

均衡局部模式DOG多尺度融合的人脸识别算法

钟理权, 王 鹏, 卢丛慧, 叶学义

杭州电子科技大学通信工程学院, 浙江 杭州

收稿日期: 2024年7月1日; 录用日期: 2024年8月5日; 发布日期: 2024年9月20日

摘 要

针对局部二值模式(LBP)及局部图结构(LGS)方法因非均衡的提取方式导致特征表达能力不强, 以及单纯提取局部特征存在对人脸信息描述不全面的局限性, 提出基于均衡局部模式的DOG多尺度融合金字塔人脸识别算法。首先针对LBP、LGS的不足, 利用双圆交叉采样和以局部宏观信息为参考的自适应阈值、中心对称的采样图结构实现均衡优化, 提出可变参的ECLBP和FLGS方法, 合称为均衡局部模式, 以增强对关键局部信息的利用; 然后将利用高斯核及图像差分生成的DOG金字塔与均衡局部模式方法融合, 以多尺度特征图在丰富样本信息的同时实现大尺度全局轮廓和小尺度局部细节的融合, 进一步捕捉类间及类内差异特征; 最后加权级联所有特征图各子块的统计直方图, 得到全面准确的特征向量, 采用最近邻分类器匹配, 通过特征图、直方图的主客观对比, 从理论上验证了均衡局部模式获得了信息更丰富、判别性更强的局部特征。在ORL、AR和LFW数据集上, 针对ORL、AR和LFW人脸库和相关典型方法的对比结果发现: 在时间耗费相当时, 均衡局部模式的识别率提升最高达15.52%; 进一步融合DOG金字塔补充多尺度特征后, 识别率再次提升, 最高可达9.24%。实验结果表明, 与现有典型提取方法相比, 均衡局部模式特征明显拥有更强的表征能力及鲁棒性, 多尺度信息进一步增强了特征性能, 尤其在少样本环境, 当样本信息有限时, 整体算法的优势更加明显。

关键词

LBP, LGS, 图论, 金字塔, 均衡局部模式, 人脸识别

Face Recognition Algorithm Based on Balanced Local Pattern and DOG Multi-Scale Fusion

Liquan Zhong, Peng Wang, Conghui Lu, Xueyi Ye

School of Communication Engineering, Hangzhou Dianzi University, Hangzhou Zhejiang

Received: Jul. 1st, 2024; accepted: Aug. 5th, 2024; published: Sep. 20th, 2024

文章引用: 钟理权, 王鹏, 卢丛慧, 叶学义. 均衡局部模式 DOG 多尺度融合的人脸识别算法[J]. 仪器与设备, 2024, 12(3): 408-432. DOI: 10.12677/iae.2024.123054

Abstract

In order to solve the problem that local binary pattern and local graph structure methods lack sufficient feature expression ability because of the unbalanced extraction method, and the limitation that only using local features can't fully describe face information, this paper proposes a method called Face recognition based on DOG multi-scale fusion of Balanced Local Pattern. Firstly, in view of the shortcomings of LBP and LGS, on the basis of balanced optimization by using double-circle cross-sampling, adaptive threshold based on local macro information, and a center-symmetric sampling graph structure, this paper proposes Extended Cross Local Binary Pattern and Four-angle star Local Graph Structure methods with variable parameters, which are collectively called Balanced Local Pattern, which can enhance the extraction of key feature information. Then, the DOG pyramid generated by the Gaussian kernel and image difference is fused with the balanced local pattern method. The supplemented multi-scale feature map enriches the sample information while achieving the fusion of large-scale global contours and small-scale local details, which can further capture inter-class and intra-class difference characteristics. Finally, the comprehensive and accurate feature vector is obtained by weighted cascading the sub-block histograms of all feature map, and the nearest neighbor classifier is used to complete the recognition. The subjective and objective comparison of feature map and histogram theoretically verifies that the balanced local pattern can obtain local features with richer information and stronger discrimination. On the ORL, AR and LFW datasets, the proposed method is compared with the relevant typical methods, and the results on the ORL, AR and LFW databases show that the recognition rate of Balanced Local patterns is improved by up to 15.52% when the time consumption is the same as that of typical methods; after further integration of DOG pyramid, the recognition rate is increased by up to 9.24% again. The experimental results show that compared with the existing typical extraction methods, the balanced local pattern features have stronger representation ability and robustness, and the multi-scale information further enhances the feature performance, especially in the small sample environment, when the sample information is limited, the advantage of the whole algorithm is more obvious.

Keywords

Local Binary Patter, Local Graph Structure, Graph Theory, Pyramid, Balanced Local Pattern, Face Recognition

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

人脸识别是一种通过对人脸图像的关键信息提取从而估计出待识别人脸身份的计算机视觉技术，已广泛应用于金融、医疗、安检等多个领域。人脸图像的动态变化和无约束环境的多重干扰，使得人脸识别在应用中面临巨大的挑战[1]。大量研究表明，精细的局部特征对光照、表情等局部变化表现出极强的鲁棒性[2] [3]，因此被广泛研究。

LBP [4]方法通过邻域像素点与中心点灰度值的快速比较，实现了对局部图像的纹理信息记录，这使得它很快成为一种流行且成功的局部特征提取方法。由于 LBP 以中心像素为固定阈值，使得它对噪声非

常敏感。Liu [5]等人提出基于邻域强度的 LBP 方法(Neighborhood Intensity based LBP, NI-LBP), 利用邻域的灰度均值取代中心点阈值, 增强了算法的抗干扰能力。Yang [6]等人提出自适应的局部三进制模式(Adaptive Local Ternary Pattern, ALTP), 根据韦伯定律选择得到自适应阈值。除以阈值为参考的 LBP 类方法以外, Abusham [7]等人受图论启发, 提出局部图结构(Local Graph Structure, LGS), 通过记录有向像素点间的内在联系, 捕获了稳定且详细的空间信息。Yunanto [8]等人提出扩展对称的局部图结构(Extended Symmetric Local Graph Structure, ESLGS), 通过改进的扩展对称拓扑采样记录了更宽广的空间结构。LBP 类和 LGS 类方法的最大特点是它们基于灰度信息高效地获得关键局部特征, 但非均衡的提取策略使得它们依然缺乏足够的表征能力。

此外, 在实际运用中, 训练样本通常会很少, 甚至只有单张样本, 只用局部特征提取方法常常会忽略了其他有效人脸特征。由于局部特征提取方法一般只是单纯提取局部特征, 获得的人脸信息比较片面, 容易导致算法出现性能瓶颈, 尤其当样本不足时, 这种劣势将会体现的更加明显。除局部特征外, 在不同尺度下得到的人脸特征也大不相同, 在单一尺度提取特征必然会损失其他尺度的人脸信息。高斯差分(Difference of Gaussian, DOG)金字塔[9]通过对不同高斯核的尺度图像作差分, 补充的多尺度人脸特征[10]可以有效扩充样本, 在丰富人脸信息时可以更加准确地描述人脸尺度变化。通过将局部特征与 DOG 金字塔结合, 这种融合型的算法可以学习得到更全面的人脸信息, 进而获得性能提升。从对局部特征的提取方式来看, 主要有利用灰度、梯度[11]和融合多特征信息[12]的特征提取方法。但是梯度或者是融合特征的提取方法计算过程复杂, 如果直接在这些方法基础进一步提取多尺度特征, 其时间耗费比较大, 难以满足识别对实时性的要求。而类似于 LBP、LGS 的算法利用灰度信息, 计算简单, 可以极快捕获判别性局部特征。因此本文考虑在利用灰度信息快速获得关键特征的基础上, 通过 DOG 金字塔补充多尺度特征, 以有效平衡好时间耗费和识别精度之间的关系。

基于上述分析, 提出基于均衡局部模式的 DOG 多尺度融合金字塔人脸识别算法, 以有效增强对人脸特征的利用。算法首先针对 LBP、LGS 的不足, 在局部采样和模式编码过程均衡优化, 提出两种可变参的改进特征提取方法: 延展交叉型局部二值模式(Extended Cross Local Binary Pattern, ECLBP)和四星型局部图形结构(Four-angle star Local Graph Structure, FLGS), 由于两种改进方法的启发思路都是尽可能从均衡提取信息的角度来改善算法性能, 因此合称为均衡局部模式。具体来说, ECLBP 利用了均衡的双圆交叉采样和以局部宏观信息为参考的自适应阈值在丰富采样信息的同时增强了编码的抗干扰能力和纹理描述能力, FLGS 方法通过精心设计的类似四角星的中心对称拓扑图结构, 均衡记录了水平与竖直多个方向的空间结构信息, 并且两种方法都设置了可调整的参数以灵活记录不同纹理信息。在利用均衡局部模式快速获得具有更强表征能力的局部特征后, 进一步将高斯核及图像差分生成的 DOG 金字塔与均衡局部模式方法融合。级联的多尺度特征图在扩充人脸样本同时将大尺度全局轮廓和小尺度局部细节融合, 更利于区分人脸的类间及类内差异, 在复杂的人脸识别环境中进一步增强特征的判别性, 并提高鲁棒性。最后利用信息熵加权, 串接各个特征图所有子块的直方图实现特征融合, 获得包含多方面信息的特征向量, 采用卡方距离实现相似性判别后, 利用最近邻分类器估计出人脸身份, 完成识别。

2. 相关算法分析

LBP 方法基于简单的邻域像素与中心像素的比较记录整体微纹纹理, LGS 方法利用图论知识记录局部图像的内部联系。两种方法各有不同, 但都可以快速的挖掘局部特征的关键信息, 有极大的研究价值。而 DOG 金字塔在扩充样本的同时补充多尺度信息, 弥补了单纯局部特征在描述人脸信息不够全面的局限性。由于所提算法以 LBP 类、LGS 类方法及 DOG 金字塔模型为基础改进实现特征增强, 因此本节进一步作了更加详细的介绍及相关的理论分析, 为后续的算法改进作理论铺垫。

2.1. LBP 类算法

在计算机视觉方向,局部二值模式(LBP)最初在纹理分类中被提出。由于其强大的纹理描述能力、对光照等变化良好的抗干扰性、极快的计算速度等一系列优点,这使得它很快被应用于指纹识别、光学字符识别、人脸识别及车牌识别等重要领域。

原始 LBP 算子的核心思路是:在一个正方形邻域范围,选取中心点的像素灰度值为阈值,将相邻点像素与阈值比较,如果附近点的像素比中心像素大就记为 1,否则记为 0。经过采样邻域点的遍历统计后,就可以得到 8 位的二进制编码串,转化为十进制后替代中心像素值作为该局部区域的 LBP 特征值。以 3×3 邻域为例的编码过程如图 1:

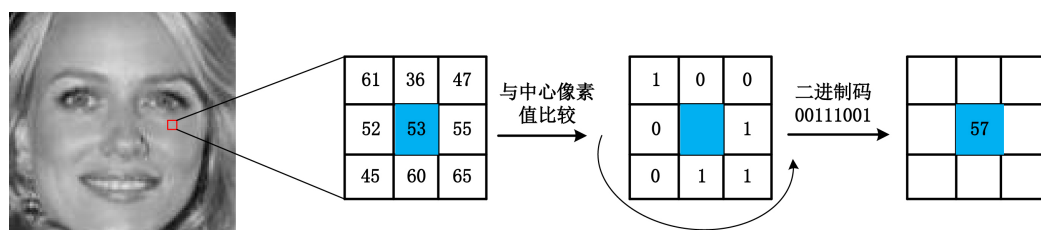


Figure 1. An example of the original LBP encoding

图 1. 原始 LBP 编码示例

需要注意的是,图 1 中的编码是以水平最左的像素为起点然后逆时针遍历得到,实际上生成 LBP 的顺序没有特定要求,只要在训练和匹配时保证处理的方式一致即可。通过邻域像素点与中心点的比较,LBP 根据它们的灰度关联描述了纹理信息。

从图 1 来看,它仅选择固定的 3×3 邻域范围来描述图像纹理,这种提取方式缺少灵活性,当面对大尺寸的高分辨率图像,利用小邻域的图像信息而产生的 LBP 编码不足以表征高分辨率图像更丰富的纹理细节。为了能够描述不同尺寸的纹理信息,Ojala 通过可变半径的圆形范围和可变采样像素数目的方式扩展了 LBP。具体的几种变体示例在图 2 中展现,其中符号 (N,R) 代表了以中心点为圆心、半径为 R 的圆形范围内均匀采样 N 个像素点:

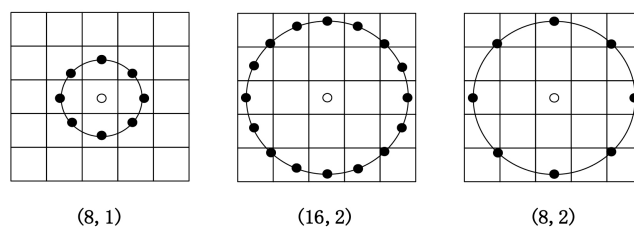


Figure 2. Variant LBP

图 2. 变体 LBP

获得变体 LBP 特征的计算公式为:

$$\text{LBP}_{(N,R)} = \sum_{j=0}^{N-1} \text{sign}(I_j - I_c) \times 2^j \quad (1)$$

$$\text{sign}(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

公式(1)和公式(2)中, I_c 是固定阈值, 即中心点的灰度值, I_j 是在以中心点为圆心, 半径为 R 的圆形邻域上采样点的灰度值, 如果采样点与真实像素点不完全匹配时, 则通过双线性插值方式拟合。 N 是采样点的个数, $LBP_{(N,R)}$ 是编码值, 由于采样点 N 和邻域半径 R 的可选性, 这种改进使得它在面对不同尺度的图像时, 描述纹理更加灵活。

另外, 参考图 3, 由于 LBP 以中心点灰度为固定阈值, 一旦中心点受到噪声等破坏, 二进制编码就可能发生改变, 这使得它对干扰极其敏感。当图像遇到噪声干扰或是图像本身质量较差时, LBP 编码不再能准确反映真实的纹理特征, 在匹配阶段会因为较大误差从而导致准确率的下降。为了尽可能减少这种影响, MBP 和 NI-LBP 方法分别选择了邻域内像素点的灰度中值和所有采样点的灰度均值作为参考阈值, 这两种编码方式与对图像做中值滤波和均值滤波类似, 进一步提高了特征的抗噪性。

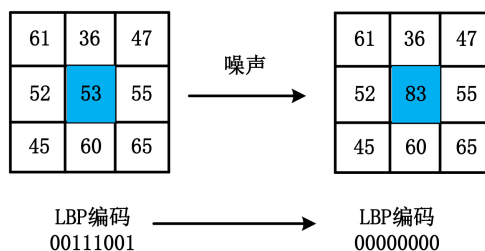


Figure 3. Noise interference causes changes in LBP encoding

图 3. 噪声干扰导致 LBP 编码的改变

2.2. LGS 类算法

与 LBP 通过单个点的参考阈值来产生编码的方式不同, LGS 根据图论直接利用像素点之间的关系来编码局部特征。LGS 背后的灵感在于: 人脸图像可以被认为是子图模式的组合, 这些子图模式相对于单调灰度变化是不变的。结合这些图案, 可以获得面部图像的整体描述。具体的实施过程在图 4 中展示。

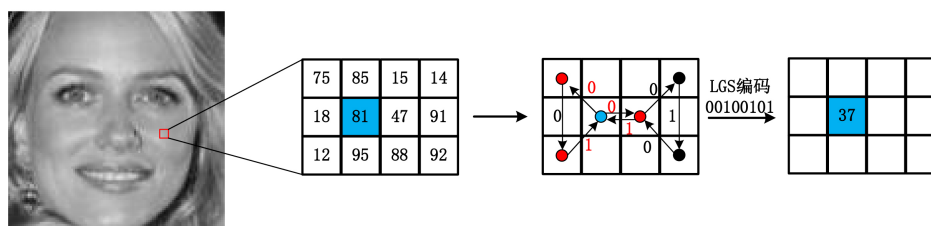


Figure 4. An example of LGS encoding

图 4. LGS 编码示例

如图 4 所示, 在选定目标像素后(图中蓝色像素, 值为 81), LGS 首先将在目标像素左侧进行逆时针的信息提取, 当相邻像素后一个的灰度值大于等于前一个时, 则会在连接两个相邻顶点的线段边缘赋 1, 否则赋 0。当左边子图模式的逆时针编码完成且回归至目标像素时, 会从目标像素继续以水平方式移至右侧子图区域进行顺时针的信息提取, 处理方式同左侧一致, 返回到目标像素后停止。通过这种在有向图路径上相邻像素值的比较, LGS 可以得到一个有序的 8 位二进制编码, 转换为十进制并替代目标像素灰度值后就可以用来提取 LGS 特征。要产生目标像素 (x_c, y_c) 的 LGS 编码, 计算方法如下式:

$$LGS(x_c, y_c) = \sum_{k=0}^7 \text{sign}(I_d - I_t) \times 2^k \quad (3)$$

式中, I_d 和 I_c 代表相邻两个像素点后一个点与前一个点的灰度值, $q = 7, 6, \dots, 0$ 代表加权序列的分配顺序。 $sign(x)$ 的计算方式同公式(2)一致。

从图 4 来看, LGS 不仅利用了目标像素点和相邻点的相对关系, 还根据目标像素的有向图结构描述了其路径上相邻点的空间联系, 这使得 LGS 算子通过两个三角形子图模式的信息组合记录了这个 3×4 窗口的整体特征。相对于 LBP 仅能捕获距目标像素一个像素点的邻域信息, LGS 利用距目标像素不同距离的分布更宽泛的采样点记录了更宽广的空间结构, 有更强的纹理描述能力。另外, 当某一点受到噪声干扰时, 仅可能改变它和它相邻点间的关系, 不可能导致所有采样点的编码同时发生改变。这种组合子图的提取方式使得 LGS 在设计之初就已具备更强的抗干扰能力。

尽管 LGS 拥有诸多优点, 但是它依然存在很多缺陷。第一, LGS 是左右不对称的, 其目标像素处于中心偏左, 如果去掉目标像素, 左边和右边子图的采样点数目分别为 2 和 3, 这使得它在提取特征时更偏向右侧信息, 这种不均衡的提取方式不利于整体信息的准确表征。由于其图结构路径中与目标像素发生联系的有 3 个点(图 4 中红色点), 一旦目标像素(图 4 蓝色点)受到干扰, 有 4 位二进制码(图 4 中红字编码)都可能因为目标像素被干扰被“间接”改变, 这会导致编码值发生较大改变, 产生错误的信息表征。第二, LGS 的编码有冗余, 计算复杂度高, 例如对于图 4 目标像素点(值为 81)和它水平向右的像素点(值为 47), 它们之间的编码必然是相对的, 一位编码已足以记录它们之间的关系, 这使得 LGS 编码出现了冗余。针对这些问题, ESLGS 算子被提出。

同 LGS 类似, ESLGS 从目标像素(图 5 蓝点)出发, 首先向左逆时针提取子图信息, 等回到目标像素后再右移顺时针捕获右侧子图信息, 最后级联左右子图信息得到编码, 示意图在图 5 展示。与 LGS 相比, ESLGS 采用了对称 5×5 的邻域和对称的图结构, 用更均衡的方式捕获了更广泛的空间结构, 并进一步消除了冗余编码。当目标像素点(图 5 蓝点)被干扰时, 仅会影响到附近值为 18、47 的 2 个像素点(图 5 红色点), 仅有 2 位二进制码(图 5 红色数字)可能发生改变, 受到影响的码字仅为 LGS 的一半。另外, 由于它更宽且对称的提取范围, 这使得: 更大的区域得到均衡, 更宽的空间结构得到表征[10]。显然, 对称的图形结构和邻域帮助 ESLGS 获得了更强的抗干扰能力和特征表达能力, 这样可以有效提升算法的识别精度。

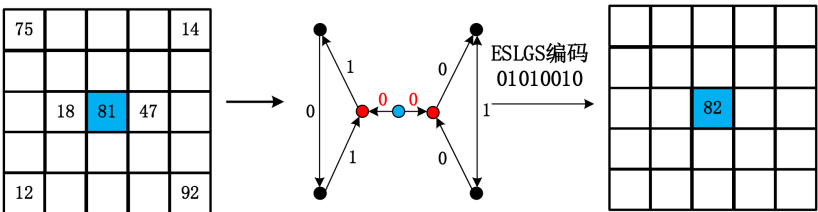


Figure 5. ESLGS encoding diagram
图 5. ESLGS 编码示意图

2.3. 高斯差分(DOG)金字塔模型

在图像处理中, 只有在一定尺度, 物体的图像结构才能被客观感知。并且, 在不同尺度下, 物体表现的特征也不尽相同, 在一种尺度下难以发现的特征在其他尺度下可能被轻松发掘。在人脸识别中, 大尺度和小尺度的图像分别反应了整体轮廓和局部细节的人脸特征, 如果在单一尺度下提取特征必然会损失其他尺度的信息。因此, 想要全面准确的描述人脸属性, 对图像的多尺度表达就显得尤为必要。

如图 6 所示, 图像金字塔是一种常见的尺度融合技术, 它通过一系列下采样(隔行隔列采样)和平滑滤波实现多尺度图像的信息融合, 目前已在模式识别、图像增强等多个计算机视觉领域应用。高斯差分(DOG)

金字塔作为一种经典的方法，它通过获取多分辨率、多尺度的差分图像，可以很好的融合多尺度信息，因此本文考虑将它与局部特征结合以全面表征人脸信息。

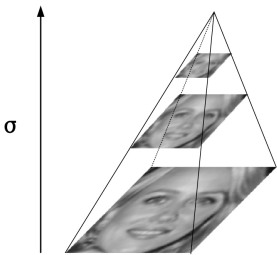


Figure 6. Image pyramid
图 6. 图像金字塔

DOG 金子塔在高斯金字塔基础上发展而来，思路是在获得一系列连续尺度的差分图像后综合获得多尺度空间下的图像特征。其构建过程如图 7 所示：

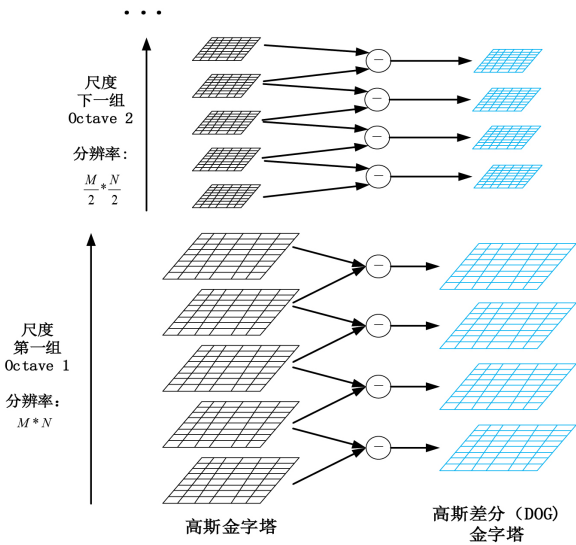


Figure 7. The construction process of the Gaussian difference pyramid
图 7. 高斯差分金字塔的构建过程

在建立高斯金字塔之前，通常将图像先扩大一倍，然后进行一系列高斯模糊处理以获得连续尺度变化的多尺度图像[13]。它包含若干组(Octave)，每一组的图像分辨率相同，由多层高斯模糊程度逐渐增强的图像形成，越上层的图像模糊越严重，尺度越大。图 8 展示了一组内各层图像尺度的变化，随着图像模糊的增大，越往后的图像细节越少。对图像进行高斯模糊的过程是将图像与高斯核作卷积以模拟尺度改变。假设原图为 $I(x,y)$ ，高斯函数是 $g(x,y,\sigma)$ ，那么计算方式为：

$$F(x,y,\sigma)=I(x,y)*g(x,y,\sigma) \tag{4}$$

$$g(x,y,\sigma)=\frac{1}{\sqrt{2\pi}\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5}$$

式中, σ 是尺度空间因子, 该值越大图像被模糊的就越严重, 尺度就越大。 $F(x, y, \sigma)$ 是被模糊后的图像。



Figure 8. Images with increased blurriness
图 8. 模糊程度增大的图像

对于同一组的图像, 主要变化在于模糊程度不同, 也就是 σ 的变化, 以第一组(Octave 1)的图像为例, σ 的变化为: $\sigma_0, k\sigma_0, k^2\sigma_0, k^3\sigma_0, k^4\sigma_0, \dots$, 相邻两张图像 σ 为 k 倍的关系, 通过对 k 的控制可以获得有不同间隔尺度的多尺度图像。为保证尺度的连续性, 下一组第一张图像(如图 7 中 Octave 2 最下面的图像)的 σ 为上一组第一张图像 σ 的 2 倍, 它由上一组的图像下采样得到。第 2 组图像 σ 的具体变化是: $2\sigma_0, 2k\sigma_0, 2k^2\sigma_0, 2k^3\sigma_0, 2k^4\sigma_0, \dots$ 。如果要获得更多不同尺度的图像, 进行类似的操作即可。通过在多个分辨率下对每一组图像使用不同参数获得多尺度图像后形成高斯金字塔。

建立好高斯金字塔后, 如图 7 所示, 将每一组内相邻尺度图像相减得到差分图像, 如 DOG 金字塔第一组第二层差分图由高斯金字塔的第一组第三层与第一组第二层作差取得。对于同一组内两张相邻图像, 假设分别为 $F_1(x, y, \sigma_1)$ 和 $F_2(x, y, \sigma_2)$, 那么差分图像的计算方式为:

$$\text{DOG}(x, y) = F_2(x, y, \sigma_2) - F_1(x, y, \sigma_1) \quad (6)$$

$$F_2(x, y, \sigma_2) = I(x, y) * g(x, y, \sigma_2) \quad (7)$$

$$F_1(x, y, \sigma_1) = I(x, y) * g(x, y, \sigma_1) \quad (8)$$

其中, $\text{DOG}(x, y)$ 是高斯差分图像, 它是对不同尺度的高斯模糊图像作差得到, 其生成方式在图 9 中展示。同样地, 对高斯金字塔所有组的相邻层图像进行类似操作后形成多个分辨率下的所有多尺度差分图像, 得到 DOG 金字塔。

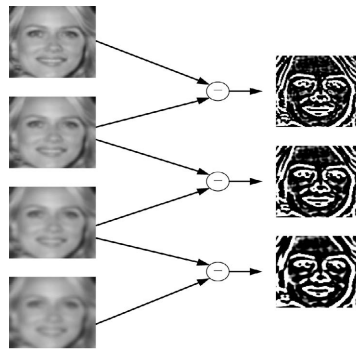


Figure 9. Example of Gaussian difference image generation
图 9. 高斯差分图像生成示例

为了直观的感受 DOG 金字塔不同图像的差异, 图 10 展示了 DOG 金字塔在不同尺度和不同分辨率条件下的 3 张高斯差分图像。从图 10 中来看, 小尺度的高分辨率图像细节更丰富, 这可以帮助我们区分人脸表情等更细腻的变化, 通过对高分辨率图像细节信息的利用可以帮助我们区分同一个人因表情、光照等导致变化的人脸图像, 可以捕捉类内差异。对于在大尺度下的低分辨率图像, 由于细节大量丢失和

分辨率的降低，我们能看到的仅仅是人脸轮廓，它反应了脸型的全局轮廓信息。有时候我们远远看到一个人可以迅速辨认出他们是谁，正是我们对这种整体轮廓的快速分析从而确定其人脸身份，因此整体轮廓使我们更加关注不同人的整体脸部差别，它可以帮助我们快速区分不同人的脸，可以捕捉类间差异。另外，高斯差分图像包含了不同尺度的差分特征，它展现出不同尺度的人脸信息间的差异，这些差异信息在间接扩大样本数量的同时，有效丰富了人脸信息，使得我们可以利用它来更加充分的描述人脸特征。综合来看，DOG 金字塔模型可以实现大尺度的全局轮廓和小尺度局部细节的融合，可以在补足更多信息量的同时帮助我们捕获类间差异和类内差异特征，进而提高算法的整体性能。



Figure 10. DOG pyramid
图 10. DOG 金字塔

3. 基于均衡局部模式 DOG 多尺度融合的人脸识别算法

3.1. 均衡局部模式

在对局部特征提取时，有两个非常关键的步骤：局部采样和模式编码。局部采样决定着采样信息的丰富程度，而模式编码则决定着从采样信息中得到的编码特征判别性及鲁棒性的强弱，两者相辅相成。LBP 在单个圆和固定领域提取特征，LGS 利用非对称的图结构偏向于提取水平方向特征，它们这种偏向性的提取方式使得它们的抗噪性能和描述能力都受到很大局限。从前文对它们的分析表明，在编码和采样之初尽量无偏向地均衡提取人脸信息，可以显著改善算法的抗干扰能力和信息表征能力。立足于这一点，我们综合优化了局部采样和模式编码过程，提出改进方法。具体而言，我们使用均衡的优化策略，分别针对经典的 LBP 与 LGS 方法作了改进，分别提出 ECLBP 和 FLGS。由于本文的启发思路是尽可能从均衡提取信息的角度来改善算法性能，因此将 ECLBP 与 FLGS 合称为：均衡局部模式。

3.1.1. 延展交叉型局部二值模式(ECLBP)

通过 2.1 节对 LBP 方法及其改进方法的分析发现，以固定阈值为参考，一旦有噪声等干扰，LBP 编码可能全部改变，这对于关键特征的稳定提取是极为不利的。另外，LBP 在固定半径的单个圆上采样无法像 LGS 一样获得更大空间结构的局部信息。针对这些局限，本文从局部采样和编码阈值两方面同时改进，提出 ECLBP 方法，以更均衡的方式有效捕获更关键的局部特征。编码过程如图 11 所示：

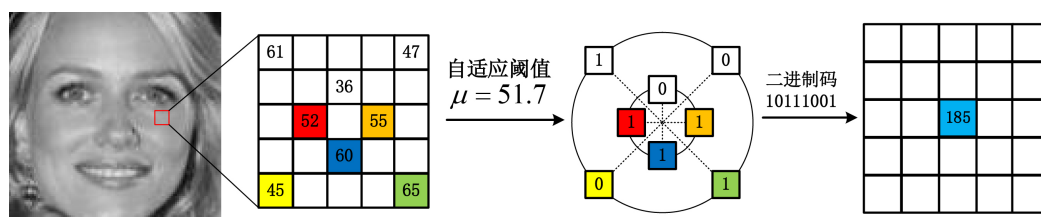


Figure 11. Example of ECLBP encoding process
图 11. ECLBP 编码过程示例

在得到 ECLBP 编码之前，首先需要计算自适应阈值。为了增强编码的多方面性能，ECLBP 利用局

部图像的整体宏观信息得到自适应阈值 μ :

$$\mu = \frac{(I_{R_1 1} + I_{R_1 3} + I_{R_1 5} + I_{R_1 7}) \times \frac{R_2}{R_1} + (I_{R_2 2} + I_{R_2 4} + I_{R_2 6} + I_{R_2 8})}{4 + 4 \times \frac{R_2}{R_1}}, R_1 < R_2 \quad (9)$$

当 $R_1 = 1$, $R_2 = 2\sqrt{2}$, 采样点空间分布如图 12 所示。式中, $I_{R_1 1}$, $I_{R_1 3}$, $I_{R_1 5}$, $I_{R_1 7}$, 是内圆(以中心点为圆心、半径为 R_1 的圆)上在 0 、 $\frac{\pi}{2}$ 、 π 和 $\frac{3\pi}{2}$ 四个方向上采样点的灰度值。而 $I_{R_2 2}$, $I_{R_2 4}$, $I_{R_2 6}$, $I_{R_2 8}$ 则是在外圆(更大半径 R_2 的圆)上在 $\frac{\pi}{4}$ 、 $\frac{3\pi}{4}$ 、 $\frac{5\pi}{4}$ 和 $\frac{7\pi}{4}$ 方向上点的灰度值。在计算 μ 时, ECLBP 根据两个圆采样点与中心点的相对关系分配权重, 由于内圆点距中心点更近, 因此设内圆上点权重为外圆与内圆半径之比, 外圆上点权重为 1。与像素值加权相乘后, 再除以所有点的权重之和得到自适应阈值。之后将采样点的灰度值与 μ 比较生成 ECLBP 编码:

$$\begin{aligned} \text{ECLBP} = & \text{sign}(I_{R_1 5} - \mu) \times 2^7 + \text{sign}(I_{R_2 6} - \mu) \times 2^6 + \text{sign}(I_{R_1 7} - \mu) \times 2^5 + \text{sign}(I_{R_2 8} - \mu) \times 2^4 \\ & + \text{sign}(I_{R_1 1} - \mu) \times 2^3 + \text{sign}(I_{R_2 2} - \mu) \times 2^2 + \text{sign}(I_{R_1 3} - \mu) \times 2^1 + \text{sign}(I_{R_2 4} - \mu) \times 2^0 \end{aligned} \quad (10)$$

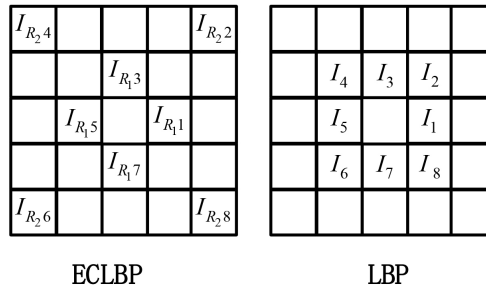


Figure 12. Sampling examples of ECLBP and LBP
图 12. ECLBP 与 LBP 的采样示例

结合图 11、图 12, ECLBP 以内圆最左侧点(如图 11 红点, 值为 52)为起点, 紧接遍历外圆上左下方点(图 11 黄点), 然后回到内圆遍历蓝点, 在逆时针方向不断交叉遍历内圆与外圆上所有采样点, 与自适应阈值比较后生成一串二进制码, 利用公式 10 转为十进制后得到 ECLBP 编码值。

参考图 3, 当中心点破坏, LBP 编码因中心点阈值的变化发生较大改变, 由于 LBP 仅以单点信息参考, 它对中心点波动极其敏感, 容易导致错误的特征编码。对此, 本文的 ECLBP 以邻域内所有采样点的宏观信息为参考并均衡分配权重得到鲁棒性更强的自适应阈值, 当某个采样点因干扰改变, 其阈值仅会发生微小变化, 干扰点距中心点越远, 这种影响则会越小。ECLBP 参考了邻域的整体信息, 因此阈值不会因为单个点的破坏而发生较大改变, 在生成编码值过程中(参考式 10), 由于它以 μ 为参考生成编码, 因此当 μ 不发生改变或变化很小时, 相比 LBP, 其编码值改变的可能性就会大大降低, 这使得它面对干扰有更强的鲁棒性。

与此同时, 由于 LBP 和 ECLBP 在阈值选择的区别, 这也使得它们在编码图像纹理时存在较大差别, 为了说明这一点, 图 13 展示了两种方法在面对不同纹理时的编码差异。图 13 中, (a)、(c)代表了表面纹理灰度变化平缓的一类图像, 这类图像通常不在关键的器官区域附近, 判别信息较少。而(b)、(d)则代表了表面纹理灰度变化剧烈的一类图像, 这类图像通常是人脸器官的边缘轮廓图像, 包含大量核心细节。

尽管(a)和(b)表面纹理差异甚大, LBP 仍然利用相同的编码错误的将不同纹理同等对待, 判别能力较弱。由于 LBP 仅考虑单点信息选择阈值, 未以整体邻域的联合分布作为参考, 这使得它在编码不同纹理图像时区分性较差, 可能导致错误匹配。而本文提出的 ECLBP 以邻域所有采样点的联合信息为参考, 其阈值伴随整体信息变化自适应改变[14]。显然, 在面对不同图像时, ECLBP 拥有区分性的编码, 可以捕捉不同图像纹理的区别, 这使得它拥有更强的判别性和纹理描述能力。

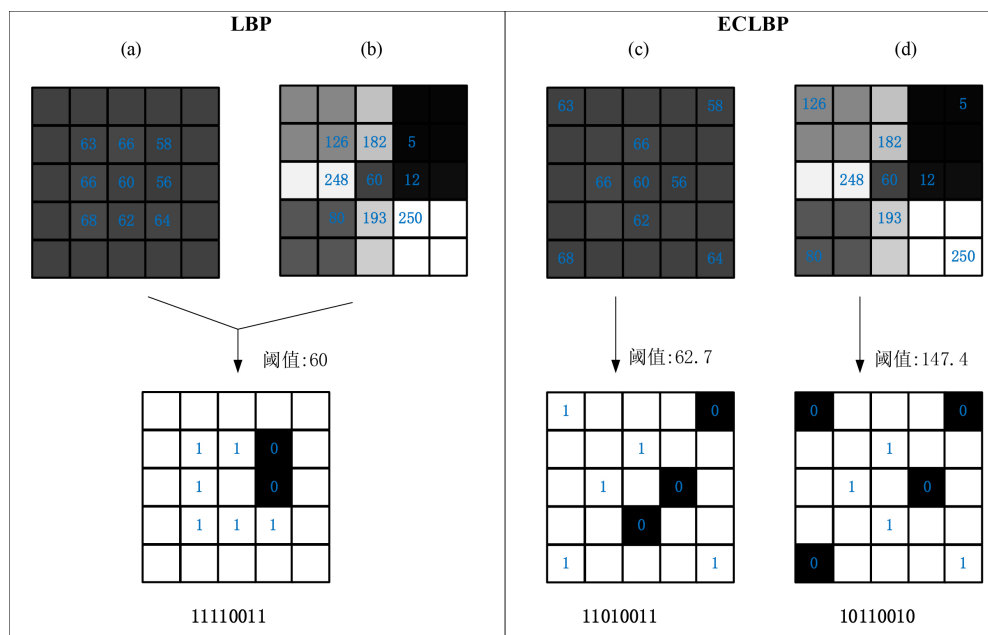


Figure 13. Encoding differences of ECLBP and LBP under different textures

图 13. 不同纹理时 LBP 与 ECLBP 的编码差异

另外, 从采样上来看, 尽管图 2 中 LBP 的变体方法可以通过调整圆形邻域半径及采样点来描述不同尺度的人脸图像, 但是通常而言, 在一个有限的邻域范围内图像的灰度分布是关联密切的, 部分采样点有效信息很少, 过多的在单个圆上捕获采样信息并不能保证有效信息量的大幅提升[15]。基于这一点, 对比图 11 与图 2 的采样方式, 可以看到: 不同于 LBP 方法仅利用单个圆的采样信息, 本文的 ECLBP 将 LBP 的圆形邻域进行了延伸, 并在保留 8 个方向信息基础上同时捕获了两个圆的采样点信息, 并且控制采样点在不同圆上尽可能均匀交叉分散, 以捕获更丰富的有效人脸特征。由于 ECLBP 均衡优化了采样点的分布及数目, 控制 ECLBP 的码长与 LBP 相同(都为 8), 同时利用 R_1 和 R_2 实现了可变尺度的采样。多方面的改进使得所提的 ECLBP 方法在不增加特征维数情况下可均衡记录更丰富的局部信息, 并可通过可变参数仍保证编码的灵活性。

3.1.2. 四星型局部图形结构(FLGS)

LGS 方法利用有向图表征图像更宽泛的空间结构关系, 但非对称的图结构和仅在水平方向提取信息的方式不利于其抗干扰能力和纹理描述能力的进一步提高。通过分析发现, ESLGS 针对 LGS 的一些缺点, 通过重新设计的对称、宽广的图形结构帮助 ESLGS 取得了多方面的性能提升, 这表明更均衡的图形设计对于方法的成功至关重要。然而, LGS 及 ESLGS 依然缺失对竖直方向信息的提取, 由于实际上人脸纹理是多方向的, 这势必会导致部分人脸特征的丢失。这种偏重于水平信息的策略很可能造成信息的不均衡提取, 容易导致核心特征丢失。此外, 无论是 LGS 还是 ESLGS 其采样点都为固

定位置,不利于人脸特征的灵活编码。基于这些考虑,本文综合 LGS 类方法的优缺点提出 FLGS,利用中心对称的可调整四角星型图形结构均衡获得多方向纹理信息,充分捕捉人脸特征。具体的示例如图 14 所示。

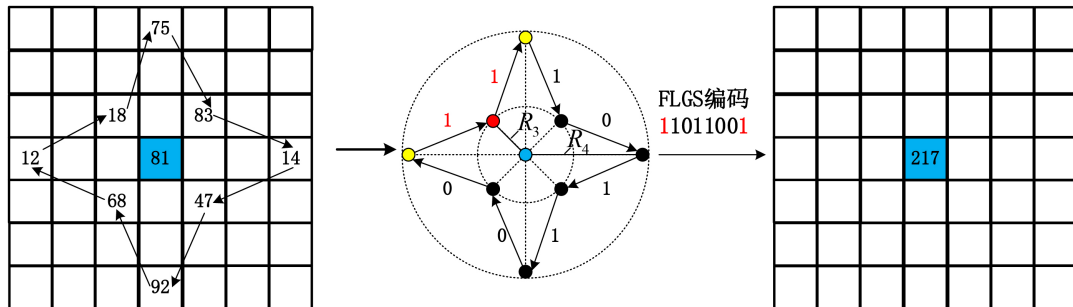


Figure 14. Example of FLGS encoding process
图 14. FLGS 编码过程示例

FLGS 根据目标像素(图 14 蓝点)的邻域信息,依据 8 个采样点同时捕获了水平、竖直多个方向的空间结构,由于其有向图路径与四角星形状极为相似,我们将方法命名为四星型局部图形结构以形象说明其特点。此外为了使 FLGS 可以灵活的编码不同空间结构,设置了两个可变参数:内圆半径 R_3 和外圆半径 R_4 ,这两个参数决定采样点的具体位置从而间接控制图形具体形状,图 14 展示了当 $R_3 = \sqrt{2}$, $R_4 = 3$ 时 FLGS 具体的采样点位置及编码过程。从图 14 可以看到:FLGS 以目标像素(图 14 蓝点,像素值 81)左上角采样点(图 14 红点,像素值为 18)为起点,然后顺时针沿四星型的图形路径遍历所有采样点,并记录路径上相邻采样点的关联信息,记录方式同 ESLGS 一致(图形路径后一个顶点与前一个顶点比较,若后一个点像素值大于前一个则记 1,否则记 0)。回到起点后则完全记录了形成四星型局部图结构的所有像素点之间关系,遍历完成后得到有序二进制编码,转换为十进制用以表征目标像素附近的关联结构,以提取人脸特征。

为了综合对比 FLGS 与 ESLGS,图 15 展示了两种方法的主要区别。从纹理描述上来看,从图 15 上看到,ESLGS 的采样区域可以看作是水平方向左右三角形子图的模式信息组合(图 15 中用框进行了标注),但它忽略了竖直方向的信息。而 FLGS 通过均衡的图形结构同时在水平方向与竖直方向捕捉纹理细节,其信息可以看作是目标像素对称的水平方向两个三角形子图与竖直方向两个三角形子图的信息融合。由于 FLGS 同时采集了在目标像素附近多个方向的特征并均匀提取了各个方向的空间信息,相比 ESLGS,其提取方式更为均衡,纹理表达更准确全面。另外,从两个编码的整体抗干扰能力来看。从 2.2 节对 ESLGS 的初步分析发现,当目标像素点受到干扰,因为有两个采样点与其相邻,因此有 2 位编码可能改变。类似的,可以分析不同采样点被破坏时编码的改变位数。为了直观的将 ESLGS 与 FLGS 进行整体的抗干扰分析,图 15 同时列出了当某个采样点干扰时,具体将有多少位编码发生改变(在图 15 中点附近用数字进行了标记),例如在 ESLGS 中,有三个点与目标像素左侧点相邻,因此当该点受到破坏将有三位编码可能改变。平均来看,ESLGS 平均干扰长度为 $(3 \times 2 + 2 \times 5) / 7$,即 2.29。而 FLGS 通过完全关于目标像素中心对称的图形路径均匀在水平与竖直方向采样图形信息,每个点仅可能有两个点相邻,因此不管哪一个点被破坏,都仅有两个编码可能改变,因此平均受干扰码字长度仅为 2。面对单点的干扰两者已有如此差距,可想而知,当面对更大的整体干扰时(如表情偏移、小范围遮挡等),两者将会有更明显的性能差距。综合上述分析,充分验证了 FLGS 更强的纹理描述能力及抗干扰能力。

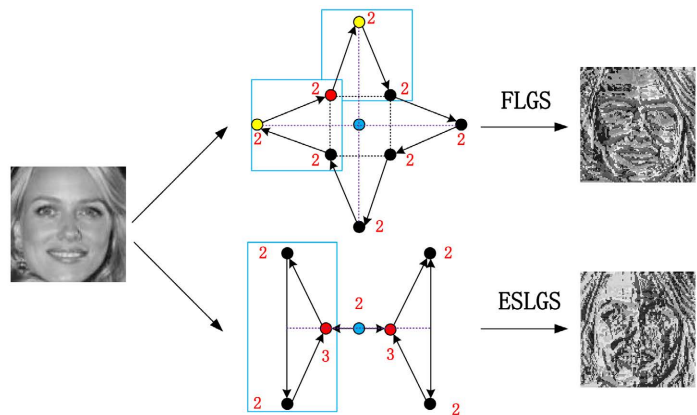


Figure 15. Comprehensive comparison of FLGS and ESLGS
图 15. FLGS 与 ESLGS 的综合对比

另外，从图 15 的特征图上来看，FLGS 有着更加精细的纹理刻画，而 ESLGS 在一些较为重要的区域已经开始出现边缘线不连续的问题，这是由于其采样过于偏向于水平方向信息导致的，部分纹理因此而被丢失。综合来看，无论是纹理描述、抗干扰能力还是特征图的表现，FLGS 都有着更加优秀的性能，都初步验证了经过改进的新算法 FLGS 有更强的识别能力和鲁棒性。

3.2. 基于均衡局部模式的 DOG 金字塔多尺度特征融合

为了补充不同尺度空间的人脸特征以全面描述人脸信息，本文进一步将均衡局部模式与 DOG 金字塔结合，通过直方图特征级联以融合多尺度特征。其中直方图参数 bin 代表统计时将灰度分布平分为多少个区间。由于灰度图像素值在[0, 255]区间分布，因此若 $\text{bin} = 256$ 时则代表将区间[0, 255]平分为 256 段，每一段区间跨度为 1。bin 数越大统计越精细，但维数也会增加，设置 bin 时应根据具体需要确定。另外考虑到金字塔层数越大，图像越多，提取时间与特征维数都会大幅增加。为了尽可能平衡好时间耗费与识别精度，通过初步实验，设 $\sigma_0 = 0.8$ ，选择 DOG 金字塔层数为 3，每一层包含一张差分图像，以在保留多尺度特征时减少时间耗费。

具体的融合步骤如图 16 所示(图 16 均衡局部模式以 FLGS 为例)：

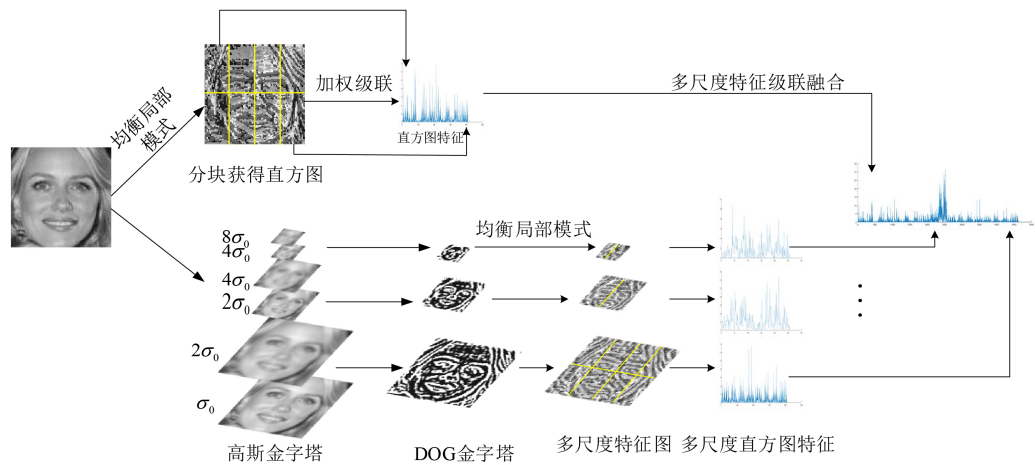


Figure 16. Comprehensive comparison of FLGS and ESLGS
图 16. 多尺度特征融合过程

(1) 首先对灰度图利用均衡局部模式(FLGS 或 ECLBP)进行特征提取获得原人脸的编码特征图,分块统计后获得各子块直方图, bin 设为 256。由于信息熵越大代表图像纹理信息越丰富,因此根据各子块信息熵与所有子块信息熵之和的比值分配权重。加权级联所有子块的直方图,获得原人脸直方图特征。

(2) 依据原始人脸图获得 3 层的高斯金字塔图像,第一、二、三层图像分辨率分别为原图的 1、1/4 和 1/8,每一层包含两张不同尺度的图像。根据初步实验,设同一层两张人脸 σ 之比为 2。具体而言,第一层的两张图像 σ 为 σ_0 和 $2\sigma_0$,为保证尺度连续性,第二层的第一张图像由上一层最后一张图像降采样得到,且第二层第一张图像与上一层最后一张图像的 σ 相等,即 $2\sigma_0$,第三层依次类推。对高斯金字塔每一层的两张相邻图像作差分后取得 3 张 DOG 图,然后利用均衡局部模式获得多个尺度的特征图,分块统计后得到多尺度直方图特征。在分块时,为减少特征维数并保留重要信息,第一层图像分块方式同步步骤(1)对原人脸图的处理方式一致,第二、三层水平与竖直方向分块数都为第一层的 1/2,第一层到第三层的 bin 设置时随数据库环境灵活调整。

(3) 最后将原灰度图的直方图特征与利用 DOG 金字塔得到的多尺度直方图特征串接级联,获得完整的人脸特征向量

从融合过程来看,由于 DOG 金字塔越上层的图像模糊越严重且分辨率越低,此时保留的为全局轮廓特征,它是在大尺度空间对人脸的粗略描述。而高分辨率特征图及原始人脸特征图人脸细节纹理清晰,保留了丰富的局部细节,是在小尺度空间对人脸的精细表征。通过补充的金字塔特征图像既扩大了样本量,且极大丰富了有效的人脸信息。相比于只获得原图的特征图,本文算法通过 DOG 融合获得了多尺度的人脸表征,实现了全局轮廓与局部细节的特征融合,可以更加全面的对人脸进行多方位描述。

3.3. 匹配分类

得到人脸特征后,考虑到卡方距离充分考虑直方图间的相对差异,可以较好辨别两张人脸,因此本文选择卡方距离来计算两张人脸的相似度,计算如下:

$$r_j(V, F_j) = \sum_{i=1}^N \frac{(V(i) - F_j(i))^2}{V(i) + F_j(i)} \quad (11)$$

式中, V 和 F_j 代表待识别测试人脸特征向量和样本库某一样本的人脸特征向量,它们的向量长度为 N 。 $r_j(V, F_j)$ 代表卡方距离,距离越小说明两个人脸相似度越高。

根据公式(11)确定待测人脸与训练样本库所有人脸相似度后,根据最近邻分类器,将与测试样本最相似的训练样本类别标签赋给测试样本,确定测试样本的人脸身份后完成分类。

3.4. 整体识别流程

整体来看,基于均衡局部模式 DOG 多尺度融合的人脸识别算法主要分为如下几步:

- 1) 输入: 测试样本人脸图片 y , 训练样本库 $T = [T_1, T_2, \dots, T_n]$ 。
- 2) 对 y 及训练样本库 T 的每一张样本 T_j , 利用均衡局部模式及 DOG 金字塔, 根据图 16 多尺度融合过程分别得到它们的人脸特征向量 V 和 F_j 。
- 3) 利用 V 和 F_j , 根据公式 11 得到测试样本与训练库每一张样本的相似度 $r_j(V, F_j)$ 。
- 4) 输出: 测试样本身份信息

$$label(y) = \arg \min_j r_j(V, F_j) \quad (12)$$

4. 初步理论分析

在设计特征提取方法时,通过实际数据库仿真,验证所提方法的实际性能是必不可少的。但同样的,

特征图、直方图、信息熵量化指标等理论结果都可以辅佐研究者预测方法的特征提取效果，可以更快初步衡量方法的理论性能。因此，本小节将独立于数据库仿真，从主客观两方面来初步验证所提方法效果。一方面，利用特征图与直方图的可视化结果可帮助我们主观层面，视觉上感知提取的人脸特征。另一方面，利用信息熵、卡方距离等一系列量化数据，我们可以通过它们的相对关系并结合可视化结果进一步客观预测方法性能。

4.1. 特征图、直方图主观对比分析

针对同一输入人脸，不同的特征提取方法，其特征图纹理与直方图分布往往各不相同。特征图人脸轮廓的清晰程度、在眼鼻嘴等关键器官区域纹理细节的丰富程度，往往反应出所提方法性能的优劣。而从特征图提取的统计直方图不再直接显示编码特征，它通过计算各灰度级灰度频数或频率来统计图像的总灰度分布。当图像过暗、过亮、对比度差时，其直方图通常集中分布于某一范围。而高质量、对比度好的图像，直方图数据在所有灰度级上尽可能平均分布，灰度分布更加均匀。在拍摄时，数码摄影者可以借助它来分析图像是否存在过曝、欠曝、对比度低等一系列问题并在后期处理时针对性地优化。因此，通过特征图与直方图的可视化结果可以帮助我们在视觉上直观了解不同方法的特征提取效果并进而分析各方法优劣。

为了直观展现各方法所获特征的区别，本节从无约束人脸库 LFW (Labeled Faces in the Wild) 中选择了娜奥米·沃茨(Naomi Watts)的一张清晰彩色人脸图作为输入，灰度化后通过 LBP、NI-LBP、ECLBP、DOG-ECLBP、LGS、SLGS、ESLGS、FLGS、DOG-FLGS 等一系列方法进行了特征编码并生成特征图，其中 DOG-ECLBP 和 DOG-FLGS 分别代表融合 DOG 金字塔的 ECLBP 和 FLGS 方法，统计各特征图灰度分布后生成直方图。图 17 详细的展示了可视化结果。

从图 17 中，我们发现，仅对原图灰度化，其直方图偏于右侧，这说明它很可能在高光环境下拍摄，存在部分区域过曝的问题，对比度较差。总体观察，经过 LBP、LGS、FLGS 等局部特征提取方法处理后，相比于未进行任何处理的灰度图，特征图凸显了眼鼻嘴附近关键图像的纹理特征，直方图分布明显更加均匀，大大改善了图像灰度的总体分布。结合特征图与直方图，这也可以帮助研究者理解局部特征提取方法为什么可以在不进行任何预处理的情况下来提取判别特征，提高识别性能。

进一步再仔细分析各个方法。ECLBP 不同于 LBP 和 NI-LBP 方法，它同时利用了多个圆上交叉采样点的灰度信息，这使得它以相同的编码长度提取出有更强表达能力的人脸特征。观察图 17，ECLBP 的特征图与直方图也验证了这一点：其特征图重点凸显了眼睛、鼻子、嘴巴核心区域的人脸特征，且拥有更丰富的图像细节和更均衡的直方图分布，视觉上的特征效果显著增强。LBP 类方法通过选择阈值提取特征，LGS、SLGS、ESLGS、FLGS 等局部图结构方法不再拘束于这种策略，它们利用相邻像素的有向图结构来捕获空间微纹理。但是 LGS、SLGS、ESLGS 方法更注重水平方向的空间图结构，这会导致其丢失了较多的竖直方向信息。而人脸中，眼睛，鼻子，嘴巴等器官上面不仅仅只有水平方向的差异，在竖直方向上也有差异，这样就会影响捕捉到的轮廓的完整性。从图中结果来看，这些方法的特征图虽然保留了基本轮廓，但在部分区域(例如在关键的眼、鼻、嘴器官周围)出现了轮廓纹理不连续和边缘细节丢失的问题。相比之下，FLGS 方法经过重新设计的四星型拓扑采样同时捕获了水平和竖直方向的空间结构，其可视化结果更佳：它的特征图拥有极为连续且完整的轮廓，同时，图像纹理细节更丰富，直方图分布也要均匀的多；视觉上看，改进后的 FLGS 和 ECLBP 方法无论是特征图还是直方图都有着更优秀的可视化效果。结合图 17，进一步观察融合高斯差分金字塔的 DOG-ECLBP 和 DOG-FLGS 方法，可以看到，它们补充了多尺度多分辨率的人脸特征：高分辨率特征图丰富了局部细节，而经过下采样后的低分辨率特征图则进一步补充了整体轮廓的全局信息。相比于仅提取单张灰度图特征的方法，融合金字塔的特征

提取方法通过扩充的多尺度特征图，融合了全局与局部信息，提取出更丰富有效的人脸特征，进一步增强了算法的整体性能。

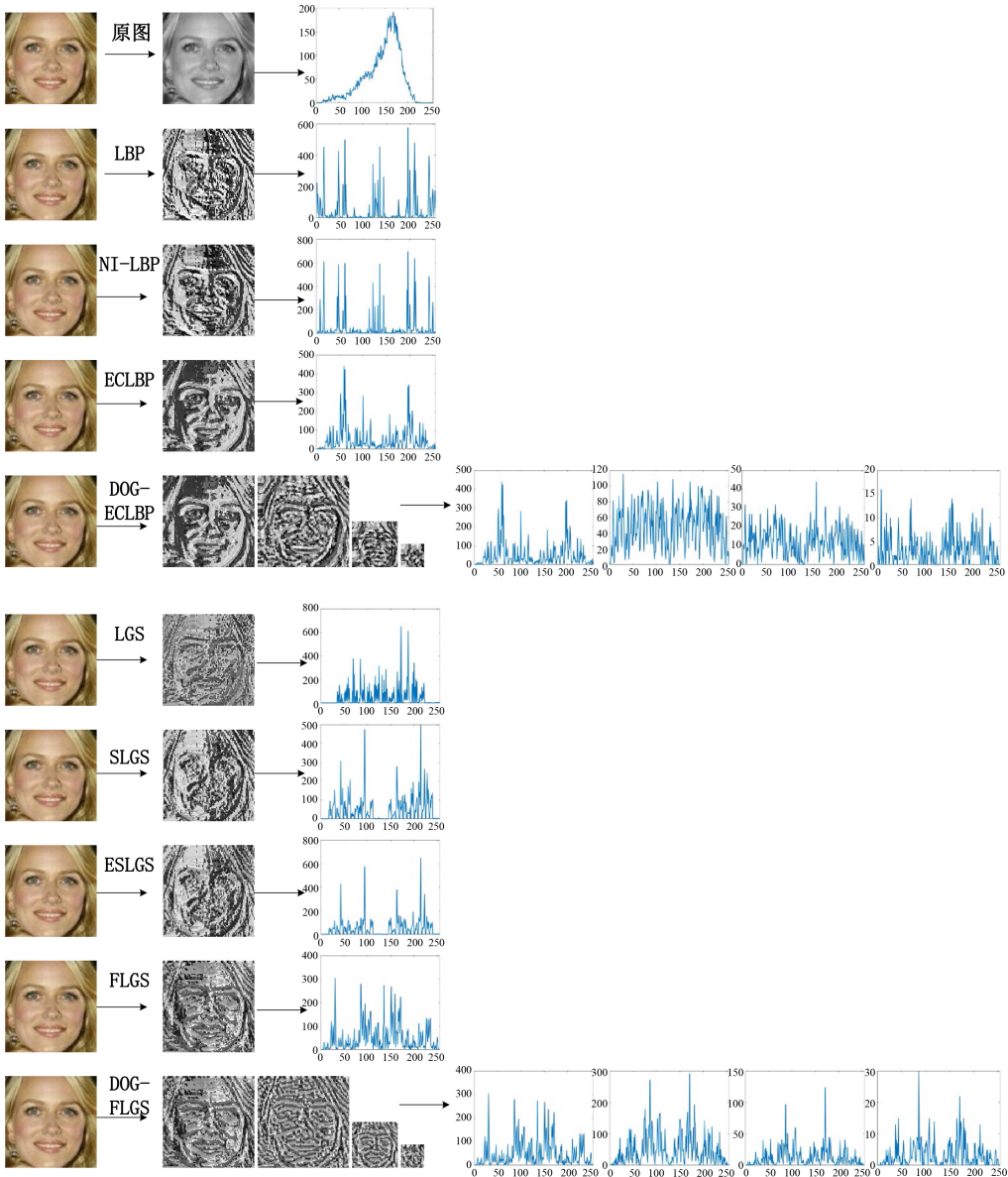


Figure 17. Visualization results of feature maps and histograms
图 17. 特征图与直方图可视化结果

4.2. 客观量化数据评估

为了使客观评估尽可能准确，我们从特征提取、特征分类两方面，分别选择了信息熵、卡方距离来评估方法性能。在特征提取方面，信息熵量化了图像所包含的平均信息量，可衡量总的信息保留特性[16]。在局部特征提取中，特征图信息熵越大往往代表图像纹理越复杂，细节越丰富。利用信息熵可以度量特征图从而客观评估特征提取的有效性；而在特征分类方面，特征图与原图直方图间的卡方距离代表了特征图与原图的相似程度[17]。值越小，相似性程度越高，那么直方图统计特征被正确分类的概率也就越大。

因此计算了图 17 中各方法特征图的信息熵、特征图与原灰度图直方图间的卡方距离来评估各个方法。由于信息熵越大，卡方距离越小，识别性能越好。所以我们根据它们的关系进一步量化了评估标准，得到评估系数 p ， p 越大说明方法越好，计算公式如下：

$$p = 0.5 * \frac{h}{10} + 0.5 * \frac{1}{d/10000} \tag{13}$$

式(13)对数据进行了转换处理，以尽可能使得两个指标在同一数量级，更适于量化评估。 h 为信息熵， d 为卡方距离， p 为评估系数。详细实验结果在表 1 展示。

Table 1. Objective data evaluation
表 1. 客观数据评估

方法\量化指标	信息熵(h)	卡方距离(d)	评估系数(p)
LBP	6.1737	17276	0.598
NI-LBP	5.9052	16165	0.605
ECLBP	7.0570	10580	0.825
LGS	6.2071	11731	0.737
SLGS	6.8110	12472	0.741
ESLGS	6.7298	11767	0.761
FLGS	7.1943	7426	1.033

从表 1 的实验结果来看，LBP 卡方距离最大且信息熵非常低，评估系数最小，说明其总体性能较差。总体来看，之后改进的方法大多都拥有更低的卡方距离和更大的信息熵，取得不同程度的性能提升。其中，NI-LBP、LGS、SLGS、ESLGS 的信息熵小于 7，卡方距离大于 11700，性能提升较少。而 FLGS 方法在卡方距离不到 LBP 卡方距离一半的同时，信息熵达到最大，评估系数远高于其他方法，表现出极大的优势。ECLBP 稍逊于 FLGS，但卡方距离也很小，信息熵超过 7，同样有着不错的提升效果。从客观度量结果来看，本文提出的 FLGS 和 ECLBP 两种均衡局部模式方法在理论数据上有着非常明显的性能提升，这与上一小节我们通过对特征图与直方图的主观分析得到的结果基本吻合。

5. 数据库仿真结果及分析

5.1. 仿真环境和参数设置

为了验证改进方法的有效性，在 ORL、AR 和 LFW 公开数据库上进行实验，这些公开数据库包含了光照、表情、遮挡、姿态、年龄变化等多方面实际生活中可能遇到的干扰。其中，在 AR 库上设置了两个实验以进行更细致的对比。选择和典型的算法如 LBP、NI-LBP、LGS、SLGS、ESLGS 等，进行识别性能及时间耗费的对比。实验所用硬件设备为 Intel I7-2600K, RAM 为 16GB，仿真环境为 Matlab R2016a。

5.1.1. 均衡局部模式相关参数设置

通过大量实验并结合实验结果，在实验过程中发现：对于 ECLBP，设置： $R_1 = 1$ ， $R_2 = 3\sqrt{2}$ 时，在 ORL、AR、LFW 数据库下都有不错的识别率；对于 FLGS，在 ORL 库下设置 $R_3 = 3\sqrt{2}$ ， $R_4 = 6$ ，在 AR 库实验一条件下，设置 $R_3 = \sqrt{2}$ ， $R_4 = 3$ ，AR 库实验二条件下，设置 $R_3 = 3\sqrt{2}$ ， $R_4 = 6$ ，LFW 库设置 $R_3 = 3\sqrt{2}$ ， $R_4 = 6$ 时，方法的识别效果较好。可以看到，在大部分情况下，当调整 FLGS 的采样点使得其分布更宽时，调整 ECLBP 方法使得采样点距中心点更近时，可以达到较好的识别效果。

5.1.2. 图像分块参数设置

由于在特征提取中，不同的分块方式会导致不同的识别效果。为了选择最佳分块，通常将分块数控制在一定范围进行对比实验，以找到每个数据库下最好的分块方式。通过大量实验发现，在 ORL 库下选择分块为 2×4 ，均衡局部模式两种方法的识别率最高；在 AR 库实验一中，在 AR 表情子集和 AR 墨镜子集下分块都设置为 12×5 ，在 AR 光照子集、AR 围巾子集下、AR 库实验二中分块都设置为 12×11 时，两种方法识别率最高；在 LFW 库，选择分块为 5×10 ，有最好的识别效果。

5.2. ORL 数据库实验结果及分析

ORL 数据库在不同的时间拍摄，光照、表情、姿态都有变化，头部旋转可达 20 度，人脸尺寸也有最多 10% 的变化。总共有 400 张图片，包含 40 个人，每人 10 张人脸样本，分辨率为 112×92 ，实验中选择将分辨率归一化到 96×96 ，部分样本如图 18 所示：



Figure 18. Partial samples from ORL database
图 18. ORL 库部分样本

为保证实验的准确性，实验中随机选择每个人的 k 张人脸图片作训练，剩余的图片作测试，取重复 10 次实验的平均值作为结果，详细的实验结果在表 2 中展示。

Table 2. ORL experimental results
表 2. ORL 实验结果

训练/测试	1/9	2/8	3/7	4/6	5/5	6/4
LBP	77.00	89.69	95.11	97.04	98.10	99.00
NI-LBP	80.31	90.16	95.17	97.5	98.55	98.75
ECLBP	82.06	91.92	96.76	98.32	99.00	99.75
LGS	77.78	90.47	95.25	97.29	98.40	99.31
SLGS	79.50	91.06	95.96	98.13	98.60	99.44
ESLGS	80.52	91.98	96.32	98.43	98.95	99.46
FLGS	83.5	92.52	96.43	98.63	99.00	99.46
DOG-ECLBP	87.28	94.59	98.63	99.29	99.38	100
DOG-FLGS	89.17	95.12	98.75	99.58	99.75	100

参考表 2，当训练样本数较多时，大部分方法在 ORL 库下表现不错，但在极少量样本训练时，大部分方法实验效果并不理想。仅 1 张训练样本时，LBP 方法和 LGS 方法的识别率仅有 77% 和 77.78%，识别率较低，ECLBP 方法利用双圆交叉像素点获得了更多信息，识别率相比 LBP 超过 5%。而相比 LGS

方法, FLGS 方法通过中心对称的四角星型拓扑采样图结构以等长编码同时获得了水平和竖直方向的结构信息, 识别率相比 LGS 提高超过 5.5%。进一步通过高斯差分金字塔与改进方法融合后, 利用金字塔实现了多尺度多分辨率图像的特征融合, 这使得低分辨率全局轮廓与高分辨率的局部细节实现融合, 获得的人脸特征拥有更强的判别能力。表 2 的结果验证了这一点, 1 张样本训练时, DOG-LBP 和 DOG-FLGS 相比于 LBP 方法和 LGS 方法提高均超过 10%, 提升效果明显。并且, 在训练样本数不相同的情况下, 融合金字塔的均衡局部模式方法始终都有着更高的识别率。

另外, 从表 2 看到, 随着训练样本数的增加, 各方法识别率逐渐提高。而融合金字塔的方法, 如 DOG-FLGS 和 DOG-ECLBP, 随着训练样本数增加, 其识别率更快的趋向最优。例如当训练样本图片为 3 张时, DOG-FLGS 就已经达到了 98.75% 的识别率, 而 FLGS 方法要保持同等识别率需要 4 张或 5 张训练样本。结合图 17 分析, 融合了 DOG 金字塔后, 在 1 张样本作训练时, DOG-FLGS 方法就可以获得 4 张特征图, 而 FLGS 方法仅获 1 张特征图。由于融合金字塔的方法扩充了虚拟样本, 补充了更多人脸信息, 因此随着样本增加, 相比于单一局部特征提取方法, 其特征信息的增长速度要快得多, 所以它们可以更快达到最优性能。

5.3. AR 数据库实验结果及分析

AR 数据库包含 126 个人的 4000 多张图片, 图片分辨率为 120×165 。每个人有 26 张图片, 在不同的时期采集。每个时期包含了 13 张图片, 分为表情、光照、墨镜遮挡和围巾遮挡四个子集, 拥有各种环境变化的情形, 它也是目前用于检验人脸识别最广泛的数据库之一。部分样本图片如图 19 所示。本小节选择了其中 50 名男性和 50 名女性的图片, 并设计了两组实验来验证在不同情况下所提算法的有效性。



Figure 19. Partial samples from AR database
图 19. AR 库部分样本

实验一: 为了精确测试在光照、表情、墨镜、围巾遮挡等变量单独干扰时, 各方法的实际性能。选择同一时期的第一张无遮挡正脸样本训练, 其余各子集的样本参与测试。表 3 展示了在面对各种变量单独干扰时各方法的仿真数据。图 20 的曲线图则直观的展示了各方法识别率的差异。

Table 3. AR single factor interference experimental results
表 3. AR 库单一因素干扰实验结果

算法	表情	光照	墨镜	围巾
LBP	96.33	92.33	91.67	85.67
NI-LBP	97.33	94.33	94.67	89.00
ECLBP	98.67	99	96.33	95.33

续表

LGS	97.33	94.67	93.67	89.67
SLGS	98.00	95.33	94.67	90.33
ESLGS	98.67	96.33	95.33	91.67
FLGS	99.33	98.67	96.67	95.67
DOG-ECLBP	100	100	99.33	98.00
DOG-FLGS	100	100	99.67	99.00

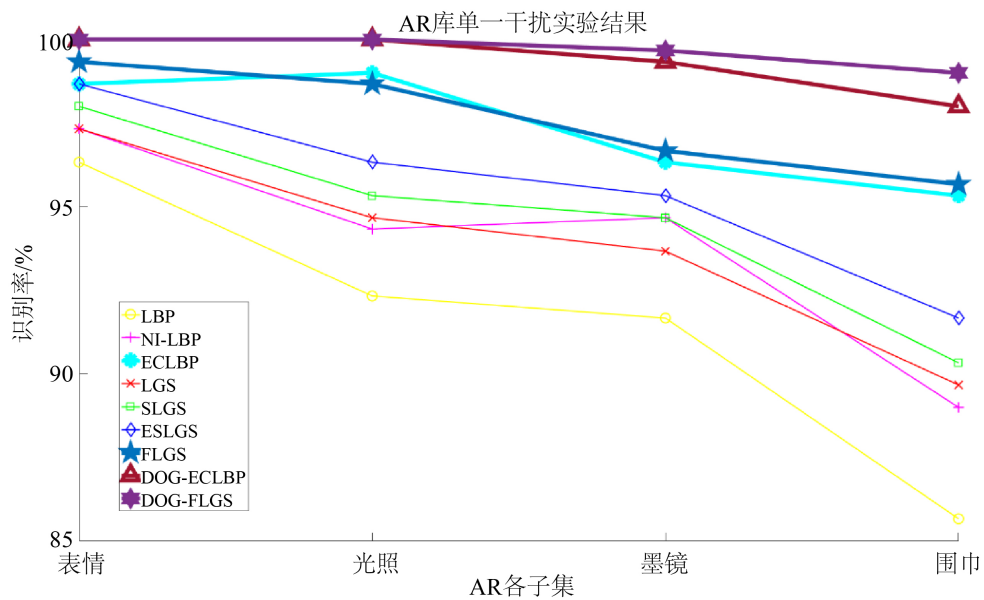


Figure 20. Experimental results of AR single interference

图 20. AR 库单一干扰的实验结果

在表 3、图 20 中可以看到，在墨镜、围巾子集下，大部分方法识别率有明显下降，这表明相比表情、光照干扰，墨镜、围巾遮挡对识别的影响要更加明显。在表情、光照等无遮挡子集下，尽管大部分方法已经有较好的性能，但 ECLBP、FLGS 仍进一步提升了识别率：ECLBP、FLGS 方法在表情子集的识别率分别达到 98.67% 和 99.33%，光照子集下，ECLBP 识别率超出 LBP 6.67%，FLGS 高出 LGS 约 4%，这验证了两种方法改进策略的有效性；在墨镜、围巾子集下，由于物品遮挡导致眼睛、嘴巴等关键图像细节丢失，LBP、ESLGS 等对比方法识别率明显下降，而 ECLBP、FLGS 的下降幅度远小于同类方法，并且，在围巾子集下，ECLBP、FLGS 方法识别率有明显提高：相比于 LBP 方法，ECLBP、FLGS 识别率提高均超过 9%。从仿真数据来看，无论是表情、光照的无遮挡干扰，还是墨镜、围巾的严重破坏，ECLBP、FLGS 方法识别率都要明显高于 LGS 等同类特征提取方法，这表明了均衡局部模式提取的人脸特征拥有更强的判别能力和良好的鲁棒性。

另外，尤其需要说明的是，通过 DOG 金字塔与均衡局部模式的融合，DOG-FLGS、DOG-ECLBP 在表情、光照子集下识别率达到了 100%，它们的曲线始终在最上方且基本趋于平稳，在围巾的严重干扰下，DOG-FLGS 方法的识别率也仅下降 1%。这表明通过金字塔与局部特征提取方法的融合，算法有着显著的性能提升并且对各种干扰有更强的鲁棒性。

实验二：为了进一步验证在混合光照、表情、遮挡、年龄等多重干扰后，所提方法是否有效。本实验根据 AR 库的样本特性，在不区分样本拍摄时间的情况下随机打乱样本顺序后，选择 26 张图片的任意 3、6、10、13 张参与训练，剩余图片参与测试。取 10 次实验的平均值作为结果。表 4 及图 21 统计了在不同比例下各方法的识别结果。

Table 4. Experimental results of mixed multiple interference
表 4. 混合多重干扰的实验结果

训练/测试	3/23	6/20	10/16	13/13
LBP	50.43	70.11	82.44	90.18
NI-LBP	51.36	71.35	84.19	90.46
ECLBP	57.88	77.89	89.13	93.68
LGS	51.54	70.30	84.18	89.66
SLGS	52.69	71.12	85.95	90.13
ESLGS	54.67	75.55	87.81	92.11
FLGS	61.80	80.44	91.15	95.29
DOG-ECLBP	68.52	85.40	94.18	96.58
DOG-FLGS	71.04	87.50	95.20	97.82

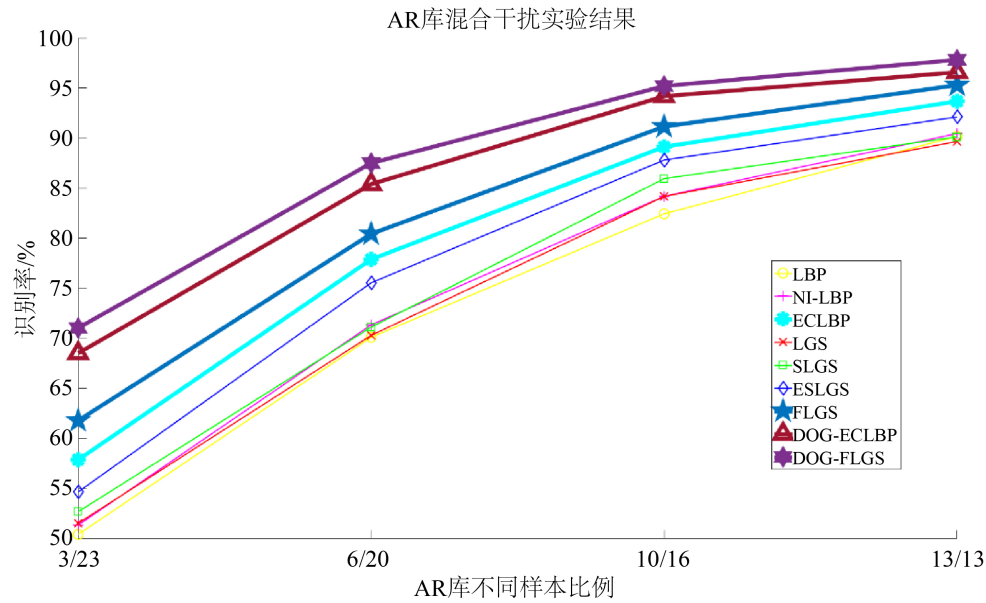


Figure 21. Experimental results of AR mixed multiple interference
图 21. AR 库混合多重干扰的实验结果

结合表 4 和图 21 发现，DOG-FLGS 方法优于所有对比方法，DOG-ECLBP 稍次，在训练样本占比较小小时，所提方法的性能提升更加显著。当训练样本数为 3 或 6 时，ECLBP 相比 LBP 识别率提高超过 7%，FLGS 相比 LGS 提升更是超出 10%，这表明了均衡局部模式提取的特征拥有更充分的人脸信息，在面对多重干扰时拥有更强的泛化性，这与我们通过初步理论分析得到的结果是一致的。融合 DOG 金字塔的多

尺度多分辨率特征后，在 3 张样本训练时，DOG-FLGS 的识别率相比于未补充多尺度特征的 FLGS 提升高达 9.24%，在 6 张训练样本数时，DOG-ECLBP、DOG-FLGS 方法已经达到 85.4%和 87.5%的识别率，相比于 LBP 和 LGS 识别率提升超过 15%。并且，随着训练样本数目的增长，融合金字塔的方法始终以更快的速度达到更高的识别率。从仿真结果来看，所提算法的识别性能有明显提升，验证了本文算法在均衡局部模式方法基础上利用金字塔实现多尺度特征融合的有效性。

5.4. LFW 数据库实验结果及分析

ORL 和 AR 库在受控的环境下获取样本，不能完全模拟真实的无约束场景。因此进一步选择 LFW 库来验证在无约束环境下方法的有效性。LFW(Labeled Faces in the Wild)数据库来源于美国的计算机视觉实验室，总共包含 5749 个人的 13000 多张图片。其照片大多来自互联网，取自于实际生活的无约束场景，存在极强的姿态、光照、表情、遮挡、年龄等变化，非常具有挑战性[18]。因此它也是目前常用于研究无约束环境下人脸识别的大型验证数据库。实验中我们对人脸图片进行了裁剪，部分样本如图 22 所示。实验中我们选择了部分样本数超过 20 张的人脸图片参与仿真，并设置训练样本数占总样本数之比为 0.3、0.5、0.7 进行对比，分辨率归一化为 100×100 ，实验结果见表 5 和图 23。



Figure 22. Partial samples from LFW database
图 22. LFW 库部分样本

Table 5. LFW experimental results and analysis
表 5. LFW 库实验结果

训练样本占比	0.3	0.5	0.7
LBP	51.70	59.42	62.88
NI-LBP	52.98	60.89	64.12
ECLBP	63.51	71.54	74.92
LGS	52.70	60.08	64.23
SLGS	55.00	61.92	67.27
ESLGS	60.16	67.90	71.99
FLGS	66.67	74.53	78.40
DOG-ECLBP	68.20	76.35	80.76
DOG-FLGS	71.07	79.40	82.45

从表 5 和图 23 中 LFW 不同比例下的仿真结果来看，本文的均衡局部模式方法 ECLBP、FLGS 在各种比例下都取得了明显的性能提升：在 0.3 比例下，ECLBP 识别率超出 LBP 约 11%，FLGS 超过 LGS

约 13%。在 0.7 比例下，相比于 LBP，性能最好的 FLGS 方法识别率提升高达 15.52%。在所有对比的方法中，DOG-FLGS 识别率优于其他算法，DOG-ECLBP 性能稍差于 DOG-FLGS，但两种融合金字塔的方法识别率都要明显高于未融合金字塔的局部特征提取方法。由于经过差分空间金字塔不同尺度的滤波图像作差，人脸的关键特征被保留，而且通过不同分辨率特征图的信息级联，高分辨率的细节纹理与低分辨率的全局轮廓信息实现特征融合，这可以更好的区分类内差异及类间差异，因此识别率进一步提高。另外，在 0.5 比例时 DOG-ECLBP、DOG-FLGS 的识别率就已经超越了 0.7 比例下的 ECLBP、FLGS 方法的识别率，这表明：金子塔的融合进一步弥补了单一局部特征提取方法因训练样本不足导致性能瓶颈的问题。补充的金字塔图像，扩充了多尺度样本，极大的丰富了人脸特征信息，使得算法的整体性能能进一步增强。

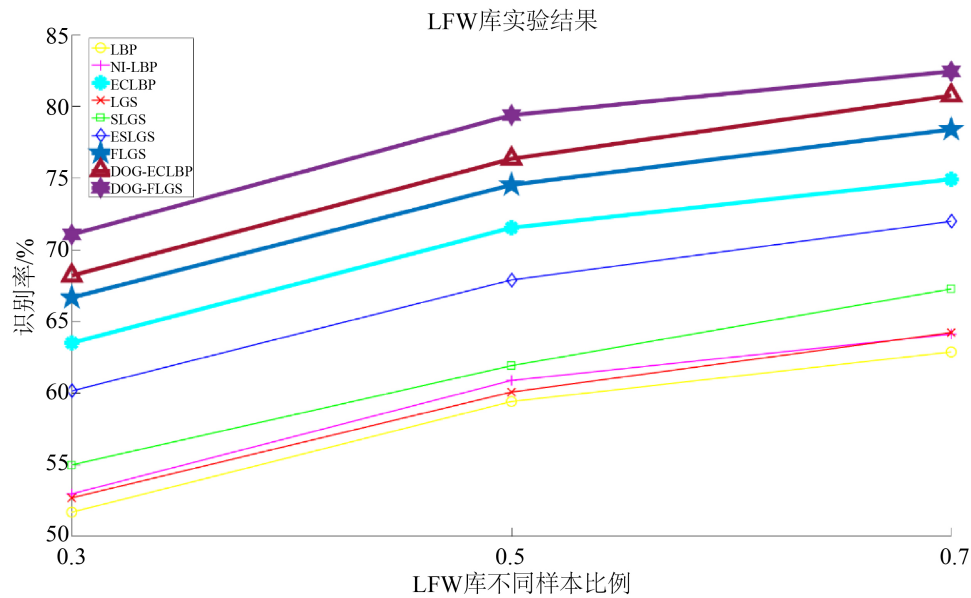


Figure 23. LFW experimental results curve
图 23. LFW 库实验结果曲线图

5.5. 时间复杂度分析

经过改进，算法的识别率上升了，那么这一小节将考虑改进后的算法在时间复杂度上的表现。为了比较各方法在时间消耗上的差异，本节在 ORL 库实验，选择以每个人的 1 张样本作训练，剩余 9 张样本作测试，统计了各方法的特征提取及特征匹配耗时结果。取 10 次实验的平均值作为最终结果，实验仿真数据见表 6。

Table 6. Time consumption comparison of different methods
表 6. 各个方法时间消耗对比

时间/s	库特征提取 总时间/s	单张样本特征 提取时间/s	库特征匹配 总时间/s	单张样本特征 匹配时间/s
LBP	3.70	0.009	0.21	0.0005
NI-LBP	3.79	0.010	0.21	0.0005
ECLBP	3.86	0.010	0.21	0.0005

续表

LGS	3.71	0.009	0.21	0.0005
SLGS	3.72	0.009	0.21	0.0005
ESLGS	3.74	0.009	0.21	0.0005
FLGS	3.76	0.009	0.21	0.0005
DOG-ECLBP	8.92	0.022	0.42	0.0011
DOG-FLGS	8.75	0.022	0.42	0.0011

结合本小节的仿真时间和前面各小节在数据库的仿真结果来看,在未融合金字塔特征时,均衡局部模式方法:ECLBP、FLGS,在特征提取和特征匹配上消耗的时间同LBP等对比方法基本一致,数量级上并没有差距,但是两种改进方法的识别率明显更高且鲁棒性也更强。这表明了改进方法在不增加时间复杂度的前提下显著增强了特征的表达能力,验证了改进策略的有效性。

从表6中结果可知,DOG-ECLBP耗时大约为ECLBP的两倍。这主要是由于融合金字塔后,增加了对多尺度多分辨率差分图像的特征提取,因此耗时有所增加。但换算到单张样本上,DOG-ECLBP的提取时间仅为0.022 s,相比于大部分特征提取方法而言(例如利用梯度特征的方法),其提取时间其实是很短的,依然符合人脸识别系统的实时性要求。综合来看,所提算法通过金字塔空间的融合以一定的时间损耗,换来了更强表达能力和更具有鲁棒性的人脸特征,这是完全可以接受的。

6. 结论

针对LBP、LGS类算法可快速提取特征但存在特征表达不强的问题,以及单纯使用局部模式存在描述人脸信息片面的局限性,由此提出一种基于均衡局部模式的DOG多尺度融合金字塔人脸识别算法。为了弥补LBP与LGS的缺点,算法均衡优化了局部采样及模式编码过程,提出可灵活记录信息的ECLBP和FLGS方法,合称为均衡局部模式。ECLBP引入的双圆交叉采样、自适应阈值和FLGS的均衡图结构,显著增强了对关键特征的均衡提取。为有效扩充样本信息,进一步融合的DOG金字塔多尺度信息,通过全局轮廓与局部细节的特征融合,进一步增强了特征性能。从特征图和直方图的可视化分析,以及信息熵、卡方距离等客观量化指标的评估,验证了所提方法在特征提取和分类性能上的优势,在保持计算复杂度不变的同时,显著提升了人脸识别的准确率,具有更好的实用价值。

参考文献

- [1] Talab, M.A., Qahraman, N.A., Aftan, M.M., Mohammed, A.H. and Ansari, M.D. (2022). Local Feature Methods Based Facial Recognition. 2022 *International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, 9-11 June 2022, 1-5. <https://doi.org/10.1109/hora55278.2022.9799910>
- [2] Malik, A.S., Boyko, O., Aktar, N. and Young, W.F. (2001) A Comparative Study of MR Imaging Profile of Titanium Pedicle Screws. *Acta Radiologica*, **42**, 291-293. <https://doi.org/10.1080/028418501127346846>
- [3] Shreedevi, P., Mohana, H.S. et al. (2021) An Improved Local Binary Pattern Algorithm. *Face Recognition Applications. 2021 IEEE Mysore Sub Section International Conference*, Hassan, 24-25 October 2021, 394-398.
- [4] Guo, Z., Zhang, L. and Zhang, D. (2010) Rotation Invariant Texture Classification Using LBP Variance (LBPV) with Global Matching. *Pattern Recognition*, **43**, 706-719. <https://doi.org/10.1016/j.patcog.2009.08.017>
- [5] Liu, L., Zhao, L., Long, Y., Kuang, G. and Fieguth, P. (2012) Extended Local Binary Patterns for Texture Classification. *Image and Vision Computing*, **30**, 86-99. <https://doi.org/10.1016/j.imavis.2012.01.001>
- [6] Yang, W., Wang, Z. and Zhang, B. (2016) Face Recognition Using Adaptive Local Ternary Patterns Method. *Neuro-computing*, **213**, 183-190. <https://doi.org/10.1016/j.neucom.2015.11.134>
- [7] Abusham, E.E.A. and Bashir, H.K. (2011) Face Recognition Using Local Graph Structure (LGS). In: Jacko, J.A., Ed.,

- Human-Computer Interaction: Interaction Techniques and Environments*, Springer, 169-175.
https://doi.org/10.1007/978-3-642-21605-3_19
- [8] Yunanto, A.A. and Herumurti, D. (2016). Face Recognition Based on Extended Symmetric Local Graph Structure. 2016 *International Conference on Information & Communication Technology and Systems*, Surabaya, 12 October 2016, 80-84. <https://doi.org/10.1109/icts.2016.7910277>
- [9] Rafiee, G., Dlay, S.S. and Woo, W.L. (2013) Region-of-Interest Extraction in Low Depth of Field Images Using Ensemble Clustering and Difference of Gaussian Approaches. *Pattern Recognition*, **46**, 2685-2699.
<https://doi.org/10.1016/j.patcog.2013.03.006>
- [10] 杨赛, 赵春霞, 刘凡, 陈峰. 一种基于多种特征融合人脸识别算法[J]. 计算机辅助设计与图形学学报, 2017, 29(9): 1667-1672.
- [11] Xie, Z. (2013) Single Sample Face Recognition Based on DCT and Local Gabor Binary Pattern Histogram. In: Huang, D.-S., Bevilacqua, V., Figueroa, J.C. and Premaratne, P., Eds., *Intelligent Computing Theories*, Springer, 435-442.
https://doi.org/10.1007/978-3-642-39479-9_52
- [12] Abdullah, M.F.A., Sayeed, M.S., Sonai Muthu, K., Bashier, H.K., Azman, A. and Ibrahim, S.Z. (2014) Face Recognition with Symmetric Local Graph Structure (SLGS). *Expert Systems with Applications*, **41**, 6131-6137.
<https://doi.org/10.1016/j.eswa.2014.04.006>
- [13] Rathgeb, C., Wagner, J. and Busch, C. (2018) Sift-Based Iris Recognition Revisited: Prerequisites, Advantages and Improvements. *Pattern Analysis and Applications*, **22**, 889-906. <https://doi.org/10.1007/s10044-018-0719-y>
- [14] Kas, M., El merabet, Y., Ruichek, Y. and Messoussi, R. (2018) Mixed Neighborhood Topology Cross Decoded Patterns for Image-Based Face Recognition. *Expert Systems with Applications*, **114**, 119-142.
<https://doi.org/10.1016/j.eswa.2018.07.035>
- [15] Guo, Y., Zhao, G. and Pietikäinen, M. (2012) Discriminative Features for Texture Description. *Pattern Recognition*, **45**, 3834-3843. <https://doi.org/10.1016/j.patcog.2012.04.003>
- [16] Roy, H. and Bhattacharjee, D. (2016) Local-Gravity-Face (LG-Face) for Illumination-Invariant and Heterogeneous Face Recognition. *IEEE Transactions on Information Forensics and Security*, **11**, 1412-1424.
<https://doi.org/10.1109/tifs.2016.2530043>
- [17] Chakraborty, S., Singh, S.K. and Chakraborty, P. (2016) Local Gradient Hexa Pattern: A Descriptor for Face Recognition and Retrieval.
- [18] Lu, J., Liong, V.E. and Zhou, J. (2018) Simultaneous Local Binary Feature Learning and Encoding for Homogeneous and Heterogeneous Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **40**, 1979-1993.
<https://doi.org/10.1109/tpami.2017.2737538>