

DOI:10.3969/j.issn.1673-4785.201206021

网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20120921.1545.002.html>

## 基于感知视觉重要性的立体图像质量评价

姜峰, 郭美思, 刘绍辉, 赵德斌

(哈尔滨工业大学 计算机科学与技术学院, 黑龙江 哈尔滨 150001)

**摘要:**视觉心理、生理因素是有效、准确评价图像质量的重要依据。尽管在计算层面已有众多视觉心理、生理计算模型及方法为其提供支持,但在图像质量评价任务中如何分析各种孤立方法之间的内在关系进而使之有效协同,是使得评价结果更符合人主观评测的关键。从图像质量评价的角度出发,功能上将人眼的视觉注意区域定义并数学形式化为视觉初期注意区域与视觉转移期的劣质区域;同时考虑人眼的感知冗余特性,结合JND感知冗余模型,进而提出了图像质量评价范畴下的视觉感知模型PVSSIM。以此为依据,将感知视觉的方法在二维图像数据库中验证其可行性,并将其引入到立体图像质量评价中。实验结果表明,提出的客观评价方法与传统方法相比,充分考虑到了图像质量评价任务中各种视觉心理、生理因素的协同,与人主观的图像质量评价相比具有更高的相关度,评价方法在立体图像库中能很好地与主观评价相吻合,达到了预期的效果。

**关键词:**图像质量评价;立体图像评价;JND;视觉心理;生理模型;感知视觉

**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2012)05-0414-09

## Quality assessment of stereoscopic images based on the significance of perceptual vision

JIANG Feng, GUO Meisi, LIU Shaohui, ZHAO Debin

(Department of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China)

**Abstract:** The visual psychological and physiological factors are crucial for the assessment on image quality. There are many visual, psychological and physiological calculation models used to support the assessment of image quality, however, how to analyze the intrinsic relations among various isolated methods and further effectively coordinate them in the task of image quality assessment is the key to making the assessment results more conform to subjective assessment of human. The definition and formalization of visual attention areas as saliency in the early vision period and poor-quality area for vision transformation periods were examined from an image quality angle. Simultaneously, the property of perceptual redundancy of human eyes, JND(just noticeable distortion) perception redundancy model, and vision perception model PVSSIM of image equality assessment will be examined. Based on the research findings, feasibility of vision perception methods are verified using a two-dimensional image database study through the introduction of image stereoscopic. The experimental results confirm that, in comparison with traditional methods, the raised objective assessment method sufficiently considers the coordination of various psychological and physiological factors in the task of image quality assessment, it is more relevant to the subjective image quality assessment of human. In the stereoscopic image database, the assessment method may properly coincide with the subjective assessment, the expectable effect is realized.

**Keywords:** image quality assessment; stereoscopic image; just noticeable distortion(JND); visual psychology; physiological model; perceptual vision

在信息、媒体、通讯高速发展的时代,人们已经

不满足于二维平面图像带来的享受。2010年我国举行的第16届亚运会,首次实现了综合运动会的三维传播,这些都为立体图像视频的欢迎做好了铺垫。随着人们生活水平的提高,3DTV、3D电影、家庭影院等都在人们的日常生活中越来越普及。在这种强大

收稿日期:2012-06-18。网络出版日期:2012-09-21。

基金项目:国家自然科学基金资助项目(61100096);国家“973”计划资助项目(2009CB320905)。

通信作者:姜峰。E-mail: fjiang@hit.edu.cn。

的市场需求的情况下,立体图像视频已经成为当前研究的热点。

立体图像的2幅图像(左视和右视)是通过2个很近的摄像机(近似两眼之间的距离)采集的。为了节省带宽和存储空间,左右视图图像会被相应压缩。虽然3D图像的获取技术在不断出现,但这些技术和图像压缩一样,极大地影响了3D的感知质量。在每个环节中,原始图像都会产生失真的可能性。与传统的2D图像/视频不同,3D图像/视频在采集、压缩、存储、传输、重建过程中需要克服新的挑战,因为三维图像能利用左、右眼观察到的具有视差的图像对,在大脑中融合出真实的立体感觉。三维图像有时会给用户带来梯形失真、鬼影效应、木偶戏效应、硬纸板效应、切形变、眩晕等特有的现象,因此研究适于人眼的3D图像质量评价方法就显得非常重要。

## 1 图像质量评价

立体技术<sup>[14]</sup>的不断涌现,为研究立体图像提供了技术保障。从当前的国内外研究现状来看,可以将立体图像质量评价方法分为以下3种类别:第1种是基于2D图像评价方法的立体图像评价;第2种是基于2D评价方法融合3D信息的立体图像评价;第3种是基于3D分析的立体图像质量评价。

1)第1种质量评价方法的思想是基于最新的2D质量评价方法来评价3D图像质量。近年来,2D图像质量评价方法主要有PSNR、MSE、SSIM<sup>[5]</sup>、MSSIM<sup>[6]</sup>、P-SSIM<sup>[7]</sup>、IWSSIM<sup>[8]</sup>、VASSIM<sup>[9]</sup>等。You等<sup>[10]</sup>已经将上述的一些方法应用到立体图像评价中,并研究了其性能。立体图像中的失真图像对被单独分开作为2D方法的输入,最后的分数是左、右图像分数的平均值。实验结果显示,以上方法对立体图像评价是有效的,虽然基于以上方法的立体图像评价性能没有2D图像性能好。

2)第2种质量评价方法的思想是将立体图像的差异图或深度图以不同的方式融合在标准框架中。You等使用3种不同的方法融合视差图来得到最后的质量评价分数<sup>[10]</sup>。Yang等根据立体图像对的绝对差异图来进行立体图像质量评价<sup>[11]</sup>。Hewage和Martini提出的方法计算原始和受损的深度图的轮廓,然后这些轮廓被二值化,继而使用PSNR方法评价图像质量<sup>[12]</sup>。Xing等提出了从立体图像中计算出深度图的方法,然后通过原始和受损的深度图像得出SSIM map被赋予相应的权值,这个方法的分数是带权深度图的平均值<sup>[13]</sup>。

3)第3种方法的主要思想是提取3D的感知特

性<sup>[13]</sup>。Roushain等提出了无参的方法评价非对称的JPEG压缩图像,基于图像对在固定块的划分作为3D人工产品的特性,获得的块被分成平坦块(无轮廓)和活跃块(有轮廓),利用这2个特征评估立体对的质量<sup>[14]</sup>。基于差异图,Kim等提出了3D立体视频质量评价方法。首先进行主观实验作为3D质量准则,然后根据色差及其在时间上的平均变化来实现该方法<sup>[15]</sup>。Olsson等提出了一个基于深度图的方法<sup>[16]</sup>,在预处理阶段,他们在不同深度层次上合成2D立体对,层次是根据相机距离和焦距信息来选择的,在每个层次中,像素根据文献[17]的算法识别,编码噪音都用MSSIM的方法评价<sup>[16]</sup>。国内的研究机构也在此方向做了很多类似的研究,并得出一些重要的结论<sup>[18-20]</sup>。

## 2 感知视觉重要性方法

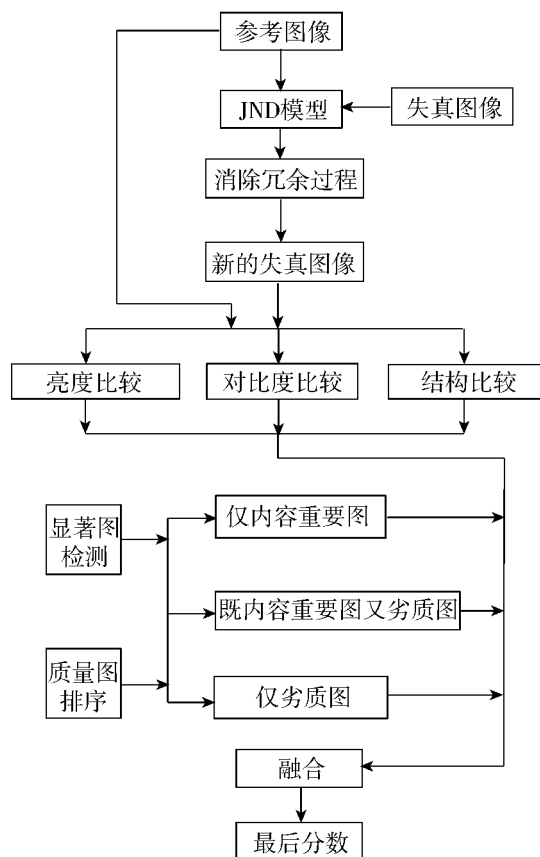


图1 基于视觉重要性的图像质量评价框架

Fig.1 The framework of image quality assessment based on the significance of perceptual vision

人眼感知冗余特性是由于图像的亮度对比和空间掩藏效应产生的。JND(just noticeable distortion)模型能模拟HVS的亮度对比度和空间掩藏效应等特性,能表示人眼的感知冗余,在感知冗余评估上与人眼相吻合。另一方面,在一副图像中,只有部分区

域能引起视觉注意. 这部分区域可以是视觉初期注意到的内容重要性区域, 也可以是视觉转移时注意到的质量不好区域. 因此, 本文从这2个方面来检测视觉重要性区域. 对于不同的区域来说, 观看者注意程度是不同的, 对不同的区域训练出不同的权值, 然后形成视觉感知模型 VPM (vision perception model). 基于感知视觉重要性的图像质量评价框架如图1所示.

### 2.1 感知冗余

图像的感知冗余主要是由于亮度对比度和空间掩盖效应引起的. 在 JND 得出的可见性门限阈值内, 图像的变化在人眼中是察觉不到的; 然而超出这个阈值, 人们就能察觉到图像的失真.

为使图像质量评价更加符合人眼特性, 如果像素值在 JND 的可见性门限阈值内, 则将待测图像的像素值修改成参考图像的像素值; 否则, 待测图像的像素值会相应地增加或减少 JND 的值. 本文采用的是空域 JND 模型<sup>[21]</sup>, 具体过程如下.

1) 利用模型计算参考图像的 JND 值  $S_{jnd}$ .

2) 利用 JND 特性修改失真图像. 如果原始图像  $O(x, y)$  和失真图像  $D(x, y)$  的差值在门限阈值内, 则  $D(x, y) = O(x, y)$ ; 否则, 失真图像的像素值减去门限阈值:

$$D(x, y) = D(x, y) - S_{jnd}(x, y).$$

如果原始图像像素值大于失真图像像素值, 则失真图像的像素值加上门限阈值:

$$D(x, y) = D(x, y) + S_{jnd}(x, y).$$

本文利用空间 JND 模型来得到能感知冗余的可见性门限阈值, 如式(1):

$$S_{jnd} = \max \{f_1(b_g(x, y), m_g(x, y)), f_2(b_g(x, y))\}. \quad (1)$$

式中:  $f_1(b_g(x, y), m_g(x, y))$  和  $f_2(b_g(x, y))$  分别是估计空间掩盖和亮度对比的函数, 而  $f_1(b_g(x, y), m_g(x, y))$  定义为

$$f_1(b_g(x, y), m_g(x, y)) = m_g(x, y) \times \alpha(b_g(x, y)) + \beta(b_g(x, y)).$$

式中:  $m_g(x, y)$  是像素  $(x, y)$  在4个方向上亮度变化的最大加权平均值, 如式(2):

$$m_g(x, y) = \max_{k=1,2,3,4} \{ |g_k(x, y)| \}, \quad g_k = \frac{1}{16} \sum_{i=1}^5 \sum_{j=1}^5 p(x-3+i, y-3+j) * G_k(i, j). \quad (2)$$

式中:  $G_k(i, j)$  是检测纹理的高通滤波器, 如图2所示.

0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	-1	0
1	3	8	3	1	0	0	8	3	0	0	0	0	0	3	8	0	0	3	0	-3	0
0	0	0	0	0	0	1	3	0	-3	-1	0	-1	-3	0	3	1	0	8	0	-8	0
-1	-3	-8	-3	-1	0	0	0	-3	-8	0	0	0	-8	-3	0	0	0	3	0	-3	0
0	0	0	0	0	0	0	0	-1	0	0	0	0	0	-1	0	0	0	1	0	-1	0

图2 4个方向的高通滤波器

Fig. 2 The high-pass filters in 4 directions

而  $\alpha(b_g(x, y))$  和  $\beta(b_g(x, y))$  被定义为:

$$\alpha(b_g(x, y)) = b_g(x, y) \times 0.0001 + 0.115,$$

$$\beta(b_g(x, y)) = \mu - b_g(x, y) \times 0.01,$$

$$b_g(x, y) = \frac{1}{32} \sum_{i=1}^5 \sum_{j=1}^5 p(x-3+i, y-3+j) * B(i, j).$$

式中:  $b_g(x, y)$  为平均背景亮度, 通过滤波 B 实现, 如图3所示.

1	1	1	1	1
1	2	2	2	1
1	2	0	2	1
1	2	2	2	1
1	1	1	1	1

图3 低通加权滤波器

Fig. 3 The low-pass filter

函数  $f_2(b_g(x, y))$  计算可见门限值, 如式(3):

$$f_2(b_g(x, y)) = \begin{cases} T_0 \times [1 - (\frac{b_g(x, y)}{127})^{1/2} + \varepsilon], & b_g(x, y) \leq 127; \\ \gamma \times [b_g(x, y) - 127] + \varepsilon, & b_g(x, y) > 127. \end{cases} \quad (3)$$

### 2.2 视觉重要性模型

视觉注意是指对图像的某一部分特别注意, 而能引起注意的区域包括内容重要性的显著区域和劣质区域. 本文方法将图像分为4个部分区域: 既是显著区域又是劣质区域、仅是显著区域、仅是劣质区域、既不是显著区域又不是劣质区域. 然后, 将这4个部分区域分别赋予不同的权值来形成视觉重要性模型 (vision significance model, VSM).

#### 2.2.1 显著性模型

显著区域是指在图像中能提取出图像的重要内容的区域, 可以依据图像在频域中的特性来提取. 根据信息论的观点, 图像信息能被分解为两部分: 新颖部分和先验部分. 不同的图像在频谱对数曲线中有共同的曲线趋势, 而频谱上的频谱冗余部分即是图像的新颖部分, 由此来构造图像的显著区域<sup>[22]</sup>.

设  $I(x, y)$  为一输入图像,  $A(u, v)$  和  $P(u, v)$  分别是经过傅里叶变换后的频谱和相谱, 对数频谱  $L(u, v)$  为

$$L(u, v) = \ln(A(u, v)).$$

式中:  $A(u, v)$  是对数频谱的一般形式, 相当于给定

的先验部分. 而  $A(u, v)$  的平均频谱  $A_1(u, v)$  可以用局部滤波器来近似  $A(u, v)$  的形状:

$$A_1(u, v) = h_n(u, v) * L(u, v).$$

滤波器  $h_n(u, v)$  的矩阵表达形式如式(4):

$$H_n(u, v) = \frac{1}{n^2} \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \quad (4)$$

综上, 频谱冗余  $R(u, v)$  为

$$R(u, v) = L(u, v) - A_1(u, v).$$

此时频谱冗余意味着图像中的新颖部分, 即显著部分. 之后再通过反傅里叶变换得到显著图  $S(x, y)$ , 如式(5):

$$S(x, y) = g(x, y) * \mathcal{F}^{-1}[\exp(R(u, v) + P(u, v))]^2. \quad (5)$$

式中:  $g(x, y)$  是一个高斯滤波器, 目的是通过平滑产生较好的视觉效果. 显著图突出了吸引人眼注意的物体, 为了检测显著图中的前物体, 可采用阈值的方法, 则前物体图  $O(x, y)$  定义为:

$$O(x, y) = \begin{cases} 1, & S(x, y) > T; \\ 0, & \text{其他.} \end{cases}$$

式中:  $T = E(S(x, y)) \times 3$ .

### 2.2.2 劣质图像模型

在观看图像时, 劣质区域同内容重要区域一样能吸引人的注意. 在本文中, 劣质图像区域的确定采用基于百分比的模型<sup>[7]</sup>. 参考图像和失真图像经过 SSIM 获得的质量图 SSIM\_map 进行从小到大排序, 若像素值  $f_{\text{ssim}}(x, y)$  落入前  $p\%$  的集合范围内, 就标记为图像劣质, 其值为 1, 否则为 0. 将前  $p\%$  的值的集合定义为集合  $A$ , 则图像劣质图  $V(x, y)$  定义为:

$$V(x, y) = \begin{cases} 1, & f_{\text{ssim}}(x, y) \in A; \\ 0, & \text{其他.} \end{cases}$$

### 2.2.3 视觉重要性模型

由于提取视觉重要性的时候将图像分为 4 个部分区域: 既是显著区域又是劣质区域、仅是显著区域、仅是劣质区域、既不是显著区域又不是劣质区域. 因此, 可以提取既是显著又是劣质的图像区域, 其被定义为双边重要区域 (both importance map, BI map):

$$B(x, y) = \begin{cases} 1, & O(x, y) = 1, V(x, y) = 1; \\ 0, & \text{其他.} \end{cases}$$

同理, 定义  $\Theta(x, y)$  为仅有显著区域没有劣质区域,  $\Psi(x, y)$  为仅有劣质区域没有显著区域. 将以上 3 个部分的区域分别赋予不同的权值,  $\Theta(x, y)$  的权值  $r_s \geq 1$ ,  $\Psi(x, y)$  的权值  $r_v \geq 1$ ,  $B(x, y)$  的权值  $r_b \geq 1$ .

根据不同区域赋予不同的权值, 权值被训练

(步长从 1 到 4 000) 以得到最好的增益, 在客观分数和主观分数之间的相关系数值最大时, 就可得到该尺度的 3 个权值, 每个尺度上都是这样训练得到权值. 本文在 LIVE 数据库中抽取各种失真类型和不同失真程度的 150 幅图像进行实验.

通过分别分析 4 个尺度上的  $\psi$ -SSIM、 $\Theta$ -SSIM、BI-SSIM 与主观分数之间相关系数的散点图得出最优的  $r_s$ 、 $r_v$ 、 $r_b$  值, 其训练的结果如表 1 所示. 实验结果表明, 在显著区域且劣质区域的权值比仅是显著区域或者仅是劣质区域的权值高.

表 1 不同尺度上的  $r_s$ 、 $r_v$ 、 $r_b$  值

Table 1 The  $r_s$ ,  $r_v$  and  $r_b$  with different scales

模 型	尺度 1	尺度 2	尺度 3	尺度 4
$\psi$ -SSIM	28	5	3	1
$\Theta$ -SSIM	20	2	1	2
BI-SSIM	40	7	4	6

因此, 视觉重要性模型为

$$V_s(x, y) = \{\Theta(x, y) + \frac{\bar{\Theta}(x, y)}{\Psi(x, y)}\Psi(x, y) + \frac{\bar{\Theta}(x, y)}{B(x, y)}B(x, y)\} / 3.$$

式中:  $\bar{\Theta}(x, y) = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \Theta(x, y)$ ,  $\bar{\Psi}(x, y) =$

$$\frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N \Psi(x, y), \bar{B}(x, y) = \frac{1}{M \times N} \sum_{x=1}^M \sum_{y=1}^N B(x, y).$$

### 2.3 感知视觉方法

感知视觉方法主要由以下 4 个步骤实现:

1) 通过参考图像和失真图像计算出 SSIM 中的质量图;

2) 再通过参考图像计算出 JND 值, 消除视觉的感知冗余, 突出失真区域;

3) 通过显著区域和劣质区域来计算视觉重要性模型 VS;

4) 最后融合到多尺度 SSIM 中得到最终的质量评价结果.

设  $M$  为最高的尺度,  $M$  的值取为 5,  $w_{j,i}$  为第  $j$  个尺度上的 VS 图在空间位置  $i$  的值. 则对于  $j=1, 2, \dots, M-1$  时, 第  $j$  个尺度的 PVSSIM 的定义如式(6):

$$F_{\text{pvssim}}(j) = \frac{\sum_i w_{j,i} \times c(X_{j,i}, Y_{j,i}) \times s(X_{j,i}, Y_{j,i})}{\sum_i w_{j,i}}. \quad (6)$$

当  $j=M$  时, 第  $j$  个尺度的 PVSSIM 的定义如式(7):

$$F_{\text{pvssim}}(j) = \frac{1}{N_j} \sum_i l(X_{j,i}, Y_{j,i}) \times c(X_{j,i}, Y_{j,i}) \times s(X_{j,i}, Y_{j,i}). \quad (7)$$

式中:  $l(X_{j,i}, Y_{j,i})$ 、 $c(X_{j,i}, Y_{j,i})$  和  $s(X_{j,i}, Y_{j,i})$  分别为

SSIM 中的 3 个分量. 最后的 PVSSIM 被定义为

$$F_{pvssim} = \prod_{j=1}^M (F_{pvssim}(j))^{\beta_j}$$

式中: 每个尺度  $\beta_j$  的权值与文献[6]中的多尺度 SSIM 保持一致, 尺度  $M=5$ .

#### 2.4 实验结果

本文采用的实验数据库图像是由 982 幅图像构成的 LIVE 数据库<sup>[23]</sup>. 该数据库有 5 种失真类型, 其中 JPEG 压缩图像有 233 幅, JPEG2000 压缩图像有 227 幅, 白噪声的图像有 174 幅, 高斯模糊的图像有 174 幅, 快速退化的图像有 174 幅, 还有 203 幅图像是参考图像即无失真图像.

本文使用 4 种 IQA 测量性能的评估标准: PLCC、SRCC、KRCC、RMSE 来衡量提出的新方法, 并和最近流行的客观评价方法比较. 在评测中, 一个较好的客观评价方法应该有较高的 PLCC、SRCC、KRCC 和较低的 RMSE. 表 2 中详细具体地显示了当前具有代表性的各个客观质量评价方法在 LIVE 图像数据库上的评估指标.

表 2 客观图像质量评价方法在 LIVE 库上的比较

Table 2 Comparison of the objective stereoscopic images assessment methods in LIVE

模 型	PLCC	SRCC	KRCC	RMSE
SSIM <sup>[5]</sup>	0.902	0.911	0.732	11.801
MSSSIM <sup>[6]</sup>	0.940	0.951	0.804	9.312
P-SSIM <sup>[7]</sup>	0.937	0.946	0.796	9.537
IWSSIM <sup>[8]</sup>	0.943	0.957	0.818	9.137
VASSIM <sup>[9]</sup>	0.948	0.957	0.817	8.668
PVSSIM	0.949	0.958	0.818	8.584

图 4 显示了 PVSSIM 方法在 LIVE 数据库上的散点图, 每一个点代表一幅图像. 从评估数据和散点图中可以看出, 基于感知视觉的方法符合人眼的特性, 与主观评价吻合.

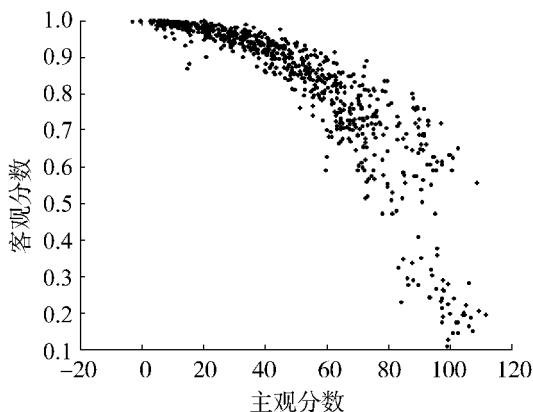


图 4 PVSSIM 方法在 LIVE 库上的散点

Fig. 4 The scatters of PVSSIM in LIVE base

### 3 立体图像质量评价

目前 3D 图像质量评价处于探索阶段, 与 2D 图像质量评价相比, 它有着独特的三维信息即立体感. 三维图像技术的基本原理是利用人的左、右眼分别观察具有视差相关性的左右视图, 在人脑中融合出真实的立体世界. 因此, 本文将 3D 图像质量评价分为质量评价和立体感评价.

当观看立体图像时, 大脑会依据双眼产生的视差信息来融合产生立体感. 绝对视差图对立体感有很重要的影响<sup>[11]</sup>, 而影响立体视觉的因素还有很多, 如阴影、重叠、双眼会聚、个人的经验和记忆等, 但是能直接影响立体图像中的立体感因素就是双眼视差. 图像对(具有标准视差)的绝对差异图像可以近似为图像的轮廓线, 由于图像的像素值不能为负值, 因而取绝对差异图. 若参考图像对如图 5 所示, 则它的绝对差异图如图 6 所示.

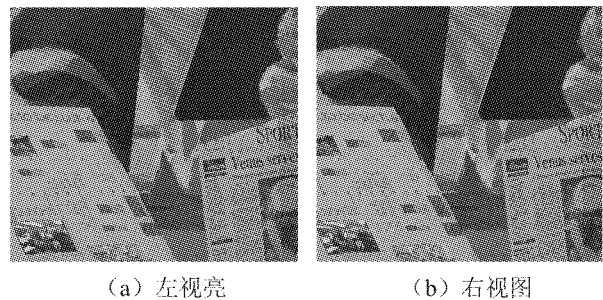


图 5 参考立体图像对

Fig. 5 The inference images



图 6 绝对视差图

Fig. 6 The absolute parallax

在研究立体感和绝对差异图之间关系时做了以下实验<sup>[11]</sup>. 参考图像对来自图 5, 绝对差异图来自图 6. 在参考图像对位于绝对差异图中值较大的区域加噪声, 这时采取主观评价的方法观察图像对, 立体感觉几乎丧失, 但质量还可以. 在参考图像对位于绝对差异图中值较小的区域加噪声, 这时如果采取主观评价的方法观察图像对, 立体感觉几乎没有损害, 但质量会很差. 如图 7~8 所示.

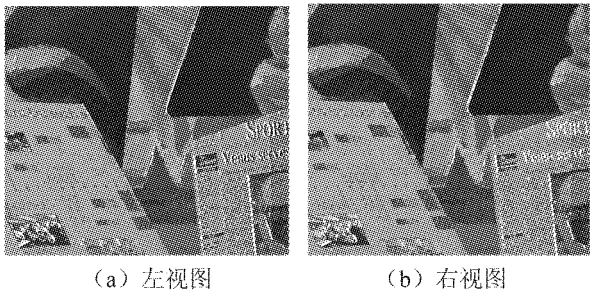


图 7 在相对值较大的区域加入噪声  
Fig. 7 Images added noise in the significant value area

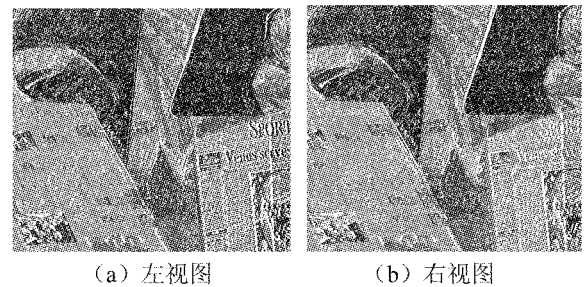


图 8 在相对值小的区域加入噪声  
Fig. 8 Images added noise in the small value area

在双眼差异值大的区域加入恒值噪声,立体感会减退,噪声把两眼的差异部分弱化.在绝对差异图中相对较小值的区域加入噪声,双眼视差变化较小,只是质量很差,双眼还是可以融合出立体感.由此,绝对差异图是人的双眼视差的直接体现,是研究 3D 图像立体感最根本的出发点.

### 3.1 质量评价

在 3D 图像质量评价中,沿用 2D 图像质量评价方法:基于感知视觉的客观质量评价方法.将左图像和右图像单独评价,然后根据最后评价结果取均值即为质量评价结果.质量评价框架如图 9 所示,图像质量评价的结果是左右图像的得分均值.

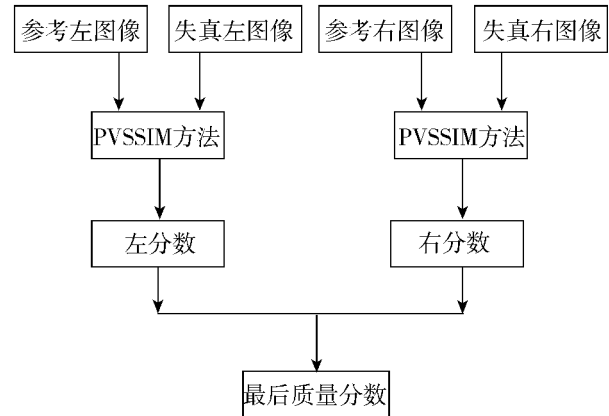


图 9 质量的评价框架

Fig. 9 The framework of images quality assessment

### 3.2 立体感评价

通过研究双眼视差,可得到绝对差异图是产生

立体图像立体感的基本原因.因此,基于左右图像的绝对差异图来评价立体感,本方法也是基于视觉重要性模型的方法.首先,根据立体图像对的左右图像得出立体图像差异图.  $L_o$  和  $R_o$  分别代表参考立体图像对的左图像和右图像,  $L_d$  和  $R_d$  分别代表失真立体图像对的左图像和右图像.则参考图像对的绝对差异图  $X_1 = |L_o - R_o|$ ,失真图像对的绝对差异图  $X_2 = |L_d - R_d|$ .然后将参考图像差异图和失真图像差异图作为立体感的输入图像,根据感知冗余和视觉注意特性融入 SSIM 中进行立体感评价.立体感评价的框架如图 10 所示.

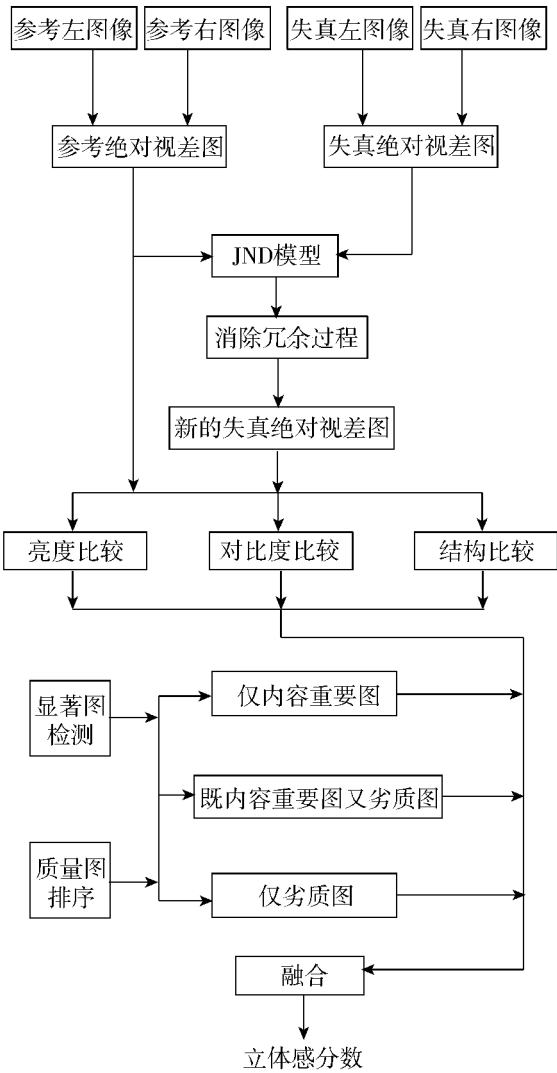


图 10 立体感的评价框架

Fig. 10 The framework of illusion depth assessment

### 3.3 立体图像库建立与实验结果

#### 3.3.1 立体图像库的建立

无失真的立体图像对即为参考立体图像对,可以根据这些图像对得到各种失真类型的图像对,从而建立立体图像库.对在网上搜集的立体图像对进行 4 种失真处理:高斯模糊、高斯白噪声、JPEG、

JPEG2000, 并有3种不同的失真等级. 将高斯模糊、高斯白噪声、JPEG这3种失真类型用Matlab处理, 而JPEG2000用的是Kakadu软件进行处理, 如图11所示.

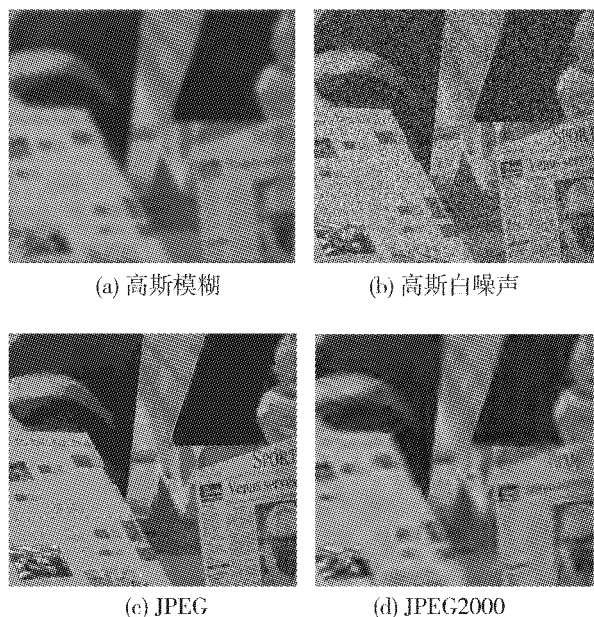


图11 失真类型

Fig. 11 The types of distortion

在立体图像库中, 立体图像对分为左图像和右图像, 因此, 失真种类应该包括左图失真右图失真、左图失真右图不失真、左图不失真右图失真、左图不失真右图不失真这4种情况, 并且每幅图像中的每种失真类型都有不同失真程度的图像. 但为了简便, 本文只对左图不失真右图失真的情况进行了主观评价. 建立的立体图像库一共有180对图像, 各有180个主观的图像质量分数和主观的立体感分数.

为了控制合适的主观评价结果, 主观评价的观看条件、评估者条件、评价的说明指示、测试的过程已经被标准化<sup>[19]</sup>, 各个参数都有标准的定义. 要求测试者都是非专业的观看者, 测试者需要带着偏振眼镜, 共有18个观测者参加了测试. 对测试者进行示范说明, 详细介绍评价方法和评分标准. 该立体图像库的评测包括2个方面: 图像质量分数和立体感分数. 实验采用双刺激评分标准, 即参考图像和失真图像交替显示观看2次, 在显示每幅图像后都会停歇5s, 这样就不会使观察者眼睛产生视觉停留的现象, 从中恢复到正常状态, 最后对评分数据进行处理. 为了获得可靠的数据, 根据测试者的评分的均值、标准差和置信区间来得到最后的主观评分.

### 3.3.2 实验结果

由于将立体图像的主观评价分为质量和立体感来评测, 这与文献[11]中的PSNR方法是相同的, 因此, 与文献[11]的PSNR方法进行实验比较. 质量评

价标准指标的比较结果如表3所示, 立体感评价标准指标的比较结果如表4所示. 从表中的结果可以看出, 本文提出的客观质量评价方法有明显的性能提升.

表3 质量评价指标的结果

Table 3 The results of quality assessment

模 型	PLCC	SRCC	KRCC	RMSE
PSNR <sup>[7]</sup>	0.935	0.915	0.750	3.92
PVSSIM	0.939	0.924	0.768	3.80

表4 立体感评价指标的结果

Table 4 The results of stereoscopic assessment

模 型	PLCC	SRCC	KRCC	RMSE
PSNR <sup>[7]</sup>	0.897	0.859	0.684	0.190
PVSSIM	0.914	0.882	0.714	0.174

本文的质量评价和主观质量分数的散点图如图12所示, 立体感的评价和主观立体分数的散点图如图13所示. 每一个点代表一幅图像, 该点的横坐标代表该图像由各方法得出的主观分数, 纵轴代表该图像的客观分数. 从这2个散点图中可以看出, 本文方法得到的客观分数和主观分数的拟合曲线更接近一条平滑曲线, 这样经过拟合后本文方法得到的图像评价分数就会更接近主观分数.

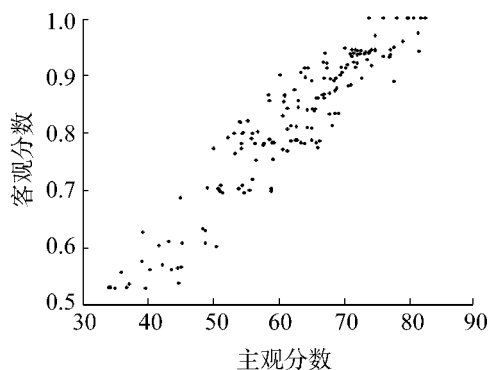


图12 质量评价和主观质量分数的散点

Fig. 12 The scatters of quality assessment and subjective evaluation

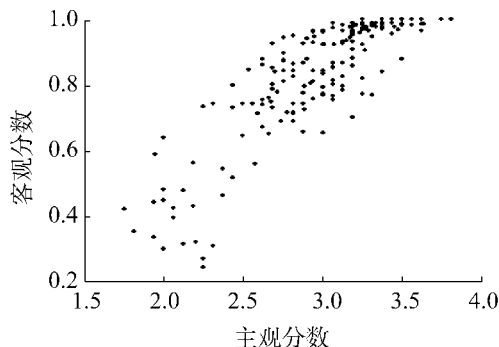


图13 立体感评价与主观立体感分数的散点

Fig. 13 The scatters of illusion depth assessment and subjective evaluation

## 4 结束语

基于人眼视觉冗余和视觉注意特性之间的关系,本文提出了基于感知视觉的立体图像质量评价方法.该立体图像质量评价分为质量评价和立体感评价2种.在质量评价中,使用了基于感知视觉的2D图像质量评价方法,将立体图像对的左右图像分别评价,再取其均值作为最后的客观质量分数.在立体感评价中,通过分析立体感产生的因素得到了基于双眼视差的绝对差异图是立体感最直接的体现,从而通过立体图像对得到参考图像绝对差异图和失真图像绝对差异图,再根据这2个绝对差异图来评价立体感.最后在立体图像数据库上进行验证,结果表明该评价方法在立体图像库中能很好地与主观评价相吻合.

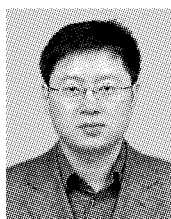
## 参考文献:

- [1] LIPTON L. Foundations of the stereoscopic cinema; a study in depth[M]. New York, USA: Van Nostrand Reinhold, 1982.
- [2] YEMEZ Y, GUDUKBAY U, ZABULIS X, et al. Scene representation technologies for 3DTV: a survey[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(11): 1587-1605.
- [3] 丁剑飞,刘永进. 三维立体显示技术综述[J]. 系统仿真学报, 2008, 20(S1): 132-135.  
DING Jianfei, LIU Yongjin. A survey on three-dimension display technologies [J]. Journal of System Simulation, 2008, 20(Suppl. 1): 132-135.
- [4] SAKAMOTO K, KIMURA R, TAKAKI M. Parallax polarizer barrier stereoscopic 3D display systems[C]//Proceedings of the 2005 International Conference on Active Media Technology. Kagawa, Japan, 2005: 469-474.
- [5] 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.
- [6] WANG Z, SIMONCELLI E P, BOVIK A C. Multi-scale structural similarity for image quality assessment[C]//Proceedings of the 37th IEEE Asilomar Conference on Signals, Systems and Computers. Pacific Grove, USA, 2003: 1398-1402.
- [7] MOORTHY A K, BOVIK A C. Visual importance pooling for image quality assessment[J]. IEEE Journal of Selected Topics in Signal Processing, Special Issue on Visual Media Quality Assessment, 2009, 3(2): 193-201.
- [8] WANG Zhou, LI Qiang. Information content weighting for perceptual image quality assessment[J]. IEEE Transactions on Image Processing, 2011, 20(5): 1185-1198.
- [9] GUO An'an, ZHAO Debin, LIU Shaohui, et al. Visual attention based image quality assessment[C]//18th IEEE International Conference on Image Processing. Brussels, Belgium, 2011: 3297-3300.
- [10] YOU Junyong, XING Liyuan, PERKIS A, et al. Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis[C]//Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics. Scottsdale, USA, 2010: 4033-4036.
- [11] YANG Jiachen, HOU Chunping, ZHOU Yuan, et al. Objective quality assessment method of stereo images[C]//3DTV-Conference: The True Vision Capture, Transmission and Display of 3D Video. Potsdam, Germany, 2009: 1-4.
- [12] HEWAGE C T E R, MARTINI M G. Reduced-reference quality metric for 3D depth map transmission[C]//3DTV-Conference: The True Vision Capture, Transmission and Display of 3D Video. Tampere, Finland, 2010: 1-4.
- [13] XING Liyuan, YOU Junyong, EBRAHIMI T, et al. A perceptual quality metric for stereoscopic crosstalk perception[C]//17th IEEE International Conference on Image Processing. Hong Kong, China, 2010: 4033-4036.
- [14] ROUSHAIN A, SAZZAD P, HORITA Z M, et al. No reference stereoscopic image quality assessment [C]//Proceedings of the SPIE Volume 7524: Stereoscopic Displays and Applications XXI. San Jose, USA, 2010: 17-21.
- [15] KIM D H, MIN D B, OH J H, et al. Depth map quality metric for three-dimensional video[C]//Proceedings of the SPIE Volume 7237: Stereoscopic Displays and Applications XX. San Jose, USA, 2009: 29-34.
- [16] OLSSON R, SJOSTROM M. A depth dependent quality metric for evaluation of coded integral imaging based 3D-images [C]//3DTV-Conference: The True Vision Capture, Transmission and Display of 3D Video. Kos, Greece, 2007: 1-4.
- [17] TAKAHASHI K, NAEMURA T. Unstructured light field rendering using on-the-fly focus measurements[C]//IEEE International Conference on Multimedia and Expo. Amsterdam, The Netherlands, 2005: 205-208.
- [18] 王翔. 数字图像缩放及图像质量评价关键技术研究[D]. 杭州:浙江大学, 2012: 105-126.  
WANG Xiang. Research on image scaling and image quality assessment[D]. Hangzhou: Zhejiang University, 2012: 105-126.
- [19] 张桦. 基于视觉感知的图像质量评价方法研究[D]. 杭州:浙江大学, 2009: 68-88.  
ZHANG Ye. Research on image quality assessment based on visual perception[D]. Hangzhou: Zhejiang University, 2009: 68-88.



- [20] 王科俊, 刘靖宇, 马慧, 等. 手指静脉图像质量评价[J]. 智能系统学报, 2011, 6(4): 324-327.  
WANG Kejun, LIU Jingyu, MA Hui, et al. A finger vein image quality assessment method[J]. CAAI Transactions on Intelligent Systems, 2011, 6(4): 324-327.
- [21] CHOU Chunhsien, LI Yunchin. A perceptually tuned sub-band 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.
- [22] HOU Xiaodi, ZHANG Liqing. Saliency detection: a spectral residual approach[C]//IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA, 2007: 1-8.
- [23] SHEIKH H R, SESHADRINATHAN K, MOORTHY A K, et al. Image and video quality assessment research at LIVE [EB/OL]. (2012-03-12) [2012-06-15]. <http://live.ece.utexas.edu/research/quality/>.

## 作者简介:



姜峰,男,1978年生,副教授,硕士生导师,博士,主要研究方向为计算机视觉、模式识别、视频编码,发表学术论文60余篇。



郭美思,女,1987年生,硕士研究生,主要研究方向计算机视觉、视频编码。



刘绍辉,男,1977年生,副教授,硕士生导师,博士,主要研究方向为计算机视觉、视频编解码,发表学术论文60余篇。



赵德斌,男,1963年生,教授,博士生导师,主要研究方向为视频编码、图像处理和视频传输。作为负责人完成和在研国家“973计划”项目、国家“863”计划项目、国家自然科学基金重点项目与面上项目、国际合作项目等科研项目

19项。共获得国家及省部级奖励13项,其中国家技术发明二等奖1项,国家科学技术进步二等奖2项,省部级科学技术进步一、二、三等奖共10项,发明专利14项。发表学术论文270余篇,其中被SCI检索62篇、EI检索202篇。

## 《应用科技》征订启事

《应用科技》杂志创刊于1974年,双月刊,工业和信息化部主管,哈尔滨工程大学主办,国内外公开发行。该刊集科学性、前瞻性、实用性于一体,以高等院校、科研院所的科研人员以及致力于科技开发、成果应用的社会各界人士为读者对象,主要栏目涉及现代电子技术、计算机应用、自动化技术、机电工程、建筑工程、材料与化学等学科。

《应用科技》杂志为双月刊,大16开本,国内统一刊号:CN23-1191/U;邮发代号:14-60。

全国各地邮局均可订阅,也可以直接与本社联系邮购。

每期定价10.00元,全年定价60.00元。

通信地址:(150001)哈尔滨市南岗区南通大街145号1号楼《应用科技》编辑部

电 话:0451-82534001

期刊网址:<http://yykj.hrbeu.edu.cn>