

Friendly Interface for Objects Selection in a Robotized Kitchen

A.Casals

Department of Automatic Control
and Computer Engineering
Polytechnical University of Catalonia
08028 Barcelona - Catalonia - Spain

X.Cufi, J.Freixenet, J.Martí, X.Muñoz

Computer Vision and Robotics Group
EIA - IliA
University of Girona
17071 Girona- Catalonia - Spain

Abstract

This paper presents an interface for the interaction between a human and an adapted kitchen where different elements, including a robot, have to be controlled. After the global structure of the interface is described, a more detailed explanation is presented on the way the interface offers the user an easy and friendly way to select the desired objects and to provide their position to the robot for their manipulation.

1 Introduction

Independent living constitutes an increasing need in our society today. Age and a great number of different disabilities prevent many people from living autonomously. For these reasons, for a long time, some research and development has been done to progress towards the possibility to provide disabled with devices or systems that allow increasing their autonomy. The kitchen is a part of the house where different technological aids can become essential for personal autonomy. There are studies about the main user's needs and ergonomic requirements in such environments [2, 6] useful in the design of such aids. Among many other projects on Robotics assistive technology, the CAPDI project [3] is centered uniquely to the assistance in the kitchen, focusing towards the main needs there. The project combines the use of a robotic arm, when needed, with other adapted elements to enable the user to gain some independence. Being the user's acceptance [1] one of the critical aspects to be considered, a friendly interface that facilitates the selection and location of the different objects in this environment has been developed. Based on the images obtained from different cameras located in the kitchen environment, a menu based interface and a computer vision system supports the user's selection

tasks. User's interaction with several controllable devices as well as with the robot can be hard and tedious. Further to the design of an easy to understand menu on the screen, the task to locate the objects to be manipulated by the robot is described. In this environment where there is a frequent human interaction the location of objects is uncertain and thus, the robot-arm can not be previously programmed.

This paper presents two different methods for selecting objects which have to be manipulated by a robot arm in a kitchen adapted to disabled. Both methods are based on normal human behaviors such as figure 1 depicts, and although they are presented as two different methods they are completely compatible and complementary. The first method consists on sequentially presenting the different objects of the scene and the user interacts with the system by a simple "yes" or "not" mechanism. This method is based on a bottom-up strategy in which the image is segmented using a non purposive algorithm that combines edge detection and region growing. The second method consists on the direct recognition of the object indicated by the user. This method is based on a previous learning phase, in which the system registers where the objects of interest are located and which are the most significant characteristics of every object.

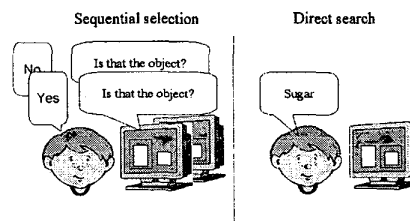


Figure 1: As a simple schema, the central idea of the two interfaces proposed for object selection

1.1 The CAPDI kitchen

The CAPDI kitchen (Adapted Kitchen for the Disabled) has been conceived as a modular system that can be adapted to users with many different degrees of disability. For the most common inability's of disabled and older to perform different tasks in the kitchen, some adapted elements, including a robot-arm, have been designed. With these basic elements, a modular kitchen can be configured according to different user's needs. The kitchen can be progressively adapted to the evolving user's needs, due to age or to a degenerative disability.

One of the most important problems in a kitchen is the accessibility to the different items. For this reason the kitchen design considers the need of designing different adapted elements. The first kind of adapted element consists of lifting wall-cupboards, to approach the desired objects to the user. Additionally, a specific cupboard carousel, similar to those used in industry, has been designed to approach the desired shelf to the user. Each shelf moves to a height accessible either for standing persons or for wheelchair ridden users. This design is to be extended to the refrigerator in the future. For more severely disabled people a robot arm is also included in the adapted kitchen. The objective of adding this new and more complex element is to enable such users to manipulate objects and to perform simple tasks that allow them to gain some autonomy. For instance, someone could bring some cooked food in the morning and place it in the refrigerator. At lunch time the user could, with the support of the robot-arm, take the food from the refrigerator, put it in the microwaves oven and carry it to the desired position. This support is not intended to provide complete autonomy, but to extend the user's autonomy. Figure 2 shows the laboratory kitchen prototype with a robot and a carousel cupboard, while figure 3 shows the kitchen interface for local area visualization and control of its elements.

To facilitate the user interaction with the robot, a vision system can be used to locate the desired objects and to guide the robot towards the target. A panoramic camera, (or more than one) can visualize the selected area (areas) where the robot has to move to pick up an object. A local camera, over the robot, can locate more precisely the object when the gripper is close and the arm occludes the panoramic view. In what follows, the structure of the user's interface as well as the method to efficiently select the objects to be manipulated by the robot arm is presented.

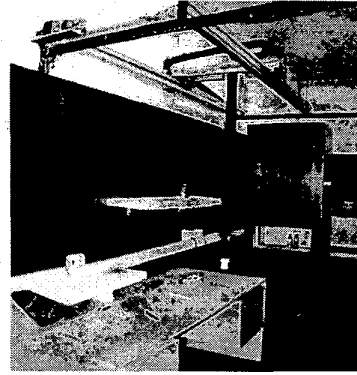


Figure 2: First prototype of the robotized kitchen in the lab environment

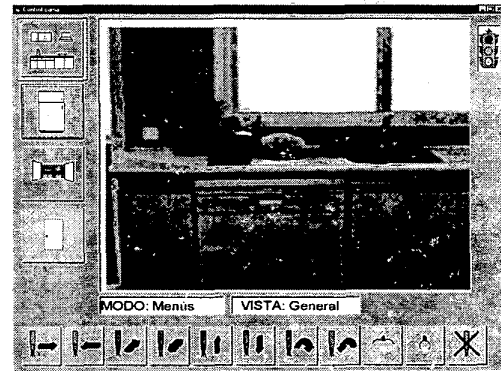


Figure 3: A view of the kitchen interface

2 Visual interface

The need for elder or disabled people to interact and control all the different kitchen elements has forced to design a friendly interface. On the one hand the user has to have available at every instant any information related to the status of each controllable element. On the other hand, he or she has to be able to control them. The interface requirements will depend on the number of controllable elements of every kitchen. For instance the cupboard carousel has its own specific keyboard since this element can be in some cases the only adapted element in the kitchen. Similarly, simple push buttons can be used to move up and down the lifting cupboards when there are few elements to control. Nevertheless, when more controllable elements are used (moving elements, the robot, timers, temperature regulators, etc), a unique interface will be the

best solution. The input device will be a set of keys, or a microphone to interact with the controllable elements, and the output is a screen with icons to select the desired options and an image that visualizes the selected kitchen area. Figure 3 shows the monitor with the different menu options. In the lateral vertical line of icons different kitchen elements or areas are indicated. Choosing one of them new options appear to control the selected element. Also, from the adequate cameras distributed around the kitchen area, the zone of interest can be visualized on the screen, for instance, inside the refrigerator. The bottom area of the screen contains the icons related to the different robot movements. From them, the robot can be positioned to any point having been explicitly programmed by the user to control the robotic arm. The vision system locates the objects in the scene by interacting with the user and permits to choose the one he desires. First, the user selects the area of interest of the kitchen by using the camera which offer a global view of the kitchen, afterwards within this area, the user selects the object of interest using a local camera located on the arm. These two steps can be achieved following different strategies, as depicted in figures 4 and 6. In the first option, named "sequential selection", the interface marks sequentially the items that could be of interest to the user. The user interacts with the system by a simple "yes" or "not" mechanism. In the second option, named "direct search", the user indicates which is the item of interest, so that the system search this item in some pre-specified location.

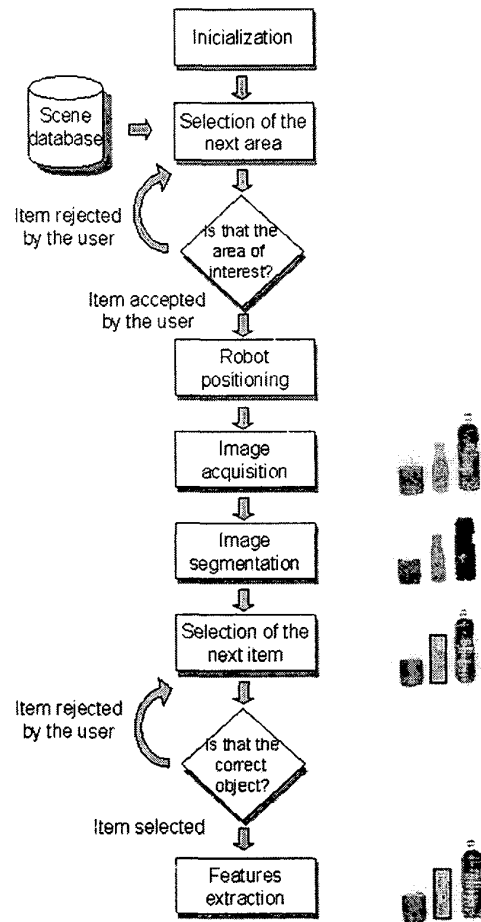


Figure 4: A scheme of the bottom-up approach for object selection

2.1 Bottom-up approach

Figure 4 shows the block diagram of the method which is based on a sequential selection of the area of interest and of the segmented imaged objects. The user selects one area of interest, the one where he supposes the object to be manipulated is placed. Once the area is selected, the robot arm is positioned in a pre-defined 3D position, from where the local camera attached to the arm can acquire the local image. The process of segmentation is the core of the bottom-up interface and consists on the partition of the imaged scene into meaningful regions. But, perfect segmentation is a difficult issue. In the following we present a new algorithm which achieves reliable results in reasonable time, as required by the application.

2.1.1 Image segmentation

Region-based techniques often fail to yield the desired structure due to the difficulty of choosing a reasonable starting "seed" point, and appropriate growing or stopping rules. Similarly, edge-based techniques often fail due to the similarity of objects within the scene or indistinct boundaries between image objects. Combining both edges and regions segmentation approaches lead to better results [5, 7]. The basic principles of the presented method are to detect contours first and then close them and obtain more precise segmentation by region growing techniques. The first step of the segmentation process is to extract the most significant contours of the image [4]. Then, the growing centers (the "seeds") are placed at both sides of the

contour and along it as illustrated in figure 5. All the seeds starts at the same time a concurrent growing algorithm. Finally, the initial regions should be merged to eliminate small regions and to obtain a valid segmentation, where each region corresponds to an object in the image.

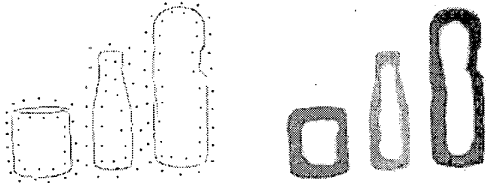


Figure 5: Scheme showing different phases of the proposed region growing algorithm. First, seeds are placed on each side of the contour. All the seeds of a side belong to the same region and for each one a growing thread is launched in a concurrent way

2.1.2 Selection of the items

The various regions are run over sequentially, with a new item offered to the user at each step. The item is offered through a visual effect highlighting the region containing the item in the original image. The user may reject the items offered until the one he/she wishes is reached. At this point the examination of the different items in the image stops.

2.1.3 Feature extraction

Finally we focus only on the item chosen and all the information the arm requires in order to effect the movement is extracted. The essential data are: the spatial location of the item and its measurements. It is also very important to try to recognize the item in such a way that a strategy for determining the best way to grasp and manipulate the item can be worked out.

2.2 Top-down approach

Figure 6 shows the block diagram of the proposed method. The strategy here is to directly search a given object by exploiting the knowledge about the scene. The information concerning the specific characteristics of objects is contained in the object model databases. On the other hand, the list of objects that are expected to be in the kitchen and their relationships, and expected positions are contained in

the scene model database. The use of scene models is required to perform several tasks such as to validate some initial results, to ensure a consistent description of the image, or even to guide the recognition process. Therefore, scene knowledge plays a significant role in this strategy. Furthermore the code is designed so that the data of the models allows to guide the recognition process and when this data changes, the recognition process changes without the need of recompiling. In order to directly recognize the object of interest two different kind of processes have been designed:

- A coordinating process. This process is in charge of reading the kitchen model in order to decide in which part of the kitchen the object of interest is located. After the robot is well positioned the coordinating process decides and calls the segmentation process. Furthermore this process has to ensure that the results of the segmentation are coherent with the scene model.
- A segmentation process. This task is based on a previous supervised learning task where the system learns how to recognize a specific object. The segmentation process is performed in two phases. First, a direct classification of pixels is performed based on a function that extracts predicted features according to the object model database. Second, after the pixels are grouped into regions, a region growing process is performed.

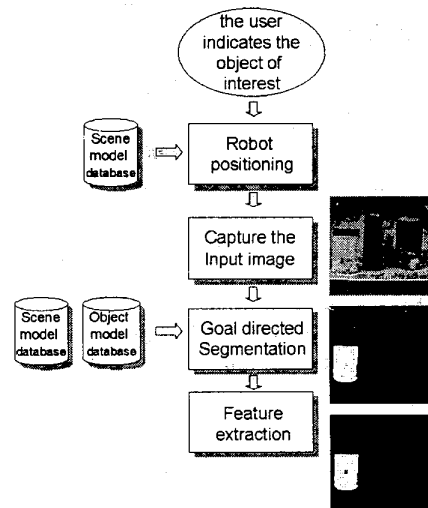


Figure 6: A scheme of the top-down interface for object selection

2.2.1 Scene and object modeling: learning the environment

The scene model is constructed with a graphic interface by relating objects that have been previously characterized in the object modeling task. Figure 7 shows a detail of the front-end of the application, where a model of a given kitchen scene appears. In the scene model the objects can be connected by three kinds of relationships: spatial (above, below, right to, left to, adjacent), KIND-OF (ako), and COMPOSED-BY (ipo). Furthermore, an object can have as special attribute the 3D position where the robot arm will be pre-positioned in front of it.

Object modeling starts with an interactive process where a teacher selects meaningful examples of objects in training images by clicking on blobs of pixels. This process enables to compute several color and texture features for every blob. Afterwards, the system selects the features that best characterize each object class, i.e. those features which separate the clusters in the feature space representing the object samples in some optimal way. The learning application gives the possibility to segment the pixels of the current image providing immediate visual feedback to the teacher. The visual feedback guides the teacher giving him the possibility to interfere in the process of feature selection, introducing new training images, or constraining the selection of some texture and color parameters.

The process of data model acquisition can be performed a priori, setting up an initial scenario. But, realistically the objects of a normal kitchen change their appearance, new objects can appear and some others disappear. Then, when a new object appears in the kitchen for the first time, and the user wants to select it, it is necessary to use the bottom-up approach. Then, the user has the possibility to indicate to the system that this object (a segmented region of the image) is new, and has a given name. Afterwards the system updates its database acquiring the object's location and extracting their most significant features. This behavior facilitates the knowledge engineering task because knowledge is acquired incrementally.

2.2.2 Segmentation based on previous learning

Following the "divide and conquer" paradigm, a decision tree separates the samples of the training images in a recursive way. The decision trees considered in our approach are binary trees with multivariate deci-

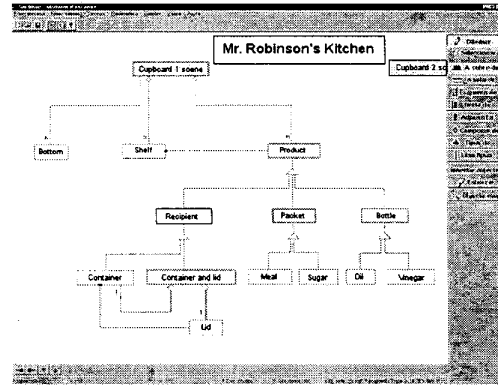


Figure 7: A detail of the graphic application for modeling scenes where a part of the kitchen model is shown. This application is used in the initial process of collecting the data model of a given kitchen.

sion functions, where each node is a binary test represented by a linear function. Each node of the tree attempts to separate, in a set of known instances (the training set), target (mapped as +) from non-target instances (mapped as -), but this is achieved only in certain ratio because realistic data is not always linearly separable. The resulting two subsets of samples are again subdivided into two parts by using two new calculated linear functions. This process is extended in a binary tree structure until an appropriate misclassification ratio is achieved. The result is a tree of hyperplane nodes that recursively try to divide the feature space into target and non-target samples.

The segmentation algorithm consists of dropping each pixel of the image down the tree, so that the pixels will be classified as + or -, representing the + a segmented pixel for that object class. Due to problems related to the fact of dealing with real images (different lighting conditions, objects can change their appearance due to its daily use, etc.) it is possible that some segmentation errors appear. Then, to improve the initial results, a final region growing process tries to fit these regions successfully.

2.3 Robot and Vision System integration

The structure of the CAPDI kitchen control system is shown in figure 8. The kitchen control unit is in charge of the control of the adapted kitchen's elements: the refrigerator, the carrousel cupboard, the lifting cup-

board, and the vision system.

As it has already been said, the vision system facilitates the object selection task to the users. Once the desired object has been chosen by the user, the vision system computes the grasping points. This is done by extracting a set of parameters related to its geometrical shape (basically the bounding rectangle corresponding to the segment associated to the chosen object) that allows to obtain the most suitable grasping points for each object or type of objects. When the coordinates of this grasping point are required by the kitchen control unit, a dynamic library call is engaged to obtain it. Afterwards, these coordinates are provided to the Robot Control module, which converts them in the appropriate control signals for the robotic arm. These signals provide an adequate and safe grasping of the selected object.

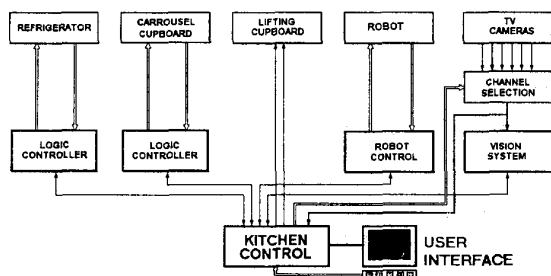


Figure 8: The CAPDI control diagram showing the interrelation between the involved modules

3 Conclusions and future trends

In the paper an interface specially designed for disabled users in a robotized kitchen has been described. In this sense, two main principles have guided the design of the interface: easy-to-use and open-interface. The goal has been to provide a friendly way to access to different objects as well as an interface which can be adapted to different kind of disabled people. Although two different methods have been presented, they have been conceived as complementary. Pre-positioning and object selection can be performed by manual, bottom-up, or top-down operation. These three options can be combined for the user in order to select the objects of interest. Furthermore, bottom-up and top-down approaches collaborate in the knowledge acquisition process that is necessary for top-down operation. This interface is being now evaluated by

users with different degrees of disabilities to extract from the user's trials the functional and operative performances and limitations. A specific four keys keyboard has been designed. A static and a remote unit have been build to control the elements at the given distance through the screen menu. By now, we are working on preparing the interface to accept voice input orders which will facilitate object selection tasks and improve the communication between the user and the whole system.

Acknowledgments

The project has been developed within the program PITER funded by IMSERSO, and CICYT (Spanish Research Agency) through the Special Actions frame. The authors wish to thanks the contribution of all the participants in the project, within the different teams.

References

- [1] J. Amat. *Technology advances and aids for disabilities*, Technological Aids for Disabled, edited by A. Casals, IEC. Barcelona, Spain, 1998.
- [2] K.A. Beasley and D. Davies. *Kitchen design for the wheelchair user*, Paraplegia News, August, 1991.
- [3] A. Casals, R. Merchan, et al. *CAPDI: A robotized kitchen for the disabled and elderly people*, Proc. of the Association for Assistive Technology Conference. Dusseldorf, Germany, 1999.
- [4] X.Cufi, A. Casals, J.Battle. *A Method to Obtain the Integrated Contour Image in Colour Scenes*, Proc. of the IEEE International Conference on Image Processing, pages 764-767. 1997.
- [5] Falah, Bolon, Cocquerez. *A region-region region-edge cooperative approach of image segmentation*, Proc. of the IEEE International Conference on Image Processing, pages 470-474. 1994.
- [6] J. Hennequin, R. Platts and Y. Hennequin. *Putting technology to work for the disadvantaged*, Rehabilitation Robotics Newsletter, vol. 4, no. 2. 1992.
- [7] T. Pavlidis, Y.T. Liow. *Integration region growing and edge detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 3, pages 225-233. March 1990.