Computer Vision

1

CS 461 – Artificial Intelligence Pinar Duygulu Bilkent University, Spring 2006

Why study Computer Vision?

- An image is worth 1000 words
- Images and movies are everywhere;
- Fast-growing collection of useful applications
 - building representations of the 3D world from pictures
 - automated surveillance (who's doing what)
 - movie post-processing
 - face finding
- Greater understanding of human vision
- Challenge: To develop human-level capabilities for computers and robots
 - Various deep and attractive scientific mysteries

What do you see in the picture?



adapted from Matial Hebert, CMU

What do you see in the picture?



Black background Two objects One teapot One toy There is a light coming from right One object is shiny the other is not

Toy:

Consists of 5 layers, in different colors There is a text : Fisher Price The layers are in donut shape Layers are plastic Bottom is wood

Teapot:

Consists of body and handle Body is metal Handle is ceramic Handle: Dark blue on white Body : golden Reflection of toy on the body

adapted from Matial Hebert, CMU

Challenge – What do you see in the picture?



adapted from Octavia Camps, Penn State

Challenge – What do you see in the picture?



A hand holding a man A hand holding a shiny sphere An Escher drawing

adapted from Octavia Camps, Penn State

People draw distinctions between what is seen

- This could mean
 - "is this a fish or a bicycle?"
 - "is this George Washington?"
 - "is this poisonous or not?"
 - "is this slippery or not?"
 - "will this support my weight?"
- Great mystery
 - How to build programs that can draw useful distinctions based on image properties.

Perception and grouping



We need some sort of perceptual organization process that tells us what "low-level" measurements might "group" together There are many different cues, including multiple views (motion, stereopsis) texture shading

What?





How many?





Tasks



- Verification
- Detection (+Localization)
- Classification / Recognition
- Grouping
- Analogy

...

Issues

- Representation
- Recognition
- Learning

Models: appearance+shape







Fischler & Elschlager, 1973

Perrett & Oram, 1993

Perona et al. '95



Schmid '99, Lowe '99, Moreels '04



Belongie et al. '02

(Interest points) Local appearance Shape / deformation (Clutter) Correspondence

Correspondence





Adapted from Pietro Perona, Object Recognition Workshop, 2004



Occlusion and `unreliable' features





occlusion









Deformations



Clutter































































































- Lavar





















































This is a pottopod





Previously learnt categories Spotted cats Smart Airplanes box airOnunio Motorcycles New category

Training example

FeiFei et al '03

Priors on geometry





റ

О

Similarity metric



ie eraquis ressemble b Louis-Philippe, vous conductiones done?





wary of landra conductor colused, gol rememble as pression.



poire, qui rescentile aux croquie printines.





Form and function



Context



Murphy et al., ICCV2003

Context



(a) Isolated object (b) Object in context (c) Low-res Object

Murphy et al., ICCV2003

Object Recognition

- Model based vision
- Object Recognition as template matching









Parameters: 3D position and orientation

David Lowe



"Filter" image to find brightness changes.

David Lowe



"Fit" lines to the raw measurements.

David Lowe



"Project" model into image and "match" to lines (solving for 3D pose).

David Lowe



Biederman's Geons

Possible approach: If line drawings are easy to recognize then maybe we should first find lines.

Line drawings





University of South Florida
... vs edge detection



University of South Florida



University of South Florida



Match "model" to measurements?



39







Figure from "Object recognition using alignment," D.P. Huttenlocher and S. Ullman, Proc. Int. Conf. Computer Vision, 1986, copyright IEEE, 1986



Figure from "The evolution and testing of a model-based object recognition system", J.L. Mundy and A. Heller, Proc. Int. Conf. Computer Vision, 1990 copyright 1990 IEEE Adapted from David Forsyth, UC Berkeley



Figure from "The evolution and testing of a model-based object recognition system", J.L. Mundy and A. Heller, Proc. Int. Conf. Computer Vision, 1990 copyright 1990 IEEE Adapted from David Forsyth, UC Berkeley



Figure from "The evolution and testing of a model-based object recognition system", J.L. Mundy and A. Heller, Proc. Int. Conf. Computer Vision, 1990 copyright 1990 IEEE Adapted from David Forsyth, UC Berkeley



Figure from "The evolution and testing of a model-based object recognition system", J.L. Mundy and A. Heller, Proc. Int. Conf. Computer Vision, 1990 copyright 1990 IEEE

Application: Surgery

- To minimize damage by operation planning
- To reduce number of operations by planning surgery
- To remove only affected tissue
- Problem
 - ensure that the model with the operations planned on it and the information about the affected tissue lines up with the patient
 - display model information supervised on view of patient
 - **Big Issue**: coordinate alignment, as above





NMI

USI

Reprinted from Image and Vision Computing, v. 13, N. Ayache, "Medical computer vision, virtual reality and robotics", Page 296, copyright, (1995), with permission from Elsevier Science Adapted from David Forsyth, UC Berkeley











- •Some objects are 2D patterns – e.g. faces
- •Build an explicit pattern matcher
- discount changes in illumination by using a parametric model
- changes in background are hard
- changes in pose are hard

Face detection



http://www.ri.cmu.edu/projects/project_271.html

Face detection



http://www.ri.cmu.edu/projects/project_320.html

Templates and relations

- e.g. find faces by
- finding eyes, nose, mouth
- finding assembly of the three that has the "right" relations



Patch Model



http://www.research.ibm.com/ecvg/biom/facereco.html

Parts and relations



How flexible are the spatial relations of the parts?

How good are our models?





Thompson, P. (1980). "Margaret Thatcher: a new illusion." Perception 9:483-484

How good are our models?





Thompson, P. (1980). "Margaret Thatcher: a new illusion." Perception 9:483-484

Is it only about matching?

What are our "models"? How good are they?



Ron Rensick

Is it only about matching?



Context



Antonio Torralba

Context



a person?

Antonio Torralba

Context



a person?

Antonio Torralba

the blob is identical to the one on the previous slide after a 90deg rotation

Prior Expectations



Giuseppe Arcimboldo

Inverse problems. Recover information that is lost. Make explicit information that is implicit.

Understand geometry and physics of light and world.

Our measurements are always ambiguous. This means perception involves *inference*. We must use our prior information about the world and the combination of evidence from multiple cues to infer what is in the world.

Understand probabilistic inference.



http://www.well.com/user/jimg/index.html







http://www.well.com/user/jimg/index.html

From two (or more) images, determine the geometry of the scene by *matching* corresponding areas of the images



adapted from Gregory Hager, JHU

3D Reconstruction



adapted from David Forsyth, UC Berkeley

Work by Paul Debevec and Jitendra Malik

3D Reconstruction


3D Reconstruction



Mosaicking





Mosaicking





Tracking



adapted from Martial Hebert, CMU







Visual Tracking



Hager & Rasmussen 98



Hager & Belhumeur 98



Bregler and Malik 98



Black and Yacoob 95



compuBaseletandialakep98ical Systems

Application

Track users' head gaze for hands-free pointing...



Model-based Brightness Constraints: on Direct Estimation of Structure and Motion





http://www.ai.mit.edu/people/gideon/Demos/DirectMethods/Demo1.html

Structure from Motion



Input

81

Structure from Motion





Applications of Optical flow







Impressionist effect. Transfer motion of real world to a painting

Adapted from Michael Black, Brown University

Applications

- Industrial inspection, quality control
- Surveillance and security
- Assisted living
- Human-computer interfaces
- Medical image analysis
- Reverse engineering
- Image databases

Applications : Document Analysis



Digit recognition, AT&T labs http://www.research.att.com/~yann/

Applications : 3D Scanning



Scanning Michelangelo's "The David"

- The Digital Michelangelo Project
 - http://graphics.stanford.edu/projects/mich/
- UW Prof. Brian Curless, collaborator
- 2 BILLION polygons, accuracy to .29mm

Applications : Motion Capture, Games



Applications : Medical Imaging



Applications : Robotics





Applications : Image Retrieval



Organizing Image Collections



Informedia Digital Video Library Project





IDVL interface returned for "El Nino" query along with different multimedia abstractions from certain documents.

Informedia Digital Video Library Project



IDVL interface returned for "bin ladin" query

The results can be tuned using many classifiers

Associating video frames with text



Query on "president"

Association problem

Associating Visual Features with Text

Solving correspondences in broadcast news for better retrieval



...tanks on the street



...start attacking on houses by helicopters and tanks...



.. fuel tank ...

Face Recognition by resolving correspondences between named entities and faces







Associating Names and faces





President George W. Bush makes a statement in the Rose Garden while Secretary of Defense Donald Rumsfeld looks on, July 23, 2003. Rumsfeld said the United States would release graphic photographs of the dead sons of Saddam Hussein to prove they were killed by American troops. Photo by Larry Downing/Reuters

British director Sam Mendes and his partner actress Kate Winslet arrive at the London premiere of 'The Road to Perdition', September 18, 2002. The films stars Tom Hanks as a Chicago hit man who has a separate family life and co-stars Paul Newman and Jude Law. REUTERS/Dan Chung





World number one Lleyton Hewitt of Australia hits a return to Nicolas Massu of Chile at the Japan Open tennis championships in Tokyo October 3, 2002. REUTERS/Eriko Sugita

German supermodel Claudia Schiffer gave birth to a baby boy by Caesarian section January 30, 2003, her spokeswoman said. The baby is the first child for both Schiffer, 32, and her husband, British film producer Matthew Vaughn, who was at her side for the birth. Schiffer is seen on the German television show 'Bet It...?!' ('Wetten Dass...?!') in Braunschweig, on January 26, 2002. (Alexandra Winkler/Reuters)

US President George W. Bush (L) makes remarks while Secretary of State Colin Powell (R) listens before signing the US Leadership Against HIV /AIDS, Tuberculosis and Malaria Act of 2003 at the Department of State in Washington, DC. The five-year plan is designed to help prevent and treat AIDS, especially in more than a dozen African and Caribbean nations(AFP/Luke Frazza)



Incumbent California Gov. Gray Davis (news - web sites) leads Republican challenger Bill Simon by 10 percentage points – although 17 percent of voters are still undecided, according to a poll released October 22, 2002 by the Public Policy Institute of California. Davis is shown speaking to reporters after his debate with Simon in Los Angeles, on Oct. 7. (Jim Ruymen/Reuters)



Berg et al. CVPR 2004



President Daniel Arap Moi

Witherspoon

Actress Jennifer Aniston

Martha Stewart

