Journal of Electronic Imaging

JElectronicImaging.org

Internal fingerprint zone detection in optical coherence tomography fingertip scans

Luke Nicholas Darlow
James Connan
Sharat Saurabh Akhoury

Internal fingerprint zone detection in optical coherence tomography fingertip scans

Luke Nicholas Darlow,[a,b,*] James Connan,[a] and Sharat Saurabh Akhoury[b]
[a]Rhodes University, Department of Computer Science, Grahamstown, Hamilton Building, Prince Alfred Street, Grahamstown 6139, South Africa
[b]Council for Scientific and Industrial Research, Modeling and Digital Science, Building 17C, 627 Meiring Naude Road, Pretoria, South Africa

Abstract. Optical coherence tomography (OCT) is a high-resolution imaging technology capable of capturing a three-dimensional (3-D) representation of fingertip skin. The papillary junction—a junction layer of skin containing the same topographical features as the surface fingerprint—is contained within this representation. The top edge of the papillary junction contains the topographical information pertinent to the internal fingerprint. Extracting the internal fingerprint from OCT fingertip scans has been shown to be possible. Currently, acquiring the internal fingerprint involves manually defining the region containing it. This manner of definition is inefficient. Perfect knowledge of the location of the papillary junction is hypothesized as achievable. This research details and tests a *k*-means clustering approach for papillary junction detection. All tested metrics are of a standard comparable to the measured human error. The technique presented in this research is highly successful in detection of the location of the papillary junction. Furthermore, high-quality internal fingerprints are acquired using the coordinates obtained. *© 2015 SPIE and IS&T* [DOI: 10.1117/1.JEI.24.2.023027]

Keywords: optical coherence tomography; internal fingerprint; biometrics; vomputer vision; clustering; digital signal processing.
Paper 14793 received Dec. 17, 2014; accepted for publication Mar. 16, 2015; published online Apr. 9, 2015.

1 Introduction

Digital privacy and security are constant concerns that need to be addressed by inventive and improved biometric solutions for access control. Improving on the established and widely accepted biometric of fingerprints is a means to this end. Fingerprints have the advantages of universality, high distinctiveness, and good false rejection and acceptance rates. However, conventional acquisition of fingerprints results in a two-dimensional (2-D) image of the surface fingerprint.[1] The security and reliability of this type of representation are known to be lacking.[1,2] Essential progress thereupon should result in the mitigation of any associated limitations.[1] These limitations are as follows:

1. **Distortion**: owing to the elasticity of the skin and varying pressure during the acquisition process, surface fingerprints are subject to distortion.[1]
2. **Damage**: any surface skin damage affects fingerprint quality and thus reduces matching capability. The fingerprint is exposed to varying degrees of wear and tear. For instance, individuals employed in heavy duty industries are often without usable fingerprints.
3. **Security**: owing to the 2-D representation of a surface fingerprint scan, fingerprint spoofing is an uncomplicated task and requires a few inexpensive household items only.[3,4]

The acquisition and use of an internal fingerprint can mitigate these disadvantages. Between the epidermis and dermis is an intermediary layer of skin known as the papillary junction. This junction grows with ridges and valleys to increase the surface area contact between the epidermis and dermis. These ridges and valleys propagate to the surface, resulting in what is known as the surface fingerprint.[5] The emergence of this structure in the papillary junction gives rise to an internal fingerprint. The direct correlation between the surface and the papillary junction results in identical topographical features (ridges and valleys). Thus, the papillary junction and surface fingerprint have the same structure.

The internal fingerprint is robust against distortion, damage, and spoofing.[6] The imaging of subsurface skin layers is accomplished with a technology known as optical coherence tomography (OCT).[7–10]

Proof-of-concept extraction of the internal fingerprint from OCT scans was accomplished by Bossen et al.,[8] while da Costa et al.[11] strengthened this hypothesis and also evaluated fingerprint deformation using OCT. However, the presence of speckle noise in OCT scans[12] makes the process of extracting the internal fingerprint challenging and also reduces the quality of the fingerprint itself.

Existing techniques for internal fingerprint acquisition from OCT scans often involve simple curvature normalization processes prior to *en face* slice averaging over a manually defined region.[7,8,13] The process of averaging over this region makes an unnecessary assumption: vital fingerprint information (i.e., ridge and valley undulation) is contained within this region. In cases where this assumption holds true, the manually defined region containing the internal fingerprint information is excessive, and thus contains peripheral information that serves to obscure the clarity, contrast, and quality of the internal fingerprint.

The problem addressed in this research is the task of comprehensively and accurately defining the region within which the internal fingerprint resides. Since the undulations of the

*Address all correspondence to: Luke Nicholas Darlow, E-mail: LDarlow@csir.co.za

top-edge of the papillary junction best describe the internal fingerprint, detection thereof is paramount. Papillary junction detection is accomplished through the use of *k*-means clustering, speckle noise reduction, and contrast enhancement.

Features are defined to describe local maxima within one-dimensional (1-D) *en face* depth profiles (also known as A-lines) of an OCT scan. Various clustering algorithms were assessed prior to choosing *k*-means clustering for its simplicity, robustness, and relevance to the data. The result of clustering is a single cluster that describes the location of the center of the papillary junction throughout the OCT scan volume. This is then used to extract smaller image regions in which the papillary junction top-edge can be found. These regions are enhanced and binarized prior to edge detection—a process where the papillary junction top-edge is found.

This paper is structured as follows. Relevant literature is reviewed in Sec. 2. The method used to detect the papillary junction is outlined in Sec. 3. The experimental setup for testing the performance of the proposed approach is detailed in Sec. 4, while the corresponding results are presented in Sec. 5. Conclusions are drawn and future work is suggested in Sec. 6.

2 Related Work

OCT is a tool for the three-dimensional (3-D) imaging of light scattering media such as biological tissue. The drive behind the consistent improvement of OCT technology is undoubtedly medicine. Huang et al.[14] introduced OCT as an emerging technology and highlighted its application in medicine. Although an advantage over other medical imaging technology is the relative hardware simplicity and lower cost of OCT, broad use OCT technology remains relatively expensive. However, its use is diversifying. Other applications of OCT include biometrics, forensics, and document security.[6–10,15]

Liu and Chen[7] captured the internal fingerprint using OCT and showed its direct relationship with the surface fingerprint. They made use of Doppler OCT for liveness detection. A live finger has blood flowing through it and can be detected in this fashion. Bossen et al.[8] showed the feasibility of OCT as a technology for extracting the internal fingerprint. They gave examples of extracted internal fingerprints and tested them against a constructed database. When compared with OCT fingerprints with similar characteristics, within the database, a false rejection rate of 5% was achieved. This was attributed to the small surface area scanned. However, the OCT fingerprints were not tested against corresponding surface fingerprints and quantitative correlation was not established. They identified the need for improved papillary junction segmentation. Zam et al.[9] also extracted internal fingerprints from OCT scans and showed that the capillary pattern yields another biometric and a possible liveness indicator.

An *en face* OCT system was developed by Nasiri-Avanaki et al.[16] This system was able to image the fingertip with dynamic focus, extracting a single *en face* slice to represent the internal fingerprint. This allowed for direct and fast acquisition. Harms et al.[10] also made use of *en face* OCT technology for fast internal fingerprint acquisition. However, the acquired internal fingerprint is not found at a fixed depth and is not uniformly straight, resulting in detail loss. This technique is efficient in terms of acquisition, but negates an important feature of OCT scans: the 3-D representation of fingertip skin. Acquiring a 2-D fingerprint representation from a 3-D scan allows for the development and use of liveness detection. Therefore, the research undertaken here endeavors to solve the issue of internal fingerprint zone detection in 3-D OCT fingertip scans.

OCT has also been used for liveness detection and artificial fingerprint detection. Liu and Buma[17] showed that spectral domain OCT clearly maps eccrine glands in a fingertip. This mapping provided another level[18] of fingerprint biometrics, as well as a means of safe-guarding against spoof attacks. They also provided evidence that the internal fingerprint is present in the papillary junction, but never attempted to locate the associated region accurately.

Cheng and Larin[19] postulated that 3-D visualizations of fingertip skin OCT scans can be used for artificial fingerprint detection and proposed an artificial fingerprint recognition system. They assumed that fake fingerprints exhibit purely homogeneous depth-resolved 1-D signal characteristics, and real fingertip skin exhibits highly inhomogeneous, yet repetitive, signal characteristics. Autocorrelation analysis was then used to identify fake fingerprints.

A later study was conducted by Meissner et al.[20] to demonstrate the capability of OCT in liveness detection. Spoof detection was twofold: (1) qualitative identification of an additional layer and (2) identification of sweat glands. This study showed the superior capability of OCT-based liveness detection when compared with conventional fingerprint liveness detection. Manual liveness estimation on OCT fingertip scans outperformed their automated approach.

Dubey et al.[15] demonstrated the usefulness of OCT in detecting latent fingerprints on poorly reflecting samples. Conventional latent fingerprint acquisition usually involves a physical or chemical process that may damage or alter the fingerprint. OCT is a touchless technology and does not incur the same damage or change. Sano et al.[21] used OCT to assess fingertip skin light scattering characteristics. This information served in the development of a new fingerprint acquisition device that is robust against surface artifacts such as wrinkles and scars.

The internal fingerprints acquired in previous research were not tested or optimized in terms of fingerprint quality. Although it is widely known that the characteristic undulation of the papillary junction describes the ridge and valley structure of the internal fingerprint, accurate detection of this undulation was never undertaken. Instead, OCT images were averaged over a predetermined, manually chosen region, or a single slice was chosen that coincided well with the papillary junction. These approaches captured the papillary junction undulation to a degree sufficient for internal fingerprint representation, but failed to define the region in which pertinent fingerprint information resides. This is inefficient. The research presented here provides an improved and accurate means of extracting an internal fingerprint by detection of the papillary junction. The approach taken is detailed in the next section.

3 Methodology

This section explains the methodology behind papillary junction detection. A system overview diagram, exemplifying the processes involved, is given in Fig. 1. The stratum

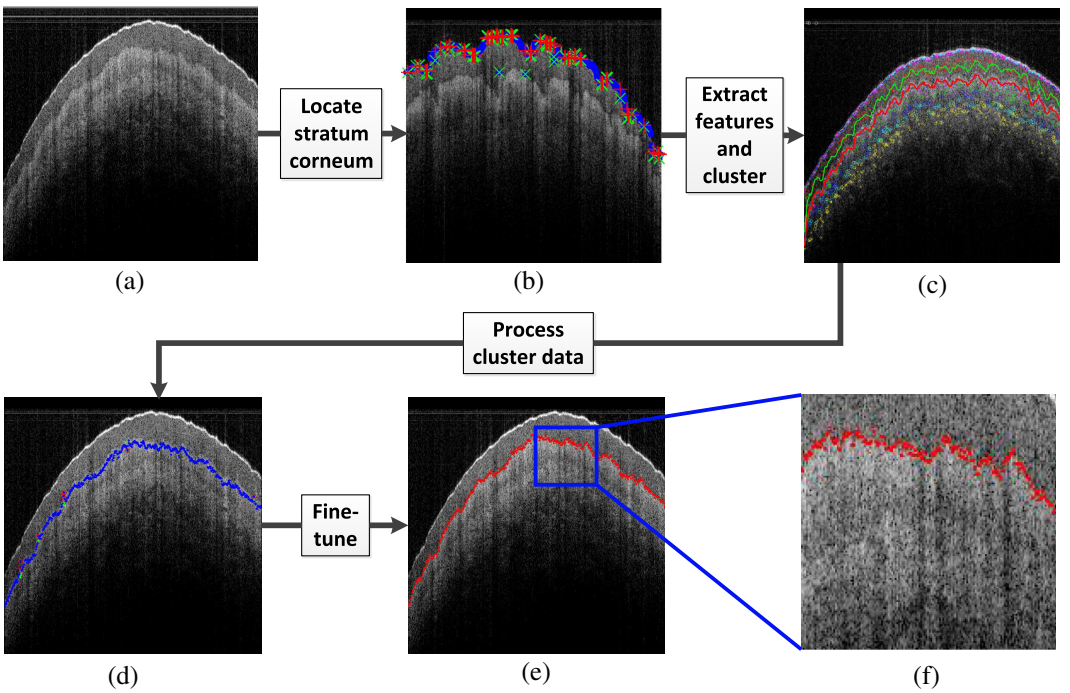

Fig. 1 A system-wide flowchart representation of the processes involved in papillary junction detection. (a) Optical coherence tomography (OCT) image; (b) example of stratum corneum detection–blue dots are the detected stratum corneum points, green crosses are outliers, and red plus' are the adjusted outliers; (c) clustering result example—differing color-shape combinations represent different clusters and the lines show the region extracted for fine-tuning; (d) processed cluster coordinates—blue dots are the points on this OCT image that are within the cluster determined to best describe the papillary junction, red encircles outliers, and green shows corrected outliers; (e) fine-tuned coordinates with denoted region of interest—the red dots show the end result of papillary junction detection; (f) highlighted region of interest in (e).

corneum contour (i.e., surface layer of the skin) is detected as the uppermost brightest contour. Using this location, features are then extracted for *k*-means clustering. Following clustering, the correct cluster is determined and processed. A region containing the top-edge of the papillary junction is then extracted and enhanced, prior to edge detection, for fine-tuning.

Section 3.1 details the process of stratum corneum detection. Section 3.2 provides details on the use of *k*-means clustering, while Sec. 3.3 identifies the problems and corresponding solutions to postprocessing *k*-means clustering coordinates. Fine-tuning is discussed in Sec. 3.4. Cluster analysis was used to determine a rough estimation of the center of the papillary junction. Column-wise intensity local maxima were used as data points. The manner in which this was accomplished is outlined in Sec. 3.2. Image enhancement was used to fine-tune this estimation to coordinates detailing the top-edge of the papillary junction.

3.1 *Stratum Corneum Detection*

In order to extract useful features for clustering from OCT scans, the location of the stratum corneum must be known with a high degree of certainty for the following reasons:

1. Clustering requires data points to have descriptive vectors on which its performance hinges.
2. An intuitive defining feature of any OCT fingertip scan is that of skin layer reflectivity [see Figs. 2(a) and 3].
3. The surface layer, or stratum corneum, is highly reflective and easily detectable.
4. The location of the stratum corneum can be found and used for feature extraction. Algorithm 1 details how the detection of the stratum corneum is accomplished.

Owing to the dimensional continuity of the stratum corneum in corresponding B-scan images (i.e., cross-sectional image slices) in a volume, its location in the previous image is used to determine its current location. The location of the stratum corneum is used to describe an important feature for *k*-means clustering.

3.2 *k-Means Clustering*

The process of classification is intuitive for human beings.[22] For a machine, data classification falls into two categories: supervised and unsupervised. The main difference between the two is human intervention. In supervised approaches, classes are established before classification begins, whereas in unsupervised approaches, classes are derived as natural groupings within the data itself.

Clustering is the unsupervised task where data are segmented into a predefined number of clusters. Data within a cluster have greater similarity to other data inside the same cluster than to data outside the respective cluster.

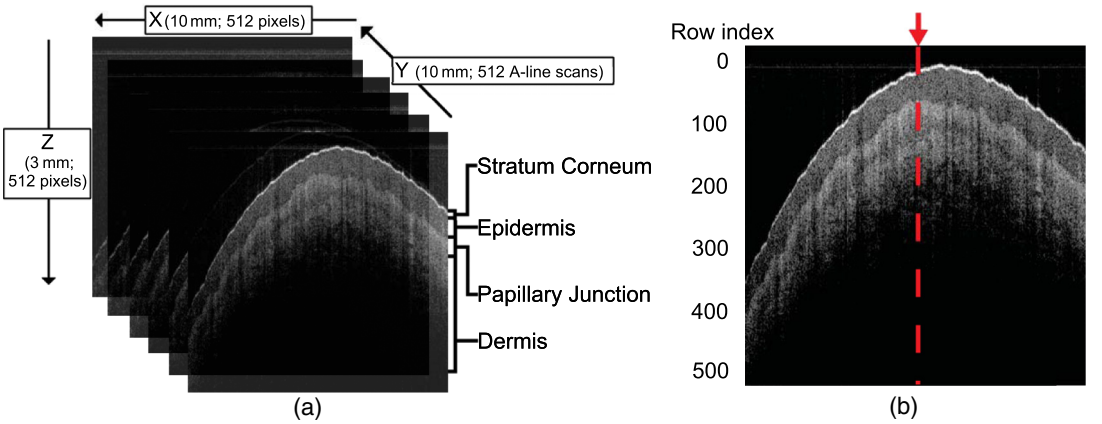

Fig. 2 An example of skin layers and the analysis of a single A-line (column) in a B-scan image: (a) example of a series of B-scans used to construct an OCT volume, with resolution constraints; and (b) the column-wise approach detailed in Fig. 3. Note: dimensions of (a) and (b) are the same.

Although there is no universal definition of a cluster, it can be described in terms of internal homogeneity and external difference.[23]

Choosing which clustering algorithm to use is complex and is dependent on the problem space, with data type and cluster model as factors defining this choice. Xu and Wunsch [23] showed how the selection of a clustering algorithm is complex. Estivill-Castro[24] attempted to explain the large number of available clustering algorithms, arguing that the lack of a decisive idea of "cluster" is to blame.

Initial experimentation was carried out to compare *k*-means,[25] *k*-medoids,[26] fuzzy *c*-means,[27] and expectation-maximization with Gaussian mixture models[28] (EM) clustering. None of these techniques showed any significant performance gain. Therefore, the *k*-means clustering algorithm was chosen for this research because of its simplicity, robustness, and data relevance. It is known to yield good results when datasets are distinct, yet determined clusters cannot overlap.[23] However, in the case of the data used in this research, distinct clusters are advantageous. Although *k* has to be predefined for *k*-means clustering, this limitation does not have an effect on the research carried out here, as a range of *k* values are tested. Further disadvantages of *k*-means clustering are that it is unable to deal with outliers and cluster sizes are assumed to be roughly equal. However, the advantages sufficiently outweigh the relevant disadvantages. Further investigation will be done to determine whether *k*-means is indeed the optimal clustering algorithm, but that is beyond the scope of this work.

The effectiveness of an unsupervised learning technique is confined by the input data and their descriptive features. The choice of *k* can vary (as shown in Sec. 5). Input data points and features were chosen using domain knowledge and logic and through experimentation.

In terms of the data points themselves, an analysis of A-lines was made. An A-line corresponds to a single column in a B-scan. In this way, a column-wise perspective was taken: A-line intensity changes can be seen to yield patterns in the data. Figure 3 exemplifies the analysis performed and shows the correspondence of intensity local maxima to interesting regions in a B-scan. The signals are individually smoothed before the strongest local

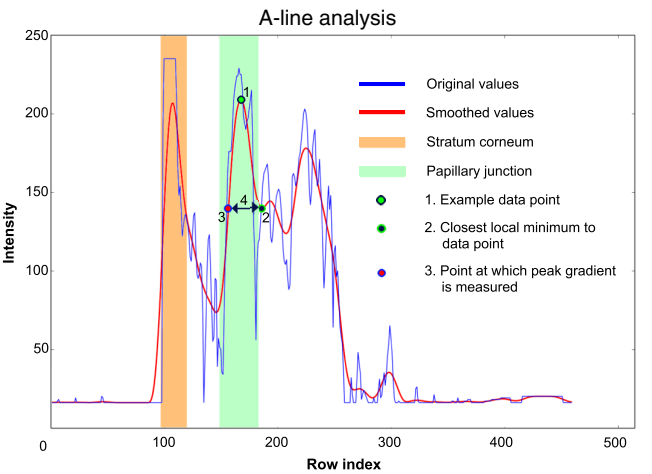

Fig. 3 An example of data used for feature extraction. This is a single A-line profile of an B-scan, as exemplified in Fig. 2.

maxima are extracted as data points. Since each B-scan in the volume contains 512 A-lines, and there are 512 B-scans in the volume, the number of 1-D signals analyzed per volume is 262,144.

Features describing the chosen data points are defined as follows:

1. **Relative distance** to the stratum corneum. The pixel-distance from each local intensity extremum (see point marked as 1 in Fig. 3) to the stratum corneum previously found (refer to Algorithm 1) is computed and normalized. Normalization occurs on a B-scan basis as the distance between the stratum corneum and papillary junction may not be consistent over an entire fingertip, but is relatively consistent in a single B-scan. Each normalization constant is defined as the approximate pixel distance between the stratum corneum and the papillary junction over a single B-scan. This is estimated as the median of all measured distances between the two largest column-wise extrema, and is used to normalize each data-point's distance to the stratum corneum.

Algorithm 1 Stratum corneum location.

```
Input: OCT_volume
Output: Stratum corneum coordinates
for scan_i ← 0 to length(OCT_volume) do
  B_scan ← OCT_volume(scan_i)
  for i ← 0 to length(B_scan) do
    A_line ← B_scan(i)
    smoothed_A_line ← 1-Dsmooth(A_line)
    local_maxima ← extrema(smoothed_A_line)
    if i==0 then
      considered_extrema ← brightest(local_maxima)
      stratum_c[scan_I][i] ← uppermost(considered_extrema)
    End
    Else
      dist_to_prev_stratum ← stratum_c[scan_i−1][i] -
      local_maxima
      index_of_closest ← index_of(min(dist_to_prev_stratum))
      stratum_c[scan_i][i] ← local_maxima[index_of_closest]
    End
  End
End
```

2. **Peak width**. The peak width is defined to be twice the distance to the closest local minimum. Figure 3 exemplifies the closest local minimum and peak width as the points marked as 2 and 4, respectively. The width of the papillary junction is greater than the width of the stratum corneum. As these are the most prominent local extrema, this feature distinguishes them.
3. **Peak standard deviation**. The standard deviation is determined from the region defined by the considered data point (for example, see the point marked as 1 in Fig. 3) and the peak width. Since speckle noise is signal-dependent, the standard deviation will be different in regions with signal (such as the papillary junction) compared with regions without signal.
4. **Peak gradient**. The gradient is measured half way between the considered local maximum and next upper local minimum. The point at which the gradient is measured is the point marked as 3 in Fig. 3. This feature gives a measure of the strength of the increase of the considered data point—a unique trait for pertinent local extrema.
5. **Relative intensity**. The height of the peak is the pixel intensity. This intensity is first normalized using an approach similar to that of the relative distance: on a B-scan basis using the median of all estimated papillary junction peaks (i.e., the second brightest peaks in each A-line). Since the structure of skin results in the papillary junction usually being the second strongest peak (after the stratum corneum), this feature is also unique.

There is a depth-dependent "roll-off" problem inherent in the OCT system used for this research. What this means is that the OCT scanner has different sensitivities at different depths. The result of this depth-dependency when imaging a curved surface such as a fingertip is differing intensity profiles for regions of the fingertip that are further away from the scanner to those close to the scanner. Figure 2(a) shows this—the left regions of these images (where the fingertip skin is spatially lower) have relatively lower pixel intensities. Another artifact present in the OCT scans is the geometric curvature of the fingertip.

One way of dealing with the pixel intensity depth-dependency and fingertip curvature is to make use of a glass slide in the scanning process. However, this mitigates the touchless capability of OCT and introduces fingerprint distortion. A simple solution is found in the form of specific feature normalization. Relative distance and relative intensity are features normalized to mitigate this issue. Although the determined depth of the papillary junction will still be subject to "roll-off," the effect thereof can be spatially normalized. However, that is not the focus of this research.

The ideal output after clustering is a single cluster corresponding to the center of the papillary junction. This cluster would have exactly one data point per column, all of which must be correct. This is not assured to be the case, necessitating the need for cluster postprocessing.

3.3 *Cluster Postprocessing*

The first step in postprocessing clusters is determining which cluster best describes the location of the papillary junction. To do so, the approximate stratum corneum to papillary junction distance, estimated when extracting features, is used along with knowledge of the stratum corneum location. It is only necessary to use a single B-scan to determine the correct cluster because there are enough data points in one B-scan to make this estimation. Furthermore, a B-scan exhibiting a clear stratum corneum and papillary junction can be used. The accuracy of an estimation for all other B-scans will be very similar or worse.

Even after the correct cluster is determined, numerous inconsistencies with this cluster may be present. The following anomalies, along with respective solutions, occur:

- **Problem**: multiple data points from the same A-line may be clustered together.
  Solution: consider only the uppermost (i.e., closest to the surface) data point. If this proves to be wrong, it will be detected as an outlier.
- **Problem**: an A-line may contain no data points in the chosen cluster and, therefore, no coordinate.
  Solution: extrapolate coordinates based on surrounding cluster points.

- **Problem**: some data points in the chosen cluster may not correspond to the papillary junction.
  Solution: identify and process these points as outliers, extrapolating correct coordinate based on surrounding correct cluster points.

The papillary junction coordinates, acquired after processing cluster output, correspond to the center of the papillary junction. The precision of these coordinates was measured relative to ground-truth coordinates in Sec. 5. Although these coordinates may be a good estimate of the papillary junction location, the richest fingerprint data are found in the papillary junction top-edge. The papillary junction center coordinates must, therefore, be adjusted. This adjustment will now be discussed.

3.4 *Fine-Tuning*

The fine-tuning process is designed to accurately determine the location of the papillary junction top-edge. This edge contains information most pertinent to the internal fingerprint. An image containing the edge is extracted from each B-scan and enhanced prior to edge detection. These steps are detailed below.

The region containing the papillary junction top-edge can be extracted using the papillary junction center coordinates. For each A-line, this region starts at the cluster-determined papillary junction center coordinate (that defines the pixel-wise depth) and ends 25 pixels toward the stratum corneum. The result is an image (of 25×512 pixels) extracted from each B-scan. An example of this region and the corresponding extracted image can be seen in Fig. 4. Detection of the line defining the top-edge is made difficult by the presence of speckle noise.

Speckle noise is multiplicative and signal degrading. Darlow et al.[29] reviewed six state-of-the-art speckle reduction techniques pertinent to OCT fingertip scans. The optimized blockwise nonlocal means (OBNLM) algorithm was found to be the best performing approach in terms of signal-to-noise ratio improvement, error minimization, and structural similarity improvement (SSIM).

Coupé et al.[30] proposed an adaptation of the nonlocal (NL)-means filter[31] to combat speckle in magnetic resonance images. An adaptive general speckle model was used to reformulate the original NL-means approach. The original result was a powerful speckle reducing filter. However, it was computationally expensive. The OBNLM[32] filter was developed as an optimization. Since the speckle model is similar to that of OCT images, this speckle reduction technique was used.

Although a product of speckle noise reduction is contrast and clarity improvement, this improvement is limited by the inherent contrast of OCT scans. Pixel intensity in an OCT scan is defined by the reflectivity of the imaged region. The stratum corneum and papillary junction have relatively higher pixel intensities compared with the dermis and epidermis. Since the goal of this research is to locate the papillary junction, contrast enhancement was used to highlight this region.

Contrast enhancement can be seen as an operation where the concerned pixel intensity range is appropriately stretched.[33,34] Because the region of interest is the papillary junction and the pixel intensities within this region are inconsistent, a local contrast enhancement approach is preferred.

The approach used in this research is local intensity normalization. A windowed region is considered in which the local intensities are stretched to fit a defined range. Following local normalization, unsharp masking was used to further clarify the papillary junction top-edge. Unsharp masking is a filter that amplifies the high-frequency components of an image. This is accomplished through comparison of a blurred version of the image to the original. In this way, unsharp masking is able to enhance edges.

Once these image regions are enhanced and binarized, the top-edge is detected. For each extracted region, the uppermost white pixels are determined to represent this edge.

Examples and results of this process can be found in Sec. 5. Experimental procedures to assess the effectiveness of papillary junction detection are discussed in the following section.

4 Experimental Setup

Figure 5 shows the layout of the OCT system (OCS1300SS, Thorlabs) used to acquire an OCT scan for this research. The central wavelength is 1325 nm. The spectral bandwidth is 100 nm. It has an axial scan rate of 16 kHz, a coherence length of 6.0 mm, and an average output power of 10.0 mW.

Figure 2 shows an example of the layout of a single OCT scan. A scan results in 512 individual image slices (called B-scans, in the direction of the Y-axis), with a lateral resolution of 512 pixels (X-axis), and a depth resolved resolution of 512 pixels (Z-axis). Each individual column in a B-scan

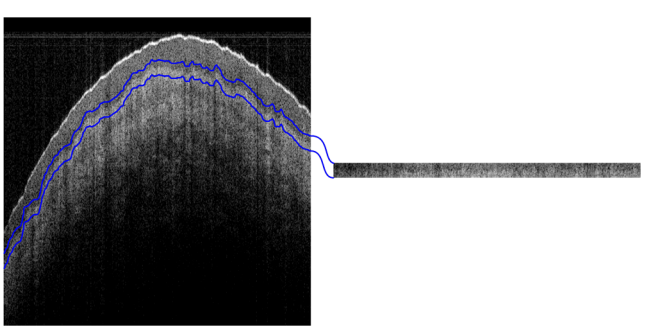

Fig. 4 An example of the region containing the papillary junction top-edge. This extracted image is given on the right, and used as input for fine-tuning.

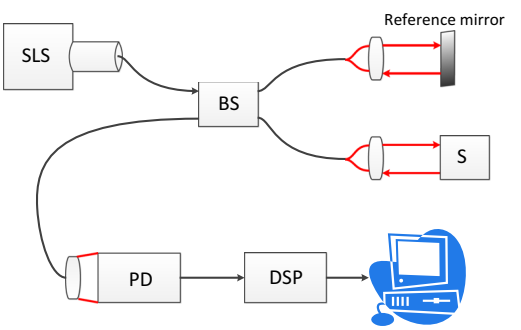

Fig. 5 The OCT system used. SLS: swept laser source; BS: beam splitter; S: sample; PD: standard photodiode detector; DSP: digital signal processor. Reproduced with permission, courtesy of Darlow et al.[29]

(i.e., an A-line) represents a depth of 3 mm. The area scanned is 10×10 mm^2 (X- and Y- directions, respectively). This results in a resolution of 19.53 μm per pixel in the X- and Y-directions, and a resolution of 5.86 μm per pixel in the Z-direction.

No glass plate was used to stabalize during scanning. Instead, the scan was carried out repetitively until such a time as a suitably stable iteration was established. Any changes induced by heartbeat and blood-flow do not induce changes on a scale that affects the papillary junction shape, and thus the internal fingerprint.

Papillary junction ground truth was established for each B-scan. The ground truth was manually established using the GNU Image Manipulation Program found in Ref. 35.

The papillary junction top-edge coordinates were processed for each image. Thirteen human volunteers were asked to provide ground-truth estimates for five B-scans. These data were used to establish an estimate of "human error" associated with ground-truth estimation.

Papillary junction coordinates found automatically (P) were tested against corresponding ground-truth coordinates (G). Mean squared error [MSE, Eq. (1)], the Hausdorff[36,37] distance metric [H, Eq. (2)], SSIM[38,39] [Eq. (4)], and RMS-contrast[40] [Eq. (5)] were used.

Both the MSE and H are known to give a good indication of the error between two datasets. MSE is defined as

$$\mathrm{MSE}(s,t) = \frac{1}{N}\sum_{i=1}^{N}(s_i - t_i)^2, \tag{1}$$

where s and t are the compared signals, N is the total number of data points, and i refers to the data point being compared.

MSE is the average squared error, indicative of the deviation of two datasets. MSE works well on a finer scale as it efficiently shows fluctuation between datasets. The Hausdorff distance is defined as

$$H(X,Y) = \max[h(X,Y), h(Y,X)], \tag{2}$$

where

$$h(X,Y) = \max_{x\in X}\min_{y\in Y}\|x-y\|, \tag{3}$$

where x and y are the points in X and Y, respectively.

H defines the degree of mismatch between two subspaces; that is, it indicates the maximum of the shortest distance between any point in one subset and any point in the other.

Although these metrics are able to give a good indication of the error between two subsets, they suffer from case-specific dependence. The ranges of MSE and H are not normalized and can deviate from what is expected. For instance, coordinates with a global offset may yield bad values of MSE and H. On the other hand, smoothed coordinates (from 2-D Gaussian blurring) may yield good values of MSE and H. However, globally offset coordinates have better structural value.

To accommodate the lack of suitable metrics for image comparison, Wang and Bovik[38] developed a case independent universal quality index. The SSIM metric was later developed by Wang et al.[39] by making use of the assumptions that the human visual system is able to detect SSIM, and that local pixels hold spatially structural information. SSIM is defined as

$$\mathrm{SSIM}(u,v) = \frac{(2\mu_u\mu_v + S_1)(2\sigma_{uv} + S_2)}{(\mu_u^2 + \mu_v^2 + S_1)(\sigma_u^2 + \sigma_v^2 + S_2)}, \tag{4}$$

where SSIM is calculated in a local, windowed fashion. u and v define the measured windows. μ_u is the mean of u, and μ_v is the mean of v. The variances of u and v are σ_u^2 and σ_v^2, respectively. σ_{uv} is the covariance of u and v. S_1 and S_2 are the stabilization variables.

Considering the context of fingerprints, another useful measure would be one of the contrasts. RMS-contrast is an accepted means of measuring the contrast of images. Therefore, it is suited to measuring the contrast of 2-D coordinates. It is defined as

$$\mathrm{RMS\_contrast}(\mathrm{Im}) = \sqrt{\frac{1}{WL}\sum_{i=0}^{W-1}\sum_{j=0}^{L-1}[\mathrm{Im}_{ij} - \mathrm{mean}(\mathrm{Im})]^2}, \tag{5}$$

where Im is a 2-D array, W and L are the dimensions, Im_{ij} are the i'th and j'th elements, while mean(Im) is the average of Im.

Table 1 Image enhancement parameters.

Enhancement technique		Parameters
OBNLM	Stage one	Search area size = 12
		Patch size = 2
		Smoothing = 2
	Stage two	Search area size = 6
		Patch size = 1
		Smoothing = 1
	Stage three	Search area size = 6
		Patch size = 1
		Smoothing = 1
Local normalization		Amplitude = 9
		Radius = 20
		Neighborhood smoothness = 40
		Average smoothness = 40
Unsharp masking		Radius = 50
		Amount = 1.5
		Threshold = 15%

The MATLAB® k-means clustering algorithm was used in this research. Predefining k (number of clusters) is an unsolved problem. Therefore, a wide range of k was tested.

The parameters for the image enhancement involved in fine-tuning are provided in Table 1. OBNLM was applied in three stages although stage three is a repetition of stage two. The application of OBNLM in this manner further clarifies the papillary junction top-edge.

Comparisons with ground-truth values were made at two stages: after cluster coordinates were processed, and after fine-tuning (i.e., against P). This allowed for a quantitative measure of the performance increase due to fine-tuning.

To test the performance of k-means clustering as applied to this problem, the number of clusters (k) and the number of data points (n) per A-line were varied. The aforementioned metrics were processed for k from 2 to 50, and for n from 2 to 33. The limitation on n is due to the fact that there are a limited number of local extrema per A-line in a B-scan.

Qualitative results of k-means clustering were used to give an indication of the process involved and its performance. With regard to the comparison made after the clustering step, the ground-truth papillary top-edge was offset by the mean difference between the top and bottom edges.

As for P, these coordinates were directly compared with G. Qualitative results were also used for assessment. Since the fine-tuning process involves many parameters regarding image enhancement, an optimization was not attempted.

A final assessment technique was considered the output fingerprint. This fingerprint was obtained by converting G and P into grayscale fingerprint images, followed by various contrast enhancement techniques. The details are not discussed as this deviates from the focus of this research. After processing a fingerprint from both P and G, minutiae detection was performed using the open source fingerprint recognition toolkit, SourceAFIS.[41]

5 Results and Discussion

The accuracy of the results depends on the accuracy of ground-truth estimation. This manual estimation was done as accurately as possible, but cannot be assumed to be perfect. Therefore, some deviation between G and P is attributable to this. In order to quantify this human-error, 13 volunteers supplied ground-truth coordinates for five B-scans.

Each volunteer's ground-truth estimations were compared to all other estimations. A mean MSE of 15.6 ± 6.0, a mean H of 157.3 ± 22.0, and an RMS-contrast of 103.1 ± 0.4 were obtained. SSIM is not given in this case as five

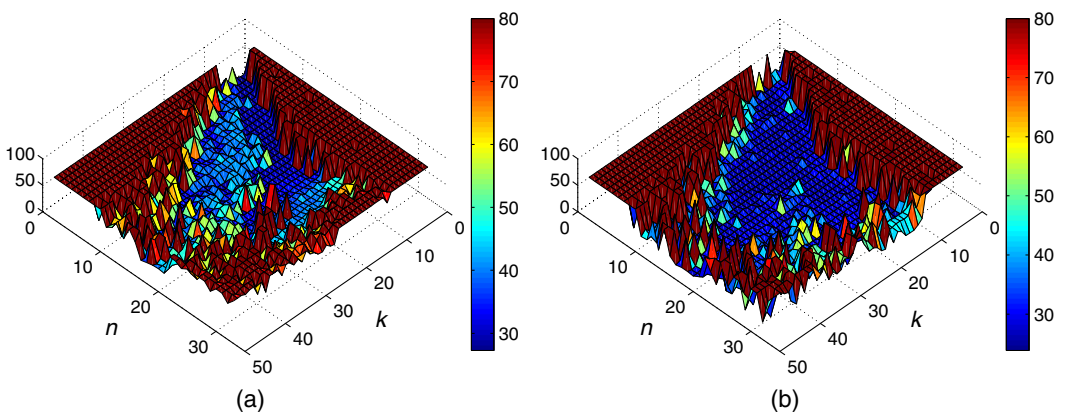

Fig. 6 MSE results: k versus n. Lower mean-squared-error (MSE) values are shown in blue, higher values in red, with intermediate values in the color ranges in between. (a) shows MSE results after clustering, while (b) shows MSE results after fine-tuning.

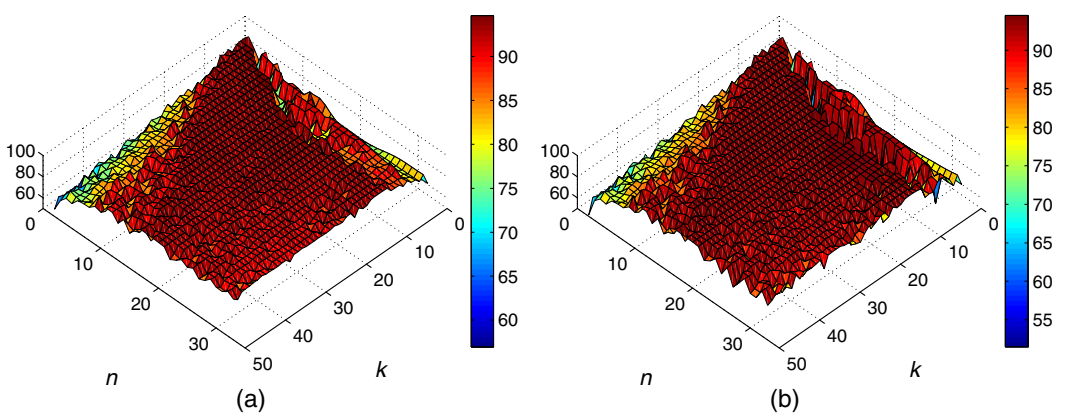

Fig. 7 SSIM results: k versus n. Lower structural similarity improvement (SSIM) values are shown in blue, higher values in red, with intermediate values in the color ranges in between. (a) shows SSIM results after clustering, while (b) shows SSIM results after fine-tuning.

ground-truth estimations does not yield a meaningful SSIM result. All human-error measured was comparable with results automatically obtained.

Figures 6 through 9 show the results through the tested ranges of k and n. In each of these figures, (a) shows results after the clustering step, while (b) shows the results after fine-tuning. When creating these graphs, differing upper and lower limits were set to exemplify the region of interest.

Lower MSE and H values, higher RMS-contrast values, and high SSIM values are seen in the region corresponding to a roughly even ratio of $k:n$. This behavior is attributable to the fact that each data point in an A-line can be unique to that signal and thus falls within a unique cluster. All four metrics correlate consistently: there is a region of stable high performance. This fact allowed for definition of a manually chosen region of interest (see Fig. 10).

The error metrics measured (MSE and H) strongly show a saddle point in which good performance is achieved. All values of n and k encompassing this region were tested. The results for SSIM also exhibit behavior akin to this. RMS-contrast results indicate similar stable-region behavior, but show an increase with combinations of high k and low n.

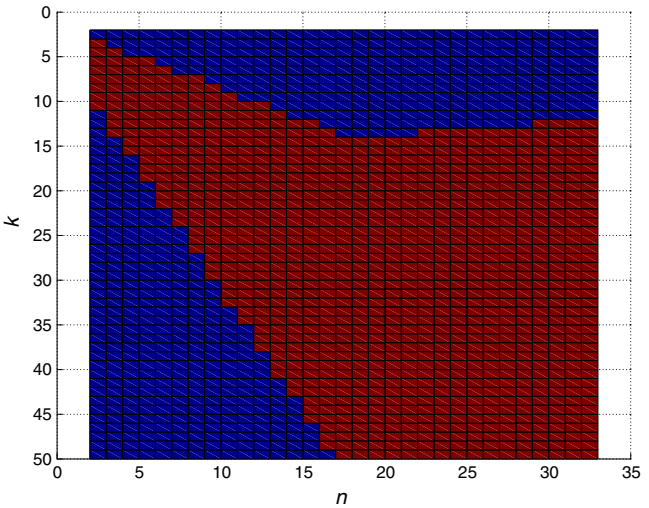

Fig. 10 The experimentally chosen region of interest. The region in which stable performance can be found is in the center (red), while poor performance is found with n and k combinations corresponding to the outer (blue) regions.

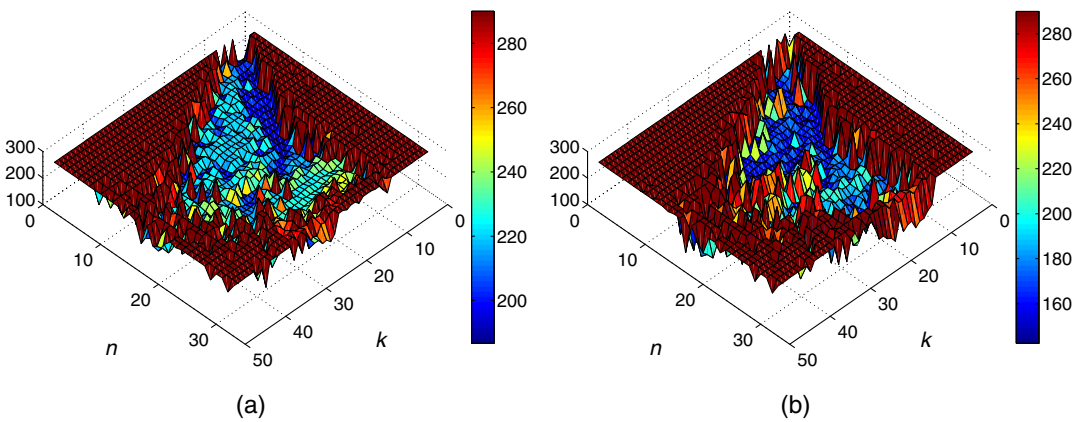

Fig. 8 H results: k versus n. Lower H values are shown in blue, higher values in red, with intermediate values in the color ranges in between. (a) shows H results after clustering, while (b) shows H results after fine-tuning.

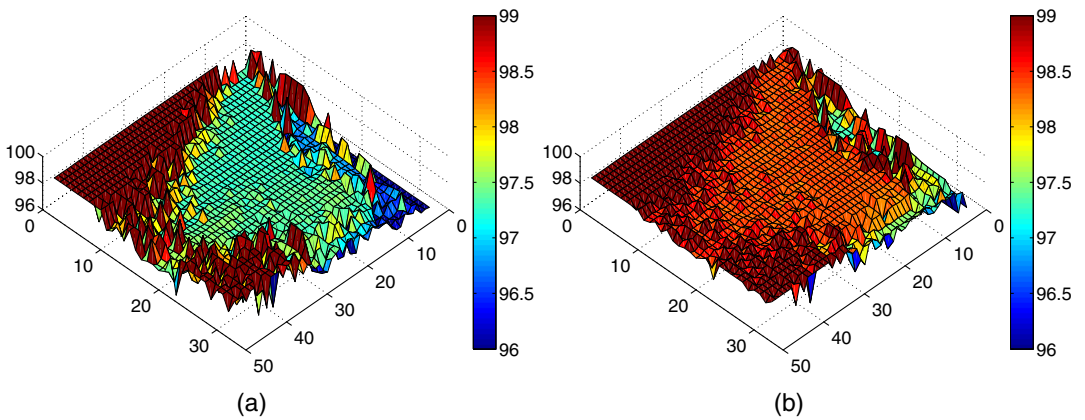

Fig. 9 RMS-contrast results: k versus n. Lower RMS-contrast values are shown in blue, higher values in red, with intermediate values in the color ranges between. (a) shows RMS-contrast results after clustering, while (b) shows RMS-contrast results after fine-tuning.

Table 2 Quantitative results after clustering. ROI denotes the determined region of interest. RMS-C is the RMS-contrast. Human results are included for comparison.

	Result after clustering			
Metric	Entire region	ROI	Best	Human
MSE	1337.6 ± 5521.9	50.7 ± 18.2	27.3	15.6 ± 6.0
H	697.0 ± 976.2	240.4 ± 34.0	186.7	157.3 ± 22.0
SSIM	$88.5 \pm 6.7\%$	$92.4 \pm 2.2\%$	94.7%	—
RMS-C	100.0 ± 13.7	97.6 ± 0.6	141.7	103.1

Table 3 Quantitative results after fine-tuning. ROI denotes the determined region of interest. RMS-C is the RMS-contrast. Human results are included for comparison.

	Result after fine-tuning			
Metric	Entire region	ROI	Best	Human
MSE	1266.1 ± 5263.7	38.4 ± 16.8	23.6	15.6 ± 6.0
H	684.6 ± 966.2	218.2 ± 53.6	142.5	157.3 ± 22.0
SSIM	$87.4 \pm 9.0\%$	$92.2 \pm 3.3\%$	94.5%	—
RMS-C	101.1 ± 13.8	98.4 ± 0.4	143.3	103.1

This is due to the high contrast error resulting from cluster coordinate processing.

What should be noted is the relationship between Figs. 6(a) and 6(b). The region of lower MSE values is larger and more stable after fine-tuning. This shows the effectiveness of the fine-tuning step. These evident improvements after fine-tuning are present in all measurements. The result is a widening and improved stability of the clearly evident stable region.

Tables 2 and 3 give the average results for all metrics over the entire tested region and the region of interest after clustering and fine-tuning, respectively. Human performance is included. The error related metrics (MSE and H) have very large standard deviations over the entire tested region. This is because their respective ranges are unbounded and potentially large. In this case, large average error results take into account combinations of k and n that are unsuitable. For this reason, results over the determined region of interest are given. The standard deviation of all metrics measured decreases after fine-tuning.

With an average MSE of 50.7 ± 18.2 and a minimum MSE of 27.3 (when $k = 4$ and $n = 2$) after the clustering step, the success of clustering for detection of the center of papillary junction is confirmed. Fine-tuning decreased the average MSE to 38.4 ± 16.8 and the minimum MSE to 23.6 (when $k = 33$ and $n = 13$). Similar improvements for H can be seen—with minimum H values of 186.7 and 142.5 after clustering and fine-tuning, respectively.

Since the structural layout of coordinates found after clustering is very similar to P, SSIM results are very similar. Maximum SSIMs are 94.7% and 94.5% after clustering and fine-tuning, respectively. The structural relationships of ridges and valleys correspond well between G and P.

After the clustering step, the maximum RMS-contrast in the stable region was 141.7, while the average RMS-contrast in the stable region was 97.6 ± 0.6. The maximum RMS-contrast after fine-tuning was 143.3, while the stable-region average was 98.4 ± 0.4. In the context of internal fingerprints, RMS-contrast gives a quantitative assessment of ridge and valley definition. The RMS-contrast for the ground-truth coordinates is 97.7—generally less than that of P.

Figure 11 gives accompanying qualitative results for clustering. Each cluster is denoted by a unique marker type and color combination. What is noteworthy here is that the 1:1 ratio of k:n gives better results, justifying the saddle regions appearing in the measured metrics' results.

Figure 12 exemplifies the process involved in fine-tuning. The contribution of OBNLM speckle noise reduction is highly evident here in both the preservation of edges and the smoothing of homogeneous areas. The application of speckle reduction in three phases increases the signal clarity, as is evident in the progression depicted in Figs. 12(b)–12(d). Local contrast enhancement and unsharp masking are shown in Figs. 12(e) and 12(f), respectively. This enhancement process dramatically improved the clarity and contrast of the papillary junction top-edge.

The contribution of this research is noteworthy: the papillary junction top-edge is difficult for a human to detect, but the results obtained (both quantitative and qualitative) show

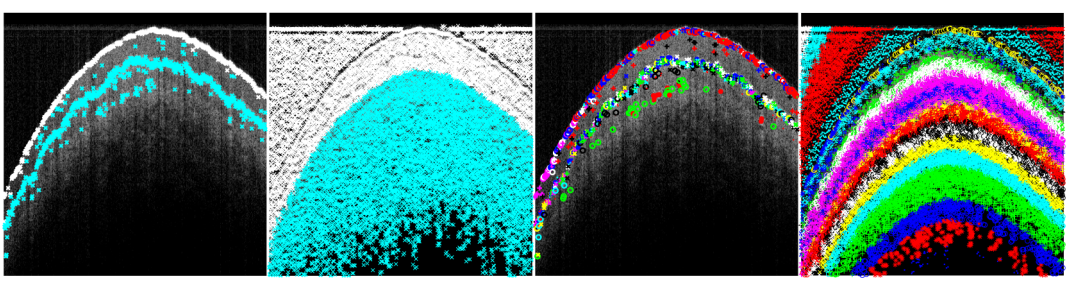

Fig. 11 Qualitative results of k-means clustering for varying combinations of n and k: for (a), $n = 2$ and $k = 2$; for (b), $n = 30$ and $k = 2$; for (c) $n = 2$ and $k = 30$; and for (d) $n = 30$ and $k = 30$. Each color-shape combination represents a different cluster.

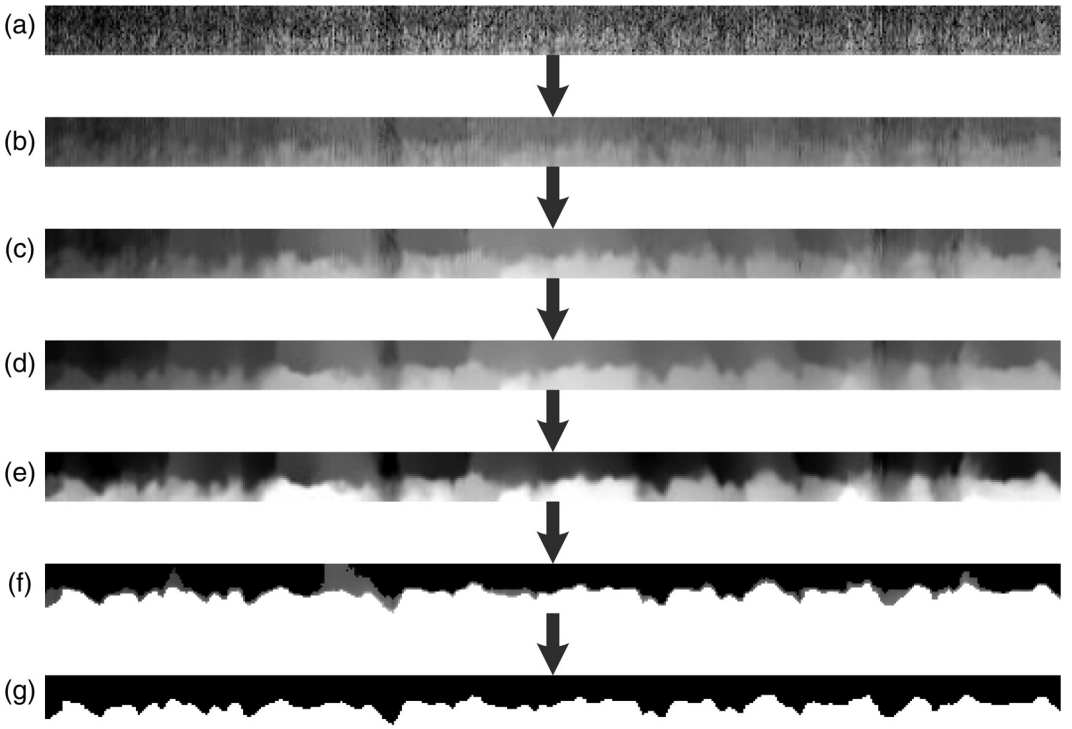

Fig. 12 Fine-tuning process. From the top down: (a) papillary junction top-edge; (b) stage one optimized blockwise nonlocal means (OBNLM) speckle noise reduction; (c) stage two OBNLM speckle noise reduction; (d) stage three OBNLM speckle noise reduction; (e) intensity local normalization; (f) unsharp masking; and (g) thresholding. The flattening process is discussed in Sec. 3.4, and exemplified in Fig. 4.

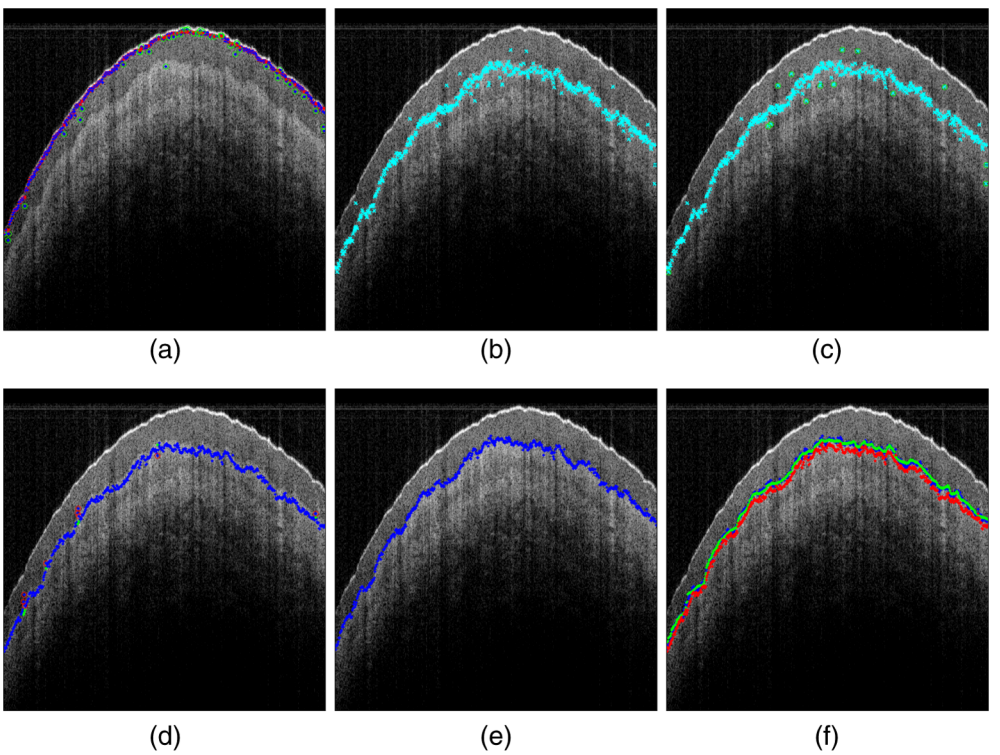

Fig. 13 Visual overview of entire process: (a) locating stratum corneum; (b) chosen cluster data; (c) finding outliers (green circles show outliers); (d) coordinates found by processing clusters; (e) adjusted coordinates after fine-tuning and (f) a comparison of coordinates [the top smooth (green) points are ground truth, the top deviating (blue) points are fine-tuned coordinates, and the bottom deviating (red) points are cluster coordinates].

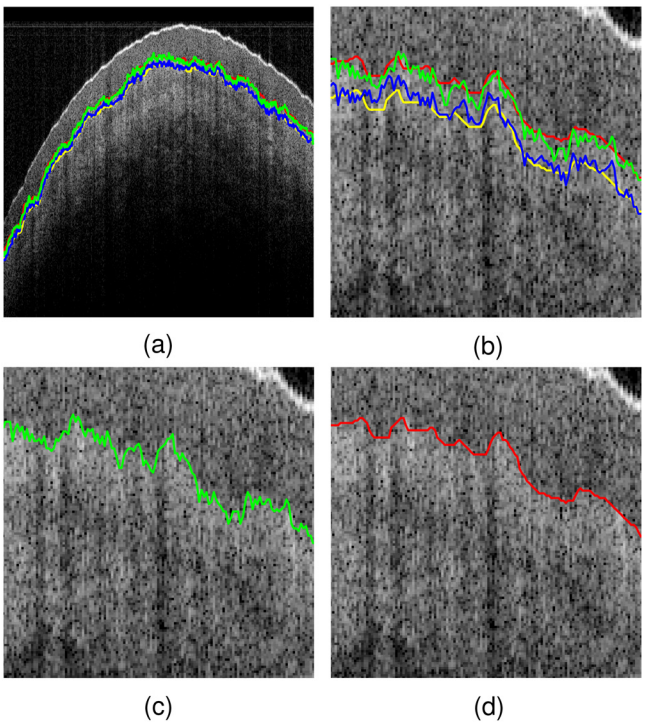

Fig. 14 Comparison of coordinates: (a) shows coordinates for one B-scan; (b) is the region of interest defined in (a). (c) and (d) also show the region of interest defined in (a), exemplifying *P* and *G*, respectively. The top deviating (green) line represents *P*; the bottom deviating (blue) is the clustering coordinates results; the top smooth (red) is *G*; and the bottom smooth (yellow) line is the adjusted *G* (for comparison with the clustering coordinates).

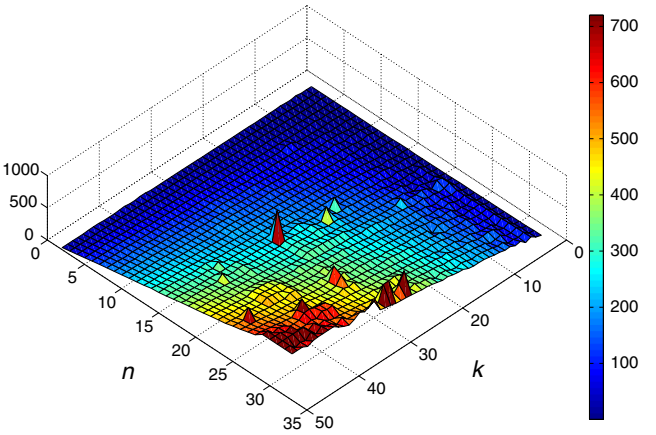

Fig. 16 Timing results cluster processing.

this process to be successful through the approach developed in this research.

An encompassing visual explanation of the entire process present in this research can be seen in Fig. 13. Figure 14 shows all coordinates found and used for a single B-scan. A comparison of Figs. 14(c) and 14(d) shows the consistency present between *P* and *G*. The evident deviation is slight and can often be attributed to human error (in *G*).

Figure 15 shows the fingerprint processed from *G* and *P*. These fingerprints constitute an area of 10×10 mm^2. The process involved in extracting a fingerprint is exemplified along with processed fingerprints and corresponding minutia. The minutia detected are very similar between fingerprints obtained using *G* and *P*, with few falsely detected minutia found in each case.

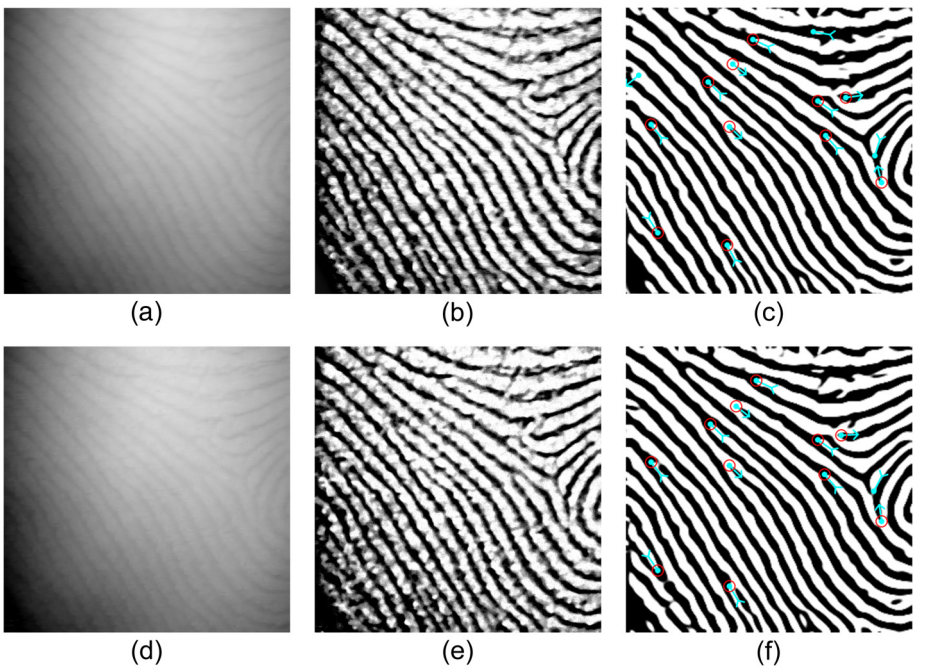

Fig. 15 Fingerprint extraction from coordinates. The top row shows ground-truth coordinates, while the bottom row shows determined coordinates. From left to right: (a) and (d) are coordinates converted into an image; (b) and (e) exemplify the process involved; (c) and (f) show the detected minutiae from processed coordinates.

Figure 16 gives the timing results for cluster processing. The minimum time taken is 0.36 s. Large values of n (resulting in exponentially more data) and k yield unsatisfactory times, upward of several minutes. However, these combinations have no advantages. Timing results for the fine-tuning step are not included as this process is independent of k and n.

Considering the extensive testing of n and k parameter combinations undertaken here, a quantitative standard for the choice thereof can be given. This is based on two observations:

1. The saddle point for a high performance is large and is defined by a roughly equal ratio of n and k (see Figs. 6 through 9);
2. Processing time increases as both n and k increase (see Fig. 16).

Equal ratios and lower values of n and k are suitable. A noteworthy constraint on this choice is that lower values of n and k are less stable. This is characterized by the shape of the identified region of stability: there is a "tapering" toward the lowest possible values. Owing to this, a balance between stability of performance and performance efficiency is reached when both n and k are roughly equal and sufficiently buffered from their corresponding lowest possible values, i.e., $4 < (k = n) < 10$.

6 Conclusion and Future Work

This research addressed the problem of automatic papillary junction detection in OCT fingertip scans using machine learning and image enhancement algorithms. Accurate knowledge of the location of the papillary junction can be used as an asset for the extraction of an internal fingerprint. This is a significant deviation from previous research, where a manually determined region is used for internal fingerprint acquisition. Using coordinates obtained through this research, an internal fingerprint of high quality (in terms of minutiae detected) was acquired.

K-means clustering was used to find the approximate location of the center of the papillary junction. A comparison was made against a ground-truth estimation, yielding a minimum MSE of 27.3 (with $k = 4$ and $n = 2$). The approximate location, determined by clustering, was improved by the use of edge detection on enhanced extracted image regions, yielding a minimum MSE of 23.6.

In combination with this low MSE, a maximum SSIM of 94.5%, a minimum H of 147.5, and an average RMS-contrast of 98.4 were achieved after fine-tuning. These results indicate low error, high SSIM, and improved contrast when compared with ground-truth coordinates.

The human error associated with ground-truth estimation was quantified. The best results obtained automatically were within ranges comparable with human error. Therefore, the papillary junction detection approach presented in this research performs relatively well when compared to humans.

All metrics measured corresponded well. There was a region, roughly defined by combinations of equal n and k, where MSE, H, SSIM, and RMS-contrast yielded stable results. Furthermore, a quantitative standard for the choice of k and n combinations was derived.

Various qualitative results were given to exemplify the entire process; details on the fine-tuning and clustering process; and resultant processed fingerprints. The qualitative results correlated well with the quantitative assessment.

All results obtained strongly showed the success of clustering and image enhancement procedures as tools for papillary junction detection.

Future work entails optimizing the computational aspects of the methods involved. The acquisition of a high quality internal fingerprint will be detailed in a future work.

The results obtained were found by processing a 3-D OCT scan. Owing to the nature of the papillary junction detection algorithm, the only assumption made is that the 1-D depth profiles analyzed are consistent with skin structure. This consistency remains regardless of the OCT scan dimensions. This will be shown in future work.

An advantageous future work is to explore different clustering techniques. However, it would be unlikely that any minor accuracy gain would be significant. This is because the fine-tuning step mitigates any minor changes due to cluster result deviations.

Acknowledgments

The authors would like to thank the CSIR National Laser Center for providing the OCT scans used in this research. The authors also would like to thank Dr. Karen Bradshaw, Mrs. Jeanne Cÿrus, and Miss Piette Cÿrus for their contribution in sharpening this work to a state of excellence.

References

1. D. Maltoni et al., *Handbook of Fingerprint Recognition*, 2nd ed., Springer, London (2009).
2. M. Tistarelli, S. Z. Li, and R. Chellappa, *Handbook of Remote Biometrics*, Springer, Switzerland (2009).
3. T. Matsumoto et al., "Impact of artificial gummy fingers on fingerprint systems," *Proc. SPIE* **4677**, pp. 275–289 (2002).
4. S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: security and privacy concerns," *IEEE Security Privacy* **1**(2), 33–42 (2003).
5. C. Champod et al., *Fingerprints and Other Ridge Skin Impressions*, IET, United Kingdom (2004).
6. S. Chang et al., "Optical coherence tomography used for security and fingerprint-sensing applications," *IET Image Process.* **2**, 48–58 (2008).
7. G. Liu and Z. Chen, "Capturing the vital vascular fingerprint with optical coherence tomography," *Appl. Opt.* **52**(22), 5473–5477 (2013).
8. A. Bossen, R. Lehmann, and C. Meier, "Internal fingerprint identification with optical coherence tomography," *IEEE Photon. Technol. Lett.* **22**(7), 507–509 (2010).
9. A. Zam et al., "Feasibility of correlation mapping optical coherence tomography (cmoct) for anti-spoof sub-surface fingerprinting," *J. Biophoton.* **6**(9), 663–667 (2013).
10. F. Harms, E. Dalimier, and A. C. Boccara, "*En-face* full-field optical coherence tomography for fast and efficient fingerprints acquisition," *Proc. SPIE* **9075**, 90750E (2014).
11. H. S. G. da Costa et al., "Evaluation of fingerprint deformation using optical coherence tomography," *Proc. SPIE* **8946**, 89460I (2014).
12. J. M. Schmitt, S. H. Xiang, and K. M. Yung, "Speckle in optical coherence tomography," *J. Biomed. Opt.* **4**(1), 95–105 (1999).
13. K. V. Larin and Y. Cheng, "Three-dimensional imaging of artificial fingerprint by optical coherence tomography," *Proc. SPIE* **6944**, 69440M (2008).
14. D. Huang et al., "Optical coherence tomography," *Science* **254**(5035), 1178–1181 (1991).
15. S. K. Dubey et al., "Simultaneous topography and tomography of latent fingerprints using full-field swept-source optical coherence tomography," *J. Opt. A: Pure Appl. Opt.* **10**(1), 015307 (2008).
16. M. Nasiri-Avanaki et al., "Anti-spoof reliable biometry of fingerprints using *en-face* optical coherence tomography," *Opt. Photon. J.* **1**, 91–96 (2011).
17. M. Liu and T. Buma, "Biometric mapping of fingertip eccrine glands with optical coherence tomography," *IEEE Photon. Technol. Lett.* **22**, 1677–1679 (2010).
18. A. K. Jain, Y. Chen, and M. Demirkus, "Pores and ridges: high-resolution fingerprint matching using level 3 features," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(1), 15–27 (2007).
19. Y. Cheng and K. V. Larin, "Artificial fingerprint recognition by using optical coherence tomography with autocorrelation analysis," *Appl. Opt.* **45**, 9238–9245 (2006).

20. S. Meissner, R. Breithaupt, and E. Koch, "Defense of fake fingerprint attacks using a swept source laser optical coherence tomography setup," *Proc. SPIE* **8611**, 86110L (2013).
21. E. Sano et al., "Fingerprint authentication device based on optical characteristics inside a finger," in *Proc. 2006 Conf. on Computer Vision and Pattern Recognition Workshop, CVPRW'06*, pp. 27–27, IEEE, New York (2006).
22. M. R. Anderberg, "Cluster analysis for applications," Technical Report, DTIC Document (1973).
23. R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Trans. Neural Netw.* **16**(3), 645–678 (2005).
24. V. Estivill-Castro, "Why so many clustering algorithms: a position paper," *ACM SIGKDD Explor. Newslett.* **4**(1), 65–75 (2002).
25. J. A. Hartigan and M. A. Wong, "Algorithm as 136: a *k*-means clustering algorithm," *Appl. Stat.* **28**, 100–108 (1979).
26. H.-S. Park and C.-H. Jun, "A simple and fast algorithm for *k*-medoids clustering," *Exp. Syst. Appl.* **36**(2), 3336–3341 (2009).
27. J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy *c*-means clustering algorithm," *Comput. Geosci.* **10**(2), 191–203 (1984).
28. T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.* **13**(6), 47–60 (1996).
29. L. N. Darlow, S. S. Akhoury, and J. Connan, "A review of state-of-the-art speckle reduction techniques," *Proc. SPIE* **9445**, 944523 (2014).
30. P. Coupé et al., "Bayesian non local means-based speckle filtering," in *Proc. 5th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro, (ISBI 2008)*, pp. 1291–1294, IEEE, New York (2008).
31. A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.* **4**(2), 490–530 (2005).
32. P. Coupé et al., "An optimized blockwise nonlocal means denoising filter for 3-d magnetic resonance images," *IEEE Trans. Med. Imaging.* **27**(4), 425–441 (2008).
33. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, Inc., Noida (1989).
34. E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities*, Elsevier, Melbourne, Australia (2004).
35. The GIMP Development team, "GIMP 2.8.10," 2013, www.gimp.org (25 November 2014).
36. D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(9), 850–863 (1993).
37. M.-P. Dubuisson and A. K. Jain, "A modified hausdorff distance for object matching," in *Proc. 12th IAPR Int. Conf. on Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing*, Vol. 1, pp. 566–568, IEEE, New York (1994).
38. Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.* **9**(3), 81–84 (2002).
39. Z. Wang et al., "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
40. E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A* **7**(10), 2032–2040 (1990).
41. V. Robert, "SourceAFIS 1.7," 2014, www.sourceafis.org (25 November 2014).

Luke Nicholas Darlow is a student at Rhodes University, South Africa. After obtaining his honours degree in computer science he became a student researcher at the Council for Scientific and Industrial Research. Currently, he is working toward his master's degree. His research involves endeavoring to find the best possible way to extract internal fingerprints from optical coherence tomography scans as a replacement for current fingerprint-scanning technology.

James Connan received his MSc degree in computer science from the Stellenbosch University and is currently a senior lecturer at Rhodes University. His research interests include computer vision, machine learning, and robotics.

Sharat Saurabh Akhoury received his MASc degree in electrical and computer engineering from the University of Ottawa, Ontario, Canada, in 2013. Currently, he is a biometrics and signal processing engineer at the Council for Scientific and Industrial Research, South Africa. His research interests include image and signal processing, computer vision, machine learning, and pattern recognition.