

文章编号: 1004-924X(2016)03-0659-09

## 结合人眼对比度敏感视觉特性的视频质量客观评价

姚军财<sup>1,2</sup>, 刘贵忠<sup>1\*</sup>

(1. 西安交通大学 电子与信息工程学院, 陕西 西安 710049;

2. 陕西理工学院 物理与电信工程学院, 陕西 汉中 723000)

**摘要:** 结合人眼对亮度、色度、对比度以及运动目标的感知特性, 提出了一种基于人眼对视频内容感知的视频质量客观评价方法。该方法将视频分为空域和时域信息分别描述, 并利用人眼感知特性, 从视频的亮度、色度、对比度以及目标运动4个方面提取特征, 计算其强度。然后以人眼对比度敏感值作为强度的权重因子求和, 构建人眼感知视频内容模型。最后, 分别以此模型模拟人眼感知源视频和失真后的视频, 计算每对应单元的所有像素之间和运动矢量之间的强度差; 以强度差作为视频质量评价的分数, 构建视频质量客观评价模型。采用LIVE数据库中的6个源视频和48个测试视频进行了质量评价实验, 并与视频质量专家组(VQEG)推荐的5个较好的视频质量客观评价模型进行了对比分析。结果表明: 提出模型的视频质量评价结果与主观评价结果之间的线性相关性系数达到0.8705, 显示了较好的一致性, 评价效果优于5个典型的模型。

**关键词:** 人眼视觉特性; 视频质量评价; 亮度; 色度; 对比度; 相关系数

**中图分类号:** TN919.81 ;R339.14      **文献标识码:** A      **doi:** 10.3788/OPE.20162403.0659

## Video quality objective assessment combined contrast sensitivity characteristics of human visual system

YAO Jun-cai<sup>1,2</sup>, LIU Gui-zhong<sup>1\*</sup>

(1. School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, 710049, China;

2. School of Physics and Telecommunication Engineering, Shaanxi University of Technology, Hanzhong 723000, China)

\* Corresponding Author, E-mail: Liugz@xjtu.edu.cn

**Abstract:** In combination of the perceiving characteristics of human eyes for brightness, chroma, contrast and moving targets, an objective assessment method of video quality based on contrast sensitivity characteristics of a human visual system was proposed. In the method, the video was divided into spatial and time domains to be described. The features of image were extracted from four aspects, brightness, chroma, contrast, and target motion based on the perceiving characteristics of human eyes and their intensities were computed. Then, the contrast sensitivity values of human eyes were used as the weight factors of the intensity to sum and to construct the model of human eye perception content of the video. Finally, original and distorted videos respectively perceived by imitating eyes with this model, and the intensity differences of the pixels and the motion vectors between arbitrary corresponding units of two videos were computed. By taking the intensity differences as the scores of video quality objective evaluation, the objective evaluation model for video quality was constructed by them. The experiments were carried out with 6 source videos and 48 test videos proposed by LIVE database, and the 5 classic

收稿日期: 2015-12-08; 修订日期: 2016-01-14

基金项目: 国家自然科学基金资助项目(No. 61301237);陕西省科技新星计划资助项目(No. 2015KJXX-42)

video quality evaluation models recommended by the Video Quality Expert Group (VQEG) were compared with the proposed model. The results show that the linear correlation coefficient between video quality evaluated by the proposed model and the subjective evaluation results reaches 0.870 5. They have good consistency, and evaluation effects are better than those of other 5 classical models.

**Key words:** human visual system characteristics; video quality assessment; brightness; chroma; contrast; correlation coefficients

## 1 引言

多媒体、信息技术和光学技术的快速发展使得视频应用得到了前所未有的发展,但是,由于网络带宽的限制,在视频传输过程中必须对其进行压缩处理,同时由于网络传输的不可靠性,不可避免地会产生各种失真,这在很大程度上降低了视频的质量,影响了消费者的观看效果。所以目前急需一种有效且方便的视频质量客观评价方法,用来指导和控制网络传输,以便获得更好的视频质量<sup>[1-3]</sup>。目前,虽然视频的相关技术取得了较多的成果,但视频质量评价方法还远远不能满足通信服务的要求,这主要是因为视频的客观质量评价结果与主观感知结果存在一定的差距,则研究一种尽可能逼近人眼主观感知的视频质量客观评价方法尤为重要<sup>[4-6]</sup>。

目前国内外相关学者对视频质量客观评价方法做了大量的研究工作,提出了较多的评价模型,诸如峰值信噪比(Peak Signal to Noise Ratio, PSNR)、结构相似性(Structural Similarity Index Measurement System, SSIM)、视觉信噪比(Visual Singal-to-Noise Ratio, VSNR)、基于运动的视频质量评价(Motion-based Video Integrity Evaluation, MOVIE)和基于感知的质量评价(Perception-based Image Metric, PIM)等模型,同时也成立了视频质量评价专家组(Video Quality Experts Group, VQEG)机构<sup>[1-13]</sup>。这些方法根据设计原理可分为:①基于像素域的方法。该方法是对原始视频和失真视频按帧进行帧内像素间的对比分析,如PSNR、MSE((Mean Square Error)等;此方法目前比较成熟,但评价结果与人眼视觉效果的差异较大<sup>[1-2, 7-8]</sup>。②基于视觉感知的方法。此方法是利用视觉模型来模拟人眼视觉感知,从而构建评价模型,以评价视频质量。此类方法与主观感知具有较好的一致性,但对视频中的运动信息的评价结果有待改进,且计算量较大<sup>[4,9-11]</sup>。③工程类方

法。该方法通过提取和分析视频中引起失真的特征因素,根据权重构建评价模型。此方法计算相对简单,效果较为明显,但是此方法针对性较强,而且为了提高评价准确度,需要提取更多的有效特征,使得计算比较复杂,计算量较大。最为关键的是其没有考虑人眼感知特性,使得主观和客观评价结果存在一定程度上的不一致<sup>[12-15]</sup>。

目前的评价方法主要是针对压缩和传输等外在因素而进行的评价,很少考虑视频内容本身的感知<sup>[7, 14]</sup>。基于以上的不足,本文结合人眼对亮度、色度、对比度以及运动目标的感知特性,提出了一种基于人眼对视频内容感知的视频质量客观评价方法。采用LIVE数据库<sup>[16]</sup>中的6个源视频和48个测试视频进行了质量评价实验,且与VQEG推荐的5个较好的视频质量客观评价模型进行了对比分析,以验证算法的先进性。

## 2 基于人眼感知的视频时域和空域描述方法

视频可以看作是图像逐帧播放,帧与帧之间的时域信息可体现在像素点的运动矢量上,每一点的运动矢量即代表了该点的运动轨迹,同时也反映了该点的运动特征。则第 $t$ 帧到 $t+1$ 帧的视频播放过程中,人眼看到的是第 $t$ 帧的图像和其图像上的像素点按照运动矢量轨迹在运动。根据此描述,则可以把人眼观察的视频信息看作是静止图像和运动的像素点的组合,其描述示意图如图1所示(彩图见期刊电子版)。将所有帧的静止图像看作是视频的空域信息,记作 $\{ \text{Inform}_{\text{spatial}}(i) \}$ ;所有帧图像对应的像素点的运动看作是时域信息,记作为 $\{ \text{Inform}_{\text{temporal}}(i) \}$ ;视频的总帧数设为 $N$ ,以任一帧图像和其相邻的时域信息作为一个单元,则整个视频可以看作 $N$ 个视频单元,单元的示意图如图2所示(彩图见期刊电子版)。在每个单元中,帧图像上所有像素点的运动时间为 $1/F$ ( $F$ :帧率,单位为frame/s),每一像素点在该时间内的运动距离记为运动矢量的大小。据此,随着视频的播放,可以

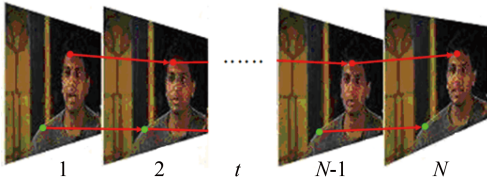


图1 视频的时域和空域描述示意图

Fig. 1 Schematic diagram for describing time and spatial domains of video

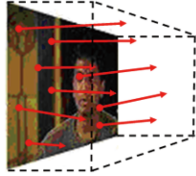


图2 视频单元示意图

Fig. 2 Schematic diagram of video unit

把人眼对视频的感知理解为人眼逐个感知视频单元的过程。人眼对于空域图像的视觉感知, 最敏感的是每一像素点的亮度、色度以及其所所在位置与周围环境形成的对比; 人眼对于时域的视觉感知是像素点从 $t$ 时刻运动到 $t+1$ 时刻的运动, 其运动特征可通过运动矢量值来描述。而且对于不同位置和不同的运动矢量, 人眼对像素点的运动感知效果不同, 则时域的感知需要考虑空间频率和时间频率。因此可以把视频内容表述为4个方面的特征值, 即亮度、色度、对比度以及目标运动。

综上所述, 将视频分为时域和空域信息, 则在客观评价视频质量时, 可以分别从空域和时域两个方面单独评价, 然后再通过权重求和得到最终的评价结果。

### 3 人眼视觉特性模型及其在视频处理中的应用

#### 3.1 视觉模型

人眼主要是依据目标的亮度、色度及其对比度

$$MCSF_{st}(f_{\theta}, f_t) = 4\pi^2 f_{\theta} \cdot f_t e^{-4\pi(2f_{\theta}+f_t)/45.9} \times \left(6.1 + 7.3 \left| \lg \frac{f_t}{3f_{\theta}} \right|^3 \right), \quad (4)$$

$$f_t = v \cdot f_{\theta}, \quad (5)$$

式中:  $MCSF_{st}$ 表示人眼对运动目标的对比度敏感阈值,  $f_{\theta}$ 为角频率,  $f_t$ 为时间频率, 其用于描述人眼视角范围内光栅完全通过的次数,  $v$ 为光栅的运动速度<sup>[19]</sup>。根据上面关于视频的时域和空域描述方法可知, 像素点从第 $t$ 帧运动到第 $t+1$ 帧的时间为

来感知对象的存在。人眼对比度敏感视觉特性模型一般被称为人眼对比度敏感函数(Contrast Sensitivity Function, CSF), 目前主要采用不同条件下的人眼对比度敏感值(人眼对比度觉察阈值 $C_T$ 的倒数)与角频率( $f_{\theta}$ )之间的函数关系来加以描述。国内外对CSF做了大量的研究, 得出了较多的CSF模型。综合考虑视频质量评价的实际要求和计算复杂度, 以及CSF模型所反映的感知特性的效果和色度空间特性, 本文选用由Barten和Nadenau提出的人眼觉察静止目标时的亮度和色度对比度敏感函数( $CSF_L$ 、 $CSF_{rg}$ 和 $CSF_{by}$ ), 以及由Kelly提出的人眼觉察运动目标的对比度敏感函数( $MCSF_{st}$ )<sup>[17-19]</sup>作为评价标准。3类CSF的表达式如下所示。

(1)亮度对比度敏感函数

$$CSF_L(f_{\theta}) = \frac{1}{C_T(f_{\theta})} = a \cdot f_{\theta} \cdot \exp(-b \cdot f_{\theta}) [1 + c \cdot \exp(b \cdot f_{\theta})]^{\frac{1}{2}}, \quad (1)$$

$$\text{式中: } a = \frac{540 \cdot (1 + 0.7/L)^{-0.2}}{1 + \frac{12}{w \cdot (1+f_{\theta}/3)^2}},$$

$$b = 0.3 \cdot (1 + 100/L)^{0.15}, \quad c = 0.66.$$

式中:  $CSF_L$ 表示人眼对亮度的对比度敏感阈值,  $C_T$ 表示人眼对亮度的对比度觉察阈值,  $f_{\theta}$ 为视觉上的空间频率, 称之为角频率, 单位为cycle/degree,  $L$ 表示目标的平均亮度, 表示空间视角<sup>[17]</sup>。

(2)对立色色度对比度敏感函数

$$CSF_{rg}(f_{\theta}) = a \cdot \exp^{b(f_{\theta})^c} \quad a=1, \quad b=-0.152, \quad c=0.893, \quad (2)$$

$$CSF_{by}(f_{\theta}) = a \cdot \exp^{b(f_{\theta})^c} \quad a=1, \quad b=-0.2041, \quad c=0.9, \quad (3)$$

式中:  $CSF_{rg}$ 和 $CSF_{by}$ 分别表示人眼对红绿和蓝黄对立色目标的色度对比度敏感阈值,  $f_{\theta}$ 为角频率<sup>[18]</sup>。

(3)人眼感知运动目标的对比度敏感函数

$T_s=1/F$ , 其与视觉测量中一个光栅通过人眼视角范围内的时间是一致的, 也等于该点在相邻两帧图像间的运动矢量( $M_v$ )上的运动时间。式(5)应用到视频质量评价中可变换为式(6),  $f_t$ 即可表示运动矢量时间频率。

$$f_t = \frac{M_v}{T_s} \cdot f_{\theta} = M_v \cdot F \cdot f_{\theta}. \quad (6)$$

### 3.2 视觉模型在视频处理中的应用

在结合人眼视觉模型的图像和视频处理技术中,如何计算人眼观察视频或图像时的 $f_\theta$ 是最为关键的一步,其计算方法如下。

图像经离散余弦变换(Discrete Cosine Transform, DCT)再逆变换到空域时,表现为以下特点:(1)若频谱图中水平( $x$ 方向)或竖直( $y$ 方向)上任意两点经IDCT变换后的空域图是竖直或水平的条纹光栅,且光栅覆盖整个频谱图。(2)若频谱图上任意一点的频谱坐标为 $(f_x, f_y)$ ,则IDCT变换后的空域图上光栅条纹的周期数目为 $\sqrt{f_x^2 + f_y^2}$ ,其中: $f_x$ 和 $f_y$ 在图像处理中被称为空间频率。上述特点表明:空域图像本身可以看成是 $m$ ( $m$ 为图像的像素个数)个大小相同、空间频率不同的光栅的合成图,或者说对于图像上任何一点,其本身就代表一个大小与原图像一样、周期数目为 $\sqrt{f_x^2 + f_y^2}$ 的光栅,则人眼觉察图像上任何一点即是觉察该点对应的光栅。

在视觉研究领域,将角频率 $f_\theta$ 定义为单位视角内刺激人眼的光栅条纹的周期数目。记人眼视角范围内的光栅的周期数目为 $N$ ,空间视角为 $\theta$ ,则 $f_\theta$ 的计算如式(7)所示:

$$f_\theta = \frac{N}{\theta}. \quad (7)$$

结合以上图像DCT变换的特征可知人眼观察视角为整幅图像对应的视角,图像不变,则视角不变。结合式(7)可得,人眼觉察图像上每一点的角频率计算式如式(8)所示:

$$f_\theta = \frac{\sqrt{f_x^2 + f_y^2}}{\theta}. \quad (8)$$

根据以上关系可知,当人眼观察视频时,视频中任一点的角频率和时间频率均可以通过式(6)和(8)计算获得,将他们代入到视觉模型(1)、(2)、(3)和(4)中,便可以计算出人眼觉察视频中任何一像素点的敏感阈值,包括亮度、色度和运动敏感值。

## 4 视频质量客观评价方法

人眼在感知视频时,最为强烈的刺激是视频的亮度、色度、对比度以及运动目标信息,因此可以把它们作为人眼感知视频的要素,并根据人眼对每个要素的敏感程度对各要素分配不同权重,再进行求和并作为人眼感知视频的内容,以此构

建人眼感知视频模型。分别计算源视频和失真视频,并计算其强度差,从而获得质量评价结果。基于人眼对视频内容的感知方法对视频质量进行评价的步骤如下:首先将视频分为空域和时域(空域即为帧图像,时域采用运动矢量表示);再从亮度、色度、对比度以及目标运动4个方面把分离的视频利用人眼感知特性模型进行特征提取,并计算其强度;接着再结合CSF计算人眼感知空域和时域信息的敏感值,并以此敏感值作为强度的权重因子求和,以此求和数据作为人眼感知的视频内容,并构建模型;最后以此模型分别感知源视频和失真后的视频,计算对应视频单元的所有像素之间的差异和运动矢量之间的差异(即强度差),以差异的大小作为视频质量的客观评价分数,从而构建视频质量客观评价模型。方法流程描述如图3所示。具体过程可分为3个部分:

### (1) 视频时域信息感知模型

结合人眼对运动目标的视觉敏感特性及其模型,构建人眼对时域信息的感知模型,时域信息主要采用运动矢量来描述。

① 将视频帧图像变为灰度图。采用全搜索算法计算每一帧每一像素点的运动矢量 $\mathbf{M}_V$ 。

② 将图像划分成 $8 \text{ pixel} \times 8 \text{ pixel}$ 大小的子块,并对其进行DCT变换,将零频分量移到频谱中心,形成对称。计算变换后的频谱图中每一子块每一点的频谱坐标 $f = \sqrt{f_x^2 + f_y^2}$ ,即空间频率。则 $f$ 为IDCT变换空域上覆盖的光栅条纹的周期数目 $N$ 。

③ 以 $8 \text{ pixel} \times 8 \text{ pixel}$ 大小的子块作为人眼每次关注的目标来计算视角,并代入式(8)计算子块上每一点的角频率 $f_\theta$ 。对于一般显示的图像,角频率不发生变化。

④ 由于每次关注的均是 $8 \text{ pixel} \times 8 \text{ pixel}$ 大小的子块,则结合式(6)可以计算出运动矢量的时间频率 $f_t$ , $f_t = \mathbf{M}_V \times F \times f_\theta / 8$ 。

⑤ 把角频率 $f_\theta$ 和时间频率 $f_t$ 代入到人眼对运动目标的视觉敏感函数模型(式4)中,计算人眼对运动矢量的敏感程度(即敏感值 $\text{MCSF}_{st}$ )。

⑥ 将人眼对运动矢量的敏感值 $\text{MCSF}_{st}$ 作为权重因子乘以运动矢量 $\mathbf{M}_V$ ,得到人眼感知运动矢量值,即为主观感知的视频时域信息( $I_{\text{temporal}}$ ),其表达式如式(9)。

$$I_{\text{temporal}} = \mathbf{M}_V \times \text{MCSF}_{st}. \quad (9)$$

### (2) 视频空域信息感知模型

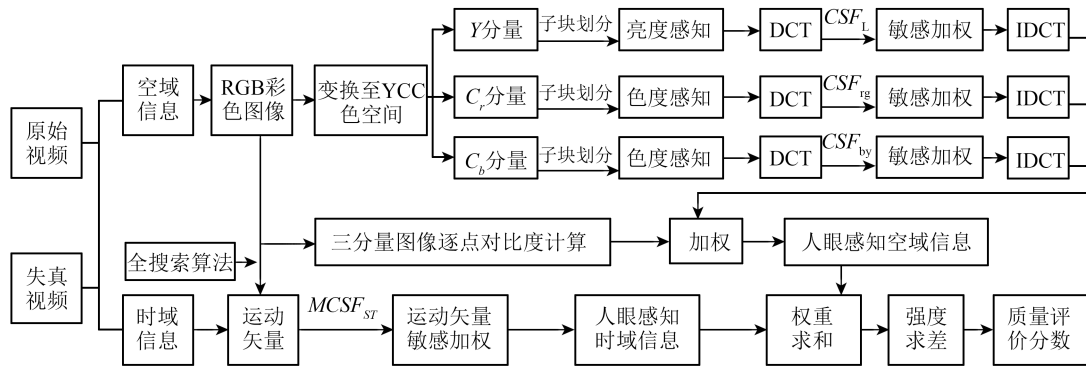


图3 基于人眼感知特性的视频质量客观评价方法示意图

Fig. 3 Schematic of objective evaluation method for video quality based on human perception characteristics

视频空域信息感知模型的构建思想为: 结合人眼对视频帧图像的亮度、色度的感知特性和数学模型, 计算人眼对其强度的感知值 $S$ ; 再利用CSF计算感知亮度和色度强度的敏感值; 最后计算每一点与其周围8个点的平均对比度 $C_{mean}$ 。以敏感值和对比度值分别作为感知亮度和色度强度的权重因子再求和, 构建人眼感知图像内容的模型, 以求和数据信息作为人眼感知视频帧图像的内容。具体方法为:

① 利用RGB和 $YCbCr$ 颜色空间的经典变换关系, 将视频帧图像转换到 $YCbCr$ 色空间, 进行亮度和色度分离; 并把3个分量图分别划分成 $8\text{ pixel} \times 8\text{ pixel}$ 大小的子块。结合人眼对强度感知的公式(10)分别计算3个分量图各子块上每一点的亮度主观感知强度值 $S$ , 分别记为 $S_L$ 、 $S_{rg}$ 和 $S_{by}$ 。

$$S = K \log I. \quad (10)$$

② 根据计算出的角频率 $f_\theta$ , 利用人眼对比度敏感函数表达式(1)、(2)和(3), 依次计算出每一点处人眼的敏感程度值 $CSF_L$ 和 $CSF_{rg}$ 、 $CSF_{by}$ 。

$$I_{\text{spatial}} = S \times CSF \times C_{\text{mean}} = \begin{bmatrix} R_{\text{subject}} \\ G_{\text{subject}} \\ B_{\text{subject}} \end{bmatrix} = \begin{bmatrix} IDCT[DCT(S_L) \times CSF_L] \\ IDCT[DCT(S_{rg}) \times CSF_{rg}] \\ IDCT[DCT(S_{by}) \times CSF_{by}] \end{bmatrix} \cdot [C_{\text{mean\_R}} \ C_{\text{mean\_G}} \ C_{\text{mean\_B}}] \quad (11)$$

(3) 基于人眼对比度感知特性的视频质量客观评价模型

对人眼感知的视频空域和时域信息按照重要性进行加权, 从而构建人眼感知视频模型, 表达式如(12)所示, 记其为LCVC(Video Perception Model based on Luminance, Contrast, Motion Vector and Sensitivity Characteristics)模型。式中,  $K$ 为加权因子, 通过多次实验得知,  $K$ 取0.9时效果较好。

③ 对主观感知强度 $S_L$ 、 $S_{rg}$ 和 $S_{by}$ 分别进行DCT变换, 再分别乘以 $CSF_L$ 和 $CSF_{rg}$ 、 $CSF_{by}$ , 再做逆离散余弦变换(Inverse Discrete Cosine Transformation, IDCT), 即得到结合人眼敏感特性的主观感知亮度和色度值; 按照 $YCbCr$ 和RGB的经典转换( $T_{YCC-RGB}$ )公式将 $YCbCr$ 转换到RGB空间显示, 得到R、G、B主观感知分量图。

④ 在人眼觉察图像时, 图像中的每一个点都会受到周围环境的影响。对于原图像分别以R、G、B三分量图上的每一点为中心, 计算该点与周围最邻近的8个点之间的对比度值, 并求其平均值 $C_{mean}$ 。

⑤ 将 $C_{mean}$ 与主观感知R、G、B分量图中的各点相应幅值相乘, 并按照颜色混合原则中R、G、B的比例关系得到主观感知视频帧图像( $RGB_{\text{subject}}$ ), 即人眼主观感知视频空域信息( $I_{\text{spatial}}$ )。计算表达式如式(11), 其即为视频空域信息的人眼感知模型。

利用LCVC模型计算原始视频和失真后的视频, 得出两个视频的人眼感知信息强度 $LCVC_{\text{original}}$ 和 $LCVC_{\text{distortion}}$ 。再将帧内对应像素点进行幅度相减, 得出强度差。再结合PSNR的定义式, 用强度差代替PSNR中的MSE值进行计算, 计算结果即为视频质量客观评价分数, 将其记为 $MOS_{LCVC}$ (Mean Option Score, MOS)。其值越大, 代表质量越好, 反之测质量越差则越小。基于人

眼感知特性的视频质量客观评价模型如式(13)所示, 记为VQOE\_BHVP(Model of Video Quality of Experience Based on the Human Visual Perception)模型。

$$LCVC = I = K \cdot I_{\text{spatial}} + (1 - K) \cdot I_{\text{temporal}} \quad (12)$$

$$MOS_{LCVC} = 10 \lg \left( \frac{255 \times 255}{LCVC_{\text{original}} - LCVC_{\text{distortion}}} \right) \quad (13)$$

### 5 实验结果和讨论

#### 5.1 实验结果

按照上述的视频质量客观评价方法, 选取LIVE数据库中提供的6个视频作为源视频(如图4)<sup>[16]</sup>, 这6个视频经过H.264和MPEG2压缩后的视频作为测试视频, 通过Matlab编程进行了实验测试。

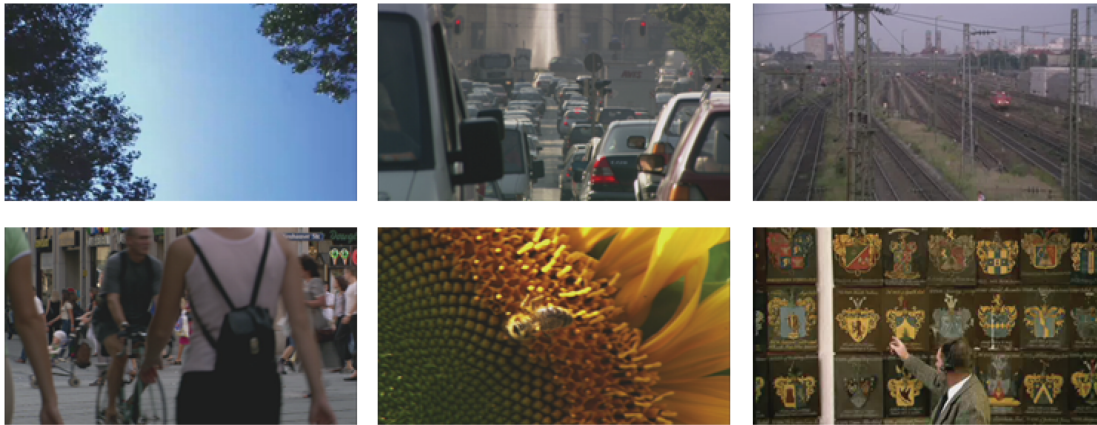


图 4 LIVE数据库中提供的6个源视频

Fig. 4 Six source videos provided by LIVE database

结果如图5所示。图5是利用所提出的质量评价模型计算出的评价结果 $MOS_{LCVC}$ 与LIVE数据库提供的主观质量评价结果MOS之间的散点图。6个源视频是未失真视频, 对每个源视频均采用H.264和MPEG2两种编码方法按4种压缩比进行压缩, 每种方法各获得24个测试视频, 共48个测试视频, 它们均由LIVE数据库提供; 并且, 该数据库还给出了各测试视频的主观质量评价分数值MOS<sup>[16]</sup>。

#### 5.2 讨论

在对视频质量客观评价的研究中, 为了说明客观评价结果的准确性, 一般将其与主观评价分数值MOS进行对比, 并计算它们之间的相关性。相关性参数主要有: 线性相关系数(Linear Correlation Coefficient, LCC)、Spearman 等级相关系数(Spearman Rank Order Correlation Coefficient,

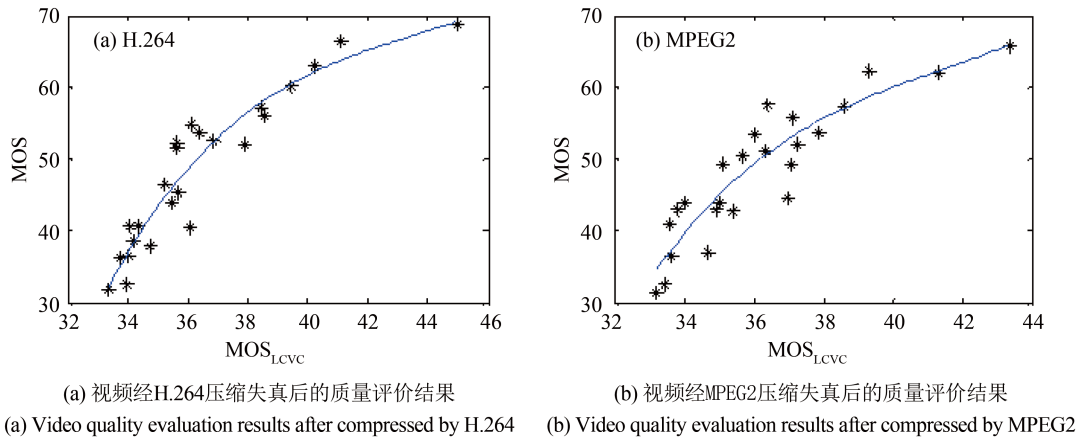


图 5 本文评价模型得到的分数 $MOS_{LCVC}$ 与LIVE数据库中提供的主观评价分数MOS值之间的散点图

Fig. 5 Scatter plots between subjective evaluation scores of MOS proposed by LIVE database and objective evaluation scores of  $MOS_{LCVC}$  obtained by proposed model

SROCC)、均方根误差(Root Mean Squared Error, RMSE)和散点图的离出率(Outlier Ratio, OR)<sup>[20]</sup>。这4个参数值能够较好地衡量客观评价结果与主观评价结果的吻合程度。为了说明提出的VQOE\_BHVP视频质量客观评价模型的优劣, 将其评价结果与LIVE数据库提供的主观评价结果MOS进行对比, 计算结果如表1所示。

表1 用于衡量视频主观评价结果与客观评价结果一致性的4个参数值

Tab. 1 Four parameters to assess consistency of subjective and objective evaluation results of video

压缩方法和参数	LCC	SROCC	RMSE	OR
H.264	0.881 6	0.871 9	7.839 6	0.228 7
MPEG2	0.859 4	0.848 7	7.952 8	0.257 3
平均	0.870 5	0.860 3	7.896 2	0.243 0

从衡量视频主观评价结果与客观评价结果吻合程度的4个参数上看, 平均线性相关系数和非线性相关系数均达到0.86以上, 而散点图的平均离出率小于0.25, 这表明本文提出的视频质量客观方法的评价结果与主观评价结果具有较好的一致性。

国内在视频质量客观评价方面做了大量的研究工作, 提出了较多的评价模型, 目前VQEG推荐的模型主要有PSNR、SSIM、VSNR、MOVIE和PIM等, 其中PSNR和SSIM是目前研究者常常使用的两种模型<sup>[20]</sup>。为了说明本文模型的优越性, 对比分析了本文评价模型和以上5种评价模型的评价分数与主观评价分数MOS之间的相关性, 并计算了衡量参数LCC、SROCC、RMSE和OR值<sup>[20]</sup>。结果如表1所示。

由表1可知: 从用于衡量客观评价结果与主观评价结果吻合程度的4个参数上看, 本文模型的评价效果优于用于对比的其他5个模型。

在视频质量评价模型的研究中, 一般采用算法运行时间来描述其模型的计算复杂性(Complexity)。Latency参数一般定义为在模型评价过程中采用了多少帧图像对视频进行评价<sup>[20]</sup>。为了更好地说明本文评价模型的优势, 对比分析了本文模型和以上

5种模型的Complexity和Latency参数, 结果如表2所示。从Complexity参数来看, 本文评价模型的算法复杂性稍高, 其只比MOVIE模型简单; 但结合Complexity和Latency两个参数进行分析可以发现, 本文评价模型每次采用10帧图像进行评价, 其评价结果比PSNR、SSIM、VSNR和PIM模型的更准确, 更有具有代表性。

表2 本文提出的视频质量客观评价模型与5种典型评价模型的评价效果对比

Tab. 2 Comparison of evaluation effects between five typical video quality objective evaluation models and proposed model

评价模型	LCC	SROCC	RMSE	OR	Complexity/s	Latency
PSNR	0.782 1	0.783 2	9.878 4	0.671 9	1	1
SSIM	0.793 3	0.765 2	9.783 1	0.681 3	11.1	1
VSNR	0.825 5	0.798 0	8.917 9	0.665 6	18.8	1
MOVIE	0.854 6	0.858 8	8.090 6	0.618 8	134	33
PIM	0.863 1	0.854 4	7.920 6	0.609 4	17	1
VQOE_BHVP	<b>0.870 5</b>	<b>0.860 3</b>	<b>7.896 2</b>	<b>0.243 0</b>	<b>109.43</b>	<b>10</b>

## 6 结论

本文结合人眼对亮度、色度、对比度以及运动目标的感知特性, 提出了一种基于人眼对视频内容感知的视频质量客观评价方法。该方法结合人眼对运动目标的敏感特性及其模型, 提出了一种人眼对视频时域信息的感知模型。结合人眼对视频帧图像的亮度、色度和对比度的敏感特性, 构建了一种人眼感知图像内容的模型。结合人眼对视频时域和空域信息感知的模型, 提出了一种基于人眼对视频内容感知的视频质量客观评价模型。最后, 采用LIVE数据库中的6个源视频和48个测试视频进行了质量评价实验, 且与VQEG推荐的5个较好的视频质量评价模型进行了对比分析。结果表明: 本文提出模型的视频质量评价结果与主观评价结果的线性相关性系数达到0.870 5, 具有较好的一致性, 评价效果优于用于对比的5个典型模型。

## 参考文献:

[1] 范媛媛, 沈湘衡, 桑英军. 基于对比度敏感度的无参考图

像清晰度评价[J]. 光学精密工程, 2011, 19(10): 2485-2493.

FAN Y Y, SHEN X H, SAN Y J. No reference image

- sharpness assessment based on contrast sensitivity[J]. *Opt. Precision Eng.*, 2011, 19(10): 2485-2493. (in Chinese)
- [2] 宁方立, 何碧静, 韦娟. 基于 $l_p$ 范数的压缩感知图像重建算法研究[J]. *物理学报*, 2013, 62(17): 174212.  
NING F L, HE B J, WEI J. An algorithm for image reconstruction based on  $l_p$  norm[J]. *Acta Phys. Sin.*, 2013, 62(17): 174212. (in Chinese)
- [3] CHEN Y J, WU K S, ZHANG Q. From QoS to QoE: A tutorial on video quality assessment [J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(2): 1126-1165.
- [4] KARTHIKEYAN R, SAINARAYANAN G, DEEPA S N. Perceptual video quality assessment in H. 264 video coding standard using objective modeling[J]. *Springerplus*, 2014, 3(1): 1-6.
- [5] 杨亚威, 李俊山, 张士杰, 等. 基于生物视觉标准模型特征的无参考型图像质量评价方法[J]. *液晶与显示*, 2014, 29(6): 1016-1023  
YANG Y W, LI J S, ZHANG S J, et al. Non reference image quality assessment approach based on standard model features of biological vision[J]. *Chinese Journal of Liquid Crystals and Displays*, 2014, 29(6): 1016-1023. (in Chinese)
- [6] 米曾真. 小波域中CSF频率与方向加权的图像质量评价方法[J]. *电子学报*, 2014, 42(7): 1273-1276.  
MI Z ZH. Image quality evaluation method based on frequency and direction weighted to CSF in wavelet domain[J]. *Acta Electronic Sinica*, 2014, 42(7): 1273-1276. (in Chinese)
- [7] LI C F, BOVIK A C. Content-weighted video quality assessment using a three-component image model[J]. *Journal of Electronic Imaging*, 2010, 19(1): 143-153.
- [8] 邱聚能, 李辉, 闫乐乐, 等. 基于图像质量评价的LCD运动模糊检测方法[J]. *液晶与显示*, 2015, 30(3): 531-537  
QIU J N, LI H, YAN L L, et al.. LCD motion blur detecting method based on image quality assessment [J]. *Chinese Journal of Liquid Crystals and Displays*, 2015, 30(3): 531-537. (in Chinese)
- [9] OU Y F, XUE Y Y, WANG Y. Q-STAR\_A perceptual video quality model considering impact of spatial, temporal, and amplitude resolution[J]. *IEEE Transactions on Image Processing*, 2014, 23(6): 2473-2486.
- [10] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [11] AKAMINE A W Y, FARIAS M C. Video quality assessment using visual attention computational models[J]. *Journal of Electronic Imaging*, 2014, 23(6): 061107.
- [12] SOUNDARARAJAN R, BOVIK A C. Video quality assessment by reduced reference spatio-temporal entropic differencing[J]. *IEEE Transactions on Circuits & Systems for Video Technology*, 2013, 23(4): 684-694.
- [13] MOORTHY A K, CHOI L K, BOVIK A C, et al.. Video quality assessment on mobile devices: subjective, behavioral and objective studies[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2012, 6(6): 652-671.
- [14] ANEGEKUH L, SUN L, JAMMEH E, et al.. Content-based video quality prediction for HEVC encoded videos streamed over packet networks[J]. *IEEE Transactions on Multimedia*, 2015, 17(8): 1323-1334.
- [15] GINESU G, MASSIDDA F, GIUSTO D D. A multi-factors approach for image quality assessment based on a human visual system model[J]. *Signal Processing: Image Communication*, 2006, 21(4): 316-333.
- [16] SESHADRINATHAN K, SOUNDARARAJAN R, BOVIK A C, et al.. LIVE Video Quality Database[DB] [http://live.ece.utexas.edu/research/quality/live\\_mobile\\_video.html](http://live.ece.utexas.edu/research/quality/live_mobile_video.html), [2016-04-20].
- [17] BARTEN P G J. Evaluation of subjective image quality with the square-root integral method[J]. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 1990, 7(10): 2024-2031
- [18] NADENAU M. *Integration of human color vision models into high quality image compression* [D]. Switzerland: École Polytechnique Fédérale de Lausanne. 2000.
- [19] KELLY D H. Motion and vision. II. Stabilized spatio-temporal threshold surface [J]. *Journal of the Optical Society of America*, 1979, 69(10): 1340-1349.



[20] ZHANG F, BULL D R. Quality assessment methods for perceptual video compression [C]. *20th IEEE*

*International Conference on Image Processing (ICIP 2013)*, 2013: 39-43.

作者简介:



姚军财 (1979-), 男, 湖北黄冈人, 博士研究生, 副教授, 2007年于云南师范大学获得硕士学位, 主要从事视觉光学和多媒体信号处理与通信方面的研究。E-mail: yjc4782@126.com

导师简介:



刘贵忠 (1962-), 陕西延安人, 教授, 博士生导师, 1982年、1985年于西安交通大学分别获得学士、硕士学位, 1989年于荷兰埃因霍温大学获得博士学位, 主要从事多媒体信号处理与通信方面的研究。E-mail: liugz@xjtu.edu.cn