

一种由粗到细的头发分割方法^{*}

王丹^{1,2}, 山世光¹, 张洪明³, 曾炜³, 陈熙霖¹

¹(中国科学院 计算技术研究所 智能信息处理重点实验室, 北京 100190)

²(中国科学院大学, 北京 100049)

³(NEC 中国研究院, 北京 100084)

通讯作者: 山世光, E-mail: sgshan@ict.ac.cn

摘要: 从图像中提取出头发区域, 能够为头发分析、发型趋势预测等任务提供有利的线索. 但是, 头发的类内模式非常复杂, 并且它与其他物体类间也常因光照复杂、表观特征相似等因素而难以分离. 因此, 头发分割是一个非常具有挑战性的问题. 为了一定程度地解决这些问题, 提出了一种由粗到细的头发分割方法. 首先, 该方法利用最新提出的利用视点进行主动分割(active segmentation with fixation, 简称 ASF)的方法, 粗略提取头发分割的候选范围, 保证头发区域的高召回率(准确率也许较低), 并由此排除大部分与头发区域难以分离的背景区域; 然后, 利用特定于当前图像的头发表类信息, 使用图割(graph cuts, 简称 GC)法在限定的范围内进行更加精细的分割. 具体地, 采用均值漂移(mean shift, 简称 MS)方法对输入图像进行区域的过分割; 然后, 利用贝叶斯方法选择一些可靠的、有较大概率属于头发或背景的“种子区域”, 针对头发和背景的种子区域, 采用支持向量机(support vector machine, 简称 SVM)在线学习头发和背景的分类器, 并将其用于预测每个像素或区域属于头发或背景的概率; 最后, 将得到的概率用以 Graph Cuts 的初始化, 求解得到最终的头发分割结果. 实验结果表明, 所提出的头发分割方法能够超越当前提出的头发分割方法. 为了验证方法的可推广性, 对其进行了一定扩展, 并在马、汽车、飞机这 3 个类别的公开数据库上作了评测, 取得了较好的性能.

关键词: 头发分割; 由粗到细; 图割; 支持向量机

中图法分类号: TP391 文献标识码: A

中文引用格式: 王丹, 山世光, 张洪明, 曾炜, 陈熙霖. 一种由粗到细的头发分割方法. 软件学报, 2013, 24(10): 2391-2404. <http://www.jos.org.cn/1000-9825/4423.htm>

英文引用格式: Wang D, Shan SG, Zhang HM, Zeng W, Chen XL. Coarse-to-Fine hair segmentation method. Ruan Jian Xue Bao/Journal of Software, 2013, 24(10): 2391-2404 (in Chinese). <http://www.jos.org.cn/1000-9825/4423.htm>

Coarse-to-Fine Hair Segmentation Method

WANG Dan^{1,2}, SHAN Shi-Guang¹, ZHANG Hong-Ming³, ZENG Wei³, CHEN Xi-Lin¹

¹(Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

³(NEC Labs China, Beijing 100084, China)

Corresponding author: SHAN Shi-Guang, E-mail: sgshan@ict.ac.cn

Abstract: Segmenting hair regions from human images facilitates many tasks like hair analysis and hair style trends forecast. However, hair segmentation is quite challenging due to large within-class pattern diversity and between-class confusion resulted from complex illumination and similar appearance. To solve these problems to some extent, this paper proposes a novel coarse-to-fine hair segmentation

* 基金项目: 国家重点基础研究发展计划(973)(2009CB320902); 国家自然科学基金(61025010, 61173065); 北京市自然科学基金(4111003)

收稿时间: 2012-08-16; 修改时间: 2012-12-27, 2013-03-18; 定稿时间: 2013-05-07

method. Firstly, the recently published "active segmentation with fixation (ASF)" is used to coarsely define a candidate region with high-recall (but possibly low-precision) of hair pixels and exclude considerable part of the backgrounds which are easily confused with hair. Then the graph cuts (GC) method is applied to the candidate regions to perform more precise segmentation, by incorporating image-specific hair information. Specifically, Bayesian method is employed to select some reliable hair and background regions (seeds) among the ones over-segmented by mean shift. SVM classifier is then learnt online from these seeds and explored to predict hair/background likelihood probability, which is used as an initialization for performing GC algorithm. Experimental results demonstrate the approach outperforms existing hair segmentation methods. To validate the generality, the paper extends the method and achieves good results on the public databases of horse, car and aeroplane classes.

Key words: hair segmentation; coarse-to-fine; graph cut; SVM (support vector machine)

由于头发关乎个人形象,头发仿真、头发编辑等问题备受研究者关注^[1-4].研究者们发现,头发在人的身份识别^[5]、性别识别^[6]方面起着至关重要的作用.由于人们一般不会频繁地更换发型,头发分割有利于身份识别,而男女发型差异使其可辅助性别识别.为了完成以上这些应用任务,从输入图像中分割出头发区域用于后续处理,是相当有价值的工作.但在大多数文献中,头发区域被假设为已分割好或者已由手动标注出完整区域^[7,8],因此,自动的头发分割仍然是一个开放式问题.头发分割问题之所以很有挑战性,是因为其类内模式复杂,且很难与相似背景(如深色树枝、深色衣服等)区分开来.目前,研究头发分割的文献还相对较少,在较早的 Liu 等人^[9]的工作中,他们利用纹理和几何特征,将像素分为头发和非头发两类;在文献[5]中,Yacoob 和 Davis 利用颜色模型进行头发分割,但是只对结果进行了主观评价,缺乏定量分析;在另一项较新的研究工作^[10]中,Rousset 和 Coulon 基于频域和颜色信息,应用 matting 方法进行头发分割;文献[11]也采用颜色模型进行头发分割,但缺乏细致的设计.值得注意的是,文献[5]和文献[10]的方法都需要首先选出足量、可靠的头发区域作为种子,以学习准确的头发模型.而既保证种子数量,又保证准确度,也是很难的任务.最近,Wang 等人^[12]和 Lee 等人^[13]提出的头发分割方法获得了令人满意的性能.Wang 等人^[12]采用贝叶斯方法进行头发的种子区域选择和头发分割,但忽略了背景信息的利用;Lee 等人^[13]利用 Graph Cuts 算法进行头发分割,但在 GC 的优化中有太多的复杂背景对结果产生了不可预测的影响,使得 GC 优化计算的结果与用户期望的结果之间存在很大差距.综上,复杂场景下的头发分割问题并没有得到很好的解决.

本文方法的早期版本发表在 FG2011^[14]上,相比于早期版本,本文将特定于头发分割的方法进行了一定的扩展,推广到解决其他物体类别的分割问题.方法扩展主要是采用了多视点的粗分割方法,使方法对于表观模式复杂的物体仍能保持较高的召回率;实验上的扩展,主要是新增 3 个类别在公开数据库的大量评测、比较和分析.

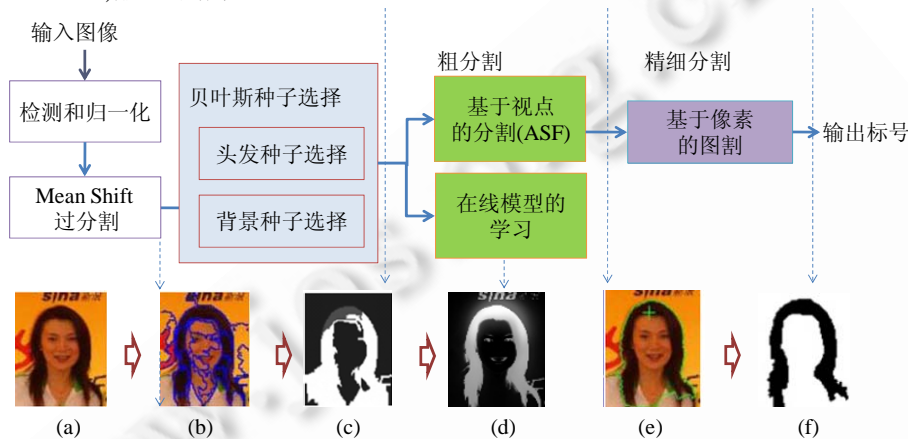
基于视点的主动分割方法由 Mishra^[15]提出,它利用物体的视点信息,自动提取物体或一致性区域的闭合轮廓,适用于凸形状物体的轮廓提取.根据这一性质,我们利用 ASF 排除大部分难与前景分离的背景区域,提供包含头发的“粗略”候选区域,提取出的候选区域常常是由全部头发像素、全部或部分面部像素以及一些周边像素共同组合成的凸形状区域.而后,在此候选区域内,应用 Graph Cuts 排除错误接受的像素,达到精细的头发分割结果.在运行 GC 之前,需要从图像中选择足量、准确的种子区域,用以对头发和背景建模.与 Wang 等人^[12]提出的方法相似,本文也采用贝叶斯方法进行种子选择.不同的是,为了保证选到更多准确的种子像素,采用了基于区域,而非基于像素的策略.得到种子区域后,采用判别式方法——支持向量机(support vector machine,简称 SVM)^[16]来学习头发和背景模型.之所以采用 SVM,是因为判别式方法往往比产生式(比如高斯混合模型)更加关注类别间的差异,这更符合分离头发和背景的目标.学习到的 SVM 分类器用以预测每个像素或者区域属于头发或者背景的概率,该概率被作为后端 Graph Cuts 的初始化.本文方法与文献[13]中的方法有 3 个不同之处:

- 1) 提出了由粗到细的头发分割框架,极大地提高了分割性能;
- 2) 采用基于区域的种子选择策略,而非基于像素的策略;
- 3) 采用判别式方法而非生成式方法进行前景和背景模型的学习.

本文第 1 节概述提出的方法,第 2 节介绍基于区域的种子选择方法和 SVM 在线学习,第 3 节详述由粗到细的分割策略,第 4 节给出实验结果,并将其与其他方法进行比较,第 5 节给出结论.

1 方法概述

图 1 给出了方法的详细步骤:给定一幅输入图像,首先进行头部检测,并依据检测框归一化,得到图 1(a)所示的归一化图像;然后,采用 Mean Shift 过分割方法^[17]将图像分割为若干形状不规则、物理性质均一的区域,如图 1(b)所示;然后,将特定于头发类别的颜色和位置模型融合在一个贝叶斯框架下,用以进行种子选择.所谓前景或背景种子,即是根据目前的模型,以较高的概率属于前景或者背景的可靠区域,选出的种子如图 1(c)所示.这些种子一方面用以学习特定于当前图像的判别前景和背景的 SVM 模型,从而计算每个点属于头发的概率,如图 1(d)所示;另一方面,随机选择其中一点作为 ASF 算法的初始视点(fixation),从而得到如图 1(e)所示的初始粗略轮廓.最后,在 ASF 得到的候选区域内应用 Graph Cuts 算法,获取如图 1(f)所示的更加精细的分割结果.图 1(e)是应用学习到的 SVM 模型得到的概率图;图 1(f)则是 SVM 的输出概率和头发的位置先验概率(hair occurrence prior probability of location)融合的结果.



(a) 归一化图像; (b) Mean Shift 过分割结果; (c) 贝叶斯方法选出的头发种子(灰色区域)和背景种子(黑色区域); (d) 融合 SVM 和 HOPP 的概率图; (e) ASF 生成的闭合轮廓; (f) 分割结果

Fig.1 Flowchart of the proposed method

图 1 本文方法的流程图

2 头发种子选择及特定于图像的 SVM 分类器

2.1 采用贝叶斯准则进行基于区域的种子选择

直观地,人的头发通常分布在面部周围,并以更高的概率分布在面部上方.因此,头发和人脸的相对位置是头发分割的重要线索.将头发位置的先验概率称为 HOPP(hair occurrence prior probability),HOPP 是从事先标注好的训练集中学习得到的.具体地,首先将所有训练图像按照检测到的人脸位置归一化为同一大小,然后计算每个像素点出现的频率.图 2 给出了多姿态头发先验模型的可视化图.在离线的训练中,还需要学习通用头发颜色模型(generic hair color model,简称 GHCM).“通用”模型反映头发样例的共性,比如,大部分人的头发颜色为黑、棕、红棕等.我们采用期望最大化(expectation-maximization,简称 EM)算法学习 GHCM.

基于学习到的两种先验模型,即 HOPP 和 GHCM,种子选择的问题就可以形式化为贝叶斯模型.与文献[12]不同,本文将贝叶斯模型应用于过分割后的区域,而非像素.基于区域的策略可以在保证准确率的条件下降低计算复杂度,保证种子数目.形式地,将 $X=\{(i,j):1\leq i\leq W,1\leq j\leq H\}$ 定义为图像栅格, I_x 是定义在栅格上的彩色图像,用 I_x 表示一个区域或者一个像素的颜色向量, $I_x=[R_x,G_x,B_x]^T$; $L_x=(l_1,l_2,\dots,l_m)^T$ 表示标号向量, $m=W\times H$, 并且 $l_x=1$ 表示 x 属于前景, $l_x=0$ 表示 x 属于背景.将一个区域或者像素是头发的概率表示为 $P(l_x=1)$, 而其在通用颜色模型下的条件概率表示为 $P(I_x|l_x=1)$.因此,依照贝叶斯准则,一个区域 R 属于头发的后验概率为

$$P(I_R=1|I_R) \propto P(I_R|I_R=1) \cdot P(I_R=1) \quad (1)$$

$$I_R = \frac{1}{|R|} \cdot \sum_{x \in R} I_x \quad (2)$$

$$P(I_R=1) = \frac{1}{|R|} \cdot \sum_{x \in R} P(I_x=1) \quad (3)$$

$P(I_x=1)$ 表示位置 x 出现头发的概率.在计算了每个区域是头发的后验概率之后,一些概率较高的区域被选择为头发种子区域,概率较低的区域选择为背景种子区域.

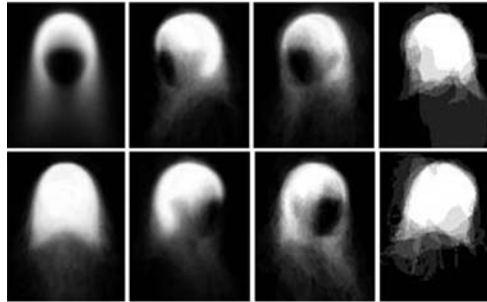


Fig.2 Visualization of multi-pose HOPP

图2 多姿态 HOPP 可视化图

2.2 头发/背景的SVM在线学习

基于已经获得的头发和背景种子,采用支持向量机(SVM)学习头发、背景判别信息.SVM 是一种用于模式分类和回归的学习算法^[16],SVM 本质的训练思想是寻求使得期望分类错误最小的线性超平面.我们在实现中使用了 LibSVM 开发包^[18].在 SVM 分类器中,RGB 矢量作为输入的特征,训练样本则是第 2.1 节中选择出的头发和背景种子像素.与之前文献[12,13]中利用 GMM 的方法相比,SVM 的优势在于其挖掘头发和背景差异的判别力更强.SVM 分类器针对每个像素都给出其属于头发或背景的概率,并提供给 Graph Cuts 进行下一步处理.

3 由粗到细的头发分割方法

由粗到细的头发分割方法主要包括两个步骤:

- (1) 粗分割:利用基于视点的主动分割方法提取出头发候选区域,该区域通常是由全部头发像素、全部或部分面部像素以及一些周边像素共同组合成的凸形状区域;
- (2) 精细分割:在由步骤(1)得到的候选区域内应用 Graph Cuts 方法,得到精确的头发区域.

以下将对这两个部分进行详细介绍.

3.1 基于视点的主动粗分割

基于视点的主动分割方法 ASF 由 Mishra 等人^[15]提出.人们关注的视点总是在图像中某区域的内部,且此区域的边界通常包络了一个完整的物体,或是完整物体的一部分.基于视点的主动分割方法能够在给定一个视点的前提下,将整个闭合的轮廓提取出来.在头发分割的过程中,头发和肤色区域可以看作是一个整体的闭合区域,而 ASF 往往可以给出这个闭合区域的轮廓.实验中使用了作者提供的源码^[19].

利用 ASF 的动机是,大部分头发区域具有相对均一的模式,并且与面部、颈部肤色连成一片,如果给 ASF 提供一个初始的“视点(fixation)”,头发区域或者头肩区域可能被作为一个整体分割出来.在上述假设成立的条件下,ASF 可以将头发区域保留,并排除大部分的非头发区域,比如其周围的深色衣服区域.图 3 给出了一些 ASF 分割的结果,可以看出,ASF 得到的闭合区域包含了几乎全部头发像素、全部或部分面部像素以及一些周边像素,即,可以保证较高的头发像素召回率.同时,ASF 的结果可以排除大量易与头发区域混淆的背景.



十字标出了“视点”,实线标出了 ASF 分割得到的轮廓

分割结果

Fig.3 ASF results examples

图 3 ASF 的结果示例

3.2 基于像素级Graph Cuts的精细分割

在第 3.1 节中,采用基于视点的主动分割方法,已经得到了粗略的头发候选区域,下一步是在此范围内构建图模型,并采用 Graph Cuts 算法求解得到精细的分割结果.

采用与第 2.1 节中一致的定义:将 $X=\{(i,j):1\leq i\leq W,1\leq j\leq H\}$ 定义为图像栅格, I_x 是定义在栅格上的彩色图像, $I_x=[R_x,G_x,B_x]^T$; $L_x=(l_1,l_2,\dots,l_m)^T$ 表示标号向量, $m=W\times H$, $l_x\in\{0,1\}$.通过最小化以下能量函数,决定最优的标号向量 L^* :

$$E(L)=D(L)+\lambda\cdot B(L) \tag{4}$$

其中, λ 用以权衡两项的重要性.公式右端的第 1 项通常称为数据项,用以描述图像局部对前景或者背景模型的拟合程度;第 2 项是平滑项,用以描述相邻像素(或区域)标号互异时的惩罚代价.

3.2.1 数据项

根据第 2.2 节的介绍,利用已经选择的可靠前景和背景种子样本,我们可以在线学习 SVM 分类器.基于该 SVM 分类器,可以计算出每个像素点属于前景或者背景的概率.为了尽可能精确地刻画前景和背景模型,我们采用贝叶斯准则,融合位置先验模型和特定于当前图像的 SVM 模型,用以描述图像对前景和背景模型的拟合程度,即

$$D(L)=\sum_{x\in X} P(I_x|l_x)\cdot P(l_x) \tag{5}$$

其中, $P(l_x)$ 表示每个像素点位置是头发或者背景的先验概率,即 HOPP;而 $P(I_x|l_x)$ 则是 SVM 的输出概率.

3.2.2 平滑项

公式(4)右端第 2 项是平滑项,用以描述相邻像素或区域标号互异时的惩罚代价.采用 4-邻域系统,选择与颜色差异相关的函数来描述平滑项:

$$B(L)=\sum_{\{p,q\}\in Neighbors} B(p,q)\cdot\delta(l_p,l_q) \tag{6}$$

其中,

$$\delta(l_p,l_q)=\begin{cases} 1, & \text{if } l_p \neq l_q \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$B(p,q)\propto \exp\left(-\frac{\|I_p-I_q\|^2}{2\sigma^2}\right)\cdot\frac{1}{Dist(I_p,I_q)} \tag{8}$$

其中,参数 σ 通常表示整幅图像的平均平滑度.

3.2.3 优化

公式(4)的能量函数最小化可以通过文献[20]提出的 Graph Cuts 算法来求解,能量函数的解对应于分割结果.通常情况下,将图像中的每个像素作为节点,在整个图像栅格上构建图^[21].本文的不同之处在于,我们只在

ASF 提供的候选区域内构建图及进行图的优化.具体地,图 4 给出了 ASF 对构建图的作用,考虑图像中的所有像素都作为图的节点,则构建的图 G 如图 4(a)示例.经过 ASF 后,图中的浅色节点被确定为前景,即头发候选节点,深色节点确定为背景.图 4(b)是由图 4(a)浅色节点及边组成的子图 G' .本方法中的 Graph Cuts 分割只考虑在子图 G' 中的节点,不考虑图 4(a)中的任何深色节点.这样做的优势在于:通过 ASF 排除了所有深色节点对优化过程带来的干扰,减少了优化的不确定性,可得到更加精确的分割结果.

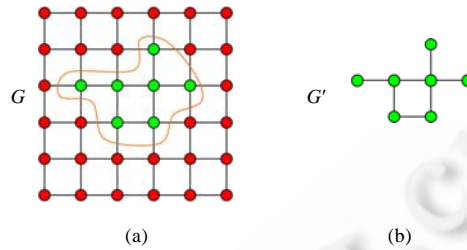


Fig.4 Effects of the ASF

图 4 ASF 的作用图示

4 实验结果与分析

本节从 3 个方面进行实验验证:首先,针对贝叶斯种子选择方法,将提出的基于区域的种子选择策略与文献 [12]中基于像素的策略进行了性能比较;之后,验证了基于视点的主动分割的效果;然后,将由粗到细的头发分割方法与文献[12,13]及由 Ning 等人^[22]提出的方法进行了性能比较.Ning 等人^[22]提出的方法是以交互的方式、基于最大相似度的区域合并(maximal similarity based region merging,简称 MSR)策略,将图像中的前景和背景分离.最后,我们在另外 3 个类别——马、汽车、飞机的公开数据库上进行了扩展实验,以验证本文方法的可推广性.

4.1 头发分割

4.1.1 数据库

我们从网上收集了两个数据库,并对每幅图片进行了头发区域的标注,用以评测分割性能.这两个数据库涵盖了各种复杂情况,如多种发型、颜色、光照条件差,姿态变化大,前景和背景区域相邻且有相似的视觉特性,容易混淆等情况.第 1 个数据库称为准正面头肩数据库 1000(near-frontal head-shoulder database 1000),简记为 NHD1000,包含了 1 000 张准正面的头肩图像,男、女性图片数目各半;涵盖了各种发型、头发颜色和不同背景.所有标注图像均被归一化为 80×100 ,两眼间距 16 像素;第 2 个数据库称为多姿态头肩数据库 3816(multi-pose head-shoulder database 3816),简记为 MHD3816.该库包含 3 816 幅图片,其中,2 886 张属于准正面图片,930 张为侧面或背面图片,所有标注图像按照头部的检测矩形框,归一化 80×100 的图片.为了训练通用的头发颜色模型和位置先验模型,我们另外收集了 300 幅图片并标注头发区域,该数据集与 NHD1000, MHD3816 均无交集.

4.1.2 种子选择实验

实验在 NHD1000 数据库上比较了基于区域和基于像素的种子选择策略.为了公平性地比较,对于两种策略,实验选择相同的头发颜色模型和位置先验模型.头发颜色的高斯混合模型采用了 13 个高斯成分,颜色模型和位置先验模型的权重分别为 0.7 和 0.3.对于基于区域的策略,采用 Mean Shift 进行图像的过分割,设置空间、颜色差异度参数分别为 5 和 2,区域最少像素数目为 50.

关于一幅图像是否选择到可靠的种子区域,我们设定条件:若选出的头发种子像素达到 95%的准确率,则认为一幅图像选到了正确的种子.实验结果表明,同样是保证 90%的图片选到正确的种子,采用基于像素的种子选择策略平均种子数目达到 320 个像素,而基于区域的种子选择策略平均种子数目达到 370 个像素.即,在保证同样种子准确率的情况下,采用基于区域的种子选择策略,可以选择到更加足量的种子像素.

4.1.3 基于视点的主动分割实验

这一节将评测 ASF 对头发分割的有效性和鲁棒性.如第 3 节所述,ASF 算法需要一个初始的物体上的“视点”.在实现中,从选择到的头发种子中,选择居中序号的种子点作为视点.给定初始视点后,期望由 ASF 给出头发的大概轮廓,涵盖头发的大部分区域.图 5 给出 ASF 在 NHD1000 和 MHD3816 两个数据库上测试头发区域的召回率(recall rate)和准确率(precision rate).从中可以看出,ASF 粗分割给出的候选区域可以保证高达 95%的召回率.另外,ASF 的结果对初始视点的位置并不敏感,如图 6 所示,随机从种子区域中选择 3 个不同的视点,得到的轮廓大体是一致的.这说明,以头发区域不同像素为视点的 ASF 可以生成稳定的闭合区域轮廓.

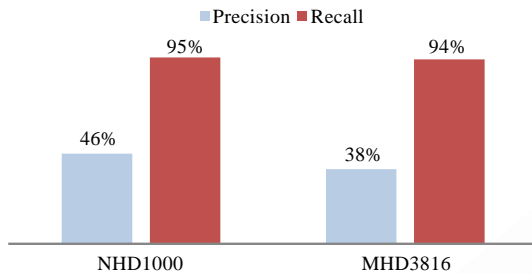


Fig.5 ASF performance on different databases
图 5 ASF 在不同数据库上的性能

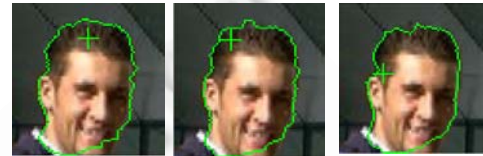


Fig.6 ASF results with different fixation points
图 6 不同视点得到的 ASF 分割结果

4.1.4 不同头发分割方法的比较

实验将本文方法与文献[12]的贝叶斯方法、文献[22]中的 MSRM 方法以及文献[13]中的 Graph Cuts 方法进行了性能比较.所有方法都涉及到种子选择的步骤,为公平起见,前端都采用基于区域的贝叶斯种子选择方法.文献[22]中的 MSRM 是一种交互式方法,我们采用了作者提供的源码^[23],其前端需要用户输入,提示部分前景和背景区域;本实验将其前端改为自动种子选择,使其成为全自动的分割方法.通过衡量分割结果和手工标定的一致性来评价分割性能,具体采用 F -value^[24]作为评测标准, F -value 定义为 $2PR/(P+R)$.其中, P (precision)代表自动分割的头发区域和手工标定区域一致的像素所占的比例, R (recall)表示手工标定的头发区域和自动分割区域一致的像素所占的比例. PR -曲线上不同的点,对应着不同的 F -value,本文给出的是针对每种方法调优后的结果.

表 1 比较了本文提出的方法与文献[12]、文献[22]和文献[13]中方法的性能.在所有的实验中,针对每种方法,我们分别调整参数使其达到最优的 F -value,包括对于文献[12]中的贝叶斯方法中概率的阈值设定、GMM 成分的数目、Graph Cuts 的能量函数中数据项和平滑项的权重 λ 的设置等.从表 1 中可以看出,由粗到细的头发分割策略在两个数据库上的性能都优于其他方法.

Table 1 Performance evaluation of different methods

表 1 不同分割方法的性能比较

Algorithm	NHD1000		MHD3816	
	F -value	P/R	F -value	P/R
Bayesian method ^[12]	0.66	0.59/0.81	0.64	0.74/0.62
MSRM ^[22]	0.74	0.77/0.78	0.70	0.81/0.70
GC based approach ^[13]	0.81	0.87/0.81	0.67	0.82/0.64
The proposed method	0.85	0.89/0.85	0.77	0.79/0.82

图 7 给出了一些可视化的实验结果比较,除了最后一行的失败例子以外,其余都是相对成功的例子,第(f)列给出了中间 ASF 的结果.由粗到细的分割方法之所以比较有效,得益于两个方面:首先,ASF 算法通常可以去除大部分难与前景分离开的背景区域;其次,Graph Cuts 只在 ASF 结果的基础上运行,即排除大部分背景影响后,全局优化更加有效地集中在少数节点上.贝叶斯方法的问题在于:无法对每幅图像都单独调整较优的阈值,并且该方法没有考虑空间上相邻像素标号的平滑性;MSRM 的不足在于:在种子区域连通的情况下(参考图 7 第 3 行

的例子),它无法分割出有两个不连通区域的物体.文献[13]的 Graph Cuts 方法中,由于无法实现预测背景的复杂程度,也就无法为每幅图像自适应地调整权重 λ ,由粗到细的分割方法虽然涉及到权重设置问题,但与文献[13]相比,由于 ASF 去除了大量的背景,为 Graph Cuts 优化去除了大量的干扰,使其对于不同权重值不敏感,因而有相对一致的优化结果.



(a) 输入; (b) 贝叶斯^[12]; (c) MSRM^[22]; (d) Graph Cuts^[13]; (e) 本文方法; (f) ASF

Fig.7 Comparison of different segmentation algorithms

图 7 不同方法的结果比较

4.2 推广性实验

4.2.1 数据库

为了验证由粗到细的分割方法的推广性,我们在另外 3 个类别——马、汽车和飞机类别上进行了扩展实验,这也是相对于早期版本^[4]增加的重要部分.这 3 个类别的实例在颜色、结构等方面的类内变化都极具多样性.在推广过程中,我们既从方法上进行了扩展(详见第 4.2.2 节),也增加了大量的实验.实验在公开的 Weizmann Horse^[25],ETHZ-Cars^[26,27],MSRC-Aeroplane^[28]数据库上进行,数据库的示例如图 8 所示.下面分别给出各数据库的介绍:

- 1) Weizmann Horse 数据库包含 328 幅马的图片,涉及到多种四肢姿态、颜色、纹理等复杂情况,每幅图片都有对应标注了马和背景的二值图;
- 2) ETHZ-Cars 包含了 100 幅汽车的图片,包括了不同款式、视角、颜色等情况,每幅图片都标定了汽车

的像素;

- 3) MSRC-Aeroplane 包括了 30 幅飞机的图片,每幅图像的飞机机身都包含多种颜色;不同图像中的飞机视角也有一定差异;而飞机的表现与背景建筑物、天空的相似性也增加了分割的难度.由于原始的标注不够准确,我们采用了 Malisiewicz 和 Efron^[29]提供的精确标注作为评测的 Ground truth.

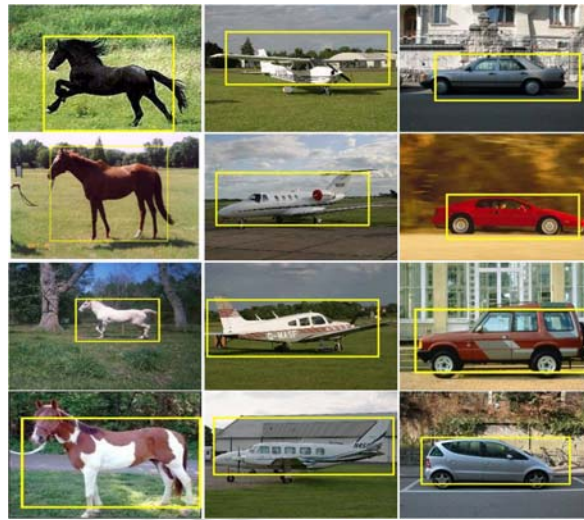


Fig.8 Detection examples of horses, aeroplanes and cars

图 8 特定类物体(马、飞机、汽车)检测的结果示例

4.2.2 方法扩展

我们将由粗到细的方法从头发分割推广到解决其他特定物体类分割的问题上,并进行了两点扩展:

第一,采用 Felzenszwal 等人^[30]提出的可变形部件模型进行物体检测,图 8 给出了一些检测结果的示例;

第二,在粗分割步骤采用了多个视点的输入,以期得到较高的召回率;为使不同视点的分布具有一定散度,采用随机的方式从种子像素中选择 $nPoints$ 个像素点作为视点.

基于以上扩展,由粗到细的物体分割方法的实验步骤为:

- 1) 采用可变形部件模型检测物体的位置,得到矩形检测框;
- 2) 根据矩形框位置,向周围作 20%的区域扩展,并将图片进行归一化;
- 3) 根据物体的颜色模型和位置模型,确定前景和背景区域的种子区域;
- 4) 根据种子区域训练特定于当前图像的 SVM 分类器,计算每个像素点是前景或背景的概率;
- 5) 在种子区域随机选择一个或多个视点,对每个视点分别采用 ASF 方法进行分割,得到粗略轮廓,将所有轮廓内区域的并集作为候选区域;
- 6) 在步骤 5)得到的候选区域内运行 Graph Cuts 算法,得到最终分割结果.

在训练集合上统计得到的位置先验模型如图 9 所示.注意,该实验采用和头发分割实验中相同的参数设置,比如选种子时颜色、位置信息的权重和 Graph Cuts 算法中的平滑项权重等.

在方法扩展中,以多视点输入得到的轮廓并集作为候选区域,是获得召回率的重要一步.这是因为飞机、汽车等实例本身表现变化较大,物体内部有较强的内边缘,ASF 通常很难以一个视点的输入获得物体的全部候选区域.如图 8 所示,多数车体部分的车窗、车身、车轮、车胎都有不同的颜色、纹理特征,且车窗边缘、车轮边缘甚至车门附近的边缘都比较显著,此时,单独一个视点并不能完全获得车体的所有候选区域.我们将方法扩展为多视点输入,每个输入分别产生候选区域,而后把所有区域的并集作为候选区域.如图 10 所示,图 10(b)中,一种颜色的视点应对应相同颜色的输出轮廓,单独一个视点得到的轮廓并不能覆盖所有车体区域,而将所有视点应对

的轮廓封闭区域并集作为候选区域,则可以基本保证车体区域较高的召回率.

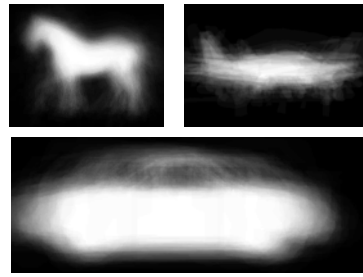
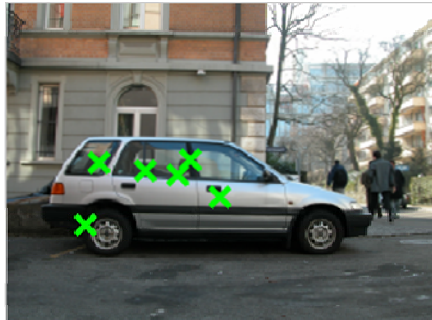


Fig.9 Location models of horses, cars, aeroplanes

图 9 马、飞机、汽车的位置先验模型



(a) 输入图像及视点(标有x处)



(b) 对不同视点,ASF 得到的粗略轮廓(每个彩色轮廓是经 ASF, 由与其相同颜色的输入视点得到封闭轮廓)

Fig.10 An example for ASF with more than one fixation points

图 10 ASF 多个视点示例

图 11 中给出了召回率(recall)随视点数目变化的规律.可以看出,当数目小于 8 时,召回率持续上升,而从 8 开始上升变缓.因此,视点选择太少,不易得到较高的召回率;而太多又容易选到过多背景,增加干扰.本文的实验中,所有视点从种子区域中自动选择得到.针对马、汽车、飞机这 3 个类别,我们根据经验分别设置 $nPoints=1,8,4$. 图 12 给出了不同数据库上的 ASF 方法的性能比较,可以看出,在新增的马、汽车和飞机类别上,ASF 保持了较高的召回率.

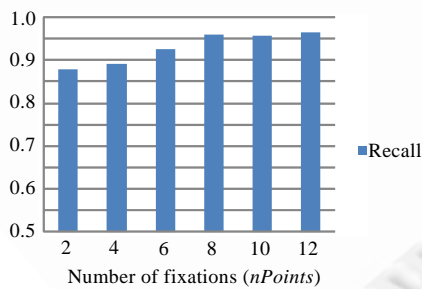


Fig.11 Recall varies with the number of fixations (ETHZ-Cars)

图 11 召回率随视点数目变化的规律(ETHZ-Cars)

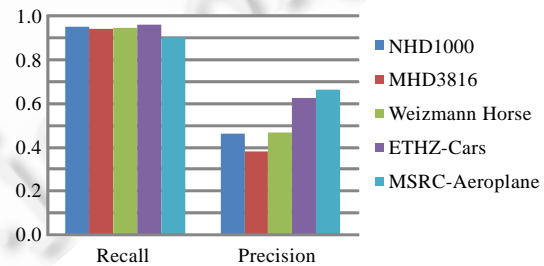


Fig.12 ASF performances on different databases

图 12 不同数据库上的 ASF 性能

4.2.3 实验对比

实验将 Graph Cuts^[21]和 GrabCut^[31]作为基准(baseline)的方法进行性能比较,并采用 F -value 对分割结果进

行评价.我们选择这两种方法作为 baseline 的原因主要有两条:

- 1) 标准的 Graph Cuts 方法没有初始的粗略分割,而是基于整幅图的所有像素构建图的结构,与之比较可以验证本文由粗到细策略的有效性;
- 2) GrabCut 是典型的交互分割方法,其前端需要用户标注含分割物体的矩形框,作为分割范围的约束,后端在此约束的范围内进行迭代分割.其框架结构和由粗到细的方法有相似之处,因此作为 baseline 方法的一种.

为公平起见,Graph Cuts 方法前端的种子选择方法以及学习前景和背景模型的方法均与本文方法一致.对于 GrabCut 方法,采用两种 bounding box 的输入:

- 1) 一种是以 ground truth 限定的矩形框作为其输入,可以使 GrabCut 方法达到其最优性能,记为 GrabCut1(GrabCut1 以 ground truth 前景边界的外接矩形为输入);
- 2) 一种是采用了 Kuettel^[32]在实验对比中的方式,即 bounding box 以图像中心为中心点,取整幅图像的 50%作为初始输入,记为 GrabCut2(GrabCut2 以图片中心向周围扩展 50%图像面积的矩形为输入).

表 2 比较了本文方法与 baseline 方法在 Weizmann Horse,ETHZ-Cars,MSRC-Aeroplane 数据库上的性能.可以看出,我们的方法超越了其他 3 种 baseline 方法,在 Weizmann Horse 数据库上的 F -value 达到了 0.81,充分证明了由粗到细分割方法的优越性.图 13 给出了本文方法和对比方法在 Weizmann Horse 数据库的分割结果示例.

Table 2 Performance comparison on different databases

表 2 不同数据库上的性能比较

Methods	Weizmann Horse		ETHZ-Cars		MSRC-V2-Aeroplane	
	F -value	P/R	F -value	P/R	F -value	P/R
Graph cuts	0.71	0.65/0.85	0.79	0.82/0.80	0.72	0.67/0.82
GrabCut1	0.77	0.73/0.80	0.84	0.81/0.85	0.73	0.69/0.82
GrabCut2	0.58	0.79/0.46	0.63	0.62/0.65	0.56	0.85/0.44
Our method	0.81	0.83/0.82	0.86	0.87/0.86	0.76	0.77/0.79

图 13 给出了本文方法和对比方法在 Weizmann Horse 数据库的分割结果示例.

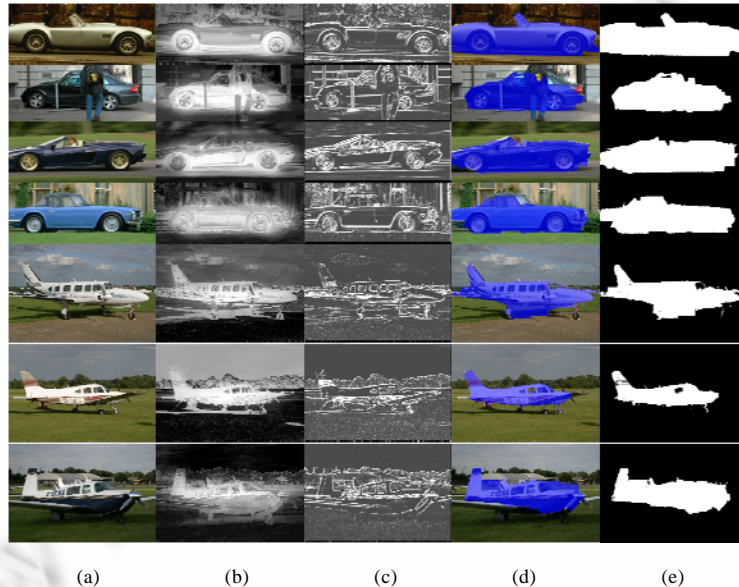


Fig.13 Comparison on Weizmann Horse database

图 13 不同方法在 Weizmann Horse 数据库上的结果比较

从中可以看出,复杂背景下,本文方法仍能较精确地分割出马体.另外,由于马腿部颜色不一致,细节部分会有所遗漏.从表 2 中还可以看出,GrabCut 和由粗到细的分割方法都比单纯采用 Graph Cuts 方法的效果要好.这是因为前两种方法均采用策略限定了优化范围,减少了后端 Graph Cuts 构建的图模型的节点数目,也即降低了复杂背景带来的不可预测的影响.ASF 限定分割范围的策略与 GrabCut 的不同之处在于:ASF 倾向于割出物理性质一致的区域,不限制边缘形状;而 GrabCut 采用直接圈定矩形框的方式,范围内混入了太多背景.

图 14 中给出了一些汽车和飞机的分割结果以及颜色和边缘等信息所起作用的示意图.大部分情况下,本文方法可以获得理想的分割结果,但在最后一幅飞机图像中,左边机翼没有很好地分割出来,主要是由于机身的黑色边缘到机翼的白色区域跳变太大,而 Graph Cuts 优化倾向于较短边缘引起的.



(a) 输入;(b) 颜色模型和位置模型得到的概率图;(c) 边缘强度图;(d) 分割结果;(e) 二值分割图

Fig.14 Effects of different cues

图 14 每个信息(颜色、位置、边缘)通道所起的作用示意图

5 结束语

本文提出了一种由粗到细的头发分割方法.在该方法中,首先,巧妙地利用 ASF 可以分割出凸形状轮廓的特性,粗略分割出头发的大致区域,保证头发区域较高的召回率;然后,在 ASF 得到的闭合区域内进行 Graph Cuts 的优化求解,去除其中的非头发区域,得到细致的边界.具体地,Graph Cuts 中的前景和背景模型是通过融合 SVM 在线学习的特定模型和通用的头发分割模型得到的,该方法的优势在于:

- 1) ASF 可以有效地提取出高召回率的闭合区域,由此为 Graph Cuts 的优化求解去除了大量不可预测的复杂背景的干扰;
- 2) 基于区域的种子选择策略,可以在保证准确率的前提下选择出足量的种子,用于学习当前图像的前景和背景模型;
- 3) 采用了 SVM 分类器进行在线学习,判别式的方法可以更加有效地捕捉前景和背景的差异.

实验结果表明,本文提出的方法优于已有的头发分割方法,并在其他物体类别的分割上有一定的可推广性.

References:

- [1] Hadap S, Magnenat-Thalmann N. Modeling dynamic hair as a continuum. *Computer Graphics Forum*, 2001,20(3):329–338. [doi: 10.1111/1467-8659.00525]
- [2] Paris S, Briceno HM, Sillion FX. Capture of hair geometry from multiple images. *ACM Trans. on Graphics (TOG)*, 2004,23(3): 712–719. [doi: 10.1145/1015706.1015784]
- [3] Bai XD. Real sense simulation of three dimensions hair [MS. Thesis]. Xi'an: Xidian University, 2006 (in Chinese with English abstract).
- [4] Ward K, Galoppo N, Lin M. Interactive virtual hair salon. *Presence: Teleoperators and Virtual Environments*, 2007,16(3):237–251. [doi: 10.1162/pres.16.3.237]
- [5] Yacoub Y, Davis LS. Detection and analysis of hair. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2006,28(7): 1164–1169. [doi: 10.1109/TPAMI.2006.139]
- [6] Ueki K, Komatsu H, Imaizumi S, Kaneko K, Imaizumi S, Sekine N, Katto J, Kobayashi T. A method of gender classification by integrating facial, hairstyle, and clothing images. In: *Proc. of the ICPR, Vol.4. 2004.* 446–449. [doi: 10.1109/ICPR.2004.1333798]
- [7] Wang L, Yu Y, Zhou K, Guo B. Example-Based hair geometry synthesis. *ACM Trans. on Graphics (TOG)*, 2009,28(3):56:1–56:9. [doi: 10.1145/1531326.1531362]
- [8] Paris S, Chang W, Kozhushnyan OI, Jarosz W, Matusik W, Zwicker M, Durand F. Hair photobooth: Geometric and photometric acquisition of real hairstyles. In: *Proc. of the ACM SIGGRAPH. 2008.* [doi: 10.1145/1399504.1360629]
- [9] Liu ZQ, Guo JY, Bruton L. A knowledge-based system for hair region segmentation. In: *Proc. of the Int'l Symp. on Signal Processing and Its Applications. 1996.* 575–576. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=615106>
- [10] Rousset C, Coulon PY. Frequential and color analysis for hair mask segmentation. In: *Proc. of the ICIP. 2008.* 2276–2279. [doi: 10.1109/ICIP.2008.4712245]
- [11] Fu WL. Image segmentation algorithm research and application of hair [MS. Thesis]. Shanghai: Shanghai Jiaotong University, 2010 (in Chinese with English abstract).
- [12] Wang D, Shan SG, Zeng W, Zhang HM, Chen XL. A novel two-tier Bayesian based method for hair segmentation. In: *Proc. of the ICIP. Cairo, 2009.* 2401–2404. [doi: 10.1109/ICIP.2009.5414215]
- [13] Lee K, Anguelov D, Sumengen B, Gokturk SB. Markov random field models for hair and face segmentation. In: *Proc. of the IEEE Int'l Conf. on Automatic Face and Gesture Recognition. 2008.* 1–6. [doi: 10.1109/AFGR.2008.4813431]
- [14] Wang D, Chai XJ, Zhang HM, Chang H, Zeng W, Shan SG. A novel coarse-to-fine hair segmentation method. In: *Proc. of the IEEE Int'l Conf. on Automatic Face and Gesture Recognition. 2011.* 233–238. [doi: 10.1109/FG.2011.5771403]
- [15] Mishra A, Aloimonos Y, Fah CL. Active segmentation with fixation. In: *Proc. of the ICCV. 2009.* 468–475. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5459254
- [16] Cortes C, Vapnik V. Support vector networks. *Machine Learning*, 1995,20(3):273–297. [doi: 10.1007/BF00994018]
- [17] Comaniciu D, Meer P. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002,24(5):603–619. [doi: 10.1109/34.1000236]
- [18] Chang CC, Lin CJ. LIBSVM—A library for support vector machines. 2011. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- [19] Mishra A, Aloimonos Y, Fah CL. Fixation-Based segmentation code w/o fixation strategy. 2012. <http://www.umiacs.umd.edu/~mishraka/activeSeg.html>
- [20] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2001,23(11):1222–1239. [doi: 10.1109/34.969114]
- [21] Boykov YY, Jolly MP. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. In: *Proc. of the ICCV. 2001.* 105–112. [doi: 10.1109/ICCV.2001.937505]
- [22] Ning J, Zhang L, Zhang D, Wu C. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 2010,43(2):445–456. [doi: 10.1016/j.patcog.2009.03.004]
- [23] Ning J, Zhang L, Zhang D, Wu C. Interactive image segmentation by maximal similarity based region merging. 2010. <http://www4.comp.polyu.edu.hk/~cslzhang/papers.htm>
- [24] Van Rijsbergen CJ. *Information Retrieval*. 2nd ed., London: Butterworths, 1979.

- [25] Borenstein E, Ullman S. Class-Specific, top-down segmentation. In: Proc. of the ECCV. 2002. 639–641. <http://dl.acm.org/citation.cfm?id=649285>
- [26] Leibe B, Leonardis A, Schiele B. Combined object categorization and segmentation with an implicit shape model. In: Proc. of the ECCV Workshop on Statistical Learning in Computer Vision. 2004. 17–32. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.6272>
- [27] Agarwal S, Awan A, Roth D. Learning to detect objects in images via a sparse, part-based representation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004,26(11):1475–1490. [doi: 10.1109/TPAMI.2004.108]
- [28] Shotton J, Winn J, Criminisi A. TextonBoost for image understanding: Multi-Class object recognition and segmentation by jointly modeling texture, layout, and context. Int'l Journal of Computer Vision, 2007,71(1):2–23. [doi: 10.1007/s11263-007-0109-1]
- [29] Malisiewicz T, Efros A. Improving spatial support for objects via multiple segmentations. In: Proc. of the BMVC. 2007. <http://130.203.133.150/showciting.jsessionid=3793017ACB6314D3C91D3599212670B8?cid=5032225>
- [30] Felzenszwalb P, Girshick R, McAllester D, Ramanan D. Object detection with discriminatively trained part based models. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2010,32(9):1627–1645. [doi: 10.1109/TPAMI.2009.167]
- [31] Rother C, Kolmogorov V, Blake A. GrabCut: Interactive foreground extraction using iterated graph cuts. ACM Trans. on Graphics (TOG), 2004,23(3):309–314. [doi: 10.1145/1015706.1015720]
- [32] Kuettel D, Ferrari V. Figure-Ground segmentation by transferring window masks. In: Proc. of the CVPR. 2012. 558–565. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6247721

附中文参考文献:

- [3] 白小丹.真实感三维头发的模拟[硕士学位论文].西安:西安电子科技大学,2006.
- [11] 傅文林.图像分割技术研究及头发分割应用[硕士学位论文].上海:上海交通大学,2010.



王丹(1985—),女,山东菏泽人,博士,主要研究领域为图像处理,模式识别.
E-mail: dan.wang@vip1.ict.ac.cn



曾炜(1973—),男,博士,研究员,主要研究领域为计算机视觉,图像处理.
E-mail: zeng_wei@nec.cn



山世光(1975—),男,博士,研究员,博士生导师,CCF 会员,主要研究领域为模式分析,机器学习.
E-mail: sgshan@ict.ac.cn



陈熙霖(1965—),男,博士,研究员,博士生导师,CCF 杰出会员,主要研究领域为计算机视觉,模式识别,人工智能.
E-mail: xlchen@ict.ac.cn



张洪明(1976—),男,博士,研究员,主要研究领域为计算机视觉,模式识别.
E-mail: hmzhang76@qq.com