Computer Vision Story

Mubarak Shah

http://www.cs.ucf.edu/courses/cap6412/2003/Lecture-1.pdf

Computer Vision

- Computer Vision deals with recovery and use of information about objects present in a scene from images of the scene.

Computer Vision

- Computer Vision emerged from:
  - Image Processing
  - Pattern Recognition

Computer Vision

- Computer Vision started as an AI problem.

AI

- Artificial Intelligence is the study of mental faculties through the use of computational models.
  - Search
  - NLU
  - Speech Recognition
  - Games
  - Computer Vision
  - Expert Systems

Image Understanding

- To understand a single image of a scene, locate and identify objects, their structure, and spatial arrangements, and relationships with other objects.

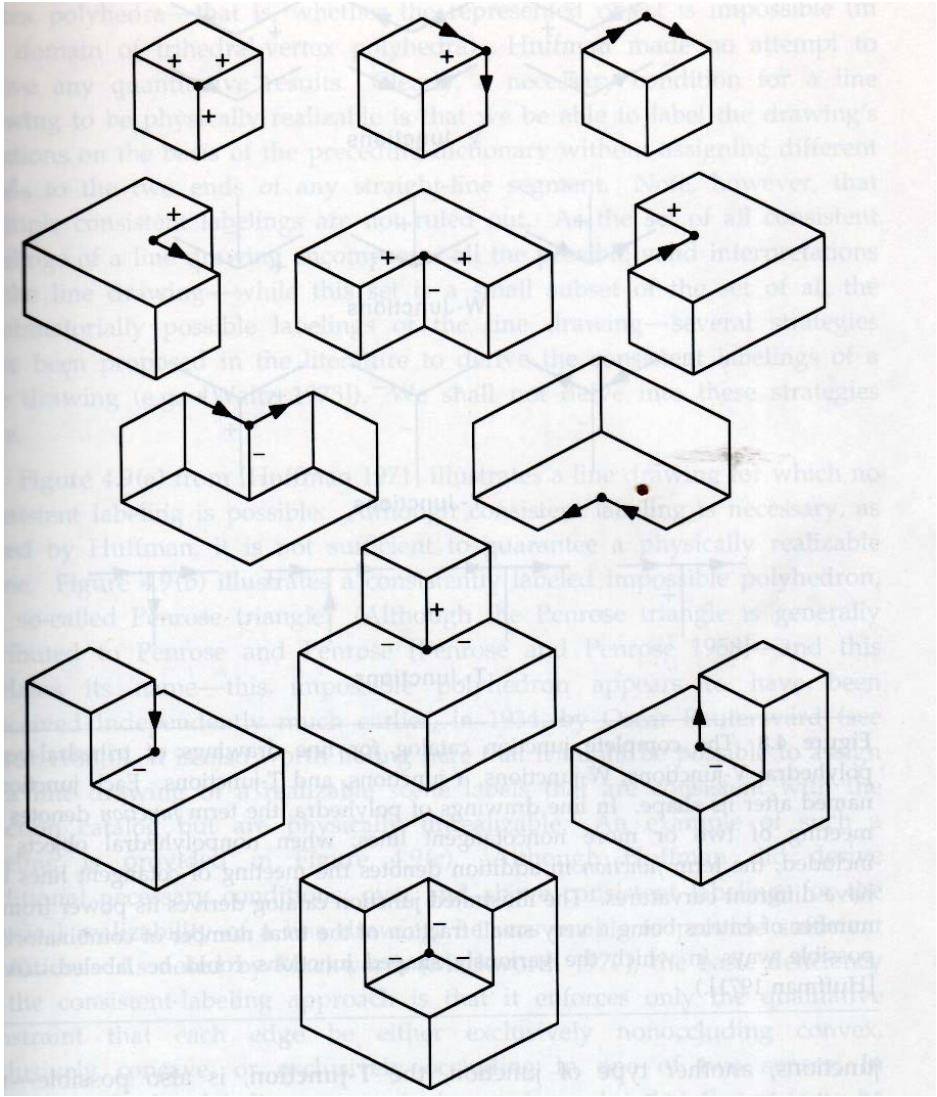
Different Levels

- Low Level: Extraction of symbolic information
- Intermediate Level
- High Level: Interpretation

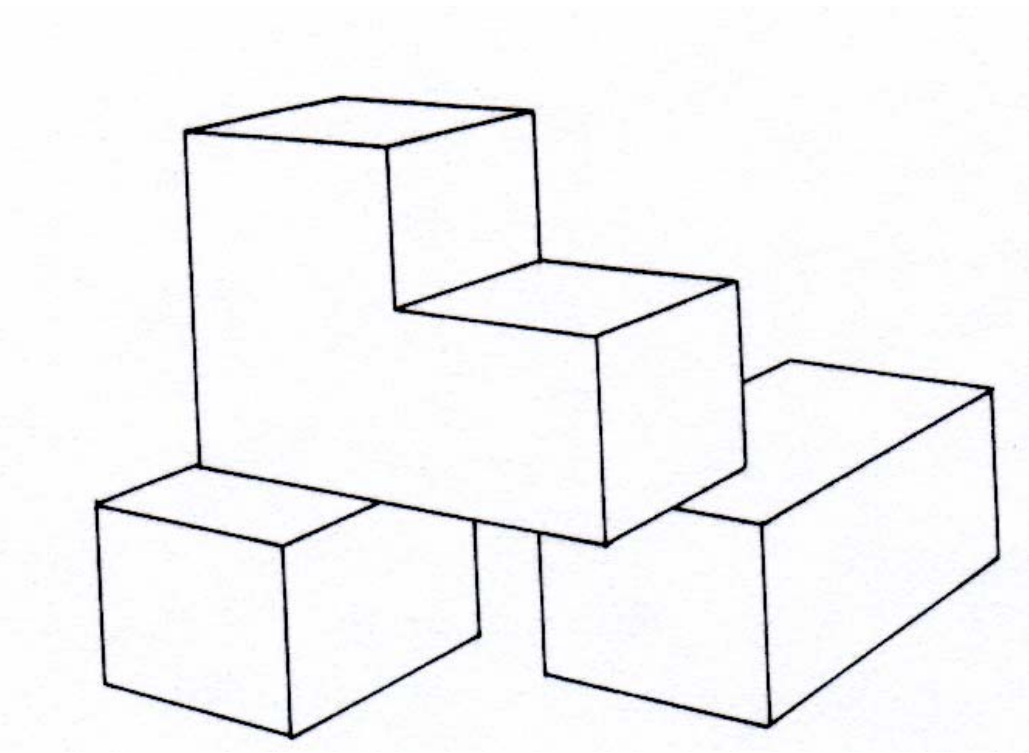
High Level Vision

- Image Understanding
- Scene Interpretation
- Line Drawings

Interpretation of Line Drawing



MIT Copy Demo

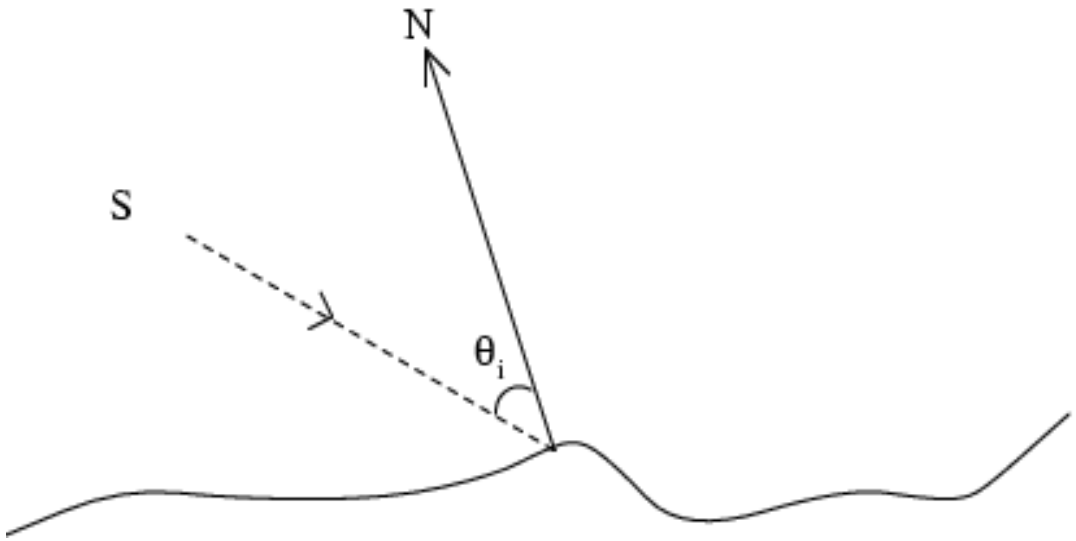


What happened?

- In order to do line interpretation, need to extract lines from images
  - Horn-Binford line finder
  - Solve low level problems before high level problems can be solved.

Horn: Physics Based Vision

- Optics
- Reflectance
- Illumination



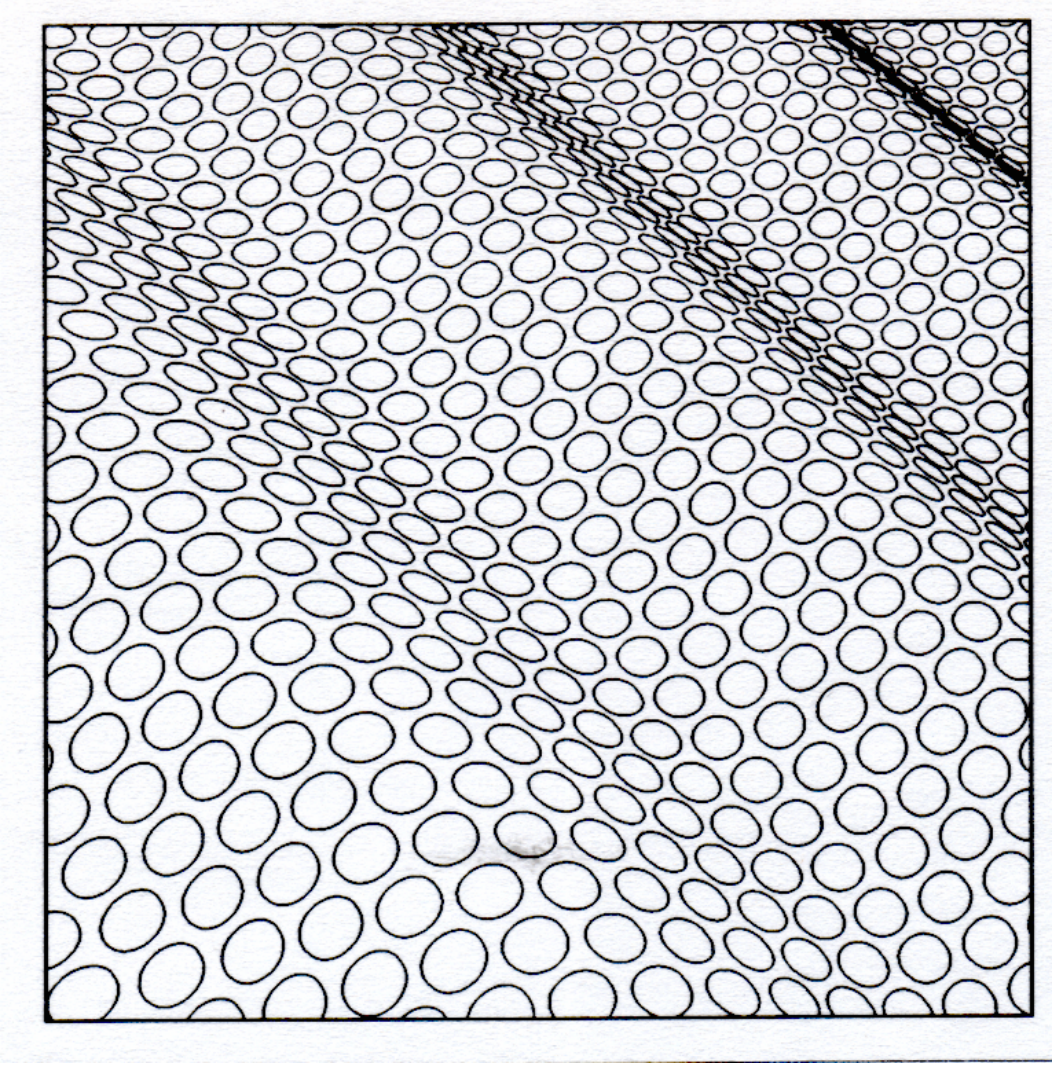
Marr Approach

- **Human vision system**
- **Shape from X: Recover 3-D from 2-D**
- **Quantitative vs Qualitative**

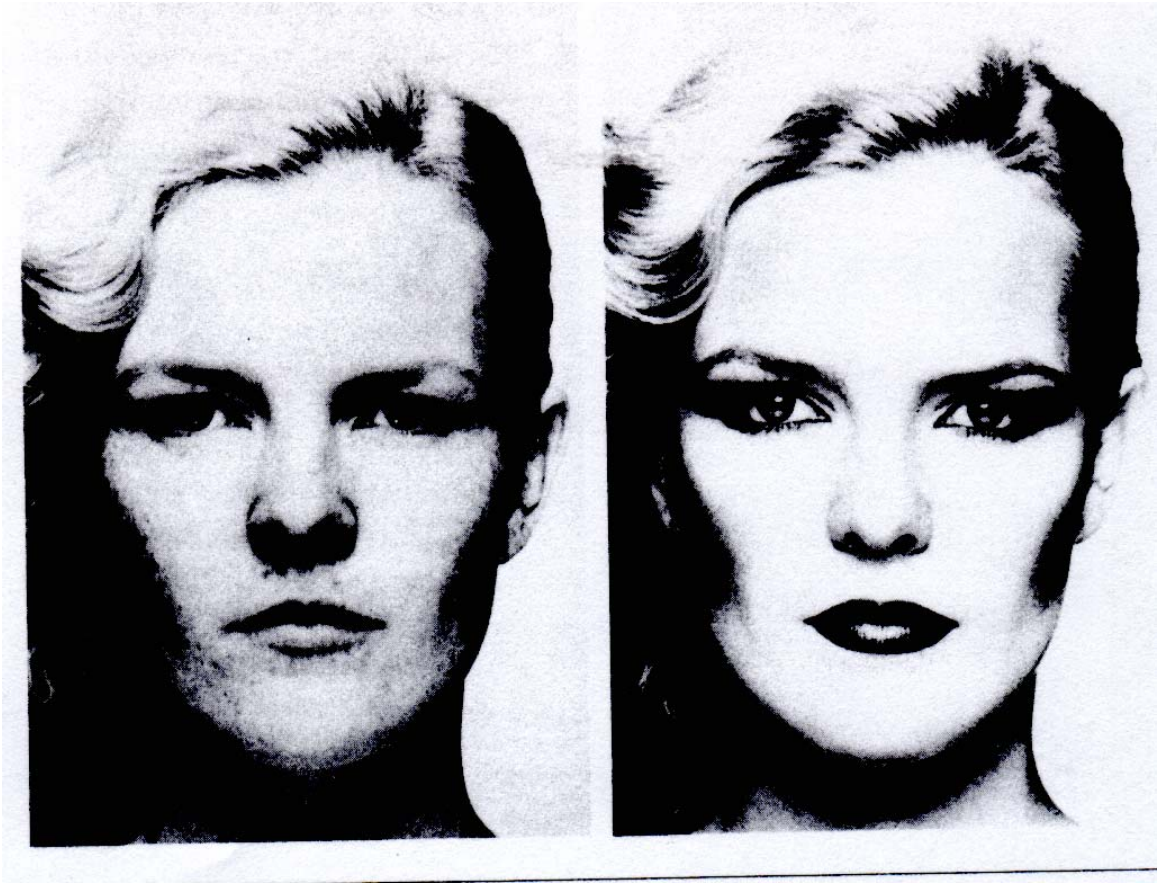
Shape from X

- Shading
- Stereo
- Texture
- Motion
- Contours

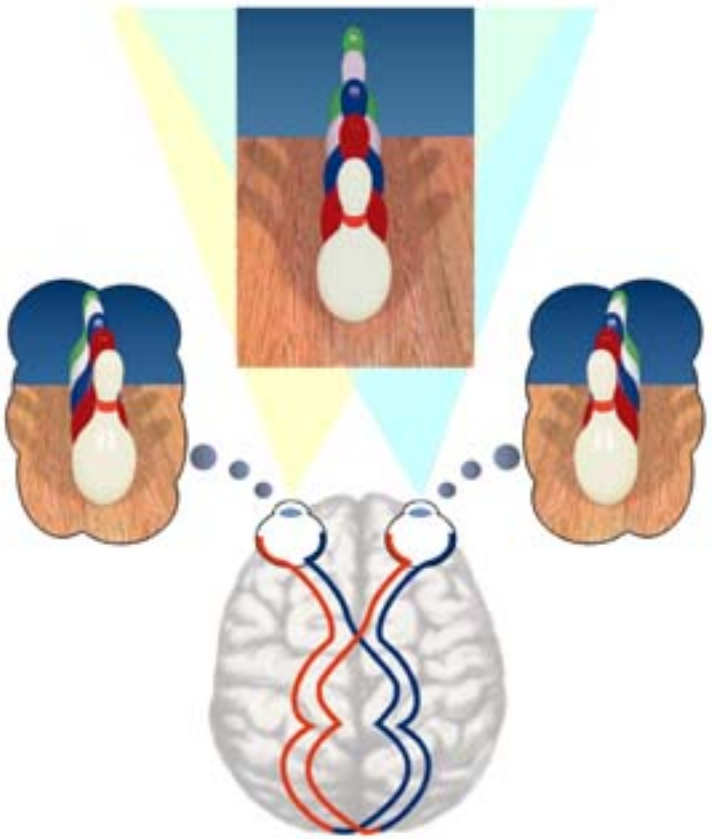
Shape from Texture



Shape from Shading



Shape from Stereo



Marr's Three Levels

- Primal Sketch
  - Marr-Hildreth edge detector
- 2.5 Sketch
  - Marr-Poggio stereo algorithm
  - Grimson's stereo algorithm
  - Ullman's structure from motion
  - Pentland, Witkin, Kass,
  - Terzopoulos: surface reconstruction
- 3-D
  - Generalized Cylinders: Nishihara

After 30 Years of Research

- Stereo is almost a solved problem
- Structure from motion is very hard
- Shape from shading is not interesting/applicable
- Range images did not help much
- Not much progress in understanding/recognition/interpretation

Motion-Based Recognition

- A longer sequence leads to recognition of higher level motions, like walking or running,which consist of a complex and coordinated series of events that cannot be understood by looking at only a few frames.
  - 3-D is not necessary for recognition
  - Use motion directly for recognition vs
    - Recognition followed by reconstruction

Video Understanding

- Gestures
- Activities
- Facial expressions
- Visual Speech

- Applications
  - Video Surveillance and Monitoring
  - Perceptual User Interface
  - Model-based Video Compression
  - Augmented Reality and Video Games
  - Synthesis of Video Sequences

Copy Demo Using A Video Sequence:

Making a Sandwich

[bread, lettuce, ham, bread]

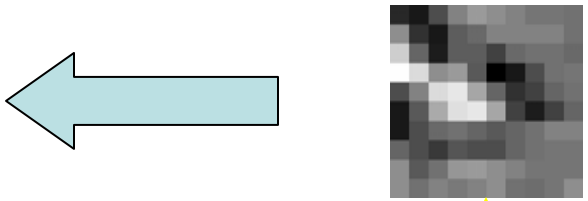
A picture is worth a thousand words.

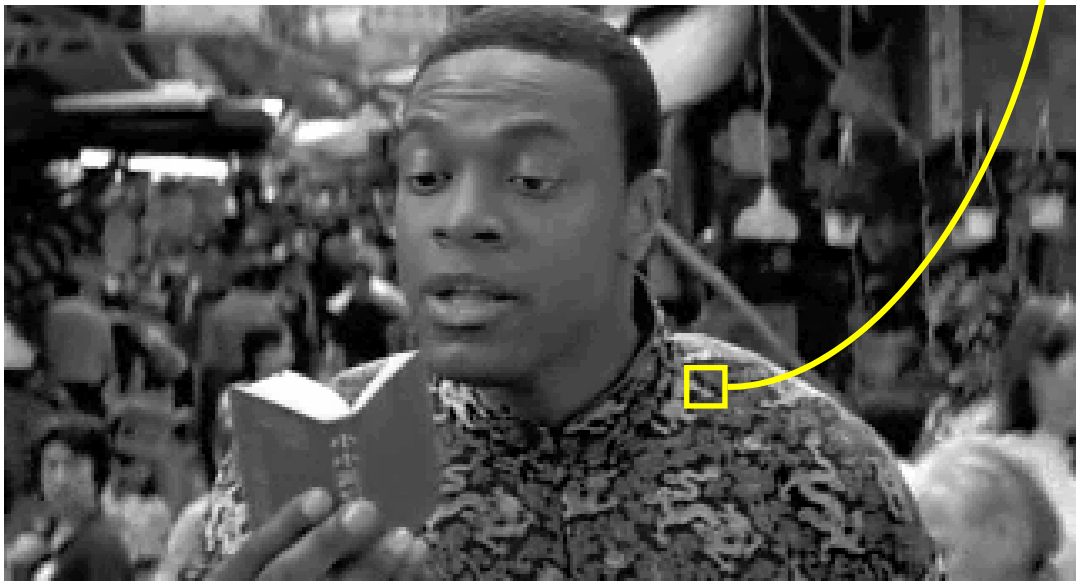
A word is worth a thousand pictures.

A H UNT

What is an Image?

34	23	58	89	106	97	89	83	83	81
97	39	23	67	75	89	89	89	89	81
139	73	26	67	67	50	75	81	81	75
171	147	97	106	64	7	23	58	81	83
56	89	147	155	114	73	40	50	73	81
23	64	115	148	155	114	48	26	48	73
23	56	74	81	73	64	73	81	89	89
73	56	45	62	57	56	73	81	82	82
97	64	81	103	106	97	89	82	82	82
97	81	89	86	89	97	81	78	82	97





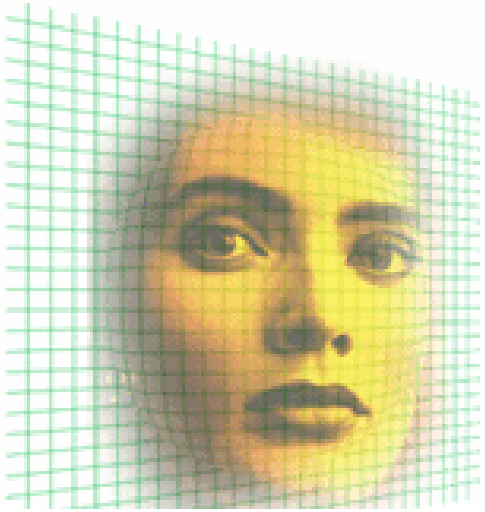
Video Clip

Sequence of Images

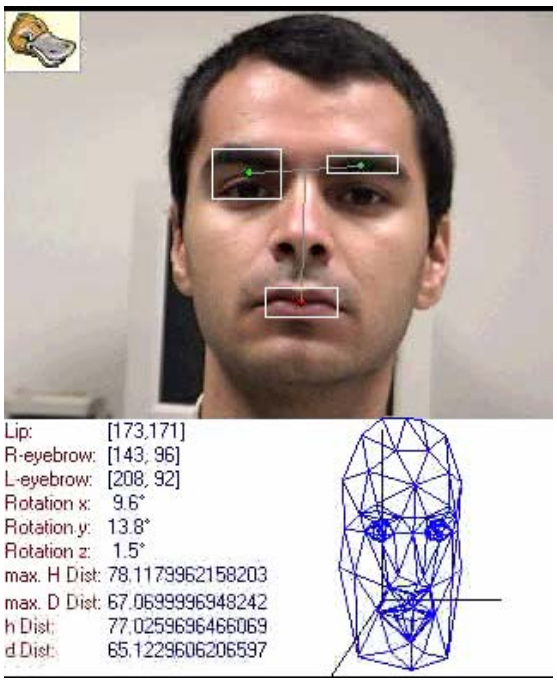
Applications

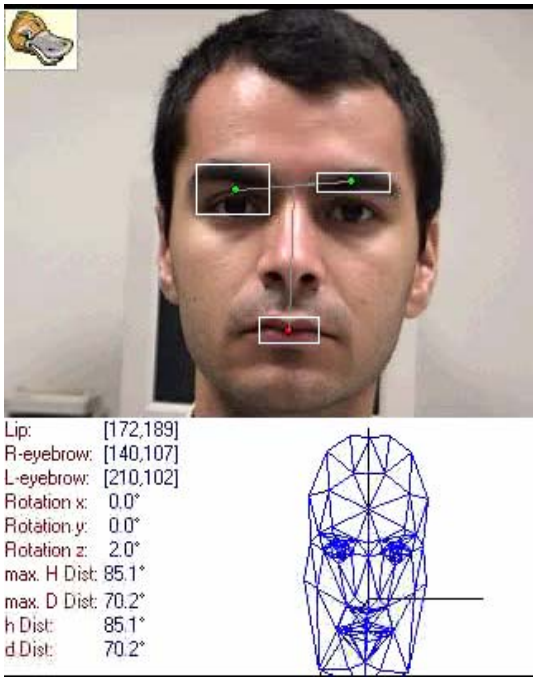
- Face Recognition
- Robotics
- Remote Sensing: UAVs
- Computer Graphics
- Video Surveillance and Monitoring
- Video Data Mining

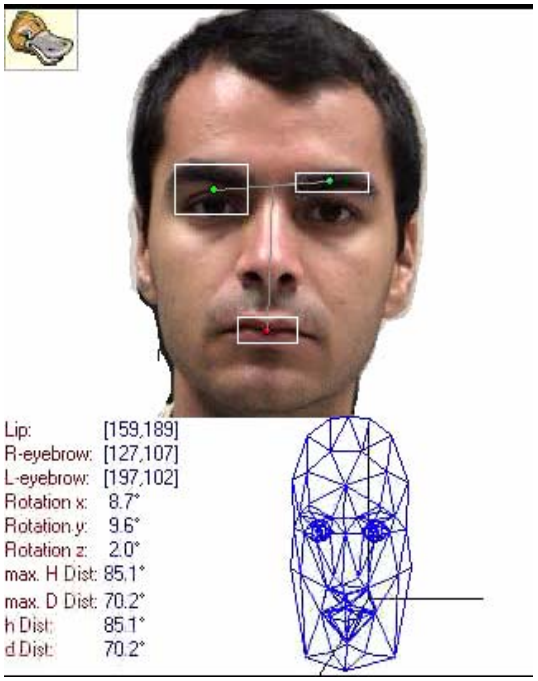
Face Recognition



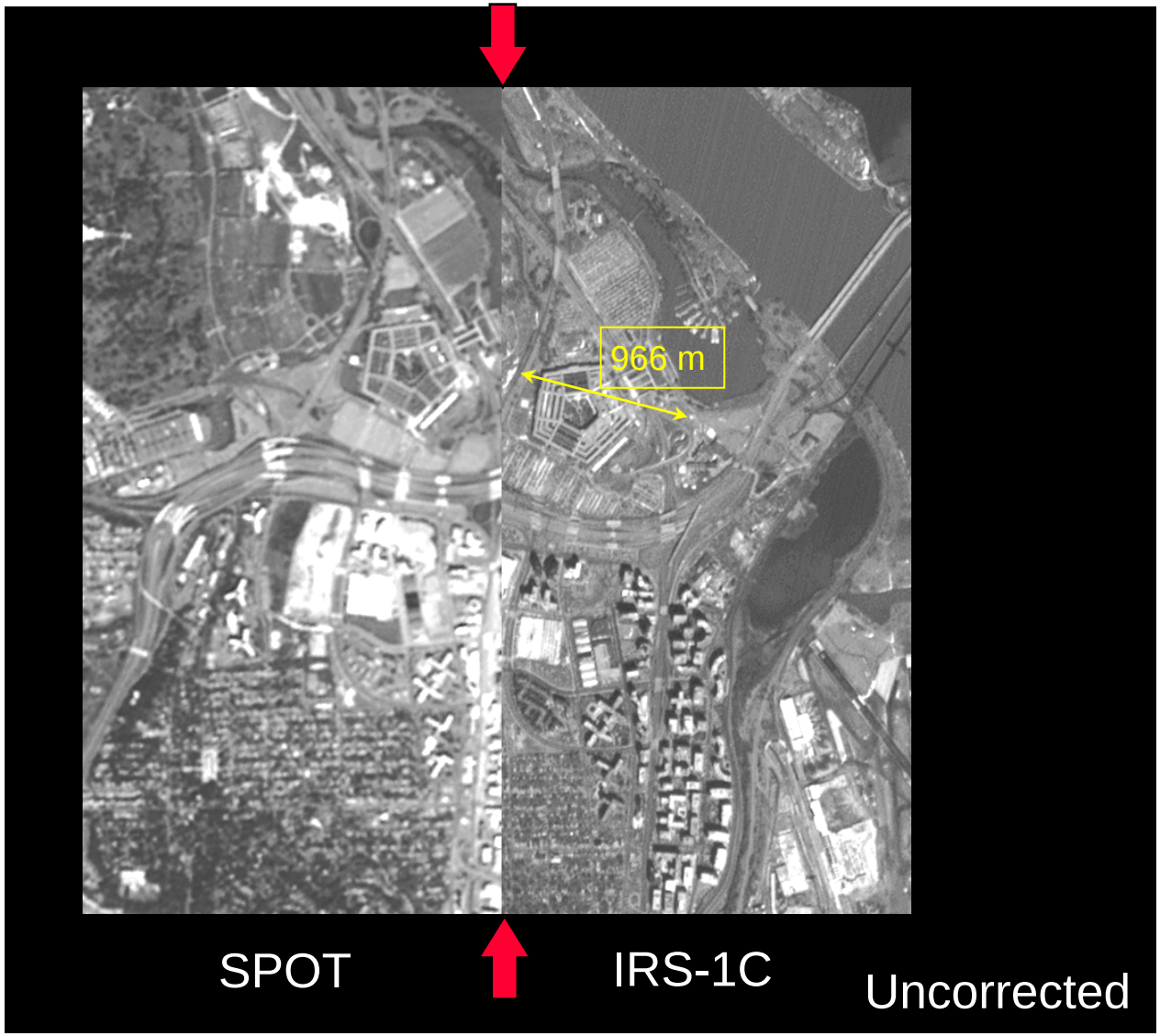
Determining Face Orientation



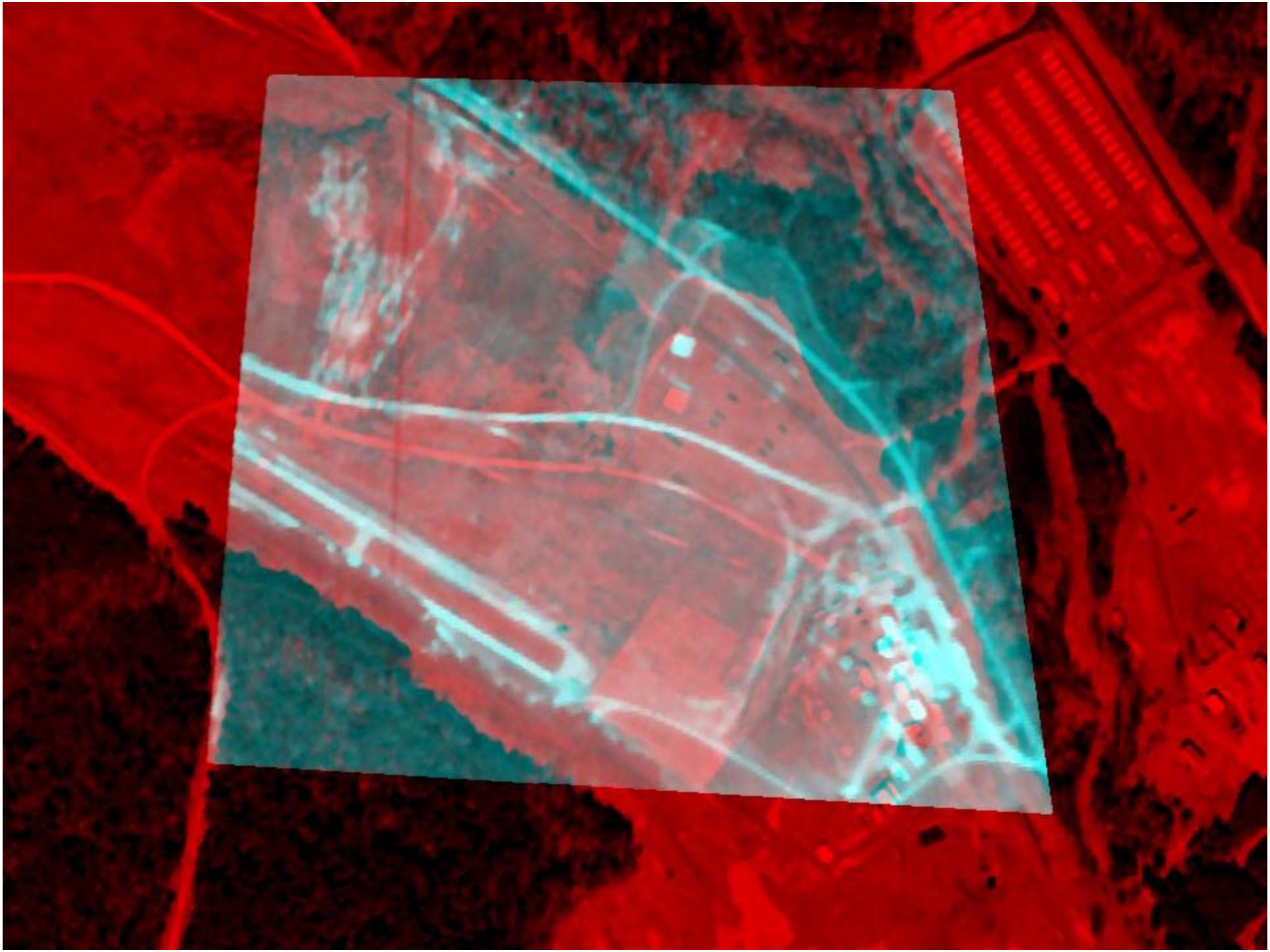




Geo-registration



Registered IRS-1C to SPOT



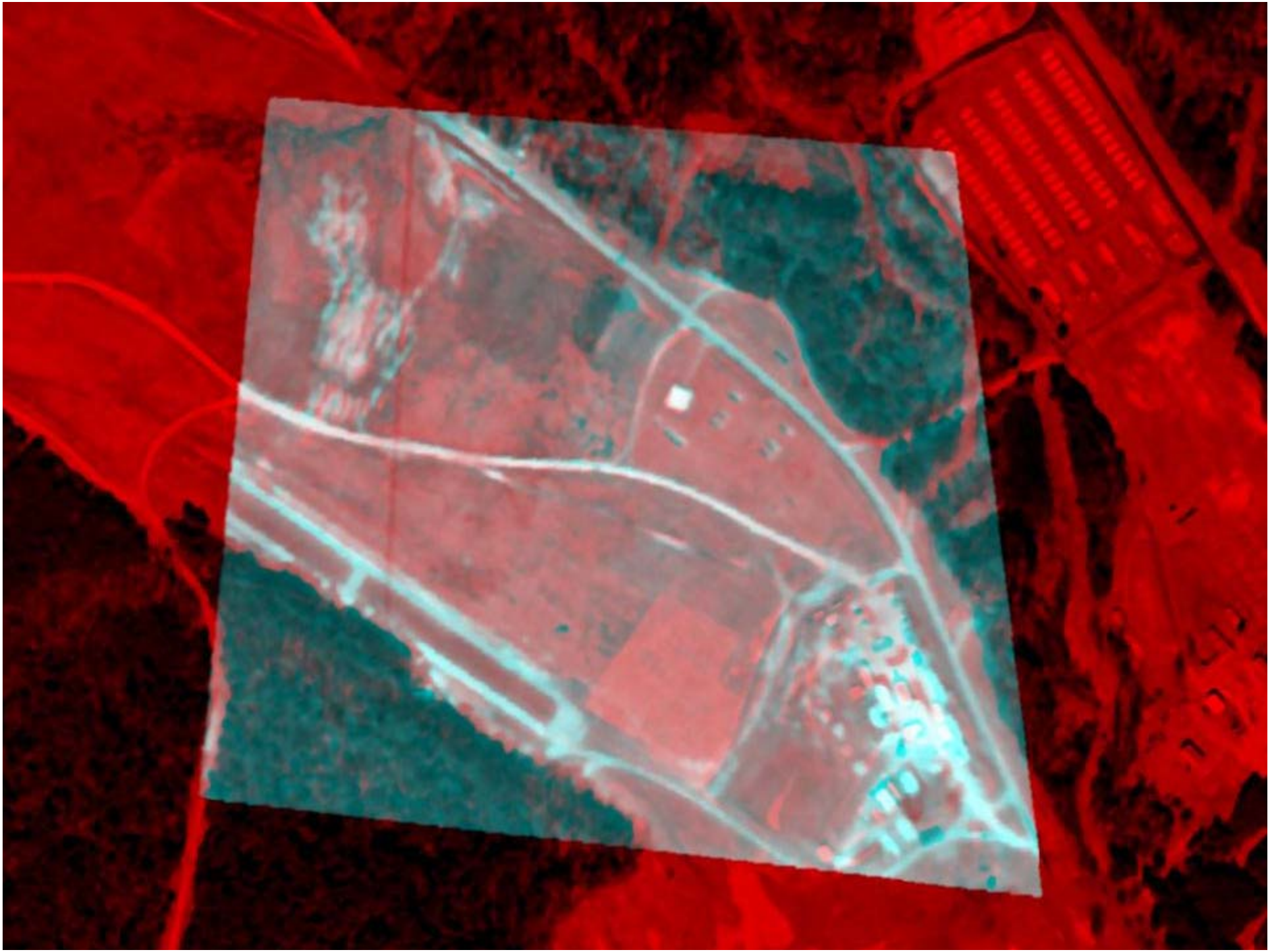


Image-Based Rendering



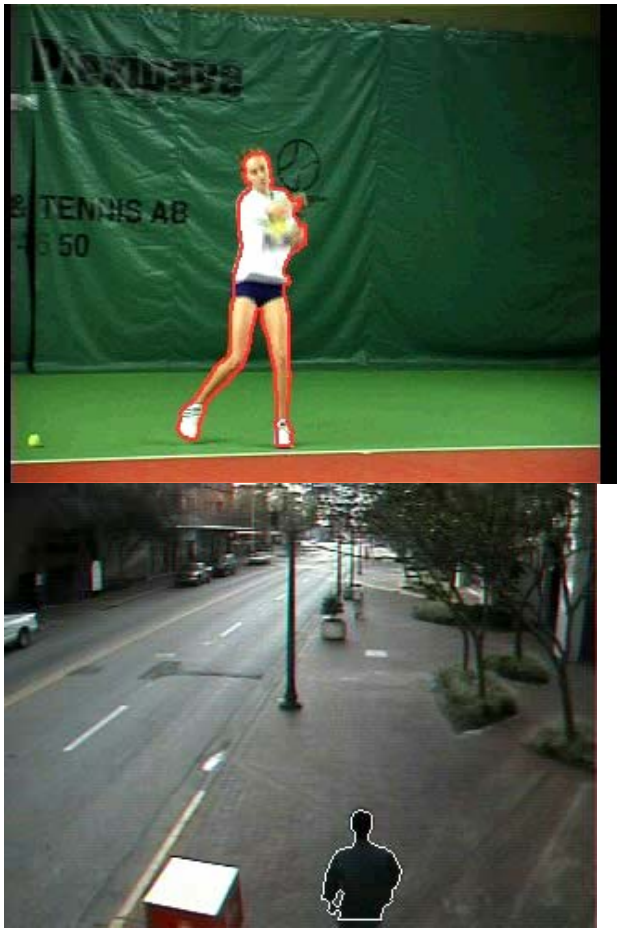



KNIGHT
Crime Scene Detection System
for The Orlando Police Department

Cameras

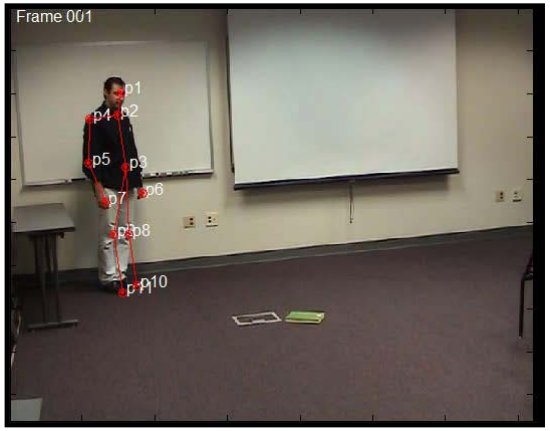
Tracking

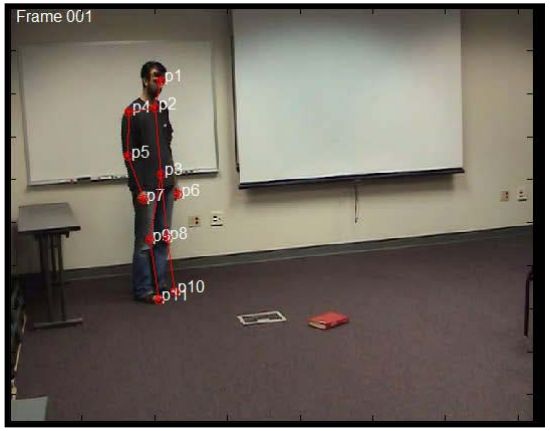
Contour-based Object Tracking Using Level Sets

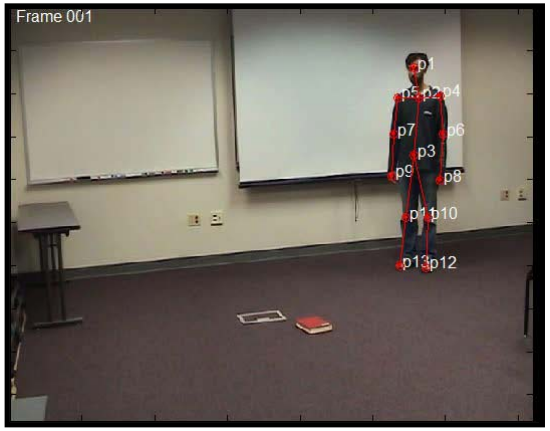


Action Detection:
Different approaches, different people, the same action

ReferencePattern

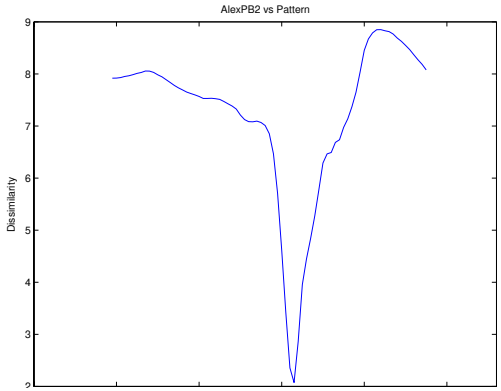


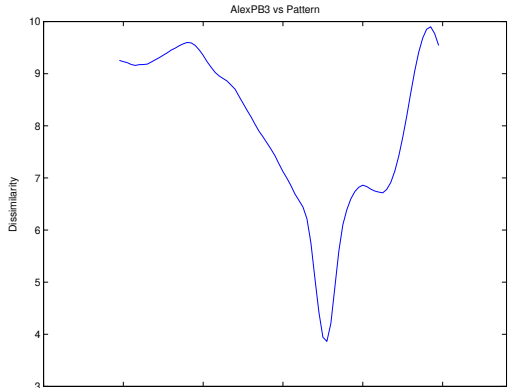


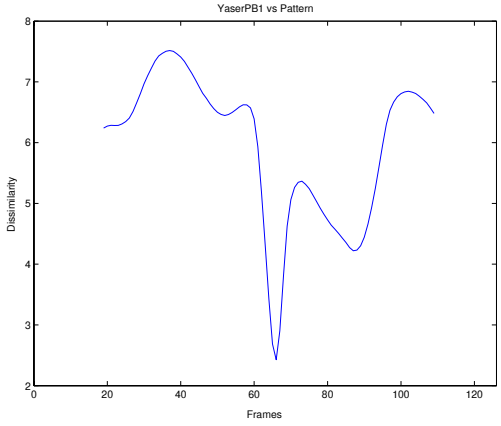


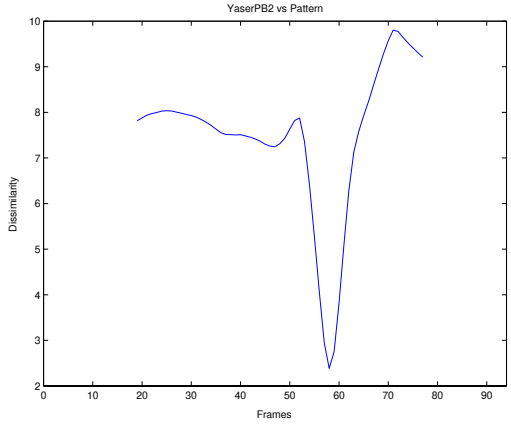
Test Sequences

Action Detection:
Different approaches, different people, the same action



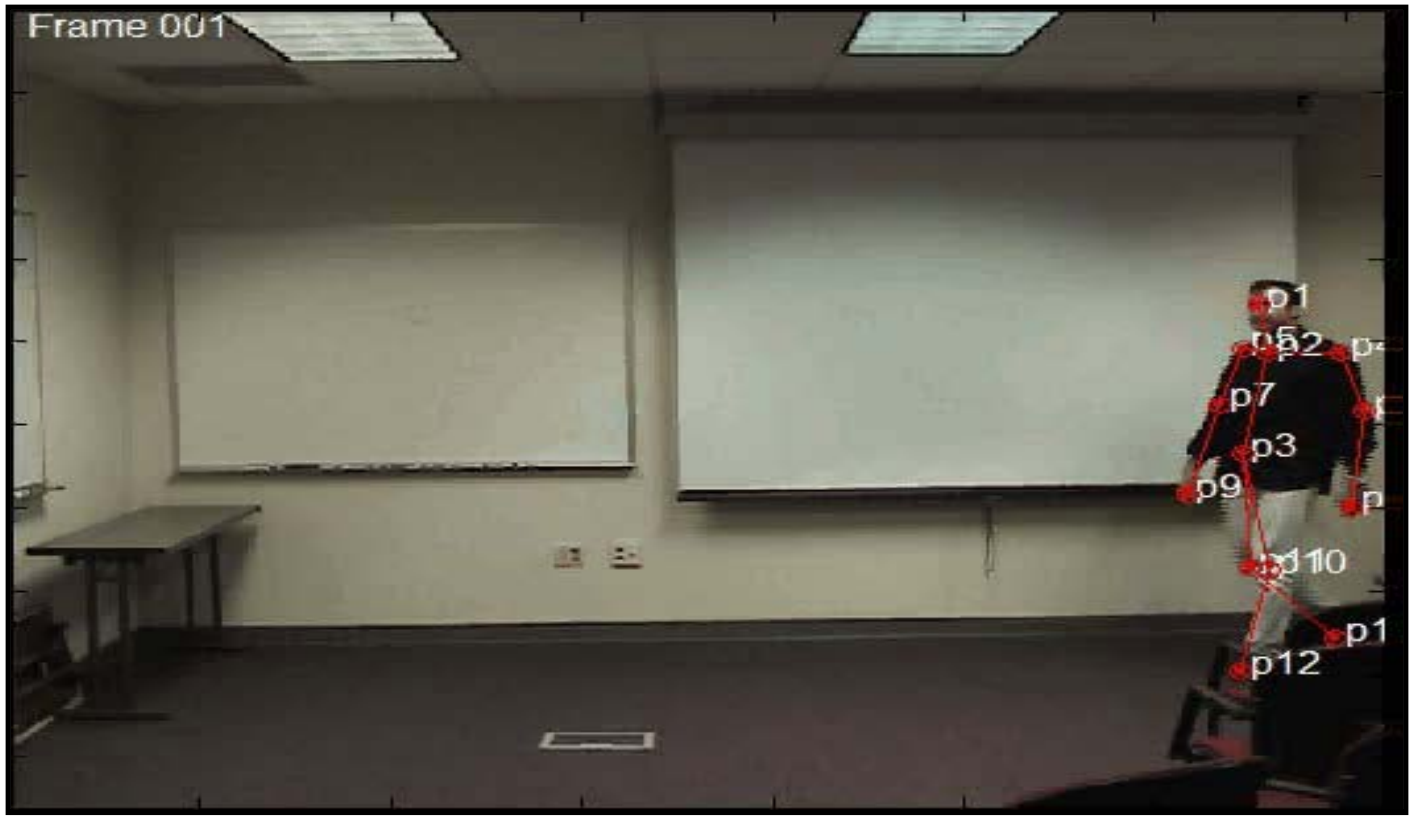




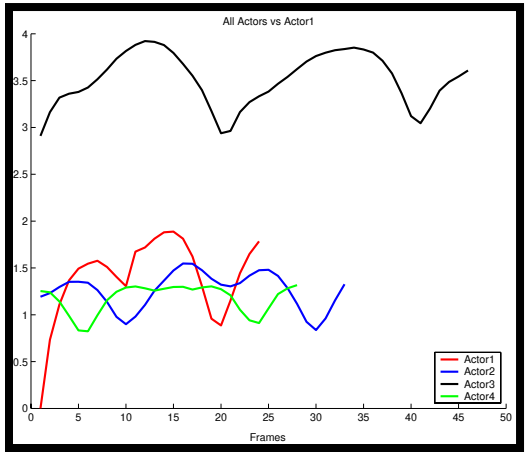


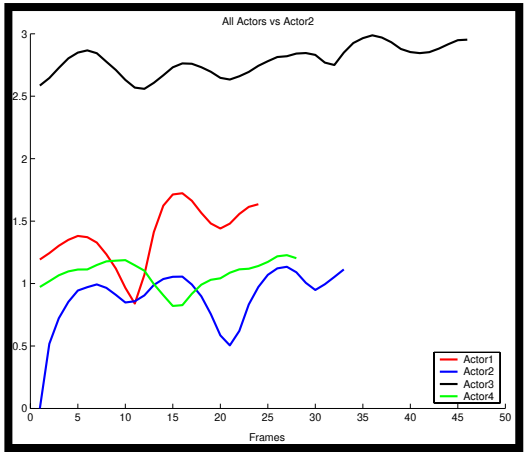
Analyzing Actions

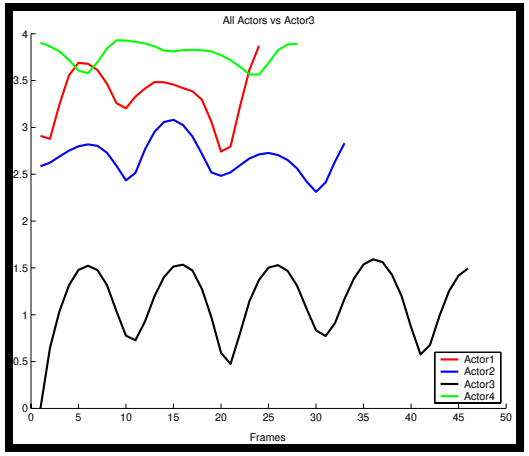
Odd One Out

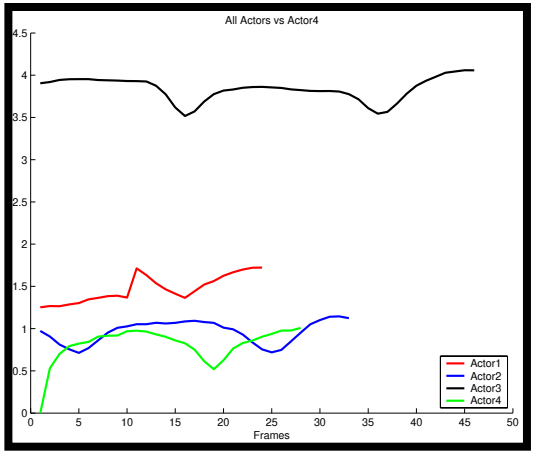


‘Odd One Out’


All Actors vs Actor1
Frames
Actor1
Actor2
Actor3
Actor4


All Actors vs Actor2
Frames
Actor1
Actor2
Actor3
Actor4


All Actors vs Actor3
Frames
Actor1
Actor2
Actor3
Actor4

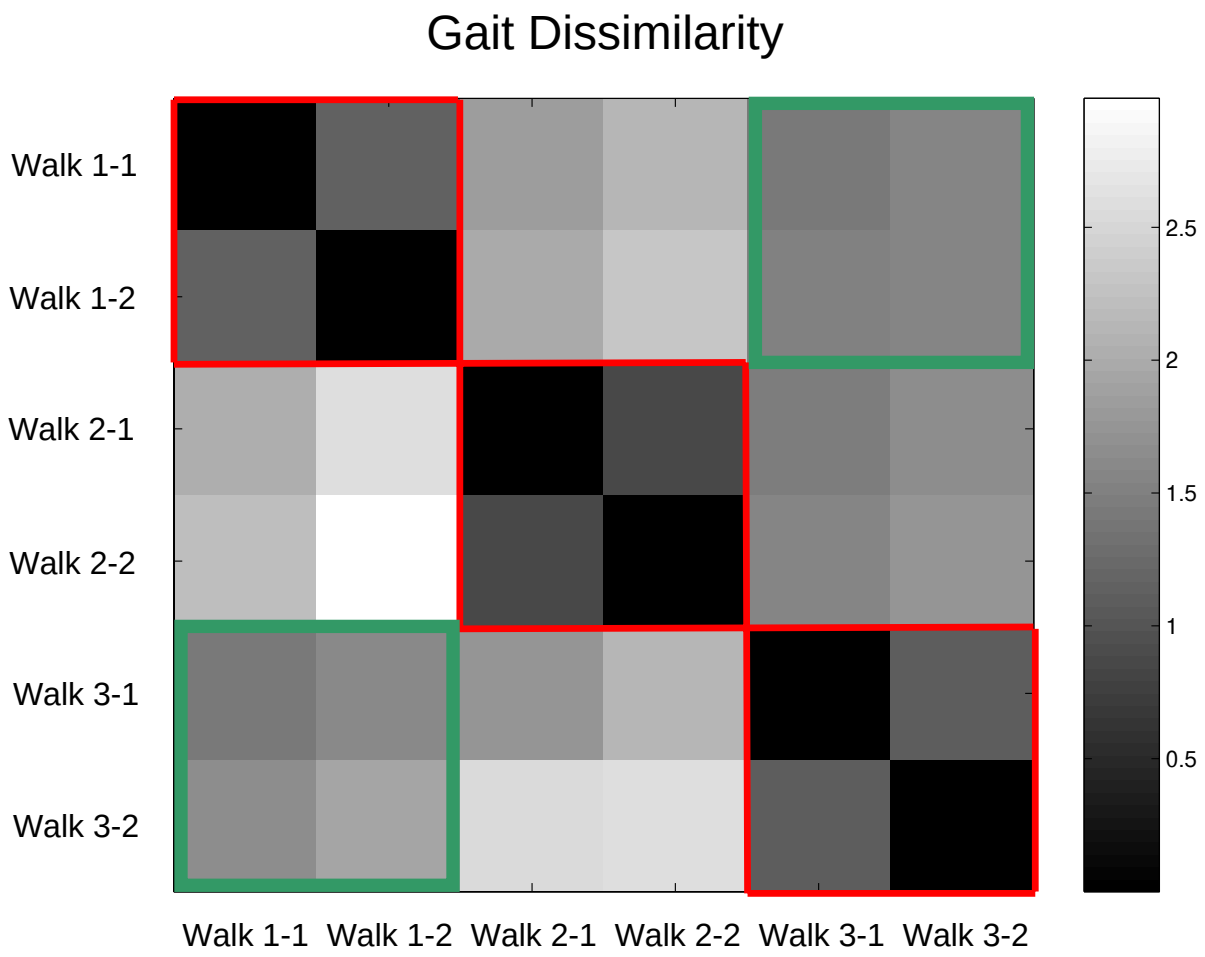

All Actors vs Actor4
Frames
Actor1
Actor2
Actor3
Actor4

Gait Analysis

- Three Actors viewed from two views each

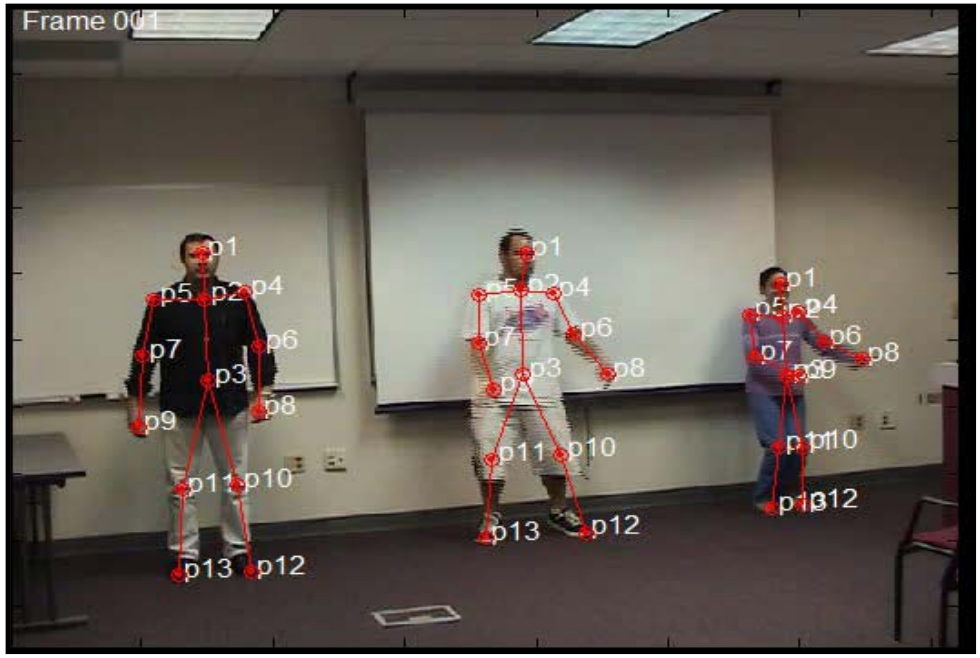
Gait Analysis: Human ID Dataset

Gait Analysis



Action Synchronization

Following the Leader



Action Synchronization

Following the Leader

Outdoor Activities

University of Central Florida VISION

Scene Representation (Terminator II)

Obtained from the DVD

Chapter 21: Syringe Point

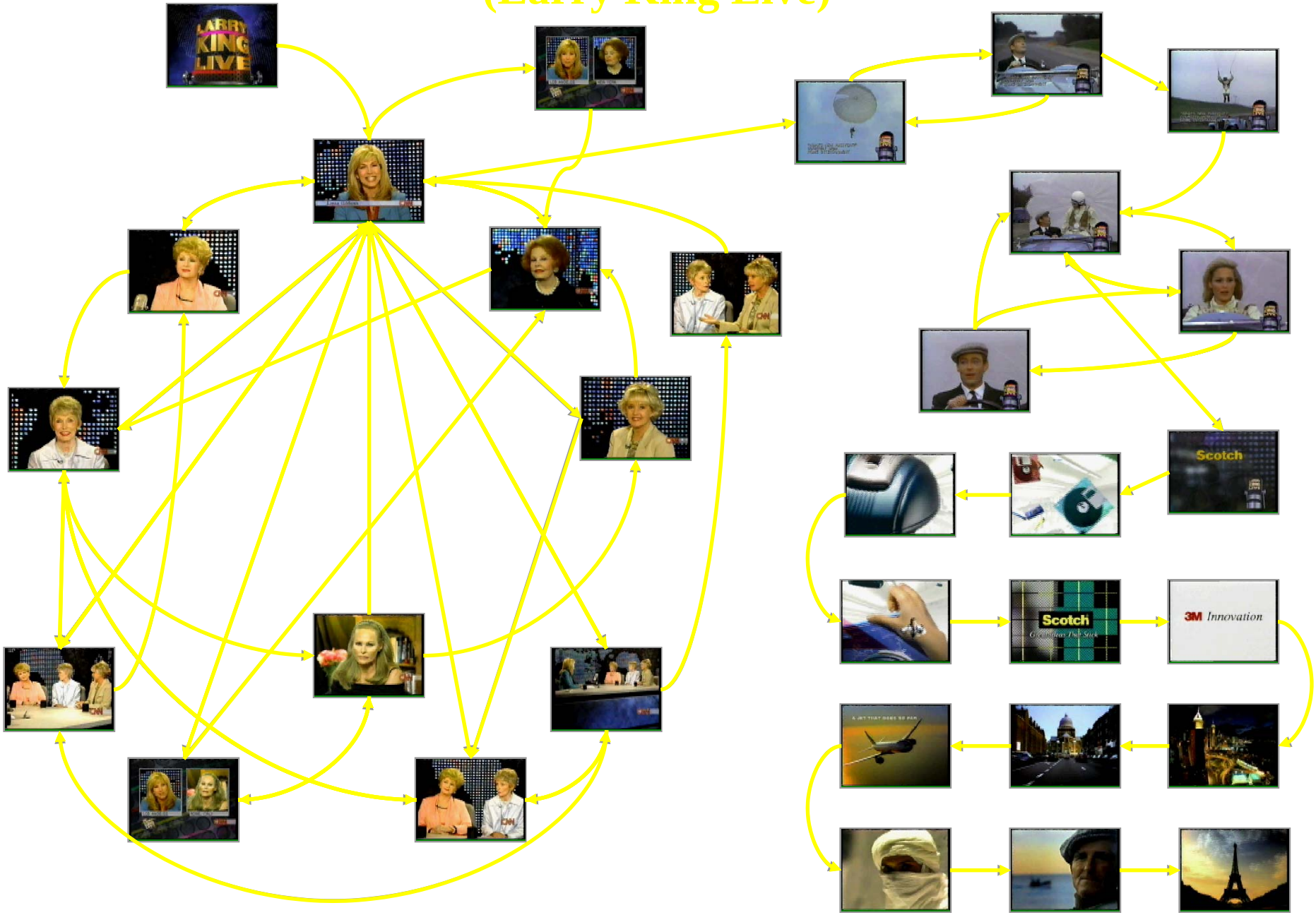

Detected by our algorithm

Chapter 26: Night Repairs


Chapter 29: Detailed Files


Chapter 30: Scalcedas Camp


A Shot Connectivity Graph
(Larry King Live)



Video Google

NIST TRECVID Competition

Face Detection

Training Clinton Detector

Some statistics

- Given:
  - Totally 128 videos (30 minutes each)
  - Totally 48,893 key frames (9,918 faces)

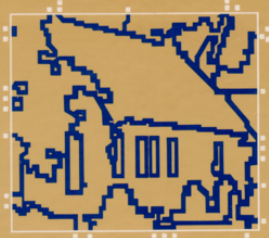
The difficult case: Albright

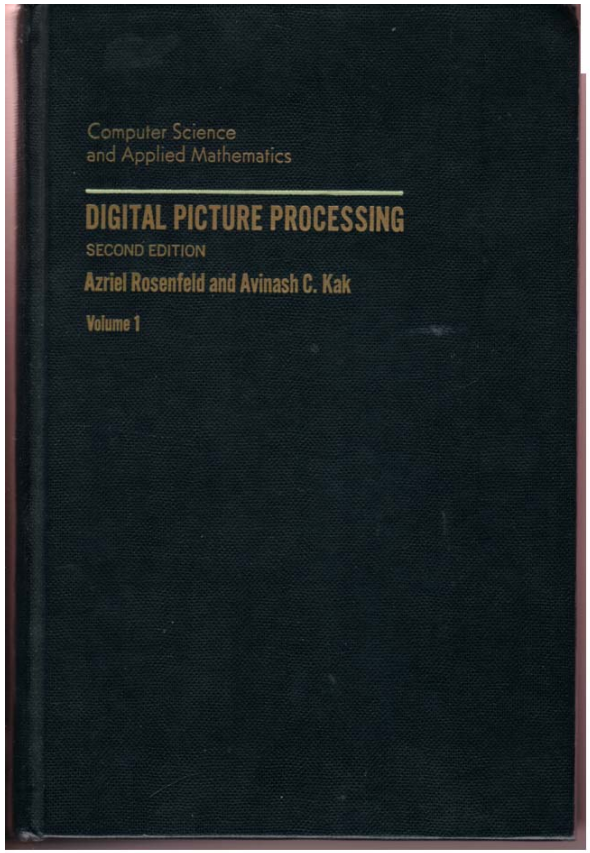
Computer Vision Text Books

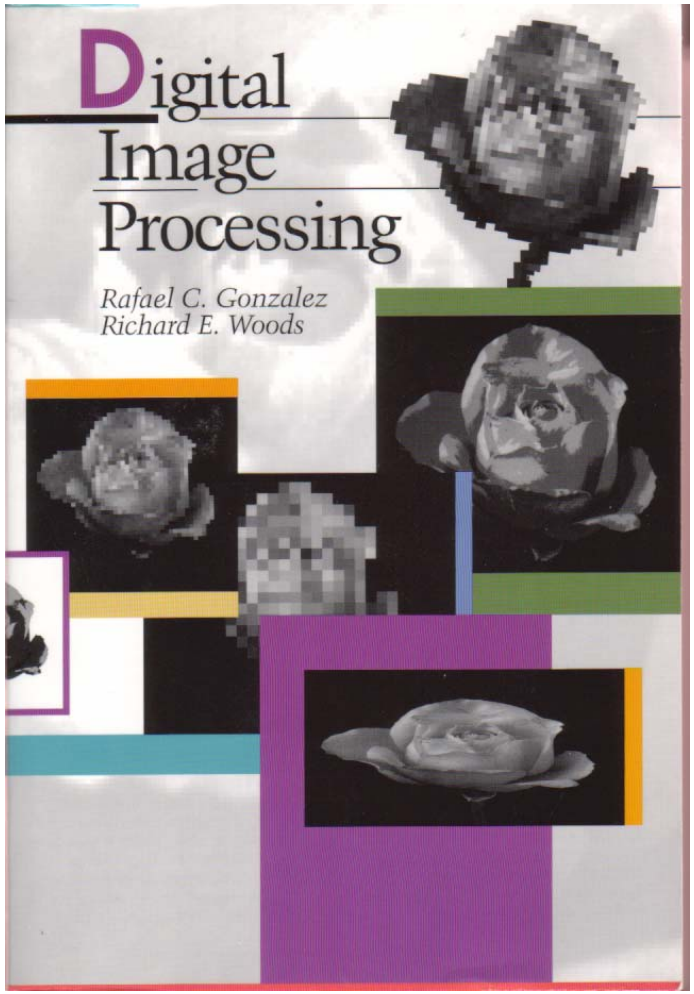
History

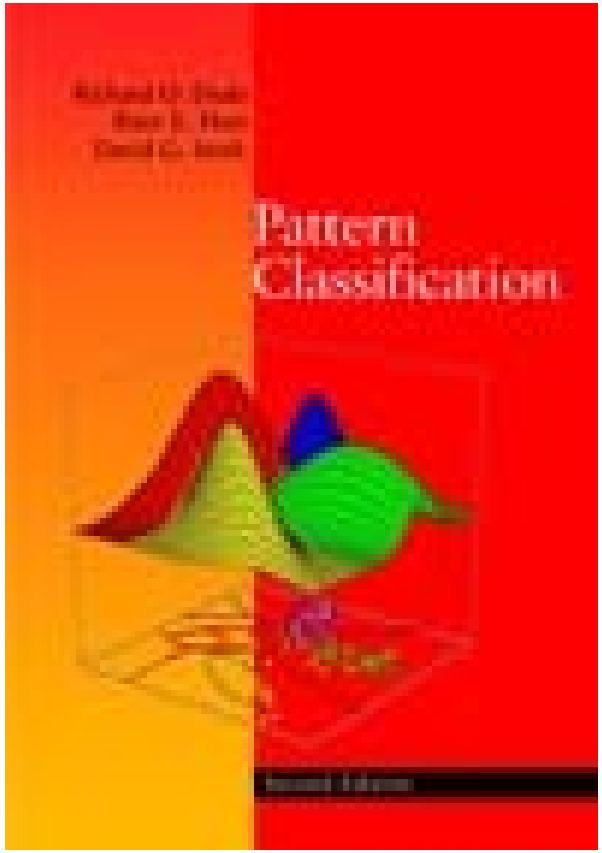
COMPUTER VISION

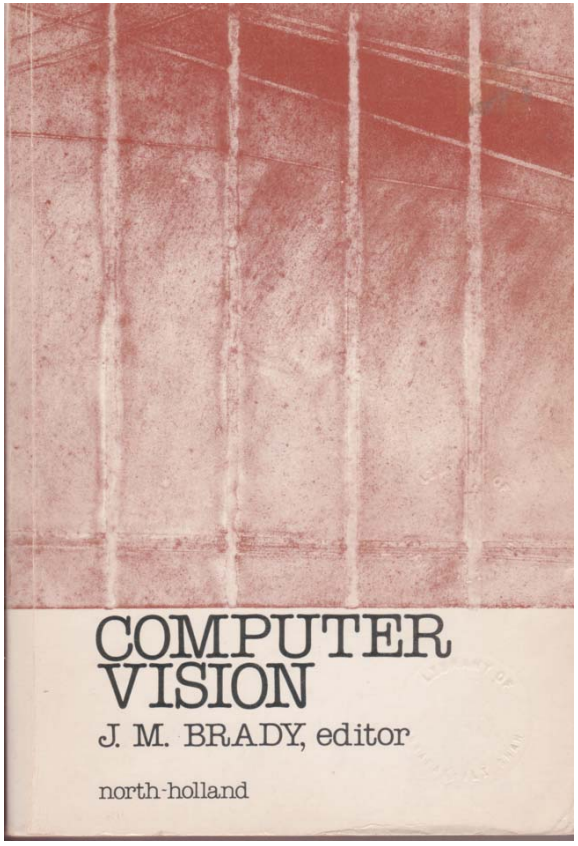
DANA H. BALLARD • CHRISTOPHER M. BROWN

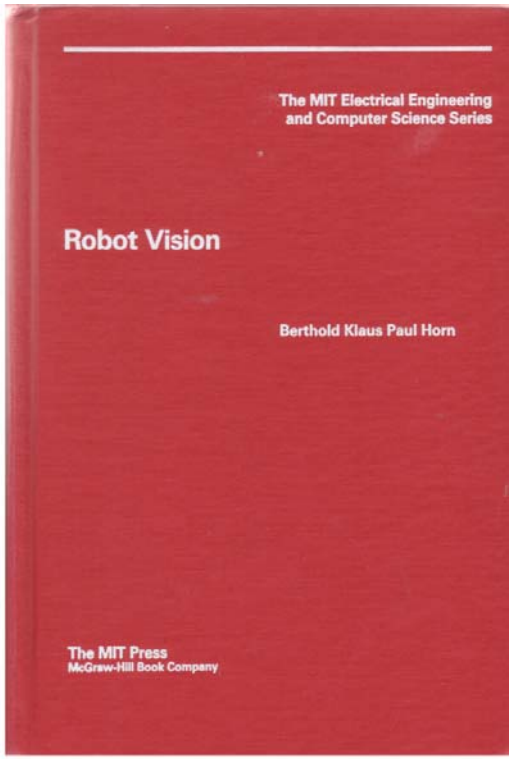



Computer Science
and Applied Mathematics
DIGITAL PICTURE PROCESSING
SECOND EDITION
Azriel Rosenfeld and Avinash C. Kak
Volume 1


Digital
Image
Processing
Rafael C. Gonzalez
Richard E. Woods

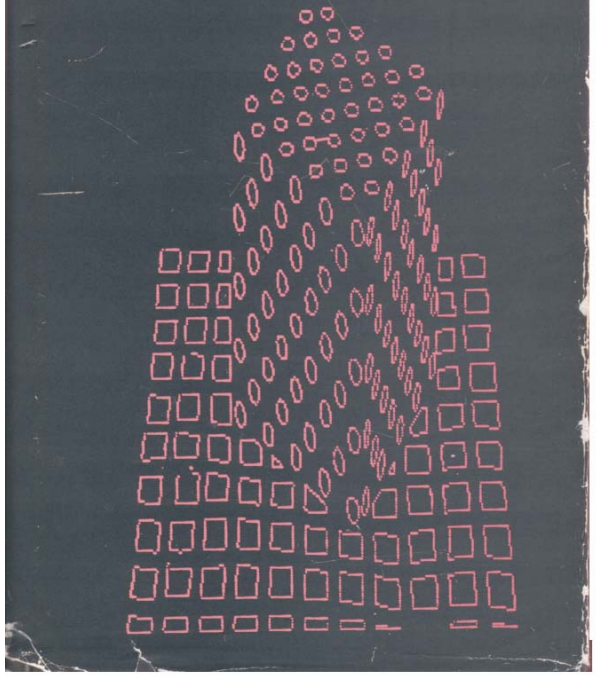

Pattern
Classification

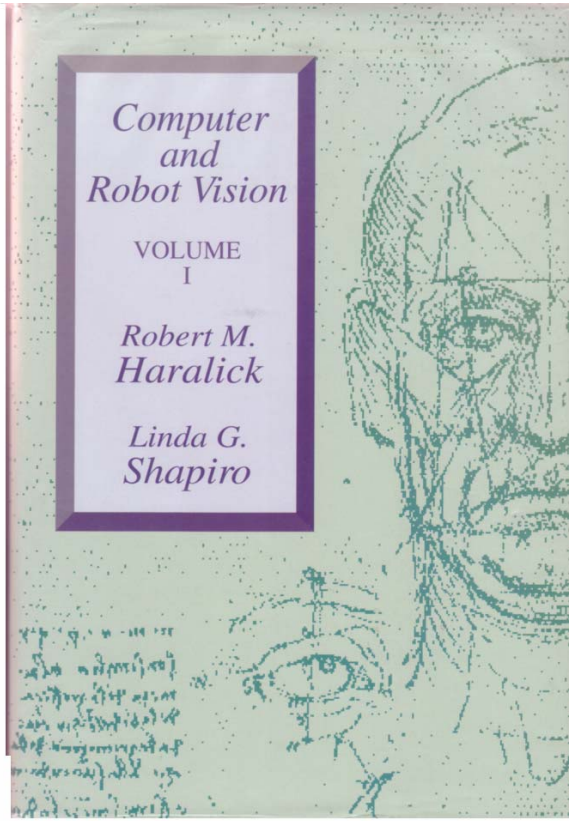

COMPUTER
VISION
J. M. BRADY, editor
north-holland

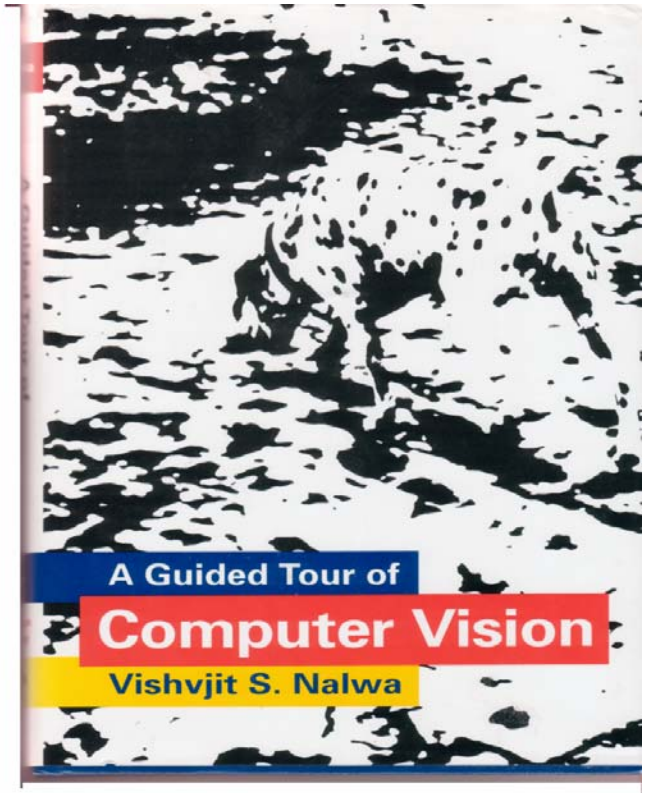

The MIT Electrical Engineering
and Computer Science Series
Robot Vision
Berthold Klaus Paul Horn
The MIT Press
McGraw-Hill Book Company

VISION IN MAN AND MACHINE

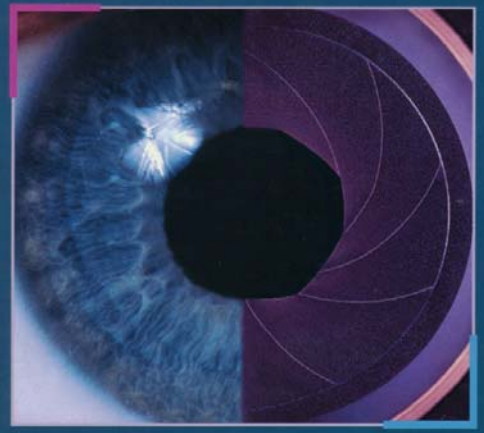
MARTIN D. LEVINE



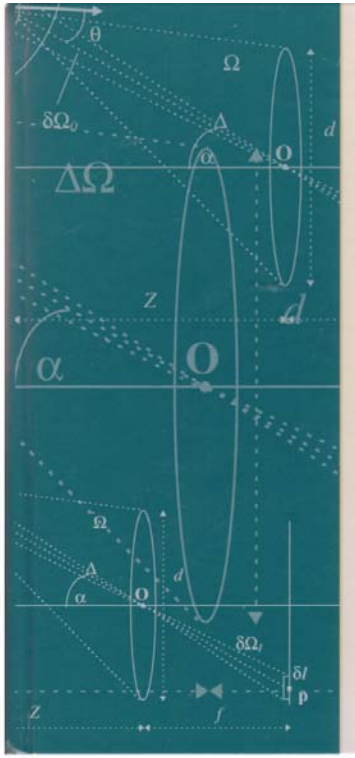

Computer
and
Robot Vision
VOLUME
I
Robert M.
Haralick
Linda G.
Shapiro


A Guided Tour of
Computer Vision
Vishvjit S. Nalwa

MACHINE VISION



Ramesh Jain • Rangachar Kasturi • Brian G. Schunck

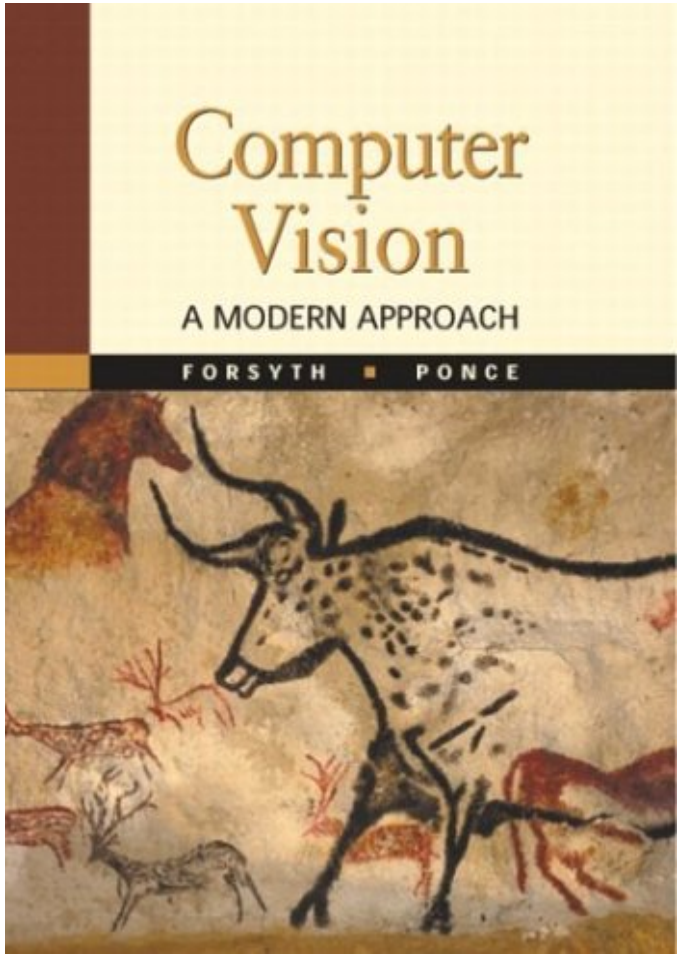


INTRODUCTORY TECHNIQUES *for* 3-D COMPUTER VISION

Emanuele Trucco
Alessandro Verri

COMPUTER VISION

Linda G. Shapiro ▪ George C. Stockman


Computer
Vision
A MODERN APPROACH
FORSYTH ■ PONCE

Computer Vision Researchers

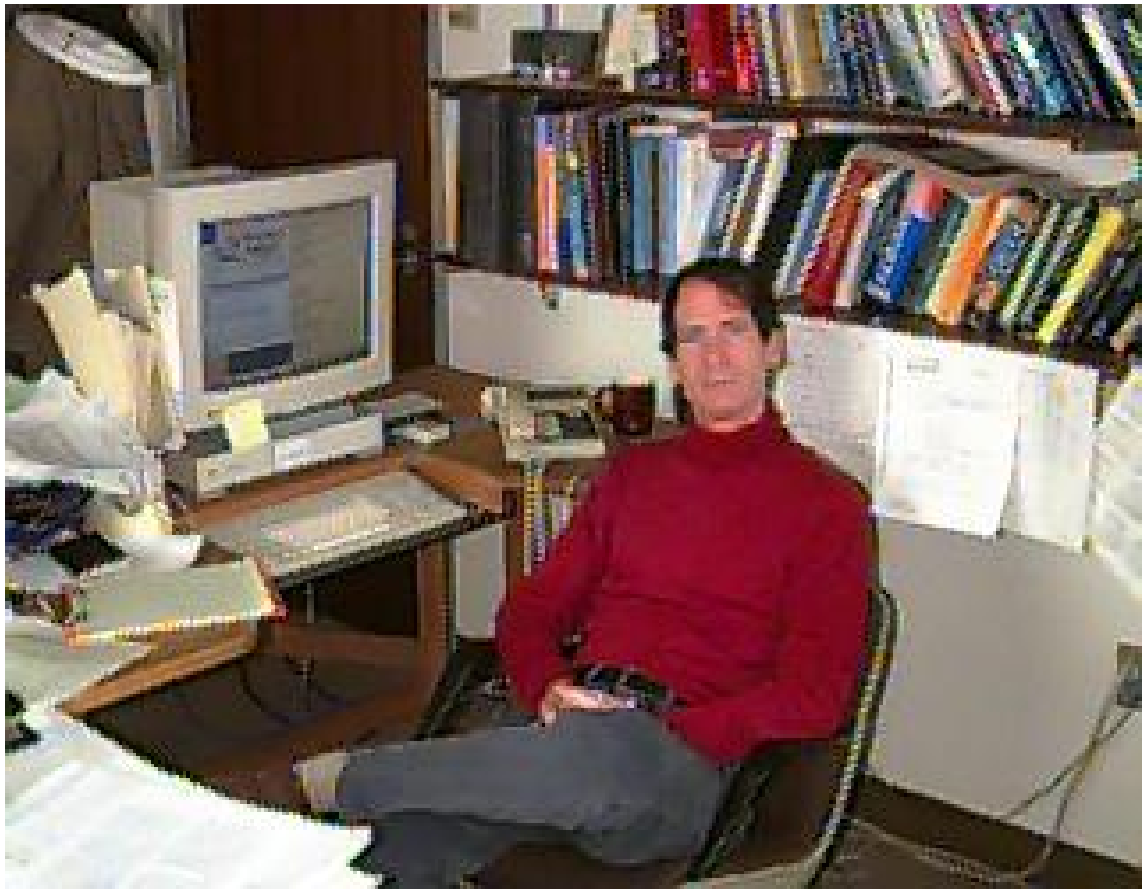
Azriel Rosenfeld

Berthold Horn

Thomas Huang

Jake Aggarwal

Chris Brown



Bob Haralick

Olivier Faugeras

Takeo Kanade

Sandy Pentland

Shree Nayar

John Canny

Demetri Terzopoulos

Ramesh Jain

Computer Vision Journals

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

JUNE 1993 VOLUME 15 NUMBER 6 ITPIDJ (ISSN 0162-8828)

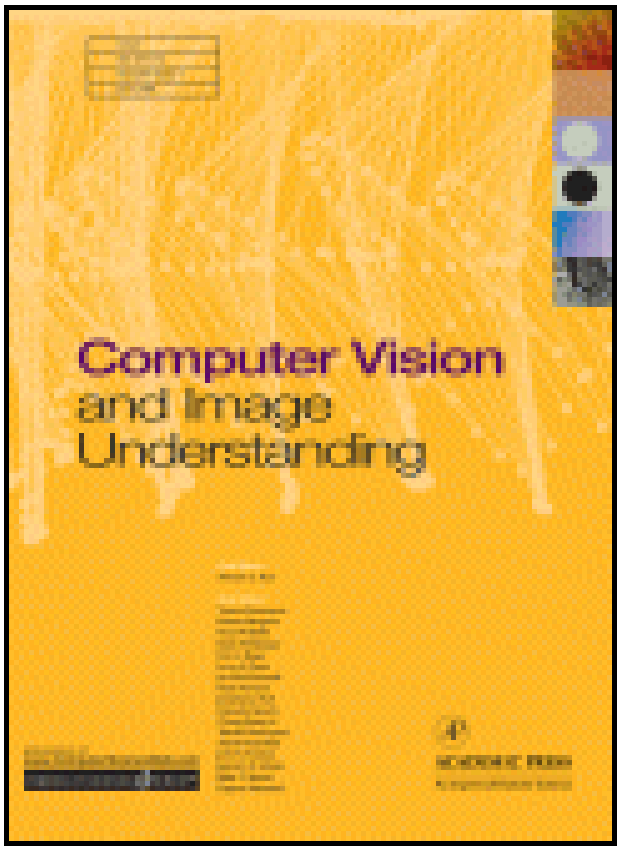
PUBLISHED BY THE IEEE COMPUTER SOCIETY

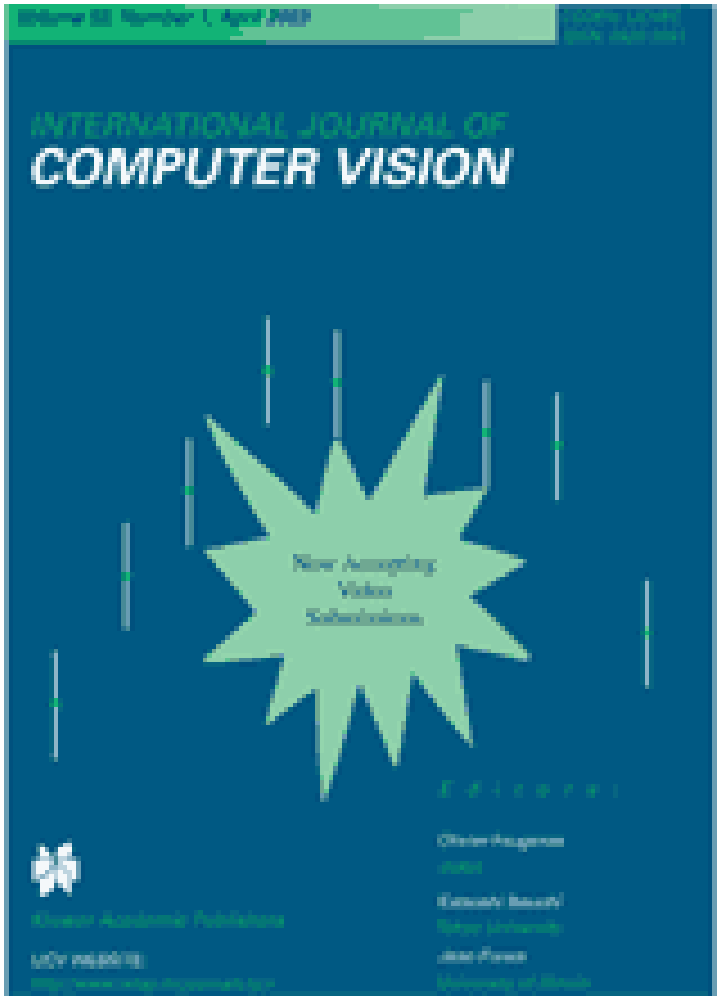
In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Society
Systems, Man, and Cybernetics Society
Ultrasonics, Ferroelectrics, and Frequency Control Society

SPECIAL SECTION ON 3-D MODELING IN IMAGE ANALYSIS AND SYNTHESIS

REGULAR SECTION


Computer Vision
and Image
Understanding


INTERNATIONAL JOURNAL OF
COMPUTER VISION
Now Accepting
Video
Submissions

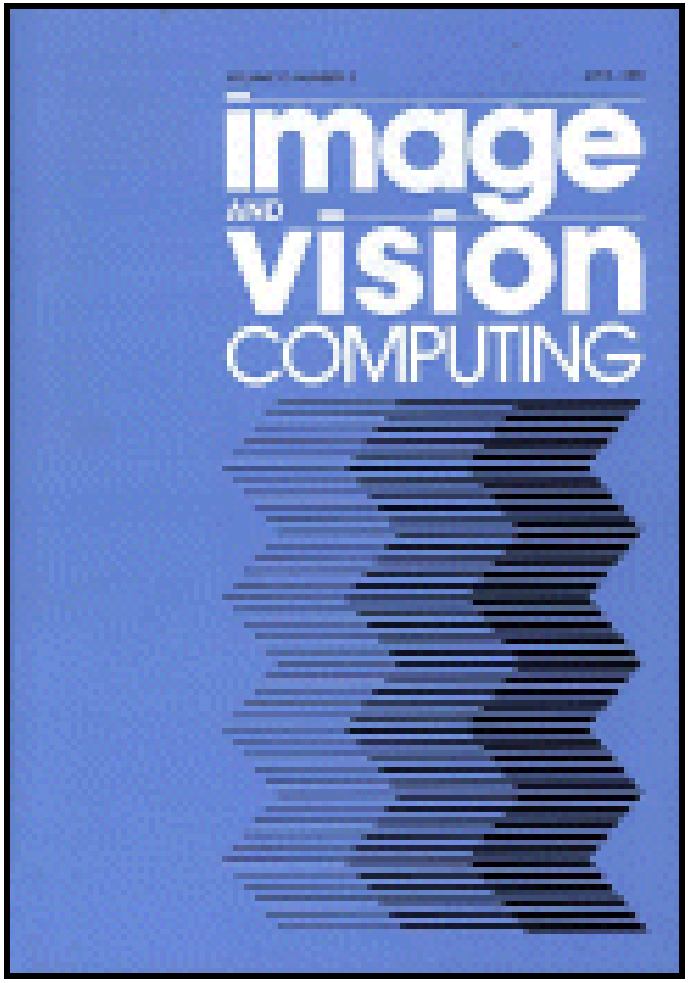


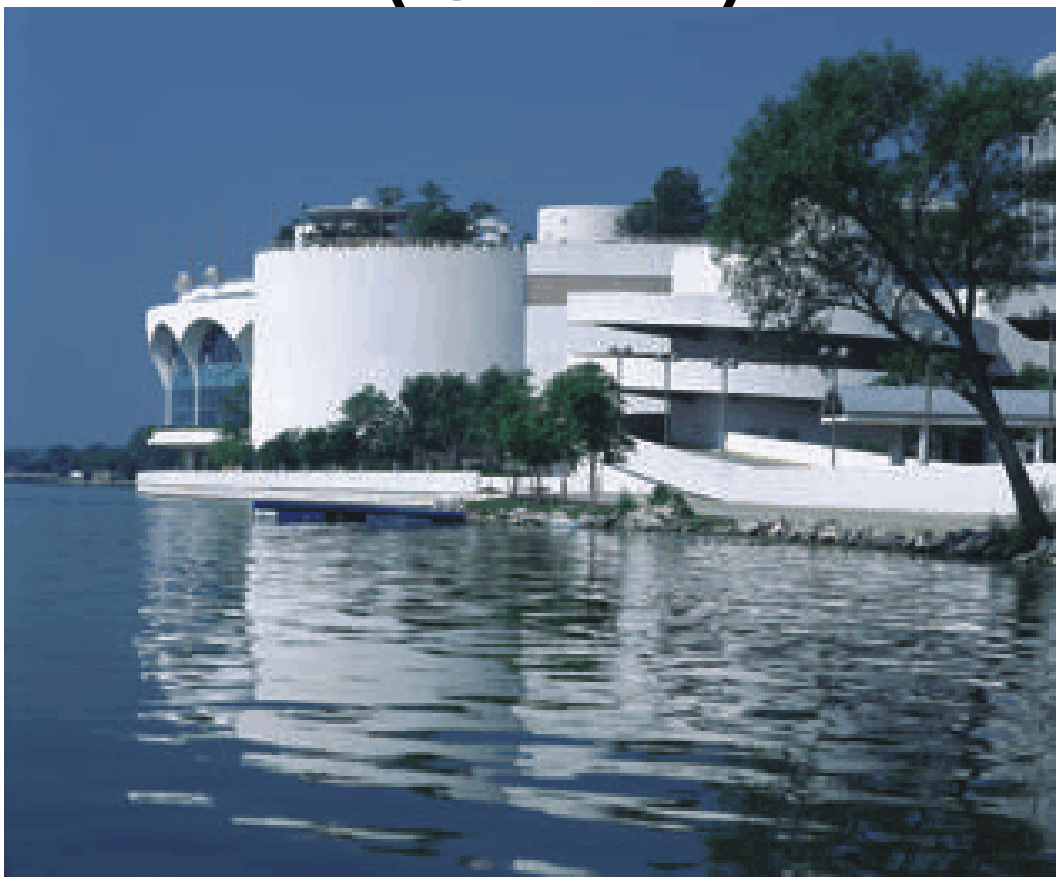
image
AND
vision
COMPUTING


Machine Vision
and Applications
An international journal

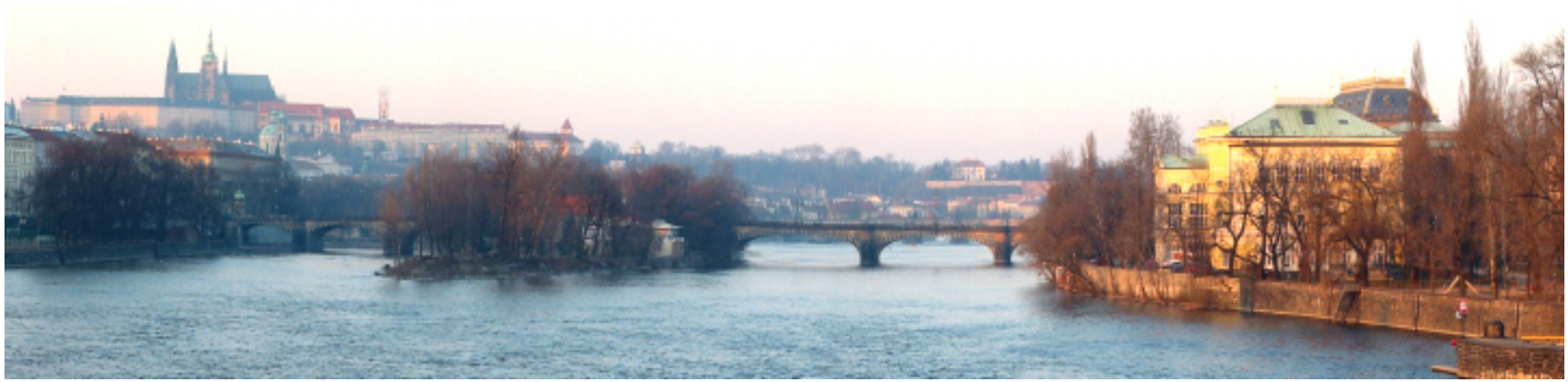
Computer Vision Conferences

International Conference on Computer Vision (ICCV)

IEEE Conference on Computer Vision and Pattern Recognition (CVPR)

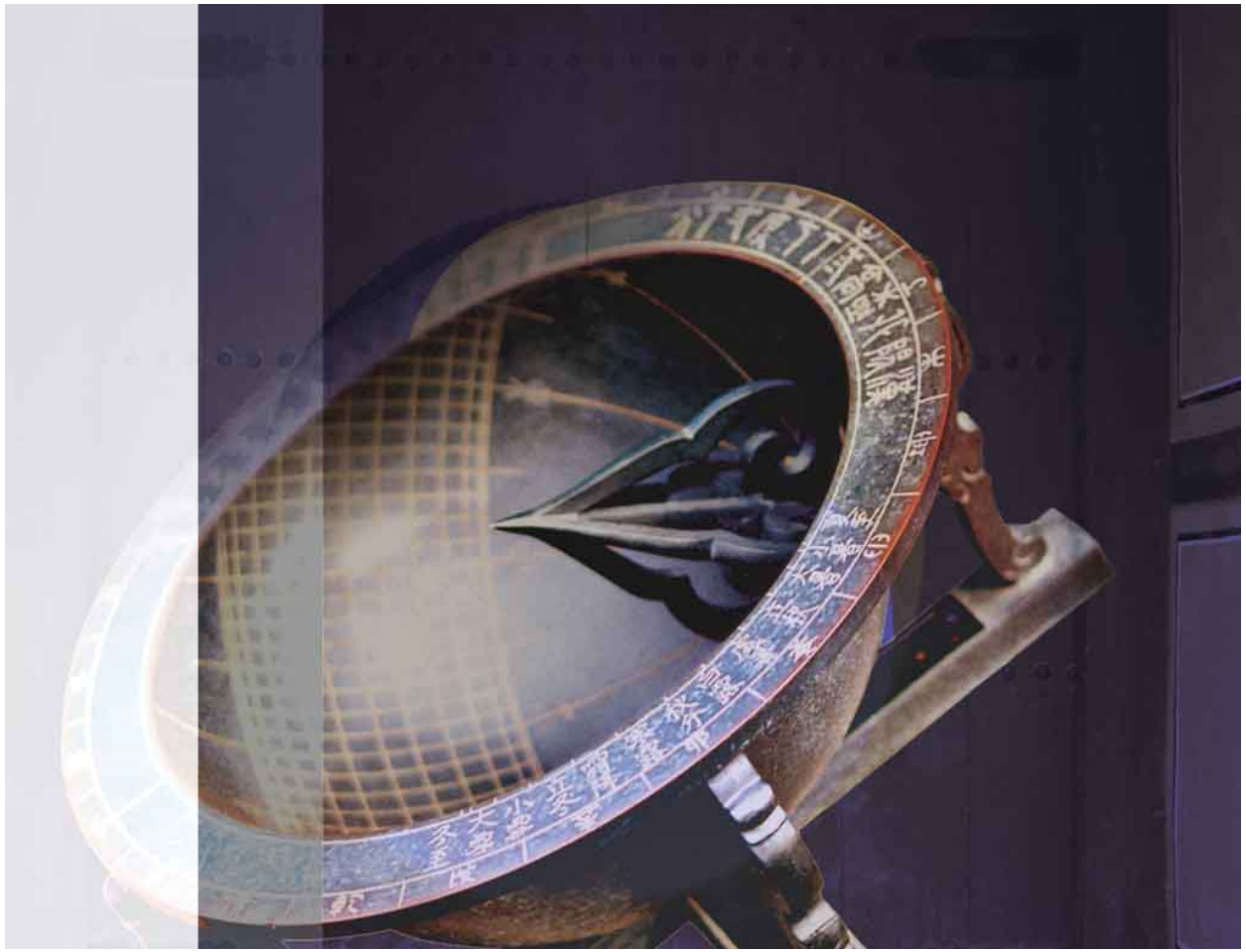


European Conference on Computer Vision (ECCV)



International Conference on Pattern Recognition (ICPR)

Asian Conference on Computer Vision (ACCV)



International Conference on Image Processing

Computer Vision at UCF

- Started in August 1986
- Developed four courses
  - Intro to Robot Vision
  - Computer Vision
  - Computer Vision Systems
  - Advanced Computer Vision

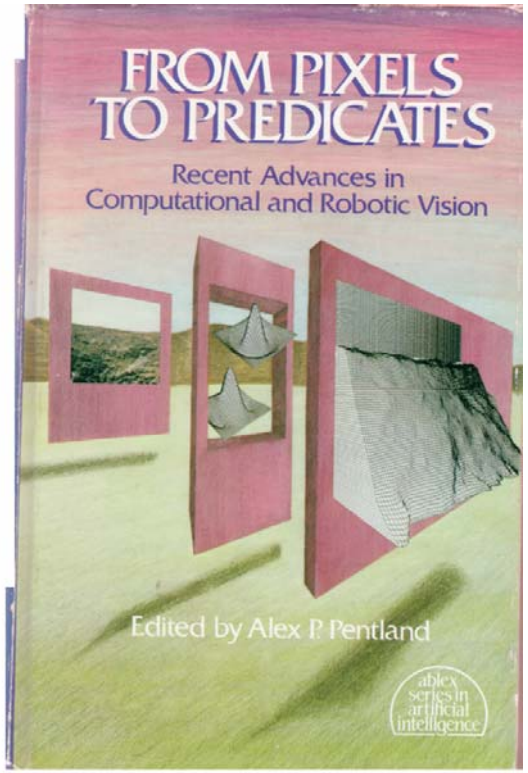
- Graduated first Ph.D. student in 1989
- Dr. Lobo joined in 1992
- Dr. Foroosh joined in 2002

Vision Books Used at UCF

FUNDAMENTALS OF COMPUTER VISION[1]

Mubarak Shah
Computer Science Department
University of Central Florida
Orlando, FL 32816

December 7, 1997

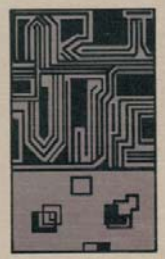

FROM PIXELS TO PREDICATES
Recent Advances in Computational and Robotic Vision
Edited by Alex P. Pentland
ablex series in artificial intelligence

VISION

DAVID MARR

Model-Based Computer Vision

Rodney Allen Brooks

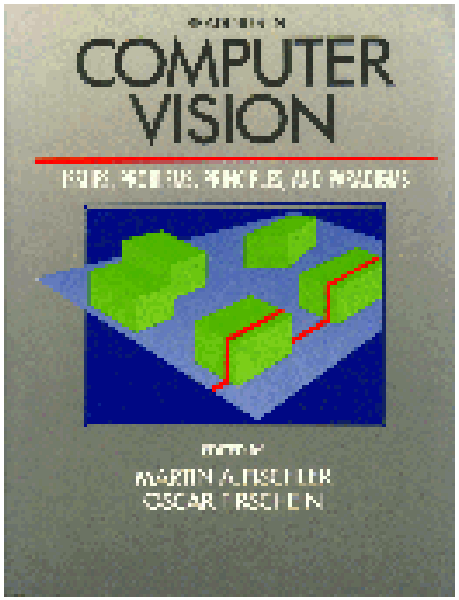


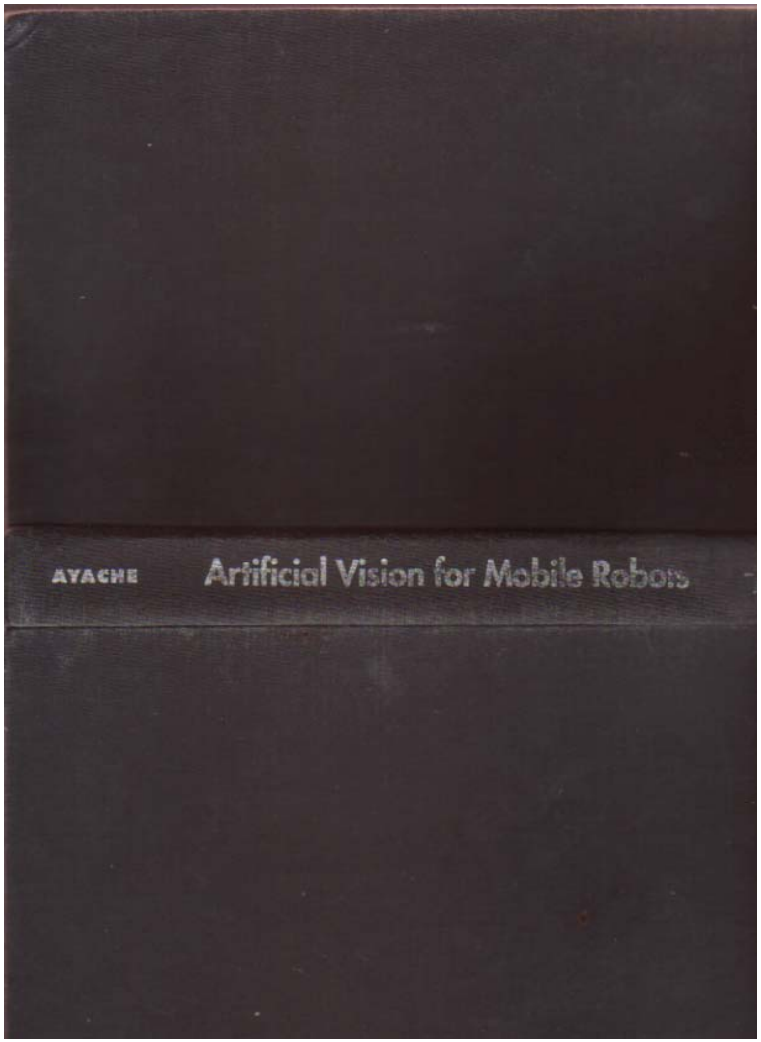
UMI Research Press
Computer Science: Artificial Intelligence

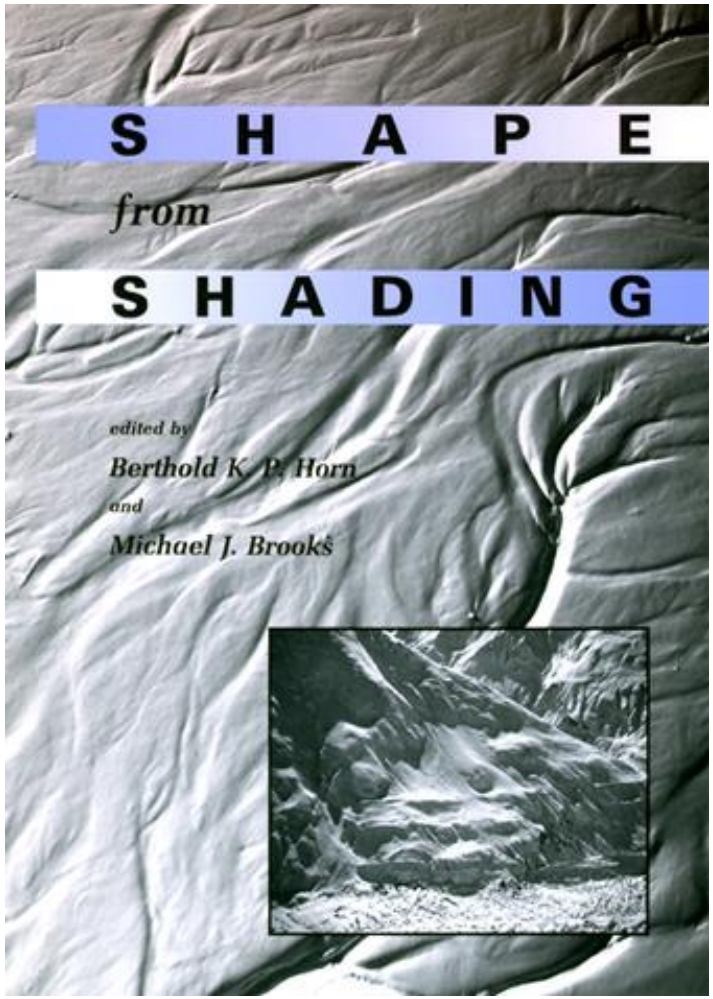
Perceptual Organization and Visual Recognition

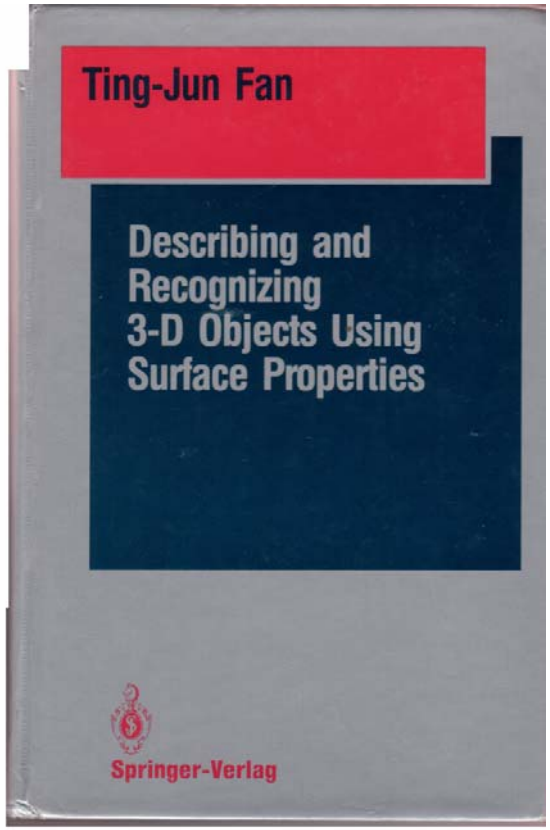
David Lowe

Kluwer Academic Publishers, 1985


COMPUTER
VISION


AYACHE
Artificial Vision for Mobile Robots


SHAPE
from
SHADING
edited by
Berthold K. P. Horn
and
Michael J. Brooks


Ting-Jun Fan
Describing and Recognizing 3-D Objects Using Surface Properties
Springer-Verlag

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

JUNE 1993 VOLUME 15 NUMBER 6 ITPIDJ (ISSN 0162-8828)

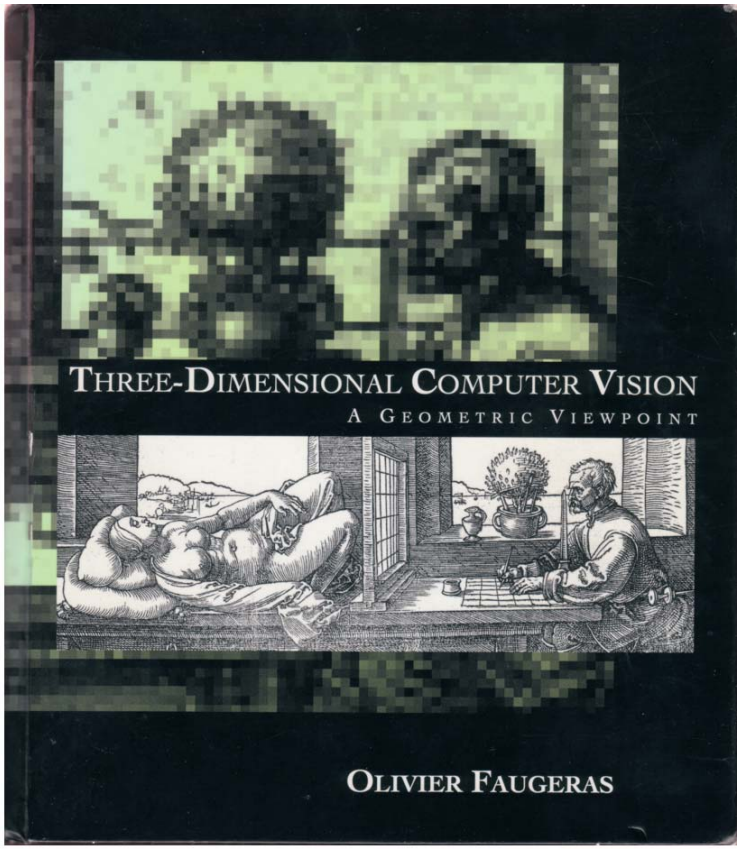
PUBLISHED BY THE IEEE COMPUTER SOCIETY

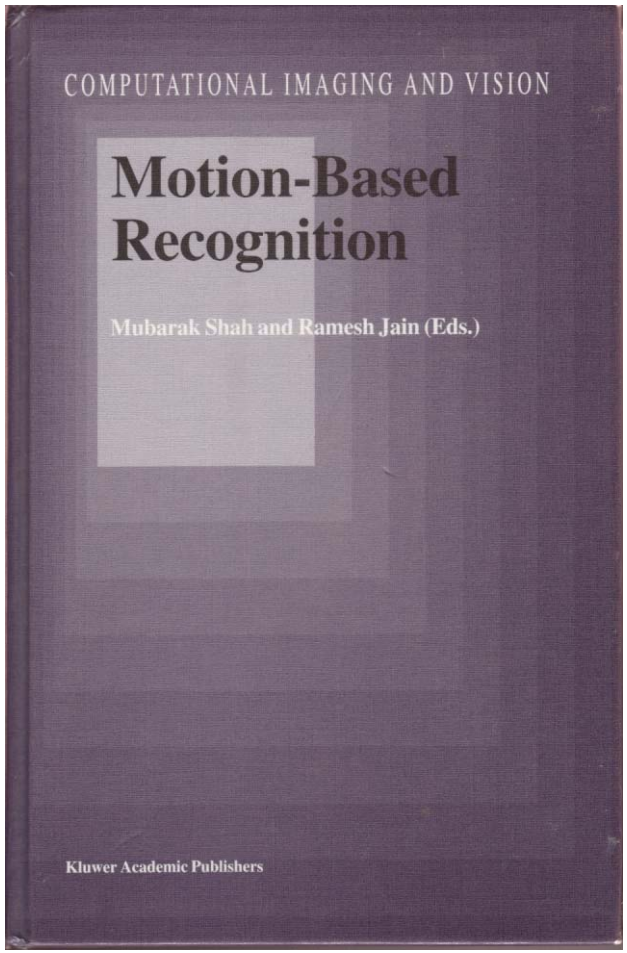
In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Society
Systems, Man, and Cybernetics Society
Ultrasonics, Ferroelectrics, and Frequency Control Society

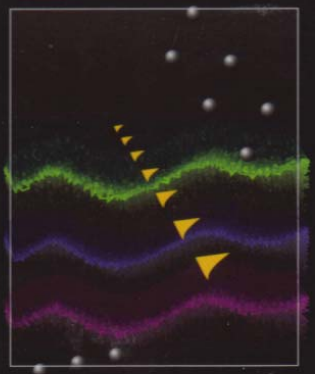
SPECIAL SECTION ON 3-D MODELING IN IMAGE ANALYSIS AND SYNTHESIS

REGULAR SECTION


THREE-DIMENSIONAL COMPUTER VISION
A GEOMETRIC VIEWPOINT
OLIVIER FAUGERAS


COMPUTATIONAL IMAGING AND VISION
Motion-Based Recognition
Mubarak Shah and Ramesh Jain (Eds.)
Kluwer Academic Publishers

digital video processing



A. MURAT TEKALP

PRENTICE HALL SIGNAL PROCESSING SERIES

Vision Ph.D. Graduates

Donna J Williams, 1989

- Donna Williams, and Mubarak Shah. “Edge Characterization Using Normalized Edge Detection”, CVGIP: Graphical Models and Image Processing, Vol. 55, No. 4, July 1993, pp 311-318.

- Donna Williams, and Mubarak Shah. “A Fast Algorithm for Active Contours and Curvature Estimation”, Computer Vision, Graphics and Image Processing, Vol 55, No.1, January 1992, pp 14-26.

- Donna Williams, and Mubarak Shah. “Edge Contours Using Multiple Scales”, Computer Vision, Graphics and Image Processing, September 1990, Volume 51, pp 256-274.

Krishnan Rangarajan, 1990

- Krishnan Rangarajan, Bill Allen, and Mubarak Shah. “Matching Motion Trajectories”, Pattern Recognition, Vol. 26, No. 4, pp 595-610, April, 1993.

- Mubarak Shah, Krishnan Rangarajan, and Ping-Sing Tsai. “Motion Trajectories”, IEEE Transaction on Systems, Man and Cybernetics, Vol. 23, No. 4, August 1993, pp 1138-1150.

- Krishnan Rangarajan, and Mubarak Shah. “Interpretation of Motion Trajectories Using Focus of Expansion", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 14, No. 12, December 1992.

- Krishnan Rangarajan, and Mubarak Shah. "Establishing Motion Correspondence”, Computer Vision, Graphics and Image Processing: Image Understanding, July 1991, pp 56-73.

- Krishnan Rangarajan, Mubarak Shah, and David Van Brackle. “Optimal Corner Detector”, Computer Vision, Graphics and Image Processing, vol. 48, pp 230-245, November 1989.

Ping-Sing Tsai, 1995

- Ping-Sing Tsai and Mubarak Shah. “Shape from Shading with Variable Albedo", Optical Engineering, pp 121-1220, April 1998.
- James Cryer, Ping-Sing Tsai and Mubarak Shah. “Shape from Shading and Stereo", Pattern Recognition, Volume 28, No. 7, pp 1033-1043, July 1995.
- Ping-sing Tsai and Mubarak Shah. “Shape From Shading Using Linear Approximation” Image and Vision Computing Journal, 1994.
- Ping-Sing Tsai, Mubarak Shah, Katharine Keiter, and Takis Kasparis. “Cyclic Motion Detection for Motion Based Recognition”, Pattern Recognition, Vol. 27, No. 12, 1994.
- Ping-Sing Tsai and Mubarak Shah. ”Shape from shading using linear approximation”, Image and Vision Computing, Vol. 12, No. 8, October 1994, pp 487-498

Ruo Zhang, 1995

- Ruo Zhang,Ping-Sing Tsai, James Cryer and Mubarak Shah. “Shape from Shading: A Survey”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 21, Number 08, August, 1999, pp 690-706.

- Ruo Zhang and Mubarak Shah. “Shape from Intensity Gradient", IEEE Transactions on Systems, Man and Cybernetics, PART A, May 1999.

- Ruo Zhang and Mubarak Shah. “Iterative Shape Recovery From Multiple Images", Image and Vision Computing, Volume 15 (1997), 801-814, November 1997.

- Ruo Zhang, Ping-Sing Tsai and Mubarak Shah. “Photomtion", CVGIP: Image Understanding, Vol. 63, No. 2, pp 221-231, March 1996.

Yu Tian, 1997

- Yu Tian and Mubarak Shah. “Recovering 3D Motion and Structure using Adaptive Hough Transform", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol19, N0. 10, October 1997, pp 1178-1183.

- Yu Tian and Mubarak Shah. “Motion Estimation and Segmentation", Machine Vision and Applications, vol 9, pp 32-42, 1995

Niels Haering, 1999

- Niels Haering, Richard Qian, and Ibrahim Sezan, *"A Semantic Event Detection Approach and Its Application to Detecting Hunts in Wildlife Video"*, IEEE Transactions on Circuits and Systems for Video Technology, 1999.

- Niels Haering and Niels da Vitoria Lobo, *"Features and Classification Methods to Annotate Images with their Deciduous Tree Content"*, Journal of Computer Vision and Image Understanding, 1999.

Sohaib Khan, 2002

- From Thesis
  - S. Khan, M. Shah, “Tracking in the Presence of Occlusion”, ACCV 2000
  - S. Khan, O. Javed, Z. Rasheed, M. Shah, “Camera Handoff: Tracking in Multiple Uncalibrated Stationary Cameras”, HUMO 2000
  - S. Khan, O. Javed, Z. Rasheed, M. Shah, “Human Tracking in Multiple Cameras”, ICCV 2001
  - S. Khan, O. Javed, M. Shah, “Tracking in Uncalibrated Stationary Cameras with Overlapping Field of View”, PETS 2001, in conjunction with CVPR 2001
  - S. Khan, M. Shah, “Consistent Labeling of Tracked Objects in Multiple Cameras with Overlapping Fields of View”, PAMI (in press)
  - S. Khan, M. Shah, “Object-Based Segmentation of Video using Color, Motion and Spatia Information”, CVPR 2001
  - S. Khan, M. Shah, “Object Based Video Segmentation using Multiple Cues”, under preparation…
- Others
  - O. Javed, S. Khan, Z. Rasheed, M. Shah, “A Framework for Segmentation of Interview Videos”, IMSA 2000
  - O. Javed, S. Khan, Z. Rasheed, M. Shah, “Visual Content Based Segmentation of Talk and Game Shows”, Intl. Journal of Computers and Applications, 24(2), 2002

Cen Rao, 2003

- "View-Invariant Representation And Recognition of Actions", International Journal of Computer Vision, Vol. 50, Issue 2, 2002
- "View Invariance in Action Recognition", Computer Vision and Pattern Recognition, CVPR 2001, Kauai, Hawaii, Dec 11-13, 2001
- "View-Invariant Representation and Learning of Human Action", IEEE Workshop on Detection and Recognition of Events in Video, Vancouver, Canada, July 8, 2001
- "A View-Invariant Representation of Human Action", International Conference on Control, Automation, Robotics and Vision, Singapore, Dec 5th-8th, 2000
- "View-invariant Alignment and Matching of Video Sequences", submitted to ICCV 2003.
- Cen Rao, Alexei Gritai, Mubarak Shah, View-invariant Alignment and Matching of Video Sequences. The Ninth IEEE International Conference on Computer Vision, Nice, France, 2003.
- Cen Rao, Mubarak Shah and Tanveer Syeda-Mahmood, Action Rectionition based onView Invariant Spatio-temporal Analysis, ACM Multimedia 2003, Nov 2-8, Berkeley, CA, USA, 2003.

Zeeshan Rasheed, 2003

University of Central Florida VISION

- Book Chapter
  - Video Categorization using Semantics and Semiotics, In Video Mining Techniques, KLUWER Acader Publishers, 2003
- Journal Publications
  - Zeeshan Rasheed and Mubarak Shah, ``Scene Segmentation of Hollywood Movies and TVShows", IEI Transactions on Multi-media, accepted.
  - Classification Of Hollywood Genres From Previews, IEEE Transactions on Circuit and Systems for Video Technology
  - Visual Content Based Segmentation of Talk and Game Show, International Journal of Computers and Applications (ACTA press), June 2002
- Conference and Workshop Publications
  - Scene Detection In Hollywood Movies and TV Shows, IEEE Computer Vision and Pattern Recognitio Conference, 2003
  - Movie Genre Classification By Exploiting Audio-Visual Features Of Previews, IEEE International Conference on Pattern Recognition, 2002
  - A Framework for Segmentation of Talk & Game Shows, IEEE International Conference on Computer Vision, 2001
  - A Framework for Segmentation of Interview Vides, IASTED International Conference on Internet and Multimedia Systems and Applications, 2000
  - Semantic Film Preview Classification Using Low-Level Computable Features, 3rd International Workshop on Multimedia Data and Document Engineering 2003
  - A Graph Theoretic Approach for Scene Detection in Produced Videos, Multimedia Information Retriev Workshop 2003

Zeeshan Rasheed, 2003

- Tracking Across Multiple Cameras With Disjoint Views, Omar Javed, Zeeshan Rasheed, Khurram Shafique and Mubarak Shah, IEEE International Conference on Computer Vision, Nice, France, 2003
- M-KNIGHT: A Real-time Surveillance System for Multiple Overlapping and Non-overlapping Camera, Omar Javed, Zeeshan Rasheed, Orkun Alatas and Mubarak Shah, IEEE International Conference on Multimedia and Expo, 2003
- Human Tracking in Multiple Camera, Sohaib Khan, Omar Javed, Zeeshan Rasheed and Mubarak Shah, IEEE International Conference on Computer Vision, 2001
- Camera handoff: Tracking in Multiple Uncalibrated Stationary Cameras, IEEE Workshop on Human Motion, 2000

Alper Yilmaz, 2004

- Journals
  - "Object Tracking: A Survey," *submitted to ACM Computing Survey, 2004*.
  - "Contour Based Object Tracking with Occlusion Handling in Video Acquired Using Mobile Cameras," accepted for publication IEEE PAMI, 2004.
  - "Target Tracking in Airborne Forward Looking Infrared Imagery," IVC, Vol. 21, No. 7, 2003, pp. 623-635.
  - "View Invariant Representation and Recognition of Actions," IJCV, Vol. 50, No. 2, 2002, pp. 203-226.
- Conferences
  - "Spatio-Temporal Volume Sketch: A Novel Representation of Actions," *submitted to NIPS, 2004*.
  - "Actions As Objects: A Novel Approach for Representation, Recognition and Retrieval of Human Actions," *submitted to ACM Multimedia Conference, 2004*.
  - "Object Tracking Using Level Sets," proc. of ACCV, Korea, 2004.
  - "Automatic Feature Detection and Pose Recovery for Faces," proc. of ACCV, 2002, pp. 284-289.
  - "Estimation of Arbitrary Albedo and Shape from Shading for Symmetric Objects," proc. of BMVC, England, 2002, pp. 728-736.
  - "Target Tracking in FLIR Imagery Using Mean-Shift and Global Motion Compensation," proc. of IEEE CVBVS, Hawaii, 2001.
  - "Estimation of Rigid and Non-rigid Motion Using Anatomical Face Model," proc. of ICPR, Vol. 1, Canada, 2002, pp. 377-380.
  - "Shot Detection Using Principal Coordinate System," proc. of IASTED IMSAC, November 2000, pp.168.

Current Ph.D. Students

- Omar Javed
- Alper Yilmaz
- Orkun Alatas
- Lisa Spencer
- Yaser Shaikh
- Jiangjian Xiao
- Yun Zhai
- Asad Hakeem
- Yunjun Zhang
- Alexei Gritai
- Paul Smith
- Imran Junejo
- Lisa Spencer
- Saad Ali
- Xiochun Cao
- Fahad Rafi

End of Story