Image and Vision Computing 27 (2009) 141–152

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

# Graph cut based panoramic 3D modeling and ground truth comparison with a mobile platform – The Wägele ☆

Sven Fleck *, Florian Busch, Peter Biber, Wolfgang Straßer

*WSI/GRIS, University of Tuebingen, Sand 14, Tuebingen, Germany*

ARTICLE INFO

*Article history:*
Received 19 January 2007
Received in revised form 13 April 2008
Accepted 9 May 2008

*Keywords:*
Graph cut
3D model acquisition
3DTV

ABSTRACT

Efficient and comfortable acquisition of large 3D scenes is an important topic for many current and future applications in the field of robotics, factory and office visualization, 3DTV and cultural heritage.

In this paper we present both an omnidirectional stereo vision approach for 3D modeling based on graph cut techniques and also a new mobile 3D model acquisition platform where it is employed. The platform comprises a panoramic camera and a 2D laser range scanner for self localization by scan matching. 3D models are acquired just by moving the platform around and recording images in regular intervals. Additionally, we concurrently build 3D models using two supplementary laser range scanners. This enables the investigation of the stereo algorithm's quality by comparing it with the laser scanner based 3D model as ground truth. This offers a more objective point of view on the achieved 3D model quality.


## 1. Introduction

High quality 3D modeling of real world environments is a hot research topic, as many emerging applications rely on the availability of such models. In mobile robotics, both surveillance of the robot's operation in the environment and path planing utilizing these models are desired applications. The visualization of office buildings, factory or safety critical environments like airports or power plants is another important class of applications. Upcoming 3D television (3DTV) also requires visually convincing 3D models.

Stereo vision is one major approach to 3D modeling relying solely on the visual modality. We present a graph cut based approach to omnidirectional stereo vision. In conjunction with this method, we designed a mobile sensor platform (called "The Wägele") for this purpose which we will first describe in Section 4. The scene is sampled by moving this platform around and acquiring several omnidirectional images in conjunction with their poses. The result of this whole process is a colored point cloud. The poses and thus the according external camera parameters are determined by matching of laser scans. However, the stereo vision algorithm does not rely on this laser range data directly, it is solely used for localization, thus any other method of localization would do.

☆ A Wägele – Swabian for a little cart.

* Corresponding author. Tel.: +49 70712978965.
*E-mail addresses:* fleck@gris.uni-tuebingen.de (S. Fleck), busch@gris.uni-tuebingen.de (F. Busch), biber@gris.uni-tuebingen.de (P. Biber), strasser@gris.uni-tuebingen.de (W. Straßer).

0262-8856/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2008.05.009

## 2. Related work and contributions

In the following, an overview of work related to omnidirectional image acquisition and to stereo vision is given, followed by a brief summary of the contributions and structure of this article.

### 2.1. Related work – omnidirectional cameras

Benosman and Kang give an comprehensive overview of panoramic vision [1]: Panoramic image capturing systems, the theory involved in the imaging process, algorithmic approaches for creating panoramic images, and applications based on panoramic images are covered. Three hundred and sixty degrees can be achieved using either a rotating camera together with mosaicing, an array of cameras (e.g., Ladybug 2 [2]), a fisheye lens or a mirror (e.g., [3]). An overview is given by Nayar [4] who especially contributed to the field of the mirror based (catadioptric) approach, e.g., [5] describes optimizing mirror design. A survey of omnistereo imaging is given by Peleg et al. [6] where the technical realization to obtain omnidirectional images is covered for the purpose of stereo panorama pairs. Besides mosaicing techniques for a rotating camera two custom systems with no moving parts are presented that are capable of capturing a stereo panorama feed: One system is based on spiral mirrors, the other system uses spiral lenses. Daniilidis illustrates most of the hardware setups available today [7].

### 2.2. Stereo vision – overview and related work

To review stereo algorithms, geometrical aspects of N-view and omnidirectional vision, [8,9] and [10] give good starting points, respectively. Stereo algorithms can be classified into four method families [8]: *Local* methods (like traditional SSD methods), *cooperative* methods, *dynamic programming* methods and *global* methods. The *graph cut* methods which were introduced by Roy and Cox [11] belong to the class of global methods. Experimental comparative studies show, that graph cut based methods perform better than simulated annealing methods [8] and show approximately four times smaller errors than traditional methods based on normalized correlation [12]. Due to their impressive performance, many recent publications cover graph cut [13–17]. Other work in point based omnidirectional reconstruction on mobile robotics can be found, e.g., in [18], although this is not based on graph cut. There is also plenty of work in the realtime reconstruction area, i.e., [19]. However, we address high quality, not realtime in the first place.

### 2.3. Contributions and paper overview

The work presented within this article is based on [20], which finds its roots in [21,35]. According to the knowledge of the authors it describes the first approach on omnidirectional stereo based on graph cuts. First, it was introduced on a PeopleBot mobile robot platform, in [20] the Wägele platform has been designed that allows for both higher quality models due to better sensors, no need for odometry information and last but not least, the possibility of acquiring approximative ground truth 3D information based on additional laser range scanners. The proposed omnidirectional graph cut stereo approach is referenced by Lhuillier [22] at CVPR 2007 and He et al. at OMNIVIS 2007 [23], Weiss and Zell at AMS 2005 [24], Skraba and Guibas at DCOSS 2007 [25], and Ariacan and Frossard at AVSS 2007 [26]. Ariacan and Frossard have proposed an extension by working directly in the spherical domain.

The first two following sections give an overview of the presented approach: Section 3 covers the model acquisition process, Section 4 describes the Wägele sensor platform. Afterwards, the components of the system are detailed. Section 5 gives a brief overview of the scan matching and localization method. Afterwards, Section 6 summarizes the graph cut algorithm itself. In Section 7 some renderings of the resulting 3D models are shown, so the reader can get an impression of the achieved quality we currently obtain. We also evaluate how sensitive the results react to varying parameters of the graph cut's energy function. In Section 8 the stereo vision results are compared to "ground truth", which is also acquired using additional laser range scanners on the Wägele platform. A surveillance application where the acquired 3D models are utilized to embed results of a distributed network of cameras is presented in Section 9 before we conclude this paper Fig. 1.

## 3. Model acquisition flow

The flexible mapping between sensors and 3D information (localization, geometry and texture) is illustrated in Fig. 3. To give an overview, both the graph cut stereo pipeline and the laser scanner based modeling pipeline are illustrated in Fig. 2. The first step is the localization using the laser scanner mounted horizontally on the Wägele platform. Both the stereo matching pipeline and the laser scanner pipeline require an exact localization. The stereo pipeline also takes three images acquired from the omnidirectional camera mounted on top of the platform and then performs stereo matching, graph cut and further post processing steps to produce a final point cloud. The laser scanner pipeline uses the positions from the localization unit to convert the data from the other two laser scanners to absolute 3D positions and then texturizes them using one image from the omnidirectional camera to produce its final point cloud which is considered "ground truth" in this paper. Details of both pipelines are described in the following.

| | L1 | L2 | L3 | C1 |
|---|---|---|---|---|
| Loc. | X | | | |
| Geo. | | | | X |
| Text. | | | | X |

| | L1 | L2 | L3 | C1 |
|---|---|---|---|---|
| Loc. | X | | | |
| Geo. | | X | X | |
| Text. | | | | X |

**Fig. 1.** Left, Stereo vision flow; Right, Laser flow for "ground truth" acquisition.

## 4. Overview of the mobile sensor platform and acquisition flow

Besides the algorithmical aspect, we designed a mobile platform that offers a very flexible sensor setup. We concentrate on the sensor aspect, for now the platform is moved around manually to acquire sensor data. However, it could also be mounted on a mobile robot to acquire environments in a teleoperated way or on a pickup truck for outdoor acquisitions. In contrast to [21] our platform does not rely on odometry data and is thus a self-contained unit that aims at delivering high quality results.

The process of acquiring a 3D model with the Wägele platform is depicted in Fig. 3. After a recording session (online phase – the so called "Wägele run") the collected data is assembled to create a consistent 3D model in an automated offline processing step. First a 2D map of the scene is built and all scans of the localization scanner are matched to this map. After this, the position and orientation of the Wägele is computed at each time step. This data is then fed into the graph cut stereo pipeline.

### 4.1. Sensor platform

The sensor acquisition platform is built using aluminium XC $44 \times 44$ mm and $88 \times 44$ mm structural beams. This combines great stiffness with limited weight and allows for a very flexible setup. So it can be easily adapted to various sensor and scene requirements. Currently it comprises both an omnidirectional camera and three 2D laser range scanners in a stiff and self contained unit. The power supply consists of 24 NiMH cells with 1.2 V/3000 mAh each and some voltage regulators (5, 6, 24 V). It is mounted by just three screws on the basic cart for easy transportation and so composes "The Wägele" platform. See Fig. 4 for an overview.

### 4.2. Flexibility of the platform

An important goal of our project is to allow 3D scanning in various scenarios. For that, the mobile platform is designed for maximal flexibility. The setup can be changed quickly and more sensors can be mounted. Fig. 5 shows the actual setup with the three laser scanners and the panoramic camera.

### 4.3. Omnidirectional camera

To achieve high quality images with a panoramic mirror a high spatial resolution is necessary. We utilize a Canon EOS-20D SLR camera with 8Mpixels ($3456 \times 2304$), mounted vertically. Its appealing properties include the very low noise CMOS sensor, the excellent optics and its great speed. It is accessible via USB2.0 by the Canon camera API. For our acquisitions, we use the manual mode, where the highest possible aperture value