

الاتحاد الدولي للاتصالات

ITU-R

قطاع الاتصالات الراديوية في الاتحاد الدولي للاتصالات

التوصية ITU-R BT.1907-0
(2012/01)

تقنيات القياس الموضوعي للجودة الفيديوية
المدركة للتطبيقات الإذاعية التي تستعمل
التلفزيون عالي الوضوح في وجود إشارة
مرجعية كاملة

السلسلة BT

الخدمة الإذاعية (التلفزيونية)



تمهيد

يضطلع قطاع الاتصالات الراديوية بدور يتمثل في تأمين الترشيد والإنصاف والفعالية والاقتصاد في استعمال طيف الترددات الراديوية في جميع خدمات الاتصالات الراديوية، بما فيها الخدمات الساتلية، وإجراء دراسات دون تحديد مدى 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 الكهروتقنية الدولية (ITU-T/ITU-R/ISO/IEC) والمشار إليها في القرار ITU-R 1. وترد الاستمارات التي ينبغي لحاملي البراءات استعمالها لتقديم بيان عن البراءات أو للتصريح عن منح رخص في الموقع الإلكتروني <http://www.itu.int/ITU-R/go/patents/en> حيث يمكن أيضاً الاطلاع على المبادئ التوجيهية الخاصة بتطبيق سياسة البراءات المشتركة وعلى قاعدة بيانات قطاع الاتصالات الراديوية التي تتضمن معلومات عن البراءات.

سلاسل توصيات قطاع الاتصالات الراديوية

(يمكن الاطلاع عليها أيضاً في الموقع الإلكتروني <http://www.itu.int/publ/R-REC/en>)

العنوان	السلسلة
البث الساتلي	BO
التسجيل من أجل الإنتاج والأرشفة والعرض؛ الأفلام التلفزيونية	BR
الخدمة الإذاعية (الصوتية)	BS
الخدمة الإذاعية (التلفزيونية)	BT
الخدمة الثابتة	F
الخدمة المتنقلة وخدمة التحديد الراديوي للموقع وخدمة الهواة والخدمات الساتلية ذات الصلة	M
انتشار الموجات الراديوية	P
علم الفلك الراديوي	RA
أنظمة الاستشعار عن بعد	RS
الخدمة الثابتة الساتلية	S
التطبيقات الفضائية والأرصاد الجوية	SA
تقاسم الترددات والتنسيق بين أنظمة الخدمة الثابتة الساتلية والخدمة الثابتة	SF
إدارة الطيف	SM
التجميع الساتلي للأخبار	SNG
إرسالات الترددات المعيارية وإشارات التوقيت	TF
المفردات والمواضيع ذات الصلة	V

ملاحظة: تمت الموافقة على النسخة الإنكليزية لهذه التوصية الصادرة عن قطاع الاتصالات الراديوية بموجب الإجراء الموضح في القرار

ITU-R 1

النشر الإلكتروني

جنيف، 2020

© ITU 2020

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

التوصية ITU-R BT.1907-0*

تقنيات القياس الموضوعي للجودة الفيديوية المدركة للتطبيقات الإذاعية التي تستعمل التلفزيون عالي الوضوح في وجود إشارة مرجعية كاملة

(2012)

مجال التطبيق

تحدد هذه التوصية طرائق لتقدير الجودة الفيديوية المدركة للتطبيقات الإذاعية التي تستعمل التلفزيون عالي الوضوح (HDTV) في وجود إشارة مرجعية كاملة.

إن جمعية الاتصالات الراديوية للاتحاد الدولي للاتصالات،

إذ تضع في اعتبارها

- أ) أن القدرة على قياس جودة الإذاعة الفيديوية أوتوماتياً طالما أقر بأنها وسيلة قيمة تنفيذ منها دوائر صناعة الاتصالات؛
- ب) أن التوصية ITU-R BT.1683 تصف طرائق موضوعية لقياس الجودة الفيديوية المدركة لتلفزيون إذاعة رقمية ذات وضوح معياري في وجود إشارة مرجعية كاملة؛
- ج) أن التوصية ITU-R BT.709 تصف قيم معلمات لمعايير التلفزيون عالي الوضوح من أجل إنتاج البرامج وتبادلها على الصعيد الدولي، وأن التوصية ITU-R BT.500 تصف طرائق تقييم ذاتية لجودة الصورة في التلفزيون عالي الوضوح؛
- د) أن التلفزيون عالي الوضوح يشجع استخدامه على نطاق واسع في مجال الإذاعة؛
- هـ) أن لجنة الدراسات 9 في قطاع تقييس الاتصالات، بناء على نتائج تقرير عن التلفزيون عالي الوضوح أرسله فريق الخبراء المعني بجودة الفيديو (VQEG)، وضعت التوصية ITU-T J.341، التي حددت موضوعياً قياس الجودة الفيديوية المدركة للتلفزيون عالي الوضوح في وجود إشارة مرجعية كاملة؛
- و) أن القياس الموضوعي للجودة الفيديوية المدركة للتلفزيون عالي الوضوح من شأنه أن يكمل طرائق التقييم الذاتية،

توصي

باستخدام نموذج الجودة الفيديوية الموضوعي الوارد في الملحقات 1 و 2 و 3 من أجل القياس الموضوعي للجودة الفيديوية المدركة للتطبيقات الإذاعية التي تستعمل التلفزيون عالي الوضوح في وجود إشارة مرجعية كاملة.

* أجرت لجنة الدراسات 6 تعديلات صياغية على هذه التوصية في فبراير 2020 طبقاً للقرار ITU-R 1.

الملحق 1

1 مقدمة

توفر هذه التوصية طريقة لقياس الجودة الفيديوية المدركة لاستخدامها في التطبيقات غير التفاعلية للتلفزيون عالي الوضوح (HDTV) عندما يمكن استخدام طريقة قياس المرجع الكامل (FR). وقد وُضع النموذج لتقدير درجات الجودة ذاتياً.

ويمكن استخدام طريقة قياس المرجع الكامل عندما تتوفر بسهولة الإشارة الفيديوية المرجعية دون تشوه عند نقطة القياس، كما قد تكون الحالة في القياسات المجرأة على فرادى المعدات أو على سلسلة في المختبر أو في بيئة مغلقة من قبيل محطة إذاعة تلفزيونية. وتحتوي الطريقة على كل من تقديرات المعايير والتقديرات الموضوعية للجودة الفيديوية ، على السواء.

وقد تضمنت مواد اختبار التحقق، على السواء، كلاً من أحوال انحطاط التشفير H.264 و MPEG-2 ومختلف أحوال أخطاء الإرسال (من قبيل أخطاء البتات والرمز المفقودة). ويمكن استخدام النموذج في هذه التوصية لرصد جودة الشبكات المنشورة لضمان جاهزيتها التشغيلية. وقد تشمل الآثار المرئية أشكالاً من الانحطاط من حيث البعد المكاني والزمني على السواء. ويمكن استخدام النموذج في هذه التوصية أيضاً لاختبار الأنظمة الفيديوية في المختبر. وعندما يستخدم لمقارنة نظامي فيديو مختلفين، من المستصوب استخدام الطريقة الكمية (كتلك الواردة في التوصية ITU-T J.149) لتحديد دقة النموذج لذلك السياق بعينه.

وتُعتبر هذه التوصية ملائمة للخدمات الإذاعية المرسله بمعدل ما بين 1 و 30 Mbit/s. وقد أخذت معدلات الاستبانة والأرتال التالية في الاعتبار في اختبار التحقق:

1080/59.94/I	–
1080/25/P	–
1080/50/I	–
1080/29.97/P	–

وقد رُوعيت الشروط التالية في اختبار التحقق بالنسبة لكل استبانة:

عوامل الاختبار
الاستبانة الفيديوية: 1080 × 1920 مشدرة وتدريجية
معدلات الأرتال الفيديوية 29,97 و 25 رتلاً في الثانية
معدل البتات الفيديوية 1 إلى 30 Mbit/s
تجمد الرتل الزمني (التوقف والتخطي) 2 ثانية كحد أقصى
أخطاء الإرسال مع فقدان الرزم
تحويل SRC من 1080 إلى 720/P و ضغط وإرسال وإزالة ضغط ثم عودة التحويل إلى 1080.
تكنولوجيات التشفير
H.264/AVC (MPEG-4 Part 10)
MPEG-2

ويلاحظ أن 720/P قد اعتبرت في خطة اختبار التحقق كجزء من شرط الاختبار في دارة مرجعية مفترضة (HRC). ولأن من الشائع التوسع في 720/P في الوقت الراهن كجزء من عملية العرض، رئي أن HRCs 720/P تلائم هذا النسق على نحو أفضل.

1.1 التطبيق

تشمل التطبيقات لنموذج التقدير الموصوف في هذه التوصية العمليات التالية، دون أن تقتصر عليها:

- (1) رصد الجودة عند المصدر، أثناء الخدمة وفي الوقت الفعلي قدر الإمكان؛
- (2) رصد الجودة في المقصد النهائي عندما تتوفر نسخة من المصدر في نقطة القياس؛
- (3) قياس الجودة لرصد نظام تخزين أو إرسال يستخدم تقنيات الضغط وإزالة الضغط الفيديوي، إما مرة واحدة أو في تسلسل من هذه التقنيات؛
- (4) اختبار الأنظمة الفيديوية في المختبر.

2.1 التحديدات

لا يمكن استخدام نموذج تقدير الجودة الفيديوية الموصوف في هذه التوصية بديلاً عن الاختبار الذاتي. وتقع عادة قيم الترابط بين اختبارين ذاتيين تم تصميمهما وتنفيذهما بعناية (أي في مختبرين مختلفين) في مدى 0,95 إلى 0,98. فإذا استُخدمت هذه التوصية لإقامة مقارنات بين الأنظمة الفيديوية (أي مقارنة نظامي كودك)، فإن من المستصوب استخدام طريقة كمية لتحديد دقة النموذج لذلك السياق بعينه.

وفي حال جمود الرتل، تتوفر في شروط الاختبار عموماً فترات من جمود الرتل لأقل من ثانيتين. ولم يتم التحقق من النموذج في هذه التوصية لقياس جودة الفيديو في ظرف إعادة تخزين احتياطي (أي الفيديو الذي يتزايد فيه التأخر أو التجمد دون التخطي). ولم يختبر النموذج في معدلات أرتال أخرى غير تلك المستخدمة في أنظمة التلفزيون (أي 29,97 رتلاً في الثانية و 25 رتلاً في الثانية، في أسلوب مشدر أو تدريجي).

ومن الجدير بالملاحظة أنه في حالة تكنولوجيات تشفير وإرسال جديدة تفضي إلى تشوهات لم تدرج في هذا التقييم، فإن النموذج الموضوعي قد يتمخض عن نتائج مغلوطة. وفي هذه الحالة يستدعي الأمر إجراء تقييم ذاتي.

2 المراجع

لا شيء.

3 التعاريف

1.3 مصطلحات معرّفة في أماكن أخرى

تستخدم هذه التوصية المصطلحات التالية المعرّفة في أماكن أخرى:

1.1.3 تقييم ذاتي (صورة): تحديد جودة أو تشوه برنامج من قبيل صور تُعرض على فريق من الأشخاص المقيمين في جلسات مشاهدة.

2.1.3 قياس إدراكي موضوعي (صورة): قياس أداء سلسلة برنامج باستخدام صور شبيهة بالبرنامج وطرائق قياس (أدوات) موضوعية للتوصل إلى دليل من شأنه تقريب سوية الدرجة التي يمكن الحصول عليها من اختبار تقييم ذاتي.

3.1.3 داعية: منظمة أو شركة تقترح نموذج جودة فيديوية لاختبار التحقق واحتمال الإدراج في واحدة من توصيات الاتحاد.

2.3 مصطلحات معرفة في هذه التوصية

تعرف هذه التوصية المصطلحات التالية:

1.2.3 معدل الأرتال: عدد الأرتال الفريدة (أي مجموع الأرتال - الأرتال المتكررة) في الثانية.

2.2.3 أخطاء الإرسال المحاكاة: أخطاء تُفرض على تدفق بتات الفيديو الرقمية في بيئة تخضع لدرجة عالية من التحكم. ومن أمثلة ذلك، محاكاة معدلات فقدان الرزم ومحاكاة أخطاء البتات. وتعرف على نحو واضح المعلمات المستخدمة للتحكم في أخطاء الإرسال المحاكاة.

3.2.3 أخطاء الإرسال: أي خطأ يُفرض على الإرسال الفيديوي. ومن أمثلة ذلك أنماط أخطاء الإرسال المحاكاة وظروف الشبكة الحية.

4 مختصرات

تستخدم هذه التوصية المختصرات التالية:

ACR	درجة الفئة المطلقة (انظر التوصية ITU-R BT.500) (<i>Absolute Category Rating (see ITU-R BT.500)</i>)
ACR-HR	درجة الفئة المطلقة مع مرجع خفي (انظر التوصية ITU-T P.910) (<i>Absolute Category Rating with Hidden Reference</i>)
AVI	تشدير سمعي فيديوي (<i>Audio Video Interleave</i>)
DMOS	متوسط فارق درجة الرأي (<i>Difference Mean Opinion Score</i>)
FR	مرجع كامل (<i>Full Reference</i>)
FRTV	تلفزيون مرجع كامل (<i>Full Reference TeleVision</i>)
HRC	دائرة مرجعية مفترضة (<i>Hypothetical Reference Circuit</i>)
ILG	مجموعة المختبرات المستقلة لفريق خبراء الجودة الفيديوية (VQEG) (<i>VQEG's Independent Laboratory Group</i>)
MOS	متوسط درجة الرأي (<i>Mean Opinion Score</i>)
MOSp	متوسط درجة الرأي، متنبأ به (<i>Mean Opinion Score, predicted</i>)
NR	لا (أو صفر) مرجع (<i>No (or Zero) Reference</i>)
PSNR	ذروة معدل الإشارة إلى الضوضاء (<i>Peak Signal-to-Noise Ratio</i>)
PVS	تتابع فيديو معالج (<i>Processed Video Sequence</i>)
RMSE	جذر متوسط الخطأ التربيعي (<i>Root Mean Square Error</i>)
RR	مرجع مخفّف (<i>Reduced Reference</i>)
SFR	معدل رتل المصدر (<i>Source Frame Rate</i>)
SRC	قناة أو دائرة مصدر المرجع (<i>Source Reference Channel or Circuit</i>)
VQEG	فريق خبراء الجودة الفيديوية (<i>Video Quality Experts Group</i>)
YUV	حيّز اللون ونسق الملف (<i>Colour Space and file format</i>)

5 اصطلاحات

لا شيء.

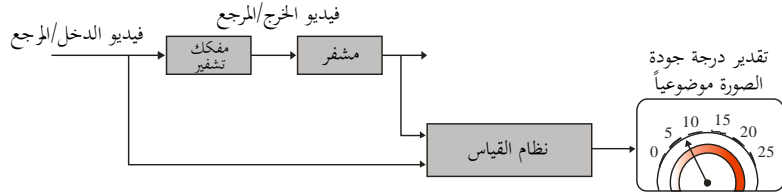
6 وصف منهجية المرجع الكامل

تعتمد طريقة القياس المزدوجة النهائية ذات المرجع الكامل، للقياس الموضوعي لجودة الفيديو المدركة، إلى تقييم أداء الأنظمة بإجراء مقارنة بين الدخل غير المشوه، أو الإشارة الفيديوية المرجعية عند دخل النظام، وبين الإشارة المنحطة عند خرج النظام (الشكل 1).

ويوضح الشكل 1 مثال تطبيق لطريقة المرجع الكامل لاختبار كودك ما في المختبر.

الشكل 1

تطبيق طريقة قياس الجودة المدركة ذات المرجع الكامل لاختبار كودك ما في المختبر



BT.1907-01

وقد تستدعي المقارنة بين إشارات الدخل والخرج عملية تسوية زمانية أو مكانية، والغرض من هذه الأخيرة هو تعويض أي زحزحة أو 'تقليم' عمودياً أو أفقياً. وقد تتطلب أيضاً التصحيح لتسوية أي تخالفات أو فروق في الكسب في كلتا قناتي الضوء واللون. ومن ثم يتم حساب تقدير درجة جودة الصورة موضوعياً، عموماً بتطبيق نموذج إدراكي لرؤية الإنسان.

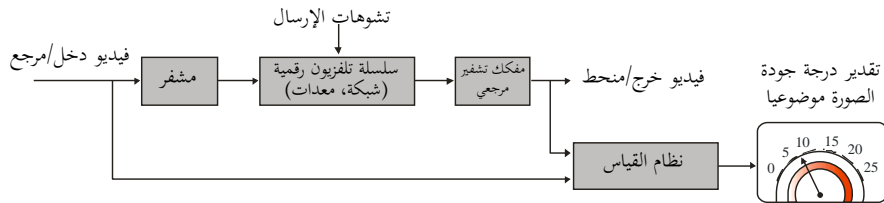
وتعرف عملية التسوية وتعديل الكسب باسم 'التسجيل'. وهذه العملية مطلوبة لأن غالبية طرائق المرجع الكامل تقارن في واقع الأمر الصور المرجعية والصور المعالجة على أساس 'بكسل تلو الآخر'. وتشمل مقاسات جودة الفيديو الموصوفة في الملحق 2 طرائق التسجيل هذه.

وبما أن قياسات الجودة الفيديوية تستند عموماً إلى تقريبات لاستجابات الإنسان البصرية وليس إلى قياس 'تشوهات' تشفير محددة، فهي من حيث المبدأ صالحة للأنظمة التماثلية وللأنظمة الرقمية على حد سواء. وهي كذلك صالحة من حيث المبدأ من أجل السلاسل التي تختلط فيها الأنظمة التماثلية والرقمية، أو حيث تكون أنظمة الضغط الرقمي سلسلية.

ويوضح الشكل 2 مثال تطبيق لطريقة المرجع الكامل لاختبار سلسلة إرسال.

الشكل 2

تطبيق طريقة قياس الجودة المدركة ذات المرجع الكامل لاختبار سلسلة إرسال



BT.1907-02

في هذه الحالة، تتم تغذية مفكك مشفر مرجعي من نقاط شتى في سلسلة الإرسال، حيث يمكن مثلاً وضع مفكك التشفير في نقطة في الشبكة، كما في الشكل 2، أو مباشرة عند خرج المشفر، كما في الشكل 1. فإذا كانت سلسلة الإرسال الرقمي شفافة، فإن قياس درجة جودة الصورة موضوعياً عند المصدر يساوي القياس في أي نقطة لاحقة في السلسلة.

ومن المقبول بصفة عامة أن طريقة المرجع الكامل توفر أعلى درجة من الدقة لقياسات جودة الصورة المدركة. وقد برهنت الطريقة على أنها تنطوي على إمكانية عالية من الترابط مع التقييمات الذاتية المجراة وفقاً لطرائق درجة الفئة المطلقة مع مرجع خفي (ACR-HR) المحددة في التوصية ITU-T P.910.

7 استنتاجات فريق خبراء الجودة الفيديوية

تجري قياسات الجودة الفيديوية الإدراكية في إطار فريق غير رسمي، يدعى فريق خبراء الجودة الفيديوية (VQEG)، يرفع تقاريره إلى لجنتي الدراسات 9 و12 في قطاع تقييس الاتصالات وإلى لجنة الدراسات 6 في قطاع الاتصالات الراديوية. وقد تمخض اختبار المرحلة الأولى للتلفزيون عالي الوضوح الذي استكماله مؤخراً فريق الخبراء VQEG عن تقييم أداء خوارزميات قياس الجودة الفيديوية المدركة ذات المرجع الكامل المقترحة.

والإحصاءات الواردة أدناه مستقاة من التقرير النهائي للفريق VQEG عن التلفزيون عالي الوضوح. وجدير بالملاحظة أن صلب التقرير المذكور يشمل قياسات أخرى منها علاقة ترابط Pearson وخطأ جذر متوسط التربيع (RMSE) المحسوبة في فرادى التجارب، وفواصل الثقة واختبار الدلالة الإحصائية في فرادى التجارب، وتحليل مجموعات فرعية من البيانات التي تشتمل على تشوهات محددة (مثل التشفير بموجب H.264 فقط)، ومخططات التشتت، ومعاملات التوافق.

التحليل الأولي

يوجز الجدول 1 أدناه أداء نموذج المرجع الكامل. ويتم حساب ذروة معدل الإشارة إلى الضوضاء (PSNR) بموجب التوصية ITU-T J.340 وهو وارد في هذا التحليل لأغراض المقارنة. وتحدد "المجموعة الكبرى لخطأ RMSE" القياس الأولي لخطأ جذر متوسط التربيع (RMSE) المحسوب على المجموعة الكبرى الكلية (أي أن نتائج التجارب الست الأولى تدون على مقياس وحيد). ويحدد "حاصل مجموعة أفضل أداء" عدد التجارب (0 إلى 6) التي كان هذا النموذج بالنسبة لها إما النموذج الأفضل أداءً أو المكافئ إحصائياً للنموذج الأفضل أداءً. ويحدد "حاصل ما هو أفضل من PSNR" عدد التجارب (0 إلى 6) التي كان هذا النموذج بالنسبة لها أفضل إحصائياً من PSNR. وتدرج المجموعة التي هي "أفضل من المجموعة الكبرى PSNR" ما إذا كان كل نموذج أفضل إحصائياً من PSNR في المجموعة الكبرى الكلية. وتحدد "علاقة الترابط في المجموعة الكبرى" علاقة ترابط Pearson المحسوبة في المجموعة الكبرى الكلية.

الجدول 1

جودة 'سويسرية'	PSNR	قياس
0,56	0,71	المجموعة الكبرى RMSE
5	1	حاصل أفضل مجموعة أداء
4	–	حاصل ما هو أفضل من PSNR
نعم	–	ما هو أفضل من المجموعة الكبرى PSNR
0,87	0,78	علاقة ترابط المجموعة الكبرى

الملحق 2

وصف النموذج

ملاحظة من المحرر: تشكل شفرة المصدر الواجب إدراجها في هذا القسم جزءاً إلزامياً من هذه التوصية، ويمكن الاطلاع عليها في العنوان: <http://ifatemp.itu.int/t/2009/sg9/exchange/q2/>.

النموذج بصفة عامة

يتنبأ النموذج بالجودة الفيديوية كما يدركها المشاركون في التجربة. ويستخدم نموذج التنبؤ عملية نمذجة إدراكية نفسانية بصرية لمحاكاة الإدراك الشخصي.

وكنهج كامل المرجع، يقارن النموذج الدخل أو المرجع الفيديوي ذا الجودة العالية مع التابع الفيديو المنحط المرتبط به قيد الاختبار. وهذه العملية مصورة بيانياً في الشكل 3.

ويقوم تقدير الدرجات على الخطوات التالية:

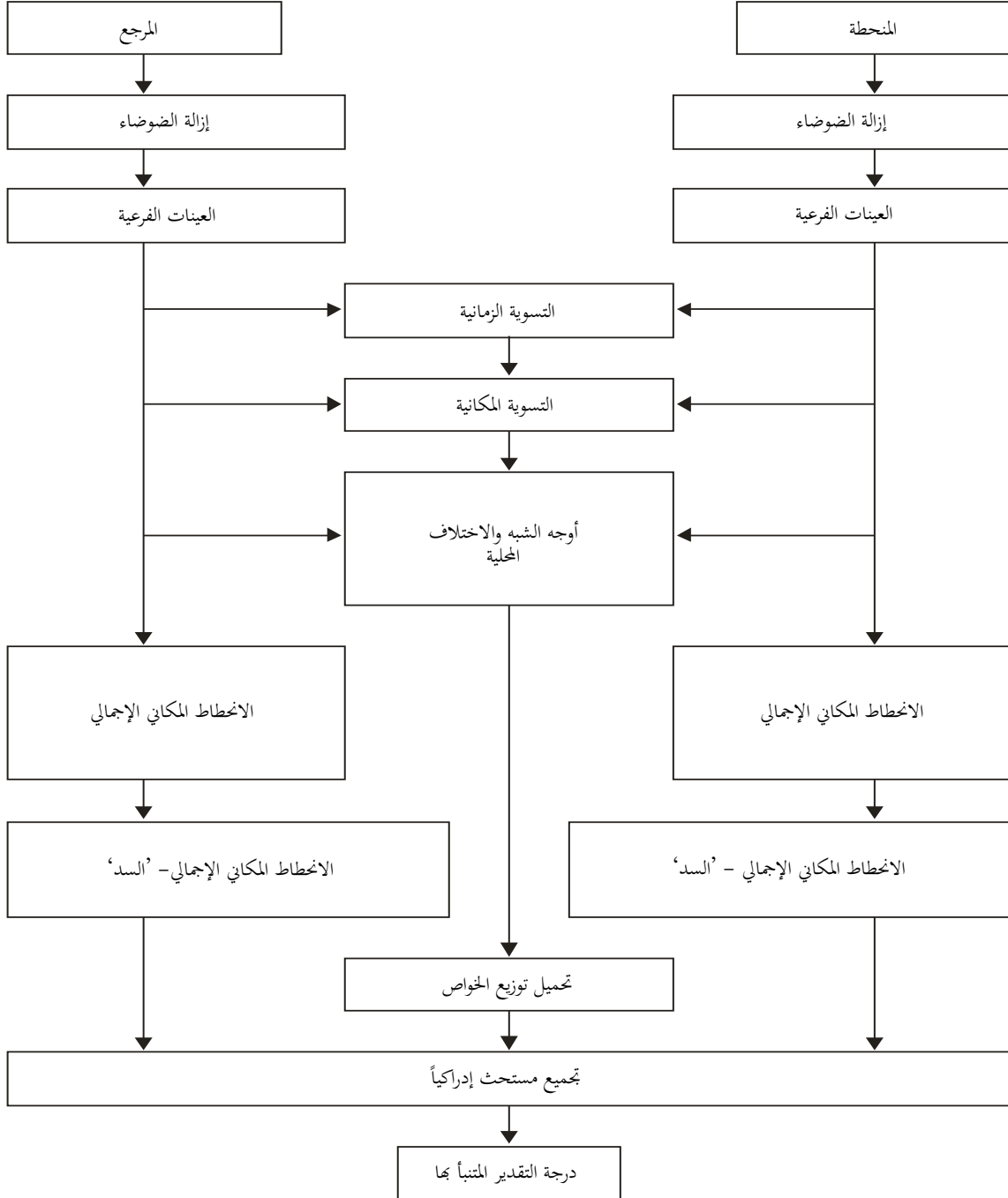
- (1) أولاً، تعالج مسبقاً التتابعات الفيديوية. وعلى وجه التحديد، تزال الضوضاء بترشيح الأرتال وتقسيمها فرعياً إلى عينات.
- (2) تتم عملية تسوية زمانية للأرتال بين التابع الفيديوي المرجع والتابع المعالج.
- (3) تتم عملية تسوية مكانية للأرتال بين التابع الفيديوي المعالج والتابع الفيديوي المرجعي المقابل له.
- (4) تحسب خواص الجودة المكانية المحلية: قياس أوجه الشبه وأوجه الاختلاف المحلية، انطلاقاً من الإدراك البصري.
- (5) يجري تحليل لتوزيع ملامح الشبه والاختلاف المحلية.
- (6) يقاس الانحطاط المكاني الإجمالي باستخدام عنصر 'السد'.
- (7) يقاس الانحطاط الزماني الإجمالي باستخدام عنصر 'الاهتزاز'. ويقاس الاهتزاز من خلال تقييم شدة الحركة المحلية والإجمالية وزمن مشاهدة الرتل.
- (8) تقدر درجة الجودة استناداً إلى تجميع غير خطي للخواص المذكورة أعلاه.
- (9) تجنباً لسوء التنبؤ في حالة سوء تسوية مكانية كبير نسبياً بين التابع الفيديوي المرجعي والتابع الفيديوي المعالج، تحسب الخطوات المتقدمة أعلاه من أجل 3 تسويات مكانية أفقية وعمودية مختلفة للتابع الفيديوي، وتكون الدرجة القصوى المنتبأ بها بين كل المواقع المكانية هي درجة الجودة المقدرة النهائية.

ويرد شرح فرادى هذه الخطوات بمزيد من التفصيل في الفقرات من 1.2 إلى 9.2 شاملة. وتنطوي الفقرة 10.2 على 'أرشيف مبيّت' له شفرة مصدر C++ تغطي الأجزاء والوظائف الأساسية من أجل وصف للنموذج يمثل للتنفيذ. وتشير أسماء الوظيفة C++ المذكورة في الفقرات من 1.2 إلى 9.2 إلى شفرة المصدر المرجعي هذا (مثلاً، تشير الفقرة 2.2 إلى:

(CFrameAnalysisFullRef::ContentTimeAlignment

الشكل 3

مخطط تدفق بياني لخطوات المعالجة في النموذج. في أعلى الشكل، الدخل هو المرجع والتتابعات الفيديوية المنحطة (أو المعالجة). وتفصي مختلف خطوات المعالجة إلى خرج النموذج الرئيسي، وتقدير الدرجة المنتبأ بها في أسفل الشكل



1.2 المعالجة المسبقة

يتم ترشيح كل رتل، في التابع الفيديوي المرجعي والتابع الفيديوي المعالج مكانياً، في تمرير منخفض ويقسم إلى عينات فرعية في ثلاث فئات مختلفة من الاستبانة، R1 و R2 و R3، كما يلي:

الرتل الأصلي	R1	R2	R3
عرض × ارتفاع	540 × 960	270 × 480	96 × 128

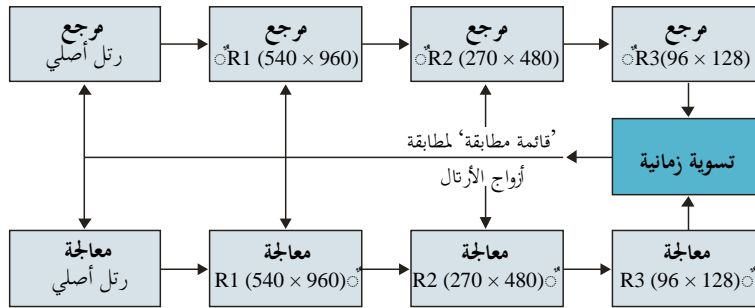
انظر الطريقة `CFrameAnalysisFullRef::ContentTimeAlignment` لتوليد الأرتال في استبانة R3 والطريقة `CFrameSeq::ReadFrame` لتوليد الأرتال في استبانتي R1 و R2.

ويلاحظ أن التنفيذ ليس من السهولة بمكان، وذلك بسبب التقييدات على الذاكرة.

ويبين الشكل 4 الاستبانات الثلاث لتقسيم العينات الفرعية.

الشكل 4

يتم بالتمرير المنخفض ترشيح أرتال التابع الفيديوي المرجعي والمعالج وتقسيم الأرتال إلى عينات فرعية في ثلاث استبانات. وتستخدم الاستبانة الأصغر R3 لتسوية الرتل زمنياً. ويمكن استخدام القائمة المتبقية من الأرتال المتطابقة لمطابقة الأرتال في أي استبانة أخرى



BT.1907-04

2.2 التسوية الزمانية

تتم التسوية الزمانية باستخدام التابع الفيديوي المرجعي والمعالج في الاستبانة المنخفضة R3.

وتتم التسوية الزمانية بشكل تكراري على النحو التالي:

(1) يُبحث عن رتل 'مرساة' في التابع المرجعي (`Ref_anchor`).

(2) يطابَق هذا الرتل مع أفضل رتل مطابقة في التابع المنحط (`Deg_best_match`).

ويؤخذ هذا الرتل الأفضل لمطابقة في التابع المنحط (`Deg_best_match`) ويطابق مع الأرتال القريبة من الرتل 'المرساة' في المرجع (`Ref_anchor`). ويُبحث عن مطابقة أفضل، تبعاً لمعيار تشابه، بين `Ref_anchor` والأرتال في بيئة `Ref_anchor` ويحتفظ بما بوصفها أفضل زوج مطابقة. وتستخدم العلاقة:

$$(1.2) \quad \text{sim} = \exp(-\text{mean_square_diff}(a*x+b, y))$$

بوصفها معيار تشابه بين المستوي Y للرتل المعالج x والرتل المرجعي y، مع اختيار المعلمتين a و b، بحيث يقل متوسط تربيع الفرق إلى الحد الأدنى بين قيم المستوي Y للرتل المعالج x والرتل المرجعي y. في إطار تنفيذ المرجع، انظر الطريقة

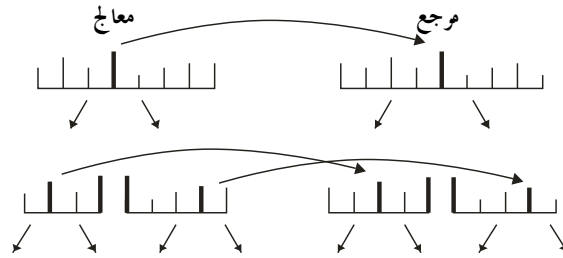
`.FrameSimilarity::similarity`

(1) إذا كان زوج الأرتال المطابق هذا حسن التطابق (أي أن معيار التشابه يتجاوز عتبة القبول) يُفصل التتابع الفيديوي المرجعي والمعالج عند زوج الأرتال المطابق، كل إلى تتابعين فيديويين قبل الأرتال المطابقة وبعدها. يُبدأ عند البند 1) بكلا زوجي التتابع المرجعي والمعالج.

(2) إذا لم يكن زوج الأرتال المطابق 'حسن التطابق'، يُبدأ مرة أخرى عند البند 1) برتل 'مرساة' مختلف من الفيديو المرجعي. وبما أن ليس هنالك من معرفة سابقة بالقيمة المرتقبة لرتل 'حسن التطابق'، فإن عتبة المطابقة تُخفض تكرارياً. وقد تحددت القيم التالية بناء على عينات بيانات تدريب كثيرة: عتبة البدء، فيما يتعلق بمعيار التشابه في المعادلة (1.2) هي 0,98. وبعد الفشل في مطابقة 10 'مراسي' تُخفض العتبة بعامل 0,98، ويعاود البدء بالمطابقة عند البند 1). وعلى هذا النحو، تجري محاولة 10 مراسي أخرى على الأكثر، فإذا فشلت تُخفض العتبة مرة أخرى. وتستمر هذه العملية حتى الوصول إلى قيمة دنيا قدرها 0,1. للاطلاع على تفاصيل التنفيذ كاملة، انظر: .SQ_TimeAlignment::findAncorAndDescend

الشكل 5

رسم توضيحي للنهج التكراري المستخدم للتسوية الزمانية. تجري مطابقة رتل مرساة في المرجع مع رتل في التتابع المعالج. ثم يفصل كلا التتابعين ويختار رتل مرساة في كل تابع ويطبق



BT.1907-05

ونتيجة التسوية الزمانية هي تتابع ('قائمة مطابقة' أساساً) ينسب إلى كل رتل في التتابع الفيديوي المعالج رتلاً من المرجع، أو مؤشراً، يشير إلى عدم العثور على رتل حسن المطابقة إلى درجة كافية. وهكذا، ومن أجل مراحل المعالجة اللاحقة، يكون لكل رتل مطابق في التتابع الفيديوي المعالج رتل مقابل في الفيديو المرجعي. وتقارن تلك الأرتال في التتابع الفيديوي المعالج المقرونة بمؤشر 'غير مطابق' مع الرتلين المرجعيين المطابقين للرتل 'المطابق' السابق واللاحق في التتابع الفيديوي المعالج. ويلاحظ أن خيار 'حد المطابقة' يكون منخفضاً جداً بحيث لا ينطبق مؤشر 'غير مطابق' إلا على الأرتال المنحطة جداً.

للاطلاع على تفاصيل التنفيذ كاملة، انظر الطريقة:

CFrameAnalysisFullRef::sqVTA_ContentFrameTimeAlignment_M

3.2 تسوية الرتل مكانياً

يسري التكرار على جميع الأرتال في التتابع الفيديوي المعالج:

(1) فإذا كان هذا الرتل غير مطابق، تستخدم التسوية المكانية السابقة. وإذا كان هذا الرتل مطابقاً، تجري تسوية مكانية بين الرتل المعالج والرتل المرجعي المقابل - تبعاً لقائمة المطابقة للتسوية الزمانية:

أ) بالنسبة للرتل الأول، يتم تدميث الزحزحة المكانية لتكون 0 (في كلا الاتجاهين الأفقي والعمودي). وبالنسبة للأرتال اللاحقة، تستخدم بمثابة 'سابق' التسوية المكانية للرتل المطابق الأسبق.

ب) يسري التكرار على كل أحوال الزحزحة المكانية الممكنة (أفقياً وعمودياً) باستخدام الحد في البند 2) أدناه. فإذا أدت زحزحة مكانية مختلفة إلى فارق أصغر كثيراً (نسبة إلى دالة تكلفة) بين الرتل المعالج والرتل المرجعي المقابل، يتم تعديل الزحزحة المكانية. وتستخدم الدالة

$$\text{rmse}(Y(dv, dh), Y_{\text{ref}}) + \text{abs}(dv) + \text{abs}(dh),$$

بمغابة دالة تكلفة، حيث تشير Y إلى المستوي Y في الرتل المعالج عند استبانة R1 وتشير Y_ref إلى الرتل المرجعي عند الاستبانة R1، وتشير Y(dv,dh) إلى الرتل Y، المزاح بمقدار dv و dh، حيث dv و dh هما الزحزحة العمودية والأفقية على التوالي. ويدخل الحد الثاني والحد الثالث في دالة التكلفة، لتيسير زحزحات مكانية صغيرة. ويلاحظ تحطبي حافة صغيرة من الأرتال من أجل حساب خطأ جذر متوسط التربيع (rmse)، وذلك لتجنب معاملة للحافة أكثر تعقيداً.

ج) وعلى هذا النحو، يمكن تعويض زحزحات مكانية متغايرة زمنياً. ويمكن تصحيح سوء التسوية في رتل ما بالتسوية في أرتال لاحقة.

(2) وتقتصر هذه الخطوة الأولى في تسوية الزحزحة المكانية الأوتوماتية على ± 4 بيكسلات. بالنسبة لزحزحات مكانية أكبر، انظر الفقرة 9.2.

(3) وبعد التسوية المكانية، يكون لكل رتل في التابع الفيديوي المعالج رتل مرجعي مقابل (أو رتلان في حالة رتل غير مطابق) وفقاً لقائمة المطابقة من التسوية الزمانية وتصحيح زحزحة مكانية معرّفة بوضوح. وهكذا يمكن مقارنة أرتال التابع الفيديوي المعالج بدقة مع أرتال التابع المرجعي. وهذا أمر أساسي بالنسبة للملامح التالي ذكرها.

للاطلاع على كامل تفاصيل التنفيذ، انظر الطريقة CFrameAnalysisFullRef::DetermineSpatialAlignment ويمكن زيادة العتبة الثابتة في الخطوة 2 (± 4 بيكسلات) لاستيعاب زحزحات مكانية أكبر.

4.2 حساب أوجه الشبه المحلية وملامح الاختلاف المحلية

تحسب، بالنسبة لكل زوج أرتال مسوّى، مجموعة من ملامح الجودة المكانية:

أولاً، يحسب مقياس تشابه محلي ومقياس اختلاف محلي بالتكرار عبر مناطق تريبعية متاخمة وموزعة بالتساوي بمساحة 13×13 من الرتل المعالج والرتل المرجعي باستبانة R2. وبما أن الاستبانة R2 لا تقبل القسمة على 13، تحمل حافة صغيرة.

وتدعى المناطق المحلية s_prc و s_ref، ويحسب التشابه s والاختلاف D بالعلاقة:

$$(1.4) \quad S = (\text{cor}(s_{\text{prc}}, s_{\text{ref}}) + 25) / (\text{var}(s_{\text{ref}}) + 25)$$

$$D = \text{sqrt}(\text{avg}((S * (s_{\text{prc}} - \text{mean}(s_{\text{prc}})) -$$

$$(2.4) \quad (s_{\text{ref}} - \text{mean}(s_{\text{ref}}))^2))$$

حيث cor هي علاقة الترابط و var هي التفاوت في قيم البيكسلات في منطقة التربيع المقابلة. وتحسب الدالة avg المتوسط عبر جميع بيكسلات منطقة التربيع، و sqrt تشير إلى جذر التربيع. والقيمتان D و S هما المساهم الرئيسي في قيمة الجودة المكانية.

وعند هذه النقطة، يكون التشابه والاختلاف S و D هما صنفية قيم لكل رتل، تقابل كل قيمة منطقة تربيع محلية. ولا يقتصر العنصر الهام للجودة المدركة على متوسط القيمة فحسب وإنما يشمل شكل توزيع S و D على التوالي.

5.2 تحليل توزيع الملامح المحلية

يبدأ هذا القسم بتقديم بعض التسميات:

لنفترض أن الكم (X, c) quantile يشير إلى c -quantile في توزيع قيم (مدخلات) متجه أو صفيقة X . وعلى نحو أدق، ومن أجل متجه X وثابت c وعلاقة $0 < c \leq 1$ ، يكون الكم

$$q = \text{quantile}(X, c)$$

هو القيمة q ، بحيث يكون جزء c من جميع القيم X أصغر أو يساوي q .

وتعرف الدالة trimmed_mean كما يلي أدناه، وسوف تستخدم هذه التسمية فيما بعد. وبالنسبة لصفيقة X تكون الدالة trimmed_mean :

$$\text{trimmedMean}(X, c)$$

هي متوسط جميع المدخلات X بين أكمات c و $(1-c)$ من الصفيقة X .

مثال ذلك، $\text{trimmedMean}(X, 0.1)$ هي متوسط قيم X ، بإهمال 10% من أصغر قيم X و 10% من أكبر قيم X .

وتشير $X(X > c)$ إلى مجموعة من جميع قيم X التي هي أكبر من c . مثال ذلك

$$\text{trimmedMean}(X, c) = \text{mean}(X(X > \text{quantile}(X, c) \text{ and } X < \text{quantile}(X, 1-c)))$$

وباستخدام هذه التسميات، تحسب قيم الملامح التالية بناء على S من المعادلة (1.4)، وبناء على D من المعادلة (2.4):

$$(1.5) \quad s_m = \text{trimmedMean}(S, c)$$

$$(2.5) \quad d_m = \text{trimmedMean}(D, c)$$

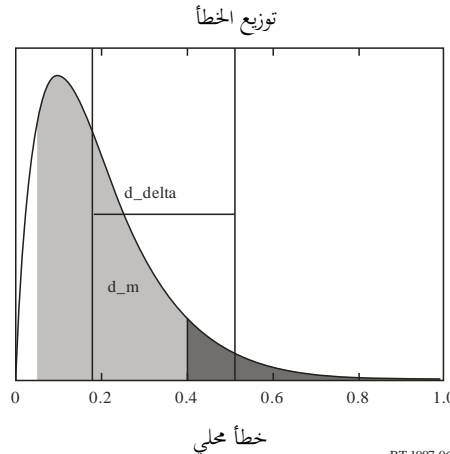
$$(3.5) \quad s_delta = s_m - \text{mean}(S(S < \text{quantile}(S, c)))$$

$$(4.5) \quad d_delta = \text{mean}(D(D > \text{quantile}(D, 1-c))) - d_m$$

باستعمال $c=0.2$. ويبدو هذا بيانياً في الشكل 6، الذي يعرض d_m و d_delta .

الشكل 6

توزيع ظاهرة D محلية. يقابل $\text{trimmed mean } d_m$ متوسط المنطقة الرمادية الفاتحة (خط أسود عمودي). وتقابل القيم d_delta الفرق في متوسط القيم في المنطقة الرمادية الداكنة والمنطقة الرمادية الفاتحة (خط أفقي)



لمعرفة حساب S و D انظر الطريقة: CFrameAnalysisFullRef::ComputeSimilarity.

6.2 حساب ظاهرة السد

تُحسب ظاهرة السد باستخدام الأرتال في استبانة R1. ويتم بذلك قياس مدى رؤية حواف السد الناجمة عن أخطاء التشفير و/أو البث. ونظراً للحساب في الاستبانة R1، يوضع أوتوماتياً تركيز على مدى رؤية الحواف المدركة. وبدءاً من نظرة مجملية، تحسب ظاهرة السد:

(1) المشتقات التوجيهية (صور الحواف) للحواف الأفقية والعمودية. وهذا يفضي إلى صفتين، واحدة للحواف الأفقية وأخرى للحواف العمودية، لكل رتل في التتابع الفيديوي، تدعيان $horGrad_n$ و $verGrad_n$ في شبه الشفرة أدناه.

(2) المجموع باتجاه الصفوف وباتجاه الأعمدة للوغاريتم الحواف الأفقية والعمودية، ما يؤدي إلى متجهين، أحدهما يقابل مجموع الحواف الأفقية والآخر مجموع الحواف العمودية، يدعيان $sumW$ و $sumH$ أدناه.

(3) متوسط عينة فرعية لكل من $sumW$ و $sumH$ ، على التوالي، بحجم خطوة n وتخالف m ، يحسب بواسطة الدالة $vq_AvgSubsample$ أدناه.

والفكرة هي أن بنية تربيعية قوية من التريعات بحجم n سوف تظهر بمثابة فارق هام $delta_edge$ في $vq_AvgSubsample$ في حجم خطوة n محسوبة لتخالفات شتى.

مثال ذلك، بنية تربيعية بحجم 4 في الرتل الأصلي لها بنية تربيعية بحجم 2 في الاستبانة R1. لذلك ينبغي أن يكشف حساب $vq_AvgSubsample(x, 2, 0)$ و $vq_AvgSubsample(x, 2, 1)$ عن فارق هام، إذا وجدت بنية تربيعية قوية. ولتجنب أي تبعية للمحتوى، بينت التجارب التي تستعمل عينة كبيرة من التتابعات الفيديوية كيفية ربط الفارق المقيس للتتابع الفيديوي المعالج مع قيم التتابع المرجعي.

وأفضل وسيلة لشرح الحساب بمزيد من التفصيل هي باستعمال شبه الشفرة التالية. وهنا، يكون $horGrad$ و $verGrad$ هما المشتقين المكانيين الأفقي والعمودي للرتل، نتيجة الفرق بين البكسلات المتاخمة،

$$verGrad_n(i, j) = Y_n(i+1, j) - Y_n(i, j),$$

$$horGrad_n(i, j) = Y_n(i, j+1) - Y_n(i, j),$$

حيث تشير $Y_n(i, j)$ إلى قيمة البكسل في الموقع (i, j) في المستوي Y في الرتل n . وتُحسب الدالة

$$vq_AvgSubsample(x, step, offset)$$

متوسط قيمة المتجه x عبر جميع العينات بحجم خطوة $step$ والبدء عند التخالف $offset$.

```
// loop over all frames and compute:
for( UINT i=0; i<horGrad.Height; i++ ){
    for( UINT j=0; j<horGrad.Width; j++ ){
        w = (double)verGrad(i, j);
        h = (double)horGrad(i, j);
        // sum edges (-2: small differences can be the result of integer
        // values used to store frames)
        sumW(i) += log(1.0+max(0.0, fabs(w)-2.0));
        sumH(j) += log(1.0+max(0.0, fabs(h)-2.0));
    }
}

double dH0 = vq_AvgSubsample( sumH, 2, 0 );
double dH1 = vq_AvgSubsample( sumH, 2, 1 );
double dW0 = vq_AvgSubsample( sumW, 2, 0 );
double dW1 = vq_AvgSubsample( sumW, 2, 1 );
```

```

edge_max = 0.5 * (vq_Max(dW0,dW1) + vq_Max(dH0,dH1) );
edge_min = 0.5 * (vq_Min(dW0,dW1) + vq_Min(dH0,dH1) );

// now: denote by edge_max(i) the value of edge_max above, corresponding to
// frame i of the processed video sequence, and by edge_max_ref(i) the values
// edge_max above corresponding to frame i of the reference video sequence,
// and analogously for edge_min(i), edge_min_ref(i). Then compute:

for( UINT i=0; i<nbOfFramesInProcessedVideo; i++ ){
    // get frame nb of ref frame (according to match-list)
    UINT i_ref = (UINT)floor(ref_frameNb_all(i)+0.5f);

    float delta_edge = edge_max(i) - edge_min(i);
    float delta_edge_ref = edge_max_ref(i_ref) - edge_min_ref(i_ref);

    x(i) = vq_Max(0.0f,delta_edge - delta_edge_ref) / (1.0f+edge_max(i));
}
// blockiness(i) is then a non-linear monotone transform of x(i) ...

```

ويلاحظ أنه نظراً لإمكانية زيادة العينات حتى 720 رتلاً، يزداد تعقيد الحساب نسبياً.

ومعرفة كامل تفاصيل تنفيذ حساب ظاهرة السد، انظر `vquad_hd::vq_BlockinessPhaseDiff` و `CQualityModelFullRef::Blockiness`.

7.2 حساب ظاهرة الاهتزاز (الجودة الزمانية)

تحسب ظاهرة الاهتزاز بتوسيط ناتج زمن العرض النسبي وتحويل لا خطي لزمن العرض وتحويل لا خطي لشدة الحركة. وتشق شدة الحركة أساساً بواسطة الفوارق بين الأرتال في فرادى مناطق الرتل. وزمن العرض هو الزمن، بالمليثانية، الذي يعرض الرتل أثناءه على الشاشة. ويلاحظ أنه لتحديد زمن العرض لكل رتل يتم تحليل شدة الحركة محلياً، إذ قد تكون الأرتال في التتابع الفيديوي المعالج تكراراً لأرتال سابقة.

وتأخذ ظاهرة الاهتزاز في الحساب كمية المعلومات المفقودة أثناء استعادة بث التتابع الفيديوي المعالج. وهي منخفضة جداً في حالة استعادة سلسلة لبث تتابع، بينما تزداد في حالات التوقف أو تخفيض معدل الأرتال. ومن جهة أخرى، وبالنسبة لتشوه زمني ثابت، يأخذ مقياس الاهتزاز قيمة أعلى في أحوال التتابع الفيديوي الذي تزداد فيه شدة الحركة.

وتبين شبه الشفرة التالية التفاصيل. ويلاحظ أن المدخلات هي متجه شدة الحركة `motionInt`، ومتجه احتمالات تكرار الأرتال `repFrame`، ومتجه زمن عرض الرتل `displayTime`. والخرج هو متجه `jerkiness`، أي الاهتزاز عند كل رتل من التتابع الفيديوي المعالج. وبمزيد من التفصيل، يشير المتجه `motionInt` إلى جذر متوسط تربيع الفرق بين الأرتال، مقيساً في المستوي Y في استبانة R2. ويشير المتجه `repFrame` إلى احتمال تكرار الأرتال، أي تبعاً لشدة الحركة يحتمل لكل رتل أن يكون تكراراً للرتل الأسبق: في حالة تكرار كامل للرتل الأسبق، يكون للرتل الحالي احتمال 1 أن يكون تكراراً. وفي حالة شدة حركة عالية، يكون للرتل الحالي احتمال 0 أن يكون تكراراً للرتل الأسبق. ويمكن أن تقع احتمالات قيم متوسطة إذا كانت شدة الحركة ضئيلة جداً ولكنها ليست صفراً. وبكل وضوح،

$$\text{repFrame}(i) = \begin{cases} 0 & \text{if } m(i) < p/2 \\ (m(i)-p/2)/p & \text{if } p/2 \leq m(i) < 3/2*p \\ 1 & \text{if } 3/2*p \leq m(i) \end{cases}$$

حيث تشير $m(i)$ إلى شدة الحركة في الرتل i . وقد تم، تجريبياً، اختيار المعلمة $p=0.01$.

```
int vq_CalcJerkiness( const CVector<float> & motionInt,
                    const CVector<float> & repFrame,
                    const CVector<float> & displayTime,
                    CVector<float> & jerkiness ){

    // -----
    // the 4 parameters of the jerkiness measure: determined using a
    // large sample of video sequences containing temporal degradations
    // only.
    float a = 0.9f;
    float b = 5.0f;

    float aT = 40.0f;
    float bT = 5.0f;
    // -----

    // get probability of new frame = 1 - prob of repeated frame:
    CVector<float> newFrame = repFrame*(-1.0f) + 1.0f;

    // count number of non-repeated frames
    float fNbRepeated = repFrame.Sum();
    float fNbNonRepeated = repFrame.Length() - fNbRepeated;

    // calculate jerkiness
    float fR = 0.0f;
    // look for frame repetition intervals (~ display time)
    // of length i
    for( UINT i=1; i<= iNbFrames; i++ )
    {
        // look for frame repetitions starting at position j
        for( UINT j=0; j<iNbFrames-i+1; j++ )
        {
            float fP = newFrame(j); // prob. : start of repetition block

            for( UINT k=1; k<i; k++ )
            {
                fP *= repFrame(j+k); // prob. : all repeated frames
            }
            if( i+j < iNbFrames )
            {
                fP *= newFrame(j+i); // prob. : end of repetition block
            }

            // calculate the display time (in s) of frame j,
            // if displayed from
            // time t_j until t_(j+i), which occurs with probability fP
            float fDispTime = displayTime.SumPart(j,j+i)/1000.0f;

            // -> measure jerking and add to result
            float fIFDiff = motionInt( j+i-1 );

            // normalisation values: such that at 0 jerkiness value is 0,
            // and saturates at 1
            float c = 1.0f / (1.0f+exp(b));
            float cT = 1.0f / (1.0f+exp(bT));

            float fJ = 1.0f / (1.0f + exp( -( a * fIFDiff - b ) ));
            float fJT = 1.0f / (1.0f + exp( -( aT * fDispTime - bT)));
            fJ = (fJ - c)/(1.0f - c);
            fJT = (fJT - cT)/(1.0f - cT);
        }
    }
}
```

```

//      // jerkiness value: propability * interframeDiffFactor *
//      displayTimeFactor

      fR = fP * fJ * fJT * fDispTime;

      // add to jerkiness vector at position j+i (corresponding to the
      // end of display time)
      jerkiness( vq_Min(j+i,iNbFrames-1) ) += fR;
    }
    return 0;
}

```

لمعرفة تفاصيل التنفيذ، انظر الطريقة `vquad_hd::vq_ProbOfRepeatedFrame` والطريقة `vquad_hd::vq_CalcJerkiness`.

8.2 المجموع التراكمي لمتوسط درجة الرأي (MOS)

إن الظواهر الموصوفة أعلاه، 'التشابه' بموجب s_m و s_{Δ} ، و'الاختلاف' بموجب d_m و d_{Δ} و $blockiness$ و $jerkiness$ ، هي أساس تقدير الدرجة، إلى جانب متجه أزمنة عرض الأرتال $displayTime$. وهي متجهات تساوي أطوالها عدد الأرتال في التابع الفيديوي المعالج.

ولتدوين هذه الظواهر على مقياس إدراكي، تستخدم دالة ذات معلمات في شكل حرف S:

$$S: \mathbb{R} \rightarrow \mathbb{R}, y = S(x).$$

وتحدد معلمات الدالة S بالقيم الثلاثية (p_x, p_y, q) . وتكون المعلمات (p_x, p_y) هما الموقع في $\mathbb{R} \times \mathbb{R}$ و q هي انحدار نقطة الانحناء، ومن باب التفصيل:

$$S(x) = a * x^b \text{ عندما } x \leq p_x$$

$$(1.8) \quad d / (1 + \exp(-c * (x - p_x))) + 1 - d \text{ خلاف ذلك}$$

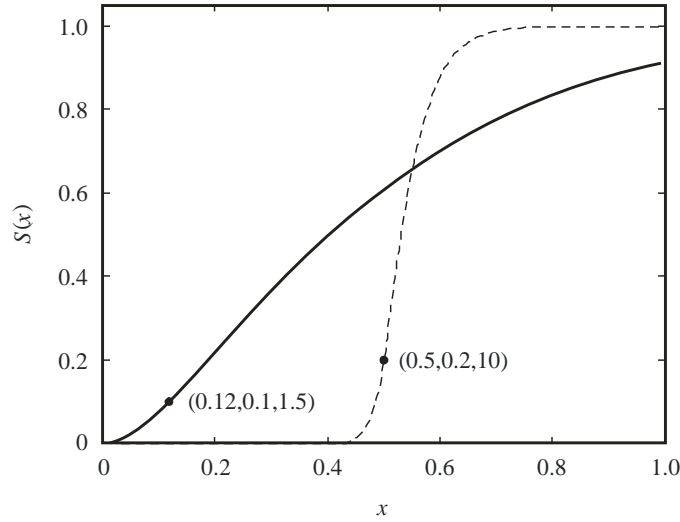
حيث:

$$\begin{aligned}
 p_y / p_x^{(q * p_x / p_y)} &= a \\
 q * p_x / p_y &= b \\
 4 * q / d &= c \\
 2 * (1 - p_y) &= d
 \end{aligned}$$

للاطلاع على المنحنى البياني للدالة s بمختلف المعلمات، انظر الشكل 7. وتبدأ الدالة ذات الشكل s عند نقطة الأصل، وتنمو متعددة الحدود حتى نقطة الانحناء، وتبلغ حد الإشباع بالتزايد هندسياً نحو القيمة 1.

الشكل 7

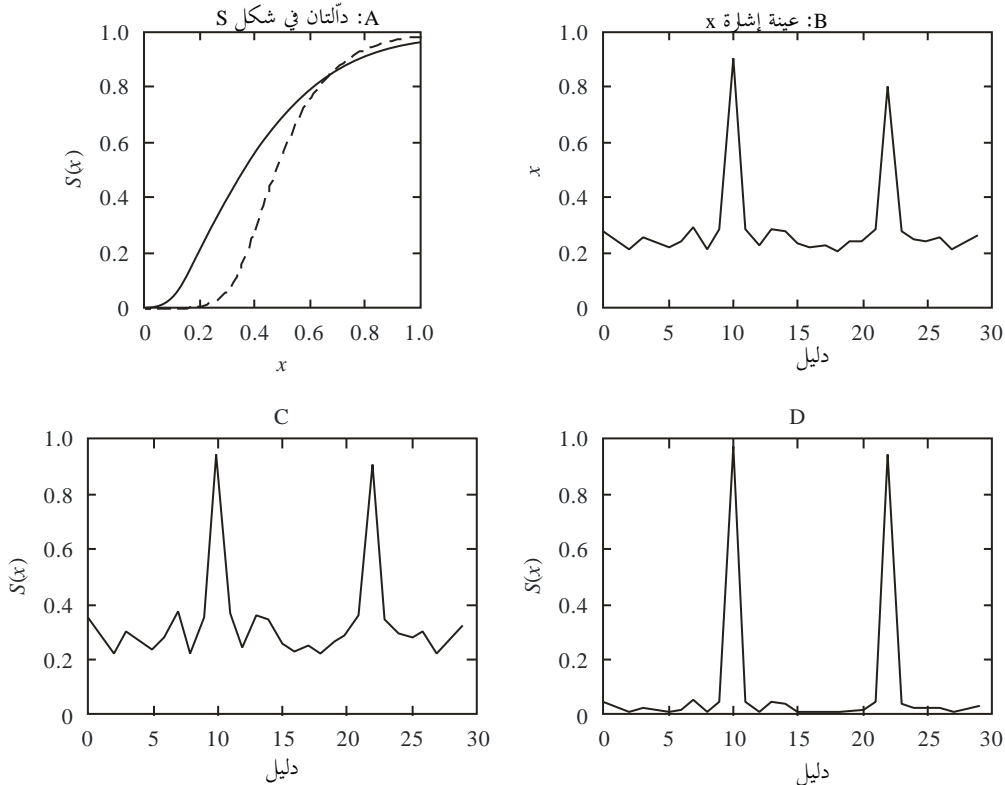
الدالة ذات الشكل S، محددة معلماتها بحكم موقع وانحدار نقطة الانحناء. ويبدو في الشكل مثال دالتين لمعلمات مختلفة



BT.1907-07

الشكل 8

يبين المنحنى (A) دالتين في شكل حرف S لمجموعتين مختلفتين من المعلمات. ويبين المنحنى (B) عينة لمتجه إشارة. جميع القيم x محولة باستخدام الدالتين في شكل حرف S للمنحنى (A). وتظهر النتيجة في (C) و (D) يستخدم التحويل في شكل حرف S لتدوين ظاهرة x في مقياس إدراكي، باستخدام معلمات مكيفة لعينات البيانات، مبينة في (C). وباستخدام دالة في شكل S (الخط المتقطع في (A)) ومعلمات تتوقف على البيانات (المعلمة الأولى موضوعة لقيمة مضاعفة لمتوسط الإشارة x)، يمكن استحداث كاشف لأحوال الانحطاط العابرة، (D)



BT.1907-08

وباستخدام دالة في شكل حرف S ومعلمات ثابتة، يمكن تدوين قيمة الظاهرة في مقياس إدراكي. وباستعمال معلمات معتمدة على البيانات، يمكن استخدام الدالة في شكل حرف S للعمل مثلاً على ضغط جميع القيم دون المتوسط لقيمة ضئيلة، وتوسيع كل القيم ما فوق ذلك. وعلى هذا النحو، يمكن استخدامها لحساب ظاهرة حساسة لأحوال الانحطاط العابرة. انظر الشكل 8. وللمضي في وصف النموذج، يعرف أولاً انحطاط يستند إلى ملامح التشابه أعلاه، ولكن مع ترجيح أهمية أشد حالات الانحطاط:

$$d_s = 1 - s_m + 1.5 s_{\text{delta}}.$$

والفكرة التالية هي استعمال دالتين في شكل حرف S، واحدة باستخدام مجموعة ثابتة من المعلمات `cod_par`، لتحويل `d_s` إلى `d_cod` على مقياس إدراكي متصل بانحطاط قاعدة، يعكس الأخطاء الناجمة عن التشفير الفيديوي.

وثمة دالة ثانية في شكل حرف S لها معلمات تعتمد على البيانات، لتحويل `d_s` إلى `d_trans` على مقياس إدراكي متصل بأحوال انحطاط عابرة يعكس الأخطاء الناجمة عن البث.

ومن قبيل التفصيل، فيما يتعلق بشبه الشفرة أدناه، تحدد الدالة `SplitCodTrans` باستخدام المتجه `d_s` بمثابة مُدخل والمتجه لزمن عرض الرتل `disp_time` ويكون الخرج `d_trans`، `d_cod`

```
SplitCodTrans(d_s, disp_time, d_cod, d_trans)
```

وتبين شبه الشفرة التالية الدالة بالتفصيل، ويلاحظ أن

```
stat.STransform(x,px,py,q)
```

تشير إلى الدالة في شكل حرف S ويكون المدخل فيها القيمة الحقيقية `x` التي سوف تحوّل، والمعلمات الثلاث `(px,py,q)` المحددة في المعادلة (1.8).

```
SplitCodTrans( const CVec& v, const CVec& dispTime,
               CVec& v_cod, CVec& v_trans ){

    // these parameters are determined empirically
    float qPosSmall = 0.55f;
    float qPosLarge = 0.65f;

    // q is the mean of the values in v between the qPosSmall and qPosLarge
    // quantiles.
    float q = r.TrimmedMean( qPosSmall, qPosLarge, v, dispTime );

    for( UINT i=0; i<v.Length(); i++ ){

        // the parameters used here are the result of fitting to sample
        // data
        v_cod(i) = stat.STransform(v(i), 0.07f, 0.1f, 2.0f);

        // Note that the STTransform is not directly applied to v, but
        // to v(i)-q . This is preferred here for numerical reasons of the
        // resulting STTransform.

        // v_trans is part of v above quantile value q
        v_trans(i) = vq_Max(0.0f, v(i)-q);

        float px = 0.5f * (q+0.2f);
        // the parameters used here are the result of fitting to sample
        // data
        v_trans(i) = stat.STransform(v_trans(i),px, 0.1f,16.0f);
    }
}
```

للاطلاع على كامل تفاصيل التنفيذ، انظر `CQualityModelFullRef::SplitCodTrans`.

وبالقياس، تشتق d_{diff_cod} من d_{diff_trans} و d_{delta} ، بتحديد

$$d_{diff} = d_m + 1.5 d_{delta}$$

وتسمية

$SplitCodTrans(d_{diff}, disp_time, d_{diff_cod}, d_{diff_trans}),$

واستخدام ثلاثية المعلمات لتحويل الدالتين S

$$cod_par = (4.0f, 0.05f, 0.2f)$$

$$trans_par = (0.5*(q+4.0f), 0.1f, 0.4f),$$

حيث تشير q إلى المتوسط بين الأكمام، كما في شبه الشفرة أعلاه. ويكون خرجا الدالة هما d_{diff_cod} و d_{diff_trans} .

وبعدئذ تحسب قيمة الظاهرة المتصلة بقيمة الاهتزاز الكبيرة العابرة بتحويل الاهتزاز في شكل S:

$$d_{t_trans} = S(\max(0, jerkiness-q))$$

حيث تعطى معلمات التحويل في شكل S بواسطة $(\max(0.048, q), 0.2, 40.0)$ وتشير q إلى المتوسط ما بين الكم 0,55 والكم 0,65 في متجه الاهتزاز. وقد تحددت المعلمات بالتكيف مع مجموعة واسعة من عينات البيانات.

للاطلاع على كامل التفاصيل، انظر $CQualityModelFullRef:: SplitTempTrans$.

وبعدئذ، تعرّف جودة القاعدة q_{cod} بوصفها

$$(2.8) \quad q_{cod} = (1 - d_{cod}) * (1 - d_{diff_cod}) * (1 - blockiness)$$

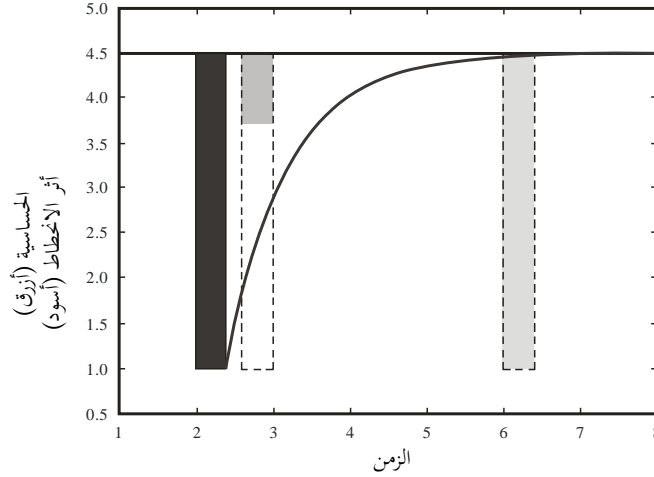
وتعرّف الجودة العابرة بوصفها

$$q_{trans} = (1 - d_{trans}) * (1 - d_{diff_trans}) * (1 - d_{t_trans}).$$

ويخفف تأثير أي انقطاعات إضافي إذا كان يحدث قريباً بعد انقطاعات أول. ولأخذ هذا التأثير في الحسبان، تحول q_{trans} إلى q_{fq} . وهذه الفكرة موضحة في الشكل 9.

الشكل 9

يشير الشريط الأسود إلى انحطاط يحدث في تابع فيديوي مفترض. ويشير المنحنى الأسود إلى الحساسية لأحوال انحطاط لاحقة. ويخفف تأثير أي انحطاط إضافي يعبر عن شدته بالخط المتقطع (الشريط الرمادي الداكن الأيسر) إذا كان يحدث قريباً بعد الانحطاط الأول. ولا يتغير تأثير أي انحطاط لاحق إذا كان يحدث منفصلاً بفترة زمنية واسعة (الشريط الرمادي الفاتح الأيمن)



BT.1907-09

وأفضل ما يكون وصف تفاصيل الحساب في القسم التالي لشبه الشفرة، حيث تستخدم $1-q_{trans}$ بمثابة مُدخل المتجه v ويكون المتجه $disp_time$ هو متجه أزمنة عرض الرتل. وقد تحدد ثابت زمن الإخماد $dT=1000$ ميللي ثانية باستخدام عينات بيانات، ثم وضع

$$(3.8) \quad q_{fq} = 1 - w,$$

حيث w هي متجه خرج الدالة:

```

DegFreq( const CVec& v, const CVec& disp_time, CVec& w, float dT ){
    // time constant for temporal integration of degradations
    float t_const = 80.0f;
    w(0) = v(0)*vq_Min(t_const,disp_time(0))/t_const;

    for( UINT i=1; i<v.Length(); i++ ){
        // integrate degradation over the last t_const milliseconds:
        // use an indicator function, which is 1 over the interval
        // [t_i-t_const, t_i] and 0 otherwise
        float dT_sum = 0.0f;
        UINT j=0;
        float v_sum = 0.0f;
        while( dT_sum < t_const && (int)i-(int)j>=0 ){
            // b is the overlap in [0,1] of the display time interval
            // of frame i-j with respect to the integration interval
            //      t_i-t_const   t_i
            //      |-----|-----> integration window
            // -----> time
            // | | | | | | | frames
            // ----->| b*dT |<-----
            //
            float b = vq_Min(t_const-dT_sum,disp_time(i-j)) / t_const;
            v_sum += v(i-j)*b;
        }
    }
}

```

```

        dT_sum += disp_time(i-j);
        j++;
    }
    // compute the fall-off factor a:
    float a = exp(-disp_time(i-1)/dT);
    // the degradation is now:
    // 1) a linear combination of the fall-off of a previous
degradation
    // and the current (integrated) degradation, or
    // 2) the integrated current degradation (if it is stronger than 1)
).
    w(i) = vq_Max(v_sum, a * w(i-1) + (1.0f-a) * v_sum);
}

```

عملية حساب المتوسط

يفترض أن تشير $\text{disp_time}(i)$ إلى زمن عرض الرتل i وأن تشير $\text{jerkiness}(i)$ إلى قيمة الاهتزاز المقابلة للرتل i . وتُستعاد q_{cod} من المعادلة (8.2) و q_{fq} من المعادلة (3.8). وتوضع القيم

$$\begin{aligned}
 Q_t &= 1 - 1/T \sum_i \text{jerkiness}(i) \\
 Q_{\text{fq}} &= 1/T \sum_i q_{\text{fq}}(i) * \text{disp_time}(i) \\
 Q_{\text{cod}} &= 1/T \sum_i q_{\text{cod}}(i) * \text{disp_time}(i)
 \end{aligned}$$

حيث تشير sum_i و $T = \sum_i \text{disp_time}(i)$ إلى المجموع عبر جميع الأدلة $i=0, \dots, \text{number_of_frames}$

ويكون تقدير الدرجة الأخير المتنبأ به هو ناتج وإعادة لمقياس مدى متوسط درجة الرأي (MOS) [1,5]:

$$s = 4 * Q_t * Q_{\text{cod}} * Q_{\text{fq}} + 1$$

لمعرفة تفاصيل التنبؤ بتقدير الدرجة، انظر الطريقة `CQualityModelFullRef::PredictScoreCodTrans`.

9.2 معاملة تناوبات فيديو سيئة التسوية مكانياً

لتجنب سوء التنبؤ في حالة سوء تسوية مكانية كبير نسبياً بين التتابع الفيديوي المرجع والتتابع الفيديوي المعالج، تحسب الخطوات المذكورة أعلاه من أجل ثلاث خطوات مختلفة للتسوية المكانية الأفقية والعمودية في التتابع الفيديوي، وتقدير الدرجة الأقصى المتنبأ بها بين جميع المواقع المكانية هو درجة الجودة المقدرة النهائية.

ويستخدم حجم خطوة من أربعة بيكسلات في كل اتجاه. وبذلك يتحقق مدى بحث مكاني قدره $8 \pm$ بيكسلات. وهذا يغطي بسهولة الزحزحة المكانية القصوى المستخدمة في مجموعة الاختبار ($5 \pm$ بيكسلات). وبما أن هذا التكبير يحدث في إطار دالة عالية المستوى في النموذج فإن من الميسور تكيف مدى التسوية مع زحزحات أكبر أو يمكن تخفيضه ($4 \pm$ بيكسلات) لتحقيق وفورات في موارد الحوسبة.

انظر `vquad_hd::vq_vquad08`.

10.2 تنفيذ شفرة المصدر المرجعي

ينطوي هذا القسم على أرشيف يتسم بشفرة مصدر C++ تشمل الأجزاء والوظائف الأساسية لوصف نموذج يمثل للتنفيذ. وتشير كل الوصلات إلى التنفيذ الفعلي الوارد في الأقسام الواردة أعلاه إلى شفرة المصدر المرجعي هذه.

الملحق 3

اختبار المطابقة

هذا الملحق مستكمل بمرفق رقمي يحتوي المعلومات والملفات التالية:

- (1) ستة عشر تتابعاً فيديوياً عالي الوضوح (تتابعات مرجعية وتتابعات منحنية لثمانى حالات اختبار). وتشمل هذه التتابعات تشوهات ومحتويات شتى. وهي متوفرة من أجل اختبار تنفيذ نموذج مطابق من قبل مستعمل لهذه التوصية إزاء التنفيذ المرجعي للنموذج. وتتألف التتابعات الفيديوية من مجرد بضعة أرتال فقط وذلك لتخفيض سعة التخزين. وهي لا ترمي إلى أي اختبار مرئي، وإنما لاختبار مطابقة فقط لتنفيذ النموذج عالي الوضوح.
- (2) تقديرات متوسط درجة الرأي (MOS) المتنبأ بها للتتابعات عالية الوضوح المذكورة في البند (1) بوصفها 'HD_ConformanceReferenceResults.xls'. وقد تم الحصول على تقدير الدرجات هذه على أساس تنفيذ مرجعي للنموذج عالي الوضوح.
- (3) تقديرات متوسط درجة الرأي (MOS) متنبأ بها لخمسة قواعد بيانات عالية الوضوح عمومية متاحة عن طريق فريق خبراء الجودة الفيديوية (VQEG). ويمكن استخدام قواعد البيانات هذه بمثابة اختبار مطابقة موسع لتنفيذ النموذج.

معايير اختبار المطابقة:

- '1' يتعين إعادة استخراج الدرجات المرجعية الثماني المذكورة في البند (2)، على وجه الدقة وذلك بتنفيذ النموذج. وتعني عبارة 'على وجه الدقة' هنا إعادة استخراج تقدير الدرجة بدقة بمقدار ثلاثة أرقام عشرية.
- '2' يتعين إعادة استخراج الدرجات المرجعية لمتوسط درجة الرأي (MOS) الخمسة لقواعد بيانات الفريق VQEG العمومية بقدر محدود جداً من الانحراف. ويمكن التساهل ببعض التفاوتات البسيطة، إذ بينت التجربة أن مختلف عمليات الاستمثال من حيث السرعة واستخدام التخزين يمكن أن تؤدي إلى بعض الانحرافات البسيطة والمهملة في درجة التقدير النهائية.

الجدول 2

التوزيع المسموح به لمدى التفاوت في كل بيانات اختبار المطابقة

نسبة التسامح	التفاوت المطلق
5,00%	0,0001 <
1,00%	0,001 <
0,50%	0,01 <
0,05%	0,1 <
0,00%	0,3 <

وبالنسبة لقواعد البيانات غير المحددة في هذا الملحق، يجب عدم تجاوز نفس توزيع الأخطاء. وبالنسبة لبيانات مجهولة يتعين أخذ مجموعة اختبار لا تقل عن 500 زوج ملفات - يـُستحسن أن تكون من تجارب ذاتية كاملة - من أجل هذه الإحصاءات.

بيليوغرافيا

VQEG Final Report of HDTV Phase I Validation Test (2010), “*Video Quality Experts Group: report on the validation of video quality models for high definition video content*”, Video Quality Experts Group (VQEG), <http://www.its.bldrdoc.gov/vqeg/projects/hdtv>

Recommendation ITU-R BT.601-7 (2011), Studio encoding parameters of digital television for standard 4:3 and wide screen 16:9 aspect ratios.

Recommendation ITU-T J.244 (2008), Calibration methods for constant misalignment of spatial and temporal domains with constant gain and offset.

Recommendation ITU-R BT.500-12 (2009), Methodology for the subjective assessment of the quality of television picture.

Recommendation ITU-T J.149 (2004), Method for specifying accuracy and cross-calibration of Video Quality Metrics (VQM).

Recommendation ITU-T J.247 (2008), Objective perceptual multimedia video quality measurement in the presence of a full reference.

Recommendation ITU-T J.144 (2004), Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference.

Recommendation ITU-T P.931 (1998), Multimedia communications delay, synchronization and frame rate measurement.

Recommendation ITU-T J.148 (2003), Requirements for an objective perceptual multimedia quality model.

Recommendation ITU-T H.264 (2012), Advanced video coding for generic audiovisual services.

Recommendation ITU-T J.340 (2010), Reference algorithm for computing peak signal to noise ratio (PSNR) of a processed video sequence with constant spatial shifts and a constant delay.

Recommendation ITU-T P.910 (2008), Subjective video quality assessment methods for multimedia applications.

Recommendation ITU-T P.911 (1998), Subjective audiovisual quality assessment methods for multimedia applications.

Recommendation ITU-T J.143 (2000), User requirements for objective perceptual video quality measurements in digital cable television.
