

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2013 4
Vol.18 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

2013 年 4 月 第 18 卷 第 4 期(总第 204 期)

目次

综述

车载视觉系统中的行人检测技术综述 许腾,黄铁军,田永鸿(359)

图像处理和编码

利用边缘匹配矢量量化和变率编码的高嵌入效率信息隐藏 王凌飞,潘志斌,邓晓曼,胡森(368)

小波变换和 SHA-1 相结合的图像压缩加密 李园园,张绍武(376)

图像分析和识别

基于 Haar 纹理的非结构化道路消失点检测 王永忠,文成林(382)

多种人群密度场景下的人群计数 覃勋辉,王修飞,周曦,刘艳飞,李远钱(392)

计算机图形学

大规模散乱点的 k 邻域快速搜索算法 杨军,林岩龙,王阳萍,王小鹏(399)

三角 Bézier 曲面的几何约束形变 陈军(407)

遥感图像处理

面向测绘任务的光学遥感影像质量分级评定 王昱,时红伟,胡闻达(415)

采用高斯归一化水体指数实现遥感影像河流的精确提取 沈占锋,夏列钢,李均力,骆剑承,胡晓东(421)

地理信息技术

图层约束下的点状自然灾害风险地图综合 潘东华,王静爱,贾慧聪,袁艺(429)

模板阴影体扩展方法 曹雪峰,万刚,李锋,李明(436)

“第 9 届中国计算机图形学大会”专栏

利用 SIFT 特征的非对称匹配图像拼接盲检测 杜振龙,杨凡,李晓丽,郭延文,沈钢纲(442)

视觉感知的图像和视频抽象 宋杰,徐丹,时永杰(450)

局部线性模型优化的灰度图像彩色化 厉旭杰,赵汉理,黄辉(460)

分区域去运动模糊 张璐,盛斌,马利庄(467)

脑血管体绘制的快速表意式增强 吴腾飞,骆岩林,田云,武仲科,闫建平(476)

第七届国际图象图形学学术会议征文通知 封 3

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 4 April 2013

Contents

Review

Survey on pedestrian detection technology for on-board vision systems Xu Teng, Huang Tiejun, Tian Yonghong(359)

Image Processing and Coding

Highly efficient information hiding using SMVQ-compressed images and variable length coding
..... Wang Lingfei, Pan Zhibin, Deng Xiaoman, Hu Sen(368)

Image compression and encryption based on DWT and SHA-1 Li Yuanyuan, Zhang Shaowu(376)

Image Analysis and Recognition

Vanishing point detection of unstructured road based on Haar texture Wang Yongzhong, Wen Chenglin(382)

Counting people in various crowd density scenes using support vector regression
..... Qin Xunhui, Wang Xiufei, Zhou Xi, Liu Yanfei, Li Yuanqian(392)

Computer Graphics

Fast algorithm for finding the k -nearest neighbors of a large-scale scattered point cloud
..... Yang Jun, Lin Yanlong, Wang Yangping, Wang Xiaopeng(399)

Shape modification of triangular Bézier surfaces with geometric constraints Chen Jun(407)

Remote Sensing Image Processing

Optical image quality grading assessment orienting to surveying & mapping mission requirements
..... Wang Yu, Shi Hongwei, Hu Wenda(415)

Automatic and high-precision extraction of rivers from remotely sensed images with Gaussian normalized water index
..... Shen Zhanfeng, Xia Liegang, Li Junli, Luo Jiancheng, Hu Xiaodong(421)

Geoinformatics

Generalization for point features in risk mapping of natural hazards based on layer constraint
..... Pan Donghua, Wang Jing'ai, Jia Huicong, Yuan Yi(429)

Extension method of stencil shadow volume Cao Xuefeng, Wan Gang, Li Feng, Li Ming(436)

Issum of Chinagraph'2012

Fogery image blind detection by asymmetric search based on SIFT
..... Du Zhenlong, Yang Fan, Li Xiaoli, Guo Yanwen, Shen Ganggang(442)

Visual perception based image and video abstraction Song Jie, Xu Dan, Shi Yongjie(450)

Local linear model optimization based grayscale image colorization Li Xujie, Zhao Hanli, Huang Hui(460)

Motion deblurring based on image segmentation Zhang Lu, Sheng Bin, Ma Lizhuang(467)

Fast illustrative enhancements on volume rendering of cerebral vessel
..... Wu Tengfei, Luo Yanlin, Tian Yun, Wu Zhongke, Yan Jianping(476)

中图分类号: TP391 文献标识码: A 文章编号: 1006-8961(2013)04-0450-10

论文引用格式: 宋杰, 徐丹, 时永杰. 视觉感知的图像和视频抽象[J]. 中国图象图形学报, 2013, 18(4): 450-459.

视觉感知的图像和视频抽象

宋杰, 徐丹, 时永杰

云南大学计算机科学与技术系, 昆明 650091

摘要: 可靠的视觉感知估计体现的是对图像中有意义的那部分结构信息的理解。基于此, 描述一个完整的图像和视频抽象化框架, 通过一种边缘保持的滤波器有区分地风格化抽象图像和视频的显著部分和非显著部分, 并保持两部分之间的和谐过渡。首先, 对于显著区域的识别引入一种自动显著性目标分割算法, 基于局部空间邻域的自动显著性计算。考虑到实际需要, 提供了交互式技术, 方便有目的地去指定图像中的显著结构信息。然后, 根据生成的显著结构信息掩模使用单尺度的各向异性滤波处理显著部分, 而用多尺度的同种滤波处理非显著部分以实现更强程度的抽象效果。本文方法生成的图像和视频的抽象化不仅呈现了视觉满意的非真实感效果, 而且在软量化后还可以应用于其他风格的非真实感渲染(NPR)效果的实现。实验结果显示, 处理一定数量的图像及视频, 本文方法可以得到期望的结果。

关键词: 边缘保持的滤波; 显著性; 交互式技术; 各向异性; 多尺度; 非真实感渲染

Visual perception based image and video abstraction

Song Jie, Xu Dan, Shi Yongjie

Department of Computer Science and Technology, Yunnan University, Kunming 650091, China

Abstract: Reliable estimation of visual perception reflects the understanding of the meaningful structure in an image. In this paper, we describe a complete abstraction framework for images and videos that explicitly responds to this goal. The method stylizes salient images or video parts and non-salient images or video parts differently by an edge preserving filter, and keeps a harmonious transition between the two parts. First, we introduce an automatic salient object segmentation algorithm to distinguish salient regions, and it is a saliency computation based local spatial neighbors. Taking into account the actual needs, we provide an interactive technology, which can be convenient for specifying image salient structure information on purpose. Based on the generated salient structure information mask, we use a single-scale anisotropic filter to process the salient parts, and use a multi-scale anisotropic filter to process the non-salient parts so that we can implement a strong abstraction effect. The proposed method generates a kind of image and video abstraction that does not only represent a preferable visual effect, but also can be applied to the implementation of another non-photorealistic rendering (NPR) result after soft image quantization. Experiments show that our algorithm could get the desired result for processing a certain number of images and videos.

Key words: edge preserving filter; saliency; interactive technology; multi-scale; non-photorealistic rendering

收稿日期: 2012-08-20; 修回日期: 2012-11-19

基金项目: 国家自然科学基金项目(61163019); 高等学校博士学科点专项科研基金项目(20095301110006)

第一作者简介: 宋杰(1986—), 男, 云南大学计算机科学与技术系2010级计算机应用专业在读硕士研究生, 主要研究方向为计算机图形图像处理及计算机视觉。E-mail: hknescafe@gmail.com

0 引言

通常在抽象化时,艺术家们会认真地操作图像细节以控制自己感兴趣和理解的结构信息,绘制出局部显著部分的风格抽象画。受此现象启发,一系列用显著性值来进行非真实感渲染的方法运用而生,并产生了有趣的结果。其中最为杰出的代表是早期 DeCarlo 等人^[1]研究工作。他们利用眼球跟踪硬件设备来获取图像显著性信息为之后抽象化提供必要参数。显然,硬件的介入给用户带来了很大的不便,他们需要的是一个完全由自己控制的流程。幸运地是,自 Itti 等人^[2]基于生物启发模型的基础首次定义图像的显著性以来,越来越多的显著性计算模型被提出和应用^[3-4],并产生了可靠的结果。根据这一基础,引入一种图像显著性定义,能均匀地高亮出相应的显著性区域,得到全分辨率的显著图。

图像风格化抽象技术中,除了文献[1]外,Wen 等人^[5]提出的彩色素描图生成技术以及 Zhao 等人^[6]提出的类似油画生成技术等都会对图像细节进行过多的平滑,使得图像纹理得不到较为完整地展现。为了让用户在抽象化图像的同时,体味到纹理所传递的一种美感,采用一种称做 Kuwahara 滤波

的边缘保持滤波。这种滤波在高对比度的区域以及低对比度区域都能有效地保护纹理信息,在处理整幅图像时,保持了大体一致的抽象程度。

图1为一个完整的图像和视频抽象化处理框架。过程始于对输入图像的两方面处理,包括构建图像金字塔以及定义图像显著性。一方面,对于显著性的定义,描述了两种方法:基于局部空间邻域的自动显著性计算和用户交互操作的显著性指定。根据显著性定义生成的掩模,将一幅完整的图像分解为两个视觉组成部分:显著部分和非显著部分;运用算法分别渲染图像的相应区域,通过使用混合单尺度、多尺度的各向异性 Kuwahara 滤波实现符合视觉分析的图像和视频风格化抽象。另一方面,构建图像金字塔的目的是将滤波层面从单尺度扩展到多尺度上。按金字塔由粗到细的顺序,在每个级别上进行平滑结构张量以及各向异性 Kuwahara 滤波的计算,然后上采样这些结果到下一较细级别。基于局部方差的近似,上采样的滤波结果与当前金字塔级别的图像数据融合,并且从这个融合结果中计算平滑结构张量;基于各向异性值的计算,得到的平滑结构张量和之前上采样的平滑结构张量(前级)进行融合,并根据结果计算出融合图像数据的各向异性 Kuwahara 滤波。重复执行这个过程直到金字塔的最细级别。

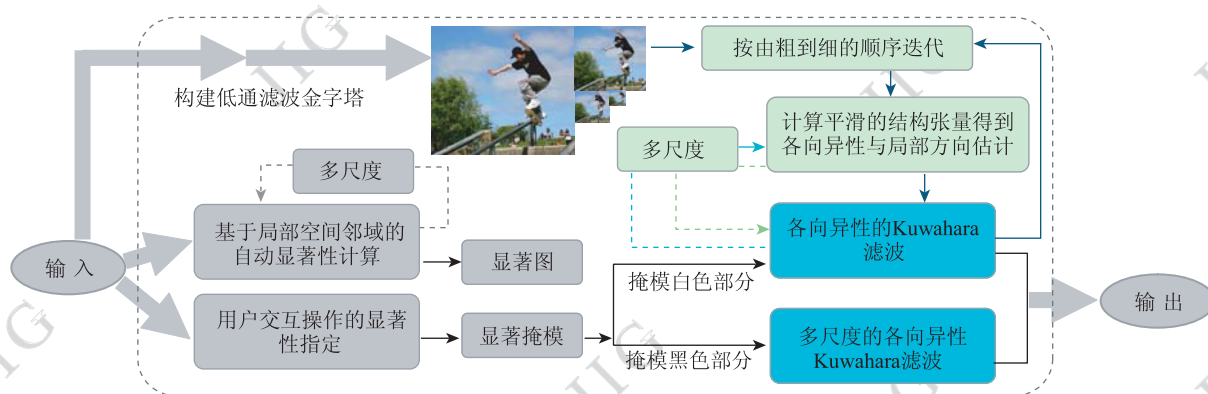


图1 技术框架流程图

Fig. 1 Framework of the proposed technique

根据这个系统框架,一个用户需要做的仅仅是在几秒钟的扫描图像之后,提供必要的参数,即给出图像中显著性区域的位置信息。不需要任何额外的硬件设施,用户依靠简单的鼠标操作就可以完成数据提交工作。通过对收集到的数据进行分析,系统就可以自动地将一幅自然图像转换成凸显显著区域

的抽象画(见图2),而做到这一切或许根本不要求用户有着任何的绘画天赋。

1 相关工作

图像显著性是图像中重要的视觉特征,体现了



(a) 近处的男孩为图像的显著结构信息

(b) 远处女人为图像的显著结构信息

(c) 近处的房屋墙壁为图像的显著结构信息

图 2 基于视觉感知的抽象化示例

Fig. 2 The abstraction example based visual perception

人眼对图像的某些区域的重视程度。自 1998 年 Itti 等人^[2]的工作以来,产生了大量的自底向上的显著性计算模型。这些方法利用亮度、颜色、边缘等底层特征属性来决定图像的某个区域与其周围环境的对比度。最近,Cheng 等人^[7]提出了一种提取全分辨率的全局显著性图像的分析方法,形成了基于区域对比度的显著性图。借助此概念在图像内容分析的基础上衍生出了基于局部空间近邻的自动显著性计算模型。

早期的边缘保持滤波包括双边滤波^[8]以及均值漂移滤波^[9]技术。两者能够平滑低对比度区域同时保持高对比度边缘。所以,对于处理高对比度图像这两种技术也许会失效,即或是不执行抽象,或是阈值的使用导致相关联的信息被抹掉。而对于处理低对比度的图像也经常如此,太多的信息会被忽略。另一种边缘保持滤波称作 Kuwahara 滤波^[10]能够克服这个局限。基于局部区域的平整性,Kuwahara 滤波甚至在高对比度的区域适当地除去细节而在低对比度区域保护结构的边缘。因此,它在处理整幅图像时,可以保持大体一致的抽象程度。各向异性的 Kuwahara 滤波^[11]技术建立在由 Papari 等人^[12]提出的广义 Kuwahara 滤波的概念基础上并且利用定义在椭圆上的权重函数来取代先前定义在圆盘扇区上权重函数。通过改变这些椭圆的形状、尺度以及方向来自适应输入图像的局部结构信息,使得方向性的图像特征被更好地保持和增强。

DeCarlo 等人^[1]的图像抽象化方法以均值漂移算法为基础,利用眼球跟踪硬件设备来获取图像的显著性信息,进而通过调整分辨率达到增强的目的。类似地,Wen 等人^[5]实现图像抽象化的目的也是将

相同算法作为基础用于预处理,只不过他们是通过调整颜色来增强图像的显著区域内容。在视频抽象化方面,主要有两类不同的方法:其一,不使用笔刷,直接从视频剪辑中抽象化帧图像,代表有 Winnemöller 等人^[13]和 Zhang 等人^[14]的方法。前者通过改变亮度及颜色对比度来抽象化区域和边界信息;后者将 DeCarlo 等人^[1]的图像抽象化方法扩展至视频处理中,使用光流将分割信息在每帧中进行传播,借助一个参考视频,利用视频流颜色体系替代方法改变目标输入视频的颜色,达到抽象化的目的。其二,基于笔刷的视频风格化抽象,使用精确的基于物理或基于样例的笔刷来表示物体外貌及结构,代表有 Hertzmann 等人^[15]和 Lin 等人^[16]的方法。虽然后者利用改变笔刷的形状与物体纹理结构达成一致以及将笔刷单元限制在每个分割区域,并提出了一个服从渲染、回退完成的策略,解决了一些由笔刷引导的抽象化带来的问题,如笔刷偏离目标物体,笔刷的闪烁,但同本文提出的抽象化方法相比,其额外的操作,如图像的分割、笔刷的建库和放置等处理势必加深了算法的复杂度。对比之下,本文方法不需要琐碎的步骤就能在输入视频的抽象化基础上很好地与原有纹理及非纹理区域保持一致,并且不会出现所谓的闪烁或抖动现象。在本文系统中,利用各向异性的 Kuwahara 滤波准确地捕获帧图像的方向性特征及结构或称之纹理特征及结构,并在此基础上通过单尺度和多尺度相交叉的方法来区别处理由显著性信息导向的帧各部分,利用对主体运动目标的识别,将显著性信息传播至其他视频帧中。确保在视频抽象时良好的时间连贯性,减小了整体视频处理的复杂度,提高了时间效率。

2 视觉显著性估计

2.1 基于局部空间邻域的自动显著性计算

认知心理学家提出了许多视觉注意的数学模型来模拟人注视点的转换过程。这些模型可以分为两大类:快速的、下意识的、自底向上的模型和慢速的、任务依赖的、自顶向下的模型。根据观察,人们会更加注意到:1)位于图像中央附近的区域,因此很有可能这块区域就定义了图像的显著结构信息;2)显著的结构信息在通常情况下总是不同于周围的图像内容。

2.1.1 单一尺度的显著性计算

依据上述观察得到的两个特性来进行自底向上、数据驱动的显著性计算,引入一种基于局部空间邻域的显著性计算思想^[17]。借助 Cheng 等人的基于区域级来计算显著图的思想,首先用基于图的图像分割方法^[18]将输入图像分割成 N 个区域 $r_i, i=1, \dots, N$,将与某个区域相关联的环境内容定义为空间邻域的一个集合。给定区域 r_i ,以及它的空间邻域 $r_k = \{r_k, k=1, \dots, K\}$,分别以 $y = -\log(1-x)$ 和 $y=x$ 作为公式模板的两个对区域 r_i 的显著性定义为

$$S(r_i) = -w(r_i) \log \left(1 - \sum_{k=1}^K \lambda_{ik} D_r(r_i, r_k) \right) \quad \text{或} \quad S(r_i) = \sum_{k=1}^K w(r_i) \lambda_{ik} D_r(r_i, r_k) \quad (1)$$

与后者的线性关系不同地是,前者的非线性关系不再表示互不相干的独立关系,而表示相互作用关系。这样就使得整体不再是简单地等于部分之和,而可能出现不同于线性叠加的增益或亏损。因此,选择 $y = -\log(1-x)$ 来定义区域的显著性。其是一个单调递增函数, x 越大, y 值越大,即颜色距离越大,显著性值也就越大。所有这些正好对应于上述的其中一个特性,即显著的结构信息在通常情况下总是不同于周围的图像内容。式(1)中 λ_{ik} 为区域 r_k 的面积和区域 r_i 所有邻域面积之和的比值, $w(r_i)$ 为区域 r_i 的权值, $D_r(\cdot, \cdot)$ 为两个区域在 CIE L*a*b* 色度空间下的颜色距离度量,类似 Cheng 等人的方法,为每个区域建立颜色直方图,并将两个区域 r_i 和 r_k 的颜色距离定义为

$$D_r(r_i, r_k) = \sum_{p=1}^{N_i} \sum_{q=1}^{N_k} f(c_{i,p}) f(c_{k,q}) D(c_{i,p}, c_{k,q}) \quad (2)$$

式中, $f(c_{i,p})$ 为第 p 个颜色在第 i 个区域 r_i 的所有 N_i 种颜色中出现的概率, $f(c_{k,q})$ 为第 q 个颜色在第 k 个区域 r_k 的所有 N_k 种颜色中出现的概率。

更进一步,通过在式(1)中引进高斯衰减函数来定义权值 $w(r_i)$,目的是为了将空间位置信息加入进来,以定义图像中央附近的区域显著性部分。因此,对任意区域 r_i ,显著性重定义为

$$S(r_i) = -\exp \left(-\alpha \left(\left(\frac{Dx_i}{X} \right)^2 + \left(\frac{Dy_i}{Y} \right)^2 \right) \right) \times \log \left(1 - \sum_{k=1}^K \lambda_{ik} D_r(r_i, r_k) \right) \quad (3)$$

式中, X, Y 分别表示图像的宽度和高度。 (Dx_i, Dy_i) 为在区域 r_i 中的像素到图像中心的空间平均距离。 α 为可调参数,通常情况下 $\alpha=3$ 。

从图3可以明显看出,与 Cheng 等人提出的方法相比较,本文算法均匀地高亮了显著目标物体。



(a) 原图 (b) Cheng 等人方法^[8] (c) 本文方法

图3 显著性图的视觉效果对比

Fig. 3 Visual effect comparison of saliency maps

2.1.2 多尺度的显著性计算

为了使算法能够在复杂的背景下更加的鲁棒,这里也将单尺度的基于区域的显著性计算扩展为多尺度的基于区域的显著性计算。首先,设定 M 组不同参数,每组参数设定了不同的规格,然后根据一组给定的参数分割图像为 N 个区域,循环地执行单一尺度的显著性计算直到第 M 组参数。基于这个多尺度的概念,同时将显著性值从区域级传播至像素级,因此任意区域 r_i 中的像素 x 显著性值定义为

$$S^m(x) = \frac{\sum_{m=1}^M \sum_{i=1}^N S(r_i^m) D(I_x, c_i^m)^{-1}}{\sum_{m=1}^M \sum_{i=1}^N D(I_x, c_i^m)^{-1}} \quad (4)$$

式中, c_i^m 为区域 r_i^m 的颜色中心, $D(I_x, c_i^m)$ 为从像素 x 到区域 r_i^m 颜色中心的颜色距离。

2.2 交互式的显著性定义

任何自动的显著性检测算法的适用范围只能局限于部分的图像数据集,正如2.1节描述的算法对于检测处在图像周边的显著性目标则会失效。一种更好的解决方案是使用交互式的方法来自定义显著结构信息,以避免上述所提到的问题。借助近来比较热的基于测地星凸的交互式图像分割技术^[19]作为此次选择。这种技术基于原始的图切割算法,并融入了形状优先约束,依靠对 Veksler^[20]的星凸优先的扩展,融入的形状约束增强了区域连通等性质,有效地限制了将图像空间分割成较小子集的可能性,消除了零碎的错误分割,使得提取的前景目标结构更加的完整。

2.2.1 星凸优先的扩展

在 Veksler 方法中,首次将星凸集合作为形状优先用于图像分割,如图4所示。图4(a)标记星中心为 c ,星凸 y 表示集合, p 和 q 是图像域 Ω 中经过点 c 射线 $\Gamma_{c,p}$ 上的像素点,并且 q 点位于点 c 和点 p 之间,如果 p 被标记为前景,那么 q 也必须被标记为前景。若在集合 y 中存在射线 $\Gamma_{c,p}$,那么点 p 对于星中心 c 来说是可见的。如果位于集合 y 中的每个点对于星中心 c 都是可见的,那么集合 y 就是关于星中心 c 的星凸。因此形式化星凸约束为能量函数

$$\forall q \in \Gamma_{c,p}, E_{pq}^*(y_p, y_q) = \begin{cases} \infty & y_p = 1 \text{ 且 } y_q = 0 \\ 0 & \text{其他} \end{cases}$$

$$E^*(y | c) = \sum_{p \in \Omega} \sum_{q \in \Gamma_{c,p}} E_{pq}^*(y_p, y_q) \quad (5)$$

式中,定义星凸 y 为函数 $y: \Omega \rightarrow \{0, 1\}$, $\forall p \in \Omega: p \in y \Leftrightarrow y_p = 1$ 。

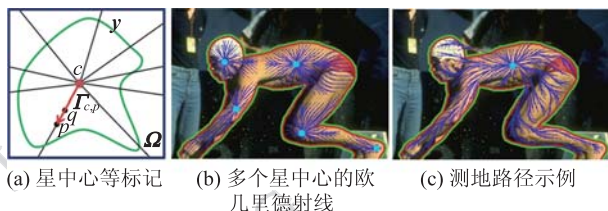


图4 关于星中心的一个星凸目标前景

Fig. 4 A star-convex object wrt star center

根据上述内容, $\Gamma_{c,p}$ 被定义为星中心 c 和任意点 p 之间的最短路径。而 Gulshan 等人^[19]则将这一概念从欧几里德泛化为测地学。因为测地路径可以任意弯曲,这样更加适应于图像数据特征,以至于在减少需要的星中心数目的情况下不会影响可见性的伸展(见图4(c))。在其方法中,测地距离被定义为

$$d_g(a, b) = \min_{\Gamma \in P_{a,b}} \sum_{i=1}^{n-1} \left((1 - \gamma_g) d(\Gamma^i, \Gamma^{i+1})^2 + \gamma_g \|\nabla I(\Gamma^i)\|^2 \right)^{\frac{1}{2}} \quad (6)$$

式中的最外面括号内定义了一个离散路径的长度。 $P_{a,b}$ 表示两个像素点 a, b 之间的所有离散路径集合。 Γ 是一个 n 像素的任意参数化离散路径,表示为集合 $\{\Gamma^1, \Gamma^2, \dots, \Gamma^n\}$, $d(\Gamma^i, \Gamma^{i+1})$ 是连续像素间的欧几里德距离。定量 $\|\nabla I(\Gamma^i)\|^2$ 是在点 (Γ^i, Γ^{i+1}) 之间的图像梯度的一个有限差分近似。系数 γ_g 控制了欧几里德距离和测地长度的权重。进一步地,定义像素点 a, b 之间的测地路径为

$$\Gamma_{a,b} = \arg d_g(a, b) \quad (7)$$

另外, Gulshan 等人假设 c 为星中心的集合,将式(6)中的两点间的测地距离扩展至点与点集 c 间的测地距离,得到了称作测地森林的结构,并且改写星凸能量函数 $E^*(y | c)$ 为 $E^*(y | x, c)$ 以表示范围扩充至整个图像。

2.2.2 分割系统下的星凸应用

文献[19]描述 Boykov 等人^[21]能量最小化函数为

$$E(y | x) = \sum_{i \in \Omega} U(y_i | x) + \lambda \sum_{(i,j) \in N} V(y_i, y_j | x)$$

$$\bar{y} = \arg \min_{y \in Y} E(y | x) \quad (8)$$

式中, x 表示图像, λ 是成对项的权重, N 表示邻域像素对集。数据项 $U(y_i | x)$ 的计算使用了一个颜色模型的负对数似然函数,记为 $U(y_i | x) = -\log h(x_i | \alpha)$, α 在每个像素上取 0 或 1 值, 0 表示背景, 1 表示前景, 这个颜色模型从用户提供的笔划学习得来。成对项定义为 $V(y_i, y_j | x) = \exp(-\beta \|x_i - x_j\|^2)$ 。进一步地,式(8)中定义的能量函数 $E(y | x)$ 被最小化并服从 y 是星凸这个条件

$$\min_y E(y | x) \Leftrightarrow \min_y E(y | x) + E^*(y | x, c) \quad (9)$$

式中, $S^*(c)$ 表示所有不同形状的星凸集合,存在一个星凸 y 属于 $S^*(c)$, 使得能量函数 $E(y | x)$ 最小化。因为图像分割能够被估计为一个能量函数的全局最小值,而通过之前的描述,有效地获得式(9)的全局最小值是可能的,因此分割可以被实现。另外,以上星凸的约束限制了形状的范围。基于这两点,就可以逐步地从图像中获取可靠的前景目标。

图5为上述方法的示例,图中分割边缘框内的笔划用来指示前景,边缘框外的笔刷指示背景。通

过观察与对比,可以清楚地发现交互式的显著性检测不仅提供了一个方便用户自定义的接口,同时分割出的目标前景结构显得更加的完整。

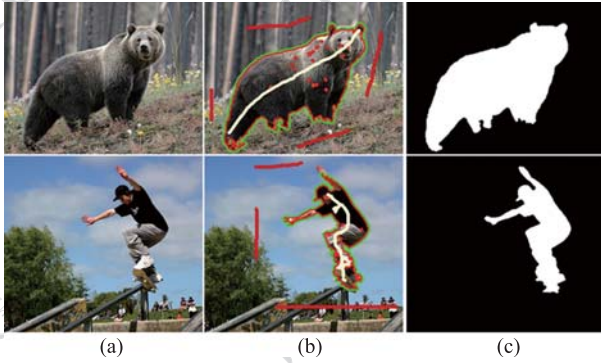


图5 交互式的图像分割

Fig. 5 Interactive image segmentation

3 基于视觉感知的滤波处理

借助各向异性的思想,Kuwahara 滤波提供了一个桥梁使得人们可以有效地生成不同的风格化抽象效果。根据 Kyprianidis^[22]所描述的多尺度概念下的滤波思想,结合实验目的,对各向异性的 Kuwahara 滤波实现过程进行改进,使得最终生成的结果符合预期效果。

3.1 图像的局部结构分析

各向异性的 Kuwahara 滤波在原有滤波的基础上融入了由计算平滑的结构张量获取的各向异性值与局部方向估计信息,使得滤波器能够自适应地根据输入图像的局部特征改变自身的形状、尺度以及方向,更好地凸显方向性的图像特征。

3.1.1 平滑的结构张量

结构张量的计算需要偏导的近似。为了获取更精确的结果,直接使用一个 Scharr 算子,通过对图像用相应的内核进行卷积操作来计算图像的梯度。选取 Scharr 算子的原因是因为若使用小的核,精度会比较高,可以有效地解决精度差的问题,同时保证速度的不变。 3×3 Scharr 滤波器定义为

$$S_x = \begin{pmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{pmatrix}, S_y = S_x^T$$

上面的 3×3 矩阵表示了 Scharr 滤波器的水平和垂直方向上的卷积掩模。这样处理完后,便可得到每个像素点的结构张量。进一步使用高斯滤波对结构张量进行平滑。假设 f 表示高斯平滑后的图像,则 f

的结构张量定义为

$$J = \begin{pmatrix} f_x f_x & f_x f_y \\ f_x f_y & f_y f_y \end{pmatrix} = \begin{pmatrix} E & F \\ F & G \end{pmatrix} \quad (10)$$

式中 f_x 、 f_y 分别定义为 $f_x = S_x * f$ 、 $f_y = S_y * f$ 。对于彩色图像,平滑的结构张量通过不同颜色通道的梯度乘积和给出。例如对于一个 RGB 色彩图像,有

$$J = \nabla R \nabla R^T + \nabla G \nabla G^T + \nabla B \nabla B^T \quad (11)$$

图像中每个像素的特征值根据 $\det(J - \lambda_i I) = 0$ 可计算得到

$$\lambda_{1,2} = \frac{E + G \pm \sqrt{(E - G)^2 + 4F^2}}{2} \quad (12)$$

相应的特征向量定义为

$$\xi = \begin{pmatrix} F \\ \lambda_1 - E \end{pmatrix} \quad (13)$$

$$\eta = \begin{pmatrix} \lambda_2 - G \\ F \end{pmatrix}$$

假设 λ_1 表示模较大的主特征值, λ_2 表示模较小的次特征值。对应的 ξ 是主特征向量表示较大变化的方向, η 是次特征向量表示较小变化的方向。

3.1.2 各向异性值与局部方向估计

平滑结构张量的特征值和特征向量提供了用来估计局部方向以及各向异性值的信息。计算各向异性的值为

$$A = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2} \quad (14)$$

取值范围从 0 到 1。由于平滑结构张量的次特征值 λ_2 衡量了梯度偏离由主特征向量定义的轴的程度,这提供了一种方式来衡量方向估计的质量。当 $\lambda_1 \gg \lambda_2$ 时,表示在考虑到的邻居上存在着一个主导方向;而另一方面,当 $\lambda_1 \approx \lambda_2$ 时,表示在梯度之间没有特定的轴。因此,A 取 0 对应于各向同性,取 1 对应于各向异性。显然,次特征向量方向上的局部方向可直接从平滑结构张量提供的信息获取

$$\varphi = \frac{1}{2} \arctan\left(\frac{2F}{E - G}\right) + \frac{\pi}{2} \quad (15)$$

3.2 构建金字塔

金字塔的构建一般使用重采样的卷积执行。为了建立 $P+1$ 级金字塔,通常情况下:1) 通过对输入图像进行滤波并以 2 为步长进行子抽样来计算输入图像的低分辨率的近似值;2) 为了生成与输入等分辨率的预测图像,对第 1 步的输出进行内插(因子仍为 2)并进行过滤。计算第 2 步的预测值和第 1 步的输入之间的差异。以 $j(j=1, 2, \dots, P+1)$ 级

预测残差进行标识的这个差异将用于原始图像的重建。在没有量化差异的情况下,预测残差金字塔可以用于生成相应的近似金字塔(包括原始图像),而没有误差。在实验中,使用双线性插值法来构建图像金字塔,对于大多数的情况,这种方法生成的效果跟其他已知的方法相比没有太大差异,同时速度较快。

3.3 各向异性的 Kuwahara 滤波处理过程

借助上面所求得各向异性值及局部方向的估计,定义两个转换矩阵

$$S = \begin{pmatrix} \left(\frac{\alpha + A}{\alpha}\right)^{-1} & 0 \\ 0 & \left(\frac{\alpha}{\alpha + A}\right)^{-1} \end{pmatrix}$$

$$R_\varphi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

式中, S 用于调整滤波器形状的离心率,由各向异性值给定,实验中设置 $\alpha = 1$ 。矩阵 R_φ 定义了 φ 弧度的旋转变换。这样, SR_φ 就定义了一个线性的坐标变换将圆转换为椭圆(图6)。各向异性的 Kuwahara 滤波将上面定义的椭圆划分成若干不同的扇区,假设 N 表示扇区数,实验中设置值为4。

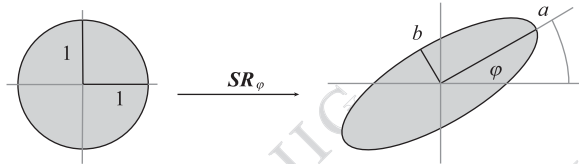


图6 转换矩阵 SR_φ 将圆转换为椭圆

Fig. 6 The mapping SR_φ maps an unit disc to the rotated ellipse

根据 Kyprianidis^[12]的定义,各向异性 Kuwahara 滤波的输出定义为椭圆所有扇区局部平均的一个权重和。因此,需要描述的情况只针对一个扇区即可。和 Kyprianidis 不同的是,本文方法只需两个步骤就可以完成上述的工作。

1) 首先令 $a = \left(\frac{\alpha + A}{\alpha}\right)^{-1}$, $b = \left(\frac{\alpha}{\alpha + A}\right)^{-1}$, 依据上述的变换矩阵 S 和 R_φ 提供的信息,定义一个 2 维 2×3 变换矩阵

$$M = \begin{pmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \end{pmatrix}$$

式中, $m_0 = a \cos \varphi$, $m_1 = a \sin \varphi$, $m_2 = a(-\cos \varphi \cdot c_x - \sin \varphi \cdot c_y) + o_x$, $m_3 = b(-\sin \varphi)$, $m_4 = b \cos \varphi$, $m_5 = b(\sin \varphi \cdot c_x - \cos \varphi \cdot c_y) + o_y$ 。 c_x 和 c_y 分别表

示每个像素点在 x 方向和 y 方向上的变换中心, o_x 和 o_y 则是对应在两个方向上的偏移量,实验中设 $o_x = o_y = 6$ 。利用获得的矩阵 M ,将输入图像仿射变换到每一个同等大小的小型椭圆上,并计算相应加权的局部平均以及标准方差。值得注意的是,加权平均和标准方差的计算是执行在不同颜色通道的各个扇区上的。因此,椭圆内一个像素点的加权平均和标准方差为

$$m_k^i = \sum_{i=1}^3 \sum_{k=1}^4 (t_i \cdot s_k)$$

$$(s_k^i)^2 = \sum_{i=1}^3 \sum_{k=1}^4 ((t_i)^2 \cdot s_k - (m_k^i)^2)$$
(16)

式中, m_k^i 定义了加权局部平均值, S_k^i 定义了标准方差。 t_i 是椭圆中一个颜色通道上的任意像素点; s_k 是预处理生成的滤波器圆盘扇区,其处理过程类似椭圆,即先计算一个 2 维旋转变换矩阵,然后通过仿射变换到圆,细节就不重述了。

2) 对于图像中的每个像素,计算出其局部平均的加权和,定义为

$$O = \frac{\sum_{i=1}^3 \sum_{k=1}^4 s_k^i \cdot m_k^i}{\sum_{i=1}^3 \sum_{k=1}^4 s_k^i}$$
(17)

式中, $s_k^i = (\max(0.02, \|s_k^i\|))^{-8}$ 。然后,参照 Kyprianidis^[22],得知多尺度涉及两个方面:基于各向异性的计算,平滑结构张量和上采样的结构张量(前级)融合;以及基于局部方差的近似,上采样的滤波结果然后与当前金字塔级别的图像数据融合。注意的是,在实验中,将图像中每个像素点的权重系数 β 的值计算为

$$\beta = \frac{1}{4} \sum_{i=1}^3 \sum_{k=1}^4 \max(0.02, \|s_k^i\|) \times 0.02 \times 1.25 - 0.1$$
(18)

由于输入图像根据生成的掩模被划分成了显著部分和非显著部分,因此基于先前生成的图像显著性掩模值,记为 $maskV$ 来有效地控制参数 β 。当 $maskV$ 的值取 255 时,对应白色所覆盖的显著区域,此时 β 的值为 1,算法执行单一尺度的 Kuwahara 滤波运算。当 $maskV$ 的值取 0 时,即黑色所覆盖的非显著区域。 β 的值采用式(18)来计算得到,此时算法迭代地执行在金字塔上实现多尺度的 Kuwahara 滤波效果。

4 实验结果

实验使用 C++ 和 OpenCV API 实现,并采用 Openmp 进行多线程程序设计。实验中使用 NPR 框架(图1)对每幅图像进行了反复的实验比较,为了验证方法的有效性,定义了完全不同的显著性区域,例如,对于有着远景和近景的对比图,分别定义近景区域和远景区域为显著性进行测试,结果显示两种选择都能较清晰地凸出显著性区域而相对模糊非显著性区域,因此得到了预期的效果。

在图7中,图7(b)可以看出图像更多细节能够得到保留。图7(c)移去了图像大多数的细节以致结果图被大幅度的简化,但图像大部分的整体结构仍然可辨识。图7(d)根据显著性的定义,图像的细节被控制在图中的高亮部分。图7(f)是用户根据自己的目的和用途来实施笔划的具体操作,其中分割边缘框内笔划用来定义显著性的区域,边缘框外的笔划用来定义非显著性的区域,边缘框中的内边缘表示指向显著性区域一侧,外边缘表示指向非显著性区域一侧。图7(g)的白色区域就是用户自己所定义的显著性部分,黑色区域为相应的非显著性部分。

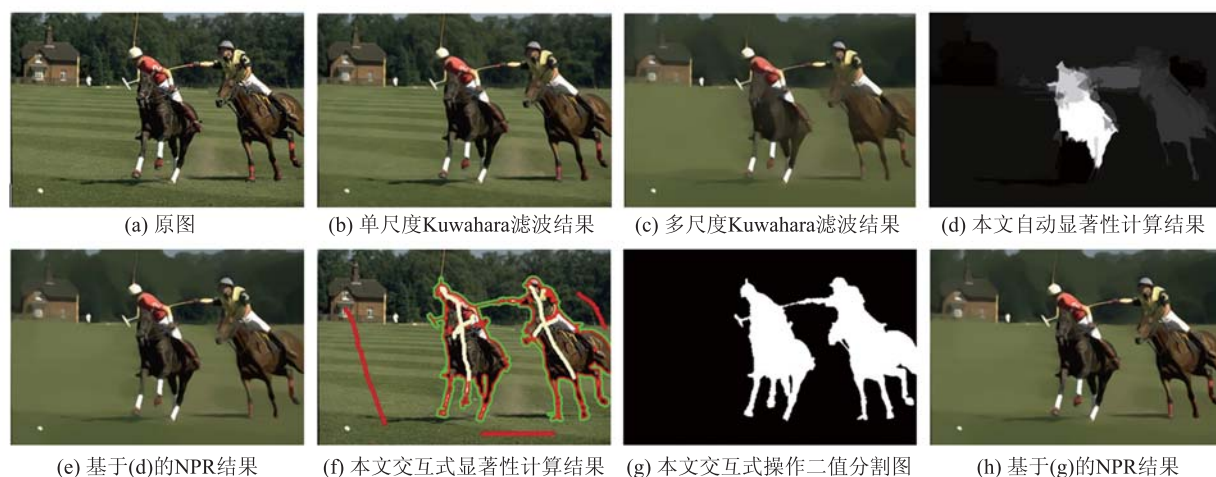


图7 对于单张图像的实验比较示例

Fig. 7 An experimental comparison of a single image

图像中的局部简化以两种方法予以呈现:一种是基于纯粹的低级特征,另一种是基于人工的自定义。图像细节的控制根据提交的显著性定义分别执行单尺度和多尺度的 Kuwahara 滤波来实现,从而生成强弱不同的抽象效果。

为了全面的测试本文方法的可靠性,还将得到的基于显著性的 NPR 效果图进行软量化并加入风格化的边缘与已有的工作做比较。有趣地是,与文献[2]所实现的效果相比,本文使用的滤波器提供了更好的边缘保持,使得在高对比度的区域适当的除去细节而在低对比度区域保护结构的边缘,同时避免了阈值使用而导致的相关信息被抹掉的缺陷(图8)。图8(c)为本文方法基于视觉感知抽象并加入了后期的必要处理生成的最终效果图。定义近处的人物为显著性区域,其余为非显著性区域。可以看出,不管是图像的显著性部分还是非显著性部分,本文方法更好地保持了原图的纹理和结构,而

且对图像光线的处理也更加得准确。因此,本文方法能够较好地体现图像中的深度,符合基于视觉感知的抽象要求。由于时间的限制,后续的工作中将加入用户的调查,对不同算法的抽象化效果进行主观上的评判。



图8 基于显著度信息的抽象化效果比较

Fig. 8 Compare the image abstract results based on saliency

图9呈现了一个视频抽象的例子。在这个示例中,定义了近处的女孩为显著性目标区域,其余为非显著性的背景区域。在视频抽象化上,考虑到连续两帧之间的运动不会太快,因此依据前一帧获取的显著

性信息作为初始估计,延用到以后的帧;同时,指定若干个可延用的帧数,当达到指定的最后一帧时,重新获取下一帧的显著性信息作为初始估计。这样,在保证视频帧间连贯性的同时,提高了时间效率。



(a) 本文方法处理后连续帧



(b) 原视频连续帧

图 9 基于视觉感知的视频风格化抽象示例

Fig. 9 Abstraction example of video based on visual perception

5 结 论

提出了一个基于视觉感知的完整图像及视频抽象化框架。其中,视觉显著性信息除了可自动地由低级特征计算得到外,也可根据用户的需要交互式地获取。在视觉感知信息的指引下,相应地执行各向异性的单、多尺度的 Kuwahara 滤波处理,从而形成预期结果。提出的技术能够灵活地指定显著性目标区域,同时也确保了在视频抽象时良好的帧间连贯性。

在许多 NPR 风格中,抽象并不意味着绝对地携带图像中的任何信息。带有目的性地对图像抽象能够使得观察者对图像中的增强信息的理解,因此这种风格化的抽象方式是有意义的。

参考文献 (References)

- [1] DeCarlo D, Santella A. Stylization and abstraction of photographs [J]. ACM Transactions on Graphics, 2002, 21(3): 769-776. [DOI: 10.1145/566654.566650]
- [2] 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. [DOI: 10.1109/34.730558]
- [3] Shi M L, Yang L, Peng G Q, et al. A content-aware image resizing method with prominent object size adjusted [C]// The 17th ACM Symposium on VRST. New York, USA: Association for Computing Machinery, 2010: 175-176.
- [4] Shi M L, Xu D. Survey on content-aware image resizing techniques [J]. Journal of Image and Graphics, 2012, 17(2): 157-324. [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): 157-324.]
- [5] Wen F, Luan Q, Liang L, et al. Color sketch generation [C]// The 4th International Symposium on Non-Photorealistic Animation and Rendering. New York, USA: ACM, 2006: 47-54. [DOI: 10.1145/1124728.1124737]
- [6] Zhao M, Zhu S C. Sisley the Abstract Painter [C]// The 8th International Symposium on Non-Photorealistic Animation and Rendering. New York, USA: ACM, 2010: 99-107. [DOI: 10.1145/1809939.1809951]
- [7] Cheng M M, Zhang G X, Mitra N J, et al. Global contrast based salient region detection [C]// Proceedings of IEEE CVPR 2011. Washington DC, USA: IEEE Computer Society, 2011: 409-416.
- [8] Tomasi C, Manduchi R. Bilateral filtering for gray and color images [C]// Proceedings of the 6th International Conference on Computer Vision. Bombay India: IEEE Computer Society, 1998: 839-846.
- [9] Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603-619. [DOI: 10.1109/34.1000236]
- [10] Kuwahara M, Hachimura K, Eiho S, et al. Processing of rian-giocardiographic images [C]// Digital Processing of Biomedical Images. New York, USA: Plenum, 1976: 187-202.
- [11] Kyprianidis J E, Kang H, Döllner J. Image and video abstraction by anisotropic kuwahara filtering [J]. Computer Graphics Forum, 2009, 28(7): 1955-1963. [DOI: 10.1111/j.1467-8659.2009.01574.x]
- [12] Papari G, Petkov N, Campisi P. Artistic edge and corner enhancing smoothing [J]. IEEE Transactions on Image Processing, 2007, 16(10): 2449-2462. [DOI: 10.1109/TIP.2007.903912]
- [13] Winnemöller H, Olsen S C, Gooch B. Real-time video abstraction [J]. ACM Transactions on Graphics, 2006, 25(3): 1221-1226. [DOI: 10.1145/1141911.1142018]
- [14] Zhang S H, Li X Y, Hu S M, et al. Online video stream abstraction and stylization [J]. IEEE Transactions on Multimedia, 2011, 13(6): 1286-1294. [DOI: 10.1109/TMM.2011.2165052]
- [15] Hertzmann A, Perlin K. Painterly rendering for video and interaction [C]// The 1st International Symposium on Non-Photorealistic Animation and Rendering. New York, USA: ACM, 2000: 7-12.
- [16] Lin L, Zeng K, Lv H, et al. Painterly animation using video semantics and feature correspondence [C]// The 8th International Symposium on Non-Photorealistic Animation and Rendering. New York, USA: ACM, 2010: 73-80.
- [17] Jiang H Z, Wang J D, Yuan Z J, et al. Automatic salient object

- segmentation based on context and shape prior [C]// Proceedings of the 22nd British Machine Vision Conference. Dundee, Britain: BMVA Press, 2011:110.1-110.12.
- [18] Felzenszwalb P F, Huttenlocher D P. Efficient graph-based image segmentation [J]. International Journal of Computer Vision, 2004, 59 (2): 167-181. [DOI: 10.1023/B:VISI.0000022288.19776.77]
- [19] Gulshan V, Rother C, Criminisi A, et al. Geodesic Star Convexity for Interactive Image Segmentation [C]//Proceedings of IEEE CVPR 2010. San Francisco, USA: Computer Society, 2010: 3129-3136
- [20] Veksler O. Star shape prior for graph-cut image segmentation [C]// Proceedings of the 10th European Conference on Computer Vision: Part III. Berlin, Germany: Springer-Verlag, 2008: 454-467.
- [21] Boykov Y, Jolly M P. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images [C]// Proceedings of International Conference on Computer Vision. Vancouver, Canada: IEEE Computer Society, 2001:105-112.
- [22] Kyprianidis J E. Image and video abstraction by multi-scale anisotropic kuwahara filtering [C]//The ACM Eurographics Symposium on Non-Photorealistic Animation and Rendering. New York, USA: ACM, 2011: 55-64.
- [23] Rother C, Kolmogorov V, Blake A. "Grabcut"-interactive foreground extraction using iterated graph cuts [J]. ACM Transactions on Graphics, 2004;23 (3):309-314. [DOI: 10.1049/el.2009.2014]
- [24] Van den Boomgaard R. Decomposition of the kuwahara-nagao operator in terms of linear smoothing and morphological sharpening [C]// The 6th International Symposium on Mathematical Morphology. Berlin, Germany: ACM, 2002:283-292.
- [25] Huang H, Fu T N, Li C F. Painterly rendering with content-dependent natural paint strokes [J]. The Visual Computer: International Journal of Computer Graphics, 2011, 27 (9): 861-871. [DOI: 10.1007/s00371-011-0596-5]
- [26] Zhang S H, Chen T, Zhang Y F, et al. Vectorizing cartoon animations [J]. IEEE Transactions on Visualization and Computer Graphics, 2009, 15 (4): 618-629. [DOI: 10.1109/TVCG.2009.9]