【第一部分】20个常考专题总结

专题一 颜色空间

1.RGB颜色空间

①三原色、加色

②视网膜有针对红、绿、蓝三种颜色的椎体，如果这三种椎体受到同等强度的刺激，那么会发生消色(achromatism)。

③与设备相关

2.HSV颜色空间

①Hue（色调）,Saturation（对比度）,Value(亮度)。

②人眼感知颜色，感知先后顺序：色调、对比度、亮度。

③人眼对亮度最敏感。

3.CMY颜色空间

①减色，Cyan（青色）,Magenta（洋红）,Yellow

②与设备相关

4.CIE颜色空间

①设备不相关

5.YUV颜色空间

色调、饱和度、亮度

6.彩色与消色

彩色：红黄蓝等单色，以及它们的混合色。对各个波长的光的反射具有选择性，所以在白光照射下呈现不同的颜色。

消色：白色、黑色，以及各种深浅不同的颜色。对各个波长的光的反射没有选择性。

7.视觉感知

①视杆细胞：数量多，对光敏感，无法区分颜色。

②视锥细胞：数量少，强光才能激发，可以区分颜色。

③视网膜上存在3种分别对红、绿、蓝光线的波长特别敏感的视锥细胞，以及相应的感光色素。

专题二 图像存储格式

1.JPEG格式

①采用行程编码，压缩率高。

②根据压缩比的要求，从高频到低频逐步削减信息。

③高频信息占用的存储空间比较大，减少高频信息，可以获得更高压缩比。

④低频信息可以保留物体的基本轮廓、色彩分布，最大程度维持图像原貌。

⑤适用于互联网图像。

2.行程编码：高压缩率图像，每一行都由子列表描述。列表的第一个元素是行号。第2-3,4-5，…个元素分别是像素的起始和结束点的列标。

比如11149，表示：行号是第1行，有一串像素从1开始，到1截止。有一串像素，从4开始，到9截止。

3.BMP图像格式：两个文件头+调色板。

专题三 二值图像

1.每个像素点像素值仅为0或1。

2.存储空间少、效率高，不适用于三维情形。

3.获取二值图像：设置阈值，亮度高于阈值即为白色，低于阈值即为黑色。

4.大津二值化：

（1）目的：最大化类间方差，最小化类内方差，获取较好的阈值。

（2）算法步骤：

①寻找图像中的最小、最大像素值。

②从最小像素值开始，每次+1，直到最大像素值，分别以此为阈值。

③在该阈值下，确定前景、背景。

④计算类内方差、类间方差。

⑤返回步骤②，直到阈值达到了图像中的最大像素值。

⑥最大类间方差、最小类内方差对应的阈值，即为最终所求阈值。

（3）寻找全局阈值可能不适用于一些情况，解决方法是：设置更小的窗口，对每个窗口分别运用大津法，寻找局部阈值进行二值化。

专题四 形态学操作

1.主要适用于二值图像，常见目的有：去噪、形状简化、分割、目标描述。

2.逻辑操作：与、或、异或、差

3.位移、对称

4.几种用结构元进行的常见形态学操作：

①膨胀：只要窗口中有1，目标像素就是1。作用：弥补图像中的空隙。

②腐蚀：窗口中全部都是1，目标像素才是1。作用：去除边界、小型噪声。

③腐蚀应用：边界提取。先得到腐蚀图像，原始图像减去腐蚀图像就是图像边界。

④膨胀应用：弥补空洞。先对图像进行膨胀，再跟A的补图像进行交运算，不断迭代直至图像不变。

⑤开操作：先腐蚀，再膨胀。作用：去除小型噪声，在较细的部分切割物体，边界平滑同时保留原有面积。

⑥闭操作：先膨胀，再腐蚀。作用：弥补小型空洞，连接邻近物体，在平滑边界的同时最大限度地保留面积。

专题五 视觉增强

1.灰度感知

①韦伯定律：

$$\frac{I_{max}}{I_{min}} = (1 + K_{weber})^{255}$$

2.视觉增强——对数化处理

注意：亮度值已经归一化，取值范围为[0,1]

①公式：

$$L_d = \frac{\log(L_w+1)}{\log(L_{max}+1)}$$

②最大亮度为1，其余亮度数值平滑变化。

③不出现梯度逆转，即原来相对较暗者，经过对数变换后仍然相对较暗。

3.对数函数曲线越靠近0，图像越陡，越往右图像越平缓。那么亮度低的部分，经过对数变换后亮度范围增大，对比度变大，实现了视觉增强

专题六 直方图均衡化与直方图拟合

1.算法步骤：

①计算Sk，也就是之前+当前灰度级像素的总比例。

$$P(r_k) = n_k / n$$

$$s_k = T(r_k) = \sum_{i=0}^{k} P(r_i) = \sum_{i=0}^{k} \frac{n_i}{n} = \frac{1}{n} \sum_{i=0}^{k} n_i$$

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02

Original histogram

Histogram equalization--Example

$$s_k = T(r_k) = \frac{1}{n} \sum_{i=0}^{k} n_i$$

Step 1: Compute s_k

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00

$$s_k = \sum_{k=0}^{k} P(r_k) = s_{k-1} + P(r_k)$$

②把Sk放入对应的灰度级。

Histogram equalization--Example

Step 2: Assign the computed s_k into one of the eight gray levels

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
$\sim s_k$	1/7	3/7	5/7	6/7	6/7	1	1	1

（1/7=0.14；2/7=0.29；3/7=0.43；4/7=0.57；5/7=0.71；6/7=0.86；7/7=1）

③根据重新放入的灰度级，对Sk进行重命名。

Step 3: rename s_k (=k/7)

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
$\sim s_k$	1/7	3/7	5/7	6/7	6/7	1	1	1
s_k	s_1	s_3	s_5	s_6	s_6	s_7	s_7	s_7

④对于重命名后的Sk,归并各个灰度级中的像素个数。

Histogram equalization--Example

Step 4: merge the pixel numbers n_k sharing the same s_k

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
$\sim s_k$	1/7	3/7	5/7	6/7	6/7	1	1	1
s_k	s_1	s_3	s_5	s_6	s_6	s_7	s_7	s_7
n_k	790	1023	850	985		448		

⑤计算各个新灰度级对应的频率。

Histogram equalization--Example $P(s_k) = n_k / n$

Step 5: Compute the corresponding probability $P(s_k)$

k	0	1	2	3	4	5	6	7
r_k	0	1/7	2/7	3/7	4/7	5/7	6/7	1
n_k	790	1023	850	656	329	245	122	81
$P(r_k)$	0.19	0.25	0.21	0.16	0.08	0.06	0.03	0.02
s_k	0.19	0.44	0.65	0.81	0.89	0.95	0.98	1.00
$\sim s_k$	1/7	3/7	5/7	6/7	6/7	1	1	1
s_k	s_1	s_3	s_5	s_6	s_6	s_7	s_7	s_7
n_k	790	1023	850	985		448		
$P(s_k)$	0.19	0.25	0.21	0.24		0.11		

2.【问题】为什么经过直方图均衡化，所得的直方图不是完全均衡的?

【答】在算法的第二步中，Sk并不是完全等价于八个灰度级之一，而是被归入离它最近的灰度级。所以多个相邻的Sk可能落入同一个新的灰度级中。因此在离散直方图均衡化操作后，得到的不同灰度级对应的概率可能不一致。

In step 2, s_k is not exactly equivalent to one of the eight gray levels. Hence, it is included by a nearby gray level. So several neighboring s_k may falls into the same gray level. After discrete histogram equalization, different gray levels' probability may be different.

3.直方图均衡化的本质：增强图像对比度，降低图像的灰度级数。

4.直方图拟合

①拟合另一个图像的直方图，或者一个预先设定好的分布。

②目的：增强感兴趣的灰度级，改善图像质量。

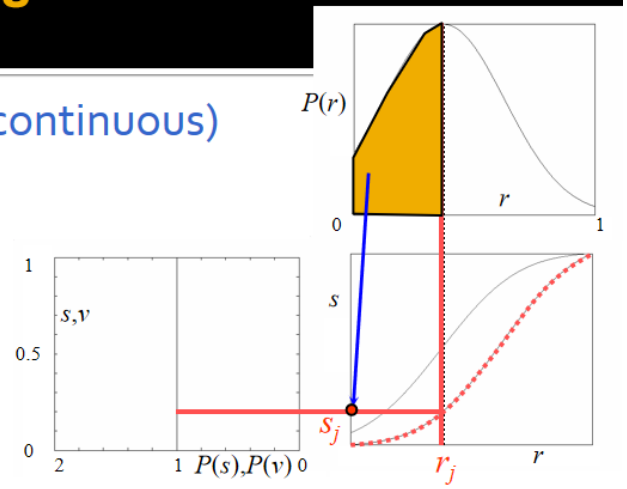
③通过直方图均衡化，实现直方图拟合。

5.连续情形下的直方图拟合算法

①把概率函数映射为和函数

Histogram fitting

Histogram fitting (continuous)



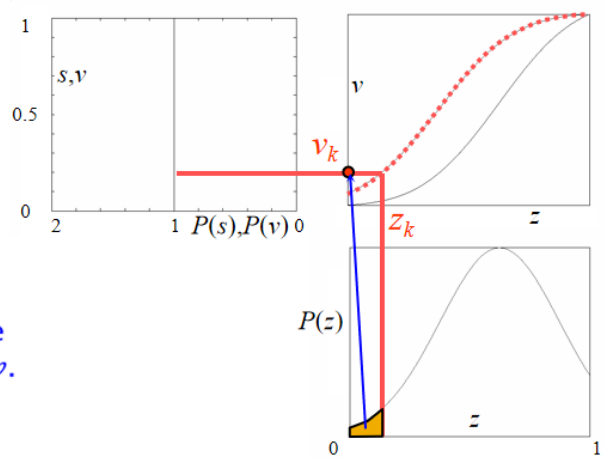
Step 1: Base on the equation

$$s = T(r) = \int_0^r P(r)dr$$

map r in the original histogram to be s.

②把灰度级z映射到目标直方图的纵坐标v。

Histogram fitting (continuous)



Step 2: based on equation

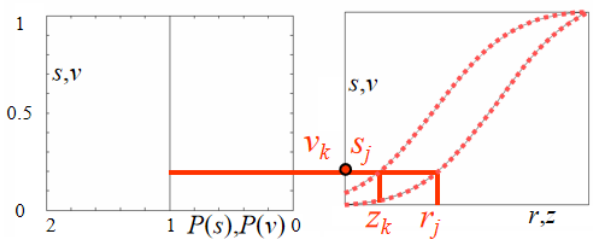
$$v = T(z) = \int_0^z P(z)dz$$

map the gray level z in the resulted histogram to be v.

③通过逆变换获取r。

Histogram fitting (continuous)

Step 3: obtain $z = G^{-1}(v)$ from $v = G(z)$.
$v = s$ because s and v have the same distribution.
Obtain $z = G^{-1}(s)$ corresponding r。



专题七 直方图变换

1.直方图变换的目的：增强图像，强化重要信息，弱化不重要信息。

2.常见的直方图变换：

①亮度变换

②对比度调整：如果使用线性变换，那么斜率 > 1时，对比度增强，对比度范围扩大；斜率 < 1时，对比度减弱，对比度范围缩小。

③色彩量化：可以用于图像分割、图像检索。具体方法是：把某一个区间内的灰度值全部映射为一个常数值。

3.直方图变换：线性变换、非线性变换

①线性灰度变换：

斜率与1的关系影响对比度，截距影响的是起始灰度。

②线性灰度变换可以实现对比度范围拉伸。

③可以进行分段变换，也就是只对图片的一部分灰度范围进行增强或抑制。目的是：对于感兴趣的灰度范围，进行对比度加强；对于不感兴趣的灰度范围，进行对比度弱化。

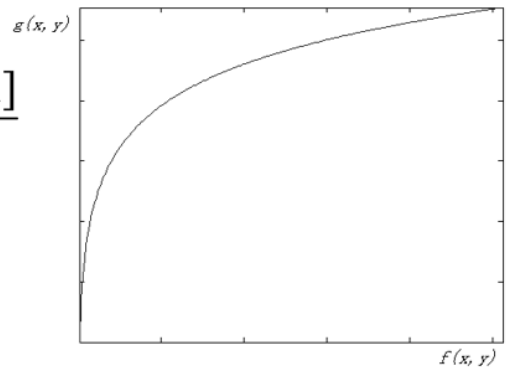
4.直方图非线性变换：指数函数、对数函数较为常用。

①对数函数变换：对于灰度低的区域，有对比度加强作用；对于灰度高的区域，有对比度削弱作用。

Logarithmic transform

$$g(x,y) = a + \frac{\ln[f(x,y)+1]}{b\ln c}$$

a,b,c are all adjustable

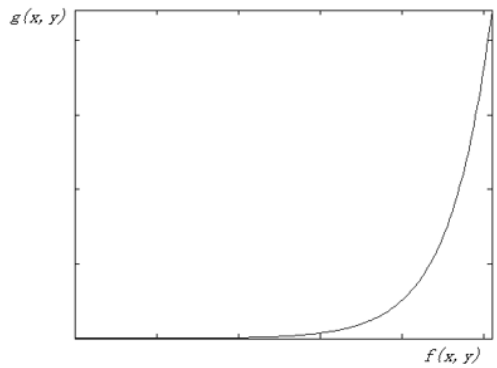


②指数函数变换：对于灰度值较高的区域，增强对比度；对于灰度值较低的区域，削弱对比度。

Exponential transform

$$g(x,y) = b^{c[f(x,y)-a]} - 1$$

a,b,c are all adjustable



专题八 几何变换

1.总述

①几何变换方程：

Equation 新的图像 原来的图像
（x,y）表示像素在二维画面上的位置

$$g(x,y)=f(x',y')=f[a(x,y),b(x,y)]$$

OR

$$\begin{cases} x'=a(x,y) \\ y'=b(x,y) \end{cases}$$

f(x,y): input image;
g(x,y): output image;
a(x,y), b(x,y): describe the spatial transformation

②简单变换、复合变换

2.位移

Translation——Equation

$$\begin{cases} x'=x+x_0 \\ y'=y+y_0 \end{cases}$$

OR

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}=\begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Each pixel in the original image is translated x_0 and y_0 respectively.

实现时，画布往往需要变大。

3.镜像变换——关于X或Y轴进行翻转

Flip around x axis:

$$\begin{cases} x' = x \\ y' = -y \end{cases}$$

Flip around y axis:

$$\begin{cases} x' = -x \\ y' = y \end{cases}$$

Mirror——Equation

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

When $S_x = 1$, and $S_y = -1$, flip around the x axis
When $S_x = -1$, and $S_y = 1$, flip around the y axis

4.旋转变换（注意旋转的中心点）

Rotation——Equation

$$\begin{cases} x' = x\cos\theta - y\sin\theta \\ y' = x\sin\theta + y\cos\theta \end{cases}$$

OR

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

单纯旋转带来的问题：会有一些像素点因为坐标计算四舍五入而覆盖不到，从而呈现白色，使得旋转后图像出现许多白色的小像素点。解决方法是插值，可用最近邻插值法、双线性插值法等。

5.放缩

放缩同样需要注意插值问题。

Scale——Equation

$$\begin{cases} x' = cx \\ y' = dy \end{cases}$$

➢Scale the image horizontally with coefficient c (enlarge when c>1, shrink when o<c<1); scale the image vertically with coefficient d (enlarge when d>1, shrink when o<d<1)
➢If c=d, the image is scaled with the same ratio. If NOT, the image is scaled horizontally and vertically with different ratios.
➢Shrink (down-sampling)——construct a new image by select pixels by predefined intervals from the original image.
➢Enlarge (a simple up-sampling)——there will be blank rows and columns. If you fill them by using interpolation, you will find some “mosaics” in the result.

6.错切

Shear——Equation

Shear on x axis

$$\begin{cases} a(x,y) = x + d_x y \\ b(x,y) = y \end{cases}$$

Shear on y axis

$$\begin{cases} a(x,y) = x \\ b(x,y) = y + d_y x \end{cases}$$

专题九 插值方法

1.最近邻插值法

通过几何变换的逆变换，寻找在原始图像中对应的像素点（一般是小数坐标），然后通过四舍五入方法找到最近邻像素点。

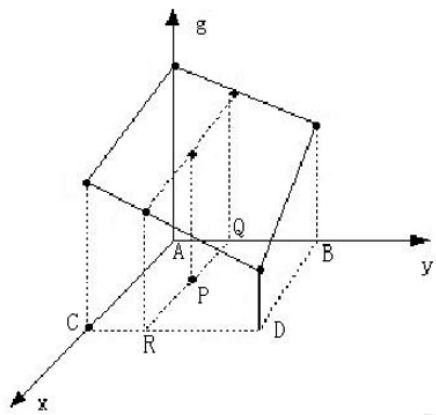
最近邻插值——计算过程

为了计算几何变换后新图像中某一点*P'*处的像素值，可以首先计算该几何变换的逆变换，计算出*P'*所对应的原图像中的位置*P*。通常情况下，*P*的位置不可能正好处在原图像的某一个像素位置上（即*P*点的坐标通常都不会正好是整数）。寻找与*P*点最接近的像素Q，把Q点的像素值作为新图像中*P'*点的像素值。

2.双线性插值法

Linear interpolation——Equation (2D)

1. Define the bilinear equation $g(x,y) = ax + by + cxy + d$;
2. Substitute the coordinates and grayscale intensity of A, B, C and D into the equation, and an equation system can be obtained.
3. Solve the equation system to obtain a, b, c, and d.
4. Substitute the coordinate of P into the equation and obtain its grayscale intensity.



首先通过几何变换逆变换计算出对应于原始图像的小数坐标，那么小数坐标对应周围有4个像素点，根据小数坐标和4个像素点的邻近程度，赋予4个像素点不同权重，然后线性求和即可。

3.径向基函数插值法（RBF）

Radial Basis Function (RBF) based interpolation

径向基函数插值

$$G(x) = \sum_{i=1}^{n} w_i G(c_i),$$

$$\text{where } w_i = \frac{\varphi(|x-c_i|)}{\sum_{i=1}^{n} \varphi(|x-c_i|)}$$

where x is a scalar or a vector. Hence, RBF can do 1D interpolation , 2D interpolation or high dimensional interpolation. It depends on the dimensionality of x.

常用的径向基函数有：

A radial basis function (RBF) is a function based on a scalar radius.

$$\varphi(r) = \varphi\left(\left|x - x_i\right|\right)$$

Gaussian and Multiquadrics functions have adjustable parameter σ.

Kernel functions available in *RBF*

Gaussian: $\varphi(r) = \exp\left(-\frac{r^2}{2\sigma^2}\right)$

Multiquadrics: $\varphi(r) = \sqrt{1 + \frac{r^2}{\sigma^2}}$

Linear: $\varphi(r) = r$

Cubic: $\varphi(r) = r^3$

Thinplate: $\varphi(r) = r^2 \ln(r+1)$

专题十 Warp & Morph

1.Warp:只改变像素的位置（几何变换）

2.Morph:形态学变换，不仅改变像素位置，而且改变像素的数值。

3.关键帧选取：在大小、结构上相似，才能获得自然的morph结果。

4.结构获取：利用特征点定义结构，对特征点进行匹配。

5.定义特征点->对两幅图像的特征点进行匹配->对图像进行分割->运用插值方法

6.Morph 的具体方法：根据变换快慢，设置步长，将图A中每一点的RGB数值逐渐编程与图B中相同位置像素的RGB。可以选择：等比、等差等方法。

- Given a color image, RGB should be taken into account together.

$$r_{i,j} = r_{i,j}^{a} + \frac{r_{i,j}^{b} - r_{i,j}^{a}}{N} * n.....n = 0...N-1$$

$$g_{i,j} = g_{i,j}^{a} + \frac{g_{i,j}^{b} - g_{i,j}^{a}}{N} * n.....n = 0...N-1$$

$$b_{i,j} = b_{i,j}^{a} + \frac{b_{i,j}^{b} - b_{i,j}^{a}}{N} * n.....n = 0...N-1$$

专题十一 空间域滤波——卷积

1.连续情形下的一维卷积

定义：乘积之和，f(x)为输入函数，h(x)为卷积函数。

Continuous 1D convolution——Definition

The convolution $g(x)$ of two 1D functions $f(x)$ and $h(x)$

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

which means the convolution of two 1D functions can be formulated as an integral of multiplication.

Usually we call $f(x)$ as input function, $h(x)$ as convolution function.

2.连续一维卷积的性质

分配律、结合律、交换律

3.离散情况下的一维卷积

积分运算变成求和运算，注意求和之后除以求和序列长度进行标准化。

Discrete 1D convolution——

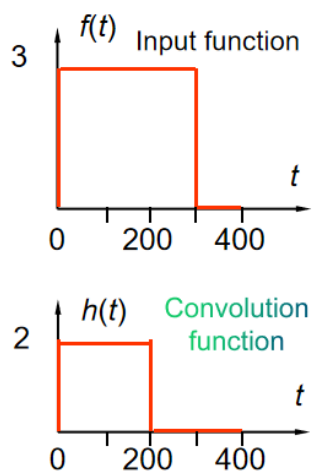
For a discrete sequence, the integral can be treated as summing up problem.

For $f(x)$ and $h(x)$ whose lengths are M, their convolution is：

$$g(x)=f(x)*h(x)=\frac{1}{M}\sum_{t=0}^{M-1}f(t)h(x-t)$$

4.举例

Example



Both $f(x)$ *and* $h(x)$ *have 400 samples, i.e.,* $x=0,1,2,\ldots,399$.

$$f(x)=\begin{cases}3 & 0\le x<300\\ 0 & 300\le x<400\end{cases}$$

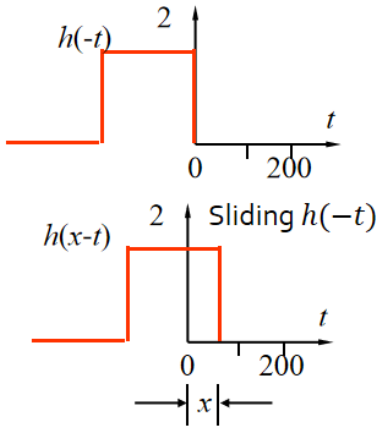
$$\mathrm{h}(x)=\begin{cases}2 & 0\le x<200\\ 0 & 200\le x<400\end{cases}$$

compute $g(x)$ at each x

$$\boxed{g(x)=f(x)*h(x)=\frac{1}{M}\sum_{t=0}^{M-1}f(t)h(x-t)}$$

给定输入函数和卷积函数，操作步骤：

①卷积函数先关于Y轴对称，再向右平移X个单位。

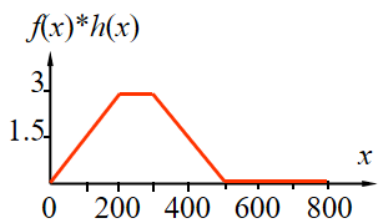


Step 2：

Slide $h(-t)$ by adding x, *i.e.*, *h*(*x*-*t*)。

$$g(x)=f(x)*h(x)=\frac{1}{M}\sum_{t=0}^{M-1}f(t)h(x-t)$$

②计算所有样本点的乘积f(t)h(x-t)，先求和，然后除以求和序列的长度M。



Step 3：

For each sample t （t = 0, 1, 2, 3, ..., 399）, compute $f(t)h(x-t)$;

Sum up all the $f(t)h(x-t)$ and divide it by M ($M = 400$) .

Now you get the convolution at x.

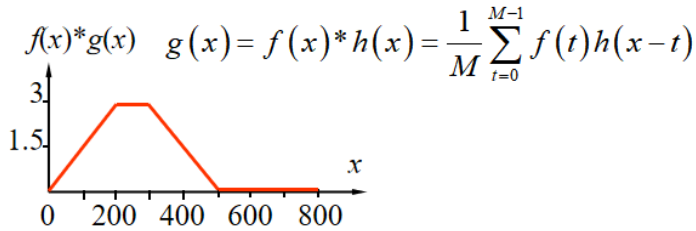
Step 4：

Compute the convolution value for each x in the “effective range”.

You get a curve of $g(x)$.

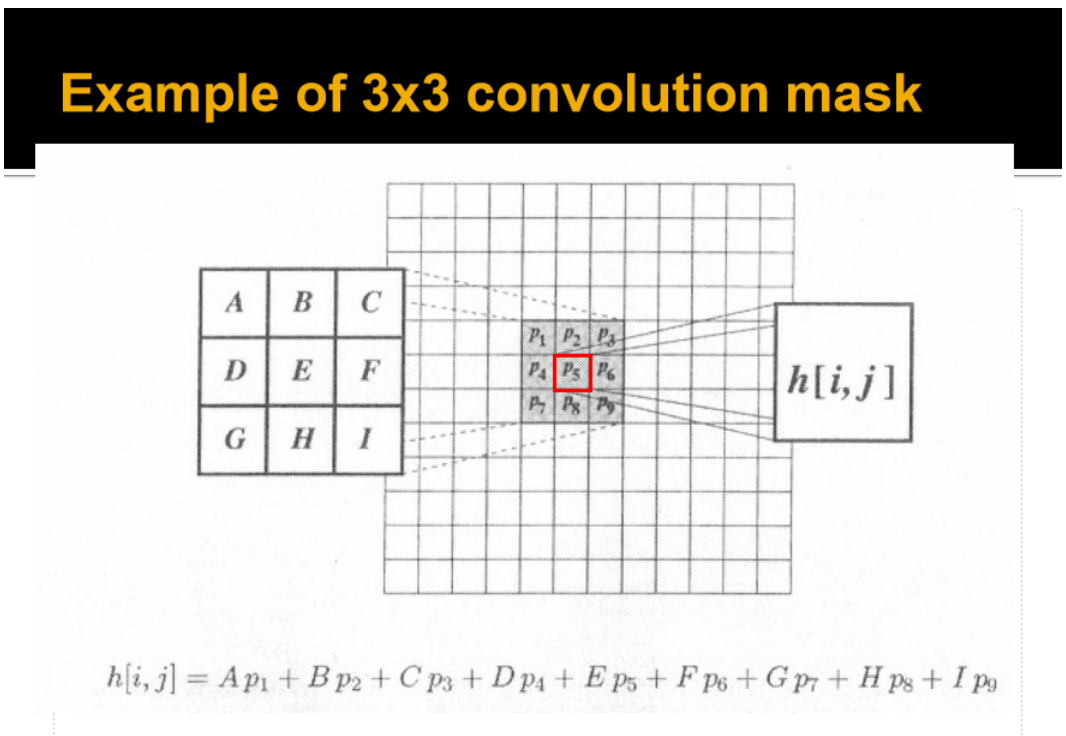
$$g(x)=f(x)*h(x)=\frac{1}{M}\sum_{t=0}^{M-1}f(t)h(x-t)$$

卷积结果：

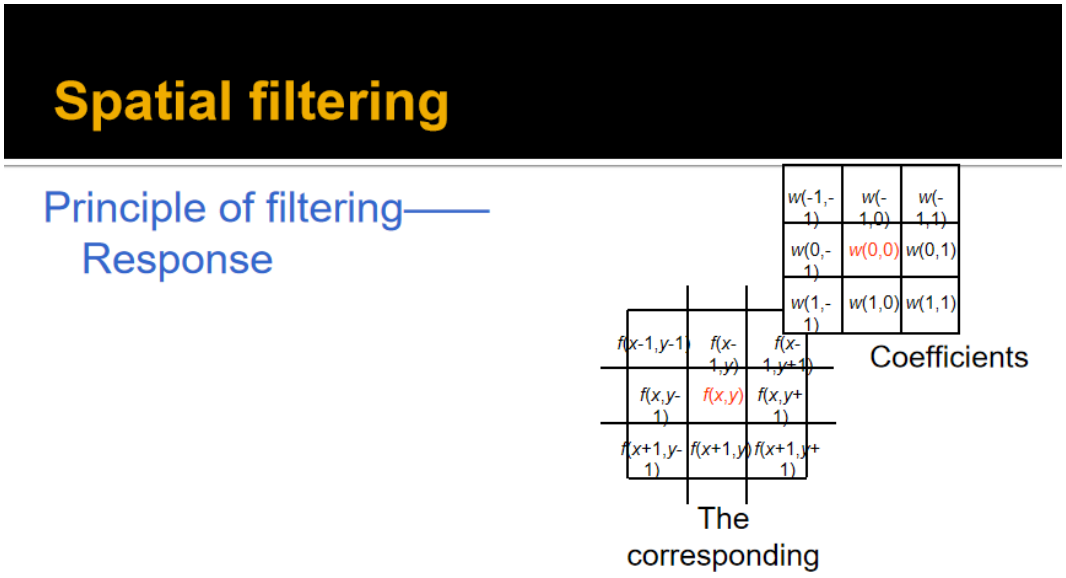


Hence the convolution $g(x)$ is also a function and its variable is x.

5.二维离散卷积



6.基于卷积操作的空间域滤波



$$R = w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + \cdots + w(0,0)f(x,y) + \cdots + w(1,0)f(x+1,y) + w(1,1)f(x+1,y+1)$$

Convolution: $g(x) = f(x) * h(x) = \frac{1}{M}\sum_{t=0}^{M-1} f(t)h(x-t)$

Principle of filtering——Response

Usually, the width and the height of the mask are odd number such as 2*a*+1 and 2*b*+1. When the mask window is centered at (*x*,*y*), the updated value of (x,y) is computed as:

$$g(x,y) = \sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)f(x+s,y+t)$$

After updating all the pixels in the image, you will obtain a new image *g*.

7.空间域滤波——平滑操作

①线性平滑滤波

本质是对窗口内的像素求加权平均值，作用是去除小型的噪声，模糊图像。

②统计排序滤波

对于窗口内的像素值进行排序，然后取中止、最大值、第K大值等，非线性滤波，得到的图像不会特别模糊。

专题十二 空间域滤波——锐化

1.锐化

利用微分算子，增强边界和变化明显的部分，包括噪声。弱化变化较为细微的部分。

2.一、二阶微分算子的离散定义

Differential——

For a function *f*(*x*)，we use difference to represent the differential operator:

$$\frac{\partial f}{\partial x} = f(x+1) - f(x)$$

Similarly, the second order differential is：

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x)$$

3.罗伯特交叉梯度算子

Gradient based method——Another approach

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Original image, z_i is the pixel value, z_5 is the center pixel.

-1	0
0	1

0	-1
1	0

Robert cross gradient operator, Gx=(z9-z5)
Gy=(z8-z6)

$$\nabla f = \left[(z_9 - z_5)^2 + (z_8 - z_6)^2 \right]^{\frac{1}{2}}$$

Or

$$\nabla f \approx |z_9 - z_5| + |z_8 - z_6|$$

4.拉普拉斯算子

Mask of Laplacian operator

$$\nabla^2 f = \left[f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) \right] - 4f(x,y)$$

0	1	0
1	-4	1
0	1	0

It is rotation invariant.

Extended mask $\nabla^2 f = \sum_{i=-1}\sum_{j=-1} f(x+i, y+j) - 9f(x,y)$

1	1	1
1	-8	1
1	1	1

It's also rotation invariant, i.e., isotropic.

对于拉普拉斯算子，把二阶导数跟原始像素值叠加，可以进行图像增强。注意符号的正负，最终要让中心元素的系数为正。

Image enhancement by Laplaician:

$$g(x,y) = \begin{cases} f(x,y) - \nabla^2 f(x,y) & \text{If the center element of the mask is negative} \\ f(x,y) + \nabla^2 f(x,y) & \text{If the center element of the mask is positive} \end{cases}$$

Fuse the original image and the Laplacian result can preserve the sharpening effect and restore the original visual information.

注意事项：

①卷积值赋值给新的图像。

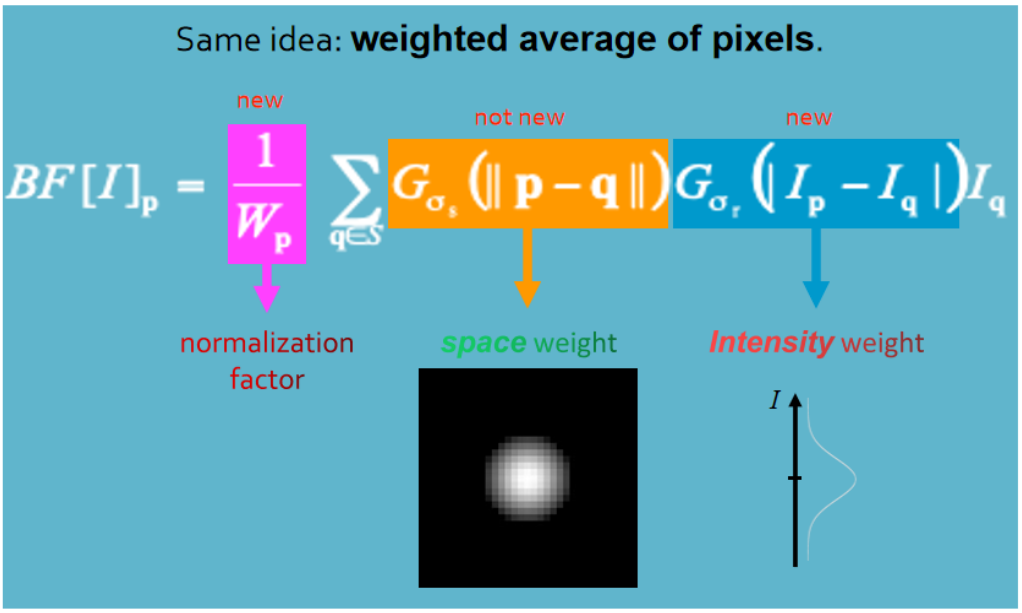
②注意卷积操作遇到图像边缘的处理方式。

③注意计算所得像素值的溢出判断。

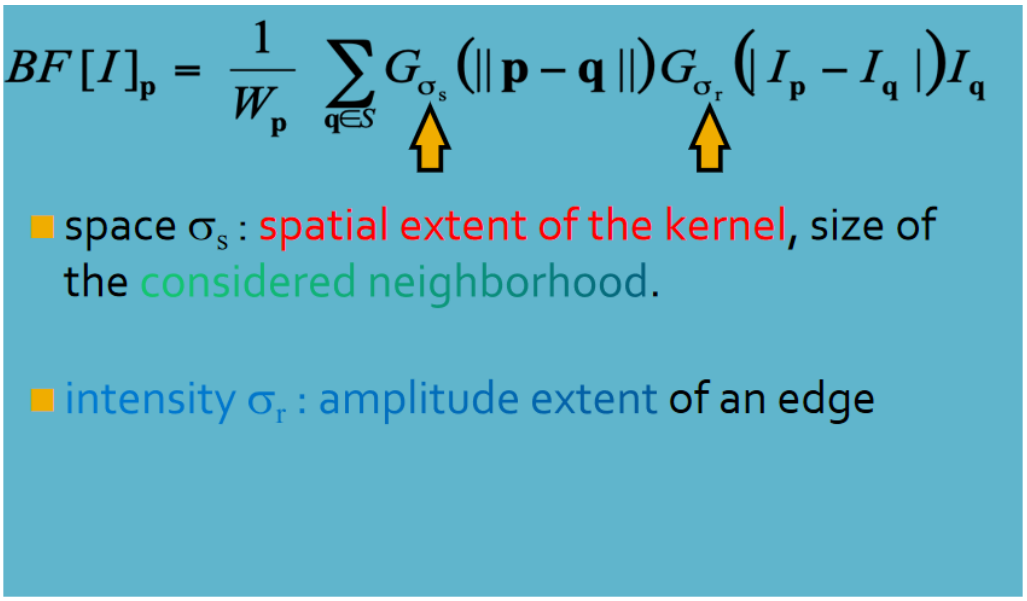
专题十三 双边滤波

1. 图像的基本特征：空间域S、亮度域R。

2.双边滤波同时考虑在两个域上进行作用，在两个域上分别寻找一个高斯函数，然后作乘积，作为像素点的权重，然后进行加权平均处理。



两个参数分别影响空间域、亮度域



3.参数设置：

Depends on the application. For instance:

- space parameter: proportional to image size
  - e.g., 2% of image diagonal
- intensity parameter: proportional to edge amplitude
  - e.g., mean or median of image gradients

空间域参数：对角线长度的2%。亮度域参数：图像梯度的中值或均值。

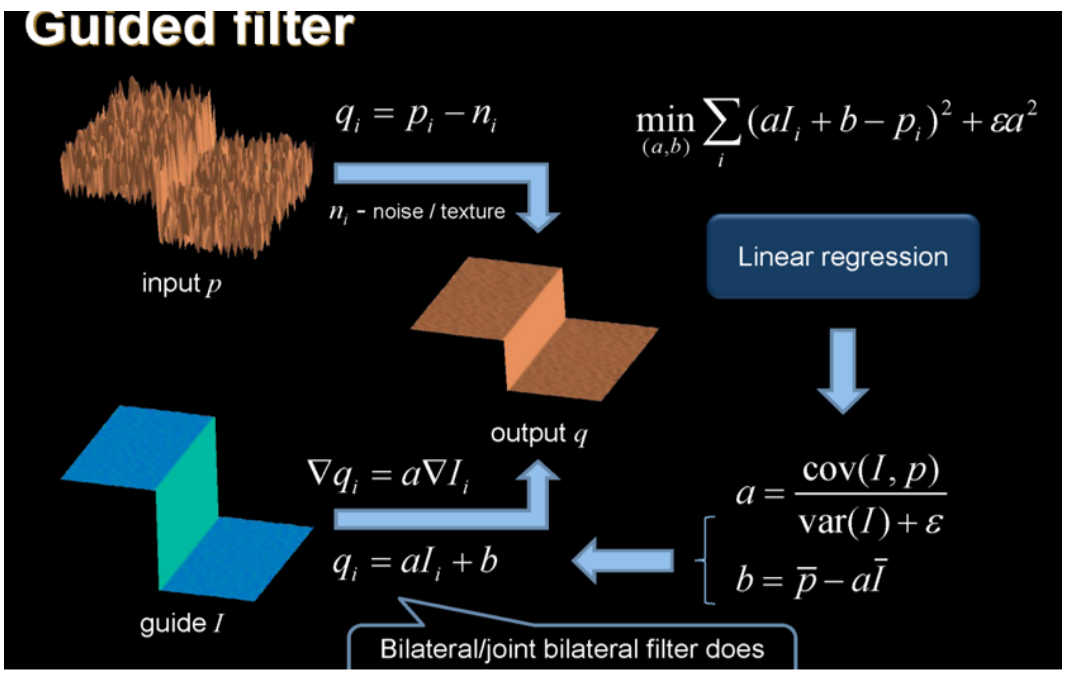
4.作用：去噪、色调映射、纹理编辑、重新打光。

5.问题：

①可能造成对比度降低。

②双边滤波保边，保的是差值的绝对值，所以会产生梯度逆转。

专题十四 引导滤波



1.双边滤波保边，保的是差值的绝对值，所以会产生梯度逆转。

2.引导滤波考虑了保梯度的方向，不会造成梯度逆转。

3.作用：细节强化、去噪、去雾。

4.优点：

①保边滤波，不会发生梯度逆转。

②非迭代。

③O（1）时间复杂度，快速准确。

④比平滑滤波器更通用。

专题十五 稀疏范数滤波（SNF）

1.平滑滤波的目的，即使得l^2范数最小。

The simplest way to smooth an image is to minimize the l^2 norm:

$$\min_{I_i^{new}} \sum_{j \in N_i} \left(I_i^{new} - I_j\right)^2 \quad \rightarrow \quad \text{Averaging Filter}$$

In order to let the filter be edge-preserving, we propose to change the norm into a sparser norm, and generalize into:

$$\min_{I_i^{new}} \sum_{j \in N_i} \left|I_i^{new} - I_j\right|^p, \ 0 < p \leq 2 \quad \rightarrow \quad \text{Sparse Norm Filter}$$

2.进行如下处理：

①为了实现保边滤波，对于范数，写成稀疏形式，即绝对值之和。

②计算的是距离P次方之和，P在（0,2）内取值。

专题十六 Fourier变换

1.泰勒展开

Example (Taylor Series)

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots.$$

constant

first-order term

second-order term

...

British mathematician Taylor invented a method to represent one analytic function by infinite linear combination of power functions.

2.Fourier级数

$$f(x) = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + \sum_{n=1}^{\infty} b_n \sin(nx),$$

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\,dx$$

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$$

Fourier series make use of the orthogonality relationships of the sine and cosine functions

思想：将函数分解为若干个频率不同、振幅可能不同的简谐振动的叠加。

3.Fourier图像

研究频率变化范围。Fourier图像中心，频率变化比较慢；图像四周，频率变化比较快。

对于正常图像，绝大多数频率集中在低频部分。

4.【问题】离散Fourier变换中，幅值和相位哪个更重要?

【答】相位更重要。相位可以重建图像，但是幅值不可以。

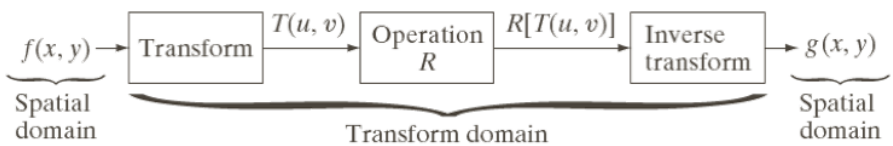
5.Fourier变换基本步骤

- Many times, image processing tasks are best performed in transformed domain other than the *spatial domain*.
- Key steps:

  (1) Transform the image

  (2) Carry out the task(s) in the *transformed domain*.

  (3) Apply *inverse transform* to return to the spatial domain.



①对图像进行Fourier变换。

②在频率域中进行处理。

③对于处理后的频率域，经过Fourier逆变换转换回空间域。

6.连续情形下的正逆Fourier变换，变为跟u有关的函数。

- Transforms a signal (i.e., function) from the spatial domain to the frequency domain.

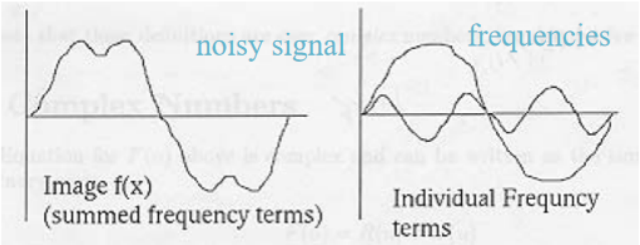
Forward FT: $F(f(x)) = F(u) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi ux}dx$

Inverse FT: $F^{-1}(F(u)) = f(x) = \int_{-\infty}^{\infty} F(u)e^{j2\pi ux}du$

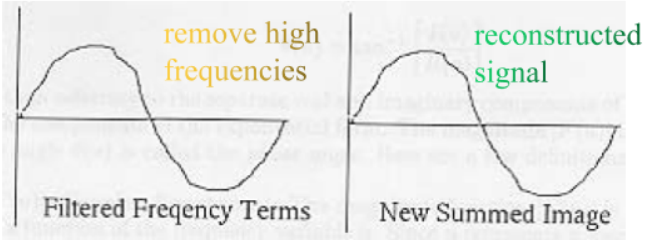
(IFT)

where $e^{\pm j\theta} = cos(\theta) \pm jsin(\theta)$

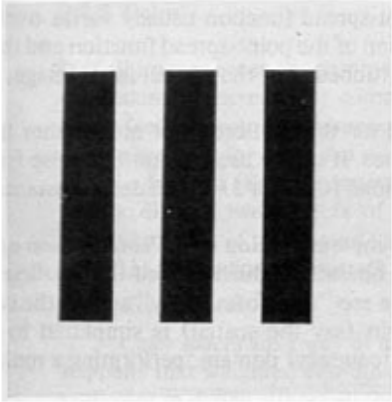
7.f(u)某些项的系数置0，可以去除一些不需要的频率。



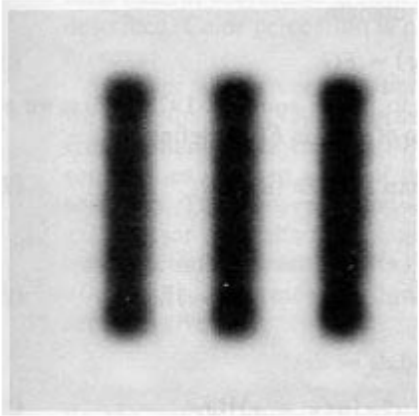
To remove certain frequencies, set their corresponding F(u) coefficients to zero!



8.频率域中的低通滤波、高通滤波、带通滤波



Original Image



Low-passed

9.离散Fourier正、逆变换

（注意逆变换的系数1/N）

- Forward DFT
- $F(u) = \sum_{x=0}^{N-1} f(x) e^{-\frac{j2\pi ux}{N}}$,
- $u = 0,1,\cdots,N-1$

- Inverse DFT
- $f(x) = \frac{1}{N}\sum_{u=0}^{N-1} F(u) e^{\frac{j2\pi ux}{N}}$,
- $x = 0,1,\cdots,N-1$

10.快速Fourier变换（FFT），降低时间复杂度。

专题十七 图像特征与Harris角点检测

1.图像特征的性质：旋转不变性、平移不变形、放缩不变性。

2.自动全景缝合（Automatic Panorama Stitching）步骤：

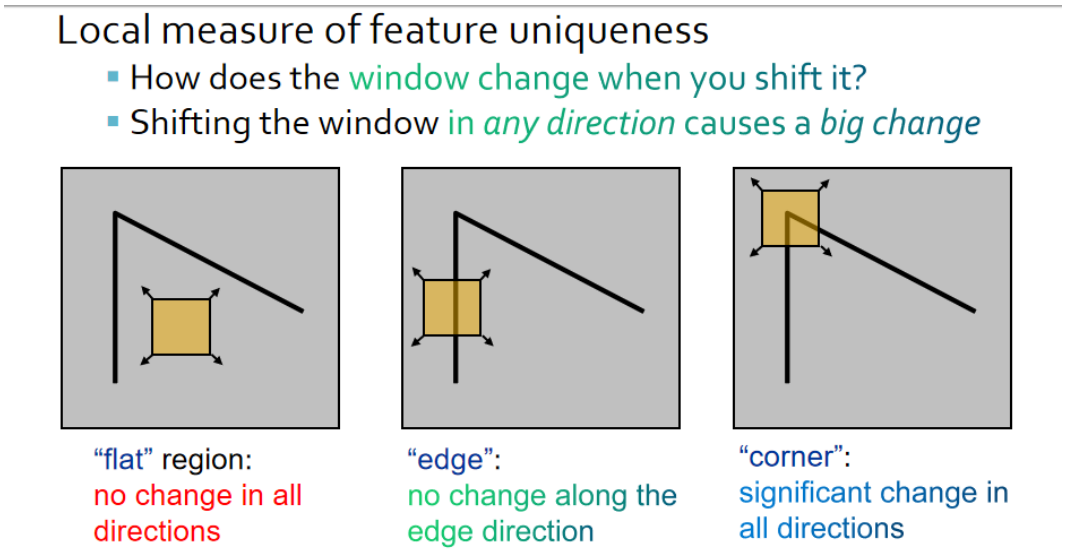
①寻找两幅图像中的特征点。

②对特征点进行匹配，形成特征点对。

③依据特征点对进行图像对齐。

3.Harris算子

①边缘、角点与平坦部分



平坦部分：窗口向任何一个方向移动，窗口内的像素都没有变化。

边缘：窗口沿着边缘方向移动，窗口内的像素没有变化。

角点：沿着任何方向移动，窗口内的像素都有明显变化。

②数学原理：

$$E(u,v)=\sum_{(x,y)\in W}[I(x+u,y+v)-I(x,y)]^2$$

③由于可以进行Tailor展开近似计算，所以像素变化可以写成矩阵乘积形式：

Taylor Series expansion of I:

$$I(x+u, y+v) = I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \text{higher order terms}$$

If the motion (u,v) is small, then first order approx is good

$$\begin{aligned} I(x+u, y+v) &\approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \\ &\approx I(x,y) + [I_x \;\; I_y] \begin{bmatrix} u \\ v \end{bmatrix} \end{aligned}$$

shorthand: $I_x = \frac{\partial I}{\partial x}$

Plugging this into the formula on the previous slide…

---

④根据特征值，可以判断变化大小。

矩阵如下：

$$\underbrace{\begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix}}_{H}$$

对于上述矩阵，变化的计算公式为：

$$\begin{aligned} f &= \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2} \\ &= \frac{determinant(H)}{trace(H)} \end{aligned}$$

其中，行列式为特征值乘积，迹为特征值之和。

4.Harris算子的一些性质

①平移、仿射、旋转不变性。

②亮度变化时，具有部分不变性。

5.Harris算子的缺点：不具有尺度不变性。

6.尺度的寻找：变换观察视野，找到一个合适的函数的极值。函数的极值应该是单峰、陡峭的。

- **A "good" function for scale detection: has one stable sharp peak**

- For usual images: a good function would be a one which responds to contrast (sharp local intensity change)

专题十八 SIFT算法与SURF算法

1.特征描述子：旋转不变性。

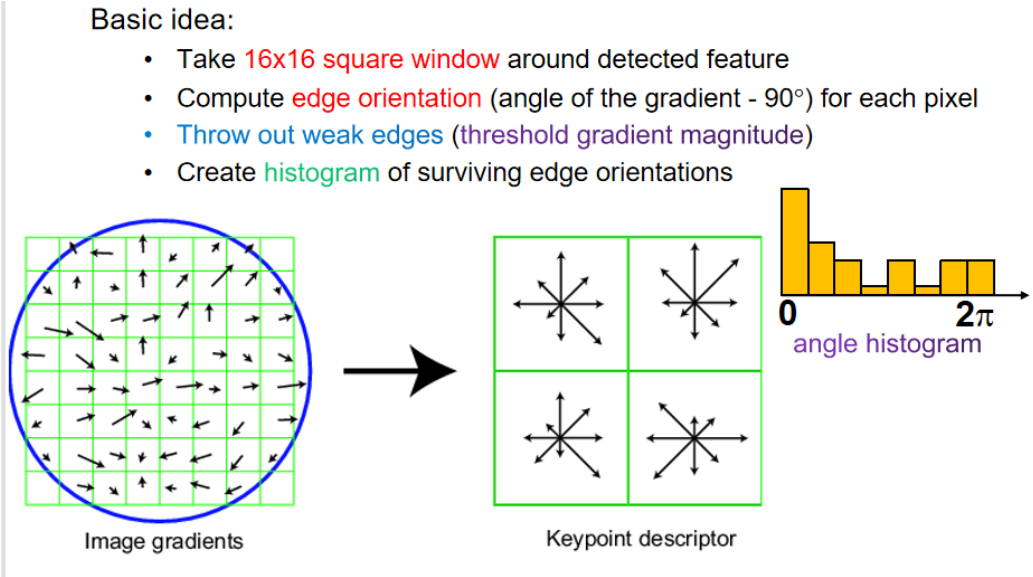
2.算法步骤

①对于检测特征，创建一个16×16的正方形窗口。

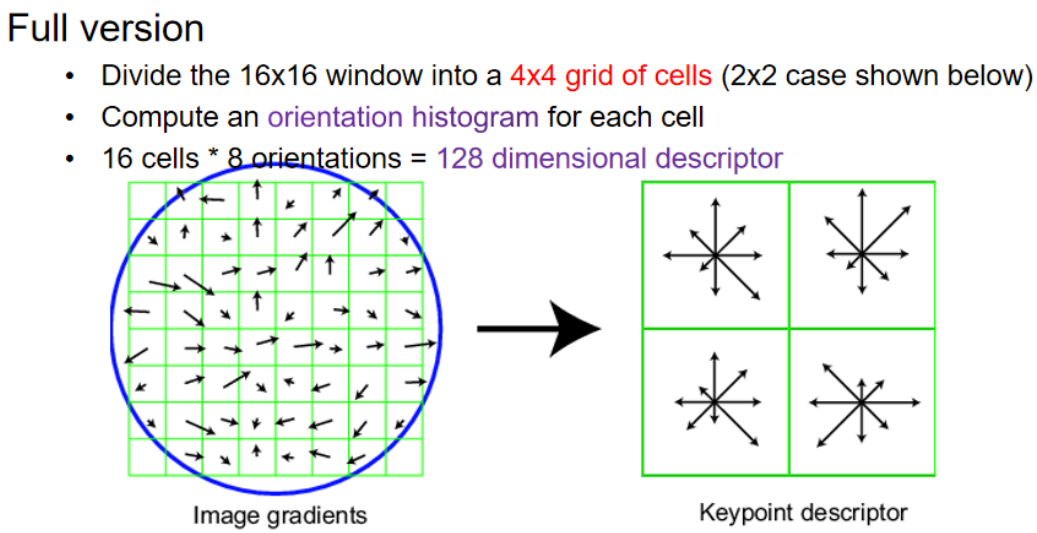
②计算边缘方向（梯度方向-90°）

③抛弃弱边（利用梯度阈值）

④利用剩余的边，创建直方图。



3.把16×16窗口分成16个4×4的窗口，每个窗口有8个方向，所以就形成了128维的描述子。



4.SIFT的优点

①尺度、旋转、亮度不变性。

②局部区域：具有高度独特性、描述性。

③对于刚体的表述特别有效。

5.SIFT缺点

①时间太长。

②对于非刚体，效果不好。

③严重扭曲下可能失效。

6.SURF——兴趣点检测

Interest point *detector*: 兴趣点检测

- Compute integral image 计算积分图
- Apply 2nd derivative (approximate) filters to image 获得图像的二阶导数（二阶差分）
- Non-maximal suppression 压缩非极大值 (Find local maxima in (x,y,σ) space)
- 寻找局部最大数值
- Quadratic interpolation 平方插值

①计算积分图。

②获取图像的二阶微分。

③压缩非极大值。

④寻找局部极大值。

⑤使用平方插值法。

7.SURF——特征描述子

Interest point *descriptor*:特征提取

- Divide window into 4x4 (16 subwindows)
- 分成16个子窗口
- Compute Haar wavelet outputs Harr小波特征
- W

$$v_{subregion} = \left[\sum dx, \sum dy, \sum |dx|, \sum |dy|\right]$$

- 一阶导数（x,y） 一阶导数绝对值（x,y）
- This yields a 64-element descriptor

(Only implement USURF – no rotation)

①把图像分成16个子窗口。

②计算小波特征：X，Y方向一阶导数及其绝对值。

③获取由64个元素组成的描述子。

专题十九 图像的拼接

1.图像拼接步骤：

①找到两张图像上的特征点。

②匹配特征点。

③利用特征点，进行图像对齐。

2.流程：

①特征点检测。

②建立SIFT描述子。

③匹配SIFT描述子。

④拟合几何变换矩阵。找N对对应的特征点，解方程。

⑤RANSAC排除离群点，只考虑非离群点。

⑥图像渲染。

专题十九 深度学习

1.深度学习过程：

参数随机初始化->计算误差->更新权重。（学习率太大，振荡幅度大，无法收敛；学习率太小，可能收敛到局部极大值，并且收敛速度慢）

2.反向传播算法

①参数随机初始化。

②计算误差E.

③利用反向传播算法，通过误差对权重系数求导×学习率，计算每个神经元应该承担的损失。

- Back-propagation (BP)
  - Step 1: Randomly initialize weight, compute $h_W(X)$;
  - Step 2: Compute error $\mathrm{E} = (h_W(X) - y)^2$
  - Step 3: $W_k = W_{k-1} - \epsilon \frac{\partial E}{\partial W}$, specifically, $w_{ij}^{(k)} = w_{ij}^{(k-1)} - \epsilon \frac{\partial E}{\partial w_{ij}^{(k-1)}}$

3.卷积神经网络：与其他类型的神经网络相比，大幅降低参数数量。

4.池化的意义：通过池化，降低卷积层输出特征的维度，不仅能有效减少网络参数，还能防止过拟合。

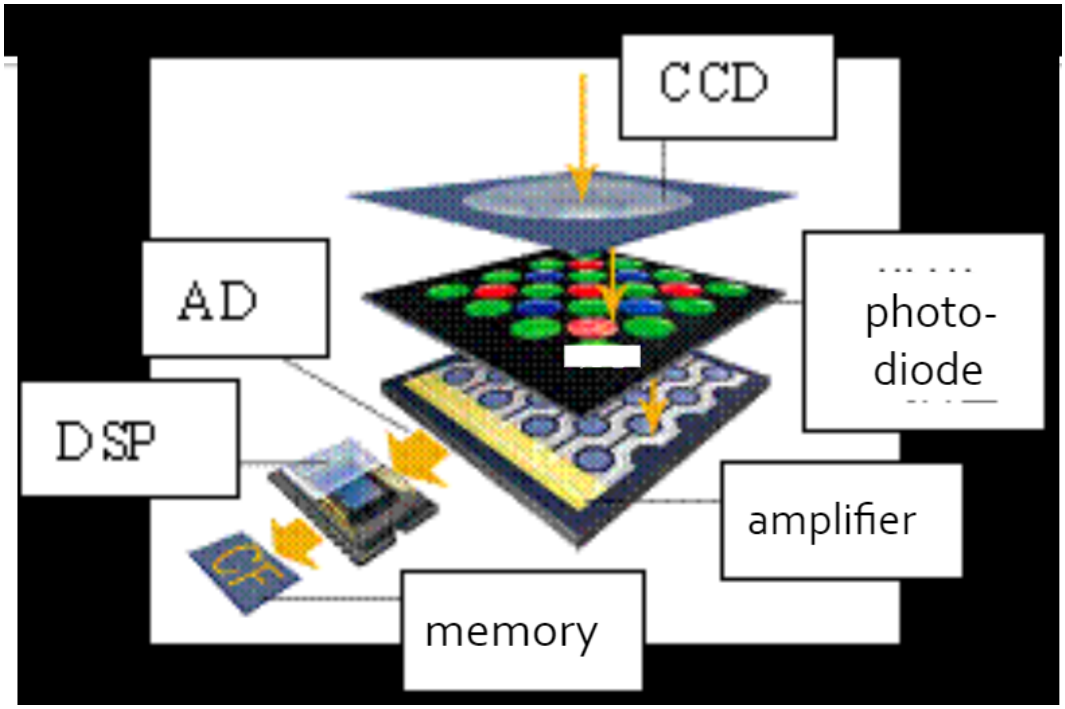
5.梯度下降法、早停法

6.池化：对不同特征进行聚合统计处理，以获得相对更低的维度。

专题二十 成像原理

1.数字相机成像原理

- （1） When taking photograph, the light from the scene goes through the lens and reaches CCD.
- （2） When CCD being exposed, the photodiode is stimulated to release charges, and produces electrical signals.
- （3） CCD controlling chip controls the electric flow through the controlling signal circuit in photoperceptive component. CCD will collect such electrical signals and output them to an amplifier.
- （4） After amplifying and filtering, the electrical signals reach ADC. And ADC transfer such electrical signals (continuous) to digital ones (discrete). The value of digital signal is proportional to the intensity of electrical signal and voltage. These values are corresponding to those of an image.
- （5） But the above data cannot be treated as image directly. They will be further processed by DSP (digital signal processing). In DSP, color correction and white balance will be performed to obtain a qualified image, and the image will be encoded into the supported format and resolution, which can be stored as a image file.
- （6） After the above steps, the image file appears on your memory card and can be previewed.



①照相时，光通过透镜，到达CCD。

②CCD曝光时，光电二极管被激发，发射电荷，产生电信号。

③CCD控制芯片通过感光元件中的控制信号电路控制电流，收集电信号，把它们输出给放大器。

④电信号经过放大、滤波，到达ADC。ADC把连续电信号转换成离散的数字信号。数字信号的数值与电信号的强度、电压成正比。这些数值对应于图像。

⑤上述数值还不能直接成为图像，而是要经过DSP（Digital Signal Processing数字信号处理）。在DSP中，执行白平衡、色彩校正，获得标准图像，图像将会被编码为支持的格式和分辨率，存储为图像文件。

⑥经过以上步骤，图像存储于存储卡中，可以预览。

2.弥散圆

①定义：在焦点前后，光线开始聚集、扩散，点的影像变模糊，形成一个扩大的圆。

②容许弥散圆：不能辨认为模糊的弥散圆，称为容许弥散圆。焦点前后最远的容许弥散圆之间的距离，叫做景深。

3.光圈对成像的影响

①光圈孔径过大，成像模糊。

②光圈孔径过小：可通过的光线少，光强不足；孔径足够小的时候，光会发生衍射现象，影响成像质量。

4.景深：焦点前后，容许弥散圆之间的距离。

①光圈越大，景深越小。

②焦距越长，景深越小。

③拍摄距离越远，景深越大。

（记忆：景深只与拍摄距离成正比，跟焦距、光圈都成反比）

【第二部分】4年真题归类及试答

专题一 成像原理

1.简述光圈（Iris）与景深（Aperture）的关系。

答：光圈越大，景深越小。

2.简述相机光圈大小对成像的影响。

答：①光圈孔径过大，成像模糊。

②光圈孔径过小：可通过的光线少，光强不足；孔径足够小的时候，光会发生衍射现象，影响成像质量。

3.介绍数码相机生成图片的过程

答：①照相时，光通过透镜，到达CCD。

②CCD曝光时，光电二极管被激发，发射电荷，产生电信号。

③CCD控制芯片通过感光元件中的控制信号电路控制电流，收集电信号，把它们输出给放大器。

④电信号经过放大、滤波，到达ADC。ADC把连续电信号转换成离散的数字信号。数字信号的数值与电信号的强度、电压成正比。这些数值对应于图像。

⑤上述数值还不能直接成为图像，而是要经过DSP（Digital Signal Processing数字信号处理）。在DSP中，执行白平衡、色彩校正，获得标准图像，图像将会被编码为支持的格式和分辨率，存储为图像文件。

⑥经过以上步骤，图像存储于存储卡中，可以预览。

4.哪些因素会对景深有影响？如何影响？

答：景深是指焦点前后，容许弥散圆之间的距离，受以下因素的影响。

①光圈越大，景深越小。

②焦距越长，景深越小。

③拍摄距离越远，景深越大。

5.举出3种成像方式。

答：数字成像、射线成像、胶片成像。

专题二 颜色空间

1.解释消色和彩色的含义。

答：彩色：红黄蓝等单色，以及它们的混合色。对各个波长的光的反射具有选择性，所以在白光照射下呈现不同的颜色。

消色：白色、黑色，以及各种深浅不同的颜色。对各个波长的光的反射没有选择性。

2.【信息不完整】将人眼对色彩的灵敏度进行排序。

答：【网络信息】人眼对绿光最为敏感，对红、蓝光敏感度较低。

3.假如RGB的颜色空间是一个正立方体，那么(0,0,0)到(1,1,1)这条对角线表示什么?

答：表示消色。（0,0,0）点为黑色，（1,1,1）点为白色，其余各点表示亮度不同的灰色。RGB三个颜色分量的强度相同时，混合成消色，消色对各个波长的光的反射没有选择性。

4.列举设备无关的颜色空间、设备相关的颜色空间。

答：RGB、CMY都是设备相关的颜色空间，LAB是设备无关的颜色空间。

5.（1）HSV各个字母的意思。

（2）画简图表示HSV。

（3）HSV比RGB好在哪里。

答：（1）分别是Hue,Saturation,Value，表示色调、饱和度、亮度。

（2）倒椎体形状。其中，旋转角表示色调，由Z轴发出的半径表示饱和度，Z坐标表示亮度。

（3）

①符合人眼感知颜色的先后顺序：色调、对比度、亮度。

②人眼对亮度最敏感。所以HSV颜色空间，专门设置一个亮度V通道，调整图像亮度时，只需调整该通道即可。而对于RGB颜色空间，调整亮度需要同时对RGB三个通道进行调整，比较麻烦。

6.用于表示加色和减色的颜色空间是什么，解释这两个空间每一维的含义，并说明它们的应用场合。

答：加色：RGB（Red,Green,Blue）。可以应用于LED显示。

减色：CMY（Cyan青色，Magenta洋红，Yellow黄色）。可以用于颜色打印。

7. 写出至少五种颜色空间，并解释两种颜色空间每一维的含义。

答：（1）颜色空间：YUV,RGB,CMY,CIE,LAB等。

(2)RGB（Red,Green,Blue）

CMY（Cyan青色，Magenta洋红，Yellow黄色）

专题三 图像存储

1.给定图像，写出行程编码，并说明行程编码的原理。

答：行程编码：高压缩率图像，每一行都由子列表描述。列表的第一个元素是行号。第2-3,4-5，…个元素分别是一串像素的起始和结束点的列标。

比如11149，表示：行号是第1行，有一串像素从1开始，到1截止。有一串像素，从4开始，到9截止。

2.JPEG图像的压缩策略是什么？有什么优点？

答：（1）①采用行程编码，压缩率高。

②根据压缩比的要求，从高频到低频逐步削减信息。

（2）①高频信息占用的存储空间比较大，减少高频信息，可以获得高压缩比。

②保留低频信息，可以保留物体的基本轮廓、色彩分布，最大程度维持图像原貌。

3.计算机里灰度用几位表示，一共几个灰度梯度。

答：8位，256个。

4.BMP图像的图像数据（bitmap data）部分中，每一行的数据存储有什么要求？如果想要存储12 34 56 78 9A BC（具体数据不是这个，忘了，不过没差- -），应该怎样存储？（我不知道描述清楚了没有- -应该问的是四字节对齐的事情...）

答：四字节对齐。推测：考虑大小端、考虑32位BMP图像中的保留字。

5.给出一张二值图像，求行进间编码，并简述你是怎么求出来的。

答：略。

专题四 空间域滤波

1.(1)给定掩膜大小和灰度图（灰度值均已知），求中值滤波后的图像。 (2)擦去灰度图中的某个像素点，对这个点进行双线性插值。

答：略。

2.(1)简述双边滤波的基本思想

(2)简述导向滤波与双边滤波相比的优点，并说说导向滤波如何避免了梯度反转

答：（1）双边滤波同时考虑在两个域上进行作用，在两个域上分别寻找一个高斯函数，然后作乘积，作为像素点的权重，然后进行加权平均处理。

（2）优点：①保边滤波，不会发生梯度逆转。

②非迭代。

③O（1）时间复杂度，快速准确。

④比平滑滤波器更通用。

避免梯度反转：①双边滤波保边，保的是差值的绝对值，所以会产生梯度逆转。

②引导滤波考虑了保梯度的方向，不会造成梯度逆转。

3.给了6*6的矩阵，用3*3的窗口算中值与拉普拉斯滤波（计算量巨大）

答：略。

4.双边滤波的基本思想与各个项的解释。

答：（1）基本思想：双边滤波同时考虑在两个域上进行作用，在两个域上分别寻找一个高斯函数，然后作乘积，作为像素点的权重，然后进行加权平均处理。

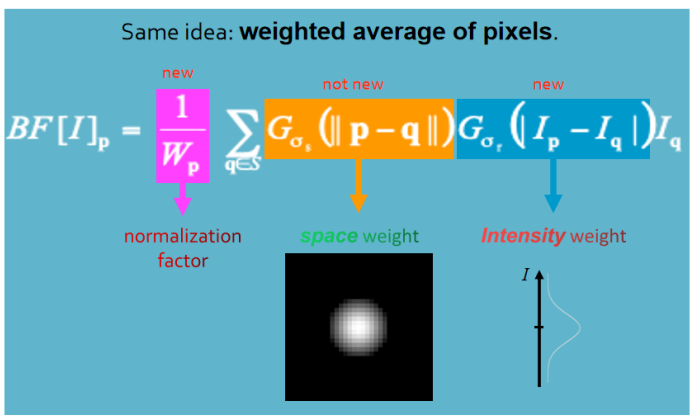
（2）各个像的解释：

①针对空间域、亮度域的高斯函数

②正则化参数。

③窗口内像素点Q与中心像素点P之间的距离。

④空间域、亮度域高斯函数对应的参数。



5.双线性插值的过程。

答：略。（第一部分有，记熟了，不想写了。）

6. (1)简述如何利用双线性插值求出某个点的灰度；

(2)给一张5×5的灰度图，求用3×3的mask进行中值滤波后的结果，忽略图像边界。

答：略。

专题五 直方图

1.给图，写直方图匹配（拟合）过程。

答：统计S、P,形成函数图像->函数图像转换->逆变换，得到新的像素值。

2.直方图均衡化的结果分布不均匀的原因

答：在算法的第二步中，Sk并不是完全等价于n个灰度级之一，而是被归入离它最近的灰度级。所以多个相邻的Sk可能落入同一个新的灰度级中。因此在离散直方图均衡化操作后，得到的不同灰度级对应的概率可能不一致。

3.直方图均衡化计算。

答：略。

4.填写一张进行直方图均衡化的表格（给了每个rank对应的像素点数，求每个rank的一些中间参数值和均衡化后对应的像素点数）

答：略。

专题六 形态学操作

1.(1)给定结构元和二值图像，画出膨胀操作的结果并简述膨胀的物理意义。

(2)画出腐蚀操作的结果，并讨论开操作和闭操作的结果是否一致。

答：（1）膨胀的意义：弥补图像中的空隙。

（2）可能不一致。举例：比如两个图形之间距离很近，先腐蚀再膨胀连不到一起，先膨胀再腐蚀可以连到一起。

2. (1)给定结构元和图像，画出腐蚀操作的结果。并说说腐蚀操作的意义。

(2)在(1)的基础上，画出膨胀操作的结果。

答：腐蚀的意义：擦除边界和小型噪声。

3.开运算和物理意义

答：去除小型噪声，在较细的部分切割物体，边界平滑同时保留原有面积。

4.给出一个结构元和一个二值图像，问这个结构元和二值图像进行膨胀操作后的结果是什么，并简述膨胀操作的物理意义。

答：弥补图像中的空隙。

【注】形态学操作的物理意义似乎是必考点。总结如下：

①膨胀：只要窗口中有1，目标像素就是1。作用：弥补图像中的空隙。

②腐蚀：窗口中全部都是1，目标像素才是1。作用：去除边界、小型噪声。

③腐蚀应用：边界提取。先得到腐蚀图像，原始图像减去腐蚀图像就是图像边界。

④膨胀应用：弥补空洞。先对图像进行膨胀，再跟A的补图像进行交运算，不断迭代直至图像不变。

⑤开操作：先腐蚀，再膨胀。作用：去除小型噪声，在较细的部分切割物体，边界平滑同时保留原有面积。

⑥闭操作：先膨胀，再腐蚀。作用：弥补小型空洞，连接邻近物体，在平滑边界的同时最大限度地保留面积。

专题八 视觉增强

1.写出对数增强的过程，解释对数变换如何实现视觉增强。

答：

（1）注意：亮度值已经归一化，取值范围为[0,1].

①公式：

$$L_d = \frac{\log(L_w+1)}{\log(L_{max}+1)}$$

②最大亮度为1，其余亮度数值平滑变化。

③不出现梯度逆转，即原来相对较暗者，经过对数变换后仍然相对较暗。

（2）对数函数曲线越靠近0，图像越陡，越往右图像越平缓。那么亮度低的部分，经过对数变换后亮度范围增大，对比度变大，实现了视觉增强。

专题九 大津二值化算法

1.如何生成高质量二值化图片

答：设置小窗口，对于每个小窗口进行局部大津二值化。具体算法流程：

①寻找图像中的最小、最大像素值。

②从最小像素值开始，每次+1，直到最大像素值，分别以此为阈值。

③在该阈值下，确定前景、背景。

④计算类内方差、类间方差。

⑤返回步骤②，直到阈值达到了图像中的最大像素值。

⑥最大类间方差、最小类内方差对应的阈值，即为最终所求阈值。

2.推导大津算法，并简述大津算法基本步骤。

答：（1）推导：主要是针对类内方差、类间方差进行推导。

（2）①寻找图像中的最小、最大像素值。

②从最小像素值开始，每次+1，直到最大像素值，分别以此为阈值。

③在该阈值下，确定前景、背景。

④计算类内方差、类间方差。

⑤返回步骤②，直到阈值达到了图像中的最大像素值。

⑥最大类间方差、最小类内方差对应的阈值，即为最终所求阈值。

3.(1)简述把灰度图像二值化的步骤。

（2）如果一张灰度图像的光照变化比较大，应该如何二值化?

答：（1）①寻找图像中的最小、最大像素值。

②从最小像素值开始，每次+1，直到最大像素值，分别以此为阈值。

③在该阈值下，确定前景、背景。

④计算类内方差、类间方差。

⑤返回步骤②，直到阈值达到了图像中的最大像素值。

⑥最大类间方差、最小类内方差对应的阈值，即为最终所求阈值。

（2）设置合理大小的小窗口，在每个小窗口内进行局部大津二值化处理。

专题十 特征检测

1.简述SURF寻找角点的步骤，并说说与Harris角点探测相比有哪些优点。

答：（1）SURF寻找角点的步骤：

①计算积分图。

②获取图像的二阶微分。

③压缩非极大值。

④寻找局部极大值。

⑤使用平方插值法。

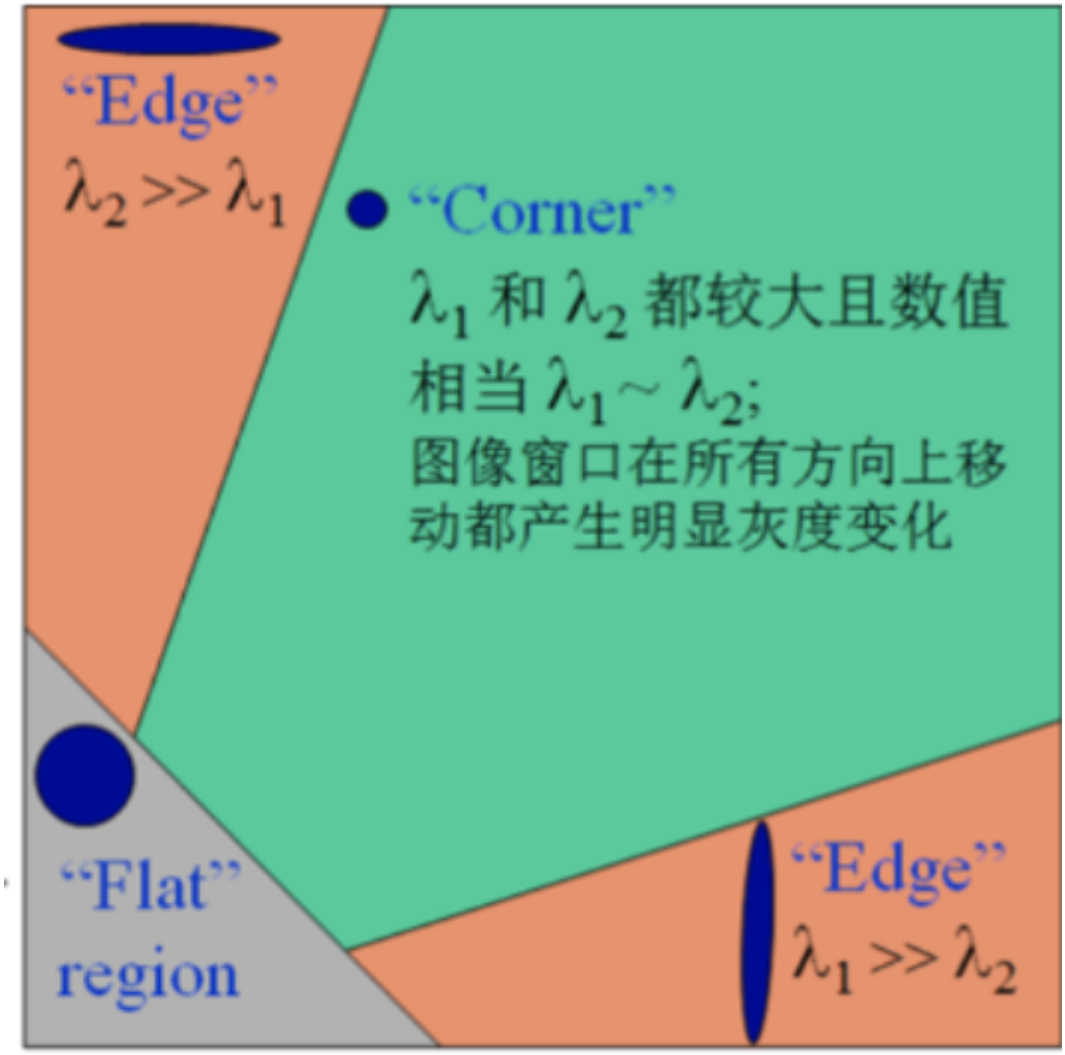
（2）优点：SURF(加速鲁棒特征)具有快速性,同时具有尺度不变性,对光照变化和仿射、透视变化也具有较强的鲁棒性。

Harris算子没有尺度不变性。

2. Harris角点检测公式推导以及如何判断。

（1）主要涉及Tailor展开进行一阶近似，由原始求和式子推导为矩阵乘积形式。

（2）两个特征值都比较大，而且数值上相当，即为角点。



专题十一 卷积神经网络

1.给图，写BP全过程。

①参数随机初始化。

②计算误差E.

③利用反向传播算法，通过误差对权重系数求导×学习率，计算每个神经元应该承担的损失。

2. CNN中的池化是什么？有什么作用?

答：（1）池化：对不同特征进行聚合统计处理，以获得相对较低的维度。

（2）作用：通过池化，降低卷积层输出特征的维度，不仅能有效减少网络参数，还能防止过拟合。

3.(1)简述反向传播训练CNN的步骤

(2)解释为什么CNN可以大规模减少权重参数训练量

(3)简述CNN中进行池化的意义

答：

（1）①参数随机初始化。

②计算误差E.

③利用反向传播算法，通过误差对权重系数求导×学习率，计算每个神经元应该承担的损失。

（2）①局部连接而非全连接。

②权值共享。

③池化处理。

（3）通过池化，降低卷积层输出特征的维度，不仅能有效减少网络参数，还能防止过拟合。

专题十二 Morph

1.给trump到obama的morph图，问如何morph。

答：①关键帧选取：在大小、结构上相似，才能获得自然的morph结果。

②结构获取：利用特征点定义结构，对特征点进行匹配。

③定义特征点->对两幅图像的特征点进行匹配->对图像进行分割->运用插值方法

④Morph 的具体方法：根据变换快慢，设置步长，将图A中每一点的RGB数值逐渐变成与图B中相同位置像素的RGB。可以选择：等比、等差等方法。

2.简述用ratio image将source人脸的表情转移到target人脸上的步骤。

答：①关键帧选取：在大小、结构上相似，才能获得自然的morph结果。

②结构获取：利用特征点定义结构，对特征点进行匹配。

③定义特征点->对两幅图像的特征点进行匹配->对图像进行分割->运用插值方法

④Morph 的具体方法：根据变换快慢，设置步长，将图A中每一点的RGB数值逐渐变成与图B中相同位置像素的RGB。可以选择：等比、等差等方法。

专题十三 图像拼接

1.给你PPT上的那两张图，问如何拼接图像。

答：①特征点检测。

②建立SIFT描述子。

③匹配SIFT描述子。

④拟合几何变换矩阵。找N对对应的特征点，解方程。

⑤RANSAC排除离群点，只考虑非离群点。

⑥图像渲染。

专题十四 数字图像与信息

1.为什么说数字图像是信息的主要展示方式?

答：在人对外部世界的感知中，视觉占比至少70%。

2.图像与图形在来源/应用方面的不同。

答：（1）来源：图形是指由外部轮廓线条构成的矢量图，即由计算机绘制的直线、圆、矩形、曲线、图表等。而图像是由扫描仪、摄像机等输入设备捕捉实际的画面产生的数字图像，是由像素点阵构成的位图。

（2）应用：图形描述的是轮廓不复杂，色彩不是很丰富的对象，如几何图形、工程图纸、CAD、3D造型软件等。图像表现含有大量细节的对象，如照片、绘图等，通过图像软件可进行复杂图像的处理以得到更清晰的图像或产生特殊效果。

【附录】2022-2023图像信息处理期末回忆卷

一、填空题

1.列举3种成像方式：光成像、__、__、__。

2.灰度图像用__个字节存储，有__个灰度级；彩色图像用__个字节存储，能表示__种颜色。

3.图像存储格式分为有压缩、无压缩，有压缩又分为__和__。其中，JPEG属于__。

4.给定一个3×3图像（中心像素缺失），分别求均值、中值、标准拉普拉斯滤波的结果。

二、简答题（顺序不一定对）

1.视觉成像系统感光的优先级和灵敏度。

2. （1）光圈大小对成像的影响。

（2）光圈和景深的关系，以及为什么是这样。

3.

4.简述双边滤波、引导滤波的原理和不同之处。

5.简述RANSAC算法。

6.（1）给定比率图，简述用比率图对目标图像进行处理的步骤。

（2）与Warping相比，该方法有哪些好处。

三、计算题

1.给定结构元和图像，计算膨胀、腐蚀、开操作、闭操作所得结果，可以用矩阵表示。

2.双线性插值法计算。

3.直方图均衡化计算。

四、证明题

1.设r,s分别是直方图均衡化前后的灰度级。证明直方图均衡化变换后，s = P(r)从0到r的积分。

2.推导Harris角点检测公式，并说明特征值与角点的关系。