ORIGINAL

NATIONAL UNIVERSITY OF SINGAPORE

EXAMINATION FOR

(Semester I: 2018/2019)

EE5731 – VISUAL COMPUTING

November / December 2018 - Time Allowed: 2.5 Hours

INSTRUCTIONS TO CANDIDATES:

1. This paper contains **FOUR (4)** questions and comprises **FIVE (5)** printed pages.
2. Answer all questions, and answer them as detailed as possible.
3. This is a CLOSED BOOK examination but candidates are allowed to bring in **ONE (1)** A4 sheet of paper, on which they may write any information, into the examination hall.
4. Total marks are 100.

Q1:

A. Regarding SIFT, answer the following questions:

i. What is the reason that SIFT descriptors are robust to lighting changes? Include in your answer the parts of the algorithm that are responsible for this and the justification.

ii. How can SIFT maintain its invariance despite change of perspective (like shown in Figure **Q1.A**)? Also, in your answer include the discussion why this invariance is robust only up to a 60-degree angle of changes, and which parts of the algorithm is responsible for this.

[12 points]

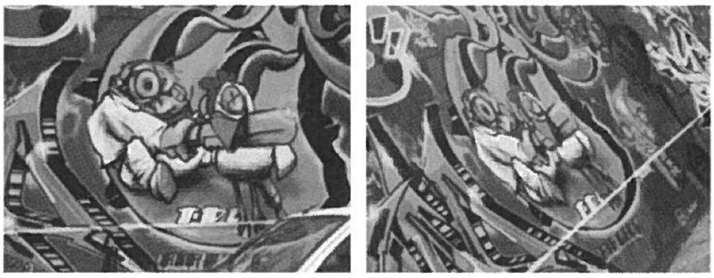

Figure Q1.A. Images taken from a different perspective

B. Regarding Viola-Jones' object detection algorithm, answer the following questions:

i. Referring to Figure **Q1.B**, how can we set the value of T? Intuitively, what does α_t represent? What are the implications of updating the weights?

ii. Demonstrate how the integral image and cascade of classifiers can speed up the algorithm. Write either a pseudocode or a step-by-step algorithm to support your answer.

[12 points]

- For $t = 1, \ldots, T$:

1. Normalize the weights, $w_{t,i} \leftarrow \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}$
2. Select the best weak classifier with respect to the weighted error

$$\epsilon_t = \min_{f,p,\theta} \sum_i w_i \,|\, h(x_i, f, p, \theta) - y_i \,| \,.$$

3. Define $h_t(x) = h(x, f_t, p_t, \theta_t)$ where f_t, p_t, and θ_t are the minimizers of ϵ_t.
4. Update the weights:

$$w_{t+1,i} = w_{t,i}\beta_t^{1-e_i}$$

where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$.

- The final strong classifier is:

$$C(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha_t = \log \frac{1}{\beta_t}$

Figure Q1.B. Part of Viola-Jones algorithm for object detection

Q2:

A. Regarding image geometric properties, answer the following questions:
   i. Justify why the homography transformation works only for planar surfaces.
   ii. Explain why the fundamental matrix is scene independent.

[14 points]

B. The following questions are about the epipolar geometry:
   i. In what situation does the fundamental matrix become a zero matrix? Prove your answer mathematically.
   ii. Referring to Figure **Q2.B**, suppose you are given two images of a scene and a point x^h, discuss what you need to do in order to form the backward projection line from $C_{t'}$ to x'^h_∞.

[12 points]

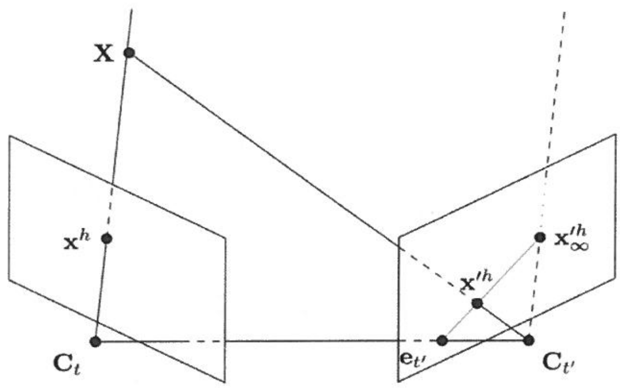

Figure Q2.B. Epipolar Geometry

Q3:

A. Regarding Markov random field (MRF), answer the following questions:

i. In the MRF graphical models, there are nodes representing the hidden states and nodes representing the observed data. Provide justification why there are no edges (connections) between nodes representing the observed data (the pixels in the input image).

ii. Suppose you are given a task to detect rural houses using MRF. A typical example of the input image is shown in Figure **Q3.A** (a), and the expected output is (b). Design the data term function of your MRF and justify your function.

Fig.Q3.A: (a) Input image. (b) Output image

[12 points]

B. Regarding Bayesian filtering, answer the following questions:

i. Provide the detailed correlation between the Hidden Markov Model and Kalman Filter, by first drawing the graphical model that is based on the formulation of Kalman Filter, and then providing the justification of the correlation.

ii. Particle filter computes the weight using the formula below. Discuss the consequences if w_{t-1}^{m} is removed from the equation.

$$w_t^m = w_{t-1}^m \frac{p(d_t|Z_t^m)p(Z_t^m|Z_{t-1}^m)}{q(Z_t^m|Z_{t-1}^m, d_t)}$$

[14 points]

Q4:

A. To deal with the lighting changes, optical flow algorithms can use either the gradient constancy constraint or structure-texture decomposition. For comparisons between these two methods, mention 3 benefits/drawbacks of them (regardless of your choice being benefits or drawbacks, in total you need to mention only 3 items), and provide justification.

[12 points]

B. Assess whether the following statements are right or wrong, and provide your justification:
   i. Rudin-Osher-Fatemi (ROF) algorithm suppresses the structure layer information based on the total brightness of the pixel values.
   ii. ROF algorithm works only when the input image is noisy or the lighting changes.
   iii. ROF algorithm can be employed to handle images suffer from a moderate presence of rain.

[12 points]

END OF PAPER