# Control Based on Image Information

Nielsen, Lars

1986

CODEN: LUTFD2/(TFRT-3182)/1-16/(1986)

# Control Based on Image Information

Lars Nielsen

Department of Automatic Control
Lund Institute of Technology
November 1986

**Department of Automatic Control**
**Lund Institute of Technology**
P.O. Box 118
S-221 00 Lund Sweden

*Document name*
Report

*Date of issue*
November 1986

*Document Number*
CODEN: LUTFD2/(TFRT-3182)/1-16/(1986)

*Author(s)*
Lars Nielsen

*Supervisor*

*Sponsoring organisation*
The Swedish Board for
Technological Development (STU)

*Title and subtitle*
Control Based on Image Information

*Abstract*

Control based on image information offers great possibilities for advanced automation both in motion control and in process control. Semi-insulating GaAs-wafers can be characterized using image analysis. The ash-line position is important for efficiency in firing in bark ovens. Algorithms that can operate in the speed required by the process has been developed for ash-line detection. Automatic guided vehicles in warehouses can be controlled from a camera viewing down on the floor, instead of being controlled by magnetic slings in the floor. The programming of automatic guided vehicles or other robots can be simplified by combining real images and computer graphics in the man-robot interface. The discovery of area-invariants under perspective projection extends the capability of existing special purpose image processing systems based on contour descriptions. A new theory for optimal digitization of images can be used as a rule of thumb, for adaptation, or for structuring of special image hardware.

*Key words*

*Classification system and/or index terms (if any)*

*Supplementary bibliographical information*

*ISSN and key title*

*ISBN*

*Language*
English

*Number of pages*
16

*Recipient's notes*

*Security classification*

71 %

# Control Based on Image Information

Final Report STU

82-3429, 83-3562, 84-3396, 85-3680

Lars Nielsen

Department of Automatic Control
Lund Institute of Technology
November 1986

Resten av sidorna
(= sid 1 t.o.m. 16)
ner till 83 %

# 1. I

Control based on image information offers great possibilities for advanced automation. The applications include both motion control and control in the process industry. Motion control is mainly concerned with robotics including manipulator arms and intelligent vehicles, but control of simpler devices like overhead cranes is also of interest. In the process industry other problems can be attacked due to slower time scales. Automatic supervision systems can provide measurements improving performance either in closed loop control or in open loop control. The development of image technology is significant and exciting and has a major impact into the area of visual servoing. There is also a current merging of the research fields of image understanding, signal processing, automatic control, and artificial intelligence.

The project "Control Based on Image Information" ("Reglering baserad på bildinformation") sponsored by the Swedish Board for Technological Development (STU-projekt) has been active from July 1, 1982, until June 30, 1986, under the project numbers STU-82-3429, 83-3562, 84-3396, 85-3680. The work has, as will be indicated, been performed in collaboration with the staff at Department of Automatic Control, other departments at Lund University, Master thesis students (examensarbetare), visiting researchers, industry, and the Department of Computer Science at Caltech, Pasadena, USA. In this report the different activities performed within the program are briefly described in Chapter 2. The conclusions are given in Chapter 3 and future work is suggested in Chapter 4. The publication list in Chapter 5 gives references to detailed descriptions of the work performed.

# 2. Projects

The theory of visual servoing has not developed far enough so that the closed loop performance can be judged only by analysis of theoretical models or simulations. A prototype system therefore needs to be built, and significant experiments with the system must be performed. In our project the philosophy has been to use simple robots and develop a flexible laboratory at a reasonable cost.

## 2.1 Experimental platform

The laboratory represents the first effort in vision control at Lund University. The project has included development of both hardware and software. A standard computer has been extended with commercially available instrumentation. The image system is built around a raster image memory interfaced to a VAX-11/780. The raster image memory has been built from plug-in boards manufactured by Matrox. The image memory uses standard video-signals to external video equipment such as video-cameras, video-recorders, or monitors. The connection to the computer is via a general bus. Two joysticks and a mouse are used for interaction. The video camera used is a black and white Intensa GPC-25 camera equipped with a high sensitivity vidicon combined with automatic aperture control, which solves the problem of getting good visual quality of the image data, when the illumination changes. A Barco CD 33 HR RGB monitor is used to present the video output. There is also a home video set with a video tape recorder, a color video camera, and a TV. The color video camera has lower quality than the GPC camera. The video set provides, however, an easy way to handle image information from other places via the exchange of video tapes.

Any process may be interfaced via the AD and DA converters. In our experiments two vehicles moving on the floor are used. One vehicle is a commercially available toy Turtle. It has two separately driven wheels individually controlled using on/off signals transferred via a cord. The Turtle has touch-sensors. The control system of the Turtle may thus combine image and touch information. The other vehicle is the Ilon car, which has four Ilon wheels, allowing complete control of the motion in the plane via control of the rotation of the four wheel motors. The Ilon car may thus be moved straight forward, straight sideways or be rotated on the spot.

The software is written in Pascal in a highly modular way. Related types, variables, and procedures are grouped together in packages. Machine dependent parts are isolated to improve portability. A preprocessor is used to produce a standard Pascal program from the packages. There are a number of basic support packages available. They contain sufficient operations to build user programs without knowing the explicit organization of the hardware. The operations include video control commands, image definitions and in/out operations, color look-up table handling, AD and DA conversion, and file handling to save and restore images.

The experimental equipment has proven comfortable and easy to handle. Important at the time (1982) was that it proved it neither necessary to build special hardware nor to make complicated file handling to deal with images. The virtual memory of the Vax computer automatically solves the problem of addressing large amounts of data in software. The speed is limited by the capability of the Vax, but all aspects of the visual servo problem may in principle be demonstrated (Nielsen and Elmqvist, 1983; Nielsen, 1985, 1986).

## 2.2 Automatic guided vehicles

One particularly interesting visual servo application is automated guided vehicles (AGV). AGVs are used in warehouses as part of a flexible manufacturing system. Currently these vehicles are controlled by magnetic trails in the floor, and the paths are essentially fixed due to the cost of rearranging the magnetic trails. A much more flexible system can be developed using a visual servo with a camera on a wall or ceiling supervising the AGV. Of course hybrid systems with visual servoing in specific areas and magnetic trails for traveling in between can be used in warehouses where flexibility is not needed everywhere.

A major experimental contribution in the project has been to develop an experiment where the Turtle simulates an AGV working on the floor in a warehouse (Johansson, 1984; Nielsen, 1984a, 1985, 1986; Nielsen and Johansson, 1984). The most complete presentation of the experiment can be found in (Nielsen, 1985). It should be emphasized that the only position measurement of the robot was provided from video image interpretation. The experiment has a lot of general interest since a simple robot (the Turtle moving on the floor) was extended with higher level capabilities such as obstacle avoidance, path planning, task-level programming, and a graphic based man-robot interface. The video camera is placed to get an overview of the scene, and is fixed in position. The Turtle moves on the floor performing a workcycle, which consists of visits to work stations. A human operator is assumed to define the workcycle and then not to be active or even present during the repeated cycles. The stations are the fundamental sites in the workspace symbolizing places where a load is collected or delivered, or where a task is performed. The requirements on the motion paths between stations are more flexible. The primary concern is that the Turtle reaches the stations. The scenario also includes obstacles, either described by the operator using the man-robot interface or detected by the program. The program checks that the motion paths of the Turtle avoid the known obstacles. When the Turtle detects new obstacles during operation the paths are replanned automatically. If the Turtle is unable to reach a station, due to new obstacles, it calls for help by sending an alarm to the operator.

The visual servo is provided with the capability to handle three different types of events when working unsupervised. Firstly, the Turtle can interact with the Ilon car symbolizing another (bigger) AGV which occasionally delivers material. The position of the Ilon car is not specified from time to time, but it is assumed to signal when it comes or leaves. The Turtle should go to the Ilon car and make a symbolic load transfer. Secondly, objects may unexpectedly appear in the work space. For example, humans may enter a risk zone. An external motion detector signals if something is moving into the scene. Thirdly, detected collisions must be handled.

The structure of the AGV scenario was formalized to clarify the aspects of the system common with general robot applications (Nielsen, 1985, 1986). The extent of the workspace includes all the points the robot can reach and the camera can see. The elements in the workspace are easily abstracted to geometrical objects in the floor plane. The elements are obstacles, stations, event points, and paths, whereas the geometrical objects are points, curves, and areas within closed curves. It is convenient to discuss aspects of the visual servo based on this abstraction of the workspace.

The experiment has formed a basis for further work and different problems have been isolated and studied in more detail. These results will be described in the following sections.

## 2.3 Robot programming

A man-robot interface should provide two main services. One is the programming where the workspace and tasks are defined. The other is supervision facilities at different levels of detail. Different approaches to robot motion programming have been identified in the litterature. In explicit programming, the user specifies all of the motions needed to accomplish a desired path by giving an explicit list of coordinates. Programming by teaching is done by guiding the robot manually and storing the path. The highest level of programming is task level programming. The user specifies geometric models and descriptions of tasks in terms of these models, and the detailed motions are derived automatically from these specifications.

The key idea in our work on robot programming is to explicitly use interaction based on real gray scale images of the scene (Nielsen, 1985, 1986). We may first note that the correspondence between image and workspace is especially simple in an AGV application. The mapping between the floor plane and the image is one to one. There is thus a unique correspondence between the geometrical objects of the workspace and points and curves in the image. The geometric description of the workspace may be entered by pointing in the image of the workspace using a graphical editor, where the output of the graphical editor is overlayed on the gray level images. Our man-robot interface is used to describe the workspace i.e. to enter stations, obstacles, paths, and path attributes. Hence the facilities of this interface is concentrated on motion. Other important aspects such as gripping and manipulation are not treated.

The images of the robot and its workspace are presented on the monitor screen. A graphical editor allowing color graphical manipulations on these images has been developed. The terminal, the mouse, the joysticks and the monitor screen are used for the interaction. The mouse is used for pointing in the image on the monitor screen. The commands to the editor are selected from menus on the terminal. The editor can handle points and curves. Two types of curves are possible: polygons and cubic B-splines. The shape of a polygon is defined from its vertices. The shape of a cubic B-spline is defined from control points. The editor can be used to interactively enter, copy, move, rotate, or delete both points and curves. The shape of a curve is edited by adding, moving, or deleting the shape defining points (a polygon vertex or a control point of a cubic B-spline). The cubic B-spline control points have local support. This means that if a control point is moved then the curve is changed only in the neighborhood of that point. It also satisfies the requirements of shape stability in editing, which means that the shape of a curve is not drastically changed if a control point is moved slightly. These properties makes them feasible for trajectory generation.

*Path programming.* The use of graphics in images of the workspace simplifies the motion trajectory generation. A spline is entered using the graphical editor. The spline is automatically transformed to a robot path, in our case by sampling it to a polygon, which is then projected on the floor. The program of course checks that the obtained path avoids the obstacles. The result is a smooth polygon path. In other current systems for trajectory generation several intermediate positions, via points, have to be defined to obtain a smooth path. Here only a few control points have to be defined. It is done by pointing in the image, the result is directly presented, and it is easy to modify interactively.

*Task level programming.* Commands are used to determine whether the graphical input is obstacles, stations or paths. When a new station is entered a path avoiding the obstacles is calculated by default, using a path finding algorithm, and presented to the operator. Hence, task level programming is done here using graphics. The operator may of course modify the calculated path using the path programming facility.

*Presentation of Status.* During operation there are three levels of detail for the presentation on the monitor screen. Firstly, the monitor can just display the working scene as seen by the camera. Secondly, the graphical description of the workspace can be overlayed the image. Thirdly, the steps of the image processing can be added to the presentation, giving a complete record of the internal status and algorithms. The actual processed part of the image is marked, the output of the edge detector is displayed, etc.

## 2.4 Motion detection

Visual motion detection is useful in different ways. The AGV servo uses it as an alarm mechanism if e.g. a human is entering a risk zone. Motion detection is also useful for focusing the attention, and in that way decrease the otherwise sometimes elaborate heuristics in general image understanding.

We have developed algorithms for detecting and tracking a moving object in a stationary scene (Nielsen, 1984b; Jansson, 1985). The method is based on knowledge about the visual system of the frog. It is known from classical measurements in the optic nerve that there are no signals from the eye to the brain if there are no changes in the scene. Further, the visual processing in the frog retina is spatially localized. These properties make the motion detection algorithms feasible for implementation in special purpose hardware, e.g. in one single chip. A by-product of our work is a software system for study of time sequences of images (Jansson, 1985).

## 2.5 Perspective area-invariants

Current state of the art industrial vision systems like ASEA Vision or Erivision, are essentially designed for and limited to the interpretation of images extracted from a primarily two-dimensional scene, and typically used in industrial applications with a camera looking down on a conveyor belt. These vision systems are based on edge detection and contour description. The invariants of two-dimensional translation and rotation, e.g. area, perimeter and shape factor, are used to describe closed contours in the image. An advantage of using these measurements is the error robustness. The extension of this technique to the interpretation of images extracted from a three-dimensional scene is a major research direction in robotics and promises to increase the applicability of the existing hardware.

The scene is three-dimensional in the AGV visual servo. The shape and the size of objects thus vary when viewed in different orientations and at different distances from the camera, and of course area and perimeter then vary. Here we have a new approach to visual recognition. The key idea is to use marking symbols, which have properties invariant under projection. The design problem consist of constructing symbols with descriptions based on perspective area-invariants. It is well-known that the quotient between two areas of a planar figure is invariant under parallel projection. Extensions of this result to find area-invariants under perspective projection is the main result (Nielsen, 1985; Nielsen and Sparr, 1985). The marking symbols designed here are black and white, which simplifies edge detection and contour description.

## 2.6 Optimal digitization

Handling large amounts of data is a general problem in a visual servo. An image may typically be 512x512 pixels with 8 bits per pixel. The possibility to represent an image with less data decreases the computational cost. On a conventional computer the computing time decreases. For special purpose parallel hardware the computer size needed is reduced.

Professor E.I. Jury was at the Department of Automatic Control as a guest professor, and his visit resulted in the work on "Optimal Digitization of 2-D Images" (Nielsen, Åström, and Jury, 1983, 1984; Nielsen, 1985). The problem of optimal digitization of 2-D images had been sporadically mentioned in several texts, but it had not been addressed in full details. Experimental investigations of the effect of coarse scan/fine print for bilevel images had been studied, but no theory was given in that work. However, there was a detailed theory of transmission of an analog signal over a fixed bit-rate channel in the 1-D case. This work motivated our extension of the theory to the 2-D case. However, our theory differs from the previous 1-D theory in the respect that our problem formulation admits an analytical solution.

Our theory formulates an optimal digitization problem. The image should be represented with MxN pixels and b bits per pixel, under the constraint of a fixed total number of bits. The answers to problems of this type obviously depend on the type of image and the information in it. Reasonable assumptions are made so that the optimization problem has a closed form solution. The solution has been tested experimentally and it agrees well with human visual perception of picture quality. The image characteristics used are expressed in terms of intensity range and second order moment of partial derivatives. The analytical solution clearly shows the dependence on image characteristics, and the solution also explains results of earlier subjective tests performed by others. The simplicity of the solution makes real-time adaptation possible, e.g. to changes in illumination.

## 2.7 GaAs-wafers

Semi-insulating boules of GaAs are grown for production of logic circuits. This fabrication is an example of process control. A joint project with the Department of Solid State Physics, Lund, on 2-D characterization GaAs-wafers using image processing has been performed (Silverberg, Nielsen, Omling, and Samuelsson, 1985). The project has resulted in a new fast method for the determination of the energy level EL2 in GaAs-wafers. This energy level, EL2, determines to a large extent the electrical properties of semi-insulating GaAs. The method is based on processing of images obtained by using infrared light to penetrate the wafer. Images are taken from two different states of the wafer, quenched and unquenched. A new image representing the concentration of EL2 is obtained by mainly point-wise division of the two images. Division is easy on a computer but otherwise difficult to implement. The result is called an EL2-map. This information can be presented in different ways by the image processing system. The information of the spatial distribution can be used to adjust growth process parameters like pull speed, partial pressures, temperature etc.

## 2.8 Ash-line control

Burning processes present important control problems. We have successfully studied one example (Dahl, 1985; Dahl and Nielsen, 1986). The work was initiated, and will also be continued, by ASEA Generation (Ideon, Lund) and Götaverken Energy Systems (Gothenburg).

Bark furnaces are used to generate steam or hot water for heating or for power generation. The control signals to a furnace are a spatial vector of fuel inputs at the top edge of the grate and a matrix of air supplies from under the grate. For efficient firing it is important that the ash line is close to the lower edge of the grate. The ash line is the border between burning fuel and fuel which is burned out (ash). This insures there is fire over a large area of the grate, and no burning material will exit unused. In many plants the process supervision is done manually. If the operator wants to determine the position of the ash line, he has to leave the control room and walk (sometimes climb) to a suitable view of the process.

Currently bark furnaces are increasingly being equipped with a video system for convenient inspection. Our investigation is a case study of extending such a video system with image processing facilities. The main purpose of the project was to develop image processing algorithms which, implemented on a small computer installed at a bark furnace, can give information about the position of the ash line as a function of time. This information can in a first step of development be presented to an operator, and in the future possibly be used for closed loop control of the process. It is believed that knowledge of the ash line position as a function of time will assist the operator to improve the manual control of the process. The process is slow, and trend curves are a good help for human memory.

The main result is an algorithm for ash-line detection. A criterion was especially designed to formalize the properties of the ash line. The criterion is designed to be based on local properties. This, together with the knowledge that the ash line goes from the left side of the image to the right side, makes the technique of dynamic programming feasible for the optimization. For each image sampled in time, the ash line is determined. The position is determined in eight segments of the image. After time filtering with a Butterworth filter the positions are presented as trend curves. The performance of the algorithm has been investigated by experiments on a 75 minute video tape recording from the Stora Kopparberg plant in Skutskär. The video tape contains particularly difficult operating conditions for ash line detection due to wet bark, and the algorithm still performs successfully. The computation time (on a VAX-11/780) is less than 30 seconds per image. This indicates that a prototype system, implemented on a small computer (IBM-PC, Macintosh etc.) will fulfill the speed requirements given by the process.

## 2.9 Visiting researcher at Caltech

I have spent the year 1985-86 working with professor Carver Mead and his group at the Department of Computer Science 256-80, Caltech, Pasadena, Ca 91125, USA. The work there is concentrated on VLSI-design of chips, and a new technology is under development based on integration of sensors and processing on the same chip. My own work is not published yet, but a brief overview is as follows. The first period of my stay I constructed a servo consisting of a vehicle under computer control. Different vision chips developed by the Mead group has been mounted on the vehicle and investigated under motion. I also made some test chips integrating optical sensors and pulse control servos. Thereafter I was mainly working together with Misha Mahowald and Carver Mead on design and test of a large complicated chip that combines visual and auditory signal processing.

# 3. Conclusions

There is not yet a problem formulation of visual servoing which captures the essential properties and admits an analytical or numerical solution. It is thus essential to build a system, and try significant experiments. Using our laboratory set-up we have obtained results in two cases of process control and in different aspects of motion control related to the experiments on automated guided vehicles.

The project on characterization of GaAs-wafers is the first example in the process control area. The Department of Solid State Physics have now built their own image processing system. They use the method for basic research on physical properties of GaAs. The method could also be of value in industry. Currently GaAs-circuits are tested after the chips have been fabricated. The possibility to determine anomalies already in the wafer but before producing circuits could save considerably on time and resources.

The second example of process control is the control of bark ovens. Visual servoing offers the possibility of direct measurements of properties like good firing in burning processes instead of indirect measurements like temperature and pressure. The project was a positive first experience in vision control for ASEA Generation. They will exploit the result and also continue to look into control of other burning processes.

Sensory feedback in motion control clearly points toward more flexible systems than the systems of today. The conclusion from our laboratory experiment is that an industrially useful automated guided vehicle (AGV) in flexible warehouses could be obtained using available technology and the principles presented in our work.

A flexible system increases the requirements on good interaction with an operator. A man-robot interface was designed, based on an operator manipulating interactive color graphics overlaid on images of the working scene of the robot. In this way it is feasible to use a new path programming technique requiring few control points instead of current methods involving several via points. Furthermore, task level programming can be done by pointing out robot destinations in the image. The use of graphics here is thus very different from other recent uses in robotics, where graphics have been used for simulation but not to simplify the real-time interaction.

Visual recognition in a three-dimensional scene has to cope with the problem of objects having infinitely many poses. Obviously all poses cannot be stored, and therefore one must rely on more dense descriptions. A key contribution here is the design of marking symbols, which are described by area-invariants under perspective. The constructed symbols can be used for robot marking or as signposts, and since they are based on areas within closed contours they are insensitive to individual pixel errors and they can be easily detected by existing special purpose hardware like ASEA Vision or Erivision.

The result on optimal digitization represents the first theory presented on a problem previously mentioned in several texts. The results can be used e.g. as a rule of thumb, for adaptation, or for structuring of special purpose image processing hardware.

# 4. Future research

Broad technology aspects are involved in visual servoing. This is especially true in robotics, and future work in robotics will have wide-spread general scientific interest. Further, robots have become an important and widely used production element in industry. Sweden has been internationally competitive for several years. Currently several manufacturers offer robots with good and reliable mechanic and basic electronic control. Future competition will be based on the capabilities of providing more flexible and intelligent functions. The possibilities of shorter production series, less fixtures surrounding the robot, and faster reprogramming are primary goals. Other more speculative ideas are moving robots in hazardous environments, for example automatic fire-brigades in nuclear power plants.

Higher robot functions need increased interaction between the robot and the environment. The sensory input to the robot could be based on touch (pressure or force), hearing, or vision. The information processing in currently developed sensors for robotics tend to be based on analog chips or hybrid chips combining analog and digital processing. New sensors and new chips are in several cases most naturally explored and tested in a robotics laboratory. Other important issues are interpretation of the sensory data and feedback from the information obtained. Here the developments in expert systems and artificial intelligence, pattern recognition, adaptation, and real time programming are important to combine.

# 5. Publications

The publications in the project are so far:

DAHL O. (1985): "Image processing techniques for ash line detection," Master thesis CODEN: LUTFD2/TFRT-5328, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

DAHL O. and L. NIELSEN (1986): "Ash-line control," First IFAC Workshop on Vision Control, June 10-12, Espoo, Finland.

JANSSON, S-O. (1985): "Rörelsedetektering i bildsekvenser," Master thesis CODEN: LUTFD2/TFRT-5327, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

JOHANSSON, K. (1984): "Ett exempel på robotpositionering med hjälp av videokamera," Master thesis CODEN: LUTFD2/TFRT-5307, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

NIELSEN, L. (1984a): "Bildtolkning i ett exempel på robotpositionering med hjälp av videokamera," Symposium Bildanalys, Svenska Sällskapet för Automatiserad Bildanalys (SSAB), Linköping 15-16 mars.

NIELSEN, L. (1984b): "Motion detection in image sequences," NSF-STU International Workshop on Computer Vision and Industrial Applications, May 14-18, Stockholm, Sweden.

NIELSEN, L. (1985): "Simplifications in visual servoing," Ph.D. thesis CODEN: LUTFD2/TFRT-1027, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

NIELSEN, L. (1986): "A Visual Servo," First IFAC Workshop on Vision

Control, June 10-12, Espoo, Finland.

NIELSEN, L. and H. ELMQVIST (1983): "An image laboratory," Report CODEN: LUTFD2/TFRT-7261, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

NIELSEN, L. and K. JOHANSSON (1984): "Robot med egna ögon," *Industriell Datateknik 1984:5*, 25–29.

NIELSEN, L. and G. SPARR (1985): "Perspective area-invariants," Report CODEN: LUTFD2/TFRT-7313, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

NIELSEN, L., K. J. ÅSTRÖM, and E. I. JURY (1983): "Optimal digitization of 2-D images," Report CODEN: LUTFD2/TFRT-7265, Department of Automatic Control, Lund Institute of Technology, Lund, Sweden.

NIELSEN, L., K.J. ÅSTRÖM, and E.I. JURY (1984): "Optimal digitization of 2-D images," *IEEE Trans. on Acoustics, Speech, and Signal Processing* **32**, 1247–1249.

SILVERBERG, P., L. NIELSEN, P. OMLING, and L. SAMUELSSON (1985): "EL2-maps from computer based image analysis of semi-insulating GaAs wafers," Symposium on Defect Recognition and Image Processing in III-V Compounds, July 2-4, Montpellier, France.