

基于潜在半径优化策略的数字图像抠图算法

赵海, 雷凯茹, 朱宏博, 朴春鹤
(东北大学信息科学与工程学院, 辽宁沈阳 110819)

摘要: 针对目标图像利用导向滤波算法进行图像抠图时参数需要根据不同图像进行人为设定的问题, 提出了一种利用隐藏的支持向量机 LSVM(latent support vector machine) 自动设定参数的潜在半径优化的数字图像抠图算法. 该方法首先是应用 LSVM 潜在性的思想, 利用已知数据库模板训练输入目标图像的样本集, 再利用样本集将导向图像和二值图像以不同半径进行分块并进行判定, 确定半径值, 从而能够自动产生合理参数. 最后利用导向滤波器对图像进行抠图, 从而优化抠图算法, 最终实现抠图算法的智能化和灵活化.

关键词: 导向滤波器; 图像抠图; LSVM; 目标图像; 训练样本集

中图分类号: TP 391 文献标志码: A 文章编号: 1005-3026(2015)11-1553-05

A Digital Image Matting Algorithm Based on the Optimization Strategy of Latent Radius

ZHAO Hai, LEI Kai-ru, ZHU Hong-bo, PIAO Chun-he

(School of Information Science & Engineering, Northeastern University, Shenyang 110819, China. Corresponding author: LEI Kai-ru, E-mail: leikairu@163.com)

Abstract: Due to parameters is needed to be set depending on the issue of human goal-oriented to filter matting algorithms, a classifier was proposed based on component detection called LSVM (latent support vector machine). The parameters of guided radius could be automatically set so as to optimize algorithm. Firstly, for determining the radius, a potential idea of LSVM was applied which using the known database templates to train the input samples of the target image, then the sample set of the guided image and the binary image were divided into blocks with different radius which could automatically generate the reasonable parameters and optimize algorithm. At last, guided filter was used to advance matting and the intelligent and flexible of the matting algorithm was achieved.

Key words: guided filter; image matting; LSVM(latent support vector machine); target image; training sample set

数字图像抠图技术是计算机视觉和图像处理中重点研究的热门课题,也是一个难点. 数字图像抠图(matting)^[1]所研究的是将一幅图像中的前景和背景信息分离开,把人们感兴趣的区域提取出来,以便更换背景或其他操作.

目前数字图像抠图技术虽然取得了一定的成功,但仍不能满足在图像处理实践中对于抠图结果精准性、灵活性的要求. Smith 与 Blinn^[2]在 1996 年首次提出了蓝屏抠图技术,该技术被广泛应用于动态影像拍摄及制作电视节目,但是该方

法需要固定的背景颜色. 由文献[3]中提出的式(11)的线性表达式可以看出,在 RGB 色彩空间中,求解 α , F 和 B 的过程实际上是 1 个方程求解 3 个未知数的过程. 显而易见,由于不存在一个确定的解,该求解过程是一个病态的问题. 因此, Ruzon 与 Tomasi^[4]引入了 α 通道的概念,通过计算周围前景或背景像素颜色分布来估计 α 的值. 但是该方法对于样本点分析较为粗略. Levin 等提出了 Closed form^[5] 抠图方法,该算法通过对局部前景 F 和背景 B 平滑性假设求解抠图结果,但

是存在漏算和误算的问题。

数字图像抠图也经常用到图像滤波器,如高斯滤波器、双边滤波器及导向滤波器等来完成,并且滤波器还具有边缘检测、特征提取^[6]和去噪等功能,进一步辅助精确抠图。但是对于目标图像,大多数滤波器算法繁琐,并具有一定指向性,即根据不同的图像进行手动设定参数^[7],降低了算法的灵活性。

针对以上问题,本文提出了一种充分利用图像隐藏信息进行模型学习的基于隐藏的支持向量机(latent support vector machine, LSVM)^[3]的潜在半径优化策略抠图算法。该算法的基本思想是,利用已知的训练部件模型对感兴趣目标的边界进行矩形标注,从而来学习模型,精确半径值以达到自动产生合理参数,最终完成利用导向滤波器(guided filter)^[8]进行图像精确抠图的目的。

1 基于目标图像的导向滤波算法

为了通过求解显而易见的图像约束条件以得到精确输出图像抠图结果求解 α 通道图像的目的,本文采用导向滤波算法很好地解决了这个问题。导向滤波算法的关键思想是在导向目标图像、二值输入图像和输出抠图图像间建立很好的联系^[9]。其中,导向图像是一个具有明确指向性的目标图像,它可以是输入图像自身或者是另外一张不同的独立图像。二值图像是目标图像的 Trimap 模板图像^[10]。

基于目标图像的导向滤波算法就是建立在导向目标图像、二值输入图像和输出抠图图像之间的一个滤波器。系统示意图如图 1 所示。

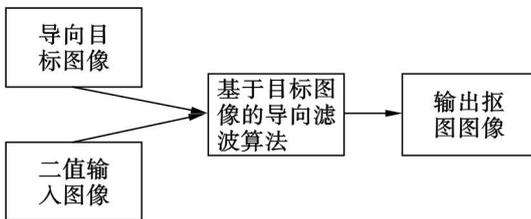


图 1 基于目标图像的导向滤波示意图
Fig. 1 The explanatory views of guided image filtering based on the target image

文献[8]对导向滤波器的参数计算进行了详尽的解释与说明。对于导向滤波器来说其关键的假设是这个导向滤波器在导向图像 I 和滤波输出 q 之间是一个局部线性模型。首先,从整体角度上看,假设导向图像为 I ,输入图像为 p ,输出图像为 q ,假设 q 是 I 内中心在像素点 k 的窗口 w_k 的线

性转换:

$$q_i = a_k I_i + b_k, \forall i \in w_k. \quad (1)$$

其中:半径 r 表示窗口; (a_k, b_k) 是假定同 w_k 相同的线性系数。由式(1)可知 $\nabla q = a \nabla I$,即这个局部线性模型确定只要 I 有一个边缘,那么 q 就有一个边缘。

其次,为了寻求一个解决办法能够最小化 q 和 p 之间的差异,同时保持线性模型(1),最小化窗口 w_k 的代价函数:

$$E(a_k, b_k) = \sum_{i \in w_k} ((a_k I_i + b_k - p_i)^2 + \lambda a_k^2). \quad (2)$$

这里, λ 是 a_k 的正则化参数。整理线性山脊回归模型式(2)可以得到:

$$a_k = \frac{\frac{1}{|w|} \sum_{i \in w_k} I_i p_i - \overline{\mu_k} \overline{p_k}}{\sigma_k^2 + \lambda}, \quad (3)$$

$$b_k = \overline{p_k} - a_k \overline{\mu_k}. \quad (4)$$

其中: μ_k 和 σ_k^2 分别是导向图像 I 的滤波窗口 w_k 的平均值和方差; $|w|$ 是 w_k 中像素个数; $\overline{p_k} = \frac{1}{|w|} \sum_{i \in w_k} p_i$, 是 p 在 w_k 上的平均值。

然而,由于一个像素 i 与所有覆盖 i 的重叠窗口 w_k 相关,所以当用不同的窗口计算时,方程 q_i 的值是不同的,因此一个简单的策略是平均所有 q_i 的可能值。于是式(1)进一步改写为

$$q_i = \frac{1}{|w|} \sum_{kl \in w_k} (a_k I_i + b_k). \quad (5)$$

进一步整理式(5)可得

$$q_i = \overline{a_i} I_i + \overline{b_i}. \quad (6)$$

其中: $\overline{a_i} = \frac{1}{|w|} \sum_{i \in w_k} a_k, \overline{b_i} = \frac{1}{|w|} \sum_{i \in w_k} b_k$.

通过上述内容可以看出,基于导向滤波器的数字图像抠图算法针对同一目标图像的不同部位来说,参数的设定是单一不变的,这就会导致抠图结果在准确性和智能性上稍显不足。因此若能针对同一物体的不同部位进行不同的参数设定,将大大提升抠图的精确性和灵活性,扩大应用范围。

2 基于 LSVM 的滤波算法参数设定

由上述内容可知,导向滤波器的窗口半径 r 和平滑系数 λ 是基于图像导向滤波算法的主要参数,它们影响着图像抠图的输出效果和算法的执行速度。但是,对于一幅导向图像来说,滤波器对于图像所有边缘部分的窗口半径 r 和平滑系数 λ 是固定的。然而对于图像边缘,不同的部位实际边缘半径 r 是不同的,因此本文提出了隐藏的支

持向量机 (LSVM) 的思想。

第一步：训练样本模型。利用对 PASCAL 数据库中的模板信息进行精确训练来学习模型。这样，样本就由带标注边界矩形的感兴趣目标组成。选取所需训练样本集 D 定义如下：

$$D = (x_1, y_1, x_2, y_2, \dots, x_n, y_n), y_i \in (+1, -1).$$

第二步：对导向图像和二值图像进行分块。目标图像通过与训练样本集比对，就可以沿着目标边界将图像沿边缘分割成不同的部分，形成边界矩形大小不同的图像模块。

第三步：预判边缘半径 r 的大小。利用图像模块的不同，依照初步的设定边界矩形与边缘半径的比例，求解每个模块的边缘半径 r 。

第四步：滤波并利用 LSVM 判定第三步中求解的边缘半径 r 是否正确。其思想是

1) 当滤波输出结果存在 $q_i = 0$ ，则边缘半径 r 包含了完整前景信息；

2) 当滤波输出结果不存在 $q_i = 0$ ，则边缘半径 r 取值不当，需要调整 r 的值。然后再进行判断，直至满足 1)。

这个判断过程就利用了 LSVM 的思想。首先假设前景用 $(i + 2)$ 元数组 $(F_0, p_1, \dots, p_i, c)$ 定义含有 i 个子模板的目标模型。其中： F_0 表示导向滤波器； p_i 是第 i 个图像模块的模型； c 是偏移量。

在图像特征金字塔中，假定模型中每个滤波器的位置为 $z = (p_1, \dots, p_i)$ ，其中 $p_n = (x_n, y_n, s_n)$ 表示第 n 个滤波器在特征金字塔第 s_n 层的位置，对应的响应是每个滤波器在各自位置上响应再加上偏移量 c 。

$$\text{Filter}(p_i) = \sum_{i=0}^n F_i \varphi(H, z) + c. \quad (7)$$

其中： H 表示特征金字塔； $\varphi(H, z)$ 表示以行优先顺序并置 H 中以 p_i 为左上角的 $w_i \times h_i$ 子窗的特征向量。

通过滤波结果式 (6) 可得出 $\varphi(H, z)$ 等价公式为

$$I_k = \alpha_k F_k + (1 - \alpha_k) \beta_k. \quad (8)$$

其中： I_k 是图像颜色； F_k 是前景图像颜色； β_k 是背景图像颜色； α_k 是透明度，其区间为 $[0, 1]$ ，其中 $\alpha_k = 0$ 或者 1 称作绝对前景或者绝对背景，否则称 k 点为混合点。 α 介于 $0 \sim 1$ 之间时像素颜色往往出现混合，通常出现在图像边缘、毛发等区域^[11]。

于是将式 (7) 简化，用模型参数向量 γ 和向量 $\varphi(H, z)$ 的点积表示为 $\gamma \cdot \varphi(H, z)$ ，

$$\gamma = (\overline{a_1}, \dots, \overline{a_n}, c). \quad (9)$$

$$\varphi(H, z) = (\varphi(H_1, p_1) + \frac{\overline{b_1}}{a_1}, \dots, \varphi(H_n, p_n) + \frac{\overline{b_n}}{a_n}, 1). \quad (10)$$

这样就可以考虑将模型与分类器结合进行学习得到模型参数，从而确定边缘半径 r 的取值。

图 2 给出了本文针对导向图像滤波算法改进的基于潜在半径优化算法的总流程图。

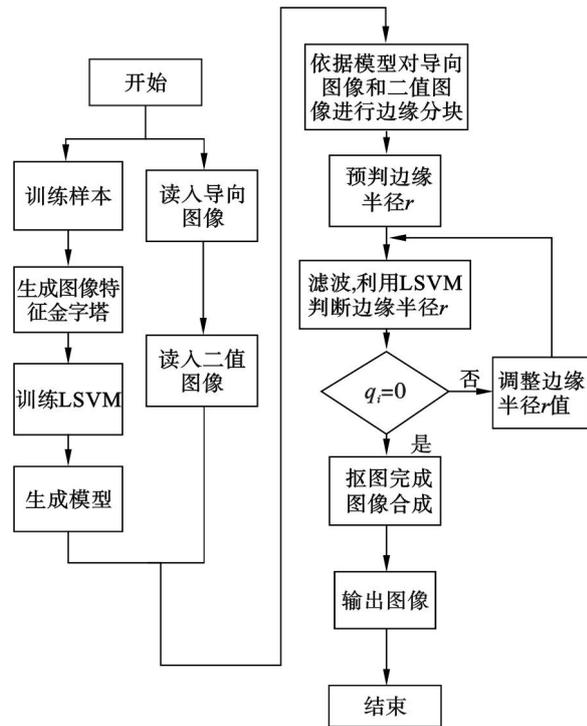


图 2 基于潜在半径优化策略的数字图像抠图算法流程图

Fig. 2 The explanatory views of matting algorithm based on the optimization strategy of latent radius

3 实 验

在抠图问题中，将图像假设成为前景颜色与背景颜色依照图像内容表达出的线性组合，其表达式见式 (8)。

根据式 (8) 可以将式 (6) 转换成 α 通道的形式：

$$\alpha_i = \frac{1}{F_i - B_i} I_i + \frac{-B_i}{F_i - B_i}. \quad (11)$$

令 $a = \frac{1}{F_i - B_i}$ ， $b = \frac{-B_i}{F_i - B_i}$ ，即可得出滤波输出的 α 值：

$$\alpha_i = a I_i + b, \forall i \in w_k. \quad (12)$$

数字图像抠图技术的关键就是求解 α 的值，导向滤波器可以很好地解决 α 的计算。

图 3 显示了 4 类图像分别采用基于 LSVM

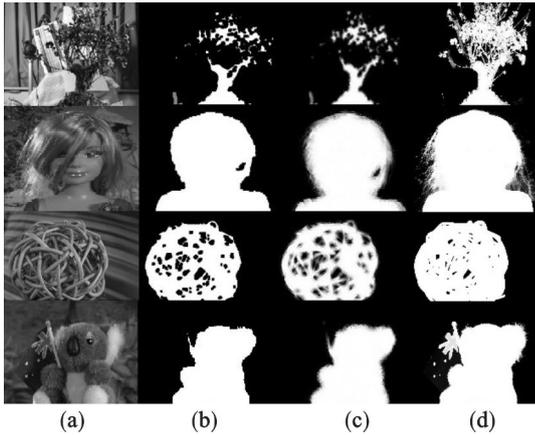


图 3 基于潜在半径优化策略的抠图算法与手动设定参数的抠图算法对比图

Fig. 3 The contrast of matting algorithm between optimization strategy of latent radius and manual set

(a)—原图; (b)—二值图像;
(c)—导向滤波器算法; (d)—LSVM.

算法的抠图算法与手动设定参数的抠图算法对比实验结果如下:

1) 图 3a 是输入的尺寸为 647 × 800 的导向目标图像;

2) 图 3b 是输入二值图像,即对前景图像与背景图像进行首次粗分割的结果;

3) 图 3c 是运用导向滤波器算法,并将图 3a 与图 3b 作为输入图像得到的抠图结果;

4) 图 3d 采用了本文所提出的利用 LSVM 自动设定参数的基于潜在半径优化策略数字抠图算法.该方法同样采用图 3a 和图 3b 作为输入图像.

通过以上结果对比可以看出,与基于手动划分参数的方法相比,本文提出的利用 LSVM 自动设定参数的基于潜在半径优化策略数字抠图算法能够使输出图像的细节部分更加丰富,甚至相对于手动划分,在某些局部的处理与跟踪效果更加生动、明显.同时本文提出的方法省去了原始算法中设定参数的步骤,增强了算法的灵活性和智能性.

本实验同时也将本文提出的抠图算法同几种经典的抠图算法进行了比对,如图 4 所示.图 4 显示了 4 类图像,分别采用以下 5 种抠图算法得到的 α 进行评价与合并.这里选取 2 幅原始图像,以及通过 5 种算法得到的 5 幅 α 结果图像 $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(5)}$.这 5 种抠图算法分别为 ①Spectral matting

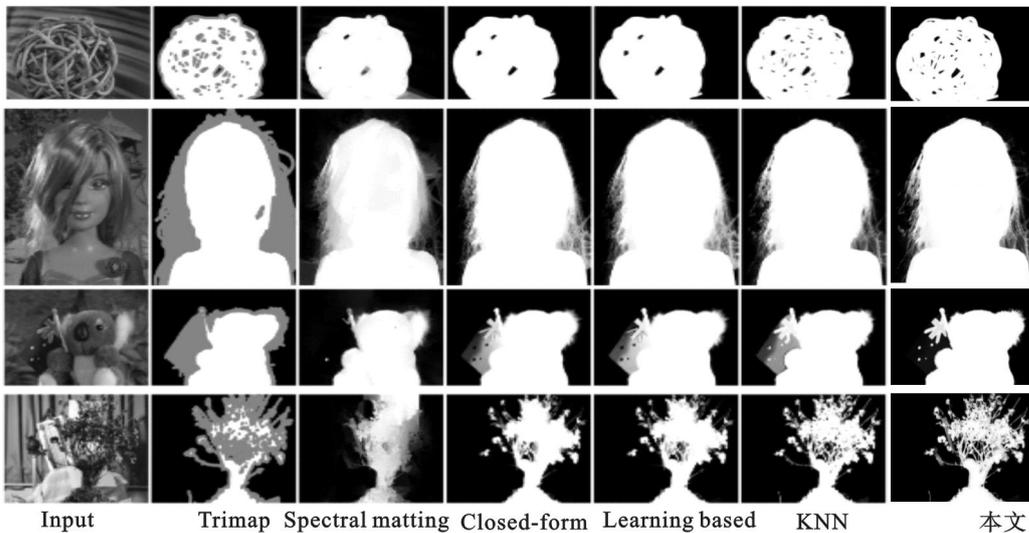


图 4 不同抠图算法的 α 结果图
Fig. 4 different algorithms to the results of alpha matting

算法;②Closed - form 算法;③ Learning based 算法;④KNN 算法;⑤ Optimization of Guided Filter Radius 算法.每幅图像通过以上算法计算得到一幅该算法下的 α 结果图.在这里统一采用 Trimap 对未知区域进行确定,然后运行抠图算法.对比以上 5 种抠图算法的实验结果不难发现: Spectral Matting 算法对于图像细节处理能力不足,主要是由于待处理图像的光照情况并不是自然光均匀照射直接导致前景与背景图像边界不明显,从而严

重地影响了抠图效果;对于 Closed - form^[4]算法,其效果已经明显优于之前的算法,但是其缺点在于对闭合区域的处理中,若前景与背景存在多个层次的变化,则不能有效实施抠图操作; Learning Based 算法与 KNN 算法相较前两者,抠图效果明显提高,但是对于前景、背景交界处颜色比较相近的图像来说,抠图效果真实度不够准确;相比较而言,由于利用 LSVM 的潜在半径优化算法进行抠图,能够很好地利用已知 PASCAL 数据库中图像

的特征,从而很好地对未知待抠图区域的系数进行预判.这种算法对于前景背景的层次多变、光照影响、边界颜色相近等问题的影响也能够很好地改进,更接近抠图效果的真实值.

4 结 语

本文提出了一种基于潜在半径优化策略的数字图像抠图技术.该算法改进了滤波器抠图的系数单一的缺点,能够确定抠图时不同目标图像边缘未知区域的不同系数,提升了算法的智能性和灵活性.结果表明:该方法在优化图像抠图技术的 α 通道值方面效果明显.

参考文献:

- [1] He K, Sun J, Tang X. Guided image filtering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6) :1397 – 1409.
- [2] Smith A R, Blinn J F. Blue screen matting [C] // Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques. New York: ACM, 1996: 259 – 268.
- [3] Yu C, Joachims T. Learning structural SVMs with latent variables [C] // Proceedings of 26th International

Conference on Machine Learning. Montreal, 2009: 1169 – 1176.

- [4] Ruzon M A, Tomasi C. Alpha estimation in natural images [C] // Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2000: 18 – 25.
- [5] Levin A, Lischinski D, Weiss Y. A closed-form solution to natural image matting [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2008, 30(2) :228 – 242.
- [6] Heilbronner R, Barrett S. Image analysis in earth sciences [M]. Berlin: Springer, 2014: 31 – 57.
- [7] Lin S Y, Pan R F, Du H, et al. A survey digital matting [J]. *Journal of Computer Design & Computer Graphic's*, 2007, 19(4) :473 – 479.
- [8] He S J, Tang X. Guided image filtering [M]. Berlin: Computer Vision – ECCV, 2010.
- [9] Li Y, Sun J, Shum H. Video object cut and paste [J]. *ACM Transactions on Graphics*, 2005, 24(3) :595 – 600.
- [10] 林生佑, 潘瑞芳, 杜辉, 等. 数字抠图技术综述 [J]. 计算机辅助设计与图形学报, 2007, 19(4) :473 – 479.
(Lin Sheng-you, Pan Rui-fang, Du Hui, et al. A survey on digital matting [J]. *Journal of Computer-aided Design & Computer Graphics*, 2007, 19(4) :473 – 479.)
- [11] Levin A, Lischinski D, Weirs Y. A closed form solution to natural image matting [C] // Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006: 61 – 68.

(上接第 1552 页)

参考文献:

- [1] 3GPP. TS 23. 228 IP multimedia subsystem (IMS) stage 2 [EB/OL]. (2014-09-22) [2014-10-09]. <http://www.3gpp.org/DynaReport/23228.htm>.
- [2] 3GPP. TS 23. 207 End-to-end quality of service (QoS) concept and architecture [EB/OL]. (2012-09-14) [2014-10-09]. <http://www.3gpp.org/DynaReport/23207.htm>.
- [3] 3GPP. TS 23. 203 Policy and charging control architecture [EB/OL]. (2014-06-24) [2014-10-09]. <http://www.3gpp.org/DynaReport/23203.htm>.
- [4] Rosenberg J, Schulzrinne H, Camarillo G, et al. IETF RFC 3261 SIP: session initiation protocol [EB/OL]. (2002-06-12) [2014-09-18]. <https://www.ietf.org/rfc/rfc3261.txt>.
- [5] Akella A, Maggs B, Seshan S, et al. A measurement-based analysis of multihoming [C] // Proceedings of ACM SIGCOMM. Karlsruhe: ACM, 2003 :353 – 364.
- [6] Thaler D, Hopps C. IETF RFC 2991 multipath issues in unicast and multicast next-hop selection [EB/OL]. (2000-11-01)

[2014-08-15]. <https://www.ietf.org/rfc/rfc2991.txt>.

- [7] Ford A, Raiciu C, Handley M, et al. IETF RFC 6182 architectural guidelines for multipath TCP development [EB/OL]. (2011-03-12) [2014-08-19]. <https://www.ietf.org/rfc/rfc6182.txt>.
- [8] Sarkar D, Paul S. Architecture, implementation, and evaluation of cmpTCP westwood [C] // Proceedings of IEEE GLOBECOM. San Francisco, 2006: 1 – 5.
- [9] Lei W M, Zhang W, Liu S. A framework of multipath transport system based on application-level relay (MPTS-AR) [EB/OL]. (2014-07-25) [2014-09-03]. <https://tools.ietf.org/html/draft-leiwm-tsvwg-mpts-ar-02>.
- [10] Lei W M, Zhang W, Liu S. Multipath real-time transport protocol based on application-level relay (MPRTP-AR) [EB/OL]. (2014-07-25) [2014-09-03]. <https://tools.ietf.org/html/draft-leiwm-avtcore-mprtp-ar-02>.
- [11] Zhang W, Lei W M, Liu S, et al. A general framework of multipath transport system based on application-level relay [J]. *Computer Communications*, 2014, 51 :70 – 80.