CARDIFF UNIVERSITY
EXAMINATION PAPER

Academic Year:	2015/2016
Examination Period:	Autumn
Examination Paper Number:	CMT107
Examination Paper Title:	Visual Computing
Duration:	2 hours

Do not turn this page over until instructed to do so by the Senior Invigilator.

Structure of Examination Paper:

There are 5 pages.
This examination paper is divided into 2 sections.
There are 4 questions in total.
There are no appendices.
The maximum mark for the exam paper is 100%, and the mark obtainable for a question or part of a question is shown in brackets alongside the question.

Students to be provided with:

The following items of stationery are to be provided:
ONE answer book.

Instructions to Students:

Answer the COMPULSORY question in Section A and TWO questions from Section B.

Important note: if you answer more than the number of questions instructed, then answers will be marked in the order they appear only until the above instruction is met. Extra answers will be ignored. Clearly cancel any answers not intended for marking. Write clearly on the front of the answer book the numbers of the answers to be marked.

The use of translation dictionaries between English or Welsh and a foreign language bearing an appropriate school stamp is permitted in this examination.

SECTION A

Q1. **Compulsory Question**

(a) Besides Computer Graphics, what other *sub-fields* are included in Visual Computing? Discuss the relations among these sub-fields based on their primary inputs and outputs. [7]

(b) Consider the homogeneous coordinate representations listed below. Convert them into Cartesian coordinates if the conversion is meaningful. Otherwise, explain why they cannot be converted.
a) $(0, 7, 2)$, b) $(3, 0, 0)$, c) $(0, 0, 0)$, d) $(0, 0, 0, 5)$. [4]

(c) Give the mathematical formula for the Phong illumination model and explain each term in detail. [8]

(d) A vector $\mathbf{w} = [1, 3, 2]^T$ is decomposed as the sum of the two orthogonal vectors $\mathbf{u}$ and $\mathbf{v}$, where $\mathbf{u}$ is parallel to the vector $\mathbf{z} = [2, 1, -1]^T$. Calculate $\mathbf{u}$ and $\mathbf{v}$. [8]

(e) Dilate the image with the structuring element given below, where the points with '$\times$' are local origins with value 0.

0	1	0	0
0	1	0	0
0	0	1	1
0	0	1	0
$\times$	0	1	0

Image to be processed

1	0
$\times$	0
1	1

Structuring element

[8]

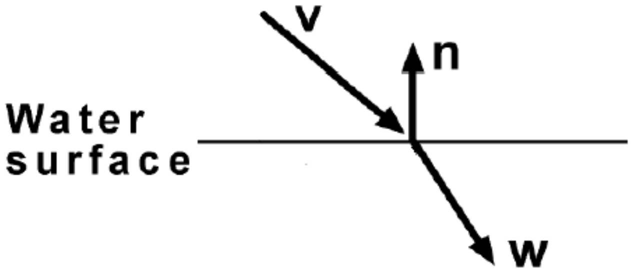
(f) With the aid of a suitable diagram, explain how to construct the vanishing point of a line. [5]

SECTION B

Q2. **Ray Tracing**

(a) Describe a basic ray tracing algorithm which handles reflections and transparency. There is no need to provide any mathematical formulae, but you should clearly state what has to be computed. [12]

(b) A light vector v passing from air into water is refracted on the water surface along the direction of vector w. See the figure below, where n is the *unit* normal vector of the water surface. Assume that the projections of v and w onto n have the same length and that the projection of w onto the water surface has half the length of the projection of v onto the surface. Explain how to compute the resulting vector w based on given vectors v and n.



[8]

(c) An extended light source is a light source which emits light from a surface, e.g. a sphere, rather than from a single point. With the help of a simple diagram, explain how extended light sources create soft shadows. Briefly describe how you might extend the basic ray tracing algorithm to handle extended light sources to give soft shadows. [10]

Q3. **Image Processing**

(a) **[Image filtering]** Answer the questions related to the following four filters: 1) *box filter*, 2) *Gaussian filter*, 3) *median filter*, 4) *unsharp mask filter*.

(i) Which of the above filters *cannot* be represented as a convolution? [4]

(ii) Which filter or combination of filters can be used to remove the following noise:
1) *Gaussian noise*, 2) *Gaussian and uniform noise* 3) *salt-and-pepper noise* 4) *Gaussian and salt-and-pepper noise*.
In the case of combination of filters, please describe the order of using each filter. Justify your answer. [10]

(iii) The pixel values of an image and a filter kernel (denoted by g) are given in the two tables below.

10	13	13	42	51
34	20	12	0	10
15	25	$\boxed{\mathbf{73}}$	13	14
55	60	100	30	15
35	40	105	20	20

The Pixel Values

0.03	0.12	0.03
0.12	0.40	0.12
0.03	0.12	0.03

The filter Kernel: g

Compute the filtering results of the pixel in the centre of the image (with value 73) using:
1) *a* 3×3 *box filter*, 2) *a* 3×3 *median filter*, 3) *the filter kernel* g, and 4) *the unsharp mask filter based on g with 100% of detail being added back to the original image*.

[8]

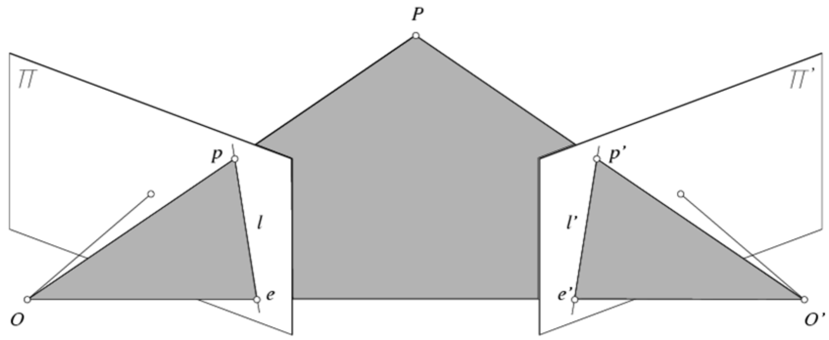
(b) **[Canny Edge Detector]**

(i) Briefly describe the main steps of the *Canny edge detector*. [4]

(ii) What are the purposes of non-maximum suppression and hysteresis that are done in the Canny edge detector. [4]

Q4. **Epipolar Geometry and Stereo**

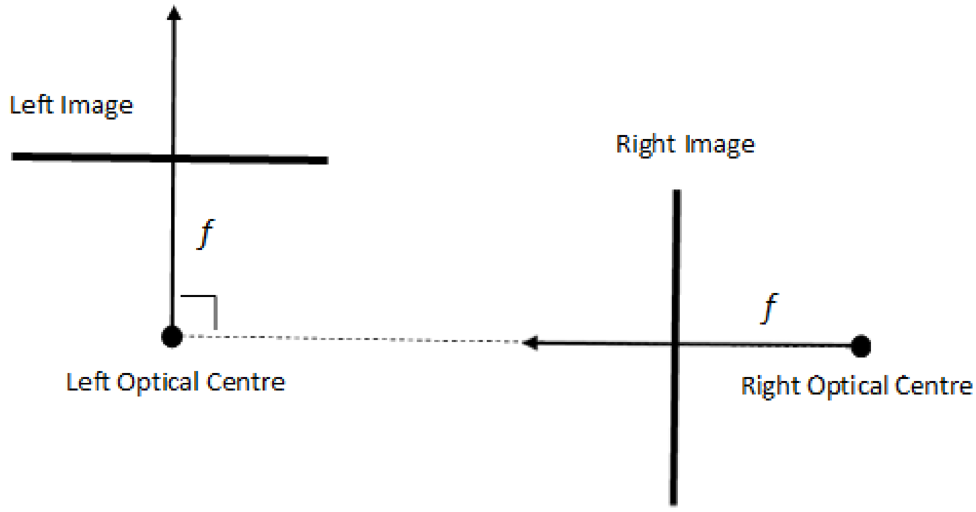
(a) The following figure shows the epipolar geometry with O and O' as the optical centres of a pair of cameras.



(i) Based on this figure, explain the following terms: *baseline*, *epipolar plane*, *epipoles* and *epipolar lines* [4]

(ii) What is the *epipolar constraint* and how can it be used in *stereo matching*? [6]

(b) Consider the following top view of a stereo device. Suppose the distance between the two optical centres is 50cm, and f = 10cm. Answer the following questions:



(i) Draw the *front views* of the two 2D images that show and clearly label the *approximate positions of the epipoles and epipolar lines* for this configuration of cameras. [6]

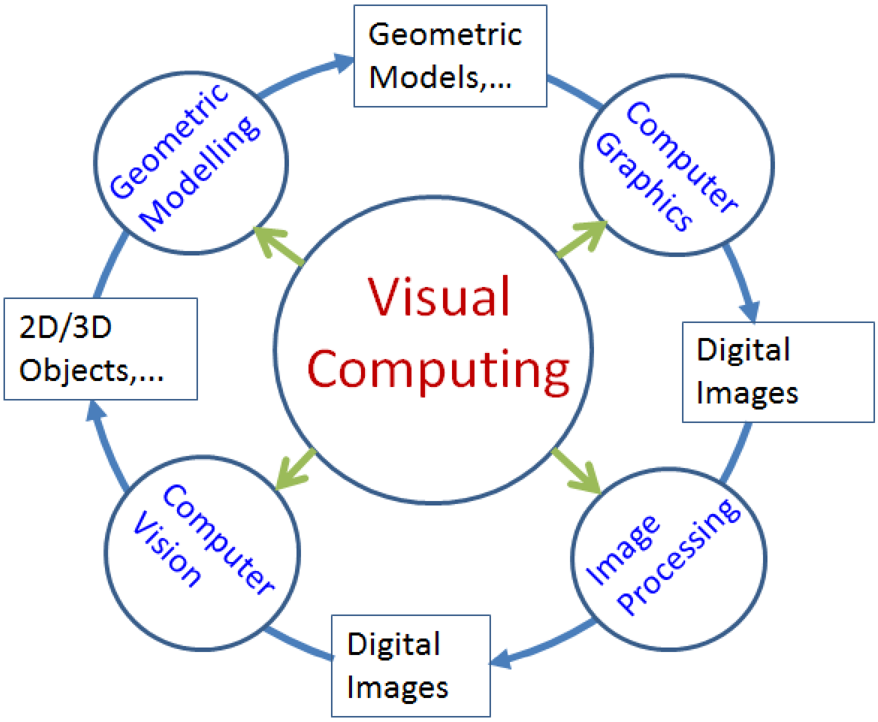
(ii) Compute the *essential matrix* for this configuration. [6]

(iii) Describe the epipolar line on the left image associated with a point $p' = (100, 200)$ on the right image. [4]

(iv) In order to simplify stereo matching, the images are rectified so that they are parallel to each other and to the baseline. Give the homography matrices for both images. [4]

SECTION A

A1. (a) The following figure clearly shows the answer.
(1 point for each sub-field except for Computer Graphics. 1 point for each input and output description)



[7]

(b) *1 point for each answer.*
a) $(0, 3.5)$, b) an infinite point on the direction $(3, 0)$, c) undefined, d) $(0, 0, 0)$. [4]

(c) • For a given monochromatic light, Phong illumination model calculates the reflected light by

$$R_a I_a + R_d(n^T d) I_d + R_s(r^T v)^\sigma I_s$$

(4)

- R_a, R_d, and R_s are surface reflectance coefficients for ambient, diffuse, and specular light, respectively. (1)
- I_a, I_d, and I_s are the intensities of the incoming ambient, diffuse, and specular light, respectively. (1)
- n is the unit surface normal; d is the unit direction from surface point to light source; r is unit direction of perfect reflection, v is unit direction towards viewer poisition, and σ is the shininess exponent. (1)
- The summation over all light sources for red, green blue gives total intensity for all colours. (1)

[8]

(d) • First calculate $\mathbf{u}$, the projection of $\mathbf{w}$ on the direction of $\mathbf{z}$.

$$\mathbf{u} = \frac{\mathbf{w}\cdot\mathbf{z}}{\mathbf{z}\cdot\mathbf{z}}\mathbf{z} = \frac{1\times 2+3\times 1-2\times 1}{2^2+1^2+(-1)^2}\begin{bmatrix}2\\1\\-1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}2\\1\\-1\end{bmatrix}$$

(5)

• Then calculate the vector $\mathbf{v}$.

$$\mathbf{v} = \mathbf{w} - \mathbf{u} = \frac{1}{2}\begin{bmatrix}0\\5\\5\end{bmatrix}$$

(3)

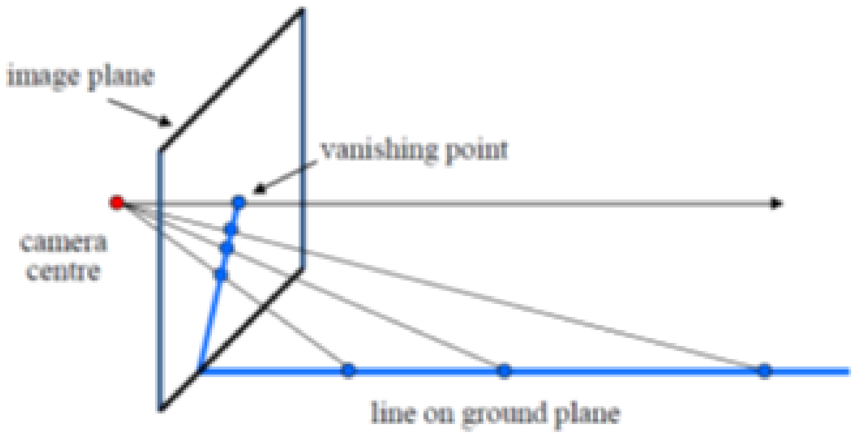
[8]

(e) See below

0	1	0	0	0
0	1	0	0	0
0	1	1	1	0
0	1	1	0	0
0	0	1	1	1
×	0	1	1	0
0	0	1	1	0

[8]

(f) See the picture below.



Suppose we have a line on the ground, and we want to find its vanishing point on the image plane. Select at least two points on the line, and find the intersection points between the image plane and the lines linking the camera centre c and these points, separately. All this intersection points form a line L on the image plane. The vanishing point is a point p on the line L, such that the line linking c and p is parallel to the line on the ground. [5]

SECTION B

A2. (a) I. Input: viewing position v, look-at point a, up vector u such that viewing plane is centered at a and orthogonal to $a - v$ and u is parallel to y-direction. (1)

II. For each pixel (x, y), $x = 0, \ldots$ `X-Resolution`, $y = 0, \ldots$ `Y-Resolution` do

(i) Create a ray from the viewing position v in direction d such that it passes through (x, y) in the viewing plane.

(ii) Set colour of (x, y) to the return value of `raytrace`(v,d). (2)

III. function `raytrace` (v,d):

(i) Initialise position t on ray l from v in direction d to infinity and the nearest object n to empty. (1)

(ii) For each object o in the scene

i. Compute intersection p of l and o closest to v

ii. If p exists and it is closer to v than t, set t to p and n to o. (2)

(iii) If n is empty, return background colour Else (1)

i. If n is reflective and we haven't reached the maximum recursion depth level, compute perfect reflection vector r of d at t and call `raytrace` (t, r) to obtain reflected colour c_r (1)

ii. If n is transparent and we haven't reached the maximum recursion depth level, compute refraction vector r' of d at t and call `raytrace` (t, r') to obtain refracted colour c_t (1)

iii. For each light source $k = 1, \ldots, m$ at position l_k, cast ray from t to l_k. If this line segment intersects with any of the other objects, t is in the shadow of this object. Otherwise compute the amount of light c_k reaching t from k. (2)

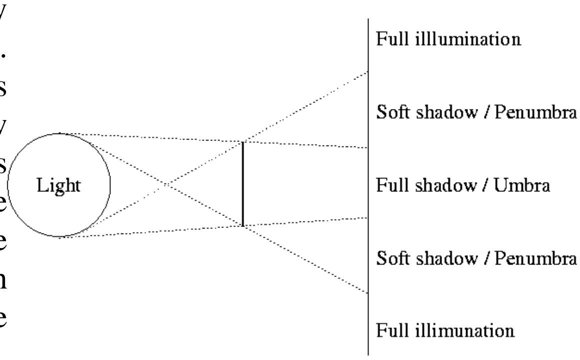
iv. Return combination of colours c_r, c_t and c_k, $k = 1, \ldots m$. (1)

[12]

(b)
- Let $m = (v^t n)n$ (3)
- Let $e = v - m$ (2)
- Then $w = m + e/2 = (v + m)/2 = (v + (v^t n)n)/2$ (3)

[8]

(c)
- An extended light source may only be partially "visible" from a point. The areas where the light source is fully occluded are in the full shadow of the occluding surface; the areas where the light source if fully visible are fully lit; the areas in between are partially lit depending on how much of the light source is visible from the position.



(4)

- For each extended light source cast multiple rays from the surface point to the extended light source to cover the whole light source surface. (3)

- The percentage of the rays that are not blocked by other surfaces gives the percentage of light reaching the point from the extended light source. (3)

[10]

A3. (a) (i) *1 point for each filter being correctly indicated as "can" or "cannot"*
3) cannot represented as a convolution. [4]

(ii) *2 points each for 1), 2) and 3), and 4 points and 4)*
1) Use Gaussian filter (or box filter) because Gaussian filter is a low pass filter, which can remove high-frequency noise like Gaussian noise.
2) Use Gaussian filter (or box filter) because both Gaussian noise and uniform noise are high-frequency noise.
3) Use median filter because it replaces a black or white pixel with the median value of its neighbouring pixels.
4) First use median filter and then use Gaussian filter to remove Gaussian and salt-and-pepper noises because in this way, we can first remove salt-and-pepper noise, and then remove Gaussian noise. If we do it the other way round, then because Gaussian filter may first blur the black and white pixels, further using median filter may have not effect. [10]

(iii) *2 point for each correct answer.*
1) 37, 2) 25, 3) 50.5, 4) 95.5. [8]

(b) (i) *1 point for each item:*

- Filter image with derivative of Gaussian
- Find magnitude and orientation of gradient
- non-maximum suppression
- Link and threshold (hysteresis)

[4]

(ii) *2 points for each item:*

- Non-maximum suppression thins wide ridges down to single pixel width
- Hysteresis uses a high threshold to eliminate weak edges that are caused by noise, and a low threshold to fill in the gaps between strong edge points where only weak edge response is detected.

[4]

A4. (a) (i) In the figure,

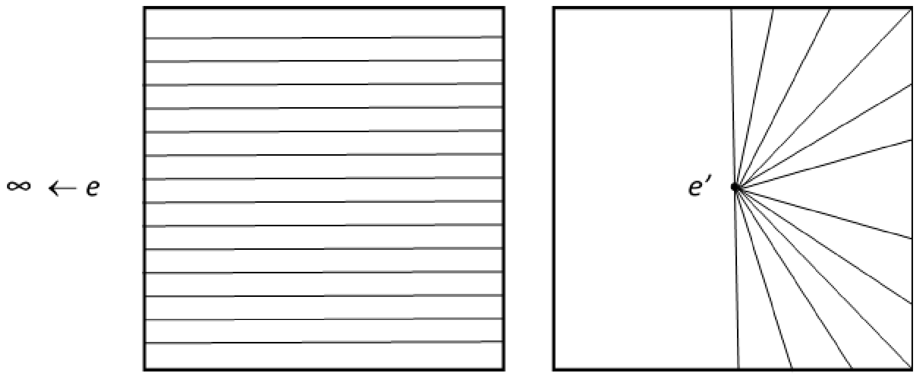
- Line $O - O'$ is the baseline; (1)
- $OO'P$ is on an epipolar plane; (1)
- e and e' are epipoles; (1)
- l and l' are epipolar lines (1)

[4]

(ii)
- The epipolar constraint represents geometry of two cameras, reduces a correspondence problem to 1D search along an epipolar line. (2)
- A point in one view "generates" an epipolar line in the other view. The corresponding point lies on this line. (2)
- Epipolar geometry is a result of co-planarity between camera centres and a world point - all of them lie in the epipolar plane. (2)

[6]

(b) (i) *3 points for each figure*



[6]

(ii)
- The translation vector $t = [50,\ 0,\ 0]^T$.
- The rotation matrix

$$R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

- the essential matrix is

$$E = [t_\times]\,R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -50 \\ 0 & 50 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 50 & 0 & 0 \\ 0 & 50 & 0 \end{bmatrix}$$

[6]

(iii) The epipolar line associated with p' is

$$l = Ep' = \begin{bmatrix} 0 & 0 & 0 \\ 50 & 0 & 0 \\ 0 & 50 & 0 \end{bmatrix} \cdot \begin{bmatrix} 100 \\ 200 \\ -f \end{bmatrix} = \begin{bmatrix} 0 \\ 5000 \\ 10000 \end{bmatrix}.$$

So, any point $p = (x, y)$ on this epipolar line can be represented by

$$\begin{bmatrix} x \\ y \\ -f \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 5000 \\ 10000 \end{bmatrix} = 0.$$

i.e., $y = 2f = 20$. [4]

(iv) The left image don't need to be rectified, so it's homography matrix is a 3×3 identity matrix. The right image camera need to be rotated around the up direction (y-axis) $-90°$, so the homography matrix is

$$R = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

[4]