

一种改进的基于 NMF 的人脸识别方法

林庆^{1,2} 李佳¹ 雍建平¹ 廖定安³

(江苏大学计算机科学与通信工程学院 镇江 212013)¹ (南京大学计算机科学与技术学院 南京 210094)²
(常州纺织服装职业技术学院信息技术系 常州 213164)³

摘要 针对 NMF(非负矩阵分解)算法基于局部特征提取的特点,提出了一种对 NMF 基矩阵的处理方法,以提高在局部遮挡环境下人脸识别系统的识别率。首先使用离散小波变换得到样本的低频信息,利用 NMF 得到基矩阵;然后通过阈值判断提取能够突出表现人脸特征的部分,得到优化后的特征子空间,并将样本在该子空间上投影;最后使用支持向量机对所得到的投影系数分类。实验结果表明,优化算法其运算时间较短,且能有效地提高人脸在部分遮挡环境中的识别率。

关键词 非负矩阵分解,离散小波变换,人脸识别,基矩阵

中图分类号 TP391.41 **文献标识码** A

Improved Face Recognition Method Based on NMF

LIN Qing^{1,2} LI Jia¹ YONG Jian-ping¹ LIAO Ding-an³

(School of Computer Science and Telecommunications Engineering, Jiangsu University, Zhenjiang 212013, China)¹

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China)²

(Department of Information Technology, Changzhou Textile Garment Institute, Changzhou 213164, China)³

Abstract Non-negative matrix factorization(NMF) is a method of parts-based feature extraction, according to this characteristic. An improved algorithm dealing with NMF basis was introduced to enhance partial occlusion face recognition rate. Firstly, discrete wavelet transformation was used to produce a representation in the low frequency domain, and basic matrix was got according to the NMF method. Secondly, parts of face features which possess outstanding performance were extracted by threshold value judgments, and they were used to form optimized facial subspace feature. The training and testing images were projected to the optimized subspace feature. Finally support vector machine was used for classification. Experiments show that the improved algorithm's computing time is short, and this method achieves remarkable effects under the partially occluded circumstance.

Keywords NMF, DWT, Face recognition, Basic matrix

人脸识别是近年来模式识别的一个热点研究领域,随着计算机技术的发展,已经取得了很大的进步。现行的很多方法都要求具有良好的环境基础,但在实际生活当中,由于职业、隐私、宗教等一系列的原因,会有眼镜、口罩等遮挡物,而这些会导致系统识别率的下降。

人脸识别就是在人脸图像范围内提取出有利于分类的特征,并据此进行分类识别,一般包括预处理、特征提取和分类 3 个部分。人脸识别是典型的高维小样本训练问题,其中关键部分是特征提取,一个有效的降维方法能够提高整个系统的效率和识别率。子空间分析法是当前人脸识别中的主流方法之一,它具有描述性强、计算代价小、易实现及可分性好等特点^[1]。目前常用的子空间分析法有主元分析法(PCA)、线性判决分析法(LDA)和独立元分析法(ICA)。这三者的共性是先整个图像转为向量形式写入一个大矩阵 V 中,再根据各自相应的运算准则,将 V 近似分解为两个低维矩阵 W 和

H 的乘积,最后针对低维矩阵进行分类。它们在分解过程中都允许出现负值,但在实际中负值并没有物理意义,而且系数之间的正负会相互抵消,从而削弱特征,导致识别率下降。

1999 年, Lee 和 Seung 提出了非负矩阵分解(Nonnegative Matrix Factorization, NMF)算法^[2],即在分解过程中对基图像的像素点和重建系数都施加非负约束,这样重建图像就由基图像加性地叠加组合而成,它是一种基于局部特征的分解方法^[3],符合人类思维中“局部构成整体”的概念。但是直接用 NMF 进行人脸特征提取, NMF 基矩阵存在大量的冗余信息,不利于分类的进行。因此,研究者们提出了几种改进算法,如局部非负矩阵分解(LNMF)、Fisher 非负矩阵分解(FNMF)、稀疏非负矩阵分解(SNMF)、加权非负矩阵分解(WNMF)、局部敏感非负矩阵分解^[4]等,这些方法能在一定程度上提高识别率,但在迭代得到基矩阵时需要更大的计算量。

到稿日期:2011-06-26 返修日期:2011-09-29 本文受国家自然科学基金(61170126),江苏省自然科学基金(BK2009199, BK2010339)资助。
林庆(1962-),男,副教授,硕士生导师,主要研究方向为数据挖掘、模式识别与人工智能, E-mail: milam0479@126.com; 李佳(1987-),女,硕士生,主要研究方向为图像处理、模式识别与数据挖掘。

为此,本文提出了一种改进的算法:首先使用离散小波变换得到样本的低频信息,利用基本的 NMF 运算得到基矩阵,再通过阈值比较对所得到的基矩阵进行优化,最后利用支持向量机进行分类。

1 离散小波变换预处理

小波变换采用的是 Mallat 塔式分解法,即将待检测图像与一个低通和一个带通滤波器在水平方向和垂直方向上交替进行卷积,其本质就是信号的时间-尺度分析法^[5]。小波分解在时域和频域都具有较强的表征信号局部特征的能力,能够多分辨率地对信号进行分析,因此被称为“数学显微镜”。设 $f(t)$ 是平方可积函数,即 $f(t) \in L^2(R)$,则连续小波变换的定义如式(1)所示。

$$w_f(a, b) = \int_{-\infty}^{+\infty} f(t) \psi_{k,b}(t) dt \quad (1)$$

式中,尺度因子 a 必须大于零, b 为定位参数,函数 $\psi_{k,b}(t)$ 为小波或母小波,定义如式(2)所示。

$$\psi_{k,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (2)$$

令上式中的 $a = a_0^m, b = ka_0^m b_0$,即可得到离散小波的定义:

$$\begin{aligned} \langle f, \psi_{m,n} \rangle &= a_0^{-m/2} \int_{-\infty}^{+\infty} f(t) \psi_{m,n}(t) dt \\ &= a_0^{-m/2} \int_{-\infty}^{+\infty} f(t) \psi(a_0^m t - nb_0) dt \end{aligned} \quad (3)$$

当原始图像经过一层小波变换后,低频子图像的尺寸变为原图像的 1/4,有效地降低了原图像的维数,减少了计算的复杂度。本文采用 haar 小波作为小波基函数,一层小波分解后得到的子图如图 1 所示。



图 1 小波分解示意图

图 1(a) 为图像小波分解后的示意图,图 1(b) 为分解前的原始图像,图 1(c) 为图像经过小波分解后得到的 4 个子图。当一副图像经过小波分解后能提取出 4 个部分(LL, LH, HL, HH),其中 LL 为低频信号,是原始图像的近似,保留了绝大部分人脸信息,同时在一定程度上消除了光照和噪声的影响;LH 和 HL 为水平和垂直方向的子图,HH 为图像的最高频子图,对应角线方向,所包含的信息量最少。因此选取低频 LL 分量作为小波特征,既能保留人脸的全局信息,又能淡化局部细节。

2 特征提取

NMF 基本思想可以直观地描述为:对于任意一个给定的 $m \times n$ 维非负矩阵 V ,都可以找出两个非负矩阵 W 和 H 满足 $V \approx WH$ 。这里 W 的维数为 $m \times r$,称之为基矩阵; H 的维数为 $r \times n$,称之为系数矩阵, r 满足条件 $(m+n) \times r < m \times n$,这样一个高维的非负矩阵就分解成两个低维非负矩阵的乘积。由于分解前、后的矩阵仅包含非负元素,因此原矩阵 V 中的列向量就能解释成基矩阵 W 中所有列向量(即基向量)的加

权和,权重系数则是系数矩阵 H 中所对应列向量的元素。这时就得到了一个原始数据的压缩模型,矩阵 H 可以看作是样本 V 在基矩阵 W 上的非负投影系数,之后的运算都能根据系数矩阵 H 进行,大大降低了矩阵的维数。

在矩阵分解近似过程中,定义 $\|V - WH\|^2$ 为目标函数来保证逼近效果,这样优化问题就转为 $\|V - WH\|^2$ 最小化的问题,此时对应的迭代规则如式(4)、式(5)所示。

$$H_{oi} \leftarrow H_{oi} \frac{(W^T V)_{oi}}{(W^T W H)_{oi}} \quad (4)$$

$$W_{ia} \leftarrow W_{ia} \frac{(V H^T)_{ia}}{(W H H^T)_{ia}} \quad (5)$$

在 NMF 分解的迭代运算中,并没有对基矩阵 W 加以条件约束,数据之间存在着大量的冗余信息。为了解决这个问题, Lee 等人在此基础上构造了局部 NMF(LNMF)算法^[6],算法在迭代过程中对 W 引入了列正交的限制,使得 W 稀疏化。这样通过增强正交性,尽可能地降低了基图像之间的相关性,使权值分布较为集中,权值矩阵稀疏性增强。然而 LNMF 在迭代过程中对 H 的优化形式比较简单,采用了较多的近似表达,因此它对数据的描述力较差。

本文根据阈值判断对基矩阵进行了优化,在原始的 NMF 算法得到基矩阵之后,选择一个阈值,对基矩阵的每一列分别进行处理,将大于这个阈值的部分保留,其它区域置零,这样既能够集中提取出能表现人脸特征的区域,又能够减少其他不相关区域的干扰。整个处理过程计算简单,处理后的矩阵稀疏化,使得图像的特征向量中包含了更多的信息。本文将改进前、后的方法和 LNMF 进行比较,并在实验中小波变换预处理相结合。结果表明,这种方法在人脸识别中的识别率高于原始 NMF 和 LNMF,当人脸图像有所遮挡时,效果更为明显。

小波分解后, NMF、LNMF 和本文的方法得到的基矩阵分别如图 2(a)、2(b) 和 2(c) 所示。

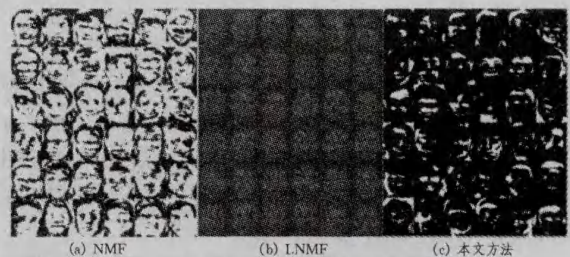


图 2 3 种方法得到的基矩阵

3 支持向量机分类

支持向量机(SVM)是建立在统计学理论基础上的,它采用一种核函数映射的方法,基本思想是通过一个非线性映射,将输入空间的数据映射到一个高维的特征空间中,使得样本在这一高维空间上线性可分。它比传统学习机器有着更好的分类能力和泛化性能,现已被公认为人工神经网络的替代方法^[7]。

SVM 实质上是一个二类分类器,但在本文实验中要将多个类别的人脸进行训练分类,这是一个典型的多分类问题。为解决这类问题,一般有两种策略:“一对一”方法和“一对多”方法。“一对一”方法是样品的 N 类进行两两分类,这时需

构造 $N(N-1)/2$ 个分类器。而“一对多”方法主要是在一个样本和其它 $N-1$ 类样本之间构造分类器(决策函数), 这样 N 类样本则只需构造 N 个分类器。当用两种方法进行分类时, 都出现了多个决策函数值相等的情况, 这时必须通过二次处理进行分类。

本文试验中人脸的样本数量为 40, 若用“一对一”方法则需构造 780 个分类器, 计算量过大。因此我们采用“一对多”方法, 当出现决策函数值相等时, 用 K 近邻法进行二次处理^[8]。

4 算法及实验分析

4.1 算法步骤

根据以上介绍, 本文提出的算法的具体步骤如下:

(1) 读取全部人脸样本并对每一幅图像用 haar 小波进行一层分解, 提取出分解后的 LL 低频子图按列存储在矩阵 V 中。

(2) 根据式(4)、式(5)的迭代规则对矩阵 V 进行非负矩阵分解, 得出 V 的基矩阵 W 。

(3) 对基矩阵 W 进行优化, 选择阈值 T , 将基矩阵按列分别二值化, 使其对应二值化矩阵 1 的位置, 并保留原矩阵的数据, 其余区域置零, 这样得到的新基矩阵记为 W_{new} 。

(4) 将 V 中的训练样本记为 V_{train} , 将其在基矩阵 W_{new} 上投影, 得到训练样本的特征空间 H_{train} 。

(5) 将 V 中的测试样本记为 V_{test} , 将其在基矩阵 W_{new} 上投影, 得到测试样本的特征空间 H_{test} 。

(6) 将 H_{train} 在支持向量机进行训练, 对 H_{test} 进行分类测试, 输出识别结果。

4.2 实验数据

本文的实验数据由两部分组成, 第一部分是剑桥大学提供的 ORL (Olivetti Research Laboratory) 图库, ORL 图库中包含有 40 个人, 每人各 10 幅图像共 400 幅, 每幅图像有 256 个灰度级, 大小为 112×92 。其中每个人的脸部表情和脸部细节都有着不同程度的变化, 如笑与不笑、眼睛睁与闭、戴与不戴眼镜等; 人脸姿态也有相当程度的变化, 深度旋转和平面旋转最大可达 20 度; 人脸的尺度也有多达 10% 的变化。我们随机选取图库中的 35 幅。第二部分是参照 ORL 图库的要求在实验室中拍摄的 5 组共 50 幅这样的图像, 试验中将两者混合在一起进行训练分类。对于人脸遮挡, 用 20×80 分别对眼睛和嘴部分进行遮挡, 如图 3 所示(图 3(a) 为数据库中的原始图像, 图 3(b) 为对眼睛进行遮挡时的图像, 图 3(c) 为对嘴进行遮挡时的图像)。

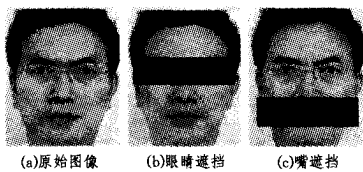


图 3 数据库中原图像及眼睛、嘴遮挡下的图像

对每个人随机选取 5 幅图像进行训练, 剩下的 5 幅作为测试图像。这样训练图库和测试图库中各有 200 幅。

4.3 实验结果及分析

经过 haar 小波提取后, 每幅图像低频子图 LL 大小为 56×46 , 将所有低频子图按列构成 2576 维向量, 即待分解矩阵

V 。选择特征向量的个数 r , 按迭代规则对 V 进行非负矩阵分解, 得到基矩阵 W , 此时 W 的维数为 $2576 \times r$ 。对 W 进行 4.1 节中步骤(3)的处理, 得到优化后新的基矩阵 W_{new} , 将训练样本和测试样本在 W_{new} 上投影, 分别得到 r 维的训练特征空间和测试特征空间, 最后采用支持向量机进行分类。

在对本文方法进行实验验证的过程中, 对特征基的个数 r 取 25、36、49 这 3 个值, 在无遮挡的环境中, 测试对应不同阈值 T 所得到的识别率。根据图 4 所示的实验结果, 本文确定后续实验中的阈值 T 为 0.3。

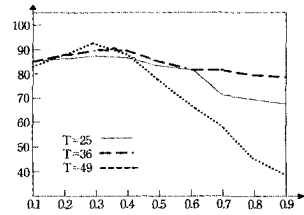


图 4 不同阈值对应的识别率

采用 NMF 对原始矩阵 V 进行分解时, 特征基的个数 r 是一个重要的待定参数, 它是分解后特征子空间的维数。但是至今还没有一个公认的 r 值计算方法, 因此我们取不同的 r 值进行试验, 以期得到一个理想的识别率。实验中, 采用 DWT+NMF 和 DWT+LNMF 来与本文的方法进行对比。在不同环境下对应不同 r 值时, 这 3 种方法的实验结果及运算时间分别如图 5—图 7 所示, 图 5(a) 表示识别率, 图 5(b) 为运算时间(单位为 s)。

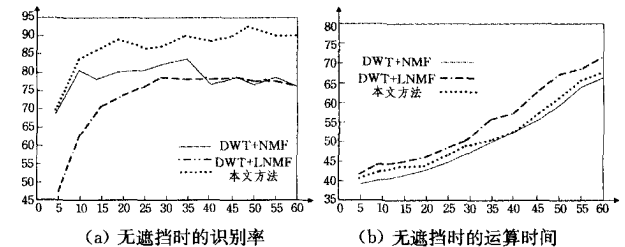


图 5

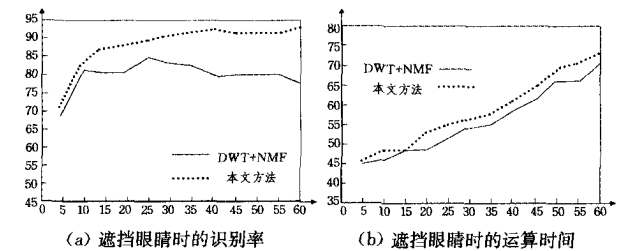


图 6

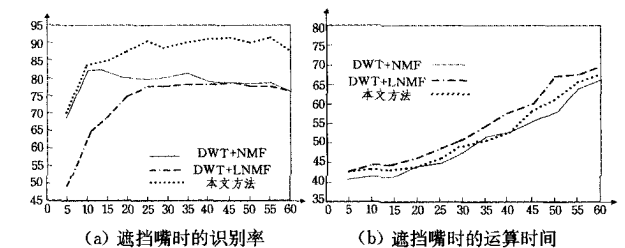


图 7

当在眼睛部分遮挡的环境下, DWT+LNMF 的识别效果很差, 故其识别率不用进行对比。从图中可以看出:

(下转第 270 页)

参考文献

- [1] 侯志强,韩崇昭. 视觉跟踪技术综述[J]. 自动化学报,2006,32(4):603-617
- [2] 高琳,唐鹏,盛鹏. 基于前景分割和特征空间自适应选择的视觉目标跟踪[J]. 控制与决策,2010,25(2):207-209
- [3] Comaniciu D,Ramesh V,Meer P. Kernel-based Object Tracking [J]. IEEE Transation on Pattern Analysis and Machine Intelligence,2003,25(5):564-577
- [4] Cheng Y. Mean Shift Mode Seeking and Clustering[J]. IEEE Transation on Pattern Analysis and Machine Intelligence, 1995,17(8):790-799
- [5] 一种具有跟踪外观变化目标能力的均值漂移算法[J]. 计算机科学,2007,34(10):244-253
- [6] 袁广林,薛模根,韩裕生,等. 基于自适应多特征融合的 mean shift 目标跟踪[J]. 计算机研究与发展,2010,47(9):1663-1671
- [7] Comaniciu D,Ramesh V,Meer P. Real-time Tracking of Non-rigid Objects Using Mean Shift[C]//IEEE Conference on Computer Vision and Pattern Recognition. Washington D C, USA, 2000:142-149
- [8] Tuzel O,Porikli F,Meer P. Region Covariance: A Fast Descriptor for Detection and Classification[C]//9th European Conference on Computer Vision. Graz, Austria,2006:589-600
- [9] Porikli F, Tuzel O, Meer P. Covariance Tracking Using Model Update Based on Lie Algebra [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. New York, NY, USA,2006:728-735
- [10] Ojala T, Pietikäinen M, Harwood D. A Comparative Study of Texture Measures with Classification Based on Feature Distributions[J]. Pattern Recognition,1996,29:51-59
- [11] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA,2005:886-893
- [12] 李晓磊,邵之江,钱积新. 一种基于动物自治体的寻优模式:鱼群算法[J]. 系统工程理论与实践,2002,11:32-38
- [13] Forstner W, Moonen B. A Metric for Covariance Matrices[R]. Department of Geodesy and Geoinformatics, Stuttgart University,1999
- [14] 黄华娟,周永权. 改进型人工鱼群算法及复杂函数全局优化方法[J]. 广西师范大学学报:自然科学版,2008,26(1):194-197
- [15] 刘耀年,李迎红,刘俊峰,等. 基于人工鱼群算法的径向神经网络的研究[J]. 东北电力大学学报,2006,26(04):24-27

(上接第245页)

(1)当 r 值较小时,这些方法的识别率都随着 r 值的增大而增大;但当 r 值增加到一定值时,识别率趋于稳定,甚至当 r 值继续增加时,识别率有下降的趋势。当 r 值较大时(特别是当 r 值大于30),所需要的运算量和时间都将大大增加。因此,在实际应用中, r 的取值必须综合考虑识别率和运算量。

(2)在正常无遮挡的条件下,三者的识别率都比较稳定,本文的方法比其他两者略有优势。但在部分遮挡的环境中,DWT+LNMF和DWT+NMF这两种方法就变得很不稳定,特别是在眼睛遮挡的环境中,本文方法的优势更为明显,这应该与眼睛部分包涵的信息量大有关系。

(3)当 r 值大于15时,基矩阵所描述的特征已经趋于稳定。在3种情况下用本文方法进行试验所得到的识别率相差不是很大,甚至有个别 r 值在有遮挡情况下的识别率大于无遮挡的情况,这表明了本文的方法具有很好的表征局部特征的能力。

(4)在3种条件下,对于3种方法所需要的运算时间,DWT+NMF最快,本文方法次之,DWT+LNMF需要的时间最长,但三者差距并不很大。这是因为本文算法在DWT+NMF的基础上多了优化的过程,而LNMF在分解过程中引入了列正交的限制,因此迭代运算更为复杂,需要的时间更多。但是相对于结果识别率的提高,本文算法运算时间的增加是可接受的。

结束语 本文对基于非负矩阵分解的人脸识别方法进行改进,首先利用离散小波变换得到图像低频信息,去除了一定的噪声和光照的影响,降低了维数;再利用基本NMF运算得到基矩阵,根据阈值比较的方法对基矩阵进行了优化,提取出了能突出表现人脸特征的部分;最后将图像在优化后的特征子空间上投影,用得到的投影系数在支持向量机上进行分类。

实验结果表明,以上的改进算法是可行的,在一定程度上有效地提高了人脸图像的识别率,特别是在人脸图片有遮挡的环境下,且与原来的算法相比,运算时间相差不明显。但是本文没有定量地分析特征向量的维数 r 值和基矩阵处理时的阈值 T 的选取,只是将通过实验数据得到的识别率作为选择的标准,因此这两个值的确定还需进行进一步的研究。

参考文献

- [1] Zafeiriou S, Petrou M. Nonlinear Nonnegative Component Analysis[C]//IEEE Conference on Computer Vision and Pattern Recognition. 2009:2860-2865
- [2] Lee D D, Seung H S. Learning the parts of objects by non-negative matrix factorization[J]. Nature,1999,401(6755):788-791
- [3] 刘维湘,郑南宁,游屈波. 非负矩阵分解及其在模式识别中的应用[J]. 科学通报,2006,51(3):241-250
- [4] 姜伟,杨炳儒,隋海峰. 局部敏感非负矩阵分解[J]. 计算机科学,2010,37:211-214
- [5] 徐佩霞. 小波分析与应用实例[M]. 合肥:中国科学技术大学出版社,2001,182-185
- [6] Kim J, Lee C, Kim J, et al. A Changeable Biometric System That Uses Parts-Based Localized Representation for Face Recognition [C] // IEEE Workshop on Automatic Identification Advanced Technologies. 2007:165-168
- [7] Li Yu-jian, Liu Bo, Yang Xin-wu, et al. Multiconltron: A General Piecewise Linear Classifier[J]. IEEE Transactions on Neural Networks,2011,22(2):276-289
- [8] Lin Qing, Hu Rui-rui, Zhan Yong-zhao. Facial Expression Feature Extraction Based on Integral Image [C] // International Conference on Advanced Computer Theory and Engineering. 2009:1345-1351