UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Final Exam 2015

Prescription Number: COSC428

Paper Title: *Computer Vision*

Time allowed: TWO hours

Number of pages: ?

- This exam is worth a total of 100 marks
- Contribution to final grade: 40%
- Length: ? questions
- Answer *all* questions.
- Calculators are *not* allowed.
- *This is a closed book test*.
- Use the separate *Answer Booklet* for answering *all* questions.

1 (12 marks total)

Briefly describe advantages and/or disadvantages of the following four different types of camera technologies **for acquiring image depth values**. [1 mark for each advantage *or* disadvantage cited]

(a) structured light camera [3 marks]
(b) time-of-flight camera [3 marks]
(c) stereo camera [3 marks]
(d) LIDAR (Light Detection And Ranging) [3 marks]

2 (8 marks)

a) **Spectral Resolution**. Humans can perceive 10 octaves of sound frequencies, from 20Hz to 20kHz. State the approximate spectral resolution (wavelength in nm) that can humans perceive. [2 marks]
b) **Dynamic Range**. The difference in intensity between the softest perceivable sound and the loudest sound that can be tolerated without pain is a ratio of 10^9:1. State the approximate visual dynamic range

(as a ratio) that humans can perceive in regards to the difference between the lowest perceptible light intensity and the highest intensity we can tolerate without glare. [2 marks]

c) **Spatial Resolution**. State the approximate number of centimetres spatial resolution that humans can perceive at 20 metres. [2 marks]

d) **Radiometric Resolution**. Regardless of our spectral resolution and dynamic range, most humans can only reliably distinguish between a limited number of colours and shades of grey. State the approximate number of colours and the approximate number of shades of grey that humans can reliably distinguish between (when placed in a graduated scale on the same page). [2 marks]

3 (7 marks)

The Hough (pronounced "huff") transform (HT) can detect a line using a "voting" scheme where points vote for a set of parameters describing a line. The more votes for a particular set, the more evidence that the corresponding line is present in the image. So it can detect MULTIPLE lines in one shot.

To find straight lines:

- A line in the image corresponds to a point in Hough space
- To go from image space to Hough space:
    - given a set of points (x,y), find all (m,b) such that y = mx + b

Basic Hough transform algorithm

1. Initialize H[d, θ]=0
2. for each edge point I[x,y] in the image
   for θ = 0 to 180
   $d = x\cos\theta + y\sin\theta$
   H[d, θ] += 1
3. Find the value(s) of (d, θ) where H[d, θ] is maximum
4. The detected line in the image is given by $d = x\cos\theta + y\sin\theta$

Describe how this Hough transform can be generalised to detect curved lines in an image, even when the curve does not have a simple analytic form.

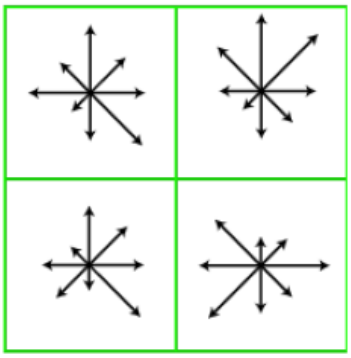
4 (9 marks total)

TextonBoost is a good algorithm for segmenting textured regions in an image.
Describe the following three steps of TextonBoost: [3 marks each]

(a) convolution and clustering
(b) boosted texture layout filters
(c) alpha-expansion graph cut

5 (8 marks)

The orientation of objects can be tracked from one frame to the next using the *scale-invariant feature transform* (SIFT) which extracts scale and rotation invariant features from images. SIFT is also used to recognise objects and match different views of a scene for stereo vision. In this algorithm, descriptors of key-points across an image are created as a set of 8 gradient orientations about a pixel as shown in the diagram to the right.

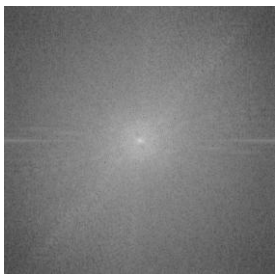


Name and describe the gradient operators you would use to create such a set of **eight** orientations as shown in the diagram to the right and also describe how you would use them to create these orientations. (Hint: If a gradient operator is only centred on a key-point, then the only result is a gradient in one direction equal to the gradient in the opposite direction – in contrast to orientations shown in the diagram to the right.)

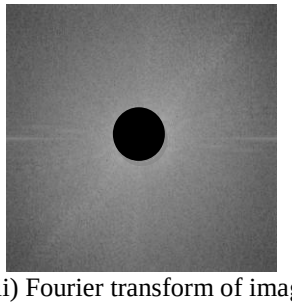
6 (12 marks)

Images can be filtered by applying an inverse Fourier transform to a Fourier transformed image. Given the following image and Fourier transforms of that original image:

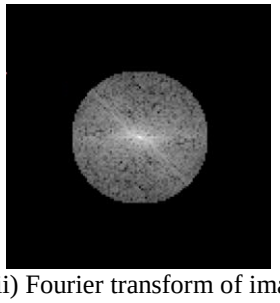
Original cheetah image



(i) Fourier transform of image



(ii) Fourier transform of image with centre blacked out



(iii) Fourier transform of image with all but the centre zeroed

(a) Is image (i) the magnitude transform or phase transform of the original image? Explain the appearance of image (i) [2 marks]
(b) If an inverse Fourier transform is applied to images (ii) and (iii), describe the resulting images and also explain why they would appear as described [4 marks]
(c) The Fourier transform is a self-inverting transform. Explain what this means. [2 marks]
(d) Describe the purpose of the main components of the following equations for a forward Fourier transform and an inverse Fourier transform in the discrete domain: [4 marks]

7 (12 marks total)

Briefly describe the following morphological operators and explain what effect they have on an image and why they have such an effect:

(a) Erosion [3 marks]
(b) Dilation [3 marks]
(c) Open [3 marks]
(d) Close [3 marks]

8 (10 marks total)

A fiducial marker can be used to find the six degree-of-freedom pose of a camera.
Describe the following five steps for finding this six degree-of-freedom pose: [2 marks each]

(a) Fiducial marker detection
(b) Rectangle fitting
(c) Pattern checking
(d) Lens undistortion
(e) Pose estimation

9 (6 marks total)

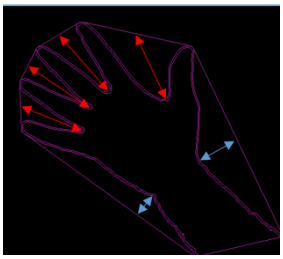
The Bundle Adjustment algorithm can enable accurate 3D reconstruction from multiple views. Describe this algorithm.

10 (16 marks)

You are to briefly describe **only four** of the following 2015 class projects [for 4 marks each] by listing at least four algorithmic steps (for each of the four projects), naming the algorithms used in the order they were used. **Do not select your own or similar project** (e.g. face recognition projects do not select other face recognition projects, etc).

(i) Find outlines of textured regions using
"Extendible Edge Detection for Real-time Systems"

[[
variance ridge detector
calculate/smooth the variance flow or texture based flow smoothing
computing the localised derivative of the variance by kernel
computing the variance image
apply to the variance flow to highlight edges at specific orientations
threshold vector magnitudes
]]



(ii) Recognise hand pose using "Hand Gesture Recognition"

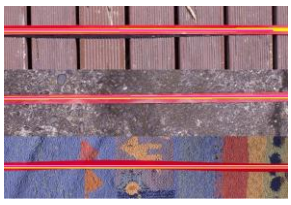
[[
background subtraction using Adaptive Gaussian Mixture Model
use the Canny edge detector to outline the hand
find the convex hull of the segmented hand
calculate distances/depths of concave hulls to convex hull (defects)
double threshold depth values (not too small, not too large)
]]

(iii) Identifying tree branches in "A Novel Approach to Tree Limb Identification"

[[
tree manually segmented from the background
bilateral filter to Gaussian smooth image while preserving edges
Shi-Tomasi corner detector (or Harris)
Canny edge detector
Suzuki border following algorithm.
]]

(iv) "A Method of 3D Object Reconstruction" using photos of an object taken with a mobile phone from different viewpoints

[[
pyramid adapted FAST for feature matching
Iterative Linear-LS for triangulation
EPnP for camera pose estimation
Levenberg-Marquardt for bundle adjustment
]]



(v) Track lines of electrical tape on the ground with "A Novel Approach to Line Extraction using Saturation Thresholding"

HSL decomposition of the image for background segmentation
Peak and Valley Histogram thresholding
Suzuki algorithm to extract a set of contours
Ramer-Douglas-Peucker algorithm
Vector decomposition on contours to find parallel lines

]]

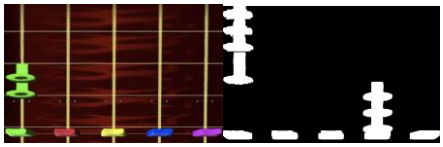
(vi) Recognize plate numbers that contain both Arabic and Latin characters using "Bilingual Number Plate Recognition"

[[
grayscale conversion, then Gaussian blur to remove noise
Sobel operator to localise number plate
threshold
support vector machine classifier to verify plate
characters are segmented
OCR classification
artificial neural network to train classifying characters
]]

(vii) Collect data from controlled intersections using a "Vision Based Traffic Light Detection System"

[[
convert to HSV colour model to detect red, orange and green lights
Canny Edge detection multistep algorithm used to find contours within colour specific images
Gaussian smoothing reduces noise
eclipse detection to identify circular shapes in colour specific images
]]

(viii) “Runway and note detection for the computer game Frets on Fire”



[[
compare a Canny edge map and a HSV threshold
Hough transform to detect the lines on the runway
Erosion and dilation to remove noise and holes
]]

END OF PAPER