

中图法分类号: TP301.6 文献标识码: A 文章编号: 1006-8961(200) - -

论文引用格式:

面向烟雾识别与纹理分类的 Gabor 网络

袁非牛^{1,2}, 夏雪¹, 李钢^{1,3}, 章琳^{1,4}, 史劲亭⁵

1. 江西财经大学 信息管理学院, 南昌 330032; 2. 上海师范大学 信息与机电工程学院, 上海 201418;
3. 宜春学院 数计学院, 江西省宜春市, 336000; 4. 江西科技师范大学 数学与计算机科学学院, 江西 南昌 330038;
5. 江西农业大学 职业师范(技术)学院, 江西 南昌 330045.

摘要: **目的** 通过烟雾检测能够实现早期火灾预警, 但烟雾的形状、色彩等属性对环境的变化敏感, 这使烟雾特征容易缺乏辨别力与鲁棒性, 最终导致图像烟雾识别、检测的误报率与错误率较高。为解决以上问题, 本文提出一种基于 Gabor 滤波的层级结构, 可视为 Gabor 网络。**方法** 首先, 构建一个 Gabor 卷积单元, 包括基于 Gabor 的多尺度多方向局部响应提取和跨通道响应浓缩。其次, 将 Gabor 卷积单元输出的浓缩响应图进行跨通道编码并统计出直方图特征。以上 Gabor 卷积单元与编码层构成了一个 Gabor 基础层, 用于提取多尺度、多方向的基础特征, 对基础层引入最大响应索引编码和全局优化能生成扩展特征。最后, 将基础和扩展特征首尾相连形成完整烟雾特征。通过堆叠上述 Gabor 基础层能形成一个前馈网络结构, 将每一层特征首尾相连即可获得烟雾的多层级特征。**结果** 实验结果表明, 此 Gabor 网络泛化性能好, 所提烟雾特征的辨别力在对比实验中综合排名第一, 所提纹理特征的辨别力在两个纹理数据集上分别排名第一与第二。**结论** 所提 Gabor 网络能够实现多尺度、多方向的多层级纹理特征表达, 能够提高烟雾识别的综合效果, 也可提高纹理分类的准确率。未来可进一步研究如何降低特征的冗余度, 探索不同层特征之间的关系并加以利用, 以期在视频烟雾实时识别中得到实际应用。

关键词: 烟雾识别; 纹理分类; 特征提取; Gabor 滤波; 层级结构

A GaborNet for Smoke Recognition and Texture Classification

Yuan Feiniu^{1,2}, Xia Xue¹, Li Gang^{1,3}, Zhang Lin^{1,4}, Shi Jinting⁵

1. School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China
2. College of Information, Mechanical and Electrical Engineering, Shanghai Normal University, Shanghai 201418, China
3. College of Mathematics and Computational Science, Yichun University, Yichun 336000, China
4. School of Mathematics and Computer Science, Jiangxi Science and Technology Normal University, Nanchang 330038, China
5. Vocational School of Teachers and Technology, Jiangxi Agricultural University, Nanchang 330045, China

Abstract: Objective Since smoke often occurs earlier than flame when fire breaks out, smoke detection provides earlier fire alarms than flame detection does. The color, shape and movement of smoke are susceptible to external environment, so existing smoke features lack discriminative ability and robustness. These factors make image-based smoke recognition or detection a difficult task. In order to decrease false alarm rates (FARs) and error rates (ERRs) of smoke recognition without dropping detection rates (DRs), we propose a Gabor-based hierarchy (termed GaborNet) in this paper. **Method** First, a Gabor convolutional unit, which consists of a set of learning-free convolutional kernels and condensing modules,

收稿日期: 2018年6月-日; 回修日期: 2018年7月-日

基金项目: 本课题得到国家自然科学基金(61363038)、江西省高校科技落地计划(KJLD12066)、江西省青年科学家培养对象(20142BCB23014)、江西省教育厅科技项目(GJJ170317)资助。

Supported by: National Natural Science Foundation of China (61363038); Science Technology Application Projection of Jiangxi Colleges and Universities (KJLD12066); Cultivated Talent Program for Young Scientists of Jiangxi Province (20142BCB23014); Science Technology Project of Jiangxi Education Department (GJJ170317)

is constructed. The Gabor filters with fixed parameters generate a set of response maps from an original image as a multi-scale and multi-orientation representation. Besides, the condensing module conducts max pooling across channels of every response map to capture subtle scale- and orientation- invariant information and thus generate condensed response map. Then, condensed maps, i.e. the outputs of the above-mentioned Gabor convolution unit, are encoded both within and across channels. LBP (Local Binary Pattern) encoding method is leveraged to describe texture distribution within every channel of a condensed map, and Hash binary encoding is used to capture the relations across map channels. The binarization during encoding helps the representation be robust to local changes. Thereafter, histogram calculation is applied to encoded maps to obtain statistical features, known as basic features. The aforementioned Gabor convolution unit, encoding module and histogram calculation, forms a basic Gabor layer. In addition, this Gabor layer is provided with two extensive modules. One is to further explore the invariance and global structures of texture distributions, and the other is to enrich the pattern of response maps. The former restores and encodes the indices of max responses in the Gabor convolutional unit. The latter holistically learns a set of projection vectors from condensed response maps to construct a feature space. Once being projected into this feature space, the texture representation not only becomes more separable, but also carries more patterns. At last, completed smoke features of a Gabor layer are generated by concatenating basic features and extensive ones. Adding extensive features help improve the robustness and discriminative ability of basic features since invariant texture structures, holistic information and more patterns are characterized. Through stacking several Gabor layers on top of each other, a feedforward network, termed GaborNet, can be built. Consequently, the concatenation of features acquired from every Gabor layer constitutes multi-scale, multi-orientation and hierarchical features. As the network goes deeper, features become higher-level and less explicable. Thus, the extension, which explicitly improves basic features, is conducted only on the first Gabor layer that carries low-level features. Besides, once implemented, the holistic learning extension is not required any more in subsequent steps. **Result** This paper conducted ablation experiments to gain insights to the extensive features. Then, comparison experiments for smoke recognition were carried out to present the performance of the proposed GaborNet. Since this algorithm utilizes texture representations to present smoke, texture classification was also conducted as a supplement to the experiment. Experimental results demonstrate that the proposed GaborNet achieves powerful generalization ability. Smoke features extracted by the GaborNet decrease FARs and ERRs without dropping DRs, thus the results of GaborNet rank first among state-of-the-art methods. While results of texture classification rank at first and second place respectively in two standard texture datasets. In summary, the GaborNet provides better texture representation than most of the existing texture descriptors in both smoke recognition and texture classification. **Conclusion** The proposed GaborNet can extract multi-scale, multi-orientation and hierarchical representations for textures, and consequently helps improve the performance of smoke recognition and increases the accuracy of texture classification. Future researches should focus on eliminating the redundancy in features to gain compactness, and on exploring and utilizing the relations between features in different layers to enhance transform invariance. Eventually, this method is expected to be practically applied in real-time video smoke recognition.

Key words: smoke recognition; texture classification; feature extraction; Gabor filtering; hierarchy

0. 引言

烟雾检测能够提供比火焰检测更及时的火灾报警，因为大多数燃烧物在燃烧初期就能产生烟雾，在与空气充分接触之后才产生明火。烟雾传感器的工作范围有限，轻烟在无风时传播速度慢等因素限制了基于传感器的传统烟雾检测方法的应用场合，如：传统方法不适于户外环境或大场景室内的烟雾

检测。基于计算机视觉的烟雾检测打破了这一限制，它可通过检测摄像头采集到的场景视频或图像序列中的烟雾来进行烟雾预警。传统视频烟雾检测基本框架是：图像采集、疑似烟雾区域提取、特征表达、烟雾识别与定位^[1]，本文关注烟雾的识别过程。虽然烟雾具有流动性，但烟雾的静态特征更有效，原因如下：(1) 在大场景、远距离、无风或轻风条件下采集的视频中，相邻帧间的差异较小，需增大对视频帧的采样间隔才能获得足以分辨烟雾的动态特征，但难以自适应地选取采样间隔；(2) 在有风条件

下,除烟雾外也有其他运动物体,难以单独捕捉烟雾分量的光流。综上所述,动态特征更适合用于辅助、增强基于静态特征的烟雾识别。因此,本文主要关注单帧图像中的烟雾静态特征表达与烟雾识别。

常用的烟雾静态特征有形状、色彩与纹理,其中前两者对环境变化敏感,表现如下:(1)烟雾的流动与扩散受风力与风向影响,导致烟雾不具有稳定的形状;(2)燃烧物种类、空气含氧量、温度等诸多因素都会影响烟雾的颜色与浓度;(3)薄烟具有半透明性,周围场景的颜色会透出来影响烟雾本身的颜色。而纹理是静态属性中最稳定的,并不受到环境与天气影响^[2]。大面积烟雾使得全图具有局部自相似性,因此仅采用场景边缘、全局结构信息无法细致地描述烟雾的纹理。研究证明,局部纹理特征能够有效地表达烟雾^[3]。基于此,本文利用纹理分类的框架实现烟雾识别,框架分为特征提取与分类两阶段,在此侧重面向烟雾的局部纹理特征描述与提取。

近年来应用最广泛的局部纹理描述子是局部二值模式(LBP)^[4],具有计算效率高、旋转不变、光照不变等优势。随后出现了一系列基于LBP的改进方法:(1)基础扩展:引入幅度信息的完整局部二值模式(CLBP)^[5],扩展模式种类的局部三值模式(LTrP)^[6],实现尺度不变的金字塔局部二值模式(PLBP)^[7]等;(2)引入高阶信息的扩展:基于去噪导数幅度和中心像素值得高阶局部三值模式(HLTPMC)^[8],捕获LBP编码间关系的子方向梯度直方图(SOHLBP)^[9]和成对旋转不变共现局部二值模式(PRICoLBP)^[10]等;(3)利用变换域实现尺度不变和旋转不变的扩展:局部二值模式直方图离散傅里叶变换(LBP-HF)^[11],引入离散小波变换(DWT)的多分辨纹理算子^[12]等;(4)与学习结合的扩展:自适应直方图的完整辨别局部特征(CDLF)^[13],对模式出现频率进行学习与建模的主旋转局部二值模式(DRLBP)^[14]等;(5)端对端的深度学习框架:将卷积核二值化的局部二值卷积网络(LBCNN)^[15],引入激活旋转滤波器组(ARF)的旋转不变二值网络(RI-LBCNN)^[16]等。

以上方法有三个特点:(1)能简便地实现尺度不变或旋转不变的方法大多基于变换域,即,对不同频率或方向的子带提取局部特征后合并成高维特征。(2)能捕获层级信息的方法大多基于深度学习,训练样本数或网络的深度直接影响了特征的表达力

或分类的结果。(3)引入学习过程的特征提取方法大多以数据驱动的策略构建投影矩阵、字典或聚类中心,以此自动地将原数据表征为具有判别力的较低维特征。能够综合以上优点——多层次结构、可解释和可扩展性、引入学习过程——的方法,目前公开发表的只有主成分分析网络(PCANet)^[17]。该网络从局部去均值像素块中学习出多个投影矩阵,将原始数据投影后利用二值化哈希编码和块直方图统计来提取分辨力强的层级特征;但未引入尺度不变性,且逐块直方图统计导致特征维度与计算量极高。

基于以上分析,本文旨在利用少量学习提取多方向、多尺度的层级特征。主要贡献如下:

- (1)本文提出了一种基于Gabor滤波的前馈网络结构。通过多尺度、多方向的特征描述实现尺度与方向不变性。此网络结构可扩展为两层(或多层),从而提取多层次结构的特征。
- (2)受深度学习中池化模块的启发,本文利用空间最大值池化获得稳定的浓缩响应图;进一步提出最大响应索引编码,记录最稳定的响应所出现的尺度与方向,增加对边缘的描述能力同时尽量保持原纹理结构。
- (3)本文提出了基于LDA学习的3D卷积,以此从所有对尺度与方向鲁棒的数据中全局地学习出有利于分类的特征。

1. 相关工作

Gabor滤波器组可用于为提取多尺度与旋转不变的特征做准备^[18],滤波响应直接作为特征,维度太高,因此更适合作为特征图用于后续特征提取。Abdulrahman等人^[19]用PCA对Gabor滤波后的特征图组降维并创新性地从中提取LBP特征。Li等人^[20]将Gabor滤波后的特征图组逐像素对应地用二范数合成一张图,并从中提取LBP特征。

早在1988年,Daugman^[21]就构建了基于Gabor的三层神经网络,其中两层固定参数,另一层以最小均方误差为目标学习参数。Kameyama等人^[22]构建了一个第一层为Gabor滤波器组的神经网络。Oh等人^[23]对Gabor卷积后的图像块进行像素级的赋权,通过学习更新权值,最终连接极限学习机(ELM)分类,此框架能通过增加卷积层或堆叠特征提取模块而扩展为深度框架。

此后一段时间, Gabor 与神经网络的结合都停留在以上框架, 即, Gabor 核用于构成网络中(往往是最浅层)的免学卷积层。

Low 等人^[24]首先将原始 Gabor 滤波核组浓缩为多方向单尺度的滤波核组, 再将其与自身或其他卷积核进行自卷积, 以此扩展滤波核组。Luan 等人^[25]利用不同方向的 Gabor 滤波核组对卷积神经网络(CNN)的卷积核进行点乘, 以此生成 Gabor 方向滤波器(GoF: Gabor orientation filters)从而构建 Gabor 卷积网络(GCN: Gabor Convolutional Network)。遗憾的是, 以上两种新颖的 Gabor 滤波核扩展——自卷积和点乘——的有效性都是通过实验而非数学推导证明的, 缺乏可解释性。

基于以上分析, 本文利用 Gabor 滤波获得多尺度和多方向的响应图; 随后, 利用沿着尺度与方向的空间池化对响应图进行浓缩, 获得具有尺度与方向不变性的浓缩图; 最后, 进行跨通道与通道内的二值编码, 得到光照不变、旋转不变的特征。以上过程经过堆叠便可扩展为 Gabor 网络。此外, 还对最浅层特征提出两种纹理扩展表达: 用基于 3D 卷积的全局学习获取有利于分类的特征图, 对浓缩图中最大值的索引进行编码与统计来记录全局结构信息。此 Gabor 网络具有可解释性, 也能加入反馈形成端对端的深度结构。

2. 本文方法

2.1. 基于Gabor滤波的局部特征提取

2.1.1. 多尺度与多方向的局部特征表达

Gabor 滤波器的本质是加高斯窗的短时傅里叶变换, 其中高斯函数经傅里叶变换或逆变换后仍为高斯函数, 因此 Gabor 滤波可实现时频的局部化。空间域上以原点(0,0)为中心的 2D Gabor 滤波核及其参数表达如下:

$$G(x, y, \theta) = \frac{1}{\sqrt{2\pi\sigma_x\sigma_y}} e^{-\frac{1}{2}\left[\left(\frac{\hat{x}}{\sigma_x}\right)^2 + \left(\frac{\hat{y}}{\sigma_y}\right)^2\right]} \times e^{j(\omega_x x + \omega_y y)} \quad (1)$$

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (2) \text{式中 } \sigma_x, \sigma_y \text{ 是高斯函数的方差, 二者值越接近则高斯窗越圆。} \omega_x \text{ 与 } \omega_y \text{ 是空间频率, } \theta \text{ 代表 Gabor 函数的方向。在此令 } \omega_x = 2\pi\cos\theta/\lambda, \omega_y = 2\pi\sin\theta/\lambda, \sigma_x = \sigma_y = \sigma, \text{ 则可将方程组重写成为以下形式。}$$

$$G_{real}(x, y, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\hat{x}^2 + \hat{y}^2}{2\sigma^2}} \cos\left(\frac{2\pi\hat{x}}{\lambda} + \varphi\right) \quad (3)$$

$$G_{imag}(x, y, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\hat{x}^2 + \hat{y}^2}{2\sigma^2}} \sin\left(\frac{2\pi\hat{x}}{\lambda} + \varphi\right) \quad (4)$$

式中 $\theta = \{\sigma, \lambda, \theta, \varphi\}$, 三角函数之前的成分代表高斯窗, 三角函数代表被调制的复正弦、余弦信号。本文只选用实部作为滤波核。改变 Gabor 滤波核的尺度因子 v 与方向因子 u 时, $\theta_u = (u-1)\pi/0.25$, 高斯函数的方差取 $\sigma_v = 2^{(v-1)/2}$, 波长 $\lambda_v = \sigma_v/0.25$ ^[26]。本文在滤波过程中没有补零, 这一滤波过程可看作步长为 1 的卷积。多尺度和多方向的 Gabor 卷积核的响应图 I^{resp} (response map)可通过公式(5)获得, 该卷积核定义如公式(6)。

$$I_{v,u}^{resp} = I * G_{v,u} \quad (5)$$

$$\begin{cases} G_{v,u}(x, y, \theta_{v,u}) = \frac{1}{\sqrt{2\pi\sigma_v^2}} e^{-\frac{\hat{x}^2 + \hat{y}^2}{2\sigma_v^2}} \sin\left(\frac{2\pi\hat{x}}{\lambda_v}\right) \\ \theta_{v,u} = \{\sigma_v, \lambda_v, \theta_u \mid u = 1, 2, \dots, 8, v = 1, 2, 3, 4\} \end{cases} \quad (6)$$

对图像 I 作 Gabor 卷积得到 $8 \times 4 = 32$ 张响应图, 通过调整卷积核的方向和尺度能捕获纹理的细小变化, 但也带来了冗余。

2.1.2. 尺度不变与方向不变的响应浓缩

响应图携带了明确的尺度与方向信息, 本方法提出沿着尺度和方向的空间池化, 以此保留最稳定的尺度与方向响应, 获得尺度与方向不变性, 池化后的图称浓缩图。

$$\begin{cases} I_{o,u} = \max_{v=1,2,3,4} (I_{v,u}^{resp}) \\ I_o = [I_{o,1}, I_{o,2}, \dots, I_{o,8}] \in \mathbb{R}^{w \times h \times 8} \end{cases} \quad (7)$$

$$\begin{cases} I_{s,v} = \max_{u=1,2,\dots,8} (I_{v,u}^{resp}) \\ I_s = [I_{s,1}, I_{s,2}, I_{s,3}, I_{s,4}] \in \mathbb{R}^{w \times h \times 4} \end{cases} \quad (8)$$

如公式(7)所示, 对第 u 个方向, 将 4 个尺度中的最大响应浓缩成一张图 $I_{o,u}$ (o: orientation); 则 8 个方向($u=1,2,\dots,8$)能得到的 8 通道的浓缩图 I_o ; 由于, 浓缩掉了尺度信息而保留了方向信息, 故称 I_o 为方向浓缩图(后简称方向图)。同理, 公式(8)浓缩了方向信息, 保留了尺度信息, 构成 4 通道的尺度浓缩图 I_s (s: scale) (后简称为尺度图)。尺度图与方向图都属于浓缩图。

深度学习中最大值池化通过记录相对稳定的最

大值获得微小平移不变性，本方法利用跨通道的空间最大值池化，从一系列响应图中浓缩出最强、最稳定的纹理响应，形成具有微小尺度不变性的尺度图和微小方向不变性的方向图。

2.1.3. 跨通道编码的纹理表达

为探索不同通道中信息的关系，本文采用跨通道的哈希二值编码与直方图统计作为浓缩图的特征提取方式。

$$h_o(i) = \sum_{x,y} \delta \left(\sum_{u=1}^8 s(I_{o,u}(x,y)) 2^{u-1}, i \right) \quad (9)$$

$$h_s(i) = \sum_{x,y} \delta \left(\sum_{v=1}^4 s(I_{s,v}(x,y)) 2^{v-1}, i \right) \quad (10)$$

$$s(a) = \begin{cases} 1, & a \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

$$\delta(a,b) = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases} \quad (12)$$

式中函数 $s(x)$ 对浓缩图 I_o 与 I_s 以 0 为阈值进行量化， $\delta(a,b)$ 是判断函数：仅当 $a=b$ 时输出 1，否则输出 0，用于统计直方图。 $h(i)$ 表示直方图第 i 个 bin 的统计值。

上述哈希二值编码中引入了量化过程，通过忽略幅度值增加特征对光照变化的鲁棒性^[4]。直方图统计将有限个模式的出现频率作为这一多通道浓缩图的一阶特征。其中方向图编码将所有卷积窗中的纹理在 8 个方向下的分布规律记录成 $2^8=256$ 种模式，方向响应直方图 $h_o \in P^{256 \times 1}$ 。而尺度图编码记录的是纹理在 4 个尺度下的 $2^4=16$ 种模式，反映纹理在不同视角下的分布，尺度响应直方图 $h_s \in P^{16 \times 1}$ 。

2.1.4. 通道内编码的纹理表达

最能够表征纹理的属性是携带方向而非尺度信息的像素梯度，这使方向信息对纹理表征的贡献度

大于尺度信息。换言之，跨通道编码足以探索尺度信息，却没能充分利用尺度图中的方向信息，故在此还对尺度图进行了通道内编码，提取每一个子通道的 LBP(局部二值模式)特征，来呈现尺度不变的纹理描述。

$$\begin{cases} LBP_{s,v}(x,y) = \sum_{p=1}^8 s(I_{s,v}(x,y) - I_{s,v}(x^p, y^p)) 2^{p-1} \\ h_{sub,v}(i) = \sum_{x,y} \delta(LBP_{s,v}(x,y), i) \end{cases} \quad (13)$$

$$h_{sub} = [h_{sub,1}^T, h_{sub,2}^T, h_{sub,3}^T, h_{sub,4}^T]^T \quad (14)$$

公式(13)中 (x,y) 代表每一个 LBP 编码的中心点， (x^p, y^p) 是中心点 (x,y) 的第 p 个邻域像素。对特征图的每一个通道进行 LBP 直方图统计，将通道内直方图首尾连接成子直方图 h_{sub} 。

$$h = [h_s^T, h_{sub}^T, h_o^T]^T \quad (15)$$

最后，特征由尺度图的跨通道直方图 h_s 和子通道直方图 h_{sub} ，及方向图的跨通道直方图 h_o 组成。

2.2. Gabor基础单元与层级结构

本文将以上的多尺度、多方向 Gabor 卷积和沿着尺度与方向的浓缩，称作一个 Gabor 卷积单元，输出的是尺度图与方向图，如图 1 所示。黑边的白框代表输入输出的变量，灰色虚线框代表处理模块。Gabor 卷积、浓缩、以及编码过程都是处理模块。

原图经过一个 Gabor 卷积单元、跨通道哈希二值编码、直方图统计而输出特征的过程定义为一个 Gabor 基础单元，如图 2 中的白底灰边框所示。此基础单元输出基础特征 h 。堆叠此基础单元能构建一个层级结构，即一个无反馈网络(后简称网络)，若加上反馈便能构成一个端对端的深度结构。

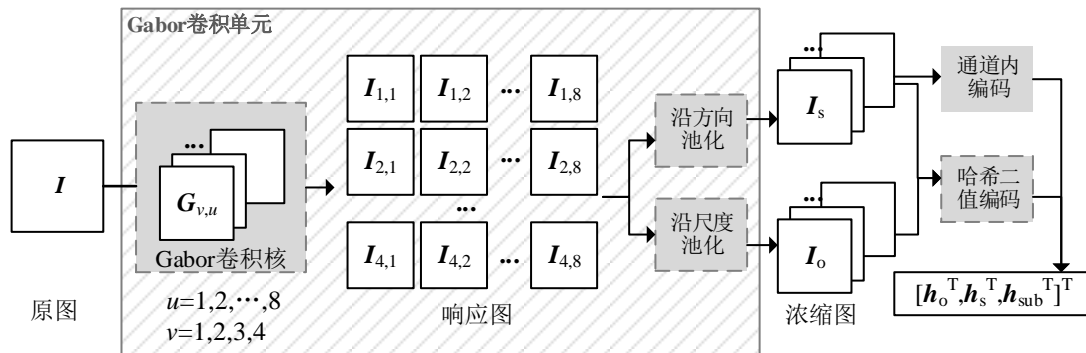


图1 Gabor 卷积单元中的操作与变量

Fig.1 The pipeline of a Gabor convolutional unit

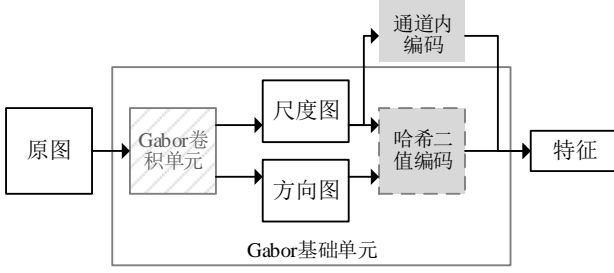


图2 Gabor 网络结构的第一层

Fig.2 The first layer of Gabor net.

随着层数的增加, 深层特征会越来越接近语义, 因此本文只在网络的第一层对最浅层特征(即第一层的尺度图)使用通道内编码。图 2 中, 灰色无边框的“通道内编码”模块不属于 Gabor 基础单元的操作, 因为它只出现在网络第一层。

2.3. 特征扩展: 全局学习与强纹理表达

2.3.1. 强纹理区域的特征表达

全局亮度发生变化时, 最大响应值(强纹理)相对其他像素值来说很可能依然最大, 即, 最大响应分布不易随光照改变。于是, 基于最大响应的特征 \mathbf{h}_s 与 \mathbf{h}_o 对光照鲁棒, 但直方图特征丢失了全局结构信息。为此, 本文通过对最大响应来自的尺度和方向进行编码, 以此加强边缘表达^[27], 并记录强纹理的全局分布信息^[28]。

$$\begin{cases} M_{\text{idx},v}(x, y) = \sum_{u=1}^8 \delta(I_{v,u}(x, y), I_{s,v}(x, y)) \times u \\ h_{\text{idx},v}(j) = \sum_{x,y} \delta(M_{\text{idx},v}(x, y), j) \end{cases} \quad (16)$$

$$\begin{cases} M_{\text{idx},u}(x, y) = \sum_{v=1}^4 \delta(I_{v,u}(x, y), I_{s,u}(x, y)) \times v \\ h_{\text{idx},u}(j) = \sum_{x,y} \delta(M_{\text{idx},u}(x, y), j) \end{cases} \quad (17)$$

式中 $\delta(\cdot)$ 是公式(12)中的判断函数, 公式(16)中, $I_{s,v}$ 代表每个尺度中方向响应的最大值, u 是方向索引号。仅当该最大值与索引号对应上时 $\delta(I_{v,u}, I_{s,v})$ 的输出才为 1, 否则为 0。因此 $\delta(I_{v,u}(x, y), I_{s,v}(x, y)) \times u$ 是在逐点地寻找最大响应所对应的方向索引。

$M_{\text{idx},v}$ 中记录的是第 v 个尺度通道中, 最大响应值所来自的方向, 故 $M_{\text{idx},v}$ 是方向响应索引图, 取值范围为 1~8。相应地, 直方图 $h_{\text{idx},v}$ 统计的是第 v 个尺度通道中最大响应来自于第 j (其中 $j=1,2,\dots,8$) 个方向的频率, 且 $h_{\text{idx},v} \in \mathbb{P}^{8 \times 1}$, 4 个尺度下的 $h_{\text{idx},v}$ ($v=1,2,3,4$) 组成方向索引特征。同理, 公式(17)记录了最大响应来自的尺度, 8 个 $h_{\text{idx},u} \in \mathbb{P}^{4 \times 1}$ ($u=1,2,\dots,8$) 能构成尺度索引特征。

将以上两种索引直方图首尾相连得最大索引特征 $\mathbf{h}_{\text{idx}} \in \mathbb{P}^{64 \times 1}$, 形成记录了强纹理结构信息的扩展特征。

2.3.2. 全局学习与纹理模式扩展

前文获得的尺度响应编码 \mathbf{h}_s 只有 16 种模式, 不足以描述纹理在不同尺度下的分布。因此, 本方法利用 3D 卷积从 4 通道的 I_s 中学习出 8 个投影向量, 将 I_s 扩展为 8 通道, 以此将模式增加为 $2^8=256$ 种。目标是将这 8 个投影向量张成一个有助于分类的特征空间, 同类数据在此空间中尽量接近, 不同类数据尽量分离, 故选用 LDA 作为学习策略。

为了将基于向量计算的 LDA 应用于本 3D 学习, 引入向量与矩阵的转换, 用步长为 1 的卷积实现向量映射。用 Ω_i 表示 I_s 中第 i 个像素处的 3D 采样区域, 若采样窗边长为 a , 则 I_s 的 4 个通道上以 (x^i, y^i) 为中心点的方形区域所组成的 Ω_i 为 $a \times a \times 4$ 。将这 $4a^2$ 个像素点张成一个向量 \mathbf{f}_i , 则投影后的向量 \mathbf{z} 可由公式(19)算得。

$$\begin{cases} \mathbf{f}_i = \text{vec}(I_s(x, y)) \in \mathbb{R}^{4a^2 \times 1} \\ \mathbf{w}_k = \text{vec}(\mathbf{W}_k) \in \mathbb{R}^{4a^2 \times 1} \end{cases} \quad (18)$$

$$\mathbf{z}_i = [\mathbf{w}_1^T \mathbf{f}_i, \mathbf{w}_2^T \mathbf{f}_i, \dots, \mathbf{w}_8^T \mathbf{f}_i]^T \in \mathbb{R}^{8 \times 1} \quad (19)$$

式中 vec (全称 vectorization) 将矩阵 \mathbf{W}_k 转化为向量 \mathbf{w}_k , 二者包含的元素个数与数值都相同, 只是形状不一样。 \mathbf{w}_k 表示第 k 个待求的投影向量, \mathbf{W}_k 是向量 \mathbf{w}_k 对应的卷积核形式。求解 \mathbf{w}_k 的过程是一个有约束的最大化问题, 如公式(20)所示。利用公式(21)构建拉格朗日方程将其转化为无约束问题。

$$\max_{\mathbf{w}} \frac{\mathbf{w}^T \mathbf{S}_b \mathbf{w}}{\mathbf{w}^T \mathbf{S}_w \mathbf{w}}, \quad \text{s.t. } \mathbf{w}^T \mathbf{S}_w \mathbf{w} = 1 \quad (20)$$

$$L(\mathbf{w}) = \mathbf{w}^T \mathbf{S}_b \mathbf{w} - \lambda (\mathbf{w}^T \mathbf{S}_w \mathbf{w} - 1) \quad (21)$$

类内散度矩阵 \mathbf{S}_w 与类间散度矩阵 \mathbf{S}_b 求解方法如公式(22)和(23)所示, 式中 c 代表第 c 类, C 是总类数; N_c 是第 c 类样本的数量, N 为总样本数, N_s 代表在每张图中卷积窗滑动的次数。

$$\mathbf{S}_w = \sum_{c=1}^C \sum_{n=1}^{N_c} \sum_{i=1}^{N_s} (\mathbf{f}_i^n - \frac{1}{N_c} \sum_{n=1}^{N_c} \mathbf{f}_i^n) (\mathbf{f}_i^n - \frac{1}{N_c} \sum_{n=1}^{N_c} \mathbf{f}_i^n)^T \quad (22)$$

$$\begin{cases} \mathbf{S}_b = \sum_{i=1}^C \sum_{c=1}^{N_s} N_c (\frac{1}{N_c} \sum_{n=1}^{N_c} \mathbf{f}_i^n - \mathbf{m}_i) (\frac{1}{N_c} \sum_{n=1}^{N_c} \mathbf{f}_i^n - \mathbf{m}_i)^T \\ \mathbf{m}_i = \frac{1}{N} \sum_{n=1}^N \mathbf{f}_i^n \end{cases} \quad (23)$$

令公式(21)的偏导为 0 可得: $\mathbf{S}_b \mathbf{w} = \mathbf{S}_w \mathbf{w}$ 。无论 \mathbf{S}_w 是否可逆, 都能够通过求解特征方程 $\mathbf{S}_b - \lambda \mathbf{S}_w = 0$ 而算出 \mathbf{w}_k ^[29], 本文取前 8 个特征值最大的特征向量

作为投影向量，即 $k = 1, 2, \dots, 8$ 。

与传统 LDA 不同：(1) 本方法中向量 f 来自一个 3D 窗口，探索了通道间的信息关系；(2) 类间散度 S_b 的计算是按滑动次数的累加而非 C 个类中心的叠加，因此矩阵 S_b 的秩远大于 $C-1$ ，这使得本方法能获得比传统 LDA 更多的有效投影向量。

如前文所述，特征投影的过程等同于用一组（8 个）卷积核卷积特征图，所求的 8 个 w_k 本质上就是 8 个 $a \times a \times 4$ 的卷积核 W_k ，最终 8 通道特征图 M 可由公式(24)计算得到。

$$\begin{cases} M_k = I_s * W_k \in \mathbb{R}^{w \times h}, k = 1, 2, \dots, 8 \\ M = [M_1, M_2, \dots, M_8] \in \mathbb{R}^{w \times h \times 8} \end{cases} \quad (24)$$

公式中*代表步长为 1 的卷积。8 通道图 M 的第 i 个位置 (x^i, y^i) 能取出 8 个点组成向量 z_i 。 $z_i = M(x^i, y^i) \in \mathbb{P}^{8 \times 1}$ ，来自于未卷积的 I_s 中的 3D 区域 Ω_i 。再用公式(25)对特征图 M 作跨通道的哈希二值编码与直方图提取，获得一个 256 维的直方图 h_{lda} 来替换 h_s 。此 h_{lda} 仅对应于最浅层网络中的 I_s 。

$$h_{lda}(i) = \sum_{x,y} \delta(\sum_{k=1}^8 s(M_k(x,y))2^{k-1}, i) \quad (25)$$

以上过程利用 8 个 3D 卷积核 W_k 从 4 通道的尺度图 I_s 中卷积出一个 8 通道的、引入了类别信息的特征图 M ，其中卷积核的参数是通过以扩大类间差异与缩小类内差异为目标全局地学习而来的。从 M 中算得的特征 h_{lda} 将纹理以更细致且更有助于分类的方式进行了表达，扩展了原 h_s 的表达力。

2.4. Gabor网络与层级特征表达

为获得更高层的特征，本文将 Gabor 基础单元堆叠形成 Gabor 网络。

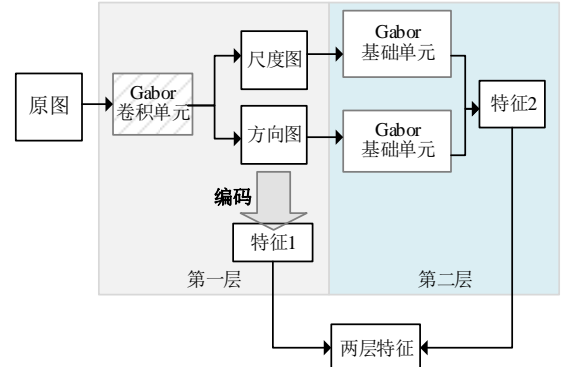


图3 两层 Gabor 网络流程图

Fig.3 The pipeline of a two-layer GaborNet

图 3 表示一个两层的 Gabor 网络，每一层包括一个或多个并行的 Gabor 基础单元（由 Gabor 卷积单元和编码模块组成）。网络第一的输入为原图，第一层的尺度图和方向图将作为下一层网络的输入。

图 4 展示了一个更深的层级结构，即，一个多层无反馈网络。其中条纹背景的方框代表一个 Gabor 卷积单元，前后交叠的白方框表示每个 Gabor 卷积单元的输出——尺度图与方向图。每层的尺度图和方向图都将分别输入到下一层网络的两个卷积单元，于是每层的卷积单元数都是上层的 2 倍，第 n 层将有 2^{n-1} 个卷积单元，对应 2^n 个浓缩图，最终有 2^n 个直方图。换言之，随着网络层数的增加，特征维度会成倍增长。

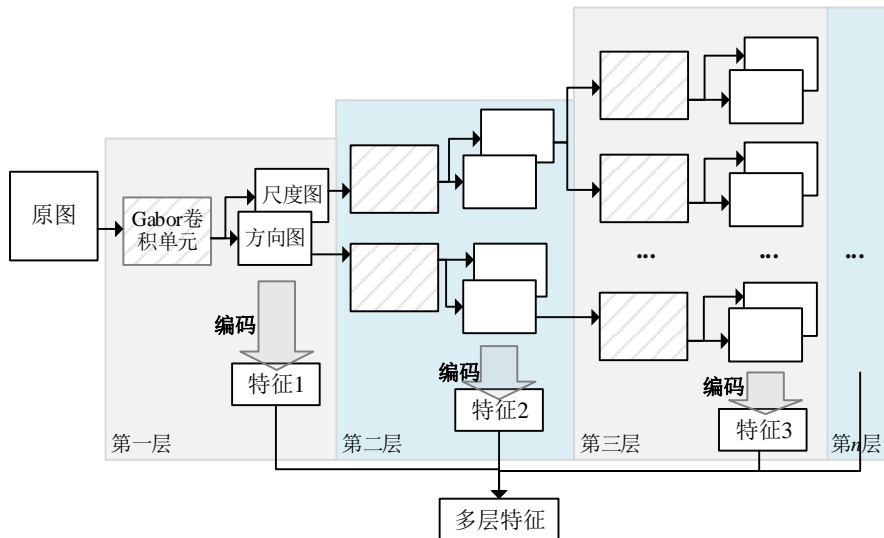


图4 多层 Gabor 网络流程图

Fig.4 The pipeline of a multi-layer GaborNet

如前文所述, 本方法采用的跨通道哈希二值编码过程与 LBP 的编码思路一致, 因此, 跨通道编码方法能够与 LBP 的三种映射模式(旋转不变模式 RI, 均匀模式 U2, 旋转不变的均匀模式 RIU2)结合起来控制特征维度。网络深层的方向图携带的信息更抽象, 其子通道对应的方向依然明确, 但方向的顺序不像浅层那样分明。于是采用 RI 映射模式能不影响表达力地对直方图特征降维, 编码方式如下:

$$\begin{cases} M_o(x, y) = \sum_{u=1}^8 s(I_{o,u}(x, y))2^{u-1} \\ M_{o_{ri}}(x, y) = \min_{u=1,2,\dots,8} \{ROR(M_o(x, y), u) | u=1,2,\dots,8\} \\ h_{o_{ri}}(i) = \sum_{x,y} \delta(M_{o_{ri}}(x, y), i) \end{cases} \quad (26)$$

式中 $ROR(a, u)$ 代表对 a 的二进制形式逐位旋转 u 次, \min 函数记录了旋转过程中出现的最小值。

随着层级结构加深, 特征越来越高层, 不适合用底层编码表达。于是, 所提 Gabor 网络只对最浅层的特征(即, 第一层浓缩图)作扩展编码。

3. 实验与分析

3.1. 烟雾识别对比实验

烟雾识别中用到的数据集共有四个, 烟雾图为正样本, 类间散度较小; 非烟雾图为负样本, 种类繁多导致类间差别很大。于是烟雾识别是正负样本不平衡的二分类问题。烟雾数据集介绍如表 1 所示, 最小的数据集 Set1 用于训练, 其他三个大数据集用于测试。所有数据集中正样本数都少于负样本数, Set3 和 4 中负样本比例在 78% 以上, 且负样本多种多样, 这为分类带来了难度。

用于评估烟雾识别效果的常用指标有三个, 分别是: 检测率(DR: detection rate)、误报率(FAR: false alarm rate)和错误率(ERR: error rate)。其中检测率(DR)就是真阳率(True positive rate), 误报率(FAR)是错判为正样本的负本数占真实负样本数的比率, 错误率(ERR)是所有的错分样本数占总样本数的比率。表达力强的特征能提高 DR, 降低 FAR 和 ERR。本实验用 LIBSVM 搭配适合直方图的卡方核作为分类器。烟雾识别实验中, 为避免分类面偏移, 正样本权重 $w_1=831/1383$, 负样本权重 $w_2=552/1383$; 其他参数同文献[10](代码已公开)。仅对 DMD^[30](dense micro-block difference)特征采用线性核, 因为

此特征维度很高, 足以表达数据的非线性分布, 无需再引入非线性核函数映射。

表1 烟雾数据集介绍

数据集	正样本数	负样本数	负样本占比(%)	用途
Set1	552	831	60.09	训练
Set2	688	817	54.29	测试
Set3	2201	8511	79.83	测试
Set4	2254	8363	78.77	测试

实验中采用两层 Gabor 网络提取特征。首先对本方法的几种版本进行剥离对比评估, 在实验中 Gabor 统一简称为 Gb。Gb 基础特征仅包含对浓缩图的跨通道及通道内编码, 写作“Gb 基础”。针对强纹理的最大响应索引特征称为最大编码, “+最大编码”表示 Gb 基础+最大编码特征。引入全局学习的特征称为学习特征, “+学习特征”表示 Gb 基础+学习特征。Gb 基础特征+响应编码+学习特征, 称为“Gb 完整”。为降低冗余同时保证方向编码的丰富信息, 除第一层网络外的所有跨通道编码均采用 RI 模式映射(对应的直方图为 36 维)。

表2 Gabor 网络所提特征的烟雾识别结果

测试库	特征	DR /%	FAR /%	ERR /%
Set2	Gb 基础特征	98.0	1.47	1.73
	+ 最大编码	98.1	1.22	1.53
	+ 全局学习	97.8	0.98	1.53
	Gb 完整特征	98.0	0.98	1.46
Set3	Gb 基础特征	97.3	2.03	2.18
	+ 最大编码	96.2	1.50	1.97
	+ 全局学习	97.2	1.84	2.04
	Gb 完整特征	96.5	1.48	1.90
Set4	Gb 基础特征	96.9	1.79	2.08
	+ 最大编码	95.9	1.46	2.02
	+ 全局学习	96.8	1.70	2.03
	Gb 完整特征	95.7	1.30	1.93

表 2 中可看出, 扩展特征的加入降低了 FAR 与 ERR, 但也影响了 DR。Set2 中, 扩展特征提升了基础特征的表达力, 但在 Set3 和 4 中, 并不能明显看出提升, 有三个原因: (1) 烟雾库中样本图尺寸很小, 基于 3D 卷积的学习中引入的扩展信息有限; (2) Set3 和 4 的样本数是训练集(Set1)的 10 倍多, 限制了学习过程的表达力和分类的效果; (3) 最大

编码保留的全局信息表达力不如局部信息（如引言所分析），故很难直观地看出效果的提升。(4) 此表格能看出扩展特征带来了稳定性，所以 FAR 与 ERR 才会明显降低，但不能全面地展现扩展特征带来的综合效果，因为不同的特征可能会在不同的分类阈值下获得最优分类面（在后文会详细解释）。

表3 烟雾识别对比结果

Table 3 Comparison results of smoke recognition

对比方法	Set2 /%			Set3 /%			Set4 /%		
	DR	FAR	ERR	DR	FAR	ERR	DR	FAR	ERR
Gb 基础	98.0	1.47	1.73	97.3	2.03	2.18	96.9	1.79	2.08
Gb 完整	98.0	0.98	1.46	96.5	1.48	1.90	95.7	1.30	1.93
HLTPMC ^[8]	97.1	1.22	1.99	94.0	1.63	2.52	95.4	1.22	1.93
SOHLBP ^[9]	97.0	1.35	2.13	95.2	1.82	2.45	94.7	1.46	2.28
PLBP-RI ^[7]	95.5	1.59	2.92	94.7	4.66	4.80	94.6	4.02	4.30
CLBP-RI ^[5]	96.2	3.06	3.39	95.0	4.78	4.84	94.6	3.66	4.02
LBPHF1 ^[11]	94.2	5.51	5.65	93.0	7.47	7.38	93.8	7.28	7.05
LBPHF2 ^[11]	95.6	2.20	3.19	95.2	3.40	3.68	95.8	2.98	3.22
NRLBP ^[31]	81.3	18.7	18.7	84.4	24.1	22.3	84.1	23.3	21.7
LTrP ^[32]	92.6	5.14	0.62	91.4	3.57	4.60	90.7	3.56	4.78
DFD ^[33] (1,256)	83.1	8.45	12.9	84.9	12.1	14.3	83.2	13.3	14.0
DFD (4,256)	93.2	6.37	6.00	91.4	6.49	7.44	94.5	3.92	4.27
DFD (7,1024)	92.0	4.28	5.98	92.1	11.9	11.1	91.4	11.3	12.0
DMD ^[30]	95.4	1.67	3.02	92.9	5.66	5.95	92.4	5.26	5.76
Gb_PCA_LBP ^[19]	95.8	5.63	4.98	96.0	2.40	2.74	96.1	2.42	2.74

表 3 中粗体表示最高 DR 和最低 FAR、ERR。最高 DR 和最低 ERR 都出现在本方法的 Gb 特征中；只有在 Set4 上，本 Gb 的 FAR 位居第二，比 HLTPMC 高 0.08%，但 Gb 完整特征的 DR 在三个

库上都明显高于 HLTPMC。LBPHF 对 LBP 直方图进行了傅里叶变换，LBPHF1 代表一个尺度的变换，2 代表两个尺度下的变换。DFD(Discriminant face descriptor)(1,256)代表不分块，聚类中心 256 个；DFD(7,1024)代表原图分为 7×7 个块，每一块 1024 个聚类中心（原文默认设置）；DFD 对人脸有很好的分类效果，但纹理图常具有局部自相似性，分块反而会带来冗余。DMD 从随机密集采样的图像块中学习出纹理的表达模型，在此选用文献默认的特征提取参数；由于烟雾图尺寸很小，不能提供足够采样块，因此 DMD 对烟雾识别的效果不如其对纹理分类的效果。Gb 与 PCA 结合的方法适用于表达人脸但不适合纹理，加上用于纹理的 LBP 模块后效果显著提升，因此本文选用适于纹理表达的 Gb_PCA_LBP 进行对比。

为了更综合地展示本文特征在烟雾识别上的辨别力，利用不同分类阈值下的 DR 和 FAR 组合绘制 ROC 曲线，选用表 3 中效果较好的方法，加上 LBP(经典)和 DMD(最新纹理描述子)作为对比。曲线越接近左上角，特征越有辨别力，图 5 显示：在三个数据集上 Gb 完整特征都比 Gb 基础特征更有辨别力。而 Gb_PCA_LBP 依赖于每一个样本自身，没有进行全局学习，因此稳定性不够，在三个数据集上效果相差较大。

本文方法在 Set2 和 3 中体现出比较明显的优势，这与表 3 呈现的结果略有不同，因为表 3 中的分类阈值为 0，而 ROC 曲线是在分类阈值从 -1 等差变化到 1 的过程中绘制的，说明最优分类面有时不在 0 点处。对于 Set4，本文方法略低于文献[8]，因为该方法对一阶和二阶信息进行了三值编码，模式更丰富。

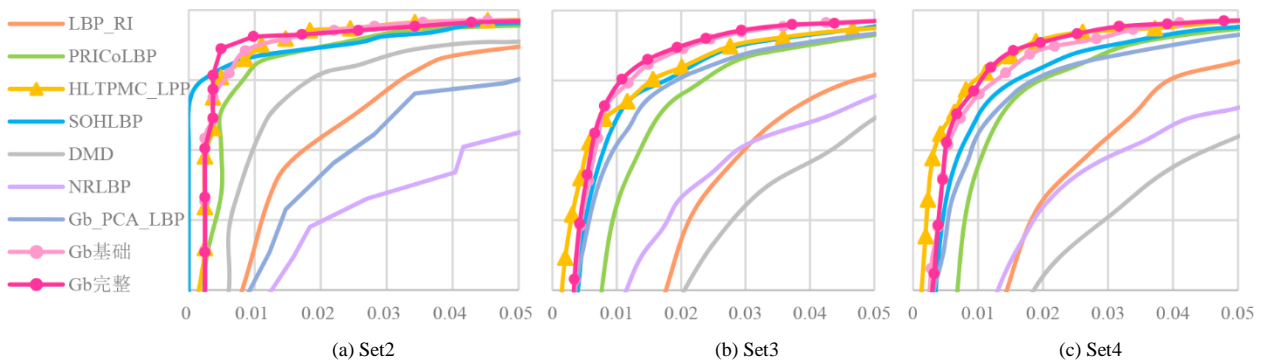


图5 烟雾识别 ROC 曲线对比

Fig.5 ROC curves of smoke recognition. (a) ROC curves on Set2; (b) ROC curves on Set3; (c) ROC curves on Set4.

3.2. 纹理检测对比实验

静态特征的烟雾识别在国际上没有统一的数据库和评估标准，同时本方通过纹理特征表达烟雾，故在此用标准纹理库 Brodatz 与 KTH-TIPS 进一步验证本纹理描述框架的有效性，采用的分类准则同文献[30]（代码已公开）。依然只对 DMD 采用线性核（原因同前），对其他直方图特征用卡方核，搭配 LIBSVM 进行分类，参数同文献[10]。

Brodatz 库共有 2048 张图，每类 64 张，样本包含原图，旋转图，缩放图以及旋转加缩放的图；随机选取 32 张训练分类器，32 张测试。KTH-TIPS 共 810 张图，每类 81 张，随机选 40 张训练分类器，41 张测试。由于存在一定的随机性，在此取 10 次交叉验证的平均准确率作为最终评估指标。

纹理图大小为 200×200 ，将近是烟雾图的 4 倍，故在此用 Gb 完整特征描述这些大纹理图。其中 3D 卷积核无需重新学习，直接沿用从烟雾 Set1 中学得的卷积核。为了公平对比，DMD 与 Gb_PCA_LBP 的学习过程也是在 Set1 上进行的，以此测试学习过程的通用性以及所学特征表达的泛化性能。

表4 纹理分类的对比结果

Table 4 Comparison results of texture classification

对比方法	Brodatz /%	KTH-TIPS /%
Gb 完整 RI	98.96±0.42	98.68±0.65
Gb 完整无映射	99.10±0.20	98.70±0.72
LBP ^[11]	97.46±1.28	96.12±1.83
CLBP ^[5]	96.27±0.79	98.22±1.10
PLBP ^[7]	92.85±1.25	98.02±1.04
PRICoLBP ^[10]	98.46±0.57	98.95±0.82
Gb_PCA_LBP ^[19]	85.70±1.07	82.51±1.69
DMD(Set1) ^[30]	95.41±0.52	95.93±1.86

表 4 记录了 10 次交叉验证的平均准确率±标准差。此处“Gb 完整 RI”与烟雾中“Gb 完整特征”一致，维度是 $1920+256+32 \times 2=2240$ ；“Gb 完整无映射”没有用 RI 映射，维度为 $3504+256+32 \times 2=3824$ 。

基于 LBP 的对比方法都没经过 RI 映射，其中 LBPHF 和 CLBP 采用论文中的完整版特征。本 Gb 特征平均准确率排名第一和第二，但标准差始终最小，说明本特征较稳定。原 DMD 特征在每一轮测试都要重新学习，现将烟雾 Set1 用于 DMD 的一次性学习，则准确率不如原文和本方法。而本 Gabor 网络能从两类小尺寸样本中学习出适用于大尺寸样

本的多分类投影模型，说明泛化性能更强。PRICoLBP 也引入了多尺度信息，且统计了成对编码的共现信息，因此具有很好的辨别力，在 KTH-TIPS 库上的准确率超越了本文方法，但在 Brodatz 库上的准确率不如本方法。而 Gb_PCA_LBP 没有采用全局学习，依赖于每个样本自身，因此泛化性能不足，用 Set1 学得模型不足以用于描述标准纹理样本。

4. 结论

本文提出了一种基于 Gabor 卷积的无反馈网络结构——Gabor 网络。首先，对图像进行多尺度和多方向的局部表达；随后，用跨通道的空间池化从局部特征中获得微小方向与尺度不变的浓缩图；接着，对此浓缩图进行跨通道与通道内的量化编码，量化过程带来了光照不变性，通道内 LBP 编码引入了旋转不变性。编码后的直方图作为基础特征，对微小方向与尺度变化、光照变化鲁棒，且具有旋转不变性。通过堆叠以上步骤可形成层级结构，随着层级加深，特征会更加接近语义，最终实现了用于多层次特征描述的 Gabor 网络。

为进一步探索浓缩图中的方向信息，本文针对 Gabor 网络的最浅层（第一层）浓缩图提出两种特征表达扩展。第一，对浓缩图中最大响应的索引进行编码，最大值的稳定性使得此编码能够记录强纹理的稳定成分，从而增加了特征的边缘描绘力度并保留了强纹理的全局结构。第二，对尺度浓缩图进行 3D 卷积，卷积的参数通过全局学习获得，学习目标为减少类内差异同时增大类间差异。此学习过程只需一次，从小尺寸的两类样本中学出的卷积核可有效应用于其他数据库；卷积后的特征能提供更丰富的纹理模式并更有利于分类。以上扩展特征通过对强纹理分布的描述增加了稳定性，通过有监督学习增加了可分性。

将具有变换不变性的基础特征与具有稳定性和可分性的扩展特征结合，能生成更具鲁棒性的完整特征。

本 Gabor 网络中需要学习的参数少，不涉及反馈，可与不同分类器搭配使用，比标准卷积网络简便、灵活，但仍有以下缺点：(1) 一些卷积与编码过程无法并行，导致计算复杂度或时间复杂度比传统方法高；(2) 特征的维度比一些传统方法高；(3) 本

方法所提特征仅对微小变换具有不变性。今后也许能通过加深网络和增加反馈来提高特征表达力与分类效果，也可采用 3D 卷积进一步控制特征维度、增加特征的变换不变性。

本方法的 Gabor 特征及烟雾识别的 ROC 曲线绘制代码可从 <http://staff.ustc.edu.cn/~yfn/index.html> 下载。

参考文献 (Reference)

- [1] Shi J, Yuan F, Xia X. Video smoke detection: a literaturesurvey[J]. *Journal of Image and Graphics*, 2018, 23(3): 0303-0322. [史劲亭, 袁非牛, 夏雪. 视频烟雾检测研究进展 [J]. *中国图象图形学报*, 2018, 23(3): 0303-0322.]
- [2] Tian H, Li W, Ogunbona P O, et al. Detection and separation of smoke from single image frames [J]. *IEEE Transactions on Image Processing*, 2018, 27(3): 1164-1177.
- [3] Yuan F. Video-based smoke detection with histogram sequence of LBP and LBPV pyramids [J]. *Fire Safety Journal*, 2011, 46(3): 132-139.
- [4] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971-987.
- [5] Guo Z, Zhang L, Zhang D. A completed modeling of local binary pattern operator for texture classification [J]. *IEEE Transactions on Image Processing*, 2010, 19(6): 1657-1663.
- [6] Tan X, Triggs B. Enhanced local texture feature sets for face recognition under difficult lighting conditions [J]. *IEEE Transactions on Image Processing*, 2010, 19(6): 1635-1650.
- [7] Qian X, Hua X, Chen P, et al. PLBP: an effective local binary patterns texture descriptor with pyramid representation [J]. *Pattern Recognition*, 2011, 44: 2502-2515.
- [8] Yuan F, Shi J, Xia X, et al. High-order local ternary patterns with locality preserving projection for smoke detection and image classification [J]. *Information Sciences*, 2016, 372: 225-240.
- [9] Yuan F, Shi J, Xia X, et al. Sub oriented histograms of local binary patterns for smoke detection and texture classification [J]. *Ksii Transactions on Internet and Information Systems*, 2016, 10(4): 1807-1823.
- [10] Qi X, Xiao R, Li C-G, et al. Pairwise rotation invariant co-occurrence local binary pattern [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, 36(11): 2199-2213.
- [11] Zhao G, Ahonen T, Matas J, et al. Rotation-invariant image and video description with local binary pattern features [J]. *IEEE Transactions on Image Processing*, 2012, 21(4): 1465-1477.
- [12] Ahmadvand A, Daliri M R. Invariant texture classification using a spatial filter bank in multi-resolution analysis [J]. *Image and Vision Computing*, 2016, 45: 1-10.
- [13] Zhang Z, Liu S, Mei X, et al. Learning completed discriminative local features for texture classification [J]. *Pattern Recognition*, 2017, 67: 263-275.
- [14] Mehta R, Egiazarian K. Dominant rotated local binary patterns (DRLBP) for texture classification [J]. *Pattern Recognition Letters*, 2016, 71: 16-22.
- [15] Juefeixu F, Boddeti V N, Savvides M. Local binary convolutional neural networks [C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA: IEEE, 2016: 4284-4293.
- [16] Zhang X, Xie Y, Chen J, et al. Rotation Invariant Local Binary Convolution Neural Networks [J]. *IEEE Access*, 2018, 6: 18420-18430.
- [17] Chan T, Jia K, Gao S, et al. PCANet: a simple deep learning baseline for image classification [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 5017-5032.
- [18] Yin Q, Kim J N. Rotation invariant texture classification using circular Gabor filter banks [C]//*Proceedings of the 7th international Conference on Computational Science*, Beijing, China: Springer Berlin Heidelberg, 2007: 149-152.
- [19] Abdulrahman M, Gwadabe T R, Abdu F J, et al. Gabor wavelet transform based facial expression recognition using PCA and LBP [C]//*Proceedings of the Signal Processing and Communications Applications Conference*, Trabzon, Turkey: IEEE, 2014: 2265-2268.
- [20] Li Z, Huang Z, Zhang T. Gabor-scale binary pattern for face recognition [J]. *International Journal of Wavelets, Multiresolution and Information Processing*, 2016, 14(05):

- 1-22.
- [21] Daugman J. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1988, 36(7): 1169-1179.
- [22] Kameyama K, Mori K, Kosugi Y. A neural network incorporating adaptive Gabor filters for image texture classification [C]//Proceedings of the International Conference on Neural Networks, Houston, TX, USA: IEEE, 1997: 1523-1528.
- [23] Oh B-S, Oh K, Teoh A B J, et al. A Gabor-based network for heterogeneous face recognition [J]. Neurocomputing, 2017, 261: 253-265.
- [24] Low C Y, Teoh B J, Ng C J. Multi-fold Gabor, PCA and ICA filter convolution descriptor for face recognition [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2017. (Online)
- [25] Luan S, Chen C, Zhang B, et al. Gabor convolutional networks [J]. IEEE Transactions on Image Processing, 2018, 27(9): 4357-4366.
- [26] Haghighat M, Zonouz S, Abdel-Mottaleb M. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification [J]. Expert Systems with Applications, 2015, 42(21): 7905-7916.
- [27] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [28] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [C]//Proceedings of the European Conference on Computer Vision, Zürich, Switzerland: Springer, 2014: 818-833.
- [29] Sun J. Morden pattern recognition[M]. 2nd ed. Beijing: Higher Education Press, 2008: 81-85. [孙即祥. 现代模式识别 [M]. 第二版 ed. 北京: 高等教育出版社, 2008: 81-85.]
- [30] Mehta R, Egiazarian K. Texture classification using dense micro-block difference [J]. IEEE Transactions on Image Processing, 2016, 25(4): 1604-1616.
- [31] Ren J, Jiang X, Yuan J. Noise-resistant Local Binary Pattern with an embedded error-correction mechanism [J]. IEEE Transactions on Image Processing, 2013, 22(10): 4049-4060.
- [32] Murala S, Maheshwari R P, Balasubramanian R. Local

Tetra Patterns: a new feature descriptor for content-based image retrieval [J]. IEEE Transactions on Image Processing, 2012, 21(5): 2874-2886.

- [33] Lei Z, Pietikainen M, Li S Z. Learning discriminant face descriptor [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(2): 289-302.

第一作者简介:



袁非牛(1976-), 男, 教授, 博士(中国科学技术大学), 主要研究方向为图像处理、模式识别、视频分析与事件检测、语义分割、三维建模与可视化, 在 SCI 检索期刊上等发表论文 60 多篇。

E-mail: yfn@ustc.edu

通信作者:



夏雪, 博士研究生, 主要研究方向为图像处理、模式识别。

E-mail: yeziandkuma@qq.com



史劲亭, 女, 博士, 讲师, 主要研究方向为图像处理、模式识别。

E-mail: icanflysjt@126.com

其他作者简介:

李钢, 男, 讲师, 主要研究方向为模式识别。E-mail: 562672612@qq.com

章琳, 女, 副教授, 主要研究方向为模式识别, 语义分割。E-mail: zymm_nc@163.com