Smoothness assessment of corneal stromal surfaces

Anca Marian, PhD, Ossama Nada, MD, FRCS(G), PhD, François Légaré, PhD, Jean Meunier, PhD, François Vidal, PhD, Sébastien Roy, PhD, Isabelle Brunette, MD, FRCSC, Santiago Costantino, PhD

PURPOSE: To assess the accuracy of the scanning electron microscopy (SEM) and present alternative approaches to quantify surface roughness based on numerical analysis.

SETTING: Department of Ophthalmology, Maisonneuve-Rosemont Hospital, University of Montreal, Montreal, Quebec, Canada.

DESIGN: Experimental study.

METHODS: Lamellar stromal cuts were performed on human corneas using a femtosecond laser or a microkeratome. The photodisrupted stromal surfaces were processed for SEM, and images were acquired at ×1000 magnification. First, images were evaluated by independent observers. Second, images were analyzed based on first-order and second-order statistics of gray-level intensities. Third, 3-dimensional (3-D) surface reconstructions were generated from pairs of SEM images acquired at 2 angles.

RESULTS: Results show that traditional assessment of roughness based on evaluating SEM images by independent observers can be replaced by computer-image texture analysis; an algorithm was developed to avoid subjective and time-consuming observations. The 3-D reconstructions allowed additional characterization of surface properties that was not possible with SEM images alone. Significant fluctuations in surface height were lost, although they could be retrieved using 3-D reconstructions.

CONCLUSIONS: Image texture analysis allowed objective and repeatable assessment of stromal surface roughness; however, full assessments of surface-height fluctuations required 3-D reconstruction. These complementary methodologies offer a more comprehensive assessment of corneal surface roughness in clinical applications.

Financial Disclosure: No author has a financial or proprietary interest in any material or method mentioned.

J Cataract Refract Surg 2012; ■:■–■ © 2012 ASCRS and ESCRS

Supplemental material available at www.jcrsjournal.org.

Whether performed with a laser or mechanically using a blunt dissector, a blade, or a microkeratome, lamellar dissection of the cornea has become routine practice for corneal surgeons worldwide. The 20 330 lamellar corneal transplants reported in 2010 by the Eye Bank Association of America[1] and the 700 000 laser in situ keratomileusis surgeries performed every year in the United States[2] illustrate well the wide acceptance of this practice. Surface smoothness appears to be one of the most important parameters in assessing the quality of a lamellar stromal cut. A smooth surface is believed to ensure a more uniform interface, with less scarring, improved light transmission, greater contrast sensitivity, less optical aberrations, and overall better vision.[3–7]

Scanning electron microscopy (SEM) images are commonly used to assess the smoothness of a corneal lamellar cut generated using a new surgical technique or new laser parameters. Most frequently, the degree of roughness of SEM images is determined subjectively by 1 or more observers,[8–11] sometimes using as a reference a series of 4 or more standard images[12,13] or assigning scores based on the dissected surface characteristics.[14–18] In this context, numerical SEM image analysis would bring an advantage in terms of accuracy, reproducibility, and speed of evaluating

0886-3350/$ - see front matter
http://dx.doi.org/10.1016/j.jcrs.2012.08.050

the quality of stromal cuts and optimizing new technologies involving corneal dissections.

An interesting first objective approach used the standard deviation (SD) of the gray levels in SEM images as a quantitative estimation of corneal-surface roughness.[8,19,20] However, a drawback of this approach is that gray levels in SEM images reflect the amount of secondary electrons detected and not the height of the surface. The amount of secondary electrons emitted in all directions by a homogeneously metal-coated SEM corneal sample largely depends on the angle subtended between the electron beam and the local sample surface and, as a first approximation, the signal is inversely proportional to the cosine of such angle.[21] Therefore, gray-level contrast is an indication of a change in height (i.e., local slope) rather than the height itself. Hence, the texture of a corneal image is correlated with the surface roughness, even if a roughness coefficient, in terms of surface height, cannot be simply calculated from single SEM images.

First-order and second-order statistics on gray levels of SEM images can provide quantitative information on surface roughness. In first-order statistics, each pixel contained in the image is regarded individually. Mean values and SDs are calculated using the gray levels associated with each pixel and are used as image descriptors. Second-order statistics go a step further; pairs of pixels with a specified relative distance and angle are compared to characterize the texture of the image. Second-order statistics basically provide information about the presence of different intensity patterns and their distribution inside the image.[22]

Submitted: January 18, 2012.
Final revision submitted: August 10, 2012.
Accepted: August 12, 2012.

From the Maisonneuve-Rosemont Hospital Research Center (Marian, Nada, Brunette, Costantino), the Department of Mathematics and Statistics (Meunier, Roy), University of Montreal, and the Department of Ophthalmology (Meunier, Brunette, Costantino), University of Montreal, Montreal, Quebec, and the Institut National de la Recherche Scientifique–Énergie, Matériaux et Télécommunications (Marian, Légaré, Vidal), Varennes, Quebec, Canada.

Supported by the Canadian Institutes of Health Research, the FRSQ Research in Vision Network, the Canadian National Institute for the Blind, and the Fond de Recherche en Ophtalmologie de l'Université de Montréal.

Christophe Chabanier and Miguel Chagnon provided technical support.

Corresponding author: Santiago Costantino, PhD, Department of Ophthalmology, Maisonneuve-Rosemont Hospital, 5415 boulevard de L'Assomption, Montreal, Quebec, H1T 2M4 Canada. E-mail: santiago.costantino@gmail.com.

The goal of this study was to improve the analysis of SEM images to obtain a quick and objective classification of corneal roughness. We first present methodologies based on numerical texture analyses of SEM images to obtain such information. This approach provides objectivity and repeatability, avoiding intraobserver and interobserver variability associated with human perception of roughness. Second, we show how 3-dimensional (3-D) reconstructions of the corneal surface from stereo SEM images can be performed to quantify surface roughness. A comprehensive study of these techniques and their correlation with traditional subjective observations is also presented.

MATERIALS AND METHODS

Corneal Sample Preparation

Human globes unsuitable for transplantation were obtained from the local eye bank (Banque d'Yeux du Québec, Montreal, Quebec, Canada) within 24 hours after death and preserved at 4°C in a humid chamber for a maximum of 48 hours. Corneas were mechanically deepithelialized, and a stromal lamellar dissection was performed with a femtosecond laser (n = 22 samples) or a microkeratome (n = 12 samples).

For the laser lamellar cuts, globes were placed in a holder with an 18.0 mm diameter round opening and pressurized at 18 mm Hg by injecting saline into the vitreous. An applanation suction ring was applied to the anterior surface of the cornea, and saline solution was pipetted to maintain humidity of the surface. A femtosecond laser (Visumax, Carl Zeiss Meditec AG) was used to perform an 8.0 mm diameter lamellar cut at an intended depth of 220 μm from the anterior surface; the cut had a superior hinge. The flap was then excised and the corneoscleral button dissected and fixed in 10% formaldehyde.

For microkeratome lamellar cuts, the corneoscleral buttons were dissected and mounted on an artificial chamber (ALTK/DSAEK CBm microkeratome system, Moria). A 300 μm footplate was used to cut a free cap. First-use and second-use blades were used to obtain a wider range of surface quality. After the cut, corneas were unmounted and fixed in 10% formaldehyde.

Scanning Electron Microscopy Images

A quarter of each corneoscleral button was processed for SEM. Tissues were rinsed, dehydrated in ascending concentrations of ethanol (25%, 50%, 75%, and 100%), infiltrated with 3 changes of hexamethyldisilazane, and dried overnight. Samples were then coated with a 20 nm layer of gold by vacuum resistive thermal evaporation.

The stromal bed of each specimen was examined using a scanning electron microscope (Jeol JSM-6300F, Jeol Ltd.) at 5 kV. Samples were placed on an xyz translation stage, which allowed tilt with respect to the beam axis. Different regions were visualized on each sample using a ×1000 magnification and 15.0 mm working distance. This magnification was chosen to match that in previous studies in which SEM images were used to judge corneal surface

quality.[8,12,20,23] All SEM images had a 100 nm/pixel scale and were saved in an 8-bit format.

Stereoscopic pairs of SEM images were then acquired for 3-D reconstruction of the corneal surfaces. Two images of the same region captured from 2 opposite angles were acquired by tilting the sample −3 degrees and +3 degrees with respect to the electron beam axis.

Surface Roughness Grading Based on Human Perception

A set of 17 SEM images representing a wide range of corneal-surface-roughness levels were presented to 60 observers who agreed to respond and would provide statistically significant data. From these answers, the ones that contained logical mistakes (eg, omission or repetitions) were discarded, yielding 52 responses for the analysis.

The 17 images were selected from a larger set, regardless of whether they were obtained from the same samples, and considering that the goal was to have a large variety in terms of roughness for testing purposes rather than classifying these particular samples. Micrographs were assembled randomly in a single picture, which allowed an ensemble view of all samples (Supplement A, available at http://jcrsjournal.org). The observers were asked to sort the images according to perceived levels of surface roughness, from the roughest to the smoothest. These observations were treated as votes. Voters assigned to images a relative roughness rank, and each image received 52 votes. For each cornea, the average position assigned by voters was calculated and a consensus ranking was computed by sorting the mean value of the voted position. Observers were blinded to the conditions under which the corneas were processed. Figure 1 shows 6 images representative of the surface roughness range observed in the montage.

Numerical Analysis of 2-Dimensional Scanning Electron Microscopy Images

The quantitative analysis of first-order (single-pixel) and second-order (pairs of pixels) statistics of pixel intensities was implemented to classify the SEM images using Matlab software (Mathworks, Inc.). First-order statistics consisted of calculating the SD of the gray-level intensites of each SEM image. For second-order statistics, the method was based on the calculation of the gray-level co-occurrence matrix (GLCM) proposed by Haralick et al.[22] The GLCM describes intensity spatial distribution of the pixels in an image and is obtained by calculating how often different combinations of pixel-pair intensities are found for a given distance (d) and direction between pixels. To build the GLCM, all possible pixel-pair intensity combinations (i, j) distant by d in the image are considered. The total number of these combinations is determined by the number of gray levels of the image, and the total number of pixel pairs separated by d (called further GLCM distance) in a set direction are counted for each (i, j). Each particular direction and distance d determines a unique GLCM, which is typically normalized to express probability.

Among the 14 parameters initially proposed by Haralick et al.,[22] contrast was chosen and analyzed as a descriptor of roughness in this study as follows:

$$\text{Contrast} = \sum_{i,j}^{N-1} (i-j)^2 \, p(i,j)$$

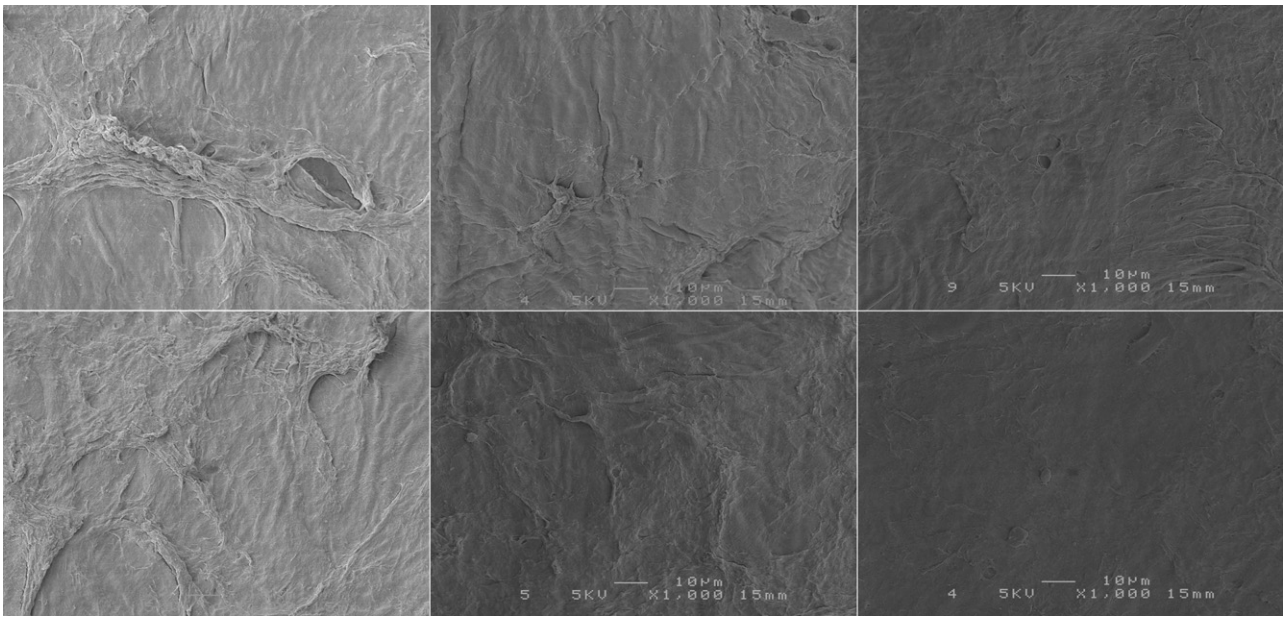

Figure 1. Corneal dissection images. Six of the 17 SEM images provided to the 52 independent observers for assessing corneal surface roughness. These representative images illustrate the range of surface roughness present in the montage of 17. Shown are 2 rough samples (*left column*), 2 intermediate samples (*middle column*), and 2 smooth samples (*right column*). The differences in image background are due to differences in metallic coatings and variation in the electron-gun intensity between batches, despite constant contrast and brightness settings of the microscope. Nevertheless, these background variations did not affect the image statistics used in this study or the 3-D reconstruction.

where $p(i, j)$ is the GLCM component for intensity pair (i, j). Images were converted to 128 intensity levels and for each distance, ranging from 1 to 800 pixels (0.1 to 80 µm), the contrast values obtained for 0 degrees, 45 degrees, 90 degrees, and 135 degrees were averaged.

First, the SD of the gray levels for each of the 17 SEM images as well as the Haralick texture contrast were calculated. The GLCMs were computed for each image, and the contrast was plotted as a function of the GLCM distance. Second, whether human perception of roughness was correlated with first-order and second-order statistics was tested.

Surface Roughness Grading Based on 3-Dimensional Reconstructions

Three-dimensional reconstructions were performed for the 17 samples. A custom algorithm developed using Mathematica (Wolfram Research, Inc.) was used to calculate the 3-D height map of each surface using stereo pairs of SEM images.[24] This was an extension of a previous work[25] in which the strict constraints on sample positioning (manual alignment of the eucentric points, tilt axis, and precise knowledge of tilt angles and working distance) were alleviated by automatically finding an average depth plane through the sample.

For that purpose, a set of similar points was automatically identified in the stereo images and the coordinates of these matching points were extracted. These correspondences yielded a linear transformation **T**, which modeled geometrically the tilt movement of the microscope stage. By registering the stereo images with the transformation **T**, a reference plane for the subsequent measurement of heights was obtained.

When comparing the registered stereo images, because the original sample was not planar, any depth deviation (from the reference plane) showed up as a residual displacement (or disparity) of any corresponding pair of pixels. To ease computation, the images were also rotated to make these displacements horizontal to establish a dense stereo matching step with an appropriate algorithm.[26] The resulting horizontal disparity map (i.e., the differences between the horizontal coordinates of the matching points in the stereo images) was finally converted into heights according to the acquisition parameters of tilt angle, magnification, and working distance. Missing or hidden pixels were estimated by interpolation. A full elevation map was computed for all samples. A calibration test was performed using as reference sample a holographic replica grating (13.33 µm step and 4.63 µm height), and the reconstructed surface was compared with manufacturer specifications. This calibration also provided precise information about the size of the pixel in the SEM images.

Statistical Analysis

The Pearson correlation coefficient (r) was calculated to study the association between roughness values obtained with the 3 objective methods. The Spearman rank correlation coefficient (ρ) was used to study the association between the ranking obtained with the different objective methods and subjective methods. A P value less than 0.05 was considered statistically significant. The analyses were performed using SPSS software (version 19, SPSS, Inc.) and Matlab software.

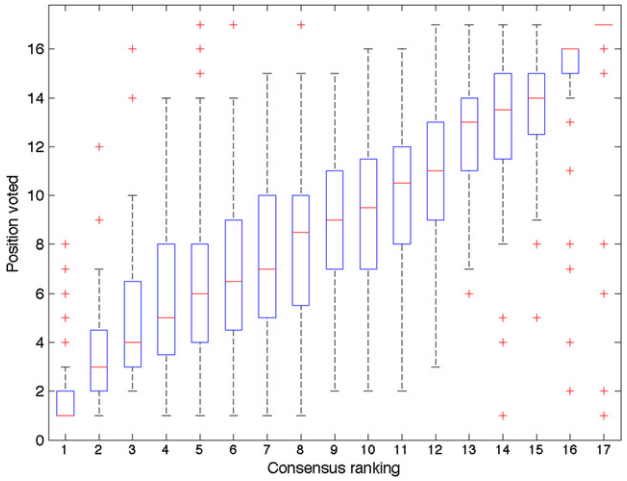

Figure 2. Human perception of surface roughness of corneal surfaces. All SEM images were obtained with identical microscope settings. Dispersion of the subjective ranking. Each of the 52 observers ranked all 17 images, from the roughest to the smoothest. The central mark of the boxes identifies the median. The length of the boxes is the interquartile range (IQR). Values more than 3 IQRs from the end of a box are labeled as outliers (+). The consensus ranking represents the mean of all 52 votes positions by the observers to each of the 17 images sorted in ascending order. Very rough images (*bottom of the ranking*) and very smooth images (*top of the ranking*) showed a low dispersion of votes, while images with intermediate roughness showed significant variations.

RESULTS

Human Perception of Roughness

The boxplot in Figure 2 shows the distribution of the observers' votes. The size of the boxes represents the interquartile range and reflects the dispersion of votes, which is lower at the extremes of the ranking and much larger in the intermediate positions, indicating better agreement between observers at the extremes of the ranking.

Numerical Analysis of Scanning Electron Microscopy Images

Table 1 shows the results of the SD and texture contrast calculations. A high value of these parameters is generally associated with a high roughness of the sample.

These first-order and second-order statistics parameters were highly correlated ($r = 0.971$, $P < .0001$) (Figure 3, *A*), and this high correlation held for a wide range of GLCM distances (data not shown).

Strong correlations were observed between the subjective consensus ranking and the gray-level intensities SD ($\rho = -0.753$, $P = .0005$) and texture contrast ($\rho = -0.775$, $P < .0005$) (Figure 3, *B*).

Table 1. All parameters calculated for the 17 samples presented to the observers. The position of a sample in the montage image is numbered as a matrix (row_column), as shown in Supplement A (available at: http://jcrsjournal.org).

Position	Mean Vote	3-D Roughness Coefficient (AU)	Texture Contrast	Image Intensity Standard Deviation
1_1	1.83	8942	228.03	21.26
1_3	3.60	3476	227.64	21.31
1_4	4.96	8131	100.33	14.47
1_2	5.90	9446	189.51	19.75
2_2	6.17	9184	109.72	16.18
4_1	7.12	5861	141.58	16.77
1_5	7.67	12 359	75.80	12.61
4_2	8.33	8039	98.90	13.94
3_5	8.60	9833	162.53	19.03
2_3	9.27	10 063	71.39	12.75
2_1	10.21	4514	78.08	13.69
3_1	11.00	4052	46.73	10.45
3_4	12.21	13 047	106.65	16.04
3_3	12.69	3155	54.53	12.23
3_2	13.38	2406	123.19	16.46
2_4	14.54	1565	41.99	9.58

AU = arbitrary units; 3-D = 3 dimensional

The roughness ranking depended weakly on the GLCM distance, except for the lowest and highest values considered (Figure 3, *C*). Figure 3, *D*, shows a histogram of the changes in the roughness ranking as a function of GLCM distance; that is, the number of order variations among the 17 samples. Images remained stably sorted for GLCM distances between 10 pixels and 700 pixels. Consistently, Figure 3, *E*, shows that the Pearson correlation coefficient remained relatively constant (better than 0.75) for GLCM distances beyond 100 pixels up to approximately 700 pixels.

Surface Roughness Based on 3-Dimension Reconstructions

An SEM image (Figure 4, *A*) and its corresponding elevation color map overlaid on the original SEM image (Figure 4, *B*) are shown as an example of the 3-D reconstructions of the 17 samples. A 3-D rendering of the surface is shown in Figure 4, *C*, from which it becomes evident that this corneal sample presented large-scale height variations that could not be appreciated using a single SEM image. To quantify sample roughness, the SD of the height was extracted from the elevation map and used as a roughness coefficient; however, a lack of correlation between the roughness coefficient and the textural analysis was observed. The correlation between this roughness coefficient and the mean votes was not statistically significant ($\rho = -0.436$; $P = .079$). It was not correlated with the gray-level SD ($r = 0.336$, $P = .187$) or with texture contrast ($r = 0.250$, $P = .332$).

DISCUSSION

To our knowledge, we present the first systematic assessment of the tools available to measure the quality of a corneal surface using SEM samples. We discuss subjective and mathematic classifications of roughness based on SEM images and compare them with roughness coefficients obtained from 3-D reconstructions of such surfaces.

Texture of an image is a relatively complex concept. In the case of interest here, texture is related to the gray level properties of pixels and their spatial fluctuations. When these fluctuations show an order, texture can be used for classification. It has been shown that human perception can easily discriminate textures that differ in their second statistics, while more cognitive effort is required to distinguish textures when second-order statistics are similar.[27]

A wide range of approaches has been used to characterize texture during the past 4 decades.[22,28,29] The method developed by Haralick et al. in 1973,[22] based on the computation of GLCM, has become to a large degree the standard approach to texture analysis and is used to study the statistics of pixel intensity distribution. Pairs of pixels are considered because single-pixel statistics do not provide enough information on textures for most practical applications. (Single-pixel statistics cannot distinguish between checkerboards with different number of squares, while second-order statistics can.) The GLCM computes the frequency at which 2 distant pixels have intensities i, j (see Materials and Methods). This way of storing gray-level changes is the basis of a variety of parameters designed to

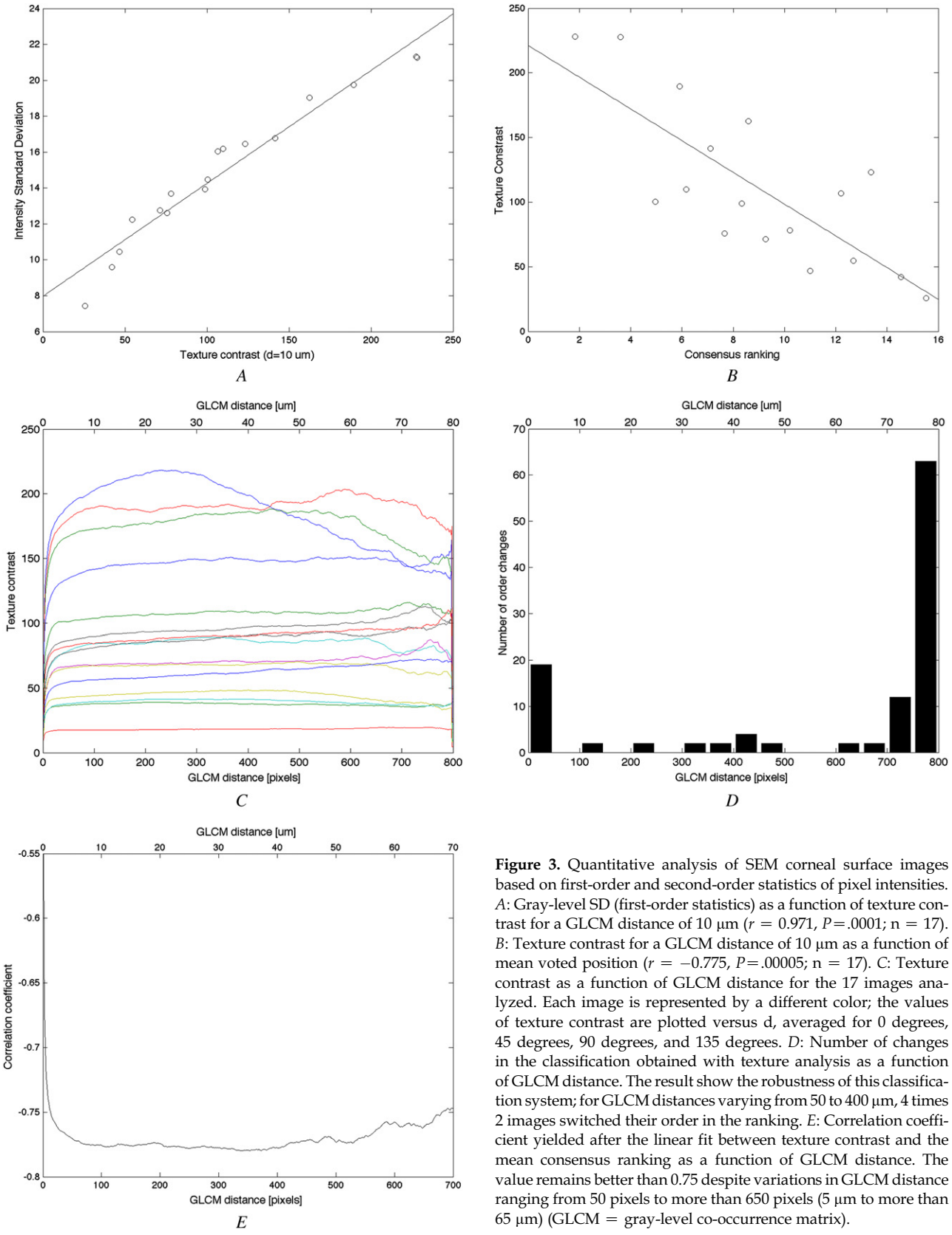

Figure 3. Quantitative analysis of SEM corneal surface images based on first-order and second-order statistics of pixel intensities. *A*: Gray-level SD (first-order statistics) as a function of texture contrast for a GLCM distance of 10 μm ($r = 0.971$, $P = .0001$; n = 17). *B*: Texture contrast for a GLCM distance of 10 μm as a function of mean voted position ($r = -0.775$, $P = .00005$; n = 17). *C*: Texture contrast as a function of GLCM distance for the 17 images analyzed. Each image is represented by a different color; the values of texture contrast are plotted versus d, averaged for 0 degrees, 45 degrees, 90 degrees, and 135 degrees. *D*: Number of changes in the classification obtained with texture analysis as a function of GLCM distance. The result show the robustness of this classification system; for GLCM distances varying from 50 to 400 μm, 4 times 2 images switched their order in the ranking. *E*: Correlation coefficient yielded after the linear fit between texture contrast and the mean consensus ranking as a function of GLCM distance. The value remains better than 0.75 despite variations in GLCM distance ranging from 50 pixels to more than 650 pixels (5 μm to more than 65 μm) (GLCM = gray-level co-occurrence matrix).

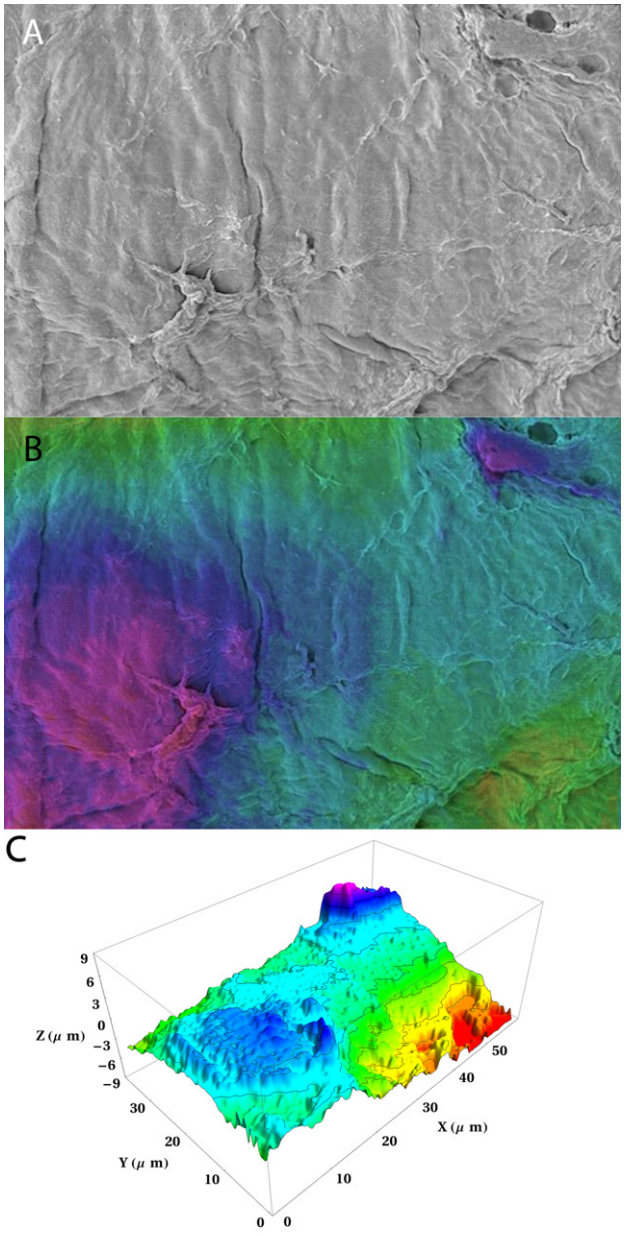

Figure 4. Representative example of SEM image 3-D reconstruction. *A*: One of the 2 stereo images required for the reconstruction. *B*: Color map obtained after the 3-D reconstruction overlying the original 2-D image. *C*: The 3-D rendering of the surface obtained after reconstruction.

describe texture and develop image segmentation algorithms. Haralick et al.[22] introduced 14 parameters based on the co-occurrence matrix used to classify textures in terms of uniformity, homogeneity, contrast, gray-level dependencies, and complexity of the image. Parameters such as energy, entropy, correlation, homogeneity, contrast, and variance have been used in biomedical sciences to show the signatures of different physiological and pathological states,[30,31] to classify protein,[32] to characterize aging in microorganisms,[33] and to perform automatic image classification based on neural networks for several other specific applications.[34,35] Texture contrast used as a parameter in the present study is considered one of the most relevant[36] of Haralick et al.'s parameters[22] and it is also one of the most commonly used.

As opposed to second-order statistics on pixel intensity, first-order statistics do not take into account the spatial distribution of intensities. Nevertheless, the SD of the gray-level intensities of our SEM corneal images still showed a good correlation with the mean voted position, although the correlation was somewhat lower than that for texture contrast. Texture contrast is a measure of intensity fluctuations over a certain GLCM distance; a rough surface shows a higher texture contrast than a smooth surface. Furthermore, the texture contrast seems to reproduce better human perception when the GLCM distance is greater than a few pixels, because at small scale, noise dominates spatial fluctuations. The gray-level variations over a scale of more than a few micrometers are expected to better reflect the clinically relevant characteristics of the tissue related to the dissection method, such as collagen lamellae disruption and tissue bridges.

As seen in Figure 2, *C*, the 17 curves representing the texture contrast of our images were well separated, which means that texture analysis distinguished roughness in ranges in which a group of independent observers could not. One can see that all the curves were relatively constant provided that the GLCM distance was greater than a few micrometers. For smaller distances, the curves intersect due to noise in the images, an effect that disappear as the distance increases.

The high correlation between the texture analyses and human observation indicates that a computer-based calculation of gray-scale SD or texture contrast can replace the current subjective evaluation of roughness (as performed in the field of corneal lamellar surgery assessment). As noted above, very smooth and very rough surfaces were classified as such by a majority of the observers; however, substantial variations were obtained for samples with intermediate roughness. This study confirms the difficulty of precisely classifying images of similar roughness and illustrates the need for an objective method able to discriminate between comparable levels of roughness. Numerical analysis simplifies and accelerates classification of the SEM images; it provides objectivity and repeatability without intraobserver or interobserver variability.

Alternative approaches for measuring the surface height still do not perform well enough to become standards for stromal-bed-roughness quantification. Atomic force microscopy (AFM) has been used to

probe elevations in corneal surfaces with nanometer accuracy.[37–40] This technology does not require sample preparation; however, the tradeoff for high-accuracy measurements of elevation is the lower field of view obtained using this imaging modality. Areas of approximately 100 $\mu m^2 \times 100\ \mu m^2$ can be scanned, provided the elevations on the surface do not surpass 5 μm. These parameters make AFM a useful tool to probe local roughness with very high accuracy in relatively flat regions; however, AFM remains inadequate for global estimations of the quality of stromal surfaces generally featuring a wide range of elevations. Conversely, optical coherence tomography (OCT) can image large corneal surfaces, but with insufficient resolution. Research-laboratory OCT configurations can determine the height of a surface with submicrometric resolution in the axial direction[41,42]; however, the resolution in the perpendicular direction is usually limited to several micrometers.[43,44] Fine details that are important for assessing surface quality, such as stromal fibrils and tissue bridges, which are typically smaller than 10 μm, cannot be resolved with this technique.

When an SEM image is acquired, the atoms of the sample surface interact with the high-energy electrons of a scanning beam. Low-energy secondary electrons resulting from inelastic scattering are collected by a detector, and the signal is displayed as a 2-dimensional (2-D) intensity distribution map. Usually, the electron beam enters the sample perpendicular to the sample holder surface and secondary electrons are collected by a detector oriented at a certain angle.

Gray levels in SEM images are determined by the amount of secondary electrons collected by the detector. The image contrast depends on a series of variables that are not exclusively related to the topography of the surface. Although the narrow electron beam allows a very large depth of focus, variations in the angle subtended by the sample and the beam induce changes in the number of secondary electrons collected. Because of this angular dependence of the signal, small structures and cracks become evident while small gradients in the height of the sample are barely visible.

The roughness coefficient in the 3-D images was calculated as the SD of the height extracted from the elevation map; however, poor or no correlations were found when these coefficients were compared with texture analysis and human observation. This shows that the real topography of the surfaces was not the same as that perceived from standard SEM images.

The 3-D reconstructions did not reproduce the fine details visible in SEM images. These details were too small compared to the resolution of the reconstructions or, in most cases, did not correspond to significant changes in height. The stereo reconstruction algorithms require that when observed from 2 different angles, the features in the sample become displaced and the magnitude of this displacement is used to assign a surface elevation. Hence, a large fraction of the details that can be visually noted in 2-D SEM images correspond to changes in height too shallow to be detected by the 3-D reconstruction at this magnification.

There are 2 main types of errors in the reconstruction of 3-D height maps. One is local and induced by point matching errors during the dense stereo-matching step. These errors show up as noise added to the height maps that could be reduced by low-pass filtering or by tuning the stereo-matching parameters (eg, cost function of the Hirschmuller algorithm[26]). The second type is global and induced by errors in tilting of the specimen. The deviation between the actual tilting and the intended tilting results in a scaling of the height map that is proportional to the angular difference; thus, the 3-D relief is fully preserved. Any other misplacement of the specimen (eg, alignment of the eucentric points) does not affect the height map because transformation **T** will account for them.

On the other hand, standard SEM images did not discriminate the wide change in height documented by the 3-D reconstructions. This is explained by the fact that the long working distance of SEM is such that large changes in surface elevation (tens of micrometers) still appear in focus. Although these large fluctuations in elevation are strikingly revealed in 3-D reconstructions, in the 2-D images they appear to be masked by small fractures and gaps. Thus, the lack of correlation between roughness coefficients computed from 3-D reconstructions with alternative methods is not surprising. These coefficients are dominated by large changes in elevation that cannot be detected in 2-D SEM as major gray-level changes.

Stereo-imaging and 3-D reconstruction procedures are more time consuming than standard 2-D SEM, both at the experimental and processing levels. Also, open-source robust algorithms for 3-D reconstructions are not user friendly for standard vision science research groups. Our software is open source and available on request.

Elevations on the order of tens of micrometers may arise from the shrinkage and other modifications that occur during the sample processing for SEM and might complicate the assessment of the quality obtained by the laser or the microkeratome. This disadvantage of using SEM images can sometimes be overlooked, and other techniques such as environmental SEM or AFM, in which no preimaging sample processing is required, could be envisaged for more precise quantitative description of the surface quality.

That gray levels in 2-D images correspond to the slope rather than the height of the sample surface has an impact on the interpretation of SEM images, and our study sheds light on this issue. The simple observation of corneal-surface SEM images does not correlate with roughness following a classic definition; that is, in terms of the SD of the surface elevation. Nevertheless, corneal surgeons typically assess the spatial distribution of these irregularities on the surface, which we show can be quantified by simple texture analysis algorithms and have clinically been shown to correlate with vision outcome.

In summary, we have shown that an automated computation can provide the results that can be obtained by asking independent observers to assess the roughness of corneal surfaces. This computer-based calculation is objective, quantitative, much faster, and robust. In addition, we found that some 3-D information is lost when single standard SEM images are considered and that 3-D reconstructions provide complementary surface information. Large fluctuations in surface height that could be discovered and quantified using 3-D reconstructions can be of much higher amplitude than the ones that can be visually perceived in standard 2-D images and sometimes dominate the calculation of the roughness coefficient.

WHAT WAS KNOWN

- The standard method to evaluate the quality of lamellar dissections is to grade SEM images in term of roughness by independent observers. There is a lack of quantitative methods based on the SEM image analysis allowing an objective roughness assessment.
- Gray levels in SEM images do not reflect the height of the surface; thus, color fluctuations cannot be correctly interpreted by visual inspection only.

WHAT THIS PAPER ADDS

- A new quantitative method to assess stromal surface roughness based on the analysis of standard SEM images was developed. The method is fast, automatic, and unbiased, and the results are in agreement with the estimations obtained by visual inspection.
- For the first time, we performed 3-D reconstructions using pairs of SEM images of such surfaces to obtain height maps that provide complementary information. We found that major surface fluctuations are not shown by the standard 2-D analysis and that both approaches should be combined for a thorough characterization.

REFERENCES

1. Eye Bank Association of America. 2009 Eye Banking Statistical Report. Washington DC, Eye Bank Association of America, 2010; 1–17
2. American Academy of Ophthalmology. Eye Health Statistics at a Glance. Am Academy Ophthalmol 2009 Available at: http://www.aao.org/newsroom/press_kit/upload/Eye-Health-Statistics-June-2009.pdf. Accessed August 20, 2012
3. Güell JL, Velasco F, Roberts C, Sisquella MT, Mahmoud A. Corneal flap thickness and topography changes induced by flap creation during laser in situ keratomileusis. J Cataract Refract Surg 2005; 31:115–119
4. Huang D, Arif M. Spot size and quality of scanning laser correction of higher-order wavefront aberrations. J Cataract Refract Surg 2002; 28:407–416
5. Lim T, Yang S, Kim MJ, Tchah H. Comparison of the IntraLase femtosecond laser and mechanical microkeratome for laser in situ keratomileusis. Am J Ophthalmol 2006; 141:833–839
6. Tran DB, Sarayba MA, Bor Z, Garufis C, Duh Y-J, Soltes CR, Juhasz T, Kurtz RM. Randomized prospective clinical study comparing induced aberrations with IntraLase and Hansatome flap creation in fellow eyes; potential impact on wavefront-guided laser in situ keratomileusis. J Cataract Refract Surg 2005; 31:97–105
7. Vinciguerra P, Azzolini M, Airaghi P, Radice P, De Molfetta V. Effect of decreasing surface and interface irregularities after photorefractive keratectomy and laser in situ keratomileusis on optical and functional outcomes. J Refract Surg 1998; 14: S199–S203
8. Sarayba MA, Ignacio TS, Binder PS, Tran DB. Comparative study of stromal bed quality by using mechanical, IntraLase femtosecond laser 15- and 30-kHz microkeratomes. Cornea 2007; 26:446–451
9. Sarayba MA, Maguen E, Salz J, Rabinowitz Y, Ignacio TS. Femtosecond laser keratome creation of partial thickness donor corneal buttons for lamellar keratoplasty. J Refract Surg 2007; 23:58–65
10. Nuzzo V, Aptel F, Savoldelli M, Plamann K, Peyrot D, Deloison F, Donate D, Legeais J-M. Histologic and ultrastructural characterization of corneal femtosecond laser trephination. Cornea 2009; 28:908–913
11. Jones YJ, Goins KM, Sutphin JE, Mullins R, Skeie JM. Comparison of the femtosecond laser (IntraLase) versus manual microkeratome (Moria ALTK) in dissection of the donor in endothelial keratoplasty; initial study in eye bank eyes. Cornea 2008; 27: 88–93
12. Cheng YYY, Kang SJ, Grossniklaus HE, Pels E, Duimel HJQ, Frederik PM, Hendrikse F, Nuijts RMMA. Histologic evaluation of human posterior lamellar discs for femtosecond laser Descemet's stripping endothelial keratoplasty. Cornea 2009; 28:73–79
13. Sanka RK, Loft ES, Randleman JB. Effect of varying microkeratome parameters on laser in situ keratomileusis interface surfaces. J Cataract Refract Surg 2010; 36:493–496
14. Wilhelm FW, Giessmann T, Hanschke R, Duncker GI, Wilhelm LH. Cut edges and surface characteristics produced by different microkeratomes. J Refract Surg 2000; 16:690–700
15. Kunert KS, Blum M, Duncker GI, Sietmann R, Heichel J. Surface quality of human corneal lenticules after femtosecond laser surgery for myopia comparing different laser parameters. Graefes Arch Clin Exp Ophthalmol 2011; 249:1417–1424
16. Heichel J, Blum M, Duncker GIW, Sietmann R, Kunert KS. Surface quality of porcine corneal lenticules after femtosecond lenticule extraction. Ophthalmic Res 2011; 46:107–112

17. Mehta JS, Parthasarthy A, Por Y-M, Cajucom-Uy H, Beuerman RW, Tan D. Femtosecond laser-assisted endothelial keratoplasty; a laboratory model. Cornea 2008; 27:706–712
18. Hammer T, Hanschke R, Wörner I, Wilhelm FW. Evaluation of four microkeratome models: quality and reproducibility of cut edge and cut surface as determined by scanning electron microscopy. J Refract Surg 2005; 21:454–462
19. Srinivasan S, Holmyard D. Corneal stromal bed quality obtained with mechanical microkeratome and femtosecond laser [letter]. Cornea 2008; 27:746; reply by T Ignacio, 746–747
20. Sarayba MA, Ignacio TS, Tran DB, Binder PS. A 60 kHz IntraLase femtosecond laser creates a smoother LASIK stromal bed surface compared to a Zyoptix XP mechanical microkeratome in human donor eyes. J Refract Surg 2007; 23: 331–337
21. Laudate T. Imaging guidelines for scanning electron microscopy. Advanced Materials & Processes 2003; 161(7):23–25. Available at: http://www.google.com/url?sa=t&rct=j&q=imaging%20guidelines%20for%20scanning%20electron%20microscopy&source=web&cd=1&ved=0CGEQFjAA&url=http%3A%2F%2Fwww.jeolusa.com%2FDesktopModules%2FBring2mind%2FDMX%2FDownload.aspx%3FEntryId%3D614%26PortalId%3D2%26DownloadMethod%3Dattachment&ei=agQyUO32EIa69gS6iICwBg&usg=AFQjCNEhQkxcOWe986DJ01ltBoOctMm9uw. Accessed August 20, 2012
22. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Systems Man Cybernetics 1973; SMC-3:610–621. Available at: http://dceanalysis.bigr.nl/Haralick73-Textural%20features%20for%20image%20classification.pdf. Accessed August 20, 2012
23. Behrens A, Seitz B, Langenbucher A, Kus MM, Rummelt C, Küchle M. Evaluation of corneal flap dimensions and cut quality using a manually guided microkeratome. J Refract Surg 1999; 15:118–123; errata, 400
24. Roy S, Meunier J, Marian AM, Nada O, Légaré F, Vidal F, Brunette I, Costantino S. Automatic 3D reconstruction of quasi-planar stereo SEM images. Montreal, Quebec, Canada, Université de Montréal, Département d'Informatique et de Recherche Opérationnelle, 2011; Technical Report number 1361
25. Ponz E, Ladaga JL, Bonetto RD. Measuring surface topography with scanning electron microscopy. I. EZEImage: a program to obtain 3D surface data. Microsc Microanal 2006; 12:170–177
26. Hirschmüller H. Accurate and efficient stereo processing by semi-global matching and mutual information. IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2005; 807–814. Available at: http://www.dlr.de/rm/en/PortalData/3/Resources/papers/modeler/cvpr05hh.pdf. Accessed August 20, 2012
27. Julesz B. A theory of preattentive texture discrimination based on first-order statistics of textons. Biol Cybern 1981; 41:131–138
28. Weszka JS, Dyer CR, Rosenfeld A. Comparative-study of texture measures for terrain classification. IEEE Trans Syst Man Cybern 1976; SMC-6:269–285
29. Troy EB, Deutsch ES, Rosenfel A. Gray-level manipulation experiments for texture analysis. IEEE Trans Syst Man Cybern 1973; SMC-3:91–98
30. Nielsen B, Albregtsen F, Danielsen HE. Statistical nuclear texture analysis in cancer research: a review of methods and applications. Crit Rev Oncog 2008; 14:89–164
31. Kassner A, Thornhill RE. Texture analysis: a review of neurologic MR imaging applications. Am J Neuroradiol 2010; 31:809–816
32. Nanni L, Shi J-Y, Brahnam S, Lumini A. Protein classification using texture descriptors extracted from the protein backbone image. J Theor Biol 2010; 264:1024–1032
33. Shamir L, Wolkow CA, Goldberg IG. Quantitative measurement of aging using image texture entropy. Bioinformatics 2009; 25:3060–3063. Available at: http://bioinformatics.oxfordjournals.org/content/25/23/3060.full.pdf. Accessed August 20, 2012
34. Chebira A, Barbotin Y, Jackson C, Merryman T, Srinivasa G, Murphy RF, Kovačević J. A multiresolution approach to automated classification of protein subcellular location images. BMC Bioinformatics 2007; 8:210. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1933440/pdf/1471-2105-8-210.pdf. Accessed August 20, 2012
35. Hamilton NA, Pantelic RS, Hanson K, Teasdale RD. Fast automated cell phenotype image classification. BMC Bioinformatics 2007; 8:110. Available at: http://www.biomedcentral.com/content/pdf/1471-2105-8-110.pdf. Accessed August 20, 2012
36. Baraldi A, Parmiggiani F. An investigation of the textural characteristics associated with gray-level cooccurrence matrix statistical parameters. IEEE Trans Geosci Remote Sens 1995; 33: 293–304
37. Nógrádi A, Hopp B, Révész K, Szabó G, Bor Z, Kolozsvari L. Atomic force microscopic study of the human cornea following excimer laser keratectomy. Exp Eye Res 2000; 70:363–368
38. Lydataki S, Lesniewska E, Tsilimbaris MK, Panagopoulou S, Le Grimellec C, Pallikaris IG. Excimer laser ablated cornea observed by atomic force microscopy. Single Mol 2002; 3:141–147
39. Lombardo M, De Santo MP, Lombardo G, Barberi R, Serrao S. Atomic force microscopy analysis of normal and photoablated porcine corneas. J Biomech 2006; 39:2719–2724. Available at: http://www.visioeng.it/documents/AFMJBiomechanics.pdf. Accessed August 20, 2012
40. Lombardo M, De Santo MP, Lombardo G, Lomoriello DS, Desiderio G, Ducoli P, Barberi R, Serrao S. Surface quality of femtosecond dissected posterior human corneal stroma investigated with atomic force microscopy. Cornea 2012 May 29 [Epub ahead of print]
41. Costantino S, Martínez OE, Torga JR. Wide band interferometry for thickness measurement. Opt Express 2003; 11:952–957. Available at: http://www.lec.df.uba.ar/administrator/components/com_jresearch/files/publications/costantino_2003.pdf. Accessed August 20, 2012
42. Singh K, Dion C, Costantino S, Wajszilber M, Lesk MR, Ozaki T. Development of a novel instrument to measure the pulsatile movement of ocular tissues. Exp Eye Res 2010; 91:63–68
43. Wagenblast PC, Ko TH, Fujimoto JG, Kaertner FX, Morgner U. Ultrahigh-resolution optical coherence tomography with a diode-pumped broadband Cr^{3+}:LiCAF laser. Opt Express 2004; 12:3257–3263. Available at: http://www.opticsinfobase.org/oe/abstract.cfm?uri=oe-12-14-3257. Accessed August 20, 2012
44. Drexler W, Kärtner FX, Pitris C, Boppart SA, Li XD, Ippen EP, Fujimoto JG. In vivo ultrahigh-resolution optical coherence tomography. Opt Lett 1999; 24:1221–1223. Available at: http://www.uta.edu/rfmems/N_041101/Reference/85.pdf. Accessed August 20, 2012