Outlier Detection for Multi-Sensor Super-Resolution in Hybrid 3-D Endoscopy

Thomas Köhler[1,2], Sven Haase[1], Sebastian Bauer[1], Jakob Wasza[1], Thomas Kilgus[3], Lena Maier-Hein[3], Hubertus Feußner[4], Joachim Hornegger[1,2]

[1]Pattern Recognition Lab, FAU Erlangen-Nürnberg
[2]Erlangen Graduate School in Advanced Optical Technologies (SAOT)
[3]Division of Medical and Biological Informatics, DKFZ Heidelberg
[4]Research Group Minimally-invasive Interdisciplinary Therapeutical Intervention, TU München
thomas.koehler@fau.de

Abstract. In hybrid 3-D endoscopy, range data is used to augment photometric information for minimally invasive surgery. As range sensors suffer from a rough spatial resolution and a low signal-to-noise ratio, subpixel motion between multiple range images is used as a cue for super-resolution to obtain reliable range data. Unfortunately, this method is sensitive to outliers in range images and the estimated subpixel displacements. In this paper, we propose an outlier detection scheme for robust super-resolution. First, we derive confidence maps to identify outliers in the displacement fields by correlation analysis of photometric data. Second, we apply an iteratively re-weighted least squares algorithm to obtain the associated range confidence maps. The joint confidence map is used to obtain super-resolved range data. We evaluated our approach on synthetic images and phantom data acquired by a Time-of-Flight/RGB endoscope. The proposed method improves the median peak-signal-to-noise ratio by 1.1 dB compared to super-resolution without outlier detection.

1 Introduction

In hybrid 3-D endoscopy, photometric information is augmented with 3-D data provided by range imaging (RI) sensors, e. g. based on Time-of-Flight (ToF) imaging [1]. While range data is essential for robotic-based interventions, photometric data provides color and texture of tissue to enhance an intuitive representation of the underlying scene. However, a serious issue towards clinical applications is the limited spatial resolution and the low signal-to-noise ratio (SNR) of todays range sensors. Super-resolution algorithms reconstruct a high-resolution (HR) image from multiple low-resolution (LR) frames by exploiting subpixel displacements present in an image sequence [2]. This approach has recently been introduced to RI in image-guided surgery [3]. For an improved reconstruction, Köhler et al. [4] proposed a super-resolution framework for a multi-sensor setup. In this approach, photometric data is utilized for optical flow estimation to derive displacement fields for the associated range images. Due to the increased

accuracy of motion estimation on photometric data this yields accurate super-resolved range data. However, the method relies entirely on optical flow estimation and is susceptible to mis-registration in difficult scenarios such as large displacements caused by sizable non-rigid deformations of tissue or independent moving objects. Additionally, range data is disturbed by outliers, e. g. due to specular highlights. There are two strategies to deal with such outliers: (i) Error models accounting for outliers can be utilized for super-resolution. This results in an optimization problem involving robust error norms [5]. (ii) Outliers can be removed before super-resolution is performed. This can be achieved using image similarity metrics to identify outliers caused by erroneous motion estimation [6].

Both approaches focus on a single modality without exploiting additional guidance by different modalities in hybrid imaging. To overcome this issue, we propose outlier detection for multi-sensor super-resolution. Our method identifies outliers in displacement fields as well as in LR range data. For robust super-resolution, we derive confidence maps using image similarity and an iterative scheme to assign less weights to imputed outliers. The performance of our method is demonstrated in hybrid 3-D endoscopy to super-resolve range data.

2 Materials and Methods

Super-resolution is applied to range images $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(K)}$ where each $\boldsymbol{y}^{(k)} \in \mathbb{R}^M$ is represented as a vector. Due to movements of the camera or motion in the underlying scene, each $\boldsymbol{y}^{(k)}$ is related to a reference frame $\boldsymbol{y}^{(r)}$ by a geometric transformation. For each range image $\boldsymbol{y}^{(k)}$, there exists an optical image $\boldsymbol{z}^{(k)}$acquired simultaneously to encode photometric information in a hybrid imaging setup.

2.1 Multi-Sensor Super-Resolution

The objective of multi-sensor super-resolution is to reconstruct a super-resolved range image $\hat{\boldsymbol{x}} \in \mathbb{R}^N$ with $N > M$ from $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(K)}$ according to:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \sum_i \beta_i \left| r_i(\boldsymbol{x}) \right|^p + \lambda R(\boldsymbol{x}), \tag{1}$$

where $r : \mathbb{R}^N \to \mathbb{R}^{KM}$ denotes a residual term employed in the underlying error model based on the L_p norm and $\boldsymbol{\beta} \in \mathbb{R}^{KM}$ is a confidence map to weight the residual element-wise. For the regularizer $R(\boldsymbol{x})$ with weight $\lambda > 0$, the edge-preserving Huber prior is employed. We set $r(\boldsymbol{x}) = (\boldsymbol{r}^{(1)}, \ldots, \boldsymbol{r}^{(K)})^\top$ and the residual of the k^{th} frame $\boldsymbol{y}^{(k)}$ is given as:

$$\boldsymbol{r}^{(k)} = \boldsymbol{y}^{(k)} - \gamma_m^{(k)} \boldsymbol{W}^{(k)} \boldsymbol{x} - \gamma_a^{(k)} \mathbf{1}, \tag{2}$$

where $\boldsymbol{W}^{(k)}$, $\gamma_m^{(k)}$ and $\gamma_a^{(k)}$ denote the system matrix and the associated range correction factors for the k^{th} frame and $\mathbf{1} \in \mathbb{R}^M$ is the all-one vector. In a multi-sensor approach, $\boldsymbol{W}^{(k)}$ encodes downsampling, the range sensor point spread function (PSF) and 2-D geometric displacements derived from optical flow estimated on photometric data. The parameters $\gamma_m^{(k)}$ and $\gamma_a^{(k)}$ are determined by a range correction scheme to account for out-of-plane motion [4].

2.2 Outlier Detection

The confidence map $\boldsymbol{\beta}$ in Eq. (1) is composed element-wise as $\beta_i = \beta_{r,i} \cdot \beta_{z,i}$. Here, $\boldsymbol{\beta}_r$ is chosen to suppress outliers in range data. The confidence map $\boldsymbol{\beta}_z$ weights outliers in the associated displacement fields. The confidence maps are derived by image similarity analysis and an iterative outlier detection procedure.

We detect outliers in displacement vector fields provided by optical flow in the domain of the photometric data. Therefore, the reference frame $\boldsymbol{z}^{(r)}$ is aligned with each frame $\boldsymbol{z}^{(k)}$ according to the estimated displacements. Afterwards, we analyze the similarity between the warped reference $\tilde{\boldsymbol{z}}^{(r)}$ and each frame $\boldsymbol{z}^{(k)}$. In this paper, we employ the normalized cross-correlation (NCC) as similarity metric patch-wise to derive the confidence map $\boldsymbol{\beta}_z$ according to:

$$\beta_{z,i} = \frac{\sum_{\boldsymbol{v}\in\mathcal{N}(\boldsymbol{u}_i)}(\boldsymbol{z}^{(k)}(\boldsymbol{v}) - \bar{\boldsymbol{z}}^{(k)})(\tilde{\boldsymbol{z}}^{(r)}(\boldsymbol{v}) - \bar{\boldsymbol{z}}^{(r)})}{\sqrt{\sum_{\boldsymbol{v}\in\mathcal{N}(\boldsymbol{u}_i)}(\boldsymbol{z}^{(k)}(\boldsymbol{v}) - \bar{\boldsymbol{z}}^{(k)})^2}\sqrt{\sum_{\boldsymbol{v}\in\mathcal{N}(\boldsymbol{u}_i)}(\tilde{\boldsymbol{z}}^{(r)}(\boldsymbol{v}) - \bar{\boldsymbol{z}}^{(r)})^2}}, \tag{3}$$

where $\mathcal{N}(\boldsymbol{u}_i)$ denotes the local neighborhood formed by the set of pixels in the photometric data associated with the i^{th} pixel $\boldsymbol{u}_i$ in the range images, and $\bar{\boldsymbol{z}}^{(k)}$ and $\bar{\boldsymbol{z}}^{(r)}$ are the local means in $\mathcal{N}(\boldsymbol{u}_i)$ for the k^{th} frame and the reference, respectively. For $\beta_{z,i} < \varepsilon_z$, we set $\beta_{z,i} = 0$ to reject $\boldsymbol{u}_i$ as an outlier where ε_z is adjusted to the noise level of photometric data. The confidence map $\boldsymbol{\beta}_z$ is transformed to the domain of LR range data. This approach is similar to [6], whereas in our work outliers are detected in photometric data instead of using LR data directly.

While displacement outliers are removed based on photometric data exploited as guidance, outliers in depth data must be detected in range images directly. Therefore, we propose an iterative re-weighted least squares (IRLS) scheme [7] to derive the confidence map $\boldsymbol{\beta}_r$. Let $\boldsymbol{x}^{(0)}$ be the super-resolved image as solution of Eq. (1) using the confidence maps $\boldsymbol{\beta}_r^{(0)} = \mathbf{1}$ and $\boldsymbol{\beta}_z$ determined according to Eq. (3). For $p = 2$, this is done by common least squares optimization such as a scaled conjugate gradient (SCG) algorithm. Then, the residual $\boldsymbol{r}^{(0)} = r(\boldsymbol{x}^{(0)})$ given by Eq. (2) is used to derive the range confidence as:

$$\beta_{r,i}^{(0)} = \begin{cases} 1 & \text{if } |r_i^{(0)}| \leq \varepsilon_r \\ \frac{\varepsilon_r}{|r_i^{(0)}|} & \text{otherwise} \end{cases}. \tag{4}$$

We set $\varepsilon_r = \sigma_r$ for the residual standard deviation σ_r. Therefore, we employ the robust estimator $\sigma_r = 1.4826 \cdot \mathrm{MAD}$ based on the median absolute deviation $\mathrm{MAD} = \mathrm{Median}_i(|r_i - \mathrm{Median}_i(r_i)|)$. The confidence map $\boldsymbol{\beta}_r$ and the super-resolved data $\boldsymbol{x}$ are iteratively updated in IRLS as depicted in Table 1.

2.3 Experiments

We compared the proposed multi-sensor super-resolution (MSR) to the approach using the L_2 norm model introduced in [4]. Additionally, we evaluated MSR

Table 1. Super-resolution using iteratively re-weighted least squares (IRLS).

1. Initialize range confidence map $\beta_{r,i}^{(0)} = 1$ for $i = 1, \ldots, KM$ and $t = 0$.
2. Determine residual term $\boldsymbol{r}^{(t)} = r(\boldsymbol{x}^{(t-1)})$ according to Eq. (2).
3. Set $\beta_i^{(t)} = \beta_{r,i}^{(t)} \cdot \beta_{z,i}$ with $\boldsymbol{\beta}_z$ according to Eq. (3) and $\boldsymbol{\beta}_r^{(t)}$ according to Eq. (4).
4. Solve for $\boldsymbol{x}^{(t)}$ with $p = 2$ according to Eq. (1) using SCG optimization: $\boldsymbol{x}^{(t)} = \arg\min_{\boldsymbol{x}} \left\{ \sum_i \beta_i^{(t)} r(\boldsymbol{x})^2 + \lambda R(\boldsymbol{x}) \right\}$
5. Set $t \leftarrow t + 1$ and proceed with step 2 until convergence.

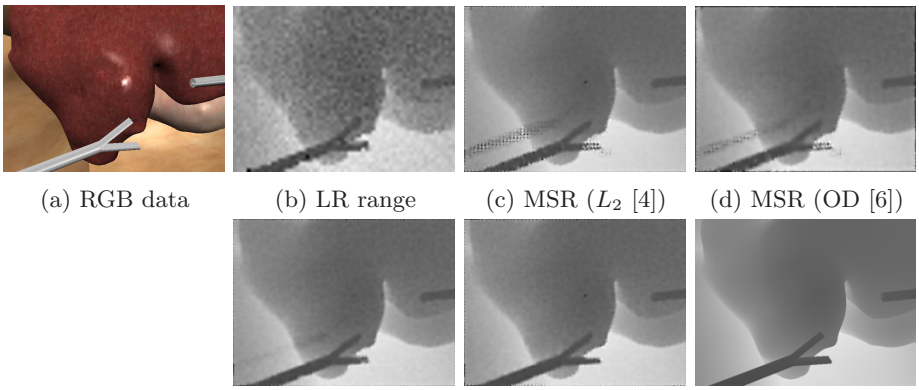
based on L_2 norm with outlier detection on LR data [6] and MSR using a robust L_1 norm model [5]. Motion estimation and regularization was realized analogous to [4]. $K = 31$ frames were used to achieve a magnification of 4. We approximated the PSF as a Gaussian of width $\sigma = 0.2$. For outlier detection, we set $\varepsilon_z = 0.8$. Supplementary material for our experiments is available on our web page[1].

First, RGB images (640×480 px) and LR range data (64×48 px) were obtained from a laparoscopic model using an RI simulator. Range data was affected by distance-dependent Gaussian noise (max. $\sigma_n = 10$ mm) and Gaussian blur ($\sigma_b = 3$ mm). Perlin noise was induced to simulate invalid ToF measurements caused by specular highlights. We simulated flying pixels in range data by randomly flipping 20% of all edge pixels. In terms of motion, we generated four data sets (S1 - S4): Small random motion of the virtual camera was used to simulate a jitter of a hand-held endoscope (S1). The endoscope was displaced to slightly different viewing directions (S2). Surgical tools were shifted (S3). Organ surfaces were moved to simulate respiratory motion (S4). Super-resolved data was assessed by comparison to a ground truth using the peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM). Second, we measured a liver phantom with a ToF/RGB endoscope prototype manufactured by Richard Wolf GmbH, Knittlingen, Germany. Image data was captured with a frame rate of 30 Hz in the same spatial resolution as synthetic data. During acquisition the endoscope was moved relative to the liver and tools were moved due to a jitter of the hand.

3 Results

Qualitative results for synthetic data is presented in Fig. 1. If no outlier detection was employed, super-resolved data contained artifacts due to mis-registrations. PSNR and SSIM was evaluated for ten sequences per data set using sliding window processing and is shown as boxplot in Fig. 2. We obtained median values of 33.4 dB for PSNR and 0.94 for SSIM in the absence of outlier detection. The proposed method improved the median PSNR (SSIM) by 1.1 dB (0.01). Our approach also outperformed outlier detection on LR data and L_1 norm based super-resolution. For phantom data, super-resolved range images are presented in Fig. 3. In our experiments, mis-registrations for endoscopic tools caused artifacts

[1] http://www5.cs.fau.de/research/data/



(a) RGB data (b) LR range (c) MSR (L_2 [4]) (d) MSR (OD [6])

(e) MSR (L_1 [5]) (f) MSR (proposed) (g) Ground truth

Fig. 1. Synthetic RGB (a) and range data (b): Multi-sensor super-resolution (MSR) using L_2 norm model without (c) and with outlier detection (d), based on L_1 norm model (e) and our method (f) compared to the ground truth (g).

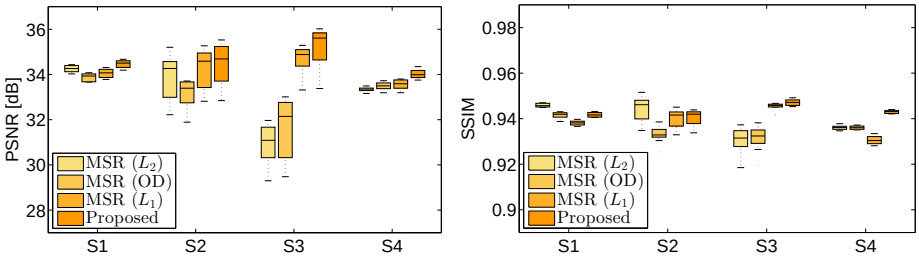
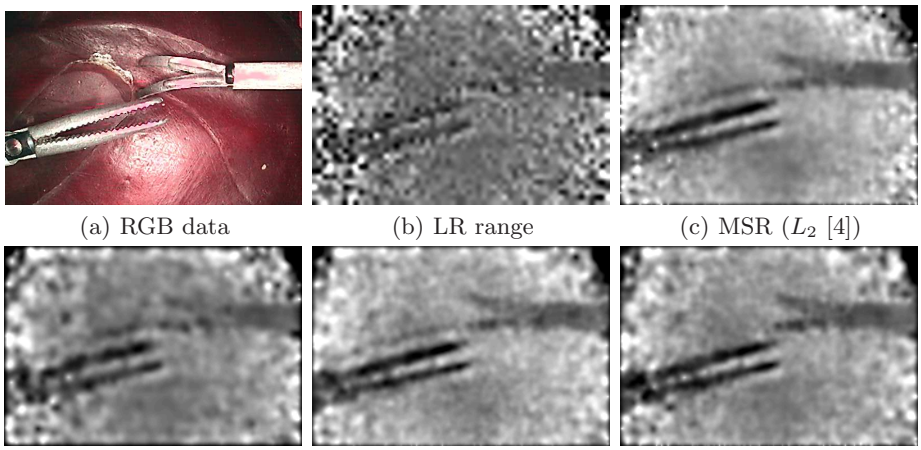


Fig. 2. Boxplots of peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM) created for ten sequences per data set using sliding window processing.

in the super-resolved image if outlier detection was not employed. These artifacts were still present in case of outlier detection on LR data and the L_1 norm approach but well suppressed using the proposed method.

4 Discussion

In this work, we proposed an outlier detection scheme for robust multi-sensor super-resolution. Our approach detects outliers in range data and the associated displacement fields to obtain reliable range data for hybrid 3-D endoscopy. The proposed method outperforms outlier detection based on LR data only and exploits photometric data as guidance. Compared to super-resolution using an L_1 norm model, our method achieves improved robustness under real conditions on phantom data and improved results in terms of PSNR and SSIM on synthetic data. Future work will focus on the integration of different weighting schemes to our outlier detection and a combination of IRLS with L_1 norm optimization.



(a) RGB data (b) LR range (c) MSR (L_2 [4])

(d) MSR (OD [6]) (e) MSR (L_1 [5]) (f) MSR (proposed)

Fig. 3. RGB (a) and range data (b) obtained from a liver phantom with the results of multi-sensor super-resolution (MSR) using L_2 norm model without (c) and with outlier detection (d), based on L_1 norm model (e) and our method (f).

Acknowledgments The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative and the support by the DFG under Grant No. HO 1791/7-1. This research was funded by the Graduate School of Information Science in Health (GSISH) and the TUM Graduate School. We thank the Metrilus GmbH for their support.

Errata for published conference proceedings

- Changes in the abstract
- Caption and border of Table 1 removed
- Acknowledgments shortened

References

1. Haase S, Forman C, Kilgus T, Bammer R, Maier-Hein L, Hornegger J. ToF/RGB Sensor Fusion for 3-D Endoscopy. Curr Med Imaging Rev. 2013;9:113–119.
2. Park SC, Park MK, Kang MG. Super-resolution image reconstruction: a technical overview. IEEE Signal Process Mag, IEEE. 2003;20(3):21–36.
3. Wetzl J, Taubmann O, Haase S, Köhler T, Kraus M, Hornegger J. GPU-Accelerated Time-of-Flight Super-Resolution for Image-Guided Surgery. In: Proc BVM; 2013. p. 21–26.
4. Köhler T, Haase S, Bauer S, Wasza J, Kilgus T, Maier-Hein L, et al. ToF Meets RGB: Novel Multi-Sensor Super-Resolution for Hybrid 3-D Endoscopy. In: MICCAI 2013. vol. 8149 of LNCS. Springer Berlin Heidelberg; 2013. p. 139–146.

5. Farsiu S, Robinson MD, Elad M, Milanfar P. Fast and robust multiframe super resolution. IEEE Transactions on Image Processing. 2004;13(10):1327–1344.
6. Zhao W, Sawhney HS. Is super-resolution with optical flow feasible? In: ECCV 2002. vol. 2350 of LNCS. Springer; 2002. p. 599–613.
7. Scales JA, Gersztenkorn A. Robust methods in inverse theory. Inverse Problems. 1988;4(4):1071–1091.