التقدير التغيرات الناجمة عن هذا العامل. وفي الانتظار ينبغي للمختبرين أن يوفروا أكثر معلومات ممكنة عن المجموعات التي تقوم بعمليات التقييم بغية التمكن من استكمال دراسة هذا العامل.

3.5.2 تعليمات خاصة بالتقدير

يجب أن تقدم بعناية إلى المقيّمين طريقة التقدير وأنماط الانحطاط أو عوامل النوعية المتوقع حدوثها وسلم إسناد الدرجات والتتابع والتوقيت. وتُستعمل تتابعات التمرينات تبين مدى الانحطاطات الواجب تقديرها وأنماطها مع صور توضيحية غير الصور المستعملة في الاختبار، لكن ذات حساسية مماثلة. ويمكن، في حالة تقديرات النوعية، أن تعرف هذه الأخيرة بنعوتها المدركة المميزة.

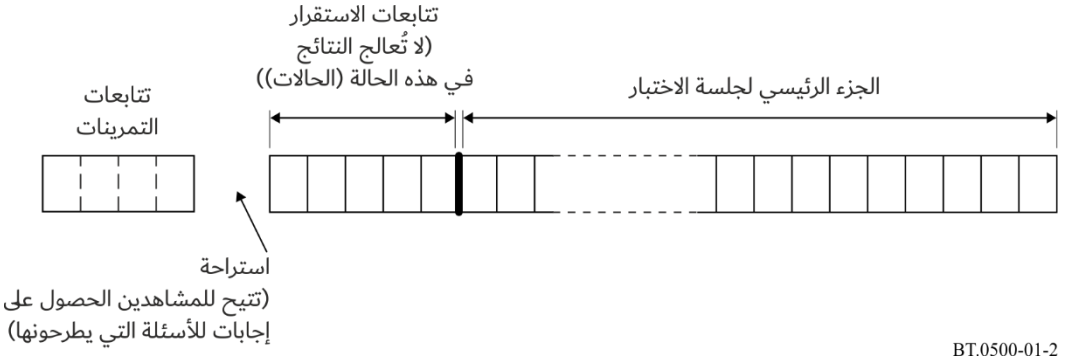
6.2 جلسة الاختبار

يجب ألا تتجاوز مدة الجلسة الواحدة نصف الساعة. تقدم بداية الجلسة الأولى، خمسة "عروض زائفة" تقريباً من أجل تثبيت آراء المراقبين. ويجب عدم مراعاة البيانات الصادرة عن هذه التقديمات في نتائج الاختبار. وإذا تطلب الأمر اللجوء إلى عدة جلسات، تكفي ثلاث عروض زائفة فقط عند بداية الجلسة التالية.

يستعمل ترتيب عشوائي للعروض (مستخلص من المربعات اليونانية–اللاتينية، على سبيل المثال)؛ ويجب أن ترتب شروط الاختبار على نحو يوازن جلسة بعد جلسة تأثيرات التعب أو التكيف في إسناد الدرجات. ويمكن تكرار بعض العروض من جلسة إلى جلسة من أجل التحقق من التماسك.

الشكل 1-2

بنية تقديم جلسة الاختبار



7.2 تقديم النتائج

لما كانت الأحكام تتغير وفقاً للمدى، فمن غير المناسب أن يصار إلى تأويل الأحكام الصادرة عن أكثرية طرائق التقدير بصورة مطلقة (مثل نوعية صورة أو تتابع من الصور).

يجب أن يعطى لكل معلمة من الاختبار متوسط التوزيع الإحصائي لدرجات التقدير وفترة الثقة البالغة 95%. فإذا تعلق التقدير بتغيير الانحطاط وفقاً لتغير قيمة المعلمة، يجب أن تُستعمل تقنيات ملاءمة المنحني. وتسمح ملاءمة المنحني والإحداثيات اللوغاريتمية برسم خط مستقيم، وهو الشكل المفضل للعرض. ويقدم الملحق 1 من الجزء 1 بهذه التوصية معلومات إضافية عن معالجة البيانات.

يجب أن تُقدَّم النتائج مع المعلومات التالية:

- تفاصيل عن تشكيلة الاختبار؛
- تفاصيل عن مواد الاختبار؛
- نمط مصدر الصورة وشاشات العرض (الملاحظة 1)؛
- عدد المقيّمين وأنماطهم (الملاحظة 2)؛
- الأنظمة المرجعية المستعملة؛
- متوسط العلامات العام للتجربة؛
- متوسط العلامات الأصلي والمصحح وفترة الثقة إلى 95% إذا ما سحب مراقب واحد أو عدة مراقبين وفقاً للإجراء المقدم أدناه.

الملاحظة 1 – نظراً لأنه من الواضح أن يؤثر حجم شاشة العرض على نتائج التقدير الشخصاني، يُطلب من المختبرين بيان حجم الشاشة وشكلها وعدد النماذج المستعملة في كل من الاختبارات.

الملاحظة 2 – من الطبيعي أن تؤثر تغييرات مستوى المهارة على مجموعة مشاهدين (حتى عند المشاهدين غير الخبراء) على نتائج التقدير الشخصاني للمشاهدة. ويطلب من المختبرين بغية تسهيل الدراسات الأخرى عن هذا المعامل أن يسجلوا قدر الإمكان السمات المميزة لمجموعات المشاهدة هذه. مثل: العمر والجنس ومستوى التعليم والمهنة.

3 انتقاء طرائق الاختبار

لقد استُعمل للتقديرات التلفزيونية مدى واسع جداً من طرائق الاختبار الأساسية. غير أنه ينبغي، في التطبيق العملي، أن تُستعمل طرائق خاصة لمعالجة مشاكل التقدير الخاصة. ويقدم الجزء 3 من هذه التوصية إرشادات للتقدير الشخصاني لنوعية الصورة في أنساق وتطبيقات الصورة ذات الصلة.

الملحق 1
بالجزء 1

تحليل النتائج وتقديمها

1-A1 مقدمة

تُجمع كمية كبيرة من البيانات في أثناء التجارب الشخصانية لتقدير أداء نظام تلفزيوني. وهذه البيانات تُقدَّم على شكل وريقات علامات المراقبين أو مكافئها الإلكتروني ويجب أن يتم تكثيفها بواسطة تقنيات إحصائية بحيث تُمثَّل النتائج بشكل بياني و/أو شكل عددي/صيغ/خوارزمية الذي يلخص أداء الأنظمة الخاضعة للاختبار.

ويطبَّق التحليل التالي على نتائج سلم الانحطاط ثنائي الحافز (DSIS) وسلم النوعية المستمرة ثنائي الحافز (DSCQ) لطريقتي تقدير نوعية الصورة التلفزيونية والتي ترد في الملحقات 1 و2 و3 بالجزء 2 من هذه التوصية وعلى الطرائق الأخرى البديلة التي تستعمل السلالم الرقمية. ويوزع الانحطاط، في الحالتين الأولى والثانية، على سلم خماسي أو متعدد الدرجات بينما يستعمل في الحالة الأخيرة سلم التقدير المستمر وتُقيَّس النتائج (الفرق بين تقديرات الصورة المرجعية والصورة الحالية الخاضعة للاختبار) عند قيم صحيحة تتراوح بين 0 و100.

2-A1 طرائق التحليل المشتركة

تؤدي الاختبارات المحققة وفقاً لمبادئ الطريقتين المذكورتين في الفقرة 2 من الجزء 1 إلى توزيع لقيم صحيحة تتراوح بين 1 و5 أو بين 0 و100. ويتضمن هذا التوزيع التباين بين أحكام المراقبين وتأثير مختلف شروط التجربة، من قبيل استعمال عدة صور أو متتابعات.

ويتضمن الاختبار عدداً من العروض L. ويمثل كل عرض، عدداً من شروط الاختبار في j، المطبقة على تتابع واحد من تتابعات/صور الاختبار K. وفي بعض الحالات، قد تكرر كل تشكيلة في تتابعات/صور الاختبار مع شروط الاختبار عدة مرات، R.

1.2-A1 حساب العلامات المتوسطة

تكمن المرحلة الأولى لتحليل النتائج في حساب العلامة المتوسطة، $\bar{u}jkr$، لكل حالة اختبار:

$$\bar{u}_{jkr} = \frac{1}{N} \sum_{i=1}^{N} u_{ijkr} \qquad (1)$$

حيث:

u_{ijkr}: علامة المراقب i المعطاة لشرط الاختبار في j، والتتابع/الصورة k، في المرات r

N: عدد المراقبين.

وبطريقة مماثلة يمكن حساب إجمالي متوسط الدرجات $\bar{u}_j$ و $\bar{u}_k$، لكل شرط اختبار وكل تتابع/صورة اختبار.

2.2-A1 حساب فترة الثقة

1.2.2-A1 معالجة البيانات الخام (التي لم تخضع لأي عملية تعويض و/أو تقريب)

عند تقديم نتائج اختبار ما ينبغي أن يكون لجميع العلامات المتوسطة فاصل ثقة يشتق من الانحراف المعياري لكل عينة وحجمها. يقترح استعمال فترة ثقة من 95% تعطيها العبارة:

$$\left[\bar{u}_{jkr} - \delta_{jkr}\,,\ \bar{u}_{jkr} + \delta_{jkr}\right] \tag{2}$$

حيث

$$\delta_{jkr} = 1.96\,\frac{S_{jkr}}{\sqrt{N}} \tag{3}$$

يحسب الانحراف المعياري لكل عرض S_{jkr}، بواسطة العبارة التالية:

$$S_{jkr} = \sqrt{\sum_{i=1}^{N} \frac{(\bar{u}_{jkr} - u_{ijkr})^2}{(N-1)}} \tag{4}$$

إذا كانت نسبة الاحتمال 95%، تكون القيمة المطلقة للفرق، بين العلامة المتوسطة في التجربة والعلامة المتوسطة "الحقيقية" (لعدد كبير جداً من المراقبين) أصغر من 95% من فترة الثقة، شريطة أن يستجيب توزيع العلامات الفردية لبعض المتطلبات.

وبطريقة مماثلة يمكن حساب الانحراف المعياري S_j، لكل شرط اختبار. لكن يلاحظ أن هذا الانحراف المعياري يتأثر، عندما تكون تتابعات/صور الاختبار قليلة، بالفروق بين تتابعات الاختبار المستعملة أكثر من تأثره بالفروق بين المقيّمين المشاركين في التقدير.

2.2.2-A1 معالجة البيانات التي خضعت لعملية تعويض و/أو تقريب

فيما يتعلق بالبيانات التي خضعت آثار الانحطاطات/التحسينات المتبقية أو آثار حدود سلم التقدير فيها لعمليات تعويض أو فيما يتعلق بالبيانات المقدمة على شكل ردود على الانحطاطات أو قاعدة إضافة الانحطاطات بعد التقريب، (بسبب ارتباط متوسط علامات متوسط النوعية الناتجة عن التجارب بهذه التشوهات) ينبغي حساب فترة الثقة باستعمال متحولات متغيرات إحصائية مع مراعاة تشتت قيم هذه المتغيرات.

وإذا عُرضت نتائج تقدير النوعية على شكل ردود فعل على انحطاط ما (أي منحني تجريبي) يكون حدا الثقة الأعلى والأدنى لفترة الثقة هما دالة كل قيمة تجريبية. ومن أجل حساب هذين الحدين ينبغي حساب الانحراف المعياري مع تقييم تقريبي لارتباطه وذلك لكل قيمة تجريبية من رد الفعل الأولي على الانحطاط.

3.2-A1 الفرز اللاحق للمراقبين

1.3.2-A1 الفرز اللاحق القائم على التفرطح للطريقتين DSIS وDSCQS والطرائق البديلة باستثناء الطريقة SSCQE

أولاً، يجب التأكد من أن توزيع العلامات لكل عرض عادي أو غير عادي بواسطة الاختبار β_2 (من خلال حساب معامل التفرطح في الدالة، أي نسبة العزم من الرتبة الرابعة إلى تربيع العزم من الرتبة الثانية). وإذا كانت β_2 بين 2 و4، فيمكن اعتبار التوزيع عادياً. ولكل عرض ويجب عندها أن تقارن العلامات u_{ijkr} لكل مراقب مع القيمة المتوسطة $\bar{u}jkr$ المصاحبة زائداً الانحراف المعياري المصاحب، $Sjkr$، مضروب بـ 2 (عادي) أو بـ $\sqrt{20}$ (غير عادي)، P_{jkr}، ومع القيمة المتوسطة المصاحبة ناقصاً الانحراف المعياري نفسه مضروباً في 2 أو في $\sqrt{20}$، Q_{jkr}. ويجب، في كل مرة تكون علامة المراقب خارج هذا المدى، أن تسجل على عداد مصاحب لكل مراقب. ويستعمل عدادان منفصلان للقيم فوق (P_{jkr}) وتحت (Q_{jkr}). وأخيراً، يجب أن تحسب النسبتان $P_i + Q_i$ إلى عدد العلامات الإجمالي لكل مشاهد في أثناء الدورة الكاملة، و$P_i - Q_i$ إلى $P_i + Q_i$ معبراً عنها بالقيمة المطلقة. فإذا كانت النسبة الأولى أكبر من 5% والنسبة الأخيرة أقل من 30%، يجب أن يُستبعد المراقب i، (انظر الملاحظة).

ملاحظة – لا يطبق هذا الإجراء أكثر من مرة واحدة على نتائج تجربة معينة. ويجب، إضافةً إلى ذلك، أن ينحصر استعمال الإجراء في الحالات التي يوجد فيها عدد قليل نسبياً من المراقبين (أقل من 20، مثلاً)، جميعهم غير اختصاصي.

يوصى بهذا الإجراء لطريقة EBU (DSIS)؛ وقد نجح أيضاً استعماله في طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) وفي الطرائق البديلة.

ويمكن التعبير رياضياً عن الإجراء المذكور آنفاً على النحو التالي:

يحسب لكل عرض اختبار المتوسط $\bar{u}jkr$، الانحراف المعياري $Sjkr$، ومعامل التفرطح، β_{2jkr} حيث يعطي القيمة β_{2jkr} بالمعادلة التالية:

$$\beta_{2jkr} = \frac{m_4}{(m_2)^2} \quad \text{with} \quad m_x = \frac{\sum_{i=1}^{N}(u_{ijkr} - \bar{u}_{ijkr})^x}{N} \qquad (5)$$

وتحسب لكل مراقب، i، القيمتان P_i وQ_i، أي:

بالنسبة إلى j، k، $r = 1$، 1، 1 إلى J، K، R

إذا $4 \geq \beta_{2jkr} \geq 2$، يكون:

إذا $u_{ijkr} \geq \bar{u}_{jkr} + 2\,S_{jkr}$ يكون $P_i = P_i + 1$

إذا $u_{ijkr} \leq \bar{u}_{jkr} - 2\,S_{jkr}$ يكون $Q_i = Q_i + 1$

غير ذلك:

إذا $u_{ijkr} \geq \bar{u}_{jkr} + \sqrt{20}\,S_{jkr}$ يكون $P_i = P_i + 1$

إذا $u_{ijkr} \leq \bar{u}_{jkr} - \sqrt{20}\,S_{jkr}$ يكون $Q_i = Q_i + 1$

فإذا كان $0{,}05 < \frac{P_i + Q_i}{J \cdot K \cdot R}$ و $0{,}3 > \left|\frac{P_i - Q_i}{P_i + Q_i}\right|$ إذن يتم رفض المراقب i.

حيث:

N: عدد المراقبين

J: عدد شروط الاختبار بما في ذلك المرجع

K: عدد صور أو تتابعات الاختبار

R: عدد مرات التكرار

L: عدد عروض الاختبار (في معظم الحالات يكون عدد العروض مساوياً J، K، R، لكن يلاحظ أن بعض التقديرات قد تفضي إلى أعداد غير متساوية من التتابعات لكل شرط اختبار).

2.3.2-A1 الفرز اللاحق القائم على التفرطح لطريقة SSCQE

فيما يتعلق بفرز المراقبين عند استعمال إجراء الاختبار بطريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE)، لم يعد مجال التطبيق واحداً من تشكيلات الاختبار (جمع شرط اختبار مع تتابع اختبار) لكن نافذة زمنية (مثال مقطع تقدير طوله 10 ثوان) من تشكيلة اختبار. وهنالك ترشيح بمرحلتين، الأولى ترشيح لكشف واستبعاد المراقبين الذين يظهرون فرقاً شديداً في تقديراتهم نسبةً إلى المنحى العام، والثانية من أجل كشف وفرز المراقبين الذين يظهرون عدم اتساق دون مراعاة أي فروق منهجية.

المرحلة 1: كشف قلب التقدير المحلي

هنا أيضاً يجب التحقق أولاً مما إذا كان توزيع العلامات هذا الخاص بكل نافذة زمنية لكل تشكيلة اختبار "عادياً" أم لا وذلك باستعمال الاختبار β_2. وإذا تراوحت القيمة β_2 بين 2 و4، يمكن اعتبار التوزيع عادياً. ويطبق الإجراء عندئذٍ على كل نافذة زمنية لكل تشكيلة اختبار وفقاً للمعادلة الرياضية الواردة لاحقاً.

ويحسب لكل نافذة زمنية لكل تشكيلة اختبار وباستعمال التقديرات u_{ijkr} لكل مراقب، المتوسط، $\bar{u}_{jklr}$، والانحراف المعياري S_{jklr}، والمعامل β_{2jklr}. وتعطى القيمة β_{2jklr} في المعادلة:

$$\beta_{2jklr} = \frac{m_4}{(m_2)^2} \quad \text{with} \quad m_x = \frac{\sum_{n=1}^{N}(u_{njklr} - \bar{u})^x}{N} \tag{6}$$

وتحسب لكل مراقب i، القيمتان P_i وQ_i، مثال:

بالنسبة إلى j، k، l، $r = 1$، 1، 1، 1 إلى J، K، L، R

إذا $2 \le \beta_{2jklr} \le 4$، يكون:

إذا $u_{njklr} \ge \bar{u}_{jklr} + 2\, S_{jklr}$ يكون $P_i = P_i + 1$

إذا $u_{njklr} \le \bar{u}_{jklr} - 2\, S_{jklr}$ يكون $Q_i = Q_i + 1$

غير ذلك:

إذا $u_{njklr} \ge \bar{u}_{jklr} + \sqrt{20}\, S_{jklr}$ يكون $P_i = P_i + 1$

إذا $u_{njklr} \le \bar{u}_{jklr} - \sqrt{20}\, S_{jklr}$ يكون $Q_i = Q_i + 1$

إذا كان $X < \frac{P_i}{J \cdot K \cdot L \cdot R}$ أو $X < \frac{Q_i}{J \cdot K \cdot L \cdot R}$ إذن يتم رفض المراقب i.

حيث:

N: عدد المراقبين

J: عدد النوافذ الزمنية مع تجميع شرط الاختبار وتتابعه

K: عدد شروط الاختبار

L: عدد التتابعات

R: عدد مرات التكرار.

وتتيح هذه العملية استبعاد المراقبين الذين قدموا تقديرات شديدة التخالف مع متوسط العلامات. ويبين الشكل 1-3 مثالين (منحنيان بيانيان لطرفي التخالف الشديد). لكن معيار الاستبعاد هذا لا يتيح كشف عمليات القلب المحتملة التي تشكل مصدراً هاماً للتباين. ولهذا السبب تُقترح مرحلة العملية الثانية.

المرحلة 2: كشف قلب التقدير محلياً

في المرحلة 2 يستند الكشف أيضاً على معادلة الفرز الواردة في هذا الملحق. وتضاف فيها تعديل طفيف يتعلق بمجال التطبيق. وتتكون مجموعة البيانات الداخلة أيضاً من علامات جميع النوافذ الزمنية (مثال 10 ثوان) لجميع تشكيلات الاختبار. لكن العلامات هذه المرة مركزة مبدئياً حول المتوسط العام من أجل تخفيف أثر التباين الذي سبق وعولج في المرحلة الأولى من العملية فقد استعملت العملية العادية.

ويجب التحقق أولاً مما إذا كان توزيع العلامات هذا لكل نافذة زمنية لكل تشكيلة اختبار "توزيعاً طبيعياً"، أم لا باستعمال اختبار $\beta 2$. فإذا كانت قيمة $\beta 2$ تقع بين 2 و4، اعتبر التوزيع طبيعياً. وعندئذٍ يستعمل الإجراء لكل نافذة زمنية لكل تشكيلة اختبار وفق المعادلة الرياضية الواردة أدناه.

والمرحلة الأولى من العملية هي حساب العلامات المتوسط لكل نافذة زمنية ولكل مراقب. ويتحدد متوسط العلامة، $\overline{u}_{klr}$، لكل من تشكيلات الاختبار على النحو التالي:

$$\overline{u}_{klr} = \frac{1}{N} \cdot \frac{1}{J} \sum_{n=1}^{N} \sum_{j=1}^{J} u_{njklr} \qquad (7)$$

وبطريقة مماثلة يتحدد متوسط علامة كل من تشكيلات الاختبار وكل مراقب على النحو التالي:

$$\overline{u}_{nklr} = \frac{1}{J} \sum_{j=1}^{J} u_{njklr} \qquad (8)$$

وتعادل القيمة u_{njklr} علامة المراقب i لنافذة زمنية j، وشرط اختبار k، وتتابع l، وعدد مرات r.

وتحسب العلامات المتوسطة u^*_{njklr} لكل مراقب على النحو التالي:

$$u^*_{njklr} = u_{njklr} - \overline{u}_{nklr} + \overline{u}_{klr} \qquad (9)$$

وتحسب لكل نافذة زمنية لكل تشكيلة اختبار المتوسط $\overline{u}^*_{jklr}$ والانحراف المعياري، S^*_{jklr} والمعامل $\beta 2^*_{jklr}$. وتحسب القيمة $\beta 2^*_{jklr}$ في المعادلة:

$$\beta_{2^*jklr} = \frac{m_4}{(m_2)^2} \quad \text{with} \quad m_x = \frac{\sum_{n=1}^{N} (u^*_{njklr})^x}{N} \qquad (10)$$

وتحسب القيمتان P^*_i وQ^*_i، لكل مراقب i كالتالي:

بالنسبة إلى j، k، l، $r = 1$، 1، 1، 1 إلى J، K، L، R

إذا $2 \le \beta_2{}^*_{jklr} \le 4$، يكون:

إذا $u^*_{njklr} \ge \bar{u}^*_{jklr} + 2\,S^*_{jklr}$ يكون $P^*_i = P^*_i + 1$

إذا $u^*_{njklr} \le \bar{u}^*_{jklr} - 2\,S^*_{jklr}$ يكون $Q^*_i = Q^*_i + 1$

غير ذلك:

إذا $u^*_{njklr} \ge \bar{u}^*_{jklr} + \sqrt{20}\,S^*_{jklr}$ يكون $P^*_i = P^*_i + 1$

إذا $u^*_{njklr} \le \bar{u}^*_{jklr} - \sqrt{20}\,S^*_{jklr}$ يكون $Q^*_i = Q^*_i + 1$

إذا كان $Y < \dfrac{P^*_i + Q^*_i}{J \cdot K \cdot L \cdot R}$ أو $Z < \left|\dfrac{P^*_i - Q^*_i}{P^*_i + Q^*_i}\right|$ إذن يتم رفض المراقب i.

حيث:

N: عدد المراقبين

J: عدد نوافذ الزمن مع تجميع شرط الاختبار وتتابعه

K: عدد شروط الاختبار

L: عدد التتابعات

R: عدد مرات التكرار.

أما القيم المقترحة للمعلمات (X وY وZ) بعد التجربة والمكيفة مع هذه الطريقة فهي 0,2، 0,1، 0,3.

3.3.2-A1 الفرز اللاحق القائم على الارتباط

يجب أن يكون لكل مراقب طريقة مستقرة ومتماسكة للتصويت العادل على انحطاط الجودة بالنسبة إلى كل مشهد وخوارزمية. وتؤكد معايير النبذ مستوىً معقولاً من التوافق بين علامات مراقب ما وفق متوسط درجات جميع المراقبين لعملية اختبار معيّنة. ويستند معيار القرار إلى ارتباط بين العلامات الإفرادية ومتوسط العلامات المقابل من جميع مراقبي الاختبار. وهذا الإجراء أبسط في التنفيذ من الإجراء المقابل الموصوف في الأقسام السابقة.

1.3.3.2-A1 ارتباط بيرسون

يُفترض أن تكون العلاقة بين سلم الجودة ومدى علامات المراقبين خطيةً لتطبيق ارتباط بيرسون.

والهدف الرئيسي هو التحقق بطريقة بسيطة مما إذا كانت علامات أحد المراقبين متسقةً مع متوسط العلامات لجميع المراقبين في مجمل عملية الاختبار. ويُعتبر المرجع المخفي وسيلة تثبيت ذات جودة عالية. وفي حال وجود وسيلتي تثبيت واحدة عالية وأخرى منخفضة، فإنهما يزيدان من ارتباط العلامات، وعلى العكس من ذلك تقل تخالفات الارتباط بين المراقبين.

$$r(x,y)=\frac{\sum_{i=1}^{n}x_i y_i-\frac{\left(\sum_{i=1}^{n}x_i\right)\left(\sum_{i=1}^{n}y_i\right)}{n}}{\sqrt{\left(\sum_{i=1}^{n}x_i^2-\frac{\left(\sum_{i=1}^{n}x_i\right)^2}{n}\right)\left(\sum_{i=1}^{n}y_i^2-\frac{\left(\sum_{i=1}^{n}y_i\right)^2}{n}\right)}} \qquad (11)$$

حيث:

x_i: متوسط علامة جميع المراقبين للمعاملات الثلاثة (الخوارزمية، معدّل البتات، المشهد)

y_i: العلامة الإفرادية لمراقب واحد للمعاملات الثلاثة ذاتها

n: (عدد الخوارزميات) × (عدد المشاهد)

i: {رقم الكودك، رقم معدّل البتات، رقم المشهد}.

2.3.3.2-A1 ارتباط رتبة سبيرمان

يمكن تطبيق ارتباط رتبة سبيرمان حتى لو لم تتحقق فرضية أن العلاقة بين سلم الجودة ومدى علامات المراقبين خطية[4]:

$$r(x,y)=\left[1-\frac{6\times\sum_{i=1}^{n}\left[R(x_i)-R(y_i)\right]^2}{n^3-n}\right] \qquad (12)$$

حيث:

xi: متوسط علامة جميع المراقبين للمعاملات الثلاثة (الخوارزمية، معدّل البتات، المشهد)

yi: العلامة الإفرادية لمراقب واحد من أجل المعاملات الثلاثة ذاتها

n: (عدد الخوارزميات) × (عدد المشاهد)

$R(xi$ أو $yi)$: ترتيب الرتبة

i: {رقم الكودك، رقم معدّل البتات، رقم المشهد}.

3.3.3.2-A1 معايير النبذ النهائي لاستبعاد مراقب اختبار

يُنفَّذ ارتباطا رتبة سبيرمان وبيرسون لاستبعاد مراقب (مراقبين) وفق الشروط التالية:

إذا كان [متوسط (r) – sdt(r)] > عتبة الارتباط القصوى (MCT).

فإن عتبة النبذ = عتبة الارتباط القصوى (MCT).

وفيما عدا ذلك فإن: عتبة النبذ = [متوسط (r) – sdt(r)]

[4] عموماً، تعتبر نتائج ارتباط بيرسون قريبةً جداً من نتائج ارتباط سبيرمان.

وإذا كان [r (مراقب i)] > عتبة النبذ.
لا يتم استبعاد المراقب "i" للاختبار.
وفيما عدا ذلك يتم استبعاد المراقب "i" للاختبار.

حيث:

r = الحد الأدنى (ارتباط بيرسون، ارتباط رتبة سبيرمان)

متوسط (r): متوسط الارتباطين لجميع مراقبي الاختبار

sdt(r): الانحراف المعياري لارتباطَي جميع مراقبي الاختبار

عتبة الارتباط القصوى (MCT) = 0,85.

وتعد القيمة MCT 0,85 صالحةً لطريقتي SAMVIQ وDSCQS، فيما عدا ذلك يجب اعتبار قيمة MCT 0,7 لطريقتي SS وDSIS.

4.2-A1 حساب متوسط العلامات وفترات الثقة في ظل ظروف الاختبار الصعبة

يتعين، في كثير من الأحيان، إجراء اختبار شخصي في ظل ظروف صعبة. على سبيل المثال، في اختبار التعهيد الجماعي، يتعرض الأشخاص لبيئة أقل انضباطاً مما هي عليه في المختبر. وفي اختبار واسع النطاق تجريه مختبرات متعددة، يمكن أن يؤدي التباين بين المختبرات إلى تغاير كبير في الدرجات التي يتم جمعها. والطرائق المقدمة في الفقرات من 1.2-A1 إلى 3.2-A1 لا تكون في الغالب مناسبةً تماماً لمثل هذه الظروف. ويقدم هذا القسم تقنيةً متقدمةً لتحليل البيانات ثبت أنها تحسن جودة البيانات الخاصة بمتوسط العلامات وفترات الثقة المستخرجة. ويمكن أيضاً العثور على تنفيذ مرجعي بلغة Python في المرفق 1 بهذا الملحق.

الحدس وراء هذه التقنية هو ما يلي. من المفيد وضع نموذج واضح لسلوك كل فرد؛ خاصةً أن تحيز الشخص واتساقه من العوامل البشرية المهيمنة التي تؤثر على أصوات الشخص. ومن خلال إجراء تكراري، تحاول هذه التقنية إجراء تقدير مشترك للجودة الحقيقية لكل عرض والتحيز والاتساق لكل شخص. ويمكن تفسير الجودة الحقيقية المقدرة لكل عرض على أنها "متوسط علامات الرأي الموزون على أساس الاتساق المزال منه التحيز". وبالمقارنة مع الفرز اللاحق للأشخاص الموصوف في الفقرة 1.3.2-A1، والذي إما يحتفظ بجميع أصوات شخص ما أو يرفضها ("الرفض القاطع")، يمكن وصف هذا الأسلوب بأنه "الرفض الناعم". وهذا يعني أنه بالنسبة لشخص ناشز يصوت بشكل غير متسق، فإن أصوات هذا الشخص سيكون لها وزن صغير، وبالتالي لا تساهم إلا قليلاً في متوسط العلامات الإجمالي. وأحد المنتجات الثانوية لهذه التقنية هو تقدير انحياز كل شخص في الاختبار واتساقه. وهذه معلومات قيمة لمدى ملاءمة الشخص لإجراء اختبارات شخصية، وبالتالي يمكن استخدامها لفحص الأشخاص في الاختبارات المستقبلية. فعلى سبيل المثال، إذا أظهر أحد المشاركين أن التصويت غير متسق إلى حد كبير، فقد يتم استبعاده/استبعادها من الجلسات المستقبلية.

وتقوم التقنية أولاً بتقدير متوسط العلامات لكل عرض لجميع الأشخاص والتكرارات:

$$\bar{u}_{jk} = \frac{1}{N \cdot R} \sum_{i=1}^{N} \sum_{r=1}^{R} u_{ijkr} \qquad (13)$$

حيث u_{ijkr} هي علامات المراقب i للحالة j، والتتابع/الصورة k، والتكرار r، وN هو عدد المراقبين، وR يرمز إلى عدد التكرارات.

وفي الخطوة الثانية، يتم تقدير انحياز كل مراقب b_i بواسطة:

$$b_i = \frac{1}{J \cdot K \cdot R} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{r=1}^{R} u_{ijkr} - \bar{u}_{jk} \qquad (14)$$

حيث J وK عدد الحالات وعدد التتابعات، على التوالي. ثم يتم تنفيذ الخطوات التالية في حلقة تكرارية.

ويتم تسجيل التقدير الحالي لمتوسط العلامات لكل عرض في صورة $\bar{u}_{jk}^{c}$

$$\bar{u}_{jk}^{c} = \bar{u}_{jk} \qquad (15)$$

ويلي ذلك حساب المتبقي في كل درجة ملاحظة لا يمكن تفسيرها بمتوسط العلامات وتحيز المراقب:

$$e_{ijkr} = u_{ijkr} - \bar{u}_{jk} - b_i \tag{16}$$

وتُستخدم هذه القيم المتبقية فيما بعد في حساب عدم الاتساق، σ_i، لكل مراقب:

$$\sigma_i = \sqrt{\frac{1}{J \cdot K \cdot R} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{r=1}^{R} (u_{ijkr} - \mu_{e_i})^2} \tag{17}$$

حيث:

$$\mu_{e_i} = \frac{1}{J \cdot K \cdot R} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{r=1}^{R} e_{ijkr} \tag{18}$$

ويمكن بعد ذلك الحصول على التقديرات الجديدة لمتوسط العلامات بالمعادلة:

$$\bar{u}_{jk} = \frac{\sum_{i=1}^{N} \sum_{r=1}^{R} \sigma_i^{-2} (u_{ijkr} - b_i)}{\sum_{i=1}^{N} \sum_{r=1}^{R} \sigma_i^{-2}} \tag{19}$$

ثم يُحدَّث الانحياز تبعاً للمعادلة (12).

وتُقفل الحلقة إذا كان:

$$\sum_{j=1}^{J} \sum_{k=1}^{K} (\bar{u}_{jk} - \bar{u}_{jk}^{c})^2 \tag{20}$$

وبعد إقفال الحلقة، يُتحصل على الانحراف المعياري لعلامات كل عرض كالتالي:

$$S_{jk} = \frac{\sigma_j}{\sqrt{N}} \tag{21}$$

حيث:

$$\sigma_j = \sqrt{\frac{1}{N \cdot R} \sum_{i=1}^{N} \sum_{r=1}^{R} (e_{ijkr} - \mu_{e_j})^2} \tag{22}$$

و

$$\mu_{e_j} = \frac{1}{N \cdot R} \sum_{i=1}^{N} \sum_{r=1}^{R} e_{ijkr} \tag{23}$$

ثم تُحسب فترة الثقة النهائية تبعاً للمعادلتين (2) و(3).

3-A1 العملية التي تسمح بتعيين علاقة بين العلامة المتوسطة والقياس الموضوعي لتشوه الصورة

عند إجراء الاختبارات الشخصانية بهدف دراسة العلاقة بين القياس الموضوعي للتشوه ومتوسط العلامات $\bar{u}$ (تحسب $\bar{u}$ وفقاً لما يرد في القسم 1.2-A1) يمكن استعمال العملية التالية التي تكمن في تعيين علاقة مستمرة بسيطة بين $\bar{u}$ ومعلمة الانحطاط.

1.3-A1 التقريب بواسطة دالة منطقية رياضية متناظرة

يبدو أن تقريب هذه العلاقة التجريبية بواسطة دالة منطقية رياضية له أهمية خاصة.

يمكن معالجة البيانات $\bar{u}$ على النحو التالي:

يقيَّس سلم القيم $\bar{u}$ بالنسبة إلى متغير مستمر p بحيث:

$$p = (\bar{u} - u_{min})/(u_{max} - u_{min}) \tag{24}$$

حيث:

u_{min}: أدنى علامة متاحة في سلم تقدير u لأسوأ نوعية

u_{max}: أدنى علامة متاحة في سلم تقدير u لأفضل نوعية.

ويُظهر التمثيل البياني للعلاقة بين p وD أن المنحني يميل إلى شكل شبيه بحرف S بتناظر مركزي شريطة أن تمتد الحدود الطبيعية لقيم D بعيداً عن المنطقة التي تتغير فيها قيمة u تغيراً سريعاً.

ويمكن الآن تقريب الدالة: $p = f(D)$ بواسطة دالة منطقية رياضية يكون اختيارها حكيماً، ويعبر عنها بالعلاقة التالية:

$$p = 1/[1 + \exp(D - D_M) \cdot G] \tag{25}$$

حيث D_M وG قيمة موجبة أو سالبة.

تُستعمل القيمة p المستخلصة من دالة التقريب المنطقية الرياضية المثلى من أجل توفير قيمة رقمية I مشتقة وفقاً للعلاقة:

$$I = (1/p - 1) \tag{26}$$

ويمكن الحصول على القيمتين D_M وG من البيانات التجريبية بعد التحويل التالي:

$$I = \exp(D - D_M) \cdot G \tag{27}$$

ويؤدي استعمال سلم لوغاريتمي لقيم I إلى علاقة خطية:

$$\log_e I = (D - D_M) \cdot G \tag{28}$$

يصبح الاستكمال الداخلي بخط مستقيم بسيطاً، وله في بعض الأحيان دقة كافية لكي يعتبر الخط المستقيم ممثلاً للانحطاط العائد إلى التأثير الذي تقيسه D.

ويعبر عندئذٍ عن ميل الخاصية بواسطة المعادلة التالية:

$$S = \frac{D_M - D}{\log_e I} = \frac{1}{G} \tag{29}$$

التي تعطي قيمة $G.\ DM$ المثلى هي قيمة D من أجل $I = 1$.

يمكن أن يعبر الخط المستقيم عن خاصية الانحطاط المصاحبة للانحطاط الخاص المعني. وتجدر الإشارة إلى أن من الممكن تعريف الخط المستقيم بواسطة القيمتين D_M وG المميزتين للدالة المنطقية الرياضية.

2.3-A1 التقريب بواسطة دالة لا تناظرية

1.2.3-A1 وصف الدالة

إن تقريب العلاقة بين علامات الاختبار والقياس الموضوعي لتشوه الصورة باستعمال دالة تناظرية رياضية عملية ناجحة في معظم الأحيان في حالة التمكن من قياس معلمة التشوه D في وحدة ذات صلة مثل النسبة S/N (dB). وإذا قيست معلمة التشوه في وحدة مادية d، مثل المهلة (ms)، يتعين إعادة وضع العلاقة (27) باستعمال المعادلة:

$$I = (d/d_M)^{1/G} \tag{30}$$

وبالتالي تصبح المعادلة (25):

$$p = 1/\left[1 + (d/d_M)^{1/G}\right] \tag{31}$$

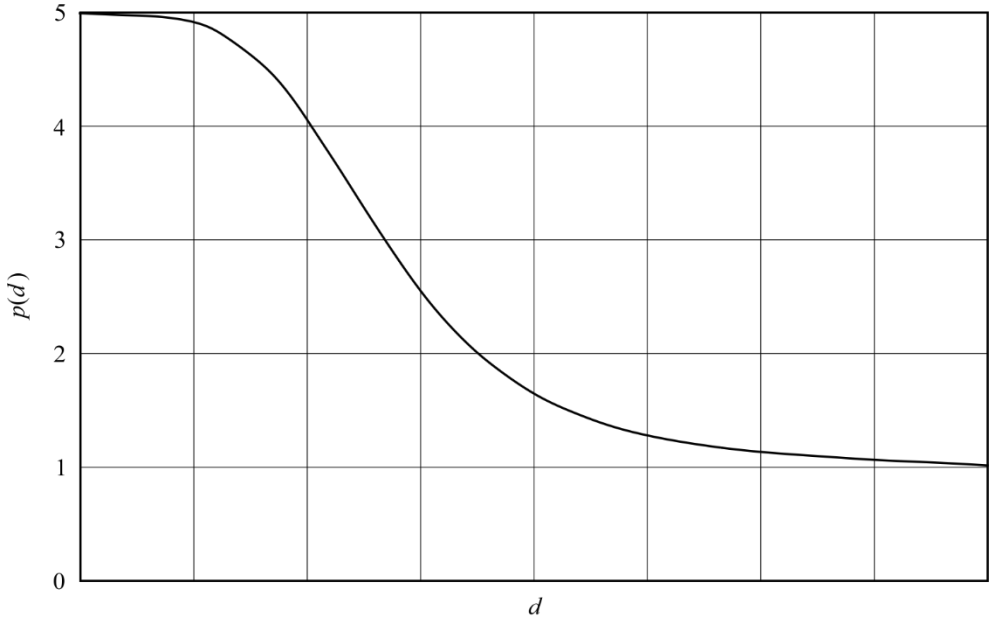
ومن شأن هذه الدالة تقريب الدالة المنطقية بطريقة لا تناظرية.

2.2.3-A1 تقدير معلمات التقريب

يمكن الحصول على تقدير المعلمات المثلى للدالة التي توفر أدنى الأخطاء المتبقية بين البيانات الفعلية والدالة، بواسطة أية خوارزمية تقدير تكراري. ويبين الشكل 1-3 مثالاً عن استعمال دالة لا تناظرية من أجل تمثيل بيانات شخصانية فعلية. ويسمح هذا التمثيل بتقدير قياسات موضوعية تقابل قيمةً شخصانيةً مهمةً: 4,5 مثلاً على السلم خماسي الدرجات.

الشكل 1-3

تقريب لا تناظري



BT.0500-01-3

3.3-A1 تصحيح الانحطاط/التحسين المتبقي وآثار حدود سلم التقدير

عملياً، لا يمكن لاستعمال الدالة المنطقية الرياضية أحياناً تفادي بعض الفروق بين بيانات التجريب والتقريب. وقد ينجم هذا التباين عن نهاية آثار سلم التقدير أو وجود متزامن لانحطاطات عدة في الاختبار من شأنها أن تؤثر على النموذج الإحصائي وتشوه الدالة النظرية المنطقية.

وهناك نوع من أنواع أثر نهايات سلم التقدير ينزع المراقبون عنده إلى عدم استعمال القيم القصوى للسلم وخاصةً علامات النوعية المرتفعة. وقد ينتج ذلك عن عدة عوامل بما فيها الإحجام النفسي عن إطلاق أحكام متطرفة. وعلاوةً على ذلك قد يسبب استعمال المتوسط الحسابي للأحكام وفقاً للمعادلة (1) قرب نهايات السلم الحصول على نتائج متحيزة بسبب التوزيع غير الغوسي للتقديرات في هذه المناطق.

وغالباً ما يرد في الاختبارات انحطاط متبقٍ (حتى في الصورة المرجعية لا يبلغ متوسط العلامات إلا القيمة $(\bar{u}_0 < u_{max})$.

وهنا بعض النُهج المفيدة لتصحيح البيانات الخام للتقديرات لمعالجة الاستنتاجات الصحيحة (انظر الجدول 1-3).

ويشكل تصحيح آثار "الحدود" إن وجدت في بيانات التجريب جزءاً بالغ الأهمية من معالجة البيانات. لذا ينبغي اختيار الإجراء بدقة بالغة. ويلاحظ أن إجراءات التصحيح هذه تضم افتراضات خاصة، لذا ينصح بالتزام الحيطة عند استعمالها؛ وينبغي تسجيل استعمالها لدى تقديم النتائج.

الجدول 1-3

مقارنة طرائق تصحيح آثار حدود سلم التقدير

طرائق التعويض عن آثار حدود سلم التقدير	العناصر		
	التعويض عن الخطأ المتبقي	التعويض عن التحسين المتبقي	تخالف في وسط سلم التقييم
عدم التعويض	لا	لا	لا
تحويل خطي للسلم	نعم	قد يكون الخطأ كبيراً	لا
تحويل لا خطي للسلم(1)	نعم	نعم	لا
طريقة قائمة على إضافة الوحدات	نعم	لا	نعم
طريقة تكاثرية	نعم	لا	نعم

(1) يتعين حساب التقديرات المصححة وفقاً للتحويل اللاخطي للسلم على النحو التالي:

$$u_{corr} = C(\bar{u} - u_{mid}) + u_{mid}$$

$$C = \frac{\bar{u} - u_{0_{min}}}{u_{0_{max}} - u_{0_{min}}} \frac{u_{max} - u_{mid}}{u_{0_{max}} - u_{mid}} + \frac{u_{0_{max}} - \bar{u}}{u_{0_{max}} - u_{0_{min}}} \frac{u_{min} - u_{mid}}{u_{0_{min}} - u_{mid}}$$

حيث:

u_{corr}: علامة مصححة

$\bar{u}$: علامة تجريبية غير مصححة

u_{min}، u_{max}: حدود سلم التقدير

u_{mid}: منتصف سلم التقدير

$u_{0\,min}$، $u_{0\,max}$: الحدان الأدنى والأعلى لاتجاه العلامات التجريبية.

4.3-A1 إدماج جانب الاعتمادية في الصور البيانية

تبنى ثلاث سلاسل من الدرجات انطلاقاً من الدرجات المتوسطة لكل انحطاط خاضع للاختبار وفترات الثقة عند 95%:

– سلسلة الدرجات الدنيا (قيم متوسطة – فترات الثقة)؛

– سلسلة الدرجات المتوسطة؛

– سلسلة الدرجات القصوى (قيم متوسطة + فترات الثقة).

ثم تقدر معلمات التقدير للسلاسل الثلاث تقديراً مستقلاً. ويمكن عندها أن ترسم الدوالّ الثلاث على الرسم البياني نفسه. ترسم دالتا السلسلتين القصوى والدنيا بالخطوط المنقطة ودالة الدرجات المتوسطة بالخط المتواصل. وتبين أيضاً القيم التجريبية على هذا الرسم البياني (انظر الشكل 1-4). ونحصل بذلك على تقدير لمنطقة الثقة المستمرة عند 95%.

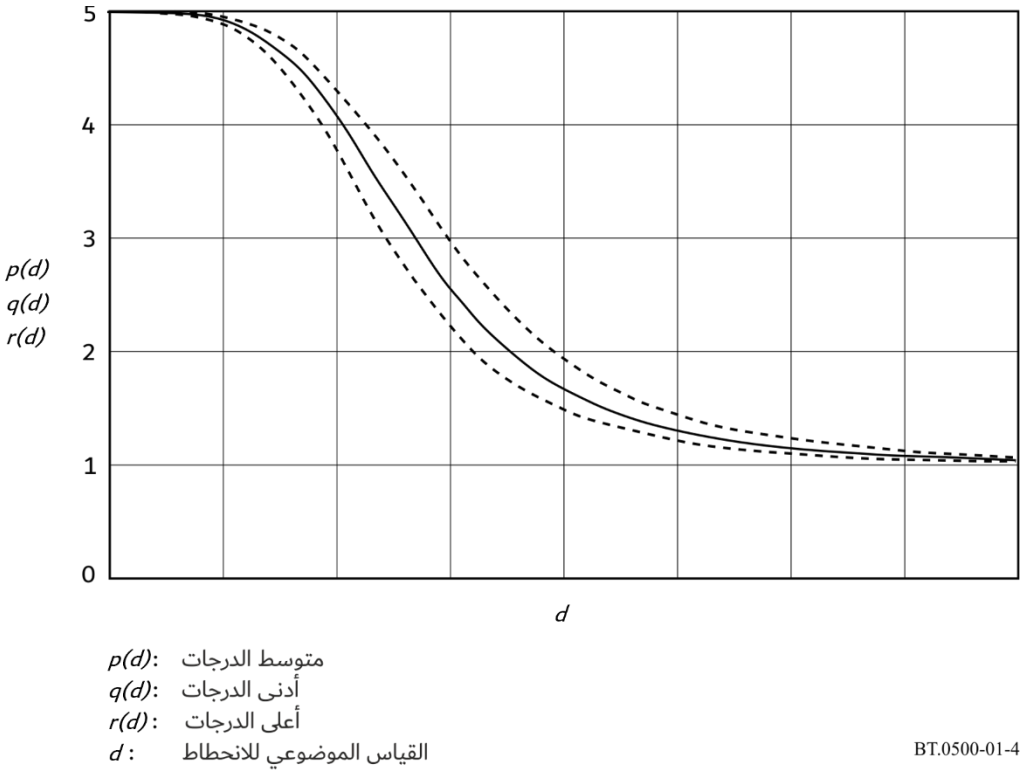
يمكن أن تقرأ من ثم للدرجة 4,5 (عتبة الرؤية الخاصة بالطريقة) قراءة مباشرة على الرسم البياني، فترة الثقة 95% المقدرة والتي يمكن أن تستعمل لتحديد مدى مسموح به.

الحيز بين منحني القيم القصوى ومنحني القيم الدنيا ليس فترة 95% بل تقديراً متوسطاً له.

يجب أن تقع 95% على الأقل من القيم التجريبية داخل منطقة الثقة؛ وإلا قد يُستنتج أن ثمة مشكلة في إجراء الاختبار أو أن نموذج الدالة المختار لم يكن النموذج الأمثل.

الشكل 1-4

حالة خاصية انحطاط لا تناظرية



4-A1 الاستنتاجات

لقد وُصِف إجراء لتقدير فترات الثقة أي دقة مجموعة من اختبارات التقدير الشخصاني.

يؤدي الإجراء أيضاً إلى تقدير كميات عامة متوسطة ليست مهمةً للتجربة المعينة فحسب، بل للتجارب الأخرى أيضاً المحققة بواسطة نفس المنهجية.

ولهذا يمكن أن تستعمل هذه الكميات لرسم مخططات سلوك فترة الثقة المفيدة للتقديرات الشخصانية، ولتخطيط التجارب المستقبلية كذلك.

المرفق 1
بالملحق 1

التنفيذ المرجعي للطريقة المأخوذة من الفقرة 4.2-A1

يتضمن هذا المرفق تنفيذاً مرجعياً بلغة Python لتقنية تحليل البيانات المقدمة في الفقرة 4.2-A1. وبيانات الشفرة والعينات المستخدمة متاحة أيضاً للجمهور في حزمة SUREAL Python على: https://github.com/Netflix/sureal/tree/master/itur_bt500_demo

ويجري إعداد بيانات الدخل على النحو التالي. يتم تنظيم الأصوات الأولية في مصفوفة ثنائية الأبعاد، مفصولةً بفاصلات. ويقابل كل صف أحد العروض (الصورة المصدر في ظل حالة الاختبار)؛ ويقابل كل عمود أحد الأشخاص.

لا يلزم قيام كل شخص بالتصويت على كل عرض. فإذا لم يقم الشخص *i* بالتصويت على العرض *jk*، فسيتم وضع "nan" (ليس رقماً) في الموقع (*jk,i*). وتُوضع بيانات الدخل في ملف csv.. ويوجد أدناه عينة صغيرة لملف csv. يضم أصوات 20 شخصاً و30 عرضاً مع تكرارين.

```
5.0,nan,5.0,4.0,2.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
1.0,3.0,5.0,2.0,5.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
3.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,3.0,4.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
1.0,4.0,3.0,4.0,5.0,5.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0
4.0,5.0,nan,3.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
4.0,3.0,2.0,5.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
1.0,3.0,4.0,5.0,1.0,4.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0,5.0,3.0,5.0,5.0,4.0,3.0
3.0,5.0,4.0,2.0,4.0,5.0,4.0,5.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0
5.0,2.0,1.0,3.0,3.0,4.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
1.0,2.0,1.0,1.0,3.0,1.0,1.0,1.0,1.0,3.0,1.0,2.0,2.0,1.0,1.0,1.0,2.0,1.0,1.0
5.0,5.0,3.0,1.0,3.0,1.0,2.0,2.0,2.0,3.0,2.0,3.0,4.0,2.0,1.0,2.0,2.0,1.0,2.0
5.0,2.0,4.0,3.0,4.0,2.0,2.0,2.0,2.0,4.0,3.0,3.0,3.0,5.0,2.0,2.0,2.0,4.0,2.0
5.0,5.0,5.0,5.0,4.0,3.0,3.0,3.0,3.0,5.0,3.0,4.0,4.0,3.0,2.0,2.0,3.0,3.0,3.0
5.0,5.0,4.0,3.0,5.0,4.0,4.0,4.0,4.0,5.0,4.0,4.0,5.0,4.0,3.0,3.0,4.0,3.0,3.0
1.0,4.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0
1.0,4.0,1.0,4.0,3.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0
4.0,2.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0
2.0,5.0,3.0,2.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
5.0,5.0,5.0,5.0,3.0,3.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0
4.0,5.0,5.0,3.0,5.0,2.0,2.0,3.0,1.0,3.0,3.0,2.0,3.0,5.0,1.0,1.0,2.0,2.0,2.0
1.0,2.0,2.0,4.0,5.0,1.0,2.0,2.0,1.0,3.0,2.0,2.0,4.0,2.0,3.0,1.0,2.0,2.0,1.0
4.0,5.0,3.0,5.0,2.0,3.0,2.0,3.0,3.0,4.0,2.0,3.0,4.0,3.0,3.0,1.0,2.0,2.0,2.0
1.0,5.0,3.0,5.0,4.0,2.0,3.0,3.0,3.0,5.0,3.0,3.0,4.0,2.0,3.0,2.0,3.0,3.0,2.0
5.0,5.0,5.0,5.0,1.0,4.0,4.0,3.0,3.0,5.0,3.0,4.0,4.0,4.0,4.0,3.0,4.0,3.0,3.0
5.0,5.0,5.0,5.0,4.0,5.0,4.0,4.0,4.0,5.0,5.0,4.0,4.0,5.0,5.0,5.0,5.0,3.0,4.0
5.0,1.0,4.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0
3.0,4.0,4.0,2.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0
4.0,1.0,3.0,5.0,3.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0
3.0,3.0,1.0,3.0,1.0,1.0,2.0,3.0,1.0,3.0,1.0,3.0,1.0,2.0,2.0,2.0,2.0,2.0,2.0
5.0,3.0,2.0,2.0,5.0,3.0,1.0,3.0,1.0,4.0,3.0,4.0,3.0,4.0,3.0,3.0,3.0,2.0,1.0,2.0
,
5.0,nan,5.0,4.0,2.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
1.0,3.0,5.0,2.0,5.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
3.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,3.0,4.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
1.0,4.0,3.0,4.0,5.0,5.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0
4.0,5.0,nan,3.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
4.0,3.0,2.0,5.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
1.0,3.0,4.0,5.0,1.0,4.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0,5.0,3.0,5.0,5.0,4.0,3.0
3.0,5.0,4.0,2.0,4.0,5.0,4.0,5.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0
5.0,2.0,1.0,3.0,3.0,4.0,5.0,5.0,3.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0
1.0,2.0,1.0,1.0,3.0,1.0,1.0,1.0,1.0,3.0,1.0,2.0,2.0,1.0,1.0,1.0,2.0,1.0,1.0
5.0,5.0,3.0,1.0,3.0,1.0,2.0,2.0,2.0,3.0,2.0,3.0,4.0,2.0,1.0,2.0,2.0,1.0,2.0
5.0,2.0,4.0,3.0,4.0,2.0,2.0,2.0,2.0,4.0,3.0,3.0,3.0,5.0,2.0,2.0,2.0,4.0,2.0
5.0,5.0,5.0,5.0,4.0,3.0,3.0,3.0,3.0,5.0,3.0,4.0,4.0,3.0,2.0,2.0,3.0,3.0,3.0
5.0,5.0,4.0,3.0,5.0,4.0,4.0,4.0,4.0,5.0,4.0,4.0,5.0,4.0,3.0,3.0,4.0,3.0,3.0
1.0,4.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,4.0,5.0,4.0,5.0,5.0
```

```
1.0,4.0,1.0,4.0,3.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0,4.0
4.0,2.0,5.0,5.0,4.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0
2.0,5.0,3.0,2.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
5.0,5.0,5.0,5.0,3.0,3.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0
4.0,5.0,5.0,3.0,5.0,2.0,2.0,3.0,1.0,3.0,3.0,2.0,3.0,5.0,1.0,1.0,2.0,2.0,2.0,2.0
1.0,2.0,2.0,4.0,5.0,1.0,2.0,2.0,1.0,3.0,2.0,2.0,4.0,2.0,3.0,1.0,2.0,2.0,1.0,3.0
4.0,5.0,3.0,5.0,2.0,3.0,2.0,3.0,3.0,4.0,2.0,3.0,4.0,3.0,3.0,1.0,2.0,2.0,2.0,3.0
1.0,5.0,3.0,5.0,4.0,2.0,3.0,3.0,3.0,5.0,3.0,3.0,4.0,2.0,3.0,2.0,3.0,3.0,2.0,3.0
5.0,5.0,5.0,5.0,1.0,4.0,4.0,3.0,3.0,5.0,3.0,4.0,4.0,4.0,4.0,3.0,4.0,3.0,3.0,4.0
5.0,5.0,5.0,5.0,4.0,5.0,4.0,4.0,4.0,5.0,5.0,4.0,4.0,5.0,5.0,5.0,5.0,3.0,4.0,4.0
5.0,1.0,4.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0
3.0,4.0,4.0,2.0,5.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,4.0,5.0,5.0,5.0,5.0,5.0,5.0,5.0
4.0,1.0,3.0,5.0,3.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0
3.0,3.0,1.0,3.0,1.0,1.0,2.0,3.0,1.0,3.0,1.0,3.0,1.0,2.0,2.0,2.0,2.0,2.0,2.0,2.0
5.0,3.0,2.0,2.0,5.0,3.0,1.0,3.0,1.0,4.0,3.0,4.0,3.0,4.0,3.0,3.0,3.0,2.0,1.0,2.0
```

توجد شفرة Python التي تنفذ الطريقة في الملف *demo_bt500.py*.

:demo_bt500.py

```
import argparse
import csv
import sys
import pprint

import numpy as np
from scipy import linalg


def read_csv_into_3darray(csv_filepath):
    """
    Read data from CSV file.

    The data should be organized in a 2D matrix, separated by comma. Each row
    correspond to a PVS; each column corresponds to a subject. If a vote is
    missing, a 'nan' is put in place.

    If some subjects evaluated a PVS multiple times, another 2D matrix of the
    same size [num_PVS, num_subjects] can be added under the first one. A row
    with a single comma (,) should be placed before the repetition matrix.
    Where the repeated vote is not available, a 'nan' is put in place.

    :param csv_filepath: filepath to the CSV file.
    :return: the numpy array in 3D [num_PVS, num_subjects, num_repetitions].
    """

    data = []
    data3dlist = []
    with open(csv_filepath, 'rt') as datafile:
        datareader = csv.reader(datafile, delimiter=',')

        for row in datareader:
            if row != ["", ""]:
                data.append(np.array(row, dtype=np.float64))
            else:
                data3dlist.append(data)
                data = []
        data3dlist.append(data)

    data3d = np.zeros([len(data3dlist[0]), len(data3dlist[0][0]), len(data3dlist)])

    for r_idx, r_mat in enumerate(data3dlist):
        data3d[:, :, r_idx] = r_mat

    return data3d


def weighed_nanmean_2d(a, wts, axis):
    """
    Compute the weighted arithmetic mean along the specified axis, ignoring
    NaNs. It is similar to numpy's nanmean function, but with a weight.

    :param a: 1D array.
    :param wts: 1D array carrying the weights.
```

```
    :param axis: either 0 or 1, specifying the dimension along which the means
    are computed.
    :return: 1D array containing the mean values.
    """

    assert len(a.shape) == 2
    assert axis in [0, 1]
    d0, d1 = a.shape
    if axis == 0:
        return np.divide(
            np.nansum(np.multiply(a, np.tile(wts, (d1, 1)).T), axis=0),
            np.nansum(np.multiply(~np.isnan(a), np.tile(wts, (d1, 1)).T), axis=0)
        )
    elif axis == 1:
        return np.divide(
            np.nansum(np.multiply(a, np.tile(wts, (d0, 1))), axis=1),
            np.nansum(np.multiply(~np.isnan(a), np.tile(wts, (d0, 1))), axis=1),
        )
    else:
        assert False


def one_or_nan(x):
    """
    Construct a "mask" array with the same dimension as x, with element NaN
    where x has NaN at the same location; and element 1 otherwise.

    :param x: array_like
    :return: an array with the same dimension as x
    """
    y = np.ones(x.shape)
    y[np.isnan(x)] = float('nan')
    return y


def get_sos_j(sig_j, u_jkir):
    """
    Compute SOS (standard deviation of score) for presentation jk
    :param sig_j:
    :param u_jkir:
    :return: array containing the SOS for presentation jk
    """
    den = np.nansum(
        stack_3rd_dimension_along_axis(one_or_nan(u_jkir) / np.tile(sig_j ** 2,
(u_jkir.shape[1], 1)).T[:, :, None],
                                        axis=1),
        axis=1)
    s_jk_std = 1.0 / np.sqrt(np.maximum(0., den))
    return s_jk_std


def stack_3rd_dimension_along_axis(u_jkir, axis):
    """
        Take the 3D input matrix, slice it along the 3rd axis and stack the resulting 2D
matrices
        along the selected matrix while maintaining the correct order.
        :param u_jkir: 3D array of the shape [JK, I, R]
        :param axis: 0 or 1
        :return: 2D array containing the values
            - if axis=0, the new shape is [R*JK, I]
            - if axis = 1, the new shape is [JK, R*I]
    """

    assert len(u_jkir.shape) == 3
    JK, I, R = u_jkir.shape

    if axis == 0:
        u = np.zeros([R * JK, I])

        for r in range(R):
            u[r * JK:(r + 1) * JK, :] = u_jkir[:, :, r]

    elif axis == 1:
        u = np.zeros([JK, R * I])

        for r in range(R):
            u[:, r * I:(r + 1) * I] = u_jkir[:, :, r]

    else:
        NotImplementedError

    return u


def run_alternating_projection(u_jkir):
    """
```

```
    Run Alternating Projection (AP) algorithm.

    :param u_jkir: 3D numpy array containing raw votes. The first dimension
    corresponds to the presentation (jk); the second dimension corresponds to the
    subjects (i); the third dimension correspons to the repetitions (r).
    If a vote is missing, the element is NaN.

    :return: dictionary containing results keyed by 'mos_j', 'sos_j', 'bias_i'
    and 'inconsistency_i'.
    """
    JK, I, R = u_jkir.shape

    # video by video, estimate MOS by averaging over subjects
    u_jk = np.nanmean(stack_3rd_dimension_along_axis(u_jkir, axis=1), axis=1)  # mean
marginalized over i

    # subject by subject, estimate subject bias by comparing with MOS
    b_jir = u_jkir - np.tile(u_jk, (I, 1)).T[:, :, None]
    b_i = np.nanmean(stack_3rd_dimension_along_axis(b_jir, axis=0), axis=0)  # mean
marginalized over j

    MAX_ITR = 1000
    DELTA_THR = 1e-8
    EPSILON = 1e-8

    itr = 0
    while True:

        u_jk_prev = u_jk

        # subject by subject, estimate subject inconsistency by averaging the
        # residue over stimuli
        e_jkir = u_jkir - np.tile(u_jk, (I, 1)).T[:, :, None] - np.tile(b_i, (JK, 1))[:,
:, None]
        sig_i = np.nanstd(stack_3rd_dimension_along_axis(e_jkir, axis=0), axis=0)
        sig_j = np.nanstd(stack_3rd_dimension_along_axis(e_jkir, axis=1), axis=1)

        # video by video, estimate MOS by averaging over subjects, inversely
        # weighted by residue variance
        w_i = 1.0 / (sig_i ** 2 + EPSILON)
        # mean marginalized over i:
        u_jk = weighed_nanmean_2d(
            stack_3rd_dimension_along_axis(u_jkir - np.tile(b_i, (JK, 1))[:, :, None],
axis=1),
            wts=np.tile(w_i, R),  # same weights for the repeated observations
            axis=1)

        # subject by subject, estimate subject bias by comparing with MOS,
        # inversely weighted by residue variance
        b_jir = u_jkir - np.tile(u_jk, (I, 1)).T[:, :, None]
        # mean marginalized over j:
        b_i = np.nanmean(stack_3rd_dimension_along_axis(b_jir, axis=0), axis=0)

        itr += 1

        delta_u_jk = linalg.norm(u_jk_prev - u_jk)

        msg = 'Iteration {itr:4d}: change {delta_u_jk}, u_jk {u_jk}, ' \
              'b_i {b_i}, sig_i {sig_i}'.format(
            itr=itr, delta_u_jk=delta_u_jk, u_jk=np.mean(u_jk),
            b_i=np.mean(b_i), sig_i=np.mean(sig_i))

        sys.stdout.write(msg + '\r')
        sys.stdout.flush()

        if delta_u_jk < DELTA_THR:
            break

        if itr >= MAX_ITR:
            break

    u_jk_std = get_sos_j(sig_j, u_jkir)
    sys.stdout.write("\n")

    mean_b_i = np.mean(b_i)
    b_i -= mean_b_i
    u_jk += mean_b_i

    return {
        'mos_j': list(u_jk),
        'sos_j': list(u_jk_std),
        'bias_i': list(b_i),
        'inconsistency_i': list(sig_i),
    }


if __name__ == "__main__":
```

```
parser = argparse.ArgumentParser()

parser.add_argument(
    "--input-csv", dest="input_csv", nargs=1, type=str,
    help="Filepath to input CSV file. The data should be organized in a 2D "
         "matrix, separated by comma. The rows correspond to PVSs; the "
         "columns correspond to subjects. If a vote is missing, input 'nan'"
         " instead.", required=True)

args = parser.parse_args()
input_csv = args.input_csv[0]

o_jir = read_csv_into_3darray(input_csv)

ret = run_alternating_projection(o_jir)

pprint.pprint(ret)
```

الملحق 2
بالجزء 1

وصف نسق مشترك لتبادل الملفات

إن الغرض من وضع نسق مشترك لتبادل الملفات هو تسهيل تبادل البيانات بين المختبرات التي تشارك جماعياً في حملات دولية للتقييم الشخصاني.

ويُعد كل تقدير شخصاني تبعاً لخمس مراحل متعاقبة ومترابطة هي: إعداد الاختبار وإجراء الاختبار، ومعالجة البيانات، وتقديم النتائج وتفسيرها. ويتوزع العمل عادةً في الحملات الدولية الواسعة على مختلف المختبرات المشاركة:

- يتولى أحد المختبرات مسؤولية إعداد الاختبار بالتعاون مع أطراف أخرى من خلال تحديد معلمات النوعية الواجب تقديرها والمواد الواجب استعمالها (حرجة عموماً ولكن دون إفراد) وإطار الاختبار (كالطريقة ومسافة المشاهدة وترتيب الجلسة وتتابع عرض بنود الاختبار مثلاً) وبيئة الاختبار (مثل شروط المشاهدة وكلمة المقدمة).
- يُطلب من المختبرات المشاركة بتوفير مادة الاختبار الخاضعة للمعالجة وفقاً للتقنيات الملائمة التي تمثل معلمات النوعية الواجب تقديرها (استناداً إلى المحاكاة أو التجهيزات المادية).
- يلزم وجود طرف آخر يكون من واجبه التعامل مع شريط الاختبار.
- تُجري الاختبار مختبرات مشاركة مختلفة باستعمال الشريط الذي سبق إعداده. ويمكن إجراء الاختبار دون تمييز. وفي هذه الحالة يقوم المختبر بإجراء الاختبار من خلال جمع تقديرات المشاهدين دون الحاجة إلى معرفة معلمات النوعية الخاضعة للتقييم.
- يطلب عادةً إلى مختبر آخر مشارك أن ينسق عملية جمع البيانات الخام الناتجة وإصدار النتائج وهي عملية يمكن إنجازها دون تمييز أيضاً.
- أخيراً، تفسر النتائج من النص/الجدول أو من التمثيل البياني ويصدر التقرير النهائي.

ويتيح النسق المقترح جمع النتائج التي تم الحصول عليها وفق إجراءات الاختبار المحددة في مرحلة إعداد الاختبار.

ويمتثل النسق لطرائق التقييم الوارد وصفها في الجزأين 1 و2 من هذه التوصية.

وتتألف من ملفات نصوص يظهر تنظيمها في الجدولين 1-4 و1-5، وتبنى قواعد تركيبها حول وسوم وميادين إضافةً إلى مجموعة محدودة من الرموز الخاصة (مثل "[" و"]" و" " و"ـ" و"=").

ليس هناك أي حدود من حيث الاستطاعة (مثل عدد المختبرات المشاركة والمراقبين وتتابعات الاختبارات ومعلمات النوعية أو حدود سلم التقدير أو نمط التقدير الحدودي).

الجدول 1-4

تحديد نتائج نسق الملف

تحديد نسق الملفات والشكل	الشرح
[Test framework]↵	[معرف القسم]
Type = *"DSCQS" or "DSIS I"*, "*DSIS II", etc.*↵	معرف الطريقة المستعملة في التوصية ITU-R BT.500
Number of sessions = *1 ≤ integer ≤ x*↵	عدد الجلسات[1] التي توزع فيها اختبارات
Scale minimum = *integer*↵	تعريف سلم التقدير (انظر متطلبات الطريقة إن وجدت)
Scale maximum = *integer*↵	
Display size = *integer*↵	طول الخط القطري لشاشة العرض (بالبوصة)
Display make and model = *chain of characters*↵	
[RESULTS] ↵	[معرف القسم]
Number of results = *1 ≤ integer ≤ y*↵	عدد ملفات النتائج[1] التي ينظر فيها
Result(*j*).Filename(s) = *character string*.DAT↵	اسم الملف Full.DAT (انظر الجدول 1-5) بما فيه المسار
....	
Result(*j*).Name = *character string*↵	اسم ملف النتائج المتداول
Result(*j*).Laboratory = *character string*↵	تعريف المختبر الذي أجرى الاختبار
Result(*j*).Number of observers = *1 ≤ integer ≤ N*↵	إجمالي عدد المراقبين
Result(*j*).Training = *"Yes" or "No"* ↵	يدل على ما إذا كانت التقديرات المجمعة في التدريبات مدرجة في الملف DAT المرفق
[Result(*j*).Session(*i*).Observers] ↵	[معرف القسم]
O(*k*).First Name = *character string*↵	[تعريف المراقب]
O(*k*).Last Name = *character string*↵	
O(*k*).Sex = *"F" or "M"* ↵	خياري
O(*k*).Age = *integer*↵	خياري
O(*k*). Occupation = *character string*↵	زمر اجتماعية اقتصادية رئيسية (مثال: عمال، طلاب)
O(*k*).Distance = *integer*↵	مسافة المشاهدة، ارتفاع الشاشة (3 *H*, 4 *H*, 6 *H*)

(1) الجلسة: يمكن تقسيم الاختبار إلى عدد من الجلسات من أجل التقيد بالشروط المحددة لطول المدة القصوى للاختبار. ويمكن لنفس المراقبين أو لمراقبين مختلفين أن يشاركوا في جلسات مختلفة يطلب منهم خلالها تقدير بنود اختبار مختلفة. ويعطي دمج التقديرات المجمعة من مختلف الجلسات مجموعة كاملة من النتائج (عدد العروض × عدد التقديرات لكل عرض). ويمكن إرفاق النتائج في ملفات DAT يمكن إرسالها عند كل اختبار.

الجدول 1-5

نسق ملف نصوص النتائج DAT للبيانات الخام

نسق وشكل الملف Filename.DAT	الشرح
integer integer integer.....↵	يتألف ملف البيانات الخام DAT من قيم التقدير يفصل بينها فاصل. وينبغي استعمال سطر واحد لكل مراقب
integer integer integer.....↵	تُسجل البيانات الخام وفق ترتيب الدخول
integer integer integer.....↵	يمكن توزيع البيانات على ملفات DAT مختلفة تحددها في الجدول 1-4 المعلمة Result(*j*). Filename(s)[1].
.....	

(1) انظر الملاحظة (1) في الجدول 1-4.

الملحق 3
(إعلامي)
بالجزء 1

خصائص العطل في محتوى الصورة

1-A3 مقدمة

سوف يخضع النظام، بعد تنفيذه، إلى عدد يمكن أن يكون كبيراً من البرنامج، وقد لا يتم ذلك دون فقدان في النوعية. ويجب، عند تحديد ملاءمة النظام، أن تُعرف نسبة مواد البرنامج التي قد تكون حرجة للنظام وفقدان النوعية المتوقع في هذه الحالات. إذ ما هو مطلوب بالفعل، هو خاصية العطل في محتوى الصورة بالنسبة إلى النظام المعني.

وتعتبر هذه الخاصية ذات أهمية خاصة للأنظمة التي قد لا يتعرض أداؤها لانحطاط منتظم عندما تزداد المواد حرجاً. فيمكن، على سبيل المثال، أن تحافظ بعض الأنظمة الرقمية والتكييفية على نوعية عالية في مدى واسع من مواد البرامج، لكنها تنحط خارج هذا المدى.

2-A3 تحديد خاصية العطل

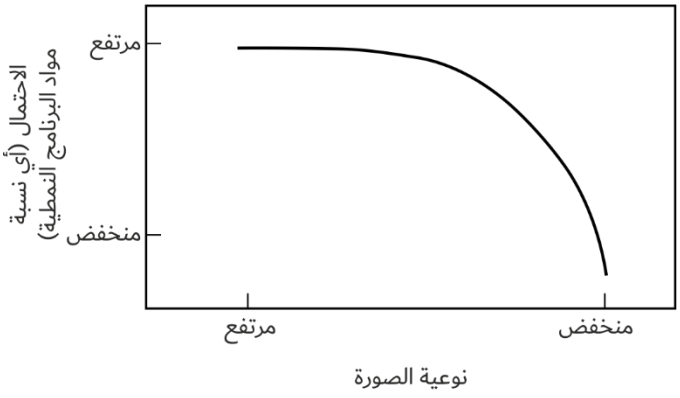
يحدد مفهوم الخاصية في محتوى الصورة نسبة المواد التي يمكن أن تظهر على المدى الطويل والتي يحقق النظام بالنسبة إليها سويات خاصة من النوعية. ويمثل ذلك في الشكل 1-5.

يمكن الحصول على خاصية العطل في محتوى الصورة على أربع مراحل:

- *المرحلة 1*: تحديد قياس خوارزمي "للحرج" يكون قادراً على ترتيب عدد من تتابعات الصور التي يتسبب لها النظام المعني أو صنف الأنظمة المعني بتشوهات، وعلى نحو يقابل فيه الترتيب ما كان يمكن أن يحصل عليه إنسان مراقب يؤدي هذه المهمة. ويمكن أن يتضمن القياس الحرج جوانب من النماذج المرئية.
- *المرحلة 2*: استناداً إلى تطبيق القياس الحرج على عدد كبير من العينات المأخوذة من برامج تلفزيونية نمطية، يتم استخلاص توزيع يقدر احتمال حدوث المواد التي توفر سويات مختلفة من الحرج بالنسبة إلى النظام أو إلى صنف الأنظمة المعني. ويبين الشكل 1-6 مثالاً لتوزيع هذا النمط.
- *المرحلة 3*: استخلاص قدرة النظام على المحافظة على النوعية من خلال وسائل تجريبية عندما تتزايد السوية الحرجة لمواد النظام. ويتطلب ذلك، في التطبيق العملي، تقديراً شخصانياً للنوعية المحققة في النظام مع مواد تُنتقى لاعتيان مدى الحرج المعرف في المرحلة 2. وينتج عن ذلك دالة تحدد العلاقة بين النوعية المحققة في النظام وسوية الحرج لمواد البرنامج. ويعطي الشكل 1-7 مثالاً عن هذه الدالة.
- *المرحلة 4*: استخلاص المعلومات من المرحلتين 2 و3 بهدف الحصول على خاصية العطل في محتوى الصورة على النحو المقدم في الشكل 1-5.

الشكل 1-5

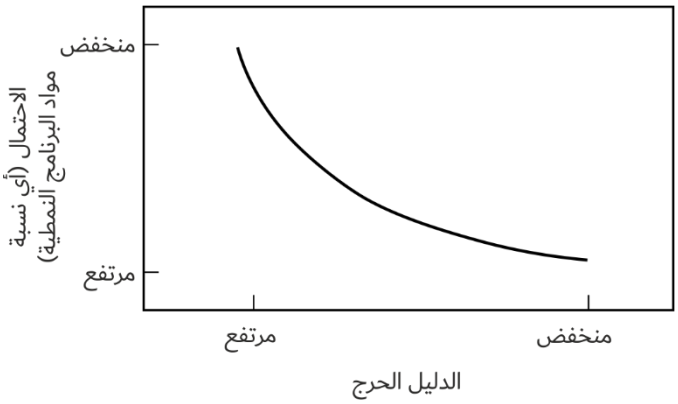
التمثيل البياني لخاصية العطل الممكن في محتوى الصورة



BT.0500-01-5

الشكل 1-6

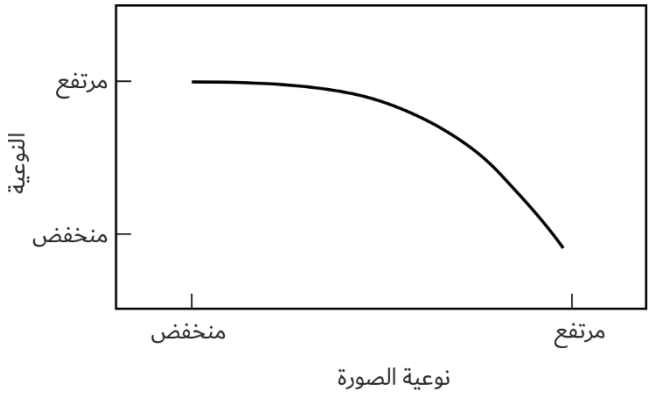
احتمال ظهور مواد ذات سويات حرج مختلفة



BT.0500-01-6

الشكل 1-7

مثال لدالة تحديد العلاقة بين النوعية وسوية حرج مواد البرنامج



BT.0500-01-7

3-A3 استعمال خاصية العطل

تعتبر خاصية العطل التي توفر صورةً شاملةً للأداء الذي يمكن تحقيقه على مدى أنماط البرامج الممكنة، أداةً مهمةً لدراسة ملاءمة الأنظمة. ويمكن استعمال خاصية العطل في ثلاث حالات:

- تحسين معلمات النظام إلى أبعد حد ممكن (مثل استبانة المصدر، ومعدل البتات، وعرض النطاق) في مرحلة التصميم لمواءمته مع متطلبات الخدمة؛
- دراسة مدى ملاءمة نظام معين (أي توقع تأثير العطل وأهميته في أثناء التشغيل)؛
- تقدير الملاءمة النسبية لأنظمة بديلة (أي مقارنة خصائص العطل وتحديد النظام الأنسب للاستعمال). وتجدر الإشارة إلى أنه من الممكن أن تستعمل عدة أنظمة من أنماط مشابهة دليل الحرج نفسه، لكن الأنظمة ذات الأنماط المختلفة لها دلائل حرج مختلفة. ولما كانت خاصية العطل تعبر فقط عن احتمال ملاحظة سويات مختلفة من النوعية في التطبيق العملي، فمن الممكن مقارنة الخصائص مباشرةً حتى عندما تنتج عن دلائل مختلفة للحرج مميزة لكل نظام.

على الرغم من أن الطريقة الموصوفة في هذه التوصية توفر طريقةً لقياس خاصية العطل في محتوى الصورة لنظام معين، فإنه من غير المؤكد أن تتمكن من التنبؤ بأن نظاماً يعتبر مقبولاً من جانب مشاهد لخدمة تلفزيونية. وقد يتطلب الحصول على هذه المعلومات، أن تقدم لعدد من المشاهدين برامج مشفرة وفقاً للنظام المعني، وأن تدرس تعليقاتهم.

ويرد مثال لخصائص عطل محتوى الصورة في التلفزيون الرقمي في الملحق 1 بالجزء 3.

الملحق 4
(إعلامي)
بالجزء 1

طريقة تحديد خاصية عطل مركبة بالنسبة إلى محتوى البرنامج وشروط الإرسال

1-A4 مقدمة

تحدد خاصية العطل المركبة العلاقة بين نوعية الصورة المدركة واحتمال ظهورها عملياً، على نحو يُدرس فيه بوضوح محتوى البرنامج وشروط الإرسال.

ويمكن مبدئياً، أن تُستخلص هذه الخاصية من دراسة تقديرات شخصانية مع عدد كافٍ من الملاحظات والاختبارات ونقاط الاستقبال التي تقدم عينةً ممثلةً لجمهور محتويات البرامج وشروط الإرسال الممكنة. غير أن من المحتمل ألا يكون هذا النوع من التجارب قابلاً للتحقيق عملياً.

يصف هذا الملحق إجراءً بديلاً سهل التحقيق يسمح بتحديد خصائص العطل المركب. وتتضمن هذه الطريقة ثلاث مراحل:

- تحليل محتوى البرنامج؛
- تحليل قناة الإرسال؛
- استخلاص خصائص العطل المركب.

2-A4 تحليل محتوى البرنامج

تتضمن هذه المرحلة عمليتين، يُستخلص في الأولى قياس مناسب لمحتوى البرنامج، وتُقدَّر في الثانية احتمالات أن تحدث عملياً قيم هذا القياس.

إن قياس محتوى البرنامج هو قياس إحصائي يطال جوانب محتوى البرنامج التي تشدد على قدرة النظام (أو الأنظمة) المعني بتوفير استعادة موثوقة لمواد البرنامج. من الواضح أن هذا القياس قد يكون أفضل لو ارتكز على نموذج إدراكي مناسب. لكن يكفي في غياب هذا النموذج اللجوء إلى قياس يلتقط بعض الجوانب من مدى التنوع الفضائي داخل أرتال/ومجالات فيديوية وفيما بينها، شريطة أن يقدم هذا القياس علاقةً شبه منتظمة مع نوعية الصورة المحسوسة. وقد يتوجب استعمال قياسات مختلفة لأنظمة (أو أصناف من الأنظمة) تستعمل مقاربات مختلفة تماماً لتمثيل الصورة.

يجب بعد انتقاء القياس المناسب، أن تقدر الاحتمالات التي يمكن أن تظهر فيها هذه القيم الإحصائية. ويمكن أن يتم ذلك بطريقتين مختلفتين:

- تحلل، بواسطة إجراء تجريبي، عينة عشوائية من 200 قطعة برنامج من 10 ثوان تقريباً في نسق استوديو مناسب للنظام أو الأنظمة المعني في الاستبانة، وتردد الرتل والنسبة الباعية. ويعطي تحليل هذه العينة الترددات النسبية لظهور القيم الإحصائية التي تعتبر تقديرات لاحتمالات الظهور في التطبيق العملي؛
- أو يستعمل في الطريقة النظرية، نموذج نظري لتقدير الاحتمالات. وتجدر الإشارة إلى أنه رغم تفضيل الطريقة التجريبية، قد يتوجب في حالات محددة أن تستعمل الطريقة النظرية (مثلاً عندما لا تتوفر معلومات كافية عن محتوى البرنامج لا سيما مع ظهور تكنولوجيات إنتاج جديدة).

وسوف تؤدي التحليلات أعلاه إلى توزيع احتمالات لقيم المحتوى الإحصائية (انظر أيضاً الملحق 3). وتُضم هذه النتائج إلى نتائج تحليل شروط الإرسال من أجل التحضير لمرحلة الإجراء النهائية.

3-A4 تحليل قناة الإرسال

تتضمن هذه المرحلة كذلك عمليتين، يُستخلص في الأولى قياس أداء قناة الإرسال، ويُقدر في الثانية توزيع احتمالات ظهور قيم القياس في التطبيق العملي.

إن قياس قناة الإرسال هو قياس إحصائي يطال جوانب أداء القناة التي تؤثر في قدرة النظام (أو الأنظمة) المعني بتوفير استعادة موثوقة لمواد المصدر. من الواضح أن هذا القياس قد يكون أفضل لو ارتكز على نموذج إدراكي مناسب. لكن يكفي في غياب هذا النموذج اللجوء إلى قياس يلتقط بعض الجوانب من التقييد الذي تفرضه القناة، شريطة أن يقدم هذا القياس علاقة شبه منتظمة مع نوعية الصورة المحسوسة. وقد يتوجب استعمال قياسات مختلفة لأنظمة (أو أصناف من الأنظمة) تستعمل مقاربات مختلفة تماماً لتشفير القناة.

يجب بعد انتقاء القياس المناسب، أن تقدر الاحتمالات التي يمكن أن تظهر فيها هذه القيم الإحصائية. ويمكن أن يتم ذلك بطريقتين مختلفتين:

- يقاس، بواسطة إجراء تجريبي، أداء القناة عند 200 لحظة ونقاط استقبال مختارة عشوائياً، على سبيل المثال. ويعطي تحليل هذه العينة الترددات النسبية لظهور القيم الإحصائية التي تعتبر تقديرات لاحتمالات الظهور في التطبيق العملي؛
- أو يستعمل في الطريقة النظرية، نموذج نظري لتقدير الاحتمالات. وتجدر الإشارة إلى أنه رغم تفضيل الطريقة التجريبية، قد يتوجب، في حالات محددة، أن تستعمل الطريقة النظرية (مثلاً عندما لا تتوفر معلومات كافية ذات أهمية عن أداء القناة خاصة مع ظهور تكنولوجيات الإرسال الجديدة).

وسوف تؤدي التحليلات أعلاه إلى توزيع احتمالات للقيم الإحصائية للقناة. وتُضم هذه النتائج إلى نتائج تحليل محتوى البرنامج من أجل التحضير لمرحلة الإجراء النهائية.

4-A4 تحديد خصائص العطل المركب

تكمن هذه المرحلة في تجربة شخصانية يتغير فيها محتوى البرنامج وشروط الإرسال معاً وفقاً للاحتمالات المحددة في المرحلتين السابقتين.

الطريقة الأساسية المستعملة هي إجراء النوعية المستمرة ثنائي الحافز، وبشكل خاص صيغة 10 ثوان الموصى بها لتتابعات الصور المتحركة. (انظر الملحق 2 بالجزء 2). المرجع هنا هو صورة ذات نوعية الاستوديو تقدم في نسق مناسب (على سبيل المثال، نسق مع استبانة، وتردد رتل، ونسبة باعية مناسبة للنظام (أو للأنظمة) المعني (المعنية)). وفي المقابل، يقدم الاختبار الصورة نفسها التي قد تُستقبل في النظام (أو الأنظمة) المعني (المعنية) ضمن شروط القناة المختارة.

تنتقى مواد الاختبار وشروط القناة وفقاً للاحتمالات المحددة في أول مرحلتين من الطريقة. وتحتوي قطع مواد الاختبار التي حللت كل منها بهدف تحديد قيمتها الأساسية وفقاً لإحصائية المحتوى، على مجموعة انتقاء. وتؤخذ من هذه المجموعة عينات من المواد تغطي مدى القيم الإحصائية الممكن، فتؤخذ متناثرةً عند السويات الحرجة الدنيا، وأكثر كثافةً عند السويات الأكثر حرجاً. وتنتقى القيم الإحصائية للقناة بطريقة مماثلة. ثم يصار إلى ضم هذين المصدرين المستقلين للتأثير عشوائياً بحيث يؤدي ذلك إلى محتويات وشروط إرسال مختلطة ذات احتمال معروف.

وتُستعمل فيما بعد نتائج هذه الدراسات التي تحدد العلاقة بين نوعية الصورة المحسوسة واحتمالات ظهورها عملياً في تقدير ملاءمة نظام معين أو في مقارنة الأنظمة بالنسبة إلى مدى ملاءمتها.

الملحق 5
(إعلامي)
بالجزء 1

الآثار السياقية

تنتج الآثار السياقية عندما يتأثر التقدير الشخصاني لصورة ما بترتيب وشدة الانحطاطات الظاهرة. وعلى سبيل المثال، إذا عُرضت صورة شديدة الانحطاط بعد سلسلة من الصور ضئيلة الانحطاط قد يقدِّر المشاهدون هذه الصورة خطأً أقل مستوى مما لو عرض في سياقٍ عادي.

وقد أجرت مجموعة من أربعة مختبرات في بلدان مختلفة بحوثاً عن الآثار السياقية المحتملة المرتبطة بنتائج طرائق ثلاث (طريقة DSCQS والصيغة 2 من طريقة DSIS وطريقة مقارنة الحوافز) مستعملة لتقييم نوعية الصور. وأُنتجت مواد الاختبار باستعمال التشفير MPEG (ML@MP) بموازاة تخفيف الاستبانة الأفقية. واستُعملت أربعة شروط اختبار أساسية (B1، B2، B3، B4) وستة شروط اختبار سياقية لكل سلسلة اختبارات، يحتوي أحدها على انحطاطات سياقية ضئيلة والأخرى على انحطاطات شديدة. وطُبقت طرائق الاختبار الثلاث على سلسلتي الاختبارات. والآثار السياقية هي الفرق بين نتائج الاختبار الذي يحتوي بصورة أساسية على انحطاطات ضئيلة ونتائج الاختبار الذي يحتوي بصورة أساسية على انحطاطات شديدة. واستُعمل شرطا الاختبار الأساسيان B2 وB3 لتحيد آثار السياق.

وتدل نتائج المختبرات مجتمعةً على عدم وجود آثار سياقية لطريقة DSCQS. لكن هذه الآثار كانت واضحةً في طريقة DSIS وطريقة المقارنة. أما أشد الآثار السياقية فظهرت في النوع II من طريقة DSIS. وتدل النتائج على إمكانية أن تسبب هيمنة الانحطاطات الضئيلة تقديرات أقل درجة للصورة بينما تسبب هيمنة الانحطاطات الشديدة تقديرات أعلى درجة.

وتوحي نتائج البحوث بأن طريقة DSCQS هي الطريقة الأفضل لتقليص آثار السياق إلى أبعد حد في عمليات التقدير الشخصاني لنوعية الصورة التي يوصي بها قطاع الاتصالات الراديوية.

ويرد مزيد من المعلومات حول البحوث المذكورة آنفاً في التقرير ITU-R BT.1082.

الملحق 6
(إعلامي)
بالجزء 1

مقياسا المعلومات المكانية والزمانية

مقياسا المعلومات المكانية والزمانية الواردان أدناه هما أحاديا القيمة لكل رتل عبر تتابع اختبار كامل. ويفضي ذلك إلى سلاسل زمنية من القيم التي ستتغيّر عموماً بدرجة ما. أما مقاييس المعلومات الإدراكية الواردة أدناه فهي تزيل إمكانية التغيّر هذه بدالّة قصوى (قيمة قصوى للتتابع). ويمكن دراسة إمكانية التغيّر بحد ذاتها بشكل مفيد، باستعمال رسوم بيانية مثلاً للمعلومات المكانية والزمانية على أساس كل رتل على حدة. كما يتيح استعمال توزيعات المعلومات عبر تتابع اختبار الفرصة للتقييم الأفضل للمشاهد بواسطة قطع المشاهد.

المعلومات الإدراكية المكانية (SI): هي مقياس يبيّن عموماً كمية التفاصيل المكانية في صورة. وهو أعلى عادةً بالنسبة للمَشاهد الأكثر تعقيداً من الوجهة المكانية. وليست الغاية منه أن يكون مقياساً للإنتروبيا ولا مصحوباً بمعلومات كتلك المعرفة في نظرية الاتصالات. وتستند المعلومات الإدراكية المكانية *SI* إلى مرشاح سوبل (Sobel). حيث يتم ترشيح كل رتل فيديوي (مستوي نصوع) في وقت n (F_n) يُرشّح أولاً بمرشاح سوبل [Sobel (F_n)]. بعدئذٍ يُحسب الانحراف المعياري عبر البيكسلات (std_{space}) في كل رتل تم ترشيحه بمرشاح سوبل. ويتم تكرار هذه العملية لكل رتل في تتابع فيديوي مما يُسفر عن سلسلة زمنية للمعلومات المكانية للمشهد. وتُختار القيمة القصوى في السلسلة الزمنية (max_{time}) بحيث تمثل محتوى المعلومات المكانية للمشهد. ويمكن تمثيل هذه العملية في صورة معادلة كما يلي:

$$SI = max_{time}\ \{std_{space}\ [\mathrm{Sobel}(F_n)]\}$$

المعلومات الإدراكية الزمانية (TI): هي مقياس يبيّن عموماً كمية التغيّرات الزمانية في تتابع فيديوي. وهو أعلى عادةً بالنسبة للتتابعات الأكثر حركة. وليست الغاية منه أن يكون مقياساً للانتروبيا ولا مصحوباً بمعلومات كتلك المعرفة في نظرية الاتصالات.

يُحسب مقياس المعلومات الزمنية *TI* باعتباره الحد الأقصى يتم حسابه على أنه الحد الأقصى بمرور الوقت (max_{time}) للانحراف المعياري عبر مكان (std_{space}) يخص $M_n(i, j)$ لكل i و j.

$$TI = max_{time}\ \{std_{space}\ [M_n(i, j)]\}$$

حيث $M_n(i, j)$ هو الفارق بين وحدتين بنائيتين للصورة في نفس الموقع من الرتل لكن تنتميان إلى رتلين متعاقبين، أي:

$$M_n(i, j) = F_n(i, j) - F_{n-1}(i, j)$$

حيث $F_n(i, j)$ هو وحدة الصورة في الصف رقم i والعمود رقم j من الرتل رقم n في الزمن.

ملاحظة – بالنسبة للمشاهد التي تحوي مقاطع للمشهد، يمكن إعطاء قيمتين: واحدة يُدرج فيها مقطع المشهد في مقياس المعلومات الزمانية، والأخرى حيث يُستبعد من القياس.

الملحق 7
(إعلامي)
بالجزء 1

مصطلحات وتعاريف

خوارزمية	واحدة أو أكثر من عمليات معالجة الصورة *(One or several image processing operations)*
AVI	إشارة سمعية فيديوية مشذَّرة *(Audio video interleaved)*
CCD	جهاز اقتران شحنات *(Charge coupled device)*
CI	الفاصل الزمني للثقة *(Confidence interval)*
CIF	نسق متوسط مشترك (معرَّف في التوصية ITU-T H.261 لهاتف الفيديو: 352 خط × 288 بكسل) *(Common intermediate format (defined in Recommendation ITU-T H.261 for video phone: 352 lines ´ 288 pixels))*
CRT	صمام الأشعة المهبطية *(Cathode ray tube)*
DSCQS	طريقة سلم النوعية المستمرة ثنائية الحافز *(Double stimulus using a continuous quality scale method)*
DSIS	طريقة سلم الانحطاط ثنائي الحافز *(Double stimulus using an impairment scale method)*
LCD	شاشة عرض بالبلورات السائلة *(Liquid crystal display)*
MOS	متوسط علامات الرأي *(Mean opinion score)*
SC	طريقة مقارنة الحافز *(Stimulus comparison method)*
PDP	لوحة عرض بلازمية *(Plasma display panel)*
PS	مقطع من برنامج *(Programme segment)*
QCIF	ربع النسق المتوسط المشترك (معرَّف في التوصية ITU-T H.261 لهاتف الفيديو: 176 خط × 144 بكسل) *(Quarter CIF (defined in Recommendation ITU-T H.261 for video phone: 176 lines ´ 144 pixels))*
SAMVIQ	طريقة التقدير الشخصاني لنوعية الفيديو المتعدد الوسائط *(Subjective assessment of multimedia video quality)*
تتابع (Sequence)	مشهد مع معالجة مدمجة أو بدون معالجة *(Scene with combined processing or without processing)*
مشهد (Scene)	محتوى صوتي مرئي *(Audiovisual content)*
S/N	نسبة الإشارة إلى الضوضاء *(Signal-to-noise ratio)*
SI	معلومات مكانية *Spatial information*
SIF	نسق متوسط معياري [معرّف في المعيار ISO 11172 (MPEG-1): 352 خط × 288 بكسل × 25 رتل/ث 352 خط × 240 بكسل × 30 رتل/ث] *(Standard intermediate format [defined in ISO 11172 (MPEG-1): 352 lines ´ 288 pixels ´ 25 frames/s and 352 lines ´ 240 pixels ´ 30 frames/s])*
SP	عرض متزامن *(Simultaneous presentation)*
SQCIF	ربع النسق المتوسط المشترك الفرعي *(Sub-QCIF)*
SS	الطرائق وحيدة الحافز *(Single stimulus method)*

SSCQE	طريقة تقييم النوعية المستمرة وحيدة الحافز *(Single stimulus using a continuous quality evaluation method)*
std	انحراف معياري *(Standard deviation)*
TI	معلومات زمانية *(Temporal information)*
TP	عرض اختبار *(Test presentation)*
TS	جلسة اختبار *(Test session)*
VTR	مسجّلة شريط فيديو *(Video tape recorder)*

الجزء 2

وصف المنهجيات الشخصانية لتقدير الصورة

جدول المحتويات

الصفحة

الجزء 2 41

وصف المنهجيات الشخصانية لتقدير الصورة 41

1 مقدمة 43

2 منهجيات تقدير الصورة الموصى بها 43

3 ملاحظات 44

الملحق 1 بالجزء 2 – طريقة سلم الانحطاط ثنائي الحافز (DSIS) ("طريقة EBU") 44

A1-1 الوصف العام 44

A1-2 الترتيب العام للاختبار 44

A1-3 تقديم مواد الاختبار 45

A1-4 سلالم إسناد الدرجات 45

A1-5 المدخل إلى التقديرات 46

A1-6 دورة الاختبار 46

الملحق 2 بالجزء 2 – طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) 47

A2-1 الوصف العام 47

A2-2 الترتيب العام 47

A2-3 عرض مواد الاختبار 47

A2-4 سلم إسناد الدرجات 48

A2-5 تحليل النتائج 50

A2-6 تفسير النتائج 50

الملحق 3 بالجزء 2 – الطرائق وحيدة الحافز (SS) 50

A3-1 الترتيب العام للاختبار 50

A3-2 انتقاء مواد الاختبار 51

A3-3 جلسة الاختبار 51

A3-4 أنماط الطرائق وحيدة الحافز 51

A3-4.1 طرائق الحكم حسب الفئات بواسطة الصفات 52

A3-4.2 طرائق الأحكام العددية حسب الفئات 52

3.4-A3 طرائق الأحكام التي لا تتوزع حسب الفئات 52
4.4-A3 طرائق الأداء 53
الملحق 4 بالجزء 2 – طرائق مقارنة الحوافز 53
1-A4 الترتيب العام للاختبار 53
2-A4 انتقاء مواد الاختبار 53
3-A4 جلسة الاختبار 53
4-A4 أنماط طرائق مقارنة الحوافز 53
1.4-A4 طرائق الحكم حسب الفئات بواسطة الصفات 54
2.4-A4 طرائق الأحكام التي لا تتوزع حسب الفئات 54
3.4-A4 طرائق الأداء 54
الملحق 5 بالجزء 2 – طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE) 55
1-A5 جهاز التسجيل وضبطه 55
2-A5 الشكل العام لبروتوكول الاختبار 55
3-A5 معلمات المشاهدة 55
4-A5 سلالم إسناد الدرجات 56
5-A5 المراقبون 56
6-A5 تعليمات إلى المراقبين 56
7-A5 عرض البيانات ومعالجة النتائج وتقديمها 56
8-A5 معايرة نتائج النوعية المستمرة واستنتاج تقييم عام للنوعية 56
الملحق 6 بالجزء 2 – طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE) 57
1-A6 إجراء الاختبار 58
2-A6 المراحل المختلفة 58
3-A6 خصائص بروتوكول الاختبار 59
4-A6 معالجة البيانات 59
5-A6 موثوقية الأشخاص 60
الملحق 7 بالجزء 2 – طريقة التقدير الشخصاني لنوعية الفيديو المتعدد الوسائط (SAMVIQ) 62
1-A7 مقدمة 62
2-A7 المرجع الصريح والمخفي والخوارزميات 63
3-A7 شروط الاختبار 63
4-A7 تنظيم الاختبار 63
5-A7 عرض وتحليل البيانات 64
1.5-A7 معلومات موجزة 64
2.5-A7 طرائق التحليل 65
3.5-A7 فرز المراقبين 65

6-A7 مثال عن سطح بيني لطريقة تقدير SAMVIQ (إعلامي) 66
الملحق 8 بالجزء 2 – بروتوكول مشاهدة الخبراء (EVP) لتقييم نوعية المادة الفيديوية 66
1-A8 إعدادات المختبر 66
1.1-A8 اختيار شاشة العرض وإعدادها 66
2.1-A8 مسافة المشاهدة 67
3.1-A8 شروط المشاهدة 67
2-A8 المشاهدون 67
3-A8 خلية الاختبار الأساسية 67
4-A8 ورقة إسناد الدرجات وسلُّم التصنيف 68
5-A8 تصميم الاختبار وإنشاء الجلسة 70
6-A8 التدريب 70
7-A8 جمع البيانات ومعالجتها 70
8-A8 شروط استعمال نتائج بروتوكول مشاهدة الخبراء 70
9-A8 قيود استعمال نتائج بروتوكول مشاهدة الخبراء (EVP) 71
المرفق 1 (إعلامي) بالملحق 8 بالجزء 2 تطبيق بروتوكول مشاهدة الخبراء وسلوكهم في حضور عدد كبير من المقيّمين الخبراء 71

1 مقدمة

يقدم هذا الجزء تفاصيل كل من منهجيات تقدير الصورة المطلوبة لإجراء تقديرات شخصانية لنوعية الصورة. وفي بعض الحالات، يؤدي ذلك إلى تغيير الخصائص المشتركة للتقدير الواردة في الفقرة 2 من الجزء 1.

ومن أجل ضمان إمكانية تفسير المختبرات الأخرى لنتائج التقديرات الشخصانية لنوعية الصورة تفسيراً صحيحاً، من المهم توفر ملاحظات تفصيلية للإجراءات وتسجيل أي تغيير في المنهجية المستعملة مع جميع المعلومات الإضافية التي قد تحتاج إليها المختبرات الأخرى التي ترغب في تكرار إجراء التقدير.

2 منهجيات تقدير الصورة الموصى بها

الملحق 1 طريقة سلم الانحطاط ثنائي الحافز (DSIS)

الملحق 2 طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS)

الملحق 3 الطرائق وحيدة الحافز (SS)

الملحق 4 طرائق مقارنة الحوافز

الملحق 5 طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE)

الملحق 6 طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE)

الملحق 7 طريقة التقدير الشخصاني لنوعية الفيديو المتعدد الوسائط (SAMVIQ)

الملحق 8 بروتوكول مشاهدة الخبراء (EVP) لتقييم نوعية المادة الفيديوية

3 ملاحظات

هنالك تقنيات أخرى مثل طرائق التقدير متعددة الأبعاد والطرائق المتنوعة ترد في التقرير ITU-R BT.1082 وما تزال قيد الدراسة.

وجميع الطرائق المتداولة حتى الآن نقاط لها قوتها ونقاط ضعفها ومن غير الممكن بعد التوصية بشكل قاطع بإحداها غير الأخرى. ولذا يعود انتقاء الطرائق الأكثر ملائمةً للحالات المطروحة إلى رغبة الباحث عنها.

أما نقاط ضعف مختلف الطرائق فتكمن في أنها لا تنصح كثيراً بالاعتماد على طريقة واحدة. ولذا فقد يكون من المناسب اتباع نُهج "كاملة" من قبيل استعمال عدة طرائق أو استعمال النهج متعدد الأبعاد.

الملحق 1
بالجزء 2

طريقة سلم الانحطاط ثنائي الحافز (DSIS)
("طريقة EBU")

A1-1 الوصف العام

قد يستدعي تقدير نمطي تقييم نظام جديد أو تقييم تأثير انحطاط في مسير الإرسال. وتكمن الخطوات الأولى التي يقوم بها منظم الاختبار في انتقاء مواد اختبار كافية لتسمح بتقدير دلالي، وفي تحديد شروط الاختبار الواجب استعمالها. وإذا كان تأثير تغير المعلمة مهماً، يجب اختيار مجموعة من قيم المعلمات تغطي درجات الانحطاط كلها في عدد صغير من الدرجات المتساوية تقريباً. وإذا كان التقدير يتعلق بنظام جديد لا يمكن أن تغير فيه قيم المعلمات، يجب عندها إضافة انحطاطات أخرى مماثلة في التقدير الشخصاني، أو استعمال طريقة أخرى مثل الطريقة المقدمة في الملحق 2 بالجزء 2.

وطريقة سلم الانحطاط ثنائي الحافز (DSIS) (طريقة EBU) دورية، أي أن المشاهد يرى في البداية مرجعاً لم يتعرض للانحطاط، ثم تقدم الصورة نفسها منحطة. ويطلب منه، بعد ذلك، أن يقدر الصورة الثانية، مع إبقاء الأولى في ذهنه. وتقدم للمشاهد في أثناء كل دورة التي لا تتجاوز نصف الساعة، سلسلة من الصور أو التتابعات وفقاً لترتيب عشوائي ومع انحطاطات عشوائية تشمل كل التركيبات المطلوبة. وتدرج الصورة التي لم تتعرض للانحطاط داخل الصور أو التتابعات المطلوب تقديرها. ويحسب، في نهاية سلسلة الدورات، متوسط العلامات لكل حالة اختبار وكل صورة في الاختبار.

تستعمل الطريقة سلم الانحطاط الذي يعطي عادةً نتائج أكثر استقراراً للانحطاطات الضعيفة من نتائج الانحطاطات الكبيرة. ومع أن الطريقة قد استعملت أحياناً مع مديات محدودة من الانحطاط، إلا أنه من الأنسب استعمالها مع مدى كامل من الانحطاط.

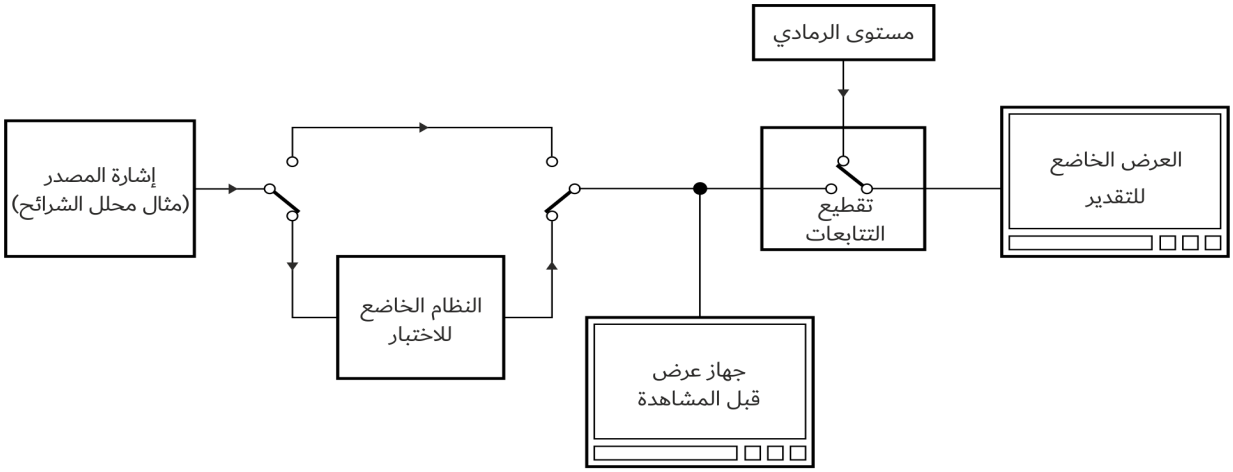
A1-2 الترتيب العام للاختبار

تعرف شروط المشاهدة، وإشارات المصدر ومواد الاختبار والمراقبون وتقديم النتائج، أو تنتقى وفقاً للفقرة 2 من الجزء 1.

ويكون الترتيب العام لنظام الاختبار على النحو المبين في الشكل 2-1.

الشكل 2-1

الترتيب العام لنظام الاختبار في طريقة سلم الانحطاط ثنائية الحافز



BT.0500-02-1

تقدم للمشاهدين شاشة عرض للتقدير تزوَّد بإشارة تمر عبر مفتاح زمني مؤقِّت. ويمكن لمسير الإشارة إلى المفتاح الزمني المؤقِّت أن يكون مباشراً من الإشارة المصدر، أو غير مباشر عبر النظام الخاضع للاختبار. تقدم للمشاهدين سلسلة من الصور أو تتابعات الاختبار. وترتب على أزواج بحيث تأتي الصورة الأولى من الزوج من المصدر مباشرة، بينما تكون الثانية الصورة نفسها تمر عبر النظام الخاضع للاختبار.

A1-3 تقديم مواد الاختبار

تحتوي دورة الاختبار على عدة تقديمات. وثمة صيغتان مختلفتان لبنية التقديمات I وII تعرض فيما يلي:

الصيغة I: تقدَّم الصورة المرجعية أو التتابع المرجعي وصورة الاختبار أو التسلسل مرة واحدة فقط كما هو مبين في الشكل 2-2(أ).

الصيغة II: تقدَّم الصورة المرجعية أو التتابع المرجعي وصورة الاختبار أو التسلسل مرتين كما هو مبين في الشكل 2-2(ب).

الصيغة III: التي تستهلك وقتاً أطول من الصيغة I، يمكن تطبيقها إذا توجب التمييز بين انحطاطات صغيرة جداً، أو إذا خضعت تتابعات متحركة للاختبار.

A1-4 سلالم إسناد الدرجات

يُستعمل سلم الانحطاط خماسي الدرجات:

5 لا يمكن إدراكه

4 يمكن إدراكه، ولكنه غير مزعج

3 مزعج قليلاً

2 مزعج

1 مزعج جداً.

يجب أن يستعمل المشاهدون استمارةً تمثل السلم تمثيلاً واضحاً جداً، مع مربعات مرقمة أو أية وسيلة أخرى لتسجيل درجاتها.

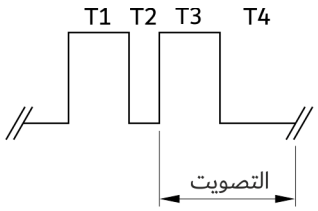
5-A1 المدخل إلى التقديرات

يقدم للمراقبين، في بداية كل دورة، تفسير عن نمط التقدير وسلم إسناد الدرجات والتتابع والتوقيت (الصورة المرجعية والرماديات وصورة الاختبار وفترة التصويت). ويجب أن يوضح مدى الانحطاطات الواجب تقديرها ونمطها على صور غير الصور المستعملة في الاختبار لكن ذات حساسية مماثلة. ويجب ألا يعني ذلك أن أسوأ نوعية مرئية تقابل بالضرورة أدنى درجة شخصانية. ويُطلب من المراقبين أن يسندوا أحكامهم إلى الانطباع العام الذي تعطيه الصورة، ويعبروا عن هذه الأحكام بواسطة العبارات المستعملة لتعريف سلم التقديرات الشخصانية.

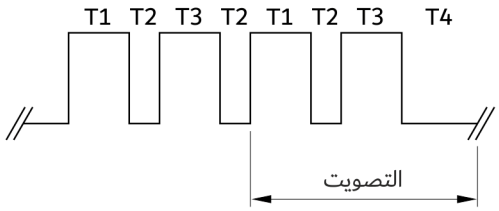
ويُطلب من المراقبين أن ينظروا إلى الصورة أثناء المدة الكاملة من أجل T1 وT3. ويسمح بالاقتراع في أثناء T4 فقط.

الشكل 2-2

بنية تقديم مواد الاختبار



أ) الصيغة I



ب) الصيغة II

BT.0500-02-2

مراحل العرض:

T1 =	10 ثوان	الصورة المرجعية
T2 =	3 ثوان	رمادي متوسط صادر عن مستوى فيديوي يقارب 200 mV
T3 =	10 ثوان	شروط الاختبار
T4 =	5-11 ثانية	رمادي متوسط

أظهرت التجربة أن تمديد الفترتين T1 وT3 إلى أكثر من 10 ثوان لا يحسِّن من قدرة المشاهدين على تقدير الصور أو التتابعات.

6-A1 دورة الاختبار

ينبغي عرض الصور والانحطاطات في تتابع شبه عشوائي ويفضل أن يكون التتابع مختلفاً في كل جلسة. وينبغي، في كل الأحوال، ألا تقدم أبداً الصورة أو التتابعات نفسها في مناسبتين متتابعتين مع نفس سوية الانحطاط أو مع سويات مختلفة.

وينبغي اختيار مدى الانحطاطات على نحو يستعمل فيه أغلبية المراقبين كل الدرجات؛ وأن يكون الهدف التوصل إلى متوسط عام للعلامات (تقدير متوسط كل الأحكام المقدمة في التجربة) قريب من 3.

وينبغي ألا تدوم الجلسة أكثر من نصف ساعة تقريباً، بما في ذلك التفسيرات والعروض التمهيدية؛ ويمكن أن يبدأ تتابع الاختبار ببعض الصور التي تدل على مدى الانحطاطات؛ ولا تؤخذ الأحكام المتعلقة بهذه الصور في الحسبان في النتائج النهائية.

ويعطي الملحق 2 بالجزء 1 توضيحات إضافية من أجل اختيار سويات الانحطاط.

الملحق 2
بالجزء 2

طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS)

A2-1 الوصف العام

قد يستدعي التقدير النمطي تقييم نظام جديد، أو تأثيرات مسير الإرسال على النوعية. ويُعتقد أن الطريقة ثنائية الحافز مفيدة بشكل خاص، إذا تعذر توفير شروط اختبار الحافز التي تعرض المدى الكامل للنوعية.

وهذه الطريقة دورية أي أنه يقدَّم للمشاهد زوج من الصور من المصدر نفسه لكن الأولى تمر عبر النظام الخاضع للدراسة بينما تأتي الأخرى من المصدر مباشرة. ويُطلب منه تقدير نوعية الصورتين.

يقدم للمشاهد، في جلسات لا تدوم أكثر من نصف الساعة، سلسلة من أزواج الصور (بترتيب داخلي عشوائي) تقديماً عشوائياً ومع انحطاطات عشوائية تشمل كل التركيبات المطلوبة. ويحسب في نهاية الجلسات متوسط العلامات لكل حالة اختبار ولكل صورة من الاختبار.

A2-2 الترتيب العام

تُعرَّف شروط المشاهدة وإشارات المصدر ومواد الاختبار والمراقبون والمدخل إلى التقدير، أو تنتقى وفقاً للفقرة 2 من الجزء 1. ويصف القسم A1.6 في الملحق 1 بالجزء 2 جلسة الاختبار.

يكون الترتيب العام لنظام الاختبار على النحو المبين في الشكل 2-3 أدناه.

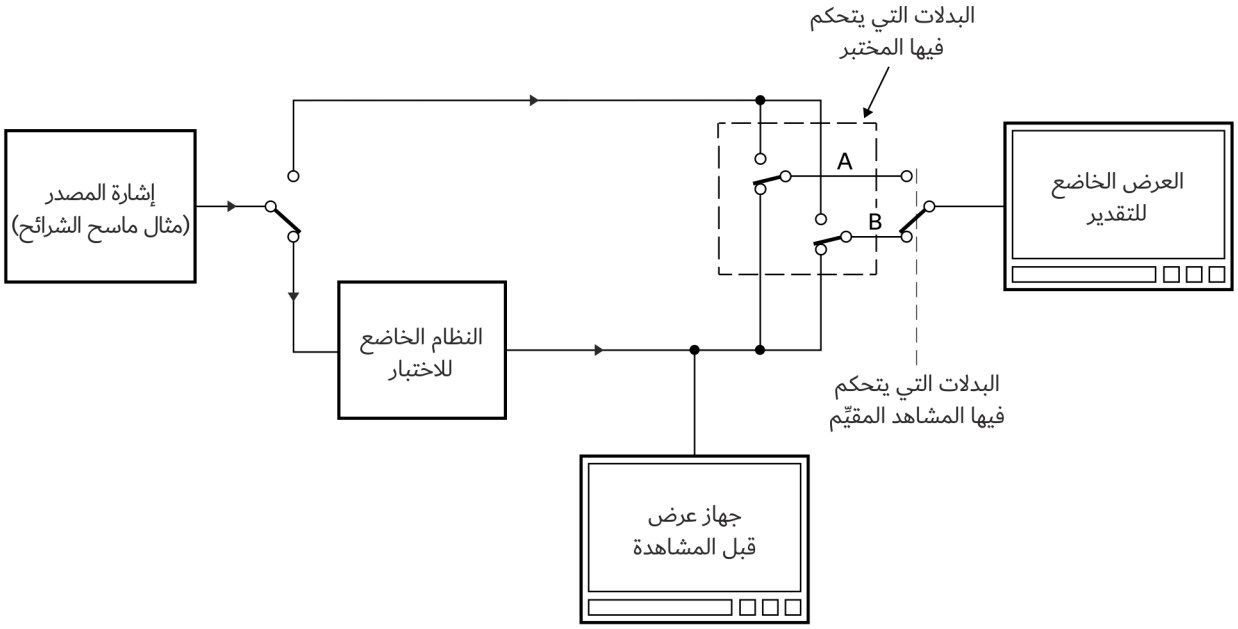
A2-3 عرض مواد الاختبار

تضم جلسة الاختبار عدة عروض. ففي حالة الصيغة I التي تتطلب مراقباً واحداً، يستطيع المشاهد أن يختار لكل عرض التبديل بين الإشارتين A وB حتى يمتلك القياس الذهني للنوعية المصاحبة لكل إشارة. ويستطيع المشاهد بشكل عام تكرار ذلك مرتين أو ثلاث مرات لفترات لا تتجاوز 10 ثوان. أما في الصيغة II التي تستعمل عدة مراقبين في آن معاً، وقبل تسجيل النتائج، يُقدَّم كل زوج من الحالات مرة أو عدة مرات لفترات زمنية متساوية لكي يسمح للمشاهد بأن يقيس ذهنياً النوعية المصاحبة لهذه الحالات، ثم يُقدَّم الزوج مرة أخرى أو عدة مرات بينما يُصار إلى تسجيل النتائج. ويتعلق عدد العروض المكررة بطول تتابعات الاختبار. ويبدو أن تتابعاً من 3-4 ثوان و5 عروض مكررة (الاقتراع أثناء العرضين الأخيرين)، هو مناسب للصور الثابتة. ويكون تتابع من 10 ثوان مع تكرارين (الاقتراع أثناء العرض الثاني) مناسباً للصور المتحركة التي تتعرض لظواهر اصطناعية تتغير زمنياً. ويبين الشكل 2-4 بنية العروض.

وعندما تحدد اعتبارات عملية مدة التتابعات المتيسرة بأقل من 10 ثوان، يمكن اللجوء إلى تركيبات تستعمل هذه التتابعات الأقصر على شكل قطع من أجل تمديد فترة العرض إلى 10 ثوان. ويمكن أن تعكس في الزمن قطع التتابعات المتتالية (يسمى أحياناً "عرضاً يمكن تقديمه طرداً وعكساً" "palindromic") لكن تخفض إلى أدنى حد ممكن عدم الاستمرارية عند التوصيلات. إلا أن من الضروري التأكد من أن حالات الاختبار المعروضة كقطع معكوسة زمنياً تمثل عمليات سببية، أي أنه يجب الحصول عليها بواسطة تمرير إشارة المصدر معكوسة زمنياً عبر النظام الخاضع للاختبار.

الشكل 3-2

الترتيب العام لنظام الاختبار في طريقة سلم النوعية المستمرة ثنائية الحافز



BT.0500-02-3

وهناك صيغتان لهذه الطريقة على النحو الوارد أدناه:

الصيغة I يُسمح للمقيّم، الذي عادةً ما يكون بمفرده، بالتبديل بين الحالتين A وB لتكوين رأي راسخ عن كل منهما. ويجري تزويد الخطين A وB بالصورة المرجعية أو صورة عبر النظام قيد الاختبار، ولكن يجري تغيير الصورة الداخلة إلى الخط عشوائياً بين إحدى حالات الاختبار والحالة التالية. ويلاحظ المختبر هذه الحالة ولكن لا يكون المقيّم على علم بها.

الصيغة II يُعرض على المقيّمين الصور الواردة من الخطين A وB على التوالي من أجل تكوين رأي. وتجري تغذية الخطين A وB لكل عرض على النحو الموضح في الصيغة I أعلاه.

4-A2 سلم إسناد الدرجات

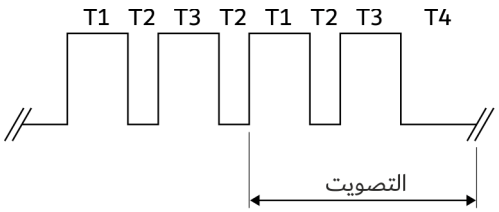
تتطلب الطريقة تقدير صيغتين لكل صورة اختبار. وفي كل زوج من الصور، تكون إحدى الصور غير منحطة بينما يمكن أن تتضمن الأخرى أو لا تتضمن انحطاطاً. تدرج الصورة غير المنحطة كصورة مرجعية، لكن المراقبين يجهلون أية صورة هي الصورة المرجعية. ويتغير موقع الصورة المرجعية في سلسلة الاختبارات على نحو شبه عشوائي.

ويُطلب فقط من المراقبين تقدير نوعية الصورة الكلية في كل تقديم من خلال إدراج علامة على سلم رأسي. تُطبع السلالم الرأسية مزاوجة لمراعاة التقديم المزدوج لكل صورة اختبار. وتوفر السلالم نظام تقدير مستمر يسمح بتجنب أخطاء التكمية، لكنها مقسمة إلى خمسة مقاطع متساوية الطول تقابل سلم النوعية العادي خماسي الدرجات الذي وضعه قطاع الاتصالات الراديوية. وتستعمل العبارات

نفسها التي تستعمل عادةً لتمييز السويات المختلفة؛ لكنها تُدرج من أجل التوجيه العام وتُطبع فقط إلى يسار أول سلم في كل صف من 10 أعمدة مزدوجة على صفحة العلامات. ويبين الشكل 5-A2 جزءاً من صفحة علامات نمطية. وتُستبعد أية إمكانية التباس بين تقسيمات السلم ونتائج الاختبار بطبع السلالم بالأزرق بينما تسجل النتائج بالأسود.

الشكل 2-4

بنية عرض مواد الاختبار



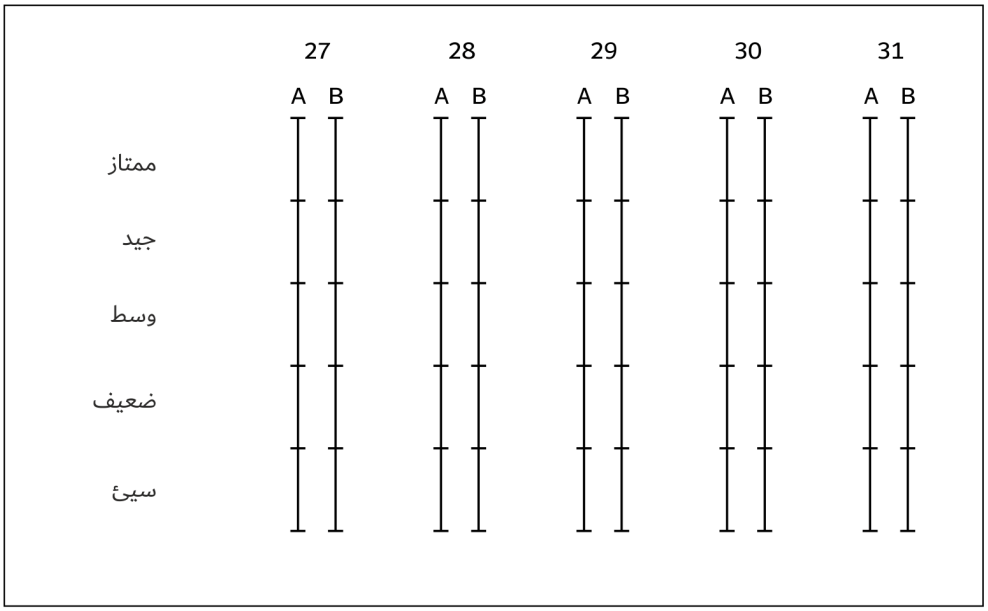
BT.0500-02-4

مراحل العرض

T1 =	10 ثوان	صورة مرجعية
T2 =	3 ثوان	رمادي متوسط صادر عن مستوى فيديوي يقارب 200 mV
T3 =	10 ثوان	شروط الاختبار
T4 =	5-11 ثانية	رمادي متوسط

الشكل 2-5

جزء من استمارة تقدير النوعية التي تستعمل سلالم مستمرة*



BT.0500-02-5

* لدى ترتيب مواد الاختبار ضمن جلسة اختبار بطريقة DSCQS، يستحسن أن يدرج المختبر بعض عمليات التحقق التي توحي بالثقة في أن الاختبار لا تشوبه أخطاء منهجية. إلا أن طريقة إجراء عمليات التحقق هذه ما تزال قيد البحث.

5-A2 تحليل النتائج

في عمليتي تقدير (الصورة المرجعية وصورة الاختبار) كل شرط اختبار يجري تحويل قياسات طول الخط المرسوم على ورقة العلامات إلى علامات معيارية تقع بين 0 و100، ثم تحسب بعد ذلك الفروق بين تقدير صورة المرجع وصورة الاختبار. وترد الإجراءات التفصيلية في الملحق 2 بالجزء 1.

وقد بينت التجربة أن العلامات الناتجة لتتابعات الاختبار المختلفة تتوقف على درجة حساسية تتابعات صور الاختبار المستعملة. ومن الممكن الحصول على معرفة أدق لأداء الكودك لو أن نتائج التتابعات المختلفة للاختبار عرضت منفصلةً وليس على شكل متوسطات مجمعة لمجمل التتابعات المستعملة في التقييم وحسب.

وفي حال ترتيب نتائج تتابعات الاختبار المتفرقة في إحداثيات سينية حسب درجة حساسية تتابع الاختبار، من الممكن تقديم وصف بياني خام بخصائص خلل محتوى الصورة في النظام الخاضع للاختبار. لكن هذا النوع من التقديم لا يصف إلاّ أداء الكودك ولا يعطي أي دلالة عن احتمال حدوث تتابعات تمثل درجة الحساسية المعنية (انظر الملحق 2 بالجزء 1). ويجب إجراء مزيد من الدراسات عن حساسية تتابعات الاختبار واحتمال حدوث تتابعات بدرجة ما من الحساسية قبل تكوين فكرة كاملة عن خصائص النظام.

6-A2 تفسير النتائج

عند استعمال طريقة DSCQS قد يكون من الخطر ولربما من الخطأ التوصل إلى استنتاجات عن نوعية الشروط الخاضعة للاختبار من خلال ربط القيم الرقمية الناتجة عن هذه الطريقة بصفات ناتجة عن بروتوكولات اختبار أخرى (مثل، "لا يمكن إدراكه" أو "يمكن إدراكه، ولكنه غير مزعج" وغيرها من الصفات المستعملة في طريقة DSIS).

ويجدر بالذكر ضرورة عدم اعتبار نتائج طريقة DSCQS علامات مطلقة ولكن تباين علامات بين الصورة المرجعية وصورة الاختبار. ولذا فمن غير المستصوب ربط العلامة بصفة جودة واحدة حتى ولو كانت من صفات هذه الطريقة (مثل ممتاز، جيد، لا بأس وغيرها).

ومن الهام عند إجراء أي اختبار اعتماد معايير القبول قبل البدء بالتقدير. وتزداد تلك الأهمية عند استعمال طريقة DSCQS بسبب احتمال سوء فهم المستعملين من غير ذوي الخبرة لمعنى قيم تدرج النوعية التي تنتجها هذه الطريقة.

الملحق 3
بالجزء 2

الطرائق وحيدة الحافز (SS)

تقدم في الطرائق وحيدة الحافز صورة وحيدة أو تتابع من الصور ويقدم المشاهد مؤشراً عن العرض الكامل. وقد لا تضم مواد الاختبار سوى تتابعات الاختبار أو قد تضم تتابعات الاختبار وتتابعها المرجعي على حد سواء. وفي هذه الحالة، يعرض التتابع المرجعي بوصفه حافزاً مفتوحاً للتقدير مثل أي حافز اختبار آخر.

1-A3 الترتيب العام للاختبار

تعرف شروط المشاهدة وإشارات المصدر، ومدى الشروط وإرسائها والمراقبون، والمدخل إلى التقديرات، وتقديم النتائج، أو تنتقى وفقاً للقسم 2 من الجزء 1.

2-A3 انتقاء مواد الاختبار

يجب أن ينتقى محتوى صور الاختبار للتجارب في المختبر، وفقاً للوصف المقدم في القسم 3.2 من الجزء 1.

بعد انتقاء المحتوى، تحضر صور الاختبار لتعكس خيارات التصميم المعنية أو مدى (أو مديات) عامل واحد أو عدة عوامل. وعندما يدرس عاملان أو عدة عوامل، يمكن تحضير الصور وفقاً لطريقتين مختلفتين. تَعرض كل صورة، في الحالة الأولى، مستوى واحداً لعامل واحد فقط. وتَعرض في الحالة الأخرى مستوى واحداً لكل عامل من العوامل المدروسة، لكن مع عرض الصور يصادف كل مستوى من كل عامل كل مستوى لجميع العوامل الأخرى. وتسمح الطريقتان بمعرفة واضحة للنتائج الخاصة بعامل محدد. بينما تسمح أيضاً الطريقة الثانية بكشف التفاعلات فيما بين العوامل (أي التأثيرات غير المضافة).

3-A3 جلسة الاختبار

تتضمن جلسة الاختبار سلسلة من اختبارات التقدير التي تقدم عشوائياً، ويفضل أن يكون التتابع العشوائي مختلفاً لكل مراقب. ولدى استعمال الترتيب العشوائي الوحيد للتتابعات توجد صيغتان لعرض البنية هما الصيغة I (SS) والصيغة II (حافز وحيد مع تكرار متعدد (SSMR)) على النحو التالي:

أ) تعرض صور الاختبار أو تتابعاته مرة واحدة في جلسة الاختبار؛ وفي بداية الجلسات الأولى ينبغي إدخال بعض التتابعات الزائفة (وفقاً لما يرد في القسم 7.2 من الجزء 1)؛ وتؤكد التجربة عادةً على أن لا تعرض نفس الصورة لمرتين متتاليتين بنفس مستوى الانحطاط.

وتتألف محاولة التقدير النمطية من ثلاثة عروض هي: صورة مكيفة رمادي متوسط وصورة حافز وصورة جديدة بالرمادي المتوسط. وتختلف مدة هذه العروض باختلاف مهمة المشاهدة والمواد والآراء أو العوامل ذات الصلة، لكن تستعمل عادةً مدد 3 و10 و10 ثوان على التوالي. وقد تُجمع مؤشرات المشاهدين خلال عرض الحافز أو صورة الثانية الجديدة.

ب) تُعرض صورة الاختبار أو تتابعاته ثلاث مرات خلال الجلسة مقسمة بذلك هذه الجلسة إلى ثلاثة عروض يحتوي كل منها على جميع الصور أو التتابعات الواجب اختبارها مرة واحدة فقط؛ ويعلن عن بدء كل عرض في رسالة على شاشة العرض (مثال العرض 1)؛ الغرض من العرض الأول هو استقرار رأي المراقب؛ ويجب عدم إدخال البيانات الواردة في هذا العرض في الحسبان في نتائج الاختبار؛ ويتم الحصول على العلامات التي تحوز عليها الصورة أو التتابعات من خلال تقدير متوسط البيانات الواردة في العرضين الثاني والثالث؛ وتؤكد التجربة عادةً على تطبيق الحدود التالية على الترتيب العشوائي للصور أو التتابعات في كل عرض:

- عدم ورود صورة أو تتابع ما في نفس الوضعية في العروض الأخرى؛
- عدم ورود صورة أو تتابع ما مباشرةً قبل نفس الصورة أو التتابع في العروض الأخرى.

وتنطوي محاولة التقدير النمطية على عرضين: صورة حافزة وصورة جديدة بالرمادي المتوسط. وقد تختلف مدة عرض هاتين الصورتين باختلاف مهمة المشاهد أو المادة المختبرة والآراء أو العوامل المكونة، لكن تقترح المادة 10 و5 ثوان على التوالي. ولا تُجمع أدلة المشاهد إلا خلال عرض الصورة الثانية.

وتُدخل الصيغة II (SSMR) سقفاً واضحاً في المدة المحددة لإجراء جلسة الاختبار (45 ثانية مقابل 23 ثانية لكل صورة أو تتابع تخضع للاختبار؛ ومع ذلك فإن ذلك يقلل كثيراً من ارتباط نتائج الصيغة 1 بترتيب الصورة أو التتابعات داخل الجلسة.

وعلاوةً على ذلك، تظهر نتائج التجارب أن الصيغة II تتيح هامشاً يقارب 20% داخل مدى التقديرات.

4-A3 أنماط الطرائق وحيدة الحافز

استُعملت، بشكل عام، ثلاثة أنماط وحيدة الحافز في التقديرات التلفزيونية.

1.4-A3 طرائق الحكم حسب الفئات بواسطة الصفات

يخصص المراقبون، في الأحكام بالصفات حسب الفئات، صورة أو تتابع صور إلى فئة واحدة أو إلى مجموعة من الفئات تعرف نمطياً وفقاً لقواعد علم الدلالات. وقد تعكس الفئات الحكم على وجود نعت معين أو غيابه (من أجل تحديد عتبة الانحطاط، على سبيل المثال). وقد استعملت سلالم الفئات التي تقدر نوعية الصورة وانحطاطها، في غالبية الحالات. ويقدم الجدول 2-1 أدناه السلالم التي وضعها قطاع الاتصالات الراديوية. وتستعمل أحياناً في العرض التشغيلي أنصاف الدرجات. وقد استعملت في بعض الحالات الخاصة سلالم تقدر إمكانية قراءة النص، وجهد القراءة وفائدة الصورة.

الجدول 2-1

سلالم النوعية والانحطاط التي وضعها قطاع الاتصالات الراديوية

سلم خماسي الدرجات			
النوعية		**الانحطاط**	
5	ممتاز	5	لا يمكن إدراكه
4	جيد	4	يمكن إدراكه، ولكنه غير مزعج
3	وسط	3	مزعج قليلاً
2	ضعيف	2	مزعج
1	سيء	1	مزعج جداً

تؤدي هذه الطريقة إلى توزيع الأحكام على فئات في السلم بالنسبة إلى كل حالة. وتتعلق طريقة تحليل الإجابات بالحكم نفسه (الكشف، إلخ.)، وبالمعلومات المطلوبة (عتبة الكشف، ترتيب الحالات أو متوسط الاتجاهات، "المسافات" النفسية فيما بين الحالات). وهناك طرائق تحليل كثيرة متيسرة.

2.4-A3 طرائق الأحكام العددية حسب الفئات

دُرِسَ الإجراء وحيد الحافز الذي يستعمل سلماً عددياً حسب الفئات يتضمن 11 درجة (SSNCS)، وقُورِن بالسلالم البيانية وسلالم النسب. وتُظهر هذه الدراسة التي يعرضها التقرير ITU-R BT.1082، تفضيلاً واضحاً للطريقة SSNCS من ناحية الحساسية والاستقرار، إذا لم يتيسر أي مرجع.

3.4-A3 طرائق الأحكام التي لا تتوزع حسب الفئات

يخصص المراقبون، وفقاً لهذه الطرائق، قيمةً لكل صورة أو كل تتابع صور معروضين. وثمة صيغتان مختلفتان لهذه الطريقة.

يخصص المشاهد، في صيغة السلم المستمر الذي يعتبر صيغة من طريقة توزيع الأحكام حسب الفئات، لكل صورة أو كل تتابع صور نقطة على خط يُرسم بين وسمين في علم الدلالات (طرفي سلم حسب الفئات كما في الجدول 2-1، على سبيل المثال). وقد يتضمن السلم وسوماً إضافية عند نقاط متوسطة، تعتبر مرجعية. وتستعمل المسافة من طرف من السلم دليلاً لكل حالة.

أما في صيغة السلم العددي فيخصص المشاهد لكل صورة أو لكل تتابع صور عدداً يعكس سوية حكمه على عامل محدد (وضوح الصورة مثلاً). ويمكن أن يكون مدى الأعداد المستعملة مقيداً، (من 0 إلى 100 مثلاً) أو غير مقيد. ويصف أحياناً العدد المخصص السوية المقدرة بعبارات "مطلقة" (دون الرجوع المباشر إلى سوية أية صورة أخرى أو تتابع آخر كما يحصل في بعض أشكال تقدير الاتساع). ويصف هذا العدد، في حالات أخرى، السوية المقدرة بالنسبة إلى سوية اعتبرت سابقاً سوية "معيارية" (مثل تقدير الاتساع والتقطيع وتقدير النسب).

وتؤدي الصيغتان إلى توزيع للأعداد لكل حالة. وتتعلق طريقة التحليل المستعملة بنمط التقدير وبالمعلومات المطلوبة (مثل الصفوف والاتجاه المركزي، و"المسافات" النفسية).

4.4-A3 طرائق الأداء

يمكن التعبير عن بعض جوانب المشاهدة العادية بالنسبة إلى أداء مهمات محددة خارجياً (البحث عن المعلومات المستهدفة، قراءة النص، تعرف هوية الأشياء، إلخ.). ويمكن بعدئذٍ أن يُستعمل قياس أداء معين، مثل الدقة أو السرعة التي تنفذ بها المهمات، دليلاً للصورة أو لتتابع الصور.

تؤدي طرائق الأداء إلى توزيعات لعلامات الدقة أو السرعة لكل حالة. ويتركز التحليل على إقامة علاقات فيما بين الحالات في الاتجاه المركزي (والتشتت) للعلامات، ويستعمل غالباً تحليل التغير أو تقنية مماثلة.

الملحق 4
بالجزء 2

طرائق مقارنة الحوافز

تُعرض في طرائق مقارنة الحوافز صورتان أو تتابعان من الصور، ويعطي المشاهد دليلاً عن العلاقة بين التقديمين.

1-A4 الترتيب العام للاختبار

تعرف شروط المشاهدة والإشارات المصدر، ومدى الشروط وكيفية إرسائها والمراقبون والمدخل إلى التقدير وتقديم النتائج أو تنتقى وفقاً للقسم 2 من الجزء 1.

2-A4 انتقاء مواد الاختبار

تولَّد الصور أو تتابعات الصور بالطريقة نفسها التي تولَّد فيها الطرائق وحيدة الحافز. ثم يصار إلى ضم الصور الناتجة أو تتابعات الصور من أجل تشكيل الأزواج المستعملة في اختبارات التقدير.

3-A4 جلسة الاختبار

يستعمل اختبار التقدير شاشة عرض واحدة أو شاشتين متوائمتين ويجري عادةً على النحو المتبع في حالات الحافز الوحيد. فإذا استُعملت شاشة عرض واحدة، يدرج الاختبار مجال حافز إضافياً مطابقاً للأول في المدة. ويفضل في هذه الحالة التأكد من أن عنصري الصورتين يتكرران على مر الاختبارات، عدداً متساوياً من المرات في الموقعين الأول والثاني. أما إذا استُعملت شاشتان، فتُقدم مجالات الحافز متآونة.

تقدر طرائق مقارنة الحوافز العلاقات فيما بين الشروط بطريقة أفضل عندما تتعلق المقارنة بكل أزواج الشروط الممكنة. لكن إذا ما تطلب ذلك عدداً كبيراً جداً من المشاهدات، فيمكن أن تقسم التقديمات فيما بين المشاهدين، أو تستعمل عينة من كل الأزواج الممكنة.

4-A4 أنماط طرائق مقارنة الحوافز

لقد استعملت في التقديرات التلفزيونية ثلاثة أنماط من طرائق مقارنة الحوافز.

1.4-A4 طرائق الحكم حسب الفئات بواسطة الصفات

يخصص المراقبون في هذه الطرائق العلاقة بين عنصري الصورتين إلى فئة أو إلى عدد من الفئات التي تعرف نمطياً بواسطة علم الدلالات. ويمكن أن تشير هذه الفئات إلى وجود فروقات يمكن إدراكها (مثل، مماثل، مختلف)، أو إلى وجود فروقات يمكن إدراكها وإلى درجتها (مثل أقل، مماثل، أكثر)، أو تقديرات عن الدرجة والاتجاه. ويعرض الجدول 2-2 أدناه سلم المقارنة الذي وضعه قطاع الاتصالات الراديوية.

الجدول 2-2

سلم المقارنة

–3	أسوأ بكثير
–2	أسوأ
–1	أسوأ بقليل
0	مماثل
+1	أفضل بقليل
+2	أفضل
+3	أفضل بكثير

تؤدي هذه الطريقة إلى توزيع الأحكام على فئات في السلم بالنسبة إلى كل زوج من الشروط. وتتعلق طريقة تحليل الإجابات بالحكم نفسه (الفرق مثلاً)، وبالمعلومات المطلوبة (مثل، فقط اختلافات ملحوظة، ورتب الشروط، و"المسافات" بين الشروط، وما إلى ذلك).

2.4-A4 طرائق الأحكام التي لا تتوزع حسب الفئات

يخصص المراقبون وفقاً لهذه الطرائق قيمةً للعلاقة بين عنصري الصورتين الخاضعتين للتقدير. وثمة شكلان مختلفان لهذه الطريقة:

– يخصص المشاهد، في طريقة السلم المستمر، لكل علاقة نقطة على خط يُرسم بين وسمين (مثل مماثل، مختلف، أو طرفي السلم حسب الفئات كما في الجدول 2-2). ويمكن أن يتضمن السلم وسوماً إضافيةً مرجعيةً عند نقاط متوسطة. تعتبر المسافة من طرف من الخط قيمة لكل زوج من الشروط.

– أما في الشكل الثاني من الطريقة، فيخصص المشاهد لكل علاقة عدداً يعكس سوية حكمه على عامل محدد (مثل الفرق في النوعية). ويمكن أن يكون مدى الأرقام المستعملة مقيداً أو غير مقيد. وقد يصف العدد المخصص العلاقة بعبارات "مطلقة" أو بالعبارات الخاصة بزوج "معياري".

يؤدي الشكلان إلى توزيع القيم على كل زوج من الشروط. وتتعلق طريقة التحليل بطبيعة الحكم وبالمعلومات المطلوبة.

3.4-A4 طرائق الأداء

يمكن في بعض الحالات استخلاص قياسات الأداء انطلاقاً من إجراءات مقارنة الحوافز. ويُحضَّر الزوج، في طريقة الاختيار القسري، على نحو يحتوي فيه العضو الواحد على سوية خاصة لنعت ما (مثل الانحطاط)، بينما يحتوى الآخر على سوية أخرى من هذا النعت أو أنه لا يحتوى على أي نعت. ويُطلب من المراقب أن يقرر أي العنصرين يحتوي على السوية الأعلى أو الأدنى من النعت أو أيهما لا يحتوي على أي نعت. وتعتبر دقة الأداء وسرعته مؤشرات للعلاقة فيما بين عنصري الزوج الواحد من الصور.

الملحق 5
بالجزء 2

طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE)

سيسيء إدخال انضغاط التلفزيون الرقمي لنوعية الصورة التي ترتبط بالمشهد وتتغير مع الوقت. وحتى في مقاطع فيديوية قصيرة مشفرة رقمياً يمكن أن تختلف النوعية اختلافاً كثيراً تبعاً لمحتوى المشهد وقد تكون فترات الخلل قصيرة جداً. والطرائق التقليدية التي وضعها قطاع الاتصالات الراديوية غير كافية بمفردها لتقدير هذا النوع من المواد. ومن ناحية أخرى لا تستطيع الطريقة ثنائية الحافز للاختبار في المختبر أن تعيد تماماً شروط المشاهدة في المنزل بالحافز الوحيد. لذا يستحسن قياس النوعية الشخصاني للإشارات الفيديوية المشفرة رقمياً بصورة مستمرة مع أشخاص يشاهدون الصور مرة واحدة دون مصدر مرجعي.

ونتيجةً لذلك، وُضعت تقنية تقييم النوعية المستمرة وحيدة الحافز (SSCQE) واختبرت.

A5-1 جهاز التسجيل وضبطه

يُستعمل نظام تسجيل إلكتروني موصول بحاسوب من أجل تسجيل التقدير المستمر للنوعية الذي يقوم به المختبرون. وينبغي تزويد هذا الجهاز بالخصائص التالية:

- آلية انزلاق دون أي وضعية استعادة؛
- مدى خطي طوله 10 cm؛
- ثابت أو يمكن تثبيته على مكتب؛
- عينات مسجلة مرتين في الثانية الواحدة.

A5-2 الشكل العام لبروتوكول الاختبار

تعرض المواضيع في جلسات اختبار بالأنساق التالية:

- *مقطع من برنامج (PS)*: ويعادل برنامجاً نمطياً (مثل برنامج رياضي أو نشرة أخبار أو تمثيلية) معالجة وفق إحدى معلمات النوعية (QP) الخاضعة للتقييم (مثل معدل البتات)؛ وينبغي ألا تقل مدة كل مقطع عن 5 دقائق؛
- *جلسة الاختبار (TS)*: وهي سلسلة من تتابع واحد أو تتابعات مختلفة من المقاطع PS/QP لا فاصل بينها وفق ترتيب شبه عشوائي. وتضم كل جلسة جميع المقاطع PS وQP مرة واحدة على الأقل ولكن ليس بالضرورة جميع تشكيلات المقاطع PS/QP؛ وينبغي أن تتراوح مدة كل جلسة بين 30 و60 دقيقة؛
- *عرض الاختبار (TP)*: ويمثل كامل أداء الاختبار. ويمكن تقسيم عرض الاختبار إلى جلسات اختبار من أجل مراعاة أقصى متطلبات المدة وتقدير النوعية في جميع الأزواج PS/QP. وإذا كان عدد الأزواج PS/QP محدوداً، يمكن أن يتألف عرض الاختبار TP من تكرار نفس المقطع TS لإجراء اختبار في فترة زمنية كافية.

وفيما يتعلق بتقييم نوعية الخدمة، يمكن إدخال إشارات سمعية. وفي هذه الحالة، ينبغي اعتبار الإشارات المرافقة بنفس درجة أهمية الإشارات الفيديوية المنتقاة قبل إجراء الاختبار.

ويستعمل أبسط نسق اختبار مقطع برنامج (PS) واحد ومعلمة نوعية (QP) واحدة.

A5-3 معلمات المشاهدة

ينبغي أن تكون شروط المشاهدة هي تلك المحددة في الجزء 1 أو الشروط الخاصة بالتطبيق الواردة في الجزء 3.

4-A5 سلالم إسناد الدرجات

ينبغي إعلام المشاهدين في تعليمات الاختبار بأن طول طريق آلية الانزلاق تقابل سلم تقييم النوعية المستمرة الوارد في القسم 1-A4 من الجزء 1.

5-A5 المراقبون

ينبغي تكليف خمسة عشر مشاهداً غير خبير على الأقل واستعمال الشروط الموصى بها عادةً والواردة في القسم 5.2 من الجزء 1.

6-A5 تعليمات إلى المراقبين

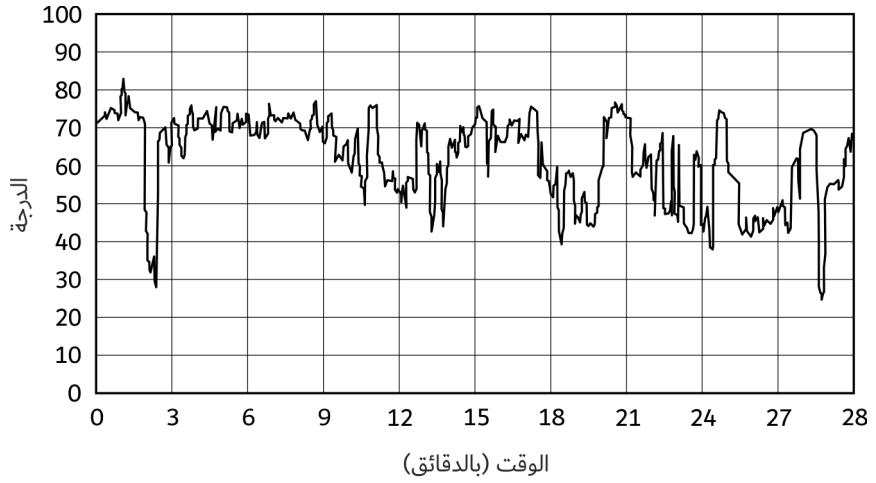
في حال تقييم نوعية الخدمة (مرفقة بإشارات سمعية)، ينبغي إعلام المراقبين بضرورة اعتبار النوعية الإجمالية وليس النوعية الفيديوية بمفردها.

7-A5 عرض البيانات ومعالجة النتائج وتقديمها

تُجمع البيانات من جميع جلسات الاختبار. لذا ينتج رسم بياني واحد لمتوسط تقييم النوعية بدلالة الوقت $q(t)$، وهو متوسط درجات تقدير النوعية التي أعطاها مجمل المراقبين لكل مقطع برنامج أو معلمة نوعية أو لكل جلسة اختبار كاملة (انظر المثال في الشكل 2-6).

الشكل 2-6

شروط الاختبار الكودك X/مقطع البرنامج: Z



BT.0500-02-6

غير أن فروق زمن تجاوب المشاهدين قد تؤثر على نتائج التقدير إذا لم يحسب المتوسط إلا استناداً إلى مقطع برنامج. وتُجرى حالياً دراسات لتقييم تأثير فرق زمن تجاوب المشاهدين على درجة النوعية الناتجة.

ويمكن تحويل هذه البيانات إلى مخطط تسجيل احتمالات، $P(q)$، حدوث مستوى النوعية q (انظر المثال في الشكل 2-7).

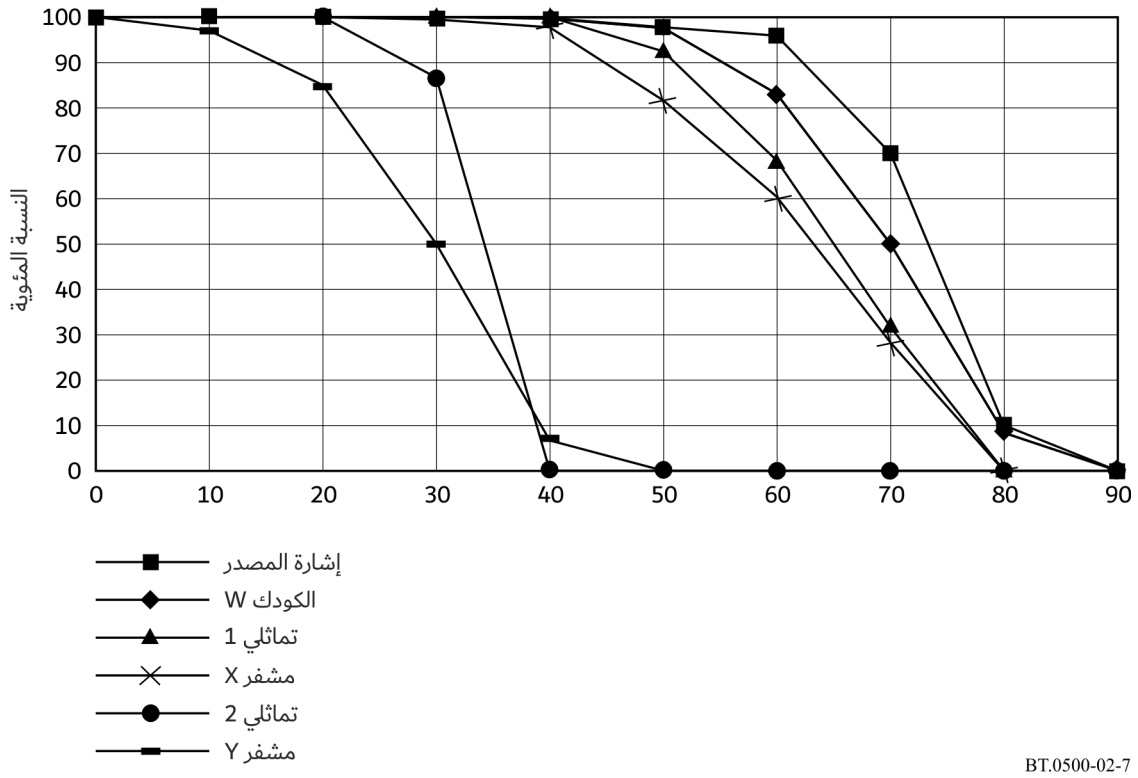
8-A5 معايرة نتائج النوعية المستمرة واستنتاج تقييم عام للنوعية

على الرغم من ملاحظة أن الأخطاء التي تعزى إلى الذاكرة تظهر خلال جلسات طويلة بطريقة DSCQS لتقييم نوعية التسجيلات المشفرة رقمياً، تم التحقق مؤخراً من أن هذه الأخطاء ليست جسيمة عند التقييم بطريقة DSCQS لمقاطع فيديوية مدتها 10 ثوان. لذا هنالك مرحلة ثانية ممكنة لطريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE) قيد الدراسة حالياً، وتتمكن من معايرة مخطط النوعية المسجلة بواسطة طريقة DSCQS القائمة باستعمال عينات تمثل مقاطع من بيانات المخطط مدة العينة منها 10 ثوان.

وأتاحت المنهجيات التقليدية التي استعملها قطاع الاتصالات الراديوية سابقاً تقييمات لنوعية تتابعات إشارات تلفزيونية. وقد أجريت تجارب لدراسة العلاقة القائمة بين التقييم المستمر لنوعية تتابع فيديوي مشفر والتقييم الشامل لنوعية المقطع ذاته. وقد تبين سابقاً أن ذاكرة الإنسان قد تخدعه وتشوِّه درجات تقييم النوعية إذا ظهرت الانحطاطات التي يمكن ملاحظتها في الثواني الأخيرة العشر حتى الخمس عشر من التتابع، لكن تبين أيضاً إمكانية إعادة نموذج الآثار الخادعة للذاكرة على شكل دالة أسية تنازلية. وهناك احتمال مرحلة ثالثة لمنهجية SSCQE تكمن في معالجة نتائج هذه التقييمات المستمرة للنوعية من أجل استنتاج قياس مكافئ للنوعية. وهذا الأمر قيد الدراسة.

الشكل 2-7

متوسط علامات التتابعات المطروحة للتقييم بشأن مقطع من البرنامج Z



الملحق 6
بالجزء 2

طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE)

وضع قطاع الاتصالات الراديوية طريقة التقييم المستمر لأن الطرائق السابقة لم تكن ملائمةً تماماً لقياس النوعية الفيديوية لأنظمة الانضغاط الرقمي. وأخطر المساوئ التي كانت تتسم بها هذه الطرائق المعيارية السابقة هي كثرة ظهور انحطاطات متصلة بسياق الصور الرقمية المعروضة. وفي البروتوكولات السابقة كانت مدة مشاهدة التتابعات الفيديوية المطلوب تقييمها تقتصر عادةً على 10 ثوان، وهي بالطبع غير كافية لأن يكوِّن المراقب حكماً يمثل حقيقة ما يحصل في الخدمة الفعلية. وترتبط الانحطاطات الرقمية ارتباطاً شديداً بالمحتوى المكاني والزمني لصورة المصدر. وينطبق ذلك أيضاً على أنظمة الانضغاط لكنه يتعلق أيضاً بسلوك مقاومة

الأخطاء في أنظمة الإرسال الرقمية. وقد كان من الصعوبة بمكان استناداً إلى المنهجيات المعيارية السابقة اختبار تتابعات فيديوية تمثل الواقع تمثيلاً صحيحاً أو على الأقل يمكن تقييم مدى تمثيلها للواقع. ولهذا السبب وضع قطاع الاتصالات الراديوية طريقة SSCQE التي تتيح قياس النوعية الفيديوية في تتابعات أطول تمثل المحتويات الفيديوية وإحصاءات الخطأ. ومن أجل إنتاج شروط مشاهدة قريبة إلى أبعد حدٍ ممكن من قبل الشروط في الواقع لا تُستعمل مرجعيات في طريقة SSCQE.

أما عند تقييم الأمانة، فيجب إدراج شروط مرجعية. وقد وضعت طريقة SDSCE استناداً إلى طريقة SSCQE مع إدخال بعض التعديلات فيما يتعلق بطريقة عرض الصور على المشاهدين وبسلم التقييم. وجرى اقتراح هذه الطريقة على الفريق MPEG لتقييم مقاومتها للأخطاء في معدل ثبات شديد الانخفاض، لكن يمكن استعمالها جيداً في جميع الحالات التي يتوجب فيها تقييم أمانة المعلومات المرئية المتأثرة بانحطاط متغير مع الزمن.

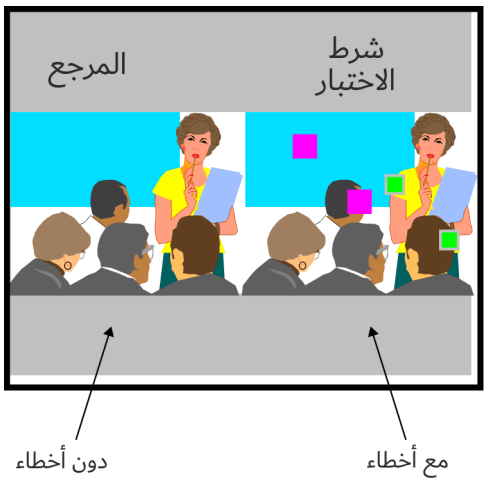
ونتيجةً لذلك، وُضِعت التقنية الجديدة SDSCE التالية وجُرِّبت.

1-A6 إجراء الاختبار

تشاهد مجموعة الأشخاص تتابعين اثنين في نفس الوقت: أحدهما مرجعي والآخر للاختبار. إذا كان نسق التتابعين هو نسق الصورة المعياري (SIF) أو أصغر، ويمكن عرض التتابعين جنباً إلى جنب في نفس شاشة العرض وإلاَّ فتستعمل شاشتا عرض متوافقتان (انظر الشكل 2-8).

الشكل 2-8

مثال لنسق العرض



BT.0500-02-8

ويُطلب إلى المشاهدين أن يرصدوا الاختلافات بين التتابعين ويقيموا مدى أمانة المعلومات الفيديوية وذلك بتحريك مزلاق جهاز التقييم. فعندما تكون الأمانة كاملة يوضع المزلاق على أعلى درجة (100) وعندما تنعدم الأمانة يوضع المزلاق على أدنى درجة (0).

ويُعلم المشاهدون بالصورة المرجعية ويطلب إليهم إبداء رأيهم لدى مشاهدة التتابعات وإلى مدى كامل الجلسة.

2-A6 المراحل المختلفة

مرحلة التدريب مرحلة حاسمة في طريقة الاختبار هذه كي يدرك المشاهدون تماماً ما يتوجب عمله. وينبغي توفير تعليمات كتابية للتأكد من أن جميع المشاهدين تلقوا نفس المعلومات تماماً. وينبغي أن تضم التعليمات شرحاً لما سيراه المشاهدون ولما يتوجب عليهم تقييمه (أي التباين في النوعية) وكيفية التعبير عن آرائهم. وينبغي الإجابة على كل سؤال يطرحه المشاهدون لتفادي قدر الإمكان أي رأي متحيز من المسؤول عن الاختبار.

وبعد توزيع التعليمات تُجرى *جلسة توضيحية* كي يألف المشاهدون إجراءات التقييم وأنواع الانخطاطات على حد سواء.

وأخيراً يستحسن إجراء محاكاة اختبار تعرض فيه بعض الشروط النمطية. وينبغي أن تكون التتابعات غير تلك المستعملة في الاختبار وينبغي عرضها واحد تلو الآخر دون أي انقطاع.

وعند انتهاء *اختبار المحاكاة*، ينبغي أن يتحقق المختبر أساساً من أن التقييم في الحالات المختبرة المماثلة للصور المرجعية قريب من درجة مئة (أي لا فرق ملحوظ): وإذا أعلن المشاهدون بدلاً من ذلك أنهم لاحظوا فروقاً، ينبغي أن يعيد المختبر كلاً من شرح الاختبار ومحاكاته.

3-A6 خصائص بروتوكول الاختبار

تطبق التعاريف التالية على وصف بروتوكول الاختبار:

- *مقطع فيديوي (VS)*: وهو تتابع فيديوي واحد.
- *شرط اختبار (TC)*: هو إما عملية فيديوية محددة وإما شرط إرسال أو الاثنان معاً. وينبغي أن يُعالَج كل مقطع VS وفق شرط TC واحد على الأقل. كما ينبغي إضافة الصور المرجعية إلى قائمة شروط الاختبار من أجل تكوين الصور المرجعية/أزواج من الصور المرجعية للتقييم.
- *جلسة (S)*: هي مجموعة من الأزواج المؤلفة من مقطع فيديوي/شرط اختبار دون فواصل ومرتبة حسب ترتيب شبه عشوائي. وتشمل كل جلسة جميع المقاطع VS والشروط TC مرة واحدة على الأقل دون تجميعاتها بالضرورة.
- *عرض الاختبار (TP)*: هو سلسلة من الجلسات التي تشمل جميع تجميعات المقاطع VS/الشروط TC. ويجب أن يتم تقييم جميع التجميعات VS/TC من نفس عدد المراقبين (وليس بالضرورة نفس المراقبين).
- *مرحلة التقييم*: يطلب إلى كل مراقب أن يدلي بتقييمه بصورة متواصلة أثناء الجلسة.
- *مقاطع التقييم (SOV)*: وهو مقطع طوله عشر ثوان من التقييم؛ ويتم الحصول على المقاطع SOV باستعمال مجموعات من 20 تقييماً متعاقباً (يعادل 10 ثوان) دون أي تراكب.

4-A6 معالجة البيانات

بعد إجراء الاختبار، يتاح ملف بيانات واحد (أو أكثر) يضم جميع أصوات التقييم لمختلف الجلسات التي تمثل إجمالي عدد درجات التقييم لعرض الاختبار (TP). ويمكن إجراء أول تحقق من صلاحية البيانات من خلال التحقق من أن كل زوج VS/TC عولج وأن عدداً متكافئاً من درجات التقييم قد أعطيت لكل من هذه الأزواج.

ويمكن معالجة البيانات المجمعة من الاختبارات التي أُجريت وفق هذا البروتوكول بطرائق ثلاث هي:

- تحليل إحصائي لكل مقطع VS على حدة؛
- تحليل إحصائي لكل مقطع TC على حدة؛
- تحليل إحصائي شامل لجميع الأزواج VS/TC.

ويُشترط إجراء تحليل متعدد المراحل لكل حالة:

- تحسب القيم المتوسطة والانحرافات المعيارية لكل تقييم يجمع أصوات المراقبين.
- تحسب القيم المتوسطة والانحرافات المعيارية لكل مقطع SOV وفقاً لما يرد في الشكل 2-9 ويمكن عرض نتائج هذه المرحلة في مخطط زمني كما هو مبين في الشكل 2-10.

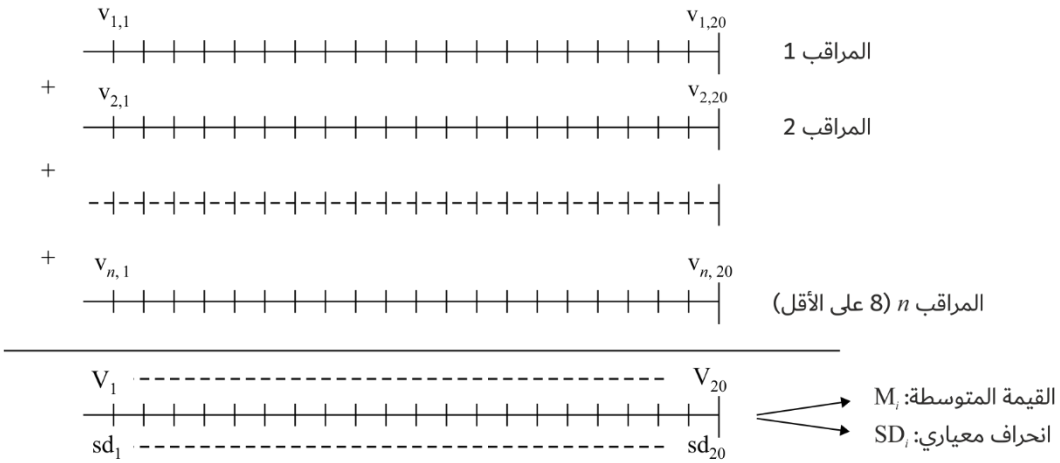
– يحلل التوزيع الإحصائي للقيم المتوسطة الناتجة عن المرحلة السابقة (أي معادل كل مقطع SOV) ووتيرة ظهورها. وسعياً لتفادي الآثار الجديدة الناجمة عن عمليات تجميع VS × TC السابقة، تستبعد أول 10 مقاطع SOV من كل عينة VS × TC.

– تحسب الخصائص العامة للمضايقات من خلال جمع مرات ظهورها. وينبغي في هذا الحساب مراعاة فترات الثقة على النحو الوارد في الشكل 2-11. وتعادل الخصائص العامة للمضايقات دالة التوزيع الإحصائي التراكمي ببيان العلاقة بين القيم المتوسطة لكل مقطع SOV ومجموع عدد مرات ظهورها.

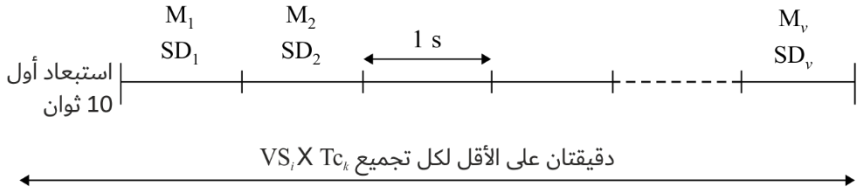
الشكل 2-9

معالجة البيانات

أ) حساب متوسط العلامات، V، والانحراف المعياري، SD، لكل حالة تقييم للمراقبين لكل مقطع تقييم لكل تجميع VS X TC



ب) حساب متوسط M والانحراف المعياري SD لكل مقطع تقييم للمراقبين لكل مقطع تقييم طوله 1 ثانية لكل تجميع VS X TC



BT.0500-02-9

5-A6 موثوقية الأشخاص

يمكن تقييم موثوقية الأشخاص نوعياً من خلال التحقق من ردود فعلهم لدى مشاهدتهم لأزواج الصور المرجعية. وفي هذه الحالات يفترض أن يعطي الأشخاص تقييماً قريباً جداً من الدرجة 100. مما يثبت على الأقل أنهم فهموا المهمة المسندة إليهم وأنهم لا يجرون تقييمات عشوائية.

كما يمكن التأكد من موثوقية الأشخاص من خلال استعمال إجراءات قريبة من تلك الواردة في القسم 2.3.2-A1 من الملحق 1 بالجزء 1 المتعلقة بطريقة SSCQE.

وفي إجراء طريقة SDSCE تتوقف موثوقية التقييم على المعلمتين التاليتين:

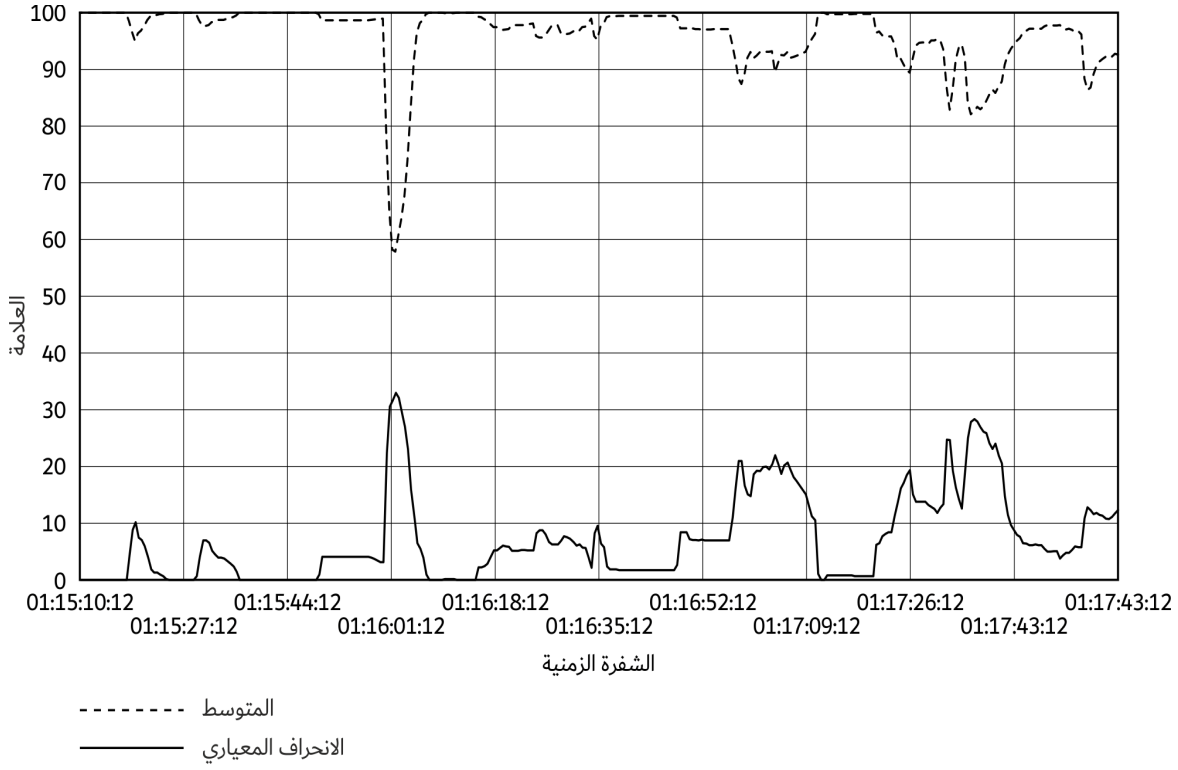
التحولات المنهجية: قد يكون المشاهد أثناء الاختبار مفرطاً بالتفاؤل أو بالتشاؤم أو قد يكون غير فاهم لإجراءات التقييم (مثال معنى سلم التقدير). ويمكن أن يؤدي ذلك إلى سلسلة من التقديرات المتخالفة نوعاً ما وبشكل منتظم مع القيم المتوسطة إن لم تتناقض تماماً معها.

ارتكاسات محلية: قد يطلق المراقبون على النحو المعروف في إجراءات اختبار أخرى تقييمهم دون إيلاء انتباه كبير للمشاهدة وتتبع نوعية تتابع الصور المعروضة. وفي هذه الحالة يكون منحني التقييم العام داخل المدى المتوسط نسبياً. ومع ذلك، يمكن ملاحظة بعض الارتكاسات المحلية.

وبالإمكان تفادي هذين الأثرين غير المرغوب بهما (السلوك غير العادي والارتكاسات). فتدريب المشاركين بالطبع أمر بالغ الأهمية. لكن ينبغي أيضاً إتاحة استعمال أداة تتيح كشف ملاحظات المشاهدين غير المتسقة واستبعادها حسب الاقتضاء. وتقدم هذه التوصية اقتراح عملية بمرحلتين تتيح تنقية من هذا القبيل.

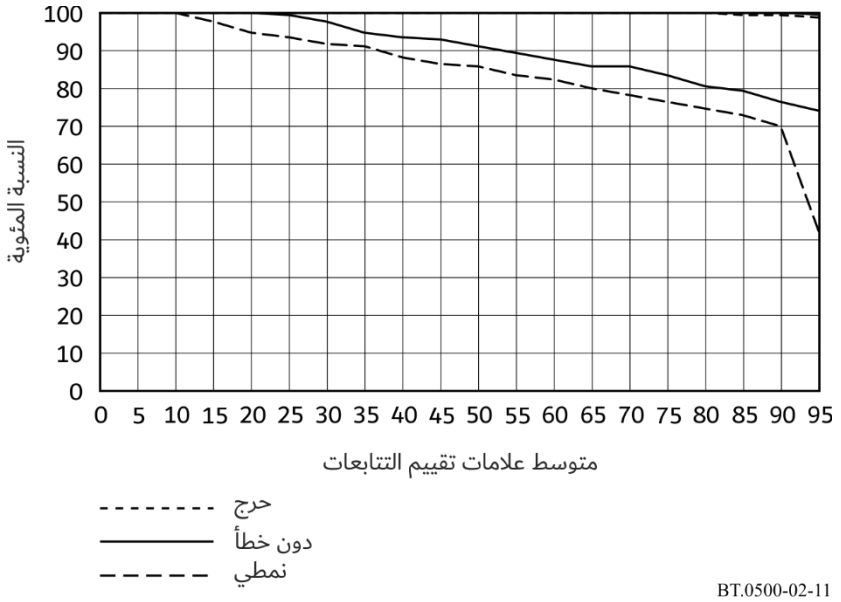
الشكل 2-10

مخطط إجمالي زمني



الشكل 2-11

تحديد خصائص المضايقة العامة استناداً إلى التوزيعات الإحصائية بما فيها فترات الثقة



الملحق 7
بالجزء 2

طريقة التقدير الشخصاني لنوعية الفيديو المتعدد الوسائط (SAMVIQ)

1-A7 مقدمة

تَستعمل طريقة تقدير النوعية SAMVIQ سلم جودة مستمرة لتقدم قياساً للنوعية الأصيلة لتتابعات الفيديو. ويحرّك كل مراقب مؤشر منزلق على سلم مستمر بتدريج من 0 إلى 100 ومزوّد بحواشٍ من 5 بنود للنوعية مرتبة خطياً (ممتاز، جيد، مقبول، ضعيف، سيء).

في طريقة SAMVIQ، يتاح للمشاهد النفاذ إلى العديد من نسخ التتابع. وبعدما يصنّف المشاهد كل النسخ، يمكن النفاذ لمحتوى التتابع التالي.

ويمكن للمشاهد اختيار النسخ المختلفة عشوائياً عبر سطح بيني لرسوم الحاسوب. ويمكن للمشاهد أن يوقف ويراجع ويعدّل علامة كل نسخة من نسخ التتابع حسب الرغبة. وتتضمن هذه الطريقة تتابعاً مرجعياً صريحاً (أي غير معالج) علاوةً على العديد من نسخ التتابع نفسه تضم تتابعات معالجة وغير معالجة (أي مرجعية مخفية) على حد سواء. وتُعرض كل نسخة بشكل منفرد وتُصنَّف باستعمال سلم جودة مستمرة مشابه لذاك المستعمل في طريقة DSCQS. ومن ثم فإن الطريقة تعتبر مماثلة وظائفياً إلى حد كبير للطريقة وحيدة الحافز ذات النفاذ العشوائي، بيد أن المراقب بوسعه مشاهدة المرجع الصريح كلما رغب بذلك، مما يجعل من هذه الطريقة شبيهةً بتلك التي تستعمل مرجعاً.

وتستعمل طريقة تقدير النوعية SAMVIQ سلم نوعية مستمرة لتقدم قياساً للنوعية الأصيلة لتتابعات الفيديو. ويحرّك كل مراقب مؤشر منزلق على سلم مستمر بتدريج من 0 إلى 100 ومزوّد بحواشٍ من 5 بنود للنوعية مرتبة خطياً (ممتاز، جيد، مقبول، ضعيف، سيء).

ويُنفَّذ تقدير النوعية *لكل مشهد على حدة* (انظر الشكل 2-12) على أن يشتمل على مرجع صريح ومرجع مخفي وخوارزميات متنوعة. ووصولاً لفهم أفضل للطريقة، تُعرّف الكلمات المحددة التالية أدناه:

مشهد: محتوى سمعي مرئي

تتابع: مشهد مع معالجة مركبة أو بدون معالجة

خوارزمية: تقنية أو عدة تقنيات لمعالجة الصورة.

2-A7 المرجع الصريح والمخفي والخوارزميات

تتضمن طريقة التقدير عموماً وسائل تثبيت للنوعية من أجل إقرار النتائج. ويتم مراعاة وسيلتين رفيعتين للنوعية في طريقة SAMVIQ للأسباب التالية. أظهرت اختبارات عديدة أُجريت أن الانحرافات المعيارية في الدرجات باستعمال مرجع صريح أقل منها عند استعمال مرجع مخفي أو عدم استعمال مرجع. ويُستحسن استعمال مرجع صريح للحصول على الاعتمادية القصوى للنتائج، لا سيما في تقدير أداء الكودك. ويضاف أيضاً مرجع مخفي بدلاً من المرجع الصريح لتقدير النوعية الأصيلة للمرجع، وذلك لأن العرض غير محدد الهوية وكذلك التتابعات المعالجة. و"للمرجع" ذي الاسم الصريح تأثيرٌ على حوالي 30% من المراقبين الذين يعطون أعلى علامة ممكنة (100) لمرجع صريح على نحوٍ مغاير تماماً للعلامة المقابلة للمرجع المخفي. وعلى نحوٍ لافتٍ، يظل الاختبار ممكناً في غياب مرجع، غير أن الانحراف المعياري يزداد بشكل كبير.

وتعد طريقة SAMVIQ مناسبةً للمحتوى متعدد الوسائط حيث يمكن دمج خصائص مختلفة لمعالجة الصورة من قبيل نمط الكودك ونسق الصورة ومعدل البتات والتحديث الزمني والتكبير والتصغير وما إلى ذلك. وتُختَزل إحدى هذه الخصائص أو تركيبة من عدد منها تحت اسم الخوارزمية.

3-A7 شروط الاختبار

إن اختيار المحتويات المتجانسة بالاتّباع الضمني لنفس القواعد التي تستعملها منهجيات أخرى توفر علامة تقييم شاملة (مثل طرائق وحيدة الحافز)، يؤدي إلى الحد من التغيير في الإلحاح أثناء المشهد. ومن ثم تكفي فترة مشاهدة تتابع قصوى قدرها 10 أو 15 ثانية للحصول على علامة للنوعية مستقرة وموثوقة. وينبغي استعمال أجهزة فك شفرة–تشغيل مسجَّلة أو نسخة تستقبل على شاشة أخرى من خرجها للحفاظ على أداء العرض المناسب.

4-A7 تنظيم الاختبار

أ) يُنفَّذ الاختبار لكل مشهد على حدة حسب الوصف في الشكل 2-12.

ب) بالنسبة للمشهد الراهن، يمكن تشغيل أي تتابع وتقييمه بعلامة بأي ترتيب. كما يمكن تشغيل كل تتابع وتقييمه بعلامة لمرات عديدة.

ج) يُرتَّب النفاذ عشوائياً من مشهد لآخر للحيلولة دون محاولة المراقبين من التصويت بطريقة متطابقة وفق ترتيب ثابت. وفي الواقع، يُحتفظ بنفس ترتيب الخوارزمية داخل الاختبار لتبسيط تحليل النتائج وعرضها. ولا يُرتَّب عشوائياً إلا النفاذ المقابل من زر مطابق.

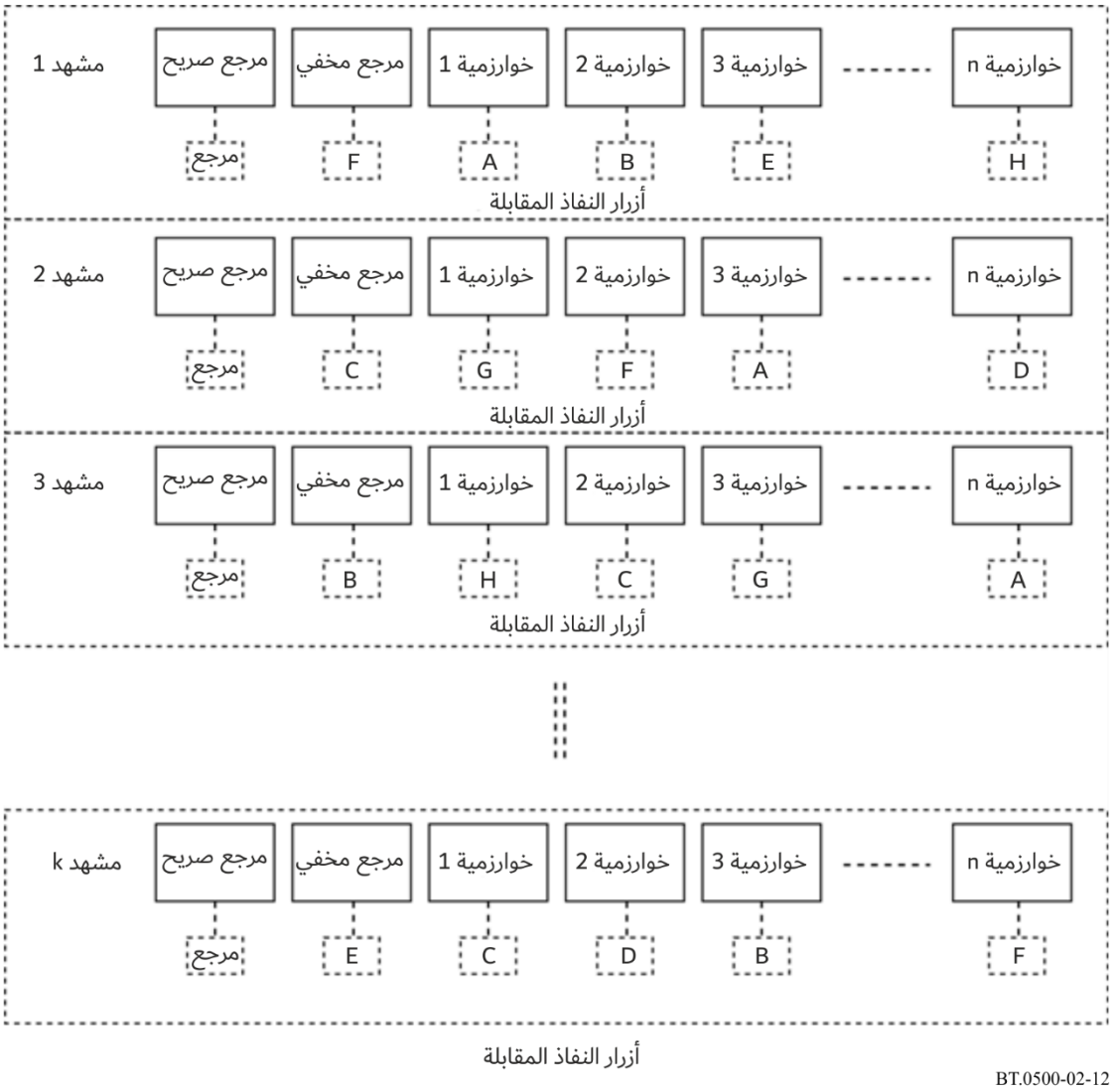
د) بالنسبة للمشاهدة الأولى، ينبغي أن يُستعرَض التتابع الراهن بالكامل قبل تقييمه بعلامة، أو يمكن تسجيل العلامة والتوقف فوراً.

هـ) لاختبار المشهد التالي، لا بد من تقييم جميع التتابعات في المشهد الراهن بإعطائها العلامة.

و) للانتهاء من الاختبار، لا بد من إعطاء علامة لجميع التتابعات في جميع المشاهد.

الشكل 2-12

مثال تنظيم اختبار لطريقة SAMVIQ



تُنفَّذ طريقة SAMVIQ عبر البرمجيات. وفضلاً عن أزرار النفاذ المبيّنة في الشكل 2-12، تلزم أزرار "تشغيل" و"إيقاف" و"المشهد التالي" و"المشهد السابق" كي يتمكن المشاهد من إدارة عرض المشاهد المختلفة (انظر الفقرة A7-6 مثلاً). وعندما يعطي المشاهد علامةً معينةً ينبغي إظهارها تحت زر النفاذ المقابل لذلك المشهد. وعند إعطاء علامات لجميع النسخ المختلفة لتتابع ما، تظل متاحةً للمشاهد المقارنة بين العلامات وتعديلها إذا لزم الأمر. وليس ضرورياً مراجعة التتابع الراهن برمته لأن فروقاً كبيرة يكون قد تم تسليط الضوء عليها في المشاهدة الأولى.

A7-5 عرض وتحليل البيانات

A7-1.5 معلومات موجزة

من الضروري وجود معلومات دقيقة عن بيئة الاختبار لاستنساخ الاختبار أو لمقارنة النتائج عبر الاختبارات المختلفة. لذا، يُقترح إبلاغ المعلومات المتعلقة ببيئة الاختبار على النحو الوارد في الجدول 2-3.

الجدول 2-3

معلومات موجزة عن الاختبار

اسم الطريقة	
تكنولوجيا العرض	
الاسم المرجعي للعرض	
مستوى ذروة النصوع (cd/m²)	
مستوى النصوع الأسود (cd/m²)	
ضبط مستوى اللون الأسود: PLUGE (العتبة المحسوسة للمسافة بين سوية الأسود وفوق الأسود = 8). وفيما عدا ذلك يبيّن قيمة العتبة	
مستوى نصوع الخلفية (cd/m²)	
الإضاءة (lux)	
مسافة المشاهدة: — غير مقيّدة: أمام شاشة العرض — مقيّدة: nH	
أبعاد شاشة العرض (قطري بالبوصة)	
النسبة عرض/طول شاشة العرض	
نسق العرض (عدد الأعمدة والخطوط)	
نسق دخل الصورة (عدد الأعمدة والخطوط)	
نسق خرج الصورة (عدد الأعمدة والخطوط)[1]	
حرارة اللون الأبيض: D_{65} وفيما عدا ذلك تعطي إحداثيات اللون الأبيض (x، y)	
عدد المراقبين المؤثرين	

[1] تلزم هذه المعلومة عند معالجة صورة الدخل، عند إعادة مقايستها، عند العرض، مثلاً.

وقد يكون لخصائص العرض أثر على نتائج الاختبار. وينبغي وجود معلومات إضافية لشاشات العرض المسطحة مثل استجابة النصوع (أمانة غاما) وأساسيات اللون.

وتعد خصائص تتابعات الفيديو مهمةً لتصميم اختبار أو تفسير نتائجه. ويُقترح الإبلاغ عن الخصائص المكانية–الزمانية حسب وصفها في الملحق 1 بالجزء 1. وينبغي النظر في هذه المعلومات ضمن مجموعة تتابعات الاختبار في مكتبة مواد الفيديو المناسبة للتقدير الشخصاني لنوعية الفيديو في التطبيقات متعددة الوسائط.

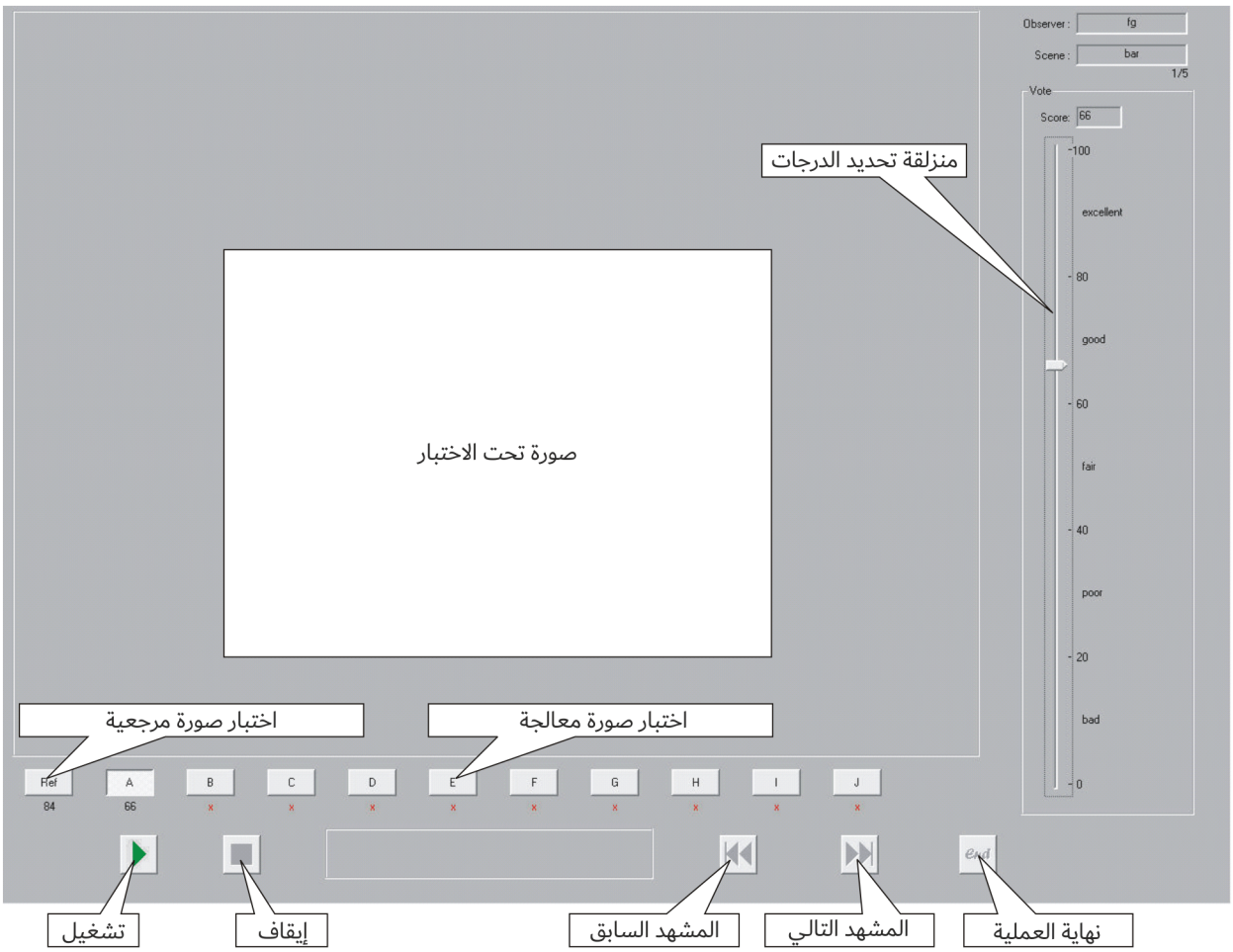
2.5-A7 طرائق التحليل

طرائق التحليل هي تلك الموصوفة في الملحق 1 من الجزء 1.

3.5-A7 فرز المراقبين

يرد وصف الفرز المتبع في طريقة SAMVIQ في الفقرة 3.3.2-A1 من الملحق 1 بالجزء 1.

6-A7 مثال عن سطح بيني لطريقة تقدير SAMVIQ (إعلامي)



BT.0500-02-12a

الملحق 8
بالجزء 2

بروتوكول مشاهدة الخبراء (EVP) لتقييم نوعية المادة الفيديوية

يصف هذا الملحق طريقة التقييم الشخصاني لنوعية الفيديو للصور المتحركة باستعمال بروتوكول مشاهدة الخبراء، وبمشاركة عدد منخفض من المشاهدين يتم اختيارهم جميعاً من بين الخبراء العاملين في مجال معالجة الفيديو ذي الصلة.

1-A8 إعدادات المختبر

1.1-A8 اختيار شاشة العرض وإعدادها

ينبغي أن تكون شاشة العرض المستعملة مسطحةً وتتميز بأداء نمطي للتطبيقات المهنية (مثل استوديوهات أو عربات الإذاعة)؛ ويمكن أن يتراوح البعد القُطري للشاشة بين 22' (كحد أدنى) و40' (مقترح)، ولكنه قد يمتد إلى 50' أو أكثر، عندما تقيّم أنظمة صور ذات استبانة عالية الوضوح أو أعلى.

ويُسمح باستعمال جزء مختصَر من منطقة المشاهدة النشطة للشاشة؛ وفي هذه الحالة ينبغي ضبط المنطقة المحيطة بالجزء النشط من الشاشة بمتوسط الرمادي. وفي ظرف الاستعمال هذا، ينبغي عدم السماح بضبط الشاشة باستبانة مختلفة عن استبانتها الأصلية.

وينبغي أن تتيح الشاشة إعدادات مناسبة ومعايرة للنصوع واللون، باستعمال أداة مهنية لقياس الضوء. وينبغي أن تتوافق معايرة شاشة العرض مع المعلمات المحددة في التوصية ذات الصلة للاختبار الجاري.

2.1-A8 مسافة المشاهدة

ينبغي اختيار مسافة المشاهدة التي يجلس عندها الخبراء وفقاً لاستبانة الشاشة، وارتفاع الجزء النشط من الشاشة، ووفقاً لمسافة المشاهدة بالتصميم الموصوفة في الفقرة 2.3.1.2 من الجزء 1 أو مسافة مشاهدة أقصر، حسب المتطلبات من حيث شروط المشاهدة الحرجة.

3.1-A8 شروط المشاهدة

لا ينبغي بالضرورة تشغيل تجربة بروتوكول مشاهدة الخبراء (EVP) في مختبر اختبار، ولكن من المهم أن يكون موقع الاختبار محمياً من الاضطرابات المسموعة أو المرئية (على سبيل المثال: يمكن أيضاً استعمال مكتب هادئ أو قاعة اجتماعات هادئة).

وتنبغي إزالة أي مصدر لضوء مباشر أو منعكس يقع على الشاشة؛ وينبغي أن يكون الضوء الآخر المحيط باهتاً، وأن يبقى بالحد الأدنى الذي يمكِّن من ملء أوراق إسناد الدرجات (إذا ما استعملت).

ويمكن أن يختلف عدد الخبراء الذين يجلسون أمام شاشة العرض، وفقاً لمقاس الشاشة، من أجل ضمان القدر نفسه من إظهار الصورة وعرض المؤثرات لجميع المشاهدين.

2-A8 المشاهدون

ينبغي أن يكون المشاهدون المشاركون في تجربة بروتوكول مشاهدة الخبراء (EVP) خبراء في مجال الدراسة.

ولا ينبغي بالضرورة فحص المشاهدين بشأن حدة البصر أو عمى الألوان، باعتبار أنه ينبغي اختيارهم من بين الأشخاص المؤهلين.

وينبغي أن يكون الحد الأدنى لعدد المشاهدين المختلفين تسعة.

وللوصول إلى الحد الأدنى لعدد المشاهدين، يمكن إجراء التجربة نفسها في الموقع نفسه وتكرار الاختبار، أو في أكثر من موقع واحد. وتمكن المعالجة الإحصائية سَوِيّةً للعلامات المستقاة من المواقع المختلفة المشاركة في جلسة مشاهدة الخبراء.

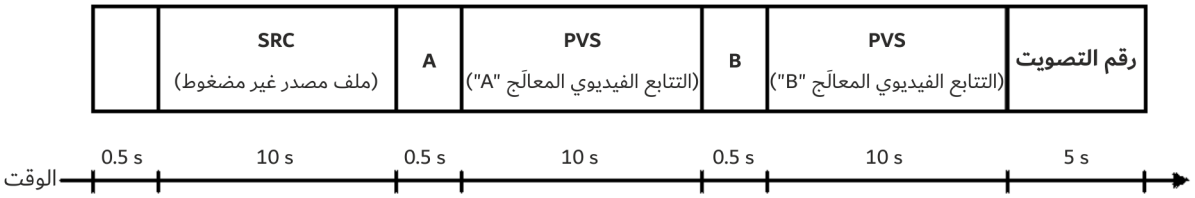
3-A8 خلية الاختبار الأساسية

ينبغي تنظيم المادة التي ستُعرض على الخبراء بإنشاء خلية اختبار أساسية (BTC) لكل زوجين من شروط التشفير التي يتعين تقييمها (انظر الشكل 2-13).

وينبغي دائماً لمقاطع تتابعات المصدر المرجعي (SRC) والتتابعات الفيديوية المعالَجة (PVS) التي ينبغي أن تنظر فيها خلية الاختبار الأساسية (BTC) أن تتصل بنفس التتابع الفيديوي لكي يتمكن الخبراء من تحديد أي تحسن في الجودة المرئية التي تقدمها خوارزميات الضغط قيد الاختبار.

الشكل 2-13

مواقيت خلية الاختبار الأساسية لبروتوكول مشاهدة الخبراء



BT.0500-02-13

وينبغي تنظيم خلية الاختبار الأساسية (BTC) على النحو التالي:

- 0,5 ثانية مع ضبط الشاشة بمتوسط الرمادي (متوسط القيمة في مقياس النصوع)؛
- عرض مقطع الفيديو المرجعي غير المضغوط لمدة 10 ثوان؛
- 0,5 ثانية بإظهار الرسالة "A" (الفيديو الأول الذي يتعين تقييمه) على خلفية بمتوسط الرمادي؛
- عرض نسخة رديئة من مقطع الفيديو لمدة 10 ثوان؛
- 0,5 ثانية بإظهار الرسالة "B" (الفيديو الثاني الذي يتعين تقييمه) على خلفية بمتوسط الرمادي؛
- عرض نسخة رديئة من مقطع الفيديو لمدة 10 ثوان؛
- 5 ثوان تظهر رسالة تطلب من المشاهدين التعبير عن رأيهم.

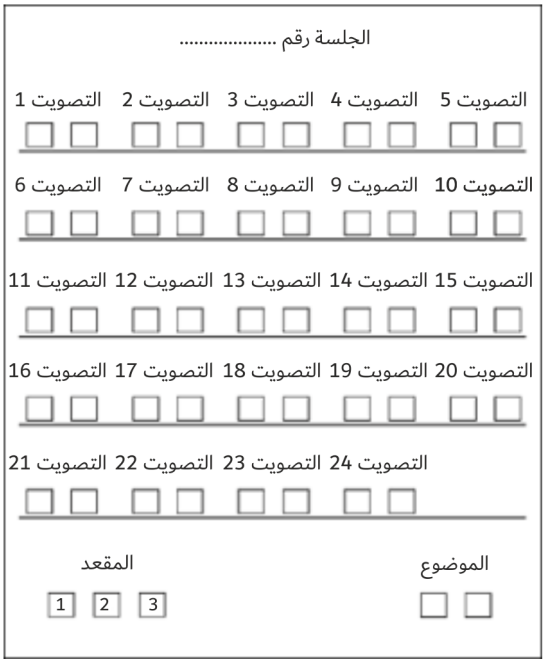
وينبغي أن تُتبع رسالة "صوِّتوا" برقم يساعد على التزامن على ورقة إسناد الدرجات.

4-A8 ورقة إسناد الدرجات وسلُّم التصنيف

على النحو المبين في الشكل 2-13، ينبغي ترتيب عرض مقاطع الفيديو بطريقة تظهر فيها تتابعات المصدر المرجعي (SRC) غير المتردية في البداية، ويليها اثنان من التتابعات الفيديوية المعالَجة (PVS) المتردية. وينبغي تغيير ترتيب عرض التتابعات الفيديوية المعالَجة بشكل عشوائي في كل خلية اختبار أساسية (BTC) وينبغي ألا يعرف المشاهدون ترتيب العرض.

الشكل 2-14

مثال على ورقة إسناد الدرجات في جلسة مشاهدة تضم 24 خلية اختبار أساسية



BT.0500-02-14

يُستعمل سلُّم رقمي ذو 11 درجة من 10 (انحطاطات لا يمكن إدراكها) إلى 0 (انحطاطات مزعجة للغاية).

ويقدم الجدول 2-4 توجيهات بشأن معنى السلُّم العددي ذي 11 درجة.

الجدول 2-4

معنى السلُّم الرقمي ذي 11 درجة

العلامة	بند الانحطاط	
10	لا يمكن إدراكه	
9	يمكن إدراكه قليلاً	في مكان ما
8		في كل مكان
7	يمكن إدراكه	في مكان ما
6		في كل مكان
5	يمكن إدراكه بوضوح	في مكان ما
4		في كل مكان
3	مزعج	في مكان ما
2		في كل مكان
1	مزعج للغاية	في مكان ما
0		في كل مكان

ويُطلب من المشاهدين ملء استبيان يتكون من مربعين (يسميان "A" و"B") في كل خلية اختبار أساسية (BTC)، وكتابة علامة، في كل من المربعين، يختارونها من السلم العددي ذي 11 درجة.

ويقدم الشكل 2-14 مثالاً على ورقة إسناد الدرجات في جلسة تتكون من 24 خلية اختبار أساسية (BTC).

وفي كل خلية اختبار أساسية (BTC)، يقوم المشاهدون بملء كل من المربع المحدد بالحرف **A** (لتقييم مقطع الفيديو المعروض على أنه المقطع الأول) والمربع المحدد بالحرف **B** (لتقييم مقطع الفيديو المعروض على أنه المقطع الثاني).

ويسمح عرض مقطع الفيديو الأصلي غير المتردي بأن يقوم الخبراء بتقييم أي ترد بسهولة أكبر.

وينبغي شرح معنى السلُّم الرقمي ذي 11 درجة بعناية خلال "الجلسات التدريبية" على النحو المبين أدناه.

5-A8 تصميم الاختبار وإنشاء الجلسة

ينبغي لمصمم الاختبار أن يرتب عروض خلية الاختبار الأساسية (BTC) ترتيباً عشوائياً، بحيث لا يظهر مقطع الفيديو نفسه ولا المقطع المتردي نفسه مرتين متتاليتين.

وينبغي أن تبدأ أي جلسة عرض "بمرحلة تثبيت الاستقرار" بما في ذلك "أفضل"، و"أسوأ" خلية اختبار أساسية (BTC) وتلك "متوسطة الجودة" بين الخلايا المدرجة في كل جلسة اختبار. وسيتيح ذلك للمشاهدين انطباعاً مباشراً عن نطاق النوعية، منذ بداية جلسة الاختبار.

وإذا طالت جلسة العرض لأكثر من 20 دقيقة، ينبغي لمصمم الاختبار تقسيمها إلى جلستي مشاهدة منفصلتين (أو أكثر)، لا تتجاوز مدة كل منهما 20 دقيقة. وفي هذه الحالة، ينبغي تقديم "مرحلة تثبيت الاستقرار" قبل كل جلسة عرض.

6-A8 التدريب

حتى إذا كان يُرتقب استعمال هذا الإجراء بمشاركة الخبراء، يفضل تنظيم جلسة عرض تدريبية قصيرة (5-6 خلايا اختبار أساسية (BTC)) قبل كل تجربة.

ويمكن أن تكون مواد الفيديو المستعملة في جلسة التدريب هي نفسها التي ستُستعمل أثناء الجلسات الفعلية، ولكن ينبغي أن يكون ترتيب العروض مختلفاً.

وينبغي تدريب المشاهدين على استعمال السلُّم العددي ذي 11 درجة بمطالبتهم النظر بعناية في مقاطع الفيديو المعروضة مباشرةً بعد الرسالة "A" و"B" على الشاشة، والتحقق مما إذا كان يمكنهم رؤية أي فرق عن مقطع الفيديو المعروض على أنه المقطع الأول (تتابع المصدر المرجعي (SRC)).

7-A8 جمع البيانات ومعالجتها

ينبغي جمع العلامات في نهاية كل جلسة وتسجيلها على جدول بيانات إلكتروني لحساب القيم المتوسطة.

ويفضَّل إجراء "فرز لاحق" للمشاهدين، باستعمال ارتباط بيرسون (Pearson) الخطي.

وينبغي تطبيق دالة "الترابط" مع الأخذ في الاعتبار جميع علامات كل موضوع فيما يتعلق بمتوسط علامات الرأي (MOS)؛ ويمكن تحديد عتبة لتحديد كل مشاهد على أنه "مقبول" أو "مرفوض" (تقترح التوصية ITU-T P.910 رفض المواضيع ما دون عتبة الاستبعاد 0,75).

8-A8 شروط استعمال نتائج بروتوكول مشاهدة الخبراء

يمكن استعمال بروتوكول مشاهدة الخبراء (EVP) عندما لا يسمح الوقت والموارد بإجراء تجربة تقدير شخصاني رسمية.

ويحتاج البروتوكول EVP إلى وقت أقل من التقدير الشخصاني الرسمي ويمكن تنفيذه في بيئة "غير رسمية" بافتراض أن البيئة المحيطة التي ينفَّذ فيها محمية من أي تشويش خارجي مرئي ومسموع.

وتتعلق الشروط الإلزامية الوحيدة بالإضاءة المحيطة وشروط المشاهدة (شاشة العرض وزاوية المراقبة ومسافة المشاهدة) على النحو المبين في الفقرات أعلاه.

9-A8 قيود استعمال نتائج بروتوكول مشاهدة الخبراء (EVP)

حتى لو أثبت بروتوكول مشاهدة الخبراء قدرته على تقديم نتائج مقبولة بتسعة مشاهدين فقط، فإن متوسط علامات الرأي (MOS) الذي تقدمه تجربة بروتوكول مشاهدة الخبراء (EVP) لا يمكن اعتباره بديلاً من النتائج التي يمكن الحصول عليها بتجربة تقدير شخصاني رسمية.

ويمكن استعمال بيانات متوسط علامات الرأي (MOS) المحصَّلة باستعمال بروتوكول مشاهدة الخبراء (EVP) للحصول على مؤشر أولي لمستوى التردي.

ويمكن استعمال بيانات متوسط علامات الرأي (MOS) المحصَّلة باستعمال بروتوكول مشاهدة الخبراء (EVP) لوضع تصنيف أولي لخطط معالجة الفيديو قيد التقييم.

وحيثما يؤخذ بتجربة بروتوكول مشاهدة الخبراء (EVP) على أنها مريحة أو ضرورية، يمكن القيام بها على التوازي في المزيد من المواقع، على افتراض تطابق شروط المشاهدة ومسافة المشاهدة وتصميم الاختبار.

وإذا كان عدد المشاهدين الخبراء المشاركين في تجربة بروتوكول مشاهدة الخبراء نفسها، وأيضاً في حال إجراء التجربة في مواقع مختلفة، يساوي 15 أو أكثر، تمكن معالجة البيانات الشخصية الخام للحصول على متوسط علامات الرأي (MOS) والانحراف المعياري وبيانات فترة الثقة، التي يمكن أن تساعد في إجراء تصنيف أدق للحالات قيد الاختبار. وفي هذه الحالة الأخيرة، يمكن إجراء تحليل إحصائي استنتاجي أدق، من قبيل اختبار توزع احتمالات T.

المرفق 1
(إعلامي)
بالملحق 8
بالجزء 2
تطبيق بروتوكول مشاهدة الخبراء وسلوكهم في حضور عدد كبير من المقيّمين الخبراء

يقدم هذا المرفق الإعلامي معلومات عن نتائج تقييمَين شخصيَين مختلفَين للقطات فيديوية بتشفير HD وUHD باستعمال بروتوكول مشاهدة الخبراء أُجريا في الاجتماع السابع عشر بعد المائة (117) لفريق الخبراء المعني بالصور المتحركة من خلال تطبيق أحكام الملحق 8 لتصنيف أسلوبَي تشفير المصدر المختلفَين بطريقة سريعة وموثوقة.

ونظراً إلى أن عدداً كبيراً من الخبراء حضروا الاجتماع السابع عشر بعد المائة لفريق الخبراء المعني بالصور المتحركة، فاق عدد المقيّمين المشاركين في الجلستين EVP بشكل كبير العدد الموصى به في الملحق 8 بالجزء 2 من هذه التوصية والبالغ 9 مقيّمين؛ وحضر 30 خبيراً جلسة الاختبار HD EVP أما جلسة الاختبار UHD EVP فحضرها 32 خبيراً.

وأتاحت المشاركة الواسعة لمقيّمين خبراء الفرصة لتحليل بيانات متوسط علامات الرأي (MOS) من أجل التحقق من مستوى الاعتمادية الملازمة لاستعمال الملحق 8 عند تصنيف اللقطات الفيديوية المشفرة.

وفي إطار هذا التقييم، تم تشكيل أربع مجموعات من المشاهدين (أي مجموعات من 9 مشاهدين و12 مشاهداً و15 مشاهداً و18 مشاهداً) قاموا بمقارنة القيم MOS التي حصلت عليها المجموعة التي تضم 9 خبراء والقيم MOS التي حصلت عليها المجموعات التي تضم 12 و15 و18 مشاهداً على التوالي.

والهدف المنشود هو مقارنة التصنيف المتحصل عليه من جانب 9 خبراء (وبالتالي وفقاً للبروتوكول EVP) مع التصنيفات المتحصل عليها من المجموعات التي تضم 12 و15 و18 خبيراً (على غرار اختبار التقييم الشخصاني الرسمي).

ويتضح من الشكل 2-15 (تجربة مع محتوى UHD) والشكل 2-16 (تجربة مع محتوى HD) أن نتائج التصنيف متشابهة للغاية فيما يخص جميع الحالات الأربع.

وباعتبار أن النتائج التي حصلت عليها المجموعة التي تضم 18 مشاهداً "حقيقة أساسية"، يمكن رسم المخططَين البيانيَين في الشكلين 2-15 و2-16 لتصنيف نقاط الاختبار وفقاً للقيم MOS التي حصلت عليها المجموعة التي تضم 18 مشاهداً (الخط الأحمر المتواصل).

وتبين الخطوط الأخرى في المخططيَن البيانيَين النتائج التي حصلت عليها المجموعة التي تضم 9 مشاهدين (الخط الأحمر المنقط)، والمجموعة التي تضم 12 مشاهداً (الخط الأزرق المتقطع) والمجموعة التي تضم 15 مشاهداً (الخط الأخضر المتواصل).

ويمكن أن يُلاحظ من النتائج الموضحة في الشكلين 2-15 و2-16 ما يلي:

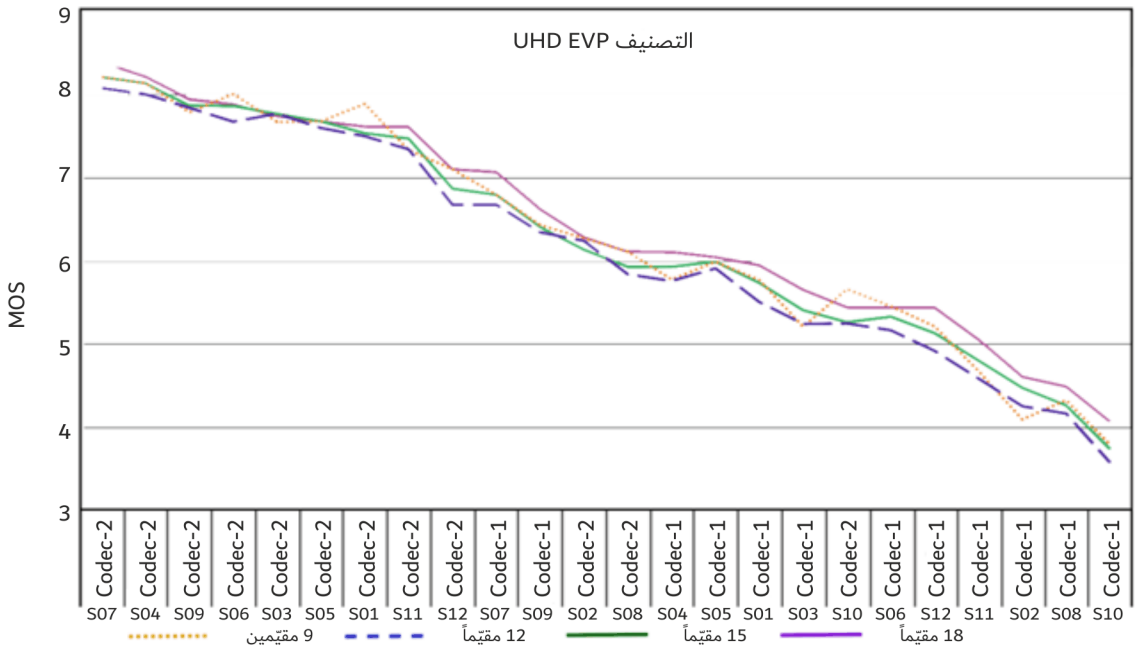
– يُظهر الشكلان المتعلقان بالمجموعتين اللتين تضمان 15 و18 مشاهداً منحنىً متجانساً لقيم MOS تتراوح بين جودة عالية وجودة منخفضة؛

– يُظهر الشكلان المتعلقان بالمجموعتين اللتين تضمان 9 مشاهدين و12 مشاهداً بعض "الانعكاسات" للتصنيف بالمقارنة مع الرسم البياني المتعلق بالمجموعة التي تضم 18 مشاهداً، حتى وإن كان اختلاف العلامات محدوداً نوعاً ما.

وفي الختام، تبين التجارب EVP الموصوفة هنا أداءً جيداً للغاية للبروتوكول EVP مما يؤكد ما جاء في نص الملحق 8 أي أن استعمال البروتوكول EVP، على الرغم من أنه لا يمكن اعتباره بديلاً بالكامل للتجربة الشخصانية الرسمية، يمكن أن يكون بمثابة إجراء تقييمي مستقر يؤدي إلى نتائج قريبة جداً من تلك المتحصل عليها عندما يكون عدد أكبر بكثير من المشاهدين متيسراً مع إجراء تقييم شخصي رسمي.

الشكل 2-15

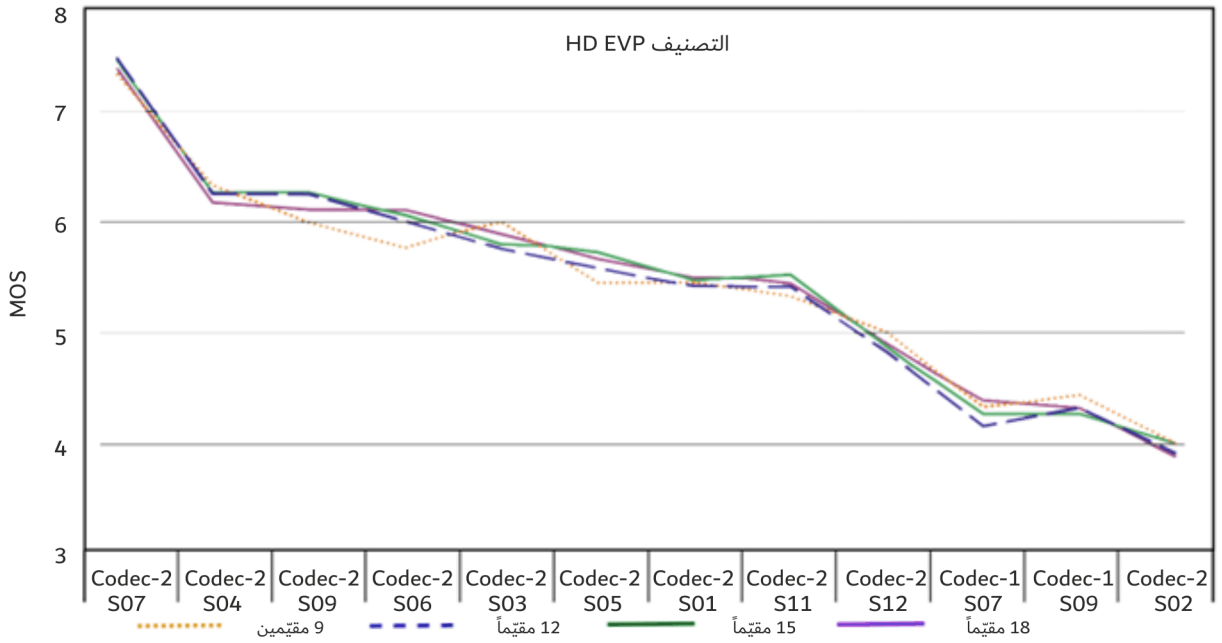
تصنيف التجربة UHD بحسب عدد المقيّمين



BT.0500-02-15

الشكل 2-16

تصنيف التجربة HD بحسب عدد المقيّمين



BT.0500-02-16

الجزء 3

منهجيات التقدير الشخصاني لنوعية الصورة الخاصة بالتطبيق

ينبغي مراعاة الاعتبارات الخاصة بالتطبيق عند تصميم اختبارات التقدير الشخصاني. ويقدم الجزء 3 إرشادات للتقدير الشخصاني للصورة في أنساق وتطبيقات الصور ذات الصلة:

الملحق 1 التقدير الشخصاني لأنظمة التلفزيون الرقمي عادي الوضوح (SDTV)

الملحق 2 التقدير الشخصاني لأنظمة التلفزيون عالي الوضوح (HDTV)

الملحق 3 التقدير الشخصاني لنوعية الصور الهجائية العددية والصور البيانية في التلتيكس وفي خدمات مماثلة

الملحق 4 التقدير الشخصاني لنوعية الصورة في الخدمات متعددة البرامج

الملحق 5 مشاهدة الخبراء لتقدير نوعية الأنظمة الخاصة بأجهزة العرض الرقمية للتصوير الرقمي على الشاشات الكبيرة في المسارح

الملحق 6 التقدير الشخصاني لنوعية الفيديو في التطبيقات متعددة الوسائط

الملحق 7 التقدير الشخصاني لأنظمة التلفزيون ثلاثي الأبعاد ومجسم الصورة

الملحق 1
بالجزء 3

التقدير الشخصاني لأنظمة التلفزيون الرقمي عادي الوضوح (SDTV)

1-A1 مقدمة

يقدم هذا الملحق، المزمع استعماله بالاقتران مع الجزأين 1 و2 من هذه التوصية، تفاصيل تتعلق بتطبيق الطرائق العامة الواردة في التوصية على التقديرات الشخصانية للأنظمة الرقمية التي تقدم مستويات نوعية مماثلة لأنظمة التلفزيون الرقمية أو قريبة منها. وتتعلق التفاصيل الإجرائية الواردة في هذه التوصية، جنباً إلى جنب مع المعلومات الأساسية ذات الصلة، باختبارات الكودكات (أو الأنظمة) المستعملة لنقل المواد الناشئة وفقاً للتوصية ITU-R BT.601 في تطبيقات المساهمة والتوزيع وكذلك تلك المستعملة في تطبيقات البث.

فبالنسبة لتطبيقات التوزيع، يمكن التعبير عن مواصفات النوعية من حيث الحكم الشخصاني للمراقبين. ولذلك يمكن نظرياً أن يتم تقدير الكودكات هذه بطريقة شخصانية مقابل هذه المواصفات. غير أنه لا يمكن من الناحية النظرية تحديد نوعية الكودك المصمم لتطبيقات المساهمة من حيث معلمات الأداء الشخصاني لأن مخرجاته ليست مخصصة للمشاهدة الفورية، ولكن للمعالجة اللاحقة في الاستوديو، و/أو التخزين و/أو التشفير لمواصلة الإرسال. وبسبب صعوبة تحديد هذا الأداء لمجموعة متنوعة من عمليات ما بعد المعالجة، فإن النهج المفضل هو تحديد أداء سلسلة من المعدات، بما في ذلك وظيفة ما بعد المعالجة، التي يُعتقد أنها ممثلة لتطبيق المساهمة العملي. وعادةً ما تتكون هذه السلسلة من كودك، وتليها وظيفة ما بعد المعالجة في الاستوديو (أو كودك آخر في حالة تقدير نوعية المساهمة الأساسية)، ويليها كودك آخر قبل وصول الإشارة إلى المراقب. ويعني اعتماد هذه الاستراتيجية لتحديد الكودكات لتطبيقات المساهمة أنه يمكن أيضاً استعمال إجراءات القياس الواردة في هذه التوصية لتقديرها.

وفي مجال التقدير الشخصاني، حيث توجد خبرة كثيرة، يمكن التوصية بشروط ومنهجيات الاختبار. غير أنه يجب أن نتذكر عند تحديد أهداف النوعية أو الانحطاط، أن الطرائق الحالية لا يمكن أن تعطي تقديرات شخصانية مطلقة بل بالأحرى نتائج تتأثر إلى حد ما باختيار المرجع و/أو شروط المرتكز. ويمكن استعمال نفس المنهجيات للكودات بطول الكلمة الثابت والمتغير، وللكودكات ضمن المجالات وبين الأرتال على الرغم من أن اختيار تتابع صور الاختبار يمكن أن يتأثر.

والطريقة الأكثر موثوقية لتقييم ترتيب الكودكات عالية النوعية هي تقدير جميع الأنظمة المرشحة في نفس الوقت تحت شروط متطابقة. ويجب استعمال الاختبارات التي تجرى بشكل منفصل، والتي تتضمن اختلافات دقيقة في النوعية، للتوجيه لا أن تكون دليلاً لا جدال فيه على التفوق.

ويمكن تحديد مقياس شخصاني مفيد عن طريق الانحطاط كدالة لنسبة خطأ البتات الذي يحدث في وصلة الإرسال بين المشفر ومفكك الشفرة. وفي الوقت الحاضر، لا توجد معرفة تجريبية كافية لإحصاءات أخطاء الإرسال الحقيقية للتوصية بمعلمات لنموذج يراعي تجمع الأخطاء أو الرشقات. وحتى تكون هذه المعلومات متاحةً يمكن استعمال الأخطاء الموزعة إحصائياً بأسلوب بواسون.

2-A1 شروط المشاهدة

شروط المشاهدة العامة للتقديرات الشخصانية هي تلك الواردة في الفقرة 2 من الجزء 1. وترد شروط المشاهدة الخاصة للتقديرات الشخصانية للأنظمة الرقمية في الفقرات التالية.

1.2-A1 البيئة المختبرية

الغرض من البيئة المختبرية هو توفير شروط الحساسية اللازمة لفحص الأنظمة. وترد شروط المشاهدة الخاصة للتقديرات الشخصانية في البيئة المختبرية في الجدول 3-1 التالي.

الجدول 3-1

شروط المشاهدة الخاصة للتقديرات الشخصانية في الأنظمة الرقمية في البيئة المختبرية

الشرط	البند	القيم
a	نسبة مسافة المشاهدة إلى ارتفاع الصورة	H 4 و H 6 [1]
b	ذروة النصوع	70 cd/m^2
c	زاوية المشاهدة التي تقع قبالة جزء الخلفية الذي يفي بالمواصفات	W °57 × H °43≤
d	العرض	شاشة عالية النوعية. الحجم ≤ 20" (50 cm) [2]

[1] H6 هي مسافة المشاهدة الخاصة بالتصميم (DVD) لتقدير الأنظمة الرقمية عادية الوضوح، ولكن يكون استعمال المقيّمين عند H4 مقبولاً أيضاً، شريطة تقديم النتائج بشكل منفصل.

[2] نظراً لوجود بعض الأدلة على أن حجم العرض يمكن أن يؤثر على نتائج التقديرات الشخصانية، يُطلب من القائمين بالتجربة الإبلاغ مباشرةً عن حجم الشاشة وأنواع ونماذج الشاشات المستعملة في أي تجارب.

2.2-A1 البيئة المنزلية

الغرض من هذه البيئة هو توفير وسيلة لتقدير النوعية عند جانب المستهلك في سلسلة التلفزيون الرقمي. وترد في الجدول 3-2 شروط المشاهدة الخاصة بالتقديرات الشخصانية لأنظمة التلفزيون الرقمية عادية الوضوح (SDTV) في البيئة المنزلية.

الجدول 3-2

شروط المشاهدة الخاصة للتقديرات الشخصانية في الأنظمة الرقمية في البيئة الرقمية

الشرط	البند	القيم
a	نسبة مسافة المشاهدة إلى ارتفاع الصورة	H 6
b	حجم الشاشة للنسق 4/3	من 25" إلى 29"[1]
c	حجم الشاشة للنسق 16/9	من 32" إلى 36"[1]
d	معيار شاشة العرض	SDTV
e	ذروة النصوع	200 cd/m^2
f	نصوع البيئة على الشاشة (ينبغي قياس الضوء الساقط من البيئة على الشاشة عمودياً على الشاشة)	200 Lux

[1] يفي حجم الشاشة هذا بقواعد مسافة المشاهدة المفضلة (PVD) لمسافة قدرها 6 H.

3-A1 طرائق التقدير

1.3-A1 تقييمات نوعية الصورة الأساسية

عند تقدير كودك ما لتطبيقات التوزيع، تشير هذه النوعية إلى الصور التي تم فك تشفيرها بعد مرور واحد من خلال زوج كودكات. وبالنسبة إلى كودكات المساهمة، يمكن تقدير النوعية الأساسية بعد عدة كودكات في سلسلة، من أجل محاكاة تطبيق مساهمة نموذجي.

وعندما يكون مدى النوعية المطلوب تقديره صغيراً، كما هو الحال عادةً في الكودكات التلفزيونية، تكون منهجية الاختبار المستعملة هي الصيغة II من طريقة سلم النوعية المستمرة ثنائية الحافز الموصوفة في هذه التوصية. وسيتم استعمال تتابع المصدر الأصلي بوصفه الشرط المرجعي. ويتعين إيلاء مزيد من الاعتبار لمدة تتابعات العرض. وفي الاختبارات الأخيرة المتعلقة بالكودكات لمكون الفيديو بسوية 4:2:2، اعتُبر من المفيد تعديل العرض عن العرض الوارد في هذه التوصية. وقد استُعملت صور مركبة كمرجع إضافي لتوفير سوية نوعية أقل للحكم على أداء الكودك مقابلها.

ويوصى باستعمال ستة تتابعات للصور على الأقل في التقدير، زائد تتابع إضافي لاستعماله لأغراض التدريب قبل بدء التجربة. وينبغي أن يتراوح مدى التتابعات بين التتابعات الحساسة إلى حد ما والتتابعات الحساسة في سياق تطبيق تخفيض معدل البتات قيد النظر.

وفي جميع أقسام هذا الملحق، يتم التشديد على أهمية اختبار الكودكات الرقمية بتتابعات للصور حساسة في سياق تخفيض معدل البتات التلفزيونية. ولذلك، من المنطقي التساؤل عن مدى حساسية تتابع معين للصور لمهمة معينة من مهام تخفيض معدل البتات، أو ما إذا كان تتابع ما أكثر حساسية من تتابع آخر. وهناك إجابة بسيطة ولكن غير مفيدة جداً وهي أن "الحساسية" تعني أشياء مختلفة جداً للكودكات المختلفة. فعلى سبيل المثال، بالنسبة إلى الكودكات ضمن المجالات، يمكن أن تكون صورة ثابتة تحتوي على الكثير من التفاصيل حساسةً فعلاً، بينما بالنسبة إلى الكودكات بين الأرتال القادرة على استغلال أوجه التشابه من رتل إلى رتل، لن يمثل هذا المشهد نفسه أي صعوبة على الإطلاق. وستكون بعض التتابعات التي تستعمل نسيجاً متحركاً وحركةً معقدةً حساسةً لجميع فئات الكودكات، وبالتالي، يكون توليد وتحديد هذه الأنواع من التتابعات أكثر فائدةً. ويمكن أن تأخذ الحركة المعقدة شكل حركات يمكن أن يتنبأ بها المراقب ولكن لا يمكن أن تتنبأ بها خوارزميات التشفير، مثل الحركة الدورية الملتوية.

وكشف أحد فحوص المقاييس الإحصائية المحتملة لحساسية الصورة، مثلاً بالطرائق الترابطية والطرائق الطيفية والطرائق الإنتروبية المشروطة وما إلى ذلك، عن إجراء بسيط ولكن مفيد يستند إلى قياس الأنتروبيا التكيفية ضمن المجالات/بين الأرتال. واستُعملت هذه الطريقة "لمعايرة" تتابعات الصور المقترحة لاستعمالها في تجارب قطاع الاتصالات الراديوية الخاصة بالكودكات في حالة 34 و45 و140 Mbit/s، وأثبتت فائدتها في اختيار التتابعات المستعملة. وتُجرى مثل هذه القياسات على تتابعات الصور بسهولة أكبر عن طريق نقلها إلى حواسيب معالجة الصور وتحليلها بواسطة البرمجيات.

وعندما لا يتوافر النفاذ إلى هذه التقنيات، يمكن استعمال الإرشادات العامة التالية بشأن كيفية اختيار المواد الحساسة.

أ) *الكودكات ضمن المجالات بطول الكلمة الثابت*

في حين أنه من الممكن والسليم تقدير الكودكات هذه على الصور الثابتة، يوصى باستعمال التتابعات المتحركة نظراً لأنه من الأسهل مراقبة عمليات ضوضاء التشفير وهذا أكثر واقعيةً في التطبيقات التلفزيونية. وإذا استُعملت الصور الثابتة في عمليات الحاسوب لمحاكاة الكودكات، ينبغي إجراء المعالجة عبر تتابع التقدير بأكمله من أجل الحفاظ على الجوانب الزمنية لأي ضوضاء مصدرية، على سبيل المثال. ويجب أن تحتوي المشاهد المختارة على أكبر قدر ممكن من التفاصيل التالية: مناطق ذات ملامح ثابتة ومتحركة (بعضها بملامح ملونة)؛ وأجسام ثابتة ومتحركة ذات حواف حادة عالية التباين في اتجاهات مختلفة (بعضها ملون)؛ ومناطق ثابتة بمتوسط الرمادي ذات اللون الواحد. وينبغي أن يُظهر تتابع واحد على الأقل في المجموعة ضوضاء مصدرية يمكن إدراكها قليلاً وينبغي أن يكون تتابع واحد على الأقل اصطناعياً (أي يولده الحاسوب) بحيث يكون خالياً من عيوب الكاميرا مثل فتحة المسح الضوئي والتأخر.

ب) *الكودكات بين الأرتال بطول الكلمة الثابت*

ينبغي أن تحتوي جميع مشاهد الاختبار المختارة على حركة وأكبر عدد ممكن من التفاصيل التالية: مناطق متحركة ذات *ملامح* (بعضها ملون)؛ وكائنات ذات حواف حادة عالية التباين تتحرك في اتجاه عمودي على هذه الحواف وفي اتجاهات مختلفة (بعضها ملون). وينبغي أن يُظهر تتابع واحد على الأقل في المجموعة ضوضاء مصدرية يمكن إدراكها قليلاً وينبغي أن يكون تتابع واحد على الأقل اصطناعياً.

ج) *الكودكات ضمن المجالات بطول الكلمة المتغير*

يوصى باختبار هذه الكودكات بتتابع متحرك للصور لنفس الأسباب المتعلقة بالكودكات بطول الكلمة الثابت. وجدير بالذكر أنه نتيجة تشفيرها بطول الكلمة المتغير ومخزن الذاكرة المرتبط بها، يمكن لهذه الكودكات أن توزع سعة بتات التشفير ديناميكياً في جميع أجزاء الصورة. وبالتالي، على سبيل المثال، إذا كان نصف الصورة يتكون من سماء بدون سمات ولا يتطلب العديد من البتات للتشفير، يتم توفير السعة للأجزاء الأخرى من الصورة التي يمكن بالتالي إعادة إنتاجها بنوعية عالية حتى إذا كانت حساسة. والاستنتاج المهم من ذلك هو أنه إذا كان لتتابع الصور أن يكون حساساً لمثل هذا الكودك، فينبغي تفصيل محتوى كل جزء من الشاشة. وينبغي أن يكون ممتلئاً بملامح متحركة وثابتة، وبأكبر قدر ممكن من اختلاف الألوان والكائنات ذات الحواف الحادة عالية التباين. وينبغي أن يُظهر تتابع واحد على الأقل في المجموعة ضوضاء مصدرية يمكن إدراكها قليلاً وينبغي أن يكون تتابع واحد على الأقل اصطناعياً.

د) *الكودكات بين الأرتال بطول الكلمة المتغير*

هذه هي الفئة الأكثر تعقيداً من الكودكات والنوع الذي يتطلب أكثر المواد تطلباً للتأكيد عليه. ولا ينبغي ملء كل جزء من المشهد بالتفاصيل كما هو الحال في حالة الكودكات ضمن المجالات بطول الكلمة الثابت فحسب، ولكن ينبغي أيضاً أن تُظهر هذه التفاصيل الحركة أيضاً. وعلاوةً على ذلك، نظراً لأن العديد من الكودكات تستعمل طرائق تعويض الحركة، فينبغي أن تكون الحركة خلال التتابع معقدة. ومن الأمثلة على الحركة المعقدة هي: المشاهد التي تستعمل الزوم والمسح البانورامي للكاميرا بشكل متزامن؛ ومشهد يحتوي كخلفية على ستارة ذات ملامح أو مفصلة تتحرك بسبب الريح؛ ومشهد يحتوي على كائنات تدور في عالم ثلاثي الأبعاد؛ والمشاهد التي تحتوي على كائنات مفصلة تتزايد سرعتها عبر الشاشة. وينبغي أن تحتوي جميع المشاهد على حركة كبيرة للكائنات بسرعات مختلفة، وملامح وحواف عالية التباين فضلاً عن محتوىً متنوع من الألوان. وينبغي أن يُظهر تتابع واحد على الأقل في المجموعة ضوضاء مصدرية يمكن إدراكها قليلاً، وينبغي أن يكون لتتابع واحد على الأقل حركة كاميرا معقدة تم توليدها بواسطة الحاسوب من صورة ثابتة طبيعية (بحيث تكون خالية من الضوضاء وتأخر الكاميرا)، وينبغي أن يكون تتابع واحد على الأقل مولد بالكامل بالحاسوب.

2.3-A1 تقييم نوعية الصورة بعد المعالجة اللاحقة

الغرض من هذا التقدير هو السماح بالحكم على مدى ملاءمة كودك ما لتطبيقات المساهمة فيما يتعلق بعملية لاحقة معينة مثل اللون غير اللامع، والحركة البطيئة، والزوم الإلكتروني. وأدنى ترتيب للمعدات لمثل هذا التقدير هو تمرير وحيد من خلال الكودك قيد الاختبار، وتليه العملية اللاحقة قيد الدراسة، ويليها المشاهد. غير أنه قد يكون أكثر تمثيلاً لتطبيق المساهمة استعمال كودكات إضافية بعد العملية اللاحقة.

ومنهجية الاختبار التي يتعين استعمالها هي الصيغة II من طريقة سلم النوعية المستمرة ثنائية الحافز. غير أن الشرط المرجعي في هذه الحالة هو المصدر الخاضع لنفس المعالجة اللاحقة التي خضعت لها الصور التي تم فك تشفيرها. وإذا ارتئيت فائدة من إدراج مرجع أقل نوعية، فينبغي أن يخضع أيضاً لنفس العملية اللاحقة.

وتخضع تتابعات الاختبار المطلوبة للتقديرات اللاحقة لنفس معايير الحساسية تماماً التي تخضع لها التتابعات للتطبيقات الرقمية الأخرى. غير أنه قد يكون من الصعب تحقيق ذلك في تتابعات مفتاح الألوان الأمامية لأنها عادةً ما تحتوي على نسبة كبيرة من الخلفية الزرقاء التي ليس لها سمات.

وبسبب القيود العملية التي تؤدي إلى تقدير كودك ما بالعديد من العمليات اللاحقة، قد يكون عدد تتابعات صور الاختبار المستعمل ثلاثاً كحد أدنى مع تتابع إضافي متاح لأغراض العرض التوضيحي. وستعتمد طبيعة التتابعات على مهمة المعالجة اللاحقة التي تجري دراستها ولكن ينبغي أن تتراوح بين الحساسة إلى حد ما والحساسة في سياق تخفيض معدل البِتات التلفزيونية وللعملية قيد النظر. ولتقدير الحركة البطيئة، قد يكون معدل عرض قدره عُشر معدل المصدر مناسباً.

3.3-A1 تقييمات خصائص الأعطال

في التقديرات الشخصانية للانحطاطات في صور الكودكات الناتجة عن عيوب في الإرسال أو قناة البث، ينبغي اختيار ما لا يقل عن خمس، ولكن يفضل أكثر، نسب لأخطاء البتات أو شروط الإرسال/البث، وأن تكون متباعدة تقريباً لوغاريتمياً وأخذ عينات مناسبة من المدى الذي يؤدي إلى انحطاط الكودكات من "لا يمكن إدراكه" إلى "مزعج للغاية".

ومن الممكن أن تكون تقديرات الكودكات مطلوبةً عند نسب أخطاء بتات الإرسال التي تؤدي إلى صور مرئية عابرة نادرة لدرجة أنه قد لا يُتوقع حدوثها خلال فترة تتابع اختبار مدتها 10 ثوان. ومن الواضح أن توقيت العرض المقترح هنا غير مناسب لمثل هذه الاختبارات.

وإذا كانت تسجيلات خرج الكودكات تحت شروط نسبة أخطاء بتات منخفضة إلى حد ما (مما ينتج عنه عدد صغير من الصور المرئية العابرة خلال فترة 10 ثوان) ستُجرى لمعالجتها لاحقاً في العروض للتقديرات الشخصانية، فينبغي توخي الحذر لضمان أن التسجيل المستعمل نموذجي لخرج الكودكات على مدى فترة زمنية أطول.

وبسبب الحاجة إلى استكشاف أداء الكودكات عبر مجموعة من نسب أخطاء بتات الإرسال، تشير القيود العملية إلى أن ثلاثة تتابعات لاختبار الصورة مع تتابع توضيحي إضافي ستكون كافيةً على الأرجح. وينبغي أن تكون مدة التتابعات 10 ثوان ولكن ينبغي ملاحظة أن مشاهدي الاختبار قد يفضلون مدة 15-30 ثانية. وينبغي أن تتراوح بين الحساسة إلى حد ما والحساسة في سياق تخفيض معدل البتات التلفزيونية.

وبالنظر إلى أن الاختبارات ستشمل المدى الكامل للانحطاط، فإن طريقة سلم الانحطاط ثنائي الحافز مناسبة وينبغي استعمالها.

4.3-A1 خصائص أعطال محتوى الصورة

يرد المفهوم العام لخصائص أعطال الصورة في الملحق 1 بالجزء 1. ولتطبيق هذا المفهوم على أنظمة التلفزيون الرقمي عادي الوضوح، ينبغي اتباع الإجراء التالي.

1.4.3-A1 تعريف الحساسية

ينبغي تحديد مقياس معين يسمى "الحساسية" ويمثل خصائص نظام التلفزيون الرقمي قيد الاختبار ويقاس بمقياس شخصاني. وكمثال على نظام التلفزيون الرقمي، يتم استعمال MPEG-2 MP@ML ويتم تطبيق طريقة تحديد الكمية الثابتة للحساسية القائمة على الإنتروبية، والتي يرد وصف لها في التوصية ITU-R BT.1210.

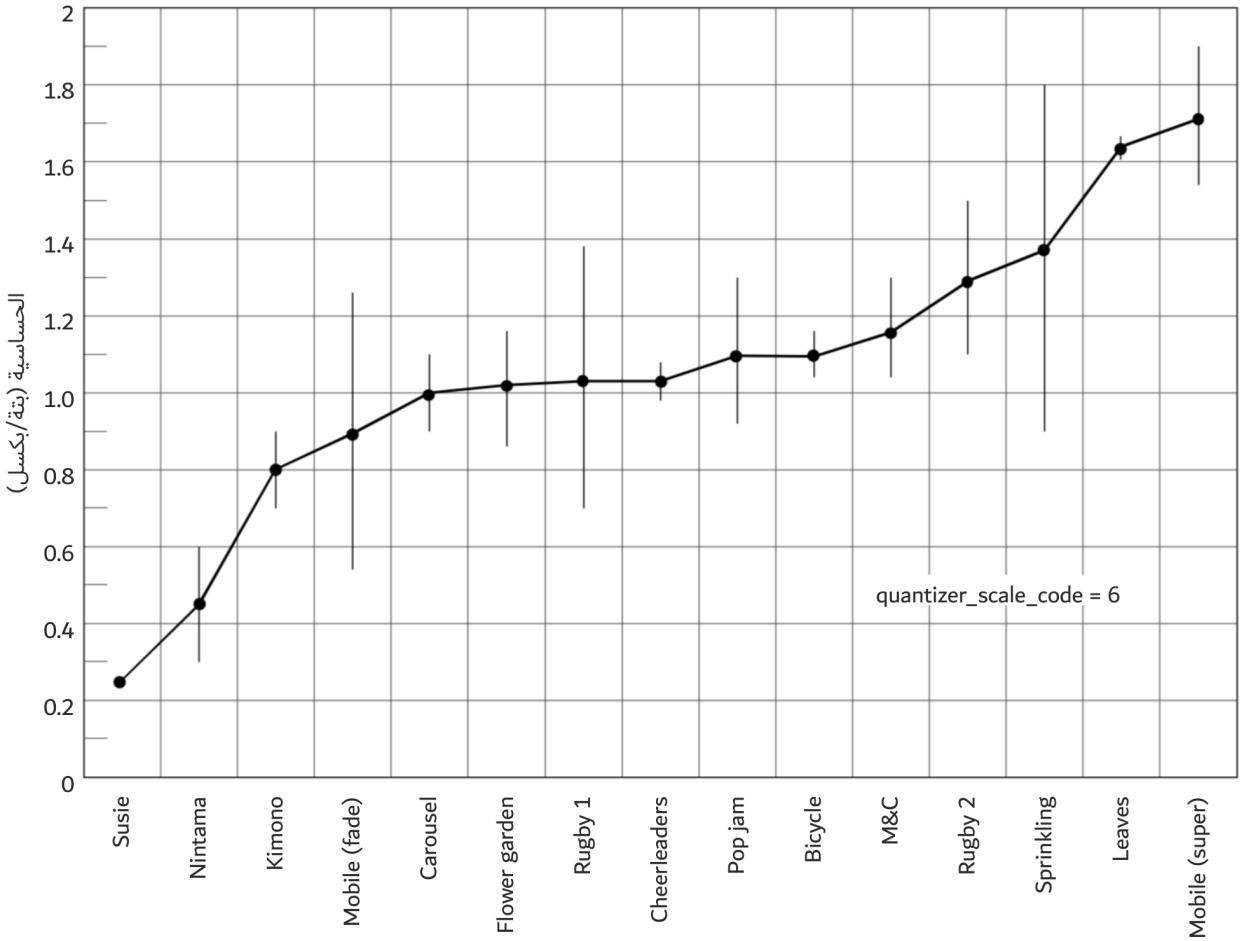
2.4.3-A1 إجراء الحصول على خصائص أعطال محتوى الصورة

– *المرحلة 1*: قياس حساسية تتابع الاختبارات المستعملة في التقديرات الشخصانية

يجري قياس حساسية تتابعات الاختبار المستعملة لإجراء التقدير الشخصاني الوارد وصفه في المرحلة 3 أدناه. ويوضح الشكل 3-1 المتوسط والانحراف المعياري لكل تتابع للنظام المستعمل كمثال. وتتراوح مقاييس الحساسية لمعظم التتابعات بين 0,8 و1,4 بتة/بكسل. ويكون الانحراف المعياري لبعض التتابعات كبيراً لأن محتوى الصورة يختلف اختلافاً كبيراً أثناء التتابع.

الشكل 3-1

المتوسط والانحراف المعياري لحساسية تتابعات الاختبار



BT.0500-03-1

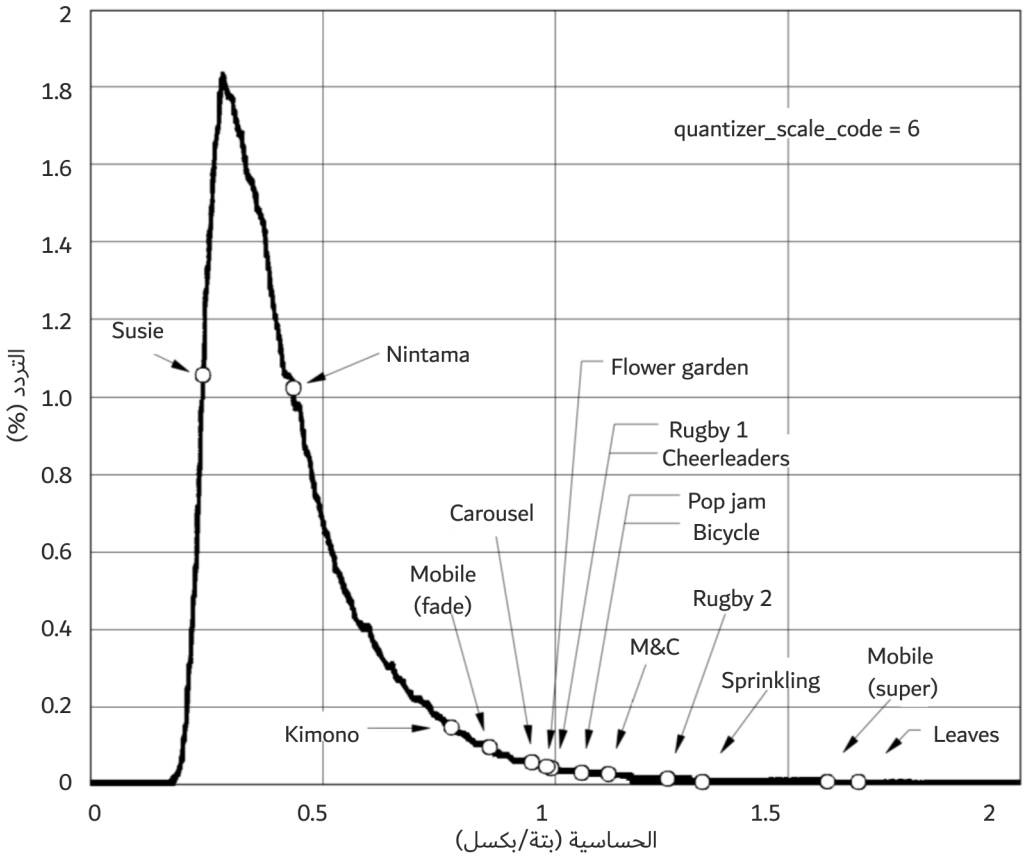
– *المرحلة 2*: قياس توزيع حساسية البرامج الإذاعية لفترة طويلة من الزمن

يُقاس توزيع حساسية البرامج التلفزيونية الإذاعية لفترة طويلة بما فيه الكفاية من الزمن، على سبيل المثال أسبوع واحد. ويوضح الشكل 3-2 مثالاً على التوزيع المقاس لمدة أسبوع واحد، أي ما مجموعه 130 ساعة للإشارات الإذاعية NTSC، والتي تم تحويلها

إلى إشارات *Y/C* المكونة من أجل القياس. وتم حساب تكرار حدوث الحساسية للبرامج التلفزيونية كل 5×10^{-3} بتة/بكسل. ويوضح هذا الشكل أيضاً أهمية حساسية تتابع الاختبار المستعمل في التقدير الشخصاني.

الشكل 3-2

توزيع حساسية البرامج الإذاعية وتتابعات الاختبار



BT.0500-03-2

– *المرحلة 3*: إجراء تقدير شخصاني لنوعية صورة النظام قيد الاختبار، والحصول على علاقة بين الحساسية والنوعية الشخصانية للصورة

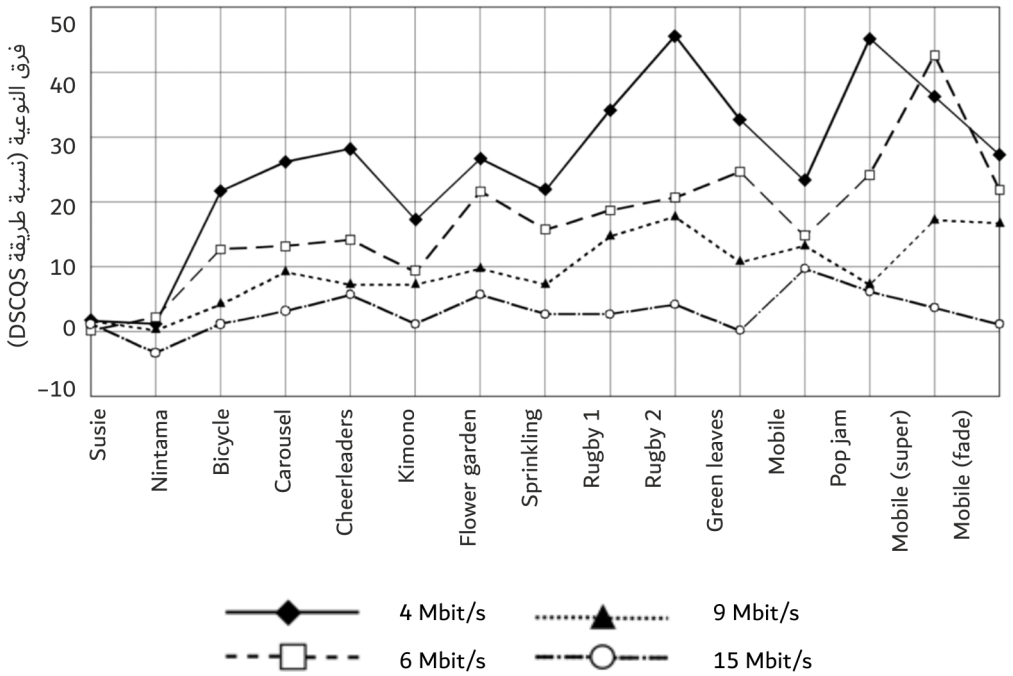
يجري تقدير نوعية صورة نظام التلفزيون الرقمي بطريقة DSCQS. ومن خلال الجمع بين نتيجة التقدير الشخصاني والحساسية التي تم الحصول عليها في المرحلة 1، يتم الحصول على العلاقة بين الحساسية وعلامات اختبار التقدير. ويوضح الشكل 3-3 نوعية الصورة للنظام المستعمل كمثال بمعدلات بتات قدرها 4 و6 و9 و15 Mbit/s. ويمثل فرق النوعية (DSCQS %) الوارد في الشكل التدهور مقارنةً بتتابع المكون الأصلي 4:2:2. ويوضح الشكل 3-4 العلاقة بين الحساسية واختلاف النوعية. وفي هذا المثال، افتُرضت علاقة خطية بين الحساسية ونوعية الصورة، وتم الحصول على خطوط الانحدار باستعمال طريقة المربعات الصغرى. ويوضح خط الانحدار عند كل معدل بتة في الشكل. وبشكل عام، يمكن تطبيق علاقة غير خطية حسب نتائج التقدير.

– *المرحلة 4*: الحصول على خصائص أعطال محتوى الصورة (النوعية مقابل تكرار الحدوث) من خلال الجمع بين نتائج المرحلة 3 (الحساسية مقابل النوعية) والمرحلة 2 (الحساسية مقابل تكرار الحدوث).

وعن طريق جمع النتائج التي تم الحصول عليها في المرحلتين 2 و3، يتم الحصول على خصائص أعطال محتوى الصورة، أي توزيع نوعية الصورة للبرامج التلفزيونية المشفرة رقمياً. ويتم تحويل تدهور الصورة في البرامج التلفزيونية الإذاعية إلى تكرار تراكمي لحالات الحدوث. ويبين الشكل 3-5 خصائص أعطال محتوى الصورة في النظام المستعمل كمثال.

الشكل 3-3

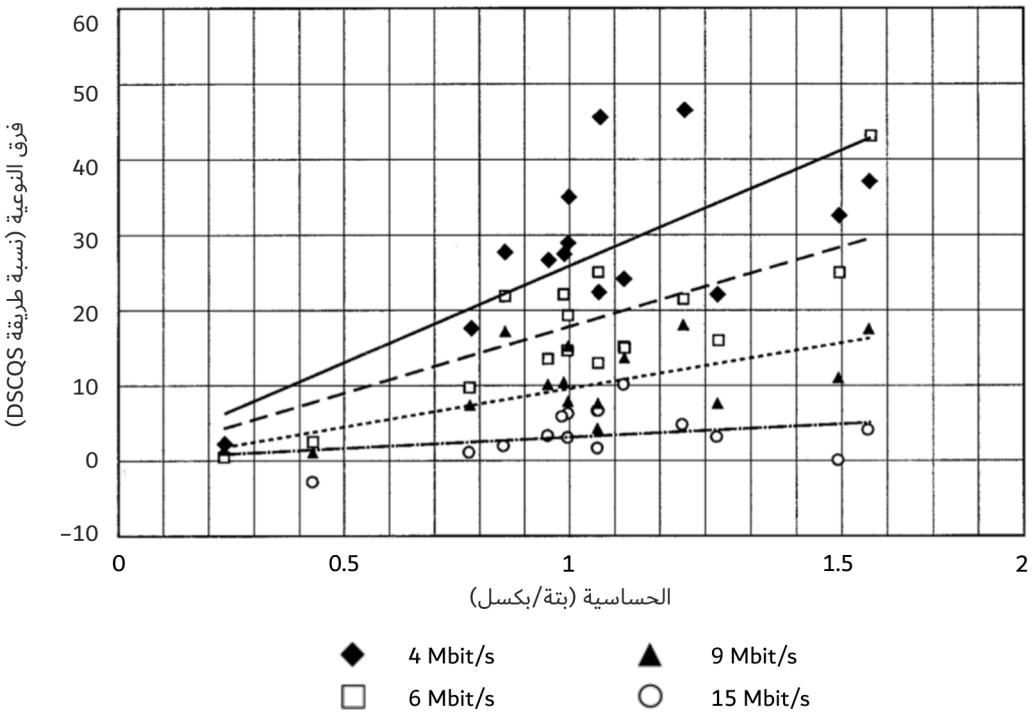
نتائج التقدير الشخصاني (MP@ML عند *H*6)



BT.0500-03-3

الشكل 3-4

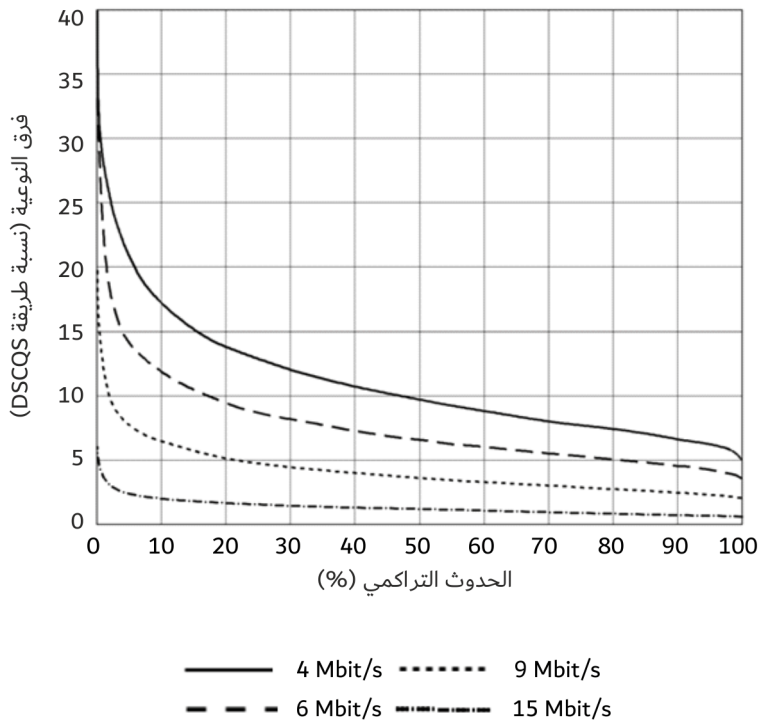
العلاقة بين الحساسية وعلامة التقدير (MP@ML عند *H*6)



BT.0500-03-4

الشكل 3-5

التكرار التراكمي لحدوث تدهور الصورة (MP@ML عند *H*6)



BT.0500-03-5

4-A1 مذكرات التطبيق

عندما لا يكون الحكم على نوعية الكودكات المطلقة أو الانحطاط مطلوباً، ولكن الترتيب فقط، أو حيث يكون تأكيد الترتيب الموجود من نتائج ثنائية الحافز مطلوباً، ينبغي استعمال طريقة المقارنات متزاوجة الحافز.

وكما هو موصوف في هذه التوصية، توفر الطريقة مقارنةً حساسةً ووسيلةً لتحديد مقياس للعلاقة بين أزواج من الأنظمة. ومن الممكن تمديد هذه الطريقة، لترتيب نوعية أو انحطاط أكثر من نظامين. وفي هذا النهج، يمكن الحصول على الترتيب العام من ترتيب جميع الأزواج المحتملة لتتابع الصور من قبل المراقبين.

ويكون التحليل معقداً لأن المراقب يستطيع أن يرتب، على سبيل المثال، الصورة A أفضل من الصورة B، والصورة B أفضل من الصورة C، ولكن أيضاً الصورة C أفضل من الصورة A. ويُطلق على ذلك اسم "العلاقة الثلاثية غير الانتقالية".

ومن مشاكل هذه الطريقة هي أن عدد العروض المطلوبة يزداد مع مربع عدد تتابعات صور الاختبار والكودكات، ويمكن أن يصبح غير عملي.

وإذا استُعملت قناة الإذاعة لتقديم تدفقات برامج متعددة أو مخططات تشفير قابلة للتوسع أو الترتيب الهرمي، فقد يكون من الضروري تعديل منهجية التقدير لمراعاة ما يلي:

– أن معيار الخدمة المقبولة قد لا يكون هو الشفافية في تشفير المصدر؛ وقد تكون، بدلاً من ذلك، قدرة النظام على توفير بديل قابل للتطبيق للخدمة التقليدية عند توزيع معين لمعدل البتات. وبناءً على ذلك، كمرجع في اختبارات النوعية، قد يكون من المناسب استعمال المواد على النحو الذي يقدمها النظام التقليدي وفق شروط الاستقبال النموذجية، بدلاً من استعمال المواد في شكل رقمي غير مضغوط. وعلاوةً على ذلك، قد يكون من المناسب استعمال مواد الاختبار المختارة لتمثيل مدى محتوى البرامج الحالية والمستقبلية (انظر الملحق 3 بالجزء 1). وفي الاختبارات، ينبغي أن تكون شروط

المشاهدة على النحو الموصوف في الجزء 1 وفي الفقرة 2-A1 من هذا الملحق، بينما ينبغي أن تكون طريقة الاختبار العامة هي طريقة سلم النوعية المستمرة ثنائية الحافز (الملحق 2 بالجزء 2)؛

– من المهم مراعاة قدرة النظام على الحفاظ على سلامة تدفقات البرامج الفردية وفق شروط التحميل الكامل للقناة وانحطاط الإرسال. وبناءً على ذلك، قد يكون من المناسب، في اختبارات الانحطاط، التأكد من التحميل الكامل للقناة واستعمال مجموعة من مستويات الانحطاط المختارة لتمثيل مجموعة شروط الاستقبال المحتملة (انظر الملحق 4 بالجزء 1). وفي الاختبارات، ينبغي أن تكون شروط المشاهدة كما هو موصوف في الجزء 1 وفي الفقرة 2-A1 من هذا الملحق، في حين ينبغي أن تكون طريقة الاختبار العامة هي طريقة سلم الانحطاط ثنائي الحافز (انظر الملحق 1 بالجزء 2).

ملاحظة – عندما يتم تقدير الأنظمة التماثلية والرقمية في نفس السياق، من المهم اختيار مجموعة من مواد الاختبار التي تعكس صعوبةً متوازنةً للأنظمة التماثلية والرقمية. وقد يكون من المفيد في هذه الحالة تطبيق إجراء تقدير متعدد الأبعاد كتحليل تكميلي.

الملحق 2
بالجزء 3

التقدير الشخصاني لنوعية صورة أنظمة التلفزيون عالي الوضوح (HDTV)

1-A2 بيئة المشاهدة

ينبغي أن تكون بيئة المشاهدة كما هو موصوف في الفقرة 2 من الجزء 1 ما لم يذكر خلاف ذلك في الجدول 3-3 أدناه.

الجدول 3-3

شروط المشاهدة للتقدير الشخصاني لنوعية صورة أنظمة التلفزيون عالي الوضوح

الشرط	البند	القيم
a	نسبة مسافة المشاهدة إلى ارتفاع الصورة	3
b	ذروة نصوع الشاشة (cd/m^2) (1)	250-150
c	نسبة نصوع الشاشة غير النشطة إلى ذروة النصوع (2)	≥ 0,02
d	نسبة نصوع الشاشة عند عرض مستوى أسود فقط في قاعة مظلمة تماماً إلى تلك المقابلة إلى ذروة بيضاء (3)	تقريباً 0,01
e	نسبة نصوع الخلفية وراء جهاز عرض الصورة إلى ذروة نصوع الصورة	تقريباً 0,15
f	النصوع الناتج عن المصادر الأخرى(4)	منخفض
g	ألوان الخلفية	D65
h	زاوية المشاهدة التي تقع قبالة جزء الخلفية الذي يفي بالمواصفات أعلاه (5). وينبغي الحفاظ عليها لجميع المراقبين.	53° ارتفاعاً × 83° عرضاً
i	ترتيب المراقبين	في حدود ± 30° أفقياً من وسط العرض. والحد الرأسي قيد الدراسة
j	حجم شاشة العرض(6)	1,4 m (55 بوصة)

(1) ذروة نصوع الشاشة المقابلة للإشارة الفيديوية باتساع قدرة 100%.

(2) يمكن أن يتأثر هذا البند بإضاءة القاعة، وكذلك بمدى تباين شاشة العرض.

(3) المستوى الأسود يقابل إشارة فيديوية باتساع قدره 0%.

(4) ينبغي تعديل إضاءة القاعة بحيث يكون من الممكن استيفاء الشرطين c وe.

(5) يوصى كحد أدنى بارتفاع 28° × وعرض 48°

(6) ينبغي استعمال قيم ≤ 76,2 سم (30 بوصة) إذا لم تكن أجهزة العرض ذات الحجم المحدد متاحة. انظر الملاحظة 3 بالجزء 1.

2-A2 طرائق التقدير

ينبغي إجراء التقديرات الشخصانية للنوعية الإجمالية لصورة التلفزيون HDTV التي يقدمها نظام بث باستعمال طريقة سلم النوعية المستمرة ثنائية الحافز (الملحق 2 بالجزء 2) وصورة ذات نوعية الاستوديو للتلفزيون HDTV كمرجع.

وينبغي إجراء تقدير لخصائص الأعطال في نظام بث التلفزيون HDTV باستعمال طريقة سلم الانحطاط ثنائي الحافز (الملحق 1 بالجزء 2) مع صورة ذات نوعية الاستوديو للتلفزيون HDTV أو صورة بث بدون انحطاط كمرجع.

وعندما يتعلق الأمر بالأداء على مدى محتوى البرنامج وشروط الإرسال المرجح مصادفتها في الممارسة العملية، ينبغي النظر في وصف خصائص الأعطال المركبة كما هو موصوف في الملحق 4 بالجزء 1.

وعند استعمال هذه الطرائق، ينبغي توخي الحذر للتمييز بين أثر نسق جهاز العرض ونسق النظام الأساسي (مثل أي تحويل لأعلى). وإذا خُلص إلى أنها قابلة للتطبيق ومناسبة، يمكن إجراء تقديرات تكميلية باستعمال أجهزة عرض مختلفة من أجل مراعاة أنساق العرض المختلفة.

وقد تشتمل بعض أنظمة بث التلفزيون HDTV على نسق التلفزيون التقليدي المدمج (التوافق الرجعي). وبالتالي، هناك حاجة إلى تقييم، من حيث نوعية الصورة، مدى ملاءمة الصور التلفزيونية التقليدية المدمجة في بث التلفزيون HDTV. وبالنسبة لهذه الأنظمة، ينبغي تطبيق شروط المشاهدة وطرائق التقدير الواردة في الملحق 1 بالجزء 3.

وينبغي تطبيق المفاهيم والإجراءات الأساسية الموصوفة في الملحق 1 بالجزء 3 على أنظمة البث الرقمي للتلفزيون HDTV التي تستعمل مخططات تخفيض معدل البتات.

3-A2 مواد الاختبار

يسرد التقرير ITU-R BT.2245 مجموعةً كبيرةً من الصور الثابتة والتتابعات المتحركة. ويفضل استعمالها بوصفها مواد الاختبار الموحدة لتقديرات نوعية صورة التلفزيون HDTV.

الملحق 3
بالجزء 3

التقدير الشخصاني لنوعية الصور الألفبائية الرقمية والصور البيانية في التلتيكس وفي خدمات مماثلة

مقدمة

هناك أنظمة تعرض الصور البيانية والألفبائية الرقمية وتنقلها عن طريق الشفرات الرقمية المناسبة. وتتسم الصور الألفبائية الرقمية والصور البيانية بخاصية محددة متميزة عن تلك الخاصة بالصور التلفزيونية التقليدية، وقد تختلف العملية العقلية المتبعة في تقديرها الشخصاني.

وتقترح هذه التوصية طرائق لإجراء التقدير الشخصاني لنوعية الصور على النحو الوارد في البرامج التلفزيونية الحالية. وهناك حاجة إلى دراسات عن نوعية الصور الألفبائية الرقمية والصور البيانية المستعملة للعديد من الخدمات الجديدة المرسلة عبر القنوات التلفزيونية والتي تستعمل شفرات رقمية لوصف الصور الألفبائية الرقمية والصور البيانية. وتؤثر بعض معلمات الإرسال على نوعية الصور المعروضة: استبانة الصفحة (عدد الصفوف في الصفحة وعدد الأحرف في الصف) في حالة التشفير الهجائي–الفسيفسائي للتلتيكس، واستبانة خلية الأحرف (عدد البكسلات والخطوط في الخلية) في حالة (تشفير مجموعة الأحرف القابلة لإعادة التعريف

ديناميكياً (DRCS) (انظر التوصية ITU-R BT.653) واستبانة الصورة في حالة صياغة الأصوات الإذاعية أو الفاكس أو التلتيكس. وعلاوةً على ذلك، ينبغي أيضاً مراعاة آثار أخطاء الإرسال التي قد تؤثر على الشفرات. وبالتالي، من الضروري قياس النوعية وتحديد العلاقات الموضوعية إلى الشخصانية لهذه المعلمات.

وقد أظهرت الدراسات أن هناك جوانب مختلفة مطلوبة لتقدير نوعية هذه الصور التي قد تكون لها خصائص مختلفة عن تلك الخاصة بالصور التلفزيونية التقليدية. وتؤثر معلمات مثل نسق البكسلات، واستبانة خلايا الأحرف، والمباعدة، والألوان، والتخطيط على سمات مختلفة للنوعية: الوضوح، والنوعية، والراحة، والانزعاج، والجهد المبذول للقراءة، والتعب والاعتبارات الجمالية. ويتم النظر في ثلاثة جوانب رئيسية هنا: شروط المشاهدة وطرائق التقدير وسياق التقدير.

ونظراً لأهمية وضع أساس التقديرات الشخصانية لنوعية الصور الألفبائية الرقمية والصور البيانية، ينبغي تقديم أكمل وصف لتشكيلات الاختبار ومواد الاختبار والمراقبين والطرائق في جميع تقارير الاختبار.

A3-1 شروط المشاهدة

يعرّف الجزء 1 شروط المشاهدة للصور التلفزيونية المقابلة لمستويات الإضاءة المنخفضة في القاعة. ومن المحتمل أن تتم مشاهدة الصور الألفبائية الرقمية والصور البيانية أيضاً في شروط الإضاءة العادية. وبالتالي، تُقترح مجموعة تكميلية من شروط المشاهدة للدراسة: إضاءة 500 لوكس، وأقصى نصوع للشاشة من 70 إلى 200 cd/m2، ونسبة تباين الشاشة من 30 إلى 50 وقيمة 1/4 لنسبة نصوع الخلفية (من جدران القاعة) إلى أقصى نصوع للشاشة. وينبغي النظر أيضاً في مسافات عرض من أربعة إلى ثمانية أضعاف ارتفاع الصورة.

A3-2 طرائق التقدير

أُجري عدد كبير من الدراسات عن الجوانب المطبعية. واستُعملت معظم الدراسات "مقاييس الأداء" مثل عتبات الكشف أو التعرف، ونسبة التعرف، وسرعة القراءة، وما إلى ذلك. واستعمل عدد قليل جداً من الدراسات "المقاييس الشخصانية" التي تُستعمل تقليدياً في تقدير نوعية الصور التلفزيونية. ويعتبر أن أداء الأنظمة الجديدة المرسلة عبر القنوات التلفزيونية جيد (على سبيل المثال، تزيد النسبة المئوية للتعرف على الأحرف بشكل جيد عن 95%). وبالتالي يمكن استعمال مقاييس النوعية والانحطاط الواردة في هذه التوصية بكفاءة على الرغم من أن هناك حاجة إلى دراسات لتحديد الطريقة التي يمكن بها ربط هذه السلالم بالوضوح. وتمت تجربة مقارنة مع طرائق تقدير نوعية الكلام (قطاع تقييس الاتصالات) واقتُرح سلم بشأن "جهد القراءة" مكون من 5 درجات لمزيد من الدراسة.

وتقارن طريقة أخرى نتائج التقديرات الشخصانية التي أجريت باستعمال سلمين اثنين مختلفين من 5 درجات يردان في الجدول 3-4.

الجدول 3-4

سلالم الوضوح وجهد القراءة

سلم النوعية والوضوح	سلم جهد القراءة
وضوح ممتاز	لا جهد في القراءة
وضوح جيد	الانتباه مهم، ولكن لا جهد كبير في القراءة
وضوح مقبول	جهد معتدل للقراءة
وضوح ضعيف	جهد كبير للقراءة
وضوح سيء	جهد كبير جداً للقراءة

وخُلص إلى أنه من المهم جعل صيغة كل سلم درجات واضحةً للغاية. وعادةً ما تكون القيم المتوسطة للعلامات التي تم الحصول عليها باستعمال سلم جهد القراءة أعلى بشكل عام من تلك التي تم الحصول عليها بسلم الوضوح ويكون مدى العلامات التي يعطيها المراقبون أعلى في حالة سلم جهد القراءة.

واستعملت تجربة أخرى سلم النوعية الموصوف في الفقرة 1.4-A3 من الجزء 2 لتقدير آراء كل من النوعية العامة والوضوح العام للنص المكتوب المرسَل بواسطة نظام تلفزيوني بمعيار متغير للخط وعرض نطاق متغير. ولكل رأي، وُجِد نموذجان، أحدهما أكثر تعقيداً ودقةً، ولكن يستعمل كلاهما مفهوم إضافة "سلم الانحطاط" يصفان التأثيرات المجمعة للوضوح الأفقي والرأسي المحدود. كما تم قياس الوضوح من حيث نسبة الأحرف المحددة بشكل صحيح. ومع ذلك، ظل الوضوح في هذه الشروط عالياً عندما كانت النوعية منخفضةً، ومن الواضح أن المعيار السابق عادةً ما يكون أقل فائدةً.

وأجرت دراسة أخرى مقارنات بين الأداء والطرائق الشخصانية للمواد النصية المطبوعة باستعمال أحرف ثابتة العرض ومتغيرة العرض. وتبين أن الطرائق الشخصانية هي الأكثر حساسية. وتم تكرار نفس نوع الدراسة باستعمال جهاز عرض أنبوبي بأشعة الكاثود، وتطبيق هذه المرة الطرائق الشخصانية فقط. وأسفر استعمال هذه الطرائق الشخصانية عن نتائج تتعامل مع الأحجام المثلى بصرياً للمصفوفات الثابتة والمتغيرة.

3-A3 سياق التقدير

ينظر نهج جديد لتقدير الخدمة في الحالة التي يمكن فيها تحديد أنشطة المستعمل في الخدمة قيد الدراسة بدقة. ولا تُجرى التقديرات وفقاً للطريقة التقليدية المتمثلة في عرض الصور ومجرد مطالبة المشاهدين بإجراء تقديرات شخصانية قياسية. وبدلاً من ذلك، يستعمل المشاهدون الصور المقدمة كما لو كانوا يستعملون الخدمة قيد الدراسة وتُجرى جميع التقديرات في هذا السياق.

ولا تحول مضاهاة استعمال الخدمة دون استعمال التدابير الشخصانية التقليدية. غير أنها تضع سياقاً للتقييمات الشخصانية أنسب للخدمة قيد الدراسة. كما أنها تسمح أيضاً باستعمال مقاييس موضوعية لأداء المشاهد ووضع مقاييس شخصانية جديدة تكون مناسبةً بشكل خاص للخدمة والمعلمات قيد الدراسة. وأخيراً، فإنها تضع أساساً أكثر أماناً لتعميم التقديرات التي تُجرى في المختبر مقارنةً بتلك التي تتم تحت شروط الخدمة.

الملحق 4
بالجزء 3

التقدير الشخصاني لنوعية الصورة في الخدمات متعددة البرامج[5]

مقدمة

من أجل إجراء تقدير شخصاني لنوعية البرامج الفردية المضغوطة والمشفرة بمعدل بتات ثابت (CBR) في خدمة متعددة البرامج، ينبغي استعمال الإجراءات الشخصانية المفصلة في الملحقين 1 أو 2 بالجزء 3 والإجراء الموصوف في الفقرة 2-A4 من هذا الملحق.

ومن أجل إجراء تقدير شخصاني لنوعية البرامج الفردية المضغوطة والمشفرة بمعدل بتات متغير (VBR) باستعمال طرائق مثل تعدد الإرسال الإحصائي أو التشفير المشترك في خدمة متعددة البرامج، ينبغي استعمال الإجراءات الشخصانية المفصلة في الملحقين 1 أو 2 بالجزء 3 والإجراء الموصوف في الفقرة 3-A4 من هذا الملحق.

1-A4 تفاصيل التقدير العامة

– ينبغي إجراء تقديرات نوعية القنوات القائمة على المواضيع باستعمال مواد اختبار ذات محتوى مماثل وحساسية مماثلة لتلك التي تنُقل عادةً على تلك القنوات.

[5] بما في ذلك مصطلح "تعدد الإرسال الإحصائي" أو خدمات "Stat-Mux".

- من أجل تقدير النوعية الشاملة الملحوظة للبرامج التي تتباين من حيث النوعية "الآنية" على مدى فترة من الزمن، ينبغي استعمال الإجراءات الموصوفة في الفقرتين 2-A4 و3-A4.
- ينبغي تطبيق سلالم النتائج للأنظمة التي تنطوي على مراجع منخفضة النوعية، وفقاً للتعليقات الواردة في وصف طريقة DSCQS، ومواصلة دراستها للاختبار الذي يقارن الخدمات متعددة البرامج بمواد منخفضة النوعية.

2-A4 إجراءات التقدير الشخصاني لنوعية الصورة في الخدمات متعددة البرامج ذات معدل بتات ثابت

يمكن إجراء التقدير الشخصاني لنوعية الصورة لكل من برامج التلفزيون SDTV والتلفزيون HDTV بشكل مستقل باستعمال الطرائق الموصوفة في الملحق 1 (SDTV) أو الملحق 2 (HDTV) بالجزء 3. ولتقدير النوعية الأساسية للنظام، ينبغي استعمال طريقة الاختبار العامة DSCQS (الموصوفة في الملحق 2 بالجزء 2). ولتقدير البرامج التي تعاني من انحطاط في الإرسال، ينبغي استعمال طريقة الاختبار العامة DSIS (الموصوفة في الملحق 1 بالجزء 2).

3-A4 إجراءات التقدير الشخصاني لنوعية الصورة في الخدمات متعددة البرامج ذات معدل بتات متغير

من أجل إجراء تقدير شخصاني لنوعية الصورة لبرامج التلفزيون SDTV والتلفزيون HDTV، يمكن تنفيذ التشفير بمعدل بتات متغير باستعمال منهجية DSCQS. وينبغي أيضاً الانتباه إلى اختيار مواد الاختبار، نظراً لأن نوعية الصورة يمكن أن تعتمد على محتوى الصورة لجميع البرامج متعددة الإرسال.

الملحق 5
بالجزء 3

مشاهدة الخبراء لتقدير نوعية الأنظمة الخاصة بأجهزة العرض الرقمية للتصوير الرقمي على الشاشات الكبيرة[6] في المسارح

1-A5 مقدمة

استُعملت مشاهدة الخبراء كثيراً في السنوات الماضية لإجراء تحقق سريع من أداء عملية الفيديو العامة.

ويصف هذا الملحق طريقة اختبار مشاهدة الخبراء التي ستضمن اتساق النتائج التي يتم الحصول عليها في المختبرات المختلفة، وذلك باستعمال عدد محدود من خبراء التقدير.

2-A5 ما سبب استعمال طريقة جديدة تستند إلى "مشاهدة الخبراء"

تجدر الإشارة إلى المزايا الناتجة عن تطبيق المنهجية المقترحة.

أولاً، يتطلب اختبار التقدير الشخصاني الرسمي عادةً استعمال 15 مراقباً على الأقل، يتم اختيارهم على أنهم "غير خبراء"، مما يتطلب اختبارات مطولة وبحثاً مستمراً عن مراقبين جدد. وهذا العدد من المراقبين ضروري لتحقيق الحساسية اللازمة بحيث يمكن تمييز الأنظمة التي يتم اختبارها وتصنيفها بثقة، أو الحكم بثقة بأنها مماثلة.

[6] الصور الرقمية التي تعرض على شاشات كبيرة (LSDI) هي مجموعة من أنظمة الصور الرقمية التي تطبق على برامج مثل الدراما والمسرحيات والأحداث الرياضية والحفلات الموسيقية والأحداث الثقافية وما إلى ذلك من مرحلة الالتقاط وحتى العرض على شاشات كبيرة ذات نوعية عالية الاستبانة في المسارح السينمائية المجهزة بالشكل المناسب والقاعات والأماكن الأخرى.

وثانياً، عن طريق استعمال مراقبين غير خبراء، قد تفشل الاختبارات التقليدية في الكشف عن الاختلافات التي قد تصبح، مع التعرض المطول، بارزة حتى لغير الخبراء.

وثالثاً، تضع التقديرات التقليدية عادةً مقاييس للنوعية (أو اختلافات في النوعية)، ولكنها لا تحدد بشكل مباشر المصنوعات أو المظاهر المادية الأخرى التي تؤدي إلى هذه المقاييس.

وتحاول المنهجية المقترحة هنا حل جميع المشاكل الثلاث.

3-A5 تعريف الخبراء في المواضيع

لغرض هذا الملحق، "المشاهد الخبير" هو شخص يعرف المواد المستعملة لإجراء التقدير، ويعرف "ما ينبغي أن ينظر إليه" وقد يكون أو لا يكون على علم عميق بتفاصيل الخوارزمية المستعملة في معالجة مواد الفيديو المراد تقديرها. وعلى أي حال، "المشاهد الخبير" هو شخص لديه خبرة طويلة في مجال التحقق من النوعية، ومنخرط بشكل مهني في المجال المحدد الذي يتناوله الاختبار. وكمثال، عند تنظيم جلسة اختبار "مشاهدة الخبراء" بشأن مادة الصور الرقمية التي تعرض على شاشات كبيرة (LSDI)، ينبغي اختيار الخبراء في مجال الإنتاج أو الإنتاج اللاحق للأفلام أو في إنتاج محتوى فيديوي عالي النوعية (على سبيل المثال، مديري التصوير الفوتوغرافي، ومصححي الألوان، وما إلى ذلك)؛ وينبغي إجراء هذا الاختيار مع وضع في الاعتبار القدرة على إصدار أحكام شخصانية فريدة لنوعية الصورة LSDI ومصنوعات الضغط.

4-A5 اختيار المقيّمين

إن اختبار مشاهدة الخبراء هو جلسة تقدير تستند إلى آراء المقيّمين، حيث تقدم أحكام على النوعية البصرية و/أو انحطاط الرؤية.

وتتكون مجموعة الخبراء الأساسية من خمسة إلى ستة خبراء. ويجعل هذا الرقم الصغير من السهل جمع المقيّمين، والوصول إلى قرار أسرع.

ووفقاً لاحتياجات التجربة، من المقبول استعمال أكثر من مجموعة أساسية من الخبراء، مجمعة في مجموعة أكبر من الخبراء (على سبيل المثال، خبراء يعملون في مختبرات مختلفة).

ومن المسلم به أن الخبراء قد يميلون إلى تحيز علاماتهم عند اختبار التكنولوجيا الخاصة بهم، وبالتالي ينبغي تجنب إدراج الأشخاص الذين شاركوا بشكل مباشر في تطوير النظام (الأنظمة) قيد الاختبار.

وينبغي فحص جميع المقيّمين لمعرفة حدة البصر الطبيعية أو المصححة إلى الطبيعية (اختبار Snellen) ورؤية الألوان الطبيعية (اختبار Ishihara).

5-A5 مواد الاختبار

ينبغي اختيار مواد الاختبار لأخذ عينات من مدى قيم الإنتاج ومستويات الصعوبة المتوقعة في السياق الحقيقي الذي سيتم فيه استعمال النظام (الأنظمة) قيد الاختبار. وينبغي أن يفضل الاختيار المواد الأكثر صعوبةً دون أن الإفراط في ذلك بشكل لا داعٍ له. ومن الناحية المثالية، ينبغي استعمال 5-7 تتابعات اختبار.

وقد تختلف طريقة اختيار المواد أيضاً وفقاً للتطبيق الذي تم تصميم النظام قيد الاختبار من أجله.

وفي هذا الصدد، لن يُقدم في هذا القسم أي إشارة أخرى هنا بشأن قواعد اختيار مواد الاختبار، وسيُترك القرار لمصمم الاختبار وفقاً للاعتبارات أعلاه.

6-A5 شروط المشاهدة

ينبغي أن تكون شروط المشاهدة، التي سيتم وصفها بالكامل في تقرير الاختبار، متوافقةً مع الجدول 3-5 وينبغي أن تظل ثابتةً أثناء الاختبار.

الجدول 3-5

نظرة عامة على شروط المشاهدة

شروط المشاهدة	الإعداد (الإعدادات)	
	الحد الأدنى	الحد الأقصى
حجم الشاشة (m)	6	16
مسافة المشاهدة[1]	H 1,5	H 2
نصوع جهاز العرض (الشاشة الوسطى، ذروة بيضاء)	34 cd/m^2	48 cd/m^2
نصوع الشاشة (وجهاز العرض مغلق)		< 1/1 000 من نصوع جهاز العرض

[1] ينبغي استعمال أسلوب عرض "الفراشة" عندما تكون مسافة المشاهدة أقل من H 1,5. وإذا كان سيتم استعمال عرض "جنباً إلى جنب"، ينبغي أن تكون مسافة المشاهدة أقرب إلى قيمة H 2.

7-A5 المنهجية

1.7-A5 جلسات التقييم

ينبغي أن تتكون كل جلسة تقييم (تعرّف بأنها مجموعة جلسات الاختبار لمجموعة معينة من المراقبين) من مرحلتين (هما المرحلة الأولى والمرحلة الثانية).

1.1.7-A5 المرحلة الأولى

تتألف المرحلة الأولى من اختبار شخصاني رسمي يتم إجراؤه في بيئة خاضعة للرقابة (انظر الفقرة 6-A5) ويسمح بالحصول على نتائج اختبار صالحة وحساسة وقابلة للتكرار. وفي هذه الحالة، يقوم الخبراء بشكل فردي بتقدير المادة المعروضة باستعمال سلالم التصنيف الموصوفة أدناه. ولا يُسمح لأعضاء اللجنة بمناقشة ما يرونه أو التحكم في العروض. وخلال هذه المرحلة، ينبغي ألا يكون الخبراء على دراية بمخطط التشفير قيد الاختبار أو بترتيب عرض المادة قيد الاختبار. وسيتم ترتيب المواد قيد الاختبار بشكل عشوائي، لتجنب أي تحيز في التقدير.

1.1.1.7-A5 عرض المواد

تجمع طريقة العرض بين عناصر طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE) (الملحق 6 بالجزء 2) وطريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) (الملحق 2 بالجزء 2). وكمرجع، يمكن أن يطلق عليها اسم طريقة الحافز الثنائي المتآون (SDS).

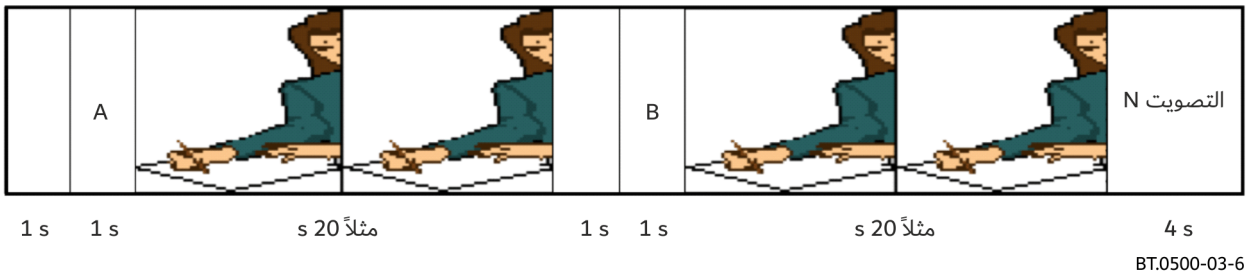
وكما هو الحال مع طريقة SDSCE، ستنطوي كل تجربة على عرض المواد بتقسيم الشاشة إلى صورتين. وفي معظم الحالات، سيكون أحد مصادر الصور هو الصورة المرجعية (أي صورة المصدر)، بينما سيكون الآخر هو صورة الاختبار؛ وفي حالات أخرى، سيتم استمداد كلتا الصورتين من الصورة المرجعية. وينبغي أن يكون المرجع مادة المصدر المعروضة بشفافية (أي لا تخضع للضغط بخلاف الضغط الضمني لوسيط تسجيل المصدر). ويجب أن تكون مادة الاختبار هي مادة المصدر التي تمت معالجتها من خلال أحد الأنظمة قيد الاختبار. وينبغي أن يكون معدل البتات و/أو مستوى النوعية على النحو المحدد في تصميم الاختبار. وعلى عكس طريقة SDSCE، لن يدرك المراقبون الشروط التي يمثلها الرقمان في زوج الصور.

وينبغي أن يتم عرض الشاشة المنقسمة إما باستعمال الشاشة المنقسمة التقليدية بدون انعكاس أو باستعمال تقنية الفراشة، حيث يتم قلب الصورة على الجانب الأيمن من الشاشة أفقياً. وبالنظر إلى أنه سيتم استعمال الصور بحجمها العرضي الكاملة، يمكن عرض نصف كل صورة فقط في المرة الواحدة. وفي كل عرض، سيتم عرض نصف الصورة نفسه على كل من جانبي شاشة العرض.

وكما هو الحال مع طريقة DSCQS، يُعرض زوج الصور مرتين على التوالي، مرة للسماح بالتعرف والتدقيق ومرة للسماح بالتأكيد والتصنيف. وستكون مدة كل تتابع 15-30 ثانية. ويمكن تسمية كل تتابع في بداية كل مقطع لمساعدة المقيّمين (انظر مثال الشاشة المنقسمة غير المنعكسة الموصوفة في الشكل 3-6).

الشكل 3-6

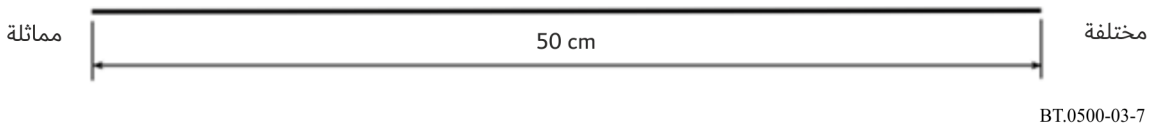
مثال على الشاشة المنقسمة غير المنعكسة



2.1.1.7-A5 سلم الحكم على الصورة

معيار القبول في تطبيقات LSDI هو أنه لا يمكن تمييز صورة الاختبار (أي المضغوطة) عن الصورة المرجعية. ويمكن استعمال العديد من طرائق تسجيل الدرجات شائعة الاستعمال لتقييم الأنظمة قيد الاختبار. وإحدى الطرائق المقترحة هي سلالم مقارنة الحافز الموصى بها (الملحق 4 بالجزء 2). ومن الأمثلة الخاصة على سلالم القياس هو سلم القياس بدون تصنيف (مستمر) SAME-DIFFERENT كما هو موصوف في الفقرة 2.4-A4 من الملحق 4 بالجزء 2:

الشكل 3-7



3.1.1.7-A5 جلسة الحكم على الصورة

ينبغي أن تتضمن الجلسة، التي قد تنطوي على أكثر من مقابلة واحدة حسب عدد شروط الاختبار، نوعين من التجارب: تجارب الاختبار وتجارب التحقق. ففي تجربة الاختبار، يعرض نصف الشاشة الصورة المرجعية بينما يُظهر النصف الآخر صورة الاختبار. وفي تجربة التحقق، يعرض نصفا الشاشة الصورة المرجعية. والغرض من تجربة التحقق هو تقدير مقياس لتحيز الحكم بشأن الصورة.

ولكل نظام يجري اختباره، يلزم إجراء تجارب الاختبار التالية لكل تتابع اختبار:

الجدول 3-6

لوحة العرض اليسرى	لوحة العرض اليمنى
الصورة المرجعية للنصف الأيسر	صورة اختبار النصف الأيسر
الصورة المرجعية للنصف الأيمن	صورة اختبار النصف الأيمن
صورة اختبار النصف الأيسر	الصورة المرجعية للنصف الأيسر
صورة اختبار النصف الأيمن	الصورة المرجعية للنصف الأيمن

ويفضل أن يكون هناك تكراران على الأقل لكل حالة من الحالات المذكورة أعلاه. وبالنسبة لكل نظام، يلزم إجراء تجارب التحقق التالية لكل تتابع اختبار:

الجدول 3-7

اللوحة اليسرى	اللوحة اليمنى
الصورة المرجعية للنصف الأيسر	الصورة المرجعية للنصف الأيسر
الصورة المرجعية للنصف الأيمن	الصورة المرجعية للنصف الأيمن

ومرة أخرى، يفضل أن يكون هناك تكراران على الأقل لكل حالة من الحالات المذكورة أعلاه.

وينبغي تقسيم جلسة الاختبار إلى مقابلات لا تزيد مدتها عن ساعة واحدة وتفصل بينها فترات راحة مدتها 15 دقيقة. وينبغي توزيع تجارب الاختبار والتحقق الناتجة عن توليفة الكودكات وتتابع الاختبار عبر المقابلات عن طريق التعيين العشوائي. ومن المفيد فرض بعض القيود على هذه العملية، وإن كانت ستصبح أكثر تعقيداً. وعلى سبيل المثال، إذا كانت هناك أربع مقابلات، يمكن تعيين بشكل عشوائي لكل تجربة من تجارب الاختبار الأربعة لكودك معين وتتابع اختبار معين موضع محدد عشوائياً في أحد المقابلات. وتتمثل فائدة هذا النهج في التأكد من توزيع تجارب الاختبار لكل نظام على مدار جلسة الاختبار بأكملها.

4.1.1.7-A5 معالجة علامات الاختبار

بالنسبة لتجربة اختبار معينة، تكون علامة الاختبار هي المسافة بين نقطة النهاية "SAME" لسلم القياس والعلامة التي وضعها المراقب، معبراً عنها على سلم من 0 إلى 100. وسيتم تحليل النتائج من حيث متوسط علامات الرأي (MOS)، وسيتم استعمال MOS لتحديد ترتيب الأنظمة التي تم اختبارها. وحسب عدد المشاهدات لكل نظام (المراقبون × تتابعات الاختبار × التكرارات)، يمكن أن تخضع البيانات لتحليل التباين (ANOVA)[7]. ويمكن استعمال الأداء في تجارب التحقق للحصول على فرق الحكم الأساسي على الصورة الذي يكون "صدفة".

2.1.7-A5 المرحلة الثانية

يتمثل أحد الأهداف الرئيسية للمرحلة الثانية في تحسين الترتيب النسبي لنتائج المرحلة الأولى، والتي يمكن أن تنخفض دقتها وموثوقيتها بسبب قلة عدد المراقبين و/أو تجارب الحكم على الصورة. وهناك هدف آخر ومهم وهو استخلاص الملاحظات بشأن الخصائص التي يُنظر إلى أن الصور تختلف على أساسها والتي استندت إليها الأحكام في المرحلة الأولى.

وينطوي هذا الجزء على استعراض من جانب لجنة الخبراء للمواد المعروضة. وهنا، يُسمح للخبراء بمناقشة المواد كما هي معروضة، وتكرار جزء أو كل المواد عدة مرات حسب الحاجة للاستعراض و/أو الإثبات، والتوصل إلى حكم توافقي ووصف لما يرونه. ويسمح باستعمال "التشغيل الخادع" (Trick Play)، بما في ذلك استعمال أوضاع مثل الحركة البطيئة والمرحلة الواحدة والرتل الثابت، إذا طلب المشاهدون الخبراء ذلك. وتتطلب هذه التقنيات بعض التفاعل مع مدير الاختبار وتدخله.

1.2.1.7-A5 تجميع المواد قيد الاختبار

لإجراء اختبار المرحلة الثانية بشكل صحيح، من الضروري تجميع المواد قيد الاختبار حسب المحتوى، والحصول على ما يطلق عليه اسم مجموعة مشاهدة الخبراء الأساسية (BES)، أي ينبغي تجميع كل التتابعات المشفرة التي تم الحصول عليها من نفس تتابع المصدر ثم ترتيبها وفقاً للترتيب المستمد من المرحلة الأولى.

وسيتم ترتيب مواد الاختبار من أدنى قيمة MOS إلى أعلى قيمة MOS. وسيكون هناك العديد من مجموعات BES مثل عدد التتابعات المستعملة للاختبار.

[7] يكفي إجراء ما مجموعه 10 مشاهدات – 20 مشاهدة في شروط أدنى ترتيب لتطبيق المعالجات الإحصائية الاستدلالية، مثل تحليل التباين.

2.2.1.7-A5 الجلسة الفرعية لاختبار مشاهدة الخبراء الأساسية

الجلسة الفرعية لاختبار مشاهدة الخبراء الأساسية (BEV) هي جلسة نقاش يفحص خلالها الخبراء جميع المواد المدرجة في مجموعة BES؛ وتتمثل إحدى المهام في تأكيد أو تعديل ترتيب التصنيف الناتج عن الاختبار الرسمي في المرحلة الأولى. ولذلك، ينبغي تأكيد أو تعديل الرؤية النسبية للاختلافات.

3.2.1.7-A5 خطة المرحلة الثانية

خلال المرحلة الثانية، يجب إجراء جميع مشاهدات BEV. وسيجري إبلاغ الخبراء بأن ترتيب العرض هو نتيجة لترتيب المرحلة الأولى. ولن يكون الخبراء على دراية بأي علاقة بين الأنظمة المؤيدة والترتيب.

وستجري المرحلة الثانية كجهد جماعي ينتج عنه آراء توافقية بين المقيّمين.

وقبل أن تبدأ المرحلة الثانية، ستُقدَّم التعليمات إلى المقيّمين، ربما باستعمال نص مكتوب، لأداء المهام التالية:

- رؤية المواد في كل مشاهدات BEV.
- مناقشة ترتيب المواد في كل مشاهدة من مشاهدات BEV؛ وفي حالة عدم موافقة المجموعة على الترتيب، يتم تحديد ترتيب جديد.
- التعليق على كل حالة مع إبداء ملاحظات تفصيلية حول طبيعة الاختلافات التي شوهدت، إن وجدت.
- توثيق ترتيبهم وتعليقاتهم وملاحظاتهم.

وسيكون من مسؤولية مدير الاختبار جمع كافة التعليقات من المجموعات والتحقق من التناقضات. وأثناء سير الاختبارات، ستبقى نتائج المرحلتين الأولى والثانية من المجموعات الفردية سريةً لتفادي التأثير على المجموعات اللاحقة. وحيثما يمكن، يُصرح لمدير الاختبار بتحديد التناقضات ودعم الاستبانة عن طريق مواصلة اختبار الترتيبات التي كان هناك خلاف بشأنها. والهدف من هذه المرحلة الأخيرة هو ضمان تحقيق توافق عام في الآراء.

8-A5 التقرير

سيكون مدير الاختبار مسؤولاً عن التقرير النهائي للاختبار.

وستُقدم المعلومات التالية في هذا التقرير:

- نتائج المرحلة الأولى (بما في ذلك جداول قيم MOS، وكذلك نتائج التحليلات الإحصائية، عند الاقتضاء).
- تعليقات الخبراء التي تم جمعها خلال المرحلة الثانية.
- تعليقات على أي إعادة تقييم للترتيب.
- جميع المعلومات ذات الصلة عن شروط المشاهدة وخصائص إشارة الدخل ومعالجة الإشارات وخصائص جهاز العرض وإعدادات جهاز العرض واللونية واختيار المشاهد وشروط الاختبار.
- توصيف كامل لأداء جهاز العرض (متوسط الوقت بين الأعطال، وما إلى ذلك).
- الملخص والاستنتاجات.

الملحق 6
بالجزء 3

التقدير الشخصاني لنوعية الصورة في التطبيقات متعددة الوسائط

1-A6 المقدمة

شرعت العديد من البلدان بنشر أنظمة الإذاعة الرقمية التي ستتيح تقديم تطبيقات إذاعة الوسائط المتعددة والبيانات المشتملة على الإشارات الفيديوية والسمعية والصور الساكنة والنصوص والرسوم.

تعتبر طرائق التقدير الشخصاني الموحدة ضروريةً لتوصيف متطلبات الأداء والتحقق من صحة الحلول التقنية قيد الاعتبار لكل تطبيق. وتعد المنهجيات الشخصانية ضروريةً لأنها توفر قياسات تسمح للصناعة بأن تتوقع ردود فعل المستعملين النهائيين بصورة مباشرة.

ويختلف النظام الإذاعي اللازم لتقديم تطبيقات متعددة الوسائط اختلافاً كبيراً عن النظام المستعمل حالياً: إذ يتم النفاذ إلى المعلومات عبر مستقبلات ثابتة و/أو متنقلة؛ وقد يكون معدل الأرتال ثابتاً أو متغيّراً؛ ويكون لحجم الصورة المحتملة مدى واسع (يتراوح بين SQCIF وHDTV)؛ وتصاحب الإشارة الفيديوية نمطياً إشارة سمعية مدمجة مع نص و/أو صوت؛ ويمكن معالجة الإشارة الفيديوية بكودكات فيديوية متقدمة؛ وتعتمد مسافة المشاهدة المفضَّلة بشكل كبير على التطبيق.

وينبغي تطبيق منهجيات التقدير الشخصاني الموصّفة في الجزء 2 في هذا السياق الجديد. فضلاً عن ذلك، يمكن تنفيذ الأبحاث الخاصة بالأنظمة متعددة الوسائط باستعمال هذه الطرائق الجديدة لتلبية متطلبات المستعمل من خصائص المجال متعدد الوسائط.

ويصف هذا الملحق التقدير الشخصاني غير التفاعلي لتقدير جودة الفيديو للتطبيقات متعددة الوسائط. ويمكن تطبيق هذه الطرائق لأغراض مختلفة تتضمن ولا تقتصر على: اختيار الخوارزميات وتصنيف أداء النظام السمعي المرئي وتقدير مستوى جودة الفيديو أثناء توصيل سمعي مرئي.

2-A6 الخصائص العامة

1.2-A6 شروط المشاهدة

ترد شروط المشاهدة الموصى بها في الجدول 3-8. وينبغي أن تكون أبعاد ونمط العرض المستعمَل مناسبةً للتطبيق قيد البحث. ونظراً لاتجاه النية لاستعمال العديد من تكنولوجيات العرض في تطبيقات الوسائط المتعددة، ينبغي إبلاغ كل المعلومات ذات الصلة بوسيلة العرض (مثل الجهة المصنعة والطراز والمواصفات) المستعملة في التقييم.

وعند استعمال أنظمة قائمة على الحاسوب الشخصي لعرض التتابعات، ينبغي كذلك إبلاغ خصائص هذه الأنظمة (مثل بطاقة العرض الفيديوية).

ويُبيِّن الجدول 3-9 مثالاً عن سجلّ البيانات لتشكيلات نظام متعدد الوسائط تحت الاختبار.

وإذا تم الحصول على صور الاختبار بواسطة توليفة معيّنة لمفكك شفرة–مشغّل، فينبغي فصل الصور عن الطبقة السطحية المسجّلة للحصول على عرض غير محدد الهوية. وهذا ضروري لضمان عدم تأثر تقييم الجودة بمعارف البيئة المصدرة.

وعندما تستعمل الأنظمة الخاضعة للتقييم في اختبار ما نسقاً مخفضاً للصورة مثل CIF أو SIF أو QCIF وما إلى ذلك، ينبغي عرض التتابعات على نافذة في شاشة العرض. وينبغي أن يكون لون الخلفية على الشاشة رمادي بنسبة 50%.

الجدول 3-8

شروط المشاهدة الموصى باستعمالها في تقييم جودة الوسائط المتعددة

المعلمة	الإعدادات
مسافة المشاهدة[1]	مقيّدة: 1-8 H غير مقيّدة: حسبما يفضّله المشاهد
ذروة نصوع الشاشة	70-250 cd/m^2
نسبة نصوع الشاشة الخاملة إلى ذروة النصوع	≥ 0,05
نسبة نصوع الشاشة عندما تعرض مستوى السواد، فقط، في قاعة مظلمة تماماً، إلى النصوع المقابل لذروة البياض	≥ 0,1
نسبة نصوع الخلفية وراء شاشة عرض الصورة، إلى نصوع ذروة الصورة[2]	≥ 0,2
لونية الخلفية[3]	D_{65}
الإضاءة الخلفية للقاعة[2]	≥ 20 lux

(1) تعتمد مسافة المشاهدة عموماً على التطبيق.

(2) تشير هذه القيمة إلى ضبط يسمح بالحد الأقصى من إمكانية كشف التشوهات. ويُسمح بقيم أعلى لبعض التطبيقات، أو تُحدد هذه القيم من قبل التطبيق.

(3) بالنسبة لشاشات الحواسيب الشخصية، ينبغي أن تقارب لونية الخلفية، قدر الإمكان، لونية "النقطة البيضاء" على شاشة العرض.

الجدول 3-9

تشكيلة النظام تعدد الوسائط الخاضع للاختبار

المعلمة	المواصفة
نمط العرض	
أبعاد العرض	
بطاقة العرض الفيديوي	
جهة التصنيع	
النموذج	
معلومات الصورة	

2.2-A6 إشارات المصدر

تقدم إشارة المصدر الصورة المرجعية مباشرةً والدخل للنظام الخاضع للاختبار. وينبغي أن تكون جودة تتابعات المصدر عاليةً قدر الإمكان. كمبدأ توجيهي، ينبغي تسجيل الإشارة الفيديوية في الملفات متعددة الوسائط باستعمال YUV (نسقي 4:2:2، 4:4:4) أو RGB (24 أو 32 بتة). وعندما يرغب القائم بالتجربة في مقارنة نتائج من مختبرات مختلفة، من الضروري استعمال مجموعة مشتركة من تتابعات المصدر تحاشياً لسبب آخر من مسببات الاختلاف.

3.2-A6 اختيار مواد الاختبار

يعتبر عدد مشاهد الاختبار وأنماطها من الأمور الحاسمة في تأويل نتائج التقدير الشخصاني. وقد تؤدي بعض العمليات إلى قدر مشابه من الانحطاط لمعظم التتابعات. وفي حالات كهذه، ينبغي للنتائج المستخلصة من عدد قليل من التتابعات (اثنين مثلاً) أن تقدم تقديراً جيداً. غير أنه كثيراً ما يكون للأنظمة الجديدة تأثير يعتمد على محتوى المشهد أو التتابع بشكل كبير. وفي حالات كهذه، ينبغي اختيار عدد مشاهد الاختبار وأنماطها بحيث توفر تعميماً معقولاً للبرنامج العادي. وعلاوةً على ذلك، ينبغي اختيار

المواد على نحو تكون فيه "حاسمة دون إفراط" بالنسبة إلى النظام الخاضع للاختبار. وتعني عبارة "دون إفراط" أن المشهد يمكن أن يظل مشكلاً جزءاً من محتوى البرنامج العادي. ويمكن للخصائص المكانية الزمانية الإدراكية للمشهد أن تقدم مؤشراً مفيداً عن تعقيد مشهد. وتُعرض قياسات الخصائص المكانية الزمانية الإدراكية بمزيد من التفصيل في الملحق 6 بالجزء 1.

4.2-A6 مدى الشروط وإرساؤها

نظراً إلى أن أغلب طرائق التقدير حساسة للتغييرات في مدى الشروط المشاهدة وتوزيعها، يجب أن تتضمن عمليات التقييم النطاقات الكاملة للعوامل المتغيّرة. غير أنه يمكن تقريب ذلك من خلال استعمال مدى أكثر تقييداً، وكذلك من خلال فرض بعض الشروط التي قد تقع عند أطراف السلالم. ويمكن تمثيلها على أنها أمثلة، وتحديدها بأنها الأكثر صرامةً (إرساء مباشر). أو توزيعها على كل العملية وعدم تحديدها بأنها الحالات الأكثر صرامةً (إرساء غير مباشر). وينبغي استعمال مدى واسع الجودة إن أمكن ذلك.

5.2-A6 المراقبون

يجب أن يكون هناك 15 مراقباً على الأقل بعد الفرز على أن يكونوا غير مختصين، أي أنهم غير معنيين مباشرةً بجودة صورة التلفزيون كجزء من عملهم العادي ولا يكونوا مقيّمين محترفين. ويجب فرز المراقبين قبل العملية وفقاً لحدة بصرهم العادية (أو المحولة إلى العادية) باستعمال مخطط اختبار "سنلين" (Snellen) أو "لندولت" (Landolt)، ووفقاً لرؤيتهم العادية للألوان بواسطة مخططات تُنتقى خصيصاً لهذا الغرض (مخطط اشيهارا (Ishihara)، مثلاً).

ويعتمد العدد المطلوب من المقيّمين على حساسية إجراء الاختبار واعتماديته على الحجم المتوقع للأثر المنشود.

يجب أن تتضمن التجارب، على قدر الإمكان، التفاصيل الخاصة بخصائص أفرقة تقييمها لتيسير إجراء مزيد من البحث لهذا العامل. ويمكن للمعلومات المقترح تقديمها أن تشمل: فئة المهنة (مثل موظف لدى هيئة إذاعية، طالب جامعي، عامل مكتبي)، والجنس والنطاق العمري.

6.2-A6 التصميم التجريبي

يُترك لمنفّذ التجربة اختيار التصميم التجريبي لتحقيق الأهداف المحددة من حيث الكلفة والدقة. ويحبذ تكرار التجربة مرتين على الأقل (أي التكرار في نفس الشروط). حيث يتيح التكرار حساب الاعتمادية الفردية، وإذا لزم الأمر، استبعاد النتائج غير الموثوقة من بعض الأشخاص. وفضلاً عن ذلك، تضمن عمليات التكرار توازن آثار التعلّم ضمن الاختبار إلى حد ما. ويمكن الحصول على تحسين آخر في معالجة آثار التعلّم، من خلال تضمين بعض "العروض الزائفة" في مستهل كل عملية اختبار. وينبغي لهذه الشروط أن تكون معبرةً عن العروض التي ستقدَّم لاحقاً أثناء العملية. ويجب ألا تؤخذ العروض التمهيدية في الحسبان عند التحليل الإحصائي لنتائج الاختبار.

ويجب ألا تستمر مدة العملية الواحدة، المؤلفة من سلسلة من العروض، لأكثر من نصف الساعة.

وينبغي استعمال ترتيب عشوائي لعرض المشاهد أو الخوارزميات عند اختبار مشاهد أو خوارزميات متعددة. وقد يُعدّل الترتيب العشوائي لضمان عدم عرض المشاهد نفسها أو الخوارزميات نفسها في تقارب زمني (أي على التعاقب).

3-A6 طرائق التقييم

يمكن فحص الأداء الفيديوي لأنظمة متعددة الوسائط باستعمال المنهجيات الوارد وصفها في الجزء 2. وتستفيد طريقة SAMVIQ من خصائص المجال متعدد الوسائط ويمكن استعمالها في تقييم أداء الأنظمة متعددة الوسائط.

الملحق 7
بالجزء 3

طرائق التقدير الشخصاني لأنظمة التلفزيون ثلاثي الأبعاد ومجسم الصورة

1-A7 أبعاد التقييم (الإدراكية)

يوظف التلفزيون ثلاثي الأبعاد مجسم الصورة خصائص نظام الإبصار البشري المزدوج عن طريق محاكاة الشروط التي تهيئ إدراك العمق النسبي للكائنات في المشهد المرئي. والمتطلب الرئيسي لتكوين الصور المجسمة هو التقاط ما لا يقل عن مشهدين لنفس المنظر من كاميرتين متحاذيتين أفقياً. ويكون لصور الكائنات الظاهرة في المنظر مواضع نسبية مختلفة في كلٍ من المشهدين الأيمن والأيسر. وهذا الاختلاف في المواضع النسبية في المشهدين هو ما يسمى تقليدياً تفاوت الصور (أو التخاطل)، ويعبر عنه عادةً بالبكسل أو بمسافات مادية (بالملليمتر مثلاً) أو بمقاييس نسبية (نسبة مئوية من عرض الشاشة مثلاً). وينبغي التفريق بين تفاوت الصور والتفاوت الزاوي (الشبكي)، بل إن من شأن نفس معلومات تفاوت الصور أن تنتج تفاوتات زاوية (شبكية) مختلفة مع اختلاف مسافة الرؤية. ويستند إدراك العمق كماً واتجاهاً إلى كميات التفاوتات الشبكية التي تستدعيها الصورة المجسمة واتجاهاتها.

ومن الممكن تطبيق عوامل التقييم التي تطبق بشكل عام على صور التلفزيون المستوية، مثل الاستبانة وتمثيل الألوان وتمثيل الحركة والنوعية الكلية والحدة وما إلى ذلك، على أنظمة التلفزيون المجسم كذلك. ويضاف إلى ذلك عوامل كثيرة خاصة بأنظمة التلفزيون المجسم الصورة تعييناً. وقد تتضمن هذه عوامل مثل استبانة العمق، وهي الاستبانة المكانية في اتجاه العمق، وحركة العمق، أي إذا كانت إعادة تكوين الحركة على اتجاه العمق تتسم بالسلاسة، والتشوهات المكانية. ومن الأمثلة المشهورة لهذا العامل الأخير ما يسمى تأثير مسرح العرائس، وهو ما يحدث عند إدراك أحجام الكائنات على أنها كبيرة أو صغيرة بشكل غير طبيعي، وتأثير الورق المقوى، أي ما يحدث عند إدراك الأشياء بشكل مجسم لكنها تظهر رفيعة بشكل غير طبيعي.

ويمكن تحديد ثلاثة أبعاد إدراكية أساسية تؤثر معاً على نوعية التجربة التي يوفرها نظام مجسم الصورة: نوعية الصورة ونوعية العمق والارتياح البصري. وقد طرح بعض الباحثين إمكانية قياس الوقع النفساني لتكنولوجيات التصوير المجسم من منطلق مفاهيم أوسع مثل الطبيعية والشعور بالتواجد.

1.1-A7 الأبعاد الإدراكية الأولية

نوعية الصورة تشير إلى النوعية المدركة للصورة التي يعرضها النظام. وهذا من المحددات الرئيسية لأداء أي نظام فيديو. وتتأثر نوعية الصورة أساساً بمعلمات تقنية وأخطاء تسببها، على سبيل المثال، عمليات التشفير و/أو الإرسال.

نوعية العمق تشير إلى قدرة النظام على تحقيق إحساس محسَّن بالعمق. ويحقق وجود بعد إيعازات الإبصار الأحادي، مثل المنظور الخطي والضبابية والتدرجات وما إلى ذلك، شيئاً من الإحساس بالعمق حتى في الصور التقليدية ثنائية الأبعاد. ومع ذلك فإن الصور ثلاثية الأبعاد المجسمة تحتوي أيضاً على معلومات تفاوت توفر معلومات عمق إضافية، وبالتالي إحساساً محسناً بالعمق مقارنةً بالصور ثنائية الأبعاد.

(عدم) الارتياح البصري يشير إلى ما قد يقترن بمشاهدة الصور المجسمة من إحساس شخصاني بالارتياح أو عدمه. ومن شأن الصور المجسمة الملتقطة بشكل غير سليم أو المعروضة بشكل غير سليم أن تكون سبباً في عدم ارتياح بالغ.

2.1-A7 أبعاد إدراكية إضافية

الطبيعية تشير إلى إدراك الصورة المجسمة وكأنها تمثيل صادق للواقع (أي الواقعية الإدراكية). وقد يدخل على الصورة المجسمة أنواع مختلفة من التشوهات تقلل من طبيعيتها. فعلى سبيل المثال، تبدو الكائنات المجسمة للرائي أحياناً وكأنها كبيرة أو صغيرة بشكل غير طبيعي (تأثير مسرح العرائس)، أو تبدو أحياناً رفيعة بشكل غير طبيعي (تأثير الورق المقوى).

الشعور بالتواجد يشير إلى الإحساس الشخصاني بالتواجد في مكان أو بيئة ما خلاف الواقع.

وتقدم هذه التوصية معلومات تتعلق بطرائق وإجراءات لتقييم الأبعاد الأولية الثلاثة المذكورة أعلاه: نوعية الصورة ونوعية العمق والارتياح البصري. ولا تتضمن هذه التوصية طرائق لتقييم الطبيعية والشعور بالتواجد، لكن من المخطط إدراجهما في مرحلة لاحقة.

2-A7 الطرائق الشخصانية

تستعرض هذه التوصية العديد من المنهجيات لتقييم نوعية الصورة. وفي جميع الطرائق، تُعرض على هيئة من المشاهدين مجموعة من تتابعات فيديو عولجت بالأنظمة محل الفحص (على سبيل المثال، خوارزمية بمعلمات مختلفة أو تكنولوجيا للتشفير بمعدلات بتات مختلفة أو سيناريوهات إرسال مختلفة وما إلى ذلك) في سلسلة من التجارب التحكيمية. وفي كل تجربة، يطلب من المشاهدين تقييم خاصية ذات صلة (مثل نوعية الصورة) لتتابع (تتابعات) الفيديو باستعمال مقياس محدد مسبقاً. وتنحصر أغلب التنوعات بين الطرائق المختلفة في أسلوب العرض، أي في كيفية عرض تتابعات الفيديو على المشاهدين، وفي المقياس الذي يستعمله المشاهدون لتصنيف تلك التتابعات.

والصور المستعملة في الاختبارات صور مجسمة بالازدواج البصري مختارة على أساس العناصر المذكورة في الفقرة 4-A7. ويقيِّم المقيِّمون العناصر الثلاثة التالية:

- نوعية الصورة: التأثير الواقع على استبانة صور ثلاثية الأبعاد مجسمة من نظام بمسير بين صور الاختبار والشاشة المستعملة لعرض الصور المطلوب تقييمها؛
- نوعية العمق: التأثير الواقع على إدراك العمق بالنسبة إلى صور ثلاثية الأبعاد مجسمة من نظام بمسير بين صور الاختبار والشاشة المستعملة لعرض الصور المطلوب تقييمها؛
- الارتياح البصري: التأثير الواقع على سهولة المشاهدة بالنسبة إلى صور ثلاثية الأبعاد مجسمة من نظام بمسير بين صور الاختبار والشاشة المستعملة لعرض الصور المطلوب تقييمها.

ويتضمن هذا الملحق مجموعةً فرعيةً من ست طرائق من هذه التوصية؛ حيث استُعملت هذه الطرائق بنجاح على مدى العقدين الماضيين لمعالجة قضايا بحثية ذات صلة كانت متعلقة بنوعية الصورة ونوعية العمق والارتياح البصري لتكنولوجيات التصوير المجسم. وهذه الطرائق هي:

- الطرائق وحيدة الحافز (SS)؛
- سلم الانحطاط ثنائي الحافز (DSIS)؛
- طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS)؛
- طريقة مقارنة الحوافز (SC)؛
- طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE)؛
- طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE).

وقد استُعملت الطرائق بشكل معدل تعديلاً طفيفاً عند الاقتضاء، كأن تستعمل مقاييس مختلفة للارتياح البصري على سبيل المثال. وتضم الجداول 3-10 و3-11 و3-12 تلخيصاً لأسلوب العرض والمقاييس المقترنة بطريقة تقييم نوعية الصورة ونوعية العمق والارتياح البصري على الترتيب.

ويرد فيما يلي من هذا القسم وصف موجز لكل منهجية. بينما تعرض الأقسام التالية العناصر المنهجية المشتركة بين جميع الطرائق.

الجدول 3-10

الطريقة الشخصانية لتقدير نوعية الصورة

أسلوب العرض	مدة التتابع	السلم الثنائي	السلم المتدرج	السلم المتواصل
الطرائق وحيدة الحافز (SS) على النحو الموصوف في الفقرة 1.6 من الملحق 1.	~10 s		5 ممتازة 4 جيدة 3 مقبولة 2 متدنية 1 سيئة	مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة سلم الانحطاط ثنائي الحافز (DSIS) على النحو الموصوف في الفقرة 4 من الملحق 1.			5 لا يمكن إدراكها 4 يمكن إدراكها، ولكنها مزعجة 3 مزعجة قليلاً 2 مزعجة 1 مزعجة للغاية	
طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) على النحو الموصوف في الفقرة 5 من الملحق 1.	~10 s			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طرائق مقارنة الحافز (SC) على النحو الموصوف في الفقرة 2.6 من الملحق 1.	~10 s	A مقابل B	3– أسوأ بكثير 2– أسوأ 1– أسوأ قليلاً 0 تطابق 1 أفضل قليلاً 2 أفضل 3 أفضل بكثير	
طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE) على النحو الموصوف في الفقرة 3.6 من الملحق 1.	~3–5 دقيقة			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE) على النحو الموصوف في الفقرة 4.6 من الملحق 1.				دقة ممتازة (تشفير 100) دقة معدومة (تشفير 0)

الجدول 3-11

الطريقة الشخصانية لتقدير نوعية العمق

أسلوب العرض	مدة التتابع	السلم الثنائي	السلم المتدرج	السلم المتواصل
الطرائق وحيدة الحافز (SS) على النحو الموصوف في الفقرة 1.6 من الملحق 1.	~10 s		5 ممتازة 4 جيدة 3 مقبولة 2 متدنية 1 سيئة	مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة سلم الانحطاط ثنائي الحافز (DSIS) على النحو الموصوف في الفقرة 4 من الملحق 1.			5 لا يمكن إدراكها 4 يمكن إدراكها، ولكنها مزعجة 3 مزعجة قليلاً 2 مزعجة 1 مزعجة للغاية	
طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) على النحو الموصوف في الفقرة 5 من الملحق 1.	~10 s			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طرائق مقارنة الحافز (SC) على النحو الموصوف في الفقرة 2.6 من الملحق 1.	~10 s	A مقابل B	3– أسوأ بكثير 2– أسوأ 1– أسوأ قليلاً 0 تطابق 1 أفضل قليلاً 2 أفضل 3 أفضل بكثير	
طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE) على النحو الموصوف في الفقرة 3.6 من الملحق 1.	~3–5 دقائق			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE) على النحو الموصوف في الفقرة 4.6 من الملحق 1.				دقة ممتازة (تشفير 100) دقة معدومة (تشفير 0)

الجدول 3-12

الطريقة الشخصانية لتقدير الارتياح البصري

أسلوب العرض	مدة التتابع	السلم الثنائي	السلم المتدرج	السلم المتواصل
الطرائق وحيدة الحافز (SS) على النحو الموصوف في الفقرة 1.6 من الملحق 1.	~10 s		5 مريحة جداً 4 مريحة 3 غير مريحة قليلاً 2 غير مريحة 1 غير مريحة للغاية	مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة سلم الانحطاط ثنائي الحافز (DSIS) على النحو الموصوف في الفقرة 4 من الملحق 1.			5 لا يمكن إدراكها 4 يمكن إدراكها، ولكنها مزعجة 3 مزعجة قليلاً 2 مزعجة 1 مزعجة للغاية	
طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) على النحو الموصوف في الفقرة 5 من الملحق 1.	~10 s			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طرائق مقارنة الحافز (SC) على النحو الموصوف في الفقرة 2.6 من الملحق 1.	~10 s	A مقابل B	3– أسوأ بكثير 2– أسوأ 1– أسوأ قليلاً 0 تطابق 1 أفضل قليلاً 2 أفضل 3 أفضل بكثير	
طريقة تقييم النوعية المستمرة وحيدة الحافز (SSCQE) على النحو الموصوف في الفقرة 3.6 من الملحق 1.	~3–5 دقائق			مريحة جدا مريحة غير مريحة قليلا غير مريحة غير مريحة للغاية
طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE) على النحو الموصوف في الفقرة 4.6 من الملحق 1.				دقة ممتازة (تشفير 100) دقة معدومة (تشفير 0)

3-A7 شروط المشاهدة العامة

ينبغي أن تكون شروط المشاهدة (بما في ذلك نصوع الشاشة والتباين وإضاءة الخلفية ومسافة المشاهدة وما إلى ذلك) متسقةً مع تلك المستعملة للصور ثنائية الأبعاد كما هو موصوف في الفقرة 1.2 من الجزء 1. ولاتباع نهج الاتساق هذا أسباب منطقية من شقين، أولهما أن المستعملين في التطبيق العملي سيشاهدون التلفزيون ثلاثي الأبعاد بنفس أجهزة العرض المستعملة للتلفزيون ثنائي الأبعاد وتحت نفس شروط المشاهدة، والثاني أن الحاجة ستنشأ كثيراً إلى قياس التقدم المحرز في تكنولوجيات الفيديو ثلاثي الأبعاد نسبةً إلى التقدم المحرز في تكنولوجيات فيديو التلفزيون عالي الوضوح العادي (أي "مقارنةً بها").

وتحدد الفقرة 1.2 من الجزء 1 معيارين محتملين لاختيار مسافة المشاهدة. والمختار منهما هو "مسافة المشاهدة التصميمية" (DVD)، وتعريفها بالنسبة إلى الأنظمة الرقمية هو المسافة التي يقابل عندها بكسلان متجاوران زاوية قياسها دقيقة قوسية واحدة عند عين المشاهد.

وينبغي ملاحظة أن تقابل بكسلين متجاورين لزاوية قياسها دقيقة قوسية واحدة عند عين المشاهد يعني أن أصغر تفاوت زاوي (شبكي) يمكن للنظام تمثيله عند مسافة المشاهدة التصميمية (أي استبانة عمق النظام) يساوي دقيقة قوسية واحدة (أو ما يعادل 60 ثانية قوسية). وقد أظهرت الأبحاث أن ما يقرب من 97% من الناس يستطيعون تمييز تفاوتات أفقية تبلغ 140 ثانية قوسية أو أقل، وأن 80% منهم على الأقل يستطيعون استشعار تفاوتات أفقية تبلغ 30 ثانية قوسية. وعلى ذلك فيُفترض أنه يسهل على معظم المشاهدين استيعاب أصغر تفاوت قابل للتمثيل على أنظمة الفيديو ثلاثي الأبعاد الحالية على مسافة المشاهدة التصميمية.

4-A7 مواد الاختبار

ينبغي توجيه مواد الاختبار حسب المسألة التجريبية التي تعالجها الدراسة. وبشكل عام، ينبغي أن يعبر محتوى تتابعات الاختبار (رياضة، دراما، فيلم، وما إلى ذلك) وخصائصها الزمانية–المكانية عن البرامج التي تبثها الخدمة الخاضعة للدراسة.

وعلاوةً على ذلك، ينبغي أن يمثل محتوى تتابعات الاختبار المجسمة المنتقاة مادةً مريحةً في المشاهدة بشكل طبيعي. ويعتمد الارتياح البصري للصور المجسمة أقصى اعتماد على تفاوتات الصورة (التخاطل) التي تنطوي الصورة عليها وعلى شروط المشاهدة. وعلى ذلك، ينبغي العمل بعناية على ضمان عدم تجاوز التفاوتات الحدود المنصوص عليها في القسم التالي، ما لم تستهدف الدراسة تعييناً قياس الارتياح البصري. وعلاوةً على ذلك، ينبغي كلما أمكن قياس العوامل الإحصائية من متوسط وانحراف قياسي ومدى (أدنى/أقصى) لتوزيع التفاوت في تتابعات الاختبار وإدراجها في التقارير.

ويمكن طرح التخاطل والتراوح بين الصورتين اليمنى واليسرى توزيع التخاطل وتغيره كعناصر ينبغي أخذها في الاعتبار عند اختيار صور اختبار بحيث تكون صوراً ثلاثية الأبعاد مجسمة سهلة المشاهدة. وتتناول الأقسام الفرعية التالية بالوصف العلاقة بين الصور ثلاثية الأبعاد المجسمة سهلة المشاهدة والتخاطل والتراوح بين الصورتين اليمنى واليسرى وتوزيع التخاطل وتغيره.

1.4-A7 استعمال مواد فيديو مرجعية

يجدر بالباحثين إدخال التتابع المرجعي، إن كان متاحاً، ضمن مجموعة تتابعات الاختبار. وتكون المادة المرجعية في العادة نسخةً من تتابع الاختبار لم تخضع لأي معالجة (أي تتابع المصدر الأصلي). وبالنسبة إلى دراسات الصور المجسمة، تكون المادة المرجعية الأساسية هي التتابع المجسم الأصلي غير المعالَج، غير أنه يجوز كذلك تضمين النسخة المستوية ثنائية الأبعاد من المادة المرجعية (أي مشهد واحد فقط لتتابع المصدر الأصلي) في خطة التجربة؛ فقد يكون من المفيد على سبيل المثال في دراسات الارتياح البصري استعمال المادة المرجعية المستوية لتكون خط الأساس. وينبغي عرض النسخة المسطحة من المادة المرجعية في وضع الأبعاد الثلاثية (كأن يُعرض المشهد الأيسر لكلا العينين اليمنى واليسرى باستعمال نفس إعدادات معدات الأبعاد الثلاثية المستعملة في التتابع المجسم الفعلي). ويتيح تضمين المادة المرجعية في خطة التجربة ميزتين مهمتين، أولهما أن ذلك يهيئ فرصة لقياس الشفافية (المعروفة

أيضاً باسم الدقة) التي توفرها الخوارزمية أو التكنولوجيا محل الفحص[8]. وأما الميزة الثانية لتضمين المادة المرجعية فهي أنها تهيئ مرتكزاً عالي الجودة من شأنه الإعانة على تنظيم التصنيفات[9].

2.4-A7 حدود الارتياح البصري

تسبب زيادة التفاوت/التخاطل فقداً للارتياح البصري لأنها تزيد من التضارب بين الاستيعاب والتجانح. ولذلك، يُقترح في سبيل تقليل التضارب بين الاستيعاب والتجانح أن تكون التفاوتات في الصورة المجسمة من الصغر بما يكفي لوقوع الأعماق المدركة للكائنات ضمن "منطقة ارتياح". وقد اقتُرحت عدة نُهج لتعريف هذه الحدود. ومن هذه النهج المقترحة استعمال قياس لتخاطل الشاشة يُعبَّر عنه كنسبة مئوية من حجم الشاشة الأفقي، وذلك لتعيين حدود المشاهدة المريحة. وقد طُرحت لذلك القيم 1% بالنسبة إلى التفاوتات المتقاطعة/السلبية و2% بالنسبة إلى التفاوتات غير المتقاطعة/الإيجابية (بقيمة إجمالية حوالي 3%). ووفق نهج آخر، تُحدَّد منطقة الارتياح حسب عمق مجال العين. وقد افترض الباحثون لشروط مشاهدة بث تلفزيوني تقليدي عمق مجال يتراوح بين ±D2,0 (ديوبتر) و±D0,3 (ديوبتر). وتناظر هاتان القيمتان بالنسبة إلى نظام استبانة صور 1920 × 1080 في أنظمة التلفزيون عالي الوضوح (التوصية ITU-R BT.709) يُشاهَد من مسافة المشاهدة التصميمية التي تبلغ 3,1 ضعفاً من الارتفاع ما يقرب من ±2% و±3% من تخاطل الشاشة. وأخيراً، يعبر نهج ثالث عن حدود الارتياح باستعمال التفاوت الشبكي، حيث يضع حديه عند ±1° من زاوية الإبصار بالنسبة إلى التفاوتات الإيجابية والسلبية.

وجدير بالذكر أن نواتج هذه النُهج المختلفة تميل إلى التقارب فيما بينها لتحدد نفس حدود الارتياح. ونذكِّر بأن بكسلين متجاورين يقابلان عند مسافة المشاهدة التصميمية زاوية قياسها دقيقة قوسية واحدة عند عين المشاهد. وعلى ذلك، يناظر 60 بكسلاً درجة واحدة من زاوية الإبصار. وهذا يتيح لنا التعبير بسهولة عن حدود الارتياح باستعمال التفاوت الشبكي (بالنسبة إلى مشاهد متوسط). فعلى سبيل المثال، تناظر 1% (حوالي 19,2 بكسلاً) في أنظمة استبانة صور 1920 × 1080 في أنظمة التلفزيون عالي الوضوح (التوصية ITU-R BT.709) حوالي 20 دقيقة قوسية، وتناظر 2% حوالي 40 دقيقة قوسية، وتناظر 3% حوالي 60 دقيقة قوسية (أو ما يعادل درجة واحدة).

وجدير بالملاحظة أنه بالرغم من تقابل بكسلين متجاورين دوماً لزاوية قياسها دقيقة قوسية واحدة عند مسافة المشاهدة التصميمية، فإن الفاصل المادي (أي ما يقاس بالمليمتر) بين هذين البكسلين يزداد كلما كبر حجم جهاز العرض (يبقى عدد البكسلات كما هو لكن الحجم المادي للشاشة يزداد). وعلى ذلك فمن شأن الحدود الأعلى (±3% مثلاً) أن تؤدي في أجهزة العرض الأكبر إلى تجاوز المسافة المادية بين النقاط المناظرة (أي تخاطل المشهدين بالمليمتر) المسافة بين الحدقتين لدى المشاهد المتوسط (حوالي 63-65 mm)، وهذا قد يسبب عدم ارتياح.

3.4-A7 الفروق بين الصورتين اليمنى واليسرى

تتكون الصورة ثلاثية الأبعاد المزدوجة بصرياً في الأنظمة ثلاثية الأبعاد المجسمة عن طريق عرض كلٍ من الصورة اليمنى واليسرى للعين المناظرة لجهتها. ومن شأن أي فروق تنشأ بين هاتين الصورتين أن تسبب إجهاداً نفسياً، بل وقد تفشل المشاهدة ثلاثية الأبعاد كليةً في بعض الحالات. فعلى سبيل المثال، قد تحدث عند تصوير أفلام تلفزيونية ثلاثية الأبعاد مجسمة بعض التشوهات الهندسية مثل عدم اتساق الأحجام والإزاحة الرأسية وأخطاء الدوران بين الصورتين اليمنى واليسرى. ويفضل أن تخلو صور الاختبار من هذه التشوهات الهندسية. انظر الفقرة 1.2.3 من الملحق 4 بالتقرير ITU-R BT.2160-2 للحصول على مزيدٍ من المعلومات.

---

[8] الشفافية (الدقة) مفهوم يصف أداء كودك أو نظام مقارنةً بنظام إرسال مثالي دون أي انحطاط. ويسهل تبين إمكانية قياس الشفافية بالمقارنة بين التصنيفات المعطاة للتتابع المرجعي وتلك المعطاة للتتابع المعالَج بالخوارزمية أو التكنولوجيا محل الفحص.

[9] من المسلم به أنه يمكن أيضاً رفع مستوى استقرار التصنيفات مكانياً (أي بين مختلف المختبرات) وزمانياً (أي في نفس المختبر عند أوقات مختلفة) باستخدام مرتكزات منخفضة الجودة. غير أن الاتحاد ليس لديه خطط فورية لوضع/تعريف مرتكزات منخفضة الجودة لتقييم تكنولوجيات التصوير المجسم.

والعناصر المتعلقة بالفروق بين الصورتين اليمنى واليسرى التي ينبغي أخذها في الاعتبار عند اختيار صور اختبار بحيث تكون صوراً ثلاثية الأبعاد مجسمة سهلة المشاهدة هي كالتالي:

– الفروق الهندسية، بما فيها الحجم والإزاحة الرأسية والدوران؛

– فروق السطوع، بما فيها مستوى الأبيض والأسود؛

– التداخل.

4.4-A7 المدى والتوزيع والتغير في التخاطل

ترتبط توزيعات التخاطل بعلاقة بينية مع الارتياح البصري للصور المجسمة.

ويكون توزيع التخاطل في الصور المجسمة متقطعاً خلال أرتال تغير المنظر. وتسبب حالات التغير الأقصى في التخاطل أو التغيرات المفاجئة في التخاطل عدم ارتياح بصري، لذا فمن المهم إدارة تخاطل صور الاختبار بعناية. انظر الفقرة 2.2.3 من الملحق 4 بالتقرير ITU-R BT.2160-2 للحصول على مزيدٍ من المعلومات.

وبشكل عام، بما أن من شأن الدراسات التي تستعمل تتابعات اختبار مجسمة أن تسبب درجة ما من عدم الارتياح البصري، فيوصى باستعمال مواد اختبار لا تتجاوز التفاوتات فيها حدود الارتياح متى ما أمكن ذلك، مع السماح في حالات متباعدة بالحياد اليسير عن تلك الحدود.

5-A7 معدات التجارب

ينبغي أن تسمح إمكانيات معدات التجارب (مخدم الفيديو وجهاز العرض وغير ذلك) بعرض تتابعات اختبار عالية الوضوح باستبانة كاملة، وذلك باستعمال نسق ترزيم أرتال HDMI على سبيل المثال. فمن شأن هذا أن يتيح قدراً أكبر من المرونة في مدى الدراسات التي يمكن إجراؤها.

ولم يستقر الأمر حتى الآن على جهاز عرض مرجعي موحد لتقييم التلفزيون ثلاثي الأبعاد. وعلى ذلك فمن المتوقع أن يستعمل معظم الباحثين أجهزة عرض التلفزيون ثلاثي الأبعاد المتاحة على مستويات المستهلكين الحالية. ونظراً للتنوع في خصائص أجهزة العرض هذه حسب الصانع، فيوصى بشدة أن يدرج الباحثون في تقاريرهم معلومات الإعدادات ذات الصلة لجهاز العرض المستعمل في الدراسة.

6-A7 المراقبون

1.6-A7 حجم العينة

يوصى بشكل عام باستعمال ما لا يقل عن 30 مشاهداً. ومع ذلك، فمن المسلم به أن الرقم الفعلي يعتمد على الأهداف المحددة للدراسة المعنية، علماً بأن اعتبارات حجم العينة لدراسات الأبعاد الثلاثية لا تختلف عن اعتبارات دراسات الأبعاد الثنائية.

2.6-A7 الفرز البصري

ينبغي أن يخضع المراقبون للتقييم من حيث حدة الإبصار وعمى الألوان والإبصار المجسم باستعمال اختبارات الإبصار الإكلينيكية الحالية، مثل مخططات سنيلين أو ما يكافئها لحدة الإبصار؛ ولوحات إيشيهارا للألوان أو ما يكافئها؛ واختبار النقاط العشوائية أو ما يكافئه للإبصار المجسم. وجدير بالملاحظة أن اختبارات الإبصار المجسم مثل اختبارات النقاط العشوائية والذبابة المجسمة وفريزبي تقيس في العادة تفاوتات شبكية من حوالي 20 إلى 400 ثانية قوسية. ويوصى بأن يدرج الباحثون في تقاريرهم الإحصائيات ذات الصلة حول قدرات الإبصار المجسم لدى المراقبين المشاركين في الدراسة. وإذا اقتضت الحاجة تحليلاً أكثر تفصيلاً لقدرات الإبصار المجسم لدى المشاركين، فللباحثين استعمال مواد الاختبار المذكورة في المرفق 1 بهذا الملحق.

7-A7 تعليمات للمراقبين

ينبغي أن تفصَّل التعليمات حسب الأبعاد الخاضعة للفحص (مثل نوعية العمق والارتياح وما إلى ذلك). ويلاحَظ أن المبادئ التوجيهية الأخلاقية في حالة دراسات الأبعاد الثلاثية أشد صرامةً من تلك المستعملة عادةً في تقييم نوعية الصور ثنائية الأبعاد نظراً لاحتمال تعرض المشاركين لحالات عدم ارتياح بصري. وبشكل عام، تتطلب دراسات الأبعاد الثلاثية مزيداً من العناية في إعلام المشاركين بدوافع الدراسة علاوةً على أي أثر سلبي محتمل ينتج عن التعرض لعوامل التحفيز المستعملة في الدراسة.

8-A7 مدة الجلسة

إذا اعتبرت المادة المعروضة مريحةً فيمكن أن تمتد جلسة الاختبار نفس المدة المستعملة لدراسات الأبعاد الثنائية (أي حوالي 20-40 دقيقة تتخللها فترات راحة). أما إن وجد أن المادة تحتوي على تخاطل مفرط، مما يعني احتمال تسببها في عدم ارتياح، فينبغي تقصير المدة.

9-A7 تنوع الاستجابات

تتسم التصنيفات التي يقررها المشاهدون في تجارب التقدير الشخصاني بقدر كبير من التنوع بشكل عام. وقد تعكس الاختلافات بين المشاهدين خصائص المجموعة المرجعية ليس إلا، مما يسهل معالجته بزيادة حجم العينة.

ومع ذلك فقد ينشأ جانب من التنوع من تغيرات في أنماط الاستجابة لآحاد المشاهدين خلال التجربة. وتقتضي هذه التغيرات تغيير معايير التقييم، مما قد يحدث بسبب زيادة التمرس في المهمة (حفظ خصائص التدخلات وما إلى ذلك). وينبغي للباحثين في سبيل تقليل الآثار السلبية لهذا التنوع أن يوفروا إجراءات تدريبية وافية (المهمة ومستوى الانخطاط وما إلى ذلك) ويستعملوا توزيعات عشوائية متعددة (بمعنى عرض تتابعات الاختبار بترتيبات عشوائية مختلفة لمختلف المشاهدين) ويستعملوا التكرار (مما قد يتيح كذلك قياس التغير المحتمل في أنماط الاستجابة).

10-A7 معايير رفض المشاهدين

يصف الجزء 1 في الفقرة 2-A7 معايير رفض المشاهدين (فرز المراقبين) للطرائق المعروضة.

11-A7 التحليل الإحصائي

التحليلات الإحصائية لفحص أنظمة التصوير ثلاثي الأبعاد هي نفسها المستعملة لأنظمة التصوير ثنائي الأبعاد.

المرفق 1
بالملحق 7

مواد اختبار الإبصار

1-A7 اختبار الإبصار

يورد الجدول 3-13 مخططات اختبارات الإبصار. وهذه الاختبارات الاثنا عشر مختارة وفق تدرج نظام الإبصار البشري من مستويات أدنى إلى أعلى. ويرد أدناه وصف لثمان اختبارات إبصار رئيسية، والأربعة الأخرى للاختبار الإكلينيكي. ويجب أن يكون الإبصار التجسيمي لدى المراقب سليماً، مما يقتضي اجتيازه اختبار الإبصار الرابع للإبصار التجسيمي الدقيق واختبار الإبصار السابع

للإبصار التجسيمي الديناميي. وأما الاختبارات الستة الباقية فهي موجهة إلى تحديد الخصائص بمزيدٍ من التفصيل. وينبغي النظر إلى مخططات الاختبار من مسافة تساوي ثلاثة أضعاف ارتفاع شاشة العرض.

الجدول 3-13

مواد الاختبارات المجسمة لاختبار الإبصار

الرقم	العنصر	محل الاختبار	المحتوى
1	الإدراك المتزامن	القدرة على إدراك صور معروضة بازدواج المجال بشكل متزامن وفي الموضع الصحيح	تعرض صورة قفص لعين وصورة أسد للعين الأخرى
2	دمج الازدواج البصري	القدرة على إدراك صورتين مزدوجتي المجال في العينين اليمنى واليسرى على أنهما صورة واحدة	تحتوي الصورة الموجهة إلى إحدى العينين على نقطتين، بينما تحتوي الصورة الموجهة إلى العين الأخرى على ثلاث نقاط، مع اشتراكهما في إحدى هذه النقاط
3	الإبصار التجسيمي الثخين	القدرة على إدراك صور معروضة بازدواج المجال بتخاطل كصورة واحدة بعمق ثخين	الصورة للعينين عبارة عن زوج تجسيمي من الصور ليعسوب جناحاه ممتدان
4	الإبصار التجسيمي الدقيق	القدرة على إدراك صور معروضة بازدواج المجال بتخاطل كصورة واحدة بعمق دقيق	تعرض تسع رقع اختبار معينة الشكل في كل منها أربع دوائر إحداها بها تخاطل ضئيل
5	حد الدمج المقطوع	القدرة على إدراك صور معروضة بازدواج المجال بتفاوتات مقطوعة كصورة واحدة	يعرض زوج تجسيمي من شريطين مع تغيير تخاطله المقطوع بمقدار 10'/ث
6	حد الدمج غير المقطوع	القدرة على إدراك صور معروضة بازدواج المجال بتفاوتات غير مقطوعة كصورة واحدة	يعرض زوج تجسيمي من شريطين مع تغيير تخاطله غير المقطوع بمقدار 11'/ث
7	الإبصار التجسيمي الديناميي	القدرة على إدراك العمق في صور مجسمة بنقاط عشوائية متحركة	صورة مجسمة بنقاط عشوائية دينامية
8	حدة الازدواج البصري	حدة الازدواج البصري، بما في ذلك أي خلل في توازن حدة الإبصار الأحادي قد يحول دون الإبصار التجسيمي الجيد	أحرف E متنوعة الاتجاه والحجم
9	الحول الأفقي	انحراف أفقي في العين لا يستطيع المريض التغلب عليه	خطوط رأسية وأفقية
10	الحول الرأسي	انحراف رأسي في العين لا يستطيع المريض التغلب عليه	خطوط رأسية وأفقية
11	تفاوت الصورتين	حالة تسبب اختلاف في صورة كائن المبصرة بإحدى العينين عن الأخرى حجماً وشكلاً	تتكون الصورة اليسرى من الحرفين "o[" بينما تتكون اليمنى من الحرفين "o]" مع اشتراكهما في موضع حرف "o"
12	الاحولال التدويري	انحراف إحدى العينين حول المحور الأمامي الخلفي عند منع الدمج	تتكون الصورة اليسرى من وجه ساعة بينما تتكون اليمنى من عقربي ساعة في موضع الساعة السادسة

الملاحظة 1 – هذه المواد موضوعة في نسق 1125/60/I (انظر التوصية ITU-R BT.709).

الملاحظة 2 – يمكن الحصول على المواد من معهد معلومات الصور ومهندسي التلفزيون باليابان (ITE)، 3-5-8 Shibakoen، Minato-ku، Tokyo 105-0011، Japan، هاتف: 4675-3432-3-81، بريد إلكتروني: ite@ite.or.jp.

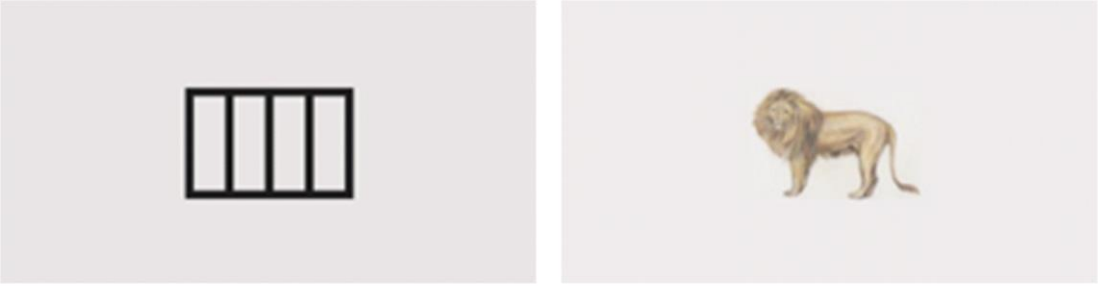
توضع صور مصغرة متجاورة في الأسفل وعلى اليمين واليسار للدمج الحر المقطوع لأغراض توضيحية.

1) اختبار الإبصار–01: الإدراك المتزامن (اختبار الأسد)

يختبر القدرة على إدراك صور معروضة بازدواج المجال بشكل متزامن وفي الموضع الصحيح. تعرض صورة قفص لعين وصورة أسد للعين الأخرى، مع تحرك موضعها بمقدار 12'/ث. ويكون حجم كل صورة مثبتاً على 10° حتى يتمكن المراقبون من التقاط الصور على الشبكية. ويستطيع المراقبون ذوو الإبصار العادي رؤية الأسد داخل القفص في وقت معين ضمن مدة العرض.

الشكل 3-8

مخطط اختبار الإبصار–01



BT.0500-03-8

2) اختبار الإبصار–02: دمج الازدواج البصري (اختبار صافي النقاط الأربع)

يختبر القدرة على إدراك صورتين مزدوجتي المجال في العينين اليمنى واليسرى كصورة واحدة. تحتوي الصورة الموجهة إلى إحدى العينين على نقطتين، بينما تحتوي الصورة الموجهة إلى العين الأخرى على ثلاث نقاط، مع اشتراكهما في إحدى هذه النقاط. ويستطيع المراقبون ذوو الإبصار العادي رؤية أربع نقاط.

الشكل 3-9

مخطط اختبار الإبصار–02

BT.0500-03-9

3) اختبار الإبصار–03: الإبصار التجسيمي الثخين (اختبار اليعسوب)

يختبر القدرة على إدراك صور معروضة بازدواج المجال بتخاطل كصورة واحدة بعمق ثخين. وتكون الصور للعينين عبارة عن زوج تجسيمي من الصور ليعسوب جناحاه ممتدان ويستطيع المراقبون ذوو الإبصار العادي إدراك الجناحين أمام شاشة العرض.

الشكل 3-10

مخطط اختبار الإبصار–03

الصورة اليسرى الصورة اليمنى

BT.0500-03-10

4) اختبار الإبصار–04: الإبصار التجسيمي الدقيق (اختبار الدائرة)

يختبر القدرة على إدراك صور معروضة بازدواج المجال بتخاطل كصورة واحدة بعمق دقيق. وتعرض تسع رقع اختبار معينة الشكل في كل منها أربع دوائر إحداها فقط بها تخاطل ضئيل. ويستطيع المراقبون ذوو الإبصار العادي إدراك الدائرة ذات التخاطل الضئيل أمام شاشة العرض. ويبين الجدول 3-14 رقم الاختبار والإجابات الصحيحة وزاوية الإبصار التجسيمي عند مسافة تساوي ثلاثة أضعاف الارتفاع.

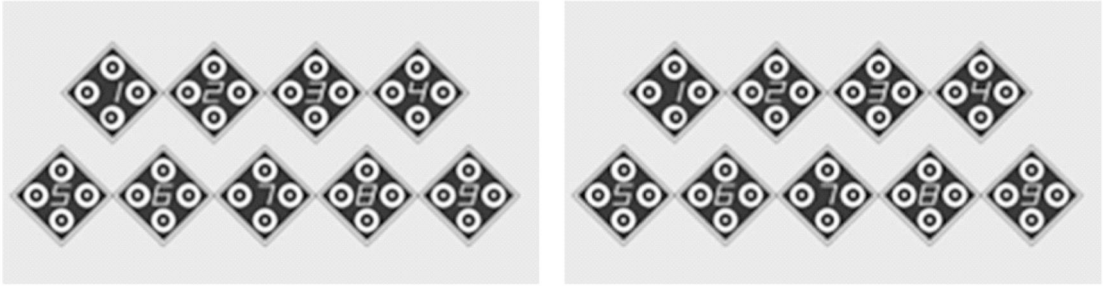
الجدول 3-14

الإجابات الصحيحة والتخاطل

رقم الاختبار	الإجابات الصحيحة	زاوية الإبصار التجسيمي عند مسافة تساوي ثلاثة أضعاف الارتفاع (")
1	أسفل	480
2	يسار	420
3	أسفل	360
4	أعلى	300
5	أعلى	240
6	يسار	180
7	يمين	120
8	يسار	60
9	–	0

الشكل 3-11

مخطط اختبار الإبصار–04



الصورة اليسرى الصورة اليمنى

BT.0500-03-11

5) اختبار الإبصار–05: حد الدمج المقطوع (اختبار الأشرطة)

يختبر القدرة على إدراك صور معروضة بازدواج المجال بتفاوتات مقطوعة كصورة واحدة. ويعرض زوج تجسيمي من شريطين مع تغيير تخاطله بمقدار 10'/ث. ويمكن قياس حدود الدمج للسلسلة المتزايدة والمتناقصة. ويوجَه المراقبون إلى الإفادة بحد الدمج لديهم بمجرد إدراكهم صوراً مزدوجة في السلسلة المتزايدة، واستعادة الدمج لديهم بمجرد إدراكهم صورتي المجال المزدوج كصورة واحدة في السلسلة المتناقصة.

الشكل 3-12

مخطط اختبار الإبصار–05

الصورة اليسرى الصورة اليمنى

BT.0500-03-12

6) اختبار الإبصار–06: حد الدمج غير المقطوع (اختبار الأشرطة)

يختبر القدرة على إدراك صور معروضة بازدواج المجال بتفاوتات غير مقطوعة كصورة واحدة. الصور المعروضة هي نفسها المستعملة للحالة المقطوعة أعلاه، مع التبديل بين الصورتين اليمنى واليسرى.

الشكل 3-13

مخطط اختبار الإبصار–06

الصورة اليسرى الصورة اليمنى

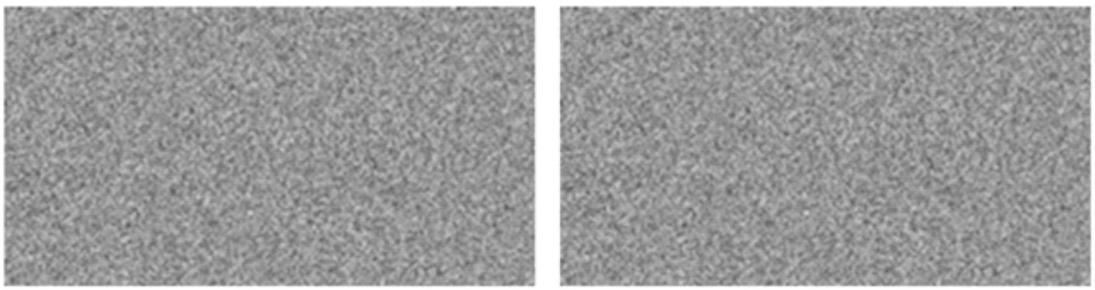
BT.0500-03-13

7) اختبار الإبصار–07: الإبصار التجسيمي الدينامي (اختبار الصورة المجسمة بنقاط عشوائية دينامية)

يختبر القدرة على إدراك العمق في صور مجسمة بنقاط عشوائية متحركة. ويستطيع المراقبون ذوو الإبصار العادي إدراك شكل مستطيل وحركة عمق جيبية في الصورة المجسمة بنقاط عشوائية متحركة.

الشكل 3-14

مخطط اختبار الإبصار–07



الصورة اليسرى الصورة اليمنى

BT.0500-03-14

8) اختبار الإبصار–08: حدة الازدواج البصري (اختبار الحدة)

يختبر حدة الازدواج البصري مع دمج الازدواج البصري، بما في ذلك أي خلل في توازن حدة الإبصار الأحادي قد يحول دون الإبصار التجسيمي الجيد. تضم الصورتان أربعة أعمدة وخمسة سطور تتألف من أحرف E متنوعة الاتجاه والحجم. ويمكن رؤية العمودين الأوسطين بكلا العينين، بينما لا يُرى عمودا اليسار إلا بالعين اليسرى ولا يُرى عمودا اليمين إلا بالعين اليمنى. ويستطيع المراقبون ذوو الإبصار العادي تمييز اتجاه أحرف E بشكل صحيح. وتناظر أحجام الأحرف درجات حدة 1,0 و0,5 و0,33 و0,25 و0,125 عند 3 × الارتفاع.

الشكل 3-15

مخطط اختبار الإبصار–08

الصورة اليسرى الصورة اليمنى

BT.0500-03-15

9 و10) اختبار الإبصار–09: الحول الأفقي (اختبار مادوكس الأفقي) واختبار الإبصار–10: الحول الرأسي (اختبار مادوكس الرأسي)

يقيس هذان المخططان الانحراف الأفقي والرأسي في العين. ويتخذ المحوران الرأسيان موضعاً نسبياً من بعضهما البعض يختلف عن الموضع الذي تتطلبه الظروف الفسيولوجية. وتتكون الصورتان من خط رأسي وآخر أفقي. ويستطيع المراقبون ذوو الإبصار العادي تمييز نقطة تقاطع الخطين في مركز الخطين. ووحدات الأرقام هي الديوبتر المنشوري مع تحديد المسافة بين الحدقتين = 65 mm عند 3,02 × الارتفاع.

الشكل 3-16

مخطط اختبار الإبصار–09

BT.0500-03-16

الشكل 3-17

مخطط اختبار الإبصار–10

BT.0500-03-17

11) اختبار الإبصار–11: تفاوت الصورتين (اختبار أحرف "[ ]")

حالة تسبب اختلاف في صورة كائن المبصرة بإحدى العينين عن الأخرى حجماً وشكلاً. وتتكون الصورة اليسرى من الحرفين "o[" بينما تتكون اليمنى من الحرفين "o]" مع اشتراكهما في موضع حرف "o" ويستطيع المراقبون ذوو الإبصار العادي تمييز حرفي "["و"]" بنفس الحجم ونفس الارتفاع.

الشكل 3-18

مخطط اختبار الإبصار–11

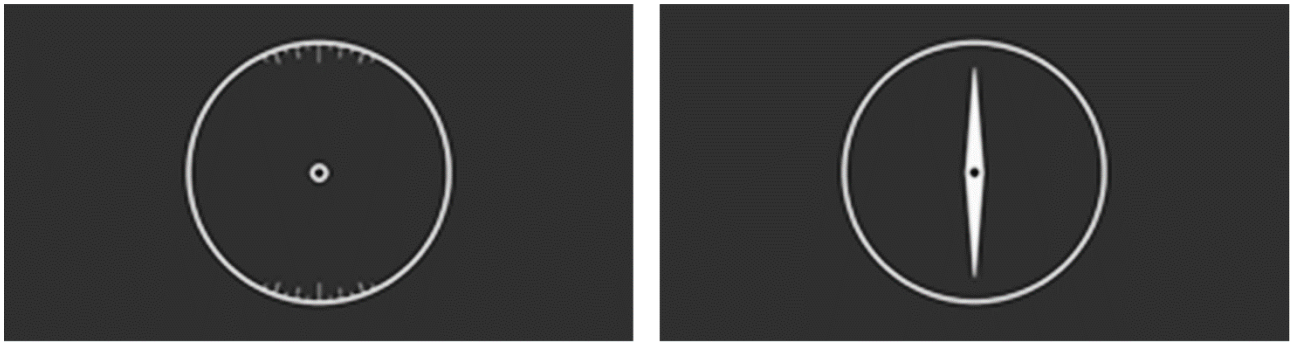
BT.0500-03-18

12) اختبار الإبصار–12: الاحولال التدويري (اختبار الساعة)

انحراف إحدى العينين حول المحور الأمامي الخلفي فقط عند تغطيتها ومنع الدمج. وتتكون الصورة اليسرى من وجه ساعة بينما تتكون اليمنى من عقربي ساعة في موضع الساعة السادسة. ويستطيع المراقبون ذوو الإبصار العادي تمييز الساعة مشيرة إلى تمام السادسة.

الشكل 3-19

مخطط اختبار الإبصار–12



BT.2021-17