1.6.1 Identify three novel applications of the Internet or multimedia applications. Discuss why you think these are novel. (5%)

Solution: edushi; online video game; telemedicine; social networking service; html5; OS for smart phones; SoundHound; Dropbox; Text to speech; Content-based image search engines; Web OS; Online PS; Internet telephony; 3D movie; Adobe Air; Google Maps; b2b; Vertical Search;

1.6.3 Your task is to think about the transmission of smell over the Internet. Suppose we have a smell sensor at one location and wish to transmit the Aroma Vector (say) to a receiver to reproduce the same sensation. You are asked to design such a system. List three key issue to consider and two applications of such a delivery system. Hint: Think about medical applications. (8%)

Solution: 1. How to express a kind of aroma in a vector; 2. How to eliminate the effect from surroundings; 3. How to setup redundancy to avoid inaccuracy; 4. How to transmit data quickly, even real-time; 5. Compress and decompress data; 6. How to reproduce the same aroma; 7. The loss of data when compress.
Two applications: Telemedicine; Remote cooking; Remote shopping; Environment monitoring; Bad environmental exploration.

3.4.4 Suppose we have available 24 bits per pixel for a color image. However, we notice that humans are more sensitive to R and G than to B—in fact, 1.5 times more sensitive to R or G than to B. How could we best make use of the bits available? (6%)

Solution: $1.5*x + 1.5*x + x = 24$, so $x = 6$. The best way is to use 9 bits for R and G, 6 bits for B.

3.4.5 At your job, you have decided to impress the boss by using up more disk space for the company's grayscale images. Instead of using 8 bits per pixel, you'd like to use 48 bits per pixel in RGB. How could you store the original grayscale images so that in the new format they would appear the same as they used to, visually? (7%)

Solution: Assign 16bits for each channel, and then multiply the original gray value by 2^8, copy the new value to each channel.

4.5.1 Consider the following set of color-related terms: (a) Wavelength (b) Color level (c) Brightness (d) Whiteness   How would you match each of the following (more vaguely stated) characteristics to each of the above terms? (e) Luminance (f) Hue (g) Saturation (h) Chrominance (8%)

Solution: (a) Wavelength↔(f) Hue;   (b) Color level↔(h) Chrominance;   (c) Brightness↔(e)

Luminance;   (d) Whiteness↔(g) Saturation;

4.5.3 "The LAB gamut covers all colors in the visible spectrum." (8%)

   (a) What does this statement mean? Briefly, how does LAB relate to color? Just be descriptive.

   (b) What are (roughly) the relative sizes of the LAB gamut, the CMYK gamut, and a monitor gamut?

Solution: (a) CIELAB (also known as L*a*b*) uses a power law of 1/3 instead of a logarithm. CIELAB uses three values that correspond roughly to luminance and a pair that combine to make colorfulness and hue (variables have an asterisk to differentiate them from previous versions devised by the CIE). The color difference is defined as

$$\Delta E = \sqrt{(L^*)^2 + (a^*)^2 + (b^*)^2} \qquad (4.21)$$

where

$$L^* = 116 \left(\frac{Y}{Y_n}\right)^{\left(\frac{1}{3}\right)} - 16$$

$$a^* = 500 \left[\left(\frac{X}{X_n}\right)^{\left(\frac{1}{3}\right)} - \left(\frac{Y}{Y_n}\right)^{\left(\frac{1}{3}\right)}\right] \qquad (4.22)$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n}\right)^{\left(\frac{1}{3}\right)} - \left(\frac{Z}{Z_n}\right)^{\left(\frac{1}{3}\right)}\right]$$

with X_n, Y_n, Z_n the XYZ values of the white point. Auxiliary definitions are

$$\text{chroma} = c^* = \sqrt{(a^*)^2 + (b^*)^2}$$

$$\text{hue angle} = h^* = arctan\frac{b^*}{a^*} \qquad (4.23)$$

Roughly, the maximum and minimum of value a^* correspond to red and green, while b^* ranges from yellow to blue. The chroma is a scale of colorfulness, with more colorful (more saturated) colors occupying the outside of the CIELAB solid at each L^* brightness level, and more washed-out (desaturated) colors nearer the central achromatic axis. The hue angle expresses more or less what most people mean by "the color" – that is, you would describe it as red or orange.

(b) LAB gamut > monitor gamut > CMYK gamut

5.5.6 We don't see flicker on a workstation screen when displaying video at NTSC frame rate. Why do you think this might be? (5%)

Solution: http://blogold.chinaunix.net/u3/111328/showart_2228954.html
NTSC video is displayed at 30 frames per sec, so flicker is possibly present. Nonetheless, when video is displayed on a workstation screen the video buffer is read and then rendered on the screen at a much higher rate, typically the refresh rate – 60 to 90 Hz – so no flicker is perceived.

(And in fact most display systems have double buffers, completely removing flicker: since main memory is much faster than video memory, keep a copy of the screen in main memory and then when this buffer update is complete, the whole buffer is copied to the video buffer.)

5.5.9 What is the advantage of interlaced video? What are some of its problems? (8%) Page114, Page116

Solution: http://hi.baidu.com/operationsystem/blog/item/e19e279ab1f68abcc8eaf4e7.html
(1) Advantages: Interlacing was invented because, when standards were being defined, it was difficult to transmit the amount of information in a full frame quickly enough to avoid flicker. The double number of fields presented to the eye reduces perceived flicker. Because of interlacing, the odd and even lines are displayed in time from each other. This is generally not noticeable except when fast action is taking place onscreen, when blurring may occur.
(2) Disadvantages: Since it is sometimes necessary to change the frame rate, resize, or even produce stills from an interlaced source video, various schemes are used to *de-interlace* it. The simplest de-interlacing method consists of discarding one field and duplicating the scan lines of the other field, which results in the information in one field being lost completely. Other, more complicated methods retain information from both fields.

6.5.4 Suppose the sampling frequency is 1.5 times the true frequency. What is the alias frequency? (6%) Page129

Solution: If the sampling frequency is less than twice the true frequency, and is greater than the true frequency, then the alias frequency equals the sampling frequency minus the true frequency.

$$f_{alias} = f_{sampling} - f_{true}, for\ f_{true} < f_{sampling} < 2 \times f_{true}.$$

Then the alias frequency is 0.5 times the true frequency.

6.5.9 Draw a diagram showing a sinusoid at 5.5 kHz and sampling at 8 kHz (show eight intervals between samples in your plot). Draw the alias at 2.5 kHz and show that in the eight sample intervals, exactly 5.5 cycles of the true signal fit into 2.5 cycles of the alias signal. (10%)

Solution:
画出用 8kHz 采样 5.5kHz sinusoid 的示意图。
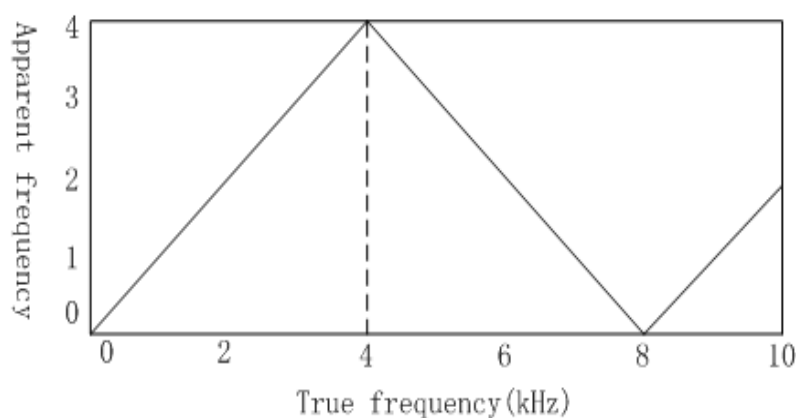
课本 P131 图 6.5 可知，采样后的信号为 2.5kHz，采样 8kHz,考虑在 1ms 内的原信号，相位为 0，这样在原来的信号上 1ms 有 5.5 个周期，画出 sin(5.5x) $x \in [0, 2\pi)$ 在原信号上采样 8 个点，每 0.25π 采样一个点。结果如下图所示。
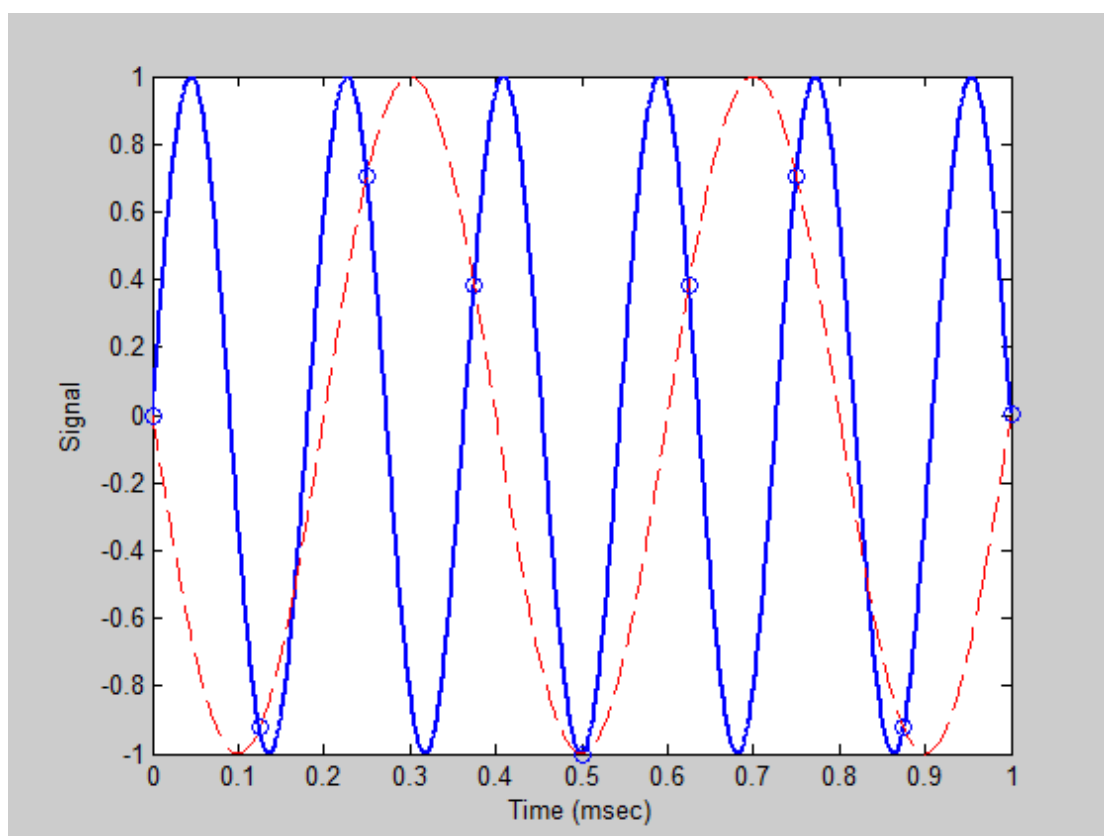
7.9.1 Suppose eight characters have a distribution A: (1), B: (1), C: (1), D: (2), E: (3), F: (5), G: (5), H: (10). Draw a Huffman tree for this distribution. (Because the algorithm may group subtrees with equal probability in a different order, your answer is not strictly unique.) (10%)

Solution:
An example:

28

18

8

5

2   3   10

A: (1)   B: (1)   C: (1)   D: (2)   E: (3)   F: (5)   G: (5)   H: (10)
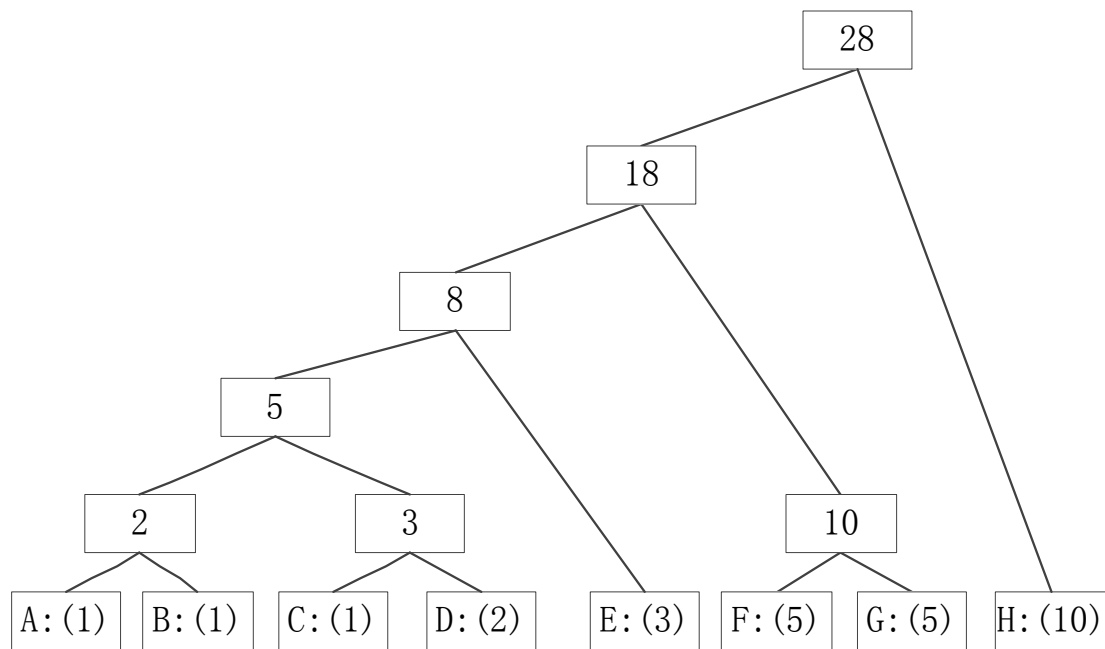
7.9.4 Extended Huffman Coding assigns one codeword to each group of k symbols. Why is *average (l)* (the average number of bits for each symbol) still no less than the entropy η as indicated in equation (7.7)? (6%) P176

Solution: Let $S^{(k)}$ denote the block source, $\bar{l}^{(k)}$ denote the average code length for each block, $\eta^{(k)}$ denote the entropy of the block source. According to $\eta \leq \bar{l} < \eta + 1$ (7.6), we get $\eta^{(k)} \leq \bar{l}^{(k)} < \eta^{(k)} + 1$; obviously, $\bar{l}^{(k)} = k\bar{l}$; From the two formulas above, $\frac{\eta^{(k)}}{k} \leq \bar{l} < \frac{\eta^{(k)}}{k} + \frac{1}{k}$; if

each symbol has a entropy of η, then it makes sense that $\eta^{(k)} = \overbrace{\eta + \eta \ldots \ldots + \eta}^{k} = k\eta$, so at last

we obtain $\eta \leq \bar{l} < \eta + \frac{1}{k}$.

8.11.4 If the block size for a 2D DCT transform is 8×8, and we use only the DC components to create a thumbnail image, what fraction of the original pixels would we be using? (5%) P207, P213

Solution: 1/64, because each 8×8 block has only one DC.

8.11.6 We could use a similar DCT scheme for video streams by using a 3D version of DCT. Suppose one color component of a video has pixels f_{ijk} at position *(i, j)* and time k. How could we define its 3D DCT transform? (8%)

Solution:

$$F[u] = \sqrt{\frac{2}{N}} C(u) \cdot \sum_{k=0}^{N-1} \cos \frac{(2k+1)u\pi}{2N} \cdot f[k]$$

$$F[u,v,w] = 2\sqrt{\frac{2}{MNT}} C(u)C(v)C(w) \sum_{i=0}^{M-1} \sum_{j=0}^{T-1} \sum_{k=0}^{N-1} \cos \frac{(2i+1)u\pi}{2M} \cdot \cos \frac{(2j+1)v\pi}{2T}$$

$$\cdot \cos \frac{(2k+1)w\pi}{2N} \cdot f[i,j,k]$$

$$F[u,v,w] = \frac{1}{4}\sqrt{\frac{2}{N}} C(u)C(v)C(w) \sum_{i=0}^{7} \sum_{j=0}^{7} \sum_{k=0}^{N-1} \cos \frac{(2i+1)u\pi}{16} \cdot \cos \frac{(2j+1)v\pi}{16} \cdot \cos \frac{(2k+1)w\pi}{2N}$$

$$\cdot f[i,j,k]$$

You must decide on the number N of frames.