X Triennial International SAUM Conference
on Systems, Automatic Control and Measurements
Niš, Serbia, November 10th-12th, 2010

Invited Paper

Robust Machine Vision for Complex Service Robotic Systems

Sorin M. Grigorescu+ and Axel Gräser++

+Department of Automation, Transilvania University of Braşov, Mihai Viteazu 5, 500174, Braşov, Romania
Phone: +40-268-418-836, E-Mail: s.grigorescu@unitbv.ro

++ the Institute of Automation, University of Bremen, Otto-Hahn-Allee NW1, 28359 Bremen, Germany
Phone: +49-421-218-62470 / 7523, E-Mail: ag@iat.uni-bremen.de

***Abstract* - In this paper, a concept for designing machine vision systems for service robotics is presented. The focus of the paper is on implementing novel vision methods and architectures based on the inclusion of feedback control at image processing level, with the purpose of improving the perceptual capabilities of service robots. A proposed vision system, entitled ROVIS, will be described along with its integration within the robotic platform FRIEND. The objective of FRIEND is to help disabled and elderly people in their daily and professional life activities.**

I. INTRODUCTION

In recent years, especially in the last two decades, the worldwide healthcare community showed a high interest on rehabilitation robotic systems that can partially overtake tasks that usually are carried out by care-giving personnel [1]. The growing interest in this field of robotics is due to the fact that in a large number of healthcare areas there is a lack of trained personnel. Parallel to this, the number of elderly and persons with disabilities is increasing every year [2]. In industrialized countries, like US, Europe, Japan or Canada, the number of estimated persons which suffer from a certain disability is approx. 75 mil., whereas the number of elderly is approx. 130 mil. [3]. A certain percentage of persons from the mentioned statistics suffer from a form of severe disability which requires a 24h/day assistance from trained personnel. In Europe alone, the number of persons which suffered *Spinal Cord Injuries* (SCI) above vertebra C5 or a stroke with similar symptoms is estimated to be at around 9000 and increasing every year by 1500. The number of patients who could use a rehabilitation robot after a stroke is approx. 10 times higher then that with SCI. This group of persons with high disabilities currently need personal support for 24h/day and long for any technical system which could give them a form of independence and functional restoration, even if limited.

Assistive robots are mainly designed to support persons with disabilities in *Activities of Daily Living* (ADL) and professional life. The care-providing robot FRIEND (*Functional Robot with dexterous arm and user-frIENdly interface for Disabled people*) illustrated in Fig. 1 and commercially available since the beginning of 2010, is a semi-autonomous robot designed to support disabled and elderly people in their daily life activities, like preparing and serving a meal, eating, drinking, or reintegration into professional life. FRIEND, which fells into the category of wheelchair-mounted manipulator arm systems, enables a disabled user to perform a large set of tasks in daily and professional life self-determined and without any help from other people like therapists or nursing staff. Usually, patients with disabilities have to rely on care-giving personnel 24 h/day. The independence given to them through the FRIEND system presently aims to a minimum of 90 uninterrupted minutes, where certain tasks, commonly performed by trained persons, are transferred to the robot. The achieved independence is a proven benefit in the social life of the patients.

FRIEND is the result of more then a decade's work in the field of assistive robotics performed at the Institute of Automation from the University of Bremen in Germany. The rehabilitation robot is the 3rd generation of assistive robots designed at the institute, after FRIEND I [4] and FRIEND II [5].

One of the key requirements in the field of assistive rehabilitation robotics is the robust perception of the robot's environment, aiming at reliable 3D environment reconstruction which is necessary for autonomous object manipulation. As a result of progress in research on robot vision and technology development, the use of vision as a primary perception sensor for providing information for controlling manipulators has grown significantly in recent years [6], [7]. As well as being robust against cluttered scenes, a robot vision system has to be robust against unpredictability in the appearance of objects due to different

Fig. 1. The care-providing robotic system FRIEND operating in a complex environment.

external influences such as variable illumination.

In this paper, we present a vision system, entitled ROVIS, which aims at giving FRIEND a reliable perceptual module to be used in the understanding of its surrounding environment. The methods used in ROVIS are based on the inclusion of feedback structures at image processing level [8], a design concept which has as goal the improvement of visual system with respect to external influences, such as cluttered scene and variable illumination conditions.

The rest of the paper is organized as follows. In Section II the concept of the ROVIS machine vision system is presented, followed in Section III by the design details of two robust feature extraction methods. Conclusions are given in Section IV.

II. THE ROVIS CONCEPT

The objective of the rehabilitation robotic system FRIEND is to help disabled patients in their daily life activities. Thus, the robot operates in a human, unstructured environment. FRIEND is equipped with a 7 *Degrees of Freedom* (DoF) manipulator arm mounted on an electrical wheelchair and a series of sensors used by the robot to understand its environment, like a stereo camera system attached to a pan-tilt head. The purpose of ROVIS is to provide reliable visual information to the robot in order to plan the 7-DoF manipulator motion and to grasp the recognized objects of interest [9].

ROVIS is integrated within the Reactive Layer of the overall control system of the robot, presented in Fig. 2. Visual information is provided to the *Sequencer* which plans sequences of operations needed to perform specific tasks and to activate the planning of the manipulator arm. The Sequencer plays the role of a Discrete Event Controller 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 Model*. The World Model defines the information produced and consumed by the operations in the Reactive Layer. 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 and translated further into machine language for interpretation [5]. The software interconnections between the processing layers are implemented using the *Common Object Request Broker Architecture* (CORBA).

Fig. 3 shows a schematic overview of the ROVIS building 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 HMI 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 FRIEND navigates the system in front of the container related to the particular scenario, for example a fridge.

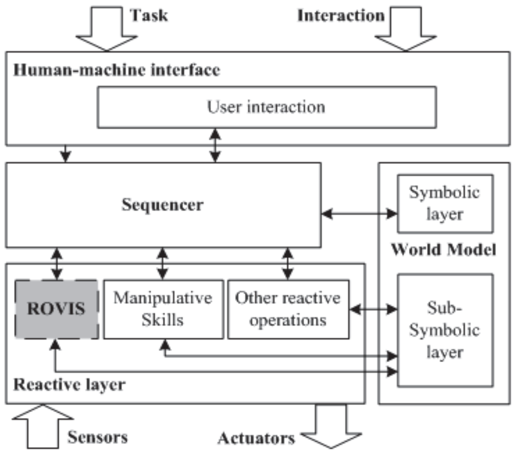


Fig. 2. Overall control architecture of the service robot FRIEND.

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 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 an image *Region of Interest* (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. Information with respect to the 3D position of objects is needed for a "look-and-move" type of robot control.

III. ROBUST REGION AND BOUNDARY FEATURE EXTRACTION

In this section, the implementations of the closed-loop mechanism used in ROVIS are introduced. Both of them deal with object recognition based on different segmentation methods, region and boundary based, respectively.

The overall goal of the components from the block diagram illustrated in Fig. 3 is to reliably extract the 3D pose of objects so that a robust autonomous manipulator action is accomplished [8]. 3D reconstruction of objects to be manipulated is strongly dependent on the result of object recognition, that is, on the result of feature extraction and data classification. One key requirement for optimal pose estimation is the robust segmentation of the objects of interest. In the following, the design of two closed-loop feature extraction modules will be detailed for two types of segmentations, region and boundary based, respectively. The final 3D reconstruction is represented as a component which takes as input 2D object feature points and the determined object class. The goal of the processing chain is to parallel process left and right stereo images in order to extract the feature points of the objects to be manipulated

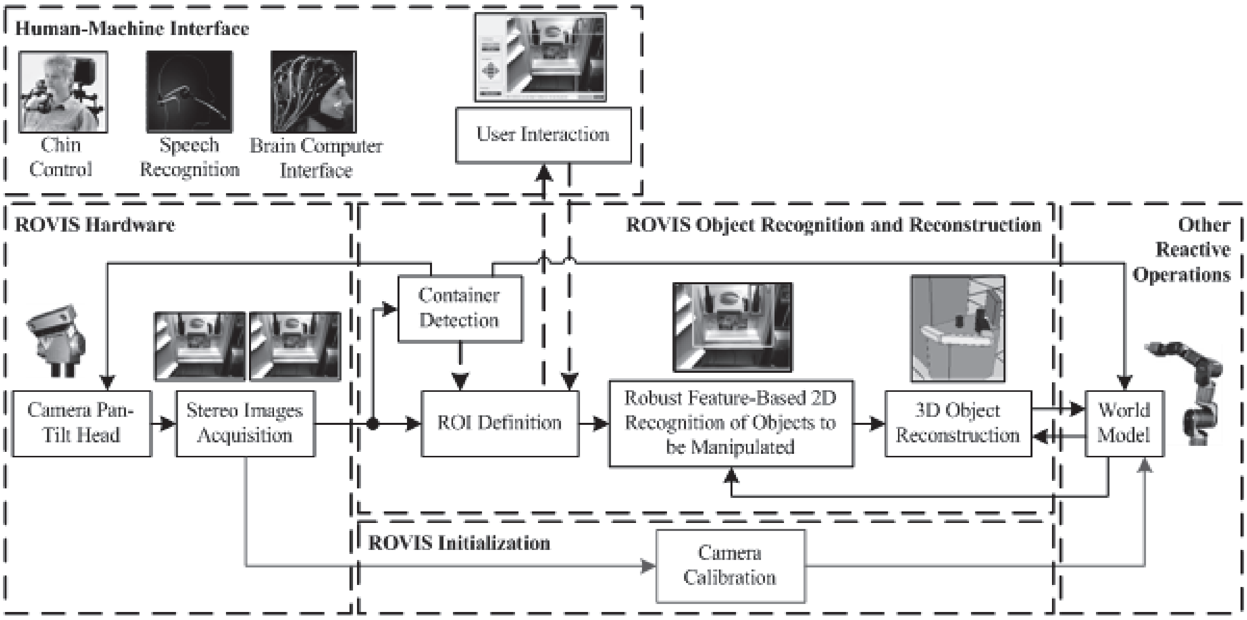
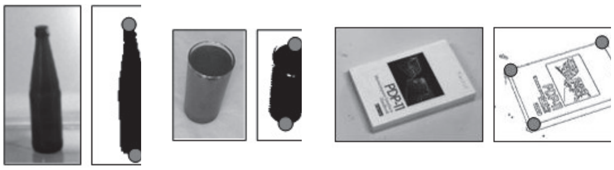


Fig. 3. Block diagram of ROVIS, the robust vision architecture of the service robotic system FRIEND.

and used them for pose reconstruction. Different feature points obtained for different segmented objects can be seen in Fig. 4.



(a) (b) (c)

Fig. 4. Feature points extraction based on different segmentation types. (a,b) Region based segmentation. (c) Boundary based segmentation.

Depending on the required number of points, the feature points are represented as:

$$\begin{cases} p_{Li} = (x_{Li}, y_{Li}), \\ p_{Ri} = (x_{Ri}, y_{Ri}), \end{cases} \quad i = 1, \dots, n. \tag{1}$$

where p_{Li} and p_{Ri} represent object feature points in 2D image coordinates (x_i, y_i) in left and right images, respectively.

A. Robust Region-Based Feature Extraction

Region segmentation based on color is one of the mostly used object detection method in robotics. Its goal is to group pixels based on their color similarity. One natural way to segment color is through the HSI (*Hue*, *Saturation*, *Intensity*) color model [3]. In the hue image, color information is represented as pixel values belonging to the interval $[0, 2\pi]$, where the hue of a pixel corresponds to an angle *H*.

For controlling the color segmentation process, the cascade control structure from Fig. 5 has been proposed. The inner-loop from Fig. 5 is responsible for finding the optimal value of the saturation segmentation interval. On the other hand, the objective of the outer-loop is to find the optimal value of the hue thresholding interval, represented by the viewed object color class. The search for the optimal thresholding intervals is performed using an extremum seeking algorithm.

The closed-loops from Fig. 5 are controlled using a measure of region segmentation goodness. A region based segmented image is said to be of good quality if it contains all pixels of the object of interest forming a "full" (unbroken) and well shaped segmented object region. Bearing in mind the qualitative definition of a segmented image of good quality given above, the quantitative measure of segmented image quality in Eq. (2) has been proposed:

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

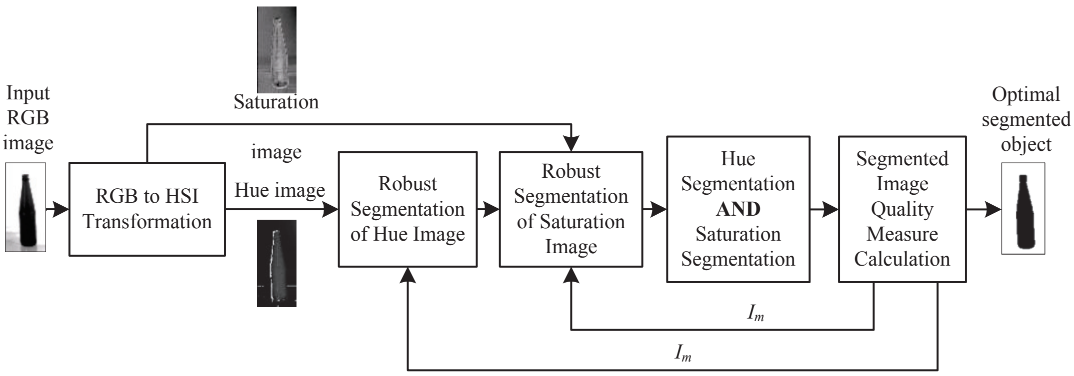


Fig. 5. Cascade closed-loop control structure for color segmentation.

where p_8 is the relative frequency, that is, the estimate of the probability of a segmented pixel to be surrounded with 8 segmented pixels in its 8-pixel neighborhood.

Robust Boundary-Based Object Detection

A second type of robust closed-loop object recognition method that we propose is an object boundary detection algorithm. In this paper, boundary segmentation is considered to be a combination of raw *Canny* edge detection and *Hough* transform.

In order to control the Canny and Hough thresholds, a measure of boundary segmentation quality had to be defined. This quality measure is to be used as a controlled variable in the proposed closed-loop system. A good boundary segmented image is one where the detected Hough lines lie on the real object's edges. Such a quality measure can be calculated at feature extraction level, where detected Hough lines are combined in order to form the objects candidate solutions vector $N_{\#}$. A solution is computed by combining parallel and perpendicular lines extracted from the Hough transform. The equation of the proposed measure is:

$$y = \begin{cases} e^{N/N_{max}} \cdot \sum_{n=0}^{N_{\#}} \frac{N_f(n)}{A}, & \text{if } N \leq N_{max}, \\ 0, & \text{if } N > N_{max}, \end{cases} \quad (3)$$

where N represents the total number of Hough lines, $N_{\#}$ the number of candidate solutions and $N_f(n)$ the number of foreground pixels covered by the Hough lines of the n^{th} object, normalized with the area of the image, A. Having in mind the computational burden of the Hough transform, the maximum number of lines allowed in an image is set to a constant value N_{max}. The exponential term in Eq. (3) is introduced in order to force feature extraction with a minimum amount of Hough lines. Hence, y decays to zero when the number of Hough lines increases.

In Fig. 6, the block diagram of the proposed closed-loop boundary segmentation method is displayed. In the presented system, the reference value of the chosen controlled variable is not explicitly known, since the goal is to develop a method able to detect objects independent of their sizes, color, or texture information. The objective of the control structure from Fig. 6 is to find the maximum value of the controlled variable y. This is achieved through a feedback optimization process using an extremum seeking algorithm. Optimal determined values of the Canny and Hough transform thresholds ensure that a reliable input is given to the feature extraction module which directly influences the precision of 3D object reconstruction.

IV. CONCLUSION

In this paper a machine vision system for service robotics was presented. The goal of the system is to provide reliable visual information needed to control the motions of the FRIEND robot. The reliability and robustness of the proposed ROVIS architecture lies in the closed-loop control of the involved image processing operation. The concept of feedback image processing was applied to both a region and a boundary based feature extraction methods. Due to the flexibility of its design, the proposed vision system can be easily transferred to different robotic systems.

REFERENCES

[1] M. Hillman, "Rehabilitation robotics from past to present – a historical perspective," in Proc. of the IEEE 8th Int. Conf. on Rehabilitation Robotics ICORR 2003, Apr. 2003.

[2] M. H. Van der Loos and D. J. Reinkensmeyer, "Rehabilitation and health care robotics," in Springer Handbook of Robotics (B. Siciliano and O. Khatib, eds.), Springer-Verlag Berlin Heidelberg, may 2007.

[3] M. H. Van der Loos, R. Mahoney, and C. Ammi, "Great expectations for rehabilitation mechatronics in the comming decade," in Advances in Rehabilitation Robotics, Lect. Notes (Z. Bien and D. Stefanov, eds.), pp. 427–433, Springer-Verlag Berlin Heidelberg, 2004.

[4] C. Martens, N. Ruchel, O. Lang, O. Ivlev, and A. Graeser, "A FRIEND for assisting handicapped people," IEEE Robotics and Automation Magazine, pp. 57–65, Mar. 2001.

[5] O. Prenzel, C. Martens, M. Cyriacks, C. Wang, and A. Graeser, "System controlled user interaction within the service robotic control architecture MASSiVE," Robotica, Special Issue, vol. 25, Mar. 2007.

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

[7] D. Kim, R. Lovelett, A. Behal, An empirical study with simulated adl tasks using vision-guided assistive robot arm, in: Proc. of the IEEE 11th Int. Conf. on Rehabilitation Robotics ICORR 2009, Kyoto, Japan, 2009.

[8] S.M. Grigorescu, *Robust Machine Vision for Service Robotics*, Ph.D. thesis, Shaker Verlag, Aachen, Germany, April 2010.

[9] C. Fragkopoulos and A. Graeser, "A RRT based path planning algorithm for rehabilitation robots," in *Proc. of the 41st ISR/Robotics 2010*, Munich, Germany, June, 2010.

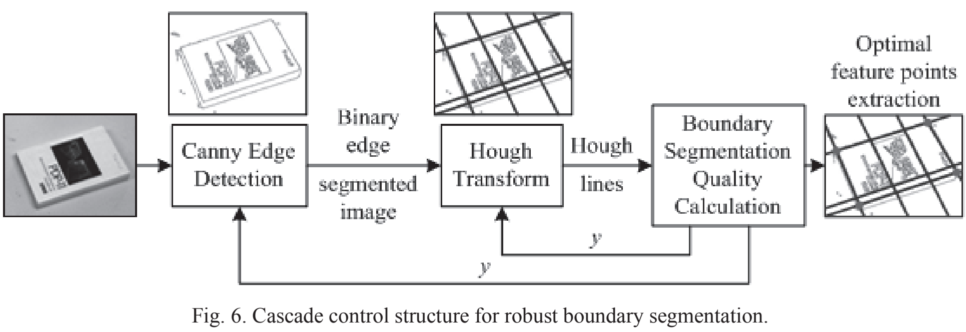


Fig. 6. Cascade control structure for robust boundary segmentation.