الاتحاد الدولي للاتصالات

ITU-R

قطاع الاتصالات الراديوية في الاتحاد الدولي للاتصالات

التوصيـــة ITU-R BT.500-13

(2012/01)

منهجية التقدير الشخصاني لنوعية الصور التلفزيونية

السلسلة BT

الخدمة الإذاعية (التلفزيونية)

الاتحاد الدولي للاتصالات ITU

تمهيـــد

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 الملحق 1 بالقرار 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 http://www.itu.int/publ/R-REC/en)

السلسلة	العنـــوان
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 بعد
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، 2012

© ITU 2012

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).

التوصيـــة ITU-R BT.500-13

منهجية التقدير الشخصاني لنوعية الصور التلفزيونية

(المسألة ITU-R 81/6)

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

مجال التطبيق

تقدم هذه التوصية طرائق لتقدير نوعية الصورة بما فيها طرائق الاختبار العامة وسلالم درجات التقدير وشروط المشاهدة. وتوصي بطريقة سلم الانحطاط ثنائي الحافز (DSIS) وطريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS) فضلاً عن طرائق تقدير أخرى من قبيل الطرائق وحيدة الحافز ومقارنة الحوافز وتقييم النوعية المستمرة وحيدة الحافز (SSCQE) وطريقة الحوافز الثنائية المتآونة للتقييم المستمر (SDSCE).

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 الاختبارات؛

ج) أن تبنّي طرائق مقيسة يعتبر مهماً في تبادل المعلومات بين مختلف المختبرات؛

د) أن تقديرات نوعية الصورة العادية أو الخاصة و/أو الانحطاطات التي تستعمل سلم النوعية خماسي الدرجات وسلم الانحطاط والتي يقوم بها بعض المهندسين المشرفين في أثناء عمليات رتيبة أو عمليات خاصة، قد تستخدم أيضاً بعض الجوانب من الطرائق الموصى بها للتقديرات في المختبر؛

هـ) أن إدخال أنماط جديدة من معالجة إشارة التلفزيون مثل التشفير الرقمي وتنقيص معدل البتات، وأنماط جديدة من الإشارات التلفزيونية التي تستعمل مكونات تعدد إرسال بتقسيم الزمن وربما خدمات جديدة من مثل التلفزيون المعزز والتلفزيون عالي الوضوح (HDTV)، قد يتطلب تغييرات في طرائق التقديرات الشخصانية؛

و) أن إدخال هذه الأنماط من المعالجة والإشارات والخدمات سوف يزيد من احتمالات أن يتعلق أداء كل قسم من سلسلة الإشارة بالعمليات المحققة في الأجزاء السابقة من السلسلة،

توصـــي

1 أن تستعمل طرائق الاختبار العامة، والسلالم وشروط المشاهدة لتقدير نوعية الصورة التي تصفها الملاحق التالية، في تجارب المختبرات وفي التقديرات أثناء التشغيل كلما أمكن ذلك؛

2 أن تستعمل في المستقبل القريب وعلى الرغم من وجود طرائق بديلة ووضع طرائق جديدة، الطرائق التي تصفها الفقرتان 4 و5 من الملحق 1 بهذه التوصية، كلما أمكن ذلك؛

3 أن يتوفر في كل تقارير الاختبارات أشمل وصف ممكن لتشكيلات الاختبارات، وموادها وللمشاهدين والطرائق، نظراً إلى أهمية تحديد قاعدة التقديرات الشخصانية؛

4 أن تعالج المعطيات المجمعة وفقاً للتقنيات الإحصائية المفصلة في الملحق 2 بهذه التوصية، من أجل تسهيل تبادل المعلومات بين المختبرات المختلفة.

الملاحظة 1 – يقدم الملحق 1 المعلومات عن طرائق التقدير الشخصاني لتحديد أداء الأنظمة التلفزيونية.

الملاحظة 2 – يقدم الملحق 2 وصفاً للتقنيات الإحصائية الخاصة بمعالجة المعطيات المجموعة في أثناء الاختبارات الشخصانية.

الملحق 1

وصف طرائق التقدير

1 المقدمة

تستعمل طرائق التقدير الشخصاني لتحديد أداء الأنظمة التلفزيونية بواسطة قياسات تأخذ مباشرة 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.

يوجد عادة صنفان من التقديرات الشخصانية: أولاً، تقديرات تحدد أداء الأنظمة في الشروط المثلى، وتسمى تقديرات النوعية. ثانياً، التقديرات التي تحدد قدرة الأنظمة على المحافظة على النوعية في شروط غير الشروط المثلى للإرسال أو البث، وتسمى تقديرات الانحطاط.

ويجب من أجل إجراء التقديرات الشخصانية المناسبة، أن تنتقى، أولاً، من بين الخيارات المختلفة المتيسرة الخيارات الأفضل لأهداف مشكلة التقدير المعنية وشروطها. وتقدم الفقرة 3، بعد الخصائص العامة المذكورة في الفقرة 2، وللمساعدة، بعض المعلومات عن مشاكل التقدير التي تثيرها كل من الطرائق المذكورة. ثم تفصّل الطريقتان الرئيسيتان الموصى بهما في الفقرتين 4 و5. وأخيراً تعرض الفقرة 6 معلومات عامة عن طرائق بديلة قيد الدراسة.

ويقتصر هدف هذا الملحق على الوصف التفصيلي لطرائق التقدير. غير أن اختيار الطريقة الأنسب تتعلق بأهداف الخدمة للنظام الخاضع للاختبار. ولهذا تعرض في توصيات أخرى لقطاع الاتصالات الراديوية الإجراءات الكاملة لتقدير تطبيقات محددة.

2 الخصائص المشتركة

ترد هنا الشروط العامة للمشاهدة لأغراض التقدير الشخصاني. أما شروط المشاهدة الخاصة المتعلقة بأنظمة محددة للتقدير الشخصاني فترد في التوصيات ذات الصلة.

1.2 الشروط العامة للمشاهدة

فيما يلي وصف بيئات مختلفة تختلف شروط مشاهدتها.

فبيئة المشاهدة المختبرية تفترض شروطاً حرجة للتحقق من الأنظمة. وترد الشروط العامة للمشاهدة للتقدير الشخصاني 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 حساسية بدرجة بسيطة من الحالات العادية للمشاهدة المنزلية.

وترد أيضاً بعض الجوانب المتعلقة باستبانة أجهزة العرض وتباينها.

1.1.2 البيئة المختبرية

1.1.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المختبرية

يجب أن ترتب شروط المشاهدة على النحو التالي:

أ)	نسبة نصوع الشاشة الخاملة إلى ذروة النصوع:	$\geq 0{,}02$
ب)	نسبة نصوع الشاشة عندما تعرض سوية السواد، فقط، في غرفة مظلمة تماماً، إلى النصوع المقابل لذروة البياض:	$\approx 0{,}01$
ج)	اللمعان والتباين في العرض:	يثبت عبر PLUGE (راجع التوصية ITU-R BT.814 والتوصية ITU-R BT.815).
د)	أقصى زاوية مشاهدة بالنسبة إلى الزاوية العادية (تنطبق هذه القيمة على الشاشات CRT، وتخضع القيم المناسبة للشاشات الأخرى للدراسة):	$30°$
هـ)	نسبة نصوع الخلفية وراء مرقاب الصورة، إلى نصوع ذروة الصورة:	$\approx 0{,}15$
و)	لونية الخلفية:	$D65$
ز)	إضاءة القاعة من مصادر أخرى:	ضعيفة

2.1.2 البيئة المنزلية

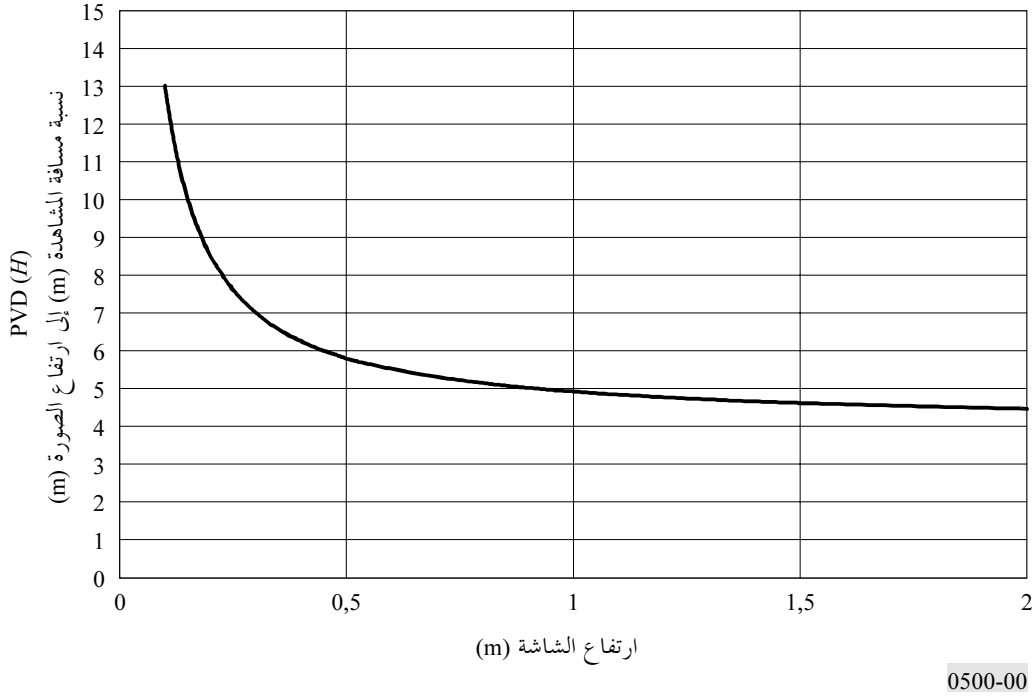
1.2.1.2 الشروط العامة للمشاهدة لأغراض التقدير الشخصاني في البيئة المنزلية

أ)	نسبة نصوع الشاشة الحاملة إلى ذروة النصوع:	$\geq 0{,}02$ (انظر الفقرة 4.1.2)
ب)	اللمعان والتباين في العرض:	يثبت عبر PLUGE (راجع التوصية ITU-R BT.818 والتوصية ITU-R BT.815).
ج)	أقصى زاوية مشاهدة بالنسبة إلى الزاوية العادية (تنطبق هذه القيمة على الشاشات CRT، وتخضع القيم المناسبة للشاشات الأخرى للدراسة):	$30°$
د)	حجم الشاشة للنسق 4/3:	ينبغي أن يفي حجم الشاشة بشروط مسافة المشاهدة المفضلة (PVD)
هـ)	حجم الشاشة للنسق 16/9:	ينبغي أن يفي حجم الشاشة بشروط مسافة المشاهدة المفضلة (PVD)
و)	معالجة جهاز العرض:	دون معالجة رقمية
ز)	استبانة جهاز العرض:	انظر الفقرة 3.1.2
ح)	ذروة النصوع:	$200\ cd/m^2$
ط)	شدة إضاءة المكان على الشاشة (ينبغي قياس الضوء الوارد من المكان على الشاشة عمودياً نسبة إلى الشاشة):	200 lux

ويتعين انتقاء مسافة المشاهدة وحجم الشاشة بحيث تستوفى الشروط PVD. وتظهر القيم PVD (بدلالة حجم الشاشة) في الجدول والرسم التاليين. والقيم الواردة صالحة للنظامين التلفزيونيين SDTV وHDTV نظراً لضآلة الفرق لدى الاختبار.

مسافة المشاهدة المفضلة	ارتفاع الشاشة (*H*)	قطر الشاشة (بالبوصة)	
(*H*)	(m)	النسبة 16/9	النسبة 4/3
9	0,18	15	12
8	0,23	18	15
7	0,30	24	20
6	0,45	36	29
5	0,91	73	60
4-3	1,53 <	120 <	100<

مسافة المشاهدة المفضلة للصور المتحركة



ويُفترض أن يوفر الجدول والرسم البياني معلومات بشأن المسافة PVD وحجم الشاشة المقابلة لها من أجل اعتمادها في التوصيات الخاصة بتطبيقات محددة.

3.1.2 استبانة جهاز العرض

تمثل استبانة أجهزة العرض المتخصصة المزودة بأنابيب الأشعة الكاثودية (CRT) المتخصصة للمعايير المطلوبة في التقدير الشخصاني ضمن مدى تشغيل نصوعها.

ولا يمكن لجميع الأجهزة بلوغ نصوع ذروة قدره 200 cd/m^2.

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

وفي حال استعمال أجهزة تلفزيونية تجارية مجهزة بشاشة CRT تجارية للتقدير الشخصاني، يمكن الحصول على نتيجة غير صحيحة للاستبانة تبعاً لقيمة النصوع.

وفي هذه الحالة يوصى بشدة التحقق من الاستبانتين القصوى والدنيا (مركز الشاشة وزواياها) وتسجيلهما بالقيمة المستخدمة للنصوع.

والنظام الأكثر عملية والمتوفر حالياً لأداء التقدير الشخصاني من أجل التحقق من استبانة أجهزة العرض أو التلفزيونات التجارية هو استعمال نموذج اختبار مسموح مولَّد كهربائياً.

ويمكن التحقق من الاستبانة بالتحليل المرئي. وتقع عتبة الرؤية تقديرياً بين –12/–20 dB. والعيب الرئيسي لهذا النظام هو الانحراف الذي ينجم عن قناع الظل الذي يجعل التقييم المرئي عسيراً، ولكن وجود الانحراف يدل من جهة أخرى على أن إشارة التردد الفيديوي تتجاوز حدود قناع الظل مما يقلل من اعتيان الإشارة الفيديوية.

ويوصى بإجراء مزيد من الدراسات للاختبارات المتعلقة بتعريف الأنابيب الكاثودية.

4.1.2 التباين في أجهزة العرض

قد تؤثر شدة إضاءة الوسط تأثيراً كبيراً على التباين.

ونادراً ما استخدم أجهزة العرض المتخصصة CRT، تقنيات لتحسين تباينها في وسط شديد الإضاءة، *ولذا فمن المحتمل ألاّ تمثل لمعايير التباين المطلوبة إن استعملت في وسط شديد الإضاءة.*

وتستخدم الأجهزة CRT التجارية تقنيات للحصول على تباين أفضل في وسط شديد الإضاءة.

ولحساب التباين في أنبوب CRT ما، يلزم معرفة معامل انعكاس الشاشة K، ومعامل انعكاس الشاشة في أفضل الحالات يساوي $K = 6\%$ تقريباً.

وفي وسط مشوش مع إضاءة شدتها 200 lux و$K = 6\%$، تعطي المعادلة التالية قيمة 3,82 cd/m^2 للنصوع المنعكس من المناطق الحاملة للشاشة:

$$L_{reflected} = \frac{I}{\pi} K$$

ومع القيم المذكورة آنفاً، يبلغ النصوع (cd/m^2) المنعكس لشدة الإضاءة الواردة (lux) نسبة 2% تقريباً.

ويفترض ألا يعطي الأنبوب CRT انعكاسات مرآوية على اللوحة الزجاجية الأمامية التي يصعب عليها تكمية التأثير الدقيق على التباين فيها لأن ذلك يتوقف إلى درجة كبيرة على شروط الإضاءة.

وفي الفقرتين 1.1.2 و2.1.2 تعطى نسبة التباين CR في المعادلة:

$$CR = L_{min} / L_{max}$$

حيث:

L_{min}: نصوع المناطق الخامدة في الإضاءة المحيطة(cd/m^2)

(مع القيم المشار إليها $L_{min} = L_{inactive\ areas} + L_{reflected}$ = 3,82 cd/m^2)

L_{max}: نصوع المناطق البيضاء في الإضاءة المحيطة (cd/m^2)

(مع القيم المشار إليها $L_{max} = L_{white} + L_{reflected}$ = 200 + 3,82 cd/m^2).

وينتج من هذه القيم CR = 0,018 وهي قيمة قريبة جداً من قيمة 0,02 الواردة في الفقرتين 1.1.1.2 و1.2.1.2 أ).

2.2 إشارات المصدر

توفر إشارة المصدر الصورة المرجعية مباشرة، 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 والتتابعات المخزنة رقمياً إشارات المصدر التي يمكن استعادتها بصورة أفضل، ولهذا تعتبر النمط المفضل. يمكن تبادلها فيما بين المختبرات، من أجل الحصول على مقاربات أكثر دلالة فيما بين الأنظمة. وتعتبر أشرطة الفيديو أو تلك الخاصة بالحاسوب أنساقاً مناسبة.

وتوفر ماسحات الشرائح 35 ملليمتر، في المدى القصير، مصدراً مفضلاً للصور الثابتة. وتكون الاستبانة مناسبة لتقدير التلفزيون التقليدي. وقد يعطي قياس اللون ومعلمات الفيلم الأخرى مظهراً ذاتياً مختلفاً لصور آلات التصوير في الأستوديو. وإذا كانت هذه المعلمات تؤثر في النتائج، فيجب استعمال مصادر الأستوديو المباشرة، مع أن ذلك يبدو غالباً أقل ملاءمة. 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.

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

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 علي أجزاء السلسلة التي يمكن أن تدخل تشوهات في المعالجة، حتى لو كانت غير مرئية، أن تسجل 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 يحافظ على هذه التسجيلات في مكتبة مواد الاختبار لاستعمال مستقبلي وفقاً للحاجة، وأن يضاف إليها تقرير مفصل عن تاريخ الإشارة المسجلة.

3.2 انتقاء مواد الاختبار

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

الجدول 1

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. راجع التذييلين 1 و2 للملحق 1 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 التعريف 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، 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 تستعمل، في هذه الحالات، 4 بنود في الأقل: فيكون نصفها مثلاً حرجاً فعلاً، والنصف الآخر حرجاً بعض الشيء.

وقد وضع عدد من المنظمات صوراً وتتابعات ثابتة للاختبار. ويؤمل أن تنظم هذه المواد في المستقبل في إطار عمل قطاع الاتصالات الراديوية. وتقترح صور محددة في التوصيات المتعلقة بتقدير التطبيقات.

يقدم التذييلان 1 و2 للملحق 1 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 المشاهدون

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

وقبل بدء الجلسة، يختار المشاهدون 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). ويتعلق عدد المشاهدين اللازم بحساسية إجراء الاختبار واعتماديته وبالحجم المتوقع للتأثير المقدر. وفي الدراسات محدودة النطاق ذات الطبيعة الاستكشافية مثلاً، يمكن استخدام أقل من 15 مشاهداً. وفي هذه الحالة، ينبغي تحديد هذه الدراسة على أنها "غير رسمية". وينبغي الإبلاغ عن مستوى خبرة المشاهدين المشاركين في تقييم جودة الصورة التلفزيونية.

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 المشاهدين. 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. فيمكنهم، على سبيل المثال، أن يوفروا تفاصيل عن مهنة الأشخاص (موظف في هيئة إذاعية، أو طالب جامعي أو موظف إداري ...) وعن جنسه وعمره.

6.2 تعليمات خاصة بالتقدير

يجب أن تقدم بعناية إلى المشاهدين طريقة التقدير وأنماط الانحطاط أو عوامل النوعية المتوقع حدوثها وسلم التقدير والتتابع والتوقيت. وتستعمل تتابعات تدريجية 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.

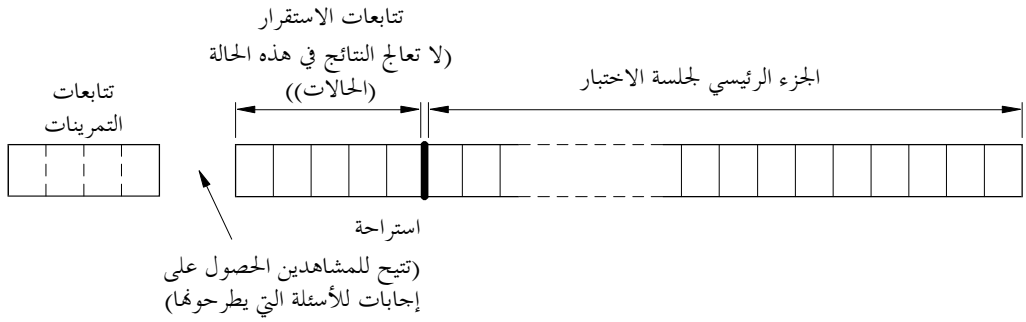
7.2 جلسة الاختبار

يجب ألا تتجاوز مدة الدورة الواحدة نصف الساعة. تقدم بداية الجلسة الأولى، خمس "عروض زائفة" تقريباً من أجل تثبيت آراء المشاهدين. ويجب ألا تؤخذ المعطيات الصادرة عن هذه التقديمات في الحسبان في نتائج الاختبار. وإذا تطلب الأمر اللجوء إلى عدة جلسات، تكفي ثلاثة 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 تدرج التقديرات. ويمكن تكرار بعض العروض من جلسة إلى جلسة من أجل التحقق من التماسك.

الشـــكل 1

بنية تقديم دورة الاختبار



BT0500-01

8.2 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 تستعمل تقنيات ملاءمة المنحني. وتسمح ملاءمة المنحني والإحداثيات اللوغاريتمية برسم خط مستقيم، وهو الشكل المفضل للعرض. يقدم الملحق 2 بهذه التوصية معلومات إضافية عن معالجة المعطيات.

يجب أن تقدم النتائج مع المعلومات التالية:

- تفاصيل عن تشكيلة الاختبار؛
- تفاصيل عن مواد الاختبار؛
- نمط مصدر الصورة ومرقاب العرض (الملاحظة 1)؛
- عدد المشاهدين وأنماطهم (الملاحظة 2)؛

- الأنظمة المرجعية المستعملة؛
- متوسط العلامات العام للتجربة؛
- متوسط العلامات الأصلي والمصحح وفترة الثقة إلى 95% إذا ما سحب مشاهد واحد أو عدة مشاهدين 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 المستخدمة في كل من الاختبارات.

الملاحظة 2 – من الطبيعي أن تؤثر تغييرات مستوى المهارة 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 استعمل للتقديرات التلفزيونية مدى واسع جداً من طرائق الاختبار الأساسية. غير أنه ينبغي، في التطبيق العملي، أن تستعمل طرائق خاصة لمعالجة مشاكل التقدير الخاصة. يقدم الجدول 2 عرضاً لمشاكل التقدير النمطية والطرائق المستعملة لمعالجة هذه المشاكل.

الجـــدول 2

اختيار طرائق الاختبار

مشكلة التقدير	الطريقة المستعملة	الوصف
قياس نوعية الأنظمة بالنسبة إلى مرجع ما	طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS)[1]	التوصية ITU-R BT.500، الفقرة 5
قياس متانة الأنظمة (أي خصائص الأعطال)	طريقة سلم الانخطاط ثنائية الحافز (DSIS)[1]	التوصية ITU-R BT.500، الفقرة 4
تكمية نوعية الأنظمة (عندما لا يتيسر المرجع)	طريقة تدرج النسب[2] أو تدرج الفئات (قيد الدراسة)	التقرير ITU-R BT.1082
مقارنة نوعية أنظمة بديلة (عندما لا يتيسر المرجع)	طريقة المقارنة المباشرة، أو طريقة تدرج النسب[2] أو تدرج الفئات (قيد الدراسة)	التقرير ITU-R BT.1082
التعرف إلى العوامل التي تسمح بإدراك فرق بين الأنظمة، وقياس تأثيرها الإدراكي	طريقة قيد الدراسة	التقرير ITU-R BT.1082
تحديد النقطة التي يصبح عندها الانخطاط مرئياً	تقدير العتبة بطريقة الاختيار القسري أو طريقة الضبط (قيد الدراسة)	التقرير ITU-R BT.1082
تحديد إمكانية ملاحظة فرق بين الأنظمة	طريقة الاختيار القسري (قيد الدراسة)	التقرير ITU-R BT.1082
قياس نوعية تشفير الصورة المجسمة	طريقة سلم النوعية المستمرة ثنائية الحافز (DSCQS)[3]	التوصية ITU-R BT.500، الفقرة 5
قياس الشبه بين تتابعين فيديويين منحطين	الطريقة المستمر ثنائية الحافز للتقييم المستمر (SDSCE)	التوصية ITU-R BT.500، الفقرة 4.6
مقارنة مختلف أدوات مقاومة الأخطاء	الطريقة المستمر ثنائية الحافز للتقييم المستمر (SDSCE)	التوصية ITU-R BT.500، الفقرة 4.6

(1) يبدو من بعض الدراسات أن هذه الطريقة أكثر استقراراً عندما يتيسر مدى كامل من النوعيات.

(2) ترى بعض الدراسات أن هذه الطريقة تكون أكثر استقراراً عندما يتاح المدى الكامل للنوعية.

(3) نظراً لاحتمال حدوث تعب شديد عند تقييم الصور المجسمة، ينبغي إيجاز إجمالي مدة جلسة الاختبار إلى أقل من 30 دقيقة.

4 طريقة الانحطاط ثنائي الحافز ("الطريقة EBU")

1.4 الوصف العام

قد يستدعي تقدير نمطي تقييم نظاماً جديداً 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 الفقرة 5.

والطريقة ثنائية الحافز (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.

تستخدم طريقة سلم الانحطاط الذي يعطي عادة 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.

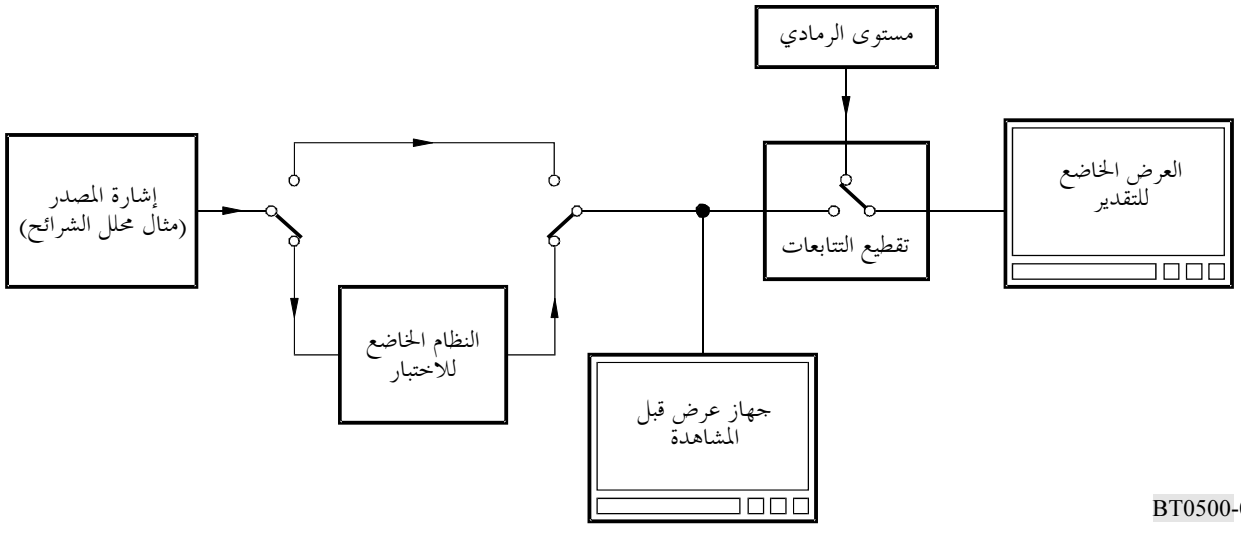
2.4 الترتيب العام للاختبار

تعرف شروط المشاهدة، وإشارات المصدر ومواد الاختبار والمشاهدون وتقديم النتائج، أو تنتقى وفقاً للفقرة 2.

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

الشـــكل 2

الترتيب العام لنظام الاختبار في طريقة سلم الانحطاط ثنائية الحافز



تقدم للمشاهدين شاشة عرض للتقدير تزوَّد بإشارة تمر عبر مفتاح زمني (موقِّت). ويمكن لمسير الإشارة إلى المفتاح الزمني (الموقّت) 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.

3.4 تقديم مواد الاختبار

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

الصيغة I: تقدم الصورة المرجعية أو التتابع المرجعي مرة واحدة فقط كما هو مبين في الشكل 3 أ).

الصيغة II: تقدم الصورة المرجعية أو التتابع المرجعي مرتين كما هو مبين في الشكل 3 ب).

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

4.4 سلالم التقدير

يستعمل سلم الانحطاط خماسي الدرجات:

5 لا يلاحظ

4 يلاحظ لكنه غير مزعج

3 مزعج قليلاً

2 مزعج

1 مزعج جداً.

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

5.4 المدخل إلى التقديرات

يقدم للمشاهدين، في بداية كل دورة، تفسير عن نمط التقدير وسلم التقدير والتتابع والتقدير (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. يطلب من المشاهدين أن يسندوا أحكامهم إلى الانطباع العام الذي تعطيه الصورة، ويعبروا عن هذه الأحكام بواسطة العبارات المستعملة لتعريف سلم التقديرات الشخصانية.

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

6.4 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.

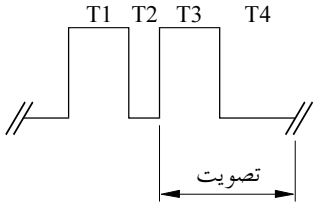
وينبغي اختيار مدى الانحطاطات على نحو يستعمل فيه أغلبية المشاهدين كل الدرجات؛ وأن يكون الهدف التوصل إلى متوسط عام للعلامات (تقدير متوسط كل الأحكام المقدمة في التجربة) قريب من 3.

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

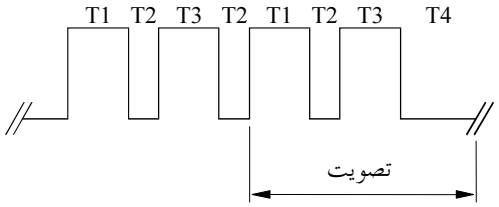
يعطي التذييل 2 للملحق 1 توضيحات إضافية من أجل اختيار سويات الانحطاط.

الشـــكل 3

بنية تقديم مواد الاختبار



أ) الصيغة I



ب) الصيغة II

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.

BT.0500-03

5 طريقة سلم النوعية المستمرة ثنائية الحافز

1.5 الوصف العام

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

وهذه الطريقة دورية أي أنه يقدم للمشاهد زوج من الصور من المصدر نفسه لكن الأولى تمر عبر النظام الخاضع للدراسة بينما تأتي الأخرى من المصدر مباشرة. ويطلب 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.

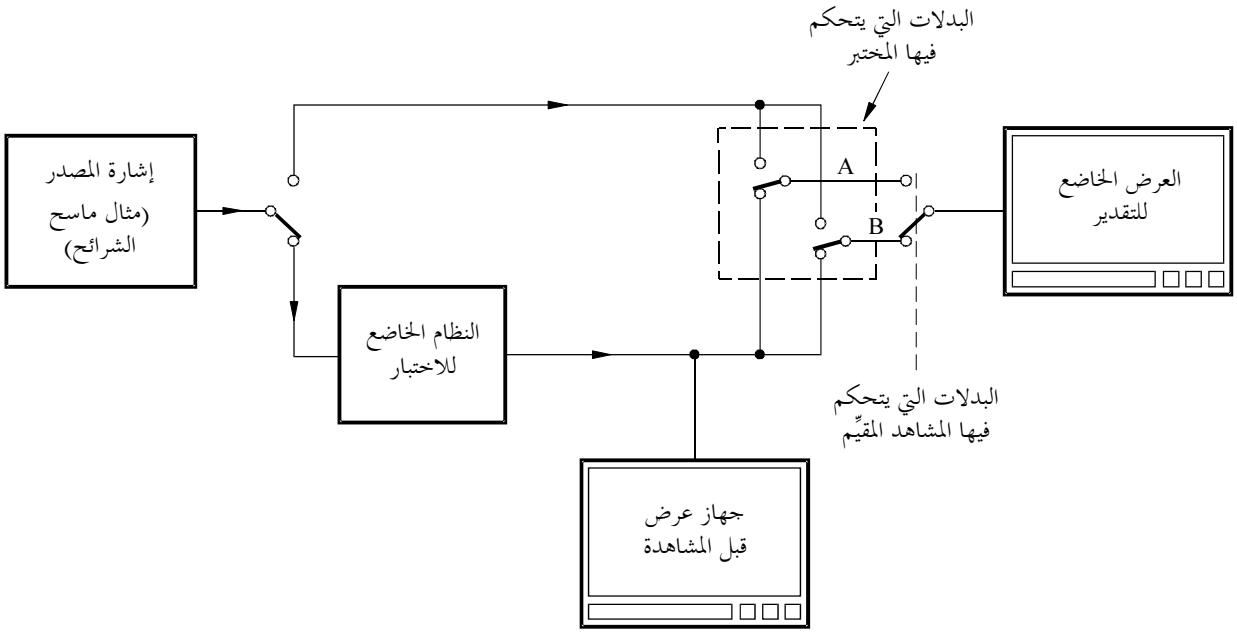
2.5 الترتيب العام للاختبار

تعرف شروط المشاهدة وإشارات المصدر ومواد الاختبار والمشاهدون والمدخل إلى التقدير، أو تنتقى وفقاً للفقرة 2. وتصف الفقرة 6.4 جلسة الاختبار.

يكون الترتيب العام لنظام الاختبار على النحو المبين في الشكل 4 أدناه.

الشـــكل 4

الترتيب العام لنظام الاختبار في طريقة سلم النوعية المستمرة ثنائية الحافز



ثمة صيغتان مختلفتان لهذه الطريقة (I) و(II) تردان فيما يلي.

(I) يسمح للمشاهد الذي يكون وحده في العادة، بالتنقل بين حالتين A وB إلى أن يقرر أنه حدد رأياً يتعلق بكل منهما. يزود الخطان A وB بالصورة المرجعية المباشرة، أو بالصورة عبر النظام الخاضع للاختبار ويتغير خط التغذية عشوائياً من حالة اختبار إلى الأخرى. فيسجل القائم بالاختبار ذلك لكنه لا يعلنه.

(II) يقدم للمشاهدين على التوالي الصور الصادرة عن الخطين A وB لكي يحددوا رأيهم بكل منها. ويزود الخطان A وB في كل تقديم على النحو الموضح في الصيغة (I) أعلاه. ويعتبر أن استقرار نتائج هذه الصيغة التي تستعمل مدى محدوداً من النوعيات لا يزال قيد الدراسة.

BT.0500-04

3.5 عرض مواد الاختبار

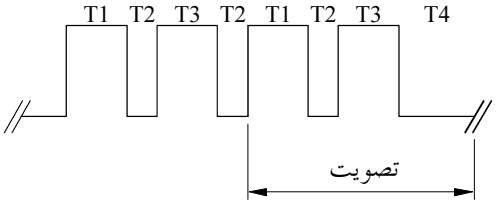
تضم جلسة الاختبار عدة عروض. ففي حالة الصيغة (I) التي تتطلب مشاهداً واحداً، يستطيع هذا الأخير أن يختار لكل عرض التبديل بين الإشارتين A وB حتى يمتلك القياس الذهني للنوعية المصاحبة لكل إشارة. ويستطيع المشاهد بشكل عام تكرار ذلك مرتين أو ثلاث مرات لفترات لا تتجاوز 10 ثوان. أما في الصيغة (II) التي تستعمل عدة مشاهدين في آن معاً، وقبل تسجيل النتائج، يقدم كل زوج من الحالات مرة أو عدة مرات لفترات زمنية متساوية لكي يسمح للمشاهد بأن يقيس ذهنياً النوعية المصاحبة لهذه الحالات، ثم يقدم الزوج مرة أخرى أو عدة مرات بينما يصار إلى تسجيل النتائج. ويتعلق عدد العروض المكررة بطول تتابعات الاختبار. يبدو 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. يبين الشكل 5 بنية العروض.

عندما 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.

الشـــكل 5

بنية عرض مواد الاختبار



مراحل العرض:

T1 =	10 s	التتابع A	
T2 =	3 s	رمادي متوسط صادر عن مستوى فيديوي يقارب 200 mV	
T3 =	10 s	التتابع B	
T4 =	5-11 s	رمادي متوسط	

BT.0500-05

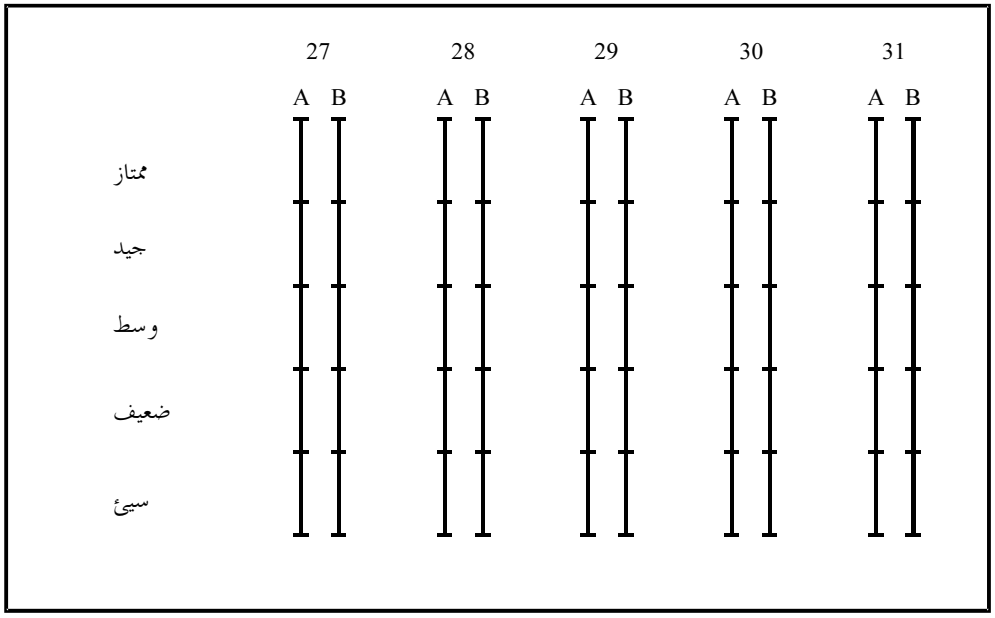
4.5 سلم التقدير

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 المشاهدين يجهلون أية صورة هي الصورة المرجعية. ويتغير موقع الصورة المرجعية في سلسلة الاختبارات على نحو شبه عشوائي.

يطلب فقط من المشاهدين تقدير نوعية الصورة الكلية في كل تقديم من خلال إدراج علامة على سلم رأسي. تطبع 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 عادة لتمييز السويات المختلفة؛ لكنها تدرج من أجل التوجيه العام وتطبع فقط إلى يسار أول سلم في كل صف من 10 أعمدة مزدوجة على صفحة العلامات. ويبين الشكل 6 جزءاً من صفحة علامات نمطية. وتستبعد أية إمكانية التباس بين تقسيمات السلم ونتائج الاختبار بطبع السلالم بالأزرق بينما تسجل النتائج بالأسود.

الشـــكل 6

جزء من استمارة تقدير النوعية التي تستعمل سلالم مستمرة*



* لدى ترتيب مواد الاختبار ضمن جلسة اختبار بالطريقة 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.

BT.0500-06

5.5 تحليل النتائج

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

وقد بينت التجربة أن العلامات الناتجة لتتابعات الاختبار المختلفة تتوقف على درجة حساسية تتابعات صور الاختبار المستخدمة. ومن الممكن الحصول على معرفة أدق لأداء الكودك لو أن نتائج التتابعات المختلفة للاختبار عرضت منفصلةً وليس على شكل متوسطات مجمعة لمجمل التتابعات المستخدمة 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 التذييل 1 للملحق 1). ويجب إجراء مزيد من الدراسات عن حساسية تتابعات الاختبار واحتمال حدوث تتابعات بدرجة ما من الحساسية قبل تكوين فكرة كاملة عن خصائص النظام.

6.5 تفسير النتائج

عند استعمال الطريقة 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 يلاحظ" أو "يلاحظ لكنه غير مزعج" وغيرها من الصفات المستخدمة في الطريقة DSIS).

ويجدر بالذكر ضرورة عدم اعتبار نتائج الطريقة DSCQS قيماً مطلقة ولكن تباين درجات بين الصورة المرجعية وصورة الاختبار. ولذا فمن غير المستصوب ربط الدرجة بصفة 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 الطريقة DSCQS بسبب احتمال سوء فهم المستعملين من غير ذوي الخبرة لمعنى قيم تدرج النوعية التي تنتجها هذه الطريقة.

6 طرائق تقدير بديلة

تستعمل في الشروط المناسبة طرائق وحيدة الحافز وطرائق مقارنة الحوافز.

1.6 الطرائق وحيدة الحافز

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 حافز اختبار آخر.

1.1.6 الترتيب العام للاختبار

تعرف شروط المشاهدة وإشارات المصدر، ومدى الشروط وإرسائها والمشاهدون، والمدخل إلى التقديرات، وتقديم النتائج، أو تنتقى وفقاً للفقرة 2.

2.1.6 انتقاء مواد الاختبار

يجب أن ينتقى محتوى صور الاختبار للتجارب في المختبر، وفقاً للوصف المقدم في الفقرة 3.2.

بعد انتقاء المحتوى، تحضر صور الاختبار لتعكس الخيارات المعنية أو مدى (أو مديات) عامل واحد أو عدة عوامل. وعندما يدرس عاملان أو عدة عوامل، يمكن تحضير الصور وفقاً لطريقتين مختلفتين. تعرض كل صورة، في الحالة الأولى، مستوى واحداً لعامل واحد فقط. وتعرض 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).

3.1.6 جلسة الاختبار

تتضمن جلسة الاختبار سلسلة من اختبارات التقدير التي تقدم عشوائياً، ويفضل أن يكون التتابع العشوائي مختلفاً لكل مشاهد لدى 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 الفقرة 7.2)؛ وتؤكد التجربة عادة على أن لا تعرض نفس الصورة لمرتين متتاليتين بنفس مستوى الانخطاط.

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 عادة مدد 3 و10 و10 ثوان على التوالي. وقد تجمع مؤشرات المشاهدين خلال عرض الحافز أو صورة الثانية الجديدة.

ب) تعرض 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 الجهاز (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 المقدرة في العرضين الثاني والثالث؛ وتؤكد التجربة عادة 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 مباشرة 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 تجمع أدلة المشاهد إلا خلال عرض الصورة الثانية.

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

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

4.1.6 أنماط الطرائق وحيدة الحافز

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

1.4.1.6 طرائق الحكم حسب الفئات بواسطة الصفات

يخصص المشاهدون 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 3 أدناه السلالم التي وضعها قطاع الاتصالات الراديوية. وتستعمل أحياناً في مراقبة التشغيل أنصاف الدرجات. وقد استعملت في بعض الحالات الخاصة سلالم تقدر إمكانية قراءة النص، وجهد القراءة وفائدة الصورة.

الجـــدول 3

سلالم النوعية والانحطاط التي وضعها قطاع الاتصالات الراديوية

سلم خماسي الدرجات			
النوعية		الانحطاط	
5	ممتاز	5	لا يدرك
4	جيد	4	يدرك لكنه غير مزعج
3	وسط	3	مزعج قليلاً
2	ضعيف	2	مزعج
1	سيئ	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.

2.4.1.6 طرائق الأحكام الرقمية حسب الفئات

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

3.4.1.6 طرائق الأحكام التي لا تتوزع حسب الفئات

يخصص المشاهدون، وفقاً لهذه الطرائق، قيمة 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 أو خطاً يرسم بين وسمين في علم الدلالات (طرفي سلم حسب الفئات كما في الجدول 3، 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 الرقمي فيخصص المشاهد لكل صورة أو لكل تتابع صور رقماً يعكس سوية حكمه على عامل محدد (وضوح الصورة مثلاً). ويمكن أن يكون مدى الأرقام المستعملة مقيداً، (من 0 إلى 100 مثلاً) أو غير مقيد. ويصف أحياناً الرقم 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 الرقم، في حالات أخرى، السوية المقدرة. بالنسبة إلى سوية اعتبرت سابقاً سوية "معيارية" (مثل تقدير الاتساع والتقطيع وتقدير النسب).

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

4.4.1.6 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 يستعمل 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، ويستخدم غالباً تحليل التغير أو تقنية مماثلة.

2.6 طرائق مقارنة الحوافز

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

1.2.6 الترتيب العام للاختبار

تعرف شروط المشاهدة والإشارات المصدر، ومدى الشروط وكيفية إرسائها والمشاهدون والمدخل إلى التقدير وتقديم النتائج أو تنتقى وفقاً للفقرة 2.

2.2.6 انتقاء مواد الاختبار

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

3.2.6 جلسة الاختبار

يستعمل اختبار التقدير مرقاباً واحداً أو مرقابين متوائمين ويجري عادة على النحو المتبع في حالات الحافز الوحيد. فإذا استعمل مرقاب واحد، 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 استعمل مرقابان، فتقدم 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 تستخدم عينة من كل الأزواج الممكنة.

4.2.6 أنماط طرائق مقارنة الحوافز

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

1.4.2.6 طرائق الحكم حسب الفئات بواسطة الصفات

يخصص المشاهدون 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 تدرك (مثل، مماثل، مختلف)، أو إلى وجود فروقات تدرك وإلى درجتها (مثل أقل، مماثل، أكثر)، أو تقديرات عن الدرجة والاتجاه. ويعرض الجدول 4 أدناه سلم المقارنة الذي وضعه قطاع الاتصالات الراديوية.

الجـــدول 4

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، أقل وفرق يدرك، وترتيب الشروط، و"المسافات" فيما بين الشروط، إلخ.).

2.4.2.6 طرائق الأحكام التي لا تتوزع حسب الفئات

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

- يخصص المشاهد، في طريقة السلم المستمر، لكل علاقة نقطة أو خطاً يرسم بين وسمين (مثل مماثل، مختلف، أو طرفي السلم حسب الفئات كما في الجدول 4). ويمكن أن يتضمن السلم وسوماً إضافية مرجعية عند نقاط متوسطة. تعتبر المسافة من طرف من الخط قيمة لكل زوج من الشروط.
- أما في الشكل الثاني من الطريقة، فيخصص المشاهد لكل علاقة رقماً يعكس سوية حكمه على عامل محدد (مثل الفرق في النوعية). ويمكن أن يكون مدى الأرقام المستعملة مقيداً أو غير مقيد. وقد يصف الرقم 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.

3.4.2.6 طرائق الأداء

يمكن في بعض الحالات استخلاص قياسات الأداء انطلاقاً من إجراءات مقارنة الحوافز. ويحضر 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. ويطلب من المشاهد أن يقرر أي العنصرين يحتوي على السوية الأعلى أو الأدنى من النعت أو أيهما لا يحتوي على أي نعت. وتعتبر دقة الأداء وسرعتها مؤشرات للعلاقة فيما بين عنصري الزوج الواحد من الصور.

3.6 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 هذه الصور. 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، وضعت التقنيات الجديدة SSCQE التالية واختبرت.

1.3.6 التقدير المستمر للنوعية الإجمالية

1.1.3.6 جهاز التسجيل

يستعمل نظام تسجيل إلكتروني موصول بموصل من أجل تسجيل التقدير المستمر للنوعية الذي يقوم به المختبرين. 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.

2.1.3.6 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.

3.1.3.6 معلمات المشاهدة

ينبغي أن تمتثل شروط المشاهدة تلك المحددة في التوصيات ITU-R BT.500 وITU-R BT.1128 وITU-R BT.1129 وITU-R BT.710.

4.1.3.6 سلالم التقدير

ينبغي إعلام المشاهدين في تعليمات الاختبار بأن طول طريق آلية الانزلاق تقابل سلم تقدير النوعية المستمرة الوارد في الفقرة 4.5.

5.1.3.6 المراقبون

ينبغي تكليف خمسة عشر شاهداً غير خبير كحد أقصى واستعمال الشروط الموصى بها عادة والواردة في الفقرة 5.2.

6.1.3.6 تعليمات إلى المراقبين

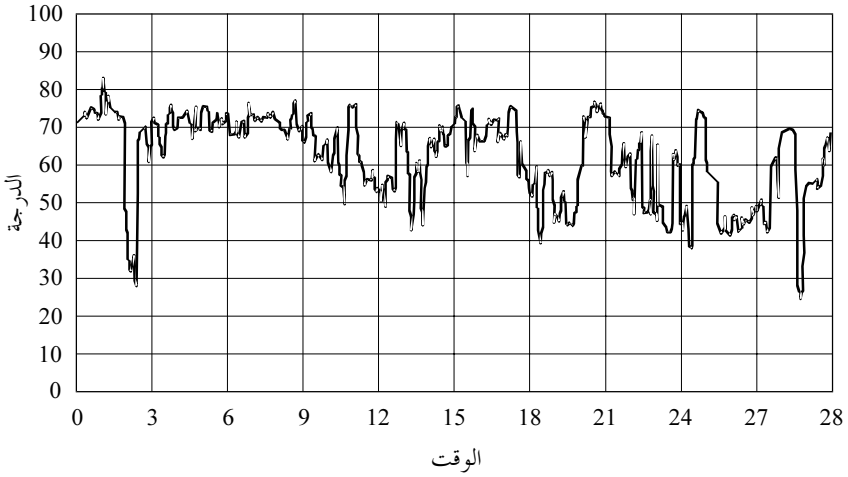
في حال تقييم نوعية الخدمة (مرفقة بإشارات سمعية)، ينبغي إعلام المراقبين بضرورة اعتبار النوعية الإجمالية وليس النوعية الفيديوية بمفردها.

7.1.3.6 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 7).

الشـــكل 7

شروط الاختبار. الكودك X/مقطع البرنامج: Z



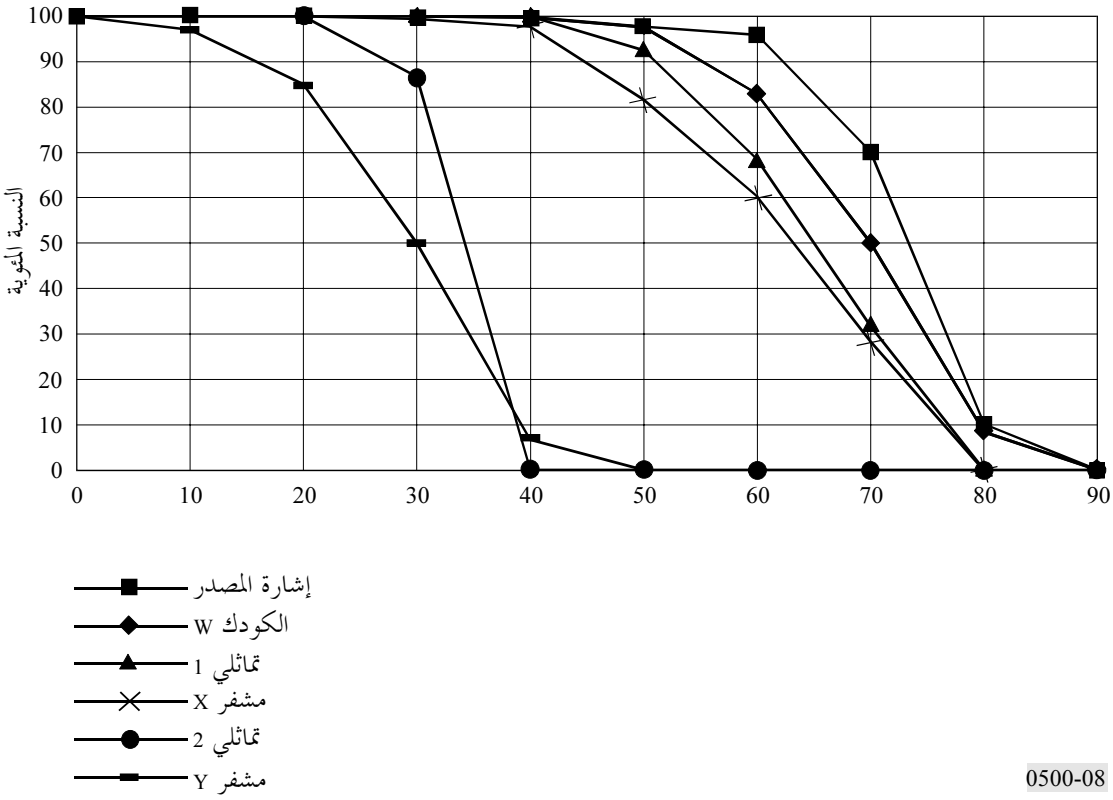
0500-07

غير أن فروق زمن تجاوب المشاهدين قد تؤثر على نتائج التقدير إذا لم يحسب المتوسط إلا استناداً إلى مقطع برنامج. وتجرى 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 8).

الشـــكل 8

متوسط علامات التتابعات المطروحة للتقييم بشأن مقطع من البرنامج Z



2.3.6 معايرة نتائج النوعية المستمرة واستنتاج تقييم عام للنوعية

على الرغم من ملاحظة أن الأخطاء التي تعزى إلى الذاكرة تظهر خلال جلسات طويلة DSCQS لتقييم نوعية التسجيلات الفيديوية الرقمية، تم التحقق مؤخراً من أن هذه الأخطاء ليست جسيمة عند التقييم بالطريقة DSCQS لمقاطع فيديوية مدتها 10 ثوان. لذا هنالك مرحلة ثانية ممكنة لعملية التقييم المستمر للنوعية بحافز وحيد (SSCQE) قيد الدراسة حالياً، وتتمكن في معايرة مخطط النوعية المسجلة بواسطة الطريقة 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 بن 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. وهنالك احتمال مرحلة ثالثة للمنهجية SSCQE تكمن في معالجة نتائج هذه التقييمات المستمرة للنوعية من أجل استنتاج قياس مكافئ للنوعية. وهذا الأمر قيد الدراسة.

4.6 طريقة تقييم النوعية المستمرة بحافزين متآونين (SDSCE)

وضع قطاع الاتصالات الراديوية طريقة التقييم المستمر لأن الطرائق السابقة لم تكن ملائمة 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 عادة على 10 ثوان، وهي بالطبع غير كافية لأن يكوِّن المشاهد 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 الطرائق المعيارية السابقة اختبار تتابعات فيديوية تمثل الواقع تمثيلاً صحيحاً أو على الأقل يمكن تقييم مدى تمثيلها للواقع. ولهذا السبب وضع قطاع الاتصالات الراديوية الطريقة SSCQE التي تتيح قياس النوعية الفيديوية في تتابعات أطول تمثل المحتويات الفيديوية وإحصاءات الخطأ. ومن أجل إنتاج شروط مشاهدة قريبة إلى ابعد حد ممكن من قبل الشروط في الواقع لا تستعمل مرجعيات في الطريقة SSCQE.

أما عند تقييم الأمانة، فيجب إدراج شروط مرجعية. وقد وضعت الطريقة SDSCE استناداً إلى الطريقة 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.

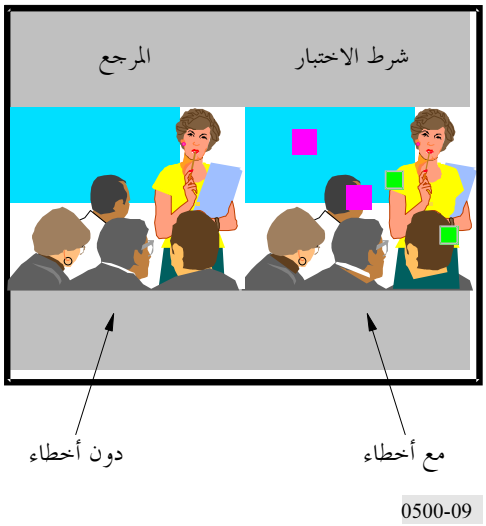
ونتيجة لذلك، وضعت التقنية الجديدة SDSCE التالية وجربت.

1.4.6 إجراء الاختبار

تشاهد مجموعة الأشخاص تتابعين اثنين في نفس الوقت: أحدهما مرجعي والآخر للاختبار. نسق التتابعين هو نسق الصورة المعياري (SIF) أو أصغر، ويمكن عرض التتابعين جنباً إلى جنب في نفس جهاز العرض وإلاَّ فيستعمل جهازا عرض متوافقان (انظر الشكل 9).

الشـــكل 9

مثال لنسق العرض



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.

2.4.6 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 المعلومات 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 قد يتسبب بخطأ من المسؤول عن الاختبار.

وبعد توزيع التعليمات تجرى *جلسة توضيحية* كيما يألف المشاهدون إجراءات التقييم وأنواع الانحطاطات 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.

3.4.6 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 يعالج كل مقطع VS وفق شرط TC واحد على الأقل. كما ينبغي إضافة الصور المرجعية إلى قائمة شروط الاختبار من أجل تكوين 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 والشروط 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.

4.4.6 معالجة البيانات

بعد إجراء الاختبار، يتاح ملف بيانات واحد أو أكثر يضم جميع علامات 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 أجريت وفق هذا البروتوكول بطرق ثلاث هي:

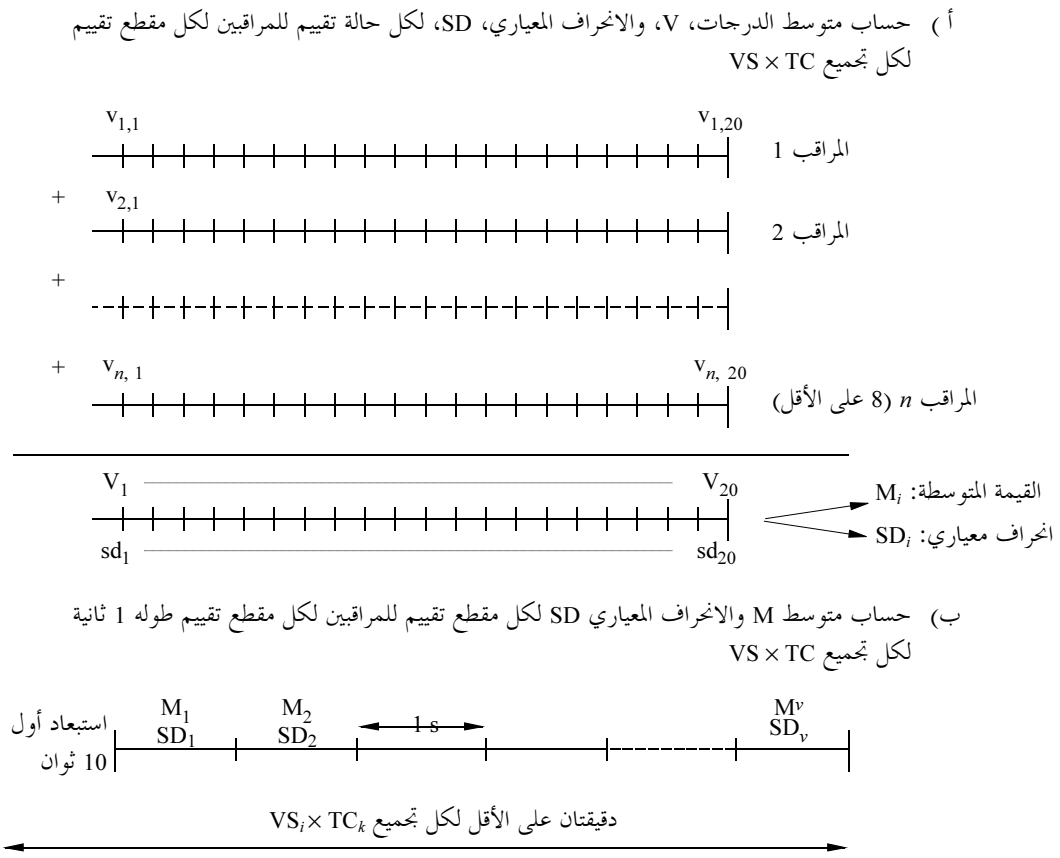
- تحليل إحصائي لكل مقطع VS على حدة؛
- تحليل إحصائي لكل مقطع TC على حدة؛
- تحليل إحصائي شامل لجميع الأزواج TC/VS.

ويشترط 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 10. ويمكن عرض نتائج هذه المرحلة في مخطط زمني كما هو مبين في الشكل 11.
- 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 فواصل الثقة على النحو الوارد في الشكل 12. وتعادل خصائص المطابقة دالة التوزيع الإحصائي التراكمي ببيان العلاقة بين القيم المتوسطة لكل مقطع SOV ومجموع عدد مرات ظهورها.

الشـــكل 10

معالجة البيانات



BT.0500-10

5.4.6 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 الفقرة 2.3.2 من الملحق 2 المتعلقة بالطريقة SSCQE.

وفي الإجراء SDSCE تتوقف موثوقية التقييم على المعلمتين التاليتين:

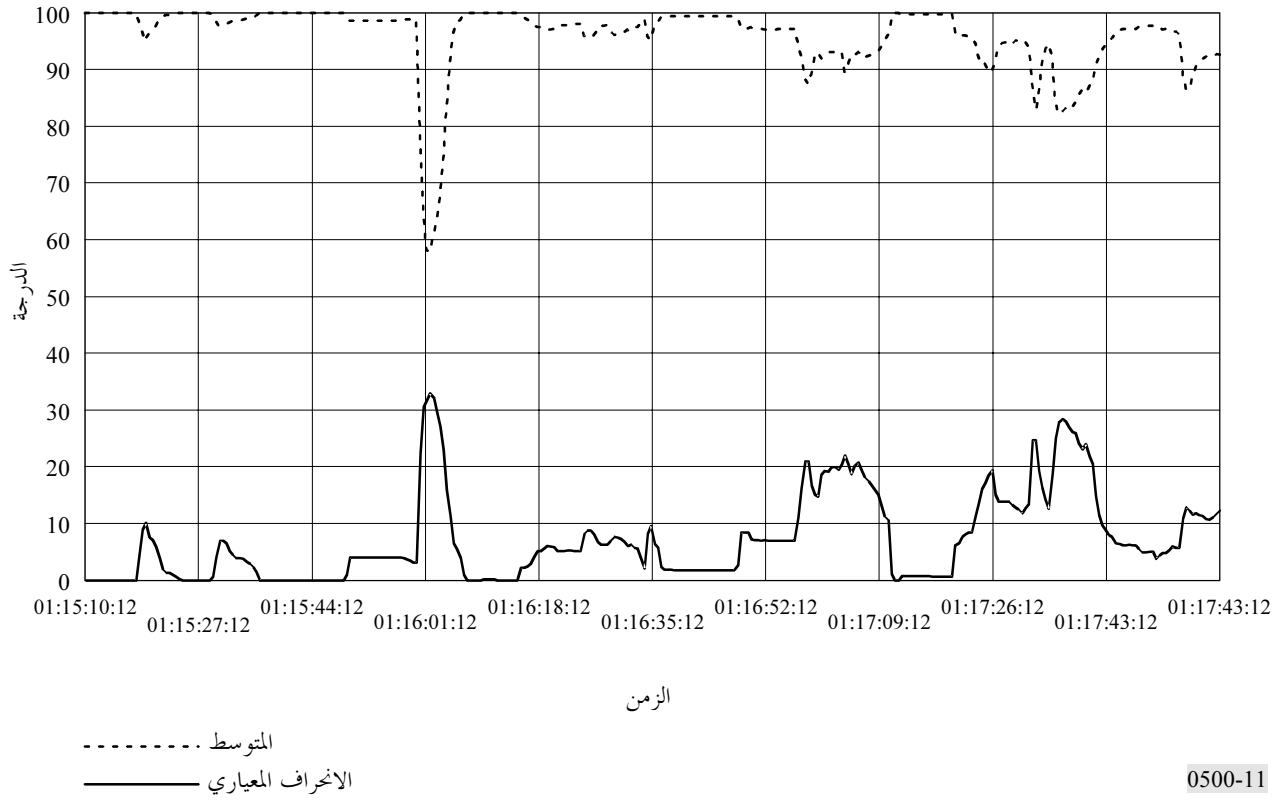
التخالف المتكرر: قد يكون المشاهد أثناء الاختبار مفرطاً بالتفاؤل أو بالتشاؤم أو قد يكون غير فاهم لإجراءات التقييم (مثال معنى سلم التقدير). ويمكن أن يؤدي ذلك إلى سلسلة من التقديرات المتخالفة نوعاً ما وبشكل منتظم مع القيم المتوسطة إن لم تتناقض تماماً معها.

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.

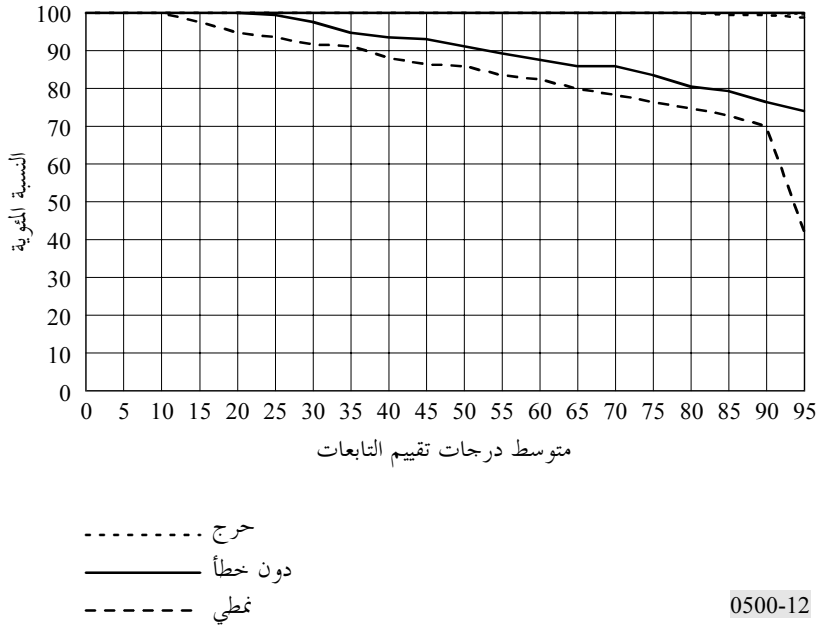
الشـــكل 11

مخطط إجمالي زمني



الشـــكل 12

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



5.6 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 بأحدها غير الأخرى. ولذا يعود انتقاء الطرائق الأكثر ملائمة للحالات المطروحة إلى رغبة الباحث عنها.

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

التذييل 1
للملحق 1

خصائص العطل في محتوى الصورة

1 المقدمة

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 تعرف 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 تحديد خاصية العطل

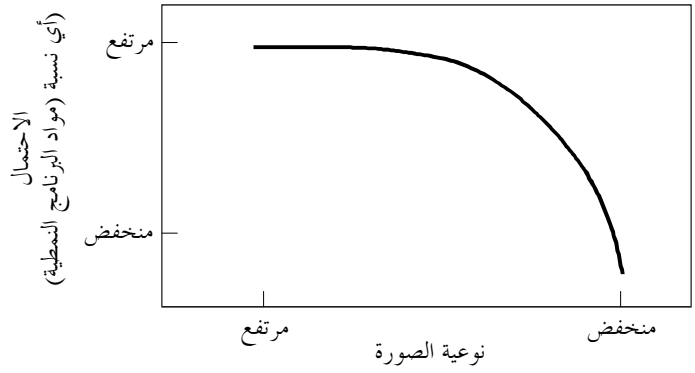
يحدد مفهوم الخاصية في محتوى الصورة نسبة المواد التي يمكن أن تظهر على المدى الطويل والتي يحقق النظام بالنسبة إليها سويات خاصة من النوعية. ويمثل ذلك في الشكل 13.

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 مشاهد 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 14 مثالاً لتوزيع هذا النمط.
- *المرحلة 3*: استخلاص قدرة النظام على المحافظة على النوعية من خلال وسائل تجريبية عندما تتزايد السوية الحرجة للنظام. ويتطلب ذلك، في التطبيق العملي، تقديراً شخصانياً للنوعية المحققة في النظام مع مواد تنتقى لاعتيان مدى الحرج المعرف في المرحلة 2. وينتج عن ذلك دالة تحدد العلاقة بين النوعية المحققة في النظام وسوية الحرج لمواد البرنامج. ويعطي الشكل 15 مثالاً عن هذه الدالة.
- *المرحلة 4:* استخلاص المعلومات من المرحلتين 2 و3 بهدف اشتقاق خاصية العطل في محتوى الصورة على النحو المقدم في الشكل 13.

الشـــكل 13

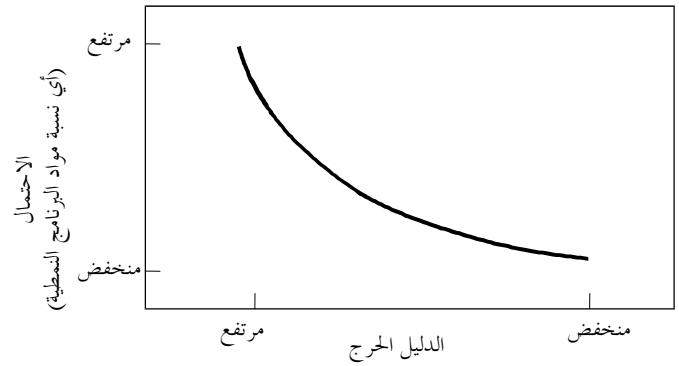
التمثيل البياني لخاصية العطل الممكن في محتوى الصورة



0500-13

الشـــكل 14

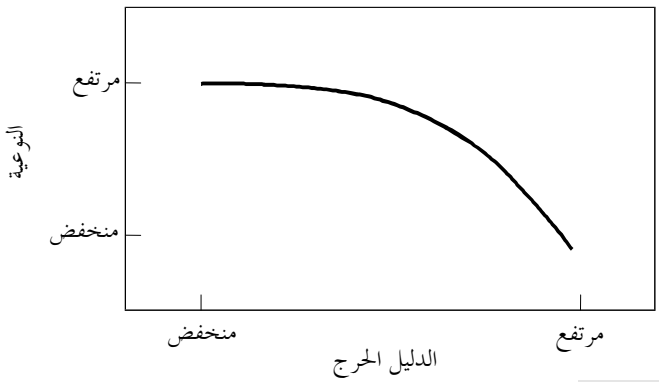
احتمال ظهور مواد ذات سويات حرج مختلفة



0500-14

الشـــكل 15

مثال لدالة تحديد العلاقة بين النوعية وسوية حرج مواد البرنامج



BT.0500-15

3 استعمال خاصية العطل

تعتبر خاصية العطل التي توفر صورة شاملة للأداء الذي يمكن تحقيقه على مدى أنماط البرامج الممكنة، أداء مهمة لدراسة ملاءمة الأنظمة. يمكن استعمال خاصية العطل في ثلاث حالات:

- تحسين معلومات 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 مباشرة حتى عندما تنتج عن دلائل مختلفة للحرج مميزة لكل نظام.

على الرغم من أن الطريقة الموضحة في هذه التوصية توفر طريقة 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 بالتوصية ITU-R BT.1129.

التذييل 2
للملحق 1

طريقة تحديد خاصية عطل مركبة بالنسبة إلى محتوى البرنامج وشروط الإرسال

1 المقدمة

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

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

يصف هذا التذييل 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 تحليل محتوى البرنامج

تتضمن هذه المرحلة عمليتين، يستخلص في الأولى قياس مناسب لمحتوى البرنامج، وتقدر 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 علاقة 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 أستوديو مناسب للنظام المعني في الاستبانة، وتردد الرتل والنسبة الباعية. يعطي تحليل هذه العينة الترددات النسبية لظهور القيم الإحصائية التي تعتبر تقديرات لاحتمالات لظهور 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 (راجع أيضاً التذييل 1 للملحق 1). وتضم هذه النتائج إلى نتائج تحليل شروط الإرسال من أجل التحضير لمرحلة الإجراء النهائية.

3 تحليل قناة الإرسال

تتضمن هذه المرحلة كذلك عمليتين، يستخلص في الأولى قياس أداء قناة الإرسال، ويقدر 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؛
- 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. وتضم هذه النتائج إلى نتائج تحليل محتوى البرنامج من أجل التحضير لمرحلة الإجراء النهائية.

4 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. (راجع الفقرة 5 من الملحق 1). المرجع هنا هو صورة بنوعية الأستوديو 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 تستقبل 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 متناثرة عند السويات الحرجة الدنيا، وأكثر كثافة عند السويات الأكثر حرجا. وتنتقى القيم الإحصائية للقناة بطريقة مماثلة. ثم يصار إلى ضم هذين المصدرين المستقلين للتأثير بحيث يؤدي ذلك إلى محتويات وشروط إرسال مختلطة ذات احتمال معروف.

وتستعمل 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.

التذييل 3
للملحق 1

الآثار السياقية

تنتج الآثار السياقية عندما يتأثر التقدير الشخصاني لصورة ما بترتيب وشدة الانحطاطات الظاهرة. وعلى سبيل المثال، إذا عرضت 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 (الطريقة DSCQS والصيغة 2 من الطريقة DSIS وطريقة المقارنة) مستعملة لتقييم نوعية الصور. وأنتجت مواد الاختبار باستعمال التشفير MPEG (ML@MP) بموازاة تخفيف الاستبانة الأفقية. واستعملت 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. وطبقت 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. واستعمل شرطا الاختبار الأساسيان B2 وB3 لتحيد آثار السياق.

وتدل نتائج المختبرات مجتمعة على عدم وجود آثار سياقية للطريقة DSCQS. لكن هذه الآثار كانت واضحة في الطريقة DSIS وطريقة المقارنة. أما أشد الآثار السياقية فظهرت في النوع II من الطريقة 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 الطريقة 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 2

تحليل النتائج وتقديمها

1 المقدمة

تجمع كمية كبيرة من المعطيات في أثناء التجارب الشخصانية لتقدير أداء نظام تلفزيوني. وهذه المعطيات تقدم على شكل وريقات علامات المشاهدين أو مكافئها الإلكتروني ويجب أن يتم تكثيفها بواسطة تقنيات إحصائية بحيث تمثل النتائج بشكل بياني و/أو يلخص أداء الأنظمة الخاضعة للاختبار.

ويطبق التحليل التالي على نتائج سلم الانحطاط ثنائي الحافز (DSIS) وسلم النوعية المستمرة ثنائي الحافز (DSCQ) لطريقتي تقدير نوعية الصورة التلفزيونية (الفقرات 4 و5 و6 من الملحق 1). ويوزع الانحطاط، في الحالة الأولى، على سلم خماسي الدرجات بينما يستعمل في الحالة الثانية سلم التقدير المستمر وتُقيَّس النتائج (الفرق بين تقديرات الصورة المرجعية والصورة الحالية الخاضعة للاختبار) عند قيم صحيحة تتراوح بين 0 و100.

2 طرائق التحليل المشتركة

تؤدي الاختبارات المحققة وفقاً لمبادئ الطريقتين المذكورتين في الملحق 1 إلى توزيع لقيم صحيحة تتراوح بين 1 و5 أو بين 0 و100. ويتضمن هذا التوزيع التباين بين أحكام المشاهدين 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 حساب العلامات المتوسطة

تكمن الخطوة الأولى لتحليل النتائج في حساب العلامة المتوسطة، $\overline{u}_{jkr}$، لكل حالة اختبار:

$$\overline{u}_{jkr} = \frac{1}{N} \sum_{i=1}^{N} u_{ijkr} \tag{1}$$

حيث:

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

N: عدد المشاهدين.

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

2.2 حساب فاصل الثقة

1.2.2 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:

$$\left[\overline{u}_{jkr} - \delta_{jkr},\ \overline{u}_{jkr} + \delta_{jkr}\right]$$

حيث

$$\delta_{jkr} = 1.96\,\frac{S_{jkr}}{\sqrt{N}} \tag{2}$$

يحسب الانحراف النمطي S_{jkr}، بواسطة العبارة التالية:

$$S_{jkr} = \sqrt{\sum_{i=1}^{N} \frac{(\overline{u}_{jkr} - u_{ijkr})^2}{(N-1)}} \tag{3}$$

إذا كانت نسبة الاحتمال 95%، تكون القيمة المطلقة للفرق، العلامة المتوسطة في التجربة والعلامة المتوسطة "الحقيقية" (لعدد كبير جداً من المشاهدين) 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 المراقبين المشاركين في التقدير.

2.2.2 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 النوعية الناتجة عن التجارب بهذه التشوهات) ينبغي حساب فاصل الثقة باستعمال متحولات متغيرات إحصائية مع مراعاة تشتت قيم هذه المتغيرات.

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

3.2 اختيار المشاهدين

1.3.2 اختيار المشاهدين للطريقتين DSIS وDSCQS والطرائق البديلة باستثناء الطريقة SSCQE

أولاً، يجب التأكد من أن توزيع العلامات لكل مشهد عادي أو غير عادي بواسطة الاختبار β_2 (من خلال حساب معامل kurtosis 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. ويجب عندها أن تقارن العلامات u_{ijkr} لكل مراقب مع القيمة المتوسطة $\overline{u}_{jkr}$ المصاحبة زائداً الانحراف النمطي المصاحب مضروب بـ 2 (عادي) أو بـ $\sqrt{20}$ (غير عادي)، P_{jkr}، ومع القيمة المتوسطة المصاحبة ناقصاً الانحراف النمطي نفسه مضروباً في 2 أو في $\sqrt{20}$، Q_{jkr}. ويجب، في كل مرة تكون علامة المشاهد خارج هذا المدى، أن تسجل على عداد مصاحب لكل مشاهد. 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 يستبعد المشاهد، (انظر الملاحظة 1).

الملاحظة 1 – لا يطبق هذا الإجراء أكثر من مرة واحدة على نتائج تجربة معينة. ويجب، إضافة إلى ذلك، أن ينحصر استعمال الإجراء في الحالات التي يوجد فيها عدد قليل نسبياً من المشاهدين (أقل من 20، مثلاً)، جميعهم غير اختصاصي.

يوصى بهذا الإجراء لطريقة EBU (DSIS)؛ وقد نجح أيضاً استخدامه في الطريقة DSCQS وفي الطرائق البديلة.

ويمكن التعبير رياضياً عن الإجراء المذكور آنفاً على النحو التالي:

يحسب لكل عرض اختبار المتوسط $\overline{u}_{jkr}$، الانحراف المعياري S_{jkr}، ومعامل kurtosis، $\beta 2_{jkr}$ حيث يغطي القيمة $\beta 2_{jkr}$ بالمعادلة التالية:

$$m_x = \frac{\sum_{i=1}^{N} (u_{ijkr} - \overline{u}_{ijkr})^x}{N} \quad \text{مع} \quad \beta_{2jkr} = \frac{m_4}{(m_2)^2} \tag{4}$$

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

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

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

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

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

غير ذلك:

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

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

فإذا كان $\frac{P_i + Q_i}{J \cdot K \cdot R} > 0.05$ و$\left|\frac{P_i - Q_i}{P_i + Q_i}\right| < 0.3$، 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 $J \cdot K \cdot R$، لكن يلاحظ أن بعض التقديرات قد تفضي إلى أعداد غير متساوية من التتابعات لكل شرط اختبار).

2.3.2 الاختيار للطريقة SSCQE

فيما يتعلق بالاختيار الخاص للمشاهدين الذين يستعملون إجراء الاختبار 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 واختيار 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 هذه العلامات الخاصة 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 زمن لكل تشكيلة اختبار وفقاً للمعادلة الرياضية الواردة لاحقاً.

ويحسب لكل نافذة زمن لكل تشكيلة اختبار وباستعمال التقديرات u_{ijkr} لكل مشاهد، المتوسط، $\overline{u}_{jklr}$، والانحراف المعياري S_{jklr}، والمعامل $\beta 2_{jklr}$. وتعطى القيمة $\beta 2_{jklr}$ في المعادلة:

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

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 $\beta 2_{jklr} \geq 2$ و≤ 4، يكون:

إذ $u_{njklr} \geq \overline{u}_{jklr} + 2\, S_{jklr}$ يكون $P_i = P_i + 1$

إذ $u_{njklr} \leq \overline{u}_{jklr} + 2\, S_{jklr}$ يكون $Q_i = Q_i + 1$

غير ذلك:

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

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

إذا كان $\frac{P_i}{J \cdot K \cdot L \cdot R} > X\%$ أو $\frac{Q_i}{J \cdot K \cdot L \cdot R} > X\%$، 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.

وتتيح هذه العملية استبعاد المراقبين الذين قدموا تقديرات شديدة التخالف مع متوسط العلامات. ويبين الشكل 17 مثالين (منحنيات بيانيان لطرفي التخالف الشديد). لكن معيار الاستبعاد هذا لا يتيح كشف عمليات القلب المحتملة التي تشكل مصدراً هاماً للتباين. ولهذا السبب تقترح مرحلة العملية الثانية.

المرحلة 2: كشف قلب التقدير محلياً

في المرحلة 2 يستند الكشف أيضاً على اختيار صيغة ترد في الملحق 2 من هذه التوصية. وتضاف فيها تعديل طفيف يتعلق بمجال التطبيق. وتتكون مجموعة البيانات الداخلة أيضاً من علامات جميع نوافذ الزمن (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 اللامات هذا لكل نافذة زمنية لكل تشكيلة اختبار "عادياً"، أم لا باستعمال الاختبار β_2 بين 2 و4، اعتبر التوزيع عاديا. وعندئذٍ يستعمل الإجراء لكل نافذة زمنية لكل تشكيلة اختبار وفق المعادلة الرياضية الواردة لاحقاً.

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:

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

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

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

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:

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

وتحسب لكل نافذة زمنية لكل تشكيلة اختبار المتوسط $\overline{u}^*_{jklr}$ والانحراف المعياري، c S^*_{jklr} والمعامل $\beta_2{}^*_{jklr}$. وتحسب القيمة $\beta_2{}^*_{jklr}$ في المعادلة:

$$m_x = \frac{\sum_{n=1}^{N} (u^*_{njklr})^x}{N} \quad \text{مع} \quad \beta_2{}^*_{jklr} = \frac{m_4}{(m_2)^2}$$

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

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

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

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

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

غير ذلك

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

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

إذا كان $\frac{P^*_i + Q^*_i}{J \cdot K \cdot L \cdot R} > Y$ و $\left|\frac{P^*_i - Q^*_i}{P^*_i + Q^*_i}\right| < Z$، 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 العملية التي تسمح بتعيين علاقة بين العلامة المتوسطة والقياس الموضوعي لتشوه الصورة

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

1.3 التقريب بواسطة دالة منطقية رياضية متناظرة

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

يمكن معالجة المعطيات $\overline{u}$ على النحو التالي:

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

$$p = (\overline{u} - u_{min}) / (u_{max} - u_{min}) \quad (5)$$

حيث:

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

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

ويظهر 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:

$$p = 1 / [1 + \exp(D - D_M) G] \quad (6)$$

حيث D_M وG قيمة موجبة أو سالبة.

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

$$I = (1/p - 1) \quad (7)$$

ويمكن اشتقاق القيمتين D_M وG من المعطيات التجريبية بعد التحويل التالي:

$$I = \exp(D - D_M)\, G \tag{8}$$

ويؤدي استعمال سلم لوغاريتمي لقيم I إلى علاقة خطية:

$$\log_e I = (D - D_M)\, G \tag{9}$$

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:

$$S = \frac{D_M - D}{\log_e I} = \frac{1}{G} \tag{10}$$

التي تعطي قيمة G المثلى. D_M 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 التقريب بواسطة دالة لا تناظرية

1.2.3 وصف الدالة

إن تقريب العلاقة بين علامات الاختبار والقياس الموضوعي لتشوه الصورة باستعمال دالة تناظرية رياضية عملية ناجحة في معظم الحيان 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 (8) باستعمال المعادلة:

$$I = (d / d_M)^{1/G} \tag{11}$$

وبالتالي تصبح المعادلة (6):

$$p = 1 \Big/ \left[1 + (d / d_M)^{1/G}\right] \tag{12}$$

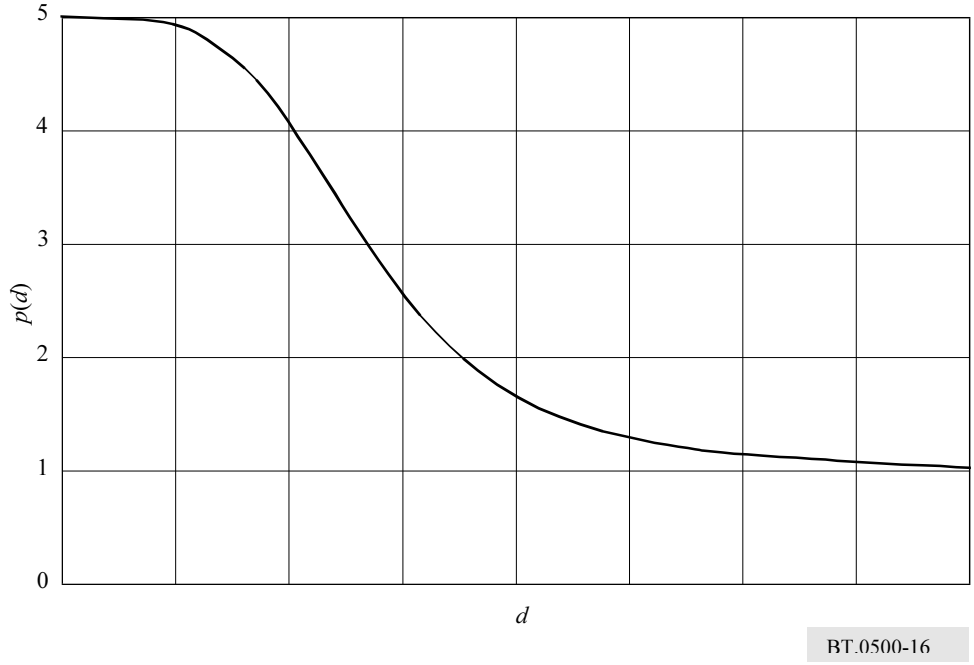
ومن شأن هذه الدالة تقريب الدالة المنطقية بطريقة لا تناظرية.

2.2.3 تقدير معلمات التقريب

يمكن الحصول على تقدير المعلمات المثلى للدالة التي توفر أدنى الأخطاء المتبقية بين المعطيات الفعلية والدالة، بواسطة أية خوارزمية تقدير تكراري. يبين الشكل 16 مثالاً عن استعمال دالة لا تناظرية من أجل تمثيل معطيات شخصانية فعلية. ويسمح هذا التمثيل بتقدير قياسات موضوعية تقابل قيمة شخصانية مهمة: 4,5 مثلاً على السلم خماسي الدرجات.

الشـــكل 16

تقريب لا تناظري



3.3 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 عدم استعمال القيم القصوى للسلم وخاصة معلومات النوعية المرتفعة. وقد ينتج ذلك عن عدة عوامل بما فيها الإحجام النفسي عن إطلاق أحكام متطرفة. وعلاوة على ذلك قد يسبب استعمال المتوسط الحسابي للأحكام وفقاً للمعادلة (1) قرب نهايات السلم الحصول على نتائج خاطئة بسبب التوزيع غير الغوسي للتقديرات في هذه المناطق.

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

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.

الجـــدول 5

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:

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

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

حيث:

u_{corr}: علامة مصححة

$\overline{u}$: علامة تجريبية غير مصححة

u_{max}، u_{min}: حدود سلم التقدير

u_{mid}: منتصف سلم التقدير

$u0_{max}$، $u0_{min}$: الحدان الأدنى والأعلى لاتجاه العلامات التجريبية.

4.3 إدماج جانب الاعتمادية في الرسوم 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 الدوال 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 17). ونحصل بذلك على تقدير لمنطقة الثقة المستمرة عند 95%.

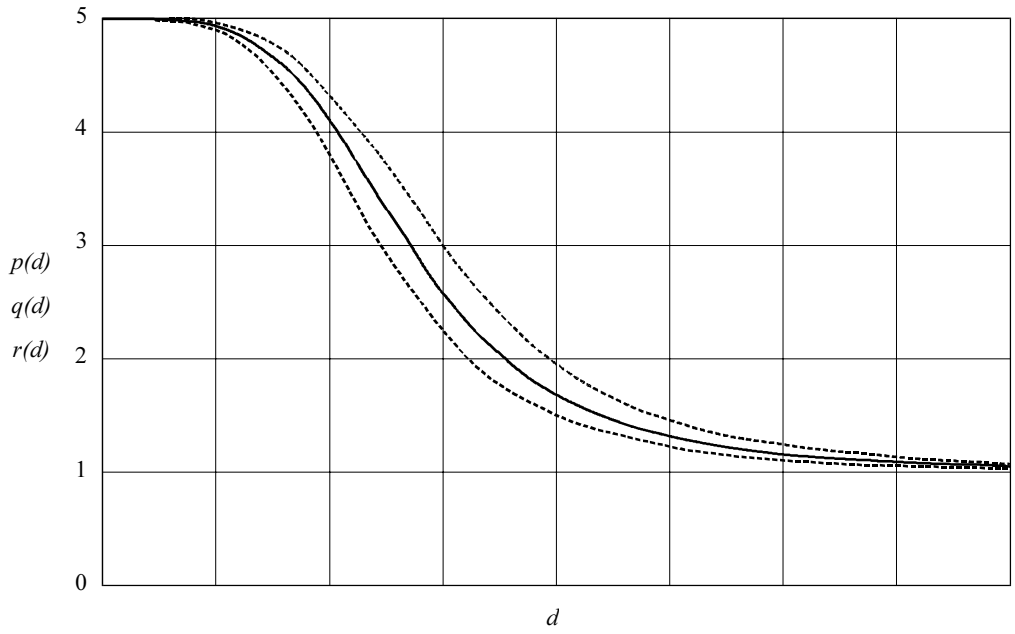
يمكن أن تقرأ من ثم للدرجة 4,5 (عتبة الرؤية الخاصة بالطريقة) قراءة مباشرة على الرسم البياني، فترة الثقة 95% المقدرة والتي يمكن أن تستعمل لتحديد مدى مسموح به.

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 يستنتج أن ثمة مشكلة في إجراء الاختبار أو أن نموذج الدالة المختار لم يكن النموذج الأمثل.

الشـــكل 17

حالة خاصية انحطاط لا تناظرية



p(d): متوسط الدرجات
q(d): أدنى الدرجات
r(d): أعلى الدرجات
d: القياس الموضوعي للانحطاط

0500-17

4 الاستنتاجات

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

يؤدي الإجراء أيضاً إلى تقدير كميات عامة متوسطة ليست مهمة 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.

الملحق 3

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.

ويعد كل تقدير شخصاني تبعاً لخمس مراحل متعاقبة ومترابطة هي: إعداد الاختبار وإجراء الاختبار، ومعالجة البيانات، وتقديم النتائج وتفسيرها. ويتوزع العمل عادة 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.
- تجري الاختبار مختبرات مشاركة مختلفة باستعمال الشريط الذي سبق إعداد. 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 عادة إلى مختبر آخر مشارك أن ينسق عملية جمع البيانات الخام الناتجة وإصدار النتائج وهي عملية يمكن إنجازها بصورة مستقرة 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 الواردة في التوصية ITU-R BT.500.

وتتألف من ملفات نصوص يظهر تنظيمها في الجدولين 6 و7، وتبنى قواعد تركيبها حول وسوم وميادين إضافةً إلى مجموعة محدودة من الرموز الخاصة (مثل "[" و"]" و" " و"↵" 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).

الجـــدول 6

تحديد نتائج نسق الملف

الشرح	تحديد نسق الملفات والشكل
[معرف القسم]	[Test framework] ↵
معرف الطريقة المستخدمة في التوصية ITU-R BT.500	Type = *"DSCQS" or "DSIS I"*, "*DSIS II", etc.*↵
عدد الجلسات[1] التي يوزع فيها اختبارات	Number of sessions = *1 ≤integer ≤x*↵
تعريف سلم التقدير (انظر متطلبات الطريقة إن وجدت)	Scale minimum = *integer*↵ Scale maximum = *integer*↵
طول الخط القطري (بالبوصة)	Monitor size = *integer*↵
	Monitor make and model = *chain of characters*↵
[معرف القسم]	[RESULTS] ↵
عدد ملفات النتائج[1] التي ينظر فيها	Number of results = *1 ≤integer ≤ y*↵
اسم الملف Full.DAT (انظر الجدول 7) بما فيه المسار	Result(*j*).Filename(s) = *character string*.DAT↵
	....
اسم ملف النتائج المتداول	Result(*j*).Name = *character string*↵
تعريف المختبر الذي أجرى الاختبار	Result(*j*).Laboratory = *character string*↵
إجمالي عدد المراقبين	Result(*j*).Number of observers = *1 ≤integer ≤ N*↵
يدل على ما إذا كانت التقديرات المجمعة في التدريبات مدرجة في الملف DAT المرفق	Result(*j*).Training = *"Yes" or "No"* ↵
[معرف القسم]	[Result(*j*).Session(*i*).Observers] ↵
[تعريف المراقب]	O(*k*).First Name = *character string*↵
	O(*k*).Last Name = *character string*↵
خياري	O(*k*).Sex = *"F" or "M"* ↵
خياري	O(*k*).Age = *integer*↵
زمر اجتماعية اقتصادية رئيسية (مثال: عمال، طلاب)	O(*k*). Occupation = *character string*↵
مسافة المشاهدة، ارتفاع الشاشة (3 *H*, 4 *H*, 6 *H*)	O(*k*).Distance = *integer*↵

[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.

الجـــدول 7

نسق ملف نصوص النتائج DAT للبيانات الخام

الشرح	نسق وشكل الملف Filename.DAT
يتألف ملف البيانات الخام DAT من قيم التقدير يفصل بينها فاصل. ينبغي استعمال سطر واحد لكل مراقب تُسجل البيانات الخام وفق ترتيب الدخول يمكن توزيع البيانات على ملفات DAT مختلفة تحددها في الجدول 6 المعلمة Result(*j*). Filename(s)[1].	integer integer integer.....↵ integer integer integer.....↵ integer integer integer.....↵ .....

[1] انظر [1] في الجدول 6.