A Digital Camera and Real–time Image correction for use in Edge Location.

D.Hutber
Sowerby Research Centre
British Aerospace NESD
P.O.Box 5 FPC 267
Filton
Bristol
BS17 3JW

S.Wright
Cambridge University Engineering Dept.
Mill Lane
Cambridge

Abstract

The accuracy of a depth map acquired using triangulation or stereo techniques is limited by the resolution of the sensor, and by the accuracy with which distortions of the image can be calibrated. This paper lists the sources of error in an imaging system, and concludes that most of them can be removed by careful camera design and calibration, with the exception of the spatial quantisation (discrete sampling). The use of a high–fidelity camera in performing experiments to calculate the errors in a lens and CCD array is described, and the results are compared with a standard technique for camera calibration. A pipeline architecture for real–time calibration of the image is proposed.

1 Introduction

One of the most important applications of computer vision research is in the area of Computer Integrated Manufacturing (CIM). This application covers a range of tasks, two of which are robot navigation and automatic inspection. These tasks demand specific requirements from a vision system, for example, robot navigation demands a speed of operation high enough to provide feed–back to a mechanical control system, while automatic inspection can require a high degree of accuracy.

The approach used by the authors in constructing systems to perform these tasks involves the acquisition of a depth map as a first step to doing any processing. The depth map is formed by looking for edges in the separate images, and then using the disparity technique to calculate depth. The relatively short ranges needed in an industrial environment mean that triangulation or stereo techniques are particularly appropriate, and moreover active lighting may be employed to solve the correspondence problem. The accuracy of the resulting depth measurements from stereo depends then on the accuracy of edge location in the individual images. The problem being addressed in this paper is the assessment and improvement of edge location accuracy in grey level images.

The work reported here is part of a collaborative effort looking at an area of joint interest to two institutions. Cambridge University Engineering Department is developing a high speed, depth mapping device suitable for robotic assembly, and previous work [1] has shown the feasibility of using an array of cameras in conjunction with a Hough transform–based algorithm for near real–time acquisition of a depth map. This work was reported at AVC '86, and is now being developed to improve its accuracy. Other work at B.Ae. Sowerby Research Centre is concentrating on industrial inspection and consequently needs high accuracy, without too much regard for the acquisition and processing time. Hence both authors have a need for this study on the sources of error in depth maps.

The structure of this paper is to first set out the possible sources of error in edge location, and then to describe the approach used to assess and calibrate them. This is followed by a description of the implementation of a real–time correction algorithm in VLSI which is implemented using a pipelined set of 2 input look–up tables and dual–port RAM buffers.

2 Errors in Imaging Systems

There are many sources of error in a vision system that constructs an edge map for use in determining depth. A more detailed list of these may be found elsewhere [2]. In this section, the causes of error and their relative importance are discussed.

2.1 Lens Distortion

For a typical lens used by a vision system, straightforward calculations may be used to measure the distortion. If the lens is assumed to be circularly symmetric, which will be true in most cases, the single parameter of apparent versus true field angle

AVC 1987 doi:10.5244/C.1.20

may be used to estimate the errors. This will adequately model the effects of barrel or pincushion distortion and spherical aberrations. The focus error may further be estimated by fitting a Gaussian to the image of a point source.

2.2 CCD Array Errors

There are two sorts of error that occur with CCD arrays; spatial noise and temporal noise. Variation on the size of individual elements, and the thickness of the polysilicon electrodes affect the sensitivity and spectral response of the pixel. (Polysilicon absorbs more light of shorter wavelengths.) These errors are principally caused by lens effects in the optical printing process used for chip manufacture, as well as process variation in generating uniformly thick layers. These variations lead to a fixed pattern dark current, which is highly temperature dependent. Both these causes result in effects which can be significant when comparing the sensitivity of widely separated pixels, and are typically of the order of 2–3% [3]. The quantum efficiency is also affected by surface reflection characteristics and material inhomogeneities and is a function of the wavelength. The spatial, or fixed–pattern noise on a CCD chip can largely be calibrated out. Another source of fixed–pattern errors is due to the charge transfer efficiency of the chip which results in a small proportion of the charge being left behind at each shift operation. This leads to a position–dependent error in the measured charge and reduces the effective resolution of the sensor.

The temporal noise arises from the shot noise, (statistical variation) in both the dark current and the incident image. This is a 'hard' limit on the signal to noise ratio of the detector, and to this must be added any noise occurring in the charge to voltage amplifier or subsequent transmission lines.

Another possible source of error is the Analogue to Digital Converter (ADC), which gives a quantisation error in the least significant bit. In typical images an 8 bit ADC is used which gives a rounding error of 0.4% . For many applications, this accuracy is sufficient (e.g. humans only have about 6 bit accuracy) and commercial cameras are usually designed with this in mind. However, the temporal noise at the output of the CCD array has been measured at around 65–68dB, [3,4], which corresponds to a useful 12 bit signal, and means that significantly higher accuracy than commercial cameras achieve should theoretically be possible.

2.3 Image Processing Limitations

The spatial quantisation of the CCD array limits the maximum spatial frequency that may be unambiguously detected to the Nyquist limit i.e. $1/(2 \times \text{CCD}$ element spacing) by the sampling theorem. If high spatial frequencies occur in the image, they appear as aliasing errors i.e. they appear as an artefact at a lower spatial frequency. This is another 'hard' limit on the maximum resolution possible, and is likely to affect the accuracy of edge location.

Having obtained a sampled image, edges are often found by some combination of blurring and differentiation. The operation of differentiation, however, amplifies the high frequencies in the image, which includes noise from the imager, and causes inaccuracies in edge location to occur. Studies in noise effects [5,6] show that the uncertainty in the position of zero crossings varies linearly with signal/ noise ratio. On typical isolated edges in synthetic data with 3% noise, the standard error on zero crossings is 0.03 pixels using an interpolative technique.

When all these effects are examined more closely, it is found that most can be calibrated out, at least theoretically. The two main sources of error that cannot be dealt with in this way are the signal/noise ratio, and the spatial quantisation of the sensor array.

3 Approach used to Calibrate Imager

In this section, the hardware used in this work is described, followed by an explanation of a series of experiments that measure the inaccuracies described in section 2. Finally, a comparison with a standard method of calibration [7] is made.

3.1 Use of a High Fidelity Camera

In order to perform the experiments described in the rest of this section, it is necessary to have the capability of accessing each individual CCD element. Conventional imaging systems that use the RS170 video format do not generally have this facility for the following reasons:

Firstly, there is a difference between the number of elements on each line of the CCD and the number of memory locations on each line of the framestore, which means that the response from a single element is smeared over more than one location in memory. Secondly, there are a variety of timing errors in the electronics which lead to apparent motion between successive frames (sometimes referred to as camera jitter). Lastly there is a degree of bandlimiting in the video circuitry that tends to smooth out the image between capture and storage.

For these reasons, a high fidelity image capture system was built that is based on a Sanyo CCD evaluation board [4] interfaced to a personal computer. The charge coming out of the CCD array is converted into digital form without first being put into a video format. The resultant data stream is then fed directly into a framestore, together with addressing information, from which it can be accessed from the host PC. The timing of this sequence of events is controlled by the camera clock which means that each CCD element is mapped into a separate location in the framestore, with no bandlimiting effects.

The next subsection describes experiments performed with this imaging system. The purpose of these were to calibrate the high-fidelity camera, and to investigate the relative magnitude of the different errors described in section 2. The inaccuracies that have been modelled and estimated here fall into two categories, and for each of these categories a separate experiment was devised to measure certain parameters. The experiments were:

i) Overall CCD response variation.

Three factors were measured that contribute towards the overall variation. These are :

a) Linearity of response to incident light.

b) Spatial variation of response of CCD.

c) Temporal variation of response of CCD (noise).

For this experiment the chip was illuminated by a white light, pinhole source with a varying number of neutral density filters to attenuate the beam. A series of images was then captured for each light level on a frame store, and the mean and standard deviation over the series of images for each pixel was recorded. The experiment was repeated after moving the chip set sideways in the beam in order to capture any gross variations in the intensity of the incident light. The results of this experiment are shown in Figure 1.

From this table it was concluded that the temporal noise was far larger than should arise purely from the CCD array. The equipment being used did not have a high enough signal/noise ratio to accurately measure the CCD noise and work is now in progress to build a low-noise version of this equipment with a 12 bit ADC.

Normalised Incident Light	Mean Detected Light (grey levels)	Std. Dev. of Detected Light in Time (grey levels)
1.0	43.1	1.1
0.73	34.8	0.76
0.62	23.2	1.1
0.49	16.8	1.0
0.38	12.5	1.0

Figure 1: Table of Variations in Detected Light

ii) Lens Calibration.

Under the assumption of circular symmetry described in section 2, it is necessary to construct a graph of real to apparent field angle to model the lens. In order to do this measurement, the lens and detector array were mounted on a highly accurate turntable. This was then illuminated with a pinhole source. By varying the angle of the turntable through a known amount, and measuring the position of the received spot of light, the graph can be calculated. An important assumption was that the graph was linear in a small neighbourhood of the optical axis, enabling the geometry of the detector array to be calculated from these points. The point of normal incidence of the beam onto the array was calculated by initially substituting a plane mirror for the array, and adjusting the turntable until the beam shone back on itself.

The result of this experiment is shown in Figure 2.

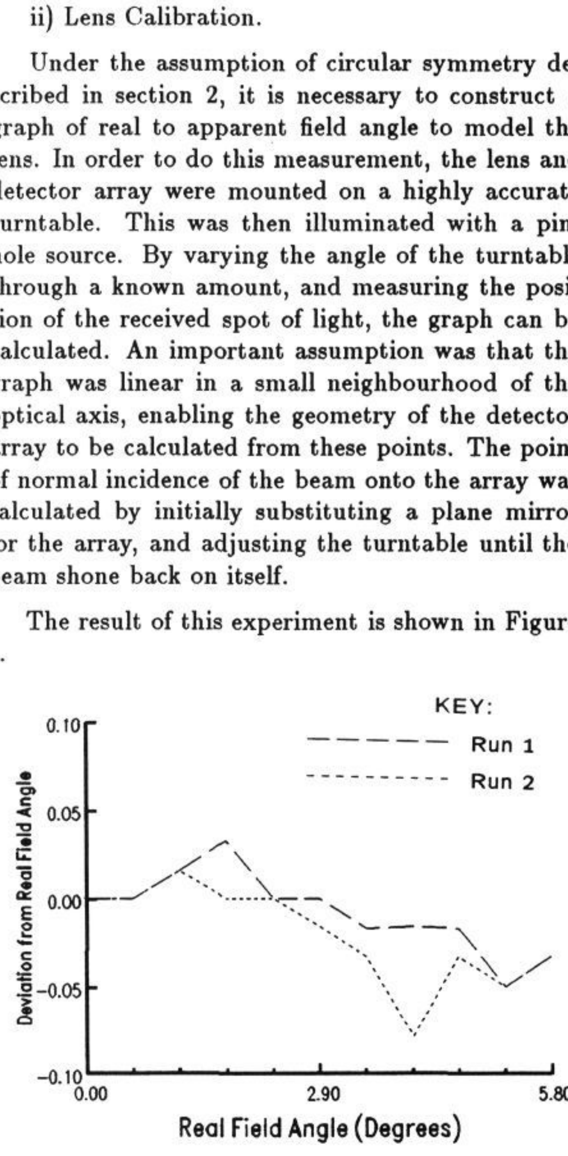

Figure 2: Graph of Deviation from Real Field Angle

From this graph it is clear that for this particular lens, which had a 13° field of view, the distortion at the outer edge of the image is less than one pixel of the CCD array used.

3.2 Use of Results

As a consequence of these experiments, the distortion due to the lens has been estimated, and by using high quality video circuitry, it should be possible to estimate the fixed–pattern noise and linearity of response of the CCD array. A conventional calibration, e.g. Tsai [7] can then be carried out to determine the relative geometry of the CCD array to lens, and the transformation between camera co-ordinates and world co-ordinates.

However, important information has been acquired which enables a more accurate calibration to be achieved. The use of a high–fidelity camera has eliminated the uncertainty previously caused by hardware timing errors for scanning and digitisation. In previous work, [6] this has been found to be a major source of error in an imaging system.

The point of intersection between the optical axis and the CCD array may be found by repeating experiment ii) with the lens and array in their usual geometry and noting the image point for normal incidence. The distortion of the lens is modelled by Tsai as a radial distortion of $\kappa_1 \times r^2 + \kappa_2 \times r^4$, where κ_1 and κ_2 are parameters determined by a steepest descent optimisation procedure. The data obtained in experiment ii) for a given lens can either be used to determine κ_1 and κ_2 , or implemented as a calibration table as described in section 4. However, for the lens used in the experiment, the distortion is negligible.

When calculated accurately, the fixed pattern noise of the CCD array can be used to improve edge accuracy. The fixed pattern noise variations in the Sanyo camera have been quoted [3] at 2–3%, and this figure is significant when using interpolative edge–finding techniques. The temporal noise estimate can be used to allocate uncertainty to edges or optical flow vectors [6], where it is important to know the signal/noise ratio.

Information gained as a result of doing the experiments may be used to construct a mapping between the incoming image and a corrected version, that takes into account lens distortion and variation in the response of the CCD array. The next section describes an implementation of this mapping at video rates.

4 Implementation of Calibration as a Video Rate Correction of the Image

The objective of implementing camera calibration and edge detection in hardware is to offload the burden of low–level vision analysis from the host computer. The architecture chosen consists of three stages :

i) Pixel to pixel intensity corrections.

ii) Geometric correction using bilinear interpolation.

iii) Convolution filter to perform edge detection.

The fixed pattern spatial noise and the distortion due to the lens can be considered as a mapping of the input image onto a corrected image. This mapping can be calculated for each pixel in the input image in advance from the model deduced from experiments i) and ii), and will remain fixed from then on. For this reason, a lookup table can be used to perform this mapping. The effect of non–integer co-ordinates for the resampling addresses can be reduced by the use of a bilinear interpolation between pixels. This will not result in significant loss of accuracy, even for sub-pixel resolution, provided that the incoming image is band–limited.

Having decided to map the image geometry and intensity correction algorithms onto a dedicated piece of hardware, it was then necessary to decide on the specification of the processing elements. The directed graph representation of the algorithm is mapped onto a pipelined sequence of 2 input/1 output look–up tables. The choice was based on the need to perform a single operation at a video scan rate of 250 ns (i.e. for a 256 by 256 window scanned at 50 Hz). A further requirement was that the system should be built from readily available components, to minimise the cost, since one of the objectives of the project was to produce an affordable, portable industrial depth sensor. Since the low level processing used for this design needs only a short data wordlength (i.e. 8 bits at most), this enabled the use of a recently developed 64K by 8 EPROM as a look–up table for the image.

Each processing cell therefore consists of a 64K by 8 EPROM followed by an octal D flip–flop so that the operations can be pipelined. The data out lines of one EPROM are connected to 8 bits of the address lines of the next EPROM, and so on. The other operations used by the algorithm are line delays for providing nearest neighbour information, and random access memory for resampling and buffering images if using a timeshared sensor for acquiring images. A dynamic RAM with dual ports was used to perform both of these functions. A block diagram of the design is shown in Figure 3.

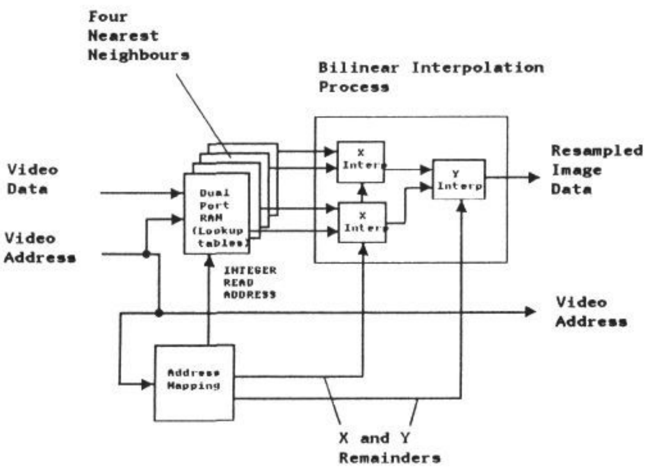

Figure 3: Geometric Correction

A point to note is that the dual-port RAM design allows data to be read into the 4 nearest neighbour buffers in parallel, while allowing simultaneous access to the previous frame of data.

5 Conclusions and Further Work

This work is continuing with the construction of a 12 bit digital camera, which it is hoped will be noise-limited by the CCD array. However, the principle of using a digital camera has been established, limited at present by the video circuitry, and the version described in this paper has been calibrated. It is envisaged that a high-fidelity camera like the one proposed will be limited by the spatial quantisation of the CCD array, and therefore techniques such as moving the array by a fraction of a pixel (jittering) are also being studied to overcome this.

6 References

1. S.Wright 'Hough Transform Analysis of Data from a Planar Array of Image Sensors' Proceedings of AVC '86.

2. A.Hutchings 'An Analysis of Edge Location Accuracy' GEC Hirst Research Centre, Computer Systems Research Memo. No. 85, 1983.

3. J.Smith, Head of Electro-optics Dept., B.Ae. Stevenage, Private Communication.

4. Sanyo Electric Co. Ltd. 'Evaluation Board for Operating Frame Transfer CCD Image Sensor' Operating Manual.

5. I.Overington 'Edge Detection and Local Motion Sensing by VISIVE in the Presence of Random Noise' B.Ae. Optical Flow Consortium Report (JS10677) August 1986.

6. D.Hutber 'A Study of Accuracy Limitations on Optical Flow' B.Ae. Optical Flow Consortium Report - to be published.

7. R.Y.Tsai. 'An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision', Proc. IEEE Conf. on Comp. Vis. and Patt. Recog., 1986.