الاتحاد الدولي للاتصالات

ITU-R

قطاع الاتصالات الراديوية في الاتحاد الدولي للاتصالات

التوصيـــة ITU-R BT.1885

(2011/03)

تقنيات القياس الموضوعي للنوعية الفيديوية الحسية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض

السلسلة BT

الخدمة الإذاعية (التلفزيونية)

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، 2011

© ITU 2011

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.1885

تقنيات القياس الموضوعي للنوعية الفيديوية الحسية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض

(2011)

مجال التطبيق

تصف هذه التوصية أساليب لتقييم موضوعي للنوعية الفيديوية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح. وتتيح هذه الأساليب قياس النوعية الفيديوية الحسي في ظروف الاستقبال المتنقل والثابت عندما تكون الخصائص المستخرجة من إشارة الفيديو المرجعية متاحة بسهولة في نقطة القياس.

إن جمعية الاتصالات الراديوية للاتحاد الدولي للاتصالات،

إذ تضع في اعتبارها،

أ) أن القدرة على القياس التلقائي لتردي إشارات الفيديو المذاعة كانت ولا تزال أمراً محبذاً؛

ب) أن النوعية الفيديوية الحسية للاستقبال المتنقل يمكن أن تتغير دينامياً تبعاً لظروف الاستقبال؛

ج) أن القياس الموضوعي للنوعية الفيديوية الحسية يمكن أن يكمل تكميلاً مفيداً أساليب التقييم الشخصي؛

د) أن ثلاث تقنيات اقتُرحت لقطاع الاتصالات الراديوية من أجل قياس النوعية الفيديوية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض، وقد تبين أن هذه التقنيات توفر نتائج مكافئة بمحاذاة جيدة؛

هـ) أن التوصية ITU-R BT.1683 تورد موصفات تقنيات القياس الموضوعي للنوعية الفيديوية الحسية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض،

توصي

1 باستخدام أي من نماذج النوعية الفيديوية الموضوعية من أجل القياس الموضوعي للنوعية الفيديوية الحسية المعدة للإذاعة التلفزيونية الرقمية عادية الوضوح بوجود عرض نطاق مرجعي مخفض.

1 مقدمة

إن هذا الاختبار التلفزيوني (RRNR-TV) يتناول الصور على النحو المحدد في التوصية ITU-R BT.601-6، وهو على نموذجين: بمرجع مخفض (RR) وبدون مرجع (NR)؛ حيث يتاح لنماذج المرجع المخفض النفاذ إلى المصدر الفيديوي بعرض نطاق محدود فيما لا يتعذر على النماذج بدون مرجع النفاذ إلى المصدر الفيديوي.

وشملت الدرات المرجعية الافتراضية (HRC) في كل تجربة مفرزات بتشفير فقط وتشفير مع أخطاء إرسال. وكانت خطتا التشفير اللتان دُرستا MPEG-2 وH.264 (MPEG-4 الجزء 10). وشُغلت مشفرات MPEG-2 بمعدلات بتات متنوعة تتراوح بين 1,0 و5,5 Mbit/s. وشُغلت مشفرات H.264 بمعدلات بتات متنوعة تتراوح بين 1,0 و3,98 Mbit/s. وشملت كل تجربة تتابعات لاثني عشر مصدراً كان اثنان من هذه المصادر سريين. وتخللت كل تجربة 34 دارة مرجعية افتراضية و156 تتابعاً فيديوياً معالَجاً، احتوى 40 من هذه التتابعات أخطاء إرسال و116 منها احتوى تشفيراً فحسب.

1.1 التطبيقات

توفر هذه التوصية تقديرات النوعية الفيديوية في أصناف الفيديو من TV3 حتى MM5B على النحو المحدد في الملحق B للتوصية ITU-T P.911. وتشمل تطبيقات نماذج التقدير التي يرد وصفها في هذه التوصية ما يلي، على سبيل الذكر لا الحصر:

1) احتمال مراقبة النوعية عند المصدر في الوقت الفعلي خلال الخدمة؛

2) مراقبة النوعية في مقصد بعيد عند توفر القنوات الجانبية للخصائص المستخرجة من تتابعات فيديو المصدر؛

3) قياس النوعية لمراقبة نظام تخزين أو إرسال يستخدم تقنيات ضغط الفيديو وفك ضغطه، إما دفعة واحدة أو بسلسلة من هذه التقنيات؛

4) فحص مختبري لأنظمة الفيديو.

2.1 القيود

يتعذر استخدام نماذج التقدير الواردة في هذه التوصية للاستعاضة بها تماماً عن الاختبار الشخصي. وعادة ما تقع قيم الارتباط بين اختبارين شخصيين جرى تصميمهما وتنفيذهما بعناية (أي في مختبرين مختلفين) في حدود 0,95 إلى 0,98. وإذا ما استخدمت هذه التوصية لمقارنة كودكين مختلفين، يُنصح باستخدام أسلوب كمي (مثل ذلك المذكور في التوصية ITU-T J.149) لتحديد دقة النموذج لذلك السياق بالذات.

وقد تم التحقق من النماذج الواردة في هذه التوصية من خلال قياس للنوعية الفيديوية الموضوعية التي تظهر تجمداً في الأطر لمدة تصل إلى 2 s.

ولم يتم التحقق من النماذج الواردة في هذه التوصية من حيث قياس النوعية الفيديوية الموضوعية ذات التأخير المتزايد باطراد (مثل إشارة الفيديو التي لا تتجاهل الأطر الناقصة بعد تجمد إطار).

وتجدر الإشارة إلى أنه في حالة إفراز التكنولوجيات الجديدة للتشفير والإرسال مفرزات لم ترد في هذا التقييم، يمكن لنماذج التقييم الموضوعي أن تعطي نتائج خاطئة. وهنا لا بد من تقييم شخصي.

2 المراجع

تتضمن التوصيات التالية لقطاع تقييس الاتصالات وغيرها من المراجع أحكاماً تشكل من خلال الإشارة إليها في هذا النص جزءاً لا يتجزأ من هذه التوصية. وقد كانت جميع الطبعات المذكورة سارية الصلاحية في وقت النشر. ولما كانت جميع التوصيات والمراجع الأخرى تخضع إلى المراجعة، نحث جميع المستعملين لهذه التوصية على السعي إلى تطبيق أحدث طبعة للتوصيات والمراجع الواردة أدناه. وتُنشر بانتظام قائمة توصيات قطاع تقييس الاتصالات السارية الصلاحية.

والإشارة إلى وثيقة في هذه التوصية لا يضفي على الوثيقة في حد ذاتها صفة التوصية.

1.2 المراجع المعيارية

التوصية ITU-R BT.500-12 – منهجية التقييم الشخصي لنوعية الصور التلفزيونية.

التوصية ITU-T P.910 (2008)– طرائق تقييم شخصية للنوعية الفيديوية في تطبيقات الوسائط المتعددة.

التوصية ITU-T P.911 (1998) – طرائق التقييم الشخصي للنوعية السمعية المرئية في تطبيقات الوسائط المتعددة.

التوصية ITU-T J.143 (2000) – متطلبات المستخدم لقياسات النوعية الفيديوية الحسية في التلفزيون الكبلي الرقمي.

التوصية ITU-T J.244 (2008) – طرائق معايرة كاملة المراجع ومختصرة المراجع في أنظمة الإرسال الفيديوي مع سوء تراصف دائم للميادين المكانية والزمنية بكسب وتخالف دائمين.

2.2 المراجع الإعلامية

التوصية ITU-T J.149 (1998) – طرائق التقييم الشخصي للنوعية السمعية المرئية في تطبيقات الوسائط المتعددة.

التوصية ITU-T J.144 (2001) – تقنيات القياس الموضوعي لما يُلمس من النوعية الفيديوية المعدة للتلفزيون الكبلي الرقمي بوجود مرجع كامل.

التوصية ITU-T P.931 (1998) – قياس التأخر والتزامن ومعدل الإطارات في الاتصالات متعددة الوسائط.

التوصية ITU-T J.148 (2003) – متطلبات النموذج الموضوعي لما يُلمس من نوعية الوسائط المتعددة.

التوصية ITU-T H.261 (1993) – الكودك الفيديوي للخدمات السمعية المرئية بمعدل kbits 64 x p.

التوصية ITU-T H.263 (1996) – التشفير الفيديوي للاتصالات بمعدل بتات منخفض.

التوصية ITU-T H.263 (1998) – التشفير الفيديوي للاتصالات بمعدل بتات منخفض (H.263+).

التوصية ITU-T H.264 (2003) – التشفير الفيديوي المتطور للخدمات السمعية المرئية العامة.

فريق الخبراء المعني بالجودة الفيديوية (VQEG): التحقق من نماذج التلفزيون عادي الوضوح بمرجع مخفض وبدون مرجع، المرحلة الأولى، 2009.

3 التعاريف

1.3 المصطلحات المعرفة في مكان آخر:

تستخدم هذه التوصية المصطلحات التالية المعرفة في مكان آخر:

1.1.3 **تقييم شخصي (للصورة)** (التوصية ITU-T J.144): تعريف مقتبس على نحو اختياري.

2.1.3 **القياس المحسوس الموضوعي (للصورة)** (التوصية ITU-T J.144): تعريف مقتبس على نحو اختياري.

3.1.3 **الجهة المؤيدة** (التوصية ITU-T J.144): تعريف مقتبس على نحو اختياري.

2.3 مصطلحات معرفة في هذه التوصية

تعرّف هذه التوصية المصطلحات التالية:

1.2.3 **تكرار الإطار الشاذ:** يعرف على أنه حدث تنتج فيه الدارة المرجعية الافتراضية (HRC) في خرجها إطاراً واحداً مراراً رداً على حدث غير عادي أو استثنائي. ويشمل تكرار الإطار الشاذ الأنواع التالية من الأحداث دون أن يقتصر عليها: خطأ في قناة الإرسال وتغير في التأخر عبر قناة الإرسال وتأثير موارد الحاسوب المحدودة على أداء وحدة فك التشفير وتأثير موارد الحاسوب المحدودة على عرض إشارة الفيديو.

2.2.3 **التخطي الثابت للأطر:** يعرف على أنه حدث تنتج فيه الدارة المرجعية الافتراضية (HRC) في خرجها أطراً ذات محتوى محدّث بمعدل أطر فعال ثابت يقل عن معدل أطر المصدر.

3.2.3 **معدل الأطر الفعال:** يعرف على أنه عدد الأطر الفريدة (أي مجموع الأطر – الأطر المتكررة) في الثانية الواحدة.

4.2.3 **معدل الأطر:** يعرف على أنه عدد الأطر الفريدة (أي مجموع الأطر – الأطر المتكررة) في الثانية الواحدة.

5.2.3 معدل الأطر المقصود: يعرف على أنه عدد الأطر الفيديوية في الثانية المخزنة مادياً من أجل عرض تتابع فيديوي. ويتعين أن يكون معدل الأطر ثابتاً. ومن أمثلة معدل الأطر المقصود الثابت شريط BetacamSP® الحاوي على 25 إطاراً في الثانية، وملف المساحة اللونية YUV ذو الستمائة وخمسة وعشرين خطاً المتوافق مع المرحلة الأولى من تجربة التحقق VQEG FR-TV والحاوي على 25 إطاراً في الثانية. ويبلغ معدل الأطر المقصود في كلا هذين المثالين 25 إطاراً في الثانية.

6.2.3 ظروف الشبكة الحية: تُعرَّف على أنها أخطاء تطرأ على تدفق البتات الفيديوية الرقمية نتيجة لظروف الشبكة الحية.

7.2.3 التوقف مع التخطي: يعرف على أنه أحداث تتوقف فيها الإشارة الفيديوية لفترة من الوقت ثم تستأنف فاقدةً بعض المعلومات الفيديوية. وفي التوقف مع التخطي، سيتفاوت التأخر الزمني عبر النظام حول متوسط تأخر النظام زيادةً أحياناً ونقصاناً في أحيان أخرى. ومن أمثلة التوقف مع التخطي، زوج من الهواتف الفيديوية العاملة ببروتوكول الإنترنت، حيث تتسبب كثافة حركة الشبكة بتجمد شاشة عرض الهاتف الفيديوي لفترة وجيزة ، وعندما يُستأنف العرض يكون بعض المحتوى قد فُقد. ويعتبر التخطي الثابت للأطر والتخطي المتغير للأطر من فروع التوقف مع التخطي. ومدة تتابع فيديوي معالَج يحوي توقفاً مع التخطي تتساوى تقريباً مع مدة التتابع الفيديوي الأصلي المرتبط بالتتابع المعالَج.

8.2.3 التوقف دون التخطي: يعرف على أنه أي حدث تتوقف فيه الإشارة الفيديوية لفترة من الوقت ثم تستأنف دون أن تفقد أي معلومات فيديوية. ومن ثم، لا بد أن يزيد التأخر الزمني عبر النظام.

9.2.3 معدل التجديد: يعرف على أنه المعدل الذي يتم به تحديث العرض.

10.2.3 أخطاء الإرسال المحاكاة: تُعرَّف على أنها أخطاء تطرأ على تدفق البتات الفيديوية الرقمية في بيئة مضبوطة للغاية. ومن الأمثلة على ذلك، معدلات فقدان الرزم المحاكى وأخطاء البتات المحاكاة.

11.2.3 معدل أطر المصدر (SFR): هو معدل الأطر المقصود للتتابعات الفيديوية الأصلية في المصدر. ومعدل أطر المصدر هو معدل ثابت. وقد بلغ معدل أطر المصدر في فحص VQEG RRNR-TV 25 إطاراً في الثانية أو 30 إطاراً في الثانية.

12.2.3 أخطاء الإرسال: تُعرَّف على أنها أخطاء تطرأ على الإرسال الفيديوي. ومن أمثلة أنواع الأخطاء، أخطاء الإرسال المحاكاة وظروف الشبكة الحية.

13.2.3 تخطي الأطر المتغير: يعرف على أنه حدث تنتج فيه الدارة المرجعية الافتراضية (HRC) في خرجها أطراً ذات محتوى محدَّث بمعدل أطر فعال يتغير مع الزمن. وسيتفاوت التأخر الزمني عبر النظام حول متوسط تأخر النظام زيادةً ونقصاناً. ومدة تتابع فيديوي معالَج يحوي تخطي الأطر المتغير تتساوى تقريباً مع مدة التتابع الفيديوي الأصلي المرتبط بالتتابع المعالَج.

4 المختصرات

تستخدم هذه التوصية المختصرات التالية:

ACR تصنيف بحسب الفئات المطلقة (انظر التوصية ITU-T P.910) *(Absolute category rating)*

ACR-HR تصنيف بحسب الفئات المطلقة ذات المرجع الخفي (انظر التوصية ITU-T P.910) *(Absolute category rating with hidden reference)*

AVI تشذير سمعي فيديوي *(Audio video interleave)*

DMOS متوسط علامة الرأي في الفارق *(Difference mean opinion score)*

FR مرجع كامل *(Full reference)*

FRTV تلفزيون ذو مرجع كامل *(Full reference television)*

HRC دارة مرجعية افتراضية *(Hypothetical reference circuit)*

NR بدون مرجع (أو بمرجع صفري) (*No (or Zero) reference*)

PSNR نسبة ذروة الإشارة إلى الضوضاء (*Peak signal-to-noise ratio*)

PVS تتابع فيديوي معالَج (*Processed video sequence*)

RMSE الخطأ الفعال (*Root mean square error*)

RR مرجع مخفض (*Reduced reference*)

SFR معدل أطر المصدر (*Source frame rate*)

SRC الدارة أو القناة المرجعية للمصدر (*Source reference channel or circuit*)

VQEG فريق الخبراء المعني بالجودة الفيديوية (*Video Quality Experts Group*)

YUV المساحة اللونية (*Colour space*)

5 الاصطلاحات

لا توجد.

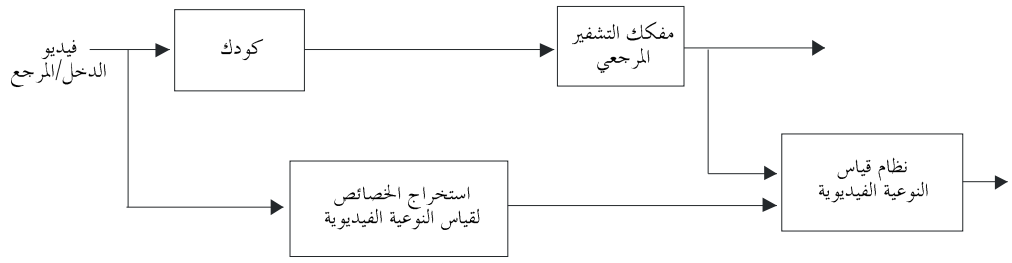
6 وصف طريقة القياس بمرجع مخفض

إن أسلوب القياس الموضوعي على الطرفين لما يُلمس من النوعية الفيديوية يقيّم أداء الأنظمة بالمقارنة بين الخصائص المستخرجة من الدخل غير المشوه أو إشارة الفيديو المرجعية في دخل النظام وبين الإشارة المتردية في خرج النظام (انظر الشكل 1).

ويظهر الشكل 1 مثالاً على تطبيق أسلوب المرجع المخفض لفحص كودك في المختبر.

الشـــكل 1

تطبيق أسلوب قياس النوعية الحسي بمرجع مخفض لفحص كودك في المختبر



BT.1885-01

إن المقارنة بين إشارات المدخلات والمخرجات قد تحتاج إلى عملية موازاة زمنية أو مكانية، حيث تعوض الأخيرة عن أي انزياحات رأسية أو أفقية في الصورة أو عن تشذيبها. وقد تتطلب أيضاً تصحيح أي تخالفات، أو فوارق في الكسب في قناتي النصوع والتلون على السواء. ثم يُحسب تصنيف النوعية الموضوعية للصورة عادةً بتطبيق نموذج الرؤية البشرية الحسي.

وتُعرف عملية الموازاة وضبط الكسب بالتقييس، وهي عملية مطلوبة لأن معظم أساليب المرجع المخفض تقارن الخصائص المستخرجة من الصور المرجعية والصور المعالَجة على أساس كل بيكسل على حدة فعلياً. ومثال ذلك، حساب نسبة ذروة

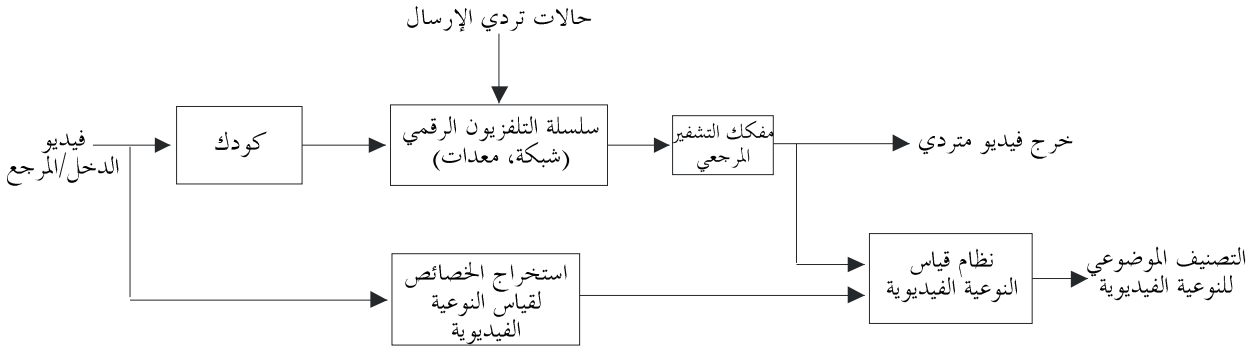
الإشارة إلى الضوضاء (PSNR). فلا تُزال إلا التغيرات الساكنة غير المتغيرة بمرور الوقت، وتقاس التغيرات الدينامية الناجمة عن النظام قيد الفحص كجزء من حساب تصنيف النوعية. وتوفر التوصيتان ITU-T J.244 وITU-T J.144 أساليب معيارية للإبلاغ عن القيم اللازمة لتقييس إشارة الفيديو قبل تقييم النوعية موضوعياً. وتشمل مقاييس النوعية الفيديوية التي يرد وصفها في الملحق بهذه التوصية أساليب التقييس ذات الصلة. ويمكن استخدام أساليب تقييس بديلة لمقاييس النوعية الفيديوية الواردة في الملحق شريطة أن توفر دقة التقييس المطلوبة.

وإذ تستند مقاييس النوعية الفيديوية عادة إلى تقريبات الاستجابات البصرية البشرية بدلاً من قياس مصنوعات تشفير محددة، فهي من حيث المبدأ صالحة بالقدر نفسه للأنظمة التماثلية والأنظمة الرقمية. كما تصلح من حيث المبدأ للسلاسل التي تختلط فيها الأنظمة التماثلية والرقمية، أو حيث تتسلسل أنظمة الضغط الرقمي.

ويظهر الشكل 2 مثالاً على تطبيق أسلوب المرجع المخفض لفحص سلسلة إرسال.

الشـــكل 2

تطبيق أسلوب قياس النوعية الحسي بمرجع مخفض لفحص سلسلة إرسال



BT.1885-02

في هذه الحالة، يغذى مفكك تشفير مرجعي من نقاط مختلفة في سلسلة الإرسال، فعلى سبيل المثال يمكن أن يكون مفكك التشفير موجوداً في نقطة في الشبكة ، كما في الشكل 2، أو في خرج المشفر مباشرة كما في الشكل 1. فإذا كانت سلسلة الإرسال الرقمي شفافة، كان تصنيف قياس النوعية الموضوعية للصورة في المصدر مساوياً للقياس في أي نقطة لاحقة في السلسلة.

ومن المسلم به عموماً أن أسلوب المرجع الكامل يوفر أفضل دقة للقياسات الحسية لنوعية الصورة. وقد ثبت أن هذا الأسلوب لديه القدرة على الارتباط الكبير مع التقييمات الشخصية المتوافقة مع أساليب التصنيف بحسب الفئات المطلقة ذات المرجع الخفي (ACR-HR) الموصفة في التوصية ITU-T P.910.

7 النتائج التي توصل إليها فريق الخبراء المعني بالجودة الفيديوية

تجرى دراسات على القياسات الحسية للنوعية الفيديوية لدى فريق غير رسمي يدعى فريق الخبراء المعني بالجودة الفيديوية (VQEG) وهو يقدم تقاريره للجنتي الدراسات 9 و12 في قطاع تقييس الاتصالات وللجنة الدراسات 6 في قطاع الاتصالات الراديوية. وقد استكمل هذا الفريق مؤخراً فحص RRNR-TV الذي قيّم فيه أداء خوارزميات القياس الحسية للنوعية الفيديوية بمرجع مخفض من أجل أنساق الصورة المذكورة في التوصية ITU-R 601-6.

واستناداً إلى الأدلة الحالية، يمكن لقطاع تقييس الاتصالات في هذه الآونة أن يوصي بستة أساليب ذات مرجع مخفض (النموذج-A 15 k، النموذج-A 80 k ، النموذج-A 256 k، النموذج-C 80 k، النموذج-B 80 k (525 خطاً فقط)، النموذج-B 256 k (525 خطاً فقط).

ويمكن الاطلاع على الوصف التقني لهذه النماذج في الملحقات من A حتى C على التوالي، علماً بأن ترتيب الملحقات يرد دون أي تعيين على الإطلاق ولا يقدم أي مؤشر على أداء التنبؤ بالنوعية.

ويظهر الجدولان 1 و2 اختبارات الدلالة في فحص VQEG RRNR-TV. ففي نسق 525 خطاً، تتفوق من الناحية الإحصائية أربعة نماذج (النموذج-A k 15، النموذج-A k 80 ، النموذج-A k 256، النموذج-C k 80) على نسبة ذروة الإشارة إلى الضوضاء (PSNR)، ويتساوى من الناحية الإحصائية نموذجان (النموذج-B k 80، النموذج-B k 256) مع نسبة ذروة الإشارة إلى الضوضاء. وتجدر الإشارة إلى أن الإدارة القومية الأمريكية للاتصالات والمعلومات (NTIA) حسبت هذه النسبة بواسطة البحث المستنفد لحدود المعايرة. وفي نسق 625 خطاً، تتساوى وتتفوق من الناحية الإحصائية أربعة نماذج (النموذج-A k 15، النموذج-A k 80 ، النموذج-A k 256، النموذج-C k 80) على نسبة ذروة الإشارة إلى الضوضاء (PSNR).

الجـــدول 1

اختبار الدلالة للنسق 525 خطاً

النسق 525 خطاً	يقارَن مع الأفضل	يقارَن مع PSNR	الترابط
النموذج–A k 15	1	1	0,906
النموذج–A k 80	1	1	0,903
النموذج–A k 256	1	1	0,903
النموذج–C k 80	1	1	0,882
النموذج–B k 80	0	1	0,795
النموذج–B k 256	0	1	0,803
PSNR_NTIA	0	1	0,826

الملاحظة 1 – تشير قيمة "1" في عمود "يقارَن مع الأفضل" أن هذا النموذج يعادل من الناحية الإحصائية النموذج الأعلى أداءً. وتشير قيمة "1" في عمود "يقارَن مع PSNR" أن هذا النموذج يعادل من الناحية الإحصائية النموذج الأعلى أداءً. وتشير قيمة "0" أن هذا النموذج لا يعادل من الناحية الإحصائية النموذج الأعلى أداءً.

الجـــدول 2

اختبار الدلالة للنسق 625 خطاً

النسق 625 خطاً	يقارَن مع الأفضل	يقارَن مع PSNR	الترابط
النموذج–A k 15	1	1	0,894
النموذج–A k 80	1	1	0,899
النموذج–A k 256	1	1	0,898
النموذج–C k 80	1	1	0,866
PSNR_NTIA	0	1	0,857

الملاحظة 1 – تشير قيمة "1" في عمود "يقارَن مع الأفضل" أن هذا النموذج يعادل من الناحية الإحصائية النموذج الأعلى أداءً. وتشير قيمة "1" في عمود "يقارَن مع PSNR" أن هذا النموذج يعادل من الناحية الإحصائية النموذج الأعلى أداءً. وتشير قيمة "0" أن هذا النموذج لا يعادل من الناحية الإحصائية النموذج الأعلى أداءً.

ويوفر الجدولان 3 و4 معلومات تفصيلية بشأن أداء النماذج في فحص VQEG RRNR-TV.

الجـــدول 3

وصف بالمعلومات لأداء النماذج في فحص VQEG RRNR-TV
(النسق 525 خطاً)

النسق 525 خطاً	الترابط	RMSE	OR
النموذج A 15	0,906	0,418	0,385
النموذج A 80 k	0,903	0,423	0,378
النموذج A 256 k	0,903	0,424	0,378
النموذج C 80 k	0,795	0,598	0,667
النموذج B 80 k	0,803	0,587	0,647
النموذج B 256 k	0,882	0,465	0,513
PSNR_NTIA	0,826	0,556	0,571

الجـــدول 4

وصف بالمعلومات لأداء النماذج في فحص VQEG RRNR-TV
(النسق 625 خطاً)

النسق 625 خطاً	الترابط	RMSE	OR
النموذج A 15	0,894	0,524	0,468
النموذج A 80 k	0,899	0,513	0,462
النموذج A 256 k	0,898	0,516	0,468
النموذج C 80 k	0,866	0,585	0,583
PSNR_NTIA	0,857	0,605	0,564

الملحق A

النموذج A: أسلوب جامعة يونسي بالمرجع المخفض

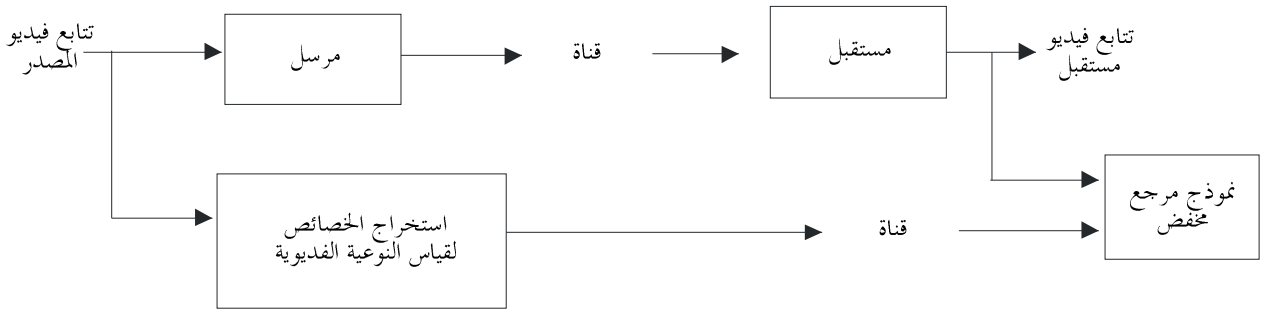
1 مقدمة

رغم استخدام نسبة ذروة الإشارة إلى الضوضاء (PSNR) على نطاق واسع كمقياس موضوعي للنوعية الفيديوية، فقد قيل أيضاً أنها لا تمثل النوعية الفيديوية على خير ما يرام. وبتحليل الكيفية التي يتحسس بها البشر بالنوعية الفيديوية، لوحظ أن النظام البصري البشري يتحسس للتردي حول الحواف. وبعبارة أخرى، عندما تُغشى بيكسلات الحواف لعرض فيديوي، يميل المقيمون لتقييمه تقييماً متدنياً حتى ولو كانت نسبة ذروة الإشارة إلى الضوضاء عالية. وبناء على هذه الملاحظة، وُضعت نماذج بمرجع مخفض تقيس تردي الحواف أساساً.

ويوضح الشكل 3 كيفية عمل نموذج المرجع المخفض. وتُستخرج الخصائص التي ستُستخدم لقياس النوعية الفيديوية في نقطة مراقبة من التتابع الفيديوي في المصدر، ويجري إرسالها. ويبين الجدول 5 عروض نطاق القناة الجانبية لهذه الخصائص، وهي عروض نطاق جرى اختبارها في فحص VQEG RRNR-TV.

الشـــكل 3

مخطط وظيفي لأسلوب المرجع المخفض



BT.1885-03

الجـــدول 5

عروض نطاق القناة الجانبية

النسق الفيديوي	عروض النطاق المختبرة
نسق 525	kbit/s 15، kbit/s 80، kbit/s 256
نسق 625	kbit/s 15، kbit/s 80، kbit/s 256

2 نماذج نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) بمرجع مخفض

1.2 نسبة ذروة الإشارة إلى الضوضاء (PSNR) في الحواف

تقيس نماذج المرجع المخفض أساساً التردي في الحواف. وفي هذه النماذج، تُطبق خوارزمية كشف الحافة أولاً على التتابع الفيديوي في المصدر لتحديد موقع بيكسلات الحواف. ثم يقاس تردي بيكسلات الحواف تلك بحساب متوسط الخطأ المربع، ومنه تُحسب نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR).

ويمكن استخدام أي خوارزمية للكشف عن الحافة على الرغم من الاختلافات البسيطة التي قد تظهر في النتائج. فعلى سبيل المثال، يمكن استخدام أي مشغّل تدرج لتحديد موقع بيكسلات الحواف. وقد اقترح عدد من مشغلي التدرج في هذا الصدد. وفي العديد من خوارزميات الكشف عن الحواف، تُحسب أولاً صورة التدرج الأفقي (m,n) وصورة التدرجي الرأسي (m,n) بواسطة مشغلي التدرج. ثم يمكن حساب صورة تدرج اتساع $g(m,n)$ على النحو التالي:

$$g(m,n)=|g_{horizontal}(m,n)|+|g_{vertical}(m,n)|$$

وأخيراً، يُطبَق تشغيل العتبة على صورة تدرج اتساع $g(m,n)$ للعثور على بيكسلات الحواف. وبعبارة أخرى، تُعتبر البيكسلات التي يتجاوز تدرج اتساعها قيمة العتبة بيكسلات الحواف.

ويتضح الإجراء في الأشكال من 4 إلى 8. فيظهر الشكل 4 صورة مصدر. ويظهر الشكل 5 صورة التدرج الأفقي $g_{horizontal}(m,n)$ التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الأفقي على صورة المصدر في الشكل 4. ويظهر

الشكل 6 صورة التدرج الرأسي ($g_{vertical}(m,n)$) التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الرأسي على صورة المصدر في الشكل 4. ويظهر الشكل 7 صورة تدرج الاتساع (صورة الحافة)، فيما يُظهر الشكل 8 صورة الحافة الثنائية (صورة قناع) التي تم الحصول عليها من خلال تطبيق العتبة على صورة تدرج الاتساع في الشكل 7.

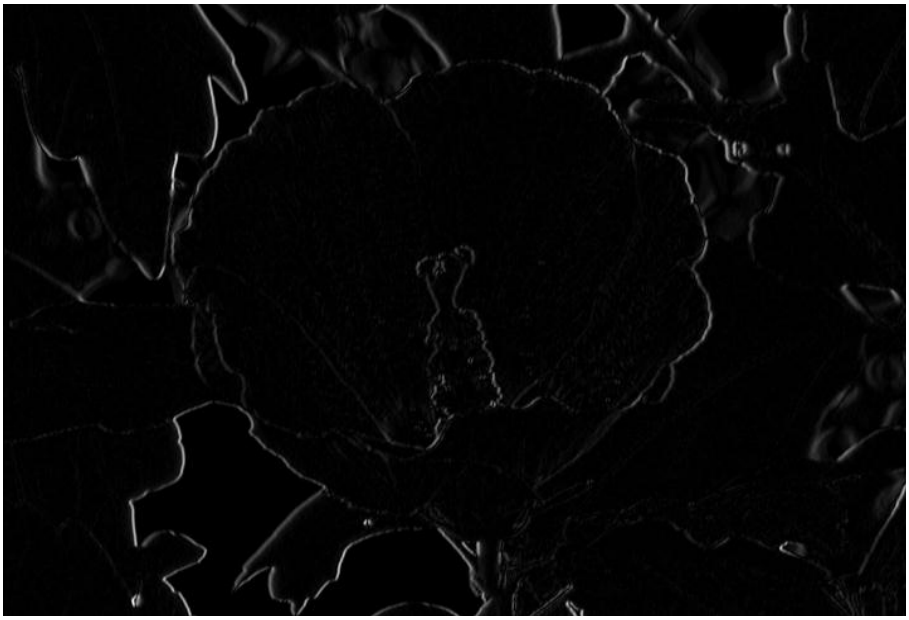
الشـــكل 4

صورة المصدر (الصورة الأصلية)

BT.1885-04

الشـــكل 5

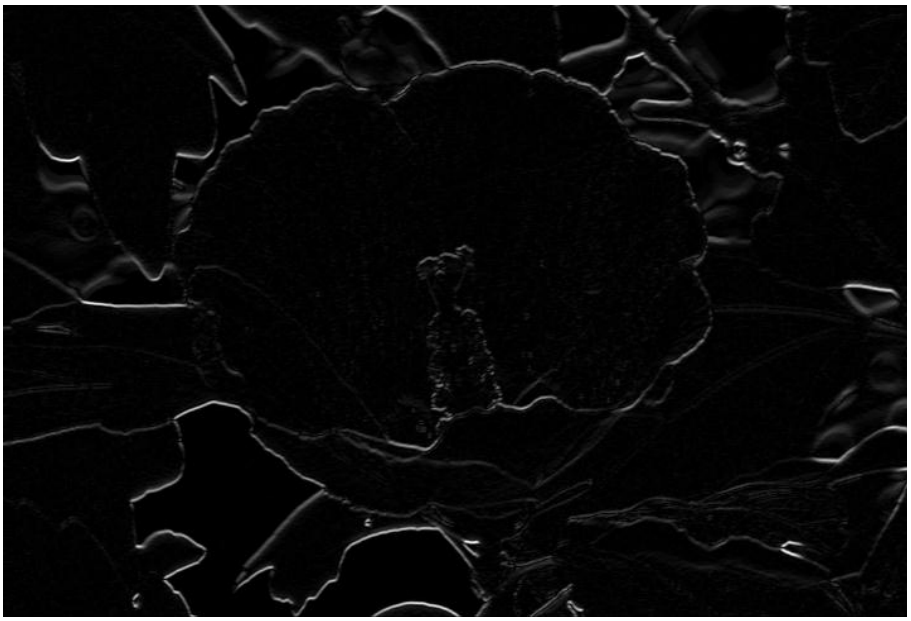
صورة التدرج الأفقي التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الأفقي على صورة المصدر في الشكل 4



BT.1885-05

الشـــكل 6

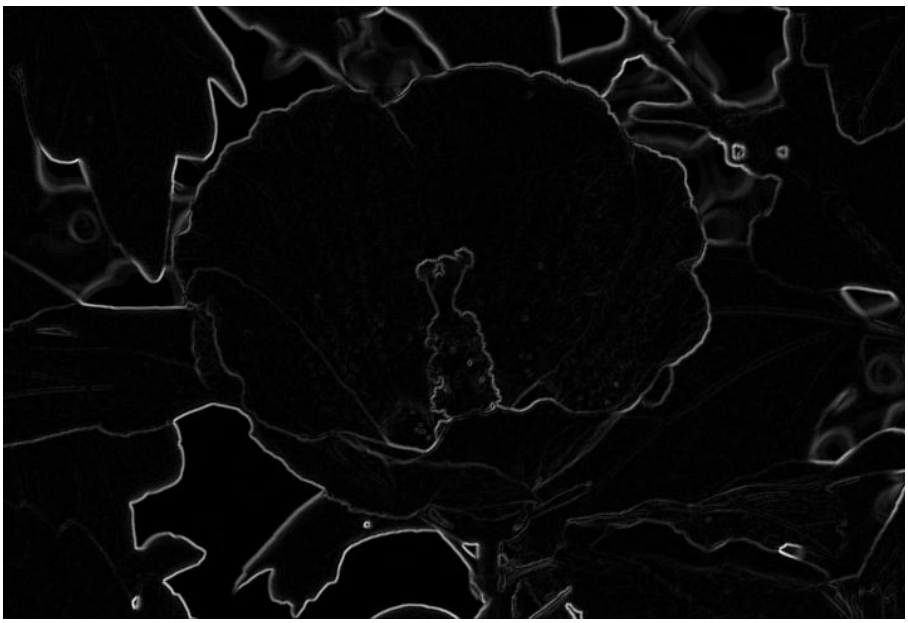
صورة التدرج الرأسي التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الرأسي على صورة المصدر في الشكل 4



BT.1885-06

الشـــكل 7

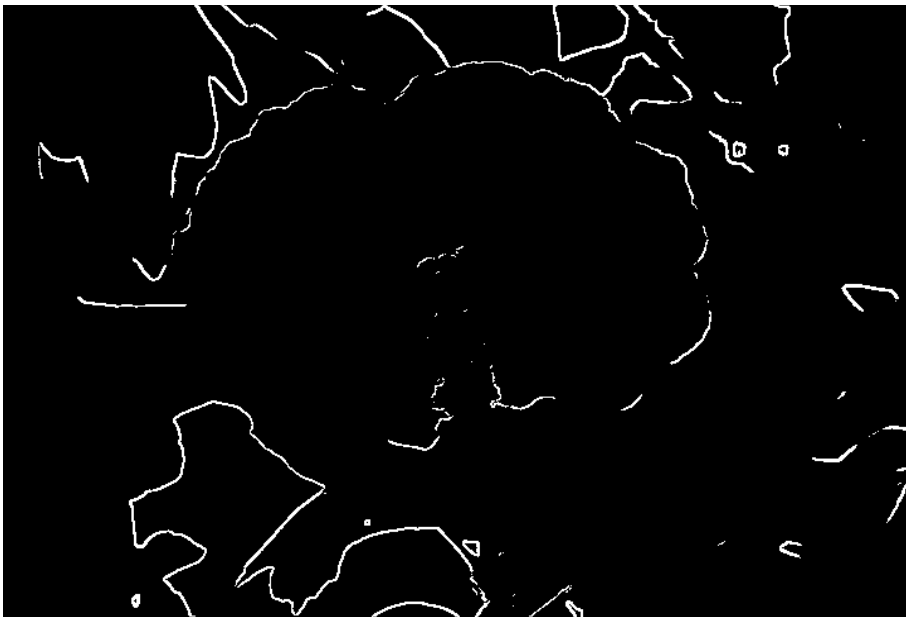
صورة تدرج الاتساع



BT.1885-07

الشـــكل 8

صورة الحافة الثنائية (صورة قناع) التي تم الحصول عليها من خلال تطبيق العتبة على صورة تدرج الاتساع في الشكل 7



BT.1885-08

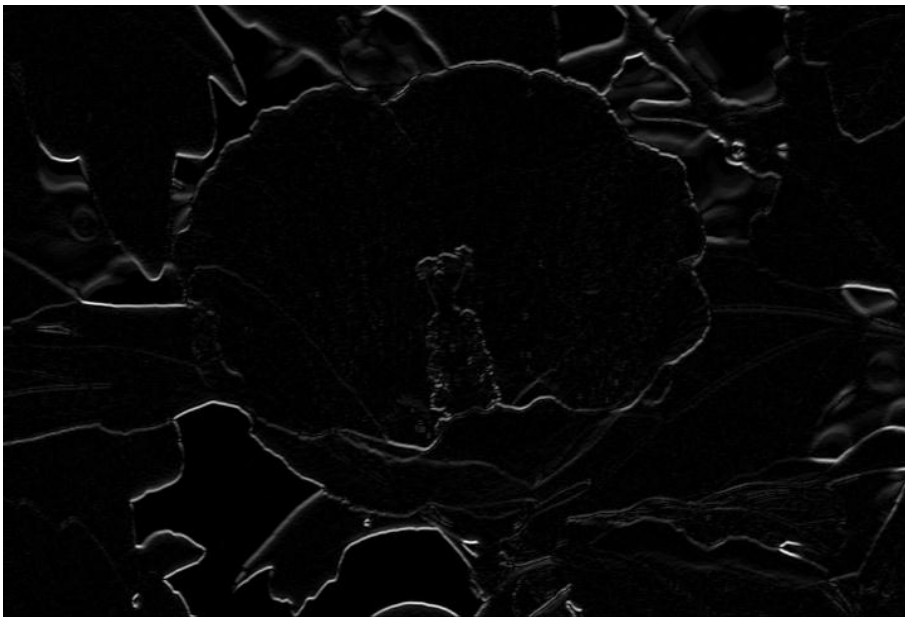
وبدلاً من ذلك، يمكن استخدام إجراء معدل لإيجاد بيكسلات الحواف. فعلى سبيل المثال، يمكن أولاً تطبيق مشغّل التدرج الرأسي على صورة المصدر لإنتاج صورة تدرج رأسي. ثم يطبَّق مشغّل تدرج أفقي على صورة التدرج الرأسي لإنتاج صورة تدرج متوالي معدلة (صورة التدرج الأفقي والرأسي). وأخيراً، يمكن تطبيق تشغيل العتبة على صورة التدرج المتوالي المعدلة لإيجاد بيكسلات الحواف. وبعبارة أخرى، تعتبر بيكسلات صورة التدرج المتوالي المعدلة التي تتجاوز قيمة، بيكسلات حواف. وتوضح الأشكال 9 حتى 12 الإجراء المعدل. ويظهر الشكل 9 صورة التدرج الرأسي $g_{vertical}(m,n)$ التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الرأسي على صورة المصدر في الشكل 4. ويظهر الشكل 10 صورة تدرج متوالي معدلة (صورة التدرج الأفقي والرأسي) التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الأفقي على صورة التدرج الرأسي في الشكل 9. فيما يُظهر الشكل 11 صورة الحافة الثنائية (صورة قناع) التي تم الحصول عليها من خلال تطبيق العتبة على صورة التدرج المتوالي المعدلة في الشكل 10.

وتجدر الإشارة إلى إمكانية فهم كلا الأسلوبين كخوارزمية كشف حواف. ويمكن اختيار أي خوارزمية لكشف الحواف تبعاً لطبيعة الإشارات الفيديوية وخوارزميات الضغط. بيد أن بعض الأساليب قد تتفوق على غيرها في الأداء.

وهكذا، يطبَّق في هذا النموذج مشغّل كشف الحواف أولاً لإنتاج صور الحواف (انظر الشكلين 7 و10). ثم تنتَج صورة القناع (صورة الحافة الثنائية) من خلال تطبيق العتبة على صورة الحافة (انظر الشكلين 8 و11). وبعبارة أخرى، تُسنَد قيمة الصفر لبيكسلات صورة الحافة التي تقل قيمتها عن العتبة te، وتسنّد قيمة غير صفرية للبيكسلات التي تساوي أو تزيد قيمتها قيمة العتبة. ويبين الشكلان 8 و11صور القناع. وبما أن مشاهدة الفيديو ممكنة في تتابع للأطر أو المجالات، فإن الإجراء المذكور أعلاه يمكن أن يطبَّق على كل إطار أو مجال فيديوي. وباعتبار أن النموذج يمكن استخدامه للفيديوهات القائمة على المجالات أو تلك القائمة على الأطر، فإن مصطلح "الصورة" سيُستخدم للدلالة على مجال أو إطار.

الشـــكل 9

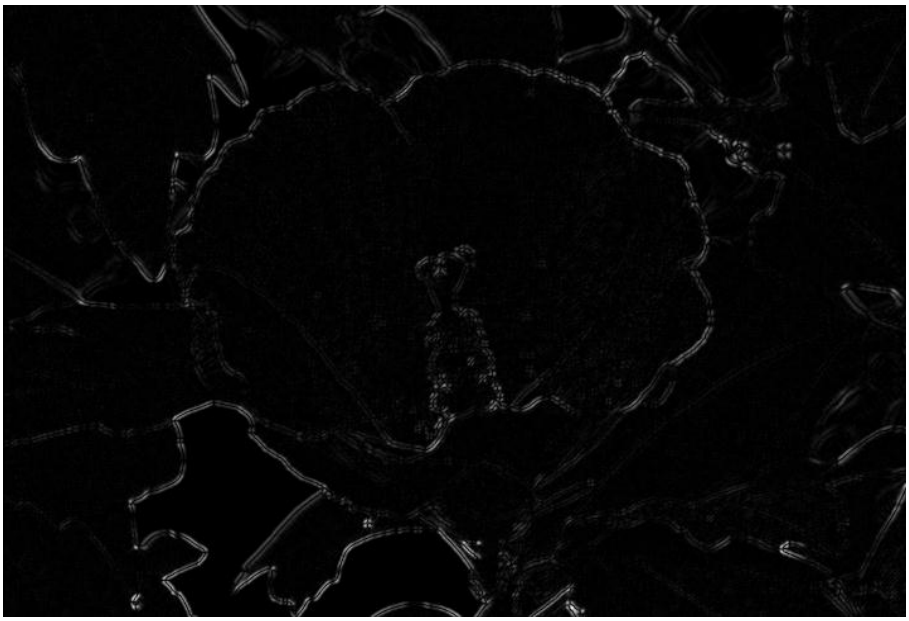
صورة التدرج الرأسي التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الرأسي على صورة المصدر في الشكل 4



BT.1885-09

الشـــكل 10

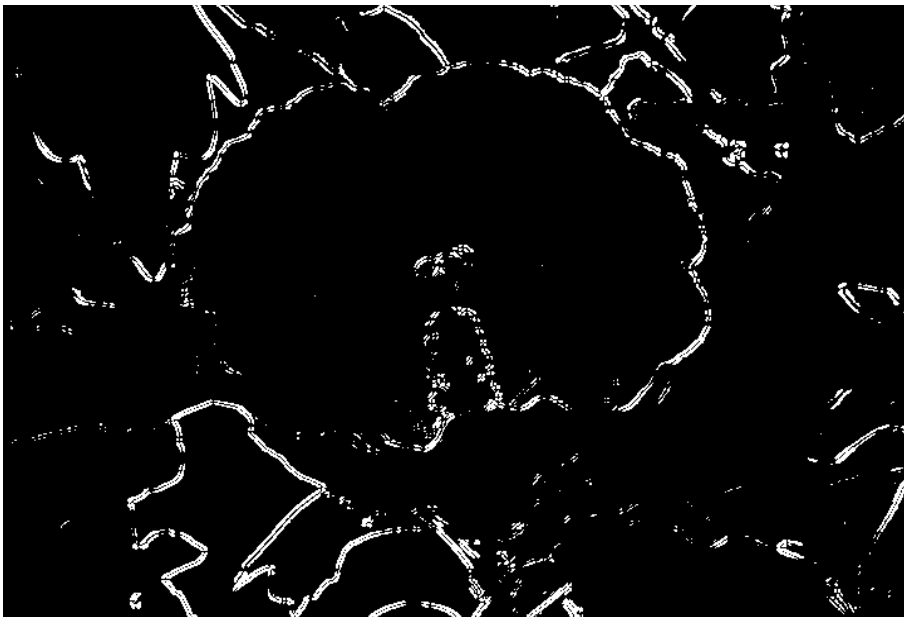
صورة تدرج متوالي معدلة (صورة التدرج الأفقي والرأسي) التي يتم الحصول عليها عن طريق تطبيق مشغّل التدرج الأفقي على صورة التدرج الرأسي في الشكل 9.



BT.1885-10

الشـــكل 11

صورة الحافة الثنائية (صورة قناع) التي تم الحصول عليها من خلال تطبيق العتبة على صورة التدرج المتوالي المعدلة في الشكل 10



BT.1885-11

2.2 اختيار الخصائص من تتابعات فيديو المصدر

بما أن هذا النموذج هو نموذج المرجع المخفض، فإن الحاجة تدعو لاستخراج مجموعة خصائص من كل صورة لتتابع فيديو المصدر. وفي نموذج نسبة ذروة الإشارة إلى الضوضاء في الحواف بمرجع مخفض (EPSNR RR)، يُختار عدد معين من بيكسلات الحافة من كل صورة. ثم تشفَّر المواقع وقيم البيكسلات وترسَل. ولكن في بعض تتابعات الفيديو، قد يقل عدد بيكسلات الحواف كثيراً عند استخدام قيمة عتبة ثابتة. وفي أسوأ سيناريو، قد ينعدم هذا العدد (صور فارغة أو صور ذات تردد منخفض جداً). ولمعالجة هذه المشكلة، إذا كان عدد بيكسلات حافة الصورة أصغر من قيمة معينة، يمكن للمستخدم خفض قيمة العتبة حتى يزيد عدد بيكسلات الحافة عن قيمة معينة. وبدلاً من ذلك، يمكن اختيار بيكسلات الحافة التي تقابل التي تقابل كبرى قيم صورة التدرج الأفقي والرأسي. وعندما تنعدم بيكسلات الحواف (كما في صور فارغة) في إطار، يمكن اختيار العدد المطلوب من البيكسلات عشوائياً أو تخطي الإطار. فعلى سبيل المثال، إذا اختيرت 10 بيكسلات حواف من كل إطار، يمكن فرز بيكسلات صورة التدرج الأفقي والرأسي وفقاً لقيمها واختيار أكبر 10 قيم. غير أن هذا الإجراء قد ينتج بيكسلات حواف متعددة في مواقع متطابقة. ولمعالجة هذه المشكلة ، يمكن أولاً اختيار العدد المرغوب فيه لبيكسلات صورة التدرج الأفقي والرأسي مراراً، ثم اختيار العدد المرغوب فيه لبيكسلات الحواف، لا على التعيين، من بين البيكسلات المختارة في صورة التدرج الأفقي والرأسي. ففي النماذج المختارة في فحص VQEG RRNR-TV، يُختار العدد المرغوب فيه لبيكسلات الحواف عشوائياً من بين مجموعة كبيرة من بيكسلات الحواف التي تُحصَّل بتطبيق عملية العتبة على صورة التدرج.

وفي نماذج نسبة ذروة الإشارة إلى الضوضاء في الحواف بمرجع مخفض (EPSNR RR)، تشفَّر المواقع وقيم البيكسلات بعد تطبق مرشاح تمرير منخفض غاوسي (Gaussian) على مواقع البيكسلات المختارة. ورغم أن مرشاح التمرير المنخفض الغاوسي (3 × 5) كان قد استُخدم في فحص VQEG RRNR-TV، يمكن استخدام مراشيح تمرير منخفض أخرى تبعاً لأنساق الفيديو. وتجدر الإشارة إلى إمكانية تطبيق التشذيب أثناء عملية التشفير. وتفادياً لاختيار بيكسلات الحواف في المساحات المشذبة، يختار النموذج بيكسلات الحواف في المساحة الوسطى (انظر الشكل 12). ويبين الجدول 6 المقاسات بعد التشذيب، كما يبين عدد البتات المطلوبة لتشفير الموقع وقيمة البيكسل في حافة.

الجـــدول 6

البتات المطلوبة لكل بيكسل في الحواف

نسق الفيديو	المقاس	المقاس بعد التشذيب	بتات للموقع	بتات لقيمة البيكسل	مجمل البتات في كل بيكسل
525	720 × 486	656 × 438	19	8	27
625	720 × 576	656 × 528	19	8	27

الشـــكل 12

مثال على التشذيب والمساحة الوسطى

BT.1885-12

يختار هذا النموذج بيكسلات الحواف من كل إطار وفقاً لعرض النطاق المسموح به (انظر الجدول 5). ويبين الجدول 7 عدد بيكسلات الحواف في كل إطار التي يمكن إرسالها ضمن عروض النطاق المختبرة.

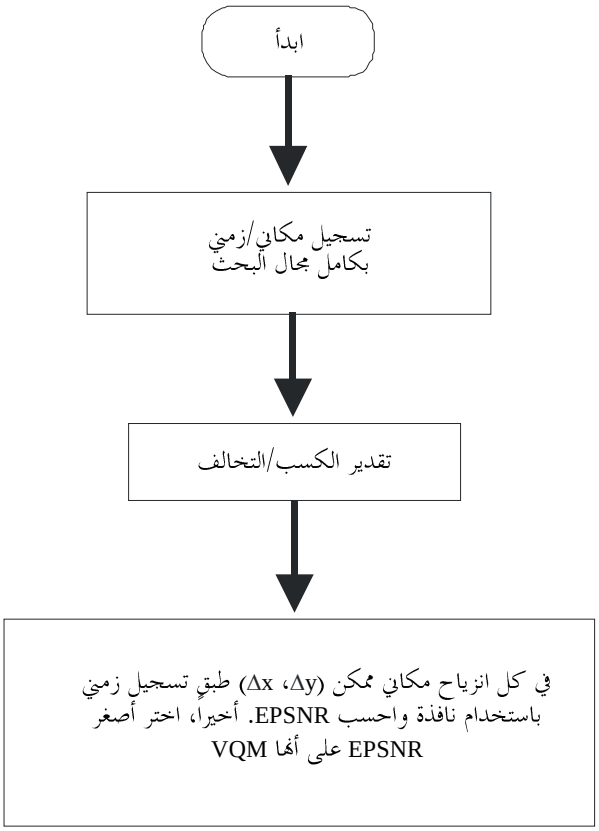
الجـــدول 7

عدد بيكسلات الحواف في كل إطار

نسق الفيديو	**15 kbit/s**	**80 kbit/s**	**256 kbit/s**
525	16	74	238
625	20	92	286

الشـــكل 13

المخطط الانسيابي للنموذج



BT.1885-13

3.2 التسجيل المكاني/الزمني وتعديل الكسب/التخالف

قبل حساب الفرق بين بيكسلات الحواف لتتابع فيديو المصدر وبيكسلات الحواف لتتابع الفيديو المعالَج للحصول على تتابع الفيديو المستقبَل في جهاز الاستقبال، يطبق النموذج أولاً التسجيل المكاني/الزمني وتعديل الكسب/التخالف. واستُخدم لهذا الغرض أسلوب المعايرة (الملحق B) المذكور في التوصية ITU-T J.244. ولنقل خصائص الكسب والتخالف المذكورة في التوصية ITU-T J.244 (الملحق B)، استُخدم في فحص VQEG RRNR-TV، 30% من عروض النطاق المتاحة. ونظراً لتشابك تتابع الفيديو، يُطبَّق أسلوب المعايرة ثلاث مرات: على المجالات الزوجية وعلى المجالات المفردة وعلى الأطر الجامعة بين المجالات الزوجية والمفردة. فإذا ما زاد الفرق بين خطأ المجال الزوجي (PSNR) وخطأ المجال المفرد عن عتبة معينة، استُخدمت نتائج التسجيل (انزياح x، انزياح y) ذات نسبة ذروة الإشارة إلى الضوضاء (PSNR) الأدنى. وخلاف ذلك، استُخدمت نتائج التسجيل في الأطر الجامعة بين المجالات الزوجية والمفردة. وفي فحص VQEG RRNR-TV، حُددت العتبة بقيمة 2 dB.

وعند نقطة المراقبة، ينبغي محاذاة تتابع الفيديو المعالَج مع بيكسلات الحواف المستخرجة من تتابع فيديو المصدر. ولكن إذا كان عرض نطاق القناة الجانبية ضيقاً، لن يتوفر إلا القليل من بيكسلات الحواف في تتابع فيديو المصدر (انظر الشكل 14). وبالتالي، قد لا يصح التسجيل الزمني إذا ما أُجري باستخدام إطار واحد (انظر الشكل 15). ولمعالجة هذه المشكلة، يستخدم النموذج نافذة للتسجيل الزمني. وبدلاً من استخدام إطار واحد من تتابع الفيديو المعالَج، يبني النموذج نافذة تتألف من عدد

من الأطر المتاخمة للعثور على الانزياح الزمني الأمثل. ويوضح الشكل 16 هذا الإجراء. ويُحسب متوسط الخطأ المربع ضمن النافذة على النحو التالي:

$$MSE_{window} = \frac{1}{N_{win}} \sum (E_{SRC}(i) - E_{PVS}(i))^2$$

حيث:

MSE_{window}: متوسط الخطأ المربع

$E_{SRC}(i)$: بيكسل تتابع الفيديو المعالج المقابل لبيكسل الحافة

$E_{PVS}(i)$: بيكسل حواف ضمن النافذة يقابل بيكسل في تتابع الفيديو المعالَج

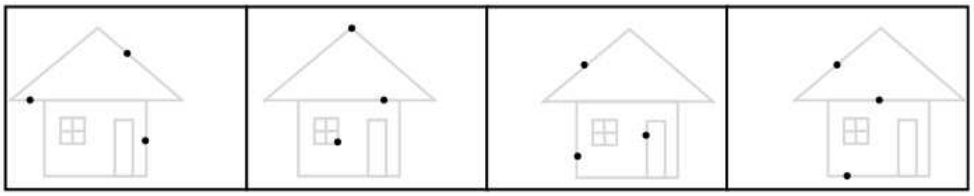
N_{win}: العدد الإجمالي لبيكسلات الحواف المستخدم لحساب MSE_{window}.

يُستخدم متوسط الخطأ المربع ضمن النافذة هذا بمثابة الفرق بين إطار التتابع الفيديو المعالَج وما يقابله من إطار تتابع فيديو المصدر.

ويمكن تحديد مقاس النافذة بالنظر في طبيعة تتابع الفيديو المعالَج. ففي تطبيق نمطي، يوصى بنافذة تقابل ثانيتين. وبدلاً من ذلك، يمكن تطبيق نوافذ زمنية ذات مقاسات مختلفة، حيث يمكن استخدام النافذة الفضلى التي توفر أصغر خطأ فعال. وعلاوة على ذلك، يمكن استخدام مراكز نافذة مختلفة للنظر في تخطي الأطر بفعل أخطاء الإرسال (انظر الشكل 20).

الشـــكل 14

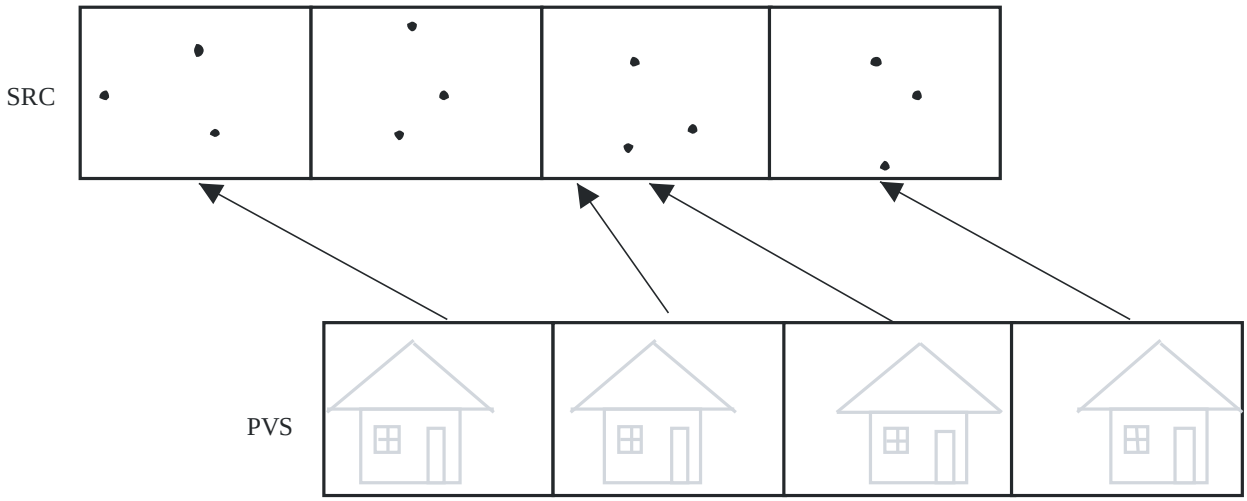
اختيار بيكسلات الحواف في تتابع فيديو المصدر



BT.1885-14

الشـــكل 15

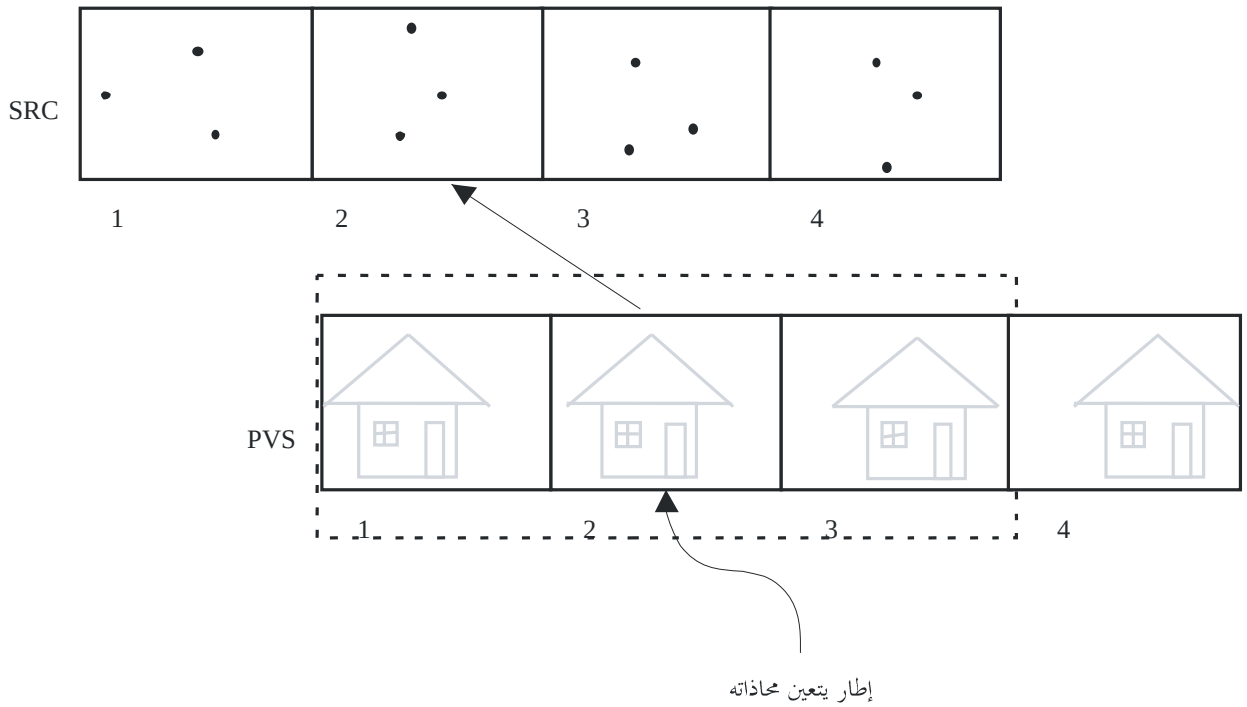
محاذاة تتابع الفيديو المعالَج مع بيكسلات تتابع فيديو المصدر



BT.1885-15

الشـــكل 16

محاذاة تتابع الفيديو المعالَج مع بيكسلات الحواف باستخدام نافذة

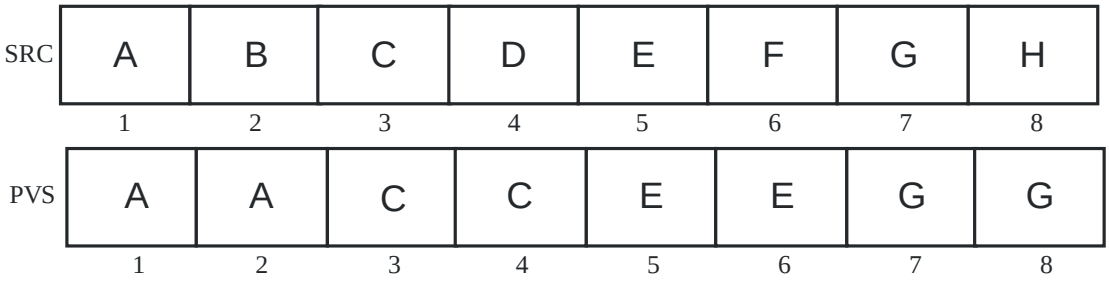


BT.1885-16

عند تشفير تتابع فيديو المصدر بنسب ضغط عالية، يمكن للمشفر أن يخفض عدد الأطر في الثانية وتتكرر الأطر في تتابع الفيديو المعالَج (انظر الشكل 17). وفي الشكل 17، ليس لتتابع الفيديو المعالَج أطرٍ تقابل بعض الأطر في تتابع فيديو المصدر (الإطار الثاني والرابع والسادس والثامن). وفي هذه الحالة، لا يستخدم النموذج أطراً متكررة في حساب متوسط الخطأ المربع. وبعبارة أخرى، فإن النموذج ينفذ التسجيل الزمني باستخدام الإطار الأول (الإطار الصالح) من كل فدرة متكررة. ومن ثم، لا تُستخدم للتسجيل الزمني ضمن النافذة في الشكل 18 إلا ثلاثة أطر (الإطار الثالثة والخامس والسابع).

الشـــكل 17

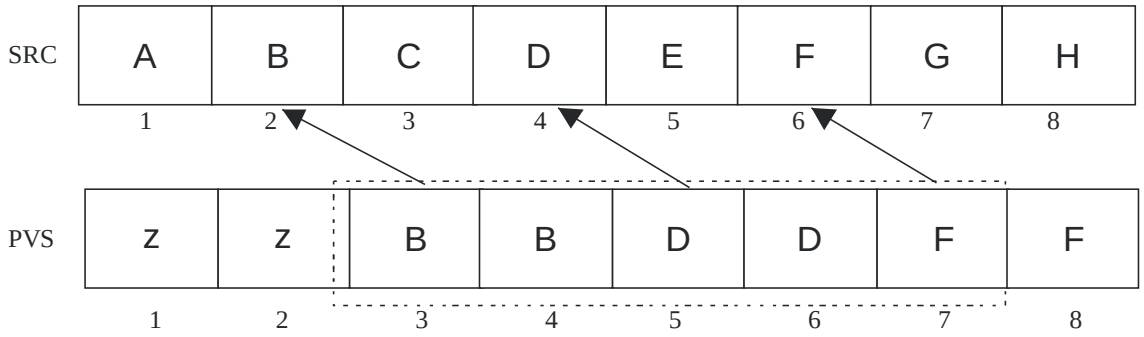
مثال على الأطر المتكررة



BT.1885-17

الشـــكل 18

تمرير الأطر المتكررة

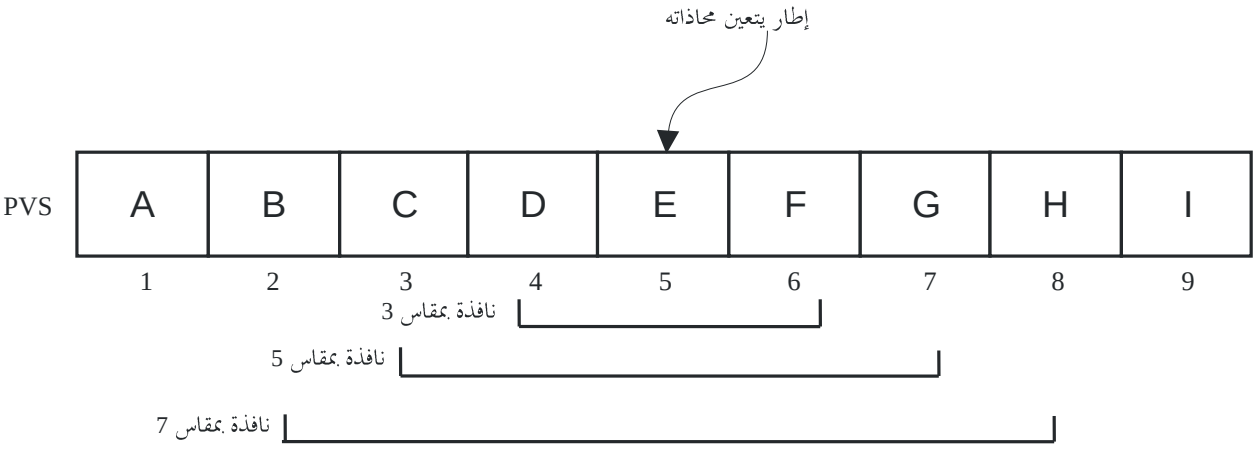


BT.1885-18

ومن الممكن أن يكون هناك تتابع الفيديو المعالَج بتكرار أطر غير منتظم مما قد يُسفر عن نتائج غير دقيقة في أسلوب التسجيل الزمني الذي يستخدم نافذة. ولمعالجة هذه المشكلة، يمكن ضبط كل إطار نافذة محلياً ضمن قيمة معينة (± 1 مثلاً) كما هو مبين في الشكل 21 بعد التسجيل الزمني الذي يستخدم نافذة. ثم، يُستخدم الضبط المحلي الذي يوفر الحد الأدنى من متوسط الخطأ المربع لحساب نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR).

الشـــكل 19

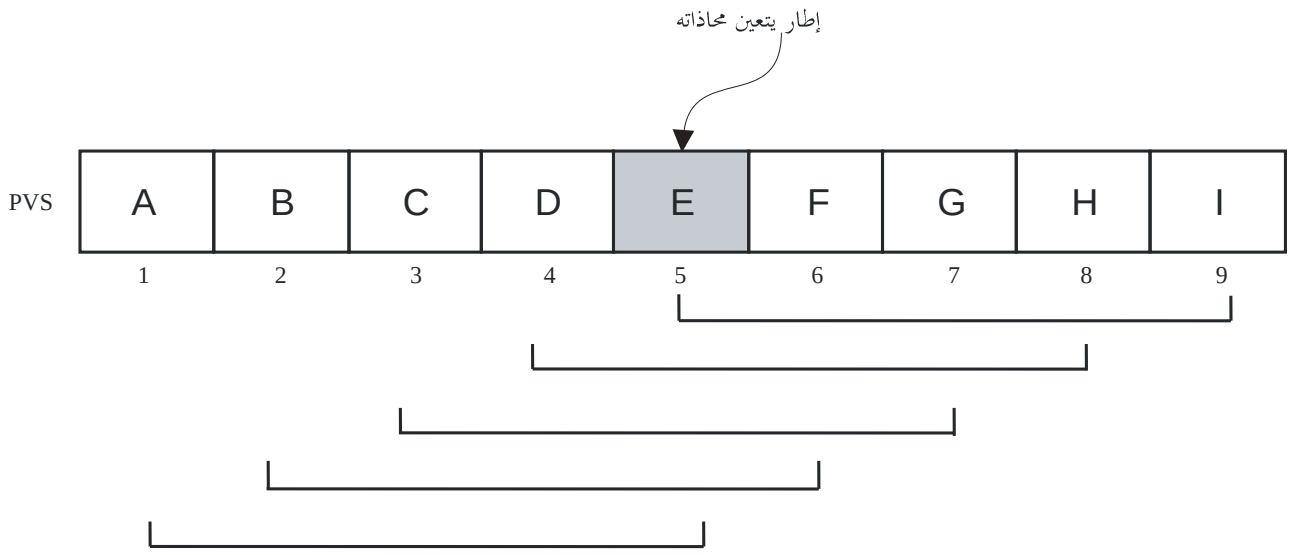
نوافذ بمختلف المقاسات



BT.1885-19

الشـــكل 20

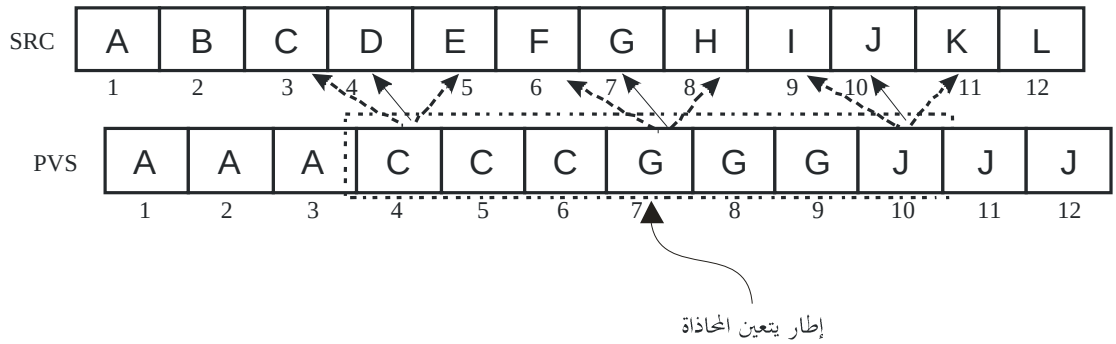
مراكز النافذة



BT.1885-20

الشـــكل 21

الضبط المحلي للتسجيل الزمني الذي يستخدم نافذة



BT.1885-21

4.2 حساب نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) وما بعد المعالجة

بعد إجراء التسجيل الزمني، يُحسب متوسط الفروق بين بيكسلات الحواف لتتابع فيديو المصدر وبين ما يقابلها من بيكسلات تتابع الفيديو المعالَج، وهو المتوسط الذي يعبر عن متوسط الخطأ المربع لتتابع الفيديو المعالَج في الحواف (MSEedge). وأخيراً، تُحسب نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) على النحو التالي:

$$EPSNR = 10\log_{10}(\frac{P^2}{MSE_{edge}})$$

حيث p هي قيمة الذروة للصورة.

1) الأطر المجمدة

قد تتكرر بعض الأطر جراء انخفاض معدلات الأطر، وقد تتجمد أطر بسبب خطأ في الإرسال، مما يؤدي إلى تردي النوعية الفيديوية الحسية. ولمعالجة ذلك، يطبق النموذج التعديل الحالي قبل حساب نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR):

$$MSE_{freezed_frame_considered} = MSE_{edge} \times \frac{K \times N_{total_frame}}{N_{total_frame} - N_{total_freezed_frame}}$$

حيث:

$MSE_{freezed_frame_considered}$: متوسط الخطأ المربع الذي يأخذ الأطر المتكررة والمتجمدة في الحسبان

N_{total_frame}: العدد الكلي للأطر، $N_{total_freezed_frame}$

K: ثابت.

وفي النموذج الذي اختُبر في فحص VQEG RRNR-TV، أُسندت قيمة 1 إلى K.

2) التردد العالي والحركة السريعة

إذا كان التتابع الفيديوي يحتوي على كمية كبيرة من الترددات العالية والحركات السريعة، تميل النوعية الحسية إلى الارتفاع في ظل متوسط الخطأ المربع نفسه. وللنظر في هذه الظاهرة، يُعرَّف فرق الأطر المقيّس (NFD) وطاقة الترددات العالية (NHFE) على النحو التالي:

$$NFD = \frac{FD}{average\ energy\ \ per\ \ pixel}$$

حيث $FD = \frac{1}{N_F} \sum_{i} \sum_{k=1}^{height} \sum_{j=1}^{width} (Frame_i[j,k] - Frame_{i-1}[j,k])^2$ و N_F هو عدد الأطر المستخدمة في الجمع. ويشار إلى استبعاد ثلاثة أكبر فروق في الأطر لدى حساب هذا العدد (FD) من أجل استبعاد تغير المشاهد من حساب متوسط فرق الأطر، على افتراض تتابعات فيديو مدة كل منها 8 ثوان. وتُحسب الطاقة المقيّسة للترددات العالية (NHFE) بحساب متوسط طاقات الترددات العالية (انظر الشكل 22) بعد تطبيق تحويل فوريه ثنائي الأبعاد:

$$NHFE = \frac{average\ \ high\ \ frequency\ \ energies}{average\ energy\ \ per\ \ pixel}$$

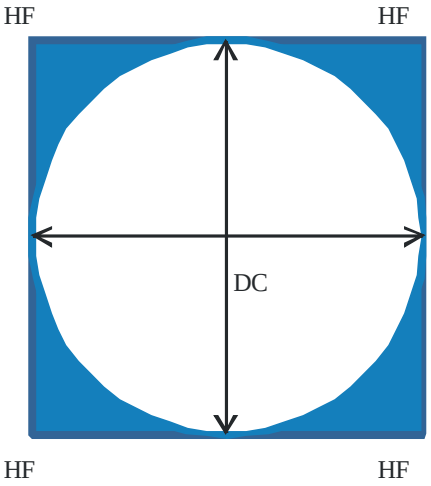
وأخيراً، تُستخدم المعادلات التالية:

```
IF(SNFD > 0.35 && SNHFE > 2.5)        {
                    IF(EPSNR < 20) EPSNR = EPSNR+3
                    ELSE IF(EPSNR < 35) EPSNR = EPSNR+5
}
ELSE IF((SNFD > 0.2 && SNHFE > 1.5) || (SNFD>0.27) && SNHFE > 1.3))      {
      I             F(28 < EPSNR < 40) EPSNR = EPSNR + 3
                    IF(EPSNR > 40) EPSNR = 40
}
```

حيث SNFD هو فرق الأطر المقيّس (NFD) في المصدر وSNHFE هو طاقة الترددات العالية (NHFE) في المصدر. ويجري حساب SNFD وSNHFE من الدارة أو القناة المرجعية للمصدر (SRC) ويرسلان كجزء من بيانات الخصائص (بايتة واحدة لكل منهما).

الشـــكل 22

حساب الطاقة المقيّسة للترددات العالية (NHFE)، تُحسب طاقات الترددات العالية من المساحة المظللة



BT.1885-22

3) الغشاوة

للنظر في مؤثرات الغشاوة، تستخدم المعادلات التالية:

```
IF (NHFE/SNHFE < 0.5)
        IF(EPSNR>26)        EPSNR = 26
ELSE IF (NHFE/SNHFE < 0.6)
        IF(EPSNR>32)        EPSNR = 32
ELSE IF (NHFE/SNHFE < 0.7)
        IF(EPSNR>36)        EPSNR = 36
ELSE IF (NHFE/SNHFE > 1.2)
        IF(EPSNR>23)        EPSNR = 23
ELSE IF (NHFE/SNHFE > 1.1)
        IF (EPSNR>25)       EPSNR = 25
```

حيث طاقة الترددات العالية (NHFE) هي هذه الطاقة في تتابع فيديوي معالَج (PVS NHFE).

4) السد

للنظر في مؤثرات السد، يُحسب متوسط فروق العمود. وعلى افتراض قابلية القسمة على 8، تُحسب علامة السد للإطار ذي الرتبة i على النحو التالي:

$$Blk[i]=\frac{largest\ column\ difference}{second\ largest\ column\ difference}$$

وتحسب علامة السد النهائية (*Blocking*) بحساب القيمة المتوسطة لعلامات سد الأطر:

$$Blocking=\frac{1}{number\ of\ \ frames}\sum_{i}Blk[i]$$

وأخيراً، تُستخدم المعادلات التالية:

```
IF(BLOCKING > 1.4)          {
                IF (20≤EPSNR<25) EPSNR = EPSNR-1.086094*BLOCKING-0.601316
                ELSE IF (EPSNR<30) EPSNR = EPSNR-0.577891*BLOCKING-3.158586
                ELSE IF (EPSNR<35) EPSNR = EPSNR-0.223573*BLOCKING-3.125441
}
```

5) الحد الأقصى للأطر المتجمدة

قد تسبب أخطاء الإرسال تجمد الأطر لفترة طويلة. وللنظر في تجمد الأطر لفترة طويلة، تُستخدم المعادلات التالية:

```
IF(MAX_FREEZE > 22 AND EPSNR>28) EPSNR = 28
ELSE IF(MAX_FREEZE > 10 AND EPSNR>34) EPSNR = 34
```

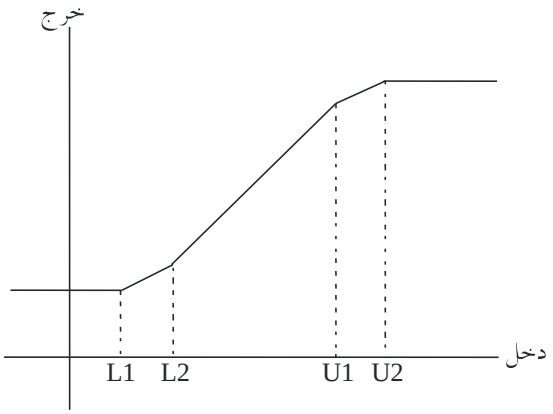
حيث MAX_FREEZE هي الأطر المتجمدة لأطول فترة. أما إذا لم تكن مدة التتابع الفيديوي 8 ثوان، فينبغي استخدام عتبات مختلفة.

6) الملاءمة الخطية المجتزأة

عندما تتجاوز نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) قيمة معينة، تصبح النوعية الحسية مشبعة. وفي هذه الحالة، يمكن ضبط الحد الأعلى من نسبة ذروة الإشارة إلى الضوضاء في الحواف. وعلاوة على ذلك، عند الرغبة بعلاقة خطية بين هذه النسبة ومتوسط الفرق في علامات الآراء (DMOS)، يمكن تطبيق دالة خطية مجتزأة على النحو الموضح في الشكل 23. وفي النموذج المختبر في فحص VQEG RRNR-TV، يحدَد الحد الأعلى بقيمة 48 والحد الأدنى بقيمة 15.

الشـــكل 23

دالة خطية مجتزأة لعلاقة خطية بين نسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) ومتوسط الفرق في علامات الآراء (DMOS)



BT.1885-23

إن النماذج ذات المرجع المخفض لنسبة ذروة الإشارة إلى الضوضاء في الحواف (EPSNR) تقوم على تردي الحواف. ويمكن تطبيق هذه النماذج في الوقت الفعلي باستخدام معتدل لقوة الحوسبة. وتلائم هذه النماذج التطبيقات التي تتطلب مراقبة النوعية الفيديوية في الوقت الفعلي عند توفر القنوات الجانبية.

الملحق B

النموذج B: نموذج NEC بمرجع مخفض

يوفر هذا الملحق وصفاً وظيفياً كاملاً لنموذج المرجع المخفض. ففي هذا النموذج، ترسَل إلى جانب العميل قيم النشاط بدلاً من قيم البيكسلات في فدر بيكسلات فردية ذات مقاس معين. وتقدَّر النوعية الفيديوية على أساس فرق النشاط بين القناة المرجعية للمصدر (SRC) وتتابع فيديوي معالَج (PVS). ويعود للترجيح البصري الفيزيولوجي فيما يتعلق بفرق النشاط تحسين دقة تقدير.

وهذا النموذج لا يحتاج إلى تسجيلات متطلبة من الناحية الحسابية للمكان والكسب والتخالف. وعلاوة على ذلك، يمكن تنفيذه ببرنامج ذي 30 خطاً وبرنامج ذي 250 خطاً على جانبي المخدم والعميل على التوالي. لذلك، فهو مناسب لمراقبة النوعية الفيديوية في الوقت الفعلي في الخدمات الإذاعية التي ستستفيد، أكثر ما تستفيد، من قلة التعقيد وسهولة التنفيذ.

1 ملخص

في نموذج المرجع المخفض، ترسَل إلى جانب العميل قيم النشاط بدلاً من قيم البيكسلات في فدر بيكسلات فردية ذات مقاس معين. وتقدَّر النوعية الفيديوية على أساس فرق النشاط بين القناة المرجعية للمصدر (SRC) وتتابع فيديوي معالَج (PVS). ويعود للترجيح البصري الفيزيولوجي فيما يتعلق بفرق النشاط تحسين دقة تقدير.

وهذا النموذج لا يحتاج إلى تسجيلات متطلبة من الناحية الحسابية للمكان والكسب والتخالف. وعلاوة على ذلك، يمكن تنفيذه ببرنامج ذي 30 خطاً وبرنامج ذي 250 خطاً على جانبي المخدم والعميل على التوالي. لذلك، فهو مناسب لمراقبة النوعية الفيديوية في الوقت الفعلي في الخدمات الإذاعية التي ستستفيد، أكثر ما تستفيد، من قلة التعقيد وسهولة التنفيذ.

2 تعاريف

نشاط – متوسط الفرق المطلق بين كل قيمة نصوع ومتوسط قيم النصوع في فدرة ذات مقاس معين.

فدرة – صفيف بيكسلات M x N (M عمود × N صف).

إطار – صورة واحدة تلفزيونية كاملة

كسب – عامل التوسع القياسي المضاعف الذي تطبقه الدارة المرجعية الافتراضية (HRC) على جميع بيكسلات مستوي صورة فردية (مثل النصوع والتلون). ويُعرف كسب إشارة النصوع عموماً بالتباين.

الدارة المرجعية الافتراضية (HRC) – نظام الفيديو تحت الاختبار مثل كودك أو نظام إرسال فيديوي رقمي.

النصوع (Y) – جزء من إشارة الفيديو التي تحمل في الغالب معلومات النصوع (أي الجزء الأبيض والأسود من الصورة).

لجنة أنظمة التلفزيون الوطنية (NTSC) – نظام الفيديو التماثلي الملون المركب ذو الخمسمائة وخمسة وعشرين خطاً [1].

التخالف أو تخالف المستوى – عامل مضاف تطبقه الدارة المرجعية الافتراضية على جميع بيكسلات مستوي صورة فردية (مثل النصوع والتلون). ويُعرف تخالف إشارة النصوع عموماً باللمعان.

نسبة ذروة الإشارة إلى الضوضاء (PSNR) – النسبة بين أقصى قدرة ممكنة لإشارة إلى قدرة الضوضاء المفسدة.

خط تبديل الطور (PAL) – نظام الفيديو التماثلي الملون المركب ذو الستمائة وخمسة وعشرين خطاً.

المسح التنقيطي – تقابل مخطط مستطيل ثنائي الأبعاد مع مخطط ذي بعد واحد بحيث يأتي المُدخل الأول في المخطط أحادي البعد من الصف الأول العلوي للمخطط ثنائي الأبعاد ممسوحاً من اليسار إلى اليمين، ويتبعه بالمثل الصف الثاني والثالث إلى آخر ما هنالك من صفوف في المخطط (هبوطاً) على أن يُمسح كل منها من اليسار إلى اليمين.

المرجع المخفض – منهجية لقياس النوعية الفيديوية تستخدم خصائص عرض النطاق الترددي الضيق المستخرج من تدفقات الفيديو الأصلية أو المعالَجة، بدلاً من استخدام فيديو ذي مرجعية كاملة يتطلب معرفة كاملة بتدفقات الفيديو الأصلية أو المعالَجة [2]. وتمتاز منهجيات المرجع المخفض بإمكانية مراقبة نوعية الخدمة من طرف إلى طرف نظراً لسهولة نقل معلومات المرجع المخفض عبر شبكات الاتصالات في كل مكان.

المنطقة المسترعية للاهتمام (ROI) – شبكة صورة (موصّفة بإحداثيات مستطيل) تُستخدم للدلالة على منطقة فرعية في مجال أو إطار الفيديو.

مشهد – تتابع من أطر الفيديو.

التسجيل المكاني – عملية تستخدم لتقدير وتصحيح الانزياحات المكانية لتتابع فيديوي معالَج بالنسبة إلى التتابع الفيديوي الأصلي.

التسجيل الزمني – عملية تستخدم لتقدير وتصحيح الانزياحات الزمنية لتتابع فيديوي معالَج (أي تأخر الفيديو) بالنسبة إلى التتابع الفيديوي الأصلي.

مقياس النوعية الفيديوية (VQM) – مقياس شامل لتردي الفيديو يعبَّر عنه كرقم واحد ويتراوح مدى خرجه الاسمي بين صفر وواحد، حيث يشير الصفر إلى انعدام التردي المحسوس والواحد إلى أقصى حد محسوس من التردي.

3 نظرة عامة على حساب مقياس النوعية الفيديوية

يرسل نموذج المرجع المخفض إلى جانب العميل قيم النشاط في فدر بيكسلات فردية ذات مقاس معين. وتشير هذه القيمة إلى تغاير قيم النصوع في الفدرة. ويلخص الشكل 1 نموذج المرجع المخفض. وكما يظهر في الشكل 24، تقدر النوعية الفيديوية على أساس فرق النشاط بين القناة المرجعية للمصدر (SRC) وتتابع فيديوي معالَج (PVS). وبالإضافة إلى ذلك، تُطبق الترجيحات البصرية الفيزيولوجية فيما يتعلق بفرق النشاط لتحسين دقة تقدير. ويجري تقدير النوعية الفيديوية في الخطوات التالية:

1) تُحسب قيمة النشاط لكل فدرة بيكسلات نصوع 16 × 16 للقناة المرجعية للمصدر (SRC) على جانب المخدم. وترسَل بعد ذلك كل قيم النشاط إلى جانب العميل. وتعرَّف قيمة النشاط لفدرة على أنها متوسط الفرق المطلق بين القيم الفردية ومتوسطها.

2) تُحسب قيم النشاط على جانب العميل فيما يتعلق بالتتابع الفيديوي المعالَج (PVS).

3) على جانب العميل، يتم تقييم كل فدرة في البداية بمربع خطأها أي مربع الفرق بين قيمتي نشاط القناة المرجعية للمصدر (SRC) وتتابع فيديوي معالَج (PVS).

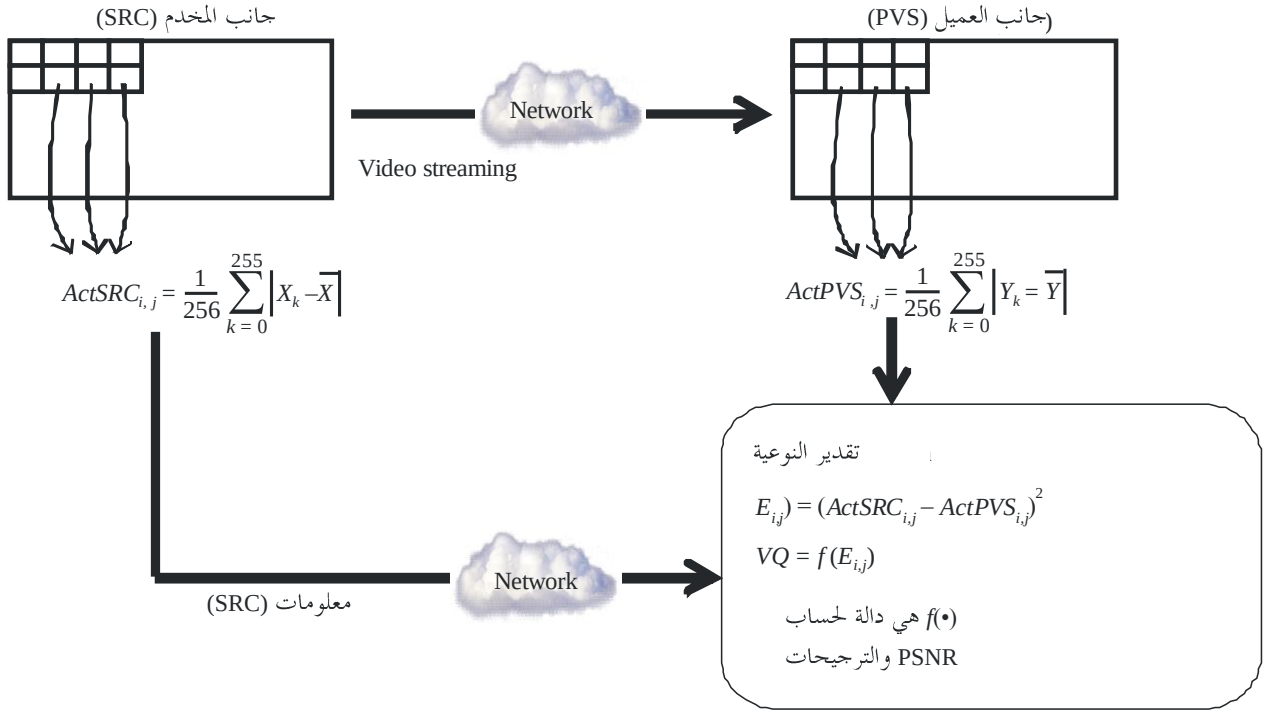
4) تُطبق الترجيحات البصرية الفيزيولوجية على الأخطاء المربعة في الفدر ذات المكونات المكانية عالية التردد واللون المحدد والفرق الواسع بين الأطر وتغيير المشهد.

5) تُشتق علامة مرحلية للنوعية الفيديوية من مجموع الأخطاء المربعة على غرار حساب نسبة ذروة الإشارة إلى الضوضاء (PSNR).

6) تُعدل العلامة لتعكس الترديات المحسوسة المدمِرة جراء الانسداد والتدهور المحلي. وأخيراً، تمثل العلامة المعدلة النوعية الفيديوية المقيسة للتتابع الفيديوي المعالَج (PVS) في نموذج المرجع المخفض.

الشـــكل 24

تقدير النوعية الفيديوية على أساس فرق النشاط



BT.1885-24

4 الخوارزمية التفصيلية

1.4 جانب المخدم

1) تقسم بيكسلات النصوع للقناة المرجعية للمصدر (SRC) إلى 16 × 16 من فدر البيكسلات في كل أطار بعد ثانية واحدة من رأس تتابع فيديوي. وفي الثانية الأولى، لا ترسَل معلومات القناة المرجعية للمصدر (SRC) إذ يصعب على نظام الرؤية البشري كشف تردي النوعية الفيديوية في المشاهد التي تلي الإطار الأول مباشرةً.

2) تُحسب قيم النشاط (نشاط SRC: $ActSRC_{i,j}$) في كل فدرة باستثناء تلك الموجودة في حافة الإطار. ويصف الشكل 25 الفدر التي تُحسب قيم نشاطها وتُرسل. ويُحسب نشاط القناة المرجعية للمصدر (SRC) على النحو التالي:

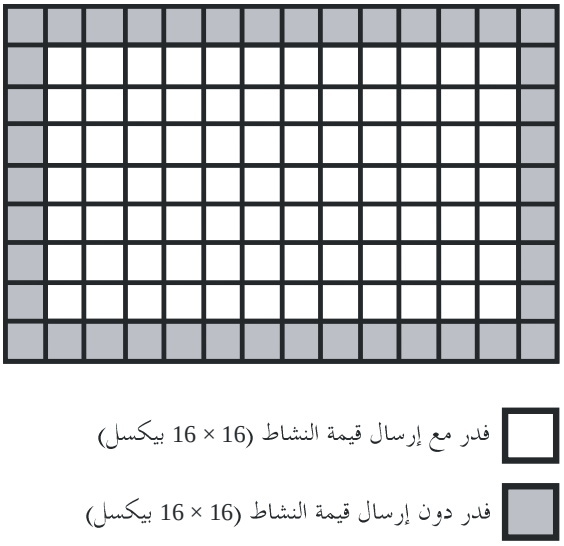
$$ActSRC_{i,j} = \frac{1}{256}\sum_{k=0}^{255}\left|X_k - \overline{X}\right|$$

حيث X_k هي قيمة النصوع في فدرة قناة مرجعية للمصدر ذات مقاس معين، و $\overline{X}$ هي قيمتها المتوسطة وi هو رقم الإطار وj هو رقم الفدرة في الإطار.

3) ترسّل قيم النشاط المعبَّر عنها بثماني بتات في كل الفدرة إلى جانب العميل في ترتيب المسح النقطي بعد ثانية واحدة من من رأس تتابع فيديوي. ولإرسال معلومات القناة المرجعية للمصدر (SRC) بمعدل 256 kbit/s، ترسَل قيم النشاط في جميع الأطر. وعند خفض المعدل إلى 80 kbit/s، ترسَل قيم النشاط في كل أربعة أطر.

الشـــكل 25

فدر مع وبدون إرسال قيمة النشاط



BT.1885-25

2.4 جانب العميل

1.2.4 حساب الخطأ المربع لقيم النشاط

1) تقسم بيكسلات النصوع للتتابع الفيديوي المعالَج (PVS) إلى 16 × 16 من فدر البيكسلات في كل أطار بعد ثانية واحدة من رأس تتابع فيديوي.

2) تُحسب قيم النشاط (نشاط PVS: $ActPVS_{i,j}$) في كل فدرة باستثناء تلك الموجودة في حافة الإطار. وفيما يخص معلومات القناة المرجعية للمصدر (SRC) بمعدل 256 kbit/s، تُحسب قيم النشاط في جميع الأطر. وعند خفض المعدل إلى 80 kbit/s، تُحسب قيم النشاط في كل أربعة أطر.

$$ActPVS_{i,j} = \frac{1}{256}\sum_{k=0}^{255}\left|Y_k - \overline{Y}\right|$$

حيث:

Y_k: قيمة النصوع لفدرة تتابع فيديوي معالَج (PVS) ذات مقاس معين

$\overline{Y}$: متوسطها

i: رقم الإطار

j: رقم الفدرة في الإطار.

3) وتحسب الأخطاء المربعة بين نشاط القناة المرجعية للمصدر (SRC) ونشاط التتابع الفيديوي المعالَج (PVS)

$$E_{i,j} = \left(ActSRC_{i,j} - ActPVS_{i,j}\right)^2$$

2.2.4 الترجيح البصري الفيزيولوجي فيما يتعلق بالخطأ المربع

تُطبق على $E_{i,j}$ ثلاثة أنواع من الترجيحات، وهي على وجه التحديد: ترجيح للفرق في التردد المكاني وترجيح للفرق في منطقة اللون المحددة وترجيح للفرق في النصوع، وذلك لمراعاة خصائص نظام الرؤية البشري.

1) ترجيح للفرق في التردد المكاني

يُستخدم في هذا الترجيح عامل الترجيح W_{SF} والعتبة Th_{SF}. (انظر الجدول 8 للاطلاع على قيم W_{SF} وTh_{SF}.)

$$E_{i,j} \Leftarrow \begin{cases} E_{i,j} \times W_{SF}, & ActPVS_{i,j} > Th_{SF} \\ E_{i,j}, & otherwise \end{cases}$$

2) ترجيح للفرق في منطقة اللون المحددة

في فدرة معينة وما يحيط بها من فدر ثماني، إذا كان عدد البيكسلات (*NumROIPixels*) ضمن المتراجحات $48 \le Y \le 224$ و$104 \le Cb \le 125$ و$135 \le Cr \le 171$ أكبر من عتبة ما، يجرى الترجيح التالي باستخدام عامل الترجيح W_{CR} والعتبة Th_{CR}.

$$E_{i,j} \Leftarrow \begin{cases} E_{i,j} \times W_{CR}, & NumROIPixels > Th_{CR} \\ E_{i,j}, & otherwise \end{cases}$$

انظر الجدول 8 للاطلاع على قيم W_{CR} وTh_{CR}.

3) ترجيح للفرق في النصوع بين الأطر

يُحسب متوسط الفرق المطلق ($MAD_{i,j}$) للنصوع بين فدرة معينة وفدرة في الإطار السابق. ويعرَّف متوسط الفرق المطلق ($MAD_{i,j}$) على أنه:

$$MAD_{i,j} = \frac{1}{256} \sum_{k=0}^{255} |Y_k - Y'_k|$$

حيث Y_k هي قيمة النصوع في فدرة بيكسلات التتابع الفيديوي المعالَج (PVS) 16×16 و Y'_k هي قيمة النصوع في الموضع نفسه في الإطار السابق.

ويجرى الترجيح التالي باستخدام عاملي الترجيح W_{MAD1} وW_{MAD2} والعتبتين Th_{MAD1} وTh_{MAD2}.

$$E_{i,j} \Leftarrow \begin{cases} E_{i,j} \times W_{MAD1}, & MAD_{i,j} > Th_{MAD1} \\ E_{i,j} \times W_{MAD2}, & MAD_{i,j} \le Th_{MAD2} \\ E_{i,j}, & otherwise \end{cases}$$

انظر الجدول 8 للاطلاع على قيم W_{MAD1} وW_{MAD} وTh_{MAD1} وTh_{MAD2}.

3.2.4 الترجيح في حال كشف تغير المشهد

يُحسب متوسط MAD_i ($MADAve_i$) لكل إطار على النحو التالي:

$$MADAve_i = \frac{1}{M} \sum_{j=0}^{M-1} MAD_{i,j}$$

حيث M هو عدد الفدر في إطار.

وإذا كان $MADAve$ أكبر من العتبة Th_{SC}، يُعتبر ذلك تغيراً في المشهد. وعند كشف تغير المشهد، يُضبط الترجيح $E_{i,j}$ بقيمة 0 على مدى 15 إطاراً بعد تغير المشهد.

$$SceneChange = \begin{cases} TRUE, & MADAve_i > Th_{SC} \\ FALSE & otherwise \end{cases}$$

$$E_{i,j} \Leftarrow \begin{cases} E_{i,j} \times W_{SC} & 15\ frames\ after\ SceneChange = TRUE \\ E_{i,j}, & otherwise \end{cases}$$

انظر الجدول 8 للاطلاع على قيم W_{SC} وTh_{SC}.

4.2.4 نسبة ذروة الإشارة إلى الضوضاء (PSNR) على أساس مربع خطأ النشاط

تُحسب نسبة ذروة الإشارة إلى الضوضاء (PSNR) على أساس فرق النشاط على النحو التالي:

$$VQ = 10 \times \log_{10} \frac{255 \times 255}{E_{Ave}}$$

$$E_{Ave} = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} E_{i,j}$$

حيث N وM هما عدد الأطر والفدر المستخدمة لحساب نسبة ذروة الإشارة إلى الضوضاء.

5.2.4 الترجيح لمفرزات الانسداد

يُستخدم عامل الترجيح W_{BL} والعتبة Th_{BL} معلومات مستوى الانسداد BL_{Ave} في هذا الترجيح. (انظر الجدول 8 للاطلاع على قيم W_{BL} وTh_{BL}.)

$$VQ \Leftarrow \begin{cases} VQ \times W_{BL}, & BL_{Ave} > Th_{BL} \\ VQ, & otherwise \end{cases}$$

وتُحسب BL_{Ave} باتباع الخطوات التالية:

الخطوة 1: تُحسب قيم النشاط لفدر بيكسلات التتابع الفيديوي المعالَج (PVS) 8 × 8. وكما هو مبين في الشكل 26، تُحسب القيمة المتوسطة (Act_{Ave}) لقيمتي النشاط في الفدرتين الأفقيتين المجاورتين ($ActBlock_1$, $ActBlock_2$) كما يلي:

$$Act_{Ave} = \frac{1}{2}\left(ActBlock_1 + ActBlock_2\right)$$

الخطوة 2: يُحسب الفرق المطلق في قيم النصوع على امتداد الحد بين فدرتين. وكما هو موضح في الشكل 26، تمثل $Y_{1,0}$ و$Y_{2,0}$ قيمتا النصوع في الفدرتين اليسارية واليمينية على امتداد الحد. ويعبَر عن متوسط قيمة فرق النصوع المطلق، $DiffBound$، على النحو التالي:

$$DiffBound = \frac{1}{8} \sum_{i=0}^{7} \left| Y_{1,i} - Y_{2,i} \right|$$

الخطوة 3: يُعرَّف مستوى الانسداد ($BL_{i,j}$) بالنسبة بين $DiffBound$ وAct_{Ave}، أي:

$$BL_{i,j} = \frac{DiffBound}{Act_{Ave} + 1}$$

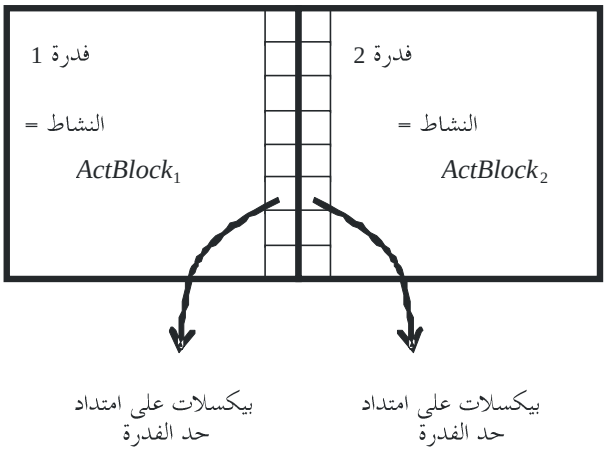
الخطوة 4: تُحسب قيمة BL المتوسطة على النحو التالي:

$$BL_{Ave} = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} BL_{i,j}$$

وتصفَّر قيمة $BL_{i,j}$ للفدر في أقصى يمين الأطر. وإذا كان BL_{Ave} أكبر من عتبة محددة مسبقاً، يُعتبر أن تتابع الفيديو يتخلله مستوىً عال من الانسداد ويطبَّق الترجيح على القيمة المحسوبة للنوعية الفيديوية.

الشـــكل 26

قيم البيكسلات والنشاط المستخدمة لحساب مستوى الانسداد



BT.1885-26

6.2.4 الترجيح لمفرزات التردي المحلي

يُستخدم عامل الترجيح W_{LI} والعتبة Th_{LI} والتردي المحلي LI في هذا الترجيح. (انظر الجدول 8 للاطلاع على قيم W_{LI} وTh_{LI}).

$$VQ \Leftarrow \begin{cases} VQ \times W_{LI}, & LI < Th_{LI} \\ VQ, & otherwise \end{cases}$$

ويُحسب التردي المحلي LI باتباع الخطوات التالية. ويُستخدم الفرق في تغاير النشاط لكشف التردي المحلي جراء أخطاء الإرسال.

1) في فدرة معينة والفدر الثماني المحيطة بها، يُحسب تغاير النشاط لكل من SRC ($ActVar_{SRC}$) و PVS ($ActVar_{PVS}$)، ويُحسب الفرق المطلق بين هاتين القيمتين للتغاير على النحو التالي:

$$\Delta ActVar = \left| ActVar_{SRC} - ActVar_{PVS} \right|$$

2) وتُحسب القيمة المتوسطة لقيم الفرق المطلق هذه في كل إطار.

3) ويُحسب LI كنسبة التغاير ($\Delta ActVar_{Max}$) الأقصى إلى التغاير ($\Delta ActVar_{Min}$) الأدنى للقيمة المتوسطة.

$$LI = \begin{cases} \Delta ActVar_{Min} / \Delta ActVar_{Max} & \Delta ActVar_{Max} \neq 0 \\ 1 & \Delta ActVar_{Max} = 0 \end{cases}$$

ويمثل الرمز VQ علامة النوعية الفيديوية.

7.2.4 معلمات الترجيحات

يبين الجدول 8 قيم معلمات الترجيحات. وتحدد هذه القيم عن طريق تجربة أولية بمجموعة بيانات تدريبية.

الجـــدول 8

معلمات الترجيحات

نمط عملية الترجيح	قيم المعلمة	
ترجيح التردد المكاني	W_{SF}	0,36
	Th_{SF}	25
ترجيح اللون المحدد	W_{CR}	4,0
	Th_{CR}	175
ترجيح الفرق بين الإطارات	W_{MAD1}	0,06
	Th_{MAD1}	17
	W_{MAD2}	25
	Th_{MAD2}	13
كشف تغير المشهد	W_{SC}	0,0
	Th_{SC}	35
ترجيح الانسداد	W_{BL}	0,870
	Th_{BL}	1,0
ترجيح التردي المحلي	W_{LI}	0,870
	Th_{LI}	1,67

8.2.4 التسجيل

1) التسجيل المكاني

لا يحتاج نموذج المرجع المخفض لأي تسجيل مكاني لأن الخطأ المربع يُحسب على أساس قيم النشاط الأكثر حصانةً ضد الانزياحات المكانية مما يقوم على أساس قيم البيكسلات.

2) تسجيل الكسب والتخالف

لا يحتاج نموذج المرجع المخفض لتسجيل الكسب والتخالف. وتخلو قيم النشاط بطبيعتها من التخالف (أي مكونات التيار المستمر) ولا تتأثر بالكسب.

3) التسجيل الزمني

ينقسم التتابع الفيديوي المعالَج (PVS).إلى تتابعات فرعية مدة كل منها ثانية واحدة. وفي كل تتابع فرعي، يُحسب متوسط الأخطاء المربعة للنشاط بخمسة تغيرات تأخر تصل حتى ±2 إطار قناة مرجعية للمصدر (SRC). وأخيراً تُستخدم القيمة الدنيا لمتوسط الأخطاء المربعة كمتوسط الخطأ المربع في هذا التتابع الفرعي. ويعدَّل التأخر الناتج في هذا الحد الدنى للخطأ المربع كتسجيل زمني.

5 عينات شفرات

ترد فيما يلي عينات شفرات بلغة–C في نموذج المرجع المخفض.

1.5 شفرة مشتركة على جانبي المخدم والعميل معاً

```
// Calcule the activity value
unsigned int CalcActivitybyRect(unsigned char * lpBuffer, int nWidth, int iRectWidth, int iRectHeight)
{
        // lpBuffer: Luminance Frame Buffer
    // nWidth: Frame Buffer Width
    // iRectWidth: Width of the rectangle to calculate an activity value.
    // iHeightWidth: Height of the rectangle to calculate an activity value.
        unsigned int i, j, nTmp, nSum;
        unsigned char *pSrc;

        pSrc = lpBuffer;   nSum = 0;
        for (j = 0; j < iRectHeight; j++){
            for (i = 0; i <iRectWidth; i++){
                nSum += pSrc[i];
            }
            pSrc += nWidth;
        }
        nSum /= (iRectWidth*iRectHeight);

        pSrc = lpBuffer; nTmp = 0;
        for (j = 0; j < iRectHeight; j++){
            for (i = 0; i <iRectWidth; i++){
                nTmp += abs(pSrc[i] - nSum);
            }
            pSrc += nWidth;
        }
        return nTmp/iRectWidth/iRectHeight;
}
```

2.5 جانب المخدم

```
// Server side
int nStart = 30; // the frame number to start transmission (30 or 25)
int nMaxFrame = 240; // the number of total video frames (240 or 200)
int nFrameIncrement = 1;  // 1 for 256kbps,  4 for 80kbps
void ReadOneFrame(unsigned char, int, unsigned char *, int, int); // function to read one frame data
int nRim = 16 // 16 or 32 (use 32 to avoid the problem in HRC9)

// nWidth: Frame Buffer Width
// nHeight: Frame Buffer Height
// lpSrc: Frame Buffer
for(int nFrame = nStart;  nFrame < nMaxFrame;  nFrame+=nFrameIncrement){
        ReadOneFrame(SRC_file_name, nFrame, lpSrc, nWidth, nHeight);
        for (j= 16; j<nHeight-32; j+=16) {
            for (i= nRim; i<nWidth- nRim; i+=16) {
                lpOrg = lpSrc + i + j * nWidth;
                nActSrc = CalcActivitybyRect(lpOrg, nWidth, 16, 16);
          // OutputSRCInfo(nActSrc);  // Output or transmission the SRC information
            }
        }
}
```

3.5 جانب العميل

```
// Client Side
int nStart = 30; // the frame number to start transmission (30 or 25)
int nMaxFrame = 240; // the number of total video frames  (240 or 200)
int nFrameIncrement = 1;  // 1 for 256kbps,  4 for 80kbps
int nFrameRate = 30;  //30 or 25
void ReadOneFrame(unsigned char, int, unsigned char **, int, int); // function to read one frame data
void ReadRRData(unsigned char, int, unsigned char *); // function to read RR-data

// nWidth: Frame Buffer Width
// nHeight: Frame Buffer Height
// lpPvsByte[3]: Frame Buffer  (0:Y, 1:Cb, 2:Cr)
```

```
// lpRRData: RR-data Buffer
// double ddActivityDifference[][]:  Store the activity-difference
// double ddActivityVariance[][]:   Store the activity-variance
// double ddBlock[][]:              Store the blockiness level
// int nSceneChange:                Scene change detection

for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){  // Changing temporal alignment
        for(int nFrame = 0; nFrame < nMaxFrames; nFrame++){
            if(nFrame+nTemporalAlign >= nMaxFrames || nFrame+nTemporalAlign < 0){
                continue;
            }
            ReadOneFrame(PVS_file_name, nFrame+nTemporalAlign, lpPvsByte, nWidth, nHeight);
            if(((nFrame-(nFrameRate+nStart)) % nFrameIncrement) == 0
                    && nFrame >= nStart ){
                ReadRRData(RR_file_name, nFrame, lpRRData);
                ddActivityDifference[nTemporalAlign+2][nFrame]
                                    = RRCalcObjectiveScore(lpPvsByte, lpRRData, nWidth, nHeight);
                ddActivityVariance[nTemporalAlign+2][nFrame] = gnActVar;
            }else{
                ddActivityDifference[nTemporalAlign+2][nFrame] = 0.0;
                ddActivityVariance[nTemporalAlign+2][nFrame] = 0.0;
            }
            // Blockiness Level
            if(nTemporalAlign ==0){
                ddBlock[nFrame] = BlockinessLevelEstimation(lpPvsByte[0], nWidth, nHeight);
            }
            // Pixel copy for inter-frame difference calculation
            memcpy(lpPrev, lpPvsByte[0], sizeof(char)*nWidth*nHeight);
            if(nSceneChange){
                nSceneChange--;
            }
        }
}

double ddSum[8][5]; // Sum of the Activity-difference for each second
double ddActVarSum[8][5]; // Sum of the Activity-variance for each second
double ddActVarMax[8][5]; // Maximum of the Sum of the Activity-variance
double ddActVarMin[8][5]; // Minimum of the Sum of the Activity-variance
int nnMin[8];
int nnNumFrames[8][5];
#define LARGENUMBER 100000
for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){
        for(int j=0;j<8;j++){ // for each one second
            nnNumFrames[j][nTemporalAlign+2] = 0;
            ddActVarMax[j][nTemporalAlign+2] = 0.0;
            ddActVarMin[j][nTemporalAlign+2] =  LARGENUMBER;
            ddActVarSum[j][nTemporalAlign+2] = 0.0;
            ddSum[j][nTemporalAlign+2] = 0.0;
            for(int i=nFrameRate*j;i< (j+1)*nFrameRate; i++){
                if(ddActivityDifference[nTemporalAlign+2][i]){
                    ddSum[j][nTemporalAlign+2] += ddActivityDifference[nTemporalAlign+2][i];
                    nnNumFrames[j][nTemporalAlign+2]++;
                }
                ddActVarSum[j][nTemporalAlign+2] += ddActivityVariance[nTemporalAlign+2][i];
                if(ddActivityVariance[nTemporalAlign+2][i]){
                                    if(ddActivityVariance[nTemporalAlign+2][i] >
                                    ddActVarMax[j][nTemporalAlign+2]){
                                            ddActVarMax[j][nTemporalAlign+2] =
                                            ddActivityVariance[nTemporalAlign+2][i];
                    }
                                    if(ddActivityVariance[nTemporalAlign+2][i] <
                                    ddActVarMin[j][nTemporalAlign+2]){
                                            ddActVarMin[j][nTemporalAlign+2] =
                                            ddActivityVariance[nTemporalAlign+2][i];
                    }
                }
            }
        }
}
```

```
// Local Impariment Level Calculation
double dSum = 0.0;
double dActMax = 0.0;
double dActMin = LARGENUMBER;
int nNumFrames = 0;
for(int j=1; j<8; j++){
        double dMin = LARGENUMBER;
        double dMinSum = LARGENUMBER;
        for(int nTemporalAlign = -2; nTemporalAlign <=2; nTemporalAlign++){
            if(ddActVarSum[j][nTemporalAlign+2] < dMin){
                dMin = ddActVarSum[j][nTemporalAlign+2];
                dMinSum = ddSum[j][nTemporalAlign+2];
                nnMin[j] = nTemporalAlign+2;
            }
        }
        dSum += dMinSum;
        nNumFrames += nnNumFrames[j][nnMin[j]];
        if(ddActVarMax[j][nnMin[j]] > dActMax){
            dActMax = ddActVarMax[j][nnMin[j]];
        }
        if(ddActVarMin[j][nnMin[j]] < dActMin){
            dActMin = ddActVarMin[j][nnMin[j]];
        }
}
double dTransError =  dActMax/dActMin;

// Blockiness Level Calculation
double dBlockinessLevel = 0.0;
for(int i=0;i<nMaxFrames; i++){
        dBlockinessLevel += ddBlock[i];
}
dBlockinessLevel = dSumBlock / (double)(nMaxFrames-nFrameRate);

// Calculating the Video Quality Score
if(nNumFrames && nNumberOfBlocks && dSum){
        dSum = dSum / (double)(nNumFrames)/(double)nNumberOfBlocks;
        dSum = 10*log10(255.0*255.0/dSum); //PSNR based on the activity difference
        if(dBlockinessLevel > dBlokinessTh){
            dSum /= nBlockinessWeighting;  // Weighting for blockiness level
        }
        if(dTransError > nErrorTh){
            dSum /=nErrorWeightin;   // Weighting for transmission errors
        }
}
return dSum;
-------------------------------------------------------------------------------------
// Calculating the MAD value
unsigned int CalcMAD(unsigned char *lpSrc, unsigned char *lpSrc2, int nWidth, int nHeight)
{
        // lpSrc:   Frame Buffer of the current frame
         // lpSrc2:  Frame Buffer of the previous frame
         unsigned int nSum = 0;
        for (y = 0; y < nHeight; y++) {
            for (x = 0; x < nWidth; x++) {
                nSrc =  lpSrc[x  + y*nWidth];
                nSrc2 = lpSrc2[x + y*nWidth];
                nSum += abs(nSrc - nSrc2);
            }
        }
        return nSum/nWidth/nHeight;
}

// Calculating a mean squared error with weightings
double RRCalcObjectiveScore(unsigned char *lpBuffer[], unsigned char *lpRRData, int nWidth, int nHeight)
{
        int    i, j, nActSrc, nActSrc2, nY, nCb, nCr, nYMin, nYMax, nCbMin, nCbMax, nCrMin, nCrMax;
        int    nMBX, nMBY, nMB, nStart;
        unsigned int nMAD;
```

```
double     e2, dMADFrame;
unsigned char *lpRec, *lpRecCb, *lpRecCr, *lpPrev1;
int nRim = 16 // 16 or 32 (use 32 to avoid the problem in HRC9)

nYMin = 48; nYMax = 224; nCbMin = 104; nCbMax = 125; nCrMin = 135; nCrMax = 171;
nMB = nMBY = nStart = 0;
e2 = dMADFrame = 0.0;

for (j=16+nStart; j<iImageHeight-32; j+=16) {
      nMBX = 0;
      for (i= nRim; i<nWidth- nRim; i+=16) {
            lpRec = lpBuffer[0] + i + j * nWidth;
            lpRecCb = lpBuffer[1] + i/2 + (j/2) * nWidth/2;
            lpRecCr = lpBuffer[2] + i/2 + (j/2) * nWidth/2;
            lpPrev1 = lpPrev + i + j * nWidth;

            nActSrc = lpRRData[nMB];  // SRC activity
            nActSrc2 = CalcActivitybyRect(lpRec, nWidth, 0, 16, 16); // PVS activity
            nActArray[nMB] = (double)nActSrc;
            nActArray2[nMB] = (double)nActSrc2;
            e2 += (double)(nActSrc - nActSrc2)*(nActSrc - nActSrc2);  // Mean squared error

            nMAD     = CalcMAD(lpRec, lpPrev1, 16, 16, nWidth);  // Inter-frame differnece
            dMADFrame += (double)nMAD;

            int nNumROIPixels=0;
            for(int jj=-16;jj<32; jj++){
                  for(int ii=-16;ii<32; ii++){
                        nY  = lpRec[ii];
                        nCb = lpRecCb[ii/2];
                        nCr = lpRecCr[ii/2];
                        if(nY >= nYMin && nY <= nYMax
                        && nCb >= nCbMin && nCb <= nCbMax
                        && nCr >= nCrMin && nCr <= nCrMax){
                                    nNumROIPixels++;
                        }
                  }
                  lpRec += nWidth;
                  if((jj & 1) == 1){
                        lpRecCb += nWidth/2;
                        lpRecCr += nWidth/2;
                  }
            }

            // Weighting for spatial frequency
            if(nActSrc2 > gdwActThHigh){
                  e2 *= dActWeightingHigh;
            }else if(nActSrc2 > gdwActThLow){
                  e2 *= dActWeightingMiddle;
            }else {
                  e2 *= dActWeightingLow;
            }

            // Weighting for specific color region
            if(nNumROIPixels > dwROITh){
                  e2 *= dROIWeighting;
            }

            // Weighting for inter-frame difference
            if(nMAD > dwMADThHigh){
                  e2 *= dMADWeightingHigh;
            }else if(nMAD > dwMADThLow){
                  e2 *= dMADWeightingMiddle;
            }else {
                  e2 *= dMADWeightingLow;
            }
            nMB++;   nMBX++;
      }
      nMBY++;
}
```

```
        // Calculating Activity-Variance for Surrounding Nine Blocks.
        double nSumActSrc, nSumActPvs, nActVar, nActVar2;
        nSumActSrc = nSumActPvs = nActVar = nActVar2 = 0.0;
        gnActVar = 0.0;
        for (j=1; j<nMBY-1; j++) {
            for (i=1; i<nMBX-1; i++) {
                nSumActSrc = 0.0;
                nSumActPvs = 0.0;
                for(int jj=-1; jj<2; jj++){
                    for(int ii=-1; ii<2; ii++){
                        nSumActSrc += nActArray[i+ii+nMBX*(j+jj)];
                        nSumActPvs += nActArray2[i+ii+nMBX*(j+jj)];
                    }
                }
                nSumActSrc /= 9.0;
                nSumActPvs /= 9.0;

                nActVar = 0.0;
                nActVar2 = 0.0;
                for(int jj=-1; jj<2; jj++){
                    for(int ii=-1; ii<2; ii++){
        nActVar +=  (nActArray[i+ii+nMBX*(j+jj)]-nSumActSrc)*(nActArray[i+ii+nMBX*(j+jj)]-nSumActSrc);
                                        nActVar2 += (nActArray2[i+ii+nMBX*(j+jj)]-
                                        nSumActPvs)*(nActArray2[i+ii+nMBX*
                                        (j+jj)]-nSumActPvs);
                    }
                }
                nActVar /= 9.0;
                nActVar2 /= 9.0;
                gnActVar += abs(nActVar- nActVar2);
            }
        }
        // Average of the Activity-Variance for the Frame
        gnActVar = gnActVar/(double)(nMBY-2)/(double)(nMBY-2);

        // Scene change detection
        if(dMADFrame/ nMB > 35){
            nSceneChange = 15;
        }
        if(nSceneChange){
            e2 = 0.0;
        }
        gnFrame++;

        return e2;
}


// Calculate Blockiness Level
double BlockinessLevelEstimation(unsigned char *lpBuffer, int nWidth, int nHeight)
{
        // lpBuffer: Frame Buffer
        int    i, j, nActSrc, nActSrc2, nDiff, nMB;
        unsigned char *lpRec = lpBuffer;
        double dBlock=0.0;

        nMB = 0;
        for (j=0; j<nHeight-16; j+=8) {
            for (i=0; i<nWidth-16; i+=8) {
                lpRec = lpBuffer + i + j * nWidth;
                nActSrc  = CalcActivitybyRect(lpRec, nWidth, 0, 8, 8); // Activity of the left block
                nActSrc2 = CalcActivitybyRect(lpRec+8, nWidth, 0, 8, 8); // Activity of the right block
                nActSrc = (nActSrc + nActSrc2)/2; // Average of the activity values
                nDiff = 0;
                for(int jj=0;jj<8; jj++){
                    nDiff += abs(lpRec[7+jj*nWidth] - lpRec[8+jj*nWidth]);
                                    // Difference of the luminance values at the boundary
                }
                nDiff/= 8;
```

```
                dBlock += (double)nDiff/(double)(nActSrc+1);
                nMB++;
            }
        }
        return (double)dBlock/(double)nMB;
}
```

6 المراجع الإعلامية

[1] SMPTE 170M, "SMPTE Standard for Television – Composite Analog Video Signal – NTSC for Studio Applications", Society of Motion Picture and Television Engineers.

[2] التوصية ITU-T J.143 – متطلبات المستخدم لقياسات النوعية الفيديوية الحسية في التلفزيون الكبلي الرقمي.

الملحق C

النموذج C: أسلوب NTIA بالمرجع المخفض

1 معلومات أساسية

في الإطار الزمني 2003-2004، وضعت الإدارة الوطنية للاتصالات والمعلومات (NTIA) الأمريكية نموذجين للنوعية الفيديوية (VQMs) بعرض نطاق مرجعي مخفض يقارب 12 إلى 14 kbits/s للعينات الفيديوية المذكورة في التوصية ITU-R BT.601 [1]. ودعي هذان النموذجان "نموذج النوعية الفيديوية بعرض النطاق الضيق" و"نموذج النوعية الفيديوية بعرض النطاق الضيق السريع". والنموذج السريع أسرع بحوالي أربع مرات لأنه يستخلص أولاً الخصائص المكانية من أطر الفيديو التي تُحسب متوسطاتها مسبقاً بدلاً من استخلاص الخصائص المكانية من أطر فيديو التوصية ITU-R BT.601 مباشرةً. وجاءت وفورات حسابية إضافية للنموذج السريع من حساب خصائص المعلومات الزمنية (أي الحركة) القائمة على أخذ عينات فرعية عشوائياً من بيكسلات قناة النصوع Y بدلاً من استعمال جميع البيكسلات في قنوات الفيديو الثلاث كلها (Y وCb وCr). ويتوفر النموذجان كليهما منذ بضع سنوات ضمن أدوات برمجيات نماذج النوعية الفيديوية (VQM) لدينا،، ويمكن استخدامهما بحرية للتطبيقات التجارية وغير التجارية. وتتوفر للتنزيل من الإنترنت إصدارات اثنينية قابلة للتنفيذ لأدوات نماذج النوعية الفيديوية هذه وما يرتبط بها من شفرة مصدر [2].

وإذ أرادت الإدارة الوطنية للاتصالات والمعلومات (NTIA) أن تُخضع نموذج النوعية الفيديوية بعرض النطاق الضيق ونموذج النوعية الفيديوية بعرض النطاق الضيق السريع كليهما لاختبارات تلفزيون المرجع المخفض (RRTV) التي يجريها فريق الخبراء المعني بالجودة الفيديوية (VQEG)، اختارت الإدارة أن تقدمهما ضمن فئات مختلفة لمعدلات البتات رغم تطابق متطلباتهما من حيث معدل بتات المرجع المخفض. فاختارت الإدارة أن تقدم نموذج النوعية الفيديوية بعرض النطاق الضيق إلى فئة 256 k ونموذج النوعية الفيديوية بعرض النطاق الضيق السريع إلى فئة 80 k، لأن النموذج الأول يُتوقع أن يتفوق على النموذج السريع. ويستعمل كلا النموذجين خوارزمية معايرة المرجع المخفض الخاصة بالإدارة الوطنية للاتصالات والمعلومات والتي ترد في التوصية ITU-T J.244 [3]. وتتطلب خوارزمية المعايرة هذه ما يقرب من 22 إلى 24 kbits/s من عرض النطاق المرجعي المخفض لإنتاج تقديرات عن التأخر الزمني والانزياح المكاني والتوسع القياسي المكاني والكسب وتخالف المستوى.

وكانت النتيجة المثيرة للاهتمام من اختبارات التقييم VQEG RRTV [4] أن نموذج النوعية الفيديوية بعرض النطاق الضيق السريع تتفوق أداءً على نموذج النوعية الفيديوية بعرض النطاق الضيق في اختباري 525 خطاً و625 خطاً على السواء. وتسترعي هذه النتيجة اهتماماً لأنها تعني ضمناً أن استخراج الخصائص المكانية من الأطر التي حُسبت قيمها المتوسطة أفضل من استخراجها من الأطر التي لم تُحسب قيمها المتوسطة. وستظهر البحوث اللاحقة ما إذا كانت هذه النتيجة ستصح لمجموعات البيانات الأخرى. وفي هذه الآونة، لا ترى الإدارة الوطنية للاتصالات والمعلومات (NTIA) سبباً يدعو لتقييس النموذجين كليهما، ومن ثم يكتفي هذا الملحق بوصف نموذج النوعية الفيديوية بعرض النطاق الضيق السريع.

2 مقدمة

يعرض هذا الملحق وصفاً ومدونة مرجعية لنموذج النوعية الفيديوية بعرض النطاق الضيق السريع. ويتشابه هذا النموذج الذي وضعته الإدارة الوطنية للاتصالات والمعلومات (NTIA) في التقنيات التي يستخدمها مع النموذج العام الذي وضعته الإدارة في هذا الصدد والذي يرد وصفه في كلتا التوصيتين ITU-T J.144 [5] وITU-R BT.1683 [6]. ويستخدم نموذج النوعية الفيديوية بعرض النطاق الضيق السريع خصائص المرجع المخفض بعرض نطاق يقل كثيراً عن النموذج العام، وإن تشابهت عملية استخراج الخصائص ومقارنتها في كلا النموذجين. ففي العينات الفيديوية المذكورة في التوصية ITU-R BT.601، يستخدم نموذج النوعية الفيديوية بعرض النطاق الضيق السريع خصائص مرجع مخفض تتطلب ما يقرب من 12 إلى 14 kbits/s من عرض نطاق الإرسال. ويقتصر وصف هذا الملحق على نموذج النوعية الفيديوية بعرض النطاق الضيق السريع لأن ما يتممه من خوارزميات معايرة عرض النطاق الضيق موثقة تماماً في التوصية ITU-T J.244 [3]. ولكن، توخياً لاكتمال العرض، تتضمن الشفرة المرجعية في هذا الملحق خوارزميات المعايرة نموذج النوعية الفيديوية بعرض النطاق الضيق السريع وما يرتبط به من خوارزميات معايرة عرض النطاق الضيق. وتتضمن الشفرة المرجعية أيضاً أمثلة على وظائف التكمية للخصائص التي تستخدمها معايرة عرض النطاق الضيق (وظائف التكمية ليست جزءاً من التوصية ITU-T J.244).

3 وصف نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع

1.3 نظرة عامة على نموذج النوعية الفيديوية (VQM)

يشمل وصف نموذج النوعية الفيديوية (VQM) ثلاثة مجالات رئيسية هي:

1) خصائص عرض النطاق الضيق المستخرجة من تدفقات الفيديو الأصلية والمعالَجة؛

2) المعلمات الناتجة عن مقارنة تدفقات الخصائص المتماثلة الأصلية والمعالَجة؛

3) حساب نموذج النوعية الفيديوية الذي يجمع بين مختلف المعلمات التي يقيس كل منها جانباً مختلفاً من النوعية الفيديوية.

ويعتمد هذا الوصف على استخدام المراجع المتوفرة بسهولة للحصول على التفاصيل التقنية.

2.3 وصف الخصائص

1.2.3 نظرة عامة على الخصائص

يستخدم نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع ثلاثة أنواع من الخصائص: لونية ومكانية وزمنية. ويحدد كل منها في ميدانه كمية التشوهات الحسية. ويوفر البرنامج الفرعي للشفرة المرجعية "model_fastlowbw_features" وصفاً رياضياً كاملاً للخصائص التي يستخدمها نموذج النوعية الفيديوية بعرض النطاق الضيق السريع.

2.2.3 الخصائص اللونية

إن الخصائص اللونية هي نفس خصائص f_{COHER_COLOR} المستخدمة في نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات (NTIA). ويرد وصف مفصّل لهذه الخصائص في الملحق 3.7.D بالتوصية ITU-T J.144. وتوفر خصائص قياس متجه ثنائي الأبعاد لكمية معلومات التلون الأزرق والأحمر (C_B, C_R) في كل منطقة مكانية زمنية (S-T). وهكذا، تتحسس خصائص f_{COHER_COLOR} لتشوهات الألوان، وهي، في نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع الذي وضعته الإدارة الوطنية للاتصالات والمعلومات (NTIA)، تُستخرج من منطقة مكانية زمنية مقاسها 30 خطاً رأسياً ضرب 30 بيكسلاً أفقياً ضرب ثانية واحدة من الزمن (أي 1 s × 30 × 30). في حين أن نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات استخدم إطار مقاسات 1 × 8 × 8 للمنطقة المكانية الزمنية.

3.2.3 الخصائص المكانية

إن الخصائص المكانية هي نفس خصائص f_{SI13} وf_{HV13} المستخدمة في نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات (NTIA). ويرد وصف مفصّل لهذه الخصائص في الملحق 2.2.7.D بالتوصية ITU-T J.144. وتقيس خصائص f_{SI13} وf_{HV13} كمية التدرجات المكانية وتوزيعها الزاوي في المناطق الفرعية المكانية الزمنية (S-T) لصور النصوع (Y). وهكذا، تتحسس خصائص f_{SI13} وf_{HV13} للتشوهات المكانية كالغشاوة والسد، وهي، في نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع الذي وضعته الإدارة الوطنية للاتصالات والمعلومات (NTIA)، تُستخرج من منطقة مكانية زمنية مقاسها 30 خطاً رأسياً ضرب 30 بيكسلاً أفقياً ضرب ثانية واحدة من الزمن (أي $30 \times 30 \times 1$ s). في حين أن نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات استخدم مقاسات $8 \times 8 \times 0{,}2$ s للمنطقة المكانية الزمنية. وبالإضافة إلى ذلك، ولاختصار الحسابات، يتم أولاً الحساب المسبق لمتوسط صور للنصوع Y عبر الزمن قبل تطبيق مراشيح تعزيز الحواف ثنائية الأبعاد 13×13 التي ترد في الملحق 1.2.7.D بالتوصية ITU-T J.144.

وتُستخرج خاصية مكانية إضافية من الحساب المسبق لمتوسط صور للنصوع Y الممتدة لثانية واحدة. وهذه الخاصية هي متوسط مستوى النصوع (Y) (في كل منطقة مكانية زمنية (S-T) $30 \times 30 \times 1$ s (يشار إليها هنا بالرمز f_{MEAN}). والغرض من هذه الخاصية هو توفير دالة ترجيح مستوى النصوع الحسي لفاقد المعلومات المكانية (SI) كما تقيسه خصائص f_{SI13} وf_{HV13}. وسيأتي وصف ذلك في فقرة وصف المعلمات.

4.2.3 الخصائص الزمنية

يمكن الحصول على تقديرات قوية للنوعية الفيديوية من مجموعة الخصائص اللونية والمكانية الموصوفة أعلاه. بيد أن المناطق المكانية الزمنية (S-T) التي تُستخرج منها هذه الخصائص تمتد عبر العديد من أطر الفيديو (أي ثانية واحدة من أطر الفيديو) فهي تميل لتجاهل الاضطرابات الزمنية القصيرة في الصورة. وقد تنتج هذه الاضطرابات عن الضوضاء أو أخطاء الإرسال الرقمي، وهي وإن كانت قصيرة بطبيعتها يمكن أن يكون لها تأثير كبير على النوعية الملموسة للصورة. إذن، تُستخدم خاصية المرجع المخفض القائم على الزمن للتحديد الكمي لآثار الاضطرابات الزمنية الملموسة. وتقيس هذه الخاصية المعلومات الزمنية المطلقة (ATI) أو الحركة في مستوي صورة النصوع Y، وهي تُحسب على النحو التالي:

$$f_{ATI} = rms\{rand5\%|Y(t) - Y(t-0.2s)|\}$$

وتوخياً للكفاءة الحسابية، تؤخذ عينات فرعية من Y بحيث تحوي 5% فقط من بيكسلات الصورة (تمثَّل هنا بدالة $rand5\%$). وتُطرح صورة Y التي تؤخذ عينات فرعية منها خلال زمن t-0,2s من صورة Y التي تؤخذ عينات فرعية منها على نحو مماثل خلال زمن t، وتُستخدم القيمة الفعالة (rms) للنتيجة كقياس للمعلومات الزمنية المطلقة (ATI). وجرياً على ما اصطُلح عليه في الملحق 8.D بالتوصية ITU-T J.144، سيشار إلى ذلك بالصيغة التالية:

$$f_{ATI} \cong \text{Y_rand5\%_ati0.2s_rms}$$

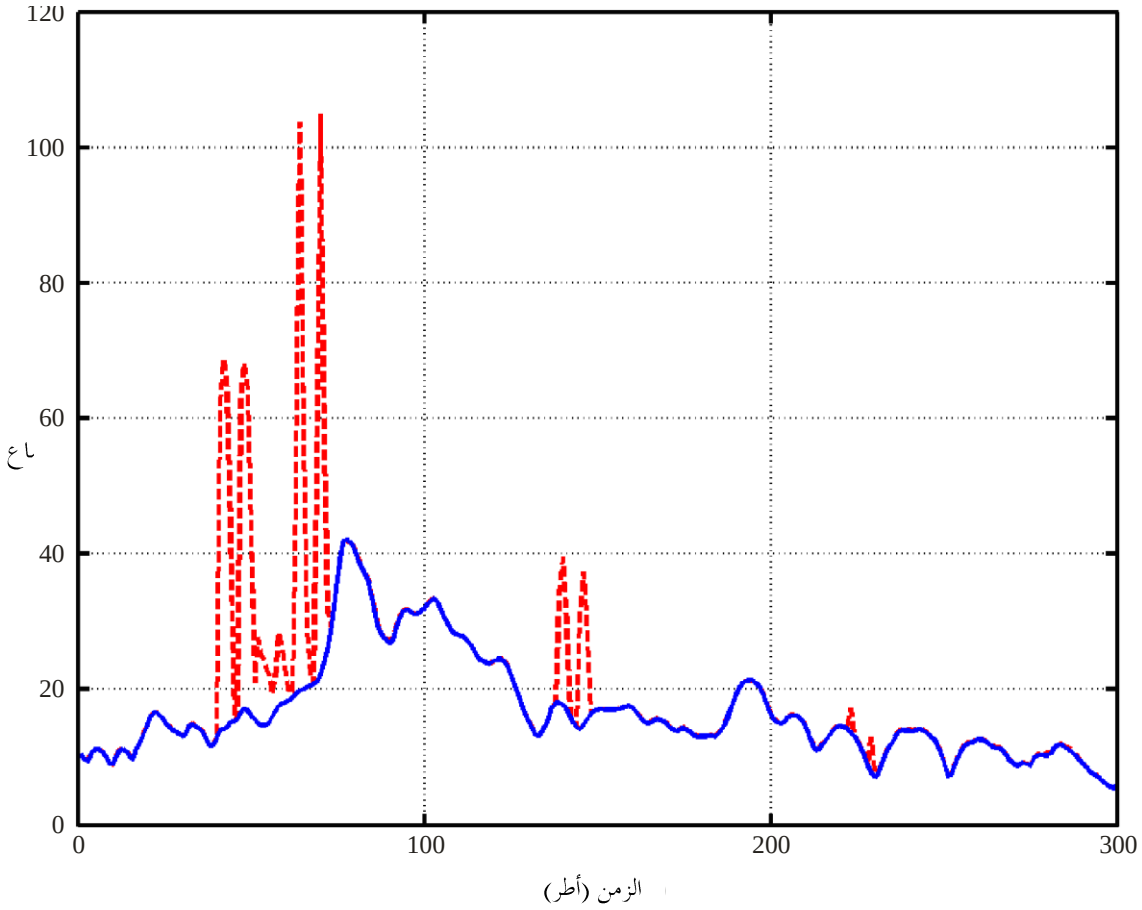
وتتحسس الخاصية f_{ATI} بالاضطرابات الزمنية. ففي المعدل الفيديوي البالغ 30 fps، تغطي فترة 0,2 s ستة أطر فيديو. أما في المعدل الفيديوي البالغ 25 fps، تغطي فترة 0.2 s خمسة أطر فيديوية. وطرح صور متباعدة بفاصل 0,2 s لا يتيح للخاصية أن تتحسس لأنظمة فيديوية بمعدل 30 fps و25 fps في الوقت الفعلي وتحدَّث أطرها بمعدلات لا تقل عن 5 fps.وتكفي الخصائص f_{SI13} وf_{HV13} وf_{COHER_COLOR} لالتقاط الجوانب النوعية لهذه الأنظمة الفيديوية ذات معدل الأطر المنخفض والتي يشيع استخدامها في تطبيقات الوسائط المتعددة. كما أن التباعد بفاصل 0,2 s أقرب مطابقةً لذروة الاستجابة الزمنية في النظام البصري البشري من مفاضلة صورتين متباعدتين بإطار واحد زمنياً.

ويوفر الشكل 27 مثالاً على رسم خاصية f_{ATI} لمشهد فيديوي أصلي (أزرق متواصل) ومعالَج (أحمر متقطع) من نظام فيديوي رقمي ذي أخطاء رشقية عابرة في قناة الإرسال الرقمية. وتُحدث الأخطاء العابرة في الصورة المعالَجة نتوءات نبضية في خاصية f_{ATI}. أما عرض النطاق اللازم لنقل هذه الخاصية فهو ضيق جداً لأنها لا تقتضي إلا 30 عينة في الثانية لفيديو بمعدل 30 fps. وستظهر أنواع أخرى من الضوضاء المضافة، كتلك التي يولدها نظام فيديوي تماثلي، بشكل انزياح موجب لتيار مستمر في السجل الزمني لتدفق الخاصية المعالَج مقارنةً مع تدفق الخاصية الأصلي. فيما تتسبب أنظمة تشفير الفيديو التي تزيل الضوضاء بانزياح سالب لتيار مستمر.

وقبل استخراج معلمة خطأ عابر من تدفقات خاصية f_{ATI} المبينة في الشكل 27، يستفاد من تعريض النتوءات النبضية للحركة (النتوءات النبضية الحمراء في الشكل 27). وذلك لأن النتوءات النبضية الضيقة للحركة الناجمة عن الأخطاء العابرة لا تمثل بشكل كاف التأثير الحسي لهذه الأنواع من الأخطاء. ويتمثل أحد أساليب تعريض النتوءات النبضية للحركة في تطبيق مرشاح الحد الأقصى على تدفقات الخاصية الأصلية والمعالَجة على السواء قبل حساب دالة معلمة الخطأ بين شكلي الموجة. وفي معلمة الخطأ القائمة على f_{ATI}، استُخدم مرشاح حد أقصى بعرض 7 نقاط (سيُرمز إليه بدالة *max7pt*) لإنتاج عينة خرج في كل إطار حد أقصى بذاته وفي الأطر الثلاثة المجاورة الأقرب على الجانبين (أي قبل وبعد عينات الوقت).

الشـــكل 27

مثال على السجل الزمني لخاصية f_{ATI}



BT.1885-27

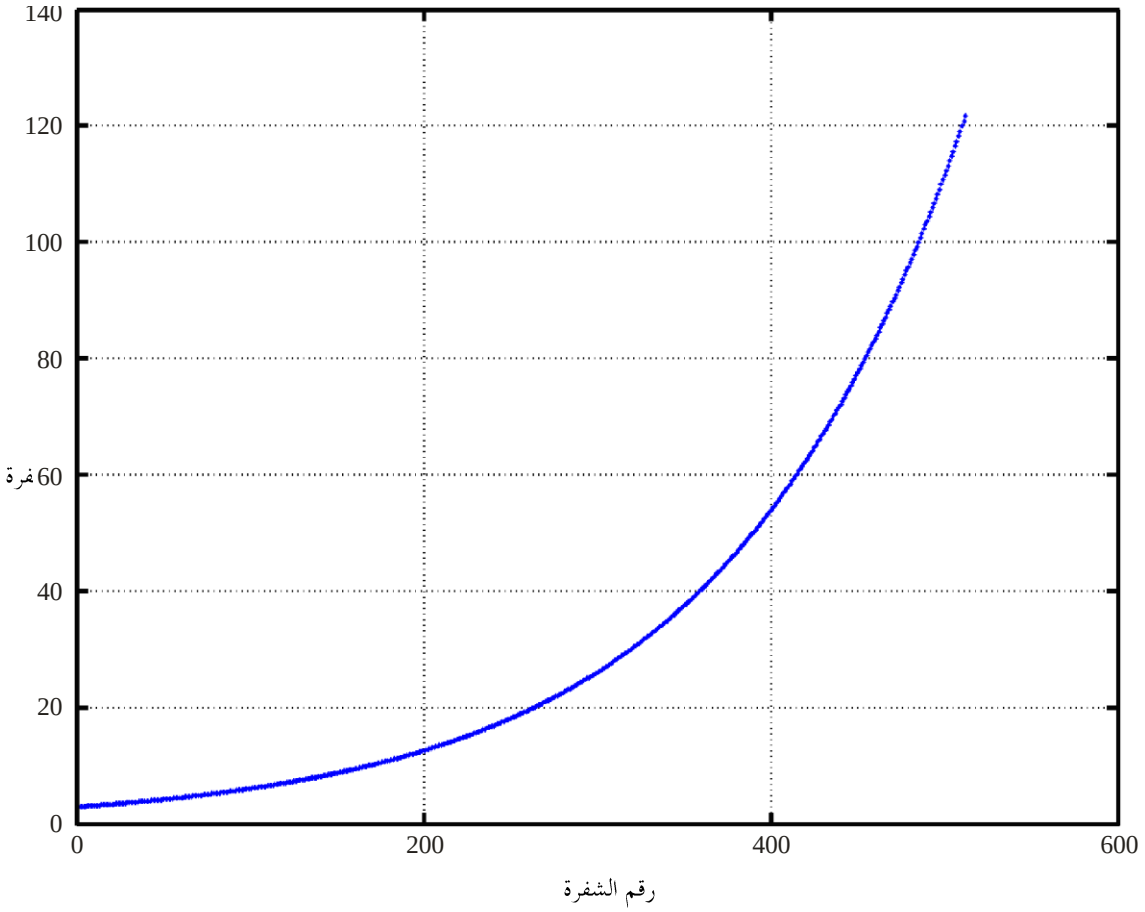
5.2.3 تكمية الخصائص

تكفي التكمية بدقة 9 بتات للخصائص Y_{MEAN} وf_{SI13} وf_{HV13} وf_{COHER_COLOR}؛ فيما ينبغي تكمية الخاصية f_{ATI} بدقة 10 بتات. وللتقليل إلى أدنى حد من التأثير على حسابات معلمة النوعية الفيديوية، ينبغي استخدام تصميم لا خطي للمكمم بحيث يتناسب خطأه مع اتساع الإشارة قيد التكمية. وتجري تكمية القيم المنخفضة جداً على نحو موحد إلى قيمة قطع معينة، تنعدم ما دونها المعلومات المفيدة بشأن تقييم النوعية. ويحد هذا التصميم من الخطأ في الحسابات المقابلة للمعلمات لأن هذه الحسابات تستند عادة إلى نسبة الخطأ أو النسبة اللوغاريتمية لتدفقات الخاصية الأصلية والمعالَجة (انظر فقرة وصف المعلمات أدناه).

ويرد في الشكل 28 منحني المكمم اللاخطي بتسع بتات لخاصية f_{SI13} الأصلية. ويوفر البرنامج الفرعي للشفرة المرجعية "model_lowbw_compression" وصفاً رياضياً كاملاً للمكممات التي يوصى باستخدامها في نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع. وإذا ما وقعت الخصائص خارج مجال المكممات الموصى بها على الطرفين المرتفع والمنخفض (وهو أمر مستبعد جداً)، تصفَّر المعلمات المكانية الزمنية المشتقة من هذه الخواص بحيث لا تؤثر على نموذج النوعية الفيديوية بوجه الإجمال.

الشـــكل 28

المكمم اللاخطي بتسع بتات لخاصية f_{SI13}



BT.1885-28

3.3 وصف المعلمات

1.3.3 نظرة عامة على المعلمات

إن حساب المعلمات التي تقتفي مختلف الجوانب الحسية للنوعية الفيديوية ينطوي على عدة خطوات يمكن أن تشمل ما يلي:

- تطبيق عتبة حسية على الخواص المستخرَجة من كل منطقة مكانية زمنية (S-T) فرعية؛
- حساب دالة الخطأ بين الخصائص المعالَجة وما يقابلها من خصائص أصلية؛
- تجميع الخطأ الناتج عبر المكان والزمان.

انظر الملحق 8.D بالتوصية ITU-T J.144 للاطلاع على وصف مفصل لهذه التقنيات وما يرافقها من ترميز رياضي لأسماء المعلمات سيُستخدم هنا أيضاً. ويوفر البرنامج الفرعي للشفرة المرجعية "model_fastlowbw_features" وصفاً رياضياً كاملاً للخصائص التي يستخدمها نموذج النوعية الفيديوية بعرض النطاق الضيق السريع. وللتبسيط، لا يحتسب وصف المعلمات في هذه الفقرة آثار تكمية الخاصية (كالتعامل مع قيم الخاصية التي يمكن أن تقع خارج مجالات التكمية الموصى بها).

2.3.3 أساليب جديدة

ستلخص هذه الفقرة الأساليب الجديدة التي تم الاهتداء إليها لتحسين الترابط الموضوعي والشخصي للمعلمات القائمة على خصائص ذات مرجع مخفض بعروض نطاق إرسال ضيقة جداً كتلك المستخدمة في في نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات (NTIA) (أي أساليب جديدة لا ترد في التوصية ITU-T J.144). ومن الجدير بالذكر عدم التوصل إلى أي تحسينات للشكل الأساسي لدالات خطأ المعلمة الواردة في الملحق 1.2.8.D بالتوصية ITU-T J.144. فدالتا الخطأ اللتان تنتجان باستمرار نتائج المعلمات الفضلى (للمعلمتين المكانية والزمنية) هما النسبة اللوغاريتمية $\{\log10\ [f_p(s,t)\ /\ \{[(\ f_o(s,t)]\}\}$ ونسبة الخطأ $\{[f_p(s,t)-f_o(s,t)]\ /\ f_o(s,t)\}$، حيث $f_p(s,t)$ و$f_o(s,t)$ هما على التوالي الخاصية المعالَجة والخاصية الأصلية المقابلة المستخرجتان من المنطقة المكانية الزمنية ذات الإحداثيات المكانية s والإحداثيات الزمنية t. ويجب فصل الأخطاء بين المكاسب والخسائر، باعتبار أن البشر يستجيبون استجابة مختلفة لتردي الزيادة (مثل السد) وتردي النقصان (مثل الغشيان). ومن شأن تطبيق عتبة حسية أدنى على الخصائص قبل تطبيق دالتي الخطأ أن يمنع القسمة على صفر.

وبعد حساب المعلمات المكانية الزمنية باستخدام إحدى دالتي الخطأ، يجب تجميع هذه المعلمات على امتداد المكان والزمان لإنتاج قيمة معلمية لمقطع الفيديو. ويمكن لتجميع الخطأ هذا أن يجري على مراحل متعددة (على امتداد المكان ثم على مر الوقت مثلاً). ويدعى أحد الأساليب الجديدة المستخدمة في نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع لتجميع الخطأ، تجميع خطأ السد الموسع (MB). إذ يجمّع هذا الأسلوب عدداً من المناطق المكانية الزمنية الفرعية المتجاورة ويطبق دالة تجميع الخطأ على هذه المجموعة. فعلى سبيل المثال، فإن الدالة التي يُرمز إليها بصيغة "MB(3,3,2)max" ستؤدي وظيفة الحد الأقصى عبر القيم المعلمية من كل مجموعة من المناطق المكانية الزمنية الفرعية الثماني عشرة المكدسة ثُلاث رأسياً بثلاثَ أفقياً بمثنى زمنياً. وفي المناطق المكانية الزمنية الفرعية 32 x 32 x 1 s للخصائص f_{SI13} وf_{HV13} وf_{COHER_COLOR} المذكورة أعلاه، تشمل كل منطقة MB(3,3,2) شطراً من تدفق الفيديو الممتد عبر 96 خطاً رأسياً ضرب 96 بيكسلاً أفقياً ضرب ثانيتين من الزمن. وقد تبينت فائدة تجميع خطأ السد الموسع (MB) في اقتفاء التأثير الحسي لحالات التردي المحصورة في المكان والزمان التي كثيراً ما تهيمن على عملية قرار النوعية. ويمكن أيضاً تنفيذ تجميع خطأ السد الموسع (MB) باعتبارها عملية اصطفاء. فبدلاً من إنتاج قيمة خرج واحد لكل تجميع MB، يستعاض عن كل عينة مكانية زمنية بقيمتها في اصطفاء تجميع MB، حيث تتمركز هذه العينة حول تجميع MB. ويدعى هذا التجميع تجميع خطأ السد الموسع المتراكب (OMB).

ويتمثل الأسلوب الثاني لتجميع الخطأ في جمع مينكوفسكي المعمم ((Minkowski(P,R)) المعرَّف كما يلي:

$$Minkowski(P,R) = \sqrt[R]{\frac{1}{N}\sum_{i=1}^{N}|v_i|^P}$$

وهنا يمثل الرمز v_i قيم المعلمات التي تم تضمينها في الجمع. وقد يشمل هذا الجمع، على سبيل المثال، جميع قيم المعلمات في لحظة معينة من الوقت(التجميع المكاني)، أو يمكن تطبيقه على الفدر الموسعة المذكورة أعلاه. وقد استخدم العديد من المطورين لمقاييس النوعية الفيديوية جمع مينكوفسكي لتجميع الخطأ حيث الأس P يساوي الجذر R. أما حيث $P \neq R$، فإن جمع مينكوفسكي المعمم يوفر المزيد من المرونة لتحقيق خطية استجابة المعلمات الفردية للتغيرات في النوعية الحسية. وهذه خطوة ضرورية قبل الجمع بين معلمات متعددة في تقدير واحد للنوعية الفيديوية الحسية، ويتم ذلك بملاءمة خطية.

3.3.3 معلمات اللون

تُستخرج معلمتان من خصائص f_{COHER_COLOR}. وتدعى إحداهما تطرف اللون وهي تقيس تشوهات اللون المتطرفة التي قد تكون ناجمة عن الفدر الملونة من أخطاء الإرسال. وتدعى المعلمة الأخرى انتشار اللون وتوفر على التغاير أو الانتشار في أخطاء اللون. فبدلاً من استخدام قياس المسافة الإقليدية لتحديد كم التشوهات (كما في الملحق 2.2.8.D بالتوصية ITU-T J.144). وتستخدم هاتان المعلمتان كلتاهما الجذر التربيعي لمسافة مانهاتن. وباتباع الترميز الرياضي في الملحق 2.2.8.D بالتوصية ITU-T J.144، حيث تمثل الخاصيتان $f_p(s,t)$ و$f_o(s,t)$ الخاصية f_{COHER_COLOR} ثنائية الأبعاد المستخرجة من المنطقة المكانية الزمنية (S-T) لتدفقات الفيديو الأصلية، تعطى دالة مقارنة الخصائص بالصيغة التالية:

$$sqrtmanhat(s,t) = \sqrt{\sum_{C_B,C_R} \left| \underline{f}_p(s,t) - \underline{f}_o(s,t) \right|}$$

يبدو قياس المسافة مانهاتن أفضل من قياس المسافة الإقليدية ويلزم أن تحقق دالة الجذر التربيعي خطية استجابة المعلمة لتغيرات النوعية. وباتباع الترميز الرياضي في الملحق 8.D بالتوصية ITU-T J.144، تعطى معلمات اللون بالصيغ التالية:

$$color_extreme = \text{color_coher_color_30x30_1s_mean_sqrtmanhat_OMB(3,3,2)above99\%_Minkosski(0.5,1)}$$

$$color_spread = \text{color_coher_color_30x30_1s_mean_sqrtmanhat_OMB(3,3,2)Minkosski(2,4)_90\%}$$

ثم تُحسب معلمة اللون المختلطة (*color_comb*) التي تحوي التوليفة المثلى لمعلمتي تطرف اللون وانتشار اللون على النحو التالي:

$$color_comb = 0.691686 * color_extreme - 0.617958 * color_spread$$

ثم تُقص معلمة اللون المختلطة (*color_comb*) هذه ذات القيمة الموجبة في الطرف المنخفض، ويُعبر عن ذلك رياضياً (باتباع الترميز في الملحق 5.8.D بالتوصية ITU-T J.144) على النحو التالي:

$$color_comb = color_comb_clip_0.114$$

وترد معلمة اللون المختلطة (*color_comb*) هذه في التوليفة الخطية لحساب نموذج النوعية الفيديوية (VQM).

4.3.3 المعلمات المكانية

تحسب معلمتان مكانيتان من الخاصية f_{SI13}، فتقيس إحداهما خسارة المعلومات المكانية (*si_loss*) وتقيس الأخرى كسب المعلومات المكانية. وباتباع الترميز الرياضي في الملحق 8.D بالتوصية ITU-T J.144، تعطى هاتان المعلمتان بالصيغ التالية:

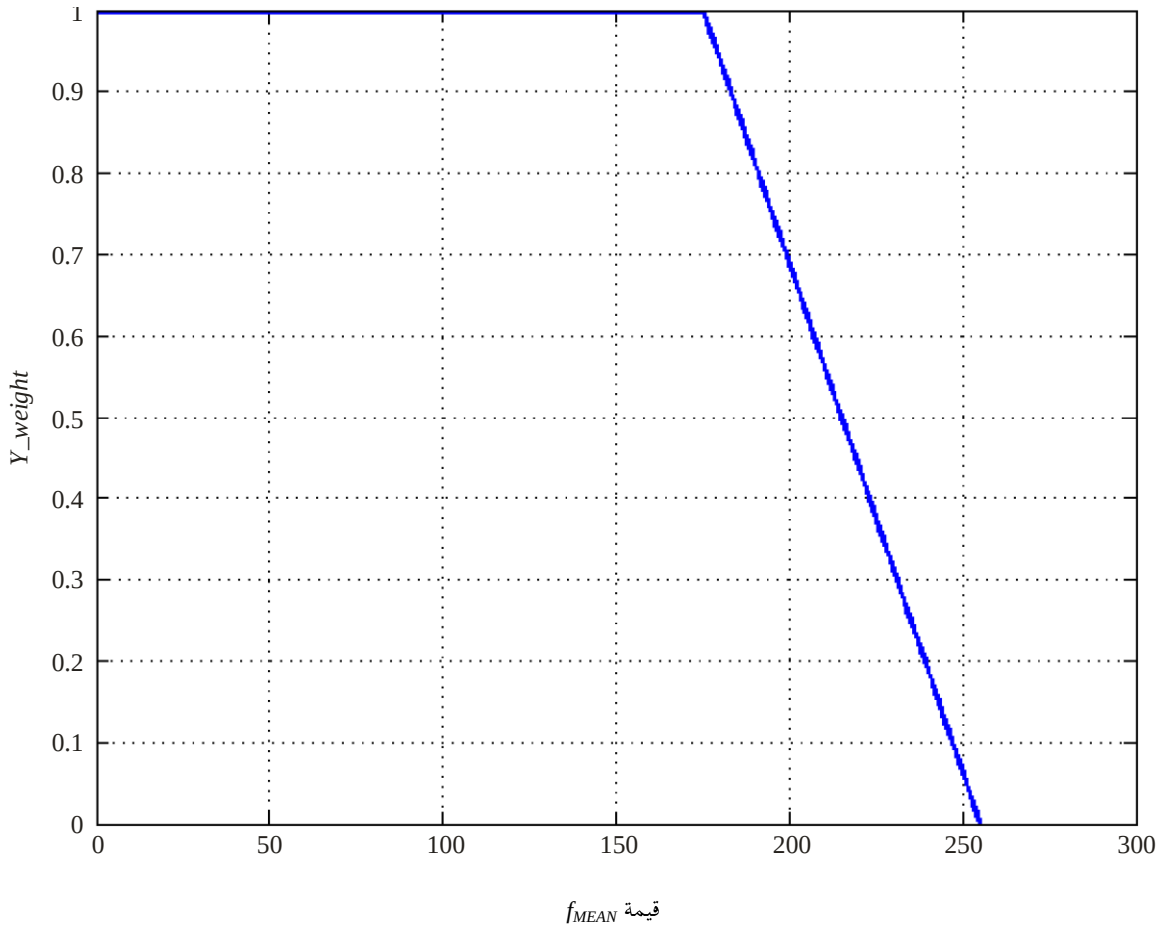
$$si_loss = \text{avg1s_Y_si13_30x30_std_3_ratio_loss_OMB(3,3,2)Minkosski(1,2)_Minkosski(1.5,2.5)_clip_0.12}$$

$$si_gain = \text{avg1s_Y_si13_30x30_std_3_log_gain_clip_0.1_above95\%tail_Minkosski(1.5,2)}$$

وإذ يرتفع متوسط مستوى النصوع (Y) في المنطقة المكانية الزمنية الفرعية (أي كما تقيسه الخاصية f_{MEAN})، تتناقص القدرة على تحسس التغيرات في التفاصيل المكانية (مثل الغشيان المقيس في *si_loss*). ويمكن استيعاب ذلك عن طريق إدخال دالة الترجيح (*Y_weight*) على النحو المبين في الشكل 29 إلى قيم *si_loss* من كل منطقة مكانية زمنية فرعية (أي قيم *si_loss* بعد إعمال دالة مقارنة فقدان النسبة في كل منطقة مكانية زمنية فرعية ولكن قبل دوال الطي الزمني). فتساوي دالة الترجيح *Y_weight* واحداً (أي كامل الترجيح) حتى يبلغ متوسط مستوى النصوع 175 ثم تنخفض خطياً إلى الصفر فيما تتزايد قيم النصوع من 175 إلى 255. ويُطبَّق التصحيح الوسيط حصراً على قيم *si_loss*، لا على قيم *si_gain*.

الشـــكل 29

دالة الترجيح *Y* – ترجيح لتعديل معلمات *si_loss* المكانية الزمنية (S-T)



BT.1885-29

وتحسب معلمتان مكانيتان من الخاصية f_{HV13}، فتقيس إحداهما خسارة المعلومات المكانية الأفقية والرأسية (HV) النسبية (*hv_loss*) وتقيس الأخرى الكسب (*hv_gain*). وباتباع الترميز الرياضي في الملحق 8.D بالتوصية ITU-T J.144، تعطى هاتان المعلمتان بالصيغ التالية:

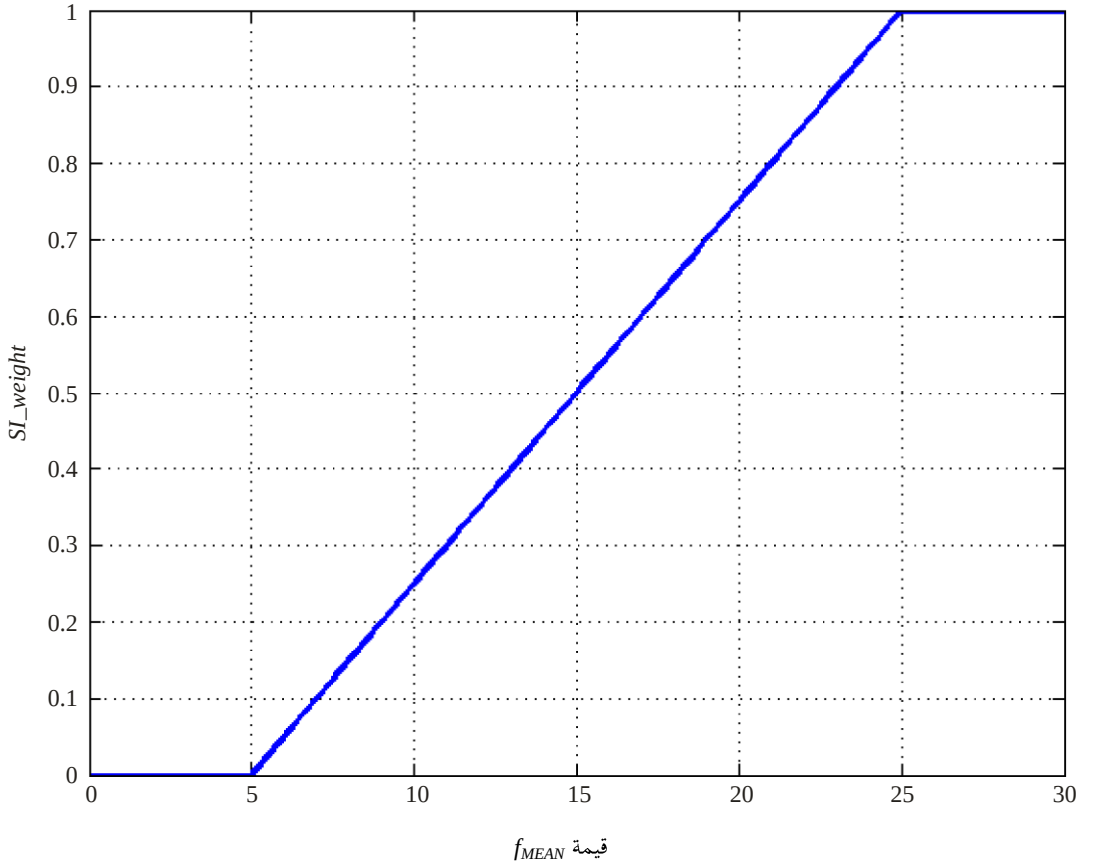
$$hv_loss = \text{avg1s_Y_hv13_angle0.225_rmin20_30x30_mean_4_ratio_loss_...}$$
$$\text{OMB(3,3,2)below1\%_Minkosski(1,1.5)_clip_0.08}$$

$$hv_gain = \text{avg1s_Y_hv13_angle0.225_rmin20_30x30_mean_4_log_gain_...}$$
$$\text{clip_0.06_OMB(3,3,2)above99\%tail_Minkosski(1.5,3)}$$

ولا يظهر في المعادلات أعلاه أن دالة *Y_weight* المبينة في الشكل 29 تطبَق أيضاً على قيم *hv_loss* و*hv_gain* على السواء من كل منطقة مكانية زمنية فرعية قبل دوال الطي المكاني والزمني (بعد حسابات خسارة النسبة والكسب اللوغاريتمي، على التوالي). وتطبَق دالة ترجيح إضافية (*SI_weight*) على النحو المبين في الشكل 30) على قيم *hv_loss* من كل منطقة مكانية زمنية فرعية تقل فيها كثيراً المعلومات المكانية (أي تنخفض فيها قيم خاصية f_{SI13} الأصلية).

الشـــكل 30

دالة الترجيح *SI* – ترجيح لتعديل معلمات *hv_loss* المكانية الزمنية (S-T)



BT.1885-30

ويمكن أن تُسحق معلمات التشوه المكاني (أي الحد من أو ضغط الجنوح بعيداً عما يصادَف في بيانات التدريب) باستخدام دوال مثل دالة سحق نموذج النوعية الفيديوية (VQM). الواردة في فقرة حساب نموذج النوعية الفيديوية.

5.3.3 المعلمات الزمنية

تحسب معلمتان زمنيتان من الخاصية f_{ATI}، فتقيس إحداهما الضوضاء العشوائية المضافة (*ati_noise*) وتقيس الأخرى اضطرابات الحركة الناجمة عن أخطاء الإرسال (*ati_error*). وباتباع الترميز الرياضي في الملحق 8.D بالتوصية ITU-T J.144، تعطى هاتان المعلمتان بالصيغ التالية:

$$ati_noise = \text{Y_rand5\%_ati0.2s_rms_5_ratio_gain_between25\%50\%}$$

$$ati_error = \text{Y_rand5\%_ati0.2s_rms_max7pt_12_ratio_gain_above90\%}$$

ولزيادة تحصين معلمتي *ati_noise* و*ati_error* ضد الانحرافات الزمنية، يتم حسابهما من أجل جميع الاصطفافات الزمنية للفيديو المعالَج التي تقع ضمن ± 0,4 s من أفضل اصطفاف زمني مقدر مع الفيديو الأصلي، ثم تُختار القيمة الدنيا للمعلمة.

4.3 حساب نموذج النوعية الفيديوية (VQM)

على غرار نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات (NTIA) الوارد في الملحق D بالتوصية ITU-T J.144، فإن حساب نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع يجمع خطياً بين معلمتين من الخاصية f_{HV13} (*hv_loss* and *hv_gain*)، ومعلمتين من الخاصية f_{SI13} (*si_loss* and *si_gain*)، ومعلمتين من الخاصية f_{COHER_COLOR} (غير أن المعلمتين قد جُمع بينهما في معلمة تشوه لون واحدة تدعى *color_comb*). وقد استعيض عن معلمة الضوضاء الواحدة في نموذج النوعية الفيديوية العام للإدارة الوطنية للاتصالات والمعلومات (NTIA) بمعلمتين (*ati_noise* و*ati_error*) على أساس خاصية f_{ATI} ذات عرض النطاق المنخفض الموصوفة هنا.

ومن ثم، فإن نموذج VQM_{FLB} (اختصار لنموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع) يتألف من توليفة خطية لثماني معلمات. ويعطى VQM_{FLB} بالصيغ التالية:

$$VQM_{FLB} = \{\ 0.38317338378290 * hv_loss + 0.37313218013131 * hv_gain +$$
$$0.58033514546526 * si_loss + 0.95845512360511 * si_gain +$$
$$1.07581708014998 * color_comb +$$
$$0.17693274495002 * ati_noise + 0.02535903906351 * ati_error\ \}$$

ويُقص مجمل نموذج VQM (بعد إضافة المساهمات من جميع المعلمات) في عتبة أدنى قيمتها 0,0 لمنع أرقام VQM سلبية. وأخيراً، تطبَّق دالة السحق التي تسمح بتطبيق تجاوز أقصاه 50% على قيم VQM الزائدة عن 1,0 تقييداً لقيم VQM في مشاهد الفيديو المشوهة للغاية التي تقع خارج نطاق بيانات التدريب.

إذا $VQM_{FLB} > 1,0$، عندئذ $VQM_{FLB} = (1 + c)*VQM_{FLB} / (c + VQM_{FLB})$، حيث $c = 0,5$

وسيكون لنموذج VQM_{FLB} المحسوب بالطريقة أعلاه قيم تساوي أو تزيد عن الصفر، وواحد كقيمة قصوى اسمية. ويمكن لنموذج VQM_{FLB} أن يتجاوز أحياناً قيمة الواحد في مشاهد الفيديو المشوهة للغاية.

ولزيادة تحصين نموذجVQM_{FLB} ضد الانحرافات المكانية، فهو يُحسب من أجل جميع الاصطفافات المكانية للفيديو المعالَج التي تقع ضمن بيكسل واحد، زيادة أو نقصان، من أفضل اصطفاف مكاني مقدر مع الفيديو الأصلي، ثم تُختار القيمة الدنيا لنموذج VQM_{FLB}.

4 المراجع

[1] التوصية ITU-R BT.601-6 (01/07) – معلمات التشفير في الأستوديو للتلفزيون الرقمي في النسبتين الباعيتين 4:3 (المعيارية) و 16:9 (شاشة عريضة).

[2] Video Quality Model (VQM) Software Tools – Binary executables and source code, available from the National Telecommunications and Information Administration (NTIA) at http://www.its.bldrdoc.gov/n3/video/VQM_software.php.

[3] التوصية ITU-T J.244 (04/08) – طرائق معايرة كاملة المراجع ومختصرة المراجع في أنظمة الإرسال الفيديوي مع سوء تراصف دائم للميادين المكانية والزمنية بكسب وتخالف دائمين.

[4] VQEG Final Report of MM Phase I Validation Test (2008), "Final report from the Video Quality Experts Group on the validation of objective models of multimedia quality assessment, phase I", Video Quality Experts Group (VQEG), http://www.its.bldrdoc.gov/vqeg/projects/multimedia, ITU-T Study Group 9 TD923, Study Period 2005-2008.

[5] التوصية ITU-T J.144 (03/04) – تقنيات القياس الموضوعي لما يُلمس من النوعية الفيديوية المعدة للتلفزيون الكبلي الرقمي بوجود مرجع كامل.

[6] التوصية ITU-R BT.1683 (06/04) – تقنيات القياس الموضوعي لما يُلمس من النوعية الفيديوية المعدة لتلفزيون الإذاعة الرقمية عادي الوضوح بوجود مرجع كامل.

5 الشفرة المرجعية لتنفيذ نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع

الغرض من هذه الشفرة المرجعية هو مساعدة المستخدم في التنفيذ السليم لنموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع. وفيما تُستخدم شفرة MATLAB® code كالشفرة المرجعية، يمكن استخدام أي شفرة برمجية لإعادة إنتاج النتائج الواردة هنا. وتحوي كل فقرة فرعية من الفقرة 5 شفرة MATLAB للدالة المسماة في رأس الفقرة (مثل حفظ محتويات الفقرة 1.5 في ملف يدعى "fastlowbw_ref.m"). وبتنفيذ دالة fastlowbw_ref بدون بيانات دخل، يمكن الحصول على معلومات مساعدة بشأن كيفية استدعاء البرنامج الفرعي. وتتسم هذه الشفرة بمرونة تشغيل النموذج على مقطع فيديو قصير (أي مدته من 5 ثوان إلى 15 ثانية) ضمن تتابع فيديو أطول (كتتابع مدته دقيقة واحدة). ويتم ذلك بزحزحة مقطع فيديو قصير ثانية واحدة ومعاودة حساب النموذج لكل انزياح زمني. وفي حين لا تبيَّن هذه الميزة الوظيفية أدناه، فإن التعقيبات ضمن الشفرة والبيانات العائدة من دالة "model_fastlowbw_parameters.m" ستحيل إلى هذه القدرة. المستترة التي يمكن أن الاستفادة منها في تنفيذ نظام لمراقبة النوعية الفيديوية ضمن الخدمة.

وعند معالجة متجهات اختبار العينة (أي مقاطع الفيديو) ذات الشفرة المرجعية لتنفيذ نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع (الدالة "fastlowbw_ref.m")، يتم إنتاج ملفات نصية تحتوي على نتائج المعايرة والنموذج. وفي المثال التالي على استدعاءات دوال MATLAB، ينبغي الحصول على ملفات خرج تماثل تلك الواردة أدناه (نظراً للعمليات العشوائية التي يستخدمها نموذج النوعية الفيديوية (VQM) بعرض النطاق الضيق السريع، قد تختلف النتائج قليلاً عن تلك المعروضة هنا):

5.Reference Code.doc

التذييل

تحاليل خطأ الإرسال

تضمنت اختبارات التحقق من VQEG لمشروع RRNR-TV نسقي (NTSC) 525 و(PAL) 625. وشملت كل تجربة 12 تتابع مصدر و156 تتابع فيديو معالَج (PVSs). وفي 156 تتابع فيديو معالَج، احتوى 40 منها على أخطاء إرسال واحتوى 116 منها على أخطاء تشفير فقط. ويظهر الجدولان 9 و10 الخطأ الفعال (RMSE) ومجال التشغيل (OR) لتتابعات الفيديو المعالَج المشوبة بأخطاء إرسال. وتجدر الإشارة إلى أن الخطأ الفعال (RMSE) ومجال التشغيل (OR) حسبا بواسطة خطوط الارتداد المحصَّلة من مجمل البيانات. وبعبارة أخرى، حُسبت خطوط الارتداد باستخدام مجمل البيانات. ثم حُسب الخطأ الفعال (RMSE) ومجال التشغيل (OR) في أخطاء الإرسال باستخدام تتابعات الفيديو المعالَج المشوبة بأخطاء إرسال.

الجـــدول 9

الخطأ الفعال (RMSE) ومجال التشغيل (OR) في اختبار التحقق من RRNR-TV (النسق 525). TE: أخطاء الإرسال

النسق 525	الكل		مع TE		بدون TE	
	RMSE	OR	RMSE	OR	RMSE	OR
النموذج_A_ 15k	0,418	0,385	0,574	0,500	0,362	0,293
النموذج_A_80k	0,423	0,378	0,582	0,475	0,366	0,293
النموذج_A_ 256k	0,424	0,378	0,584	0,475	0,367	0,293
النموذج_B_ 80k	0,598	0,667	0,768	0,650	0,544	0,586
النموذج_B_256k	0,587	0,647	0,763	0,600	0,530	0,578
النموذج_C_80k	0,465	0,513	0,557	0,550	0,440	0,405
النموذج_C_256k	0,511	0,609	0,584	0,450	0,495	0,578
PSNR_NTIA	0,556	0,571	0,549	0,500	0,568	0,491

الجـــدول 10

الخطأ الفعال (RMSE) ومجال التشغيل (OR) في اختبار التحقق من RRNR-TV (النسق 625). TE: أخطاء الإرسال

النسق 625	الكل		مع TE		بدون TE	
	RMSE	OR	RMSE	OR	RMSE	OR
النموذج_A_ 15k	0,524	0,468	0,597	0,450	0,508	0,414
النموذج_A_80k	0,513	0,462	0,594	0,500	0,494	0,379
النموذج_A_ 256k	0,516	0,468	0,593	0,500	0,499	0,379
النموذج_B_ 80k	0,887	0,724	0,545	0,500	0,986	0,716
النموذج_B_256k	0,864	0,744	0,523	0,600	0,962	0,716
النموذج_C_80k	0,585	0,583	0,282	0,200	0,663	0,647
النموذج_C_256k	0,657	0,590	0,292	0,175	0,747	0,638
PSNR_NTIA	0,605	0,564	0,338	0,250	0,678	0,517