

RobotVision@ImageCLEF 2013 Task overview

Organizers: J. Martínez Gómez, I. García-Varea, M. Cazorla and B. Caputo



Robot Vision - Motivation

- The behaviour of a robot r at timestamp t depends on:
 - Not the topological $\langle x, y, z \rangle$ location
... but
 - Then semantic category of the place
&
 - The objects that are suitable for manipulation

Robot Vision - The Task

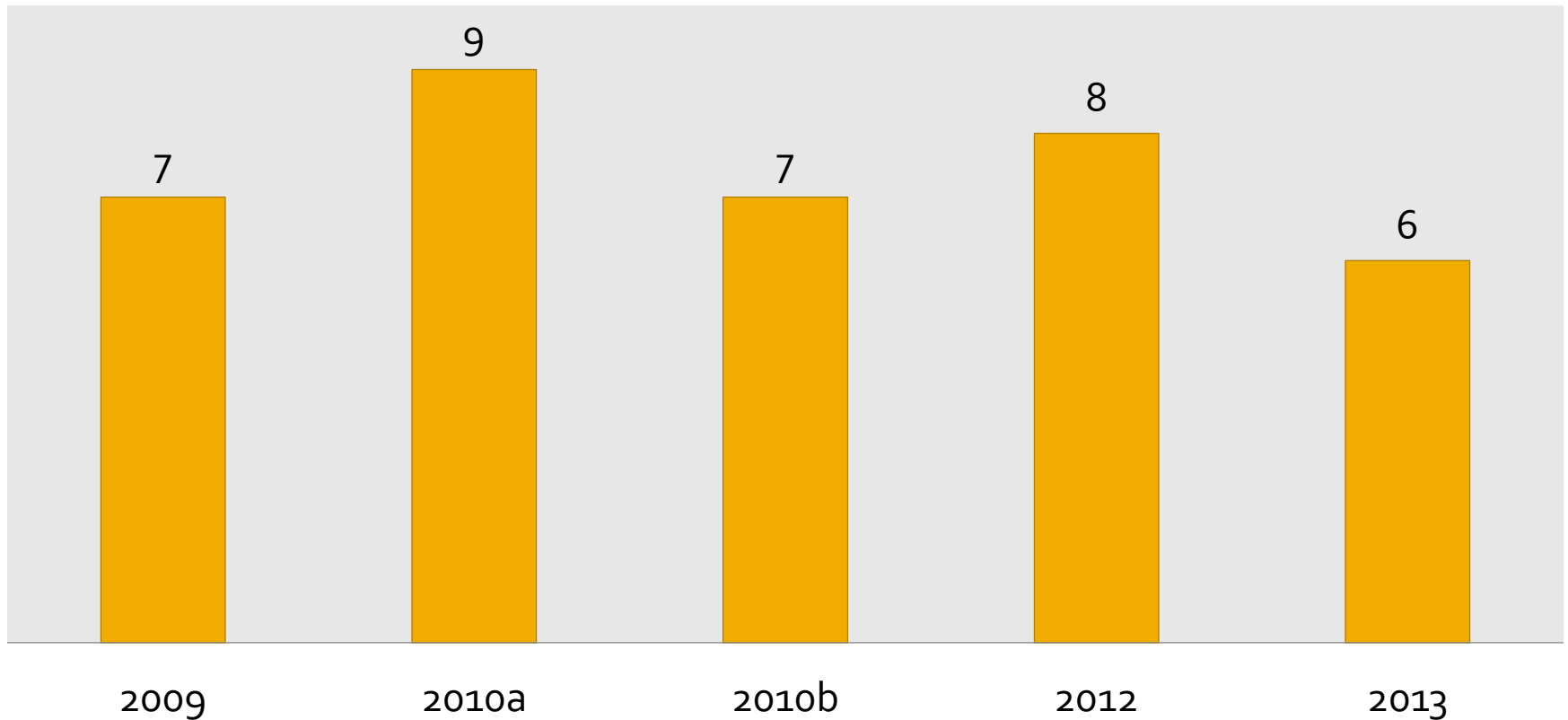
- Proposed in 2009 (5th edition)
 - ImageCLEF 2009 – Greece
 - ImageCLEF@ICPR 2010 – Turkey
 - ImageCLEF 2010 – Italy
 - ImageCLEF 2012 – Italy
 - ImageCLEF 2013 – Spain
- Considerable attention
 - ImageCLEF 2009 – 7 groups
 - ImageCLEF@ICPR 2010 – 9 groups
 - ImageCLEF 2010 – 7 groups
 - ImageCLEF 2012 – 8 groups
 - ImageCLEF 2013 – 6 groups
- Organizers
 - Jesus Martínez Gómez, Ismael García Varea, Miguel Cazorla and Barbara Caputo



Robot Vision - The Task

Evolution on the Participant Number

■ Participants

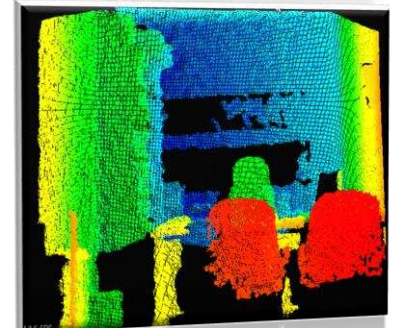


Robot Vision - Organizers

Organizer	2009	2010	2010	2012	2013
<u>B. Caputo</u>	X	X	X	X	X
A.Pronobis	X	X	X	-	-
P. Jensfelt	X	-	-	-	-
H.I. Christensen	-	X	X	-	-
M. Fornoni	-	-	X	-	-
J. Martínez-Gómez	-	-	-	X	X
I. García-Varea	-	-	-	X	X
M. Cazorla	-	-	-	-	X

Robot Vision - The Task

- Multimodal information retrieval
 - Two sources of information
 - Visual Images
 - Range Images
 - Two problems to solve
 - Presence or lack of objects in the scene
 - Semantic category of the scene
- In between computer vision and robotics



Robot Vision – The Task



- Supervised classification problem
 - Participants are provided with labelled training sequences
- Each training frame contains
 - Visual Image
 - Range Image (.pcd format)
 - Semantic category of the scene where the frame was acquired from
 - List of objects appearing in the scene
- Training and test sequences
 - Same building but with variations in the lighting conditions and acquisition procedure (counter)clockwise

Robot Vision – The Task

- The problem
 - Place classification and object detection
 - Question a.- where are you?
 - Question b.- list the objects that are in the scene
- 10 room categories
 - Corridor, Hall, Professor Office, Student Office, Technical Room, Toilet, Secretary, Visio conference, Elevator area and Warehouse
- 8 objects
 - Extinguisher, Computer, Chair, Printer, Urinal, Screen, Trash and Fridge

Robot Vision – The Task

- Performance evaluation
 - Maximum ... 2 points by frame
 - Place classification – 1 nominal problem
 - Category correctly classified → +1 point
 - Category wrongly classified → -0.5 point
 - Category not classified → +0.0 points
 - Object detection – 8 binary problems
 - Each object correctly classified → +0.125 points
 - Each object misclassified → -0.125 points
 - Each object not classified → +0.0 points

Robot Vision – The Task

Corridor



Hall



ProfessorOffice



TechnicalRoom



TechnicalRoom



Toilet



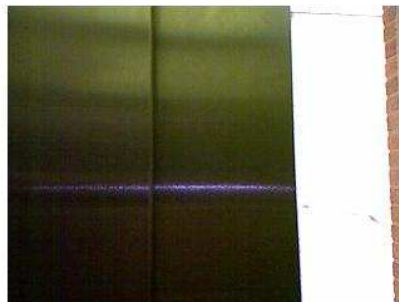
Secretary



VisioConference



ElevatorArea



Warehouse



Robot Vision – The Task

Exting.

Computer

Chair

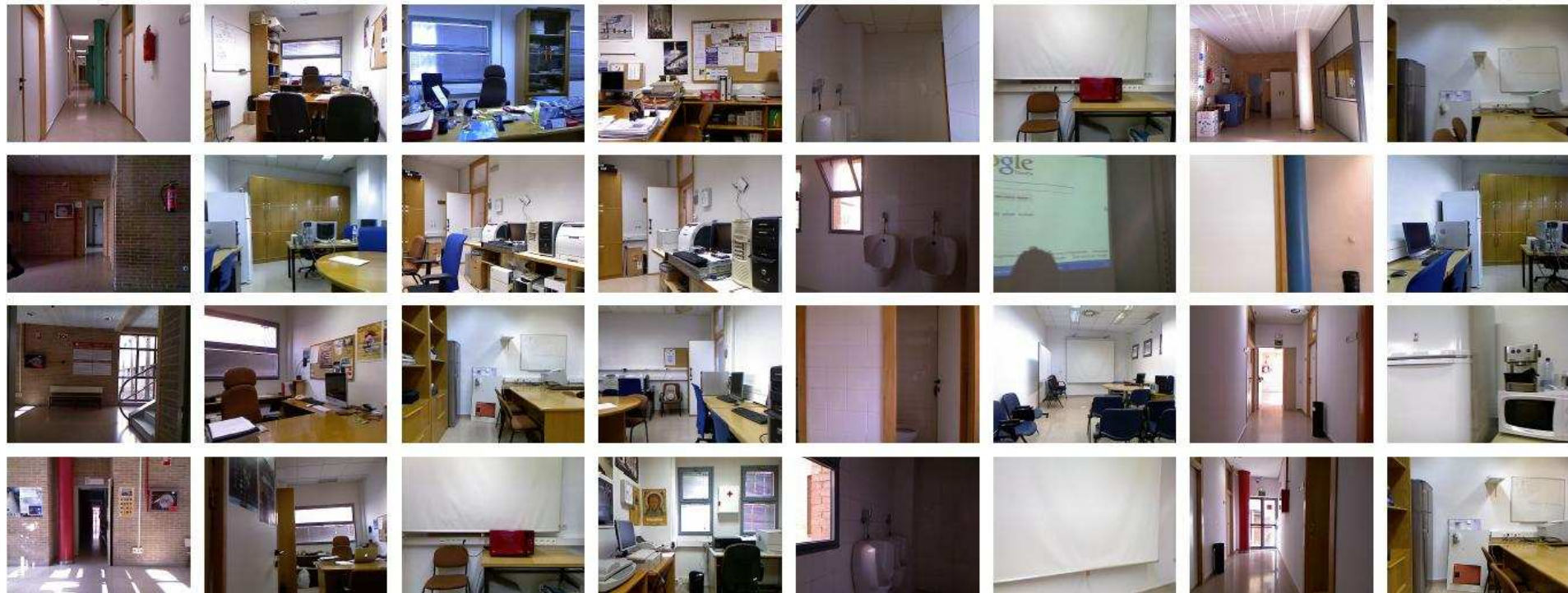
Printer

Urinal

Screen

Trash

Fridge



Robot Vision – Major Changes

Topics	2009	2010	2010	2012	2013
Place Classification	X	X	X	X	X
2 subtasks	X	X	X	X	-
Stereo Images	-	X	-	-	-
Unknown Labels	-	X	X	-	-
Object Detection	-	-	-	-	X
3D Images	-	-	-	X	X

Robot Vision – Performance Eval

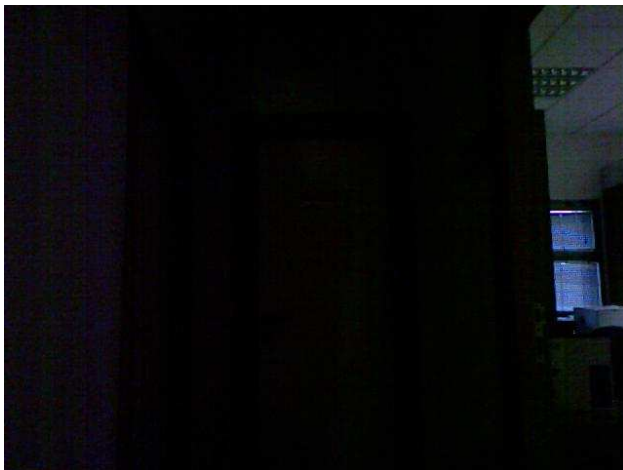
- Test frame
 - Real labels (not provided)
 - Class: Technical Room
 - Objects present: Computer and Printer
 - Participant decision
 - Class: Technical Room
 - Objects present: Computer and Trash
 - Objects not present: Chair, Printer, Urinal
 - Objects not classified: Extinguisher, Screen and Fridge

Robot Vision – Performance Eval

Class	Exting.	Comp.	Chair	Printer	Urinal	Screen	Trash	Fridge
Real labels								
Technical Room	NO	YES	NO	YES	NO	NO	NO	NO
Participant Decision								
<i>Technical Room</i>	-	YES	NO	NO	NO	-	YES	-
Right/Wrong Classification								
Hit	-	Hit	Hit	Miss	Hit	-	Miss	-
Points								
+1.0	0.0	0.125	0.125	-0.125	0.125	0.0	-0.125	0.0
Total: $1.0 + 0.125 + 0.125 - 0.125 + 0.125 - 0.125 = 1.125$								

Robot Vision – The Data

- O-Vida dataset
 - 4 sequences (2 for training, test and validation)
 - More than 10.000 frames
 - Extreme lighting conditions in the test sequence
 - Range images are highly recommended



Robot Vision – The Data

<u>Room Category</u>	Number of Frames			
	<u>Training 1</u>	<u>Training 2</u>	<u>Validation</u>	<u>Test</u>
Corridor	891	1262	764	1317
Hall	103	228	0	297
Professor Office	124	192	200	222
Student Office	155	276	282	318
Technical Room	136	281	214	240
Toilet	121	242	188	198
Secretary	98	195	181	201
Visioconference	149	300	0	306
Warehouse	70	166	0	127
Elevator Area	100	174	40	289
All	1947	3316	1869	3515

Robot Vision – The Data

Number of Presences / Lacks				
<u>Objects</u>	<u>Training 1</u>	<u>Training 2</u>	<u>Validation</u>	<u>Test</u>
Extinguisher	259 / 1688	529 / 2787	286 / 1583	520 / 2995
Computer	289 / 1658	466 / 2850	416 / 1453	473 / 3042
Chair	470 / 1477	767 / 2549	567 / 1301	889 / 2626
Printer	210 / 1737	292 / 3024	255 / 1614	279 / 3236
Urinal	054 / 1893	110 / 3206	070 / 1799	090 / 3425
Screen	081 / 1866	190 / 3126	000 / 1869	151 / 3364
Trash	406 / 1541	451 / 2865	253 / 1616	662 / 2853
Fridge	057 / 1890	101 / 3212	099 / 1770	114 / 3401
All	1826 / 13750	1909 / 23610	1946 / 13006	3178 / 24942

Robot Vision – Participation

- 39 participants but only 6 submitted one run
 - NUDT: National University of Defense Technology, Changsha, China
 - MIAR ICT: Beijing, China
 - MICA: Hanoi university of Science and Technology, Hanoi, Vietnam
 - REGIM: University of Sfax National School of Engineers, Tunisia
 - GRAM: University of Alcalá de Henares, Spain
 - SIMD: University of Castilla-La Mancha, Spain
 - Out of competition contribution using proposed techniques

Robot Vision – Overall Results

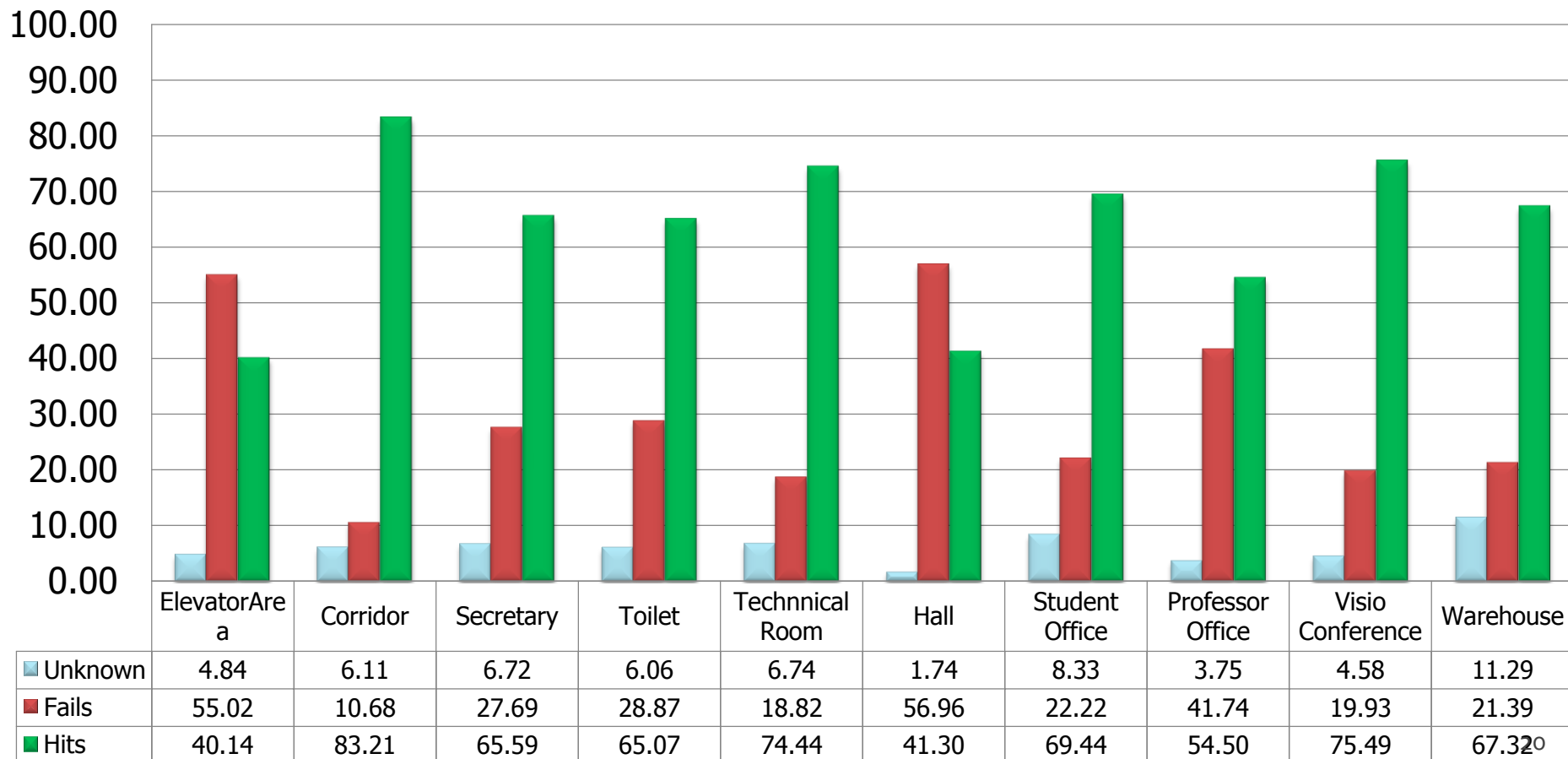
- Ranking of the best run submitted by group

Rank	Group Name	Score Class	Score Object	Score Total
1	MIAR ICT	3168.5	2865.0	6033.5
2	NUDT	3002.0	2720.5	5722.5
3	SIMD*	1988.0	3016.75	5004.75
4	REGIM	2223.5	2414.75	4368.25
5	MICA	2063.0	2416.875	4479.875
6	GRAM	-487.0	0.0	-487.0

- Winner: MIAR ICT group

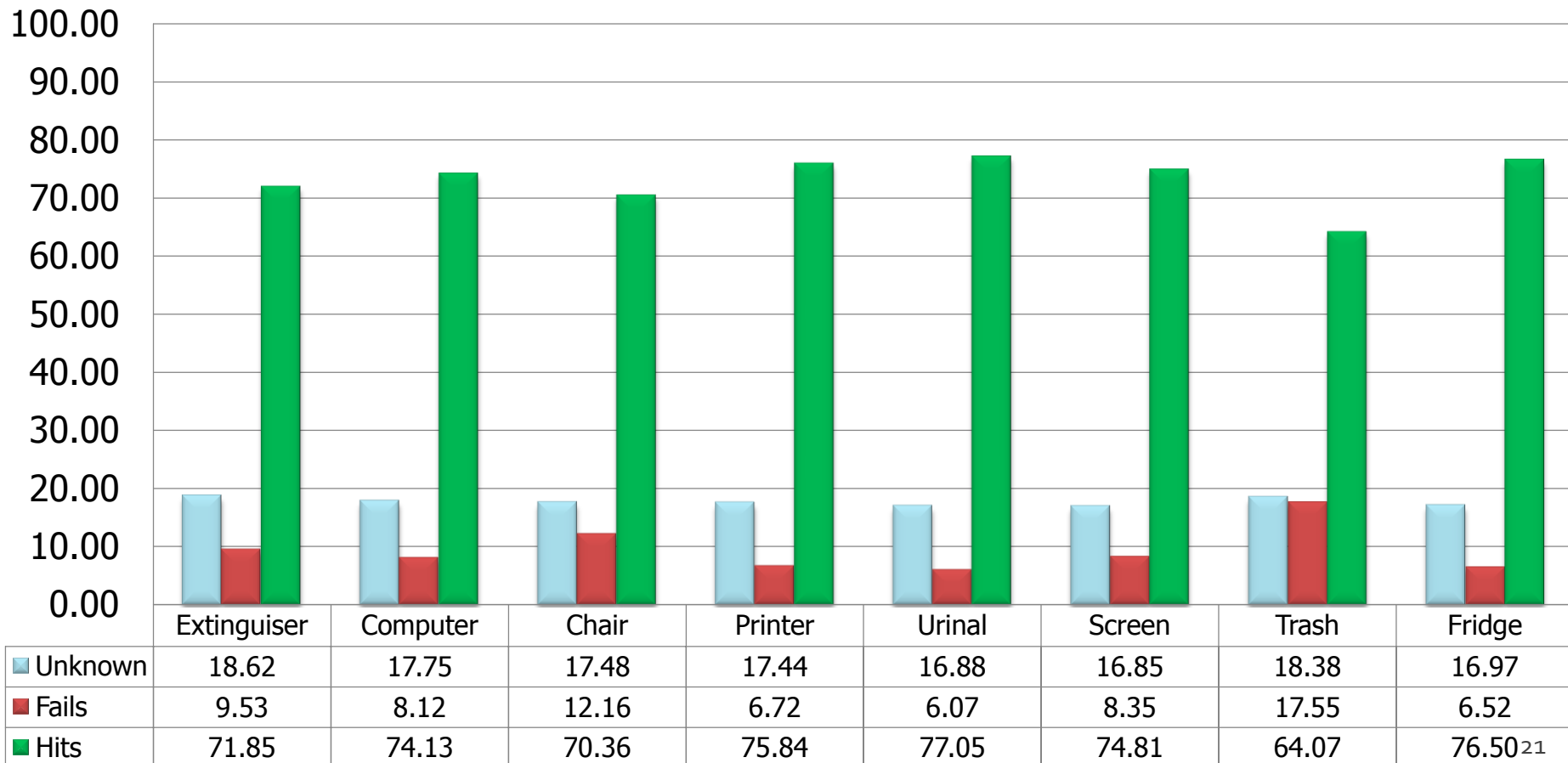
Robot Vision – Detailed Results

■ % of hits, fails and unknowns for Rooms (avg)



Robot Vision – Detailed Results

■ % of hits, fails and unknown for Objects (avg)



Robot Vision –Results Remarks

- Higher differences for room classification than for object detection
 - Percentage of hits in room classification
 - From 40.14% to 83.21%
 - Percentage of hits in object detection
 - From 64.07 to 77.05%

Robot Vision –Results Remarks

- Higher differences for room classification than for object detection
- Explanation A
 - Some room categories are still challenging
 - The number of rooms in the database is unbalanced
 - Several room categories can be confused
 - E.g. Hall and Corridor

Robot Vision –Results Remarks

- Higher differences for room classification than for object detection

- Explanation B

- The appearance ratio for objects is very low (<30%)
 - Negative detections are considered as hits
- A simple Zero-R classifier would obtain high scores
 - For all frames: there are no objects in the scene
 - NUDT used it and ranked 3rd in object classification

Robot Vision – Proposals

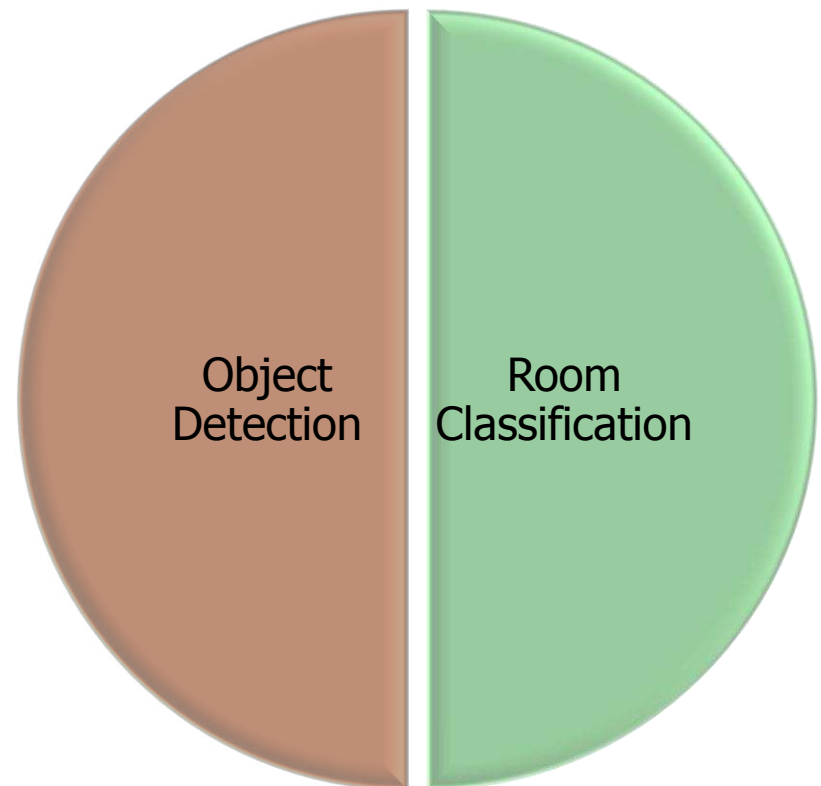
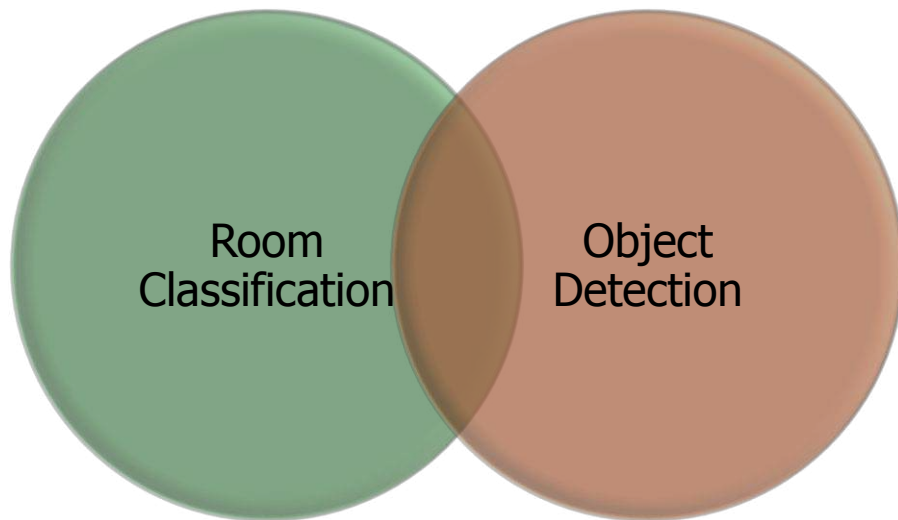
- MIAR-ICT
 - Ranked 1st
 - Visual and Depth images
 - Kernel Descriptors + Efficient Match Kernels
 - Features generation
 - Image feature size = 42000
 - Linear SVN
 - Classification

Robot Vision – Proposals

- REGIM
 - Ranked 4th
 - Visual images
 - PHOW
 - Feature generation
 - Pegasos - SVM
 - Classification
 - Post-processing
 - Threshold selection

Robot Vision – Proposals

■ Expectations versus Reality



Robot Vision – Conclusions

- Less attention than expected
 - New groups should be encouraged to participate
- The 2 problems were managed separately
 - No high level layers are used

Robot Vision – Future work

- Hard environment changes
 - Detect objects in an environment when trained the algorithm in a different environment
 - Export the model
 - An extinguisher will be an extinguisher in every environment it appears
- Adjust the performance evaluation formula
 - How to evaluate object detections should be carefully studied

RobotVision@ImageCLEF 2013 Task overview

Thanks for your attention Questions?

Organizers: J. Martínez Gómez, I. García-Varea, M. Cazorla and B. Caputo

