

# 最优化网格形变的图像非均匀映射方法

郑庆阳<sup>1</sup> 涂玲<sup>2</sup> 毋立芳<sup>1</sup> 王丽娟<sup>1</sup> 肖鹏宇<sup>1</sup>

(1. 北京工业大学电子信息与控制学院, 北京 100124; 2. 中日友好医院, 北京 100029)

**摘要:** 本文针对内容感知的图像非均匀映射问题, 建立了一种有效的基于网格形变的全局最优化模型。引入了一系列保护重要网格形变的能量约束项, 使得当图像尺寸比例发生改变时, 视觉显著的重要区域不发生形变, 并且尽可能地保留图像中的全局信息, 给予用户最佳的图像浏览体验。此外, 也通过建立网格初始划分问题的最优化数学模型, 实现覆盖重要区域的网格密集分布, 减少关键对象边界信息的扭曲失真。通过与其他主流方法的结果比较与分析, 证明本文方法能够更好地保证图像关键区域不发生形变, 得到质量更优的结果图像。

**关键词:** 图像非均匀映射; 内容感知; 显著图; 网格变形; 全局最优化模型

**中图分类号:** TN911.73 **文献标识码:** A **文章编号:** 1003-0530(2014)09-1064-07

## Content-Aware Image Retargeting Using Optimization Grid Deformation

ZHENG Qing-yang<sup>1</sup> TU Ling<sup>2</sup> WU Li-fang<sup>1</sup> WANG Li-juan<sup>1</sup> XIAO Peng-yu<sup>1</sup>

(1. School of electronic information and control engineering, Beijing

University of Technology, Beijing 100124, China;

2. China-Japan Friendship Hospital, Beijing 100029, China)

**Abstract:** We built an effective global optimization model of image retargeting based on grid deformation. This model introduces a series of energy functions that prevent important grid deforming. The proposed approach allows resizing images into arbitrary aspect ratio while protecting visually prominent area and retaining the global information as much as possible, in order to give users the best experience of browsing image. In addition, we achieved that goal that important grid distribute densely by establishing the optimization model of partitioning the initial grid in the original image, reducing the artifacts on the boundary of the prominent object. A number of comparison experiments show that our method performs better than state-of-the-art approaches.

**Key words:** image retargeting; content aware; saliency; grid deformation; global optimization model

## 1 引言

随着科技的进步,适用于不同网络的终端设备不断更新换代。电视、笔记本电脑、PDA、手机等,它们通常具有不同的显示尺寸。通常,数字图像均以某一固定尺寸和宽高比例制作而成。然而,在不同类型的数字设备上显示它们时,往往需要改变它们的尺寸,来适应各种不同的显示设备。传统的均匀缩放以及简单的裁切方法都没有考虑图像本身的

内容,会导致拥有各自的缺陷。当图像的长宽比例发生改变时,均匀缩放方法导致结果图像中重要区域发生拉伸变形。而裁剪方法会带来图像大量的信息丢失,很有可能裁掉关键区域。那么,如何将图像内容更好地显示在多样化的数字设备上,基于内容感知的图像非均匀缩放技术成为最近几年一大热点研究话题。

基于内容感知的图像非均匀缩放技术的共同思想是尽可能保护人眼所关注的重要区域不发生

形变,尽量保留原始图像中的重要结构信息。此外,需要保证结果图像中没有引起明显的视觉上的人造伪影。

细缝裁剪方法<sup>[1]</sup>作为一种离散的图像非均匀缩放技术被 Avidan 等人在 2007 年首次提出,该方法通过不断地删除或复制不重要的细缝来改变图像尺寸。他们将一条细缝定义成只包含一个像素宽度、连续的、自顶向下或自左向右,并且拥有最小能量和的连通路。但是由于该类方法的离散性,它更容易破坏图像中的结构信息,并且引入不平滑的人造伪影,有时不能够很好地保护重要区域的结构信息,并且算法计算时间较长。之后,针对原始细缝裁剪方法所带来的速度和结果质量问题,一些改进算法也相继被提出<sup>[2-5]</sup>。另一类,连续的图像非均匀映射方法发展的更为迅速<sup>[6-9]</sup>,并且能够得到更为平滑的结果。这类方法的主要思想是建立源图像到目标图像的最优映射关系,映射过程中加入一些约束条件,来保护图像中的重要内容。具体做法是根据图像内容确定各像素点的重要性,然后根据重要性对图像进行非均匀缩放,使得图像重要区域形变较小,将形变扩散在非重要区域。wolf<sup>[6]</sup>通过聚合和扩张相邻低能量像素同时尽量完整的保留重要特征区域来改变图像尺寸,他们建立一个稀疏线性方程组来描述该最优化问题。应用基于网格形变非均匀映射方法<sup>[7-8]</sup>能够得到质量更好结果图像。Wang<sup>[7]</sup>的方法提出的最优化模型更侧重于达到局部最优化,它的结果能够较好的保留重要区域的比例形状,但有时不能够获得一个全局最优的结果图像,而且当缩放后的图像尺寸超过原始图像中重要物体的尺寸时,结果不是很理想。Cheng<sup>[8]</sup>的方法将全局最优化模型归纳成二次规划的形式,并且鼓励放大重要区域以及将直线保留的功能融入到最优化模型中,取得了不错的效果。但是,算法的复杂度较高,而且该最优化模型中保护图像中重要区域比例的约束条件需要给定经验值作为调整参数,一定程度上限制了算法的通用性。文献<sup>[9]</sup>的方法向最优化模型中加入拉普拉斯平滑能量约束项,用来平滑网格的变化程度,避免相邻网格缩放程度差别较大,改善了结果图像质量。

所有图像的非均匀缩放方法都必须经过两个

主要步骤:第一、获得原始图像对应的重要度图;第二、惩罚不重要的信息来保护图像中的重要区域。那么,针对第一个步骤,如何获得图像的重要度图来表征图像每个像素的重要度值呢?一些研究者使用简单、快速、低级别的梯度信息来获取图像的重要度图。还有一些人利用基于视觉注意力模型的分析方法来获得。这种基于视觉显著度的衡量标准使用定量的方法去分析人眼在观看一幅图像时注意力的集中位置。此类方法基于图像的颜色、亮度、纹理、边缘方向等多方面图像低级别特征与背景的对比度计算得来。图像中该点像素对比度值越大,说明显著度值就越高,越重要。文献<sup>[10]</sup>提出了基于内容上下文的显著度模型,该模型不仅适合处理仅含有单一重要物体的图像,并且对于复杂图像也能够获得不错的结果。

这篇文章提出了一种新的基于网格的图像非均匀映射的全局最优化模型。首先,根据文献<sup>[10]</sup>中提出的图像显著模型计算图像中各个像素的显著度值。然后,在全局最优化模型中,通过定义三种保护重要网格形变的有效能量约束项,尤其在于将保护重要区域比例的能量项融入该最优化模型的目标函数中。有效地保证当图像的尺寸比例改变时,在没有引起图像明显视觉失真的前提下,关键对象没有发生扭曲形变。同时,我们也尝试了一种有效的自适应初始网格划分方法,有效地保护了重要区域的边缘结构信息,进一步减少了结果图像中引入的人造伪影,提高了结果图像视觉质量。

## 2 基于网格形变的最优化模型

本文方法的大体流程,如图 1 所示,首先对原始图像进行初始网格划分,然后在图像尺寸改变的过程中,根据图像内容的重要度对网格进行不规则的变形,最后向变形后的网格内插入对应的原始图像内容,得到目标尺寸的结果图像。在这一部分,将主要阐述本篇论文的核心部分,针对网格形变问题,建立对应的全局最优化数学模型。在该最优化模型当中,要达到保护重要区域不发生拉伸变形的目的,需注意以下三点:第一、保护覆盖重要区域网格的比例不发生过多改变。第二、保护覆盖重要区域网格的全局方向,避免网格线过度弯曲;第三、希望覆盖重要区域网格经历相似变换。

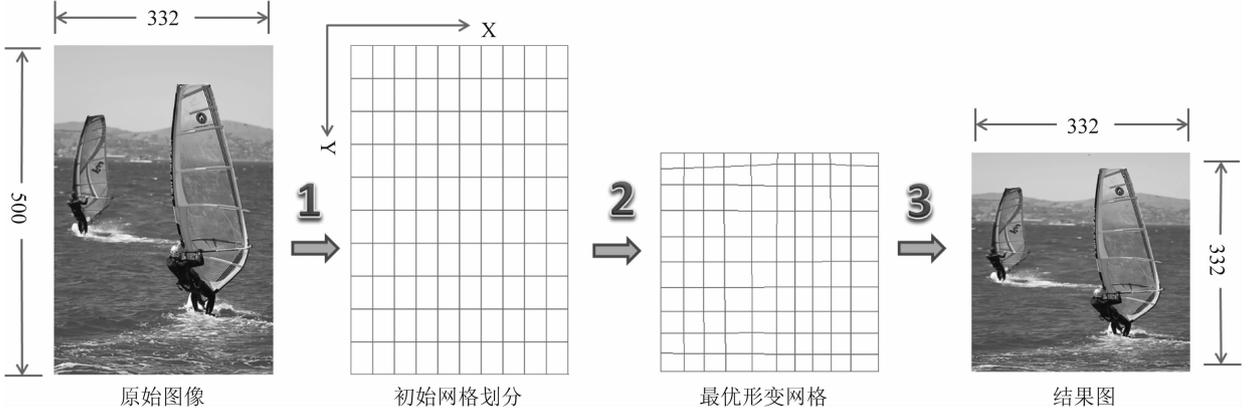


图1 算法整体流程图

Fig.1 Algorithm Overview

## 2.1 目标函数

假设,定义一幅大小为  $W \times H$  的原始输入图像,设定其目标尺寸为  $W_T \times H_T$ ,将原始图像划分为  $M \times N$  个网格,  $DG_{ij}$  代表第  $(i, j)$  个网格  $g_{ij}$  的形变量,  $IMP_{ij}$  代表网格  $g_{ij}$  的重要度值。图像在缩放的过程中,某一网格形变示意图如图2所示,  $g_{ij}$  表示原始图像中某一个初始划分的网格,  $g'_{ij}$  表示对应变形后的网格,  $(v_{i,j,x}, v_{i,j,y}), (v_{i+1,j,x}, v_{i+1,j,y}), (v_{i+1,j+1,x}, v_{i+1,j+1,y}), (v_{i,j+1,x}, v_{i,j+1,y})$  分别为网格  $g_{ij}$  的四个顶点坐标,为已知量。  $(v'_{i,j,x}, v'_{i,j,y}), (v'_{i+1,j,x}, v'_{i+1,j,y}), (v'_{i+1,j+1,x}, v'_{i+1,j+1,y}), (v'_{i,j+1,x}, v'_{i,j+1,y})$  分别为变形后网格  $g'_{ij}$  的四个顶点坐标位置,为所求量。那么,设定该问题中,最优化模型的目标函数为:

$$\min \sum_i^M \sum_j^N DG_{ij} \cdot IMP_{ij} \quad (1)$$

即通过网格重要度约束最小化网格总体形变量,尽量避免覆盖重要区域的网格发生变形。其中,  $IMP_{ij}$  根据文献<sup>[10]</sup>中得到的显著图计算得到,并进行归一化处理,使其值在  $[0, 1]$  范围之内。网格  $g_{ij}$  的形变量  $DG_{ij}$  定义如下:

$$DG_{i,j} = \lambda_1 \cdot DR_{ij} + \lambda_2 \cdot DD_{ij} + \lambda_3 \cdot DS_{ij} \quad (2)$$

式(2)中,  $\lambda_1, \lambda_2$  以及  $\lambda_3$  分别为各能量项的比重参数,本方法实验中取  $\lambda_1 = 3, \lambda_2 = 2, \lambda_3 = 1$ 。  $DR_{ij}$  为保护网格比例的能量项,  $DD_{ij}$  为保护网格线方向的能量约束项,  $DS_{ij}$  保护网格形状的能量约束项。

### 2.1.1 保护网格比例的能量约束项 $DR_{ij}$

当改变图像的尺寸或者宽高比时,为了保护用

户所关注的重要对象不发生拉伸变形,尽量保持它的原始结构比例,需要在网格形变的过程中,保证覆盖重要区域的网尽量不改变长宽比例。因此,定义用来保护网格比例的能量项约束  $DR_{ij}$  为:

$$DR_{ij} = [ (v'_{i+1,j,x} - v'_{i,j,x}) - (v'_{i+1,j+1,y} - v'_{i+1,j,y}) \cdot \text{ratio}_{ij} ]^2 + [ (v'_{i+1,j+1,x} - v'_{i,j+1,x}) - (v'_{i+1,j+1,y} - v'_{i+1,j,y}) \cdot \text{ratio}_{ij} ]^2 + [ (v'_{i+1,j,x} - v'_{i,j,x}) - (v'_{i,j+1,y} - v'_{i,j,y}) \cdot \text{ratio}_{ij} ]^2 + [ (v'_{i+1,j+1,x} - v'_{i,j+1,x}) - (v'_{i,j+1,y} - v'_{i,j,y}) \cdot \text{ratio}_{ij} ]^2 \quad (3)$$

式(3)中,  $\text{ratio}_{ij}$  为对应原始网格  $g_{ij}$  的宽高比例,计算如下:

$$\text{ratio}_{ij} = (v_{i+1,j,x} - v_{i,j,x}) / (v_{i+1,j+1,y} - v_{i+1,j,y}) \quad (4)$$

相对于文献<sup>[8]</sup>的方法,本方法更简单、有效地保护重要区域不发生形变。而且,不必花更多的精力去设定和调整各种阈值参数,具有更好的通用性。

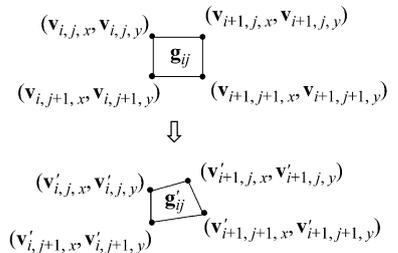


图2 单一网格形变说明示意图

Fig.2 The deformation diagram of a mesh

### 2.1.2 保护网格线方向的能量约束项 $DD_{ij}$

浏览图像的用户通常对于图像中重要物体方

向的改变也非常敏感,重要物体方向的变化也会造成图像结构信息的改变,且在该最优化模型中弯曲的网格线有时会给结果图像带来较多的人造伪影。为避免网格线过度扭曲,同时保护结果图像中重要物体的全局方向,定义保护重要网格线方向的能量约束项  $DD_{ij}$  为:

$$DD_{ij} = (\nu'_{i+1,j,y} - \nu'_{i,j,y})^2 + (\nu'_{i+1,j+1,y} - \nu'_{i,j+1,y})^2 + (\nu'_{i+1,j+1,x} - \nu'_{i+1,j,x})^2 + (\nu'_{i,j+1,x} - \nu'_{i,j,x})^2 \quad (5)$$

式(5)中,前两项保护了网格水平两条边的方向,后两项保护了网格竖直两条边的方向。

### 2.1.3 保护网格形状的能量约束项 $DS_{ij}$

为保证重要区域不发生失真变形,定义保护网格形状的能量约束项  $DS_{ij}$ :

$$DS_{ij} = [(\nu'_{i+1,j,x} - \nu'_{i,j,x}) - (\nu_{i+1,j,x} - \nu_{i,j,x})]^2 + [(\nu'_{i+1,j+1,y} - \nu'_{i+1,j,y}) - (\nu_{i+1,j+1,y} - \nu_{i+1,j,y})]^2 + [(\nu'_{i+1,j+1,x} - \nu'_{i,j+1,x}) - (\nu_{i+1,j+1,x} - \nu_{i,j+1,x})]^2 + [(\nu'_{i,j+1,y} - \nu'_{i,j,y}) - (\nu_{i,j+1,y} - \nu_{i,j,y})]^2 \quad (6)$$

## 2.2 约束条件

定义边界约束条件来保证得到的结果图像是矩形:

$$\begin{cases} \nu'_{i,j,x} = 0 & \nu_{i,j} \text{ is on the left boundary} \\ \nu'_{i,j,x} = TW & \nu_{i,j} \text{ is on the right boundary} \\ \nu'_{i,j,y} = 0 & \nu_{i,j} \text{ is on the top boundary} \\ \nu'_{i,j,y} = TH & \nu_{i,j} \text{ is on the bottom boundary} \end{cases} \quad (7)$$

其中, TW、TH 分别为图像的目标宽度和目标高度。

为了避免由于过度缩放而造成变形后的网格翻转问题,本方法引用了文献<sup>[8]</sup>的约束条件,有效解决了由于过度变形导致网格翻转造成结果严重失真的问题:

$$\begin{cases} \nu'_{i+1,x} \geq \nu'_{i,x} \\ \nu'_{i+1,y} \geq \nu'_{i,y} \end{cases} \quad (8)$$

### 2.3 网格初始划分

最初选择的初始网格划分是均匀一致的,然而当划分的网格数增多时,得到的结果图像能够更好的保护重要物体的局部边缘结构信息。因为当重要物体被密集的网格覆盖时,重要物体的边缘特征能够更好的被表征出来,网格内的像素重

要度值越一致。这样,在一定程度上能够减轻重要对象边缘扭曲失真。反之,当网格数量很少时,如图3所示,覆盖重要物体边缘的大网格(红色矩形框标注)内可能由于含有重要的像素不够多,而被当作不重要的网格而进行拉伸变形,进而导致重要区域(轮子部分)边缘信息变形失真。另一方面,每增加一个网格,最优化系统需要添加8个网格顶点变量,这将给算法带来很大的计算量,导致运行速度大大降低。基于以上2点,提出这样一个划分初始网格的策略:重要区域网格划分的精细,不重要区域的网格划分的粗犷。使得数量较少的自适应初始网格划分同样可以达到数量较多的网格划分效果。这样,无论在质量和计算速度上都得到了很好地提高。

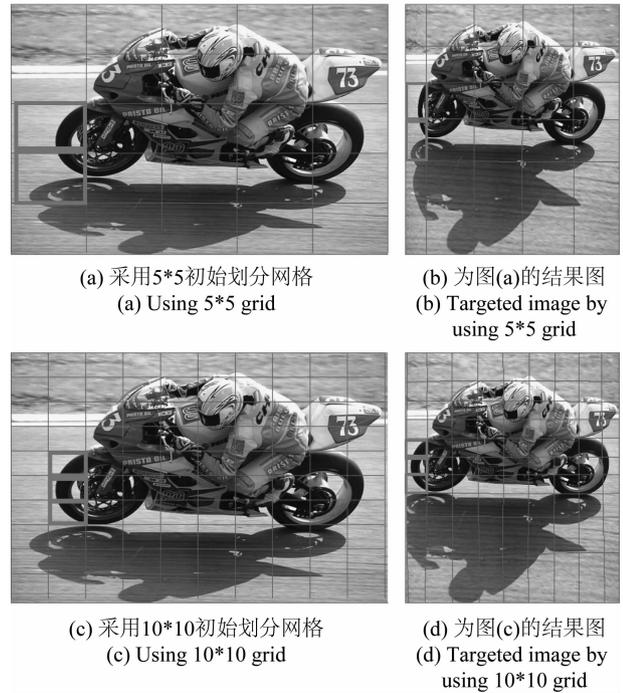


图3 采用不同网格数量的结果对比图

Fig. 3 Comparison of results by using different grid

区别于文献<sup>[7]</sup>的初始网格划分策略,我们定义初始划分网格为矩形网格,这样能够有效的将最优化系统中的变量个数由  $2(M+1)(N+1)$  减小到  $M+N$ ,大大提高了初始划分网格的运行时间。仍然需要建立初始网格自适应划分的最优化模型,我们希望越重要的网格覆盖的面积越小,定义该问题的目标函数为:

$$\min \sum_i^M \sum_j^N (\alpha + \text{IMP}_{ij}^\beta) \cdot \text{Area}(g_{ij}) \quad (9)$$

式(9)中,  $\text{Area}(g_{ij})$  为网格  $g_{ij}$  的近似面积;  $\alpha, \beta$  分别权重比例调整参数, 本实验中取  $\alpha=1, \beta=2$ 。同样采用 2.2 节中的边界条件(7)、(8), 不断的迭代求出新的网格顶点, 并且每次迭代后, 重新计算网格的重要度值。如图 4 所示, 网格在重要区域分布密集, 不重要区域分布稀疏。

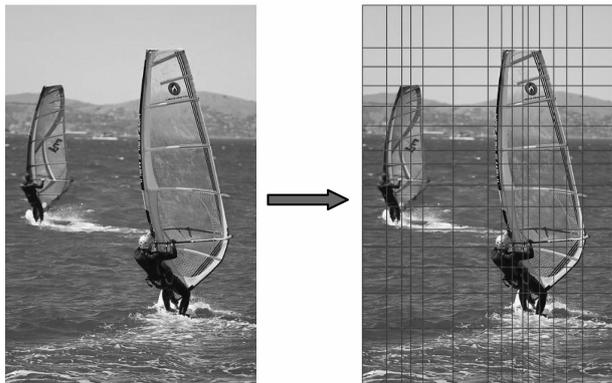


图 4 自适应网格初始划分结果图

Fig. 4 The result of adaptive initial partition

### 3 实验结果及分析

本方法采用 C++ 完成算法实现, 由于该算法所建立的最优化模型属于带有不等式约束条件的非线性最优化问题, 所以采用 ALGLIB 库中的 BLEIC 算法<sup>[11]</sup>来求解该最优化模型。运行的电脑配置为: CPU: 酷睿 i3, 主频: 3.3GHz, 内存: 2G。本文采用  $15 \times 15$  的初始网格划分策略, 将图 6 中的 7 组原始图像缩放到目标尺寸的运行时间如表 1 所示:

采用本文算法将一幅  $600 \times 450$  大小的原始图像缩放到不同尺寸比例(图 5), 得到的结果图像中, 重要对象(鸟)没有因为图像尺寸比例的改变而发生拉伸变形。从 Retargeting Me<sup>[12]</sup> 库中选取 7 组不同类型的结果图, 将我们的方法与其中几种最具代表性的连续方法进行比较分析。如图 6 所示, 发现传统的均匀缩放方法由于没有考虑图像本身内容各个区域的重要程度, 对整幅图像进行均匀缩放, 当图像比例改变时, 图像中重要对象发生拉伸变形。Wolf<sup>[6]</sup> 的方法仅仅在一定程度上

减轻重要区域发生形变。而且, 在某些情况下, 该方法可能会出现丢失图像中重要内容的情况, 如第三行的蝴蝶部分以及第四行的荷花部分。Wang<sup>[7]</sup> 的方法更注重得到局部最优的结果。Cheng<sup>[8]</sup> 的方法鼓励放大重要区域, 当图像中没有足够的背景区域来吸收形变时, 导致背景区域发生严重形变, 例如最后一行 Cheng 的方法导致荷花周围引入视觉失真。而且有时不能够很好的保护关键对象不发生形变, 如图第二行船的部分发生明显的拉伸变形。本方法的结果无论在关键对象的比例的保持以及图像全局信息的保留方面都取得不错的效果。

表 1 本文算法的运行时间

Tab. 1 The runtime of our method

	原始尺寸	目标尺寸	运行时间/s
1	1024 * 646	512 * 646	0.028
2	819 * 541	410 * 541	0.106
3	1024 * 700	512 * 700	0.184
4	1024 * 768	512 * 768	0.059
5	460 * 300	230 * 300	0.168
6	1024 * 564	512 * 564	0.152
7	500 * 313	250 * 313	0.060

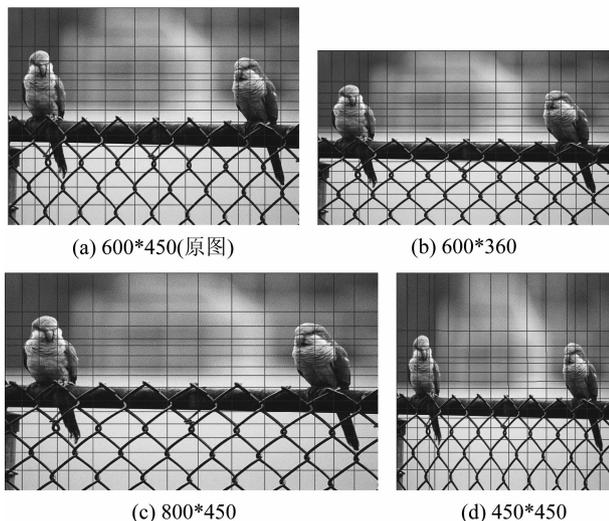
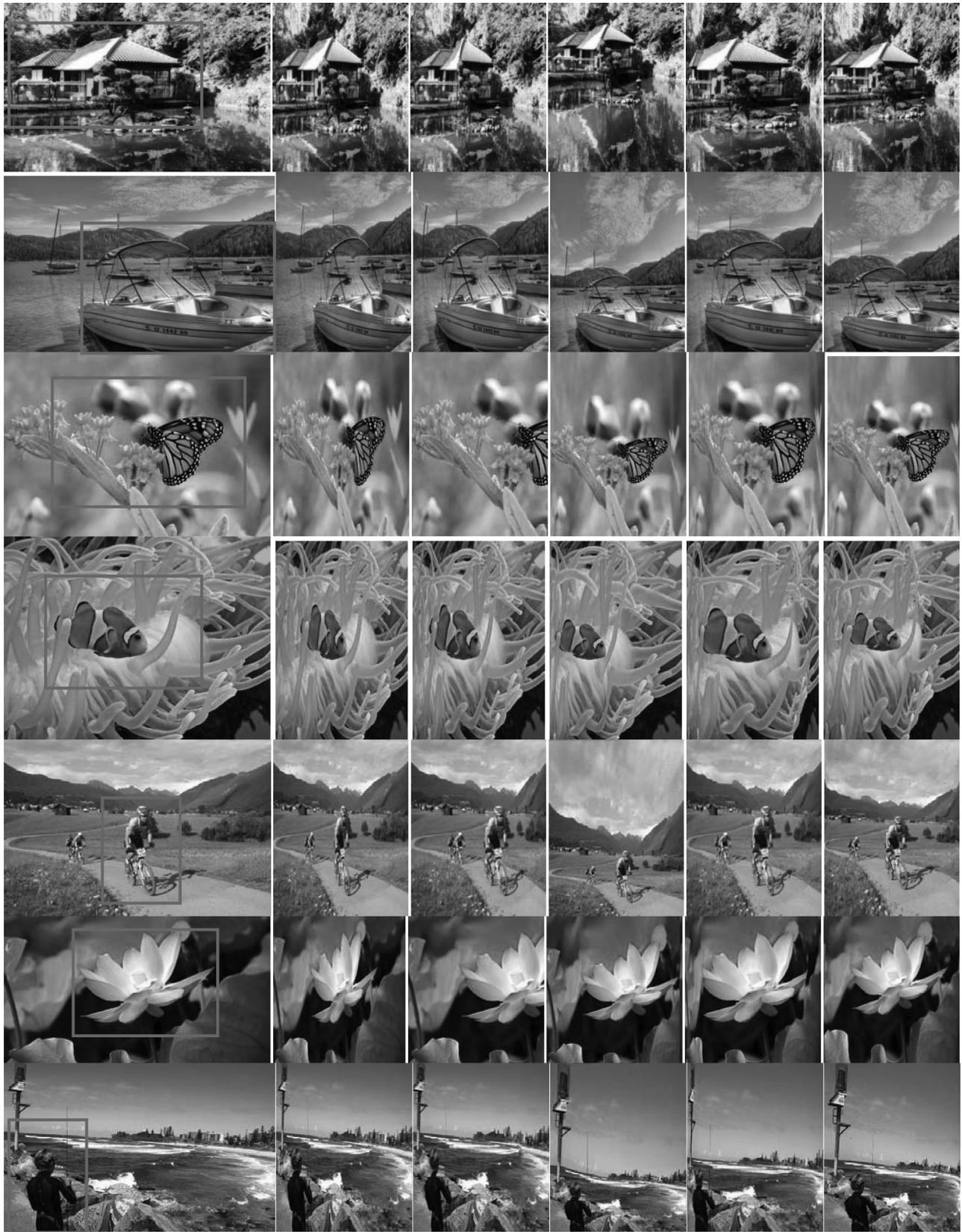


图 5 采用本文算法得到不同尺寸的变形网格结果图

Fig. 5 Our retargeting results with different sizes



(a) 原图

(b) 均匀缩放

(c) Wolf[6]

(d) Wang[7]

(e) Cheng[8]

(f) 我们的方法

图 6 我们的方法与其他方法的实验结果对比图

Fig. 6 Comparison of our results with other methods

## 4 结论

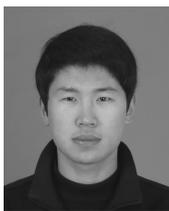
本文针对图像非均匀缩放问题建立了一种有效的全局最优化数学模型,保证图像在尺寸改变的过程中,覆盖重要区域的网格发生的形变小,不重要的网格形变大。最后向变形后的网格内插入原始图像内容,得到关键对象不发生形变的结果图像。同时,也建立了一种针对初始网格划分的最优化数学模型,以此达到覆盖重要区域的网格密集分布的目的,减少关键对象边缘引入的人工伪影以及算法的运行时间。该方法相于其他主流连续方法,能够更好的保证图像中的关键对象不发生形变得到视觉质量更优的结果图像,并且具备较快的运行速度。

### 参考文献

- [1] Avidan S, Shamir A. Seam carving for content-aware image resizing[J]. ACM Transactions on Graphics, 2007, 6(3): 10-18.
- [2] Rubinstein M, Shamir A, Avidan S. Improved seam carving for video retargeting[J]. ACM Transactions on Graphics, 2008, 27(3): 1-9.
- [3] Grundmann M, Kwatra V, Han M, et al. Discontinuous seam-carving for video retargeting[C]//2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, 2010: 569-576.
- [4] Wu L, Cao L, Chen C W. Fast and improved seam carving with strip partition and neighboring probability constraints[C]//2013 IEEE International Symposium on Circuits and Systems (ISCAS). 2013: 2812-2815.
- [5] Botao Wang, Hongkai Xiong, Zhiqian Ren, Chang Wen Chen. Deformable Shape Preserving Video Retargeting With Salient Curve Matching[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2014, 4(1): 82-94.
- [6] Wolf L, Guttman M, Cohen-Or D. Non-homogeneous content-driven video-retargeting[C]//In Proceedings of the Eleventh IEEE International Conference on Computer Vision, Rio de Janeiro, 2007: 1-6.
- [7] Wang Y S, Tai C L, Sorkine O, et al. Optimized scale-and-stretch for image resizing[C]//ACM Transactions on Graphics (TOG), 2008, 27(5): 118.

- [8] Chen R, Freedman D, Karni Z, et al. Content-aware image resizing by quadratic programming[C]//IEEE Computer Vision and Pattern Recognition Workshops (CVPRW), 2010: 1-8.
- [9] Hu W, Luo Z, Fan X. Image Retargeting Via Adaptive Scaling With Geometry Preservation[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2014, 4(1): 70-81.
- [10] Goferman S, Zelnik-Manor L, Tal A. Context-aware saliency detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(10): 1915-1926.
- [11] ALGLIB Project. Bound and linear equality/inequality constrained optimization [EB/OL]. <http://www.alglib.net/optimization/boundandlinearlyconstrained.php>, 2014.
- [12] Vaquero D, Turk M, Pulli K, et al. A survey of image retargeting techniques [C]//In Proceedings of SPIE, 2010, 7798: 779814.

### 作者简介



**郑庆阳** 男, 1989年生于辽宁, 北京工业大学硕士研究生。主要研究方向为基于内容感知的图像视频缩放技术研究。  
E-mail: 817huang@163.com

**涂玲** 女, 1985年生于江西, 中日友好医院信息技术和管理部助理工程师。主要研究方向为医疗数据分析、图像处理。  
E-mail: too0@sohu.com



**毋立芳** 女, 1970年生于山西, 教授, 现就职于北京工业大学电子信息与控制工程学院, 担任科研副院长, 主要研究方向为视频分析及应用、人脸检测与识别、生物特征加密及应用、图像处理。  
E-mail: lfwu@bjut.edu.cn

**王丽娟** 女, 1989年生于河北, 北京工业大学硕士研究生。主要研究方向为基于内容感知的图像视频缩放技术研究、图像的质量评价。  
E-mail: wanglijuan\_891014@163.com

**肖鹏宇** 男, 1991年生于北京, 北京工业大学本科毕业生, 密歇根大学硕士研究生。主要研究方向为信号处理。