AN OMNIDIRECTIONAL STEREO SYSTEM FOR LOGISTIC PLANTS.
PART 1: CALIBRATION AND MULTI-CHANNEL RECTIFICATION

Marius DRULEA, Andrei VATAVU, Szilárd MANDICI, Sergiu NEDEVSCHI

Technical University of Cluj-Napoca, Faculty of Automation and Computer Science, Romania
Corresponding author Sergiu NEDEVSCHI, E-mail: `sergiu.nedevschi@cs.utcluj.ro`

Abstract. This work presents a fisheye lens based omnidirectional stereo sensor. The sensor is installed on the top of the vehicle such that it can observe the entire area around it. The vehicles we use in the experiments are automated forklifts. We propose a multi-channel rectification to divide the original fisheye images into three pairs of rectified perspective images. The complexity of the fisheye stereo processing therefore reduces to a series of standard image processing techniques. This is especially useful for the efficient generation of the 3D information from images, where the epipolar lines have to correspond to the lines in the images. We use a standard dense stereo matching engine to generate three disparity images from the pairs of rectified images. The stereo reconstruction, the processing of the disparity images for obstacles detection and the detection accuracy evaluations are presented in the second part of the paper.

Key words: omnidirectional stereovision, fisheye lens, multi-channel rectification.

1. INTRODUCTION

Omnidirectional 3D perception of the environment is desired in most computer vision applications. The most well-known applications where omni-stereo sensors are used are robot applications. The robot can easily turn around or move backward and therefore the sensors are required to provide 360 degree perception. In driving assistance applications a large field of view of perception is also desirable, especially in intersections, as it was presented in [1–3]. Small field of view cameras are still used here because they provide a higher angular resolution. It is also more difficult to mount a 360 degree sensor on a car. 3D environment reconstruction and mapping provides maps that can be later used for navigation or to simply present them to the users. The map might consist of the cloud of 3D points or a list of extracted primitives like corridor, wall, static obstacle and other entities depending on the application.

The literature offers several types of omnidirectional sensors. Catadioptric systems consist of a camera, a rectilinear lens and a convex mirror [4]. Two types of mirrors can be used: hyperbolic and convex mirrors. Perspective cameras are combined with hyperbolic mirrors [5], while orthographic cameras are combined with parabolic mirrors [6]. Other combinations are also possible, but these two provide a single effective viewpoint [7–9]. Omnidirectional sensors can be obtained by combining an imaging sensor with omni-lenses like wide angle and fisheye lenses. These sensors are smaller than the catadioptric systems and provide a better angular resolution. Recently, Aikio *et al*. [10] have invented a new lens which is significantly more compact than a catadioptric system and provides a field of view of 360 degrees. Existing applications use omnidirectional stereovision mainly for video surveillance and driving assistance [40–42].

For AGV surrounding 3D perception, a novel solution is proposed by using two omnidirectional imaging sensors. The two sensors can be mounted horizontally or vertically, depending on the application. In the horizontal configuration the imaging planes are almost coplanar. In the vertical configuration one sensor is mounted on the top of the other. The condition to have two physical sensors to obtain 3D perception is not always necessary, as we can see in [11–13]. Southwell *et al*. [11], Yi and Ahuja [12] presented an interesting way to get the 3D information by using a single camera, a lens and two convex mirrors. Li and Li [13] presented a stereo sensor consisting of a fisheye lens and a convex mirror. Instead of an omnidirectional stereo sensor, several perspective stereo sensors can be used, as in [14]. Such a system is more expensive, but the accuracy of the reconstruction is higher. For offline 3D mapping an omnidirectional stereo sensor is

not mandatory. Here, a single camera with a small field of view suffices because the acquisition and reconstruction process can be employed at multiple locations, like in [15].
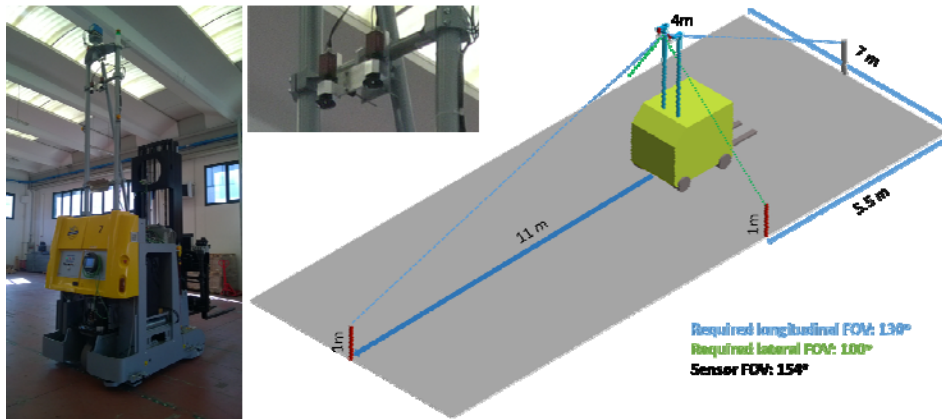

Fig. 1 – *Left*: The AGV and the stereo sensor. *Top middle*: Close-up look of the stereo sensor installed on the AGV. *Right*: The mounting position, range and field of view of the fisheye stereo sensor

There is no clear winner among the omnidirectional sensors. The best sensor is decided based on the application requirements.

The PAN-Robots project [16] proposes a new generation of autonomous guided vehicles for autonomous logistics. The automated guided vehicle (AGV) used in this project is an automated forklift. In a warehouse there are several AGVs whose tasks are to carry goods to and from the storage locations [17]. The map used for navigation is static and it is extracted prior to the installation of the AGVs [18]. The AGV's task consists in a starting point, the type of good to be transported and a final destination point. Besides this, the AGV also receives a route from the starting point to the destination point where the goods have to be unloaded [19–23]. Within the PAN-Robots project, infrastructure sensors are proposed in order to improve the efficiency of AGVs in warehouse intersections [24]. A camera based approach for loading and unloading the pallets is also proposed [25]. While the AGV moves along its route, it can meet obstacles like other AGVs, manual forklifts, workers, pallets of goods and other possible objects. For a safe navigation, these obstacles have to be detected. Prior to PAN-Robots, the automated forklift used several safety laser scanners in order to detect the obstacles around it. The forklift decreases its speed if obstacles appear on its route. The safety laser scanners issue an emergency brake command whenever a collision with an obstacle might occur. Unfortunately, the scanning area of laser-scanners is limited to a single horizontal plane. Hanging and protruding objects are not detected and might cause stacks of pallets to fall. It is also more difficult to distinguish between humans and other obstacles. The purpose of this work is to develop a camera based stereo sensor which is able to detect, track and recognize all the obstacles around the AGV. The onboard system of the AGV fuses the list of object detected by our system with the list of objects provided by the laser-scanners. The system will benefit from the dense information of the stereo sensor in order to provide a safer navigation.

Choosing the right stereo sensor was an important decision in the project. The candidate stereo sensors were the following: perspective cameras, catadioptric systems, omni-lenses, fisheye lenses and wide angle lenses. The sensing devices were compared w.r.t. angular resolution, field of view, light collecting capability, production cost, endurance and maintenance costs. They also need to fit onto the automated forklifts. If perspective stereo cameras are considered, several pairs have to be used to cover the surrounding of the forklift. This solution is expensive. With respect to the production cost, the catadioptric lenses are slightly less expensive than the fisheye lenses. However, all the other criteria are in favor of the fisheye objectives. The fisheye lenses can easily provide the field of view required by the project. For the catadioptric systems, this is problematic. The angular resolution of the fisheye lenses is better since they cover a bigger part of the imager. They are more compact, easier to clean and they have a better resistance to vibrations. We have decided to use a fisheye stereo sensor mounted on top of the vehicle, at 4m above the ground. It was mounted on the existing two poles of the forklift, which are used for the navigation laser-scanner. Fig. depicts this setup. The optical axes of the lenses point toward the ground. In this manner, a single omni-

stereo sensor can monitor the entire area around the AGV. Mayra *et al.* [26] built fisheye objectives with better specifications than the off-the-shelf fisheye objectives. The lenses were designed to fit into the logistic factories targeted by the PAN-Robots project. The objectives have a maximized angular resolution w.r.t. 1'' type camera sensor for a field of view of $360^o \times 154^o$. They have a focal length of 4.7 mm and a light collecting capability of 1.2. The light collecting capability is important because the environment is darker in logistic factories.

The fisheye stereo sensor has to detect, track and classify all the objects around the AGV. In this first part of paper, we present the calibration and the rectification modules. The stereo reconstruction and obstacle detection modules are presented in the second part of the paper [27]. The tracking and classification of the objects will be available in the next phases of the project.

The present paper extends a preliminary conference version [28] by providing additional information about the system and detailed descriptions of the software components. The software components are enhanced and give a significantly higher obstacle detection rate. In order to evaluate the system, a new test session took place in the test warehouse. The experimental results are presented in the second part of the paper [27].

The rest of the paper is organized as follows. Section 2 presents the related work. Section 3 presents an overview of our system. Section 0 presents the camera and lens calibration and our approach for fisheye image rectification. We conclude the paper in Section 5.

2. RELATED WORK

Gehrig *et al.* [2] implemented a 6D stereo sensor for a driving assistance application. A pair of fisheye lenses were used to monitor large areas in intersections. The 6D positions of the image features consist of their 3D position and their 3D motion vector. The authors have used the calibration toolbox presented in [29], which was originally designed for pinhole cameras with large distortions. The rectification model involves the projection of the image onto a cylinder. A standard pinhole rectification was also considered, but this significantly increased the size of objects at the edges of the rectified images. A number of features were detected in the rectified images. The features were tracked over time using the Kanade-Lucas-Tomasi tracker [30, 31]. The disparities of all features were computed using a local stereo matching module. These two steps enable the computation of the 6D state of each feature w.r.t. to the ego system. The 6D state is then compensated with the ego motion of the system, which is computed by matching the clouds of raw 3D points over time [32]. For numerical evaluations, the authors measured the depth of several selected image features with a laser pointer. The ground-truth depths were compared with those provided by the system. The disparity errors for the selected features were proven to be within one pixel. In this experiment, a baseline of 30 cm was used. The authors also performed an experiment in an intersection using a baseline of 50 cm. No numerical evaluations were provided for this experiment.

As in [2] we also use fisheye lenses in order to reach a large field of view perception. However, our application targets the logistics sector, the fisheye stereo sensor being mounted on automated forklifts. We use a multichannel fisheye image rectification, each channel being a pinhole-rectified image. We use a real-time SGM GPU implementation for computing dense stereo disparities of the channels of rectified images. Besides the computation of the 3D position of image pixels, we also develop an obstacle detection module using a grid-based approach. In contrast to [2], we perform more advanced evaluations of the system. Our evaluation takes place at object level. We measure the object detection rate, the accuracy of obstacle localization and the accuracy of the size of these objects. The evaluations are performed on sequences where the automated vehicle navigates among the objects.

3. SYSTEM OVERVIEW

The software architecture of the system is given in Fig. 2. The acquisition module receives images from the fisheye stereo sensor. The transmission and the reception of the images are not guaranteed to be synchronous. The frame synchronization module correctly pairs the input images based on their timestamps. The fisheye images are then sent to the multi-channel rectification module, presented in section 0. Here, the

images are divided into three pairs of perspective stereo images. The rectification module uses the intrinsic and extrinsic parameters of the cameras. The calibration module computes these parameters offline. The stereo matching and 3D reconstruction engine takes all the pairs of rectified images and generates a cloud of 3D points for each pair. The obstacle detection module merges the 3D clouds into a unified digital elevation map. It then proceeds with the ground plane estimation and then with the grid classification and object detection. The second part of the paper [27] presents details about the obstacle detection module and its sub-modules. Finally, the detected list of obstacles is sent to the application layer where it is used for collision avoidance and local-path planning [33, 34]. The application layer is not part of our system.

On the hardware side, we use a custom pair of fisheye lenses [26]. Any other fisheye lenses can be used instead. However, the custom-made lenses are optimized for the PAN-Robots specifications. They optimally cover 1'' type camera sensor for a field of view of $360^{o} \times 154^{o}$ and a light collecting capability of F/1.2. We use two Gigabit Ethernet Manta G-419 B NIR cameras. The baseline between the cameras is 20 cm. In order to facilitate the synchronization between the left and right images we simultaneously trigger the cameras at a frequency of 10 Hz using an Arduino Uno microcontroller. Even if the trigger is synchronous, the transmission and the reception of the image data are not guaranteed to be so. The frame synchronization module solves this issue. We use an NVIDIA GeForce GTX 780 Ti to carry the intensive computation of the stereo matching and 3D reconstruction engine.
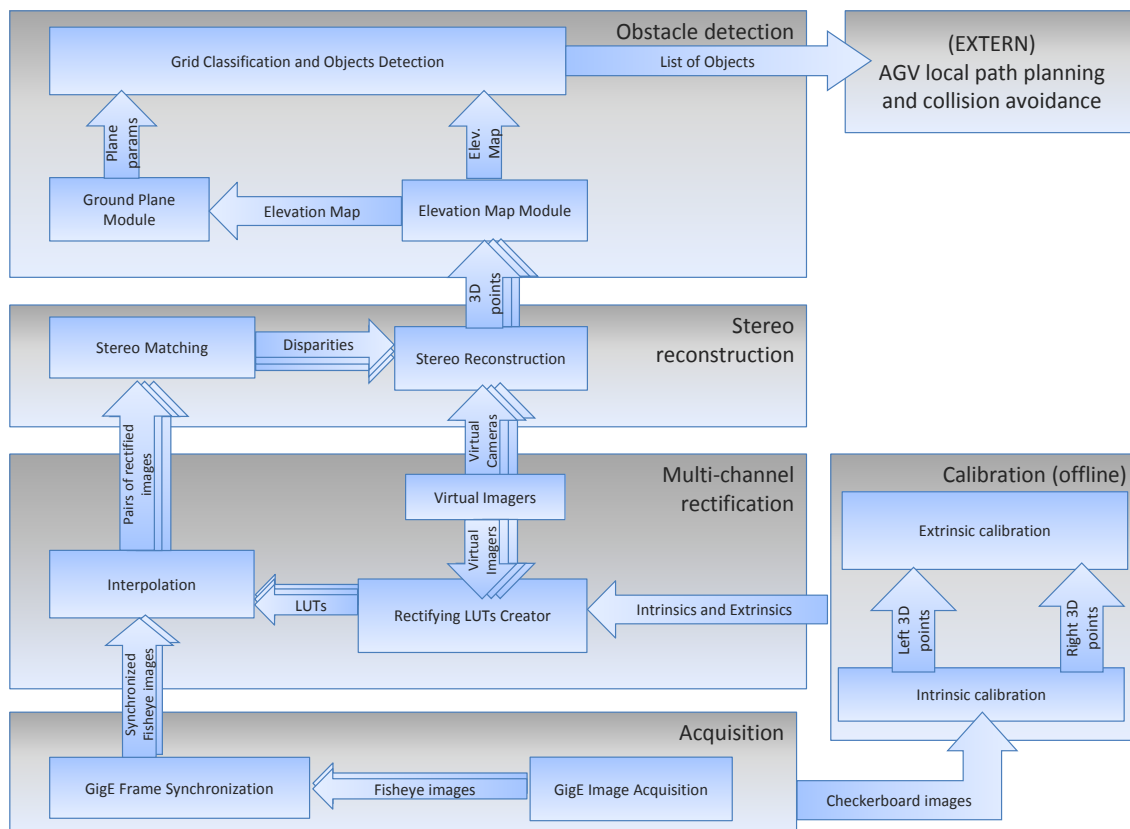

Fig. 1 – The software architecture of the omnidirectional stereo system.

4. CAMERA CALIBRATION AND MULTI-CHANNEL RECTIFICATION

4.1. Fisheye lens description and calibration

The fisheye lenses can provide field of views up to 180°. There are four main types of fisheye lenses based on the characteristic projection function: equidistance, stereographic, equisolid and orthographic projection functions [35]. The most well-known fisheye lenses are the equiangular (equidistance) lenses. The lenses we use in the project also belong to this category. The equidistance projection function is:

$$r = f \cdot \alpha, \tag{1}$$

where f is the focal length of the lens, α is the angle of the incoming ray with the optical axis and r is the projection of the ray onto the imaging plane (Fig. 2). The equiangular fisheye lenses have the simplest projection function among the fisheye lenses. Due to inaccuracies in lens production and assembly there are deviations from this rule. The mounting of the fisheye lens and of the electronic imager is also a source of errors. For a precise calibration and a good rectification, these deviations have to be modelled as well.

Kannala and Brandt [36] presented a general projection function that can accurately model all single effective viewpoint lenses: the fisheye, ultra-wide, catadioptric objectives and rectilinear lenses. The projection function is a polynomial:

$$r(\alpha) = k_1 \alpha + k_2 \alpha^3 + k_3 \alpha^5 + k_4 \alpha^7 + k_5 \alpha^9 . \tag{2}$$

For a complete characterization, the authors also considered a distortion profile. The radial and tangential profiles also apply to all single effective viewpoint lenses. The radial distortion has the following form:

$$\Delta_{r(\alpha,\varphi)} = \left(l_1 \alpha + l_2 \alpha^3 + l_3 \alpha^5 \right)\left(i_1 \cos(\varphi) + i_2 \sin(\varphi) + i_3 \cos(2\varphi) + i_4 \sin(2\varphi) \right), \tag{3}$$

where φ is the angle of the incoming ray in the XOY plane. See Fig. 2 for a graphical description. The angles (α, φ) are directly computed from the coordinates (X, Y, Z). Similarly, the tangential distortion is given by:

$$\Delta_{t(\alpha,\varphi)} = \left(m_1 \alpha + m_2 \alpha^3 + m_3 \alpha^5 \right)\left(j_1 \cos(\varphi) + j_2 \sin(\varphi) + j_3 \cos(2\varphi) + j_4 \sin(2\varphi) \right). \tag{4}$$

Including the distortions leads us to the following projection function:

$$\begin{bmatrix} x \\ y \end{bmatrix} = F_{mm}(X,Y,Z) = r(\alpha) \cdot \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} + \Delta_r(\alpha,\varphi) \cdot \begin{bmatrix} \cos(\varphi) \\ \sin(\varphi) \end{bmatrix} + \Delta_t(\alpha,\varphi) \cdot \begin{bmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{bmatrix}. \tag{5}$$

$p(x,y)$ is the projection of the point $P(X,Y,Z)$ onto the imager plane. Both p and P are in millimeters. The final step is to transform the metric coordinates into pixel coordinates. Let (u_0, v_0) be the principal point and let t_u and t_v be the number of pixels per unit mm in horizontal and vertical direction respectively. The final projection function is:

$$\begin{bmatrix} u \\ v \end{bmatrix} = F(X,Y,Z) = \begin{bmatrix} t_u & 0 \\ 0 & t_v \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}. \tag{6}$$

The projection function has 23 intrinsic parameters $(k_{1-5}, f_u, f_v, u_0, v_0, l_{1-3}, i_{1-4}, l_{1-3}, i_{1-4})$. We use the calibration procedure described in [36] and a 8×8 checkerboard pattern to calculate these parameters. Let ck^i be the 3D coordinate of the i-th point of the checkerboard pattern, w.r.t. to the local coordinate system of the pattern. We denote R_m and T_m the rotation and translation of the pattern w.r.t. the camera system, in the m-th calibration image. We have the following relationship:

$$\begin{bmatrix} u_m^i \\ v_m^i \end{bmatrix} = F\left(R_m \cdot ck^i + T_m \right), \tag{7}$$

where $[u_m^i, v_m^i]$ is the corresponding image coordinate of the point ck^i. The image coordinates $\left[u_m^i, v_m^i \right]$ of the checkerboard corners are extracted using the RADDOC checkerboard detector tool [37, 38].
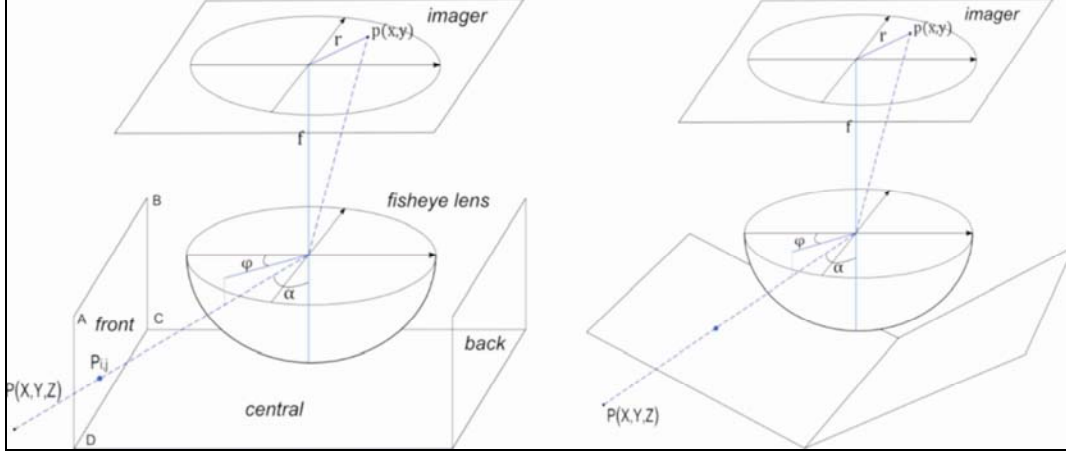
Fig. 2 – *Left*: The projection model of the fisheye lens. The front, central and back virtual imagers used for rectification.
Right: candidate solution with two virtual imagers.

The calibration procedure has to find the intrinsic parameters, rotations and translations that minimize the relations (7) in a least-square sense:

$$\sum_{m,i} \left[\begin{array}{c} u_m^i \\ v_m^i \end{array} \right] - F\left(R_m \cdot ck^i + T_m \right)^2 . \tag{8}$$

The function to be minimized 0 is non-linear and the solution is found using the Levenberg-Marquardt algorithm. The readers can find further details about the calibration procedure in [36].

In our experiments, the average reprojection error of the 8×8 corners in all the images was about 0.11 pixels. The maximum error was about 1 pixel and this occurs for the corners at the image border when the calibration pattern is very close to the fisheye lens.

For the extrinsic calibration we use the coordinates $CK_m^i = R_m \cdot ck^i + T_m$ which are available once the intrinsic calibration is done. Let's denote the CKL_m^i and CKR_m^i these coordinates for the left and right camera respectively. They are related by a rotation R and a translation T of the right camera w.r.t. to the left camera:

$$CKL_m^i = R \cdot CKR_m^i + T . \tag{9}$$

The rotation and the translation can be easily computed using the least squares method. Although R and T can be used directly in the rectification procedure, we prefer to use the matrices R_L and R_R that minimizes the distortions of both images. Therefore, instead of rotating the right system to the left one, we slightly rotate the left system and slightly rotate the right system such that they become aligned. We have the following equality:

$$R_L \cdot CKL_m^i = R_R \cdot CKR_m^i + t, \, t = \left[T, 0, 0 \right]^T . \tag{10}$$

The matrices R_L, R_R and the translation vector t are directly computed from R and T. This technique can be found in [39]. In the experiments, the mean extrinsic error, which is the mean deviation of the equation 0, was 3.46 mm. The maximum extrinsic error was 1 cm.

4.2. Multi-channel stereo rectification

The purpose of the stereo rectification is to correct the image distortions and to virtually align the left and right imaging planes such that they are coplanar. Our rectification method divides the original fisheye image into multiple perspective rectified images. We call them channels. We generate rectified channels for both left and right fisheye images. The complexity of the fisheye stereo processing therefore reduces to a series of standard image processing techniques.

In order to generate the rectified images we create several virtual pinhole imagers, as depicted in Fig. 3. One question is how many virtual imagers we should use. In this paper, we have performed experiments for up to three imagers. If a single central imager is used, the objects at the image borders become very large due to the large field of view of the sensor. We have found that the solution with three channels (Fig. 2, left) is superior to the one with two channels (Fig. 2, right). The two-channel imagers deviate more from the geometry of the fisheye lens and they provide fewer reconstructed 3D points. We do not use more than three virtual imagers because the objects would split over multiple images. This creates problems if we want to recognize the objects using the gray level features in the generated images. Therefore, we have chosen to use three channels. We create a front, a back and a central virtual pinhole imagers, as depicted in Fig. 2, left. The imagers are in fact rectangles in the 3D space relative to the objective optical center. We choose the central rectangle such that it corresponds to a field of view of 90º both horizontally and vertically. The front and back imagers correspond to a field of view of 90º horizontally and 45º vertically.

The rectified images are obtained by sampling the virtual imagers into a uniform grid of virtual pixels. Each virtual pixel is then projected onto the original fisheye image to get its intensity. In the following, we exemplify this process for the front left rectified image I^L_{front} of size $N{\times}M$. Let us denote the corners of the front imager by A, B, C and D. The 3D coordinates of the corners are known. We also denote the projection functions of the left and right cameras with F_L and F_R respectively. The virtual pixel $P_{i,j}(X, Y, Z)$ at location (i, j) is a linear combination of the points A, B, C, D

$$S_i = ((N-i)\cdot A + i\cdot D)/N \,, \ T_i = ((N-i)\cdot B + i\cdot C)/N \,, \ P_{i,j} = ((M-j)\cdot S_i + j\cdot T_i)/M \,. \tag{11}$$

We take into account the rotations R_L and R_R which transform the fisheye system into a canonical system. Therefore, the 3D point $P_{i,j}$ is rotated with the matrix R^T_L. Now we simply project it onto the original fisheye image: $F_L\left(R^T_L \cdot P_{i,j}\right)$. Finally, the intensity of the rectified pixel (i, j) is calculated by interpolating the fisheye image at the location $F_L\left(R^T_L \cdot P_{i,j}\right)$

$$I^L_{front}(i, j) = I_{fisheye}\left(F_L\left(R^T_L \cdot P_{i,j}\right)\right). \tag{12}$$

$I_{fisheye}$ represents the original fisheye image. For efficiency purposes, we store the mapping between the rectified coordinates and their corresponding points in the fisheye images in a look-up-table. The look-up-tables are computed only once, when the application is started. We have the following relationships:

$$LUT^L_{front}(i, j) = F_L\left(R^T_L \cdot P_{i,j}\right). \tag{13}$$

For the right image we have

$$LUT^R_{front}(i, j) = F_R\left(R^T_R \cdot P_{i,j}\right). \tag{14}$$

The final rectification equality is:

$$I^X_{front}(i, j) = I_{fisheye}\left(LUT^X_{front}(i, j)\right), X \in \{L, R\} \,. \tag{15}$$

In a similar manner we calculate the rectified images $I^X_{central}$ and I^X_{back}, $X \in \{L, R\}$. This process results three rectified left images and three rectified right images. An example of such images is depicted in Fig. 3c, e, g. We can apply the algorithms developed in the classical mono and stereo-vision for each pair of rectified images. In our implementation, we allow a slight overlap of the rectified images because the stereo matching algorithm we use does not provide disparities at the image borders.
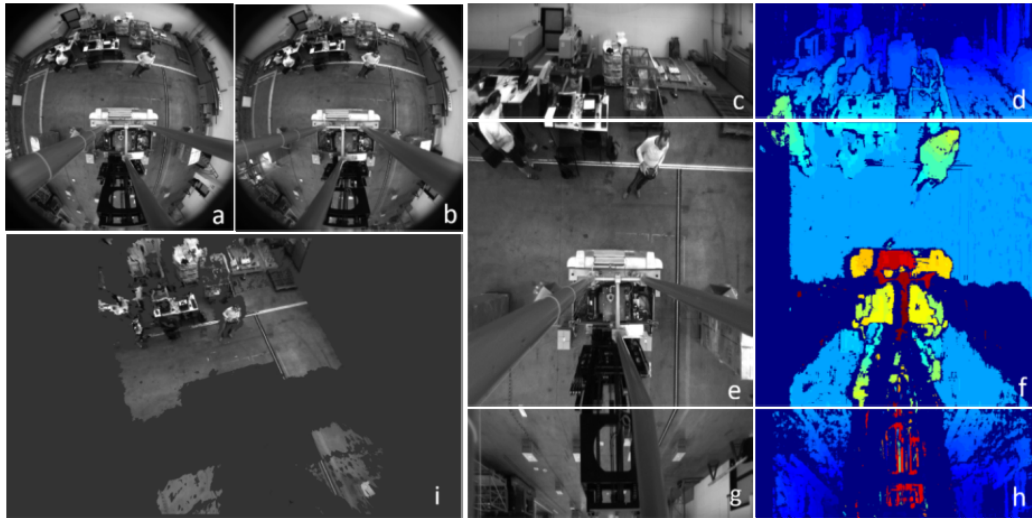
Fig. 3 – a) Left fisheye image; b) right fisheye image; c), e), g) the division of the fisheye image into: the front, central and back rectified images; d), f), h) the disparities for each left-right pair of rectified images; i) the common 3D view of all the channels.

5. CONCLUSIONS

In this work we have presented an omnidirectional stereo system based on fisheye lenses. Catadioptric systems, wide angle lenses and rectilinear lenses were also considered but fisheye lenses fit the requirements of this project better. They provide a higher angular resolution and light collecting capability, have a better resistance and they can cover a higher area around the AGV. The sensor is installed onto an automated forklift at 4m height, pointing to the ground. In this manner, a single sensor suffices for monitoring the area around the automated forklift. We have used an existing calibration toolbox [36] for the intrinsic calibration of the cameras. The relative rotation and translation of the right cameras w.r.t. the left camera is based on the 3D coordinates of the checkerboard pattern, returned by the intrinsic calibration. We have used a multi-channel rectification. The fisheye images were divided into three rectified perspective images. This rectification enables the use of a standard stereo matching algorithm for each rectified channel. The stereo reconstruction and obstacle detection using digital elevation maps are detailed in the second part of the paper [27].

ACKNOWLEDGMENTS

REFERENCES

1. NEDEVSCHI, S., MARITA, T., DANESCU, R., ONIGA, F., BOTA, S., HALLER, I., PANTILIE, C., DRULEA, M., GOLBAN, C,, *On-Board 6D Visual Sensor for Intersection Driving Assistance*, A Microsystems for Automotive Applications 2010 SE-**25**, Meyer, G., Valldorf, J., Eds., VDI-Buch, Springer Berlin Heidelberg, 2010, pp. 253–264.
2. GEHRIG, S., RABE, C., KRUEGER, L., *6D Vision Goes Fisheye for Intersection Assistance*, Proc. of CRV, 2008, pp. 34–41.
3. GEHRIG, S.K. *Large-field-of-view stereo for automotive applications*. Omnivis, **1**, 2005.
4. HAMIT, F. *New video and still cameras provide a global roaming viewpoint*, Adv. Imaging, **12**, *3*, 1997.
5. YAGI, Y., NISHII, W., *Rolling motion estimation for mobile robot by using omnidirectional image sensor hyperomnivision*, Proc. of 13th Int. Conf Pattern Recognition, **1**, 1996, pp. 946–950.
6. NAYAR, S.K., *Catadioptric omnidirectional camera*, Proc. of CVPR 1997, pp. 482–488.
7. BAKER, S., NAYAR, S.K., *A theory of catadioptric image formation*, Proc. of IEEE Int. Conf. on Computer Vision, 1998, pp. 35–42.
8. BAKER, S., NAYAR, S.K., *A theory of single-viewpoint catadioptric image formation*, Int. J. Comput. Vis., **35**, pp. 175–196, 1999.
9. SVOBODA, T., PAJDLA, T., HLAVÁČ, V., *Epipolar geometry for panoramic cameras*, Computer Vision (ECCV'98), Springer, 1998, pp. 218–231.
10. AIKIO, M., MAKINEN, J.-T., YANG, B., *Omnidirectional camera*, Proc. IEEE ICCP, 2013, pp. 217–221.

11. SOUTHWELL, D., BASU, A., FIALA, M., REYDA, J., *Panoramic stereo*, Proc. of IEEE Int. Conf. on Pattern Recognition, **1**, 1996, pp. 378–378.
12. YI, S., AHUJA, N., *An omnidirectional stereo vision system using a single camera*, Proc. of ICPR, **4**, 2006, pp. 861–865.
13. LI, W., LI, Y.F., *Single-camera panoramic stereo imaging system with a fisheye lens and a convex mirror*, Opt. Express, **19**, 2011, pp. 5855–67.
14. ZHU, J., HUMPHREYS, G., KOLLER, D., STEUART, S., WANG, R., *Fast omnidirectional 3D scene acquisition with an array of stereo cameras*, 3DIM 2007 – Proc. of 6th Int. Conf. on 3-D Digital Imaging and Modeling, 2007, pp. 217–224.
15. KANG, S. B., SZELISKI, R., *3-D Scene Data Recovery using Omnidirectional Multibaseline Stereo*, Proc. of CVPR'96, pp. 364–370.
16. PAN-Robots – Plug&Navigate Robots for Smart Factories www.pan-robots.eu.
17. ESTEFANIA, J. A., RAMOS, B., *Warehouse logistics from Casbega point of view*, Proc. of IEEE 9th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2013, pp. 195–202.
18. BEINSCHOB, P., REINKE, C., *Strategies for 3D data acquisition and mapping in large-scale modern warehouses*, Proc. of IEEE Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2013, pp. 229–234.
19. DIGANI, V., SABATTINI, L., SECCHI, C., FANTUZZI, C., *Towards decentralized coordination of multi robot systems in industrial environments: A hierarchical traffic control strategy*, Proc. of IEEE ICCP, 2013, pp. 209–215.
20. DIGANI, V., SABATTINI, L., SECCHI, C., FANTUZZI, C., *Hierarchical traffic control for partially decentralized coordination of multi AGV systems in industrial environments*, Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA), 2014, pp. 6144–6149.
21. SABATTINI, L., DIGANI, V., SECCHI, C., COTENA, G., RONZONI, D., FOPPOLI, M., OLEARI, F., *Technological roadmap to boost the introduction of AGVs in industrial applications*, Proc. of IEEE ICCP, 2013, pp. 203–208.
22. DIGANI, V., SABATTINI, L., SECCHI, C., FANTUZZI, C., *An automatic approach for the generation of the roadmap for multi-AGV systems in an industrial environment*, Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, 2014, pp. 1736–1741.
23. OLEARI, F., MAGNANI, M., RONZONI, D., SABATTINI, L., *Industrial AGVs: Toward a pervasive diffusion in modern factory warehouses*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 233–238.
24. BOEHNING, M., *Improving safety and efficiency of AGVs at warehouse black spots*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 245–249.
25. ARGA, R., NEDEVSCHI, S., *Vision-based autonomous load handling for automated guided vehicles*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 239–244.
26. MAYRA, A., AIKIO, M., KUMPULAINEN, M., *Fisheye optics for omnidirectional perception*, Proc. of IEEE Int. Conf. on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2014.
27. DRULEA, M., A. VATAVU, MANDICI, S., NEDEVSCHI, S., *An omnidirectional stereo system for logistic plants*, *Part 2*: *Stereo reconstruction and obstacle detection using digital elevation maps*, Proc. Rom. Acad., 2016.
28. DRULEA, M., SZAKATS, I., VATAVU, A., NEDEVSCHI, S., *Omnidirectional stereo vision using fisheye lenses*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 251–258.
29. KRUGER, L. E., WOHLER, C., WURZ-WESSEL, A., STEIN, F., *In-factory calibration of multiocular camera systems*, Photonics Europe, International Society for Optics and Photonics, 2004, pp. 126–137.
30. TOMASI, C., KANADE, T., *Shape and motion from image streams under orthography: a factorization method*, Int. J. Comput. Vis., **9**, 1992, pp. 137–154.
31. SHI, J., TOMASI, C., *Good features to track*, Proc. of CVPR'94, 1994, pp. 593–600.
32. BADINO, H., FRANKE, U., RABE, C., GEHRIG, S., *Stereo-vision based detection of moving objects under strong camera motion*, Int. Conf. on Computer Vision Theory and Applications, 2006, pp. 253–260.
33. DIGANI, V., CARAMASCHI, F., SABATTINI, L., SECCHI, C., FANTUZZI, C., *Obstacle avoidance for industrial AGVs*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 227–232.
34. CARDARELLI, E., SABATTINI, L., SECCHI, C., FANTUZZI, C., *Multisensor data fusion for obstacle detection in automated factory logistics*, Proc. of IEEE 10th Int. Conf. on Intelligent Computer Communication and Processing (ICCP), 2014, pp. 221–226.
35. *Fisheye lens* – Wikipedia http://en.wikipedia.org/wiki/Fisheye_lens.
36. KANNALA, J., BRANDT, S.S., *A Generic Camera Model and Calibration Method for Conventional, Wide-Angle, and Fish-Eye Lenses*, IEEE Trans. Pattern Anal. Mach. Intell., **28**, 2006, pp. 1335–1340.
37. KASSIR, A., PEYNOT, T., *Reliable Automatic Camera-Laser Calibration*. Australas. Conf. Robot. Autom. 2010.
38. *Matlab RADOCC Toolbox* http://www-personal.acfr.usyd.edu.au/akas9185/AutoCalib/AutoCamDoc/index.html.
39. *Jean-Yves Bouguet calibration toolbox* http://www.vision.caltech.edu/bouguetj/calib_doc/.
40. SHIH, S., TSAI, W.H., *A Two-Omni-Camera Stereo Vision System With an Automatic Adaptation Capability to Any System Setup for 3-D Vision Applications*, IEEE Transactions on Circuits and Systems for Video Technology, 2013.
41. YUAN, PH., YANG, KF., TSAI W.H., *Real-time security monitoring around a video surveillance vehicle with a pair of two-camera omni-imaging devices*, IEEE Transactions on Vehicular Technology, 2013.
42. GANDHI, T., TRIVEDI, M., *Vehicle surround capture: Survey of techniques and a novel omni-video-based approach for dynamic panoramic surround maps*, IEEE Transactions on Intelligent Transportation Systems, 2006.