

基于 GrabCut 改进算法的服装图像检索方法

胡玉平 肖 行 罗东俊
(广东财经大学信息学院 广州 510320)

摘 要 为了消除服装图像背景的影响,针对目前的 GrabCut 算法存在对图像局部像素值的变化敏感、时间开销大、边缘不准确等问题,提出了改进的 GrabCut 算法。在改进算法中,通过对梯度图像使用多尺度分水岭去噪增强了图像的边缘信息,减少了后续处理的计算量;通过采取熵惩罚因子最优能量函数减少了检索图像的有效信息丢失。将改进后的 GrabCut 算法引入基于内容的服装图像检索系统中,实验结果表明与同类方法相比,所提方法在检索显示准确性以及检索的平均查准率和查全率方面均有明显的提升。

关键词 图像检索, GrabCut 算法, 图像分割, 图像背景
中图法分类号 TP391 文献标识码 A

Clothing Image Retrieval Method Based on Improved GrabCut Algorithm

HU Yu-ping XIAO Hang LUO Dong-jun

(School of Information, Guangdong University of Finance & Economics, Guangzhou 510320, China)

Abstract In order to eliminate interference of clothing image background, GrabCut algorithm was introduced. But the current GrabCut algorithm is sensitive to local noise and time-consuming, and its segmentation edge is not accurate. To solve these problems, the multi-scale watershed algorithm to de-noise gradient image was employed, enhancing the image edge points and reducing the subsequent processing computation. To reduce the loss of image key features, we employed entropy penalty factor optimal segmentation energy function, reducing the effective information loss of image retrieval. And then we introduced the improved GrabCut algorithm to content based clothing image retrieval system. The experimental results show the method has obvious advance on the accuracy of retrieval effect than the existing algorithms.

Keywords Image retrieval, GrabCut algorithm, Image segmentation, Image background

1 引言

随着电子商务的发展,特别是在线购衣的兴起,对图像检索技术的要求不断提高。现有的服装图像检索方法大多是基于图像的一些底层特征,如纹理、颜色、形状等。这些图像检索方法具有普适性,任何类别的图像检索都可以使用这些检索方法。如果没有针对服装固有特征进行检索,必然导致服装检索结果不理想,检索效率不高^[1,2]。

服装图像按照有无背景可以分为两类:无背景的服装图像和有背景的服装图像。无背景的服装图像主要是电子商务网站中的服装图像,其只包含服装,人工拍摄后经过了一定的后期处理,背景单一,如图 1(a)中的服装图像,在图像检索时识别率较高;对于另外一类有背景的服装图像,一些电子商务网站为了展示服装的美观,直接拍摄穿在模特上的服装图像,其参杂了较多的背景(噪声),如图 1(b)的服装图像,这一类图像在检索时较难识别,因此如果能够利用图像分割算法把服装图像分割出来,继而进行检索,识别率一定会大大提升。



图 1 无背景与有背景服装图像

服装图像分割的主要目的是找出服装的区域部分,过滤掉无用的信息,为服装图像的下一步分类识别做准备。目前的图像分割算法按照人工参与的程度可以分为部分人员参与和无人参与;按照采用的策略可以分为交互式图像分割和全自动图像分割^[3]。全自动图像分割由于可以大量节省时间开销,提高处理效率,因此是图像分割的完美解决方案。而交互式图像分割方法与全自动图像分割算法所使用的分割算法几乎相同,两种方法的主要区别在于交互式的分割方法需要人工进行干预,比如人为地选择一些参数、附加条件或其它准

本文受广东省自然科学基金(2016A030313717),国家自然科学基金课题(61472135)资助。

胡玉平(1969—),男,博士,教授,主要研究方向为网络安全与电子商务技术;肖 行(1980—),男,硕士生,主要研究方向为电子商务技术, E-mail:okhyp@gdufe.edu.cn;罗东俊(1975—),男,博士,主要研究方向为网络安全。

则。由于图像分割一般都是由人的主观因素决定的,因此交互式图像分割方法依据人的先验信息能获得更加准确的图像分割效果。常用的一些交互式图像分割方法大多是基于图论的,如 Live wire^[4], Intelligent scissors^[5], Graph Cuts^[6] 以及 GrabCut^[7]等。传统的交互式图像分割方法虽然在效果上更胜一筹,但因它需要人机交互而增大了人工工作量,同时也很容易受人主观对图像判断的影响。

本文在分析经典 GrabCut 算法的基础上,对其能量函数进行了优化,同时对其分水岭算法算子进行了改进,通过对梯度图像使用多尺度分水岭去噪,增强了图像的边缘信息,减少了后续处理的计算量,通过采取熵惩罚因子最优能量函数减少了检索图像的有效信息丢失。将改进后的 GrabCut 算法引入基于内容的服装图像检索系统中,实验结果表明:与同类方法相比,本文方法在检索显示准确性以及检索的平均查准率和查全率方面均有明显的提升。

2 经典 GrabCut 算法介绍

2.1 主要思想

GrabCut 算法^[7]的主要思想是把图像中的所有像素映射为图论中 S-T 网络,其中 S 表示源点,主要用来表示目标物;T 表示汇点,主要用来表示背景部分。如图 2 所示,在图中有两种类型的边(Edge):图像中相邻像素之间的边,它表示相邻像素之间的相关性,如果相邻像素之间的相关性大,那么这条边对应的值就较大;像素点到 S 和 T 之间的边,它的值主要通过计算其属于前景及背景的概率得出。

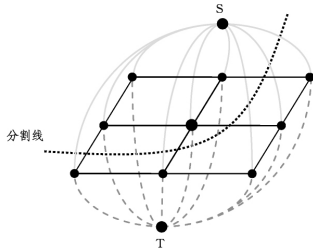


图 2 S-T 网络模型

GrabCut 算法是一种需要用户交互的图像分割方法,用户只需在图像中用一个矩形框把要分割的目标图像标注出来,如图 3(a)所示的图像,矩形框通过人工凭先验知识画出。矩形框以外表示背景区域 T_B ;矩形框以内包括我们需要提取的目标图像,同时也包括一些不希望有的背景像素,所以为位置区域 T_U 。

假设整张服装图像一共有 z_n 个像素点,本文所用的每张服装图像中都有 65536 个像素点,即 $n \in [1, 2, \dots, 65536]$,在实际处理图像像素排列时,按照图像每行从上往下逐个像素扫描。这里用 α_n 表示服装图像的分割,初始分割为用户手工画出的矩形框,背景区域 T_B 中 α_n 中的值为 0,未知区域 T_U 中 α_n 的值为 1。再利用混合高斯模型 GMM 分别对 T_B 和 T_U 建模,以判断未知区域 T_U 中的像素分别属于 T_B 和 T_U 的概率, T_B 和 T_U 的 GMM 模型分别由 5 个高斯模型混合而成,每个像素分别属于的模型用 k_n 表示,其中 $n \in [1, 2, \dots, 65536]$ 。则像素值与源点、汇聚点之间的权值可以表示为:

$$D_n(\alpha_n, k_n, \theta, z_n) = -\log \pi(\alpha_n, k_n) + \frac{1}{2} \log \det \Sigma(\alpha_n, k_n) +$$

$$\frac{1}{2} [z_n - \mu(\alpha_n, k_n)]^T \Sigma(\alpha_n, k_n)^{-1} [z_n - \mu(\alpha_n, k_n)] \quad (1)$$

其中, $\theta = \{\pi(\alpha, k), \mu(\alpha, k), \Sigma(\alpha, k)\}$, $\pi(\cdot)$ 表示高斯混合系数, $\mu(\cdot)$ 表示均值, $\Sigma(\cdot)$ 表示协方差。下式表示像素值 z_m 与 z_n 之间的权值:

$$V(\alpha, z) = \gamma \sum_{(m,n) \in C} [\alpha_m \neq \alpha_n] \exp(-\beta \|z_m - z_n\|^2) \quad (2)$$

其中, z_m 与 z_n 分别表示这两个像素点的灰度值,由于只在人工初始化边界处继续分割图像,因此需要满足 $\alpha_m \neq \alpha_n$,即一个像素点位于未知区域,另外一个像素点位于背景区域;如果 $\alpha_m = \alpha_n$,那么这两个像素之间的权值为 0,否则权值为 $\exp(-\beta \|z_m - z_n\|^2)$,即底数为 e,指数为像素值差的绝对值。

整幅图像的 Gibbs 能量可以表示为:

$$E(\alpha, k, \theta, z) = U(\alpha, k, \theta, z) + V(\alpha, z) \quad (3)$$

其中, $U(\alpha, k, \theta, z)$ 表示像素值与源点、汇点之间的权值 $D_n(\alpha_n, k_n, \theta, z_n)$ 之和。图像分割的过程就是使式(3)不断减小的过程,当它不能再减小,即趋于某一个恒定的值时,就表示图像分割完成。

2.2 算法流程

整个算法的具体流程如下。

Step1 输入图像 I。

Step2 提取目标图像。

1) 通过用户初始化的目标边界线来区分三元图 T 初始轮廓线。 T_B 为背景区域, T_U 为未知区域, T_F 为目标区域。

2) 将 T_U 内全部像素点的 α_n 置为 1, 将 T_B 内全部像素点的 α_n 置为 0, 使用 α 初始化前景和背景的 GMM 模型参数。

Step3

1) 为每一个 T_U 区域的像素计算出 GMM 分量:

$$k_n := \arg \min D_n(\alpha_n, k_n, \theta, z_n) \quad (4)$$

2) 利用数据 Z 训练 GMM 参数:

$$\theta := \arg \min U(\alpha, k, \theta, z) \quad (5)$$

3) 用 Max-Flow Min-Cut 算法估计初始分割:

$$\min_{\{\alpha_n, n \in T_U\}} = \min_k E(\alpha, k, \theta, z) \quad (6)$$

4) 回到 1), 直到 Gibbs 能量函数收敛。

Step4 获取目标图像 I_{result} 。

图 3 为 GrabCut 分割目标的示意图,首先选取分割图像,再通过人工添加一个预分割目标的矩形框,添加的矩形框至少要包括预分割的目标,矩形框包括的背景越少,目标分割就会越准确,目标的边缘也越精准。

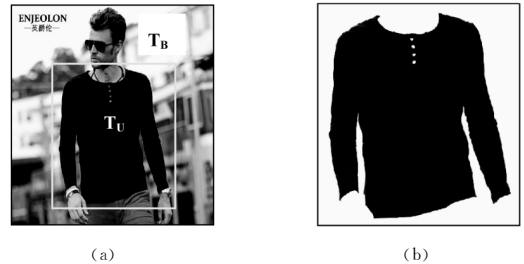


图 3 GrabCut 分割示意图

3 GrabCut 分割算法的改进

3.1 能量函数的优化

假设整幅服装图像一共有 z_n 个像素点, $n \in [1, 2, \dots,$

65536], 维数为 d , 使用 GMM 模型描述背景和前景像素之间的分布情况。像素点都是由 K 个 GMM 的混合模型表示, 且 $k_N \in \{1, 2, \dots, 5\}$; α_n 为像素点的 Transparency, 其中 $\alpha_n = 1$ 表示前景, $\alpha_n = 0$ 表示背景。

$$p_k(z_i, \theta) = \frac{\exp[-\frac{1}{2}(z_i - u_k)^T \sum_{i=1}^k (z_i - u_k)]}{(2\pi)^{\frac{d}{2}} |\sum_k|^{-\frac{1}{2}}} \quad (7)$$

其中, x_i 表示第 k 个高斯模型下的密度函数。

$$p_k(z_i/\alpha_n, \theta) = \sum_{k=1}^K \pi_k p_k(z_i, \theta) \quad (8)$$

其中, $\theta = \{\pi_k, u_k, \Sigma_k\}$ 为参数, π_k 为比例系数, u_k 为平均值, Σ_k 为协方差。

似然估计公式为:

$$\begin{aligned} L(x) &= -\log \prod_{i=1}^N \sum_{k=1}^K \pi_k p_k(z_i, \theta) \\ &= -\sum_{i=1}^N \log [\sum_{k=1}^K \pi_k p_k(z_i, \theta)] \end{aligned} \quad (9)$$

原 GrabCut 中 Gibbs 的能量函数形式如下:

$$\begin{aligned} E(\alpha, k, \theta, x) &= U(\alpha, k, \theta, z) + V(\alpha, z) \\ &= \sum_n D(\alpha_n, k_n, \theta, z_n) + V(\alpha, z) \end{aligned} \quad (10)$$

其中, (α, k, θ, z) 为区域项, $V(\alpha, z)$ 为平滑项。由上面两式得到:

$$\begin{aligned} D(\alpha_n, k_n, \theta, z_n) &= -\log [\pi_k p_k(z_n, \theta)] \\ &= -\log \pi_k - \log p_k(x_n, \theta) \\ &= -\log \pi_k + \frac{1}{2} \log \det \Sigma(\alpha_n, k_n) + \frac{1}{2} [x_n - \\ &\quad u(\alpha_n, k_n)]^T \Sigma(\alpha_n, k_n)^{-1} [x_n - u(\alpha_n, k_n)] \end{aligned} \quad (11)$$

通过更新 GMM 模型的参数得到下面两个公式:

$$\theta = \arg \min U(\alpha, k, \theta, z) \quad (12)$$

$$V(\alpha, z) = \gamma \sum_{(m,n) \in C} [a_m \neq a_n] \exp(-\beta \|z_m - z_n\|^2) \quad (13)$$

其中, C 为相邻像素点对的数量, $\gamma = 50$, $[a_m \neq a_n]$ 为 0 或 1 的二值函数。当且仅当 $\alpha_m \neq \alpha_n$ 时, $[a_m \neq a_n] = 1$; 当且仅当 $\alpha_m = \alpha_n$ 时, $[a_m \neq a_n] = 0$ 。 $\beta = (2 \langle \|x_m - x_n\|^2 \rangle)^{-1}$ 为数学期望。

熵在信息论的定义为:

$$H(\pi) = -\sum_{k=1}^K \pi_k \log(\pi_k) \quad (14)$$

其中, π 为信息量的大小。能量函数经过优化后如下:

$$\begin{aligned} E(\alpha, k, \theta, x, H) &= U(\alpha, k, \theta, z) + V(\alpha, z) + \rho [-H(\pi)] \\ &= \sum_n D(\alpha_n, k_n, \theta, z_n) + V(\alpha, z) + \rho \sum_{k=1}^K \pi_k \log(\pi_k) \end{aligned} \quad (15)$$

其中, ρ 表示惩罚因子。分割能量函数经过优化后, 不仅图像的细节信息可以最大限度地保留, 而且可以消除掉多余的高斯分量, 最终提高图像的分割精度。

3.2 分水岭算子的改进

经典的分水岭图像分割算法不仅可以较好地保存图像原始边缘的细节, 还可以让图像分割开的每个小区域之间的差异特别小。在实际应用过程中图像量化、图像高频信息和图像渐变部分所产生的影响很容易导致过分割的现象。针对这些问题, 本文提出使用多尺度形态的梯度算子替代单尺度的梯度算子来对图像进行预分割。以下为单尺度的梯度算子公式:

$$G(f) = [f \oplus B] - [f \odot B] \quad (16)$$

式中, $\oplus$ 表示膨胀, $\odot$ 表示腐蚀, B 代表单元结构元素。从式 (16) 可以看出, 单尺度形态的梯度算子的优劣由 B 的值决定。事实上单尺度形态的梯度算子在进行图像分割时还是会出现过分割的现象。为了解决这个问题, 文献[9]中定义的多尺度形态的梯度算子如下:

$$MG(f) = \frac{1}{n} \sum_{i=1}^n \{ [f \oplus B_i] - (f \odot B_i) \} \odot B_{i-1} \quad (17)$$

其中, 元素 B_i ($0 \leq i \leq n$) 为正方形结构, B_i 的尺寸为 $i \times i$ 。

3.3 改进的 GrubCut 算法

改进的 GrubCut 算法的具体步骤如下:

Step1 预处理。首先对输入服装使用二次分水岭的分割算法, 求出分割后得到的小区域颜色平均值, 并把它作为像素节点, 以便进行后续的处理。

Step2 初始化。初始化过程如下:

- 1) 用户初始化 T_B , 令 $T_F = \emptyset$, $T_U = \bar{T}_B$ 。通过设定位置区域 T_U 和背景区域 T_B , 把三元图 T 转换为二元图的标号问题, 以此对图像进行非完全编号。
- 2) 将背景像素透明度 α 置 0; 对未知的区域 T_U , 将 α 设置为 1。
- 3) 对于 $\alpha = 0$ 与 $\alpha = 1$ 这两个集合, 使用 k-means 聚类方法对背景和前景的 GMM 进行初始化, 以获取 GMM 的参数 (π_n, u_n, Σ_k) 初值。

Step3 最小化迭代。过程如下:

- 1) 计算出 x_n 的 GMM 参数 k_n , $k_n = \arg \min_{k_n} E(\alpha, m, \theta, x, H)$ 。
- 2) 所有像素 α 执行标记并优化, $\alpha_n = \arg \min_{\alpha} E(\alpha, m, \theta, x, H)$ 。
- 3) 从目标图像的节点求出 GMM 的参数 $\theta = \arg \min_{\theta} E(\alpha, m, \theta, x, H)$, 以此求出 GMM 参数 (π_n, u_n, Σ_k) 的估计值。这个过程不仅可以得到节点到背景和前景之间的距离 D , 同时还可以计算出基于混合比例熵的惩罚 $H(\pi)$ 。
- 4) 各个像素的 GMM 模型的数据项 $u(\alpha, m, \theta, x)$ 经过 EM 算法估计得到, 再通过熵惩罚获取索引号, 最后对其约束平滑项部分, 最终得到能量项 $E(\alpha, m, \theta, x, H)$ 。
- 5) 对 T_U 构造 S-T 的网络图, 采用 Max-Flow Min-Cut 算法初始化分割图像: $\min_{\{\alpha_n, x_n \in T_U\}} \min_{k_n} E(\alpha, m, \theta, x, H)$ 。
- 6) 重复执行 Step3 中的步骤 (1) 一步骤 (5), 直到收敛。
- 7) 执行边界的优化。

Step4 输出目标。分别获取 $\alpha = 0$ 与 $\alpha = 1$ 像素点, 把 $\alpha = 1$ 的像素输出, 即得到前景目标。

4 实验结果与分析

本文的所有实验都是在 CPU 为 Pentium(R) Dual-Core 2.60GHz、内存为 2GB、操作系统为 Windows7 的计算机上进行的。为了验证所提方法的有效性, 在 Matlab R2013a 上做了一些测试实验, 实验所用数据集是从淘宝、亚马逊、京东等电子商务网站上收集的。所选图像都是 256×256 的 JPG 格式, 包括短 T-恤、长 T-恤、短寸衫、长寸衫、Polo 衫, 每类含有 50 副, 共有 2000 幅图像。实验对比了 3 种不同检索方法: 基于文献[8]的图像检索方法, 基于文献[7]的图像检索方法, 以及本文提出的基于 GrabCut 改进的图像检索方法。

图 4—图 6 为这 3 种方法检索的结果对比, 其中的第一幅图像为检索的目标图, 后面 29 幅为检索结果中的前 29 幅服装图像。图 4 所示为使用文献[8]中算法检索的结果, 可以

看出从匹配结果 3 开始就出现了一些不希望得到的结果,检索目标图为长袖的衬衫,而匹配结果 3 为短袖衬衫,在 29 个检索结果中出现了 13 个不希望得到的结果。图 5 为文献[7]中算法检索得将结果,去除褶皱后,在 29 个检索结果中,从匹



图 4 文献[8]中算法检索结果



图 5 文献[7]中算法检索结果



图 6 改进后的 GrabCut 算法检索结果

检索策略的性能定性评价采用两个比较通用的评价准则:查准率和查全率^[10]。查准率为检索结果中检索出的相似图像数量占全部检索出的图像总数的百分比。查全率为检索结果中相似图像占图像库全部相似图像的百分比。定义如下:

$$\text{查准率} = \frac{\text{所有检测到的相关结果}}{\text{所有检索到的结果}} = \frac{S}{S+U} \quad (18)$$

$$\text{查全率} = \frac{\text{所有检测到的相关结果}}{\text{库中所有相关的结果}} = \frac{S}{S+V} \quad (19)$$

其中, S 为一次查询中检索到的相关图像的数目, U 为一次检索过程中检索到的不相关的图像数目, V 为图像库中与检索图像相关但未被检索到的图像数目。

为了定性评价本文提出的方法的有效性,从服装图像库中对每一种款式随机抽取 10 幅服装图像进行实验。

配结果 13 开始出现一些不匹配的服装,一共出现了 5 个不匹配的结果,检索的准确性有了很大的提升。图 6 为本文改进后的 GrabCut 算法检索所得结果,在 29 个检索结果中,只有匹配结果 22 不符合要求,检索的准确性有了更大的提升。

表 1 3 种算法查准率的比较

服装款式	长袖衬衫	长袖 T 恤	Polo 衫	短袖衬衫	短袖 T 恤
文献[8]	0.497	0.345	0.3975	0.415	0.615
文献[7]	0.525	0.585	0.43	0.505	0.695
本文方法	0.655	0.852	0.61	0.58	0.715

表 2 3 种算法查全率的比较

服装款式	长袖衬衫	长袖 T 恤	Polo 衫	短袖衬衫	短袖 T 恤
文献[8]	0.454	0.314	0.328	0.364	0.526
文献[7]	0.470	0.536	0.370	0.478	0.568
本文方法	0.572	0.738	0.536	0.542	0.574

表 1 和表 2 为采用文献[8]中算法、文献[7]中算法、本文改进后的 GrabCut 算法分别对长袖衬衫、长袖 T 恤、Polo 衫、短袖衬衫和短袖 T 恤的平均查准率和查全率的比较。从表 1 和表 2 中可以明显看出本文的检索方法比文献[8]中算法和

文献[7]中算法有更好的检索效果。

结束语 本文在分析传统 GrabCut 算法的基础上,对其能量函数进行了优化,对其分水岭算法算子进行了改进,提出了改进的 GrabCut 算法,将改进后的 GrabCut 算法引入基于内容的服装图像检索系统中,实验结果表明:本文方法与同类方法相比,在检索显示准确性以及检索的平均查准率和查全率方面均有明显提升。

参考文献

[1] Forczmański P, Czapiewski P, Frejlichowski D. Comparing Clothing Styles by Means of Computer Vision Methods[J]. Computer Vision and Graphics, 2014, 86(71):203-211

[2] Wu Xiao, Zhao Bo, Liang Ling-ling, et al. Clothing Extraction by Coarse Region Localization and Fine Foreground/Background Estimation[J]. Advances in Multimedia Modeling, 2013, 77(33):316-326

[3] 林瑶, 田捷. 医学图像分割方法综述[J]. 模式识别与人工智能, 2002, 15(2):192-204

[4] Mortensen E N, Barrett W A. Intelligent scissors for image composition[C]// Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, 1995:191-198

[5] 郝凯. 图像及图像序列上的交互抠图技术研究[D]. 南京:南京理工大学, 2012

[6] Falcao A X, Udupa J K, Samarasekera S, et al. User-steered image segmentation paradigms: Live wire and live lane[J]. Graphical Models and Image Processing, 1998, 60(4):233-260

[7] Rother C, Kolmogorov V, Blake A. GrabCut: interactive foreground extraction using iterated graph cuts[J]. ACM Transactions on Graphics, 2004, 23(3):309-314

[8] Lee T S. Image representation using 2D Gabor wavelets[J]. Pattern Analysis and Machine Intelligence, 1996, 18(10):959-971

[9] Arifina A Z. Image segmentation by histogram thresholding using hierarchical cluster analysis[J]. Pattern Recognition Letters, 2006, 27(13):1515-1521

[10] 韦娜, 耿国华, 周明全. 基于内容的图像检索系统性能评价[J]. 中国图形图像学报, 2004, 9(11):1271-1276

(上接第 207 页)

表 2 CNN 特征信息及检索结果

模型特征	Compressed Net	Compressed Net	AlexNet
	Fc6	Fc6	Fc6
Norm	No	Yes	Yes
dim	2048	2048	4096
PR(%)	32.31	34.84	33.68
RR(%)	86.95	87.1	85.13
AP(%)	73.50	78.66	71.41
Time(s)	8.7	7.8	60

表 3 检索结果

模型特征	Compressed Net Fc6	HOG	Proposed
PR(%)	34.84	24.2	36.4
RR(%)	87.1	69.4	94.32
AP(%)	78.66	56.45	90.45
Time(s)	7.8	3	11

结束语 鉴于治安卡口、及准卡口图像数据的不断增加,待检索图片的数量也在不断飙升,如何在更大的数据集上进行快速有效的个性化图像检索,以及其中检索的精度、检索耗时的提升、跨平台的特征匹配等问题都有待研究并解决。因此,本文提出了一种基于全局及个性化感兴趣区域特征的图像检索算法,通过对 AlexNet 优化加速得到 Compressed Net 的 fc6 层特征并结合多个个性化感兴趣区域的 HOG 特征,利用欧氏距离计算出待检索图库中与模板图像相似度最高的相关图像。通过对 50 辆目标车辆图像的检索验证,结果表明,该算法的检索率、平均准确率都达到 90%,呈现出较好的检索效果,计算速度快,具有较强的鲁棒性和实用性。

参考文献

[1] Smeulders A W M, Worring M, Santini S, et al. Content Based Image Retrieval at the End of the Early Years[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(12):1349-1380

[2] Zhou X S, Huang T S. Image Retrieval: Feature Primitives, Feature Representation, and Relevance Feedback[C]// Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries, 2000:10-14

[3] Breiteneder C, Eidenberger H. Content-Based Image Retrieval in Digital Libraries[C]// Proceedings of IEEE International Conference on Digital Libraries: Research and Practice, 2000:288-295

[4] Sun J D, Zhang X M, Cui J T, et al. Image Retrieval Based on Color Distribution Entropy[J]. Pattern Recognition Letters, 2006, 27(10):1122-1126

[5] Hinton G E, Qsindero S, Teh Y W. A Fast Learning Algorithm For Deep Belief Nets[J]. Neural Computation, 2006, 18(7):1527-1554

[6] Simonyan K, Zisserman A, Simonyan K, et al. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. Eprint Arxiv, 2014

[7] He K, Zhang X, Ren S, et al. Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 37(9):1904-1916

[8] Dong C, Chen C L, He K, et al. Learning a Deep Convolutional Network for Image Super-Resolution[C]// European Conference on Computer Vision, 2014:184-199

[9] Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[J]. Advances in Neural Information Processing Systems, 2012, 25(2):2012

[10] Lowe D. Distinctive Image features from scale-invariant key points[J]. International Journal of Computer Vision, 2004, 60(2):91-110

[11] Bay H, Tuytelaars T. SUFR: Speeded Up Robust Features[C]// Proc. of European Conference on Computer Vision, 2006:404-417

[12] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA, 2005:886-893

[13] Manjunath B S, Ohm J R, Vasudevan V V, et al. Color and texture descriptors[J]. IEEE Trans. Circ. Syst. Video Technol., 2001, 11(6):703-715

[14] Datta R, Joshi D, Li J, et al. Image Retrieval: Ideas, Influences, and Trends of the New Age[J]. ACM Computing Surveys, 2008, 40(2):1-60