Medical Image Analysis 24 (2015) 220–234

Contents lists available at ScienceDirect

Medical Image Analysis

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

Multi-sensor super-resolution for hybrid range imaging with application to 3-D endoscopy and open surgery

Thomas Köhler[a,b,*], Sven Haase[a], Sebastian Bauer[a], Jakob Wasza[a], Thomas Kilgus[c], Lena Maier-Hein[c], Christian Stock[d], Joachim Hornegger[a,b], Hubertus Feußner[e]

[a] *Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany*
[b] *Erlangen Graduate School in Advanced Optical Technologies (SAOT), Germany*
[c] *Division of Medical and Biological Informatics Junior Group: Computer-assisted Interventions, German Cancer Research Center (DKFZ) Heidelberg, Germany*
[d] *Institute of Medical Biometry and Informatics, University of Heidelberg, Germany*
[e] *Research Group Minimally-invasive interdisciplinary therapeutical intervention, Klinikum rechts der Isar of the Technical University Munich, Germany*

ARTICLE INFO

Article history:
Received 23 July 2013
Revised 27 January 2015
Accepted 25 June 2015
Available online 3 July 2015

Keywords:
Hybrid range imaging
Sensor fusion
Super-resolution
3-D endoscopy
Open surgery

ABSTRACT

In this paper, we propose a multi-sensor super-resolution framework for hybrid imaging to super-resolve data from one modality by taking advantage of additional guidance images of a complementary modality. This concept is applied to hybrid 3-D range imaging in image-guided surgery, where high-quality photometric data is exploited to enhance range images of low spatial resolution. We formulate super-resolution based on the maximum a-posteriori (MAP) principle and reconstruct high-resolution range data from multiple low-resolution frames and complementary photometric information. Robust motion estimation as required for super-resolution is performed on photometric data to derive displacement fields of subpixel accuracy for the associated range images. For improved reconstruction of depth discontinuities, a novel adaptive regularizer exploiting correlations between both modalities is embedded to MAP estimation. We evaluated our method on synthetic data as well as ex-vivo images in open surgery and endoscopy. The proposed multi-sensor framework improves the peak signal-to-noise ratio by 2 dB and structural similarity by 0.03 on average compared to conventional single-sensor approaches. In ex-vivo experiments on porcine organs, our method achieves substantial improvements in terms of depth discontinuity reconstruction.

1. Introduction

Hybrid imaging is an emerging field of research in medical imaging describing the fusion of different modalities. From a general perspective, sensor fusion enables an augmented representation of complementary information. Common setups are the combination of positron emission tomography (PET) in nuclear medicine with computed tomography (CT) or magnet resonance imaging (MRI) to visualize metabolism and internal structures simultaneously. A novel hybrid imaging setup addressed in this paper combines range imaging (RI) technologies with RGB sensors to augment 3-D range data with photometric information.

1.1. Applications

In recent years, RI has been proposed for several applications in healthcare (Bauer et al., 2013). In this work, we examine two different applications of hybrid RI in the field of image-guided surgery.

In terms of minimally invasive procedures, various approaches to gain intra-operative range data have been introduced and evaluated with respect to their clinical usability (Maier-Hein et al., 2014). Stereoscopy (Field et al., 2009) has been proposed as a passive technique to capture 3-D range data for interventional imaging. On the other hand, active sensor technologies based on structured light (SL) (Schmalz et al., 2012) or Time-of-Flight (ToF) (Penne et al., 2009) have been examined. These sensors can be augmented with photometric information to enable hybrid RI within one single endoscope, e.g. to facilitate ToF/RGB endoscopy (Haase et al., 2013a). This fusion of complementary modalities provides the surgeon a comprehensive view of a scene including 2-D and 3-D information. In addition, sensor fusion is beneficial to enhance robustness and reliability of many image processing tasks, e.g. for localization and tracking of laparoscopic instruments (Haase et al., 2013b).

In open surgery, one common workflow is to register pre-operative 3-D planning data acquired, e.g. by CT, with intra-operative range data gained by means of marker-less RI technologies

* Corresponding author. Tel.: +49 9131 85-27775; fax: +49 9131 303811.
E-mail address: thomas.koehler@fau.de (T. Köhler).

http://dx.doi.org/10.1016/j.media.2015.06.011
1361-8415/

(Bauer et al., 2013). As for minimally invasive procedures, stereo vision is a common technique for intra-operative imaging evaluated, e.g. for brain shift compensation in image-guided neurosurgery (Sun et al., 2005). In the field of active sensor technologies, laser scanning (Cash et al., 2007) and ToF (Kilgus et al., 2014; Mersmann et al., 2011) have been introduced as marker-less approaches to facilitate augmented reality. Similarly to 3-D endoscopy, range data acquired in open surgery can be augmented by photometric information to enhance the intuitive representation of the underlying scene.

1.2. Technical challenges

As demonstrated in recent studies (Maier-Hein et al., 2014), active and passive approaches are complementary RI technologies. While passive stereo vision is able to provide highly accurate range information under ideal situations, it might be error-prone on surfaces without texture or with repetitive structures. Active sensors are less influenced by texture but are limited in their spatial resolution and the signal-to-noise ratio (SNR) due to inherent physical or economical limitations. In particular, the resolutions of low-cost ToF or SL sensors are rather low compared to photometric information. For minimally invasive procedures or open surgery, this means a major limitation and restricted the integration of RI to many clinical workflows.

In order to enhance the spatial resolution of digital images, *super-resolution* (Milanfar, 2010) is a technique to reconstruct high-resolution (HR) data from the acquired raw images. One common approach is to fuse multiple low-resolution (LR) frames into a new HR image (Greenspan, 2008). Most conventional super-resolution algorithms exploit only images from a single modality and are termed as *single-sensor* methods below. Opposed to this approach, *multi-sensor* techniques take advantage of additional guidance by multiple modalities. In terms of RI, high-quality photometric data can be exploited for range super-resolution.

1.3. Contribution and outline

We propose a multi-sensor super-resolution framework and present its application for hybrid RI in image-guided surgery. This paper is an extension of a conference proceeding (Köhler et al., 2013) introducing this concept for 3-D endoscopy for the first time. In this work, we examine sensor data fusion of range and photometric information for two RI setups applicable to 3-D endoscopy and open surgery as generalization of (Köhler et al., 2013). This concept is used to derive subpixel displacement fields for range super-resolution under the guidance of photometric information. As extension of our prior work, we also introduce a novel adaptive regularization technique, where photometric data is exploited for improved edge reconstruction in range super-resolution. In our experimental evaluation, we demonstrate the performance of our method on ex-vivo porcine organs qualitatively and quantitatively.

The remainder of this paper is organized as follows: Section 2 discusses relevant work on upsampling of range data. Section 3 introduces our system calibration approach to perform sensor data fusion for multi-sensor super-resolution. In Section 4, we introduce MAP super-resolution as computational framework for our approach. Section 5 introduces our super-resolution approach for hybrid RI. In Section 6, our method is evaluated for image-guidance in minimally invasive as well as open surgery. A discussion of our method is given in Section 7.

2. Related work

Multi-frame super-resolution algorithms exploit relative movements of a camera with respect to a 3-D surface. Due to subpixel displacements in the associated image sequence, an HR image of finer spatial sampling can be reconstructed (Irani and Peleg, 1991). Unlike single-frame upsampling techniques (He et al., 2010; Kopf et al., 2007; Park et al., 2011), image deblurring and denoising may be treated in a joint approach. In this paper, we distinguish between single-sensor and multi-sensor methods.

2.1. Single-sensor approach

Early approaches divide super-resolution in a motion estimation and a reconstruction stage formulated in the frequency domain (Tsai and Huang, 1984) or the spatial domain using set theoretic or statistical parameter estimation (Elad and Feuer, 1997; Farsiu et al., 2004). The latter approaches are attractive since prior knowledge can be easily incorporated to regularize the reconstruction process. However, most approaches are based on simplified motion models, e.g. a rigid transformation *in the image plane* ignoring the projective mapping of a real camera. For this reason and to handle non-rigid motion, optical flow estimation has been also widely employed (Mitzel et al., 2009; Zhao and Sawhney, 2002). As accurate motion estimation is a major challenge for super-resolution, improved registration schemes have been proposed (Vandewalle et al., 2006). Similarly, image reconstruction and motion estimation can be treated in a *joint* optimization procedure, e.g. based on joint MAP or Bayesian approaches (Hardie et al., 1997; Pickup et al., 2007), variable projection (Robinson et al., 2009) or expectation maximization (Fransens et al., 2007). Even if the reliability of super-resolution is improved, joint optimization is a computationally demanding and highly non-convex problem (Pickup et al., 2007) or limited to simplified motion models.

Recently, such methods have been adopted to RI by Schuon et al. (2009) or Bhavsar and Rajagopalan (2012). However, these techniques super-resolve a single modality, i.e. range data, without exploiting sensor fusion with guidance images, i.e. photometric data. Our work shows that sensor fusion can be employed to enhance super-resolution reconstruction in terms of motion estimation as well as regularization in statistical, spatial domain algorithms.

2.2. Multi-sensor approach

There are only few multi-sensor methods, which are often limited to specific applications. Multi-sensor super-resolution for single RGB/infrared images has been proposed by Zomet and Peleg (2002). One approach designed for PET/CT scanners has been proposed by Kennedy et al. (2007), where PET scans are super-resolved and augmented by anatomical information gained from high-resolution CT scans. Even if this approach is beneficial for PET resolution enhancement, it does not generalize to different applications. In terms of hybrid 3-D endoscopy, a first multi-sensor super-resolution framework has been introduced, recently (Köhler et al., 2013). In this approach, photometric information is employed as guidance for range super-resolution. Motion estimation is performed on high-quality photometric data in order to derive accurate displacement fields for range images. However, this framework is based on a simplified setup for sensor fusion exploiting a homographic mapping and has not been investigated for more general systems, yet.

This work generalizes the approach of Köhler et al. (2013) and demonstrates its application in minimally invasive and open surgery.

3. System calibration

Our framework exploits sensor data fusion between range images and complementary photometric information. The LR range data of size $M_1 \times M_2$ is denoted as $\boldsymbol{Y}$ defined on a rectangular grid $\Omega_r \subset \mathbb{R}^2$. We consider a sequence of K range images denoted as $\boldsymbol{Y}^{(1)}, \ldots, \boldsymbol{Y}^{(K)}$. For convenience, each image $\boldsymbol{Y}$ is linearized to a vector $\boldsymbol{y} \in \mathbb{R}^M$ with $M = M_1 \cdot M_2$. For each $\boldsymbol{Y}^{(k)}$ there exist an associated image $\boldsymbol{z}^{(k)} \equiv \boldsymbol{Z}^{(k)}$