doi:10.3969/j.issn.1001-893x.2016.04.003
引用格式:施静兰,常侃,张智勇,等.适用于彩色图像人脸识别的字典学习算法[J].电讯技术,2016,56(4):365-371.[SHI Jinglan, CHANG Kan, ZHANG Zhiyong, et al. A dictionary learning algorithm for color face recognition[J]. Telecommunication Engineering, 2016, 56(4):365-371.]

适用于彩色图像人脸识别的字典学习算法*

施静兰[1],常 侃**[1,2,3],张智勇[1],覃团发[1,2,3]

(1. 广西大学 计算机与电子信息学院,南宁 530004;
2. 广西高校多媒体通信与信息处理重点实验室(广西大学),南宁 530004;
3. 广西多媒体通信与网络技术重点实验室培育基地(广西大学),南宁 530004)

摘 要:现有的基于稀疏表示的人脸识别算法在识别前需要将彩色人脸图像转换成灰度人脸图像,这样虽然提高了运算速度,但忽视了不同色彩通道数据本身所包含的信息及它们之间的相关性。为了利用不同通道间相关性,基于标签一致的 K 奇异值分解(LC-KSVD)字典学习算法,提出了一种适用于彩色图像人脸识别的字典学习算法。该算法将 RGB 通道数据顺序排列成列向量,并在稀疏编码的环节中,对正交匹配追踪(OMP)算法的内积计算准则进行修正,以此提高字典原子的色彩表达能力。在彩色人脸数据库上进行实验,结果表明:所提出的字典学习算法能够有效地提高识别率。

关键词:彩色图像;人脸识别;稀疏表示;字典学习;稀疏编码
中图分类号:TP391.4 **文献标志码**:A **文章编号**:1001-893X(2016)04-0365-07

A Dictionary Learning Algorithm for Color Face Recognition

SHI Jinglan[1], CHANG Kan[1,2,3], ZHANG Zhiyong[1], QIN Tuanfa[1,2,3]

(1. School of Computer and Electronic Information, Guangxi University, Nanning 530004, China;
2. Guangxi Colleges and Universities Key Laboratory of Multimedia Communications and Information Processing, Guangxi University, Nanning 530004, China; 3. Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base), Guangxi University, Nanning 530004, China)

Abstract: The existing sparse-representation-based face recognition algorithms usually transform the color face images into gray images. Although this procedure increases the recognition speed, it ignores the information of the different color channels and the correlation among them. In order to utilize the correlation among different channels, based on the label consistent K-Singular Value Decomposition (LC-KSVD) algorithm, a new dictionary learning method for color face recognition is proposed. To improve the representing ability of each atom for color images, this algorithm concatenates R, G and B values into a single vector, and then introduces a new inner product into orthogonal matching pursuit (OMP) during sparse coding procedure. Experiments on different color face images datasets demonstrate that the proposed algorithm can achieve a higher recognition rate.

Key words: color images; face recognition; sparse representation; dictionary learning; sparse coding

1 引 言

人脸识别的研究始于 20 世纪 60 年代,是一项具有挑战性的任务[1]。随着图像处理、计算机视觉和模式识别等领域理论与技术的成熟,与之相关的人脸识别技术经过半个多世纪的发展,也得到了广泛应用。

* **收稿日期**:2015-10-28;**修回日期**:2016-03-11 Received date: 2015-10-28; Revised date: 2016-03-11
基金项目:国家自然科学基金资助项目(61401108, 61261023);广西自然科学基金资助项目(2013GXNSFBA019272)
Foundation Item: The National Natural Science Foundation of China (No. 61401108, 61261023); The Natural Science Foundation of Guangxi (2013GXNSFBA019272)
** **通信作者**:pandack0619@163.com **Corresponding author**: pandack0619@163.com

近年来,稀疏编码已经成功地应用于计算机视觉、机器学习和图像分析领域,包括图像去噪[2]、图像修复[3]、图像压缩[4]和图像分类[5]等。文献[6]中提出了基于稀疏表示的分类算法(Sparse Representation-based Classification,SRC),并应用于人脸识别。该方法直接使用整个训练图像集来构成字典,对于一张测试人脸图像,通过计算其在该字典上的稀疏编码系数来对其进行分类。与传统的人脸识别方法不同的是,SRC 人脸识别方法不需要像 Eigenface[7]和 Fisherface[8]方法一样进行明确的特征提取,而且对光照变化、表情变化和局部遮挡等问题鲁棒性高,这些优点使其成为了一个新兴的研究热点。

虽然基于 SRC 的人脸识别算法取得了一定成功,但是,文献[6]利用整个训练样本集作为字典进行稀疏编码,字典的维度高且不能有效表征训练样本,因此进行分类时复杂度高且识别率也有待进一步提升。直接有效的改进方法是采用精简的、由自适应学习得到的字典。著名的字典学习算法包括最优方向字典学习算法[9]和 K 奇异值分解(K-Singular Value Decomposition,K-SVD)字典学习算法[10]等,其中,K-SVD 算法循环进行稀疏编码和字典更新,字典学习的效率很高。但是,K-SVD 算法本身并没有考虑字典的区分能力,因此由该算法学习而得的字典不能直接应用于人脸分类。文献[11]通过根据输出的线性预测分类器进一步迭代更新由 K-SVD 算法训练出的字典,由此获得一个除了具有表示能力外还同时有利于分类的字典。文献[12]提出了区分性的 K-SVD(Discriminative K-SVD,D-KSVD)算法,通过在原始 K-SVD 算法的目标函数上加入区分项,保证学习出的冗余字典同时具有表示能力和区分能力,使字典和分类器的学习过程统一起来。文献[13]提出一种监督性算法——标签一致的 K-SVD(Label Consistent K-SVD,LC-KSVD)算法来学习出一个精简的和具有区分性的字典用于稀疏编码,通过在目标函数中加入“区分性”稀疏编码误差准则和“最佳”分类性能准则,并利用 K-SVD 算法进行优化求解,同时学习出具有区分性的字典和线性分类器。除了以 D-KSVD 和 LC-KSVD 为代表的改进的 K-SVD 算法外,还涌现了一些其他的字典学习算法,例如:香港理工大学的 Yang 和 Zhang 等人陆续提出了 Metaface 字典学习(Metaface Learning,MFL)算法[14]、Fisher 判别准则字典学习算法[15]等适用于人脸识别的字典学习算法。

虽然上述基于 K-SVD 的各类人脸识别算法均取得了较好的识别效果,但是现有的基于字典学习的人脸识别方法均是直接将彩色人脸图像转成灰度人脸图像后再进行人脸识别,这样在识别的过程中就仅应用了灰度图像的纹理信息。较少文献对彩色人脸图像进行处理的主要原因在于:首先,相比于灰度人脸图像,单纯通过增高数据维度带来的识别率提升有限;其次,数据维度增高会引入过高的复杂度,不利于算法的实际应用。但是,一方面,随着处理器计算能力的不断增强,高维度数据的处理难度在不断降低;另一方面,可以通过引入并行计算技术,例如图像处理单元(Graphics Processing Unit,GPU)等,进一步降低所需的计算时间。因此,与其他文献不同,本文讨论高维度的彩色人脸图像识别问题,希望通过探索不同色彩通道之间的相关性达到理想的识别率。

最为直接的彩色图像人脸识别方法是将每张彩色人脸图像的 R、G、B 3 个通道数据按顺序排列为一个列向量代替灰度图像向量进行人脸识别。以 LC-KSVD 算法[13]为例,可以应用此思想进行直接扩展,对应的方法在本文中称为 CE1-LC-KSVD(Color-Extension 1 of LC-KSVD)。但是,此方法是简单地直接增高数据维度,各色彩通道之间的相关性是完全被忽略的。为了合理利用色彩通道之间的相关性,从而进一步提升算法识别率,本文提出一种适用于彩色图像人脸识别的字典学习算法——CE2-LC-KSVD(Color-Extension 2 of LC-KSVD)。与 CE1-LC-KSVD 算法不同,在稀疏编码阶段,CE2-LC-KSVD 算法通过修正正交匹配追踪(Orthogonal Matching Pursuit,OMP)算法中的内积计算准则,对各个通道之间的相关性加以利用。这样使得学习出的字典能够体现通道间相关性,从而减少重构图像和原始图像之间的误差,最终提高识别率。

2 LC-KSVD 字典学习算法[13]

LC-KSVD 是一种监督字典学习算法,它旨在利用输入信号的监督信息(例如标签)来学习出一个具有重建能力和区分性的字典。为了利用训练样本的类别标签信息,将标签信息与每一个字典原子关联起来,加入一种新的标签一致的约束,并且结合重建误差和分类误差以构建最优化问题。文献[13]提出了两类 LC-KSVD 算法,分别记为 LC-KSVD1 和 LC-KSVD2。

2.1 LC-KSVD1

令 $\boldsymbol{Y}$ 表示一组维度为 n 的 N 个输入信号,即

$\boldsymbol{Y}=[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_N]\in\mathbf{R}^{n\times N}$,$\boldsymbol{D}=[\boldsymbol{d}_1,\boldsymbol{d}_2,\cdots,\boldsymbol{d}_K]\in\mathbf{R}^{n\times K}$为学习字典(其中,$K>n$ 以保证字典为冗余字典),$\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N]\in\mathbf{R}^{K\times N}$为输入信号 $\boldsymbol{Y}$ 的稀疏编码。线性分类器的性能决定于输入的稀疏编码 $\boldsymbol{X}$ 的区分性。为了利用学习得到的字典 $\boldsymbol{D}$ 获得具有区分性的稀疏编码 $\boldsymbol{X}$,定义如下字典构造的目标函数:

$$\begin{cases}<\boldsymbol{D},\boldsymbol{A},\boldsymbol{X}>=\underset{\boldsymbol{D},\boldsymbol{A},\boldsymbol{X}}{\arg\min}\|\boldsymbol{Y}-\boldsymbol{D}\boldsymbol{X}\|_2^2+\alpha\|\boldsymbol{Q}-\boldsymbol{A}\boldsymbol{X}\|_2^2\\ \text{s.t. }\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{1}$$

式中:α 是权重参数;T 为稀疏度;$\boldsymbol{x}_i$ 为第 i 个测试样本 $\boldsymbol{y}_i$ 在字典 $\boldsymbol{D}$ 上的稀疏编码系数;$\|\boldsymbol{Y}-\boldsymbol{D}\boldsymbol{X}\|_2^2$ 为重建误差;$\boldsymbol{Q}=[\boldsymbol{q}_1,\boldsymbol{q}_2,\cdots,\boldsymbol{q}_N]\in\mathbf{R}^{K\times N}$为输入信号 $\boldsymbol{Y}$ 的具有区分性的稀疏编码,如果 $\boldsymbol{q}_i$ 中的非零值位于输入信号 $\boldsymbol{y}_i$ 和字典原子 $\boldsymbol{d}_k$ 具有相同类别标签的地方,那么称 $\boldsymbol{q}_i=[q_i^1,q_i^2,\cdots,q_i^K]^{\mathrm{T}}=[0\cdots1,1,\cdots0]^{\mathrm{T}}\in\mathbf{R}^K$ 为与输入信号 $\boldsymbol{y}_i$ 相应的"区分性"稀疏编码;$\boldsymbol{A}$ 是一个线性变换矩阵,定义一个线性变换 $g(\boldsymbol{x},\boldsymbol{A})=\boldsymbol{A}\boldsymbol{x}$,它将原始的稀疏编码 $\boldsymbol{x}$ 转换成在稀疏特征空间 $\mathbf{R}^K$ 中最具区分性;$\|\boldsymbol{Q}-\boldsymbol{A}\boldsymbol{X}\|_2^2$ 项表示具有区分性的稀疏编码误差,它强制了稀疏编码 $\boldsymbol{X}$ 近似为具有区分性的稀疏编码 $\boldsymbol{Q}$,使得来自同一个类别的信号具有非常相似的稀疏表示(例如:令稀疏编码结果中标签一致),这样一来,即使使用一个简单的线性分类器也能获得好的分类结果。

2.2　LC-KSVD2

为了使得在分类中字典最优化,在字典学习的目标函数中加入分类误差项。在 LC-KSVD2 中,使用一个线性预测分类器 $f(\boldsymbol{x},\boldsymbol{W})=\boldsymbol{W}\boldsymbol{x}$。一个同时具有重建能力和区分能力的字典 $\boldsymbol{D}$ 的学习目标函数定义如下:

$$\begin{cases}\langle\boldsymbol{D},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}\rangle=\underset{\boldsymbol{D},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}}{\arg\min}(\|\boldsymbol{Y}-\boldsymbol{D}\boldsymbol{X}\|_2^2+\alpha\|\boldsymbol{Q}-\boldsymbol{A}\boldsymbol{X}\|_2^2+\\ \qquad\beta\|\boldsymbol{H}-\boldsymbol{W}\boldsymbol{X}\|_2^2)\\ \text{s.t. }\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{2}$$

式中:$\|\boldsymbol{H}-\boldsymbol{W}\boldsymbol{X}\|_2^2$ 项表示分类误差;$\boldsymbol{W}$ 为分类器参数;$\boldsymbol{H}=[\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_N]\in\mathbf{R}^{m\times N}$为输入信号 $\boldsymbol{Y}$ 的类别标签;$\boldsymbol{h}_i=[0,0,\cdots,1,\cdots,0,0]^{\mathrm{T}}\in\mathbf{R}^m$ 为与输入信号 $\boldsymbol{y}_i$ 相对应的标签向量,其中非零值的位置表明 $\boldsymbol{y}_i$ 的类别;α 和 β 为权重参数。

假设具有区分性的稀疏编码 $\boldsymbol{X}'=\boldsymbol{A}\boldsymbol{X}$ 和 $\boldsymbol{A}\in\mathbf{R}^{K\times K}$是可逆的,则 $\boldsymbol{D}'=\boldsymbol{D}\boldsymbol{A}^{-1}$,$\boldsymbol{W}'=\boldsymbol{W}\boldsymbol{A}^{-1}$。式(2)的目标函数可以重新写为

$$\begin{cases}\langle\boldsymbol{D}',\boldsymbol{W}',\boldsymbol{X}'\rangle=\underset{\boldsymbol{D}',\boldsymbol{W}',\boldsymbol{X}'}{\arg\min}(\|\boldsymbol{Y}-\boldsymbol{D}'\boldsymbol{X}'\|_2^2+\alpha\|\boldsymbol{Q}-\boldsymbol{X}'\|_2^2+\\ \qquad\beta\|\boldsymbol{H}-\boldsymbol{W}'\boldsymbol{X}'\|_2^2)\\ \text{s.t. }\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{3}$$

式中:第一项表示重建误差;第二项表示具有区分性的稀疏编误差能够使各个类别之间的稀疏编码具有区分性;第三项表示分类误差,通过 $\|\boldsymbol{H}-\boldsymbol{W}'\boldsymbol{X}'\|_2^2$ 使学习得到的分类器最优。

3　彩色图像的稀疏表示人脸识别

3.1　彩色图像 LC-KSVD 字典学习算法

在人脸识别中,可获取的样本数量通常远小于人脸图像样本的维数,因此由训练样本组成的字典原子色彩表示能力不高。已有的 LC-KSVD 字典学习算法[13]仅应用于灰色图像,而忽略了原始彩色图像中的大量信息。本文希望在 LC-KSVD 算法的基础上,有效地利用不同色彩通道数据之间的相关性(而不仅是简单地通过数据维度的增加),提高字典原子色彩表示能力,从而最终提高识别率。

传统的 K-SVD 算法包括稀疏编码和字典更新两个步骤。因为 OMP 算法高速有效,所以其被应用于 K-SVD 算法的稀疏编码阶段。文献[3]指出,将灰度图像的 K-SVD 字典训练扩展到彩色图像的一个简单方法是将 RGB 色彩空间 3 个通道的值顺序连接成一个列向量后再进行训练。但是,由于原始的 OMP 算法在重构过程中缺少一定的约束,并不能保证重建图像会包含原始图像的平均色,因此训练过程会产生伪色彩和假象,这是彩色图像处理中会遇到的典型问题。为了利用不同色彩通道数据之间的相关性提高字典原子的色彩表达能力,可以对 OMP 的内积计算准则进行适当的修正,使其适应于彩色图像的稀疏编码。

假设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 是两张不同的彩色人脸图像,表示为 $\boldsymbol{x}=[\boldsymbol{x}_{\mathrm{R}}^{\mathrm{T}},\boldsymbol{x}_{\mathrm{G}}^{\mathrm{T}},\boldsymbol{x}_{\mathrm{B}}^{\mathrm{T}}]^{\mathrm{T}}$,$\boldsymbol{y}=[\boldsymbol{y}_{\mathrm{R}}^{\mathrm{T}},\boldsymbol{y}_{\mathrm{G}}^{\mathrm{T}},\boldsymbol{y}_{\mathrm{B}}^{\mathrm{T}}]^{\mathrm{T}}$,其中 $\boldsymbol{x}_{\mathrm{R}}$、$\boldsymbol{x}_{\mathrm{G}}$、$\boldsymbol{x}_{\mathrm{B}}$ 和 $\boldsymbol{y}_{\mathrm{R}}$、$\boldsymbol{y}_{\mathrm{G}}$、$\boldsymbol{y}_{\mathrm{B}}$ 分别为图像 $\boldsymbol{x}$ 和图像 $\boldsymbol{y}$ 的 R、G、B 通道值列向量,维度为 n。定义一种新的 OMP 内积形式如下:

$$\langle\boldsymbol{y},\boldsymbol{x}\rangle_\gamma=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{x}+\frac{\gamma}{n^2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{K}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{T}}(\boldsymbol{I}+\frac{\gamma}{n}\boldsymbol{K})\boldsymbol{x}。\tag{4}$$

式中:γ 是权重参数;

$$\boldsymbol{K}=\begin{pmatrix}\boldsymbol{J}_n & 0 & 0\\ 0 & \boldsymbol{J}_n & 0\\ 0 & 0 & \boldsymbol{J}_n\end{pmatrix}。\tag{5}$$

式中:$\boldsymbol{J}_n$ 是一个 $n\times n$ 的矩阵,其元素全部为 1。等式第一项为原始的 Euclidean 内积,它能够保证获得当前最好的结果。包含矩阵 $\boldsymbol{K}$ 的第二项计算每一个通道中的 $\boldsymbol{E}(\boldsymbol{y})$ 和 $\boldsymbol{E}(\boldsymbol{x})$ 的估计量并将它们相乘,因此强制选择的原子考虑平均色。为了方便求解,令 $\gamma=2a+a^2$,则

$$\boldsymbol{I}+\gamma/n\boldsymbol{K}=(\boldsymbol{I}+a/n\boldsymbol{K})^{\mathrm{T}}(\boldsymbol{I}+a/n\boldsymbol{K})。\tag{6}$$

因此,一个简单的实现方法是,将每一张人脸图像和字典中的每一个原子都乘以 $\boldsymbol{I}+a/n\boldsymbol{K}$,并将其进行归一化,这样便可以实现新的内积准则。

由于只有训练图像和字典原子中存在 3 个色彩通道之间的相关性,因此,在彩色图像 LC-KSVD 字典学习算法中,采用 OMP 算法时,将训练图像和字典原子的 RGB 通道值顺序联结成一个列向量,然后分别乘以 $\boldsymbol{I}+a/n\boldsymbol{K}$,并采用新的 OMP 内积准则求解稀疏编码系数。

令 $\hat{\boldsymbol{Y}}=(\boldsymbol{I}+\frac{a}{n}\boldsymbol{K})\boldsymbol{Y}$、$\hat{\boldsymbol{D}}=(\boldsymbol{I}+\frac{a}{n}\boldsymbol{K})\boldsymbol{D}$,则 LC-KSVD1 算法的目标函数式(1)变为

$$\begin{cases}<\hat{\boldsymbol{D}},\boldsymbol{A},\boldsymbol{X}>=\arg\min\limits_{\hat{\boldsymbol{D}},\boldsymbol{A},\boldsymbol{X}}\|\hat{\boldsymbol{Y}}-\hat{\boldsymbol{D}}\boldsymbol{X}\|_2^2+\alpha\|\boldsymbol{Q}-\boldsymbol{A}\boldsymbol{X}\|_2^2\\ \text{s.t.}\quad\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases},\tag{7}$$

LC-KSVD2 的目标函数式(2)变为

$$\begin{cases}\langle\hat{\boldsymbol{D}},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}\rangle=\arg\min\limits_{\hat{\boldsymbol{D}},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}}(\|\hat{\boldsymbol{Y}}-\hat{\boldsymbol{D}}\boldsymbol{X}\|_2^2+\alpha\|\boldsymbol{Q}-\boldsymbol{A}\boldsymbol{X}\|_2^2+\\ \qquad\qquad\beta\|\boldsymbol{H}-\boldsymbol{W}\boldsymbol{X}\|_2^2)\\ \text{s.t.}\quad\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{8}$$

由此可见,彩色图像的 LC-KSVD 算法与原始 LC-KSVD 算法形式相同,仅在 OMP 步骤采用新的内积计算准则来处理伪色彩和假象问题。

3.2 彩色图像 LC-KSVD 字典学习算法的优化

本文同时考虑 LC-KSVD1 算法与 LC-KSVD2 算法的彩色图像扩展算法。由于两者的优化过程相同,接下来仅以彩色图像的 LC-KSVD2 算法为例描述优化过程。

为了求解式(8),将其重新写为

$$\begin{cases}\langle\hat{\boldsymbol{D}},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}\rangle=\arg\min\limits_{\hat{\boldsymbol{D}},\boldsymbol{W},\boldsymbol{A},\boldsymbol{X}}\left\|\begin{pmatrix}\hat{\boldsymbol{Y}}\\ \sqrt{\alpha}\boldsymbol{Q}\\ \sqrt{\beta}\boldsymbol{H}\end{pmatrix}-\begin{pmatrix}\hat{\boldsymbol{D}}\\ \sqrt{\alpha}\boldsymbol{A}\\ \sqrt{\beta}\boldsymbol{W}\end{pmatrix}\boldsymbol{X}\right\|_2^2\\ \text{s.t.}\quad\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{9}$$

引入两个新的变量如下:

$$\boldsymbol{Y}_{\text{new}}=(\hat{\boldsymbol{Y}}^{\mathrm{T}},\sqrt{\alpha}\boldsymbol{Q}^{\mathrm{T}},\sqrt{\beta}\boldsymbol{H}^{\mathrm{T}})^{\mathrm{T}};\tag{10}$$

$$\boldsymbol{D}_{\text{new}}=(\hat{\boldsymbol{D}}^{\mathrm{T}},\sqrt{\alpha}\boldsymbol{A}^{\mathrm{T}},\sqrt{\beta}\boldsymbol{W}^{\mathrm{T}})^{\mathrm{T}}。\tag{11}$$

式中:$\boldsymbol{D}_{\text{new}}$ 为 l_2 范数列归一化矩阵。则式(9)的最优化等价于求解如下问题:

$$\begin{cases}\langle\boldsymbol{D}_{\text{new}},\boldsymbol{X}\rangle=\arg\min\limits_{\boldsymbol{D}_{\text{new}},\boldsymbol{X}}\|\boldsymbol{Y}_{\text{new}}-\boldsymbol{D}_{\text{new}}\boldsymbol{X}\|_2^2\\ \text{s.t.}\quad\forall i,\|\boldsymbol{x}_i\|_0\leqslant T\end{cases}。\tag{12}$$

上述问题与传统 K-SVD 的目标函数形式相同,因此可以直接通过 K-SVD 算法求解。

3.3 分类方法

通过上述学习过程最终得到字典 $\boldsymbol{D}_{\text{new}}$ 后,便可利用其对待测人脸图像进行分类。从 $\boldsymbol{D}_{\text{new}}$ 中获得 $\hat{\boldsymbol{D}}=\{\hat{\boldsymbol{d}}_1,\hat{\boldsymbol{d}}_2,\cdots,\hat{\boldsymbol{d}}_K\}$,$\boldsymbol{A}=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_K\}$,$\boldsymbol{W}=\{\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_K\}$。由于 $\hat{\boldsymbol{D}}$、$\boldsymbol{A}$ 和 $\boldsymbol{W}$ 共同进行过 l_2 范数归一化,因此还不能直接使用它们进行分类。分类操作所需的 $\widetilde{\boldsymbol{D}}$、变换参量 $\widetilde{\boldsymbol{A}}$ 和分类器参量 $\widetilde{\boldsymbol{W}}$ 计算方法如下:

$$\widetilde{\boldsymbol{D}}=\left\{\frac{\hat{\boldsymbol{d}}_1}{\|\hat{\boldsymbol{d}}_1\|_2},\frac{\hat{\boldsymbol{d}}_2}{\|\hat{\boldsymbol{d}}_2\|_2},\cdots,\frac{\hat{\boldsymbol{d}}_K}{\|\hat{\boldsymbol{d}}_K\|_2}\right\};\tag{13}$$

$$\widetilde{\boldsymbol{A}}=\left\{\frac{\boldsymbol{a}_1}{\|\hat{\boldsymbol{d}}_1\|_2},\frac{\boldsymbol{a}_2}{\|\hat{\boldsymbol{d}}_2\|_2},\cdots,\frac{\boldsymbol{a}_K}{\|\hat{\boldsymbol{d}}_K\|_2}\right\};\tag{14}$$

$$\widetilde{\boldsymbol{W}}=\left\{\frac{\boldsymbol{w}_1}{\|\hat{\boldsymbol{d}}_1\|_2},\frac{\boldsymbol{w}_2}{\|\hat{\boldsymbol{d}}_2\|_2},\cdots,\frac{\boldsymbol{w}_K}{\|\hat{\boldsymbol{d}}_K\|_2}\right\}。\tag{15}$$

对于测试图像 $\boldsymbol{y}_i$,首先通过求解以下优化问题计算它的稀疏表示系数 $\boldsymbol{x}_i$:

$$\boldsymbol{x}_i=\arg\min_{\boldsymbol{x}_i}\{\|\boldsymbol{y}_i-\widetilde{\boldsymbol{D}}\boldsymbol{x}_i\|_2^2\}\ \text{s.t.}\ \|\boldsymbol{x}_i\|_0\leqslant T,\tag{16}$$

然后简单地使用线性预测分类器 $\widetilde{\boldsymbol{W}}$ 来判断图像 $\boldsymbol{y}_i$ 的标签 j:

$$j=\arg\max_j(\boldsymbol{l}=\widetilde{\boldsymbol{W}}\boldsymbol{x}_i)。\tag{17}$$

式中:$\boldsymbol{l}\in\mathbf{R}^m$ 为类别标签向量。

4 实验结果与分析

在实验部分,将提出的彩色图像 LC-KSVD 字典学习算法称为 CE2-LC-KSVD。为了全面分析算法性能,与灰度图下的 LC-KSVD 算法[13]进行比较(需注意,灰度图下有 LC-KSVD1 和 LC-KSVD2,因此在彩色人脸图像下对应的改进算法分别为 CE2-LC-KSVD1 和 CE2-LC-KSVD2)。此外,将 R、G、B 通道数据排列成列向量,直接代替原有灰度图像进行字典学习的方法称为 CE1-LC-KSVD。

由于许多主流的人脸图像库为灰度图像,本文仅选择 AR 和 CMU_PIE 两个彩色人脸图像库进行实验。其中,AR 人脸库下包含与人脸色彩差距鲜明的遮挡物,CMU_PIE 包含更为复杂的光照情况。

实验中,α 和 β 与文献[13]中取相同值,分别为

16 和 4，OMP 新内积计算准则中的 a 取 0.001。实验中使用的是台式计算机，Intel(R)Core(TM)2 Duo CPU，主频3 GHz，4 GB内存。

4.1　AR 人脸库实验结果

AR 人脸库[16]包含来自 126 个人超过4 000张的正面人脸图像，每张图像尺寸为 120 pixel×165 pixel。每类有 26 张图像，选取两个不同时期，分别在不同光照、表情以及面部遮挡条件下拍摄。遮挡包括墨镜遮挡和围巾遮挡。AR 人脸库中部分人脸图像如图 1 所示。实验中，我们选择一组包含 50 个男性对象和 50 个女性对象的人脸子集，并将人脸图像缩小为 40 pixel×55 pixel。实验分以下两部分进行：

(1)对于每个对象，随机选择 N_1 张人脸图像作为训练样本，其余作为测试样本，字典维度为 500；

(2)对于每个对象，选择第一时期和第二时期拍摄的只有表情变化和光照变化无遮挡的 14 张人脸图像，从中随机选择 N_2 张作为训练样本，其余作为测试样本，字典维度为 100。

图 1　AR 人脸库部分人脸图像
Fig. 1 Part of AR Database

实验 1 的结果列于表 1。从表 1 可以观察到：与 LC-KSVD 算法相比，CE1-LC-KSVD 算法的识别率有大幅提升，说明通过增加数据维度，能够获得更高的识别率；与 CE1-LC- KSVD 算法的识别率相比，CE2-LC-KSVD 有 1.7% ~2.8% 的提升，说明新的 OMP 内积计算准则能够有效地利用不同色彩通道数据之间的相关性信息来提高字典原子的色彩表达能力。此外，还可以观察到：与 CE1-LC-KSVD 算法相比，CE2-LC-KSVD 的字典学习时间并没有明显增加，也进一步说明了本文算法的优势。

表 1　AR 人脸库中各算法识别率与字典学习时间(实验 1)
Tab. 1 Recognition rates and associated dictionary learning times by different methods on the AR Database(Experiment 1)

算法	识别率/%		时间/s	
	$N_1=15$	$N_1=20$	$N_1=15$	$N_1=20$
LC-KSVD1	83.1	86.0	226.0	249.5
LC-KSVD2	83.5	86.2	234.4	260.6
CE1-LC-KSVD1	90.7	91.5	524.5	569.4
CE1-LC-KSVD2	90.9	93.2	533.8	581.4
CE2-LC-KSVD1	92.4	93.7	547.0	598.3
CE2-LC-KSVD2	93.2	96.0	554.3	612.4

实验 2 的结果列于表 2。与表 1 的结果类似，CE1-LC-KSVD 算法和 CE2-LC-KSVD 算法均能够取得比 LC-KSVD 算法更高的识别率，但是识别率提升幅度低于表 1 的结果。原因在于 AR 人脸库中的部分人脸图像具有遮挡，而遮挡物与人脸肤色差距较大，这些色彩信息在字典学习中发挥了重要的作用。实验 1 的样本中包含具有遮挡物的人脸图像，而实验 2 的样本中去除了此类图像，因此可利用的色彩信息较少。另外，从字典学习时间上看，CE2-LC-KSVD 算法的速度明显快于 CE1-LC-KSVD 算法，说明在训练样本色彩接近的情况下，本文修正的 OMP 内积方法有利于字典学习过程的快速收敛。

表 2　AR 人脸库中各算法识别率与字典学习时间(实验 2)
Tab. 2 Recognition rates and associated dictionary learning times by different methods on the AR Database(Experiment 2)

算法	识别率/%		时间/s	
	$N_2=7$	$N_2=10$	$N_2=7$	$N_2=10$
LC-KSVD1	90.6	91.5	78.6	97.5
LC-KSVD2	90.4	92.2	81.6	101.4
CE1-LC-KSVD1	91.6	92.2	222.7	275.2
CE1-LC-KSVD2	91.7	93.0	226.0	279.9
CE2-LC-KSVD1	91.9	92.7	78.7	106.7
CE2-LC-KSVD2	92.0	93.0	79.9	108.9

4.2　CMU_PIE 人脸库实验结果

CMU_PIE 人脸库[17]包含 68 个人的超过4 000张人脸图像，每张图像尺寸为 140 pixel×200 pixel。每个人有 13 种姿势、43 种不同的光照条件和 4 种不同表情，包括“illumination”和“light”两个子集：“illumination”子集背景光关闭，仅有不同方向的闪光灯照亮，每个类有 21 幅图像；“light”子集同时包含背景光和闪光灯，每个类有 22 幅图像，其中第一张图像只有背景光。“illumination”子集部分人脸图像如图 2 所示，“light”子集部分人脸图像如图 3 所示。

图 2　CMU_PIE 人脸库“illumination”子集部分人脸图像
Fig. 2 Part of CMU-PIE Database “illumination” subset

图 3　CMU_PIE 人脸库“light”子集部分人脸图像
Fig. 3 Part of CMU-PIE Database “light” subset

实验选择 CMU_PIE 人脸库中的“illumination”

子集和“light”子集,将每张图像缩小为 35 pixel×50 pixel。对于每个对象,随机选择 N_3 张人脸图像作为训练样本,其余作为测试样本,字典维度为 340。

CMU_PIE 的实验结果列于表 3。从表 3 中可以看出:与表 1 和表 2 中的结果类似,CE2-LC-KSVD 算法的识别率最高,CE1-LC- KSVD 算法次之,LC-KSVD 算法最低,从而验证了彩色图像人脸识别的有效性。需要注意的是:首先,彩色图像下识别率的提升幅度低于 AR 人脸库实验 1(详见表 1),原因在于 CMU_PIE 两个子集中的人脸图像没有出现色差很大的遮挡物,因此色彩信息少于 AR 人脸库实验 1;其次,彩色图像下识别率的提升幅度高于 AR 人脸库实验 2(详见表 2),原因在于 CMU_PIE 两个子集中的光照情况较为复杂,因此色彩信息相对 AR 人脸库实验 2 而言更丰富。

表 3 CMU_PIE 人脸库中各算法识别率与字典学习时间

Tab. 3 Recognition rates and associated dictionary learning times by different methods on the CMU_PIE Database

算法	识别率/%		时间/s	
	$N_3=6$	$N_3=7$	$N_3=6$	$N_3=7$
LC-KSVD1	89.9	94.7	43.8	49.7
LC-KSVD2	93.0	98.2	46.6	53.9
CE1-LC-KSVD1	94.9	96.3	169.0	174.3
CE1-LC-KSVD2	95.0	98.7	172.3	173.8
CE2-LC-KSVD1	98.1	98.6	172.9	179.5
CE2-LC-KSVD2	98.8	98.8	175.0	187.6

4.3 综合讨论与分析

CE1-LC-KSVD 算法同时利用了 R、G、B 3 个色彩通道的数据值,导致数据维度大大增加,所以其识别效果会比灰度图像的 LC-KSVD 算法更好。但是,CE1-LC-KSVD 算法仅通过单纯的数据维度的增加来获取更高的识别率,并没有考虑不同色彩通道之间的相关性。因此,这种高维度数据的应用方法效率不高,算法识别率也还有进一步提升的空间。而 CE2-LC-KSVD 算法采用了新的 OMP 内积计算准则,在稀疏编码中考虑原子的平均色,有效地解决了训练过程中产生的伪色彩和假象等问题,所以提高了字典原子的色彩多样性。因此,CE2-LC-KSVD 算法能够有效地利用 3 个色彩通道之间的相关性,进一步提高识别率。

另一方面,因为两种彩色图像下的字典学习算法 CE1-LC-KSVD 和 CE2-LC-KSVD 将 3 个色彩通道的数据排列成列向量处理,所以每个训练样本以及字典原子的维度是灰度图像情况下的 3 倍,导致字典学习时间明显增加。为了解决此问题,一种较为有效的方法是应用并行计算技术,例如 GPU 等降低运算时间。需要注意的是,CE2-LC-KSVD 算法在稀疏编码步骤中采用了新的内积计算准则,仅通过将每一张人脸图像和字典中的每一个原子同时乘以 $\boldsymbol{I}+a/n\boldsymbol{K}$ 来实现。因此,在大部分情况下,CE2-LC-KSVD 算法学习时间仅比 CE1-LC-KSVD 算法略长。根据 AR 人脸库实验 2 的结果可知,在训练样本间色彩接近的情况下,所采用的修正的 OMP 内积方法有利于字典学习过程的快速收敛,因此该情况下所提出的 CE2-LC-KSVD 算法优势更为明显。

最后需要提出的是,在实际应用中,不可避免地会存在人脸图像颜色不正的问题。为了解决上述问题,可以增大训练集,让训练集涵盖尽可能多的光照、遮挡等情况,从而最大程度地保证彩色图像字典的鲁棒性。

5 结束语

为了利用 R、G、B 色彩通道的信息及不同通道之间的相关性,以提高字典原子的色彩表示能力,本文将 3 个色彩通道的数据排列成列向量,并在 OMP 稀疏编码中采用新的内积计算准则。在 AR 和 CMU_PIE 两个彩色人脸库上进行实验,并与灰度图像下的 LC-KSVD 算法以及直接将其扩展到彩色图像下的 CE1-LC-KSVD 算法进行比较。实验结果表明本文算法能有效提高识别率。但是本文算法同时应用了 3 个色彩通道的数据,提高了算法对内存的需求,字典训练时间也有明显增加。在未来的工作中,研究重点之一是通过并行化技术,例如 GPU 等降低字典学习时间,以提高算法的实用性。此外,也将探索 Lab 色彩空间下的人脸识别算法性能。与 RGB 色彩空间不同,Lab 色彩模型是均匀的。可以尝试将彩色图像转换到 Lab 空间,再根据需要采用灰度图像(仅 L 通道)或彩色图像(L、a、b 通道)的字典学习方法。

参考文献:

[1] ZHAO W, CHELLAPPA R, PHILLIPS P J, et al. Face recognition: a literature survey[J]. ACM Computing Surveys, 2003, 35(4): 399-458.

[2] ELAD M, AHARON M. Image denoising via sparse and redundant representations over learned dictionaries[J]. IEEE Transactions on Image Processing, 2006, 15(12): 3736-3745.

[3] MARIAL J, ELAD M, SAPIRO G, et al. Sparse represen-

tation for color image restoration[J]. IEEE Transactions on Image Processing,2008,17(1):53-69.
[4] BRYT O,ELAD M. Compression of facial images using the K-SVD algorithm[J]. Journal of Visual Communication and Image Representation,2008,19(4):270-282.
[5] YANG J C,YU K,GONG Y H,et al. Linear spatial pyramid matching using sparse coding for image classification[C]//Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL: IEEE,2009:1794-1801.
[6] WRIGHT J,YANG A Y,GANESH A,et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2009,31(2):210-227.
[7] TURK M,PENTLAND A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience,1991,3(1):71-86.
[8] BELHUMEUR P N, HESPANHA J P, KRIEGMAN D J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1997,19(7):711-720.
[9] ENGAN K,AASE S O,HUSOY J H. Frame based signal compression using method of optimal directions (MOD)[C]//Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS). Orlando, FL: IEEE, 1999:1-4.
[10] AHARON M,ELAD M,BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation[J]. IEEE Transactions on Signal Processing,2006,54(11):4311-4322.
[11] PHAM D S,VENKATESH S. Joint learning and dictionary construction for pattern recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Anchorage, AK: IEEE, 2008:1-8.
[12] ZHANG Q,LI B X. Discriminative K-SVD for dictionary learning in face recognition[C]//Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, CA: IEEE, 2010: 2691-2698.
[13] JIANG Z L,LIN Z,DAVIS L S. Label consistent K-SVD: learning a discriminative dictionary for recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2013,35(11):2651-2664.
[14] YANG M,ZHANG L,YANG J,et al. Metaface learning for sparse representation based face recognition[C]// Proceedings of 2010 IEEE International Conference on Image Processing (ICIP). Hong Kong: IEEE, 2010: 1601-1604.
[15] YANG M,ZHANG L,FENG X C,et al. Fisher discrimination dictionary learning for sparse representation[C]//Proceedings of 2011 IEEE International Conference on Computer Vision (ICCV). Barcelona: IEEE, 2011:543-550.
[16] MARTINEZ A,BENAVENTE R. The AR face database[R]//CVC Technical Report #24. Barcelona: Centre de Visióper Computador,1998.
[17] SIM T,BAKER S,BSAT M. The CMU pose,illumination,and expression (PIE) database[C]//Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition. Washington, DC: IEEE, 2002:46-51.

作者简介:

施静兰(1990—),女,广西南宁人,2013 年于广西大学获工学学士学位,现为硕士研究生,主要研究方向为人脸识别、稀疏表示、图像处理;

SHI Jinglan was born in Nanning, Guangxi Zhuangzu Autonomous Region, in 1990. She received the B. S. degree from Guangxi University in 2013. She is now a graduate student. Her research concerns face recognition, sparse representation and image processing.

Email: shijinglan@mail.gxu.cn

常　侃(1983—)男,广西南宁人,2010 年于北京邮电大学获博士学位,现为广西大学计算机与电子信息学院副教授,主要研究方向为压缩感知、稀疏表示、图像处理;

CHANG Kan was born in Nanning, Guangxi Zhuangzu Autonomous Region, in 1983. He received the Ph. D. degree from Beijing University of Posts and Telecommunications in 2010. He is now an associate professor. His research interests include compressed sensing, sparse representation and image processing.

Email: pandack0619@163.com.

张智勇(1991—),男,广西扶绥人,2014 年于广西大学获工学学士学位,现为硕士研究生,主要研究方向为稀疏表示、低秩矩阵及应用;

ZHANG Zhiyong was born in Fusui, Guangxi Zhuangzu Autonomous Region, in 1991. He received the B. S. degree from Guangxi University in 2014. He is now a graduate student. His research concerns sparse representation, low rank matrix and its application.

Email: zhangzhiyong1160@163.com

覃团发(1966—),男,广西宾阳人,1997 年于南京大学获博士学位,现为广西大学教授、计算机与电子信息学院副院长、中国电子学会高级会员、中国通信学会高级会员,主要研究方向为无线多媒体通信、网络编码。

QIN Tuanfa was born in Binyang, Guangxi Zhuangzu Autonomous Region, in 1966. He received the Ph. D. degree from Nanjing University in 1997. He is now a professor and the Vice Dean of School of Computer and Electronic Information, Guangxi University, and also a senior member of China Institute of Electronics and senior member of China Communications Institute. His research interests include wireless multimedia communications, network coding.

Email: tfqin@gxu.edu.cn