Eric L. Wisotzky*, Jost Triller, Benjamin Kossack, Brigitta Globke, Philipp Arens, Anna Hilsmann, and Peter Eisert

From Multispectral-Stereo to Intraoperative Hyperspectral Imaging: a Feasibility Study

https://doi.org/10.1515/cdbme-2023-1078

Abstract: Spectral imaging allows to analyze optical tissue properties that are invisible to the naked eye. Different spectral setups have been introduced in the last years to allow intraoperative image-guidance. All systems suffer from different disadvantages, such as lack of real-time capability, limited spectral coverage or low resolution. We present a novel approach using two multispectral snapshot cameras covering different spectral ranges as a stereo-system. The proposed method allows for continuous hyperspectral imaging of the surgical scene, while also enabling 3D scene reconstruction. The study demonstrates the feasibility of this approach and its potential applications in real-time monitoring and tissue identification of surgical interventions. The paper also discusses the technical challenges and future directions for improving the imaging system.

Keywords: Hyperspectral Imaging, Multispectral Imaging, Stereoscopic Analysis, Image Fusion.

1 Introduction

Hyperspectral imaging is a contact-less, radiation-free and non-destructive method to analyze optical surface properties that are invisible to the naked eye [5]. This can be of great benefit in medical intraoperative image analysis. Therefore, hyperspectral imaging has experienced a strong interest in clinical research in recent years [2].

Hyperspectral cameras are used in the spectral range of $\lambda = 400 - 1000$ nm to retrieve important information from the image data, such as physiological parameters, specific tissue characteristics and their changes. To acquire hyperspectral data cubes with 100 or more spectral bands, different capturing techniques exists, e.g., line-scanning, filter-wheel, or push-broom based scanning. However, these techniques are based on sequential scanning and thus need several seconds to acquire one complete hyperspectral data cube [8], hindering real-time processing and interactive applications.

In contrast, multispectral mosaic cameras acquire spectral data within one frame [4]. However, this comes with the drawback of much lower spectral resolution and decreased spatial resolution. Snapshot cameras use mosaic patterns similar to a standard RGB-Bayer pattern, but with increased pattern size and different filters on each pixel [9].

This work combines stereoscopic image analysis with the benefits of mosaic-snapshot cameras by using mosaic-snapshot sensors with different spectral coverage in a stereo-setup to increase spectral resolution without loosing additional spatial resolution. With the aid of stereoscopic image processing, we transfer the spectral information between the different cameras, enabling a real-time capable hyperspectral setup without loosing spatial resolution and additionally providing depth information about the scene.

2 Method

In order to achieve a wide spectral coverage with good sampling rate, we assume that both cameras of a multispectral-stereo system have no overlap in their spectral ranges. Otherwise, the stereoscopic image analysis could be performed on the overlapping spectral bands. To provide the full spectral information from $\lambda = 400 - 1000$ nm for each point in the scene, the captured spectral data of both cameras have to be fused into one consistent data cube. To achieve that, the following steps are performed:

1. **Demosaicking** to increase images size and spatial resolution. Classical methods, e.g., bilinear or identity difference interpolation [6], or network-based methods, e.g., 3D-CNNs [9] can be used.
2. **Band selection** in each image to extract two bands showing the best similarity in terms of intensity distribution. The aim of this step is to extract two bands having the lowest normalized difference for later matching between the 2 cameras. To improve the outcome of the next steps, both selected bands are enhanced in contrast.
3. **Feature detection** in both images and **matching** of these features [14]. Optionally, it is possible to combine step 2

***Corresponding author: Eric L. Wisotzky,** Fraunhofer Heinrich-Hertz-Institute HHI, Humboldt-University & Charité, Berlin, Germany, e-mail: eric.wisotzky@hhi.fraunhofer.de
Jost Triller, Benjamin Kossack, Anna Hilsmann, Fraunhofer Heinrich-Hertz-Institute HHI, Berlin, Germany
Brigitta Globke, Philipp Arens, Charité – Universitätsmedizin, Berlin, Germany
Peter Eisert, Fraunhofer Heinrich-Hertz-Institute HHI & Humboldt-University, Berlin, Germany

Open Access. © 2023 The Author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

and 3. The feature detection and matching is then performed for all spectral bands. Then, all found feature pairs are used for the following steps.

4. **Homography** estimation and calculation of the warping matrix.
5. **Rectification** of the camera views based on multi-view geometry to achieve that matching features are on the same image row in both views. This reduces the corresponding problem by one dimension.
6. **Disparity** calculation results in a sparse disparity map [1]. By looping over the image pair for several times to enhance feature matching outcome, a dense disparity map is achieved.
7. **Transferring spectral information** from one camera view into the other. All spectral data can be fused for pixels with existing disparity information.

This results in a hyperspectral data cube holding in addition depth information in the form of the disparity map. With a stereo calibration of the camera system, accurate depth information can be derived [12].

3 Stereo-Setup and Data

To show the functionality and intraoperative usability of our multispectral image fusion, we captured intraoperative stereo data with two different multispectral snapshot cameras[1] mounted on a rigid trail to realize a stereo setup, see Fig. 1. The left camera is a 4×4-VIS camera covering the spectral range from $\lambda = 436$ nm to $\lambda = 650$ nm with 16 spectral bands and a sensor resolution of 1024×2048 pixels. The right camera is a 5×5-NIR camera covering the spectral range from $\lambda = 675$ nm to $\lambda = 975$ nm with 25 spectral bands on a 1080×2045 pixels sensor. Both cameras hold a F$2.8/75$ mm optics[2] and are described in more detail in Mühle et al. [7].

The distance between both cameras is $a = 6$ cm. The cameras are rotated around the image plane for about 10 degree. This allows the same scene content shown on the sensor at a working distant (WD) of about $d = 25 - 35$ cm. Both cameras are frame-wise synchronized and connected to a laptop providing power and fast data transfer with capturing rates of up to 60 Hz, depending of the needed acquisition parameter like exposure settings.

We captured data during different surgeries in the surgical fields: parotidectomy (ENT surgery), cf. Fig. 2 and kidney transplantation, cf. Fig. 5. Both types of surgery are performed as open surgeries, in which the surgeon usually wears magnifying glasses to enlarge the anatomical structures. For illumination, the present surgical light is used, which always comprehensively covers the interesting parts of the scene with high brightness. To correct the present LED-illumination setting, a white reference image is acquired directly before or after the actual capturing. The image correction is performed as described in Wisotzky et al. [10]



Fig. 1: The stereo-HSI-setup. Left: the setup mounted on a surgical retractor system in the operation room. Right: detailed view of the stereo-cameras mounted on the rigid trail.

After the demosaicking step, both images are cropped to identical sizes of 1020×2000 pixels for better image handling. Due to performance reasons, we warped the 16 bands of the 4×4-camera and concatenated the 16 spectral points for all pixels having disparity information at the beginning of the 25 bands of the 5×5-camera to achieve 41 spectral bands per pixel.

4 Results

The captured image data of seven surgeries in two different surgical fields have been analyzed. It was possible to reconstruct a dense disparity map and transfer the spectral data for all cases.

The final hyperspectral data cube is always smaller than the multispectral-stereo input images, since the baseline of the stereo system leads to a small scene shift as well as to occlusions in the two multispectral-snapshot images, see Fig. 5. The rectified and warped image shows empty image borders due to the described stereoscopic effects, see Fig. 2, but due to the dimension reduction, a fast determination of the disparity information is achievable. The dense disparity map of the case shown in Fig. 2 is presented in Fig. 3. No information are present on the left side of the disparity map (white area in Fig. 3) due to the stereoscopic baseline shift.

Using the calculated disparity map, the captured spectral data can be taken pixel-wise from the 4×4-camera and stitched in front of the 5×5-camera data. Thus, for each pixel having

1 XIMEA GmbH
2 Ricoh

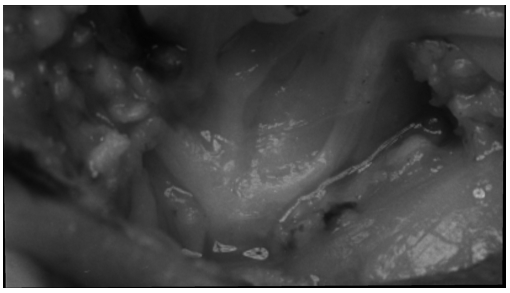


Fig. 2: One spectral channel of the right 5×5 camera view after step six (warping) of our processing algorithm. The disparity calculation is now reduced by one dimension as all corresponding image features are located on the identical image row on both camera views. Due to the warping, some areas at the image borders hold no data (black borders).

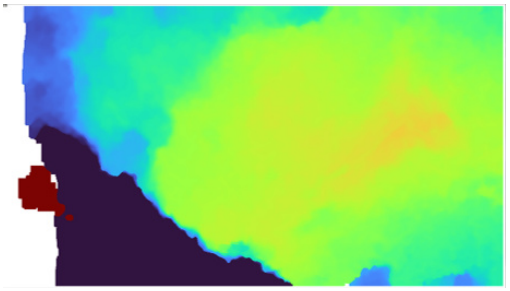


Fig. 3: The disparity map of the case shown in Fig. 2. Red pixels refer to regions far away from the camera and blue pixels refer to regions closer to the camera. The white area on the left holds no information due to the stereoscopic baseline shift. In addition, due to the baseline shift and missing information, there is also an incorrect prediction of a region distant from the camera system (red pixels) that is actually close to the camera system.

a valid disparity correspondence a full spectrum from the captured scene can be calculated, see Fig. 4.

The measured latencies of the entire algorithm are presented in Tab. 1. The processing steps are implemented in C++ and partially implemented for GPU usage. The latencies were measured on a DELL Precision 5570 laptop with NVIDIA RTX A2000 GPU. While the pipeline is running, incoming images are ignored.

5 Discussion and Conclusion

In this work, we present an image/spectral fusion method for stereoscopic multispectral snapshot images. This allows a continuous capturing (60 fps) and processing of intraoperative images to achieve a hyperspectral data cube as video-like stream (>1 fps). The hyperspectral data cube shows high spatial resolution with about 2000×1000 pixels and a good spectral resolution with 41 spectral bands in the interval of

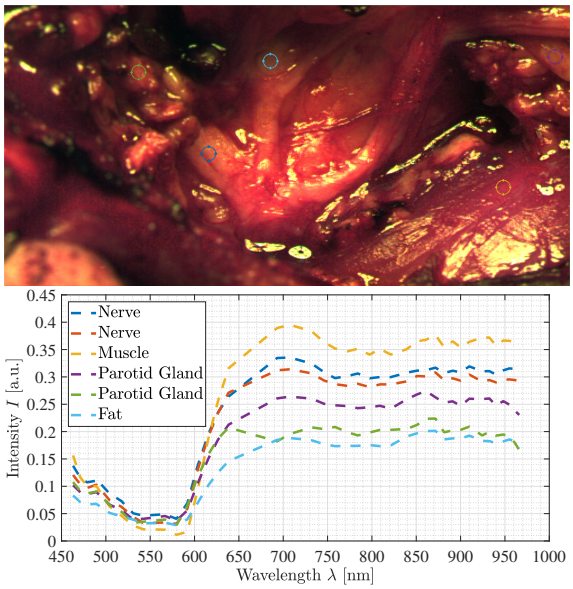


Fig. 4: Due to the fusion of the spectral information into one camera view, a realistic RGB representation can be calculated (top view). Further for all regions with disparity information, the spectral information in the VIS-NIR range can be extracted, cf. colored circular ROIs in the top view and spectral plots in the bottom view.

$\lambda = 400 - 1000$ nm. This spectral resolution with the present sampling is sufficient to reconstruct a spectrum in the VIS-NIR region for robust analyses, which is comparable to spectra extracted from other hyperspectral cameras [7]. Furthermore, our system combines advantages of different HSI and MSI systems from the literature without showing their disadvantages [2–4]. In addition, it allows a real-time reconstruction of the 3D surface texture of the situs, which thus creates a significant value increase. This makes it possible to continuously monitor the present anatomical structures for physiological changes and at the same time, if necessary, to measure their sizes or register them precisely on preoperative image data. Thus, the surgeon can be comprehensively provided with objective data. Other comparable systems only allowed to reconstruct few spectral bands and 3D surface and were not real-

Tab. 1: Computational times of individual modules if feature detection is performed on all 41 spectral bands.

Module	Latency [ms]
Demosaicking	250
Band Selection	120
Feature Detection	300
Homography	100
Disparity Map	30
Final Fusion	50
Total Latency	850

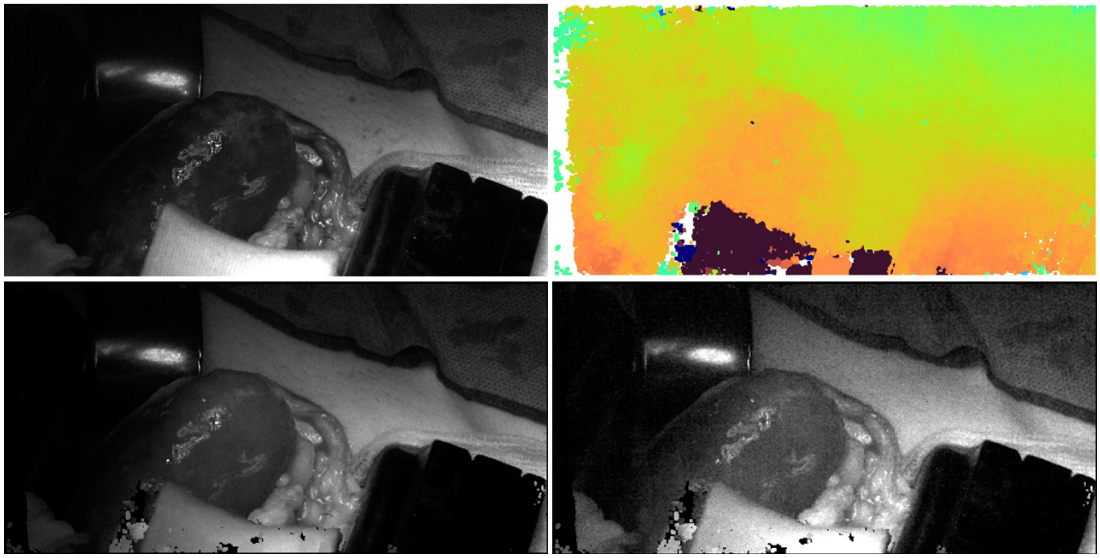


Fig. 5: The final data fusion of data captured during kidney transplantation. Top left: left camera view, demosaicked. Top right: calculated disparity map. Bottom row: two spectral bands of the fused data from right view into the left view.

time capable [13]. Our system shows the potential of real-time analysis as the pipeline is only partially implemented for GPU usage.

Author Statement

This work was funded by the German BMBF under Grant No. 16SV8602 (KIPos) and by the German BMWK under Grant No. 01MT20001D (Gemimeg II). The authors state no conflict of interest. Informed consent has been obtained from all individuals included in this study. The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the Charité ethics board.

References

[1] Allan M, Mcleod J, Wang C, Rosenthal JC, Hu Z, Gard N, et al. Stereo correspondence and reconstruction of endoscopic data challenge. arXiv preprint, 2021, arXiv:2101.01133.

[2] Clancy NT, Jones G, Maier-Hein L, Elson DS, Stoyanov D. Surgical spectral imaging. Med Image Anal, 2020, 63:101699.

[3] Clancy NT, Soares AS, Bano S, Lovat LB, Chand M, Stoyanov D. Intraoperative colon perfusion assessment using multispectral imaging. Biomed Opt Express, 2021, 12(12):7556-7567.

[4] Ebner M, Nabavi E, Shapey J, Xie Y, Liebmann F, Spirig JM, et al. Intraoperative hyperspectral label-free imaging: from system design to first-in-patient translation. J Phys D: Appl Phys, 2021, 54(29):294003.

[5] Lu G, Fei B. Medical hyperspectral imaging: a review. J Biomed Opt, 2014, 19(1):010901.

[6] Mihoubi S, Losson O, Mathon B, Macaire L. Multispectral demosaicing using intensity-based spectral correlation. Int Conf Image Proc Theory, Tools and Appl (IPTA). IEEE, 2015, pp. 461-466

[7] Mühle R, Markgraf W, Hilsmann A, Malberg H, Eisert P, Wisotzky EL. Comparison of different spectral cameras for image-guided organ transplantation. J Biomed Opt, 2021, 26(7):076007.

[8] Shapey J, Xie Y, Nabavi E, Bradford R, Saeed SR, Ourselin S, Vercauteren T. Intraoperative multispectral and hyperspectral label-free imaging: A systematic review of in vivo clinical studies. J Biophotonics, 2019, 12(9):e201800455.

[9] Wisotzky EL, Daudkane C, Hilsmann A, Eisert P. Hyperspectral Demosaicing of Snapshot Camera Images Using Deep Learning. Pattern Recognition: 44th DAGM German Conference. DAGM GCPR, 2022, pp. 198-212.

[10] Wisotzky EL, Kossack B, Uecker FC, Arens P, Hilsmann A, Eisert P. Validation of two techniques for intraoperative hyperspectral human tissue determination. J Med Imaging, 2020, 7(6):065001.

[11] Wisotzky EL, Rosenthal JC, Hilsmann A, Eisert P, Uecker FC. A multispectral 3D-endoscope for cholesteatoma removal. Curr Dir Biomed Eng, 2020, 6(3):257-260.

[12] Wisotzky EL, Rosenthal JC, Meij S, van den Dobblesteen JJ, Arens P, Hilsmann A, et al. Telepresence for surgical assistance and training using eXtended reality (during and after pandemic periods). J Telemed Telecare, 2023, in publication. doi: 10.1177/1357633X231166226

[13] Wisotzky EL, Rosenthal JC, Wege U, Hilsmann A, Eisert P, Uecker FC. Surgical guidance for removal of cholesteatoma using a multispectral 3D-endoscope. Sensors, 2020, 20(18):5334.

[14] Zilly F, Müller M, Eisert P, Kauff P. The stereoscopic analyzer—an image-based assistance tool for stereo shooting and 3D production. Int Conf Image Proc (ICIP). IEEE, 2010, pp. 4029-4032.