

基于特征融合的 WCE 图像检索

邹月娴¹ 霍佳森¹ 刘吉¹ 李奕² 邓文军²

(ELIP/ADSPLAB 北京大学深圳研究生院 深圳 518000)¹

(深圳市资福技术有限公司 深圳 518057)²

摘要 无线胶囊内镜(WCE)是一种新型的、具有极大市场价值的肠道疾病检测技术。该技术产生的图像数据量极大,为医生开展疾病诊断带来了困难。WCE 图像快速检索技术是 WCE 后端智能图像分析软件的基本功能之一。本文针对 WCE 图像中相似图像的快速检索问题开展研究。通过对大量 WCE 图像本身特征的分析,提出了一种融合颜色特征与纹理特征的图像检索算法。首先分别提取图像的色调饱和度二维相关直方图以及图像的局部二值模式(LBP)纹理特征,然后采用高斯归一化的方法对不同的特征进行融合,最后采用融合特征为待检索图像建立索引,实现了输入目标 WCE 图像的快速准确查找。实验结果表明本文所提方法计算复杂度低和检索结果良好。

关键词 基于内容的图像检索,无线胶囊内镜图像,色调饱和度相关直方图,LBP 纹理特征,高斯归一化,特征融合
中图分类号 TP391 **文献标识码** A

WCE Image Retrieval by Using Fused Feature

ZOU Yue-xian¹ HUO Jia-sen¹ LIU Ji¹ LI Yi² DENG Wen-jun²

(ELIP/ADSPLAB, Peking University Shenzhen Graduate School, Shenzhen 518000, China)¹

(ShenZhen Jifu Technology Ltd, Shenzhen 518057, China)²

Abstract Wireless capsule endoscopy (WCE) is a new technique for detecting the gastrointestinal diseases which has a great market value. This technique will release a mass of image data and lead a big challenge for doctors to review the WCE images. Hence, WCE image retrieval is a fundamental function of the WCE image analysis system. This paper investigates the WCE image retrieval problem by using the similarity information of the WCE images. With the analysis of the property of WCE image, we develop an image retrieval algorithm by fusing the color feature and texture feature. Firstly, the color feature hue-saturation correlation histogram and texture feature LBP (local binary pattern) have been extracted; Then the color and texture features are fused by Gaussian normalization; Finally the index of the WCE image dataset can be built accordingly using the resultant fused feature vector. The experiments show that the proposed method has low computational complexity and a good retrieval result.

Keywords Content-based image retrieval, Wireless capsule endoscopy, Hue-Saturation correlation histogram, LBP, Gaussian normalization, Fused feature

1 引言

消化道疾病是一种高发病症,但传统的疾病诊断方法痛苦大且无法诊断小肠部位。胶囊内镜图像(Wireless Capsule Endoscopy)^[1]简称 WCE,是一种新型的消化道疾病医疗诊断技术,如图 1 所示。患者吞服胶囊,胶囊在人体肠胃蠕动的作用下向前运动并连续对人体消化道壁进行拍摄,所拍摄的图片通过无线通信方式传输到体外接收设备,最终将其下载到电脑上由阅片医生进行检查并作出诊断。由于胶囊在患者体内停留 8 个小时左右,将产生 50000 张以上的 WCE 图片。

WCE 技术具有无痛苦、免侵入等一系列优点,因此有很好的市场前景,但是每个病例产生的 WCE 图片数据量过大,医生一般需要 3~4 个小时才能完成一个病例的诊断,造成极为昂贵的医生成本,直接限制了 WCE 技术的推广。采用图

像处理和机器学习技术来提高 WCE 图片质量、消除 WCE 冗余图片,进行快速检测等成为 WCE 技术的关键核心技术和研究重点。其中,香港大学孟庆虎团队针对图像三维重建^[2]、视频摘要^[3]和病灶自动检测^[4]等技术开展研究;伊利诺斯大学的 Sae Hwang 团队在聚类码本^[5]、器官分割^[6]以及出血检测^[7]方面开展了相关研究;东英格兰大学的 Michal Mackiewicz 团队专注于 WCE 图像分类的理论研究,在 WCE 图像序列器官分段和异常图片检测方面取得进展,如文献[8]中采用离散余弦变换和主成分分析的方法提取图像特征并通过 SVM 分类器和 K-NN 分类器分别对特征进行分类判断。

图像检索技术是一个重要的研究热点,在医学领域有重要的应用价值,医学图像往往具有高分辨率、高精度和大数据量的特点,通过典型图像快速查找其相似图像,将极大方便医生进行医疗诊断。

本文受深圳市科技创新委员会项目(JCYJ20130329175141512),公司合作项目资助。

邹月娴(1964-),女,博士,教授,主要研究方向为模式识别、信号与信息处理, E-mail: zouyx@pkusz.edu.cn; 霍佳森(1987-),男,硕士,主要研究方向为图像分析。



图1 胶囊内镜

目前已经有许多针对医学图像的检索方法研究成果,在文献[9]介绍的针对不同医学图像的检索方法中,包括针对胸透图像使用的小波多尺度特征,针对骨骼关节图像的聚类结构码本特征以及针对细胞显微镜图像的显著纹脉特征。卡内基梅隆大学的 Jia-Yu 提出的 ViVo 检索码本^[10]对 X 光胸透图像具有良好的检索效果。

根据我们的调研显示,迄今尚未见专门针对 WCE 图像检索的技术报道。我们的初步研究结果表明,采用著名的 sift 和 surf 特征并不能取得满意的检索结果,如对于同一位置的胃壁发生收缩、伸张所对应的 WCE 图像,其 surf 特征有明显的不同,如图 2 所示,没有很好地反映 WCE 图像的相似性。此例说明,局部不变特征针对的是图像中仿射不变的斑块特征区域,对于 WCE 图像检索并不适用。

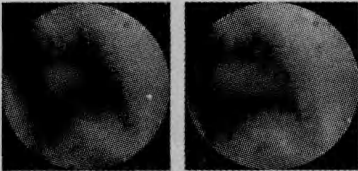


图2 相似图像 surf 特征提取对比(其中图像方块标识处为图像的 surf 特征区域)

本文开展 WCE 图像检索研究工作,针对 WCE 图像所具有的独特性质提出一种基于融合特征的图像检索方法,即提取颜色特征色调饱和度相关直方图和纹理特征 LBP 算子,采用高斯归一化方法调整特征的尺度,通过加权对特征进行有效融合,实验证明本文提出的算法具有很好的效果。

2 图像特征提取算法

2.1 颜色特征

颜色特征是一种重要且有效的图像特征表示。颜色特征具有显著、稳定性好及良好的鲁棒特性。由图 3 所示的 WCE 图像,可以发现不同器官的 WCE 图像颜色特征区分性显著,主要体现在色调与饱和度的差异。通常 WCE 图像是 RGB 格式,为了提取色调和饱和度,可将 WCE 图像从 RGB 空间转换到 HSV 空间,在此基础上,采用色调(Hue)饱和度(Saturation)相关直方图特征来表示 WCE 图像的颜色特征。

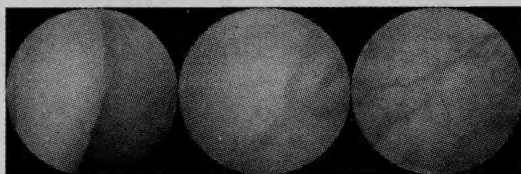


图3 WCE 图片示意图

具体方法描述如下:首先,对每一幅 WCE 图像进行 RGB 空间到 HSV 空间的变换,从而得到色调单分量的图像和饱和度单分量的图像,如式(1)和式(2)中所示。

$$s = \begin{cases} 0, & \text{if } \max(r, g, b) = 0 \\ 1 - \frac{\min(r, g, b)}{\max(r, g, b)}, & \text{otherwise} \end{cases} \quad (1)$$

$$h = \begin{cases} \text{undefined}, & \text{if } \max(r, g, b) = \min(r, g, b) \\ \frac{60^\circ \times (g-b)}{\max(r, g, b) - \min(r, g, b)} + 0^\circ, & \text{if } \max(r, g, b) = r, g > b \\ \frac{60^\circ \times (g-b)}{\max(r, g, b) - \min(r, g, b)} + 360^\circ, & \text{if } \max(r, g, b) = r, g < b \\ \frac{60^\circ \times (b-r)}{\max(r, g, b) - \min(r, g, b)} + 120^\circ, & \text{if } \max(r, g, b) = g \\ \frac{60^\circ \times (r-g)}{\max(r, g, b) - \min(r, g, b)} + 240^\circ, & \text{if } \max(r, g, b) = b \end{cases} \quad (2)$$

式(1)和式(2)中, r, g, b 分别代表一个像素的红、绿、蓝 3 个分量。 h 与 s 分别表示一个像素的色调、饱和度分量。根据式(1)和式(2)获得的 h 和 s , 计算二维相关直方图:首先将色调(h)区间和饱和度(s)区间分成若干子区段。本研究中,将色调的原始区段 $[0-180]$ 平均分为 4 个子区段:分别为 $[0-44]$ 、 $[45-89]$ 、 $[90-134]$ 和 $[135-180]$,记做 H_i ($i \in [1, 2, 3, 4]$)。饱和度的原始区段为 $[0-255]$,同样将饱和度区段平均分为 4 个子区段,分别为: $[0-63]$ 、 $[64-127]$ 、 $[128-191]$ 和 $[192-255]$,记做 S_j ($j \in [1, 2, 3, 4]$)。将色调区间与饱和度区间建立笛卡尔二维平面坐标,形成 16 个二维子区间 $\{H_i, S_j\}$,例如, $\{[0-44], [64-127]\}$ 指的是色调值落在子区间 $[0-44]$ 且饱和度值落在区间。根据色调单分量图像与饱和度单分量图像,计算相应的二维直方图,表示为 HS_{ij} ,将二维直方图展开得到颜色特征定义为 F_c 。

2.2 纹理特征提取

图像纹理特征反映的是图像的一种局部结构化特征,可以认为是灰度或颜色在空间以一定的形式变化而产生的模式。常见的纹理分析方法包括:统计方法、结构方法、变换域方法。LBP(Local Binary Pattern, 简称 LBP)^[11]方法是一种基于局部统计信息的纹理表示方法。本研究采用局部二值模式的 LBP 纹理特征实现 WCE 图像纹理特征的提取。分析 WCE 图像,我们发现 WCE 图像反映了胃肠壁表面结构的分布信息,包括了肠壁褶皱、斑块与黏着物等。采用 LBP 算子能够有效地提取 WCE 图像的纹理特征。为了降低计算复杂度,我们计算 LBP 的统计直方图来描述一幅 WCE 图像的纹理分布。

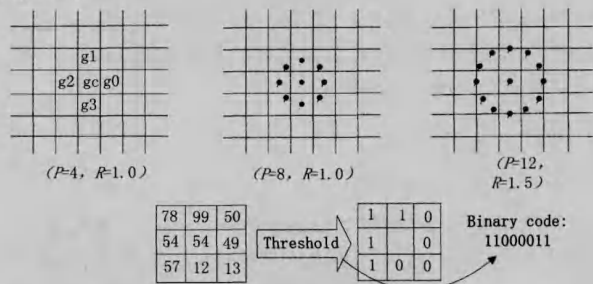


图4 LBP 算子提取示意图

WCE 图片的 LBP 纹理特征的计算方法具体描述如下:计算 WCE 的灰度图像,取一个像素点(中心像素点)半径 R 范围内的像素点为参照像素点,根据参照像素点与中心像素点的比较值确定二值编码 S_p ,若参照像素点的值大于中心像

素点,则编码值为 1,若参照像素点的值小于中心像素点,则编码值为 0,基本思想如图 4 所示。显然,LBP 方法需要选择参照像素个数与参照像素半径。本研究采用 8 个参照像素且参照像素半径为 1。

根据上述方法确定 WCE 图像二值编码值,则 WCE 图像 I 中的每一个像素的 LBP 值可以计算如下:

$$I_{LBP}(i,j) = \sum_{p=0}^{P-1} s_p 2^p, i=1, \dots, K, j=1, \dots, L \quad (3)$$

式中, $I_{LBP}(i,j)$ 代表图像 I 的第 i 行第 j 列的像素所对应的 LBP 值, s_p 为第 p 个位置上的 LBP 编码值, K 和 L 分别为 WCE 图像 I 的维度。通过式(3)的计算,我们获得了 WCE 图片的灰度图像的 LBP 编码图像。计算 LBP 编码图像的统计直方图作为 WCE 图像的纹理特征向量,如图 5 所示。



图 5 原始图像、LBP 编码与纹理特征直方图

3 特征融合算法与相似度比较

研究表明,采用 WCE 图像的单一特征往往很难完整地表示一幅 WCE 图像的全部信息,基于单一特征的 WCE 图像检索其性能受到局限。采用特征融合的思想,本文研究融合 WCE 图像的纹理特征与颜色特征来有效表示 WCE 图像特征。基于特征融合的思想首先需要考虑不同的特征具有不同的尺度范围,在不同的特征空间反映图像的重要性不同的情况。本文采用高斯归一化^[12]方法将 WCE 图像特征和颜色特征进行融合。

基于原始图像提取的颜色特征与纹理特征之间的数值尺度存在较大的差异性。例如,若一个特征 F_c 的取值范围在 10~100 范围,而另一个的取值则在 100~1000 之间。基于直接融合特征的方法,则第二个特征的作用要比第一个特征显著。针对此问题,我们提出基于高斯归一化的特征融合算法。具体方法描述如下:

用 I_1, \dots, I_M 表示 WCE 图像库中的 M 张 WCE 图像,则第 i 幅 WCE 图像可以用提取的 N 维特征向量 F_i 来表示,其中, $F_i = [f_{i1}, \dots, f_{iN}]$ 。则基于 M 幅 WCE 图像提取的特征可以构成一个 $M \times N$ 的特征矩阵 $F = \{f_{ij}\}, i=1, \dots, M, j=1, \dots, N$,且 f_{ij} 表示第 i 幅图像的特征向量的第 j 分量。每一维上的均值可以通过式(4)计算,每一维上的方差可以通过式(5)计算,形成 M 张 WCE 图像特征的均值向量 $\mu = [\mu_1, \dots, \mu_N]$ 以及方差向量 $\sigma = [\sigma_1, \dots, \sigma_N]$ 。

$$\mu_j = \sum_{i=1}^M f_{ij} / M \quad (4)$$

$$\sigma_j = \sum_{i=1}^M (f_{ij} - \mu_j)^2 / M \quad (5)$$

对于一幅图像 I 提取的原始颜色特征(为 N_c 维向量)表示为 $F_c = [c_1, c_2, \dots, c_{N_c}]$,原始纹理特征(为 N_t 维向量)表示为 $F_t = [t_1, \dots, t_{N_t}]$,对图像特征进行高斯归一化如式(6)所示:

$$F_c^{norm} = (F_c - \mu_c) / \sigma_c$$

其中, $c_j^{norm} = (c_j - \mu_{c_j}) / \sigma_{c_j}, j=1, \dots, N_c$

$$F_t^{norm} = (F_t - \mu_t) / \sigma_t, \quad (6)$$

其中, $t_j^{norm} = (t_j - \mu_{t_j}) / \sigma_{t_j}, j=1, \dots, N_t$

式中, c_j 与 t_j 为对应的第 j 维颜色特征和纹理特征分量。 F_c^{norm} 与 F_t^{norm} 为高斯归一化后的颜色和纹理特征向量, c_j^{norm} 与 t_j^{norm} 为相应的第 j 维分量, μ_c 与 σ_c 为图像颜色特征的均值向量和方差向量, μ_{c_j} 与 σ_{c_j} 为对应的第 j 维分量, μ_t 与 σ_t 为图像纹理特征的均值向量和方差向量, μ_{t_j} 与 σ_{t_j} 为对应的第 j 维分量。

设查询图像与待检索图像分别为 Q 和 I ,提取 Q 和 I 图像的颜色特征与纹理特征并进行高斯归一化,得到 Q 的颜色特征向量表示为 $F_c^Q = [c_1^Q, c_2^Q, \dots, c_{N_c}^Q]$,纹理特征向量表示为 $F_t^Q = [t_1^Q, t_2^Q, \dots, t_{N_t}^Q]$ 。同理,得到图像 I 的颜色特征向量 $F_c^I = [c_1^I, c_2^I, \dots, c_{N_c}^I]$,纹理特征向量为 $F_t^I = [t_1^I, t_2^I, \dots, t_{N_t}^I]$ 。因此,定义 Q 和 I 图像的归一化颜色特征向量之间的距离为:

$$Dist_{color}(Q, I) = \sum_{i=1}^{N_c} |c_i^Q - c_i^I| \quad (7)$$

同理, Q 和 I 图像的归一化纹理特征向量之间的距离定义为:

$$Dist_{texture}(Q, I) = \sum_{i=1}^{N_t} |t_i^Q - t_i^I| \quad (8)$$

式(7)和式(8)中特征分量皆为图像 Q 和 I 的归一化颜色特征分量和归一化纹理特征分量。

本文采用的融合策略不是在特征层进行融合,而是在决策层的特征融合。本文提出采用加权融合方法,即将颜色特征距离与纹理特征距离进行加权求和作为两幅特征之间相似性的判据,具体表示为:

$$Dist(Q, I) = w_1 Dist_{color}(Q, I) + w_2 Dist_{texture}(Q, I) \quad (9)$$

式中, w_1 与 w_2 分别表示颜色距离分量与纹理距离分量的权值, $w_1 + w_2 = 1$ 。如何选择权值将通过实验来说明。

本文提出的算法为基于特征融合的图像检索算法(Feature Fusion based Image Retrieval,简称 FF-IRA),算法总结如下:

基于特征融合的图像检索算法(FF-IRA)

输入:包含 M 张图像的待检索图像集合 $D = \{I_1, I_2, \dots, I_M\}$,待查询图像 I_{input}

1. 对图像集 D 中的每一幅图像完成下面的计算
 - a) 利用式(1)和式(2)计算图像 I_i 的色调和饱和度
 - b) 计算图像的颜色特征, F_{c_i} 表示是第 i 幅图像
 - c) 计算图像的纹理特征:由式(3)计算图像的 LBP 编码,对 LBP 编码图像计算统计直方图,作为图像纹理特征,记为 F_{t_i} , i 表示是第 i 幅图像
 - d) 采用式(6)对图像颜色特征和纹理特征向量进行归一化,获得其归一化特征向量,记为 $F_{c_i}^{norm}$ 与 $F_{t_i}^{norm}$
2. 根据上述计算方法,计算待查找图像 I_{input} 的归一化颜色特征和纹理特征,记为 $F_{c_{input}}^{norm}$ 与 $F_{t_{input}}^{norm}$
3. 根据式(7)计算待测试图和 D 中 M 幅图的颜色特征距离
4. 根据式(8)计算待测试图和 D 中 M 幅图的纹理特征距离
5. 根据式(9)计算待测试图和 D 中 M 幅图的融合特征距离
6. 将所有计算的按照由小到大进行排序(距离最小者为与 I_{input} 最相似的图片)

返回结果。

算法结束

4 实验及结果评价

为了评估本文提出的算法的性能,进行以下实验。实验采用挑选处的图像共 1000 张。从中选择 800 张 WCE 图像作为待检索图片集,其中口腔食道图像、胃部图像、小肠图像和大肠图像各 200 张。采用 200 张 WCE 图片作为查询图像,其

中口腔食道图像、胃部图像、小肠图像以及大肠图像各 50 张。采用的实验环境为：操作系统为 Windows7, CPU 为 Intel(R) Core(TM)2 Duo CPU T9400 2.53GHZ, 内存为 4.00GB, 在 Visual Studio2008 平台上使用 c++ 实现本算法。

常用的检索结果评价方法使用召回率(R)和精度(P)作为评价标准, 召回率表示检索结果返回的相关图片的比例, 精度是检索结果的纯净程度, 如图 6 所示, 对于一幅图像可以在客观上将其分为相关与不相关, 而根据算法返回的结果可以分为检索到与没有检索到, 因此可以获得一个交叉矩阵将图像分为 4 种类型。



图 6 召回率与精度示意图

因此召回率的定义就是检索到的相关图像与相关图像的比:

$$R = \frac{A}{A+C} \quad (10)$$

而精度的定义就是检索到的相关图像与检索到的全部图像的比:

$$P = \frac{A}{A+B} \quad (11)$$

式中, A 为检索结果中相关图像的数量, A+C 为全部相关图片的数量, A+B 为检索返回的全部结果的数量。

此外, 考虑到医生查询 WCE 相似图片最为关注的是第一页返回的前十个结果, 即可采用 P@10 的评价方法。P@10 方法^[13]是根据检索对结果排在前十的结果进行评价的一种方法, 通过计算前十个返回结果的精度来评估检索的结果, 本文采用前十个结果每一个排名位置的精度的平均值来进行检索算法性能评估:

$$P_{mean} = \sum_{i=1}^{10} P_i \quad (12)$$

式中, P_i 为前 i 个检索结果中符合查询结果图像个数与 i 的比值, 即前 i 个检索结果的精度。

本文进行两组实验: 第一组实验是测试 FF-IRA 算法中权值对 WCE 图像检索性能的影响, 并通过实验获得最佳权值; 第二组实验则是对比本文提出的基于特征融合的 WCE 图像检索方法与单纯使用纹理或单纯使用颜色特征进行 WCE 图像检索的方法的性能。

实验 1 在 200 张待检索图像中随机挑选 30 张胃部图像和 30 张小肠图像作为查找图像, 变动 w_1 与 w_2 的权值比例由 (0.1, 0.9) 到 (0.9, 0.1), 统计每张图像查询返回结果的 P@10 精度, 统计加权平均值结果如表 1 所列。

表 1 不同特征权值条件下的 Pmean

组号	W1	W2	胃部图像精度	小肠图像精度
1	0.1	0.9	0.895	0.882
2	0.2	0.8	0.833	0.848
3	0.3	0.7	0.897	0.893
4	0.4	0.6	0.928	0.925
5	0.5	0.5	0.945	0.946
6	0.6	0.4	0.957	0.952
7	0.7	0.3	0.912	0.895
8	0.8	0.2	0.897	0.875
9	0.9	0.1	0.865	0.861

由表 1 的实验室结果可见, 当权值 w_1 与 w_2 变化时, 检索结果的精度也随之发生变化。当特征权值 w_1 与 w_2 分别取 0.6 与 0.4 时, 检索精度达到最大值。

实验 2 本实验采用包含 4 类图像的全部 200 张 WCE 图像作为查询图像。对于本文提出的基于特征融合的图像检索(FF-IRA)算法, 根据实验 1 中得到的结果, 设定特征权值 w_1 与 w_2 为 0.6 与 0.4。图像检索算法性能对比算法包括仅采用纹理特征的图像检索算法(Texture-Feature based image Retrieval Algorithm; TF-IRA)以及仅采用基于颜色特征(Color Feature-based Image Retrieval Algorithm; CF-IRA), 实验结果如表 2 所列。

表 2 3 种图像检索算法的结果对比

	TF-IRA	CF-IRA	FF-IRA
食道	0.834	0.882	0.933
胃	0.855	0.895	0.957
小肠	0.853	0.875	0.952
大肠	0.875	0.865	0.926
平均	0.854	0.879	0.942

由表 2 可知, 本文提出的 FF-IRA 算法在测试图像集上的平均检索精度达到了 94.2%, 而 TF-IRA 算法的平均检索精度为 85.4%, CF-IRA 算法的平均检索精度为 87.9%。可以认为本文提出的基于融合纹理与颜色特征的 FF-IRA 算法获得了较为满意的检索结果。

结束语 本文提出了一种基于颜色特征与纹理特征相融合的图像检索算法, 该算法采用高斯归一化方法对颜色特征和纹理特征进行归一化处理, 并通过不同权值进行特征融合, 算法的特征维数低, 计算复杂度低。基于 1000 张 WCE 图像测试集的实验结果表明, 本文提出的算法在召回率和检索精度性能上都获得了较好的结果, 为解决 WCE 相似图片检索提供了方法。

参考文献

- [1] Swain P. Wireless Capsule Endoscopy[J]. Gut, 2003, 52: 48-50
- [2] Fan Y, Meng M Q H. 3D Reconstruction of WCE images by affine SIFT method[C]//Proc. 9th World Congr. Intell. Control Autom. 2011; 943-947
- [3] Meng M Q H, Zhao Qian. A Strategy to Abstract WCE Video Clips Based on LDA[C]//International Conference on Robotics and Automation, May 2011; 454-459
- [4] Chen D, Meng M-H, Wang H, et al. A novel strategy to label abnormalities for wireless capsule endoscopy frames sequence [C]//International Conference on Information and Automation, ICIA. 2011; 379-383
- [5] Hwang, Sae. Bag-of-Visual-Words Approach to Abnormal Image Detection in Wireless Capsule Endoscopy Videos. Advances in Visual Computing [M]. Berlin Heidelberg: Springer, 2011; 320-327
- [6] Hwang S, Celebi M. Multilevel Wireless Capsule Endoscopy Video Segmentation[C]//Proc. of SPIE Vol. Vol. 7623. 2010
- [7] Hwang, Sae, et al. Blood Detection in Wireless Capsule Endoscopy Using Expectation Maximization Clustering[M], Medical Imaging, International Society for Optics and Photonics, 2006
- [8] Mackiewicz M, Berens J, Fisher M. Wireless Capsule Endoscopy Video Segmentation Using Support Vector Classifiers And Hid-

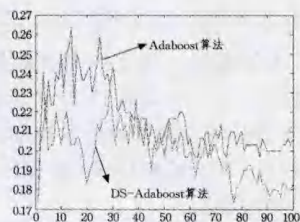
den Markov Models[C]//Proc. International Conference Medical Image Understanding and Analyses, June 2006

- [9] Muller H, Michoux N, Bandon D, et al. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions[J]. International Journal of Medical Informatics, 2004, 73: 1-23
- [10] Bhattacharya A, Ljosa V, Pan Jia-yu, et al. ViVo: Visual Vocabulary Construction for Mining Biomedical Images[C]//Proceedings of the Fifth IEEE International Conference on Data Mining (ICDM), 2005

- [11] Ojala T, Pietikainen M. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-987
- [12] 周明全, 耿国华, 韦娜. 基于内容图像检索技术[M]. 北京: 清华大学出版社, 2007
- [13] 陆建江, 张亚非, 徐伟光. 智能检索技术[M]. 北京: 科学出版社, 2009

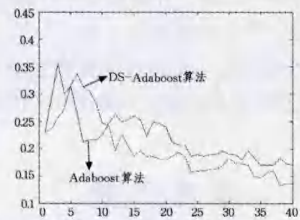
(上接第 319 页)

数据集, 其中, 训练样本中包括 200 个人脸样本和 400 个非人脸样本, 测试样本中包括 200 个人脸样本和 400 个非人脸样本, 样本全部归一化为 20×20 像素。通过训练得到由 40 个弱分类器组成的强分类器, 对测试集合进行测试。图中的曲线表示随着弱分类器个数的增加, 分类器对测试集合的错误率波动曲线。



横坐标: 弱分类器个数 纵坐标: 检测率(%)

图 3 检测率与弱分类器数目的关系图



横坐标: 弱分类器个数 纵坐标: 检测率(%)

图 4 检测率与弱分类器数目的关系图

从上面的两个波动曲线图可以清楚地看到, 绿色曲线代表的是 DS-Adaboost 算法, 蓝色曲线代表连续 Adaboost 算法, 改进的算法随着分类器个数的增加, 错误率下降得更快。最后利用网上下载的 MIT 样本库中的 2706 个人脸样本和 4481 个非人脸样本作为实验数据集, 其中训练样本集合包括 1500 个人脸样本和 3000 个非人脸样本, 训练得到由 100 个弱分类器组成的强分类器, 对测试集合进行了检测, 结果如表 1 所列。

表 1 DS-Adaboost 算法与连续 Adaboost 算法的实验结果对比

方法	检测率	误检率
连续 Adaboost 算法	84.66%	2.10%
DS-Adaboost 算法	86.57%	1.32%

4.3 算法的局限性

由上面的实验结果对比图可以很清晰地看出本文的算法不但在检测率方面有一定的提高, 在误检率方面也有大幅度地降低。所以从总体上来说本文所改进的算法取得了比较好的结果, 跟我们前面所分析的基本吻合, 但是依然没有解决训练时间的问题, 本文和其它同类型的检测方法一样, 依然要花

很大一部分时间在训练弱分类器上面, 这也正是这种算法的局限性, 如何能够有效地提高训练速度才是至关重要的一个环节, 就目前来看只能寄希望于好的电脑配置来解决这个问题。

4.4 人脸检测系统

本文利用 matlab 语言构建了一个由 15 层强分类器组成的级联分类器的人脸检测系统, 每层强分类器的误检率不超过 0.5, 检测率不低于 0.95, 最后我们训练的级联分类器的最终检测率为 $0.95^{15} \approx 0.928$, 误检率为 $0.5^{15} \approx 0.003\%$ 。

图 5—图 7 是一些检测结果的效果图。

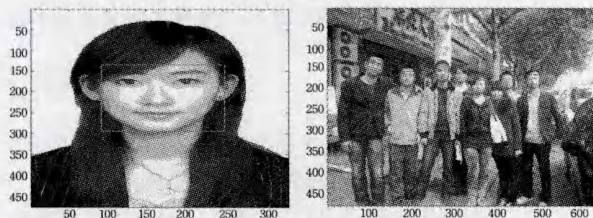


图 5

图 6

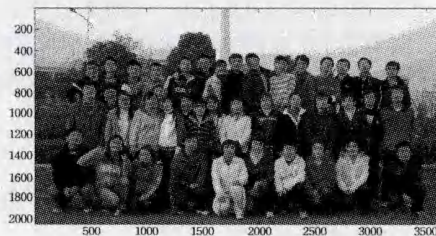


图 7

结束语 针对连续 Adaboost 算法中平滑因子选取的不足: 本文提出了一种动态选取平滑因子的 DS-Adaboost 算法, 该算法对弱分类器输出中的平滑因子 ϵ 进行了动态选取, 根据 $\frac{W_{t+1}^+}{W_{t+1}^-}$ 比值的大小动态地选择平滑因子。实验表明, DS-Adaboost 算法能较好地起到平滑作用, 使得在同一区间里面的正负样本比例都在可以比拟的范围内。

参考文献

- [1] 李月敏, 陈杰, 高义, 等. 快速人脸检测技术综述[C]//全国 16 届计算机科学与技术应用学术论文集. 2004: 771-774
- [2] Viola P, Jones M J. Robust Real-Time Face Detection[J]. International Journal of Computer Vision, 2001, 57(2): 137-154
- [3] Freund Y, Schapire R E. Experiments with a New Boosting Algorithm[J]. IEEE Computer Society, 1996, 43(4): 148-156
- [4] Schapire R E, Singer Y. Improved boosting algorithms using confidence-rated predictions[C]// Proceedings of 11th Annual Conference on Computational Learning Theory. 1998: 80-91