Supplemental Document

Biomedical Optics EXPRESS

Probing variations of fibrous structures during the development of breast ductal carcinoma tissues via Mueller matrix imaging: supplement

YANG DONG,[1,6] SHAOXIONG LIU,[2,6,7] YUANXING SHEN,[3,4] HONGHUI HE,[3,*] AND HUI MA[1,3,5,8]

[1]*Center for Precision Medicine and Healthcare, Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen 518071, China*
[2]*Shenzhen Sixth People's Hospital (Nanshan Hospital) Huazhong University of Science and Technology Union Shenzhen Hospital, Shenzhen 518052, China*
[3]*Guangdong Research Center of Polarization Imaging and Measurement Engineering Technology, Shenzhen Key Laboratory for Minimal Invasive Medical Technologies, Institute of Optical Imaging and Sensing, Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*
[4]*Department of Biomedical Engineering, Tsinghua University, Beijing 100084, China*
[5]*Department of Physics, Tsinghua University, Beijing 100084, China*
[6]*These authors contributed equally to this work*
[7]*liusx20088@sina.com*
[8]*mahui@tsinghua.edu.cn*
[*]*he.honghui@sz.tsinghua.edu.cn*

Supplement DOI: https://doi.org/10.6084/m9.figshare.12760481

Parent Article DOI: https://doi.org/10.1364/BOE.397441

Probing variations of fibrous structures during the development of breast ductal carcinoma tissues via Mueller matrix imaging: supplemental document

1. Six statistical characteristic parameters of polarimetric image blocks

Frequency distribution tells how frequencies are distributed over values, and its mathematical definitions and formulas can be found in [1]. Experimental results of various tissue samples have shown that, the FDH, a semi-quantitative analysis tool, can display in a much clearer graphics form the dominant microstructural features of complex biological tissues than the 2D images of Mueller matrix elements or derived parameters. By analyzing the peak positions, widths and shapes of the FDHs, abundant microstructural information and optical properties can be decoded from Mueller matrix. For example, the anisotropic degree of samples, the origin of anisotropy, the depolarization power of samples, and the orientation direction of anisotropic structures. Moreover, the central moment analysis of FDHs can provide us a group of orientation insensitive parameters representing the dominant microstructural features of tissues quantitatively [2].

In this work, the Mueller matrix element $|m43|$, MMT parameter t and MMPD parameter δ images of 990×1260 pixels are all divided into 55×60 pixels patches in sequence. Of note, it has been proved that an image block of 18×21 pixels can be used as a guiding standard for general problems in the application of face recognition considering the tradeoff between accurate description ability and feature complexity [3]. Inspired by the method of dividing the image into blocks with the same size in Multi block local binary patterns (MB-LBP), the traditional image block size of 18×21 pixels can be expanded by approximate 1, 2, 3 or more times [4]. Appropriate image block size has to be used according to practical requirements in different applications. In this study, considering that polarimetric images of breast ductal carcinoma tissue pays more attention on the robustness and the coarse granularity of image information, we tried to respectively divide the 2D parameters images of 990×1260 pixels into image blocks of 18×21, 33×42, 55×60 and 66×84 pixels. The experimental results have shown that the image block size of 55×60 pixels has the best performance in discriminating different pathological stages of breast ductal carcinoma considering the tradeoff between accurate description ability and feature complexity. The proposed method works among breast ductal pathological tissue samples at different stages under different experiment parameters. When the experiment conditions changed, the size settings of polarimetric images need to be adjusted accordingly. For example, when the Mueller matrix images were obtained under a 40× objective instead of 4×, the size of each polarimetric image block should be enlarged to 550×600 pixels, in order to retain the same physical dimension. Meanwhile, to cover the same imaging region as that of under 4×, one must also increase the size of 2D Mueller matrix parameters images of breast tissue samples 100 times for comprehensive and accurate pathological analysis.

Fig. S1 (a) shows that the 2D image of parameter $|m43|$ is divided into 378 image blocks. In FDHs of each image block, as shown in Fig. S1 (b), the horizontal axis represents the pixel values of the corresponding Mueller matrix parameter, and the vertical axis represents the distributing probability. The areas under the curves are normalized to 1. Here, the values of retardance-related Mueller matrix parameters of breast ductal carcinoma pathological samples are distributed between 0 and 1, and this range is divided into 400 parts, denoted as L in Eq. (S1). Therefore, for FDH of each 2D image block of Mueller matrix parameter, the horizontal axis ranges from 0 to 1, in which every 0.0025 is an interval of value, recorded as z_i. And the vertical axis is the proportion of pixels with the values in z_i range to the total number of

pixels, denoted as $p(z_i)$. Then six statistical characteristic parameters of the image block can be exactly calculated according to Eq. (S1) [5].

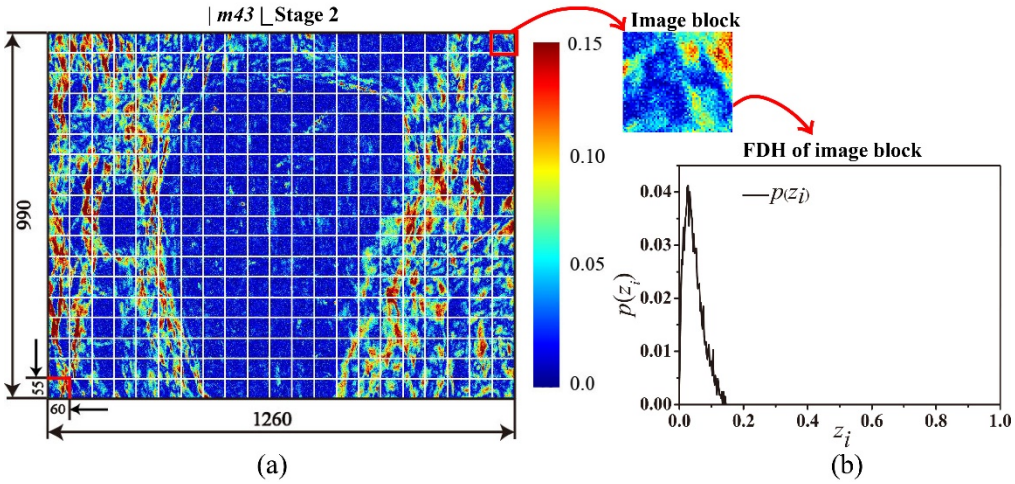

Fig. S1. (a) Pseudo-color image of Mueller matrix parameter |*m43*| of breast ductal carcinoma *in situ*, which is divided into 378 image blocks in sequence. Different image blocks are marked by white squares; (b) FDHs of the image block marked by red square.

$$
\begin{aligned}
m &= \sum_{i=0}^{L-1} z_i p(z_i) \\
std &= \sqrt{\sum_{i=0}^{L-1} (z_i - m)^2 p(z_i)} \\
R &= \frac{1}{\left(1 + std^2\right)} \\
u3 &= \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \\
U &= \sum_{i=0}^{L-1} p^2(z_i) \\
e &= -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)
\end{aligned}
\tag{S1}
$$

Here, the expected value m is the mean value, and std is the standard deviation of all pixel values in this image. R is supposed to be a measurement of relative smoothness of image texture brightness. Smoothness value ranges from 0 to 1, and for regions with identical pixel values, the value of R is 1. Similarly, U represents the consistency of the image. When the pixel values of the image are the same, U has the maximum value of 1. The fourth statistical characteristic parameter $u3$ is called skewness, which can be positive or negative. A positive (or negative) skewness value represents that the tail on the right side (or the left side) of the FDHs is longer or fatter than the left side (or the right side). The last statistical characteristic parameter e, known as entropy, has the potential to measure randomness of the image. The larger value of entropy means greater uncertainty and randomness of the image.

2. Six classification characteristic values obtained from normal breast ductal tissue samples

We calculate the mean value of 378 components in each statistical feature vector for Mueller matrix parameters of the normal samples in stage 1, and the total mean values of six statistical

feature vectors–*f_m*, *f_std*, *f_R*, *f_u3*, *f_U* and *f_e*–for 10 normal samples are used as a criterion denoted as classification characteristic values, to distinguish polarimetric images between normal and carcinoma samples. The six classification characteristic values can be acquired by Eq. (S2), where $m(i)$ is the i-th (i = 1,2,…,378) component in statistical feature vector $\overline{m}$, i.e. mean value of i-th image block of the polarimetric image, and $f\_m(j)$ is the mean value of 378 components in statistical feature vector $\overline{m}$ for one of Mueller matrix parameters of normal sample j (j = 1,2,…,10) in stage 1. Therefore, *f_m* is the total mean value for 10 normal samples used as a classification characteristic value to classify 378 components in statistical feature vector $\overline{m}$ for breast ductal carcinoma samples at different stages. The other five classification characteristic values (*f_std*, *f_R*, *f_u3*, *f_U* and *f_e*) can be calculated using the same definition as *f_m* respectively.

$$
\begin{aligned}
f\_m(j) &= \frac{\sum_{i=1}^{378} m(i)}{378}; \quad f\_m = \frac{\sum_{j=1}^{10} f\_m(j)}{10} \\
f\_std(j) &= \frac{\sum_{i=1}^{378} std(i)}{378}; \quad f\_std = \frac{\sum_{j=1}^{10} f\_std(j)}{10} \\
f\_R(j) &= \frac{\sum_{i=1}^{378} R(i)}{378}; \quad f\_R = \frac{\sum_{j=1}^{10} f\_R(j)}{10} \\
f\_u3(j) &= \frac{\sum_{i=1}^{378} u3(i)}{378}; \quad f\_u3 = \frac{\sum_{j=1}^{10} f\_u3(j)}{10} \\
f\_U(j) &= \frac{\sum_{i=1}^{378} U(i)}{378}; \quad f\_U = \frac{\sum_{j=1}^{10} f\_U(j)}{10} \\
f\_e(j) &= \frac{\sum_{i=1}^{378} e(i)}{378}; \quad f\_e = \frac{\sum_{j=1}^{10} f\_e(j)}{10}
\end{aligned}
\tag{S2}
$$

References

1. K. Pearson, "Contributions to the mathematical theory of evolution. II. Skew Variation in homogeneous material," Phil. Trans. R. Soc. Lond. A. doi: 10.1098/rsta.1895.0010 (1895).
2. C. He, H. He, X. Li, J. Chang, Y. Wang, S. Liu, N. Zeng, Y. He, and H. Ma, "Quantitatively differentiating microstructures of tissues by frequency distributions of Mueller matrix images," J. Biomed. Opt. **20**(10), 105009 (2015).
3. A. Timo, H. Abdenour, and P. Matti, "Face description with Local Binary Patterns: Application to Face Recognition," IEEE T. Pattern Anal. **28**(12), 2037–2041 (2006).
4. Z. Zhang and Z. Zhao, "Expression Recognition based on multi-scale block local gabor binary patterns with dichotomy dependant weights," J. Netw. Comput. Appl. **30**, 964–966 (2010).
5. A. Materka and M. Strzelecki, "Texture analysis methods: a review," Technical Report; University of Lodz, Institute of Electronics COST B11 Report (1998).