

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **2**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年2月 第17卷 第2期(总第190期)

目次

综述

内容感知图像缩放技术综述 施美玲,徐丹(157)

图像处理和编码

引入纹理相似性的纺织品图像增强 杨学志,田晓梅,方静,卢洁(169)

保持边缘特征和增强对比度的图像缩放算法 计忠平,方美娥,王毅刚,吴向阳(178)

以图像认证为目的的分形编码数字水印算法 黄晓晴,于盛林(183)

结合运动矢量和像素递归的全局运动估计方法 赵亚湘,刘少强,樊晓平(191)

无缝栅格数据小波金字塔构建 马伯宁,冷志光,汤晓安,匡纲要(197)

图像分析和识别

伪彩色空间完全非结构化道路检测方法 胡晓辉,孙苗强,苏晓许(203)

未知环境下机器人障碍物检测技术 王文格,武凯宾,朱江,段文彬,许重阳(209)

局部 GAC 模型在医学图像分割中的应用 张建伟,方林,陈允杰,詹天明,罗春燕(215)

判别割(Dcut)的图像分割及其快速分割算法 邹小林,陈伟福,冯国灿(222)

改进的遮挡条件下瞳孔检测方法 潘林,魏丽芳,郑炳锟,余轮(229)

木材显微图像的局部水平集分割方法 汪杭军,祁亨年(236)

关键点检测的线要素综合算法 黄志坚,张金芳,徐帆江(241)

GPU 加速的高分辨率 DEM 图像地形特征线提取算法 刘洲俊,胡包钢(249)

改进 GVF 的自动 Snakes 模型 周亚男,程熙,骆剑承,沈占锋,胡晓东(256)

曲波变换用于磨粒图像不变矩的提取 张云强,张培林,任国全,王国德,徐超,李兵(263)

图像理解和计算机视觉

深度图像中基于轮廓曲线和局部区域特征的3维物体识别 吴莉婷,张宇,杨一平,田原(269)

机载 LiDAR 数据的 LZD 航带平差 王丽英,宋伟东(279)

计算机图形学

基于 Z 曲线的瓦片地图服务空间索引 聂云峰,周文生,舒坚,许虎(286)

虚拟现实和增强现实

GPU 加速的台风可视化方法 秦绪佳,张勤锋,陈坚,郑红波,徐晓刚(293)

中国图象图形学报

刊名题字: 宋 健 月刊(1996 年创刊)

第 17 卷 第 2 期

2012 年 2 月 16 日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文

编辑出版 《中国图象图形学报》编辑出版委员会

北京 9718 信箱 邮编 100101

电子信箱: jig@irsa.ac.cn

电话: 010-68407995 010-82614429

网 址: www.cjig.cn

印刷装订 北京北林印刷厂

广告经营许可证 京朝工商广字第 0346 号

总 发 行 北京报刊发行局

订 购 全国各地邮局

国外发行 中国国际图书贸易总公司

(中国国际书店)

(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
CAS China Society of Image and Graphics
Institute of Applied Physics and Computational
Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
of Journal of Image and Graphics
(P. O. Box 9718, Beijing 100101, China)
E-mail: jig@irsa.ac.cn
Distributed by Beijing Bureau for Distribution of Newspapers
and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
(P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 2 February 2012

Contents

Review

Survey on content-aware image resizing techniques Shi Meiling, Xu Dan (157)

Image Processing and Coding

Enhancement of textile image with texture similarity Yang Xuezhi, Tian Xiaomei, Fang Jing, Lu Jie (169)

Edge-preserving and contrast-enhancement image scaling Ji Zhongping, Fang Mei'e, Wang Yigang, Wu Xiangyang (178)

Watermarking method on fractal coding for image authentication Huang Xiaoqing, Yu Shenglin (183)

Global motion estimation method with motion vectors and pixel recursion Zhao Yaxiang, Liu Shaoqiang, Fan Xiaoping (191)

Seamless wavelet pyramid construction for raster data Ma Boning, Leng Zhiguang, Tang Xiao'an, Kuang Gangyao (197)

Image Analysis and Recognition

Completely unstructured road detection for pseudo-color space Hu Xiaohui, Sun Miaoqiang, Su Xiaoxu (203)

Obstacle detection for robot in unknown environment
..... Wang Wenge, Wu Kaibin, Zhu Jiang, Duan Wenbin, Xu Chongyang (209)

Application of local GAC model for medical image segmentation
..... Zhang Jianwei, Fang Lin, Chen Yunjie, Zhan Tianming, Luo Chunyan (215)

Fast image segmentations of Deut Zou Xiaolin, Chen Weifu, Feng Guocan (222)

Improved method for the pupil measurement under occlusion Pan Lin, Wei Lifang, Zheng Bingkun, Yu Lun (229)

Segmentation method of wood microscopic image based on local level set Wang Hangjun, Qi Hengnian (236)

Algorithm of line generalization with key point detection Huang Zhijian, Zhang Jinfang, Xu Fanjiang (241)

High resolution DEM topographic feature line extraction algorithm using GPU Liu Zhoujun, Hu Baogang (249)

Automatic snakes model based on modified GVF Zhou Yanan, Cheng Xi, Luo Jiancheng, Shen Zhanfeng, Hu Xiaodong (256)

Invariant moment extraction by curvelet transform for wear particle images
..... Zhang Yunqiang, Zhang Peilin, Ren Guoquan, Wang Guode, Xu Chao, Li Bing (263)

Image Understanding and Computer Vision

Three-dimensional free-form object recognition based on contour curve and local surface patches in range images
..... Wu Liting, Zhang Yu, Yang Yiping, Tian Yuan (269)

Least Z-difference strip adjustment of airborne LiDAR data Wang Liying, Song Weidong (279)

Computer Graphics

Spatial index for tile map service based on Z curve Nie Yunfeng, Zhou Wensheng, Shu Jian, Xu Hu (286)

Virtual Reality and Augmented Reality

GPU accelerated typhoon visualization method Qin Xujia, Zhang Qinfeng, Chen Jian, Zheng Hongbo, Xu Xiaogang (293)

中图分类号: TN911.73 文献标志码: A 文章编号: 1006-8961(2012)02-0157-12
论文引用格式: 施美玲, 徐丹. 内容感知图像缩放技术综述 [J]. 中国图象图形学报, 2012, 17(2): 157-168

内容感知图像缩放技术综述

施美玲^{1,2} 徐丹¹

1. 云南大学信息学院, 昆明 650091; 2. 曲靖师范学院物理与电子工程学院, 曲靖 655011

摘要: 随着不同分辨率和纵横比的显示设备的迅猛增长, 内容感知图像缩放技术逐渐成为图像处理领域新的研究热点之一。内容感知图像缩放的目标是在任意改变图像大小时保持图像中的主体特征不变。围绕其关键步骤: 图像重要度识别和基于重要度的缩放, 先概述重要度识别的相关方法, 然后重点综述基于重要度的缩放技术。根据缩放是在像素级上操作还是亚像素级上操作, 或者两者兼有, 将其分为基于线裁剪缩放、基于图像变形缩放和多操作缩放 3 类, 并比较各类方法的优缺点, 同时给出各类方法所适合处理的图像类型。最后在分析各类研究方法的基础上, 给出了内容感知缩放技术的可能发展方向。

关键词: 内容感知图像缩放; 重要度图; 显著度图; 线裁剪; 基于图像变形缩放; 多操作缩放

Survey on content-aware image resizing techniques

Shi Meiling^{1,2}, Xu Dan¹

1. School of Information Science and Engineering, Yunnan University, Kunming 650091, China;

2. School of Physics and Electrical Engineering, Qujing Normal University, Qujing 655011, China

Abstract: With the rapid growth of the diversity and versatility of display devices which all come in different aspect ratios and resolutions, the content-aware image resizing has become one of the hot research fields in image processing. The main objective of such a technique is to preserve the image features when changing its size. The most content-aware image resizing methods have two basic steps: content significance recognition and image resizing based on a significance map. Firstly, the classic methods of the significance recognition are surveyed. Secondly, the resizing technologies based on significance maps are summarized. According to resizing methods based on pixel level discrete operating, sub-pixel level continuous operation or combination of the discrete operating and continuous operation, it can be categorized into image resizing based on seam carving, image resizing based on warping or multi-operator resizing. Thirdly, the algorithm effect comparisons between classes are given and their suitable image types are presented. Finally, future directions are discussed.

Key words: content-aware image resizing; importance map; saliency map; seam carving; image warping; multi-operation

0 引言

传统图像缩放如最近邻域、双线性或双三次插值等技术在改变图像长宽比缩放时会引起图像中的

主体对象产生压扁或拉伸的失真畸变。而另一技术——裁剪, 不可避免地产生了图像内容的丢失。为了解决这些问题, 在 2007 年 Siggraph 上, Avidan 和 Shamir^[1] 提出内容感知图像缩放技术。不同于传统方法忽视图像内容, 内容感知图像缩放通过自动

收稿日期: 2011-02-22; 修回日期: 2011-06-15

基金项目: 国家自然科学基金项目(60663010); 高等学校博士学科点专项科研项目(20095301110006)

第一作者简介: 施美玲(1982—), 女, 云南大学信息学院通信与信息系统专业博士研究生, 主要研究方向为数字图像处理和计算机视觉。E-mail: yuling5581@163.com

识别图像中各个区域的重要度,对重要度高的主体区域尽量保持不变或保持统一缩放,而让重要度低的同构背景区域尽可能多地吸收因缩放引起的形变。因此,内容感知图像缩放方法一般都包括两个步骤:图像重要度识别和基于重要度的缩放。图像重要度反映的是人眼对图像中不同内容区域变化的敏感程度,通常基于图像梯度和图像显著度来计算。其中,图像显著度反应的是图像中不同内容区域对人眼的吸引程度,属于视觉感知的范畴。许多基于视觉感知的应用如图像检索、图像分割、目标检测、图像感兴趣区域提取等均对图像显著度有深入地研究,因此,现有的内容感知图像缩放技术方面的文献主要侧重于第二阶段的缩放技术。

1 图像重要度识别方法

内容感知图像缩放属于不适宜问题,因为其结果完全依赖于对图像内容及内容重要性的定义。显然,不同的观察者会对图像有不同的理解,同时根据具体应用也会对图像缩放有不同的要求。因此,通过用户交互学习定义重要度是获取图像重要度最理想的方式,但是缩放在图像处理中属于相对简单且常用的操作,用户通常不愿意参与交互。因此,大部分内容感知图像缩放方法都采用一些启发式的操作来自动生成每个像素对应的重要度值。

起初,图像梯度由于其计算简单而被直接用作重要度,即基于梯度的重要度图^[1-3]。在该重要度图中,梯度大的像素被赋予高的重要度值,相反,梯度小的区域就被认为是不重要的背景同构区域。该模型的优势是计算简单、速度快;不足之处是:1)只对主体对象的边缘敏感,致使对象内部有较低的重要度值而被错误地标识为背景同构区域;2)包含大量琐碎的细节信息如树叶、波纹等,影响对背景同构区域的正确识别。

针对上述问题,许多研究者纷纷引入视觉显著度作为进一步的衡量标准,用定量的方式衡量人眼注意力焦点在图像中的位置。在其计算模型中,有由数据驱动、独立于任务的自底向上的注意模型,也有受意识支配、依赖于任务的自顶向下的注意模型。其中,自底向上模型的经典代表是 Itti 等人^[4]结合多尺度图像特征提出的以生物认知体系为指导的视觉关注度计算模型。而 Ma 等人^[5]的计算模型由于其速度快,效果与 Itti

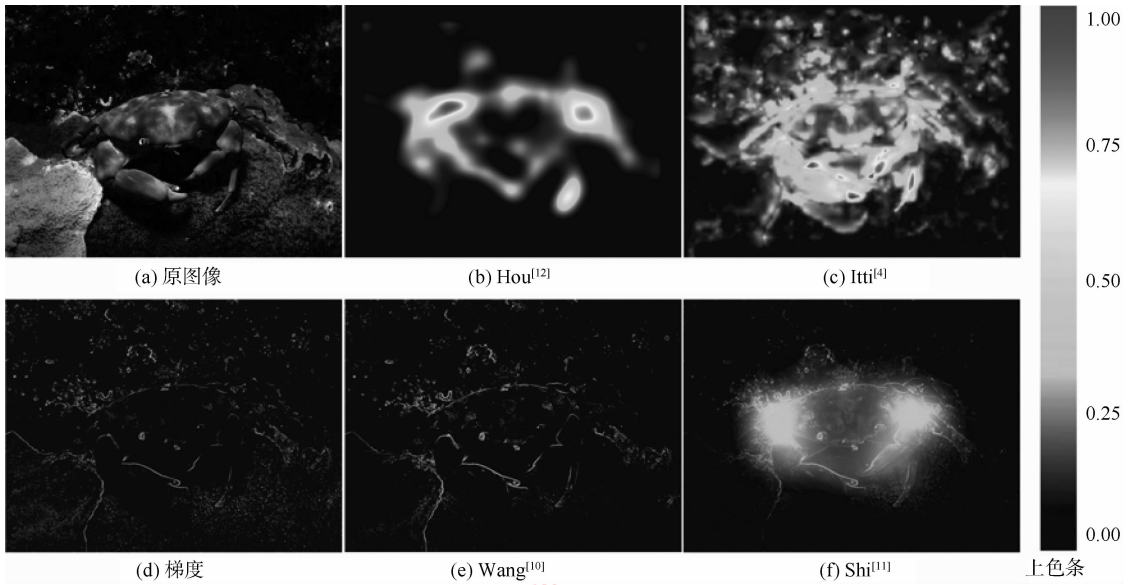
等人的不相上下而多次被用在内容感知缩放^[6-8]中。文献[7-8]合并 Ma 等人的低层显著度图及基于 Adaboost 模型^[9]的人脸检测结果来生成图像重要度。从整体上来看,基于显著度生成的重要度图能相对精确地识别出图像中的主体对象,但忽略了背景区域中的结构信息。

目前,结合梯度和视觉显著度生成重要度成为一种趋势。Wang 等人^[10]提出以梯度和显著度的乘积作为图像重要度。本质上,该方法只对梯度值和显著度值都大的像素才赋予较高的重要度值,等价于从基于梯度的重要度图中滤除琐碎的细节信息。因此,该方法同样无法正确识别出主体对象的内部区域。而 Shi 等人^[11]认为用梯度和显著度的加权平均来计算重要度更合理,并进一步改进基于稀疏特征的单分辨率显著度模型^[12]为多分辨率模型。Shi 等人方法可以相对较好地识别出主体对象的内部区域同时又保存背景区域中的结构信息。当然,对琐碎细节信息的滤除没有 Wang 等人的好。

上述提到的重要度都是以单一能量的方式作为缩放阶段的约束条件。但是,文献[13]认为这种简单的能量融合方式会混淆各种能量的作用,提出采用各个能量相互独立的多能量模板约束策略,即对于优化方程中的同一个变量,根据与目标函数或约束条件相关联的图像内容性质而选取相应的能量进行约束。

在处理主体对象与背景区域在颜色、亮度、纹理等特征上比较相似的图像时,所有自动的重要度模型很难得到精度的结果。此时,手动或半自动的交互是一种很好的补充。如文献[14]由用户在 Photoshop 中交互指定需要保护的特征区域。文献[15]提供可选的自动检测^[16]和手动交互方式来指定需保护的直线。

图 1 对几种典型的图像重要度生成方法进行了比较。图 1(a)是原图像,螃蟹是其主体对象,应该有较高重要度值。图 1(d)是基于梯度的重要度图,图 1(b)(c)是基于显著度的重要度图,图 1(e)(f)是结合梯度和显著度的重要度图。从图中可以清楚地看出,基于梯度的重要度图主体不突出且具有较多的琐碎细节信息;基于显著度的重要度图中主体相对较突出,但没识别出背景中的边缘线条。结合梯度和显著度的结果是最好的,主体和边缘线条相对突出。

图1 图像重要度图比较^[11]Fig. 1 Compare the image significance maps^[11]

2 基于重要度的缩放技术

在缩放阶段,内容感知缩放技术的核心是尽量让图像中具有高重要度值的主体对象区域保持不变,同时改变重要度较低的同构背景区域来满足缩放目标大小的几何约束。若仅仅考虑如何满足目标大小的几何约束,传统缩放和裁剪提供了两种典型的技术。裁剪是一种离散的基于像素的0-1式开关操作,即裁剪窗口内的像素完全保留而窗口外的像素完全删除。传统缩放是一个连续的亚像素级的插值重采样过程,会产生由邻近像素合并得到的新像素值。目前,绝大多数内容感知缩放技术都是基于这两种操作。其区别在于:1)用何种策略来选择0-1操作方式中需删除(放大时采用复制)像素的位置;2)用什么优化准则来确定图像中不同区域对应的采样率;3)在两种技术结合的方法中用何种标准来选择每个操作的数量及操作顺序。由此,根据不同的技术特点,将主要的内容感知缩放方法分为3类:基于线裁剪缩放、基于图像变形缩放和多操作缩放。

2.1 基于线裁剪缩放方法

基于线裁剪缩放技术是由 Avidan 和 Shamir^[1] 在 2007 年 Siggraph 会议上提出的。所谓线裁剪是指通过连续删除(插入)连通的水平(垂直)穿过整

幅图像的裁剪线(seam)来实现图像缩放的技术。其中,裁剪线的宽度为1个像素,删除裁剪线时缩小图像,反之,复制裁剪线则放大图像。处理水平(垂直)裁剪线则改变图像高度(宽度)。为了保持图像内容的主要特征,线裁剪算法每次都删除或插入能量最小的裁剪线。所谓裁剪线能量是指该裁剪线所包含像素的重要度值的总和。

文献[1]中线裁剪算法也被称为 Backward 方法。不失一般性,以垂直裁剪线为例介绍该算法思想。Backward 方法以像素梯度值作为图像的重要度。对大小为 $n \times m$ (n 为列数, m 为行数)的图像,垂直裁剪线 S^v 表示为

$$S^v = \{S_i^v\}_{i=1}^m = \{(i, x(i))\}_{i=1}^m \quad (1)$$

$$\text{使得 } \forall i, |x(i) - x(i-1)| \leq 1$$

式中, $x(i)$ 是映射: $x[1, \dots, m] \rightarrow [1, \dots, n]$ 。从式(1)可以看出,垂直裁剪线是从第1行到最后一行,且每行提供一个像素组成的8连通路经。

在假设图像中位于 (i, j) 位置的重要度为 $e(i, j)$ 的情况下,文献[1]给出长度为 m 的垂直裁剪线能量为: $E(s) = \sum_{i=1}^m e(i, x(i))$, 并将能量最小的垂直裁剪线 S^* 定义为

$$S^* = \min_s E(s) = \min_s \sum_{i=1}^m e(i, x(i)) \quad (2)$$

文献[1]中能量最小裁剪线的寻找采用动态归化算法。其中像素点 (i, j) 的累积能量 $M(i, j)$ 为

$$M(i, j) = e(i, j) + \min(M(i-1, j-1), M(i, j-1), M(i+1, j-1)) \quad (3)$$

文献[2]注意到 Backward 方法只考虑删除像素具有的能量而忽略了删除像素时增加的能量。于是,提出 Forward 线裁剪算法,该算法增加了由于删除像素而使非相邻像素点成为新相邻像素点时产生的能量。如图 2 所示,删除垂直裁剪线时将出现 3 种新的相邻像素,对应增加的能量分别为

$$\begin{aligned} C_L(i, j) &= |I(i, j+1) - I(i, j-1)| + \\ &\quad |I(i-1, j) - I(i, j-1)| \\ C_U(i, j) &= |I(i, j+1) - I(i, j-1)| \quad (4) \\ C_R(i, j) &= |I(i, j+1) - I(i, j-1)| + \\ &\quad |I(i-1, j) - I(i, j+1)| \end{aligned}$$

式中, $I(i, j)$ 为像素点 (i, j) 位置的颜色值。相应地, Forward 方法中寻找最小能量裁剪线的累积能量项由 Backward 方法中的能量项修改为

$$M(i, j) = e(i, j) + \min \begin{cases} M(i-1, j-1) + C_L(i, j) \\ M(i, j-1) + C_U(i, j) \\ M(i+1, j-1) + C_R(i, j) \end{cases} \quad (5)$$

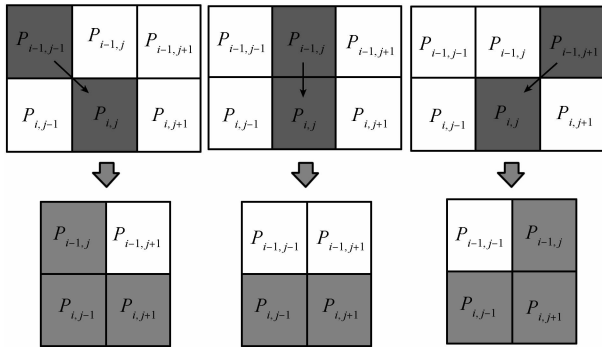


图 2 删除像素时新增的能量^[2]

Fig. 2 The newly added energy when delete pixel^[2]

文献[11]提出双向线裁剪(Bi-direction)算法。该算法将沿图像单一方向(如宽度)的缩小(放大)操作转化为两个方向上的操作之和:1)删除(复制)垂直裁剪线;2)复制(删除)水平裁剪线之后保持长宽比统一地缩小(放大)图像。其中,处理水平裁剪线和垂直裁剪线的数量则根据图像内容来自适应地产生。

文献[17]提出连续线裁剪算法。该算法总是按一定比例先找出多于欲删除或复制数量的裁剪线,然后让每条裁剪线根据其能量大小来承担删除(复制)量,即决定重采样的密度。文

献[18]可以一次删除宽度大于 1 的裁剪线,能够实行对主体对象平移的功能,因此该算法也被称为 shift-map 方法。文献[19]考虑到同时删除或复制多条相邻裁剪线容易破坏图像的全局视觉效果。故在删除某条裁剪线时,将删除像素的能量按一定的权重加到相邻像素上,如此裁剪线的能量在一定程度上记忆了已删除像素的信息。文献[20]则通过局部调整穿越图像特征点的裁剪线来达到保持图像特征的目的。

在上述基于线裁剪的缩放方法中, Backward 方法^[1]的地位是毋庸置疑的,它的开创性、创新性是其他方法无法比拟的。目前, Backward 方法的作者 Avidan 已加入 Adobe 公司, Backward 方法以内容识别比例工具成为 Photoshop CS4 的新功能。 Forward 方法^[2]对 Backward 方法的改进效果是显著的,能有效消除该方法产生的人工噪声点,见图 3(a)^[2],房子的结构线条和车头在 Backward 方法的结果中出现弯曲噪点,在 Forward 方法的结果中得到了改善。但是,这两种方法在缩放图像时都只利用了一个方向的背景同构信息。因此,对于垂直向(水平向)背景同构信息较少的图像,这两种方法都无法得到好的缩放宽度(高度)的结果,见图 3(b)^[11],两种方法的结果在车身上出现很多噪点,车身的整体形状也发生了畸变。 Bi-direction 方法^[11]通过充分利用图像中两个方向上的背景同构信息而使结果得到了很好的改进。 Backward、 Forward 和 Bi-direction 方法处理的裁剪线都是一个像素宽的,操作非常精细,可以很好地利用小面积的背景同构区域,但是容易破坏图像中的细节信息,见图 3(c)^[18], Forward 方法的栅栏已变得破碎,没能保持其原有的结构,同时人物的体形也被明显地拉长。 Shift-map 方法通过一次性处理多像素宽的裁剪线,能够很好地保持图像中原有物体的结构,不足之处是会出现明显的内容丢失,见图 3(d), 7 个人最后只剩下 4 个。

2.2 基于图像变形缩放方法

基于图像变形的缩放^[3,6-8,10,13-15,21-22]是另一种有效的内容感知图像缩放技术。该技术先将图像划分为网格,然后在目标图像大小约束下,以保持图像中重要度高的主体区域内网格尽量不变或尽量做长宽等比统一缩放而让拉伸变形尽可能发生在重要度较低的背景同构区域的准则下进行优化生成新的变形网格,最后在对应网格间进行逆向映射生成目标图像。

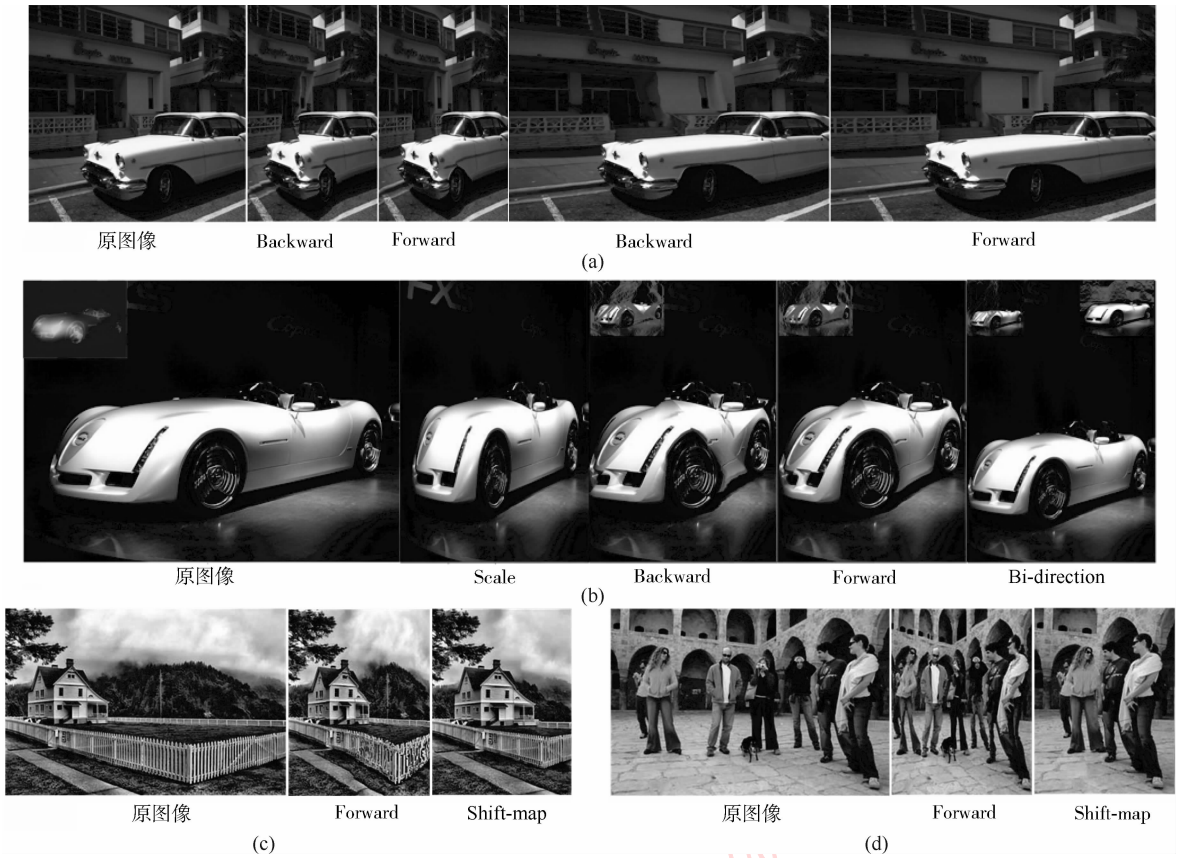


图 3 基于线裁剪方法结果比较

Fig. 3 Compare the image resizing results by different methods based on seam carving

在 2008 年 Siggraph Asia 会上, Wang 等人提出 optimized scale-and-stretch (OSS) 图像缩放技术^[10]。该技术先将原图像划分为矩形网格 (V, E, F), V 为顶点集, E 为边集, F 为四边形网格集。 V 定义为 $V = [v_0^T, v_1^T, \dots, v_{end}^T]$, $v_i \in \mathbf{R}^2$, 表示初始顶点坐标。在改变图像大小时, OSS 方法固定左上角顶点 v_0 位置, 让用户移动右下角顶点 v_{end} 到 v'_{end} 。接着, OSS 方法将缩放问题转换为寻找满足几何约束的新网格顶点位置 $V' = [v_0^T, v'_1, \dots, v'^T_{end}]$ 的问题。对于单个网格 $f(f \in F)$ 的形变量, OSS 方法计算变形后的网格 V' 与原始网格 V 的某个统一缩放版本 $s_f v + t$ (其中, s_f 为缩放因子, t 为平移量) 之间的距离来衡量网格形状上的改变量 $D_u(f)$:

$$D_u(f) = \sum_{\{i,j\} \in E(f)} \| (v'_i - v'_j) - s_f(v_i - v_j) \|^2 \quad (6)$$

最终, 每个网格形变量对总体网格形变量 D_u 的影响与对应网格的重要度 w_f 有关。其中, w_f 被定义为网格所包含像素的重要度平均值。 D_u 为

$$D_u = \sum_{f \in F} w_f D_u(f) \quad (7)$$

根据大部分图像中主体对象通常占据多个连续网格的特点, OSS 方法通过最小化网格线弯曲度能量来保持主体对象的特征。其优化策略是在保持网格边方向的同时调整其长度。OSS 方法定义的网格线弯曲度能量为

$$D_l = \sum_{\{i,j\} \in E} \| (v'_i - v'_j) - l_{i,j}(v_i - v_j) \|^2$$

$$l_{i,j} = \frac{\| (v'_i - v'_j) \|}{\| (v_i - v_j) \|} \quad (8)$$

最后最小化网格形变量及网格线弯曲度能量之和

$$D = D_u + D_l \quad (9)$$

在计算时, OSS 方法从网格顶点 v' 的初始猜测开始, 局部优化计算每个网格的统一缩放因子 s_f , 之后在目标图像矩形特性的边界条件约束下全局优化更新网格顶点 v' , 以此迭代求解。

在网格形变量模型中, OSS 方法着重考虑了网格形状的变化, 而 Shi 等人^[21]增加了衡量网格大小

变化的能量项并将网格形变量模型修改为

$$D_u(f) = \sum_{|i,j| \in E(f)} \left(\|(\mathbf{v}'_i - \mathbf{v}'_j) - s_f(\mathbf{v}_i - \mathbf{v}_j)\|^2 + \|s_f(\mathbf{v}_i - \mathbf{v}_j) - \alpha(\mathbf{v}_i - \mathbf{v}_j)\|^2 \right) \quad (10)$$

通过考虑网格大小的变化量, Shi 等人^[21]方法

可以有效避免 OSS 方法中出现的主体对象大小随图像大小本身的缩放而缩放的问题, 并且通过简单的调整参数 α , 能在缩放过程中根据用户需求来控制主体对象的大小(图 4)。

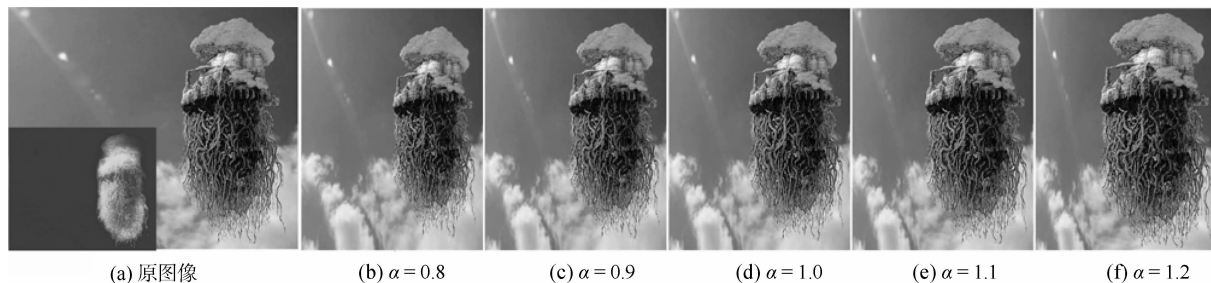


图 4 Shi 方法通过调整参数控制主体对象大小的结果^[2]

Fig. 4 The image reduction results with different size of the prominent object created by Shi method^[2]

按 OSS 方法划分网格时, 无法做到网格顶点与图像特征完全重合, 极易出现同时包含主体对象及背景区域像素的网格, 不利于局部特征的保留。针对这种情况, 文献[22]通过识别主体对象的边界特征点, 将包含特征点的网格进一步与特征点相连生成不同的控制结点, 最后通过优化控制结点变形来保持图像特征。文献[23]改进文献[22]的优化函数来避免重要区域的扭曲和旋转。

文献[3, 15]同 OSS 方法相似, 也将图像划分为矩形网格。但在描述网格长宽等比缩放时不像 OSS 方法使用统一的缩放因子 s_f , 而是区分 x 轴、 y 轴的缩放因子 $a_{i,j}^1$ 、 $a_{i,j}^2$, 然后根据网格重要度, 调整相应网格中两轴缩放因子的比值 $r_{i,j} = a_{i,j}^1/a_{i,j}^2$ 与 1 的接近程度来控制对应网格的形变量。对于重要度高的网格, 尽量使其 r 与 1 接近来保持网格的长宽比。相反, 低重要度网格的 r 则允许与 1 背离较大, 如此来满足图像大小改变的需要。

文献[3, 15]的算法思想及网格能量项都是相似的, 主要区别在于优化网格能量项的求解上。文献[3]以迭代方式先局部优化单个网格能量项求解缩放因子 $a_{i,j}^1$ 、 $a_{i,j}^2$, 再全局优化所有网

格能量项总和来求解网格顶点变形后的坐标位置, 被简称为 LG (local-global) 方法。而文献[15]则使用 convex quadratic program 方法来求解, 简称为 QP 方法。

不同于上述方法用矩形网格划分图像, 文献[7-8]用三角网格来划分图像。三角网格中的顶点由等长取样的边界点、图像特征点及平滑区域内通过最长边长限制添加的顶点组成。网格变形时, 主体对象区域内的三角形顶点根据其重心做刚体变换, 其余顶点在边界及结构约束下, 根据平面网格完全由其所有边的边长决定的思想下通过参数化来求解。

不同于基于矩形和三角形网格的缩放方法将区域边缘分解成小段来近似表示, 文献[13]考虑从整体上保持区域边缘的形态, 提出一种根据均质区域边缘构建网格的曲边梯形网格表示方法。该方法首先基于 mean-shift 算法^[24]划分出来的均质区域构建曲边梯形(图 5), 然后通过最小化网格重要度权重约束下的网格面积变化来达到突出重要物体的目的和最小化梯度能量约束下的网格变形量来达到消除明显视觉扭曲的目标。



图 5 图像的网络剖分^[13]

Fig. 5 The different kinds of image mesh^[13]

文献[6]根据鱼眼镜头的视觉特点提供了一种既强调重要区域又保存图像内容的缩放技术。该技术先在原图像中确定包含重要内容的矩形 ROI 区域,并根据目标图像大小以保持 ROI 区域长宽比的方式缩放 ROI 区域后确定 ROI 区域在目标图像中的位置,接着以 ROI 区域的中心为中心,生成与中心距离成反比的变形函数。该变形函数是一个在 3 个确定对应关系约束下拟合而成的非线性函数,即原图像的边界与目标图像边界对应、原图像中 ROI 区域的中心与目标图像 ROI 区域的中心对应、原图像中 ROI 区域的边界与目标图像中 ROI 区域的边界对应。最后,根据变形函数逆向映射生成目标图像。

在基于图像变形的缩放技术中,主要用到规则的矩形网格、不规则的三角形网格和曲边梯形网格。其中,三角形和曲边梯形网格的优势在于可以根据图像内容特征点来划分网格,有利于达

到同一网格内像素的重要度值相对平滑的优化前提,但是对网格形变量的衡量不如矩形网格方便、简单,常常要寻找重心位置。相反,用矩形网格划分图像时,网格顶点很少与图像特征点重合,导致主体对象边界附近的网格经常由重要度比较大的像素组成,使得由像素重要度平均得到的网格重要度不能有效表示实际像素的重要度而影响优化效果。但是,矩形网格只要保持长宽等比缩放就能保持其形状,网格形变量优化函数相对简单、直观。在效果上,OSS 方法^[10]通过使重要区域中的网格尽量保持长宽等比的缩放策略,能够很好地保持图像中主体对象的形状。但是该方法没有考虑不同缩放因子时网格大小差异带来的影响,因此结果图像中的主体对象大小随图像大小本身的缩放而缩放,如图 6 中的啤酒酒店在结果图中被缩得很小而与图像中其他建筑物的比例失调,影响图像的全局视觉效果。



图 6 基于图像变形缩放方法结果比较^[15,21]

Fig. 6 Compare the image resizing results by different methods based on image warping^[15,21]

LG 方法^[3]由于采用局部全局优化迭代的方法产生结果,所以生成结果在局部已达最优时无法保持全局也最优,故生成结果的效果不稳定。如啤酒酒店被拉伸非常明显,没能保持原来的高宽比(图 6)。QP 方法^[15]用非凸二次规划法求出确定的全局优化解,同时又通过增加阻止网格折叠、放大重要区域等能量项,获得了相对较好的

缩放结果(图 6)。Shi 方法^[21]在衡量网格形变量时同时考虑其形状和大小的变化,得到了较好的结果,在图 6 中,Shi 方法的效果最好。另一方面,QP 方法^[15]在自动检测或手动指定图像中的直线特征后,通过在优化函数中增加相应能量项求解能很好地保持图像中的直线特征,如图 7 所示。



图 7 QP 方法放大结果^[15]

Fig. 7 The image magnifying results of QP method^[15]

基于鱼眼镜头视觉效果缩放图像^[6]的结果如图 8 所示。在该方法的结果中,ROI 区域的特征得到了很好的保持,但对其余图像内容的处理只考虑与 ROI 区域中心点距离的关系。如图 8 中明暗相间的纹理图所示,ROI 区域外的图像内容,距 ROI 区域的中心点越远,内容越不清晰,到边界点时,甚至会被完全压缩。此外,该方法只适合于单一 ROI 区域图像,不能处理含有多个 ROI 区域的图像。



图 8 基于鱼眼镜头视觉效果缩小结果^[6]

Fig. 8 The image reduction results with fisheye-view warping^[6]

文献^[13]主要有两个创新,多能量模板约束策略和基于曲边梯形网格及区域关系图的缩放方法。该算法的缩放效果很大程度上取决于前期 mean-shift 方法^[23]分割的结果。

2.3 多操作图像缩放方法

多操作图像缩放方法^[25-29]结合线裁剪、传统插值缩放和裁剪 3 种方法中的前两种或三种来实现内容感知缩放。该类方法的重点是确定各个操作的顺序及数量。最直接的思路是穷举所有不同操作顺序及操作数量下的缩放结果,然后从中选择最优的结果。倘若允许缩小和放大交叉进行,则可列举地执行序列是无穷的。通常在缩小(放大)图像时,都预先规定每个子操作都只能做缩小(放大)处理。在这一前提下,若采用 m 种操作在一个方向上改变图像大小为 w 时,可选的操作序列为 m^w , 仍然比较大,因此现有的多操作缩放方法都会给出一些额外假设来减少计算量,常用的假设有两种:一是寻找常规路

径(regular path)^[25],根据操作数量对结果的影响比操作顺序对结果的影响大的特点,通过固定操作顺序将问题简化为寻找最优操作数量;另一种是寻找混合路径(mix path)^[25],先固定好每次子操作的数量 $k(k \geq 1)$,从而进一步确定操作序列的长度 n ,然后用动态归化法确定出操作序列中每步的操作类型。此外,在找最优操作类型时对各个子操作数量相同但顺序不同的序列只保留最优值以此来减少计算量。在确定了寻找操作序列的顺序及数量的方法后,多操作方法的难点转化为如何设计出一种好的度量准则来衡量缩放后的图像与原图的相似度,这也是各种多操作缩放方法的主要差别所在。

2009 年的 SIGGRAPH 上,Rubinstein 等人提出了多操作图像缩放方法^[25]。该方法通过统计分析用户对同一幅图像用不同的缩放方法(单一的线裁剪、传统插值缩放、裁剪及 3 种方法结合)进行操作后所得结果的满意程度,得出一般情况下多操作的结果优于单一操作结果的结论。

Rubinstein 等人的多操作方法^[25]运用双向变形量(BDW)来衡量缩放后的目标图像与原图像间的差异度。BDW 同时包括目标图像到原图像的差(反应目标图像与原图像间的偶合性,即目标图像中带入了多少不属于原图像的噪声)及原图像到目标图像的差(反应目标图像对原图像内容保留的完整度,即原图像中有多少内容在目标图像中没有)。BDW 先将图像分块,然后以行为单位再对原图像(目标图像)中的每块用动态归化法找出目标图像(原图像)中对应行上最相似的块。其中,最优匹配块的查找遵循对应块的顺序与原来块的顺序一致,且只允许多对一,不允许一对多的匹配原则。最后,计算所有块中对应像素色彩差的平均作为图像间的差异度(图 9)。根据该差异度值,文献^[25]提出规则路径和混合路径来找出最优多操作序列并得出相应缩放结果。

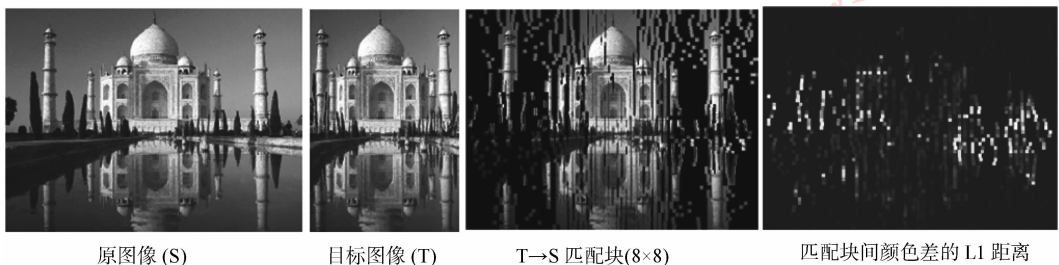


图 9 使用 BDW 找优化匹配块^[25]

Fig. 9 Find optimal matched patches by BDW^[25]

在国内,董未名等人^[26-28]对多操作也有非常深入的研究。其中,前期文献[26-27]仅结合线裁剪和传统插值统一缩放两种技术来实现多操作。文献[26]始终从线裁剪开始缩放图像,当遇到以下两种情况时关闭线裁剪实施一定数量的插值统一缩放,然后转到线裁剪,开始新一轮的迭代直到缩放到目标大小。关闭线裁剪的条件为:1)遇到裁剪线的能量突变时,可能遇到了主体对象的边界所以停止线裁剪操作;2)裁剪线能量变化均匀但连续有多条裁剪线被裁剪时,为了避免破坏全局视觉效果甚至完全删除某一对象而停止线剪裁操作。至于对突变能量及裁剪线数量的阈值划分则由实验得到的经验参数(放大3个,缩小4个)预先设定。文献[27]与文献[25]类似,也对图像分块,用对应块间能量差的和给出两幅图像间的差异度,但块间能量用双向图像欧几里得距离来计算。

文献[28]通过加入裁剪将操作类型增加到了3种。该方法采用操作能量及主色调(DCD)的损失来定义当前每种操作的成本,并选择成本最低的操作作为下一步的操作。文献[28]将操作能量 $E(s)$ 定义为

$$E(s) = \frac{1}{2} \left(\frac{1}{N_s} \sum_{i=1}^{N_s} (e(I))(s_i) + \max_{1 \leq i \leq N_s} (e(I))(s_i) \right) \quad (11)$$

式中,能量 $e(I)$ 通常为当前像素位置的梯度, s_i 为操作域 S 上的像素, N_s 为操作域上像素的个数。对于线裁剪,操作域 S 为当前具有最小能量的裁剪线包含的像素;裁剪的操作域为边界上的像素;传统缩放的操作域为整幅图像。

另一方面,文献[28]将DCD描述为表示主色

调颜色矢量的 v_i ,分布概率的 p_i 和方差 c_i 组成的三元组 $\{v_i, p_i, c_i\}, i = 1, 2, \dots, N_{\text{DCD}}, N_{\text{DCD}}$ 是主色调的数目,一般取16或32。定义主色调的损失 $D(s)$ 为

$$D(s) = 1 - \frac{1}{N_s} \sum_{i=1}^{N_s} d(I(s_i) - v_{j^*(i)}) \times (1.0 - p_{j^*(i)}) \quad (12)$$

式中, d 是 L_2 距离,操作域 S 上的像素 s_i 对应的主色调的索引 $j^*(i)$ 为

$$j^*(i) = \arg \min d(I(s_i) - v_j) \quad (13)$$

$$1 \leq j \leq N_{\text{DCD}}$$

最后,操作 O_i 的操作成本被定义为

$$C(O_i(I)) = (1.0 - w_{\text{DCD}})E(s') + w_{\text{DCD}}D(s') \quad w_{\text{DCD}} = 0.2 \quad (14)$$

2008年Simakov等人^[29]提出的SSD缩放技术虽不属多操作,但与文献[25,27]较相似也使用了相似块匹配的方法来优化缩放结果,并且其能量项也能用作多操作中的优化能量。

上述提到的多操作方法在效果上优于或等价于包含于其中的任何一种单操作的结果,如图10的男孩和雪人图所示,后两张多操作的结果从全局视觉信息的保留及维持主体特征形状的角度综合考虑,效果比单操作的好。当图像中的重要区域分布相对集中时,裁剪可以得到不错的缩小效果,因此,增加裁剪的多操作方法^[25,28]在效果上比只包含线裁剪和传统插值缩放的方法^[26-27]好。文献[26]方法对缩小用到了4个(放大3个)阈值参数,虽然根据实验经验,文中给出了一些相对稳定的参数值,但经常要根据图像内容来调整,并且参数间耦合性较大,调整起来不方便。



图10 多操作方法结果比较^[28]

Fig. 10 Compare the image resizing results by different methods based on multi-operator^[28]

文献[25,27,29]应用匹配块间的差异来衡量两个图像间的偏离度,这种度量相对精确些,但速度

非常慢。文献[28]给出缩放 500×333 大小的图像,使用2~4个操作时,所需时间是10~40s,而

文献[25]方法所需时间却是 2~10 min。

3 内容感知缩放技术效果比较

上述介绍的 3 类图像缩放技术中,基于线裁剪和基于图像变形的方法在速度上明显快于多操作,尤其快于基于块匹配的多操作。从效果上看,每种方法各有其优势,各有适合的图像类型,没有一种方法是通用的,并且也不能相互取代。

线裁剪的优劣均来自于其像素级离散操作的特点。因为是像素级的,其结果是精细的;因为是离散的开关操作,其结果中保留的颜色值与原图是最接近的。但也因为是离散的,很容易破坏图像中的结构信息,产生人工噪声。因此,线裁剪比较适合于背景区域过渡平滑、结构性不强的图像,如图 11 中女孩图像的缩放结果对女孩区域特征的保留相对较好。此外,线裁剪方法也可用于对象删除。

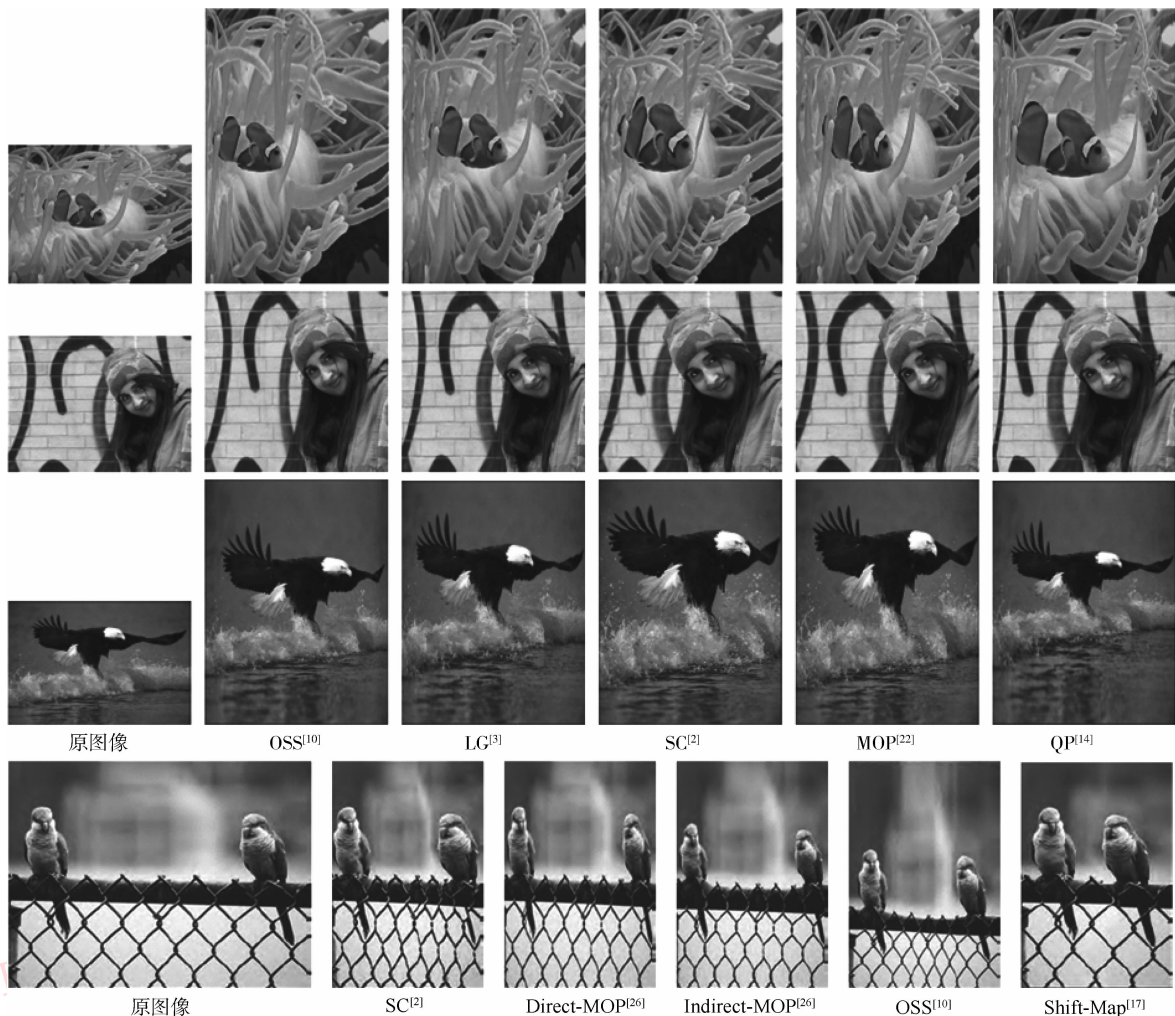


图 11 典型缩放方法结果比较^[15]

Fig. 11 Compare the image resizing results by typical methods^[15]

多操作的实质是在权衡主体对象变形、结构信息破坏、产生人工噪声、图像内容丢失等情况对图像整体效果的影响程度,然后再折中选择。通常,多操作较理想的情况是先对图像做线裁剪操作,当线裁剪破坏图像结构信息、引入人工噪声较多或过度删除某区域的背景信息而影响图像全局视觉效果时,

放弃线裁剪,转用统一缩放。此时,统一缩放的连续性可以保持图像中各个部分的相对比例,也可抑制人工噪声的产生。但是,绝大部分多操作的优化过程都找不到最优结果,只能得到相对较优结果。因此,多操作结果经常给人一种不是最理想,但还行的感觉。在多数情况下,多操作是线裁剪结果相对较

差时的一种不错选择,如图 11,多操作结果中鱼形状的保持比线裁剪结果好。

基于图像变形的缩放的重要度及边界条件约束下,通过优化网格形变量的总和,给出每个局部区域网格不同的缩放因子,具有插值缩放的连续性,能够很好地保持图像中主体对象的特征。如图 11 中鹰的特征保持得最好。但是当网格退化为一个点或一条线时,对应区域的过度压缩也会产生图像内容的不连续,如图 11 中金鱼图像的左边区域。

上述 3 类方法都很难较好地保持图像中任意方向的直线特征,如图 11 中的鸚鵡。对于背景同构信息所占比例相对较少的图像,也很难得到好的结果。此外,最终缩放结果的好坏也在很大程度上依赖于图像重要度识别的精确程度。至于如何评估图像缩放结果,目前没有也难设计出定量的标准,毕竟图像质量评估与个人主观感觉和具体的应用密切相关。在此,给出对 3 种图像缩放方法的一些定性比较(如表 1 所示)。

表 1 3 种缩放方法的定性比较

Table 1 The qualitative comparison to three kinds of image resizing method

名称特性	基于线裁剪	基于图像变形	多操作
解性质	离散	连续	离散 + 连续
重要度图	需要	需要	需要
可利用的 冗余信息	与大小改变方 向垂直的方向	水平、垂直 两方向	水平、垂直 两方向
人工噪声	较多	较少	中间
主体变形	较少	较多	中间
内容丢失	较多	较少	中间
速度	快	快	慢
适合图像	平滑背景区域 垂直于图像大 小改变方向;缩 放比例小	主体与背景对 比鲜明;主体纹 理复杂,背景纹 理简单	线裁剪引 入大量人 工噪声的 图像

4 结 论

经过对各种图像缩放方法的分析研究,认为缩放方法的效果在很大程度上取决于重要度识别的精确程度,即现有的基于重要度的缩放方法基本上能满足缩放要求,但重要度识别的精确度不高。因此,

重要度识别算法的研究空间比缩放阶段算法的研究空间大。对于重要度识别算法的研究,目前常用的自底向上识别方法的精度还很难满足图像缩放的需求,融合目标驱动的自顶向下的识别结果是种不错的选择。除此,加入一些简单的用户交互或是图像分割结果也能作为很好地补充。另一方面,内容感知的视频缩放也是目前的研究热点之一。如果将内容感知图像缩放技术扩展到视频缩放上,考虑帧间内容的连贯性,避免视频抖动和提高处理速度是其关键。

参考文献 (References)

- [1] Avidan S, Shamir A. Seam carving for content-aware image resizing [J]. ACM Transactions on Graphics, 2007, 26(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] Karni Z, Freedman D, Gotsman C. Energy-based image deformation [J]. In Computer Graphics Forum, 2009, 28(5): 1257-1268.
- [4] Iti L, Koch C, Neibur 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.
- [5] Ma Y F, Zhang H J. Contrast-based image attention analysis by using fuzzy growing [C] // Proceedings of the ACM Multimedia. New York, NY: Association for Computing Machinery, 2003: 374-381.
- [6] Liu F, Gleicher M. Automatic image retargeting with fisheye-view warping [C] // Proceedings of the 18th Annual ACM Symposium on UIST. New York, NY: Association for Computing Machinery, 2005: 153-162.
- [7] Shi J, Guo Y W, Du Z L, et al. A mesh parameterization-based image retargeting method [J]. Journal of Software, 2008, 19(sup.): 19-30. [时健, 郭延文, 杜振龙, 等. 一种基于网格参数化的图像适应方法 [J]. 软件学报, 2008, 19(增刊): 19-30.]
- [8] Guo Y, Liu F, Shi J, et al. Image retargeting using mesh parameterization [J]. IEEE Transactions on Multimedia, 2009, 11(5): 856-867.
- [9] Viola P, Jones M. Rapid object detection using Adaboost cascade of simple features [C] // Proceedings of the 2001 IEEE Computer Society Conference on CVPR. Florida, USA: IEEE, 2001: 511-518.
- [10] Wang Y S, Tai C L, Sorking O, et al. Optimized scale-and-stretch for image resizing [J]. ACM Transactions on Graphics, 2008, 27(5): 1-8.

- [11] Shi M L, Peng G Q, Yang L, et al. Optimal Bi-directional seam carving for content-aware image resizing [C] // Proceedings of the 5th International Conference on E-learning and Games, Edutainment. Berlin Heidelberg: Springer-Verlag, 2010: 456-467.
- [12] Hou X D, Zhang L Q. Dynamic visual attention: Searching for coding length increments [C] // Proceedings of the 22nd Annual Conference on NIPS. Cambridge, USA: MIT Press, 2008: 681-688.
- [13] Ren T W, Liu Y, Wu G S. Image retargeting based on region relation graph [J]. Journal of Software, 2010, 21 (0): 2237-2249. [任桐炜, 刘焱, 武港山, 等. 基于区域关系图的图像显示适配技术软件学报, 2010, 21(0): 2237-2249.]
- [14] Gal R, Sorkine O, Cohen-Or D. Feature aware texturing [C] // Proceedings of the Eurographics Symposium on Rendering. Goslar, Germany: Eurographics Association, 2006: 297-303.
- [15] Chen R J, Freedman D, Kami Z, et al. Content-aware image resizing by quadratic programming [C] // Proceedings of the 2010 IEEE CVPR WorkShop on NORDIA. Florida, USA: IEEE, 2010: 1-8.
- [16] Fernandes L, Oliveira M. Real-time line detection through an improved hough transform voting scheme [J]. Pattern Recognition, 2008, 41(1): 299-314.
- [17] Liu Z, Yan H B, Shen L Q, et al. Adaptive image retargeting using saliency-based continuous seam Carving [J]. Optical Engineering, USA, 2010, 49(1): 017002. 1-10.
- [18] Pritch Y, Kav-Venaki E, Peleg S. Shift-map image editing [C] // Proceedings of the 12th IEEE ICCV. Florida, USA: IEEE, 2009: 151-158.
- [19] Cho S H, Choi H, Matsushita Y, et al. Image retargeting using importance diffusion [C] // Proceedings of the ICIP. Florida, USA: IEEE, 2009: 977-980.
- [20] Deng H P. A Feature-Preserve Image Resizing Method Based on Seam Carving [D]. Zhejiang: Zhejiang University, 2008. [邓会鹏. 基于线裁剪的保特征图像缩放算法 [D]. 浙江: 浙江大学, 2008].
- [21] Shi M L, Yang L, Peng G Q, et al. A content-aware image resizing method with prominent object size adjusted [C] // Proceedings of 17th ACM Symposium on VRST. New York, NY: Association for Computing Machinery, 2010: 175-176.
- [22] Zhang G X, Cheng M M, Hu S M, et al. A shape-preserving approach to image resizing [J]. Computer Graphics Forum Special Issue of Pacific Graphics, 2009, 28 (7): 1897-1906.
- [23] Lei L X. Content-aware image resizing based on hybrid energy [J]. Chinese Journal of Computer, 2010, 33(10): 2015-2021. [雷励星. 基于混合能量的内容敏感图像缩放新方法 [J]. 计算机学报, 2010, 33(10): 2015-2021.]
- [24] Comanicu D, Meer P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619.
- [25] Rubinstein M, Shamir A, Avidan S. Multi-operator media retargeting [J]. ACM Transactions on Graphics, 2009, 28(3): 23.
- [26] Dong W M, Paul J C. Adaptive content aware image resizing [EB/OL]. [2011-01-08]. <http://liama.ia.ac.cn>.
- [27] Dong W M, Zhou N, Paul J C, et al. Optimized image resizing using seam carving and scaling [J]. ACM Transactions on Graphics, 2009, 28(5): 125.
- [28] Dong W M, Zhang X P, Zhou N, et al. Fast multi-operator image resizing [EB/OL]. [2011-01-08]. <http://liama.ia.ac.cn>.
- [29] Simakov D, Caspi Y, Shechtman E, et al. Summarizing visual data using bidirectional similarity [C] // Proceedings of the IEEE Conference on CVPR. Florida, USA: IEEE, 2008: 1-8.