

双目立体视频最小可辨失真模型及其在质量评价中的应用

张艳^{*①②} 安平^{①③} 张秋闻^① 张兆杨^{①③}

^①(上海大学通信与信息工程学院 上海 200072)

^②(安徽财经大学计算机科学与技术系 蚌埠 230030)

^③(新型显示技术及应用集成教育部重点实验室 上海 200072)

摘要: 该文针对目前最小可辨失真(JND)模型只能应用于评价单视点视频的问题,探索一种以双眼亮度关系为基本出发点的双目立体视频的JND模型。首先根据立体视频特性引入双眼亮度关系模型,然后考虑了背景亮度掩盖、纹理掩盖、帧间掩盖、空间时间对比灵敏度、眼睛运动等因素,提出了双目JND模型,最后将该双目JND模型应用到立体视频质量评价方法中。实验结果表明该文提出的质量评价方法符合人类的视觉感知特性,更接近于主观测试结果。

关键词: 图像处理; 双目立体视频; 最小可辨失真模型; 质量评价; 人类视觉系统

中图分类号: TN919.82

文献标识码: A

文章编号: 1009-5896(2012)03-0698-06

DOI: 10.3724/SP.J.1146.2011.00799

Binocular Stereoscopic Video Just Noticeable Distortion Model and Its Application to Quality Assessment

Zhang Yan^{*①②} An Ping^{①③} Zhang Qiu-wen^① Zhang Zhao-yang^{①③}

^①(School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China)

^②(Department of Computer Science and Technology, Anhui University of Finance & Economics, Bengbu 230030, China)

^③(Key Laboratory of Advanced Displays and System Application, Shanghai University, Shanghai 200072, China)

Abstract: Since current video quality assessment method based on Just Noticeable Distortion (JND) model can only assess monoscopic video, a binocular JND model considering binocular luminance relationship is explored. First, binocular luminance relationship model is used according to stereo video characteristic. Then, many perception factors such as the background luminance masking effect, texture masking effect, intra-frame masking effect, spatiotemporal contrast sensitivity function, eye movement are considered. Finally, binocular JND model is applied to stereo video quality assessment. Experimental results prove that the proposed method can reflect the human visual perception characteristics and has a more consistency with the subjective judgments.

Key words: Image processing; Binocular stereo video; Just Noticeable Distortion (JND); Quality assessment; Human Visual System (HVS)

1 引言

为了满足人们对场景真实和自然再现的需求,能显示立体视频的3维电视(3DTV)应运而生^[1]。我们所看到的自然界的真实场景是具有深度信息的立体物体,但2DTV却丢失了作为第3维的深度信息^[2]。3DTV可分为两大类。第1类3DTV指采用双摄像机(双视)拍摄3维场景的和基于立体对合成显示的3DTV系统。第2类3DTV指采用多摄像机阵

列拍摄3维场景的和基于多视(或深度加多视)自由视点显示的3DTV系统。本文在以下章节中所说的3DTV主要指由两路视频组成的立体对。当人们在观看立体视频时,左右两路视频质量的变化对合成的立体视频是有影响的。对于每路的单视点视频,当视频失真小于某一阈值时,人们就难于感知到;而对于两路视频合成的立体视频,当两路视频或其中一路视频失真到什么程度,立体视频质量的变化才会使人们所感知是有待进一步研究的问题。

在对立体视频评价过程中,为了反映人眼的主观特性,要建立基于人类视觉系统(Human Visual System, HVS)模型的评价系统。最小可辨失真(Just Noticeable Distortion, JND)是指人眼对最小差别量的感知能力,也就是人类视觉系统能够感觉到的最小失真变化量。JND模型较好地反映了人眼的视

2011-08-03收到, 2011-11-15改回

国家自然科学基金重点项目(60832003), 国家自然科学基金(60902085, 61172096, 61171084), 上海市科委重点项目基金(10510500500)和安徽高校省级自然科学基金项目(KJ2011Z008)资助课题

*通信作者: 张艳 zhangyan101@shu.edu.cn

觉系统,如亮度掩盖、纹理掩盖和时间掩盖等HVS特征已应用在JND模型中。JND模型在单视视频编码和评价中都有成功的应用,许多文献都对单视JND模型进行了阐述。文献[3]提出了基于图像像素的JND模型,主要考虑了背景亮度和纹理掩盖效应两个因素,亦提出以感知的峰值信噪比(Peak Signal to Perceptible Noise Ratio, PSPNR)进行感知评价的方法。文献[4]从空间和时间相关性以及感知特性建立带有时空特性的JND模型。文献[3, 4]的纹理掩盖均用平均背景亮度和像素周围的平均亮度差异来决定,认为空间JND是纹理掩盖或者亮度适应的主要因素,用时间掩盖(运动掩盖)来评价平均帧内亮度差异。利用视频中连续帧的相似性,通过加权的亮度可适应阈值和相对的时空运动幅度来计算JND。Kelly^[5]在稳定的条件下从可视阈值实验上建立了时空对比灵敏度函数(CSF)模型。Daly^[6]改善了Kelly的模型以适应考虑到眼睛运动的无约束的自然观看条件。文献[7]提出了一种基于离散余弦变换(DCT)的JND模型,考虑了空间对比灵敏度、亮度可适应性和对比掩盖效应。文献[8]建立了在DCT变换域的JND估计方法,综合考虑了空间时间对比灵敏度、环境、亮度可适应性、帧间及帧内的对比掩盖效应,适用于图像和视频。文献[9]在文献[3]基础上提出了新的彩色视频的JND估计模型,用掩盖非线性相加模型来集成空间掩盖效应,除了考虑文献[3]中亮度掩盖、纹理掩盖和时间掩盖这些影响JND的因素外,还引入了边缘区域和非边缘区域的差异等因素。文献[10]将图像拆分为结构图像和纹理图像来进行边缘掩盖和纹理掩盖估计。然而,以上相关研究主要是针对单视点视频或者图像的JND。

由于立体视频具有不同于单视点视频的特性,例如视点融合等,使得传统的JND模型并不能直接用于立体视频编码和评价。文献[11]提出了基于自由视点显示器的立体融合系数和聚焦加权模型,并且将其应用到立体JND模型中,但此模型要依赖于显示器的类型。文献[12]提出基于心理的非对称失真立体图像的双目JND模型,文中给出了两个实验,第1个实验根据亮度掩盖和双目噪声组合设定了联合的阈值;第2个实验对于立体图像,测试了由于对比度掩盖效应而引起的在双目视觉中可视敏感度的降低情况。然而,此方法主要建立在特定的实验模型和条件基础上,对于实际应用有一定的局限性;而且文献[11]和文献[12]没有提出相关的基于所给出模型的立体视频的评价方法。

本文针对双视点视频,建立了基于背景亮度掩盖、纹理掩盖、帧间掩盖、空间时间对比灵敏度、

眼睛运动等因素的双目最小可辨失真模型(Binocular Just Noticeable Distortion, BJND),并将BJND模型引入到立体视频的评价中,提出一种基于BJND模型的立体视频评价方法。实验结果证明本文提出的评价方法更接近于主观测试结果。

2 双目最小可辨失真(BJND)模型

当视频失真小于某一范围时,人眼不能够感觉到此种影响,基于此人们提出了JND模型。视觉生理、心理等方面的研究发现人类视觉系统特性和掩盖效应对视频编码起着非常重要的作用。常见的掩盖效应包括:(1)亮度掩盖特性,人们对高亮区所附加的噪声其敏感性较小。(2)纹理掩盖特性,HVS对图像的平滑区的敏感性远远高于纹理区。背景亮度和纹理掩盖是影响JND的主要因素,此外视频的空间时间对比灵敏度也起着非常重要的作用。单路视频在融合成立体视频时,背景亮度和纹理掩盖等这些因素与单路视频之间的关系如何是本文要研究的内容之一。由于评价视频的基本元素是视频中每帧中的像素亮度,因此本文首先分析单路视频像素亮度与立体视频像素亮度的关系,并将得到的立体视频的像素亮度应用到BJND模型中。本文以背景亮度、纹理掩盖、帧间掩盖效应的JND模型和以空间时间对比灵敏度为主的基本可视阈值组合建立了BJND模型。主要流程图如图1所示。

2.1 双眼亮度关系模型

由于图像的亮度在失真图像的评价中起着非常重要的作用,因此,人们在建立JND模型时,首要的是考虑图像的亮度。立体视频由左右两路视频融合而成,对于左右视点亮度对立体视频亮度的影响,Curtis等人^[13]提出了向量和模型。给出左右视点的亮度 L_L , L_R 与组合双眼亮度 L_B 的关系为

$$L_B(x, y) = \sqrt{L_L(x, y)^{2k} + L_R(x, y)^{2k} + 2(L_L(x, y)^k L_R(x, y)^k \cos \alpha)} \cdot \lambda \quad (1)$$

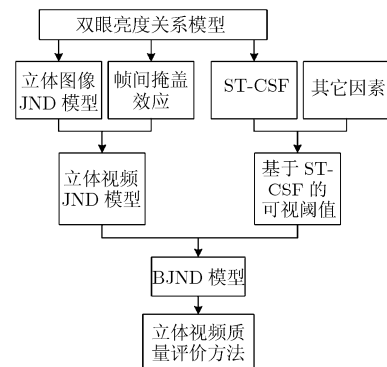


图1 本文方法流程图

经实验得到: $k \geq 0.33$, $90 \leq \alpha \leq 120$ 。这里我们取 $k = 1$, $\alpha = 120$ 。 λ 为与显示器有关的亮度校正系数, $0 < \lambda \leq 1$, 当图像的亮度与显示器显示亮度相一致时 $\lambda = 1$ 。

本文将式(1)作为立体JND模型的像素亮度, 以 $p_B(x, y)$ 表示:

$$p_B(x, y) = L_B \quad (2)$$

2.2 基于背景亮度、纹理掩盖、帧间掩盖效应的立体视频 JND 模型

对于JND模型, 首先研究基于背景亮度、纹理掩盖的立体图像的JND, 我们称为空间JND。然后通过研究立体视频的时间掩盖效应, 建立立体视频的JND, 即: 立体视频的JND为立体图像的JND与帧间掩盖效应的乘积。

当仅仅考虑空间域上的无色图像时, 主要有两个因素影响每个像素的错误可视阈值, 一个是平均背景亮度, 另一个是背景亮度的空间非均匀性即纹理掩盖效应。

考虑两个决定因素即背景亮度和纹理掩盖的立体图像的JND模型表示式为

$$\text{JND}_S(x, y) = \begin{cases} \max \{L(mg(x, y)), T(bg(x, y))\}, \\ \text{边缘像素} \\ \min \{L(mg(x, y)), T(bg(x, y))\}, \\ \text{非边缘像素} \end{cases} \quad (3)$$

人眼对边缘处像素的变化要比非边缘处像素的变化更加灵敏, 因此边缘处的可视阈值更小一些。求图像边缘像素如下:

$p_B(x, y)$ 代表由式(1)测得的像素亮度信号, 沿着每个水平方向的差分信号为

$$d_h(x, y) = p_B(x, y + 1) - p_B(x, y) \quad (4)$$

$x \in [1, M]$, $y \in [1, N - 1]$, M, N 表示图像的水平 and 垂直方向的最大像素个数。

对于水平零跨区:

$$\text{sign}d_h(x, y) = \begin{cases} 1, & d_h(x, y) > 0 \\ -1, & d_h(x, y) < 0 \\ 0, & d_h(x, y) = 0 \end{cases} \quad (5)$$

$\text{sign}d_h(x, y)$ 为水平方向的差分信号的特征符号。

与像素 $(x, y + 1)$ 沿水平方向左右相邻像素差分信号的特征符号的乘积为

$$d_{h \text{ sign}}(x, y) = \text{sign}d_h(x, y) \times \text{sign}d_h(x, y + 1) \quad (6)$$

那么, 对于 $y \in [1, N - 2]$, 判断水平方向边缘像素的因子为

$$z_h(x, y) = \begin{cases} 1, & d_{h \text{ sign}}(x, y) < 0, \text{水平方向的边缘像素} \\ 0, & \text{其它, 水平方向的非边缘像素} \end{cases} \quad (7)$$

同样方法得到判断垂直方向边缘像素的因子为 $x \in [M - 2, N]$, $z_v(x, y)$

$$= \begin{cases} 1, & d_{v \text{ sign}}(x, y) < 0, \text{垂直方向的边缘像素} \\ 0, & \text{其它, 垂直方向的非边缘像素} \end{cases} \quad (8)$$

那么, 式(3)为

$$\text{JND}_S(x, y) = \begin{cases} \min \{L(mg(x, y)), T(bg(x, y))\}, \\ z_h(x, y) = 1 \text{ 或 } z_v(x, y) = 1 \\ \max \{L(mg(x, y)), T(bg(x, y))\}, \\ z_h(x, y) = 0 \text{ 或 } z_v(x, y) = 0 \end{cases} \quad (9)$$

式(3)中的 $L(mg(x, y))$ 由式(10)表示:

$$L(mg(x, y)) = mg(x, y) \cdot \beta \quad (10)$$

β 是线性函数 L 的斜率, 经主观实验得到 $\beta = 2/17$ 。 $mg(x, y)$ 代表亮度在像素 (x, y) 周围的最大平均加权。

$$mg(x, y) = \max_{k=1,2,3,4} \{|\text{grad}_k(x, y)|\} \quad (11)$$

$$\text{grad}_k(x, y) = \frac{1}{16} \sum_{i=1}^5 \sum_{j=1}^5 p_B(x - 3 + i, y - 3 + j) \cdot G_k(i, j) \quad (12)$$

$G_k(i, j)$ 是计算因子^[3], 表示在 (x, y) 周围 4 个方向的平均加权, $k=1,2,3,4$, $G_k(i, j)$ 随着离像素中心距离的增加而减少。

式(3)中的 $T(bg(x, y))$ 由式(13)表示:

$$T(bg(x, y)) = \begin{cases} T_0 \cdot (1 - (bg(x, y) / 127)^{1/2}) + 3, \\ bg(x, y) \leq 127 \\ \gamma \cdot (bg(x, y) - 127) + 3, \\ bg(x, y) > 127 \end{cases} \quad (13)$$

$bg(x, y)$ 是平均背景亮度。 T_0 和 γ 分别代表当背景灰度值为 0 时的可视阈值和在高背景亮度时的模型的线性斜率。 $bg(x, y)$ 由加权的低通算子计算。

$$bg(x, y) = \frac{1}{32} \sum_{i=1}^5 \sum_{j=1}^5 p_B(x - 3 + i, y - 3 + j) \cdot B(i, j) \quad (14)$$

其中 $B(x, y)$, $x, y = 1, 2, 3, 4, 5$ 为平均背景亮度算子^[3]。

建立立体视频的 JND 模型除了要考虑背景亮度、纹理掩盖的立体图像的 JND 外, 还要考虑帧间的掩盖效应。用空间 JND 即立体图像的 JND 模型和帧间亮度差异函数 ILD(Interframe Luminance Difference)得到立体视频时空域上的最小可视阈值 JND_{S-T} 。帧间亮度差异函数 ILD 用第 n 帧和 $n - 1$ 帧之间的平均亮度差异函数 $d_{al}(x, y, n)$ 表示。 JND_{S-T} 的表示式为

$$\text{JND}_{S-T}(x, y, n) = \text{JND}_S(x, y, n) \cdot \text{ILD}(d_{al}(x, y, n)) \quad (15)$$

其中

$$d_{al}(x, y, n) = [p_B(x, y, n) - p_B(x, y, n-1) + bg(x, y, n) - bg(x, y, n-1)]/2 \quad (16)$$

ILD 以曲线函数表示为^[14]

$$ILD(d_{al}) = \begin{cases} 4.8 - \frac{3.6}{128}(d_{al} + 255), & d_{al} \leq -127 \\ 1.2, & -127 < d_{al} \leq 127, d_{al} = [-255, 255] \\ \frac{1.2}{128}(d_{al} - 128) + 1.2, & d_{al} \geq 128 \end{cases} \quad (17)$$

2.3 基于空间时间对比灵敏度(ST-CSF)的可视阈值

双目 JND 模型除了考虑以上提到的背景亮度及纹理和帧间掩盖效应外, 还要考虑基本的可视阈值。基本的可视阈值主要考虑灵敏度模型即空间时间对比灵敏度(ST-CSF)。另外还要考虑用亮度阈值表达的对比灵敏度函数需要延伸到数字图像的灰度值。ST-CSF 以 CSF(x, y, n) 表示, 则其可视阈值表达式如下:

$$T(x, y, n) = \frac{1}{CSF(x, y, n) (L_{MAX} - L_{MIN})} \frac{M \cdot \lambda}{M \cdot \lambda} \quad (18)$$

其中 L_{MAX} 和 L_{MIN} 分别代表与最大和最小灰度值相对应的显示器的亮度值, M 是图像系统的灰度级别数(大多数图像系统是 256), λ 为与显示器有关的亮度校正系数。

ST-CSF 是描述人眼视觉系统空间-时间特性的主要指标之一, 反映了不同条件下的对比灵敏度与空间-时间频率之间的关系, 也可指与图像通过人眼视网膜速度相关的人类视觉系统的空间敏锐性, 以视网膜上的图像成像速度表示时间频率。典型的 ST-CSF 函数曲线表示式如下^[5]:

$$CSF(\alpha, v) = C_0(6.1 + C_1 \cdot |\log(k \cdot v / 3)|^3) \cdot k \cdot v \cdot \alpha^2 \exp(C_3 \cdot \alpha \cdot (k \cdot v + 2) / 45.9) \quad (19)$$

其中 C_0, C_1, C_3, k 为常数, 这里我们取 $C_0 = 1, C_1 = 7.3, C_3 = -2, k = 1, v$ (degrees per second) 是视网膜成像速度, α (cycles per degree) 表示空间频率。

空间频率是反映图像空间变化程度的一个量。设尺寸为 $M \times N$ 的立体图像在 (x, y) 处的像素亮度为 $p_B(x, y)$, 则立体图像的空间频率定义如下:

$$\alpha = \sqrt{(\alpha_R)^2 + (\alpha_C)^2} \quad (20)$$

α_R 为行频率:

$$\alpha_R = \sqrt{\frac{1}{MN} \sum_{x=1}^M \sum_{y=2}^N [p_B(x, y) - p_B(x, y-1)]^2} \quad (21)$$

α_C 为列频率:

$$\alpha_C = \sqrt{\frac{1}{MN} \sum_{j=1}^N \sum_{x=2}^M [p_B(x, y) - p_B(x-1, y)]^2} \quad (22)$$

图像的空间频率随着图像分辨率的不同而不同, 图像越清晰分辨率越高其空间频率越大。于是, 在第 n 帧上的点 (x, y) 的 ST-CSF 函数为

$$CSF(x, y, n) = (6.1 + 7.3 |\lg(v(n)/3)|^3) \cdot v(n) \cdot ((\alpha_R)^2 + (\alpha_C)^2) \exp(-2 \cdot \sqrt{(\alpha_R)^2 + (\alpha_C)^2} \cdot (v(n) + 2) / 45.9) \quad (23)$$

ST-CSF 与观察者眼睛的运动相关, 且要考虑运动图像的运动特征。人们在观看视频的时候, 人眼会追踪视频中的运动物体, 而忽略静止物体的微弱的失真。由于眼睛运动, 在视网膜上的感知速度与通过运动估计得到的图像平面速度不同, 在 Daly^[6] 设计的眼睛运动的初级模型上, 视网膜中图像的速度可以表示为

$$v(n) = v_I(n) - v_E(n) \quad (24)$$

$v_I(n)$ 表示第 n 帧中如果没有眼睛运动的视网膜中图像平面物体的速度; $v_E(n)$ 表示在第 n 帧中眼睛运动的速度。其中,

$$v_E(n) = \min[gv_I(n) + v_{MIN}, v_{MAX}] \quad (25)$$

g 表示跟踪物体的效率, v_{MIN} 和 v_{MAX} 分别指眼睛运动的最小和最大速度。而式(24)中的 v_I 可近似表示为

$$v_I(n) = \frac{\text{图像从第 } n-1 \text{ 帧到第 } n \text{ 帧的运动距离}}{\text{运动所需时间}} = \frac{\sqrt{(MV_x(n))^2 + (MV_y(n))^2}}{1/f} \quad (26)$$

f 指帧率(frames per second), $MV_x(n)$ 和 $MV_y(n)$ 指在第 n 帧中图像的运动向量。

这里假设立体视频的左右两路视频的运动速度保持一致。因此 $v(n)$ 也可代表立体视频的运动速度。

2.4 双目最小可辨失真模型

综合 2.1 节, 2.2 节和 2.3 节, 本文提出了双目最小可辨失真(BJND)模型, 可表示为

$$JND_B(x, y, n) = JND_{S-T}(x, y, n) T(x, y, n) = JND_S(x, y, n) ILD(d_{al}(x, y, n)) \cdot \frac{1}{CSF(x, y, n) (L_{MAX} - L_{MIN})} \frac{M\lambda}{M\lambda} \quad (27)$$

3 应用于立体视频的质量评价

本文将双目最小可辨失真模型引入传统的单视点感知质量评价方法 PSPNR(Peak Signal to Perceptible Noise Ratio)^[4]中, 由式(28)可以得到我们提出的评价立体视频质量的表示式 BPSNR(Binocular Peak Signal to Perceptible Noise Ratio)。

$$\text{BPSPNR} = 10 \cdot \lg \frac{255 \cdot 255}{\frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N \left\{ \left[|p_B(x, y, n) - \hat{p}_B(x, y, n)| - \text{JND}_B(x, y, n) \right]^2 \cdot \delta(x, y, n) \right\}} \quad (28)$$

$$\delta(x, y, n) = \begin{cases} 1, & |p_B(x, y, n) - \hat{p}_B(x, y, n)| > \text{JND}_B(x, y, n) \\ 0, & |p_B(x, y, n) - \hat{p}_B(x, y, n)| \leq \text{JND}_B(x, y, n) \end{cases}, \quad 1 \leq x \leq M, 1 \leq y \leq N \quad (29)$$

其中 $p_B(x, y, n)$ 为原始视频第 n 帧左右视点重构后的像素亮度, $\hat{p}_B(x, y, n)$ 代表失真视频第 n 帧左右视点重构后的像素亮度, $\text{JND}_B(x, y, n)$ 为双目最小可辨失真模型。

4 实验结果与分析

用本文算法(BPSPNR)对MEPG提供的立体测试序列“Book_arrival”, “Champagne_tower”, “Lovebird1”^[15]和微软研究所给出的“Ballet”^[16]进行了测试,并作了相关的主观实验进行验证。每个序列分别取其中相邻的两路视频,以4个不同的QP(22, 28, 34和40)进行立体视频的失真处理。

主观实验参照国际标准ITU-R BT.500-11^[17]。显示器是16:10(宽为40.9 cm,高为25.5 cm)的SuperD自由立体显示器,分辨率为1440×900。15个测试者参加了本次测试,包括专业测试者和非专业测试者。测试等级分为5级,为了便于精确记录结果,用百分制来代替5分制,即1[bad:0-20], 2[poor: 20-40], 3[fair:40-60], 4[good:60-80], 5[excellent: 80-100],测试者在进行测试时所在位置为其最佳观测位置。

客观实验结果与PSNR, PSPNR^[4]进行了比较。文中的PSNR和PSPNR方法,均由左右眼的平均值得到。BPSPNR是本文提出的方法。图2显示了PSNR, PSPNR, BPSPNR这3种客观评价结果和主观评价价值MOS(Mean Opinion Score)的关联曲线图,图中所得曲线由软件Origin7.5提供的非线性函数中的Logistic函数拟合而成。

在建立非线性拟合之后,我们采用如下视频质量专家组(Video Quality Expert Group, VQEG)推荐的4个度量指标来比较各种方法的优劣^[18,19]。(1)

Pearson相关系数(Pearson Correlation Coefficient, PCC),反映预测的精确性,区间为[-1, 1]上的值,其绝对值越接近于1,表明主客观评价间的相关性越好;(2)Spearman等级相关系数(Spearman Rank Order Correlation Coefficient, SROCC),表征了评价方法的单调性,也是区间[-1, 1]的值,同PCC一样,其绝对值越接近于1,表明主客观评价间的一致性越好;(3)离出率(Outlier Ratio, OR),反映客观评价的一致性,数值越小则表明模型预测越好,由通过主客观所得数据的非线性拟合后处于误差大于标准差2倍的点的比例而得;(4)均方根误差(Root Mean Square Error, RMSE),表征了数据的离散程度,是对主客观所得数据的非线性拟合后所得误差计算而得。表1为各方法的比较结果,可以看到,本文提出的BPSPNR评价方法其PCC和SROCC指标较大,而OR和RMSE指标较小。也就是说无论是从与主客观评分的相关性还是从数据样本的分散度来分析,本文所提出的BPSPNR方法均优于其它两种方法,更加接近于人类的视觉感知。

表1 客观评价模型性能指标比较

MODEL	PCC	SROCC	OR	RMSE
PSNR	0.8737	0.8882	0.3750	4.6074
PSPNR	0.9474	0.9324	0.6250	2.2008
BPSPNR	0.9829	0.9765	0.3125	1.8572

5 结论

本文针对于立体视频不同于单路视频的特性建立了基于背景亮度掩盖、纹理掩盖、帧间掩盖效应、空间时间对比灵敏度、眼睛运动等因素的双目最小

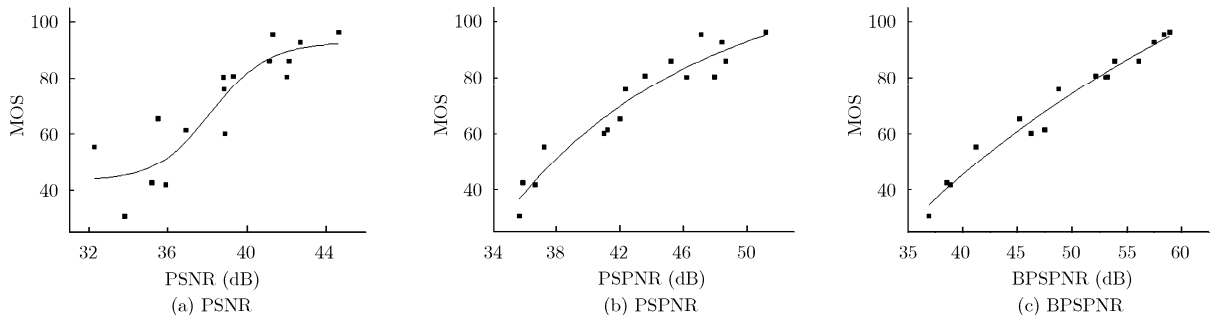


图2 客观质量评价方法与主观MOS值的关联图

可辨失真模型(BJND), 提出了基于BJND模型的立体视频质量评价方法。对提出的BPSPNR方法进行了主观验证, 并和PSNR, PSPNR评价方法进行了客观评价与主观评价一致性比较。实验结果验证了本文所提出的方法更加接近于人类的视觉感知。

参 考 文 献

- [1] Müller K, Merkle P, and Wiegand T. 3-D video representation using depth maps[J]. *Proceedings of the IEEE*, 2011, 99(4): 643-656.
 - [2] Nicolas S H, Neil A D, Gregg E F, *et al.* Three-dimensional displays: a review and applications analysis [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 372-383.
 - [3] Chou C H and Li Y C. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1995, 5(6): 467-476.
 - [4] Chou C H and Chen C W. A perceptually optimized 3-D subband codec for video communication over wireless channels[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 1996, 6(2): 143-156.
 - [5] Kelly D H. Motion and vision II: stabilized spatiotemporal threshold surface[J]. *Optical Society of America*, 1979, 69(10): 1340-1349.
 - [6] Daly S. Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models, Vision Models and Applications to Image and Video Processing [M]. Norwell, MA: Kluwer Academic Publishers, 2001, Chapter 9.
 - [7] Zhang X H, Lin W S, and Xue P. Improved estimation for just-noticeable visual distortion[J]. *Signal Processing*, 2005, 85(4): 795-808.
 - [8] Jia Y T, Lin W S, and Kassim A A. Estimating just-noticeable distortion for video[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2006, 16(7): 820-829.
 - [9] Yang X K, Lin W S, Lu Z K, *et al.* Just noticeable distortion model and its applications in video coding[J]. *Signal Processing: Image Communication*, 2005, 20(7): 662-680.
 - [10] Liu A M, Lin W S, Paul M, *et al.* Just noticeable difference for images with decomposition model for separating edge and textured regions[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2010, 20(11): 1648-1652.
 - [11] Zhang L, Peng Q, Wang Q H, *et al.* Stereoscopic perceptual video coding based on just-noticeable-distortion profile[J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 572-581.
 - [12] Zhao Y, Chen Z Z, Zhu C, *et al.* Binocular just-noticeable-difference model for stereoscopic images[J]. *IEEE Signal Processing Letters*, 2011, 18(1): 19-22.
 - [13] Curtis D W and Rule S J. Binocular processing of brightness information: a vector-sum model[J]. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, (4): 132-143.
 - [14] Zhao Y and Yu L. Perceptual measurement for evaluating quality of view synthesis[S]. ISO/IEC JTC1/SC29/WG11/M16407, Maui, USA, April 2009.
 - [15] ISO/IEC JTC1/SC29/WG11. Draft call for proposals on 3D video coding technology[S]. MPEG2011/N11830, Daegu, Korea, January 2011.
 - [16] Zitnick C L, Kang S B, Uyttendaele M, *et al.* High-quality video view interpolation using a layered representation[J]. *ACM Transactions on Graphics*, 2004, 23(3): 600-608.
 - [17] ITU-R Recommendation BT.500-11. Methodology for the subjective assessment of the quality of television pictures[S]. 2002.
 - [18] Video Quality Experts Group (VQEG). Final report from the video quality experts group on the validation of objective models of video quality assessment, Phase II[R]. <http://www.vqeg.org/>, USA, August 25, 2003.
 - [19] Chikkerur S, Sundaram V, Reisslein M, *et al.* Objective video quality assessment methods: a Classification, review, and performance comparison [J]. *IEEE Transactions on Broadcasting*, 2011, 57(2): 165-182.
- 张 艳: 女, 1979年生, 博士生, 研究方向为立体视频质量评价。
安 平: 女, 1968年生, 教授, 博士生导师, 研究方向为图像及视频信号处理、编码及应用。
张兆杨: 男, 1938年生, 教授, 博士生导师, 研究方向为数字视频处理及多媒体通信。