

国 际 电 信 联 盟

ITU-R

国际电联无线电通信部门

ITU-R BT.1908-0 建议书

(01/2012)

**在有减弱参考信号的情况下
采用HDTV的广播应用的
客观视频质量测量技术**

BT 系列

广播业务

(电视)



国际电信联盟

前言

无线电通信部门的职责是确保卫星业务等所有无线电电信业务合理、平等、有效、经济地使用无线电频谱，不受频率范围限制地开展研究并在此基础上通过建议书。

无线电通信部门的规则和政策职能由世界或区域无线电通信大会以及无线电通信全会在研究组的支持下履行。

知识产权政策（IPR）

ITU-R的IPR政策述于ITU-R第1号决议中所参引的《ITU-T/ITU-R/ISO/IEC的通用专利政策》。专利持有人用于提交专利声明和许可声明的表格可从<http://www.itu.int/ITU-R/go/patents/en>获得，在此处也可获取《ITU-T/ITU-R/ISO/IEC的通用专利政策实施指南》和ITU-R专利信息数据库。

ITU-R系列建议书

（也可在线查询<http://www.itu.int/publ/R-REC/en>）

系列	标题
BO	卫星传送
BR	用于制作、存档和播出的录制；电视电影
BS	广播业务（声音）
BT	广播业务（电视）
F	固定业务
M	移动、无线电定位、业余和相关卫星业务
P	无线电波传播
RA	射电天文
RS	遥感系统
S	卫星固定业务
SA	空间应用和气象
SF	卫星固定业务和固定业务系统间的频率共用和协调
SM	频谱管理
SNG	卫星新闻采集
TF	时间信号和频率标准发射
V	词汇和相关问题

说明： 该ITU-R建议书的英文版本根据ITU-R第1号决议详述的程序予以批准。

电子出版
2020年，日内瓦

© 国际电联 2020

版权所有。未经国际电联书面许可，不得以任何手段复制本出版物的任何部分。

ITU-R BT.1908-0 建议书*

在有减弱参考信号的情况下采用HDTV的广播应用的
客观视频质量测量技术

(2012年)

范围

本建议书规定了在有减弱的参考信号可以利用的情况下，评估采用HDTV的广播应用的感知视频质量的方法。

国际电联无线电通信全会，

考虑到

- a) 自动测量广播视频质量的能力长期以来被认为是该行业的一个重要优势；
- b) ITU-R BT.1683建议书描述了在有减弱参考信号的情况下，测量标准清晰度数字广播电视的感知视频质量的客观方法；
- c) ITU-R BT.709建议书描述了适用于制作和国际节目交流的HDTV标准的参数值，ITU-R BT.500建议书描述了高清晰度电视中图像质量的主观评价方法；
- d) HDTV正在广播领域得到广泛地应用；
- e) ITU-T第9研究组已经根据VQEG发布的HDTV报告的结果，制定了ITU-T J.341建议书，该建议书规定了在有减弱参考信号的情况下HDTV的客观视频质量测量方法；
- f) HDTV感知视频质量的客观测量可以作为主观评价方法的补充，

建议

1 在附件1中给出的客观视频质量模型应该用于在有减弱参考信号的情况下采用HDTV的广播应用的感知视频质量的客观测量。

* 无线电通信第6研究组于2020年2月根据ITU-R第1号决议对此建议书进行了编辑性修正。

附件1

1 引言

本建议书提供了在能够采用减弱参考信号（RR）测量方法的情况下，适用于高清晰度电视（HDTV）非交互式应用的视频质量测量方法。该模型与使用ITU-R BT.500建议书获得的主观质量得分进行比较，分析表明该模型的准确度与PSNR的准确度相当。

为了正确地操作RR模型，该模型应利用无损信源视频提取参数。这些提取的参数以及劣化的视频序列是RR模型的输入，评估方法既进行校准（例如增益/偏移和空间/时间配准），又进行客观视频质量评估。

验证测试的素材既包括H.264和MPEG-2编码劣化，又包括各种传输差错情况（例如比特错、丢失数据包）。本建议书中的模型可以用于监视已部署网络的质量以确保它们运行准备就绪。劣化的视觉效果可能包括空间以及时间的劣化，本建议书中的模型也能用于视频系统的实验室测试，当该模型用于比较不同的视频系统时，对于这种特殊情形，建议采用一种定量方法（例如ITU-T J.149建议书中描述的方法）来确定模型的准确度。

本建议书被认为适用于传输速率在1 Mbit/s到30 Mbit/s之间的广播业务，验证测试时设想采用下列分辨率和帧频：

- 1080/59.94/I
- 1080/25/P
- 1080/50/I
- 1080/29.97/P。

对于每一种分辨率，进行验证测试时要考虑到以下条件：

测试因素
视频分辨率：1920×1080隔行和逐行
视频帧频：每秒29.97帧和每秒25帧
视频比特率：1到30Mbit/s
最长2秒的短暂定格（带跳动的暂停）
丢失数据的传输错误
从1080到720/P的SRC转换，压缩、传输、解压缩，再转换回到1080。
编码技术
H.264/AVC（MPEG-4 第10部分）
MPEG-2

注意到在验证测试计划中720/P被当作是测试条件（HRC）的一部分。由于目前720/P通常被提升为显示的一部分，因此认为720/P HRC会更适合处理这种格式。

1.1 应用

适合于本建议书描述评估模型的应用包括但不限于：

- 1) 有侧信道可利用时的视频质量监视。
- 2) 在位于发送点和接收点之间的测量节点的视频质量监视。

本建议书中描述的模型提供了与PSNR在统计上相似的性能，然而，它能够在测量点可以利用减弱参考信号的情况下用于视频质量的评估。

1.2 限制

本建议书中描述的视频质量评估模型不能用于取代主观测试，两个精心设计和实施的主观测试（即在两个不同的实验室）之间的相关值通常在0.95到0.98的范围内，本建议书不能用于进行视频系统比较（例如比较两个编解码器，比较相同压缩算法的两种不同实现），本建议书中描述的视频质量评估模型的性能在统计上不会优于PSNR。

当出现帧定格时，测试条件通常是帧定格的持续时间小于2秒。本建议书模型不适用于重新缓存条件下（即时延或者无跳动定格总是递增的视频）视频质量的测量，该模型没有在除TV系统所使用那些帧频（即每秒29.97帧和每秒25帧，采用隔行或者逐行模式）以外的帧频上被测试过。

应该注意到如果新的编码和传输技术产生的人为现象没有包含在这次评估中，客观模型可能会产生错误的结果，这时就需要主观评价。

注意到本建议书模型不能评估视频会议场景特有的发言人头部特写内容。

2 参考文献

下列ITU-T建议书和其他参考文献的条款，通过在本建议书引用而构成本建议书的条款。在出版时，所指出的版本是有效的。所有的建议书和其他参考文献均会得到修订；因此本建议书的使用者应查证是否有可能使用下列建议书和其他参考文献的最新版本。当前有效的ITU-T建议书清单定期出版。

本建议书引用的文件自成一体时不具备建议书的地位。

ITU-T J.244建议书(2008)，《用于视频传输系统的完全基准和简化基准校验法，该系统空间域和临时域有常数增益和偏移但常数未校准》。

3 定义

3.1 在其他地方定义的术语

本建议书使用了下列在其他地方定义的术语：

3.1.1 主观评价（图像）：确定在观看期间呈现给评估小组的节目类图像的质量或者损伤。

3.1.2 客观感知测量（图像）：利用节目类图像和客观（采用仪器的）测量方法测量系列节目的性能，从而获得一个与主观评价测试得到的评定值相接近的指示。

3.1.3 支持者：在某一个国际电联建议书中提出适用于验证测试的视频质量模型以及可能包含物的组织或者公司。

3.2 本建议书中的定义的术语

本建议书中定义了下列术语：

3.2.1 帧频：每秒不同帧的数量（即全部帧–重复的帧）。

3.2.2 模拟传输错误：在高度受控的环境下发生在数字视频比特流中的错误，错误举例包括模拟包丢失率和模拟比特错误，明确规定了用于控制模拟传输错误的参数。

3.2.3 传输错误：发生在视频传输中的任何错误，错误类型举例包括模拟传输错误和实际网络条件。

4 缩写和首字母缩略语

本建议书采用了下列缩写和首字母缩略语：

ACR	绝对等级评定值（见ITU-R BT.500建议书）
ACR-HR	具有隐藏参考的绝对等级评定值（见ITU-T P.910建议书）
AVI	音频视频交错
DMOS	差分平均评价得分
FR	完整参考信号
FRTV	完整参考电视
HRC	假设基准电路
ILG	VQEG的独立实验组
MOS	平均评价得分
MOS _p	预测的平均评价得分
NR	无（或者零）参考
PSNR	峰值信噪比
PVS	处理过的视频序列
RMSE	均方根误差
RR	减弱的参考信号
SFR	信源帧频
SRC	信源参考信道或电路
VQEG	视频质量专家组
YUV	颜色空间和文件格式

5 惯例

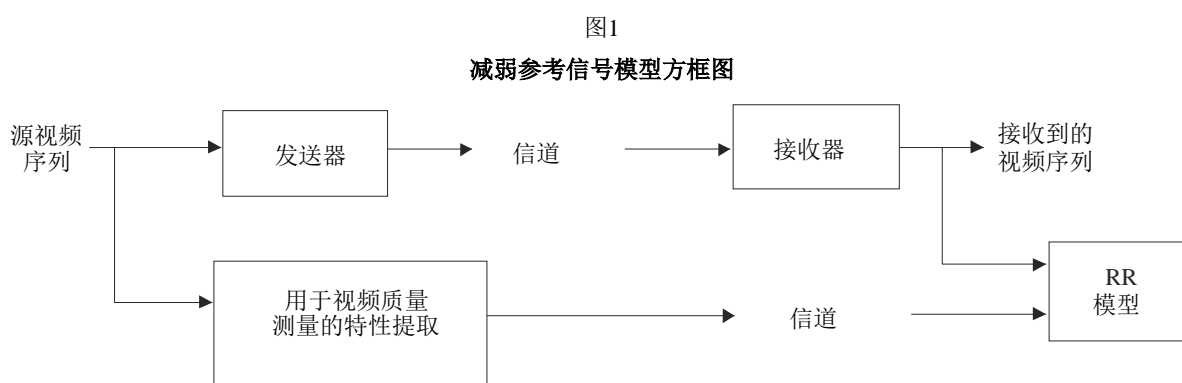
无。

6 减弱参考信号测量方法描述

6.1 引言

尽管PSNR已经作为一种客观视频质量测量方法得到广泛使用，但据报告PSNR不能很好地表现感知视频质量，通过分析人类如何感知视频的质量，可以看出人类视觉系统对边缘周围的劣化敏感，换言之，当视频的边缘像素变模糊时，评估者倾向于给视频低分，即便PSNR较高。基于这一观察，已经开发出了主要测量边缘劣化的减弱参考信号模型。

图1说明的是一个减弱参考信号模型是如何工作的，从信源视频序列中提取并发送将要用于在监视点测量视频质量的特性，表1显示的是用于传输这些特性的侧信道带宽，该带宽已经在VQEG HDTV测试中被测试过。



BT.1908-01

表 1
侧信道带宽

视频格式	已测试的带宽
1080/60 Hz (29.97 fps) 1080/30Pp (29.97 fps)	56 kbit/s, 128 kbit/s, 256 kbit/s
1080/25Pp (25 fps) 1080/50I Hz (25 fps)	56 kbit/s, 128 kbit/s, 256 kbit/s

6.2 EPSNR减弱参考信号模型

6.2.1 边缘PSNR (EPSNR)

RR模型主要测量边缘上的劣化，在该模型中，首先对信源视频序列应用边缘检测算法以定位边缘像素，然后，通过计算均方误差来测量那些边缘像素的劣化，由这个均方误差计算边缘PSNR。

人们可以使用任何边缘检测算法，尽管结果可能会有细小的差异，例如，人们能够使用任何梯度算子来定位边缘像素。已经提出了许多的梯度算子，在很多边缘检测算法中，首先使用梯度算子计算水平梯度图像 $g_{horizontal}(m,n)$ 和垂直梯度图像 $g_{vertical}(m,n)$ ，然后可以按下式计算绝对梯度图像 $g(m,n)$ ：

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

最后，对绝对梯度图像进行门限操作来寻找边缘像素，换言之，那些绝对梯度超过某一门限值的像素被认为是边缘像素。

图2-6说明的是这个过程。图2显示的是源图像，图3显示的是通过对图2中的源图像应用水平梯度算子获得的水平梯度图像 $g_{horizontal}(m,n)$ ，图4显示的是通过对图2中的源图像应用垂直梯度算子获得的垂直梯度图像 $g_{vertical}(m,n)$ ，图5显示的是绝对梯度图像（边缘图像），图6显示的是通过对图5中的绝对梯度图像应用门限操作获得的二进制边缘图像（掩模图像）。

图 2

源图像（原始图像）



图3

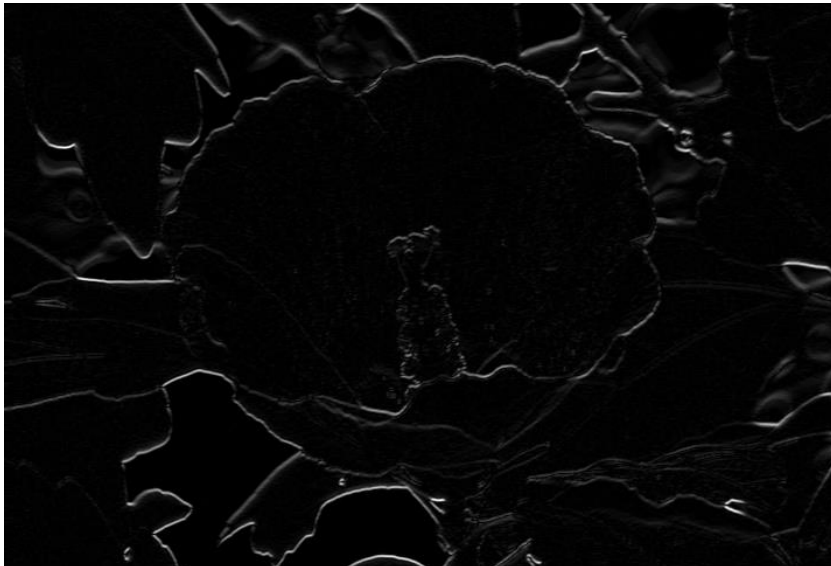
对图2中的源图像应用水平梯度算子得到的水平梯度图像



BT.1908-03

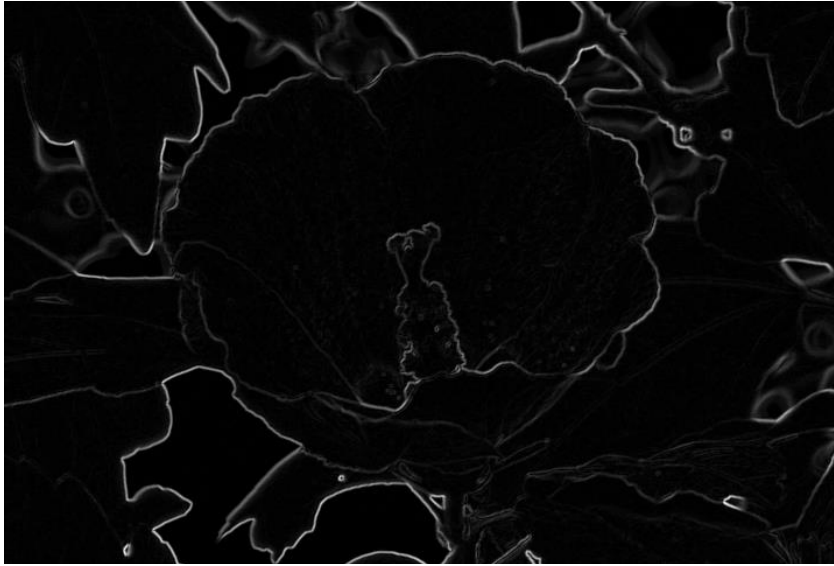
图4

对图2中的源图像应用垂直梯度算子得到的垂直梯度图像



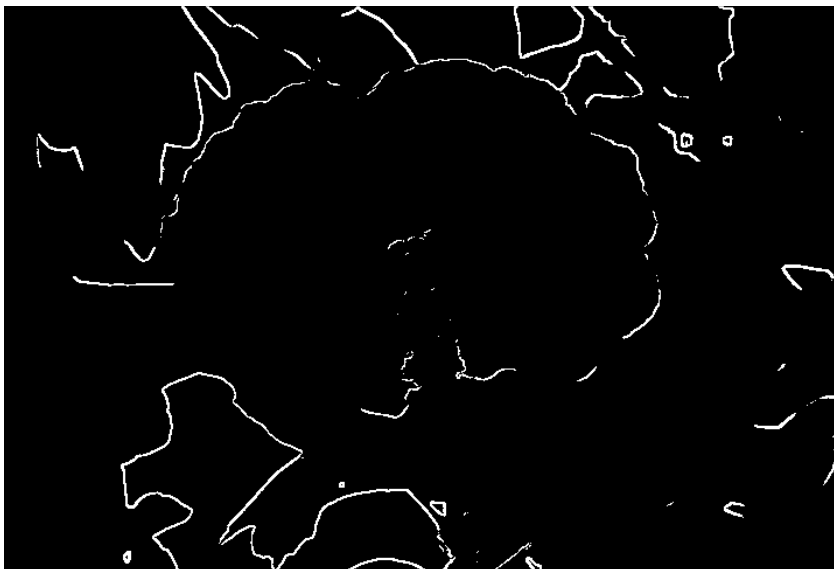
BT.1908-04

图5
绝对梯度图像



BT.1908-05

图6
对图5中的绝对梯度图像应用门限得到的二进制边缘图像（掩模图像）

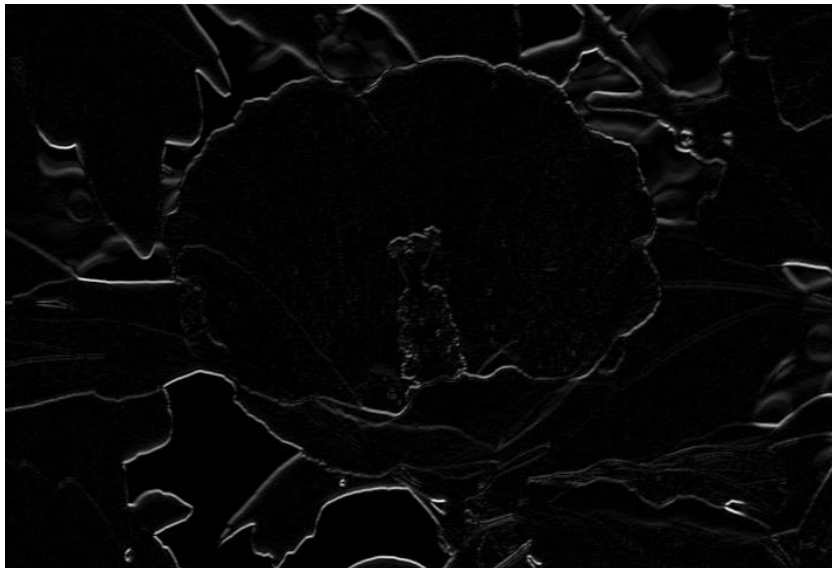


BT.1908-06

或者，人们可以采用一个改进的程序来寻找边缘像素，例如，人们可以首先对源图像应用垂直梯度算子，产生一个垂直梯度图像。然后，对垂直梯度图像应用水平梯度算子，产生一个改进的连续梯度图像（水平和垂直梯度图像），最后，可以对改进的连续梯度图像应用门限操作来寻找边缘像素，换言之，改进的连续梯度图像中的像素，超过门限值的被认为是边缘像素。图7-9说明了这个改进的程序，图7显示的是通过对图2中的源图像应用垂直梯度算子得到的垂直梯度图像 $g_{vertical}(m,n)$ ，图8显示的是对图7中的垂直梯度图像应用水平梯度算子得到的改进的连续梯度图像（水平和垂直梯度图像），图9显示的是对图8中的改进的连续图像应用门限操作得到的二进制边缘图像（掩模图像）。

图7

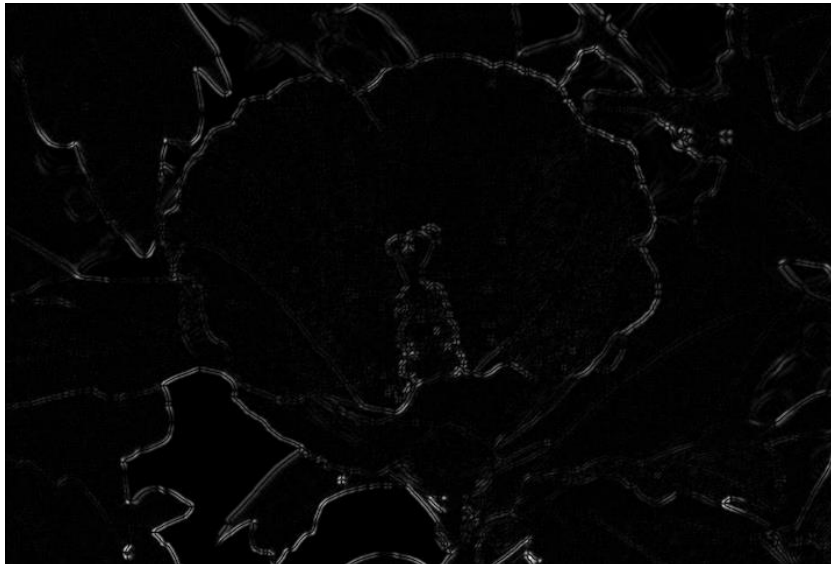
对图2中的源图像应用垂直梯度算子得到的垂直梯度图像



BT.1908-07

图8

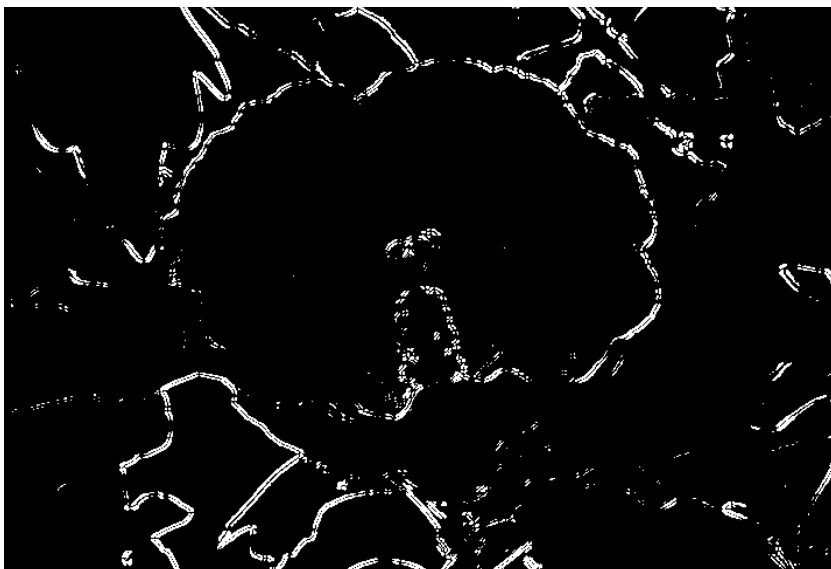
对图7中的垂直梯度图像应用水平梯度算子得到的改进的连续梯度图像（水平和垂直梯度图像）



BT.1908-08

图9

对图8中的改进的连续梯度图像应用门限得到的二进制边缘图像（掩模图像）



BT.1908-09

注意到这两种方法均可作为边缘检测算法，人们可以根据视频和压缩算法的性质选择任意一种边缘检测算法。然而，方法会有优劣之分。

因此，在该模型中，首先采用边缘检测算子产生边缘图像（图5和图8），然后，对边缘图像应用门限产生掩模图像（二进制边缘图像，图6和图9），换言之，边缘图像中的像素，如果其数值小于门限 t_c ，则设为0，如果其数值大于或等于门限，则设为非零的数值。图6和图9显示的是一些掩模图像，由于视频可以看作是帧或场的一个序列，上述的程序能够应用于视频中的每一帧或每一场。由于该模型能够用于基于场的视频或者基于帧的视频，术语“图像”将用于表示一场或者一帧。

6.2.2 从源视频序列中选择特性

由于模型是减弱参考信号（RR）模型，因此需要从源视频序列中每一个图像中提取一些特性。在EPSNR RR模型中，要从每一幅图像中选择一定数量的边缘像素，然后，对位置和像素数值进行编码和发送，然而，对于一些视频序列，当采用一个固定的门限数值时，边缘像素的数量会非常少。在最差的情况下，边缘像素的数量可以为0（空白图像或者极低频率图像）。为了解决这个问题，如果图像的边缘像素的数量小于一个给定的数值，用户可以降低门限值直到边缘像素的数量大于一个给定的数值，或者，人们能够选择与水平和垂直梯度图像的最大数值相对应的边缘像素，当一帧中没有边缘像素时（例如空白图像），人们可以随机选择要求数量的像素或者跳过这一帧。例如，如果要从每一帧中选择10个边缘像素，人们可以根据像素的数值对水平和垂直梯度图像的像素进行分类，并选择最大的10个数值，然而，这个方法可能会在同一个位置上产生多个边缘像素，为了解决这个问题，人们可以首先选择几倍于要求数量的水平和垂直梯度图像的像素，然后从选出的水平和垂直梯度图像像素中随机地选择要求数量的边缘像素。在VQEG HDTV测试被测试过的模型中，要求数量的边缘像素是从大量的边缘像素中随机选出的，而大量的边缘像素是通过梯度图像应用门限操作获得的。

在EPSNR RR模型中，选择的像素位置经过高斯低通滤波器之后，对位置和边缘像素值进行编码。尽管VQEG HDTV测试中使用的是高斯LPF（ 7×3 ），但可以根据视频格式采用不同的低通滤波器。注意到在编码过程期间，可能要进行裁剪。模型要选择位于中间区域中的边缘像素（图10），以避免选择位于被裁剪掉区域中的边缘像素。表2显示的是经过裁剪的图像的大小，还显示了对边缘像素的位置和像素数值进行编码所需要的比特数。

表2

每个边缘像素的比特数要求

视频格式	大小	裁剪后的大小	用于位置的比特数	用于像素数值的比特数	每个像素的总比特数
HD 逐行的	1 920 × 1 080	1 856 × 1032	21	8	29
HD 隔行的	1 920 × 540	1 856 × 516	20	8	28

图10
裁剪和中间区域举例



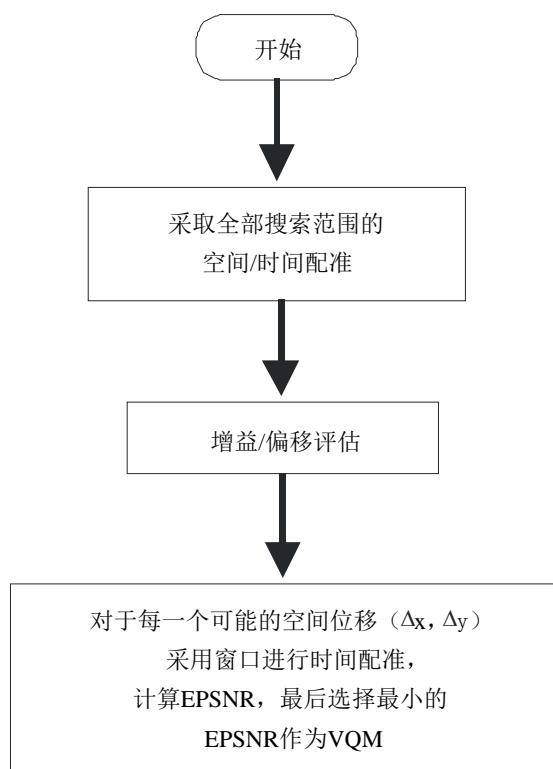
BT.1908-10

模型根据允许的带宽从每一帧中选择边缘像素（表1），表3显示的是已测试过的带宽能够发送的每一帧中边缘像素的数量。

表3
每一帧/场边缘像素的数量

视频格式	56 kbit/s	128 kbit/s	256 kbit/s
HD逐行的	46	105	211
HD隔行的	24	54	109

图11
模型流程图



BT.1908-11

6.2.3 空间/时间配准和增益/偏移调整

在计算源视频序列和那些处理过的视频序列之间的差值之前，处理过的序列是接收器接收到的视频序列，模型首先要进行空间/时间配准和增益/偏移调整，采用ITU-T J.244建议书中的校准方法（附件B）。为了发送ITU-T J.244建议书中的增益和偏移特性（附件B），在VQEG HDTV测试中使用了30%的有效带宽。当视频序列为隔行时，校准方法要使用三次：偶数场、奇数场和混合帧，而校准方法只对逐行视频序列中的帧使用。如果偶数场错误（PSNR）和奇数场错误之间差值超过某一门限，就使用具有较小PSNR的配准结果（x-位移，y-位移），否则，就使用混合帧的配准结果。在VQEG HDTV测试中，门限设为2dB。

在监视点，处理过的视频序列应与从源视频序列中提取的边缘像素对准，然而，如果侧信道带宽较窄，则只能使用源视频序列中的少数边缘像素（图12）。因此，如果使用单一帧进行时间配准，时间配准可能会不准确（图13）。为了解决这个问题，模型在时间配准时使用了一个窗口。模型构建了一个由大量相邻帧组成的窗口，取代使用处理过的视频序列中的单帧，来寻找最佳的时间偏移。图14说明的是这个过程。窗口内的均方误差计算如下：

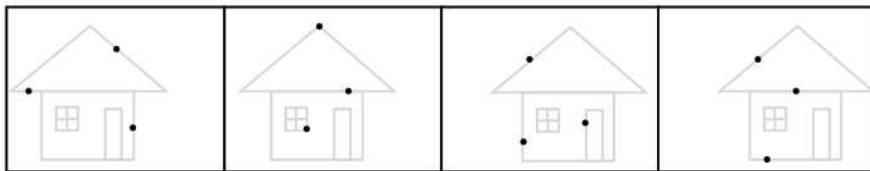
$$MSE_{window} = \frac{1}{N_{win}} \sum (E_{SRC}(i) - E_{PVS}(i))^2$$

其中 MSE_{window} 是窗口均方误差， $E_{SRC}(i)$ 是窗口内的边缘像素，该像素在处理过的视频序列中有一个对应的像素， $E_{PVS}(i)$ 是处理过的视频序列中与边缘像素对应的像素， N_{win} 是用于计算 MSE_{window} 的边缘像素的总数，这个窗口均方误差用作处理过的视频序列中的帧和源视频序列中的对应帧之间差值。

窗口的大小由处理过的视频序列的特性确定，对于典型的应用，建议采用相当于2秒钟的窗口。或者，运用不同大小的窗口，采用能提供最小均方误差的最佳的窗口大小，此外，考虑到由传输误差产生的帧跳动，可以采用不同的窗口中心（图18）。

图12

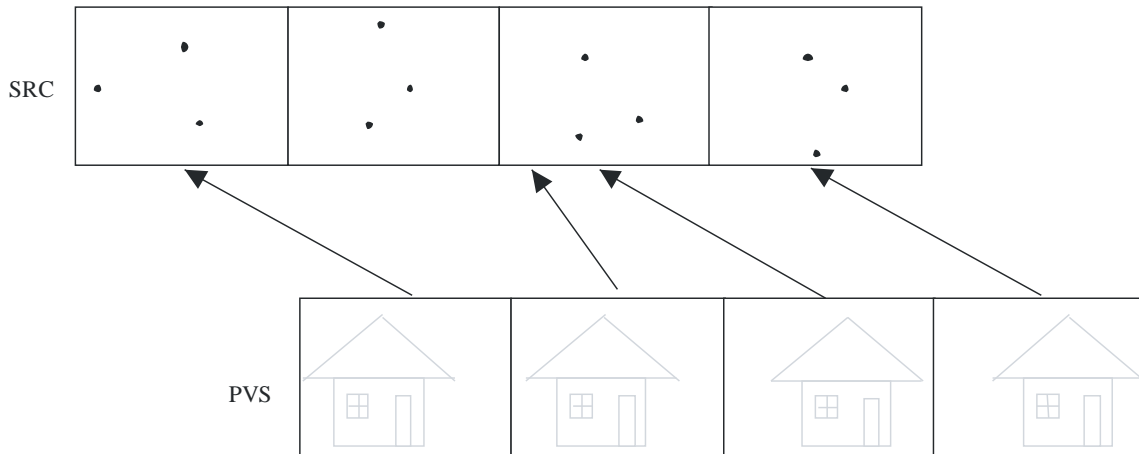
源视频序列的边缘像素选择



BT.1908-12

图13

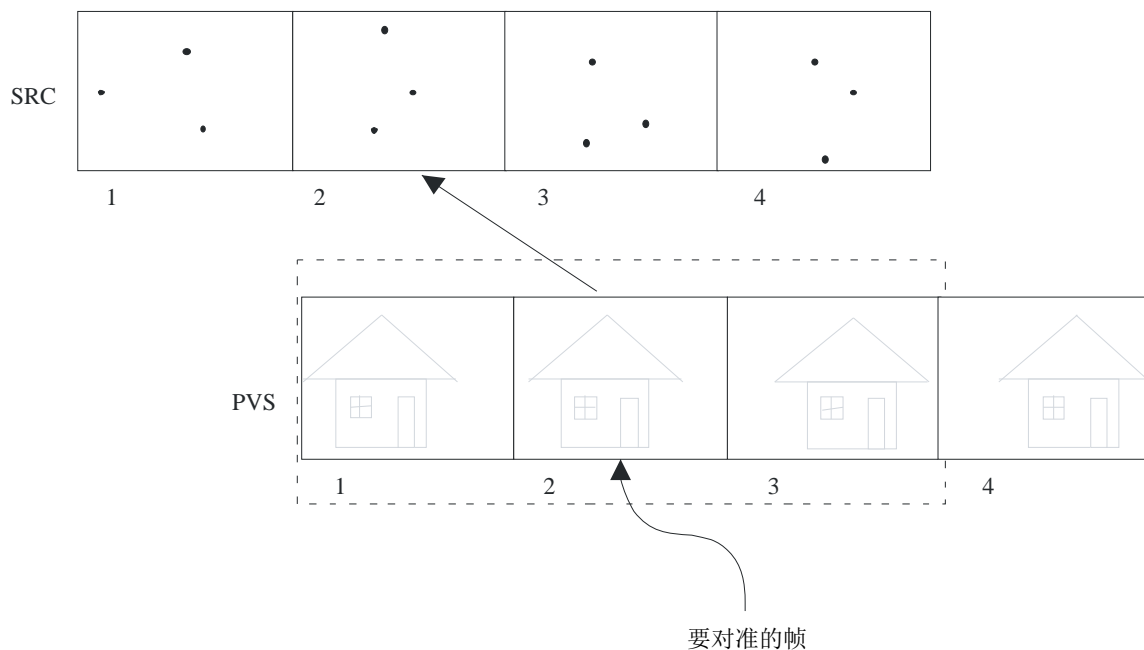
让处理过的视频序列对准源视频序列中的边缘像素



BT.1908-13

图 14

利用窗口让处理过的视频序列对准边缘像素

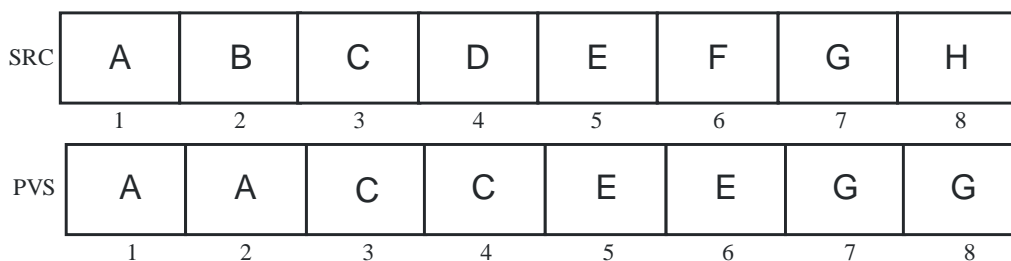


BT.1908-14

当源视频序列以高压缩比被编码时，编码器可能会减少每秒的帧数，处理过的视频序列有重复的帧（图15），在图15中，处理过的视频序列中没有与源视频序列中的某些帧（第2、4、6、8帧）相对应的帧，在这种情况下，模型在计算均方误差时不采用重复的帧，换言之，模型利用每个重复块的第一帧（有效帧）进行时间配准，因此，图16中，窗口中只有3个帧（第3、5、7帧）用于时间配准。

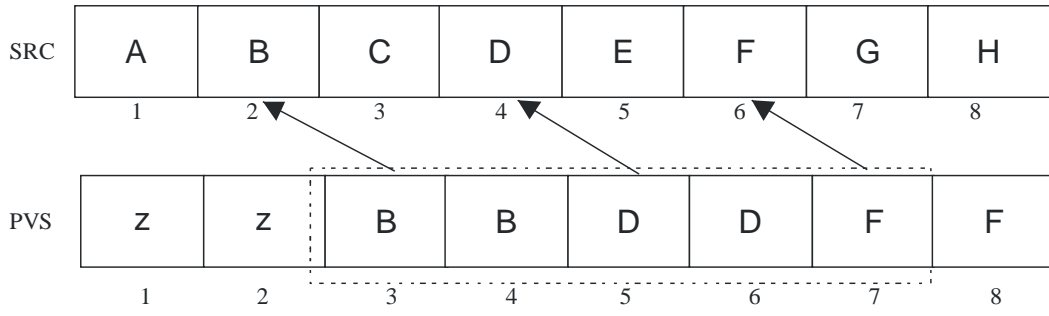
图15

重复帧举例



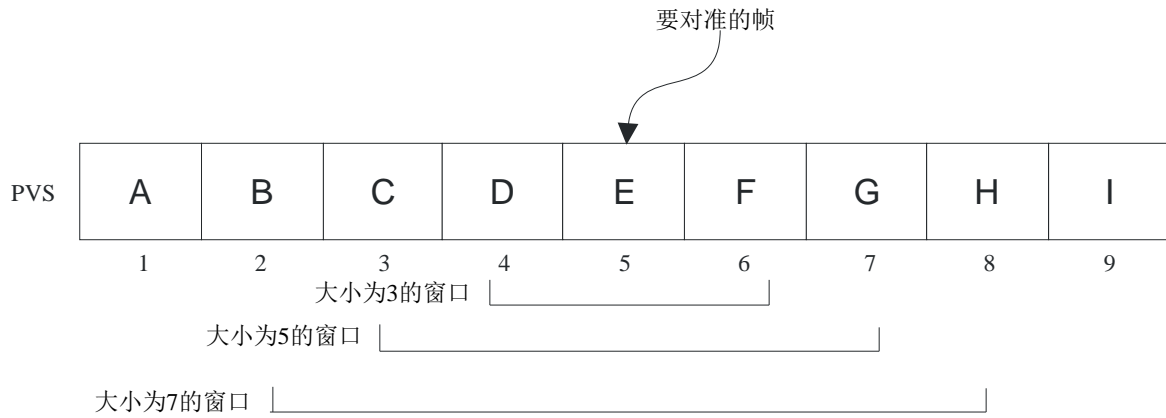
BT.1908-15

图16
处理重复帧

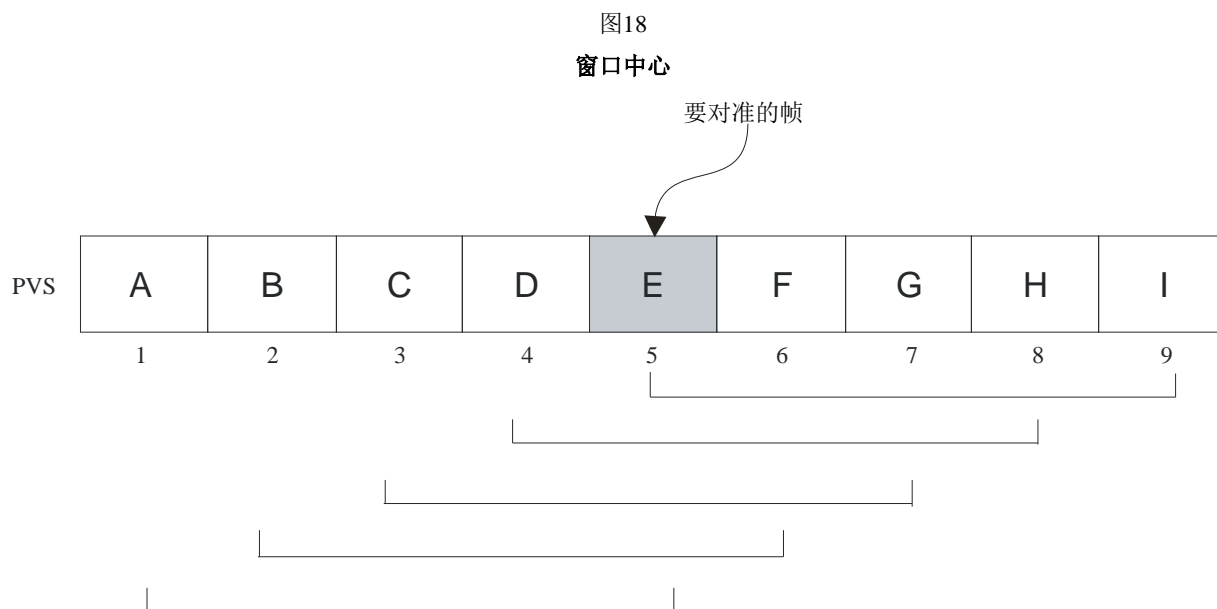


BT.1908-16

图17
不同大小的窗口



BT.1908-17



BT1908-18

6.2.4 计算EPSNR和后处理

在时间配准完成以后，计算源视频序列中的边缘像素和处理过的视频序列中的相应像素之间差值的平均值，这个平均值可以被理解为处理过的视频序列的边缘均方误差 (MSE_{edge})。最后，计算EPSNR（边缘PSNR）如下：

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

其中 p 是图像的峰值。

由于各种损伤能够降低视频的质量，因此，要针对这些在下一个小节中被量化的影响，调整EPSNR的数值。

1) 块效应度量I

针对块效应的影响，要计算列差值的平均值。假设以8为模，第 i 帧的块效应得分计算如下：

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

对帧块效应得分取平均计算最后的块效应得分（*Blocking*）。

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

最后，使用下列方程式：

$$\begin{aligned} \text{IF}(\text{BLOCKING} > 12 \text{ and } 25 \leq \text{EPSNR} < 30) & \quad \text{adjust_EPSNR_blk1}=3 \\ \text{IF}(\text{BLOCKING} > 5 \text{ and } 30 \leq \text{EPSNR} < 35) & \quad \text{adjust_EPSNR_blk1}=5 \end{aligned}$$

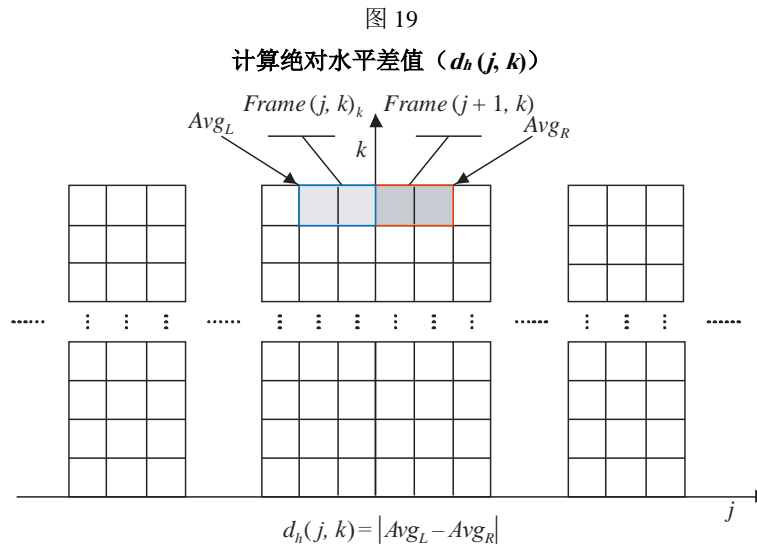
2) 块效应度量II

假设块效应损伤每隔8列现出一次（例如在MPEG2中），也可使用第2个块效应度量。为了计算第2个块效应度量，首先要计算绝对水平差值如下（图19）：

$$d_h(j, k) = |Avg_L - Avg_R|$$

其中：

$$Avg_L = \frac{1}{2} \sum_{p=-1}^0 Frame(j+p, k), \quad Avg_R = \frac{1}{2} \sum_{p=1}^2 Frame(j+p, k)$$



BT.1908-19

然后， j 位置上的水平块效应 (SB_h) 之和定义如下：

$$SB_h[j] = \left(\sum_{1 \leq k \leq height} (|Frame(j, k) - Frame(j+1, k)| \times u(d_h(j, k) - \Phi(Avg_L))) \right)^2$$

其中 $u(\cdot)$ 表示单位阶跃函数，且：

$$\Phi(s) = \begin{cases} 17(1 - \sqrt{s/127}) + 3 & \text{当 } s \leq 127 \text{ 时} \\ 3(s - 127)/128 + 3 & \text{其它} \end{cases}$$

在对全部帧重复这个过程之后，帧水平块效应 (FB_h) 计算如下：

$$FB_h = \left(\sum_{\substack{1 \leq j \leq width \\ j \equiv 0 \pmod{8}}} SB_h(j) \right)^{1/2}$$

对于每一帧，除每一个第8列以外的列差值计算如下：

$$NFB_h = \frac{1}{7} \sum_{l=1}^7 \left(\sum_{\substack{1 \leq j \leq \text{width} \\ j=l(\text{mod}8)}} \left(\sum_{1 \leq k \leq \text{height}} (|Frame(j,k) - Frame(j+1,k)| \times u(d_h(j,k) - \Phi(Avg_L))) \right) \right)^2 \Bigg)^{1/2}$$

然后，最后的水平块效应特性 BLK_H 计算如下：

$$BLK_H = \ln(FB_h / NFB_h)$$

类似地，计算垂直块效应特性 BLK_V ，对于隔行的视频序列，计算场序列中的垂直块效应特性，第 i 帧块效应得分计算如下：

$$FrameBLK(i) = 0.5 \times BLK_H + 0.5 \times BLK_V$$

对帧块效应得分中10%的高分取平均来计算最后的块效应得分（ $BLOCKING2$ ）。

最后，使用下列方程式：

IF(BLOCKING2 > 1.5 and 25 ≤ EPSNR<30)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1.3 and 30 ≤ EPSNR<35)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1.5 and 35 ≤ EPSNR<40)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 1 and 40 ≤ EPSNR<45)	adjust_EPSNR_blk2=2
IF(BLOCKING2 > 0.5 and 45 ≤ EPSNR<55)	adjust_EPSNR_blk2=2

从上面的方程式可以看出，这个调整对于最后EPSNR数值的影响很小，如果块效应人为现象不是每隔8列就出现，人们可以跳过这个调整，或者首先寻找块效应的位置。同样地，人们可以使用 $\Phi(s)$ 的不同函数。

3) 最长的定格帧和总的定格帧

传输错误可能引起长时间的定格帧，针对长时间的定格帧，使用下列方程式：

IF(MAX_FREEZE ≥ 8 and 25 ≤ EPSNR<30)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 6 and 30 ≤ EPSNR<35)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 3 and 35 ≤ EPSNR<40)	adjust_EPSNR_max_freeze=3
IF(MAX_FREEZE ≥ 1.5 and 40 ≤ EPSNR<45)	adjust_EPSNR_max_freeze=2
IF(MAX_FREEZE ≥ 1 and 45 ≤ EPSNR<95)	adjust_EPSNR_max_freeze=2

其中MAX_FREEZE是定格帧的最长持续时间，注意到如果视频序列不是10秒钟，应使用不同的门限。

同样地，总的定格帧计算如下：

IF(TOTAL_FREEZE ≥ 80 and 25 ≤ EPSNR<30)	adjust_EPSNR_total_freeze=3
IF(TOTAL_FREEZE ≥ 40 and 30 ≤ EPSNR<35)	adjust_EPSNR_total_freeze=4
IF(TOTAL_FREEZE ≥ 10 and 35 ≤ EPSNR<40)	adjust_EPSNR_total_freeze=3.5
IF(TOTAL_FREEZE ≥ 2 and EPSNR ≥ 40)	adjust_EPSNR_total_freeze=1.5

其中TOTAL_FREEZE是定格帧的总的持续时间，注意到如果视频序列不是10秒钟，应使用不同的门限。

4) 传输错误块

由于传输错误，可能会出现局部的定格块，同样地，在静态场景中，一些块与以前帧在相同位置上的块相同，考虑到由传输错误引起的局部定格块，将包含发送的边缘像素的块分为相同的块（即块与以前帧中的块相同）或者不同的块。然后，对于相同的块和不同的块计算两个EPSNR。如果两个EPSNR之间的差异（EPSNR_diff）较大，表明可能会出现传输错误。基于这一观察，调整EPSNR如下：

IF($8 \leq \text{EPSNR_diff} \leq 30$ and $25 \leq \text{EPSNR} < 30$)	adjust_EPSNR_diff= 3
IF($9 \leq \text{EPSNR_diff} \leq 30$ and $30 \leq \text{EPSNR} < 35$)	adjust_EPSNR_diff= 4
IF($10 \leq \text{EPSNR_diff} \leq 30$ and $35 \leq \text{EPSNR} < 40$)	adjust_EPSNR_diff= 6
IF($9 \leq \text{EPSNR_diff} < 10$ and $35 \leq \text{EPSNR} < 40$)	adjust_EPSNR_diff= 2
IF($9 \leq \text{EPSNR_diff} \leq 30$ and $40 \leq \text{EPSNR} < 45$)	adjust_EPSNR_diff= 4

然而，如果相同块的总数小于100，不作调整。

5) EPSNR的最后调整

最后，EPSNR数值调整如下：

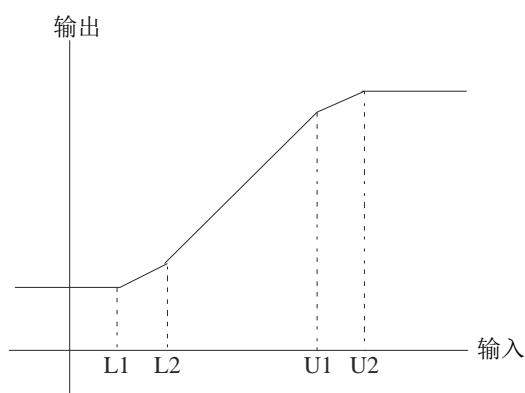
$$\text{EPSNR} \leftarrow \text{EPSNR} - \text{MAX}(\text{adjust_EPSNR_blk1}, \text{adjust_EPSNR_blk2}, \text{adjust_EPSNR_max_freeze}, \text{adjust_EPSNR_total_freeze}, \text{adjust_EPSNR_diff})$$

6) 分段线性拟合

当EPSNR超过某一数值时，感知质量不再提高，在这种情况下，可能要设立EPSNR的上界，此外，当希望获得EPSNR和DMOS（差异平均评估得分）之间的线性关系时，人们可以采用图20所示的分段线性拟合。在VQEG HDTV测试被测试的模型中，上界设为50，下界设为19。

图20

适用于EPSNR和DMOS之间线性关系的分段线性函数



附录1 (资料性的)

视频质量专家组 (VQEG) 的成果

一个被称作视频质量专家组 (VQEG) 的非正式团体在进行感知视频质量测量的研究工作，该团体向ITU-T 第9、12研究组和ITU-R第6研究组报告情况。VQEG最近完成的高清晰度电视阶段I测试评估了提议的完整参考信号感知视频质量测量算法的性能。

下列统计数字来自最终的VQEG HDTV报告 (VQEG报告)，注意到VQEG HDTV报告的主体包含了其他的度量，包括：以独立实验为基础计算得到的皮尔森相关性和RMSE，置信区间，基于独立实验的统计显著性测试，分析包含特定损伤（例如只有ITU-T H.264编码）在内的数据子集，散布图，以及适合的系数。

主要分析

RR模型的性能总结见表4。按照ITU-T J.340建议书计算PSNR，并将PSNR纳入到用于对比目的的这个分析中，“超集RMSE”表示以汇总超集（即映射到同一个范围内的全部6次实验）为基础计算得到的基本度量（RMSE），“最佳表现组的总数”表示其模型是最佳表现的模型或者在统计上相当于最佳表现模型的实验的数量（0到6），“优于PSNR的总数”表示其模型在统计上优于PSNR的实验的数量（0到6），“优于超集PSNR”显示各个模型在统计上是否优于汇总超集上的PSNR，“超集相关性”表示以汇总超集为基础计算得到的皮尔森相关性。

表4

度量	PSNR	Yonsei56k	Yonsei128k	Yonsei256k
超集RMSE	0.71	0.73	0.73	0.73
最佳表现组总数	6	4	4	4
相当于或者优于PSNR的总数	6	4	4	4
相当于超集的PSNR	是	是	是	是
超集相关性	0.78	0.77	0.77	0.77

由于对于三种带宽，模型的性能在统计上相当，建议使用侧信道带宽至少为56kbit/s的模型。

辅助分析

表5列举了对于超集的分部，RR模型的RMSE，这些分部按编码类型（ITU-T H.264或者MPEG-2）以及出现传输错误（错误）或者HRC是否只包含编码人为现象（编码）来划分数据。由于实验不能获得这些均匀跨越整个质量范围的变量，只给出了这些分部的RMSE。

表 5

HRC类型	PSNR	Yonsei56k	Yonsei128k	Yonsei256k
H.264编码	0.75	0.65	0.65	0.65
H.264 错误	0.67	0.86	0.85	0.86
MPEG-2 编码	0.78	0.81	0.81	0.80
MPEG-2 错误	0.66	0.68	0.68	0.68
编码	0.75	0.69	0.69	0.69
错误	0.67	0.79	0.78	0.79

参考资料

ITU-T P.910建议书(2008), 《多媒体应用的主观视频质量评估方法》。

ITU-T P.911建议书 (1998), 《多媒体应用的主观视听质量评估方法》。

ITU-T J.143建议书(2000), 《数字有线电视中用于客体感知视频质量衡量的用户要求》。

ITU-R BT.500建议书, 《主观评估电视画面质量的方法》。

ITU-T J.340建议书(2010), 《补偿常数空间偏移、常数临时偏移和常数亮度增益和偏移的已处理视频序列的计算峰值信噪比的基准算法》。
