

一种定位人脸部特征的光学聚类方法

周 冕^{1,2,3}, 王向军¹

- (1. 天津大学 精密测试技术及仪器国家重点实验室, 天津 300072;
2. 英国雷丁大学 系统工程学院, 雷丁 英国 RG6 6AY;
3. 天津外国语大学 教育技术与信息学院, 天津 300204)

摘 要: 人脸识别技术需要对脸部特征进行定位,从而有助于确保图像一致和建立人脸模型。提出了一种新的脸部特征定位方法,通过 Gabor 滤波器处理得到人脸图像的强度响应,其中,脸部特征表现为强响应,而其他部分表现为弱响应,如面颊和额头。通过保留强响应以及过滤弱响应,可以获得属于脸部特征的所有像素点。采用了聚类算法—— k 均值算法将不同的像素点分配到不同的簇里面,每一个簇都代表一个脸部特征。通过在 ORL 人脸数据库上的测试表明:此方法能精确、快速地定位诸如眼睛、鼻子、嘴等脸部特征。此外,此方法能够在有浓密胡须的对象上成功定位脸部特征,表现出较高的鲁棒性。

关键词: Gabor 滤波器; 聚类; 人脸识别

中图分类号: TN29 **文献标志码:** A **文章编号:** 1007-2276(2011)03-0576-05

Optical clustering method for locating facial features

Zhou Mian^{1,2,3}, Wang Xiangjun¹

- (1. State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China;
2. School of Systems Engineering, University of Reading, Reading RG6 6AY, United Kingdom;
3. School of Educational Technology and Information, Tianjin Foreign Studies University, Tianjin 300204, China)

Abstract: The localization on facial features is needed for face recognition since it helps keeping accordance between face images and building face model. In this paper, a novel method for locating facial features was presented which included two steps: filtering and clustering. Face images were firstly processed by Gabor filter into magnitude responses. In the responses, facial features demonstrated relatively high magnitude responses than other facial parts, such as cheek and forehead. By reserving high magnitude responses and removing low magnitude responses, the pixel points belonging to facial features were collected. The method adopted a clustering approach— k -means for separating pixel points into different clusters. Each cluster represented a facial feature. By testing on the ORL face database, the method shows its accuracy and speed on locating facial features, such as eyes, nose and mouth. It also exhibits high robustness in locating features on faces which have thick beard or mustache.

Key words: Gabor filter; clustering; face recognition

收稿日期:2010-06-15; 修订日期:2010-07-20

基金项目:精密测试及仪器国家重点实验室开放基金资助项目

作者简介:周冕(1979-),男,副教授,博士,主要从事计算机视觉方面的研究。Email: zhoumian@hotmail.com

0 引言

近10年来,人脸识别技术得到了飞速地发展,越来越多被应用于通关、企业安全和管理、刑侦、人机交互等众多领域。人脸识别技术按处理方式可以分为两种:基于外表的人脸识别和基于模型的人脸识别。基于外表的人脸识别是利用人脸图像中像素值的统计分布规律来实现的,主要代表为基于主成分分析的人脸识别方法^[1]。而基于模型的人脸识别需要先建立人脸模型,然后根据模型对人脸图像进行匹配,得出模型的参数,从而实现最终识别。其典型代表为弹性图匹配^[2]和主动外观模型^[3]。

虽然这两类方法有着不同的处理模式,但是他们几乎都需要先定位脸部特征。例如,在基于外表的人脸识别中,需要找到眼睛、鼻子、嘴巴等特征的位置,然后根据所得到的位置来预处理所有的人脸图像,如所有的眼睛必须在同一高度。而在基于模型的人脸识别中,模型是由一系列的关键点组成,而这些关键点就是脸部特征,通过自动搜寻脸部特征,即可建立起模型。由此可见,对脸部特征的定位是人脸识别技术的关键。该过程处理的好坏将决定最终识别的成功与否。

文中提出的是一种基于聚类脸部特征定位方法。首先通过 Gabor 滤波器对人脸图像进行转换,将面部的信息以强度响应的形式表现出来。通过保留强响应,过滤弱响应的方法,使得脸部特征得以全部保留下来。最后根据 k 均值聚类的方法,将所保留的强响应分为多个不同的簇,每一个簇代表一个脸部特征。这种方法具有快速、准确、强壮等优点,尤其适合于定位眼睛、鼻子、嘴等常见的脸部特征。

1 Gabor 滤波器

二维的 Gabor 滤波器最早是由 J. Daugman^[4]在1985年提出来的。现已广泛地应用于纹理处理^[5]和目标检测^[6]。

Gabor 滤波器由傅里叶级数方程与高斯方程相乘得到的,为受三角函数调制的高斯函数,它由两部分组成:实数部分和虚数部分。实数部分是高斯函数和余弦函数的乘积,具有对称结构;虚数部分是高斯函数与正弦函数的乘积,为非对称结构。通常 Gabor 滤波器表示为指数型式:

$$H(x,y,\sigma,U,V)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}e^{j2\pi(Ux+Vy)} \quad (1)$$

式中:第一个指数代表高斯函数;第二个指数表示三角函数的调制。Gabor 滤波器为具有 5 个变量的方程: x 和 y 分别表示图像上的横坐标和纵坐标; σ 是二维高斯函数的标准偏差(x 坐标上和 y 坐标上的标准偏差相同); U 和 V 分别决定滤波器在 x 和 y 坐标上的空间频率。

实际处理中,Gabor 滤波器表示为 $N \times N$ 的方阵模板, N 一般为奇数,如: 7×7 、 13×13 、 19×19 等。 x 和 y 变量的取值范围为 $[-(N-1)/2, (N-1)/2]$ 。模板中心的坐标为 $(0,0)$ 。

Gabor 滤波器的射频 F 由 U 和 V 两个空间频率决定:

$$F=\sqrt{U^2+V^2} \quad (2)$$

Gabor 滤波器的方向也是由 U 和 V 来决定。方向的角度表示为:

$$\theta=\arctan \frac{U}{V} \quad (3)$$

因此,两个空间频率又可以表示为 $U=F \times \cos \theta$ 和 $V=F \times \sin \theta$ 。变量 σ 和 F 之间的关系源于: $\sigma=\frac{\kappa}{F}$, 其中

$$\kappa=\sqrt{2 \ln 2}\left(\frac{2^{\varphi}+1}{2^{\varphi}-1}\right) \quad (4)$$

式中: φ 是以倍频程来衡量的带宽。根据参考文献[7], Gabor 滤波器最好选择 1~1.5 的倍频程。这里,带宽采用的倍频程为 1.0。因此,标准偏差 σ 为 π/F 。空间频率将分别为:

$$U=\frac{\pi}{\sigma} \cos \theta, \quad V=\frac{\pi}{\sigma} \sin \theta \quad (5)$$

因此,可以将 Gabor 滤波器转换为以 4 个变量表示的形式:

$$H(x,y,\sigma,\theta)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}e^{j2\pi^2\left(\frac{x\cos\theta+y\sin\theta}{\sigma}\right)} \quad (6)$$

式中: σ 不仅表示标准偏差,而且还决定 Gabor 滤波器的大小; θ 为空间方向。

滤波的结果通过 Gabor 滤波器与数字图像进行离散卷积得到。如上所述,Gabor 滤波器可以离散化为 $N \times N$ 的模板形式。然而,对于不同的标准偏差 σ , N 不是唯一值。如果 σ 的取值很小而 N 的取值很大,那

么 Gabor 滤波器只在模板中占有很小的部分, 其他部分都为空白。这样造成在卷积中, 大部分的计算都被浪费掉了。因此, 需要根据标准偏差 σ 来动态地决定 N 的取值。Dunn 和 Higgins^[6]提出了 Gabor 滤波器的宽度应为 $6\sigma+1$, 如标准偏差 σ 为 24 pixel, 那么模板的大小将为 145×145 。图 1 所示为一个 Gabor 滤波器的模板, σ 为 8, 方向 θ 为 $\pi/4$ 。其中, 图(a)为其实部, 图(b)为其虚部。

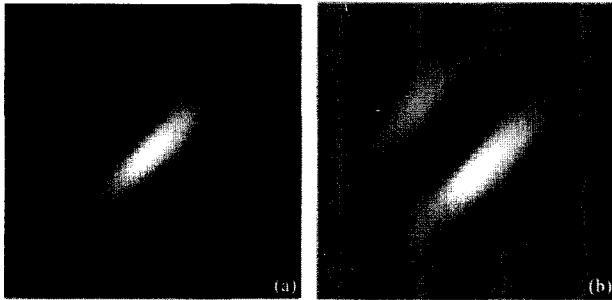


图 1 一个方向为 $\pi/4$, 标准偏差为 8 的 Gabor 滤波器
Fig.1 An example of Gabor filter, whose orientation is $\pi/4$ and σ is 8

如上所述, Gabor 滤波器由实部和虚部两部分组成, 因此, 需将实部和虚部分别与图像进行卷积。卷积的结果包括两种响应, 即实部响应和虚部响应。而强度响应是实部响应和虚部响应平方和的开方, 强度响应广泛应用于图像分割和边缘检测中, 因此, 文中的人脸识别采用的是强度响应。图 2 为人脸图像与 Gabor 滤波器($\sigma=1, \theta=\pi/2$)卷积之后强度响应的结果。



图 2 人脸图像(a)和其强度响应的结果(b)
Fig.2 Human face image (a) and its magnitude response (b) by a Gabor filter

2 k 均值算法

k 均值算法是一种基于样本间相似度的间接聚

类方法。聚类是将一系列样本重组或分割为不相连的簇。由聚类所生产的簇是一组样本的集合, 这些样本与同一个簇中的样本彼此相似, 而与其他簇中的样本相异。聚类方法已经广泛应用于数据挖掘、人工智能和图像处理等领域。

k 均值算法属于传统聚类分析中划分方法的一种。在此算法中, 需要预先假设存在 k 个簇, 然后将 n 个样本分别划分到 k 个簇中以便使得聚类满足: 同一簇中的样本相似度较高; 而不同簇中的样本相似度较低。样本之间相似度是利用簇中样本的均值——“中心对象”(引力中心)来进行计算得到的。

首先, 从 n 个样本($i_1, i_2, \dots, i_n$)选择任意 k 个样本($\omega_1, \omega_2, \dots, \omega_k$)作为初始簇($C_1, C_2, \dots, C_k$)的中心(原型); 而将剩下的样本根据它们与这些簇中心的相似度(距离), 分别分配给与其最相似的(簇中心所代表的)簇; 然后再计算每个所获新簇的中心(该簇中所有样本的均值); 不断重复这一过程直到误差值开始收敛为止。误差值 E 的计算方式为:

$$E = \sum_{j=1}^k \sum_{i_i \in C_j} (i_i - \omega_j)^2 \quad (7)$$

由公式(7)可以看出, 误差值 E 为所有样本与其簇中心的均方差之和。 k 个簇具有以下特点: 各簇本身尽可能地紧凑, 而各簇之间尽可能地分开。

每一次递归结束之后, 都需计算误差值 E , 如果误差值不收敛, 那么 k 均值算法仍将继续。实验中误差随着每一次递归的变化如图 3 所示, 其中有 398 个样本, 每个样本包括两个值。从图中可以观察到, 当运行到第 8 次递归的时候, 误差开始收敛, 之后误差基本不变。因此, 可在第 9 次递归时结束 k 均值算法的计算。

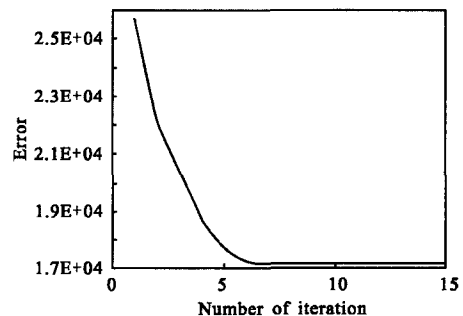


图 3 k 均值算法的误差变化

Fig.3 Error's change with the iterations in k -means algorithm

由于 k 均值算法需要的计算时间与样本数和簇数的乘积成正比,可见,在数据较为庞大时,此算法将非常耗时。但是,由于文中的样本数限制为500~1000,簇数为4~5,因此, k 均值算法仍然适用。

3 实验与分析

3.1 实验数据

文中采用的是 ORL 人脸数据库^[9],此数据库由剑桥 AT&T 实验室开发,库中包含 40 个人,每个人有 10 副人脸图像,总计 400 副人脸图像。10 副图像分别摄于不同的时间、具有不同的光照条件、不同的面部表情等。所有图像都是 92×112 像素大小,256 级的脸部正面灰度图像。

不同于实部响应和虚部响应,强度响应输出仅为正数。实验表明,最强响应为 40.2398,最弱响应为 5.4026×10^{-4} 。在人脸分析中,强响应往往代表诸如眼睛、鼻孔、嘴唇等特征区域;而弱响应往往是指那些平滑变化的区域,如脸颊。因此,为了定位脸部特征,需要采用强响应,而忽略弱响应。

3.2 Gabor 滤波器的选取

经过实验发现,模板越小(标准偏差 σ 越小),强度响应表现出的人脸局部细节的变化则越多,人脸的特征也就越容易被找到。在 Gabor 滤波器的众多方向中,当 $\theta = \pi/2$ 时,强度响应比其他方向的滤波器含有更多的强响应。最后,经过反复的实验求证,发现当滤波器取 $\sigma=1, \theta=\pi/2$ 时,能够检测出脸部最精细的特征变化。如图 2 所示,在眼部(两个)、鼻子、和嘴部都展现出了几个高亮度的簇。在响应图像中,这些高亮度的点即为聚类分析中的样本,每个样本含有两个变量 x 和 y ,即横坐标值和纵坐标值。

进行聚类分析时,不能将响应图像中所有的点都视为样本。从图 2 中可以发现,大部分的区域都是很弱的响应(几乎为 0),只有脸部特征区域才有很强的响应。基于这一点,可以通过保留强响应,忽略弱响应来降低 k 均值算法中的样本数。经过实验发现,通过滤去响应小于 3.5 的样本,样本数能够从原来的 $92 \times 112 = 10304$ 下降到 500~1000 的范围内。

3.3 聚类结果

针对每一幅图像进行如上处理后, k 均值算法即应用在响应图像的样本上。这里, k 被设定为 5,即假

设含有 5 个簇。一般 k 均值算法最多经过 12 次递归计算误差收敛。如图 4 所示,5 个簇分别为左眼、右眼、鼻尖、嘴(左)和嘴(右)。每个簇的中心点用来表示脸部特征的位置。通过此聚类算法,可以找出眼睛、鼻子和嘴的位置。尽管图中一些人脸具有很浓密的胡须,此算法仍然能够准确地定位到相应的脸部特征。



图 4 脸部特征的定位结果

Fig.4 Results of locating facial features

4 结 论

文中提出了一种定位脸部特征的方法。此方法包含两个部分:Gabor 滤波和 k 均值聚类算法。Gabor 滤波将人脸图像转换为强度响应信息。脸部特征越鲜明的部分,响应就越强,反之响应就越弱,如脸颊部分和额头部分的响应几乎为零。通过收集强响应的像素点,同时忽略弱响应的像素点,可以发现脸部特征是以簇的形式来显现的。经过 k 均值聚类算法,不同的像素点被划归于不同的簇。不同的簇就表现为不同的脸部特征。

通过 ORL 数据库的测试表明,此算法能够成功地定位眼、鼻子和嘴等特征。同时,具有运算速度快的优点,因为一般只需要 8~12 次递归运算就能完成聚类。另外,此方法还有较高的鲁棒性,对于含有浓密胡须的人脸,仍然能够成功地找出脸部特征。

虽然用聚类的方法只能找出人脸特征的大致位置,然而在人脸识别中,通过此方法可以将特征搜索和匹配的区域限制在一个足够小的范围,避免了全局的搜索和匹配,从而大大降低了计算时间。

综上所述,这种 Gabor 滤波加聚类的方法用于定位脸部特征是可行的。

参考文献:

- [1] Turk Matthew, Pentland Alex. Eigenfaces for recognition[J]. *Journey of Cognitive Neuroscience*, 1991, 3(1): 71-86.
- [2] Wiskott Laurenz, Fellous Jean Marc, Kruger Norbert, et al. Face recognition by elastic bunch graph matching [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 775-779.
- [3] Cootes Timothy, Edwards Gareth, Taylor Chris. Active appearance models[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(6): 681-685.
- [4] Daugman John. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters [J]. *Journal of the Optical Society of America A*, 1985, 2(7): 1160-1169.
- [5] Shen Wen, Xia Bin. Texture segmentation method based on Gabor ring filtering [J]. *Infrared and Laser Engineering*, 2003, 32(5): 484-488. (in Chinese)
- [6] Qin Hanlin, Liu Shangqian, Zhou Huixin, et al. Background suppression for dim small target with Gabor kernel non-local means [J]. *Infrared and Laser Engineering*, 2009, 38(4): 737-741. (in Chinese)
- 秦翰林, 刘上乾, 周慧鑫, 等. 采用 Gabor 核非局部均值的弱小目标背景抑制[J]. *红外与激光工程*, 2009, 38(4): 737-741.
- [7] Lee Tai-Sing. Image representation using 2d Gabor wavelets [J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1996, 18(10): 959-971.
- [8] Dunn Dennis, Higgins William. Optimal Gabor filters for texture segmentation [J]. *IEEE Transactions on Image Processing*, 1995, 4(7): 947-964.
- [9] Samaria Ferdinando, Harter Andy. Parameterization of a stochastic model for human face identification[C]//Proceedings of IEEE 2nd Workshop on Applications of Computer Vision. Florida USA: IEEE Computer Society Press, 1994: 138-142.

书 讯

《光学系统设计》(内部资料)

主要包括:基础光学与光学系统技术要求;光阑、光瞳和基本原理;衍射、像差和像质;光程差的概念;几何像差概述及其消除方法;玻璃的选择;球面和非球面;光学设计过程;计算机性能评价;高斯光束成像;3~5 μm 和 8~12 μm 红外成像基础和紫外光学系统;衍射光学;照明系统的设计;性能评价与光学测试;公差与生产工艺性;光学加工;光学设计中的偏振问题;光学薄膜;硬件设计问题;镜头设计优化实例;光学系统设计常见的错误和失误;经验法则和提示等。

《衍射光学元件的计算机设计方法》

本书由俄罗斯科学院萨马拉图像处理系统研究所的研究人员撰写,北京清华大学金国藩院士主持翻译,《红外与激光工程》编辑部编辑、天津科学技术出版社出版。涵盖了衍射光学元件的基本特点、衍射光学元件常用迭代设计方法、基于电磁场理论的亚波长元件的设计方法、衍射光学元件的加工方法等,并介绍了几大类衍射光学元件的实例。例如聚焦元件、激光模式转换器、产生周期性光束的元件、波面校正元件、用于照明系统的元件等,并举出一些衍射光学元件在光学信息处理系统中的应用。本书可作为各类大学物理系、光学工程系、信息科学系本科生和研究生的教学参考书,也可供有关专业从事研究、开发的科技人员参考。

有需求者请与《红外与激光工程》发行部联系。电话:(022)58625210。