

华中科技大学

研究生课程考试答题本

考生姓名 单凯

考生学号 M201772881

系、年级 武汉光电国家实验室 2017 级

类 别 非定向

考试科目 流媒体技术

考试日期 2017 年 12 月 18 日

评 分

题 号	得 分	题 号	得 分
1			
2			
3			
4			
5			

总 分：	评 卷 人：
------	--------

注：1、无评卷人签名试卷无效。

2、必须用钢笔或圆珠笔阅卷，使用红色。用铅笔阅卷无效。

2017年秋季《流媒体技术》试题

此页必须单面打印, 装订在封一封二之后. 其它页面必须双面打印.

- 请写出 JPEG 压缩编码的量化矩阵. 针对不同频率分量的 DCT 系数, 分析不同的量化策略.
 - 在提供的图像中任选一幅, 量化参数为 0.5, 请给出 Y 分量经过 DCT 变换、量化以及反变换后的结果. (要求贴原图, 保持版面美观整洁)
 - 对原图的 RGB 分量分别做 DCT 变换、量化、反量化、反变换, 量化参数选 1、2、4、8, 计算平均每个像素的 bit 数与 PSNR. (用电子表格列出)
 - 对原图的 YUV 分量分别做 DCT 变换、量化、反量化、反变换, 量化参数选 1、2、4、8, 计算平均每个像素的 bit 数与 PSNR. (用电子表格列出)
 - 在量化参数 1、2、4、8 下分别计算 RGB 分量三者的平均 bit 数与 PSNR, 并画出一条 Rate Distortion 曲线. 在量化参数 1、2、4、8 下分别计算 YUV 分量三者的平均 bit 数与 PSNR, 并画出一条 Rate Distortion 曲线. 通过对 RD 曲线的分析, 说明为什么在压缩编码时需要把 RGB 格式的图像转换到 YUV 格式.
 - 请借助于图像的显著性分析, 找出图像中的感兴趣区域, 并贴出原图与检测结果. 设计一种有差别的量化方案, 在量化参数选 4 时, 分析 Y 分量在有差别量化与无差别量化时的 RD 性能与主观质量的差别.
- 请画出任意一种尺度函数与小波函数. (Matlab 画图, 存为 EMF 图像格式再贴入答题纸)
 - 请通过程序分析, 写出画以上波形的基本算法与流程.
- 在提供的 3 层径向基函数网络基础上, 实现 4 层的学习, 并贴出增加层的反向传播代码.
 - 贴出 1 万次迭代后的学习误差图 (Matlab 画图, 存为 EMF 图像格式再贴入答题纸)?
 - 如果把径向基函数网络改成深度卷积网, 请问需要对径向基函数网络做哪些改变, 从输入数据的格式、核函数的形式、核函数与特征图之间的操作、各层 δ 在反向传播时候与核函数间的操作等方面展开.
 - 以提供的软件 DemoCNN 为平台, 在 RGB 与 Lab 颜色空间下预测某幅图像的关注区域(贴原图与检测结果图), 并进一步说明哪种颜色空间更适合于解决视频内容的优先传输问题.
- 根据提供的论文, 简述 BPS 显著性检测的基本原理. 并回答一般有哪些通用的测试数据集?
 - 任选一幅提供的图像, 用嵌入在 DemoSeamCarving 中的显著性检测算法做图像显著性检测, 并贴出原图与显著性检测的结果. 以 Ground truth 为依据, 画出检测结果的 ROC 曲线.
- 请回答视频在 YUV 颜色空间中的 4:2:0 与 4:2:2 格式是什么意思?
 - 请回答分辨率格式为 CIF 的视频画面的边长, 并分析它们为什么会是 16 的倍数.
 - 请用 JM 参考软件, 任选一个 CIF 大小的 4:2:0 的 YUV 格式的视频序列, 在量化参数为 20、24、28、32 时, 计算在各个量化参数下的 PSNR 与 Bite rate, 并画出 RD 曲线?
 - 请分析在不同量化参数下的编码的速度. 并用率失真公式 $J = D + \lambda R$ 解释量化参数与编码速度之间的关系.

题号	回 答 内 容	得分
----	---------	----

1(1) 请写出 JPEG 压缩编码的量化矩阵. 针对不同频率分量的 DCT 系数, 分析不同的量化策略.

JPEG 标准提供的两种量化矩阵作为参考, 一种为低压缩, 另一个为高压缩。

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

左图为低压缩, 右图为高压缩。

系数量化是一个十分重要的过程, 是造成 DCT 编解码信息损失 (或失真) 的根源。量化的作用是在一定的主观保真度图像质量的前提下, 丢掉那些对视觉影响不大的信息, 以获得较高的压缩比。由于 DCT 系数包含了空间频率信息, 可充分利用人眼对不同频率敏感程度不同这一特性来选择量化表中的元素值的大小, 对视觉重要的系数采用细量化 (量化步长较小), 如低频系数被细量化, 对高频系数采用粗量化 (量化步长较大)

量化的目的是减小非“0”系数的幅度以及增加“0”之系数的数目, 一般情况下都使用均匀量化器进行量化, 量化步长是按照系数所在的位置和每种颜色分量的色调值来确定, 因为人眼对亮度信号比对色差信号更敏感, 因此使用了两种量化表, 一个是亮度量化表, 一个是色度量化表。此外, 由于人眼对低频分量的图像比对高频分量的图像更敏感, 因此量化表中的左上角的量化步长一般要比右下角的量化步长小, 量化机要使得大部分数据得以压缩, 同时又要保证通过量化和编码之后能输出一个与信道传输速率匹配的比特流。

如果 DCT 系数舍去 (归 0) 数量越多, 则图像的压缩率越高, 但是图像的质量越差。图像质量跟舍去的数量呈单调关系, 即舍去的点越多, 图像质量越差。

1(2) 在提供的图像中任选一幅, 量化参数为 0.5, 请给出 Y 分量经过 DCT 变换、量化以及反变换后的结果。(要求贴原图, 保持版面美观整洁)

原图:



Y 分量:

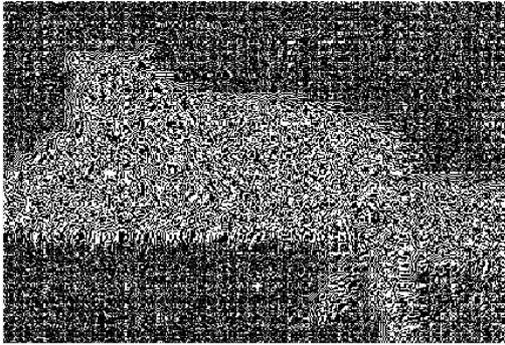


题号

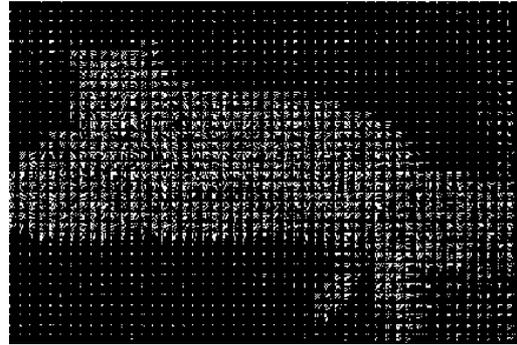
回 答 内 容

得分

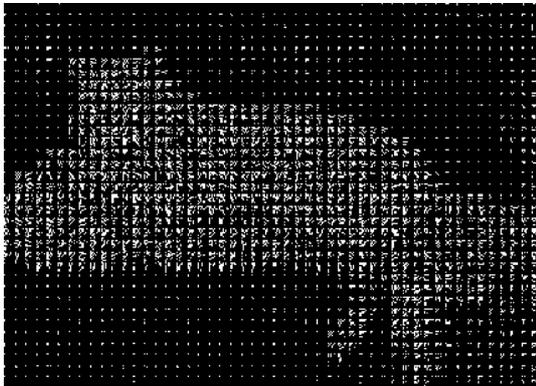
DCT 变换:



量化:



反量化:



逆 DCT 变换:



1(3)

对原图的 RGB 分量分别做 DCT 变换、量化、反量化、反变换, 量化参数选 1、2、4、8, 计算平均每个像素的 bit 数与 PSNR. (用电子表格列出)

颜色分量	R				G			
量化参数	1	2	4	8	1	2	4	8
bit/pixel	1.14034	0.820121	0.561692	0.371439	1.12864	0.804483	0.54404	0.354963
PSNR	33.9559	31.5728	29.1053	26.3768	33.9542	31.5426	29.0887	26.3353

颜色分量	B			
量化参数	1	2	4	8
bit/pixel	1.08848	0.755351	0.492716	0.301854
PSNR	33.9401	31.5468	29.1509	26.4619

1(4)

对原图的 YUV 分量分别做 DCT 变换、量化、反量化、反变换, 量化参数选 1、2、4、8, 计算平均每个像素的 bit 数与 PSNR. (用电子表格列出)

颜色分量	Y				U			
量化参数	1	2	4	8	1	2	4	8
bit/pixel	1.1268	0.804356	0.544249	0.355267	0.317392	0.223194	0.161935	0.133549
PSNR	33.963	31.5584	29.0889	26.3383	47.3799	43.0724	38.9104	35.0151

颜色分量	V			
量化参数	1	2	4	8
bit/pixel	0.339853	0.234774	0.174382	0.138665
PSNR	46.7588	43.1235	38.982	35.1507

1(5) 在量化参数 1、2、4、8 下分别计算 RGB 分量三者的平均 bit 数与 PSNR，并画出一条 Rate Distortion 曲线。在量化参数 1、2、4、8 下分别计算 YUV 分量三者的平均 bit 数与 PSNR，并画出一条 Rate Distortion 曲线。通过对 RD 曲线的分析，说明为什么在压缩编码时需要把 RGB 格式的图像转换到 YUV 格式。

答：

表格 1 RGB 和 YUV 在不同参数下的均值

	RGB 的均值				YUV 的均值			
量化参数	1	2	4	8	1	2	4	8
bit/pixel	1.119	0.793	0.533	0.343	0.595	0.421	0.294	0.209
PSNR	33.950	31.554	29.115	26.391	42.701	39.251	35.660	32.168

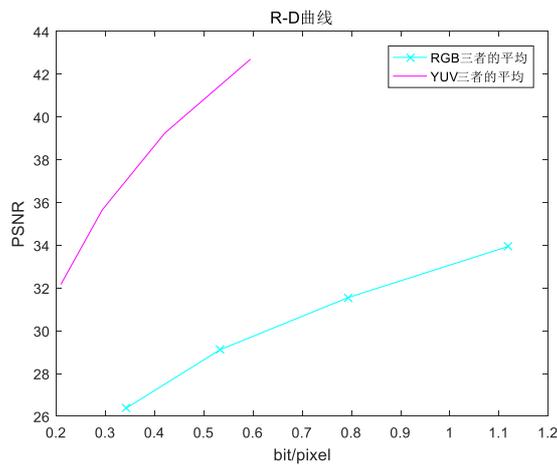


图 1 RGB 和 YUV 的比较

Y 分量经 DCT 变换重构后的图像比 U 分量经 DCT 变换重构的图要更加清晰；分析上图中 RD 曲线，Y 分量含的信息量要大一些。人眼对于亮度信号非常敏感，而对色差信号的敏感度相对较弱，也就是说，图像的主要信息包含在 Y 分量中。

图像压缩要利用人对图像的感觉的生理特性，人对亮度很敏感，对色彩信息不太敏感，对色差信号 U 和 V 可以采用大面积着色的方法，使用比亮度信号 Y 低的采样频率也可使人感觉不到图像的失真，可以把色度的分量减少，然后进行其他压缩算法， YUV 格式把图像用色度和亮度分开，可以很容易去掉色度的数据。 YUV 格式的图像能量少，压缩后再量化时失真损失少，可以更好保护图像的视觉效果。

另，附平均之前的 RD 曲线：

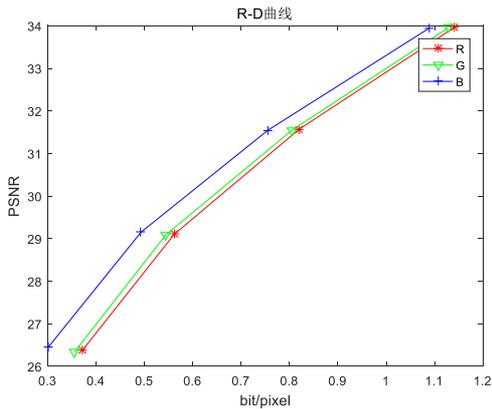


图 2 RGB 分量下的率失真曲线

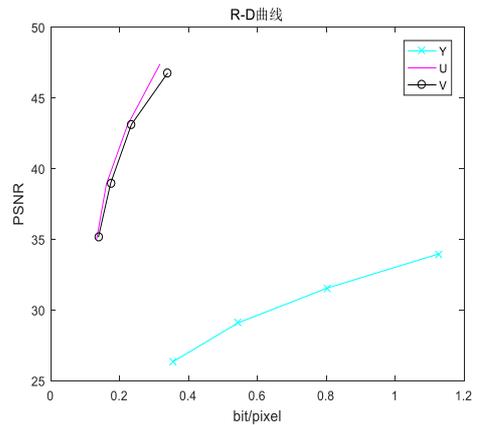


图 3 YUV 分量下的率失真曲线

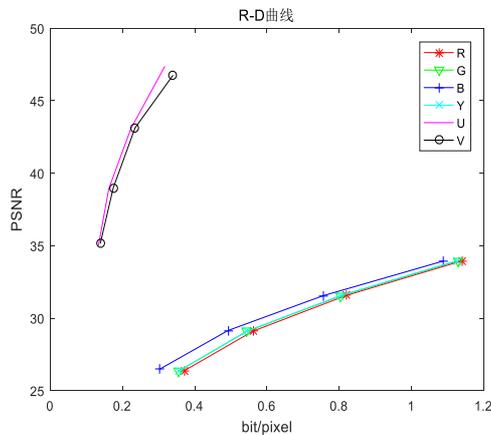


图 4 RGB 和 YUV 分量下的率失真曲线

1(6) 请借助于图像的显著性分析，找出图像中的感兴趣区域，并贴出原图与检测结果。设计一种有差别的量化方案，在量化参数选 4 时，分析 Y 分量在有差别量化与无差别量化时的 RD 性能与主观质量的差别。

答：原图同第一问中的图。

题号	回 答 内 容	得分												
	<div data-bbox="473 192 935 480" data-label="Image"> </div> <div data-bbox="548 486 856 519" data-label="Caption"> <p>图 5 检测出的感兴趣区域</p> </div> <div data-bbox="140 525 287 558" data-label="Text"> <p>量化方案:</p> </div> <div data-bbox="140 564 1267 901" data-label="Table"> <table border="1"> <thead> <tr> <th>量化方案</th> <th>无差别量化</th> <th>有差别量化</th> </tr> </thead> <tbody> <tr> <td>具体方法:</td> <td>对图像的 R 通道,量化系数为 4,量化时不计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。</td> <td>对图像的 R 通道,量化系数为 4,量化时计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。</td> </tr> <tr> <td>bit/pixel</td> <td>0.561692</td> <td>0.595532</td> </tr> <tr> <td>PSNR (DB)</td> <td>29.1053</td> <td>29.3176</td> </tr> </tbody> </table> </div> <div data-bbox="140 905 1267 1029" data-label="Text"> <p>由上表可以看出:有差别量化每像素占用的 bit 数会增多一点(6.02%),但是峰值信噪比会提高 0.73%,也就是说,图像经过压缩之后信息损失会减少,但是压缩后占用的空间会多一点。</p> </div> <div data-bbox="140 1038 672 1348" data-label="Image"> </div> <div data-bbox="193 1362 641 1395" data-label="Caption"> <p>图 6 采用无差别量化压缩还原后的图</p> </div> <div data-bbox="735 1038 1225 1348" data-label="Image"> </div> <div data-bbox="781 1362 1234 1395" data-label="Caption"> <p>图 7 采用有差别量化压缩还原后的图</p> </div> <div data-bbox="140 1401 1267 1564" data-label="Text"> <p>主观上,仔细观察两幅图,应该是有差别量化的豹子(显著性区域)相较于无差别量化的要清晰些,背景部分要比无差别量化压缩的图片的背景要模糊,但是如不仔细看,看不出差别来。但是经过显著性检测的方法处理时间明显要久些,毕竟多了一个显著性检测的步骤。</p> </div> <div data-bbox="49 1587 1267 1695" data-label="Text"> <p>2(1) 请画出任意一种尺度函数与小波函数。(Matlab 画图,存为 EMF 图像格式再贴入答题纸)</p> </div> <div data-bbox="140 1705 721 1744" data-label="Text"> <p>答:在此画 Haar 的尺度函数和小波函数。</p> </div>	量化方案	无差别量化	有差别量化	具体方法:	对图像的 R 通道,量化系数为 4,量化时不计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。	对图像的 R 通道,量化系数为 4,量化时计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。	bit/pixel	0.561692	0.595532	PSNR (DB)	29.1053	29.3176	
量化方案	无差别量化	有差别量化												
具体方法:	对图像的 R 通道,量化系数为 4,量化时不计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。	对图像的 R 通道,量化系数为 4,量化时计算显著性,进行 DCT 变换,量化,反量化,逆 DCT 变换。												
bit/pixel	0.561692	0.595532												
PSNR (DB)	29.1053	29.3176												

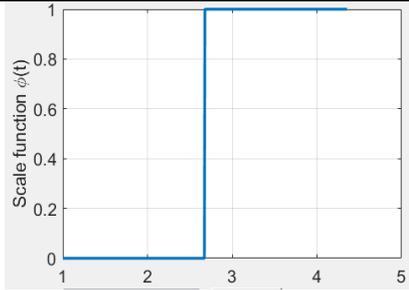


图 8 尺度函数

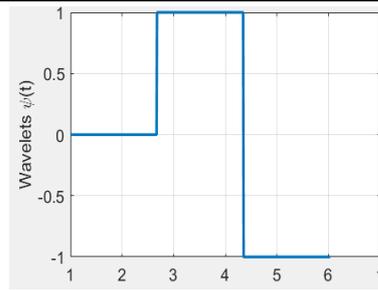


图 9 小波函数

2(2)

请通过程序分析，写出画以上波形的的基本算法与流程。

初始条件为知道两组滤波器 h_n 与 g_n ，需要得到正交小波 $\psi(t)$ 。我们可以通过构造尺度函数 $\phi(t)$ 来得到小波函数 $\psi(t)$ ，具体公式如下：

$$\varphi(t) = \sum_k h_n \varphi(2t - n)$$

$$\psi(t) = \sum_k (-1)^k g_{1-n} \varphi(2t - n)$$

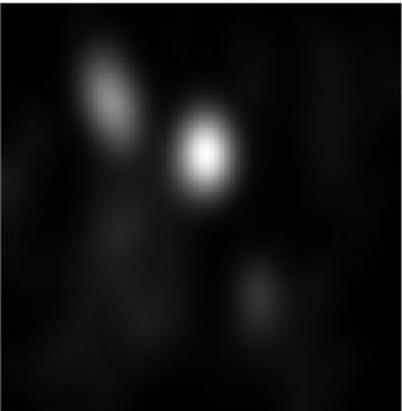
其中 g_{1-n} 表示 g_n 的倒置，即 $g_{1-n}(1)=g_n(n)$ 。

画 Haar 尺度函数和小波的算法如下：

1. $h=[1 \ 1]$; $g=[1 \ 1]$;
2. 求出 h 和 g 的长度 length_h 和 length_g ; $\text{extent}=\text{length}_g\%2$; 初始化向量 $\text{phi}=1$;
3. for $k_iteration=1:\text{iteration_times}$
4. 初始化上采样向量 $\text{phi_upsample}=\text{zeros}(1,2*\text{length}(\text{phi})+\text{extent})$;
5. phi_upsample 的偶数位填 phi 对应的值
6. $\text{phi}=\text{phi_upsample}$; $\text{length_phi}=\text{length}(\text{phi})$;
7. $\text{phi_update}=\text{zeros}(1,\text{length}_h+\text{length_phi}-1)$;
8. for $k_h=1:\text{length}_h$
9. for $k_update=1:\text{length_phi}$
10. $\text{phi_update}(k_h-1+k_update)=\text{phi_update}(k_h-1+k_update)+h(k_h)*\text{phi}(k_update)$;
11. end
12. end
13. $\text{phi}=\text{phi_update}$;
- 14.end
- 15.计算得出 length_phi
16. $\text{shift}=\text{round}(\text{length_phi}/\text{length}(g))$;
17. $\text{psi_update}=\text{zeros}(1,(\text{length}_g-1)*\text{shift}+\text{length_phi})$;
- 18.for $k_g=1:\text{length}(g)$
19. for $k_update=1:\text{length_phi}$
20. $\text{psi_update}((k_g-1)*\text{shift}+k_update)=$
 $\text{psi_update}((k_g-1)*\text{shift}+k_update)+(-1)^{(k_g+1)}*g(\text{length}(g)-k_g+1)*\text{phi}(k_update)$;
21. end
- 22.end
23. $\text{psi}=\text{psi_update}$;

题号	回 答 内 容	得分
3(1)	<p>在提供的 3 层径向基函数网络基础上, 实现 4 层的学习, 并贴出增加层的反向传播代码。</p> <p>答:</p> <p>增加层的反向传播代码:</p> <pre> % 第0层学习 dY1_dA0=[W1(1,1) W1(1,2); W1(2,1) W1(2,2)]; % W2到底在之前还是之后更新呢? dEtotal_dY1=[dEtotal_dA1(1)*dA1_dY1(1) dEtotal_dA1(2)*dA1_dY1(2)]; %这两句话可用这一句替代 dEtotal_dA1=[Delta_2*dY2_dA1(:,1) Delta_2*dY2_dA1(:,2)]; dEtotal_dA0=[dEtotal_dY1*dY1_dA0(:,1) dEtotal_dY1*dY1_dA0(:,2)]; dA0_dY0=[A0(1)*(1-A0(1)) A0(2)*(1-A0(2))]; dY0_dW0=[x(1) x(2);x(1) x(2)]; Delta_0=[dEtotal_dA0(1)*dA0_dY0(1) dEtotal_dA0(2)*dA0_dY0(2)]; dEtotal_dW0=[Delta_0(1)*dY0_dW0(1,1) Delta_0(1)*dY0_dW0(1,2);Delta_0(2)*dY0_dW0(2,1) Delta_0(2)*dY0_dW0(2,2)]; 已有两层反向传播的修改: % 第2层学习 E1=0.5*(L(1)-A2(1))^2; % Square error E2=0.5*(L(2)-A2(2))^2; % E1=L(1)*log2(A2(1)+eps); % Cross entropy % E2=L(2)*log2(A2(2)+eps); Etotal(i)=E1+E2; dEtotal_dA2=[-(L(1)-A2(1)) -(L(2)-A2(2))]; % derivative of Square error % dEtotal_dA2=1/log(2)*[L(1)/(A2(1)+eps) L(2)/(A2(2)+eps)]; % derivative of Cross entropy dA2_dY2=[A2(1)*(1-A2(1)) A2(2)*(1-A2(2))]; % 默认为Sigmoid函数 dY2_dW2=[A1(1) A1(2);A1(1) A1(2)]; Delta_2=[dEtotal_dA2(1)*dA2_dY2(1) dEtotal_dA2(2)*dA2_dY2(2)]; dEtotal_dW2=[Delta_2(1)*dY2_dW2(1,1) Delta_2(1)*dY2_dW2(1,2);Delta_2(2)*dY2_dW2(2,1) Delta_2(2)*dY2_dW2(2,2)]; % 第1层学习 dY2_dA1=[W2(1,1) W2(1,2); W2(2,1) W2(2,2)]; % W2到底在之前还是之后更新呢? dEtotal_dY2=[dEtotal_dA2(1)*dA2_dY2(1) dEtotal_dA2(2)*dA2_dY2(2)]; dEtotal_dA1=[dEtotal_dY2*dY2_dA1(:,1) dEtotal_dY2*dY2_dA1(:,2)]; dA1_dY1=[A1(1)*(1-A1(1)) A1(2)*(1-A1(2))]; dY1_dW1=[A0(1) A0(2);A0(1) A0(2)]; Delta_1=[dEtotal_dA1(1)*dA1_dY1(1) dEtotal_dA1(2)*dA1_dY1(2)]; dEtotal_dW1=[Delta_1(1)*dY1_dW1(1,1) Delta_1(1)*dY1_dW1(1,2);Delta_1(2)*dY1_dW1(2,1) Delta_1(2)*dY1_dW1(2,2)]; </pre>	
3(2)	<p>贴出 1 万次迭代后的学习误差图 (Matlab 画图, 存为 EMF 图像格式再贴入答题纸)?</p> <p>答:</p>	

题号	回 答 内 容	得分
3(3)	<div data-bbox="445 225 935 597" data-label="Figure"> </div> <div data-bbox="454 607 944 642" data-label="Caption"> <p>图 10 四层神经网络迭代 10000 次误差图</p> </div> <p>如果把径向基函数网络改成深度卷积网，请问需要对径向基函数网络做哪些改变，从输入数据的格式、核函数的形式、核函数与特征图之间的操作、各层 δ 在反向传播时候与核函数间的操作等方面展开。</p> <p>答：（1）输入数据的格式：RBF 的输入是两个向量，输入向量 x 和待预测的目标向量 y；CNN 中的输入是两幅图像（两个矩阵），也是一个输入和一个待预测的目标。</p> <p>（2）激活函数：RBF 的激活函数一般是单调的，神经网络中常用的 sigmoid 函数和双曲正切函数容易饱和，造成梯度消失；在 CNN 中采用的是另外一种激活函数：ReLU：$f(x) = \max(0, x)$。</p> <p>（3）CNN 输入数据是一个 3 通道数据，因此卷积层每个深度的节点都有 3 个不同的卷积核，分别与输入的 3 个通道进行卷积，每个通道卷积后都得到一个值，也就是 3 个通道卷积完成后，有 3 个值，再把这 3 值和偏置求和就得到节点的值。每移动卷积核都进行这样的计算。得到了卷积层每层深度的特征图后，需要将这些特征输入到激活函数 ReLU 中。</p> <p>（4）CNN 的误差传播：①对于输入数据对 (x,y)，经过 CNN 网络处理，其输出为 a。根据上面 RBF 神经网络的分析可知，层的残差满足递推关系。但是卷积神经网络在结构上和 RBF 神经网络不同，需要对 CNN 网络做一些处理，使得可以套用神经网络的公式。②当前层 l 为卷积层：卷积层的输出是由若干个卷积核与上一层输出卷积之和，并将这些和通过激活函数而生成的；假设卷积层的下一个是池化层 $l+1$。池化层输出的每一个像素都对应上一层（卷积层）的一个区域，即这两层之间不是全连接。因此需要对池化层的残差进行上采样（upsample）构造一个新的残差 map，使得残差 map 和卷积层的 map 一致，然后再根据新池化层（$l+1$ 层）的残差计算卷积层的残差；可采用后向传播来计算该权重的偏导。因为权重是共享的，因此需要对与该核有联系的所有点对进行求梯度，然后对这</p>	

题号	回 答 内 容	得分
3(4)	<p>些梯度进行求和 ③当前层为池化层：池化层的输入 maps 和输出 maps 个数相同，但是输出 maps 维数都减小了；池化层的下一层是卷积层，必须要找到对应的区域，即找到当前层残差 map 对应下一层残差 map 的像素，这样就可以利用 ANN 方法实现 δ 的递推；得到残差，就可以对偏置和权值进行偏导</p> <p>(5) BP 神经网络中，每一层都是全连接的；而卷积神经网络的卷积层每个节点与上一层某个区域通过卷积核连接，而与这块区域以外的区域无连接。CNN 的同一个卷积核生成卷积层上的所有节点。</p> <p>以提供的软件 DemoCNN 为平台，在 RGB 与 Lab 颜色空间下预测某幅图像的关注区域(贴原图与检测结果图)，并进一步说明哪种颜色空间更适合于解决视频内容的优先传输问题。</p> <p>答：</p> <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;">  <p>图 11 原图</p> </div> <div style="text-align: center;">  <p>图 12 DT 图</p> </div> </div> <div style="display: flex; justify-content: space-around; margin-top: 20px;"> <div style="text-align: center;">  <p>图 13 RGB 空间下的预测图</p> </div> <div style="text-align: center;">  <p>图 14 Lab 空间下的预测图</p> </div> </div>	

题号	回 答 内 容	得分
4(1)	<p>LAB. 不像 RGB 和 CMYK 色彩空间, Lab 颜色被设计来接近人类视觉。它致力于感知均匀性, 它的 L 分量密切匹配人类亮度感知。因此可以被用来通过修改 a 和 b 分量的输出色阶来做精确的颜色平衡, 或使用 L 分量来调整亮度对比。这些变换在 RGB 或 CMYK 中是困难或不可能的——它们建模于物理设备的输出, 而不是人类的视觉感知。</p> <p>因为 Lab 空间比电脑屏幕、印表机甚至比人类视觉的色域都要大, 表示为 Lab 的位图比 RGB 或 CMYK 位图获得同样的精度要求更多的每像素数据。在 1990 年代, 这时的电脑硬件和软体通常受限于存储和操纵 8 位/通道的位图, 从 RGB 图象到 Lab 之间的来回转换是有损耗的操作。对于现在常见的 16 位/通道支持, 这就不是问题了。</p> <p>此外, Lab 空间内的很多“颜色”超出了人类视觉的视域, 因此纯粹是假想的; 这些“颜色”不能在物理世界中再生。通过颜色管理软件, 比如内置于图象编辑应用程序中的那些软件, 可以选择最接近的色域内近似, 在处理中变换亮度、彩度甚至色相。在图象操作的多个步骤之间使用假想色是很有用的。</p> <p>根据提供的论文, 简述 BPS 显著性检测的基本原理. 并回答一般有哪些通用的测试数据集?</p> <p>答: 图像的感兴趣区域通常由一些封闭的区域组成。一般来说, 轮廓和形状是重要的低级别图像的特征。HVS 可以感知和识别图像轮廓和形状的图像内容。对于图像区域, 即使其他地区造成的遮挡使得其形状不完整, 人们还可以感知其整个轮廓。研究人员经常发现高阶位平面包含大部分视觉上重要的数据, 即轮廓特征, 而低阶位平面有助于图像更细微的细节。</p> <p>基于感兴趣区域的封闭外观以及整个区域轮廓的 HVS 感知除了对比先验和背景先验, 封闭先验也被提出。closure prior 意味着一个图像区域有一个大的如果封闭, 概率是显着的。另一方面, 图像区域和背景有很强的对比度那么它就有很高的概率表现为一个封闭的区域。这里我们可以认为, 如果一个图像区域在所有位平面上出现一个闭合区域的机会更多, 那么这个图像区域有更大的显著性概率。</p> <p>一般使用的数据集有 MIT, ImgSal, Toronto, Kootstra 和 Cerf。MIT 包括 1003 幅带户外生活主题的图像。Toronto 收集 120 个图像。这个数据集中的一些图像的关注区域没有对焦。Kootstra 包含 100 张图片。此数据集中的图像分为六个主题: 动物, 车辆, 街道, 建筑物, 花卉和植物, 以及自然场景。Cerf 由 181 个图像组成。它的主题是人脸, 电话, 玩具等。ImgSal 有 235 图像。它的主题涵盖了场景, 户外生活, 动物, 昆虫, 鲜花和水果。</p>	

题号

回 答 内 容

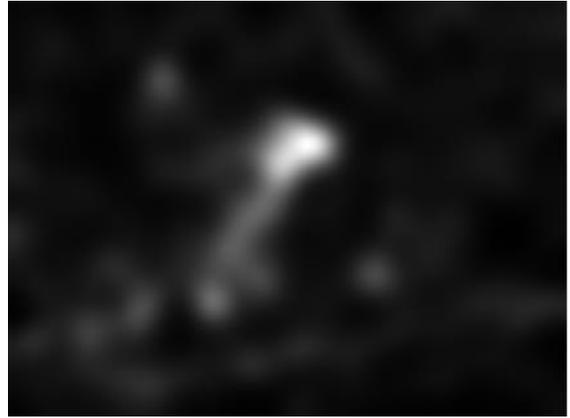
得分

4(2) 任选一幅提供的图像，用嵌入在 DemoSeamCarving 中的显著性检测算法做图像显著性检测，并贴出原图与显著性检测的结果。以 Ground truth 为依据，画出检测结果的 ROC 曲线。

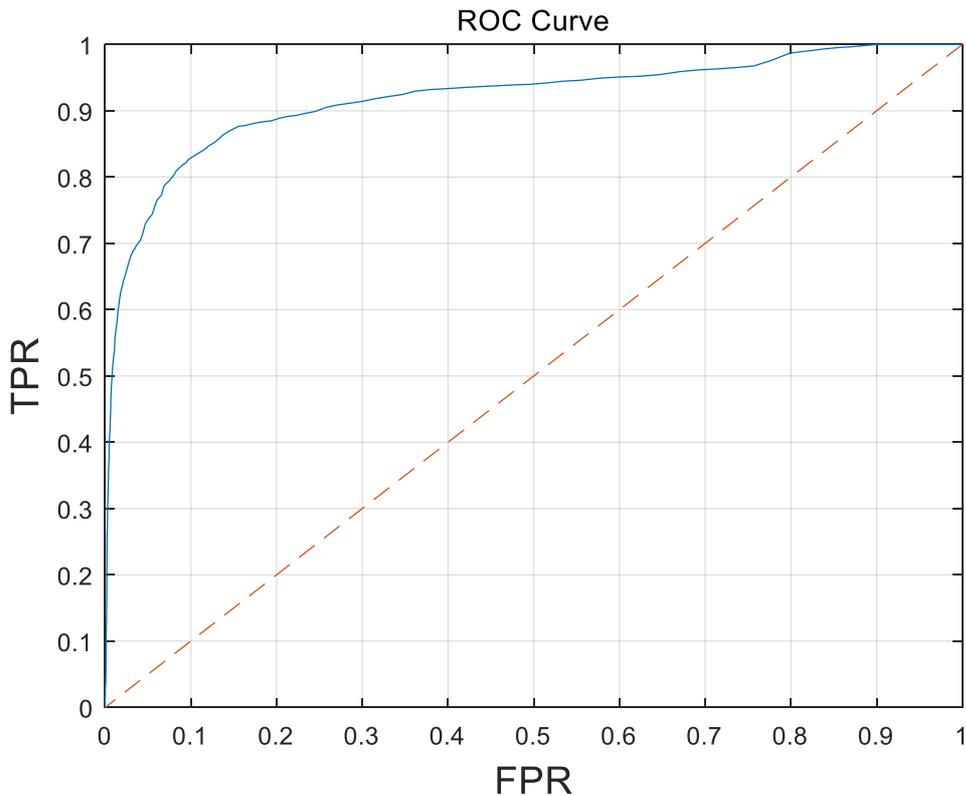
原图：



BPS 显著性检测的结果：



ROC 曲线：



下面是生成曲线的 matlab 代码：

题号	回 答 内 容	得分
	<pre> clc; clear; epsilon=0.0000001; gt_raw=imread('E:\Program Files\MATLAB\R2016a\bin\DemoSeamCarving\images\animals_00_gt.jpg'); gt_raw=normalization_0_1(gt_raw); [gt_height gt_width dimention]=size(gt_raw); x_RGB=imread('E:\Program Files\MATLAB\R2016a\bin\DemoSeamCarving\images\animals_00.jpg'); img_temp=callBPS(x_RGB); img_temp=normalization_0_1(img_temp); FPR=zeros(256,1); TPR=zeros(256,1); for threshold = 0:255 gt=im2bw(gt_raw,0+threshold*(1/255)); saliency_map=im2bw(img_temp,0+(threshold*(1/255))); TP=0; FP=0; FN=0; TN=0; for i=1:gt_height for j=1:gt_width if saliency_map(i,j)==1 if gt(i,j)==1 TP=TP+1; else FP=FP+1; end else %SAP(j)==0 if gt(i,j)==1 FN=FN+1; else TN=TN+1; end end end end end %横坐标FPR index=threshold+1; FPR(index,1)=FP/((FP+TN)+epsilon); TPR(index,1)=TP/((TP+FN)+epsilon); end plot(FPR,TPR,'LineWidth',1); xlim([0,1]); grid on; ylabel('TPR','FontSize',15);xlabel('FPR','FontSize',15); </pre>	

题号	回 答 内 容	得分
5(1)	<p>请回答视频在 YUV 颜色空间中的 4:2:0 与 4:2:2 格式是什么意思?</p> <p>答: 原理: 由于储存及传送的限制, 讯号处理中大多数会偏向被减少 (或被压缩) 以减低负荷。由于人眼对色度的敏感度不及对亮度的敏感度, 图像的色度分量不需要有和亮度分量相同的清晰度, 所以许多视讯系统在色差通道上进行较低 (相对亮度通道) 清晰度 (例如, 抽样频率) 的抽样。这样在不明显降低画面质量的同时降低了影像讯号的总带宽。因抽样而丢失的色度值用内插值, 或者前一色度值来替代。在压缩影像中, 以 4:2:2 Y'CbCr 为例, 它只需 R'G'B' (4:4:4) 三分之二的带宽。带宽的减少在肉眼上几乎没有影像上差别。</p> <p>抽样系统及比例: 视频系统的抽样系统中通常用一个三分比值表示: $J:a:b$ (例如 4:2:2), 形容一个以 J 个像素宽及两个像素高的概念上区域, 有时候会以四分比值表示 (例如 4:2:2:4)。依序列出为:</p> <p>J: 水平抽样参照 (概念上区域的宽度)。通常为 4。</p> <p>a: 在 J 个像素第一行中的色度抽样数目 (Cr, Cb)。</p> <p>b: 在 J 个像素第二行中的额外色度抽样数目 (Cr, Cb)。</p> <p>α: 水平因数 (与首数值有关连)。若没有此部分者可忽略, 或存在时与 J 相同。</p> <p>4:2:2: 每个色差通道的抽样率是亮度通道的一半, 即水平色差清晰度分半。这样从无压缩视频中可减少了三分之一的带宽, 当中的视觉差别仅仅少量甚至不存在。大多数高端数码视频格式采用这一比率, 如: AVC-Intra 100、Digital Betacam、DVCPRO50 及 DVCPRO HD、Digital-S、CCIR 601 / 串列数位界面 / D1、ProRes (HQ, 422, LT, and Proxy)、XDCAM HD422、Canon MXF HD422。</p> <p>4:2:0: 又称 I420。I420 是 YUV 格式的一种, 属于 planar format。4:2:0 并不意味着只有 Y,Cb 而没有 Cr 分量。它指的是对每行扫描线来说, 只有一种色度分量以 2:1 的抽样率存储。相邻的扫描行存储不同的色度分量, 也就是说, 如果一行是 4:2:0 的话, 下一行就是 4:0:2, 再下一行是 4:2:0...以此类推。对每个色度分量来说, 水平方向和垂直方向的抽样率都是 2:1, 所以可以说色度的抽样率是 4:1。PAL 制式和 SECAM 制式的色彩系统特别适合于用这种方式来存储。绝大多数视频编解码器都采用这种格式作为标准的输入格式。对非压缩的 8 比特量化的视频来说, 每个由 2x2 个 2 行 2 列相邻的像素组成的宏像素需要占用 6 字节内存。</p> <p>映射:</p> <p>码流 $Y_{00} U_{00} Y_{01} Y_{02} U_{02} Y_{03}$ $Y_{e0} V_{e0} Y_{e1} Y_{e2} V_{e2} Y_{e3}$</p> <p>将被映射为下面的两行各四个像素:</p> <p>$[Y_{00} U_{00} V_{e0}] [Y_{01} U_{00} V_{e0}] [Y_{02} U_{02} V_{e2}] [Y_{03} U_{02} V_{e2}]$ $[Y_{e0} U_{00} V_{e0}] [Y_{e1} U_{00} V_{e0}] [Y_{e2} U_{02} V_{e2}] [Y_{e3} U_{02} V_{e2}]$</p>	

5(2)

使用这种方法的质量很接近于 4:1:1,通常应用于下面的格式: 所有版本的 MPEG, 包括如 DVD 等 MPEG-2 成品(虽然有些 MPEG-4 profile 容许如 4:4:4 的更高质素采样), PAL DV 及 DVCAM, HDV, AVCHD, 最常见的 JPEG/JFIF、H.261, 及 MJPEG 成品, VC-1 等。

请回答分辨率格式为 CIF 的视频画面的边长, 并分析它们为什么会是 16 的倍数。

答: 在 H.264 码流中图像是以序列为单位进行组织的, 一个序列是由多帧图像被编码后的数据流, 以 I 帧开始, 到下一个 I 帧结束; 一帧图像可以分成一个或多个片(slice), 片由宏块组成, 宏块是编码处理的基本单位, 一个宏块由一个 16×16 的 YUV 亮度像素和附加的一个 8×8 Cb 和一个 8×8 Cr 彩色像素块组成, 当片编码之后会被打包进一个 NALU, 也就是一帧图像对应于一个 NALU。宏块(16*16)是编码的基本单位, 所以, CIF 的视频画面的边长都是 16 的倍数。

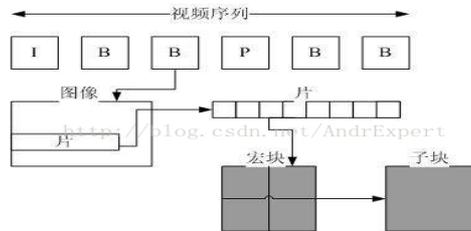


图 15 H.264 视频编码组织形式

另外, CIF 是常用的标准化图像格式(Common Intermediate Format)。在 H.323 协议簇中, 规定了视频采集设备的标准采集分辨率。CIF = 352×288 像素。对亮度和两个色差信号(Y、Cb 和 Cr)分量分别进行编码, 它们的取值范围同 ITU-R BT.601。即黑色=16, 白色=235, 色差的 最大值等于 240, 最小值等于 16。所以, 从这个方面也说明 CIF 的视频画面的边长为 16 的倍数。

5(3) 请用 JM 参考软件, 任选一个 CIF 大小的 4:2:0 的 YUV 格式的视频序列, 在量化参数为 20、24、28、32 时, 计算在各个量化参数下的 PSNR 与 Bite rate, 并画出 RD 曲线?

答: 对 flower_cif.yuv 视频进行测量, 具体数据如下表所示:

表格 2 JM 视频压缩结果

帧	I				P			
量化参数	20	24	28	32	20	24	28	32
bit/frame	42299	28830	18820	12890	139032	97699	77242	63785
SNR Y(dB)	45.995	42.891	39.915	37.308	42.58	36.891	33.001	30.198
SNR U(dB)	44.204	41.235	38.451	36.719	47.662	44.912	42.493	42.385
SNR V(dB)	42.618	38.886	35.829	33.775	44.77	42.389	40.44	38.239
SNR avg(DB)	44.27233	41.004	38.065	35.934	45.004	41.39733	38.64467	36.94067

帧	B			
量化参数	20	24	28	32
bit/frame	149419	114212	87493	63280
SNR Y(dB)	40.966	36.949	33.405	29.743
SNR U(dB)	40.131	36.075	32.263	28.905
SNR V(dB)	40.084	35.985	32.123	28.731
SNR avg(DB)	40.39367	36.336333	32.597	29.12633

RD 曲线如下图所示:

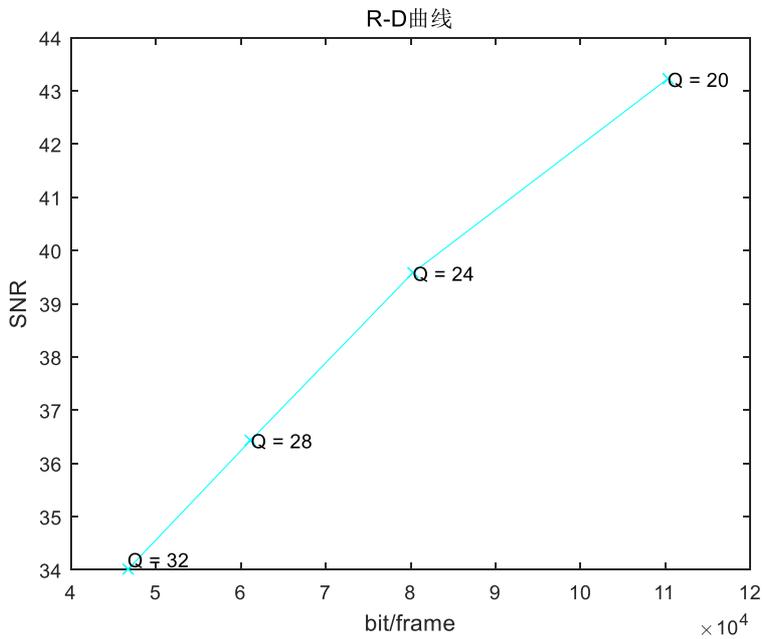


图 16 视频压缩的 RD 曲线

5(4) 请分析在不同量化参数下的编码的速度. 并用率失真公式 $J = D + \lambda R$ 解释量化参数与编码速度之间的关系.

答: 编码速度随时间的变化关系如下表和下图所示:

表格 3 编码速度随时间的关系表

量化参数	20	24	28	32
编码时间(s)	1.639	1.526	1.277	1.162
编码时间(fps)	1.83	1.97	2.35	2.58

题号

回 答 内 容

得分

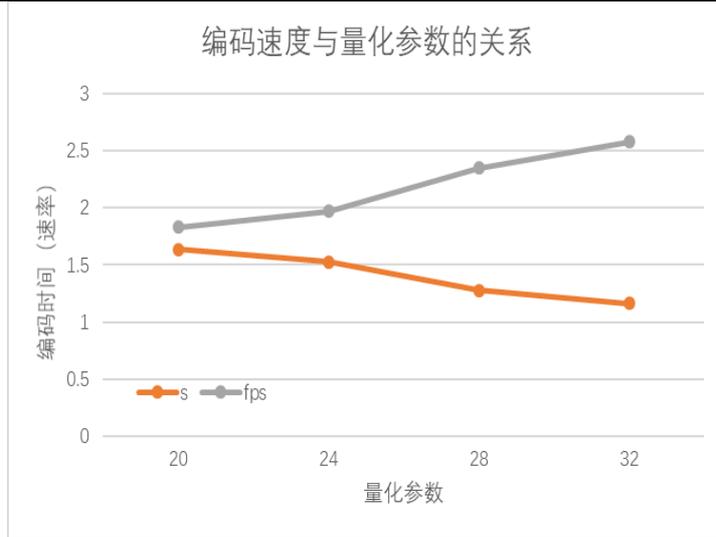


图 17 编码速度与量化参数的关系

从上图和上表中可以看到，总体上，量化参数选择越大，编码速度就加快，但同时信噪比也在降低，意味着更高的失真率。