TITLE: Vision-Based Intelligent Robots

TITLE: Input/Output and Imaging Technologies II. Taipei, Taiwan, 26-27 July 2000

Vision-Based Intelligent Robots

Minh-Chinh Nguyen

Institute of Measurement Science, LRT6
Federal Armed Forces University Munich
85577 Neubiberg - Germany
E-mail: Minh.Chinh@Unibw-muenchen.de

ABSTRACT

Vision is an ideal sensor modality for intelligent robots. It provides rich information on the environment as required for recognizing objects and understanding situation in real time. Moreover, vision-guided robots may be intelligent and largely calibration-free, which is a great practical advantage. Together with it, a new concept for intelligent robot control, that enables realization of the calibration-free visual robots, is introduced.

Keywords: Calibration-Free Robots, Vision-Guided Intelligent Robots, Robot Vision, Situation-Oriented and Behavior-Based Robot Control.

1. INTRODUCTION

Industrial robots are of great economic and technological importance. Until 1996 approximately 860,000 robots had been installed worldwide. At that time 680,000 of them were still being used, for the most part in automobile and metal-manufacturing [IFR. 1997].Typical applications include welding card, spraying paint on appliances, assembling printed circuit boards, loading and unloading machines and placing cartons on a pallet. Experts estimate that by the year 2000 about 950.000 industrial robots will be employed word-wide.

Although present robots contribute very much to the prosperity of the industrialized countries they are quite different from the robots that researchers have in mind when they talk about "intelligent robots". Today's robots are not creative or innovative, do not think independently, do not make complicated decision, do not learn from mistakes and do not adapt quickly to changes in their surroundings. They rely on detailed teaching and programming and carefully prepared environments. It is costly to maintain them and it is difficult to adapt their programming to slightly hanged environmental conditions or modified tasks.

Although the vast majority of robots today are used in factories, advances in technology are enabling robots to automate many tasks in non-manufacturing industries such as agriculture, construction, health care, retailing and other services. These so-called "field and service robots" aim at the fast growing service sector and promise to be a key product for the next decades.

From a technical point of view service robots are intermediate steps towards a much higher goal: "personal robots" that will be as indispensable and ubiquitous as personal computers today. Personal robots must operate in varying and unstructured environments without needing maintenance or programming. They must cooperate and coexist with humans who are not trained to cooperate with robots and who are not necessarily interested in them. Advanced safety concepts will be as indispensable as intelligent communication abilities, learning capabilities, and reliability. It will be a long way of research to achieve this goal, but undoubtedly vision - the most powerful sensor modality known - will enable these robots to perceive their environments, to understand complex situation and to behave intelligently.

This paper present some of the underlying concept and principle that were key to the design of our research robots. It is organized as follows: in the next chapter will be briefly described the vision and its potential for robots. The third chapter will describe the new concept for intelligent robot control. The experiments and results as well as conclusions will be discussed in the fourth and fifth chapter respectively.

In *Input/Output and Imaging Technologies II*, Yung-Sheng Liu, Thomas S. Huang, Editors,
Proceedings of SPIE Vol. 4080 (2000) • 0277-786X/00/$15.00

2. VISION AND ITS POTENTIAL FOR ROBOTS

2. 1. Advantages of the visual sensors and a conceptual structure of robot's vision systems

When a human drives a vehicle he depends mostly on his eyes for perceiving the environment. He uses his sense of vision not only for locating the path to be traversed and for judging its condition, but also for detecting and classifying external object, such as other vehicles or obstacles, and for estimating their state of motion. Entire situations may thus be recognized, and expectations, as to their further development in the "foreseeable" future, may be formed.

The same is true for almost all animals. With the exception of those species adapted to living in very dark environments, they use vision as the main sensing modality for controlling their motions. Observing animals, for instance, when they are pursuing prey or trying to escape a predator, may give an impression of the performance of organic vision system for motion control.

In some modern factory and office buildings mobile robots are operating, but almost all of them are blind. Their sensors are far from adequate for supplying all the information necessary for understanding a situation. Some of them have only magnetic or simple optical sensors, allowing them merely to follow an appropriately marked track. They will fail whenever they encounter an obstacle and they are typically unable to recover from a condition of having lost their track. The lack of adequate sensory information is an important cause making these robots move in a comparatively clumsy way and restricting their operation to the simplest of situations.

Other mobile robots are equipped with sonar systems. Sonar can, in principle, be a basis for powerful sensing systems, as evidenced by certain animals, such as bats or dolphins. But the sonar systems used for mobile robots are usually rather simple ones, their simplicity and low cost being the very reason for choosing sonar as a sensing modality. It is then not surprising that such system are severely limited in their performance by low resolution, specular reflections, insufficient dynamic range, and other effects.

Nevertheless, even when comparing the most highly developed organic sonar systems with organic vision systems, it is obvious that in all environments where vision is physically possible animals endowed with a sense of vision have, in the course of evolution, prevailed over those that depend on sonar. This may be taken as an indication that vision has, in principle, a greater potential for sensing the environment than sonar. Likewise, it may be expected that advanced robots of the future will also rely primarily on vision for perceiving their environment, unless they are intended to operate in other environments, e.g. under water, where vision is not feasible.

One apparent difficulty in implementing vision as a sensor modality for robots is the huge amount of data generated by a video camera: about 10 million pixels per second, depending on the video system used. Nevertheless, it has been shown (e.g., by [Graefe 1989]) that modest computational resources are sufficient for realizing real-time vision systems if a suitable system architecture is implemented.

As a key idea for the design of efficient robot vision systems the concept of object-oriented vision was proposed. It is based on the observation that both the knowledge representation and the data fusion processes in a vision system may be structured according to the visible and relevant external objects in the environment of the robot (Figure 1). For each object that is relevant for the operation of the robot at a particular moment the system has one separate "object process". An object process receives image data from the video section (camera, digitizers, video bus etc.) and generates and updates continuously a description of its assigned physical object. This description emerges from a hierarchically structured data fusion process which begins with the extraction of elementary features, such as edges, corners and textures, from the relevant image parts and ends with matching a 2-D model to the group of features, thus identifying the object.

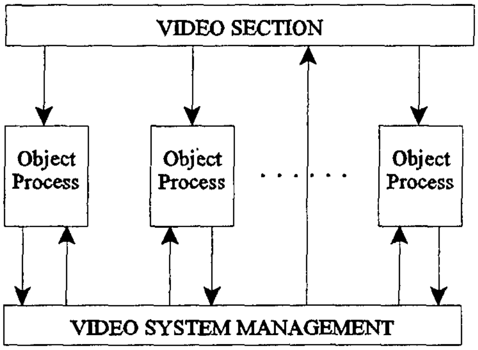

Figure 1: Conceptual structure of object-oriented robot vision system

This concept is practical because it was found that in any given moment only a small number of objects are relevant and that, consequently, only a small number of processes need to be active simultaneously. In the next moment, however, different objects may be relevant; therefore, the ability to switch the system's focus of attention quickly is crucial. The switching of attention and the control of the cameras is performed by a vision system management process that dynamically generates appropriate object processes upon request.

The potential of object-oriented vision systems was first demonstrated in high-speed autonomous highway driving applications [Graefe, Kuhnert 1988], [Graefe 1992]. Later the same concept has proved its value in mobile and stationary indoor robots.

2. 2. Perception

Model-based robot control depends on a continuous flow of numerical values describing the current state of the robot and its environment. These values are derived from measurements performed by the robot's sensors. One problem here is that the quantities that are needed for updating the numerical models may be difficult to measure, e.g., the distance, mass and velocity of some external object that is posing a collision danger. Also, there are certain important decisions that cannot be made on the basis of measurements alone; the hypothetical decision whether in a particular situation a collision with a parked car should be brought about in order to avoid a collision with a pedestrian is an example.

Humans and other organisms, on the other hand, do not depend on measurements for controlling their motions. If, for instance, we want to sit down on a chair or pass through an open door, we do not first measure the size of the chair, the door, or our body; rather, we make a qualitative judgement whether the chair is high or low, or whether the door is wide or narrow, and then execute a sequence of motions that is adequate for the situation. In short, we substitute perception for measurement.

According to Webster's Dictionary "perception" is:

- a result of perceiving;
- reaction to sensory stimulus;
- direct or intuitive recognition;
- the integration of sensory impressions of events in the external world by a conscious organism;
- awareness of the elements of the environment.

"To perceive" means, according to the same source,

- to become aware of something through the senses;
- to become conscious of something;
- to create a mental image;
- to recognize or identify something, especially as a basis for, or as recognized by, action.

Typical questions to be answered by perception are:

- Which objects exits?
- What is the relationship between objects?
- Is it necessary to react? How?

Perception, rather than measurement, is thus a prerequisite for, and a complement of, situation assessment. Vision is the ideal sensing modality for perception because it is capable of supplying very rich information on the environment.

The actual design and implementation of a behavior pattern and of related perceptual processes depend on the robot's environment and task. A mobile robot navigating in a network of passageways needs different behaviors and recognition modules than a walking robot intended to explore rough terrain.

However, advantage sensor system will be got their efficiency fully, if and only if they are combined with a sensible control concept. In the sequel we will, so that, represent a new concept of "behavior-based and situation-oriented robot control"for intelligent vision-guided robot control.

3. CONCEPT OF SITUATION-ORIENTED AND BEHAVIOR-BASED VISUAL ROBOT CONTROL

3.1. Behavior

Biological behaviors could be defined as any thing that an organism does involving action and response to stimulation, or as the response of an individual, group, or species to its environment. Behavior-based robotics has become a very popular field in robotics research because biology proves that even the simplest creatures are capable of intelligent behavior: They survive in the real world and compete or cooperate successfully with other beings. Why should it not be possible to endow robots with such an intelligence? By studying animals behavior, particularly their underlying neuroscientific, psychological and ethological concepts, robotic researchers have been enabled to build intelligent behavior-based robots according to the following principles:

- complex behaviors are combinations of simple ones, complex actions emerge from interacting with the real world
- behaviors are selected by arbitration or fusion mechanisms from a repertoire of (competing) behaviors
- behaviors should be tuned to fit the requirements of a particular environment and task
- perception should be actively controlled according to the actual situation

Many system architecture and control methods, which ware introduced in recent years, interest in realizing of behavior-based robots. Its main characteristics are active perception of the robot's dynamically changing environment, recognition and evaluation of its current situation, and dynamic selection of behaviors appropriate for the actual situation. Animals simplest capabilities, i.e., to perceive and act within an environment in a meaningful and purposive manner, can thus be imitated by our robots to a certain degree.

3. 2. Situation Assessment

According to the classical approach, robot control is model-based. Numerical models of the kinematics and dynamics of the robot and of the external object that the robot should interact with, as well as quantitative sensor models, are the basis for controlling the robot's motions. The main advantage of model-based control is that it lends itself to the application of classical control theory and, thus, may be considered a straight-forward approach. The weak point of the approach is that it breaks down when there is no accurate quantitative agreement between reality and the models. Differences between models and reality may come about easily; an error in one of the many coefficients that are part of the numerical models suffices. Among the many possible causes for discrepancies are initial calibration errors, aging of components, changes of environmental conditions, such as temperature, humidity, electromagnetic fields or illumination, maintenance work and replacement of components, to mention only a few. Consequently, most robots work only in carefully controlled environments and need frequent recalibrations, in addition to a cumbersome and expensive initial calibration.

Organisms, on the other hand, are robust and adapt easily to changes of their own conditions and of the environment. They never need any calibration, and they normally do not know the values of any parameters related to the characteristics of their "sensors" or "actuators". Obviously, they do not suffer from the shortcomings of models-based control which leads us to the assumption that they use something other than numerical models for controlling their motions. Perhaps their motion control is based on a holistic assessment of situation and the selection of behaviors to be executed on that basis, and perhaps robotics could benefit from following a similar approach.

According to Webster's Third New International Dictionary [Babcock 1976] the term "situation" describes among others "the way in which something is placed in relation to its surroundings", a "state" , a "relative position or combination of circumstances at a given moment" or "the sum of total internal and external stimuli that act upon an organism within a given time interval". We define the term "situation" in a similar way, but with a more operational aim, as the set of all decisive factors that should ideally be considered by the robot in selecting the correct behavior pattern at given moment. These decisive factors are:

- perceivable objects in the environment of the robot and their suspected or recognized states;
- the state of the robot (state of motion, presently executed behavior pattern, ...);
- the goals of the robot, i.e., permanent goals (survival, obstacle avoidance) and transient goals emerging from the actual mission description (destination, corridor to be used, ...);

- the static characteristics of the environment, even if they cannot be perceived by the robots's sensors at the given moment;
- the repertoire of available behaviors and knowledge of the robot's abilities to change the present situation in a desired way by executing appropriate behavior patterns.

Figure 2 illustrates the definition of the term "situation" by embedding it in the action- perception loop of a behavior-based and situation-oriented robot. The actions of the robot change the state of the environment, and some of these changes are perceived by the robot's sensors. After assessing the situation an appropriate behavior is selected and executed, thus closing the loop. The role of a human operator is to define external goals via a man machine interface and to control behavior selection, e. g., during supervised learning.

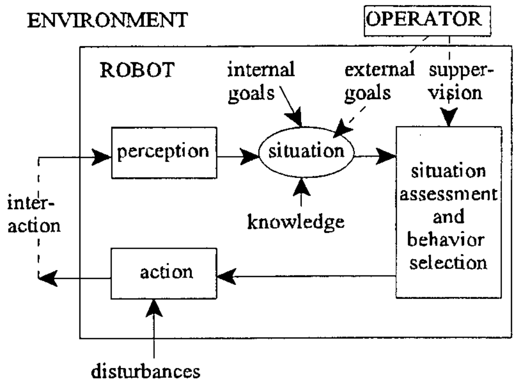

Figure 2 :
The role of "situation" as a key concept in the perception action loop of a situation- oriented behavior-based robot.

Although situation-oriented robot control has proven much more robust and flexible under real-world conditions than classical model-based control it is not perfect. One reason is that, obviously, the robot cannot base its behavior selection on a "true" or "real" situation, but only on an internal image of the situation as created by the robot according to its sensor information and its - always imperfect - knowledge of the world and of its own characteristics. Also, disturbances during the behavior execution can lead to non-expected situations. Although the disturbances may be corrected by either adjusting behavior-immanent parameters or selecting a different behavior, they will usually cause the robot to move in a non-ideal way.

4. IMPLEMENTATION

The described concepts were implemented on the calibration-free vision-guide manipulator Mitsubishi Movemaster RV-M2 with 5 degree of freedom (Figure 3) for grasping objects of various shapes (Figure 4). It eliminates the need for a calibration of the robot and of the vision system, it uses no world coordinates, and it comprises an automatic adaptation to changing parameters. The concept is based on the utilization of laws of projective geometry that always apply, regardless of camera characteristics, and on machine learning for the acquisition of knowledge regarding system parameters. Different forms of learning and knowledge representation have been studied, allowing either the rapid adaptation to changes of the system parameters or the gradual improvement of skills by an accumulation of learned knowledge.

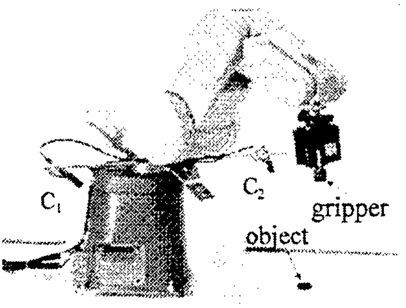

Figure 3: External view of the Movemaster RV-M2 with mounted cameras.

The images from the two cameras are processed by an object-oriented vision system described in 2.1 above, which consists of two frame grabbers, each containing a TMS320C40 Digital Signal Processor.

The situation process receives and assesses the information about the position and orientation of gripper and of object to be grasped to decide which behaviors of the robot [Nguyen 1999] will be used to achieve the grasp, and to generate appropriate motion control commands.

5. EXPERIMENTS AND RESULTS

The described concepts has been evaluated in a series of real-world experiments.

Objects of various shapes were successfully grasped. It requires no knowledge regarding:

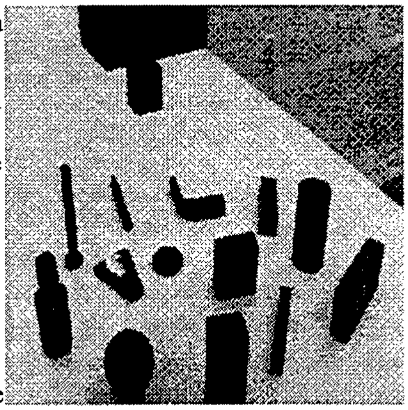

Figure 4: Objects used in our experiments

- The parameters of the robot arm
- The internal camera parameters, i.e., optical characteristics
- The exact locations of the cameras
  (except that the cameras should be located some distance away from the work plane of the robot in an opposite arrangement)
- The exact viewing directions of the cameras
  (except that both cameras should have the actual work space of the robot in their fields of view)
- The dimensions, kinematics, and joint angles of the robot
  (except that, for practical reasons, we presently assume that the robot is of an articulated arm type, and that the general type of the gripper and the number of degrees of freedom of the system are known)
- The quantitative relationships between the control words sent to the motor controllers and the resulting motions
  (except that these relationships are assumed to be "smooth")
- The surrounding environment, e.g., lighting, surrounding landmarks
  (except that it should be within reason)

In addition, even severe disturbances, such as arbitrary changes of the cameras' orientations, that would make other robots fail, are tolerated while our robot is operating.

We state that the concepts proposed in this work will be especially valuable for mobile and service robots operating in unstructured environments.

6. CONCLUSIONS

A fundamental concepts and principles for realization of intelligent robots have been presented. We strongly believe that vision - the sensor modality that predominates in nature - is also an eminently useful and practical sensor modality for intelligent robots. It provides rich and timely information on the environment and allows real-time recognition of dynamically changing situations. Situation-dependent perception and behavior selection rather than measurement and control based on quantitatively correct models are additional key factors for advanced robots. Motor control commands should be derived directly from sensor data, without using world coordinates or parameter-dependent computations, such as inverse perspective or kinematic transforms.

Building robots according to these rules and testing them intensively in the real world lead to robust and intelligent robots with the ability to adapt themselves to modified environmental conditions and tasks.

REFERENCES

1. P. Babcock, *Webster's Third New International Dictionary of the English language*, G. & C. Merriam Company, Springfield, MA, USA, 1976.
2. J. R. Cooperstock, E. E. Milios, "Self-supervised learning for docking and target reaching,"*Robotics and Autonomous Systems 11*, pp 243-260, 1993.
3. V. Graefe, K.-D. Kuhnert, "Towards a Vision-based Robot with a Driver's License,"*Proc. IEEE Int. Workshop on Intelligent Robots and System*, IRO "88. Tokyo, pp. 627 - 632, 1988.
4. V. Graefe, "Dynamic Vision Systems for Autonomous Mobile Robots,"*Proc. IEEE/RSJ International Workshop on Intelligent Robots and Systems*, IROS "89, Tsukuba, pp. 12-23, 1989.
5. V. Graefe, "Visual Recognition of Traffic Situations by a Robot Car Driver,"*Proceedings, 25th ISATA ; Conference on Mechatronics*, Florence, pp 439 - 446, 1992.
6. V. Graefe, Q.-H. Ta, "An Approach to self-learning Manipulator Control Based on Vision,"*Proc. International Symposium on Measurement and Control in Robotics*, pp 409-414, Smolenice, 1995.

7. IFR International Federation of Robotics “Key data for the world robot market,”*Http://www.ifr.org/stat.htm.* 1997.

8. I. Kamon, T. Flash, S. Edelman, "Learning to Grasp Using Visual Information," *Proc. IEEE International conference on Robotics and Automation,* vol. 2, pp 2470-2476, 1996.

9. M.-C. Nguyen, V. Graefe, "Object Manipulation Controlled by Uncalibrated Stereo Vision". *The Second Chinese Congress on Intelligent Control and Intelligent Automation; Proceeding of the CWCICIA'97*, Vol. 1, Xian-China, pp. 77-83, 1997.

10. M.-C. Nguyen, "Situation-oriented and Behavior-based Stereo Vision to Gain Robustness and Adaptation in Manipulator Control," *In D. Casasent (ed.): Intelligent Robots and Computer Vision XVIII:Algorithms, Techniques, and Active Vision, Proceedings of the SPIE,* Vol. 3837, pp. 90-97, Boston, USA, 1999.

11. K. Vollmann, M.-C. Nguyen, "Manipulator Control by Calibration-Free Stereo Vision". *In D. Casasent (ed.): Intelligent Robots and Computer Vision XV, Proceedings of the SPIE,* Vol. 2904. Boston-USA, pp. 218-226, 1996.