

一种基于 SIFT 算子的人脸识别方法

周志铭 余松煜 张瑞 杨小康

(上海交通大学电子工程系图像通信与信息处理研究所, 上海 200240)

(上海交通大学上海市数字媒体处理与传输重点实验室, 上海 200240)

摘要 高独特性特征的选择以及合适匹配策略的选用是人脸识别技术的关键。讨论了基于仿射不变的几何特征 SIFT 算子进行人脸识别的方法。SIFT 算子的计算复杂度较高, 并且不同的人脸表情和图像模糊会加大特征匹配的难度。为克服上述缺点, 提出了一种新的算法, 将选择 6 个人脸上感兴趣子区域进行描述, 并根据各自的独特性赋予不同的权值, 最后在匹配过程中使用相似度的平方来减小偏差数据造成的影响。实验结果表明, 该方法能有效减轻表情变化对于身份识别率急剧下降的影响, 并可显著减少计算复杂度和特征匹配时间。

关键词 独特性 人脸识别 SIFT 算子 感兴趣子区域

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(2008)10-1882-04

An Algorithm for Face Recognition Based on SIFT Descriptor

ZHOU Zhi-ming, YU Song-yu, ZHANG Rui, YANG Xiao-kang

(Institute of Image Communication and Information Processing, Department of Electronic Engineering,
Shanghai Jiaotong University, Shanghai 200240)

(Shanghai Key Laboratory of Digital Media Processing and Transmissions, Shanghai Jiaotong University, Shanghai 200240)

Abstract Choosing a distinctive feature and matching criterion is key to developing a reliable face recognition system. This paper discusses the availability of one of geometric feature invariants, scale invariant feature transform (SIFT) descriptor based face recognition. The SIFT feature description of an image is typically complex. In most cases, the difficulty of feature matching problem is aggravated when the different face expressions and image blur exist. For abovementioned issues, in this paper we proposes a new method that six interest sub-regions from the face are selected to be described and later be calculated through different weights according to their distinctiveness. The square of the similarity is used to solve the problem of data deviation. The experimental results demonstrate that our method does effectively moderate the face expression effect. It also successfully reduces the complexity and matching time of SIFT feature sets.

Keywords distinctive, face recognition, SIFT(scale invariant feature transform) descriptor, sub-interest region

1 引言

人脸识别是当今模式识别和人工智能领域的一个活跃的研究方向。而人脸识别技术也已经被运用到犯罪识别、视频监控、档案管理、视频会议和照片检索等各个领域, 如图1所示。

人脸识别的研究内容主要包括以下几个方面^[1]:

- (1) 人脸检测: 在输入的图像中寻找人脸区域。
- (2) 人脸的规范化: 校正人脸在尺度、光照和旋转等方面的变化。
- (3) 特征提取: 从人脸图像中映射提取一组反映人脸特征的数值表示样本。
- (4) 特征匹配: 将待识别人脸与已知人脸比较,

基金项目: 国家自然科学基金委项目(60502034, 60625103); 国家 863 计划项目(1006AA01Z124)

收稿日期: 2008-07-12; 改回日期: 2008-07-20

第一作者简介: 周志铭(1984 ~), 男。上海交通大学电子工程系图像通信与信息处理研究所硕士研究生。主要研究领域为图像内容分析, 以及基于几何特征不变量的图像检索和物体识别。E-mail: babazhouzhiming@sjtu.edu.cn

得出相关信息。

当人脸的光照、表情发生比较大的变化,或者图像存在部分模糊的情况时,其识别准确率会明显降低,这是人脸识别中需要克服的一大技术难题。

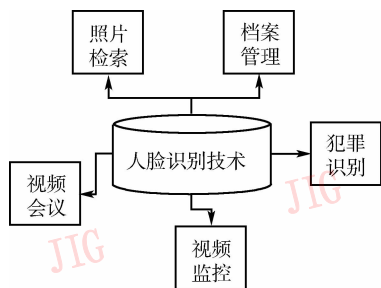


图 1 人脸识别技术的运用领域

Fig. 1 Application fields of face recognition

目前,一些研究者尝试将刚性物体识别领域中有

着良好运用的几何特征不变量逐步引入人脸识别领域,其中 SIFT (scale invariant feature transform) 几何特征不变量尤为广泛应用^[2]。

本文提出了一种基于 SIFT 算子进行人脸识别的方法。实验结果表明,该算法能有效减轻表情变化对于身份识别率急剧下降的影响,并可显著减少计算复杂度和特征匹配时间。

2 基于 SIFT 算子的人脸识别算法

本文先对目标图像进行人脸检测,再通过定位五官选定感兴趣的子区域。在 SIFT 算子对相关区域进行特征提取及描述后,通过向量间的相似度测量,与不同人脸所提取的特征向量进行比较,从而确定目标图像的人脸身份。图 2 给出了算法步骤的示意图。

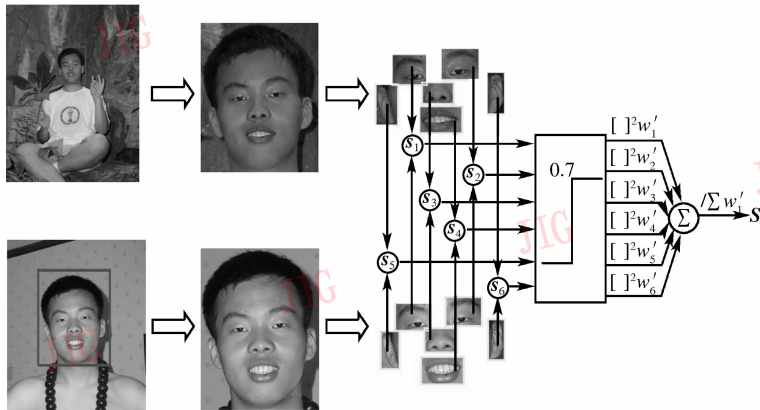


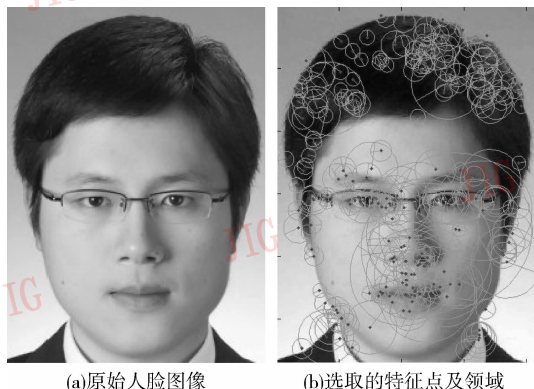
图 2 人脸识别算法步骤

Fig. 2 The steps of face recognition algorithm

目前,人脸检测算法发展得比较成熟,可以使用 OpenCV 自带的 Adaboost 和 Haar 特征对图像中的人脸区域进行提取^[3]。在检测出人脸区域后,再进行人脸的特征提取、子区域选择和匹配。

2.1 人脸特征提取

本文选择 SIFT 算子做人脸图像的特征提取。SIFT 算法在不同尺度空间上搜索稳定的特征点,然后每一特征点的邻域由一个 128 维的矢量来描述,其中每 1 维对应着一个特征方向统计量^[4],如图 3 所示,图 3(a)为原始人脸图像,图 3(b)中的白色圆圈表示了特征点的提取以及领域大小的选择。SIFT 算法检测出的特征点及其描述子能够在缩放和旋转变换下保持不变,并对光照变化不敏感,适于解决人脸图像中在尺寸角度光照情况有变化情况下的识别问题。



(a)原始人脸图像 (b)选取的特征点及领域

图 3 SIFT 算子提取特征

Fig. 3 Feature extracted by SIFT method

一幅人脸图像在经过提取特征点,对邻域进行描述后,可以得到一个 $n \times 128$ 维的矩阵 A ,其中 n 为

总共提取描述的特征点数目。一幅 200×200 左右大小的人脸,通常可以从中提取 300 个左右的特征点。

2.2 人脸子区域选择

SIFT 算法在提取匹配基元时必须遍历尺度空间中的所有像素点^[5],用高维(128 维)向量进行描述,并将生成的每个匹配基元与已知向量集中的每个基元做交叉比较,因此,具有较高的计算复杂度。而这些高昂计算的代价却并不总是有价值的。一方面,人脸的独特性主要体现在五官上,而额头、脸颊等信息并不具有很好的区分度;另一方面,全脸特征向量间的交叉比较是没有意义的,例如衡量一个人眼睛和另一个人鼻子的相似度。

为了降低计算复杂性,本文选择了人脸上的 6 个感兴趣的子区域,进行特征提取和向量的描述,这些区域分别是左眼、右眼、鼻子、嘴巴、左耳和右耳,分别定义为 1-6 号区域,如图 4 所示。然后,采用基于图的混合高斯树方法来完成对于人脸器官的定位任务^[6]。

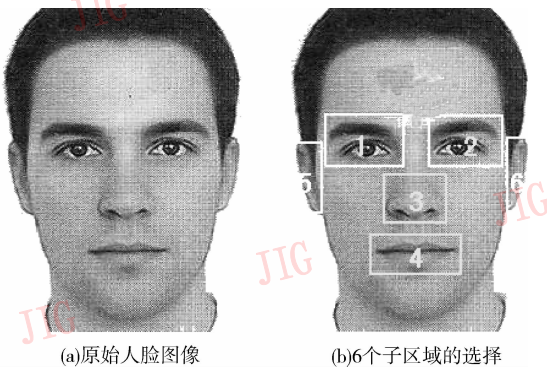


图 4 人脸感兴趣子区域

Fig. 4 Face interest sub-regions

因此,在特征描述后,用于比较的两张人脸图将各自得到 6 个子矩阵 $A_{i(m_i \times 128)}$ 和 $B_{i(n_i \times 128)}$,其中 $i = 1, 2, \dots, 6$, m_i 和 n_i 分别为两幅人脸图的第 i 个子区域中提取描述的特征点数目。从图 2 中也可以看出,进行子区域提取后,所需处理图像的“面积”显著变小了。

2.3 匹配算法

本文选用两向量间的内积来衡量其相似度。即,两向量 f_1 与 f_2 之间的相似度 d 可由式(1)得到

$$d = \frac{\langle f_1, f_2 \rangle}{\|f_1\| \cdot \|f_2\|} \quad (1)$$

通过先将 128 维的向量进行归一化,再用表示向量集的矩阵 $A_{i(m_i \times 128)}$ 和 $B_{i(n_i \times 128)}$ 进行转置相乘,记作

$$C_{i(m_i \times n_i)} = A_i B_i^T \quad (2)$$

即可得到向量集之间交叉比较的相似度。用 s_i 表示两人脸关于第 i 个子区域的相似度。定义 s_i 为

$$s_i = \max(c_{ij}) \quad (1 \leq i \leq m_i; 1 \leq j \leq n_i) \quad (3)$$

根据定义, s_i 的取值范围在 0 到 1 之间,越接近 1,表明相似度越大,则两人脸为相同身份的概率越大。

很显然,6 个子区域的独特性应该是不一样的,眼睛信息可能比嘴巴信息更容易进行身份的识别,因此,对于每个区域的相似度,应赋予不同的权值 w_i ,其中 $i = 1, 2, \dots, 6$ 。将人脸 6 个子区域的相似度单独拿来进行人脸识别实验,得到区分的准确率作为置信度 P_i 。则权值 w_i 可通过式(4)计算而得

$$w_i = \frac{1}{1 - P_i} \quad (4)$$

在先验等概,且各区域统计独立的前提假设下,式(4)所得到的权值 w_i 能很好地反映出各子区域的独特性。置信度越接近 1,所得到的权值就越大。

如果人脸的一部分存在着遮挡,或人脸图像的部分区域存在模糊,则将导致受影响的子区域出现相似度数值偏低的情况。很显然,这样的非期望数据值是不应该被引入的。因此,本文对于 6 个区域的相似度将设置一个阈值,相似度小于该阈值的区域,将不会在最后的匹配中计入。实验结果表明,即使是不同人物的人脸图像,其相对应区域的相似度一般也在 0.7 以上,故本文中设定的阈值为 0.7。则修正后的权值 w'_i 可由式(5)得到

$$w'_i = \begin{cases} w_i & s_i \geq 0.7 \\ 0 & s_i < 0.7 \end{cases} \quad (5)$$

从图 2 的算法步骤示意图中,也能看到,门限为 0.7 的比较器正好起到了式(5)的效果。

人脸识别的另一个困难在于丰富的人脸表情。人的喜怒哀乐将使得面部结构发生变化,进而影响匹配的结果。本文假设:对于人脸的 6 个子区域,当丰富表情出现的时候,一部分子区域会受到影响,而一部分子区域仍然相对保持稳定。例如:人笑的时候,嘴巴可能会张开,但此时耳朵和鼻子的图像信息应该还是保持稳定。因此,降低这种影响的办法是,让相对稳定进而得到较高相似度的区域所起到的作用更大;反之,令受影响区域进而得到较低相似度的区域所起到的作用小一些。本文选取了将相似度进行平方的策略来达到这一目的。从图 2 的算法步骤图中可以看到,相似度平方的运算将在通过比较器后进行。

最后,两人脸间的相似度 s 将通过式(6)计算得到

$$s = \frac{\sum_i^6 s_i^2 w_i'}{\sum_i^6 w_i'} \quad (6)$$

这个相似度 s 恰恰就反映了两人脸表征同一身份的可能性大小。

3 实验结果

本文选择了 400 多张数码相机所拍摄的生活照作为实验数据,每张照片大小为 $1\ 728 \times 2\ 304$ (该分辨率为当下数码照最常规的尺寸大小),人脸所占区域大小一般都在 200×200 以上,且大多为正面、倾角不超过 $\pm 15^\circ$ 。实验数据中共包含有 5 个不同身份的人。

实验分别统计了提取整幅人脸图像进行特征描述以及匹配的计算量与提取 6 个人脸感兴趣子区域的特征并进行描述以及对应区域匹配的计算量。表 1 给出了后者计算量与前者的比值。本文选取 50 对人脸图像的比较匹配过程进行计算时间的统计,结果改进方法后,平均时间可以降低 68.9%。

表 1 改进方法前后计算量的比值

Tab. 1 The improvement of the proposed method

单位: %

	遍历尺度空间	描述向量数	向量匹配对
改进后计算量较原先方法计算量之比	25.72	27.30	1.62

文中取人脸 6 个子区域的相似度单独进行人脸识别实验,得到区分的准确率作为置信度 P_i ,进而计算得到各子区域相似度的权重 w_i ,结果如表 2 所示。

表 2 人脸 6 个子区域的置信度测定与权重计算

Tab. 2 Confidence factor and weight of sub-regions

区域编号	区域位置	置信度 P_i	权重 w_i
1	左眼	0.91	11.11
2	右眼	0.88	8.33
3	鼻子	0.76	4.17
4	嘴巴	0.81	5.26
5	左耳	0.65	2.86
6	右耳	0.69	3.23

将改进的匹配算法与在使用 SIFT 算子后、各区

域相似度做简单平均的算法识别率进行了比较,测试数据中包括了通过人脸进行身份识别过程中出现的一些非期望情况,如人脸表情夸张、图像存在模糊、人脸部分存在形变等等。虽然都会导致识别率的下降,但改进后的算法识别率下降的幅度仅为未改进方法的约三分之一,明显具有更好的鲁棒性。表 3 的数据表明,该改进方法是行之有效的。

表 3 改进方法提高的识别百分比

Tab. 3 The improvement of the proposed method

匹配策略	识别率提高百分比
按独特性赋予权重	25.72
设置相似度最低域值	8.23
采用相似度平方策略	13.11
上述 3 种策略一并使用	31.02

4 结论

本文以 SIFT 算子为基础,根据其特点,提出了一种改进匹配策略的人脸识别方法。理论分析与实验结果均表明,该方法既可以显著降低运算的复杂度,又可以使人脸识别的准确率有较大的提高,并对各种非期望情况有着较好的鲁棒性。如何有效地把几何特征不变量运用到有一定倾角的侧面人脸识别中,将是下一步研究的重点。

参考文献 (References)

- 1 Harmon L D, Kuo S C, Raming P F, et al. Identification of human face profiles by computer[J]. Pattern Recognition, 1987, 10(5~6): 301~312.
- 2 Liu Bo-wen, Yu Song-yu, Xu Yi, et al. Correspondence of Object-of-Interest in Wide Baseline Active Vision [J]. Journal of Image and Graphics, 2007, 10(2): 1917~1921. [刘博文,余松煜,徐奕等.宽基线主动视觉中感兴趣目标的对应技术[J].中国图象图形学报, 2007, 10(2): 1917~1921.]
- 3 Viola P, Jones M. Robust real-time face detection[J]. International Journal of Computer Vision, 2004, 57(5):137~154.
- 4 Lowe D. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision, 2004, 60(2):91~110.
- 5 Lowe, David G. Object recognition from local scale invariant features [A]. In: Proceedings of International Conference on Computer Vision [C], Corfu, Greece, 1999:1150~1157.
- 6 Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors[A]. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition [C], Washington DC, USA, 2004:506~513.