K-means clustering of zebrafish embryos images acquired with AOTF-based hyperspectral microscope

A B Burlakov¹, S V Shirokov², C C Huang³ and D D Khokhlov²

¹ Department of ichthyology, Faculty of Biology, Moscow Lomonosov State University, Leninskie Gory GSP-1, Moscow, 119991, Russian Federation
² Laboratory of acousto-optic spectroscopy, Scientific and Technological Center of Unique Instrumentation of Russian Academy of Sciences, 15 Butlerova, Moscow, 117342, Russian Federation
³ Department of Biomedical Engineering, Medical Device Innovation Center, National Cheng Kung University, 1 University Road, Tainan City 70101, Taiwan

E-mail: khokhlov.dd@ntcup.ru

Abstract. Model organism studies are widely implemented in biomedical research fields. Zebrafish is a common and convenient model organism. To provide in vivo investigation of living zebrafish the non-invasive imaging methods are implemented. Hyperspectral imaging utilizing acousto-optic tunable filters is a perspective modality for zebrafish embryos and larvae automated observation. In this paper, the hyperspectral microscope based on the acousto-optical tunable filter is described. Using the hyperspectral image arrays obtained with the described setup, the K-means clustering algorithm is tested. The results obtained for different number of clusters are presented and discussed.

1. Introduction

Many aspects of modern biomedical research are closely related with model organism studies [1]. In recent years, zebrafish (Danio rerio) has become a common vertebrate model organism [2–4] due to its small size, ease of husbandry, short period (3 days) of embryonic development. These features lead to the convenience of continuous observation and experimental research of diseases and genetic modification mechanisms being passed to the offspring.

In vivo model organism studies require non-invasive techniques. Imaging of adult zebrafish is only possible by means of ultrasound techniques [5, 6] as they are not transparent for the optical radiation (except several transgenic lines). Acoustic imaging is also utilized for embryo investigation [7, 8], but moderate spatial resolution and specimen scanning time are the reasons for the use of alternative approaches. Since zebrafish embryos and early larvae are optically transparent, the efficient imaging providing high spatial resolution may be achieved by implementation of different optical imaging techniques as laser-scanning [7] and spinning-disk confocal microscopy [9, 10], Mueller-matrix microscopy [11] and microscopic multispectral imaging [12, 13]. The latter technique enables transmission spectra measurements for different tissues and organs of embryo by acquisition of spectral intensity dependencies for any desired pixel located in the field of view. Multispectral and hyperspectral data may be implemented for automatic tissue recognition and mapping.
2. Proposed method
Staring hyperspectral imaging approach utilizing tunable optical spectral filters is a versatile and common solution for biomedical applications [14, 15]. In this case, to form an array of spectral images the tunable filters (liquid crystal tunable filters, acousto-optical tunable filters, etc.) provide sequential spectral scanning. At the same time, tunable filters are free of moving parts, thus they do not require any mechanical adjustment and are controlled electronically. Acousto-optic tunable filters (AOTFs) provide high optical throughput, rather high spectral and spatial resolution, adjustable transfer function and fast arbitrary tuning. AOTFs may serve as the spectral filtering element of the hyperspectral imager due to the mentioned features. Their operation principle is based on anisotropic Bragg wideband optical radiation diffraction on the ultrasound wave [16]. The ultrasonic wave excited in the birefringent crystal acousto-optic cell (AOC) by the piezotransducer induces the Bragg grating due to elasto-optic effect. Since that, central transmission wavelength of AOTF is defined by the period of grating which is controlled by the ultrasonic wave frequency \( f \) tuning. Electronic adjustment of frequency leads to the arbitrary spectral scanning. AOTF operation principle is illustrated in figure 1. Crossed polarizers P1 and P2 provide the light polarization plane orientation required for certain configuration of anisotropic interaction and eliminate non-diffracted light.

![Figure 1. AOTF operation principle.](image)

Hyperspectral imaging (HSI) in biomedical applications provides enhanced contrast of tissues and substances with specific spectral features and act as a powerful quantitative analysis tool due to the ability of precise spectral measurements in arbitrary spatial points across the field of view [14, 15]. At the same time, interpretation of hyperspectral data lacking predefined spectral labels is a challenging task. The promising solution for automated hyperspectral data processing and interpretation is unsupervised classification or clustering, that has become a topic of significant interest in recent research associated with HIS [17–19]. The scope of this paper is to test HSI data clustering using state-of-art cluster analysis tools included in MATLAB Software for zebrafish hyperspectral images.

3. Experimental setup
The experimental setup implemented in the described study for data acquisition consists of an upright brightfield microscope with wide-band Koehler illumination system (IS), an AOTF-based hyperspectral imaging module and a PC with original self-developed software controlling AOTF and image sensor. Figure 2 demonstrates the scheme (a) and the image (b) of the experimental setup. During the data acquisition process specimen S (zebrafish embryo or larva) is placed in the Petri dish filled with water and located in the front focal plane of microscopic objective MO forming an intermediate infinity-located image. Custom developed optical telescopic coupling system CS is used for conjugation of MO and AOTF. The use of CS provides matching of MO exit pupil and AOTF entrance pupil matching and reduction of energy losses during image transfer. Objective of the monochrome CMOS camera CAM forms the narrow-band images on the image sensor as the spectral selection and scanning is provided by AOTF. In the described setup we implemented the AOTF providing the spectral bandwidth FWHM of 2.5 nm at a central wavelength \( \lambda_c = 650 \) nm and the working spectral range between 500 nm and 750 nm. The mentioned AOTF consists of crossed polarizers and two identical AOCs in order to improve spectral resolution and provide correction of image distortions inherent for AOTFs. The disadvantage of tandem AOTFs is a reduced light throughput in comparison with single-cell ones.
4. Image acquisition and processing
To form a hyperspectral image array, the AOTF working spectral range was scanned by sequential tuning of the bandwidth central wavelength with 2 nm step. The example raw spectral images for a 90-hpf zebrafish larva are presented in figure 3. For a demonstrated image series, hyperspectral image arrays were captured with a 2-minute interval within 48 minutes until the larva escaped the field of view. After raw data acquisition the flat field correction and background subtraction were implemented. To achieve motion correction and exclude the background pixels from clustering the image zone corresponding to the larva image was automatically selected by a pixel mask.

After processing and mask selection clustering procedure was performed for the obtained hyperspectral arrays. In this paper, the K-means algorithm with the default metric (squared Euclidean distance) was used for the clustering procedure testing. This relatively easy algorithm is commonly used for multispectral data clustering [19, 20]. Though, the high computational expense of this algorithm [21] that inevitably increases with multidimensional data is an obstacle for its widespread use in hyperspectral clustering. To accelerate the computation the Parallel Computing Toolbox was running on a GPU.

The initial partitions for K clusters were obtained by iterative running of the algorithm with random centroid positions. The centroid positions calculated for the first hyperspectral array acquired at the time series beginning were kept constant for the consequent hyperspectral arrays acquired 2-minute intervals. Thus, the relation between clusters for the consequent arrays was conserved.
5. Results
Clustering results for cluster number values $K$ of 14, 20, 30 and 40 for the initial hyperspectral array are presented in figure 4. To indicate the temporal changes the cluster maps for the initial array are shown in figure 5 in the left column, and for the arrays acquired after 24 minutes and 48 minutes are presented in the middle and right columns, respectively.

![Figure 4. Clustering results for the assigned $K$ values.](image1)

![Figure 5. Clustering results for the hyperspectral arrays acquired at the beginning of the experiment (left column), after 24 minutes (middle column) and after 48 minutes (right column).](image2)

The image zones corresponding to the skeletal muscles demonstrate the least number of different clusters. The most inhomogeneous image zones correspond to head and yolk of the larva, and, as the number of clusters rises, the new partitions mostly appear in these image zones. In general, cluster mapping remains stable for the different data acquisition moments, but the obtained data with higher number of clusters is more sensitive for the temporal changes. The implementation of the predefined spectral labels may improve data reliability as this leads to transformation of clustering task into
supervised classification, but this approach requires the presence of stable spectral features in the utilized wavelength range.

The proposed technique allows non-invasive in vivo spectral imaging of living embryos. At the same time, the implementation of the described hyperspectral clustering algorithm does not seem to provide reliable tissue mapping. Unfortunately, it is not clear, whether the origin of the new clusters and partition deformation are associated with changes in tissue spectral properties or caused by the noise and image artefacts. To achieve robust spectral target detection or spectral unmixing based on unsupervised learning, one should consider the implementation of the improved variations of K-means algorithm or other clustering algorithms capable of tackling high-dimensional data [21].

Ethics
The authors confirm that all experimental methods were carried out in accordance with relevant guidelines and regulations and were approved by the Lomonosov Moscow State University Bioethic Committee (Protocol #108-0).

Acknowledgements
This study is supported by Ministry of Science and Higher Education of the Russian Federation (project 0069-2019-0010). Experiments were performed using the equipment of the Center for Collective Use of the Scientific and Technological Center of Unique Instrumentation of the Russian Academy of Sciences.

References
[1] Rine J 2014 Molecular Biology of the Cell 25 549–53
[2] Lawson N D and Weinstein B M 2002 Nature Reviews Genetics 3 674–82
[3] Kawasaki T, Maeno A, Shiroishi T and Sakai N 2017 Scientific Reports 7 16508
[4] Konno M, et al. 2020 Diagnostics 10 392
[5] Goessling W, North T E and Zon L I 2007 Nature Methods 4 551–3
[6] Ho-Chiang C, Huang H and Huang C-C 2020 Quantitative Imaging in Medicine and Surgery 10 66–75
[7] Spitsbergen J 2007 Nature Methods 4 548–9
[8] Burlakov A B, Titov S A and Bogachenkov A N 2020 Journal of Physics: Conference Series 1679 022028
[9] Weber M and Huisken J 2015 Swiss Medical Weekly 145 w14227
[10] Logan S L, Dudley C, Baker R P, Taormina M J, Hay E A and Parthasarathy R 2018 PLOS ONE 13 e0198705
[11] Le Gratiet A, D’Amora M, Duocastella M, Marongiu R, Bendandi A, Giordani S, Bianchini P and Diaspro A 2019 Scientific Reports 9 19974
[12] Burlakov A B, Khokhlov D D, Machikhin A S, Titov S A, Lomonov V A and Vinogradov A V 2020 Journal of Communications Technology and Electronics 65 851–7
[13] Chen C, et al. 2021 Nature Communications 12 1118
[14] Li Q, He X, Wang Y, Liu H, Xu D and Guo F 2013 Journal of Biomedical Optics 18 100901
[15] Lu G and Fei B 2014 Journal of Biomedical Optics 19 010901
[16] Goutzoulis A, Pape D R and Kulakov S 1994 Design and Fabrication of Acousto-Optic Devices (Boca Raton, FL: CRC Press)
[17] Zhong Y, Zhang L and Gong W 2011 International Journal of Remote Sensing 32 5461–83
[18] Zhang H, Zhai H, Zhang L and Li P 2016 IEEE Transactions on Geoscience and Remote Sensing 54 3672–84
[19] Zhao Y, Yuan Y and Wang Q 2019 Remote Sensing 11 399
[20] James P T and Galen G 1997 Proc SPIE 3159 108–18
[21] Xu R and Wunschii D 2005 IEEE Transactions on Neural Networks 16 645–78