ECSE-626

Statistical Computer Vision

Bayesian Object Recognition

Object Recognition

- We have examined various approaches to the ill-posed problem of object recognition. Most of the approaches were:
  - Not general, easily portable
  - Deterministic
  - Static
- We noted that most approaches seek a single object label.

Object Recognition

- We examine a probabilistic solution to the recognition problem:
  - Sources of uncertainty explicitly represented by probability density functions.
  - Mathematically sound recipe for combining information – don't need to impose heuristics.
  - Subjective priors can be factored – their effect reduced with data collected.
  - Qualification of recognition result.

Bayesian Recognition Problem

- Recognition problem: Given a set of measurements of an unknown object, compute the *degree of likelihood* of it matching each of the objects in a stored database.
- Define the set of K possible object hypotheses:

$$\{O_i\} \mid i = 1..K$$

Bayesian Recognition Problem

- Posterior beliefs over entire set of K object hypotheses, given a set of measurements $\mathbf{d}$ in a scene (i.e. from a camera, laser rangefinder), is represented by a posterior, discrete, conditional probability density function:

Bayesian Recognition Problem

- Posterior beliefs over entire set of *K* object hypotheses, given a set of measurements **d** in a scene (i.e. from a camera, laser rangefinder), is represented by a posterior, discrete (conditional) probability density function:

$$p(O_i \mid \mathbf{d}) \mid_{i = 1..K}$$

Feature-based Recognition

- Consider the problem where objects are represented by models (e.g. parametric models, features) inferred from the measurements.
- Inference still takes the form of describing the belief in a set of object labels.
- There is another level of inference – involving a mapping from model space M and the space of all possible labels O.

Object Space

- O space can be visualized as a separate space, where each point is assigned a particular label i - object corresponding to the label is O_i.
- Alternately, O space can be thought of as simply a subspace of the set of all possible model parameters M – object labels can be linked to a particular set of model parameters.

Object Space

- For the particular case of object recognition, *O* space is discrete, with each point in *O* denoting a particular object in the database.

Inference Chain

- The general inference chain for this type of inverse problem can be represented by:

$$\mathbf{d} \rightarrow \mathbf{m} \rightarrow O_i$$

- This implies a level of inference from raw data, **d**, to model parameters, **m**.
- Another level of inference takes us from parameters **m** to label O_i.

Sources of Information

- Let's examine each of the sources of information available and represent each as probability density functions…

Forward Problem

- We first examine the forward problem – information from physical theory.
- For each object label, O_i, the forward problem consists of predicting the values of the observations.
- In this context, we have an additional level of inference so the forward problem consists of predicting the values of the observed parameters, $\mathbf{m}$, for each model label, O_i.

Forward Problem

- In the case of parametric shape recognition, a measurement produces a vector description of object features.
- The theory is rarely exact due to modelization uncertainties. This information is represented by a conditional probability density function,

representing how the observed parameters **m** vary given a particular object, O_i .

Forward Problem

- In the case of parametric shape recognition, a measurement produces a vector description of object features.
- The theory is rarely exact due to modelization uncertainties. This in formation is represented by a conditional probability density function,

$$p(\mathbf{m} \mid O_i)$$

representing how the observed parameters **m** vary given a particular object, O_i.

Forward Problem

- This information is usually obtained during a training or learning phase of the recognition process.
- The general form of the function over the entire space of objects is:

$$p(\mathbf{m} \mid O)$$

Prior Information

- Information from unspecified sources about the kinds of objects which exist in the world is explicitly represented by the probability density function:

Prior Information

- Information from unspecified sources about the kinds of objects which exist in the world is explicitly represented by the probability density function:

$$p(O)$$

- For the particular case of object recognition, this implies *a priori* belief in each of the objects in the database, $p(O_i)$, i.e. a discrete distribution.

Measurement Information

- In the first stage of inference, data extracted from the scene lead to model descriptors. Estimating the parameters of the underlying model that generated the data set is an ill-posed inverse problem.
- The resulting probability density function over models parameters from a set of measurements is represented by:

Measurement Information

- In the first stage of inference, data extracted from the scene lead to model descriptors. Estimating the parameters of the underlying model that generated the data set is an ill-posed inverse problem.
- The resulting probability density function over models parameters from a set of measurements is represented by:

$$p(\mathbf{m} \mid \mathbf{d})$$

Posterior Information

- The goal of the system is to compute the degree of confidence in a set of object labels, O_i, in a predetermined database, given a set of measurements, **d**.

$$p(O \mid \mathbf{d}) = \frac{p(\mathbf{d} \mid O)\, p(O)}{p(\mathbf{d})}$$

- However, we have an additional level of inference, that inferring model parameters or features from the raw data.

Posterior Information

- As we are interesting in estimating $p(O|\mathbf{d})$, we refer to the model parameter or features, **m**, as nuisance parameters.
- These can be marginalized out as follows:

Posterior Information

- As we are interesting in estimating $p(O|\mathbf{d})$, we refer to the model parameter or features, **m**, as nuisance parameters.
- These can be marginalized out as follows:

$$p(O \mid \mathbf{d}) = \int_M p(O, \mathbf{m} \mid \mathbf{d}) d\mathbf{m}$$

Posterior Information

- We wish to simplify this further given our sources of information.
- We wish to express the posterior in terms of the measurement information, for example.
- How can we split the joint density function into components that we know?

Posterior Information

- We wish to simplify this further given our sources of information.
- We wish to express the posterior in terms of the measurement information, for example.
- How can we split the joint density function into components that we know?

$$p(O \mid \mathbf{d}) = \int_M p(O \mid \mathbf{m}, \mathbf{d})\, p(\mathbf{m} \mid \mathbf{d})\, d\mathbf{m}$$

Posterior Information

- Can this be simplified further?
- We postulate that **m** is a sufficient statistic for *O*. This simplifies the term:

$$p(O \mid \mathbf{m}, \mathbf{d}) = ?$$

Posterior Information

- Can this be simplified further?
- We postulate that **m** is a sufficient statistic for *O*. This simplifies the term:

$$p(O \mid \mathbf{m}, \mathbf{d}) = p(O \mid \mathbf{m})$$

Posterior Information

Thus the posterior reduces to:

$$p(O \mid \mathbf{d}) = \int_M p(O \mid \mathbf{m})\, p(\mathbf{m} \mid \mathbf{d})\, d\mathbf{m}$$

Given the information that we have, we can invoke Bayes' law further:

Posterior Information

$$p(O\,|\,\mathbf{d}) = \int_M \frac{p(\mathbf{m}\,|\,O)\,p(O)\,p(\mathbf{m}\,|\,\mathbf{d})}{p(\mathbf{m})}d\mathbf{m}$$

$$= p(O)\int_M \frac{p(\mathbf{m}\,|\,O)\,p(\mathbf{m}\,|\,\mathbf{d})}{p(\mathbf{m})}d\mathbf{m}.$$

Posterior Information

Should the posterior information from measurements, $p(\mathbf{m} \mid \mathbf{d})$, not be available, the solution can be equivalently expressed in terms of the physical theory or forward solution for measurements, $p(\mathbf{d} \mid \mathbf{m})$:

Posterior Information

Should the posterior information from measurements, $p(\mathbf{m}\,|\,\mathbf{d}),$ not be available, the solution can be equivalently expressed in terms of the physical theory or forward solution for measurements, $p(\mathbf{d}\,|\,\mathbf{m})$:

$$p(O\,|\,\mathbf{d}) = \frac{p(O)}{p(\mathbf{d})}\int_M p(\mathbf{m}\,|\,O)\,p(\mathbf{d}\,|\,\mathbf{m})d\mathbf{m}$$

where $p(\mathbf{d})$ is a constant of proportionality such that:

$$p(O\,|\,\mathbf{d}) \propto p(O)\int_M p(\mathbf{m}\,|\,O)\,p(\mathbf{d}\,|\,\mathbf{m})d\mathbf{m}.$$

Evidence

- The *evidence* for the hypotheses is determined by the $p(\mathbf{d}|O)$ term.
- In this formulation the evidence refers to:

Evidence

- The *evidence* for the hypotheses is determined by the $p(\mathbf{d}|O)$ term.
- In this formulation the evidence refers to:

$$p(\mathbf{d} \mid O) = \int_M p(\mathbf{m} \mid O)\, p(\mathbf{d} \mid \mathbf{m})\, d\mathbf{m}.$$

- The evidence is the Bayesian transportable quantity for comparing alternate models.

Bayesian Solution to Object Recognition

- This equation is a Bayesian solution to the object recognition problem. It's form is general and can be applied to inverse problems where the posterior belief in model hypotheses is desired.
- Rather than a solution in the form of a single object identity, the system produces a probability density function describing the likelihood of the various objects in the database having led to the measurements.

Bayesian Solution to Object Recognition

- Given the ill-posedness of the problem, the importance of this result is that the *existence* of a solution is guaranteed should $p(O|\mathbf{d})$ not be identically null.
- The solution $p(O|\mathbf{d})$ is *unique*.
- In this sense, the ill-posedness of the problem can be considered alleviated.
- The shape of the distribution determines the ill-posedess of the recognition result in the classic sense (e.g. many peaks).

Sequential Recognition

- Most solutions in the literature are *static* – base interpretation on a single data set.
- Difficulty is that, even with strong *a priori* knowledge, there are still ambiguous cases where a "significant" belief in more than one model exists.
- This is further complicated by noise and quantization error inherent to most practical algorithms.

Sequential Recognition

- Even with direct surface measurements provided by a laser rangefinder, ambiguities still exist as a result of:
  - Occlusion
  - Projective singularities
  - Errors in parametrization and modeling
- Robustness should increase if decisions could be deferred until a sufficient confidence level is established.

Sequential Recognition

- How do we accumulate evidence for each model hypothesis, O_i, over a sequence of images?
- Evidence is in the form of a conditional probability density function. Let $\mathbf{d}_t$ denote the data acquired at time t.

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) = ?$$

Sequential Recognition

- How do we accumulate evidence for each model hypothesis, O_i, over a sequence of images?
- Evidence is in the form of a conditional probability density function. Let $\mathbf{d}_t$ denote the data acquired at time t.

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) = \frac{p(\mathbf{d}_t, \mathbf{d}_{t+1} \mid O)\, p(O)}{p(\mathbf{d}_t, \mathbf{d}_{t+1})}$$

Sequential Recognition

- In order to obtain an inexpensive solution to this problem, we make the following assumption:

Data acquired at time t is statistically independent of data sets acquired at other times.

$$p(\mathbf{d}_{t+1} \mid \mathbf{d}_t) = p(\mathbf{d}_{t+1})$$

Sequential Recognition

- Based on this assumption, information can easily be merged at the level of probabilities by using a recursive Bayesian chaining strategy.

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) = \frac{p(\mathbf{d}_t, \mathbf{d}_{t+1} \mid O)\, p(O)}{p(\mathbf{d}_t, \mathbf{d}_{t+1})},$$

$$= \frac{p(\mathbf{d}_t \mid O)\, p(\mathbf{d}_{t+1} \mid O)\, p(O),}{p(\mathbf{d}_t)\, p(\mathbf{d}_{t+1})}$$

$$= \frac{p(\mathbf{d}_{t+1} \mid O)}{p(\mathbf{d}_{t+1})}\, p(O \mid \mathbf{d}_t)$$

Sequential Recognition

- Thus the posterior at time t, $p(O|\mathbf{d}_t)$ can be fed back to the system as the prior at time $t+1$.
- Can we derive what this implies in our context?

Sequential Recognition

$$\frac{p(\mathbf{d}_{t+1} \mid O)}{p(\mathbf{d}_{t+1})} = \frac{p(O \mid \mathbf{d}_{t+1})}{p(O)}$$

In our context, this gives:

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) = \frac{p(O \mid \mathbf{d}_t)}{p(O)} \, p(O \mid \mathbf{d}_{t+1}).$$

We know that:

$$p(O \mid \mathbf{d}_{t+1}) \propto p(O) \int_M p(\mathbf{m} \mid O) \, p(\mathbf{d}_{t+1} \mid \mathbf{m}) \, d\mathbf{m}.$$

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) \propto \frac{p(O \mid \mathbf{d}_t)}{p(O)} \, p(O) \int_M p(\mathbf{m} \mid O) \, p(\mathbf{d}_{t+1} \mid \mathbf{m}) \, d\mathbf{m},$$

Sequential Recognition

$$p(O \mid \mathbf{d}_t, \mathbf{d}_{t+1}) \propto p(O \mid \mathbf{d}_t) \int_M p(\mathbf{m} \mid O)\, p(\mathbf{d}_{t+1} \mid \mathbf{m})\, d\mathbf{m}$$

Sequential Recognition

- By propagating evidence in this manner, evidence in the true hypothesis should grow over a short number of views, while the confidence in the others should decline.
- By quantifying the level of confidence in the various hypotheses at each stage, an active agent can gather evidence until the composite belief associated with a particular hypothesis exceeds a prescribed level of merit or until a clear winner emerges.