23

Vision Survey System, a tool for providing 3D product definition data of large constructions

W.H.A. Hendrix, P. Dunias, M. Hanajík, N. Kouwenberg
Eindhoven University of Technology
Dept. of Electrical Engineering
Measurement and Control group
P.O.Box 513, NL-5600 MB Eindhoven, The Netherlands
fax: +31 40 2434582; tel: +31 40 2473309
email: W.H.A.Hendrix@ele.tue.nl

Abstract

This paper describes a Vision Survey System, a tool for providing Product Definition Data for automatic manufacturing of large unique construction elements. The system uses multiple cameras to scan large workpieces. Scene interpretation and model-based image processing results in 3D-geometry data of the workpiece and its position and orientation in an absolute coordinate system. The achieved accuracy is satisfactory for the application, typically a robotic production environment. Here the Vision Survey System can be considered as a global (off-line) sensor to provide input data for initial robot trajectory generation, eliminating the burden of manually teaching the robot's TCP paths. Local sensors are used (on-line) for fine-tuning of the robot movements.

The system development was funded by the EC within the 3rd Framework R & D programme in the field of Information Technologies. The ESPRIT Projects nr. 5369 HEPHAESTOS I and nr. 6042 HEPHAESTOS II were carried out with the objective of constructing an intelligent robot arc welding system for thick-steel unique fabrications, see ESPRIT (1992, 1995). The projects involved international research cooperation with partners from Greece, Spain, UK, France, Sweden and the Netherlands.

1 INTRODUCTION

To perform 3-dimensional measurements the Vision Survey System (VSS) uses images acquired by CCD-cameras mounted on a transport system together with remote controlled lights, and a highly automated processing of the camera images. Within the ESPRIT-

L. M. Camarinha-Matos et al. (eds.), *Balanced Automation Systems II*

project HEPHAESTOS II, the development was guided by experiments in a real industrial environment and the system was tuned to being applicable at a Ship Repair Yard in Piraeus, Athens.
There, the VSS was planned to be used for identifying and obtaining the 3D geometry of large steel structures, typical for ship repair. The output data of the VSS was used for programming a sensor guided arc-welding robot. In particular for the path planning unit which calculates collision free trajectories and determines the starting point of a seam the 3D geometry of the workpiece must be known.
This paper describes the principles of operation and system performance.

2 PRINCIPLES OF OPERATION

2.1 Input Data

The basic input data used for calculating the 3D properties of a workpiece are a number of digitized camera images obtained with one or more cameras placed at different locations above the workpiece (workpiece 'scanning'). The number of viewpoints and the required camera locations depend on the type, complexity and size of the workpiece. To obtain 3D information from 2D acquired images, each digitized camera image requires the calibration data specifying the position, orientation and optical properties of the camera that was used for the image acquisition. The calibration principles will be explained later. The digitized images have a 256 kByte size of 512(H) x 512(V) pixels, each pixel having an 8-bit resolution, thus enabling 256 grey levels. See figure 1.
Within the ESPRIT-projects the workpiece scanning was done by mounting 4 cameras on a robot gantry with 3 orthogonal axes supporting a hanging welding robot with 6 axes . The robot gantry provided a 12x2 metres xy-translation of the camera system.

2.2 Data reduction

The digitized grey value camera images that were obtained by 'scanning' the workpiece will be input for an edge detection and line extraction process. The goal is to convert the 256 kByte size grey value image into a line image, i.e. an image consisting of straight and curved line segments, see figure 1. A typical line image contains about 300 lines (1 kByte). In the ideal case the lines in a line image represent the closed contours of the objects that make up the workpiece. However, due to shadows, light reflections and non-homogenous colouring of the workpiece surface, the contours will not be closed, some contour lines may be absent, and non-contour lines will be present in the line image.
An edge detection algorithm as described in Lee (1989) will produce edge pixels and gradient information with sub-pixel precision. Next, a line extraction algorithm connects chains of edge pixels into straight line segments or elliptical curves. Only straight lines and elliptical curves are needed to describe the 2D-projections of the contours of the workpieces to be handled by the welding robot. Parameters affecting the behaviour and performance of the edge detection and line extraction processes have default values and can be adjusted by the VSS-operator.

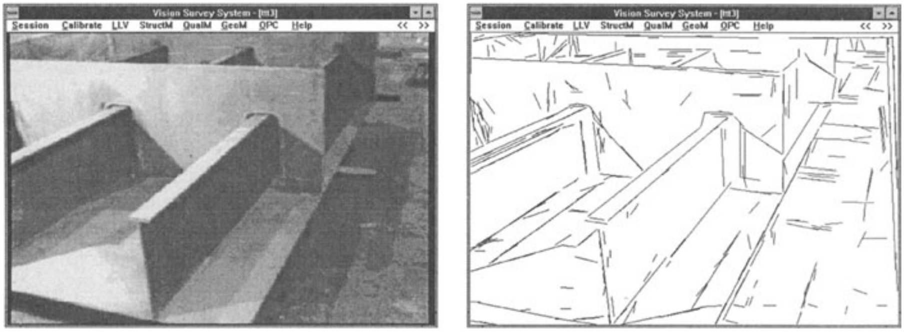

Figure 1 Typical workpiece (left) and Line Image (right).

2.3 Object identification

As explained above, each line image contains non-perfect 2D-projections of object contours and additional 'noise' lines. The perspective projection does not disturb the straightness of line segments, however, circle segments are transformed into curves which can be approached by ellipse segments. Object identification as described in Kaptein (1993) and Dunias (1994) will be used to find in each line image a subset of lines that might be the 2D projections of contours of a specific object. A knowledge base describing the structure of objects in terms of 3D- and 2D-properties of its contour lines, will be loaded and applied to the line image.

Due to the knowledge base concept and its open structure it is easy to extend the number of different objects that can be recognized. A knowledge base is a readable ascii-file in a specific format. The knowledge base can easily be adapted. Using the implemented set of line properties and bearing in mind the file format, suitable lines can be added to obtain a new knowledge base. There is no need to change the code and re-compile the VSS-software.

After object recognition, each line image is accompanied by the information that specifies groups of lines representing the contour lines of an object, and the correspondence between these lines and a predefined object model.

2.4 3D reconstruction

After the identification of the structure of the workpiece, the position, orientation and dimensions of each identified object must be calculated. This is done by matching a parameterized three-dimensional (3D) model of the identified object to the two-dimensional (2D) contour lines. A parameterized model is a model, where the position and orientation of the so called 3D features with respect to the absolute coordinate system is specified in terms of a number of unknown parameters. The features suitable for matching are e.g. contour lines of the object. Once the values of the parameters are given, the object is specified, and the 3D features of the object can be projected onto the 2D camera image plane. The location of projected 3D features is compared with the 2D features (e.g. lines), which were identified in the previous step. A value of some objective function, which expresses the deviation of location of the projected 3D features from the 2D features is calculated, and in an iterative process the value of parameters is subsequently updated, so that the minimum of the objective function, corresponding to the best match between the projected 3D features the 2D features, is achieved. Such computed values of the object model parameters specify the object. In the VSS system, the chosen objective function is the sum of squares of distances of end-points of the projected object model contours from the corresponding identified 2D image lines. The parameters can not always be computed. If this is the case, we can say that the parameters are not observable. Whether the parameters are or are not observable depends on the number and kind of unknown object parameters, on the available identified 2D features and other conditions. Several constraints can be employed to make the parameters observable. In the VSS system, a ground plane constraint is employed: It is assumed, that the features on the bottom of the workpiece lie on the ground plane with the z-coordinate equal to zero.

2.5 Object part clustering

After all images have been processed up to 3D-reconstruction, we have obtained from every image a set of partial 3D object descriptions. In order to create one 3D-description for the whole workpiece the processing results of the separate camera images must be combined. This is done by examining all the partial 3D object descriptions aiming to group all the line subsets found during object identification in such a way that each group contains only line subsets belonging to one physical object. After that, these line subsets are once more input for the 3D-reconstruction module which will now produce complete 3D-object descriptions.

2.6 Camera calibration

Camera calibration is basically the computation of the parameters that specify the position and orientation of the camera with respect to a world coordinate system, and the camera parameters affecting the projection properties. The position and orientation of the camera is specified by six so-called extrinsic camera parameters: three of them specifying the position and three specifying the orientation of the camera. The intrinsic camera properties are specified by ten so-called intrinsic camera parameters: One parameter specifies the distance of the image plane from the optical center. Two intrinsic parameters specify the resolution, i.e. the spacing of picture elements on the image plane. Two intrinsic parameters specify the offset of the origin of the image plane (the location of the picture element with coordinates (0,0)) from the location where the optical axis intersects the image plane. Finally, five intrinsic parameters specify the distortion caused by the lens and the camera geometry imperfection (image plane is not exactly perpendicular to the optical axis, etc.).

The principle of the camera calibration is in much the same as the principle of the 3D reconstruction, where also the parameter values of the parameterized model are computed. When calibrating camera parameters, an image of the reference-pattern or reference-object is acquired, and the camera parameters are computed, so that the features extracted from the acquired image match the features obtained as projections of the reference object/pattern 3D features using the camera model.

Two calibration routines are available. One, the 'intrinsic calibration', computes all camera parameters including the intrinsic parameters, and a planar grid of calibration marks is used as reference pattern. Another calibration routine, the 'extrinsic calibration', computes only extrinsic camera parameters, and uses a 3D object, an accurate cube with known sizes and at a well-known position and orientation, as the reference object. The calibration techniques used are based upon the methods described in Tsai (1987) and Weng (1992).

As long as adjustments of the camera, namely the aperture and the focal distance of the camera lens, are not changed, the camera's intrinsic parameters do not have to be re-calibrated.

The extrinsic parameters must be re-calibrated after each undefined replacement of the camera. The cameras are mounted on a transport system. When the transport system performs a well-defined movement, the extrinsic camera parameters do not have to be re-calibrated, however they must be updated appropriately according to the transport displacement.

2.7 Manual mode

Especially for the object identification the VSS user interface provides a means for manual object identification. Here the operator will examine the line images and use the computer pointing device for selecting lines in order to create sets of lines that represent object contours, in the same way as this would be done by the automatic object identification module. In this way it is always possible, even in difficult environmental conditions or with complex workpieces, to create a VSS output.

3 SYSTEM PERFORMANCE

The Vision Survey System performance and quality can be expressed by the following features:

- robustness of the automatic object identification
- 3D measurement accuracy
- processing speed
- required operator skills
- level of universality

3.1 Robustness of the automatic object identification

As described previously in section 2, automatic object identification is the intermediate step from converting grey value images into line images and processing these line images to identify object contours. Because of the non-ideal line images (incomplete contours and false contours) the process of automatic object identification can result in missing and/or false object identifications. The robustness of this process depends on the quality of the line images and in relation to that the intelligence of the knowledge base. However, at this moment the performance of the knowledge based system is limited. On the one hand because of the fact that mainly 2D-properties of the contour lines are known, and on the other hand because the process of object identification operates only locally i.e. the system doesn't have the ability to make a scene overview using its own previous results. Consequently the automatic-mode of object identification must be accompanied by a manual-mode in order to achieve a 100% correct output result.

3.2 3D Measurement accuracy

The actual accuracy of the 3D measurement depends on the following items:

- *Accuracy of camera calibration.* The adopted camera model and the calculated position and orientation of the camera specify the 2D-projection process that is assumed by the VSS. Any inaccuracy of these intrinsic and extrinsic camera parameters affects the final 3D-measurement accuracy.
- *Camera mounting.* The cameras must be mounted very rigidly. There must be no undefined changes in their position and orientation due movements of the transport system.
- *Image resolution.* As the size of the digitized camera images is fixed to 512x512 pixels, the sensor resolution (pixels/mm) can only be changed by changing the field of view of the camera. A higher sensor resolution will result in a smaller field of view which in turn implies a larger number of images needed to have a complete workpiece 'scan'.
- *Workpiece alignment.* As described previously in section 2, for 3D-reconstruction it is assumed that all objects are placed on a plane with equal z-coordinate (z=0). If in reality this is not the case, 3D-reconstruction after object part clustering will be inaccurate.

At this moment it is not possible to give figures for all aspects listed above. Experiments in our laboratory showed an average of 25 mm for the total accuracy, here the workpiece was 6x3x1 metres and the cameras 2 metres above the workpiece. For the ESPRIT-project objective this result meets the required accuracy regarding the large workpiece dimensions, its considerable deformation during welding and the range of the local path correcting sensor systems.

3.3 Processing speed

The time needed to make a full 3D-reconstruction of a workpiece by the VSS is defined by the following components:

- *Time needed for intrinsic camera calibration.*This requires every camera to be mounted in a special measurement set-up to view the reference pattern. After calibration the camera can be mounted in the working area (on the robot gantry).
- *Time needed for extrinsic camera calibration.* Extrinsic calibration of each camera is required if the position and orientation of a camera with respect to the world coordinate

system has changed in an undefined manner. This situation should be prevented during manufacturing.

- *Time to scan the workpiece and process all the camera images.* Here, we distinguish:

  - Time for image acquisition. For each image this means positioning of the transport system, updating calibration according to the new transport system's position, light adjustment, image digitizing and image file saving. Average time 2 min/image.
  - Time for automatic processing. Average time 2 minutes per image.
  - Time for manual correction. Average time 2 minutes per image.
  - Time for clustering and creating output file. Average time 2 minutes.

3.4 Required operator skills

The VSS operator will handle the system using a specially designed user interface. For normal operation this user interface is easy to use and has an on-line HELP facility. For more advanced operations such as camera calibration and knowledge base creation an operator manual is required. A person with a technical background can be trained to use the system in a few days.

3.5 Level of universality

Although the VSS has been developed with the aim to applying it at a ship repair yard, some design concepts have been chosen to achieve a more or less flexible, universal 3D-identification and 3D-reconstruction tool. Some VSS features that express this flexibility and universality are listed here:

- The VSS has its own built-in camera calibration methods, so cameras can be placed at arbitrary positions above a workpiece. Furthermore each camera's field of view can be changed easily by changing the lens for one with another focal length. All this results in a wide range of workpiece dimensions and workpiece construction complexity that can be handled by the VSS.
- The VSS allows a variable number of cameras, from 1 to 4 (or 1 to 8 when changing the video multiplexer hardware). So in order to scan workpieces one can take a large number of fixed cameras or a smaller number of moveable ((x,y,z) translation) cameras. If for a certain application fixed cameras are sufficient this will in most cases give a cheaper solution, as moveable cameras require an accurate positioning device (like the robot gantry used in this ESPRIT-project).
- An important aspect is the built-in knowledge base system. This concept is used for describing objects which are used to construct workpieces. The set of objects which can be handled by the VSS can be extended easily. To do this the user has to create a knowledge base (write a separated readable ascii text file with specific format) which describes the new object in terms of a limited set of (mainly) 2D-predicates.

As a 3D-measurement tool a further developed Vision Survey System can be used to achieve robotic automation of 'dirty' manual labour like for example sandblasting and painting of steel constructions or cleaning of ship rooms and oil containers.

4 CONCLUSIONS

Experiences obtained in numerous experimental tests in our laboratory as well as under heavy manufacturing conditions in the industrial environment of the ship repair yard have shown the performance, the reliability and quality of the Vision Survey System, as detailed in section 3. However, the VSS can still be refined for real industrial production. A user who wants to operate the vision system is required to have much knowledge on details of software implementation. Moreover, the VSS is restricted at this moment to handle only straight line contours, although the implementation of circle segment contours has already been explored. Furthermore we can say that the introduction of the automatic arc-welding robot into the ship repair yard in Greece was quite a big step. Before the projects concerned started, no arc welding robots in this country were known, and finally it became the first arc welding robot in this factory where almost all the welding and steel cutting work was done manually. It was also a very sophisticated robot, equipped with many options for arc welding and with several environment sensor systems in order to adapt the robot path and the welding process to actual workpiece deviations caused by the welding heat, and to realize a flexible production environment. In fact, the project goals can be classified as not very well-balanced automation systems. A lot of the manual labour at the ship yard could be automized with less advanced systems than the robotic system developed within the ESPRIT-projects, but obviously there was no good mixture of economical need and innovative management to establish this. Furthermore it is clear that the EC constrains on funded projects caused this non balanced automation at the Greek ship yard. After all, the EC funded projects aim to innovate were the state-of-the-art is found in the high industrialized northern countries of Europe. The demonstrated successful results of the complete project and the punctually delivered document reports convinced the reviewers of a productive and close cooperation between partners. After the second ESPRIT-project has finished, the ship yard will put efforts into integrating the robot in the process of repair part production, using the complete set, or some of the implemented sensor systems.

REFERENCES

Final Project Reports (1992,1995) *ESPRIT Projects 5369 and 6042 HEPHAESTOS 1 and 2, Intelligent Robotic Welding Systems for Unique fabrications.*

Dunias, P. and Kouwenberg, B.(1994) Knowledge-Based 3-D Object Reconstruction *IEEE Instrumentation and Measurement Technology Conference (IMtc/94)*, 1337-4.

Kaptein, E., Dunias, P. and Kouwenberg, N. (1993) Structural matching in knowledge-based 3D scene reconstruction *Journal A, No.4, Dec. 1993*, 17-8.

Lee, D.(1989) Edge detection, classification, and measurement *IEEE Comput. Soc. Conf. on Computer Vision and Pattern Recognition*, 2-9.

Tsai Roger Y.(1987) A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses *IEEE journal of Robotics and Automation*, 3:323-22.

Weng, J., Cohen, P. and Herniou, M.(1992) Camera calibration with distortion models and accuracy evaluation *IEEE Trans. Pattern Anal. Machine Intell.*, 14:965-16.