

基于结构相似的三维模型视觉降质度量

石振锋^{1,3}, 金 凯², 牛夏牧¹

(1. 哈尔滨工业大学 计算机科学与技术学院, 哈尔滨 150001, szfhit@126.com;

2. 哈尔滨工业大学 建筑学院, 哈尔滨 150001; 3. 哈尔滨工业大学 数学系, 哈尔滨 150001)

摘 要: 针对三维网格简化操作经常会引起一定的视觉降质和失真的问题, 为了更准确评估这种视觉降质, 利用三维网格显著特性信息作为网格顶点的属性, 采用包围盒空间剖分的方法描述三维局部窗口, 基于结构相似性和信息熵以定量的方式给出了三维网格模型的视觉感知降质的客观度量方法. 理论分析和实验结果表明: 此度量能非常有效地评估网格简化过程中引入的视觉降质, 可用于同源不同分辨率的三维模型之间的相似性度量, 并提供了多尺度特性的视觉降质分析.

关键词: 人类视觉系统; 感知; 结构化降质; 网格显著性; 熵

中图分类号: TP391.41

文献标志码: A

文章编号: 0367-6234(2010)09-01444-05

Evaluation of visual perception degradation for 3D mesh based on structural similarity

SHI Zhen-feng^{1,3}, JIN Kai², NIU Xia-mu¹

(1. School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China, szfhit@126.com;

2. School of Architecture, Harbin Institute of Technology, Harbin 150001, China; 3. Department of Mathematics,

Harbin Institute of Technology, Harbin 150001, China)

Abstract: Aimed at the problem that some visual perceptual degradation and distortion are usually generated during 3D mesh simplification process, we take the salient information as the property of the vertex and partition the bounding box of the mesh into some 3D local window. Then the mean perceptual structural degradation (MPSD) metric based on salient information and information theory is presented and used as a quantitative and objective metric to evaluate the visual perceptual degradation. Both analytical and experimental results show that the MPSD metric can effectively evaluate the visual perceptual structural degradation induced by mesh simplification operations and provide the functionality of multi-scale visual perceptual degradation.

Key words: human visual system; perception; structural degradation; mesh saliency; entropy

在计算机图形学应用中多边形网格常用于表示任意拓扑结构的三维模型, 其中三角形网格是最常用的表示方式之一, 其表示精度取决于三角形和顶点数目. 三维扫描技术、基于图像的重构等

技术的发展使得在给定的适当误差范围内, 描述三维模型所使用的三角形个数越来越多, 其表示精度也越来越高. 然而, Y. Pan 和 I. Cheng 等^[1]的视觉感知实验表明, 由于人类视觉系统 (Human Visual System, HVS) 视觉分辨率的限制, 对于给定的三维模型, 当其顶点数超过一定的数目后, 这些多余的三角形顶点所表示的模型的几何信息对于 HVS 是不可感知的^[2]. 同时, I. Cheng^[3]基于视觉心理学上注意到的差别 (Just Noticeable Difference, JND) 这一理论, 从实验而非理论的角度提供了一种用于评估不同的三维网格模型简化算法的实验方法.

收稿日期: 2010-07-05.

基金项目: 国家自然科学基金资助项目 (60703011, 60832010, 60671064); 国家高技术研究发展计划资助项目 (2007AA01Z458); 教育部重点基金资助项目 (20070213047); 哈工大科研创新基金资助项目 (HIT-NSRIF.2009050).

作者简介: 石振锋 (1976—), 男, 博士研究生, 讲师;
牛夏牧 (1961—), 男, 教授, 博士生导师.

为了减少三维网格的视觉冗余,实现三维模型的快速绘制渲染,国内外研究人员和学者已提出很多技术来达到这一目标,如网格简化、LOD等^[4]多分辨率技术. 这些技术通常采用面合并、递增式采样或重新网格化等方法来实现三维网格模型的简化,然而这些技术都将引入一定的视觉降质. 因此,评估或评价不同分辨率模型之间的视觉相似性或视觉差异是衡量网格简化算法好坏的关键问题. 三维模型最终是由人来观察和评估他们之间的相似性或视觉降质的,因此那些不考虑HVS因素的纯几何度量方法是不合适的. 正是基于这种考虑,国内外学者^[5-6]提出了一些基于人类视觉感知特性的度量方法,实验表明这种度量是更实际、更有效的方法.

基于HVS能高度自适应地提取场景的结构信息的假设,Z. Wang等^[7]提出了由亮度比较函数、对比度比较函数和结构化比较函数组成的结构化相似性(Structural Similarity, SSIM)索引算法,并利用平均结构化(Mean Structural Similarity, MSSIM)索引对两幅图像在全局质量上的相似性进行评估. 如果两幅图像完全相同则 $MSSIM = 1$, 当两副图像非常不同时, $MSSIM$ 的值趋近于0. 所以 $MSSIM$ 较好地以定量的方式给出了两幅图像之间的差别. 受SSIM的启发,G. Lavoué^[8]将该工作推广到三维模型,提出了网格结构化失真度量(Mesh Structural Distortion Metric, MSDM),并将MSDM用于评估三维模型数字水印算法中引入的失真,取得了较好的结果.

在三维网格感知或场景分析应用中,注意力机制是最常用的视觉感知属性之一. 显著特性是HVS的一种重要的低级特征,常用于描述HVS的注意力特性. 在本文中,首先定义了视觉感知结构降质模型,然后从信息论的角度,以熵作为主要工具给出了平均感知结构化降质(Mean Perceptual Structural Degradation, MPSD)的计算方法. 最后分析了MPSD算法在不同视距下的多尺度视觉降质特性. 实验结果表明,MPSD能有效地评估同源三维网格模型之间的视觉相似性和视觉降质,同时具有依赖于视距的多尺度结构相似的特性.

1 三维网格视觉显著特性

Itti等^[9]指出视觉注意力机制是与显著特性密切相关. Koch和Ullman^[10]提出了一种区分于其周边环境的图像显著特性位置计算模型. Lee等^[11]利用基于高斯加权曲率的中心-周围滤波器提出了可用于度量区域重要性的网格显著特性

算法. M. Feixas等^[12]给出了多边形互信息(Polygonal Mutual information, PMI)的定义,并对视点选择和网格显著特性建立了统一的信息论框架模型.

1.1 网格视觉显著特性计算模型

Lee等提出的网格显著特性计算模型能较好地模拟HVS自底向上的注意力机制,其计算流程如图1所示.

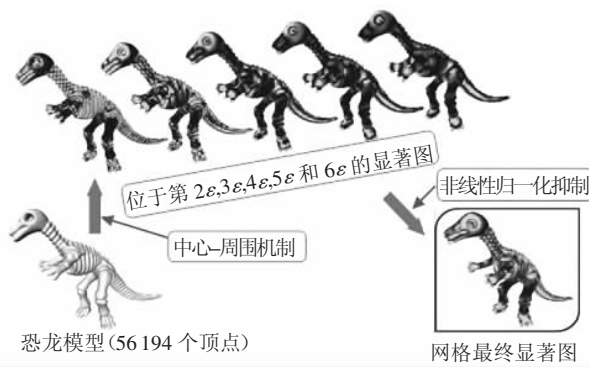


图1 三维网格视觉显著特性计算流程图

图2给出了不同分辨率的恐龙模型的视觉显著特性图,其中,三角形数目从左至右分别为112 384, 78 668, 44 954和11 238个.

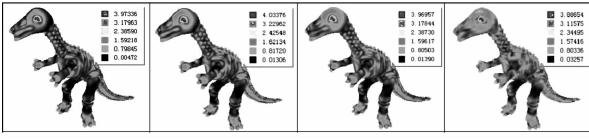


图2 恐龙模型的显著特性图

1.2 三维网格视觉显著特性归一化

给定三维网格 M , 其顶点数为 C_v , 由文献[11]计算得到的显著特性图为 $S = \{s_i\}$, $s_i \in [a, b]$, 为方便起见, 需将其归一化到单位区间 $[0, 1]$. 三维网格 M 的平均显著特性值 μ_s 为

$$\mu_s = \frac{1}{C_v} \sum_{i=1}^{C_v} s_i. \tag{1}$$

M 的方差 σ_s 为

$$\sigma_s = \left(\frac{1}{C_v} \sum_{i=1}^{C_v} (s_i - \mu_s)^2 \right)^{\frac{1}{2}}. \tag{2}$$

利用非线性映射将显著特性图 S 映射到 $S' = \{s'_i\}$, $s'_i \in [0, 1]$.

$$s'_i = \frac{1 - e^{-\frac{(s_i - a)}{2 \cdot \sigma_s}}}{1 - e^{-\frac{(b - a)}{2 \cdot \sigma_s}}}, s_i \in [a, b]. \tag{3}$$

2 基于结构相似的三维网格视觉降质

保持视觉上最大相似性或感知相似和最小的视觉降质是所有三维网格简化算法追求的目标. 由于HVS对几何特性并不敏感, 直接利用三维网格的几何信息来定义度量是不适合的. HVS具有

视觉掩蔽效应、视觉注意力机制和视觉自适应性等特点. 本文仅考虑以显著特性作为三维网格的重要视觉属性, 基于结构相似性给出一种评估此类操作引起的视觉降质的定量计算方法.

2.1 基于结构的视觉感知降质

与文献[7]提出的框架相似, 基于结构的视觉感知降质度量也是三维网格局部窗口内的显著特性进行均值、标准方差和协方差等统计计算的. 为此, 采用如图3所示的方式定义三维网格局部窗口. 设三维网格 M 的包围盒为 B , 在 X 、 Y 和 Z 等3个方向上将 B 分别均匀剖分为 N_x 、 N_y 和 N_z 等份, 则剖分后的每一个子包围盒 B_i 都被视为一个三维局部窗口. 本文中为简化这种剖分操作, 取 $N_x = N_y = N_z = N$, 其中, $1 \leq N < \infty$.

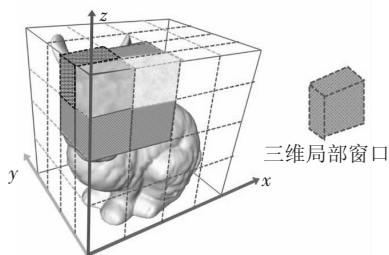


图3 三维网格局部窗口的定义

首先, 利用每个子包围盒窗口 B_i 内归一化的网格顶点显著特性计算其平均的显著特性值为

$$\mu_{B_i} = \frac{1}{C_{B_i}} \sum_{j=1}^{C_{B_i}} s(v_j).$$

式中: C_{B_i} 为第 i 个子包围盒中有效的三维网格顶点的个数, $s(v_j)$ 为该局部窗口内顶点 v_j 的归一化的显著特性值.

其次, 每个子包围盒窗口 B_i 的标准方差 σ_{B_i} 为

$$\sigma_{B_i} = \left(\frac{1}{C_{B_i} - 1} \sum_{j=1}^{C_{B_i}} (s(v_j) - \mu_{B_i})^2 \right)^{\frac{1}{2}}.$$

定义基于源网格 M_1 的协方差 $\sigma_{B_1^1 B_2^1}^{B_1^1}$ 为

$$\sigma_{B_1^1 B_2^1}^{B_1^1} = \frac{1}{C_{B_1^1} - 1} \sum_{j=1}^{C_{B_1^1}} (s(v_j) - \mu_{B_1^1}) (s(u_j) - \mu_{B_2^1}).$$

式中: B_i^1 和 B_i^2 分别为源网格 M_1 和目标网格 M_2 的第 i 个子包围盒窗口, $s(v_j)$ 为 B_i^1 中网格顶点 v_j 的归一化的显著特性值, $s(u_j)$ 为 B_i^2 中依欧氏距离度量离 B_i^1 最近的顶点. 类似地可以定义基于 M_2 的协方差 $\sigma_{B_1^2 B_2^2}^{B_2^2}$. M_1 与 M_2 之间的协方差为

$$\sigma_{B_1^1 B_2^2}^{B_1^1 B_2^2} = \frac{(\sigma_{B_1^1 B_2^1}^{B_1^1} + \sigma_{B_1^2 B_2^2}^{B_2^2})}{2}.$$

类似地, 引入网格 M_1 和 M_2 的第 i 个子包围盒 B_i^1 和 B_i^2 的3个比较函数: 显著特性比较函数 L 、对比度比较函数 C 和结构比较函数 S , 其表达式为

$$L(B_i^1, B_i^2) = \frac{2\mu_{B_i^1} \mu_{B_i^2} + C_1}{\mu_{B_i^1}^2 + \mu_{B_i^2}^2 + C_1},$$

$$C(B_i^1, B_i^2) = \frac{2\sigma_{B_i^1} \sigma_{B_i^2} + C_2}{\sigma_{B_i^1}^2 + \sigma_{B_i^2}^2 + C_2},$$

$$S(B_i^1, B_i^2) = \frac{\sigma_{B_i^1 B_i^2} + C_3}{\sigma_{B_i^1 B_i^2}^{B_1^1 B_2^1} + C_3}.$$

式中: 常量 C_1 用于避免当 $\mu_{B_i^1}^2 + \mu_{B_i^2}^2$ 非常小(趋向于0)时显著特性比较函数 L 可能引起的不稳定性. 特别地, 在本文中选择

$$C_1 = \max_{v_j \in B_i^1, v_k \in B_i^2} \{s(v_j), s(v_k)\} \varepsilon, \varepsilon \rightarrow 0,$$

$$C_2 = 3C_1, C_3 = 0.5C_2.$$

式中: ε 为任意非常小的正数常量.

最后, 利用上述3个比较函数 L 、 C 和 S 定义网格 M_1 和 M_2 之间的视觉感知结构降质为

$$PSD(B_i^1, B_i^2) = L(B_i^1, B_i^2)^\alpha C(B_i^1, B_i^2)^\beta S(B_i^1, B_i^2)^\gamma.$$

式中: 参数 $\alpha > 0, \beta > 0, \gamma > 0$ 用于分别调整3个分量 L 、 C 和 S 对视觉感知结构降质的贡献程度. 显然, $PSD(B_i^1, B_i^2)$ 满足下述4个条件:

1) 对称性: $PSD(B_i^1, B_i^2) = PSD(B_i^2, B_i^1)$.

2) 有界性: $0 \leq PSD(B_i^1, B_i^2) \leq 1$.

3) 最小值唯一性: 当且仅当三维局部窗口 B_i^1 和 B_i^2 中包含的顶点特性完全相同时 $PSD(B_i^1, B_i^2) = 0$.

4) 极大值唯一性: 当 B_i^1 和 B_i^2 差别非常大时, $PSD(B_i^1, B_i^2) \rightarrow 1$, 且理论上限为1.

本文中将 $PSD(B_i^1, B_i^2)$ 简记为 PSD_i .

2.2 平均视觉感知结构降质度量

事实上, 三维网格模型 M_1 和 M_2 在包围盒内视觉差异或相似性可由序列 $\{PSD_i\}$ 来表示, 可以将序列 $\{PSD_i\}$ 视为在各子包围盒窗口内的视觉感知特性的空间随机分布. 所以, 根据 Shannon 信息论, M_1 和 M_2 的平均视觉感知结构降质可由各序列分量的概率来确定. 设 p_i 为第 i 个相应的子包围盒 B_i^1 和 B_i^2 的视觉降质概率, 则 p_i 为

$$p_i = \frac{PSD(B_i^1, B_i^2)}{\sum_{i=1}^N PSD(B_i^1, B_i^2)}.$$

式中: N 为有效的子包围盒的个数. 在本文中子包围盒 B_i^1 和 B_i^2 中至少包含一个三维网格顶点时才被认为是有效的, 或者说只有在这种情况下, 子包围盒才会对视觉降质贡献一定的不确定性, 其物理含义是显然的. 利用熵来描述两个网格之间视觉降质的平均不确定性为

$$MPSD(M_1, M_2) = - \sum_{i=1}^N p_i \log p_i.$$

3 实验结果及分析

为了获得不同分辨率的三维三角化网格模型,利用开源软件 MeshLab V1. 1. 1 提供的二次边坍塌采样滤波器生成同源的不同分辨率的网格模型. 对于初始网格 M_0 ,按照每次简化源网格顶点数的 5% 来生成不同分辨率的网格模型序列 $M_1, M_2, M_3, \dots$.

3.1 三维网格显著特性信息分布

以三维三角化的恐龙模型为例,给出其归一化前后的显著特性信息的直方图分布. 对于归一化前后的区间 $[a, b]$ 和 $[0, 1]$,将区间按照三维模型顶点数的 2% 将区间分成 $N = 1\ 124$ 等份. 统计位于每个子区间内的显著特性值,得到 N 个概率. 恐龙模型的显著特性信息直方图分布如图 4 所示. 可以看出,归一化前具有较大显著特性值的分布是非常小的,而非线性归一化后其相应的分布明显得到了增强,这表明将对平均视觉降质的不确定性形成更多的贡献,或者说对人类视觉感知系统将产生更多的视觉刺激. 因此,非线性归一化操作将使得三维网格模型的视觉显著特性的分布在平均视觉感知降质量度中更趋合理.

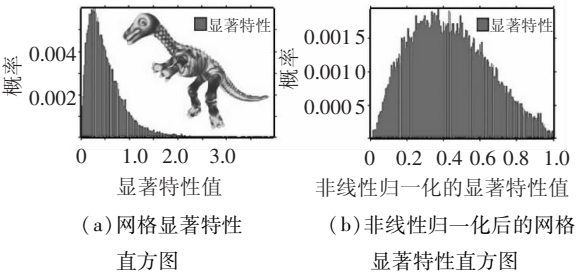


图 4 视觉显著特性在非线性归一化前后的直方图

3.2 平均视觉感知结构降质分析

选取三维网格模型包围盒的剖分参数 $N = 20$ 和 3 个比较函数分量的指数加权值为 $\alpha = \beta = \gamma = 1$. 图 5 给出了平均视觉感知降质值与网格简化率之间的关系. 当从原始恐龙模型(56 194 个顶点)中删除 90% 的顶点时,平均视觉感知降质值依然是 < 0.05 ,或者说简化了 90% 的顶点后的模型与原始模型之间依然是非常相似的. 当简化了原始模型的 95% 的顶点时,本文给出的算法表明其平均视觉感知降质值是 < 0.1 的,但明显大于简化率为 90% 时的情况,这表明能够明显感觉到简化带来的视觉降质. 而在简化率较小(如 $< 50\%$)时,平均视觉感知降质值非常小,随着简化率的增加,其变化也是非常慢的,这说明对于具有较多顶点的恐龙模型,HVS 对前期较小的简化率所能感知到的差别是不明显的. 而在简化率很

大时,平均视觉感知降质曲线明显变化加快,因此 HVS 也能明显地感知到与原始模型之间的差别. 这一点与主观上观察原始模型和简化后的模型时感受到的差别及差别的变化规律是一致的. 所以,本文提出的平均视觉感知降质量度能非常有效地用于评估网格简化过程中的视觉降质.

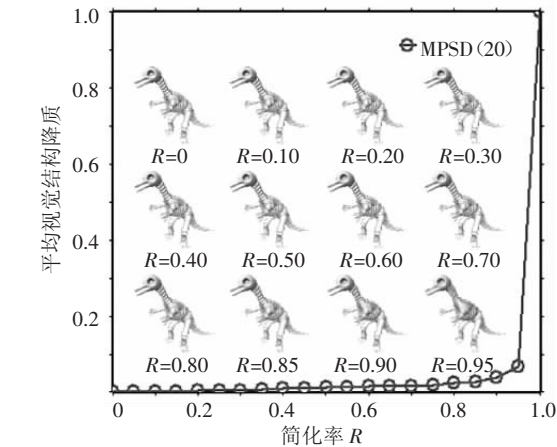


图 5 平均视觉感知降质与网格简化率之间的关系

3.3 多尺度的平均视觉感知结构降质分析

人类视觉感知系统具有很强的多分辨率、多尺度特性,图 6 给出了兔子模型在相同视点方向、不同的视距下,不同简化率的网格模型的视觉效果. 显然,当简化率固定时,随着视距的增大,对模型简化引起的视觉降质感受越来越不明显;当视距固定时,随着网格简化率的增大,越来越明显地感受到简化引起的视觉差别. 这就是 HVS 的多尺度与多分辨率特性.

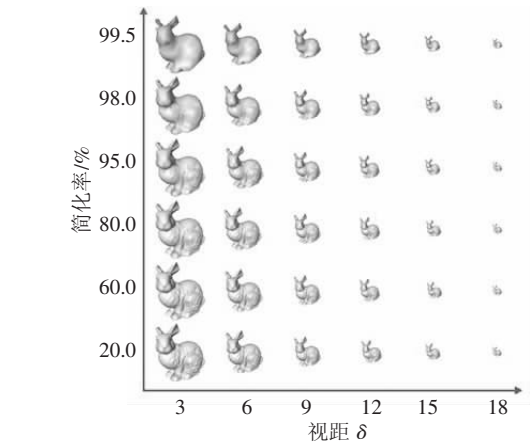


图 6 不同简化率下的网格模型在不同视距下的关系

本文提出的平均视觉感知降质量度可通过设置不同的包围盒参数 N 来模拟人类视觉感知系统的多分辨率与多尺度特性. 本文改变包围盒剖分参数 N 的值,即可模拟不同的视距. 图 7 给出了平均视觉感知降质随不同的剖分参数 N 的变化. 在该实验中取 $N = 1, 2, 4, 5, 10, 20, 50, 100$ 来说明 HVS 的多尺度特性. 显然,提出的平均视觉感

知降质度量较好地、客观地模拟了该特性.

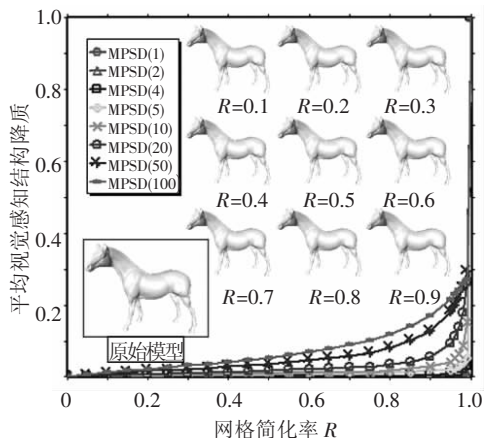


图 7 多尺度平均视觉感知降质分析:马模型
(源网格:48 485 个顶点,96 966 个三角形)

4 结 论

1) 基于信息论和结构相似的同源三维模型视觉降质度量方法能很好地定量地分析三维网格简化过程中引入的视觉差异和降质. 同时,该方法还能很好地模拟 HVS 的多尺度和多分辨率特性.

2) 该方法运用于三维网格简化等操作,能很好地度量简化引起的视觉降质,尽可能地保持网格在包围盒内的视觉降质分布.

3) 基于结构相似的三维模型视觉降质度量还能用于比较所有网格简化算法在视觉相似性上的保持程度,从而对所有网格简化算法进行定量地、客观合理地比较. 也能用于评价三维网格模型数字水印算法在嵌入后的视觉降质和变化及三维网格模型压缩算法引起的视觉降质等.

参考文献:

[1] PAN Y X, CHENG I, BASU A. Quantitative metric for estimating perceptual quality of 3D objects [J]. IEEE Transactions on Multimedia, 2005, 7(2):269-279.
[2] CHENG I, BASU A. Reliability and judging fatigue reduction in 3D perceptual quality estimation [C]//Proceedings of the 2nd International Symposium on 3D Data Processing, Visualization, and Transmission. Washington, DC:IEEE Computer Society, 2004:797-804.

[3] CHENG I, SHEN R, YANG X D, *et al.* Perceptual analysis of level-of-detail: The JND approach [C]//Proceedings of the Eighth IEEE International Symposium on Multimedia. Washington, DC:IEEE Computer Society, 2006:533-540.
[4] GARLAND M, HECKBERT P S. Surface simplification using quadric error metrics [C]//Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques. New York, NY:ACM, 1997:209-216.
[5] HOPPE H. Progressive meshes [C]//Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New York, NY:ACM, 1996:99-108.
[6] GARLAND M, WILLMOTT A, HECKBERT P S. Hierarchical face clustering on polygonal surfaces [C]//Proceedings of the 2001 Symposium on Interactive 3D Graphics. New York, NY:ACM, 2001:49-58.
[7] 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.
[8] LAVOUÉ G, GELASCA E D, DUPONT F, *et al.* Perceptually driven 3D distance metrics with application to watermarking [C]//SPIE Applications of Digital Image Processing XXIX. 2006. San Diego, CA:SPIE, 2006: 63120L.
[9] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11):1254-1259.
[10] KOCH C, ULLMAN S. Shifts in selective visual attention: Towards the underlying neural circuitry [J]. Human Neurobiology, 1985, 4(4):219-227.
[11] LEE C H, VARSHNEY A, JACOBS D W. Mesh saliency [J]. ACM Trans Graph, 2005, 24(3):659-666.
[12] FEIXAS M, SBERT M, GONZÁLEZ F. A unified information-theoretic framework for viewpoint selection and mesh saliency [J]. ACM Transactions on Applied Perception, 2009, 6(1):1-24.

(编辑 张 红)