SMOKEBOT

Mobile Robots with Novel Environmental Sensors for Inspection of Disaster Sites with Low Visibility

Project start: January 1, 2015

Duration: 3.5 years

Deliverable 3.3

Thermal Stereo Vision

Due date: month 30 (June 2017)

Lead beneficiary: ORU

Dissemination Level: Public

Main Authors:

Rasoul Mojtahedzadeh (ORU)

Version History:

0.1: initial version, rm, June, 2017

0.2: first revision, es, June 2017

0.3: second revision, rm, June 2017

Contents

A Introduction and purpose of this document

This document is part of the Horizon 2020 project SmokeBot. The project's goal is to develop mobile robots for applications in disaster scenarios with low visibility conditions. A key element in the development of the project is the perception of such harsh environment in the presence of smoke and fire in order to support the cooperation of humans (e.g. fire fighters) and robots in search and rescue missions. A multitude number of sensors - as a multimodal sensor unit - including radar, gas, thermal and traditional vision sensors for this purpose are employed. This deliverable describes the preliminary research results carried out in the applicability of a Thermal Stereo Vision setup.

Detection of thermal targets for search and rescue robots is a critically important task that its fast and reliable response could potentially save more lives. Some research work reported with a focus on human detection [1] [2], developing of exploration algorithms for thermal targets [5], and hazard detections [3] [4] from thermal imaging sensors. However, the application of stereo thermal vision in the task of search and rescue missions by robots has received less attention.

In an environment with smoke and hot surfaces, the traditional vision and laser range finders cannot perceive the true structure of the environment due to a low visibility and erroneous range measurements. A potential alternative for a 3D perception of such harsh environments is to use a stereo thermal vision system.

As of the first step of using a stereo vision system, the calibration of intrinsic, extrinsic and lens distortion parameters of the thermal cameras is carried out, and rectification maps of the stereo thermal images are estimated. Stereo matching of the rectified thermal images is the next challenging step to obtain a 3D perception of the environment.

B Stereo Thermal Vision Calibration

In this task, two FLIR thermal cameras of the same type are used in order to obtain stereo thermal images for a 3D perception of the environment under harsh conditions such as smoke and fire. The first step is to calibrate the cameras to estimate the intrinsic, extrinsic and lens distortion parameters. The focal length of the cameras are fixed to 13 millimeters with manual fine tuning. The field of view of the thermal cameras is 45° H x 37° V. The trigger of the cameras can be synchronized, and they are capable of capturing thermal images of resolution 640 × 512 at maximum 30 frames per second.

In order to obtain rectified thermal images, which is required for stereo matching, the intrinsic parameters, extrinsic rotation and translation as well as rectification matrices are required to be estimated. To achieve the calibration goal, a proper calibration pattern that can be perceived by thermal imaging sensor needs to be used. Following we examined a few different approaches to create such calibration pattern.

In the first attempt, a chessboard calibration pattern with 10 by 6 corners and a square width size of 65 millimeters printed on a glossy paper and mounted on a frame made of wood was first used for the experiment. An infrared light source illuminated the surface of the pattern while test thermal images were captured. Figure depicts some thermal images of the chessboard pattern taken by the left and right cameras.

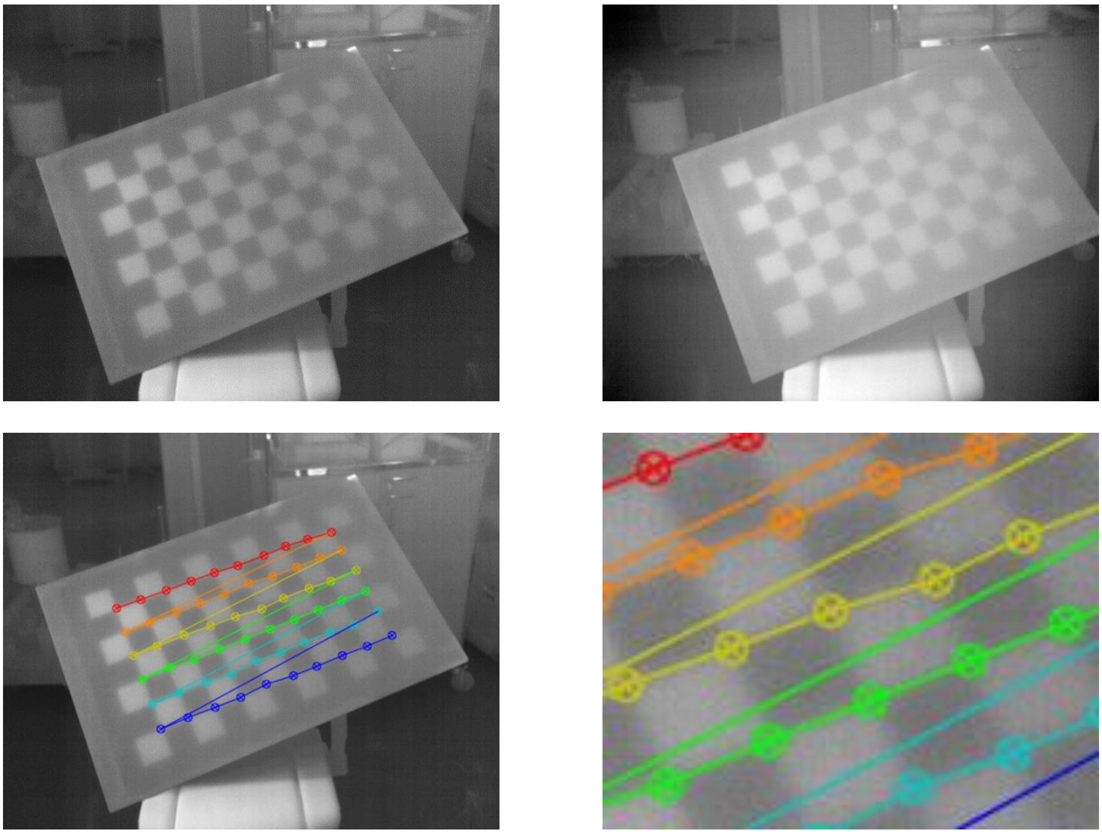

Figure 1: (up row) A sample of a stereo pair of thermal images of the chessboard calibration patterns printed on a glossy paper (bottom left) A sample of chessboard corners detection of the thermal image. (bottom right) A close-up of the detected corners that shows poor corner detected due to low sharpness in the image.

We run a chessboard detection algorithm on the test images shown in Figure, and observed the detected corners of poor quality (see Figure) due to low sharpness in the image. This experiment showed that although a traditional chessboard pattern printed on a glossy paper and lighted by an infrared source can be visible to the thermal cameras but the sharpness of the chessboard edges is not adequate for a reliable calibration.

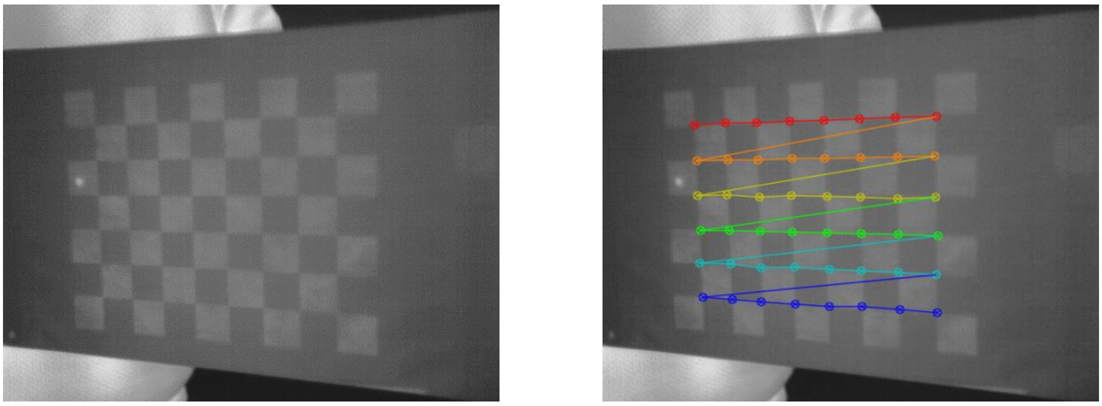

Figure 2: (left) A sample of a thermal image of the chessboard calibration patterns printed on a typical A4 paper and mounted on an aluminium plate with powder coating. The edges of the pattern are sharper in the image. (right) A sample of chessboard corners detection of the thermal image which represents a higher accuracy in the detection.

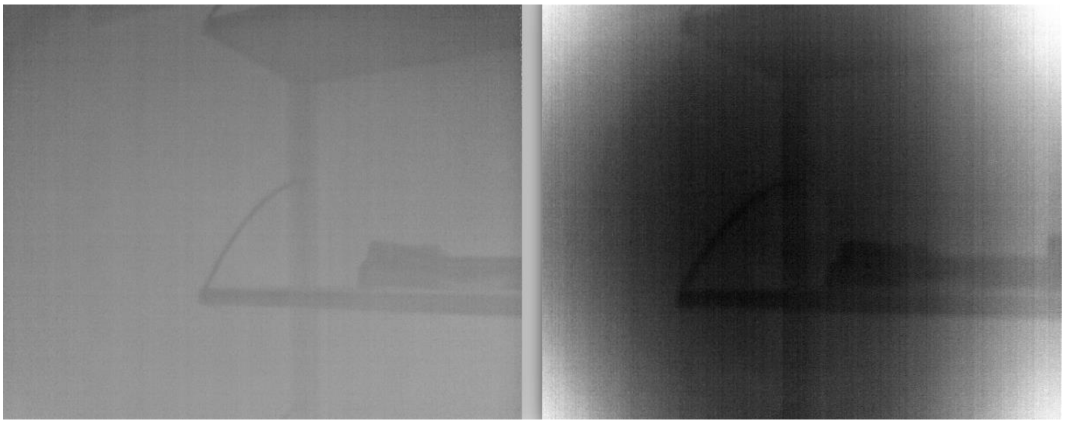

Figure 3: (left) A sample of the thermal image taken by the intact camera. (right) The corresponding thermal image taken by the damaged camera. A radial darkness can be observed on the image which is due to a malfunction of the camera.

In the second attempt, we printed a chessboard pattern with 8 by 6 corners and a square width size of 22 millimeters on a typical A4 paper and stuck it on a piece of aluminum plate with powder coating and lighted the pattern with an infrared source. The sharpness of the edges of the chessboard squares in this case is significantly higher than the previous setup, and as a result the quality of detection of the chessboard corners is better.

Figure shows a sample of thermal image of the chessboard with the detected corners. A visual comparison of the images in Figure and one can observe that the corners in the latter case can be recognized with a higher accuracy.

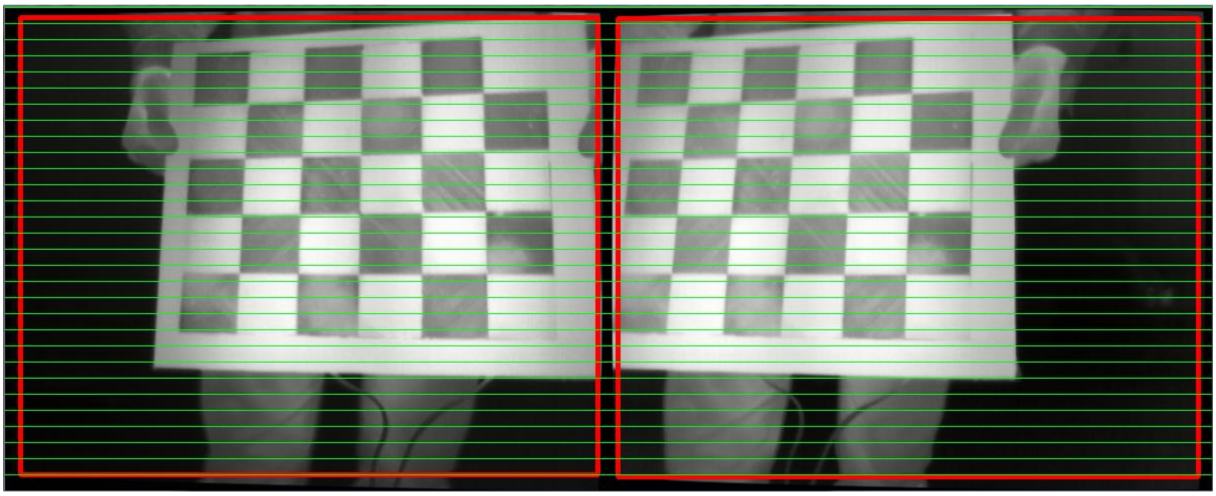

Figure 4: A sample pair of the thermal stereo images rectified by the calibration. The rectified images reduces the search for matching points on corresponding horizontal lines. The chessboard in this experiment is heated pads which gives the best results.

During the experiments with the thermal cameras, we observed that one of the cameras produce a radial darkness on its output image due to some malfunction. This camera needs to be repaired or replaced in order to be able to continue working on the task T3.3. Figure depicts the problem where two thermal images of the left and right cameras are shown.

Having the thermal camera factory recalibrated, in order to find the corners of the chessboard pattern more accurately, the same heated pattern shown in Figure 3 was used for stereo calibration. For the calibration process, tools developed within OpenCV package were used. Figure 8 shows a sample pair of thermal stereo images rectified by the calibration data.

C Stereo Matching

Stereo matching is the correspondence problem – for a point in the world which is projected on a pixel in the left (right) image, where is the corresponding pixel location in the right (left) image? In general, the search for the correspondence pixel has to be performed in entire image, while there is no guarantee that such correspondence projection does exist due to occlusion and different field of view of the cameras. The image rectification reduces the search space from the entire image to the corresponding row or column (epipolar line) of the image depending on horizontal or vertical rectification. Stereo disparity is the difference in position between correspondence points in two images, based of which the depth of the projected point in the world can be estimated.

Existing stereo matching algorithms can be divided into two major categories – local methods based on correlation and global (semi-global) methods based on minimization of some pixel energy function. In local methods, for every pixel the disparity is selected independently by means of a similarity measure (correlation) defined in a fixed sized rectangular window (block matching). Global methods, on the other hand, define an objective energy function that unlike local methods, attempts to penalize discontinuities of disparities of adjacent pixels to obtain a more coherent stereo matching results.

Among the similarity measures for local methods of stereo matching are Sum of Squared Differences (SSD), Sum of Absolute Differences (SAD), Locally-scaled Sum of Squared Differences (LSSD), Normalized Cross Correlation (NCC) and Sum of Hamming Distances (SHD), where it is shown that the performance of the mentioned similarity measures is comparable [6]. The window or block size, however, determines the behavior of the local stereo matching performance. Larger block size results in smoother, though less accurate disparity maps, whereas smaller block size produces more detailed disparity map, but also increases the wrong correspondences. Block matching approach assumes constant depth within the window, where this assumption is not valid at depth discontinuities neither is valid for non-planar surfaces, and repetitive textures and uniform areas in the images makes it difficult for the block matching method to be robust.

The global and semi-global matching methods cast the problem of estimating disparities of the pixels as an energy minimization problem, where the objective function to minimize is the sum of similarity of the pixels plus a penalty term for disparity variations. Dynamic programming, scanline optimization, graph cuts and belief propagation are well-known algorithms to solve such optimization problems. Dynamic programming finds the *edit distance* between two corresponding lines of rectified images [7]. Scanline optimization [8] is a search for a globally optimal assignment of disparity values to pixels in the horizontal or vertical scanline. Graph cuts method cast the problem of stereo correspondence as a pixel labeling problem by solving for a maximum flow in a graph [9]. Belief propagation formulates the stereo matching problem as a Markov network consisting of three coupled Markov random fields (MRF’s) - a smooth field for depth (disparity), a line process for depth discontinuity and a binary process for occlusion [10]. The global matching methods produce dense disparity maps with good performance while computationally are very expensive.

An alternative approach that represent a trade-off between computational time and robustness for the problem of stereo matching is semi-global block matching (SGBM) [11]. SGBM method is based on the idea of pixelwise matching of mutual information and approximating a global, 2D smoothness constraint by combining many one dimentional constraints.

The implementation of Stereo Block Matching (StereoBM) and Semi-global Block Matching (StereoSGBM) algorithms were examined for finding corresponding points between two rectified thermal images. The best results obtained by StereoSGBM with parameters shown in Table 1, with setting the full dynamic programming enabled.

preFilterCap	12	numberOfDisparities	240
SADWindowSize	19	uniquenessRatio	5
P1	8*19*19	speckleWindowSize	56
P2	32*19*19	speckleRange	39
minDisparity	10	disp12MaxDiff	10

Table 1. The values of StereoSGBM parameters used for stereo matching of thermal cameras.

The performance of the algorithm on SmokeBot's robotic platform is about 5 frames per second. In general there are a multitude challenges for the task of stereo matching -- naming a few, image sensor noise, specularities, foreshortening and the uniqueness constraint, perspective distortions and reflections are difficulties to address. In Figure 5, a sample of a matching result is shown.

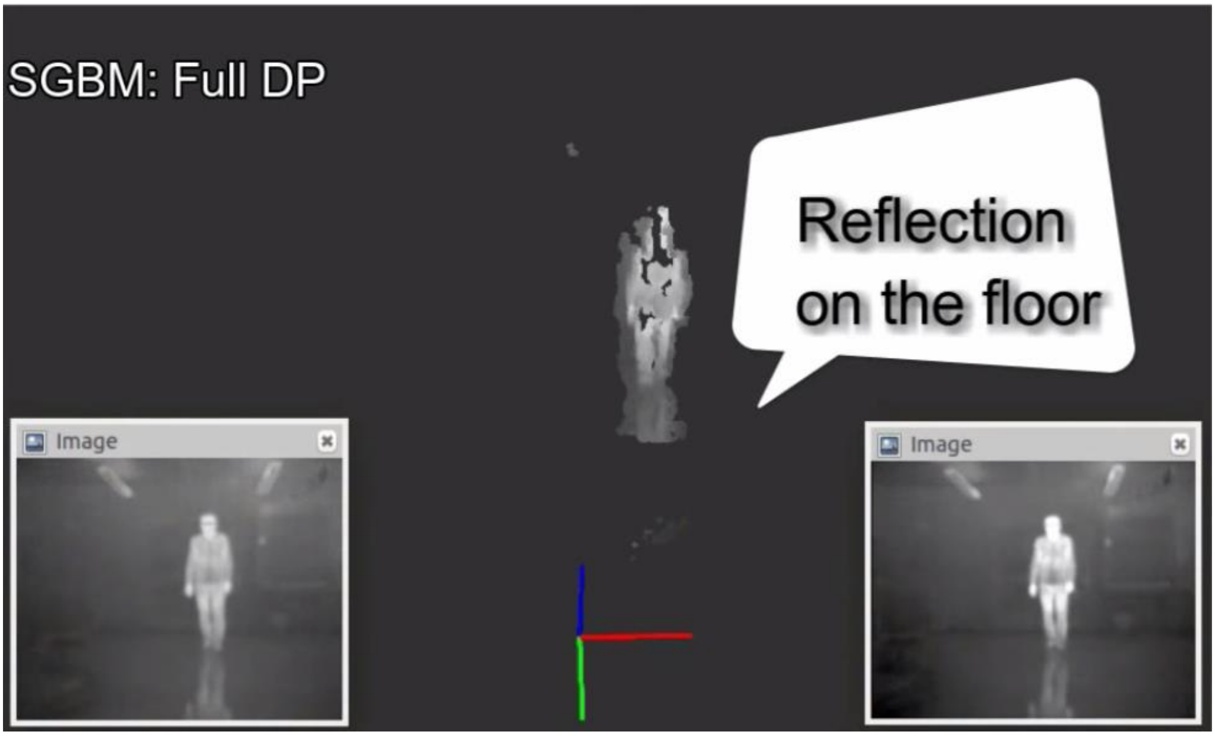

Figure 5: A sample point cloud derived from stereo matching using StereoSGBM algorithm with parameters shown in Table 1. The matching points of the body heat reflections on the floor are highlighted.

D Summary and Outlook

While stereo thermal camera imaging is quite interesting, the benefits compared to single stereo imaging are minor. In the presence of several features in thermal images e.g. those of a person, it is possible to perform the depth estimation reliably. Yet, in typical scenes considered in the project, which excludes explicitly people detection, the number of features to perform a proper stereo matching between the two thermal images is too low. Another effect to consider was shown in D3.1. The “smoothing” effect of smoke in thermal images reduces the number of useful features, which further complicates the stereo matching process in the presence of smoke. Hence, it is only very rarely possible to perform a robust depth estimate.

The stereo matching comes at two major costs: 1) the computational costs to perform the stereo matching and 2) the amount of data through-put on the internal network of the robot (where image data is the most demanding information sent), that increases considerably with a second thermal camera. In order to preserve computational power and robot internal bandwidth for the radar camera under development and in light of the difficulties encountered to practically compute the depth estimate. It was decided to only use a single thermal camera on the SmokeBot demonstrator.

References

[1] H. Mano, K. Kon, N. Sato, M. Ito, H. Mizumoto, K. Goto, R. Chatterjee, and F. Matsuno, “Treaded control system for rescue robots in indoor environment,” in Proceedings of the IEEE International Conference on Robotics and Biomimetics, Guilin (China), 2008, pp. 1836–1843.

[2] M. Z. Aziz and B. Mertsching, “Survivor search with autonomous UGVs using multimodal overt attention,” in Proceedings of IEEE Safety Security and Rescue Robotics, Bremen (Germany), 2010, pp. 1–6.

[3] L. Merino, F. Caballero, J. R. Martínez-de Dios, J. Ferruz, and A. Ollero, “A cooperative perception system for multiple UAVs: Application to automatic detection of forest fires,” J. Field Robot., vol. 23, no. 3–4, pp. 165–184, 2006.

[4] D. W. Casbeer, R. W. Beard, T. W. McLain, S.-M. Li, and R. K. Mehra, “Forest fire monitoring with multiple small UAVs,” in Proceedings of the American Control Conference, Portland (OR, USA), 2005, pp. 3530–3535.

[5] F. Amon, A. Hamins, and J. Rowe, “First responder thermal imaging cameras: establishment of representative performance testing conditions,” in Proceedings of SPIE - The International Society for Optical Engineering, Orlando (FL, USA), 2006, p. 62050Y–62050Y–12

[6] Roma, Nuno & Santos-Victor, José & Tomé, José. (2002). A Comparative Analysis Of Cross-Correlation Matching Algorithms Using a Pyramidal Resolution Approach. . 10.1142/9789812777423_0006.

[7] S. Forstmann, Y. Kanou, Jun Ohya, S. Thuering and A. Schmitt, "Real-Time Stereo by using Dynamic Programming," 2004 Conference on Computer Vision and Pattern Recognition Workshop, 2004, pp. 29-29. doi: 10.1109/CVPR.2004.154

[8] C. Zach, M. Sormann, K. Karner, Scanline Optimization for Stereo On Graphics Hardware, 3DPTV, 2006

[9] Yuri Boykov, Olga Veksler, and Ramin Zabih. 2001. Fast Approximate Energy Minimization via Graph Cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23, 11 (November 2001), 1222-1239. DOI=http://dx.doi.org/10.1109/34.969114

[10] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum. 2003. Stereo Matching Using Belief Propagation. IEEE Trans. Pattern Anal. Mach. Intell. 25, 7 (July 2003), 787-800. DOI=http://dx.doi.org/10.1109/TPAMI.2003.1206509

[11] Heiko Hirschmuller. 2008. Stereo Processing by Semiglobal Matching and Mutual Information. IEEE Trans. Pattern Anal. Mach. Intell. 30, 2 (February 2008), 328-341. DOI=http://dx.doi.org/10.1109/TPAMI.2007.1166