

# 基于最大Gabor响应的局部二元模式人脸特征提取方法

蔡新元, 张荣国, 肖柏华, 王春恒

中国科学院自动化研究所, 北京 100190

E-mail: {xinyuan.cai, rongguo.zhang, baihua.xiao, chunheng.wang}@ia.ac.cn

**摘要:** 本文中, 我们提出了一种新的基于最大Gabor响应二元模式的人脸特征提取方法。该方法借鉴了生物视觉中HMAX模型, 用Gabor滤波模拟简单细胞响应, 选取各个尺度上Gabor响应最大的方向。在最大Gabor响应图上用局部二元模式(LBP)描述人脸的局部信息, 然后用空间金字塔描述人脸的全局分布信息。在ORL和FERET数据集上的实验表明, 本文的方法与现有的最新技术相比, 在较低的特征维数上, 达到了与其相近的识别率。

**关键词:** 生物视觉, 最大Gabor响应, 二元模式, 空间金字塔

## Local Binary Pattern of Maximum Gabor Response: A Novel Face Representation

Xinyuan Cai, Rongguo Zhang, Baihua Xiao, Chunheng Wang

Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

E-mail: {xinyuan.cai, rongguo.zhang, baihua.xiao, chunheng.wang}@ia.ac.cn

**Abstract:** In this paper, we proposed a novel face representation method: Local Binary Pattern of Maximum Gabor Response. This method is inspired by the HMAX model in biological vision. We first use multi-scales & multi-orientations Gabor filters to simulate the response of simple cells in visual cortex. For a given scale, we get the maximum responses maps from all orientations. Then LBP operator is used to describe the local features, and spatial pyramids are constructed to describe the spatial distribution information of face images. The proposed method is applied to face recognition, and the experiment results on ORL and FERET databases show that, compared with the state of art methods, the proposed algorithm still could achieve the comparable recognition rate with lower dimensional features.

**Key Words:** biological vision, Maximum Gabor Response, Local Binary Pattern, Spatial Pyramid

### 1 引言

人脸是人体重要的生物特征之一, 并以其方便, 非侵入的特点, 在人机交互, 智能监控等领域受到了青睐。人脸识别的一个关键的挑战在于找到一种有效的, 并具有鉴别性的人脸外观的描述方法, 并且这种描述要能够对于光照, 表情, 年龄以及部分遮挡等变化具有一定的稳定性[1]。人脸识别虽然已经研究了很多年, 并且也有了许多商用的系统, 但是它仍然是一个活跃的, 具有挑战性的研究课题。现在仍有很多的研究者在一些公共的数据集上, 如 The Facial Recognition Technology (FERET)<sup>1</sup>, Labeled Faces in the Wild (LFW)<sup>2</sup>等, 测试验证人脸识别算法的鲁棒性。

现有的人脸识别方法可以分为两类: 基于统计和非统计的方法。(1) 基于统计的方法, 如 Eigenface[2], Independent Component Analysis(ICA)[3], LDA(Linear

Discriminate Analysis)[4]等。这些方法都需要有大量的训练本来学习出具有判别性的子空间, 所以判别子空间的学习非常依赖于训练集, 如果测试集的样本分布与训练集不同, 那么该方法的泛化性能就不好。而且基于统计的方法要求每个人要有多张训练图片, 这个要求太严格, 在实际系统中, 可能很难满足, 例如在公安部抓捕嫌疑犯的系统, 嫌疑犯的图片可能只有一张。在只有一张训练图片的情况, 基于统计的方法可能就失效了。(2)非基于统计的方法, 如Local Gabor Binary Pattern Histogram Sequence (LGBPHS)[5], Gabor Binary Codes(GBC)[6], Histogram of Gabor Phase Pattern(HGPP)[7]等, 这些方法都是在挖掘单张图片的信息, 不仅适用于单训练样本, 也适用于多训练样本的情况。

本文提出了一种基于非统计的方法—基于局部最大Gabor响应局部二元模式的人脸特征提取方法(Local Binary Pattern of Maximum Gabor Response)。该方法借鉴了视觉生理学中的HMAX模型[8], 即用多尺度多方向

<sup>1</sup> <http://face.nist.gov/colorferet/>

<sup>2</sup> <http://vis-www.cs.umass.edu/lfw/>

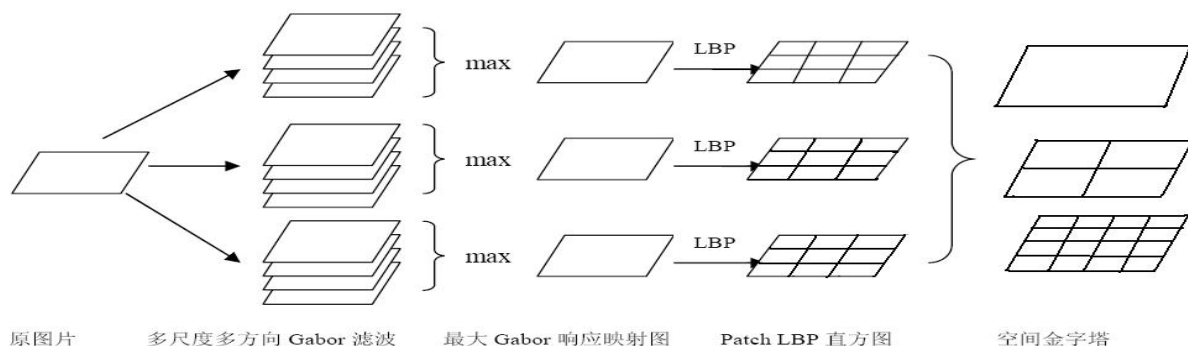


图 1 特征提取流程框图

的Gabor滤波器来模拟简单细胞的响应，然后用局部二元模式(Local Binary Pattern[9])对每个尺度上的最大的Gabor响应图进行编码。为了考虑人脸各个区域的空间分布信息，我们对最大Gabor响应图进行了空间金字塔划分。在公共数据集FERET和ORL上的实验结果表明，该特征表示对人脸表情，光照变化和部分遮挡具有一定的鲁棒性，在小样本的情况下取得了与最新技术水平相当的识别率，而且我们的方法的特征维数是现有最好方法的维数的1/9~1/4。

本文其余部分安排如下：第2节讲述本文方法的主要思想和各个步骤细节；第3节是实验设置以及实验结果与分析；最后是对本文的总结。

## 2 基于最大Gabor响应的局部二元模式人脸表示方法

Hubel&Wiesel在1959年提出细胞等级假设[10]，认为视觉皮层中较复杂的细胞感受野是由处于较低层次的，较简单的细胞感受野有序综合而成的。Riesenhuber & Poggio在细胞等级假设的基础上提出了原始的HMAX模型[8]。本文的方法正是借鉴了HMAX模型中简单细胞对视觉信息的处理过程，用多尺度多方向的Gabor滤波器来模拟简单细胞的响应。在每一尺度的Gabor滤波图像上的每一个位置，记录该位置滤波响应最强的那个方向的能量响应值，从而得到每个尺度上的最大的Gabor响应映射图(Max Gabor response map),这个映射图在每个位置上的响应只对最重要的那个方向敏感，而不是对所有方向都敏感，因此更能体现简单细胞对空间位置和方向的选择性。

与LGBPHS[5]相比，本文的方法去除了一些冗余的方向信息，能够更紧凑的表示人脸。在得到每个尺度的最大Gabor响应映射图后，我们把它划分为不重叠的小区域(Patch)，对每个小区域用LBP算子编码，得到其LBP直方向向量。为了考虑图像的空间分布信息，通过对最大Gabor响应图构建空间金字塔，对落入每一层的大区域(Block)中的小块区域的LBP直方向向量的每一维采用取平均、最大值和能量均值等三种策略进行聚合。然后将每个区域的向量连接起来，形成图像的空间金字塔特

征表示，作为图像的最终特征表示形式。特征抽取的流程框架图如图1所示。

### 2.1 最大Gabor响应映射图

Daugman在1980年发表的文章[11]，用Gabor小波来模拟视觉皮层中简单细胞的接受野，因为Gabor小波具有一些很好的性质，如空间定位、方向选择和空间频率选择等。在计算Gabor响应时，就是把原图像与一组不同尺度，不同方向的Gabor滤波器进行卷积。Gabor滤波器的形式如下：

$$g(x, y, \sigma, \theta, \varphi) = \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \cos(2\pi x' + \varphi) \quad (1)$$

其中  $x' = x \cos \theta - y \sin \theta$ ， $y' = x \sin \theta + y \cos \theta$ ， $\sigma$  表示空间尺度， $\theta$  表示方向， $\varphi$  表示初始相位。

在实验中，我们采用五个不同的尺度，和四个不同的方向的Gabor滤波器来表示简单细胞感受野的响应。

$$g_{u,v} = g(x, y, \sigma = u, \theta = v, \varphi) \quad (2)$$

那么每个尺度的最大Gabor响应映射图(MGRM)为：

$$MGRM_u = \max(\{g_{u,v} * I(x, y) | v = \{\theta\}\}) \quad (3)$$

其中\*表示卷积， $I(x, y)$ 表示一幅图像。这样对每一幅人脸图像，我们可以得到五幅最大Gabor响应映射图。



图2 从左至右，从上到下分别为原始图像，五个尺度上最大的Gabor响应映射图

## 2.2 局部二元模式 (Local Binary Pattern)

Ojala等人提出的LBP算子[9]最初用于纹理分析和分类,它具有很好的灰度和尺度不变性。为了增强最大Gabor响应映射图的信息,我们采用LBP算子对其进行编码。LBP算子就是对每个像素点,以该像素点的值为阈值,与其 $3 \times 3$ 邻域内的八个点比较,对得到的0和1序列值顺时针连接起来,并转换为十进制数,即0~255之间,这个数就是该像素点的LBP编码。LBP算子公式如下:

$$LBP(x_c, y_c) = \sum_{n=0}^7 2^n f(i_n - i_c) \quad (4)$$

其中 $i_c$ 表示 $(x_c, y_c)$ 处的像素值,  $i_n$ 表示 $(x_c, y_c)$ 周围邻域的像素值。 $f(i)$ 是阶跃函数,如果 $i \geq 0$ ,那么 $f(i) = 1$ , 否则 $f(i) = 0$ 。

LBP算子示意图如下:

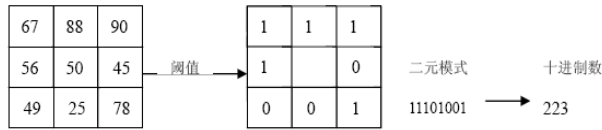


图3 LBP算子示意图

为了描述局部区域的特征,我们把最大Gabor响应映射图划分为 $L \times L$ 大小的小块(patch),在每个小块内计算均匀局部二元模式(Uniform LBP)直方图。这样,一幅 $H \times W$ 的图像的Gabor响应图就可以被表示为 $N = \text{ceil}(H/L) * \text{ceil}(W/L)$ 个Uniform LBP直方图特征。

## 2.3 空间金字塔

图像的不同尺度空间分布信息是图像分析过程中所需要的一种重要的信息,因此,我们将最大的Gabor响应映射图空间粗略的划分为 $2^l \times 2^l$  ( $l = 0, 1, 2$ )个大块(block)。对落入每一个大块(block)内的所有小块(patch)提取UniformLBP直方图特征,然后进行聚合。在实验中,采用三种聚合策略:

$$\begin{aligned} \max: z_j &= \max\{s_{j1}, s_{j2}, \dots, s_{jN}\} \\ \text{average}: z_j &= \frac{1}{N} \sum_{i=1}^N s_{ji} \\ \text{energy}: z_j &= \sqrt{\frac{1}{N} \sum_{i=1}^N s_{ji}^2} \end{aligned} \quad (5)$$

其中 $N$ 表示第 $j$ 个大块(block)中小块(patch)的个数,  $s_{ji}$ 表示第 $j$ 个block中第 $i$ 个patch的LBP直方图特征。每个block中的局部特征经过聚合后,形成一个向量,将所有block的特征向量归一化,连接起来形成图像的空间金字塔特征,作为图像的最终特征表示形式。

## 3 实验设置与结果分析

利用第二部分的特征提取方法,一幅人脸图像被表示为一个高维的特征向量。因为实验中选取的数据集中,每个人的参考样本很少,基本上都只是一幅图像,所以我们选取了三种相似度量度的方法,来计算测试图像与参考图像的相似度,然后利用最近邻分类。

$$\text{dis\_L1}(A, B) = \sum_{i=1}^n |a_i - b_i| \quad (6)$$

$$\text{dis\_Chi2}(A, B) = \sum_{i=1}^n (a_i - b_i)^2 / (a_i + b_i) \quad (7)$$

$$\text{dis\_Eul}(A, B) = \sum_{i=1}^n (a_i - b_i)^2 \quad (8)$$

公式(6), (7), (8)分别为两个向量的L1,Chi2及Eul相似度量度。

### 3.1 实验数据集

#### (1) ORL人脸数据集

ORL数据集是由AT&T实验室提供的<sup>3</sup>,该数据集包含了40张不同的人脸。每张人脸包含了10幅在不同的条件下拍摄的图片,例如:不同的时间、光照的变化、面部表情的变化以及眼镜的遮挡等,图片的分辨率为 $92 * 112$ 。

#### (2) FERET人脸数据集

FERET数据集被广泛的用于检验人脸识别算法。它共包含五个图片集,一个参考集Fa,四个测试集Fb, Fc, DupI, DupII。Fa包含1196幅正面图像, Fb包含1195幅表情变化的图片, Fc是在不同光照条件下拍摄的194幅图片, DupI是722幅在不同时间拍摄的图片, DupII包含234幅图片,是DupI的一个子集。该数据集中的原始人脸图片是未经切割的,在实验中,根据数据集提供的人工标定的眼睛,鼻子以及嘴唇的位置来校正并切割出人脸图像,并把图像归一化到 $128 * 128$ 大小。

ORL以及FERET数据集中部分人脸样本如图4所示。

### 3.2 实验参数设置

在实验中,我们用了五个尺度,四个方向的Gabor滤波器组,最大的空间频率设为 $\pi/2$ 。对ORL数据集,每张人脸图片中的一部分作为参考图片,剩下作为测试图片。对FERET数据集,就按照它标准划分的数据集进行实验。对于小块(patch)的大小,尺寸太大会失去了人脸的局部判别信息,尺寸太小最后形成的特征维数又会很高,为了平衡这两方面,对ORL图片, patch的大小设

<sup>3</sup> <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

为25\*25，对FERET图片，patch的大小设为21\*21。空间金字塔的层数设为3层。



图4: ORL和FERET数据集中人脸样本示例

### 3.3 实验结果分析

#### (1) 空间金字塔聚合策略与相似度度量选择

在实验中我们选择了公式(5)所示的三种聚合策略，以及公式(6)所示的三种相似度度量。在FERET数据集的Fa和Fb两个图片集上，我们验证了聚合策略和相似度度量的组合的人脸识别性能。识别正确率如表1所示：

表1 聚合策略与相似度度量组合的识别率

	Eul	Chi2	L1
average	90.67%	94%	95.5%
max	90.67%	94.3%	94%
energy	90%	93.6%	94%

从表1的实验结果可以看出，在相似度度量的选择上，Chi2和L1要比Eul度量高4~5个百分点。其原因在于本文中人脸特征表示是由各个小块的LBP直方图所连接而成的，在直方图特征上，Chi2和L1度量要比Eul度量更能体现样本在特征空间的分布。在空间聚合策略的选择上，三种策略都取得了相近的正确率。average聚合策略在L1度量时，比其他的组合稍高了1.5个百分点，因此在以下的实验中，如无特别说明，我们都是采用的average聚合策略和L1相似度度量。

#### (2) 对数据集大小的敏感性

传统的人脸识别算法如Eigenface[2]，LDA[4]等对训练集的大小都非常敏感。而在一些实际的应用系统中，每张人脸的图像是非常有限的。为了验证本文的方法对数据集大小的敏感性，我们在ORL数据集上做试验，从每张人脸的10幅图片中选 $n$ 幅( $n=1, 2, \dots, 9$ )作为训练图像，剩下的 $10-n$ 幅作为测试图像，与LBP[12]，LGBPHS[5]方法对比。实验结果如图5所示。

在ORL数据集中每个人的10幅图片包含了人脸的各种变化：表情、光照、遮挡等等。从实验结果可以看出，我们的方法和LGBPHS[5]方法在样本数为3和4时，都已经达到了95%~98%的识别率。因此在样本数较少的情况下，我们的方法和LGBPHS方法都充分的表示了人脸图像信息，从而为正确识别奠定了基础。对小样本的不敏感

性，也为该方法能够应用于实际的人脸鉴别系统提供了重要的前提条件。

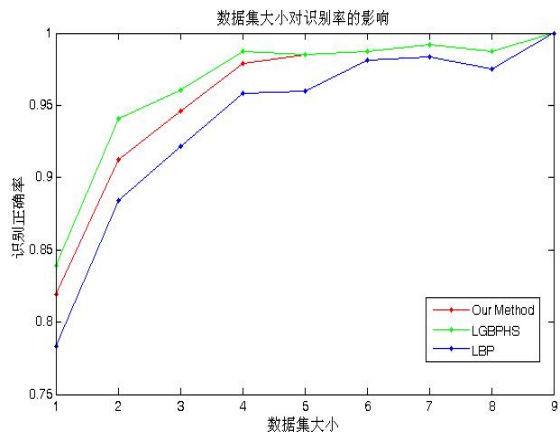


图5 数据集大小对识别率的影响

#### (3) 与现有方法比较

为了验证本文的人脸特征表示的有效性，我们在FERET数据集上，与HGPP[7]，LGBPHS[5]，LBP[12]，USC EBGM[13]的算法的第一识别率进行对比，识别结果表2所示：

表2 与现有方法在FERET数据集上第一识别率比较

	Fb	Fc	DupI	DupII
HGPP[7]	97.6%	98.9%	77.7%	76.1%
LGBPHS[5]	94.0%	97.0%	68.0%	53.0%
LBP[12]	93.0%	51.0%	61.0%	50.0%
USC EBGM[13]	95.0%	82.0%	59.1%	52.1%
Our Method	95.5%	97.0%	69%	52.0%

从上面的实验结果可以看出，我们的方法在FERET数据集上得到了与当前最新技术相当的识别率，对光照和表情的变化具有一定的鲁棒性。与其他的方法相比，我们的方法的特征维数却大大降低了，例如HGPP[7]采用5个尺度，8个方向的Gabor滤波后，得到5个GGPP maps和40个LGPP maps；LGBPHS[5]采用5个尺度，4个方向的Gabor滤波后，得到20个Gabor响应图(GMP)。我们的方法也是采用5个尺度，4个方向的Gabor滤波，得到了5个最大Gabor响应映射图(MGRM)。可见，我们的方法的特征维数是HGPP方法的1/9，是LGBPHS方法的1/4。

## 4 总结

本文在借鉴了生物视觉理论中HMAX模型[8]的基础上，提出了一种新的人脸特征表示的方法，目的是为了解决在人脸样本数目较少的情况下，如何充分挖掘图像信息，以提高人脸识别的准确率与效率。我们用多尺度多方向的Gabor滤波来模拟简单细胞的响应，在每个尺度上，取Gabor响应最大的那个方向的Gabor响应映



射图,然后用局部二元模式(LBP)来描述人脸的局部信息,用空间金字塔模型来描述人脸的全局信息,这样很好的刻画了人的脸部信息。从ORL数据集上的实验来看,本文的方法在训练样本较少(3~4)的时候,也能达到较高的识别率(95%~98%)。从FERET数据集上与现有的识别方法比较结果来看,我们的方法在特征维数是现有最好方法的1/9~1/4的基础上,仍能达到与其相当的识别率。

## 致谢

感谢芬兰拉普兰塔理工大学的机器视觉与模式识别研究组无私提供的快速有效计算Gabor特征的代码<sup>4</sup>。另外本文的工作还得到了基金项目No.60835001 和 No.60802055的资助,在此表示感谢。

## 参考文献

- [1] Xiaoyang Tan and Bill Triggs. Fusing Gabor and LBP Feature Sets for Kernel-Based Face Recognition. *Analysis and Modelling of Faces and Gestures* - Oct 2007
- [2] M. Turk, A. Pentland, Eigenfaces for Recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, 1991, pp. 71-86.
- [3] M.S. Bartlett, J.R. Movellan, T.J. Sejnowski, Face Recognition by Independent Component Analysis, *IEEE Trans. on Neural Networks*, Vol. 13, No. 6, November 2002, pp. 1450-1464.
- [4] K. Etemad, R. Chellappa, Discriminant Analysis for Recognition of Human Face Images, *Journal of the Optical Society of America A*, Vol. 14, No. 8, August 1997, pp. 1724-1733.
- [5] Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H., 2005. Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: *Proc. IEEE ICCV*, vol. 1, pp. 786-791
- [6] He, L., Hu, D., Jiang, C., 2006. Gabor binary codes for face recognition. In: *Proc. Mexican Int. Conf. Artificial Intell.*, pp. 53-60.
- [7] Zhang, B., Shan, S., Chen, X., Gao, W., 2007. Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition. *IEEE Trans. Image Process.* 16 (1), 57-68.
- [8] T Poggio M Riesenhuber. Hierarchical model of object recognition in cortex; *Nature Neuroscience* 2(11): 1019 -1025 .Nov .1999.
- [9] Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (2002) 971-987
- [10] D.H.Hubel and T.N.Wiesel. Receptive fields of single nerrones in the cat's striate cortex. *The Journal of Physuology*, 148(3):574-591, 1959.
- [11] Daugman, J.G. (1980), "Two-dimensional spectral analysis of cortical receptive field profiles", *Vision Res.* 20 (10): 847-56, PMID 7467139
- [12] Ahonen T, Hadid A, Pietikäinen M. Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell.* 2006 Dec;28(12):2037-41
- [13] S. Ravela and R. Manmatha, "Retrieving images by appearance," in *Proc. 6th IEEE Int. Conf. Computer Vision*, 1998, pp. 608-613.

---

<sup>4</sup> <http://www2.it.lut.fi/project/gabor/>