The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

ROVIS: RObust machine VIsion for Service robotic system FRIEND

Sorin M. Grigorescu, *Student Member*, *IEEE*, Danijela Ristić-Durrant, and Axel Gräser

***Abstract*—In this paper the vision architecture, named ROVIS, of the robotic system FRIEND is presented. The main concept of the ROVIS is the inclusion of feedback structures between different components of the vision system as well as between the vision and other modules of the robotic system to achieve high robustness against external influences of the individual system units as well as of the system as whole. The novelty of this work lies in the inclusion of feedback control at different levels of the 2D object recognition system to provide reliable inputs to the 3D object reconstruction and object manipulation modules of the robotic system FRIEND. The idea behind this approach is to change the processing parameters in a closed-loop manner so that the current image processing result at a particular processing level is driven to a desired result. The effectiveness of the ROVIS system is demonstrated through the presentation of experimental results on 3D reconstruction of different objects from FRIEND environment.**

I. INTRODUCTION

ONE of the key requirements in the field of service rehabilitation robotics is the robust perception of the robot environment. As a result of progress in research on robot vision and technology development, the use of vision as a primary perception sensor for controlling manipulators has grown significantly in recent years [1,2,3]. A crucial requirement of a robot vision system is the achievement of a human-like robustness against complexity of the robot's environment in order to provide reliable visual information for autonomous assistance of human beings. A robot vision system is used to robustly analyze the images of complex scenes where the objects to be recognized are surrounded by a variety of other objects. As well as being robust against cluttered scenes, a robot vision system should be robust against unpredictability in the appearance of objects due to different external influences such as variable illumination. However, in spite of the significant work on the development of robot vision systems in recent years, robustness has remained a major problem. In order to concentrate on challenging physical tasks such as object grasping, the designers of complex service robotic systems usually simplify the 2D object recognition task by choosing objects to be manipulated which are convenient for 2D recognition by a state-of-the-art recognition method. One approach is to use different, a-priori known, colors for different objects classes so that objects can be recognized by an established color based recognition method. Another approach is to use objects which have sufficient texture characteristics to allow the application of methods exploiting local texture features such as SIFT model based recognition method [4][5]. The commonly used color based and texture based recognition methods usually use default parameters in an open-loop manner so that they give good results for specific working conditions; they are however sensitive to external influences such as variations in lighting conditions. The required robustness of the robot vision system is often achieved using additional sensors [6], which may increase costs and mechanical complexity. In this paper a novel robust robot vision system in which the necessary object recognition robustness is achieved by introduction of closed-loop control structures at image segmentation level of the 2D object recognition system is presented. The main idea behind this is the automatic adjustment of the processing parameters instead of using their default values. This is a novel alternative to the above discussed conventional approach of using additional sensors or introducing a more controlled environment. The presented vision system ROVIS (*RObust machine VIsion for Service robotics*) is integrated into the semi-autonomous rehabilitation robotic system FRIEND (*Functional Robot arm with frIENdly interface for Disabled people*) which has been developing at the Institute of Automation of University of Bremen since 1997 within different projects [7]. Within the research project AMaRob (*Autonomous Manipulator control for rehabilitation Robots*) the goal has been achieved of supporting disabled people with impairments of their upper limbs in Activities of Daily Living (ADL) and professional life and giving them independence from nursing staff for at least 1,5 uninterrupted hours. In different robot working scenarios a large number of different action sequences like "pour and serve a drink", "take, prepare and serve a frozen meal", "fetch and handle a book" are necessary to fulfil the robotic system user's demands. Hence, the ROVIS system has to deal with a variety of objects, including a bottle, a glass, a

Manuscript received July 28, 2009.
The authors are with the Institute of Automation, University of Bremen, Otto-Hahn-Allee NW1, 28359 Bremen, Germany (phone: +49-421-218-3490 / 7523; fax: +49-421-218-4596 / 4707; e-mail: {grigorescu, ristic, ag} @iat.uni-bremen.de).

book, a meal tray, fridge and the microwave oven. Some of these objects are uniformly coloured like the bottle, the glass and the handle of the meal tray, while some of them are textured such as the book or some bottle types. Furthermore, some of the objects to be recognized may be located in clustered environments, for example there may be several objects in the fridge or on the book shelf. Hence, the robot vision system must be robust enough to cope with the clustered environment (complex scenes) and with a variety of different objects as well as with different appearances of the same object in different lighting conditions that arise during the robot functioning.

From the image processing point of view, the objects to be recognized in the FRIEND system are classified into two categories: 'container' objects such as the fridge, microwave oven and book shelf, and objects to be manipulated such as bottles, glasses, meal trays and books. The focus in this paper is on robust recognition of objects which have appropriate size and shape for manipulation by the robot arm. The recognition of containers, achieved by a SIFT model based method [8], is taken for granted. Moreover, the result of recognition of containers is used as a possible starting point for definition of the image *Region Of Interest* (ROI) as will be explained in Section III. The rest of the paper is organized as follows. The FRIEND's control architecture, including ROVIS, is presented in Section II. ROVIS itself is described in Section III. Since this paper concentrates on one ROVIS module, robust 2D object recognition, the other modules of the ROVIS structure are only briefly described. The closed-loop segmentation of the image ROI for the purpose of reliable feature-based object recognition and reconstruction is explained in Section IV. The experimental results on 3D reconstruction of different objects from FRIEND environment are given in Section V.

II. The Robotic System FRIEND

The system FRIEND consists of a 7 DoF (Degrees of Freedom) manipulator mounted on an electrical wheelchair can only be achieved using an appropriate control framework. The architecture used, entitled MASSiVE (*MultiLayer Architecture for SemiAutonomous Service Robots with Verified Task Execution*), represents a distributed robotic control system which combines reactive behaviour with classical artificial intelligence based task planning capabilities [9]. The MASSiVE architecture is represented in Fig. 2 with its structure divided in four specific modules. The *Human-Machine Interface* (HMI) operates at the user interaction level.

The user commands are acquired with the help of different input methods such as speech recognition, chin control and Brain-Computer Interface (BCI) and translated further into machine language for interpretation [9][10]. The processing algorithm that converts a user request into robot actions resides in the *Reactive Layer*. Here, the data collected from different *Sensors,* such as the stereo cameras and a tactile tray, are processed in order to "understand the environment". The data is further converted into actions by the available *Actuators* such as 7DoF manipulator. The sequence of operations needed to perform a specific task is generated by the *Sequencer* module. The *Sequencer* plays the role of a *Discrete Event Controller* (DEC) that plans sequences of operations by means of predefined task knowledge. Through the functioning of the system, the computed data is shared between the modules with the help of the *World Mode*. The World Model defines the information produced and consumed by the operations in the Reactive Layer. The software interconnection between the processing layers is implemented using CORBA (*Common Object Request Broker Architecture*).

As displayed in Fig.2, the vision framework ROVIS is placed inside the Reactive Layer where it provides visual information for the Sequencer which further activates the manipulator. ROVIS communicates with the sub-symbolic layer of the World Model to where it outputs the reconstructed 3D environment. This information is used

Fig. 1. The rehabilitation robotic system FRIEND operating in complex environment.

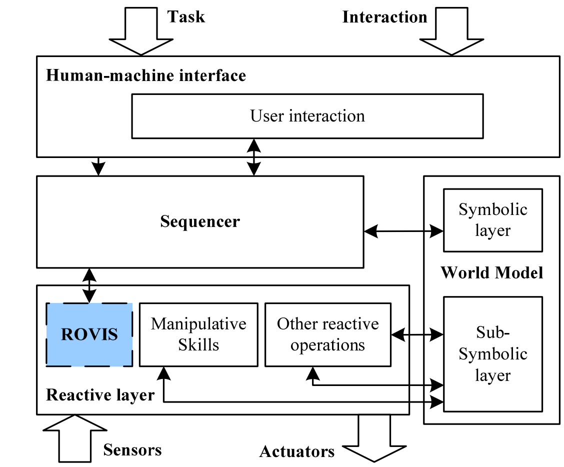

Fig. 2. The MASSiVE control architecture.

further by the manipulative skills for virtual modeling of the scene, consisting of objects to be manipulated within containers as well as obstacles, viewed by the robotic camera system. The virtual 3D scene model is used further for collision free path planning [11]. On the other hand, for performing the ROI definition, object recognition and reconstruction tasks ROVIS uses necessary information from the World Model such as the features of an object class needed for object classification.

III. The ROVIS Concept

Fig. 3 shows a schematic overview of the ROVIS buildi- ng blocks. Arrows connecting the blocks illustrate the flow of information through the ROVIS system as well as the connections of the ROVIS components with the external modules, the Human-Machine Interface and other reactive operations of the system FRIEND. As can be seen, there are two main ROVIS components: hardware and the object recognition and reconstruction chain.

The ROVIS hardware consists of a Bumblebee® stereo-camera system mounted on a pan-tilt head placed on a special rack behind the user, above his head, as illustrated in Fig. 1. Using a special input device such as a chin joystick, the user of the semi-autonomous system FRIEND navigates the system in front of the container related to the particular working scenario. The stereo cameras view the scene in front of the robotic system including the manipulator and the tray mounted on the wheelchair in front of the user. In the ROVIS initialization phase the extrinsic camera parameters are calculated through camera calibration. The viewing angle of the sensors can be changed through the pan-tilt control so that the container as initialized by the Sequencer, can be detected in the image. This is illustrated in Fig. 3 by the feedback from Container Detection to the Camera Pan-Tilt Head block.

The ROVIS object recognition and reconstruction chain consists of a sequence of image processing operations used for the extraction of features needed for both 2D recognition and 3D reconstruction of the objects present in the manipulator's environment. The main concept of ROVIS is to apply the image processing operations on the image ROI rather than on the whole image. This is motivated by the observation that people focus their visual attention on the region around an object when they grasp it as illustrated in Fig.1. 3D reconstruction data are needed for a "look-and-move" type of robot control. In order to achieve satisfactory precision of 3D stereo reconstruction the high image resolution of 1024x768px is used.

A. ROI definition

An image ROI can be defined for two cases which differ with respect to the level of a-priori knowledge about the location of the object to be manipulated within the image. In the first case only a partial knowledge about the object environment is available. For example, in the FRIEND system the available information is of the form: "the object is in the fridge" or "the object is on the shelf". Bearing in mind that the container objects in the FRIEND environment are a permanent feature of the scenarios, the SIFT method [8] is used for their recognition. This method uses a model image to train a classifier off-line. During on-line system operation the SIFT algorithm searches for the model image in the scene through a matching based algorithm. Once the model image has been detected its pose (position and orientation in 3D space) can be reconstructed. Knowing the position of the model image placed on or in a container (e.g. in the fridge) the

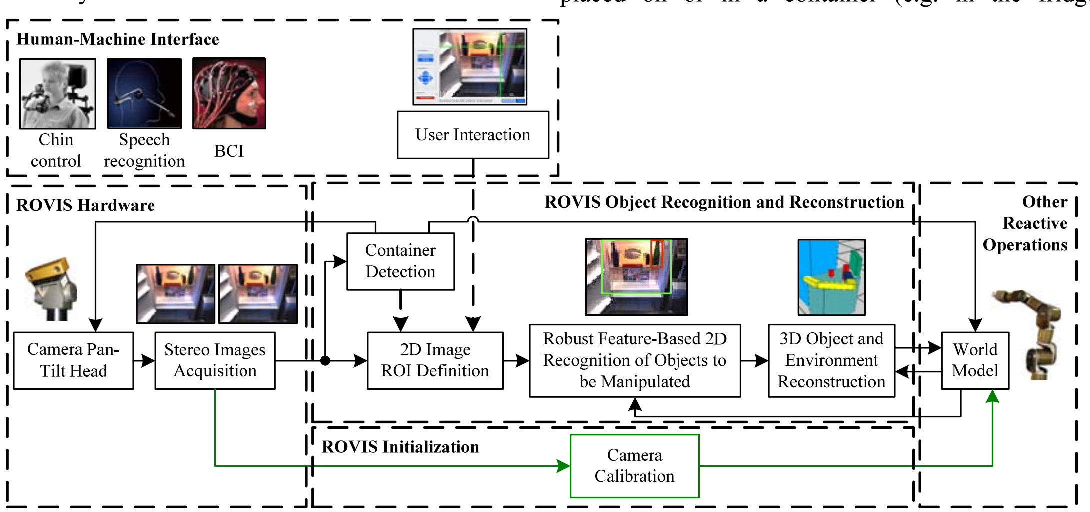

Fig. 3. Block diagram of ROVIS, the robust vision framework of the rehabilitation robotic system FRIEND.

container pose can be further reconstructed. Starting from the reconstructed 3D pose of the container, the container region in the image is obtained using 3D to 2D mapping.

The resulting image region enclosing the container, in which the object of interest is located, represents the image ROI. Hence, in this case, the defined ROI encloses all the objects present in the container and not just the object of interest. For example in the “serve a drink” scenario, where the task of the manipulator is to fetch a bottle with a drink from the fridge, such situation corresponds to a user’s command “I want a drink”.

The second possible case regarding the ROI definition is the case where precise information on the object position within the image is available through the human-machine interface (HMI). For example, the user can locate the object of interest by using a particular action, such as clicking on the displayed image using a special input device, such as a chin joystick, as illustrated in Fig. 3 for the case of “serving a drink” scenario. Starting from the user’s command “I want this drink” and an interest image point defined by the user, the size of the rectangular image ROI is automatically adjusted in order to fully bind the object of interest.

The automatic adaptation of the size of the ROI when it is defined through the HMI as well as defining the ROI using 3D to 2D container mapping is beyond the scope of this paper. In this paper the focus is on the robust recognition of uniformly colored objects of interest within the defined ROIs without using a-priori knowledge about the objects characteristics such as color and size.

IV. Feedback Control of image segmentation For reliable Object Recognition

A crucial requirement for reliable feature-based 2D object recognition and subsequent 3D object reconstruction is that the object segmented image is of good quality. A segmented image is said to be of good quality if the pixels of the object of interest forms a well shaped segmented object region. The image segmentation algorithm presented in this paper employs the idea of the inclusion of feedback structures to control the quality of the binary segmented image ROI. The idea behind this approach is to adjust the parameters of image segmentation in a closed-loop manner so that the current segmented image ROI is driven to the one of reference quality independently of external influences. This idea is suggested and investigated in detail both generally and within the context of specific gray level image processing applications in [12]. The authors published first results on the extension of this approach to color image processing in service robotics, when a-priori information about the color of objects of interest is available in [13]. In this paper, further improvement of the closed-loop method to be used for recognizing uniformly colored objects in a clustered scene without a-priori information is presented. The novelty of this paper also concerns the performance evaluation of the proposed closed-loop segmentation method. Namely in previous authors work [e.g. 12], the benefit of the proposed method over the traditional adaptive thresholding in image segmentation is demonstrated. In the section V of this paper it is shown that the proposed closed-loop thresholding contributes directly to the robustness of 3D reconstruction.

A. Choice of actuator and controlled variables

The use of closed-loop control in image processing differs significantly from its use in conventional industrial control, especially concerning the choice of the actuator and the controlled variables. Generally, the actuator variables are those that directly influence the result of image processing. In the system presented here, the image segmentation is done by thresholding of the so-called *Hue* image, which contains the pure color information of the original RGB image of a scene from a FRIEND working scenario. In thresholding, each pixel from the Hue image to be segmented is set to the foreground black color in the output segmented image if its pixel value belongs to a particular interval of the color values [13]. To further explain the thresholding operation the Hue image is defined as a 2D function $f(x,y)$ and the object color interval as $C_l = [T_{min}, T_{max}]$ where T_{min} and T_{max} are the minimum and maximum color values across the object’s pixels. Then, the thresholding operation is defined as:

$$t(x,y) = \begin{cases} 1, & \text{if } f(x,y) \in C_l, \\ 0, & \text{if } f(x,y) \notin C_l. \end{cases} \quad (1)$$

where $t(x,y)$ is the segmented binary image, 1 and 0 represent black and white color respectively, and x and y are the Hue image pixel coordinates. For the sake of clarity an object color interval C_l in the following is referred to as an *object thresholding interval.*

The thresholding operation is highly sensitive to the illumination condition. Due to the pixel color uncertainty arising from changes in illumination during image acquisition, different thresholding intervals are needed to segment the same object at different time instances. This can be seen from Fig. 4 which shows two images of the same scene from a FRIEND working scenario which contains a fridge with various objects placed inside it,

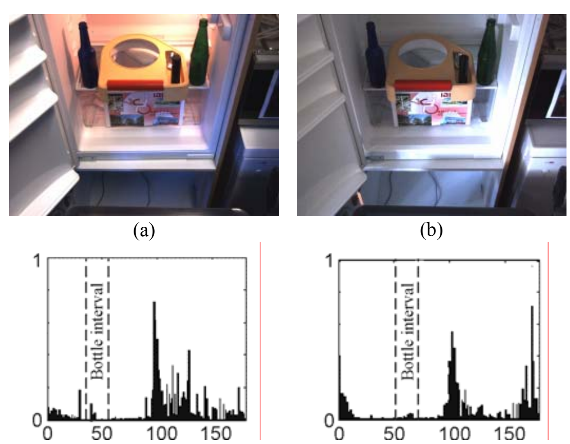

Fig. 4. RGB image of the scene containing a green bottle captured in artificial (a) and daylight illumination condition. (c) and (d) histograms of the corresponding Hue images overlaid with the thresholding interval of the green bottle.

captured in different illumination conditions. Fig. 4 shows also the histograms of the Hue planes of the considered images from which it can be seen that the thresholding object interval in the image taken in daylight conditions is shifted to the right with respect to the thresholding interval of the same object in the image captured in artificial illumination conditions. Therefore, in order to achieve good object segmentation it is necessary to adjust the object thresholding interval according to illumination. For this reason, the object thresholding interval, that is the threshold increment i, $[T_{min}+i, T_{max}+i]$, is considered as the actuator variable $u = i$ in the presented system

In order to automatically adjust the thresholding interval so that the current quality of segmented image ROI is driven to the desired, reference, value a controlled variable has to be defined. The controlled variable has to be appropriate from the control, as well as from the image processing, viewpoint. From the image processing viewpoint, a feedback variable must be an appropriate measure of image ROI quality. Two basic requirements for control are: it should be possible to calculate the chosen quality measure easily from the image and the closed-loop should satisfy the input-output controllability conditions. Input-output controllability primarily means that for the selected output (controlled variable) an input (actuator variable) which has a significant effect on it must exist in the image processing chain.

Bearing in mind the qualitative definition of a segmented image ROI of good quality given above, the following quantitative measure of ROI quality has been proposed:

$$I = -\log_2 p_8,\ I(0) = 0\,, \tag{2}$$

where p_8 is the estimate of the probability of a segmented pixel surrounded with 8 segmented pixels in its 8-pixel neighbourhood:

$$p_8 = \frac{\text{number of segmented pixels surrounded with 8 segmented pixels}}{\text{total number of segmented pixels in the image ROI}}. \tag{3}$$

Having in mind that a good segmented image ROI contains a "full" (free of holes) segmented object region, it is evident from (3) that a small probability p_8 corresponds to a large disorder in a segmented image ROI and consequently a large *uncertainty I*, defined by (2), is assigned to the segmented image ROI. Therefore, for a reliable segmentation the goal is to achieve the segmented image ROI having as small as possible uncertainty measure *I*.

To investigate the system input-output controllability when considering the threshold increment as the input variable, and the proposed uncertainty measure *I* as the output variable, the thresholding of the ROI, containing only the green bottle, in the images shown in Fig. 4(a) and Fig. 4(b) was done. The thresholding interval was set to an initial state $[T_{min}, T_{max}] = [0,\ 20]$. To this interval the increment $u = i$ was added as $[T_{min}+i, T_{max}+i]$. For each segmented image corresponding to the increment $i \in [0,179]$, the uncertainty measure *I* was calculated. The resulting input-output characteristics are presented in Fig. 5(a) and Fig. 5(c). As can be seen, the uncertainty *I* is sensitive to the chosen actuator variable across its effective operating range. Also, it is clear that each input value is mapped to at most one output value and that it is possible to achieve the minimum of *I*, corresponding to the segmented object image of reference good quality, by changing the thresholding boundaries. The satisfaction of these prerequisites for a successful control action demonstrates the pair "thresholding increment i – uncertainty measure I" as a good "actuator variable – controlled variable" pair.

The same experiment, as described above, was done also for the case of the ROI enclosing the whole fridge in Fig. 4(a) and Fig 4(b) thus containing more objects, 2 bottles of different colors and a meal tray. The resulting input-output characteristics are shown in Fig. 5(b) and Fig. 5(d). As can be seen, the characteristics have more local minima. Each minimum corresponds to a good segmentation of a particular object. For example, the

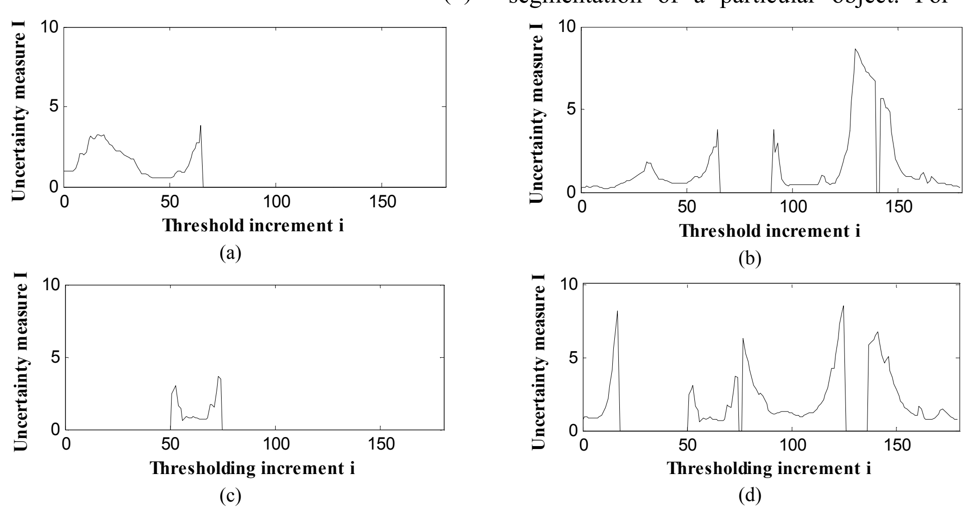

Fig. 5. The uncertainty measure *I* of pixels segmented from the images taken in different illumination conditions vs. thresholding increment i, $[T_{\min}+i, T_{\max}+i] = [0+i, 20+i]$. (a) and (c) ROI containing only one object. (b) and (d) ROI containing more objects.

minimum at [125, 145] in Fig. 5(b) corresponds to a blue object while the minimum achieved for the optimal threshold interval [35, 55] corresponds to the green bottle, as in the case of the previously described experiment and characteristic shown in Fig. 5(a). Also, it can be seen that the minima which correspond to optimal object thresholding intervals in the image taken in daylight conditions are shifted with respect to the thresholding intervals of the same object in the image captured in artificial illumination conditions, indicating the need to adjust thresholding intervals to different illumination conditions.

Based on the above discussion it can be said that the original problem, that of finding the optimal object threshold interval that provides a segmented object image of good quality, appropriate for subsequent object feature extraction, can be interpreted and converted to the problem of finding the minimum uncertainty I of the object region in the binary segmented image.

B. Closed-loop control design

In the presented system, the reference value of the chosen controlled variable is not explicitly known as the intention is to develop an object recognition method which does not uses a-priori knowledge about the object characteristics such as object size and color. However, the selection of an image ROI quality measure whose minimal value corresponds to the image ROI of good quality has been suggested for the controlled variable. Hence, the optimal value of the chosen controlled variable can be achieved by an optimization process using an appropriate extremum seeking algorithm through a control structure, as shown in Fig. 6. Here the feedback information on the image ROI quality is used to choose the optimal value u_{opt} of the actuator variable u, that is, to drive the current image ROI to one with reference optimal quality.

An optimal threshold interval ensures that a reliable input is given to the feature extraction step where different object features needed for object classification are extracted. Such features are the Hu moments [13] which uniquely describe the shape of objects.

C. Object classification and recognition

As discussed above, the goal of the presented closed-loop ROI segmentation system is to extract the optimal threshold intervals for all the objects present in the image ROI. The input image of a scene from the FRIEND environment is then segmented using the extracted threshold intervals. Therefore the segmentation result is either one binary image, in the case of an image ROI containing only one object, or as many binary images as there are objects present in the defined image ROI. The latter case is shown in Fig. 7. The input-output characteristic shown in Fig. 7 is obtained in the experiment described in Section IV A. As explained, each local minima of the input-output characteristic corresponds to an object of a particular color. The classification of objects as belonging to the class "bottle", "handle" or "noise" is done based on the object's shape descriptors extracted from the resulting binary segmented images and general knowledge on object classes stored in the World Model of system FRIEND.

Once the object of interest is correctly identified in the image, the image coordinates of the object feature points are calculated at the object recognition level. A so-called object feature point is used for solving the stereo correspondence problem for 3D object reconstruction. For example, for a bottle the object feature point is the top neck point while for the meal tray the feature point is the middle of its handle [14].

V. Performance Evaluation

The ultimate goal of the ROVIS system is reliable 3D reconstruction of objects to be manipulated which should assure correct 3D modeling of the FRIEND environment for the purpose of collision free path planning [9]. Therefore, the evaluation of the ROVIS effectiveness is done through the comparison of manually measured and automatically calculated 3D features of three different bottle objects. Two different methods are used for the automatic calculation of the 3D features: the proposed 3D reconstruction based on closed-loop object segmentation and 3D reconstruction based on traditional open-loop segmentation. In contrast to the closed-loop method, which uses feedback information on the segmentation result to adjust the thresholding parameter, the open-loop method uses a constant reference thresholding parameter. This threshold is determined offline by manual thresholding of the object image captured in reference illumination condition.

A scene from the FRIEND working scenario "serve a drink" shown in Fig. 4 was imaged in different

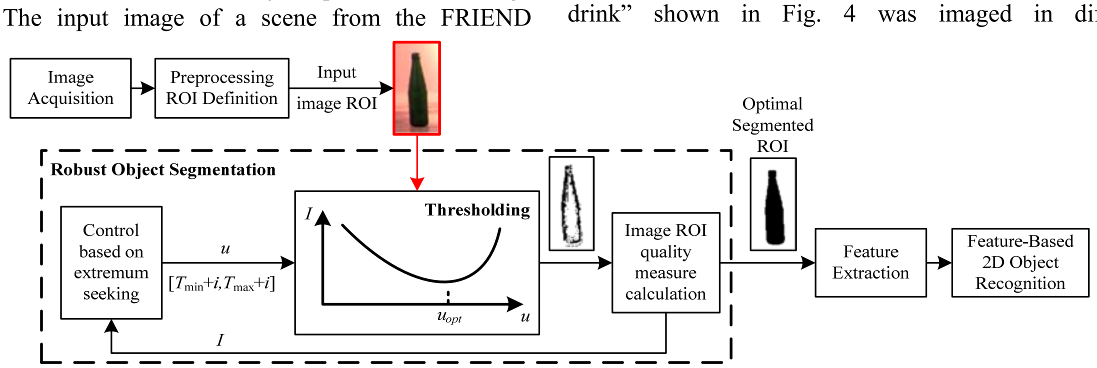

Fig. 6. Block-diagram of the proposed feature-based object recognition system with closed-loop image ROI segmentation.

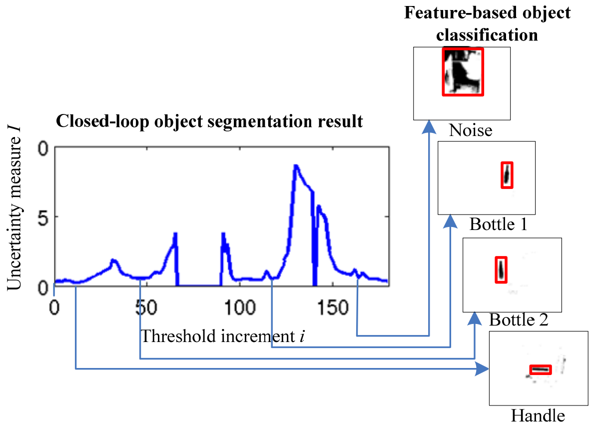

Fig. 7. Object classification based on the proposed closed-loop image segmentation.

illumination conditions ranging from 15lx to 570lx. This range of illumination corresponds to a variation of the light intensity from a dark room lighted with candles (15lx) to the lighting level of an office according to the European law UNI EN 12464 (500lx). Each captured image was segmented using the two tested segmentation methods. The object feature points were extracted from each resulting segmented image and subsequently the 3D object coordinates were calculated and compared to the real measured 3D localtions in order to calculate coordinates errors X_e, Y_e and Z_e. Also the heights of the bottles and the width of the mealtray handle were estimated based on extracted top neck and bottom feature points, that is based on extracted right and left end feature points, and compared to the real bottle heights, as the error H_e, and real mealtray handle width, as error W_e, respectively. The comparison results are shown in Fig. 8. The statistical measures of achieved error in experiments performed in different illumination conditions are given in Table I.

As it can be seen, the 3D objects features calculated using the segmented images resulting from the proposed closed-loop method only differs slightly from the real coordinates over the whole considered illumination range, thus demonstrating the robustness of ROVIS. However, the 3D objects features calculated from the segmented images resulting from the open-loop method which uses constant segmentation parameters significantly differs from the real coordinates for a number of illumination conditions which differ from the reference illumination of 200lx. This indicates the importance of using feedback information from the current segmentation result to adapt the segmentation parameters to different environmental conditions.

VI. CONCLUSION

In this paper the novel robust vision system ROVIS of the rehabilitation robotic system FRIEND is presented. One of the main ROVIS concepts is the inclusion of feedback structures between different components of the vision system as well as between the vision and other components of the robotic system in order to achieve high robustness of the individual units as well as of the overall system against external influences such as variable illumination. The emphasis is on the feedback control of image segmentation for providing reliable input to higher vision levels, 2D object recognition and 3D object reconstruction, independently of different external influences. The presented experimental results on 3D reconstruction of features for two different objects from the FRIEND environment demonstrate the benefit of using feedback information on the current segmentation result to adjust the segmentation parameters in order to provide the necessary robustness of the robot vision system.

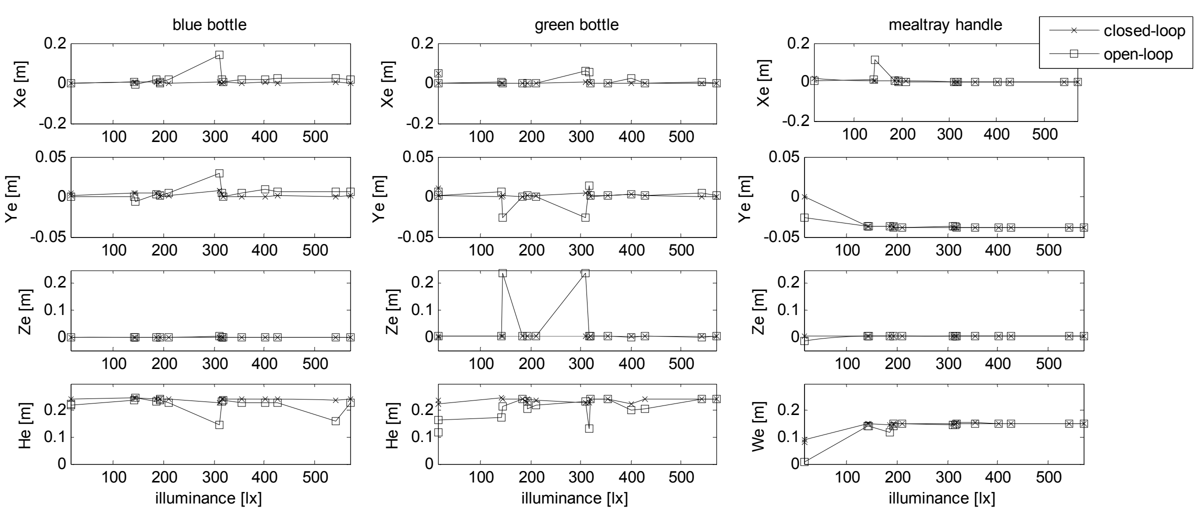

Fig. 8. Difference between the real 3D object features and the 3D features calculated from the segmented images resulting from the proposed closed-loop segmentation and from the traditional open-loop segmentation of bottle objects images captured in different illumination conditions.

TABLE I

STATISTICAL RESULTS OF OPEN-LOOP VS. THE ROVIS OBJECT RECONSTRUCTION METHOD.

	Open-loop				Closed-loop			
	X_e [m]	Y_e [m]	Z_e [m]	H_e/W_e [m]	X_e [m]	Y_e [m]	Z_e [m]	H_e/W_e [m]
Max error	0.1397	0.0391	0.2357	0.1130	0.0049	0.0086	0.0029	0.0341
Mean	0.0146	0.0083	0.0121	0.0359	0.0024	0.0051	0.0017	0.0051
St. deviation	0.0331	0.0071	0.0001	0.0282	0.0016	0.0021	0.0001	0.0044

REFERENCES

[1] D. Kragic and H. I. Christensen, "Advances in robot vision", *Robotics and Autonomous Systems*, 52, pp. 1-3, 2005.

[2] D.J. Kim, R. Lovelett and A. Behal, "An Empirical Study with Simulated ADL Tasks using Vision-Guided Assistive Robot Arm", in *Proc. 11th International Conference on Rehabilitation Robotics - ICORR'09,* 2009.

[3] C. Dune, C. Leroux, E. Marchand, "Intuitive human interaction with an arm robot for severely handicapped people – a one click approach", in *Proc. 10th International Conference on Rehabilitation Robotics - ICORR'07,* 2007.

[4] T. Asfour, P. Azad, N. Vahrenkamp, K. Regenstein, A. Bierbaum, K. Welke, J. Schröder and R. Dillmann, "Toward Humanoid Manipulation in Human-Centred Environments", *Robotics and Autonomous Systems*, pp. 54-65, 2008.

[5] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, 60(2), 2004.

[6] D. Kragic, M. Björkman, H. Christensen and J.-O. Eklundh, "Vision for Robotic Object Manipulation in Domestic Settings", *Robotics and Autonomous Systems*, 52, pp. 85-100, 2005.

[7] O. Ivlev, C. Martens and A. Graeser, "Rehabilitation robots FRIEND-I and FRIEND-II with the dexterous lightweight manipulator," *Restoration of Wheeled Mobility in SCI Rehabilitation*, vol. 17, pp.111–123, 2005.

[8] F. Alhwarin, C. Wang, D. Ristic-Durrant and A. Gräser, "Improved SIFT-Features Matching for Object Recognition", *in Proc. BCS International Academic Conference 2008 -- Visions of Computer Science*, 2008.

[9] O. Prenzel, C. Martens, M. Cyriacks, C. Wang and A. Gräser, "System controlled user interaction within the service robotic control arhitecture MASSiVE", *Robotica, Special Issue*, 2007.

[10] T. Lüth, D. Ojdanic, O. Friman, O. Prenzel and A. Gräser, "Low Level Control in a Semi-autonomous Rehabilitation Robotic System via a Brain-Computer Interface" in *Proc. 10th International Conference on Rehabilitation Robotics – ICORR'07*, 2007.

[11] D. Ojdanic and A. Gräser, "Improving the Trajectory Quality of a 7 DOF Manipulator", in *Pro. Robotik 2008, Munich, Germany*.

[12] D. Ristic, *Feedback Structures in Image Processing*, Shaker Verlag, 2007.

[13] S. K. Vuppala, S. M. Grigorescu, D. Ristić and A. Gräser, "Robust Color Object Recognition for a Service Robotic Task in the System FRIEND II", in *Proc. 10th International Conference on Rehabilitation Robotics – ICORR'07,* 2007.

[14] S. K. Vuppala and A. Gräser, "An Approach for Tracking the 3D Object Pose Using Two Object Points", in *Proc. 6th International Conference on Computer Vision Systems, Santorini, Greece*; 2008.