## Contents

8 Scene Representation ........................................ 1
  8.1 Model-Based Object Representation .................... 2
    8.1.1 Recognition Paradigm .............................. 3
    8.1.2 Feature Matching .................................. 3
  8.2 Graph Representation ................................... 4
    8.2.1 Graph Matching .................................... 4
    8.2.2 Aspect Graphs .................................... 5
List of Figures

8.1 Recognition paradigm .............................................. 3
8.2 Feature matching ................................................ 4
8.3 Breath-first search of a match tree. .......................... 4
8.4 Graph Matching ................................................... 5
8.5 Different views of an object. .................................... 6
8.6 Different views of a L-shaped object. ........................ 6
8.7 Aspect graph of a L-shaped object. ............................. 7
Chapter 8

Scene Representation

In this chapter, we consider approaches that aim to analyze scenes in an image. Scene analysis is the ability to make useful descriptions of the given scene.

- Identify objects and relationships between them
- Locate a certain object
- Make measurements for inspection

This is a very ambitious task and to date, it still remains unsolved. Issues that make this problem difficult are:

- Representation
- Occlusion
- Image complexity
- Scene complexity
- A priori knowledge is essential

Representation is a difficult issue since it must incorporate the following

- Levels of abstraction
  - Specific objects
  - Generic objects
  - Functional representations

- Properties of objects: Shape, color, texture, function

- Shape representation: Collection of points, boundaries, surfaces, volumes, collection of parts
Projection and occlusion also create difficulties. Although scenes are 3-D, images are 2-D. Projected shapes depend on the viewpoint. Objects occlude each other and hence only parts of objects are visible. This causes problems in segmentation.

Images are complex. Scenes are also complex due to:
- Large number of objects
- Occlusion
- Similar objects
- Complex shapes
- Complex patterns
- Illumination

World knowledge is often required to disambiguate the incoming information. Furthermore, as objects are often in specific relationship to other objects, this information is useful in analysis. The general approach is model-based object recognition.

8.1 Model-Based Object Representation

The general approach to model-based object recognition consists of the following stages:
- Choose a representation model
- Find instances of known models in the scene
- Estimate their pose

The complexity of the task depends on:
- The number of models in the database
- Complexity of the image
- Complexity of the scene

In recognition based on matching models, the following steps are followed:
- Template matching
  - Correlate a template with the image
  - Invariant to translation, sensitive to rotation, scaling and viewpoint
  - Useful for 2D binary patterns (character recognition)
- Feature vectors
  - Describe by features – Area, moments, Fourier coefficients
  - Solve segmentation
  - Viewpoint invariance is hard
- Features and the relationships between them
8.1.1 Recognition Paradigm

The recognition paradigm is as follows:

- Could the image be a projection of the model?
  - Match part of a scene with a model → Search for best fit
  - Estimate object pose
  - Project model and verify

- What types of features to match and how?
  - Feature matching
  - Structured description matching

![Recognition Paradigm Diagram](image)

Figure 8.1 Recognition paradigm

8.1.2 Feature Matching

In feature matching, the goal is to match a given model to the data contained in the image.

- Match image and model features → Search for best fit
- Check for consistency using constraints that depend on representation

One of the difficulties arose due to the fact that multiple matches are possible. Hence the search procedure embeds a breadth-first search over a tree where each node of the tree corresponds to a match.
8.2 Graph Representation

One approach to structural shape representation is graph representation – as used in robot path planning. A graph consists of two entities: nodes $P$ and edges $E(P)$ between nodes. Here, the model is based on a graph that represents the objects’ parts and their relationships. Given an incoming image, a similar model is constructed and matched with those in the model database. Some examples are:

- Line drawings consisting of vertices (nodes) and lines (edges)
- Surfaces (nodes) and their relationships (edges)
- Complex shape represented by parts (nodes) and the relationships (edges)

8.2.1 Graph Matching

Given $G = (P, E(P), R(P))$ and $G' = (P', E'(P'), R'(P'))$
8.2. GRAPH REPRESENTATION

- Two nodes \( n \in P \) and \( n' \in P' \) form an assignment if \( R(n) = R'(n') \)
- Two assignments \((n_1, n'_1)\) and \((n_2, n'_2)\) are compatible if \( E(n_1, n_2) = E(n'_1, n'_2) \)
- Isomorphism: Compatibility for all nodes
- Sub-graph Isomorphism: Compability for a subset of nodes

Finding an isomorphism is NP-hard - time required is exponential with the size of the nodes. The usual approach is to conduct a tree search. Furthermore isomorphism may be too strict of a criterion.

![Figure 8.4 Graph Matching](image)

8.2.2 Aspect Graphs

One approach to recognizing objects is to obtain a series of two-dimensional views of a known object, maintain them in some convenient representation in storage, and then match one or more two-dimensional views of an unknown object against the stored views of the known object, thereby reducing the recognition problem to a series of two-dimensional matching problems.
Figure 8.5 Different views of an object.

Figure 8.6 Different views of a L-shaped object.
Figure 8.7 Aspect graph of an L-shaped object.
Bibliography
