

一种基于网格参数化的图像适应方法^{*}

时 健¹, 郭延文¹⁺, 杜振龙^{2,3}, 张福炎¹, 彭群生²

¹(南京大学 计算机软件新技术国家重点实验室,江苏 南京 210093)

²(浙江大学 CAD&CG 国家重点实验室,浙江 杭州 310058)

³(南京工业大学 信息科学与工程学院,江苏 南京 210009)

A Mesh Parameterization-Based Image Retargeting Method

SHI Jian¹, GUO Yan-Wen¹⁺, DU Zhen-Long^{2,3}, ZHANG Fu-Yan¹, PENG Qun-Sheng²

¹(National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

²(State Key Laboratory of CAD&CG, Zhejiang University, Hangzhou 310058, China)

³(College of Information Science and Engineering, Nanjing University of Technology, Nanjing 210009, China)

+ Corresponding author: E-mail: ywguo@nju.edu.cn

Shi J, Guo YW, Du ZL, Zhang FY, Peng QS. A mesh parameterization-based image retargeting method. *Journal of Software*, 2008,19(Suppl.):19–30. <http://www.jos.org.cn/1000-9825/19/s19.htm>

Abstract: Image retargeting is the process of adapting images to display terminals with small sizes and different aspect ratios, such as cellular phones and PDAs. This paper presents a novel image retargeting method using saliency-based mesh parameterization. Specifically, this paper formulates retargeting an image to desired size as a constrained mesh parameterization problem which aims at finding a homomorphous target mesh with desired size. This method first constructs a mesh image representation that is consistent with the underlying image structures. To emphasize salient objects and minimize visual distortion, this paper associates image saliency into the mesh and defines image structure as mesh parameterization constraints. Through a stretch-based mesh parameterization process, this paper finally achieves the homomorphous target mesh, which is then used to render the target image by texture mapping. This method generates satisfactory retargeting effects for images with complex background structures. Also it works well for images with multiple salient objects. Experimental results demonstrate the effectiveness of the proposed method.

Key words: image retargeting; mesh parameterization; attention model

摘 要: 图像适应是指将高分辨率的数字图像显示在手机、PDA 等屏幕较小的显示终端上的过程。提出一种全新的基于网格参数化的图像适应方法,该方法的关键在于把图像表示为特征网格,从而将图像适应问题转化为网格的参数化,即求取一个与该特征网格同拓扑,且具有目标屏幕尺寸的网格。为了突出图像中的重要物体,该方法建立了源图像对应的特征网格与图像显著度的关系;通过优化基于显著度伸长的网格参数化的能量来求解适应图像的网格;然后借助纹理映射生成适应图像。另外,该方法在参数化的过程中增加了对显著区域和背景结构的约束,能够在保持并增强图像中重要物体的同时,使适应图像的结构不发生明显形变。该方法能够方便地处理具有复杂背景和包含多目标物体的图像的适应问题。实验结果显示了该方法的有效性。

* Supported by the National Natural Science Foundation of China under Grant Nos.60703084, 60723003 (国家自然科学基金); the Foundation of Jiangsu Province of China under Grant No.BK2007571 (江苏省基金)

Received 2008-04-20; Accepted 2008-11-14

关键词: 图像适应;网格参数化;用户关注度

随着信息时代的来临,数码相机、摄像机等电子产品正日益普及,人们能够非常方便地利用它们拍摄出高分辨率的数字图像.为了信息共享和交流的需要,人们通常需要把这些图像显示在手机、PDA 等移动设备上.但由于源图像分辨率和显示终端屏幕尺寸的差异,一个关键的问题是如何在快速显示图像内容的同时突出图像中的显著区域或重要物体信息.这个问题被称作图像适应,即如何将高分辨率的图像更好地显示在较小的屏幕上.

相关工作

图像适应的一个传统方法是直接缩放图像尺寸以适应目标屏幕,这也是目前大部分手机、PDA 等移动终端广泛采用的解决方案.然而当原始图像分辨率和目标屏幕尺寸比例不匹配时,通过直接缩放所获得的适应图像会产生整体的比例失真,导致重要物体的严重扭曲.

有效的图像适应算法应该能够在反映图像基本内容的基础上突出图像中的重要信息,从而保证观察者对图像的正确理解.为了突出图像中的重要物体,现有的方法一般基于用户的视觉关注度(visual attention)来标识图像上各个区域的重要度,然后针对各区域重要度的差异分别进行不同的处理.

现有的图像适应方法可以大致分为两类.一类方法在不同的图像区域间使用导航功能^[1,2],根据各区域重要度的优先级通过手工交互或自动序列化显示这些区域,重要的区域首先显示,然后顺序显示非重要区域.但是这类方法不适合对大量图像的快速浏览.为了克服这个缺点,另外一类方法利用图像信息的冗余性,尝试通过裁剪或非均匀区域变形一次显示整幅图像,例如文献[3-5]中提出的方法仅仅关注图像中最重要的区域,裁剪掉周围的冗余信息而生成适应图像;此类方法往往完全丢弃图像中大量的背景信息,不利于用户对图像的理解.还有一种选择是对图像应用类似鱼眼效果的非均匀变换^[6],从而增强图像中的显著物体,同时压缩其余部分,该方法的不足之处在于只能处理仅包含单个显著物体且位于图像中心的情况,并且容易扭曲边界附近的图像结构.

图像的背景信息有助于用户对图像的正确理解,因此一个行之有效的图像适应方法应当能够在适当增强图像重要物体的同时尽量保持其背景信息,对于背景复杂的图像还应当保持其所包含的显著结构信息,目前还不存在这样的方法.另外,如何有效地对包含多个显著物体的图像进行适应也是图像适应问题的一个难点,其主要困难在于有效定位所有的显著物体并对它们按比例地进行增强.一种方法是通过图像分割技术首先提取出图像中的重要区域,然后将它们组合到放缩过的背景图像上来实现^[7],但该方法中所采用的技术如图像分割和修复等往往比较复杂且不够鲁棒.

本文工作

针对现有方法存在的不足,本文提出一种全新的基于网格参数化的图像适应方法.该方法根据图像区域的重要度能够自适应、非均匀地适应图像中多个重要区域,显著的区域相对于其他区域被明显增强和突出.该方法首先将源图像表示为特征相关的网格;然后将网格与图像的重要度检测结果相关联,从而提取出对应图像中重要物体区域的网格;进而通过基于重要度伸长的参数化方法求解出一个具有目标屏幕尺寸的网格;最后借助纹理映射技术生成适应图像.另外,该方法还将对边界、重要区域和背景结构的保持作为约束,并将其融合到参数化的能量方程中,从而在保持并增强图像重要物体的同时使适应图像背景结构不发生明显变形.

本文方法的主要优势在于:一方面,本方法可以在增强显示图像重要区域的同时尽量保持背景所包含的重要结构信息,这有助于人们对图像正确的视觉感知和理解;另一方面,该方法在不借助图像分割技术的前提下,可以方便地对包含多个重要物体的图像进行有效适应.

本文第 1 节介绍图像区域重要度的计算方法.第 2 节详细说明基于网格参数化的图像适应方法.第 3 节给出实验结果和分析.最后总结本文所提出的算法并展望未来工作.

1 关注度计算

图像适应应当在保持图像基本信息完整的同时充分突出图像中存在的重要物体或显著区域.而图像区域的重要程度可以通过视觉关注度模型计算评估或应用人脸检测等高层检索算子检测获得.

视觉关注度本质上是神经生物学方面的概念,被用来刻画观察者对于图像的认知程度^[8].建立评估视觉关注度的可计算模型已有大量的研究工作^[3,9-11],这些模型一般基于低层的图像亮度、对比度并结合高层的图像内容和场景理解,通过一些启发式规则来计算图像区域的重要度.Itti 等人结合多尺度图像特征提出了一种以生物认知体系为指导的视觉关注度计算模型.Ma 等人借助一系列关于人眼视觉的启发式规则提出了一种基于图像对比度的关注度评价算法,由于该算法简单实效而被广泛采用.本文利用该算法为图像中的每个像素计算一个规则化的视觉关注度值,具有高关注度值的图像区域即对应用户感兴趣的重要物体.

人脸和人体通常是我们日常生活中所拍摄的照片中最常见、最显著的事物主体,因此对于此类图像的适应应当保持图像中存在的人体.本文方法利用 Viola-Jones 提出的基于 Adaboost 的算法^[12]检测人脸;基于此进一步利用文献^[13]的算法提取大致的人体区域,该算法借助人脸检测的先验知识可以自动地从包含正面头肩信息的图像中把人体提取出来,其主要步骤如下:

- 利用改进的 Adaboost 算法检测人脸及人脸特征;
- 确定覆盖主要人体区域的逼近头肩轮廓掩体;计算掩体和周围背景的颜色分布;
- 采用迭代的图切割算法提取人体区域.

由于我们只需要确定近似的人体区域,为了加速计算,算法在下采样的图像上进行处理,对于一般尺寸的图像能够非常快速地确定大致人体区域.当确定人体区域后,所有人体内部像素的关注度值被赋值为 1.此外,系统还允许用户交互指定一些不能自动识别或关注度模型检测不准确的特殊重要物体.我们称通过关注度计算得到的反映图像不同区域重要程度的图为图像的显著度图,显著度图上每点的值介于 $[0,1]$ 之间,值越高表示用户的关注度越高.

2 基于网格参数化的图像适应

图像适应尝试把高分辨率的源图像显示在分辨率和横宽比例不同的显示屏幕上,适应后的图像被称为目标图像.本文算法的主要思想是把图像适应归结为对源图像网格进行参数化的过程,通过该过程使图像中的重要物体被严格保持并相对放大突出,而周围的背景区域不发生明显的变形扭曲.图 1 给出了本文方法的主要流程,我们首先建立能够反映源图像特征结构的网格并把该网格和检测源图像得到的显著度图建立关联(第 2.1 节);图像适应相应地转化为求取一个与该网格同拓扑且具有目标屏幕尺寸大小的网格.为了提升适应图像的视觉质量,方法进一步把边界约束、显著度约束和结构约束融合到待优化的目标方程中(第 2.2 节);然后通过优化一个基于显著度伸长的网格参数化能量求解目标图像的对应网格(第 2.3 节);最后借助理映射获取目标图像.

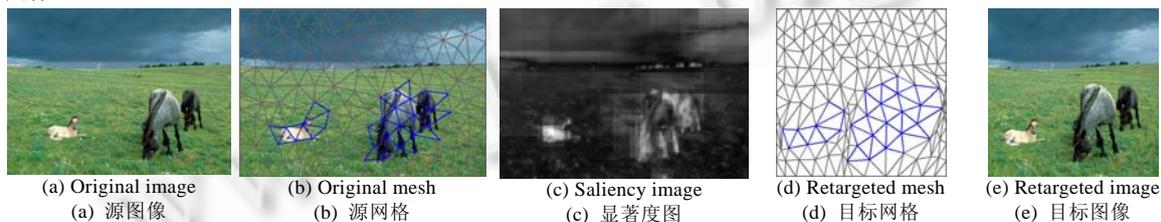


Fig.1 Algorithm overview
图 1 算法流程

2.1 图像的网格表示

我们首先生成与源图像特征结构相关的控制网格,一般来讲,图像上颜色梯度较大的点被近似视为特征点.为此,算法首先采用等间距的点离散化源图像 I_S 的边界;然后使用边缘检测算子如 Canny 算子提取图像中

的部分特征点,同时为了保持网格点密度的均匀分布,进一步在特征点稀疏的地方通过距离阈值的控制添加一些辅助点;最后采用约束的 Delaunay 三角化算法产生特征一致的网格 M_S (如图 2 所示).



Fig.2 Mesh generation
图 2 源图像的网格表示

为了描述方便,我们定义一些与网格相关的符号表示. M_t 表示待求解的目标网格; $\{P_i=(x_i,y_i)|i=1,\dots,n\}$, $\{Q_i=(u_i,v_i)|i=1,\dots,n\}$ 分别表示 M_S 和 M_t 的顶点, Q_i 表示 M_t 中的待求点, Q_i 对应于 M_S 中的点 P_i ; $\Delta_P=\Delta(P_i,P_j,P_k)$ 和 $\Delta_Q=\Delta(Q_i,Q_j,Q_k)$ 分别表示 M_S 和 M_t 中对应的三角形;由于我们要求 M_t 和 M_S 具有相同的拓扑结构,因此 M_t 中的边 $e(Q_iQ_j)$ 对应 M_S 中的边 $e(P_iP_j)$.如图 1 所示,(b)为源网格 M_S ,而(d)为参数化生成的目标网格 M_t .

为了区别对待图像中具有不同显著度的各个区域,我们将源网格 M_S 和源图像 I_S 的显著度图相关联,其中三角形 Δ_P 的显著度值 S_{Δ_P} 通过平均 Δ_P 内所有像素的显著度值求得,这样在网格参数化过程中赋予显著三角形的边较小的收缩程度,即可达到突出显著物体的目的.

2.2 参数化过程中的约束

图像适应应该完全保持图像中的显著物体,避免对它们的扭曲形变,同时尽可能地对其进行突出显示.另外,图像其他区域的结构,尤其是一些强边也是非常重要的视觉特征,应尽可能地保持它们.为了达到这些目的,我们定义边界约束、显著度约束和结构约束,并将其融合到参数化过程中分别用以保持边界的规则性、重要物体的严格变换和强边结构不发生严重变形.

2.2.1 边界约束

源图像的边界经过适应后应当被映射到目标显示屏的边缘,相应的适应图像的边界应该贴附于屏幕边缘.为此,目标网格 M_t 上与源网格 M_S 四个角点对应的点其位置需要固定在屏幕的角点;对应于 M_S 上下边界的 M_t 上的点,其 v 坐标应赋为屏幕上、下边界的纵坐标,而通过参数化求取其 u 坐标;同样地,对应于 M_S 左右边界的 M_t 上的点,其 u 坐标应赋为屏幕左、右边界的横坐标,而其 v 坐标通过参数化求解(如图 3 所示).

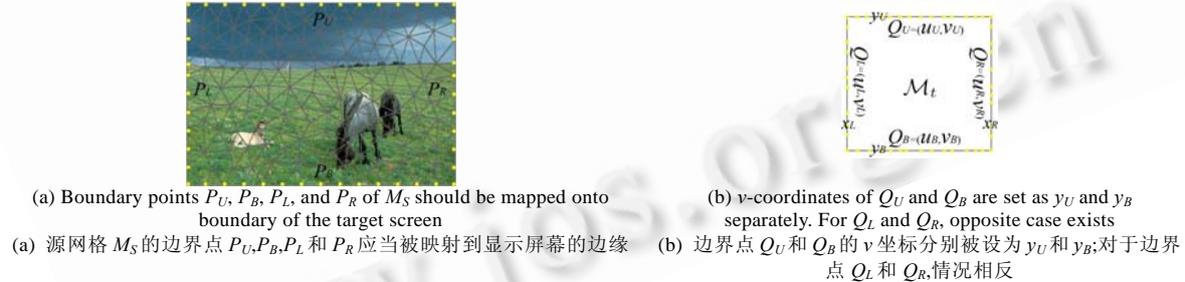


Fig.3 Boundary constraint
图 3 边界约束

假设目标网格 M_t 上对应于源网格 M_S 的上、下、左、右边界的点集分别为 $Q_{U,i|i=1..n_U}, Q_{B,i|i=1..n_B}, Q_{L,i|i=1..n_L}$ 和 $Q_{R,i|i=1..n_R}$, 目标屏幕的上、下边界的 y 坐标为 y_U 和 y_B , 其左、右边界的 x 坐标为 x_L 和 x_R , 则上述的边界约束可以表示为

$$F_B = \sum_{i=1}^{n_U} |v_{U,i} - y_U| + \sum_{i=1}^{n_B} |v_{B,i} - y_B| + \sum_{i=1}^{n_L} |u_{L,i} - x_L| + \sum_{i=1}^{n_R} |u_{R,i} - x_R| = 0 \quad (1)$$

2.2.2 显著度约束

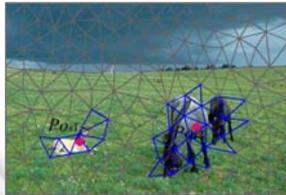
显著区域或重要物体,例如家庭照片中的人体、足球比赛中的球员和足球通常吸引较强的用户关注,在图像适应过程中应该被完全保持,避免其内部形变,同时对其进行适当增强.由于已经得到了图像的显著度图,在网格表示下可以利用聚类显著度值较高的邻接三角形近似逼近得到显著区域,采用如下形式表示:

$$\left\{ O_{sj} = \bigcup_{i=1}^{n_j} (S_{\Delta_i} > S_{\mu}) \mid j=1, \dots, J \right\} \quad (2)$$

O_{sj} 表示源图像中一个包含 n_j 个三角形的重要物体; J 表示重要物体的个数, $J=1$ 表示图像只有一个重要物体,而 $J>1$ 表示多个重要物体.其中 S_{μ} 为显著度阈值,显著度值大于此阈值的三角形视为显著区域;经过反复实验,在本文算法中 S_{μ} 取 0.6.另外,我们忽略网格中一些孤立琐碎的重要三角形区域.

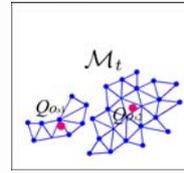
为了避免重要物体的形变扭曲,我们定义硬性约束:对每一个显著物体 O_{sj} ,在图像适应过程中只能进行按比例缩放和平移变换.为此,我们首先计算源网格 M_S 上 O_{sj} 区域内所有网格顶点的质心 P_{Os_j} ;然后将 P_{Os_j} 作为极点以求取 O_{sj} 区域内任意点 P_i 的极坐标 (r_i, θ_i) .最后在参数化的过程中 O_{sj} 区域以比例 z_{Os_j} 进行刚体变换,这样可以得到图像适应之后点 P_i 的极坐标 $(z_{Os_j} \cdot r_i, \theta_i)$;假设 P_{Os_j} 经过适应后的点为 Q_{Os_j} ,一旦给定其坐标,我们可以将上述极坐标转换为 P_i 在适应图像上的对应点 Q_i 的欧式坐标(如图 4 所示).假设将极坐标转换为欧式坐标的函数为 f ,则下式成立:

$$Q_i = f(Q_{Os_j}, z_{Os_j} \cdot r_i, \theta_i) \quad (3)$$



(a) Input image contains two salient objects

(a) 输入图像包含两个显著物体



(b) Retargeted object is determined by a mass point's position and a scale factor

(b) 适应的显著物体由其质心位置和物体的缩放比例确定

Fig.4 Saliency constraint

图 4 显著度约束

上述的显著度约束可以表示为

$$F_{SA} = \sum_{j=1}^J \left(\sum_{i=1}^{n_j} |Q_i - f(Q_{Os_j}, z_{Os_j} \cdot r_i, \theta_i)| \right) = 0 \quad (4)$$

这里 O_{sj} 的适应部分包含 n_j 个点 $Q_{i|i=1..n_j}$.

我们采用如下方式计算缩放比例参数 z_{Os_j} .如果图像中只包含一个显著物体, z_{Os_j} 可以通过严格限制适应后的显著物体尺寸不超过目标屏幕尺寸来决定.如果存在多个显著物体,则适应后所有显著物体的面积和应小于目标屏幕面积,此外,适应后显著物体的宽度总和以及高度总和也应小于目标屏幕的宽度和高度.

参数化过程中 Q_{Os_j} 作为未知参数通过参数化求得.一旦得到 Q_{Os_j} 的坐标,目标网格中显著物体上的所有网格点都可以通过公式(3)求得.

2.2.3 结构约束

图像中的强边,例如建筑物的侧墙、树木的树干是重要的视觉元素,对于图像的理解亦至关重要,在图像适应中应尽可能地保持这些强边.上一小节已经讨论了如何通过显著度约束完全保持重要物体,这里只讨论如何保持图像其余区域的强边结构.我们通过 Hough 变换检测图像背景中的强边,假设检测出的强边如下:

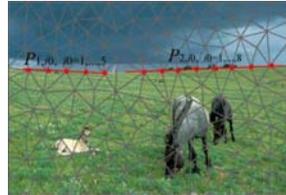
$$\left\{ L_k = \left(\bigcup_{i=1}^{m_k} P_{k,i} \right) \mid k=1, \dots, K \right\} \quad (5)$$

其中 L_k 为源图像网格 M_S 中通过点集 $P_{k,i|i=1, \dots, m_k}$ 的边, K 是强边的个数(如图 5(a)所示).

为了基本保持图像背景结构,应保持点集 $P_{k,i|i=1,\dots,m_k}$ 在适应后的对应点集 $Q_{k,i|i=1,\dots,m_k}$ 仍然为直线.假设适应后的直线集合表示为

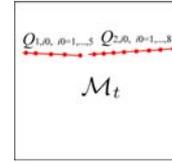
$$\left\{ RL_k = \left(\bigcup_{i=1}^{m_k} Q_{k,i} \right) \mid k=1,\dots,K \right\} \quad (6)$$

其中 RL_k 为 L_k 适应后得到的直线, L_k 的方程为 $y=a_k \cdot x+b_k$. $Q_{k,i}$ 满足 $v_{k,i}=a_k \cdot u_{k,i}+b_k$, 这里 a_k 和 b_k 可以通过 $Q_{k,i}$ 上任意两点表示.



(a) Two feature point sets $P_{1,i_0,i_0=1,\dots,5}$ and $P_{2,i_0,i_0=1,\dots,8}$ adhere to strong edges

(a) 输入图像中的特征点集 $P_{1,i_0,i_0=1,\dots,5}$ 和 $P_{2,i_0,i_0=1,\dots,8}$ 构成强边



(b) Retargeted sets $Q_{1,i_0,i_0=1,\dots,5}$ and $Q_{2,i_0,i_0=1,\dots,8}$ should pass through lines

(b) 适应后的对应点集 $Q_{1,i_0,i_0=1,\dots,5}$ 和 $Q_{2,i_0,i_0=1,\dots,8}$ 应共线

Fig.5 Structure constraint

图 5 结构约束

事实上,当图像背景中包含过多结构边的时候,结构约束不能被同时满足,由此我们定义了下面的能量约束并将其作为参数化过程中的软约束:

$$E_{ST} = \sum_{k=1}^K \left(\sum_{i=1}^{m_k} (v_{k,i} - a_k \cdot u_{k,i} + b_k)^2 \right) = 0 \quad (7)$$

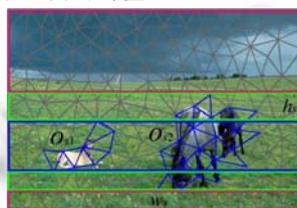
2.3 约束条件下的参数化

建立源图像的网格 M_S 后,图像适应就转化为求解与 M_S 同拓扑同时具有目标屏幕尺寸的网格 M_T 的网格参数化问题.网格参数化在计算机图形学领域已有广泛研究,其目的是用来建立无序 3D 网格和平面^[14,15]等参数域的对对应关系.大多数方法通过优化参数化过程中涉及的形变来求解,目前研究者已提出了许多不同的形变度量标准,其中基于边长伸展的方法在减小参数化过程中引起的全局网格形变方面非常有效^[15],其基本的思想是:平面网格完全由其所有边的边长决定.这里我们关注的是求解 2D 到 2D 的参数化.

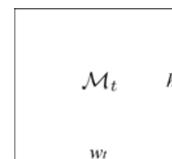
2.3.1 目标网格中边的理想长度计算

目标网格 M_T 的形状由其所有边的边长唯一确定.由于显著物体的边由其缩放比例决定,因此我们仅需计算 M_T 中的其余边.直观来看,图像适应过程中目标网格 M_T 由源网格 M_S 经过变形得到, M_T 中的边由 M_S 中的对应边经过缩放得到,缩放因子受适应过程中图像分辨率变化的影响,同时由于突出重要物体会相对挤压其邻接边,因此边的缩放因子还受到该边相对于重要物体位置的影响.

我们在 x 和 y 方向独立求解目标网格 M_T 中每条边相对于源网格 M_S 中边的缩放比例.为了求解 x 方向上的缩放比例,首先对源网格按照如图 6 所示划分,进而为每一个网格点计算一个虚拟缩放比例,而边的缩放比例为其两个顶点缩放比例的均值.



(a) Division of image according to object
(a) 图像划分



(b) Target screen
(b) 目标屏幕

Fig.6 Calculation of scale factor

图 6 缩放比例的计算

假设 $w_s \times h_s$ 和 $w_t \times h_t$ 分别表示源图像和目标屏幕的分辨率.对于 M_S 中的网格点 $P_i=(x_i, y_i)$,如果 P_i 位于图 6 中的最上方或下方矩形内,其理想的缩放比例为 w_t/w_s ;而如果 P_i 位于中间 3 个矩形内,其缩放比例还与显著物体的缩放比例相关.假设直线 $y=y_i$ 与显著物体相交形成的线段长度为 $\{l_{Os_j} | j=1, \dots, J\}$,则 $P_i=(x_i, y_i)$ 的理想缩放比例为

$$\left(w_t - \sum_{j=1}^J z_{Os_j} \times l_{Os_j} \right) / \left(w_s - \sum_{j=1}^J l_{Os_j} \right) \quad (8)$$

这里, z_{Os_j} 为 Os_j 的缩放因子.例如图 6(a)中的网格,如果 P_i 在最上方或下方矩形内部,则 $l_{Os_1} = l_{Os_2} = 0$;而如果 P_i 在中间 3 个矩形内部,则 $l_{Os_1} = 0$ 而 $l_{Os_2} > 0$;否则 l_{Os_1} 和 l_{Os_2} 均 < 0 .

我们可以类似地计算 P_i 在 y 方向上的虚拟缩放比例.每一条边 $e(P_i P_j)$ 的缩放比例通过求取 x 和 y 方向上两个端点 P_i 和 P_j 缩放比例因子的均值得到.

假设源网格 M_S 中的边 $e(P_i P_j)$ 的缩放比例为 s_{xij}, s_{yij} , $e(P_i P_j)$ 在 x 和 y 方向上的长度为 (l_{xij}, l_{yij}) ,则目标网格 M_t 中的边 $e(Q_i Q_j)$ 的合理长度为

$$l_{ij} = l_{e(Q_i, Q_j)} = \sqrt{(s_{xij} \times l_{xij})^2 + (s_{yij} \times l_{yij})^2} \quad (9)$$

2.3.2 网格参数化

由于目标网格 M_t 完全由其所有边的长度决定,故可以通过求解下列基于边长的能量方程的最小值确定 M_t :

$$E_t = \sum_{(i,j) \in \text{edges}} (\|Q_i - Q_j\|^2 - l_{ij}^2)^2 / l_{ij}^2 \quad (10)$$

对于源网格中的每一个显著物体 Os_j ,其所有内部边的边长伸展的能量形式都不在方程(10)中体现,只有连接重要物体边界点与外部点的边的边长伸展能量在求解过程中发挥作用,而 Os_j 的边界点通过第 2.2.2 节所述的方法由其适应后极点的待求坐标和 Os_j 的缩放比例因子通过公式(3)表示.

显见上述方程的最小化容易求解,但当 M_t 网格密度过大或规则性较差时,求得的目标网格 M_t 可能是拓扑无效的.这主要是由于 M_t 中的三角形相对于 M_S 中的对应三角形朝向发生了反转,为了解决这个问题,我们通过配置符号函数 $sign$ 惩罚可能引起反转的三角形的相关边来保证参数化的有效性.



Fig.7 Consistency of triangles' orientation

图 7 三角形朝向一致性

假设 $\Delta_{Q_i} = \Delta(Q_i Q_{k1} Q_j)$ 和 $\Delta_{Q_j} = \Delta(Q_j Q_{k2} Q_k)$ 是两个邻接三角形,公共边为 $e(Q_i Q_j)$. $\Delta_{P_i} = \Delta(P_i P_{k1} P_j)$ 和 $\Delta_{P_j} = \Delta(P_j P_{k2} P_k)$ 为上述两个三角形在源网格 M_S 中的对应三角形.对于每一对对应的三角形,顶点的走向应该一致(如图 7 所示).由此定义:

$$\omega_{ij} = sign \min(\det(\overline{Q_i Q_{k1}}, \overline{Q_j Q_{k1}}) \cdot \det(\overline{P_i P_{k1}}, \overline{P_j P_{k1}}), \det(\overline{Q_i Q_{k2}}, \overline{Q_j Q_{k2}}) \cdot \det(\overline{P_i P_{k2}}, \overline{P_j P_{k2}})) \quad (11)$$

故能量方程(10)可以转化为

$$E_t = \sum_{(i,j) \in \text{edges}} (\omega_{ij} \cdot \|Q_i - Q_j\|^2 - l_{ij}^2)^2 / l_{ij}^2 \quad (12)$$

其中系数 ω_{ij} 惩罚 M_t 中相对于 M_S 中对应三角形发生反转的三角形.如果出现反转, ω_{ij} 的值取 -1 ,否则取 $+1$.利用这个能量方程,可以在纹理平面上得到正确的拓扑有效的目标网格.

综上所述并结合第 2.2 节的约束,图像适应的约束参数化过程可以表示为

$$\arg \min_{Q_i, i=1, \dots, n} (E_t + \lambda \cdot E_{ST}), \text{ 满足 } F_B, F_{SA} = 0 \quad (13)$$

上式中边界约束和显著度约束是硬性约束,而结构约束视为软性约束. λ 是一个用来平衡参数化能量和结

构约束影响的权重因子。

我们通过多维牛顿方法求解等式(13)的最小值。在牛顿方法的每一次迭代过程中通过一个多重网格^[16]来求解稀疏线性方程组,可以非常快速地得到目标网格。

M_t 中显著物体之外的点和显著物体的质心通过最小化(13)求得,进而目标网格中显著物体上的点通过公式(3)计算得到。经参数化过程求得目标网格 M_t 后,我们借助对源图像的纹理映射^[17]生成适应图像。

3 实验结果与分析

3.1 实验结果

我们在一台配置 Intel P4 3.0GHz CPU,1G 内存的机器上对多种类型的图像上实验了本文算法。一些具有代表性的结果如图 8~图 12 所示。

从图 8 和图 9 中可以看出,与直接缩放方法所产生的较明显的物体扭曲和重要物体尺寸太小的效果相比较,本文算法可以有效地保持图像结构并突出重要物体。图 9~图 11 展示了一些包含多个重要区域的图例,本文方法通过非均匀地收缩图像区域有效地突出了所有的重要物体,而产生的扭曲比较细微。图 10 和图 11 显示了源图像适应不同分辨率和横宽比的例子,例如 1:1,3:4,4:3 以及 16:9。



Fig.8 Experimental results and the comparison with scaling

图 8 实验效果及与直接缩放的比较

图 12 比较了本文算法生成效果和其他图像适应方法产生的效果。源图像尺寸为 1024×768,结果图像除图 12(b)外尺寸均为 320×320。从这个例子可以看出,直接缩放方法的效果(图 12(b))中人脸部分太小难以辨认;而裁剪的效果(图 12(c))可以有效突出重要内容,但是很多信息被丢失;鱼眼效果^[6](图 12(d))可以突出重点并保持信息完整,但是产生了较明显的背景扭曲;图 12(e)为基于分割-重组图像的效果,对于前景和背景颜色分布相似的图片,准确的分割是一大挑战,而且对包含复杂结构的背景进行修复也是非常困难的;抽丝方法^[18]的效果(图 12(f))扭曲了图像中的人物;我们进一步在抽丝方法中加入人脸检测,得到的效果(图 12(g))完整地保持了图像中的人物,但是背景中的一些物体例如路标被破坏了;在本文算法的结果中,人物部分更突出并且没有明显扭曲,同时较好地保持了其他重要物体,如金门大桥的物体结构。

本文算法的效率主要取决于优化参数化目标函数(13)的效率,而优化该函数得到的方程组维数由源网格的节点数量减去重要物体上的点数决定。通常情况下,我们将源图像划分为较为稀疏的网格,这样算法的计算效率比较高。例如图 1(b)中的网格包含大约 150 个节点,去除重要物体上的网格点并添加相应的重要物体中心的坐

标参数,优化方程(13)得到的方程组关于网格横或纵坐标的维数约为 120,在我们的计算机上通过多维牛顿方法^[16]求解方程组的时间为 0.87s.在求得结果网格之后,借助图形卡通过标准的纹理映射函数可以实时生成目标图像.在手机或 PDA 上应用时可以通过 OpenGL ES 图形标准的硬件加速功能进一步优化方程,以期达到实时的图像适应.

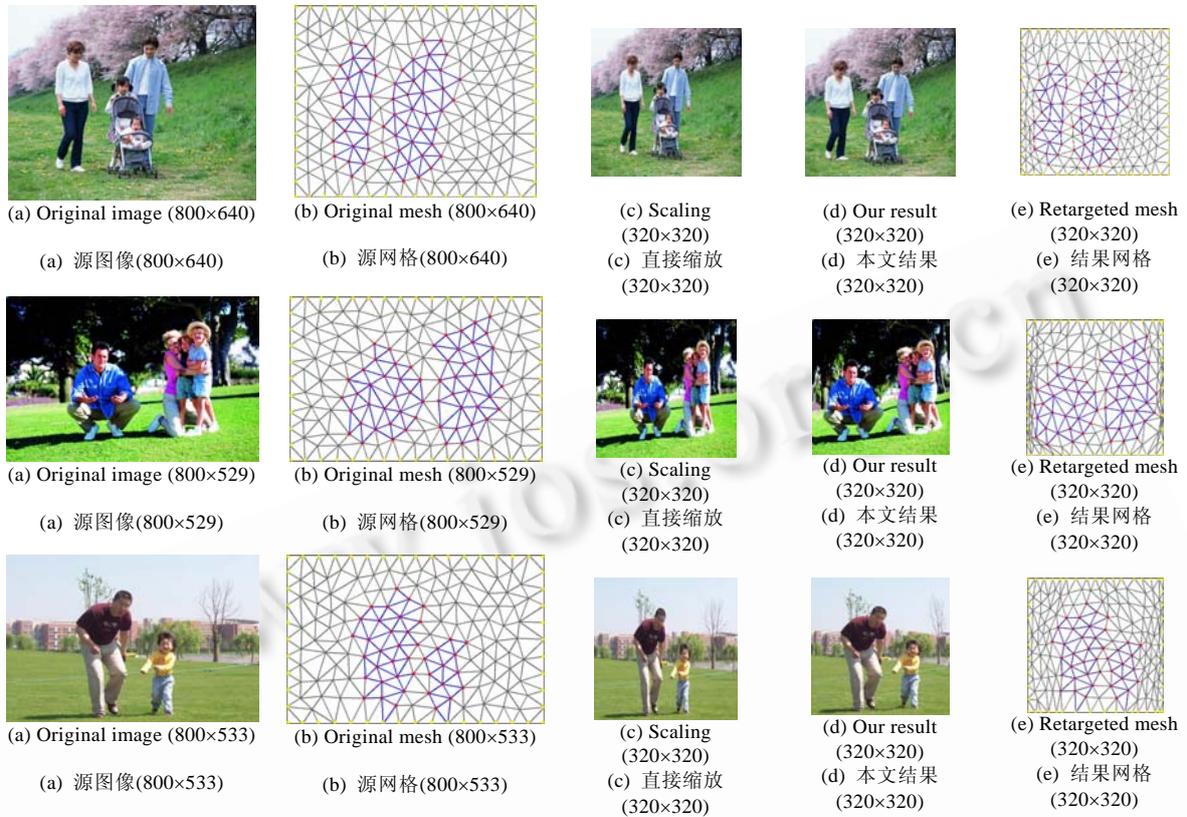


Fig.9 Experimental results
图 9 实验结果

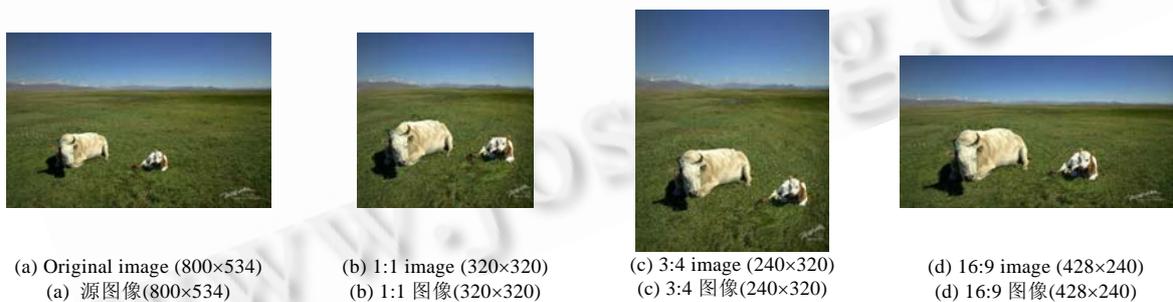


Fig.10 Retargeted images with different aspect ratios
图 10 不同横宽比的适应图像

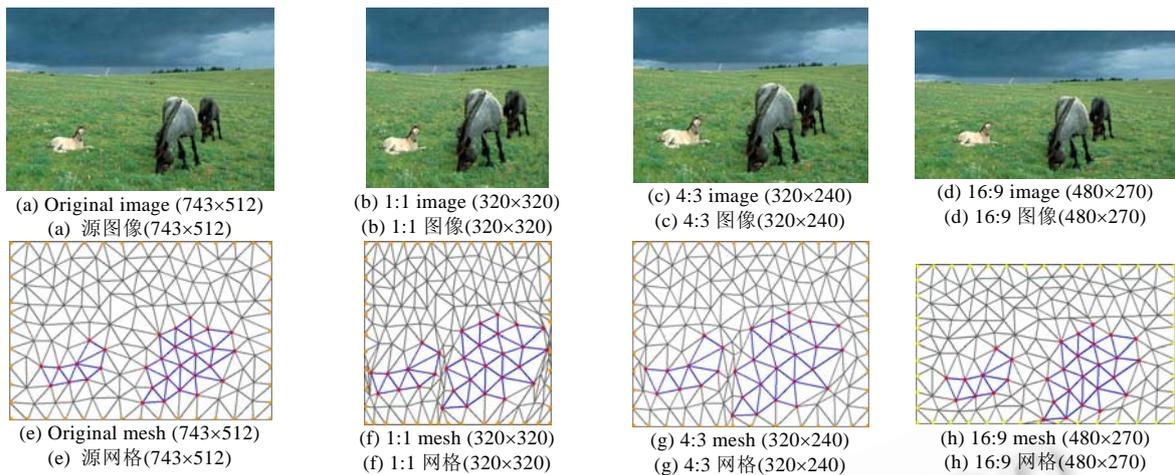


Fig.11 Retargeted images and meshes with different aspect ratios
图 11 不同横宽比的适应图像及其对应网格



Fig.12 Comparison between our approach and previous methods
图 12 本文算法和其他图像适应方法的比较

3.2 用户调查

为进一步评估我们算法的有效性,我们开展了两项用户调查(测试 1,测试 2)。测试 1 用来衡量本文算法效果相比其他目前手机、PDA 上广泛采用的方法,如直接缩放和裁剪是否更能为用户所接受,测试 2 的目的是评估我们算法结果的信息可靠性和用户满意度。

调查使用了总共 20 张具有代表性的图像及其适应效果,适应图像的尺寸均为适合 PDA 播放的 320×320 像素分辨率。测试 1 共有 17 人参与,首先比较本文算法效果和线性缩放,20 张图片的两种不同方法所产生的效果被排成 20 组,每组中本文效果和缩放的效果随机排放,参与者从每组图片中选出他们更满意的结果并写在答题板上;本文算法效果和裁剪效果的对比测试类似进行。统计数据显示:与直接缩放的结果相比,本文结果被选择的次数占 88.3%。与裁剪的结果相比为 72.5%,我们进一步检查了测试图像,发现对于简单背景的图像,用户更倾向于选择裁剪效果,但对于背景结构复杂的图像,由于本文结果保持了更多的图像信息并尽量地避免了扭曲,

测试者对本文效果更加满意.这一现象表明:当图像背景复杂时,前景物体之外的信息对于视觉效果的影响非常重要.

测试 2 中有 12 人参与了我们的问卷调查,参与者分别查看了 10 幅测试图像以及本文算法生成的适应图像并回答了两个问题:结果图像所包含的信息相对于源图像是否可靠?和如果将结果图像显示在您的手机上,是否对该效果满意?总体上,第 1 个问题的肯定回答占 96%,而第 2 个问题的肯定回答占 84%.

4 总结和展望

本文提出了一种基于网格参数化的图像适应方法,该方法首先将源图像划分为能够反映其特征结构的网格,进而通过带约束的参数化求解目标图像网格并借助理映射获取最终的适应图像.实验证明,应用该方法对图像进行适应,在有效突显图像中重要物体或显著区域的同时很好地保持了其背景内容;其中的网格化表示机制能够基本保持图像结构,适应后的图像仅有细微的视觉变形.

图像网格化使得图像中的重要物体可以通过三角形聚类结构进行表示,从而可以更方便地控制和突出显著区域.本文方法主要采用较为稀疏的网格对图像进行三角化剖分,这样计算效率比较高,但同时三角形聚类对重要物体的表示不够精细,这使得图像适应在突出重要物体的同时也略微放大了其邻近区域.为了解决这个问题,未来我们计划采用多分辨率的网格表示图像,对重要区域采用稠密网格,而非重要区域通过稀疏网格划分,借此,在不降低算法效率的同时对重要物体进行更精确地表示和控制.

此外,本文方法能够很好地保持图像结构,在适应图像上所产生的形变非常细微,但由于适应过程中没有明确考虑图像语义^[19],尤其是图像中重要物体的布局、空间关系和拓扑关系等,如不能保持这些信息可能会使用户对适应图像的理解产生偏颇.例如在足球比赛的图像中,球员的相对位置以及球员和足球之间的位置关系对用户主观视觉及其对图像的正确理解就非常关键,应该被保持.今后,我们将考虑这方面的研究并探索切实可行的解决方案.另外,将目前基于网格参数化的图像适应方法推广到视频领域也是我们未来的一个研究方向.

致谢 在此,我们向对本文工作给予宝贵建议的审稿专家,以及南京大学计算机科学技术系的宋传鸣博士、任桐伟博士表示感谢.

References:

- [1] Jiang S, Liu H, Zhao Z, Huang Q, Gao W. Generating video sequence from photo image for mobile screens by content analysis. In: IEEE Int'l Conf. on Multimedia and Expo. Beijing: IEEE Press, 2007. 1475–1478.
- [2] Liu H, Xie X, Ma WY, Zhang HJ. Automatic browsing of large pictures on mobile devices. In: ACM Multimedia. 2003. 148–155.
- [3] Chen LQ, Xie X, Fan X, Ma WY, Zhang HJ, Zhou HQ. A visual attention model for adapting images on small displays. *Multimedia Systems*, 2003,9(4):353–364.
- [4] Santella A, Agrawala M, DeCarlo D, Salesin D, Cohen M. Gaze-Based interaction for semi-automatic photo cropping. In: Proc. of the CHI. New York: ACM Press, 2006.
- [5] Suh B, Ling H, Bederton B, Jacobs D. Automatic thumbnail cropping and its effectiveness. In: ACM UIST. New York: ACM Press, 2003. 95–104.
- [6] Liu F, Gleicher M. Automatic image retargeting with fisheye-view warping. In: ACM UIST. New York: ACM Press, 2005. 153–162.
- [7] Setlur V, Lechner T, Nienhaus M, Gooch B. Retargeting images and video for preserving information saliency. *IEEE Computer Graphics and Applications*, 2007,27(5):80–88.
- [8] Itti L, Koch C. Computational modeling of visual attention. *Nature Reviews Neuroscience*, 2001,2(3):194–203.
- [9] Hu Y, Rajan D, Chia LT. Robust subspace analysis for detecting visual attention regions in images. In: ACM Multimedia. New York: ACM Press, 2005. 716–724.
- [10] Itti L, Koth C, Niebur E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1998,20(11):1254–1259.

- [11] Ma YF, Zhang HJ. Contrast-Based image attention analysis by using fuzzy growing. In: ACM Multimedia. New York: ACM Press, 2003. 374–381.
- [12] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: IEEE Conf. CVPR. IEEE Press, 2001. 511–518.
- [13] Wang J, Ying Y, Guo Y, Peng QS. Automatic foreground extraction of head shoulder images. In: Computer Graphics Int'l. LNCS 4035, Springer-Verlag, 2006. 385–396.
- [14] Guo Y, Wang J, Sun H, Peng QS. A novel constrained texture mapping method based on harmonic map. Computers & Graphics, 2005,29(6):972–979.
- [15] Zhou K, Snyder J, Guo B, Shum HY. Iso-Charts: Stretch-Driven mesh parameterization using spectral analysis. In: ACM Symp. on Geometry Processing. New York: ACM Press, 2004. 45–54.
- [16] Press W, Teukolsky S, Vetterling W, Flannery B. Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, 1992. 379–383.
- [17] Hearn D, Baker MP. Computer Graphics with OpenGL. Pearson Prentice Hall, 2004.
- [18] Avian S, Shamir A. Seam carving for content-aware image resizing. ACM Trans. on Graphics, 2007,26(3):267–276.
- [19] Ren TW, Guo YW, Wang GS, Zhang FY. Constrained sampling for image retargeting. In: ICME 2008. 1397–1400.



时健(1985—),男,江苏人,硕士生,主要研究领域为多媒体技术,图像视频处理.



郭延文(1980—),男,博士,讲师,主要研究领域为计算机图形学,图像视频处理,多媒体技术.



杜振龙(1971—),男,博士,副教授,主要研究领域为多媒体信息处理,计算机图形学.



张福炎(1939—),男,教授,博士生导师,主要研究领域为多媒体技术及其应用系统,计算机图形学,数字化图书馆.



彭群生(1947—),男,教授,博士生导师,主要研究领域为真实感图形学,虚拟现实,分子生物可视化技术.