Appeared in Int'l Conf. on Pattern Recognition, Vol. 1, pp. 896-899, Barcelona, 2000

Estimating Sensor Orientation in Cameras

Manoj Aggarwal and Narendra Ahuja
University of Illinois at Urbana-Champaign, Beckman Institute
405 N. Mathews Ave, Urbana, IL 61801, USA
E-mail: {manoj,ahuja}@vision.ai.uiuc.edu
Internet: http://vision.ai.uiuc.edu/

Abstract

For most imaging cameras, it is desirable that the sensor plane be perpendicular to the optical axis. Such an orientation ensures that the imaging configuration is perspective and planar scene objects perpendicular to the optical axis can be focussed in their entirety. In this paper, we have presented an image processing based method to estimate and subsequently correct the sensor tilt with precision. We propose to measure the tilt by measuring the variation of defocus in an image of a planar calibration chart placed perpendicular to the optical axis. We have shown that the proposed defocus based method is inherently more accurate than geometry based techniques which estimate tilt by measuring the deviation in the geometry of a scene pattern. We have analyzed the sensitivity of the tilt estimates to errors in the experimental setup and have shown that the proposed technique is quite robust to errors even as large as 1 degree in the orientation of the calibration chart.

1. Introduction

Accurate positioning of the sensor in an imaging system such as a CCD camera is critical for a number of computer vision tasks. Specifically, it is desirable that the sensor plane be perpendicular to the optical axis. This considerably simplifies the relationship between the 3-D world coordinates of the scene and the 2-D coordinates of the corresponding image, which, in turn, makes the camera calibration task less complex. In addition, for cameras with finite depth of field, a sensor at such normal orientation ensures that planar scene objects placed perpendicular to the optical axis can be focussed in their entirety.

In this paper, we present a method to place the sensor in the normal orientation with precision. The method, involves taking image of a special planar calibration pattern,

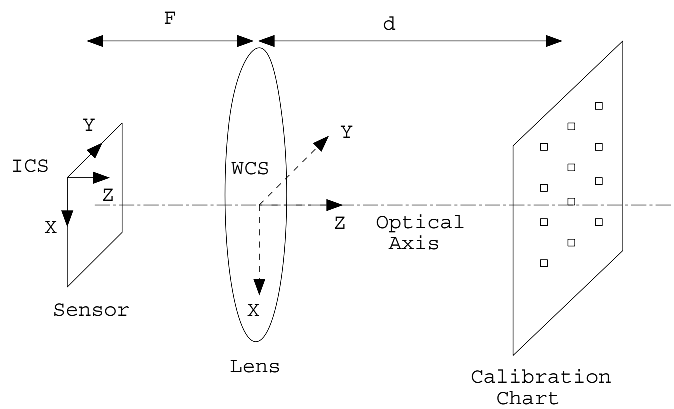

Figure 1. Experimental setup

placed perpendicular to the optical axis. The basic idea used is that the image of a plane perpendicular to the optical axis, should be uniformly defocussed for a sensor in normal orientation. If the sensor is not in the normal orientation, the image blur varies linearly over the image. If the linear blur variation can be estimated, it would yield an estimate of the sensor orientation or *tilt*. The tilt in the sensor can then be corrected by using actuators or manually. Geometric techniques, for example, simple extensions of techniques proposed by Faugeras [4], Maybank [5] and Triggs [11] can also be used to estimate the sensor tilt. A non-linear geometric camera calibration technique, which includes estimation of sensor tilt has been proposed and implemented by Castano [1]. However, as we have shown in this paper, geometric techniques may not give very accurate results due to the discrete nature of the CCD sensor.

The parameters of the linear variation in image blur can be estimated using a modification of the depth from defocus techniques. The depth from defocus techniques have been abundantly researched in literature [2, 3, 6, 7, 8, 9, 10, 12, 13]. Most of the techniques use multiple images (usu-

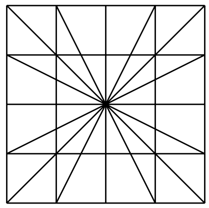

Figure 2. A sample textured pattern

ally 2) of the same scene, under different camera settings such as focal distance, aperture and distance between sensor and lens. The difference in the images due to known change in camera settings can be used to estimate the unknown depth. In this paper, the defocus method is turned around to estimate the *unknown* change in the camera settings from the images, when the object depth is *known*. Consider the experimental setup shown in Fig. 1. The planar calibration chart consists of a small but complex/textured pattern placed repetitively on a regular grid. A sample textured pattern that can be easily generated is shown in Fig. 2. The calibration chart serves as a grid of planar identical objects, and by placing the chart perpendicular to the optical axis, the objects will also be equidistant from the lens. The image of these objects on the sensor will be composed of similar subimages placed repetitively on a regular grid. These subimages are the images of each of textured patterns and they will differ mainly in their extent of defocus. These subimages can be considered as multiple instances of the same object under different camera settings, which in this case is the distance of the sensor from the lens. Any depth from defocus algorithm can now be applied to estimate the blur parameters and these can be subsequently used to find the perpendicular distance between the location of the subimage and the lens. The array of perpendicular distances can then be used to fit a planar surface, which is an estimate of the orientation of the sensor surface.

This paper is organized as follows. In Section 2, we present the proposed algorithm to estimate sensor surface orientation. In Section 3, we compare the performance of geometrical techniques and depth from defocus techniques in estimating the sensor tilt. In Section 4, we analyze the sensitivity of sensor tilt estimates to errors in the experimental setup. Section 5 presents concluding remarks.

2. Estimating the sensor tilt

Consider the experimental setup shown in Fig. 1. Let $f(x,y)$ denote the image obtained of the calibration pattern. Let $\{f_{i,j}(x,y),\ \ 0 \le i < M, 0 \le j < N\}$ be the set of subimages obtained by cropping regions of equal size corresponding to each of the textured patterns. Let (x_i, y_j) denote the centers of the $(2P+1) \times (2Q+1)$ textured patterns in $f(x,y)$ then

$$f_{i,j}(x,y) = f(x-x_i, y-y_j) \quad -P \le x \le P\ , -Q \le y \le Q \tag{1}$$

where each pattern contains a $2P+1 \times 2Q+1$ subimage to be analyzed for defocus. If we assume that the maximum radius of blur circle in the image does not exceed k pixels, it may be adequate to include a border of k pixels around every cropped subimage. These subimages are then processed by a depth from defocus algorithm. In our experiments we used Pentland's method to estimate the blur parameters [6].

The subimages $f_{i,j}(x,y)$ are images of identical objects except for the amount of defocus. If we neglect that the amount of blur variation within a subimage and represent the defocus over the entire subimage by a single kernel then $f_{i,j}$ can be modeled as

$$f_{i,j}(x,y) = \hat{f}(x,y) * H(x,y,\sigma_{i,j}) \tag{2}$$

where $\hat{f}(x,y)$ would be the object image if there was no blur, '$*$' represents 2-D convolution, and $H(x,y,\sigma_{i,j})$ is the blurring kernel with one parameter $\sigma_{i,j}$. Under the above blurring model and assuming a Gaussian blurring kernel, Pentland has shown that given two images $f_{i,j}(x,y)$ and $f_{k,l}(x,y)$, the amplitudes of their Fourier transforms are related as

$$2\pi^2(u^2+v^2)(\sigma^2_{i,j} - \sigma^2_{k,l}) = \ln \mathcal{F}_{i,j}(u,v) - \ln \mathcal{F}_{k,l}(u,v) \tag{3}$$

where $\mathcal{F}_{i,j}(u,v)$ represents the amplitude of the Fourier transform of $f_{i,j}(x,y)$. The set of equations in (3), for different spatial frequencies u and v, has only one unknown $(\sigma^2_{k,l} - \sigma^2_{i,j})$ and can be solved for easily, for example, by using singular value decomposition to get a best fit solution for it.

Using the method outlined above, we can compare each of the subimages $f_{i,j}(x,y)$ to a fixed subimage denoted by $f_{i_o,j_o}(x,y)$, to obtain an array of values $e_{i,j} = (\sigma^2_{i,j} - \sigma^2_{i_o,j_o})$. Since the blur σ varies linearly with distance from the lens, which in turn varies linearly across the image, the variation of $e_{i,j}$ over the grid of subimages is quadratic. Let the variation of blur parameter be given by

$$\sigma_{i,j} = ax_i + by_j + c, \tag{4}$$

then,

$$\sigma^2_{i,j} - \sigma^2_{i_o,j_o} = a^2x_i^2 + b^2y_j^2 + 2c(ax_i + by_j) + 2abx_iy_j - \hat{c}, \tag{5}$$

where

$$\hat{c} = (ax_{i_o} + by_{j_o})^2 + 2c(ax_{i_o} + by_{j_o}) + 2abx_{i_o}y_{j_o}$$

We can estimate a^2 and b^2 by fitting a quadratic surface to the values in array $e_{i,j}$ over the grid of subimages denoted by (x_i, y_j). The estimates of a and b yield the sensor tilt.

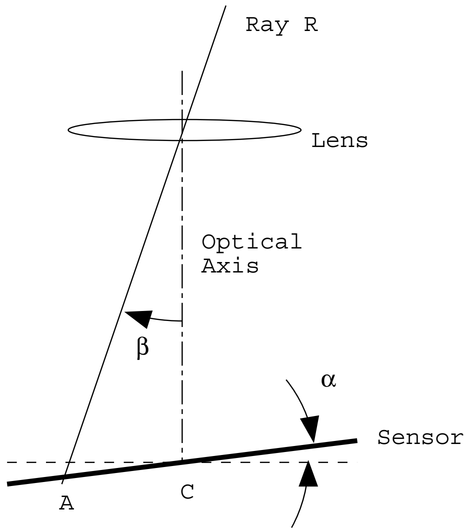

Figure 3. Geometry of a 2-D camera

3. Accuracy of geometric v/s defocus based techniques

In this section, we will compare the performance of geometric techniques and the proposed technique. Geometric techniques measure the amount of tilt by the amount of geometric distortion it causes in the image of a known pattern. Consider a simple 2-D (linescan) camera configuration shown in Fig. 3. It shows a ray R passing through the lens center at an angle β to the optical axis, and touching the sensor at position A. The sensor makes an angle of $(90+\alpha)$ with the optical axis. Let F be the distance between the lens center and the optical center of the sensor. Then by simple geometric analysis it can be shown that the distance $AC = F\frac{\sin\beta}{\cos(\beta+\alpha)}$. Thus the amount of geometric distortion (δ), i.e., the displacement in the image plane location of R, caused by tilting the sensor by angle α is given by

$$\delta = F(\frac{\sin\beta}{\cos(\beta+\alpha)} - \tan\beta) \tag{6}$$

Due to the discrete nature of the CCD sensor, a minimum distortion must occur in order to be detected. Let us assume that the minimum distortion required is 1 pixel which is typically 10 microns on a CCD sensor. Then, for $\beta = 10^o$, which is the typical half-angle of view for a 25mm CCD camera, the value of $\alpha = 0.744^o$. This implies the sensor tilt should be at least 0.744^o for the geometric techniques to detect it !

Now consider the proposed method. Assume that we are estimating the tilt using the calibration pattern at a distance d from the lens. Then, the distance d, blur circle diameter σ,

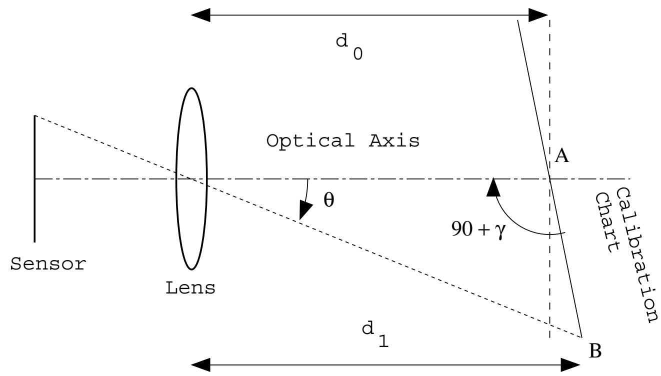

Figure 4. 2-D imaging geometry with a tilted calibration plane

and distance v of the sensor from the lens are related by [6]

$$d = \frac{Fv}{v - F - \sigma f} \tag{7}$$

where F is the focal length and f is the f-number. Equation (7) can be rewritten as

$$v = \frac{dF + d\sigma f}{d - F} \tag{8}$$

Thus when $\sigma = 0$, $v_o = v = \frac{dF}{d-F}$. The expression $(v - v_o)$ denotes the depth of field of the camera for an object at a distance d and is given by

$$v - v_o = \frac{d\sigma f}{d - F} \tag{9}$$

Assuming that the depth from defocus algorithm is capable of detecting defocus when $\sigma \geq 1$ pixel, then for $F = 25$mm, $f = 2$ and $d = 400$mm, $v - v_o = 21.3$ microns. Typical size of 2/3" CCD is 10mm x 10mm and thus the depth from defocus algorithm will be able to detect sensor tilt even when α is as low as $\tan^{-1}\frac{21.3}{10000} = 0.12^o$. Comparing this number with that obtained using geometric techniques, the proposed method is significantly more accurate. In fact, one can further improve the accuracy by increasing the distance between the calibration chart and the lens or even the aperture size.

4. Sensitivity of defocus based technique to errors in experimental setup

In this section, we derive the variation in blur due to a tilt in the calibration plane itself. In order to decouple the variation in blur due to the sensor tilt and the tilt in the calibration plane, assume that the sensor is perpendicular to the optical axis, while the calibration plane makes an angle

$(90+\gamma)$ with the optical axis, as shown in Fig. 4. For simplicity, we assume a 2-D geometry. Let d_0 be the distance of the calibration plane from the lens along the optical axis and θ be the half-angle subtended by the calibration pattern at the lens center. The relationship between the blur σ_0 in the image of point A and its distance d_0 from the lens is given by [6]

$$d_0 = \frac{Fv}{v - F - \sigma_0 f} \quad (10)$$

Let d_1 be the perpendicular distance between the farthest point (B) on the calibration chart and the lens. Let the blur associated with the image of point B be σ_1. Then,

$$d_1 = \frac{Fv}{v - F - \sigma_1 f} \quad (11)$$

Using equations (10) and (11), it can be shown that

$$\sigma_1 - \sigma_0 = \frac{Fv}{f}\left(\frac{d_1 - d_0}{d_1 d_0}\right) \approx \frac{Fv\tan(\theta)\tan(\gamma)}{f d_1} \quad (12)$$

By using simple lab instruments one can easily align the calibration pattern to within 1 degree error, thus $\gamma < 1^o$. For $F = 25$mm, $d_1 = 400$mm, $f = 2$ and $\theta = 10^o$, the maximum variation in the blur parameter over the sensor is 2.45 microns, which is one-quarter of size of pixel in a typical 2/3" CCD. This variation can be further reduced by increasing the distance between the calibration pattern and the lens. This analysis shows that the proposed technique is quite robust to experimental errors.

5. Conclusions

We have proposed a new technique to estimate the orientation of the sensor plane. This technique is based on the fact that for a planar object placed perpendicular to the optical axis, its image exhibits a range of blur determined by the orientation of the sensor. This orientation can thus be determined using any method that can estimate blur. For example a depth from defocus algorithm can be used for estimating the variation in blur. We have also shown that the proposed technique is significantly more accurate than geometric techniques. An analysis of the sensitivity of estimated orientation to errors in the experimental setup proves that the proposed technique is robust.

References

[1] A. Castano. *Range from Non-frontal Imaging Camera*. PhD thesis, University of Illinois at Urbana-Champaign, 1998.

[2] S. Das and N. Ahuja. Performance analysis of stereo, vergence, and focus as depth cues for active vision. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 17(12):1213–1219, December 1995.

[3] J. Ens and P. Lawrence. An investigation of methods for determining depth from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(2):97–107, February 1993.

[4] O. D. Faugeras. Stratification of the three-dimensional vision: projective, affine, and metric representations. *Journal of Optical Society of America A.*, 12(3):465–484, March 1995.

[5] S. J. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

[6] A. Pentland. A new sense of depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):523–531, July 1987.

[7] A. N. Rajagopalan and S. Chaudhuri. A block shift-variant blur model for recovering depth from defocused images. In *International Conference on Image Processing*, volume 3, pages 636–639, 1995.

[8] A. N. Rajagopalan and S. Chaudhuri. A variational approach to recovering depth from defocused images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(10):1158–1164, October 1997.

[9] M. Subbarao and G. Surya. Depth from defocus: A spatial domain approach. *International Journal of Computer Vision*, 13(3):271–294, 1994.

[10] M. Subbarao, T.-C. Wei, and G. Surya. Focused image recovery from two defocused images recorded with different camera settings. *IEEE Transactions on Image Processing*, 4(12):1–15, December 1995.

[11] B. Triggs. Autocalibration from planar scenes. In *European Conference on Computer Vision*, volume 1, pages 89–105, 1998.

[12] M. Watanabe and S. K. Nayar. Minimal operator set for passive depth from defocus. In *Conference on Computer Vision and Pattern Recognition*, pages 431–438, 1996.

[13] Y. Xiong and S. A. Shafer. Depth from focusing and defocusing. In *Conference on Computer Vision and Pattern Recognition*, pages 68–73, 1993.