2019 Summer Entrance Examination

Department of Creative Informatics
Graduate School of Information Science and Technology
The University of Tokyo

Programming

INSTRUCTIONS

1. Do not open this problem brochure until the signal to begin is given.
2. Write your examinee ID number below on this cover page.
3. An answer sheet and a draft sheet accompany this brochure. Write down your examinee ID number on these sheets.
4. You should have received a USB flash drive. *Before the examination starts, copy the files from the USB flash drive to your PC.* Verify that you can see its contents and then take your hands off your PC. If you have some problems, consult the exam supervisor.
5. You may choose any programming language to answer.
6. You may consult only one printed manual of a programming language during the examination. *You can use or copy any libraries or program fragments stored in your PC, but you may not connect to the Internet.*
7. By the end of the examination, make a directory/folder on your PC, whose name is the same as your examinee ID number, and put your program files and related files into the directory/folder. Copy the directory/folder onto the USB flash drive that you received.
8. At the end of the examination, the USB flash drive, the answer sheet and the draft sheet will be collected.
9. After these are collected, stay at your seat, until all the examinee program results have been checked briefly by the test supervisor.
10. After the brief check, try to save your program execution environment on the PC so that you can run your program as soon as possible during the oral examination in the afternoon.
11. *Leave your PC and this brochure together in the room* for the oral examination and stay out of the room until you are called.

Examinee ID ______________________

This page is blank.

This page is blank.

Programming

A text file contains integers from 0 to 255 which are separated by a single white-space character. Assume that the number of these integers is a multiple of three. Group every three numbers into a triplet. Each triplet represents the intensities of red, green, and blue of a pixel. For example,

```
19 7 0 17 13 1 29 3 27 5 11 23
```

are grouped into four triplets (or pixels), which are (19, 7, 0), (17, 13, 1), (29, 3, 27), and (5, 11, 23). Every pixel has an index. The index of the pixel is i if it is the i-th triplet ($i \geq 0$) in the file. In the example above, (19, 7, 0) is the 0th triplet.

We can construct an image by placing these integer triplets, or pixels, from left to right, and when reaching the width of the image, placing them in the next line below, and so forth. Assume that the image is rectangular. Answer the following questions by writing a program if necessary.

(1) Write on the answer sheet the number of the pixels stored in the file `image1.txt`.

(2) We construct an image from the pixels stored in the file `image1.txt`. All the rightmost pixels of the image are white, that is, the triplet (255, 255, 255). The image includes no other white vertical line from the top to the bottom than the rightmost one. Write the width of the image on the answer sheet.

(3) Write a program that prints the $\frac{n}{2}$-th triplet (or pixel) and its index when sorting the pixels stored in the file `image1.txt` in the ascending order of intensities. Write down the printed triplet and index on the answer sheet, too. Here, n is the number of the pixels and it is an even number. The first pixel is the zeroth triplet. The intensity of a pixel is $r^2 + g^2 + b^2$ when the triplet (r, g, b) denotes the pixel. If there are two pixels with the same intensity, the pixel with a larger index has a lower intensity.

(4) Write a program that selects k pixels e_i, where $0 \leq i < k$, from the pixels stored in the file `image2.txt`. The program prints the triplets and the indices denoting these k pixels. Here, e_i is the $\frac{ni}{k}$-th pixel when the pixels are sorted in the order mentioned in (3). n is the number of pixels and n is a multiple of k.

Write on the answer sheet all the triplets of the pixels and their indices selected when $k = 4$.

(5) Write a program that selects k colors representing all the pixels stored in the given file. k is an input. The k representing colors are selected as follows:

1. Select k pixels e_i as mentioned in (4). Let them initial representative pixels $p_i^{(0)} = e_i$.
2. Categorize all the pixels into k clusters. For each pixel q_j, find the nearest representative pixel $p_i^{(t)}$. Then the pixel q_j belongs to a cluster $C_i^{(t)}$, where $t \geq 0$. The representative pixel $p_i^{(t)}$ belongs to the cluster $C_i^{(t)}$.

3. Compute the centroid of each cluster $C_i^{(t)}$. In $C_i^{(t)}$, the nearest pixel to that centroid is a new representative pixel $p_i^{(t+1)}$. Here, the centroid of pixels is a triplet of the averages (use the floor function after division) of each element (r, g, b) of the pixels.

4. Categorize pixels again into k clusters $C_i^{(t+1)}$ by using the new representative pixels $p_i^{(t+1)}$.

5. Repeat this ten times and obtain k representative pixels $p_i^{(10)}$. The colors we want to obtain are the triplets of these representative pixels.

The distance between two triplets (r_i, g_i, b_i) and (r_j, g_j, b_j) is $|r_i - r_j| + |g_i - g_j| + |b_i - b_j|$. For 2 and 3, if multiple pixels have the same distance, select the pixel with the largest index.

Then, compute the representative pixels $p_i^{(10)}$ of the pixels stored in the file `image2.txt` when $k = 128$. Write on the answer sheet the triplets of $p_i^{(10)}$ where $i = 40, 80, 120$. Do the same thing for `image3.txt` when $k = 8, i = 2, 4, 6$.

(6) Write a program that reduces the number of colors used in the given image in the method mentioned in (5). The program reads the image from a file and stores the resulting image into a file `image.tif` in the format shown below. Assume that the shape of the image is square. After the reduction, the pixels in a cluster $C_i^{(10)}$ are set to the color of its representative pixel $p_i^{(10)}$. Then, assuming $k = 32$, reduce the number of colors for the image in the file `image2.txt` and save the obtained `image.tif` into the USB flash drive.

The format of the file `image.tif` is as follows. It consists of 104-byte attribute information and pixel data. Each byte of the first 104 bytes in the file is the following number (in decimal notation), respectively.

```
77 77  0 42  0   0  0  8  0  7  1  0  0  4  0  0
 0  1 w0 w1 w2  w3  1  1  0  4  0  0  0  1 h0 h1
h2 h3  1  2  0   3  0  0  0  3  0  0  0 98  1  6
 0  3  0  0  0   1  0  2  0  0  1 17  0  4  0  0
 0  1  0  0  0 104  1 21  0  3  0  0  0  1  0  3
 0  0  1 23  0   4  0  0  0  1 s0 s1 s2 s3  0  0
 0  0  0  8  0   8  0  8
```

Here, `w0 w1 w2 w3`, `h0 h1 h2 h3`, and `s0 s1 s2 s3` denote the 4 byte big-endian values representing the width, the height, and (width)×(height)×3.

After these 104 bytes, the pixels of the image are stored from the top to the bottom line. For each line, the pixels are stored from left to right. For each pixel, each element of the triplet (r, g, b) is stored in this order as a 1 byte value. For example, when the width is 100 pixels and the height is 50 pixels, then 15104 bytes data in total are stored.

This page is blank.