

中国科学技术大学

学士学位论文



基于谱的网格显著性分析

Spectral-based Mesh Saliency Analysis

作者姓名:	朱飞达
学号:	PB10210217
学科专业:	自动化
导师姓名:	刘利刚 教授
完成时间:	二〇一四年六月

致谢

完成这篇毕业论文，首先要感谢我的指导老师刘利刚教授。虽然不是刘老师的研究生，但是刘老师给予了我无私的帮助和悉心的指导。大四上学期，通过实现图形学领域的一些算法，让我对计算机图形领域有了初步的研究基础和实践能力，在做毕业设计的过程中，刘老师也鞭策我积极进取，广泛了解显著性检测领域。刘老师的严谨求实的治学态度、孜孜以求的工作作风和大胆创新的进取精神对我产生了重要影响。

我还要感谢我的班主任陈金雯老师、自动化系的所有任课老师，在他们指导下，学习了自动化领域的很多知识，锻炼了科研能力。感谢母校为我们创造了良好的学习环境，让我们可以全身心投入到科研中。

最后，感谢父母的养育之恩，感谢他们对我生活上的关心和照顾。

目录

摘要	2
Abstract	3
第一章 绪论	5
1.1 研究背景和意义	5
1.2 论文内容安排	6
第二章 图像显著性分析	7
2.1 视觉注意力机制与显著性模型机制	7
2.2 评价方法	7
2.2.1 线性相关系数法	8
2.2.2 ROC 曲线下面积法 (AUC)	8
2.2.3 主观评价法	9
2.3 图像显著性检测模型	9
2.3.1 符合感知的模型	9
2.3.2 谱分析模型	10
2.3.3 机器学习模型	11
第三章 网络显著性分析	12
3.1 基于感知的显著性检测方法	12
3.1.1 基于曲率的局部显著性检测	12
3.1.2 加入整体考量的显著性检测	13
3.2 基于谱的显著性检测方法	14
第四章 网络谱分析介绍	15
4.1 光滑边框的例子	15
4.2 网络谱分析	18
4.2.1 基本概念	18
4.2.2 谱分析的优点	18
4.2.3 谱分析的难点和挑战	19
第五章 基于谱的网络显著性分析	20
5.1 基于特征分解的显著性分析	20
5.1.1 算法与分析	20
5.1.2 显著性检测结果	23
5.2 基于谱变换的显著性分析	24
5.2.1 算法与分析	24
5.2.2 显著性检测结果	27
5.3 基于谱残差成分的显著性分析	29
5.3.1 算法与分析	29
5.3.2 显著性检测结果	31
第六章 总结与展望	33
参考文献	34

摘要

人在进化过程中发展出了一套高效的视觉处理机制，即使是从没见过的场景，人类也能迅速去关注场景中更为显著的部分。显著性是一个反映场景中目标吸引人注意力能力的术语，网格显著性分析就是尝试检测出网格中引起人们关注的区域。

图像显著性领域的研究已经有了很大的发展，了解图像显著性检测，对研究网格显著性有帮助和启迪作用，本文简要回顾了图像显著性检测领域的发展。谱分析的手段在网格几何处理中得到了很大的应用。大体上说，谱算法是通过研究使用特征值、特征向量、特征空间的投影等来解决某一问题。本文通过介绍谱在网格平滑中的应用来引入网格谱分析的基本概念。

本文尝试将谱分析的手段用于检测网格显著性，用了基于特征分解、谱变换、谱残差成分的方法。其中基于特征分解是已经提出的算法，本文新提出了基于谱变换与谱残差成分的算法。本文先实现并分析了基于特征分解的算法，这种方法的想法是谱中存在冗余信息，去除冗余，得到的偏差成分也许能反映显著性信息。然后本文提出基于谱变换的检测方法，在这种方法中，采用了一个模仿butterworth滤波器设计的光滑低通滤波器对谱滤波，通过反变换，根据顶点在新旧网格的位移大小来决定显著性图，本文同时指出了谱变换的方法所依据的原理，这种算法其实是比较原网格和光滑后的网格，那些坐标变化比较大的顶点，给予它们较大的显著性值。网格上比较突出的部分在光滑过程中会变化比较大，因此，这种算法趋向于能找到网格上的突出部分。谱残差成分方法的想法也是基于谱的残差成分也许能反映显著性信息，然而，这种方法与基于特征分解的方法也有不同，这种方法先求取网格顶点处的曲率，作为顶点处的特征，然后对曲率谱变换，对谱变换后的系数求取偏差成分，然后反变换后的值即作为显著性成分。本文在不同模型上检验了这三种算法，从结果可以看出，每种方法都有优点和局限性。

关键词：视觉注意力，显著性，谱分析，特征分解，谱变换，谱残差

Abstract

People have developed a set of efficient visual processing mechanism in evolution. Even if the scene is new, human beings are able to detect the salient regions quickly to obtain the detailed information of these regions. Saliency is a term which reflects the regions' abilities of attracting people's attention. Mesh saliency analysis attempts to detect the salient regions in the mesh.

Image saliency detection methods have improved greatly during the last decades. It's helpful and enlightening for us to learn those methods when we do research on mesh saliency. This paper makes a brief overview of image saliency detection. Spectral analysis has been widely used in mesh geometry processing. Generally speaking, spectral algorithms solve problems by examining and using the eigenvalues, eigenvectors, eigenspace projections, etc. This paper tries to explain a simple application of incorporating spectrum into mesh smoothing. Some basic concepts of mesh spectral analysis will be introduced next.

This paper attempts to incorporate spectral analysis into mesh saliency detection. The eigendecomposition-based, spectral-transform-based and spectral-residual-based methods are discussed in this paper. The eigendecomposition-based method has been proposed. This paper originally proposes the spectral-transform-based and spectral-residual-based methods. This paper first realizes and analyses the detection method based on eigendecomposition. The idea behind this method is that there is redundancy in the spectrum and the deviations derived from removing the redundancy may be able to reflect saliency information. Then, another detection method based on spectral transform is proposed. A low-pass smoothing filter which mimics the Butterworth filter is manipulated on the spectrum. The saliency of a vertex is determined by the magnitude of displacement between the original mesh and the new mesh derived from inverse transform. This paper also discusses the mechanism behind the spectral transform method. This method actually compares the original and the new mesh after smoothing. The vertexes which have relatively large changes in the coordinates have high saliency values. The global prominent regions will have larger changes during the smoothing process. Therefore, this method tends to find the

prominent parts of the mesh. The spectral-residual method is based on the idea that the residual of the spectral may be able to reflect saliency. However, it's different from the eigendecomposition-based method. This method first calculates the curvatures regarded as the features at the vertices of the mesh. Then, spectral transform is manipulated on the curvatures. We calculate the residual components on the spectral and the saliency can be obtained from the inverse transform. This paper examines these three methods in different meshes. From the results, we can see that each method has advantages and limitations.

Key words: visual attention, saliency, spectral analysis, eigendecomposition, spectral transform, spectral residual

第一章 绪论

1.1 研究背景和意义

视觉是人类获取外界信息的最主要方式，当我们睁开双眼，就会捕获大量的信息，如何选择更重要的信息显得尤为重要。在进化过程中，人类发展出了一套高效的视觉处理机制去处理视觉信息，给予那些引起人注意的、比较显著的部分更多关注，从而获取那些显著区域更多的细节信息，而略看其他不太显著的部分。那些引起人注意的称为显著性区域。

视觉注意力，是指观察者对视觉场景显著、重要部分选择和处理的能力，确保对输入的视觉信息选择、优化。人类视觉注意能力的生理基础为中央凹入高分辨率部分（通常称为黄斑）及周围低分辨率部分组成的视网膜组织[10]。这一项能力其实每时每刻都在被我们使用，然而这一项看似自然而然的过程却蕴含着很多人脑的计算处理，人类大脑迅速运行一系列判断策略，选定视觉场景中想要关注的区域，通过眼动扫描，将该区域移到具有高分辨率的视网膜中央凹区，实现更细致的观察和分析。目前还没有能与人类视觉注意能力相匹敌的计算机系统。人类对某个场景的视觉关注由一系列因素决定，其中物体显著性是一个重要的决定性因素。物体显著性是指视觉场景中能引起人注意的物体所具有的特性，也可以指物体上能引起人注意的局部。

一方面，图像领域的显著性检测已经取得很大的发展，Itti 等[4]提出了基于感知的检测方法，Hou 和 Zhang[7]从频域分析，近几年也出现了利用机器学习的检测方法[9-10]，图像显著性在目标检测、图像压缩、图像检索、视觉界面设计等等领域有很广泛的应用，受图像显著性分析的激发，网格显著性检测也在发展，成为网格分析方面的一个研究课题。另一方面，网格的分析、表现和处理，如网格简化、网格光滑等等[1-2]，在十多年来有了很多研究，涌现了大量应用。提取出网格的显著性，也会对一些网格处理应用有指导作用，改善结果，比如 Lee 等[3]将显著性结果用在了网格简化和视角选择中，Song 等[11]将显著性结果用于网格分割、扫描融合中。因此，网格的显著性检测是有意义的。

研究显著性的难点在于：物体显著性的确定，除了受低层次的颜色、亮度、曲率等特征影响外，还受人类高层次的经验的影响。人类的视觉注意系统涉及心理学、生物学等多个领域，目前的生物基础理论还不完善，计算机系统要在已有

的理论知识基础上,构造出模仿人类视觉注意机制的系统,使得计算机得到显著性结果尽可能的接近人类视觉注意区域的实际结果。

在信号处理领域,将时域空域信号通过傅里叶变换、小波变换等方式,得到频域信号,一些在时域不容易处理的问题,比如信号去噪等等,能在频域得到很好的解决。谱分析的手段在图形几何处理领域也有了越来越多的应用,比如形状匹配、网格分隔、聚类等等。大体上说,网格谱算法是通过研究使用特征值、特征向量、特征空间的投影等来解决某一问题,Zhang 等[17]在 2010 年比较系统的介绍了谱分析在网格处理中的应用。

本文希望能从谱分析的角度,分析网格的显著性成分。

1.2 论文内容安排

本文一共有六章。

第一章绪论,介绍了视觉注意力机制,阐明了显著性的概念以及将谱分析的手段用于网格显著性检测的意义。

第二、三章,分别介绍近几十年来图像、网格显著性检测领域的发展,除了总结模型,也说明了原理、评价方法等内容,从而对显著性检测领域有个整体的把握。

第四章介绍了网格谱分析。网格谱分析初次接触比较抽象,本文从一个例子来引入网格中谱的概念,并介绍了谱分析的优点和挑战。

第五章,分别从特征分解、谱变换、谱残差成分的角度,尝试将谱分析用于网格显著性检测,其中基于特征分解是已经提出的算法[11],本文新提出了基于谱变换和谱残差成分的算法,每节都会对算法进行分析,分析了检测结果和方法的局限性。

第六章是对论文工作的总结以及对未来工作的展望。

第二章 图像显著性分析

图像显著性检测和网格显著性检测机制有很多相似的地方,比如与周围区域不同的突出部分显著性程度会比较高,具体来说,一副黄色背景图像有一块红色区域,那么人会被红色区域吸引,网格中一块平坦的区域有一块突出的区域,那么人会被突出区域吸引。不仅背后的检测机理有相似的地方,而且图像显著性检测的计算模型也可以给网格检测带来启发。比如 Lee 等[3]就用到了 Itti 等[4]提出的中心环绕算子。所以在进入网格显著性分析之前,本文先对图像显著性分析进行概述。

2.1 视觉注意力机制与显著性模型机制

图像视觉注意力机制主要有自顶向下(top-down)和自底向上(bottom-up)。自顶向下是基于已有的知识或经验,比如人会比较关注图像中的人脸,虽然人脸可能不是图像中最突出的地方。自底向上是指没有用到先验知识,给出一副没见过的图像,人会对突出区域比较敏感。

据此显著性检测模型也可以分为两大块,即 top-down(TD)和 bottom-up(BU)。比如 Itti 等[4]提取图像中颜色、亮度、方向特征,得到特征图,特征图经过叠加后得到显著图,这便是 BU 的例子。TD 的例子比如, Judd 等[5]用人眼注视的数据集训练了线性支持向量机,用来检测图像中的特征区域。

2.2 评价方法

根据某种计算模型,我们可以计算出图像的显著图,将显著图与人眼追踪数据比较,来判断结果的好坏。为了方便表示,在下述内容中,模型计算得到的显著图称为 Estimated Saliency Maps(ESM),将根据人眼注视数据或结合了人类标注显著区域得到的显著图称为 Ground-truth Saliency Maps(GSM)。

虽然评价方法各有不同,但是他们均是建立在有参照显著图的基础上。图像显著性检测领域,已存在通过人眼追踪、鼠标跟踪等方式建立的数据集,其中的图片都有真实的参考显著图。不过,目前检验模型效果的数据集还不算很多。

而在网格显著性检测领域,更是缺乏对模型的评价工作。就目前所看的相关论文中,没有提到真实的参考显著图,论文中仅仅是在比较不同模型之间的结果。

本文觉得建立数据集、完善评价指标也应成为一项研究内容。

具体的评价方法有很多种，在此列举几条。

2.2.1 线性相关系数法

线性相关系数法即为 Linear correlation coefficient，这种方法被广泛用来比较两幅显著图之间的相关性。

$$CC(G, S) = \frac{\sum (G(x, y) - u_g) * (S(x, y) - u_s)}{\sqrt{\sigma_g^2 \sigma_s^2}} \quad (2.1)$$

G 代表 GSM, S 代表 ESM。 u 和 σ 分别表示显著图各自的平均值和标准差。CC 方法的好处是有个单一的计算值，且在-1 到+1 之间，结果越靠近-1 获+1，说明两幅显著图线性相关性越好。

2.2.2 ROC 曲线下面积法 (AUC)

AUC 表示 area under curve，ROC 表示 Receiver Operating Characteristic。这种方法在机器学习中很常用，这里把显著性检测模型看作一个分类器。通过模型计算得到显著图后，通过设定阈值，将显著性值高于阈值的点标记为预测阳性点，低于阈值的点记为预测阴性点。而根据人眼追踪得到的正确显著图，可以将原图中每一个点标记为真正阳性点或真正阴性点。假设真正阳性点有 A 个，真正阴性点有 B 个。其中预测阳性点的确是真正阳性点的个数记为 a 个，预测阳性点其实是真正阴性点的个数记为 b 个。

真阳性率 true positive rate(TPR)即为 a/A ，假阳性率 false positive rate(FPR)即为 b/B ，TPR 作为纵坐标，FPR 作为横坐标。设定好一个阈值，可以计算出对应的 TPR, FPR，对应着平面中一个点。从小到大不断增大阈值，便能得到 ROC 曲线（离散情况下是折线）。一个较好的分类器的 ROC 曲线类似图 2.2.2 所示。

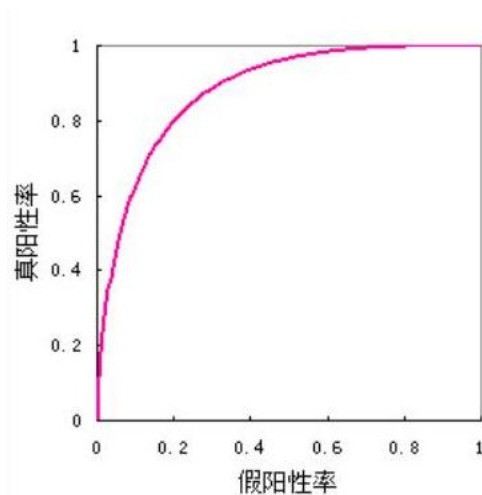


图 2.2.2 ROC 曲线示例，来自[6]

AUC 即为曲线下的面积。AUC 一般在 0.5~1 之间。AUC 越大，说明分类效果越好，显著性检测模型越精确。

2.2.3 主观评价法

人们主观的对显著性检测模型得到的显著图判断优劣，比如可将评价等级分为好、中、差，对大量人员进行调查，通过统计，来判断模型的优劣。然而这种方式耗费人力，评价少量图片还行，但难以将这种方式扩展到评价大的数据集。

2.3 图像显著性检测模型

目前已有很多检测模型，了解已有的图像检测模型，也会给网格显著性检测带来启发。这里本文简要介绍三种模型。

2.3.1 符合感知的模型

最经典的就是 Itti 等[4]于 1998 年提出的检测方法。这种方法先对原图用高斯滤波器下采样，得到不同尺度下的图像，从 1: 1 到 1: 256 共 9 个尺度。对每一个尺度下的图像，提取颜色、强度、方向信息。 r 、 g 、 b 是原图像的三个颜色通道。

该方法将强度表示为

$$I = \frac{r + g + b}{3} \quad (2.2)$$

将颜色表示为四通道的颜色特征

$$\begin{cases} R = r - (g + b)/2 \\ G = g - (r + b)/2 \\ B = b - (r + g)/2 \\ Y = \frac{r + g}{2} - \frac{|r - g|}{2} - b \end{cases} \quad (2.3)$$

利用 Gabor 滤波器对图像滤波，得到 0° 、 45° 、 90° 、 135° 四个方向的方向特征。

对颜色、强度、方向三个信息分别用中心环绕算子进行计算得到特征图，归一化后，线性加权求和就得到最后的显著性图。

这种方法的计算原理符合人的视觉感知规律，即与周围区域不同的突出区域引人注意，然而这种方法计算局部对比度，不能从整体上分析显著性，而且也不能识别出图像中存在哪些对象。

2.3.2 谱分析模型

Hou 和 Zhang[7]在 2007 年提出一种从频域分析的模型，而不是在空域分析。这种方法基于的想法就是：谱中的相似性表明存在冗余，去掉谱中的冗余成分，得到的残差成分也许能反应原图像中突出的区域。

整体的计算步骤很简单：

- Step1: 对原图像傅里叶变换得到频域空间的振幅图和相位图。
- Step2: 对振幅图取对数得到谱图
- Step3: 用一个 $n \times n$ 的平均滤波器对谱图求平均，和原谱图作差，得到谱图的残差图
- Step4: 对残差图进行指数运算，然后和相位图一起，做反傅里叶变换，得到初步的显著性图。
- Step5: 为了使显著性图视觉效果好一点，用高斯滤波器对显著性图滤波，得到最后的显著性图。

Matlab 只用 5 行代码就可以实现[8]。

虽然这种方法易于实现,效果也挺成功,但是缺少必要性或者充分性的证明,而且生物学机理上的解释也不清楚。

2.3.3 机器学习模型

机器学习的手段也被应用在了图像显著性检测领域。通过学习已有的人眼注视数据或者标注过的显著区域,模型会学习到一种“刺激-显著性”机制,即如果输入图像中存在那些学习到的刺激因素,则会提高显著性值。

Kienzle 等[9]介绍了一种非参数的自底向上的手段,直接从人眼追踪数据学习。这个模型由一个非线性映射构成,从一块图像分区映射到一个具体的值。它被训练成如果是显著区域则输出正值,如果是随机选取的非显著区域则输出负值。显著性函数可以通过使观测数据的预测结果最大化确定[10]。

随着可用的人眼追踪数据增多,机器学习的模型得到更广泛的应用。然而,这些模型的预测结果是基于数据的,而且预测原理在某种意义上是个黑盒(black-box)。

第三章 网格显著性分析

网格显著性分析在十多年来也有了一定的发展,但是缺少像图像显著性检测中的那些数据集,目前对结果的评价还比较主观。尽管这样,研究网格显著性检测技术仍然很有意义,显著性结果也广泛的用在了形状匹配、网格简化、网格分割等领域[11-12]。这里介绍两类显著性检测方法

3.1 基于感知的显著性检测方法

从人类感知看,图像中最关键的是颜色、强度、方向等信息,网格中则是形状。这类方法提取网格形状特征,经过一定计算,使得与周围区域不同的突出区域显著性值高。

3.1.1 基于曲率的局部显著性检测

人们一般认为在曲率大的地方,容易引起关注,而平滑的地方不会引起注意。这里用曲率来描述形状,可以用 Taubin[13]中的方法计算离散网格中每一个点处的曲率。但是只把曲率作为显著性会有图 3.1.1 (1) 的问题。

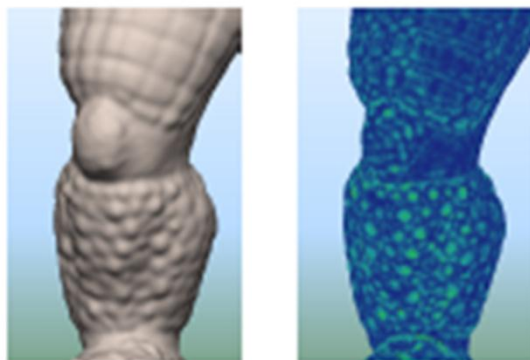


图 3.1.1 (1) 穿山甲模型的腿部, 只根据曲率得到的显著性图, 来自[3]。

小腿部分虽然曲率比较高,但小腿表面普遍比较粗糙,人们不会关注小腿表面,因此小腿部分的显著性值应该比较低。

在一块平坦区域,有一块突出的地方,人们会认为突出的地方显著;在一块粗糙区域,有一块平坦的地方,人们会认为平坦的地方显著。可见与周围区域不同的地方会引起人的注意。基于这种人类感知机制, Lee 等[3]提出了改进方法,

他以每个点处的曲率作为基础，之后用一个中心环绕的高斯加权平均算子滤波，得到显著性值。高斯算子的尺度半径不同，最后得到的显著性图也会不同。对不同尺度下求到的显著性图非线性加权叠加，得到最后的显著性图。

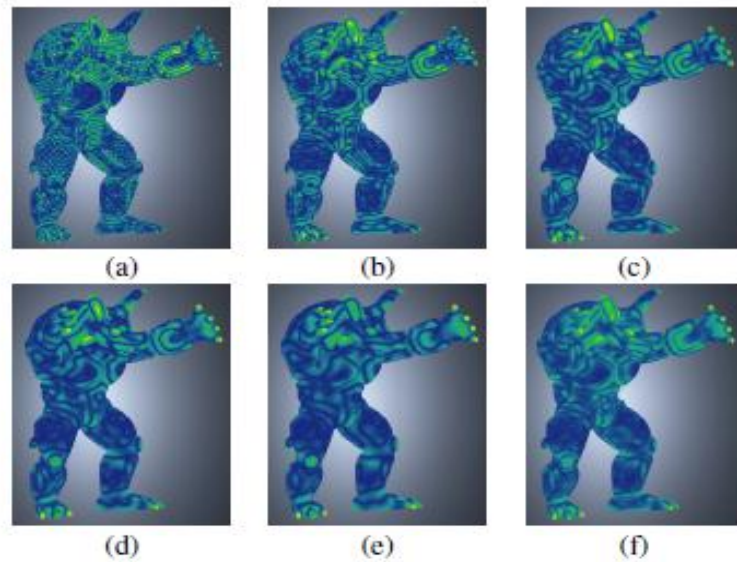


图 3.1.1(2) (a)-(e)为不同尺度下，得到的显著性图。(f)为叠加后的结果。来自[3]

这种方法，并没有从整体上求显著区域，而且需要先检测到当前点周围区域的点，才能再用高斯加权算子滤波，当模型比较大时，这种处理计算量非常大。

3.1.2 加入整体考量的显著性检测

全局看来，人们会对较少出现的区域敏感，所以显著的区域不仅仅是局部凸显，而且是全局稀少的。基于这种人类视觉机制，Wu 等[14]提出了加入全局稀有性检测的方法。

本文觉得 Wu 等[14]提出的方法有两点值得学习。

第一，不用曲率作为点的特征描述，而是采用了 Maximo 等[15]提出的局部高度图描述方法。如图 3.1.2 所示，在顶点处根据法向量作切平面，选取一定大小的区域。区域内的值由切平面与网格表面的垂直高度确定，改变区域的大小，便可以得到多尺度空间。

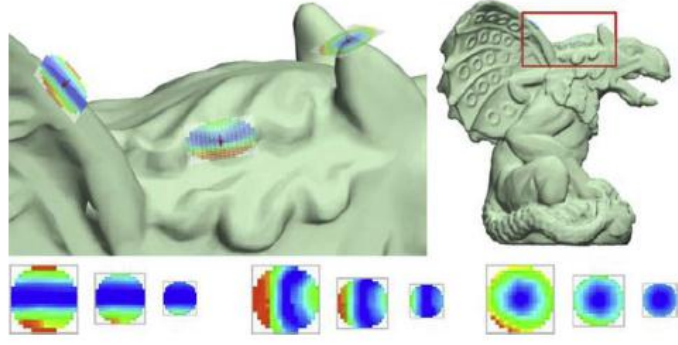


图 3.1.2 局部高度图的示例，来自[14]，下方为三种尺度下的局部高度图。

局部高度图的描述方法包含了更多点和邻域之间关系的信息，对局部高度图 Zernike 分解得到的系数向量作为点的特征描述。

第二，全局稀有性检测。理论上可以比较点与所有其他点的差异，然而这种算法复杂度太大。所以 Wu 等[14]先根据点的特征描述，用 K-means 方法聚类，先求类的显著性。

$$S_g(c_i) = \sum_{j=1}^k d(c_i, c_j) n_j \quad (3.1)$$

C_i 表示第 i 个类， $d(c_i, c_j)$ 表示两个类之间的特征描述的欧式距离， n_j 表示类 j 中的顶点个数。可见如果类 i 在全局稀有，则其他类包含的顶点个数会比较多，因此最后求得的类 i 的整体显著性也会比较大。

3.2 基于谱的显著性检测方法

Song 等[11]提出了谱分析检测网格显著性的方法。他其实用到了 Hou 和 Zhang[7]在图像显著性检测时的猜想，即频域中存在着冗余信息，去除这些冗余信息，得到偏差成分，这些偏差成分可能会反映显著性信息。在第五章将详细介绍方法的实现。

这种方法基于的原理还是一种猜想，缺少证明，Song 等[11]也没有在论文中解释清楚。

第四章 网格谱分析介绍

在用谱分析检测网格显著性之前，本章先简要介绍网格谱分析的概念，比如特征分解、谱变换等等。

4.1 光滑边框的例子

从一个例子来引入谱的概念。如图 4.1(1)所示，有一副海马图，边框由一系列的顶点和顶点之间的连线构成，现要光滑边框，可以怎么做呢？

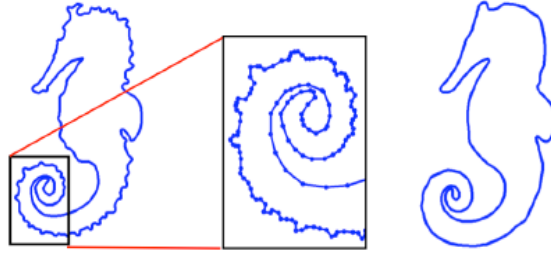


图 4.1(1) 海马图，来自[16]

方法可以是如图 4.1(2)

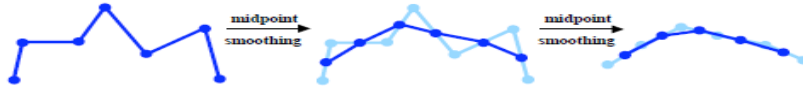


图 4.1(2) 边框光滑方法，先取原边框折线的中点，得到新折线图，然后再取中点，得到光滑后的边框，来自[16]

那么新的顶点

$$v_i^* = \frac{1}{4}v_{i-1} + \frac{1}{2}v_i + \frac{1}{4}v_{i+1} \quad (4.1)$$

定义拉普拉斯算子为

$$\delta(v_i) = 2 * (v_i^* - v_i) = \frac{1}{2}v_{i-1} - v_i + \frac{1}{2}v_{i+1} \quad (4.2)$$

在海马边框定义一个起点， \mathbf{X} 表示所有顶点的 x 坐标向量，我们可以得到如下等式。

$$\delta(X) = LX = \begin{bmatrix} 1 & -\frac{1}{2} & 0 & \dots & \dots & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \dots & \dots & 0 & -\frac{1}{2} & 1 \end{bmatrix} X.$$

图 4.1(3) 拉普拉斯矩阵, 来自[16]

可见 L 矩阵是对称矩阵, 而对称矩阵有很好的性质, 即它的所有特征向量构成一组正交基。 L 的最小的 8 个特征值对应的特征向量的绘图如图 4.3(4), 特征向量像正弦曲线一样震荡, 特征值大, 对应的特征向量的频率高。

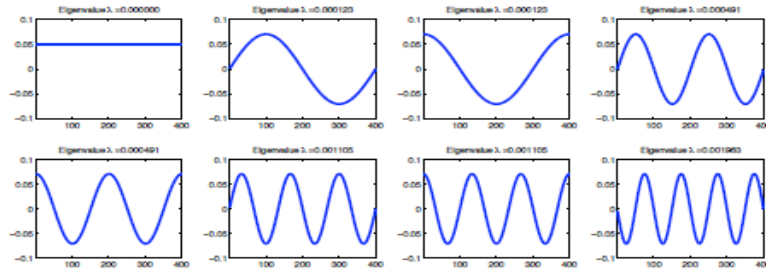


图 4.1(4), 特征向量的绘图。来自[16]

因为所有特征向量相互正交, 构成一组正交基, 因此任意 n 维向量都可以由 L 矩阵的 n 个特征向量线性表示。假设 $V_1 \dots V_n$ 表示特征向量

$$X = c_1 * V_1 + c_2 * V_2 + \dots c_n * V_n \quad (4.3)$$

类似于傅里叶变换, 我们可以用系数向量 $C = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ 表示 X , 称之为谱变换。

系数向量的绘图如图 4.1(5)

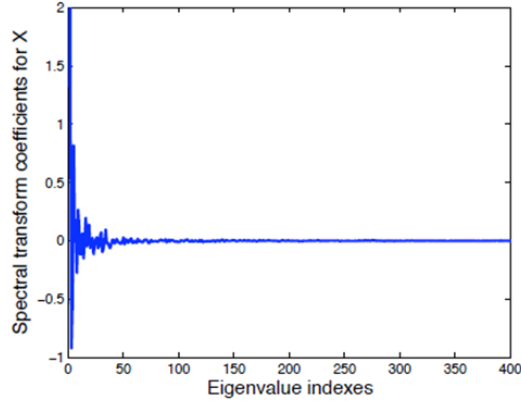


图 4.1(5) 海马边框 X 坐标对应的系数向量的绘图。来自[16]

可见，在低频部分谱变换的系数比较大，也可以说 X 坐标信号的能量集中在低频部分。那么用谱变换后的前 k 个系数也能大体表示边框。一种边框光滑算法即用前 k 个系数重构边框。

$$X = c_1 * V_1 + c_2 * V_2 + \cdots c_k * V_k \quad (4.4)$$

光滑结果如图 4.1(6)所示。

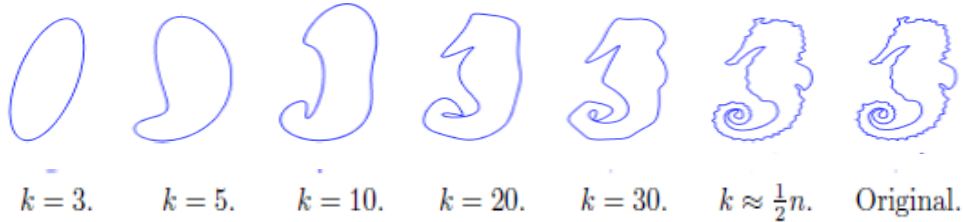


图 4.1(6) 前 k 个系数重构，得到的边框光滑结果，来自[16]

扩展到 3D 网格中，拉普拉斯矩阵 A 的元素与顶点之间是否相邻有关。

$a_{i,j} = 1$, 如果顶点 i 和顶点 j 相邻

$a_{i,j} = 0$, 如果顶点 i 和顶点 j 不相邻

$a_{i,i} = Degree$, Degree 为顶点 i 的度数

矩阵 A 也是对称阵，顶点的 x、y、z 坐标向量也可以由矩阵 A 的特征向量表示，按照相似的方式，对马模型的光滑效果如图 4.2(7)所示

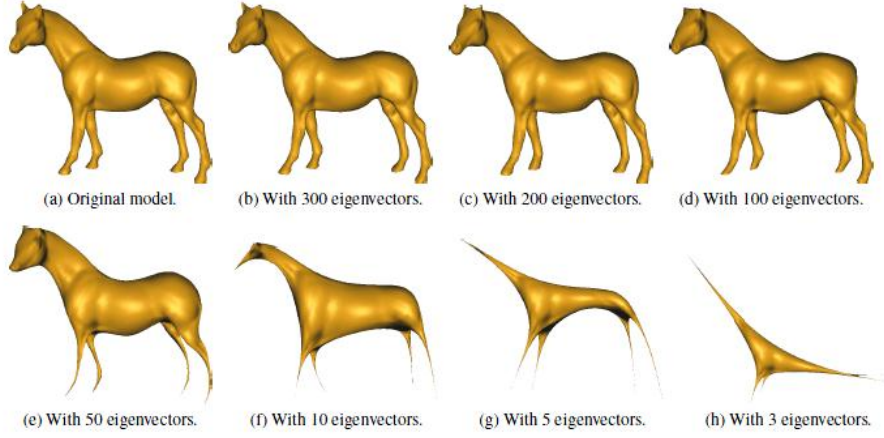


图 4.2(7) 马模型，由前 k 个系数重构结果。来自[17]

4.2 网格谱分析

4.2.1 基本概念

网格谱分析即用适当定义过的矩阵（算子）的特征结构进行几何分析。特征结构包括：

- 特征值和特征向量

$$Au = \lambda u \quad (4.5)$$

- 特征分解

$$A = ULU^T \quad (4.6)$$

- 谱变换

$$y = U\hat{y} \quad (4.7)$$

- Y 在前 k 个特征子空间的投影

$$y' = U_{(k)} \hat{y}_{(k)} = U_{(k)} U_{(k)}^T Y \quad (4.8)$$

其中， A 表示拉普拉斯矩阵， u 代表特征向量， λ 表示特征值， U 表示所有特征向量构成的矩阵， $U_{(k)}$ 表示 U 矩阵的前 k 列构成的矩阵， \hat{y} 表示谱变换系数向量。

4.2.2 谱分析的优点

通过谱变换，可以在另一个域中处理问题。相比于用坐标表示网格，算子可以蕴含网格的全局的内在的几何性质，适当的谱分析可以提取这些性质。而且有

时不需要所有的特征函数，利用那些主要的特征函数就可以，可以将维数简化。

图谱理论认为特征值和主要的全局图形形状密切相关，经常将它用作简洁的全局形状描述。而特征向量可以用于谱变换，这与传统的傅里叶变换有类似的地方。

4.2.3 谱分析的难点和挑战

谱分析与离散傅里叶变换有很大的不同

- DFT 的基在给定维度 n 后是确定的
- 谱变换与不同的拉普拉斯算子有关，特征函数随着算子而变化。
- 特征函数的规律性很有可能消失，难以描述特征函数的性质，Jain[18]证明了循环矩阵的特征向量才有正弦震荡的性质，如之前的图 4.1(4)所示
- 不容易比较两组特征函数的优劣
- 没有哪个拉普拉斯算子可以满足一系列令人满意的性质

第五章 基于谱的网格显著性分析

5.1 基于特征分解的显著性分析

本节主要实现 Song 等[11]的算法，尝试分析了原理，并且对论文中的某些地方提出了质疑，对公式进行了修改。

5.1.1 算法与分析

经典拉普拉斯矩阵 A ，元素按如下方式构造。

- (1) $a_{i,j} = -1$, 如果顶点 i 和顶点 j 相邻
- (2) $a_{i,j} = 0$, 如果顶点 i 和顶点 j 不相邻
- (3) $a_{i,i} = Degree$, $Degree$ 为顶点 i 的度数

然而，经典拉普拉斯矩阵只能反映网格的拓扑结构，我们要检测网格的显著性，同一拓扑结构的几何形状可以有很大不同，显然用经典拉普拉斯是不够的。所以，我们要加入顶点之间距离权重。顶点距离较远，说明他们相互间的影响较小，根据此，拉普拉斯矩阵 L 按如下方式构造。

- (1) $L_{i,j} = -\frac{1}{distance(i,j)^2}$, 如果顶点 i 和顶点 j 相邻， $distance$ 为欧式距离，
- (2) $L_{i,j} = 0$, 如果顶点 i 和顶点 j 不相邻
- (3) $L_{i,i} = -\sum_{j=1, j \neq i}^n L_{i,j}$

然后，一般还要对矩阵归一化，本文对原论文归一化方法存在疑问，他即要求每一行的和为 1，并且保持归一化后的矩阵是个对称阵，本文觉得不能实现，原文也没有给出具体的归一化公式。在此，本文采用 Zhang 等[17]介绍的带距离权重的拉普拉斯矩阵的一般归一化方法，再构造一个 D 矩阵

- (1) $D_{i,i} = L(i, i)$,
- (2) $D_{i,j} = 0$, 如果 $i \neq j$

拉普拉斯矩阵去掉对角线的矩阵 W 为

$$W = L - D \quad (5.1)$$

归一化后的拉普拉斯矩阵

$$L = D^{-1/2} L D^{-1/2} \quad (5.2)$$

即保证拉普拉斯矩阵 L 对角线上的元素为 1.

本文用 2000 个顶点的马模型构造了拉普拉斯矩阵，对原矩阵和归一化后的

矩阵分别计算了特征值，发现归一化后矩阵的特征值很小，在 0 到 2 之间，原矩阵的最大特征值达到 10^4 数量级。论文给出了 2000 个顶点的牛模型的拉普拉斯矩阵的特征值绘图（特征值从小到大排列），如图 5.1(1)所示，最大特征值达到了 10^4 数量级，因此本文怀疑，原论文没有对矩阵进行归一化操作，不管如何，后文用到的拉普拉斯矩阵没有归一化。

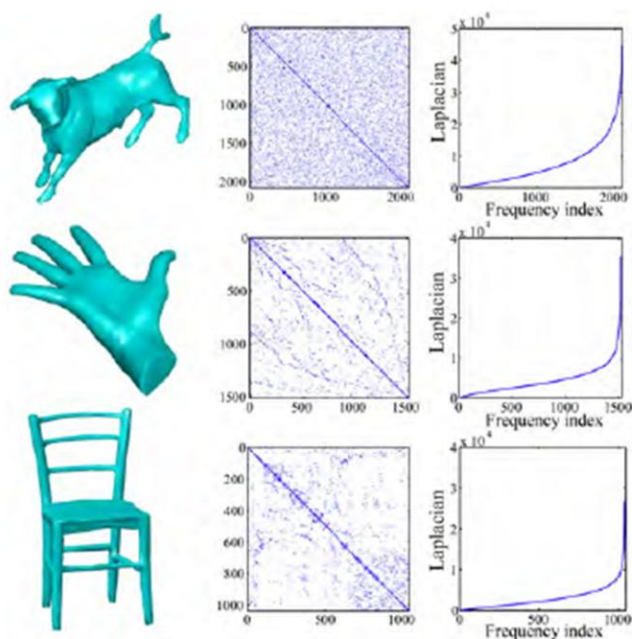


图 5.1(1) 一些模型的拉普拉斯矩阵，和对应的特征值的绘图。来自[11]

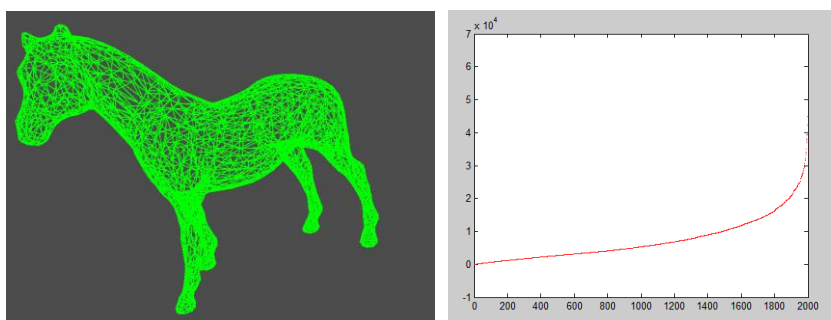


图 5.1(2) 2000 个顶点的马模型，与它的拉普拉斯矩阵的特征值的绘图

确定好拉普拉斯矩阵后，对拉普拉斯矩阵特征分解

$$L = BUB^T \quad (5.3)$$

U 对角线上的元素即为 L 的特征值。

构造一个新的离散函数

$$h(k) = \lambda_k, \quad (5.4)$$

特征值从小到大排列 $\lambda_1 < \lambda_2 < \dots < \lambda_n$ ，我们就将 $h(k)$ 称为拉普拉斯谱。

由图 5.1 (1) (2) 可见，虽然网格模型有差异，拉普拉斯矩阵互相之间也有很大的不同，但是拉普拉斯谱有相似的形状。谱间的类似性说明谱存在冗余信息，那么，去除冗余性，得到谱的偏差成分，也许能反应网格的显著性。

论文先对谱取对数，然后再求偏差成分。为什么要求对数呢？在第四章介绍到，较小特征值对应的特征向量能反映网格的主要信息，所以较小特征值的与常态值的偏差值也许更能反映网格主要的显著性信息，取对数可以增大较小特征值之间的变化幅度，而让减小较大特征值之间的变化幅度，所以，在求谱的偏差成分之前，本文先对谱取对数是有道理的。

$$h(k) = \ln(|h(k)|) \quad (5.5)$$

然而，拉普拉斯矩阵存在特征值 0，不能直接取对数。这里本文也对论文直接取对数的方法产生怀疑。对这种方法稍作修改，先对特征值加 1，再取对数。

$$h(k) = \ln(|h(k)| + 1) \quad (5.6)$$

得到谱后，用一个 n 维的算子 $[\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}]$ ，对谱滤波来求取谱的常态值，相当于求局部平均值，这里 n 取 9。

$$h(k)^* = \frac{h(k-4) + h(k-3) \dots \dots + h(k+3) + h(k+4)}{9} \quad (5.7)$$

$$r(k) = h(k) - h(k)^* \quad (5.8)$$

得到偏差成分后，构造 R 矩阵

$$(1) R_{i,i} = \exp(r(k)),$$

$$(2) R_{i,j} = 0, \text{ 如果 } i \neq j$$

显著性矩阵 S 按如下方式构造

$$S = BRB^TW \quad (5.9)$$

W 矩阵由公式(5.1)构造，乘矩阵 W 是为了减少相距远的点以及不相连的点之间的相互影响

注意到拉普拉斯矩阵 L 的每一行对应着网格中的一个顶点，类似的， S 矩阵的每一行也对应着网格中的一个顶点，顶点 i 的显著性值可以用 S 矩阵第 i 行的和表示。

论文然后用类似 Lee 等[3]中的多尺度模型进一步改善显著性结果，但是，算法比较复杂，本文还没能实现。

5.1.2 显著性检测结果

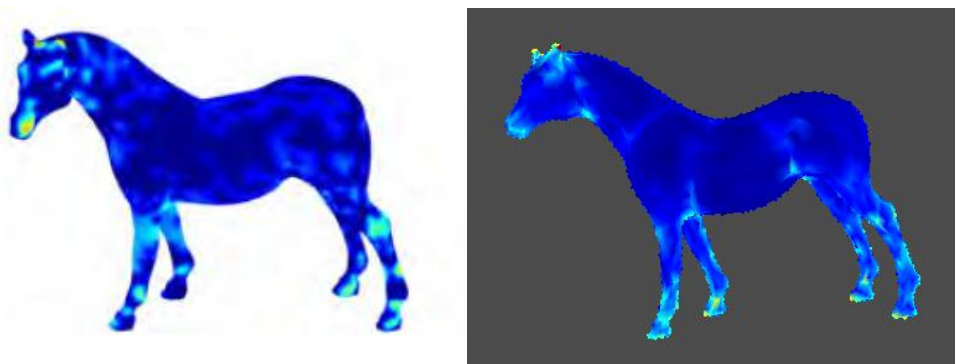


图 5.2.1(1) 马模型的显著性结果，左图来自[11]，右图是实现结果

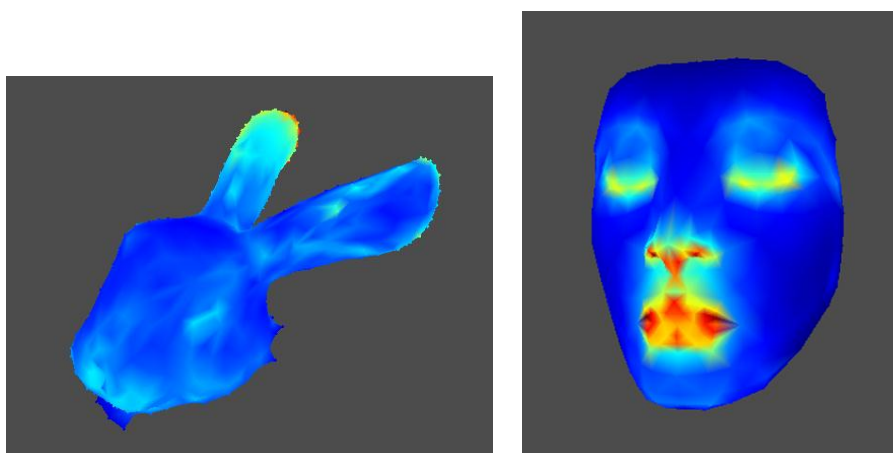


图 5.2.1(2) 左图兔子头部模型的显著性结果，右图是人脸模型

由图 5.2.1 (1-2)可见，马模型的头部、脚部显著性较高，兔子头部的耳朵、嘴、眼睛显著性较高，人脸的嘴、鼻子、眼睛显著性较高，基于特征分解的显著性检测方法在这些模型上取得较好的结果。

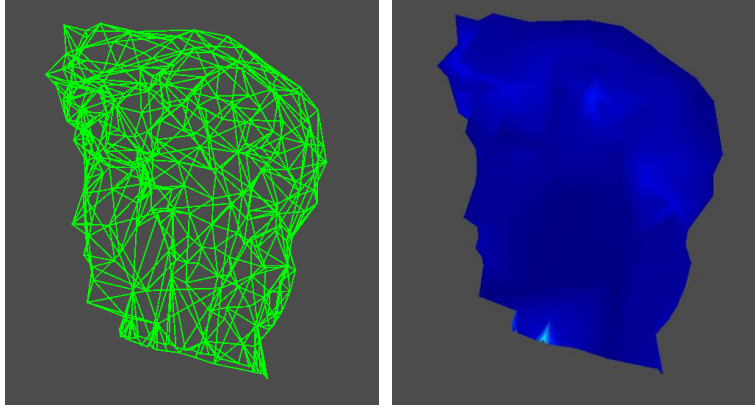


图 5.2.2 人类头部模型

由图 5.2.2 所示，在另一个人类头部模型上，显著性检测结果很差，原因可能是，这个模型头部整体都比较突出，包括发型、鼻子、下巴等等，人们可以根据常识判断它是一个人的头部，然而单纯靠拉普拉斯矩阵的几何信息，不能很好的检测显著性。

5.2 基于谱变换的显著性分析

本节提出一种新的基于谱变换的显著性分析算法。

5.2.1 算法与分析

在 5.1 节已经说明，经典拉普拉斯算子只能反映网络的拓扑结构，单单用拓扑结构是不能分析显著性的，所以本章的拉普拉斯矩阵 L 按如下方式构造。

(1) $L_{i,j} = -\frac{1}{distance(i,j)^2}$, 如果顶点 i 和顶点 j 相邻, $distance$ 为欧式距离,

(2) $L_{i,j} = 0$, 如果顶点 i 和顶点 j 不相邻

(3) $L_{i,i} = -\sum_{j=1, j \neq i}^n L_{i,j}$

L 是对称阵, 所以所有的特征向量构成一组正交基。记 L 矩阵的特征值为 λ_k , 从小到大排列。 V_k 是特征值对应的单位特征向量。 X 是所有顶点的 x 坐标向量, Y 是所有顶点的 y 坐标向量, Z 是所有顶点的 z 坐标向量。 X, Y, Z 均可以由 V_k 线性表示。

$$X = \sum_{i=1}^N a(i) * V(i) \quad (6.1)$$

$$Y = \sum_{i=1}^N b(i) * V(i) \quad (6.2)$$

$$Z = \sum_{i=1}^N c(i) * V(i) \quad (6.3)$$

谱变换后，X 对应的系数向量为 A，Y 对应的系数向量为 B，Z 对应的系数向量为 C，马模型的系数向量的绘图如图 6.1(1)所示。

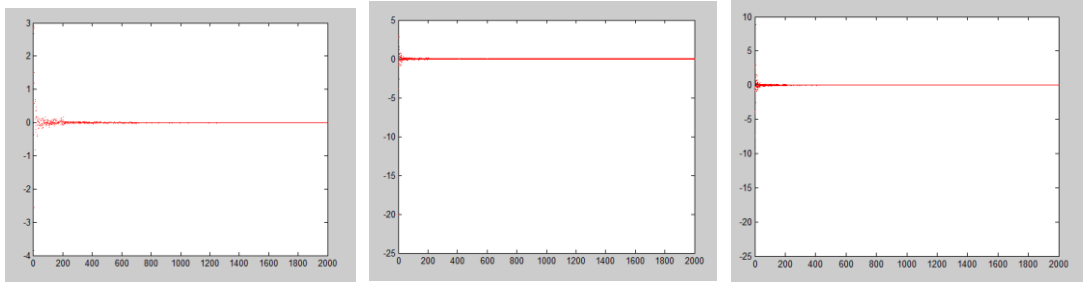


图 6.1(1) 左、中、右图分别对应系数向量 A、B、C

这里把系数向量称为谱，可见，谱的能量集中在低频部分。在第四章中介绍到，可以用前 k 个系数重构网格，假设顶点 $V(i)$ 原坐标为 (x,y,z) ，重构后得到的坐标为 (x^*,y^*,z^*) ，新旧坐标间的欧氏距离就作为顶点的显著性值。K 取 100，结果如图 6.1(2)所示

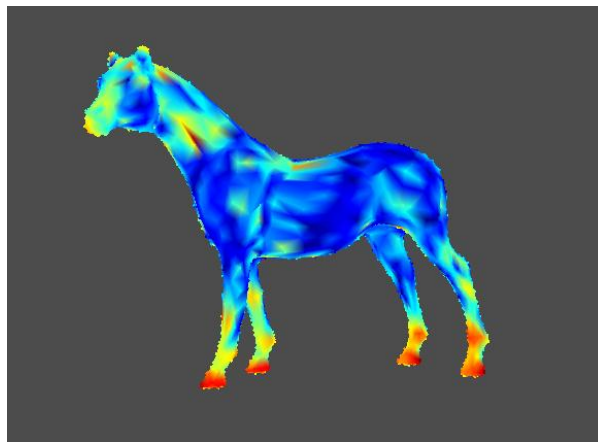


图 6.1(2) k 取 100，显著性结果，结果比较粗糙。

结果能反应一部分显著性信息，但是显著性结果不光滑。本文认为，用前 100 个系数重构，相当于对谱用了一个理想低通滤波器 H 滤波。其中当 i 小于等于 100 时， $h(i)=1$ ，当 i 大于 100 时， $h(i)=0$ 。

$$X = \sum_{i=1}^N a(i) * h(i) * V(i) \quad (6.4)$$

理想低通滤波器 H 的绘图如图 6.1(3)所示

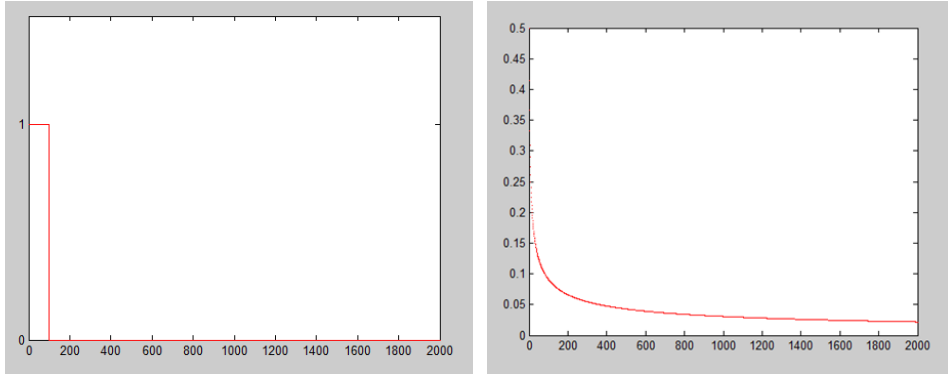


图 6.1(3) 左图是理想低通滤波器，右图是光滑低通滤波器

为了使显著性结果也较为光滑，我们可以选择光滑的低通滤波器。模仿 Butterworth 滤波器的设计，本文把低通滤波器设置为

$$H(i) = \frac{1}{1 + i^{0.5}} \quad (6.5)$$

光滑低通滤波器的绘图如图 6.1(3)所示

确定好滤波器后，按公式(6.6)-(6.8)重构顶点的三维坐标，依旧把新旧顶点间的欧氏距离作为显著性值，显著性值可用公式(6.9)求取。

$$X = \sum_{i=1}^N a(i) * h(i) * V(i) \quad (6.6)$$

$$Y = \sum_{i=1}^N b(i) * h(i) * V(i) \quad (6.7)$$

$$Z = \sum_{i=1}^N c(i) * h(i) * V(i) \quad (6.8)$$

$$S(V_i) = [(x(i) - x(i)^*)^2 + (y(i) - y(i)^*)^2 + (z(i) - z(i)^*)^2]^{1/2} \quad (6.9)$$

构造带距离权重的拉普拉斯矩阵，然后进行谱变换，对谱滤波，重构网格，其实就是对网格进行光滑操作。这种基于谱变换的算法，其实是比较原网格和光滑后的网格，那些坐标变化比较大的顶点，给予它们较大的显著性值。

网格上比较突出的部分在光滑过程中会变化比较大，因此，这种算法趋向于能找到网格上的突出部分。

5.2.2 显著性检测结果

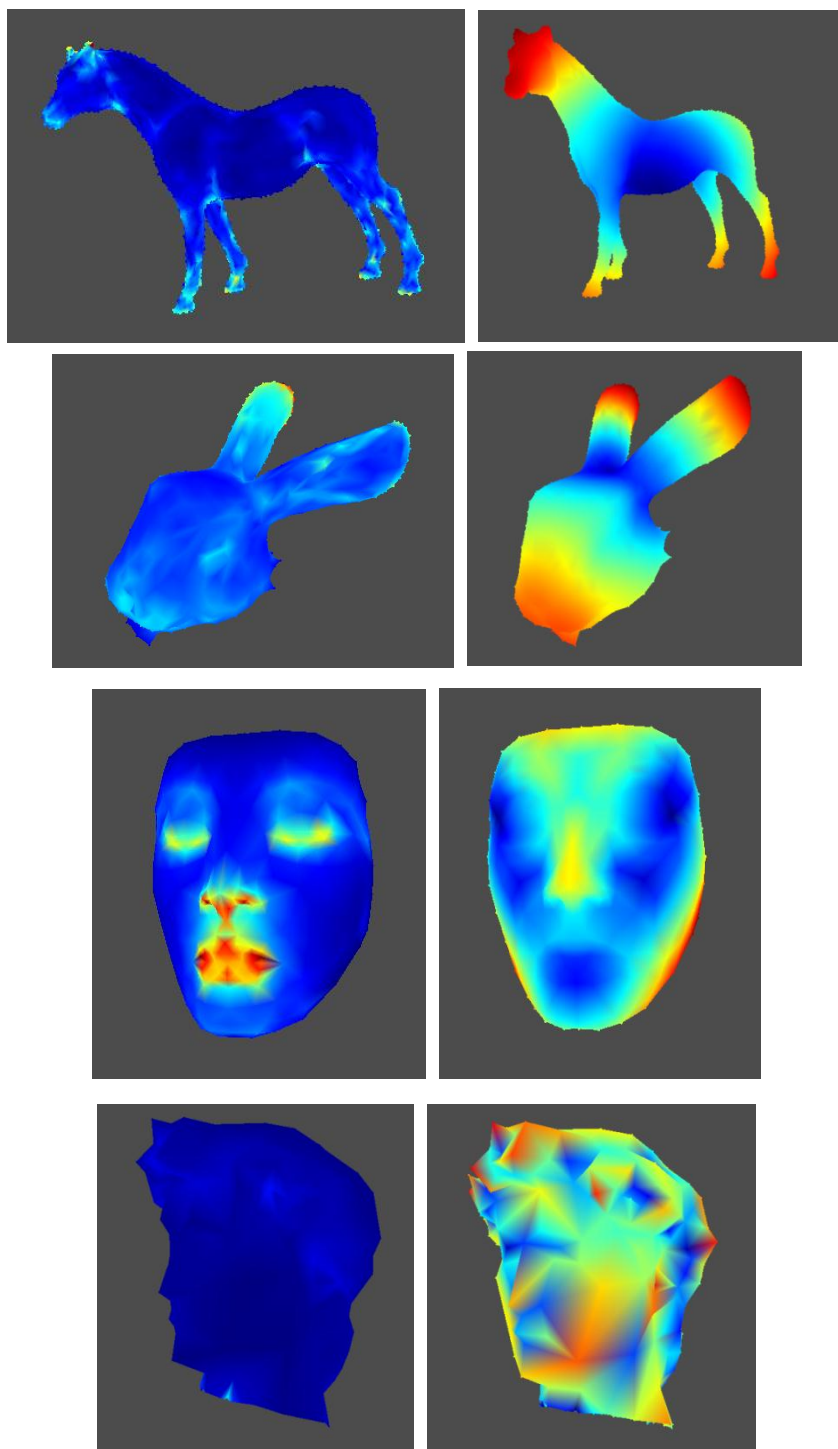


图 6.2 左边是按 Song[11]实现的结果，右边是基于谱变换算法的结果

可见，马模型、兔子头部模型、包括人的头部模型，基于谱变换的算法能得到比较符合直观的显著性结果。然而这种算法的局限性也十分明显，它不能检测出比较细节的部位，比如马的耳朵、兔子的眼睛等等，而且它在人脸模型的效果并不好，算法检测出突出的鼻子和边缘部分，没能检测出眼睛嘴巴的显著性。

5.3 基于谱残差成分的显著性分析

本节再提出一种新的算法，在谱的层面进行网格显著性分析。

5.3.1 算法与分析

Hou 和 Zhang[7]提出了从频域分析图像显著性的方法。这种方法基于的想法就是：谱中的相似性表明存在冗余，去掉谱中的冗余成分，得到的残差成分也许能反应原图像中突出的区域。

图像中的特征主要有颜色、亮度、方向等，网格中的特征是网格的几何形状，在本章中，我们用顶点处的曲率表示顶点处的形状特征。当前已经有很多方法计算网格中顶点处的曲率，本文参考了 Dong and Wang[19]的方法，实现了曲率计算。在马模型上的曲率结果如图 7.1(1)所示

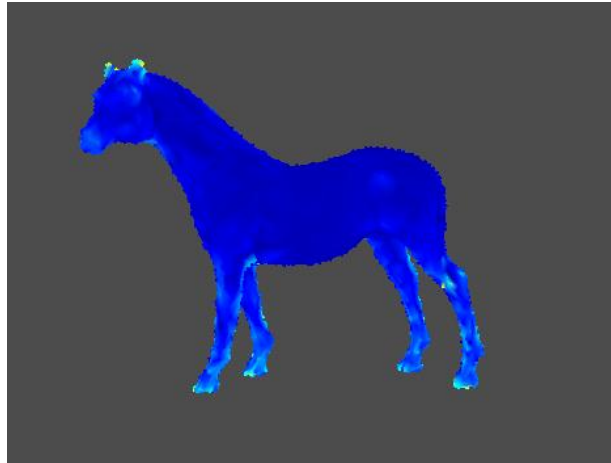


图 7.1(1) 马模型表面的曲率

本文已经在 3.1.1 节指出，只把曲率当作显著性成分，会造成图 3.1.1(1)的问题，即某个区域的曲率都比较大时，这个区域的显著性不一定大，那些与周围区域不同的、突出的局部才会引起人的注意。

Lee 等[3] 对曲率用了中心环绕的高斯加权算子来改善结果，本文不从空域分析的角度，而是尝试用谱的残差成分来检测网格的显著性。

本章的拉普拉斯矩阵 L 按如下方式构造。

$$(1)L_{i,j} = -\frac{1}{\text{distance}(i,j)^2}, \text{如果顶点 } i \text{ 和顶点 } j \text{ 相邻, distance 为欧式距离,}$$

(2) $L_{i,j} = 0$, 如果顶点 i 和顶点 j 不相邻

(3) $L_{i,i} = -\sum_{j=1, j \neq i}^n L_{i,j}$

L 是对称阵, 所以所有的特征向量构成一组正交基。记 L 矩阵的特征值为 λ_k , 从小到大排列。 V_k 是特征值对应的单位特征向量。 C 是顶点处的曲率向量, 那么可以由 V_k 线性表示。

$$C = \sum_{i=1}^N a(i) * V(i) \quad (7.1)$$

通过谱变换, 得到曲率向量对应的谱变换系数 a 。

得到谱后, 用一个 n 维的算子 $[1/n, 1/n \cdots 1/n]$, 对谱滤波来求取谱的常态值, 相当于求局部平均值, 这里 n 取 9。

$$a(k)^* = \frac{a(k-4) + a(k-3) \dots \dots + a(k+3) + a(k+4)}{9} \quad (7.2)$$

求取谱中的残差成分。

$$r(k) = a(k) - a(k)^* \quad (7.3)$$

通过反变换, 将谱中的残差成分映射为网格中的显著性值

$$S = \sum_{i=1}^N r(i) * V(i) \quad (7.4)$$

马模型的显著性结果如图 7.1(2)所示

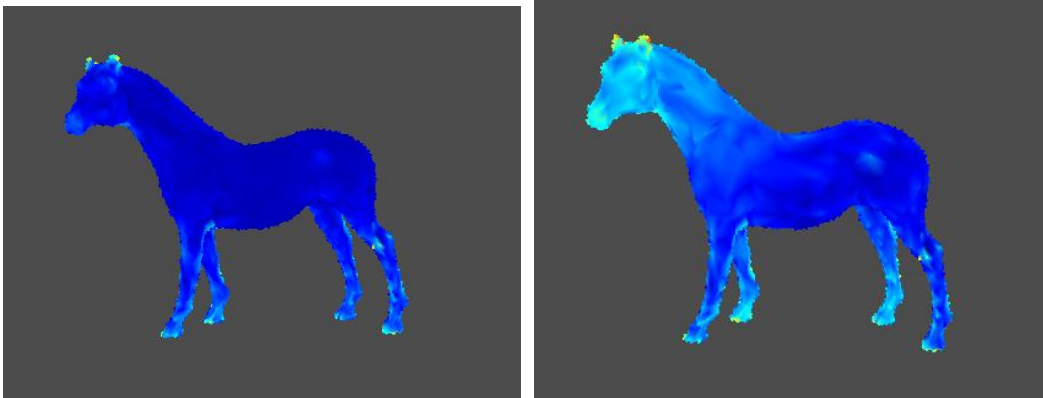


图 7.1(2) 左图是马模型表面的曲率, 右图是显著性结果

5.3.2 显著性检测结果

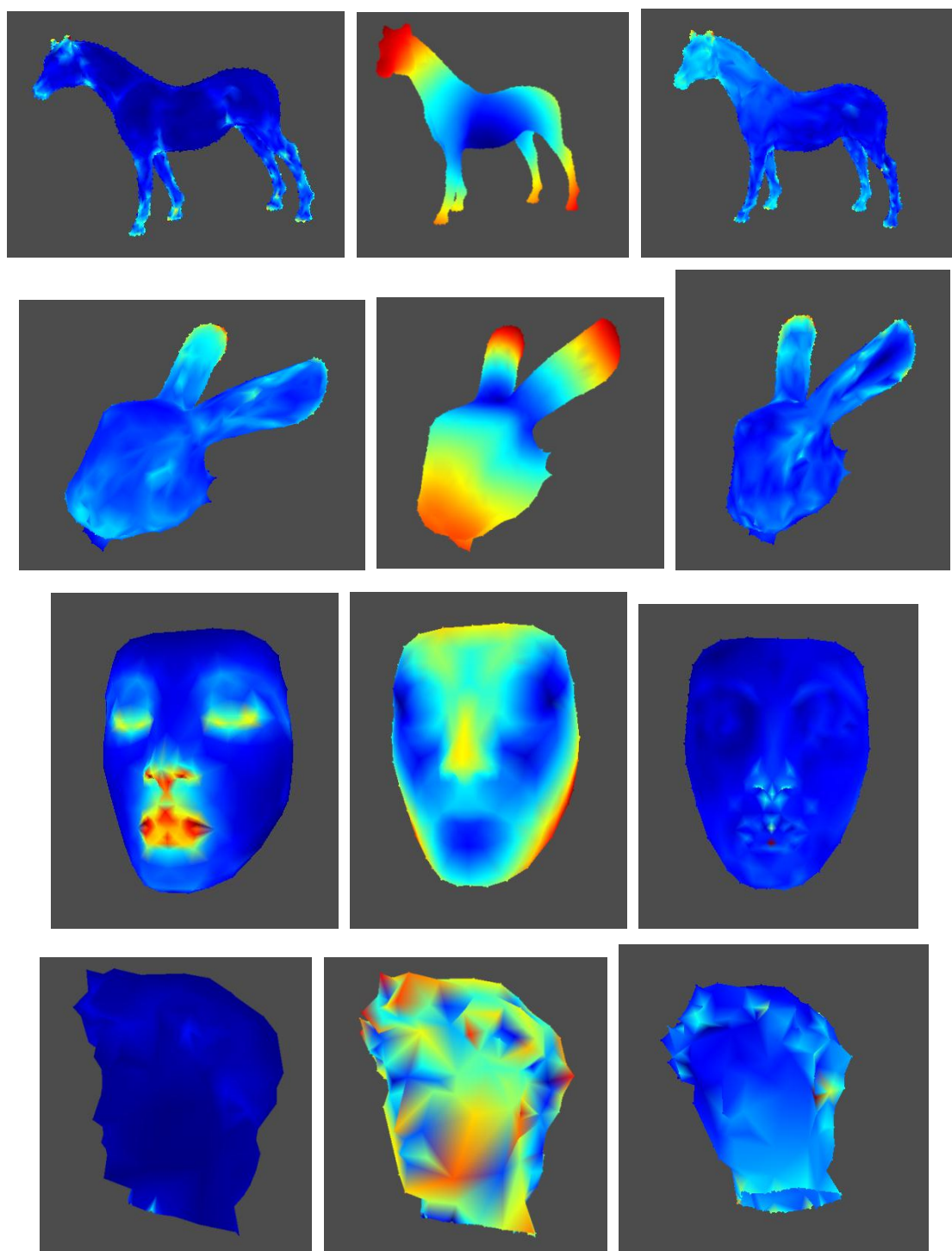


图 7.2 左边是按 Song 等[11]实现的结果，中间是基于谱变换算法的结果，右边是基于残差成分算法的结果

我们将结果和按论文[11]实现的结果比较，虽然方法不同，但是马模型和兔子头部模型上的结果较为接近。在马模型上，新提出方法的检测结果在整体头部上的显著性都比身体部分高，这也是符合直观感觉的，而且也能分辨出嘴、眼睛、耳朵。[11]的方法并没有给予头部相对于身体较高的显著性成分，从这点上来说，新提出的方法更符合直观感受。然而在人脸模型上，新方法的结果并不满意。

第六章 总结与展望

本文先调研了图像、网格显著性分析在近几十年来的发展，解释了传统的显著性检测模型所基于的原理，即人的视觉注意力机制。采用谱分析的显著性检测方法的原理还是一种猜想，即谱中存在冗余信息，与常态值的偏差信息也许能反应显著性。同时，本文也指出在网格显著性检测领域，缺乏对模型的评价工作。根据人眼追踪等方式建立真实的数据集、完善评价指标也应成为一项研究内容。

谱分析已经在网格几何处理中得到很大的应用，本文用了基于特征分解、谱变换、谱残差的方法，尝试检测网格的显著性。从结果可以看出，每种方法都有优点和局限性。基于特征分解的方法能检测比较细节的部分，整体上的检测结果也比较好，然而在一些模型上没能检测出显著性。基于谱变换的方法的检测结果比较明显，但是不能检测出网格中有特点的的细节部分，而且有的模型的显著性检测结果不符合人的直观感受。基于谱残差成分的方法的检测结果与基于特征分解方法的结果较为接近。本文指出了谱变换的方法所依据的原理，这种算法其实是比较原网格和光滑后的网格，那些坐标变化比较大的顶点，给予它们较大的显著性值。网格上比较突出的部分在光滑过程中会变化比较大，因此，这种算法趋向于能找到网格上的突出部分。谱残差成分的方法的原理还是一种猜想，它认为谱中与常态值的偏差信息也许能反应显著性，怎样证明以及清楚得解释背后的原理是未来的一项研究内容。

显著性检测是以人的直观感受为衡量标准，有时候人们会根据已有的知识或经验去关注网格的显著区域。目前已有的网格显著性检测方法只是根据网格的几何性质判断显著性，怎样将机器学习的手段用于网格显著性检测，会是未来一项有挑战的研究内容。检测出显著性后，怎样用显著性结果指导网格简化、网格分割、视角选取等等应用也很有研究价值。

参考文献

- [1] P. Cignonia, C. Montania and R. Scopigno. A comparison of mesh simplification algorithms. *Computers & Graphics*, 22(1): 37–54, 1998,
- [2] T.R. Jones, F. Durand and M. Desbrun. Non-iterative, feature-preserving mesh smoothing. *ACM Transactions on Graphics (TOG)*, 22(3): 943-949, 2003,
- [3] C.H. Lee, A. Varshney and D.W. Jacobs. Mesh saliency. *ACM Transactions on Graphics* 24(3): 659–666, 2005.
- [4] L. Itti, C. Koch and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(11): 1254-1259, 1998.
- [5] T. Judd, K. Ehinger, F. Durand and A. Torralba. Learning to predict where humans look. *IEEE International Conference on Computer Vision*, 2106 -2113, 2009.
- [6] <http://www.baik.com/wiki/ROC%E6%9B%B2%E7%BA%BF>
- [7] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. *IEEE Conference on Computer Vision and Pattern Recognition*, 1-8, 2007
- [8] <http://www.its.caltech.edu/~xhou/>
- [9] W. Kienzle, M.O. Franz, B. Scholkopf, and F.A. Wichmann. Center-Surround Patterns Emerge as Optimal Predictors for Human Saccade Targets. *Journal of Vision*, 9:1-15, 2009
- [10] A. Borji, and L. Itti. State-of-the-Art in Visual Attention Modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185-207, 2013
- [11] R. Song, Y. Liu, R. Martin and P. Rosin. Mesh saliency via spectral processing. *ACM Transactions on Graphics*, 33(1): 1-17, 2014
- [12] P. Shilane, T. Funkhouser. Distinctive regions of 3D surfaces. *ACM Transactions on Graphics*, 26(2): 7, 2007
- [13] G. Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. *IEEE International Conference on Computer Vision*, 902-907, 1995.
- [14] J. Wu, X. Shen, W. Zhu, L. Liu. Mesh saliency with global rarity. *Graphical*

- Models, 75(5):255-264, 2013.
- [15]A. Maximo, R. Patro, A. Varshney and R. Farias. A robust and rotationally invariant local surface descriptor with applications to non-local mesh processing. Graphical Models, 73(5):231-242, 2011
- [16]B. Levy and H. Zhang. Spectral Mesh Processing. SIGGRAPH 2010 Course.
- [17]H. Zhang, O. Van Kaick and R. Dyer. Spectral Mesh Processing. Computer Graphics Forum, 29(6):1865-1894, 2010.
- [18]A. K. Jain. Fundamentals of Digital Image Processing. Prentice Hall, Inc., 1989.
- [19]C. Dong and G. Wang. Curvatures estimation on triangular mesh. Journal of Zhejiang University Science, 6(1): 128-136, 2005.