

基于 Lucas-Kanade 算法的最大 Gabor 相似度 大姿态人脸识别

程超^{1,2}, 达飞鹏^{1,2}, 王辰星^{1,2*}, 姜昌金^{1,2}

¹东南大学自动化学院, 江苏 南京 210096;

²东南大学复杂工程系统测量与控制教育部重点实验室, 江苏 南京 210096

摘要 在人脸识别科学研究和实际应用领域中,大角度姿态是影响人脸识别结果的主要因素之一,成为限制人脸识别技术进步的难点,而姿态的校正归一化是解决该问题的常用手段。首先通过加权的 LK(Lucas-Kanade)算法得到侧脸块和对应正脸块的仿射变换参数,基于最大 Gabor 相似度寻找校正人脸姿态的最优参数。然后,以每一人脸块最优参数得到的平均 Gabor 相似度作为这一块人脸的识别权重,可以增加大姿态人脸识别的精度和稳健性。在 FERET 人脸数据库中进行了实验,当水平偏转角度为 45°时,准确率达到 97.3%,证明本文提出的以最大 Gabor 相似度作为加权 LK 算法参数提取的依据是有效的,得到的最优参数具有较好的光照无关性,而将平均 Gabor 相似度作为识别权重,有助于使算法的应用更加稳健和有效。

关键词 图像处理; 人脸识别; 最大 Gabor 相似度; 加权 LK 算法; 人脸分块; Gabor 特征

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/AOS201939.0715005

Pose Invariant Face Recognition Using Maximum Gabor Similarity Based on Lucas-Kanade Algorithm

Cheng Chao^{1,2}, Da Feipeng^{1,2}, Wang Chenxing^{1,2*}, Jiang Changjin^{1,2}

¹*School of Automation, Southeast University, Nanjing, Jiangsu 210096, China;*

²*Key Laboratory of Measurement and Control of Complex Systems of Engineering, Ministry of Education, Southeast University, Nanjing, Jiangsu 210096, China*

Abstract In the field of face recognition, pose variation is one of the significant challenges that affects the recognition performance and has been one of the major obstacles hindering the improvement of the face recognition technology. In this study, affine transformation parameters of side face and full-frontal face patches are obtained by applying the weighted Lucas-Kanade (LK) algorithm. We further propose that an optimal parameter for correcting face pose can be obtained based on the maximum Gabor similarity. Furthermore, the average Gabor similarity acquired from the optimal parameter of each face patch can be considered to be the face recognition weight, improving the recognition rate and enhancing the robustness of the pose invariant face recognition. Finally, the experimental results obtained based on the FERET face database denote that the recognition rate for the image with a pose of 45° can reach up to 97.3%, indicating that the usage of the maximum Gabor similarity as a basis for parameter extraction of the weighted LK algorithm is valid. This method can also handle illumination variations. Considering the average Gabor similarity as the recognition weight will ensure the robust and effective application of this algorithm.

Key words image processing; face recognition; maximum Gabor similarity; weighted Lucas-Kanade algorithm; face patches; Gabor feature

OCIS codes 100.3008; 100.3010; 100.5010

1 引 言

人脸识别技术是当代人工智能、模式识别和计

算机视觉领域中热门的研究课题之一。在实际生活和科学应用中,通过人脸识别检测,得出大角度姿态是影响人脸识别结果的主要因素结论。当输入的

收稿日期: 2018-12-12; 修回日期: 2019-01-29; 录用日期: 2019-04-01

基金项目: 国家自然科学基金(61828501,61462072,61628304)

* E-mail: cxwang@seu.edu.cn

人脸图像是一张水平旋转角度较大的侧脸图像时,即使是同一个人的两幅图像,不同姿态反映的图像也会有很大的差异,这种差异可能会比相同姿态而不同身份的人的差异还要大^[1]。这使得很多常规人脸识别算法的性能大幅下降,进而导致识别率也明显降低。

针对大角度人脸姿态的主流研究方法主要集中在三维人脸模板和二维平面技术。在三维人脸模板方面,Tang等^[2]利用二维人脸重构三维虚拟人脸,然后将三维人脸进行二维投影,构建二维多姿态人脸库,将多姿态人脸识别问题转化为某一角度下两幅相同姿态照片的对比问题^[3]。这种采用二维人脸重构三维虚拟人脸的思路,在一定程度上改善了姿态变化对人脸识别的影响,但三维人脸重建运算较复杂,计算量较大,难以实现实时性的要求。

在二维平面技术方面,Chai等^[4]提出了局部线性回归算法(LLR),该算法假设人脸侧脸与正脸呈局部线性关系,用侧脸图像与其对应的正脸图像进行训练,得到侧脸到正脸之间的线性映射关系,进而由侧脸估计出其对应的正脸图像。为解决侧脸到正脸的非线性问题,LLR算法把人脸分成了很多个小区域,并对每个小区域分别进行线性映射,这必然会导致非线性信息的丢失。Ashraf等^[5]基于LK(Lucas-Kanade)算法提出了人脸区域的流堆栈(StackFlow)算法,该算法首先将人脸划分为若干块,用仿射变换进行侧脸到正脸的校正;训练过程中使用LK算法,经仿射变化后的侧脸与正脸的误差最小;作者提出采用流堆栈算法来获得每块人脸的全局最优,即某一块的全局最优应该是使所有训练图像的对应块和正面人脸的对应块差值和最小的仿射变换参数,但是这种方法是根据像素寻找全局最优参数,所以对人脸区域的划分准确度要求较高,同时由于姿态、光照和表情等因素不完全相同,合成的正面人脸比较模糊,导致识别率较低。类似地,Ho等^[6]提出了一种通过马尔可夫随机场(Markov Random Fields, MRF)寻找仿射变换参数的方法,该方法的目标函数与流堆栈算法相同,但是其每块人脸全局最优参数的确定是通过马尔可夫随机场实现的。通过构建代价函数,实现代价函数的最小化来选择最优参数,并运用BP(Belief Propagation)算法求出最优参数。但是在确定每一块最优参数时,运行BP算法非常耗时,同时确定最优参数是在某一张人脸面上运行的,因此不可避免地会带上该人脸的特性信息。Li等^[7]提出了基于神经网络的深度

多视角特征学习(MVRL)方法,通过不同的深度学习模型来得到不同特征,进而融合为一个整体特征作为此人脸特征的代表。Yim等^[8]提出多任务的深度神经网络(CPF)方法来实现人脸校正,主任务用于从输入人脸图像构建任意姿态的人脸图像,副任务用于从得到的人脸图像再反过来构造输入图像,以实现整个神经网络的反馈。Peng等^[9]提出了基于深度学习学习的姿态校正(Multi-Source Multi-Task, MSMT)方法,首先通过正面的二维人脸生成一个三维模板,然后投影生成不同姿态的二维人脸,再在训练识别模型时将身份信息与非身份信息(如姿态和特征点)放在一起训练,得到一个包含姿态、特征点和身份的多特征,最后把侧脸的姿态特征和正面人脸的身份特征拼接在一起重新构建,以保证不同姿态的人脸得到的特征是相似的。但是基于神经网络的算法需要的训练集大,同时需要手工标定特征点,在训练阶段比较耗时。

本文主要研究人脸姿态的二维平面校正方法,采用基于加权的LK算法得到每块人脸的最优仿射变换参数,提出每一块的最优仿射变换参数应使所有训练集中这一块校正的人脸块与对应正脸块Gabor^[10]相似度最大,并以每一人脸块最优参数得到的平均Gabor相似度作为这一块人脸的权重进行识别。最后,在FERET人脸数据库中进行了每块最优参数的评估和侧脸姿态的校正实验,同时进行了识别率实验,证明了所提方法的有效性。

2 基于最大 Gabor 相似度的人脸姿态校正

2.1 用加权 LK 算法求每块人脸的仿射变换参数

LK算法^[11]是一种有效的、实现图片对齐的算法。但是在侧脸角度情况下,人脸各个区域变化的形态不同,如果对整个人脸都运用LK算法,则各个区域的对齐效果较差,所以首先要对某一个特定角度的侧面和对应的正面人脸进行分块。通过Gabor滤波器的尺度和方向对图像进行滤波,将其结果作为纹理特征来选择全局最优参数。该过程基于提取到的人脸局部空间信息而不是每个点的像素差值信息,因此用于校正的侧脸和正脸分块并不需要精确的点对点,而是大致的局部区域对应即可。侧面和正面人脸的分块示意图如图1所示,首先将正脸划分成 N 块矩形区域,然后根据正脸区域的划分,在侧脸上标定得到大致的对应区域。

用 I 表示侧脸的图像,用 T 表示对应的正脸图

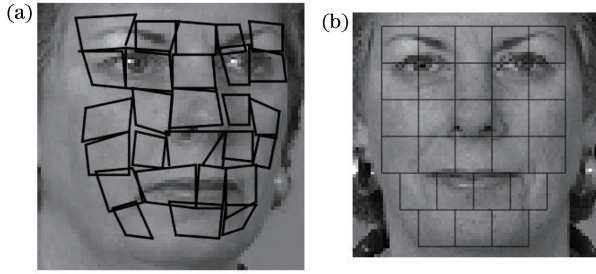


图 1 (a)侧面和(b)正面人脸的分块示意图

Fig. 1 Diagrams of side face and full-frontal face patches

像,下标 r 表示块的编号。采用仿射变换来实现侧脸到正脸各部分的对齐。设仿射变换矩阵为

$$\mathbf{A} = \begin{bmatrix} 1 + p_1 & p_3 & p_5 \\ p_2 & 1 + p_4 & p_6 \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$

仿射变换参数 $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5, p_6)^T$, 用 W 表示仿射变换。则一个坐标 x 仿射变换后得到 x' , 则 $x' = W(x, \mathbf{p})$ 。目标函数可表示为

$$E = \sum_x \{I_r[W(x, \mathbf{p})] - T_r(x)\}^2. \quad (2)$$

使目标函数最小时对应的 \mathbf{p} , 即可认为是最优的姿态校正参数。(2)式的求解可采用加权的 LK 算法^[12]来实现。加权的 LK 算法引入二维 Gabor 变换^[10]来计算权值,以适应光照变化。二维 Gabor 变换的定义为

$$g_{\mu, \nu}(\mathbf{z}) = \frac{\|\mathbf{k}_{\mu, \nu}\|^2}{\sigma^2} \exp\left(-\frac{\|\mathbf{k}_{\mu, \nu}\|^2 \|\mathbf{z}\|^2}{2\sigma^2}\right) \cdot \left[\exp(i\mathbf{k}_{\mu, \nu} \cdot \mathbf{z}) - \exp\left(-\frac{\sigma^2}{2}\right) \right], \quad (3)$$

式中: $\mathbf{z} = (x, y)^T$ 为空间位置坐标; 参数 $\sigma = 2\pi$; $\mathbf{k}_{\mu, \nu}$ 用来确定 Gabor 内核的方向和尺度, μ, ν 表示 Gabor 变换方向和尺度变化, 每一组 μ, ν 参数都会进行一次 Gabor 变换, 从而得到相应的 Gabor 系数, 经过 $\mu \times \nu$ 个 Gabor 变换后, 可在不同粗细粒度上得到图像的有效特征。

假设有 M 个 Gabor 变换, 用 g_i 表示第 i 个 Gabor 变换, 则目标函数为

$$E = \left\| \{g_1 * I_r[W(x, \mathbf{p})], \dots, g_M * I_r[W(x, \mathbf{p})]\} - [g_1 * T_r(x), \dots, g_M * T_r(x)] \right\|^2, \quad (4)$$

式中: $*$ 代表卷积。将(4)式变换到傅里叶频域, 即

$$E = \{I_r[W(x, \mathbf{p})] - T_r(x)\} \mathbf{F}^T \mathbf{S} \mathbf{F} \cdot \{I_r[W(x, \mathbf{p})] - T_r(x)\}, \quad (5)$$

式中: $\mathbf{S} = \sum_{i=1}^M [\text{diag}(\hat{g}_i)]^T \text{diag}(\hat{g}_i)$, 其中 \hat{g}_i 为 g_i 对应的傅里叶变换; \mathbf{F} 为包含傅里叶基向量的矩阵; $\mathbf{Q} = \mathbf{F}^T \mathbf{S} \mathbf{F}$ 为加权值。则最后的误差函数写为

$$E = \sum_x \{I_r[W(x, \mathbf{p})] - T_r(x)\} \cdot \mathbf{Q} \{I_r[W(x, \mathbf{p})] - T_r(x)\}. \quad (6)$$

(6)式中将 \mathbf{p} 换为 $\mathbf{p} + \Delta\mathbf{p}$, $\Delta\mathbf{p}$ 表示迭代步长, 通过 $\Delta\mathbf{p}$ 来求取最终的参数 \mathbf{p} , 而不是直接求取参数 \mathbf{p} 。然后进行一级泰勒展开, 并令其关于 $\Delta\mathbf{p}$ 的偏导数等于 0, 得到

$$\Delta\mathbf{p} = \mathbf{H}^{-1} \sum_x \left\{ \frac{\partial I_r[W(x, \mathbf{p})]}{\partial \mathbf{p}} \right\}^T \cdot \mathbf{Q} \{I_r[W(x, \mathbf{p})] - T_r(x)\}, \quad (7)$$

$$\text{式中: } \mathbf{H} = \sum_x \left\{ \frac{\partial I_r[W(x, \mathbf{p})]}{\partial \mathbf{p}} \right\}^T \mathbf{Q} \frac{\partial I_r[W(x, \mathbf{p})]}{\partial \mathbf{p}}.$$

每一次迭代 $\Delta\mathbf{p}$, \mathbf{p} 参数从零向量开始迭代 $\mathbf{p} = \mathbf{p} + \Delta\mathbf{p}$, 直到达到收敛, 最后将收敛值看作这一块人脸的仿射变换参数 \mathbf{p} 。

2.2 基于最大 Gabor 相似度寻找每块全局最优的仿射变换参数

如果基于像素值寻找全局最优参数, 由于人脸块划分不标准、光照条件和姿态误差等因素的影响, 很难保证校正参数的准确性, 最终造成人脸合成效果不理想。因此, 利用 Gabor 特征提高光照和姿态的稳健性, 通过使 Gabor 特征向量余弦相似度最大, 来寻找一个全局最优参数。

假设共有 B 张侧脸, 第 k 个图片得到的第 i 块参数为 $\mathbf{p}_i^{(k)} \{k=1, 2, \dots, B\}$, 则第 i 块仿射变换参数候选集 $P_i = \{\mathbf{p}_i^{(k)}, k=1, 2, \dots, B\}$ 。选择第 i 块最优参数的原则如下:

求库中第 k 个侧面人脸中的第 i 块区域在参数 \mathbf{p}_i 下的 Gabor 特征, 得到一个特征向量 $\mathbf{b}_{k,i}(\mathbf{p}_i)$:

$$\mathbf{b}_{k,i}(\mathbf{p}_i) = \{g_1 * I_{k,i}[W(x, \mathbf{p}_i)]; \dots; g_M * I_{k,i}[W(x, \mathbf{p}_i)]\}. \quad (8)$$

求库中对应第 k 个正面人脸图片第 i 块的 Gabor 特征 $\mathbf{a}_{k,i}$:

$$\mathbf{a}_{k,i} = [g_1 * T_{k,i}(x); \dots; g_M * T_{k,i}(x)]. \quad (9)$$

对于第 i 块区域的每一个参数 \mathbf{p}_i , 计算其 Gabor 余弦相似度:

$$f_{\text{gabor-s}}(\mathbf{p}_i) = \sum_{k=1}^B \frac{\mathbf{a}_{k,i} \cdot \mathbf{b}_{k,i}(\mathbf{p}_i)}{\|\mathbf{a}_{k,i}\| \|\mathbf{b}_{k,i}(\mathbf{p}_i)\|}, \quad (10)$$

则第 i 块区域全局最优参数 $\hat{\mathbf{p}}_i = \max_{\mathbf{p}_i \in P_i} f_{\text{gabor-s}}(\mathbf{p}_i)$ 。

基于上述提出的 Gabor 相似度求取的全局最优姿态校正参数对光照变化有良好的适应性。此外, 全局最优参数的选择是基于 Gabor 变换提取的局部空间特征, 而不是利用每个点的像素信息, 所以人脸校正过程中侧脸的划分不需要很精确, 只需进

行粗略的侧脸区域划分即可通过全局最优参数得到清晰的正面人脸。

2.3 基于加权的分块 Gabor 向量人脸识别

考虑到姿态校正过程中采用的是基于人脸分区域的 Gabor 特征校正方法,所以加权的局部 Gabor 二值模式直方图序列(LGBPHS)算法^[13]更适用于上述姿态校正后进一步的人脸识别。由于本方法基于局部分块,而每个局部块在识别整个人脸中所起的作用大小不一,且根据最优校正参数拟合的局部块也有相对的优劣之分,因此提出为每块人脸设定一个权重值,以提高最终的识别率。

按照上述人脸分块,将不同尺度、不同方向的 Gabor 滤波器分别与侧脸校正的图片块和库中正脸对应块进行卷积,取其幅值部分组成 Gabor 幅值图谱(GMM),对每个人脸区域的 GMM 应用局部二值模式(LBP)算法得到局部 Gabor 二值模式图谱并计算直方图,将所有区域直方图串接为一个直方图序列,并将其作为这一块人脸图像的描述符。

假设两个直方图是 \mathbf{H}^1 和 \mathbf{H}^2 ,则两个直方图的相似度可表示为

$$\Psi(\mathbf{H}^1, \mathbf{H}^2) = \sum_{i=1}^L \min(h_i^1, h_i^2), \quad (11)$$

式中: h_i^1 表示第一个直方图中的第 i 个条目; h_i^2 表示第二个直方图中的第 i 个条目; L 为直方图中的条目数,衡量的标准为对应条目中共有的部分。

采用加权的局部 Gabor 二值模式直方图序列算法进行人脸识别,即对上述每块人脸添加权重系数。假设灰度区间为 $[0, L-1]$, $\mathbf{H}_{\mu,\nu,i}$ 表示第 i 块人脸在 Gabor 变换 $g_{\mu,\nu}$ 下得到的局部二值模式直方图,即

$$\mathbf{H}_{\mu,\nu,i} = (h_{\mu,\nu,i,0}, h_{\mu,\nu,i,1}, \dots, h_{\mu,\nu,i,L-1}), \quad (12)$$

式中: $h_{\mu,\nu,i,r'} = \sum_{z \in I} I\{\text{LBP}[g_{\mu,\nu}(z)] = r'\}$;LBP 表示图像进行局部二值模式算法; r' 表示灰度等级; I 可表示为

$$I\{A\} = \begin{cases} 1, & A \text{ is true} \\ 0, & A \text{ is false} \end{cases} \quad (13)$$

最终每个人脸的直方图可表示为

$$\mathbf{H} = (\mathbf{H}_{0,0,1}, \dots, \mathbf{H}_{0,0,N}, \mathbf{H}_{0,1,1}, \dots, \mathbf{H}_{0,1,N}, \dots, \mathbf{H}_{7,4,N}), \quad (14)$$

式中: N 为人脸区域划分的数量。则衡量两个人脸相似度的加权 LGBPHS 为

$$S(\mathbf{H}_1, \mathbf{H}_T) = \sum_{\mu} \sum_{\nu} \sum_i \omega_i \Psi(\mathbf{H}_{\mu,\nu,i}^1, \mathbf{H}_{\mu,\nu,i}^T), \quad (15)$$

式中: \mathbf{H}_1 表示侧脸的直方图; \mathbf{H}_T 表示正脸的直方图; ω_i 为第 i 块人脸的权重值。

设定第 i 块的权重值为通过此人脸块的最优参数 \hat{p}_i 得到的平均 Gabor 相似度,即

$$\omega_i = \frac{\sum_{k=1}^B \frac{\mathbf{a}_{k,i} \cdot \mathbf{b}_{k,i}(\hat{p}_i)}{\|\mathbf{a}_{k,i}\| \|\mathbf{b}_{k,i}(\hat{p}_i)\|}}{B}. \quad (16)$$

最后,找到使 $S(\mathbf{H}_1, \mathbf{H}_T)$ 最大的 \mathbf{H}_T 即为此侧脸图片的身份。

3 实验结果与分析

本实验采用 FERET 人脸数据库^[14]进行实验。该数据库包含约 200 人,共 14051 张灰度人脸图像,同一个人包含不同的表情、光照和姿态图片。在训练过程中,每一种角度都采用前 60 个人(身份 ID 从 1 到 60)的 $B=60$ 张侧脸图片和对应的正脸图片来得到每一块人脸区域的全局最优参数,并用于后续的实验。将人脸图像裁剪成 80×80 的灰度图像,并按照人脸形状特征(图 1)将正脸划分为 $N=27$ 块,每一块的大小为 12×12 (尺寸过大会降低校正效果,尺寸过小会增大计算量),并将正脸分块对应到侧脸的相应位置,使每个侧脸也是 27 块。划分的人脸块不需要正脸与侧脸严格准确地一一对应,仅需大致对应即可。所有正脸与侧脸的局部区域划分及位置都保持一致。

提取 Gabor 特征时,Gabor 内核定义为 $k_{\mu,\nu} = k_{\nu} \exp(i\phi_{\mu})$,其中 $k_{\nu} = k_{\max}/f^{\nu}$ 为采样尺度, $\phi_{\mu} = \pi\mu/8$ 为采样方向, k_{\max} 为最大频率, f 为频域内最大间隔因子。为了获得较好的小波特征和辨识效果,实验中令参数 $k_{\max} = \pi/2$, $f = \sqrt{2}^{[10]}$,取 5 个尺度和 8 个方向,分别标为 $\nu = (0, 1, \dots, 4)$ 和 $\mu = (0, 1, \dots, 7)$ 。因此,每个局部区域可获取 $M=40$ 个 Gabor 变换系数来提取 Gabor 特征,这样可以充分利用图像的方向信息和尺度,在不同粗细粒度上得到图像的特征。训练最终得到每个对应局部区域块的最优 Gabor 相似度参数。

根据每一块得到的最优参数进行了三部分实验,分别是关于最优参数的评估、侧脸的姿态校正和侧脸识别实验。本实验使用 MATLAB 语言,在 Intel(R) i5-3230M 处理器上运行。

3.1 评估得到的每块最优参数

本实验测试了同一角度下(40°)60 个人(身份 ID 从 61 到 120)的侧脸区域经过最优参数校正后的

侧脸块与正脸块 Gabor 余弦相似度,其中实验的第 1,7,9,15 块(根据图 1 的划分,从上往下数,从左往右数)结果如图 2 所示。横坐标表示测试的 60 个侧脸图像,纵坐标表示每个图像对应区域的 Gabor 相似度。每一个人脸区域中的水平虚线即为用这一块

最优参数求得的平均 Gabor 余弦相似度,这个值也将作为每一块人脸的识别权重。从图中可以看到,有的人脸块拟合得好,而有的人脸块拟合得一般,验证了每块人脸需要加权值的必要性。

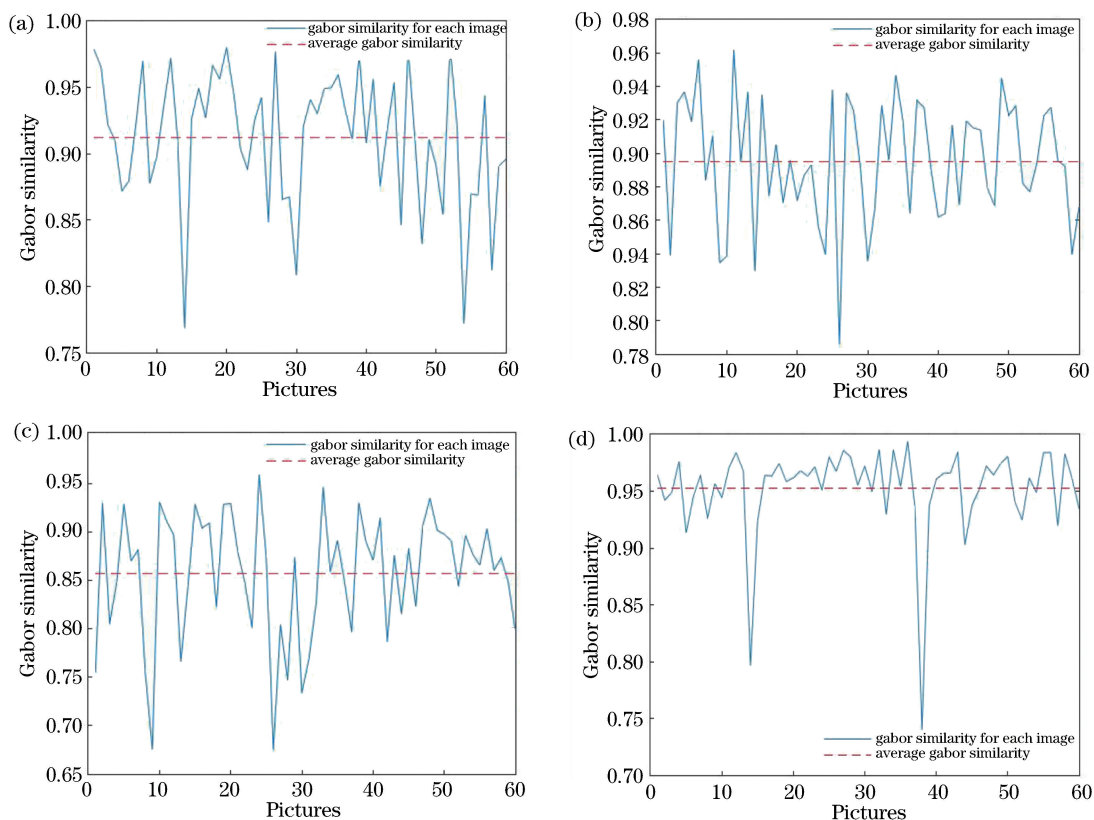


图 2 基于最优参数的 Gabor 相似度表现。(a)第 1 块;(b)第 7 块;(c)第 9 块;(d)第 15 块

Fig. 2 Performance of Gabor similarity based on optimal parameter. (a) 1st patch; (b) 7th patch; (c) 9th patch; (d) 15th patch

3.2 通过每块最优参数校正人脸

本部分实验不仅使用 FERET 人脸数据库,还选取了 CMU_PIE 人脸数据库^[15]的图片进行测试。CMU_PIE 数据库规模不大,仅有 68 个人,共 41368 张人脸图片,但含有更多丰富光照变化的图片。实验对人脸库中标定为 $\pm 40^\circ$ 的侧脸进行校正,此外还实际拍摄了侧脸图片进行实验。效果如图 3 所示,第 1~3 行为 FERET 库中图片,第 4、5 行图片来自 CMU_PIE 库,最后一行为实拍图片。

从上述结果可以看出,在光照不均匀的条件下,根据 Gabor 特征得到的最优变换参数依然适用,所以得到的全局最优参数对光照有一定的稳健性。另外,虽然数据库中标定的角度有误差(前后相差 10° 左右),但是恢复出来的正面人脸基本是清晰的。从这一点可以看出,全局最优参数对姿态也具有一定

的稳健性。同样情况下(人脸分块也是粗略划分的),如果是根据像素差最小得到的最优参数对侧脸进行校正,得到的效果图非常模糊,这是因为依据像素差值最小寻得的最优参数对人脸分块的准确性有很大的依赖。

3.3 侧脸的识别率

侧脸识别实验在 FERET 人脸数据库中进行,测试了身份 ID 从 61 到 190 共 130 个人的对应角度的侧脸图片,即每个角度测试数据为 130 张图片,并与其他方法进行了对比。实验对比方法主要选择与本文分块提取相似度思想相似的 3 种经典方法,即文献[4-6]中的方法,同时对比了近年来较热门的基于深度学习的方法,即文献[7-9]中的方法。实验结果如表 1 所示。

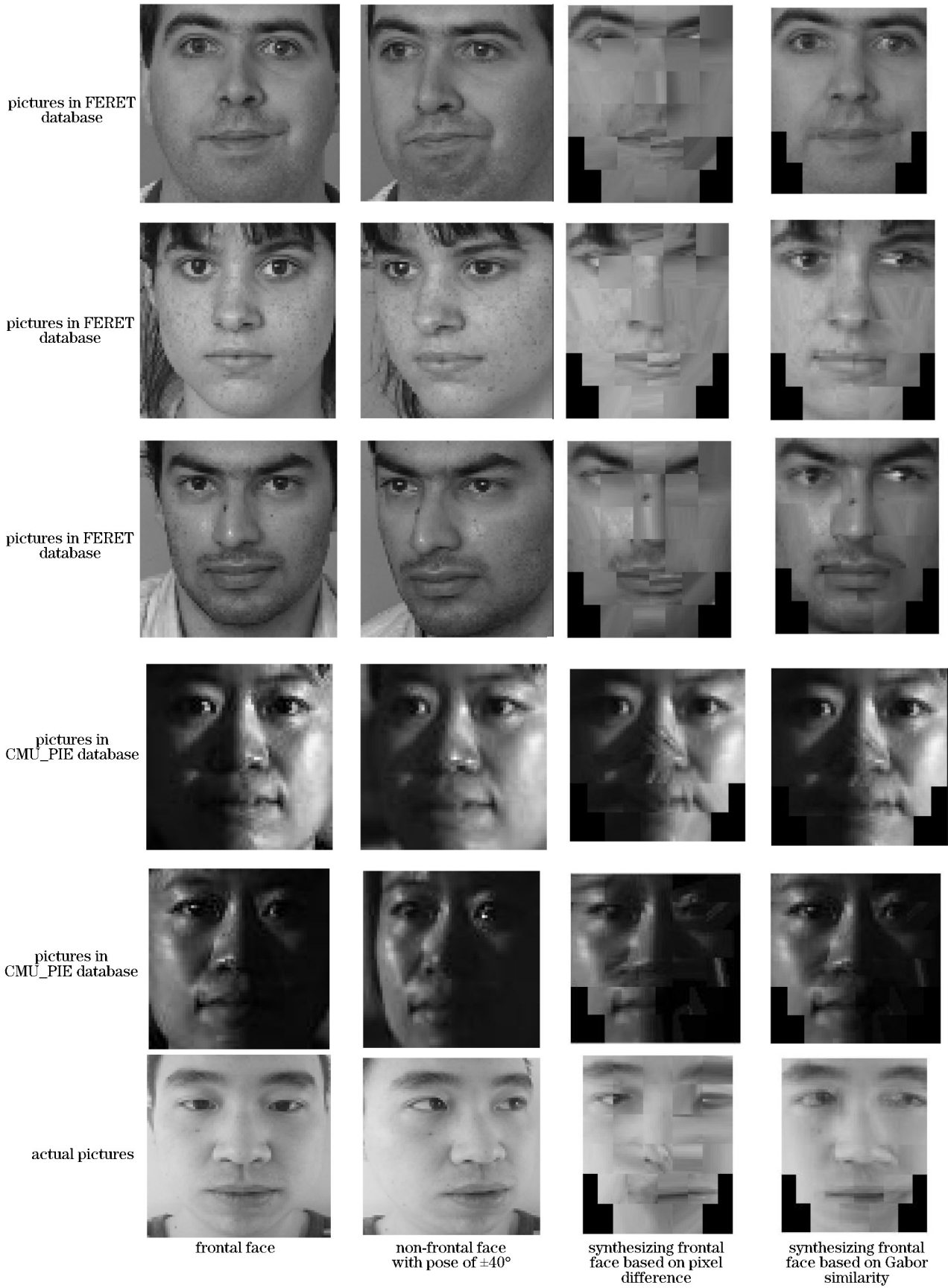


图 3 基于 Gabor 最优参数和像素差最优参数校正后的人脸对比

Fig. 3 Synthesizing frontal face comparison based on Gabor optimal parameters and pixel difference optimal parameters

表 1 若干方法在 FERET 人脸数据库中的识别率对比
Table 1 Comparison of recognition rates of different approaches on FERET database

Method	Recognition rate / %			
	-45°	-30°	30°	45°
LLR ^[4]	55.0	89.5	77.0	53.0
StackFlow ^[5]	70.0	89.0	82.0	62.0
MRF ^[6]	91.0	97.3	96.5	91.5
MVRL ^[7]	96.2	99.2	100.0	96.9
CPF ^[8]	98.4	100.0	99.2	97.7
MSMT ^[9]	99.6	100.0	100.0	99.2
Our approach	96.9	98.4	99.2	97.7

从表 1 可以看出,传统的 LLR 方法识别率最低,这是由于传统方法只提取了人脸线性信息而忽略了非线性信息,所以得到的校正图像比较模糊,从而严重影响了识别率。StackFlow 和 MRF 算法,尤其是后者,在大姿态人脸识别率上已经有了较大提升,但这两种方法均是根据像素点的灰度值寻找最优参数,而这种方式非常依赖人脸区域划分和映射的准确性,并且灰度值受光照的影响。基于最大 Gabor 相似度寻找最优参数则不受上述因素的影响和限制,因此,识别率也得到了进一步的提高,所提方法的操作更加简便,也更适用于实际应用。

对于基于深度学习的几种方法,整体识别率都较高,但这些方法对硬件设备和训练数据规模的要求均较高。对于 CPF 方法(训练集用了 14000 张图片),虽然整个神经网络有反馈,可以得到任意姿态的人脸图像,但是图片的校正过程需要对图像另外添加一些编码信息,而且提取的各个角度特征的区分度并不太好。MVRL 方法(训练集用了 30000 张图片)通过不同深度学习模型提取的特征有些冗余,并且只利用了纹理特征而没有对形状特征进行区分,因而需要的训练图像数量要求规模更大。而最新的基于深度学习神经网络学习的 MSMT 方法,训练集用了 106402 张图片,计算量极大。相比而言,所提方法操作简单,对硬件设备要求不高,在数据规模小的情况下具有高效率和高精度的优势。

4 结 论

大姿态已成为限制人脸识别率提高的主要因素。采用基于加权的 LK 人脸分块算法分析可知,每一人脸块的全局最优参数是使得 Gabor 特征的

余弦相似度最大的参数,同时将根据最优参数得到的平均 Gabor 相似度作为姿态识别中的权重可以提高识别率。所提方法的优点在于数据规模小的情况下具有独特优势,以少量训练样本和较短训练时间即可得到理想的结果,并且无需进行任何标定点的检测,对人脸的分块进行粗略划分即可。下一步的研究应将深度学习学习与本文方法所提取的 Gabor 参数相结合,以便更加高效地得到清晰的正脸图片和更高的识别精度。

参 考 文 献

- [1] Zhang X Z, Gao Y S. Face recognition across pose: a review [J]. *Pattern Recognition*, 2009, 42 (11): 2876-2896.
- [2] Tang H, Yin B, Sun Y, *et al.* Pose-invariant face recognition based on a single view [J]. *Journal of Information and Computer Science*, 2010, 7 (12): 2369-2379.
- [3] Tong Y, Wei Y M, Shen Y H. Supervised sparsity preserving projection based on global constraint [J]. *Acta Optica Sinica*, 2018, 38(9): 0910001.
童莹, 魏以民, 沈越泓. 基于全局约束的监督稀疏保持投影降维方法研究 [J]. *光学学报*, 2018, 38(9): 0910001.
- [4] Chai X J, Shan S G, Chen X L, *et al.* Local linear regression (LLR) for pose invariant face recognition [C] // 7th International Conference on Automatic Face and Gesture Recognition (FGR06), April 10-12, 2006, Southampton, UK. New York: IEEE, 2006: 631-636.
- [5] Ashraf A B, Lucey S, Chen T. Learning patch correspondences for improved viewpoint invariant face recognition [C] // 2008 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2008, Anchorage, AK, USA. New York: IEEE, 2008: 4587754.
- [6] Ho H T, Chellappa R. Pose-invariant face recognition using Markov random fields [J]. *IEEE Transactions on Image Processing*, 2013, 22 (4): 1573-1584.
- [7] Li J, Zhao J, Zhao F, *et al.* Robust face recognition with deep multi-view representation learning [C] // Proceedings of the 24th ACM International Conference on Multimedia, October 15-19, 2016, Amsterdam, Netherlands. New York: ACM, 2016: 1068-1072.
- [8] Yim J, Jung H, Yoo B, *et al.* Rotating your face using multi-task deep neural network [C] // 2015 IEEE Conference on Computer Vision and Pattern

- Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 676-684.
- [9] Peng X, Yu X, Sohn K, *et al.* Reconstruction-based disentanglement for pose-invariant face recognition [C] // 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE, 2017: 1632-1641.
- [10] Liu C J, Wechsler H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition [J]. IEEE Transactions on Image Processing, 2002, 11(4): 467-476.
- [11] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision [C] // Proceedings of 7th International Joint Conference on Artificial Intelligence, August 24-28, 1981, Vancouver, British Columbia. San Francisco: Morgan Kaufmann Publishers Inc., 1981: 674-679.
- [12] Ashraf A B, Lucey S, Chen T. Fast image alignment in the Fourier domain [C] // 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June 13-18, 2010, San Francisco, CA, USA. New York: IEEE, 2010: 2480-2487.
- [13] Zhang W C, Shan S G, Gao W, *et al.* Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition [C] // Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, October 17-21, 2005, Beijing, China. New York: IEEE, 2005: 786-791.
- [14] Phillips P J, Moon H, Rizvi S A, *et al.* The FERET evaluation methodology for face-recognition algorithms [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1090-1104.
- [15] Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression database [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12): 1615-1618.