ERROR MODELING OF DEPTH MEASUREMENT USING FIR STEREO CAMERA SYSTEMS

Jun Ho Oh[1], Jungsik Park[1], Sang Hwa Lee[2], Boo Hwan Lee[3], and Jong-Il Park[1,*]

[1] Dept. of Computer Eng. Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, S. Korea
[2] INMC, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, S. Korea
[3] Agency for Defense Development, 488 Beongil, Bugyeseong-daero, Yuseong-gu, Daejeon 305-152, S. Korea
{jhoh, nangsik}@mr.hanyang.ac.kr, lsh529@snu.ac.kr, bhlee@add.re.kr, jipark@hanyang.ac.kr

ABSTRACT

This paper deals with an error model and its evaluation in estimating the 3-D depth information using stereo camera system. This paper defines an error model caused by stereo matching, and evaluates the error bound empirically in estimating the depth information. The proposed error model is that the disparity values in the pixel domain have ±0.5 pixel errors maximally even though the disparity values are correctly estimated. Thus, the depth information also has the same amount of errors with respect to the focal length and baseline of stereo cameras. To evaluate and apply the proposed model to the military applications, we use uncooled FIR cameras to set the stereo camera system and consider the small high-temperature objects as the targets. We first detect the high-temperature moving object in the stereoscopic IR images, and estimate the disparity field in the pixel domain. Then, we calculate the physical depth between the cameras and moving object by using triangulation of disparity, baseline and focal length. According to the experiments with various distances and baselines, we evaluate the error bound in the depth estimation. This error bound and stochastic parameters are useful for estimating the 3-D position of moving objects probabilistically.

KEYWORDS

FIR camera, stereo camera, stereo matching, depth estimation, small target.

*Corresponding author.
This work was supported by the Research fund of Survivability Technology Defense Research Center of Agency for Defense Development of Korea (No. UD090090GD)

1 INTRODUCTION

Recently, stereo camera systems are introduced to estimate depth information in various applications. Stereo matching is to estimate the disparity map between images taken at the slightly difference viewpoints. When we know the intrinsic parameters such as camera focal length, baseline (distance of two cameras), and pixel pitch (distance of consecutive pixels), we can calculate the depth or distance of object from camera using the estimated disparity values [1]. Fig. 1 shows the stereo camera system based on parallel camera setup. The epipolar geometry is the horizontal line, thus the disparity is the difference between projection positions in the pixel domain. The depth (z) or distance between camera and 3-D object point is calculated as

$$z = \frac{b \cdot f}{d \cdot p} \qquad d = D_L + D_R \tag{1}$$

where f is the focal length, b is the baseline, p is the pixel pitch, and d is the disparity in the pixel domain. Since the disparity is usually estimated in the pixel domain, we have to convert the pixel dimension into physical dimension in mm by multiplying the pixel pitch value of the imaging sensor.

ISBN: 978-0-9853483-3-5 ©2013 SDIWC

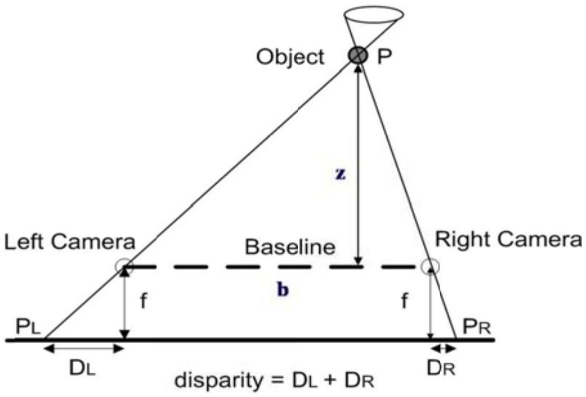


Figure 1. Stereo camera system and depth calculation. D_L and D_R are the projection pixel position of 3D point P. And the difference between two pixel positions is the disparity. This figure is based on the parallel camera setup.

We focus on the military applications for depth estimation of high-temperature moving objects. We exploit the IR camera system since the important objects are usually high-temperature and fast targets. Furthermore, IR camera detects the invisible IR radiation from the objects, thus it is not exposed to the opposite sides [2]. On the contrary, since the radar system has to send radar signals to the targets, it is undesirably detected by the opposite sides. Some military applications using IR camera systems have been studied. These researches are divided into two groups, one is to detect the targets from cluttered scene [3-5], and the other is to estimate the depth of targets [6, 7].
This paper deals with the depth estimation using IR stereo cameras, and proposes an error model in the depth information. The depth error model results from disparity estimation in the pixel domain even though the result of stereo matching is correct. We construct far IR (FIR) stereo camera system for near range environment, and estimate the depth information. Using the ground truths of depth information for indoor and outdoor environments, this paper evaluates the depth errors and bounds.
The rest of paper is organized as follows. In Section 2, we propose an error model in estimating depth information from stereo matching. We will show experimental results and evaluation of error model in Section 3. Finally, we conclude this paper in Section 4.

2 ERROR MODEL

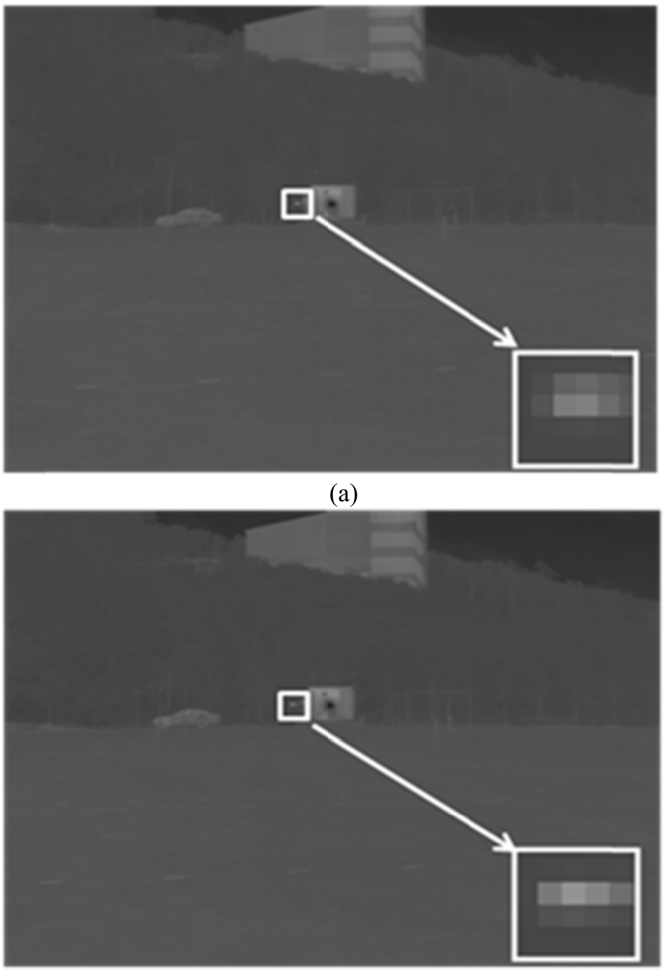
In the stereo matching to estimate disparity values in the pixel, domain, there are some reasons to cause estimation errors. We concentrate on the pixel domain in estimating disparity values. Since the disparity values are assigned to a pixel position, there is an error from -0.5 to +0.5 pixels inevitably even though the disparity values are correctly estimated. If an object point is perfectly located on the pixel grid both in the left and right images, then the correct disparity value is an integer number and has no error at all. In the general cases, however, the object points are not exactly located on the pixel grid. An object point is blurred or overlapped over the multiple pixels. Thus, when we assign the disparity values to a pixel grid, there are maximum errors from -0.5 to +0.5 pixels. These errors cause the uncertainties of depth estimation. We assume that the pixel errors in disparity values are uniform distributed from -0.5 to +0.5. Thus, the depth information also has the same amount of error bound with respect to the focal length and baseline. In the next section, we show the experimental results on the error model using IF stereo camera.

3 EXPERIMENTS AND EVALUATION

We first setup the FIR stereo camera system with parallel camera configuration. The image size is 384x288, the focal length f is 25mm, view angle is 30° x 23°, the pixel pitch p is 0.035mm, and the base line is 300mm on the outdoor experimental. We have experiments on the indoor and outdoor environments respectively.

3.1 STEREO MATCHING

We detect the small object regions in the IR stereo images using temperature threshold and background subtraction. Then, we calculate the mean position of object region. The temperature of object region is considered as the probability, thus, the mean position is the highest temperature position of the object. This enables us to match the object regions between stereo images at a single point. Fig. 2 shows an example why the mean position is necessary. As shown in Fig. 3, the same object in the stereo images looks different in the pixel domain, so it is not easy to match the corresponding object. We localize the object region at a point by the mean position. Thus, it is possible to eliminate the errors due to blurring and ambiguous matching points. This localization of object region eliminates the errors in the disparity estimation except the pixel structure.



(a)

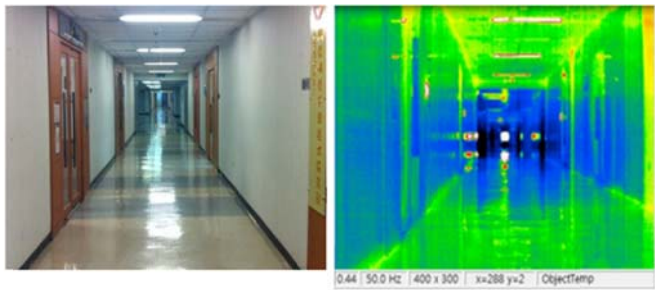
(b)

Figure 2. Detected object regions in stereo images. The detailed patterns in the pixel domain are different despite the regions are the same object. (a)left image, (b)right image.

3.2 INDOOR EXPERIMENTAL RESULTS

Stereo camera system has been built through 3 states of camera calibration, stereo rectification and stereo matching. Camera calibration by Zang method searches for camera internal matrix, distortion factor, rotation vector and moving vector [1].

We set a high-temperature iron panel at 40m, 60m and 80m distance respectively, and estimated the depth using stereo matching. Also, we changed the baseline 0.5 m and 1.0m respectively. Fig. 3 shows the indoor scene of IR camera and visual one.



(a) (b)

Figure 3. Indoor scene for depth estimation. (a) visual camera image, (b) IR camera image.

Fig. 4 shows a captured image of detected object regions, depth estimation using 3 IR cameras. The differences between the mean positions in the pixel domain are calculated as the disparity values.

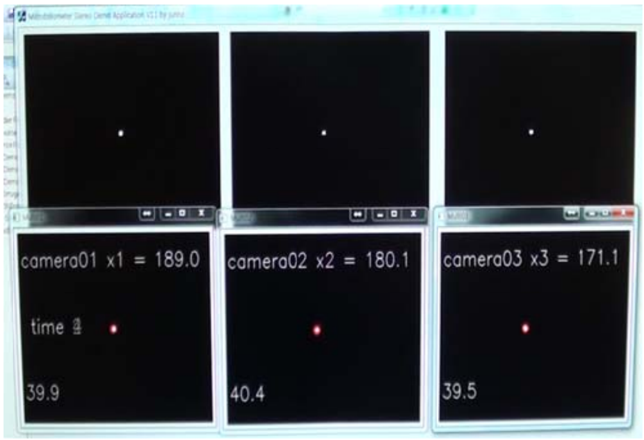
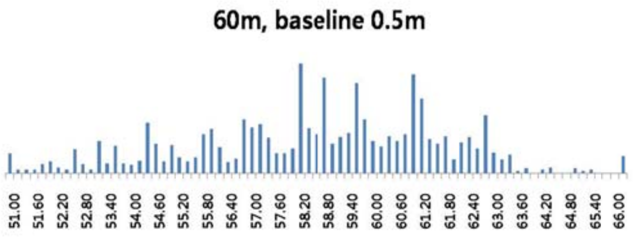


Figure 4. Object detection and stereo matching. The mean positions of object regions are shown on the upper area in the images, and the estimated depth information is shown in the lower area.

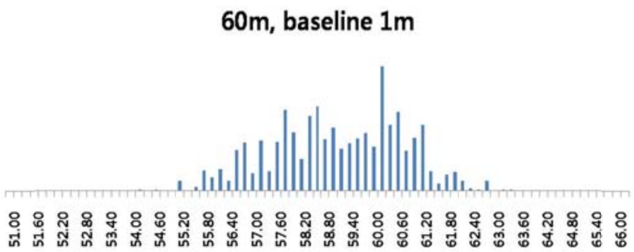
Table 1 shows the experimental results on estimating depth information from stereo matching for indoor environment. The leftmost distances are the ground truths that have been measured. The theoretical depth ranges are calculated using (1) and intrinsic camera parameters. The range or depth bound is derived by the maximal pixel error in the disparity values by ±0.5 pixels. In the table, d is the disparity values in pixels, and σ and mean are the standard deviation and mean value of estimated depth information from more than 2200 frames respectively. As shown in the table, the estimated depth information exists in the theoretical ranges with higher than 95% confidence. Thus, the proposed error model is appropriate to define the error bound in estimating the depth information.

Note that there are jitter error and non-synchronization problem in the IR stereo cameras. To avoid these errors, we move the object very slowly, and find the statistic of depth information. Thus, the experimental results are reliable to evaluate the disparity estimation errors by ±0.5 pixels.

Fig. 5 and 6 shows the histograms of estimated depth information for 60m and 40m ground truths. Two baselines, 0.5m and 1m were used to obtain the histograms. In the case of 80m distance, the baseline 1m was not enough to resolve the depth information. As we can see in Fig. 5 and, 6, the distributions of depths are not regular as we assumed. When we concentrate on the theoretical depth ranges, the histograms of estimated depths are close to the uniform distribution. For example, consider the histogram from 57m to 62m in Fig. 5, and from 38m to 41m in Fig. 6. As we can see, there is no special form of histograms in the ranges. Therefore, we can model the depth errors from disparity values as the uniform distribution in the range cause by ±0.5 pixel error. This result is caused by disparity errors of ±0.5 pixels even though the disparity values are correctly estimated.

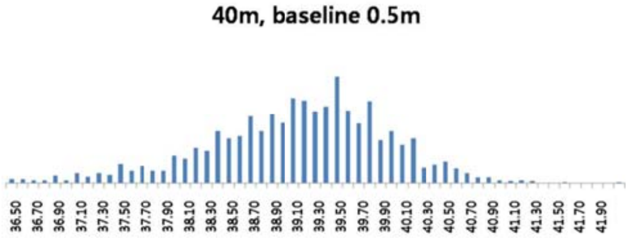


(a)

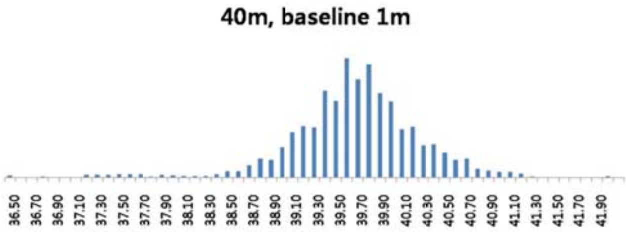


(b)

Figure 5. Histogram of depth information for 60 m. (a) 0.5 m baseline, (b) 1m baseline.



(a)



(b)

Figure 6. Histogram of depth information for 40 m. (a) 0.5 m baseline, (b) 1m baseline.

3.3 OUTDOOR EXPERIMENTAL RESULTS

We had another experiment on the fast moving object for outdoor environment. The ground truth is measured by 3-D trajectory tracker which measures the 3-D position of moving object at every 2ms. The tracker gave the ground truth from 30m to 52m, and we extended the depth range by regression analysis from 15 m up to 80m. The target moves at the speed of 172m/s and approaches the stereo camera aside by 5m. Fig. 7 shows the trajectory of moving object. Fig. 2 shows a stereo image pair for outdoor environment. The very small object is detected and stereo matching is performed on the mean positions of object positions as mentioned before. Fig. 2 shows the results of depth estimation, ground truth, and error bound at the same time.

Table 2 shows the experimental results on estimating depth information (z) from stereo matching for outdoor environment. The proposed error model could find the small target's position in high speed, approaching within near range and this method shows the possibility of the target measuring by using FIR stereo camera.

Fig. 8 shows the estimated depth information for outdoor object on the object trajectory. The depth is estimated on the error bound from disparity estimation.

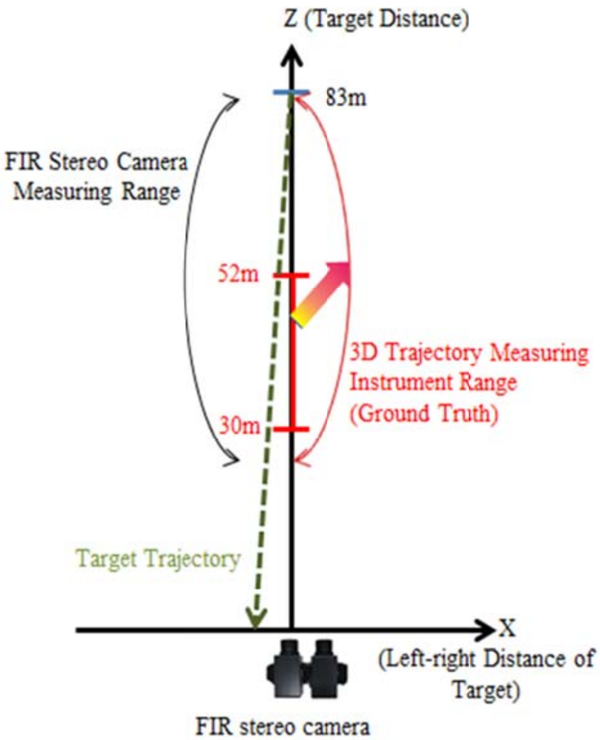


Figure 7. Outdoor experiments using FIR stereo camera.

4 CONCLUSIONS

This paper has proposed an error model in the stereo matching and evaluated it in estimating the 3-D depth information using stereo camera system. The proposed error model is that the disparity values in the pixel domain have ±0.5 pixel errors maximally even though the disparity values are correctly estimated. Thus, the depth information also has the same amount of errors with respect to the focal length and baseline of stereo cameras. To evaluate and apply the proposed model to the military applications, we use uncooled FIR cameras to set the stereo camera system and consider the small high-temperature objects as the targets. We first detect the high-temperature moving object in the stereoscopic IR images, and estimate the disparity field in the pixel domain. Then, we calculate the physical depth between the cameras and moving object by using triangulation of disparity, baseline and focal length. According to the experiments with various distances and baselines, we evaluate the error bound in the depth estimation. This error bound and stochastic parameters are useful for estimating the 3-D position of moving objects probabilistically.

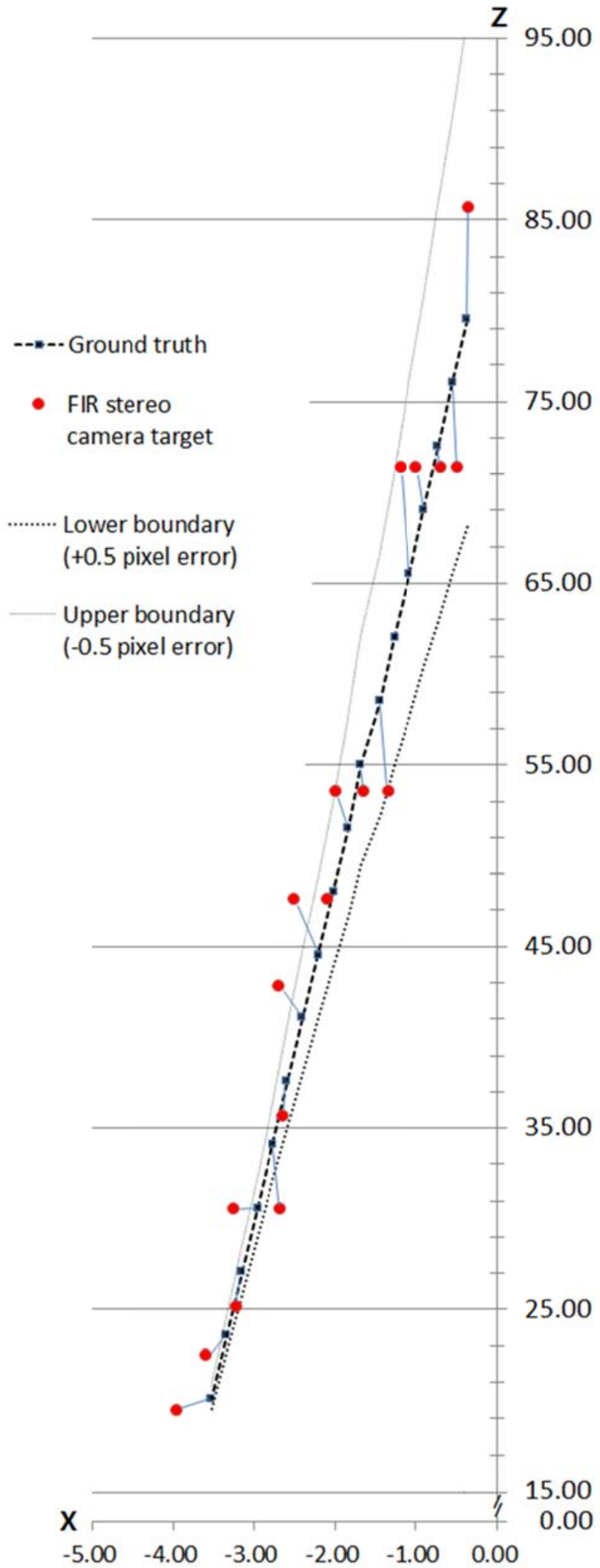


Figure 8. Position error (±0.5pixel) boundary of ground truth and target point using FIR stereo camera.

Table 1. Results of depth estimation for indoor environment

Distance (m)	Baseline(m)	Position Error				
		Theoretical Depth Range (±0.5 pixel error)		Experimental Depth Range (95.9% Confidence)		
		d (pixel)	Depth range(m)	σ(m)	mean(m)	Depth range(m)
79.4	1	9.0	75.2 ~ 84.0	2.0	77.9	74.0 ~ 81.9
59.5		12.0	57.1 ~ 62.1	1.6	59.0	55.8 ~ 62.3
39.7		18.0	38.6 ~ 40.8	0.6	39.7	38.5 ~ 40.8
79.4	0.5	4.5	71.4 ~ 89.3	5.7	77.6	66.2 ~ 89.0
59.5		6.0	54.9 ~ 64.9	3.1	58.5	52.2 ~ 64.8
39.7		9.0	37.6 ~ 42.0	0.8	39.1	37.5 ~ 40.8

Table 2. Target position using FIR stereo camera and position error of ground truth in outdoor experiment.

Frames No.	Depth	
	Depth Range of Ground Truth(m) (±0.5 Position error)	FIR Stereo Target(m)
1	68.1 ~ 95.9	85.71
2	65.5 ~ 90.8	71.43
3	60.3 ~ 80.8	71.43
4	57.7 ~ 76.0	71.43
5	55.0 ~ 71.3	not measured
6	52.3 ~ 66.6	53.57
7	49.5 ~ 62.1	53.57
8	46.7 ~ 57.6	53.57
9	43.9 ~ 53.2	47.62
10	41.0 ~ 48.9	47.62
11	38.0 ~ 44.7	42.86
12	35.1 ~ 40.5	35.71
13	32.0 ~ 36.4	30.61
14	29.0 ~ 32.4	30.61
15	25.9 ~ 28.5	25.21
16	22.7 ~ 24.6	22.56
17	19.5 ~ 20.8	19.48

5 REFERENCES

1. Z. Zhang, "A flexible camera calibration by viewing a plane from unknown orientations," *Proc. of International Conference on Computer Vision*, pp. 666-673, Corfu, 1999.
2. O. Riou, S. Berrebi, and P. Bremond, "Non-uniformity correction and thermal drift compensation of thermal infrared camera," *Proc. of SPIE*, volume 5405, pp. 294-302, 2004.
3. D. Maltese, "New generation of naval IRST," *Proc. of SPIE*, volume 7660, pp. 766004-1-76604-12, 2010.
4. R. Saran, and A. K. Sarje, "Robust long range target detection algorithm using adaptive selective top-hat transform," *Proc. of the International Conference on Image Information Processing*, pp. 6108852-1-6108852-5, 2011.
5. S.H. Kim, "Target attribute-based false alarm rejection in small infrared target detection," *Proc. of SPIE*, volume 8537, pp. 85370G-1-85370G-12, 2012.
6. E. R. Parshall, J. M. Mooney, and A. Anselmo, "Infrared stereo imaging for 3D tracking of point targets," *Proc. of SPIE*, volume 2847, pp. 139-148, 1996.
7. B. Jutzi, R. Gabler, and K. Jaeger, "Stereo vision for small targets in IR image sequences," *Proc. of SPIE*, volume 4473, pp. 361-370, 2001.