



(12)发明专利申请

(10)申请公布号 CN 105959684 A

(43)申请公布日 2016.09.21

(21)申请号 201610363321.7

(22)申请日 2016.05.26

(71)申请人 天津大学

地址 300072 天津市南开区卫津路92号

(72)发明人 李素梅 丁学东 刘富岩 侯春萍

(74)专利代理机构 天津市北洋有限责任专利代理事务所 12201

代理人 刘国威

(51)Int.Cl.

H04N 17/00(2006.01)

H04N 13/00(2006.01)

权利要求书2页 说明书16页 附图7页

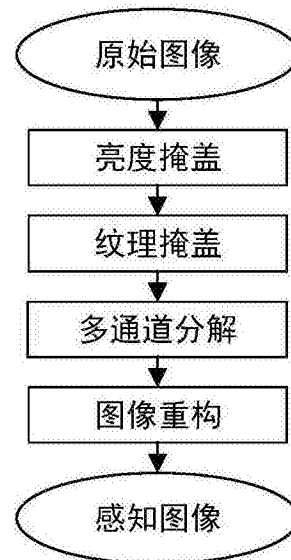
(54)发明名称

基于双目融合的立体图像质量评价方法

(57)摘要

本发明属于视频和图像处理领域,为提出一种符合人类视觉系统相关特性且更加有效的立体图像客观质量评价方法。使新的立体图像客观质量评价方法更加准确高效的评价立体图像的质量,同时在一定程度上推动立体成像技术的发展。本发明采用的技术方案是,基于双目融合的立体图像质量评价方法,包括以下步骤:步骤1:构建符合人类视觉特性的单目感知模型;步骤2:计算由步骤1得到的感知图像 $X_v$ 的图像失真度 $Q_D1$ )计算左右视图图像子块的结构相似度SSIM2)构建最终的图像失真度 $Q_D$ 结合人眼立体视觉双目融合特性和视觉中心显著性特性,对上一步所得的图像子块结构相似度指标进行加权求和计算最终的图像失真度评价分数 $Q_D$ 。本发明主要应用于视频和图像处理。

CN 105959684 A



1. 一种基于双目融合的立体图像质量评价方法,其特征是,步骤如下:

步骤1:构建符合人类视觉特性的单目感知模型,该模型由亮度掩盖模块、纹理掩盖模块和多通道分解模块组成,原始图像 $X_0$ 经过单目感知模型转变成为感知图像 $X_v$ ;

步骤2:计算由步骤1得到的感知图像 $X_v$ 的图像失真度 $Q_b$

1)计算左右视图图像子块的结构相似度SSIM

首先分别将原始立体图像和失真立体图像的左右视图分别分为 $8 \times 8$ 的图像子块,然后对每一个图像子块计算相应的结构相似度 $SSIM(X_i, Y_i)$ ,其中 $X_i, Y_i$ 分别代表原始立体图像和失真立体图像的第 $i$ 个图像子块;

2)构建最终的图像失真度 $Q_b$

结合人眼立体视觉双目融合特性和视觉中心显著性特性,对上一步所得的图像子块结构相似度指标进行加权求和计算最终的图像失真度评价分数 $Q_b$ 。

2. 如权利要求1所述的基于双目融合的立体图像质量评价方法,其特征是,原始图像 $X_0$ 经过单目感知模型转变成为感知图像 $X_v$ ,具体步骤是:

1)原始图像 $X_0$ 经过亮度掩盖模块成为亮度掩模图像 $X_1$

首先计算原始图像 $X_0$ 中每个像素点的亮度掩盖门限 $T_1(x, y)$ 和背景亮度 $I_0(x, y)$ ;然后计算原始图像 $X_0$ 中每个像素点的亮度 $I(x, y)$ 与背景亮度 $I_0(x, y)$ 的差值 $d_1(x, y)$ ,如果 $d_1(x, y)$ 小于掩盖门限 $T_1(x, y)$ ,则表明人眼无法从背景中准确识别到该像素点,因此将该像素点的亮度设置为该点背景亮度值;如果 $d_1(x, y)$ 大于该像素点掩盖门限 $T_1(x, y)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点亮度值不变;重复上述计算过程直到整幅图像计算完成,最后就得到了亮度掩模图像 $X_1$ ;

2)亮度掩模图像 $X_1$ 经过纹理掩盖模块成为纹理掩模图像 $X_2$

首先分别将参考图像和失真图像的亮度掩模图像分为 $8 \times 8$ 的图像子块;然后计算每个图像子块的纹理掩盖门限值 $T_1(\mu)$ ;最后依据所获得的纹理掩盖门限值计算失真图像中每个图像子块的纹理掩模图像子块,从而获得完整的失真图像纹理掩模图像 $X_2$ ;

3)纹理掩模图像 $X_2$ 经过多通道模块成为最终的感知图像 $X_v$

首先,利用Le Gall 5/3小波对纹理掩模图像 $X_2$ 进行三级小波分解得到1个低频分量系数矩阵 $cA(p)$ 和9个高频系数矩阵,分别为垂直方向高频细节分量矩阵 $cV_\lambda(p)$ ,水平方向高频细节分量矩阵 $cH_\lambda(p)$ 和对角方向高频细节分量矩阵 $cD_\lambda(p)$ ,其中 $\lambda$ 为小波分解级数, $\lambda = \{1, 2, 3\}$ ;然后,计算各通道对比敏感度权重 $\omega$ ;最后,使用各通道对比度敏感度权重 $\omega$ 对高频系数矩阵进行加权重构,从而获得最终的感知图像 $X_v$ 。

3. 如权利要求2所述的基于双目融合的立体图像质量评价方法,其特征是,步骤2)亮度掩模图像 $X_1$ 经过纹理掩盖模块成为纹理掩模图像 $X_2$ 进一步地具体过程为:计算失真图像第 $i$ 个子块中像素点 $(x, y)$ 处的像素值 $\tilde{I}^i(x, y)$ 与参考图像中对应图像块像素值 $I^i(x, y)$ 的差值 $d_i^i(x, y)$ ,如果 $d_i^i(x, y)$ 小于该图像块纹理掩盖门限 $T_i^i(\mu)$ ,则表明人眼无法感受到失真图像与参考图像对应像素点的差异,因此将失真图像该像素点的像素值设置为参考图像对应点的像素值;如果差值 $d_i^i(x, y)$ 大于该区域掩盖门限 $T_i^i(\mu)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点像素值不变,依据以上计算过程对失真图像中每个子块每个像素点进行计算从而获得完整的失真图像纹理掩模图像 $X_2$ 。

4. 如权利要求1所述的基于双目融合的立体图像质量评价方法,其特征是,加权求和中权重因子的构建步骤是:

①视觉显著性权重因子

采用图像中心点的各向异性高斯核函数计算不同图像子块的权重,进而模拟视觉心理学的中央偏移特性;

②增益权重因子

使用增益控制理论构建增益权重因子,模拟双目融合特性。

## 基于双目融合的立体图像质量评价方法

### 技术领域

[0001] 本发明属于视频和图像处理领域,涉及立体图像质量评价方法改进优化,尤其是涉及人类视觉系统中的掩盖效应、多通道效应和对比度敏感度在立体图像质量客观评价中的应用。

### 背景技术

[0002] 立体图像技术是一种能够提供视觉深度感的新型多媒体技术,与平面图像相比,立体图像主要利用双目视差原理,通过双目分别接收左右两个通道的图像,给观看者带来较好的临场感,但同时也需要更多的存储空间和更大的传输带宽。因此,对立体图像进行高效压缩是降低存储成本、缓解网络带宽压力的关键技术。但通过压缩编码减少立体图像的数据量,势必会引起各种降质。除此之外,在立体图像采集、处理及传输过程中,都难免会引起失真。因此,探索影响立体图像质量的主要因素,建立一个有效可靠的立体图像质量评价模型对立体信息行业的发展尤为重要。立体图像评价系统不仅可以评价接收端立体图像质量的优劣,还可以为立体图像压缩、传输及其他处理方法提供新的理论依据,从而促进立体图像压缩算法和传输方式的优化。

[0003] 立体图像质量评价可分为主观评价和客观评价。由于主观评价方法对实验环境、测试流程、被试者素质等有较高要求,所以完成一次主观评价既耗时又耗力,且结果的精确度易受外界条件影响,结果的可移植性和稳定性难以达到应用要求。客观评价方法较主观评价方法快速、稳定且简单易行。因此,研究合理且高效的客观评价方法意义重大。

[0004] 目前,在立体图像质量的客观评价方面,经典且常用的方法有基于统计的峰值信噪比<sup>[1]</sup>、结构相似度方法及其衍生出的一系列方法<sup>[2]</sup>。文献[3]在参考平面图像质量评价的基础上,将工程学评价方法峰值信噪比与结构相似度相结合,使用两种方法分别评价立体图像左右视图的质量,然后通过四种不同的方法计算得到绝对差值信息用于评价立体感,最后分别使用局部结合和全局结合的方法将图像质量与立体感质量整合为统一的立体图像质量指标。最后将所得分数取平均值,作为评价立体图像质量的指标。该文章还验证了单纯的平面图像客观评价方法无法简单地适用立体图像质量评价,在立体图像质量评价过程中需要考虑立体感的因素。文献[4]将立体图像中的一些特性,如深度特性,结构特性和梯度特性等。与其他客观图像质量评价指标相比,该文献提出的方法与主观评价有较高的一致性。文献[5]提出了一种双目感知质量模型,首先基于双目不对称性分割立体图像,然后对不同区域设置不同的感知权重,最后计算立体图像质量,该文献证明了结合双目视觉不对称性可以提高立体图像客观评价准确性。文献[6]认为,人眼视觉系统对立体图像的边缘信息极其敏感,因此其考虑通过边缘信息对经典结构相似度指标加以改进,提出了一种基于边缘的结构相似度评价方法,使用该方法评价立体图像左右视图质量。然后作者通过基于自适应权重匹配算法计算左右视图的视差图,通过判断失真图像视差图与参考图像视差图的差异计算立体图像的立体感指标。最后将左右视图质量与立体感质量进行拟合,得到评价立体图像质量的综合指标。文献[7]参考2D平面图像的评价方法,从质量属性、质量倾

向、图像物理特征和计数变量四方面计算立体图像左右视点图像质量,然后计算两者均值作为立体图像整体质量。文献[8]分析了高斯噪声对立体图像的影响,作者发现,对于受到高斯噪声影响的图像,PSNR的评价效果比SSIM评价效果更好,这是因为作为一种统计学方法,PSNR能够准确的判断图像加性噪声失真。因此,作者使用PSNR作为评价指标对立体图像对的绝对差值图进行评价,判断立体图像立体感的降质程度。文献[9]在立体图像主观实验的基础上对立体图像串扰进行分析,确定了影响图像串扰的因素,其中相机基线、串扰等级、图像对比度以及应变效应会增强图像的串扰程度,而立体图像的纹理特征,图像场景细节信息等会降低立体图像串扰。因此作者将立体图像的深度图与结构相似度方法相结合,提出了一种基于立体图像串扰的客观评价方法。作者在文献[10]中对文献[9]的评价方法做出了改进,通过立体图像的视差图对串扰程度做进行评判。在试验中作者发现,使用视差图评价串扰等级要比使用深度图效果更好。文献[11]根据人眼视觉的心理学特性,将立体图像进行分割成不相关区域、双目融合区域以及双目抑制区域,提取各区域中局部相位和局部幅值特征,将两者整合为立体图像评价指标。最后作者结合双眼恰可察觉失真对评价指标进行优化,从而判断立体图像质量。文献[12]提出了一种基于感知图像质量的立体图像质量评价方法。首先参考人眼的视觉显著性特性,将图像分为显著部分和非显著部分,然后对各部分按照失真程度再次进行划分,分为失真部分和非失真部分。接着根据人眼视觉特性对各部分的重要程度进行建模,以此为权重计算各区域的相位相似性。最后将各部分的评价结果进行加权平均,作为评价立体图像质量的指标。文献[13]提出一种立体图像质量评价模型,首先通过最小可察觉失真和图像显著性将被试图像分成四种区域,对不同区域设置不同的权重,计算单张图像结构相似度指标,最后将左右视图结构相似度取平均,作为立体图像质量指标。文献[14]提出了一种基于人类视觉系统的质量评价方法,包括使用Weber-Fechner定律模拟亮度幅度非线性特性,使用可控金字塔模型模拟多通道效应,使用对比度敏感度函数对各频带图像进行感知加权,最后计算各频带结构相似度,作为评价立体图像质量的指标。文献[15]提出了一种基于人眼视觉系统的基本特性的立体图像质量评价方法。该方法首先对立体图像左右视图进行色彩空间变换,提取亮度信息,然后通过感知分解将图像分为多个子块,计算每个子块的对比度增益权重,最后将各部分失真程度进行综合,计算最终的质量评价指标。同时作者提取立体图像视差图,以视差图作为输入按照上述过程评价视差图质量,作为立体图像立体感的评价指标。文献[16]构建了感知图像失真度模型,将左右视图结构相似度、亮度平均值和匹配点个数比与峰值信噪比进行综合,得到最终的立体图像质量分数。文献[17]基于双目立体图像感知特性提出立体图像评价方法,根据局部信息为每个视点定义加权因子,并且在评价过程中考虑人眼视觉系统灵敏度的特性,根据双目显著性差异调整不同区域的质量分数。文献[18]针对传统的统计学方法SNR因没有考虑人类视觉特性导致其不能真正或全面地表征图像质量的缺陷,提出了一种基于人类视觉特性的信噪比方法。通过模拟人类视觉系统中亮度非线性特性、多通道特性、对比度敏感度特性和掩盖效应等特性,对人眼视觉系统进行建模,并重构了经过得到通过了人眼视觉系统模型的参考图像与失真图像,从而改进了传统的统计学方法SNR,使之适用于立体图像质量评价。文献[19]提出了一种彩色图像质量评价方法,对立体图像质量评价提出新思路。

[0005] 上述各种客观评价方法通过不同方式对立体图像质量评价进行研究,得到了较好

的评价结果。然而,大部分方法都是<sup>[3,6,7]</sup>通过分别计算左右视图的质量分数,然后平均加权或是按某种规则进行加权,作为评价立体图像质量的指标。但是人眼在捕捉立体图像时是根据视觉的双目融合特性将左右视点进行融合形成单一的像,因而仍然不能取得较为理想的评价效果。

## 发明内容

[0006] 为克服现有技术的不足,提出一种符合人类视觉系统相关特性且更加有效的立体图像客观质量评价方法。使新的立体图像客观质量评价方法更加准确高效的评价立体图像的质量,同时在一定程度上推动立体成像技术的发展。本发明采用的技术方案是,基于双目融合的立体图像质量评价方法,包括以下步骤:

[0007] 步骤1:构建符合人类视觉特性的单目感知模型,该模型由亮度掩盖模块、纹理掩盖模块和多通道分解模块组成,原始图像 $X_0$ 经过单目感知模型转变成为感知图像 $X_v$ ;

[0008] 步骤2:计算由步骤1得到的感知图像 $X_v$ 的图像失真度 $Q_D$

[0009] 1)计算左右视图图像子块的结构相似度SSIM

[0010] 首先分别将原始立体图像和失真立体图像的左右视图分别分为 $8 \times 8$ 的图像子块,然后对每一个图像子块计算相应的结构相似度 $SSIM(X_i, Y_i)$ ,其中 $X_i, Y_i$ 分别代表原始立体图像和失真立体图像的第 $i$ 个图像子块;

[0011] 2)构建最终的图像失真度 $Q_D$

[0012] 结合人眼立体视觉双目融合特性和视觉中心显著性特性,对上一步所得的图像子块结构相似度指标进行加权求和计算最终的图像失真度评价分数 $Q_D$ 。

[0013] 原始图像 $X_0$ 经过单目感知模型转变成为感知图像 $X_v$ ,具体步骤是:

[0014] 1)原始图像 $X_0$ 经过亮度掩盖模块成为亮度掩模图像 $X_1$

[0015] 首先计算原始图像 $X_0$ 中每个像素点的亮度掩盖门限 $T_1(x, y)$ 和背景亮度 $I_0(x, y)$ ;然后计算原始图像 $X_0$ 中每个像素点的亮度 $I(x, y)$ 与背景亮度 $I_0(x, y)$ 的差值 $d_1(x, y)$ ,如果 $d_1(x, y)$ 小于掩盖门限 $T_1(x, y)$ ,则表明人眼无法从背景中准确识别到该像素点,因此将该像素点的亮度设置为该点背景亮度值;如果 $d_1(x, y)$ 大于该像素点掩盖门限 $T_1(x, y)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点亮度值不变;重复上述计算过程直到整幅图像计算完成,最后就得到了亮度掩模图像 $X_1$ ;

[0016] 2)亮度掩模图像 $X_1$ 经过纹理掩盖模块成为纹理掩模图像 $X_2$

[0017] 首先分别将参考图像和失真图像的亮度掩模图像分为 $8 \times 8$ 的图像子块;然后计算每个图像子块的纹理掩盖门限值 $T_1(\mu)$ ;最后依据所获得的纹理掩盖门限值计算失真图像中每个图像子块的纹理掩模图像子块,从而获得完整的失真图像纹理掩模图像 $X_2$ ;

[0018] 3)纹理掩模图像 $X_2$ 经过多通道模块成为最终的感知图像 $X_v$

[0019] 首先,利用Le Gall 5/3小波对纹理掩模图像 $X_2$ 进行三级小波分解得到1个低频分量系数矩阵 $c_A(p)$ 和9个高频系数矩阵,分别为垂直方向高频细节分量矩阵 $c_{V\lambda}(p)$ ,水平方向高频细节分量矩阵 $c_{H\lambda}(p)$ 和对角方向高频细节分量矩阵 $c_{D\lambda}(p)$ ,其中 $\lambda$ 为小波分解级数, $\lambda = \{1, 2, 3\}$ ;然后,计算各通道对比敏感度权重 $\omega$ ;最后,使用各通道对比度敏感度权重 $\omega$ 对高频系数矩阵进行加权重构,从而获得最终的感知图像 $X_v$ 。

[0020] 步骤2)亮度掩模图像 $X_1$ 经过纹理掩盖模块成为纹理掩模图像 $X_2$ 进一步地具体过程

为:计算失真图像第*i*个子块中像素点(*x*,*y*)处的像素值 $\tilde{I}^i(x,y)$ 与参考图像中对应图像块像素值 $I^i(x,y)$ 的差值 $d_i^i(x,y)$ ,如果 $d_i^i(x,y)$ 小于该图像块纹理掩盖门限 $T_i^i(\mu)$ ,则表明人眼无法感受到失真图像与参考图像对应像素点的差异,因此将失真图像该像素点的像素值设置为参考图像对应点的像素值;如果差值 $d_i^i(x,y)$ 大于该区域掩盖门限 $T_i^i(\mu)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点像素值不变,依据以上计算过程对失真图像中每个子块每个像素点进行计算从而获得完整的失真图像纹理掩模图像 $X_2$ 。

[0021] 加权求和中权重因子的构建步骤是:

[0022] ①视觉显著性权重因子

[0023] 采用图像中心点的各向异性高斯核函数计算不同图像子块的权重,进而模拟视觉心理学的中央偏移特性;

[0024] ②增益权重因子

[0025] 使用增益控制理论构建增益权重因子,模拟双目融合特性。

[0026] 本发明的特点及有益效果是:

[0027] 本立体图像质量客观评价方法由于考虑了人眼单目视觉中掩盖效应、多通道效应、对比度敏感度效应等视觉特性,并与人眼双目融合特性结合,因此对模糊图像的评价表现出了较好的主客观一致性。实验结果说明,本方法的准确性和可靠性优于目前常见的其它立体图像质量客观评价方法。

#### 附图说明:

[0028] 图1参考实验素材图,图中:

[0029] (a)girl (b)tree2

[0030] (c)flower (d)river

[0031] (e)family (f)boy

[0032] (g)mountain (h)ox

[0033] (i)river2 (j)TJUsquare

[0034] (k)tree3 (l)woman。

[0035] 图2失真立体图像,图中:

[0036] (a)压缩率为8% (b)压缩率为2%

[0037] (c)均值为0,方差为0.02的高斯白噪声 (d)均值为0,方差为0.06的高斯白噪声

[0038] (e)模板为[15,15],sigma为1.5的高斯模糊 (f)模板为[15,15],sigma为3.8的高斯模糊

[0039] 图3单目感知模型框图。

[0040] 图4亮度掩盖优化前后的girl左视点参考图像,图中:

[0041] (a)优化前 (b)优化后。

[0042] 图5 girl纹理掩盖效果示意图,图中:

[0043] (a)参考图像 (b)失真图像。

[0044] 图6被测试者在观看点H的观看角度。

[0045] 图7优化前后的girl失真左视点效果图。图中:

[0046] (a)优化前 (b)优化后

- [0047] 图8小波行变换和列变换。
- [0048] 图9 girl参考左视点图的灰度图与二级小波分解高频系数矩阵,图中:
- [0049] (a)参考图像 (b)水平方向
- [0050] (c)垂直方向 (d)对角方向。
- [0051] 图10重构的参考图像图
- [0052] 图11立体图像失真度框图
- [0053] 图12五种失真度客观评价方法散点图
- [0054] (a)SSIM (b)PSNR
- [0055] (c)文献[18] (d)文献[30]
- [0056] (e)本文算法。

### 具体实施方式

[0057] 本发明在衡量图像失真程度时采用双目融合的特性将左右视点质量进行融合得到立体图像的失真度,实验结果表明本文方法具有较好的准确性。

[0058] 本发明采用的技术方案是,基于双目融合的立体图像客观质量评价方法,包括以下步骤:

[0059] 步骤1:构建符合人类视觉特性的单目感知模型,该模型由亮度掩盖模块、纹理掩盖模块和多通道分解模块组成。原始图像 $X_0$ 经过单目感知模型成为感知图像 $X_v$ ,具体包括以下内容:

[0060] 1.原始图像 $X_0$ 经过亮度掩盖模块成为亮度掩模图像 $X_1$

[0061] 首先计算原始图像 $X_0$ 中每个像素点的亮度掩盖门限 $T_1(x, y)$ 和背景亮度 $I_0(x, y)$ ;然后计算原始图像 $X_0$ 中每个像素点的亮度 $I(x, y)$ 与背景亮度 $I_0(x, y)$ 的差值 $d_1(x, y)$ ,如果 $d_1(x, y)$ 小于掩盖门限 $T_1(x, y)$ ,则表明人眼无法从背景中准确识别到该像素点,因此将该像素点的亮度设置为该点背景亮度值;如果 $d_1(x, y)$ 大于该像素点掩盖门限 $T_1(x, y)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点亮度值不变;重复上述计算过程直到整幅图像计算完成,最后就得到了亮度掩模图像 $X_1$ 。

[0062] 2.亮度掩模图像 $X_1$ 经过纹理掩盖模块成为纹理掩模图像 $X_2$

[0063] 首先分别将参考图像和失真图像的亮度掩模图像分为 $8 \times 8$ 的图像子块;然后计算每个图像子块的纹理掩盖门限值 $T_1(\mu)$ ;最后依据所获得的纹理掩盖门限值计算失真图像中每个图像子块的纹理掩模图像子块,从而获得完整的失真图像纹理掩模图像 $X_2$ 。具体过程为:计算失真图像第 $i$ 个子块中像素点 $(x, y)$ 处的像素值 $\tilde{I}^i(x, y)$ 与参考图像中对应图像块像素值 $I^i(x, y)$ 的差值 $d_i^i(x, y)$ ,如果 $d_i^i(x, y)$ 小于该图像块纹理掩盖门限 $T_i^i(\mu)$ ,则表明人眼无法感受到失真图像与参考图像对应像素点的差异,因此将失真图像该像素点的像素值设置为参考图像对应点的像素值;如果差值 $d_i^i(x, y)$ 大于该区域掩盖门限 $T_i^i(\mu)$ ,则表明人眼能够正确识别出该像素点,因此保持该像素点像素值不变,依据以上计算过程对失真图像中每个子块每个像素点进行计算从而获得完整的失真图像纹理掩模图像 $X_2$

[0064] 3.纹理掩模图像 $X_2$ 经过多通道模块成为最终的感知图像 $X_v$

[0065] 首先,利用Le Gall 5/3小波对纹理掩模图像 $X_2$ 进行三级小波分解得到1个低频分



量系数矩阵 $cA(p)$ 和9个高频系数矩阵,分别为垂直方向高频细节分量矩阵 $cV_{\lambda}(p)$ ,水平方向高频细节分量矩阵 $cH_{\lambda}(p)$ 和对角方向高频细节分量矩阵 $cD_{\lambda}(p)$ ,其中 $\lambda$ 为小波分解级数, $\lambda = \{1, 2, 3\}$ ;然后,计算各通道对比敏感度权重 $\omega$ ;最后,使用各通道对比度敏感度权重 $\omega$ 对高频系数矩阵进行加权重构,从而获得最终的感知图像 $X_v$ 。

[0066] 步骤2:计算由步骤1得到的感知图像 $X_v$ 的图像失真度 $Q_d$ 。

[0067] 1. 计算左右视图图像子块的结构相似度SSIM

[0068] 首先分别将原始立体图像和失真立体图像的左右视图分别分为 $8 \times 8$ 的图像子块,然后对每一个图像子块计算相应的结构相似度 $SSIM(X_i, Y_i)$ ,其中 $X_i, Y_i$ 分别代表原始立体图像和失真立体图像的第 $i$ 个图像子块。

[0069] 2. 构建最终的图像失真度 $Q_d$

[0070] 本方法结合人眼立体视觉双目融合特性和视觉中心显著性特性,对上一步所得的图像子块结构相似度指标进行加权求和计算最终的图像失真度评价分数 $Q_d$ 。其中,权重由两部分构成:视觉显著性权重因子和增益权重因子。

[0071] 权重因子的构建:

[0072] ①视觉显著性权重因子

[0073] 本方法采用图像中心点的各向异性高斯核函数计算不同图像子块的权重,进而模拟视觉心理学的中央偏移特性;

[0074] ②增益权重因子

[0075] 增益控制理论准确的描述了双目融合过程并能够同时反映Fechner悖论和单目感知理论。因此本方法使用增益控制理论构建增益权重因子,模拟双目融合特性。

[0076] 本发明所使用的立体图像参考素材均取自天津大学宽带无线通信与立体成像研究所提供的立体图像信息库。立体图像库包含未进行任何失真的参考图像共12幅,包括人物、近景、远景等图像内容,分别为标准立体素材girl、tree2、flower、river、family、boy、mountain、ox、river2、TJUsquare、tree3、woman,图像的分辨率为 $1280 \times 1024$ 。本设计所用的显示设备需要水平翻转立体图像右视点图,才能正确显示立体图像。因此实验素材中右视点图为镜像放置的,如图1所示。此外,在参考图像的基础上,本文模拟在实际生活中立体图像可能遭受的各种失真,对立体图像数据源进行了不同类型的降质,包括JPEG压缩、高斯白噪声与高斯模糊,每种失真共得到待测立体图像180幅,共计540幅图像。以实验素材girl为例,几种不同程度不同失真类型的失真立体图像如图2所示。对于每一张失真图像,数据库记录着立体图像失真度分值。

[0077] 下面结合具体实例进一步详细说明本方法:

[0078] 本方法首先构建人眼单目视觉感知模型,模拟人眼视觉系统中掩盖效应、多通道效应和对比度敏感度等三大特性;然后对人眼单目视觉感知模型的输出结果进行立体图像失真程度的客观评价。具体步骤如下:

[0079] 1建立单目感知模型

[0080] 在采集图像信号的过程中人眼是最终的接收者,无论在立体图像质量评价还是在平面图像质量评价过程中,人眼的单目视觉特性都发挥着重要的作用。因此为了获取人眼实际感知到的立体图像,本方法提出了一种单目感知模型,模拟人眼采集图像信号的一系列过程,并最终计算人眼实际感受到的图像信号。本模型共包含掩盖效应、多通道效应与图

像重构三部分,如图3所示。

[0081] 1.1掩盖效应

[0082] 在数字图像处理过程中,当多种激励共同刺激人眼时,一种激励的视觉感知阈值会因受到其他激励干扰而发生变化。这种现象就称为视觉掩盖效应<sup>[20]</sup>(Visual Masking Effect)。掩盖效应是人眼视觉系统中的一个重要组成部分,但是因为人眼结构的复杂性,现阶段尚不能系统全面的定义视觉掩盖效应。因此本方法选取空间域掩盖效应中较为典型的亮度掩盖和纹理掩盖。

[0083] 1.1.1亮度掩盖

[0084] 亮度掩盖是一种空间域掩盖效应,其体现了图像背景像素的平均亮度对人眼视觉系统可见门限值的影响。当像素点背景区域亮度改变时,视觉感知最低门限会发生相应的变化,从而视觉敏感程度也会随之发生变化。因此,本方法构建亮度掩盖门限,使用该门限值计算参考图像和失真图像中每个像素点是否能被人眼感知。

[0085] 亮度掩盖的理论最早源于Weber实验,Yang在Weber实验基础上,提出了空间域NAMM模型<sup>[21]</sup>(Nonlinear Additivity model for Masking)。在该模型中对亮度掩盖做出了准确定义,如式(1)所示,

$$[0086] \quad T_l(x, y) = \begin{cases} 17(1 - \sqrt{\frac{I_0(x, y)}{127}}) + 3, I_0(x, y) \leq 127 \\ \frac{3}{128}(I_0(x, y) - 127) + 3, I_0(x, y) \geq 127 \end{cases} \quad (1)$$

[0087] 其中 $T_l(x, y)$ 为像素点 $(x, y)$ 处的亮度掩盖门限, $I_0(x, y)$ 为图像在像素点 $(x, y)$ 处的背景亮度, $I_0(x, y) = \sum_{i=-2}^2 \sum_{j=-2}^2 \omega_{i,j} I(x+i, y+j)$ , $I(x, y)$ 为图像在像素点 $(x, y)$ 处的亮度值,

$\omega_{i,j}$ 为背景亮度加权矩阵,如式(2)所示

$$[0088] \quad \omega_{i,j} = \begin{bmatrix} 0.0256 & 0.0324 & 0.0362 & 0.0324 & 0.0256 \\ 0.0324 & 0.0512 & 0.0724 & 0.0512 & 0.0324 \\ 0.0362 & 0.0724 & 0.0000 & 0.0724 & 0.0362 \\ 0.0324 & 0.0512 & 0.0724 & 0.0512 & 0.0324 \\ 0.0256 & 0.0324 & 0.0362 & 0.0324 & 0.0256 \end{bmatrix} \quad (2)$$

[0089] 然后根据各个像素点的掩盖门限 $T_l(x, y)$ 对参考图像和失真图像进行优化,具体优化方法为:如果图像像素点 $(x, y)$ 处的亮度 $I(x, y)$ 与背景亮度 $I_0(x, y)$ 的差值 $d_l(x, y)$ 小于掩盖门限 $T_l(x, y)$ ,则表明人眼无法从背景中准确识别到该像素点,因此将该像素点的亮度值设置为该点背景亮度值;如果差值 $d_l(x, y)$ 大于该像素点掩盖门限 $T_l(x, y)$ ,则表明人眼能够正确识别出该像素点,保持该像素点亮度值不变。计算公式如(3)所示

$$[0090] \quad d_l(x, y) = |I(x, y) - I_0(x, y)|$$

$$I_l(x, y) = \begin{cases} I_0(x, y) & d_l(x, y) < T_l(x, y) \\ I(x, y) & d_l(x, y) > T_l(x, y) \end{cases} \quad (3)$$

[0091] 以girl左视点参考图像为例,经过亮度掩盖优化后,优化图像如图4所示,经过亮度掩盖优化后,原始girl参考图像中人眼无法感知的像素点被其背景亮度值所代替。

[0092] 1.1.2纹理掩盖

[0093] 纹理掩盖表示的是视觉阈值受到了图像内容的影响<sup>[22]</sup>。它表明人眼感知图像纹理复杂区域的失真要比感知平滑区域失真的难度大得多。纹理复杂区域是指图像亮度变化较大的区域,平滑区是指亮度值变化程度相对比较平缓的区域。图5为纹理掩盖示意图,其中5(a)为无失真的girl左视点参考图像,5(b)为加入方差为0.016高斯白噪声。观察5(b)可以发现,在属于平滑区域的人脸、柜子、门等部分,人眼可以很轻易的辨别出噪声失真,然而在属于纹理复杂区域的毛毯部分,则难以轻松辨别图像失真,这就是由人眼视觉纹理掩盖效应造成的。因为纹理掩盖描述的是某一图像区域内人眼视觉系统可见门限值会受该区域纹理复杂程度的影响,因此本方法首先将图像分割为 $8 \times 8$ 的图像子块,分别计算每个图像块的纹理掩盖门限值,从而对图像进行优化。

[0094] 对比度敏感度函数(CSF)是通过门限实验得到的,其倒数的物理意义就是该空间频率下人眼分辨能力的阈值。该阈值越低,则表明人眼对此处信号失真的感知越灵敏。因此本方法使用对比度敏感度函数的倒数构建图像纹理掩盖判决门限,使用门限值判断失真图像中失真信号是否被掩盖,从而模拟纹理掩盖效应。

[0095] 在这里需要注意,因为对比度敏感度函数中的输入的频率指的是单位视角内图像周期性变化的程度,因此计算出每个图像块的空间频率后,还应除以被试者在观看过程中的观看角度 $\eta$ ,如图6所示。本文实验使用的立体图像分辨率为 $1280 \times 1024$ ,根据ITU-R BT.1438建议,观看距离 $L_1$ 应为屏幕高度 $L_2$ 的6倍,即 $L_1 = 6L_2$ 。所以,可以按照公式(4)计算图像观看角度 $\eta$ 。

$$[0096] \quad \eta = 2 \arctan\left(\frac{L_2}{2L_1}\right) = 2 \arctan\left(\frac{1}{12}\right) = 4.7636^\circ \quad (4)$$

[0097] 以加入方差为0.016高斯白噪声的girl左视点图像为例,首先将图像分为 $8 \times 8$ 大小的图像块,按照公式(5)分别求取每个图像块的空间频率 $f$ 。

$$[0098] \quad f = \sqrt{f_R^2 + f_C^2} \quad (5)$$

[0099] 其中 $f_R$ 是水平方向空间频率, $f_C$ 为垂直方向的空间频率。水平方向的空间频率表征图像水平方向的细节变化,垂直方向的空间频率表征图像垂直方向的细节变化,计算公式如式(6)、(7)所示。

$$[0100] \quad f_R = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=2}^N [g(i,j) - g(i,j-1)]^2} \quad (6)$$

$$[0101] \quad f_C = \sqrt{\frac{1}{MN} \sum_{j=1}^N \sum_{i=2}^M [g(i,j) - g(i-1,j)]^2} \quad (7)$$

[0102] 其中,图像块大小为 $M \times N$ , $g(i,j)$ 为第 $i$ 行、第 $j$ 列像素的亮度值, $g(i,j-1)$ 为第 $i$ 行、第 $j-1$ 列像素的亮度值, $g(i-1,j)$ 为第 $i-1$ 行、第 $j$ 列像素的亮度值。

[0103] 然后计算每个图像块单位视角范围内的空间频率 $\mu$ ,如式(8)所示

$$[0104] \quad \mu = \frac{f}{\eta} = \frac{f}{4.7636} \quad (\text{cycle/degree}) \quad (8)$$

[0105] 接着使用对比度敏感度函数构造纹理掩盖门限值。在文献[23]中作者详细的介绍

了目前研究领域比较常用的几种对比度敏感度函数。本方法选取其中使用最为广泛的对比度敏感度函数构建纹理掩盖门限值,该函数如式(9)所示。

$$[0106] \quad CSF(\mu) = 2.6(0.192 + 0.114\mu)e^{-(0.114\mu)^{1.8}} \quad (9)$$

[0107] 根据对比度敏感度函数的物理意义可知,对比度敏感度函数的倒数即为纹理掩盖门限 $T_t(\mu)$ 。该函数定义了参考图像与测试图像间,人眼可以观察到的像素偏差的最小值。

$$[0108] \quad T_t(\mu) = \frac{1}{CSF(\mu)} \quad (10)$$

[0109] 最后根据失真图像每个图像块的掩盖门限 $T_1(\mu)$ 对失真图像进行优化,具体优化方法为:按照式(10)计算每个图像块的掩盖门限 $T_i^t(\mu)$ 。如果第 $i$ 个失真图像块中像素点 $(x, y)$ 处的像素值 $\tilde{I}^i(x, y)$ 与参考图像中对应图像块的像素值 $I^i(x, y)$ 的差值 $d_i^t(x, y)$ 小于该图像块纹理掩盖门限 $T_i^t(\mu)$ ,则表明人眼无法感受到失真图像与参考图像对应像素点的差异,因此将失真图像该像素点处的像素值设置为参考图像对应点的像素值;如果差值 $d_i^t(x, y)$ 大于该区域掩盖门限 $T_i^t(\mu)$ ,则表明人眼能够正确识别出该像素点,保持该像素点像素值不变。优化失真图像的过程如式(11)所示。

$$[0110] \quad \begin{aligned} d_i^t(x, y) &= |\tilde{I}^i(x, y) - I^i(x, y)| \\ \tilde{I}_t^i(x, y) &= \begin{cases} I^i(x, y) & d_i^t(x, y) < T_i^t(\mu) \\ \tilde{I}^i(x, y) & d_i^t(x, y) > T_i^t(\mu) \end{cases} \end{aligned} \quad (11)$$

[0111] 以girl的失真图像左视点为例,经过纹理掩盖优化前后的效果如图7所示。

[0112] 1.2多通道分解特性

[0113] 视觉心理学和生理学实验<sup>[24]</sup>表明人类视觉系统存在独立的空间频率处理单元,它通过多个并列的视觉通道来传输观看到的信息,不同通道的处理单元整体上相互独立。这些通道把视觉激励划分到不同的子带进行处理,相当于一组方向性带通滤波器,这些通道会对特定空间频率和方向上的视觉激励做出响应。通常使用低通与带通滤波器来模拟实现

[0114] 小波分析是一种空间频率局部化的分析方法,它具有良好的“空间域-频域”定位能力<sup>[25]</sup>。通过设置不同的滤波器组,小波变换可以极大的减少或消除不同通道之间的相关性,这与人眼视觉系统的多通道效应不谋而合,与此同时,小波变换通过在低频段使用宽分析窗口,在高频段使用窄分析窗口,从而能够准确的提取图像不同频率间的细节信息,这对立体图像质量评价是非常重要的。因此本文选取小波变换,来模拟人眼视觉系统中的多个通道。

[0115] 对于二维平面图像而言,它的一次小波分解是由行变换和列变换两部分组成的,如图8所示。首先经过一次行变换后,图像分成低频矩阵L和低频矩阵H,之后再对所得系数矩阵进行列变换。分别将低频矩阵L和低频矩阵H中的系数进行列变换,从而得到四个系数矩阵,分别为LL、LH、HL、HH。其中,LL表示图像的低频近似分量,其中包含着图像大部分能量,呈现为原始图像的缩略图;LH表示图像在水平方向( $0^\circ$ )高频细节分量,HL表示图像在垂

直方向(90°)高频细节分量,HH表示图像在对角方向(45°、135°)高频细节分量。经过一次完整的小波变换后,所得系数矩阵的尺寸变为分解前的四分之一,即矩阵的行数和列数均减半。对于一次小波分解后的系数矩阵,因为低频系数矩阵中仍然包含着原始图像的大部分能量,所以可以对其再进行小波分解,进一步提取图像的高频信息。重复上述过程,即可以得到经过多级小波分解的系数。

[0116] 根据对比度敏感度理论,人眼视觉系统对于不同频率的图像呈现一种带通特性,即当图像频率过低或过高的情况下,人眼对图像的识别能力都会降低。同时为了降低计算复杂度,本方法选取Le Gall 5/3小波对参考图像和失真图像进行三级小波分解。Le Gall 5/3小波是D.Le Gall与A.Tabatabai在1988年提出的,目前是JPEG 2000无损压缩标准中的核心编码系统默认采用的可逆小波。其滤波器组系数如表1所示。

[0117] 表1 Le Gall 5/3小波滤波器组系数表

i	分解滤波器系数		合成滤波器系数	
	低通滤波器	高通滤波器	低通滤波器	高通滤波器
	Hl	Hh	Gl	Gh
[0118] 0	6/8	1	1	6/8
±1	2/8	-1/2	1/2	-2/8
±2	-1/8	0	0	-1/8
其他值	0	0	0	0

[0119] 图像p经过三级小波分解后,共得到1个低频近似分量系数矩阵cA(p)和9个高频系数矩阵,分别为垂直方向高频细节分量矩阵cV<sub>λ</sub>(p),水平方向高频细节分量矩阵cH<sub>λ</sub>(p)和对角方向高频细节分量矩阵cD<sub>λ</sub>(p),其中λ为小波分解级数,λ={1,2,3}。图9为girl参考图像左视点经过小波分解后所得的系数矩阵图,其中9(a)为girl左视点图的灰度图,图9(b)~(d)为水平、垂直和对角方向高频细节系数矩阵。

[0120] 根据对比度敏感度带通理论,在识别静态图像时,对于空间频率不同的图像人眼对其的感知程度存在差异。因此,可以用对比度敏感度函数对小波分解后得到的各个高频系数矩阵进行加权,从而模拟人眼视觉系统在识别静态图像时的多通道特性。具体做法如下:首先计算各高频系数矩阵空间频率。

[0121] 对于水平方向高频细节系数矩阵cH<sub>λ</sub>(p),因为其描述的是图像在水平方向的细节变化,所以在计算空间频率的过程中,只需要计算水平方向的频率,如式(12)所示。同理垂直方向高频细节系数矩阵cV<sub>λ</sub>(p),因为其描述的是图像在垂直方向的细节变化,所以只需要计算垂直方向的频率,如式(13)所示。对于对角方向的系数矩阵cD<sub>λ</sub>(p),因为其反映的是45°和135°两个方向上的边缘细节信息,因此需要按照式(14)计算空间频率。

$$[0122] \quad f_H^\lambda = f_R^\lambda = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=2}^N [cH_\lambda(i, j) - cH_\lambda(i, j-1)]^2} \quad (12)$$

$$[0123] \quad f_V^\lambda = f_C^\lambda = \sqrt{\frac{1}{MN} \sum_{i=2}^M \sum_{j=1}^N [cV_\lambda(i, j) - cV_\lambda(i-1, j)]^2} \quad (13)$$

$$f_D^\lambda = \sqrt{(f_R^\lambda)^2 + (f_C^\lambda)^2}$$

$$[0124] \quad = \sqrt{\frac{1}{MN} \sum_{i=2}^M \sum_{j=2}^N [cD_\lambda(i, j) - cD_\lambda(i-1, j-1)]^2 + \frac{1}{MN} \sum_{i=2}^M \sum_{j=2}^N [cD_\lambda(i, j-1) - cD_\lambda(i-1, j)]^2} \quad (14)$$

[0125] 其中,  $M$  为系数矩阵行数,  $N$  为系数矩阵列数,  $\lambda$  为小波分解级数,  $\lambda = \{1, 2, 3\}$ ,  $(i, j)$  为所得系数矩阵第  $i$  行、第  $j$  列元素值。

[0126] 然后将各个高频系数矩阵的空间频率与观看角度  $\eta$  的商  $\mu$  作为对比度敏感度函数 CSF( $f$ ) 的输入, 计算各通道对比度敏感度权重  $\omega$ 。

$$[0127] \quad \omega = \text{CSF}(\mu) \quad (15)$$

[0128] 最后使用各通道对比度敏感度权重  $\omega$  对高频系数矩阵进行加权重构, 计算重构后图像  $P$ , 如式(16)所示。其中  $\omega_H^\lambda$ 、 $\omega_V^\lambda$ 、 $\omega_D^\lambda$  分别代表第  $\lambda$  层小波分解后水平、垂直和对角方向权重系数,  $cA(p)$  表示低频近似分量系数矩阵,  $cH_\lambda(p)$ 、 $cV_\lambda(p)$ 、 $cD_\lambda(p)$  分别代表水平、垂直和对角方向高频细节分量系数矩阵。iWavelet 代表小波逆变换。

$$[0129] \quad P = iWavelet(cA(p), \omega_H^\lambda \times cH_\lambda(p), \omega_V^\lambda \times cV_\lambda(p), \omega_D^\lambda \times cD_\lambda(p)), \quad (\lambda = 1, 2, 3) \quad (16)$$

[0130] 重构后的  $girl$  左视点参考图像如图10所示, 该图像即为人眼实际感知到的图像, 所以本方法使用重构后的图像计算失真度。

[0131] 2 立体图像失真度  $Q_D$

[0132] 首先将立体图像进行分块, 计算每个图像块的结构相似度分数; 然后根据人眼视觉显著性特点, 对各图像块分数进行加权; 最后借助人眼的双目融合特性直接求取立体图像整体失真度指标, 从而准确的对立体图像失真度进行评价, 如图11所示。

[0133] 2.1 子块结构相似度

[0134] 本方法先分别将原始立体图像和失真立体图像分为  $8 \times 8$  的图像子块, 然后分别按照公式(17)、(18)、(19)计算亮度比较函数、对比度比较函数和结构比较函数, 其中  $x$ 、 $y$  分别代表参考图像和失真图像对应位置的图像块,  $\mu_x$ 、 $\mu_y$  分别为图像块  $x$ 、 $y$  的均值,  $\sigma_x$ 、 $\sigma_y$  为图像块  $x$ 、 $y$  的方差,  $\sigma_{xy}$  为图像块  $x$ 、 $y$  的协方差。当三个比较函数的分母为零或者接近零时, 会影响评价结果, 于是设置很小的正数  $C_1$ 、 $C_2$  和  $C_3$  来避免这种现象, 本方法中  $C_1 = (K_1 L)^2$ ,  $C_2 = (K_2 L)^2$ ,  $C_3 = C_2 / 2$ ,  $L = 255$ ,  $K_1 = 0.01$ ,  $K_2 = 0.03$ 。

$$[0135] \quad I(x, y) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (17)$$

$$[0136] \quad c(x, y) = \frac{2\sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (18)$$

$$[0137] \quad s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x + \sigma_y + C_3} \quad (19)$$

[0138] 最后将亮度函数  $I(x, y)$ 、对比度函数  $c(x, y)$  与结构比较函数  $s(x, y)$  结合, 得到每个图像块的结构相似度指标  $SSIM(x, y)$ 。其中  $\alpha$ 、 $\beta$ 、 $\gamma$  为权重因子, 本发明一个实例取  $\alpha = \beta = \gamma = 1$ 。

$$[0139] \quad SSIM(x, y) = [I(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (20)$$

[0140] 2.2 双目融合

[0141] 立体图像虽然由两幅平面图像构成,但是人观察立体事物的过程并不是左右两幅图像简单的线性加权,而是左右视图相互影响、共同作用的产物。因此,本方法结合人眼立体视觉双目融合特性和视觉中心显著性特性,对上一步左右视图图像块结构相似度指标进行加权求和计算最终的评价分数。如公式(21)所示。

$$[0142] \quad Q_D = \sum_i \omega_i (g_i^L \times SSIM(V_i^L, V_i^{L'}) + g_i^R \times SSIM(V_i^R, V_i^{R'})) \quad (21)$$

[0143] 其中,  $\omega_i$  为第  $i$  个图像子块的视觉显著性权重因子,  $g_i^L$ 、 $g_i^R$  分别为左右视点图像中第  $i$  个图像子块所对应的增益权重因子,  $V_i^L$ 、 $V_i^R$  分别为参考图像左右视图中第  $i$  个图像块,  $V_i^{L'}$ 、 $V_i^{R'}$  分别为失真图像左右视图中第  $i$  个图像块。 $Q_D$  取值范围为  $[0, 1]$ , 其数值越接近 1, 表明失真图像的失真程度越低。

[0144] 2.2.1 视觉显著性权重因子

[0145] 根据视觉心理学实验,当人类在感知一副图片时,注意力会首先集中在图像的中间部位,从图像的中心点寻找视觉注视点。然后他的注意力范围由中央部位向四周递减。这种现象被称为中央偏移特性(Center Bias, CB)<sup>[26]</sup>。因此,本方法采用图像中心点的各向异性高斯核函数计算不同图像子块的权重<sup>[27]</sup>,进而模拟视觉心理学的中央偏移特性:

$$[0146] \quad CB(x, y) = \exp \left\{ - \left( \frac{(x-x_0)^2}{2\sigma_h^2} + \frac{(y-y_0)^2}{2\sigma_v^2} \right) \right\} \quad (22)$$

[0147] 其中,  $CB(x, y)$  表示图像块  $(x, y)$  对中心图像块  $(x_0, y_0)$  的偏移信息。 $\sigma_h$  和  $\sigma_v$  分别表示图像块水平方向和垂直方向的标准差,本方法取  $\sigma_h = 1/3W$ 、 $\sigma_v = 1/3H$ , 其中  $W$  和  $H$  表示图像的水平方向子块和垂直方向子块数。最后,使用  $CB(x, y)$  作为视觉显著性权重因子,如式(23)所示,  $i$  为图像块  $(x, y)$  的序号。

$$[0148] \quad \omega_i = CB(x, y) \quad (23)$$

[0149] 2.2.2 增益权重因子

[0150] Fechner<sup>[28]</sup> 在深入研究视觉心理学的基础上,提出了著名的Fechner悖论:当左眼右眼感知到的亮度信息存在差异时(如左眼看到亮度值的比较大的图像  $I$ , 右眼看到亮度值比较小的图像  $\delta$ ), 亮度较大的图像在感知过程中会受到亮度较暗图像的影响,从而使人眼对图像  $I$  的感知亮度达不到图像本身的亮度,如式(24)所示,其中  $I$  代表亮度较大的图像,  $\delta$  代表亮度较暗的图像,  $f_B(I, 0)$  表示只看到图像  $I$ , 另一只眼睛不接收图像信号时对图像的感知程度,  $f_B(I, \delta)$  表示一只眼睛观看图像  $I$ , 另一只眼睛观看图像  $\delta$  时对图像的感知程度。

$$[0151] \quad f_B(I, 0) > f_B(I, \delta) \quad (24)$$

[0152] 单目感知理论(cyclopean perception)描述的是生活中一种常见的现象:我们双目同时观看到的事物有时候与闭上一只眼睛看到的事物基本相同。这是因为当双眼同时受到相同等级的强刺激时,通过大脑的融合过程,给人感觉到的是一个强刺激;而当只有一只眼睛受到强刺激而另一只眼没有接受到任何刺激时,大脑会主观的抑制掉没有接收到刺激的眼睛,因此当双眼看到两幅亮度相似的图像信号时,双眼感知到的图像与单眼感知到的图像基本相同,如(25)所示。

$$[0153] \quad f_B(I, 0) \cong f_B(0, I) \cong f_B(I, I) \quad (25)$$

[0154] 在此基础上, Jian Ding<sup>[29]</sup> 提出了一种基于增益控制理论(gain-control

theory)的双目融合模型,该模型通过左右眼对于图像的感知增益权重获取人眼实际感知到的立体图像,如式(26)所示。其中 $I^L$ 、 $I^R$ 代表立体图像左右视图, $E_L$ 、 $E_R$ 分别代表左右视图的图像能量, $g^L$ 、 $g^R$ 为左右视图的增益权重, $f_B(I^L, I^R)$ 代表人眼实际感知到的立体图像。

$$\begin{aligned}
 [0155] \quad f_B(I^L, I^R) &= \left( \frac{1+E_L}{1+E_L+E_R} \right) \times I^L + \left( \frac{1+E_R}{1+E_L+E_R} \right) \times I^R \\
 &= g^L \times I^L + g^R \times I^R
 \end{aligned} \tag{26}$$

[0156] 增益控制理论准确的描述了双目融合过程并能够同时反映Fechner悖论和单目感知理论。因此本方法使用增益控制理论对左右视图各图像块结构相似度指标SSIM进行加权求和,求取平均结构相似度,模拟双目融合特性。

[0157] 3立体图像失真度评价结果与分析

[0158] 本方法使用5个评价指标来判断所提算法的性能。这五个指标分别是皮尔森线性相关系数(Pearson Linear Correlation Coefficient, PLCC)、斯皮尔曼等级相关系数(Spearman's Rank Correlation Coefficient, SRCC)、肯德尔等级相关系数(Kendall's Rank Correlation Coefficient, KRCC)、平均绝对误差(Mean Absolute Error, MAE)和均方根误差(Root-Mean-Square Error, RMSE)。在这五个指标中,PLCC用于评估预测的客观评价结果与主观评价MOS值的线性相关性。SROCC和KROCC用于评估的是客观评价结果与主观评价MOS值的秩相关性,即评价结果的单调性, RMSE和MAE则是用于评判客观评价结果的稳定性。当客观评价结果与主观评价结果相关性高时, Pearson线性相关系数(PLCC)、Spearman等级相关系数(SROCC),和肯德尔等级相关系数(KROCC)的值会很高,而平均绝对偏差(MAE)和均方根误差(RMSE)的值会很小。对于理想状态下的立体图像质量客观评价算法,客观评价分数与主观评价MOS值完全一致时,  $PLCC = SROCC = 1$ , 并且  $RMSE = MAE = 0$ 。

[0159] 表2为本方法与其余四种评价方法的主客观相关性比较。在JPEG压缩实验中,本方法失真度评价方法的皮尔森线性相关系数远高于其它评价算法,拥有较好的准确性。高斯模糊失真情况下,本方法考虑人眼单目视觉中掩盖效应、多通道效应、对比度敏感度效应等视觉特性,并与人眼双目融合特性结合,因此对模糊图像表现出了较好的相关性。对于高斯模糊图像,文献[18]提出的改进信噪比评价方法也表现出较好的相关性,这是因为作者在评价图像失真度的过程中将人眼视觉特性与客观评价模型进行结合。因此在评价模糊立体图像过程中,将评价模型与人眼视觉特性相结合是提高评价准确性的一种方式。特别的,在高斯白噪声失真实验中,PSNR评价性能出众。这是因为高斯白噪声是一种全频域加性噪声,而PSNR作为误差统计的评价方法,能够直接计算图像像素值之间的差异,因此其能更准确的评价高斯白噪声对立体图像失真程度的影响。

[0160] 表2本文方法与其余四种评价方法相关性比较



[0161]

失真类型	评价方法	相关系数				
		PLCC	KRCC	SRCC	MAE	RMS
JPEG 压缩 失真	SSIM	0.899384	0.749441	0.909754	0.090794	0.126526
	PSNR	0.872532	0.794557	0.938381	0.082388	0.112829
	文献[18]	0.946471	0.831395	0.950411	0.076265	0.103072
	文献[30]	0.913616	0.736615	0.900287	0.099377	0.133612
	本文方法	0.948139	0.834392	0.954819	0.09873	0.130912
高斯模糊 失真	SSIM	0.808599	0.696004	0.860218	0.101727	0.135788
	PSNR	0.826123	0.716751	0.873726	0.280313	0.324622
	文献[18]	0.936907	0.836031	0.961283	0.101827	0.130274
	文献[30]	0.924988	0.848856	0.966403	0.103434	0.138397
	本文方法	0.959315	0.855417	0.970046	0.104641	0.128247
高斯白噪 声失真	SSIM	0.882993	0.728369	0.863624	0.100752	0.130481
	PSNR	0.937013	0.858273	0.97107	0.153876	0.179794
	文献[18]	0.927396	0.825891	0.94171	0.10194	0.129097
	文献[30]	0.908572	0.799616	0.910489	0.156939	0.182394
	本文方法	0.928017	0.837824	0.933382	0.107288	0.131087
全部失真 类型	SSIM	0.878959	0.730171	0.880578	0.107296	0.139924
	PSNR	0.864947	0.724965	0.893947	0.172192	0.223932
	文献[18]	0.943743	0.832191	0.955233	0.123344	0.121467
	文献[30]	0.903507	0.762804	0.917705	0.119917	0.153051
	本文方法	0.954226	0.844143	0.966186	0.103553	0.110088

[0162] 为验证本方法的普适性,表2最后展示了全部失真类型下五种客观评价算法主客观相关系数。可以看到,本方法分数的皮尔森线性相关系数(0.954226)、斯皮尔曼等级相关系数(0.966186)、肯德尔等级相关系数(0.844143)、平均绝对误差(0.103553)和均方根误差(0.110088)均为所有客观评价方法中的最优值。

[0163] 图12为五种失真度客观评价分数与主观分数(MOS)的散点图。散点图能够直观反映评价算法的评价结果,其中的每个点都代表着立体图像数据库中的对应图像。散点图的横轴为失真度客观评价算法结果,纵轴为主观质量MOS值,散点的密集程度代表客观评价方法评价的准确性。从图中可以看到,本方法散点图中横坐标和纵坐标之间的相关性较高,且评价方法稳定性较好,这说明本方法更加准确有效。

[0164] 参考文献

[0165] [1]成于庆,姜秀华,3D电视视频质量主客观评价最新技术研究[J],中国传媒大学学报自然科学版[J],中国传媒大学学报自然科学版,2012,19(1):31~36

[0166] [2]Donghyun Kim,SeungchulRyu and Kwanghoon.Depth perception and motion cue based 3D video quality assessment[C],2012IEEE International Symposium on Broadband Multimedia System and Broadcasting,Seoul,2012,1~4

[0167] [3]You J,Xing L,Perkis A,et al.Perceptual quality assessment for stereoscopic images based on 2D image quality metrics and disparity analysis [C].Proc.of International Workshop on Video Processing and Quality Metrics for Consumer Electronics,Scottsdale,AZ,USA.2010.

[0168] [4]Kang K,Liu X,Lu K.3D Image Quality Assessment Based on Texture

Information[C].Computational Science and Engineering(CSE),2014IEEE 17th International Conference on.IEEE,2014:1785~1788.

[0169] [5]Jung Y J, Kim H G, Ro Y M. Critical binocular asymmetry measure for perceptual quality assessment of synthesized stereo 3D images in view synthesis[J]. Circuits and Systems for Video Technology, IEEE Transactions on, 2015, 99(3):1~14

[0170] [6]田浩南. 基于边缘和局部匹配的立体图像质量评价[D]. 天津: 天津大学, 2013.

[0171] [7]Marc Lambooi, Wijnand Ijsselstein, Don G. Bouwhuis, et al. Evaluation of Stereoscopic Images: Beyond 2D Quality[J]. IEEE TRANSACTIONS ON BROADCASTING, JUNE 2011, 57(2):432~444.

[0172] [8]Shen Y, Lu C, Xu P, et al. Objective Quality Assessment of Noised Stereoscopic Images[C]. Measuring Technology and Mechatronics Automation (ICMTMA), 2011 Third International Conference on. IEEE, 2011, 2:745~747.

[0173] [9]Xing L Y, You J Y, Ebrahimi T, Perki A. A perceptual quality metric for stereoscopic crosstalk perception[C]. Proceedings of 2010 IEEE 17th International conference on Image Processing, 2010:4033~4036.

[0174] [10]Xing L, You J, Ebrahimi T, et al. Assessment of stereoscopic crosstalk perception[J]. Multimedia, IEEE Transactions on, 2012, 14(2):326~337.

[0175] [11]Shao F, Lin W, Gu S, et al. Perceptual full-reference quality assessment of stereoscopic images by considering binocular visual characteristics[J]. 2013, 22(5):1940~1953.

[0176] [12]段芬芳, 邵枫, 蒋刚毅等. 基于感知重要性的立体图像质量评价方法[J]. 光电工程, 2013, 40(10):70~76.

[0177] [13]郭美思. 立体图像质量评价[D]. 哈尔滨: 哈尔滨工业大学, 2012.

[0178] [14]卫津津, 李素梅, 刘文娟等. 基于可控金字塔的立体图像质量评价方法[J]. 计算机应用, 2012, 32(03):710~714.

[0179] [15]Yang J, Hou C, Xu R, et al. New metric for stereo image quality assessment based on HVS[J]. International Journal of Imaging Systems and Technology, 2010, 20(4):301~307.

[0180] [16]赵瑞超. 立体图像质量客观评价方法的研究与实现[D]. 天津: 天津大学, 2010.

[0181] [17]Fezza S A, Larabi M C, Faraoun K M. Stereoscopic image quality metric based on local entropy and binocular just noticeable difference[C]. Image Processing (ICIP), 2014 IEEE International Conference on. IEEE, 2014:2002~2006.

[0182] [18]佟晓煦, 李素梅, 刘富岩等. 基于人类视觉的感知立体图像质量评价方法[J]. 光电子·激光, 2015, 26(2):378~386.

[0183] [19]孔真真. 基于视觉感知的立体图像质量评价方法研究[D]. 宁波: 宁波大学, 2014.

- [0184] [20]赵辉,鲍歌堂,陶卫.图像测量中自动调焦函数的实验研究与分析[J].光学精密工程,2004,12(5):531~536。
- [0185] [21]Yang X K,Lin W S,Lu Z H,et al.Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile[J].IEEE Transactions on circuits and systems for video technology,2005,15(6):742~752。
- [0186] [22]王怡影.基于掩盖效应的图像质量综合评价方法[D].合肥:安徽大学,2012。
- [0187] [23]R L P.Orientation tuning of the spatial-orientation mechanisms of the red-green channel[J].Journal of Optical Society of America A,1997,14(10):2622~2632。
- [0188] [24]Watson A B,Solomon J A.Model of visual contrast gain control and pattern masking[J].JOSA A,1997,14(9):2379~2391。
- [0189] [25]宋化镜.基于小波分析和虚拟仪器技术的滚动轴承测试分析系统的研究[D].上海,上海大学,2011。
- [0190] [26]P.Tseng,R.Carmi,I.G.M.Cameron,et al.Quantifying center bias of observers in free viewing of dynamic natural scenes[J].Journal of Vision,2009,9(7):4。
- [0191] [27]Le Meur O.,Le Callet,P.,Barba,et al.A coherent computational approach to model bottom-up visual attention[J].Pattern Analysis and Machine Intelligence,IEEE Transactions on,2006,28(5):802~817。
- [0192] [28]Grossberg S,Kelly F.Neural dynamics of binocular brightness perception[J].Vision Research,1999,39(22):3796~3816。
- [0193] [29]Ding J,Sperling G.A gain-control theory of binocular combination [J].Proceedings of the National Academy of Sciences of the United States of America,2006,103(4):1141~1146。
- [0194] [30]佟晓煦.立体图像质量客观评价算法的研究[D].天津:天津大学,2014。

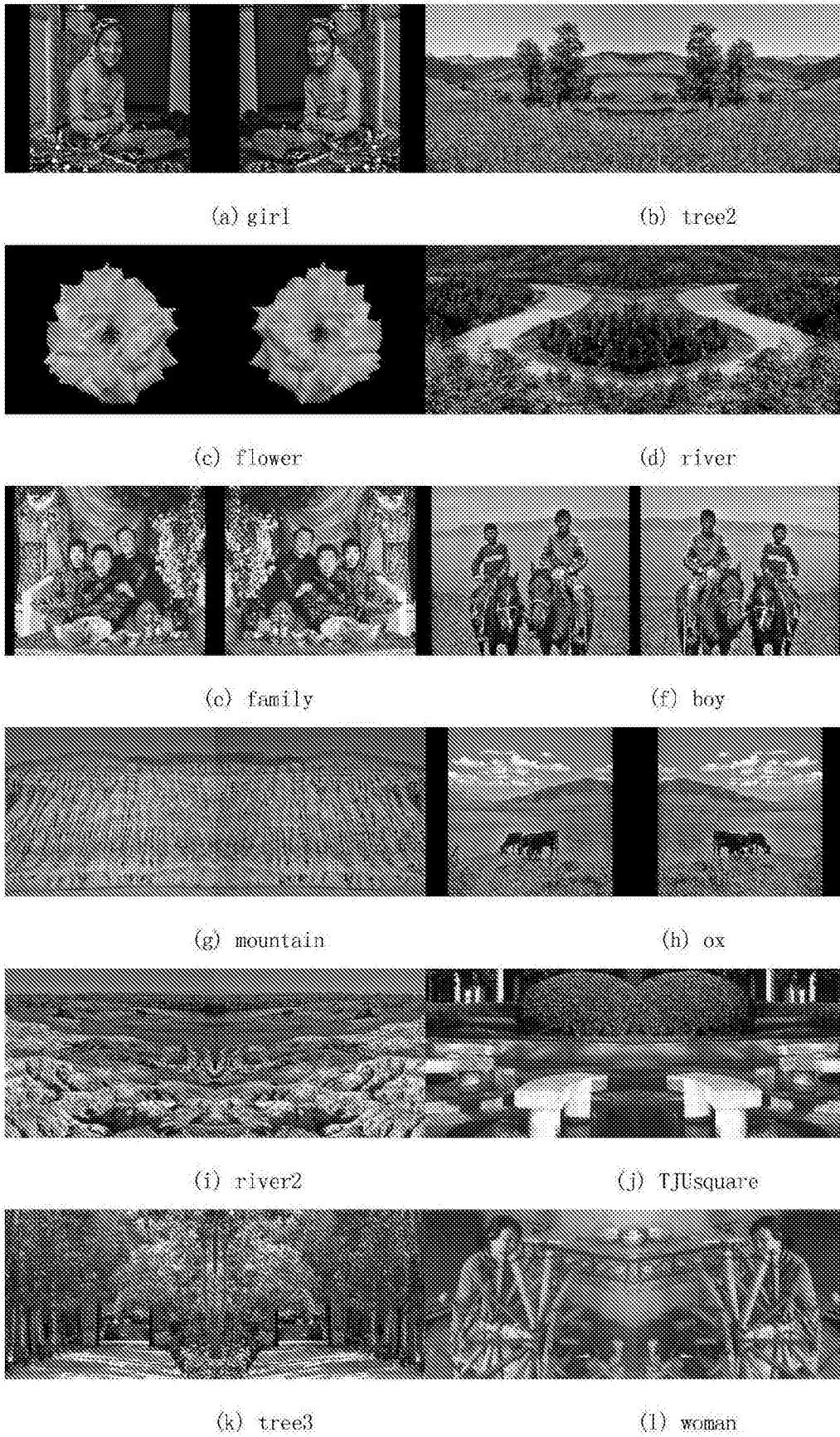
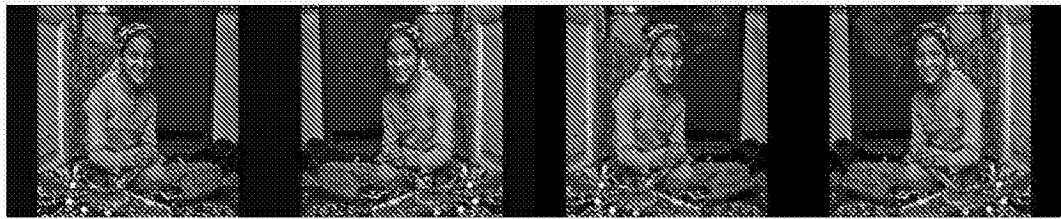
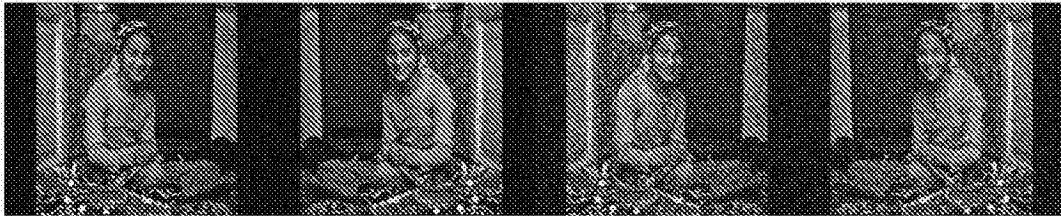


图1



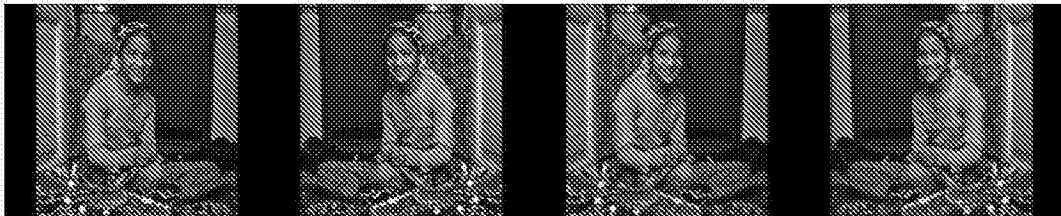
(a) 压缩率为 8%

(b) 压缩率为 2%



(c) 均值为 0, 方差为 0.02 的高斯白噪声

(d) 均值为 0, 方差为 0.06 的高斯白噪声



(e) 模板为 [15, 15], sigma 为 1.5 的高斯模糊

(f) 模板为 [15, 15], sigma 为 3.8 的高斯模糊

图2

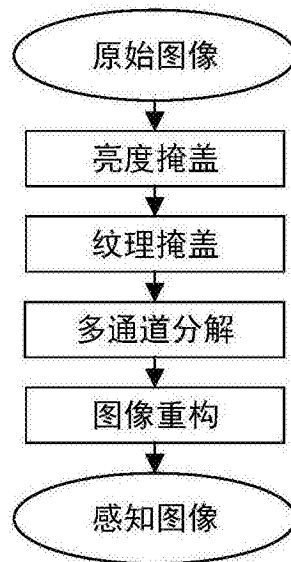


图3

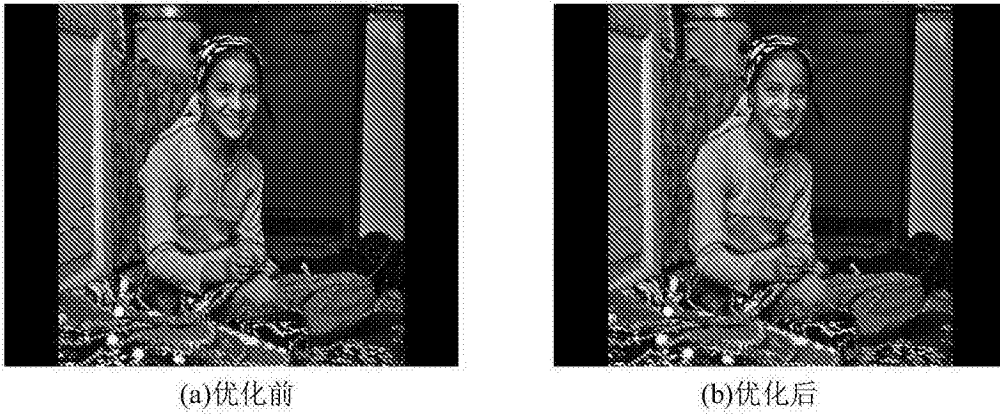


图4

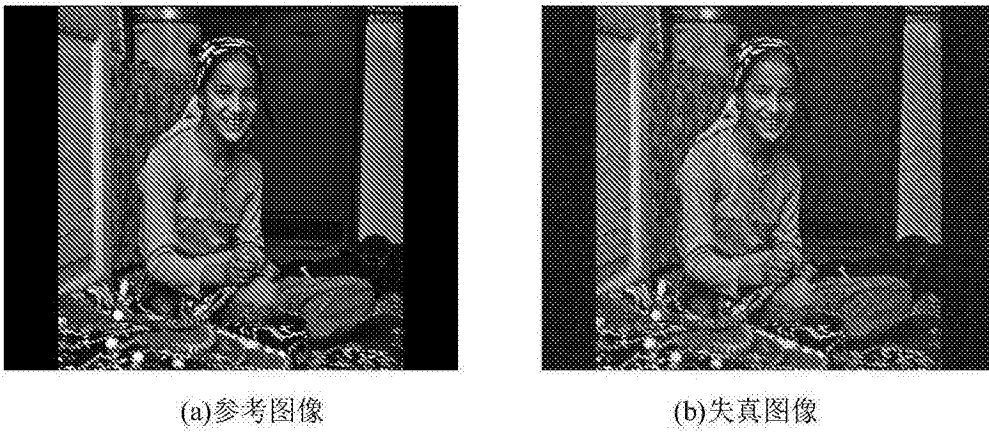


图5

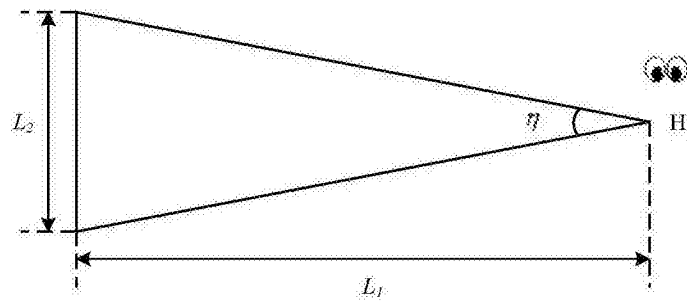


图6



(a)优化前

(b)优化后

图7

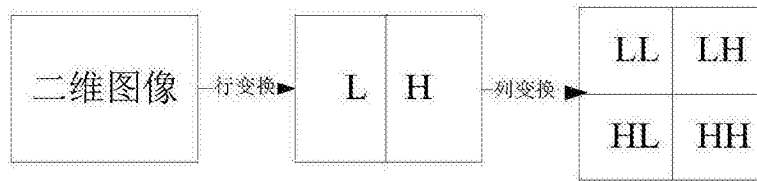


图8

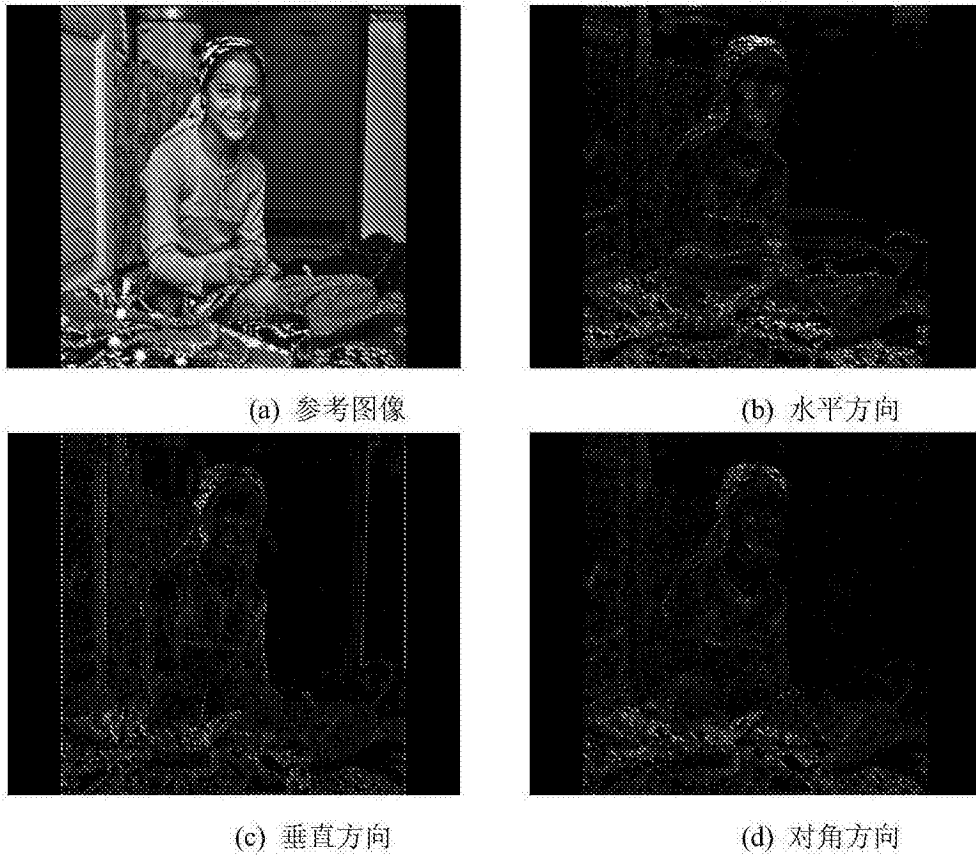


图9



图10



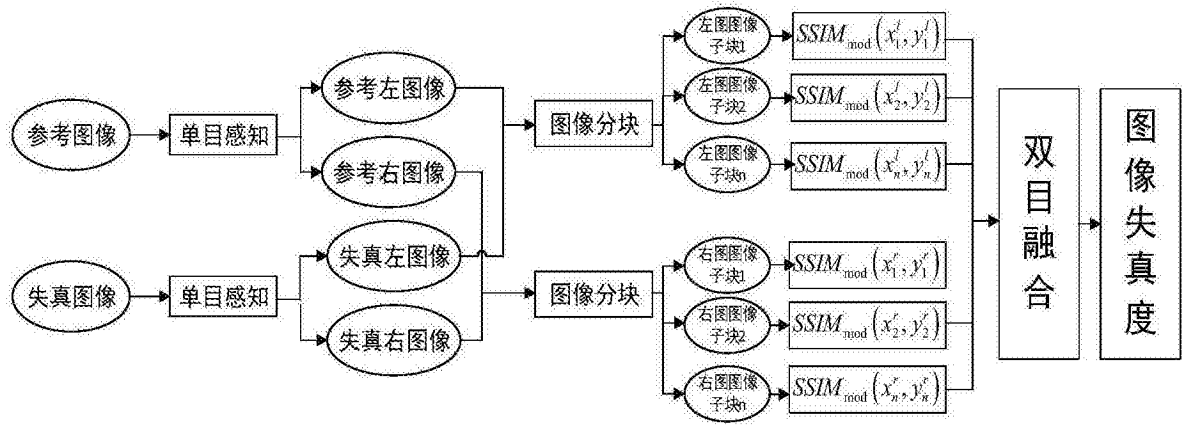
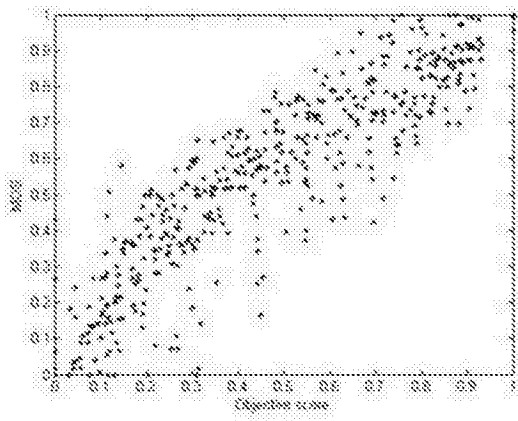
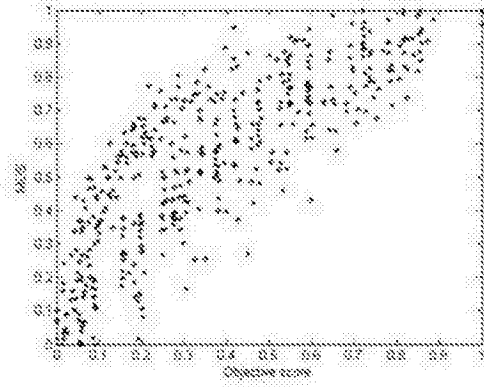


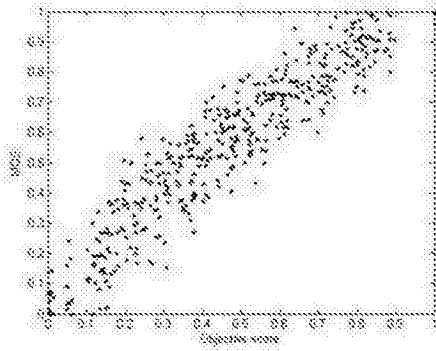
图11



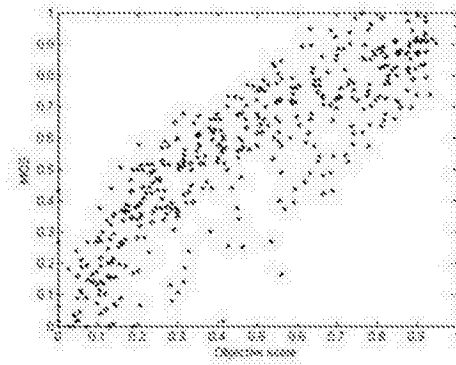
(a) SSIM



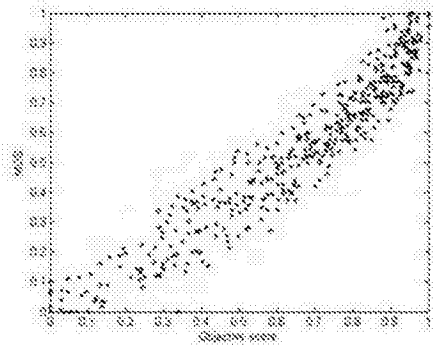
(b) PSNR



(c) 文献[18]



(d) 文献[30]



(e) 本文算法

图12