Bias reduction for stereo triangulation

J. Ferrer and R. Garcia

Stereo triangulation lays at the basis of 3D scene recovery and it is used in a wide variety of areas ranging from urban modelling to robot localisation and mapping. However, triangulation produces non-Gaussian 3D estimates from Gaussian image measurements owing to its nonlinear nature. While previous work demonstrates the presence of statistical bias and how to correct the depth estimate, in this presented report, proposed and proven in a Monte Carlo test, is an enhancement for correcting the full 3D position given the image projection noise variance.

Introduction: When compared to other sensors, cameras provide information of high spatial and temporal resolution. For this reason, the recovery of a 3D scene structure by means of images acquired by digital cameras has been for a long time a topic of interest in many fields such as photogrammetry, remote sensing and computer vision. Moreover, 3D reconstruction has become a very important step in many different areas such as urban modelling, underwater mapping, or robot localisation. Although it is well known that stereo-vision is more accurate than monocular setups for scene reconstruction, many successful approaches have been proposed to perform 3D Euclidean reconstruction using monocular cameras [1]. The most important reasons for using less-precise bearing-only sensors instead of stereo-vision are simplicity, price and computational cost. However, nowadays commercial off-the-shelf stereo systems such as Fujifilm FinePix 3D and modern computers with high computational resources can be jointly used as low-cost stereo-vision-based mapping sensors.

Related work: Stereo reconstruction is the problem of determining the structure of a scene from more than one camera given the fact that a physical point projects only at a single position in a camera image. Initial work of Blostein and Huang [2] and Matthies and Shafer [3] in stereo error modelling used image measurements in terms of discrete pixels which produced diamond shaped uncertainties in a fronto-parallel stereo setup (non-verged geometry). While the work of Blostein and Huang derived very complex probability density functions (PDFs) for the estimated uncertainty of the reconstructed points, Matthies and Shafer proposed to approximate them by Gaussian PDFs and the use of the propagation rule from independent identically distributed (IID) Gaussian noise in the left and right image measurements to the 3D estimates. Recently, Sibley *et al.* have shown in [4] that the nonlinearity in the triangulation equation produces non-Gaussian estimates from Gaussian image measurements. While close-range measurements can be well approximated as Gaussian, long-range ones are statistically biased. Sibley *et al.* propose a method for correcting the depth from a disparity variance estimate in long-range measurements. In this Letter we propose an extension to correct the statistical bias not only in the depth (z), but also in the x and y co-ordinates in Euclidean 3D reconstruction using a stereo-rig.

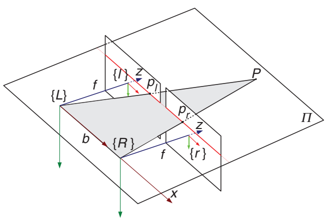

Fig. 1 *Non-verged stereo setup where left {L} and right {R} axes of aligned cameras are separated by baseline b in the collinear x axis. A world point P is projected to the left {l} and the right {r} image planes at positions p_l and p_r. Image plane frames {l} and {r} are at distance f (focal length) from their origin in the optical axis direction z. In this configuration the 3D problem can be reduced to a bidimensional one in the epipolar plane Π*

Problem statement: A stereo-vision system consists of a pair of fixed cameras separated by a constant distance, known as the baseline. This baseline constraint allows us to compute 3D estimates from image measurements. Given a pair of corresponding points, every one in each camera, we can triangulate [1] the position of the point in 3D with respect to the stereo-rig.

Let us consider the fronto-parallel stereo system illustrated in Fig. 1. In such a configuration, a stereo correspondence pair lies along the same horizontal axis (epipolar line) in both image planes. In a practical setup we achieve this configuration by standard calibration techniques [1]. High quality open source toolboxes such as Bouguet's Calibration Toolbox for Matlab® are publicly available (http://www.vision.caltech.edu/bouguetj/calib_doc) for obtaining this calibration. Given this non-verged configuration, the 3D position of a world point observed in both frames is obtained by stereo triangulation using the following equation:

$$P = \left(\frac{x_l/b}{d}, \frac{y_l/b}{d}, \frac{f/b}{d}\right)^\top \tag{1}$$

where $(x_l, y_l)^\top$ are the co-ordinates of the projection in the left image frame, $d = x_l - x_r$ is the disparity, b corresponds to the baseline of the stereo-rig, f the focal length of the camera and P is a 3D point in the scene, expressed with respect to the left camera frame. Because of the obvious nonlinearity in the triangulation equation, if we assume that image projections are corrupted by IID Gaussian noise [5], the 3D estimates are non-Gaussian. It is shown in [4] that the depth PDF is heavy tailed and produces statistically biased estimates. To correct this bias, Sibley *et al.* formulate P as three functions depending on the disparity $s_0(d)$, $s_1(d)$ and $s_2(d)$, respectively. Assuming that the measured d is Gaussian, which is reasonable since we model image noise in x_l, y_l and x_r, as IID Gaussian [5], and provided that the measured $\hat{d}$ is close to the real d, the following second-order Taylor series expansion in the depth $s_2(d)$ may provide a better estimate:

$$\tilde{s}_2(d) \simeq s_2(d) + \left.\frac{\partial s_2}{\partial d}\right|_d (\hat{d} - d) + \frac{1}{2}(\hat{d} - d)^2 \left.\frac{\partial^2 s_2}{\partial d^2}\right|_d \tag{2}$$

Since expectation $E[\hat{d} - d] = 0$ and $E[(\hat{d} - d)^2]$ is the variance, replacing d by $\hat{d}$, we obtain the unbiased estimate:

$$\tilde{s}_2(\hat{d}) \simeq \tilde{s}_2(\hat{d}) - \frac{1}{2}\mathrm{var}(\hat{d}) \left.\frac{\partial^2 s_2}{\partial d^2}\right|_{\hat{d}}$$

where $\tilde{s}_2(\hat{d})$ is the biased triangulation using (1). However, in this way, the same correction cannot be applied to $s_0(d)$ and $s_1(d)$ since these expressions contain other variables (i.e. x_l and y_l) that must be taken into account. Therefore we propose to express P as functions of x_l, y_l and x_r instead of d, obtaining:

$$P = \begin{pmatrix} s_0(x_l, x_r) = \frac{x_l \cdot b}{x_l - x_r} \\ s_1(x_l, x_r, y_l) = \frac{y_l \cdot b}{x_l - x_r} \\ s_2(x_l, x_r) = \frac{f \cdot b}{x_l - x_r} \end{pmatrix} \tag{3}$$

The multivariate second-order Taylor series expansion of (4) may lead to unbiased estimates for the full 3D position P [6], i.e.:

$$f(x) \simeq f(s) + Df(a)(x - a) + \frac{1}{2}(x - a)^\top D^2 f(a)(x - a) \tag{4}$$

where x is a variable vector $(x_1, x_2, \ldots, x_n)$, a is a point $(a_1, a_2, \ldots, a_n)$, $Df(a)$ is the Jacobian of $f(x)$ evaluated at point a and $D^2 f(a)$ is the Hessian of $f(x)$ evaluated at a. Then, applying (4) to (3) and following the same reasoning carried out in (2) about expectation, the first-order term in (4) is 0. Assuming independence between the variables x_l, y_l and x_r, the off diagonal elements in the second-order term are scaled by the covariance between variables, therefore they are 0. Finally, assuming v the equal variance in x and y image measurement directions, yields the following bias corrected 3D point estimation equation:

$$s_0(x_l, x_r) = \tilde{s}_0(x_l, x_r) - bv\frac{x_l + x_r}{(x_l - x_r)^3}$$

$$s_1(x_l, y_l, x_r) = \tilde{s}_1(x_l, y_l, x_r) - \frac{2vby_l}{(x_l - x_r)^3}$$

$$s_2(x_l, x_r) = \tilde{s}_2(x_l, x_r) - \frac{2vbf}{(x_l - x_r)^3}$$

Results: In Fig. 2 we prove, in a Monte Carlo test, that the method drastically reduces the error in the shortest disparities. The parameters used in this simulation were obtained from a real stereo calibrated system with a resolution of 640 × 480 pixels. It can be clearly observed in the plot that we significantly correct a valuable percentage (around 7%) of all the possible disparities.

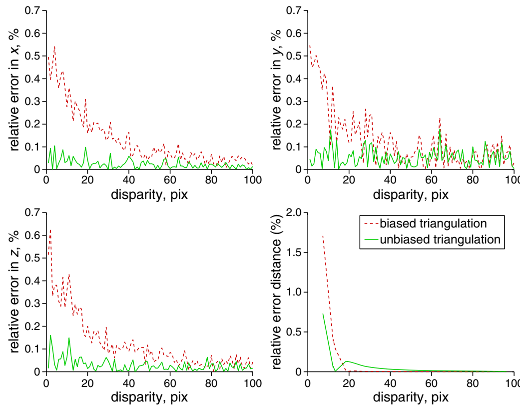

Fig. 2 *Top row and bottom left pictures demonstrate bias correction of image co-ordinate (2,200) in 10000-sample Monte Carlo test adding Gaussian noise* $\sim\mathbb{N}(\mu = 0, \sigma^2 = 1.5$ *pixels). In this setup the bias in longer range estimates (disparities smaller than 40 pixels) is clearly corrected (solid) compared to standard triangulation (dashed). Bottom right plot shows correction for full 3D position in real experiment using as ground truth the 3D estimates resulting from a nonlinear optimisation*

Conclusion: We have presented a method for correcting the statistical bias in stereo triangulation for the full 3D point co-ordinates that only requires the image measurement variance. We carried out experiments over both synthetic and real data showing the correctness and the performance of the method. Because of the simplicity of the derived correction, it could be very easily incorporated into any algorithm that uses standard triangulation, thus increasing its accuracy.

Acknowledgments: This work has been partially funded by the MICINN under grants CTM2010-15216 and PI08/9087. J. Ferrer has been funded by MICINN under FPI grant BES-2006-12733.

2 September 2010
doi: 10.1049/el.2010.2455
One or more of the Figures in this Letter are available in colour online.

J. Ferrer and R. Garcia (*Computer Vision and Robotics Group, University of Girona, Spain*)

E-mail: jferrerp@eia.udg.edu

References

1 Hartley, R., and Zisserman, A.: ‘Multiple view geometry in computer vision’ (Cambridge University Press, 2004, 2nd edn.)
2 Blostein, S.D., and Huang, T.S.: ‘Error analysis in stereo determination of 3-D point positions’, *IEEE Trans. Pattern Anal. Mach. Intoll.*, 1987, **9**, (6), pp. 752–765
3 Matthies, L., and Shafer, S.: ‘Error modeling in stereo navigation’, *IEEE J. Rob. Autom*, 1987, **3**, (3), pp. 239–248
4 Sibley, G., Sukhatme, G.S., and Matthies, L.: ‘The iterated sigma point Kalman filter with applications to long range stereo’. Robotics science and systems, Philadelphia, PA, USA, August 2006
5 Schmid, C., Mohr, R., and Bauckhage, C.: ‘Comparing and evaluating interest points’. Int. Conf. on Computer Vision, Bombay, India, 1998, Vol. 36, pp. 230–235
6 Cox, D.R., and Hinkley, D.V.: ‘Theoretical statistics’ (Chapman and Hall, 1979)