Chapter 19 Knowledge-Based Vision

19.1 Introduction

knowledge-based vision system: uses domain knowledge to analyze images
knowledge: might be very general about 3D objects or extremely specific
urban scene knowledge: appearance of houses, roads, office buildings
airport scene knowledge: runways, terminals, hangars

19.2 Knowledge Representations

knowledge representation structures:

- feature vectors
- relational structures
- hierarchical structures
- rules
- frames

19.2.1 Feature Vectors

feature vector: simplest form of knowledge representation in computer vision
feature vector: tuple of measurements of features with numeric values

feature vectors: attribute-value tables: property lists
feature-vector representation: uses global features of object
some useful features in character recognition:

- number of straight strokes
- number of loops
- width-to-height ratio of character

two hand-printed characters and their feature vectors
=====Fig. 19.1=====

19.2.2 Relational Structures

complex objects and scenes: often composed of recognizable parts
full description of complex entity consists of:

1. global features
2. global features of each of its parts
3. relationships among parts

global attributes to represent line drawing extracted from gray tone image:

- total number of line segments

- density of line segments (average number of segments per unit area)
- size (number of rows and number of columns)

for each line segment:

- start coordinates: Row_Start, Col_Start
- endpoint coordinates: Row_End, Col_End
- length: Length
- angle: Angle

three relationships to define perceptual groupings of line segments:

- proximity
- parallelism
- collinearity

relational description of a simple line drawing
=====Fig. 19.2=====

19.2.3 Hierarchical Structures

hierarchical structure represents better: entity broken into parts recursively
hierarchical structures have: both hierarchical and relational component

HD: Hierarchical Description
LEFT, ABOVE: relational description
atomic parts: can be broken down no further
hierarchical, relational description of an outdoor scene
=====Fig. 19.3=====

19.2.4 Rules
rule-based systems: encode knowledge in form of rules
$< antecedents >$: list of conditions must be met to invoke rule
$< actions >$: set of operations or actions performed when rule invoked
rule has the form:

$$< antecedents > \rightarrow < actions >$$

production associated with context-free grammars:

$$< expression > \rightarrow < expression > + < term >$$

=====picture grammars for submedian chromosome=====
shape of chromosome that can be generated from given rules
=====Fig. 19.4=====

19.2.5 Frames and Schemas
frame: data-structure for representing stereotyped situation

relational structure whose terminal nodes consist of:

- slots: attributes
- fillers: values for those attributes

schema in database terminology: model or prototype
frame describing generalized cylinder model of electric motor:
NODE: ELECTRIC_MOTOR_CONE

CLASS:	SIMPLE_CONE
SPINE:	Z0014
SWEEPING_RULE:	CONSTANT_SWEEPING_RULE
CROSS_SECTION:	Z0013

19.3 Control Strategies

control strategy of system: dictates how knowledge will be used
KS: Knowledge Source
four kinds of control used in machine vision systems
=====Fig. 19.5=====

19.3.1 Hierarchical Control

hierarchical control: most common control structure in computer programming
bottom-up control scenario:

- preprocessing routines to convert original image to extract primitives
- feature extraction: locates features of interest
- decision-making procedure: performs some recognition task

hybrid control with feedback: neither bottom-up nor top-down control

19.3.2 Heterarchical Control

heterarchical control: lets data themselves dictate order of operations
knowledge sources: knowledge embodied in set of procedures
blackboard approach: tries to add some order to heterarchy
blackboard: global database shared by set of independent knowledge sources
blackboard scheduler: controls access to blackboard by knowledge sources
blackboard scheduler: decides execution order of competing knowledge sources
=====Garfield 17:66======

19.4 Information Integration

hypothesis: proposition or statement either true or false
hypothesize-and-test paradigm: commonly used within control structures
hypothesis: generated on the basis of some initial evidence
number indicating certainty that hypothesis is true
two main approaches to information integration problem:

- Bayesian belief network
- Dempster-Shafer theory of evidence

19.4.1 Bayesian Approach

Bayesian belief network: directed acyclic graph
Bayesian belief network: with nodes representing propositional variables
Bayesian belief network: with arcs representing causal relationships
variables $X_1, X_2, ..., X_N$: nodes of graph
finite set of possible values $\{x_i\}$: for each propositional variable X_i
arc (X_j, X_i) from node X_j to X_i: X_j directly causes X_i
conditional probability $P(x_i|x_j)$: strength of causality
portion of belief network to be used in determining object identities
=====Fig. 19.6=====

$BEL(x)$: belief value
$\pi_X(u)$: current strength of causal support by incoming link to X
$\lambda_Y(x)$: current strength of diagnostic support by outgoing link from X
propagation of beliefs through a piece of a belief network
=====Fig. 19.7=====

belief updating formula:

$$BEL(x) = \alpha\lambda_Y(x)\lambda_Z(x)\sum_{u,v} P(x|u,v)\pi_X(u)\pi_X(v)$$

assume causal support functions:
$\pi_T(desk) = .9$
$\pi_T(workstation) = .4$
$\pi_T(pushbuttons) = .6$
$\pi_T(keyboard) = .5$
=====BEL(telephone),BEL(terminal)=====
=====sea dragon=====
=====body paint=====

19.4.2 Dempster-Shafer Theory
Bayesian model: allows positive belief in proposition but not disbelief
Dempster-Shafer theory: information integration allowing belief and disbelief

Representation of Belief
belief interval: represents belief in proposition
unit interval $[0, 1]$: further demarcated by j, k where $k \geq j$
$Belief(A)$: subinterval $[0, j)$
$Belief(A)$: degree to which current evidence supports proposition A

$Disbelief(A)$: subinterval $(k, 1]$
$Disbelief(A)$: degree to which current evidence supports negation of A
$Uncertainty(A)$: subinterval $[j, k]$
$Uncertainty(A)$: degree to which we believe nothing + or − about A
plausibility of A: sum of $Belief(A)$ and $Uncertainty(A)$
$Doubt(A)$: sum of $Disbelief(A)$ and $Uncertainty(A)$
belief interval
=====Fig. 19.8=====

frame of discernment Θ: set of mutually exclusive atomic proposition, one true
ϕ: represents the empty set
$m(A)$: degree of belief committed exactly to subset A by function m
basic probability assignment: function m defined on 2^{Θ} satisfying

1. $0 \leq m(A) \leq 1$, for every $A \subseteq \Theta$
2. $m(\phi) = 0$
3. $\Sigma_{A \subseteq \Theta} m(A) = 1$

belief function Bel over a frame of discernment Θ: assigns belief
Bel must satisfy the following conditions:

1. $Bel(\phi) = 0$

2. $Bel(\Theta) = 1$

3. $A \subseteq B \subset \Theta$ implies $Bel(A) \leq Bel(B)$

belief functions: easily constructed from basic probability assignment

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

knowledge sources KS's: entities producing evidence for or against proposition

Θ: set of all possible objects found on image
KS_1 computes the following probabilities:
$m_1(\{CHAIR\}) = .3$
$m_1(\{TABLE\}) = .1$
$m_1(\{DESK\}) = .1$
$m_1(\{WINDOW\}) = .15$
$m_1(\{PERSON\}) = .05$
$m_1(\{\Theta\}) = .3$
$m_1(\{\Theta\}) = .3$: probability .3 that something occurred but doesn't know what
second knowledge source KS_2: shares same frame of discernment as KS_1
$m_2(\{CHAIR\}) = .2$
$m_2(\{TABLE\}) = .05$
$m_2(\{DESK\}) = .25$

$m_2(\{WINDOW\}) = .1$
$m_2(\{PERSON\}) = .2$
$m_2(\{\Theta\}) = .2$

Belief Combination

Dempster's rule of combination defines function m_{12}:

$$m_{12}(C) = \frac{\Sigma_{A\subseteq\Theta, B\subseteq\Theta, A\cap B=C}\, m_1(A)m_2(B)}{\Sigma_{A\subseteq\Theta, B\subseteq\Theta, A\cap B\neq\phi}\, m_1(A)m_2(B)}$$

$m_{12}(\{CHAIR\})$: combined probability assigned to label $CHAIR$:

$$m_{12}(\{CHAIR\})$$

$$= \frac{m_1(\{CHAIR\})m_2(\{CHAIR\}) + m_1(\{CHAIR\})m_2(\{\Theta\}) + m_1(\{\Theta\})m_2(\{CHAIR\})}{\Sigma_{A\subseteq\Theta, B\subseteq\Theta, A\cap B\neq\phi}\, m_1(A)m_2(B)}$$

$$= \frac{(.2)(.3) + (.2)(.3) + (.2)(.3)}{\begin{array}{c}(.3)(.2) + (.3)(.2) + (.1)(.05) + (.1)(.2) + (.1)(.25) + \\ (.1)(.2) + (.15)(.1) + (.15)(.2) + (.05)(.2) + (.05)(.2) + (.3)(.2 + .05 + .25 + .1 + .2 + .2)\end{array}}$$

$$= \frac{.18}{.555} = .32$$

how beliefs from KS_1 and KS_2 are combined by Dempster's rule

=====Fig. 19.9=====

A, B, C: three exhaustive and mutually exclusive propositions
KS_1, KS_2: two knowledge sources capable of contributing belief
belief intervals from both KS's
=====Fig. 19.10=====
belief in proposition A
=====Fig. 19.11=====
result of combined m-values
=====Fig. 19.12=====
total areas renormalized to ignore useless areas (each item divided by .781)
=====Fig. 19.13=====
=====Oldie 33:118=====

19.5 A Probabilistic Basis for Evidential Reasoning

random proposition has two states: assertable and not assertable
operational probabilistic model: generalized Shafer's and Bayesian approaches

19.5.1 Legal Court Paradigm

h: argument or question
A: attorney paid to uncover, and present all supportive evidence of h
B: attorney paid to uncover, and present all supportive evidence of $\bar{h}$

theory for evidential reasoning following pattern of legal paradigm:

1. each piece of evidence: proposition
2. question to be decided: proposition
3. proposition in evidence body has measure of assertability
4. logic calculus for inferring propositions
5. belief calculus for computing degree of belief for each proposition

19.5.2 Degree of Belief as Chance Probability of Being Inferred

probability: measure of belief of proposition

19.5.3 Belief Calculus

belief calculus: set of rules to compute degree of belief in proposition
degree of belief in inferred proposition: conditional probability of inference

19.5.4 Examples

A: state "asserted"
NA: state "not asserted" different from state "denied"
proposition denied: whose negation is asserted

q and $q \to r$: propositions
four possible states of assertion:

q	$q \to r$	
A	A	$\Rightarrow r$
A	NA	
NA	A	
NA	NA	

question is proposition $q \vee r$:

q	r	
A	A	$\Rightarrow q \vee r$
A	NA	$\Rightarrow q \vee r$
NA	A	$\Rightarrow q \vee r$
NA	NA	

body of evidence consists of $q, r, q \to \bar{r}$, question is $q \vee r$:

$$
\begin{array}{llll}
q & r & q \to \bar{r} & \\
A & A & A & \Rightarrow false \\
A & A & NA & \Rightarrow q \vee r \\
A & NA & A & \Rightarrow q \vee r \\
A & NA & NA & \Rightarrow q \vee r \\
NA & A & A & \Rightarrow q \vee r \\
NA & A & NA & \Rightarrow q \vee r \\
NA & NA & A & \\
NA & NA & NA &
\end{array}
$$

19.6 Example Systems

applications of these systems include

- aerial image analysis
- bin picking
- automatic mail sorting

19.6.1 VISIONS

VISIONS: Visual Integration by Semantic Interpretation of Natural Scenes

VISIONS: has hierarchical strategy to control employment of knowledge source
VISIONS: uses both declarative and procedural forms of knowledge
VISIONS: provides bottom-up and top-down paths for hypothesis development
3D shapes: represented in terms of B splines and Coon's surface patches

19.6.2 ACRONYM

ACRONYM system: uses symbolic reasoning to aid in analyzing scenes
ACRONYM: performs iterations of prediction, description, and interpretation
ACRONYM: uses stored models whose primitives are generalized cones
ACRONYM: uses constraints on size, structure, spatial relationships
object graph: hierarchical structure, representing objects from coarse to fine
restriction graph: also a hierarchical structure
constraints: stored in restriction graph
observation graph: observables and their relationships
ACRONYM: important for its intelligent use of models, constraints, feedback

19.6.3 SPAM

SPAM: uses map and domain-specific knowledge to interpret airport scenes
SPAM has three main components:

- image/map database: contains facts about airport

- set of image-processing tools: segmentation, linear feature extraction...
- rule-based system

rules: manipulate 3 kinds of primitives: region, fragments, functional areas
regions: classified as linear, compact, small blob, large blob
linear regions: may be runway, taxiway, access road
compact regions: may be terminal buildings or hangars
small-blob regions: may be parking lots or parking aprons
large-blob regions: may be tarmacadam or grassy areas

build phase: selects regions and creates fragment interpretation
local evaluation phase: processes fragment interpretations
consistency phase: checks consistency of interpretations of neighborhood
functional area phase: tries to find functional areas
global evaluation phase: find complete airport interpretation

19.6.4 MOSAIC

MOSAIC: to build and incrementally improve 3D model of urban scene
MOSAIC: uses sequence of both monocular views and stereo pairs
three main tasks executed by the system:

- stereo analysis

- monocular analysis
- construction and modification of scene model

3D wire frame: from ground-plane knowledge, camera geometry, heuristics
structure graph: relational structure to represent current 3D scene model
topological primitives: faces, edges, vertices, objects, edge groups
geometric primitives: planes, lines, points
primitives confirmed: if they come from one sequence of input images
primitives unconfirmed: if they are only hypothesized
knowledge-based system: can be quite powerful in very limited domain
=====joke=====

19.6.5 Mulgaonkar's Hypothesis-Based Reasoning

hypothesis inconsistent: if two inference engines compute incompatible values
system works with: lines, points, arcs, planes

19.6.6 SCERPO

SCERPO: Spatial Correspondence, Evidential Reasoning, and Perceptual Organization
SCERPO: performs 3D object recognition from single 2D images
SCERPO's models: 3D wire-frame objects

three relationships among line segments: proximity, parallelism, collinearity

19.6.7 Ikeuchi's Model-Based Bin-Picking System

Ikeuchi's model-based vision system: for bin-picking using range data
compile mode: geometric modeler used to generate views of object
attitude group: views with identical vectors
attitude group: represented by representative attitude
work model consisting of attribute set:

- inertial moment of faces
- relative positions of faces
- face shapes
- edge information
- extended Gaussian image

interpretation tree: a decision-tree classifier
run mode: system generates 3 edge maps, 2 needle maps, and 1 depth map

19.6.8 Jain's Evidence-Based Recognition System

3D object recognition system: uses range data for input

surface patches: classified as planar, convex, concave
jump edges: have discontinuous depth
crease edges: have continuous depth and discontinuous normals
morphological information: global information concerning entire 3D shape
morphological information: includes object perimeter, number of regions
patch information: local information about each primitive
patch information: patch size, piecewise linear fit of boundary
relational information: describes relationships between pairs of patches
relational information: includes normal angle between patches, jump gap

19.6.9 ABLS

ABLS system: locates address block on mail pieces
knowledge sources: organized in dependency graph
each knowledge source in system: assigned a utility value
utility of knowledge source: based on efficiency, effectiveness, processing
evidence combination: performed via Dempster-Shafer theory
besides address block: return address, stamp, extraneous printing and graphics
ABLS: quality of segmentation strongly affects the results

19.7 Summary

knowledge methodology to integrate all levels of vision: important direction
migrant mother: more to our interpretation than geometrical properties
evoking emotions of understanding, sympathy, anguish, hope, despair
=====Nalwa, *A Guided Tour of Computer Vision*, Fig. 10.1=====
=====joke=====