Software Module for Objects Shape Tracing and Recognition Before Gripping

Ionel Staretu [1] and Alexandru Itu [1]

[1] Product Design and Robotics Department ,Transilvania University of Brasov, Brasov, Romania [+]

Abstract. This paper presents the general structure, implementation features and a software module for tracing, visualization, shape recognition, measurement and efficiency evaluation for 3D model generation of objects to grip by an anthropomorphic gripper. Viewing is possible with an advanced video camera and shape recognition is possible through classifiers method. After setting the object shape, the gripper will approach the target, and its displacement is measured. Based on measurements of the distance between the initial point and a second intermediate point, you can do calculations aimed at object size and remaining distance to it. Implementation, and how to work with the software designed are possible in several steps that are specified and described. The performance evaluation of the software system is based on a series of external factors, the most important being: light conditions; quality of video device; capture size; the human factor through a series of experiments focused on the system's ability to cope with external factors above mentioned.

Keywords: image processing, shape recognition, software module, anthropomorphic gripping.

1. Introduction

The new perspective on the gripping strategy in this paper enables an innovative gripping process applicable to an anthropomorphic gripper that can easily grip different objects, similar to the human manner of gripping. In a first stage, the object is video detected, and thanks to visualization, an appropriate software module is able to recognize the type of the object, to measure it, and to generate its 3D model. After setting the object shape, the gripper will approach the target, and its displacement is measured. Based on measurements of the distance between the initial point and a second intermediate point, you can do calculations aimed at object size and remaining distance to it. This information is further transferred to the pre-configuration module. The shape of the object will determine the gripper pre-configuration and the size of the object is the gripper aperture. In normal situations the objects shape is often complex. For the research envisaged in this paper, we will start from four primitives that the visualization module will be able to recognize (parallelepipeds, spheres, cones and cylinders). Thus, we will provide a library for the four primitives, which can solve the problem of the gripper pre-configuration. This paper presents the particular software module RoboVISION, especially on the structure and implementation, specifically designed for tracing, shape recognition, and 3D model generation, for the object to grip, and an experiment, which evaluates the software module effectiveness in this object shape recognition, is described [1,2].

2. Robovision Software Module For Graspable Objects Image Processing

---
[+] Corresponding author. Tel.: + 40268412921; fax: +40268418967.
 E-mail address: istaretu@yahoo.com.

2.1. Structure and functions

RoboVISION stands for implementing a software module, tracing sequence, shape recognition, measurement, and 3D model generation of an object, in a gripping strategy structure destined to anthropomorphic grippers for robots[3,4,5]. This application has several tasks as seen in "Fig. 1". By fulfilling the above objectives, RoboVISION application allows as well a connection between the real environment and the virtual environment simulation of a gripping process. This task is achieved by generating a 3D model of the object to be gripped. This software module is based on Microsoft DirectShow technology to manage video devices and a series of image processing techniques for the objects recognition and measurement.
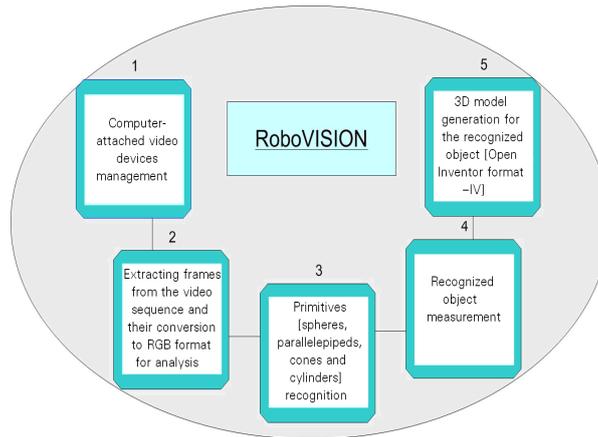

Fig. 1: RoboVISION application structure.

The general architecture of the classes used in RoboVISION application development is shown in "Fig. 2".
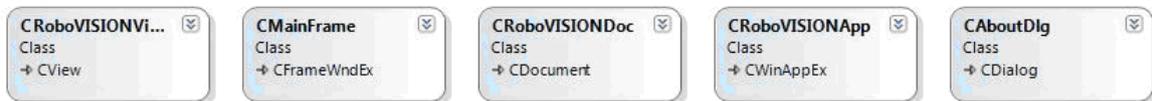

Fig. 2: UML chart of the main classes within RoboVISION.

In the UML (Unified Modeling Language) diagram of the fundamental classes developed within the application, three main issues are addressed: (1) object recognition by classifiers - CObjectDetector class, (2) CPrimitiveWriter class that generates the 3D model of the object detected and (3) interaction with video devices - CDShow class. The rest of the classes that make up the application are not presented, as RoboVISION is a standard MFC (Microsoft Foundation Classes) application, the basic structure of the application being generated by the tool available in Visual Studio 2008. In addition, it is noteworthy the latest WIMP paradigm using human-computer interaction. It is about the ribbon concept [6], which is already available in Office 2007 package and represents the latest paradigm used to create graphical user interfaces within the applications developed for Windows. The ribbon is an interactive extensive menu to activate/deactivate according to the option indicated by the user, through the mouse. The latest studies carried out on graphical user interfaces [7] have shown that interaction options layout in an interactive and accessible menu eases significantly the users' cognitive effort.

2.2. Implementation - Features

To extract a frame / image from a video sequence DirectShow technology provided by Microsoft within DirectX module was used, created especially for developing multimedia and 3D applications. DirectShow is framework software (known as API - Application Programming Interface) able to provide support aimed at multimedia applications. The core component is the filter which is a structure able to achieve a specific operation

with a multimedia file (conversion, display, compression, decoding) or to facilitate working with equipment registered through drivers in Windows (XP, VISTA, 7).

RoboVISION application multimedia functionality is based on four interrelated components, embedded in a filter. All four components of the filter are COM-based objects. COM is a technology developed by Microsoft that enables the use of software components made by a binary standard in applications developed in different programming languages. This filter will allow communication with any video device that has driver for Windows. In addition, under this component, the capture is set at a fixed resolution of 640x460 pixels. The second component is very important, being responsible to extract images captured from the video device. Its name, Sample Grabber, is very suggestive; it is responsible for taking video frames from the video device and sending them on for analysis. Once the analysis done, the video is delivered to the third component, which is responsible for colour space, YUY2 conversion from (description based on intensity) in space RGB32 (ARGB) needed for display. The next step, object shape recognition is extremely important in determining the gripper pre-configuration to generate the 3D model. During the step, the object is classified (sphere, parallelepiped, cylinder or cone) and its size in pixels is calculated. To solve this phase Open Inventor software module developed by Silicon Graphics Inc. (SGI) was selected. One of the main reasons why this format was chosen is that it allows a complete description of the 3D scene graphs with polygons rendered in the form of objects. In addition, having a standard description of virtual scene, interoperability between applications that have an Open Inventor files interpreter implemented can be achieved. Open Inventor standard enables primitives to be described through a class called Shapes. Thus, when the object was recognized and measured, by the results, the CPrimitiveWriter class developed in the application will create an Open Inventor file where it will add the Shape section and it will save this file to disk. Then the virtual simulation module will load the object saved by RoboVISION to the virtual scene and will use it for various tests and simulations.

3. Using Robovision Software - Performance Testing And Assessment

3.1. How to work with RoboVISION Application

To work with RoboVISION application the following steps are necessary: Open communication with the video device("Fig. 3,a"); Select the work area of the screen that plays the video sequence, captured by the video device, holding down the left button of the mouse, the selection being marked by a red rectangle ("Fig. 3,b"); Set the appropriate intensity and contrast using two dedicated sliders ("Fig. 3,c"); Choose the desired effect, being able to choose different combinations of processing, the final effect checked being the object detection ("Fig. 3,d");
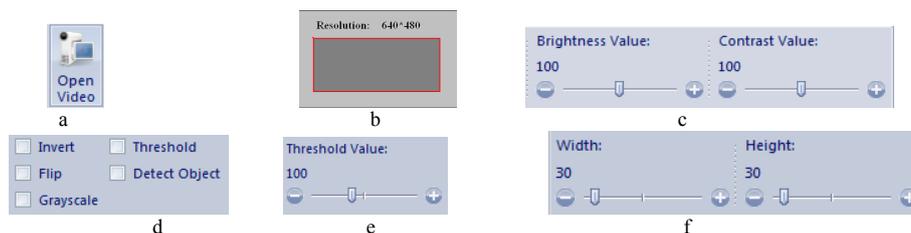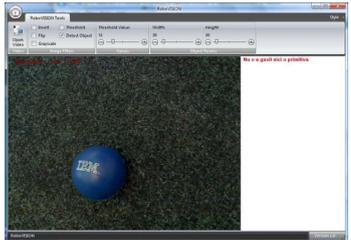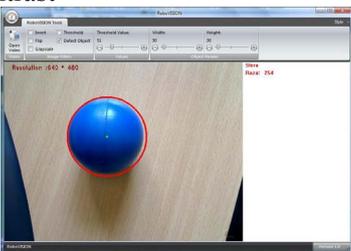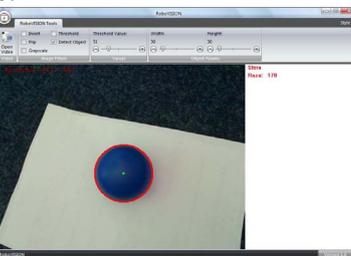

Fig. 3: RoboVISION application steps.

If the threshold effect is chosen (image binarization) the user can set the threshold value through a slider ("Fig. 3,e"); You can also opt for a filtering of recognized object through two dedicated sliders that change the minimum dimension (in pixels) of length and width of the rectangular area that contains the object ("Fig. 3,f"). When the object was detected in the right side will be written with the primitive recognized together with its size. The application will generate a directory called Detected Primitives where Open Inventor files will be saved.

3.2. RoboVISION Software Module Effectiveness Testing and Evaluation

The software system presented in this paper is responsible for the following operations: 1. recognition of the class to which the object belongs (sphere, parallelepiped, cylinder or cone); 2. measuring its size (the size of the object projection onto the surface where it is seated); 3. generating 3D model of the object recognized and measured. Evaluation of an advanced visualization system is a difficult task because the system itself is dependent on a number of external factors, the most important being: light conditions; video device quality; capture size; human factor. To evaluate RoboVISION application presented in this paper a series of experiments have been conducted aimed at the system's ability to cope with external factors mentioned above. The Experiment Concept: the aim of the experiment was to test the effectiveness of the application where light conditions vary greatly, and the background of the surface where the object is placed varies from the contrast point of view. Gripping strategies are aimed at attaching the video device which will be mounted on a robotic arm, thereby excluding the last external factor - the human factor. The experiment conducted aimed to assessing the recognition percentage of a sphere where the lighting conditions change abruptly. Thus all the frames captured from the video device were numbered and compared with the number of frames where the object was recognized, Table I.

TABLE I.   RoboVISION APPLICATION EFFICIENCY ANALYSIS

Background	Recognition percent
Low contrast	1% because of low contrast, recognition rate is very low regardless of lighting conditions.
Average contrast	46% average contrast causes large oscillations when strong fluctuations of light occur Brightness change also causes camera auto-focus activation and therefore the tendency is to stabilize the recognition after focusing.
High contrast	97%-for high contrast recognition rate is very high regardless of light conditions. However, if fluctuations are great camera auto-focus is activated which causes loss of information for 1-2 seconds as it takes to focus.

Analysis of Data and Results: The impact of variable light on the analysis leads to a more difficult object differentiation process from the background. By increasing the contrast is increased default rate of analysis accuracy. Thus, it was observed that best results are obtained when the background and the object to be recognized tend to opposite extremes of RGB colour space as observed in "Fig. 4".
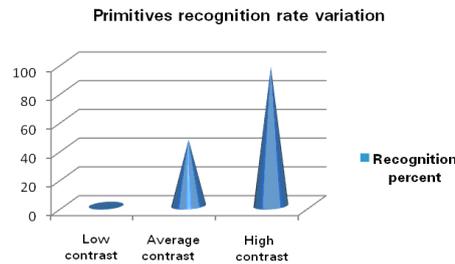
Fig. 4: Recognition percentage change by contrast.

4. Conclusions

RoboVISION is a complex application that consists of several components. Thus a software module specialized in the management of video devices attached to a computer interaction was created to provide access to every frame of the video sequence for analysis. To achieve this module, Microsoft DirectShow technology was used available in the DirectX package, created especially for multimedia and 3D. Video analysis is performed by a module specialized in the recognition of four primitives: spheres, parallelepiped, cones and cylinders. For object recognition, the algorithm based on classifiers was used. Another RoboVISION component is represented by software module specialized in objects measurement. To achieve it we used an algorithm based on the comparison of two video frames containing the object to be measured, taken from different distances. The last RoboVISION module application was created to generate the 3D model of the object recognized and measured. The 3D model is saved in an Open Inventor file that will subsequently be used in functional simulation application of the anthropomorphic gripper. RoboVISION application can interact with other software modules too because the result of the video analysis is available in the Open Inventor file generated. RoboVISION thus facilitates substantially the collection of 3D models used in functional simulation, thus simulation being strongly centred in reality.

5. Acknowledgements

6. References

[1]   A. Itu. *Contributions to Gripping Strategies in Real and Virtual Environment using a Three-Fingered Anthropomorphic Gripper*. Doctorate Thesis, Transilvania University of Brasov, Romania, 2010.

[2]   I. Staretu. *Gripping Systems*. Derc Publishing House, Tewksbury, Massachusetts, U.S.A., 2011.

[3]   C.V. Baysal and A.M. Erkmen. Preshape induced grasp controller for dextrous manipulation with a multifingered robot hand. *Proceedings of the IEEE International Conference on. 07/2004; Mechatronics, 2004*. ICM '04, Issue, 3-5 June 2004, pp. 78 – 83.

[4]   T. Marler, K. Farrell, J. Kim, S. Rahmatallah, and K. Abdel-Malek. Vision Performance Measures for Optimization-Based Posture Prediction. *Proceedings of the 2006 SAE DHM Conference*. Lyon, France, 2006.

[5]   L. D. Efrain. *Grasp planning for object manipulation by an autonomous robot*. Doctorate Thesis, Analysis and Architecture of the Systems Laboratory of the CNRS, France,2006.

[6]   www:http://www.microsoft.com/downloads/details.aspx?FamilyId=D466226B- 8DAB-445F-A7B4-448B326C48E7&displaylang=en.

[7]   www : http://www.cs.unc.edu/~geom/V_COLLIDE/.