



Graz University of Technology (Austria)

I am in charge of the Vision Based Measurement Group at Graz University of Technology. The research group is focused on two main areas: Object Category Recognition & Detection; and Structure & Motion Analysis. We have excellent research results in vision-based categorization and localization of objects in still images, and we have tried to extend this to «active object categorization». Taking controlled, additional views of an object may help to disambiguate between various categories, for instance: «Is this a cow or a horse?», «Maybe it helps to take a closer look of its head...», etc. Besides developing novel representations and inference mechanisms for active categorization, we explore them under the “active, qualitative, purposive vision” paradigm, on the humanoid robot NAO.



The benefit of active, multiple-view-based techniques has been demonstrated for the recognition of specific objects from an object database a decade ago. But this research is the first work about:

**a. Active object categorization.**

In a first step, salient features of a particular class of objects are learnt (in this case toy horses, cows, card, and soccer players). Next, a previously unseen object is handed to Nao, and the task is to classify the according object category. If one view of the object is not sufficient, the robot autonomously plans and takes more views to disambiguate between competing categories.

**“ Hand-held object manipulation by a humanoid robot is particularly challenging, because we cannot expect precise, reproducible hand trajectories, or perfect pose for a grasped object.”**

**b. Active categorization on a humanoid robot.**

Hand-held object manipulation by a humanoid robot is particularly challenging, because we cannot expect precise, reproducible hand trajectories, or perfect pose for a grasped object. For the challenging task of object category recognition by a humanoid, it is evident that we need algorithms that can tolerate imprecise manipulations as well as significant ambiguity due to intra-class variability and inter-class.

We present a novel solution on the NAO robot, extending known concepts of active recognition and view planning of specific objects to the much harder task of active object category recognition. The entire algorithm is implemented on the slim computational platform of the NAO robot, and it can tolerate the robot's limited arm motion and grasping capabilities. An active approach for foreground feature separation makes the system robust to background clutter.



Nao's view (homogeneous background)



Original



Pose 1



Pose 2



Pose 3



Pose 4



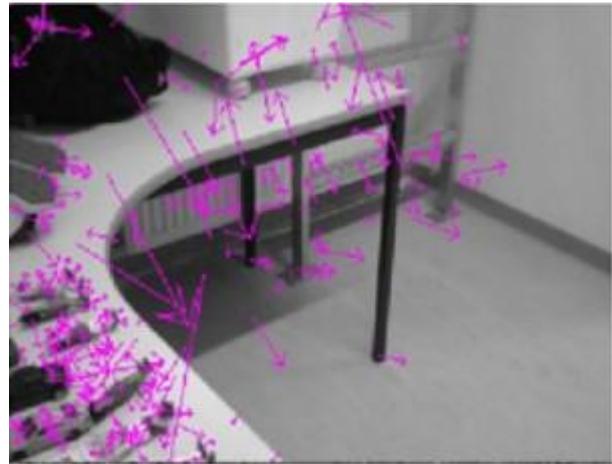
Pose 5

Finally, we extend this scheme towards active categorization in the case of unknown pose of the grasped object.

**Background elimination:**

NAO inspects the object (original image) and calculates SIFT descriptors (magenta arrows).

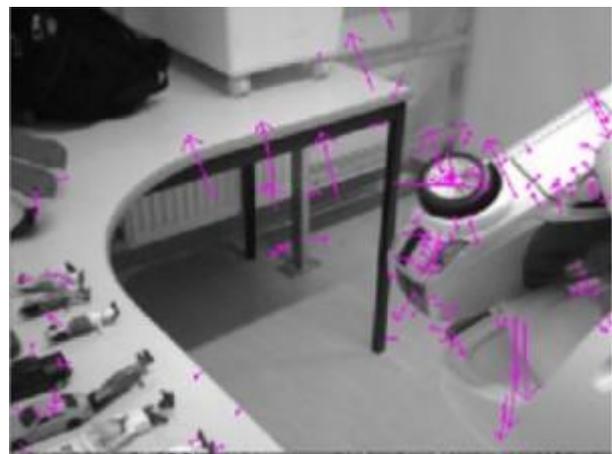
Next, NAO moves the object out of view (background image) and again calculates SIFT descriptors.



Background

Finally (result), those SIFTs that are similar in both images are removed.

NAO is ideally suited for the research performed in our group. Work with humanoids is attractive to students, and we offer an excellent working environment in our image-based measurement lab. The NAO hardware platform is sufficiently stable, and it can be rather easily accessed and programmed. At the same time, compared to industrial robots, NAO is inherently imprecise.



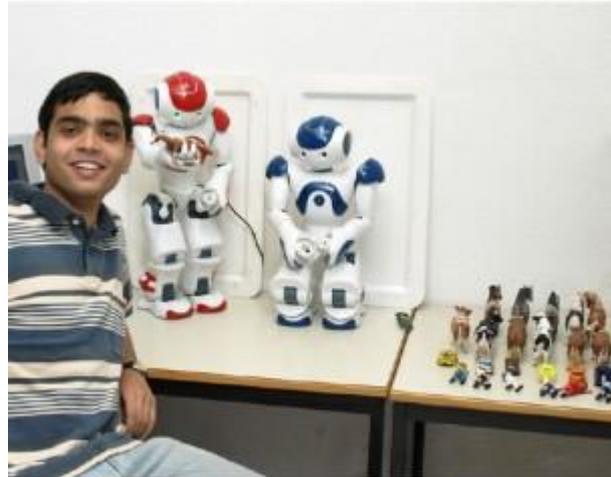
Result

This calls for qualitative, goal-oriented solutions. Furthermore, such a robotic system will force the student to use sensorial input (in our case mostly visual input), and to find ways to communicate processing results. While in Computer Vision and Image Understanding research, the output is often in the form of processed/ augmented images, NAO is well suited to demonstrate success of computer vision algorithms by actions, motion, or by voice output.

**“ NAO is ideally suited for the research performed in our group. Work with humanoids is attractive to students, and we offer an excellent working environment in our image-based measurement lab. ”**

While in Computer Vision and Image Understanding research, the output is often in the form of processed/augmented images, NAO is well suited to demonstrate success of computer vision algorithms by actions, motion, or by voice output.

I am working on NAO with groups of my students (up to 6 students per semester). Inside our vision-based measurement lab, we have set up a "living room" for NAO. This "room" is about 2x2 meters, with white walls of 60cm height, so that NAO's visual input will be constrained to the interior of its room, while it is easy to watch from outside.



Vignesh Ramanathan, our NAOs, and the four object categories

We have equipped the room with various color-coded furniture that can be used as chairs and tables of variable height, and also (turned upside-down) as color-coded baskets. Current projects include the autonomous vision-based collection of small, colored foam cubes from one of the tables, and from the floor. We have also built a flexible maze that can be put into the room in various configurations and we wish to make NAO find his way through the maze.

Nao is mainly programmed using C++ and OpenCV. Behaviours are naturally and easily implemented in Choregraphe.

The novelty of our approach is twofold: First, we demonstrate the viability of active recognition of categories by generalizing previous concepts on active recognition of specific, individual objects. Second, we use the particular abilities of NAO's humanoid arm to efficiently eliminate background features. Compared to categorization from a single view, our experimental results clearly demonstrate a significant gain in correctly categorized test objects by this "active" approach.

**“There certainly is an extremely high potential for this kind of hand-held, active inspection of objects by a humanoid robot, including human-robot interaction, home and service robotics, edutainment, active inspection, and active surveillance.”**

Furthermore, view planning can reduce the number of active steps needed to produce the final categorization result. Finally, the overall computational complexity of object categorization is significantly reduced by the integration of various views so that we can hope to see various applications on rather slim computing platforms like NAO in the near future. There certainly is an extremely high potential for this kind of hand-held, active inspection of objects by a humanoid robot, including human-robot interaction, home and service robotics, edutainment, active inspection, and active surveillance.

Improvements in autonomous grasping capabilities will be the key to further progress in this area.

For further information about Axel Pinz' research group and his research activities, please visit: [www.emt.tugraz.at](http://www.emt.tugraz.at)

Published papers:

V. Ramanathan and A. Pinz.

Active object categorization on a humanoid robot.

In Proc. VISAPP'2011.