Sensors and Actuators A, *31* (1992) 241–244

An inexpensive real-time 3-D camera with a smart optical sensor

J. Kramer and P. Seitz
Paul Scherrer Institute Zurich, Badenerstrasse 569, CH-8048 Zurich (Switzerland)

H. Baltes
Swiss Federal Institute of Technology, Physical Electronics Laboratory, ETH-Hönggerberg HPT, CH-8093 Zurich (Switzerland)

Abstract

A 3-D camera has been developed, capable of measuring 50 complete depth images per second of arbitrary, diffusely reflecting objects. The output of this real-time 3-D camera is a CCIR video signal, in which the depth is linearly coded as the local gray-value of the video image. The measurement principle is active triangulation using a solid-state laser and two galvanometer mirrors for beam and scene scanning. No complicated electronics or computers are required because the essential 'calculations' (centre-of-gravity of light profiles) are carried out purely optically with a special sensor consisting of wedge pixels. The use of this smart sensor makes the 3-D video camera relatively inexpensive, and the output in video-format allows the connection with any standard video framestore for direct 3-D image processing. It is concluded that this development can be of appreciable practical interest for many applications in industry.

1. Introduction

Standard imaging devices create two-dimensional optical reflectance maps of three-dimensional scenes. These are very suitable for human visual perception as long as they contain familiar structures which can readily be associated with known three-dimensional objects. However, reflectance depends on different local physical conditions, such as illumination, reflectivity, texture and distance. The most important parameter for the description of unknown objects and for manufacturing is the local distance from a reference plane, i.e., the surface shape of an object, which cannot fully be extracted from a reflectance map. Devices which are able to acquire depth maps of objects are called 3-D cameras. Their main applications include medical diagnosis, optical inspection of an automated production process, and robot guidance. Several 3-D cameras are available on the market. The choice of a system for a particular task depends on different criteria such as spatial and temporal resolution and price.

It was decided to develop a system with the emphasis on low cost and high speed, suitable for the observation of moving objects (parts on a conveyor belt, robotics) [1]. A useful format for recording and display of moving images is defined by the video standard. This requires a system which is capable of recording and processing 50 3-D images per second (CCIR video field rate). The resulting data rates of several MHz place high demands on its electronics. In order to keep costs low some preprocessing has therefore to be done in the optical domain.

2. Measurement principle

The system is based on the light sectioning method (Fig. 1). The object is illuminated with a plane sheet of light, and viewed from a direction out of the light plane. The image consists of a profile line containing the information about the position of the observed surface points for a given illumination and imaging geometry. Two space coordinates are specified by the two-dimensional structure of the line image, the third one is fixed by the position of the light sheet. Through scanning of the light sheet across the object, i.e., variation of this third coordinate, the whole part of the object surface which can simultaneously be illuminated and observed in the image may be recorded.

0924-4247/92/$5.00

© 1992 — Elsevier Sequoia. All rights reserved

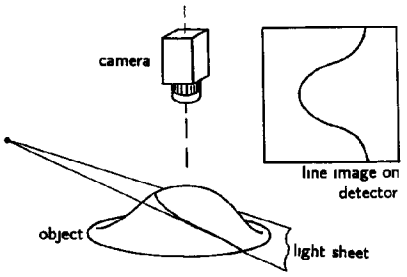


Fig 1 Principle of light sectioning A plane sheet of light is projected onto a three-dimensional, diffusely reflecting object A camera positioned out of the illumination plane records the resulting profile line from which the shape of the object can be inferred

3. Sensor

Most light sectioning devices employ a video camera to image the profile line With this method a large amount of redundant data is produced and the acquisition of a surface map takes several seconds The processing of a profile line, however, only requires imaging in one dimension and detection of the optical centroid in the other Thus, the optimum sensor for this task is a one-dimensional array of position-sensitive devices These can either be lateral-effect photodiodes [2] or pairs of identical triangular vertical photodiodes, arranged in such a way that the total length of a light line segment on the photosensitive area is constant but its partition depends linearly on its lateral position [3, 4] In both cases the position of the optical centroid is given by the ratio of the difference and the sum of the two diode currents Neither of the two sensor array types is commercially available Nevertheless we were able to implement the latter by glueing a simple stripe mask at an angle onto a standard photodiode array with rectangular pixels A detector element consists

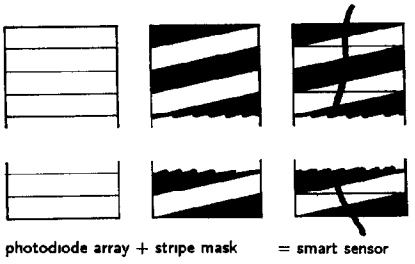


Fig 2 A stripe mask glued at an angle on top of a one-dimensional array of rectangular photodiodes produces an array of position-sensitive detectors capable of optically "calculating" the centre-of-gravity of a light profile

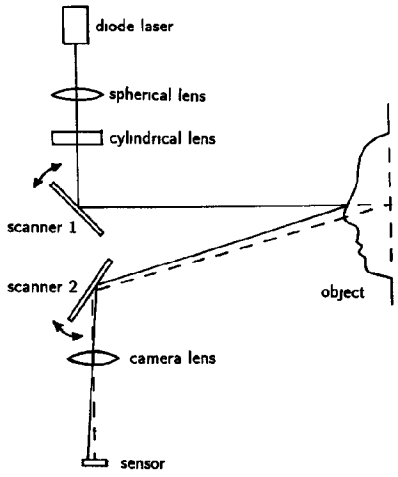


Fig 3 The set-up of the 3-D camera consists of an IR laser diode projector, two synchronously moving galvanometer mirrors, one in the illumination path and one in the observation path, and a standard camera lens imaging the profile lines onto the smart sensor

of two adjacent diodes, each of which is half-covered by the mask, as illustrated in Fig 2

4. Experimental set-up

The implementation of the system is shown in Fig 3 The light sheet is generated by a 500 mW IR semiconductor laser (wavelength 810 nm), a spherical and a cylindrical lens The resulting profile line on the object is imaged by a camera lens onto the sensor, which has a resolution of 128 pixels Two galvanometer scanners are used in the set-up, the first to sweep the light line across the object and the second to decouple the depth signal from the movement of the first one This means that a given lateral position of a line element on the sensor always corresponds to the same depth, independent of the position of the light line on the object The scanners are driven at CCIR television field rate (50 Hz) with tangent-corrected ramp functions ensuring a constant sweep speed on the object

5. Electronics

Figure 4 shows a block diagram of the electronics A 10 MHz quartz oscillator crystal provides the master clock generating the 5 MHz clock for the sensor readout, the pixel clock running at 2 5 MHz, the horizontal and vertical video timing,

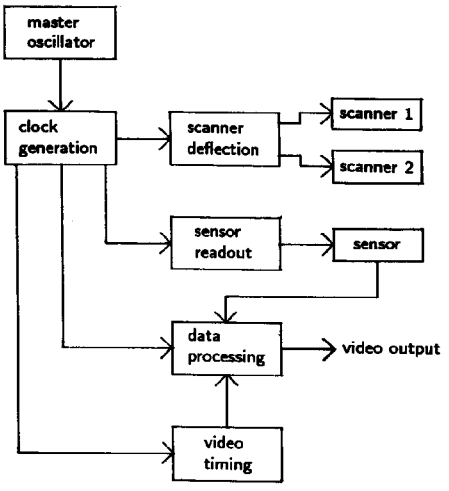


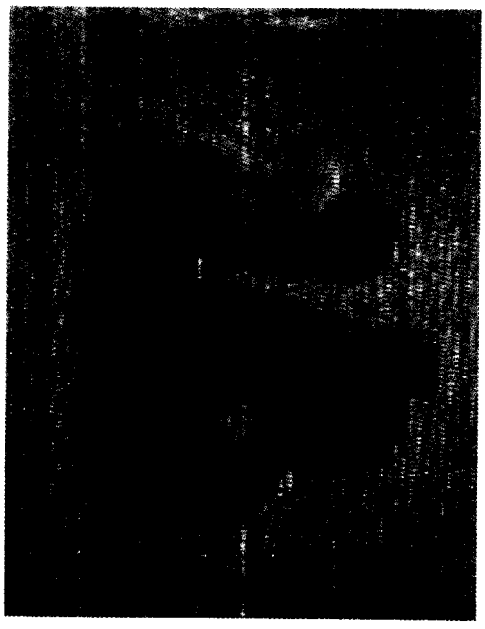
Fig. 4. Schematic diagram of the 3-D camera's electronics. A central quartz oscillator provides the clocks for the ROMs driving the galvanometer mirrors, for the readout of the smart linear sensor, and for the generation of the video depth signal according to the CCIR norm.

and the 62.5 kHz clock used to produce the drive signals for the deflection of the galvanometer mirrors from PROM look-up tables. The programmation of these takes into consideration the geometrical configuration of the device. Data processing is performed in the digital domain with 8 bit resolution. The output data is provided in CCIR video format. Each recorded profile line is converted into a video line, the signal voltage being proportional to the local depth. The output can directly be displayed on a television monitor, where the distance is encoded in gray-levels.

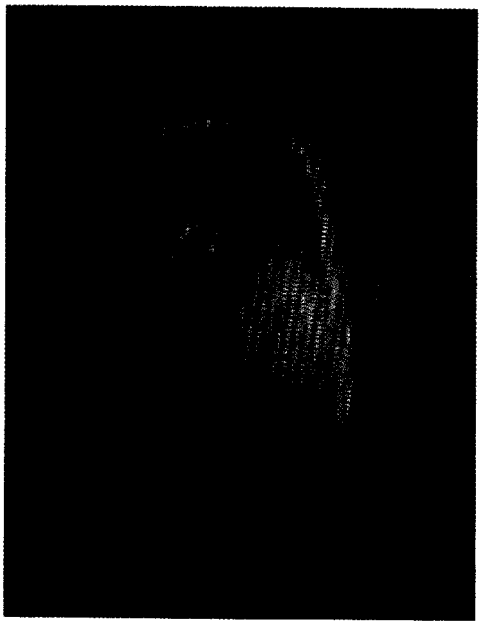
6. Results and discussion

Several objects have been imaged with the triangulation angle (between the plane of illumination and the direction of observation) ranging from 15° to 25° (Fig. 5). A single image (TV field) consists of approximately 128 × 250 pixels, but the vertical resolution is doubled through field interlacing. Depth accuracy is around 1% of the depth range for non-critical object points at slow speeds (1.4 kHz data rate). The accuracy at video rate has not yet been evaluated.

Different problems can degrade the performance of the system considerably. The so-called shadowing effect encountered in all triangulation systems is due to the fact that the part of the surface which



(a)



(b)

Fig. 5. Two examples of gray-level depth images taken in video real-time: (a) a human hand (upper) and (b) a face mask (lower). Both images are photographed directly from a monitor displaying the 3-D camera's output.

can be illuminated by the light sheet is not necessarily identical to the part seen by the sensor. This results in the recording of incomplete data sets. Shadowing is reduced by diminishing the triangulation angle. Multiple reflections on concave surface areas cause inaccuracies as the optical centroid of the line image does not normally correspond to the one of the primary reflection. A strong directionality of the scattered light may produce bright points in the line image giving rise to errors due to the geometrical separation of the two photodiodes defining one pixel. Intentional defocusing reduces this effect considerably. Such directionality in conjunction with large variations in reflectivity requires a sensor with a wide dynamic range. Some fixed pattern noise is introduced by imperfect shaping and misalignment of the mask on the photodiode array. The distortions due to the slightly varying direction of illumination could be avoided through the use of a telecentric system.

7. Conclusions

A camera has been presented which is able to generate depth maps of object surfaces at a rate of 50 Hz and display them in a standard video format. It is cost-effective, because it mainly consists of commercially available optical components and only a minimum of electronic hardware. High-speed operation makes it suitable for applications in medicine, robotics and for optical inspection (e.g., of objects on a conveyor belt where no scanning system is needed). Accuracy is estimated to be generally a few percent of the measurement range and might probably be improved through optimization of the sensor.

References

1 J. Kramer, P. Seitz and H. Baltes, Integrierter 3D-Sensor für die Tiefenbild-Erfassung in Echtzeit, in R. E. Grosskopf (ed.), *Informatik-Fachberichte*, Vol. 254, Springer, Berlin, 1990, pp. 22–28.

2 A. Kawasaki, M. Goto, H. Yashiro and H. Ozaki, An array-type PSD (position-sensitive detector) for light pattern measurement, *Sensors and Actuators, A21–A23* (1989) 529–533.

3 G. Häusler and H. Weissmann, Heterodyne triangulation system for low cost distance measurement, *Jahresber. Phys. Inst. der Universität Erlangen, Angew. Optik*, (1987) 14.

4 P. R. Haugen and C. Bocchi, High speed line sensor for mobile platforms, *Optics, Illumination, and Image Sensing for Machine Vision III, Proc. SPIE, 1005* (1988) 145–152.