Introduction

Selim Aksoy Department of Computer Engineering Bilkent University saksoy@cs.bilkent.edu.tr

What is computer vision?

- Analysis of digital images by a computer.
- Stockman and Shapiro: making useful decisions about real physical objects and scenes based on sensed images.
- Trucco and Verri: computing properties of the 3D world from one or more digital images.
- Ballard and Brown: construction of explicit, meaningful description of physical objects from images.
- Forsyth and Ponce: *extracting descriptions of the world from pictures or sequences of pictures.*

Why study computer vision?

- Possibility of building intelligent machines is fascinating.
- Capability of understanding the visual world is a prerequisite for such machines.
- Much of the human brain is dedicated to vision.
- Humans solve many visual problems effortlessly, yet we have little understanding of visual cognition.

Why study computer vision?

- An image is worth 1000 words.
- Images and videos are everywhere.
- Fast growing collections and many useful applications.
- Goals of vision research:
 - Give machines the ability to understand scenes.
 - Aid understanding and modeling of human vision.
 - Automate visual operations.

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



Subjective contours



Subjective contours



CS 484, Spring 2007



a b c

FIGURE 2.8 Examples of simultaneous contrast. All the inner squares have the same intensity, but they appear progressively darker as the background becomes lighter.

Adapted from Gonzales and Woods





Occlusion

* types of "junctions" give cues about surfaces, occlusion, and light.

I. Rock, The Logic of Perception, 1983.



- The shape of junctions constrains the possible interpretations of the scene.
- Ambiguous: paint and surface boundaries can be confused.

- How can different cues such as color, texture, shape, motion, etc., can be used for recognition?
 - Which parts of image should be recognized together?
 - How can objects be recognized without focusing on detail?
 - How can objects with many free parameters be recognized?
 - How do we structure very large model bases?

Color







Adapted from Martial Hebert, CMU

Texture



Segmentation



Adapted from Linda Shapiro, U of Washington

Segmentation







Adapted from Jianbo Shi, U Penn

Shape



Recognized objects

Adapted from Enis Cetin, Bilkent University

CS 484, Spring 2007

Model database

©2007, Selim Aksoy

Motion







Parameters: 3D position and orientation



"Filter" image to find brightness changes.



"Fit" lines to the raw measurements.



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







©2007, Selim Aksoy





Detection



Detection



Detection

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



Parts and relations



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

Parts and relations



How flexible are the spatial relations of the parts?



Adapted from Antonio Torralba, MIT



Adapted from Antonio Torralba, MIT







Adapted from Derek Hoiem, CMU



Adapted from Derek Hoiem, CMU

Stages of computer vision

- Low-level
 - image \rightarrow image
- Mid-level

 image → features / attributes
 Image analysis / image understanding

 High-level

 features → "making sense", recognition
Low-level

sharpening



blurring

Adapted from Linda Shapiro, U of Washington

Low-level



Mid-level



Adapted from Linda Shapiro, U of Washington

Low-level to high-level



Adapted from Linda Shapiro, U of Washington

Applications

- Industrial inspection, quality control
- Medical image analysis
- Remote sensing
- Security
 - Surveillance
 - Biometrics
 - Target recognition
 - Tracking

- Robotics
- Document analysis
- Multimedia
- Assisted living
- Human-computer interfaces

Medical image analysis



CT image of abdomen

Adapted from Linda Shapiro, U of Washington

Medical image analysis



Cancer detection and grading

CS 484, Spring 2007

©2007, Selim Aksoy

Medical image analysis



Slice of lung

Adapted from Linda Shapiro, U of Washington

Land cover analysis



Object recognition



Recognition of buildings and building groups



Finding similar regions: airports

Tracking



Adapted from Octavia Camps, Penn State

Tracking



Adapted from Martial Hebert, CMU

Tracking



Adapted from Martial Hebert, CMU

Biometrics



Biometrics



Adapted from Anil Jain, Michigan State

Robotics



Adapted from Linda Shapiro, U of Washington

Robotics



Adapted from Steven Seitz, U of Washington

Document analysis



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

Adapted from Steven Seitz, U of Washington

Document analysis



I looked as hard as I could see, beyond 100 plus infinity an object of bright intensity – it was the back of me!

Figure 1.5: (Left) Chinese characters and (right) English equivalent. Is it possible that a machine could automatically translate one into the other? Chinese characters and poem courtesy of John Weng.

Adapted from Shapiro and Stockman

Document analysis



CMS/Fisher HealthCare

Blood Bank / Dylmbans



Model 145 Isotemp* Dry Bath Incubator

Holds 1 to 4 heating blocks with choice of 11 well sizes Maintains even; asronie lo within til 12 C of lemoeratur

* tolek /S 5 mp

une reactions. Inactive: Leave Ob studies, slove any machine and the aster limanelosis 51, x 15 W x 4° H (28 x 28, x 13 pm). With the pare and i in Healing blocks sold separately (see lower on Each

ATE 34 995-25

m

Crd No. Rectrical Requirements телик талан не, элеки (/ SA, арритите) элеккөздер на нарад 11-715-100

Average risk ador. Lom mean of 3<mark>71 O</mark> Polama jotal Madel

Incu-Block* Partial Immersion Thermometers

For all standard bas; no blocks and water baths. Critical temperatures [251, 001, 371, 581 G) are marked with a now. Available with shattencrept, contamination proof Teller's coating. Total, ength: 176 mm. Im-

	Banga 30	DM, 3C	Tellen, France	KORL MOD	Esc.				
	25.57	0.5*	Nic	14-992.	45.a				
	25.57	D.51	Med	14-993	45.1				
	More Thermometers For more thermometers, including digital types								
		F	ee nage sozi	-					



Model 147 Isotemp" Dry Bath

 Holds single heating block with choice of 11 well sizes Similar in Model 145, but win 352 (bick 72.0 mm) mater lides, for Jahs w and HN1 C (2041 PL Ob In satiadjust estrical services hole in troat ea Supplies with strong ratios case, thermostatics indicator arro, line core and nitio, and instructions, Dimensions; B L . Rév W x X 11.15 x 17 x 8 cont CSA arranged Healing blocks solu

refeir /coa hains Rectilent Requirements Cat. No. 120V 50/90 Hz, 120V 11-715-102 223.58

Interchangeable Heating Blocks for Isotemp*. Dry Baths



The 11-715-120 clock is rche Diagnosties Planenteré programe rest tubes. This stocial shallow well clock is similar to the other block with 10 mm tokes, but semple velis are noted by detail (1,0 pr) to prest test requirements. Wells in all other blocks are 1941 been 14.4 cm.

Thins Blas, mm	WellsFinel	Cat No.	Each
6	35	11-715-105	71 19
ha	83	11-716-107	AL 15
15	2D (sht itter	11-715-120	71.19
12	12	11-715-109	71.18
po sa	12	11-716-121	71.15
13	12	11-715-111	71 12
15	12	11-715-113	71.18
16	8	11-715-123	71.9
ta	12	11-715-115	71 .8
20	li	11-716-117	(1.15
27	5	11-715-119	71.19
Secure coveries Encloseding a secure	amens (internal classify)	n wer lij	

Adapted from Linda Shapiro, U of Washington

Scene classification



Organizing image archives



CS 484, Spring 2007



UW ISL	Relevant Images:	5050B		Quit	UW	ISL	Relevant Images:			Quit
Image Database		Contraction of the second			Image D	Database				
	image006, d=0.0000	image590, d=0.0000	image005, d=0.0000	image009, d=0.0000			image274, d=-38.2099	image248, d=-37.6420	image232, d=-37.4671	image281, d=-37.0549
Query Image: image245 pqm Load Random	AMADER PR		Sector Sciences		Query image2 Load	Image: 92 ppm Random			Ø	
Database:	image252, d=0.0000	image588, d=0.0000	image084, d=0.0000	image015, d=0.0000	Data	base:	Image267, d=-36.7654	Image236, d=-36.6278	Image242, d=-36.5269	image228, d=-36.4105
Similarity Model LAR + COOC + MVG LAR + COOC + FIT LAR + COOC + Lp J Graph Theoretic Clust	THE TALL	A Constant of the second of th	anna an Aranatar an Aranatar	summerin nummerin U latti	LAR + COO LAR + COO LAR + COO	Maddel: V Model: C + MVG A iC + FIT C + Lp 7 meoretic Clust				- F
Combined Classifiers Bayes Network	image220, d=0.0000	image090, d=0.0000	image583, d=0.0000	image012, d=0.0000		d Classifiers	image273, d=-36.3317	image294, d=-36.2073	image292, d=-36.1373	image266, d=-35.9102
MARS Model ETHZ Model Relevance Feedback	Irrelevant Images:				MAR ETH: Relevance	S Model Z Model ce Feedback	Irrelevant Images:			int
Change Working Dir. Num. Retrieved: (12) ⊲11 >> << Search >>	image208,	image053,	image 539,	image049,	Change W Num. Retri	arch	image1065, d=410,5549	image752, d=401.2940	image723, d=383.1098	image736, d=385,4378



Adapted from David Forsyth, UC Berkeley

Face detection and recognition



CS 484, Spring 2007

Object recognition



CS 484, Spring 2007

©2007, Selim Aksoy

64

3D scanning



Adapted from Linda Shapiro, U of Washington

CS 484, Spring 2007

3D reconstruction



Adapted from David Forsyth, UC Berkeley

3D reconstruction



CS 484, Spring 2007

Motion capture



Adapted from Linda Shapiro, U of Washington

CS 484, Spring 2007

Visual effects



Adapted from Linda Shapiro, U of Washington

Mozaic





Adapted from David Forsyth, UC Berkeley

Mozaic



Adapted from David Forsyth, UC Berkeley

Critical issues

- What information should be extracted?
- How can it be extracted?
- How should it be represented?
- How can it be used to aid analysis and understanding?