



计算机科学

COMPUTER SCIENCE

基于网格与超像素的图像重定向方法

陈美颖, 毕秀丽, 刘波

引用本文

陈美颖, 毕秀丽, 刘波. 基于网格与超像素的图像重定向方法[J]. 计算机科学, 2023, 50(11A): 221100153-8.

CHEN Meiyong, BI Xiuli, LIU Bo. [Image Retargeting Method Based on Grids and Superpixels](#)[J]. Computer Science, 2023, 50(11A): 221100153-8.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于改进的YOLOv5s列车轨道线检测方法](#)

Railway Track Detection Method Based on Improved YOLOv5s

计算机科学, 2023, 50(11A): 230200101-6. <https://doi.org/10.11896/jsjcx.230200101>

[基于深度学习的钢轨光带检测算法](#)

Rail Light Band Detection Algorithm Based on Deep Learning

计算机科学, 2023, 50(11A): 230200146-6. <https://doi.org/10.11896/jsjcx.230200146>

[基于多编码器的多模态MRI脑肿瘤分割](#)

Multimodal MRI Brain Tumor Segmentation Based on Multi-encoder Architecture

计算机科学, 2023, 50(6A): 220200108-6. <https://doi.org/10.11896/jsjcx.220200108>

[基于球簇聚类的超像素分割迭代算法](#)

Superpixel Segmentation Iterative Algorithm Based on Ball-k-means Clustering

计算机科学, 2023, 50(6A): 220600114-7. <https://doi.org/10.11896/jsjcx.220600114>

[结构化混合注意力网络的图像超分辨率重建](#)

Image Super-resolution Reconstruction Based on Structured Fusion Attention Network

计算机科学, 2023, 50(6A): 220600240-5. <https://doi.org/10.11896/jsjcx.220600240>

基于网格与超像素的图像重定向方法

陈美颖 毕秀丽 刘波

重庆邮电大学图像认知重庆市重点实验室 重庆 400065

(chenmy0712@163.com)

摘要 图像是人与人之间进行交流的重要媒介,在信息高速发展的今天,利用图像重定向技术使图像能满足各式各样的设备尺寸具有重要意义。基于网格的图像重定向算法首先对输入图像生成对应的规则矩形网格,然后根据该网格内的图像内容来评估图像像素权重以此决定此网格的变形程度,对图像全局进行不断迭代直到图像重定向终止条件。此类算法仍存在对图像内容评估不全面的问题,进而导致输出图像结构扭曲、难以保持结果图像的对角线特征以及整体结构等问题。针对以上问题,提出了一种基于超像素、梯度以及显著性的图像重定向方法。首先利用超像素方法对输入图像做预处理,然后用超像素块作为后续处理单位,随后利用基于梯度和显著性的图像像素权重评估方法对超像素处理输出图像进行权重度量,输出一幅图像重定向权重热力图,最后根据此重定向权重热力图对网格进行迭代优化,实现对图像的重定向处理。实验结果表明,相比对比方法所提方法在6种无参考图像质量评估指标上都有一定优越性,在语义合理性、信息准确性和视觉自然性上都具有一定优势,在图像重定向领域有较大的应用价值。

关键词: 图像处理;图像重定向;图像显著性检测;超像素

中图分类号 TP751

Image Retargeting Method Based on Grids and Superpixels

CHEN Meiyang, BI Xiuli and LIU Bo

Chongqing Key Laboratory of Image Cognition, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

Abstract Image is an important medium for communication between people. With the rapid development of information today, it is of great significance using image retargeting technology to make images adapt to a variety of device sizes. The grid-based image retargeting algorithm first generates a regular rectangular grid corresponding to the input image, and then determines the deformation degree of the grid by evaluating the weight of image pixels according to the image content in the grid. The global iteration of the image is carried out until the termination condition of image retargeting. However, such algorithms still have the problem of incomplete evaluation of image content, which leads to the distortion of the output image structure, and it is difficult to maintain the diagonal features and overall structure of the result image. In order to solve the above problems, this paper proposes an image retargeting method based on superpixels, gradients and saliency. Firstly, the input image is preprocessed by the superpixel method, and then the superpixel block is used as the subsequent processing unit, and the image pixel weight evaluation method based on gradient and saliency is used to measure the weight of the superpixel output image, and an image retargeting weight heat map is output. Finally, the grid is iteratively optimized according to the retargeting weight heat map and realize the retargeting of the image. Experimental results show that the proposed method has certain advantages in the six no-reference image quality assessment indicators, and has certain advantages in semantic rationality, information accuracy and visual naturalness, and has great application value in the field of image retargeting.

Keywords Image processing, Image retargeting, Image significance detection, Superpixels

1 引言

在信息的高速发展下,图像技术也随之不断发展,数字图像已经成为媒体不可分割的一部分。人们通常会使用不同尺寸的显示设备浏览图像,比如手机、笔记本电脑、投影仪等,因此图片需要能够自动适应不同的尺寸以及分辨率来满足不同需求。

现有的图像重定向方法根据在重定位过程中是否感知图像内容信息可以分为两大类。第一类为传统方法,这类方法通过直接裁剪或拉伸使图像达到目标大小,主要包括手动

裁剪(manual Cropping, CR)、单一方向的规则缩放(uniform Scaling in one dimension, SCL)等。这些方法由于没有考虑图像内容信息,虽然能够得到目标大小图像,但结果图像通常存在过度拉伸和明显扭曲的问题。第二类方法是根据图像内容信息进行图像像素评估度量,对不同内容进行不同程度地放缩,即尽量保持图像中重要内容的原始长宽比例,而对图像中相对不重要的内容进行较大程度的放缩,使其长宽比例的变化更大,主要包括接缝雕刻(Seam Carving, SC)^[1], Shift-Map(SM)^[2], Non-homogeneous Warping (WARP)^[3], Streaming Video (SV)^[4]和 SNS(Optimized Scale-and-Stretch)^[5]

等。第二类方法又可以根据对图像处理方式的不同细分为两类:离散方法和连续方法。离散方法在图像处理过程中将图像看作离散的像素点,这种方式在缩放过程中不能保证相邻像素的连续性,易造成图像内容和结构的局部扭曲。在前文提到的方法中,SC,SM属于离散方法,WARP,SV,SNS属于连续方法。其中SNS是一种基于网格的连续方法,该方法首先将图像划分为规则方形网格,其次根据变形约束函数计算每个网格的形变程度,此形变程度与网格内容的图像像素权重以及网格线弯曲程度有关,随后通过优化求解约束函数以得到最终的目标网格,最后根据最终网格与图像之间的映射关系对输入图像进行缩放,得到最终结果图像。此方法由于没有丢失图像原始内容,能够很好地保持源图像中的水平和竖直的特征,因此备受青睐。

人类在面对各种复杂的图像时,会无意识地精准识别图像中的特殊区域,这个被识别出的特殊区域常被称为显著性区域,而人们所拥有的这种能力被称为视力注意机制。为了模拟人们的这种能力,图像显著性检测技术应运而生,它凭借捕捉图像特征对图像信息进行图像像素权重评估,进而达到模拟人类的视力注意机制的目的。目前,常用的图像显著性检测方法有ITTI^[6],AC^[7],FT^[8],LC^[9],HC^[10]和RC^[10]。频率调谐显著区域检测(Frequency-tuned salient region detection, FT)^[8]利用图像在频率域中不同频段的信息,高频部分代表图像中的纹理等细节特征信息,低频部分代表图像大致轮廓等整体特征信息,将对图像显著性的分析过程从空间域转化到频率域,使用图像的颜色和亮度特征估计中心周围的对比度,输出结果具有均匀的显著性区域、明确的边界,与输入图像分辨率相同以及较高的计算效率等特点。

本文针对SNS方法存在的问题提出了一种基于超像素、梯度和视觉显著度图共同形成的度量图像内容图像像素权重的图像重定向方法,实验结果表明,其输出图像结构纹理清晰、连续均匀、符合审美要求,并在图像客观评价指标上相比SNS,SC等图像重定向方法都有一定的提升。

本文第2章详细介绍本文所提出的图像重定向方法流程;第3章设计了3组实验,进行消融实验以及对比实验,对比分析本文方法和主流图像重定向方法在相同的3个图像数据集下的效果,以验证本文方法的性能;最后总结全文。

2 算法概述

现有的传统重定向算法普遍存在无法完整保留图像的对角线特征的问题,因此图像整体结构会受到影响,即图像在视觉上存在明显可察觉的扭曲现象。除此之外,针对基于网格的重定向方法,当划分的网格中同时包含显著区域和背景区域时,这类网格极易出现严重扭曲的现象。基于网格的图像重定向经典算法SNS使用了显著性图ITTI^[6],该显著性方法严格地模仿了生物结构模型,计算复杂度较高,且输出图像的分辨率会降低,导致目标轮廓无法被清晰辨识。更值得注意的是,此方法对图像显著值的评估是基于图像中的局部对比度,往往会产生显著对象不均匀的现象,即显著对象的边缘有较高的显著值,而其内部的显著值往往较低,这种现象在输入图像的显著对象特征较为单一的情况下尤为明显。为了解决上述问题,本文引入超像素以及FT显著性算法形成一种新的图像重定向算法。

本文提出的图像重定向方法的总流程如图1所示,其

主要包括图像重定向权重热力图生成和网格迭代优化两个部分,具体过程如下:

1)利用超像素对输入图像进行处理,生成由超像素块组成的图像,将后续操作的处理单位由单个像素点转换成超像素块,随后分别计算图像梯度图和图像显著性图,最后对图像梯度图和显著性图进行加权作为重定向权重热力图。

2)根据由重定向权重热力图 M 和由网格形变程度决定的网格能量函数对初始网格进行迭代优化,最终得到最优网格,随后根据网格与像素之间的映射关系得到输出图像。

2.1 重定向图像权重热力图获取

本文方法将图像 I 定义为 $I = (V, E, F)$ 其中 V 为顶点集、 E 为边集、 F 为网格面集,由 V, E 形成的垂直和水平的线将图像分成网格,由此将图像大小从 $m \times n$ 缩放为 $m' \times n'$ 的任务转化成由最初顶点集 $V = [v_0^T, v_1^T, \dots, v_{\text{end}}^T]$ 通过全局优化得到最终顶点集合 $V = [v_0'^T, v_1'^T, \dots, v_{\text{end}}'^T]$ 。

计算重定向权重热力图的输入图像数据格式是多样的,若将一张高分辨率图像作为输入数据,使处理过程会过于复杂。为解决此问题,本文引用超像素方法对原始输入图像进行像素数量级的降级处理,即在计算图像像素权重前先进行超像素处理,这样可以在超像素块基础上做后续对图像权重的计算。

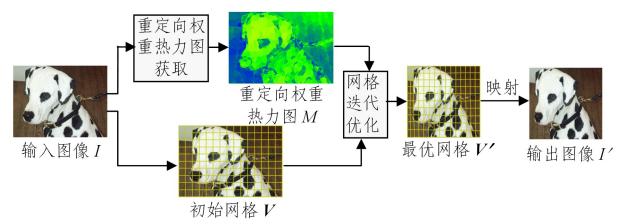


图1 算法总流程图。

Fig. 1 Overall flow chart of algorithm

超像素指利用图像相邻像素的纹理、颜色、亮度等特征进行分组,并构成的有一定视觉意义的不规则像素块,从而将一幅像素级的图像转换为区域级的图像,即用少量的超像素代替大量的像素来表达图片特征,很大程度上降低了图像后处理的复杂度,极大提高了计算的效率。简单的线性迭代聚类(Simple Linear Iterative Clustering, SLIC)^[11]是超像素算法中在算法效率、内存使用以及直观性上都表现得较好的一种经典算法。该算法基于 k -means^[12]聚类算法,将图像从RGB颜色空间转换至CIE-Lab颜色空间。将每个像素的 (L, a, b) 颜色值及其 (i, j) 坐标合并组成一个5维向量 (L, a, b, i, j) ,然后对这5维特征向量构造距离度量标准,在局部对图像像素进行聚类。由相邻像素对应的向量之间的距离决定它们之间的相似性,距离越大,相似性越小。

由输入图像生成图像重定向权重热力图的主要流程如图2所示,描述性过程为:首先对输入图像 I 利用SLIC超像素方法进行预处理,定义一个超像素块 P 内每一个像素点 (i, j) 的RGB空间坐标为 (R_p, G_p, B_p) ,其数学计算式如下:

$$R_p = \frac{\sum_{(i,j) \in P} r_{(i,j)}}{m} \quad (1)$$

$$G_p = \frac{\sum_{(i,j) \in P} g_{(i,j)}}{m} \quad (2)$$

$$B_p = \frac{\sum_{(i,j) \in P} b_{(i,j)}}{m} \quad (3)$$

其中, m 为超像素块 P 中的像素的总数。输入图像在经过超像素处理后得到图像 I_{sup} , 图像中的特征更明显, 并且可以降低后续操作的复杂度, 提高计算效率。

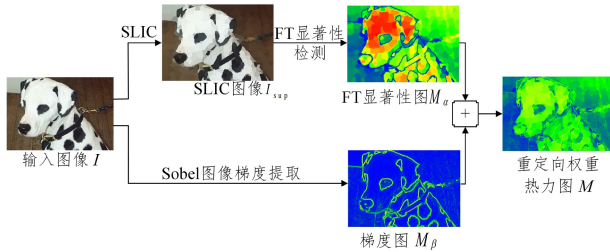


图2 重定向权重热力图获取流程图

Fig. 2 Flow chart of image retargeting weight heat map acquisition

本文所提出的图像重定向权重热力图 M 主要包括两部分, 分别为图像的梯度图 M_α 和图像显著图 M_β , 总的价值块图为两项的加权和。

$$M = \lambda_1 M_\alpha + \lambda_2 M_\beta \quad (4)$$

其中, λ_1 和 λ_2 为权重系数, 根据经验, λ_1 与 λ_2 均取 0.5 时实验效果较好, 图像的梯度图 M_α 的计算式如式(5)所示:

$$M_\alpha = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (5)$$

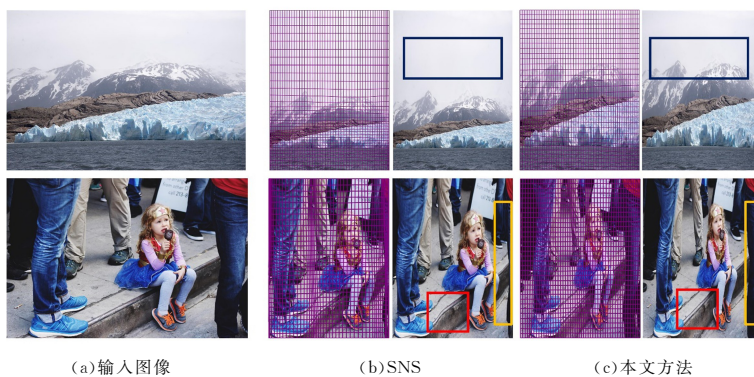
本文选择 FT 显著性探测方法来计算图像显著图 M_β , 考虑到显著性区域检测常用的是反映图像整体信息, 如物体的轮廓和基本的组成区域的低频信息, 首先使用高斯平滑实现对图像高频信息的舍去, 随后转换图像颜色空间, 由 RGB 颜色空间转换到 CIE-Lab 颜色空间, 先计算整幅图像的 Lab 颜色特征均值 $(\bar{L}, \bar{a}, \bar{b})$, 再计算每个像素 (L, a, b) 值与图像 $(\bar{L}, \bar{a}, \bar{b})$ 均值之间的欧氏距离, 最后进行归一化得到最终的显著图。 M_β 的计算式如式(6)所示:

$$M_\beta = \| I_u - I_{uhc}(i, j) \| \quad (6)$$

其中, I_u 为整幅图像的 Lab 颜色平均特征 $(\bar{L}, \bar{a}, \bar{b})$, I_{uhc} 为图像经过高斯平滑后像素 (i, j) 的 Lab 颜色特征, $\| \cdot \|$ 为 L2 范式。

对输入图像进行超像素处理可以简化图像处理的运算级, 将原先的像素级处理转变成像素集合级处理, 即将单个像素点作为人们关注的重点, 因为单个孤立的像素点并不能给人们带来有语义的信息, 因此, 只有将多个像素聚集成一片连续区域时, 这些像素才具有具体的意义, 并且这样的处理可以降低后续计算图像梯度与显著性过程处理单元的数量级, 以降低处理任务的复杂度。图像重定向权重热力图 M 的两部分都是必不可少的, 图像梯度代表了图像像素的变化速度, 它能获取图像中的边缘信息, 即图像中的物体轮廓边缘灰度变化明显且层次感强的信息, 图像梯度图 M_α 的存在可以尽可能地保留图像中结构信息, 保留图像的整体结构; 而图像显著图 M_β 可以确定图像中因自身的独特性质与周围环境形成对比的、吸引人们注意力的物体, 满足人们对图像美学的要求。

从图 3 第一行图像中可以观察到, SNS 方法的输出图像顶部存在较大区域的空白, 如图 3(b) 中蓝框所示, 而相比 SNS 方法, 本文方法图像各部分的比例与源图像更为一致, 如图 3(c) 中蓝框所示。观察并对比两种方法的最终网格可以发现, 本文方法相较于 SNS 方法的网格更加规则, 特别是图像边缘的网格, 从图 3(b) 和图 3(c) 中红色框出的具有水平结构的阶梯可以观察到, 此结构在 SNS 方法中存在明显扭曲, 而在本文方法的输出图像中阶梯的水平结构得到了完整的保留, 从橙色框出的图像内容部分可以观察到, 在 SNS 方法中人的腿在网格变形中被舍弃了, 而此部分信息在本文方法中也被完整地保留了下来, 满足了重定向技术输入图像与输出图像之间保留突出特征的基本要求。对比图 3(b) 和图 3(c) 中 SNS 方法与本文方法的最终网格可以发现, 本文方法的网格比 SNS 的网格更加均匀, 不存在集中缩放某一网格的现象, 这是由于加入了超像素预处理, 使图像中重要区域更加突出, 在网格迭代过程中, 降低了同一网格同时包含重要区域的可能性, 这样一来, 网格根据权重进行放缩的过程更加有效可靠, 图像主体部分不会产生较大形变, 满足审美原则。



注: 图(b)与图(c)分别为 SNS 方法与本文方法的最终网格(左图)和输出图像(右图)。

图3 SNS方法与本文方法的对比(电子版为彩图)

Fig. 3 Comparison between SNS and the proposed method

2.2 基于图像权重热力图的图像重定向

在获得图像重定向权重热力图 M 后, 就可以在其基础上对初始规则网格进行非均匀放缩, 建立源图像到目标图像的最优映射。

为了更好地控制每个网格的重定向程度, 本文定义了一个二次函数能量函数, 通过此公式将图像重定向问题转化为

一个二次最小化问题, 求解此最小化问题获得最优解, 从而由源图像划分的初始规则网格得到目标最优网格的位置, 然后通过映射关系得到最终目标图像。二次函数能量函数 D 的定义如式(7)所示:

$$D = D_u + D_l \quad (7)$$

其中, D_u 是与图像像素内容权重有关的形变分量, D_l 是控制

网格线弯曲程度的分量。

图像内容权重相关形变分量 D_u 的存在可以更好地根据图像内容对图像进行重定向,即权重越低的网格形变程度将更显著, D_u 由图中每个网格的形变分量 $D_u(f)$ 加权和组成,计算式如式(8)所示:

$$D_u = \sum_{f \in F} \omega_f D_u(f) \quad (8)$$

其中, F 代表网格集合,图像像素权重因子 ω_f 为某一网格 f 内像素权重的均值。 $D_u(f)$ 的计算式如式(9)所示:

$$D_u(f) = \sum_{(i,j) \in E(f)} \| (v_i' - v_j') - s_f (v_i - v_j) \|^2 \quad (9)$$

其中, E_f 表示 f 网格的边的集合, s_f 是一个比例因子,对于四边形的每个顶点 v 和变形后顶点 v' ,有 $v' = s_f * v + t$ (其中 t 是一个恒定的平移向量),如果原始网格顶点 v 和变形网格顶点 v' 是固定的,则可以通过对式(9)求微分得到 s_f ,如式(10)所示:

$$\frac{\partial D_u(f)}{\partial s_f} = 0 \quad (10)$$

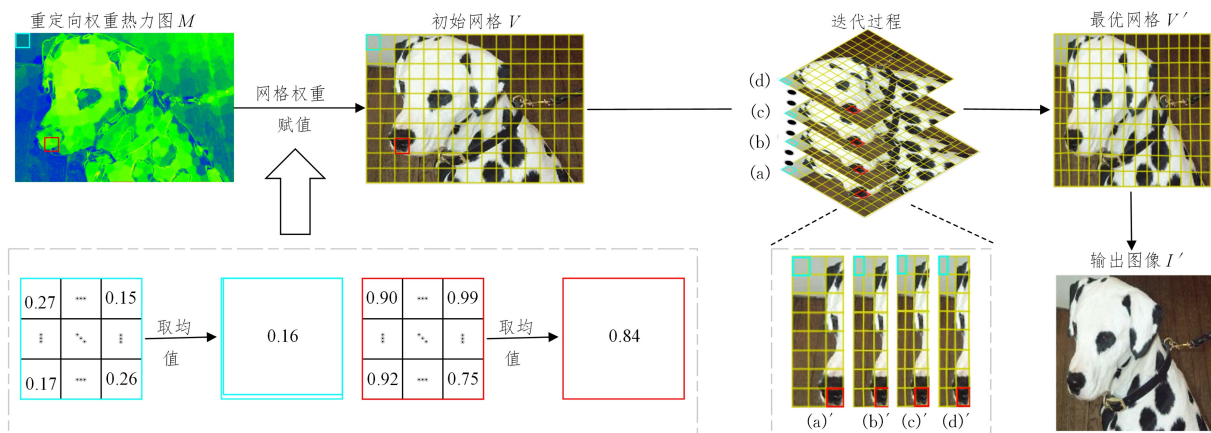
对式(10)求解可知 $D_u(f)$ 的比例因子 s_f 仅与原始网格顶点 v 和变形网格顶点 v' 有关, s_f 的计算式如式(11)所示:

$$s_f = \frac{\sum_{(i,j) \in E(f)} (v_i - v_j)^T (v_i' - v_j')}{\sum_{(i,j) \in E(f)} \| v_i - v_j \|^2} \quad (11)$$

在对图像划分网格的过程中,权重较高的内容常常跨越多个网络,这种现象会使网格在优化过程中失去其规则结构。为了减小此影响,本文定义了网格线弯曲程度分量 D_l ,通过此分量来控制网格变形过程中网格线的弯曲程度,其计算式如式(12)所示:

$$D_l = \sum_{(i,j) \in E} \| (v_i' - v_j') - l_{ij} (v_i - v_j) \|^2 \quad (12)$$

其中, $l_{ij} = \| v_i' - v_j' \| / \| v_i - v_j \|$ 是形变前后边长度比例。



注:(a),(d)分别是网格迭代过程的初态、两个中间态以及终态;(a'),(d')为与(a),(d)对应的细节放大图,蓝框部分是图像中权重较低的网格,受到了不规则缩放,而红框是权重较高网格块,被成比例地缩放。

图4 网格优化流程图

Fig. 4 Flow chart of grid optimization

通过最小化二次函数能量函数(7)就能获得最优的图像重定向结果,即目标函数可定义为式(13):

$$\min D = \min(D_u + D_l) \quad (13)$$

目标网格的边界限制的网格左上角顶点 $v_0' = (0, 0)^T$ 与右下角顶点 $v_{\text{end}}' = (n', m')^T$,此外同时约束与之对应的同行同列的像素点以保证最后的输出目标网格是矩形,这些像素点的数学表达如下所示:

$$v'_{i,y} = \begin{cases} 0, & v_{i,x} = 0 \\ m', & v_{i,x} = n' \end{cases} \quad (14)$$

$$v'_{i,x} = \begin{cases} 0, & v_{i,y} = 0 \\ n', & v_{i,y} = m' \end{cases}$$

通过求解目标函数(13)的最小值就能得到最优的图像重定向结果,这是一个线性规划的过程,也是一个凸二次规划问题,通过不断迭代可以解决这个非线性问题,求解出局部最优解时,全局解决方案也能随即得出。

从图4中可以观察到,在整个网格优化迭代过程中,网格的矩形结构得到较好地保留,没有发生明显的弯曲与扭曲。从细节放大图中红框部分可以发现,当网格内包含较多同质信息即非重要信息时,网格发生程度较大的扭曲,而对于包含图像中特征信息即重要信息较多的网格,因为这些网格中的内容对表达整幅图像的意义有重大作用,因此得到了规则的缩放,在最大程度上保留了其特征,使图像整体完整,语义

信息得到较完整地表达。

3 实验结果与分析

为了讨论和验证本文提出的图像重定向方法,将其与已有的图像重定向方法进行对比和分析。本章将从3个角度进行实验,首先设计消融实验验证本文所提出新的图像像素权重度量方法的有效性,然后对比本文方法和现有的图像重定向算法。为了确保实验的可信度,实验所用的3个数据库数据图像均来自公用数据库,其中2个为人工分组数据库,1个为完整公用数据库。所有的实验在 Intel (R) Core(TM) i5-9300H CPU@2.40 GHz 的 PC(Matlab R2017a, win11)平台上进行。对比实验中其他重定向方法的参数都设置为复现算法最佳实验结果时的参数值。

3.1 实验数据库及对比方法

实验部分共使用3个数据库,其中2个数据库为在3个图像处理常用数据库 Kodak24, VOC2012, RetargetMe^[13] 中根据是否含有前景主体分为两组图像数据库,在这个过程中除去了过于相似的部分图像,其中 Kodak24 中包含了24张大小均为 500×500 彩色图,包含自然风景、建筑物、人物、动物。VOC2012 是 PASCAL VOC 挑战赛中使用的基准数据,共有17125张尺寸不一的图像,包含20个类别,例如人、常见动物、交通车辆等,横向图的尺寸大约在 500×375 ,纵向图的

尺寸大约在 375×500 ,但长宽均不会超过 512。Retarget-Me^[13]包含了 37 张不同分辨率的彩色图像,包含从 333×240 到 640×640 多种分辨率。为了保证实验的可信度,还使用了一个完整数据库 BDS500^[14]。本文所使用的数据库分别如表 1 所列。

数据集	图像数
有主体(Mo)	635
无主体(N-Mo)	109
BDS500	500

本文采用以下 4 种主流的图像重定向方法与本文提出的方法进行对比。

1) Optimized Scale-and Stretch(SNS)^[5]:一种基于连续网格的内容感知图像重定向方法,通过对网格进行不同程度的形变来调整图像大小。

2) Seam-Carving(SC)^[1]:一种经典的离散型内容感知图像重定向方法,通过增加或删除一条垂直或水平的像素点来达到放缩的目的。

3) Weakly supervised Semantics and Aesthetics aware Multi-operator Image Retargeting(SAMIR)^[15]:一种多算子图像重定向方法,其中多算子包括裁剪、放缩和解封雕刻,将多算子重定位作为马尔可夫决策过程,并应用强化学习(RL)实现全局最优。

4) Summarizing Visual Data Using Bidirectional Similarity(SVDBS)^[16]:一种基于双向相似性的图像重定向方法,利用图像模式中的冗余信息将源图像中的重复块映射到目标图像中的一些代表性块,从而保证它们以原始比例出现,以达到预期结果。

3.2 评价指标

虽然图像质量评价方法已经相对成熟,但是图像重定向质量评价仍处于起步阶段^[17],这是由于图像经重定向技术处理之后,会产生不同尺度大小的图像,使原始图像与结果图像的像素之间失去一一对应的关系,无法直接使用适用于同尺度的质量评价方法。

为了客观评价本文方法与 SNS、SC 等主流方法在图像重定向上的效果,本文在相同的实验条件下,采用以下几种无参考类型客观评价指标来评价输出图像的质量。

1) 基于自然场景统计的两步无参考图像质量评估框架(简称 BIQI)^[18]:通过小波变换提取图像的统计特征,再利用分类算法(SVM)求出失真概率以判断失真类型,然后评估质量,最后将所有失真类型的概率与相对应失真质量分数进行加权求和得到最终分数。

2) DCT 域中的自然场景统计方法(简称 BLINDS-ID)^[19]:利用朴素贝叶斯模型来预测图像中基于 DCT 系数的 NSS 模型特征的图像质量分数,与人类对质量的判断高度相关。

3) 基于空间光谱熵的图像评估指标(Spatial-Spectral Entropy-based Quality,简称 SSEQ)^[20]:利用图像的局部空间和光谱熵特征,实现跨多种失真类别评估失真图像的质量。

4) 基于可感知模糊概念的无参考客观图像清晰度度量

(简称 JNB)^[21]:将可感知模糊集成到概率求和模型中,能够高精度地预测具有不同内容的图像中的相对模糊量。

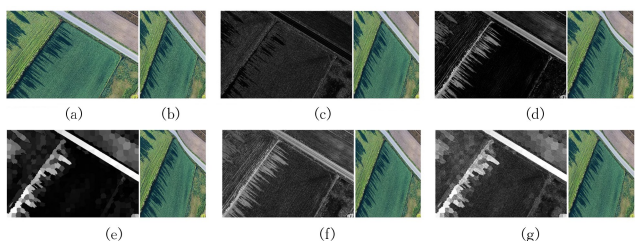
5) 基于模糊检测累积概率的无参考感知图像(No-reference blur assessment based on edge modeling,简称 EM-BM)^[22]:通过结合参数化边缘模型来描述和检测边缘来为每个边缘像素提供宽度和对比度估计,最后通过累积模糊检测的概率来获得模糊度量。

6) 清晰度度量(简称 CPBDM)^[23]:根据人类视觉系统(HVS)对模糊失真的响应来进行评估。该指标的感知意义通过主观实验进行验证。该指标对于具有不同前景和背景感知模糊质量的图像特别敏感。

3.3 消融实验

为验证本文提出的图像像素权重度的有效性,本节设计消融实验分别验证 SLIC 超像素预处理、图像显著性部分和图像梯度部分在图像重定向过程中的作用。

从图 5(c)中可以观察到,当仅采用图像梯度去评价图像内容权重时,虽然图像中的结构信息得以较好的保留但却失去了较多图像内容,如图像左上角处树影部分丢失,而当过分关注于人眼注重的图像内容信息,即只采用了图像显著性作为图像像素权重的评价指标,将会出现图像结构扭曲的现象,如图 5(d)中的道路和水流都出现了一定程度的结构形变,这说明在评价图像像素权重时仅考虑图像内容和图像结构其中某一方面的做法是片面不可取的,都会使图像重定向结果达不到预期结果。从图 5(e)可以观察到,引入超像素处理后,图像中结构的不规则形变有一定程度的改善,没有出现突变的情况,图像像素的连续性得到保证,而当同时使用图像梯度和图像显著性评价图像像素权重时,图像重定向结果能在保持图像结构的同时保留图像大量信息,如图 5(f)所示,在图像对主体内容有完整保留的同时河流和道路没有发生明显扭曲。从图 5(g)可以观察到,对于经过超像素预处理后的图片,再用同时含有梯度和图像显著性的评估方法进行评估图像像素权重,可以使图像重定向结果在保持图像基本结构的同时满足人类的美学欣赏,在视觉感官上保持连贯性。为了更好地验证本文方法的优越性,本节还对比了基于最近邻插值的放缩方法,如图 5(a)所示,虽然此方法从视觉上也能较完整地保留源图像的基本结构与信息,但是其结果存在明显的噪声点,影响图像的平滑性。



注:(a)输入源图像;(b)基于最近邻插值方法的缩小结果;(c)~(g)不同成分组成图像权重图(左)和与之对应的图像缩小结果(右,缩小比例为 0.5);(c)中包含图像梯度;(d)中包含图像显著性;(e)中包含超像素与图像显著性;(f)中包含图像梯度与图像显著性;(g)中包含超像素、图像梯度和图像显著性。

图 5 消融实验对比图

Fig. 5 Comparison of ablation experiments

如表 2 所列,在 3 个数据集上,本文方法在 6 项指标上相较于其他利用各种不同成分组成的图像权重图的重定向方法有一定程度的提升。

表 2 消融实验重定向结果客观评价

Table 2 Objective evaluation values for retargeting results of ablation experiments

	BIQI ↓	BLINDS-II ↓	SSEQ ↓	CPBDM ↑	JNB ↑	EMBM ↑
Mo	(b) 25.071	14.841	16.444	0.670	2.029	0.468
	(c) <u>24.787</u>	14.483	<u>16.191</u>	<u>0.681</u>	2.060	<u>0.469</u>
	(d) 24.661	<u>14.666</u>	<u>16.191</u>	0.683	<u>2.059</u>	<u>0.469</u>
	(e) 24.889	14.893	16.142	0.680	2.056	0.470
	(f) 24.884	14.725	16.222	0.680	2.060	0.470
	N-Mo	(b) 26.585	19.683	13.229	0.721	4.333
(c) 27.040		19.803	12.438	0.739	<u>4.327</u>	<u>0.472</u>
(d) <u>26.742</u>		<u>19.326</u>	12.353	<u>0.741</u>	4.345	<u>0.472</u>
(e) 28.485		20.165	14.851	0.756	2.901	0.489
(f) 27.170		19.266	<u>12.431</u>	0.740	4.342	0.472
BDS500		(b) 26.480	12.892	15.719	0.685	1.503
	(c) 26.604	12.814	15.197	0.695	1.514	0.474
	(d) 26.568	12.498	15.254	0.695	<u>1.520</u>	<u>0.473</u>
	(e) <u>26.402</u>	13.060	15.064	<u>0.694</u>	1.518	0.474
	(f) 26.389	<u>12.791</u>	<u>15.092</u>	0.695	1.522	0.474

其中在 BDS500 数据集上,本文方法在 BLINDS-II, SSEQ 达到次优,在 BIQI, CPBDM, JNB, EMBM 达到最优;对于 N-Mo 数据集来说,本文方法在 SSEQ, EMBM 达到次优,在 BLINDS-II 达到最优,剩余指标上位列第三;而在 Mo 数据集上,指标 BIQI, BLINDS-II, SSEQ, CPBDM 并不是最优,原因可能是 Mo 数据集包含的图像均为有突出主体的图像,使图像特征相比无主体图像来说更难以提取,而指标 BIQI, BLINDS-II, SSEQ, CPBDM 均是根据图像某些特征进行评估的方法,导致并不能很好地对图像质量进行评估。但从数据上可以看出,本文方法与 BIQI, BLINDS-II, SSEQ, CPBDM 的最优数据分别相差 0.223, 0.242, 0.031 和 0.003。

3.4 对比实验

本节在现有的图像重定向方法中,选取了基于内容感知的连续型代表算法 SNS、离散型代表算法 SC、基于强化学习在不同位置特征采用不同算法的 SAMIR,以及根据图像双向相似性基于块对图像进行处理的 SVDBS,在 3 个数据集上进行对比,并用 6 种客观评价指标来验证本文方法的有效性。

3.4.1 针对不同数据集的同比例对比试验

本节将对对比各方法对 3 个数据集的重定向操作效果,其放缩比例为 0.5,图 6 给出了各方法的重定向结果。



(a)原图 (b)SNS (c)SC (d)SAMIR (e)SVDBS (f)本文方法

图 6 各方法相同缩放比例重定向结果对比图

Fig. 6 Comparison of retargeting results of each method with the same scaling ratio

各方法缩放结果客观评价指标如表 3 所列,在 Mo 数据集上本文方法在 BIQI 和 JNB 上表现次优,在 BLINDS-II, SSEQ, CPBDM, EMBM 上表现最优;N-Mo 数据集上,本文方法在 BIQI 和 EMBM 上表现次优,在 BLINDS-II, SSEQ, CPBDM 和 JNB 上表现最优。在这两个数据集上次优结果相较于最优结果只有小数点级的差异。而在 BDS500

数据集上,本文方法在 BIQI, BLINDS-II, SSEQ 以及 JNB 表现次优,在 CPBDM 和 EMBM 上表现最优。

虽然本文方法没有在所有指标上达到最优表现,次优表现较多,但根据数据可以发现,本文方法从整体上来看指标结果要优于其他对比方法,在各个指标上表现均比较稳定。

表 3 各种方法的定量评价结果

Table 3 Quantitative evaluation results of various methods

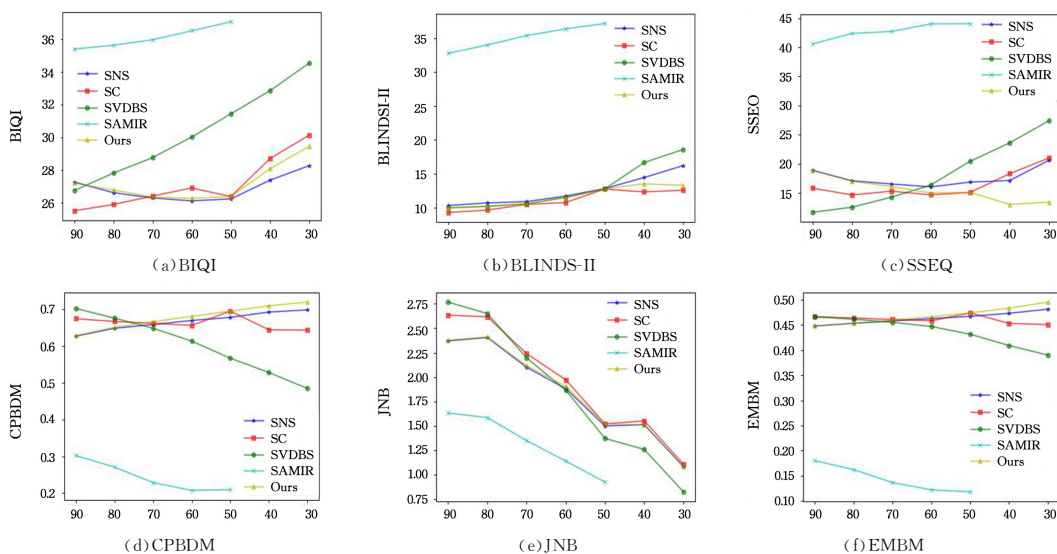
		BIQI ↓	BLINDS-II ↓	SSEQ ↓	CPBDM ↑	JNB ↑	EMBM ↑
Mo	SNS	24.683	<u>14.777</u>	16.894	0.666	2.031	<u>0.463</u>
	SC	29.641	15.848	<u>16.470</u>	<u>0.673</u>	2.213	0.461
	SAMIR	34.156	34.702	44.010	0.306	1.254	0.141
	SVDBS	28.039	17.579	20.204	0.574	1.845	0.428
	本文方法	<u>24.884</u>	14.725	16.222	0.680	<u>2.060</u>	0.470
N-Mo	SNS	26.103	<u>20.018</u>	<u>14.054</u>	<u>0.719</u>	<u>0.470</u>	<u>4.345</u>
	SC	35.427	27.147	15.568	0.711	0.461	4.466
	SAMIR	37.325	38.500	41.351	0.255	0.171	2.656
	SVDBS	30.407	24.106	16.211	0.610	0.437	3.886
	本文方法	<u>27.170</u>	19.266	12.431	0.740	0.472	4.342
BDS500	SNS	26.128	11.752	16.031	<u>0.669</u>	1.883	<u>0.463</u>
	SC	26.916	10.801	14.670	0.656	1.972	0.458
	SAMIR	36.508	36.350	44.014	0.207	1.142	0.122
	SVDBS	30.027	14.243	16.353	0.613	1.867	0.447
	本文方法	<u>26.286</u>	<u>11.537</u>	<u>14.983</u>	0.681	<u>1.907</u>	0.466

如图 6(a)所示, SNS 方法可以较好地保护图像中的重要内容和结构,如前两张图像中主体部分的比例都被很好地保留且较符合实际情况,但当图像中重要区域较多时,在划分网格时,将不可避免地出现同一网格中同时包含重要区域和非重要区域的情况,导致部分网格出现严重形变,影响最终结果,如在第三张书房图片中天花板处产生了较大扭曲。除此之外,我们从图片也可以观察到此方法对水平垂直结构的保留效果较差,在结果图像中都有不同程度的形变。如图 6(d)所示, SC 方法的结果中普遍出现锯齿形伪影,可以从后三张结果图像中明显观察到,这是由于大量接缝穿过较小面积的重要区域或集中于一个区域中,带来明显的图像质量下降和图像平滑度下降。从图 6(c)中可以观察到,图像的内容和结构都被较好地保留下来,但由于神经网络中间层的影响,图像的分辨率有一定的下降,且由于 SAMIR 方法是将多种缩放算法组合在一起,花费时间较长。从图 6(d)中可以清晰地观察到, SVDBS 对图像的内容和结构都被较好地保留下来,但由于神经网络中间层的影响,图像的分辨率有一定的下降,且由于 SAMIR 方法是将多种缩放算法组合在一起,花费时间较长。从图 6(d)中还可

以清晰地观察到 SVDBS 对图像结构保留效果较好,图像中的水平和垂直结构不存在明显的形变,但图像中保留了大量同质信息,使图像主体部分比例存在一定程度的变形,且由于此方法运用块映射,图像中出现了明显的块状模糊伪影,在第三张图像中最为明显。而由图 6(e)可知,本文方法在保留图像重要信息的同时也能够较好地保留图像结构信息,保证图像的连贯性和平滑性,没有明显的伪影,调整了主体部分在整体中的布局,更符合人眼视觉中心原则,较其他方法有一定的优越性。

3.4.2 针对不同数据集的多比例对比实验

为进一步测试本文方法在不同放缩比例下的重定向结果,本节将使用各种方法对 BDS500 数据集进行 90%~30% 的放缩测试分析,其中 SAMIR 由于神经网络结构无法实现 40%~30% 放缩。各方法缩放结果指标随放缩比例变化如图 7 所示,其中,本文方法在 SSEQ, CPBDM, EMBM 指标上为最优,且随放缩比例的增加变化都较为平滑。值得一提的是,在 SSEQ 指标上,本文方法在放缩程度较大时的效果较其他方法优越性十分明显,而在其他指标上本文方法为次优且与最优方法相差较小。



注:横坐标为缩放比例因子。

图 7 各方法多缩放比例重定向结果指标图

Fig. 7 Comparison chart of retargeting results of each method with multiple ratios

结束语 本文提出了一种基于超像素的内容感知图像重定向方法。首先对图像进行初始网格化操作,并使用超像素对输入图像进行处理,使输入图像特征更明显;接着利用图像梯度信息和图像显著性信息构成图像重定向权重热力图,以达到同时保留图像结构和内容信息的目的;随后根据图像重定向权重热力图对初始网格进行优化,其中包含形变分量和网格变形分量,前者保证图像在放缩过程中图像内容的平滑度与完整性,后者保证图像不会产生明显扭曲伪影,不断迭代优化,在达到终止调节后完成整个放缩过程,输出最后的结果图像。

本文方法效果稳定,能够适用于各类图像的不同比例的高质量图像重定向,符合人类审美原则。

参 考 文 献

- [1] SHAMIR A S, TOG A J. Seam carving for content-aware image resizing[J]. ACM Trans. Graph., 2007, 26:1-10.
- [2] PRITCH Y, KAV-VENAKI E, PELEG S. Shift-Map Image Editing[C]// 2009 IEEE 12th International Conference on Computer Vision, 2009.
- [3] WOLF L, GUTTMANN M, COHEN-OR D. Non-homogeneous Content-driven Video-retargeting[C]// IEEE International Conference on Computer Vision, 2007.
- [4] KRAEHENBUEHL P, LANG M, HORNING A, et al. A system for retargeting of streaming video[J]. ACM SIGGRAPH Asia 2009 Papers, 2009, 28(5):1-10.
- [5] WANG Y S, TAI C L, SORKINE O, et al. Optimized Scale-and-Stretch for Image Resizing[J]. ACM SIGGRAPH Asia 2008 Papers, 2008, 27(5):118.
- [6] ITTI L, KOCH C, NIEBUR 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.
- [7] ACHANTA R, ESTRADA F, WILS P, et al. Salient Region Detection and Segmentation[C]// Computer Vision Systems(ICVS 2008). 2008.
- [8] ACHANTA R, HEMAMI S F, ESTRADA† F, et al. IEEE[J]. Requery-tuned salient region detection, 2009.
- [9] ZHAI Y, SHAH M. Visual attention detection in video sequences using spatiotemporal cues[C]// Proceedings of the 14th ACM International Conference on Multimedia. Santa Barbara, CA, USA, 2006:23-27.
- [10] CHENG M, MITRA N J, HUANG X, et al. Global Contrast Based Salient Region Detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 37(3):569-582.
- [11] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11):2274-2282.
- [12] PETERSON L E. K-nearest neighbor [J]. Scholarpedia, 2009, 4(2):1883.
- [13] RUBINSTEIN M, GUTIERREZ D, SORKINE O, et al. A comparative study of image retargeting[J]. ACM Transactions on Graphics(TOG), 2010, (6):1-10.
- [14] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Machine Intelligence[J]. Contour Detection and Hierarchical Image Segmentation, 2011, 33(5):898-916.
- [15] ZHOU Y, CHEN Z, LI W, et al. Weakly Supervised Reinforced Multi-Operator Image Retargeting [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2020, 31(1):126-139.
- [16] SIMAKOV D, CASPI Y, SHECHTMAN E, et al. Summarizing visual data using bidirectional similarity[C]// IEEE Conference on Computer Vision and Pattern Recognition(CVPR 2008). 2008.
- [17] ZHANG T, YU M, GUO Y, et al. Content-Aware Retargeted Image Quality Assessment[J]. Information, 2019, 10(3):111.
- [18] MOORTHY A K, BOVIK A C. A Two-Step Framework for Constructing Blind Image Quality Indices[J]. IEEE Signal Processing Letters, 2010, 17(5):513-516.
- [19] SAAD M A, BOVIK A C, CHARRIER C. Blind image quality assessment: A natural scene statistics approach in the DCT domain[J]. IEEE Transactions on Image Processing, 2012, 21(8):3339-3352.
- [20] LIU L, LIU B, HUANG H, et al. No-reference Image Quality Assessment Based on Spatial and Spectral Entropies[J]. Signal Processing Image Communicator, 2014, 29(8):856-863.
- [21] FERZLI R, KARAM L J. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB) [J]. IEEE Transactions on Image Processing, 2009, 18(4):717-728.
- [22] GUAN J, ZHANG W, GU J, et al. No-reference blur assessment based on edge modeling[J]. Journal of Visual Communication and Image Representation, 2015, 29:1-7.
- [23] NARVEKAR N D, KARAM L J. A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection[C]// International Workshop on Quality of Multimedia Experience. 2009.



CHEN Meiyang, born in 2001, bachelor. Her main research interest is image processing.



BI Xiuli, born in 1982, Ph.D, associate professor, Ph.D supervisor, is a member of China Computer Federation. Her main research interests include image processing and multimedia information security.