1999

pg 12
JGD

Model Answer, Computer Vision, Question 2.

1. Linear operators in computer vision perform pre-conditioning tasks such as filtering and segregation of image structure by scale. There are usually implemented by the convolution integral, or equivalently (because of linearity!) by multiplication in the 2D Fourier domain. An example is to use a bandpass filter to enhance edges in an image, at a certain scale of analysis, prior to detecting and extracting such edges.

   Non-linear operators cannot be inverted, as can linear operators. They involve passing a signal (image) through some fundamental non-linear stage such as a threshold, a logical/Boolean operation, or the computation of a modulus (sum of the squares of the real and imaginary parts of the outcome of a complex-valued filtering step). They are used for detection and classification of structure, and higher-level pattern recognition tasks, rather than front-end conditioning or pre-filtering.

   The fundamental limitation of linear operators is that their outcome is never symbolic nor logical, but rather just another version of input signal. The fundamental limitation of non-linear operators is that none of the powerful tools of linear systems analysis can be applied to them, and so they are often beyond any general theoretical analysis. The benefit gained, at the cost of mathematical clarity or transparency, is that they can sometimes be configured as a kind of "signal-to-symbol converter."

   [8 marks]

2. Shape descriptors such as "codons" or Fourier boundary descriptors encode the properties of a closed 2D shape's boundary over a domain from 0 to 2π (the total perimeter of a closed curve, in radians). This means that the same shape in different sizes would always produce the same code, and so this achieves size invariance. Likewise, the description is always relative to the center of the closed curve, and so it achieves invariance to position in two dimensions. Finally, a rotation of the 2D shape in the plane amounts to merely a scrolling in angle of the code for the shape in terms of the boundary descriptors; and in the case of codons, the lexicon entry for the shape (defined by a grammar of zeroes of curvature and inflexion points) is completely unaffected by rotations.

   The achievement of these sorts of invariances, and indeed of invariance to some non-affine distortions and deformations of the shape, are important steps in pattern recognition and classification. They serve to diminish the unimportant elements of "within-class variability," and to give a compact description that sharpens the "between-class variability." These are central goals for pattern classification.

   [6 marks]

3. Superquadrics are 3D mathematical solids having a low-dimensional parameterization. They represent 3D objects as the unions and intersections of generalized superquadric solids defined by equations of the form:

$$Ax^{\alpha} + By^{\beta} + Cz^{\gamma} = R$$

Examples include "generalized cylinders" and cubes (large exponents); prolate spheroids (footballs) and oblate spheroids (tomatoes), when $\alpha = \beta = \gamma = 2$ and when only two of (A, B, C) are equal to each other.

These simple, parametric descriptions of solids, when augmented by Boolean relations for conjoining them, allow one to generate object-centered, "volumetric" descriptions of the objects in a scene (instead of an image-based description) by just giving a short list of 3D parameters and relations, rather like the codon descriptors for closed 2D shapes.

Their main limitation is that any given superquadric is always a convex shape, and the only way to build more complicated shapes is to conjoin them and this always generates a cusp. Thus the repertoire of 3D solids they can represent is rather limited, and they tend to have a rather puffy appearance. Their main advantage is that a shape is approximated extremely economically by such a short list of functional parameters; this economy can aid in the recognition or classification of the object.

[6 marks]