

研究生课程考试成绩单

(试卷封面)

院 系	计算机科学与工程	专业	计算机科学与技术			
学生姓名	卓旭	学号	212138			
课程名称	图象分析与理解					
授课时间	2022 年 3 月至 2022 年 5 月	周学时	3	学分	3	
简 要 评 语						
考核论题	《Gabor 滤波器及其应用举例》					
总评成绩 (含平时成绩)						
备注						

任课教师签名：_____

日期：

- 注：1. 以论文或大作业为考核方式的课程必须填此表，综合考试可不填。“简要评语”栏缺填无效。
2. 任课教师填写后与试卷一起送院系研究生秘书处。
3. 学位课总评成绩以百分制计分。

Gabor 滤波器及其应用举例

《图象分析与理解》课程报告，212138 卓旭

一、Gabor 滤波器简介

本节对应代码在 `gabor_1.ipynb`

Gabor 滤波器^[1]以 Dennis Gabor 命名。它的形态使得它可以对点及其周围区域检测是否在特定方向上存在特定的频率组分特征。Gabor 滤波器在数字图像处理领域的纹理分析、特征识别、特征构造、稀疏表示等方面有许多应用。

在二维空域中，Gabor 滤波器可视为被一高斯函数调制的复正弦波

$$g(x, y) = g(x, y)s(x, y)$$

其中

$$s(x, y) = \exp(j \left(2\pi \frac{x'}{\lambda} + \phi \right))$$

是复正弦波， j 是虚数符号。在进行数字图像处理时，我们通常只使用 $s(x, y)$ 的实部^[2]，即余弦

$$s(x, y) = \text{Re}[\exp(j \left(2\pi \frac{x'}{\lambda} + \phi \right))] = \cos \left(2\pi \frac{x'}{\lambda} + \phi \right)$$

另外

$$g(x, y) = \exp(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2})$$

是零均值的高斯函数包络。上面两式中， x', y' 是 x, y 进行旋转后得到的，

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

θ 是逆时针为正的旋转角。参数 ϕ 称为 Gabor 滤波器的相移，参数 λ 可调节余弦波的频率，参数 γ 和 σ 可联合调节高斯函数的尺度。

综上，Gabor 滤波器 $g(x, y)$ 由五个参数共同确定， $g(x, y) = g(x, y; \theta, \phi, \lambda, \gamma, \sigma)$ 。在一维上直观地想象，Gabor 滤波器的幅度类似下图：

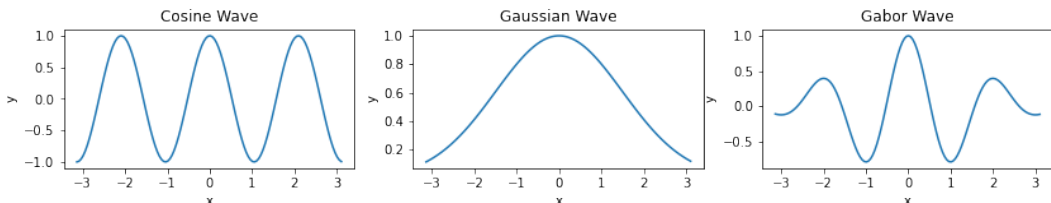


图 1 – 一维下余弦波、高斯函数以及调制后得到的 Gabor 波形

推广向二维，Gabor 滤波器的形态是中间明亮、向外呈波动衰减的条带状。

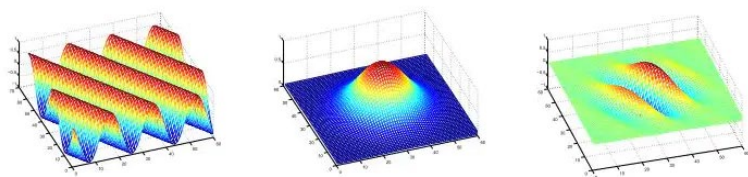


图 2 – 二维下余弦波、高斯函数以及调制后得到的 Gabor 波形

[<https://blog.csdn.net/lbelievesunshine/article/details/105113458>]

为了更直观地理解各参数下产生的 Gabor 滤波器的不同效果，下面采用控制变量的方式进行探究。请注意下方生成的滤波器的离散化尺寸是根据 x 、 y 方向的等效 σ 值按 $\pm 3\sigma$ 标准决定的。

旋转角 θ

固定 $\sigma = 3, \lambda = 7, \phi = 0, \gamma = 1$ ，调节 θ ，结果如下，可见 Gabor 滤波器发生了旋转。

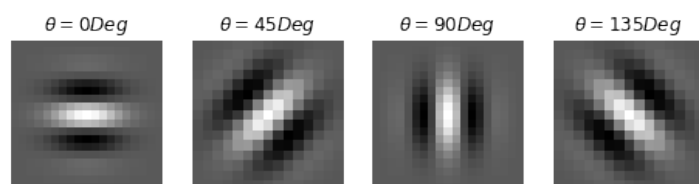


图 3 – 不同旋转角下的 Gabor 滤波器

相移 ϕ

固定 $\sigma = 3, \lambda = 7, \theta = \pi/2, \gamma = 1$ ，调节 ϕ ，结果如下，可见 Gabor 滤波器的余弦波组分发生了相位变化，特别是当 $\phi = \pi/2$ 时，偶函数的余弦波变为奇函数的正弦波。

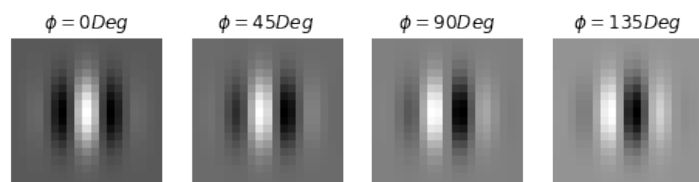


图 4 – 不同相移下的 Gabor 滤波器

余弦波频率相关参数 λ

注意 λ 在分母上，故与余弦波频率呈反比。固定 $\sigma = 3, \theta = \pi/2, \gamma = 1, \phi = 0$ ，调节 λ ，结果如下，可见 Gabor 滤波器的余弦波组分发生了频率变化。

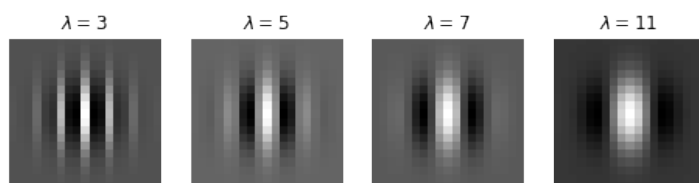


图 5 – 不同余弦波频率下的 Gabor 滤波器

y 方向展宽相关参数 γ

固定 $\sigma = 3, \lambda = 5, \theta = \pi/2, \phi = 0$ ，调节 γ ，结果如下，可见随着 γ 增大，在 y

方向上高斯函数的展宽变窄，Gabor 滤波器的能量分布范围也变窄。

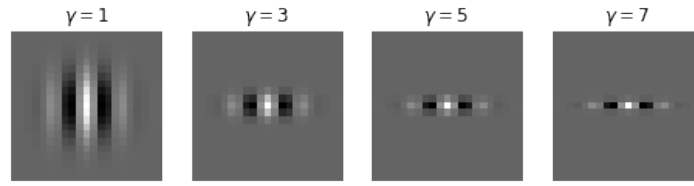


图 6 – 不同 y 方向展宽参数下的 Gabor 滤波器（离散化到同一尺寸）

高斯函数展宽相关参数 σ

固定 $\lambda = 5, \theta = \pi/2, \gamma = 1, \phi = 0$ ，调节 σ ，结果如下。可见随着 σ 的增大，高斯函数的展宽变大，有效囊括的余弦波的周期也就更多了。

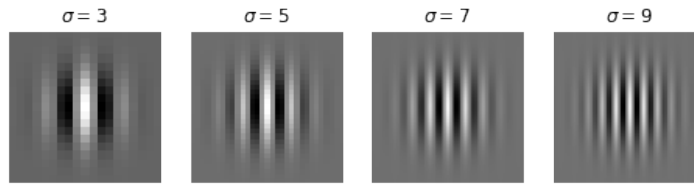


图 7 – 不同高斯函数展宽参数下的 Gabor 滤波器（离散化尺寸不一）

Gabor 滤波器丰富的参数化使得它的形态能在许多方面富于变化，从而能够检测各种不同密集度、不同方向、不同尺度、不同宽度、不同亮暗相位的交替条带状纹理。特别地，将多个不同参数的 Gabor 滤波器组成一组 Filter Bank，能实现更多、更精细的纹理相关操作。具体应用将在后文各节介绍。

我们还可以借助傅里叶变换从频域角度理解 Gabor 滤波器的特性，这里简单讨论。对于 $s(x, y)$ ，其傅里叶变换为余弦函数频率处的双边冲激 $\frac{1}{2}[\delta(\Omega - \Omega_0) + \delta(\Omega + \Omega_0)]$ ；对于 $g(x, y)$ ，其傅里叶变换仍为高斯函数。在空域相乘等价于在频域卷积，因此 Gabor 滤波器的频谱是双边冲激与高斯函数卷积的结果，即两个高斯函数被搬移到 $\pm\Omega_0$ 位置。

取一离散化的 Gabor 滤波器作离散傅里叶变换，得到对数幅度频谱如下图，符合预期。这同时表明 Gabor 滤波器一般情况下是一个带通滤波器：

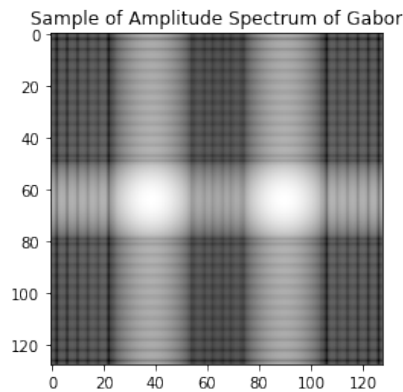


图 8 – Gabor 滤波器的幅度频谱示例

二、Gabor 滤波器与纹理分析——以牛仔布瑕疵检测为例

本节对应代码在 `gabor_2.ipynb`

牛仔布料上存在着大量重复出现的纹理特征。在牛仔布的加工过程中，有时会出现缺陷、瑕疵。为了检测出这些瑕疵，可以利用 Gabor 滤波器的响应特点，分析这些条纹状的纹理。

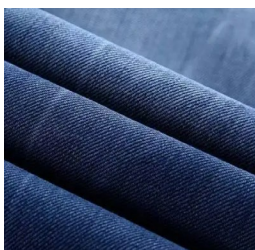


图 9 – 牛仔布料 [https://www.sohu.com/a/160691169_368281]

从 AITEX FABRIC IMAGE 数据集^{[3][4]}中获取一张有瑕疵和一张无瑕疵的牛仔布料的拍摄结果（0081_006_04.png，同一台设备拍摄同种布料），如下所示：

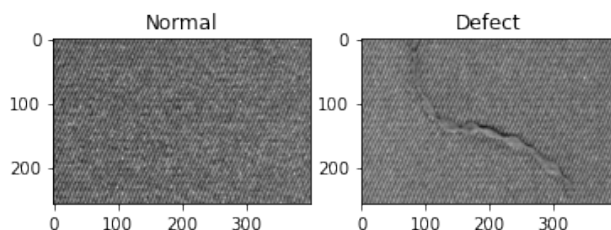


图 10 – 正常牛仔布料与有缺陷牛仔布料

使用 $\sigma = 0.5, \theta \in \{45^\circ, 135^\circ\}, \lambda = 2, \phi = 0, \gamma = 0.5$ 的两个 Gabor 滤波器对两张布料图进行滤波（由于牛仔布的纹理比较密集，因此选用较小 σ 和 λ ），得到结果如下，然后对滤波后图进行方差计算和大津阈值分割：

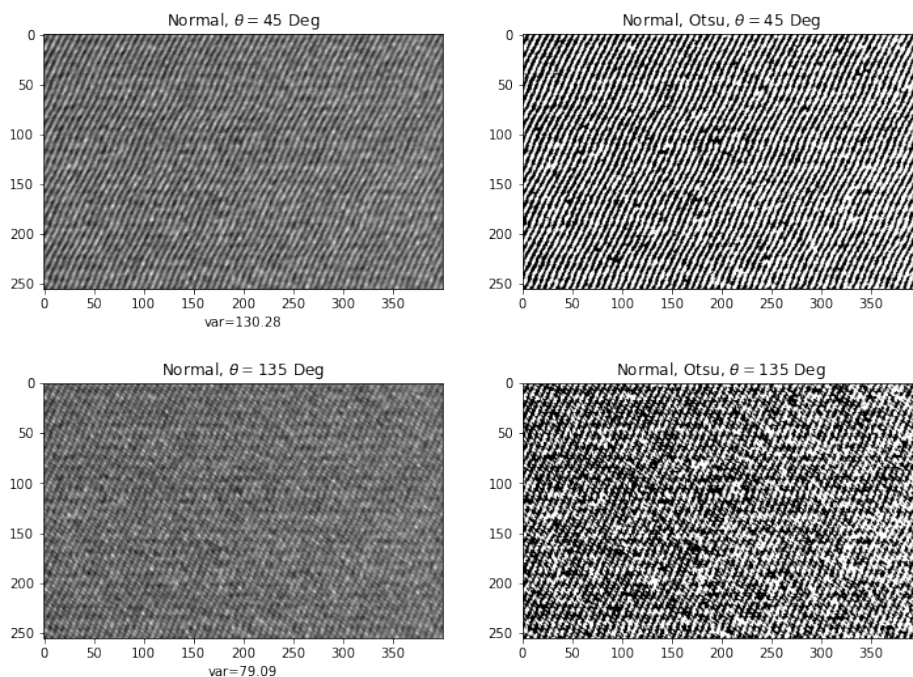


图 11 – 正常牛仔布的滤波结果

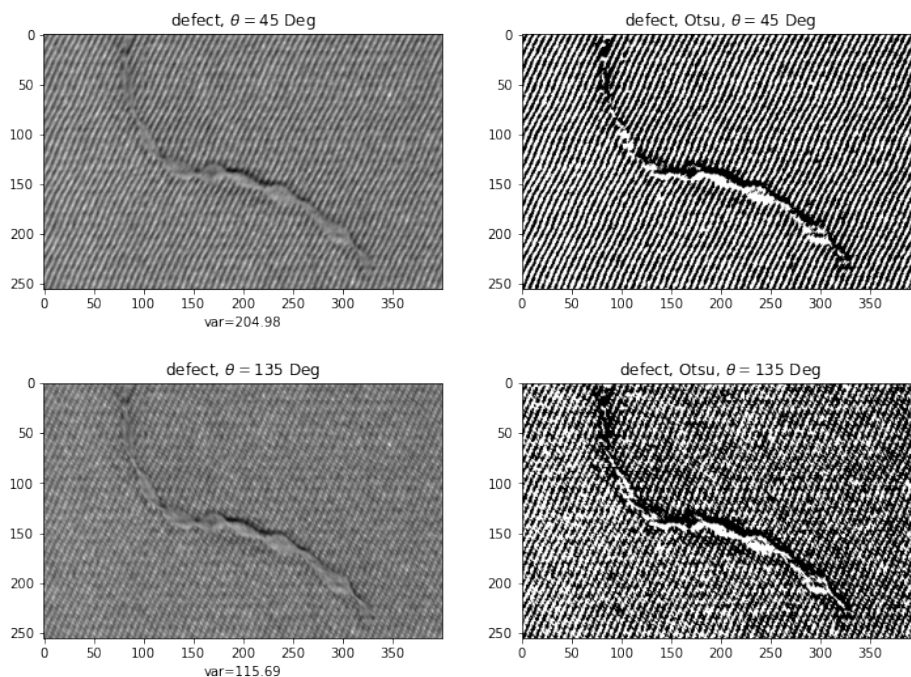


图 12 – 有缺陷牛仔布的滤波结果

可见，在应用 Gabor 滤波器后，牛仔布某一方向的纹理被提取出来，而缺陷处的响应明显不同。这样的特性便于我们进行后续分析判断。例如，使用一些简单的数值标准，如对比提取后得到的图像和参考图像的方差，超出一定量时就可认为布料存在缺陷。进一步地，在得到的图像上应用大津阈值分割，进一步增强缺陷部位，可为其他判别算法作预处理准备。更复杂、具体的做法有许多其他文献研究^[5]。

三、Gabor 滤波器与纹理提取——以图像汉字检测为例

本节对应代码在 `gabor_3.ipynb`

汉字相比于英文字母，大部分在笔画结构上更加密集且复杂，这使得使用 Gabor 滤波器进行纹理特征提取成为可能。本节，我们尝试使用 Gabor 滤波器进行图像中的汉字提取，并说明 Gabor 滤波器对汉字进行特征构建的原理。



图 13 – 一张综艺节目的截图，转为灰度图像

[https://www.youtube.com/watch?v=6ZWLKsp_DmU, 《向往的生活 6》第 1 期]

以上图为例，构造 $\sigma = 3, \theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}, \lambda = 5, \phi = 0, \gamma = 1$ 的 Gabor 滤波器组，得到四张滤波后结果图如下：

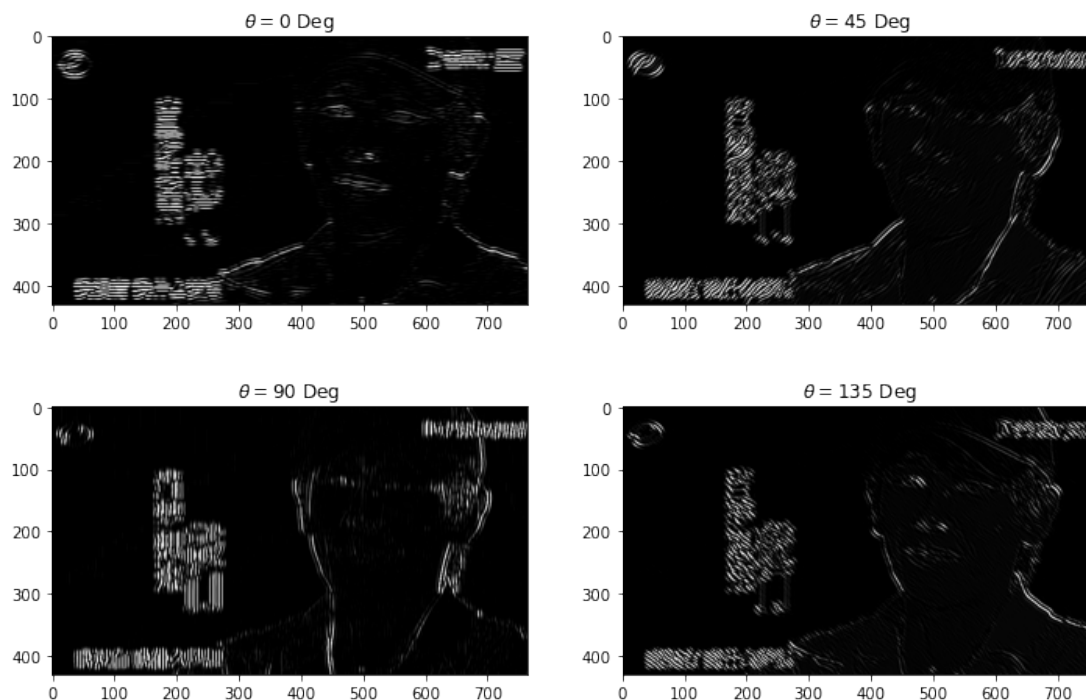


图 14 – 对综艺节目截图应用 Gabor 滤波器后的结果图

可见，对于比较丰满的汉字“当梦想来敲”以及下方的字幕和右上角的说明，在四张图上的响应都较大；对于简单的物体轮廓，如人脸、衣着等，响应很稀疏或者很低；对于不属于汉字的图形，如左上角的电视台台标，只在某一或两张结果图上有较大的响应。据此，可以设计后续算法截取出图像中的汉字部分。^[6]

Gabor 滤波器组的特征检出能力还可以用于光学字符识别（OCR）。以下图的“惑”字为例，右边四张子图是不同 θ 下的 Gabor 滤波器响应，左上子图是原图，左下子图是右边四张子图的重组结果。可见，Gabor 滤波器有效地将该汉字编码为四张子图描述的特征。可对大量汉字重复该操作，即可为每个汉字构建四个特征向量。后续采用一定判别法，如 K 近邻，来求取新来的汉字与数据库中的特征向量相近的汉字，即可完成字符识别任务。同样的处理方法与思路还可用在如指纹、掌纹等其他纹理特征的识别上。^[1]

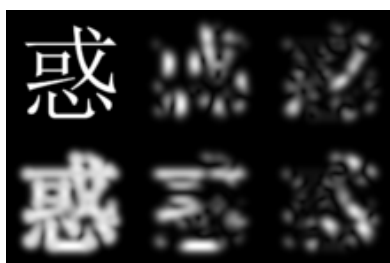


图 15 – 汉字“惑”及其 Gabor 滤波器结果 [1]

四、Gabor 滤波器与深度学习——以卷积神经网络为例

本节对应代码在 `gabor_4.ipynb`

有一些研究指出，Gabor 滤波器在频率和方向表示上的特征提取能力与动物

的视觉系统相似。例如，一篇发表在 *Nature* 上的论文通过对哺乳动物的初级视觉的简单细胞的感受野进行建模、优化后得到的滤波器的基函数及系数表示如下图所示^[7]。它们与 Gabor 滤波器的形态很相似。

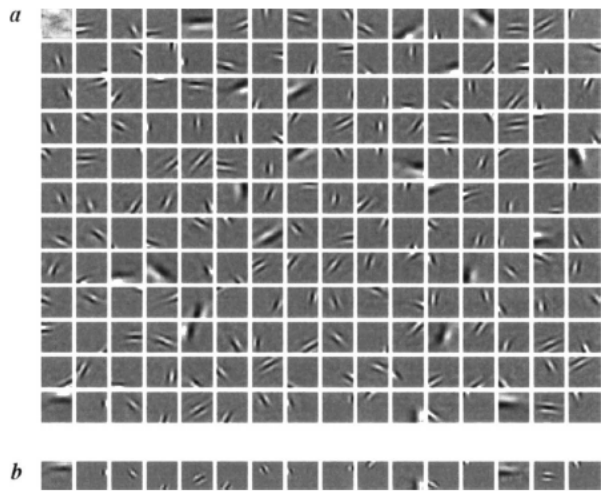


图 16 – 简单视觉细胞的感受野的基函数(a)及系数(b) [7]

深度学习中的卷积神经网络作为试图模拟人类视觉特征的一种人工智能手段，来完成计算机视觉、图像处理方面的任务，其中也有 Gabor 滤波器的身影。我们以一个在 MNIST 数据集^[8]上训练后达到 98%准确率的用于手写数字识别的经典卷积神经网络 LeNet-5^{[9][10]}为例，可视化出第一层卷积层的权重，如下图所示。

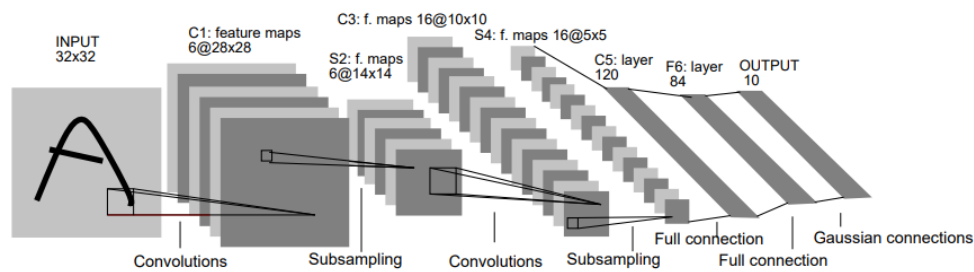


图 17 – LeNet-5 的网络结构 [9]

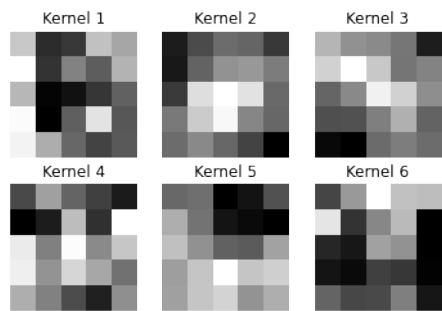


图 18 – LeNet-5 的第一层卷积核权重图

在这些卷积核中，核 2、核 3、核 6 都在不同程度上表现出 Gabor 滤波器的明暗相间、逐渐衰减的特点。同样的现象也出现在 AlexNet 等卷积神经网络中^[11]。亦有研究将 Gabor 滤波器与卷积神经网络显式结合，以取得更好效果^[11]。

五、总结

本报告首先介绍了 Gabor 滤波器的数学表达式、获取流程以及各参数的作用。接着,通过 Gabor 滤波器在布料瑕疵检测、汉字区域检测、汉字特征提取与编码方面的作用,以及其与卷积神经网络的关系,使读者能进一步了解 Gabor 滤波器的应用场景与内在特点。尽管卷积神经网络等工具使得特征的构造、图像的处理变得自动化,传统的滤波器与特征描述手段仍然值得我们学习,并从中挖掘与可与新技术相结合的创新点。

参考资料

- [1] Wikipedia. Gabor filter[EB/OL].2022.https://en.wikipedia.org/wiki/Gabor_filter.
- [2] OpenCV. cv::getGaborKernel[EB/OL].2022.
<https://github.com/opencv/opencv/blob/4.x/modules/imgproc/src/gabor.cpp>.
- [3] aitex. AITEX FABRIC IMAGE DATABASE[DB/OL].2019.
<https://www.aitex.es/afid/>.
- [4] Javier Silvestre-Blanes, Teresa Albero-Albero, Ignacio Miralles, Rubén Pérez-Llorens, Jorge Moreno. AFID: a public fabric image database for defect detection.[J]. AUTEX Research Journal, No. 4.2019.
- [5] A. Kumar, Grantham Kwok Hung Pang. Defect detection in textured materials using Gabor filters[J]. IEEE Transactions on Industry Applications.2002.
- [6] 温素年. Gabor 滤波器及其在 OpenCV 中的使用[EB/OL].2018.
<https://www.jianshu.com/p/f1d9f2482191>.
- [7] Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive-field properties by learning a sparse code for natural images[J]. Nature.1996.
- [8] Yann LeCun, Corinna Cortes and Christopher J.C. Burges. THE MNIST DATABASE of handwritten digits[DB/OL].2022 (fetch).<http://yann.lecun.com/exdb/mnist/>.
- [9] Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE.1998.
- [10] ChawDoe. LeNet5-MNIST-PyTorch[CP/OL].2022 (fetch).
<https://github.com/ChawDoe/LeNet5-MNIST-PyTorch>.
- [11] S Luan, C Chen, B Zhang, J Han and J Liu. Gabor Convolutional Networks[J]. IEEE Transactions on Image Processing.2018.

对于图片引用,已在图片题注处标明。