Conference Proceedings

東京
Tokyo

ACEAIT
Annual Conference on Engineering and Information Technology

ISFAS
International Symposium on Fundamental and Applied Sciences

Computer Engineering and Technology I

704 **13:00~14:30** **Saturday, March 29**

Session Chair: *Prof. Chaur-Chin Chen*

ACEAIT-3036

Comparison of the Effects of Some Voip Codecs on the Bandwidth and Quality of Service

Sami Bashir Mohamed Abugharsa *Misurata University-Libya*
Ahlam Mahamed b *College Of Industrial Technology*

ACEAIT-2976

The Framework of Vision-Based Grasping for Manipulator using Object Features

Widodo Budiharto *Bina Nusantara University*
Bayu Kanigoro *Bina Nusantara University*
Anita Rahayu *Bina Nusantara University*

ACEAIT-3055

High-Capacity Steganography Using MRF-Synthesized Cover Images

Chaur-Chin Chen *National Tsing Hua University*
Wei-Ju Lai *National Tsing Hua University*

ACEAIT-2790

Developing Animated Pedagogical Agents in Java Concept Learning

Adib Sarkawi *Universiti Teknologi Mara*
Aiza Johari *Universiti Teknologi Mara*
Awang Rozaimi Awang Shuib *Universiti Teknologi Mara*

ACEAIT-2905

Real-time Acquisition for Microscopic Image Analysis

Thanaphong Wattanaphong *Thammasat University*
Chalie Charoenlarpnopparut *Thammasat University*
Tsuyoshi Isshiki *Tokyo Institute of Technology*

ACEAIT-2976
The Framework of Vision-Based Grasping for Manipulator using Object Features

Widodo Budiharto[a*], Bayu Kanigoro[a] and Anita Rahayu[a]
[a]School of Computer Science, Bina Nusantara University,
Jl. KH Syahdan no.9, Jakarta-Indonesia

*Corresponding Author: wbudiharto@binus.edu

Abstract

The ability for a manipulator to grasp an object based on vision is very important. Vision-based grasping using stereo vision is proposed in this paper in order a manipulator able to pick and put an object in a good manner. We propose a framework, fast algorithm and object features for an object and color marking at the gripper in order the system knows the position of the gripper and the object (pose estimation). Experimental result presented and we analyze the result.

Keyword: manipulator, stereo vision, grasping

1. Introduction

One of the most common of manipulation tasks is grasping an object. Tasks performed by humans involve some form of grasping action. In the absence of feedback, the grasping action cannot be completed effectively. A human being grasps an object almost invariably with the aid of vision. We use visual information to identify and locate the object, and then decide how to grasp them.

Most work in robotic manipulation assumes a known 3-D model of the object and the environment, and focuses on designing control and planning methods to achieve a successful and stable grasp in simulation environments. Grasping is usually preceded by a number of tasks that effect the final grasping action. The sequence of steps involved is:

1. The movement of the end-effector from a given position to within a reaching position from the object.
2. The estimation of grasp points and orientation of the end-effector to perform the grasp operation.
3. The grasping action, once the end effector is in the appropriate position.

Based on the previous literature (visual-servoing) is huge and largely unorganized. In image based visual servoing, 2D image measurements are used directly to estimate the desired movement of the robot. Typical tasks like tracking and positioning are performed by reducing the image distance error between a set of current and desired image features in the image plane. A variety of methods have been proposed to solve vision-based manipulation [1-5].

They use vision to aid just one of the above mentioned steps. In the past, most approaches to robotic grasping [6][7] assume availability of a complete 3-D model of the object to be grasped. In practice, however, such a model is often not available—the 3D models obtained from a stereo system are often noisy with many points missing, and 3-D models obtained from a laser system are very sparse. This makes grasping a hard problem in practice. In more general grasping, Kamon et al. [8] used Q-learning to control the arm to reach towards a spherical object to grasp it using a parallel plate gripper.

For grasping 2D planar objects, most prior work focuses on finding the location of the fingers given the object contour, which one can find quite reliably for uniformly colored planar objects lying on a uniformly colored table top. Using local visual features (based on the 2-d contour) and other properties such as force and form closure, the methods discussed below decide the 2D location at which to place the fingertips (two or three) to grasp the object.

Edsinger and Kemp [9] grasped cylindrical objects using a power grasp by using visual servoing and do not apply to grasping for general shapes. An inverse kinematic solver is proposed in [10] to find all joint angles for given position of the effectors on the manipulator.

The target object is recognized by color segmentation. The 3D position is computed by the stereo vision system after contour extraction.

Inverse kinematics of manipulator and object location are the key technology for arm robot. We study various visual feedback methods from previous literature and develop a new model for grasping an object. We know that Extraction of image information and control of a robot are two separate tasks where at first image processing is performed followed by the generation of a control sequence. A typical example is to recognize the object to be manipulated by matching image features to a model of the object and compute its pose relative to the camera (robot) coordinate system.

In this paper we develop a simple framework that uses stereo vision for all the steps that result in the grasping and movement action. First the target object is recognized by the vision system which than estimates the 3D pose of the object.
Based on this information, the controller coordinates to move the arm robot to grasp the object/bottle. The framework proposed in this experiment shown in fig. 1 below, where the stereo camera for pose estimation attached about 50cm at the side of manipulator.

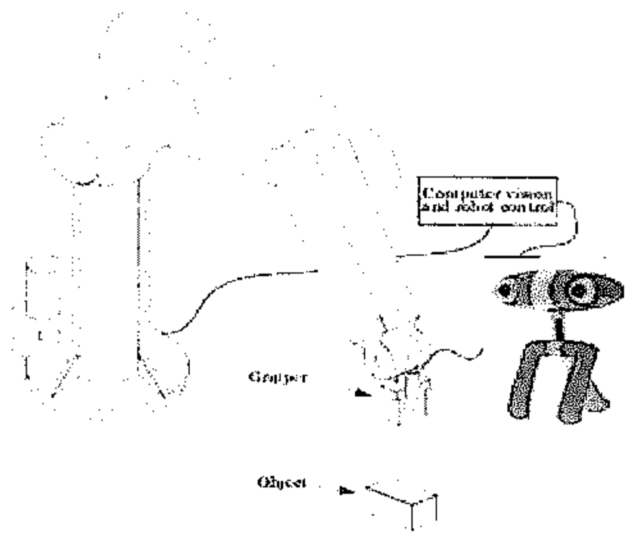

Figure 1. Framework for vision-based grasping manipulator using stereo vision

2. Framework of Vision-Based Manipulator

2.1 4 DOF Manipulator

We developed a framework of vision-based arm robot using 4 DOF (Degree of Freedom) arm robot from Lynxmotion that able to delivers fast, accurate, and repeatable movement. The robot features: base rotation, single plane shoulder, elbow, wrist motion, a functional gripper, and optional wrist rotate as shown in figure 2. This robotic arm is an affordable system with a time tested rock solid design that will last and last.

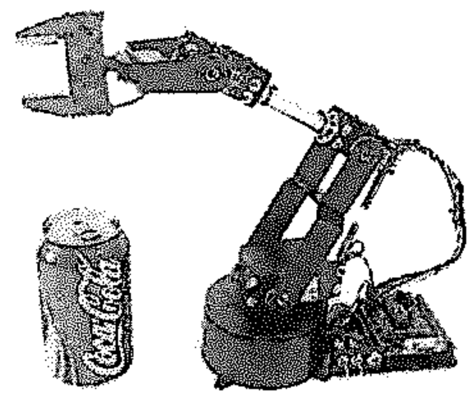

Figure 2. 4 DOF arm robot using stereo vision used in the experiment

The specification of this arm robot ;
Base: Height = 6.9 cm
Hand/Grip: Max Length (No Wrist Rotate) = 8.7 cm
Hand/Grip: Max Length (With LW Rotate) = 11.3 cm
Hand/Grip: Max Length (With HD Rotate) = 10.0 cm
Length: Forearm = 12.7 cm
Length: Arm = 11.9 cm

2.2 Stereo Vision

We have developed a system for object detection using Haar cascade classifier and depth estimation for measuring distance of object using stereo camera Minoru 3D. We use OpenCV 2.4.6[11] for image processing. OpenCV and a color filter algorithm are used to extract the specific color features of the object. Then, the 3D coordinates of the object to be grasped are derived by the stereo vision algorithm, and the coordinates are used to guide the robotic arm to the approximate location of the object using inverse kinematics. In the stereo imaging model, the tree-dimensional points in stereo camera frame are projected in the left and the right image frame [13]. On the contrary, using the projection of the points onto the left and right image frame, the three-dimensional points positions in stereo camera frame can be located. Figure 3 shows the stereo imaging model using the left front image frame *LF* and right front image frame *RF* [6].

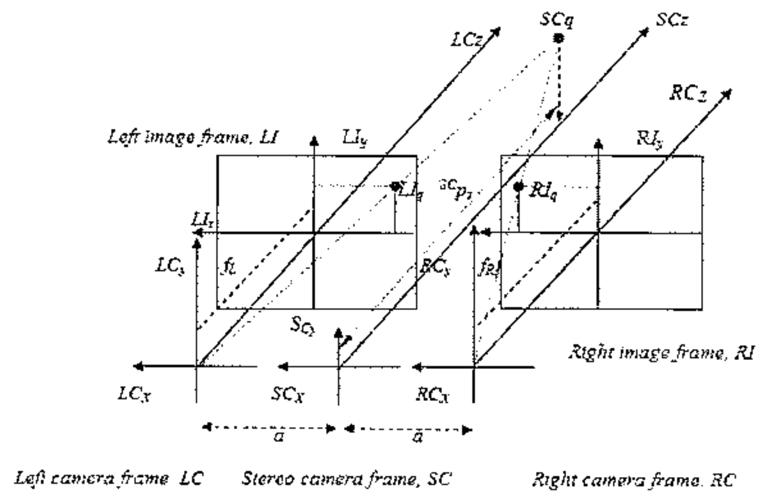

Figure 3. Stereo Imaging model[6]

By using stereo vision, we can obtain the position of object, then we can calculate and estimate the distance of the object. The three-dimensional point in stereo camera frame can be reconstructed using the two-dimensional projection of point in left front image frame and in right front image frame using formula:

$$ {}^{SC}\mathbf{q} = \begin{bmatrix} {}^{SC}q_x \\ {}^{SC}q_y \\ {}^{SC}q_z \end{bmatrix} = \frac{2}{{}^{RI}q_x - {}^{LI}q_x} \begin{bmatrix} \frac{1}{2}a({}^{RI}q_x + {}^{LI}q_x) \\ a\,{}^{RI}q_y \\ f\,a \end{bmatrix} \qquad (1) $$

Note that ${}^{LI}q_y = {}^{RI}q_y$

2.3 Grasping Model

Grasp determination is probably one of the most important questions from manipulation point of view. Usually, the object is of unknown shape. The purpose of control design is to make the robot respond in a predictable and desirable fashion to a set of input signals. In our model, we propose a simple way using pose estimation principle, if we know the distance of the bottle then move the arm to that position, then when the center point of a bottle exactly meet the center point of the gripper, then it means that is the time for grasping the bottle/object as shown in figure 3 :

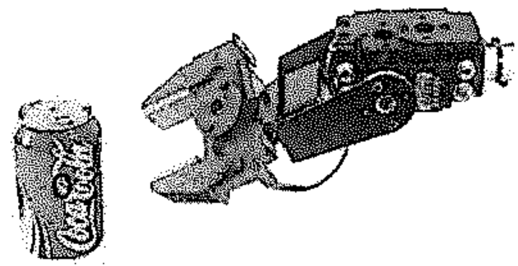

Figure 3. Finding the center of the bottle using object features (a) and color marking for indicating the position of gripper with an object (b)

The proposed algorithm to detect an object/bottle, grasp it and move to the destination is shown below:

Algorithm 1: vision-based grasping for manipulator

do
detect the object
if object/bottle detected then
begin
find position of the object
move arm robot to the object
move the gripper to the center of the object
if the position equal
grasp the object
else
move to gripper to the center of the object
end
endif
move the object to destination
go to the initial position
loop

An environment consists of a variety of objects, such as the robot itself, walls, floor, tables, objects to be grasped, etc. In order to successfully move the arm without hitting an obstacle, we provide 1 distance sensor at the gripper.

3. Experimental Result

This paper has studied the vision-based arm robot with the goal of this project is to build a general purpose household robot. To determine the state of the object, the "good" grasp position should be first determined. The experiment conducted at our lab to grasp and move a bottle to the destination. We use Lynxmotion RIOS (Robotic Arm Interactive Operating System) SSC-32 software to configure and control the arm robot as shown in figure 4 :

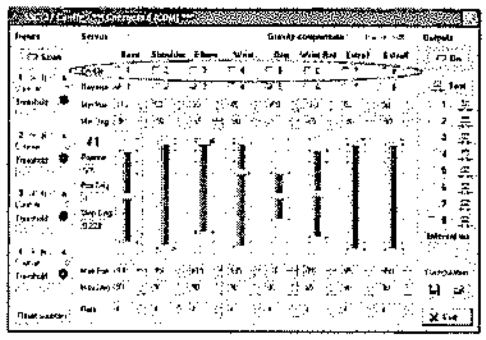

Figure 4. Configuring and control the board of arm robot using RIOS SSC-32 Software[12]

After configuring and calibrating the servos of manipulator. We put an object/bottle in front of the arm robot to be grasped and move to other position. Program for object detection of an object successfully developed with OpenCV and the manipulator able to grasp it.

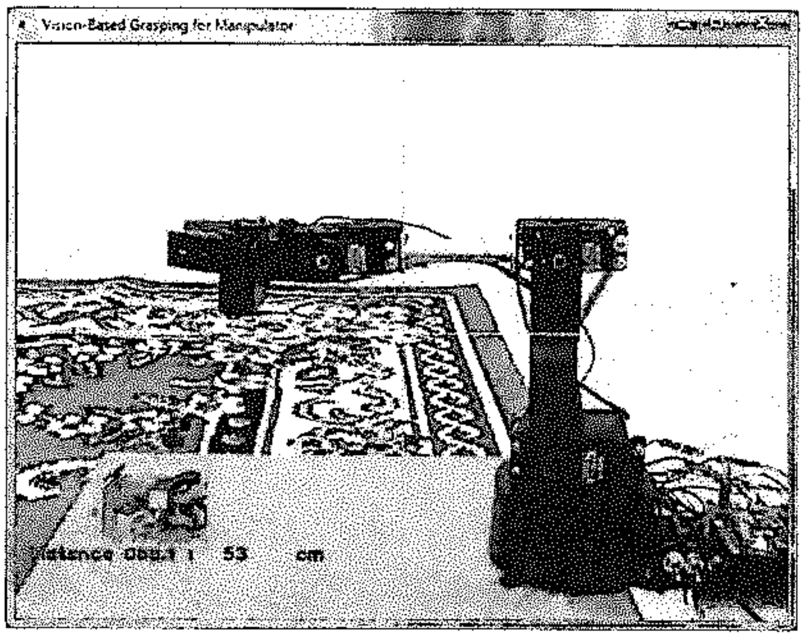

Figure 5. Object detection using OpenCV for grasping

Based on our experiment we get the expected result as shown in table 1.

Table 1. Result of detecting and grasping an object in 10x

No	*Action*	*Success*	*failure*
1	Identify the object as bottle	100%	0%
2	Grasping an object correctly	90%	10%
3	Estimate the distance of the object/bottle	90%	10%

Based on table 1, shows that with the action identify the object as bottle, percentage for success is 100% and percentage for failure is 0%. With the action grasping an object correctly and estimate the distance of the bottle, percentage for success is 90% and percentage for failure is 10%. The accuracy and robustness of the system and the algorithm were tested and proven to be effective in real scenarios.

4. Conclusion

We have proposed a model for vision-based grasping for manipulator using stereo vision. The position of the object estimated using stereo vision and color marking at the gripper used as indicator for the system to deciding grasp the object. Based on the experiment we see that our model able to be implemented for vision-based arm robot.

5. Acknowledgment

We would like to say thank you to the BINUS University for providing the lab for conduct this experiment.

References

[1] Bendiksen A and Hager G. D., 1994. *A vision-based grasping system for unfamiliar planar object*, . In ICRA, 2844–2849.

[2] Christensen H. I. and Corke P. L., 2003. *Visual Servoing*, International J.Robotic Res., 22(10-11):779–780.

[3] Kragic D., and Christensen. I., 2003. *A framework for visual servoing*, In ICVS, pages 345–354.

[4] Bicchi A. and V. Kumar V., 2000. *Robotic grasping and contact: a review*, in ICRA.

[5] Jiao J., 2012. *Embedded Vision-Based Autonomous Move-to-Grasp Approach for a Mobile Manipulator*, International Journal of Advanced Robotics System, vol. 9, 1-6.

[6] K. Shimoga, 1996. *Robot grasp synthesis: a survey*, IJRR, vol. 15, pp. 230–266.

[7] Purwanto. D., 2001. *Visual Feedback Control in Multi-Degrees-of-Freedom Motion System*. PhD thesis at Graduate School of Science and Technology - Keio University, Japan.

[8] Kamon I., Flash T., and Edelman S., 1996. *Learning to grasp using visual information*, in ICRA.

[9] Edsinger A. and Kemp C., 2006. *Manipulation in human environments*, in IEEE/RAS Int'l Conf on Humanoid Robotics (Humanoids06).

[10] Y. Yang, 2011. *Binocular Stereo Vision based Humanoid Manipulator Control*, 30th International Conference on Control, 3996 - 4001, China.

[11] OpenCV.org

[12] Lynxmotion.com

[13] Samarth B., 2013. *Practical OpenCV*, Technology in Action Publisher.