2 Problems in Robot Vision: Recognizing objects and visual localization

> Guest Lectures by David Meger McGill CS 417 November 4 & 6, 2013



Motivation





Problem 1: Object recognition













What is object recognition?

- Input:
 - Training information about objects provided by users or the system designer
 - Perceptual data: Ideally just a single image, maybe a point cloud, video, user assistance, etc
- Output:
 - Object labels: Ideally accurate match to what you were told in training. Maybe includes "ontological" information.
 - Object locations: Ideally 3D, but maybe 2D. Maybe includes pose.



Organizations that want to recognize objects







Why?



Organizations that want to recognize objects















How?

User: This is contact lens solution



User: What objects are here?

System: Here is the contact lens solution!





Example Solutions

- 1)Template Matching: Find the area of an image that matches examples as exactly as possible
- 2)Shape Matching: Perform the same operation only on "edges"
- 3)Color matching: Forget image positions, only consider R,G,B description



Challenges of Image Recognition on Realistic Images

- Lighting
- Scale
- Position
- Orientation
- Projection
- Deformation and articulation
- Occlusion



Solution: local image features

- Some local patches are "recognizable"
- Feature detectors find repeated locations
- Feature descriptors provide invariant local representations
- Groups of features that agree on geometry are kept as object guesses

137 tentative matches



35 final matches



Estimated Object





What makes a good local feature?

- High Recall: They can be located again
- High Precision: They are not too confused with other patches
- Several Alternative Examples:
 - Harris Corners
 - SURF

_1

- -SIFT
- FREAK
- BRIEF
- etc



Harris Corners

- Operation based on strong gradients in multiple directions
 - Not edges, those can not be reliably located
 - Not flat regions, those cannot be described easily







Scale Invariant Feature Transform Descriptors



128-element SIFT feature vector



Feature Matching Algorithm

- For each feature, find the nearest neighbor (closest 128 dimensional descriptor)
- If matches are not "bi-directional", discard
- Threshold on ratio of best match score to 2nd best match score
- Everything remaining is a candidate match to check geometry



Geometric Consistency

- Features on object will be a subset of all matches, with much noise
- Idea: search for a large set of features that agree on geometry
- Algorithm: Randomized Sampling and Consensus (RANSAC)

137 tentative matches



35 final matches





Example of Live Robot Recognition: SRVC Contest

 Phase 1: Automated analysis of web to learn previously unseen objects





SRVC Contest

- Phase 1: Automated analysis of web to learn previously unseen objects
- Phase 2: Explore contest environment, collect imagery







SRVC Contest

- Phase 1: Automated analysis of web to learn previously unseen objects
- Phase 2: Explore contest environment, collect imagery
- Phase 3: Perform recognition









Curious George Robot

Powerbot mobile base with sensor tower: provides excellent indoor autonomy, long battery life, numerous sensors and ability to carry very significant computation onboard







Tilting laser rangefinder: captures >180 degree 3D scans of the environment

High-res digital camera: adapts zoom and grabs 10MP images with consumer-grade quality

Bumblebee stereo camera: high framerate visual information and narrow field-ofview depth sensing



Contest Environment







(a)



(b)



(c)



(d)







Curious George Results

- Scoring:
 - 8 of 12 specific instances
 - 4 of 8 generic categories





Result Summary

Year	Туре	Name	Result
2009	category	pumpkin	incorrect
		orange	correct
		red ping pong paddle	incorrect
		white soccer ball	incorrect
		laptop	incorrect
		dinosaur	incorrect
		bottle	correct
		toy car	incorrect
		frying pan	correct
	instance	book "I am a Strange Loop" by Douglas Hofstadter	correct
		book "Fugitive from the Cubicle Police"	incorrect
		book "Photoshop in a Nutshell"	correct
		CD "And Winter Came" by Enya	correct
		CD "The Essential Collection" by Karl Jenkins et al.	correct
		DVD "Hitchhiker's Guide to the Galaxy" widescreen	correct
		game "Call of Duty	correct
		toy Domo	correct
		Lay's Classic Potato Chips	correct
		Peperidge Farms Goldfish Baked Snack Crackers	correct
		Peperidge Farm Milano Distinctive Cookies	correct



Major challenges?

- 1)Safety
- 2)Exploration
- 3)Control of the camera to get useful images4)Comparing web images to the robot's5)Modeling a pumpkin



Navigation Approaches

- 2D occupancy grids for safe and effective navigation
- Video
- How to relate these maps to visual tasks?
- What type of algorithms are needed to provide the answers?



Area Coverage

 An uncertain map boundary is detected, and a plan to explore the **frontier** is executed (Yamauchi CIRA97)





Visual Coverage



Visual Coverage

 Use known robot model to determine area captured with each image. Plan to achieve full coverage.





Which pictures to take?

- Humans "frame" pictures to capture the desired content at the desired angle, focus, etc
- How can we replicate that with a robot?



Visual Saliency

 One way for intelligent systems to focus on useful content based on analysis of "interesting-ness"





Multi-Scale Saliency





Using 3D Data for attention

s 🙆 💽 🔳 💽 i	🗾 🕎 /canon_cam/image	🛅 bart :	🞼 Ksnapshot	4 /online_results	👯 RViz	Wicd Network Mar	KSnapshot) 🔘 🗋 🖬	. A 🗋	57% 06:31 pm
O R¥iz										-	00 8
le <u>V</u> iew <u>P</u> lugins <u>H</u> elp											
Aove Camera Select 2D Na	w Goal 2D Pose Estimate										
-love califeral bolocci 2014d	and the second sec			SNOV ARTRING OF ALL				MINING AL			
90. Electr Econ (Locor E											
10. Transforms (TE)			1035								
11. Stereo Cloud (Point Clc						GUITE STORES					
12. FilteredTiltCloud (Point						MUNITER STREET					
13. New Ground Plane Clou	-				CONTRACTOR OF CONTRACTOR						
14. AMCL Cloud (Pose 4	7				WILLIAM COLUMN						
15. Obstacles (Grid Cel										and the second	
16. Inflated Obstacles											
17. Robot Footprint (Pc	✓		Store College								
18. Global Plan (Path)	2										
19. Local Plan (Path)		CHARLEN (FILLING)									
20. Planner Plan (Path)	~	III									
21. Camera (Camera)		MULLING TO									and the second
Status: Disabled	P	The state			18 24						
Image Lopic /	canon_cam/image			2000		Total Agendan		attended and a second and a second attended attende			
Alpha U					A CONTRACT			and the second sec			
22. Shapshot Cloud (Pu						 Apple 19 (2) 	A CONTRACT OF THE OWNER OWNER OF THE OWNER OWNE				
24. Table Markers (Ma							5				
Status: OK					1997 - 1997 -	Section 1	~				
Marker Topic /	table_object_detector/vis								and the second se		
Namespaces					A REAL PROPERTY.						
poses	v					1 4 6 1 1 K K K	23				
map_interest_regions [50	1.1.1.1.1.1.1.1	7				
clusters [✓				and the second						
objects_on_table_1 [✓ 3		S.S.		alter and the						
tabletop_shapes [Contraction of the second	に見ていた。	and the second sec				
objects_on_table_2				1							
25. Pose (Pose)						Sec. Sec.					
Status: OK				A CONTRACTOR		1. Con 1					
V Zanon_cam/imag											
ne ill Time: 1259721118.95240	06	Wall Elapsed: 25	70.712957	R)5 Time: 1259721118	.952395	ROS Ela	sed: 2570.712960			Reset

Parts-Based Detection









More info at: http://people.cs.uchicago.edu/~rbg/latent/



Questions

