Evaluation of Cost Functions for Stereo Matching

Heiko Hirschmüller
Institute of Robotics and Mechatronics Oberpfaffenhofen
German Aerospace Center (DLR)
heiko.hirschmueller@dlr.de

Daniel Scharstein
Middlebury College
Middlebury, VT, USA
schar@middlebury.edu

Abstract

Stereo correspondence methods rely on matching costs for computing the similarity of image locations. In this paper we evaluate the insensitivity of different matching costs with respect to radiometric variations of the input images. We consider both pixel-based and window-based variants and measure their performance in the presence of global intensity changes (e.g., due to gain and exposure differences), local intensity changes (e.g., due to vignetting, non-Lambertian surfaces, and varying lighting), and noise. Using existing stereo datasets with ground-truth disparities as well as six new datasets taken under controlled changes of exposure and lighting, we evaluate the different costs with a local, a semi-global, and a global stereo method.

1. Introduction and Related Work

All stereo correspondence algorithms have a way of measuring the similarity of image locations. Typically, a *matching cost* is computed at each pixel for all disparities under consideration. The simplest matching costs assume constant intensities at matching image locations, but more robust costs model (explicitly or implicitly) certain radiometric changes and/or noise. Common pixel-based matching costs include absolute differences, squared differences, sampling-insensitive absolute differences [2], or truncated versions, both on gray and color images. Common window-based matching costs include the sum of absolute or squared differences (SAD / SSD), normalized cross-correlation (NCC), and rank and census transforms [23]. Some window-based costs can be implemented efficiently using filters. For example, the rank transform can be computed using a rank filter followed by absolute differences of the filter results. Similarly, there are other filters that try to remove bias or gain changes, e.g., LoG and mean filters.

More complicated similarity measures are possible, including mutual information [7, 9, 11] and approximative segment-wise mutual information as used in the layered stereo approach of Zitnick et al. [24].

Recent stereo surveys [5, 17] and the Middlebury online evaluation [14] compare state-of-the-art stereo methods on test data with complex geometries and varied texture. Other evaluations focus on certain aspects like aggregation methods for real-time matching [21]. However, the insensitivity of matching costs is not evaluated since the stereo test sets are typically pairs of radiometrically very similar images.

The term *radiometrically similar* means that pixels that correspond to the same scene point have similar or ideally the same values in the images. Radiometric differences can be caused by the camera(s) due to slightly different settings, vignetting, image noise, etc. Further differences may be due to non-Lambertian surfaces, which make the amount of reflected light dependent on the viewing angle. Finally, the strength or positions of the light sources may change when images of a static scene are acquired at different times, as is the case when matching aerial or satellite images. In all cases, methods are required that can handle radiometric differences.

The scope of this paper is the evaluation and comparison of some widely used stereo matching costs on images with several common radiometric differences. The focus is on matching costs that explicitely or implicitly handle radiometric differences. This excludes popular methods like the correlation-based weighting according to proximity and color similarity [22], as this is an aggregation approach that uses the truncated absolute difference as matching cost. Furthermore, only methods that work on a single stereo pair with unknown radiometric distortions and light sources are evaluated, according to the considered applications. This excludes methods that explicitly handle non-Lambertian surfaces by taking at least two stereo images with different illuminations [6] or methods that require calibrated light sources.

2. Matching Costs and Stereo Methods

It is important to distinguish between matching costs and methods that use these costs. In this paper we compare 6 costs and 3 stereo methods. We consider all possible combinations to fully evaluate the insensitivity of each cost.

2.1. Matching Costs

Our first cost function is the commonly-used *absolute difference*, which assumes brightness constancy for corresponding pixels, and which serves as a baseline performance measure of our evaluation. Local stereo methods usually aggregate the sum of absolute differences (SAD) over a window, while global methods use the differences pixel-wise. In both cases we use the sampling-insensitive calculation of Birchfield and Tomasi (BT) [2].

Our next three cost functions can be implemented as filters that are applied separately to the input images. The transformed images are then matched using the absolute difference. The first filter is the *Laplacian of Gaussian* (LoG), which is often used in local methods for removing noise and changes in bias [10, 13]. Here we use a LoG filter with a standard deviation of 1 pixel, which is applied by convolution with a 5×5 kernel. The second filter is the *rank filter*, which replaces the intensity of a pixel with its rank among all pixels within a certain neighborhood. It was originally proposed [23] for robustness to outliers within the neighborhood, which typically occur near depth discontinuities and leads to blurred object borders. Since the method only depends on the ordering of intensities and not their values, it compensates for all radiometric distortions that preserve this ordering. Here we use a rank filter with a square window of 15×15 pixels centered at the pixel of interest. While there are other rank-based matching methods [1, 16], we chose the rank transform since it can be efficiently implemented as filter, without changing the stereo method itself. The third filter is a *mean filter*, which aims to compensate a change in bias by subtracting the mean intensity of a certain neighborhood. We again use a square window of size 15×15 that is centered at the pixel of interest.

Our next matching cost is *mutual information* (MI), a powerful method for handling complex radiometric relationships between two images [20]. The MI of two images is calculated by summing the entropy of the histograms of the overlapping parts of each image and subtracting the entropy of the joint histogram of pixel-wise correspondences. The MI value directly expresses how well images are registered. This follows from the observation that the joint histogram of well-registered images has just a few high peaks in contrast to poorly registered images where the joint histogram is rather flat. Thus, for well-registered images, the entropy of the joint histogram is low, while the entropy of the individual histograms changes little. MI has been used for local [7] and global [11] stereo methods. In the latter case, its calculation is changed by Taylor expansion for getting a pixel-wise matching cost. The costs are stored for each combination of intensities in a cost matrix. This lookup table is required for matching, but can only be created from known correspondences. The solution is an iterative design in which the disparity image of the previous loop serves for creating the cost matrix for matching intensities in the next loop [11]. The process is started with a random disparity image and requires typically only 3 to 4 iterations.

In this paper we use the efficient Hierarchical MI (HMI) method of [9], which works as follows. First, both input images are downscaled by factor 16 and MI is calculated by matching the stereo images using a random disparity image. The process is iterated a few times before the disparity is upscaled for serving as initial guess for matching at $\frac{1}{8}$th of the full resolution. Upscaling and matching is repeated until the full resolution is reached. It should be noted that the disparity image of the lower-resolution level is used only for calculating the matching costs of the higher-resolution level, but not for restricting the disparity range. The hierarchical calculation has a runtime overhead of just 14% if the runtime of the stereo method depends linearly on the number of pixels and disparities [9].

Finally, we also include normalized cross-correlation (NCC) in our evaluation. NCC is a standard method for matching two windows around a pixel of interest. The normalization within the window compensates differences in gain and bias. NCC is statistically the optimal method for compensating Gaussian noise. However, NCC tends to blur depth discontinuities more than many other matching costs, because outliers lead to high errors within the NCC calculation. MNCC has been introduced as a common variant by Moravec [15]. We selected the standard NCC as MNCC gave slightly inferior results in our experiments. In contrast to all other matching costs we consider here, NCC can only be used with local methods due to its window-based design.

In all of the above costs, we only use the image intensity (luminance) and not the color for matching. The reason is that several of the considered costs (e.g., rank and MI) are naturally defined on intensity images, and for fairness we want to compare all costs on the same input data. However, we also found that that those costs that easily extend to color only perform marginally better on our data sets. Clearly, future research is needed on robust color matching.

To summarize, we compare six costs: sampling-insensitive absolute differences (**BT**), three filter-based costs (**LoG**, **Rank**, and **Mean**), hierarchical mutual information (**HMI**), and normalized cross-correlation (**NCC**).

2.2. Stereo Algorithms

The performance of a matching cost can depend on the algorithm that uses the cost. We thus consider three different stereo algorithms: a local, correlation-based method (**Corr**), the semi-global method of [9] (**SGM**), and a global method using graph cuts [4] (**GC**). We implemented each of the six matching costs for each stereo method, except for NCC which is only used with the local method.

Our local stereo method (Corr) is a simple window-based approach [10, 13, 17]. After aggregating the matching cost

Figure 1. The left images of the Tsukuba, Venus, Teddy, and Cones stereo pairs.

with a square window of 9×9 pixels, the disparity with the lowest aggregated cost is selected (winner-takes-all). This is followed by subpixel interpolation, a left-right consistency check for invalidating occlusions and mismatches, and invalidation of disparity segments smaller than 160 pixels [8]. Invalid disparity areas are filled by propagating neighboring small (i.e., background) disparity values. The reason we perform these post-processing steps, as opposed to comparing the "raw" results, is to reduce the overall errors, which in turn yields improved discrimination between costs.

Our second stereo algorithm is the semi-global matching (SGM) method [9]. We selected it as an approach in-between local and global matching. There are other approaches in this category, e.g., dynamic programming (DP), but SGM outperforms DP and yields no streaking artefacts. SGM aims to minimize a global 2D energy function $E(D)$ by solving a large number of 1D minimization problems. Following [9], the actual energy used is

$$E(D) = \sum_{\mathbf{p}} \Big(C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N_{\mathbf{p}}} P_1 \, \mathrm{T}[|D_{\mathbf{p}} - D_{\mathbf{q}}| = 1] + \sum_{\mathbf{q} \in N_{\mathbf{p}}} P_2 \, \mathrm{T}[|D_{\mathbf{p}} - D_{\mathbf{q}}| > 1] \Big). \quad (1)$$

The first term of (1) calculates the sum of a pixel-wise matching cost $C(\mathbf{p}, D_{\mathbf{p}})$ (as defined in Section 2.1) for all pixels $\mathbf{p}$ at their disparities $D_{\mathbf{p}}$. The function $\mathrm{T}[]$ is defined to return 1 if its argument is true and 0 otherwise. Thus, the second term of the energy function penalizes small disparity differences of neighboring pixels $N_{\mathbf{p}}$ of $\mathbf{p}$ with the cost P_1. Similarly, the third term penalizes larger disparity steps (i.e., discontinuities) with a higher penalty P_2. The value of P_2 is adapted to the local intensity gradient by $P_2 = \frac{P_2'}{|I_{b\mathbf{p}} - I_{b\mathbf{q}}|}$ for the neighboring pixels $\mathbf{p}$ and $\mathbf{q}$. This results in sharper depth discontinuities as they mostly coincide with intensity variations.

SGM calculates $E(D)$ along 1D paths from 8 directions towards each pixel of interest using dynamic programming. The costs of all paths are summed for each pixel and disparity. The disparity is then determined by winner-takes-all. Subpixel interpolation is performed as well as a left-right

consistency check. Disparity segments below the size of 20 pixels are invalidated for getting rid of small patches of outliers. Invalid disparities are again interpolated.

Finally, we use a graph-cuts (GC) stereo algorithm as a representative of a global method [3, 4, 12]. Our implementation is based on the MRF library provided by [19]. We tried to use the same energy function $E(D)$ as for SGM. However, we found that for GC it gives better results to adapt the cost P_2 not linearly with the intensity gradient, but rather to double the value of P_2 for gradients below a given threshold. Like SGM, GC only approximates the global minimum of $E(D)$, but it utilizes the full 2D connectivity for the smoothness term in contrast to SGM, which optimizes separately along 1D paths. Our GC implementation, unlike Corr and SGM, neither includes subpixel interpolation nor accounts for occlusions.

We manually tuned the smoothness parameters of SGM and GC individually for each cost using images without radiometric differences. After the tuning phase, all parameters were kept constant for all images and experiments. This approach allows to concentrate on the performance of the matching cost rather than the stereo method.

3. Evaluation

We tested all combinations of all matching costs with the local, semi-global, and global stereo algorithms on images with simulated and real radiometric changes.

3.1. Simulated Radiometric Changes

For our first set of experiments, we use the standard Middlebury stereo datasets Tsukuba, Venus, Teddy, and Cones [17, 18]. Figure 1 shows the left images of each set. All images were carefully taken in a laboratory with the same camera settings and under the same lighting conditions. Therefore radiometric changes are expected to be minimal. We used a disparity range of 16 pixels for Tsukuba, 32 pixels for Venus and 64 pixels for Teddy, and Cones.

The first experiments consist of artificially changing the global brightness linearly (i.e., gain change) and nonlinearly (e.g., gamma change). Only the right stereo images were changed, while leaving the left images untouched.

Furthermore, we applied a local brightness change that mimics a vignetting effect, i.e., the brightness decreases proportionally with the distance to the image center. This transformation was performed on both stereo images. Finally, we contaminated both stereo images with different levels of Gaussian noise.

After computing disparity images for all transformations and all combinations of matching costs and stereo algorithms, we evaluate the results by counting the number of pixels with disparities that differ by more than 1 from the ground truth. In our statistics we ignore occluded areas because the GC implementation does not consider occlusions (in contrast to Corr and SGM). For the correlation results we also ignore an area of 4 pixels (half of the correlation window) at the image border. Our final error measure is the mean error percentage over all four datasets. Figure 2 plots these errors as a function of the amount of intensity change for each combination of matching cost and stereo method. We now discuss the individual results.

Figure 2a compares the matching costs when used with correlation on images with decreasing brightness. The errors of BT increase very quickly with decreasing brightness. This can be expected, because the absolute difference is based on the assumption that corresponding pixels have the same values, which is violated. The Mean and LoG filters can compensate some of the differences, but degrade quickly when $s < 0.5$. Both filters are designed for compensating a bias (i.e., constant offset), but not a gain (i.e., scale) change. NCC, HMI and Rank show a quite constant performance, until the errors suddenly increase. Theoretically, all three methods should be able to fully compensate the brightness change. The reason for the increased error is that the transformed images are stored into 8 bits. Thus, there is also an information loss, with low values of s.

Moving on to the next two plots, one can see that SGM and GC (Figure 2b–c) generally perform better than correlation. The relative performance of the different matching costs remains similar, although for SGM the LoG cost is now slightly better than Rank on the non-transformed images (i.e., for a scale factor of 1). A more important observation is that HMI performs worse than Rank with correlation, but much better with SGM and GC. The likely reason is that Rank also reduces the effect of outliers near depth discontinuities. This is important for a window-based method, but less so for pixel-based methods like SGM and GC. It is interesting that on the non-transformed images, HMI performs better than BT, especially for SGM and GC (Figure 2b–c). One might assume that BT should be best for images without any radiometric differences. However, even though the images have been taken under controlled conditions, some radiometric differences are inherent and surfaces are not Lambertian, and the brightness constancy assumption is still violated. HMI relaxes this assumption

and only expects a globally consistent mapping.

The next three plots (Figure 2d–f) show the effect of a gamma change as an example of a non-linear change of brightness. The results are similar to the case of a linear change, although the performance of NCC degrades with increasing gamma changes.

The artificial vignetting effect (Figure 2g–i) gives very similar curves compared to the global brightness changes, except for HMI. The reason for the rather bad performance of HMI is that its cost is explicitly based on the assumption of a complex, but *global* radiometric transformation. The vignetting effect locally changes the brightness. The filter solutions LoG and Mean also assume global changes, but only inside their rather small windows. Furthermore, Rank only requires an unchanged order, which is maintained. Therefore, the filter solutions and especially Rank are best in case of strong local radiometric variations.

Finally, the results for additive Gaussian noise with varying signal-to-noise ratios (SNR) are shown in the last three plots (Figure 2j–l). Higher SNR numbers mean lower noise. For correlation the different costs perform quite similar, probably since summing over a fixed window acts like averaging, which reduces the effect of Gaussian noise. The situation is different for SGM and GC, where LoG, Rank, and Mean perform even worse than BT. HMI performs consistently best for SGM and GC on all noise levels.

To summarize the above experiments, Rank appears to be the best matching cost for correlation based methods. HMI appears to be best for pixel-based matching methods like SGM and GC in the presence of global brightness changes and noise. In the case of local brightness variations such as vignetting, Rank and LoG appear to be better alternatives than HMI.

3.2. Real Exposure and Light Source Changes

As noted in the introduction, existing stereo test datasets are unusually radiometrically "clean" and do not require robust matching costs necessary for real-world stereo applications (unless, as in the previous section, changes are introduced synthetically). To remedy this situation we have created several new stereo datasets with ground truth using the structured lighting technique of [18], which are available at http://vision.middlebury.edu/stereo/data/. In this paper we use the six datasets shown in Figure 3: Art, Books, Dolls, Laundry, Moebius, and Reindeer. Each dataset consists of 7 rectified views taken from equidistant points along a line, as well as ground-truth disparity maps for viewpoint 2 and 6. In this paper we only consider binocular methods, so we use images 2 and 6 as left and right input images. Also, we downsample the original images to one third of their size, resulting in images of roughly 460×370 pixels with a disparity range of 80 pixels. When creating the datasets, we took each image using three different expo-
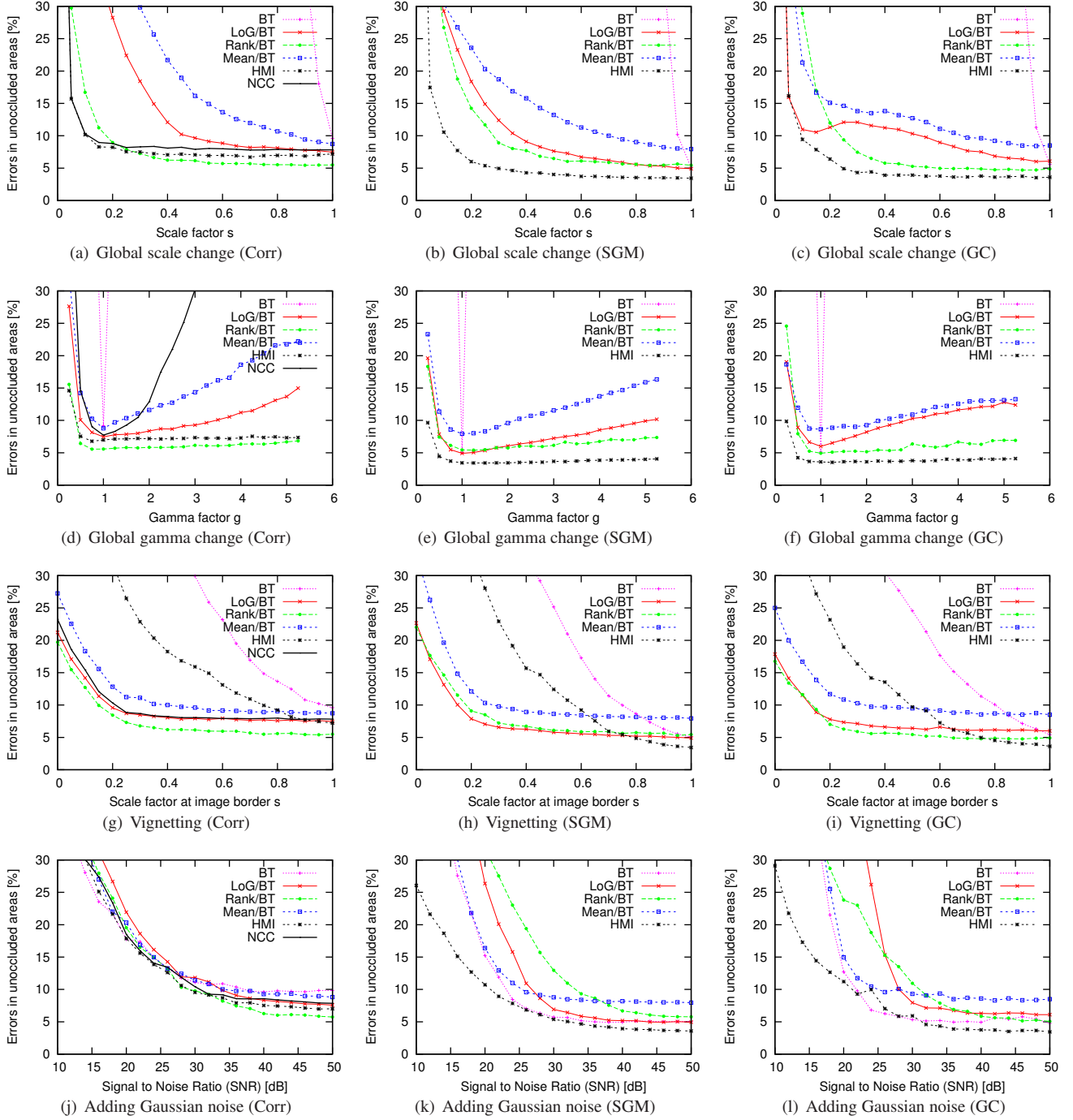
Figure 2. Effect of applying radiometric changes or noise on the Tsukuba, Venus, Teddy, and Cones datasets. The columns correspond to the three stereo methods, while each row examines a different type of intensity change.

sures and under three different configurations of the light sources. We thus have 9 different images from each viewpoint that exhibit significant radiometric differences. Figure 4 shows both exposure and lighting variations of the left image of the Art dataset.

We tested all combinations of matching costs and stereo algorithms over all 3×3 combinations of exposure and light changes. The total matching error is calculated as before as the mean percentage of outliers (disparity error > 1) over all six datasets. The resulting curves are shown in Figure 5. It should be noted that our new images are more challenging than the images used in Section 3.1, due to the increased

Figure 3. The new Art, Books, Dolls, Laundry, Moebius, and Reindeer stereo test images with ground truth.

Figure 4. The left image of the Art dataset with three different exposures and under three different light conditions.

(a) Different exposure (Corr)

(b) Different exposure (SGM)

(c) Different exposure (GC)

(d) Different lighting (Corr)

(e) Different lighting (SGM)
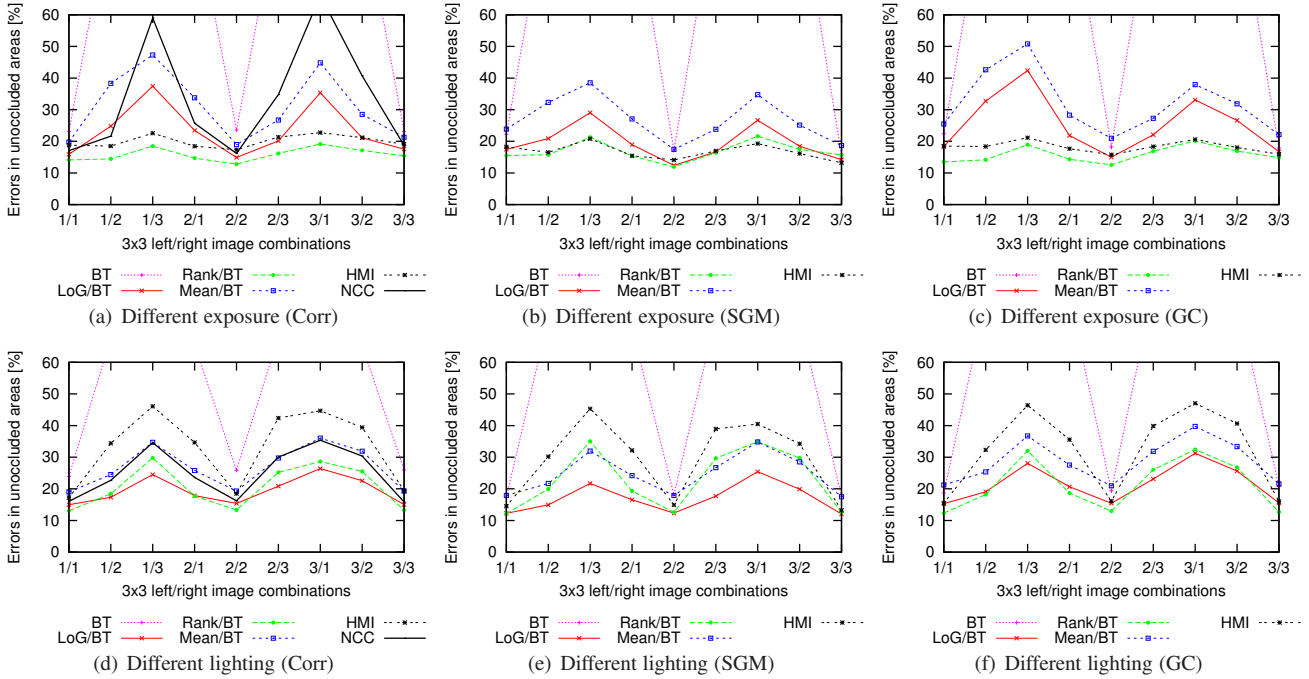
(f) Different lighting (GC)

Figure 5. Matching 3×3 left/right image combinations that differ in exposure or lighting conditions. Given is the mean error.

disparity range, lack of texture, and the more complicated scene geometry. This is reflected in the higher matching errors: the best methods now have errors of about 10%, as opposed to about 3% before.

Figure 5a shows the result of using all matching costs with correlation on pictures with different exposure settings. The change of exposure is a global transformation, which is similar to a global change of brightness. It is therefore not surprising that the observations are similar to that of Fig-

ures 2a and 2d, i.e., Rank and HMI perform best, while LoG and Mean have a bit more problems since they were not designed for compensating gain changes. The main difference is the bad performance of NCC. The performance of the matching costs with SGM and GC is shown in Figure 5b–c. In contrast to Section 3.1, Rank performs slightly better than HMI, especially in combinations with the same exposure settings. Furthermore, the performance of Rank in case of correlation is just slightly worse than Rank in com-

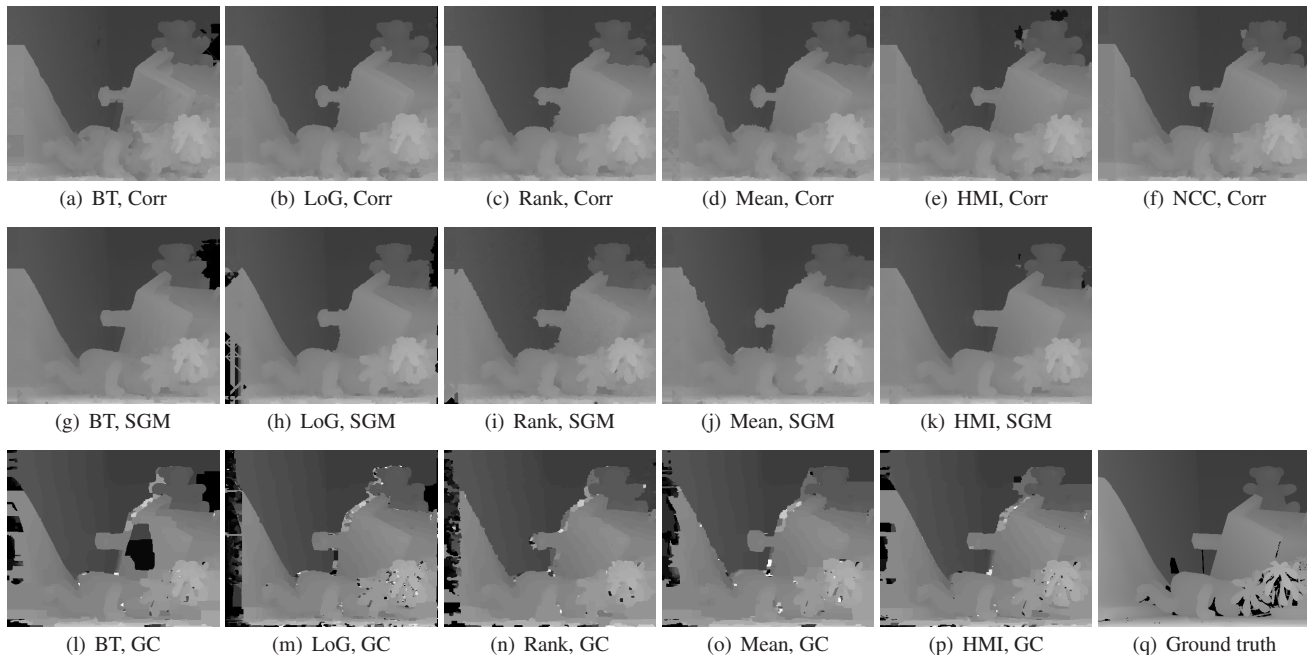|					
(a) BT, Corr	(b) LoG, Corr	(c) Rank, Corr	(d) Mean, Corr	(e) HMI, Corr	(f) NCC, Corr
(g) BT, SGM	(h) LoG, SGM	(i) Rank, SGM	(j) Mean, SGM	(k) HMI, SGM	
(l) BT, GC	(m) LoG, GC	(n) Rank, GC	(o) Mean, GC	(p) HMI, GC	(q) Ground truth

Figure 6. Disparity images of the Teddy pair without radiometric transformations.

bination with global methods. The reason is probably that the robustness of Rank against outliers helps in both cases for these challenging scenes.

Changing the position of light sources results in many local radiometric differences. The correlation results (Figure 5d) confirm that all matching costs have problems with these severe changes. However, Rank and LoG are again best, while HMI is only better than BT. This is essentially the same finding as from the artificial vignetting effect (Figure 2g). The situation is similar for SGM and GC (Figure 5e–f), where also Rank and LoG are the best-performing costs. HMI has obviously more problems with changed lighting which results in local brightness changes that cannot be expressed as a global transformation.

3.3. Qualitative Evaluation

We noted in Section 3.1 that there are performance differences among the matching costs even on the original images without any additional radiometric transformations. For a qualitative analysis of these differences we examined the disparity maps of all combinations of matching costs and algorithms on all image pairs.

Due to space limitations we show the complete set of disparity maps only for the Teddy image pair (Figure 6). At first glance, all correlation results appear similar (Figure 6a–f). However, a few more errors are visible in the BT result. Furthermore, Rank and HMI seem to cause less distortions at depth discontinuities when observing the chimney. This confirms the findings of Section 3.1 that Rank and HMI perform best. For SGM, the differences are higher (Figure

6g–k). BT and HMI produce the best object borders, while the LoG, Rank, and especially the Mean filter cause distortions at object borders. The effect is similar for GC (Figure 6l–p). Recall that the GC implementation does not include a treatment for occlusions; thus, errors left to object borders should be ignored. We found similar behavior on the other image pairs.

3.4. Runtime

Many applications demand not only accurate disparities, but also fast runtime. Table 1 gives the runtime of the LoG, Rank, and HMI computation (excluding the runtime of the stereo method) for the Teddy image pair. All methods were measured on a Pentium 4 with 2.8 GHz. The Teddy images have a size of 450×375 pixels. A 5×5 kernel was used for the LoG filter and a 15×15 window for the Rank filter. The runtime of 100 ms of HMI includes the time for upscaling and downscaling the images as well as the MI calculations on all hierarchical levels, starting with $\frac{1}{16}$th of the full image size. Additionally, the stereo method needs to be run not only on the full resolution images, but also for downscaled images. For Corr and SGM, the overhead has been measured with about 15% of their runtime at full resolution.

Table 1. Runtime of cost computation on Teddy images.

Method	LoG	Rank	HMI
Overhead	20ms	35ms	$\approx$ 100ms+15%
Implementation	C/MMX	C/MMX	C

The times required for applying the Mean filter or NCC

are not included since these methods were implemented in Java and were not optimized.

4. Conclusion

We compared several different cost functions for stereo matching on radiometrically different images. Each cost was evaluated with three different stereo algorithms: a local correlation method, a semi-global matching method, and a global method using graph cuts. We found that the performance of a matching cost function depends on the stereo method that uses it. On images with simulated and real radiometric differences, the Rank transform appeared to be the best cost for correlation-based methods. In tests with global radiometric changes or noise, hierarchical mutual information performed best for pixel-based global matching methods like SGM and GC. In the presence of local radiometric variations Rank and LoG performed better than HMI for SGM and GC.

A qualitative evaluation of the disparity images from images without radiometric transformations indicated that the filter-based costs (LoG, Rank and Mean) tend to blur object boundaries. This does not affect the results of correlation as the fixed-sized correlation window leads to blurred discontinuities anyway. On the contrary, these filters can actually remedy this problem, in particular the Rank filter, which reduces the weight of outliers near discontinuities. However, the blurring effect is clearly visible for pixel-based matching methods such as SGM and GC. For such methods, the results of BT and HMI appeared best.

None of the matching costs we compared was very successful at handling strong local radiometric changes caused by changing the location of the light sources. It would be nice if the advantages of the different costs could be combined to get a matching cost that is able to handle local radiometric transformations like Rank and LoG while still maintaining sharp depth discontinuities like HMI.

Future work includes testing other matching costs that can handle radiometric differences, e.g., the census transform [23] and the approximation of MI of Zitnick et al. [24].

Acknowledgments

References

[1] D. Bhat, S. Nayar, and A. Gupta. Motion estimation using ordinal measures. In *Proc. CVPR*, 1997.

[2] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *TPAMI*, 20(4):401–406, 1998.

[3] Y. Boykov and V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *TPAMI*, 26(9):1124–1137, 2004.

[4] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *TPAMI*, 23(11):1222–1239, 2001.

[5] M. Z. Brown, D. Burschka, and G. D. Hager. Advances in computational stereo. *TPAMI*, 25(8):993–1008, 2003.

[6] J. Davis, R. Yang, and L. Wang. BRDF invariant stereo using light transport constancy. In *Proc. ICCV*, 2005.

[7] G. Egnal. Mutual information as a stereo correspondence measure. Technical Report MS-CIS-00-20, Comp. and Inf. Science, U. of Pennsylvania, 2000.

[8] H. Hirschmüller. *Stereo Vision Based Mapping and Immediate Virtual Walkthroughs*. PhD thesis, School of Computing, De Montfort University, Leicester, UK, 2003.

[9] H. Hirschmüller. Accurate and efficient stereo processing by semi-global matching and mutual information. In *Proc. CVPR*, volume 2, pages 807–814, 2005.

[10] H. Hirschmüller, P. R. Innocent, and J. M. Garibaldi. Real-time correlation-based stereo vision with reduced border errors. *IJCV*, 47(1/2/3):229–246, 2002.

[11] J. Kim, V. Kolmogorov, and R. Zabih. Visual correspondence using energy minimization and mutual information. In *Proc. ICCV*, 2003.

[12] V. Kolmogorov and R. Zabih. What energy functions can be minimized via graph cuts? *TPAMI*, 26(2):147–159, 2004.

[13] K. Konolige. Small vision systems: Hardware and implementation. In *Proc. ISRR*, pages 203–212, 1997.

[14] Middlebury stereo website. *www.middlebury.edu/stereo*.

[15] H. Moravec. Toward automatic visual obstacle avoidance. In *Proc. Joint Conf. on Artif. Intell.*, pages 584–590, 1977.

[16] R. Sara and R. Bajcsy. On occluding contour artifacts in stereo vision. In *Proc. CVPR*, 1997.

[17] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1/2/3):7–42, 2002.

[18] D. Scharstein and R. Szeliski. High-accuracy stereo depth maps using structured light. In *Proc. CVPR*, volume 1, pages 195–202, 2003.

[19] R. Szeliski, E. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agrawala, M. Tappen, and C. Rother. A comparative study of energy minimization methods for Markov random fields. In *Proc. ECCV*, volume 2, pages 16–29, 2006.

[20] P. Viola and W. M. Wells. Alignment by maximization of mutual information. *IJCV*, 24(2):137–154, 1997.

[21] L. Wang, M. Gong, M. Gong, and R. Yang. How far can we go with local optimization in real-time stereo matching. In *Proc. 3DPVT*, 2006.

[22] K.-J. Yoon and I.-S. Kweon. Adaptive support-weight approach for correspondence search. *TPAMI*, 28(4):650–656, 2006.

[23] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondance. In *Proc. ECCV*, pages 151–158, 1994.

[24] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski. High-quality video view interpolation using a layered representation. In *SIGGRAPH*, 2004.