

Sensor Errors and the Uncertainties in Stereo Reconstruction

Gerda Kamberova and Ruzena Bajcsy
E-mail: kamberov@cis.upenn.edu, bajcsy@cis.upenn.edu
Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104

Abstract

An important objective in the evaluation of algorithms with sensory inputs is the development of measures characterizing the intrinsic errors in the results. Intrinsic are those errors which are caused by noise in the input data. The particular application which we consider is 3-D reconstruction from stereo. We demonstrate that a radiometric correction of the images could improve significantly the accuracy. We propose a confidence interval approach for quantifying the precision. We also illustrate the use of the confidence intervals for the rejection of unreliable 3D points.

1 Introduction

1.1 Motivation

Our work is inspired by the call in the computer vision community for performance evaluation methods [11, 12, 32]. In this paper we discuss errors in 3-D reconstructions from stereo induced by sensor noise. Such errors we call *intrinsic*. Any undesired feature causing discrepancies in the digital image is considered noise. Systematic noise (errors) effect the **accuracy** of the vision algorithms, i.e., the closeness of the results to the true values. Random errors primary effect the **precision**, i.e., the variability in the results due to random noise in the digital images. Both the accuracy and the precision of vision systems should be investigated.

For establishing the **accuracy**, the results must be compared with ground truth models. General ground truth models are difficult to obtain. Usually simple test examples are used. The best strategy is to compare results with widely accepted ground truth models, if such models exist.

A complete characterization of the **precision** consists of a probability distribution for the output noise. Such distributions are even more difficult to get than ground truth models since there are usually too many factors influencing the results. A popular approach is to report the variance in the output. In that approach, and others related to it (factor analysis, performance analysis, variance components), it is implicitly assumed that the input noise is normally distributed, and that the input-output transformation is well approximated by a linear transformation. Unfortunately, one often uses these methods without verifying the validity of the model assumptions.

A methodology for performance evaluation should include characterization not only of the errors depending on the environmental conditions (i.e. extrinsic errors), but also on the sensor characteristics. *For specifying the performance metrics of a vision system with various hardware configurations, it is necessary to quantify the intrinsic errors.* Our goal in this paper is to stress the importance of understanding the sensor and its effects on the performance of the algorithms.

In the rest of this section we, first, review the video sensor and the stereo algorithm (Section 1.2), and second, overview our contribution and the results presented in this paper (Section 1.3). Section 1.4 outlines the organization of the paper.

1.2 The Sensor and the Stereo Algorithm: An Overview

1.2.1 The Video Sensor

The video sensor consists of a lens, a CCD camera, and a framegrabber. The image is formed on the CCD array of identical sensing elements (*sels* [23]), and then transferred by the framegrabber, in a linear order (line-by-line, sel-by-sel), to the computer memory. The geometric and radiometric uncertainties and discrepancies in the digital image are due to the optics, the CCD camera, and the joint operation of the camera, the framegrabber, and other electronic components.

Camera related noise

The *total random noise* of the CCD has three major components: photon (shot) noise, read noise, and fixed pattern noise. The total noise is dominated by the read noise at low illumination levels, by the fixed pattern noise at high levels, and by the photon shot noise in between [18].

The source of the *photon (shot) noise* is external. This noise is due to fluctuations of the photon flux, and is always present in the data.

The sources of the *read noise* are internal to the CCD camera. The read noise is related to the physics of the camera and the measurement process. Major factors contributing to the read noise are the background noise and output amplifier noise. Noticeable components of the background noise are the dark current (thermally generated charges), and the internal luminescence. In general, dark current doubles with every 8° increase in temperature [5]. To reduce dark current cameras are cooled. Dark current can originate at different locations in the CCD but has, in all cases, to do with the irregularities in the crystal structure of the silicon. This contributes to a fixed pattern noise (FPN) in dark current. Because the background noise is temperature dependent, cameras must be warmed up before use.

The individual sels have their own physical characteristics. The difference in these characteristics is manifested in the FPN which can be observed in dark images¹ and flat fields². The photoresponse nonuniformity (PRNU) of the sels is observed in flat fields. In this paper, we correct for systematic, FPN, in dark current and PRNU. Flat fields and dark images are used for a radiometric calibration of the sensors, Section 4.

Framegrabber related noise

Geometric discrepancies due to the digitization in the framegrabber are aliasing, linejitter,

¹A dark image is the response (readout) of the camera with no access of light to the CCD array.

²A flat field is the response of the sensor to uniform light illumination.

and systematic fall-off of the intensity in a line. An effect in radiometric distortion due to interlacing³ is the shift in gray level between odd and even fields in a frame. Severe distortions occur when the sampling theorem⁴ is violated. The sel spacing puts limitation on the highest frequency in the input beyond which severe distortions occur, and the area of the CCD chip limits the lowest frequencies which can be detected. Another radiometric noise component is modeled by the quantization error [39].

Modeling the sensor

The detailed modeling of the sensor noise is a very hard problem. For that reason, an approximation using a linear model with normally distributed noise is usually used. As we demonstrate, in Section 5, even with this simple model, the precision of the stereo algorithm can be satisfactory quantified, and the accuracy can be significantly improved.

Other radiometric models are presented in [38, 14]. The model in [38] is very attractive with its simplicity: the noise is divided into three components: (a) signal-dependent; (b) signal-independent and position-dependent; and (c) signal- and position-independent. It is used in the context of Bayesian framework for image restoration. The model in [14] is detailed. It capturing the various camera and framegrabber related noise sources. But, the actual procedure for the practical estimation of the various noise parameters, and for the radiometric correction of images is very much simplified. After the simplifications, the model is practically reduced the popular class of linear models with additive noise (similar to the one we use).

References

CCD array, cameras, and displays are analyzed in [5, 40, 16, 18]. Distortions related to the optics are addressed in [37]. Discretization effects are in the center of the signal processing literature [17]. A brief and clear overview of the video signal formats is given in [23, 2]. Various sources of geometric distortions in the video sensor are analyzed in [23]. A detailed study of different framegrabber architectures and synchronization mechanisms, and related geometric and radiometric distortions is given in [4], and a comprehensive discussion of linejitter problem, its sources and detection, in [23, 2, 4].

1.2.2 The Stereo System

By the late 17th century the use of a pair of images obtained from different view points (from both eyes) in perceiving 3-D scenes was recognized [15]. But it was not until the 1970s that the power of stereo in estimating 3-D, even without recognizable features in the images, was established [19], and a computational theory of stereo proposed [25, 26]. Since, stereo reconstruction has been an active area of research in computer vision. For a review on stereo and matching see [9], or more recently [33, 30, 3].

The main principle underlying 3-D reconstruction from stereo is a simple one. Given the pinhole camera model⁵, the camera parameters, and the configuration geometry of the pair of cameras: if a correspondence is established between two points of interest, F_L and F_R , in the left and the right images, the position of the 3-D feature corresponding to F_L and

³Interlacing is a standard mode of operation of cameras.

⁴The input signal can be fully reconstructed from the samples, if the input signal frequency, f , is at most half of the sampling frequency.

⁵Images are formed on the image plane by perspective projection with center of projection the optical center of the camera.

F_R is the intersection of the optical rays $F_L C_L$ and $F_R C_R$, where C_L and C_R denote the optical centers of the two cameras. Depending on the types of features used for establishing correspondences, matching algorithms, sensor suites, and environmental assumptions vary.

The stereo algorithm we used was developed and implemented by Radim Šára, [34]. It uses: area-based matching; multiple pairs of verging, strongly calibrated cameras; assumes lambertian, opaque surfaces, and enforces the usual ordering and disparity gradient limit constraints [9, 41]. The features in area-based matching are the pixel intensities.

An intrinsic geometric property of stereo images is that the projections of the same 3-D point in left and right images lie on corresponding *epipolar lines* (resulting from the intersection of the two image planes with the plane defined by the 3-D point and the optical centers of the cameras), [27]. If the cameras are in a normal (standard) configuration⁶, the epipolar lines coincide with the horizontal (u) coordinate lines⁷. Thus for a given pixel at position (u, v) in the left image, the search for its corresponding pixel in the right image is in 1-D, among the pixels at positions (u, i) , where only i varies. Single pixel values are unstable as matching primitives, so small windows (neighborhoods) around the pixels are selected. For a fixed pixel (u, v) in the left image, the matching procedure establishes the pixel, (u, v^*) , in the right image whose neighborhood intensity pattern is most “similar” to that of (u, v) . This correspondence is established provided that the 3-D feature projected at (u, v) is in the field of view of both cameras. The similarity is measured by a modified normalized cross correlation [29],

$$\text{Cor}(\mathbf{I}_L^w, \mathbf{I}_R^w) \stackrel{\text{def}}{=} \frac{2\text{Cov}(\mathbf{I}_L^w, \mathbf{I}_R^w)}{\text{Var}(\mathbf{I}_L^w) + \text{Var}(\mathbf{I}_R^w)}, \quad (1)$$

where \mathbf{I}_L^w and \mathbf{I}_R^w denote the pixel intensities over the correlation windows in the left and right images respectively, and Var and Cov denote the spatial sample variances and covariance over the windows. The *integer disparity*, $\mathbf{D}(u, v)$, at (u, v) is defined as the shift $(v - v^*)$.

The major steps in the stereo reconstruction are: (1) integer disparity map computation; (2) subpixel disparity map recovery; and (3) reconstruction.

(1) The result of the matching procedure is an *integer disparity map*, \mathbf{D} : for each (u, v) in the left image if the correspondence is resolved, $\mathbf{D}(u, v)$ is assigned the disparity value, otherwise it is marked as *undefined*.

(2) For many applications, \mathbf{D} is too coarse (a surface recovered from integer disparities would be step-like), so a subpixel disparity, $\tilde{\mathbf{D}}$, is approximated. In [34] $\tilde{\mathbf{D}}$ is estimated based on an approach given in [8]. The mapping between the intensity values of corresponding positions in the image planes is modeled by:

$$I_R(\mathbf{x}) = BI_L(T_x(\mathbf{x} - d(\mathbf{x}))) + M \quad (2)$$

$\mathbf{x} = (u, v)$ denotes image position, $d(\mathbf{x})$ the subpixel disparity, T_x models the local perspective distortion between corresponding neighborhoods, and M and B are scalar parameters accounting for radiometric differences in the cameras over the windows of interest. Under the assumption of local planarity of the surface being imaged, and linearization of the

⁶i.e., the cameras are oriented fronto-parallel in the same direction, the image planes are coplanar, and cameras are aligned so that the horizontal coordinate lines in both image planes are collinear, and the optical centers are at the same distance from the corresponding image planes.

⁷Since we use verging cameras, the images were first rectified, [1], thus brought to a standard configuration.

functions involved, in local coordinates, for (u, v) in the window, (2) is reduced to

$$I_R(u, v) = (d_\varepsilon B, \gamma, M) \left(\frac{\partial I_R}{\partial v}(u, v), I_L(u, v), \frac{\partial I_L}{\partial v}(u, v)v, 1 \right)^t, \quad (3)$$

where d_ε is the error in the subpixel disparity correction for the center of the window in the left image, γ relates to the distortion T_x (as in (2)), and t denotes transpose of a vector. The derivatives $\frac{\partial I_R}{\partial v}(u, v)$ and $\frac{\partial I_L}{\partial v}(u, v)$ are approximated by finite differences. Equation (3) is applied over the reference window, and then solved by least squares for $\theta = (d_\varepsilon, B, \gamma, M)$. Thus, $\tilde{D}(u, v) = D(u, v) + d_\varepsilon$.

(3) The 3-D reconstruction is a linear transformation in a projective space. Let Q be the projective 4x4 matrix recovering the 3-D point in projective space, $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{K})$, i.e., $(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{K})^t = Q * (u, v, v - \tilde{D}(u, v), 1)^t$. We rewrite this in the form

$$(\tilde{X}, \tilde{Y}, \tilde{Z}, \tilde{K})^t = S * (u, v, 1, \tilde{D}(u, v))^t \quad (4)$$

where the i th row of S is $S(i, :) = (Q(i, 1), Q(i, 2) + Q(i, 3), Q(i, 4), -Q(i, 3))$, $i = 1, 2, 3, 4$. Then the Euclidean coordinates of the reconstructed point which was projected at (u, v) are calculated from

$$(X, Y, Z)^t = \frac{1}{\tilde{K}} (\tilde{X}, \tilde{Y}, \tilde{Z})^t. \quad (5)$$

1.3 Characterizing the Intrinsic Errors in Stereo: Our Contributions

The literature on the errors in 3-D reconstruction from stereo has been primarily focused on the effects of the systematic geometric discrepancies in the digital images (not radiometric), the camera configuration geometry, and the camera calibration parameters, [6, 7, 10, 36]. We address two new issues:

1. The effect of the *random sensor noise on the precision* which we quantify using confidence intervals. (Some necessary terminology and a summary of the approach are given in Section 1.3.1, and details in Section 3.)
2. The effect of the *systematic radiometric sensor errors on the accuracy*. (We show that the accuracy can be improved by the procedure for radiometric correction given in Section 4.)

1.3.1 The Confidence Intervals Approach in Stereo

Definition A confidence interval for a parameter of interest, δ , is an interval $C = [C_{min}, C_{max}]$ with bounds C_{min} and C_{max} which are random variables. The values of the bounds are obtained as a result of an experiment. The theoretical probability of capture (TPC), $\Pr[\delta \in C]$, gives the reliability of the interval, while the size, $C_{max} - C_{min}$, represents the uncertainty in δ . The TPC characterizes the likelihood of computing δ within tolerance bounds even prior to observing the outcome of the experiment.

For example, in the context of stereo, the experiment is the sampling of left and right images and the computation of the 3-D points. If δ denotes a true depth of a 3-D point, the bounds C_{min} and C_{max} could be functions of the disparity. They will be random because of the random noise in the disparity induced from the image noise.

Previous uses of confidence intervals in stereo for quantifying the accuracy

In stereo vision confidence intervals have been used for characterizing the *accuracy* of the estimate of an *average* distance to a plane [24]. The restricted geometric domain, and the particular application in this case, made possible: the development of a ground truth model for the true depth measurement and the design of a probabilistic model for the errors in the average depth estimate from stereo. Under these conditions, it is possible to derive optimal (in minimax sense) confidence intervals for the true depth [21]. These intervals are of fixed size, for a given depth range, and have guaranteed TPC.

The use of confidence intervals for quantifying the precision

In the current paper, we are interested in *dense point set* recovery, and a *wide class of scenes* (like general office scenes, human faces, hands, machine parts, etc), thus the method used in [24] is infeasible. We take a different approach here.

In order to isolate the effect of sensor errors from other noise sources, we induce a *strong random dot texture* on the scenes by using an overhead projector. Thus we restrict the domain of inputs, but do not impose farther conditions on the geometry and the material of the surfaces. In the restricted domain, the integer disparity map based on (1) is stable. For strong random texture scenes the main component of the intrinsic errors is the random one. Our experiments clearly show that the random noise in the digital images has little effect the integer disparity (for 70 – 95% of the pixels, the integer disparities were constant over multiple runs of the matching algorithm on fixed scenes each time taking new pairs of images; the integer disparity is unstable in areas where the fronto-parallel planarity assumption of matching is violated).

We characterize the precision in the reconstruction with confidence intervals (one in each coordinate direction).

Algorithmic steps in deriving the intervals:

- We start with a simple image model with normal noise (described in Section 2).
- Next: (a) we propagate the variance from the images to the disparity using the method from [13]; (b) for a selected TPC, we compute confidence intervals for the disparity variance; and (c) we map these intervals to intervals in the coordinates of the reconstructed points. In Section 3 we review the steps (a)-(c).

1.4 Organization

In Section 2 we present the models, and in Section 3, the confidence intervals quantifying the uncertainties in the reconstruction. Section 4 outlines the radiometric correction procedure used for reducing the intrinsic systematic errors. In Section 5 we show the experimental results, and in Section 6 we summarize the result and draw conclusions regarding the use of the intervals for data reduction and fusion.

2 The Models

2.1 The Sensor Model

In this section we present a sensor model which is consistent with the assumptions made in the stereo algorithm, and which can be used to account for the intrinsic errors in stereo.

We do not attempt to give a general sensor model capturing all noise components⁸.

We assume that the random noise per pixel is spatially varying over the whole image, but locally constant (over the correlation window size). For a pixel (u, v) , $1 \leq u \leq M$, $1 \leq v \leq N$, where MN is the image size, the intensity is modeled by

$$I(u, v; \theta(u, v)) = \beta(u, v)\theta(u, v) + m(u, v) + W(u, v), \quad (6)$$

where $\theta(u, v)$ denotes the true, unobservable (ideal) intensity of the pixel, the gain factor $\beta(u, v)$ model the photoresponse nonuniformity (PRNU), the offset $m(u, v)$ models the fixed pattern noise (FPN), and $W(u, v)$ is a zero-mean normally distributed random variable with variance $\sigma^2(u, v)$. Note that this is the model under which in practice most of the radiometric correction procedures are done [14, 4].

We use this model in two ways: (1) to obtain estimates of the input variance, which later is used in the confidence intervals derivations (Section 3); and (2) to justify theoretically, in [22], the radiometric correction procedure given in Section 4.

Let F_0^i , $i = 1..n$ denote a sequence of dark images, and F_θ^i , $i = 1, ..(n + 1)$ — a sequence of flat fields taken at the same level of illumination. Then the estimates of $m(u, v)$ and $\sigma^2(u, v)$ are obtained using sample means and sample variances as follows,

$$\hat{m}(u, v) = \frac{1}{n} \sum_{i=1}^n F_0^i(u, v), \quad \hat{\sigma}^2(u, v) = \frac{1}{2(n-1)} \sum_{i=1}^n (F_\theta^i(u, v) - F_\theta^{n+1}(u, v))^2, \quad (7)$$

since the mean of the differences of two flat fields is $0_{M \times N}$, and the variance of the difference is two times the variance of the original flat fields.

In experiments, we have used sequences of 100-1000 images for the estimates. Under this noise model, the flat fields used for radiometric correction in Section 4 should be taken at highest possible illumination level. The reason being that the actual signal-dependent component of the noise has higher variance at higher illumination levels which we want to bound.

2.2 The Disparity Model and the Disparity Variance

As discussed in Section 1.2, given the left and right images, I_L and I_R , and an integer disparity map, \mathbf{D} , the subpixel disparity at $\tilde{\mathbf{D}}(u, v)$ is obtained by

$$\tilde{\mathbf{D}}(u, v) = \mathbf{D}(u, v) + \mathbf{D}_\varepsilon(u, v) \quad (8)$$

where $\mathbf{D}_\varepsilon(u, v)$ is the disparity correction computed as described in Section 1.2.2. As stated previously, we aided the matching by projecting strong random texture onto the scene, and have established that under this conditions, the first set of intermediate results where the random sensor noise has a significant effect is the subpixel disparity map. In (8), the integer disparity is deterministic, only the disparity correction is random, thus $\text{Var}\tilde{\mathbf{D}}(u, v) = \text{Var}\mathbf{D}_\varepsilon(u, v)$. In [20], using the image variance estimates, and a propagation of variance method described in [13], we derived the estimates $\sigma_D^2(u, v)$ of $\text{Var}\mathbf{D}_\varepsilon(u, v)$. Thus, in our augmented algorithm, for every pixel in the left image we compute the disparity and its variance.

Under normal assumptions for the image noise, it follows that the subpixel disparity is normally distributed, i.e., $\tilde{\mathbf{D}}(u, v)$ has a normal distribution with variance $\sigma_D^2(u, v)$.

⁸We would like to pursue such a model in the future.

Remark 2.1 Experimentally, we have confirmed that for the areas corresponding to pixels with a stable integer disparity, this is a reasonable model (the empirical subpixel disparity distributions are unimodal, symmetric). For about 5 – 30% of the pixels in various experiments, we have observed bimodal empirical disparity distributions. We have concluded that the bimodality had resulted from ambiguities in matching since the shift between the modes was more than a pixel wide, and the locations of these pixels were in the areas where the integer disparity was unstable. This suggests that the matching procedure should be reexamined and improved.

3 Confidence Intervals and the Variability in 3-D Coordinates

Under the resulting normal model in disparity, for a given α , a confidence interval $C(\tilde{D}(u, v))$ for the subpixel disparity at the pixel (u, v) with theoretical probability of capture $TPC = 1 - \alpha$ is given by:

$$C(\tilde{D}(u, v)) = [\tilde{D}(u, v) - e(u, v), \tilde{D}(u, v) + e(u, v)], \quad (9)$$

where $e(u, v) = \zeta(\alpha/2)\sigma_D(u, v)$, and $\zeta(\alpha/2)$ denotes the $\alpha/2$ quantile of the standard normal distribution, i.e.,

$$\alpha/2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta(\alpha/2)} \exp(-x^2/2) dx. \quad (10)$$

We remind the reader that $TPC = Pr[\tilde{D}(u, v) \in C(\tilde{D}(u, v))]$. In our experiments, we selected the TPC, and then computed the intervals. In general, either TPC (equivalently α) or the size of the intervals should be specified based on the requirements of the application. When one of these parameters is specified, the other is calculated from (9-10). Note that these intervals characterize the variability (precision) in the results, not accuracy (i.e. deviations from an absolute ground truth).

Next, we propagate the bounds of the intervals, carefully accounting for the singularities and for the sign of the Jacobian of the reconstruction transformation (4). Details are given in [20]. As a result of the interval propagation, we obtain confidence intervals along with the coordinates of the reconstructed points. For a pixel (u, v) , for which the matching was resolved successfully with a disparity value $\tilde{D}(u, v) = \tilde{d}$, the confidence interval for the Z coordinate of the reconstructed 3-D point is $C_Z(u, v) = [L_Z(u, v), U_Z(u, v)]$. The lower and upper confidence bounds are defined by

$$L_Z(u, v) = \frac{S(3, :) * (u, v, 1, C_l(\tilde{d}))^t}{S(4, :) * (u, v, 1, C_l(\tilde{d}))^t} \quad (11)$$

$$U_Z(u, v) = \frac{S(3, :) * (u, v, 1, C_u(\tilde{d}))^t}{S(4, :) * (u, v, 1, C_u(\tilde{d}))^t}, \quad (12)$$

where, $C_l(\tilde{d})$ and $C_u(\tilde{d})$ are the bounds of the confidence interval in disparity, $C(\tilde{D}(u, v))$. The confidence bounds in X and Y are obtained by substituting $S(3, :)$ in (11-12) with $S(1, :)$ and $S(2, :)$, respectively. Note that the intervals for the disparity and for each of the coordinates X, Y and Z have the same probabilities of capture (when well defined). Although the intervals in disparity are symmetric around the computed subpixel disparity values, the intervals for the 3-D coordinates are not symmetric around the reconstructed coordinates. Since the subpixel disparity is normally distributed, the the coordinates of the

3-D point in projective space are normally distributed (that is each of \tilde{X} , \tilde{Y} , \tilde{Z} , and \tilde{K} is normal, with variances $S(i, 4)^2 \sigma_D^2(u, v)$, $i = 1, 2, 3, 4$, respectively), but the coordinates in Euclidean space are not normally distributed. Thus reporting a variance for the Euclidean coordinates is not a reasonable description of the variability (in a most dramatic example, the ratio of two independent random variable with standard normal distribution is a Cauchy distribution which has very heavy tails and nonexistent, i.e. infinite, variance).

The confidence intervals which we construct characterize the variability in the results, not the accuracy of the reconstruction algorithm. Thus experiments with ground truth models for 3-D positions are not a suitable test for these intervals. Such experiments will address directly the accuracy of the reconstruction algorithm, but not the precision. We present some quantitative data on the intervals in Section 5. The rationale which motivates us to use these intervals, despite the fact that they do not characterize the accuracy is based on the following observation: reconstruction results may be precise and inaccurate (the intervals do not characterize those), but we definitely want to reject as unreliable points which are not precise (have large variability), i.e. large interval sizes.

4 Radiometric Correction

The variation of intensity in dark images and in flat fields for different sensors clearly shows that sensors are not “equal”, contrary to the assumptions made in many multisensor computer vision applications. Given, even, absolutely the same scene and illumination, physically different sensors “see” differently. Depending on the application and the algorithm these differences may or may not have a significant effect on the results. As we demonstrate (Section 5), for area-based stereo matching, if the scene imaged has a weak texture it is important to correct radiometricly. Even for a single sensor system, if the main cue is intensity (for example in shape from shading methods) radiometric correction is necessary.

There are different radiometric correction procedures in use, but they all assume a linear sensor model (even if not explicitly stated). In some of these procedures, the model parameters are estimated off-line, and images preprocessed on line. In others, parameters are not recovered directly rather the images are preprocessed so that the expected offset in the processed images is zero and the photoresponse is uniform. The process of such radiometric correction is called *flat fielding* or *shading correction* [31]. Intuitively, in order to achieve a corrected image with zero offset and constant gain, we subtract from the original image the offset observed in averaged dark images, and scale the resulting image inversely proportional to the photoresponse observed in averaged flat fields, i.e., [31]:

$$I_{correctd}(u, v) = \frac{I(u, v) - \hat{m}(u, v)}{F(u, v) - \hat{m}(u, v)} \left[\frac{1}{MN} \sum_{u=1}^M \sum_{v=1}^N (F(u, v) - \hat{m}(u, v)) \right]$$

where F is a pixelwise average of flat fields taken under fixed illumination (an estimate of PRNU at a fixed illumination close to saturation level), and $\hat{m}(u, v)$ is the estimate of the FPN in dark current, (7). The expression in the brackets on the right-hand side is the average response (over all pixels) after F has been adjusted to zero offset.

5 Experiments

In this section, first, we illustrate the sensor noise in dark images and flat fields and the variation in the noise characteristics between different sensor configurations (Section 5.2); second, we demonstrate with ground test scenes the effect of the radiometric correction procedure in improving the accuracy of the stereo reconstruction results (Section 5.3); and third, we show series of experiments pertaining to the use of the confidence intervals in assessing the precision and rejecting unreliable points (Section 5.4).

5.1 Experimental setup

We used the following types of black and white analog CCD cameras and 8-bit framegrabbers in our experiments:

Camera	CCD array size	Pixel clock frequency
HITACHI KP 230 (231)	384(H)x485(V)	8MHz
SONY XC-77	768(H)x493(V)	14.318MHz
Framegrabber	Effective digital image size	Sampling frequency
DT1451	512(H)x480(V)	about 10MHz
S1V	640(H)x480(V)	not available
TIM40	756(H)x480(V)	set to 14.318MHz

Each camera was equipped with 25mm lens. For the stereo experiments: a pair of cameras were usually 6-10cm apart, non-verging, or verging at approximately 40° ; the cameras were viewing a volume of approximately $30 \times 30 \times 30\text{cm}^3$ at 80cm from the base line of the pair; each time a camera pair was reconfigured, it was strongly calibrated, [34]. In the experiments where strong texture was needed we induced it by using a slide projector and a slide representing random dots pattern.

5.2 Sensor Noise Observed in Dark Images and Flat Fields

In Section 5.2.1 we illustrate the noise in dark images, and in Section 5.2.2, the noise in flat fields.

5.2.1 Dark Images and the Background Noise

Figure 1 shows a typical dark image and a pixelwise average of 100 dark images for the camera/framegrabber configuration SONY XC-77/DT1451. In the left image (a representative dark image) the background random noise and internal luminance are noticeable. Systematic components in the background noise (offset and stripes) are prominent in the right image (the average of 100 dark images). The vertical stripes are primary due to aliasing resulting from the mismatch between the camera pixel clock and the framegrabber sampling frequency. When the camera pixel clock and the framegrabber sampling frequencies are well matched such aliasing does not occur (for example, in the configuration SONY XC-77/TIM 40, or for digital cameras where the digitization is done on the camera).

Figure 2 shows for 4 different sensor configurations, data based on the amplitude of the power spectra of the mean column intensities in dark images for each sensor. To enhance the display, we have applied the standard log-transform and have centered the results. For the configuration SONY XC-77/DT1451 the high frequencies are prominent, the camera lines

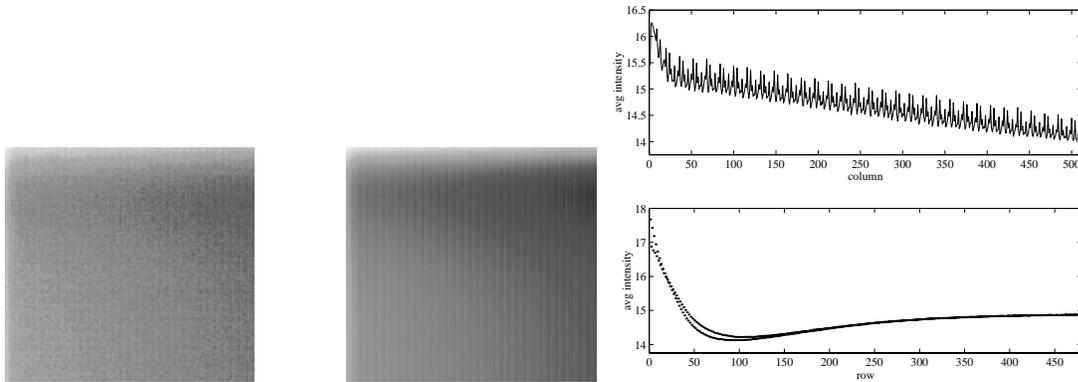


Figure 1. SONY XC-77/DT1451: **A single dark image (left) and a pixelwise average of 100 dark images (right).** The graphs represent the average intensity in the right image: **per column (top), and per row (bottom).** The offset in intensity of odd and even fields, a result of interlacing, is clearly visible in the bottom graph.

are under-sampled. Contrast this to SONY XC-77/TIM40, where the sampling is almost one-to-one. For each configuration, we tested several combinations of physically different cameras of the same type and different framegrabbers. Each physical configuration had its own dark image signature, but within a particular configuration type (same make), the background noise could be parametrized.

5.2.2 Flat Fields and Photoresponse Nonuniformities

The scene imaged was a white plane completely covering the field of view of the camera. To obtain flat fields we used a diffuser (a white glass filter) instead of the lens, achieving about 65% uniformity. Ideally, an integrating sphere should be used [28] which achieves 95% uniformity of the illumination.

The right pair of images in Figure 3 shows the pixelwise averages of 100 flat fields for each of two different cameras, configuration SONY XC-77/DT1451. The right pair of graphs underneath, Figure 4, represents the corresponding histograms. The acquisition of the flat fields without the use of a lens removes any nonuniformity which could have arisen from lens vignetting (fall off in intensity from the center of the image to the boundary) or other lens distortions. The bright square frames noticeable in the two intensity images, are due to reflections from the square aperture in front of the sensor array (a part of the camera architecture) and the adapter holding the diffuser. Apart from that artifacts, the variation in intensity in the images is mostly due to fixed pattern noise of the photoresponse nonuniformities of the individual sel (the read noise is negligible at that high level, a gray value of about 220)⁹.

5.3 The Effect of the Radiometric Correction on the Accuracy

When the texture is weak, correcting radiometrically the images reduce the number of false and ambiguous matches, and improves the accuracy of the stereo reconstruction. The purpose of the experiments in this section was to demonstrate that.

⁹The printing of the images has introduced minor defects as well.

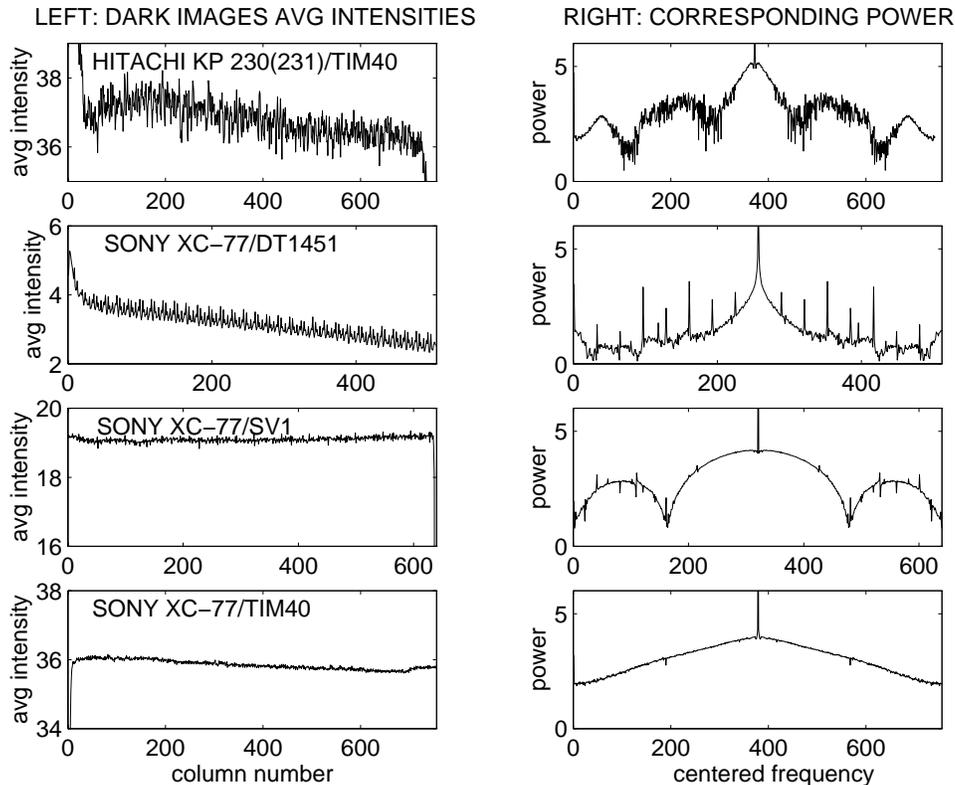


Figure 2. Data for single dark images obtained with different camera/framegrabber configurations. The graphs represent average intensities per columns (left), and the corresponding amplitudes of the power spectra (right).

For the radiometric correction procedure, we used average dark images and flat fields which were computed based on sequences of 100 images each. During the image acquisition process, caution should be paid to eliminate a possibility of image residues. The average dark image and the average flat field computation is one-time, off-line process (Figures 3-4).

On the day of the experiment, the cameras were warmed up; the lens parameters and illumination were controlled to prevent saturation; and the cameras were calibrated geometrically (intrinsic and extrinsic camera parameters, using least squares, were recovered).

Experiment 5.1 Flat Field Correction — Planar Example The target was a planar white card, lambertian reflectance surface¹⁰. A stereo pair of raw images was acquired (Figure 5, left pair of images), radiometrically corrected, and disparity maps based on the rectified uncorrected and on rectified corrected images were calculated. The rectification is a necessary step since the cameras were verging. Figure 6, shows the disparity maps as 3D meshes. The full size of the disparity maps is 256x256 since the rectification was done in half resolution. For the purpose of displaying the result, the 3D plots show the subsampled maps. “Holes” in the disparity maps can be observed, where the algorithm failed (matches were rejected). “Spikes” are outliers erroneously accepted as valid. The disparity map based on the radiometrically corrected images (right) has less holes and spikes. The two subpixel disparity maps were evaluated quantitatively by fitting a plane to each

¹⁰The material of the card was a standard poster board, there was some roughness texture, although very week, so the matching could be done.

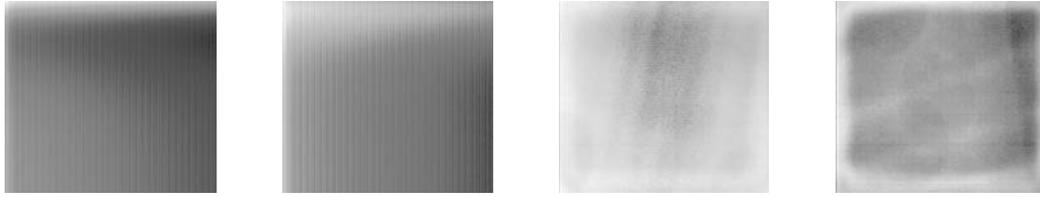


Figure 3. A pair of SONY XC-77RR/DT1451 sensors: the pair of average dark images (left) and the pair of average flat fields (right).

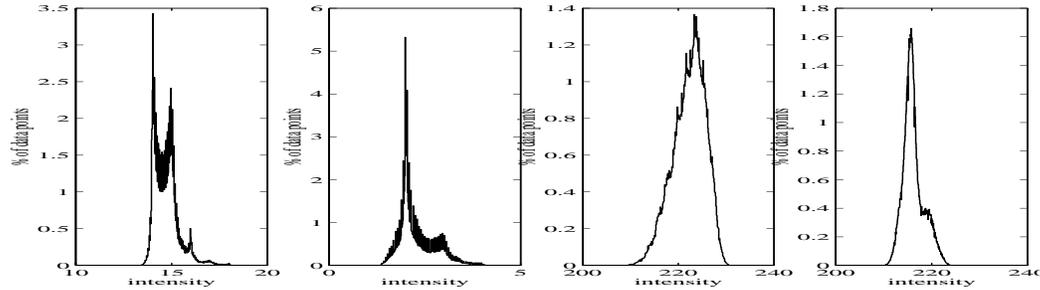


Figure 4. SONY XC-77RR/DT1451: the histograms of the corresponding images given in Figure 3 (same left to right order).

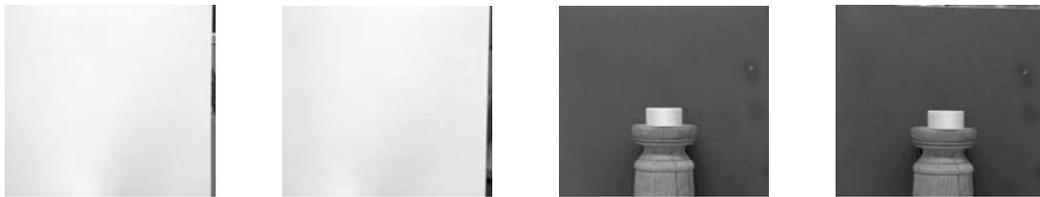


Figure 5. Original stereo pairs of images: a planar scene (left), a scene with a cylinder (right).

one of them, and comparing statistics of the residuals from the planar fit. Figure 7, shows the histogram for the residuals in the two cases. The standard deviation of the residuals, for the “uncorrected” map was 0.1668, and for the “corrected” was 0.0679, so a relative improvement of about 59% was obtained.

The experiment was repeated, testing the radiometric correction procedure at different flat field levels, approximately 60, 110, 160, 220. Subpixel disparity maps without and with flat-field correction were computed based on these, planes were fitted to each one, and statistics of residuals were compared. The dominance of the performance of the matching algorithm (in relative rms) with flat-field correction over the matching without flat-field correction is observed: from 26%, for the low intensity level flat field (level 60), to 59%, for the matching with flat field correction at level 220.

Experiment 5.2 Flat Field Correction – Non-planar Example. In this experiment, we used a scene which contained a cylindrical object, Figure 5, right pair of images. The target was a wooden cylinder, natural light color, it is the brightest object approximately in the center of the images. As in the previous experiment, the images were corrected radiometricly, and the 3D reconstructions based on the rectified uncorrected and on rectified corrected images were calculated. Figure 8, shows the reconstructed (x,y,z) points for the

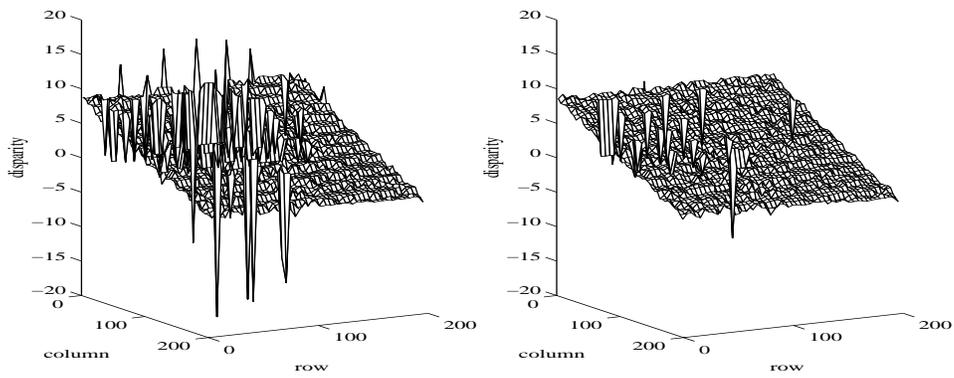


Figure 6. The planar scene: subpixel disparity maps shown as 3D mesh, based on original image pair (left), and on the radiometricly corrected pair (right).

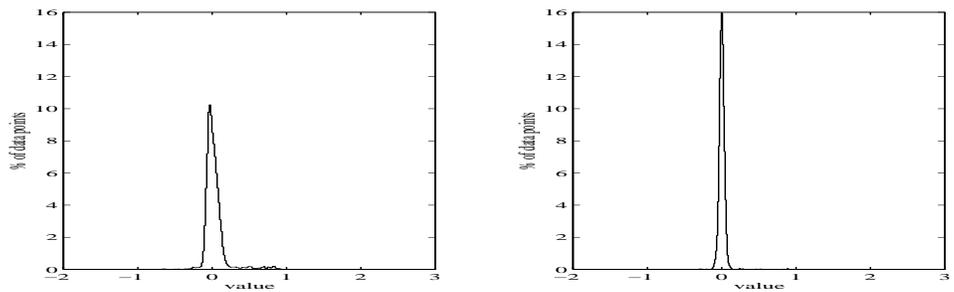


Figure 7. The planar scene: the histograms of the residuals from the planar fitting for the corresponding disparity maps shown in Figure 3 (same left to right order, uncorrected – left, and corrected – right).

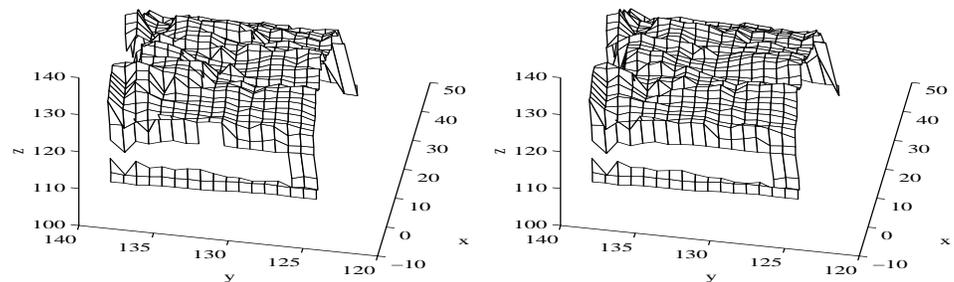


Figure 8. Cylinder: reconstruction results, shown as 3D mesh, based on the original image pair (left), and on the radiometricly corrected images (right).

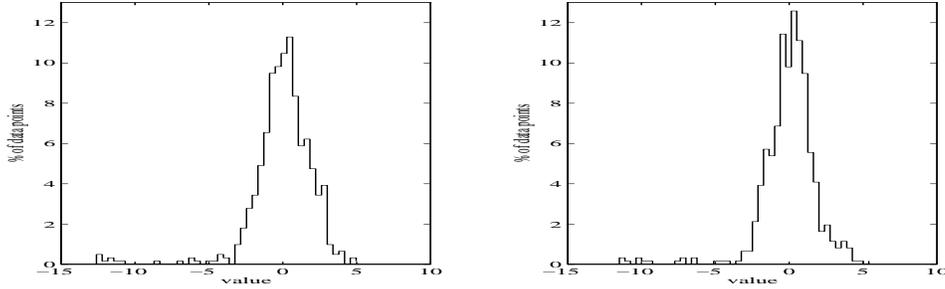


Figure 9. Cylinder: histograms of the residuals from fitting a cylinder to the reconstruction results given in Figure 8 (uncorrected — left, and corrected — right).

cylinder (only) as 3D meshes. The reconstruction based on the radiometrically corrected images (right) has smoother appearance. The two reconstructions were evaluated quantitatively by fitting a cylinder to each one of them, and comparing statistics of the residuals from the planar fit. Figure 9, shows the histogram for the residuals in the two cases. The standard deviation of the residuals, for the “uncorrected” map was 2.1477, and for the “corrected” was 1.9074, so a relative improvement of about 11% was obtained in this case.

5.4 Quantifying the Precision with Confidence Intervals

In this section we present a series of experiments regarding the confidence intervals for the precision. First we evaluated the intervals by comparing theoretical probabilities of capture (TPC) with empirical capture frequencies calculated during multiple runs of the stereo algorithm on tests scenes which were planar surfaces. We have also done experiments with more complex scenes, i.e. human faces and hands. Examples of the scenes are given on Figure 10. For the planar surfaces we present the quantitative data. For the human face and hands we did not evaluate the intervals quantitatively (the main reason being that the subjects could not keep still while sequences of 100 images needed for the empirical comparisons could be taken). We give the statistics of the intervals in the Z component of the reconstructed points, since this is of main interest in stereo. We have used standard normal distribution model for the image noise which amounts to an uniform scale of the intervals in subpixel disparity. The goal of these series of experiments was to assess the precision of the reconstruction algorithm in terms of the relative confidence interval sizes in depth, Z , and to compare these for several different stereo pairs of cameras. The test scene was a vertical planar surface coinciding with the coordinate plane ($X - Y$). The plane was positioned at precise, and known position for Z .

A random texture was projected on the planar surface (Figure 10, left pair of images). A sequence of 100 images was acquired with each of four cameras SONY XC-77RR. During the evaluation process of the confidence intervals, camera pairs were formed (using the corresponding images which were already taken). This way the performance of the intervals for varieties of stereo configurations was examined. The reconstruction algorithm, [34], augmented with the confidence intervals estimation procedure described in this paper was run for one stereo pair images. Confidence intervals, C_X, C_Y, C_Z with $TPC = 0.68$ were computed. The reconstruction algorithm was run 100 times, on each of the 100 stereo pairs. For each stereo pair and for each pixel position in the left image of the pair, we counted how many times the newly computed depth Z for that pixel position fell within C_Z . The

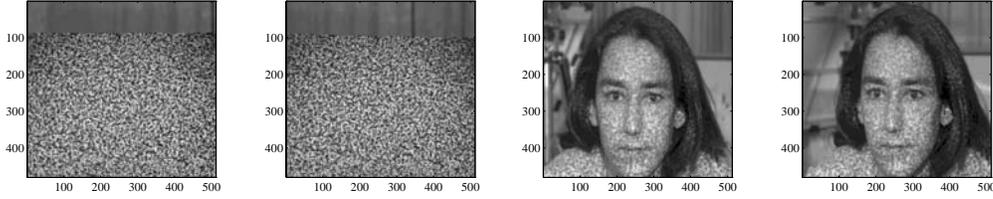


Figure 10. Test scene stereo pairs: plane (left), a face (right).

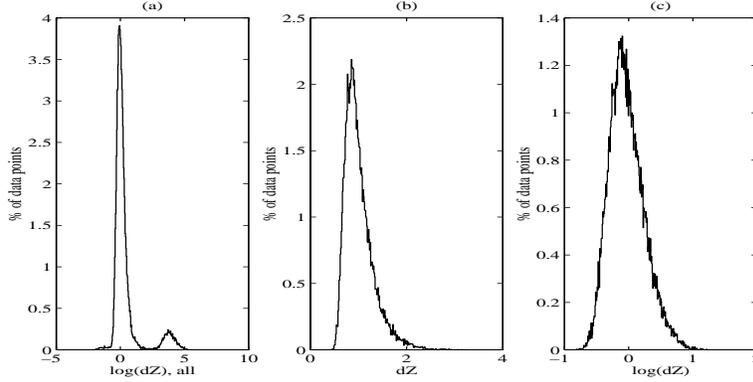


Figure 11. Pair(1,4): Histograms of interval sizes, dZ , in mm at 0.68 TPC. (a) a histogram over all intervals, log scale; (b) a histogram, only over intervals of reconstructed points, corresponding to image projection of the plane; (c) same as (b), but in log scale. We differentiate between (a) and (b)-(c) since the image of the plane does not coincide with the whole digital image.

final counts were divided by the total number of sample runs, thus the *empirical capture frequencies* (ECF) were computed.

Experiment 5.3 Quantitative data for the interval size in depth, Z , for a single camera pair. Figure 11 shows the histograms of the sizes, dZ , in mm for the confidence intervals with *theoretical probability of capture* (TPC) of 0.68, for one fixed stereo pair (labeled *Pair*(1,4)) and the planar scene, Figure 10 left images. Figure 12 shows various statistics for the empirical capture frequency for *Pair*(1,4). The pixels which corresponds to the projections of the plane (with random texture) have better ECF. In the bottom graph the pixel positions for which the ECF is below TPC have non-satisfactory performance.

Experiment 5.4 Comparisons over 4 different camera pairs. Next we compare results for four different stereo pairs which we label *Pair*(1,2)', *Pair*(1,4)', *Pair*(2,3)' and *Pair*(3,4)'. The random texture plane was covering the complete view of all cameras, so consistent comparison could be drawn across pairs. Confidence intervals, with probability of capture 0.68, were computed for the reconstructed points for each pair, and the ECF were calculated during 100 runs of the polynocular stereo reconstruction algorithm [34]. Statistics for ECF, and for the confidence interval sizes, dZ , in mm, are summarized in Table 1. Note that the TPC is less than the actual ECF. TPC underestimates the performance of the intervals. Under the same model, deriving intervals with higher TPC also leads

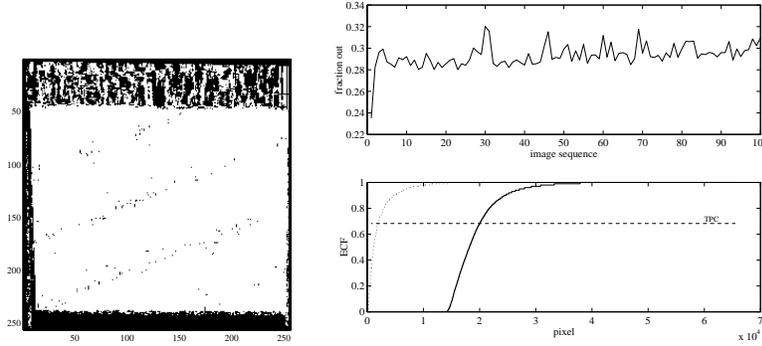


Figure 12. Pair(1,4): Data for ECF, planar scene. The image on the left shows in white the pixels for which the intervals for the corresponding 3-D depth have $ECF \geq TPC = 0.68$. The top graph on the right shows for each of the 100 reconstructions, the fraction of pixels for which the corresponding reconstructed Z is not captured by C_Z . The bottom graph shows sorted ECF of the intervals: the solid line is for the intervals corresponding to all reconstructed points; the dotted line is for those corresponding to pixels projections of the plane; and the dashed line represents the $TPC=0.68$.

	ECF			dZ mm				
<i>Pair</i>	<i>mean</i>	<i>std</i>	<i>median</i>	<i>min</i>	<i>max</i>	<i>mean</i>	<i>std</i>	<i>median</i>
(1,2)'	0.6838	0.3214	0.8000	0.1598	4.2830	0.9234	0.3043	0.8722
(1,4)'	0.7549	0.2842	0.8700	0.1180	226.5228	0.7082	0.9669	0.6666
(2,3)'	0.7327	0.2912	0.8500	0.1122	80.9347	0.6886	0.9303	0.6180
(3,4)'	0.6931	0.3154	0.8100	0.1725	4.9167	0.8375	0.2730	0.8035

Table 1. Statistics of ECF, and of confidence interval sizes, dZ in mm, across all points, for confidence intervals with TPC of 0.68, for 4 different stereo pairs of cameras

to intervals with larger sizes, meaning worse resolution and more uncertainty in the true position.

6 Conclusions

6.1 Summary

We demonstrated that by modeling the video sensor one can gain better understanding of the results of vision algorithms. We used as an illustration an algorithm for 3-D reconstruction from stereo. We showed that by modeling the sensor, and accounting for systematic and random sensor noise: (1) the accuracy could be improved, in particular, in cases where stereo usually fails, i.e. in weak texture areas (for example, one color, relatively smooth man made objects, like office walls, door, cabinets); and (2) that precision bounds could be provided in terms of confidence intervals. We have verified experimentally that the intervals meet the specified theoretical bounds. The interval sizes represent the degree of uncertainty, and the TPC degree of reliability.

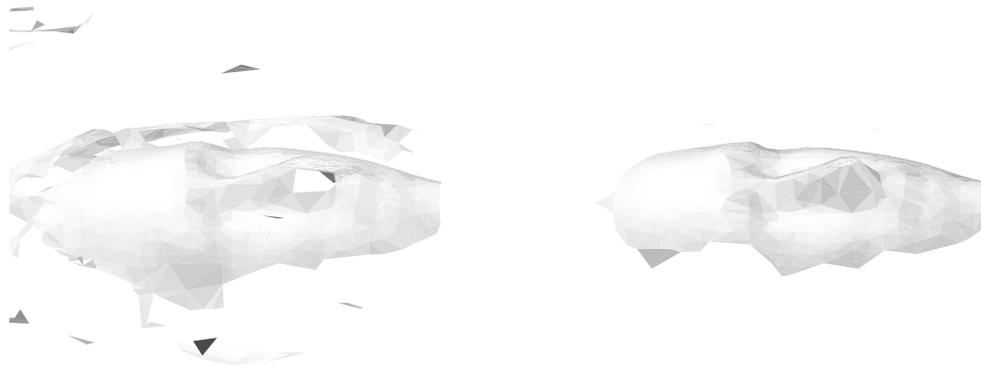


Figure 13. Left: A surface reconstruction based on all 3-D points. Right: A surface reconstruction based only on 50% of the 3-D points, only those with confidence interval in Z which have sizes smaller than the median interval size were used in the surface recovery.

6.2 Relevance of the Results to Other Vision Research Areas

The radiometric correction procedure is applicable not only to stereo, its effects are even more significant in other vision methods like shape from shading or photometric stereo where the intensity itself is of most importance. In fact, in any application, sensors should be carefully examined (using dark images and flat fields) to reveal possible defects and systematic and random noise characteristics. To obtain the intervals, we used a variance propagation method which is acceptable as far as algorithms or their modules are reasonably well approximated by linear transformations (i.e. in our case the subpixel disparity correction procedure was such). When linear approximations are not satisfactory, the possibilities are: (i) direct distribution model propagation (which could be very hard), or (ii) distribution models built directly for the results using ground test scenes.

6.3 Use of the Intervals in Surface Recovery

A feature of the intervals which supports their validity as quantifiers of the precision, is the fact that we have used them successfully for rejection of reconstructed points which have higher degree of uncertainty. We conclude here by presenting results from one such an experiment.

The test scene was a face. We have projected the random dot texture on it (Figure 10, right pair of images), and run the stereo algorithm augmented with the computation of the confidence intervals. Then, we have computed the median statistic of the interval sizes in the depth, Z , over all reconstructed points. Finally, we reconstructed the surface geometry and topology using the algorithm from [35] based on: (1) all reconstructed points, and (2) only those points for which the corresponding confidence interval sizes were less than the median. The results of the two reconstructed surfaces are shown on Figure 13. By the use of the confidence intervals for the rejection of points with high degree of uncertainty, we have reduced the set of the reconstructed points by 50%, and have still preserved the important features. We have removed outlier artifacts (particles hanging in front and around the

face), and also have to resolve ambiguities in the surface reconstruction stage (for example, the hole in the center of the face on the left is covered). The size of the confidence intervals is related to the scene attributes (texture and geometry) and the sensor noise parameters. The intervals are large when: (a) the signal-to-noise ratio of the image is locally low (as in weak texture areas, in areas overlaid by highlights where the texture contrast is low, and in dark or shadowed areas); and (b) when the match is wrong (as in case of highlights, in areas which does not image as a regular surface, like hair for example, and near occlusions where the matches are notoriously bad in all area-based stereo matching algorithms [33]). The intuition which may explain (a) is the dependence of the intervals size not only on the pixel noise levels, but also on the image gradients.

Currently we are exploring the use of the intervals for data fusion of redundant 3-D points reconstructed from different views or from multiple stereo pairs.

Acknowledgements

This work is supported by, or in part by, The National Science Foundation under grant numbers: MIP94-20397 A03 SUB, IRI93-3980, IRI93-07126, GER93-55018, and the U.S. Army Research Office under grant numbers: P-34150-MA-AAS DAAH04-96-1-0007 ARO/DURIP DAAG55-97-1-0064.

We thank Radim Šára and Reyes Enciso for helping in the preparation of this document.

References

- [1] N. Ayache and C. Hansen, "Rectification of Images for Binocular and Trinocular Stereovision", *Proc. of 9th International Conference on Pattern Recognition*, 1, pp. 11-16, Italy, 1988.
- [2] H. Beyer, "Linejitter and geometric calibration of CCD cameras", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 45, pp. 17-32, 1990.
- [3] P. Belhumeur, "A Bayesian Approach to Binocular Stereopsis", *Intl. J. of Computer Vision*, 19(3), pp 237-260, 1996.
- [4] H. Beyer, "Geometric and radiometric analysis of a CCD-camera based photogrammetric close-range system", *Dissertation, Institut für Geodäsie und Photogrammetrie, Zürich*, 1992.
- [5] E. Beynon and D. Lamb, ed., *Charge-coupled devices and their applications*, McGraw-Hill, 1980.
- [6] S. Blostein and T. Huang, "Error Analysis in Stereo Determination of 3-D Point Positions", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9(6), pp. 752-765, 1987.
- [7] Das and Ahuja, "Performance Analysis of Stereo, Vergence, and Focus as Depth Cues for Active Vision", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 17(12), pp 1213-1218, 1995.
- [8] F. Devernay, "Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models", INRIA, RR-2304, Sophia Antipolis, 1994.
- [9] U. Dohond and J. Aggrawal, "Structure from Stereo: a Review", *IEEE Transactions on Systems, Man, and Cybernetics*, 19(6), pp 1489-1510, 1989.
- [10] G. Florou and R. Mohr, "What accuracy for 3D measurements with cameras?" *Proc. of ICPR '96*, A80.1, pp.354-358, 1996.

- [11] W. Förstner, "10 Pros and cons against performance characterization of vision algorithms", *Workshop on performance characterization of vision algorithms*, Robin College, Cambridge, 1996.
- [12] R. Haralick, "Performance characterization in computer vision", *Performance versus methodology in computer vision*, Haralick and Meer, editors, University of Washington, Seattle, 1994.
- [13] R. Haralick, "Propagating Covariance in Computer Vision", *Workshop on performance characterization of vision algorithms*, Robin College, Cambridge, 1996.
- [14] E. Healey and R. Kondepudy, "Radiometric CCD Camera Calibration and Noise Estimation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(3), pp. 267-276, 1994.
- [15] J. Hoggs, *The Microscope: its history, construction and application*, London, 1854.
- [16] G. Holst, *CCD arrays, cameras, and displays*, Winter Park, FL : JCD Pub., 1996.
- [17] B. Jähne, *Digital image processing: concepts, algorithms and scientific applications*, Springer-Verlag, 1993.
- [18] J. Janesick et al, "The future scientific CCD", *Proc. of SPIE State-of-the-art imaging arrays and their applications, San Diego, California*, Vol. 501, pp. 2-33, 1984.
- [19] B. Julesz, *Foundations of Cyclopean Perception*, Chicago, 1971.
- [20] G. Kamberova and R. Bajcsy, "Precision of 3-D Points Reconstructed from Stereo", GRASP Lab, Technical Report, 1997, <http://www.cis.upenn.edu/~kamberov/doc/papers.html>
- [21] G. Kamberova and M. Mintz, "Minimax Rules Under Zero-one Loss for a Restricted Location Parameter", *Journal of Statistical Planning and Inference*, accepted, 1997.
- [22] G. Kamberova, "The Effect of Radiometric Correction on Multicamera Algorithms", Technical Report, MS-CIS-97-17, CIS Department, University of Pennsylvania, 1997
- [23] R. Lenz and D. Fritsch, "Accuracy of videometry with CCD sensors", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 45, pp. 90-110, 1990.
- [24] R. Mandelbaum, G. Kamberova and M. Mintz, "Stereo Depth Estimation: a Confidence Interval Approach", *Proc. Intl. Conf. Computer Vision ICCV98*, 1998.
- [25] D. Marr and T. Poggio, "Cooperative Computation of Stereo Disparity", *Science*, 194, pp 283-287, 1976.
- [26] D. Marr and T. Poggio, "A Computational Theory of Human Stereo Vision", *Proc. of the Royal Soc. of London*, B204, pp 301-328, 1979.
- [27] S. Maybank and O. Faugeras, "A Theory of Self-Calibration of a Moving Camera", *Intl. J. of Computer Vision*, 8(2), pp 123-151, 1992.
- [28] I. Mclean, *Electronic and computer-aided astronomy: from the eyes to electronic sensors*, Ellis Horwood Limited, England, 1989.
- [29] H. Moravec, "Robot Rover Visual Navigation", *Computer Science:Artificial Intelligence*, pp. 13-15, 105-108, UMI Research Press 1980/1981.
- [30] M. Okutomi and T. Kanade, "A Multiple-baseline Stereo", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(4), pp. 353-363, 1993.
- [31] Photometrics Homepage, *Photometrics High performance CCD Imaging*, <http://www.photomet.com>, 1996.
- [32] K. Price, "Anything you can do, I can do better (No you can't)...", *Computer Vision, Graphics, and Image processing*, Vol. 36, pp. 387-391, 1986.

- [33] R. Šára and R. Bajcsy, "On Occluding Contour Artifacts in Stereo Vision", *Proc. Int. Conf. Computer Vision and Pattern Recognition*, IEEE Computer Society, Puerto Rico, 1997.
- [34] R. Šára, "Reconstruction of 3-D Geometry and Topology from Polynocular Stereo", <http://cmp.felk.cvut.cz/~sara>
- [35] R. Šára and R. Bajcsy, "Fish-Scales: Representing Fuzzy Manifolds", *Proc. Intl. Conf. Computer Vision ICCV98*, 1998.
- [36] H. Sahabi and A. Basu, "Analysis of Error in Depth with Vergence and Spatially Varying Sensing", *Computer Vision and Image Understanding*, 63(3), pp. 447-461, 1996.
- [37] C. Slama, ed., *Manual of photogrammetry, IVth edition*, American Society of Photogrammetry, 1980.
- [38] D. Snyder *et al.*, "Compensation for readout noise in CCD images", *Journal of the Optical Society of America, A*, Vol. 12, No. 2, pp. 272-283, 1995.
- [39] A. Sripad and D. Snyder, "A necessary and sufficient condition for the quantization error to be uniform and white", *IEEE Trans. on acoustic, speech, and signal processing*, Vol. ASSP-25, No. 5, pp. 442-448, 1977.
- [40] A. Theuwissen, *Solid-State imaging with charge-coupled devices*, Kluwer Academic Publishers, 1995.
- [41] A. Yuille and T. Poggio, "A Generalized Ordering Constraint for Stereo Correspondence", *MIT, Artificial Intelligence Laboratory Memo*, No. 777, 1984.