Omnidirectional sensors for mobile robots

Sebastian Blumenthal[1], David Droeschel[1], Dirk Holz[1], Thorsten Linder[1], Peter Molitor[2], Hartmut Surmann[2]

[1] Department of Computer Science, University of Applied Sciences Bonn-Rhein-Sieg, St. Augustin, Germany
[2] Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS), St. Augustin, Germany

Abstract. This paper presents the different omnidirectional sensors developed at Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS) during the past decade. It envelops cheap USB / Firewire webcams to expensive high end cameras, continuously rotation 3D laser scanner, rotating infrared cameras and a new panoramic camera *Spherecam*. This new spherical camera is presented in detail compared to the other sensors. Beside the camera hardware we also categorize the omni cameras according to their price, connection and operating platform dependences and present the associated camera drivers and possible applications. Consider this paper as hardware descriptions of vision systems. Algorithms for ominidirectional vision systems are not discussed.

1 Introduction

Traditional imaging devices (photographic and video) are limited in their field of view. Existing devices typically use a photographic camera or a video camera, in conjunction with build in off-the-shelf lenses. The omnidirectional camera configuration allows to view the world from the center of projection of the lens in all directions (360° field of view). Such sensors have direct applications in a variety of fields, including, video surveillance [1], teleconferencing, computer vision, Simultaneous Localization And Mapping (SLAM) [2, 3] and virtual reality.

This paper presents the different omnidirectional sensors and the experiences made at our Institute during the past decade with a special focus to a new developed panoramic vision system *Spherecam*. The paper is organized at follows. Section 2 reviews the current state and know how of the different omnidirectional systems with a special focus to catadioptric systems. Therefore we categorize the system and with it the associated application according the prize (small, medium, high), connection (USB / Firewire) and supported operating system (Linux / Windows). Section 3 presents the new developed spherical camera and section 4 concludes the paper. Links to videos which show the behavior of the systems are referenced in the footnote at each section or subsection.

2 Omnidirectional vision system

Our basic catadioptric vision system consists of one camera mounted into a framework and aiming towards a hyperbolic mirror. This concept allows to make use of a varied range of different cameras (e.g. Quickcam pro 5000 / 9000, Apple iSight, AVT MARLIN) and is successfully in use on a variety of application [4, 5, 6]. Without the camera the complete system weight is about $300\,\mathrm{g}$ and has a dimension of $200 \times 100 \times 100$mm (L×W×H). The main component is the hyperbolic mirror (Fig. 2).

2.1 The hyperbolic mirror

The hyperbolic mirror is made of aluminum and is optionally coated with glass (see Fig. 2(a)). The standard hyperbolic geometry version has a diameter of $70\,\mathrm{mm}$, a pitch from $0°$ to $-30°$ and a mirror function of $\frac{z^2-x^2-y^2}{1000} = 1$. Since the mirror is produced at our laboratory we are able to implement different mirror functions. Therefore, the optimal mirror function of the catadioptric vision systems optimize the application benefit.

2.2 A cheap Firewire camera

An interesting alternative to USB web cameras is the Apple iSight Firewire camera ($\sim$ 200 US). It has a manual / auto focus and delivers pictures with 30 fps at a VGA resolution. Firewire cameras are supported by both operation systems. Linux use the ”libdc1394” library [7] and for Windows the ”CMU 1394” driver can be used [8]. Figure shows an extension of an iSight omni camera. The iSight is combined with a pan and tilt unit mounted underneath the mirror. Image data for searching and mapping appli-

(a) (b) (c)

Fig. 1. Different views of the iSight Firewire camera with a pan and tilt unit under the mirror

cations are obtained with 30 fps (VGA) while the camera is aiming towards the mirror. The focus is set manually to the mirror. After detecting Points Of Interests (POIs) the servos move the camera to that position to acquire undistorted pictures for example for object classification or recognition with a higher resolution then the omni picture. During pan and tilt movements the auto focus is helpful. A video which shows the iSight and the pan and tilt unit can be found at `http://de.youtube.com/watch?v=SPTIpnQw0eA`.

2.3 A high resolution omnidirectional vision system

This system consists of a high resolution IEEE1394 firewire camera aiming towards a hyperbolic mirror. The ”AVT Marlin F-145-C” camera is equipped with a ”Fujinon DF6HA-1B” object lens. This fixed focal lens comes with a focal length of 6mm and an aperture between $f/1.2$ and $f/16$. The minimum object distance is about 0.1 meters and the maximum field of view is $56°$ in horizontal and $43°$ in vertical alignment. The camera is able to deliver up to 10 fps in high resolution mode (1392×1038 pixels) if DMA transmission and Format_7 is used. Several image formats like RGB, RAW and YUV are build in. Drivers for Windows are available (e.g. CMU1394 library [8]). In Linux the library ”libdc1394” delivers the full functionality [7].

If mounted in the IAIS-Vision system, we achieve circular shaped cut-out images from the mirror with a diameter of about 1000 pixels[3]. This system delivers the capabilities of capturing image data for several illumination conditions, because the camera is able to auto adjust the shutter speed and other parameters like white balance and gain control. Additionally the aperture can be adjusted manually between the described values.

In comparision to the cheap iSight camera system, the images from the Marlin yield advantages regarding the detection of features and for visual bearing only SLAM due its higher resolution. But the computational costs increase proportionally with the number of pixels. Depenting on a specific application those properties must be taken in to account.

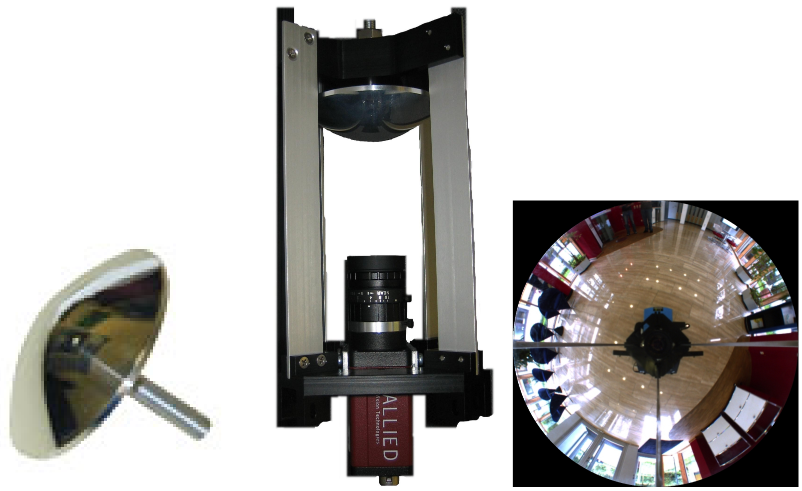
2.4 Consumer digital cameras

Beside the webcams, traditional consumer cameras are interesting for robotic applications. They have a high quality and are cheap compared with other robotic vision systems. High resolution images from 5 - 12 million pixels allow applications like environmental OCR e.g. reading door plates. Figure 3(a) shows an omnidirectional picture made with a Canon A620. The camera is connected via USB 2.0 to the robot. The operating systems Linux and Windows are supported. The *libptp* library is used for Linux and the remote capture software from Canon is used for Windows [9, 10]. The fast *view finder* mode supports a frame size of 320×160 pixels with 30 fps which is similar to the camera display resolution. When a mobile robot reaches a POI for example a door then high resolution images are taken. A high resolution image with 7.2 million pixels takes 5 seconds to be transferred to the robot. Figure 3(b) shows a canon IXUS 400 together with an infrared camera on a pan and tilt unit.[4]

The recently published software extension CHDK (Canon Hacker's Development Kit) further increase the use of standard digital cameras [11]. The CHDK is a firmware enhancement that operates on various models of Canon cameras. It provides additional functionality for example for experienced photographers beyond that currently provided by the native camera firmware. Beside the enhanced ways of recording images or additional photographic settings, the CHDK can run program scripts written in a

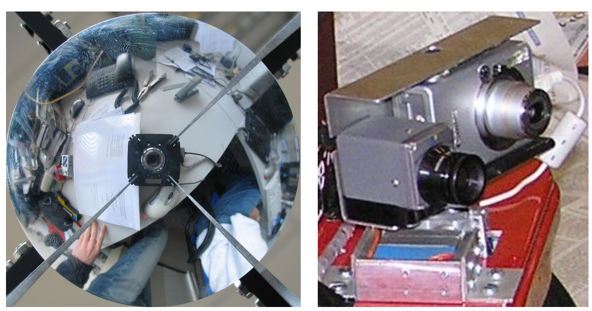
[3] `http://de.youtube.com/watch?v=up0ky2S9V04`

[4] See `http://de.youtube.com/watch?v=tAEV1lhRpVQ` for a video.



(a) (b) (c)

Fig. 2. (a) The hyperbolic mirror in standard geometry. (b) The high resolution IEEE1394 camera AVT-marlin camera with an omnidirectional mirror and (c) Example image taken with the AVT in the foyer of the Fraunhofer IAIS at campus Castle Birlinghofen.



(a) (b)

Fig. 3. (a) Omni picture made with the consumer digital camera Cananon A620. (b) Combined infrared camera A10 (Flir) and a consumer digital camera (IXUS 400) at a continously turning pan and tilt unit.

micro-version of the BASIC language for example for motion detection or interval self-triggering.

A great benefit from very high resolution images is the use for text recognition in particular for extracting information on door plates in office environments. This enables to enritch maps with semantic information.

2.5 Continuously rotating 3D Laser scanner

For building metric maps and omnidirectional pictures, a continuously rotating 3D laser scanner (e.g. 3DLS-K2) is appropriate (Fig. 4(a)) [12]. Beside the 3D point cloud (omnidirectionl distance values), the laser scanner delivers the amount of reflected light (remission values) Fig. 4(b)). This leads to black/white textured maps. The 3DLS-K2

(a)

(b)

Fig. 4. (a) The continuously rotating laser scanner 3DLS-K2 consists of a rotation unit with 2 SICK-LMS laser scanner. (b) Remission image of castle Birlinghoven

consist of a rotation unit to reach the 3^{rd} dimension and of two 2D time-of-flight SICK laser scanner (scanning range 180^{o}). They can provide distance measurement in a range of up to 80m with a centimeter accuracy (see Fig. 5). The continuously yawing angle scan scheme [13] of the unit generates two 360 degree dense point clouds every 1.2 seconds. The hundreds of 3D point clouds acquired during a robot run are registered into a common coordinate system with the well known Iterative Closest Point (ICP) algorithm described in several previously published papers[5] (e.g. [14, 15]). According to our categories the laser scanner is a very expensive omnidirectional sensor with an update frequency of 1.2 fps. The scanner is operating system independent and needs 3 USB connections.

The remission images in combination with the depth information are useful to build metric maps. In particular it is possible to place the remission images as textues in a virtual 3D map.

[5] See `http://de.youtube.com/watch?v=xr28pX9ZkXw`

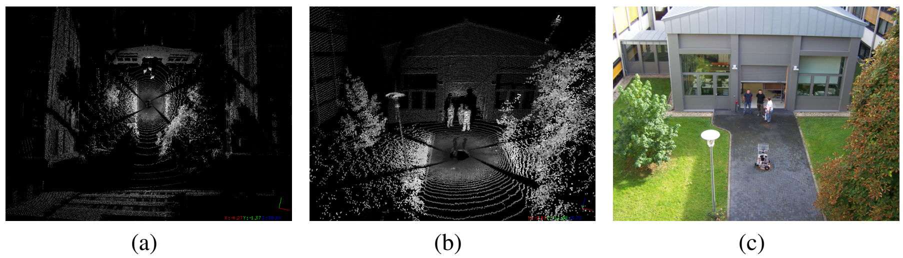


Fig. 5. (a) Top view of a point cloud acquired with the IAIS 3DLS-K2 (points are colorized according to the measured remission value), (b) a detailed view and (c) a photo showing the scene in (b).

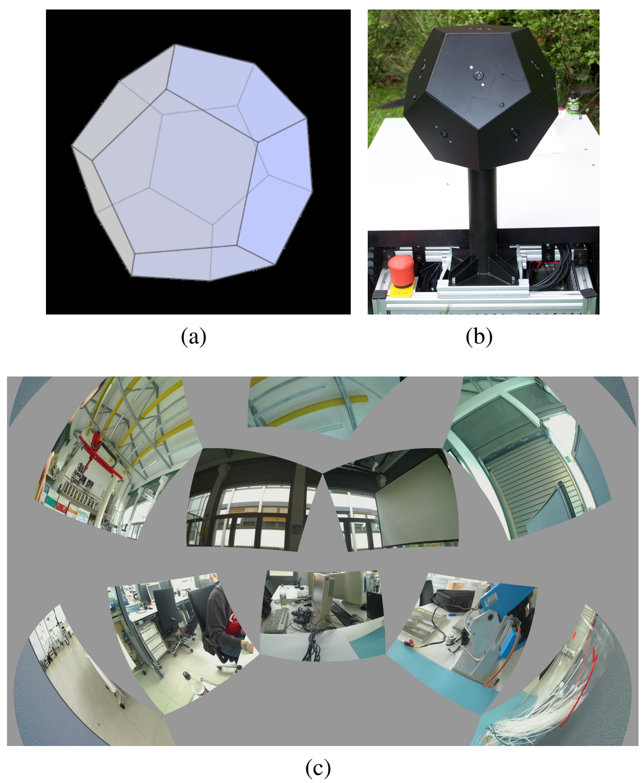


Fig. 6. Panoramic visions system. (a) Scheme of a penta-dodecahedron. [16] (b) The dodecahedron shaped camera system with eleven cameras. Each aims in another direction. (c) demonstrates an impression from panoramic images of the dodecahedron shaped camera.

3 Panoramic vision system *Spherecam*

For visual inspection and SLAM applications a new vision system *Spherecam* is presented here. It is build up from eleven off-the-shelf USB *Logitech QuickCam 9000* web cameras aiming in different directions. They are mounted in penta-dodecahedron shaped chassis with a size of $220 \times 220 \times 380$mm (L×W×H). (see Fig. 6(a) and 6(b)). Each camera delivers up to 15 frames per second at a resolution of 800×600 pixels. At a lower framerate (5 fps), pictures with a resolution of 1600×1200 pixels can be acquired. All eleven cameras are connected to one up to four computers (e.g. MacMinis) via USB 2.0 according to the desired computational power. Complete scenes of eleven images with VGA resolution can be acquired with a frame rate of approx. 3 fps due to the serial readout of the cameras. More computers speed up the frame rate but require a precise synchronization. The main advantages of the Spherecam are its inherent robustness since it has no mechanical moving units and its high resolution undistorted images which allow direct feature detection at the single images pointing in an interesting direction.[6]

Dependant of the field of view of the used cameras a near panoramic overview of the environment can be obtained. Fig. 6(c) shows a panoramic (fisheye) image obtained from all cameras.

The USB video device class or USB video class (UVC) is a special USB class which defines functionality for video streaming devices like web cams camcorders or television tuners over the Universal Serial Bus. The *Logitech QuickCam Pro 9000* web cam is an UVC compliant camera device. Drivers for the Windows operating systems are shipped with the camera. A Linux implementation of the UVC standard can be found at [17].

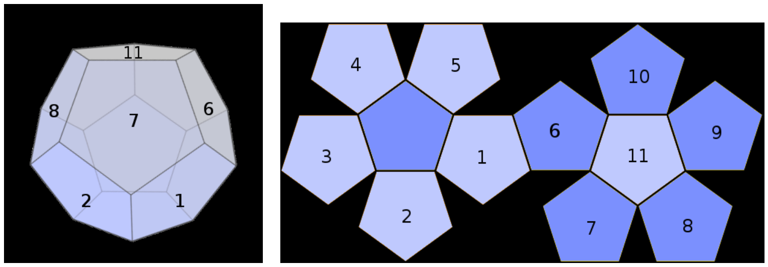
The proposed vision system has a fixed numbering strategy. One camera from the lower row is chosen as starting point and it is assigned the first index. The next four clockwise cameras (seen from a top view perspective) are labeled in ascending order. The first camera of the upper row which is directly left hand from the starting camera is assigned the sixth index. The other cameras of the upper row are aligned assigned with ascending numbers clockwise. The camera which is directly aligned to the top gets the last index. The remaining bottom surface is reserved for the mounting parts. Fig. 7 sketches the numbering.

Beside the camera numbering the exact position and orientation of each camera must be known. The camera lenses are located in the middle of each pentagon. The height h of a pantagon is characterized by

$$h = \frac{a}{2}\sqrt{5 + 2\sqrt{5}} \quad [18] \tag{1}$$

where a is the length of one vertices. The coordinate frame for the whole *Spherecam* camera system has its origin in the center of the inscribed sphere of the dodecahedron.

[6] `http://de.youtube.com/watch?v=MkkqP5Ycq0g`



(a) (b)

Fig. 7. Camera numbering for the panoramic vision system. (a) Scheme of the numbering of the penta-dodecahedron. (b) The same scheme displayed on an unfolded dodecahedron. [16]

The inner radius r of this sphere can be determined by:

$$r = \frac{a}{4}\sqrt{\frac{50 + 22\sqrt{5}}{5}} \quad [19] \tag{2}$$

For each camera alignment pitch, yaw and roll values are needed. The **pitch value** α_{pitch} can be deduced as follows. A slice plane is positioned as shown in Fig. 8(a). The resulting Fig. is depicted in 8(b). It consists of four heights which touch the inscribed sphere and two vertices. The dihedral angle (the angle between two surfaces) respectively the angle between two heights γ is given by:

$$\gamma = \frac{1}{2}\ arccos(-\frac{1}{5}\sqrt{5}) \quad [20] \tag{3}$$

In Fig. 8(b) the angle ε is exactly half of the dihedral angle

$$\varepsilon = \frac{1}{2}\gamma \tag{4}$$

where the mirror axis x is the bisecting line of γ.

The angle β which is opposite to ε in the right angle triangle and can be easily calculated:

$$\beta = 180° - 90° - \varepsilon = 90° - \varepsilon \tag{5}$$

The horizontal dotted line in Fig. 8(b) indicates a base line which is parallel to the mounting surface of the vision system. r is perpendicular to this baseline b and reaches from the center to the middle point of the height h of the upper pentagon surface. Due to the axis of symmetry x the segment r' also reaches from the center to the middle of the adjacent pentagon surface, exactly where the camera lens is located. r and b are perpendicular. Hence this perpendicular angle consists of two times β and one times α_{pitch}.

$$\alpha_{pitch} = 90° - 2\beta \tag{6}$$

$$\alpha_{pitch} = 90^\circ - 2\,(90^\circ - \varepsilon) = -90^\circ + 2\varepsilon \quad using\ \ (5) \tag{7}$$

$$\alpha_{pitch} = -90^\circ + 2\,(\frac{1}{2}\gamma) = -90^\circ + \gamma \quad using\ \ (4) \tag{8}$$

$$\alpha_{pitch} = -90^\circ + \frac{1}{2}\ \ arccos(-\frac{1}{5}\sqrt{5}) = 26.6^\circ \quad using\ \ (3) \tag{9}$$

For the **yaw angles** the symmetry of the dodecahedron is used. Seen from the top view five cameras in one horizontal plane are equidistant distributed (see Fig. 8(c)) and the angle between two adjacent cameras is:

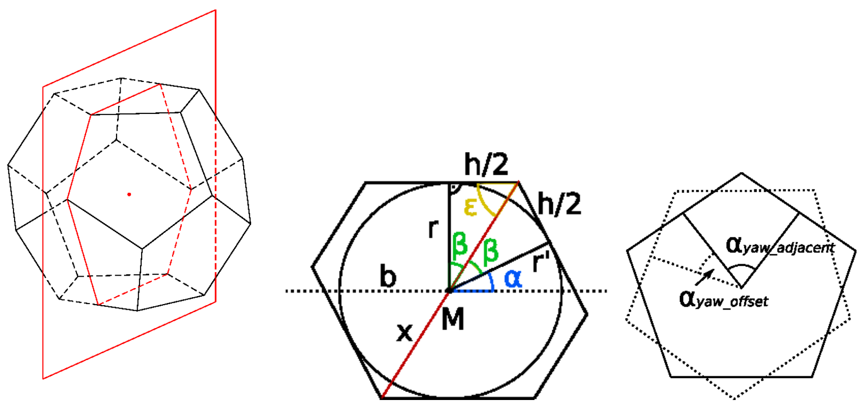
$$\alpha_{yaw\_adjacent} = \frac{360^\circ}{5} = 72^\circ \tag{10}$$

The shift between the upper and the lower row is half this angle.

$$\alpha_{yaw\_offset} = \frac{alpha_{yaw\_adjacent}}{2} = \frac{72^\circ}{2} = 36^\circ \tag{11}$$

The **roll angle** for each camera is equal to zero because the rectangular field of view of the cameras are parallel aligned to the base line.

$$\alpha_{roll} = 0^\circ \tag{12}$$



(a) (b) (c)

Fig. 8. Geometry of a penta-dodecahedron. (a) Slice plane through the dodecahedron. [18] (b) Dodecahedron with angles α, β and ε. x denotes the axis of symmetry. (c) Yaw angles seen from top perspective.

With the above considerations it is possible to set up all pitch, yaw and roll angles for all cameras in the vision system. With respect to the numbering strategy camera number one is defined with a pitch value equals zero $\alpha_{yaw\_start} = 0$. It belongs to the lower row, and therefore the yaw angle is $-\alpha_{pitch} = -26,6^\circ$. Roll is equal to zero.

The second camera has the same roll and pitch but yaw is increased by $\alpha_{yaw\_adjacent}$ therefore yaw is $72°$. For the third camera yaw is again increased by α_{yaw} and so on until camera five.

The sixth camera belongs to upper row and has a pitch of $\alpha_{pitch} = 26,6°$. The difference regarding yaw to the first camera is equal to $\alpha_{yaw\_offset} = 36°$. The next cameras until the tenth add again each up $\alpha_{yaw\_adjacent}$ for yaw.

The last camera which is mounted on the top of the vision system is directly heading upwards. Therefore the pitch angle is perpendicular to the base line: $\alpha_{pitch\_\#11} = 90°$. The yaw angle was determined experimentally and has $152°$ Table 1 summarizes all pitch, yaw and roll angles for each camera.

The distance of the displacement of each camera is described by the radius of the inscribed sphere. The length of one vertices a of the here used vision system is $75\,\mathrm{mm}$. Substituted in Equation (2) r is equal to $83.51\,\mathrm{mm}$.

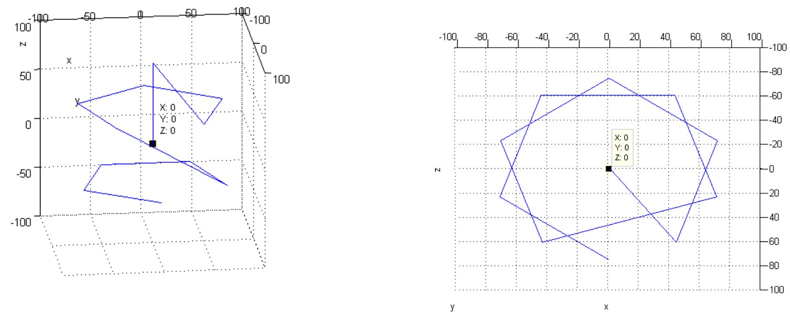
Number	Pitch	Yaw	Roll
1	-26.6	0	0
2	-26.6	72	0
3	-26.6	144	0
4	-26.6	216	0
5	-26.6	288	0
6	26.6	36	0
7	26.6	108	0
8	26.6	180	0
9	26.6	256	0
10	26.6	324	0
11	90	152	0

Table 1. Pitch, yaw and roll angles for each camera in the *Spherecam* vision system. All angles are given in degrees.

With the pitch, yaw and roll angles and the inner radius it is possible to deduce *homogeneous transformation matrices* [21] for each camera with respect to the center of the vision system. With those matrices it is easily possible to transform coordinates from a camera coordinate system into a global system for the *Spherecam*. The used coordinate system follows the Left Hand System (LHS) rule. X and y denote the width and height whereas z is the depth relative to the x-y-plane.

A transformation matrix to a specific camera coordinate frame is created by first setting up an auxiliary transformation matrix ${}^{Caux}_{G}T$ with respect to pitch, yaw and roll in the origin of the global vision system frame (no displacement). Then the distance of the inner radius is traveled along the z axis to get point ${}^{Caux}P$. The transformation matrix is inverted to transform the point ${}^{Caux}P$ from the auxiliary camera coordinate frame into the global frame: ${}^{G}P = {}^{G}_{Caux}T^{-1} \quad {}^{Caux}P$. ${}^{G}P$ describes the displacement vector to the origin of a camera coordinate frame. The values for the displacement in

${}^{G}_{Caux}T$ are replaced by the values of the calculated displacement vector ${}^{G}P$ to form the homogeneous transformation matrix ${}^{C}_{G}T$ for one camera frame. A plot of the calculated displacement vectors is depicted in Fig. 9.



(a) (b)

Fig. 9. Plotted displacement vectors. (a) Displacement vectors connected with vertices in ascending order (ends in origin). (b) Same plot seen from top perspective.

This panoramic vision system is adequate for visual inspection, surveillance, exploration and visual SLAM applications. It is still under developement and comparision of performance regarding visual SLAM is open to future work.

4 Conclusion

We have presented a variety of vision and range sensors especially designed for the application on mobile robot platforms and for a wide variety of tasks and application spaces, like for instance

- catadioptric omnidirectional camera systems for obtaining a hemispherical view,
- continuously rotating 3D laser scanners for digitalizing and reconstructing real world scenarios, buildings etc. or
- the *Spherecam* to get a near panoramic view of a robot's surroundings.

The drivers for all of the sensors are available for Unix and Windows machines. Cheap solutions with webcams as well as high end solution with expensive cameras are realized. Especially the new *Spherecam* provides new interesting possibilities because it delivers undistorted high resolution images as well omni images at the same time. It is more robust than solutions with pan and tilt or rotation units.

Acknowledgments

The authors would like to thank our Fraunhofer colleagues for their support namely Viatcheslav Tretyakov, Reiner Frings, Jochen Winzer, Stefan May, Paul Plöger and Kai Pervölz. This work was partly supported by the DFG grant CH 74/9-1 (SPP 1125).

References

[1] Murphy, R.: Trial by fire [rescue robots]. Robotics & Automation Magazine, IEEE **11**(3) (Sept. 2004) 50–61

[2] Durrant-Whyte, H., Bailey, T.: Simultaneous localization and mapping: part I. In: IEEE Robotics & Automation Magazine. Volume 13. (2006)

[3] Bailey, T., Durrant-Whyte, H.: Simultaneous localization and mapping (slam): part II. In: IEEE Robotics & Automation Magazine. Volume 13. (2006)

[4] Bredenfeld, A., Wisspeintner, T.: Rapid Prototyping für mobile Roboter. Elektronik-Informationen **12** (2004) 68–71 ISSN: 0343-6675.

[5] Calabrese, F., Indiveri, G.: An Omni-Vision Triangulation-Like Approach to Mobile Robot Localization. In: Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation. IEEE (June 2005) 604–609 ISBN: 0-7803-8936-0.

[6] Wisspeintner, T., Bose, A.: The VolksBot Concept - Rapid Prototyping for real-life Applications in Mobile Robotics. it **5** (May 2005) 274 – 281

[7] Douxchamps, D., Peters, G., Urmson, C., Dennedy, D., Moore, D.C., Lamorie, J., Moos, P., Ronneberger, O., Antoniac, D.P., Evers, T.: 1394-based dc control library. http://sourceforge.net/projects/libdc1394/ (September 2008)

[8] Baker, C.: CMU 1394 digital camera driver. http://www.cs.cmu.edu/ iwan/1394/ (September 2008)

[9] Danecek, P.: Canon capture. http://capture.sourceforge.net/ (September 2008)

[10] Canon Europe Ltd: Remote capture. http://software.canon-europe.com/software/0019449.asp?model= (September 2008)

[11] CHDK project group: CHDK Wiki. http://chdk.wikia.com/wiki/Main_Page

[12] Fraunhofer Institut Intelligent Analyse- und Informationssysteme (IAIS): Volksbot. `http://www.3d-scanner.net/` (2008)

[13] Wulf, O., Wagner, B.: Fast 3D-Scanning Methods for Laser Measurement Systems. In: International Conference on Control Systems and Computer Science (CSCS14). (2003)

[14] Nüchter, A., Lingemann, K., und Hartmut Surmann., J.H.: 6D SLAM for 3D Mapping Outdoor Environments. Journal of Field Robotics (JFR) Special Issue on Quantitative Performance Evaluation of Robotic and Intelligent Systems **24** (2007) 699 – 722

[15] Surmann, H., Nüchter, A., Lingemann, K., Hertzberg, J.: 6D SLAM - Preliminary Report on closing the loop in Six Dimensions. In: Proc. of the 5th IFAC Symposium on Intelligent Autonomous Vehicles. (June 2004)

[16] Wikimedia Commons: Dodecahedra. http://commons.wikimedia.org/wiki/Category:Dodecahedra (September 2008)

[17] Linux-UVC project: Linux UVC driver and tools. http://linux-uvc.berlios.de/ (August 2008)

[18] Fendt, W.: Das Dodekaeder. http://www.walter-fendt.de/math/geo/dodekaeder.pdf (February 2005)

[19] Gullberg, J.: Mathematics: From the Birth of Numbers. W W Norton & Co (April 1997)

[20] Köller, J.: Pentagondodekaeder. http://www.mathematische-basteleien.de/pentagondodekaeder.htm (September 2008)

[21] Craig, J.J.: Introduction to Robotics - Mechanics and Control. Addison-Wesley (1989)