

ADSP HW 3

羅政捷 R08942112

1. Write a Matlab or Python program for the 4:2:0 image compression technique.

Input image



Reconstructed image



2. Write two concepts you learned from the oral presentation on 5/1.

1. The principle of particle filter: represent the pdf as a set of weighted samples.
2. There are different map projections to align panorama photos.

3. (a) Suppose that, for an instrument, the frequency of Do is 300 Hz. What are the frequencies of So and La for the instrument?

$$So : 300 \times (2^{1/12})^7 \approx 449.5 Hz$$

$$La : 300 \times (2^{1/12})^9 \approx 504.5 Hz$$

3. (b) Why the music signal is easier to recognition than a speech signal? (Write at least 2 reasons)

1. Music signal has significantly less syllable than speech signal.
2. Context is often required to determine the speech signal.

3. (c) In the noiseless case, in what condition we cannot use the variation of amplitude to separate a speech signal into several syllables?

If a syllable lacks consonant, its amplitude might not go down from the previous syllable.

3. (d) Is it possible to hear ultrasound or infrasound? Why?

The commonly stated range of human hearing is 20 to 20,000 Hz, and since ultrasound and infrasound are outside of the range, we can't hear them.

4. (a) Write two possible methods to compress a cartoon image efficiently.

1. Only encode the edge.
2. Edge can be represented by some lines or arcs.

4. (b) Write two possible ways to compress a song more efficiently.

1. $X(f)$ concentrates on $n \times f_0$.
2. Frequencies that are $f \times Do \times 2^{1/12}$.

☞

5. (a) Why we always use the DCT instead of the DFT and the KLT to image compression? (Write two reasons).

1. DCT is near optimal for most images.
2. DCT is independent of the input.

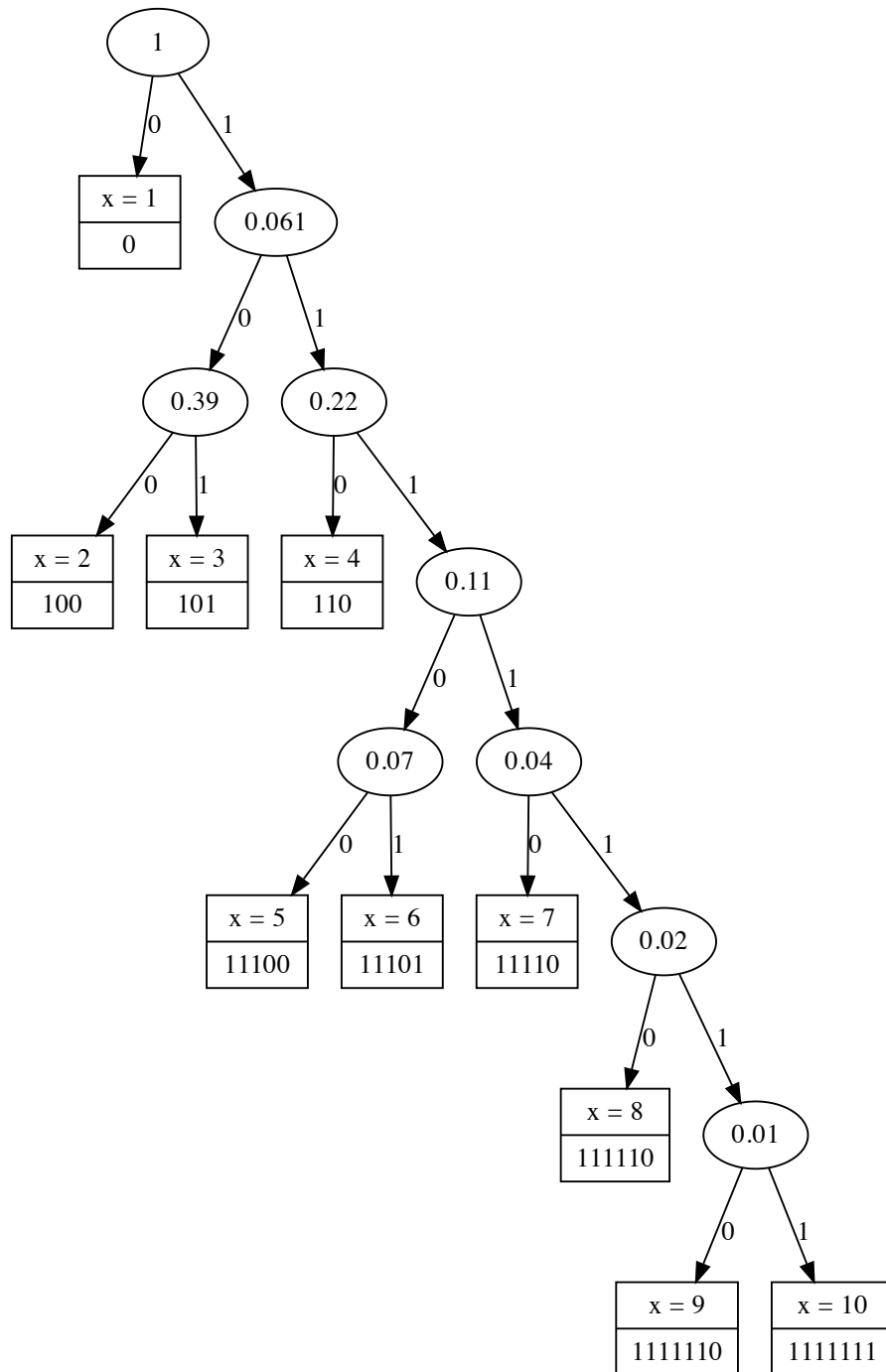
5. (b) Why we always use the 8x8 DCT instead of performing the DCT on the whole image for compression? (Write three reasons)

☞

1. Different parts of an image have different frequency distribution.
2. 8x8 DCT has lower complexity: $\mathcal{O}(MN \log_2 MN)$ vs $\mathcal{O}(MN)$
3. Less memory consumption.

6. (a) Determine the coding tree of x when using the Huffman code in the binary system.

$$\begin{aligned} P(x = 1) &= 0.4, P(x = 2) = 0.22, P(x = 3) = 0.17, P(x = 4) = 0.1, \\ P(x = 5) &= 0.04, P(x = 6) = 0.03, P(x = 7) = 0.02, P(x = 8) = 0.01, \\ P(x = 9) &= 0.006, P(x = 10) = 0.004, P(x = n) = 0 \text{ otherwise.} \end{aligned}$$



6. Suppose that $\text{length}(x) = 100,000$. Estimate the range of the total coding lengths in the k -ary system when using (1) the Huffman code and (2) the arithmetic code. Express the solution in terms of k .

$N = 100,000$

Total coding length: b

$$1. \text{ceil}(N \cdot \frac{\text{entropy}}{\log k}) \leq b \leq \text{floor}(N \cdot \frac{\text{entropy}}{\log k} + N)$$

$$2. \text{ceil}(N \cdot \frac{\text{entropy}}{\log k}) \leq b \leq \text{floor}(N \cdot \frac{\text{entropy}}{\log k} + \log_k 2 + 1)$$

Extra. What is the format for uncompressed audio files?

WAV is the audio format for storing uncompressed audio.