

# 基于因素分解模型的两步人脸识别

程载和

(无锡职业技术学院 无锡 214121)

**摘要** 为了减轻人脸识别中表情以及姿态等因素变化对识别结果的影响, Xu 提出了利用原始样本和对称样本的两步人脸识别算法。但当人脸图像受外在因素干扰产生较大变化时, 该方法的识别结果并不理想。因此提出了一种基于因素分解模型的两步人脸识别算法。新算法在特征提取过程中利用因素分解模型将“身份因素”和“表情因素”从人脸图像中分离出来, 加以控制。然后提取测试集图像中的新身份和新表情, 并将其与训练集中的旧身份或旧表情相互作用, 合成新的人脸图像。同时为了保证分类精度, 在识别阶段针对原始样本和合成样本分别采用两步人脸识别的方法, 充分利用了分数层次融合的优势, 进一步提高了算法的识别效果。

**关键词** 人脸识别, 表情因素, 因素分解模型

**中图分类号** TP391 **文献标识码** A

## Two-step Face Recognition Based on Factorization Models

CHENG Zai-he

(Wuxi Institute of Technology, Wuxi 214121, China)

**Abstract** To overcome the effect of facial expression and pose on face recognition, Xu proposed using the original and “symmetrical face” training samples to perform representation based two-step face recognition. However, it usually gets bad symmetrical face samples based on mirror image with the changes of facial poses, which may affect the accuracy of recognition. A two-step face recognition algorithm based on factor decomposition model was proposed. Firstly, to reduce the effect of light, pose and other factors, the factorization model is used to separate the “expression factors” and “identity factor” from the face images which can be controlled, to construct new image samples. Meanwhile, the two-step face recognition method is adopted in the recognition stage, which makes full use of the advantage of the score level fusion, and further improves the recognition rate of the proposed algorithm.

**Keywords** Face recognition, Expression factor, Factorization model

## 1 引言

人脸识别是模式识别的一个重要研究课题, 该领域的研究人员做了大量的工作<sup>[1-7]</sup>。Seitz 等人将投影几何 (Projective Geometry) 的基本原理与 Image Morphing 相结合, 提出了 View Morphing 技术<sup>[1]</sup>。Mukaigawa 等人在人脸合成方面的研究也做出了一定贡献。Cootes 等人在研究中发现, 人脸头部姿态的变化可以通过人脸图像统计模型中的一些变化表示出来<sup>[2-3]</sup>。Tenenbaum 和 Freeman<sup>[4]</sup>提出了双线性因素模型, 并将其用于人脸图像研究中, 取得了比较理想的实验结果。

通常, 更多有效的训练样本则更容易描述图像由于外在因素影响所引起的变化, 从而在特征提取过程中保留更有效的分类信息。但在现实生活中, 可用于训练的样本图像往往远小于样本图像的空间维数, 容易出现“小样本”问题, 导致一些典型的特征提取算法不能全面地表示人脸图像特征, 影响最终的分类结果。

为了缓解上述问题, Xu 等人<sup>[5]</sup>提出了基于人脸对称性的

两步人脸识别算法。该算法在训练过程中不仅充分利用了原始样本信息, 同时结合人脸的对称性特征, 生成新的训练样本。这样不仅可以增加有效训练样本的数量, 同时利用人脸对称性, 也可以真实地反映人脸图像中的一些变化。进而 Xu 等人还在识别阶段引入两步人脸识别的方法, 充分利用了分数层次融合的优势, 降低了算法的识别错误率。但当人脸图像受外在因素干扰较大时, 利用图像镜像原理生成的新样本的效果并不理想, 一定程度上影响了算法的识别效果。为此, 我们在 Xu 等人的算法的基础上提出了一种新的研究思路: 用双线性形式的因素分解模型来解决人脸的识别问题。即将“人的身份”和“面部表情”两个因素看作影响因素, 将人脸图片表示成由这两个因素构成的双线性函数形式, 实现基于“人的身份”和“面部表情”两个因素构成的双线性模型的人脸的图像合成。进而再将原始样本和合成样本一起当作训练本来训练, 采用两步人脸识别算法<sup>[5]</sup>实现最终的分类。

## 2 因素分解模型

本文主要讨论通过双线性因素分解模型从人脸图像中分

离出“身份因素”和“表情因素”并加以控制,从而生成新的人脸图像。

## 2.1 模型原理及其双线性形式

为了更加直观地介绍双线性模型的原理,本文分别用“内容”与“形式”来表示影响待测数据的两个隐含因素。将所有待测数据组成一个数据集,“形式  $s$ ”用  $I$  维向量表示,“内容  $c$ ”用  $J$  维向量表示,  $y^{sc}$  是“形式  $s$ ”和“内容  $c$ ”下的  $K$  维观测向量,它的第  $k$  维分量可以表示成  $a^s$  和  $b^c$  的双线性函数:

$$y_k^{sc} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ijk} a_i^s b_j^c = (a^s)^T W_k b^c \quad (1)$$

其中,  $\omega_{ijk}$  是两种因素相互作用的权值,  $W_k$  是由  $\{\omega_{ijk}\}$  组成的  $I \times J$  维矩阵。此外,式(1)还可以表示为:

$$y^{sc} = \sum_{i=1}^I \sum_{j=1}^J \omega_{ij} a_i^s b_j^c \quad (2)$$

其中,  $\omega_{ij}$  表示  $\{\omega_{ijk}\}$  组成的  $K$  维向量。

观察式(2)可以发现,模型中的“内容”和“形式”两个因素是相互对称的,简称为“对称模型”。如果将式(2)中“形式”的所有项进行合并,记作  $p_j^s = \sum_{i=1}^I \omega_{ij} a_i^s$ ,则模型的非对称形式可以表示为:

$$y^{sc} = \sum_{j=1}^J p_j^s b_j^c = P^s b^c \quad (3)$$

其中,  $P^s = [p_1^s, p_2^s, \dots, p_J^s]$  是  $K \times J$  维矩阵,它描述了从内容空间到观测空间的线性变化。

同理,若将式(2)中的“内容”相关项合并在一起,记作  $p_i^c = \sum_{j=1}^J \omega_{ij} b_j^c$ ,则式(2)可以表示为:

$$y^{sc} = \sum_{i=1}^I p_i^c a_i^s = P^c a^s \quad (4)$$

其中,  $P^c = [p_1^c, p_2^c, \dots, p_I^c]$  是  $K \times I$  维矩阵,它描述了从形式空间到观测空间的线性变化。

## 2.2 模型训练

对于任意给定的一组数据观测集,都可通过 SVD 方法进行模型训练,求出交互项  $W$ 。假设在  $C$  种内容、 $S$  种形式下,观测向量组成的阵列,将所有  $K$  维观测向量写成  $(KS) \times C$  维堆叠矩阵的形式:

$$Y = \begin{bmatrix} y^{11} & \dots & y^{1C} \\ \dots & \dots & \dots \\ y^{S1} & \dots & y^{SC} \end{bmatrix}_{(KS) \times C} \quad (5)$$

则模型非对称形式,即式(3)可以简记为:

$$Y = PB \quad (6)$$

其中,  $(KS) \times J$  维矩阵  $P$  和  $J \times C$  维矩阵  $B$  分别定义如下:

$$P = \begin{bmatrix} p^1 \\ p^2 \\ \dots \\ p^S \end{bmatrix}_{(KS) \times J}, B = [b^1, b^2, \dots, b^C]_{J \times C} \quad (7)$$

对观测矩阵  $Y$  作奇异值分解,可得:

$$Y = USV^T \quad (8)$$

其中,对角阵  $S$  中的元素按照从大到小的顺序排列,将  $US$  的

前  $J$  列和  $V^T$  的前  $J$  行分别取作式(6)中的  $P$  和  $B$ 。这样,非对称形式的模型训练过程完成。

对称模型训练要借助于非对称模型进行。记  $Y^{VT}$  为  $Y$  的向量转置(Vector Transpose),即:

$$Y^{VT} = \begin{bmatrix} y^{11} & \dots & y^{S1} \\ \dots & \dots & \dots \\ y^{1C} & \dots & y^{SC} \end{bmatrix}_{(KC) \times S} \quad (9)$$

定义  $(KI) \times J$  维堆叠权重矩阵  $W$  为:

$$W = \begin{bmatrix} \omega^{11} & \dots & \omega^{1J} \\ \dots & \dots & \dots \\ \omega^{I1} & \dots & \omega^{IJ} \end{bmatrix}_{(KI) \times J} \quad (10)$$

同样  $W^{VT}$  为  $W$  的向量转置。借助这些定义,通过将所有观测向量写在一起,对称模型表达式(2)可以写为下面两个等同的形式:

$$Y = [W^{VT} A]^{VT} B \quad (11)$$

$$Y^{VT} = [WB]^{VT} A \quad (12)$$

其中,  $I \times J$  维矩阵  $A$  和  $J \times C$  维矩阵  $B$  分别是形式向量和内容向量的集中写法,其中:

$$A = [a^1, a^2, \dots, a^S]_{I \times S}, B = [b^1, b^2, \dots, b^C]_{J \times C} \quad (13)$$

这样,  $A$  和  $B$  可根据式(11)和式(12)以迭代方式求解出来。至此,对称模型训练过程完毕。

当然,本文主要是利用因素分解模型,通过给定人脸图像的训练集,当提供一个新表情、新人脸的测试图像时,求出新表情下训练集内人脸的样子或者新人脸在训练集已有表情下的样子,简称为因素分解模型的“转移”功能。可描述如下:

对于一个未知的测试向量  $\tilde{y}$ ,要使已训练好的因素分解模型同样适应于新内容  $\tilde{C}$  和新形式  $\tilde{S}$ ,且同时保持从原有训练集得到的交互项  $W$  不变。参考式(12),“转移”问题实际上是以  $a^{\tilde{S}}$  和  $b^{\tilde{C}}$  为自变量的代价函数的最小化问题:

$$f(a^{\tilde{S}}, b^{\tilde{C}}) = \|\tilde{y} - [Wb^{\tilde{C}}]^{VT} a^{\tilde{S}}\|^2 \quad (14)$$

对式(14)进行求导即可得到最优解为:

$$a^{\tilde{S}} = \text{pinv}([Wb^{\tilde{C}}]^{VT}) \tilde{y} \quad (15)$$

$$b^{\tilde{C}} = \text{pinv}(W^{VT} a^{\tilde{S}})^{VT} \tilde{y} \quad (16)$$

其中,  $\text{pinv}(\dots)$  表示矩阵的广义逆。可以一次轮流通过式(15)和式(16)来估算参数  $a^{\tilde{S}}$  和  $b^{\tilde{C}}$ ,直到收敛。这里,迭代初值可以取为训练集中形式向量或内容向量的均值。迭代过程结束后,可以预测未知形式  $\tilde{S}$  下的已知内容  $C$  以及未知内容  $\tilde{C}$  下的已知形式  $S$ ,可分别通过下式表示:

$$y^{\tilde{S}C} = [W^{VT} a^{\tilde{S}}]^{VT} b^C \quad (17)$$

$$y^{S\tilde{C}} = [W^{VT} a^S]^{VT} b^{\tilde{C}} \quad (18)$$

本文主要通过因素分解模型将“人的身份”和“面部表情”看作影响人脸图像两种变化因素并从人脸图像中分离出来,即分别对应内容矩阵  $B$  和形式矩阵  $A$ 。然后利用因素分解模型的“转移”功能合成旧人脸的新表情图像或者新人脸的旧表情图像。具体可以根据实验中数据库的实际情况择优选择。

### 3 基于因素分解模型的两步人脸识别

#### 3.1 算法的主要步骤

Xu 提出的两步人脸识别方法中充分利用了分数层次融合的优势,进一步提高了算法的识别效果。基于因素分解模型的 TSFR 算法主要可以概括为以下几个步骤<sup>[5]</sup>:

步骤 1 利用双线性因素分解模型的方法合成与原始人脸图像相对应的新的人脸样本图像。

步骤 2 利用两步人脸识别算法(TSFR)对原始人脸图像进行训练分类,得到待测试样本属于第  $j$  类的分值  $s_j^1$ (TSFR 将在 3.2 节详细介绍)。即若待测样本属于第  $j$  类,  $s_j^1$  的值为 1;反之则为 0。

步骤 3 利用 TSFR 对合成的新样本进行训练分类,得到任意一个测试样本属于第  $j$  类的分值  $s_j^2$ ,  $s_j^2$  的定义与  $s_j^1$  的定义类似。

步骤 4 分数融合,通过对步骤 2 和步骤 3 分别得到的权值,按照公式  $s_j = \omega_1 s_j^1 + \omega_2 s_j^2$  (其中实验中需要事先设定  $\omega_1$  和  $\omega_2$ , 满足  $\omega_1 > \omega_2$  且  $\omega_1 + \omega_2 = 1$ ) 进行加权计算,得到最后的分类结果。

#### 3.2 两步人脸识别算法的原理(TSFR)

我们定义样本类别总数为  $C$ ,  $N_i$  表示第  $i$  个类中包含的训练样本数,则训练样本总数可以表示为  $N = \sum_{i=1}^C N_i$ 。  $x_i$  表示原始样本向量,  $y_i$  代表重构样本向量,  $z$  表示测试样本向量。为了方便描述,本文只通过原始样本介绍 TSFR 算法的原理,合成样本的 TSFR 训练分类过程只需将算法中的原始样本用重构样本替换即可。

假设待测样本可以通过  $N$  个训练样本表示如下:

$$z = a_1 x_1 + a_2 x_2 + \dots + a_N x_N \quad (19)$$

其中,  $a_i$  表示系数。为了方面表述,分别用矩阵  $X$  和  $A$  来标记样本矩阵和系数矩阵,即  $X = [x_1, x_2, \dots, x_N]$ ,  $A = [a_1, a_2, \dots, a_N]^T$ , 则待测样本可以重新表示为:

$$z = XA \quad (20)$$

借助式(20)可以推出:系数矩阵  $A$  可以通过  $\bar{A} = (X^T X + \mu I)^{-1} X^T z$  计算得到,其中  $\bar{A} = [\bar{a}_1, \dots, \bar{a}_N]^T$ ,  $\mu$  是一个正的小数,  $I$  为单位矩阵。

这样,待测样本与第  $k$  类样本中心之间的距离可表示为:

$$d_k = \left\| z - \sum_{i=(k-1)n+1}^kn \bar{a}_i x_i \right\|, k=1, 2, \dots, C \quad (21)$$

其中,  $x_{bn}$  表示第  $k$  类的  $n$  个训练样本,而  $\bar{a}_{bn}$  表示与之相对应的特征系数。求解式(21)即可计算出待测样本  $z$  到所有  $C$  类样本的欧氏距离。

若  $d_{r_1} \leq d_{r_2} \leq \dots \leq d_{r_c}$ , 则与待测样本距离最小的前  $t$  个类别标签用  $r_1, r_2, \dots, r_t$  表示,那么可以确定,待测样本的所属类别一定分布在  $r_1, r_2, \dots, r_t$  中。

借助式(19)一式(21)的第一步模糊分类过程,可以实现把待测样本判定在一个较小的范围内。为了得到最终的分类结果,需要在已确定的  $t$  个类别范围内进行进一步的分类训练。

假设前  $t$  个类别内训练样本分别为  $x_1', \dots, x_{bn}'$ , 则任意一

个随机变量  $\tilde{z}$  可表示为:

$$\tilde{z} = f_1 x_1' + \dots + f_m x_m' \quad (22)$$

其中,  $f_i$  表示特征系数,同式(19),可以将  $\tilde{z}$  重新描述为:

$$\tilde{z} = X'F \quad (23)$$

其中,  $F = [f_1, \dots, f_m]^T$ ,  $X' = [x_1', \dots, x_m']$ 。  $F$  可以通过下式计算得到:

$$\bar{F} = (X'^T X' + \gamma I)^{-1} X'^T z \quad (24)$$

其中,  $\bar{F} = [\bar{f}_1, \dots, \bar{f}_m]^T$ ,  $\gamma$  表示一个正的小数,  $I$  同样表示单位矩阵。

我们用  $x_g', \dots, x_k'$  表示第  $r$  (其中  $r \in C_{r_1}, C_{r_2}, \dots, C_{r_t}$ ) 类的所有训练样本,  $\bar{f}_g, \dots, \bar{f}_k$  表示与其相对应的系数,则测试样本到第  $r$  类的欧氏距离可表示为:

$$u_r = \left\| \tilde{z} - \sum_{i=g}^k \bar{f}_i x_i' \right\| \quad (25)$$

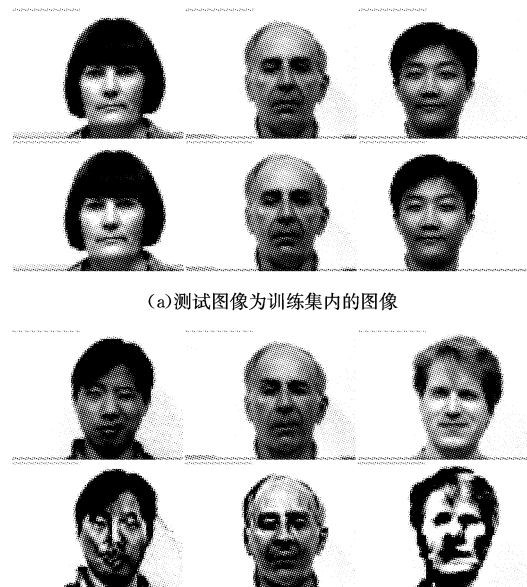
若  $k = \arg \min u_r$ , 则  $k$  即为与测试样本最相似类别,即最终的分类结果。

### 4 实验结果及分析

#### (1) 利用因素分解模型合成的样本图像

为了验证双线性因素分解模型中的“人的身份”和“人脸表情”之间的相互转移功能,本文首先在 Yaleface 人脸库上进行了合成实验。

图 1 和图 2 的第一行表示原始图像,第二行表示重构图像。观察图 1 和图 2 可以看到,当测试图像为训练集内的图像时,无论是新人脸的旧表情合成还是新表情的旧人脸合成,其合成结果都十分令人满意,总体而言,后者优于前者。但当测试图像为训练集以外的图像时,可以发现重构的新样本图像中有模糊、斑驳的情况,这主要是因为双线性因素分解模型容易产生局部极小值。因此双线性模型对于训练集中不存在的测试图像还有待改进。



(b) 测试图像为训练集外的图像

图 1 训练集内人脸在新表情下的合成实验



图2 新人脸在已有表情下的合成实验

## (2)在不同人脸库上的识别结果

为了避免因素分解模型算法中局部最小值对人脸合成效果的影响,确保识别阶段两步人脸识别算法的准确性和精度,实验中将数据库的所有训练样本统一训练。为了降低计算的复杂度,实验过程将在 FERET 和 UMIST 两个数据库中进行,人脸图像分别裁剪为  $40 \times 40$  以及  $56 \times 40$  大小。在不同人脸库上的识别结果如表 1、表 2 所列。

表1 FERET人脸库上的识别结果/%

原始训练样本数		1	2	3
$\omega_1=0.85$	Xu 的方法	51.58	37.40	44.38
	本文方法	<b>50.96</b>	<b>36.20</b>	<b>43.38</b>
$\omega_1=0.75$	Xu 的方法	50.75	36.40	42.88
	本文方法	<b>50.75</b>	<b>35.60</b>	<b>41.75</b>
$\omega_1=0.65$	Xu 的方法	49.50	35.30	40.87
	本文方法	<b>48.65</b>	<b>34.40</b>	<b>40.17</b>

表2 UMIST人脸库上的识别结果/%

原始训练样本数		1	2	3
$\omega_1=0.85$	Xu 的方法	54.22	50.93	50.75
	本文方法	<b>53.33</b>	<b>50.13</b>	<b>50.75</b>
$\omega_1=0.75$	Xu 的方法	54.22	51.73	50.45
	本文方法	<b>53.02</b>	<b>50.90</b>	<b>49.55</b>
$\omega_1=0.65$	Xu 的方法	55.18	52	51.04
	本文方法	<b>54.47</b>	<b>51.15</b>	<b>50.44</b>

表 1 和表 2 分别列出了本文改进算法和 Xu 的方法的比较结果,其中  $\omega_i$  表示所选取的权值。比较上述实验结果可以发现:

本文算法在人脸重构过程中,利用因素分解模型将“人的

身份”和“人脸表情”从人脸图像中分离出来,减少了其他外部因素对人脸图像的影响,提高了人脸合成的效果。然后进一步采用两步人脸识别算法,充分利用了分数层次融合的优势,提高了分类的准确性。因此,通过在不同人脸库上的实验结果可以发现,与基于对称性 Xu 的算法相比,本文算法在不同人脸库上的识别错误率均有不同程度的下降,识别效果更好。

**结束语** 为了进一步提高人脸识别算法在表情等因素干扰下的识别效果,本文在 Xu 等人提出的结合人脸对称性的两步人脸识别算法的基础上,提出了一种新的研究思路。首先,本文通过双线性因素分解模型把“个人身份”和“人脸表情”从人脸图像中分离出来,减少其他外部因素对人脸图像的影响,合成新的人脸图像。通过将原始图像和合成的新图像同时当作训练样本,结合两步人脸识别算法,充分利用了分数层次融合的优势。实验结果证明,改进的算法不仅在人脸合成方面更有优势,而且更有助于提高分类的准确性。但由于双线性因素分解模型在人脸合成中存在局部极小解,为了确保识别效果,需要对所有样本提取“个人身份”和“人脸表情”因素,这样容易增加计算复杂度。为了缓解这一问题,后续将对双线性因素分解模型做进一步的改进。

## 参考文献

- [1] SEITZ S M, DYER C R. View morphing[C]//Proceedings of Computer Graphics. 1996:21-30.
- [2] LANITIS A, TAYLOR C J, COOTES T F. Automatic interpretation and coding of face images using flexible models[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1997, 19(7):743-756.
- [3] COOPER D H, COOTES T F, TAYLOR C J, et al. Active shape models-their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1):38-59.
- [4] TENENBAUM J B, FREEMAN W T. Separating style and content with bilinear models[J]. Neural Computation, 2000, 12(6):1247-1283.
- [5] XU Y, ZHU X J, LI Z M, et al. Using the original and ‘symmetrical face’ training samples to perform representation based two-step face recognition[J]. Pattern Recognition, 2013, 46(4):1151-1158.
- [6] 徐欢. 基于双线性模型的人脸表情识别技术[J]. 计算机与现代化, 2014, 227(7):89-93.
- [7] XU Y, ZHANG Z, LU G, et al. Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification [J]. Pattern Recognition, 2016, 54(C):68-82.