Systems Engineering for Computer Vision - Winter 2011/12

Bernstein Focus: Neurotechnology Frankfurt

LECUTRE 6: BAYESIAN NETWORKS FOR VISION – (MODEL-B ASED INTERPRETATION, W.MANN (1995))

PREPARED BY
PROF. DR. VISVANATHAN RAMESH

(VERSION 1.0)

Objective of this Lecture

- Bayesian Networks for Vision (Discuss W. Mann's thesis)
- Discuss relationships between Mann, 1995 to Ramesh, 1995

Why is Interpretation Hard?

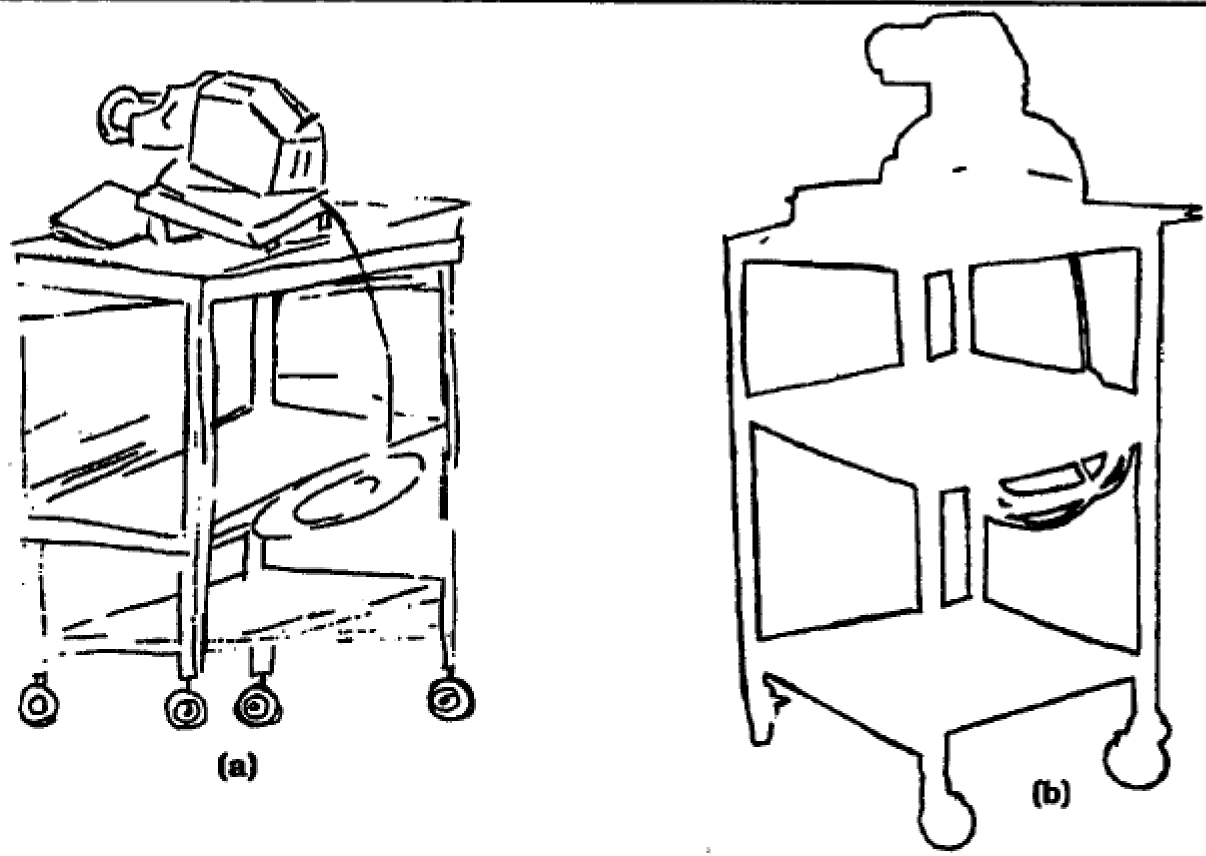


Figure 1-1: Drawing Interpretation vs. Reality

As a beginner, the artist drew figure (a) focusing his attention on what he knew the object to be. Notice that he drew four wheels even though only three were visible. After training to "see as an artist", the same artist drew figure (b) more true to how it actually appears. ([Edwards 1994], figures used by permission of the author)

Goal of Mann's Thesis

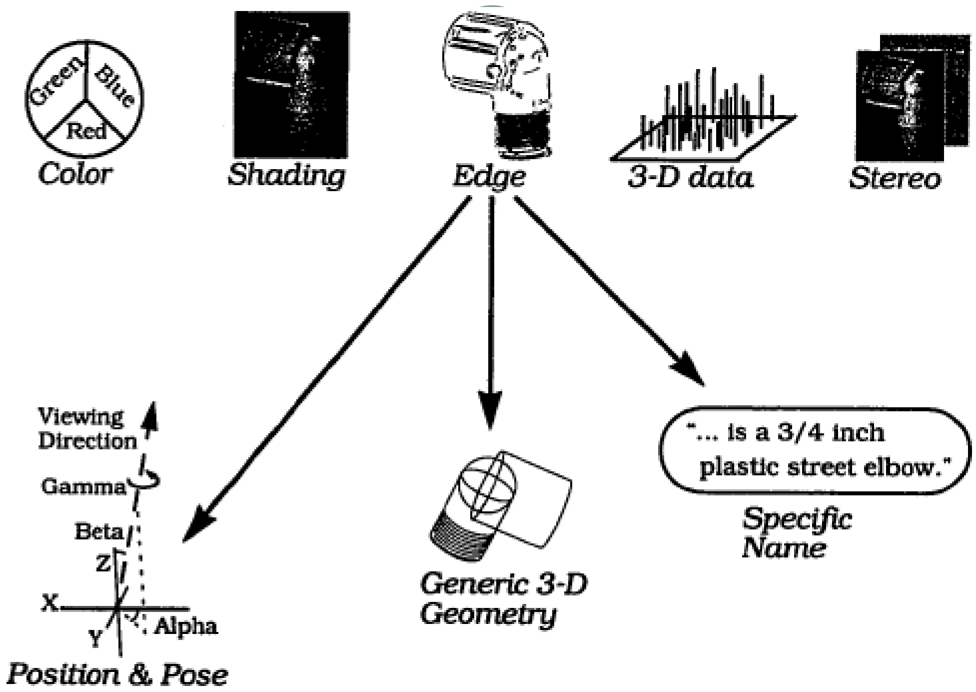


Figure 1-2: Problem Statement for Thesis

Given

- one single image (not stereo)
- with grey scale data (not color)
- of relatively simple real world objects
- using edge information (ignoring shading)
- and assuming orthographic projection (not perspective)

Find

- the geometry of the scene
- the name of the object (if known)
- and the viewing direction.

Random Process of ObjectEdge Fragments

Figure 1-5: Computer's View of the Elbow Image

These are the edges of the elbow displayed with random position and orientation. Human observers take for granted the structure and relationships they see in an image. Without parallelism, proximity, coincidence, colinearity and other relationships, the edges appear as pictured here. Until these relationships are specifically identified in the data, this is the view a simple computer program has of the information. And even in this image we already have the structure of edgels grouped into linked edges.

Object’s Edges

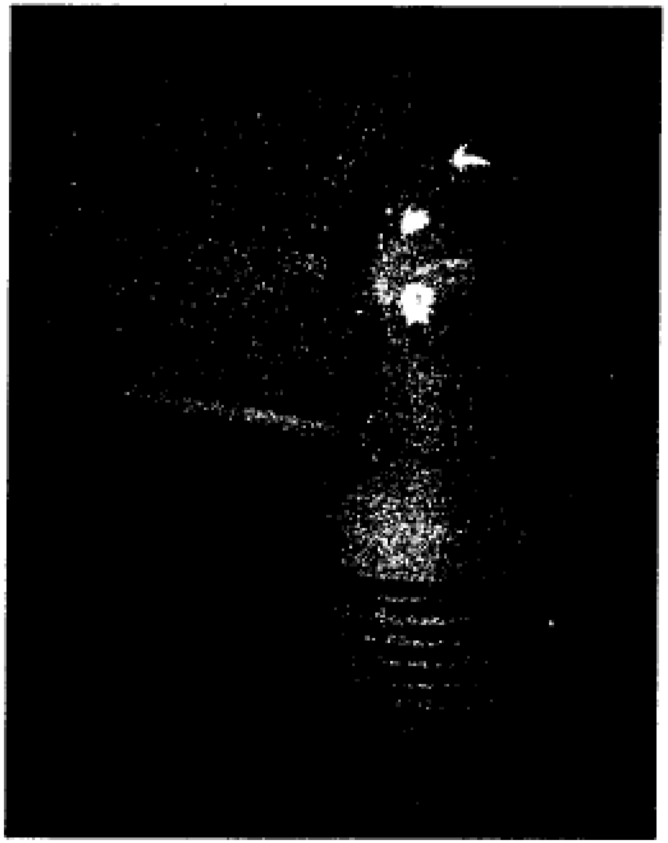


Figure 1-3: Elbow Image

The elbow is a plastic plumbing part.

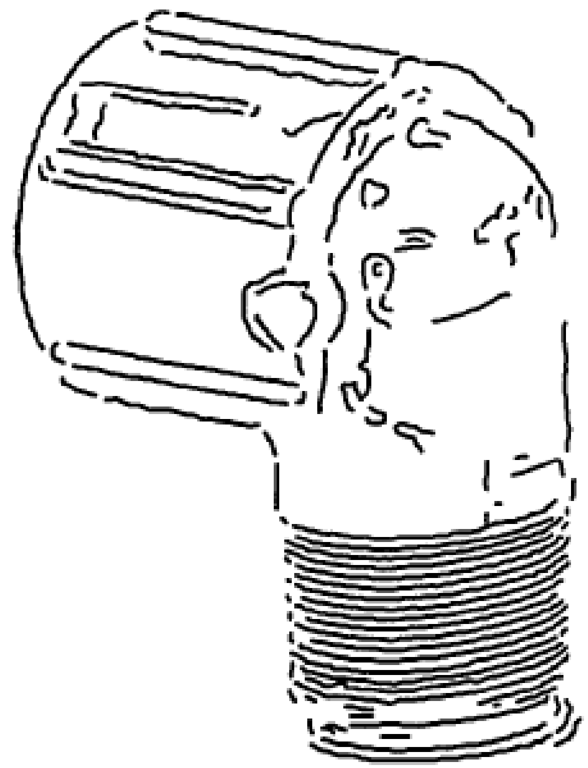


Figure 1-4: Edge Image of the Elbow

Using the Wang-Binford edge detector.

Difficulties in Grouping/Indexing

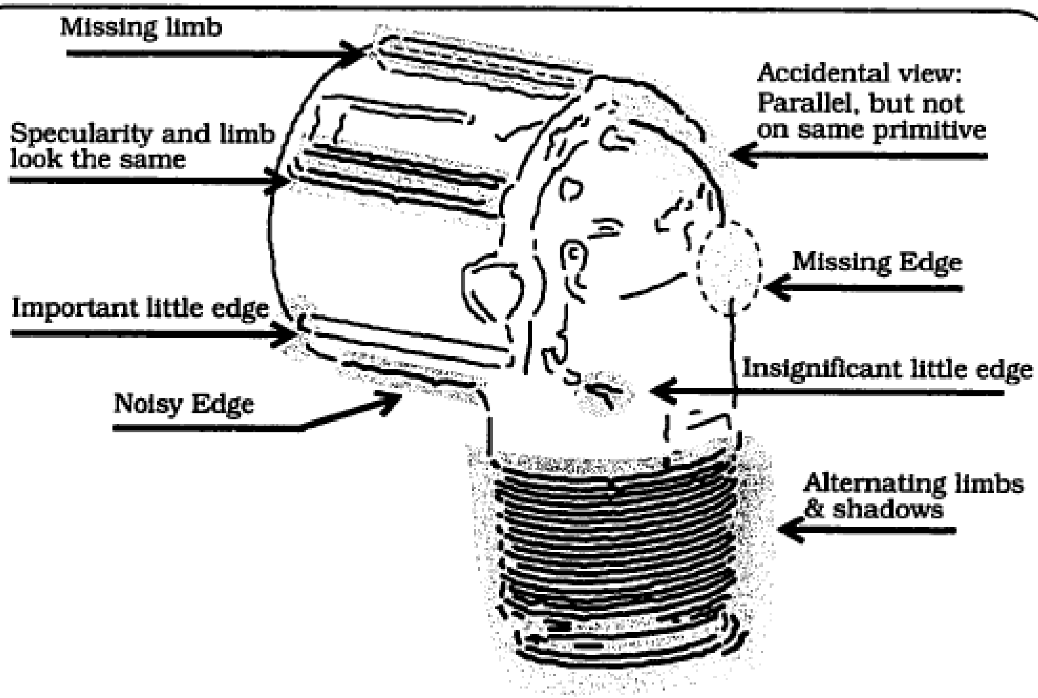


Figure 1-6: Some Difficulties in Choosing Significant Edges

An important limb is occluded by the rib; an important edge was not detected because of low contrast. Parallelism deceives where two curves that are parallel by accident lie on different primitives and the parallelism of limbs and specularities is indistinguishable. One tiny edge is the result of noise; the other is a significant termination to a cylindrical rib. An edge with high noise is a significant limb to the female part, so noise alone can not be used as a criterion to ignore an edge. And what looks like an obvious series of edges from the threads turns out to be an alternating series of limb edges and shadow edges.

Edge Image Information Content

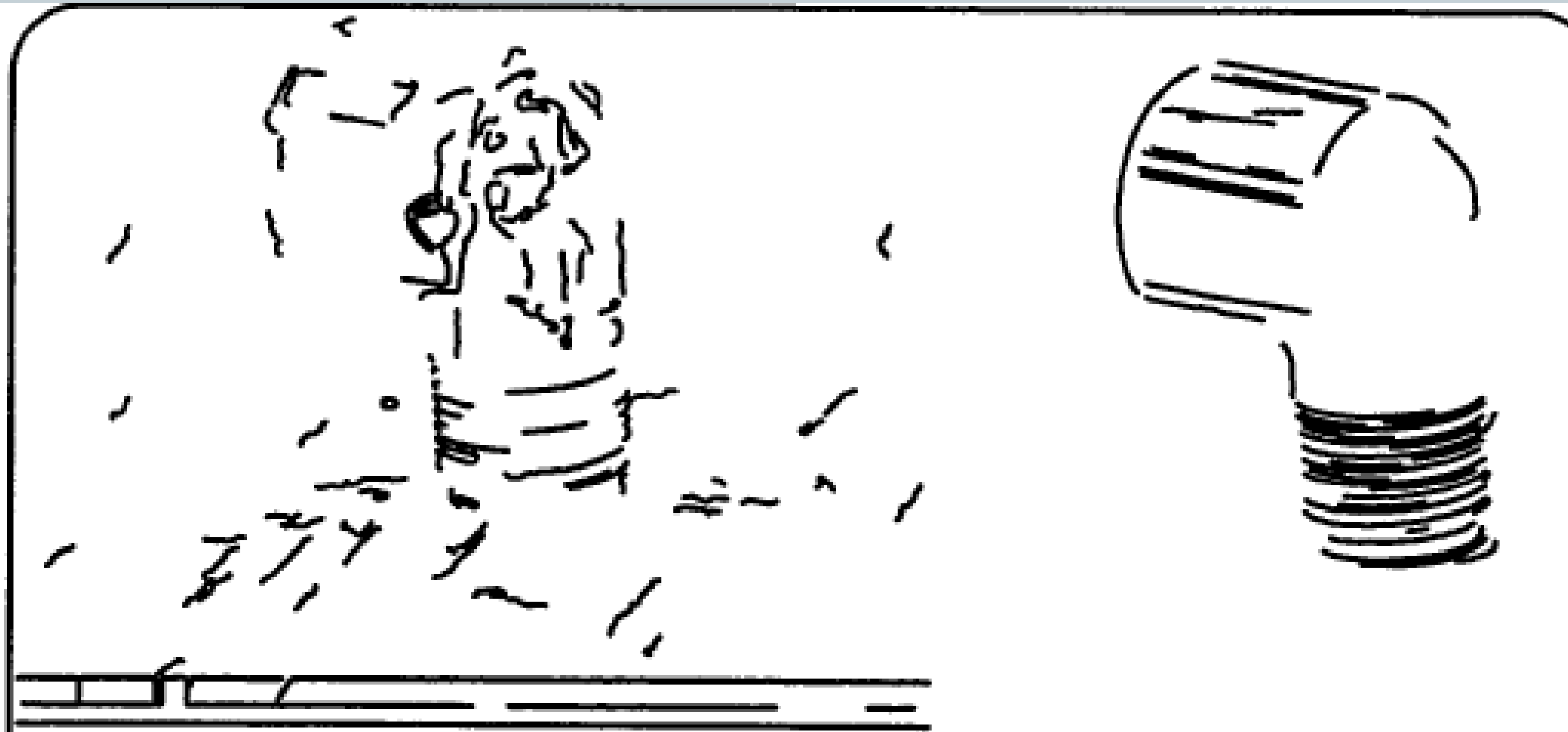


Figure 1-7: Edge Image Information Content

The image on the left contains 2/3 of the edge information, yet the elbow is still hard to see. With only 1/3 of the information, the elbow is seen clearly on the right.

Object Centered vs View Centered Representations

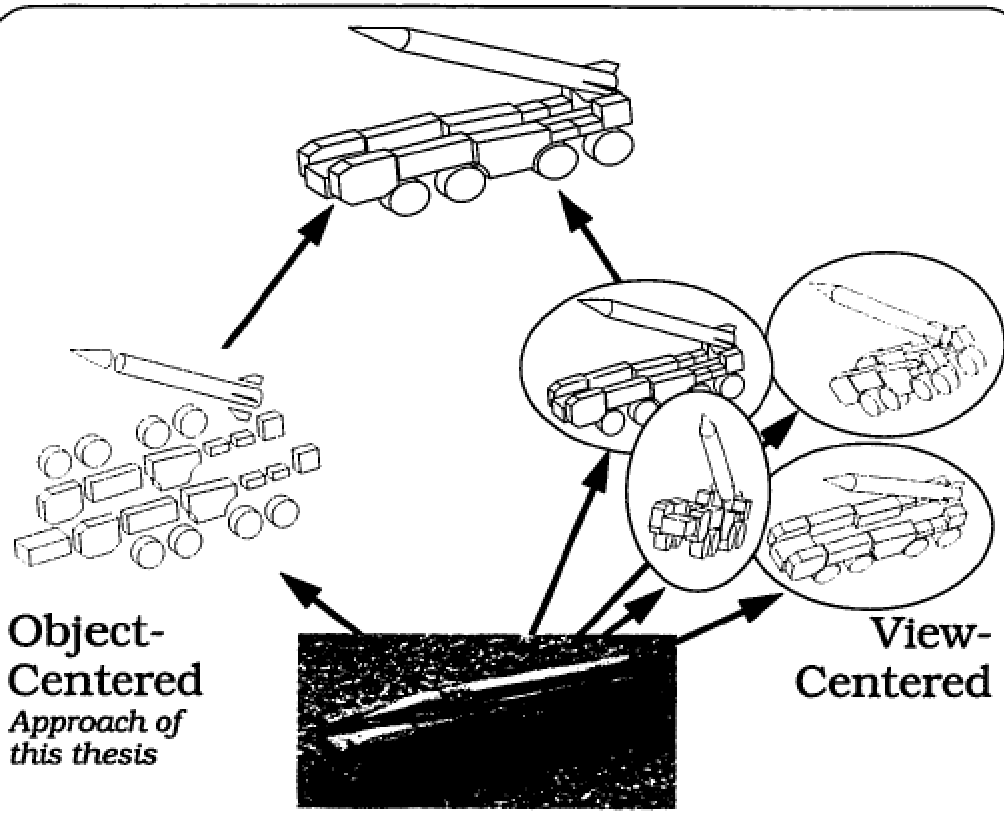


Figure 1-8: Object-centered vs. View-centered

In Object-centered vision, the image is interpreted in terms of geometric component parts. A geometric description is constructed, regardless of whether or not the specific image object exists in the model data base. In View-centered vision, the image is compared with all possible views of specific object models. View-centered vision suffers from higher combinatoric complexity and is best for finding specific object models. Object-centered algorithms are more difficult to design.

Contributions of Mann (1995)

SUCCESSOR

- Implementation of end-to-end interpretation system
- Successful interpretation of some simple, 3-D objects

Classics

- Design and implementation of highly typed constraint system
- Theoretical design of mathematically-based object system

VSCP

- Detailed framework of VSCP representation for object modeling
- Implementation of large modeling data base

Bayesian Networks

- Modular framework for automated, dynamic Bayes net construction
- Details of automated conditional probabilities for Bayesian networks
- Theoretical analysis of some 3 and 4 point quasi-invariants for vision
- Implementation of distributed network connection to HUGIN and SPI

Edge Grouping

- Simple edge segmentation, classification and grouping algorithms

Step by Step System Overview (1)

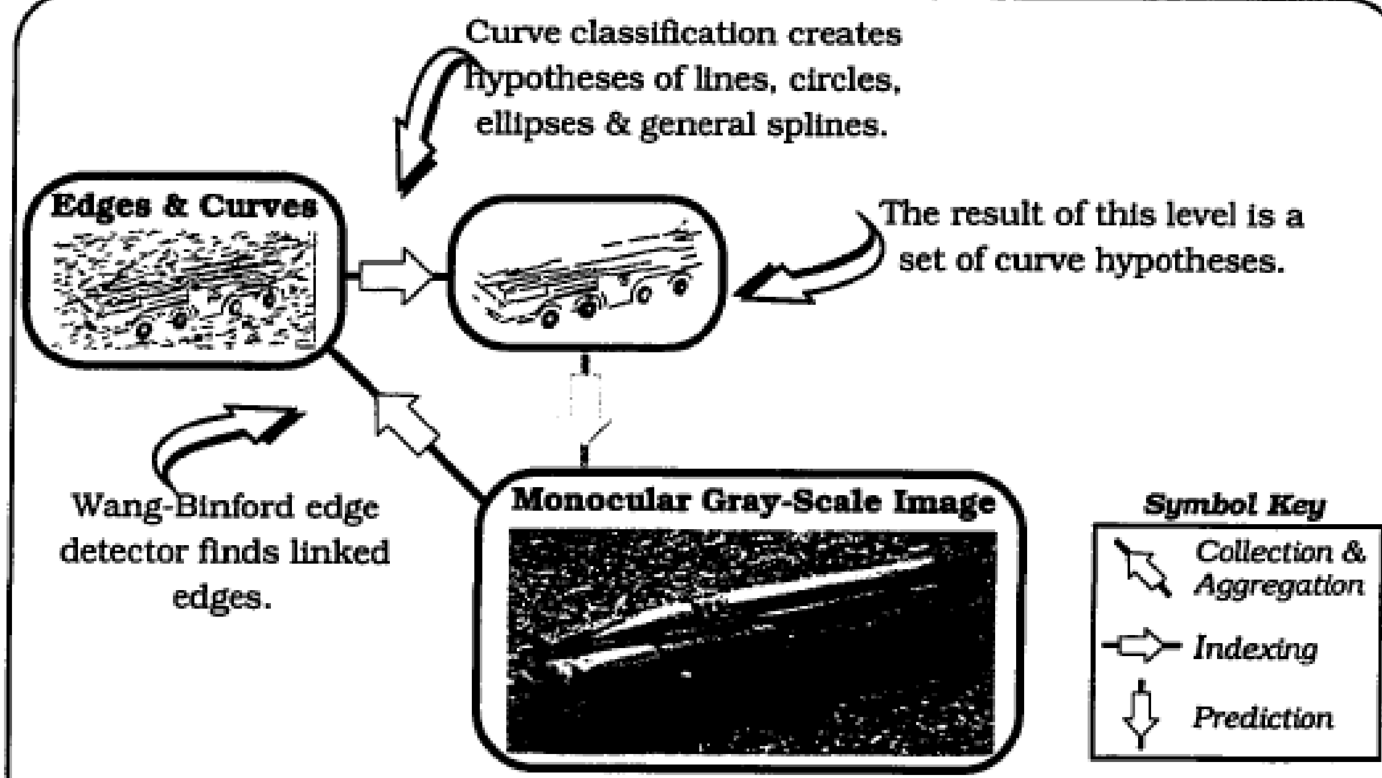


Figure 2-1A: Step One in Pictorial View of Algorithm

Curve Level

This shows the first cyclical step of interpretation. The Wang-Binford edge detector is itself a major development presented in [Wang 1994A] and [Wang 1994B]. Prediction is not used at this level. The MAZ-543 example pictured here is our aim, but more complex than our current implementation will recognize.

Step by Step System Overview (2)

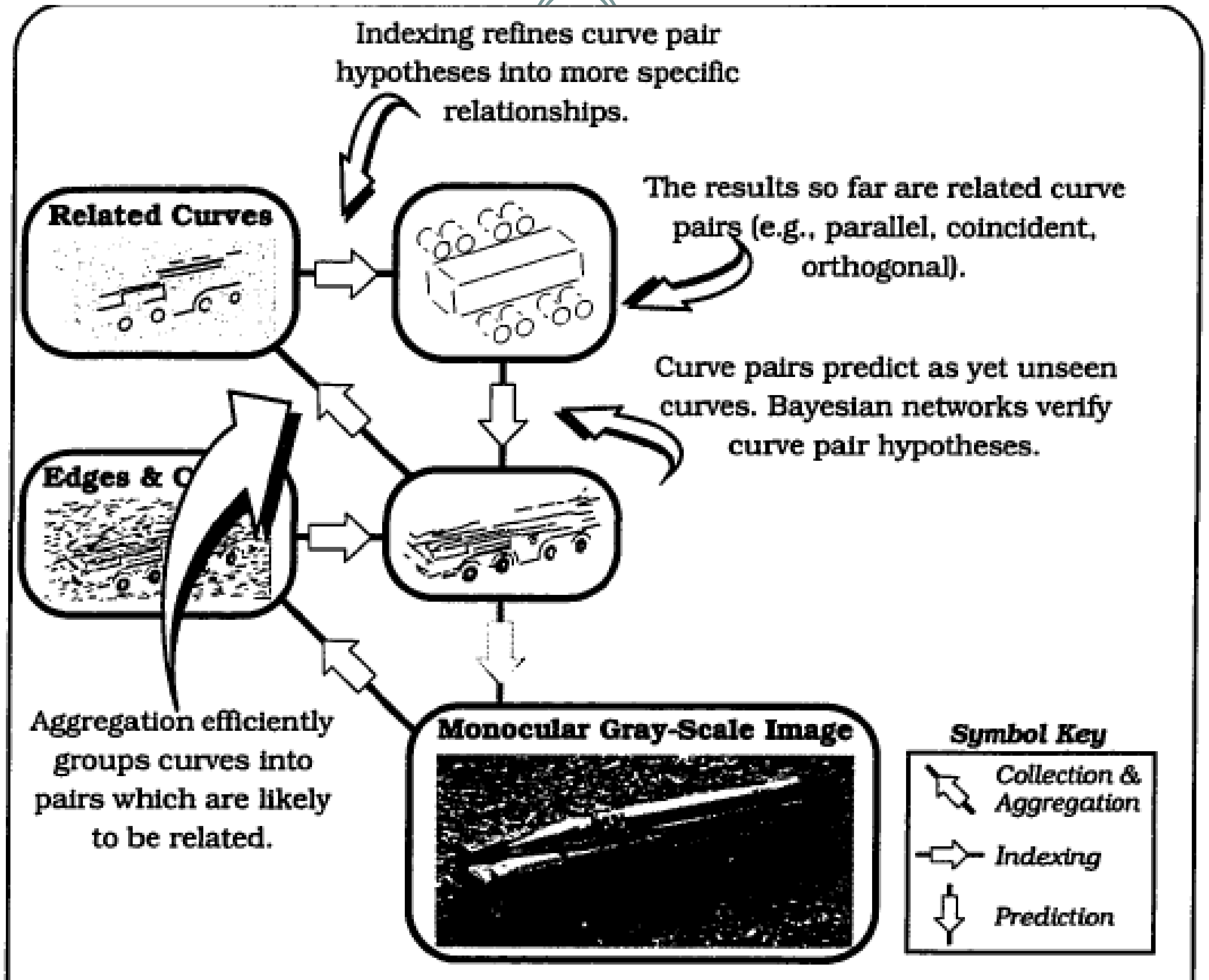
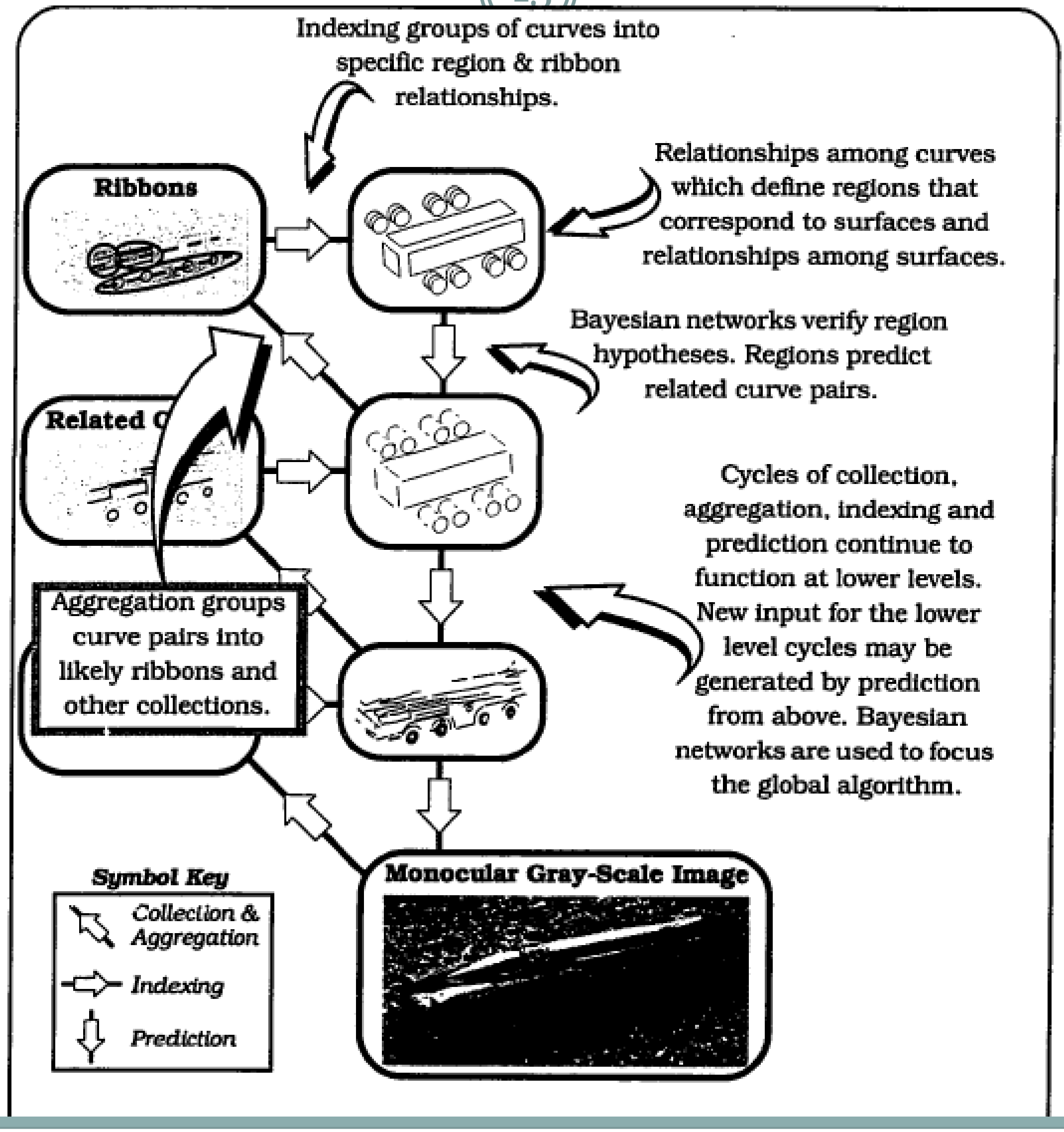


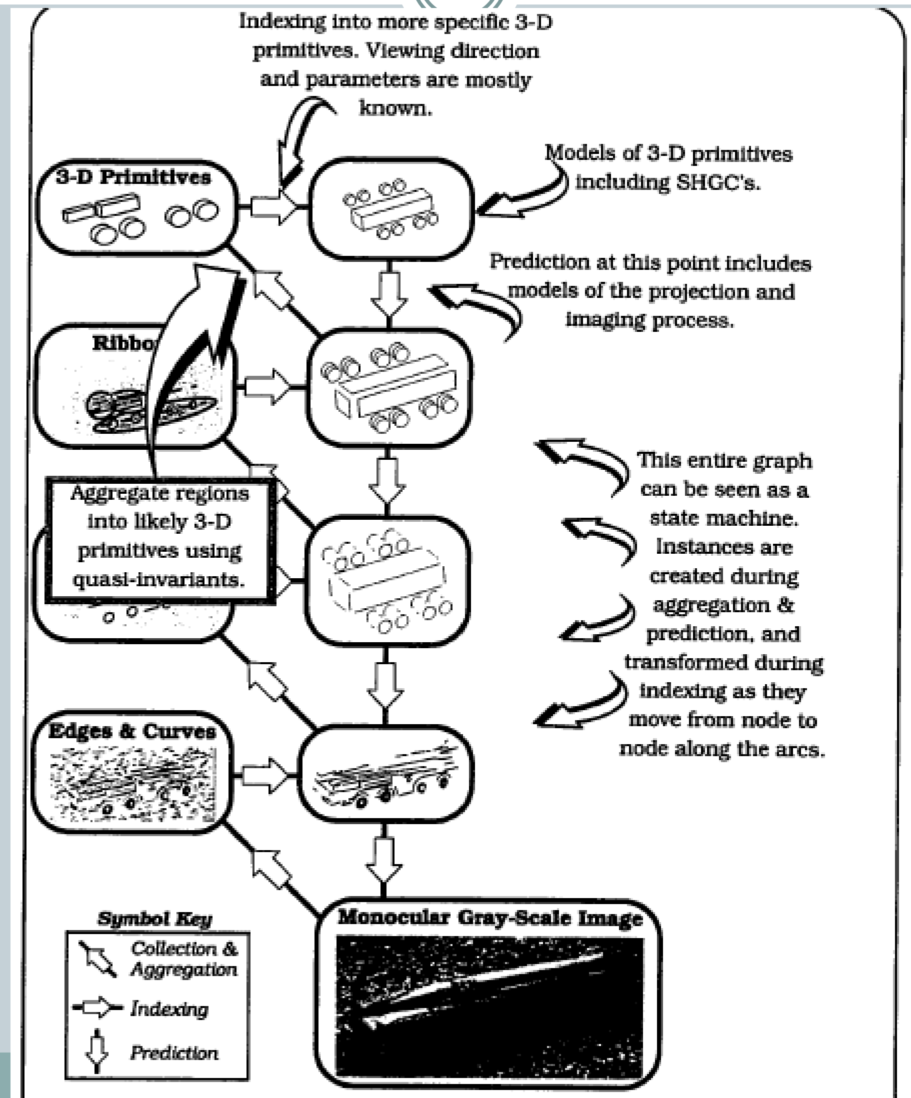
Figure 2-1B: Step Two in Pictorial View of Algorithm

Corresponding Curve Level

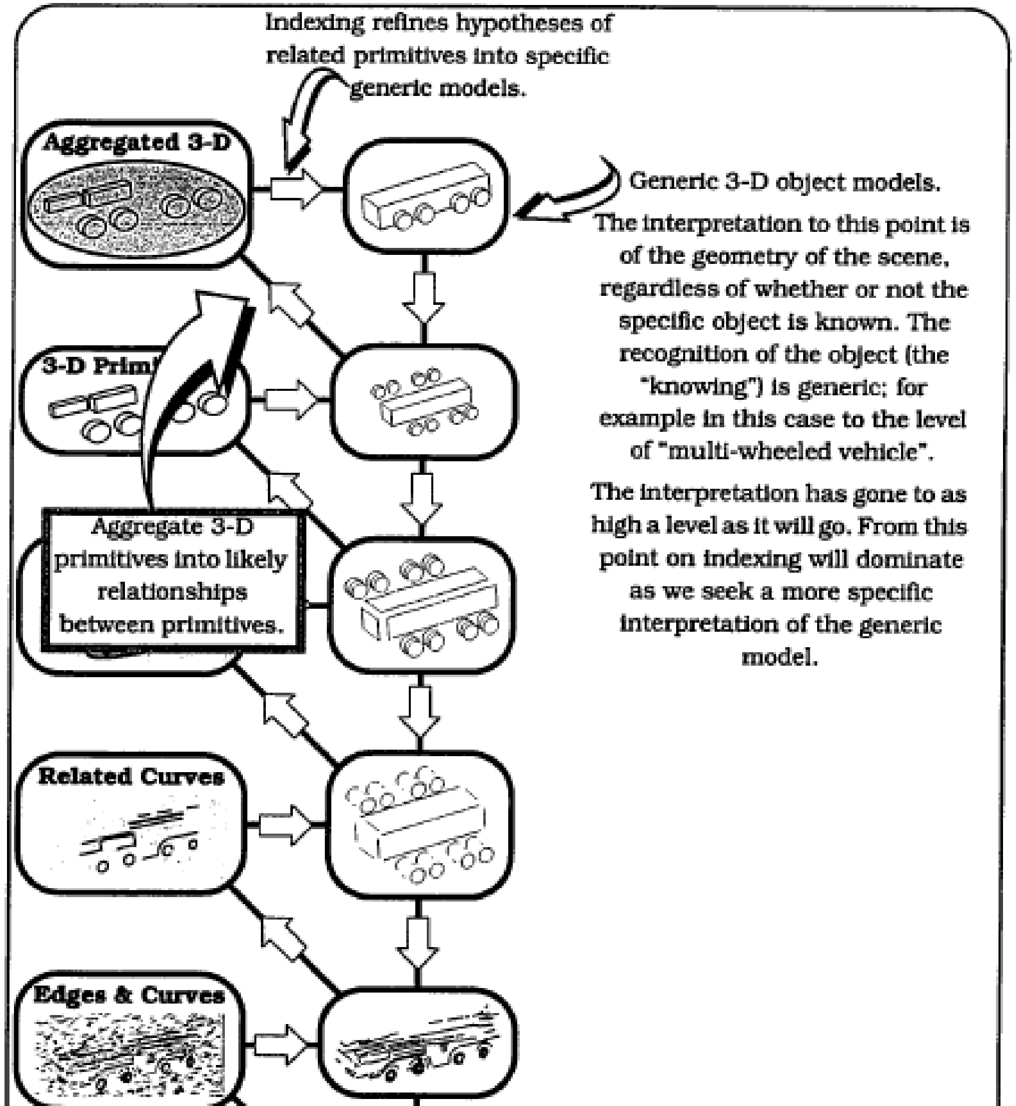
Step by Step System Overview (3)



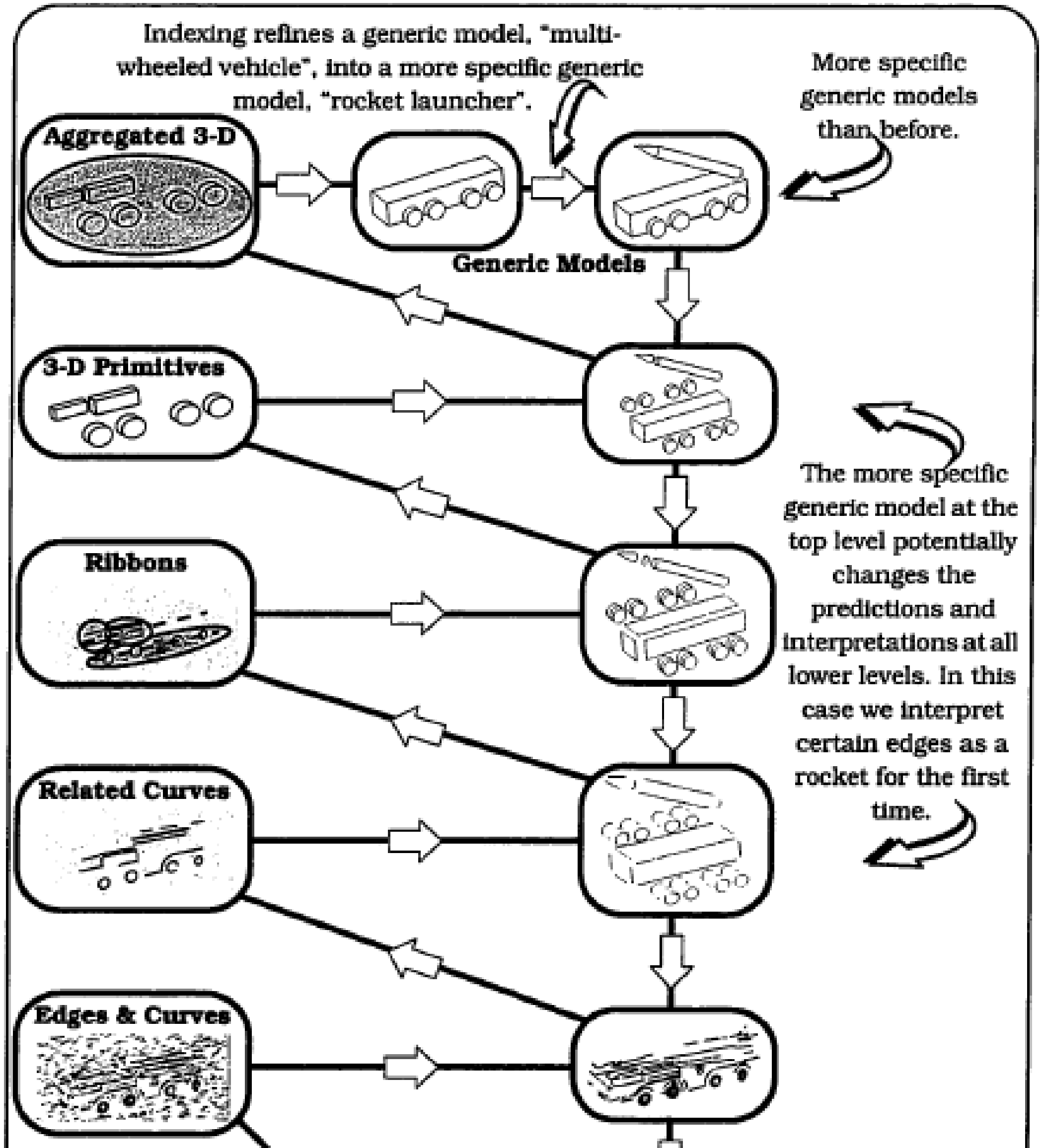
Step by Step System Overview (4)



Step by Step System Overview (5)



Step by Step System Overview (6)



Overall Example

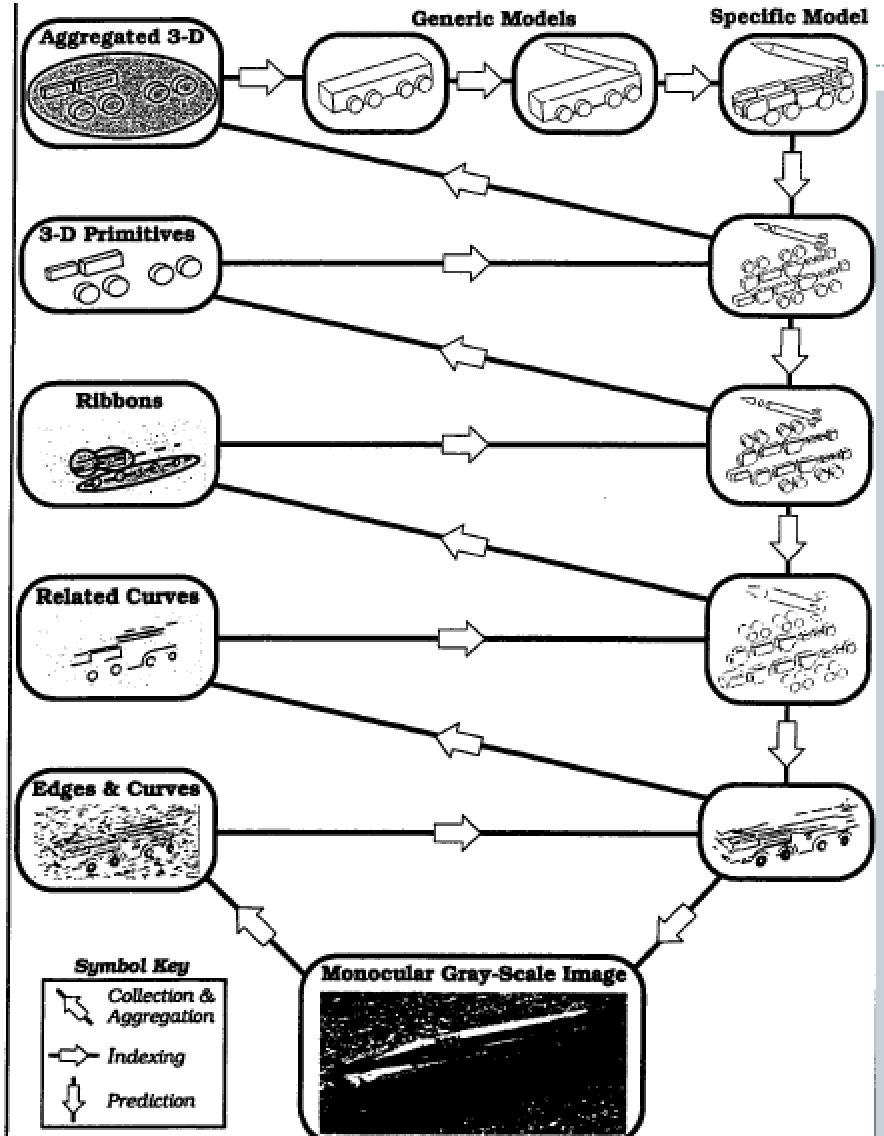
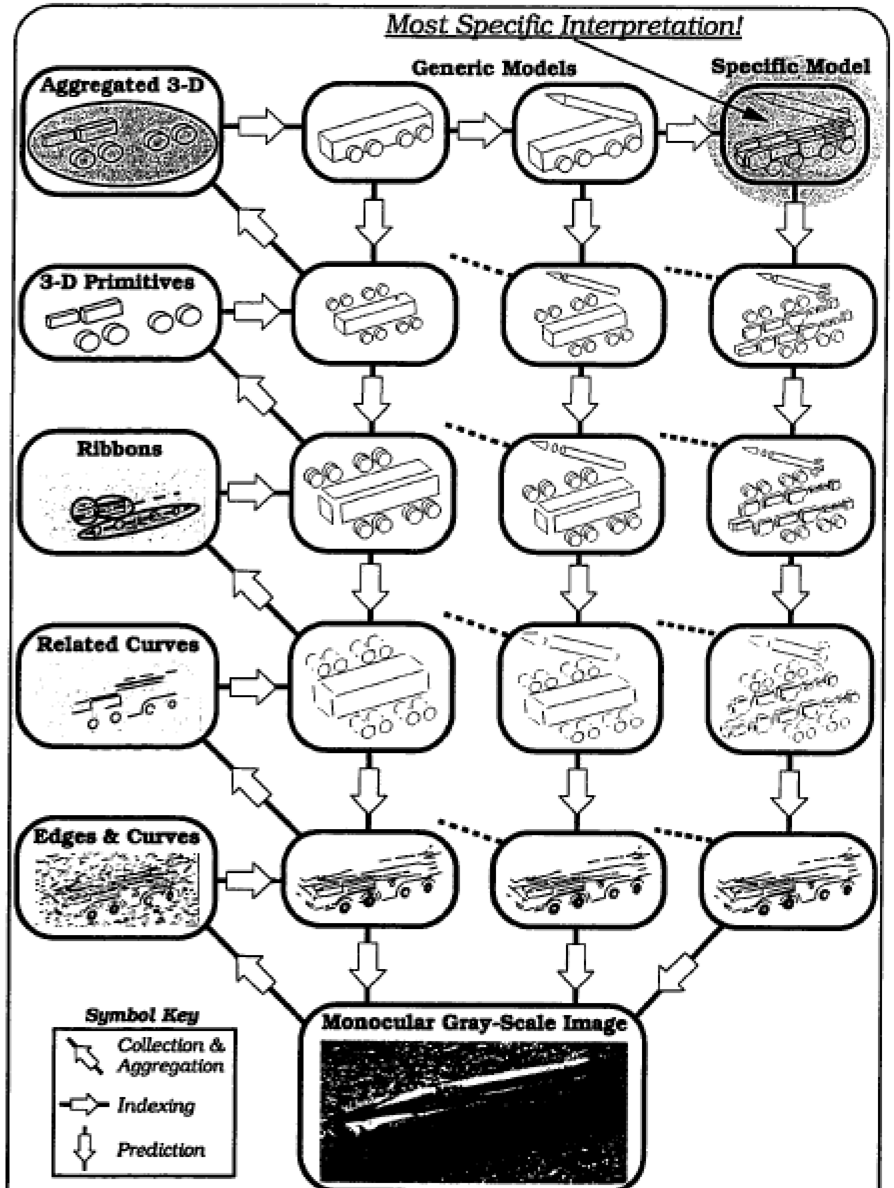


Figure 2-1G: Final Step in Pictorial View of Algorithm - Specific Models

System Summary



Basic Cycle in the Algorithm

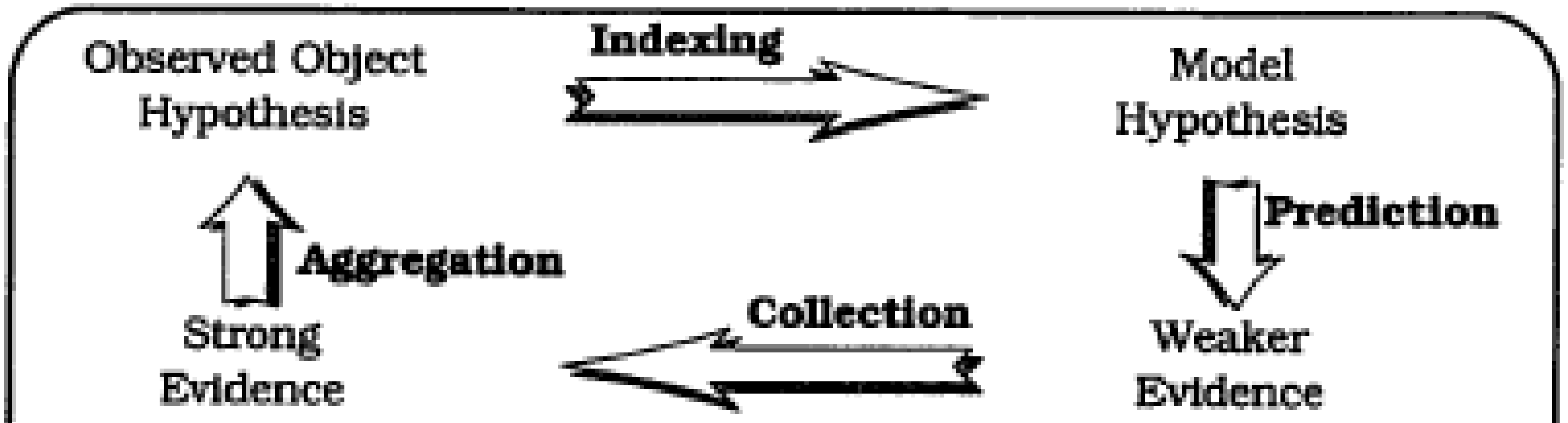


Figure 2-3: Fundamental Cycle of Interpretation

This is the basic cycle of the algorithm in this work.

Intrepretation in "Successor" System

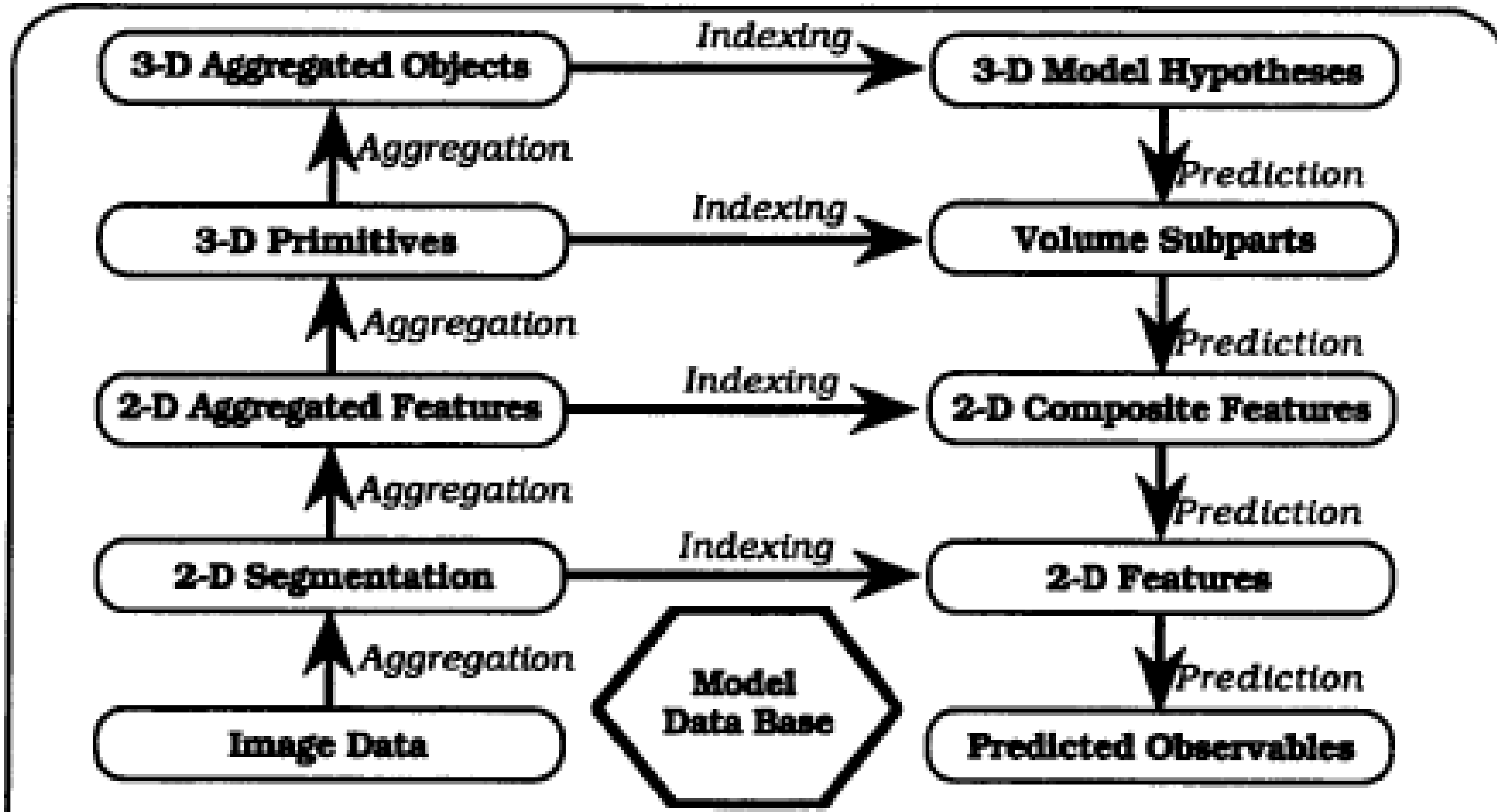


Figure 2-4: Interpretation in Successor

Elbow Image – Edge Detection Results

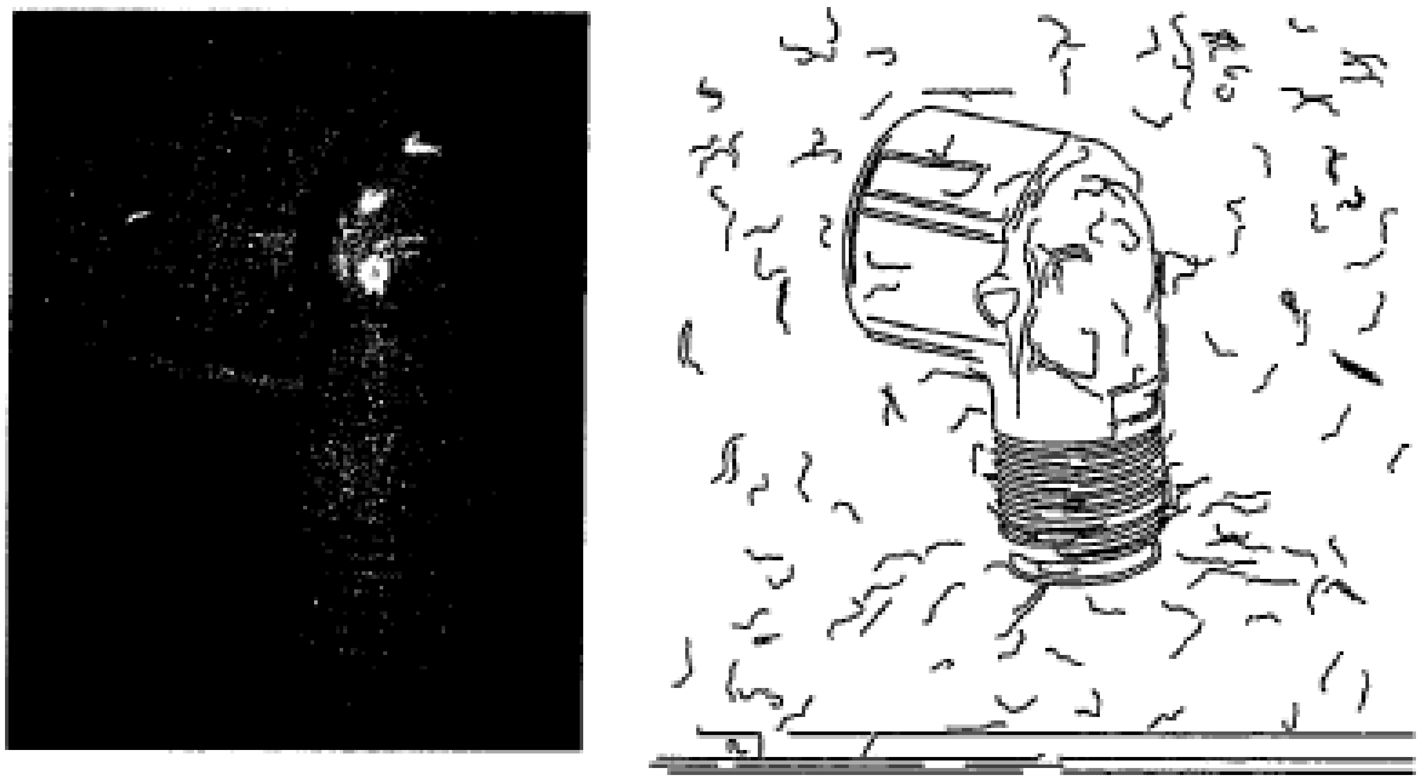


Figure 2-5: Elbow Image & Wang-Binford Edges

Classification of Edges in Curves

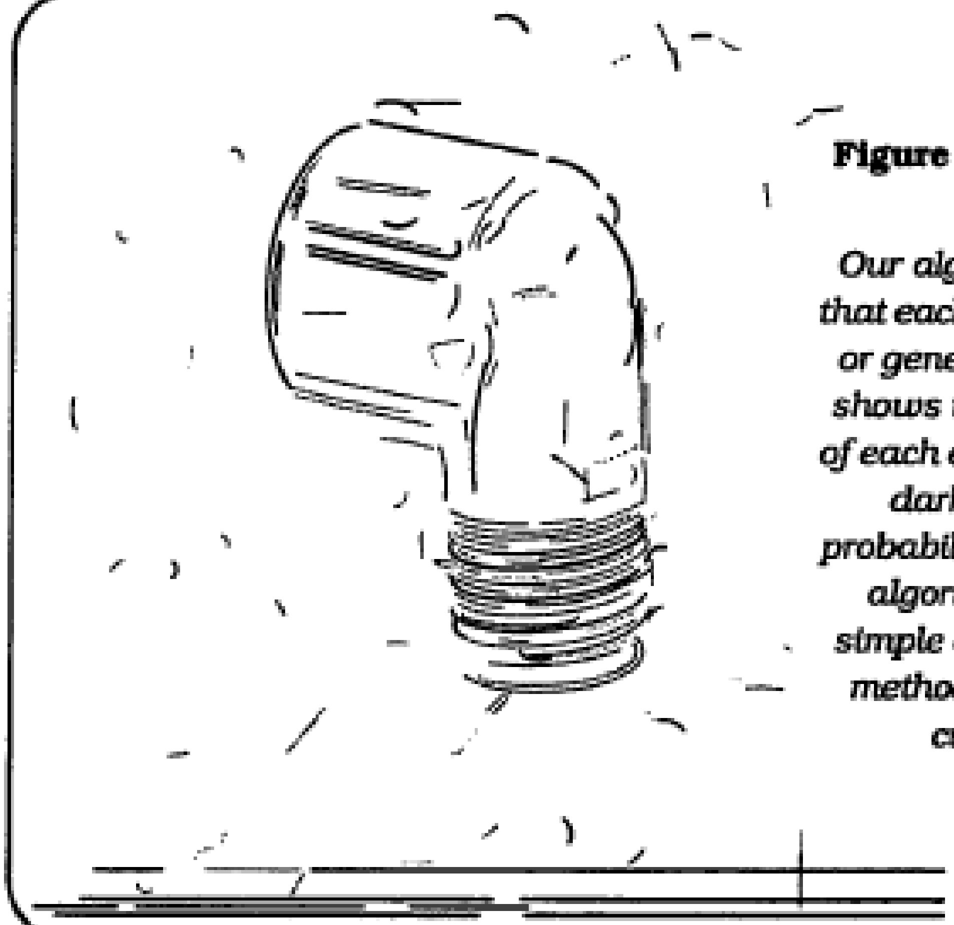


Figure 2-6: Edges Classified into Curves

Our algorithm finds the probability that each edge is a line, circle, ellipse, or general spline curve. This figure shows the most likely classification of each edge, with the thickness and darkness proportional to the probability of that classification. Our algorithm uses classification of simple curve shape, but the overall methodology does not depend on curve specific methods.

Elbow – Interpretation Example

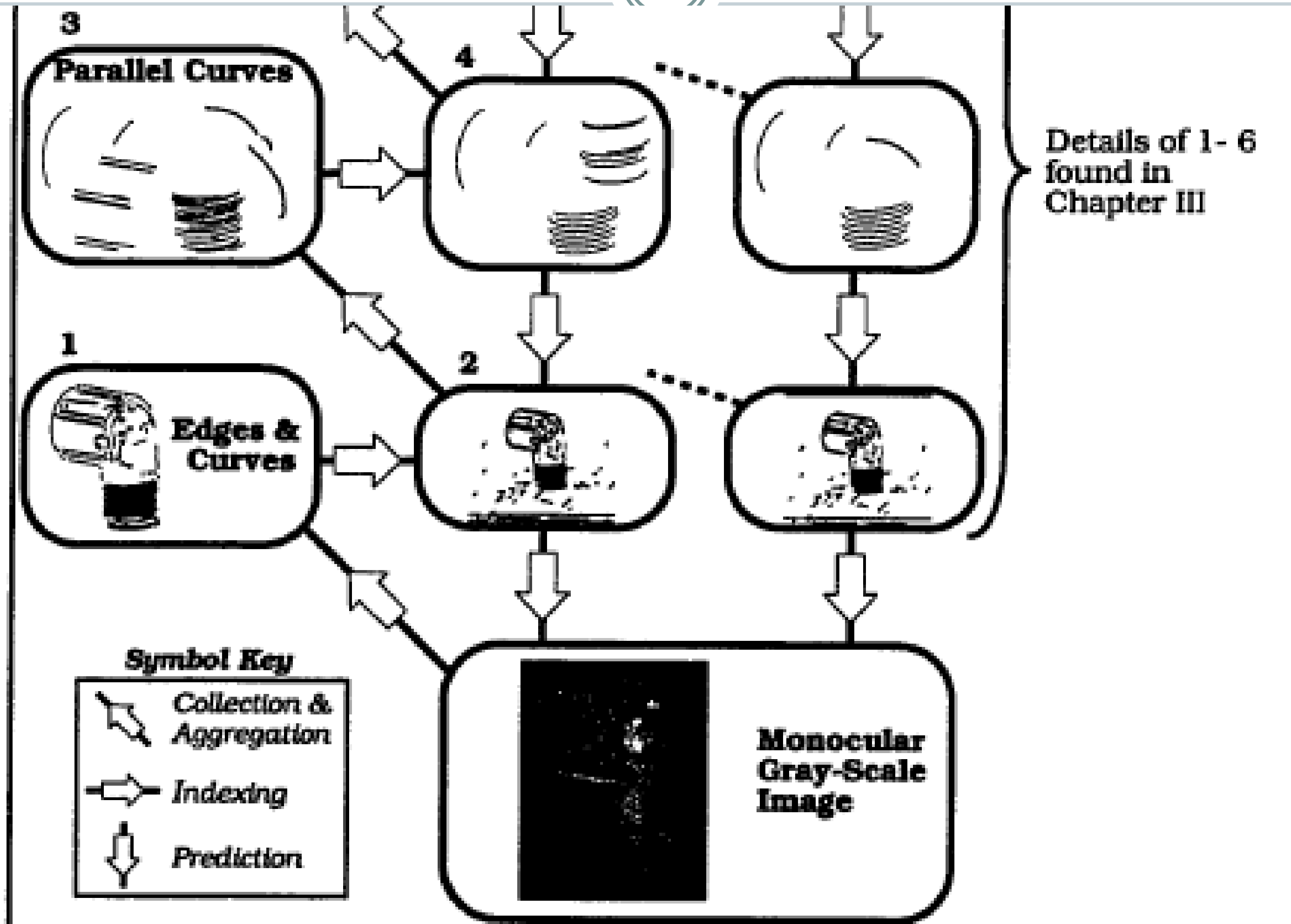
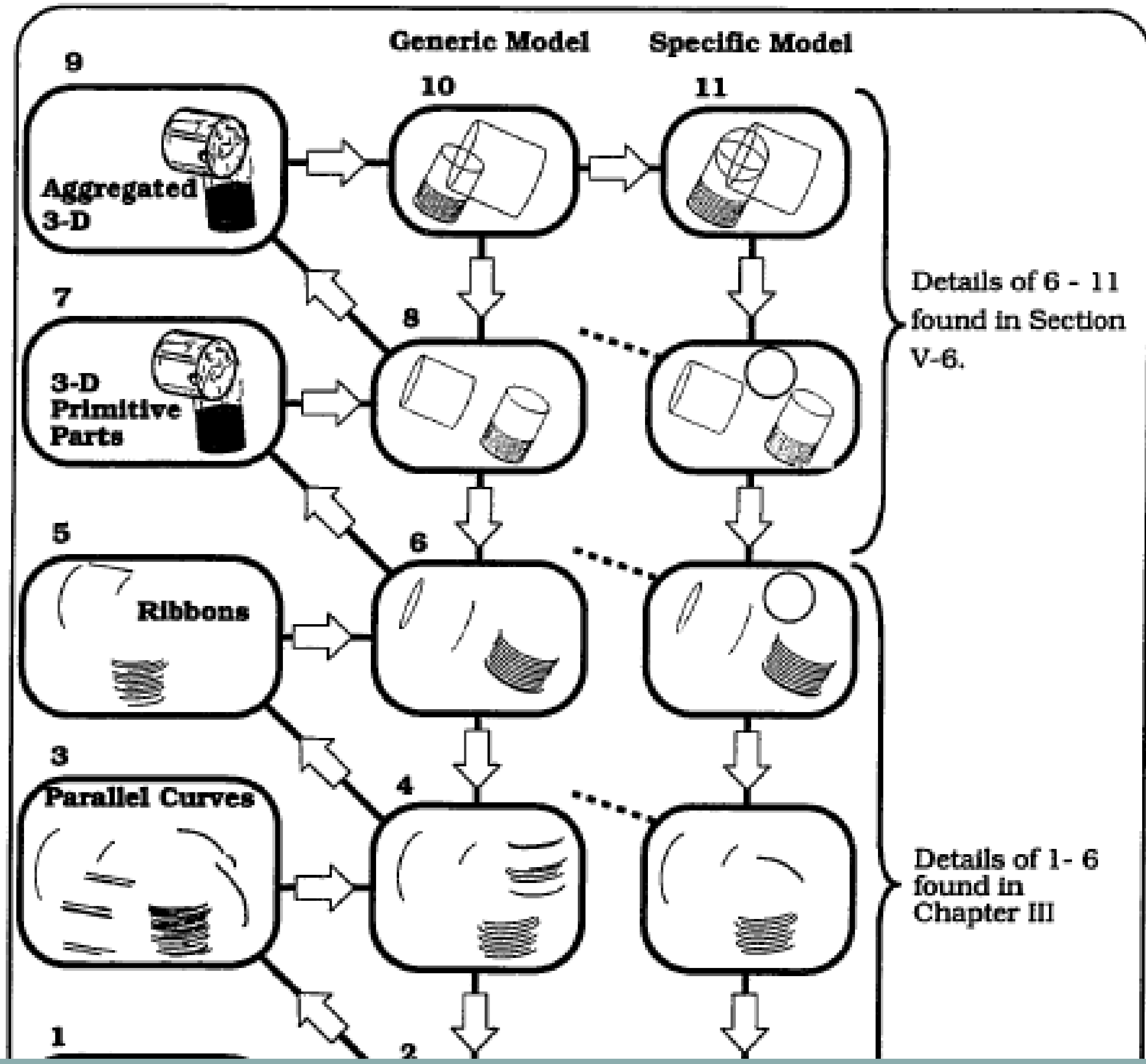


Figure 2-7: Summary View of Street Elbow Interpretation

Elbow – Interpretation Example (2)



Extraction of Parallel Curve Groups

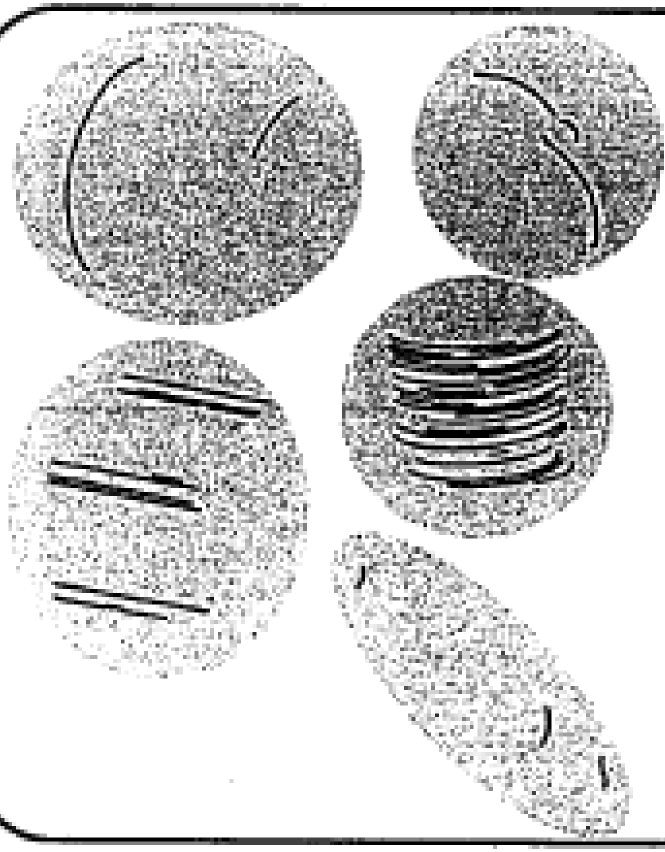


Figure 2-8 Groups of Possible Parallel Curves

We find translational parallelism for any direction of translation of general curves. This corresponds to point 3 of Figure 2-7. An efficient algorithm finds these groups of curves which are deemed likely to be parallel. Model constraints, similarity measures and regularity measures are used to further group these curves into well defined collectives

Ribbons from Parallel Curves

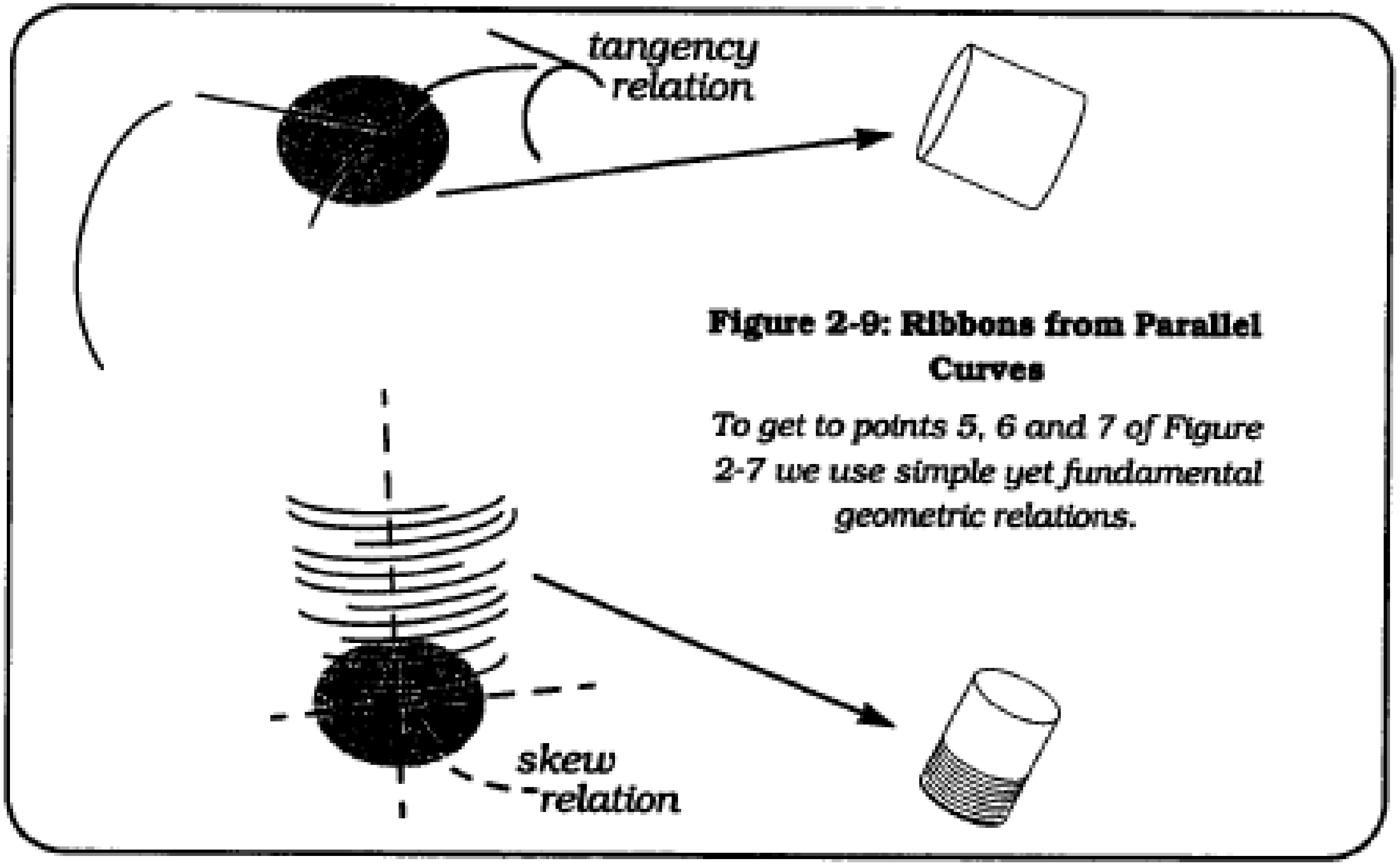


Figure 2-9: Ribbons from Parallel Curves

To get to points 5, 6 and 7 of Figure 2-7 we use simple yet fundamental geometric relations.

3D Hypotheses

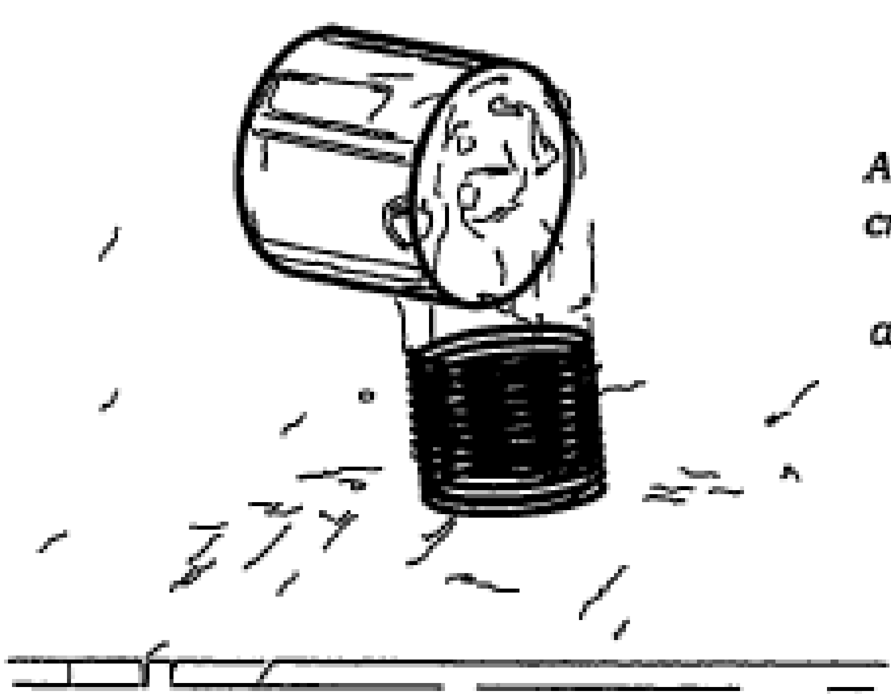


Figure 2-10: 3-D Hypotheses

At points 7 and 8 of Figure 2-7 we have created the hypothesis of two cylinders in the scene, one of which has an associated helix. No further refinement occurs in this image during indexing from 7 to 8.

Finding the Generic Model

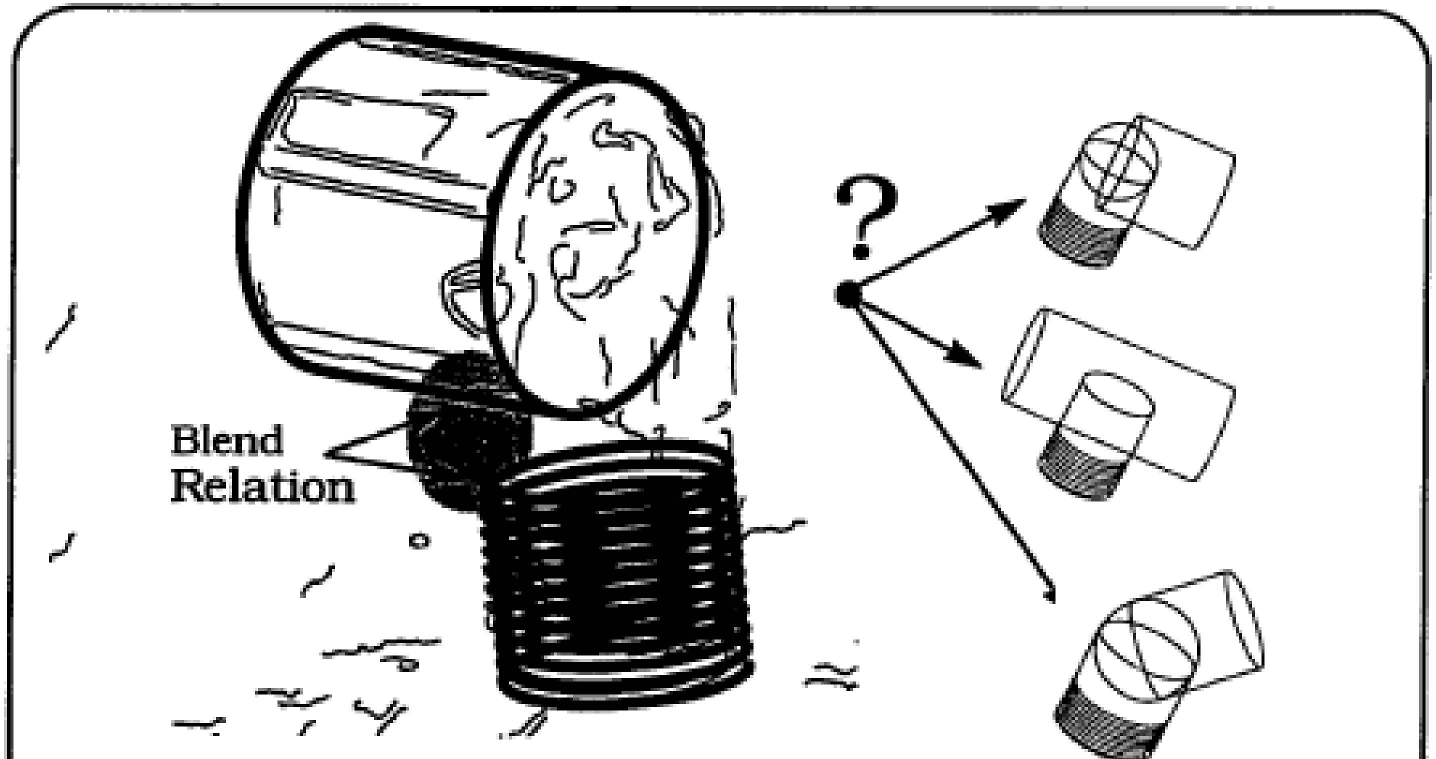


Figure 2-11: Finding the Generic Model

These are three specific object models from the generic class of cylinder pairs. The blend relation takes the process from point 8 to 9 of Figure 2-7.

Model Structure Examples

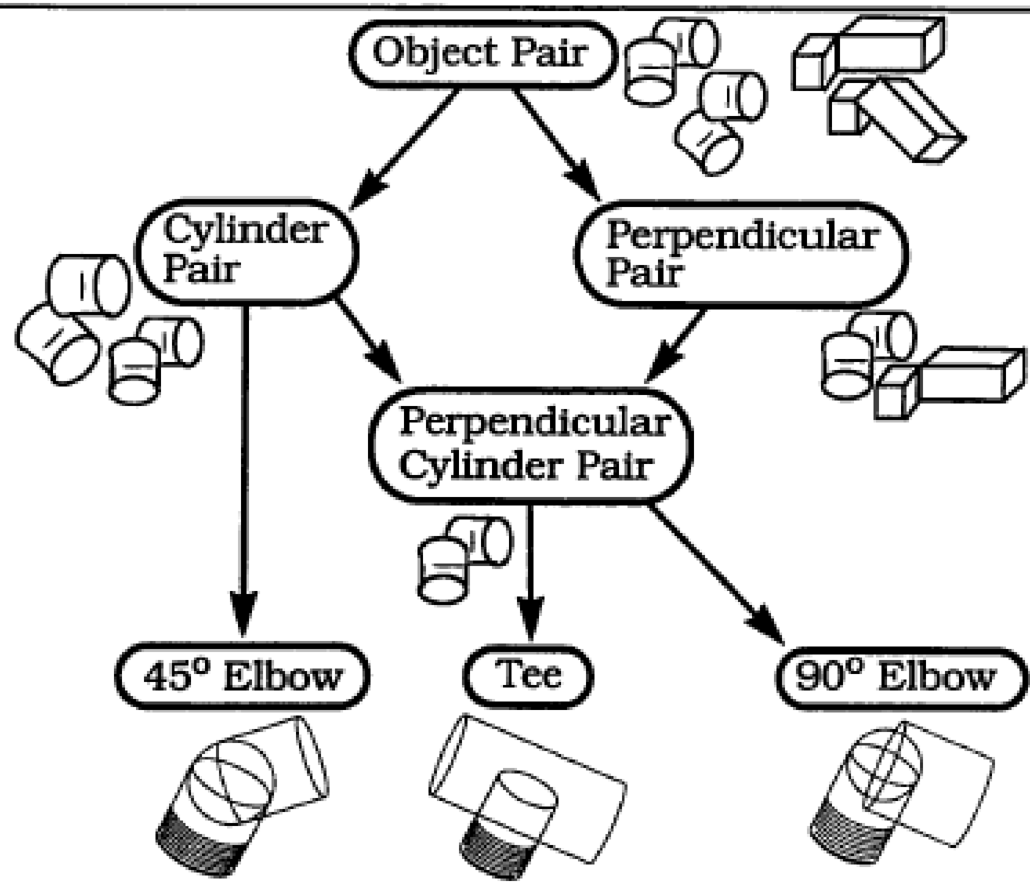


Figure 2-12: Portion of the Generic Model Structure

This shows the generic models immediately relevant to the elbow. It is this portion of the model data base that guides the algorithm from steps 9 - 11 pictured in Figure 2-7.

Probabilities of Hypotheses

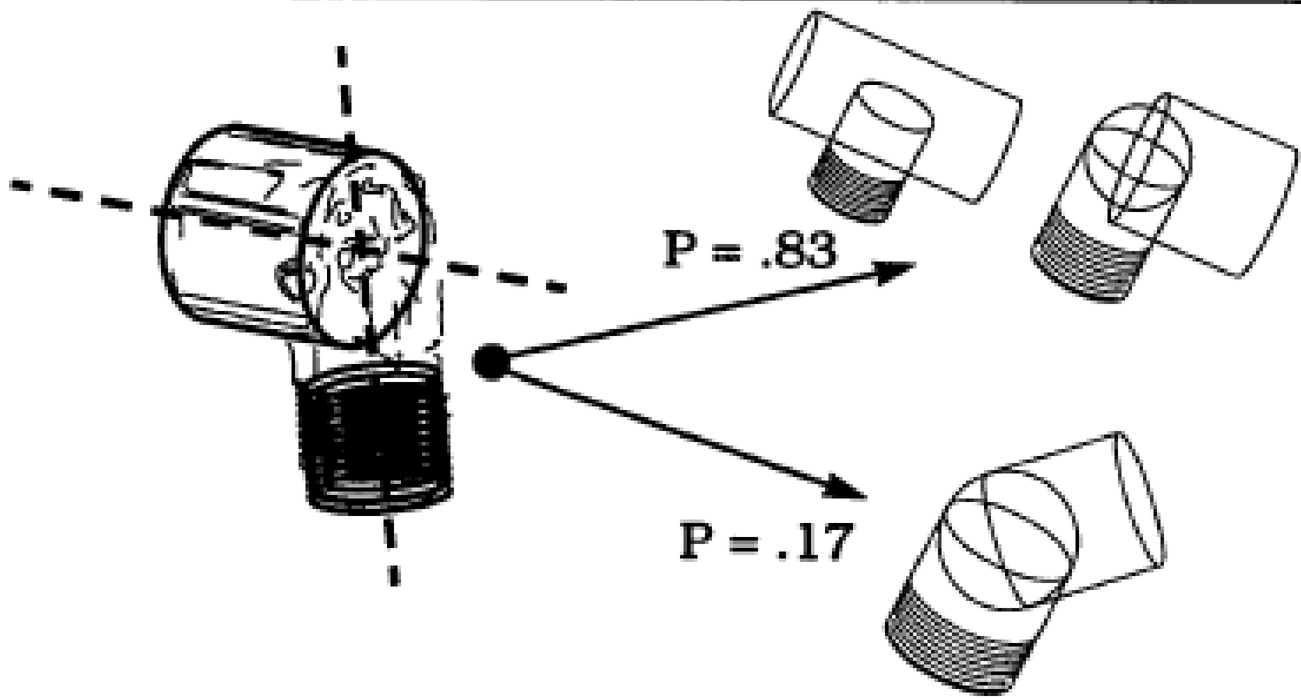


Figure 2-14: Using Probabilities to Refine Hypotheses

If we assume the actual angles take on only the two discrete values in the model data base (45 and 90), this gives the posterior probability distributions on the actual angle between two cylinders, given an observed angle. This figure shows the results for the elbow as it appears in Figure 1-3. This corresponds to step 10 in Figure 2-7.

Another Example View

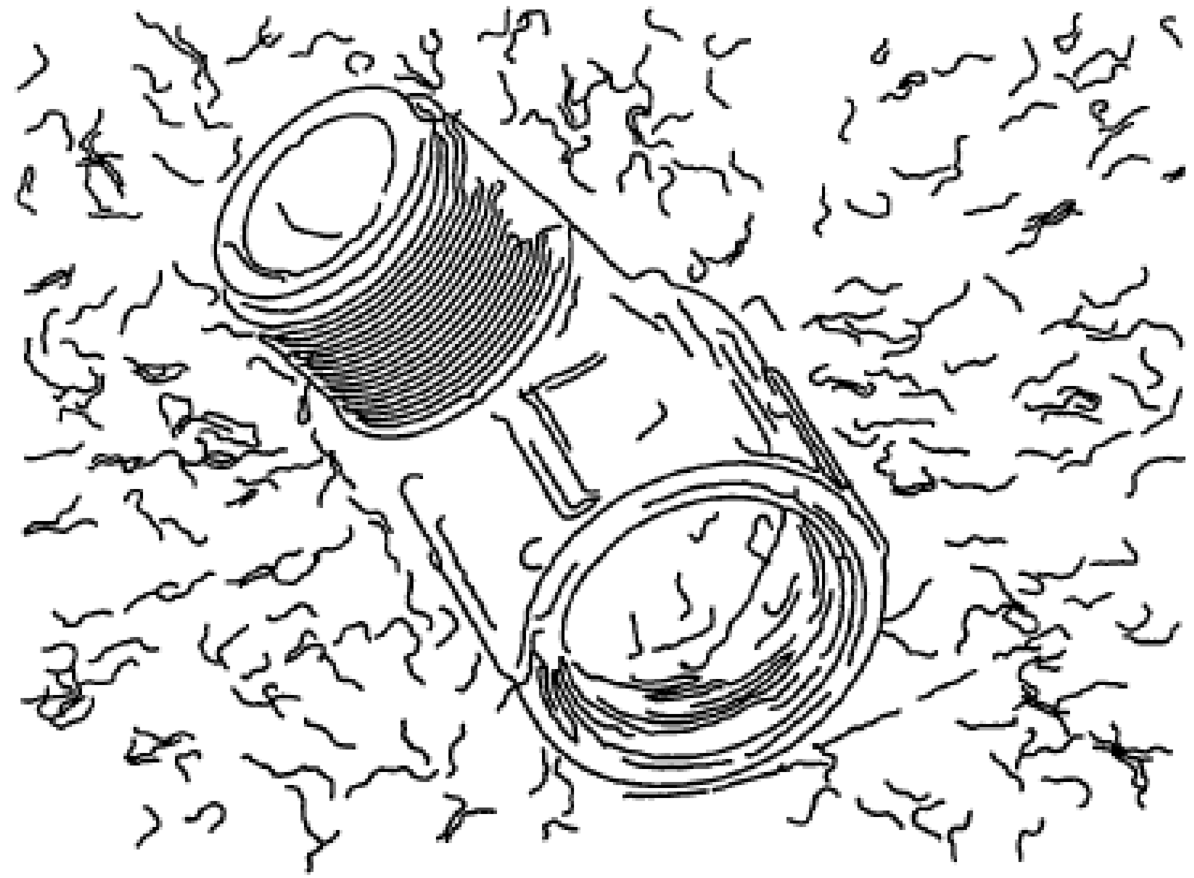


Figure 2-17: Edges for Second Elbow View

Curves Extracted

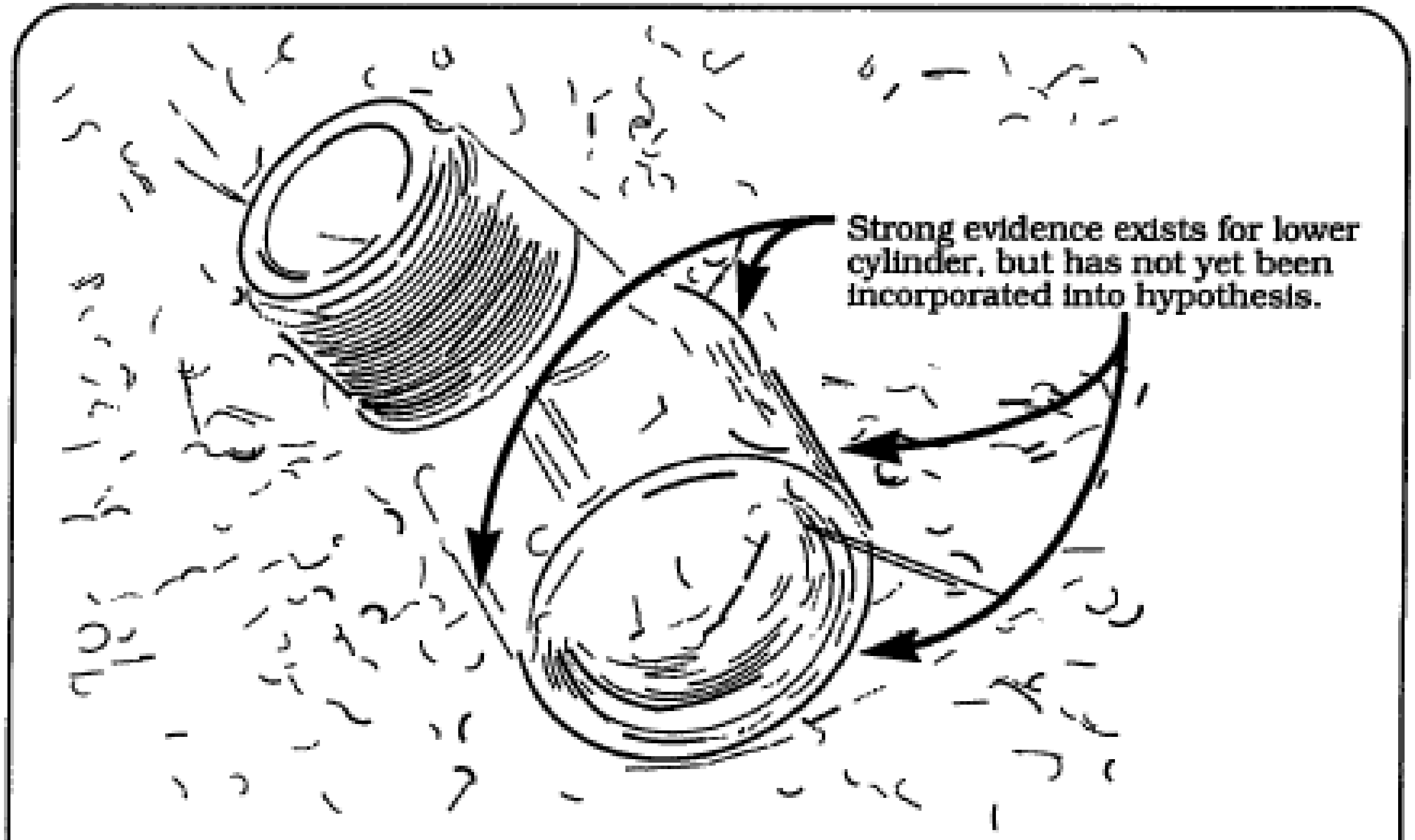


Figure 2-18: Curves for Second Elbow View

Another Example

Figure 2-21: Edges for Two Cylinders

More Complex Example

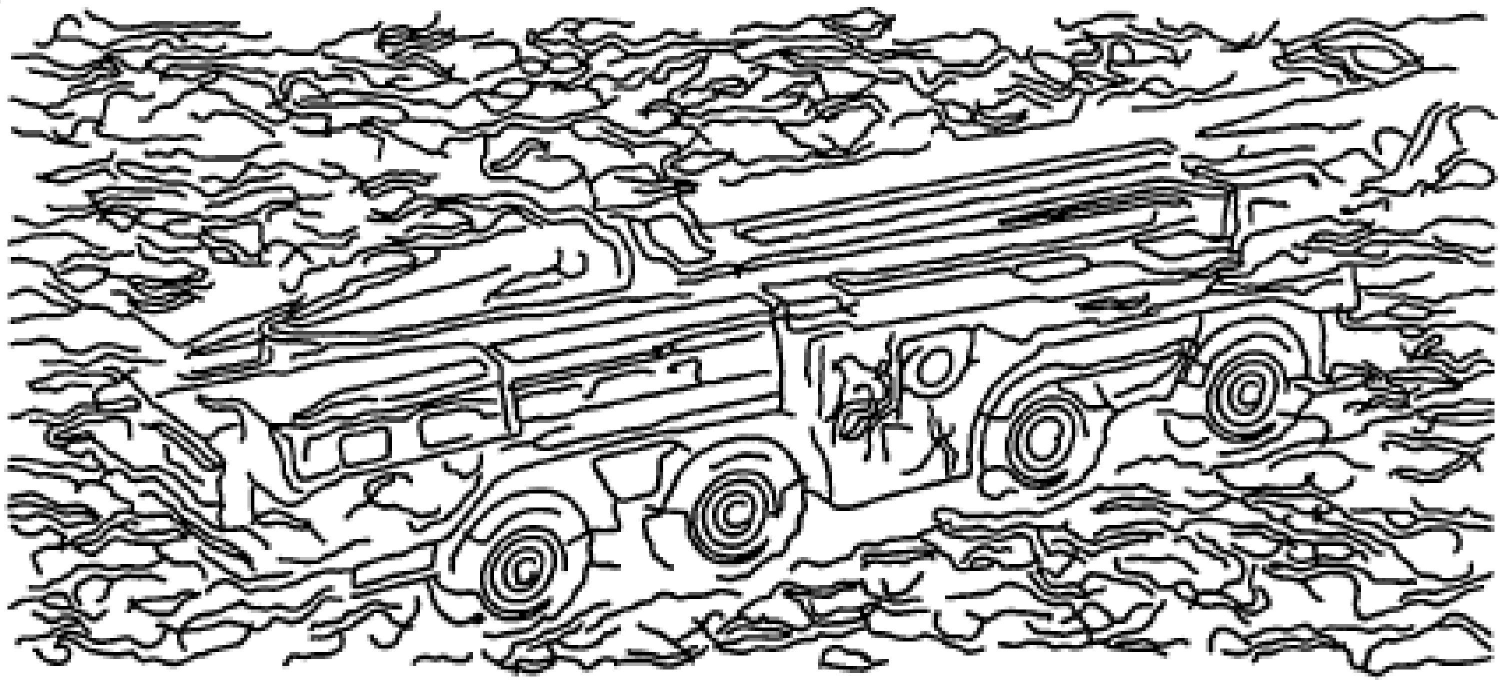


Figure 2-25: Edges for MAZ-543

Although we have not completed model-based recognition of the MAZ-543, this represents the level of complexity of object that we aim to achieve. The system generates Bayesian networks automatically for this complex example.

Parallel Curve Finding

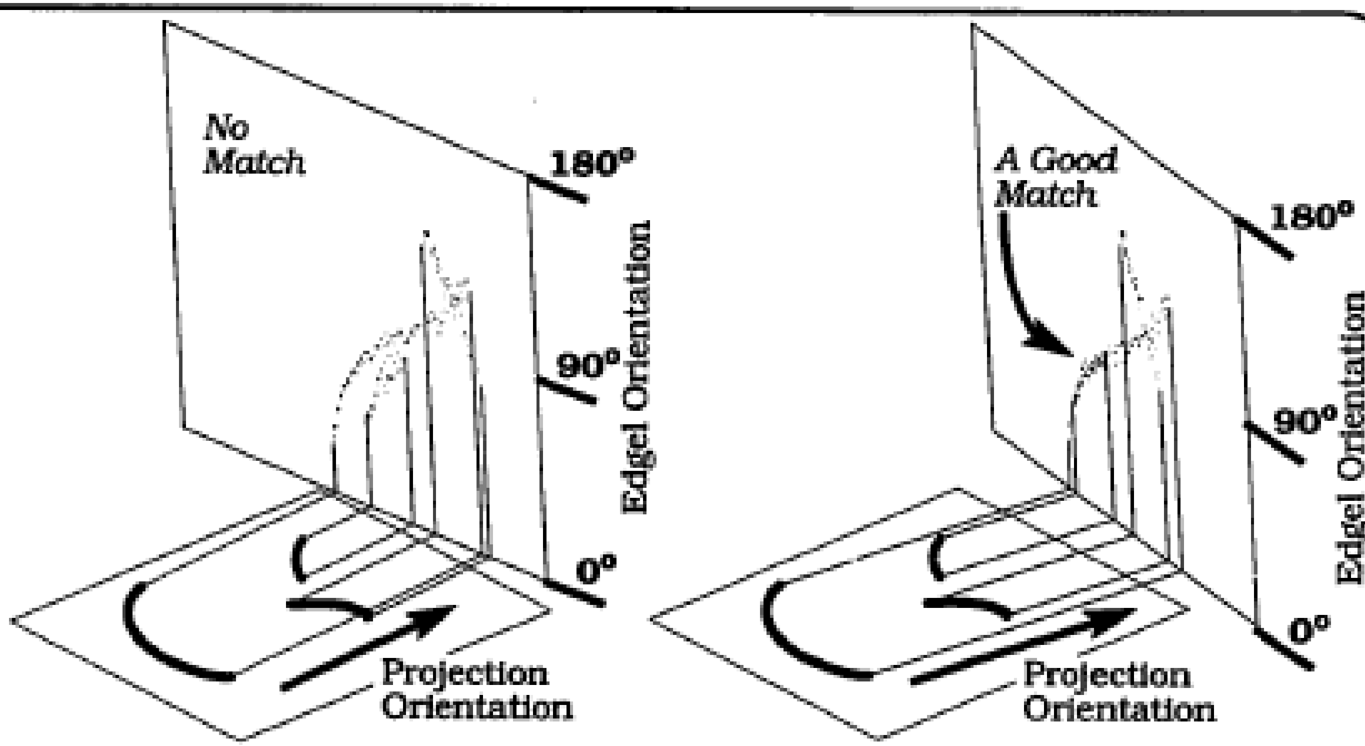


Figure 3-11: Projection of Edges in Parallel Curve Finding

In the left image, the curves are not found to be parallel for the shown projection direction. In the right image, they are parallel, as indicated by the large degree of overlap of the projections.

Measure for Parallelism

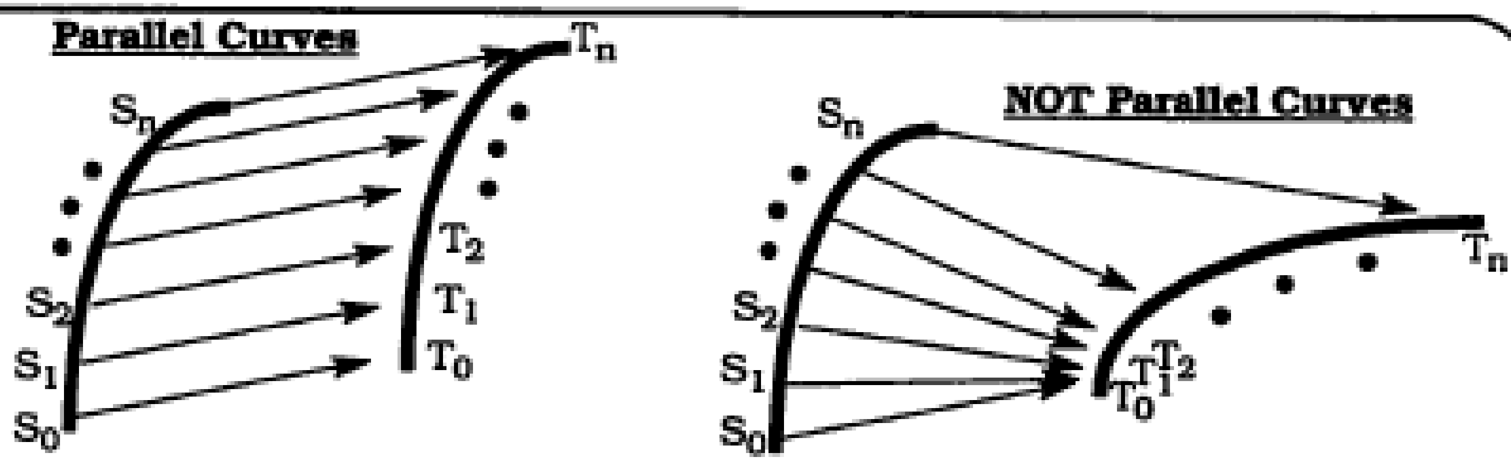


Figure 3-12: Mapping Arc Length to Distinguish Parallelism

The curves on the left are parallel and thus have a linear mapping between their arc length parameters. The curves on the right are not parallel, and have a highly nonlinear map between arc lengths. In particular, this algorithm finds only curves that are replicas of each other (i.e., the mapping has a unity scale factor).

Parallel Curve Extraction – General Version

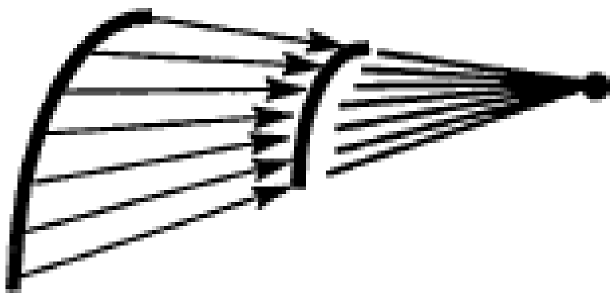


Figure 3-13: Scaled, Parallel Curves

A non-implemented extension to the parallel curve finding algorithm would be to find the parallel curves pictured here. All projection directions converge at a single point.

Groups of Curves

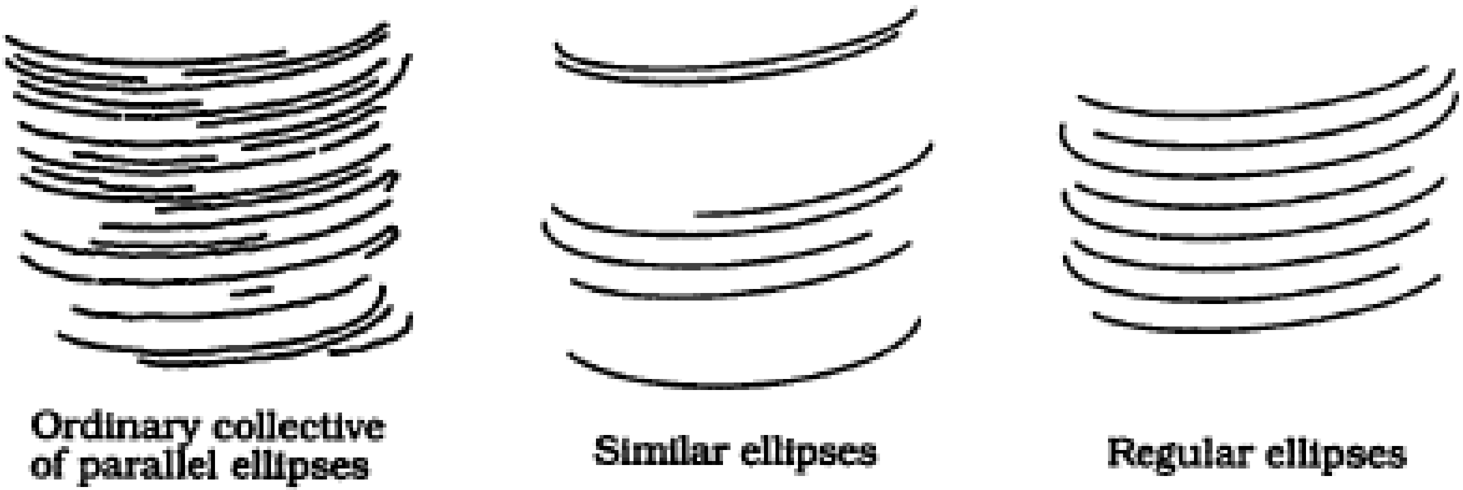


Figure 3-14: Example of Collectives

These curves are from the threads of the elbow edge image. On the left is the set of ellipses found to be parallel. One subset of those parallel ellipses is the set of similar ellipses shown in the middle. They have similar major and minor axis lengths, and similar orientations. On the right is a regular collective of ellipses found to be not only similar, but whose centers are spaced at regular intervals.

Finding Regularity

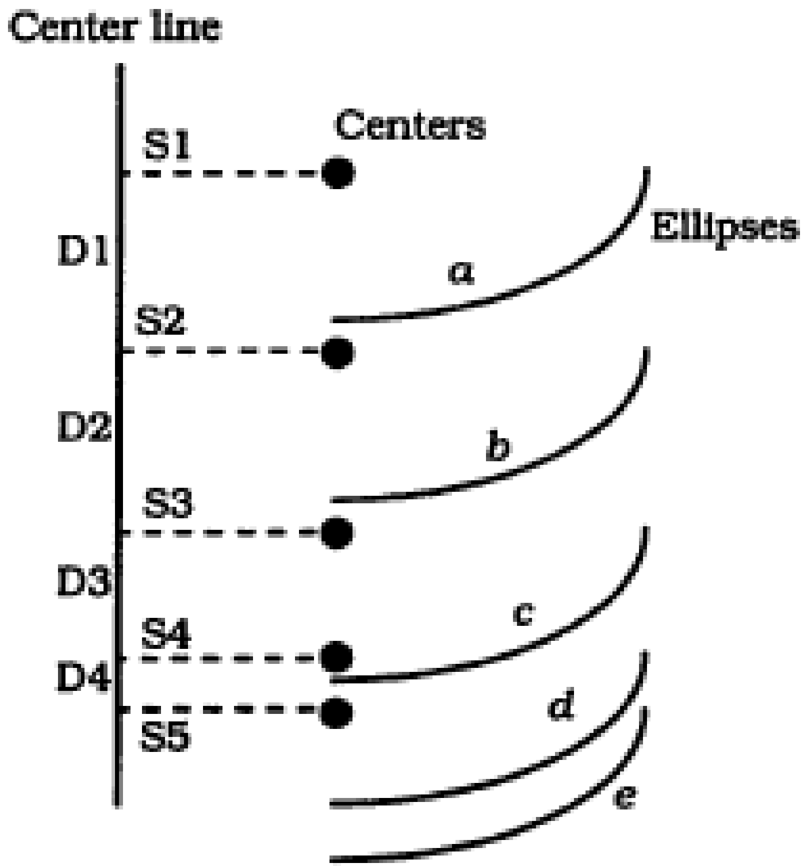


Figure 3-15: Finding Regularity

The centers of the ellipses are projected onto the best fit line through the centers. The arc length values of those projections form the metric. The D1-D4 are the distances between the arc length values. D1 & D2 are equal and thus form a regular triple. D2 & D3 are irregularly spaced, as are D3 & D4. Thus D1 = D2 = D becomes the interval of regularity. Finally, only ellipse d will be eliminated because its center is not close to an integer multiple of D away from the end.

VSCP Representation for Objects

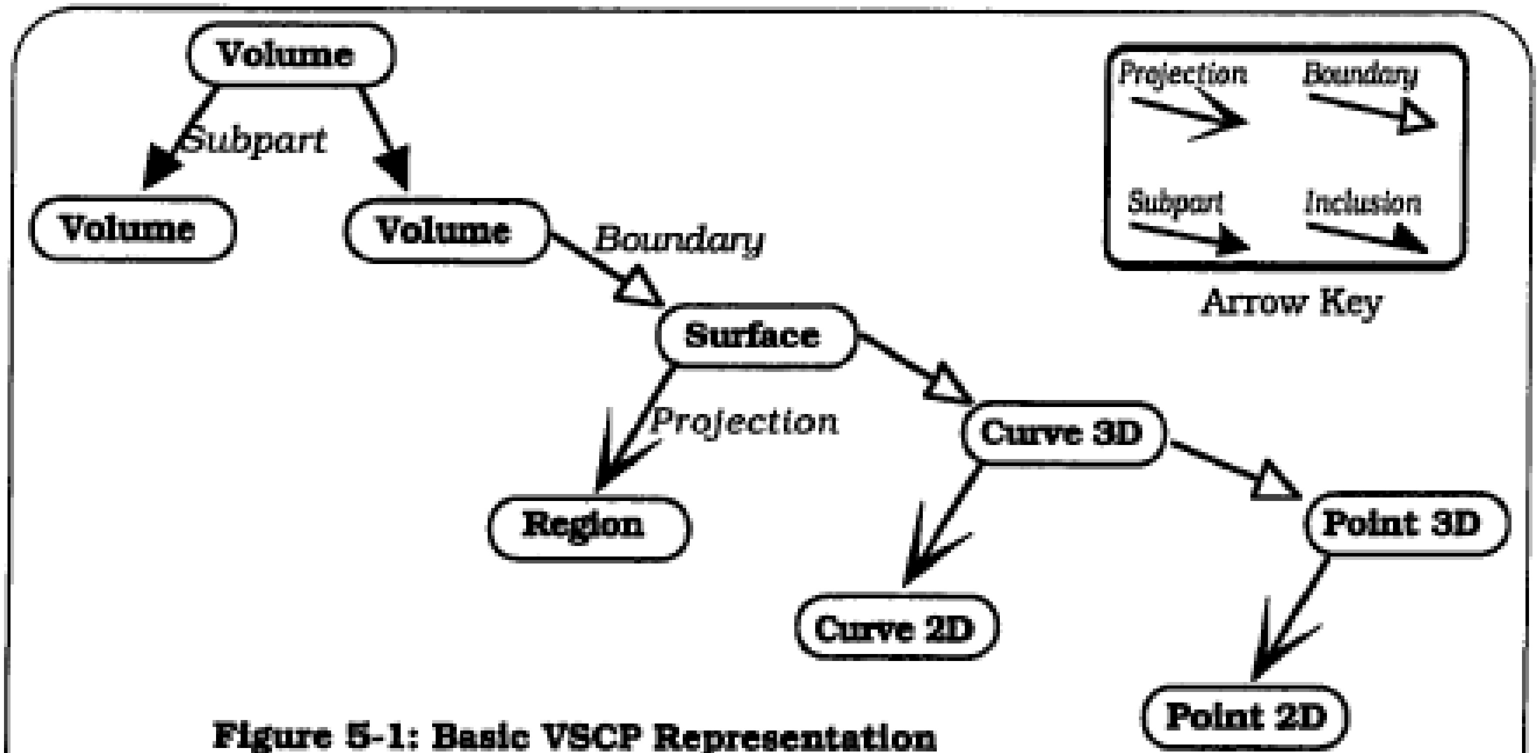
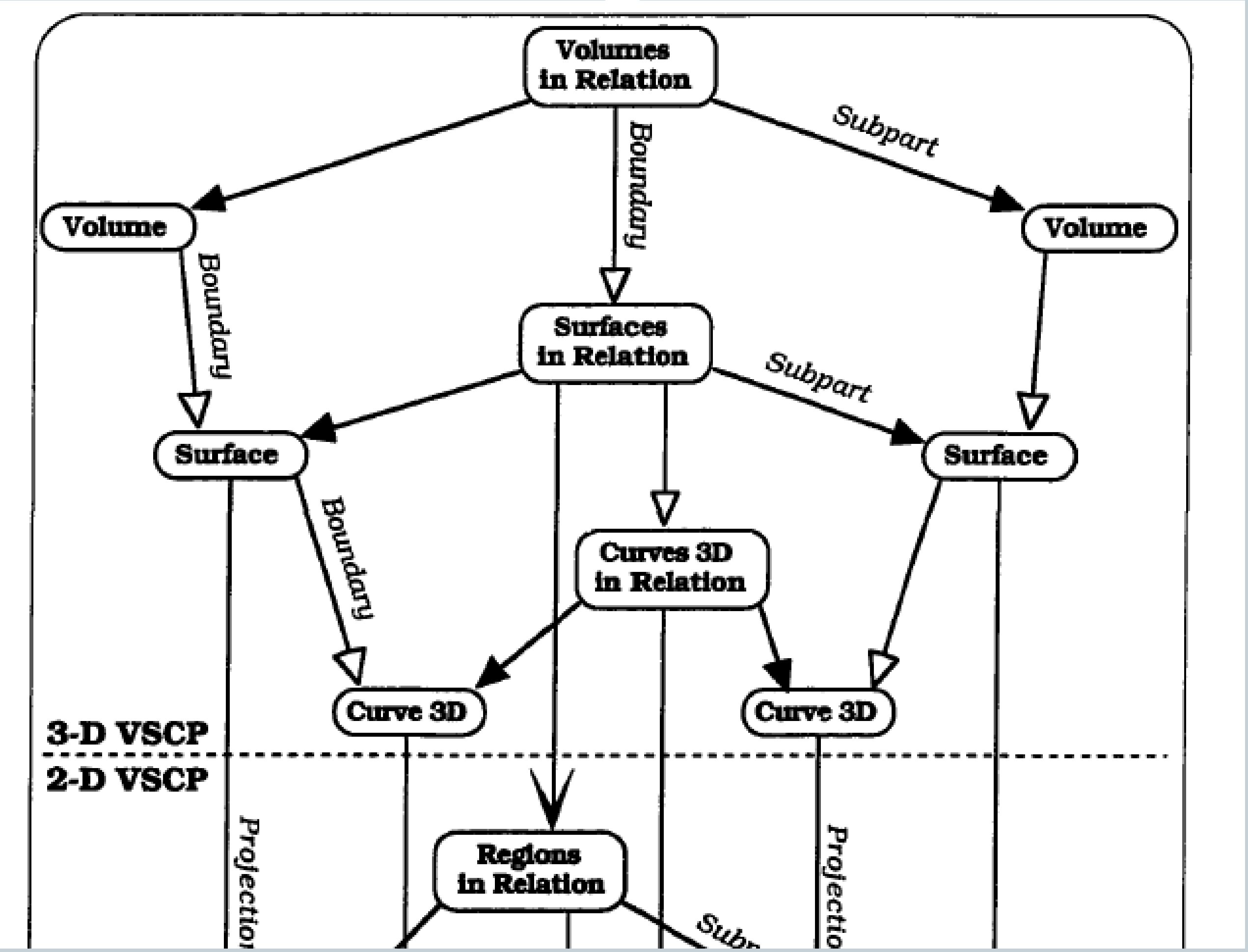


Figure 5-1: Basic VSCP Representation

The VSCP incorporates volumetric, surface interior, boundary curves and vertex geometric and observability into a single, coordinated model.

Relationships in VSCP (1)



Relationships in VSCP (2)

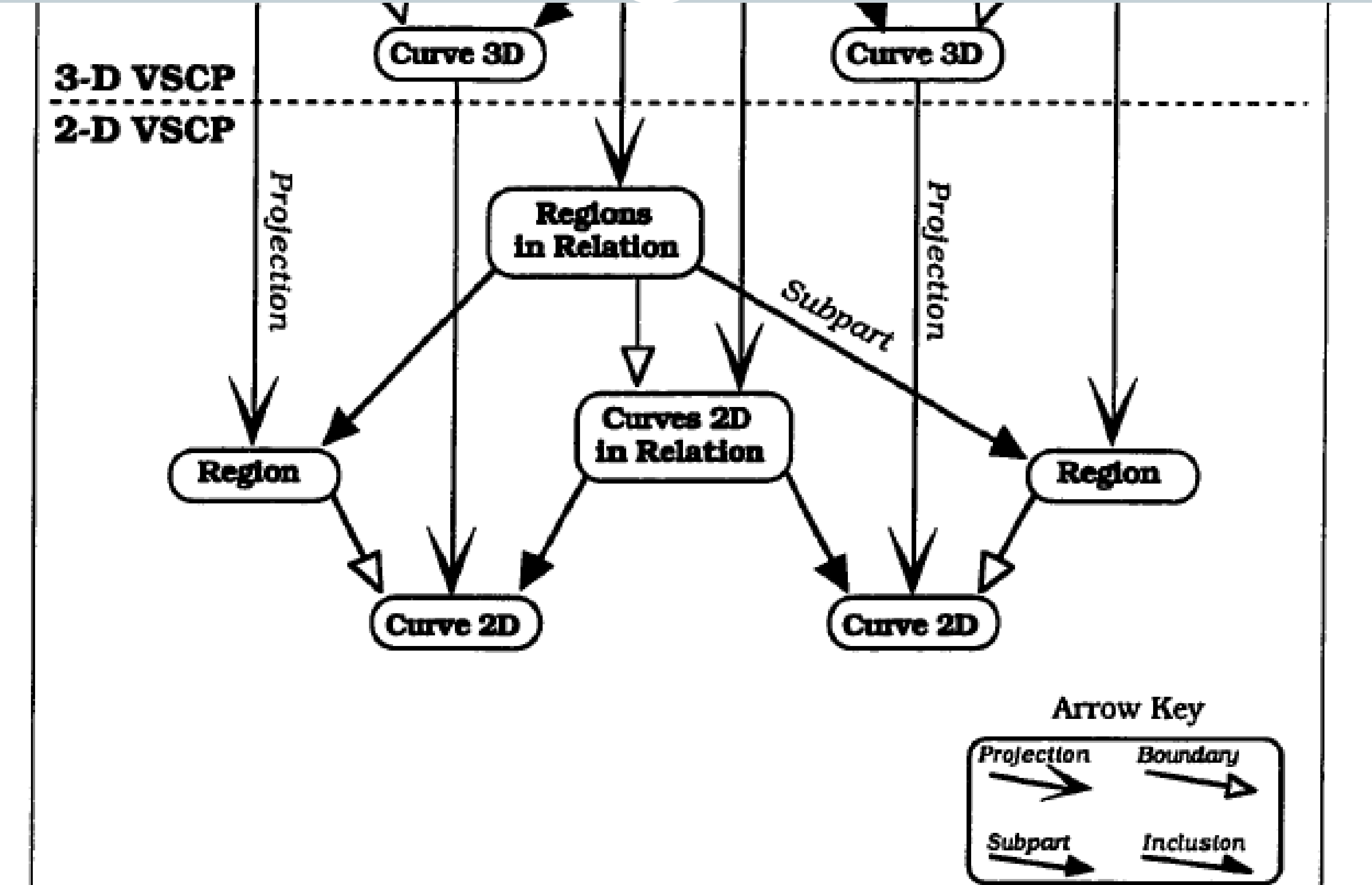


Figure 5-3: Relationships in VSCP

Bayesian Model

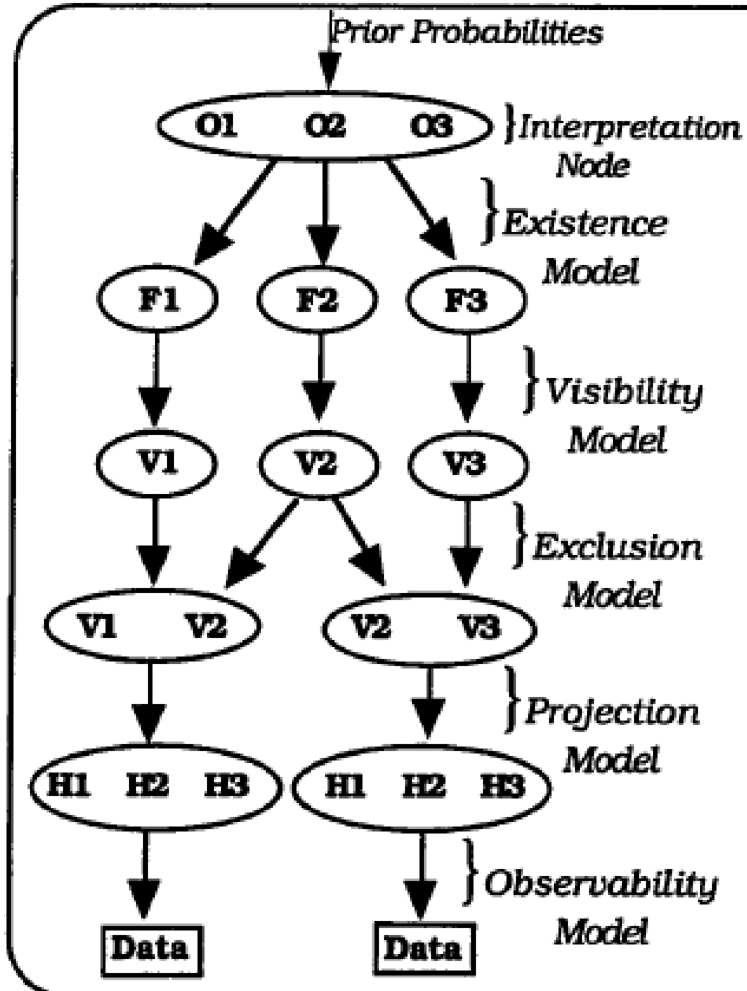


Figure 6-1: Overall Bayes Model

This shows how the models are connected together. Not all models are always constructed. The existence model may be quite complex, reflecting the subpart hierarchy and other relationships of the geometric model. Existence, exclusion and projection tend to be boolean conditional distributions, while visibility and observability contain more general distributions.

Existence Model

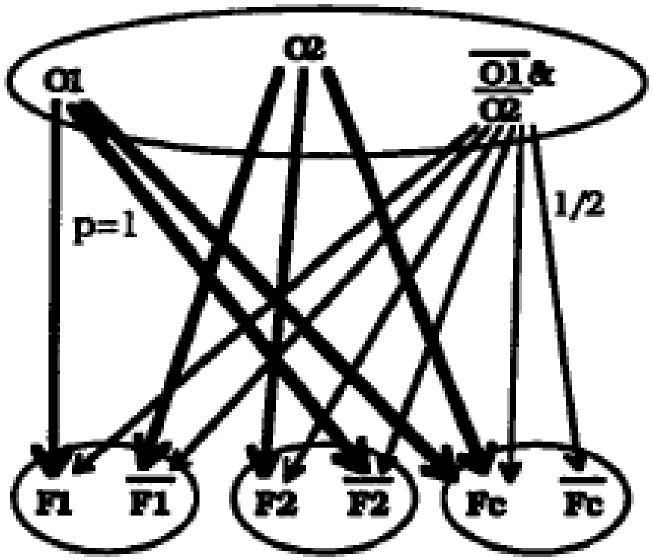


Figure 6-3: Existence Model

Only one of the object hypotheses can be true. A feature exists only if an object gives rise to its existence. F1 is a feature of object O1, and F2 is a feature of O2. Fc is a feature common to both objects.

Visibility Model

Visibility Model

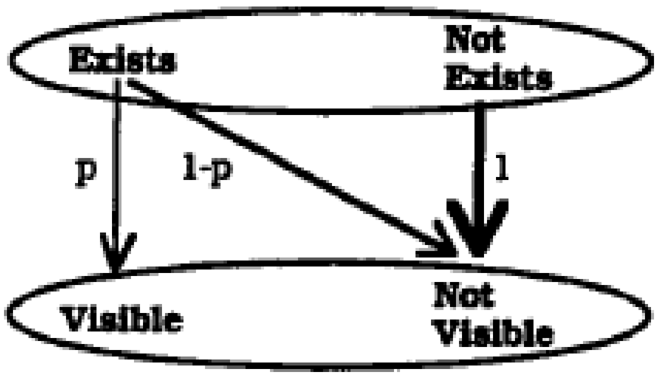


Figure 6-4: Visibility Model

A nonexistent feature is not seen, to the extent that it is caused by this object.

Exclusion Model

Exclusion Model

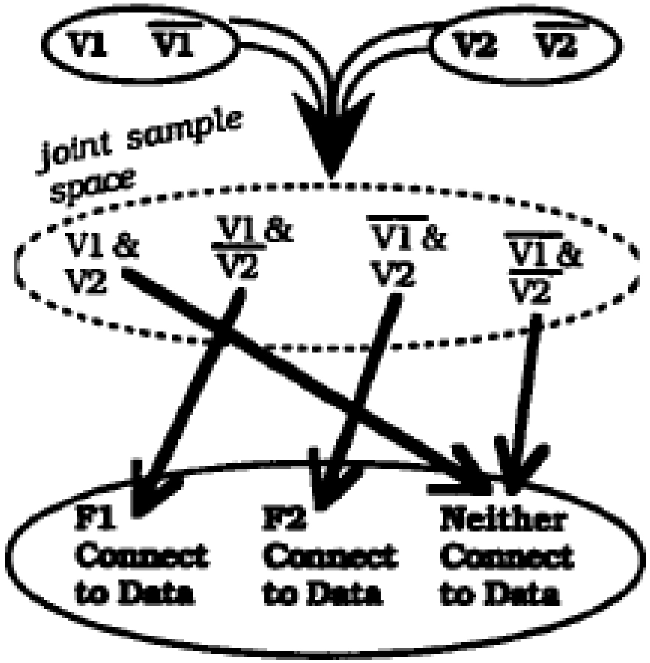


Figure 6-5: Exclusion Model

The two parents are visibility nodes. The child node expresses which of the two features, if either, is actually visible as the particular data associated with the child node. Only one of the two features can be connected to the data. That is, the Data will provide evidential support to only one of hypothesis V1 or V2. The hypotheses V1 and V2 are not mutually exclusive as if they were in the same state space. Only the evidential support is exclusive.

Projection Model

Projection Model

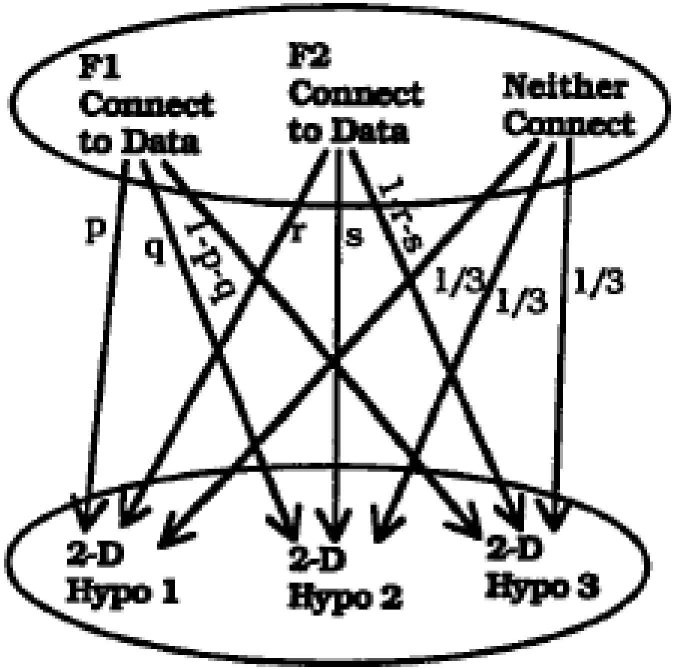


Figure 6-7: Projection Model

The probabilities are assessed as for the observability models. If neither feature connects to the data, a uniform conditional removes the influence of the data on the object belief.

Observability Model

Observability Model

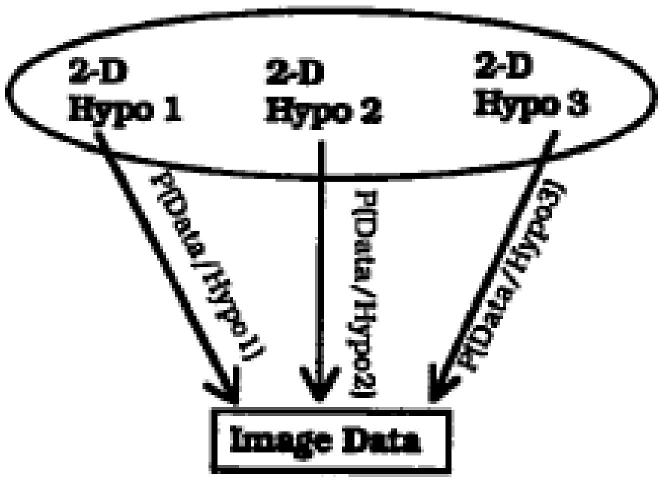
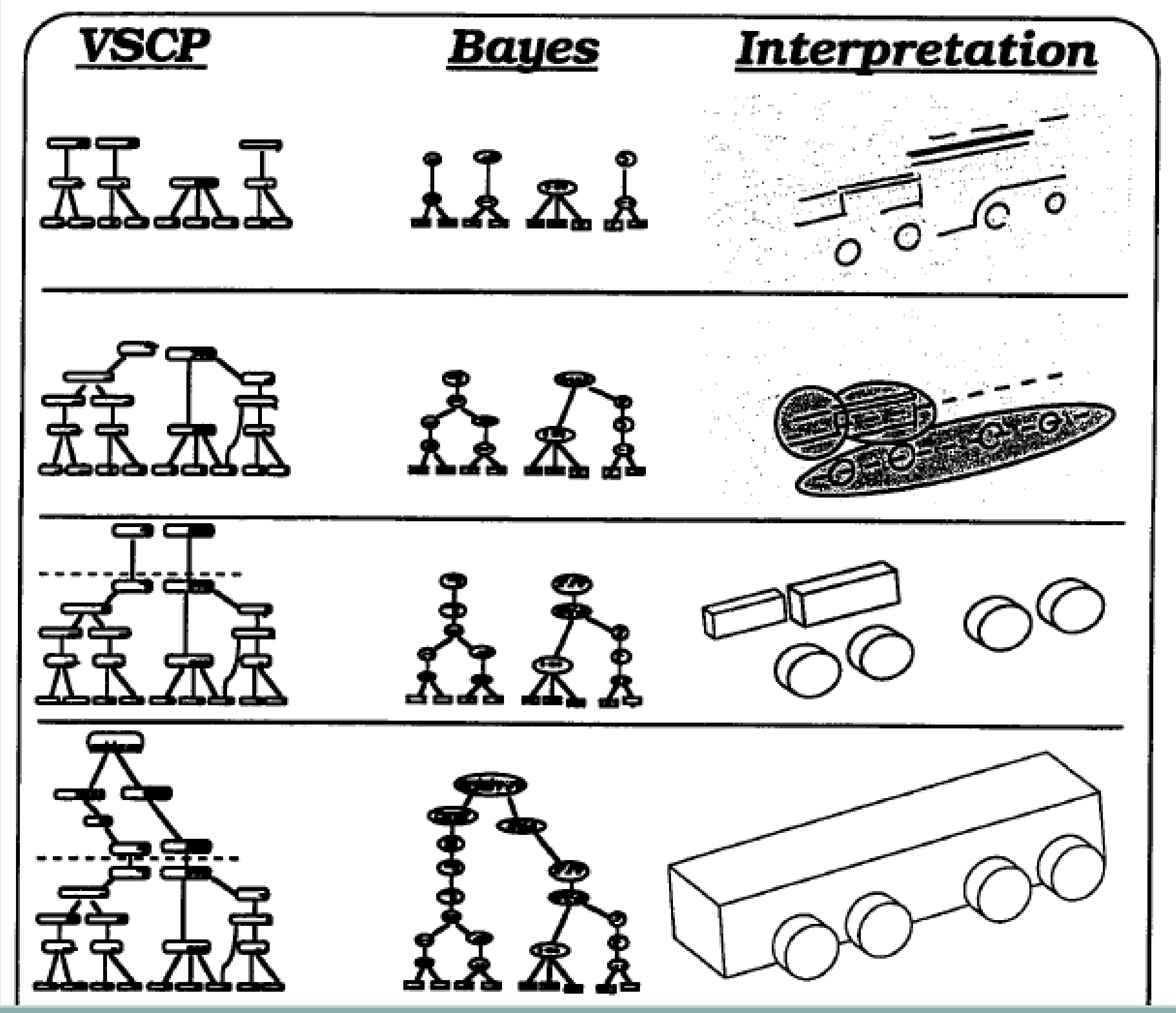


Figure 6-8: Observability Model

The image data are instantiated as being true. Implicitly there exist likelihoods P(Not-Data/Hypo) as well. These form the leaf nodes.

Final Example



Final Example (2)

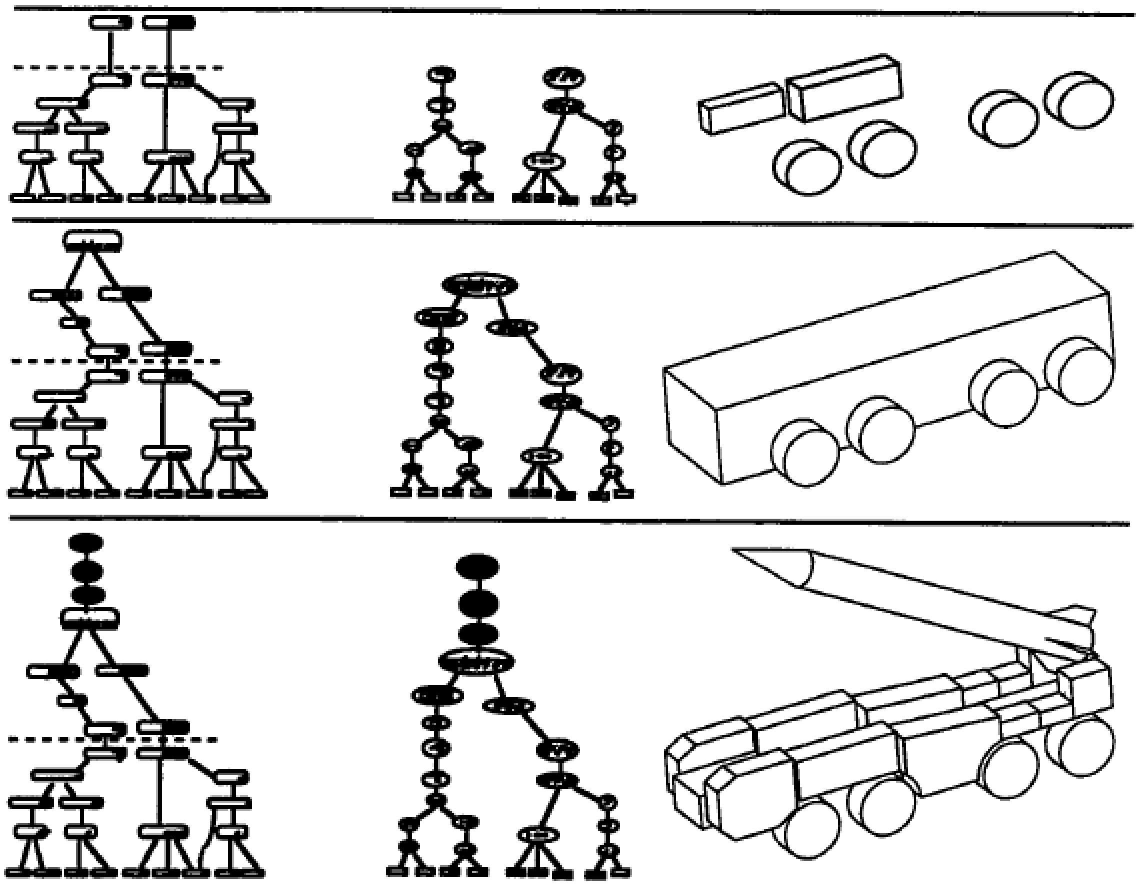


Figure 7-1: Putting it all together

This shows how the instantiated VSCP representation, the Bayes network, and the interpretation all grow in complexity as the image interpretation proceeds from the bottom up. What is not shown is the prediction and verification steps that also occur along the way.

DISCUSSION

"What is the relationship between this dissertation and Ramesh, 1995 ?"

Other Related References

- Error analysis for Feature Extraction, Geometric Inference (Foerstner, 1992 –)