RobotVision@ImageCLEF 2014 Task overview

Organizers: J. Martínez Gómez, I. García-Varea, M. Cazorla and V. Morell

CLEF Sheffield 2014

Robot Vision - Motivation

- The behaviour of a robot r at timestamp t depends on:
  - Not the topological <x,y,z> location

    … But
  - the semantic category of the place

    &
  - the objects that are suitable for manipulation

Robot Vision - Motivation

Semantic Localization

Topological Localization

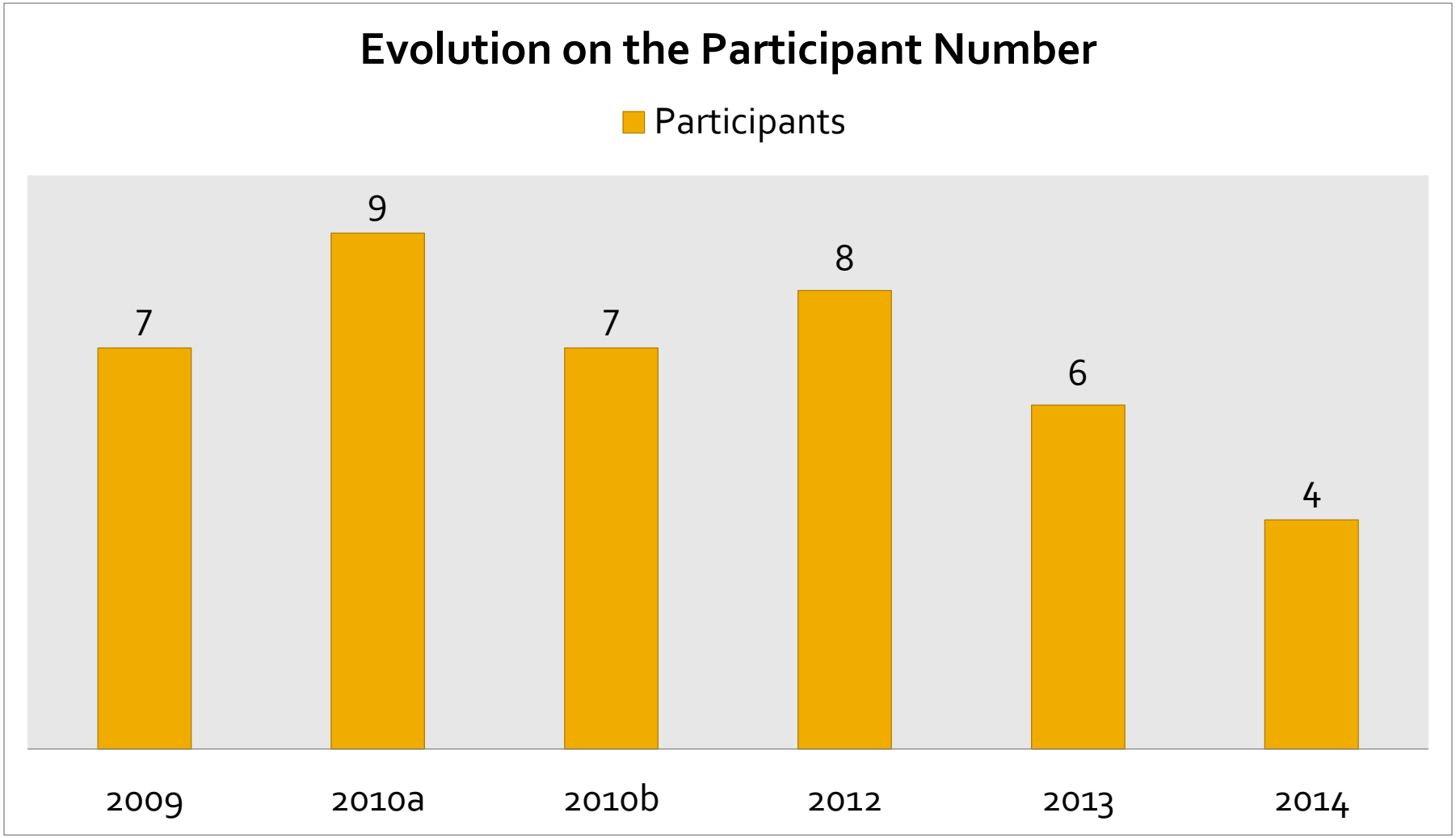
I am at the kitcken and I can see a Fridge and a table

I am placed at position <1.52,2.57, 3.51>

Robot Vision - The Task

- Proposed in 2009 (6th edition)
  - ImageCLEF 2009 – Greece
  - ImageCLEF@ICPR 2010 – Turkey
  - ImageCLEF 2010 – Italy
  - ImageCLEF 2012 – Italy
  - ImageCLEF 2013 – Spain
  - ImageCLEF 2014 – England
- Considerable attention
  - ImageCLEF 2009 – 7 groups
  - ImageCLEF@ICPR 2010 – 9 groups
  - ImageCLEF 2010 – 7 groups
  - ImageCLEF 2012 – 8 groups
  - ImageCLEF 2013 – 6 groups
  - ImageCLEF 2014 – 4 groups
- Organizers
  - Jesus Martínez Gómez, Ismael García Varea, Miguel Cazorla and Vicente Morell

Robot Vision - The Task

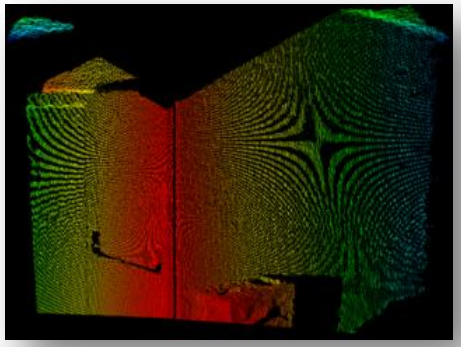


Robot Vision - Organizers

Organizer	2009	2010	2010	2012	2013	2014
B. Caputo	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	X
I. García-Varea	-	-	-	X	X	X
M. Cazorla	-	-	-	-	X	X
V. Morell	-	-	-	-	-	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
  - Different buildings but with similar structure

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, Phone, Chair, Printer, Urinal, Bookshelf, Trash and Fridge

Robot Vision – The Task

- Performance evaluation
  - 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 detected (TP) → +1.0 points
    - Each object incorrectly detected (FP) → -0.25 points
    - Each object detected as not present (TN)→ +0.0 points
    - Each object incorrectly detected as not present (FN)→ -0.25 points

Robot Vision – The Task

Robot Vision – The Task

Robot Vision – Major Changes

Topics	2009	2010	2010	2012	2013	2014
Place Classification	X	X	X	X	X	X
2 subtasks	X	X	X	X	-	-
Stereo Images	-	X	-	-	-	-
Unknown Labels	-	X	X	-	-	-
Object Detection	-	-	-	-	X	X
3D Images	-	-	-	X	X	X
New Buildings (test)	-	-	-	-	-	X

Robot Vision – Performance Eval

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

Robot Vision – Performance Eval

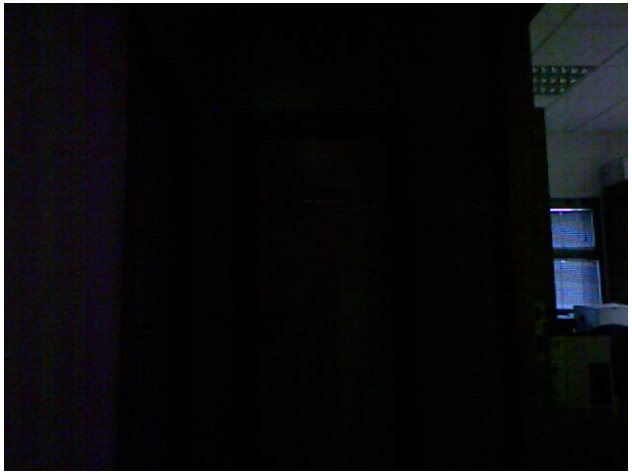
Class	Exting.	Phone	Chair	Printer	Urinal	Books	Trash	Fridge
Real labels								
Technical Room	NO	YES	NO	YES	NO	NO	NO	NO
Participant Decision								
Technical Room	*NO*	*YES*	*NO*	*NO*	*NO*	*NO*	*YES*	*NO*
Right/Wrong Classification								
Hit	TN	TP	TN	FN	TN	TN	FP	TN
Points								
+1.0	0.0	+1.0	0.0	-0.25	0.0	0.0	-0.25	0.0
Total: 1.0 + 1.0 -0.25 – 0.25 = 1.5								

Robot Vision – The Data

- Unreleased frames from the ViDRILO dataset (to be published)
  - 3 sequences
    - Training: 5000 frames
    - Validation: 1500 frames
    - Test: 3000 frames

Robot Vision – The Data

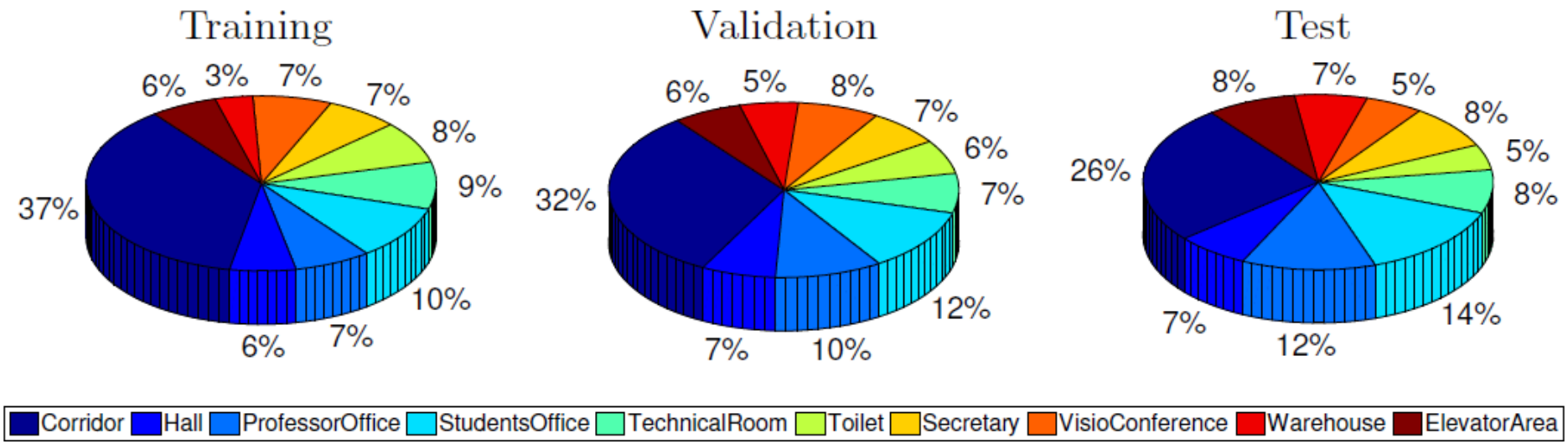
- Unreleased frames from the ViDRILO dataset (to be published)
  - Extreme lighting conditions in the test sequence
    - Range images are highly recommended



Robot Vision – The Data

	Number of Frames		
Room Category	Training (Building A)	Validation (Build. A+B)	Test (Build. B)
Corridor	1833	479	772
Hall	306	103	202
Professor Office	355	149	372
Student Office	498	174	419
Technical Room	437	110	242
Toilet	389	094	141
Secretary	336	102	245
Visioconference	364	113	159
Warehouse	174	081	201
Elevator Area	308	095	247
All	**5000**	**1500**	**3000**

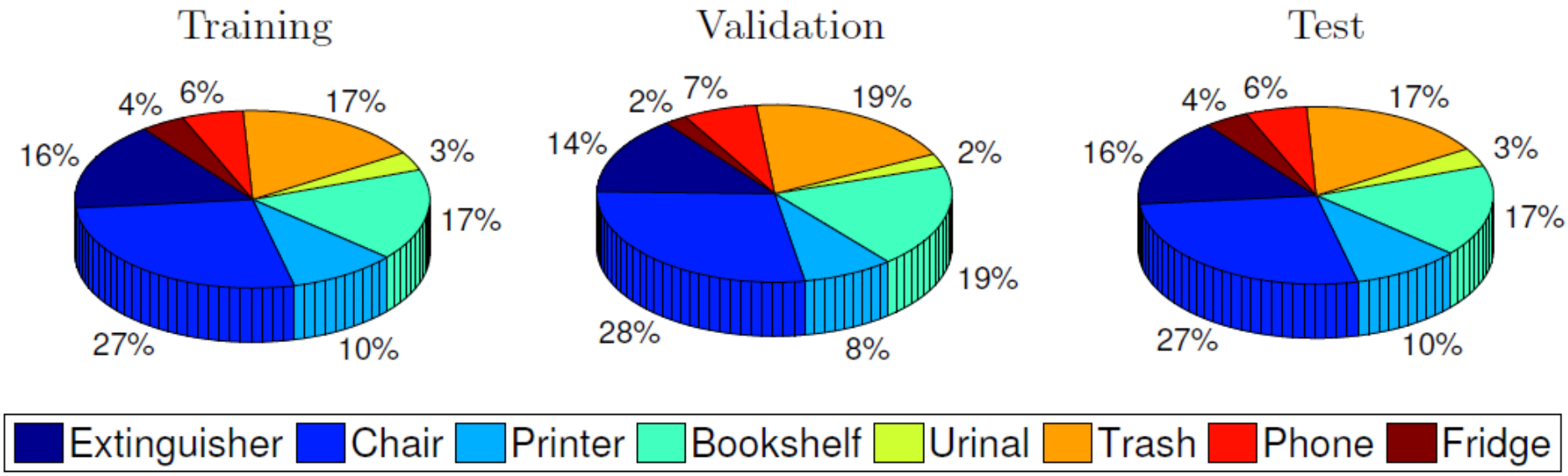
Robot Vision – The Data



Robot Vision – The Data

	Number of Presences / Lacks		
Objects	Training	Validation	Test
Extinguisher	770 / 4230	238 / 1262	566 / 2434
Chair	1304 / 3696	471 / 1029	1070 / 1930
Printer	473 / 4527	139 / 1361	265 / 2735
Bookshelf	802 / 4198	317 / 1183	896 / 2104
Urinal	162 / 4838	040 / 1460	060 / 2940
Trash	813 / 4187	323 / 1177	797 / 2203
Phone	267 / 4733	113 / 1387	303 / 2697
Fridge	190 / 4810	034 / 1466	047 / 2953
All	**4781 / 35219**	**1675 / 10325**	**4004 / 19996**

Robot Vision – The Data



Robot Vision – Participation

- 28 participants but only 4 submitted one run
  - NUDT: National University of Defense Technology, Changsha, China
  - UFMS CPPP: Federal University of Mato Grosso do Sul, Ponta Pora, Brazil
  - AEGEAN: University of the Aegean Karlovassi, Greece
  - SIMD: University of Castilla-La Mancha, Spain
    - Out of competition contribution using proposed techniques

Robot Vision – Script

- The sequences of the challenge are released with a useful MATLAB script that can be used as template for participant proposals, including:
  - Features extraction
  - Descriptor generation
  - Training
  - Classification
  - Performance evaluation

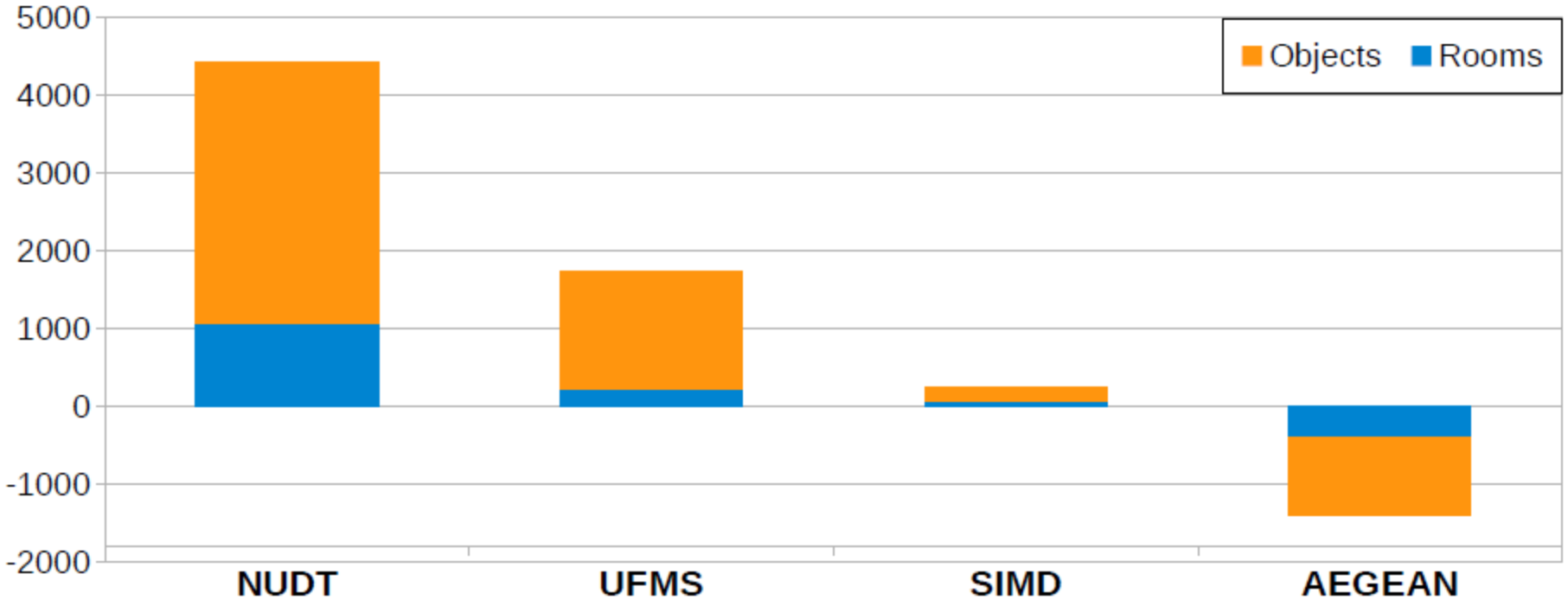
Robot Vision – Overall Results

- Ranking of the best run submitted by group

Rank	Group Name	Score Class	Score Object	Score Total (% Max)
1	NUDT	1075.5	3354.75	4430.25 (63.25)
2	UFMS	219.0	1519.75	1738.75 (24.83)
3	SIMD*	67.5	186.25	253.75 (3.62)
4	AEGEAN	-405	-995.00	-1400 (<0)

- Winner: NUDT group

Robot Vision – Detailed Results



Robot Vision –Results Remarks

- Room classification remains as an open problem when generalization is requested
  - Current approaches
    - Perform well when the test environment has been previously image (past editions)
- Object recognition is not affected by environment changes

Robot Vision – Proposals

- MIAR-ICT
  - Ranked 1st
  - Visual images
  - Bag of Words → SIFT and PHOG
  - Multiclass SVN (one versus all)

Robot Vision – Proposals

- UFMS
  - Ranked 2nd
  - Visual images
  - Bag of Words → dense SIFT descriptors
  - K-nearest neighbors
    - Classification

Robot Vision – Conclusions

- Again … less attention than expected
- None of the groups used 3D images
- Problems with room classification
- Room classification and object recognition problems were managed separately

RobotVision@ImageCLEF 2014 Task overview

Organizers: J. Martínez Gómez, I. García-Varea, M. Cazorla and V. Morell

CLEF Sheffield 2014