Video Processing and Analysis (EEE6086) – Model Answers 2012/13

1.

a. (5)

```
1  0  1
   3  0  3
      1  0  1
-----------------------
1  3  2  3  1
```

The cascade is a 5-tap horizontal low-pass filter: [1, 3, 2, 3, 1]. The effect is smoothing/blurring.

b. (5)

```
 1   0  -1
 2   0  -2
 1   0  -1
```

The cascade is a 2D high-pass filter. The effect is enhancement of vertical edges.

c. (4)

An example of a 3x3 low-pass filter can be:

1	2	1
2	3	2
1	2	1

An edge enhancement filter can be obtained by the following procedure:

0	0	0
0	1	0
0	0	0

+ (

0	0	0
0	1	0
0	0	0

-1/15

1	2	1
2	3	2
1	2	1

)= 1/15

-1	-2	-1
-2	27	-2
-1	-2	-1

d. (6)

Box filter output for the bold pixel is: 1/9 * (90+60+30+30+150+60+30+30+60) = 60;
Median filter output for the bold pixel is: $(30, 30, 30, 30, 60, 60, 60, 90, 150)^T \cdot (0, 0, 0, 0, 1, 0, 0, 0, 0) = 60$;
α-trimmed-mean filter (α=5) output is: $(30, 30, 30, 30, 60, 60, 60, 90, 150)^T \cdot 1/5(0, 0, 1, 1, 1, 1, 1, 0, 0) = 48$;

2.

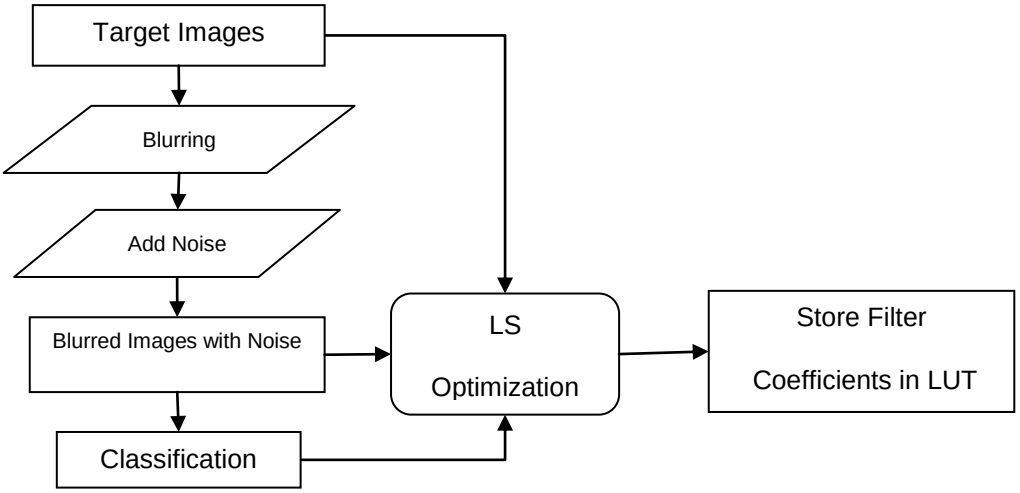
a. (4)

Enhancement techniques do not need motion estimation include (any two of the following):
Sharpness enhancement
Contrast enhancement
Noise reduction
Colour enhancement

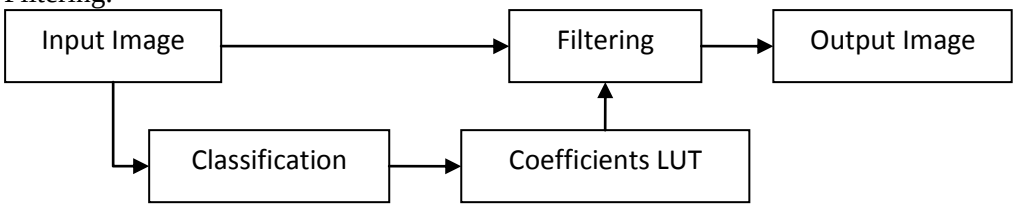
Enhancement techniques need motion estimation include:
Frame-rate upconversion
De-interlacing

b. (6)

Training:



Filtering:



c. (4)

In ADRC, each pixel value is compared to the average. For a pixel value higher than the average, 1 is assigned as the class code for that pixle; otherwise, 0. The final code is the concatenation of the class codes in all the pixels in the neighbourhood.

1	1	1
1	0	0
1	0	0

➔ 1 1 1 1 0 0 1 0 0

d. (6)

Other examples of classification besides ADRC include (any two of the following):
Local sharpness
Local contrast
Local noise/artefact level
Colour

The advantages and disadvantages of the trained filter compared with the New Edge Directed Interpolation (NEDI) are:
The coefficient optimisation in the trained filter is based on an offline training process and classification. So the online processing of the trained filter is more efficient but requires extra memory to store the LUT. The coefficient optimisation in NEDI is done during filtering and based on the assumption that interpolation is scale invariant. So NEDI is less efficient but requires no memory for storing filter coefficients.

3.

a. (4)

Motion estimation in video coding tries to minimise the differences between frames, but the motion does not have to be true motion in the video sequence;
Picture rate conversion requires a true motion estimation method.

b. (3)
The major drawback of a full search motion estimator is it is very slow. Methods to tackle this problem:
Sub-sampled search
One-at-a-time search

c. (5)
The two assumptions that 3DRS is based on:
Objects are larger than blocks
Objects have inertia

The candidate set of the 3DRS blocking matching algorithm:
Spatial candidates
Temporal candidates
Update candidates

d. (4)
Given a high frame-rate video sequence, a low frame-rate sequence can be generated by down-sampling. The down-sampled sequence can be up-converted using a true motion estimator. Then the performance of the true motion estimator can be evaluated by measuring the error between the original frames and the interpolated frames.

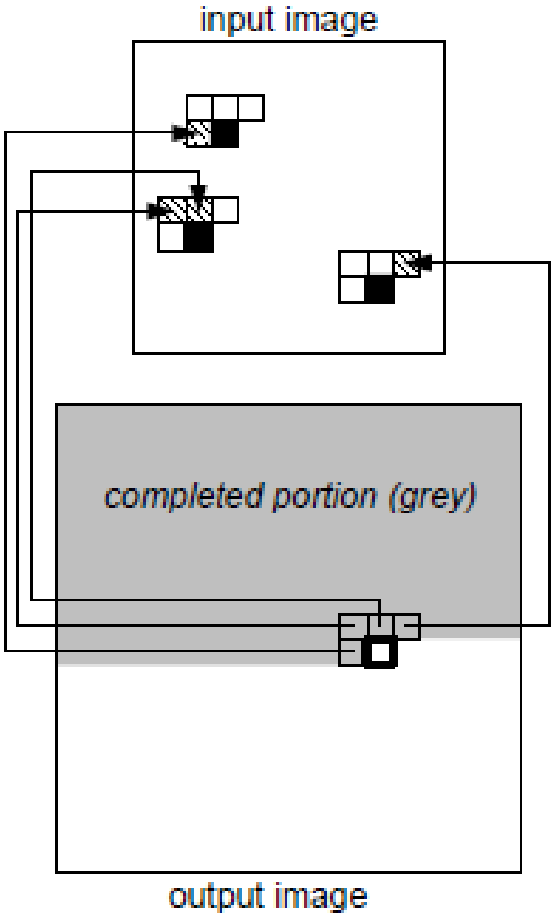
e. (4)
A simple motion adaptive de-interlacing method works as follows:
If motion is detected, intra-field edge dependent interpolation is used; otherwise, use field insertion (copy pixels from the previous field).

4.
a. (6)
The bilateral filter is an edge-preserving smoothing filter. The weighting coefficients are dependent on both space and range.
The three difficulties of the bilateral filter are:
Poor smoothing in high gradient regions
Smoothes and blunts cliffs, valleys and ridges
Can combine disjoint regions.

b. (5)
The non-local means filter searches for similar patches as the current patch and takes advantage of the correlation of image content. The output is the weighted average of those similar patches.
The non-local means filter would work the best on images containing repetitive patterns, such as texture.

c. (4)
The bilateral filter is more effective for the Gaussian noise. Because compression artifacts, especially the blocking artifacts, have large distortions on the pixel values which would be more difficult for the bilateral filter to distinguish from real edges in the image.

d. (5)
The coherency search in pixel-based texture synthesis is illustrated as follows:


input image
completed portion (grey)
output image