Lecture 3: Introduction to feature+geometry models

P. Perona
California Institute of Technology

EE/CNS148 - Spring 1999

1 Signal variability as an effect of illumination

Can you prove (using the Lambertian approximation) that for a light source at infinite distance from the scene and no cast shadows the space of possible images is 3-dimensional?

2 Representation error and classification error

Let's briefly go back to the principal components analysis. The representation of the signal is as good as the norm of the first principal values used (see calculations in class).

It has to be remarked that the classification error is not necessarily proportional to the representation error - this will be further explored in Lecture 4.

3 When multidimensional filtering is not sufficient

Consider the example shown in class (see `lecture3.m`). The signal is composed of two 'blips' with variable height and variable mutual position. Observe that when the mutual position of the two blips is highly variable the principal components (the singular values of the matrix composed of examples) decrease very slowly. Correspondingly observe that the first 16 principal vectors are 'meaningful', as opposed to the case where the mutual position of the 'blips' varies little.

It is quite clear that, as the signal deforms more and more, matched filtering works less and less well. There are two problems:

1. The number of principal components required for representing the signal increases more and more.

2. The filters are broad, thus capturing lots of noise, while the useful part of the signal is small.

How could we improve the situation?

4 Features + geometry

First idea: construct independent detectors for the two 'blips'. Two advantages:

1. Small support for the filters. Low computational cost and low amount of noise.
2. Small number of principal components.

How do we use the output of the two detectors:

First idea. Two stages: first detect the blips, then count all the responses within a window of appropriate width. Pro: computationally efficient. Con: does not encode precise mutual location.

Second idea. Model the geometry explicitly with a probability density. Two stages: first detect the features, then evaluate the goodness of the shape. Pro: takes into account the geometry. Con: (a) possibly combinatorial explosion. (b) This is not optimal...

Third idea. Model jointly the detector's output and the geometry. Detect by finding local-maximum hypotheses and then deciding whether they are good enough. Pro: optimal. Con: possibly computationally unfeasible.

Issues:

1. Translation, rotation, scale invariance?
2. Fitting of model parameters?
3. Details of implementation?
4. Automatic model building?