

融合显著与深度信息的缝切割重定向方法

吴加莹¹, 杨 赛^{1,2}, 堵 俊¹, 董 宁¹

(1. 南通大学电气工程学院, 江苏南通 226019; 2. 南通先进通信技术研究院, 江苏南通 226019)

摘 要: 针对当前内容感知的重定向方法中可能出现的变形和失真问题, 提出一种融合显著与深度信息的缝切割重定向方法. 首先利用 GBVS 算法获取图像显著信息, 结合图像梯度信息与通过 SIFT 匹配方法获取的图像深度信息构建更精确的重要度图; 其次, 根据重要度图的能量分布, 对原始图像进行处理, 得到最终的重定向结果. 基于公开数据库在两个不同评价标准下与多种重定向方法的对比表明, 本文方法能够最大程度的保留图像的显著部分.

关键词: 显著信息; 深度信息; 缝切割; 图像重定向

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112 (2019)07-1547-04

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2019.07.020

Image Retargeting with Seam-Carving Combining Saliency and Depth Information

WU Jia-ying¹, YANG Sai^{1,2}, DU Jun¹, DONG Ning¹

(1. School of Electrical Engineering, Nantong University, Nantong, Jiangsu 226019, China;

2. Nantong Research Institute for Advanced Communication Technologies, Nantong, Jiangsu 226019, China)

Abstract: To solve the problem of deformation and distortion that may occur in the current content-aware redirection method, a retargeting method with seam-carving combining saliency and depth information is proposed. Specifically, the GBVS algorithm is used to obtain the saliency information, and the image gradient information is combined with the image depth information acquired by the SIFT matching method to construct a more accurate importance map. Then, according to the energy distribution of the importance map, the original image is processed to obtain the final results. The comparison on the public database with different methods under two evaluation criteria shows that the proposed method can preserve the significant part of the image to the greatest extent.

Key words: saliency information; depth information; seam carving; image retargeting

1 引言

随着多媒体技术的不断发展, 不同尺寸的显示设备数目及种类急剧增长, 各种图像信息与电子设备屏幕的直接匹配是个亟需解决的问题. 将同一幅图像显示在具有不同尺寸和分辨率的不同设备上, 即图像重定向问题, 是目前数字图像处理领域中的重要研究内容. 传统的方法有很多, 但均未取得理想效果. 均匀缩放^[1] SCL (Scaling) 只适用于重定向比例一致の場合, 比例不一致时便会发生严重扭曲; 变形^[2] WARP (Warping) 方法较为理想, 但也容易发生扭曲; 变形裁剪^[1] CR (Cropping) 虽然没有使图像发生变形, 但在重定向的时候丢失了大量重要信息, 因而不能满足实际需求. 为

了弥补传统方法的不足, Avidan^[3] 等人首先提出了基于内容感知的缝切割图像重定向方法, 使用梯度图作为重要性图, 通过增加或者移除具有最小能量的八连通缝的方式来达到重定向的目的, 并将该方法用于视频中^[4]. 但是该方法并没有考虑到增加或移除缝的像素之间的差异对视觉效果上的影响, 所以仍然具有一定的局限性; 为了解决这个问题, Pritch^[5] 等使用图割的方法对其进行改进, 定义了前向能量函数.

长期发展以来, 基于内容感知的重定向方法大致分为离散、连续以及两者结合的三类方法. 离散的方法主要以 Avidan^[3] 等人提出的 SC 算法为代表, 该方法以梯度图作为重要性图进行重定向, 由于梯度信息对目标边缘的能量变化比较敏感, 重定向后的图像中的重

收稿日期: 2018-08-12; 修回日期: 2018-12-28; 责任编辑: 李勇锋

基金项目: 国家自然科学基金 (No. 61602150); 江苏省普通高校自然科学基金 (No. 16KJB520037); 南通大学-南通智能信息技术联合研究中心 (No. KFKT2017A02)

要部分还是存在一定程度的变形. 显著信息能够感知图像中的重要部分, 因而被广泛应用到基于内容感知的重定向方法中. 例如, Radhakrishna^[6]等使用图像显著性来计算像素能量进行重定向; Daniel^[7]等人使用显著性信息与梯度信息相结合来得到重要性图应用到流雕刻算法中. 为了进一步提高重定向效果, 最近一些重定向方法试图使用深度信息来改善缝切割算法的能量函数. Shen^[8]等引入了 Kinect 深度相机来得到较为准确的深度信息, 还使用了最小可觉差^[9] (Just Noticeable Difference, JND) 模型来改善缝切割算法的接缝选择, 获得了较好的效果; Shafieyan^[10]等人使用自适应算法估计每个像素的重要性来对梯度、显著性和深度图像分配权重. 连续的方法主要以均匀缩放 (SCL) 算法^[1]和非均匀变形^[2] (WARP) 算法为主, 这类方法根据显著图在所有空间维度上所产生的不同程度的扭曲来减小对重要区域的影响. Wang^[11]等人提出的 Scale-and-Stretch (SNS) 算法, 以梯度值和显著值的乘积作为图像重要性值, 将图像分为重要和非重要区域, 重要区域尽量保持不变或是做长宽等比缩放, 非重要区域则是进行较大的拉伸变形; Philipp^[12]等人提出 Streaming Video (SV) 算法, 则是使用线检测和用户标记线等得到重要性图; Panozzo^[13]等人在轴对称变形空间中对图像进行操作变形, 提出 ASAP (As Similar as Possible) 和 ARAP (As Rigid as Possible) 两个能量函数; Das^[14]等人在上述能量函数的基础上引入了以对象为中心的显著性概念以提高图像重定向的质量. 连续与离散相结合的方法主要是将缝切割、均匀缩放以及手工裁剪等相结合以达到最好的视觉效果的目的. 比如, Rubinstein^[1]等提出的 Multi-operator (MULTIOP) 算法, 将缝切割、均匀缩放与裁剪相结合, 使用双向变形的的方法逐步在三者之间作出选择; Zhang^[15]则是引入了一种结合视觉显著性和梯度幅度生成稀疏图像网格的补丁方法, 优化了三种能量函数以使网格变形并产生重定向图像.

综上所述, 如何构造关于输入图像的能量函数, 是目前基于内容感知重定向方法中共同研究的问题. 本文提出了一种新颖的组合能量函数用于更好地估计图像中像素的重要性从而保留图像的重要部分. 首先, 通过采取深度信息以及视觉显著信息相结合, 改进了重定向算法中的能量函数; 其次, 提出一种不使用深度相机提取深度信息的方法; 最后, 使用多种客观评价标准对本文方法与其他主流重定向方法进行评测.

2 本文方法

2.1 重要度图的计算

对于输入图像 I , 首先通过 GBVS^[16] 算法得到显著结果, 记为 $Sal(I)$. 图像 I 梯度能量函数一般定义为其

在 x 和 y 两个方向上的梯度绝对值之和, 即

$$Gra(I) = e(I) = \left| \frac{\partial I}{\partial x} \right| + \left| \frac{\partial I}{\partial y} \right| \quad (1)$$

将公用数据库 RGBD 分为彩色图库 A 与其对应的深度图库 B, 首先, 提取输入图像 I 以及图库 A 中所有图像的 HOG 特征, 利用 K-近邻算法将所有图像按所提取特征进行分类, 在彩色图库 A 中得到 n_1 个与图像 I 属于同一类的相似图像 I_1, I_2, \dots, I_{n_1} ; 其次, 将输入图像与所得彩色图像进行超像素分割, 利用 SIFT^[17] 算法的图像匹配功能, 得到 n_2 个与输入图像最为相似的超像素区域 R_1, R_2, \dots, R_{n_2} . 最后将上述过程所得的彩色图像区域坐标与其对应深度图像进行对应, 提取深度图库 B 中各图像的子区域 r_1, r_2, \dots, r_{n_2} , 各个子区域合成再优化后得到关于图像 I 的深度图像 $Dep(I)$, 即

$$Dep(I) = f(r_1, r_2, \dots, r_{n_2}) \quad (2)$$

式(2)中, $f(\cdot)$ 为其融合功能函数.

将显著图 $Sal(I)$ 、梯度图 $Gra(I)$ 以及深度图 $Dep(I)$ 进行线性融合得到重要度图 I_{Map} .

$$I_{Map} = \alpha Sal(I) + \beta Gra(I) + (1 - \alpha - \beta) Dep(I) \quad (3)$$

2.2 基于重要度图的缝切割算法

对于 $m \times n$ 的输入图像 I , 缝切割算法中其垂直方向的缝一般定义一条连接图像顶部到底部的路径, 即

$$s^x = \{s_i^x\}_{i=1}^m = \left\{ \begin{array}{l} (i, f_x(i))_{i=1}^m \\ |\forall i, |f_x(i) - f_x(i-1)| \leq \delta \end{array} \right\} \quad (4)$$

式(4)中, $f_x(i)$ 为 $[1, m]$ 到 $[1, n]$ 的一个映射, $(i, f_x(i))$ 为切割缝 s^x 中某个像素的坐标, 下标 x 代表 x 轴方向, 当 $\delta=1$ 时, 可以得到一个八连通的缝. 同理, 水平方向的缝定义为

$$s^y = \{s_j^y\}_{j=1}^n = \left\{ \begin{array}{l} (f_y(j), j)_{j=1}^n \\ |\forall j, |f_y(j) - f_y(j-1)| \leq \delta \end{array} \right\} \quad (5)$$

同样的, $f_y(j)$ 为 $[1, m]$ 到 $[1, n]$ 的一个映射, $(f_y(j), j)$ 为切割缝 s^y 中某个像素的坐标.

计算 2.1 中所得的重要度图在 x 和 y 两个方向上的梯度, 得其能量函数 $e(I_{Map})$, 则其一条垂直切割缝上的能量和为

$$E(s^x) = E(I_{Map}^{s^x}) = \sum_{i=1}^m e(I_{Map}(s_i^x)) \quad (6)$$

因而, 最佳裁剪线为所有八连通裁剪线中能量最小的裁剪线, 即

$$s_{best} = \min_{s^x} E(s^x) = \min_{s^x} \sum_{i=1}^m e(I_{Map}(s_i^x)) \quad (7)$$

设矩阵 M 存储垂直缝切割线上每一个点 (x, y) 的累积最小能量值, 矩阵 N 存储当前能量函数中的能量

值,使用动态规划的方法来确定最佳裁剪线,即从图像第 2 行遍历到最后一行,得到最后一行中的最小值

$$M[i][j] = N[i][j] + \min(M[i-1][j-1], M[i-1][j], M[i-1][j+1]) \quad (8)$$

M 矩阵中最后一行中的最小值即为最优缝合线中的能量最低值,在所得最小能量值的像素点 8 邻域中继续选取最小值进行逆推,得到最终重定向结果;同理,遍历图像第 2 列到最后一列获得的最小值得到水平切割线. 根据重定向尺寸要求,通过不断增加或减少所得的水平和垂直切割线达到重定向的目的.

3 实验与分析

3.1 数据库及评价标准

本文在公开数据库 MIT RetargetMe 上进行实验,该数据库是由 Michael Rubinstein^[18] 等创建的,共有 80 幅图像,是目前重定向领域最广泛使用的数据库. 将本文方法与 SCL^[1]、WARP^[2]、CR^[1]、SC^[3]、SM^[5]、SNS^[12]、SV^[13]、AA^[15]、MULTIOP^[1] 等算法进行比较,使用开源代码得到算法的运行结果.

为了检测重定向方法性能,我们采用 SIFT Flow 和重要性图双向相似 BIMS (Bi-directional Importance Map Similarity) 两个客观评价标准对其进行评价,SIFT Flow 方法通过 SIFT 描述符来捕捉图像中保持不变的局部信息,定义图像的距离即为能量本身,能量最低的匹配即为最佳匹配,即能量越低代表两幅图像越相似. BIMS^[19] 方法使用重要性图的双向相似性从重定向过程中提取特征,并在融合过程中进行双向质量评估,对于输入图像 I 和重定向后的图像 I' ,其基本表达式表示如下:

$$BIMS(I, I') = F(\alpha_l, \alpha_r, \rho, \beta) \quad (9)$$

其中, α_l 和 α_r 为双向质量得分,主要是对重要内容的损失率以及视觉效果的评价; ρ 为重定向质量得分,即重定向内容占原图的百分比; β 为关键点匹配得分,用于测量图像与基数之间的关键点匹配程度; $F(\cdot)$ 为其融合功能函数. 对于所得到的 $\alpha_l, \alpha_r, \rho, \beta$ 四类不同参数,该方法采用平均 (AVG)、乘法 (MULT)、线性回归 (LIN) 和非线性回归 (NLIN) 四种融合策略对重定向结果进行分析.

3.2 定量对比

本文算法与其他算法 SIFT Flow 标准下评价数据值如图 1 所示,本文算法在 SIFT Flow 标准下具有最低的能量值,说明本文算法能够在移除最低能量线的基础上,最大程度的与原图结构保持一致性.

本文算法与其他算法 BIMS 标准下评价数据值如表 1 所示,排名前两位的数值分别加粗和倾斜,本文算法在对四类参数的融合上,均取得了较好的效果,其中,在 LIN 和 NLIN 标准下,取得了最好的效果,说明本文



图1 SIFT Flow标准下本文算法与各算法评价数据值

算法最大程度的保留了图像的重要信息. 在 AVG 以及 MULT 标准下,略低于 CR 算法,这是因为 CR 算法仅通过手动裁剪方式进行重定向,并未对图像重要性图作改变.

表 1 BIMS 标准下本文算法与其他算法

算法	AVG	LIN	MULT	NLIN
SCL	0.6801	54.3444	0.1663	53.9582
WARP	0.7360	55.6308	0.2739	55.9089
CR	0.8067	57.8193	0.4577	58.7272
SC	0.7610	55.5703	0.3386	53.8669
SM	0.7764	57.2619	0.3596	56.8358
SNS	0.7385	56.0634	0.2570	56.4500
SV	0.7369	55.7858	0.2489	55.2480
MUL	0.7385	55.8667	0.2595	55.9013
本文	0.7970	60.2672	0.3653	60.7919

3.3 视觉效果对比

本文算法与其他算法的视觉效果如图 2 所示,各个图像变形失真和缺失区域已在图中用红框标出,图 2 (b) 中,最左边和最右边女生的腿都出现变形;由图 2 (c) 可以看出,在 CR 算法图中,近一半区域被裁剪去除,丢失了原图大部分边缘信息;由图 2 (d) 所示,SC 算法变形区域较为明显;SM 算法使用图割对 SC 进行改进,由图 2 (e) 可见,该算法重定向后,图像部分区域信息丢失;由图 2 (f) (g) 可以看出,SNS 和 SV 算法变形并不明显,但是图像主体区域有些缩小;AA 算法在主体区域并未出现变形和失真,但是非重要区域变化较大;如图 2 (i) 所示,在对图像进行重定向时,本文算法并未出现明显的变形和失真,也未丢失图像中重要区域和非重要区域,获得了较好的重定向结果.

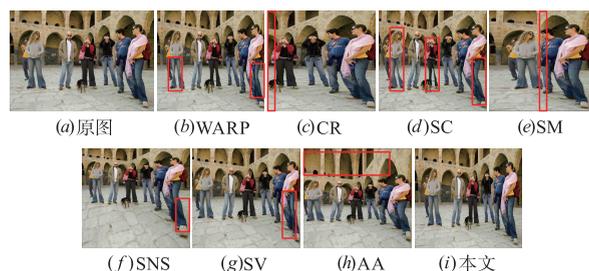


图2 图像Umdan各算法重定向视觉效果图

4 总结

本文提出一种通过 SIFT 匹配方法获取图像深度信息的方法,将图像的显著信息和深度信息应用到缝切割算法的能量函数中,改进了重定向算法中的能量函数,进行图像重定向.在公开的重定向数据库上进行了非常详实的实验,实验结果表明,在两个不同标准下与多种重定向方法的对比中表现出明显的优越性,而且能够更好的保留图像的重要部分,提高图像重定向的质量,改善当前内容感知重定向方法中的扭曲及失真问题.

参考文献

- [1] Rubinstein M, Shamir A, Avidan S. Multi-operator media retargeting [J]. *ACM Transactions on Graphics*, 2009, 28 (3): 341 – 352.
- [2] Wolf L, Guttman M, Cohenor D. Non-homogeneous content-driven video-retargeting [A]. *Proceedings of the 11th International Conference on Computer Vision [C]*. Rio de Janeiro, Brazil: IEEE, 2007. 1 – 6.
- [3] Avidan S, Shamir A. Seam carving for content-aware image resizing [J]. *ACM Transactions on Graphics*, 2007, 26 (3): 1 – 10. DOI:10.1145/1276377.1276390.
- [4] Rubinstein M, Shamir A, Avidan S. Improved seam carving for video retargeting [J]. *ACM Transactions on Graphics*, 2008, 27 (3): 1 – 10. DOI:10.1145/1360612.1360615.
- [5] Pritch Y, Kav-Venaki E, Peleg S. Shift-map image editing [A]. *Proceedings of the 12th International Conference on Computer Vision [C]*. Kyoto, Japan: IEEE, 2010. 151 – 158.
- [6] Radhakrishna A, Sabine S. Saliency detection for content-aware image resizing [A]. *Proceedings of the 16th International Conference on Image Processing [C]*. Cairo, Egypt: IEEE, 2009. 1005 – 1008.
- [7] Daniel D, Alexandre A, Pierre V. Stream carving: an adaptive seam carving algorithm [A]. *Proceedings of the 17th International Conference on Image Processing [C]*. Hong Kong, China: IEEE, 2010. 901 – 904.
- [8] Shen J, Wang D, Li X. Depth-aware image seam carving [J]. *IEEE Transactions on Cybernetics*, 2013, 43 (5): 1453 – 1461.
- [9] Nguyen P B, Beghdadi A, Luong M. Perceptual watermarking using a new just-noticeable-difference model [J]. *Signal Processing Image Communication*, 2013, 28 (10): 1506 – 1525.
- [10] Shafieyan F, Karimi N, Mirmahboub B, et al. Image retargeting using depth assisted saliency map [J]. *Signal Processing Image Communication*, 2017, 50 (C): 34 – 43.
- [11] Wang Y S, Tai C L, Sorkine O, et al. Optimized scale-and-stretch for image resizing [J]. *ACM Transactions on Graphics*, 2008, 27 (5): 1 – 8.
- [12] Philipp K, Manuel L, Alexander H, et al. A system for retargeting of streaming video [J]. *ACM Transactions on Graphics*, 2009, 28 (5): 1 – 10. DOI:10.1145/1618452.1618472.
- [13] Panozzo D, Weber O, Sorkine O. Robust image retargeting via axis-aligned deformation [J]. *Computer Graphics Forum*, 2012, 31 (2pt1): 229 – 236.
- [14] Das B C, Gopalakrishnan V, Iyer K N, et al. Similarity and rigidity preserving image retargeting [A]. *Proceedings of the International Conference on Image Processing [C]*. Phoenix, AZ, USA: IEEE, 2016. 1584 – 1588.
- [15] Zhang Y, Ngan K N. Fast patch-wise image retargeting [A]. *Proceedings of the International Conference on Image Processing [C]*. Phoenix, AZ, USA: IEEE, 2016. 1813 – 1817.
- [16] Schölkopf B, Platt J, Hofmann T. Graph-based visual saliency [J]. *Proc of Neural Information Processing Systems*, 2006, 19: 545 – 552.
- [17] Liu C, Yuen J, Torralba A, et al. SIFT flow: dense correspondence across different scenes [A]. *Proceedings of the 10th European Conference on Computer Vision [C]*. Marseille, France: Springer Verlag, 2008. 28 – 42.
- [18] Rubinstein M, et al. A comparative study of image retargeting [J]. *ACM Transactions on Graphics*, 2010, 29 (6): 160.
- [19] Oliveira S A F, Alves S S A, Gomes J P P, et al. A bi-directional evaluation-based approach for image retargeting quality assessment [J]. *Computer Vision and Image Understanding*, 2018, 168: 172 – 181.

作者简介



吴加莹 女. 1994 年出生, 江苏连云港人. 现为南通大学电气工程学院控制科学与工程系研究生. 主要研究方向为机器视觉与模式识别.
E-mail: 172251887@qq.com



杨赛(通信作者) 女. 1981 年出生, 山东滨州人. 2015 年获南京理工大学博士学位, 现为南通大学讲师, 研究方向为计算机视觉与机器学习.
E-mail: yangsai166@126.com