Predicting spatial uncertainties in stereo photogrammetry: achievements and intrinsic limitations

André Jalobeanu

Centro de Geofísica de Évora - Universidade de Évora
Rua Romão Ramalho, 59, 7002-554 Évora, Portugal
jalobeanu@uevora.pt

Abstract

We present a new probabilistic method for digital surface model generation from optical stereo pairs, with an expected ability to propagate errors from the data to the final result, providing spatial uncertainty estimates to be used for quantitative analysis in planetary or Earth sciences. Existing stereo-derived surfaces lack rigorous, quantitative error estimates, and we propose to address this issue by deriving a method of error prediction, rather than error assessment as usually done in the area through the use of reference data. We use only the information present in the available data and perform the prediction using Bayesian inference. We start by defining a forward model, using an adaptive radiometric change map to achieve robustness to noise and reflectance effects. A priori smoothness constraints are introduced to stabilize the solution. Solving the inverse problem to recover a surface from noisy data involves fast deterministic optimization techniques. Though the reconstruction results look satisfactory, we conclude that uncertainty estimates computed from two images only are unreliable, which is due to major limitations of stereo, such as non-Lambertian reflectance and incorrect spatial sampling, which violate our underlying assumptions and cause biases that cannot be accounted for in the predicted error budget.

Keywords: Bayesian inference, probabilistic modeling, digital photogrammetry, stereo, DSM generation, image processing.

1 Introduction

Stereo optical images are still widely used to generate digital surface models (DSM). Commercial cameras exhibit increasingly higher resolution, signal-to-noise ratio and dynamic range, so that the uncertainty in topographic measurements is expected to shrink accordingly. In this study, we show that not only this is not happening, but there are also fundamental limitations that make error prediction unreliable in practice, despite a rigorous probabilistic treatment of the problem.

We wish to predict the DSM accuracy, rather than assess it using reference data sets as done usually. Moreover, we want to capture the spatial variability of this error and its spatial correlation which has an impact in most applications as noticed by (Wechsler and Kroll, 2006). Many 3D reconstruction methods have been developed in the computer vision community (Brown et al., 2003), however they do not provide quantitative error estimates. Some attempts have been made to predict the
accuracy (Davis et al., 2001) based on local terrain characteristics and a qualitative matching quality measure, however these approaches do not directly take into account the data and the spatial variability of stereo cues. Bayesian approaches (Bernardo and Smith, 1994) to stereo disparity estimation have been developed (Cheng and Caelli, 2007) but do not explicitly produce error maps, and the data terms are not really adapted to photogrammetry as the method was mainly designed to handle computer vision problems with indoors imagery. We propose to use the image information content to compute the uncertainty, as the presence of edges, texture and noise can have a dramatic impact on it. A rigorous probabilistic modeling of data formation followed by Bayesian inference enables us to build a probability density function (pdf) of the disparity map, that helps provide a DSM with an associated error map via a geometric transform (known or estimated via calibration).

2 Bayesian stereo disparity inference

We extend here the approach first presented in (Jalobeanu et al., 2010). Within a probabilistic framework, all the parameters are random variables, which helps to account for the randomness of phenomena affecting observations (noise, radiometry) and the underlying variability of the object of interest (DSM or disparity map). Bayesian inference makes use of available knowledge expressed by specifying a priori pdfs, and combines it with a data formation model to derive the a posteriori pdf of the object given the data. The idea is to use Bayesian networks (Jordan, 1998) to model all variables and causal relations between them, in order to form a joint pdf by simply multiplying all the prior and conditional pdfs. Fig. 1 shows the proposed network, where nodes represent variables, converging arrows conditional pdfs and terminal nodes prior pdfs. Observed variables or data are in gray, fixed variables (whose estimation is beyond the scope of this paper) are in blue. One typically has to integrate out unwanted variables, thus performing a marginalization, ending up with a marginal pdf proportional to the sought posterior.

Figure 1. Directed graphical model or Bayesian network that helps to build the joint pdf.

We define a local area matching (Brown et al., 2003) method as follows, considering a fixed patch $I^1$ on image $Y^1$ and a moving patch $I^2$ extracted from $Y^2$ by assuming a uniform shift (determined by the local disparity), the resampling being done via Spline interpolation (Thévenaz et al., 2000) in order to minimize sampling artifacts or aliasing (Jain, 1989). To account for additive and multiplicative radiometric changes between windows $I^1$ and $I^2$, we assume a linear transformation of the pixel values with local parameters $a$ and $b$. We also assume the noise to be Gaussian of mean 0 and variance $\sigma^2$. These parameters are assumed constant over
the window support. The conditional Gaussian pdf writes:

\[
P(I_1 | I_2, a, b, \sigma) = \prod \frac{1}{\sigma \sqrt{2\pi}} e^{-\left(\frac{I_1^2 - aI_2^2 - b}{2\sigma^2}\right)}
\]

(1)

Integrating with respect to the three local change parameters gives the following expression of the likelihood, where \(c\) is the normalized correlation coefficient:

\[
P(I_1 | I_2) = \int \int \int P(I_1 | I_2, a, b, \sigma)\, da\, db\, d\sigma \propto (1 - c^2)^{-\alpha} F_\alpha(c)
\]

(2)

where \(\alpha = (n - 3)/2\), \(n\) being the number of pixels of the patch. The attenuation function \(F_\alpha\) arises from the constraint \(a > 0\) and is involves the error function erf:

\[
F_\alpha(x) = 1 + \text{erf}\left(x \sqrt{\frac{\alpha}{1 - x^2}}\right)
\]

(3)

This mapping from correlation to likelihood is the main originality of the proposed method. It is illustrated in Fig. 2 which shows how a cost function (the -log likelihood) behaves, for a 5x5 window, penalizing negative and small values of the correlation.

Now we form the joint likelihood as the product of local likelihoods, assuming independence, where the \(k\)-th patches explicitly depend on the data and the local disparity parameter \(\Delta_k\):

\[
P(Y_1 | Y_2, \Delta, \theta) \propto \prod_k P(I_{1k} | I_{2k}, \Delta_k, \theta)
\]

(4)

The sought posterior is proportional to the joint likelihood times the prior:

\[
P(\Delta | Y_1, Y_2, \theta, \omega) \propto P(Y_1 | Y_2, \Delta, \theta) P(\Delta | \omega)
\]

(5)

The prior model for the disparity (or DSM) is defined by a first order Markov Random Field (Li, 1995) which helps constrain the smoothness of the solution, with a global regularization parameter \(\omega\) (assumed fixed in this work):

\[
P(\Delta | \omega) = \frac{1}{Z_\omega} e^{-\omega \Phi(\Delta)} \quad \text{with} \quad \Phi(\Delta) = \sum_{i, k} (\Delta_i - \Delta_k)^2
\]

(6)

where \(Z_\omega\) is a normalizing constant.

We refer to (Jalobeanu et al., 2010) for a description of the inference algorithm. A fast, deterministic optimization algorithm based on Loopy Belief Propagation (Sun et al., 2003), applied to the posterior (5), is used to generate the DSM. The input is a set of lowpass-filtered, sampled likelihood terms, and the processing is done in a multiscale framework using a pyramid decomposition (Jain, 1989). The uncertainties are essentially derived from the shape of the likelihood functions at the optimum, and in the following we examine the issues related to these functions.

3 Intrinsic limitations: unpredictable biases

3.1 Fundamental assumptions and their consequences

The ability to predict local uncertainties depends on the consistency of the likelihood terms. These terms embed the local information carried by the data as a pdf of the local disparity without any prior information. A bias is judged significant when the true disparity lies outside a predefined confidence interval. Unfortunately
we observed enough significant biases (outside 95% and even 99% confidence intervals) to raise concerns about the reliability of the method. Indeed, a number of spatial locations suffers from error underestimation as the bias can not be predicted. After analysis it appears that the limitations are mainly due to the intrinsic quality of the data and not to algorithmic simplifications. For this analysis we went back to the basic assumptions and investigated their effects (see Table 1 for a summary).

Figure 3 illustrates the two major types of artifacts that affects stereo photogrammetry that we want to point out. The likelihood pdfs $P(I^1 | I^2)$ computed using (2) are shown and the bias is clearly visible in both cases, as the bulk of the pdf is far from the true elevation, checked during field work. Artifacts are due to undersampling (Jain, 1989) or image aliasing are strongest on high spatial frequency objects such as truck tracks in the sand, and the pdfs are so narrow that the final optimization stage using the smoothness prior (6) does not help to remove them. Aliasing is an intrinsic property of data and reflects a deliberate choice in the optical design (people like sharp images!). It can be reduced if a frequency space lowpass filter is employed (Jain, 1989) but with a loss in spatial resolution (factor 2 or higher). A drawback of such a filtering is the amplification of the radiometric errors described below, which was confirmed on simulations.

Radiometric changes were accounted for in a simple manner with a reasonable number of parameters (3, for about 20 data points) as defined in (1). Obviously this is insufficient in some cases, as on the edge of the tennis court (the reflectance depends on the viewing angle and changes within the patch area, despite the small size of the patch). Such cases are not uncommon in nature, as reflectance properties are spatially variable and rarely Lambertian. It is difficult to address this issue without significantly increasing the number of parameters of the radiometric model and ending up overfitting the data. The uniform parameter assumption is unavoidable and yields yet another intrinsic limitation of stereo.

Compression noise is neither white nor Gaussian; its spatial structure can induce outliers, and is commonly found in satellite or planetary data due to communication bandwidth constraints, however aerial cameras do not have such concerns. On the other hand, the uniform motion assumption is independent of data quality, and causes well-known fattening effects when the patch covers multiple objects or complex terrain. Such biases are related to slope, curvature and roughness, but not in a simple way that would help cancel or identify them. Finally the independence assumed between area-based data terms in (4) is questionable, as some overlap might be tolerated in order to achieve a satisfactory spatial resolution of the final DSM. With 5x5 patches one typically estimates a disparity every 2x2 pixels. While the dependence is obvious in this case, it is not clear how to evaluate it, or if joint likelihoods of two neighboring areas could be integrated at all in this methodology. In any case the related biases are negligible compared to radiometric or aliasing biases, as simulations have shown.
Table 1. Main assumptions made, actual real data properties and related unpredictable bias.

<table>
<thead>
<tr>
<th>Assumption</th>
<th>Real data properties</th>
<th>Observed bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform radiometric change</td>
<td>Spatially-varying</td>
<td>Artifacts</td>
</tr>
<tr>
<td>parameters</td>
<td>non-Lambertian brdf</td>
<td></td>
</tr>
<tr>
<td>Correct sampling</td>
<td>Undersampling (aliasing)</td>
<td>Aliasing artifacts</td>
</tr>
<tr>
<td>Uniform motion</td>
<td>Strong topography variations,</td>
<td>Fattening effect</td>
</tr>
<tr>
<td></td>
<td>multiple objects</td>
<td></td>
</tr>
<tr>
<td>White Gaussian noise</td>
<td>Correlated noise (compression)</td>
<td>Isolated outliers</td>
</tr>
<tr>
<td>Independent data terms</td>
<td>Overlapping windows</td>
<td>Uncertainty errors</td>
</tr>
</tbody>
</table>

3.2 Bayesian DSM generation from digital aerial images: an illustration

A series of aerial images were acquired in 2009 over the Portuguese coastline between Troia and Sines, at 20 cm ground sampling distance (GSD) and using an Intergraph DMC camera, with 60% overlap to allow for a stereo processing. One of the images is shown in Fig. 4, with the corresponding DSM reconstruction at 40 cm GSD in shaded relief and color. The direct georeferencing was insufficient, so control points had to be used to correct the orientation. The DSM was generated in the image space with the proposed method, and then resampled in UTM coordinates to be used within a cliff erosion monitoring project (Jalobeanu et al., 2010). Computing the topography variation with an error map is easily done by subtracting two multidate DSMs and adding the error variances. This enables one to check the statistical significance of the computed elevation changes, provided that the uncertainties are consistent. The DSM looks satisfactory after a preliminary inspection; a validation procedure using RTK GPS tracks as ground truth is in progress.

The trench on top left of the area is an aliasing artifact, but could easily be interpreted as an actual change if we were to trust the estimated errors (see Fig. 3). We notice that the vegetation appears in the DSM, as expected, but raises an important question about the relevance of surface-related error maps (even when they are consistent with the true surface) when most users are interested in topography.

Figure 4. Results: from left to right, image 1, shaded-relief DSM and color-coded DSM (UTM/WGS84 projection, coordinates in meters, arbitrary origin).
Conclusion

Fundamental limitations impede the robust estimation of the spatially variable uncertainty related to surface models. They are due to departures from assumptions, some of which are related to intrinsic data quality issues (sampling, noise), while the others reflect the insufficiency of stereo images to constrain all the physical parameters involved in image formation (radiometry and geometry). Filtering the images, multiplying the number of observations and reducing the parallax might help reduce some of the biases but could also increase some others. In any case, if camera manufacturers avoided undersampling and compressing, two important sources of bias would be eliminated and the error prediction would be improved.

Nowadays, to obtain reliable and topographically consistent error bars, one may consider using LiDAR data (when available), for which the error propagation is more straightforward to perform, as range measurements are made directly.

We have shown examples of failures in error estimation, however the frequency of these failures in real data is still unknown; experiments are in progress to assess the robustness of error estimation using dense reference data sets and determine in what cases the proposed uncertainty map can be used, especially when only stereo data are available for practical reasons.

Acknowledgments

This work was partially funded by the French Research Funding Agency (ANR) (SpaceFusion project, “Jeunes Chercheurs 2005” JC05_41500) and by the Portuguese Funding Agency (FCT) (AutoProbaDTM project PTDC/EIA-CCO/102669/2008, FCOMP-01-0124-FEDER-010039).

References

Bernardo, J., Smith, A. (1994), Bayesian Theory, John Wiley and Sons