

LGT[61]和 LLGP[128]这三种方法中, 模式图像的计算采用了如下的方式: 计算每个像素所在 Jet (LVC/LGT) 或块 (LLGP) 上的特征到相应码本中所有码字的距离, 选择距离最小的码字对应的索引赋值为当前像素 (参见第 4.2 节中的“表示阶段”)。

在这三种模式学习方法中, 构建码本的元素是不同的。图 4-5 (a) 和 (b) 分别示出了 LVC/LGT 和 LLGP 方法中的码本学习过程。具体地讲, 在 LVC/LGT 方法中, 码本是基于 Jet 学习得到的: 人脸图像被划分为不同的子块, 每个子块根据该子块中的 Jet 特征学习相应的码本, 如图 4-5 (a) 所示; 在 LLGP 方法中, 码本是基于图像块的且面向不同的 Gabor 小波, 即: 针对每个 Gabor 小波单独学习其对应的码本, 如图 4-5 (b) 所示。因此, LVC/LGT 方法中的码本数目等于子块的数目, 而 LLGP 方法中的码本数目则与采用的 Gabor 小波数目相同。

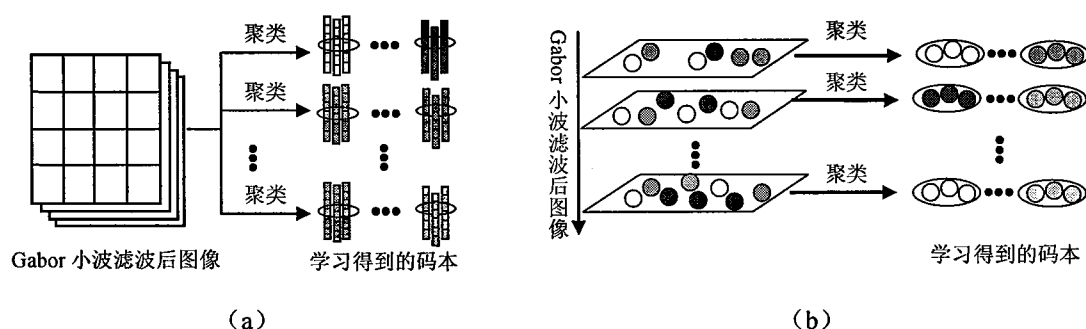


图 4-5 模式学习方法中的码本学习过程 (a) LVC/LGT (b) LLGP

从码本的学习过程来看, LVC/LGT 方法的码本包含了人脸特定区域 (即子块) 的局部特征, 而 LLGP 方法的码本包含了人脸图像特定 Gabor 小波上的局部特征。当采用分块直方图表示时, LLGP 表示反映了人脸图像上不同尺度、不同方向的特征分布, 通常情况下会取得比 LVC/LGT 方法具有更好的表示能力。

4.4.2 模式定义流程及参数对比

为了进一步地区分不同局部 Gabor 模式方法之间的差异, 图 4-6 示出了这 11 种方法的模式定义流程。如图所示, 这些方法的差异体现在 Gabor 特征 (即幅值、相位、实部和虚部)、模式定义域 (即块、Jet 和“体”) 及模式定义算子 (即 LBP、量化、XOR 和聚类) 这三个方面。如图 4-6 (a) 所示, LGBP_Mag 和 E-GV-LBP/V-LGBP 方法都基于幅值特征, 但利用 LBP 算子分别计算块和“体”上的模式; 如图 4-6 (b) 所示, GGPP 和 LGPP 都基于实部/虚部特征, 但前者是定义在 Jet 上并采用了量化算子, 而后者则定义在块上并利用了量化和异或算子; 虽然 LGBP_Pha、LGPDP 和 LGXP 都基于相位特征并在块上计算局部模式, 但 LGBP_Pha 方法利用了 LBP 算子, LGPDP 方法利用了差分 and 量化算子, LGXP 方法则利用了量化和异或算子, 如图 4-6 (c) 所示; 如图 4-6 (d) 所示, LVC/LGT 和 LLGP 方法都属于模式学习方法, 但前者的模式定义在 Jet 上, 后者

的模式则定义在块上。

这 11 种局部 Gabor 模式的共同之处在于，它们都采用了分块的直方图表示，即：每幅模式图像被划分为若干个子块并计算每个子块的直方图，最后将不同尺度、不同方向上所有子块的直方图连接起来表示人脸图像。这里提到“子块”的原因是区分判别性特征提取方法中块和子块的概念，详情请参考第二章 2.4.1 节。

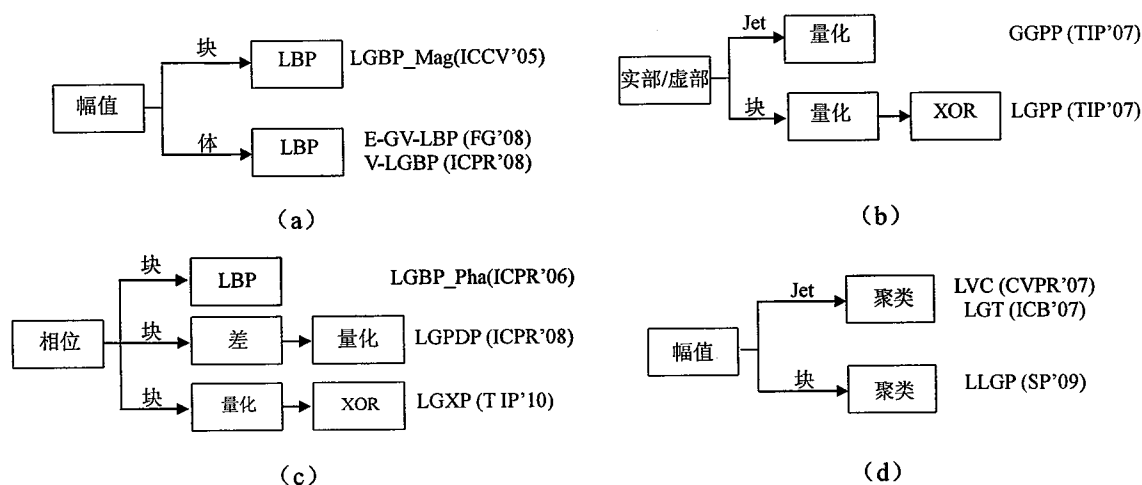


图 4-6 不同局部 Gabor 模式表示的模式定义流程
(a) ~ (c) 分别为 Gabor 幅值、实部/虚部、相位的局部模式 (d) 模式学习方法

表 4-1 不同局部 Gabor 模式表示方法的参数对比

注：P 表示块内邻域像素点数目；L 表示 E-GV-LBP/V-LGBP 方法中 Jet 上邻域大小；O 表示 GGPP 方法中 Gabor 小波的方向数目；b 表示 LGXP 方法中相位角的量化区间；C 表示 LVC/LGT 和 LLGP 方法中的码本大小。

方法分类	Gabor 特征	方法	模式定义的参数	模式数目	子块数目	模式图像数目	特征维数
模式设计方法	幅值	LGBP_Mag	P	2^P	m	40	$2^P \times m \times 40$
		E-GV-LBP/V-LGBP	P, L	2^{P+L}	m	40	$2^{P+L} \times m \times 40$
	实部/虚部	GGPP	O	2^O	m	10	$2^O \times m \times 10$
		LGPP	P	2^P	m	80	$2^P \times m \times 80$
	相位	LGBP_Pha	P	2^P	m	40	$2^P \times m \times 40$
		LGPDP	P	2^P	m	40	$2^P \times m \times 40$
		LGXP	P, b	2^P	m	40	$2^P \times m \times 40$
模式学习方法	幅值	LVC/LGT	C	C	m	1	$C \times m$
		LLGP	C	C	m	12	$C \times m \times 12$

表 4-1 进一步列出了这 11 种方法在直方图表示下的参数对比。可以看到，在模式设计方法中，模式数目通常是定义模式参数（如 P, L 和 O）的指数次幂。虽然较多数目的模式编码了图像上更多的微结构，但这也意味着更高的特征维数（表 4-1 的最右边一列）以及由此产生的比较大的相似度计算代价；同时，由于子块的大小固定，在模式数目增加时直方图会变得更稀疏，不利于图像之间的特征匹配。因此，模式数目通常不

能设置得太大（例如，不超过 64）。对于模式学习方法而言，码本的大小即是模式的数目，其数值同样也不能太大（例如，LVC 方法中码本大小设置为 64[155]）。

从表 4-1 可以看到，这些方法产生的模式图像的数目是不同的。当采用了 40 个 Gabor 小波（即 5 个尺度 8 个方向）时，LGBP_Mag、E-GV-LBP/V-LGBP、LGBP_Pha、LGPD 和 LGXP 等 6 种方法的模式图像数目都是 40 幅，即：针对每个 Gabor 小波产生一幅对应的模式图像；在 GGPP 方法中，根据实部和虚部特征分别得到 5 幅（即尺度的个数）模式图像，从而一共得到 10 幅模式图像；在 LGPP 方法中，根据实部和虚部特征分别得到 40 幅（即 Gabor 小波的个数）模式图像，最后一共得到 80 幅模式图像；在 LVC/LGT 方法中，属于同一个块的 Jet 学习得到一个相应的码本，并将每个块映射为相应模式图像的块，因此，最后只得到 1 幅模式图像；在 LLGP 方法中，每个 Gabor 小波对应了 1 个码本，因而，模式图像的数目等于采用的 Gabor 小波的个数。由于采用了 3 个尺度 4 个方向的 Gabor 小波（参见第 4.2 节），LLGP 方法中得到 12 幅模式图像。

对大多数局部 Gabor 模式方法来讲，它们高维的特征（表 4-1 的最右边一列）给图像相似度的计算带来了很高的代价。一种改进的思路是对这些特征表示进行降维。这里采用了第 2 章中的判别性特征提取方法（即 BFLD），并在实验部分对不同局部 Gabor 模式表示与判别性特征提取结合的方法进行测试。

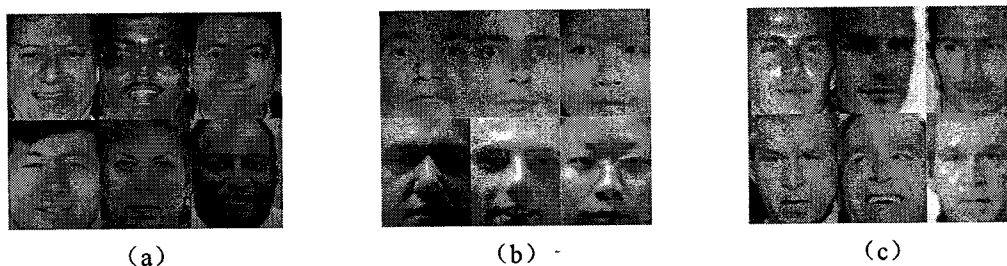


图 4-7 实验中归一化图像示例 (a) FERET (b) FRGC 2.0 (c) LFW

4.5 实验结果对比及分析

本部分采用 FERET、FRGC 2.0 和 LFW (Labeled Faces in the Wild, LFW) [42] 三个人脸数据库来测试基于学习的局部 Gabor 模式（即 LLGP）以及不同局部 Gabor 模式方法。实验设置如下：在 FERET 和 FRGC 2.0 人脸库上，根据提供的眼睛位置将所有人脸图像统一归一化到 64×80 个像素，归一化图像的两眼间距分别设置为 28 和 35 个像素；在 LFW 人脸库上，利用一个由商业人脸对齐软件处理过的图像数据[112]并根据两只眼睛中心的位置（即 (103,113)、(148,113)）将所有人脸图像的大小统一归一化到 64×80 个像素（两眼间距为 28 个像素）。图 4-7 (a) ~ (c) 分别示出了这三个人脸库上的归一化图像示例。在 FERET 数据库上，实验中采用了最近邻分类器。在 FRGC 2.0 和 LFW 数据库上，实验中将遍历相似度阈值的范围来得到 ROC 曲线。此外，实验中不采取任何光照预处理方法。

本部分内容安排如下：第 4.5.1 节测试了不同参数以及不同训练集对 LLGP 方法结果的影响；第 4.5.2 节测试了加权策略对 LLGP 方法的影响并给出了相应的分析；第 4.5.3 节对比了不同局部 Gabor 模式方法在三个数据库上的结果并进行分析；第 4.5.4 节根据第 4.5.3 节中实验结果提出了几个有趣的问题，并进行分析讨论。

4.5.1 参数对结果的影响及分析

下面将在 FERET 人脸库上系统地测试参数对 LLGP 方法的影响，并给出了相应的分析。由于 LLGP 方法需要一个训练集来构建码本，本部分也测试了不同训练集产生的码本对实验结果的影响。

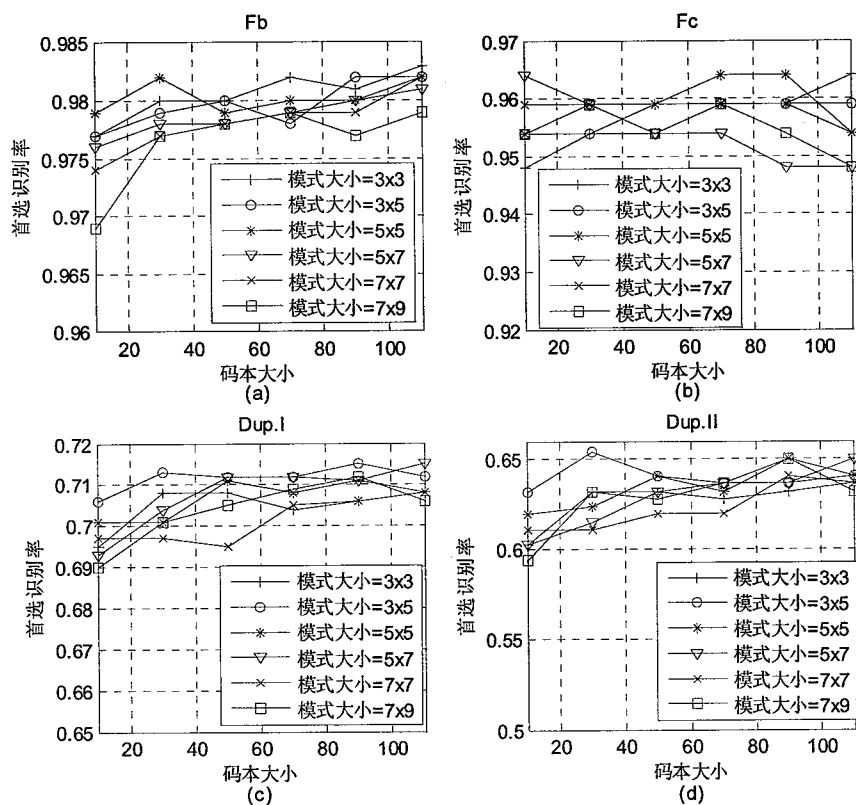


图 4-8 不同参数对 LLGP 方法结果的影响

1. 不同参数对结果的影响

根据第 4.2 节的分析，LLGP 方法主要有三个参数：模式大小（即采样图像块的尺寸），码本大小（即模式数目）和直方图表示的子块大小。这三个参数对 LLGP 方法的影响分析如下：模式大小的值越大，图像块的语义特征越强，其特征变化的范围越大，要有效地表示图像块所需要的模式数目就越多。相反，它的值越小，其语义信息就越弱，图像块之间的表现差异比较小，从而使不同模式之间的区分性变弱；码本大小的值越大，对不同变化的图像块的表示越精细，但同时也带来了高维的特征问题；直方图表示的子块大小需要在保留不同区域的相对位置信息和降低图像的特征维数之间进行均衡（参见

第二章 2.5.2 节)。下面将重点测试码本大小和模式大小这两个参数对结果的影响,并将模式大小的范围设置为 3×3 到 7×9 个像素,码本大小的范围设置为 $[10, 110]$,直方图表示的子块数目设置为 64 (即 8×8)。

图 4-8 (a) ~ (d) 分别示出了 LLGP 方法在不同参数取值时在 Fb、Fc、Dup.I 和 Dup.II 子集上的结果。可以看到,在码本大小从小变大的过程中,实验结果基本呈现了上升趋势。这说明较多数目的模式能够刻画更细粒度的特征;尽管如此,由于图像子块的尺寸是固定的,这也使直方图变得太稀疏而不利于特征的匹配。因此,当模式数目增加到一定数值时结果会保持不变甚至下降的趋势。此外,当模式大小取不同数值时,实验结果并没有呈现一致的变化趋势:从 Fb 和 Fc 子集上的结果来看,模式大小取 5×5 时 LLGP 方法取得了最高的结果,而在 Dup.I 和 Dup.II 子集上,模式大小取 3×5 时取得了相对较高的性能。

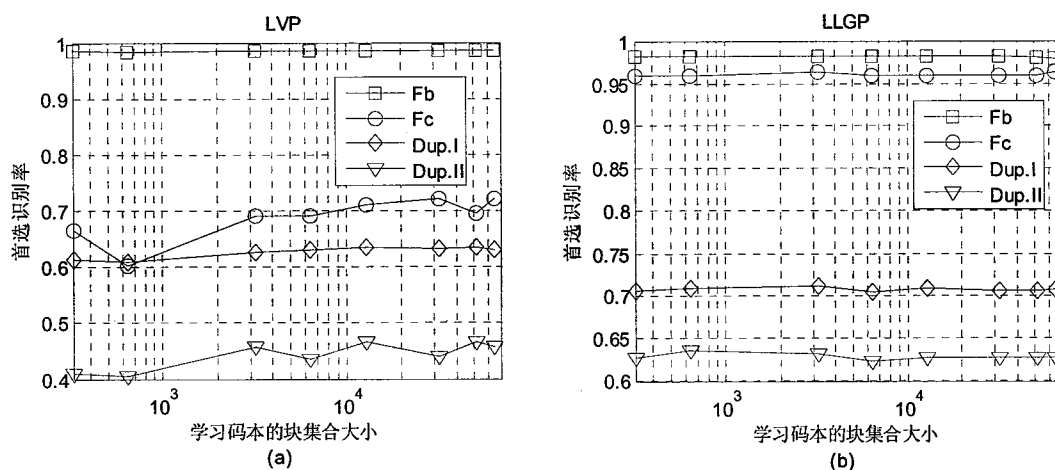


图 4-9 学习码本的块集合大小对不同方法结果的影响 (a) LVP (b) LLGP

图 4-9 (a) 和 (b) 中进一步绘出了学习码本的块集合大小分别对 LVP 和 LLGP 这两种方法结果的影响。根据图 4-8 中的实验结果,考虑到 LLGP 表示中不同参数产生的计算代价以及所取得的实验性能,这里采用了如下的参数设置:码本大小设置为 70,模式大小设置为 5×5 个像素,图像划分为 64 (即 8×8) 个子块,块集合的大小从 320 增加到 64,000。可以看到,在块集合从小变大的过程中,LVP 表示在 Fc、Dup.I 和 Dup.II 子集上的性能基本呈上升趋势;LLGP 表示在 Fb 和 Fc 子集上的结果基本上保持不变,在 Dup.I 和 Dup.II 子集上的性能先出现小幅度的上升,然后保持不变。LLGP 表示之所以呈现与 LVP 表示不同的变化趋势,原因在于:Gabor 幅值特征的平滑性导致了从其上采样得到的图像块之间的差异要小于灰度特征上采样的图像块之间的差异。换言之,单个尺度、单个方向上 Gabor 特征的图像块包含的变化类型比较少。因而,学习码本的块集合大小对产生码本的影响比较小。

2. 不同训练集产生的码本对结果的影响

LLGP 方法需要一个训练集来学习相应的码本,在上面 FERET 人脸库的实验中,

它是根据同质训练数据上产生的码本来表示相应数据库上的人脸图像,即:根据 FERET 人脸库提供的训练集学习 LLGP 码本,进而来表示 FERET 测试集中的图像。一个有趣的实验是测试异质数据上产生的码本对 LLGP 方法的影响。下面进行了如下的实验:利用 FERET 人脸库的训练集产生的 LLGP 码本来表示 FRGC 2.0 人脸库的目标集和查询集中的图像,然后进行人脸确认实验;利用 FRGC 2.0 人脸库中的训练集产生的 LLGP 码本来表示 FERET 测试集上的图像,然后进行人脸识别实验。

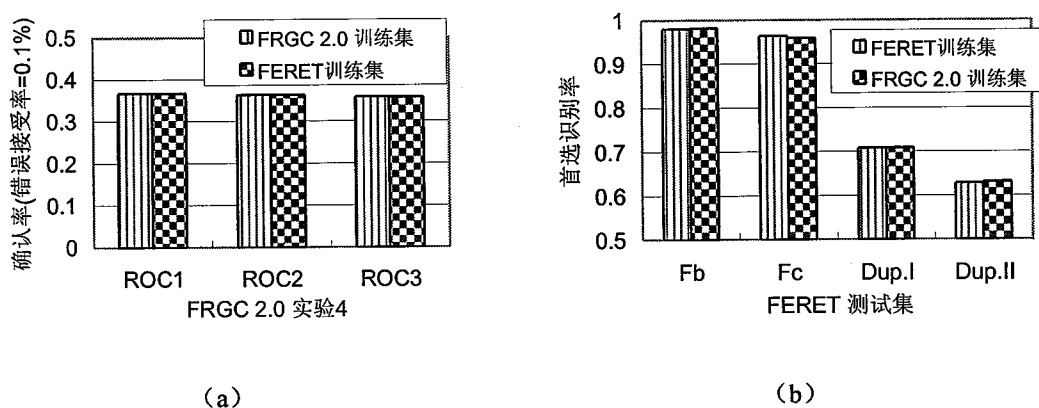


图 4-10 不同训练集产生的码本对 LLGP 方法的影响 (a) FRGC 2.0 (b) FERET

图 4-10 (a) 和 (b) 分别示出了 FRGC 2.0 实验 4 和 FERET 人脸库上 LLGP 方法在异质(即: FRGC 2.0 训练集 vs FERET 测试集, FERET 训练集 vs FRGC 2.0 测试集)和同质(即: FERET 训练集 vs FERET 测试集, FRGC 2.0 训练集 vs FRGC 2.0 测试集)测试条件下的结果对比。从图中可以看到,在这两个数据库上,LLGP 方法在异质测试条件下的结果与同质测试条件下的结果基本一致。除此之外,作者也在 FERET 和 CAS-PEAL-R1 人脸库上测试过不同训练集产生的码本对结果的影响,也得出了与此相同的结论[128]。其原因可能在于,虽然这些人脸数据库(如 FERET 和 FRGC 2.0)由不同条件下的人脸图像构成,但是,较小尺寸的图像块(例如, 5×5)具有比较弱的语义特征,这导致不同训练集上的模式构成相差不大。从这一点来看,LLGP 方法受不同训练集的影响较小。

4.5.2 加权策略及判别性特征提取方法对结果的影响

本节将在 FERET 和 FRGC 2.0 人脸库上验证加权策略的有效性。由于判别性特征提取方法(即 BFLD)(参见第二章 2.4.1 节)可对局部 Gabor 模式直方图表示进行特征提取,这里也测试了 LLGP 表示与其结合方法的结果。图 4-11 示出了 LBP、LVP 和 LLGP 方法以及它们结合不同策略时在 FERET 四个子集上的结果。相关的参数设置如下:在 LLGP 方法中,模式大小设置为 5×5 个像素,码本大小设置为 70,图像划分为 64 个子块;在 LLGP 表示与 BFLD 结合的方法中,模式图像划分为 16 个块,每个块划分为 2

个子块，每个块的特征维数（即 $2 \times 70 \times 12 = 1,680$ ）经 FLD 降至 200 维；LVP 方法采用了与 LLGP 方法中相同的参数设置；LBP 方法中子块数目设置为 64，其模式数目设置为 59 种（即 Uniform 模式）。

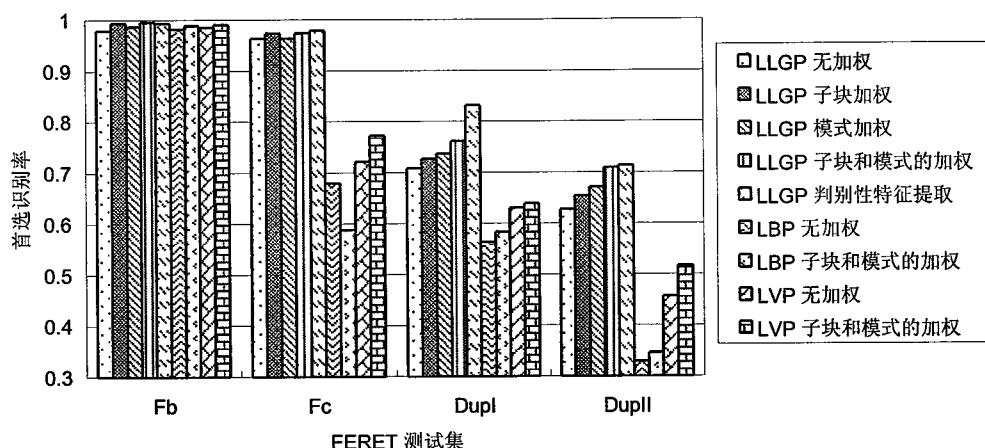


图 4-11 不同方法在 FERET 人脸库上的结果对比

从图 4-11 中可以看到，子块加权与模式加权的策略均提高了 LLGP 方法的结果，特别是在 Dup.I 和 Dup.II 这两个子集上：通过结合子块与模式的权重，LLGP 方法在 Dup.I 和 Dup.II 子集上分别提高了 6% 和 8%；其与判别性特征提取结合的方法在 Fb、Fc 和 Dup.II 子集上的结果与利用子块和模式权重的加权方法取得了相当的结果，但在 Dup.I 子集上取得了比较高的分类精度；相比无加权方法，LBP 和 LVP 的加权方法在精度上也均有提高。从图中也可观察到，LVP 方法取得了比 LBP 方法更好的结果。这表明了模式学习方法（即 LVP）可以取得比模式设计方法（即 LBP）更优的性能。此外，利用 Gabor 特征的 LLGP 方法在这四个子集上的结果要优于基于亮度特征的方法（即 LBP, LVP）。当然，该方法的不足之处在于它具有过高的计算代价。

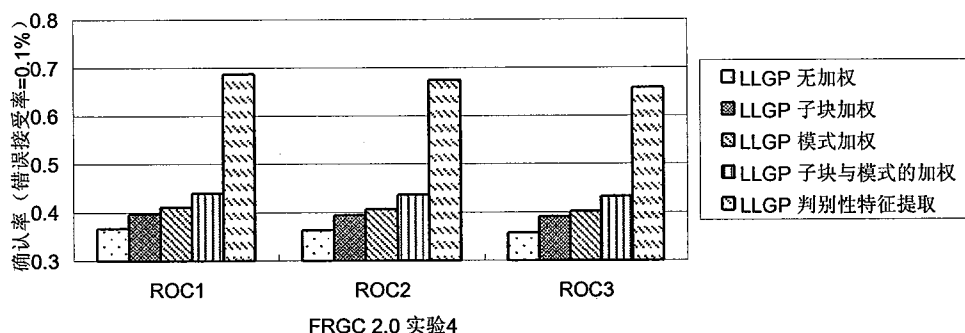


图 4-12 LLGP 表示与不同策略结合的方法在 FRGC 2.0 人脸库上的结果对比

为了进一步验证加权策略的有效性，这里在 FRGC 2.0 人脸库的实验 4 上进行了测试。在报告实验 4 的结果时，这里采用了第三章 3.4.4 节提到的相似度归一化策略（等

式 (3-13))。图 4-12 示出了 LLGP 方法在结合不同策略下的实验结果。可以看到, 与无加权方法的结果相比, 利用子块和模式权重的加权方法在 ROC1、ROC2、ROC3 上分别提高了 7.2%、7.3%、7.5%, 而其与判别性特征提取结合的方法则取得了更高的结果 (比无加权方法的结果提高了接近 30%)。这与图 4-11 中的实验结果基本一致, 验证了加权策略及判别性特征提取方法的有效性。

4.5.3 不同局部 Gabor 模式方法的结果对比

本节将在 FERET、FRGC 2.0 和 LFW 这三个人脸数据库上对不同局部 Gabor 模式方法进行测试。虽然利用块和模式权重的加权策略有效地提高了原来方法的识别精度, 但是, 该方法具有与原来方法基本相当的图像相似度计算代价。由第 4.5.2 节的结果对比可知, 基于块的 Fisher 线性判别分析 (即 BFLD) 极大地提高了识别精度; 而且, 它提取了低维的特征, 因而具有较低的相似度计算代价。因此, 下面进一步测试了不同局部 Gabor 模式与 BFLD 结合的方法, 并给出了相应的分析。

表 4-2 局部 Gabor 模式及其与 BFLD 结合方法的参数设置

方法	局部 Gabor 模式			局部 Gabor 模式与 BFLD 结合的方法				
	模式参数	模式数目	子块数目	模式参数	模式数目	块数目	子块数目	块 FLD 维数
LGBP_Mag	P=6	64	8×8	P=4	16	4×4	2×2	200
E-GV-LBP/V-LGBP	P=4,L=2	64	8×8	P=2,L=2	16	4×4	2×2	200
GGPP	O=8	256	8×8	O=6	64	4×4	2×2	200
LGPP	P=4	64	8×8	P=3	8	4×4	2×2	200
LGBP_Pha	P=6	64	8×8	P=4	16	4×4	2×2	200
LGPDP	P=6	64	8×8	P=4	16	4×4	2×2	200
LGXP	P=6,b=4	64	8×8	P=4,b=4	16	4×4	2×2	200
LVC/LGT	C=64	64	8×8	C=64	64	8×8	1×1	40/20
LLGP	C=70	70	8×8	C=70	70	4×4	1×2	200

表 4-2 中列出了不同局部 Gabor 模式以及它们与 BFLD 方法结合时的参数设置。除 LLGP 方法外, 在其他方法中, Gabor 小波参数设置为: 5 个尺度 $\nu \in \{0, 1, \dots, 4\}$, 8 个方向 $\mu \in \{0, 1, \dots, 7\}$, $\sigma = 2\pi$, $k_{\max} = \pi$, $f = \sqrt{2}$ 。表 4-2 中所有的对比方法都由本文作者实现, 并且在完全相同的测试协议下进行测试。

在局部 Gabor 模式方法中, 以下几点需要注意: 第一, 在幅值“体”的局部模式 (即 E-GV-LBP/V-LGBP) 中, Jet 上的模式根据相同尺度不同方向的滤波器响应计算得到; 第二, 在实虚部的全局 Gabor 相位模式 (即 GGPP) 中, 作者测试了不同模式数目的影响, 结果发现 256 种 (即 O=8) 模式时性能最优, 因而采用了该参数设置; 第三, 在实虚部的局部 Gabor 相位模式 (即 LGPP) 中, 由于其模式图像的数目 (即 80) 要高于其

他方法, 这里将其模式数目设置为 16 种 (即 $P=4$) 以避免过高的特征维数。

第 2 章 2.5.2 节的测试结果表明, 局部 Gabor 模式与 BFLD 结合的方法在块数目和子块数目分别取 16 (即 4×4) 和 4 (即 2×2) 时能够取得良好的精度。因此, 大多数方法采用了该参数设置 (参见表 4-2)。在 GGPP 和 LGPP 方法中, 为了避免过高的特征维数给 FLD 矩阵学习带来巨大的计算代价, 其模式数目分别设置为 64 种和 8 种。在 LVC/LGT 方法中, 码本的数目与图像块的数目相同, 因此在结合 BFLD 方法时采用了相同数目的块 (即 8×8); 由于每个块的尺寸太小 (即 8×10 个像素), 为避免直方图太稀疏, 每个块不再划分为多个子块。在 LLGP 方法中, 由于模式数目较多, 为避免每个块过高的特征维数带来更大的计算代价, 子块的数目设置为 2 (即 1×2)。除 LVC/LGT 方法外, 大多数方法中每个块的特征维数经 FLD 后降至 200 维。由于 LVC/LGT 每个块的原始特征维数比较低 (64 维), 每个块的特征维数经 FLD 后降至 40 维 (FERET 人脸库) 或 20 维 (FRGC 2.0 和 LFW 人脸库)。

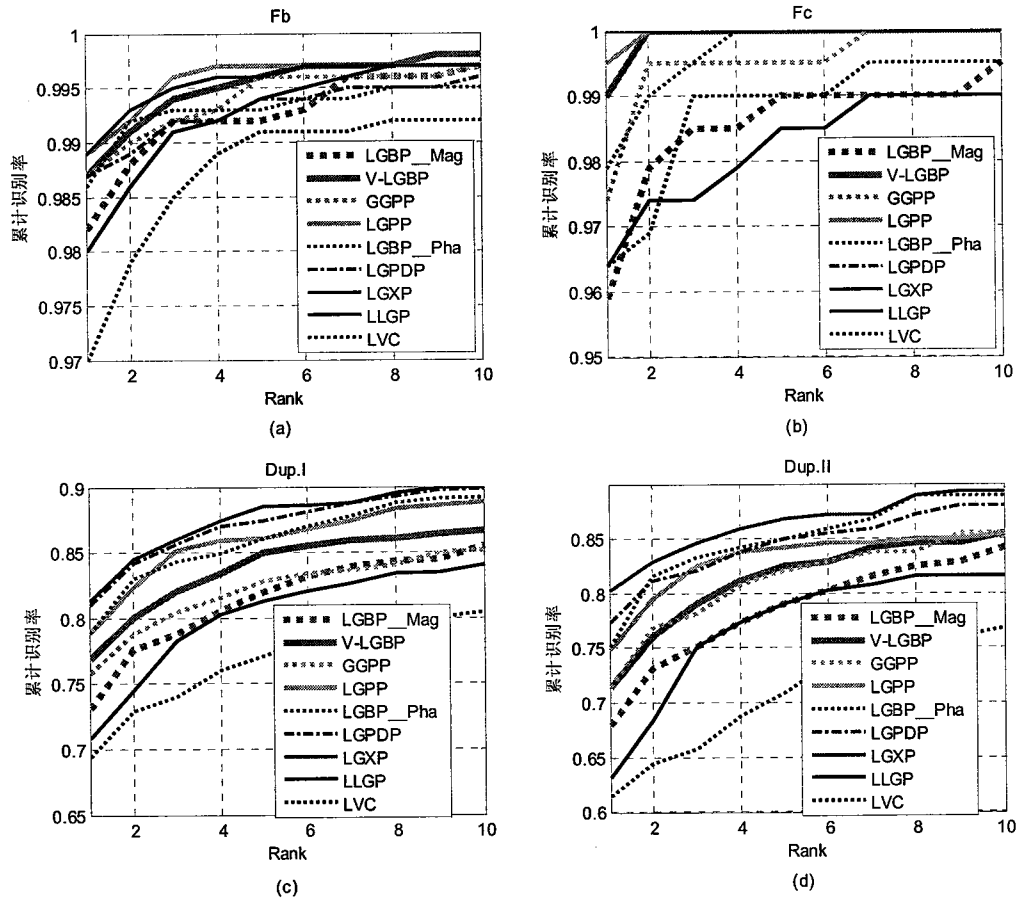


图 4-13 不同局部 Gabor 模式方法在 FERET 人脸库上的结果对比

1. FERET 人脸数据库上的结果对比与分析

图 4-13 示出了不同局部 Gabor 模式在 FERET 人脸库四个子集上的累计识别率曲线

(Cumulative Match Curves, CMCs)。从图中可以看到,大多数方法在 Fb 和 Fc 子集上都取得了非常高的结果(首选识别率高于 96%),但在 Dup.I 和 Dup.II 子集上取得了相对较低的识别率。这表明 Dup.I 和 Dup.II 子集是更富挑战性的集合。因此,下面将关注不同方法在这两个子集上的结果,并得到如下的观察:第一,大多数模式设计的方法(如 LGXP, V-LGBP)的识别精度要高于模式学习的方法(如 LVC 和 LLGP)。第二,在模式设计方法中,相位特征上局部模式(如 LGXP, LGBP_Pha)的结果比其他 Gabor 特征(即实部、虚部、幅值)的局部模式(如 V-LGBP, LGBP_Mag)表现更好。这表明如果根据 Gabor 相位特征的特点设计合理的表示方法,这类方法也可能会取得很好的分类性能。

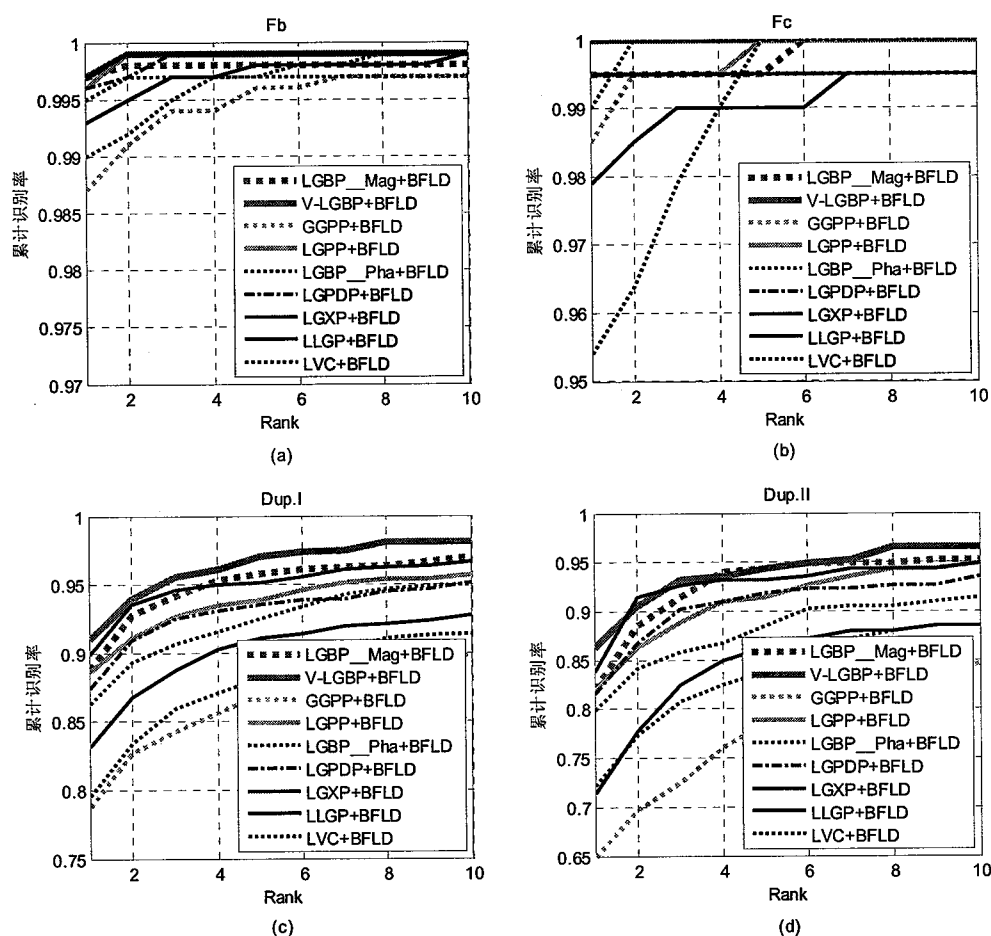


图 4-14 局部 Gabor 模式与 BFLD 结合的方法在 FERET 人脸库上的结果对比

图 4-14 绘出了不同局部 Gabor 模式与 BFLD 结合的方法在 FERET 人脸库四个子集上的累计识别率曲线。同图 4-13 中的结果相比,大多数局部 Gabor 模式与 BFLD 结合的方法取得了比局部 Gabor 模式更高的分类精度。例如,幅值“体”的局部模式(即 V-LGBP)在 Dup.I 和 Dup.II 子集上分别取得了 76.9%和 71.4%的首选识别率,而其与 BFLD 结合的方法分别取得了 91.0%和 86.3%的首选识别率。这主要归因于 BFLD 方法

具有的判别性特征提取的性质以及“分而治之”的策略。具体来讲,通过建模类内差异与类间差异,FLD 可以有效地学习到区分不同类别的特征变换;此外,通过“分而治之”的策略,BFLD 方法整合了不同人脸区域的决策(即相似度)。这种整合策略利用了这些区域之间的互补性,进一步提高了识别性能。

对比图 4-13 和图 4-14,有如下几点意外的观察:第一,在局部 Gabor 模式方法中,幅值“体”局部二值模式(即 V-LGBP)的性能要低于相位的局部 Gabor 异或模式(即 LGXP)。然而,当结合 BFLD 方法时,前者基本上取得了最好的结果。同样的情形也存在于幅值局部 Gabor 二值模式(即 LGBP_Mag)和实虚部的局部 Gabor 相位模式(即 LGPP)中。原因可能是幅值局部模式表示具有较大的冗余度,其在结合 BFLD 方法后能够提取那些适合于分类的更紧致的特征;第二,从局部 Gabor 模式及其与 BFLD 结合方法的结果对比中,可以看到,全局 Gabor 相位模式(即 GGPP)在 Dup.I 子集上的首选识别率从 75.8%提高到 78.8%,但在 Dup.II 子集上的识别率则从 71.4%下降到 65.0%。并且,其与 BFLD 结合方法的识别率要低于基于学习的局部模式(即 LVC 和 LLGP)与 BFLD 结合的方法。这表明了并不是所有的模式设计方法总优于模式学习的方法。

2. FRGC 2.0 人脸数据库上的结果对比与分析

这里在 FRGC 2.0 实验 4 上对不同局部 Gabor 模式以及它们与 BFLD 结合的方法进行测试。在报告实验 4 的结果时,这里采用了第三章 3.4.4 节提到的相似度归一化策略。图 4-15 和图 4-16 分别绘出了局部 Gabor 模式及其与 BFLD 方法结合的结果。可以看到,通过结合 BFLD 方法,图中所有局部 Gabor 模式方法的性能提升很大。例如,在错误接受率=0.1%时,幅值“体”的局部二值模式(即 V-LGBP)在 ROC1、ROC2、ROC3 上的确认率分别是 49.0%、48.9%、48.8%,而其与 BFLD 结合方法的确认率分别是 87.1%、86.9%、86.7%,性能提高了接近 40%。

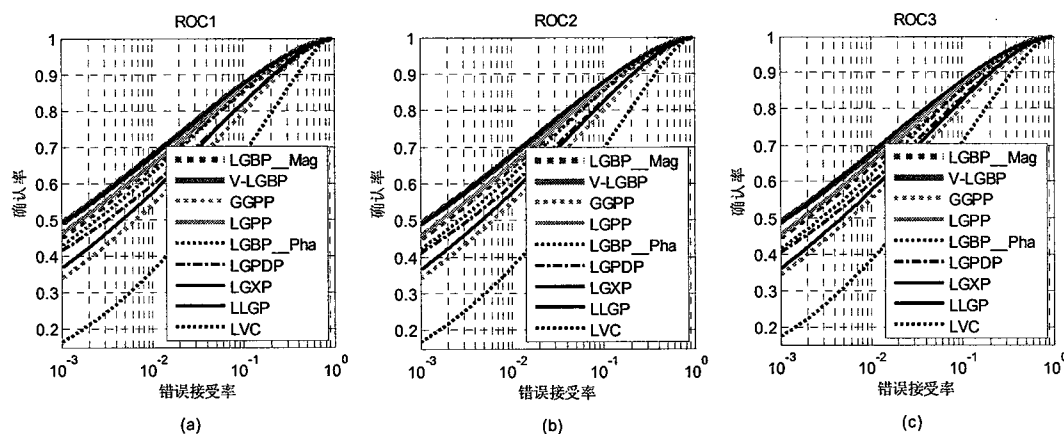


图 4-15 局部 Gabor 模式方法在 FRGC 2.0 人脸库实验 4 上的结果对比

从图 4-15 和图 4-16 中,还得到如下观察:第一,绝大多数的局部 Gabor 模式方法在实验 4 上取得了较低的结果:在错误接受率=0.1%时最好方法的确认率也低于 50%。

这表明这些方法在实验 4 上受到的影响很大: 局部 Gabor 模式直接从单幅图像中提取得到, 很难处理由复杂光照以及图像模糊等引起的较大的类内变化; 第二, 与 FERET 人脸库上的观察类似, 幅值局部 Gabor 二值模式 (即 LGBP_Mag) 的性能要低于相位的局部 Gabor 异或模式 (即 LGXP), 但其与 BFLD 结合方法的结果则要优于后者; 第三, 无论是否与 BFLD 方法结合, 实虚部的全局 Gabor 相位模式 (即 GGPP) 的结果总是介于基于学习的局部 Gabor 模式表示 (即 LLGP) 与基于学习的视觉码本 (即 LVC) 的结果之间。

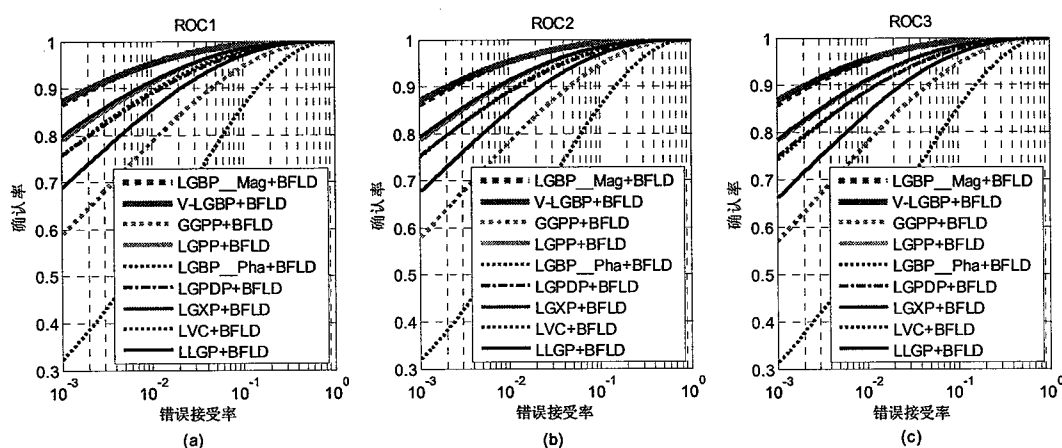


图 4-16 局部 Gabor 模式结合 BFLD 的方法在 FRGC 2.0 人脸库实验 4 上的结果对比

3. LFW 人脸数据库上的结果对比与分析

这里在富有挑战性的 LFW 人脸库上测试了不同局部 Gabor 模式及其与 BFLD 方法结合的结果。LFW[42]人脸库主要用来研究非可控条件下的人脸识别问题, 近年来受到广泛的关注[9][42][112]。构建该数据库的一个主要动机是提供大量非可控条件下的人脸图像, 主要包括姿态、光照、表情、背景、种族、年龄、性别、服饰、发型、成像质量、彩色饱和度、聚焦等类型的变化。该数据库主要关注给定图像对的匹配问题。图 4-17 (a) 和 (b) 分别示出了该数据库的匹配图像对与非匹配图像对的部分示例。

LFW 人脸数据库包括了 5,749 个人的 13,233 幅图像。其中, 1,680 个人有两幅或多幅图像, 剩余的 4,069 个人则只有一幅图像。所有图像的大小是 250×250 个像素, JPEG 格式。大多数图像是彩色格式, 只有少数图像是灰度格式。该数据库的评测协议定义了两种“视图”：“视图 1”用于调整算法参数, “视图 2”用于报告性能。通过使用视图 1, 可以避免实验者将分类器“过拟合”到最后的测试数据。按照 LFW 人脸库的测试协议, 它提供了训练数据的两种使用方法, 即: 图像限制的训练和无限制的训练。具体地讲, 在图像限制的训练中, 实验者不能根据人的姓名来推测训练集之外两幅图像的等价性和非等价性; 在非限制的训练中, 实验者可以根据人的姓名来构建尽可能多的匹配对和非匹配对。按照该测试协议, 局部 Gabor 模式方法没有利用图像类别信息, 属于图像限制的训练; 而结合 BFLD 的表示方法需要利用图像的类别信息来学习 FLD 投影矩阵,

属于无限制的训练。

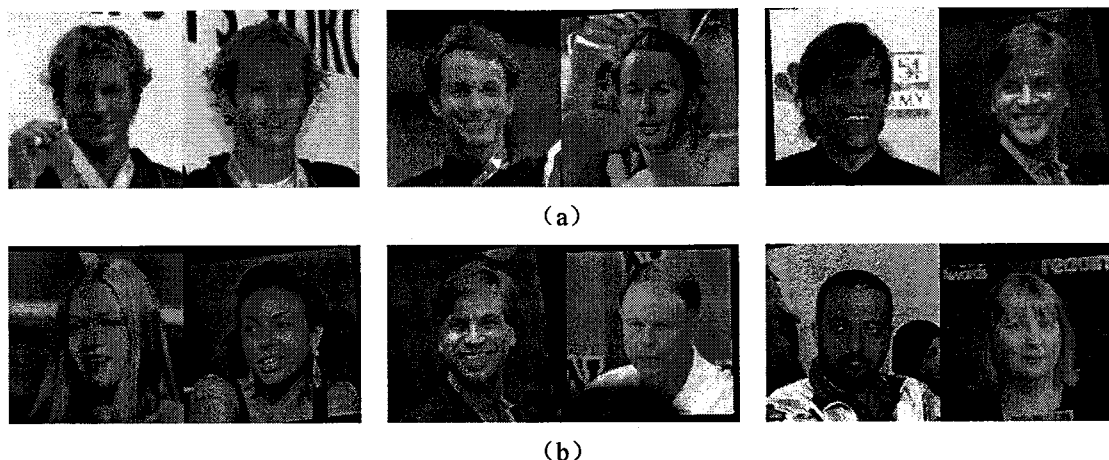


图 4-17 LFW 人脸库的示例图像对 (a) 匹配图像对 (b) 非匹配图像对

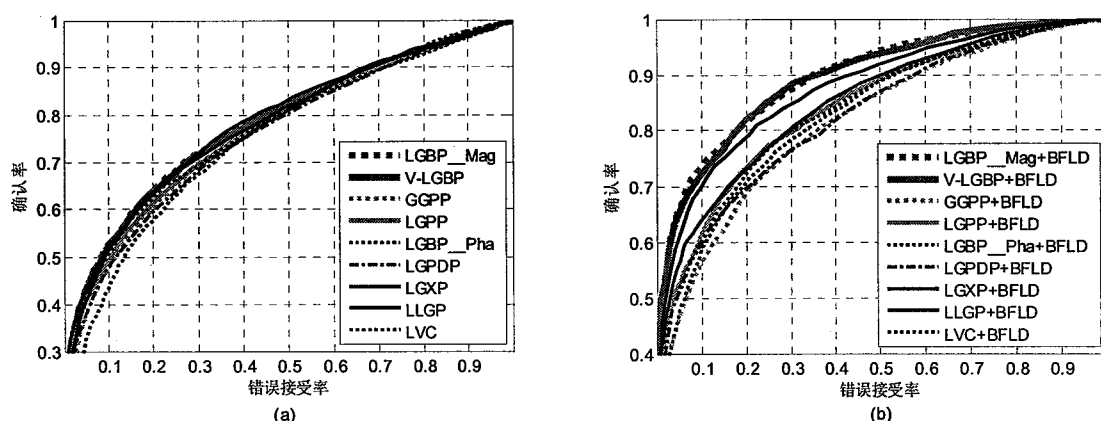


图 4-18 不同方法在 LFW 人脸库上的结果对比
(a) 局部 Gabor 模式 (b) 局部 Gabor 模式与 BFLD 结合的表达方法

图 4-18 (a) 和 (b) 分别绘出了局部 Gabor 模式表示及其与 BFLD 结合方法的实验结果。可以看到, 在这样富有挑战性的数据库上, 大多数方法的结果比较低。在图中的局部 Gabor 模式方法中, 最好的确认率 ($FAR=0.1$) 只达到了 53.2%; 当结合 BFLD 方法时, 最高的确认率 ($FAR=0.1$) 仅达到 73.4%。这与 FRGC 2.0 实验 4 上的结果有比较大的差距 (在错误接受率=0.1%时最好方法的确认率接近 90%)。这说明了 LFW 人脸数据库比 FRGC 2.0 人脸数据库更具有挑战性。此外, LFW 人脸库中图像的变化非常大, 这导致利用两眼中心归一化的图像不能精确对齐, 对这些方法的性能也有比较大的影响。从图中可以注意到, 在图像限制训练和无限制训练的这两种条件下, 幅值局部 Gabor 二值模式 (即 LGBP_Mag) 和幅值“体”的局部二值模式 (即 V-LGBP) 基本都取得了最好的结果。原因可能在于幅值特征对于非可控条件下引起的变化具有更好的鲁棒性。在图像限制训练的条件下, 基于学习的视觉码本 (即 LVC) 方法取得了最差的结果; 在非限制训练的条件下, 实虚部的全局 Gabor 相位模式 (即 GGPP) 的性能最差。原因

可能在于，这两种模式表示没有区分人脸图像在不同尺度方向（LVC）或者不同尺度（GGPP）上的结构特征，因而具有比较弱的表示能力。

4.5.4 讨论

根据第 4.5.3 节中不同局部 Gabor 模式及其结合 BFLD 方法的结果对比，可以看到一些有趣的现象。下面将提出几个有趣的问题，并进行分析与讨论。

1. 在这 11 种局部 Gabor 模式方法中，哪一种方法的性能最好？

为了回答这个问题，这里将这 11 种方法根据它们在不同人脸库上的实验结果进行排序并计算平均排名，如表 4-3 所示。从表中可以看到，在局部 Gabor 模式方法中，Gabor 幅值“体”的局部模式（即 E-GV-LBP/V-LGBP）和相位特征的局部 Gabor 异或模式（即 LGXP）方法基本上取得了最好的结果。这说明了它们对幅值和相位特征的局部模式进行了有效地建模表示。此外，在模式设计的方法中，在非可控条件下的人脸库（如 LFW）上，幅值特征的局部模式取得了比其他特征（如实部、虚部、相位）的局部模式表示更好的结果。原因可能在于，幅值特征对非可控条件下的图像变化鲁棒，因而利用幅值特征的局部模式表示是更有效的方法。

表 4-3 局部 Gabor 模式及其与 BFLD 结合的方法在三个人脸库上的性能排名

数据库 排名			局部 Gabor 模式				局部 Gabor 模式与 BFLD 结合的方法			
			FERET	FRGC 2.0	LFW	平均排名	FERET	FRGC 2.0	LFW	平均排名
模式设计方法	幅值	LGBP_Mag	7	4	2	4.33	3	2	1	2
		E-GV-LBP/V-LGBP	5	1	1	2.33	1	1	1	1
	实部/虚部	GGPP	6	8	6	6.67	9	8	9	8.67
		LGPP	4	3	8	5	4	4	5	4.33
	相位	LGBP_Pha	3	5	3	3.67	6	5	7	6
		LGPD	2	6	7	5	5	5	6	5.3
		LGXP	1	2	4	2.33	2	3	4	3
模式学习方法	幅值	LVC/LGT	9	9	9	9	8	9	8	8.33
		LLGP	8	7	5	6.67	7	7	3	5.67

表 4-3 中进一步列出了局部 Gabor 模式与 BFLD 结合的方法在三个人脸库上的排名。可以看到，利用幅值特征的模式设计方法（即 E-GV-LBP/V-LGBP 和 LGBP_Mag）是表中所有方法中性能最好的，局部 Gabor 异或模式（即 LGXP）与 BFLD 结合的方法排到第三位。此外，无论是否与 BFLD 结合，Jet 上的模式表示（即 GGPP，LVC/LGT）都几乎取得了最差的结果。

2. 这些局部 Gabor 模式方法的计算代价是怎样的？

在这 11 种局部 Gabor 模式方法中，8 种模式设计的方法具有基本相当的计算代价，而 3 种模式学习方法由于需要学习码本，因而其模式计算代价不同于模式设计的方法。这里选择了 3 种代表性的方法：幅值局部 Gabor 二值模式（即 LGBP_Mag），基于学习的视觉码本（即 LVC）和基于学习的局部 Gabor 模式（即 LLGP）。此外，本节还对比了 LGBP_Mag 方法以及“LGBP_Mag+BFLD”方法的计算代价。这些方法的参数设置见表 4-2。下面选择了如下配置的 PC 机来测试这些方法的计算代价：Intel(R) Core(TM)2 Duo CPU E6550@2.33GHz, 2GB 内存。测试数据选择了 FERET 人脸库的一部分数据，包括训练集（429 个人的 1,002 幅图像）、注册集（1,196 个人的 1,196 幅图像）和测试子集 Fb（1,195 个人的 1,195 幅图像）。

表 4-4 不同局部 Gabor 模式及其与 BFLD 结合方法的计算时间比较

注：“平均耗时”=共计的耗时时间/测试图像数目

项目 \ 耗时（毫秒）	LGBP_Mag、LVC 和 LLGP			LGBP_Mag 和 “LGBP_Mag+BFLD”	
	LGBP_Mag	LVC	LLGP	LGBP_Mag	LGBP_Mag+BFLD
码本	--	5,021,078	451,672	--	--
直方图特征	327,126	375,078	857,922	327,126	325,954
BFLD 矩阵	--	--	--	--	1,143,905
FLD 特征	--	--	--	--	88,468
相似度计算	1,075,359	30,406	386,094	1,075,359	24,203
共计	1,402,485	5,426,562	1,695,688	1,402,485	1,582,530
平均	1,173.63	4,541.06	1,418.99	1,173.63	1,324.29

表 4-4 中列出了不同方法在各个阶段的计算代价。同 LGBP_Mag 方法相比，LVC 和 LLGP 方法都需要一个学习过程来构建码本。LVC 方法需要学习 64 个码本而 LLGP 方法只需要学习 12 个码本，因此，前者在码本学习阶段的耗时高于后者。此外，在计算局部模式时，LVC 和 LLGP 方法都需要将 Jet 或块的特征映射到相应的码本，而 LGBP_Mag 方法只需要简单的算术操作（如加法和乘法）。因此，从计算图像直方图特征的过程来看，LVC 和 LLGP 方法的计算代价高于 LGBP_Mag 方法（参见表 4-4 中的“码本”和“直方图特征”对应的行）。当计算图像的相似度时，LGBP_Mag 方法比 LVC/LLGP 方法具有更高的特征维数。因此，LGBP_Mag 方法具有更高的计算代价（参见表 4-4 中的“相似度计算”行）。表 4-4 中列出了这三种方法在每幅测试图像上的平均耗时。可以看到，LGBP_Mag 和 LLGP 方法的平均耗时相当，二者均小于 LVC 方法。尽管如此，随着测试图像数目的增加，LGBP_Mag 方法的相似度计算代价将占很大的比例。当测试图像数目增加到一定规模时，就每幅图像的平均耗时而言，LVC 方法可能是最低的。

直觉上, 由于学习多个 FLD 矩阵以及利用这些矩阵计算低维特征, LGBP_Mag 表示结合 BFLD 方法的计算代价必定要高于 LGBP_Mag 方法。然而, 该方法的特征维数要远远低于其 LGBP_Mag 表示的特征维数, 因而具有很低的相似度计算代价 (参见表 4-4 的“相似度计算”行)。所以, LGBP_Mag 方法与“LGBP_Mag+BFLD”方法的在每幅测试图像上的平均耗时基本相当, 如表 4-4 所示。当测试图像数目增加到一定规模时, 对 LGBP_Mag 方法而言, 相似度计算的耗时占据了其耗时的大部分。在这种情况下, BFLD 方法会加快相似度的计算速度, 从而使“LGBP_Mag+BFLD”方法的平均耗时要低于 LGBP_Mag 方法。

3. 局部 Gabor 模式方法与其他方法的结果对比

这里选择了幅值局部 Gabor 二值模式 (即 LGBP_Mag) 及其结合 BFLD 的方法 (即“LGBP_Mag+BFLD”) 作为两种代表性的方法, 并相应地选择了两种基准方法: “Gabor Jet”与“Gabor+BFLD”。具体来讲, “Gabor Jet”方法在图像上进行均匀下采样得到多个 Gabor Jet 特征 (由相同尺度上不同方向的 Gabor 小波响应构成), 然后计算相应 Jet 特征之间的相似度, 最后将所有的相似度通过和规则进行融合[159]; “Gabor+BFLD”方法将图像的 Gabor 特征按照空间位置分块之后在块内进行下采样, 然后每个块的特征利用 FLD 计算其低维特征表示, 最后利用和规则来融合所有块之间的相似度[108]。这两种基准方法的参数设置如下: 在“Gabor Jet”方法中, 采样间隔设定为 4×4 像素; 在“Gabor+BFLD”方法中, 每幅 64×80 的图像划分为 20 (即 4×5) 个块, 每个块的特征维数通过下采样降到 640 维, 再利用 FLD 降至 200 维。

表 4-5 LGBP_Mag, “LGBP_Mag+BFLD”与基准算法的结果及其他方法结果之间的对比(%)

方法 \ 测试数据库	FERET (首选识别率)				FRGC 2.0 实验 4 (错误接受率=0.1%时的确认率)			LFW (错误接受率=10%时的确认率)
	Fb	Fc	Dup.I	Dup.II	ROC1	ROC2	ROC3	
Gabor Jet	97.8	97.9	74.7	66.2	36.8	37.4	38.0	48.1
LGBP_Mag	98.2	95.9	73.1	67.9	45.2	44.9	44.5	52.9
Gabor+BFLD	99.7	98.5	87.5	79.9	87.7	86.3	84.7	74.3
LGBP_Mag+BFLD	99.8	99.5	88.8	82.1	86.5	86.2	85.9	73.5
State-of-the-art results	99.3 ^[129]	99.5 ^[129]	93.8 ^[129]	92.7 ^[129]	93.91 ^[23]	93.55 ^[23]	93.12 ^[23]	≈87.0 ^[112]

表 4-5 中列出了这两种代表方法及基准方法的实验结果。如表中所示, 在 FRGC 2.0 和 LFW 人脸库上, LGBP_Mag 方法取得了比“Gabor Jet”方法更好的结果。这与 LGBP_Mag 方法采用了有效的模式编码方法和鲁棒的直方图表示有很大关系。这也揭示了局部 Gabor 模式方法在人脸识别中受到广泛关注的原因。然而, 对于这两种表示与判别性特征提取结合的方法 (即“LGBP_Mag+BFLD”与“Gabor+BFLD”), 二者基本上取得了相当的结果。直觉上, 既然“LGBP_Mag”方法比“Gabor Jet”方法有效, 那

么, 在结合相同的特征提取方法(即 BFLD)时, “LGBP_Mag+BFLD”方法应当取得比“Gabor+BFLD”方法更好的结果。原因可能在于, 基于直方图的距离度量对于度量局部 Gabor 模式的直方图特征之间的距离是有效的, 但在 FLD 中并没有采用同样的距离度量。尽管如此, 同局部 Gabor 模式表示相比, BFLD 方法已经极大地提升了原有方法的性能。

此外, 表 4-5 中列出了当前最好方法在这三个数据库上的结果。这三种方法都是利用了两种或多种特征进行融合: 文献[129]融合了幅值和相位特征的局部模式表示; 文献[23]利用了人脸图像两个模板上的彩色特征和 Gabor 小波表示; 文献[112]利用 SVM 分类器对多种特征计算得到的相似度进行融合。可以看到, 局部 Gabor 模式方法中最好的结果与这些结果之间仍然有一定差距。需要指出的是, 局部 Gabor 模式表示是一种单一的特征表示, 它们完全可以与其他类型的特征结合以进一步提高识别的精度。

4.6 本章小结

本章从模式学习的角度研究了人脸图像 Gabor 特征上局部模式的学习及特征表示方法, 并进一步提出了利用子块和模式权重的加权方法。针对文献中存在的 11 种局部 Gabor 模式, 文中按照模式的定义方式将它们划分为两大类: 模式设计方法和模式学习方法, 并从模式定义流程以及方法参数这两个方面进行了对比分析。此外, 由于这些模式表示具有高维的特征, 本章采用了判别性特征提取的方法来获得低维的判别特征。所有这些方法以及它们结合判别性特征提取的方法在三个大规模数据库上进行了对比, 并对一些有趣的问题进行了讨论。

本章的实验结果表明, 基于 Gabor 特征块的模式学习方法要优于利用 Gabor Jet 特征的模式学习方法以及基于亮度特征的 LBP 和 LVP 方法, 特别地, 该模式学习方法受不同训练集产生码本的影响不大; 提出的加权方法利用子块和模式的权重, 在 FERET 和 FRGC 2.0 人脸库上一致地优于无加权的方法; 该局部 Gabor 模式与判别性特征提取结合的表示方法提取了有利于分类的特征, 取得了比基准方法和加权方法更优的性能。

从模式设计方法与模式学习方法的结果对比来看, 许多模式设计方法的分类性能要优于模式学习的方法。在与判别性特征提取方法结合后, 大多数模式设计方法的性能仍然要优于模式学习的方法。这里的原因在于, 模式学习方法利用了聚类方法(即 K 均值)来学习人脸图像上的模式, 这些模式反映了人脸“共性”的特征而非“判别性”特征。如果能够将判别性准则引入到局部模式的学习中, 这类方法仍可能有很大的空间来提升分类精度。

不同局部 Gabor 模式方法在非可控条件下(如 FRGC 2.0 人脸库的实验 4, LFW 人脸库)与可控条件(FERET 人脸库)的对比结果说明, 它们在前者上的性能要远远低于后者。主要的原因在于, 在非可控条件下, 人脸库上的多种变化使得根据直方图特征计算的样本可分性非常低。幸运的是, 这些挑战性的问题可以通过结合判别性特征提取

的方法进行解决。此外，在所有的模式设计方法中，幅值特征的局部模式方法在非可控条件下的人脸数据库上取得了比其他 Gabor 特征的局部模式方法更好的结果。这一点在与判别性特征提取方法结合时尤为明显。这表明幅值特征的局部模式表示对非可控条件下的一些变化更加鲁棒。当然，也不能忽略其他 Gabor 特征的模式表示方法在人脸识别中的作用。如果可以合理地利用不同类型的 Gabor 特征，仍然有可能提出更有效的人脸识别方法（如第二章提出的“融合 Gabor 幅值和相位局部模式的表示方法”）。

从局部 Gabor 模式与基准算法的结果对比来看，很多种局部 Gabor 模式的结果要优于相应的基准算法；然而，在与判别性特征提取方法结合后，局部 Gabor 模式的最好结果只与对应的基准算法相当。从这一点来看，还需要对直方图表示的特征提取方法进行深入的研究。

第五章 结束语

5.1 本文工作总结

人脸识别技术作为一项最容易被人类所接受的生物特征识别技术,具有重要的理论价值和广阔的应用前景。人脸识别面临的主要挑战在于,人脸表观易受光照、表情、姿态等变化的影响,这导致了较大的类内差异及较小的类间差异。特征表示作为人脸识别算法的关键,对提高识别算法的精度起着非常重要的作用。

本文围绕着人脸图像的 Gabor 小波表示,研究了图像 Gabor 特征的局部模式表示以及其与判别性特征提取相结合的方法。概括地讲,本文所取得的主要研究成果如下:

1. 提出了融合 Gabor 幅值和相位局部模式的表示方法。

针对图像的 Gabor 相位特征具有对位置变化敏感性的特点,该方法提出了相位特征的局部 Gabor 异或模式表示:该表示先将相位特征量化到不同的区间,然后利用异或算子计算图像像素空间邻域内的局部模式,并采用分块的直方图来表示人脸图像;进一步地,为了利用 Gabor 幅值和相位特征之间的互补性,该方法采用了基于块的 Fisher 线性判别分析方法对 Gabor 幅值和相位的局部模式特征进行融合。实验结果表明,提出的局部 Gabor 异或模式表示具有很强的表示能力,可以取得与幅值的局部模式表示相当甚至更好的结果;基于块的 Fisher 线性判别分析方法提取了低维的判别特征,不仅减少了图像相似度的计算代价,而且极大地提高了识别精度;幅值和相位的局部模式特征之间具有互补性,二者的融合进一步提高了识别算法的精度。

2. 提出了 Gabor 特征“体”局部模式的表示方法。

为了高效地建模图像 Gabor 特征的局部模式,该方法将多尺度、多方向的 Gabor 特征组织为一个三维“体”:该 Gabor 特征“体”的前两维是图像平面,第三维对应了不同尺度、不同方向的 Gabor 小波。然后,该方法采用了局部二值模式算子来计算 Gabor 幅值“体”上的模式,并采用分块的直方图来表示人脸图像。进一步地,文中提出了该方法的两种扩展:基于 Fisher 准则的子块权重学习与判别性特征提取。实验结果表明, Gabor 幅值“体”的局部模式表示是一种高效的表示方法,取得了比幅值特征的局部 Gabor 二值模式表示更高的识别精度;利用子块权重的识别方法利用了人脸不同区域的权重,在训练集与测试集中图像的变化保持一致时能够提高识别的性能; Gabor 特征“体”局部模式表示与判别性特征提取结合的方法提取了有利于分类的投影方向,在测试的三个人脸库上都表现出了更优的分类性能。

3. 提出了基于学习的局部 Gabor 模式表示方法。

为了学习人脸专属的模式表示,针对某个尺度、某个方向的 Gabor 小波,该方法从

人脸图像的 Gabor 特征上采样得到图像块的集合,进而学习得到对应的码本:该码本中的每个码字对应了一种模式;然后,根据这些码本去表示人脸图像。进一步地,根据人脸的不同区域和不同模式判别能力的不同,文中提出了利用子块和模式权重的加权方法。此外,针对目前文献中存在的 11 种 Gabor 特征的局部模式表示方法,文中从模式定义流程以及方法参数两个方面对这些方法进行了对比分析,并测试了这些方法及其结合判别性特征提取方法在不同人脸数据库上的结果。大量的实验结果表明,基于学习的局部 Gabor 模式表示方法是一种有效的人脸表示,特别地,该方法受不同训练集上产生的码本的影响比较小;加权方法同时利用了子块和模式的权重,进一步提高了识别精度;结合判别性特征提取的方法提取了有利于分类的特征,取得了比基准方法和加权方法更高的识别性能。

5.2 后续工作

本文主要围绕人脸图像 Gabor 特征的局部模式表示方法进行了深入而细致的研究,并取得了一定的研究成果。然而,作为一项极具挑战性的研究课题,人脸识别中仍有很多关键问题亟待解决。后续的工作包括:

1. 面向直方图表示的特征提取方法。本文的实验结果表明,尽管局部 Gabor 模式表示方法要优于相应的基准算法,但在结合判别性特征提取方法时其性能仅与相应基准算法的结果相当甚至更低。因此,需要考虑面向直方图表示的特征提取方法,如核 Fisher 线性判别分析。

2. 适应“误配准”以及一定姿态变化的图像。局部 Gabor 模式及其结合判别性特征提取的识别算法适用于面部特征点位置准确标注下的准正面人脸图像。当由于姿态发生变化或者面部特征点定位不准确而导致人脸图像无法精确对齐时,这些方法会受到一些影响。下一步的工作可以从鲁棒的特征提取和精确的图像匹配这两个方面来考虑。

3. 在基于学习的局部模式表示方法中引入判别性准则。在人脸局部模式的学习方法中,利用聚类算法得到的模式反映了人脸“共性”的特征而非“判别性”的特征。如何将判别性准则引入进来并建立有效的特征表示,这将是下一步的研究课题。