

Deep learning-based hyperspectral image reconstruction from emulated and real computed tomography imaging spectrometer data

Markus Zimmermann,^{a,*} Simon Amann,^a Mazen Mel[©],^b Tobias Haist,^a
and Alexander Gatto[©]^c

^aUniversity of Stuttgart, Institut für Technische Optik, Stuttgart, Germany

^bUniversity of Padua, Department of Information Engineering, Padua, Italy

^cSony Europe B.V., Stuttgart Technology Center, Stuttgart, Germany

Abstract. The computed tomography imaging spectrometer (CTIS) is a hyperspectral imaging (HSI) approach where spectral and spatial information of a scene is mixed during the imaging process onto a monochromatic sensor. This mixing is due to a diffractive optical element integrated into the underlying optics and creates a set of diffraction orders. To reconstruct a three-dimensional hyperspectral cube from the CTIS sensor image, iterative algorithms were applied. Unfortunately, such methods are highly sensitive to noise and require high computational time for reconstruction thus hindering their applicability in real-time and high frame-rate applications. To overcome such limitations, we propose a lightweight and efficient deep convolutional neural network for hyperspectral image reconstruction from CTIS sensor images. Compared with classical approaches our model delivers considerably better reconstruction results on synthetic as well as real CTIS images in under 0.17 s, which is over 60 times faster compared with the standard iterative approach. In addition, the reshaping method we have developed enables a lightweight network architecture with over 100 times fewer parameters than previously reported. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.OE.61.5.053103](https://doi.org/10.1117/1.OE.61.5.053103)]

Keywords: hyperspectral imaging; neural networks; computed tomography imaging spectrometer; hyperspectral reconstruction.

Paper 20220334G received Mar. 31, 2022; accepted for publication May 12, 2022; published online May 28, 2022.

1 Introduction

Hyperspectral imaging (HSI) refers to the acquisition of multiple images corresponding to more spectral bands than just the typical three color channels red, green, and blue in RGB imaging. Early applications of HSI included determining the spectra of stars and their composition¹ or the identification of clay minerals from airborne measurements.² During the last decade HSI has been becoming more important in areas such as medicine,³ food inspection,⁴ environmental monitoring,⁵ and many others. Multiple approaches for hyperspectral data acquisition have been proposed in the literature and the reader might refer to an extended historical overview of HSI in Ref. 6 and different HSI systems in Ref. 7.

Computed tomography imaging spectrometer (CTIS) has been first proposed by Okamoto and Yamaguchi⁸ and independently by Bulygin and Vishnyakov.⁹ A three-dimensional (3D) data cube (x , y , and wavelength) is determined on the basis of projections taken from different “directions.” The advantages of CTIS are its hyperspectral snapshot capability and simple setup.¹⁰ The disadvantages are a large unused sensor area and a time-consuming iterative algorithm that is required to reconstruct the hyperspectral image.

*Address all correspondence to Markus Zimmermann, markus.zimmermann@ito.uni-stuttgart.de

2 Computed Tomography Imaging Spectrometer

A standard CTIS design is shown in Fig. 1. The object is illuminated by a broadband light source. A field stop is used to block the outer part of the intermediate image and thereby determines the systems' field of view. For our work, we use an aperture with a quadratic shape which results in a quadratic hyperspectral image. Two further lenses first collimate the light coming from the field stop and then image it onto the sensor. In between is a dispersive element that separates the light depending on the wavelength. Commonly, a diffractive optical element (DOE) (e.g., made of several crossed linear gratings) is employed.¹² We use a binary computer-generated hologram (CGH) that leads to an image shown in Fig. 2(a). In the central part of the image, the original broadband scene is located. Outside of this central area, we have 12 dispersed copies due to the first diffraction order of the CGH. It can be seen that the longer wavelengths, i.e., red region, are diffracted more than the shorter ones, i.e., blue region. This way the spectral information is spatially encoded on the sensor. Since we use a monochromatic sensor we obtain an image as shown in Fig. 2(b). The CGH is calculated using the direct binary search algorithm.¹³

Research on CTIS has focused on reducing the unused sensor space with different diffraction patterns,^{14–17} accelerating the reconstruction process,^{18–23} and finding applications for CTIS in, e.g., fluorescence microscopy,²⁴ infrared spectroscopy,^{25,26} and astronomy.^{27,28}

2.1 Hyperspectral Image Reconstruction from CTIS Data

Spectral information encoded on the sensor can be retrieved using iterative reconstruction algorithms. To reconstruct the hyperspectral cube, usually, the expectation-maximization (EM)

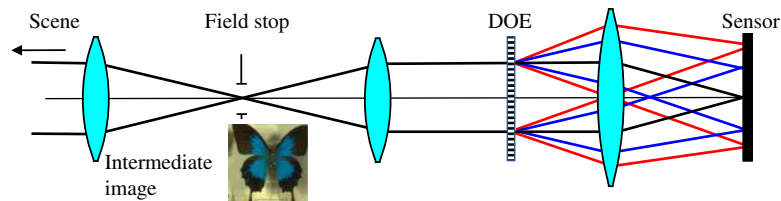


Fig. 1 Optical schematic of a standard CTIS setup. An intermediate image of an object located to the left is created and cropped by a field stop (the butterfly image shown here is taken from Ref. 11). The cropped image is then diffracted into several directions by the DOE and re-imaged onto the sensor on the right. The CTIS sensor image is shown in Fig. 2.

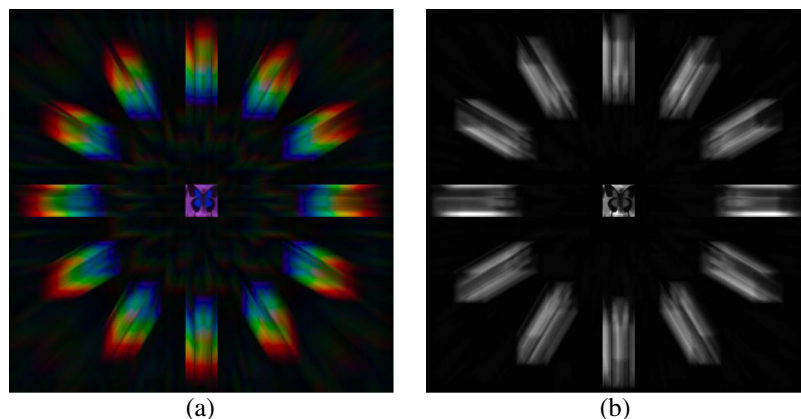


Fig. 2 CTIS sensor image: (a) RGB representation (butterfly image taken from Ref. 11) is generated using the methods described in Sec. 3.2.1. The colors of the zeroth-order are changed compared with the ground truth since the spectral efficiencies of the DOE are taken into account. (b) monochromatic sensor image as used to extract the hyperspectral information. Reconstructions from this image are shown in Sec. 4. The diffraction orders are slightly brightened here for illustration purposes.

algorithm is used.¹⁸ EM uses calibration data in the form of recorded or simulated point spread functions (PSFs) of each object point and each spectral band to be reconstructed.

Iterative reconstruction usually requires extremely high memory requirements and computation times. There are methods to reduce those requirements. However, despite increasing computing power, these algorithms still fail to achieve sufficient reconstruction speed for real-time applications. Our implementation of the EM algorithm as utilized for benchmarking purposes assumes a spatially invariant PSF and uses GPU acceleration which significantly reduces the computation time per iteration. Recently, White et al.²³ published their iterative reconstruction algorithm with GPU acceleration. Depending on the number of iterations, computation times from 0.5 (10 iterations) to 153 s (3100 iterations) were achieved for 24-spectral channels.

The reconstruction of hyperspectral images from CTIS images has lately been demonstrated by Huang et al.²⁹ using a convolutional neural network. Also Douarre et al.³⁰ introduced CTIS-Net for image classification using CTIS images. In this paper, we present an application-independent approach using precise reshaping inside a neural network architecture for the reconstruction of hyperspectral images from CTIS images. Our method demonstrates a clear superiority over EM with respect to processing time and robustness under noisy conditions. By applying prior knowledge about the system's PSF the complexity of the network can be significantly reduced.

3 Neural Networks for CTIS

In CTIS, the spatial information (zeroth diffraction order) and the spectral information (higher diffraction orders) are projected to different locations on the image sensor whereas large areas remain unused. These circumstances should be considered when designing a neural network for processing CTIS images.

Douarre et al.³⁰ took into account the characteristics of a CTIS image and separated the zeroth diffraction order from the higher diffraction orders. The higher diffraction orders are cropped and rotated to align them next to each other. The realigned diffractions are then processed by rectangular convolution kernels. The zeroth diffraction order is processed independently from the higher ones and is concatenated later. However, by rotating the higher diffraction orders, the underlying spatial information is also rotated and the neural network had the burden to learn to connect the right areas from each diffraction order.

Huang et al.²⁹ cropped the different diffraction orders and stacked them on top of each other as a new dimension. Since they only worked with horizontal and vertical diffractions, five sub-images were cropped. The cropping eliminates most of the empty sensor space and no spatial information is altered due to rotation. However, the corresponding spectral information for one pixel is still far apart on the different layers, which requires additional and unnecessary work for the neural network and could be avoided by applying prior system knowledge to the CTIS image.

3.1 Proposed Architecture

In our proposed projection to cube (p2cube) architecture, the higher diffraction orders are reshaped into cubes before convolutions are applied. The reshaping is done in a way that classic convolution can connect areas where the corresponding spectral information is distributed. The idea is adapted from deformable convolutions where the kernels can connect different regions of the input.³¹ While the deformation in deformable convolutions is learned during backpropagation we apply a fixed deformation since the pattern is known from the PSF. Even though this reshaping is rather simple, it plays a pivotal role in a robust reconstruction of the hyperspectral cube from the CTIS image.

The p2cube architecture is shown in Fig. 3 with a brief overview of the applied reshaping in the red box. Here, the input image is reshaped into 13 cubes. One cube for every higher diffraction order and one cube, where the zeroth diffraction order is replicated to fit the spectral dimensions of the other cubes and it provides the spatial information. The reshaped cubes are then processed first by a 3D deconvolution layer and then four convolutional layers which together form the baseline architecture of our approach. We added a U-Net³² like extension

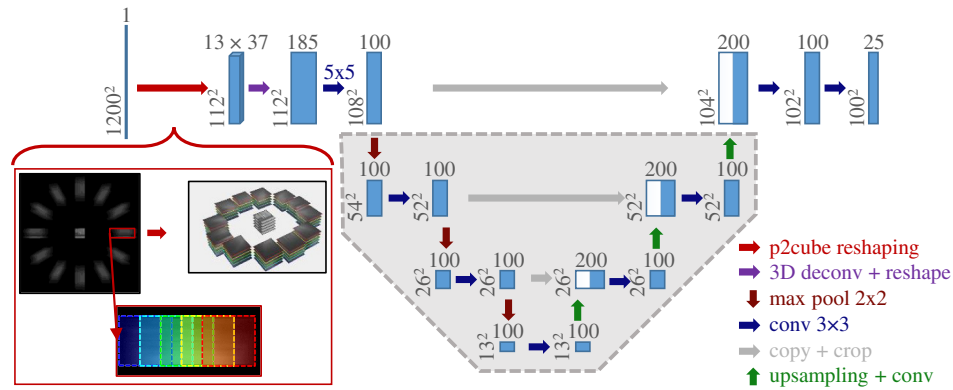


Fig. 3 Schematic visualization of the p2cube architecture with the reshaping layer visualized in more detail in the red box, and the U-Net like extension in the gray box, the combination of both architectures is referred to as p2cubeU.

Table 1 Trainable parameters of the neural networks and number of iterations for the EM algorithm with the resulting computation time. The EM algorithm is running on an Nvidia RTX 2070 SUPER and the neural networks are tested on an Nvidia GTX 1060 6 GB.

	EM	p2cube	p2cubeU
Parameters	—	575,291	1,566,091
Iterations	200	—	—
Computation time	11 s	0.14 s	0.17 s

shown in the gray box and the combined architecture will hereafter be referred to as p2cubeU. The number of trainable parameters from the two architectures is shown in Table 1. In comparison, the architecture from Huang et al.²⁹ has over 85 million parameters used to reconstruct hyperspectral images with a comparable size and the architecture from Douarre et al.³⁰ uses around 16 million parameters for a classification task on CTIS images.

The reshaping is visualized in more detail in Fig. 4 for two higher diffraction orders. Several squares are cropped and stacked on top of each other. The step size between the squares can be controlled in the model. A step size of five pixels shows good results in combination with a five-by-five kernel size. On the right side of Fig. 4, the reshaped cubes are shown. The white squares visualize the area which is connected by a convolution kernel. The connected area from the cube is projected back to the original diffraction order. It can be seen that the classic convolution, which is applied to the cube, connects exactly the area where the spectral information from a pixel in the object space is distributed along. This way even the diagonal diffraction orders can be processed by rectangular convolution kernels. By applying a 3D deconvolution to all thirteen cubes, all areas where the spectral information is distributed are connected.

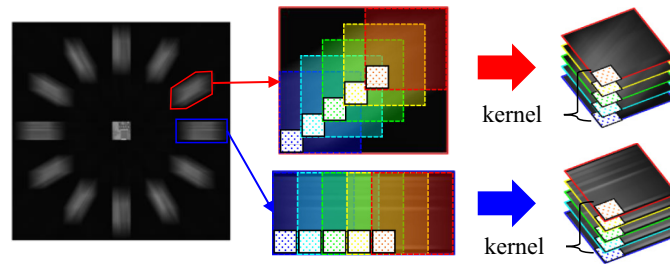


Fig. 4 p2cube reshaping in more detail: only five layers are shown for better visualization. The area connected by a convolution kernel corresponds to the area in which the spectral information is distributed. The size of the kernel is exaggerated for better visualization.

3.2 Data Generation

The training data consists of sensor images as they are captured with a CTIS system together with the hyperspectral representation of the scenes (ground truth). In the following, we describe the synthetic data generation pipeline and our CTIS system that was used to acquire a real dataset with sensor images and corresponding hyperspectral ground truth.

3.2.1 CTIS-Emulation for synthetic data generation

We created a synthetic data set by emulating the CTIS system using Fourier optics to calculate the PSF for every wavelength (color) channel as described in Ref. 33.

A monochromatic point source is considered as a scene in the far-field, which results in a uniform wavefront in the Fourier space. The CGH introduces a spatially dependent phase shift according to the optical path length. Our binary CGH creates a phase shift of π at 550 nm. The image in the Fourier space is cropped differently for every wavelength to keep the pixel size in the image plane constant and, therefore, allows the simulation of the CTIS image without interpolation in the spatial dimension. We use a sensor resolution of 1200×1200 pixels. An inverse Fourier transformation leads to the wavefront in the image plane. By taking the square of the complex magnitude from the wavefront, the intensity at the sensor (PSF) is obtained. This has to be done for every wavelength. Afterward, each wavelength-dependent PSF is convolved with the corresponding spectral band from the hyperspectral cube. By summing up all spectral channels the final grayscale sensor image is obtained. We use a spectral resolution of one nanometer to create the PSF. The hyperspectral images are interpolated in the spectral dimension to fit the calculated PSF. It has to be noted that this method does not take optical aberrations into account and assumes a spatially invariant system. Only diffraction due to the finite lens diameter is considered.

Hyperspectral images from three different publicly available datasets are used: BGU iCVL Hyperspectral Image Dataset,³⁴ CAVE Multispectral Image Database,³⁵ and TokyoTech 31-band Hyperspectral Image Dataset.¹¹ We extracted the bands from 450 to 690 nm with a spectral resolution of 10 nm, which leads to 25 spectral channels. The original images are first randomly split into training (76.5%), validation (19%), and test (4.5%) sets. Then, using a sliding window, nine subimages were extracted from a given image and then resized to a spatial resolution of 100×100 pixels. A tenth image is the whole original one but resized to 100×100 pixels also. This way the amount of samples used to train the network is increased by a factor of ten to a total number of 2585 hyperspectral images.

3.2.2 Real world CTIS data

Even though the emulation of the CTIS system using Fourier optics takes into account effects such as wavelength dependent diffraction efficiency, it does not include all optical aberrations from the lenses and internal reflections in the setup. To test the performance of the proposed neural network for real-world data, a real dataset of CTIS images paired with hyperspectral ground truth cubes was captured. A schematic of the expanded optical setup is shown in Fig. 5.

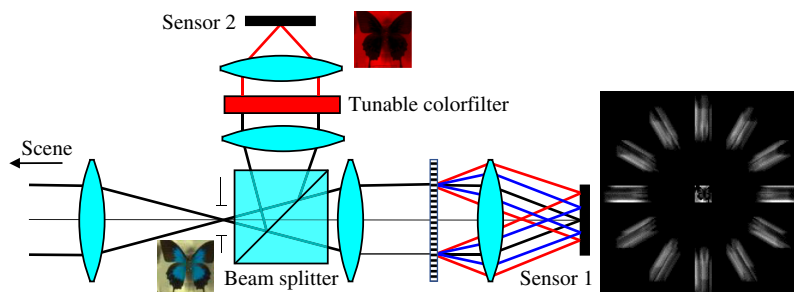


Fig. 5 Schematic of our setup to capture CTIS images and corresponding hyperspectral cubes. By inserting a beam splitter after the field stop, both sensors capture the intermediate image during the field stop. The tunable color filter allows to record the spectral channels from 455 nm to 695 nm with a 10-nm step size.³⁶

A second optical path is added by inserting a beam splitter after the field stop. There, the intermediate image is again collimated and spectrally filtered by a tunable color filter.³⁶ A second sensor records the spectral channels sequentially. All images were recorded from 455 to 695 nm with a step size of 10 nm resulting in 25 spectral channels. The CTIS images have a resolution of 1088×1088 pixels and the hyperspectral ground truth images were downsampled to match the resolution of the zeroth diffraction order of 139×139 pixels. The recorded real-world dataset consists of 495 images that were split into 400 images for training, 80 for validation, and 15 images for testing.

3.3 Training Details

Keras with TensorFlow 2 and GPU support is used to build, train, and test the neural network.³⁷ The employed loss function for training and validation is the mean squared error (MSE). Adam optimizer³⁸ is used for training. The last convolutional layer uses a rectified linear unit as an activation function whereas the other layers use the default configuration of TensorFlow. The images are normalized to $[0, 1]$.

After splitting the images as described in Sec. 3.2, no further data augmentation is applied. However, a rotation by 90 deg or a multiple of 90 deg would be possible for the synthetic data set. Since the real-world data set contains aberrations that are not rotationally symmetric, the training would probably not benefit from such augmentation.

To simulate realistic camera noise for the synthetic dataset, a dynamic shot noise layer is added.³⁹ A quantum full-well capacity of 1000 is set. To do this, the normalized image is scaled to 1000. Then the shot noise is applied and the image is then divided by 1000 so that the values are again between zero and one.

4 Results

In the following, only results that are part of the test data set and thus not utilized during training or validation are discussed. In addition, to assess the influence of noise, we use one data set without noise and another data set consisting of the same images with shot noise added to the detector images. The neural network is trained with and without noise accordingly. For the EM algorithm, a noise-free calibration (PSF image) is used for both sets. The number of iterations for the EM algorithm was fixed to 200 since the performance stagnated afterward. Hyperspectral reconstruction from real CTIS images with the EM algorithm was not successful and is therefore not presented. The failure of the EM algorithm is due to a higher noise level in the sensor images and the requirement of a spatially invariant PSF, which is not fulfilled by the real CTIS setup.

To evaluate the reconstruction performance, we used three different quality criteria as presented in Tables 2 and 3. An individually optimized scaling factor is applied to the reconstructed images from the EM algorithm to deliver the best result. This is necessary since the sensor images are normalized and the total energy is lost. This is valid because only relative and not absolute values are of interest anyway. As the first quantitative criterion, we used the root mean squared error (RMSE). Since the data range is between 0 and 1, the error can be directly

Table 2 Quantitative comparison between EM, p2cube, and p2cubeU architectures for synthetic data. The values labeled with noise are derived from the noisy sensor images. For RMSE and MAE a small value close to 0 shows good agreement of the reconstruction with the ground truth data. For MeanSSIM, in contrast, a value close to 1 shows a good structural similarity of the data.

Metrics	EM	p2cube	p2cubeU	EM noise	p2cube noise	p2cubeU noise
RMSE	0.014	0.017	0.012	0.090	0.021	0.016
MAE	0.009	0.012	0.008	0.066	0.015	0.011
MeanSSIM	0.950	0.940	0.971	0.437	0.901	0.939

Note: Best metrics are highlighted in bold.

Table 3 Quantitative results on real CTIS images using p2cube and p2cubeU architectures. Reconstruction with EM failed in this case since the PSF is not spatially invariant and the CTIS image is too noisy.

Metrics	p2cube	p2cubeU
RMSE	0.042	0.028
MAE	0.026	0.018
MeanSSIM	0.892	0.938

Note: Best metrics are highlighted in bold.

interpreted. Similar to the RMSE is the mean absolute error (MAE), which, in contrast to the RMSE, takes individual strong deviations (peaks) less into account. The mean structural similarity index measure (MeanSSIM) on the other hand is a completely different metric that is usually used to determine the quality of color images. In our case, we use the mean values over all channels. This, therefore, does not account for spectral characteristics anymore.⁴⁰ However, we consider it is still a valuable metric since it is different than the others. Respective results are shown in Table 2.

It can be seen that p2cubeU delivers the best results for all metrics for noise-free and noisy images. The EM algorithm is only slightly worse for the noise-free case. For noisy images, however, the quality of the images reconstructed with the EM algorithm decreases significantly. Here, the p2cube and p2cubeU architectures show their strength and can deliver results that are close to the noise-free ground truth images.

Qualitative comparisons are shown in Figs. 6 and 7. The RGB representation of the different images is obtained via the CIE 1931 filter response functions. Notice that the reconstruction from noise-free CTIS images works pretty well with all three methods. However, all images show some imperfections. For the EM reconstruction, one can see straight artifacts that correspond to the direction of diffraction of the DOE. This is typical and can often be observed with such kind of algorithms. With the neural network, it is sometimes noticeable that the color of small details is not reproduced correctly, e.g., the small stripes in the lower left part of the butterfly wing. For the emulated images with noise, it can be seen that the neural network-based approach delivers better results since the images from the EM algorithm become very noisy in the spectral dimension.

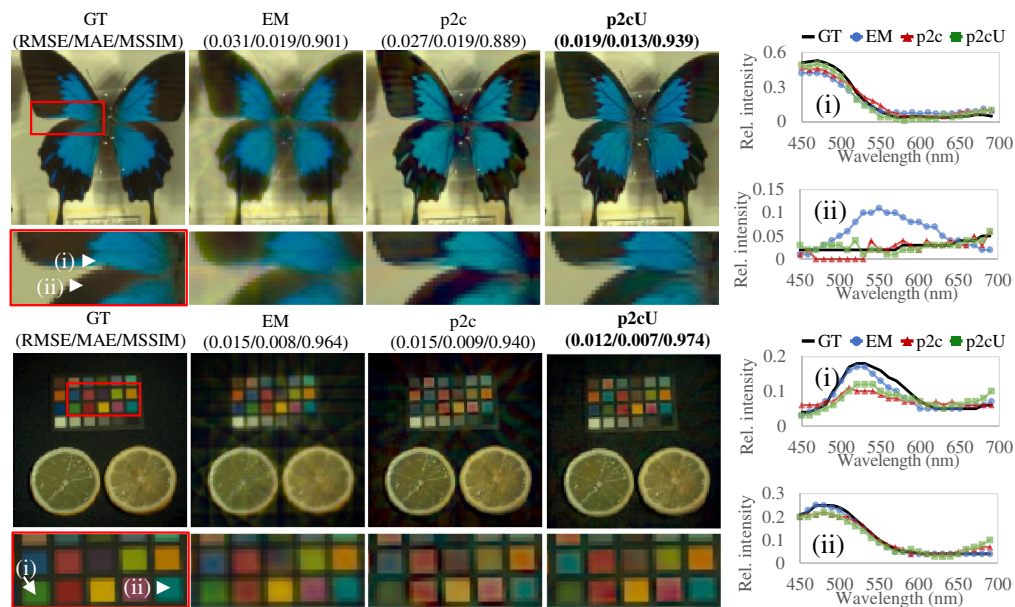


Fig. 6 RGB representation of the ground truth and the reconstructed hyperspectral images from the synthetic noise-free sensor images. The images are 100×100 pixels and have 25 channels. Spectra of two points per image are shown on the right.

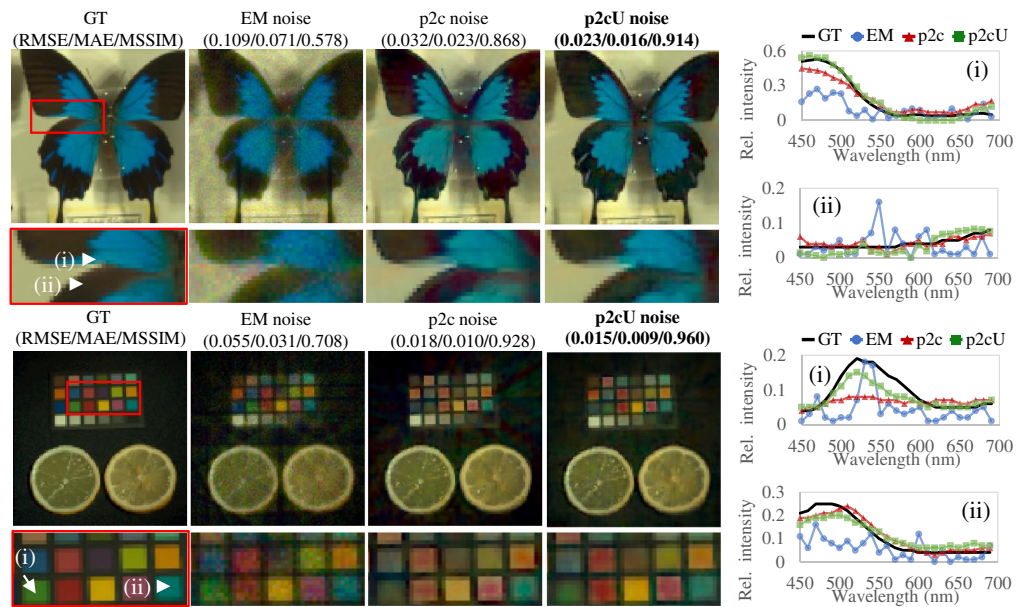


Fig. 7 RGB representation of the ground truth and the reconstructed hyperspectral images from the synthetic noisy sensor images. The images are 100×100 pixels and have 25 channels. Spectra of two points per image are shown on the right.

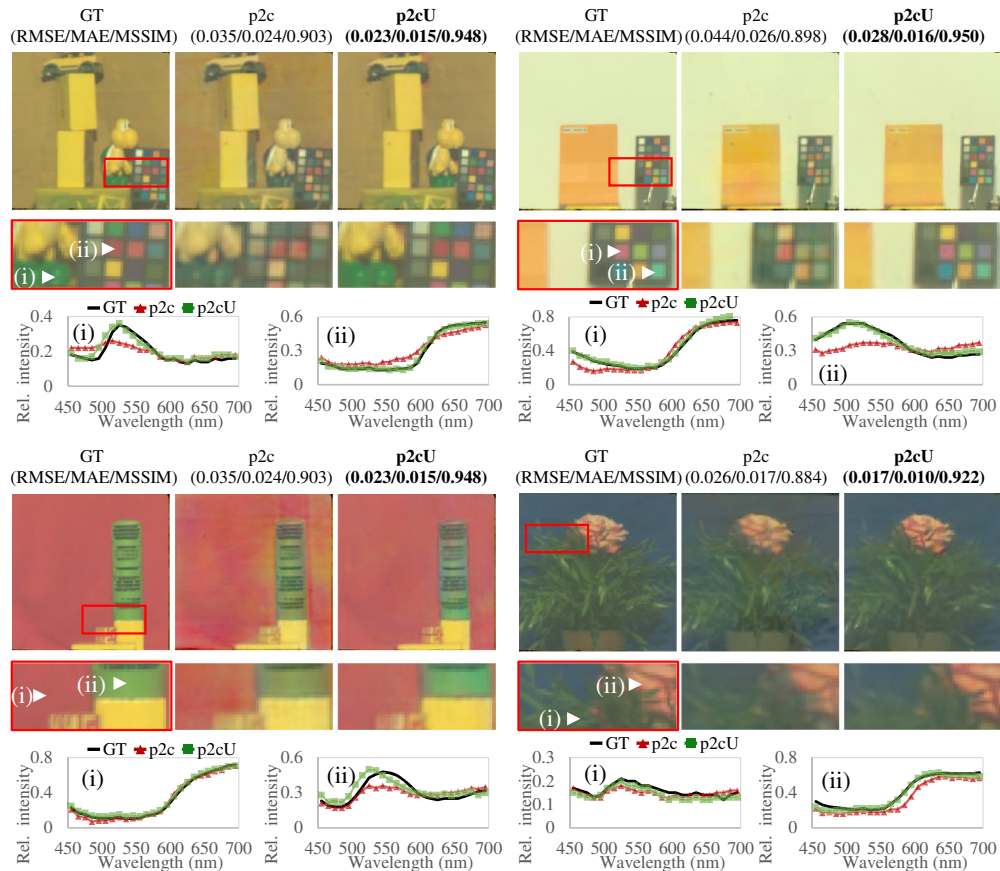


Fig. 8 RGB representation of the ground truth and the reconstructed hyperspectral images from the real-world images. The images are 139×139 pixels and have 25 channels.

In Fig. 8, reconstructed samples from the real CTIS data are shown. Notice that the images from p2cube and p2cubeU show very few artifacts in the spatial dimension. Furthermore, p2cubeU achieves better performance and produces well-reconstructed images inline with the corresponding ground truth.

5 Conclusion and Outlook

In this paper, we presented a deep convolutional neural network for hyperspectral image reconstruction from synthetic as well as real CTIS data. Since spectral and spatial information overlap on the sensor image, it is rearranged in a first step so that the subsequent convolutional layers can better link spatially related areas. This reshaping organizes the CTIS image in a more efficient way for the network than others have demonstrated so far and it ensures a robust training/reconstruction with over 30 times less parameters than Douarre et al.³⁰ for their classification task and 100 times less parameters than Huang et al.²⁹ for their hyperspectral image reconstruction. Our approach produces high-quality reconstruction results on synthetic and real CTIS data.

The advantages of the proposed neural network are the reduced computation time for inference by a factor >60 compared with the conventional EM algorithm and the reduced complexity compared with other networks dealing with CTIS data.^{29,30} With a computation time between 0.14 and 0.17 s per image on a GTX 1060 6 GB, it enables real-time performance with common hardware.

In our experiments, the spatial resolution of the hyperspectral images is the same as the spatial resolution of the zeroth diffraction order. However, it should be possible to generate higher spatial resolution images from CTIS images since the spatial information is also multiplexed in the higher diffraction orders. This will be the focus of future work.

Acknowledgments

This work was carried out during research cooperation between the Computational Sensing Group at the Stuttgart Technology Centre of Sony Europe B.V., the Institut für Technische Optik at the University of Stuttgart, and the Department of Information Engineering of the University of Padua.

References

1. C. R. Kitchin, *Optical Astronomical Spectroscopy*, Institute of Physics Pub, Bristol; Philadelphia (1995).
2. G. Vane and A. F. Goetz, "Terrestrial imaging spectroscopy," *Remote Sens. Environ.* **24**, 1–29 (1988).
3. G. Lu and B. Fei, "Medical hyperspectral imaging: a review," *J. Biomed. Opt.* **19**, 010901 (2014).
4. A. Gowen et al., "Hyperspectral imaging – an emerging process analytical tool for food quality and safety control," *Trends Food Sci. Technol.* **18**, 590–598 (2007).
5. M. B. Stuart, A. J. S. McGonigle, and J. R. Willmott, "Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems," *Sensors* **19**(14), 3071 (2019).
6. A. F. Goetz, "Three decades of hyperspectral remote sensing of the earth: a personal view," *Remote Sens. Environ.* **113**, S5–S16 (2009).
7. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Opt. Eng.* **52**, 090901 (2013).
8. T. Okamoto and I. Yamaguchi, "Simultaneous acquisition of spectral image information," *Opt. Lett.* **16**, 1277 (1991).
9. T. V. Bulygin and G. N. Vishnyakov, "Spectrotomography: a new method of obtaining spectrograms of two-dimensional objects," *Proc. SPIE* **1843**, 315–322 (1992).
10. R. Habel, M. Kudenov, and M. Wimmer, "Practical spectral photography," *Comput. Graph. Forum* **31**, 449–458 (2012).

11. Y. Monno et al., "A practical one-shot multispectral imaging system using a single image sensor," *IEEE Trans. Image Process.* **24**, 3048–3059 (2015).
12. M. Descour and E. Dereniak, "Computed-tomography imaging spectrometer: experimental calibration and reconstruction results," *Appl. Opt.* **34**, 4817 (1995).
13. M. A. Seldowitz, J. P. Allebach, and D. W. Sweeney, "Synthesis of digital holograms by direct binary search," *Appl. Opt.* **26**, 2788–2798 (1987).
14. N. Hagen, E. L. Dereniak, and D. T. Sass, "Maximizing the resolution of a CTIS instrument," *Proc. SPIE* **6302**, 168–178 (2006).
15. N. Hagen and E. L. Dereniak, "New grating designs for a CTIS imaging spectrometer," *Proc. SPIE* **6565**, 216–224 (2007).
16. N. Hagen and E. L. Dereniak, "Analysis of computed tomographic imaging spectrometers I Spatial and spectral resolution," *Appl. Opt.* **47**, F85 (2008).
17. M. W. Kudenov, "Faceted grating prism for a computed tomographic imaging spectrometer," *Opt. Eng.* **51**, 044002 (2012).
18. N. Hagen, E. L. Dereniak, and D. T. Sass, "Fourier methods of improving reconstruction speed for CTIS imaging spectrometers," *Proc. SPIE* **6661**, 15–25 (2007).
19. M. D. Horton, "A novel technique for CTIS image-reconstruction," PhD thesis, University of Tennessee, Knoxville (2010).
20. L. Sethaphong, "Large format ctis in real time: parallelized algorithms and preconditioning initializers," Master's thesis, Vanderbilt University, Nashville (2007).
21. T. J. Thompson, "Accelerated CTIS using the cell processor," Master's thesis, University of Tennessee, Knoxville (2009).
22. M. D. Vose and M. D. Horton, "A heuristic technique for CTIS image reconstruction," *Appl. Opt.* **46**, 6498–6503 (2007).
23. L. White, W. B. Bell, and R. Haygood, "Accelerating computed tomographic imaging spectrometer reconstruction using a parallel algorithm exploiting spatial shift-invariance," *Opt. Eng.* **59**, 055110 (2020).
24. B. K. Ford et al., "Computed tomography-based spectral imaging for fluorescence microscopy," *Biophys. J.* **80**, 986–993 (2001).
25. J. M. Mooney et al., "High-throughput hyperspectral infrared camera," *J. Opt. Soc. Am. A* **14**, 2951 (1997).
26. C. E. Volin et al., "Midwave-infrared snapshot imaging spectrometer," *Appl. Opt.* **40**, 4501–4506 (2001).
27. C. C. Kankelborg and R. J. Thomas, "Simultaneous imaging and spectroscopy of the solar atmosphere: advantages and challenges of a 3-order slitless spectrograph," *Proc. SPIE* **4498**, 16–26 (2001).
28. E. K. Hege et al., "Hyperspectral imaging for astronomy and space surveillance," *Proc. SPIE* **5159**, 380–391 (2004).
29. W.-C. Huang et al., "The application of convolutional neural networks for tomographic reconstruction of hyperspectral images," *Displays* **74**, 102218 (2022).
30. C. Douarre et al., "CTIS-Net: a neural network architecture for compressed learning based on computed tomography imaging spectrometers," *IEEE Trans. Comput. Imaging* **7**, 572–583 (2021).
31. J. Dai et al., "Deformable convolutional networks," in *IEEE Int. Conf. Comput. Vision (ICCV)*, pp. 764–773 (2017).
32. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
33. A. Vijayakumar and S. Bhattacharya, *Design and Fabrication of Diffractive Optical Elements with MATLAB*, SPIE (2017).
34. B. Arad and O. Ben-Shahar, "Sparse recovery of hyperspectral signal from natural RGB images," *Lect. Notes Comput. Sci.* **9911**, 19–34 (2016).
35. F. Yasuma et al., "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.* **19**, 2241–2253 (2010).
36. P. Inc., "Varispec effortlessly tune to any wavelength in the vis or NIR range—without moving parts." https://www.perkinelmer.com.cn/CMSResources/Images/46-140156DTS_010053A_01_VariSpec_DTS.pdf (2013) (accessed 12 November 2020).

37. M. Abadi et al., “TensorFlow: large-scale machine learning on heterogeneous systems,” Software available from [tensorflow.org](https://www.tensorflow.org) (2015).
38. D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” *CoRR* **1412.6980** (2015).
39. E. Posner, “How to create awesome noise that is actually real.” <https://medium.com/datadriveninvestor/cf178c9f0ae0> (2019) (accessed 12 November 2020).
40. R. Zhu, F. Zhou, and J.-H. Xue, “Mvssim: a quality assessment index for hyperspectral images,” *Neurocomputing* **272**, 250–257 (2018).

Markus Zimmermann is a researcher at the Institut für Technische Optik of the University of Stuttgart. He received his master’s degree from the University of Stuttgart in Photonic Engineering in 2020. His main research interests are in the field of hyperspectral imaging and holographic display technologies.

Simon Amann is a PhD student at the Institut für Technische Optik of the University of Stuttgart. Here, he completed his master’s degree in photonic engineering in 2017. His area of expertise is optical metrology with the primary focus on hyperspectral imaging.

Mazen Mel received his engineering degree in telecommunications from the Higher School of Communications of Tunis in 2019 and his MSc degree in telecommunication from the University of Padova in 2021. He is currently a PhD student at the Department of Information Engineering of the University of Padua. His research focuses on computational photography, in particular hyperspectral imaging, depth estimation, and depth-of-field extension.

Tobias Haist studied physics and received his PhD in engineering from the University of Stuttgart. Currently, he is leading the group 3D surface metrology at the Institut für Technische Optik, where he is working on new applications for spatial light modulators and 3-D measurement systems. His main research interests include optical and digital image processing, computer-generated holography, and optical measurement systems.

Alexander Gatto studied physics and computer science and received his PhD in physics from the University of Bonn. Currently, he is leading an R&D group at the Stuttgart Technology Center of Sony Europe B.V. His research focus is on computational photography and new imaging approaches.