

人脸的民族特征抽取及其识别

段晓东<sup>1,2</sup> 王存睿<sup>1,2</sup> 刘向东<sup>1</sup> 刘 慧<sup>2</sup>  
(大连民族学院非线性信息技术研究所 大连 116600)<sup>1</sup> (东北大学研究生院 沈阳 110004)<sup>2</sup>

**摘 要** 人脸的民族特征是人脸信息描述的重要特征之一。首先构建了中国多民族人脸数据库,利用人脸识别技术提取民族面部特征和民族识别。在特征抽取方法中,采集人脸中的代数特征和几何特征,采用 LDA(线性判别分析)算法提取人脸图像的代数特征。还构建了能够抽取人脸几何特征的弹性模板,并利用 Gabor 小波进行特征点定位。实验采用 KNN 和 C5.0 分别学习训练集中的代数特征和几何特征,并对测试集进行预测分类。实验结果表明,利用代数特征方法和几何特征方法对藏族、维吾尔族、壮族 3 个民族的平均识别准确率分别达到 79%和 90.95%。  
**关键词** 面部民族特征,人脸识别,LDA,PCA,Gabor  
中图法分类号 TP391.41 文献标识码 A

Minorities Features Extraction and Recognition of Human Faces

DUAN Xiao-dong<sup>1,2</sup> WANG Cun-rui<sup>1,2</sup> LIU Xiang-dong<sup>1</sup> LIU Hui<sup>2</sup>  
(The Research Institute of Nonlinear Information Technology,Dalian Nationalities University,Dalian 116600,China)<sup>1</sup>  
(The Graduate School,Northeastern University,Shenyang 110004,China)<sup>2</sup>

**Abstract** Minorities feature of face is one of the most important features of face features. We created a face database of ethnic minorities and extracted facial features using face recognition technology. In the feature extraction method, we adapted the algebra and geometry from face database, used LDA algorithm to extract the algebraic features of human face images, this paper also constructed a new face templates to extract the geometric features and used gabor wavelet to locate the points of face templates. KNN and C5.0 Classifiers were used to learn the train dataset. The result indicates that the average recognition accuracy rates of Tibetan, Uyгур and Zhuang reach 79% by algebraic features and 90.95% by geometry features.  
**Keywords** Minority characters of face, Face recognition, LDA, PCA, Gabor Wavelet

人脸的民族特征研究对于科学技术和促进民族团结都具有重要的意义。中国是一个多民族国家,各民族在几千年历史中相互影响融合,各个民族的面部特征也在发生着变化,这也从一个侧面见证了多民族团结发展的历史进程。各民族在相互融合的过程中,面部的民族特征不仅没有消失,而且演化出新的特征。及时数字化保留中国各民族的面部特征,研究和发现在民族融合中产生的新的特征,对于民族文化保护和研究中华民族发展历程等方面都具有重要的现实意义。

1 人脸的民族特征

面部民族特征是继人的性别、年龄、表情等特征后又一重要表述人脸信息的集合特征。相关研究对于人脸识别技术是有益的补充和丰富。研究表明,不同民族之间面部特征存在一定差异,而且同一民族具有一定的地区特点。表 1 列出了本文涉及的 3 个民族的一些面部民族特征<sup>[1-3]</sup>。

表 1 中国几个主要少数民族的面部差异

民族	面部特征
----	------

藏族	眼睛:眼裂开度中等趋窄,多数人显示内眦褶
	鼻子:鼻梁较直、细楞,鼻根高度大,鼻根凹小且靠近眉间
	嘴唇:中等唇厚,唇形稍凸;上红唇多数为中等厚,下红唇比上红唇厚
	面型:面部多呈长梭型及长梯型
	头发:发质较硬;发黑直,间有波型及卷曲型
维吾尔族	耳朵:1/3 耳壳上显示达尔文结节,耳垂形状多为圆形,部分为方形,三角形甚少
	肤色:黄褐、黑褐或古铜色
	眼睛:眼裂开度中等,眼为楔形结构,上脸皱褶明显,脸沿宽,脸裂大
	鼻子:鼻根中等偏高,鼻尖高
	嘴唇:上唇皮肤部高度多为中等和偏高,红唇多为中唇和薄唇
壮族	面型:卵圆形
	头发:发黑间有黑褐色,头发有直、波、卷曲三型
	耳朵:大多有达尔文节,耳垂多为圆形和方形
	肤色:肤色呈黄、白或棕褐。
	眼睛:眼裂开度较宽,蒙古褶较弱育,多数呈微显型
壮族	鼻子:鼻梁硬骨部稍下凹软骨部略上凸
	嘴唇:上唇皮肤部一般为中等高度,明显前凸;红唇厚度多数为中等偏厚
	面型:中面型和阔面型为主
	头发:发形直、发色黑
	耳朵:耳壳上多数没有达尔文结节,耳垂向下,悬挂成舌状,即圆形耳垂

到稿日期:2009-11-18 返修日期:2010-01-30 本文受国家自然科学基金,中央高校自主科研基金,国家民委重点项目(GM-2009-66)和辽宁省高校科研计划项目(2009A157)资助。  
段晓东(1963—),男,教授,主要研究方向为图像处理等;王存睿(1980—),男,讲师,主要研究方向为智能优化与模式识别, E-mail: cunrui@gmail.com;刘向东(1969—),男,教授,主要研究方向为图像处理等;刘 慧(1983—),女,硕士生,主要研究方向为图像处理等。

面部的民族特征抽取和识别研究在国内外开展得还相对较少。目前,相关研究包括 Shakhn-arovich<sup>[4]</sup> 利用 boosting 算法对亚洲人以及非亚洲人进行分类识别。Lu Xiaoguang 和 Anil K. Jain<sup>[5]</sup> 利用 LDA 技术对亚洲人以及非亚洲人种进行多维度面部特征提取,然后利用最近邻分类器对其进行识别分类。Satoshi Hosono 和 Erina Takikawa 等<sup>[6]</sup> 对蒙古人种、高加索人以及黑种人利用小波进行特征提取,然后利用 SVM 进行分类识别。

这些研究主要集中于世界上几个主要人种之间的面部特征对比研究,对于同一人种下的不同民族人脸特征研究在国内外开展得还相对较少。国内在 20 世纪 90 年代,才得到了国家和社会的关注,但专家及学者主要关注汉族人群,较少针对其他民族面部特征进行系统研究。

2 中国多民族人脸数据集

本文采用自己构建的中国多民族样本库。高质量的多民族人脸数据库是相关研究的基础和前提。通过学校的大力支持,采集了那些居住在民族地区或少数民族长期聚集地的少数民族的人脸样本信息,这些地区的民族更具有典型的民族特点。数据库中各民族样本数量相对均衡,使得识别过程中分类器具有较好的泛化能力。本研究已经初步建立起 17 个中国民族的人脸库。图 1 所示的是部分数据。



图 1 中国多民族人脸数据集部分样本

3 人脸民族特征提取

本文对数据集提取两种特征:代数特征和几何特征。首先对图像进行预处理,然后利用 LDA 算法提取代数特征,利用民族人脸弹性模板提取几何特征。

3.1 图像预处理

首先对民族人脸库进行图像预处理,去除图像中的噪声,加强民族特征信息以便于特征提取和识别。图像预处理分别采用了灰度变换和几何变换。灰度变换是对图像进行灰度均衡化,并且去除影响图片质量的背景;几何变换将图像大小归一化以及校正姿态。图像大小归一化就是将图像缩放到统一大小,以保证人脸的平移和尺度不变性;校正姿态是将图片的人脸双眼中心在同一条直线上。

为便于采集人脸几何特征,本文将左右眼宽平均值作为归一化基向量,对图像进行等比例变换。在数据采集过程中,如表 2 所列,本文采集数据表明眼睛的眼宽差异相对方差最小,为  $2.23 \times 10^{-3}$ 。图像经过归一化之后,可以提高数据质量,相应地提高分类算法识别效能。

表 2 3 个少数民族的面部差异(方差量级为  $10^{-3}$ )

民族	性别	眼宽		嘴宽		脸宽		鼻长	
		平均值	方差	平均值	方差	平均值	方差	平均值	方差

藏	男	0.477	3.3	0.654	3.6	1.85	9.1	0.864	1.8
	女	0.485	0.77	0.701	3.7	1.91	7.7	0.869	3.6
维吾尔	男	0.518	3.6	0.719	7.5	1.98	14.	0.845	5.9
	女	0.502	2.6	0.679	5.0	1.95	13.	0.866	6.7
壮	男	0.491	1.7	0.702	5.4	1.98	7.0	0.821	5.5
	女	0.487	1.4	0.702	5.3	2.03	11	0.859	5.9
平均值		0.493	2.23	0.693	5.08	1.95	10.3	0.85	4.90

3.2 LDA 提取代数特征

LDA 方法又称应用线性判别分析 (Linear Discriminant Analysis) 方法<sup>[7]</sup>,可从高维特征空间里提取出最具有判别能力的低维特征,即选择使类间和类内离散度的比值最大的特征。

该方法可从人脸图片信息选择出散布于类内的正交向量,并将其作为人脸代数特征的特征向量。假设人脸库中包含  $N$  个样本数据,每个图像有  $n$  个像素点,可将人脸分为  $c$  个民族类,每类的样本数为  $N_i$ , $x_{ij}$  表示第  $i$  类中的第  $j$  个人脸样本,则各类的人脸图像均值为

$$m_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{ij}$$
 (1)

因此,总体人脸均值为

$$m = \frac{1}{N} \sum_{i=1}^c \sum_{j=1}^{N_i} x_{ij}$$
 (2)

然后求人脸样本类间散度矩阵  $S_B$  和类内散度矩阵  $S_W$ :

$$S_B = \sum_{i=1}^c N_i (m_i - m)(m_i - m)^T$$
 (3)

$$S_W = \sum_{i=1}^c \sum_{j=1}^{N_i} (x_{ij} - m_i)(x_{ij} - m_i)^T$$
 (4)

LDA 算法的目标是找到一个最佳投影  $W_{opt}$ :

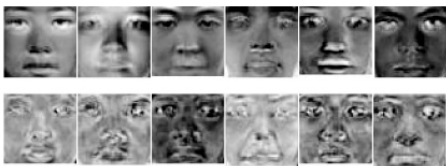
$$W_{opt} = \operatorname{argmax} \left| \frac{W^T S_B W}{W^T S_W W} \right| = [\omega_1, \omega_2, \dots, \omega_m]$$

若  $S_W$  是非奇异的,则  $\{\omega_i \mid i=1,2,\dots,m\}$  满足下式对应的  $m$  个最大特征值  $\{\lambda_i \mid i=1,2,\dots,m\}$  所对应的特征向量:

$$S_B \omega_i = \lambda_i S_W \omega_i, i=1,2,\dots,m$$
 (5)

本文将每一个民族的人脸图像看作一类。给定属于  $c$  个民族的  $N$  幅人脸图像训练样本,由于  $S_W$  的秩最多是  $N-c$ ,并且训练样本的数目小于人脸图像的维数,因此  $S_W$  是奇异概率较大,需通过相应方法变为非奇异。这里采用 PCA<sup>[8]</sup> 算法对图像进行降维至  $N-c$ ,使得  $S_W$  非奇异。

图 2 所示为训练样本特征向量的图像。可见第一排图片对应较大维度特征值,含高维人脸信息;后面对应最小特征值的特征向量则反映了图像集中的噪声情况。



(a) 前 6 个特征向量的图像 (b) 后 6 个特征向量的图像

图 2 PCA 算法降维后的人脸图像

具体步骤如下:

- Step 1 根据式(1)计算各民族人脸图像均值;
- Step 2 根据式(2)计算总图像均值;
- Step 3 利用 PCA 算法对图像进行降维;
- Step 4 根据式(3)和式(4)计算样本的类间散布矩阵  $S_B$  和类内散布矩阵  $S_W$ ;
- Step 5 根据式(5)计算广义特征值和特征向量;

Step 6 将广义特征值由大到小排列,取前  $d$  个较大特征值对应的特征向量组成特征空间;

Step 7 将每一幅人脸图像向上面得到的特征空间投影,得到一组坐标系数,就对应于子空间中的一个点。如下的向量表示  $y_i = W_f x_i$ 。  $y_i$  即为输入人脸  $x_i$  的  $d$  维向量特征描述,它可以用来在图像空间内恢复原图像或者作为提取的人脸特征进行分类识别。

3.3 民族特征弹性模板提取几何特征

基于几何特征的民族面部识别是指通过各民族人脸面部拓扑结构的先验知识,利用基于结构的方法在知识的层次上提取各民族人脸面部特征。几何特征通常包括面部主要特征点的位置以及指定点间的欧式距离或曲率、角度等,由这些特征便可以得到描述每张脸的特征矢量,将各民族人脸用一组几何特征矢量来表示。这要求选取的几何特征矢量具有一定的独特性,能够反映不同民族之间的差别,同时具有一定的弹性,以消除光照、姿态等的影响。

人类学相关研究表明,少数民族具有很多独特的面部特征,比如眉长轴和眼长轴及其之间夹角(眉长轴为眉毛最低点的连线,眼长轴为眼内外毗间的连线,例如蒙、藏族多为平行或外相交,维吾尔族多为平行或内相交)。同时,不同民族之间侧面人脸中的鼻根高度、位置和形状存在显著差异<sup>[9]</sup>。

本研究采用弹性模板进行特征提取。但由于民族特征不同于传统特征,需要构建与常规弹性模板不同的民族特征弹性模板。在弹性模板中引入了民族特征量(如图 3 所示),构建了提取民族特征的弹性模板。可见不同民族在此弹性模板下差异是明显的。

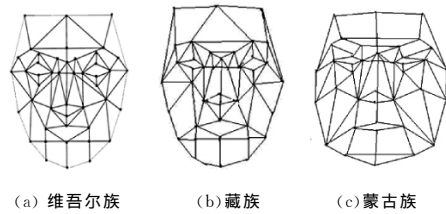


图 3 不同民族的弹性模板

由于 Gabor 小波<sup>[10]</sup>系数是通过计算一组不同中心频率与方向的 Gabor 滤波器和图像上给定位置附近像素灰度值的卷积来实现的,其具有对位置和光照不敏感的特点,因此本文利用 Gabor 小波<sup>[11,12]</sup>提取图像不同区域的不同频率成分和方向的纹理信息,如图 4 所示。本文建立人脸的特征点和弹性模板的对应部位的特征之间的联系,抽取能表达人脸各器官拓扑关系的几何特征。

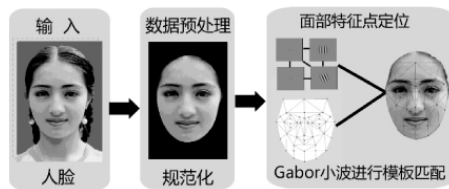


图 4 基于 Gabor 的弹性模板匹配

在图像预处理以及保证所有特征点能够相对均匀地覆盖图像的人脸区域时,利用 Gabor 小波通过弹性束图匹配来确定图 3 所示基本特征点的位置、基本特征点的选取(如眉、眼睛、鼻子、嘴巴以及脸部轮廓等)。其它特征点可通过计算若干个基本特征点的平均位置得到,这部分点称为派生特征点。

定位弹性模板的特征点后,将特征模板的坐标提取出来,本文开发了相关软件,如图 5 所示,将图像库中样本的弹性模板存储在 model 文件中,通过模板信息可以计算出很多民族特征的比例和角度。



图 5 人脸识别系统及生成的 .model 文件

4 实验及分析

本文选取数据库中的藏族、维吾尔族和壮族作为民族面部特征研究对象。在提取民族人脸数据库中的代数特征和几何特征后,将数据集分为训练集合测试集。为验证代数特征和几何特征各自的识别效果,本文分别进行了测试。

4.1 脸代数特征实验及结果

通过 LDA 方法对样本进行特征提取,得到最终的图像特征表示。由于代数特征提取的是数值向量,因此本文采用  $K$  近邻分类器(KNN)对测试集进行预测。KNN 算法中的距离采用欧式距离作为距离函数。

KNN 算法<sup>[13,14]</sup>的识别精度依赖  $K$  的取值。本文在实验的基础上得出  $K=5$  效果较好。本文共进行了 10 次实验。为了使结果不倾向于任何一个民族,每个类别的样本的训练集的数据样本个数相同。从 3 个民族中每个民族随意抽取 150 个样本作为一个民族的训练集,在剩余样本中随机抽取 70 个作为测试集。表 3 为平均识别率。

表 3 代数特征的识别结果

数据集	测试结果		
	藏族	维吾尔族	壮族
藏族测试集	0.771429	0.142857	0.085714
维吾尔族测试集	0.157143	0.785714	0.057143
壮族测试集	0.114286	0.071429	0.814286

实验结果表明,藏族识别率达到 77.1%,将 14.2%藏族误判为维吾尔族,8.5%误分类为壮族;维吾尔族识别率为 78.5%,15%误分类为藏族,5.7%误识别为壮族。壮族识别率为 81.4%,11.4%误识别为藏族,7.1%误识别为维吾尔族。平均识别准确率为 79%。

4.2 人脸代数几何特征实验及结果

将提取的人脸模板信息经过预处理,计算出相关的民族特征。从 3 个民族中每个民族随意抽取 150 个样本弹性模板作为一个民族的训练集,在剩余样本中随机抽取 70 个作为测试集。

定理 1 设  $S$  为数据集,包含类别集合为  $\{C_1, C_2, \dots, C_m\}$ ,  $P(S_i)$  表示类别  $C$  在数据集  $S$  中出现的概率。

定理 2 设属性  $A$  具有  $V$  个不同的值  $\{a_1, a_2, \dots, a_V\}$ ,可用属性  $A$  将  $S$  划分为  $V$  个子集  $\{S_1, S_2, \dots, S_V\}$ ,  $|S_V|$  分别为子集  $V$  中各个类别样本数的和,  $|S|$  为  $S$  中的样本总数。

为得到最终的分类规则,本文采用 C5.0 作为训练算法对训练集进行学习<sup>[15-17]</sup>。C5.0 算法是采用信息增益率(Gain ratio)衡量属性分裂数据的广度和均匀性,构造决策树。增益率采用综合熵值来度量,算法过程如下:

Step1 计算信息熵值  $E(S)$ ,衡量数据集  $S$  的无序程度:

E(S) = - \sum\_{i=1}^m P(S\_c) \times \log\_2 P(S\_c) \tag{6}

Step2 计算信息增益(Information Gain),用来度量序数改进的结果:

Gain(S,V) = E(S) - \sum\_{V \in Values(V)} \frac{|S\_V|}{|S|} \times E(S\_V) \tag{7}

Step3 增益中包含把数据集划分为更小有序子集的属性A的偏差。要减少这个偏差,采用 SplitInfor 来计算每个变量相对于它的 m 个变量值的熵:

SplitInfor(S,V) = \sum\_{V \in Values(V)} - \frac{|S\_i|}{|S|} \times \log\_2 \frac{|S\_i|}{|S|} \tag{8}

Step4 增益率采用式(9)计算,可减少较大值数据集的偏差。

GainRatio(S,V) = \frac{Gain(S,V)}{SplitInfor(S,V)} \tag{9}

利用 C5.0 学习训练集,得到的部分分类规则如下:

- 规则 1 规则准确率 90% 鼻宽<=0.589[类别:藏族]=>藏族  
鼻宽>0.589[类别:维吾尔族]  
    嘴角到同侧内眼角<=2.030[类别:壮族]=>壮族  
    嘴角到同侧内眼角>2.030[类别:维吾尔族]  
        鼻翼处脸宽<=4.462[类别:维吾尔族]  
        鼻宽<=1.481[类别:藏族]=>藏族  
        鼻宽>1.481[类别:维吾尔族]=>维吾尔族  
        鼻翼处脸宽>4.462[类别:壮族]=>壮族
- 规则 2 规则准确率 72.59%  
    下巴到嘴角<=1.526[类别:壮族]  
    下巴到鼻尖<=1.913[类别:藏族]=>藏族  
    下巴到鼻尖>1.913[类别:壮族]=>壮族  
    下巴到嘴角>1.526[类别:维吾尔族]  
        鼻翼处脸宽<=4.220[类别:维吾尔族]=>维吾尔族  
        鼻翼处脸宽>4.220[类别:藏族]=>藏族
- 规则 3 规则准确率 87.18%  
    嘴角到同侧内眼角<=2.030[类别:壮族]  
    下巴到鼻尖<=1.863[类别:藏族]=>藏族  
    下巴到鼻尖>1.863[类别:壮族]=>壮族  
    嘴角到同侧内眼角>2.030[类别:维吾尔族]  
        鼻孔到同侧嘴角<=0.797[类别:藏族]=>藏族  
        鼻孔到同侧嘴角>0.797[类别:维吾尔族]  
            右眼宽度<=0.962[类别:维吾尔族]=>维吾尔族  
            右眼宽度>0.962[类别:藏族]  
                鼻翼处脸宽<=4.089[类别:藏族]=>藏族  
                鼻翼处脸宽>4.089[类别:壮族]=>壮族
- 规则 4 规则准确率 86.95%  
    鼻宽<=0.589[类别:藏族]=>藏族  
    鼻宽>0.589[类别:壮族]  
        鼻宽<=0.733[类别:维吾尔族]  
        嘴角到同侧内眼角<=2.030[类别:壮族]=>壮族  
        嘴角到同侧内眼角>2.030[类别:维吾尔族]  
            鼻翼处脸宽<=4.462[类别:维吾尔族]=>维吾尔族  
            鼻翼处脸宽>4.462[类别:壮族]=>壮族  
                鼻宽>0.733[类别:藏族]=>藏族
- 通过这些规则对测试集进行分类预测,最终的预测结果

如表 4 所列。

表 4 几何特征识别结果

数据集	测试结果		
	藏族	维吾尔族	壮族
藏族测试集	0.885714	0.071429	0.042857
维吾尔族测试集	0.071429	0.900000	0.028571
壮族测试集	0.042857	0.014286	0.942857

实验结果表明,藏族识别率达到 88.6%,将 7.1%的藏族误判为维吾尔族,4.3%误分类为壮族;维吾尔族识别率为 90.0%,7.1%误分类为藏族,2.8%误识别为壮族。壮族识别率为 94.3%,4.2%误识别为藏族,1.4%误识别为维吾尔族。平均识别准确率为 90.95%。

4.3 实验分析

表 3 的实验结果表明,各民族面部在代数特征上是可分辨的,并且各民族均有一定的特征交叉。实验还表明,LDA+KNN 方法对于民族分类问题具有较好的判别能力。表 4 的实验结果表明基于几何特征的分类准确率要高于代数特征,可见人脸的民族特征在几何特征上区分度要远大于代数特征的区分度。

结束语 本文通过提取各民族的代数特征和几何特征来研究人脸的民族特征。利用 LDA+PCA 提取代数特征,利用 Gabor 小波和提取民族几何特征的弹性模板提取几何特征。实验结果表明,各民族人脸无论是代数特征还是几何特征都是可以区分的,几何特征的识别效果要优于代数特征,而且各民族之间均具有一定的交叉现象。

民族特征识别是人脸识别技术有益的补充,在理论研究和应用方面都具有广阔的前景。判断面部具备哪些民族特征或地域特征,并且在外貌特征上与哪些民族相近,是一件有意义的事情。实验也证明,本研究方法和技术路线是可行的。随着多民族人脸数据库的建立和完善,还有许多相关内容有待进一步的研究。此外,面部民族特征识别技术还可作为大规模人脸搜索时的弱分类器<sup>[18-20]</sup>,加快大规模人脸识别系统的匹配速度。民族特征识别可以扩展为多个国家的民族特征研究,对国际交往日益频繁背景下的国家安全也具有重要意义,同时可以更及时、全面地了解及掌握多民族发展现状。

参 考 文 献

[1] 张振标. 藏族的体质特征[J]. 人类学学报,1985(8)  
[2] 张振标,张建军. 广西壮族体质特征研究[J]. 人类学学报,1983(8)  
[3] 艾琼华,肖辉,赵建新,等. 维吾尔族的体质特征研究[J]. 人类学学报,1993(11)  
[4] Shakhnarovich G. A Unified Learning Framework for Real Time Face Detection and Classification[C]// Proceedings of the fifth International Conference on Automatic Face and Gesture Recognition. 2002;16-23  
[5] Lu Xiaoguang,Jain A K. Ethnicity Identification from Face Images[J]. Bull. Inst. Math. Acad. Sinica,2005(33):77-87  
[6] Hosoi S,Takikawa E,Kawade M. Ethnicity Estimation with Facial Image[C]// Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition. 2004: 10-16

(下转第 301 页)

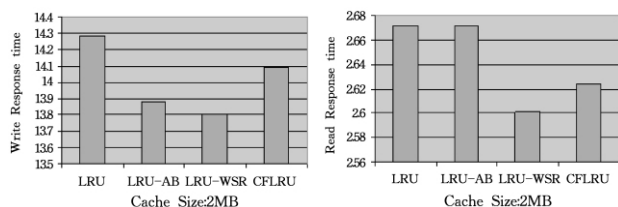


图 10 Postmark 写平均响应时间

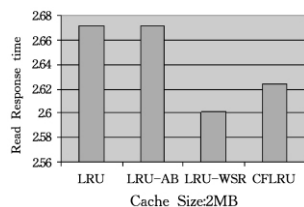


图 11 Postmark 读平均响应时间

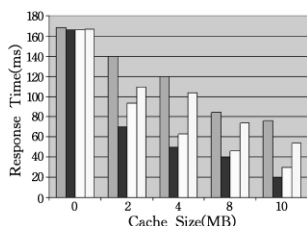


图 12 Financial 平均响应时间

### 3.3.3 清除效率的比较

图 13 给出了 3 类负载的清除效率。可以看出,对于算法 LRU-AB 而言,效率最低的是 Financial。这也再次验证了图 12 的结果。在 Cache 大小的所有变化情况下,LRU-AB 始终最优。而在 Iozone 和 Postmark 负载下,有技不如人的情况出现。

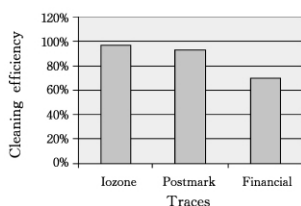


图 13 3 类负载的清除效率

**结束语** 本文详细介绍了我们提出的 SSD 置换算法 LRU-AB。它通过考虑数据访问的频度,参考 WOLF 算法进行有效的数据聚类,实现了对 SSD 读写性能的优化。通过仿真实验验证了该算法的可行性,同时建立了衡量清除效率的指标。通过对指标的仿真实验,得出结论如果负载在 SSD 中该指标数值偏高,那么 LRU-AB 会在某些情况下性能有所削弱。

## 参考文献

- [1] Wikipedia. Solid-state drive[EB/OL]. [http://en.wikipedia.org/wiki/Flash\\_drive](http://en.wikipedia.org/wiki/Flash_drive)
- [2] Agrawal N, Prabhakaran V, Wobber T, et al. Design Tradeoffs for SSD Performance[C]// Proceedings of the USENIX Technical Conference, June 2008
- [3] Wu M, Zwaenepoel W. eNVY: A Non-Volatile, Main Memory Storage System[C]// Proceedings of the 6th International Conference on Architectural Support for Programming Languages and Operating Systems, 1994:86-97
- [4] CompactFlash association. Information about CompactFlash [EB/OL]. <http://www.ssfcd.or.jp/>
- [5] Kim Jesung, Kim Jong Min, Noh S H, et al. A Space Efficient Flash Translation Layer for Compact Flash Systems[J]. IEEE Transactions on Consumer Electronics, 2002, 48:366-375
- [6] Intel corporation. Understanding the flash translation layer (FTL) specification[EB/OL]. <http://developer.intel.com/>
- [7] MTD. Memory Technology Device (MTD) subsystem for Linux [EB/OL]. <http://www.linux-mtd.infradead.org>
- [8] Park Seon-Yeong, Jung Dawoon, Kang Jeong-Uk, et al. CFLRU: a Replacement Algorithm for Flash Memory[C]// Proceedings of the 2006 International Conference on Compilers, Architecture and Synthesis for Embedded Systems, CASES'06, 2006:234-241
- [9] Wang Jun, Hu Yiming. WOLF-A Novel Reordering Write Buffer to Boost the Performance of Log-Structured File[C]// Proceedings of the FAST 2002 Conference on File and Storage Technologies, January 2002
- [10] Ganger G, Worthington B, Patt Y. The disksim simulation environment version 4.0 reference manual[R]. CMU-PDL-08-101, Carnegie Mellon University, May 2008
- [11] IOzone. org. IOzone Filesystem Benchmark[EB/OL]. <http://www.iozone.org>
- [12] Katcher J. PostMark: a New File System Benchmark[R]. TR 3022, Network Appliance, October 1997

(上接第 279 页)

- [7] Liu C, Wechsler H. Enhanced fisher linear discriminant models for face recognition[C]// Proceedings of Fourteenth International Conference on Pattern Recognition, Brisbane, Qld., Australia, 1998, 2:368-1372
- [8] Yambor W S. Analysis of PCA-based and fisher discriminant-based image recognition algorithms[R]. CS-00-103, Computer Science, Colorado State University, 2001
- [9] 肖明. 民族形象与解剖结构[M]. China Academic Journal Electronic Publishing Housing, 1994:86-93
- [10] Daugman J G. Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1988, 36(7):1169-1179
- [11] Dunn D, Higgins W E. Optimal Gabor filters for texture segmentation[J]. IEEE Transactions on Image Processing, 1995, 4(7):947-964
- [12] Jain A, Healey G. A multiscale representation including opponent color features for texture recognition[J]. IEEE Transactions on Image Processing, 1998, 7(1):124-128
- [13] Teknomo K. What is K Nearest Neighbors Algorithm? [EB/OL]. <http://people.revoledu.com/kardi/tutorial/KNN/Contents.htm>

- [14] D'Amato C, Malerba D, Esposito F, et al. Extending the K-nearest Neighbour classification algorithm to symbolic objects[C]// Convegno Scientifico Intermedio SIS. Università degli Studi di Napoli "Federico II", 9-11 Giugno 2003
- [15] Friedl M A, Brodley C E. Decision tree classification of land cover from remotely sensed data[J]. Remote Sensing of Environment, 1997, 61(3):399-409
- [16] Yang C-C, Prasher S O, Enright P, et al. Application of decision tree technology for image classification using remote sensing data[J]. Agricultural Systems, 2003, 76(3):1101-1117
- [17] Han Jiawei, Kamber M. 数据挖掘概念与技术[M]. 范明, 孟小峰, 等译. 北京:机械工业出版社, 2001:185-195
- [18] Campadelli P, Lanzarotti R. Automatic Facial Feature Extraction for Face Recognition [M]. Face Recognition, Kresimir Delac and Mislav Grgic, 2007:558
- [19] Brunelli R, Poggio T. Feature recognition: features versus template[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15(10):1042-1052
- [20] 龚雯, 陈丽华, 沈建国. 基于几何特征的人脸正面图像特征提取[J]. 现代计算机, 2005(9):61-63