$$\mathrm{JetLBP}^{i}_{\mu,\nu}=\begin{cases}1, if\ \mathrm{A}_{\mu,\nu}(z_c)\geq \mathrm{A}_{neighbor(\mu,\nu)}(z_c)\\ 0, otherwise\end{cases} \tag{3-3}$$

这里，$\mathrm{A}_{\mu,\nu}(z_c)$和$\mathrm{A}_{\mu,\nu}(z_i)$（$i=1,2,\cdots,P$）分别表示图像块内的中心像素$z_c$及其邻域像素$z_i$对应的幅值特征，$\mathrm{A}_{neighbor(\mu,\nu)}(z_c)$表示像素$z_c$对应的Jet上与第$\nu$个尺度、第$\mu$个方向的Gabor小波相邻的幅值响应。为了同时反映图像块与Jet上的模式，V-LGBP方法将这两种模式组合在一起来表示中心像素。

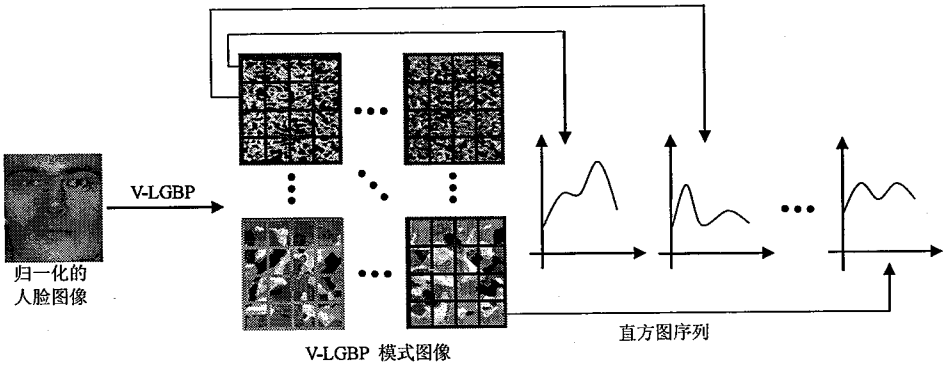
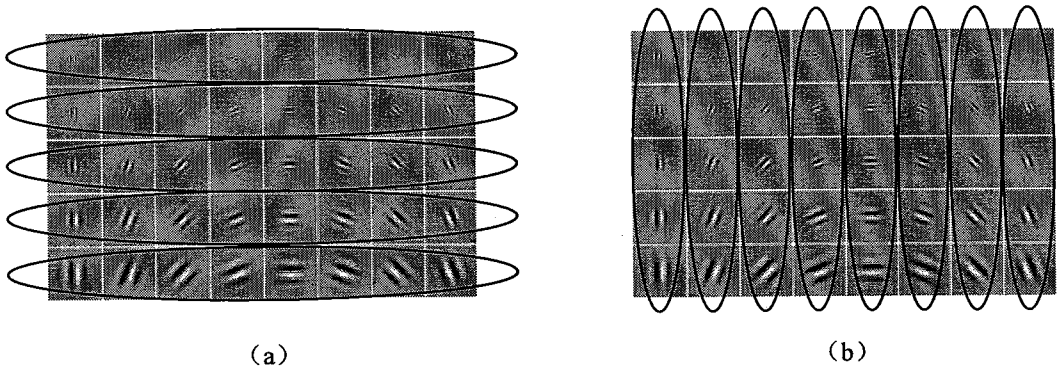


图 3-2 V-LGBP方法的分块直方图表示

根据上面的模式计算方法，每个Gabor小波对应的幅值图像都会得到一幅对应的模式图像。与幅值的局部Gabor二值模式表示类似，这里采用了分块的直方图表示来描述人脸图像：如图 3-2 所示，每幅 V-LGBP 模式图像被划分为m个不重叠的子块$\Re_{\mu,\nu,r}$（$r=1,2,\cdots,m$）并计算每个子块的直方图$H_{\mu,\nu,r}$，最后将所有尺度、所有方向上各个子块的直方图连接起来得到该幅人脸图像的特征表示：$H=\left[H_{\mu_0,\nu_0,1},\cdots,H_{\mu_0,\nu_0,m},\cdots,H_{\mu_{o-1},\nu_{s-1},1},\cdots,H_{\mu_{o-1},\nu_{s-1},m}\right]$。



（a）　　（b）

图 3-3 V-LGBP表示中Jet上滤波器的选择

（a）相同尺度不同方向的滤波器 （b）相同方向不同尺度的滤波器

在基于“体”的局部Gabor二值模式表示中，Jet上的滤波器主要有两种选择：相

同方向不同尺度的滤波器与相同尺度不同方向的滤波器，分别如图 3-3（a）和（b）所示。这两种滤波器响应上的模式分别反映了图像的局部特征对尺度和方向的显著性。如图 3-1 所示，当相同方向不同尺度（或相同尺度不同方向）的滤波器响应来计算 Jet 上的模式时，“10”反映了图像上该点的局部结构与二进制位“1”对应的 Gabor 小波刻画的形状更吻合，即：该局部结构在对应尺度（或方向）上的 Gabor 小波的响应更显著。当计算块上的模式时，在 3×3 空间邻域内像素数目固定的条件下，不同的采样对结果影响不大，因此文中采用了均匀采样的方式。

同幅值的局部 Gabor 二值模式（即 LGBP_Mag）[150]相比，在图像块内的像素数目（即 P）取相同数值时，由于引入了 Jet 上的模式，基于“体”的局部 Gabor 二值模式（即 V-LGBP）表示中的模式数目（即 2^{P+L}）要大于前者中的模式数目（即 2^{P}）。这导致了幅值“体”上的表示具有更高的特征维数和更大的计算代价。幸运的是，通过调整空间域内的像素数目，V-LGBP 表示可以取得与 LGBP_Mag 表示相同的模式数目，如：{P=2，L=2}（V-LGBP），P=4（LGBP_Mag）。此时，V-LGBP 表示具有与 LGBP_Mag 表示几乎相同的计算代价，却表现出了更优的性能。原因在于，块内邻域像素的 Gabor 幅值响应之间具有比较大的相关性，这导致了块内的模式具有较高的冗余度；另一方面，Jet 上的模式与块内的模式有比较强的互补性，二者的组合反映了图像上更多的“微结构”。因此，幅值“体”的局部模式表示具有比较强的表示能力。

图像的 Gabor 小波表示包括了幅值、相位、实部和虚部四个部分。除幅值特征的局部 Gabor 二值模式表示外，文献中还存在一些利用其他 Gabor 特征的局部模式表示，如相位特征的局部 Gabor 二值模式（LGBP_Pha）[151]以及局部 Gabor 异或模式（LGXP）[129]，实部特征的局部 Gabor 相位模式（Re_LGPP）[142]和虚部特征的局部 Gabor 相位模式（Im_LGPP）[142]。与幅值“体”的局部二值模式类似，这些方法（如 LGBP_Pha，LGXP，Re_LGPP 和 Im_LGPP）也可以将局部模式的定义从图像块内扩展到相应的特征“体”上，即：将对应的 Gabor 特征组织为一个三维“体”，然后利用相应的局部算子（如 LBP 算子、异或算子）来计算该特征“体”上的模式。然而，本章的实验结果表明，这些方法在“体”上扩展表示的性能要低于其相应块上的模式表示。原因可能在于，不同于幅值特征，Gabor 实部、虚部和相位特征具有各自的特点：实部和虚部容易在物体的边缘附近产生振荡，相位则对图像的位置变化敏感。在这些特征构成的 Jet 上，利用 LBP 或异或算子计算的模式不能有效地刻画图像上的微结构。因而，这些表示在特征“体”上的扩展甚至没有块上的模式表示有效。

3.3 进一步扩展

下面对 Gabor 特征“体”的局部模式表示提出了两种扩展：利用 Fisher 准则的子块权重学习以及判别性特征提取，分别在第 3.3.1 节和第 3.3.2 节进行介绍。

3.3.1 利用 Fisher 准则的子块权重学习

大量的研究表明[1][150]，人脸的不同区域在识别过程中的作用是不同的：眼睛、眉毛、鼻子等区域要比额头、脸颊等区域对人脸识别的贡献更大。由于 Gabor 特征“体”的局部模式表示采用了分块的直方图表示，一种扩展的思路是根据训练集[142][150]或校验集[1]学习得到各个子块的权重，然后在计算图像相似度时根据这些权重进行加权融合。

本节根据 Fisher 准则从训练集上来学习得到图像上各个子块的权重[142][150]。该方法包括如下几个步骤：首先，将训练集中的多类样本转化为两类样本：同一人不同样本之间的相似度构成类内差样本，不同人样本之间的相似度构成类间差样本；然后，分别计算出类内差样本和类间差样本的均值和方差；最后，根据这两类样本的均值和方差按照 Fisher 准则计算得到每个子块的权重。

为了清楚起见，下面将子块权重的学习过程描述如下：假定 Gabor 特征“体”的局部模式图像 $\Re_{\mu,\nu}$（$\mu=\mu_0,\mu_1,\cdots,\mu_{o-1}$，$\nu=\nu_0,\nu_1,\cdots,\nu_{s-1}$）划分为 m 个子块 $\Re_{\mu,\nu,r}$（$r=1,2,\cdots,m$），根据如下等式分别计算子块 $\Re_{\mu,\nu,r}$ 在类内差和类间差样本集合上的均值 $m_{W(\mu,\nu,r)}$ 和 $m_{B(\mu,\nu,r)}$：

$$m_{W(\mu,\nu,r)}=\frac{1}{C}\sum_{i=1}^{C}\frac{1}{N_i(N_i-1)/2}\sum_{k=2}^{N_i}\sum_{l=1}^{k-1}S(H_{\mu,\nu,r}^{i,k},H_{\mu,\nu,r}^{i,l}) \tag{3-4}$$

$$m_{B(\mu,\nu,r)}=\frac{1}{C(C-1)/2}\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\frac{1}{N_iN_j}\sum_{k=1}^{N_i}\sum_{l=1}^{N_j}S(H_{\mu,\nu,r}^{i,k},H_{\mu,\nu,r}^{i,l}) \tag{3-5}$$

其中，C 表示训练集中的类别数目，N_i（$i=1,2,\cdots,C$）表示第 i 类包含的样本数目，$H_{\mu,\nu,r}^{i}$ 表示第 μ 个方向、第 ν 个尺度、第 r 个子块的直方图，$S(\bullet,\bullet)$ 表示两个样本根据其直方图特征计算的相似度（这里采用直方图交）。然后，根据 $m_{W(\mu,\nu,r)}$ 和 $m_{B(\mu,\nu,r)}$ 分别按照式（3-6）和（3-7）计算子块 $\Re_{\mu,\nu,r}$ 在类内差与类间差样本集合上的方差 $S_{W(\mu,\nu,r)}^2$ 和 $S_{B(\mu,\nu,r)}^2$：

$$S_{W(\mu,\nu,r)}^2=\sum_{i=1}^{C}\frac{1}{N_i(N_i-1)/2}\sum_{k=2}^{N_i}\sum_{l=1}^{k-1}(S(H_{\mu,\nu,r}^{i,k},H_{\mu,\nu,r}^{i,l})-m_{W(\mu,\nu,r)})^2 \tag{3-6}$$

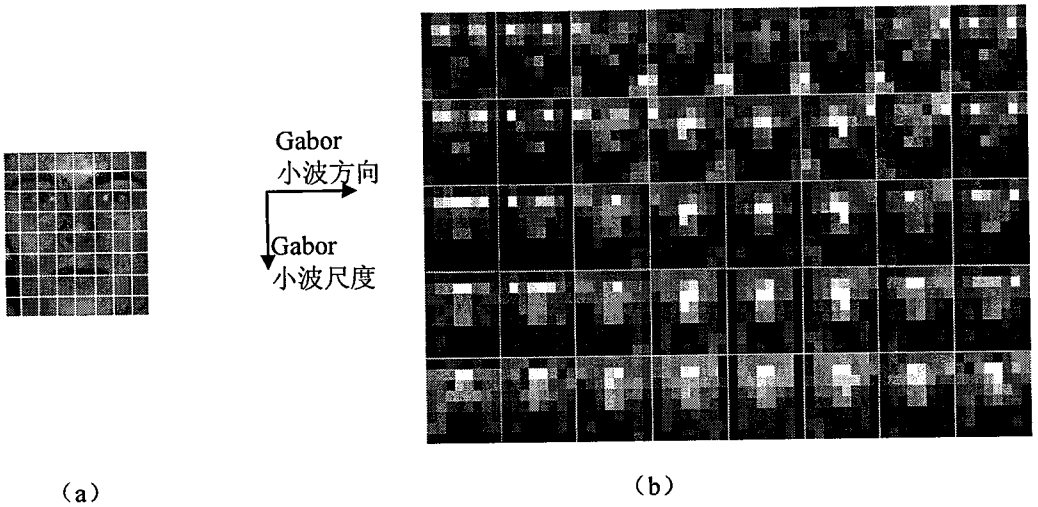
$$S_{B(\mu,\nu,r)}^2=\sum_{i=1}^{C-1}\sum_{j=i+1}^{C}\frac{1}{N_iN_j}\sum_{k=1}^{N_i}\sum_{l=1}^{N_j}(S(H_{\mu,\nu,r}^{i,k},H_{\mu,\nu,r}^{i,l})-m_{B(\mu,\nu,r)})^2 \tag{3-7}$$

最后，子块 $\Re_{\mu,\nu,r}$ 的权重 $w_{\mu,\nu,r}$ 按照下式计算得到：

$$w_{\mu,\nu,r}=\frac{(m_{W(\mu,\nu,r)}-m_{B(\mu,\nu,r)})^2}{S_{W(\mu,\nu,r)}^2+S_{B(\mu,\nu,r)}^2} \tag{3-8}$$

根据学习得到的子块权重 $w_{\mu,\nu,r}$（$\mu=\mu_0,\mu_1,\cdots,\mu_{o-1}$，$\nu=\nu_0,\nu_1,\cdots,\nu_{s-1}$，$r=1,2,\cdots,m$），图像 I_1 和 I_2 之间的相似度由其不同尺度、不同方向上子块间的相似度加权融合得到：

$$S(I_1,I_2)=\sum_{\mu=\mu_0}^{\mu_{o-1}}\sum_{\nu=\nu_0}^{\nu_{s-1}}\sum_{r=1}^{m}w_{\mu,\nu,r}\bullet S(H^1_{\mu,\nu,r},H^2_{\mu,\nu,r}) \tag{3-9}$$



（a） （b）

图3-4 根据FERET训练集学习得到的5个尺度、8个方向上各个子块的权重
（a）图像划分为64（即8×8）个子块 （b）子块权重示意图（子块越亮，其权重越大）

根据上面的子块权重学习方法，这里绘出了在 FERET 人脸库的训练集上（429 个人的 1,002 幅图像）学习得到的 Gabor 幅值“体”局部模式的各个子块权重，如图 3-4 所示。从图中可以看到，眉毛、眼睛和鼻子的区域的权重几乎在所有尺度、所有方向上都高于其他区域，这与预期的结果相一致；嘴巴区域的权重要小一些，主要原因在于该区域的形变比较大，导致了比较大的类内变化；此外，在小尺度的图像上（图 3-4（b）第一行），在下巴附近、靠近图像边界的区域分配了较大的权重，这可能与这些区域包含了人脸的轮廓信息，在区分不同人脸时起到了一定的作用有关。

3.3.2 判别性特征提取

由于采用了多尺度、多方向的 Gabor 小波，Gabor 特征“体”的局部模式表示具有高维的特征表示。对于 5 个尺度 8 个方向的 Gabor 小波，在图像划分为 64 个子块且模式数目设置为 16 种时，该特征“体”模式表示的特征维数高达 40,960(40×64×16)维。为了从高维表示中提取低维的判别性特征，本节采用了第二章 2.4.1 节的判别性特征提取（即 BFLD）方法。

图 3-5 示出了该判别性特征提取方法的特征提取流程。首先，计算图像 Gabor 特征“体”局部模式的分块直方图表示：每个尺度、每个方向的模式图像被划分为 M 个不重叠的块并将每个块划分为 K 个不重叠的子块，计算每个子块的直方图并将属于同一

个块的所有尺度、所有方向、各个子块的直方图连接起来得到每个块的直方图表示；然后，根据这 M 块的直方图特征在训练集上学习得到 M 个 FLD 投影矩阵 W_i（$i=1,2,\cdots,M$）；最后，计算图像上每个块的低维特征 F_i（$i=1,2,\cdots,M$）。详细的算法求解过程请参考第二章 2.4.1 节中算法 2.1 与 2.2。

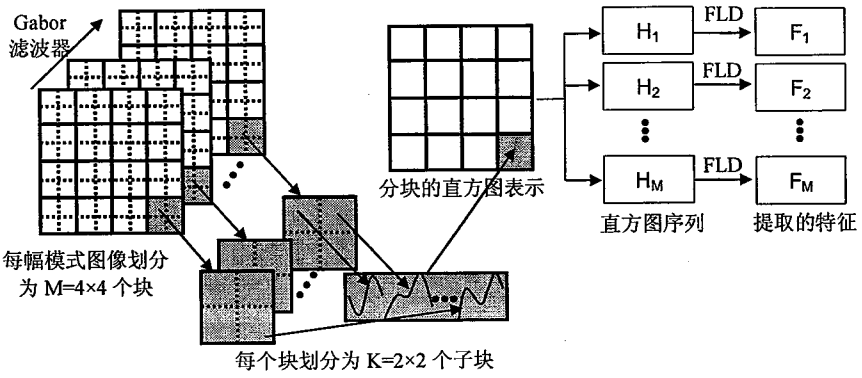


图 3-5 利用 BFLD 的特征提取过程

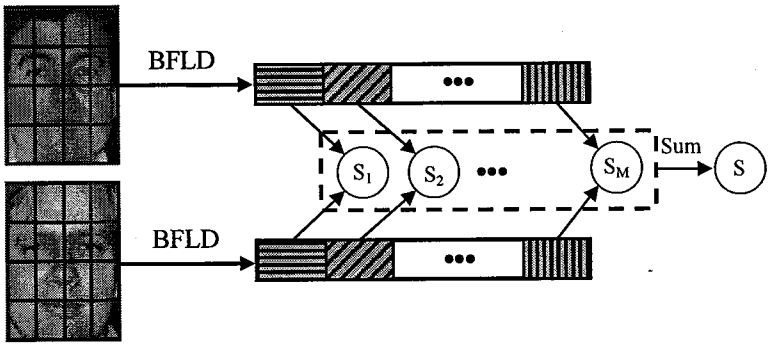


图 3-6 利用和规则的相似度计算

假定图像 I_1 和 I_2 根据该判别性特征提取方法分别得到其低维的特征 $F_1=[F_{1,1},F_{1,2},...,F_{1,M}]$ 和 $F_2=[F_{2,1},F_{2,2},...,F_{2,M}]$，则这两幅图像之间的相似度 $S(I_1,I_2)$ 通过和规则融合各个块的相似度得到，如图 3-6 所示：

$$S(I_1,I_2)=\frac{1}{M}\sum_{i=1}^{M}S\left(F_{1,i},F_{2,i}\right) \tag{3-10}$$

其中，$S\left(F_{1,i},F_{2,i}\right)$（$i=1,2,\cdots,M$）表示图像 I_1 和 I_2 第 i 个块的特征 $F_{1,i}$ 与 $F_{2,i}$ 之间的相似度，由等式（3-11）计算得到：

$$S\left(F_{1,i},F_{2,i}\right)_i=\frac{F_{1,i}\bullet F_{2,i}}{\left\|F_{1,i}\right\|\bullet\left\|F_{2,i}\right\|},\ i=1,2,...,M \tag{3-11}$$

进一步地，考虑到人脸的不同区域在识别过程中作用是不同的，这里为每个块分配一个反映其分类能力强弱的权重 w_i（$i=1,2,\cdots,M$），则两幅图像的相似度由各个块的相似度进行加权融合得到：

$$S(I_1,I_2)=\frac{1}{M}\sum_{i=1}^{M}w_i\bullet S\left(F_{1,i},F_{2,i}\right) \tag{3-12}$$

3.4 实验结果对比及分析

本部分采用 FERET、CAS-PEAL 和 FRGC 2.0 三个人脸数据库来测试 Gabor 特征“体”的局部模式表示及其扩展方法。相关内容安排如下：第 3.4.1 节介绍了实验设置；第 3.4.2 节以基于“体”的局部 Gabor 二值模式表示为例，测试了不同模式定义方法对结果的影响并给出了相应的分析；第 3.4.3 节对比不同 Gabor 特征“体”模式表示方法的结果并进行分析；第 3.4.4 节主要对比分析了基于“体”的局部 Gabor 二值模式表示与两种扩展方法的结果。

3.4.1 实验设置

由于第二章2.5.1节已经介绍了FERET和FRGC 2.0人脸库，这里将简要介绍CAS-PEAL人脸库。CAS-PEAL[29]人脸数据库是一个大规模的东方人脸库，包括1,040人（男：595，女：445）的99,594幅图像。目前发布的版本包含了1,040人的30,863幅图像，称为“CAS-PEAL-R1”人脸库。CAS-PEAL-R1人脸库中提供了三个标准的图像集用以测试不同识别算法的性能，即：训练集（Training Set）、注册集（Gallery）和测试集（Probe Set）。其中，训练集用来进行算法训练和参数学习，包含300人，每人从正面子集中选择了4幅图像；注册集是为待识别人建立的注册图像子集，包含1,040人，每人1幅标准正面图像；测试集是提供给算法进行识别的图像子集。对于正面子集，测试集包含除去训练集和注册集以外余下的6,992幅图像，分为饰物变化、表情变化、距离变化、背景变化、时间跨度和光照变化6个子集。图3-7（a）~（f）分别绘出了这六种变化下的示例图像。表3-1中列出了CAS-PEAL-R1人脸库中各个子集的统计情况。

在 FERET 和 CAS-PEAL-R1 人脸数据库上，所有人脸图像的大小被统一归一化到 80×88（宽度×高度）个像素（两眼中心分别设置为（21, 25）和（59, 25））。Gabor 小波的相关参数设置如下：5 个尺度 $\nu\in\{0,1,...,4\}$，8 个方向 $\mu\in\{0,1,...,7\}$，$\sigma=2\pi$，$k_{\max}=\pi$，$f=\sqrt{2}$。由于这两个数据库主要测试人脸识别的精度，实验中采用了最近邻分类器。在 FRGC 2.0 人脸库上，所有人脸图像的尺寸被统一归一化到 128×168 个像素（两眼中

心分别设置为(29, 65)和(100, 65))。Gabor小波参数k_{max}也进行相应的调整：$k_{max}=\frac{\pi}{2}$。由于FRGC 2.0数据库上进行人脸确认实验，下面的实验将通过遍历相似度阈值的范围来得到ROC曲线。此外，实验中不采用任何的光照预处理方法。

(a) (b) (c) (d) (e) (f)

图3-7 CAS-PEAL-R1正面人脸库示例图像

(a) 饰物变化 (b) 表情变化 (c) 距离变化 (d) 背景变化 (e) 时间跨度 (f) 光照变化

表3-1 CAS-PEAL-R1人脸库中各个子集的情况统计

图像类别	子集	变化种类	人数	图像数
正面	标准	1	1,040	1,040
	表情	5[a]	377	1,884
	饰物	6	438	2,616
	光照	>=9	233	2,450
	时间	1	66	66
	背景	2-4	297	651
	距离	1-2	296	324
	共计:			9,031
非正面		21(3*7)	1,040	21,832[b]
共计:				30,863

[a]中性表情没有包括在内

[b]8幅图像已被损坏

3.4.2 不同模式定义对实验结果的影响

在基于"体"的局部Gabor二值模式方法(即V-LGBP)中，"体"上的模式包括块内的模式和Gabor Jet上的模式。在计算Jet上的模式时，滤波器主要有两种选择：相同方向不同尺度的滤波器与相同尺度不同方向的滤波器(参见图3-3)。为了区分不同模式定义的方法，这里采用了如下的符号表示："JetLBP_Scale"与"V-LGBP_Scale"方法表示Jet上的模式根据相同方向不同尺度的滤波器响应计算得到；"JetLBP_Orient"与"V-LGBP_Orient"方法表示Jet上的模式是由相同尺度不同方向的滤波器响应计算得到。当块内的邻域像素数目P和Jet上的邻域大小L取不同数值时，"体"上的模式(即V-LGBP)、Jet上的模式(即JetLBP)与块上的模式表示(即LGBP_Mag)方法之间存在如下关系：当P=0，L>0时(即只根据Jet上的响应计算局部模式)，"V-LGBP_Scale"与"V-LGBP_Orient"方法分别等价于"JetLBP_Scale"与"JetLBP_Orient"方法；当

P>0, L=0 时（即只根据块上的响应计算局部模式），“V-LGBP_Scale”与“V-LGBP_Orient”方法都等价于图像块上的模式表示（即 LGBP_Mag）。

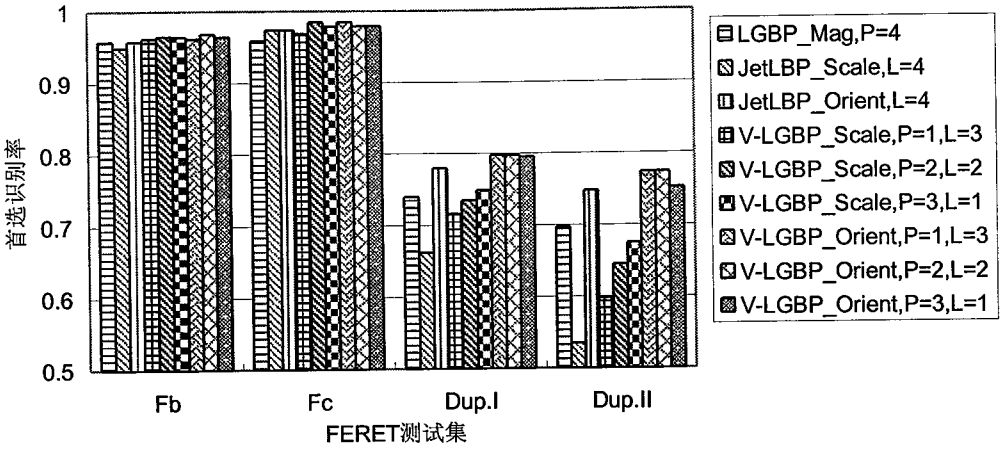


图 3-8 不同模式定义方法在 FERET 人脸库上的结果对比

图 3-8 示出了在模式数目设置为 16（即 P+L=4）种时不同模式定义方法的结果对比（每幅图像划分为 64 个子块）。可以看到，在模式数目相等的条件下，块上的模式（即 LGBP_Mag）、Jet 上的模式（即 JetLBP）和“体”上的模式（即 V-LGBP）表示方法在 Fb 和 Fc 子集上的性能相差不大，但在 Dup.I 和 Dup.II 子集上的结果呈现了明显的差异。对比 Jet 上的模式表示方法在 Dup.I 和 Dup.II 子集上的结果，JetLBP_Orient 方法的结果要远远高于 JetLBP_Scale 方法的结果；同样的关系也存在于方法 V-LGBP_Orient 和 V-LGBP_Scale 的结果之间。这表明相同尺度不同方向的滤波器响应之间的模式要比相同方向不同尺度的滤波器响应上的模式更适合于人脸表示。原因可能在于，5 个尺度 8 个方向的 Gabor 小波在方向上的粒度划分要比在尺度上的粒度划分更精细。因而，定义在前者上的模式表示更有效地反映了人脸图像上的微结构。从图中还可以看到，在这四个测试子集上，Jet 上的模式（即 JetLBP_Orient）表示方法可以取得与块上的模式（即 LGBP_Mag）表示方法相当甚至更好的结果。此外，与第 3.2 节所分析的一致，体上的模式表示方法（即 V-LGBP_Orient）组合了 Jet 上的模式与块上的模式，即使在模式数目相同的条件下也能够反映更多的信息，因此取得了比 Jet 上和块上的模式表示方法更优的性能。

3.4.3 不同 Gabor 特征块与“体”上模式表示的结果对比

为了对比不同 Gabor 特征“体”的模式表示，本节将 Gabor 实部、虚部和相位特征的局部模式表示分别扩展到对应的特征“体”上，然后利用相应的局部算子来计算“体”上的模式。具体地讲，针对如下四种 Gabor 特征的局部模式：实部的局部 Gabor 相位模式（即 Re_LGPP）、虚部的局部 Gabor 相位模式（即 Im_LGPP）、相位的局部 Gabor 二

值模式（即 LGBP_Pha）和相位的局部 Gabor 异或模式（即 LGXP），这里将它们的模式定义分别扩展到相应的 Gabor 实部、虚部和相位特征“体”上，并将这些“体”上的模式表示方法分别记作：“V-Re_LGPP”，“V-Im_LGPP”，“V-LGBP_Pha”和“V-LGXP”。

表 3-2 不同 Gabor 特征块与“体”上的模式表示在 FERET 人脸库上结果对比（%）

Gabor 特征	方法	FERET 测试集			
		Fb	Fc	Dup.I	Dup.II
实部	Re_LGPP	**98.3**	**99.0**	**80.1**	**77.8**
	V-Re_LGPP	97.6	97.9	79.1	77.4
虚部	Im_LGPP	**98.1**	**99.0**	75.6	72.2
	V-Im_LGPP	97.2	97.9	**78.0**	**76.5**
相位	LGBP_Pha	**97.4**	97.9	**77.8**	**75.2**
	V-LGBP_Pha	97.2	**99.0**	75.9	74.4
相位	LGXP	**98.3**	**100.0**	**81.7**	**82.1**
	V-LGXP	97.8	98.5	79.5	79.5
幅值	LGBP_Mag	95.8	95.9	74.0	69.7
	V-LGBP	**96.9**	**97.9**	**79.8**	**77.4**

表 3-2 中列出了不同 Gabor 特征块与“体”上的模式表示方法在 FERET 人脸库上的结果对比。这些方法的参数设置如下：每幅图像划分为 64（即 8×8）个子块；模式数目设置为 16 种（P=2，L=2）；由于相同尺度不同方向滤波器响应上的模式更有效，因此被用来计算 Jet 上的模式。从表中可以看到，除基于“体”的局部 Gabor 二值模式表示（即 V-LGBP）之外，其他 Gabor 特征“体”局部模式表示在性能上大多要低于相应“块”上的表示方法。这基本上与第 3.2 节的分析一致：Gabor 实部、虚部和相位特征不能在物体的边缘附近产生显著的响应，因而，在这些特征构成 Jet 上的模式不能有效地刻画图像上的微结构；与之相反，幅值特征反映了人脸图像上的显著性特征，在幅值响应构成的 Jet 上的模式表示是一种对尺度或方向的显著性度量，因而其与块内模式的组合具有更强的表示能力。所以，基于“体”的局部 Gabor 二值模式表示方法取得了比块上的模式表示方法更好的结果。

本节进一步在 CAS-PEAL-R1 人脸库上对 Gabor 特征块与“体”上的模式表示方法的结果进行了对比。表 3-3 中列出了这些方法分别在饰物、表情和光照三个子集上的识别结果。其中，“V-Re_LGPP”、“V-Im-LGPP”、“V-LGBP_Pha”和“V-LGXP”等方法的含义与表 3-2 中相同符号的含义一致，并采用了与表 3-2 中相同的参数设置。从表 3-3 中可以得到与表 3-2 中一致的观察：在饰物、表情和光照子集上，幅值“体”上局部模式表示（即 V-LGBP）的结果比块上的局部模式表示（即 LGBP_Mag）分别提高了 4.7%、5.3%和 10.6%，而其他特征“体”的表示方法大多呈现了比块上的模式表示方法更低的精度。此外，从表 3-2 和表 3-2 中可以看到，相位特征的局部 Gabor 异或模式表示（即 LGXP）在这两个数据库上都取得了最优的结果。

表 3-3 不同 Gabor 特征块与“体”上的模式表示在 CAS-PEAL-R1 人脸库的结果对比（%）

Gabor 特征	方法	CAS-PEAL-R1 测试集		
		饰物	表情	光照
实部	Re_LGPP	**93.5**	**98.1**	**64.0**
	V-Re_LGPP	91.9	97.1	60.7
虚部	Im_LGPP	**92.6**	**97.7**	50.1
	V-Im_LGPP	91.9	97.2	**53.6**
相位	LGBP_Pha	**89.8**	**96.2**	**48.6**
	V-LGBP_Pha	90.5	96.1	43.1
相位	LGXP	**94.0**	**98.6**	**65.8**
	V-LGXP	92.6	97.7	52.8
幅值	LGBP_Mag	88.3	92.6	49.6
	V-LGBP	**93.0**	**97.9**	**61.2**

3.4.4 Gabor 特征“体”局部模式表示与两种扩展方法的结果对比

本节将在 FERET、CAS-PEAL-R1 和 FRGC 2.0 三个人脸数据库上对 Gabor 特征“体”的局部模式表示及它的两种扩展方法进行对比分析：利用 Fisher 准则的子块权重学习和判别性特征提取（即 BFLD）。根据第 3.4.3 节的结果可知，与其他 Gabor 特征的局部模式表示相比，基于“体”的局部 Gabor 二值模式表示（即 V-LGBP）比图像块上的模式表示更有效，因而这里主要关注该方法。为了简洁起见，下面将这两种扩展方法分别记为“子块加权方法”和“结合 BFLD 的表示方法”，并用“基准方法”指代基于“体”的局部 Gabor 二值模式（或其他局部模式）表示方法。为了学习各个子块的权重以及判别性特征提取方法的投影矩阵，在这三个数据库上分别利用了如下的训练集，即：从 FERET 人脸库的训练光盘中选择的 1,002 幅图像（429 个人）、CAS-PEAL-R1 人脸库提供的 1,200 幅图像（300 个人）和 FRGC 2.0 人脸库提供的 12,776 幅图像（222 个人）。

1. FERET 人脸数据库上的结果对比

本小节在 FERET 人脸库上测试 V-LGBP 方法及其两种扩展方法。相关参数设置如下：在 V-LGBP 方法中，模式数目设置为 16 种（即 P=2，L=2），每幅图像划分为 64 个子块，在计算 Jet 上的模式时采用了相同尺度不同方向的滤波器响应；在块加权方法中，它需要根据训练集学习得到 5 个尺度、8 个方向上 64 个子块的权重；在 V-LGBP 表示结合 BFLD 的方法中，模式数目设置为 8 种（P=2，L=1），每幅图像划分为 16 个块（即 4×4），每个块划分为 4 个子块（即 2×2），每个块的特征维数（即 40×4×8=1,280 维）利用 FLD 降至 200 维，所有块上的相似度按照和规则（式(3-10)）融合得到图像间的相似度。在 V-LGBP 表示结合 BFLD 的方法中，选择 8 种模式数目的原因是避免每个块具有过高的特征维数，从而减少学习 FLD 投影矩阵的计算代价。

图 3-9 中绘出了 V-LGBP 表示及其两种扩展方法在 FERET 测试集上的识别结果对比。可以看到，同 V-LGBP 方法相比，它的两种扩展的方法在结果上有了比较大的改进：

在 Dup.I 和 Dup.II 两个子集上，块加权方法比基准方法在结果上分别提高了 2%和 4%；在 Fb、Fc、Dup.I、Dup.II 这四个子集上，结合 BFLD 的表示方法比基准方法分别提高了 3%、2%、13%、13%。这说明了两种扩展方法在该测试集上的有效性。

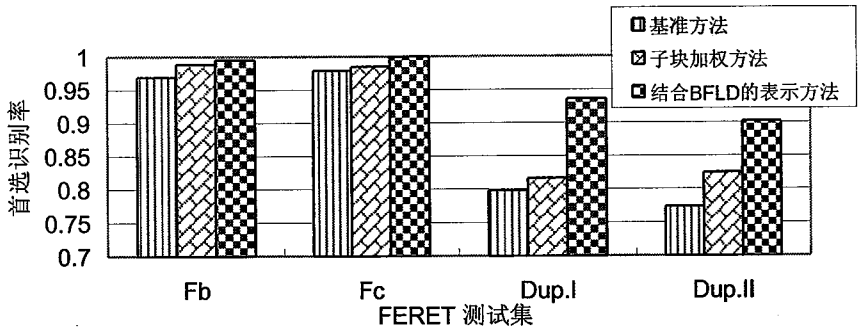


图 3-9 V-LGBP 表示及其两种扩展方法在 FERET 人脸库上的结果对比

与基准的 V-LGBP 表示方法相比，子块加权方法和结合 BFLD 的表示方法都需要一个离线的学习过程，这确实增加了一些计算代价：子块加权方法需要学习每个子块的权重；BFLD 方法需要在训练集上学习每个块对应的 FLD 矩阵。当这两种方法在线应用时，子块加权方法具有与基准方法基本相当的计算代价，而结合 BFLD 的表示方法由于提取了低维的特征，其图像相似度的计算代价甚至低于基准的 V-LGBP 表示方法。比较而言，判别性特征提取方法是一种更合适的扩展方法。

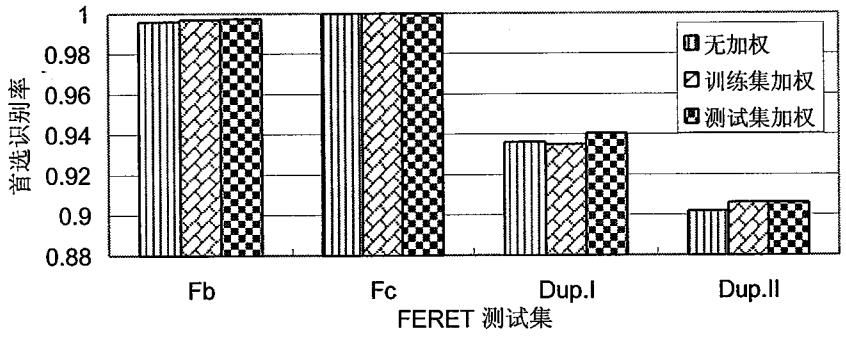
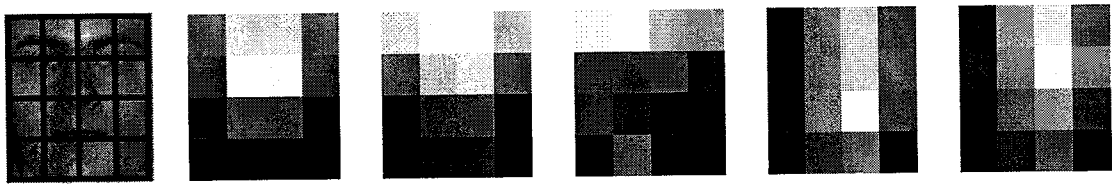


图 3-10 V-LGBP 与 BFLD 结合的表示方法在 FERET 人脸库上不同加权策略下的结果对比

在 V-LGBP 结合 BFLD 的表示方法中，一种改进的策略是按照加权和规则来融合不同块的相似度（参见等式（3-12））。这里采取了两种策略来得到每个块的权重：第一种策略与利用 Fisher 准则的子块权重学习方法类似，不同之处在于类内差与类间差样本是根据每个块经 FLD 降维后的特征利用余弦相似度计算得到（参见等式（3-11）），称之为“训练集加权”；第二种策略是计算每个块在测试集合上的识别率，并将识别率转化为每个块的权重，称之为“测试集加权”。显然，在第二种策略中，每个块在不同测试集合上的识别率会发生变化，权重也相应地发生变化。

图 3-10 示出了 V-LGBP 与 BFLD 结合的表示方法在无加权策略下以及结合“训练集加权”、“测试集加权”策略下的结果。从图中可以看到，对于结合 BFLD 的表示方法，这两种加权策略下的结果与无加权方法的结果差异很小。这主要有两方面的原因：第一，V-LGBP 结合 BFLD 的表示方法在这些测试集上已经取得了很高的精度，这导致了识别性能提升的空间不大；第二，在该方法中，图像块的数目（即 16）远远小于 V-LGBP 表示中的子块数目（即 5×8×64），这使得加权策略的性能提升空间较小。因此，对局部 Gabor 模式与 BFLD 结合的方法而言，利用加权和规则的相似度融合策略很难在识别结果上有大幅度的提升。



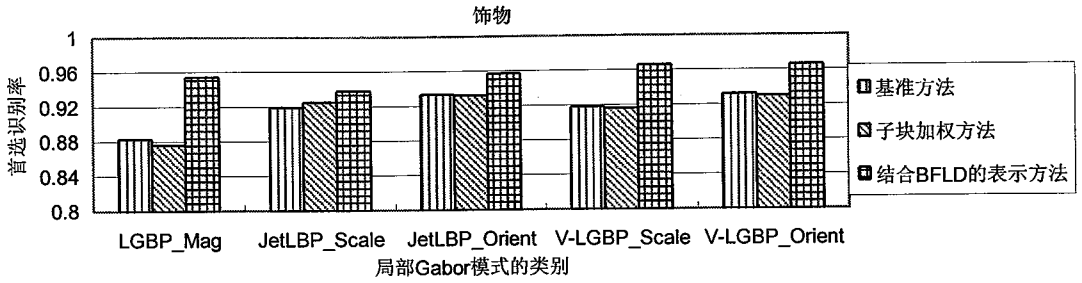
（a） （b） （c） （d） （e） （f）

图 3-11 不同加权策略的权重可视化结果（区域越亮，其权重越大）
（a）图像划分为 16 个块 （b）从训练集学习得到的权重 （c）~（f）分别是根据 Fb、Fc、Dup.I 和 Dup.II 子集上的分类精度得到的权重

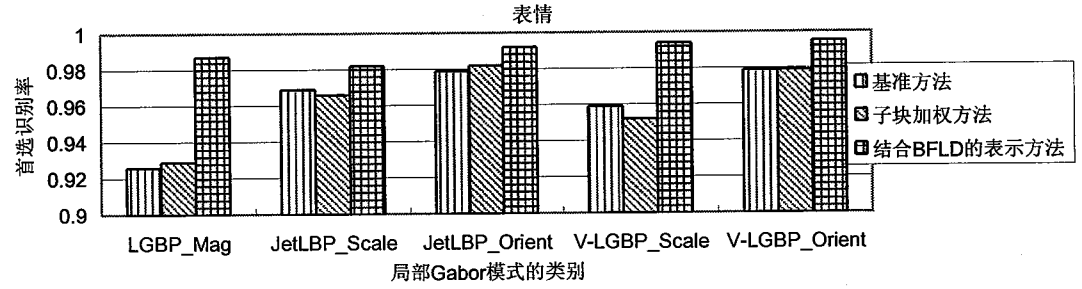
图 3-11 绘出了在图像划分为 16 个块时，V-LGBP 结合 BFLD 的表示方法在“训练集加权”和“测试集加权”这两种策略下的权重可视化结果。从图中可以看到，在训练集学习得到的人脸区域的权重中（图 3-11(b)），权重最大的前 4 个区域包含了眼睛、鼻子、眉毛等区域，这表明了这些区域在训练集上具有比较大的可分性；在不同测试集上（图 3-11(c)~(f)），各个块的权重呈现了不一致的变化趋势：在 Fb、Dup.I 和 Dup.II 子集上，权重较大（即精度较高）的区域包括眼睛、眉毛、鼻子等区域，而 Fc 子集上眉毛、额头等区域赋予了较大的权重。

2. CAS-PEAL-R1 人脸数据库上的结果对比

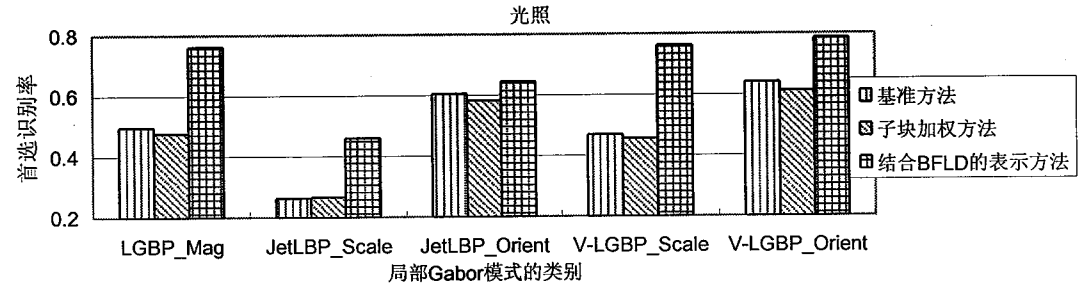
本小节在 CAS-PEAL-R1 人脸库上对 Gabor 幅值特征“块”上的模式（即 LGBP_Mag）、Jet 上的模式（即 JetLBP）和“体”上的模式（即 V-LGBP）表示方法进行测试，并进一步对比了它们在两种扩展策略下的结果。这里采用了与第 3.4.2 节相同的符号来区分不同的方法。相关参数设置如下：在基准方法和子块加权方法中，每幅图像被划分为 64（即 8×8）个子块，模式数目设置为 16 种：P=4（LGBP_Mag），L=4（JetLBP_Scale, JetLBP_Orient），{P=2，L=2}（V-LGBP_Scale，V-LGBP_Orient）；在结合 BFLD 的表示方法中，图像划分为 16（即 4×4）个块，每个块划分为 4（即 2×2）个子块，局部模式数目设置为 8 种：P=3（LGBP），L=3（JetLBP_Scale, JetLBP_Orient），{P=2，L=1}（V-LGBP_Scale，V-LGBP_Orient），每个块的特征维数利用 FLD 降到 200 维，并利用和规则来融合各个块上的相似度。



（a）



（b）



（c）

图 3-12 不同方法在 CAS-PEAL-R1 人脸库上的结果对比

图 3-12（a）~（c）分别绘出了不同模式表示方法及其两种扩展方法在饰物、表情和光照子集上的测试结果。从图中可以得到如下观察：第一，在基准方法中，Jet 上的模式表示方法（如 JetLBP_Orient）的结果能够达到甚至超过了图像块内的模式表示方法（即 LGBP_Mag），Gabor 幅值“体”上的模式则在这三个子集上都达到了最好的识别结果；第二，不同模式表示的子块加权方法在饰物和表情子集上的结果与基准方法的结果相比基本没有提高，在光照子集上某些模式表示（如 JetLBP_Orient, V-LGBP_Orient）的块加权方法甚至低于基准的局部 Gabor 模式方法的结果。原因可能在于，CAS-PEAL-R1 人脸库提供的训练集涵盖了较多的图像变化而每种变化下的图像数目较少。因此，基于训练集学习得到的权重不能提高只包含某种特定变化（如光照）的测试集上的分类精度；第三，与基准方法相比，局部 Gabor 模式结合 BFLD 的表示方法在这三个子集上的结果都有了比较大的提升。特别是在光照子集上，很多方法的结果都能够提升接近 20%。此外，在所有与 BFLD 结合的表示方法中，基于“体”的局部

Gabor 二值模式表示取得了最优的结果。

3. FRGC 2.0 人脸数据库上的结果对比

从 FERET 和 CAS-PEAL-R1 人脸库上的结果对比及分析可知，相比子块加权方法，结合判别性特征提取（即 BFLD）的方法具有更低的相似度计算代价以及更高的识别性能。因此，这里将只测试不同模式表示及其与 BFLD 结合的方法。相关参数设置如下：在基准方法中，图像划分为 64（即 8×8）个子块，相同尺度不同方向的滤波器响应用来计算 Jet 上的模式，模式数目设置为 16 种，即：P=4（LGBP_Mag），L=4（JetLBP），{ P=2, L=2}（V-LGBP）；在其结合 BFLD 的表示方法中，图像划分为 16（即 4×4）个块，每个块划分为 8（即 2×4）个子块，模式数目设置为 8 种，即：P=3（LGBP_Mag），L=3（JetLBP），{P=2, L=1}（V-LGBP），每个块的特征维数利用 FLD 降到 200 维。

(1) 相似度归一化对结果的影响

在人脸确认问题中，不同测试图像在注册图像集合上的相似度分布范围是不同的。在设定统一阈值的条件下，这一方面将导致相似度值较小的某些测试图像容易被拒识，另一方面会导致相似度值较大的某些测试图像容易被错误接受。因此，这里采用了“z-score”[45]的方法对不同测试图像的相似度进行归一化，即：根据测试图像 Q 和 N 幅注册图像 $T_i\,(i=1,2,...,N)$ 的相似度向量 $[S_1,S_2,...,S_N]$，计算该相似度向量的均值 μ 和标准差 σ；然后，每个相似度 S_i 根据均值 μ 和标准差 σ 进行归一化：

$$S_i^* = \frac{S_i - \mu}{\sigma},\quad i=1,2,...,N \tag{3-13}$$

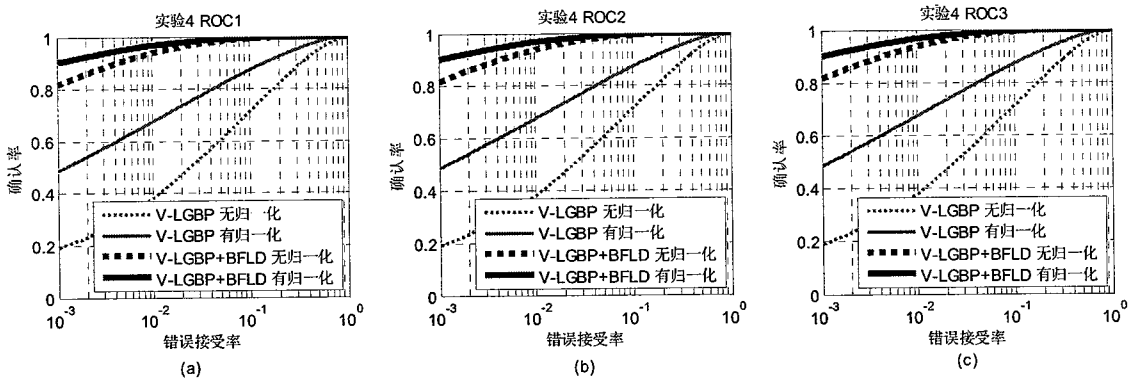


图 3-13 相似度归一化对不同方法在 FRGC 2.0 人脸库上的结果影响

经过归一化操作，不同测试图像的相似度范围将位于近似一致的区间内。然而，如果注册图像集合中并不包括与测试图像属于同一类的样本，计算得到的均值和标准差不能准确地反映该测试图像的相似度分布范围，这将导致比较高的错误接受率。尽管如此，该归一化操作极大地提高了局部 Gabor 模式及其与 BFLD 结合的方法在 FRGC 2.0 人脸库上的确认性能。如图 3-13 所示，在错误接受率等于 0.1%时，经相似度归一化操作后，

Gabor幅值“体”的模式表示方法（即V-LGBP）在ROC1、ROC2、ROC3上的确认率分别从19.22%、19.27%、19.35%提高到48.52%、48.53%、48.50%，而其与BFLD结合的方法（即“V-LGBP+BFLD”）的确认率则分别从81.63%、82.56%、83.39%提高到90.15%、89.99%、89.78%。从图中也注意到，不同方法在ROC1、ROC2和ROC3上的结果变化趋势基本上是一致的。因此，下面的实验将主要报告ROC3上采用了相似度归一化操作后的实验结果。

(2) 不同方法的结果对比

为了进一步比较不同局部Gabor模式之间的差异，图3-14（a）和（b）分别绘出了Gabor幅值块上的模式（即LGBP_Mag）、Jet上的模式（即JetLBP）与“体”上的模式（即V-LGBP）表示方法以及它们结合BFLD方法时的结果对比。从图（a）中可以看到，在基准的局部Gabor模式方法中，“体”上模式表示方法的精度要优于“Jet”上的模式表示，“块”上模式表示方法的性能最低；当结合BFLD方法时，如图（b）所示，“体”上模式表示方法的精度要优于“块”上的模式表示，并均优于“Jet”上的模式表示。但无论在哪种情况下，“体”上的模式表示方法都取得了最优的结果。

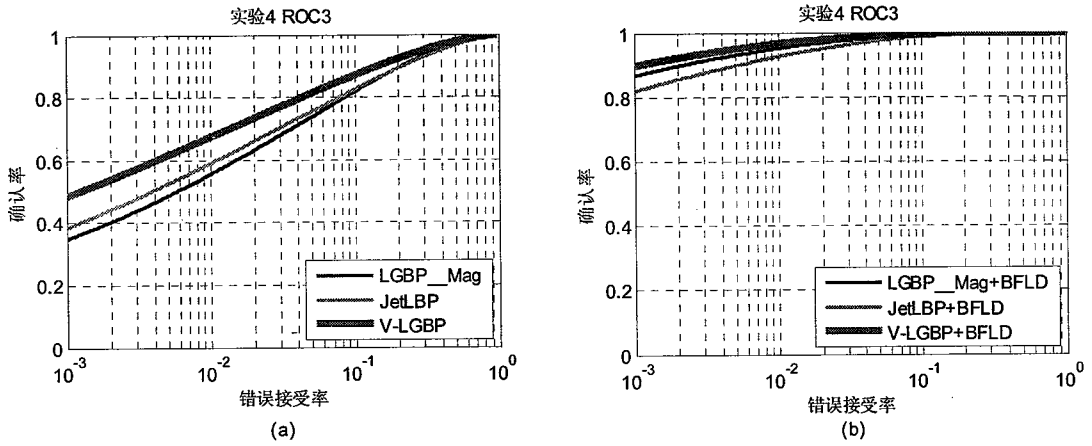


图3-14 不同方法在FRGC 2.0实验4上的结果对比

（a）基准方法（b）局部Gabor模式与BFLD结合的方法

表3-4 不同方法在FRGC 2.0人脸库上的确认率（错误接受率=0.1%）对比（%）

方法	实验4		
	ROC1	ROC2	ROC3
Liu和Liu的方法[77]	82.3	81.9	81.3
Tan和Triggs的方法[117]	N/A	N/A	83.6
Deng等的方法[23]	**93.9**	**93.6**	**93.1**
V-LGBP与BFLD结合的方法	90.1	90.0	89.8

表3-4中列出了基于“体”的局部Gabor二值模式表示结合BFLD的方法与其他方法的结果对比。表中其余的三种方法都利用了两种或多种特征：Liu和Liu的方法[77]

利用了不同人脸模板的彩色图像信息和 DCT 特征；Tan 和 Triggs 的方法[117]利用了 Gabor 特征和 LBP 特征；Deng 等的方法[23]利用了不同人脸模板的彩色信息和 Gabor 小波表示。表中这三种方法的结果直接摘自相应的参考文献。可以看到，本章利用单一特征的方法取得了比文献[77]和[117]中方法更好的结果，但是其结果要低于文献[23]中的方法。尽管如此，如果 Gabor 幅值“体”的模式表示能够与其他类型的特征进行融合，识别精度仍可能有一定的提升空间。

3.5 本章小结

为了获得图像 Gabor 特征上高效的局部模式表示，本章将多尺度、多方向的 Gabor 特征组织为一个三维“体”，然后利用 LBP 算子对幅值“体”上的局部模式进行建模表示。实验结果表明，Gabor 幅值“体”的模式表示反映了人脸图像上更多的“微结构”，从而能够取得更好的分类性能；幅值特征构成的 Jet 上的模式表示刻画了图像上对尺度或方向显著的局部特征，其分类精度能够达到甚至超过块上的模式表示。

本章研究了其他 Gabor 特征（即实部、虚部、相位）的局部模式扩展到“体”上的表示方法。在 FERET 和 CAS-PEAL-R1 人脸库的实验结果表明，与块上的局部模式表示相比，这些“体”上的表示方法取得了较低的结果。原因可能在于，这些特征（即实部、虚部、相位）不能像幅值特征一样来反映人脸图像的显著性特征，因而 Jet 上的模式表示不能够有效地刻画人脸图像的微结构。

本章进一步地研究了 Gabor 特征“体”局部模式表示的两种扩展方法：利用 Fisher 准则的子块权重学习以及判别性特征提取。在 FERET 和 CAS-PEAL-R1 人脸库上的实验结果表明，子块加权方法并不能一致地取得比 Gabor 特征“体”局部模式表示方法更好的结果：当训练集与测试集中的样本呈现相对一致的变化类型时，该方法能够提高识别精度（如 FERET 人脸库）；相反，如果训练集中涵盖了多种变化时，该方法不能提高只包含某种特定变化的测试集上的性能，甚至出现下降的现象（如 CAS-PEAL-R1 人脸库）。对于判别性特征提取方法，只要训练集与测试集中的样本是同质的（如 FERET 训练集 vs FERET 测试集），它对识别性能的提高总有帮助。特别地，在富有挑战性的 FRGC 2.0 人脸库的实验 4 上，基于“体”的局部 Gabor 二值模式与判别性特征提取结合的方法取得了优异的性能（当错误接受率=0.1%时确认率接近 90%）。

第四章　人脸局部 Gabor 模式的学习

4.1 引言

LBP表示由于其简单、有效的特点而在人脸识别中受到了广泛的关注[1][68][65][88][102][116]。它通过人工设计的算子（即 LBP）来计算所有类别（如人脸、纹理）图像上的模式，并不区分不同种类的图像。例如，在 3×3 邻域上的 LBP 算子根据中心像素的亮度值对邻域像素执行阈值化操作，并将得到的“0”、“1”等二进制位组合为一个二进制串（对应为一种模式）。因此，在所有图像的 3×3 邻域上，LBP 表示的模式数目共计 256（即 2^8）种。换句话说，对于任意图像上 3×3 大小的块，LBP 表示有一个包含 256 种码字（即“模式”）的“码本”。在计算图像上的局部模式时，LBP 表示根据该预定义的码本将所有 3×3 大小的图像块映射为某种模式（即“码字”）。

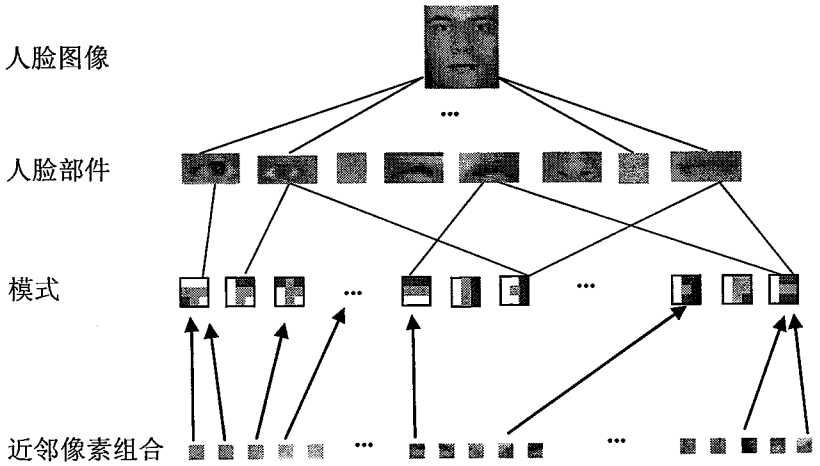


图 4-1 人脸图像的层级结构

人脸具有非常相似的结构特征，如图 4-1 所示，人脸图像的构成可看作是一个由近邻像素组合、模式、人脸部件等构成的层级结构。具体地，人脸图像是由一些人脸部件（如眉毛、眼睛、鼻子和嘴巴等）组成，这些部件进一步分解为一些模式。这里的模式与 LBP 表示中的模式具有类似的含义，是构成具有不同变化的图像块（即近邻像素的组合）的基本元素。在某种意义上，它们可看作介于单个像素和人脸部件之间的“中间层”表示。这意味这它们的语义信息要比单个像素强，却又弱于人脸部件。此外，人脸图像可以利用少数的面部特征点（如两只眼睛的中心位置）进行归一化，这种特性减弱了尺度、旋转、平移等变化对图像块的影响，从而减少了模式的种类。

人脸作为一类结构非常相似的物体，应当有其专属的模式构成。从这一角度出发，文献[83]提出了局部视觉基元（Local Visual Primitives, LVP）的表示方法。该方法通过对大量图像块的聚类学习得到构成人脸区域的基元（即 LVP），并利用这些基元来进行人脸图像的重构与识别。本章进一步扩展了该项工作，主要围绕着人脸图像的 Gabor 特征从模式学习的角度来获得人脸表示的基元，称为“基于学习的局部 Gabor 模式”（Learned Local Gabor Patterns, LLGP）。不同于第二章和第三章中基于设计的模式表示方法，本章的方法侧重于模式的学习。此外，考虑到人脸的不同区域以及不同模式在识别中的不同作用，文中提出了利用二者加权的识别方法。

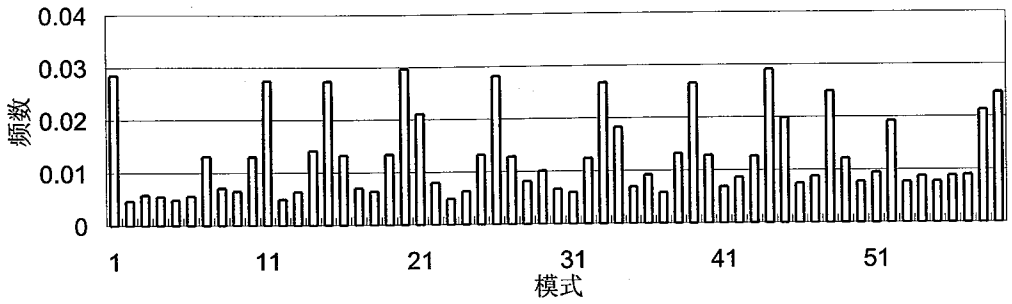
目前，图像 Gabor 特征的局部模式表示方法在人脸识别中吸引了一些研究者的目光[35][62] [127][128][129][142][150][151][155]。为了简洁起见，本章将这类方法统称为局部 Gabor 模式表示方法。为了弄清楚这些方法之间的差异，文中按照模式的定义方式将它们划分为两大类，并从模式定义流程以及方法参数这两个方面进行对比分析。由于这些表示具有高维的特征，文中采用了判别性特征提取方法来提取低维的特征。

本章内容安排如下：第 4.2 节主要介绍了人脸局部 Gabor 模式的学习及特征表示；第 4.3 节提出了利用块和模式权重的加权方法；第 4.4 节系统地对比了不同局部 Gabor 模式表示方法；第 4.5 节是实验部分，在三个大规模人脸库上测试了 LLGP 方法以及不同局部 Gabor 模式方法的结果，并分析讨论了一些有趣的问题；第 4.6 节总结了本章的工作并指出了后续的改进方向。

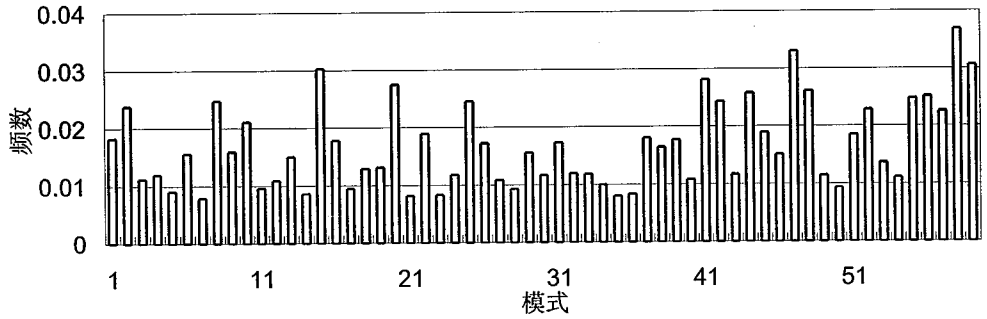
4.2 局部 Gabor 模式的学习及特征表示

基于学习的局部 Gabor 模式方法（即 LLGP）可看作局部视觉基元方法（即 LVP）在图像 Gabor 特征上的一种扩展。不同于 LBP/LGBP 方法，LVP/LLGP 方法需要一个学习过程来获得码本。由于码本从人脸图像上学习得到，它包含了构成人脸图像上的基本模式，是一种专门针对人脸图像的表示。图 4-2 示出了在 FERET 人脸库上统计的 LBP 与 LVP 表示的模式频数。可以看到，在 LBP 表示中，某些模式在人脸图像上很少出现，其频数比例要低于 1%；在 LVP 表示中，尽管不同模式的频数之间存在差异，但大多数模式的频数比例要高于 1%。因而，从人脸图像上学习得到的模式更多地刻画了人脸图像上常见的特征。

与 LVP 方法类似，LLGP 方法主要包括两个阶段：学习阶段和表示阶段，如图 4-3（a）和（b）所示。具体地讲，学习阶段根据训练集得到对应不同 Gabor 小波的多个码本：对每个尺度、每个方向的 Gabor 小波，从训练集图像的 Gabor 特征上采样图像块来构建相应的图像块集合，然后利用聚类算法从图像块的集合上学习得到相应的码本；表示阶段根据学习阶段产生的码本来表示人脸图像：对输入人脸图像提取不同尺度、不同方向的 Gabor 特征，然后根据码本计算该图像的 Gabor 特征在不同尺度、不同方向上的模式图像，最后利用分块的直方图来表示。下面将对这两个阶段分别进行详细的介绍。



（a）



（b）

图 4-2 LBP 与 LVP 表示的模式频数统计 （a）LBP （b）LVP

1. 学习阶段

LLGP 方法根据人脸图像的 Gabor 小波表示、以图像块为单位学习得到码本。对于 T 个 Gabor 小波，它将得到 T 个对应的码本。针对第 ν 个尺度、第 μ 个方向的 Gabor 小波，其码本的构建过程如下：

(1) 对训练集中的每幅图像 I，计算它的 Gabor 小波表示 $G_{\mu,\nu}(I)$。为了减少平移、旋转、尺度等变化对图像块的影响，所有人脸图像根据两只眼睛的中心位置归一化到某个固定的尺寸（如 64×80 个像素）。

(2) 从训练集中每幅图像的 Gabor 小波表示 $G_{\mu,\nu}(I)$ 中采样大小为 m×n 的图像块，这些图像块构成集合 $S_{\mu,\nu}$。如果训练集包括 N 幅图像，每幅图像上采样得到 k 个图像块，那么集合 $S_{\mu,\nu}$ 包含的图像块数目为 $k\times N$。

(3) 在图像块集合 $S_{\mu,\nu}$ 上利用 K 均值聚类算法学习得到 C 个聚类中心：$P_{\mu,\nu}=\{P_{\mu,\nu,0},P_{\mu,\nu,1},\cdots,P_{\mu,\nu,C-1}\}$。其中，$P_{\mu,\nu,i}$（$i=0,1,\cdots,C-1$）表示第 i 个聚类中心。$P_{\mu,\nu}$ 称为码本，聚类中心 $P_{\mu,\nu,i}$（$i=0,1,\cdots,C-1$）称为码字（对应某种模式）。

通常情况下，40 个 Gabor 小波（即 5 个尺度 8 个方向）用来提取人脸图像的 Gabor 特征。然而，这经常带来高维的特征表示以及较大的计算代价。为了减少特征维数，LLGP 方法采用了 3 个尺度 $\nu \in \{2,3,4\}$、4 个方向 $\mu \in \{0,2,4,6\}$ 的 Gabor 小波（即 12 个）。由于幅值特征反映了人脸图像上的显著性特征并对一些局部变化（如光照、表情）鲁棒，这里基于幅值特征来学习码本和表示人脸。

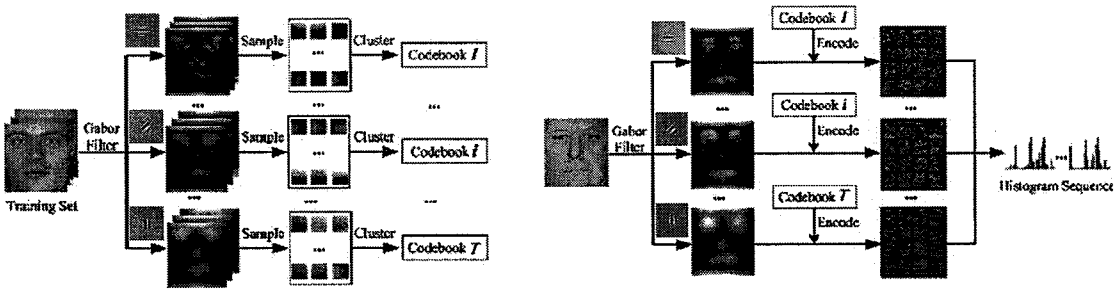


图 4-3 基于学习的局部 Gabor 模式表示的两个阶段 （a）学习阶段 （b）表示阶段

在学习码本时，如下几点需要注意：第一，尽管种子点的选择对 K 均值聚类结果的影响比较大，然而，本章的测试经验表明，它对最终的识别精度影响并不大。因此，这里选择了从人脸图像的不同位置采集的图像块作为种子点；第二，为了减弱光照变化对图像块的影响，所有采样的图像块都进行了 0 均值 1 方差的归一化处理；第三，图像块集合由从训练集中图像的 Gabor 特征上均匀采样得到的图像块构成。例如，对一幅 64×80 个像素的人脸图像，如果采样间隔设置为 8×8 个像素，那么该图像上一共得到 80 个图像块。对于包含 1,000 幅图像的训练集合，大约 80,000 个图像块可以得到。

2. 表示阶段

表示阶段根据学习阶段产生的码本来计算不同尺度、不同方向上 Gabor 小波滤波后图像的模式图像，并采用分块直方图来表示这些模式图像。与学习阶段类似，对于 T 个 Gabor 小波，一幅图像共得到 T 幅模式图像。

简要地讲，针对输入图像 I 的第 ν 个尺度、第 μ 个方向上 Gabor 小波表示 $G_{\mu,\nu}(I)$，其对应模式图像 $L_{\mu,\nu}(I)$ 的计算过程如下：

(1) 对 $G_{\mu,\nu}(I)$ 中的每个像素位置 z，以位置 z 为中心采样得到大小 m×n 的图像块 $p_{\mu,\nu}(z)$。

(2) 对图像块 $p_{\mu,\nu}(z)$，计算它到相应码本 $P_{\mu,\nu}$ 中所有码字 $P_{\mu,\nu,i}$（$i=0,1,\cdots,C-1$）的距离，并将位置 z 标记为与块 $p_{\mu,\nu}(z)$ 距离最小的码字所对应的索引，即：$L_{\mu,\nu}(z)=\left\{i \middle| \forall j: dist\left(p_{\mu,\nu}(z),P_{\mu,\nu,j}\right) \geq dist\left(p_{\mu,\nu}(z),P_{\mu,\nu,i}\right)\right\}$。其中，$dist(\bullet,\bullet)$ 表示两个向量

之间的欧氏距离。

在模式图像中，每个像素代表了与其所在块的特征距离最小的码字在相应码本中的索引，而索引之间的距离不能反映其对应码字的距离。例如，码字索引“1”和“60”之间的欧氏距离要大于码字索引“1”和“2”之间的距离，但是，前者对应码字的视觉相似度可能要大于后者。因此，传统的距离度量（如欧氏距离）并不适合于度量模式图像之间的距离。与 LVP 表示类似，LLGP 方法中采用了分块的直方图表示，即：模式图像 $L_{\mu,\nu}(\bullet)$ 被划分为若干个不重叠的子块并计算每个子块的直方图，然后将所有子块的直方图连接起来得到它的表示：

$$H_{\mu,\nu}(\bullet)=[H_{\mu,\nu,0},H_{\mu,\nu,1},...,H_{\mu,\nu,m-1}] \tag{4-1}$$

其中，m 表示模式图像 $L_{\mu,\nu}(\bullet)$ 划分的子块数目，$H_{\mu,\nu,i}(i=0,1,\cdots,m-1)$ 表示模式图像 $L_{\mu,\nu}(\bullet)$ 第 i 个子块的直方图，其定义如下式所示：

$$H_{\mu,\nu,i}(j)=\sum_{z}B\{L_{\mu,\nu,i}(z)=j\},j=0,1,\cdots,C-1 \tag{4-2}$$

这里，j 表示码字的索引，C 表示码本的大小，$B\{\bullet\}$ 是一个布尔函数，其定义如式(4-3)所示：

$$B\{A\}=\begin{cases}1, & if\ A\ is\ true\\ 0, & else\end{cases} \tag{4-3}$$

最后，一幅图像表示为由不同尺度、不同方向上各个子块的直方图连接得到的序列：

$$H(I)=[H_{\mu_0,\nu_0}(I),H_{\mu_0,\nu_1}(I),\cdots,H_{\mu_{o-1},\nu_{s-1}}(I)] \tag{4-4}$$

4.3 利用子块和模式权重的加权方法

根据提取的直方图特征，两幅图像 I_1 和 I_2 之间的相似度按照下式进行计算：

$$S(I_1,I_2)=\sum_{\mu=\mu_0}^{\mu_{o-1}}\sum_{\nu=\nu_0}^{\nu_{s-1}}\sum_{r=1}^{m}S(H^1_{\mu,\nu,r},H^2_{\mu,\nu,r}) \tag{4-5}$$

其中，$H^1_{\mu,\nu,r}$ 和 $H^2_{\mu,\nu,r}$ 分别表示图像 I_1 和 I_2 第 μ 个方向、第 ν 个尺度上第 r 个子块的直方图；$\mu=\mu_0,\cdots,\mu_{o-1}$ 和 $\nu=\nu_0,\cdots,\nu_{s-1}$ 分别表示采用的 Gabor 小波的尺度与方向；$S(\bullet,\bullet)$ 表示直方图之间的相似度，这里采用了直方图交的相似度度量。

大量的学术研究[1][107][151]表明，人脸不同区域（如眼睛，眉毛，鼻子等）在识别过程中具有不同的作用。因此，应当根据人脸的不同区域在识别过程中的作用大小而

赋予不同的权重。此外，最近的研究[46][97]表明，人脸皮肤上的斑点、胎记等特征对于人脸身份的判别也起到了重要的作用。在 LLGP 方法中，码本是在图像块集合上通过聚类学习得到，其包含的码字反映了所有图像块的一种“共性”的特征。尽管如此，码字的频数一定程度上反映了该码字在人脸图像上属于“常见”还是“少见”的特征。直觉上，频数低的码字对应了人脸图像上那些少见的特征，在识别过程中的作用要比频数高的码字大。为了同时利用不同区域和不同模式的表示能力，本节提出了一种利用二者权重的加权方法。

在该加权方法中，两幅图像 I_1 和 I_2 之间的相似度按照下式计算：

$$S(I_1,I_2)=\sum_{\mu=\mu_0}^{\mu_{O-1}}\sum_{\nu=\nu_0}^{\nu_{S-1}}\sum_{r=1}^{m}w_{\mu,\nu,r}\bullet S(H_{\mu,\nu,r}^{1,w},H_{\mu,\nu,r}^{2,w}) \tag{4-6}$$

这里，$w_{\mu,\nu,r}$ 表示模式图像 $L_{\mu,\nu}(\bullet)$ 第 r 个子块的权重，其学习过程参见第三章 3.3.1 节（即“利用 Fisher 准则的子块权重学习”）；$H_{\mu,\nu,r}^{1,w}$ 和 $H_{\mu,\nu,r}^{2,w}$ 分别表示图像 I_1 和 I_2 第 μ 个方向、第 ν 个尺度上第 r 个子块的加权直方图，定义如式（4-7）所示：

$$H_{\mu,\nu,r}^{w}(j)=\varpi_{\mu,\nu,j}\bullet\sum_{z}B\{L_{\mu,\nu,r}(z)=j\},j=0,1,\cdots,K-1 \tag{4-7}$$

其中，$\varpi_{\mu,\nu,j}$ 表示模式 $P_{\mu,\nu,j}$ 的权重，其值根据在学习阶段统计的码字频数计算得到：

$$\varpi_{\mu,\nu,j}=\frac{1}{\mathrm{Percent}(P_{\mu,\nu,j})}\Big/\sum_{j=0}^{C-1}\frac{1}{\mathrm{Percent}(P_{\mu,\nu,j})} \tag{4-8}$$

这里，$\mathrm{Percent}(P_{\mu,\nu,j})$ 表示在学习阶段第 j 个簇（即码字 $P_{\mu,\nu,j}$）的大小占采样图像块集合大小的比例，Σ 表示求和操作。显然，如果一个簇包含图像块的数目很多，该簇对应的模式可能反映了人脸区域上常见的特征，其赋予的权重就低；相反，如果一个簇包含图像块的数目很少，它对应的模式可能反映了人脸区域上那些少见的特征，其赋予的权重就高。通过将较高的权重赋值给那些少见的特征，这提高了 LLGP 直方图的表示能力。

同非加权方法相比，利用子块和模式权重的加权方法需要学习不同块和不同模式的权重。但是，这些权重的学习过程都是离线的；并且，模式权重根据学习得到的码本直接计算得到。此外，在计算图像相似度时，该方法具有与无加权方法基本相当的计算代价。总体而言，与识别精度的提高相比，该加权方法增加的计算代价是可以接受的。

4.4 局部 Gabor 模式方法的分类及对比

局部 Gabor 模式表示是一类基于图像 Gabor 特征局部模式的表示方法。到目前为止，文献中的局部 Gabor 模式表示方法多达 11 种，包括：幅值局部 Gabor 二值模式（即 LGBP_Mag）[150]、相位局部 Gabor 二值模式（即 LGBP_Pha）[151]、实虚部的局部 Gabor 相位模式（即 LGPP）[142]、实虚部的全局 Gabor 相位模式（即 GGPP）[142]、

基于学习的视觉码本（Learned Visual Codebook, LVC）[155]、局部 Gabor 纹理基元（Local Gabor Textons, LGT）[61]、有效的 Gabor 幅值“体”的局部二值模式（Effective Gabor Volume based Local Binary Patterns, E-GV-LBP）[62]、基于“体”的局部 Gabor 二值模式（即 V-LGBP）[127]、相位局部 Gabor 差分模式（Local Gabor Phase Difference Patterns, LGPDP）[35]、基于学习的局部 Gabor 模式（即 LLGP）[128]和相位的局部 Gabor 异或模式（即 LGXP）[129]。这些方法都在 Gabor 特征上根据某种模式定义算子（如 LBP）计算得到局部模式，进而来表示人脸图像。

关于这 11 种局部 Gabor 模式，如下几点需要注意：第一，在幅值“体”的局部二值模式中，E-GV-LBP[62]与 V-LGBP[127]方法的基本思路是一致的，但在相同阶段分别发表在不同的会议上。因此，下面采用 E-GV-LBP/V-LGBP 来表示这两种方法；第二，基于学习的视觉码本（即 LVC）[155]和局部 Gabor 纹理基元（即 LGT）[61]方法都利用聚类方法得到图像 Gabor 特征的局部模式，但二者分别应用于三维（LVC）和二维（LGT）人脸识别，因而文中采用 LVC/LGT 来表示这两种方法；第三，尽管文献[142]称全局 Gabor 相位模式（即 GGPP）和局部 Gabor 相位模式（即 LGPP）为“相位模式”，这二者本质上利用了 Gabor 特征的实部和虚部。此外，文献[142]提出的 Gabor 相位模式直方图（Histogram of Gabor Phase Patterns, HGPP）表示本质上是 GGPP 和 LGPP 两种表示在特征层的融合。为了公平地比较这些方法，这里忽略了 HGPP 表示而只考虑 GGPP 和 LGPP 这两种表示。

既然这 11 种方法都基于图像 Gabor 特征的局部模式，它们之间的区别有哪些？这些区别对它们有什么影响？针对这些问题，下面按照模式定义方式对这些方法进行分类，详细对比了同一类中不同方法的模式定义，并从模式定义流程和方法参数这两个方面进行了对比分析。

4.4.1 方法分类

根据模式是如何定义的，局部 Gabor 模式方法可划分为两大类：模式设计方法与模式学习方法。具体地讲，模式设计方法在图像 Gabor 特征上根据人工设计的算子（如 LBP）来计算得到其对应的模式，模式学习方法则通过学习的方法自动地得到这些模式。根据该分类准则，这 11 种局部 Gabor 模式方法中，8 种方法属于模式设计方法，即：LGBP_Mag、LGBP_Pha、GGPP、LGPP、V-LGBP、E-GV-LBP、LGPDP 和 LGXP；其余 3 种方法属于模式学习方法，即：LVC、LGT 和 LLGP。下面将分别概述这两类局部 Gabor 模式方法的模式定义，并对这些方法的异同进行分析。

为了清楚地区分不同局部 Gabor 模式之间的差异，这里引入了三种模式定义域：块（patch）、Jet 和“体”（volume）。具体地讲，“块”指的是图像平面内的空间邻域（如 3×3），Jet 是由同一点的不同 Gabor 小波响应构成的，“体”可看作是块和 Jet 的一种组合，它的前两维是图像平面（即块），第三维对应了不同尺度、不同方向的 Gabor 小波（即 Jet）。这三种定义域上局部模式的相关说明如下：块上的模式由空间邻域内的滤波

器响应计算得到，Jet 上的模式只与同一点的不同滤波器响应有关，“体”上的模式则由块和 Jet 上的滤波器响应共同决定。除模式定义域之外，不同局部 Gabor 模式之间的差异还在于：它们或者基于不同的 Gabor 特征，或者采用了不同的模式定义算子。

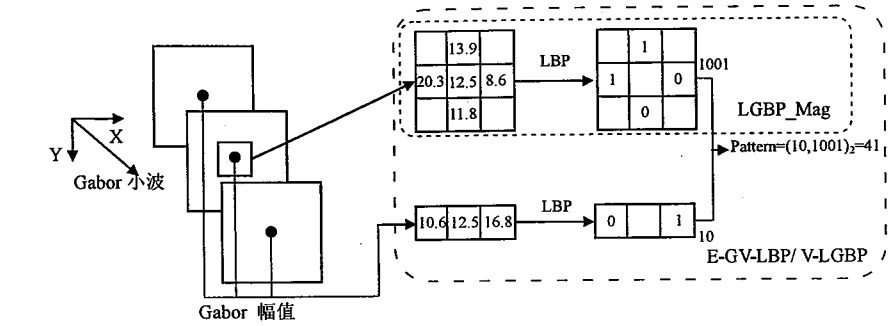
1. 模式设计方法

在模式设计方法中，每个像素的模式定义为一个二进制串的形式，它的每个二进制位根据该像素的特征与其所在的块、Jet 或者“体”上的邻域像素特征按照相应的模式定义算子计算得到。图 4-4（a）~（c）分别示出了基于 Gabor 幅值、实部/虚部、相位的模式设计方法中的模式定义示例。

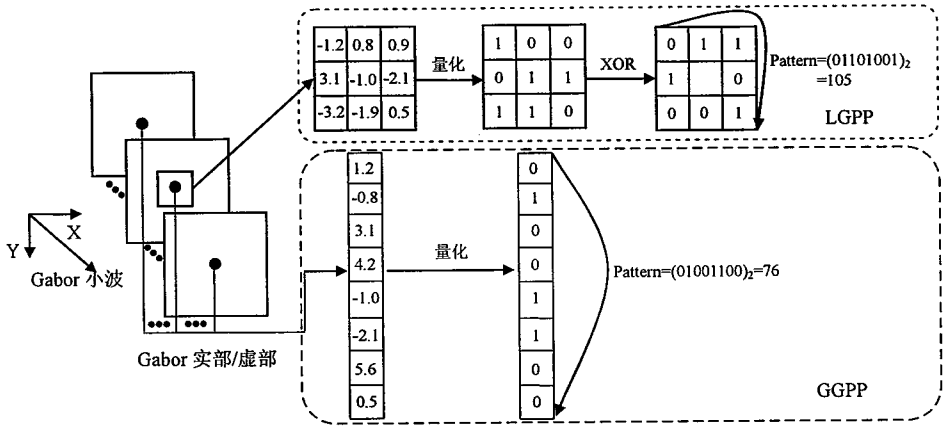
Gabor 幅值的模式设计方法包括幅值局部 Gabor 二值模式（即 LGBP_Mag）和“体”的局部二值模式（即 E-GV-LBP/V-LGBP）。如图 4-4（a）所示，在 LGBP_Mag 方法中，模式定义在块上并通过 LBP 算子计算得到；在 E-GV-LBP/V-LGBP 方法中，模式利用 LBP 算子在幅值“体”上计算得到。正如第三章 3.2 节所分析的，Jet 上滤波器响应之间的模式刻画了图像在尺度或方向上显著的局部结构，其与块内模式的组合反映了更多的微结构。因此，“体”上的模式表示（即 E-GV-LBP/V-LGBP）比块上的模式表示（即 LGBP_Mag）更有效。

Gabor 实部/虚部的模式设计方法包括全局 Gabor 相位模式（即 GGPP）和局部 Gabor 相位模式（即 LGPP）。如图 4-4（b）所示，在 GGPP 方法中，先将实部或虚部的响应按照符号量化为“0”（>0）或“1”（≤0），然后将给定尺度所有方向上实部或虚部构成 Jet 上的量化值（即“0”或“1”）组合起来得到局部模式；在 LGPP 方法中，先将实部或虚部特征根据符号量化为“0”（>0）或“1”（≤0），然后利用异或（XOR）算子在块内计算局部模式。因此，LGPP 表示涵盖了不同尺度、不同方向上图像的微结构，而 GGPP 表示只包含了不同尺度上的微结构。因此，LGPP 方法具有比 GGPP 方法更好的表示能力[142]。

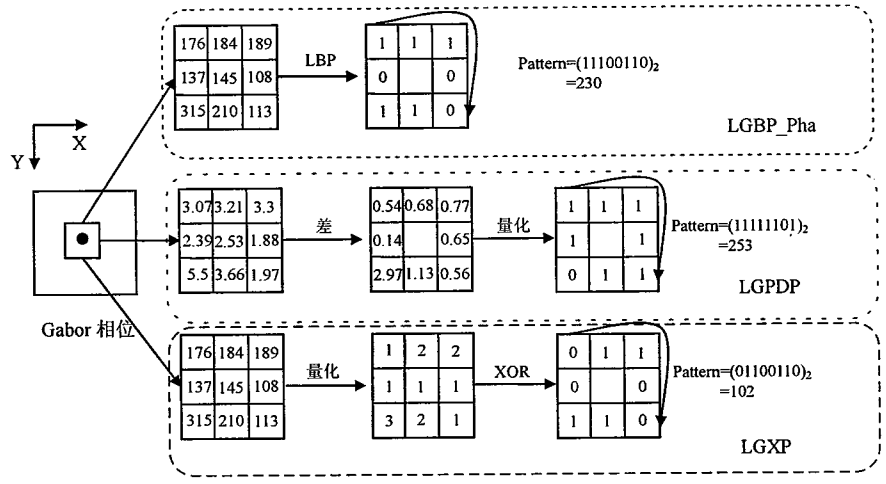
Gabor 相位的模式设计方法包括三种：相位局部 Gabor 二值模式（即 LGBP_Pha）、相位局部 Gabor 差分模式（即 LGPDP）和局部 Gabor 异或模式（即 LGXP）。如图 4-4（c）所示，这三种模式都定义在块上的：LGBP_Pha 方法直接利用 LBP 算子计算相位的局部模式；LGPDP 方法根据相邻像素间相位角差 Δ 的绝对值所处的范围得到“0”（$\pi/2 < |\Delta| \le 2\pi$）或“1”（$0 \le |\Delta| \le \pi/2$）；LGXP 方法先将相位角量化到不同的区间（如 4 个区间：[0°, 90°)→0, [90°, 180°) →1, [180°, 270°) →2, [270°, 360°) →3），然后利用异或（XOR）算子根据相邻像素的相位量化值计算得到“0”或“1”。在图 4-4（c）的示例中，LGBP_Pha 和 LGXP 方法采用了相位的角度表示，LGPDP 方法则采用了相位的弧度表示。由于采用了“软”（soft）模式计算方法，LGPDP 和 LGXP 方法中的模式受相位特征对位置变化敏感性的影响比较小。当这两种方法选择了合适的参数值时，二者将取得比较好的分类精度。



（a）



（b）



（c）

图 4-4 模式设计方法中的模式定义示例

（a）~（c）分别表示 Gabor 幅值、实部/虚部、相位的局部模式定义方法

2. 模式学习方法

不同于模式设计方法，模式学习方法需要一个学习阶段来构建码本（即模式的集合），然后根据学习得到的码本来计算图像上每个像素特征对应的模式。在 LVC[155]、