

A STEREO MATCHING DATA COST ROBUST TO BLURRING

Colin Doutre and Panos Nasiopoulos

The University of British Columbia, Vancouver, Canada

ABSTRACT

Most modern stereo matching algorithms involve solving an optimization problem where the objective function includes a data cost term and a smoothness term. The data cost term measures how well corresponding pixels match between the left and right images. In this paper a new stereo matching data cost is proposed which is robust to variations in blurring between the images caused by camera focus. In our method, each image is blurred once with a large filter. By comparing the original and blurred versions of each image we obtain a range of possible values each pixel could take on for different levels of blurring. Based on this range we construct a blur robust data cost for comparing pixels between two images. Experimental results show our proposed method greatly improves stereo matching accuracy when the left and right images in a stereo pair are focused differently.

Index Terms— blurring robust, stereo matching, focus, data cost, disparity

1. INTRODUCTION

Most stereo matching datasets, such as those from the popular Middlebury stereo page [1], consist of images captured under very carefully controlled conditions using well calibrated cameras. However it is not always possible to capture high quality well calibrated images, for example when cameras are mounted to a robot or simply due to low cost cameras being used. In other cases, such as performing matching on sets of photos found on the internet [2], photos are taken at different times with different cameras, with no attempt to calibrate the cameras. Hence stereo images often suffer from inconsistencies between cameras that can make the images differ in brightness, contrast, resolution, blurring, etc.

A large amount of research has been performed on making stereo matching robust to variations in brightness between images. A through overview and comparison of different methods is presented in [3]. Far less work has been done on making stereo matching robust to variations in blurring between images.

One recent blurring-robust stereo matching method is

based on matching phase values [4]. That method involves comparing quantized Discrete Fourier Transform (DFT) phase values of local regions of the two images. Since phase values are not affected by convolution with a centrally symmetric point spread function, their method is robust to blurring such as out-of-focus blur or Gaussian blur. A problem with this method is that the DFT must be taken on fairly large square windows for reliable matching, which leads to the well known foreground fattening effect [5].

Another stereo matching method that provides robustness to small amounts of motion blur is presented in [6]. In their method, each pixel is classified as either being affected by motion blur or not, and different weightings of the data cost and smoothness cost are used in an optimization stage for the pixels which are estimated as being affected by blurring. This method is only affective if the portion of the image affected by blurring is quite small.

A preprocessing method to correct sharpness variations in stereo images prior to matching is proposed in [7]. That method can correct global blurring, but is not applicable if the blurring varies spatially within an image, for example when part of an image is out of focus.

In this paper we propose a stereo matching data cost that is robust to blurring caused by camera focus. In our method we filter each image once in order to establish a range of values each pixel may take for different levels of blurring. Experimental results show our method significantly improves the quality of stereo matching relative to existing methods.

2. PROPOSED DATA COST

Most modern stereo matching algorithms involve solving an optimization problem considering a data cost term and smoothness cost term [8][9]. The data cost measures how well corresponding pixels match between the left and right images, whereas the smoothness cost penalizes discontinuities in the disparity map in order to favor smoother solutions.

In this work we consider the data cost used in stereo matching. The simplest data cost functions are absolute differences (AD) and squared differences (SD). When comparing a pixel in the left image (i_L) to a pixel in the right image (i_R) the absolute difference data cost is given by:

This work was supported in part by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC).

$$D_{AD} = |i_L - i_R| \quad (1)$$

Data costs such as the absolute difference perform poorly when the left and right images are blurred by different amounts. The images may be blurred differently if the cameras are focused differently (which is particularly a problem if auto-focus is used), or if there is motion blur. Blurring can change the value of a pixel significantly, particularly around object edges and in textured regions of an image. If the blurring is caused by part of the image being out of focus, the blurring can be modeled by a disk-filter with a point spread function of:

$$f(x, y) = \begin{cases} \frac{1}{\pi r^2} & x^2 + y^2 \leq r^2 \\ 0 & x^2 + y^2 > r^2 \end{cases} \quad (2)$$

In equation (2), r is the radius of the blurring, which is dependent of the depth of the image point and the depth at which the camera is focused [10]. A discrete version of the filter in (2) can easily be made by sampling on a grid and integrating over the area of each pixel, which is done in the MATLAB `fspecial()` function.

A straightforward method for making a data cost robust to blurring would be to compare the left image with a series of blurred versions of the right image that have different blurring radii (Fig. 1). If the left pixel closely matches any of the blurred versions of the right image a blur robust data cost should return a low value, since the left and right pixels match for some level of blurring.

While a blur robust data cost could be made based on the idea of Fig. 1, it would not be very practical. Each image would have to be filtered several times, which has a high computational cost, and there is redundancy between the blurred versions of the image that is not being taken into account.

In our method, instead of blurring the image several times, as in Figure 1, we blur the image once with a large filter that corresponds to the maximum level of blurring that could be expected in an application. Define $i(r)$ as the value of a pixel when the image is blurred with a disk filter of radius r . A property of blurring with a disk filter (equation (2)) is that as the radius of the filter is increased, each pixel value in the image varies continuously, i.e. the function $i(r)$ is continuous. In other words, a small change in r results in a small change in the filter $f(x, y)$ and also a small change in each pixel value in the image filtered by $f(x, y)$. Therefore by the intermediate value theorem, for every possible value i_{mid} between $i(0)$ and $i(r_{max})$, there is some level of blurring, r_{mid} , for which $i(r_{mid}) = i_{mid}$, with $0 \leq r_{mid} \leq r_{max}$.

This observation can be applied to design a blur robust data cost. Suppose we are comparing a pixel in the left image with value i_L to a pixel in the right image with value i_R , and that after blurring the right image with a filter of

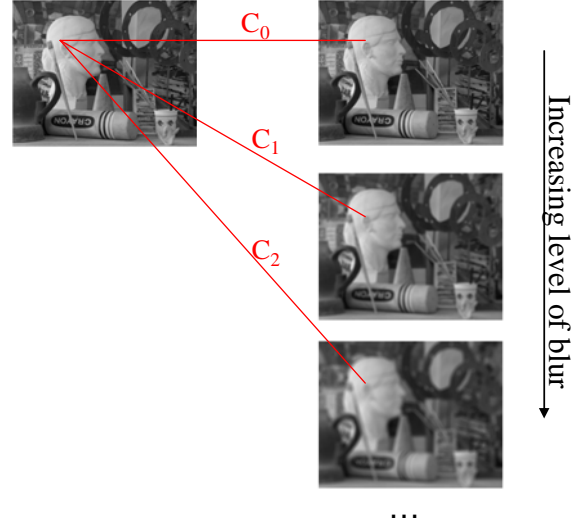


Fig. 1. Comparing a pixel left image with pixels from a series of blurred versions of the right image

radius r_{max} , the right pixel changes in value to i_{RB} . If the value i_L falls between i_R and i_{RB} then we know that applying some level of blurring less than r_{max} to the right image would make the right pixel exactly match the left pixel. Therefore, a blur robust data cost should return a low value (i.e. consider the pixels a match) since they match for some level of blurring. This is the basis for our proposed method.

In our proposed data cost, we consider three cases: i) the images have the same amount of blurring, ii) the right image is blurred more, and iii) the left image is blurred more. In our method, we calculate a separate cost for each of these three cases and then combine them to form a single, blur robust data cost.

The first case is the easiest, where we test how closely the images match assuming they are consistent in blurring. Denoting this case consistent-blurring (CB), we can use a standard cost such as the absolute difference:

$$D_{CB} = |i_L - i_R| \quad (3)$$

In the second case, we test how closely the left image matches a blurred version of the right image, which we denote the right-blurred (RB) case. Using the intermediate value theorem as discussed earlier, our data cost returns zero if i_L falls between the right pixel value (i_R) and the blurred right pixel value (i_{RB}). If i_L falls outside that range, the traditional absolute difference between i_L and i_{RB} is used.

$$D_{RB} = \begin{cases} 0 & \min(i_R, i_{RB}) \leq i_L \leq \max(i_R, i_{RB}) \\ |i_L - i_{RB}| & \text{otherwise} \end{cases} \quad (4)$$

The third case tests how well the right pixel matches a blurred version of the left pixel. We call this the left-blurred (LB) case. It is exactly equivalent to the second case with the role of the left and right pixels reversed; the cost is zero if i_R falls between i_L and i_{LB} .

$$D_{LB} = \begin{cases} 0 & \min(i_L, i_{LB}) \leq i_R \leq \max(i_L, i_{LB}) \\ |i_{LB} - i_R| & \text{otherwise} \end{cases} \quad (5)$$

With the three individual data costs computed, we merge them into a final, blur robust, single data cost. This could be done by simply taking the minimum of the three costs, but that would treat pixels that match without any blurring equally to pixels that match with blurring. Therefore we apply a small cost penalty to the second and third cases, in order to favor pixels that match without blurring, which can be considered a stronger match. Our final data cost is calculated as:

$$D = \min(D_{CB}, D_{LB} + P, D_{RB} + P) \quad (6)$$

where P is the penalty applied to the blurring cases. Experimentally we have found a penalty of $P = 2.5$ works well for 8-bit images.

3. EXPERIMENTAL RESULTS

In order to objectively test our method, we performed experiments using images pairs from the Middlebury stereo page [1], which have ground truth disparities. We introduced blurring differences by synthetically applying out-of-focus blur to each image. In each stereo pair, we

made the left image near-focused (i.e. objects close to the camera appear sharp, object further away are blurred) and the right image far-focused (objects close to the camera are blurred). In each case the image was blurred with a spatially varying disk-filter, where the radius of the filter (r) varies with the depth of the pixel as [10]:

$$r = \frac{f}{2wa} \left| \frac{z(z_F - f)}{z_F(z - f)} - 1 \right| \quad (7)$$

where f is the cameras focal length, w is the width of a pixel, a is the aperture (f-number), z is the depth of the pixel, and the camera is focused at depth z_F . Further details on this focus-blurring model can be found in [10][11]. Examples of two stereo pairs generated with near-focus and far-focus are shown in Figure 2.

Five stereo pairs from the Middlebury page were used for our tests, Art, Baby1, Cones, Dolls and Reindeer. The half resolution versions of the images, available at [1], were used in all tests.

We performed stereo matching on these differently focused stereo pairs using the belief propagation algorithm described in [9]. Three different data costs were compared: absolute difference, which is used in [9] and serves as a reference point; the blur-invariant phase quantization

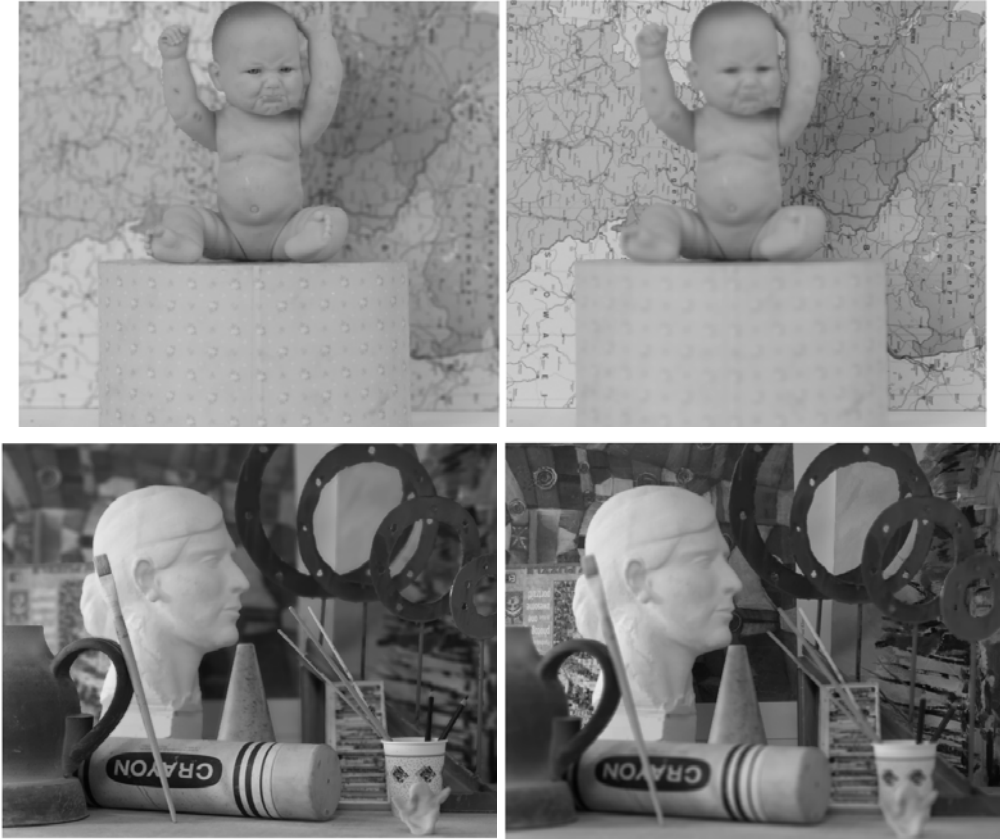


Fig. 2. Examples of synthetically focused stereo image pairs used in our tests. The left images are near focused and the right images are far focused. The images are reproduced at approximately one third resolution here, at full resolution the blurring is more prominent.

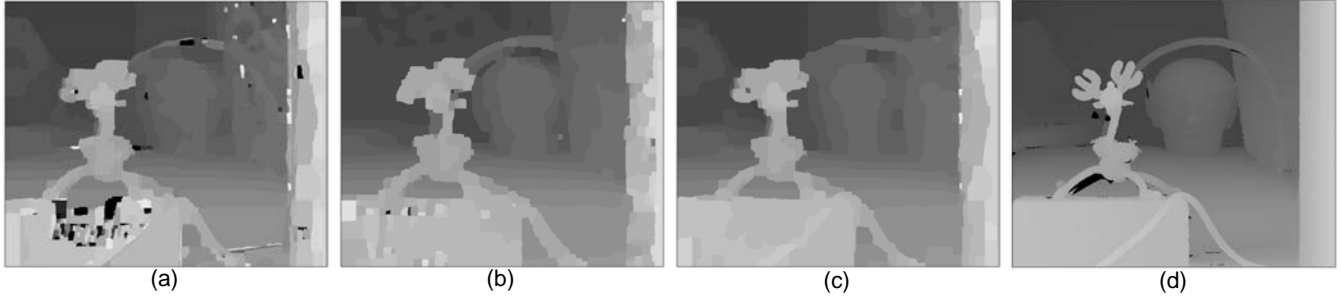


Fig. 3. Disparity maps obtained using different data costs. (a) Absolute differences, (b) Phase quantization [4], (c) Proposed data cost (d) Ground truth

TABLE I: PERCENTAGE OF ERRORS IN DISPARITY MAPS OBTAINED USING DIFFERENT DATA COSTS

Image	Absolute Differences	Phase Quantization	Proposed
Art	25.3	20.5	15.9
Baby1	8.0	4.4	3.2
Cones	24.7	10.7	11.3
Dolls	7.4	5.4	4.0
Reindeer	16.9	12.8	7.7
Average	16.4	10.8	8.4

matching proposed in [4]; and our proposed blur robust data cost. In the tests of our proposed method, we have used a value of $r_{\max} = 4$ pixels for computing our blur robust data cost.

The quality metric we use for comparing disparity maps is the percentage of errors in the un-occluded regions of the image. An error is defined as a point at which the estimated disparity differs from the ground truth disparity by more than one pixel.

In Table I, we show the percentage of errors in the disparity maps obtained using the three different data costs. Both the phase quantization cost and our proposed data cost result in fewer errors than absolute differences, with our proposed method giving the fewest errors on average. Our proposed method gives a 37% to 60% reduction in the number of errors compared to absolute differences. In Figure 3, we show the disparity map for the reindeer image generated using each data cost. Subjectively our proposed method yields a better disparity map compared to both absolute differences and phase quantization.

Our proposed method also has lower complexity than the phase quantization based method [4]. Using the parameters suggested by the authors of [4], their method requires 14 convolutions with *complex* numbers in order to calculate the phase at each point (seven for each image). Our proposed method requires only two convolutions with real numbers (one for each image).

4. CONCLUSIONS

In this paper, we have proposed a blur robust data cost for stereo matching. Our method is based on blurring each

image once with a large filter and comparing the original and blurred versions of each pixel in order to obtain a range of values the pixel may take with different levels of blurring. Experimental results show our proposed method improves stereo matching accuracy compared to both traditional data costs and a state-of-the-art blur robust data cost when the left and right images are focused differently.

5. REFERENCES

- [1] D. Scharstein and R. Szeliski, Middlebury stereo vision page, <http://vision.middlebury.edu/stereo>.
- [2] M. Goesele, N. Snavely, B. Curless, H. Hoppe, S. Seitz, "Multi-view stereo for community photo collections," Proc. International Conference on Computer Vision (ICCV), Oct. 2007.
- [3] H. Hirschmuller, and D. Scharstein, "Evaluation of stereo matching costs on images with radiometric differences," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 31, no. 9, pp.1582-1599, Sept. 2009.
- [4] M. Pedone and J. Heikkila, "Blur and Contrast Invariant Fast Stereo Matching," Proc. Advanced Concepts for Intelligent Vision Systems, pp. 883-890, Oct. 2008.
- [5] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," International Journal of Computer Vision, vol. 47, pp. 7-42, Apr. 2002.
- [6] W. Wang, Y. Wang; L. Huo, Q. Huang, and W. Gao, "Symmetric segment-based stereo matching of motion blurred images with illumination variations," International Conference on Pattern Recognition, Dec. 2008.
- [7] C. Doutre and P. Nasiopoulos, "Correcting Sharpness Variations in Stereo Image Pairs," European Conference on Visual Media Production (CVMP), pp. 45-51, Nov. 2009.
- [8] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," Proc International Conference on Computer Vision, p 508-515, 2001.
- [9] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," International Journal of Computer Vision, vol. 70, no. 1, Oct. 2006.
- [10] M. McGuire, W. Matusik, H. Pfister, J. Hughes, and F. Durand. "Defocus video matting," Proceedings of ACM SIGGRAPH 2005, vol 24, no. 3, pp. 567-576, 2005.
- [11] A.S. Glassner, Principles of Digital Image Synthesis, Morgan-Kaufmann, San Francisco, 1995.