

基于小波变换域的人脸弹性识别方法

Facial Recognition Algorithm Based on Wavelet Transform

马 燕

(上海师范大学计算机系 上海200234)

Abstract This paper aims at the application of wavelet theory in human face recognition, focuses on elastic matching which is in common use. To counter the deficiency of the methods now available, such as the complication of calculation and the redundancy of information, while reserving the stability and high efficiency of elastic matching, a human face model is proposed based on the high frequency band in the light of wavelet theory. The method of elastic matching is improved and developed through the low dimensionality approximation of images, the use of the symmetry of face, the selection of weight in faces, the calculation of similarity and cost function and the combination with high and low frequency band. And the result shows it is effective.

Keywords Face recognition, Elastic matching, Wavelet transform

1 引言

利用计算机进行自动、快速的人脸识别,一直是模式识别领域的前沿课题。目前的研究已不仅仅局限于人脸的几何特征,还包括统计特征和频谱特征等。许多方法考虑将多种特征结合以形成人脸的多元特征表达^[1]。Lades 提出的弹性图匹配方法^[2]是一种基于整体的研究方法,它通过能量函数来实现匹配,其整体识别性能要优于特征脸方法^[3]。当向人脸库加入新的人脸时,特征脸法需要重新计算特征库中的特征脸,而弹性匹配法不需要改变已有的数据,直接加入新的模板数据即可。但其匹配过程是一个反复比较过程,计算量巨大,因此,该方法的缺点就是识别速度慢。本文利用二维小波变换^[4]的时频特性和多分辨率特性,对现有的弹性匹配法进行改进,在小波变换域上利用脸部对称性准确抽取人脸最具代表性的特征,且赋以不同权重,在充分利用小波高低频子图特点的基础上,根据本文提出的相似度和代价函数的计算方法进行人脸识别实验,在保证总识别率的前提下,使得识别时间同弹性图匹配法相比大大减少。

2 基于小波高频子图的人脸模型抽取方法

1993年, Martain Lades 提出了一种基于 DLA 模型的弹性图匹配方法,它以可变形的二维网格作为模板来表示人脸:在脸部区域抽取一个二维网格,由 7×10 个点组成(图像大小为 128×128 像素)。这些网格可被视为一个二维拓扑图,其中网格的格点对应二维拓扑图的结点,将拓扑图中每个结点通过一个 Gabor 小波^[5]滤波器组进行滤波,利用70个结点的特征矢量构成的特征矩阵来表达人脸的特征信息并进行匹配比

较。这里,二维拓扑图构成了一个人脸模型,这种人脸模型存在无法区分脸部各区域重要性,不能适应较大姿势改变等缺点。Lades 将两幅图像的匹配过程分成两步:粗略匹配和精确匹配。这两步都是利用最小化能量函数的方法来实现的。在粗略匹配过程中,将待识人脸的二维网格 S_1 定位在已知人脸网格 S 的大致位置上,所以不允许网格的形变,而只进行整个网格的匹配。因此,这一步所得到的最小能量函数还不是一个极小值。在精确匹配过程中,允许网格 S_1 中各个结点的细微移动,使 S 中的各个结点较为精确地同 S_1 中各个结点进行配准,目的是为了使能量函数最终收敛到一个极小值来实现精确匹配。这种结点的移动使人脸模板能产生形变,从而容忍人脸姿势、表情、动作等的细微变化。

在大多数人脸图像中可看到,能体现人脸细节特征的区域主要集中在眉毛、眼睛、嘴巴、鼻子等处,因此,应从这些区域中抽取结点构成人脸模型。根据脸部各个特征区域在识别中所占比重的实验统计比较,本文选取如图1(a)所示的由11个结点构成的人脸模型。



图1 人脸模型

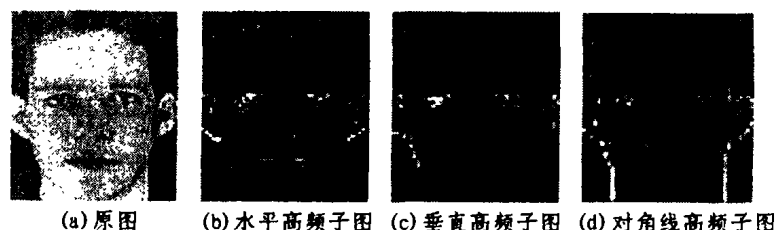


图2 人脸图象及其三个方向的高频子图

这里使用的脸部模型除眉毛和嘴巴中心外,还需抽取其它一些区域:左右眼角、左右眼帘、鼻梁、鼻尖、左右嘴角等。由

于存在对称性,人脸的左右部分在识别中所起的作用是相当的,因此,可利用这种对称性来简化人脸模型,如图1(b)所示的左半边人脸模型中去掉右边眉心、右眼角、右眼帘、右嘴角等4个点,仅剩7个结点。这些结点的位置可通过人脸图像小波变换后的高频子图去抽取。

从图2可看出,小波变换后的高频子图尤其是水平高频子图(图2(b))中,人脸模型中各结点对应区域的亮度大且较其它区域集中,因此,可以在去掉孤立亮点的二值化小波水平高频子图中选取面积最大的若干区域,根据人脸五官的自然分布来查找各结点的位置。考虑到具体查找时的技术因素,本文对于识别的人脸图像作以下几个假定:①图像中只有一个人脸,且占40%以上的幅面;②双眼和两个嘴角可见;③头部在图像中的倾斜度不超过45度。由于本文利用小波高频信息抽取结点以构成人脸模型,故称其为小波高频脸部模型。利用该模型能基本定位脸部特征,并取代了DLA方法中的粗略匹配过程,减少了计算量。实验证明,这种方法对光照不敏感,具有较强的稳定性和抗干扰能力,因此不影响粗略匹配的准确率。实验中还发现,利用小波变换后各层高频分量间存在的相关性,可将不同层小波分量查找的结果作平均,所得的定位准确率比单独在一层中查找要高。

3 基于小波变换域的人脸弹性匹配方法

3.1 相似度与代价函数的计算

在训练和识别过程中,需要对人脸图像的特征模板进行

表1 左半边脸部模型中各结点对应的权重

脸部各结点	左眉心	左眼角	左眼帘	左嘴角	嘴巴中心	鼻尖	鼻梁
权重取值	1.8	1.8	1.6	1.6	1.4	0.9	0.9

在进行精确匹配时,本文选择以脸部模型的各个结点为中心的、一个小矩形为弹性匹配区域,计算区域内每个点对应的特征矢量同另一图像模板相应结点特征矢量之间的相似度。为表述方便,引入代价函数的概念。由于训练和识别过程的要求不同,在这两个过程中采取的弹性匹配法也不尽相同,因此在训练和识别过程中采用不同形式的代价函数。

在训练过程中,相应结点之间的代价函数为

$$C_i = \frac{1}{\max_k (S(\vec{J}_i, \vec{J}_k^t))} \quad (3)$$

这里, $\vec{J}_i$ 表示图像的某一 Gabor 特征矢量, $\vec{J}_k^t$ 表示图像库中已有图像的相应特征矢量。 $\max_k (S(\vec{J}_i, \vec{J}_k^t))$ 表示 $\vec{J}_i$ 同库中所有已有图像相应特征矢量进行比较得到的最大相似度。

在识别过程中,用相似度 S 的倒数作为弹性匹配中相应结点之间的代价函数,即

$$C_i = \frac{1}{S(\vec{J}_i, \vec{J}_i^t)} \quad (4)$$

进行弹性匹配就是最小化代价函数的过程。通过迭代运算,在各结点的匹配区域内选择代价函数最小点的位置作为各结点的最佳位置,这样,就在通过小波高频子图脸部模型确定各结点大致位置的基础上,实现了精确匹配。

3.2 特征模板库的形成

本算法的主要思路是利用小波变换的高频子图对人脸图像进行大致模型定位,然后利用多通道的 Gabor 小波滤波器组从图像的低频子图中抽取特征信息,对脸部模型进行弹性匹配,得到一个代表脸部的精确特征模板,进而将其用于识别。基本步骤为:

匹配,计算它们的相似度。基于上述对弹性匹配法的改进,本文提出以下形式的相似度计算方法。

首先抽取图像中各个结点的特征矢量,然后计算对应特征矢量之间的相似度,再把各个区域的相似度按照一定的权重相加构成总的相似度。

设 $\vec{J}_i^t$ 和 $\vec{J}_i^{MOD}$ 分别为一幅人脸图像的某一结点的特征矢量和对应模板相应区域的特征矢量,二者的相似度为:

$$S(\vec{J}_i^t, \vec{J}_i^{MOD}) = \frac{\vec{J}_i^t \cdot \vec{J}_i^{MOD}}{\|\vec{J}_i^t\| \cdot \|\vec{J}_i^{MOD}\|} \quad (1)$$

实验证明,(1)式所示的相似度计算简单,便于处理,且对图像的光照等全局变化有较强的稳定性。根据脸部模板各个结点代表的区域在识别中所起的作用,将各个结点之间的相似度按照一定的权重相加得到总相似度,即为这个匹配过程的能量函数。设 $S = \{S_1, S_2, \dots, S_N\}$ 为各对应结点之间的相似度, $W = \{W_1, W_2, \dots, W_N\}$ 为各结点的权重,则整个脸部匹配过程的总相似度,即能量函数为

$$E = \sum_i S_i W_i \quad (2)$$

(2)式所示的能量函数反应了人脸图像和人脸特征模板之间的相似程度,可用它对人脸图像进行分类和识别。

本文通过实验,将左半边脸部模型中的7个结点单独进行识别,根据各自得到的识别率换算出如表1所示的权重值,总权重等于10。从表1可见,左眉心和左眼角所占的比重最大,其次是左眼帘和左嘴角,再次是嘴巴中心,最后是鼻尖和鼻梁。

1)对图像库中的样本图像进行二层 db1 小波变换,利用小波高频脸部模型对各结点定位。

2)利用1)中小波变换的第二层低频子图代替原图,根据1)所得的结点位置,用一个 M 维的 Gabor 小波滤波器组对该低频子图中的各结点进行滤波,得到各结点的特征矢量,从而形成各样本图像的脸部特征模板。

3)对新样本图像的脸部模型和库中已有图像的模型计算相似度和代价函数,进行精确匹配,得到图像库中每个人的精确特征模板,保存起来以便于比较和识别。

4)根据需要对本样本图像库中同一个人的不同照片的特征模板进行平均,得到代表这个人的平均特征模板。

实验中,对图像库中的几幅正面照片进行手工定位,其余照片则由机器自动训练完成。最初一批自动训练图像经过上述方法后仍需手工干预,进行修正,但当训练图像达到一定数目后,机器自动训练得到的特征已经较为可靠,基本无需人工干预。在利用上述方法对图像库进行训练时,可以随时向图像库中加入新图像,而不用重新计算已得到的特征量。图3所示的是自动训练图像数目同匹配准确率之间的关系。从图3可看出,图像库中新添加的图像数目超过60幅之后,自动匹配的效果就比较可靠了。

3.3 识别过程及小波高频分量的利用

识别时,对待识图像进行类似样本图像的人脸模型和人脸模板抽取,将得到的特征模板同训练模板库中所有模板进行弹性匹配,从而得到待识图像同所有样本图像的相似度,以及该待识图像同每个人的平均特征模板之间的相似度,本文选取匹配结果中同待识图像相似度最大的图像或样本作为识

别结果。本文规定：

系统识别率=判断正确的图像数/待识图像的总数；

平均识别时间=所有待识图像的识别时间之和/待识图像的总数。

根据 Campbell 和 Robson^[6] 的心理学实验证实,人类的识别过程主要依靠低频信息,在低频信息的基础之上再来作高频判断。当低频信息比较接近时,为使识别更精确,可抽取高频分量作进一步比较。本文在前面对人脸图像低频子图进行识别的结果中选取相似度最大的几幅图像或模板,然后利用小波高频子图(若参与比较的是一个人的几幅样本图像的平均模板则高频特征同样取平均)进行高频信号的匹配。具体算法为:

图像数	匹配较好图像数	准确率
15	9	60%
30	22	73%
45	36	80%
60	50	83%
75	64	85%
90	79	88%
105	94	90%

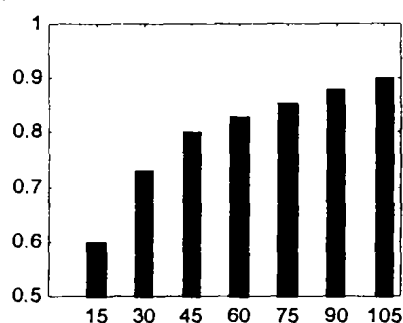
图3 自动训练图像数目同匹配准确率之间的关系

1)选取上面方法得到的与待识图像相似度最大的若干样本图像。

2)以上面方法得到的脸部模板各结点为中心,取图像第二层小波垂直子图和水平子图中的一个矩形区域,用它们的值构成一个高频特征矢量。

3)求待识图像和样本对应结点高频特征矢量之间的欧几里德距离,选取距离最小者作为最相似图像。

实验证明,本算法对前面讲述的针对低频信息的弹性匹配法起到了很好的补充作用,可以根据具体识别要求来判断是否需要进行高频分量的匹配比较。如果要求较高的识别率,则采用该方法作进一步判别;否则仅对低频信息作分析即可。



4 实验研究

4.1 实验条件

本文图像库包含40个人,每人10幅图像,大小为112×92的灰度图像,包括不同人的不同尺度、姿势、表情、光照等的照片,且所有这些图像均满足前面对用于识别的人脸图像所作的若干假定。其中大部分图像来自 ORL 标准图像库,下载自: www-rocq.inria.fr/cgi-bin/imedia/surimage.cgi,为使识别结果更具说服力,我们向图像库中添加了 FERET 图像库中一些复杂背景的人脸图像和用数码相机拍摄的若干幅人脸图像。我们随机选取每个人5幅图像,共200幅图像作为样本图像用于训练,而另外200幅图像就作为待识图像用于识别和收集实验数据。

4.2 实验结果

实验1 利用本文方法所得识别结果及部分示例

根据前面提供的实验条件,将所有200幅待识图像输入识别系统进行三次识别,并对识别结果进行统计,表2是统计结果。

表2 总的识别率和平均识别时间

	判断正确的图像数	识别率	平均识别时间 (s)
第一次	167	84%	9.1224
第二次	166	83%	9.3418
第三次	166	83%	9.4774
平均	166	83%	9.3138

在实际识别中,仅抽取同待识图像最相似的样本图像往往不能说明什么问题,进行识别的最终目的是为了知道待识图像属于图像库中哪个人的照片。为此取图像库中每个人的若干张图像的特征模板的平均值,作为图像库中每个人的平均特征模板,用于同待识图像进行比较,表3给出了采用同样的实验条件得到的识别结果:

表3 抽取平均特征方法所得识别结果

	判断正确的图像数	识别率	平均识别时间 (s)
第一次	184	92%	4.6241
第二次	183	92%	4.6021
第三次	183	92%	4.6123
平均	183	92%	4.6195

比较表2和表3,发现采用抽取平均特征的方法比仅比较图像的方法在识别率和平均识别时间上都有效得多,因此在下面的比较中,都采取这种方法进行识别。

实验2 同基于 DLA 模型的方法进行的比较

为便于比较,利用本文中的图像库,模拟 Lades 的基于 DLA 模型的方法进行了训练和识别。分别对每个待识图像及其第二层 db1 小波变换的低频分量分别进行处理,识别时同样先求平均再识别,表4给出了识别结果:

表4 基于 DLA 模型的方法所得识别结果

	判断正确的图像数	识别率	平均识别时间 (s)
原图像	190	95%	181.3622
小波低频子图	186	93%	53.6745

比较表3和表4的结果可发现,本文提出的基于小波变换域的人脸弹性识别方法在总识别率上略低于 Lades 的方法,但在识别速度上有很大提高,识别时间仅为 Lades 方法的1/40。

实验3 识别稳定性研究

为研究本文所述算法在各种类型照片下的识别效果,选取姿势、表情、光照、大小等变化较大的若干张图片,以及一些干扰较强的图像,如戴眼镜、闭眼、大胡须、复杂背景等,同图像库中训练图像进行识别比较。本文选取每种情况下各10幅图像分别用本文方法和 DLA 方法进行识别,图4给出了这两种方法识别结果的比较:

从以上统计结果看,本文方法同基于 DLA 模型的方法相比,对人脸姿势、表情变化有较强的稳定性,而对光照以及

(下转第95页)

检索主题的扩展查询和兴趣学习。

假设条件:a)用户要求查询关键字为W(可以是一个或多个词语,为便于描述算法,这儿假设就一个词语);b)经过4.1节聚类算法得到的满足条件的文献文档集合S,初始值为空;c)提供聚类中心词语的文档集合S',初始值为空;d)作为聚类中心的词语集合C',初始值为空。

算法步骤:

```

a)检索文档集合S中每一篇文档
If 文档标题中含有检索关键字W,把该文档加入到集合S'中
Then
{对该文档中标为关键字的每个词语(注意这儿要排除W)
进行以下处理:
If 该词语不在聚类中心集合C'中
Then 把它加入到集合C'中,并把C'中该词语的初始权值
记为0
Else 把C'中该词语的权值自动加1//即该词语已存在于聚
类中心集合C'中
}
Else //文档标题中不含有检索关键字W
{If 文档关键字词语中含有W,把该文档加入到集合S'中
Then
{对该文档中标为关键字的每个词语(注意这儿要排
除W)进行以下处理:
If 该词语不在聚类中心集合C'中
Then 把它加入到集合C'中,并把C'中该词语的初始
权值记为0
Else 把C'中该词语的权值自动加1//即该词语已存在
于聚类中心集合C'中
}
Else //该文档的标题和关键字词语中均不包含W
{对该文档中标为关键字的每个词语(注意这儿要排除
W)进行以下处理:
If 该词语在聚类中心集合C'中
Then 则把C'中该词语的权值自动加1
}
}
b)最后把聚类中心集合C'中词语的权值求平均作为阈
值,把C'中那些权值大于阈值的词语作为用户扩展的兴趣主
题。

```

以上就是根据科技文献文档所固有的规范格式特点,通过聚类统计用户感兴趣的主体。这里为了减少计算复杂度,没

有利用同义词典和近似词典。虽然不够精确,但因为只要文档数据库库容较大,就不妨碍产生用户的扩展兴趣主题。

总结 本文先就文档聚类的现状及常用的模型——向量空间模型进行了介绍,并提出基于BP神经网络的文档词语权重计算模型。基于这种计算模型,针对科技文献文档所固有的规范格式特点,给出了两种新的文档聚类算法,一种用来产生和用户查询相关的文档集合,一种用来产生用户的扩展兴趣主题。这两种聚类算法对设计个性化智能化的文档检索系统是有借鉴意义的。

当然,还有一些问题需要解决,例如:在训练BP网络获取文档词语权重时,网络的初值选取问题;在对科技文档进行聚类时,不同格式文档中词语的提取问题等。

参考文献

- (1) 韩(Han, J.), 数据挖掘: 概念和技术. 北京: 高等教育出版社, 2001
- (2) 闻新, 周露, 王丹力, 熊晓英. MATLAB 神经网络应用设计. 北京: 科学出版社, 2000
- (3) Steinbach M, Karypis G, Kumar V. A Comparison of Document Clustering Techniques. www.acm.org
- (4) Huang L. A survey On Web Information Retrieval Technologies. www.acm.org
- (5) Bollacker K D, Lawrence S, Lee Giles C. Discovering relevant scientific literature on the web. IEEE Intelligent systems, 2000, 15(2): 42~47
- (6) Martin J D. Clustering Full Text Documents. www.acm.org
- (7) Information Retrieval Survey -1997. www.acm.org
- (8) 王实, 高文. 数据挖掘中的聚类方法. 计算机科学, 2000, 27(4): 42~45
- (9) 赵仲孟, 张蓓, 沈均毅. 对搜索引擎未来发展的探讨. 计算机科学, 2001, 28(3): 60~61
- (10) 鲁松. 自然语言处理中词相关性知识无导获取和均衡分类器构建. [博士论文]. 中国科学院, 2001

(上接第108页)

图像大小的稳定性其次。对于闭眼、戴眼镜等干扰较大的图像有一定的稳定性,而对于大胡子、复杂背景的照片则识别效果稍差。

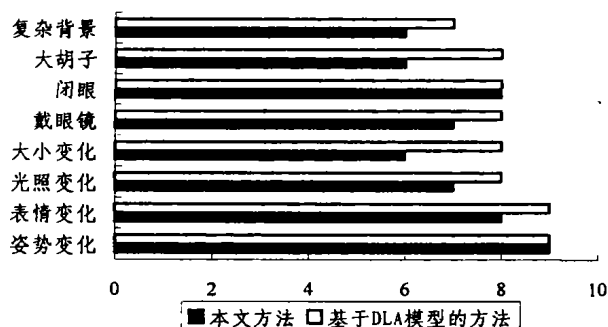


图4 各类情况下两种识别方法的结果比较

结论 本文运用弹性匹配的思想,主要从脸部对称性的利用、脸部模型的选取、人脸图像的各层小波子图的分析利用、相似度和代价函数的计算等几个方面对现有弹性图识别方法作了改进,提出了基于小波变换域的人脸弹性识别方法。实验证明,本文方法在总体识别率上接近于基于DLA模型

的方法,但在识别时间上仅为原方法的1/40,可大量节省查询时间。并且从具体实例上看,本文方法对人脸姿势、表情、光照以及大小变化等有较为稳定的识别效果。对于戴眼镜、闭眼、大胡须等易混淆的照片也具有较好的稳定性。

参考文献

- (1) Valentin D, Abdi H, et al. Connectionsist Models of Face Processing: A Survey, Pattern Recognition, 1994, 27(9): 1209~1229
- (2) Lades M, Vorbuggen J, Buhman J et al. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. on computer, 1991, 42(3): 300~311
- (3) Zhang Jun, et al., Face Recognition: Eigenface, Elastic Matching, and Neural Nets, Proc. of the IEEE, 1997, 85(9): 1422~1435
- (4) Averbuch, Lazar D, et al. Image Compression Using Wavelet Transform and Multiresolution Decomposition, IEEE Trans. On IP 1996, 5(1): 4~15
- (5) 田金文, 等. 变窗 Gabor 变换理论及其在图像处理中的应用. 红外与激光工程, 1998, 27(4): 1~5
- (6) Campell FW, Robson J G. Application of Fourier analysis to the visibility of gratings, J. Physiol. 1968, 19(7): 551~556