

Fast disparity estimation algorithm using the property of stereo matching

Jiyong Park
 Wonjae Lee, MEMBER SPIE
 Jaeseok Kim
 Yonsei University
 Department of Electrical and Electronic
 Engineering
 Seoul, Korea

Abstract. A fast disparity estimation algorithm based on a pixel-based approach for intermediate view reconstruction (IVR) is proposed. The proposed algorithm uses the matching property of stereo imaging: in stereo images, the disparity vector of a pixel is located between those of previous and subsequent pixels. According to the experiments, the proposed algorithm can reduce the computational complexity by about 8 times and achieves competitive or slightly better peak signal-to-noise ratio (PSNR) in comparison with the pixel-based full search approach. The algorithm is proven to be effective in the implementation of the real-time disparity estimation processor for arbitrary 720×480 stereo images. © 2005 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1925049]

Subject terms: disparity estimation; stereo matching; IVR.

Paper 050045LR received Jan. 19, 2005; revised manuscript received Mar. 30, 2005; accepted for publication Apr. 4, 2005; appeared online Apr. 11, 2005; published online Jun. 10, 2005.

Recently, 3-D imaging systems using multiview images that allow viewers to take any viewing posture by providing full parallax images have been gaining popularity. IVR provides a solution for the problem presented by large data involved in displaying natural 3-D images.¹ The disparity estimation module is a key processor for IVR. A pixel-based approach is better than a block-based approach in terms of quality, but it demands a larger computational complexity because the disparities of all pixels have to be estimated.² For that reason, a fast algorithm with good performance is required.

Fast algorithms^{3,4} that are based on the block-based approach are for stereoscopic image coding and can process only stereo images that are taken from parallel camera. Therefore, they aren't suitable for IVR applications, which demand high-quality images. Fast algorithms⁵ are for estimating dense disparities. However, the algorithm in Ref. 5 isn't useful in implementing a real-time disparity estimation processor for 720×480 stereo images because of insufficient reduction in the computational complexity.

In this letter, we propose a new fast algorithm that is based on the pixel-based approach and reduces computational complexity in comparison with the conventional pixel-based full search approach. The algorithm can process stereo images taken from parallel and convergent cam-

eras in implementing a real-time disparity estimation processor.

Figure 1 shows parallel and convergent camera geometries.⁶ For parallel cameras, the respective locations of image points X_{cl} and X_{cr} on the left and right imaging sensor can be given by

$$X_{cl} = \frac{f(t+2X_0)}{2Z_0} - h, \quad X_{cr} = -\frac{f(t-2X_0)}{2Z_0} + h \quad (1)$$

where f is the lens focal length, t is the camera separation, h is the sensor axial offset.⁶

For convergent cameras, imaged points on sensor X_{cl} and X_{cr} are

$$X_{cl} = f \times \tan(\alpha - \varphi), \quad X_{cr} = -f \times \tan(\beta - \varphi) \quad (2)$$

where $\alpha = \arctan(t+2X_0)/2Z_0$ and $\beta = \arctan(t-2X_0)/2Z_0$.⁶

If $X_0 > X_1 > X_2$ and $Z_0 < C$ are assumed in the parallel and convergent camera geometries as shown in Fig. 1, the conditions of $X_{0cl} > X_{1cl} > X_{2cl}$ and $X_{0cr} > X_{1cr} > X_{2cr}$ are met. Also, if $Z_0 > C$ is assumed in the convergent cameras, the conditions of $X_{0cl} < X_{1cl} < X_{2cl}$ and $X_{0cr} < X_{1cr} < X_{2cr}$ are met. We realize the important property of stereo matching from the above example: in stereo images, the disparity vector of a pixel is located between those of previous and subsequent pixels.

The search range can be predicted by using the property of the referred stereo matching. Figure 2 shows an example. First, the disparities of the 0'th and λ 'th pixels are estimated by Eq. (3). Then the disparity of $(\lambda/2)$ 'th pixel is estimated by Eq. (4). Since the predicted search range (PSR) is usually smaller than the search range, computational complexity is reduced.

$$d^f(i) = \arg \min_r [U\{I_{cur}(i), I_{ref}(i+r)\}] \quad (3)$$

$$d^p(i) = \arg \min_{\bar{r}} [U\{I_{cur}(i), I_{ref}(i+\bar{r})\}] \quad (4)$$

where d^f is the disparity estimated using the full search approach, d^p is the disparity estimated using the predictive search approach, I_{cur} is the intensity of current image, I_{ref} is the intensity of reference image, r is the search range, \bar{r} is PSR, and U is the cost function, such as the sum of absolute difference.

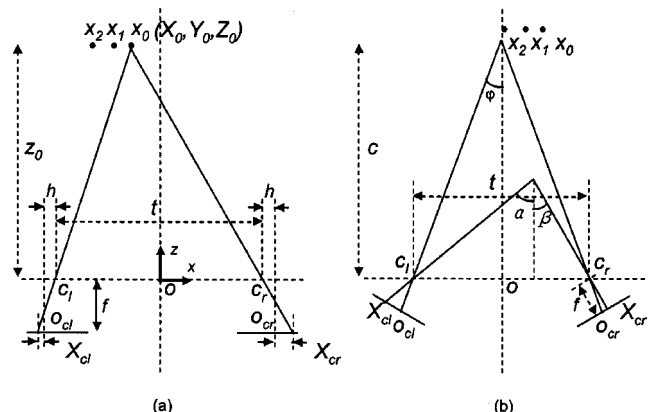


Fig. 1 Camera geometries: (a) parallel camera, (b) convergent camera.

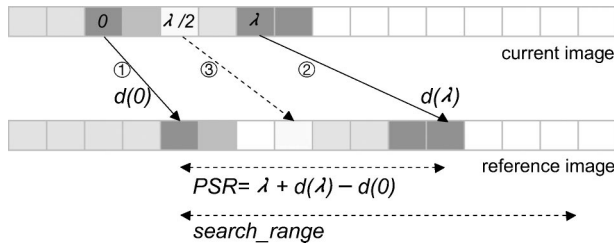


Fig. 2 Predictive search approach.

The procedure of the predictive search approach of the proposed algorithm, which is based on Eqs. (3) and (4), is as follows. First, x , y , and n are initialized to 0, and λ is determined. Second, the disparities of the pixels of the λ space are estimated by using the full search approach until x is smaller than the image size. Third, if x is larger than the image size, x is changed to $\lambda/2^{n+1}$. Fourth, the disparities of the pixels of the $\lambda/2^n$ space are estimated by using the predictive search approach until x is smaller than the image size. Then if PSR is smaller than 0, PSR is changed with the full search range to prevent wrong disparity estimation. Fifth, if x is larger than the image size, n is increased by 1. Sixth, we repeat steps 3 to 5 until the disparity estimation of all pixels of a line is finished. Seventh, if y is smaller than the image height, y is increased by 1 and both x and n are initialized to 0. Eighth, we repeat steps 2 to 7 until y is equal to the image height.

The proposed algorithm predicts the search range but it almost doesn't cause the propagation and the prediction errors because the disparities for the pixels of the λ space are estimated by using the full search approach, and the PSR that is smaller than 0 is changed with the full search range. As a result, the proposed algorithm is as exact as the full search approach.

Both the full search approach and the proposed algorithm have been tested for the 450×375 "Teddy," 434×380 "Sawtooth," and 434×383 "Venus" stereo images,⁷ which are taken from a parallel camera, and for the 720×480 "Puppy" stereo images recorded by the Realistic Broadcasting Research Team of the Electronics and Telecommunications Research Institute, Daejeon, Korea, which are taken from a convergent camera. The full search range for disparity estimation is ± 32 pixels. The window size of 16 by 16 has been used. In view of optimal trade-off between quality and computational complexity, λ of 16 is chosen.

Figure 3 compares the proposed algorithm with the full search approach in views of the computational complexity, the PSNR, and the reconstructed images. The reduction of computational complexity is evaluated by dividing the number of searched windows in the full search approach to the number of searched windows in the proposed algorithm. The PSNR is computed by the reconstructed image and the original image.

In Fig. 3, the proposed algorithm achieves competitive or slightly better PSNR and reduces the computational complexity by about 8 times in comparison with the full search approach. Therefore, the proposed algorithm is useful in implementing the real-time disparity estimation processor, which achieves over 30 frames/s for arbitrary



	PSNR		The reduction of computational complexity
	The full search approach	The proposed algorithm	
teddy 3	32.2	32.2	7.5
teddy 4	33.1	33.0	7.5
puppy 1	26.95	26.75	7.42
puppy 2	26.96	26.78	7.42
sawtooth1	32.08	32.05	7.69
sawtooth2	32.65	32.65	7.67
venus1	33.79	33.79	7.75
venus2	35.35	35.37	7.76

Fig. 3 Comparison of the full search approach with the proposed algorithm: (a) IVR image using the full search approach, (b) IVR image using the proposed algorithm, (c) PSNR vs the reduction of computational complexity.

720×480 stereo images. The number of frames can be given by $(\text{maximum frequency} \times \text{reduced computational complexity}) / (\text{image size} \times \text{search range}) = [(\text{over } (90 \times 10^6) \times 7.4) / (720 \times 480 \times 64)] \geq 30.11$ frames/s.

We have proposed a predictive search algorithm based on a pixel-based approach for fast and efficient disparity estimation. By using the matching property of stereo imaging, we reduce the computational complexity. The experimental results show that the proposed algorithm is about 8 times faster than the full search approach while achieving competitive or slightly better PSNR. The algorithm is proven to be effective in the implementation of the real-time disparity estimation processor for arbitrary 720×480 stereo images. Hence, the algorithm can be applied to stereo image applications that require real-time and high-performance solutions, such as 3-D TV.

Acknowledgments

This research was supported by University IT Research Center Project and partly by the Digital Media R&D Center, Samsung Electronic Co., Ltd.

References

1. A.-R. Mansouri and J. Konrad, "Bayesian winner-take-all reconstruction of intermediate views from stereoscopic images," *IEEE Trans. Image Process.* **9**, 1710–1722 (2000).
2. D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *Int. J. Comput. Vis.* **47**(1), 7–42 (2002).
3. W. H. Kim and S. W. Ra, "Fast disparity estimation using geometric properties and selective sample decimation for stereoscopic image coding," *IEEE Trans. Consum. Electron.* **45**, 203–209 (1999).
4. J. W. Bae, H. C. Park, E. S. Kim, and J. S. Yoo, "Efficient disparity estimation algorithm based on spatial correlation," *Opt. Eng.* **42**, 176–181 (2003).
5. O. S. Uddin, T. T. Son, and S. Mita, "Fast implementation of window-based methods for stereo correspondence," in *Applications of Digital Image Processing, Proc. SPIE* **5203**, 590–598 (2003).
6. A. Woods, T. Docherty, and R. Koc, "Image distortions in stereoscopic video system," in *Stereoscopic Display and Applications IV, Proc. SPIE* **1915**, 36–47 (1993).
7. Middlebury College Stereo Vision Research Page, <http://www.middlebury.edu/stereo/>.