التوصيـــة ITU-R BT.1788

منهجية من أجل التقييم الذاتي لجودة الفيديو في التطبيقات متعددة الوسائط

(المسألة ITU-R 102/6)

(2007)

مجال التطبيق

ستسمح أنظمة الإذاعة الرقمية بتقديم تطبيقات إذاعة الوسائط المتعددة والمعطيات التي تشمل الإشارات الفيديوية والسمعية والصور الساكنة والنصوص والرسوم. وتوصّف هذه التوصية طرائق تقييم ذاتي غير تفاعلية لتقدير جودة الفيديو للتطبيقات متعددة الوسائط.

إن جمعية الاتصالات الراديوية التابعة للاتحاد الدولي للاتصالات (ITU)،

إذ تضع في اعتبارها

أ) أن أنظمة الإذاعة الرقمية يجري استعمالها في العديد من البلدان؛

ب) أن خدمات إذاعة الوسائط المتعددة والمعطيات، التي تشمل الإشارات الفيديوية والسمعية والصور الساكنة والنصوص والرسوم وما إلى ذلك، إما مطروحة للاستعمال أو يخطط لطرحها باستخدام أنظمة الإذاعة الرقمية؛

ج) أن الخدمات المتعددة الوسائط ستضم بنية تحتية إذاعية تتسم بالاستعمال المحتمل للمستقبلات الثابتة والمتنقلة ولمعدلات الأرتال المتغيّرة والثابتة وأنساق الصورة المختلفة والكودكات الفيديوية المتقدمة وخسارة رزم وما إلى ذلك؛

د) أنه سيكون ضرورياً توصيف متطلبات الأداء والتحقق من ملاءمة الحلول التقنية قيد الاعتبار لكل خدمة مع متطلبات الأداء لتلك الخدمة؛

هـ) أن مثل هذا التحقق سيضم في الأساس التقييم الذاتي لجودة الفيديو في ظل ظروف متحكَّم بها؛

و) أن منهجيات التقييم الذاتي المحددة في التوصية ITU-R BT.500 يمكن أيضاً استعمالها للتطبيقات متعددة الوسائط؛

ز) أن منهجيات التقييم الذاتي غير تلك المحددة في التوصية ITU-R BT.500 يمكن أن استعمالها أيضاً؛

ح) أن اعتماد طرائق موحدة من الأهمية بمكان في تبادل المعلومات بين شتى المختبرات،

توصـــي

1 بأنه ينبَغي استعمال طرائق الاختبار العامة، أي سلالم العلامات وشروط المشاهدة لتقييم جودة الصورة الموضحة في الملحق 1، من أجل تجارب المختبرات وللتقييمات التشغيلية في التطبيقات متعددة الوسائط كلما أمكن ذلك؛

2 بأن يُقدم في كل تقارير الاختبارات الوصفٌ الشامل لتشكيلات الاختبار وموادها المراقبين والطرائق؛

3 بأن تعالج المعطيات المجمّعة وفقاً للتقنيات الإحصائية المفصّلة في الملحق 2، من أجل تسهيل تبادل المعلومات بين المختبرات المختلفة.

الملاحظة 1 – يتعيّن على لجنة الدراسات 6 التابعة لقطاع الاتصالات الراديوية الاستمرار في متابعة تطوير مكتبة من المواد الفيديوية المناسبة للتقييم الذاتي لجودة الفيديو في التطبيقات متعددة الوسائط.

الملحق 1

وصف طرائق التقييم

1 المقدمة

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

تعتبر طرائق التقييم الذاتي الموحدة ضرورية لتوصيف متطلبات الأداء والتحقق من صحة الحلول التقنية قيد الاعتبار لكل تطبيق. وتعد المنهجيات الذاتية ضرورية 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 حجم 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 التقييم الذاتي الموصّفة في التوصية ITU-R BT.500 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.

وتوضح هذه التوصية طرائق تقييم ذاتي غير تفاعلية لتقدير جودة الفيديو للتطبيقات متعددة الوسائط. يمكن تطبيق هذه الطرائق لأغراض مختلفة تتضمن ولا تقتصر على: اختيار الخوارزميات وتصنيف أداء النظام السمعي المرئي وتقدير مستوى جودة الفيديو أثناء توصيل سمعي مرئي.

ترد المصطلحات والتعاريف المتعلقة بهذه التوصية في التذييل 3 للملحق 1.

2 الخصائص العامة

1.2 شروط المشاهدة

ترد شروط المشاهدة الموصى بها في الجدول 1. وينبغي أن تكون أبعاد ونمط العرض المستعمَل مناسبة 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 2 مثالاً عن سجلّ المعطيات 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 1

شروط المشاهدة الموصى باستعمالها في تقييم جودة الوسائط المتعددة

المعلمة	القيمة
مسافة المشاهدة [1]	مقيّدة: H 8-1 غير مقيّدة: حسب ما يفضّله المشاهد
ذروة نصوع الشاشة	70-250 cd/m2
نسبة نصوع الشاشة الخاملة إلى ذروة النصوع	≥ 0,05
نسبة نصوع الشاشة عندما تعرض مستوى السواد، فقط، في غرفة مظلمة تماماً، إلى النصوع المقابل لذروة البياض	≥ 0,1
نسبة نصوع الخلفية وراء مرقاب الصورة، إلى نصوع ذروة الصورة [2]	≥ 0,2
لونية الخلفية [3]	D_{65}
إضاءة الغرفة الخلفية [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 2

تشكيلة النظام المتعدد الوسائط الخاضع للاختبار

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

2.2 إشارات المصدر

تقدم إشارة المصدر الصورة المرجعية مباشرةً الدخل للنظام الخاضع للاختبار. وينبغي أن تكون جودة تتابعات المصدر عالية قدر الإمكان. كمبدأ توجيهي، ينبغي تسجيل الإشارة الفيديوية في الملفات متعددة الوسائط باستخدام 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 اختيار مواد الاختبار

يعتبر عدد مشاهد الاختبار وأنماطها من الأمور الحاسمة في تأويل نتائج التقييم الذاتي. 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. وعلاوة 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 جزءًا من مضمون 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 التذييل 1 للملحق 1.

4.2 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 استخدام مدى أكثر تقييداً، وكذلك من خلال فرض بعض الشروط التي قد تقع عند أطراف السلالم. ويمكن تمثيلها على أنها أمثلة، وتحديدها بأنها الأكثر صرامة (إرساء مباشر). أو توزيعها على كل العملية وعدم تحديدها بأنها الحالات الأكثر صرامة (إرساء غير مباشر). وينبغي استعمال مدى واسع الجودة إن أمكن ذلك.

5.2 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 مقيمين محترفين. ويجب فرز المراقبين قبل العملية وفقاً لحدة بصرهم العادية (أو المحولة إلى العادية) باستخدام مخطط اختبار "سنلين" (Snellen) أو "لندولت" (Landolt)، ووفقاً لرؤيتهم العادية للألوان بواسطة مخططات تُنتقى خصيصاً لهذا الغرض (مخطط اشيهارا (Ishihara)، مثلاً). يعتمد العدد المطلوب من المقيمين 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 إرشادات خاصة بالتقييم

يجب أن تقدم بعناية إلى المقيّمين، طريقة التقييم وأنماط الانحطاط أو عوامل الجودة المتوقع حدوثها، وسلم التقدير والتوقيت وما إلى ذلك. وتُستعمل تتابعات تدريبية تبيّن مدى الانحطاطات الواجب تقييمها وأنماطها مع مشاهد غير تلك المستعملة في الاختبار، لكن ذات حساسية مقاربة.

7.2 التصميم التجريبي

يترك لمنفّذ التجربة اختيار التصميم التجريبي لتحقيق الأهداف المحددة من حيث الكلفة والدقة. ويحبذ تكرار التجربة مرتين على الأقل (أي التكرار في نفس الظروف). 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 معبرة 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 طرائق التقييم

يمكن اختبار الأداء الفيديوي لأنظمة متعددة الوسائط باستعمال منهجيات التوصية ITU-R BT.500. وترد قائمة بالطرائق المختارة في القسم 1.3.

ويرد في القسم 2.3 وصف لمنهجية إضافية، تدعى SAMVIQ، تستغل خصائص المجال متعدد الوسائط ويمكن استعمالها في تقييم أداء الأنظمة متعددة الوسائط.

1.3 منهجيات التوصية ITU-R BT.500

ينبغي استعمال المنهجيات التالية للتوصية ITU-R BT.500 لتقييم جودة الفيديو في الأنظمة متعددة الوسائط:

- طريقة سلم الانحطاط ثنائية الحافز (DSIS) حسب وصفها في التوصية ITU-R BT.500، الملحق 1، القسم 4.
- طريقة سلم الجودة المستمرة ثنائية الحافز (DSCQS) حسب وصفها في التوصية ITU-R BT.500، الملحق 1، القسم 5.
- الطرائق وحيدة الحافز (SS) حسب وصفها في التوصية ITU-R BT.500، الملحق 1، القسم 1.6.
- طرائق مقارنة الحوافز (SC) حسب وصفها في التوصية ITU-R BT.500، الملحق 1، القسم 2.6.
- طريقة وحيدة الحافز لتقدير الجودة المستمرة (SSCQE) حسب وصفها في التوصية ITU-R BT.500، الملحق 1، القسم 3.6.

2.3 التقييم الذاتي لجودة الفيديو المتعدد الوسائط (SAMVIQ)

في هذه الطريقة، يتاح للمشاهد النفاذ إلى العديد من نسخ التتابع. وبعدما يصنّف المشاهد كل النسخ، يمكن النفاذ لمضمون التتابع التالي.

ويمكن للمشاهد اختيار النسخ المختلفة عشوائياً عبر سطح بيني لرسوم الحاسوب. ويمكن للمشاهد أن يوقف ويراجع ويعدّل نتيجة كل نسخة من نسخ التتابع حسب الرغبة. وتتضمن هذه الطريقة تتابع مرجعي صريح (أي غير معالج) علاوة 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 شبيهة بتلك التي تستعمل مرجعاً.

وتستعمل طريقة تقدير الجودة SAMVIQ سلم جودة مستمرة لتقدم قياساً للجودة الأصيلة لتتابعات الفيديو. ويحرّك كل مراقب مؤشر متزلق على سلم مستمر بتدريج من 0 إلى 100 ومزوّد بحواشي من 5 بنود للجودة مرتبة خطياً (ممتاز، جيد، مقبول، ضعيف، سيئ).

يُنفَّذ تقدير الجودة لكل مشهد على حدة (انظر الشكل 1) 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.

1.2.3 المرجع الصريح والمخفي والخوارزميات

تتضمن طريقة التقدير عموماً وسائل تثبيت للجودة من أجل إقرار النتائج. يتم مراعاة وسيلتين رفيعتين للجودة في طريقة SAMVIQ للأسباب التالية. أظهرت اختبارات عديدة أجريت أن الانحرافات المعيارية في النتائج 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 الجودة 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 درجة ممكنة (100) لمرجع صريح على نحوٍ مغاير تماماً للدرجة المقابلة للمرجع المخفي. وعلى نحوٍ لافتٍ، يظل الاختبار ممكناً في غياب مرجع، غير أن الانحراف المعياري يزداد بشكل كبير.

وتعد طريقة SAMVIQ مناسبة 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.

2.2.3 شروط الاختبار

إن اختيار المحتويات المتجانسة بالاتّباع الضمني لنفس القواعد التي تستعملها منهجيات أخرى توفر درجة تقييم عالمية (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 درجة للجودة مستقرة وموثوقة. وينبغي استعمال أجهزة فك شفرة–تشغيل مسجَّلة أو نسخة تستقبل على شاشة أخرى من خرجها للحفاظ على أداء العرض المناسب.

3.2.3 تنظيم الاختبار

أ) يُنفَّذ الاختبار لكل مشهد على حدة حسب الوصف في الشكل 1.

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

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

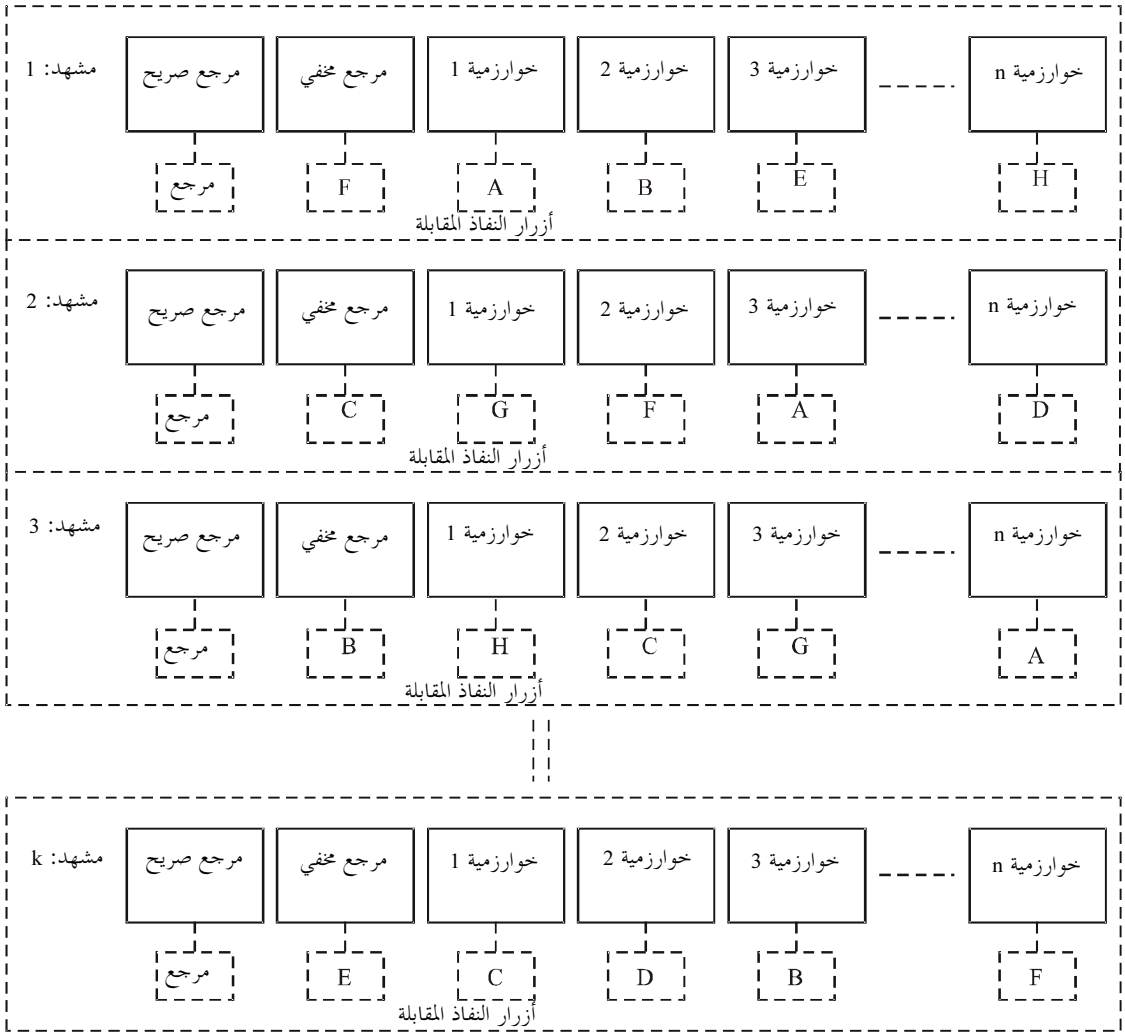
د) بالنسبة للمشاهدة الأولى، ينبغي أن يُستعرَض التتابع الراهن بالكامل قبل تقييمه بدرجة، أو يمكن تسجيل الدرجة والتوقف فوراً.

هـ) لاختبار المشهد التالي، لا بد من تقييم جميع تتابعات في المشهد الراهن بإعطائها الدرجة.

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

الشكل 1

مثال تنظيم اختبار لطريقة SAMVIQ



1788-01

تُنفَّذ طريقة SAMVIQ عبر البرمجيات. وفضلاً عن أزرار النفاذ المبيّنة في الشكل 1، تلزم أزرار "تشغيل" و"إيقاف" و"المشهد التالي" و"المشهد السابق" كي يتمكن المشاهد من إدارة عرض المشاهد المختلفة (أنظر التذييل 2 للملحق 1 مثلاً). وعندما يعطي المشاهد درجة معينة ينبغي إظهارها تحت زر النفاذ المقابل لذلك المشهد. وعند إعطاء درجات لجميع النسخ المختلفة لتتابع ما، يظل متاحاً للمشاهد المقارنة بين الدرجات وتعديلها إذا لزم الأمر. وليس ضرورياً مراجعة التتابع الراهن برمته لأن فروقاً كبيرة يكون قد تم تسليط الضوء عليها في المشاهدة الأولى.

الملحق 2

عرض وتحليل المعطيات

1 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 3.

الجدول 3

معلومات موجزة عن الاختبار

اسم الطريقة	
تكنولوجيا العرض	
الاسم المرجعي للعرض	
مستوى ذروة النصوع (cd/m^2)	
مستوى النصوع الأسود (cd/m^2)	
ضبط مستوى اللون الأسود: PLUGE (العتبة المحسوسة للمسافة بين سوية الأسود وفوق الأسود = 8). وفيما عدا ذلك يبيّن قيمة العتبة	
مستوى نصوع الخلفية (cd/m^2)	
الإضاءة (lux)	
مسافة المشاهدة: - غير مقيّدة: أمام شاشة العرض - مقيّدة: nH	
أبعاد 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: D65 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 مهمة لتصميم اختبار أو تفسير نتائجه. ويُقترح الإبلاغ عن الخصائص المكانية–الزمانية حسب وصفها في التذييل 1 للملحق 1. وينبغي النظر في هذه المعلومات ضمن مجموعة تتابعات الاختبار في مكتبة مواد الفيديو المناسبة للتقييم الذاتي لنوعية الفيديو في لتطبيقات متعددة الوسائط.

2 طرائق التحليل

طرائق التحليل هي تلك الموصوفة في التوصية ITU-R BT.500، الملحق 2، القسم 2.

3 فرز المراقبين

بالنسبة للطرائق المدرجة في الملحق 1، القسم 1.3، يرد وصف إجراءات الفرز في التوصية ITU-R BT.500، الملحق 2، القسم 3.2.

ويرد وصف الفرز خاصة SAMVIQ في القسم التالي. بيد أن هذا الإجراء يمكن استعماله من أجل طرائق SS وDSIS وDSCQS. حيث يعد هذا الإجراء أبسط في تنفيذه من الإجراء المقابل في التوصية ITU-R BT.500 لتلك الطرائق.

1.3 إجراء الفرز خاصة الطريقة SAMVIQ

يجب أن يكون لكل مراقب طريقة مستقرة ومتماسكة للتصويت العادل على انحطاط الجودة بالنسبة إلى كل مشهد وخوارزمية. وتؤكد معايير النبذ مستوى معقول من التوافق بين درجات مراقب ما وفق متوسط درجات جميع المراقبين لعملية اختبار معيّنة. وفي طريقة SAMVIQ كما في طريقة DSQCS، يمكن وضع جميع الخوارزميات (المرجع المخفي أو الضمني، علامة التثبيت المنخفضة، التتابعات المشفرة) في الاعتبار. ويستند معيار القرار إلى ارتباط بين الدرجات الإفرادية ومتوسط الدرجات المقابل من جميع مراقبي الاختبار.

2.3 ارتباط بيرسون

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

والهدف الرئيسي هو التحقق بطريقة بسيطة مما إذا كانت درجات أحد المراقبين متسقة مع متوسط الدرجات 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 الدرجات، وعلى العكس من ذلك تقل تخالفات الارتباط بين المراقبين.

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

حيث:

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

y_i: الدرجة الإفرادية لمراقب واحد للمعاملات الثلاثة ذاتها

n: (عدد الخوارزميات) × (عدد المشاهد)

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

3.3 ارتباط رتبة سبيرمان

يمكن تطبيق ارتباط رتبة سبيرمان حتى لو لم تتحقق فرضية أن العلاقة بين سلم الجودة ومدى درجات المراقبين خطية[1]:

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

حيث:

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

y_i: الدرجة الإفرادية لمراقب واحد من أجل المعاملات الثلاثة ذاتها

n: (عدد الخوارزميات) × (عدد المشاهد)

$R(x_i$ أو $y_i)$: ترتيب الرتبة

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

4.3 معايير النبذ النهائي لاستبعاد مراقب اختبار

يُنفَّذ ارتباطا رتبة سبيرمان وبيرسون لاستبعاد مراقب(ين) 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)]

[1] عموماً، تعتبر نتائج ارتباط بيرسون قريبة 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 لارتباطي جميع مراقبي الاختبار

عتبة الارتباط القصوى (MCT) = 0,85.

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

التذييل 1

للملحق 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 للمشاهد الأكثر تعقيداً من الوجهة المكانية. وليست الغاية منه أن يكون مقياساً للانتروبي 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 عناصر الصورة (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 بحيث تمثل مضمون المعلومات المكانية للمشهد. يمكن تمثيل هذه العملية في صورة معادلة كما يلي:

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

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 للانتروبي ولا مصحوباً بمعلومات كتلك المعرفة في نظرية الاتصالات.

يُحسب مقياس المعلومات الزمنية *TI* باعتباره الحد الأقصى 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:

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

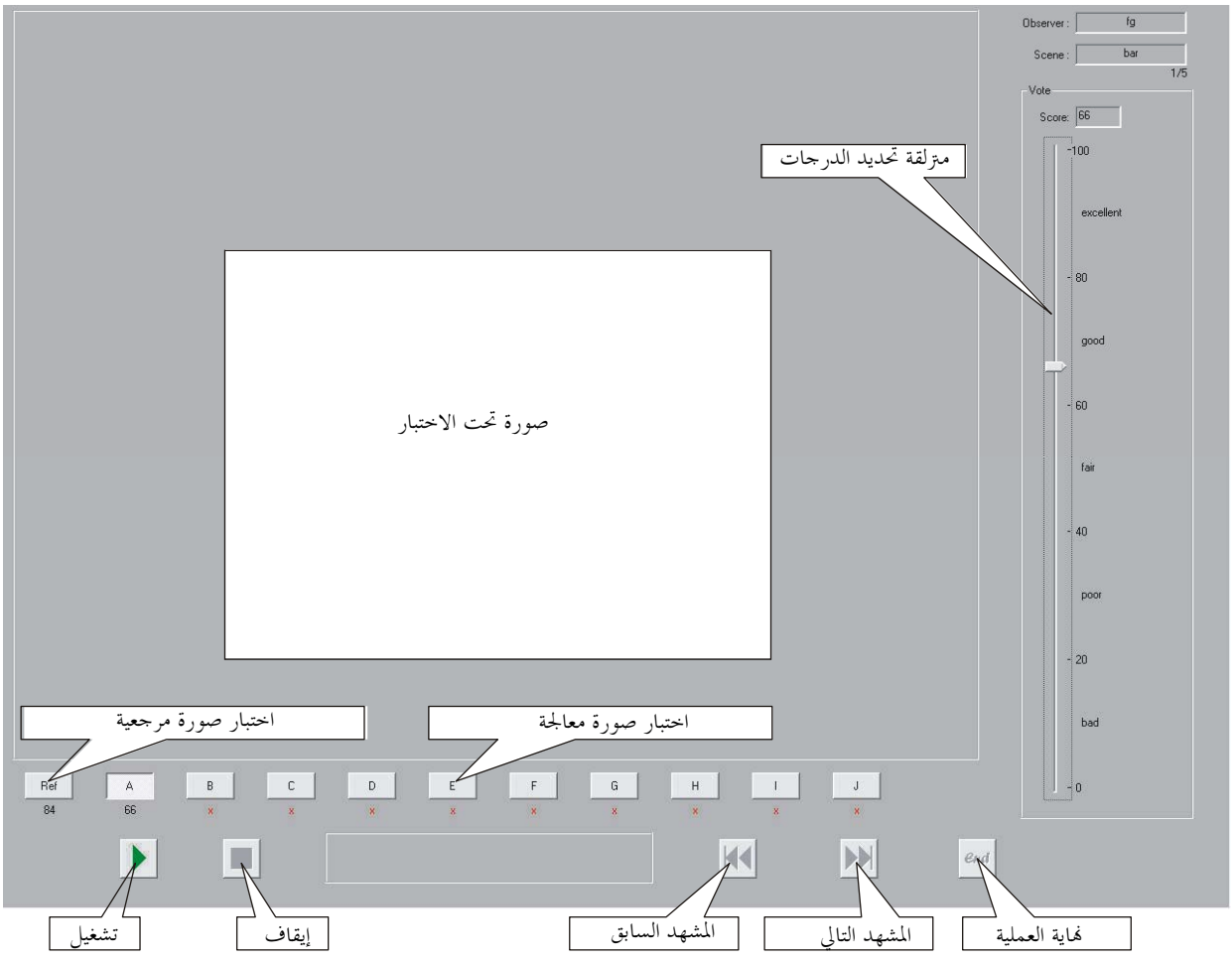
حيث $F_n(i,j)$ هو وحدة الصورة في الصف رقم i والعمود رقم j من الرتل رقم n في الزمن.

الملاحظة 1 – بالنسبة للمشاهد التي تحوي مقاطع للمشهد، يمكن إعطاء قيمتين: واحدة يُدرج فيها مقطع المشهد في مقياس المعلومات الزمانية، والأخرى حيث يُستبعد من القياس.

التذييل 2

للملحق 1

مثال عن سطح بيني لتقييم SAMVIQ



1788-02

التذييل 3

للملحق 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 (نسق صورة معرَّف في التوصية H.261 لهاتف الفيديو: 352 خط × 288 بيكسل) *(Common intermediate format (picture format defined in Recommendation H.261 for video phone: 352 lines × 288 pixels))*
CRT	صمام الأشعة المهبطية *(Cathode ray tube)*
DSCQS	طريقة سلم الجودة المستمرة ثنائية الحافز *(Double stimulus using a continuous quality scale method)*
DSIS	طريقة سلم الانحطاط ثنائية الحافز *(Double stimulus using an impairment scale method)*
LCD	شاشة عرض بالبلورات السائلة *(Liquid crystal display)*
MOS	متوسط الدرجة المعطاة *(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 (نسق صورة معرَّف في التوصية H.261 لهاتف الفيديو: 176 خط × 144 بيكسل) *(Quarter CIF (picture format defined in Recommendation H.261 for video phone: 176 lines × 144 pixels))*
SAMVIQ	تقييم ذاتي لجودة فيديو متعدد 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)	مضمون صوتي مرئي *(Audiovisual content)*
S/N	نسبة الإشارة إلى الضوضاء *(Signal-to-noise ratio)*
SI	معلومات مكانية *Spatial information*
SIF	نسق متوسط معياري [أنساق الصورة المعرّفة في المعيار ISO 11172 (MPEG-1): 352 خط × 288 بيكسل × 25 رتل/ث و352 خط × 240 بيكسل × 30 رتل/ث] *(Standard intermediate format [picture formats 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	الطريقة وحيدة الحافز *(Single stimulus method)*
SSCQE	طريقة وحيدة الحافز لتقدير الجودة المستمرة *(Single stimulus using a continuous quality evaluation method)*
std	انحراف معياري *(Standard deviation)*
TI	معلومات زمانية *(Temporal information)*
TP	عرض اختبار *(Test presentation)*
TS	عملية اختبار *(Test session)*
VTR	مسجّلة شريط فيديو *(Video tape recorder)*