IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2014-2015

PAPER C316

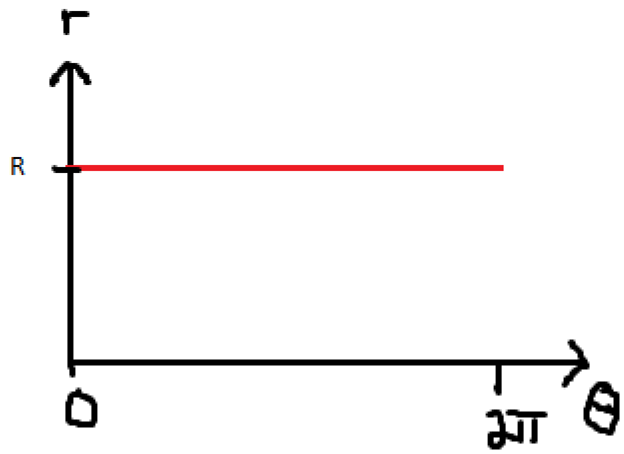
COMPUTER VISION - SOLUTIONS

Tuesday 16 December 2014, 10:00
Duration: 120 minutes

1. a) Template matching can be performed by applying the inverse fourier transform to the correlation of $F(u)G(u)*$ where $F(u)$ is the signal of the input image, $G(u)$ is the signal of the template with $G(u)*$ being the complex conjugate. Although the technique is fast, the continuous space doesn't generalise to discrete space well.

   b) i) By finding the centre of each shape, we can plot the distance from the centre to the perimeter at every angle from 0 to 2π. By taking the fourier transform, this signal can then be represented as a sum of sines and cosines with coefficients for every frequency. Each shape will have different co-efficients which distinguishes them.

   ii) The fourier descriptor for a circle is the **fourier transform** of the horizontal line with radious R (shape descriptor for the circle)



   iii) 1. Easy to store shape as we just need to store the co-efficients of the sines/cosines.
   2. If we have two shapes that are scaled, with the fourier descriptor we can just normalise the co-efficients by dividing by the perimeter and then compare them.
   3. If two shapes are rotated and we want to match them, rotation causes a change in phase to occur so we can just match against the phase shift.

   c) i) $f(x) = \sum_{u=0}^{N-1} F(u)exp(2\pi ju\frac{x}{N})$

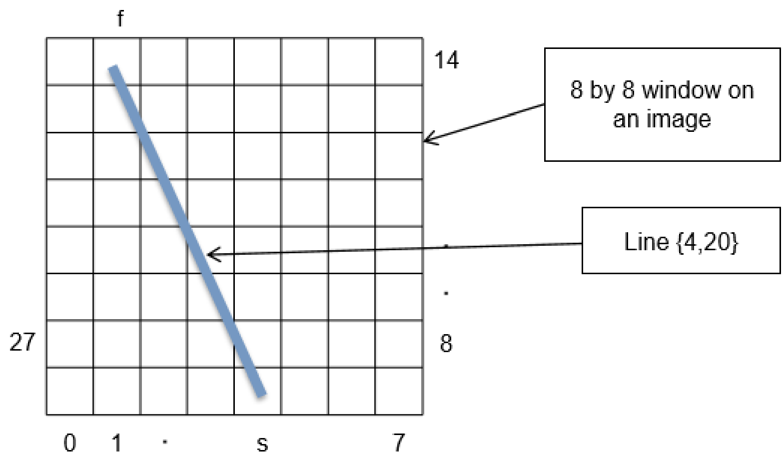
   ii) $f(0) = 1$
   $f(1) = 2$
   $f(2) = 3$

$f(3) = 2$
$N = 4$
$F(u) = \frac{1}{4}\sum_{x=0}^{3} f(x)exp(-2\pi ju\frac{x}{4})$
$= \frac{1}{4}(f(0)exp(0)+f(1)exp(-2\pi ju\frac{1}{4})+f(2)exp(-2\pi ju\frac{2}{4})+f(3)exp(-2\pi ju\frac{3}{4}))$
$= \frac{1}{4}(1 + 2exp(-\frac{\pi}{2}ju) + 3exp(-\pi ju) + 2exp(-\frac{3}{2}\pi ju))$
$F(0) = \frac{1}{4}(1 + 2 + 3 + 2) = 2$
$F(1) = \frac{1}{4}(1 + 2exp(-\frac{\pi}{2}j) + 3exp(-\pi j) + 2exp(-\frac{3}{2}\pi j))$
$= \frac{1}{4}(1+2(\cos(-\frac{\pi}{2})+i\sin(-\frac{\pi}{2}))+3(\cos(-\pi)+i\sin(-\pi))+2(\cos(-\frac{3}{2}\pi)+$
$i\sin(-\frac{3}{2}\pi)))$
$= \frac{1}{4}(1 + 2(0 - i) + 3(-1 + 0i) + 2(0 + i))$
$= -\frac{1}{2}$
$F(2) = 0$
$F(3) = -\frac{1}{2}$

2. a) i) Intrisics are useful because they are independant from any other source and can be used with priori knowledge to classify an object. For example a heart has four chambers and the general shape is circular. We can use this when doing image processing.

   ii) One heuristic to reduce side lobes is to remove edge points from the parameter space that don't agree with the line direction.

   b) If the image space is already quantised then you can use the wallace hough transform. A line would be represented by its edge points; starting with 0 the bottom left of the grid then increasing going anti-clockwise:



c) part c
   i) Algorithm 1 goes here
   ii) Algorithm 2 goes here

3. a) i) Local features have locality which are associated with object features. Also there are an abundance of them.
      ii) For edges, movement in a direction won't produce significant changes but since corners are joined by two directions, movement in a direction will cause a significant change to be detected.
   b) i) Brightness constancy: Consistency is where the intensities of an object remain the same across frames.
      ii) Temporal persistence: Persistence is where the motion between frames is small and the camera position changes slowly.
      iii) Spatial coherence: Neighbouring features belong to the same physical surface and have similar motion (common fate)
   c) A feature descriptor "describes" a feature in an image via a characteristic vector and is important because if can be used for feature matching independant of the image (SIFT)
   d) Harris Corner Detector...

4. a) i) The single point on the projected image where the object appears to be coming from. All objects that have the same 3D velocity in space appear to originate from the same point.
      ii) Velocity components: $u = \frac{\delta x}{\delta t}$ ...
   b) i)
         - $D(t)$ is the distance from the FoE to a moving image point.
         - $V(t)$ is the velocity of the point.
         - $z(t)$ is the distance from the image plane.
         - $w(t)$ is the velocity towards the image plane.

         Time to adjacency represents the time it takes the moving object to collide with the image plane.
      ii) The relative depth of two points can be calculated using the TTA ...
   c) The Aperture problem is where the local image intensity gradient is in certain configurations, it is not possible to recover the direction of motion. This problem can be avoided by using the global technique using the square of the error at each pixel (smooth constraint)