



研究与开发

基于吸引模式的局部二阶梯度轮廓人脸识别算法

叶学义, 钱丁炜, 应娜, 王涛

(杭州电子科技大学模式识别与信息安全实验室, 浙江 杭州 310018)

摘要: 针对现有 LDP 类算法在特征提取有效性和特征编码稳定性之间难以平衡的问题, 提出一种吸引局部二阶梯度轮廓 (ALSGC) 模式, 以提升人脸识别性能。首先, 利用 Kirsch 算子计算人脸的邻域边缘响应图; 其次, 引入吸引描述子, 参考边缘响应图的局部、全局平均灰度值和邻域中心灰度值完成局部吸引模式编码; 再次, 遍历整幅图像, 得到人脸 ALSGC 特征图并对 ALSGC 特征图分块分别计算, 得到各个分块中不同模式的统计直方图; 最后, 级联所有分块的统计直方图后生成对应的特征向量, 以支持向量机完成分类识别。所提算法克服了 LBP、LDP、LDN 等算法提取一阶特征有效性的不足, 以及 DLDP、CSLDP、GCSLDP 等算法提取的二阶特征对表情、姿态、饰物遮挡、光照、随机噪声等变化敏感的缺点, 较好地实现了特征提取有效性与特征编码稳定性的平衡, 兼顾了识别率和稳健性。

关键词: 局部吸引模式; 二阶梯度轮廓; Kirsch 算子; 特征提取; 人脸识别

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2021140

Face recognition algorithm based on attractive local second gradient contours

YE Xueyi, QIAN Dingwei, YING Na, WANG Tao

Laboratory of Pattern Recognition & Information Security, Hangzhou Dianzi University, Hangzhou 310018, China

Abstract: In order to solve the problem that traditional LDP algorithms is difficult to balance the effectiveness of feature extraction and the stability of feature encoding, an attractive local second gradient contours (ALSGC) face feature extraction algorithm was proposed. Firstly, the Kirsch operator was used to calculate the neighborhood edge response image of the face image. Secondly, an attraction descriptor was introduced and the local average gray value and global average gray value of the edge response image and the neighborhood center gray value was combined to complete the local attraction pattern encoding. The entire image was traversed to get the ALSGC face feature map, the ALSGC feature map was divided to blocks and the statistical histograms of different patterns for each block was obtained by calculation. Finally, the statistical histograms of all the blocks were cascaded to generate the corresponding feature vector and support vector machine was used to complete classification and recognition. The proposed algorithm not only overcame the insufficient effectiveness of LBP, LDP, LDN and other algorithms who extract

收稿日期: 2020-12-03; 修回日期: 2021-07-10

基金项目: 国家自然科学基金资助项目 (No.60802047, No.60702018); 国家自然科学基金重点基金资助项目 (No.U19B2016)

Foundation Items: The National Natural Science Foundation of China (No.60802047, No.60702018), Major Program of the National Natural Science Foundation of China (No.U19B2016)

first-gradient features, but also reduced the sensitivity to changes in expression, posture, occlusion, lighting and random noise of methods such as DLDP, CSLDP, GCSLDP who extracting second-gradient features. It better achieved the balance between the effectiveness of feature extraction and the stability of feature encoding, and took into account the recognition rate and robustness.

Key words: local attractive pattern, second gradient contours, Kirsch operator, feature extraction, face recognition

1 引言

人脸识别因其广泛的应用领域、非侵入式且便利的采集方式^[1], 一直是被深入研究但又亟待突破的研究热点。实际应用中广泛存在的, 诸如光照、遮挡、姿态、表情等因素, 依然是造成识别性能剧烈变化的主要原因, 这些因素严重影响了人脸特征提取和表征的稳定性。因此, 寻求解决该问题的关键在于如何稳定、有效地提取并表征人脸特征, 提高人脸识别算法的稳健性。其中, 有效性是描述特征提取效果好坏的重要指标之一, 体现了提取的特征信息对于分类的作用大小, 最直观的表现是识别率的高低。

人脸特征提取的研究主要分为全局特征提取和局部特征^[2]提取两大类。全局特征提取着眼于人脸图像的整体属性, 虽然具有较高的计算效率, 但忽视了人脸细节的刻画, 对局部边缘特征提取不充分。典型的方法有主成分分析 (principal component analysis, PCA)^[3]、线性判别分析 (linear discriminant analysis, LDA)^[4]等; 局部特征提取专注于刻画人脸的局部或者微纹理结构, 对原始图像进行特定模式编码, 获得对应的特征图, 对于光照、表情、姿态等干扰的稳健性较好。经典方法有局部二值模式 (local binary pattern, LBP)^[5]、局部方向模式 (local directional pattern, LDP)^[6]等。这类方法识别速度非常快, 能在速度与精度上取得一个有效平衡, 得到了广泛的研究。Ojala 等^[5]提出了局部特征提取经典的 LBP 方法, 该方法侧重于利用中心像素与邻域像素的灰度值差异, 很好地刻画了图像纹理。由于仅利用灰度域二元像素点之间的大小关系编码, 梯度信息缺

失, 对光照、噪声敏感。为了解决 LBP 存在的问题, Jabid 等^[6]提出了一种基于图像边缘响应之间大小关系的局部方向模式, 通过将图像块与 Kirsch 模板算子卷积, 得到边缘响应, 记录边缘响应最大的几个位置并进行编码。该方法有效利用了一阶梯度信息, 对光照和噪声的稳健性有所提高, 但识别时间较长, 识别效果一般。为了使 LDP 编码更加紧凑, Rivera 等^[7]提出了一种局部方向数 (local directional number, LDN) 模式, 通过抓取最大和最小的边缘响应信息进行编码, 减少了编码位数, 提升了识别速度, 编码效率更高。但由于该方法只选取两个极端边缘响应位置, 特征不稳定。在光照情况下, 一阶梯度信息对识别效果的提升不明显。为了提取更深层次的梯度信息, 李照奎等^[8]提出了差值局部方向模式 (difference local directional pattern, DLDP) 算法, 杨恢先等^[9]提出了一种基于中心对称局部方向模式 (center-symmetric local directional pattern, CSLDP) 的算法。其中, DLDP 对邻域内边缘响应相邻作差, 找出最大的几个边缘响应差值对应的位置进行编码, 利用了二阶梯度信息, 但编码较复杂, 仅考虑了近邻边缘响应之间的强度变化, 导致识别时间较长, 识别效果不佳。CSLDP 比较邻域内中心对称边缘响应之间的强度关系, 提升了识别速度, 但特征信息太少, 编码稳定性一般。王晓华等^[10]提出了梯度中心对称局部模式 (gradient center symmetric local directional pattern, GCSLDP), 同时利用了相邻边缘响应差值信息与中心对称边缘响应强度信息, 识别率有所提升, 但存在信息提取不完全和稳定性较差的缺点。基于深度学习的人脸识别方法在大规模的数据库中



达到了很高的识别精度，但这类方法需要大量的训练成本（包括时间成本和算力成本），对于数据库的样本数量要求也很高，适用于样本数量充足、硬件算力强大的场景。而手工描述符方法属于相对灵活的轻量级方法，对样本数目和算力基本没有要求，识别的实时性也很高，适用于大多数样本数量不多、硬件算力一般的场景。

LBP、LDP、LDN 等方法均利用一阶灰度或梯度信息进行编码，对姿态、表情的稳健性较高，但特征提取有效性有所欠缺。二阶梯度带来深层次信息的同时，也增大了噪声的影响。DLDP、CSLDP、GCSLDP 等利用二阶梯度编码的方法，编码稳定性不足，对随机噪声、表情、姿态敏感。为了提升上述算法的稳定性，Khadiri 等^[11]提出了基于排斥吸引模式的局部二阶梯度轮廓（repulsive- and-attractive local binary gradient contours, RALBGC）算法，采用吸引模式局部微结构编码，使稳定性得到了很大的提升，在抗干扰能力方面表现出优异的性能，但人脸识别的有效性有所欠缺。考虑 LBP、LDP、LDN 等方法对干扰的稳健性较高，但特征提取有效性不足^[8]。DLDP、CSLDP、GCSLDP 等方法提升了特征提取的有效性，但明显降低了对干扰的稳健性^[12]。

吸引局部二阶梯度轮廓（attractive local second gradient contours, ALSGC）模式引入二阶梯度信息提升方法的有效性、提升方法对干扰的稳健性，同时满足有效性和对干扰的稳健性要求。首先，将人脸图像 3×3 邻域与 Kirsch 模板算子卷积，得到对应的边缘响应图，再分别对边缘响应图进行吸引模式编码，结合局部平均灰度值、邻域中心灰度值和全局平均灰度值的吸引模式编码，生成 ALSGC 码。遍历整个人脸图像，最终将其转化为与原图大小一致的 ALSGC 人脸图。然后，对 ALSGC 人脸图进行分块，统计每个分块中不同模式数的个数，形成多个统计直方图。最后，将统计直方图中所有子块的模式数级联，

生成人脸图像对应的特征向量，用支持向量机分类器进行分类识别。

2 吸引局部二阶梯度轮廓模式

2.1 吸引排斥模式介绍

排斥和吸引模式描述了像素的局部微结构，通过比较相邻 3 个元素组成的三元组之间的大小关系，对图像进行吸引和排斥模式编码，如图 1 所示。一位二进制编码包含 3 个元素之间的细节信息，提取的信息更丰富，特征提取有效性更高；结合局部均值和全局均值，在纹理识别领域取得了不错的效果，初步证实了吸引排斥模式的理论可行性^[11]。

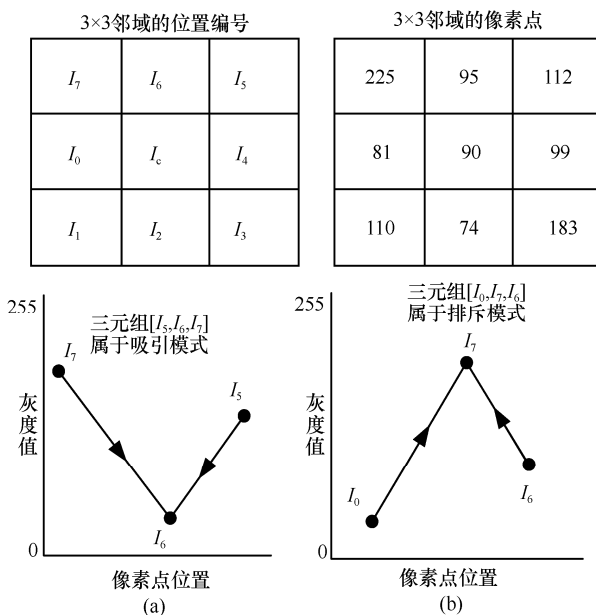


图 1 吸引和排斥的微结构（模式编码）

图 1 中， I_c 表示局部邻域中心像素的灰度值， I_k 表示 3×3 局部邻域的灰度值 ($k \in \{0, 1, \dots, 7\}$)。排斥和吸引算子分别将 3×3 局部邻域中围绕中心像素 I_c 的所有相邻三元组划分为两类，即排斥与非排斥模式、吸引与非吸引模式。图 1 举例说明了两个分别代表排斥和吸引模式的三元组，在图 1(a)中，由于 I_6 像素点的灰度值比 I_5 和 I_7 的灰度值都大，三元组 $[I_5, I_6, I_7]$ 被定义为吸

引模式。相反,在图 1(b)中,由于 I_7 像素点的灰度值比 I_0 和 I_6 都小,三元组 $[I_0, I_7, I_6]$ 被定义为排斥模式。

基于上述排斥和吸引模式的概念,定义排斥和吸引模式的二进制阈值函数 $v(\cdot)$ 和 $\nu(\cdot)$, 具体定义如式 (1)、式 (2) 所示。

吸引二元阈值函数 (attractive binary thresholding function, ABTF) $\nu(\cdot)$:

$$\nu(a, b, c) = \varphi(a, b) \otimes \varphi(c, b) = \begin{cases} 1, & a \geq b \\ 0, & \text{其他} \end{cases} \otimes \begin{cases} 1, & c \geq b \\ 0, & \text{其他} \end{cases} = \begin{cases} 1, & a \geq b, c \geq b \\ 0, & \text{其他} \end{cases} \quad (1)$$

排斥二元阈值函数 (the repulsive binary thresholding function, RBTF) $v(\cdot)$:

$$v(a, b, c) = \bar{\varphi}(a, b) \otimes \bar{\varphi}(c, b) = \begin{cases} 1, & a \leq b \\ 0, & \text{其他} \end{cases} \otimes \begin{cases} 1, & c \leq b \\ 0, & \text{其他} \end{cases} = \begin{cases} 1, & a \leq b, c \leq b \\ 0, & \text{其他} \end{cases} \quad (2)$$

其中, $\varphi(a, b) = \bar{\varphi}(b, a)$, $\otimes$ 表示乘法运算符, a 、 b 、 c 分别表示排斥或吸引三元组对应的元素值, $\varphi(a, b)$ 表示二元阈值函数, 若 $a \leq b$, 结果为 1, 反之为 0。

2.2 ALSGC 分析

传统的 LDP 特征提取算法对于特征提取的有效性以及对遮挡、表情、光照等干扰的稳健性不能取得一个较好的平衡效果^[12], 基于这两个方面, 提出了 ALSGC 算法。

(1) 对于特征提取的有效性: 将人脸灰度图转变为边缘响应图, 对边缘响应图进行吸引模式编码, 得到吸引局部二阶梯度轮廓特征图, 充分利用二阶梯度空间的图像细节信息, 同时考虑中心像素灰度值、局部平均灰度值以及全局平均灰度值之间的关系, 有效提取并表征图像特征。

(2) 对于干扰的稳健性: 该方法编码有着较强的稳定性。首先, 一位编码由 3 个边缘响应点决定, 随机噪声与光照较难对编码造成影响。其次, 对相邻边缘响应采用不对称编码, 突出吸引

模式编码, 在较大面积遮挡的情况下, 该方法编码受到的影响明显小于其他对比方法。该方法对遮挡、光照、随机噪声等干扰的稳健性明显高于其他对比方法。

2.3 ALSGC 编码

由于吸引排斥模式^[11]对于局部微结构的表征非常稳定, 抗干扰能力很强, 本文选择吸引模式, 在梯度域借鉴 RALBGC 在纹理识别中的编码方式, 以提取更深层次且稳定的梯度判别特征, 改进同样在梯度域编码的 LDP 类人脸识别方法。

ALSGC 编码具体步骤如下。

首先, 将人脸图像 3×3 邻域块 I 与 Kirsch 算子^[13]卷积 $M_a (a = 0, 1, 2, 3, 4, 5, 6, 7)$, 对得到的 8 个方向的边缘响应取绝对值, 得到 $m_a (a = 0, 1, 2, 3, 4, 5, 6, 7)$ 。Kirsch 算子如图 2 所示。再引入吸引模式对 m_a 进行编码, 沿着中心点 m_c 找出所有相邻的边缘响应三元组, 对其进行吸引模式编码, 得到一串二进制编码, 描述了二阶梯度轮廓信息, 提升了特征提取的有效性。为了提升对噪声和单调灰度变化的稳健性, 引入局部平均灰度值 μ :

$$\mu = \frac{1}{S+1} \left(\sum_{s=0}^{S-1} I_s + I_c \right) \quad (3)$$

$$\begin{array}{cccc} \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} & \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\ M_0 & M_1 & M_2 & M_3 \\ \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \\ M_4 & M_5 & M_6 & M_7 \end{array}$$

图 2 Kirsch 算子

全局平均灰度值 $\bar{\mu}$:

$$\bar{\mu} = \frac{\sum_{p=0}^{P-1} \sum_{q=0}^{Q-1} I(p, q)}{P \times Q} \quad (4)$$

其中, S 表示以 I_c 为中心的 3×3 邻域像素点数, s 表示对应的像素序号, I_s 表示对应序号像素值。 P 表示分块内横向像素点的个数, Q 表示分块内纵向像素点的个数, $I(p, q)$ 表示以分块



左下角像素点为坐标原点，坐标为 p 和 q 的像素点灰度值。

其次，结合局部平均灰度值 μ 、邻域中心灰度值 I_c 和全局平均灰度值 $\bar{\mu}$ 组成的吸引模式三元组，得到最终的二进制编码：

$$g_{\text{ALSGC}} = \begin{bmatrix} v(\mu, I_c, \bar{\mu}) \\ v(m_0, m_1, m_2) \\ v(m_1, m_2, m_3) \\ v(m_2, m_3, m_4) \\ v(m_3, m_4, m_5) \\ v(m_4, m_5, m_6) \\ v(m_5, m_6, m_7) \\ v(m_6, m_7, m_0) \\ v(m_7, m_0, m_1) \end{bmatrix} \quad (5)$$

最后，将二进制编码 g_{ALSGC} 转化为十进制数，即中心像素点 x 的 ALSGC 编码值：

$$\text{ALSGC}(x) = \omega_9^T g_{\text{ALSGC}} \quad (6)$$

其中， ω_9 定义为：

$$\omega_9^T = [2^8 \ 2^7 \ \dots \ 2^1 \ 2^0] \quad (7)$$

如图 3 所示，举例展示了 ALSGC 的编码过程。

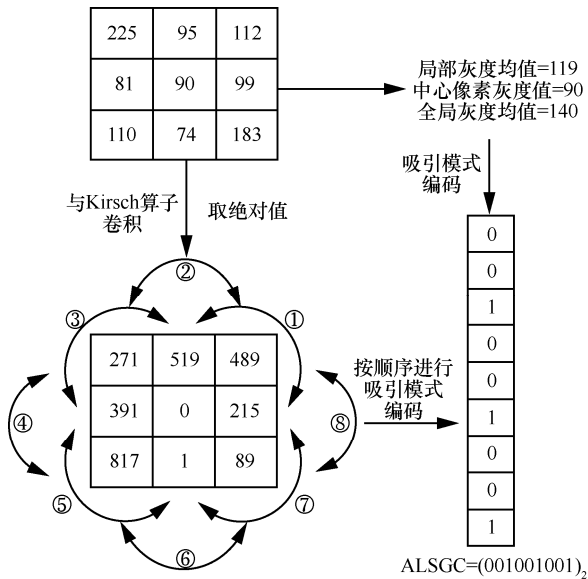


图 3 ALSGC 编码过程

对应排斥的局部二阶梯度轮廓描述 (repulsive

local second gradient contour, RLSGC) 算法的编码过程将 ALSGC 的吸引模式编码改为排斥模式编码，其他与 ALSGC 算法相同。

根据上述编码过程可知，三元组之间的大小关系一共有 6 种情况，两种属于排斥模式，两种属于吸引模式，两种既不属于排斥模式又不属于吸引模式。既不属于排斥又不属于吸引的两种情况，在 ALSGC 和 RLSGC 两种编码的情况下都为 0。因此，在样本较大的情况下，同一张图片对应 ALSGC 的编码必然存在长度为编码总长度 1/3 的编码为 0 的子编码段，RLSGC 的编码也是如此，前后这两段子编码数值均为 0，位置完全重合；而 ALSGC 编码中剩余的长度为编码总长度 2/3 的子编码段，在零一取反后，则成为了 RLSGC 对应位置的编码段，两种编码方式之间信息高度冗余。

对于同类的不同样本，结合 ALSGC 和 RLSGC 的编码方式，相对于 ALSGC 或 RLSGC 单一的编码方式，增加了同类的不同样本之间的差异，增加了类间距离，降低了识别效果。对于不同类的不同样本，不一定能增加类间距离。综合来看，结合 ALSGC 和 RLSGC 的编码方法在人脸特征提取中无法获得比单独使用两种编码方法更好的识别效果。

2.4 ALSGC 编码的稳定性

ALSGC 吸引模式占比为 1/3，编码为 1，非吸引模式占比为 2/3，编码为 0，随机噪声可能会使三元组在不同的非吸引模式间变化，相对于其他对比方法，不易导致 ALSGC 编码发生变化。

在黑块极限遮挡情况下，白块同理，人脸图像前后对比如图 4 所示。

为客观描述不同编码方法受遮挡影响程度的大小，定义编码信息损失函数如下：

$$\Delta C = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{l=1}^L \delta(m, n, l)}{M \times N \times L} \quad (8)$$



图4 人脸图像在黑块极限遮挡情况下对比

其中, M 和 N 分别表示遮挡块的纵向像素点数和横向像素点数; m 、 n 分别表示在遮挡块内, 以遮挡块左上角像素点为原点, 竖直方向与原点相隔的距离和水平方向与原点相隔的距离; L 表示当前方法每个像素点编码串的长度; l 表示编码串中每个编码的位置; $\delta(m,n,l)$ 表示坐标为 (m,n) 的像素点中第 l 位编码在遮挡前后的变化情况, 具体计算式为:

$$\delta(m,n,l) = \begin{cases} 0, & \text{Code}(m,n,l) = \text{Code}'(m,n,l) \\ 1, & \text{其他} \end{cases} \quad (9)$$

其中, $\text{Code}(m,n,l)$ 和 $\text{Code}'(m,n,l)$ 分别表示遮挡前后坐标为 (m,n) 的像素点中第 l 位编码的值。

LBP 算法是根据中心像素点灰度值和 3×3 邻域像素点灰度值之间的大小关系进行编码, 在极限遮挡和样本足够大的情况下, LBP 编码从 0、1 各占一半的情况变为全 0 或全 1, $\Delta C_{\text{LBP}} = 1/2$ 。

LDP、LDN、DLDP、CSLDP、GCSLDP 等算法根据 3×3 邻域边缘响应之间的二元组大小关系进行编码, 在极限遮挡的情况下, 依靠相邻边缘响应

或对称边缘响应二元组大小关系的编码几乎遭到了完全破坏, 趋于随机分布, $\Delta C \approx 1$ 在不考虑第一位均值编码的情况下, ASLGC 算法由于不对称编码的特性, 约 $1/3$ 的编码为 1, $2/3$ 的编码为 0; 在遮挡前后, 改变的编码位数只占编码总长度的 $1/3$, 即 $\Delta C \approx 1/3$ 。因此, ASLGC 编码与其他对比方法相比, 对于较大面积的遮挡有着较强的稳定性。

通过图 5 例子来验证 ASLGC 的稳定性。图 5 中, 将原始图像块随机增减一个数表示噪声, 将第 3 行像素值都增加 5 表示光照变化, 将第 1 行和第 1 列设置为近似相等的像素值表示遮挡 1 和遮挡 2。

由图 5 可知, 与原始图像块中各种算法的编码进行对比, 在添加噪声情况下, LDP、LDN、DLDP、CSLDP、GCSLDP 编码值发生了变化; 在光照变化情况下, LBP 编码值发生了变化; 在遮挡 1 情况下, LBP、LDP、LDN、DLDP、CSLDP、GCSLDP 编码值都发生了变化; 在遮挡 2 情况下, CSLDP 编码值发生变化。在所有情况下, 只有 ASLGC 编码值不发生变化, 稳定性较强, 体现了 ASLGC 对遮挡、光照、随机突变噪声的稳健性。

2.5 ASLGC 人脸识别的流程

人脸识别流程如图 6 所示, 具体流程如下。

(1) 使用 Kirsch 算子对人脸图像进行边缘检测并通过 ASLGC 编码对得到的边缘响应进行特

<table><tr><td>85</td><td>32</td><td>26</td></tr><tr><td>53</td><td>50</td><td>10</td></tr><tr><td>60</td><td>38</td><td>45</td></tr></table>	85	32	26	53	50	10	60	38	45	<table><tr><td>90</td><td>45</td><td>37</td></tr><tr><td>49</td><td>55</td><td>11</td></tr><tr><td>64</td><td>40</td><td>41</td></tr></table>	90	45	37	49	55	11	64	40	41	<table><tr><td>85</td><td>32</td><td>31</td></tr><tr><td>53</td><td>50</td><td>15</td></tr><tr><td>60</td><td>38</td><td>50</td></tr></table>	85	32	31	53	50	15	60	38	50	<table><tr><td>51</td><td>52</td><td>51</td></tr><tr><td>53</td><td>50</td><td>10</td></tr><tr><td>60</td><td>38</td><td>45</td></tr></table>	51	52	51	53	50	10	60	38	45	<table><tr><td>85</td><td>32</td><td>26</td></tr><tr><td>80</td><td>50</td><td>10</td></tr><tr><td>80</td><td>38</td><td>45</td></tr></table>	85	32	26	80	50	10	80	38	45
85	32	26																																															
53	50	10																																															
60	38	45																																															
90	45	37																																															
49	55	11																																															
64	40	41																																															
85	32	31																																															
53	50	15																																															
60	38	50																																															
51	52	51																																															
53	50	10																																															
60	38	45																																															
85	32	26																																															
80	50	10																																															
80	38	45																																															
LBP=(00011100) ₂	LBP=(00011100) ₂	LBP=(00011101) ₂	LBP=(01111100) ₂	LBP=(00011100) ₂																																													
LDP=(11001000) ₂	LDP=(10001001) ₂	LDP=(11001000) ₂	LDP=(10001001) ₂	LDP=(11001000) ₂																																													
LDN=4×8 ¹ +2×8 ⁰	LDN=4×8 ¹ +6×8 ⁰	LDN=4×8 ¹ +2×8 ⁰	LDN=7×8 ¹ +6×8 ⁰	LDN=4×8 ¹ +2×8 ⁰																																													
DLDP=(01011000) ₂	DLDP=(00011010) ₂	DLDP=(01011000) ₂	DLDP=(00001011) ₂	DLDP=(01011000) ₂																																													
CSLDP=(0111) ₂	CSLDP=(0110) ₂	CSLDP=(0111) ₂	CSLDP=(1110) ₂	CSLDP=(0100) ₂																																													
GCSLDP=1×4 ¹ +1×4 ⁰	GCSLDP=4×4 ¹ +1×4 ⁰	GCSLDP=1×4 ¹ +1×4 ⁰	GCSLDP=6×4 ¹ +3×4 ⁰	GCSLDP=1×4 ¹ +1×4 ⁰																																													
ASLGC=(01000100) ₂	ASLGC=(01000100) ₂	ASLGC=(01000100) ₂	ASLGC=(01000100) ₂	ASLGC=(01000100) ₂																																													
(a)原始编码	(b)噪声编码	(c)光照编码	(d)遮挡1编码	(e)遮挡2编码																																													

图5 算法编码稳定性



征提取, 获得对应的特征图。

(2) 对特征图进行不重叠分块处理, 并对每个分块逐一统计分块内不同模式编码的个数, 形成对应子块的统计直方图(统计直方图的横坐标为不同的模式编码值, 纵坐标为不同模式编码值在子块中的出现次数)。

(3) 将统计直方图中所有子块的模式数进行级联, 作为该幅人脸图像的特征向量。

(4) 利用支持向量机分类器进行分类, 完成识别。

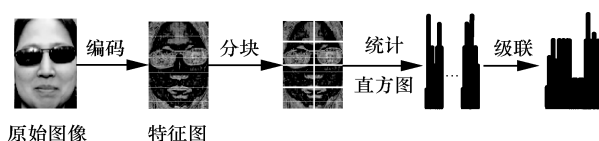


图6 人脸识别流程

3 实验结果及分析

3.1 实验环境与参数设置

为了验证本算法的性能, 本研究选择在CAS-PEAL^[14]、ORL^[15]、GT^[16]人脸数据库上进行实验, 这些数据库分别包含了表情、遮挡、光照、姿态、突变噪声等干扰。为客观比较本文算法和其他算法, 本实验均采用 10×10 的分块(即将待编码图像分为100幅子图, 每幅子图大小为待编码图像大小的 $1/100$)。为验证算法有效性, 本文选择与LBP、LDP、LDN、DLDP、CSLDP、GCSLDP等进行比较。实验所用硬件设备为AMD Ryzen 54600U, RAM为16 GB, 仿真环境为MATLAB R2018b。

3.2 基于CAS-PEAL人脸库的结果和分析

CAS-PEAL人脸库包含了1 040名中国人共99 450幅头肩部图像, 包含了姿态、表情、背景、饰物、光照、距离、时间7个子库。考虑实际应用中对面脸识别性能影响最明显的干扰因素, 本文选取其中的表情、饰物、光照子库进行实验, 其中表情子库包含376人共1 880幅图像, 饰物子

库包含433人共2 598幅图像, 光照子库包含232人共2 088幅图像, 部分样本如图7所示。

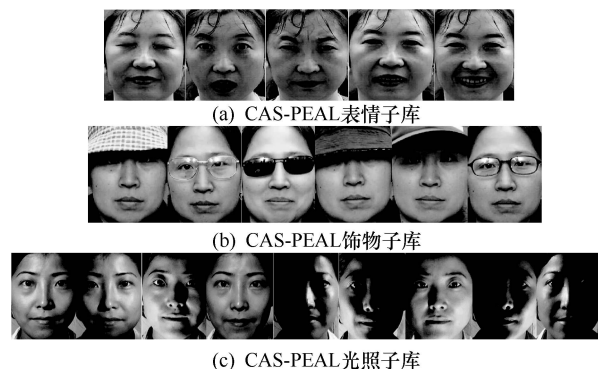


图7 CAS-PEAL人脸库部分样本

为保证实验的准确性, 在每类人脸样本中随机选取 k 张图片作为训练样本, 剩下的图片作为测试样本, 取重复计算50次的均值作为最终识别率。实验结果见表1~表3。

表1 在CAS-PEAL表情子库上的识别率

算法	每类随机选择样本数 k			
	1	2	3	4
LBP	96.61%	98.76%	99.32%	99.51%
LDP	96.54%	98.59%	99.37%	99.56%
LDN	96.34%	98.67%	99.47%	99.67%
DLDP	96.43%	98.31%	98.89%	99.06%
CSLDP	95.21%	98.13%	98.91%	99.11%
GCSLDP	96.87%	98.49%	99.2%	99.37%
本文	96.88%	99.02%	99.73%	99.94%

由表1可知, 在表情子库下, 大部分算法识别效果较好, 但本文算法有着明显的优势。LDP、LDN等方法通过比较边缘响应的大小进行编码, 受到的影响较小, 识别效果普遍较高, DLDP、CSLDP、GCSLDP等方法通过对边缘响应计算差值后进行编码, 容易受到表情、随机噪声的影响, 编码不稳定, 识别效果一般。本文方法依靠三元组不对称编码, 编码稳定, 对表情与随机噪声稳健性好。

表 2 在 CAS-PEAL 饰物子库上的识别率

算法	每类随机选择样本数 k				
	1	2	3	4	5
LBP	54.49%	70.98%	78.76%	84.38%	87.13%
LDP	55.92%	70.92%	79.29%	85.91%	88.07%
LDN	57.27%	73.9%	81.45%	86.26%	89.56%
DLDP	56.52%	74.13%	82.91%	89.84%	92.03%
CSLDP	58.72%	73.73%	81.37%	87.51%	90.93%
GCSLDP	58.82%	74.36%	82.9%	88.34%	91.24%
本文	65.45%	79.33%	87.61%	93.07%	96.81%

表 3 在 CAS-PEAL 光照子库上的识别率

算法	每类随机选择样本数 k			
	3	4	5	6
LBP	51.49%	57.93%	66.82%	73.28%
LDP	53.32%	59.4%	67.78%	74.15%
LDN	59.78%	65.43%	73.6%	79.31%
DLDP	59.31%	65.34%	73.28%	78.45%
CSLDP	67.29%	74.91%	80.28%	84.63%
GCSLDP	71.73%	78.97%	84.82%	88.21%
本文	72.85%	80.86%	86.31%	90.52%

在饰物子库下,人脸的一些部位被遮挡了,对识别效果造成了不小的影响。此时,降低遮挡对于算法编码的影响就尤为重要。对于大部分对比算法,遮挡区域的编码与该区域无遮挡时的编码相差很大,对于分类识别的贡献大大降低,但对于依靠吸引模式三元组进行不对称编码的 ALSGC 算法,遮挡区域的编码与该区域无遮挡时的编码相差较小。ALSGC 编码受遮挡的影响更小,与未遮挡时的编码更接近,因此识别效果更好。由表 2 可知,在随机选取 2 张人脸图像作为训练样本的情况下,ALSGC 算法相比 LBP、LDP、LDN、DLDP、CSLDP、GCSLDP 分别提高了 8.7%、8.79%、7.35%、7.01%、7.6%、6.68%。

在光照子库下,人脸图像会受到来自不同方向、不同程度的光照影响。LBP 在灰度域进行编

码,主要利用图像灰度值差异,刻画图像纹理,侧重于描述局部图像整体信息,表现欠佳。LDP 主要利用局部图像的边缘响应,抓住边缘响应绝对值最大的几个方向进行编码,效果有所提升,但无法精确定位最主要的纹理方向。LDN 在 LDP 的基础上,抓住最大和最小的边缘响应值,相比之下,编码不容易受到影响。但在人脸图像的平滑区域对噪声比较敏感,会对识别效果产生负面影响。DLDP 方法通过对相邻边缘响应计算差值进行编码,提取了更深层次的梯度特征。CSLDP 通过对中心对称方向上的边缘响应计算差值进行编码,可以体现边缘的大致纹理方向,但是由于编码信息太少,难以大幅提升识别效果。GCSLDP 则结合了 CSLDP 提取的中心对称差值信息和 DLDP 提取的相邻差值信息,提升了识别率。本文算法对相邻边缘响应采用了吸引模式三元组不对称编码,结合全局均值、中心灰度值与局部均值的三元组编码,不仅有效利用了梯度信息,还极大提升了编码的稳定性。由表 3 可知,在随机选取 4 张人脸图像作为训练样本的情况下,ALSGC 算法相比 LBP、LDP、LDN、DLDP、CSLDP、GCSLDP 分别提高了 39.58%、36.13%、23.58%、23.75%、7.9%、2.3%。

综合表 1、表 2、表 3 可以发现,在 CAS-PEAL 表情、饰物、光照 3 个人脸子库上,随着训练样本与测试样本比例的增加,所有算法的识别率均不断增加,但 ALSGC 算法相对于其他算法,由于特征提取有效性最高,特征编码最为稳定,识别率始终处于领先地位,尤其是在饰物子库、光照子库,本文算法的优势更加明显。实验数据也验证了第 1.3 节中编码稳定性分析,只有 ALSGC 在噪声、光照、遮挡 1、遮挡 2 的情况下编码保持稳定不变,体现了 ALSGC 对于光照和遮挡的稳健性。

3.3 基于 ORL 人脸库的结果和分析

ORL 人脸库由英国剑桥大学的 Olivetti 实验



室创建, 包含 40 名志愿者的 400 张人脸样本, 主要包含姿态和表情变化, 分辨率为 112 dpi×92 dpi, 实验中选择归一化到 100 dpi×100 dpi 分辨率下, 部分样本如图 8 所示。



图 8 ORL 人脸库部分样本

为保证实验的准确性, 在每类人脸样本中随机选取 k 张图片作为训练样本, 剩下的 $10-k$ 张图片作为测试样本, 取重复计算 50 次的均值作为最终识别率。实验结果见表 4。

表 4 各种算法在 ORL 人脸库上的识别率

算法	每类随机选择样本数 k				
	3	4	5	6	7
LBP	82.07%	87.89%	91.25%	94.17%	95.08%
LDP	80.53%	87.08%	90.52%	93.27%	95.13%
LDN	82.48%	88.68%	91.77%	94.43%	96.12%
DLDP	81.42%	87.21%	90.63%	93.15%	94.46%
CSLDP	81.46%	86.23%	89.83%	92.22%	93.37%
GCSLDP	79.35%	84.35%	88.15%	91.34%	93.52%
本文	83.5%	90.5%	93.83%	95.65%	97.29%

ORL 人脸库中, 表情和姿态的变化考验算法对于纹理变化的稳健性。由表 4 可以发现, DLDP、CSLDP、GCSLDP 通过对边缘响应计算差值后进行编码, 相当于再次计算了一次梯度, 虽然能提取更深层次的信息, 但是也放大了噪声变化带来的影响。LDP、LDN 等方法通过比较边缘响应的大小进行编码, 受到的影响较小。本文算法利用吸引模式进行编码, 同时利用中心像素的灰度信

息, 引入局部二阶梯度轮廓信息, 相较对比方法, 可以更好地抓住纹理信息, 特征提取更有效, 稳定性更高, 性能更好。在训练样本与测试样本比例为 1:1 时, 本文算法相比 LBP、LDP、LDN、DLDP、CSLDP、GCSLDP 分别提高了 2.83%、3.66%、2.24%、3.53%、4.45%、6.44%。

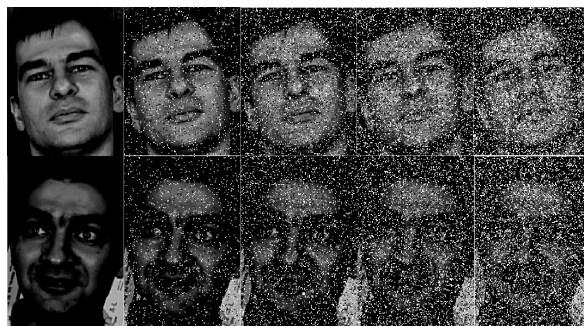
3.4 基于 GT 人脸库的噪声实验

为进一步验证 ALSGC 算法对噪声的稳健性, 在 GT 人脸库进行了噪声实验, GT 人脸库包含 50 个人的 750 张不同角度、不同表情的正面照片, 部分样本如图 9 所示。



图 9 GT 人脸库部分样本

本文在每类人的 15 张样本中随机选取 5 张作为训练样本, 剩余 10 张作为测试样本。随后对每幅图像添加椒盐噪声, 添加噪声的比例为 0~40%。不同百分比噪声情况下的样本图像如图 10 所示。



(a) 灰度图 (b) 10%噪声 (d) 30%噪声 (c) 20%噪声 (e) 40%噪声

图 10 不同百分比噪声下的图像

同时为了客观描述不同算法识别效果受噪声影响的程度, 定义识别率下降速度函数^[17], 计算式为:

$$\delta = \frac{r-f}{r} \times 100\% \quad (10)$$

其中, r 表示未加噪声的识别率, f 表示受到噪声影响后的识别率, δ 表示识别率的下降速度, δ 越高表示算法抗噪声能力越差。各种算法在不同百分比噪声下的识别率见表 5。

表 5 GT 人脸库噪声实验结果

算法	添加椒盐噪声比例					δ
	0	10%	20%	30%	40%	
LBP	80.27%	76.7%	71.02%	61.87%	48.9%	39.08%
LDP	76.44%	76.15%	72.14%	69.15%	66.49%	13.02%
LDN	79.27%	77.06%	72.73%	65.91%	57.56%	27.39%
DLDP	79.62%	77.1%	73.14%	64.77%	58.28%	26.80%
CSLDP	71.57%	65%	62.97%	61.55%	60.57%	15.37%
GCSLDP	73.06%	66.87%	59.97%	50.17%	42.16%	42.29%
本文	80.89%	79.84%	78.17%	77.59%	73.26%	9.43%

4 结束语

为了提升人脸特征提取的有效性与特征编码的稳定性, 进一步提高人脸识别性能, 本文提出了一种基于吸引二阶梯度局部轮廓模式 (ALSGC) 的人脸识别算法。该算法通过吸引模式编码, 更加深入、有效地提取二阶梯度轮廓信息, 克服了 LDP、LDN 等算法提取一阶梯度信息导致的特征提取有效性不足的缺点, 同时 ALSGC 采用三元组编码, 提升了编码的稳定性, 避免了 DLDP、CSLDP、GCSLDP 等算法由于引入二阶梯度带来的噪声敏感问题。而且由于 ALSGC 编码不对称性, 对遮挡尤其是大面积遮挡, 编码受遮挡的影响程度是上述算法中最小的。实验数据也表明, ALSGC 算法既保证了特征提取的有效性, 又对干扰有较强的稳健性。在光照、随机噪声、表情、遮挡、姿态等变化下, 本文算法能够达到更高的识别率, 证明了算法的纹理特征描述能力与稳健性。但是三元组编码带

来稳定性的同时, 提高了计算复杂度, 后续将针对该问题进行研究。

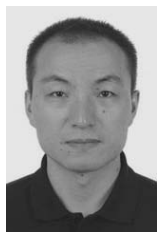
参考文献:

- [1] JAYARAMAN U, GUPTA P, GUPTA S, et al. Recent development in face recognition[J]. Neurocomputing, 2020(408): 231-245.
- [2] ROUSSEUW P J, HUBERT M. Anomaly detection by robust statistics[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2018, 8(2): e1236.
- [3] KIM C, KLABJAN D. A simple and fast algorithm for L1-norm kernel PCA[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 1842-1855.
- [4] YE Q L, ZHAO H H, FU L Y, et al. Underlying connections between algorithms for nongreedy LDA-L1[J]. IEEE Transactions on Image Processing, 2018, 27(5): 2557-2559.
- [5] OJALA T, VALKEALAHTI K, OJA E, et al. Texture discrimination with multidimensional distributions of signed gray-level differences[J]. Pattern Recognition, 2001, 34(3): 727-739.
- [6] JABID T, KABIR M H, CHAE O. Local Directional Pattern (LDP) for face recognition[C]//Proceedings of 2010 Digest of Technical Papers International Conference on Consumer Electronics (ICCE). Piscataway: IEEE Press, 2010: 329-330.
- [7] RIVERA A, ROJAS CASTILLO J, CHAE O O. Local directional number pattern for face analysis: face and expression recognition[J]. IEEE Transactions on Image Processing, 2013, 22(5): 1740-1752.
- [8] 李照奎, 丁立新, 王岩, 等. 基于差值局部方向模式的人脸特征表示[J]. 软件学报, 2015, 26(11): 2912-2929.
LI Z K, DING L X, WANG Y, et al. Face feature representation based on difference local directional pattern[J]. Journal of Software, 2015, 26(11): 2912-2929.
- [9] 杨恢先, 翟云龙, 蔡勇勇, 等. 基于中心对称梯度幅值相位模式的单样本人脸识别[J]. 光电子·激光, 2015, 26(5): 969-977.
YANG H X, ZHAI Y L, CAI Y Y, et al. Face recognition based on center-symmetric gradient magnitude and phase pattern[J]. Journal of Optoelectronics-Laser, 2015, 26(5): 969-977.
- [10] 王晓华, 李瑞静, 胡敏, 等. 融合局部特征的面部遮挡表情识别[J]. 中国图象图形学报, 2016, 21(11): 1473-1482.
WANG X H, LI R J, HU M, et al. Occluded facial expression recognition based on the fusion of local features[J]. Journal of Image and Graphics, 2016, 21(11): 1473-1482.
- [11] KHADIRI I, KAS M, EL MERABET Y, et al. Repulsive-and-attractive local binary gradient contours: New and efficient feature descriptors for texture classification[J]. Information Sciences, 2018, 467: 634-653.
- [12] 杨恢先, 刘建, 张孟娟, 等. 双空间局部方向模式的人脸识别[J]. 中国图象图形学报, 2017, 22(11): 1493-1502.
YANG H X, LIU J, ZHANG M J, et al. Face recognition based on double-space local directional pattern[J]. Journal of Image and Graphics, 2017, 22(11): 1493-1502.
- [13] SHABAT A M, TAPAMO J R. A comparative study of the use of local directional pattern for texture-based informal settlement classification[J]. Journal of Applied Research and Technology, 2017, 15(3): 250-258.



- [14] CHEN L, LI Q, JIANG J J. Noisy practical facial super-resolution method via deformable constrained model with small dataset[J]. Multimedia Tools and Applications, 2020, 79(3/4): 2577-2600.
- [15] HJOUJI A, EL-MEKKAoui J, JOURHMANE M, et al. New set of non-separable orthogonal invariant moments for image recognition[J]. Journal of Mathematical Imaging and Vision, 2020, 62(4): 606-624.
- [16] ZHANG G Y, ZOU W B, ZHANG X J, et al. Singular value decomposition based virtual representation for face recognition[J]. Multimedia Tools and Applications, 2018, 77(6): 7171-7186.
- [17] 王鹏, 叶学义, 王涛, 等. 双偏差双空间局部方向模式的人脸识别[J]. 计算机工程与应用, 2021, 57(4): 91-99.
- WANG P, YE X Y, WANG T, et al. Face recognition based on double variation and double space local directional pattern[J]. Computer Engineering and Applications, 2021, 57(4): 91-99.

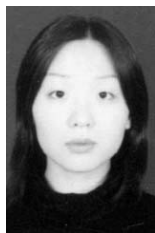
[作者简介]



叶学义 (1973-), 男, 博士, 杭州电子科技大学模式识别与信息安全实验室副教授, 主要研究方向为图像处理、模式识别与信息隐藏等。



钱丁炜 (1996-), 男, 杭州电子科技大学模式识别与信息安全实验室硕士生, 主要研究方向为图像处理。



应娜 (1978-), 女, 博士, 杭州电子科技大学讲师, 主要研究方向为语音信号处理与图像处理等。



王涛 (1990-), 男, 杭州电子科技大学模式识别与信息安全实验室硕士生, 主要研究方向为图像处理。