Improving Stereo Sub-Pixel Accuracy for Long Range Stereo

Stefan K. Gehrig and Uwe Franke
HPC 050-G024
71059 Sindelfingen
Germany
{Stefan.Gehrig,Uwe.Franke}@DCX.Com

Abstract

Dense stereo algorithms are able to estimate disparities at all pixels including untextured regions. Typically these disparities are evaluated at integer disparity steps. A subsequent sub-pixel interpolation often fails to propagate smoothness constraints on a sub-pixel level. The determination of sub-pixel accurate disparities is an active field of research, however, most sub-pixel estimation algorithms focus on textured image areas in order to show their precision.

We propose to increase the sub-pixel accuracy in low-textured regions in three possible ways: First, we present an analysis that shows the benefit of evaluating the disparity space at fractional disparities. Second, we introduce a new disparity smoothing algorithm that preserves depth discontinuities and enforces smoothness on a sub-pixel level. Third, we present a novel stereo constraint (gravitational constraint) that assumes sorted disparity values in vertical direction and guides global algorithms to reduce false matches, especially in low-textured regions. Our goal in this work is to obtain an accurate 3D reconstruction. Large-scale 3D reconstruction will benefit heavily from these sub-pixel refinements, especially with a multi-baseline extension.

Results based on semi-global matching , obtained with the above mentioned algorithmic extensions are shown for the Middlebury stereo ground truth data sets. The presented improvements, called ImproveSubPix, turn out to be one of the top-performing algorithms when evaluating the set on a sub-pixel level while being computationally efficient. Additional results are presented for urban scenes. The three improvements are independent of the underlying type of stereo algorithm and can also be applied to sparse stereo algorithms.

1. Introduction

Recent work on stereo vision has focused on dense global algorithms that are able to obtain very accurate results. For a survey of the current state of the art refer to [13] and to the Middlebury stereo website that hosts the list of top-performing stereo algorithms compared to ground truth data [12].

The applications for dense, accurate disparity estimations are numerous, ranging from view interpolation to 3D reconstruction. Our main interest lies in obtaining high accuracy disparity images for 3D reconstruction at small disparities, i.e. at large distances. In this paper we present three improvements for sub-pixel stereo computation, focusing on low-textured image regions. Textured regions benefit from the improvements as well. Our scope in this paper is limited to obtaining accurate 3D reconstructions. 3D Data storage and view synthesis are not addressed. The 3D results are shown from perspectives that allow us to assess the accuracy of the methods - they are not deemed suitable for view synthesis.

Most stereo algorithms obtain disparities on an integer level. A simple sub-pixel interpolation can be done via parabola fitting of the disparities in the vicinity of the best disparity [19]. There has also been a growing interest in obtaining accurate sub-pixel disparities since the parabola fitting approaches exhibit artifacts known as pixel-locking [15]. These approaches perform a careful interpolation of the integer disparity results [10] or work on a disparity space image that is sampled at fractional disparities [17].

How is this paper organized? Section 2 gives an overview of stereo sub-pixel interpolation techniques. In Section 3, the technique of sampling the disparity space at fractional disparity steps is described. The following section introduces our novel disparity smoothing algorithm. In Section 5 a new weak stereo constraint is introduced. Section 6 gives detailed results on fractional disparity sampling, on the disparity smoothing and on our gravitational stereo constraint. In our evaluation we use ground truth data and

urban scenes. The final section comprises conclusions and future work.

2. Related Work

We limit our review to papers addressing sub-pixel estimation of stereo correspondences.

Most dense global algorithms deliver integer disparities. These disparities can be refined via parabola fitting when a matching cost is available for adjacent disparities [19]. A few global algorithms deliver sub-pixel disparity estimates directly: Belief propagation using an MMSE (minimum mean squared error) estimator yields sub-pixel disparities implicitly [18]. In this paper, marginal probabilities per pixel and disparity step constitute weights resulting in sub-pixel disparity values. Another recent belief propagation variant also obtains sub-pixel disparities by fitting surfaces on small segments [8]. An energy-based formulation of the disparity estimation problem is presented in [1], which also results in sub-pixel disparity estimates. There, several ideas from the optical flow estimation problem were applied to stereo vision. These algorithms require large computational resources.

Most other sub-pixel algorithms get their integer disparity values with a correlation-based stereo method (e.g. [13]) and perform a refinement subsequently. Psarakis et al. [11] conduct a linear interpolation of adjacent correlation values and solve a one-dimensional optimization problem. Nehab et al. [10] observe a bias in stereo matching due to the use of typically the left image as the reference image. Nehab et al. treat both images symmetrically and a two-dimensional fit in the disparity space yields accurate sub-pixel disparities. Stein et al. [16] take the integer disparities and perform a gradient-descent in the spirit of [9]. All three methods present results free of the pixel-locking effect. The computational burden of these methods is quite low, especially Psarakis method. A fast method to remedy the pixel-locking effect is also presented by Shimizu [15]. There, the integer correlation values are resampled shifted by a half pixel and both results are averaged. The result exhibits a slight locking to integer and to half pixel values. All these methods rely on sufficient structure to successfully correlate image patches.

A rigorous analysis of the sub-pixel effects in stereo matching by Scharstein and Szeliski has been conducted in [17]. There, a Fourier analysis shows that using a sinc interpolator is in theory the best interpolation to evaluate the disparity space image at fractional disparities. Chen [4] comes to the same conclusion in the context of optical flow estimation.

The analysis of [17] neglected effects of the sensor geometry, especially the pixel fill ratio. When approaching such accuracies the physical layout of the camera becomes relevant. An investigation of the effect of sensor geometry on correspondence measurements is presented in [20]. This analysis gets even more complicated considering recent trends using micro-lenses on every pixel which gives a complicated photo-sensitivity distribution for every pixel.

We are interested in obtaining accurate sub-pixel estimates for all pixels in the scene, including areas of low or no texture, at little computational expense. We implemented several of the above sub-pixel algorithms, confirmed their accuracy provided sufficient structure, however, obtained high disparity noise in low-textured areas. This is due to the fact that the correlation values or other similarity costs do not have sufficient information in low-textured regions. Such areas can only be filled with meaningful disparities using global algorithms with smoothness constraints. One method that performs disparity smoothing similar to our algorithm is introduced in [21]. There, bilateral filtering (adaptive filter coefficients depending on adjacency and on photoconsistency to the central pixel) is applied to obtain an improved disparity map, resulting in depth-discontinuity-preserving smoothing. This non-iterative method yields excellent results but is computationally quite expensive.

Besides smoothness, the epipolar constraint and the ordering constraint are known to be helpful in establishing stereo correspondences[13]. We propose another helpful constraint for low-textured regions which we call the gravitational constraint introduced in Section 5. It develops its full potential in outdoor scenes.

2.1. Semi-Global Matching

One of the many dense stereo algorithms listed on the Middlebury web page is semi-global matching (SGM) [5]. Our investigations are conducted with this algorithm due to its computational efficiency. Among the top performing algorithms, Hirschmueller's SGM is the fastest.

Roughly speaking, SGM performs an energy minimization in a dynamic-programming fashion on multiple (8 or 16) 1D paths approximating the 2D image. The energy consists of three parts: a data term for photo-consistency, a small smoothness energy term for slanted surfaces that change the disparity slightly (parameter P_1), and a smoothness energy term for depth discontinuities (parameter P_2).

SGM is computationally efficient and has been used very successfully in aerial image 3D reconstruction. The algorithm has a real-time potential on pipe-line architectures such as graphic cards (GPU) or field-programmable gate-arrays (FPGA) and processes VGA image pairs in few seconds on standard CPUs.

A multi-baseline extension of SGM is straightforward and described in [5]. Instead of working on rectified images one walks along slanted epipolar lines of the undistorted image which can be done for all match images with one reference image. The resulting disparity map results are combined with a median filter. The algorithms described

below can be combined for a multi-baseline application in the same way, allowing large-scale reconstructions.

3. Sampling the Disparity at Fractional Disparities

As shown by Szeliski et al. [17], the sinc interpolator is the best interpolation to sample the disparity space. For image areas with sufficient texture this results in excellent subpixel disparities. However, in real images, a good portion of the image consists of low-textured or untextured areas. Due to noise, these areas do not benefit from this interpolation strategy but need support from a smoothness constraint. When working on discrete disparity steps, the number of subdivisions of one disparity step is the dominant factor for improvement in low-textured areas. We extend the analysis from [17] and perform an evaluation of the disparity steps at integer pixel, at half pixel and at quarter pixel level, considering the full image including low-textured regions. Using semi-global matching we present results that show a significant improvement in stereo accuracy for fractional disparity sampling. These results can be extended to any global stereo algorithm that incorporates smoothness constraints provided sufficient available memory.

The evaluation of the disparity space at fractional disparities is straightforward, we only need to interpolate our matching costs. Then, the original SGM algorithm is applied using the cost matrix with the disparity dimension increased by the sampling factor 1,2,or 4. We consider two similarity metrics: the Birchfield-Tomasi (BT) metric [3], and pixel-wise mutual information [7].

Taking the BT similarity metric as described in [3], we obtain the interpolated BT values at half pixel resolution by interpolating the grayvalues of the left and right image and then compute the BT metric with the interpolated values. In our investigations with ground truth data, we obtained a slight performance gain when evaluating BT with interpolated half-pixel values. When evaluating the disparity space at quarter pixel resolution, we interpolate the grayvalues linearly and compute the BT matching costs with interpolated half-pixel values.

For pixel-wise mutual information, which is in essence a lookup table, we implement a linear interpolation of the grayvalues in the match image. In Section 6 we show an analysis of the different sampling steps. Notice that the increase in computation time is linear with the number of (fractional) disparities to check.

4. Disparity Smoothing

Our second technique for improving the sub-pixel accuracy is based on the smoothness constraint. This constraint is widely used to obtain good disparity estimates but often on a pixel level only. All commonly used local interpolation schemes neglect this constraint while estimating depth with sub-pixel accuracy. Imagine a house at 50m distance and a stereo rig that yields a disparity of 5 pixels. An uncertainty of $\pm\frac{1}{4}px$ then corresponds to an uncertainty of $\pm\ 2.5m$ in depth. Obviously, such a reconstruction is not very likely to be accurate.

Our approach performs a depth-edge-preserving smoothing on the disparity image, similar to [21] where bilateral filtering was used. Bilateral filtering and adaptive smoothing are closely related to each other as shown in [2]. Our approach is similar to adaptive smoothing, however, unlike other methods we also exploit the confidence of an established disparity value.

We treat the disparity estimation as an energy minimization problem similar to [1] with:

$$E_{tot} = \sum_{x,y}(E_{data}(x,y) + \lambda E_{smooth}(x,y)). \qquad (1)$$

Let d_0 be the disparity of the central pixel of a given small patch. The empirical local disparity variance,

$$E_{smooth}(d_0) = \frac{1}{N-1}\sum_{i=0}^{N-1}(d_i - \overline{d})^2 = \sigma^2(d_0), \qquad (2)$$

is used as a smoothness energy contribution of each pixel. N is the number of pixels within the considered patch and $\overline{d}$ its the average disparity. The smoothness energy of d_0 becomes minimal for $d_0 = \overline{d}$.

How can an appropriate data term be formulated? Let d_{int} be the integer disparity computed by a stereo algorithm of your choice. The standard parabolic fit delivers not only an improved estimate d_{sub}, but also the curvature a of the parabola. The curvature is small in low-textured regions, whereas it is large in textured regions. If we define the data term according to

$$E_{data}(d_0) = a(d_0 - d_{sub})^2, \qquad (3)$$

the cost of choosing d_0 unequal to the locally estimated d_{sub} depends on the confidence of the fit. One could also incorporate the image gradient or the grayvalue variance as a confidence measure of the stereo fit.

It is easy to compute the best solution d_0 for a certain image point. Partial derivation $\partial E_{tot}/\partial d_0 = 0$ yields

$$d_0 = \frac{a d_{sub} + \frac{\lambda}{N}\overline{d}}{a + \frac{\lambda}{N}}. \qquad (4)$$

The higher that λ is chosen, the smoother is the resulting disparity image. Applying Equation 4 to the disparity image favors smooth reconstructions everywhere and ignores

the fact that the disparities can be discontinuous at object boundaries. Therefore, we modify Equation 2 according to

$$E_{smooth} = \sigma_{norm}^2 \arctan(\frac{\sigma^2(d_0)}{\sigma_{norm}^2}), \quad (5)$$

where σ_{norm} denotes the expected variance due to noise. In areas with small disparity variances ($\sigma \ll \sigma_{norm}$), the characteristic remains unchanged, whereas at disparity edges the slope of the function is small. That means that a change of d_0 has only little effect on the total energy, thus switching off the smoothing at disparity discontinuities.

Thanks to the easy derivative of the $\arctan$ function, the optimal d_0 becomes:

$$d_0 = \frac{ad_{sub} + H\overline{d}}{a + H} \; with \; H = \frac{\lambda}{N} \frac{1}{1 + (\sigma/\sigma_{norm})^4}. \quad (6)$$

H expresses the homogeneity (approaching 0 for depth discontinuities, 1 for homogeneous areas) of the disparity values in the considered patch. λ is constant and gives equal weight to all disparities. For 3D reconstructions at varying distances, we to set $\lambda(d_{sub}) = \frac{\lambda}{d_{sub}}$ since this causes equal smoothing in Euclidean space for all distance ranges.

The above equations are applied to small patches within the disparity image. In order to get close to the optimal solution of the above stated problem, we need to iterate Equation 6 to propagate the updated disparity values. d_{sub} remains the original value of the input disparity image, whereas d_0 is updated in every iteration.

This disparity smoothing algorithm favors solutions that are planar in 3D, i.e. fronto-parallel or slanted planes. This way, the algorithm is especially helpful for reconstructing buildings.

See Section 6.2 for results of this disparity smoothing algorithm.

5. Gravitational Constraint

In Figure 1 an urban scene is shown where the sky yields some erroneous correspondences due to matching ambiguities. For cameras where the optical axes point to the horizon, the following observation is made: When traversing the image from the bottom to the top, the distance to the 3D points on the rays tend to increase (see Figure 2). This is due to the fact that objects in the scene are connected to the ground (due to gravity, therefore we refer to the constraint as gravitational constraint). Summarized, we assume sorted depths for every column in the image. If the optical axes and the ground have a small angle, the constraint is still valid for most scenes. The constraint is absolutely useless for scenes viewed from the bird's eye perspective.

Even for the scenario with the optical axes parallel to the ground, scene elements who violate the constraint can

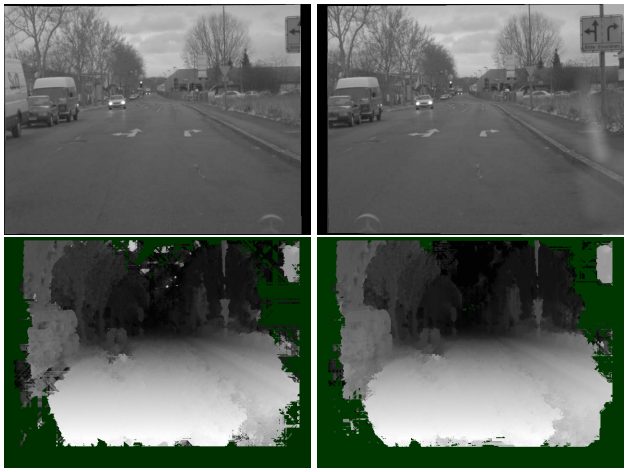

Figure 1. Street scene with traffic sign - left image shown on the top left, right image on the top right. The disparity image using SGM with BT is shown at the bottom left, the bottom right image shows the result with gravitational constraint. Note the wrong depth estimation in the sky on the left. Dark green marks unmatched areas.

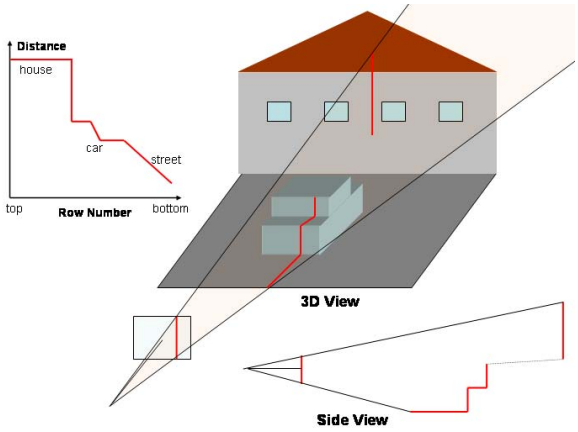

Figure 2. Principle of the Gravitational Constraint.

easily be found: Traffic signs, bridges, or ceilings in indoor environments. So, this can only be incorporated as a weak constraint to disambiguate matches in low-textured areas similar to the smoothness constraint.

For semi-global matching this constraint is easy to apply since the algorithm works on multiple 1D paths. For the bottom-up path the original cost accumulation of (see [5], Equation 12)

$$L'_r(\mathbf{p}, d) = C(\mathbf{p}, d) + \min(L'_r(\mathbf{p} - r, d), L'_r(\mathbf{p} - r, d - 1) + P_1, L'_r(\mathbf{p} - r, d + 1) + P1, \min_i L'_r(\mathbf{p} - r, i) + P_2) \quad (7)$$

is modified to the following term:

$$
\begin{aligned}
&if\ d_{best}+1 \geq d: \\
&L'_r(\mathbf{p},d) = C(\mathbf{p},d) + \min(L'_r(\mathbf{p}-r,d), L'_r(\mathbf{p}-r,d-1) \\
&+P_{G1}, L'_r(\mathbf{p}-r,d+1) + P1, \min_i L'_r(\mathbf{p}-r,i) + P_2) \\
&else: \\
&L'_r(\mathbf{p},d) = C(\mathbf{p},d) + min(L'_r(\mathbf{p}-r,d), L'_r(\mathbf{p}-r,d-1) \\
&+P_{G1}, L'_r(\mathbf{p}-r,d+1) + P1, \min_i L'_r(\mathbf{p}-r,i) + P_{G2}),
\end{aligned}
\tag{8}
$$

with $P_{Gi} = P_G \cdot P_i$. L'_r represents a possible path ($L'_0..L'_{15}$), r is the previous pixel to $\mathbf{p}$ along the path, d_{best} is the disparity position of the lowest cost from the previously visited pixel, $P_G > 1$ is the penalty factor for violating the gravitational constraint. The top-down path is modified in the same way. All other paths keep their original shape of Equation 7. In Section 6.3 we show the effect of this constraint compared to ground truth data and on urban scenes.

6. Results

6.1. Results for Fractional Disparity Sampling (FDS)

The benefit of evaluating the disparity space at fractional disparities is shown in Table 1. The parameters P_1 and P_2 of Equation 7 were kept constant at 10 and 20, respectively. Adaptive P_2 and hierarchical mutual information as explained in [5] was used. Discontinuity preserving interpolation as explained in [6] was applied in the post-processing step. The remaining disparities are filled with the disparities from the neighbors to the left or right, so no intelligent extrapolation is performed.

The computation time for Cones and Teddy is less than 7s for the quarter pixel resolution. We count all pixels exceeding 0.5px from ground truth as error for the remainder of the paper. The RMS is computed based on all pixels, so a high value occurs due to a few mismatched pixels, but the reduction due to sub-sub-pixel estimation (parabola interpolation and fractional disparity evaluation) can clearly be seen. Excluding outliers in the RMS computation has the problem that the number of outliers change. The RMS at quarter pixel accuracy slightly degrades for the Cones and Teddy set which could be accounted for the ground truth being accurate to a quarter pixel only [14].

The evaluation of the disparity space at fractional disparities is especially beneficial at large distances where one disparity step corresponds to several meters in 3D. One example is shown in Figure 3 and 4.

One can see clearly that the rear of the truck is excellently reconstructed for the quarter pixel case. This accuracy gain in stereo facilitates a subsequent 3D analysis to detect objects. The separation of two objects at similar distances becomes much easier.

FDS	1px		0.5px		0.25px	
	% err	RMS	% err	RMS	% err	RMS
Tsu	23.72	1.117	20.30	0.966	14.99	0.919
Ven	10.80	0.643	8.60	0.502	8.17	0.330
Ted	23.71	8.782	21.16	8.731	19.96	9.674
Con	15.04	5.194	12.98	4.861	12.57	5.302

Table 1. Error percentages and RMS (in px^2) for stereo ground truth data sets. The results get better with higher fractional disparity sampling (FDS). Tsu=Tsukuba, Ven=Venus, Ted=Teddy, Con=Cones image pair.

Figure 3. Van scene (rectified left image shown on the left) and corresponding disparity scene (right). Here, SGM with BT is used. The cameras deliver 12 bits per pixel and have a 55cm baseline with a 26° field of view. Dark green marks unmatched areas.

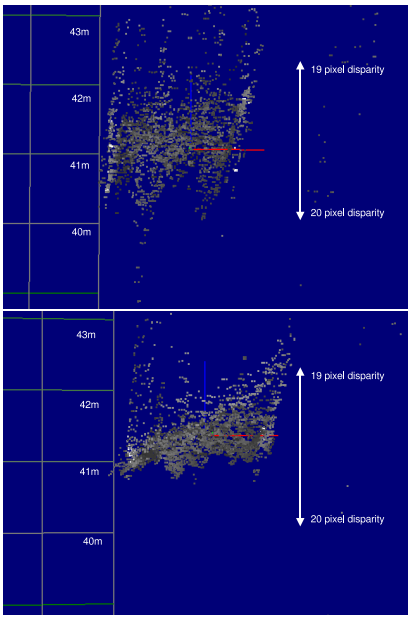

Figure 4. 3D reconstruction of the rear of the van from the previous figure viewed from the bird's eye view. We used pixel disparity sampling at the top and quarter pixel sampling at the bottom image. The result with fractional disparity sampling exhibits significantly less noise. The van has a perfectly planar rear.

6.2. Disparity Smoothing Results

In Table 2 the RMS and error percentages for the disparity smoothing algorithm is shown on the stereo ground truth

FDS	1px		0.5px		0.25px	
	% err	RMS	% err	RMS	% err	RMS
Tsu	22.82	1.083	19.87	0.943	15.75	0.857
Ven	10.21	0.413	8.46	0.342	8.70	0.328
Ted	24.17	7.701	21.97	7.183	22.17	7.046
Con	15.74	4.912	13.86	4.334	14.78	4.596

Table 2. Error percentages and RMS (in px^2) for stereo ground truth data sets. The RMS results are better with disparity smoothing compared to Table 1.

data set. The RMS decreases for all image and sampling configurations compared to Table 1 while the error percentage occasionally increases. We use parameters $\lambda = 50000$, $\sigma_{norm} = 0.01$, a mask size of 3x3 and 10 iterations for the disparity smoothing. The computation time is 130ms for the Cones images.

The benefit of disparity smoothing becomes obvious when looking at slanted surfaces. We investigated a part of the Venus stereo pair where an accuracy of 1/8 pixel is available. Taking the lower right image portion of the Venus image pair, we observe an improvement from 7% false pixels to 4% and an RMS drop from 0.080 to 0.057 px^2.

Figure 5 shows a 3D reconstruction of the van rear shown in Figure 3. The rear is reconstructed to an almost planar surface keeping smoothing artifacts minimal in the surrounding street or sky.

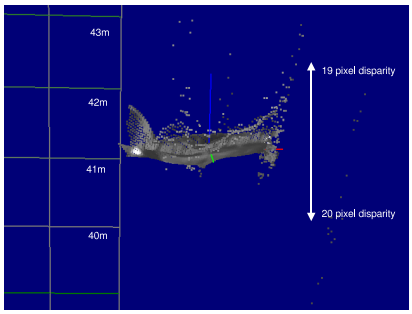

Figure 5. 3D reconstruction of the van rear shown above viewed from the bird's eye perspective. Disparity smoothing yields an almost planar surface. A video of the three presented 3D reconstructions can be found in the supplementary material.

In Figure 6 a facade reconstruction using this algorithm is shown. The walls appear very planar and the two walls meet at a 90° angle. Note the clean separation between the staircase and the wall. The eaves gutter on the right wall is also well preserved in its shape.

6.3. Gravitational Constraint Results

Table 3 shows the performance gain when applying the gravitational constraint to the ground truth data set. Here we did not apply the parabola interpolation and used a penalty factor of $P_G = 3$.

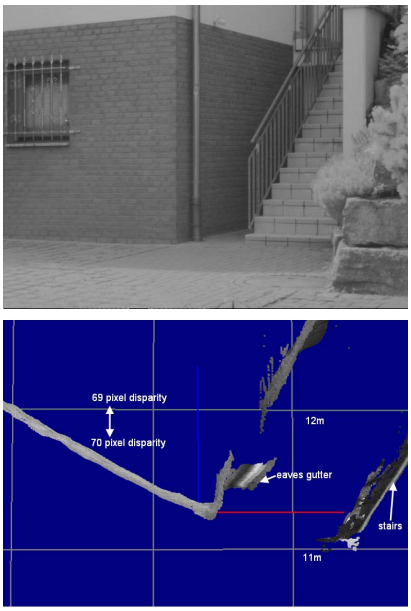

Figure 6. Top: Left image of an image pair viewing a building at 12m distance. Bottom: 3D reconstruction of the building viewed from the bird's eye perspective. Disparity smoothing yields almost planar building surfaces. The cameras deliver 12 bits per pixel and have a 55cm baseline with a 26° field of view.

FDS	0.25px		0.25px + Gravitation	
	% err	RMS	% err	RMS
Tsu	7.83	0.9203	7.66	0.9682
Ven	7.35	0.3493	6.64	0.3115
Ted	17.54	9.0838	17.43	7.1768
Con	10.76	4.9547	10.42	4.3229

Table 3. Comparison of error percentages and RMS (in px^2) of stereo ground truth data. Incorporating the gravitational constraint yields consistently better results.

At time of publication our algorithm using the gravitational constraint and fractional sampling, *ImproveSubPix*, was among the top-performing stereo algorithms on subpixel level [12], being best for the Cones image pair. The improved SGM by Hirschmueller ([6] performs similarly, but uses several additional post-processing steps.)This evaluation is based on a ground truth accuracy of a quarter pixel or better.

The gravitational constraint is very helpful in urban scenes to disambiguate matches in low-textured areas, such as the street and the sky. One example is shown in Figure 1. The original SGM result produces wrong matches in the sky (BT metric was used). These mismatches can be corrected with the gravitational constraint. Often, such wrong estimations are removed by the right-left consistency check [5], which is performed here. Note that the traffic sign at the

right side is still correctly measured although it violates the gravitational constraint. The computational overhead incorporating this constraint is negligible.

7. Conclusions and Future Work

Summarizing, we have shown three ways to improve sub-pixel stereo computation:

- Evaluate the disparity space image at fractional disparities.
- Apply the disparity smoothing algorithm to obtain a spatial connection between adjacent pixels on a sub-pixel level.
- Exploit the gravitational constraint if the camera setup is designed that way.

All three improvements can be combined independently of each other. To save computation time, it might be appealing to use the disparity smoothing algorithm without fractional disparity sampling. The gravitational constraint is especially helpful to disambiguate matches in low-textured areas. In addition, these improvements are independent of the choice of stereo algorithm. For stereo algorithms exploiting constraints only along the epipolar line, it might be beneficial to incorporate the gravitational constraint and disparity smoothing to obtain a vertical coupling.

For multi-baseline matching, the same improvements shown for two-view stereo are expected.

Finally, looking at our evaluation against ground truth, it becomes obvious that more accurate ground truth data would be extremely valuable to further push the limits of stereo accuracy.

References

[1] L. Alvarez, R. Deriche, J. Sanchez, and J. Weickert. Dense disparity map estimation respecting image discontinuities: A pde and scale-space based approach. Technical report, Research Report 3874, INRIA Sophia Antipolis, France, January 2000. 2, 3

[2] D. Barash and D. Camiciu. A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift. *Image and Vision Computing*, 22(1):73–81, 2004. 3

[3] S. Birchfield and C. Tomasi. Depth discontinuities by pixel-to-pixel stereo. In *Proceedings of Int. Conference on Computer Vision 98*, pages 1073–1080, 1998. 3

[4] J. Chen and J. Katz. Elimination of peak-locking error in piv analysis using the correlation method. *Meas.Sci.Technol.*, 16:562–574, 2005. 2

[5] H. Hirschmueller. Accurate and efficient stereo processing by semi-global matching and mutual information. In *Proceedings of Int. Conference on Computer Vision and Pattern Recognition 05, San Diego, CA*, volume 2, pages 807–814, June 2005. 2, 4, 5, 6

[6] H. Hirschmueller. Stereo vision in structured environments by consistent semi-global matching and mutual information. In *Proceedings of Int. Conference on Computer Vision and Pattern Recognition 06, New York, NY*, pages 2386–2393, June 2006. 5, 6

[7] J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. In *Proceedings of Int. Conference on Computer Vision 03*, pages 1033–1040, 2003. 3

[8] A. Klaus, M. Somann, and K. Kramer. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *International Conference on Pattern Recognition*, 2006. 2

[9] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Int. Joint Conference on Artificial Intelligence*, pages 674–679, 1981. 2

[10] D. Nehab, S. Rusinkiewicz, and J. Davis. Improved sub-pixel stereo correspondences through symmetric refinement. In *Proceedings of Int. Conference on Computer Vision 05*, pages 557–562, 2005. 1, 2

[11] E. Psarakis and G. Evangelidis. An enhanced correlation-based method for stereo correspondence with sub-pixel accuracy. In *Proceedings of Int. Conference on Computer Vision 05*, pages 907–912, 2005. 2

[12] D. Scharstein and R. Szeliski. Middlebury online stereo evaluation. http://www.middlbury.edu/stereo. 1, 6

[13] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1):7–42, 2002. 1, 2

[14] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proceedings of Int. Conference on Computer Vision and Pattern Recognition 03, Madison, Wisconsin*, pages 195–202, 2003. 5

[15] M. Shimizu and M. Okutomi. Precise sub-pixel estimation on area-based matching. In *ICCV*, pages 90–97, 2001. 1, 2

[16] A. Stein, A. Huertas, and L. Matthies. Attenuating stereo pixel-locking via affine window adaption. In *International Conference on Robotics and Automation*, pages 914–921, May 2006. 2

[17] R. Szeliski and D. Scharstein. Sampling the disparity space image. *PAMI*, 26(3):419–425, 2004. 1, 2, 3

[18] M. Tappen and W. Freeman. Comparison of graph cuts with belief propagation for stereo, using identical mrf parameters. In *Proceedings of Int. Conference on Computer Vision 03*, pages 900 – 907, 2003. 2

[19] Q. Tian and M. Huhns. Algorithms for subpixel registration. *Computer Vision, Graphics, and Image Processing*, 35:220–233, 1986. 1, 2

[20] J. Westerweel. Effect of sensor geometry on the performance of piv interrogation. In *Laser Techniques Applied to Fluid Mechanics*, pages 37–55. Springer-Verlag, 2000. 2

[21] Q. Yang, R. Yang, J. Davis, and D. Nister. Spatial-depth super resolution for range images. In *Proceedings of Int. Conference on Computer Vision and Pattern Recognition 07*, June 2007. 2, 3