11/5/88. Artificial Intelligence 3. Vision Revision.

① PATTERN RECOGNITION FEATURE VALUES

Area = $\frac{1}{2} \sum_{i=1}^{N} (x_{i+1} y_i - x_i y_{i+1})$

$\begin{vmatrix} x_i & x_{i+1} \\ y_i & y_{i+1} \end{vmatrix}$ $\begin{vmatrix} A & B \\ C & D \end{vmatrix}$ ~~AD-BE~~.

x_i, y_i in perimeter

Perimeter

$$\sum_{i=1}^{N} \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}$$

Eccentricity

$$\frac{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}{\text{Area } 4.} = m_2.$$

Compactness. $\frac{(\text{perimeter})^2}{(\text{Area})}$. $\frac{m^2}{m^2}$.

Mamments.

$\hat{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$ $\hat{y} = \frac{1}{N} \sum_{i=1}^{N} y_i$ over whole mass.

$$\mu_{ab} = \sum_{i=1}^{N} (x_i - \hat{x})^a (y_i - \hat{y})^b$$

$m_1 = (\mu_{20} + \mu_{02}) / \text{Area}^2$

LINEAR CLASSIFIERS.

NEAREST NEIGHBOUR.
MINIMUM DISTANCE

$\underset{\sim}{x}$ = Feature vector.
$\underset{\sim}{M_i}$ = Model vector.

Pick i that minimises $\|x - M_i\|$

From hand classified training.

LINEAR CLASSIFIERS

$V_i(x) = \underset{\sim}{W_i} \circ \underset{\sim}{x} + W_{i\phi}$

= evaluation of x for class i.

$\underset{\sim}{W_i}$ = weight vector of class i.

class i if $V_i(\underset{\sim}{x}) > V_j(\underset{\sim}{x})$ all $j \neq i$.

BAYES CLASSIFIER.

$$= \frac{PR(\underset{\sim}{x} | \alpha) \cdot PR(\alpha)}{PR(\underset{\sim}{x})} = \text{Prob}(\text{class } \alpha \text{ given } \underset{\sim}{x})$$

11/5/88. **AI3.** **Vision**

DERIVING LINEAR CLASSIFIERS

min $d_i(\underline{x})$ over i $\equiv$ min $d_i^2(\underline{x})$

$$d_i^2(\underline{x}) = \| \underline{x} - \underline{m}_i \|^2$$
$$= (\underline{x} - \underline{m}_i) \circ (\underline{x} - \underline{m}_i)$$
$$= \cancel{\underline{x} \circ \underline{x}} - 2(\underline{x} \circ \underline{m}_i) + \underline{m}_i \circ \underline{m}_i$$
$$= -2(\underline{x} \circ \underline{m}_i) + \underline{m}_i \circ \underline{m}_i$$

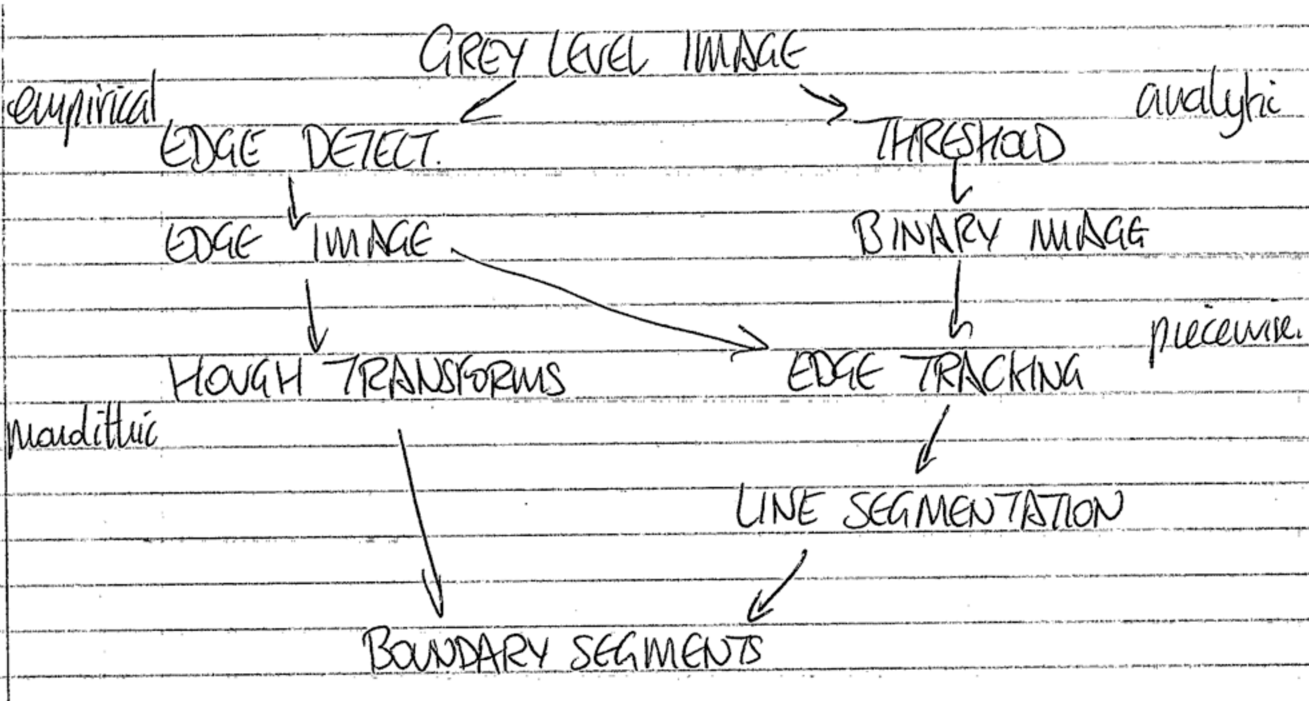
Maximise $\underline{x} \circ \underline{m}_i - (\underline{m}_i \circ \underline{m}_i / 2)$

use $\underline{w}_i = \underline{m}_i \Rightarrow w_{i\phi} = -(\underline{m}_i \circ \underline{m}_i)/2$

Maximise $\underline{\underline{w}_i \circ \underline{x} + w_{i\phi}}$ over i

---

DATA DESCRIPTION



HOUGH TRANSFORMS

Principle: Model directed feature detection to find the parameters that describe a line.

DATA: set of pairs $\{(x_i, y_i)\}$ on many lines.

Model: $x\sin\theta + y\cos\theta = r$ $[r, \theta]$

Algorithm

clear 2D histogram $-\infty \to \infty : R$
$0 \to \pi : \theta$

1 vote per ~~the~~ point ~~on~~ determining a line.

for each (x_i, y_i) evaluate θ and R for each and increment histogram at ~~R,θ~~ (θ, R).

Repeat until enough points.

Many high peaks shows an $(\hat{R}, \hat{\theta})$ of a line has many points running through it.

Verify line at (r, θ) $\hat{x}\sin\hat{\theta} + y\sin\hat{\theta} = \hat{R}$ by checking other point pairs. Problem as where are the end points!

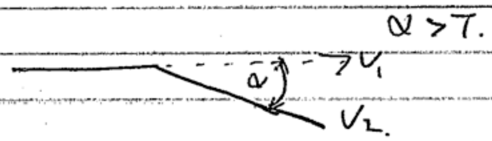
Can have coincidental alignments.

③ 11/5/88

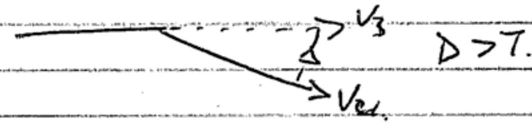
Artificial Intelligence 3. Vision.

FEATURE SEGMENTATION. T = Threshold.

(1) Tracking deviation : (1) angle

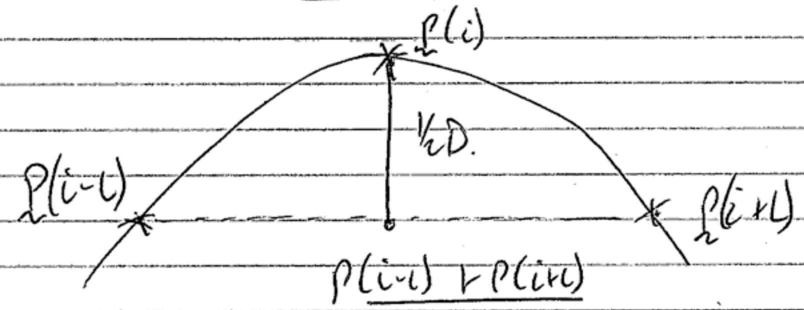


(2) distance



~~Eg~~ $\alpha = \cos^{-1}(v_1 \circ v_2)$.

(2) Malcolm Corner Detector. (Worm).

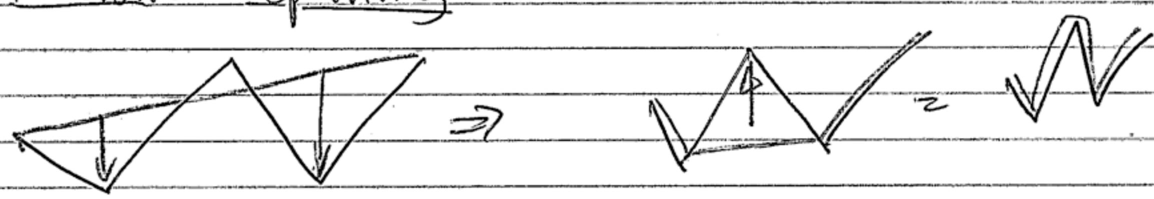


$p(i) = (x, y)$. ~~point~~

Deviation = $D(i) = |P(i-L) - 2(P(i)) + P(i+L)|$

$D(i) > T$.
$D(i-1) < D(i) > D(i+1)$

(3) Recursive splitting.



(4) Merlo's: SEGMENTED BOUNDARY POINTS.

Model θ_i, d_i $i = 1..n$

Data $\hat{\theta}_j\ \hat{d}_j$

Point pairing ~~fst~~ $i \Leftrightarrow j$. mismatch error =

$c_1 \sum |\theta_i - \hat{\theta}_{f(i)}| + c_2 \sum |d_i - \hat{d}_{f(i)}|$

Work thru list of points.

GEOMETRIC MODEL BASED

Line $(0,0) \to (1.0,0)$
hole $(0,2.5)$ @ $(0.5, 0.5)$

Binary Image → tracked Boundary Points
→ Segmented data features → model or data pairings
⇒ object location (translation, rotation).

CO-ORDINATE GEOMETRY.

$\underline{p} = (x_A, y_A)$

Translation : $\underline{p} + \underline{t} = (x_A + t_x, y_A + t_y)$

Rotation : $R(\theta)\underline{p}$ $\quad R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \underline{p}$

$= (x_A\cos\theta - x_A\sin\theta,\ y_A\sin\theta + y_A\cos\theta)$

Combinations : $R(\theta)\underline{p} + \underline{t}$

Different coords:

$R(\theta_w)\ (R(\theta)p + t) + \underline{t}_w$

world object. $\quad$ obj → subcomp

Homogeneous $(R(\theta)p + t) = M,$

$= \begin{pmatrix} \cos\theta & -\sin\theta & t_x \\ \sin\theta & \cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} = M(\theta, \underline{t})$

Inverse M^{-1} $\begin{pmatrix} \cos\theta & \sin\theta & -t_x\cos\theta - t_y\sin\theta \\ -\sin\theta & \cos\theta & t_x\sin\theta - t_y\cos\theta \\ 0 & 0 & 1 \end{pmatrix}$

(4)

11/5/88 Artificial Intelligence 3. Vision.

Line (start co-ord, End co-ord)
hole (centre, size)
corner C (A, B), θ size, θ orientation.

POLYCURVES.
Line Line Arc hole corner. . . _
Hierarchical

subcomponent @ Rotation @ translation
A = line(10) @ (0, π/2)
line (20) @ (π, 1) etc.

Recursive: Build Models with components A.

Matching Algorithm.

REPEAT. - Generate next compatible feature pairing.
- estimate reference frame transformation (1)
- predict other feature positions (2).
- verify other feature positions. (3)
UNTIL SUCCESSFUL MATCH FOUND.

(1): $\hat{\underset{\sim}{t}} = \underset{\sim}{P}_d - R(\hat{\theta})\, \underset{\sim}{P}_m$ $(\hat{\theta}, \underset{\sim}{\hat{t}})$ given.

(2): $R(\hat{\theta})\, \underset{\sim}{P}_m + \hat{t} = \hat{P}_d$

(3): $|\hat{\underset{\sim}{d}}_d \circ \underset{\sim}{d}_d| \geq 1 - e_1 \quad \| \hat{P}_d - \underset{\sim}{P}_d \| \leq e_2.$

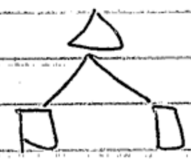
data feature

Predicted image direction & position.

GRAPH MODEL REPRESENTATION.

nodes: = structures
arcs := relations between structures.

△ Node
□ "
— Arc. : "touches"



etc.

Feature type - line, triangle curve.

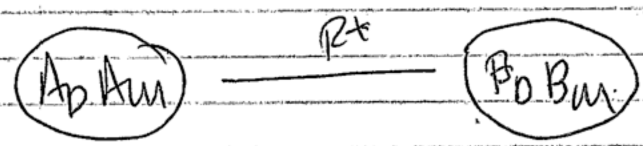
ABSTRACTION: relevant properties.

= "PARALLEL RELATION".

ie B & C can be swopped!?!

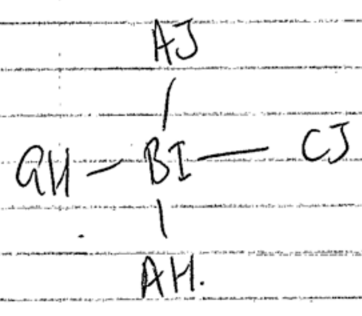
Use of "association" graph. for matching.

Revise "maximal cliques".



- Connect nodes if arcs between original nodes are compatible
- connect if no ~~nodes~~ arcs between original nodes.

Maximal cliques: ~~the~~ fully connected nodes, "cannot add another to set".



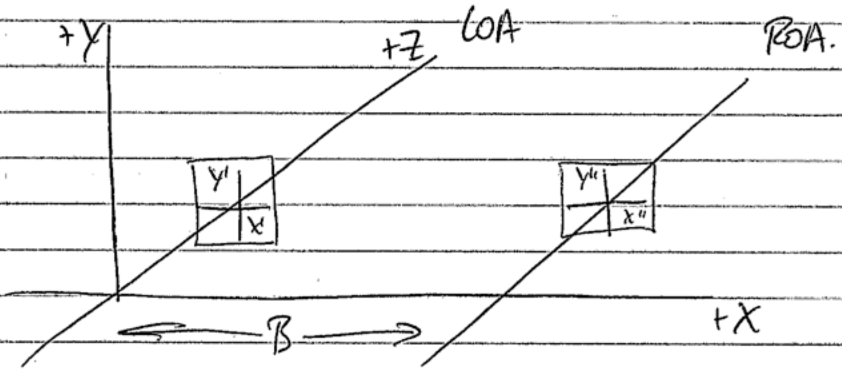
11/5/88 Artificial Intelligence 3. Vision.

BINOCULAR STEREO

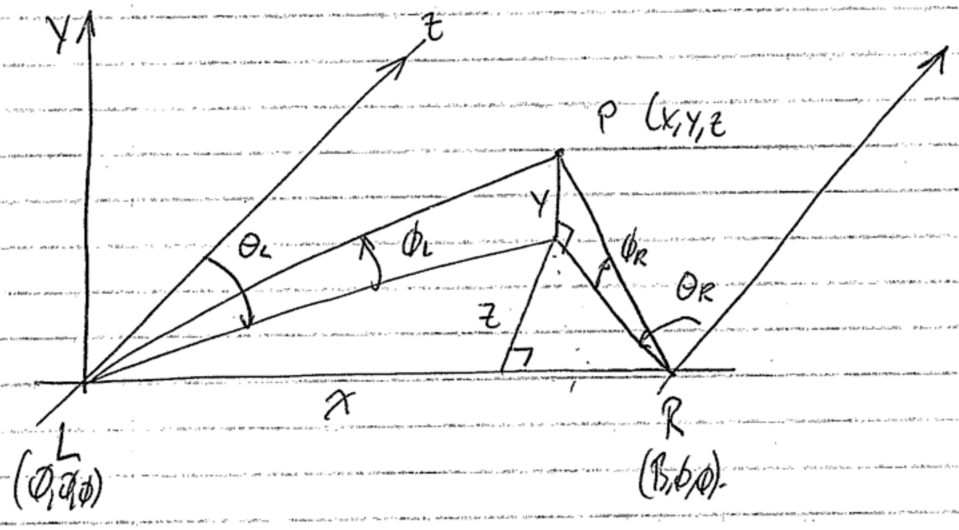
Principle: Triangulation.

Simplifying Assumptions.

- Optical axes aligned with +z plane.
- Image plane axes parallel to global x and y axes
- Each pixel represents constant angular displacement from optical.



Problem: Given a feature in both images, where is it in 3D.



$\tan \theta_L = \frac{X}{Z}$ $\tan \theta_R = \frac{X-B}{Z}$.

$\therefore \quad Z = \frac{B}{\tan\theta_L - \tan\theta_R}$

$\therefore \quad X = Z \tan\theta_L = \frac{B \tan \theta_L}{\tan\theta_L - \tan\theta_R}$

$\tan \phi_L = \frac{Y}{\sqrt{X^2+Z^2}}$?

$\therefore \quad Y = \frac{\tan \phi_L \, B}{\cos\theta_L \left(\tan\theta_L - \tan\theta_R\right)}$?

(6)

11/5/88 Artificial Intelligence 3 — Vision.

Correspondence Problem

Which feature in the left image matches which feature in the right image?

- often considered the key problem in stereo vision.

Correspondence Constraints.

Epipolar:- Feature in left image must pair with a feature lying on a given line in right image.

Uniqueness: A feature in one image can match at most one ~~image~~ feature in the other image.

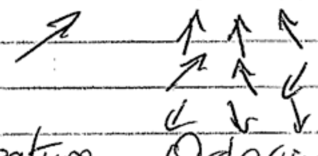
100/50   100/50/90

L   R

EDGE CONTRAST MAGNITUDE.

Match edge fragments with nearly same edge strength

Edge Direction. edge fragments with nearly same orientation match.



Feature Ordering | | | | | |

1 2 3   A B C.

Maintain left -> right relative order of paired features (can miss some).

Local surface smoothness.

Only allow pairings whose resulting z values are nearly those of nearby neighbours.

MARR'S COMPUTATIONAL THEORY

I/O description; formal description between inputs & outputs; Constraints; maths (triangulation)

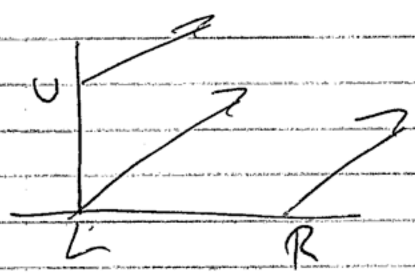
3 Levels of Program.

	multiplication		m/c stereo	
Computational Theory	number theory		I/O signals Constraints.	Biological Stereo
Algorithm	successive addition	mult, shift, add	Shelfield PMF	Parallel/ co-operative.
Implementation	μ progs.	VLSI.	C Prog	wetware

Application

Tri-nocular Stereo.

Reduced error.



Algorithm: ① find feature f_1 in image$_L$.

② predict epipolar line e_1 in image$_2$ find feature f_2 on it

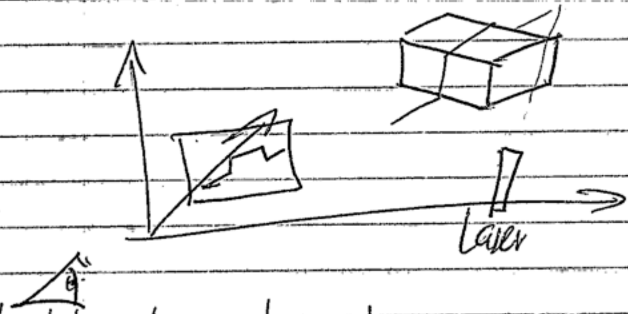
③ predict epipolar e_1 & e_2 in image$_3$. See if an image is there.

f_3 must be there if f_1 and f_2 are two views of the same feature.

11/5/88. Artificial Intelligence 3 Vision

Marr's Computational Theory: Application.

Laser Stripper (Triangulation).



- light plane has known 3-D equation.
- each point on observed stripe has known 3D equation
- intersection gives 3D point location.

Marr's Theory: Summary.

- Can deduce 3D position by triangulation of same point as seen in 2 images.
- Can build up 3D image by triangulating many points.
- Key problem is how to pair points from 2 images.

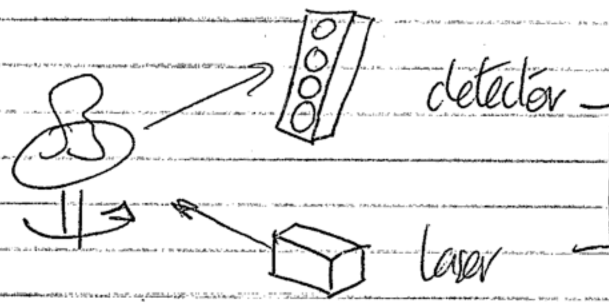
3-D OBJECT RECOGNITION

FAUGERAS - Geometric models (GM) 1983

ACRONYM (Brooks) GM's Graph matching 1981
symbolic algebra
intensity data.

IMAGINE (FISHER) GM's
Model Directed Matching 1983.
OCCLUSION.

FAUGERAS



rotate & elevate table for full scan.

PRIMITIVES.

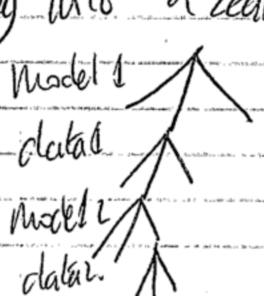
Collection of patches: $\{(c_i, n_i)\}$.
(center, normal)

least squares fit of local neighbourhoods of pixel

MATCHING. Model $\{(C_i, N_i)\}$

1: form patches for data $\{(\hat{c}_j, \hat{n}_j)\}$. (SCAN)

2: Combinatorial paring of 1st 2 model-data pairing into a search tree.



model patch a → data patch b
" " c → " " d.

8

11/5/88

Artificial Intelligence 3 — Vision

3: Use local heuristics to prune search space.

a). Similar orientation between patches.

$n_a \circ n_i \doteq \hat{n}_b \circ \hat{n}_d$

$\cos\theta_m = \cos\theta_d$

$\theta_m \approx \theta_d$

b) Similar distance between patches.

$\|c_a - c_c\| \approx \|\hat{c}_b - \hat{c}_d\|$

4: Estimate Rotation & Translation $\boxed{R(\theta)_R + t}$

Least squares estimate of T.

$E = \sum \| T_{ni} - \hat{n}_{f(i)} \|$

$\uparrow$ data normal.

5: Predict n-2 other patch positions.

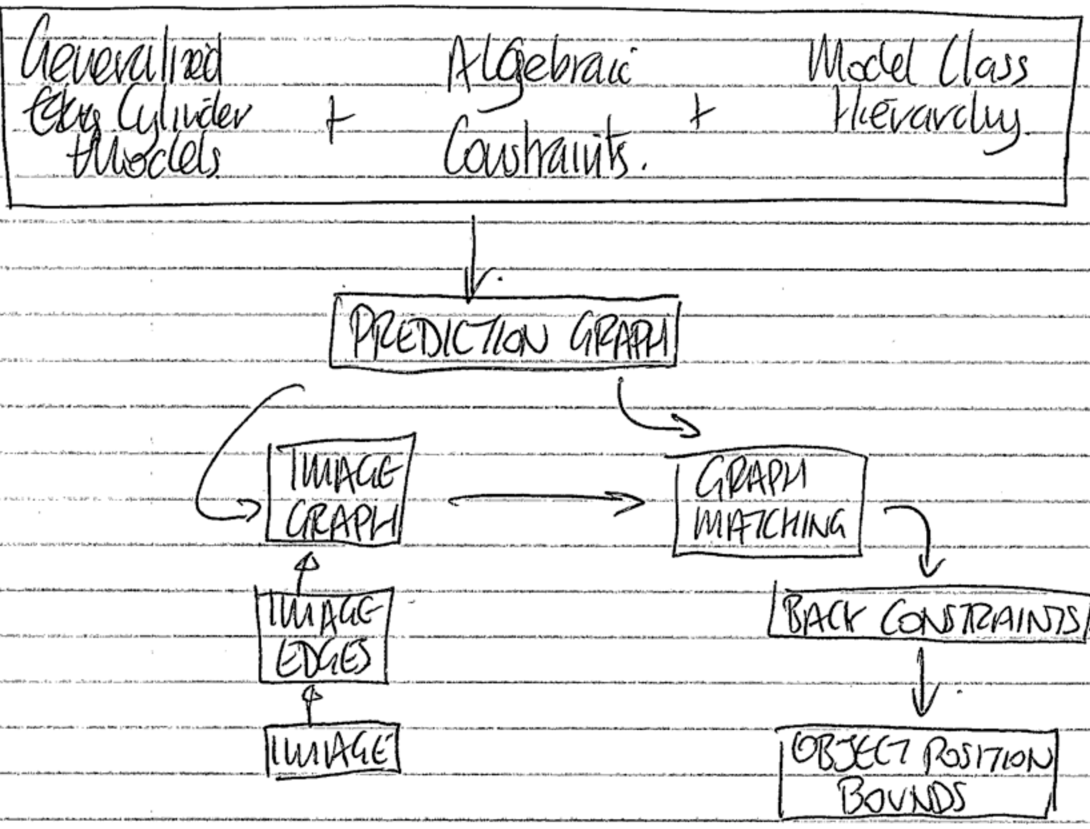
$\text{Pred}_i = (T_{ci}, T_{ni})$

6: for each Pred_i find some $(\hat{c}_j, \hat{n}_j)$ that is close enough

7: If enough pairings & cumulative error low $\Rightarrow$ match

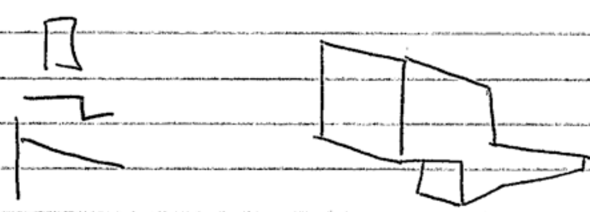
8: Refine ~~old~~ position estimate with all pairings.

ACRONYM (Brooks).



PRIMITIVES

Cross-section
Axis
Sweeprule



Structured Models.

Airplane: Body @ T_1, (Wing @ T_2 Rowing @ T_3 Engine @ . —

Each object has local reference frame, T_i maps from Global to local

T_i = (Rotation$_i$, Translation$_i$) $\boxed{(\theta, t) \text{ like in}}$

Model variables

Parameters of generalized cylinder for body.
(Bodylength, bodywidth)
L. W.

Constraints → $20 \le L \le 300$.
→ $6 \le W \le 30$ etc.

9
11/5/88 Artificial Intelligence 3. Vision.

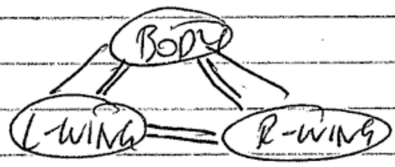
PREDICTION GRAPHS.

Tries to deduce main appearance features (2D) from the 3D Model.

Nodes: - Ribbon
- projection of gen cylinder
- spine length & cross-section width f^n.

ARCS $\overline{\overline{\theta}}$ relative spine orientation (θ)
— connectedness.

Tries to make prediction holding over many viewpoints.



L: $\frac{\hat{L}.z}{L \cdot \cos\theta_y}$

W: $\frac{\hat{\omega}.z}{w}$

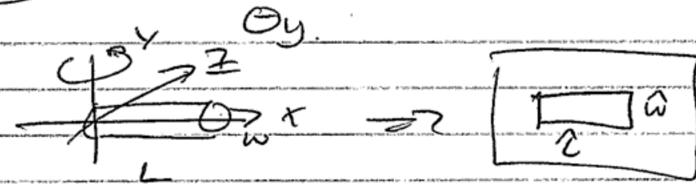
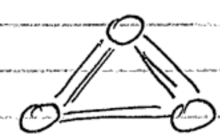


IMAGE GRAPHS

Ribbons: (NODES): length $[\hat{l}_i, \hat{l}_s]$
width $[\hat{w}_i, \hat{w}_s]$
axis direction $[\hat{d}_i \; \hat{d}_s]$
expected error included in data values.

ARCS : — connects
= angle between.



CONSTRUCTION

1: Edge detection (Canny Operator) & linking
2: Ribbons hypothesized by fitting to edges.
- Ribbon shapes & sizes to be searched for from prediction graph (INVARIANTS).
3: Arcs from adjacent ribbons.

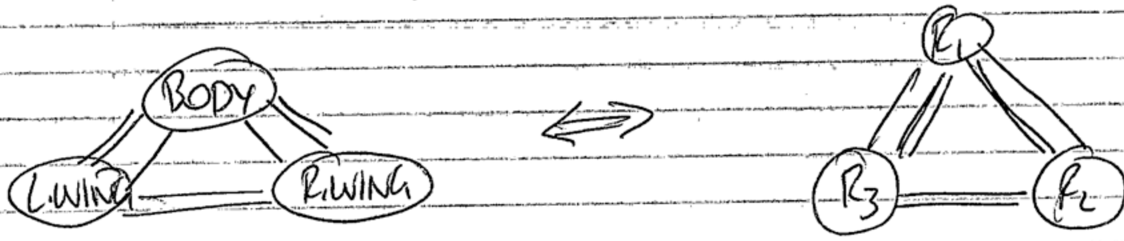
$L: \frac{\hat{L}Z}{L\cos\theta_y} \Leftrightarrow$

$\hat{L}: [98, 102]$

MATCHING (Subgraph isomorphism).

1- Additional Arc labelling:-
  MUST any match must have this node.
  EXCLUDE no match may have BOTH ENDS

2- Invariants paired with measurements & sent to constraint analysis.



PREDICTION GRAPH   IMAGE GRAPH.

CONSTRAINT ANALYSIS

- used to ensure geometric relations are consistent.
- deduce or refine parameters

- Use algebraic relations to estimate ranges for all variables

$$X \to [\mathrm{INF}(X), \mathrm{SUP}(X)].$$

Reject if $\mathrm{INF}(X) > \mathrm{SUP}(X)$.

ALGEBRA

$\mathrm{SUP}(A+B) \le \mathrm{SUP}(A) + \mathrm{SUP}(B)$
$\mathrm{SUP}(A-B) \le \mathrm{SUP}(A) - \mathrm{INF}(B)$
$\mathrm{SUP}(A*B) \le \mathrm{SUP}(A) * \mathrm{SUP}(B)$
$\mathrm{SUP}(A/B) \le \mathrm{SUP}(A) / \mathrm{INF}(B)$

$\mathrm{INF}(A+B) \ge \mathrm{INF}(A) + \mathrm{INF}(B)$
$\mathrm{INF}(A-B) \ge \mathrm{INF}(A) - \mathrm{SUP}(B)$
$\mathrm{INF}(A*B) \ge \mathrm{INF}(A) * \mathrm{INF}(B)$
$\mathrm{INF}(A/B) \ge \mathrm{INF}(A) / \mathrm{SUP}(B)$

BACK CONSTRAIN ensure that values given are confined correctly between the constraints of the values for the model parameters respectively.

for L, W, α, Z. $\frac{\hat{L}\cdot Z}{L\cos\theta_y} = \alpha \qquad \frac{\hat{W}\cdot Z}{W} = \alpha.$ $A \le L \le B$
$C \le W \le D$
$V \le W/Z \le W$

13/5/88.

Artificial Intelligence 3 — Vision

CONSTRAINT ANALYSIS

- may be tighter bounds not ~~given~~ found due to algorithmic limits.

This is a "CONSTRAINT ~~DECIS~~ MANIPULATION SYSTEM" which is a "PARTIAL DECISION PROCEDURE"

Algebraic techniques are not powerful enough to deduce tighter bounds that will show inconsistency.

SUP(x) < INF(x) ⇒ INCONSISTENT ⇒ DEFINITELY INCONSISTENT.

SUP(x) $\not<$ INF(x) ⇒ CONSISTENT ⇒ MAYBE CONSISTENT.

INTRINSIC IMAGES.

Image composed from raw image contain the properties intrinsic to the scene.

INTENSITY DEPTH SURFACE ORIENTATION
REFLECTANCE ILLUMINATION TEXTURE DENSITY
RELATIVE MOTION.

A useful intermediate representation => scene understanding.
Also for practical theory.

MARR'S PHYSIOLOGY BASED THEORY OF HUMAN VISION

13/5/88 Artificial Intelligence 3 Vision

RAW PRIMAL SKETCH

Significant events of image.
- local changes of intensity
- reflectance
- illumination
- surface depth
- surface orientation (HIGHLIGHTS)

In theory can re-construct image from this data.

REPRESENTATIONS

Edgelet	Position
Bar	orientation
Blob	contrast
Terminations	size (scale).

edgelet @ (60,30), contrast (40) orient (↑) etc.

Mask sizes: 4x4, 9x9, 18x18, 36x36.

CONSTRUCTION:

IMAGE → GAUSSIAN SMOOTHING ② → LAPLACIAN ③ → ZERO CROSSINGS ④ → MULTI-SCALE SUPPRESSION ⑤ → DESCRIPTION. ⑥

2 – noise reduction averaging by 2D Gaussian convolution.
3 – Omni-directional gradient operator.
4 – edges at "maximum" changes of intensity

FULL PRIMAL SKETCH

- larger scale structure of image.

① Grouping & description at multiple scales.

② Detect boundaries in image that may correspond with shape changes in scene.

Intrinsic properties of token: brightness, size, orientation

spatial arrangement of token: distance, density, relative

The essence of the higher primitives in the full primal sketch is their ability to capture a wide range of image items as a group or token and their ability to be arranged into groups and boundaries.

lowest level - intensity changes & terminations
second " - oriented tokens grouped together
third " - boundaries formed by differences of groups at second level &

complexity & degree of organization of 3 scales above.

2½D SKETCH.

	FROM
- Depths	stereo, motion, optic flow.
- Orientation	shading, optic flow, texture
- Occluding contours	full primal sketch
- orientation discontinuities.	

STEREOPSIS

Full primal sketch entities.
Constrained by:-
- Compatible descriptions
- uniqueness 1:1 map
- continuity depth variance smooth.

ALGORITHM.

coarse matches, estimate local depth
reduces range of matches at finer scale (4).
⇒ sparse depth at zero crossings.

Surface approximation

Surface shape can't vary much if no zero crossings.
"Thin plate model" - tries to minimize bending.

13/5/88 Artificial Intelligence 3 Vision.

DECOMPOSITION OF 2½D SKETCH

: Generalised Cylinders.
- segment contour at concavities
- connect neighbour points => regions
- describe region by major axis
- assembly model using major axes together

MODEL REPRESENTATION

Criteria.

- Accessibility — Model features
- Scope — distinct modeled objects
- Uniqueness — " models.
- stability — small obj changes -> small model change
- sensitivity. — small & large model features.

Scale based subcomponent refinement.

Human -> arm -> forearm -> Hand

catalogue linked via

Specificity index — more detailed models.
Adjunct " — subcomponents
Parent " — super "

RECOGNITION.

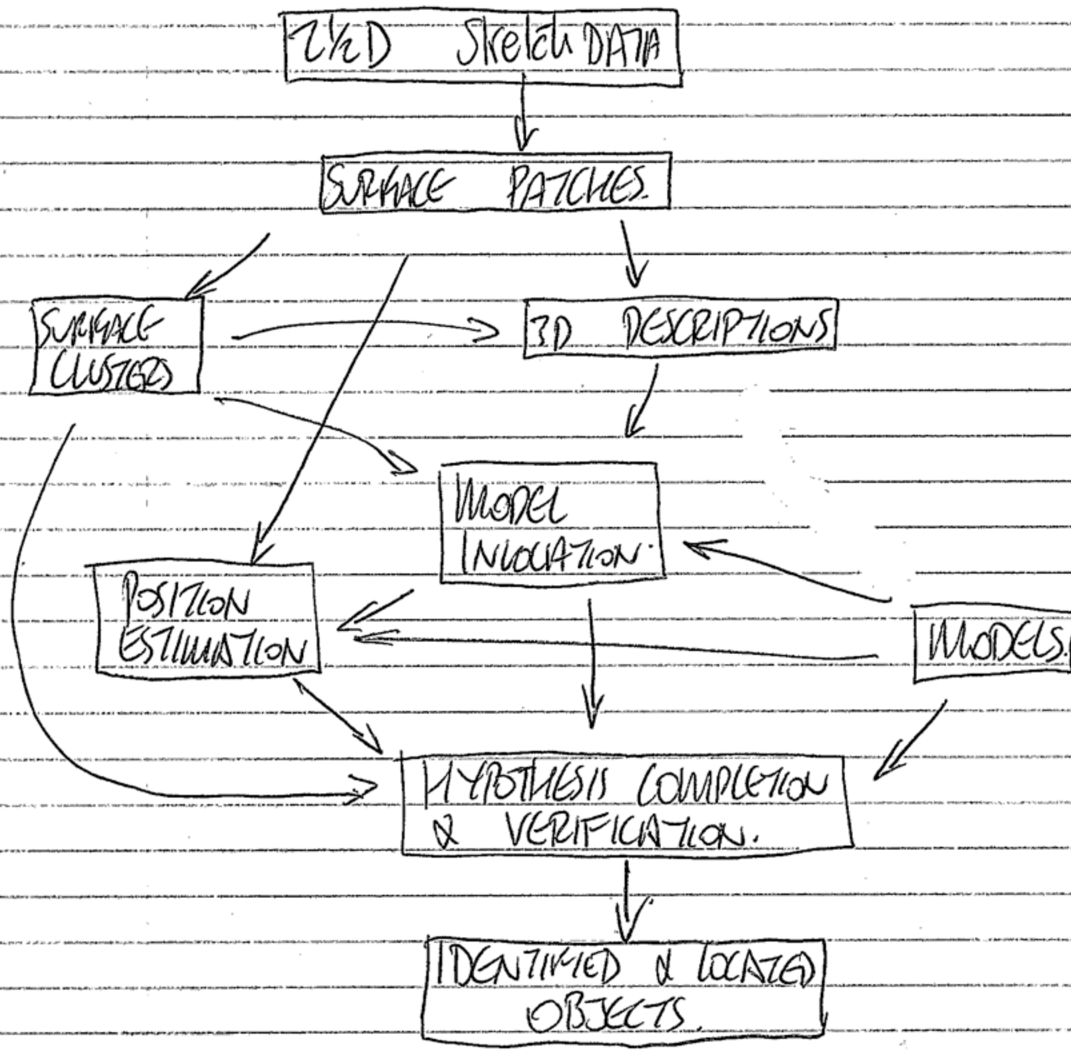
-> Match likely features to model
refine position given pairings
segment or derive new features

11/5/88 Artificial Intelligence 3 Vision

~~Subclass specialization~~

~~Specialise by adding new constraints to refine and split classes. (c.f. classifier trees.)~~

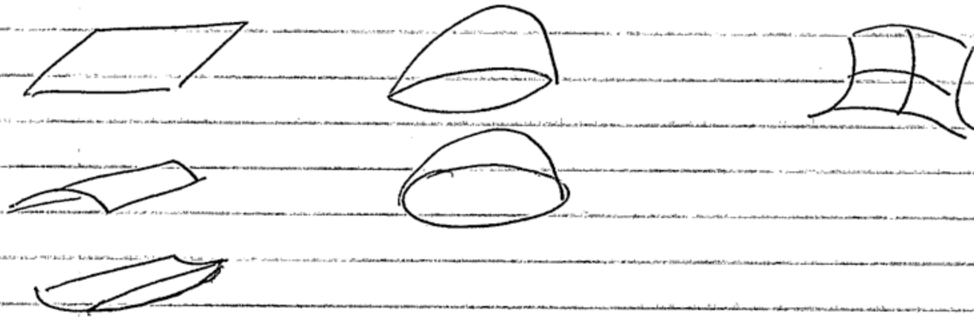
IMAGINE. (BLOB FISHER)



PRIMITIVES

Surface patches with local reference frames
- fully hierarchical.

variables represent degrees of freedom in ref. fr. transformation.



DATA REP. SURFACES

Depth & surface normals.

Patches:

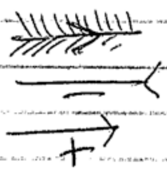
depth discontinuities
orientation "
curvature "

Descriptions:

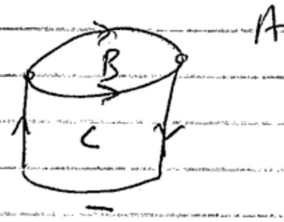
curvature axes
surface area, elongation.

Surface Clusters:

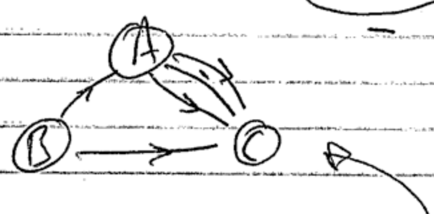
Graphs: nodes = regions
arcs = boundaries



occluding
concave
convex



1: Form:-



Stp 2: Splitup:- 1: Remove concave arcs. —
2: remove occluding arcs except those corresponding to

↑ Guzman/Waltz/Clowes.

13/5/88 Artificial Intelligence 3 Vision.

Q – Want larger surface clusters as CONTEXTS for recognizing larger model assemblies

A:– Group combinations of overlapping surface clusters.
⇒ good structural correspondence
(possibly combinatorially explosive!)

MODEL INVOCATION

– Usually selection in large d'base.

USUAL METHODS
- exhaustive comparison
- property-based discrimination
- probabilistic measures

BUT
- Non incremental
- doesn't integrate different evidence types.

WE WANT:

- multiple evidence types integrated
  * simplification, elaboration, inhibition, association properties etc
- "Fail soft"
  exclusion
- Efficient, pre-attentive.
  * no control flow
  * parallelizable.
- context focussing

MODEL MATCHING

(1) Position Estimation

Initial pairings of model & data surfaces from invocation
3 rotation & translation degrees of freedom

(2) Visibility Reduction

Status of all features predicted by generating an image using models & position estimate, then analysing.

Status:- backfacing
tangential
fully self obscured
partially self obscured
fully visible.

6(3) <u>Evidence location</u>

Model + predicted position
⟹ constrained search
⟹ verified external occlusion.

13/5/88 Artificial Intelligence 3 — Vision.

Systems Discussion:

Image Processing

- lots of image level operation.
- seldom perfect.
  * can lose or corrupt data.
  * expensive.

Pattern Recognition

- many usable 2D features. (more better).
  * cost of extraction clear.
- many classifiers as none give decent results for incorrect data.
- must minimise data inaccuracy through collection.

Geometric Model Recognition (vs Hough transforms).

- structural model ⇒ less ambiguous
- general reusable model primitives.
- model position invariant.

Graph Model Matching.

- represents matchable features
- formal criteria
- match theory, elegant & complex
- Symbol manipulation ideal for computers

MARR's Computational theory

- can deduce 3D position by triangulation of same point as seen in 2 images.
- can build up 3D image by triangulating many points.
- key problem is correspondence problem

FAUGGRAS

+ Good
  - 3D model for any viewpoint
  - irregular objects
  - automatic model building.

- Weak
  - planar observations
  - non structured objects
  - Matching very incomplete
    ⇒ weak 3D understanding.

ACRONYM

+ Good
  - shape models predict image features.
  - graph & geometric consistency.
  - model subclass & subcomponent hierarchies.
  - constraint analysis
- Bad
  - generalised cylinder not universal primitive
  - symbolic algebra very slow. & no good for complex non-linear constraints
  - prediction techniques weak on occlusion
  - not demonstrated. (Airplane scene only)

MARR's Theory of Human Vision.

+ - relates well to psychological & physiological evidence.
  - explicit assumptions & constraints
  - complete, sophisticated
  - complete object representation
- - many missing details
  - not fully tested
  - generalised cylinders not complete representation (face)
  - colour & motion?

IMAGINE

+ - variety of structured primitives in structured models
  - surface clusters contexts for evidence accumulation
  - investigation of model invocation
  - attempt to find ALL evidence
  - occlusion analysis.
- - position estimation - Hand segmented test data.
  - failure when evidence weak

P3 for notes.

AI3 Vision Module

Robert Fisher
Department of Artificial Intelligence
Forrest Hill C11 phone 2553

rbf@edai.

Objective

To extend the student's understanding of the visual process as related to both human and machine vision. In particular, to extend the material presented in AI1 and AI2 to cover: Marr's theories, practical machine vision and theoretical high-level vision systems.

Necessary prerequisites

- AI 2 vision module
- high school: geometry, trigonometry, physics, programming concepts

Helpful prerequisites: AI 1 vision module and elementary calculus, signals, systems, photography

Texts and Other Materials - from Thins (53 South Bridge) or Department Library
[main text]: Charniak and McDermott, "Artificial Intelligence", chapter 3
[main text]: Frisby, "Seeing" illusion brain + mind.
70. [reference] Ballard & Brown, "Computer Vision"
Additional reading list - see handouts
Lecture note photocopies

Assessment

75% - 1 1/2 hour exam (in June)
25% - team practical projects (6 units)
Lab notebooks should be kept, with periodic requests for review.

Course Timetable

2nd term: 9 weeks class
Thurs 1-3 - Lecture
Tutorials - Fri 12-1 on alternate weeks, handed out in advance

Planned work per week (hours)

lecture: 2
tutorial: 1
reading: 1-20 hrs
practicals: 3 (average over 8 weeks)

Practical Notebook:

Contents
- What was done
- How it was done (briefly)
- Results (Numbers, Pictures, Program Listings)
- Answers to questions

Review
- Give to me after each practical finished. I'll mark & return quickly.

Lecture	*Topics*
1	Introduction: Motivations Images Physics Optics
2	Pattern Recognition: Image Processing review Canny Edge Detector Binary Pattern Recognition vision: Measurements, Features, Classifiers
3	2D Model-Based Image Understanding Geometry Boundary segmentation, Model representation & matching LFF
4	Image Intensity Understanding: Intrinsic Images
5	Marr: Primal, 2 1/2D sketch, High level
6	Stereo: Constraints, Practicalities, Structured Light
7	3D Model-Based Image Understanding
8	ACRONYM
9	Student Presentations: Model Invocation Physiology for Visual processing: retina, Hubbel & Weisel, Visual Cortex Hardware for Vision Illusions & link to physiological processes Practical 2D vision systems: Mero & Perkins

Reading List
AI 3 Vision Module

Main Texts

Charniak and McDermott, "Introduction to Artificial Intelligence"
Frisby, "Seeing", Oxford University Press, Oxford, 1979,

Papers (in South Bridge Library)

Bolles, R.C., Horaud, P., Hannah, M.J., "3DPO: A Three-Dimensional Part Orientation System", IJCAI-8, pp1116-1120, 1983.
Bolles, R., "Locating Partially Visible Objects: The Local Feature Focus Method", AAAI 80 & SRI TN #223, Aug 1980.
Brooks, R., "Symbolic Reasoning Among 3-D Models and 2-D Images", AI vol 17, pp285-348, 1981.
Faugeras, O.D. et al, "Towards A Flexible Vision System", in Pugh (ed) Robot Vision, pp 129-142, 1983.

Other General Vision Texts

Ballard, Brown, "Computer Vision", Prentice-Hall, New Jersey, 1982
Horn, B.K.P., "Robot Vision", MIT Press, 1986
Batchelor, B. G., Hill, D. A. and Hodgson, D. C., "Automated Visual Inspection", IFS (Publications) Ltd, UK, North Holland, 1985

Other References - Image Processing

Pratt, "Digital Image Processing", John Wiley & Sons, New York, 1978
Rosenfeld, Kak, "Digital Picture Processing", Academic Press, New York, 1976

Other References - Artificial Intelligence Vision

Duda, Hart, "Pattern Classification and Scene Analysis", John Wiley & Sons, New York, 1978
Cohen, Feigenbaum, "The Handbook of Artificial Intelligence, Vol III", Pitman, 1982
Faugeras, O. (ed), "Fundamentals In Computer Vision - an advanced course", Cambridge University Press, 1983.
Nevatia, "Machine Perception", Prentice Hall, New Jersey, 1982

Other References - Robot Vision

Pugh, A. (ed), "Robot Vision", IFS (Publications) Ltd, Springer-Verlag, 1983.
Aleksander, I. (ed), "Artificial Vision For Robots", Chapman & Hall, 1983.

Dear students:

Professional Ethics

I am concerned about sharing of practical results and problem solutions. At times I have found that:

(1) no acknowledgements given for who did the actual measurements, and
(2) minor alterations to the data were made (or clumsy copying).

I appreciate that you are all busy and there is not time to do everything, but:

(1) professional ethics demands acknowledgement of borrowed material, and you are all mature enough to follow this discipline,
(2) fiddled results helps no future users of the results and gets detected eventually pretty often anyway (especially in such a small class), and
(3) here, you get more credit for understanding than you lose from not doing one part of a problem. And usually little credit is lost when you acknowledge using other's results. In the research/working environment, there is usually no concern at all provided you ask the source and acknowledge them.

Lastly, by not actually doing the practicals, you are missing many important learning experiences not directly related to the exercise itself. So, let me encourage you to do them.