Endorsed by

IPEM
Institute of Physics and
Engineering in Medicine

PAPER • OPEN ACCESS

Focus stacking in non-contact dermoscopy

View the article online for updates and enhancements.

Biomedical Physics & Engineering Express

PAPER

Focus stacking in non-contact dermoscopy

Lennart Jütte[1] , **Zhiyao Yang**[1], **Gaurav Sharma**[1] and **Bernhard Roth**[1,2]

1   Leibniz University Hannover, Hannover Centre for Optical Technologies, Hannover, Germany
2   Leibniz University Hannover, Cluster of Excellence PhoenixD, Hannover, Germany

E-mail: lennart.juette@hot.uni-hannover.de

Abstract

Dermoscopy is the main tool for early detection of skin cancer. Non-contact dermoscopes often suffer from a small depth of field leading to images of skin topographies with regions that are not in focus. We aim to provide an easy-to-implement focus stacking-based approach to ensure all-in-focus images from a non-contact dermoscope. Further, we aim to extract additional information about the skin topography from the image stacks. The focus stacking procedure itself is implemented in a non-contact dermoscope with an electrically adjustable focus realized by using a tunable liquid lens. We show that all-in-focus imaging is possible for non-contact dermoscopy and deliver a method to extract topographical information for dermatologists from the acquired image stacks. Our finding indicate that the approach can be valuable for non-contact dermoscopic examination as well as for the early detection of skin diseases such as cancer as it possible to derive hyperfocus images and information on the skin topography. With this, we were able to develop a software for the acquisition of the raw image data and its processing into a high resolution hyperresolution dermoscopic image. In the next steps, we plan to apply the approach in the clinical environment for skin cancer diagnostics or imaging of inflammatory skin diseases.

1. Introduction

Skin cancer is one of the most common types of cancer. In 2021, approximately 1.9 million cancer cases were diagnosed in the USA alone. Of these, about 5.6% are melanoma, and more than 7000 deaths were reported in conjunction with the latter [1]. In Europe, the number of new cases amounts to more than 140,000 per year. Early detection is vital for melanoma treatment. According to statistical data, the 5-year relative survival rate of melanoma drops rapidly from 99% to below 20% as stages evolve [2].

Dermoscopes are currently the widely used equipment for lesion detection and are usually operated in contact to the skin. Compared to the naked eye, a dermoscope can extract information from below the surface of the skin. There are two types lighting-systems of dermoscopes used in the clinical environment: non-polarized dermoscopes (NPD) and polarized dermoscopes (PD). The main difference between NPD and PD is the depth of visualized structures. While NPD aims to inspect the lesion on the skin surface, PD can filter out the reflection of the skin's surface and reveal subcutaneous structures [3].

When examining the patient's lesion, a contact-based dermoscope is pressed against the skin. The contact increases the risk of lesion rupture and distorts the skin geometry. Furthermore, the pressure may cause change of blood perfusion and therefore might change the color of the lesion as well. However, the geometry and color of the lesion are considered important criteria in melanoma diagnosis according to the well-known ABCDE rule (asymmetry, border, color, diameter, evolution) [4].

Non-contact dermoscopy is proposed to counteract the problems mentioned above. With the help of a liquid lens with tunable focus, non-contact dermoscopy provides a less invasive approach for lesion examination [5–7]. The skin is in its natural state when imaged. A downside to non-contact dermoscopy can be that, for topographies of a depth greater than the depth of field of the imaging system, the skin under study is not always fully in focus. Another possibility for the occurrence of this problem is failure of integrated auto-focus systems or movement of the patients. The problem of some regions being out of focus is visualized in figure 1.
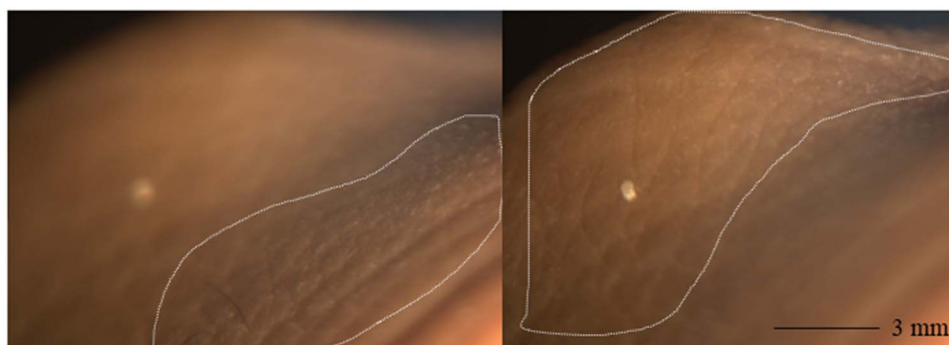
Figure 1. Visualization of in-focus regions in a stack of two images of an ex vivo melanoma obtained with the non-contact dermoscope described in figure 6. The regions marked by the white contour are in focus while the other regions are not. The countours were determined subjectively. Automated detection is possible as well and will be implemented in a later stage.

In this work, we report on a method of obtaining dermoscopic images with full focus for all skin topographies by the employment of focus stacking. Furthermore, we show the possibility of extracting topographical information from the same stack of images used for all-in-focus imaging.

1.1. State of the art

The approaches for skin disease diagnosis from dermoscopic images usually follow three steps: image segmentation, feature extraction, and classification [8]. Image segmentation labels every pixel to a category and is often used to determine the lesion's boundaries. Traditional segmentation methods include thresholding [9], the watershed-method [10] and graph cuts [11]. With the development of computer vision, many deep learning-based methods were proposed [12–14]. Feature extraction for melanoma is currently an active field of research as well. The extracted features can be based on e.g. the ABCDE rule [4], the ELM 7 point checklist [15] or the pigment network [16]. The skin lesion classification is nowadays dominated by Convolutional Neural Network (CNN) based methods [17–21]. All these approaches rely on a sharp contrast of the features in the dermoscopic data which underlines the need for all-in-focus imaging in non-contact dermoscopy.

Focus stacking is often employed in microscopic bioimaging. It can occur that objects in a scene inhibit different distances to the camera's focal point. This can lead to a blur of the objects that are outside the depth of field [22]. In dermoscopy, reliable diagnosis would be prevented by this limitation.

1.2. Overview

This work structures as follows: In section 2 we explain the working principle of the liquid lens and its implementation in the focus stacking imaging system. In section 3 we describe the proposed method step by step. In section 4 we present the details of the experiments, including the composition of the setup for the acquisition of focus stacks based on a liquid lens. In addition, we present the experimental results,

showing the effectiveness and limitations of the proposed method. Finally, in section 5, we conclude the main findings and discuss possible improvements.

2. Principle of image stacking with liquid lens

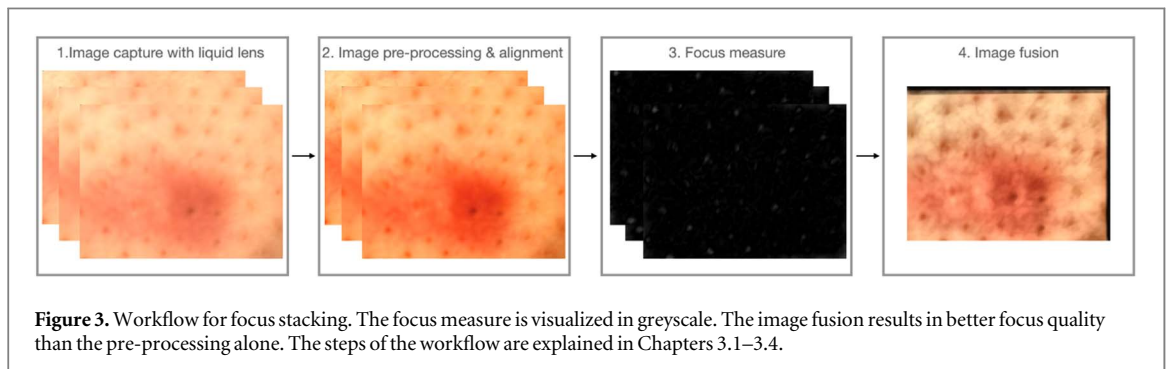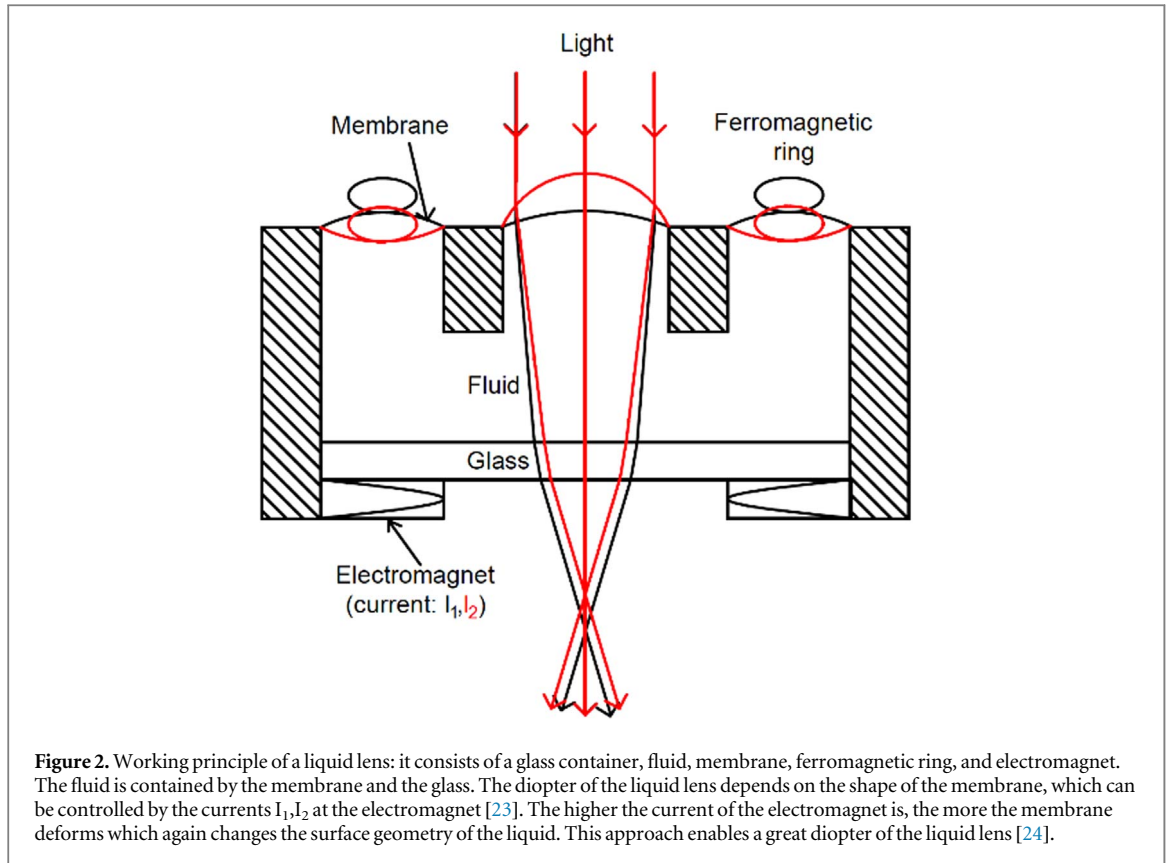### 2.1. Working principle of liquid lens

A liquid lens is an optical lens manipulating the light with different shapes of a liquid surface. It is possible to modulate its focal length by electrical controls. According to the working principle, liquid lenses can be divided into different groups. In this work, an electromagnetically actuated liquid lens was employed. Figure 2 displays the working principle.

2.2. Principle of focus stacking with liquid lens

The depth of field (DOF) of an image is limited, which means that for some skin topographies only parts of the lesion are in focus. In order to obtain an all-in-focus image of the lesion, the focused regions from every captured image are extracted and stacked together. This method is called focus stacking [25]. Figure 3 shows the workflow for focus stacking as implemented in this work. Through changing of the current of the electromagnet in the liquid lens, differently focused images are captured in one sequence. Because of the existence of misalignment within the images of the stack, the images need to be aligned first with respect to each other, so that they are overlaid pixel by pixel. Extracting the in-focus area from each image is called focus measure. The focus measure is calculated for every pixel in the stack to evaluate its focus. Based on the calculated focus results, the images are fused to create an all-in-focus image.

2.3. Image fusion metrics

Multi-focus image fusion metrics can be categorized into four groups [26]. In order to evaluate the fusion result objectively, we selected one metric from each group to assess the fusion results. The chosen metrics in this work are:

Figure 2. Working principle of a liquid lens: it consists of a glass container, fluid, membrane, ferromagnetic ring, and electromagnet. The fluid is contained by the membrane and the glass. The diopter of the liquid lens depends on the shape of the membrane, which can be controlled by the currents I_1, I_2 at the electromagnet [23]. The higher the current of the electromagnet is, the more the membrane deforms which again changes the surface geometry of the liquid. This approach enables a great diopter of the liquid lens [24].

Figure 3. Workflow for focus stacking. The focus measure is visualized in greyscale. The image fusion results in better focus quality than the pre-processing alone. The steps of the workflow are explained in Chapters 3.1–3.4.

(a) Normalized mutual information-based metric: Q_{MI} from the group of information theory-based metrics. It quantifies the distance between the fused image and input images [27].

(b) Spatial-frequency-based metric: Q_{SF} from the group of image feature-based metrics. It measures the first-order gradient error between the fused image and input images in four directions [28].

(c) Yang's metric: Q_{Yang} from the group of image structural similarity-based metrics. It evaluates the fusion result by a structural similarity index measure (SSIM) [29].

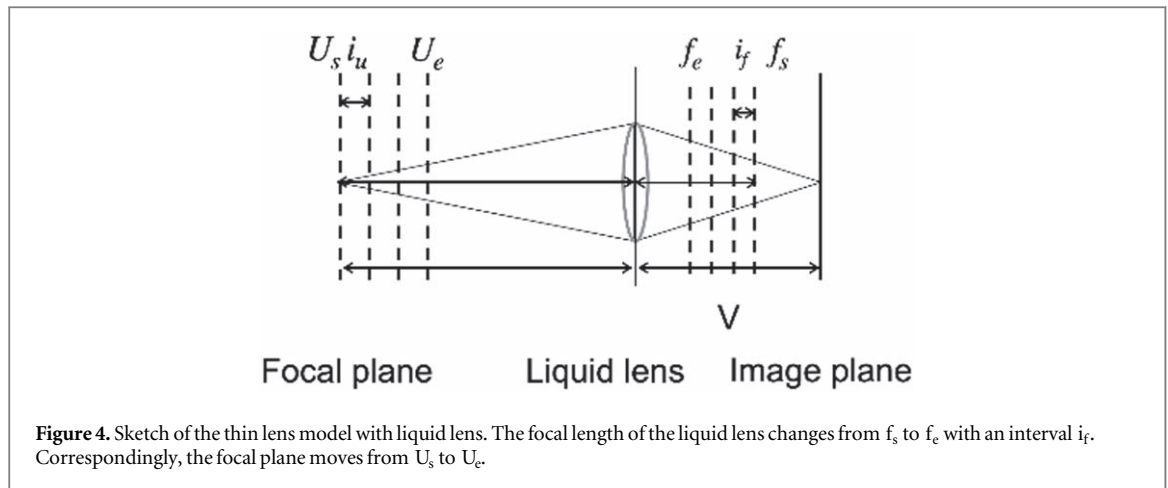(d) Chen-Varshney metric: Q_{Chen} from the group of Human perception inspired fusion metrics. This metric calculates a global quality measure based on the edge information, local region saliency, and similarity [30].

3. Proposed method

3.1. Image acquisition

In the following the basics for the image acquisition are being described. As shown in figure 2, the diopter of the liquid lens is controlled by the current of the electromagnet. A certain current corresponds to a certain focal plane. Figure 4 sketches the optics for collecting a stack of images from different focal planes by using a thin lens model.

From left to right, the three vertical planes are the focal plane, liquid lens, and the image plane. The focal length can be calculated by the thin lens equation.

Figure 4. Sketch of the thin lens model with liquid lens. The focal length of the liquid lens changes from f_s to f_e with an interval i_f. Correspondingly, the focal plane moves from U_s to U_e.

$$\frac{1}{U} + \frac{1}{V} = \frac{1}{f} \tag{1}$$

Equation (1) represents the thin lens equation, where U is the distance to the object, V is the image distance and f is the focal length. During the acquisition of the stack, the focal length of the liquid lens changes from f_s to f_e with a constant interval i_f. According to equation (1), the focal plane changes from U_s to U_e by an interval of i_u. As a result, images with different focal planes are captured sequentially.

3.2. Image alignment

An unavoidable cause for misalignment between the frames is that objects at different distances are magnified by different degrees. This is because of the change of the focal length between frames. Another reason for misalignment between frames is a slight movement of the patient (e.g., breathing) or the camera. An Enhanced Correlation Coefficient (ECC) [31] based method was employed on the captured images to eliminate the misalignment from these two sources.

Equation (2) uses the Euclidean norm to quantify the error between the reference image I_r and the warped input image $I_w(p)$. P are unknown parameters. The alignment problem is to estimate p.

$$E_{ECC}(p) = \left\| \frac{\bar{i}_r}{|\bar{i}_r|} - \frac{\bar{i}_w(p)}{\|\bar{i}_w(p)\|} \right\|^2 \tag{2}$$

In the experiment, the most magnified image in the image stack is used as the reference image. The other images of the stack are being correlated to the reference image sequentially. The ECC method aims to minimize $E_{ECC}(\mathbf{p})$. An estimated affine transformation matrix is calculated by minimizing the difference between the reference image and the warped input image.

3.3. Focus measure

The focused areas of each image are calculated by a Fast Fourier Transform (FFT) [32, 33] based method in the frequency domain. In this approach, we assume

that the focused area of an image has clear edges and texture. The aligned images are transferred from the spatial domain to the frequency domain by FFT. The focused edges and texture have a large gradient in the spatial domain, which means they exhibit high-frequency signals in the frequency domain. A Gaussian high-pass filter is employed on the aligned image in the frequency domain to filter out low-frequency signals. The amount of the residual signal depends on the size of the applied Gaussian kernel. By choosing a suitable kernel-size, the in-focus areas of each image can be optimized.

3.4. Image fusion

The aligned images are fused by a weight-based method [34]. A focus value f_{m_i} of a pixel in the input image i can be obtained by applying a focus measure to each input image. A high focus value indicates that the pixel is in-focus. On the contrary, a low focus value indicates that the pixel is not in-focus. As described in equation (3), for each pixel (u,v), the value of the fused image I_{fused} is the weighted sum of input images I_i. The number of images in the stack is k.
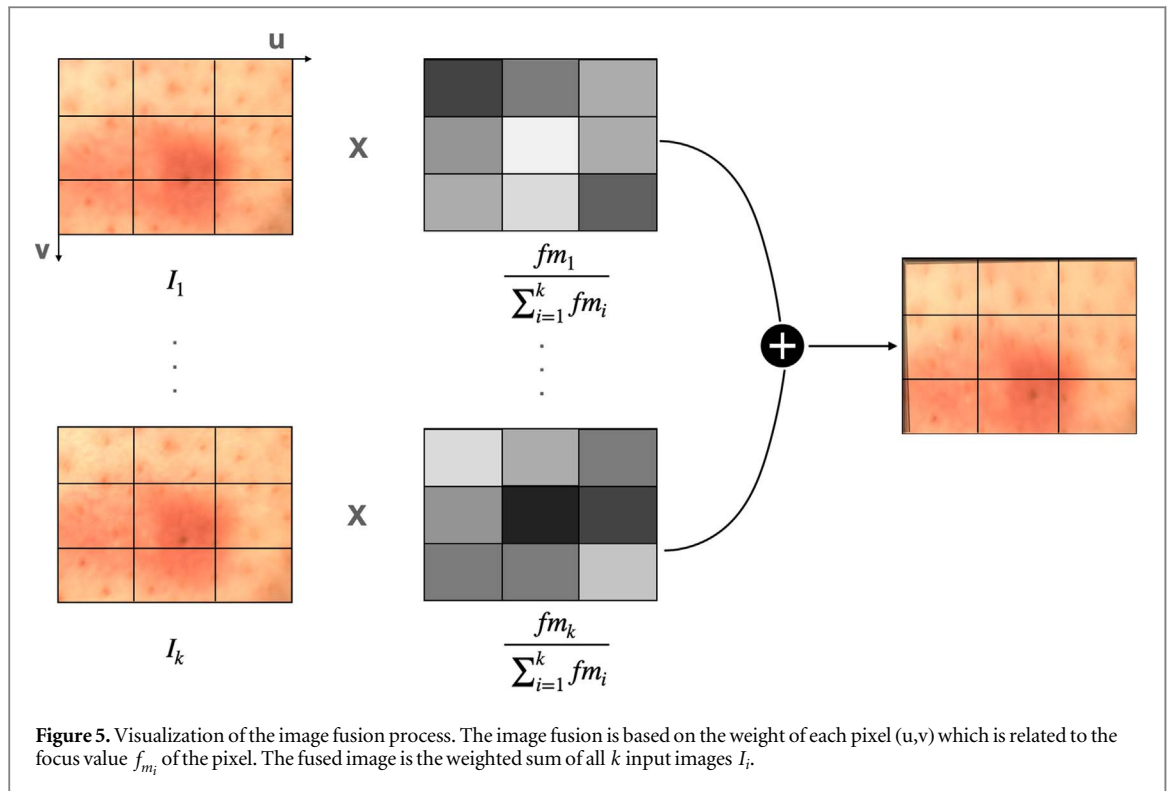
$$I_{fused}(u, v) = \sum_{i=1}^{k} \frac{f_{m_i}(u, v)}{\sum_{i=1}^{k} f_{m_i}} I_i(u, v) \tag{3}$$

Figure 5 visualizes the fusion process. The pixel on the fused image is a weighted sum of the pixels in the input image stack.

3.5. Topography measurement

In this work, we assume that the in-focus areas in one image are in the same plane. The focus measure helps to determine the in-focus areas of each image, and the corresponding image distances of those areas are calculated.

Equation (1) describes the relationship between the object distance U and the focal length f. Furthermore, the focal length f can be tuned by the current value c of the liquid lens. In the experiment, the correlation between the object distance and the current value of the liquid lens is obtained in a calibration process. We

Figure 5. Visualization of the image fusion process. The image fusion is based on the weight of each pixel (u,v) which is related to the focus value f_{m_i} of the pixel. The fused image is the weighted sum of all k input images I_i.

obtain the object distance U as a function of the current value c. Afterwards, the in-focus current c can be calculated by focus measure in equation (4).
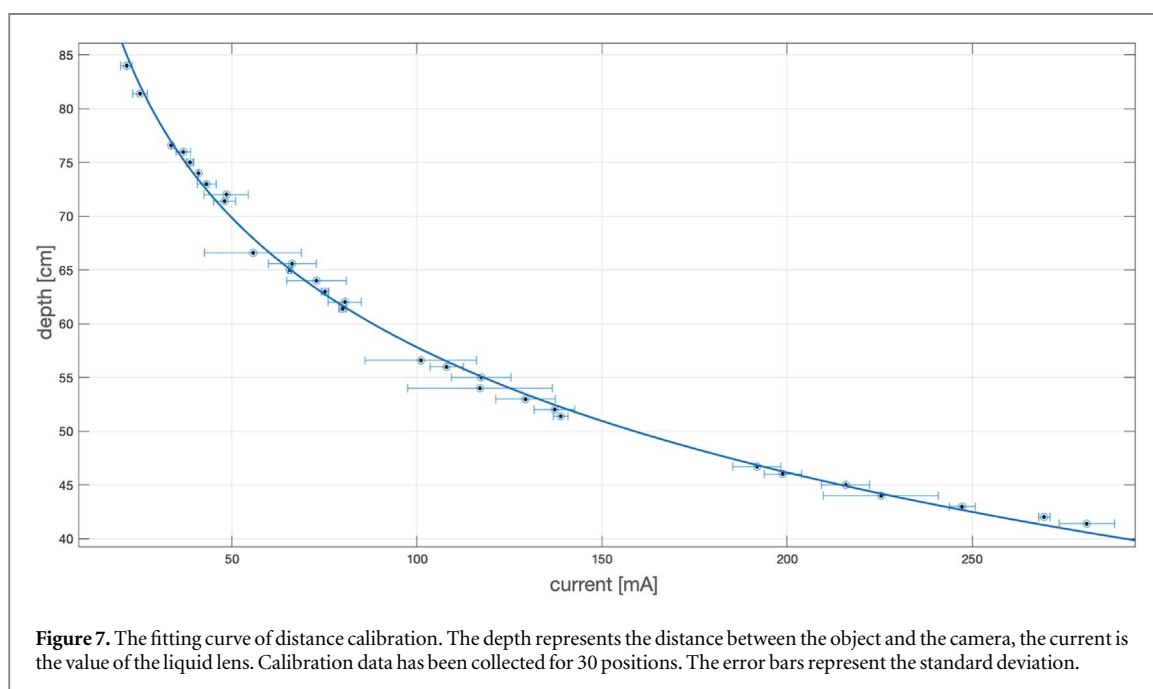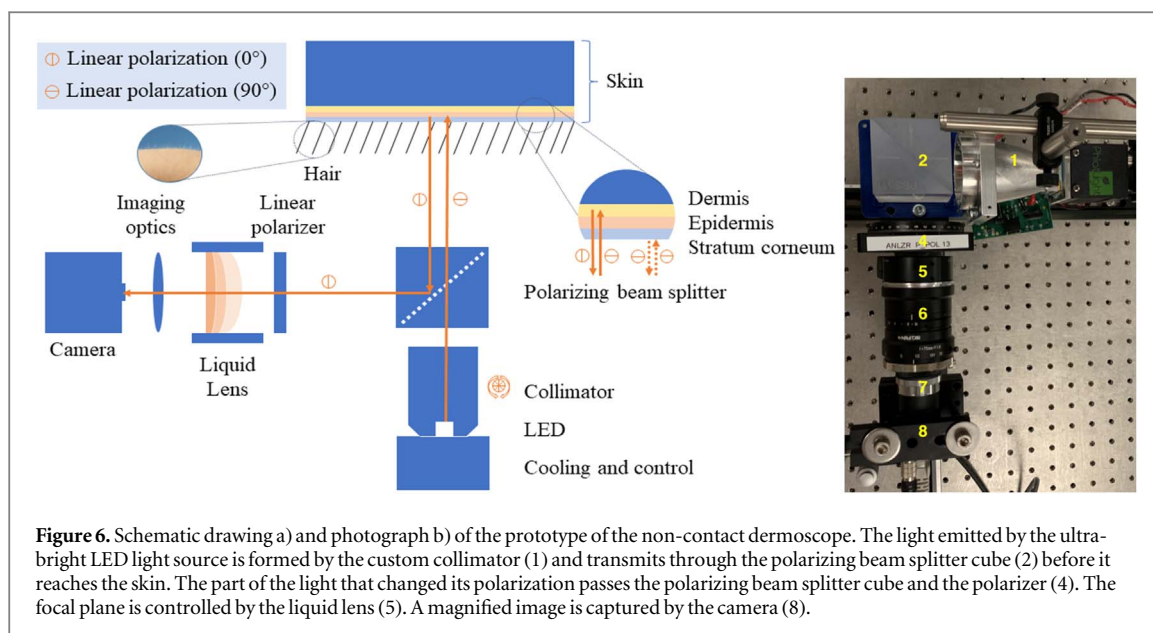
The pixel with maximal focus value is used for the topography estimation of the lesion. For example, given k input images, through focus measure, the calculated focus values of pixel (u,v) on each input image are $f_{m_1} \ldots f_{m_k}$. Among these, $f_{m_{max}}$ is the maximal focus value. The corresponding input image was captured under the current value c. Based on the fitting curve derived from the distance calibration, the depth of this pixel can be determined. Iterating through the above process for all pixels of the whole image can determine the topography map for all pixels.

For absolute topography measurements we employ a distance calibration method for focus stacking. In this approach, a flat object is placed at several working distances and a set of images under different currents is captured for each distance. The best focused image among all captured images is selected by a focus measure of the whole image. The operating current for the image with the largest focus measure is recorded as the in-focus current for the corresponding distance. By adjusting the distance to the camera, different current values are recorded. In the distance calibration, 30 working distance and current value pairs are measured. Figure 7 shows the fitting curve to the distance calibration results. With the derived regression function, equation (4), we can quantify the distance to the object U by the current of the liquid lens c as follows:

$$U = f(c) = 453.7 \frac{\text{cm}}{\text{mA}} \times c^{-0.04671} - 308 \text{ cm} \quad (4)$$

3.6. Instrumentation and image stack acquisition

Figure 6 shows the setup of the liquid lens-based non-contact dermoscope used in this work. The setup consists of an ultra-bright white light-emitting diode (CBT-90 White LED, Luminus Inc., Sunnyvale, California, USA) (LED), a custom-designed collimator (1) and a polarizing beam splitter cube (PBS511, Thorlabs, Newton, USA) (2) (PBS) for the illumination of the skin. The imaging part contains a polarizer (4), a liquid lens (EL-16-40-TC, Optotune AG, Dietikon, Schweiz) (5), a fixed-focus lens (6), a magnifier (NMV-75M1, Navitar, Rochester, New York, USA) (7), and a charge-coupled device (CCD) camera (BFS-U3-32S4M-C, FLIR Integrated Imaging Solutions Inc., Richmond, British Columbia, Canada) (8). The light source has a color rendering index (CRI) of 76 which makes it comparable to daylight [35]. Together with the custom-made collimator it provides an evenly distributed illumination to the skin. The emitted light partly transmits through the PBS, is being polarized linearly and illuminates the skin. Only the light that is scattered within the skin changes its polarization while the light reflected at the surface maintains its polarization [36]. Therefore, the polarizer filters out the surface reflections in a cross polarization configuration of the PBS and the polarizer. In addition, the liquid lens is employed to adjust the focal plane of the imaging system electrically, rapidly and without moving parts. It has an aperture of 16 mm, and the diopter of the liquid lens can be adjusted from -10 to 10 by changing the shape of the fluid. An overall reproducibility of $-/+0.05$ diopters is achievable. The response and

Figure 6. Schematic drawing a) and photograph b) of the prototype of the non-contact dermoscope. The light emitted by the ultra-bright LED light source is formed by the custom collimator (1) and transmits through the polarizing beam splitter cube (2) before it reaches the skin. The part of the light that changed its polarization passes the polarizing beam splitter cube and the polarizer (4). The focal plane is controlled by the liquid lens (5). A magnified image is captured by the camera (8).

Figure 7. The fitting curve of distance calibration. The depth represents the distance between the object and the camera, the current is the value of the liquid lens. Calibration data has been collected for 30 positions. The error bars represent the standard deviation.

settling times are 5 and 25 ms respectively [24]. The aperture of the fixed lens was adjusted between f/1.8 to f/16 while the shutter of the camera was adjusted correspondingly from 3 ms to 80 ms to maintain an adequate exposure. Image stacks were captured with automatically changing focal lengths, which are controlled by changing the diopter of the liquid lens. The diopter adjustment range of the liquid lens is −10 to 10. The corresponding control current is from −290 mA to 290 mA. During the acquisition, images were captured with an interval of 1 mA. The initial current value is the value that has the farthest point of the object in focus, and the end current value is the value that makes the nearest point of the object in focus. The aim is that every point of the object is focused on during the capture at least once. Together

with the liquid lens, a lens with fixed focal length is used for the imaging. That latter lens has an equivalent focal length of 75 mm. The aperture is adjustable from f/1.8 to f/16. Finally, the magnified image is captured by the camera, which is equipped with a 1/1.8′ CCD sensor and has a resolution of 1928 × 1448 pixels. From center to edge, the resolution changes from 120 l p mm^{-1} to 80 lp/mm. The captured images are transferred via a 1000 Mbit s^{-1} GigE interface from the camera buffer to a computer. The computer controls the camera and the liquid lens, to adjust the shutter of the camera and the diopter of the liquid lens.

The image stack acquisition was implemented in Python (version: 3.6.4) with the libraries PyCapture2 for the camera and Opto for the liquid lens. For rapid and convenient capture, a graphical user interface was
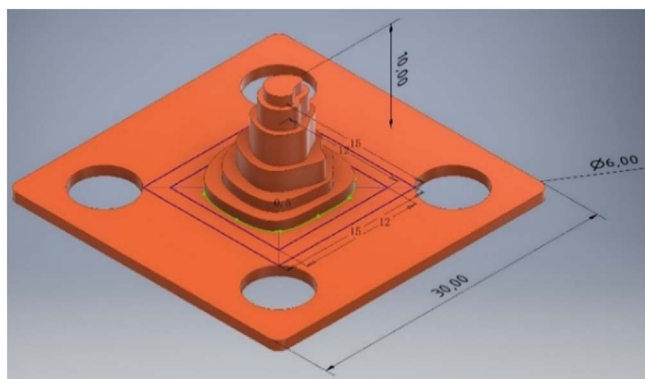
Figure 8. CAD sketch with of the designed melanoma phantom with dimensions given in mm.

Figure 9. Captured input images for different currents of the liquid lens (liquid lens current from left to right image: 103 mA, 109 mA, 115 mA, 121 mA, 127 mA).
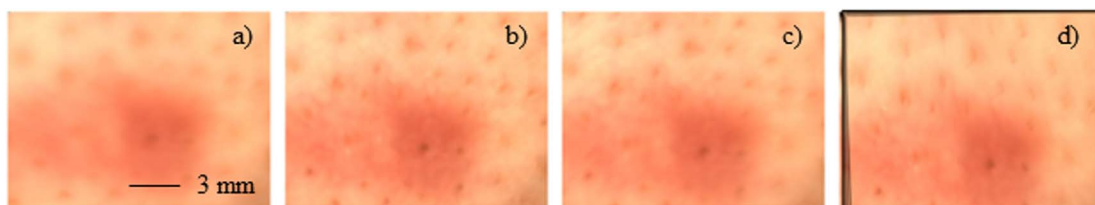

Figure 10. Captured images (a)–(c) and the fused image (d) of a lesion under cross polarization. The image stack contained 25 images.

designed based on Qt (5.15.4) while the enhancement of the depth of field and the topography measurements are implemented in MATLAB (MATLAB, 2021. version 9.11.0 (R2021b), Natick, Massachusetts: The MathWorks Inc.).

Figure 8 Shows a 3D rendering of a phantom of melanoma designed to be the sample for the focus stacking approach.

The phantom design is based on the ABCD rule for melanoma diagnosis. Therefore, it is asymmetric, has irregular borders due to the printing process, and a diameter and height of 10 mm. The phantom consists of steps with different heights. The height is relatively high to show the effect of the approach on extreme skin topographies.

4. Experimental results and discussion

4.1. Enhanced depth of field

Figure 9 shows the captured image stack of a skin lesion with changing focal planes. From the stack of the lesion, it is visible that the focal plane changes from one image to the other. The focus is moving through the image showing that possible body movement in the optical axis does not affect the approach in this case. Figure 10 shows the fusion results for the same lesion (right panel). The first three images are images from the stack while the very right image shows the result after fusion. The used image stack contains more than the three displayed images. Usually, the image stacks contain 15 to 25 images. Compared to each input image, the DOF of the fused image is extended. In the enhanced depth of focus image, the details in the background and foreground are shown in one image. Figure 11 shows the results for the same lesion illuminated without cross polarized light. Without cross polarized illumination, more surface features are visible in the human skin. Therefore, the approach with non-polarized light shows a slightly improved performance for the all-in-focus image. In figure 12 we evaluate the effectiveness of focus stacking for an *ex vivo* melanoma. The fused image shows all the regions of the melanoma in focus while the example
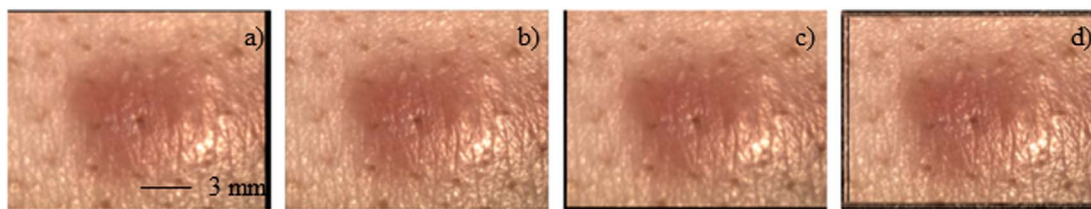
Figure 11. Captured images (a)– (c) and the fused image d) of a lesion without cross polarization. The image stack contained 25 images.
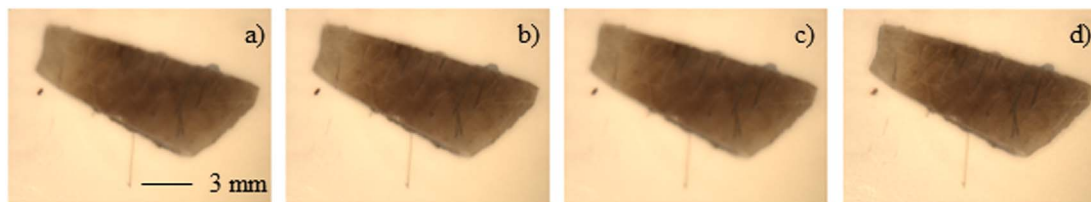

Figure 12. Captured images (a), (b) and (c) and the fused image (d) of an ex vivo melanoma under cross polarization. The image stack contained 25 images.
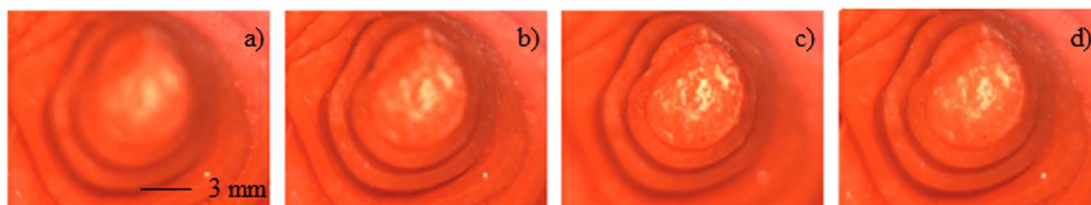

Figure 13. Captured images (a)–(c) and the fused image (d) of a custom-designed and 3D printed skin phantom under cross polarization. The image stack contained 25 images.

images of the stack hold blurred areas. A similar result is presented for the custom-designed phantom in figure 13. Figure 13 shows three images from the image stack of the phantom. The increased current moves the focal plane from the background to the foreground. Each image contains a part of the phantom, which is in-focus. In contrast to the images of the stack, the fused image on the very right displays all the regions of the phantom in-focus. Focus stacking has a better performance on the phantom than on the skin. In comparison, the image of the lesion observed with unpolarized light shows a better result compared with the observation with cross polarized light. This might be because the unpolarized illumination leads to more reflections from the skin surface which are easily detected as in-focus features. The results for the enhanced depth of field depends on several acquisition settings, i.e., aperture, illumination, polarization, samples, and parameters in the post-processing, like the size of the Gaussian kernel for focus measure. For the acquisition of the stacks we used the parameters optimized for skin imaging with our setup as the dermoscopic images are the most valuable data for the

dermatologist. Especially the aperture and illumination have a strong effect on the depth of field. The fusion result is mainly influenced by the quality of the image alignment and the application of the most-suitable focus measure. During the stack acquisition, in contrast to the phantom, it is difficult for the patient to remain static. There are two kinds of possible movements: movements within the focal plane and movements in the direction of the optical axis. Movements within the focal plane impose difficulties for the image alignment. The black bars in the images of figure 11 show the degree of misalignment that has been corrected. If the lesion is textured, like in the case of the mole observed under unpolarized illumination, automated image alignment can easily solve this type of motion and alignment problem. However, for dermatologic diagnosis, images of the lesion under cross polarization are typical. Under cross polarization the skin shows the subcutaneous information of the lesion, which has less features. Unfortunately, this limits the alignment performance. Due to the small field of view (FOV), the lesion may even move out of view during the acquisition as displayed in figure 14.
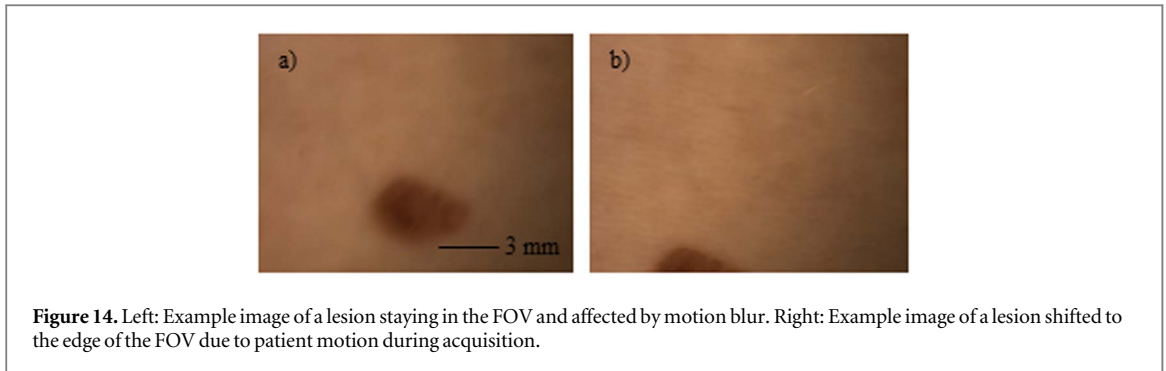
Figure 14. Left: Example image of a lesion staying in the FOV and affected by motion blur. Right: Example image of a lesion shifted to the edge of the FOV due to patient motion during acquisition.
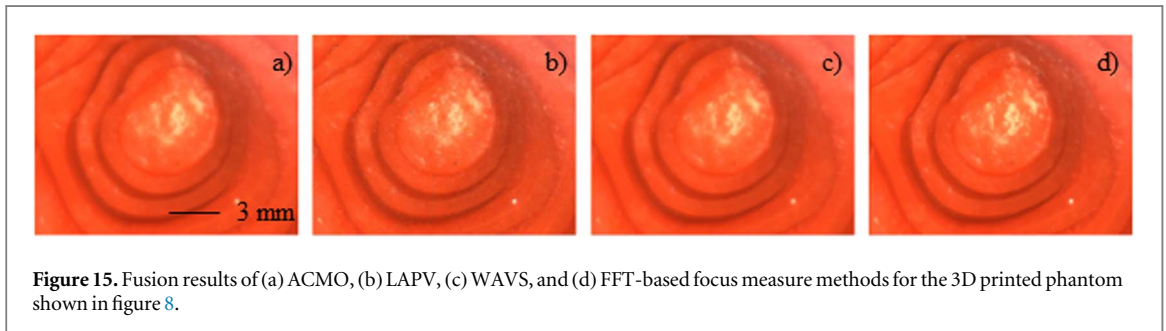

Figure 15. Fusion results of (a) ACMO, (b) LAPV, (c) WAVS, and (d) FFT-based focus measure methods for the 3D printed phantom shown in figure 8.

Table 1. Objective evaluation of the fusion results.

Metrics Methods	Q_{MI}	Q_{SF}	Q_{Yang}	Q_{Chen} Q_{Chen}
Absolute central moment (ACMO)	0.9353	0.0164	0.3085	21.5563
Variance of Laplacian (LAPV)	0.1727	0.0386	0.0192	28.2701
Sum of Wavelet coefficients (WAVS)	0.9265	0.0172	0.3045	13.0465
FFT-based	0.8687	0.0222	0.3287	8.1528

Movements in the depth direction hamper the effectiveness of the focus measure. In some settings, the depth of field in an image can be as small as 1–2 mm. Here, a slight movement may cause the lesion to lose focus. Moreover, it can occur that not each region of the object can be in-focus at least once. Except for our FFT-based fusion method, many spatial-based or wavelet-based focus measure methods have been proposed for focus stacking [37]. Figure 15 shows the fusion result of Absolute central moment (ACMO) [38], Variance of Laplacian (LAPV) [39], Sum of wavelet coefficients (WAVS) [40], and our FFT-based method.

From subjective criteria the FFT-based method shows the best performance for the all-in-focus imaging of the phantom. The ACMO and WAVS methods both generate blurred all-in-focus images, and the image fused by LAPV suffers the most noise. Compared to that, table 1 shows the results when applying the objective fusion assessment criteria explained in section 2.

The best method for each metric is given in bold values. Q_{MI}, Q_{SF} and Q_{Yang} are positive metrics, which means the larger value represents better fusion quality.

Q_{Chen} is a negative metric with a smaller value indicating better fusion quality. The FFT-based method gets the best score in Q_{Yang} and Q_{Chen}. This validates the effectiveness of our FFT-based multi-focus fusion method.

4.2. Topography measurement

Figure 16 shows the experimental results for the topography maps of the phantom. The topography map of the phantom not only has clear edges but also shows the protrusions well. Still, the topographical information is affected by noise. In figure 17 we overlaid the depth information of the phantom derived from the focus stack with the CAD model of the phantom. Figure 17 shows that the dimensions of the phantom in X- and Y-direction exceed the FOV of the imaging system. Furthermore, the depth information derived from the focus stack shows good agreement with the designed depth information of the phantom. In addition, the measured height of the phantom is matches the height of the phantom at each of the steps. Compared to that, the topography map of the lesion displayed in figure 18 is flatter as the protrusions of the lesion are shallow.
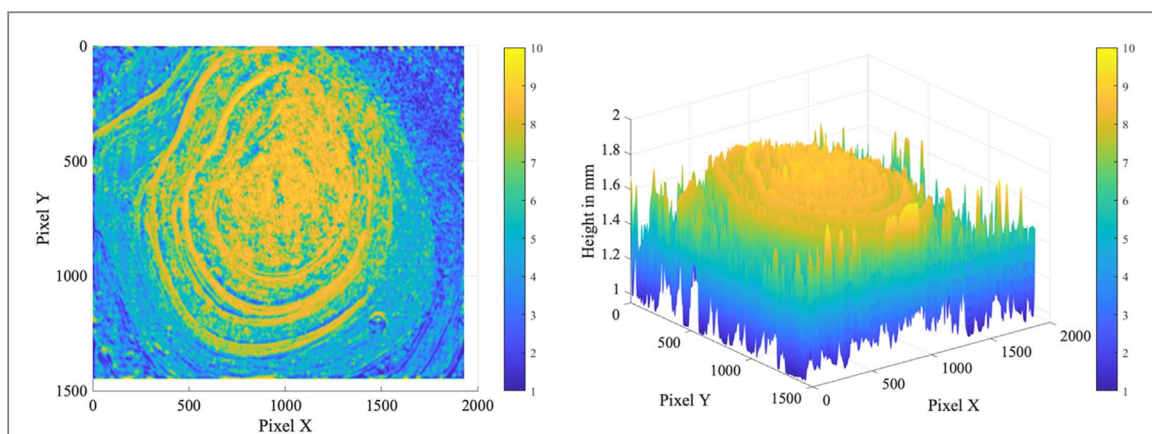
Figure 16. Topography maps of the phantom shown in figure 13, left: top view, right: side view. The topographical information is affected by noise. The color bars represent the height.
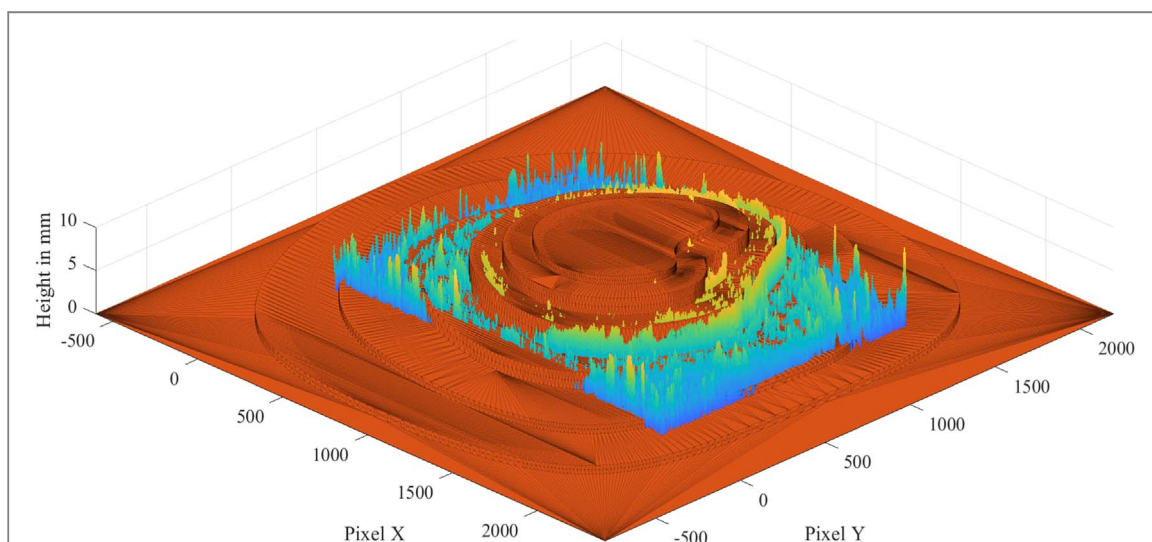

Figure 17. Comparison of generated topographical information by focus stacking with CAD model.
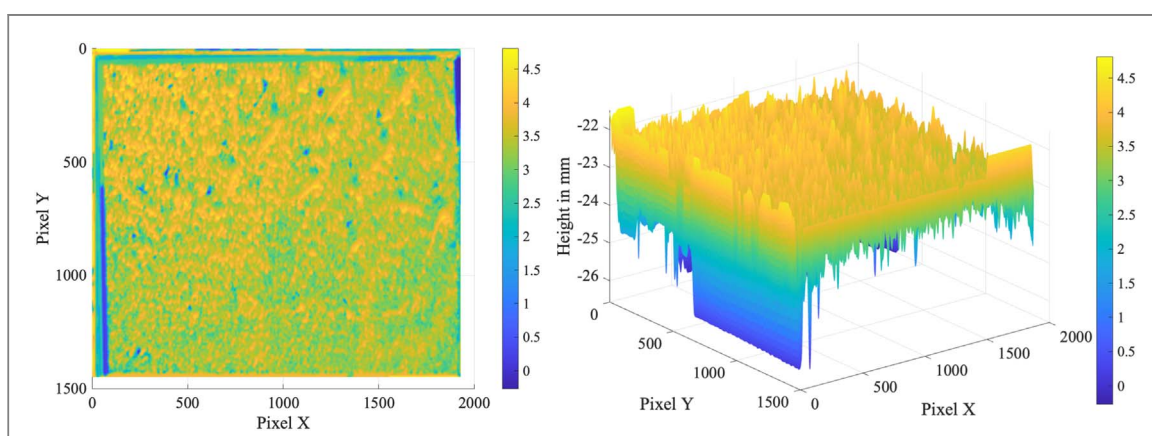

Figure 18. Topography maps of the lesion shown in figure 10, left: top view, right: side view. The topographical information is affected by noise. The color bars represent the height.

It can be observed that the mole does not have a strong elevation. It has to be noted, that movement during the image stack acquisition did not move any of the captured images out of focus. In addition, the in-focus areas might not be detected in the out-of-focus images because of the motion blur. Meanwhile,
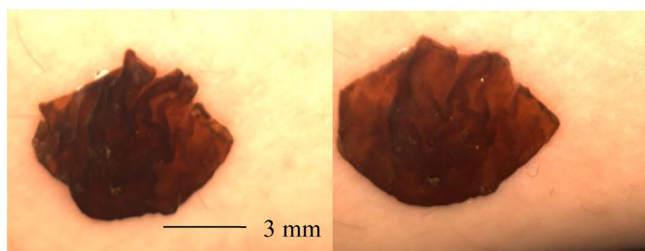
Figure 19. Melanoma phantom based on a raisin attached to human skin. The left phantom has a different geometry compared to the right phantom. The approximated average heights of the phantoms are 1,5 mm and 0,5 for the left and right phantom, respectively.
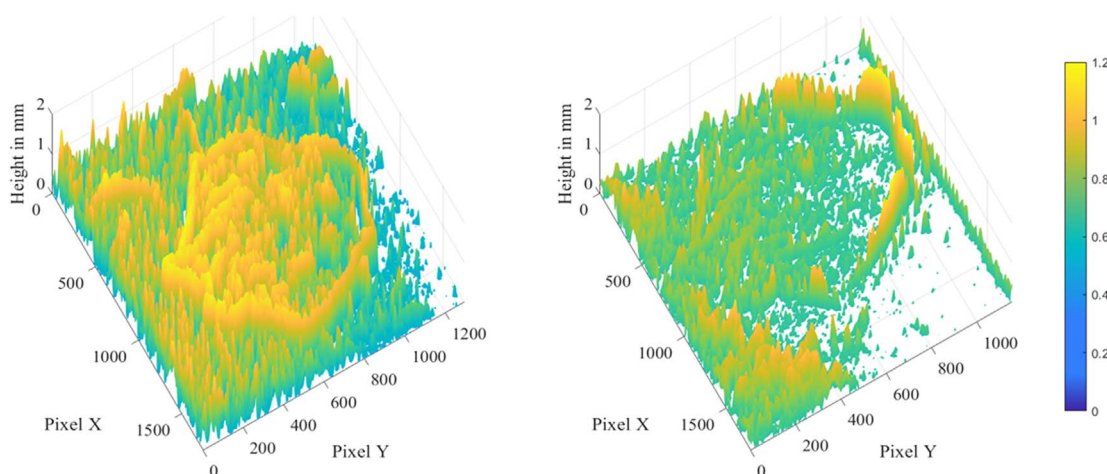

Figure 20. Comparison of two topography maps generated from two raisin phantoms with different geometries. The topography information is affected by noise.
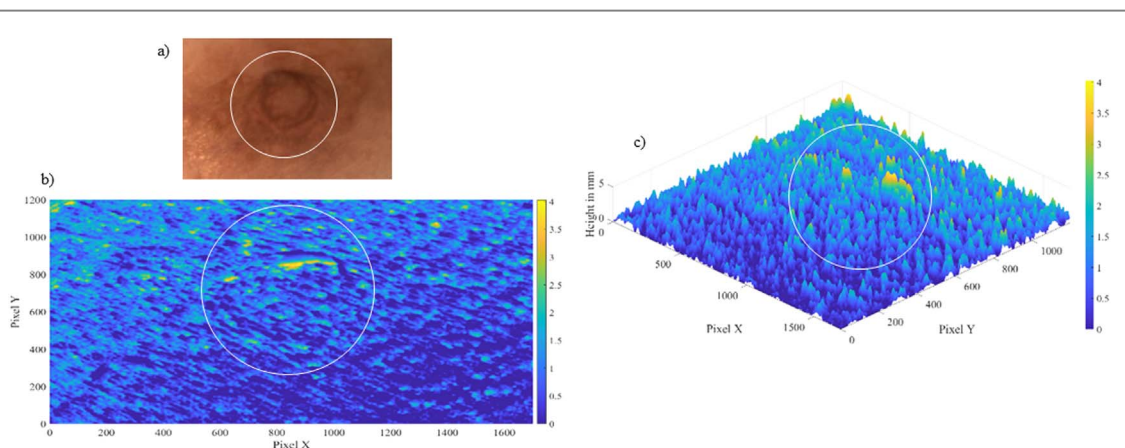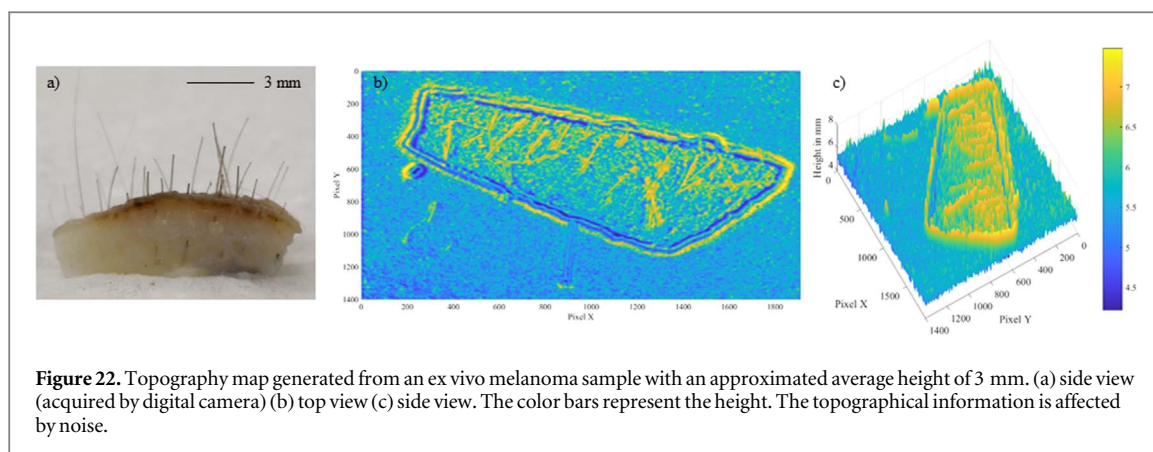

Figure 21. Topography map generated from a human acromastium. (A) RGB image of the acromastium under study. (B) Top-view with the region of the acromastium marked. (C) side-view. The approach thus works for *in vivo* measurements on human skin. The color bars represent the height. The topographical information is affected by noise.

because of the lack of ground truth of the topography map for *in vivo* measurements, an objective evaluation is not available.

In the following we discuss the results for another type of phantom for which two differently shaped raisins which have been attached to human skin. The raisin resembles the optical appearance of a lesion as displayed in figure 19.

Figure 20 shows the depth maps derived from the focus stacks of the two phantoms displayed in figure 19. The two different geometries of the raisins are clearly distinguishable in the depth maps. This

Figure 22. Topography map generated from an ex vivo melanoma sample with an approximated average height of 3 mm. (a) side view (acquired by digital camera) (b) top view (c) side view. The color bars represent the height. The topographical information is affected by noise.

shows that it is possible to detect changes in the geometry of a lesion via focus stacking. Furthermore, we show that it is possible to obtain the topology of the human skin *in vivo* by the example of a human acromastium with the depth map displayed in figure 21. In this example, the topography of the acromastium and its contours are roughly visible. Compared to the raisin phantoms, this *in vivo* measurement performs worse. It indicates that the color contrast might aid the topography measurements derived by focus stacking. The depth maps for the ex vivo melanoma are displayed in figure 22. The depth maps in figure 22 of the melanoma prove that the geometries of the melanoma are measurable by focus stacking.

5. Conclusion & outlook

In this work, we have proposed a focus stacking method for non-contact dermoscopy based on a liquid lens. An all-in-focus image and the corresponding depth map can be obtained at the same time. In addition, the FFT-based method is evaluated via both subjective and objective metrics. The estimated depth for the *in vivo* measurements map could only be assessed by subjective criteria because of the lack of ground truth. The body movements during the capture limits the performance of focus stacking in non-contact dermoscopy. For patients that cannot remain static during a non-contact dermoscopic imaging procedure (e.g. due to Parkinson's disease), focus stacking could help to obtain at least one image from the stack that shows sufficient focus. For total body dermoscopy scanners that employ focus stacking to ensure that a focused image is obtained in the case of patient movement, the focus stacks will be a valuable by-product with additional information. The hyperfocus images are the main result from the optical system of the non-contact dermoscope. The resulting improvement in image quality can potentially aid the diagnosis of the dermatologist. Furthermore, the data quality that intelligent machines for computer aided diagnostics are build on can be enhanced. This can lead to potentially higher classification accuracy.

In further works, improved and more reliable registration algorithms need to be developed to aid the difficulties in the image alignment. In addition, the depth resolution of the proposed depth resolution can be determined and the clinical application can be evaluated. Furthermore, the potential of machine learning on the depth measurement by focus stacking and its noise reduction needs to be explored. Considering the limitations of monocular depth estimation, a stereo system may provide depth information for dermatologic diagnosis as well.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Materials

The performed experiments were approved by the Ethics Committee of the University Medical Center Rostock (A 2016-0115). The results in figures 9, 10, 11, 14, 18–21 were measured on the skin of the authors.

Disclosures

The authors declare no conflict of interests.

ORCID iDs

Lennart Jütte https://orcid.org/0000-0002-5564-7472
Bernhard Roth https://orcid.org/0000-0001-9389-7125

References

[1] Sung H *et al* 2021 Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries *CA: A Cancer Journal for Clinicians* 71 (Atlanta: American Cancer Society) (https://doi.org/10.3322/CAAC.21660)

[2] Siegel R L *et al* 2021 Cancer Statistics, 2021 *CA: A Cancer Journal for Clinicians* 71 (Atlanta: American Cancer Society) 7–33

[3] Jacques S L, Ramella-Roman J C and Lee K 2002 Imaging skin pathology with polarized light *Journal of Biomedical Optics* **7** 329–40

[4] Rigel D S *et al* 2005 ABCDE—an evolving concept in the early detection of melanoma *Archives of Dermatology* **141** 1032–4

[5] Fricke D *et al* 2019 Non-Contact Dermatoscope with ultra-bright light source and liquid lens-based autofocus function *Applied Sciences* **9** 2177

[6] Meinhardt-Wollweber M *et al* 2017 A non-contact remote digital dermoscope to support cancer screening and diagnosis of inflammatory skin disease *Biomedical Physics & Engineering Express* 3 (Bristol: IOP Publishing) (https://doi.org/10.1088/2057-1976/aa84d3)

[7] Heratizadeh A *et al* 2020 Non-contact remote digital dermoscopy—new perspectives on differential diagnosis of inflammatory skin diseases *Journal of the European Academy of Dermatology and Venereology* **34** 125–6

[8] Mishra N K and Celebi M E 2016 An overview of melanoma detection in dermoscopy images using image processing and machine learning arXiv: 1601.07843

[9] Otsu 1979 A threshold selection method from gray-level histograms *IEEE Transactions on Systems, Man, and Cybernetics* **9** 1

[10] Najman L and Schmitt M 1994 Watershed of a continuous function *Signal Processing* **38** 99–112

[11] Boykov Y, Veksler O and Zabih R 2001 Fast approximate energy minimization via graph cuts *IEEE Transacions on Pattern Analysis and Machine Intelligence* **23** 1222–39

[12] Dhanachandra N, Manglem K and Chanu Y J 2015 Image segmentation using K-means clustering algorithm and subtractive clustering algorithm *Procedia Computer Science* **54** 764–71

[13] Zhou H *et al* 2008 Spatially constrained segmentation of dermoscopy images *2008 5th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro* (https://doi.org/10.1109/ISBI.2008.4541117)

[14] Guo X, Chen Z and Yuan Y 2020 Complementary Network with Adaptive Receptive Fields for Melanoma Segmentation *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (https://doi.org/10.1109/ISBI45749.2020.9098417)

[15] di Leo G *et al* 2004 ELM image processing for melanocytic skin lesion diagnosis based on 7-point checklist: a preliminary discussion *The Open Dermatology Journal* **4** 110–5

[16] Sadeghi M *et al* 2010 A novel method for detection of pigment network in dermoscopic images using graphs *Computerized Medical Imaging and Graphics* **35** 137–43

[17] Schaefer G *et al* 2014 An ensemble classification approach for melanoma diagnosis *Memetic Computing* **6** 233–40

[18] Barata C, Celebi M E and Marques J S 2015 Improving dermoscopy image classification using color constancy *IEEE Journal of Biomedical and Health Informatics* **19** 1146–52

[19] Celebi M E *et al* 2007 A methodological approach to the classification of dermoscopy images *Computerized Medical Imaging and Graphics* **31** 362–73

[20] Namozov A and Cho Y I 2018 Convolutional neural network algorithm with parameterized activation function for melanoma classification *International Conference on Information and Communication Technology Convergence (ICTC)* (https://doi.org/10.1109/ICTC.2018.8539451)

[21] Esteva A *et al* 2017 Dermatologist-level classification of skin cancer with deep neural networks *Nature* **542** 115–8

[22] Sigdel M S *et al* 2016 Focus ALL: focal stacking of microscopic images using modified harris corner response measure *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **13** 326–40

[23] Chen L *et al* 2021 Electrically tunable lenses: A review *Frontiers in Robotics and AI* **8** 678046

[24] 2022 Optotune Electrically tunable large aperture lens EL-16-40-TC (20D) (https://www.optotune.com/el-16-40-tc-lens)

[25] Zhang X 2020 Multi-focus image fusion: a benchmark arXiv: 2005.01116

[26] Liu Z *et al* 2012 Objective assessment of multiresolution image fusion algorithms for context enhancement in Night vision: a comparative study *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34** 94–109

[27] Guihong Q, Dali Z and Pingfan Y 2022 Information measure for performance of image fusion *Electronics Letters* **38** 313–5

[28] Zheng Y *et al* 2007 A new metric based on extended spatial frequency and its application to DWT based fusion algorithms *Information Fusion* **8** 177–92

[29] Yang C *et al* 2008 A novel similarity based quality metric for image fusion *Information Fusion* **9** 156–60

[30] Chen H and Varshney P K 2007 A human perception inspired quality metric for image fusion based on regional information *Information Fusion* **8** 193–207

[31] Evangelidis G D and Psarakis E Z 2008 Parametric image alignment using enhanced correlation coefficient maximization *IEEE Transactions on Pattern Analysis and Machine Intelligence* **30** 1858–65

[32] Frigo M and Johnson S G 1998 FFTW: an adaptive software architecture for the FFT *Int. Conf. on Acoustics, Speech and Signal Processing* (https://doi.org/10.1109/ICASSP.1998.681704)

[33] De K and Masilamani V 2013 Image sharpness measure for blurred images in frequency domain,' procedia engineering *Procedia Engineering* **64** 149–58

[34] Le S *et al* 2009 A novel automatic weighted image fusion algorithm *International Workshop on Intelligent Systems and Applications* (https://doi.org/10.1109/IWISA.2009.5072656)

[35] Basu C *et al* 2015 Light source design for spectral tuning in biomedical imaging *Journal of Medical Imaging* **2** 044501

[36] Günther A *et al* 2013 An ultra-bright white LED based non-contact skin cancer imaging system with polarization control *Clinical and Biomedical Spectroscopy and Imaging III*

[37] Kaur H, Koundal D and Kadyan V 2021 Image fusion techniques: a survey *Archives of Computational Methods in Engineering* **28** 4425–47

[38] Shirvaikar M v 2004 An optimal measure for camera focus and exposure *Proc. of the Annual Southeastern Symp. on System Theory* 36 (https://doi.org/10.1109/SSST.2004.1295702)

[39] Pech-Pacheco J L *et al* 2000 Diatom autofocusing in brightfield microscopy: a comparative study *International Conference on Pattern Recognition* (https://doi.org/10.1109/ICPR.2000.903548)

[40] Yang G and Nelson B J 2003 Wavelet-based autofocusing and unsupervised segmentation of microscopic images *IEEE International Conference on Intelligent Robots and Systems* (https://doi.org/10.1109/IROS.2003.1249188)