COGNITIVE VISION SYSTEM FOR AN ECOLOGICAL MOBILE ROBOT

Ana Pavel, Cristian Vasile, Cătălin Buiu

Faculty of Automatic Control and Computers
"POLITEHNICA" University of Bucharest
Spl. Independentei 313
Bucharest, Romania
email: cbuiu@ics.pub.ro

Abstract: Ecology is a challenging application area for mobile service robots. They should be able to automate tasks that are too tedious or dangerous for humans to execute, such as collecting waste material. Such a robot must make use of several senses, of which the most important and difficult to implement is vision. This paper presents the cognitive vision system of ReMaster One, an autonomous service robot that is able to recognize and sort waste in an indoor environment. A first prototype has been built and tested with success.

Keywords: service robot, ecology, cognitive vision

1. INTRODUCTION

Over the last 15 years there has been significant progress in the fields of artificial intelligence, computer perception, machine learning and robotics. Yet there has been only minor progress on truly cognitive systems.

There are numerous definitions for a "cognitive system". A recent debate on this topic has attracted more than 40 different answers to "What is a cognitive system?" question (http://www.eucognition.org/wiki/).

For example, "cognition is a process of search for an appropriate action by an intelligent agent. An (artificial) cognitive system is one that uses intelligent control, generally modelled on high-level biological intelligent systems; common features are memory, learning, and a capacity for planning" (Joanna Bryson). Or, "cognition is the ability to plan, reason, adapt and act according to high level motivations or goals and using a range of senses, typically including vision, and may be communicate" (Patrick Courtney).

Cognition can also be interpreted as "generation of knowledge on the basis of perception, reasoning, learning and prior models" (Christensen, 2003, pp. 17-18).

The same Wiki resource lists more than 40 application areas for cognitive systems, ranging from autonomous artificial cells to robot guide dogs.

A preferred application area for artificial cognitive systems is robotics. Robots are situated complex systems which are capable to perceive the environment using a wide range of artificial senses and to act upon it after reasoning about their goals, environment, other robots and people.

Ecological applications of robotics are generated by a growing interest in keeping the nature clean and the desire to replace humans in tedious, boring, and often dangerous tasks, such as collecting waste. An ecological service robot should be able to search, recognize, pick up and store a significant number of wastes in an indoor or outdoor environment.

This paper presents the cognitive vision system of ReMaster One, a prototype of a mobile service robot that is able to recognize, collect and sort empty plastic bottles, empty cans and batteries. ReMaster One has been developed at the Autonomous Robotics Lab of the POLITEHNICA University from Bucharest. Section 2 details the problem of cognitive vision and its relevance to mobile robotics field. Section 3 will briefly present the robot and its task. Section 4 gives more details about the cognitive vision system, algorithms and performances. Section 5 concludes the papers and presents directions for further improvements of the robot and of the cognitive vision system.

2. COGNITIVE SYSTEMS AND VISION

There is a high interest in cognitive systems, both from a theoretical and practical point of view. Numerous related research projects have ambitious goals.

COSY (www.cognitivesystems.org) is a cognitive science European project which aims (2004-2008) to advance the science of cognitive systems through a multi-disciplinary investigation of requirements, design options and trade-offs for human-like, autonomous, integrated, physical (e.g., robot) systems, including requirements for architectures, for forms of representation, for perceptual mechanisms, for learning, planning, reasoning and motivation, for action and communication.

The euCognition European project (www.eucognition.org) gathers people doing research in the many disciplines that address the issues of creating artificial cognitive systems including (but not limited to):

- Neuroscience
- Psychology
- Cognitive science
- Machine Learning
- Autonomous systems theory
- Cognitive robotics
- Mathematical modeling
- Cognitive Vision

Cognitive vision is a very important and delicate sub-field of cognitive systems research. A "cognitive vision system" is defined as a system that uses visual information to achieve (Cohn, *et al.*, 2003):

- Recognition and categorization of objects, structures and events
- Learning and adaptation
- Memory and representation of knowledge
- Control and attention

Related cognitive vision concepts are fundamental questions addressed in the EU project “Cognitive Vision Systems” (www.ecvision.org) in which issues related to categorization, recognition, learning, interpretation and integration were addressed in relation to vision systems for intelligent embodied systems (such as robots).

A Joint Research Project funded by the Austrian Science Fund researches cognitive vision models and techniques to enable cognitive vision abilities for “Personal Assistance” scenarios. Such a scenario is characterized by a combination of mobile devices and distributed ambient spaces which unobtrusively support users by being aware of the present situation and by responding to user requests.

A detailed survey on object categorization can be found in (Pinz, 2006). This article presents foundations, original research and trends in the field of object categorization by computer vision methods. The research goals in object categorization are to detect objects in images and to determine the object’s categories. Categorization aims for the recognition of generic classes of objects, and thus has also been termed ‘generic object recognition’. This is in contrast to the recognition of specific, individual objects. While humans are usually better in generic than in specific recognition, categorization is much harder to achieve for today’s computer architectures and algorithms (Pinz, 2006).

3. REMASTER – MOBILE SERVICE ROBOT FOR ECOLOGICAL APPLICATIONS

Mobile service robots are being developed in numerous research labs and companies worldwide. A detailed survey is presented in (Buiu, *et al.*, 2007).

ReMaster One (Fig. 1) is a prototype of a service robot for ecological applications. The robot is a first result of the ReMaster project (remaster.pub.ro) currently under development at the Laboratory of Autonomous Robotics (AUR) of the POLITEHNICA University of Bucharest. This project aims to develop an intelligent autonomous and social service robot to be used for collecting and sorting waste material.

The starting idea was given by the Eurobot 2007 contest (www.eurobot.org), whose task was to build a robot that is able to search, recognize, collect and sort waste (bottles, cans and batteries) that are disposed on a table like that one in Fig. 2.

The artificial vision system of the robot was developed to recognize two generic types of objects: objects with a label of a different color than the object itself (e.g. bottles) and uniformly monocoloured objects (e.g. cans, batteries).

Fig. 1. ReMaster One.

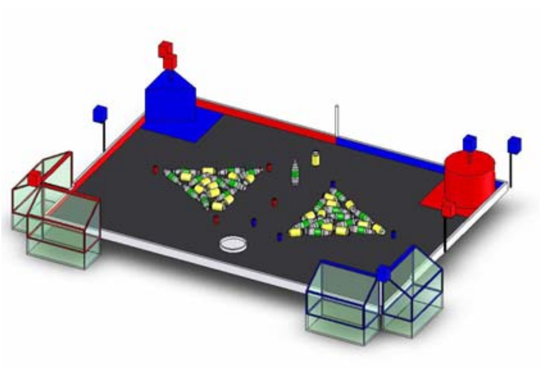

Fig.2. Eurobot 2007 playground.

The task required the detection of 0.5 liters bottles with a green label, 0.33 liters yellow colored cans and blue and red colored D batteries on an anthracite grey table, but the vision system can be easily adapted for other applications.

After recognizing and collecting the object, the robot must transport it to a collecting bin at the corner of the table. There are dedicated bins for cans and bottles, and an additional basket for batteries which is located in the middle of the table.

More details about the robot can be found in (Buiu, *et al.*, 2007). This paper will continue by presenting the cognitive vision abilities of the robot.

4. COGNITIVE VISION SYSTEM

CMUcam2+ sensor is a low cost artificial vision system (Rowe, *et al.*, 2002) which was developed at Carnegie Mellon University, and is very adequate for robotics projects. The camera has the advantages that it is a low cost vision system, it is capable of colour detection up to 16.7 fps and it has ready-on implemented basic processing operations.

Due to the CMUcam2+ restrictions, such as low transfer rate, low resolution (160x255) and the absence of a programming package for the image processing unit of the camera, the algorithms developed are based only on the camera’s basic operations, so that processing is done real-time.

4.1 Hardware

The CMUcam2+ (Fig. 3) consists of a SX52 microcontroller interfaced with a 0V6620 Omnivision CMOS camera that allows high-level data to be extracted from the camera's streaming video. Its primary function is to track and monitor highly contrasting color regions. It can also detect motion, provide color statistics, and transmit image information to a computer, via a TTL serial port, for additional processing. The CMUcam2+ supports 5 servos and 4 IO ports. The camera also has a sleep mode to conserve power.

Fig.3. CMUcam2+ camera.

4.2 Object detecting algorithms

In this section, the algorithms developed for solving the given task will be described. It will be shown how to use a camera to detect and grab objects when it is attached to a gripper, to distinguish objects by color and to detect movement.

A very difficult problem in designing the object detecting algorithms was the way the camera tracks color (Fig. 4). It returns the minimal rectangle area containing all the spots of the searched colour in the image, making it difficult to isolate one single spot.

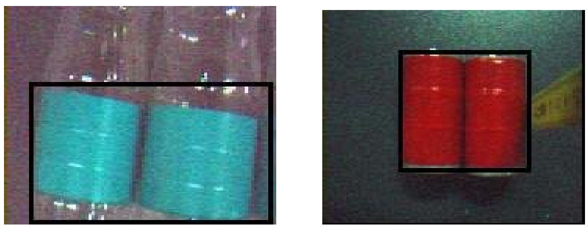

Fig.4. CMUcam2+ Color Tracking.

While solving the task described above, two generic types of objects were defined: labeled and monocoloured objects. The bottles belong to the first type, while the batteries and cans, which differ only by size, belong to the second one. It is easy to notice that distinguishing between bottles is easier than between cans or batteries. The difference of difficulty between the two cases arises when trying to distinguish a single object from a cluster of the identical objects (Fig. 5 and 6).

Fig.5. Cans cluster.

Fig.6. Bottles cluster.

A cluster of cans can form a monocoloured block which makes it difficult to select only one object to be picked up by a gripper. The worse situation for the bottles appears when their labels are aligned which is quite an improbable scenario.

The positioning algorithm for a gripper. The positioning algorithm for the gripper (Fig. 7) is a nine-step procedure which was developed for a gripper having at least three degrees of freedom (two joints for moving in the xOz plane and the possibility of moving the hand round its axe). The camera is attached to the center of the gripper's hand.

The algorithm is based on three basic operations of the camera: creating a virtual window with specified dimensions of the camera's view (VW - Virtual Window), framing a specific color in the virtual window (TC – Track Color) and determining the mean color of the virtual window (GM – Get Mean). Due to the color detection method (TC), this algorithm is very efficient for distinguishing bottles and almost precise for cans.

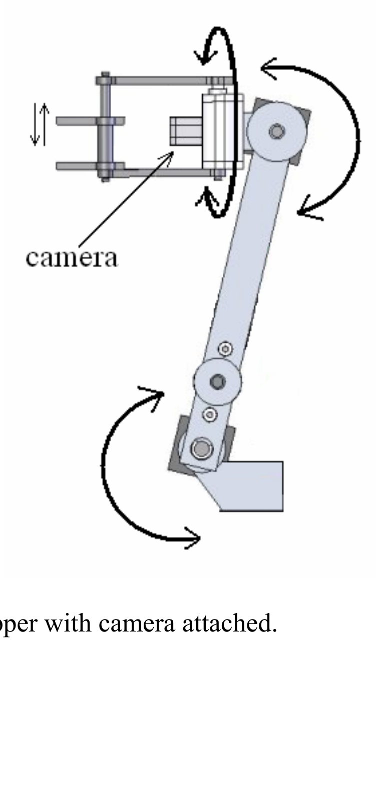

Fig.7. Gripper with camera attached.

The advantages of using this algorithm are: it uses only four movement commands of the gripper and/or robot, and it is capable of positioning the gripper even when there are several objects of the same color in the image.

The algorithm disadvantages are: it is not very easy to be implemented and it is dependent on a good color calibration.

The algorithm operates as follows:

0. The robot goes on the table looking for objects. The camera looks for color with the whole frame until it detects at least the number of pixels equal to the number of pixels of a complete single label in the frame (TC).
When the color is found:
1. Resize the frame to a square with the dimensions equal to the width of the label in pixels, fixed and centered.
2. Look for the color in the perimeter from step **0.** until the square from step **1.** contains color. Center the gripper on the square.
3. Go with the virtual window towards color until: the center of the square contains color, maximum color has been reached and there exist only two opposite corners without that color or none at all. Center the gripper on the square again.
4. Rotate the gripper until the square contains only the specified color.
5. Move the square up the frame to check the position of the bottle: vertical or horizontal.
6. If the square is not filled with the color, rotate the gripper 90 degrees and move the square to its initial position.
7. Go up and down with the square until it reaches the edge of the label and return half length of the label.
8. Give the command to the gripper to grab.

The algorithm that finds the color with the square from step **3.** is as follows:

Look at every corner. There will be one of the cases:
A: There is only one corner filled with the specified color → go towards that corner.
B: There are two corners filled with the specified color:
B.1: opposite:
B.1.1: if there is color in the center of the square → go towards the center of color.
B.1.2: if the center of color corresponds to the center of the square → stop.
B.1.3: if the center doesn't contain the color → go towards a corner with the color.
B.2: not opposite → go in the direction of the two squares with color.
C: There are three corners filled with the specified color → go towards the opposite corner of the unfilled corner.
D: There are four corners filled with the specified color:
D.1: if the square is filled → stop.
D.2: if the square is not filled → go towards the center of color.
D.3: else → go in an arbitrary direction.

Observation: Lengths in pixels correspond to the real ones (mm) in an inverse proportional way to the height the camera's lens position, h.

$$lPix = \frac{lReal \times c}{h} \tag{1}$$

where c is a constant of conversion specific to the camera (350 pixels for CMUcam2+).

The positioning algorithm for a gripper when a single object is in the camera's frame. The algorithm can be used for only one object situated entirely in the frame. Based on a mathematical formula, the algorithm is capable of positioning the gripper on any monocoloured rectangular shaped object with previously known dimensions (length and width). For the bottles it is considered the rectangular shape to be the bottle's label.

The algorithm operates as follows:

1. Center the gripper on the object, using its center of color.
2. Resize the virtual window to fit exactly the whole object.
3. Determine the length of the virtual window (on the Ox axe of the frame).
4. Calculate the angle:

$$\alpha = \begin{cases} \arccos\frac{L}{\sqrt{L^2+l^2}} - \arccos\frac{D}{\sqrt{L^2+l^2}}, D \in \left[L \to \sqrt{L^2+l^2}\right] \\ \arccos\frac{L}{\sqrt{L^2+l^2}} + \arccos\frac{D}{\sqrt{L^2+l^2}}, D \in \left[\sqrt{L^2+l^2} \to l\right] \end{cases} \tag{2}$$

where:
D is the length of the virtual window,
L is the length of the rectangular object,
l is the width of the rectangular object.
5. Divide the virtual window in four equal rectangles and determine which bottom rectangle contains least of the object's color.
6. Rotate the gripper with α degrees towards the rectangle found at step **5.** and then grab the object (Fig. 8).

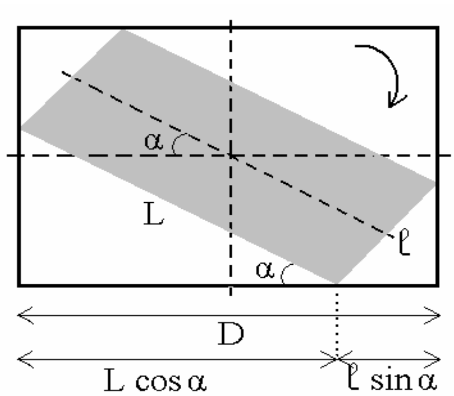

Fig.8. Gripper positioning to grab one object.

Observation: In Fig. 8 the grey rectangle represents the object to be picked up by the gripper, while the circumscribed rectangle is the camera's virtual window.
The gripper will be rotated depending on its initial position and type.

Deduction of the rotation angle formula:

$$L\cos\alpha + l\sin\alpha = D \Rightarrow$$
$$\frac{L}{\sqrt{L^2+l^2}}\cos\alpha + \frac{l}{\sqrt{L^2+l^2}}\sin\alpha = \frac{D}{\sqrt{L^2+l^2}} \Rightarrow$$
$$\cos\varphi\cos\alpha + \sin\varphi\sin\alpha = \frac{D}{\sqrt{L^2+l^2}},$$
$$\text{where } \varphi = \arccos\frac{L}{\sqrt{L^2+l^2}}, \varphi \in \left[0, \frac{\pi}{2}\right]$$
$$\cos(\varphi - \alpha) = \frac{D}{\sqrt{L^2+l^2}} \Rightarrow$$
$$\alpha = \varphi \pm \arccos\frac{D}{\sqrt{L^2+l^2}} + 2k\pi, k \in Z$$

Calibration and object detecting algorithms by color. Object identification can be performed in a light controlled environment, in this case, for example, inside the collecting system of the robot.

Because the CMOS sensor is very sensitive to the ambient light and it can be very easily influenced by the type of light (incandescent or fluorescent), a vision system that works inside the robot and that relies only on colour identification has been developed. White LEDs were used to control the intensity and type of illumination.

In order to detect colour it is necessary to implement colour calibration procedures, to ensure the best results of the colour tracking operation and the algorithms.

The calibration procedure consists of saving the colors that the camera has to detect. More precisely, the colors to be saved are put in the front of camera's lens and then the camera determines the mean value of each RGB channel (a value between 16 and 240) of the image and stores the limits of each channel. By experiment, it was determined that the best results are achieved when taking the limits:

RedMeanValue ± 20,
GreenMeanValue ± 20,
BlueMeanValue ± 20.

Taking the limits: MeanValue ± Δx (absolute maximum deviation of the values from each RGB channel of the image) gives poorer results. If the limits obtained are below 16 or greater than 240, they are taken 16, respectively 240.

The method used for detecting objects by their color is the following:

1. Create a procedure which returns the number of pixels which are between some given RGB limits.
2. Determine by experiment the minimum number of pixels of color needed for ensuring the existence of each object. For this application, these limits are: 600 pixels for the cans and bottles and 300 pixels for the batteries.
3. To detect the objects test if the number of pixels for each color previously stored in the calibration procedure is enough to identify the object.

Using this method, yellow cans, green labeled bottles, red and blue D batteries were successfully distinguished and identified because the limits of the colors of the objects were almost disjoint.

Another method with poorer practical results is to distinguish between objects by comparing the mean color (obtained by taking the mean values of the RGB channels from the image) to the colors stored in the calibration procedure. The closest color selected (the minimum sum of the differences between the mean value saved and the mean value detected for each RGB channel) gives the type of the object.

Frame differencing and detecting motion. For service robots detecting moving objects is very important in the identifying and collecting process (e.g. avoiding little animals or other robots). The task required to ignore moving objects. CMUcam2+ is capable of detecting motion by using frame differencing which is a method of identifying changes in a series of images. This capacity will be exploited in future versions of the robot and cognitive vision system.

5. CONCLUSIONS AND FUTURE IMPROVEMENTS

ReMaster One is a mobile service robot which is being developed for ecological applications. The robot uses artificial vision to recognize and categorize objects which will be then transported to recycling bins. Some cognitive vision algorithms for object recognition and categorization have been presented.

The plan is to increase ReMaster's performance by improving its artificial vision system. Future work will include changing the hardware platform, to have more facilities and greater performance. An interesting experiment will be to program a web cam, because of its high resolution and high transfer rate. An important feature is that it can process the image directly allowing to develop more performant algorithms and to optimize cognitive processes. Another option to upgrade the hardware platform is to use the CMUcam3 which offers similar facilities, but runs the basic algorithms on its own integrated microcontroller.

The software and the artificial vision algorithms will also be improved. New ideas, methods and procedures will be studied and experimented in order to help develop an advanced artificial cognitive system for future robots. The plan is to make them able to gather visual information and integrate it into complex knowledge. Their learning capacities will make them capable of planning efficient strategies for collecting waste and to create a database with known objects. Artificial learning behavior can be modeled with neural networks whose implementation and development are also part of future projects.

ACKNOWLEDGEMENTS

We gratefully acknowledge the support of Ericsson Romania who sponsored part of the development of the robot and of the Romanian Government through the Excellence Research program (Contract 83-CEEX-II-03/31.07.2006). The Robotics Masters team has contributed to the development of the robot (see the full list of team members on remaster.pub.ro).

REFERENCES

Buiu, C., F. Cazan, R. Ciurlea (2007). Developing of a Service Robot to Recognize and Sort Waste, In: *Proceedings of the 16th International Conference on Control Systems and Computer Science*, Bucharest, May 22-26, 2007.

Christensen, H (2003). Cognitive (vision) systems. ERCIM News, pages 17–18, No. 53, April 2003.

Cohn, A. G., D. Magee, A. Galata, D. Hogg, S. Hazarika (2003). Towards an architecture for cognitive vision using qualitative spatio-temporal representations and abduction, In: *Freksa, C, Brauer, W, Habel, C & Wender, K F (editors) Spatial Cognition III, Routes and Navigation, Human Memory and Learning, Spatial Representation and Spatial Learning,* pp. 232-248, Springer-Verlag, 2003.

Pinz, A. (2006). Foundations and Trends in Computer Graphics and Vision, Vol 1(4), pp. 255-353, 2006.

Rowe, A., C. Rosenberg, I. Nourbakhsh (2002). A Low Cost Embedded Colour Vision System, In: *Proc. of International Conference on Intelligent Robots and Systems (IROS),* 2002.