Data in Brief 48 (2023) 109094

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

RGB-D microtopography: A comprehensive dataset for surface analysis and characterization techniques

Stefan Siemens*, Markus Kästner, Eduard Reithmeier

Institute of Measurement and Automatic Control, Leibniz University Hannover, An der Universität 1, 30823 Garbsen, Germany

A R T I C L E   I N F O

Article history:
Received 14 February 2023
Revised 20 March 2023
Accepted 21 March 2023
Available online 28 March 2023

Dataset link: Micro-Topo-Dataset (Original data)

Keywords:
Microtopography
Surface roughness
Surface classification
Optical metrology
Confocal laser scanning microscopy
Computer vision

A B S T R A C T

The dataset presented contains microtopographies of various materials and processing methods. These microtopographies were measured using a Confocal Laser Scanning Microscope, which provides RGB-D data. This means the dataset comprises accurate height maps for each measurement and microscopic RGB images. The height maps can be used to quantify and characterize small-scale surface features such as pits and grooves, surface roughness, texture direction, and surface anisotropy. These features can significantly impact a material's properties and behavior, making them essential in many fields, such as biomaterials and tribology. Additionally, the dataset contains metadata about the specimens and the measurement conditions, such as material, surface processing method, roughness, and optical magnification. Therefore, this dataset provides an opportunity to develop and test surface classification and characterization algorithms.

* Corresponding author.
E-mail address: stefan.siemens@imr.uni-hannover.de (S. Siemens).
Social media: @S_Siem (S. Siemens)

https://doi.org/10.1016/j.dib.2023.109094

Specifications Table

Subject	Engineering, Metrology, and Computer Vision.
Specific subject area	Development and validation of algorithms to characterize and classify microtopographies and train deep neural networks.
Type of data	Metadata and the individual raw measurements. Each measurement consists of the following: • an RGB image (8 bits per channel) • a height map (32-bit) • a laser-intensity image (16-bit) • and a combined RGB+laser-intensity image (8 bits per channel) A demo Jupyter notebook for opening the CSV metadata and loading a measurement.
How the data were acquired	The data were acquired in a climate-controlled laboratory at 20°C. The measurement system is a Confocal Laser Scanning Microscope. The model is Keyence VK-X210. The measurement software of the manufacturer (VK Viewer) was used for data acquisition. Results were saved in the proprietary .vk4 data format. The .vk4 data was loaded into Python using [1] and [2] and then saved in the open .npy format using a self-written Python script.
Data format	CSV (Metadata) NPY (Raw measurements) PNG (Figures) IPYBN (Demo script) Link to the repository: https://data.uni-hannover.de/dataset/micro-topo
Description of data collection	The dataset includes 11,757 measurements of 167 microtopography specimens of 8 materials and 23 processing methods. For some specimens, the material and or processing method are not specified. For specimens manufactured to a specific roughness, the values are listed in the metadata.
Data source location	• Institution: Institute of Measurement and Automatic Control • City/Town/Region: Hannover • Country: Germany • Latitude and longitude: 52.39020000296517, 9.706996618887981
Data accessibility	Raw data were deposited at the Leibniz University Hannover Research Data Repository. Repository name: Micro-Topo-Dataset Data identification number: https://doi.org/10.25835/pngok8op Direct URL to data: https://data.uni-hannover.de/dataset/micro-topo

Value of the Data

Confocal Laser Scanning Microscopy (CLSM) is a well-established technique for characterizing and measuring surface features, such as roughness [3]. Therefore, the dataset can be useful to researchers in a variety of fields, including:

- Surface metrology:
  - Evaluation and improvement of conventional surface characterization techniques.
  - Development of new approaches for the characterization and measurement of microtopographies [4].
  - Comparison of algorithms with those defined in standards, such as ISO 25178 [5].
  - Comparison of different models or manufacturers of confocal laser scanning microscopes.
  - Comparison of a commercially available CLSM with a custom-built system.
  - Comparison of different measurement methods like focus variation or fringe light projection.
- Materials science and engineering:
  - Study of the effects of different surface processing methods on microtopography.
- Computer vision and machine learning:
  - Development and testing of algorithms for surface classification and reconstruction.

- Training of neural networks using labels provided with the dataset (e.g., material, processing method, or roughness) [6].
- Development of novel approaches for problems such as monocular depth estimation using the height and RGB measurement data.
- Testing the robustness or generalizability of algorithms developed on different datasets for tasks such as image registration.

1. Objective

The dataset was created as part of the research project *Registration of Optically Acquired Microtopographies Using Machine Learning* at the Institute of Measurement and Automatic Control. It provides a comprehensive resource for researchers in fields like surface metrology by offering a large, diverse, high-quality set of microtopographies typical for different technical applications. The initial intention of creating the dataset was to improve the registration of microtopography measurement data in terms of robustness and accuracy by training neural networks. Beyond that, the dataset can be used for various purposes, including evaluating conventional techniques, such as areal parameters defined in ISO 25178, and new approaches to surface characterization, like learning-based solutions.

2. Data Description

This dataset comprises five parts: the metadata, an Excel file explaining the metadata, the measurement data, a demo script, and supplementary files.

2.1. Metadata

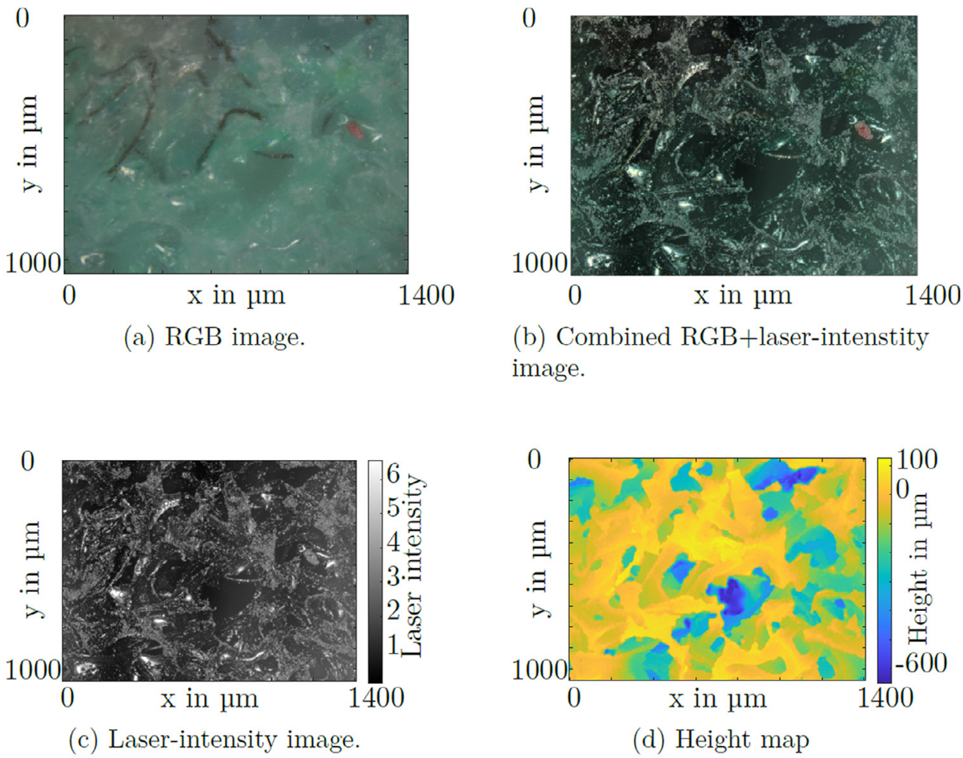
The metadata is stored in *dataset.csv*. This file can be opened with any text editor, but Pandas, for example, is recommended when Python is used for processing the data. Each row represents a single measurement, while the columns hold the corresponding metadata. The first four columns contain the raw measurement data file names (see section Measurement data), while the remaining columns store information about the measurement condition and the specimen. The naming and meaning of each entry are explained in *dataset-layout.xlsx*. Of particular importance are entries regarding the objective lens and the specimen. These entries are required to obtain metrical dimensions from the measurements and get the specimen's nominal roughness values.

2.2. Measurement Data

The measurement data is stored in 12 ZIP files named *measurement_data_part_1.zip* through *Measurement_data_part_12.zip*. These files' contents must be unpacked into the same directory as *dataset.csv*. As described above, each measurement consists of four data or image types, see Fig. 1. Each of these data types is stored in a separate .npy file. The naming convention for the files is

$$\{specimen\}\_\{material\}\_\{processing\}\_\{magn.\}x\_\{NA\}\_\{yyyy\} - \{mm\} - \{dd\}\_\{img\,type\}\_\{k\}.npy,$$

where k is an iterator. The .npy files can be opened and displayed, for example, with Python or Matlab.



(a) RGB image.
(b) Combined RGB+laser-intenstity image.
(c) Laser-intensity image.
(d) Height map

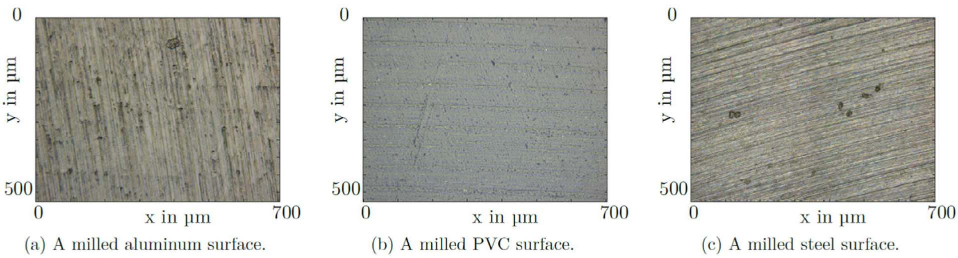
Fig. 1. Four image types acquired during a CLSM measurement.

2.3. Demo Script

The Jupyter notebook called *display_dataset.ipynb* can be used to load and display *dataset.csv*, select and display measurements, and show the corresponding metadata.

Fig. 1 shows the four data types resulting from each measurement: an RGB image, a laser-intensity image, a combined RGB+laser-intensity image, and a height map. This measurement was taken with a 10x magnification and NA = 0.30.

Fig. 2 shows RGB images of three exemplary specimens with the same processing method (milling) but different materials (aluminum, PVC, and steel). All images were taken with 20x magnification and NA = 0.46.



(a) A milled aluminum surface.
(b) A milled PVC surface.
(c) A milled steel surface.

Fig. 2. Exemplary RGB images of milled specimens with different materials.

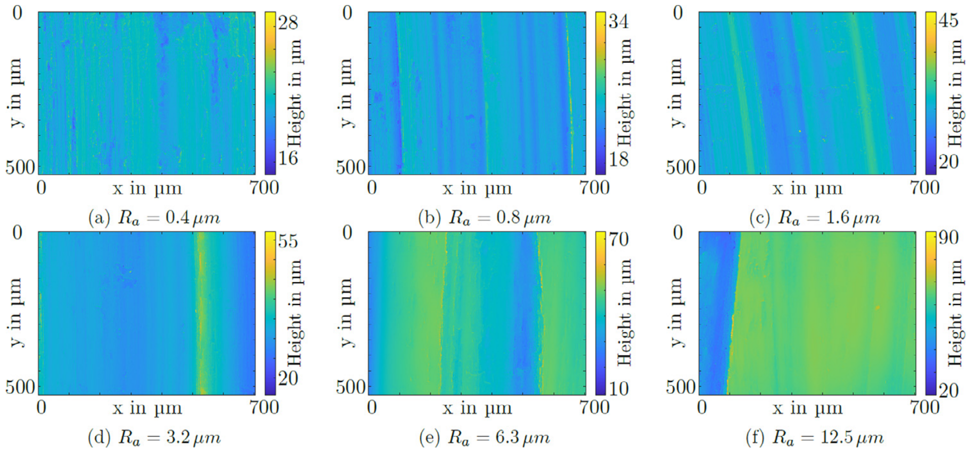


Fig. 3. Exemplary height maps of face milled surfaces with different roughness values.

Fig. 3 shows height maps of exemplary specimens with the same material (nickel-plated steel) and processing method (face milling) but different roughness values. All images were taken with a 20x magnification and NA = 0.46.

Table 1 shows the specifications of the microscopic lenses used for data acquisition.

Table 1
Specifications of the microscopic lenses.

Objective	Magn.	NA	Working dist.	Focal len.	Depth of focus	xy res.
CF IC EPI Plan 2.5xA	2.5x	0.075	8.8 mm	80 mm	48.89 µm	5.274 µm
CF IC EPI Plan 10xA	10x	0.30	16.5 mm	20 mm	3.06 µm	1.318 µm
CF IC EPI Plan 20xA	20x	0.46	3.1 mm	10 mm	1.3 µm	0.659 µm
CF IC EPI Plan ELWD 50xA	50x	0.55	8.7 mm	4 mm	0.91 µm	0.264 µm
CF IC EPI Plan Apo 50xA	50x	0.95	0.35 mm	4 mm	0.3 µm	0.264 µm
CF IC EPI Plan Apo 150xA	150x	0.95	0.2 mm	1.33 mm	0.3 µm	0.088 µm

Table 2 shows exemplary materials and processing methods contained in the dataset.

Table 2
Exemplary specimens.

Processing Method	Material	Number of measurements
Circumferential milling (R_a = 0.7 µm)	Steel	156
Circumferential milling (R_a = 1.2 µm)	Steel	129
Horizontal milling (R_a = 0.4 µm … R_a = 12.5 µm)	Nickel-plated-steel	226
Thermal spraying	Al_2O_3	2,053
Turning	Aluminum	414
Turning	PVC	547
Turning	Steel	412

3. Experimental Design, Materials and Methods

3.1. Image Acquisition

For this dataset, a Keyence VK-X210 Confocal Laser Scanning Microscope (CLSM) was used to acquire microtopographies. The measurements were taken between 2017 and 2021. The microscope was placed on an optical table to provide a stable, vibration-free platform. The entire setup was located in a climate-controlled laboratory, where the temperature was maintained at 20°C. The CLSM had six objectives with magnifications between 2.5x and 150x. The specifications of the objectives can be seen in Table 1. The microscope has two light sources: a white light source for collecting RGB images with a CCD sensor and a photomultiplier tube combined with a laser (wavelength of 408 nm) for acquiring height data. The measurements typically have xy dimensions of 1024 by 768 pixels. Measurements with differing dimensions are in the dataset as well. For example, the microscope supports a fine mode, whereby pixel shift increases the dimensions to 2048 by 1536 pixels.

3.2. Materials

The CLSM was used to examine specimens with various materials and processing methods and specimens with the same processing method but varying roughness levels (from $R_a = 0.0125$ µm to $R_a = 25$ µm). In total, the dataset consists of 11,757 individual measurements. Examples are listed in Table 2 and depicted in Figs. 2 and 3.

3.3. Calculating Height (z-Axis)

The height data is expressed relative to the measurement start point or a defined lower limit, the origin (zero point). To determine the height difference between two points, *A* and *B*, the height data has to be multiplied by the "*z calibration*" value:

$$\text{Height difference between } A \text{ and } B \text{ in nm} = (\text{intensity at point } A - \text{intensity at point } B) \cdot (z\ calibration)$$

The "*z calibration*" for each measurement can be found in *dataset.csv* under *z_length_per_digit*. The default is $1 \cdot 10^{-1}$.

3.4. Calculating Distance (x-Axis and y-Axis)

To determine the distance (length) between two points, *A* and *B*, in a vertical (y) or horizontal (x) direction, the number of pixels between the two points has to be multiplied by the "*xy calibration*" value:

$$\text{Distance between two points in nm} = (\text{No. of pixels between } A \text{ and } B) \cdot (xy\ calibration).$$

The "*xy calibration*" for each measurement can be found in *dataset.csv* under *x_length_per_pixel* and y*_length_per_pixel*. Notice that these values are the same, as the pixels are square.

Ethics Statements

This work did not include work involved with human subjects, animal experiments, or data collected from social media platforms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Micro-Topo-Dataset (Original data) (Leibniz University Research Data Repository).

CRediT Author Statement

Stefan Siemens: Conceptualization, Methodology, Software, Investigation, Data curation, Writing – original draft, Visualization; **Markus Kästner:** Writing – review & editing, Supervision, Project administration; **Eduard Reithmeier:** Resources, Supervision, Project administration, Funding acquisition.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2023.109094.

References

[1] J.G. Wendelberger, Extracting the data from the LCM vk4 formatted output file. http://www.osti.gov/servlets/purl/1419732/. Accessed January 19, 2023, doi:10.2172/1419732.
[2] W. Gunn, B. Torkian, vk4-Python-driver, 2018. https://github.com/torkian/vk4-python-driver. Accessed January 19, 2023.
[3] A. Mínguez Martínez, J. de Vicente y Oliva, Industrial calibration procedure for confocal microscopes, Materials 24 (2019) 12, doi:10.3390/ma12244137.
[4] S. Siemens, M. Kästner, E. Reithmeier, Texture direction analysis of micro-topographies using fractal geometry, Surf. Topogr.: Metrol. Properties 10 (2022) 4, doi:10.1088/2051-672X/ac9ad0.
[5] International Organization for Standardization, ISO 25178-3:2012, geometrical product specifications (GPS) — surface texture: areal — part 3: specification operators (2012). doi:10.31030/1876351.
[6] S. Siemens, M. Kästner, E. Reithmeier, Super-resolution for 2.5D height data of microstructured surfaces using the vdsr network, EPJ Web Conf. 238 (2020), doi:10.1051/epjconf/202023806014.