

# 历年卷

---

- [2022-2023](#)
- [2021-2022](#)
- [2020-2021](#)
- [2019-2020](#)

## 基础知识

---

### 1. 数字图像

- 数字图像是呈现信息主要形式，是计算机视觉和计算机图形学的基石，60%信息由视觉获取。
- 应用：计算机动画、视频监控、视觉效果增强、医学图像处理
- 图像信息处理：获取、展示、储存传输、增强修复、识别理解

### 2. 图像数据结构

- 图像格式BMP（无压无损）、JPG（有损）、GIF（无损）、PNG（无损）、TIFF（无压无损）
- 无压缩BMP
  - 组成：位图文件头、位图信息头、调色板、图像数据
  - 位图文件头（2、4、2、2、4）
    - 2 bfType：说明文件的类型，该值必需是0x4D42，也就是字符'BM'。
    - 4 bfSize：说明该位图文件的大小，用字节为单位
    - 2 bfReserved1：保留，必须设置为0
    - 2 bfReserved2：保留，必须设置为0
    - 4 bfOffBits：说明从文件头开始到实际的图象数据之间的字节的偏移量。位图信息头和调色板的长度会根据不同情况而变化。一般为54。
  - 位图信息头：
    - 4 biSize：位图信息头所需要的字数。
    - 4 biWidth：图象的宽度，以像素为单位。
    - 4 biHeight：说明图象的高度，以像素为单位。注：这个值除了用于描述图像的高度之外，它还有另一个用处，就是指明该图像是倒向的位图，还是正向的位图。如果该值是一个正数，说明图像是倒向的，如果该值是一个负数，则说明图像是正向的。大多数的BMP文件都是倒向的位图，也就是时，高度值是一个正数。
    - 2 biPlanes：为目标设备说明位面数，其值将总是被设为1。
    - 2 biBitCount：一个像素所用比特
    - 4 biCompression：说明图象数据压缩的类型，我们只讨论没有压缩的类型：BI\_RGB。0
    - 4 biSizeImage：说明图象的大小，以字节为单位。当用BI\_RGB格式时，可设置为0。
    - 4 biXPelsPerMeter：说明水平分辨率，用像素/米表示。
    - 4 biYPelsPerMeter：说明垂直分辨率，用像素/米表示。

- 4 biClrUsed: 说明位图实际使用的彩色表中的颜色索引数（设为0的话，则说明使用所有调色板项）。
- 4 biClrImportant: 说明对图像显示有重要影响的颜色索引的数目，如果是0，表示都重要。
- 调色板：
  - 四字节为一个单位：蓝绿红0
  - 作用：这样当我们表示一个像素的颜色时，只需引用该颜色在表中的索引值，节省储存空间
- 图像信息：
  - 行字节数需为4倍数，若不是需补齐
  - 图像储存原点为左下角，即倒置储存
- 景深：
  - 光圈：光圈孔径过大，会导致模糊，孔径过小，导致光强太弱
- 弥散圆：在焦点前后，光线开始聚集和扩散，点的影像变模糊，形成一个扩大的圆。不能辨认的直径叫做容许弥散圆
- 景深：焦点前后容许弥散圆的距离
  - 光圈越大，景深越小
  - 焦距越长，景深越小
  - 距离越近，景深越小
- 数码相机成像：
  - 景物反射光透过镜头，照射到CCD.
  - CCD的光电二极管受光刺激释放电荷，生成电信号
  - CCD控制芯片通过感光元件控制信号线，收集控制电流，输出到放大器
  - 电信号经放大和滤波后，输送到ADC，ADC将电信号，转换为数字信号，数值大小和电信号强度和电压有关。
  - 数字信号被输送到DSP，DSP对其进行色彩校正和白平衡处理，并将其编码为合适的图像格式。
  - 编码后的数据作为文件被储存
- 颜色空间
  - 彩色：红黄蓝等色及其混合色。消色：黑白灰。消色物体对不同波长反射无选择性
  - 视锥细胞：颜色。视杆细胞：光强。
  - 感知优先级：色调H>饱和度S>亮度V，敏感度：亮度V最敏感
  - 设备有关：RGB,CMY,HSV。设备无关：XYZ,Lab,YUV
  - 加色与减色（RGB与CMY）

R, G, B	C, M, Y
R+G=Y	Y+M=R
R+B=M	C+Y=G
G+B=C	C+M=B
R+G+B=W	C+M+Y=K
Color + light, energy increases.	Light-Pigment mixture, energy decreases.
Lightness increases	Lightness decreases
Additive color	Subtractive color
Color movies, TV, etc.	Painting, Photography, printing, etc.

- 有损压缩PG格式

- 思想：根据压缩比要求，从高频到低频逐步削减信息
- 优点：
  - 高频信息占用空间大，先消减可以获得大压缩比
  - 低频信息维持图像的基本轮廓和颜色分布，尽量保持了图像质量
  - 适合网络传输
- 行程编码：
  - 二值图像可以只记录0或1的位置，则图像可以用列表的列表表示，每一个子列表第一个元素为行号，后面一对数字  $(n_1, n_2)$  表示该行  $n_1$  到  $n_2$  列为该元素（0或1）占据。

## 二值化和形态学

### 1. 大津算法：

- 基本思想：将二值化后的图像视为前景和后景，找到阈值使得前景和内景内部方差最大且它们之前的方差最大。
- 令  $w_b w_f$  为前后景像素比例， $\mu_f \mu_b$  为前后景平均灰度值。则内部协方差：  

$$\sigma_{between} = w_b w_f (\mu_f - \mu_b)^2$$
- 操作步骤：
  1. 确定原始图像灰度最大值最小值
  2. 最小值加1作为thershhold进行二值化
  3. 计算此时二值化的外部协方差
  4. 回到第二步，直到进行到灰度最大值为止
  5. 找到外部协方差最大的thershhold
- 局部自适应：设定窗口大小和步长，对窗口内的像素找到局部thershhold进行二值化，直到遍历整个图像。

### 2. 形态学

- 膨胀： $A \oplus B = \{z | (B)_z \cap A \neq \emptyset\}$ 
  - 物理意义：将与物体边界接触的所有背景点合并到物体，边界扩张，填补空洞
- 侵蚀： $A \ominus B = \{z | (B)_z \subseteq A\}$ 
  - 物理意义：将边界点消除，边界收缩，去除小物体
  - 侵蚀与膨胀互补  $(A \oplus B)^c = A^c \ominus B$
- 填空算法：
  - 目的：对A填空
  - 操作  $X_k = (X_{k-1} \oplus B) \cap A^c$  在  $X_k = X_{k-1}$  时停下。
  - 结果： $X_k \cup A$
- 开运算：先腐蚀后膨胀
  - 消除小物体，分离物体，平滑边界
- 闭运算：先膨胀再腐蚀

- 连接物体，填补空洞，平滑边界

## 直方图

### 1. 对数可视增强

- $L_d = \frac{\log(L_r+1)}{\log(L_{max}+1)}$ , 灰度 $L$ 均已归一化
- 优势：使得最大亮度达到1，且其它亮度值较为平滑的改变

### 2. 直方图均匀化

- 令 $r$ 为改前灰度， $s$ 为改后像素， $\int_0^r P(r)dr = \int_0^s P(s)ds = \int_0^s 1 \cdot ds = s = T(r)$
- 均匀化后概率不等原因：所得 $s_k$ 不可能完全正好位于灰度级中的一个，只能就近归入最近灰度级，多个 $s$ 可能归入一个灰度级，最后在进行累加，结果就会导致灰度级概率不完全一样。
- 直方图匹配：本质上是减少灰度级增大对比度，可以突出重要的灰度范围，只需改变上面推导过程中的 $P(s)$ 即可。

## 几何变换

### 1. 简单几何变换

- 主要有：平移、镜像、旋转、缩放、剪切（矩形变平行四边形）
- 统一表示形式：
 
$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ 0 & 0 & g \end{pmatrix} \cdot \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix}$$
- 最近邻插值：将新图像像素点逆变换回原图像像素点，计算出坐标，找到最近的原图像像素点，将该像素的值赋给新图像像素
- 线性插值：同上计算原图像坐标，但是根据周边旁边几个像素点线性逼近此像素点
  - 一维：两点连线
  - 二维：如下
    1. 定义双线性方程 $g(x,y)=ax+by+cxy+d$ 。
    2. 分别将A、B、C、D四点的位置和灰度代入方程，得到方程组。
    3. 解方程组，解出a、b、c、d四个系数。
    4. 将P点的位置代入方程，得到P点的灰度。

### 2. Morph

- warp只改变位置，morph既改变位置也改变颜色
- 简单方法：
  - $I_i = I_{begin} = i * \frac{I_{end} - I_{begin}}{n}$
- 表情比例图复杂方法（差分向量关键点+同步亮度变化ERI）
  - 步骤：
    1. 对于图像 $A, A', B$ , 目标是仿照 $A$ 到 $A'$ 生成 $B'$ 。首先人工标记 $A, A', B$ ——对应的三组特征点
    2. 对 $B$ 的特征点，利用 $v_{b'} = v_b + v_{a'} - v_a$ 进行warp，令生成的图像为 $B_g$ 。
    3. 将 $A, A'$ 的特征点移动到与 $B_g$ 相同的位置，进行warp，即对齐
    4. 计算比例图： $R = \frac{A'}{A}$

$$5. B' = R \cdot B_g$$

# 滤波

## 1. 卷积

$$\circ g(x) = f(x) * h(x) = \int_{-\infty}^{+\infty} f(t) * h(x-t)dt = h(x) * f(x)$$

## 2. 滤波

### ○ 线性平滑滤波

$$\frac{1}{9} \cdot \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

### ○ 中值滤波：选取mask中的中位数填充进中心像素

### ○ 拉普拉斯滤波

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

#### ■ 拉普拉斯图像增强,令拉普拉斯滤波的结果为B

$$A' = \begin{cases} A + B & \text{如果卷积核中心为正} \\ A - B & \text{如果卷积核中心为负} \end{cases}$$

### ○ 双边滤波

#### ■ 核心思想：像素由周围像素的加权平均所取代，权值与像素点与目标像素相近程度和相似程度正相关，并且权值被归一化以保持局部均值。

$$\blacksquare BF[I]_p = \frac{1}{W_p} \sum_{q \in S} G_{\sigma_s}(\|p - q\|) G_{\sigma_r}(|I_p - I_q|) I_q$$

#### ■ p:目标像素点

#### ■ $\frac{1}{W_p}$ :归一化因子

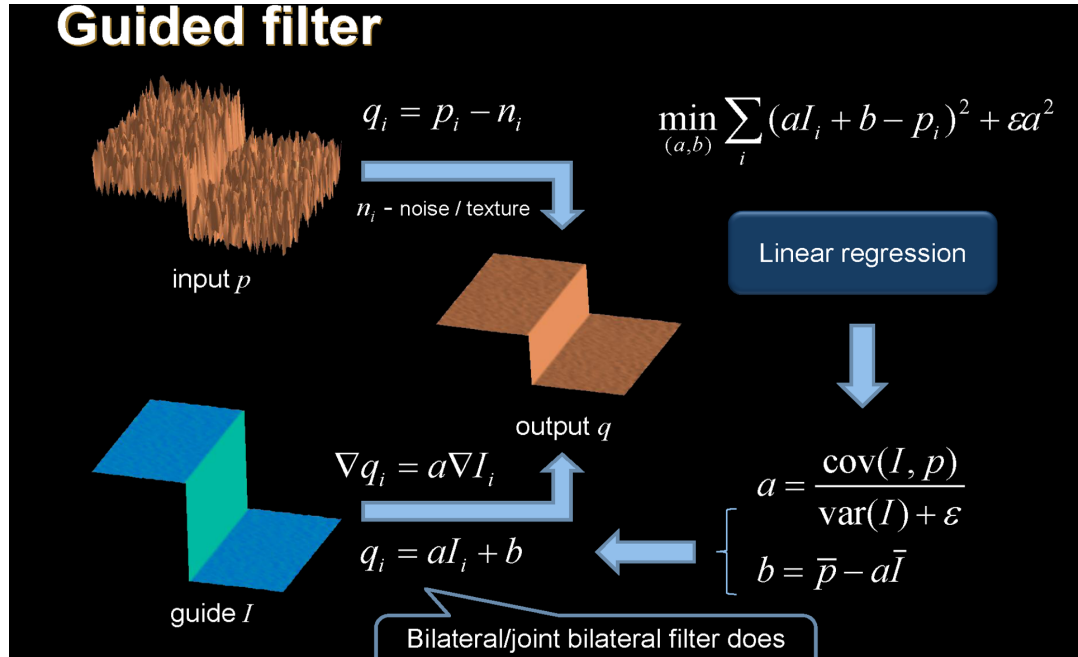
#### ■ q:核中一个像素点

#### ■ S: 卷积核

#### ■ $\sigma_s$ :空间系数，表征所考虑的邻域大小

#### ■ $\sigma_r$ :颜色系数，表征颜色变化振幅大小

- 导向滤波



- 边缘保持
- 梯度保持
- 时间复杂度 $O(1)$ ，快速
- 无需迭代
- 比平滑更通用

- 稀疏范数滤波

- $\min_{I_i^{new}} (\sum_{I_k \in D_i} (I_i^{new} - I_k)^q)$
- $q=1$ : 中值滤波;  $q=2$ : 线性平滑滤波

### 3. 傅里叶变换

- 正变换:  $F(u) = \sum_{x=0}^{N-1} I(x) e^{-j\frac{2\pi ux}{M}}$  for  $u = 0 \dots M-1$
  - 逆变换:  $I(x) = \sum_{u=0}^{M-1} F(u) e^{j\frac{2\pi ux}{N}}$  for  $x = 0 \dots N-1$
  - 快速傅里叶变换, 假设  $N = M$ , 且  $N = 2^n = 2 \cdot M$
- $$F(x) = \begin{cases} F_e(k) + F_o(k) e^{-j\frac{2\pi k}{2M}} & (x = k) \\ F_e(k) - F_o(k) e^{-j\frac{2\pi k}{2M}} & (x = k + M) \end{cases}$$
- 扩展到二阶

#### Forward DFT

$$F(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)},$$

$$u = 0, 1, \dots, M-1, v = 0, 1, \dots, N-1$$

#### Inverse DFT:

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi \left( \frac{ux}{M} + \frac{vy}{N} \right)}$$

- 傅里叶变换中相位比幅值更重要

## 特征点检测

## 1. Harris 角点检测

- W平方差:

$$\begin{aligned} E(u, v) &= \sum_{(x,y) \in W} [I(x+u, y+v) - I(x, y)]^2 \\ &\approx \sum_{(x,y) \in W} [I(x, y) + (I_x \ I_y) \cdot \begin{pmatrix} u \\ v \end{pmatrix} - I(x, y)]^2 \\ &= (u \ v) \cdot \sum_{(x,y) \in W} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} \end{aligned}$$

- 令H为中间的矩阵和，H的两个特征向量和特征值分别为： $x_+, x_-, \lambda_+, \lambda_-$ ，则：

- $x_+$ : 最大增加方向
- $x_-$ : 最小增加方向
- $\lambda_+$ : 最大增加倍数
- $\lambda_-$ : 最小增加倍数
- $\lambda_-$  最为重要

- Harris检测子

$$f = \frac{\lambda_+ \lambda_-}{\lambda_+ + \lambda_-} = \frac{H \text{的行列式}}{H \text{的迹}}$$

- 具有旋转不变性，和部分光照不变性 ( $I = aI$  或  $I = I + b$ )
- 不具有规模不变性

## 2. 规模不变检测

- 基本思想: 设计  $f = \text{Kernel} * \text{Image}$ , 找到f的极大值点

- $LoG = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$

- $DoG = G(x, y, k\sigma) - G(x, y, \sigma)$

- Harris-Laplacian检测

- 步骤

1. 多尺度Harris角点检测，保留任何尺度均为角点的点，找到局部极大值。
2. 对检测结果进行拉普拉斯操作，找出局部极大值点。在图像中给定的一个点，定义拉普拉斯响应值达到峰值的尺度为特征尺度。

- SIFT

- 检测

- 对图像进行DoG卷积操作
- 找出在三维空间中的极大值即为特征点，即待测点和它3\*3邻域空间（三个尺度）内的极大值

- 描述

- 特征点的方向

- 像素的方向

$$\begin{aligned} m(x, y) &= \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) &= \arctan 2((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \end{aligned}$$

- 根据上述公式，算出特征点特征尺度范围内所有像素的模长和角度，将角度划分为多个区间按模长投票得出像素方向。

- 基本思想:

- 在监测点周围选取16\*16检测窗口
- 计算每个像素的朝向（梯度角度-90）
- 去除较弱朝向
- 为剩下的朝向创建直方图
- 完整版本
  - 将16 \* 16的窗口划分为4 \* 4单元格，按特征点导向旋转窗口以确保旋转不变性
  - 计算每个单元格的八维直方图
  - 16个八维向量组成128维描述子
- 总输出
  - 128维描述子：16块区域\*8维直方图
  - 位置
  - 规模
  - 导向

### 3. SURF

#### ○ 基本思想

##### ■ 兴趣点检测

1. 计算积分图像
2. 使用二阶导数滤波器到图像上
3. 非极大值抑制，找到Hessian矩阵行列式本地极大值
4. 二次插值逼近

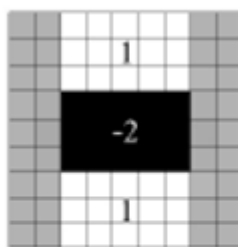
##### ■ 兴趣点描述

1. 将窗口划分为16个子窗口
2. 计算每个像素点haar小波输出
3. 在每个子窗口内计算 $[\sum dx, \sum dy, \sum |dx|, \sum |dy|]$
4. 生成64位描述子

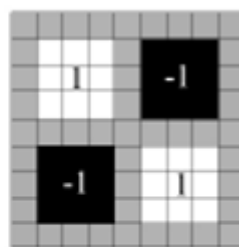
#### ○ 积分图像：

- 积分图像： $S(x, y) = I(x, y) + S(x - 1, y) + S(x, y - 1) - S(x - 1, y - 1)$
- 计算任何矩形的积分和 $V(l, r, d, u) = S(l, d) + S(r, d) + S(l, u) - S(r, u)$

#### ○ 二阶导数



**D<sub>yy</sub>**



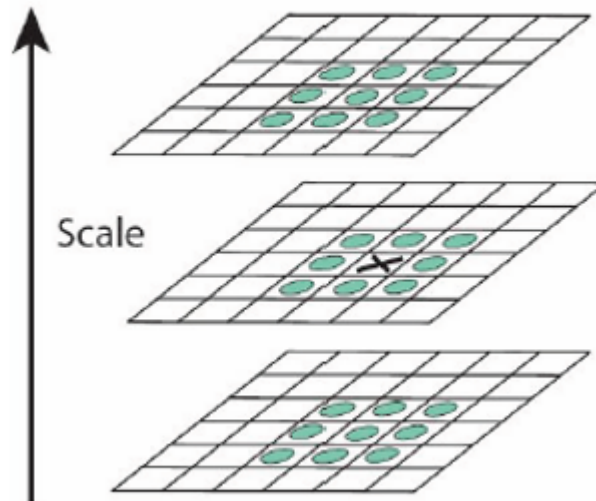
**D<sub>xy</sub>**

#### ○ 非极大值抑制

- Hessian矩阵 $\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2$



- 局部最大值



- 二次插值

- 使用泰勒展开求导逼近

Taylor expansion:

$$H(\mathbf{x}) = H + \frac{\partial H^T}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 H}{\partial \mathbf{x}^2} \mathbf{x}$$

Solution using Newton's method:

$$\hat{\mathbf{x}} = -\frac{\partial^2 H^{-1}}{\partial \mathbf{x}^2} \frac{\partial H}{\partial \mathbf{x}} \quad \frac{\partial^2 H}{\partial \mathbf{x}^2} = \begin{bmatrix} d_{xx} & d_{yx} & d_{sx} \\ d_{xy} & d_{yy} & d_{sy} \\ d_{xs} & d_{ys} & d_{ss} \end{bmatrix} \quad \frac{\partial H}{\partial \mathbf{x}} = \begin{bmatrix} d_x \\ d_y \\ d_s \end{bmatrix}$$

- haar小波

和计算二阶导数思路相同



#### 4. 图像拼接

- 主要步骤

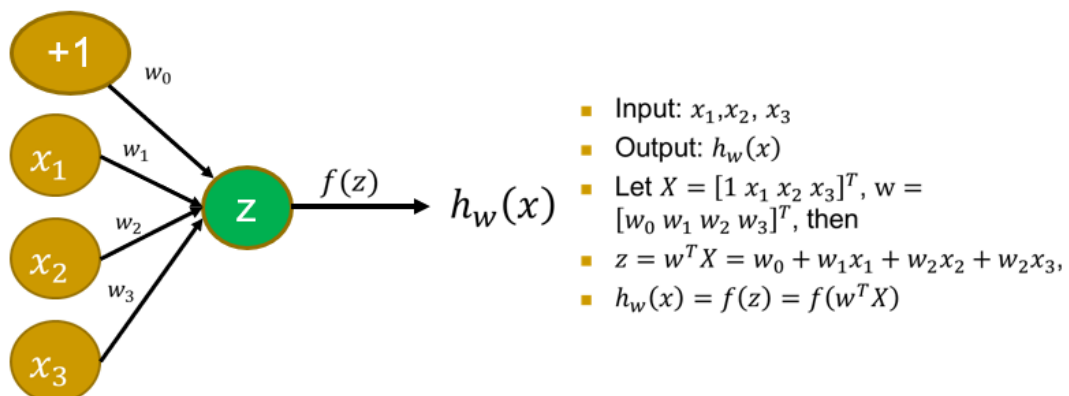
1. 计算特征点
2. 构建SIFT描述子
3. 匹配SIFT描述子
4. 计算描述子对变换矩阵
5. RANSEC
6. 图像融合

- RANSEC

- 基本思想
  - 目标：避免异常值的影响，寻找并只使用内点
  - 直觉：如果用一个异常值拟合模型，得到结果不会被许多其余数据支持
- 步骤
  1. 任选n个点
  2. 用n个点拟合变换模型
  3. 找到用该模型估计误差在允许范围内的点，记为内点
  4. 如果内点数量超过阈值，用所有内点重新计算变换模型，并更新最小误差变换模型
  5. 重复以上循环n次，存储内点支持数量最多的最小误差模型
- 迭代次数
  - 令w为内点比例，n为模型计算所需点数，k为迭代次数，则最小误差模型有外点混入的概率  $P = (1 - w^n)^k$
- 优缺点
  - 优点
    - 能处理各种拟合问题的通用方法
    - 容易计算，且故障率也易估计
  - 缺点
    - 在成本有限时，只能处理中等比例异常值
    - 许多实际问题有高异常值
- 图像融合
  - 拉普拉斯金字塔：保留图像高频细节，进行融合
  - 高斯金字塔：保留图像基本结构和特性，进行融合

## 深度学习

### 1. 神经元



### 2. 反向传播

1. 随机化权重  $W_0$
2. 计算出估计值  $h_W(x)$
3. 计算误差  $E = (h_W(x) - y)^2$
4. 更正  $W$ ,  $W_{k+1} = W_k - \epsilon \frac{\partial E}{\partial W}$

### 3. 卷积与池化

- 卷积:

12	3	19
25	10	1
9	7	17

 $*$ 

1	2
3	4

 $=$ 

133	75
100	101

$$f[n, m] * h[n, m] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f[k, l] h[n - k, m - l]$$

- 卷积加速原因:

- 共享卷积核, 多个像素对应一个神经元
- 卷积核数量固定, 卷积核参数数量固定

- 池化: 空间上, 卷积核之间

- 处理后图像大小计算公式

$$a' = \frac{a + 2 * \text{边缘填充} - \text{卷积核尺寸}}{\text{步长}} + 1$$

### 4. 交叉熵

- $L = -\log(p_c)$ ,  $p_c$ 是实际所属类对应的估计概率。

### 5. 区别

- 传统算法: 卷积完成图像到特征, 分类算法完成特征到分类
- CNN: 直接从图像到分类