Enhancement of Stereo at Range Discontinuities

Adnan Ansar, Andres Huertas, Larry Matthies and Steve Goldberg
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive, Pasadena, California

ABSTRACT

Stereo is a key component in many autonomous navigation tasks. These applications demand real-time performance and consequently, the state-of-the-art uses local correlation-based algorithms that lend themselves to algorithmic and hardware optimization. These systems perform well in simple terrain or on open ground. However, when discrete objects such as trees are present in the scene, correlation-based approaches exhibit inherent difficulties. Some of these difficulties are introduced during the preprocessing stage that attempts to compensate for photometric variations between the cameras. Other difficulties occur during the correlation stage due to occlusion. As a result, object portions appear enlarged, contracted, or missing, as the range data bleeds between the foreground object and the background. This complicates subsequent obstacle detection, representation and modeling. These problems have been addressed by more sophisticated stereo algorithms based on energy minimization and global optimization schemes. Such complex algorithms, however, are computationally demanding and not amenable to real-time implementation. Our solution uses a better preprocessing method, intelligent use of edge cues, and a variation of the traditional shiftable window approach to enhance the stereo correlation at and near depth discontinuities. There is additional computational overhead involved, but we are able to maintain real-time performance. We present details of our new algorithm and several results in complex natural environments.

Keywords: Real-time stereo, stereo pre-process, edge-assisted stereo

1. INTRODUCTION

Stereo systems have applications in many environments. In this paper we focus on autonomous navigation, robotics applications, and other tasks that require not only fast performance but also increasingly, better range data, in complex environments. Stereo provides a real-time, low cost method to infer 3D structure for use in hazard avoidance, path planning and other decision-making processes. Where such factors are a consideration, it is also a low visibility passive sensor. Key to the real-time aspect of stereo is the use of algorithms based on sum of absolute difference (SAD) correlation or basic variants. In general, SAD [1] is easier to compute and is less sensitive to outliers than other measures such as sum of square difference (SSD) and cross correlation [2,3]. Stereo by SAD correlation has proven a robust and reliable tool in moderately complex environments. However, as more demanding tasks are required of autonomous systems, the need to operate in more complex environments increases. While more elaborate approaches [4,5] often produce better results they are not yet at a stage to be used in real-time systems. Thus, any improvements to existing stereo algorithms for use in navigation are likely to be based on local correlation approaches.

The key advantage of local approaches is speed and suitability for hardware implementation. Global optimization algorithms commonly require 2 to 3 orders of magnitude more time than even the software versions of the local methods [6,7]. Our standard SAD implementation runs at 30 fps on a Pentium M at 1.4 GHz for images of size 320x 240 pixels, optimized for Pentuim 3 instructions. We anticipate that further optimization to Pentium 4 instructions will increase performance to 40 fps. In general, if such an algorithm runs in the order of tenths of seconds in software implementations, it can comfortably reach video rates using DSP and FPGA implementations [8,9]. At the moment, there is no technique for achieving simultaneously the high quality range data obtained from global optimization with the fast run-times of local schemes. All such local techniques must account in some way for photometric variation between cameras in the stereo rig. One method commonly employed with cross correlation is image normalization, in which each image is modified to have local statistics with zero mean and standard deviation equal to one. For SSD/SAD, some form of band-pass filtering is typically used. This may take the form of a Laplacian of Gaussian (LoG) convolution, a difference of Gaussians (DoG) or a difference of averaging boxes (DoB) filters. These amount to a spatial filtering in which texture information is preserved while low frequency background intensity and very high frequency noise are suppressed. In practice, only the high pass component that accounts for the photometric balance is needed. A fundamentally different approach is found in the Rank and Census algorithms [10]. Here the original image is replaced by one that directly encodes local image statistics. In the Rank case, each pixel is replaced by the number of neighboring pixels of lower intensity. In the Census case, each pixel is replaced by a bit string encoding the intensity of all neighboring pixels relative to the central pixel.

Real-time stereo can be improved by modifying the correlator, by modifying the pre-processing step to supply the correlator with better information, or by some combination of the two. Various adaptations of the basic correlation scheme have been proposed. These include shiftable [11], overlapping [12] and adaptive [13] windows. However, any of these techniques will benefit from a better pre-processing of the image. In this paper we propose a technique for improved pre-filtering of imagery for SAD-based stereo. The technique consists of replacing the normal band-pass stage, which introduces an inherent image smoothing, with an adaptive process based on the bilateral filter, first introduced by Tomasi and Manduchi [14]. We show that the results are superior to band-pass filtering with SAD as well as to normalized cross correlation. We also compare to Rank and Census, which are also known to suffer from the same problems at discontinuities as SSD/SAD and normalized cross correlation [9]. Furthermore, in our experience Rank suffers from low information content relative to the other algorithms and performs poorly on fine structures. A fair comparison with Census requires computation on imagery with bit-depth equal to the size of a filter window. Here we use the same size as the correlation window. In Section 2 we discuss how the smoothing effect of the standard SAD pre-filter is partially responsible for the low quality at range discontinuities. In Section 3 we provide the background on the bilateral filter and describe our adaptations for its use in real-time stereo. In Section 4 we describe the use of edge cues to further improve disparities near object boundaries. In Section 5 we provide experimental results with real data. Finally, we draw some conclusions in Section 6.

2. STANDARD SSD/SAD PRE-PROCESSING

Stereo algorithms must compensate for photometric variations between the cameras. The usual approach for SSD/SAD algorithms is to apply a LoG filter, which suppresses high frequency noise (intrinsic Gaussian smoothing) while simultaneously normalizing the intensity information and preserving texture information. This can be well approximated by a DoG [15] in which the original intensity image is replaced by the difference of its convolution with a large and small Gaussian kernel *G*, as follows:

$$I^{'} = I * G(\sigma_{small}) - I * G(\sigma_{large})$$

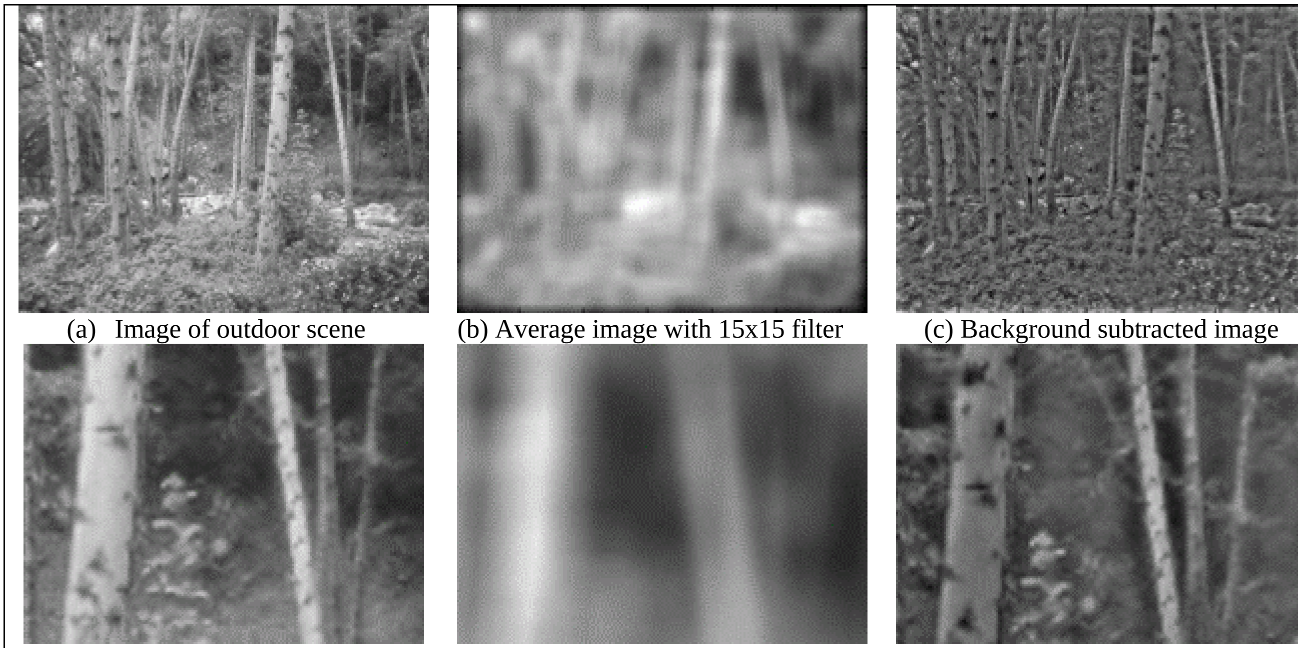
The small Gaussian serves as a low pass filter and the differencing serves as a high pass filter. For imagery of good quality, the noise suppression provided by the low pass filter is generally unnecessary. Thus, we only require a high-pass filter, which can be achieved by background subtraction:

$$I^{'} = I - I * G(\sigma_{large})$$

We have found the difference in stereo quality between background subtraction, an averaging bandpass filter, and convolution with a LoG to be negligible. In the remainder of this paper, we will use background subtraction as the basis for comparison with our new approach.

Regardless of the variant used, any of the above methods introduces a blurring across image discontinuities. The effect is a ringing around foreground objects that results in a weakening of correlation match and a bleeding of range data across the discontinuity. We illustrate this in Fig. 1 using a 15x15 kernel for background subtraction. This is a kernel size we frequently use for real applications and is not intended solely to highlight the ringing phenomenon. Note the ringing or halo effect near the trees. This does not correspond to any real image content and is simply a side effect of the background subtraction. However, it does result in misestimation of disparity near the trees. A pre-processing step that does not blur across range discontinuities therefore is an obvious step towards improved stereo. However, until recently there has been no low-cost mechanism for smoothing in homogeneous image regions while sharply preserving discontinuities. Complex schemes to extract this information would conflict with the real-time requirement. In the next section, we introduce the bilateral filter and show that it solves this problem without incurring high computational costs.

Note, however, that correlation itself introduces an additional error at range discontinuities. We discuss this aspect below in Section 4.



(a) Image of outdoor scene (b) Average image with 15x15 filter (c) Background subtracted image

(d) Detail of image in (a) (e) Detail of image in (b) (f) Detail of image in (c)

Fig. 1. Pre-processing step using averaging. Note the ringing artifact around the trees boundaries.

3. PREPROCESSING USING THE BILATERAL FILTER

The bilateral filter [14,16] computes the weighted average of the pixels within a neighborhood. The weights depend on both the spatial and intensity difference between the central pixel and its neighbors. Formally, the filter takes a signal $f(x)$ and returns:

$$h(x) = \frac{\int_{\Omega} f(\xi)\, c(\xi, x)\, s(f(\xi), f(x))\, d\xi}{\int_{\Omega} c(\xi, x)\, s(f(\xi), f(x))\, d\xi}$$

where Ω is the filter support. The weight functions *c* and *s* are typically Gaussian distributions of the form

$$c(\xi, x) = e^{-\frac{1}{2}\left(\frac{|\xi - x|}{\sigma_d}\right)^2} \quad \text{and} \quad s(f(\xi), f(x)) = e^{-\frac{1}{2}\left(\frac{|f(\xi) - f(x)|}{\sigma_r}\right)^2}.$$

Here $f(x)$ is the intensity at pixel x, σ_d is the standard deviation of the spatial component of the blurring function and σ_r is the standard deviation of the intensity component.

The bilateral filter can be used as an edge-preserving smoother, removing high-frequency components of an image without blurring its edges. We can control the spatial support of the filter, and thus the level of blurring, by varying σ_d. By varying σ_r, we can adapt the sensitivity of the filter to changes in image intensity. In Fig. 2 we show the same image of an outdoor scene as in Fig. 1, but now using a 15x15 bilateral filter with $\sigma_r = 15$; $\sigma_d = 5$. Observe that tree edges are preserved while homogeneous regions are blurred. In the background-subtracted image, texture is apparent without the noticeable ringing of the standard background subtraction. For stereo, the bilateral filter takes the place of Gaussian averaging in the background subtraction step. The resulting process achieves the same normalization effect as background subtraction in homogeneous areas, but minimizes the blurring artifact at discontinuities. We have studied the effect on stereo with a synthetic example which allows us to control ground truth and which is explicitly designed to illustrate the effect of the new pre-filter on edges. In Fig. 3 we show the left image of a stereogram (inset) consisting of uniform random noise. The image is of an 11 pixel wide column in front of a background plane. The background has a disparity of 1 pixel from left to right image and the column has a disparity of 10 pixels. The column is on average brighter than the background. We compute stereo using background subtraction and bilateral filtering, both with 15x15 kernels. In the case of the bilateral filter, we use σ_d = 5 and σ_r is computed by a heuristic described below. For comparison, we also compute stereo using normalized cross correlation and SAD correlation using DoB background subtraction with kernel sizes of 11 and 3. A 7x7 window is used for correlation, and left-right line of sight checking is enabled. Fig. 3 shows the result of averaging computed subpixel disparities over all rows using the four algorithms mentioned. We also show ground truth. Observe that bilateral pre-filtering with SAD is less susceptible to edge effects than the other background subtraction methods. In Section 5 we show that the bilateral approach also preserves the range data density typical of SAD and performs better than cross correlation on homogeneous regions.

(a) Image of outdoor scene (b) Bilateral smoothing (c) Background subtracted image

Fig. 2. Pre-processing step using a bilateral filter with σ_d = 5 and σ_r = 15. Note the preserved tree boundaries in (b)

We now address the crucial issue of runtime. The bilateral filter is not a filter in the traditional sense because the kernel actually depends on the pixel intensities. A fast implementation is challenging because the bilateral filtering process is not separable. However, we have found [16] that approximating with a separable filter is adequate. We apply a separable approximation consisting of a pair of 1-D bilateral filters, one horizontal and one vertical. This approach has been incorporated in our run time systems, and has been tested for several months with good results. In terms of timing, the figures given below correspond processing times for both images of the stereo pair. Different methods have been implemented at several levels of optimization and are not directly comparable, but help give an idea of the performance that can be achieved and/or improved. A C language implementation of the full 2-D filter using a 15x15 window on a 320x240 image takes 250 ms on average. The 1D x 1D variant requires only 46.5 ms. Both tests used un-optimized implementations of the filters. We expect at least a 2-fold speedup with optimization. For comparison, our optimized standard background subtraction method using a 15x15 averaging filter takes only 2.1ms, and optimized LoG filtering takes about 9 ms. With this “separable” version of the bilateral filter, our real-time stereo system requires 74.6 ms (13.4 fps) on 320x240 imagery using a Pentium M at 1.4 GHz. The stereo systems developed by Hirschmueller [12] running on a Pentium III at 1.13GHz for our comparisons below, require the following average times: C-optimized SAD with LoG pre-processing, 186 ms (5.5 fps). C-optimized Rank, 197.6 ms (5.06 fps). C-optimized Census, 530ms (1.88 fps).

Selection of σ_d , the standard deviation of the spatial distribution, is dictated in part by the correlation window size and is largely independent of image content. However, σ_r necessarily depends on the image and we apply a simple heuristic to derive it automatically: For each pixel, we compute local image variance. We then take square root of the mode of this variance over the whole image as a reasonable candidate for σ_r .

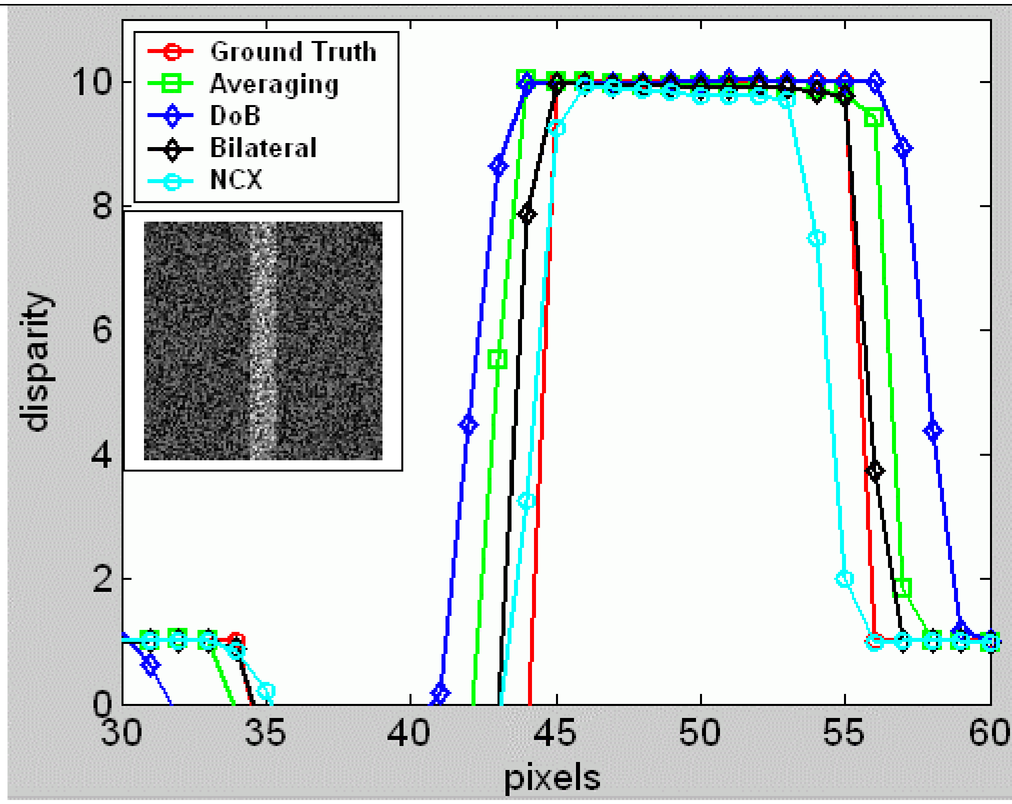


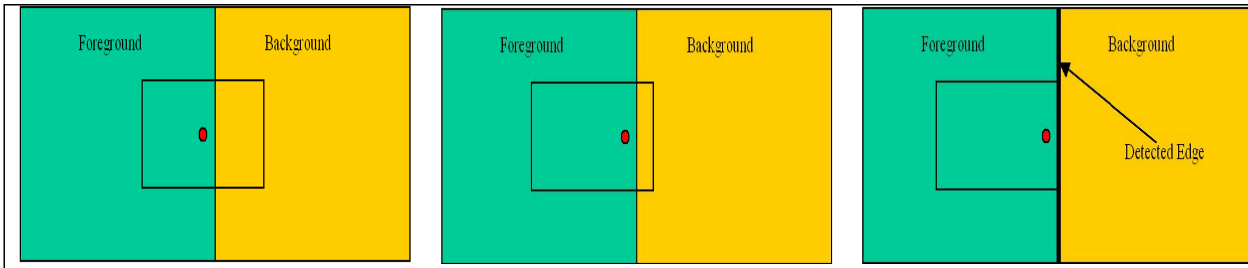
Fig. 3: Stereogram of pillar in front of background plane (inset) and profile of averaged subpixel disparities for various background subtraction methods. Bilateral filtering conforms best to shape of pillar.

4. ASSISTED CORRELATION USING EDGE CUES

Correlation itself introduces an additional error at range discontinuities. The correlation window spans objects at two different depths if they are adjacent in the image. The result is an averaging of correlation scores across boundaries. Larger correlation windows result in greater density in the range image because they provide a larger support for the correlation function. However, this increased density is at the expense of accuracy, especially at range discontinuities. To minimize this effect as much as possible we restrict ourselves to 7x7 correlation windows.

Correlation across discontinuities results in bleeding of range data. We have attempted in the past to address this issue with shiftable windows. However, shifting correlation windows proved only moderately successful at edges. It introduces range artifacts in areas of uniform texture also but has a very easy and fast implementation. Additionally we explored using adaptive windows near edge cues with good results but it is difficult to implement in real-time. We have reached a reasonable compromise by merging the two approaches as follows: First, as with the adaptive algorithm, we shift the correlation window only in the neighborhood of edge cues. Second, instead of using the standard shift, in which the window is allowed to slide arbitrarily within 1/2 the correlation width, we shift the correlation window exactly the number of pixels required to align the edge of the window with an intensity edge cue. Thus, correlation values are taken from a meaningful interior point for either the foreground or background object (See Fig. 4). The current result has all of the benefits of the adaptive approach with none of the drawbacks.

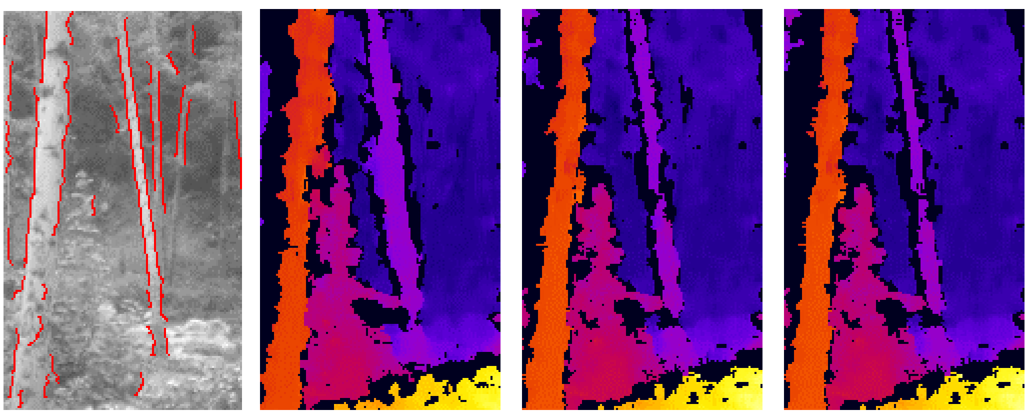
In forests and for many situations involving high vegetation, the primary range discontinuities will be vertical (or near vertical) rather than horizontal. We thus restrict ourselves to near vertical edge components and use these detected edges to modify the computation of correlation scores near edges in the horizontal direction. The availability of edge information is not guaranteed as it is image dependent, but we also restrict the edge boundaries to have reasonable contrast and extent. This helps improve quality mainly for significant objects but also helps reduce the additional time needed for recalculation of correlation scores near edges. The additional effort in extraction of edge cues seems excessive compared to the added benefit. However, we use these same extracted boundaries for other tasks, not described here, such as actual detection and measurement of tree trunk diameters for hazard determination and traversability analyses.



(a) Centered window (b) Normal Shift (c) Edge Assisted Shift

Fig. 4. Shiftable correlation window with edge assistance

Fig. 5 shows detail of stereo output for the images in Fig. 1 and Fig. 2 above using a 7x7 correlation window, with bilateral high pass filtering, and with correlation-assisted edge cues. Note that the tree diameters are preserved much more faithfully in the bilateral version, and even further with the assistance from edge cues.



(a) Image edges (b) Averaging (c) Bilateral (d) Edge-assisted

Fig. 5. Object boundary disparity refinement using prominent edge cues.

5. EXPERIMENTAL RESULTS

Next we illustrate comparisons and results using imagery of real scenes. In particular, we show that the use of bilateral filtering with SAD has advantages over other SAD background subtraction pre-processing filters with SAD, as well as over methods a such as NCX, Rank and Census. We have limited our comparisons to systems that use fixed size correlation windows, and applied the various algorithms using the same common parameters. The stereo rig consists of two color cameras with a baseline of 30 cm. The input consists of 320x240 rectified stereo pairs; correlation windows are 7x7; a maximum disparity search of 64 pixels; left-right consistency check applied with a maximum of one pixel disparity difference. All stereo techniques include a blob-filtering step that removes small isolated areas that may contain wrong disparities. In the experiments below this threshold was set to 300 pixels. The particular parameters used for the various methods are as follows:

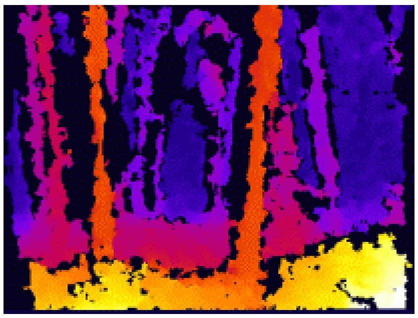
- SAD with LoG preprocessing uses σ =1.
- SAD with DoB pre-processing uses a large kernel of size 15x15 and a small kernel of size 3x3.
- SAD with averaging pre-processing uses a filter size of 15x15.
- SAD with bilateral pre-processing uses σ_d = 5; σ_r = 15; 1-D bilateral filter.
- SAD with edge-assisted correlation uses a maximum of 3-pixel horizontal shift. Edges (Canny contours) are directionally filtered (up to 22.5^o from vertical.)
- Rank and Census measure windows are 7x7. LoG pre-processing applied using σ =1.
- NCX uses zero mean and a standard deviation of 1 computed within the correlation window (C4 metric in [2]). No pre-processing.

Fig. 6, Fig. 7 and Fig. 8 illustrate the disparities computed using these various methods. In Fig. 6 the two trees in the foreground are 7.5 m away and have diameters of 26 cm (left tree) and 39 cm. The thinner tree on the right is 16 m away and has a diameter of 25 cm. The two "adjacent" trees on the left are actually 12 m (left tree) and 9 m away, and their diameters are 27.8 cm 36.7 cm respectively. These correspond to actual hand measurements. In Fig. 7, a winter scene with snow on the ground, the trees are farther away. The closest tree, in the middle, is about 18 m away, estimated by the stereo system. In Fig. 8 the large tree on the left is about 6 m away. The five thin trees lined up from left to right in the middle are approximately 12 m, 19 m, 20 m, 30 m and 33 m away. These ranges are also estimated by the stereo system. These three scenes represent a variety of tree sizes, textures, appearances, and are located over range of distances.

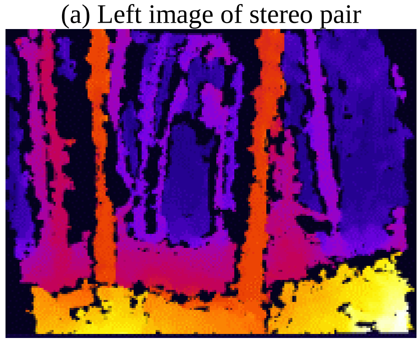
Note in particular the improved results obtained from using the approaches we have proposed. In Fig. 6h, Fig. 7h and Fig. 8h we show the results from SAD using bilateral filter background subtraction. The improvement occurs primarily at the boundaries of objects and accounts for gross errors rather than subpixel errors. The errors reflect pixels that have an inconsistent disparity estimate. The absolute difference of disparities between standard SAD and SAD with bilateral filtering is typically on the order of 10s of pixels disparity. The majority of pixels corrected by bilateral are accounted for in this difference. Notice also that more usable range data is recovered at the extreme side ends of the image. This error in the standard SAD algorithm results from the background subtraction averaging over the usable edge of the rectified image. The improvement of the bilateral filter pre-process over background subtraction is independent of kernel size and we have performed experiments (not included here, see [16]) for kernel sizes of 7x7, 11x11 and 15x15 to demonstrate this. In each case, the bilateral filter produces better stereo at edges and on fine structures. Except for the result showed in Fig. 6 (and earlier in Fig. 1 and Fig.2) the images were acquired from a robot vehicle during an autonomous navigation trial. These images present several scenarios likely to be encountered by a system for which real-time stereo is crucial. The near trees usually present a challenge to conventional SAD stereo. Ground truth is not available for these scenes to actually quantitatively characterize the improvement. However, because of the larger disparity differences involved and the sharper intensity variations between foreground and background, the advantages of the bilateral filter are, nevertheless, readily apparent. Overall these show that bilateral filtering can reproduce the trees accurately while minimizing the loss of texture in uniform regions. Note that the results pictured are typical of the whole sequences from which the current images are taken. See additional summary results for some other frames in Fig. 9 from Ft. Polk, LA, and in Fig. 10 and Fig. 11 for images of scenes from Ft. A. P. Hill, VA.

Bilateral pre-processing can also benefit modified correlators. We illustrate this fact by using shiftable correlation windows with a 3 pixel horizontal shift in conjunction with bilateral filtering. These results are shown in Figures 6j, 7j, and 8j. We extract contours using a Canny edge detector modified to compute high contrast contours with edges having a near-vertical orientation. These are shown in Figures 6i, 7i and 8i. The locations of these edges guide the correlator to recalculate correlation scores near object boundaries. The effect is noticeable along boundaries that are separated by more than the width of the correlation window. There is no effect when the window overlaps more than one contour, as when groups of trees appear close to one another in the image. Note however that in many cases portions of the trunks of thin trees that are not present in the disparity images from other methods, do appear in the results illustrated in Figures 6j, 7j, 8j, and in the additional results shown on the right column of Figures 9 and 10. For comparison the middle column of Figures 9 and 10 show SAD with averaging pre-processing. The edge-assisted SAD shiftable window correlator has not been incorporated into our real-time system yet pending optimization of our contour detector and shiftable correlator. The unoptimized contour detector in our real-time stereo system takes 40.8 ms (on a Pentium M at 1.4 GHz), and is currently used for experiments to detect and measure tree diameters in images of outdoor scenes.

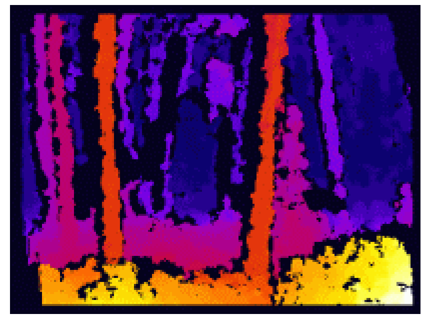
(a) Left image of stereo pair



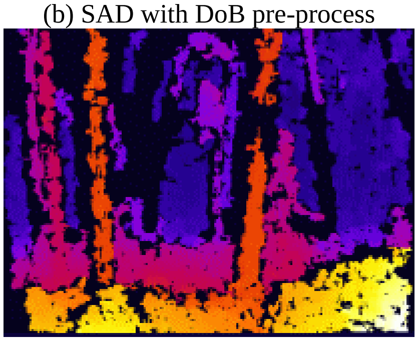
(b) SAD with DoB pre-process



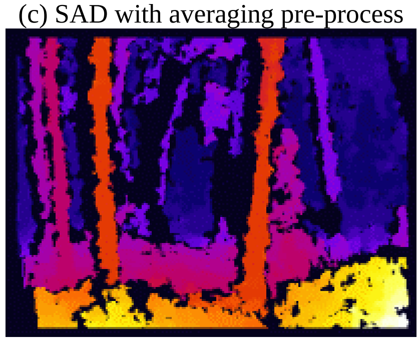
(c) SAD with averaging pre-process



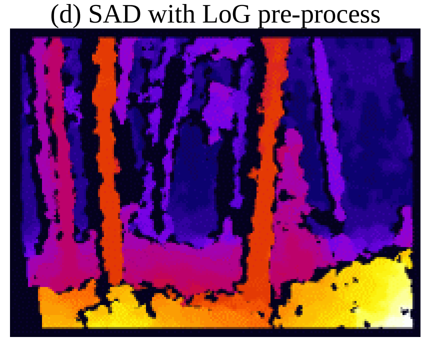
(d) SAD with LoG pre-process



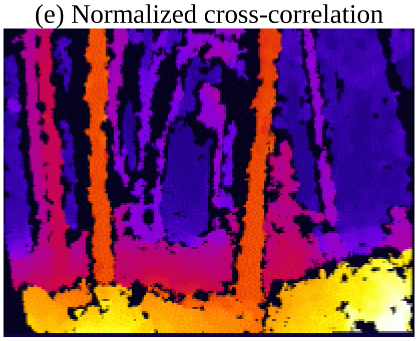
(e) Normalized cross-correlation



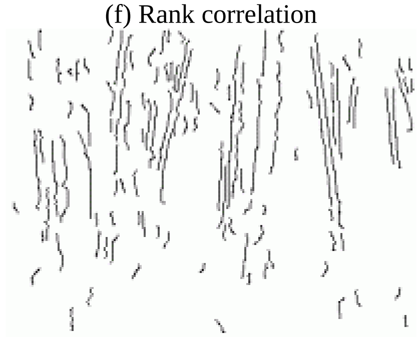
(f) Rank correlation



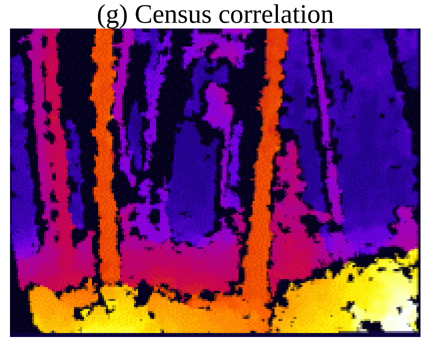
(g) Census correlation



(h) SAD with bilateral pre-process



(i) Edge cues for assisted shift



(j) SAD with bilateral pre-process and edge-assisted correlation

Fig. 6. Disparity maps for several pre-processing options. Note the improved sharpness of the tree boundaries in (h,j).

6. SUMMARY AND CONCLUSIONS

Real-time stereo algorithms typically rely on a simple correlation mechanism applied to imagery processed in some way to account for photometric variations between cameras. Improvements to such algorithms address the correlation step and/or the preprocessing step. As we have shown, those modifications, which target both, are likely to benefit the subsequent steps. We have presented in this paper an improvement to the filtering step employed by most SAD based correlation algorithms which replaces the conventional bandpass filters with background subtraction of a bilaterally filtered image. The result suppresses photometric variation between cameras, much like other bandpass filters, while maintaining much greater fidelity of data at discontinuities in intensity, hence in most discontinuities in range. This produces better stereo at these range discontinuities. We have also shown that our solution has some advantages over the alternative cross correlation approaches in that it has less loss of range data in uniform regions. Furthermore, our method does not sacrifice the real-time performance that drives current correlation based algorithms. We are currently working to compensate for the loss of data in the background seen in some experiments. We believe that a hybrid approach, using bilateral subtraction only in certain regions dictated by image statistics and normal background subtraction elsewhere, will solve this problem. We have also shown that the correlation scores can be improved near range discontinuities by using edge cues. We expect to make optimization improvements that enable us to compute and use edge cues, both to improve correlation, and for other tasks where object edges become useful, such as in urban environments.

ACKNOWLEDGEMENTS

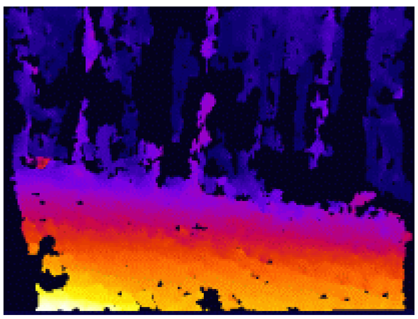
The work presented was sponsored in part by the PerceptOR program of DARPA of the U.S. Department of Defense, under a sub-contract from SAIC, Denver, and in part by the U.S. Army Robotics Collaborative Technology Alliance. Arturo Rankin and Paolo Bellutta of JPL carried out JPL stereo imagery data collections under the PerceptOR program. Optimized implementations of SAD with LoG pre-filter, and Rank and Census stereo were provided by Heiko Hirschmueller of the Centre for Computational Intelligence at DeMonfort University in England.

REFERENCES

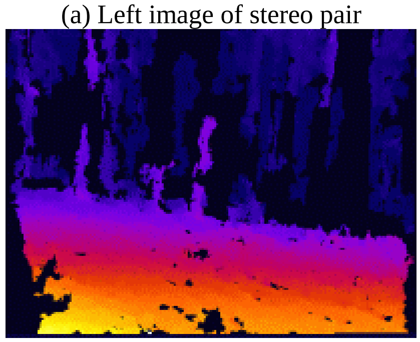
1. L. Matthies, “Stereo vision for planetary rovers: Stochastic modeling to near real-time implementation,” *International Journal of Computer Vision*, vol. 8, pp. 71–91, 1992.
2. O. Faugeras, B. Hotz, H. Mathieu, T. Vieville, Z. Zhang, P. Fua, E. Theron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, “Real time correlation based stereo: algorithm implementations and applications,” Tech. Rep. 2013, INRIA, 1993.
3. K. Muhlmann, D. Maier, J. Hesser, and R. Manner, “Calculating dense disparity maps from color stereo images, an efficient implementation,” *IJCV*, vol. 47, pp. 79–88, April 2002.
4. V. Kolmogorov and R. Zabih, “Computing visual correspondence with occlusions via graph cuts,” in *Proc. IEEE Intl. Conf. Computer Vision*, (Vancouver, Canada), pp. 508–515, 2001.
5. M. Agrawal and L. Davis, “Window-based, discontinuity preserving stereo,” IEEE Computer Vision and Pattern Recognition Conference, Madison, Wisconsin, June 2003.
6. Y. Ohta and T. Kanade, “Stereo by intra- and interscanline search using dynamic programming,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 139–154, 1985.
7. D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *International Journal of Computer Vision*, vol. 47, no. 1/2/3, pp. 7–42, 2002.
8. K. Konolige, “Small vision systems: Hardware and implementation,” in *Proc. Intl. Symposium of Robotics Research*, (Hayama, Japan), 1997.
9. J. Woodfill and B. V. Herzen, “Real-time stereo vision on the PARTS reconfigurable computer,” in *Proc. IEEE Symp. FPGAs for Custom Computing Machines*, (Napa, CA), pp. 242–250, 1997.
10. R. Zabih and J.Woodfill, “Non-parametric local transforms for computing visual correspondence,” in *Proc. IEEE European Conf. Computer Vision*, (Stockholm, Sweden), pp. 150–158, 1994.
11. A. Bobick and S. Intille, “Large occlusion stereo,” *International Journal of Computer Vision*, vol. 33, no. 3, pp. 181–200, 1999.
12. H. Hirschmueller, “Improvements in real-time correlation-based stereo vision,” in *IEEE Conf. Computer Vision and Pattern Recognition*, (IEEE Conf. Computer Vision and Pattern Recognition), pp. 141–148, 2001.
13. M. Okutomi and T. Kanade, “A locally adaptive window for signal matching,” *International Journal of Computer Vision*, vol. 7, no. 2, pp. 143–162, 1992.
14. C. Tomasi and R. Manduchi, “Bilateral filtering for gray and color images,” in *Proc. IEEE Intl. Conf. Computer Vision*, pp. 836–846, 1998.

15. D. Marr, *Vision*. New York: W.H. Freeman and Company, 1982.
16. A. Ansar, A. Huertas and L. Matthies, "Improvements to Real-Time Stereo Processing," New Technology Report No. 40813, Jet Propulsion Laboratory, Pasadena, CA, Nov. 25, 2003.

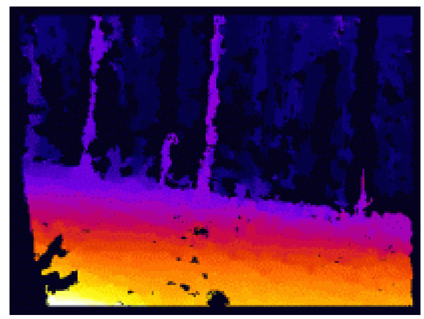
(a) Left image of stereo pair



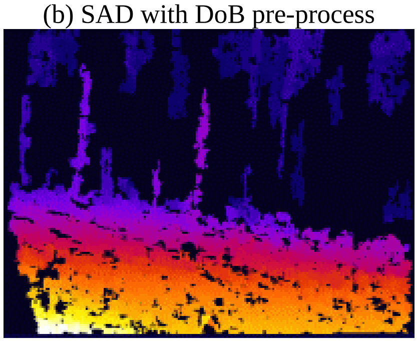
(b) SAD with DoB pre-process



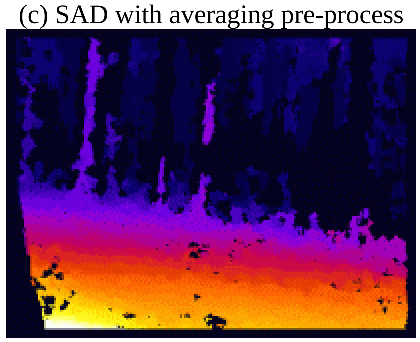
(c) SAD with averaging pre-process



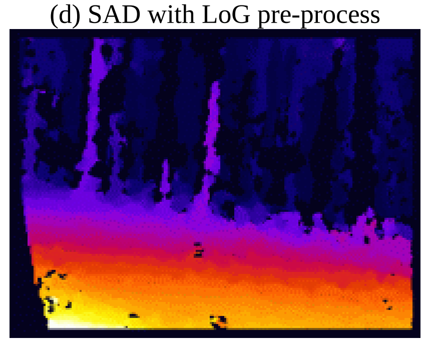
(d) SAD with LoG pre-process



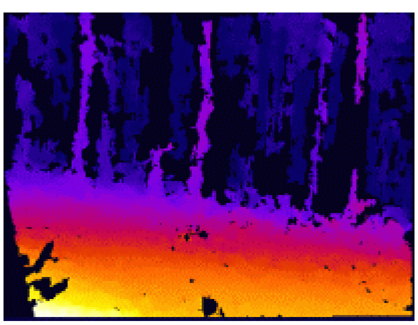
(e) Normalized cross-correlation



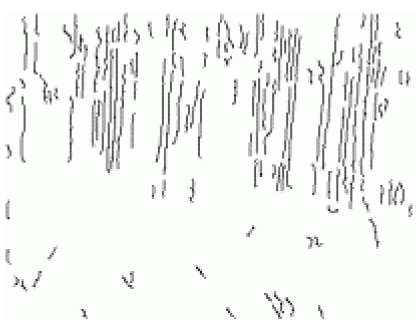
(f) Rank correlation



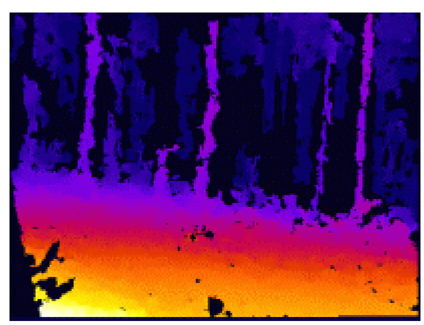
(g) Census correlation



(h) SAD with bilateral pre-process



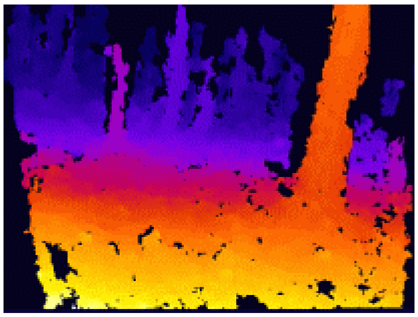
(i) Edge cues for assisted shift



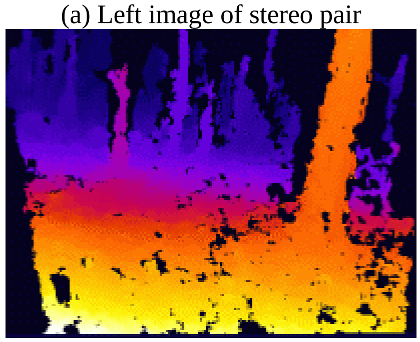
(j) SAD with bilateral pre-process and edge assisted correlation

Fig. 7. Disparity maps for several pre-processing options. Note the improved sharpness of the tree boundaries in (h,j).

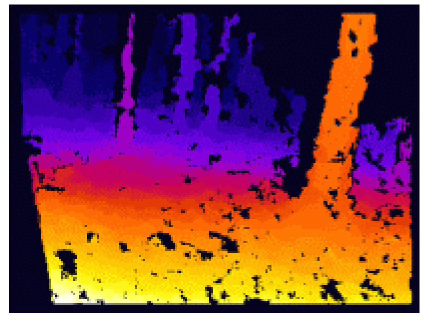
(a) Left image of stereo pair



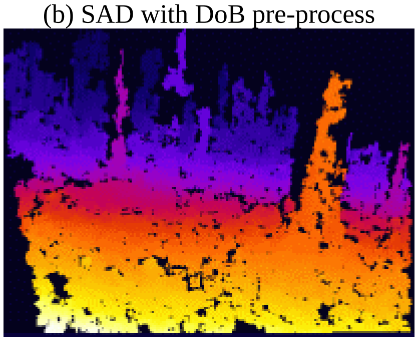
(b) SAD with DoB pre-process



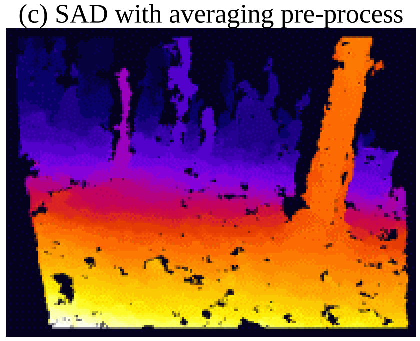
(c) SAD with averaging pre-process



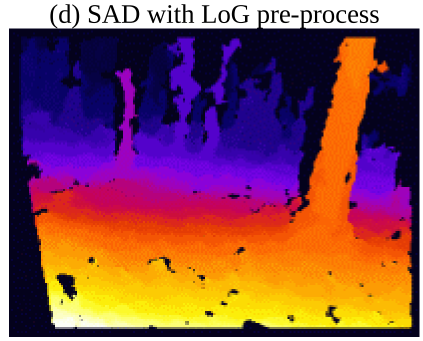
(d) SAD with LoG pre-process



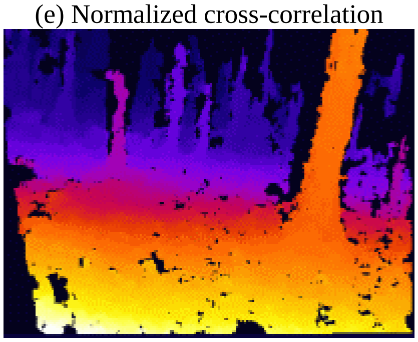
(e) Normalized cross-correlation



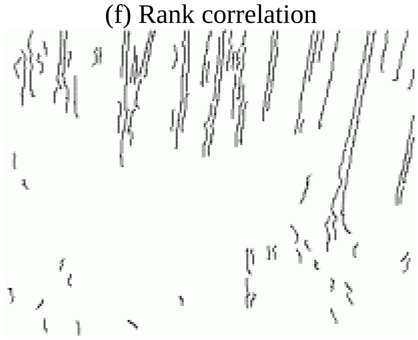
(f) Rank correlation



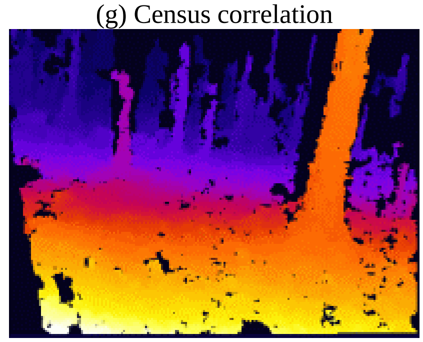
(g) Census correlation



(h) SAD with bilateral pre-process



(i) Edge cues for shift assist.



(j) SAD with bilateral pre-process and edge-assisted correlation.

Fig. 8. Disparity maps for several pre-processing options. Note the improved sharpness of the tree boundaries in (h,j).

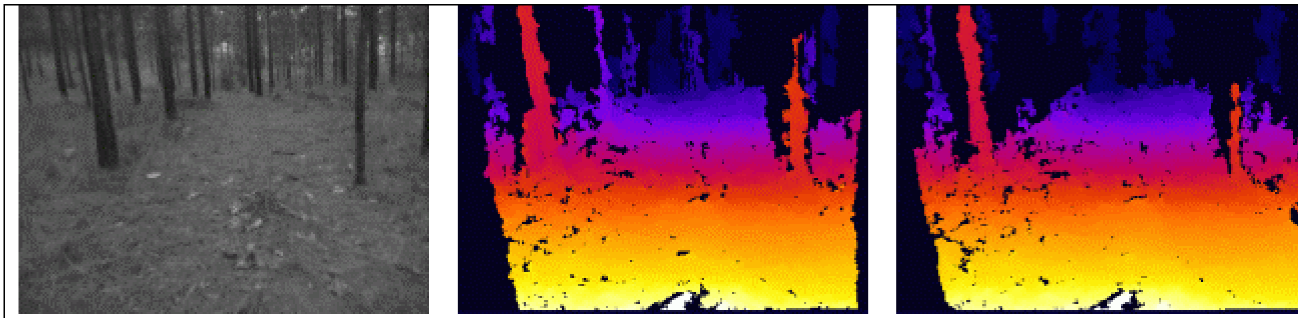


Fig. 9. Ft. Polk scene. Middle: averaging pre-process. Right: bilateral pre-process and edge-assisted correlation.

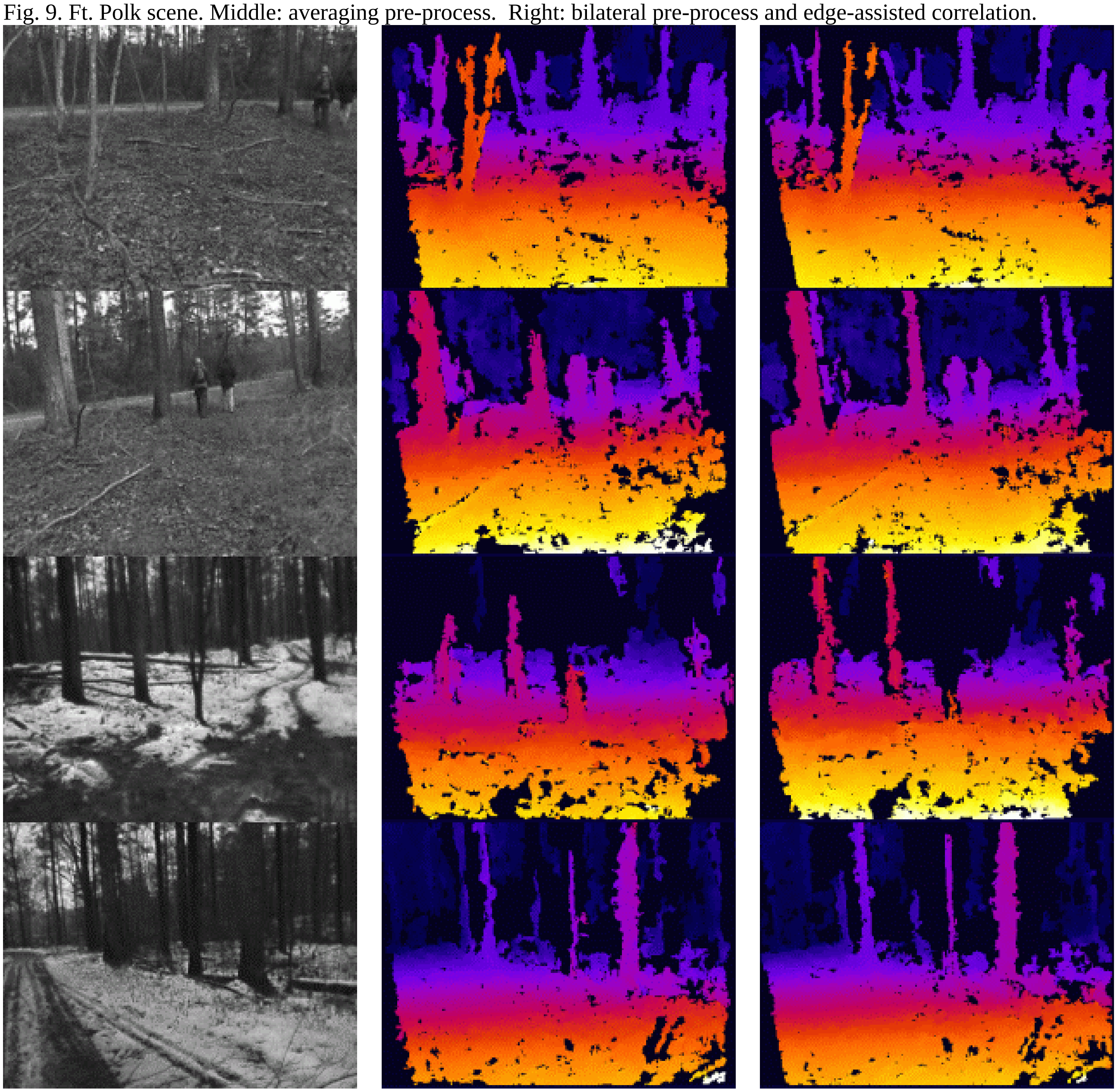


Fig. 10. Ft. A.P. Hill scenes. Middle: averaging pre-process. Right: bilateral pre-process and edge-assisted correlation.