Published in final edited form as:

J Microsc. 2013 December ; 252(3): 275–285. doi:10.1111/jmi.12090.

Automated image segmentation of haematoxylin and eosin stained skeletal muscle cross-sections

F. LIU^{*,‡}, A.L. MACKEY^{‡,§}, R. SRIKUEA^{||}, K.A. ESSER[#], and L. YANG^{*,†}

^{*}Division of Biomedical Informatics, Department of Biostatistics, University of Kentucky, Lexington, Kentucky, 40536, U.S.A. [†]Department of Computer Science, University of Kentucky, Lexington, Kentucky, 40536, U.S.A. [‡]Institute of Sports Medicine, Department of Orthopaedic Surgery M, Bispebjerg Hospital, Copenhagen, Denmark [§]Centre for Healthy Ageing, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark ^{||}Department of Physiology, Faculty of Science, Mahidol University, Bangkok, 10400, Thailand [#]Department of Physiology, Center for Muscle Biology, University of Kentucky, Lexington, Kentucky, 40536, U.S.A.

Summary

The ability to accurately and efficiently quantify muscle morphology is essential to determine the physiological relevance of a variety of muscle conditions including growth, atrophy and repair. There is agreement across the muscle biology community that important morphological characteristics of muscle fibres, such as cross-sectional area, are critical factors that determine the health and function (e.g. quality) of the muscle. However, at this time, quantification of muscle characteristics, especially from haematoxylin and eosin stained slides, is still a manual or semi-automatic process. This procedure is labour-intensive and time-consuming. In this paper, we have developed and validated an automatic image segmentation algorithm that is not only efficient but also accurate. Our proposed automatic segmentation algorithm for haematoxylin and eosin stained skeletal muscle cross-sections consists of two major steps: (1) A learning-based seed detection method to find the geometric centres of the muscle fibres, and (2) a colour gradient repulsive balloon snake deformable model that adopts colour gradient in *Luv* colour space. Automatic quantification of muscle fibre cross-sectional areas using the proposed method is accurate and efficient, providing a powerful automatic quantification tool that can increase sensitivity, objectivity and efficiency in measuring the morphometric features of the haematoxylin and eosin stained muscle cross-sections.

Keywords

Cross-sections; learning; muscle; segmentation

Introduction

Skeletal muscle is the largest tissue in the body and accounts for approximately 40% of body mass. There is growing recognition that weakness of skeletal muscle is a significant biomedical health issue associated with many different chronic diseases and ageing (Peterson *et al.*, 2012). Although the loss of muscle mass clearly contributes to the loss of strength, it is becoming increasingly clear that a decline in muscle quality also plays a role. Similarly, muscle weakness that is associated with heart failure, chronic obstructive pulmonary disease (COPD) and cancer cachexia, often persists long after disease treatment (Dall'Ago *et al.*, 2006). There is agreement within the muscle research and clinical communities that important morphological characteristics of muscle fibres, such as cross-sectional area (CSA) (Maughan *et al.*, 1983), fibre shape, fibre type, the number and position of myonuclei, cellular infiltration and fibrosis and minimum Feret diameter are among many critical factors that determine the health and functionality of the muscle (Maughan *et al.*, 1983; Schantz *et al.*, 1983; Lexell *et al.*, 1988; Pernus & Erzen, 1994).

Although there are several morphological criteria for assessing muscle specimens (Lindboe & Platou, 1982; Schantz *et al.*, 1983; Lexell *et al.*, 1988; Pernus & Erzen, 1994), currently the quantification of these important muscle morphometric characteristics is still largely based on manual or semi-automatic methods that require extensive human interventions. For example, fibre CSA is commonly measured from selected areas instead of the entire muscle cross-sections. The operator uses a standard microscope and counts the fibres from sampled fields to represent the CSA of the whole section. The major issue here is that there is considerable regional variability within a given muscle specimen with respect to fibre size, so that the measurement or sampling of only a few small areas will not necessarily provide representative measurements of the specimen. In order to capture larger regional variations, biologists often crop a large number of images in the regions of interest to compensate for the region variations. If this approach is chosen, it would require a large number of images to be quantified, which often takes plenty of time and effort in the absence of an efficient, accurate and high throughput tool. For example, in one of our pilot studies, it took an experienced technician 2 months to determine fibre CSA on 900 muscle cross-sections, where a mean of 100 fibres per cross-section (90 000 muscle fibres in total) was analysed.

Automatic muscle image segmentation can greatly release the human labour and time from morphological quantification. However, some major technical difficulties exist: (1) Appearance variances due to staining; (2) Weak object boundaries of the muscle fibre; (3) Sample defects such as folding or stretching of the muscle tissue and (4) freeze artefacts in preparing muscle specimens. All these difficulties (shown in Fig. 1) bring big challenges to current automatic segmentation algorithm.

In this study, we report the development of an automatic, robust and high throughput image analysis algorithm for skeletal muscle segmentation. As a general framework, the algorithm contains two steps: seed detection and contour refinement. We provide experimental data to demonstrate the accurate segmentation results compared with human annotation.

Materials and methods

Haematoxylin and eosin (H&E) stained muscle specimens

H&E stained human muscle specimens are utilized to evaluate the performance of the automatic image segmentation algorithm. Approval for the human study was obtained from The Research Ethics Committees of the Capital Region of Denmark and the study conformed to the standards set by the Declaration of Helsinki. All volunteers were given the written informed consent before inclusion. Muscle biopsies were obtained from the vastus lateralis muscles of young healthy men under local anaesthetic (1% lidocaine), using the percutaneous needle biopsy technique with a 5-mm biopsy needle with manual suction. Immediately after extraction, the specimen was aligned, embedded in Tissue-Tek, frozen in isopentane precooled by liquid nitrogen and stored at -80°C . The $10\text{ }\mu\text{m}$ sections were cut at -25°C using a cryostat, placed on SuperFrost Plus glass slides (Menzel-Gläser, Braunschweig, Germany), and stored at -80°C . Sections were stained using a standard H&E staining method as follows: Sections were incubated in filtered Mayer's Haematoxylin solution for 1 min and washed under running water for 5 min. After incubation in an Eosin solution containing one drop of acetic acid (per 50 mL solution), sections were rinsed quickly in water, dehydrated in a series of diluted ethanols (70% and 90%), cleared in xylene and mounted in Pertex (Histolab, Gothenburg, Sweden) before being covered with cover slips. The images were captured with a 20X objective on an Olympus BX51 microscope, using a digital camera (Olympus DP71) controlled by a software developed by Olympus Soft Imaging Solutions (Munster, Germany).

Automatic image segmentation of H&E stained muscle cross-sections

As mentioned before, muscle fibres exist as closely apposed cells due to their unique anatomy. To accurately quantify morpho-metric features of muscle, the first prerequisite is to segment each individual muscle fibre. Recently, many automatic and semi-automatic methods have been proposed for segmenting closely apposed cells (touching cells). The watershed family of algorithms has become one of the most common and standard choices. However, the primary limitation of watershed is oversegmentation. Methods such as marker-controlled watershed (Roerdink & Meijster, 2001; Zhou *et al.*, 2005) and rule-based strategies (Wählby *et al.*, 2002; Lin *et al.*, 2005; Yu *et al.*, 2009) were developed to address this problem by merging oversegmented patches, but it is usually difficult to derive a general rule for merging. Other methods, such as Voronoi diagrams (Jones *et al.*, 2005; Zhou *et al.*, 2005), double threshold-based watershed and statistical analysis (Wählby *et al.*, 2002), concavity detection (Kothari *et al.*, 2009), template matching (Díaz *et al.*, 2007), graph-based methods (Chen *et al.*, 2008; Nasr-Isfahani *et al.*, 2008; Faustino *et al.*, 2009), maximum intensity linking (Elter *et al.*, 2006) and original level set-based methods (Osher & Sethian, 1988; Malladi *et al.*, 1995; Zhao *et al.*, 1996) with different terms (Chan & Vese, 2001; Zhang *et al.*, 2004; Yan *et al.*, 2008), have been proposed, but none of them were specifically designed to process heavily apposed cells like muscle fibres.

Because the number and location of the muscle fibres are unknown, finding the geometric centre of the fibres is the first critical step, and the segmentation accuracy is largely dependent on this step (e.g. missing a small fibre can lead to a bigger measurement of CSA).

We have designed a general image segmentation framework that contains the following two steps: (1) Seed detection, which involves the detection of the geometric centres (seeds) of muscle fibres and (2) contour evolution, a deformable model that will be initiated from the contours of the seeds and will evolve to converge to the boundaries of the muscle fibres. The algorithm framework is shown in Figure 2.

Learning-based seed detection

For automatic muscle image segmentation, the first step is to find the geometric centre of each individual muscle fibre. Image textures are used as features to distinguish muscle fibres and boundaries. Image texture represents the image intensity distribution patterns and can be modelled as textons (Julesz, 1981). Textons are defined as repetitive local features that humans perceive as being discriminative between textures. We use the multiple scale Schmid filter bank (Schmid, 2001) composed of 13 rotation invariant filters (shown in Fig. 3):

$$F(r, \sigma, \tau) = F_0(\sigma, \tau) + \cos\left(\frac{\pi r \tau}{\sigma}\right) e^{-\frac{r^2}{2\sigma^2}}. \quad (1)$$

The image filtering responses are clustered using *K*-means to generate a large code book. A texton library is constructed from the corresponding cluster centres. The pixelwise segmentation of digitized muscle specimen is performed by classification. Based on the labelled ground truth masks, 2000 positive and negative pixels are extracted from the muscle fibre and nonmuscle fibre regions in the image. The appearance of the neighbours of each training pixel is modelled by a compact quantized description – texton histogram, where each pixel is assigned to its closest texton using the following equation:

$$h(i) = \sum_{j \in I} \text{count}(T(j) = i). \quad (2)$$

Here *I* denotes digitized muscle image, *i* is the *i*th element of the texton dictionary and *T*(*j*) returns the texton assigned to pixel *j*. The windowed texton histogram is computed around each individual training pixel.

After normalization, the texton histogram actually represents the texton channel frequency distribution in a local neighbourhood around the centred image pixel. In order to compensate for scale changes, the texton histogram is extracted from five different window sizes (4, 8, 16, 32, 64 pixels, respectively) and concatenated into one large feature vector. This concatenated texton histogram is used as features to train the classifiers. The integral histogram (Porikli, 2005) is used to calculate the windowed texton histogram. The algorithm starts by exploiting the spatial arrangement of data points. It then recursively propagates an aggregated histogram. The aggregated histogram starts from the origin and traverses through the remaining points along a scan-line. At each step, a single bin is updated using the values of the integral histogram at the previous visited neighbouring data points. The integral histogram method speeds up feature extraction significantly.

After the texton histograms are calculated, machine learning is applied to detect the seeds of the muscle fibres. The machine learning unit contains both the training and testing components. We have proposed to use asymmetric online boosting for seed detection in muscle image. Adaboost (Friedman *et al.*, 2000) works by sequentially applying a simple classification method on a reweighted version of the training data and produces a sequence of weak classifiers $h_i(x)$, $i = 1, 2, \dots, T$, where T is the number of weak learners. Each weak classifier will use one image feature, such as one texture feature in our case. The final strong classifier $f(x)$ is assembled from all weak classifiers $h_i(x)$ to accomplish the assigned learning tasks (automatic detection of the seeds of muscle fibres). After the training is done, the final learned expert system will contain a set of weak classifiers. Because each weak classifier represents one image feature, the final classifier not only gives us an expert system, but also denotes a combination of representative image features for seed (geometric centre of muscle fibre) detection.

Similar to standard Adaboost algorithms, asymmetric online boosting constructs a strong classifier from a set of weak learners with the smallest classification errors. The difference is that, with asymmetric online boosting, the set of weak learners are online-updated after receiving the feedbacks from the users. The sample weights are then updated based on the classification results asymmetrically taking into consideration the imbalanced training set.

The entire machine learning procedure is outlined in Figure 4. The pseudo-code for asymmetric online boosting is summarized in Algorithm 1. The strong classifier $H(x)$ is updated incrementally using users' feedback in the current iteration. Each weak learner $h_i(x)$ and its corresponding weight α_i are updated with a learning rate γ (shown in Algorithm 1). This online updating schema enables the incrementally trained classifier in order to avoid making the same type of errors in the future.

In standard Adaboost, accurate performance requires a large and balanced training set. However in muscle image analysis, we will have a limited amount of labelled training samples, and practically it is impossible to guarantee a balanced training set. For example, for seed detection for muscle fibre segmentation, a balanced training set would require an equally distributed pixels from the centres and boundaries of the muscle fibres. Instead of weighting a type of muscle with less training samples (we called positive) and more training samples (we called negative) evenly, we force the penalty of a false positive to be k times larger than a false negative to compensate for the imbalance data size during the training. Compared with the standard loss function $\exp(-y_i * H_i(x_i))$ defined in Adaboost, the loss function in our algorithm is

$$\epsilon(H) = E(ALoss(H(x), y)), \quad (3)$$

where

$$ALoss(H(x), y) = \begin{cases} k^{1/2} & \text{if } y = -1, H(x) = 1 \\ k^{-1/2} & \text{if } y = 1, H(x) = -1 \\ 0 & \text{otherwise,} \end{cases} \quad (4)$$

which is proven to be effective in (Viola & Jones, 2002). This asymmetric loss function can be integrated into an online boosting algorithm by multiplying the original weights $\exp(-y_i * H_f(x_i))$ with $\exp(y_i \sqrt{k})$. For an online boosting algorithm with T iterations, samples are weighted by $\exp(y_i \sqrt{k}/T)$ at each iteration to prevent the asymmetric weights from being absorbed by the first selected weak learner.

Given a test image, we will apply the trained asymmetric online boosting classifier for each pixel and separate the image into muscle fibre seeds and nonmuscle fibre seeds regions. Using the multiscale texton histogram (three different window sizes), integral histogram and asymmetric online boosting, a fast and accurate pixelwise segmentation algorithm can be implemented for detecting geometric centres of muscle fibres.

Deformable model-based segmentation using colour gradients

After learning-based automatic detection of the geometric centres (seeds), each seed will represent one muscle fibre. The boundaries of these seeds will be utilized to initiate the repulsive colour gradient balloon snake (BS) model to evolve and converge to the boundaries of muscle fibres.

A snake is an active curve as $v(s) = (x(s), y(s))$, $s \in [0, 1]$, moving through the image domain to minimize its energy functional, under the influence of internal and external forces. To enforce snakes to inflate or deflate, Cohen (1991) introduced a pressure force to propose the BS model. The external force F_{ext}^B is calculated by

$$F_{ext}^B = \gamma n(s) - \lambda \frac{\nabla E_{ext}(v(s))}{\|\nabla E_{ext}(v(s))\|}, \quad (5)$$

where $n(s)$ represents the normal vector (pressure force) to the curve at the specific point on $v(s)$ and $\nabla E_{ext}(v(s))$ is defined as image force, where $E_{ext}(v(s)) = -\|\nabla I(v(s))\|^2$ [$I(v(s))$ is the original image]; γ and λ are the weighting parameters controlling pressure and image forces, respectively.

The E_{ext} in Eq. (5) involves the calculation of image gradients. In grey-level images, the gradient is defined as the first derivative of the image luminance. It has a high value in those regions exhibiting high luminance contrast. We adopt the definition of gradients for colour images (Di Zenzo, 1986; Sapiro & Ringach, 1996; Gevers, 2002). By contrast to previous approaches, we define the colour gradient in Luv colour space rather than RGB colour space because Euclidian metrics and distances are perceptually uniform in Luv colour space, which is not the case in RGB colour space (Sapiro & Ringach, 1996).

Let $(x, y) \in R^3$ be a colour image, based on classical Riemannian geometry results (Kreyszig, 1991), the $L2$ norm can be written in matrix form:

$$d\Gamma^2 = \begin{bmatrix} dx \\ dy \end{bmatrix}^T \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix}, \quad (6)$$

where $g_{11} = \left[\frac{\partial \Gamma}{\partial x} \right]^2$, $g_{12} = g_{21} = \frac{\partial \Gamma}{\partial x} \frac{\partial \Gamma}{\partial y}$, $g_{22} = \left[\frac{\partial \Gamma}{\partial y} \right]^2$.

The quadratic form (6) achieves its extrema changing rates in the directions of the eigenvectors of matrix $[g_{ij}]$, and the changing magnitude is decided by its eigenvalues λ_+ and λ_- . In our approach, we select $\sqrt{\lambda_+ - \lambda_-}$ (Sapiro & Ringach, 1996; Gevers, 2002) to define the colour gradient

$$\nabla \Theta = \sqrt{\lambda_+ - \lambda_-}, \quad (7)$$

where

$$g_{11} = \begin{bmatrix} \left| \frac{\partial L}{\partial x} \right|^2 \\ \left| \frac{\partial u}{\partial x} \right|^2 \\ \left| \frac{\partial v}{\partial x} \right|^2 \end{bmatrix}, \quad g_{22} = \begin{bmatrix} \left| \frac{\partial L}{\partial y} \right|^2 \\ \left| \frac{\partial u}{\partial y} \right|^2 \\ \left| \frac{\partial v}{\partial y} \right|^2 \end{bmatrix}, \quad (8)$$

and

$$g_{12} = g_{21} = \begin{bmatrix} \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 v}{\partial x \partial y} \end{bmatrix}. \quad (9)$$

Here L, u, v correspond to the three channels in Luv colour space.

BS model cannot be directly used for touching object segmentation. If all BSs move independently, they will cross with one another. We introduce an interactive scheme to form a repulsive balloon snake (RBS) model for touching cell segmentation. The intrinsic idea of RBS we designed is based on the following observations. The cell contour should be driven by its own forces as well as extrinsic forces from other deformable contours; both amplitude and direction of these extrinsic forces should vary with the corresponding distance between snakes. When two snakes are far away, their movement should be dominantly controlled by their own driven forces (internal and external forces); when they get closer, each snake should receive repulsive forces from all other snakes. As a result, the extrinsic force can prohibit snakes from crossing or merging.

Given an image I with N cells (denoted by N curves v_i , $i = 1, \dots, N$), the new repulsive external force F_{ext}^{RB} for v_i is defined as

$$F_{ext}^{RB} = \gamma n_i(s) - \lambda \frac{\nabla E_{ext}(v_i(s))}{\|\nabla E_{ext}(v_i(s))\|} + \omega \sum_{j=1, j \neq i}^N \int_0^1 f(d_{ij}(s, t)) n_j(t) dt, \quad (10)$$

where $d_{ij}(s, t) = \|v_i(s) - v_j(t)\|_2$ is the Euclidian distance between $v_i(s)$ and $v_j(t)$. $f(x) > 0$, $x \in (0, +\infty)$, represents a monotonic decreasing function ($f(x) = x^{-2}$ in our case), and ω weights

the repulsive force. For a specific point $v_i(s)$, the closer it moves to other snakes, the more repulsive forces it will receive. Unlike the original BS, RBS moves curves under the influence of their own driven forces and extrinsic repulsive forces from other snakes. When these two types of forces achieve a balance, snakes stop evolving. The contour evolution results using colour gradient RBS is shown in Figure 5.

Experimental results

In order to validate the accuracy of the automatic algorithm, 30 randomly selected H&E stained human muscle tissues (over 3000 muscle fibres) from the cohort of 900 images are utilized for man-machine evaluation. Pixelwise segmentation accuracy is calculated to measure the difference between manual and automatic methods. All the testing images are manually segmented by an experienced technician who has 5 years experience in analysing the morphological characteristics of skeletal muscle. In Figure 6, we have provided an image of one of the automatic segmentation results. The original image is shown in Figure 6(A). The manual annotation of each muscle fibre is shown in Figure 6(B), with green contours denoting the manually delineated boundaries of muscle fibres and white numbers representing the manually calculated CSA (in μm^2). The automatic segmentation results are shown in Figure 6(C), with blue contours representing the automatically segmented muscle fibre boundaries. The automatically calculated CSA is also labelled with white fonts on each individual muscle fibre in Figure 6(C).

As a learning-based method, our proposed algorithm can produce robust segmentation results even for challenging H&E images:

1. Muscle images with weak fibre boundaries, illustrated in Figure 7(A): This is a challenging case due to the weak boundaries between muscle fibres. Even with the help of some automatic image segmentation software, it still requires extensive manual postprocessing efforts to correct the automatic segmentation results. For example, it took an experienced technician more than 40 min to calculate the CSA for an image shown in Figure 7(A). On the contrary, our algorithm can generate automatic image segmentation results in less than 1 min. The direct output from our automatic software is shown in Figure 7(C). Some zoom-in image fields are shown below Figure 7(C) for illustration purposes.
2. Muscle images with artefacts, illustrated in Figure 7(B): During muscle specimen preparation, freeze artefacts can be introduced. Air bubbles, tears and folds might also occur during sectioning, staining and mounting procedures. All of these artefacts and defects inevitably introduce challenges for automatic image analysis. This is demonstrated in Figure 7(B). The freeze artefacts are highlighted by the dotted green rectangle, whereas the sample defects are highlighted with a solid green rectangle. Even for this challenging case, our proposed algorithm can still provide robust automatic segmentation results. The results obtained from the automatic segmentation algorithm are shown in Figure 7(D). Some zoom-in image fields are shown below Figure 7(D) for illustration purposes as well.

In order to better evaluate the proposed algorithm, a quantitative, pixelwise comparative experiment is conducted between the results generated using the automatic method and manual annotations in this testing data set. The pixelwise difference for each image is summarized in Table 1. The percentage difference between automatic and manual results for CSA measurement in H&E stained digitized muscle cross-sections range from -1.71% to 2.09% , $+2.92\%$, with an average difference of Some automatic segmentation results are presented in Figure 8.

The algorithm also exhibits good scalability, which means that it can handle a muscle image with a large number of muscle fibres in an efficient manner. In Figure 8, we demonstrate the automatic segmentation results on a large image patch that contains hundreds of muscle fibres. Our automatic image analysis algorithm can still accurately segment the fibre boundaries in less than 1 min, compared with 1 h manual annotation required for this big image patch with hundreds of muscle fibres.

Discussion

Some of the technical challenges for automatic muscle cross-section image segmentation result from the unique morphological characteristics of muscle fibres. Skeletal muscle is composed of long, multinucleated cells (fibres) tightly grouped into fascicles, interspersed with other mononucleated cell types and surrounded by connective tissue and fat. This structural organization, coupled with aforementioned technical challenges, can lead to confusing and overlapping muscle cell boundaries that cannot be accurately segmented using existing image analysis methods, thus requiring extensive human correction and postprocessing. The specific morphological characteristics of skeletal muscle pose big challenges in biomedical image analysis known as touching cell segmentation. Each apposed fibre needs to be precisely separated before accurately measuring its morphometric features.

Current image analyses methods (manual and semiautomated) are relatively inefficient and labour intensive, especially for challenging images containing biological variations and/or technical defects that arise during sample processing, including staining variations, weak cell boundaries, freeze artefacts and sample defects. Although there are some automatic programs available, most of them are based on relatively simple thresholding-based algorithm or CSA estimation based on ellipse fitting, which cannot produce similar accurate algorithm as our proposed method. In addition to CSA, we also want to point out that our algorithm can also produce minimum Feret diameter through some extensions. Minimum Feret diameter is less dependent on fibre orientation, and is often preferred in many morphological measurements.

In this paper, we have developed an automatic image segmentation approach that can accurately and efficiently quantify H&E stained muscle fibre size in cross-sections. Previously, it often take a muscle research lab a tremendous amount of time to obtain accurate CSA data from H&E stained sections for one project. The advantage of this software is its potential for a high-throughput analysis which will transform days of labour-intensive work into minutes. The algorithm can also produce an objective measurement with

less bias. Furthermore, the distribution of muscle fibre sizes can later be used to evaluate muscle quality and provide biological correlation for issues such as muscle strength. Finally, as a data-driven method, the algorithm can be applied to other tissues, such as adipose, where the quantification of cell size, shape, fibrosis and cellular infiltration are quite time-consuming.

Acknowledgments

The project described was supported by the National Center for Research Resources, UL1RR033173, and the National Center for Advancing Translational Sciences, UL1TR000117. The project was also supported by the seed funding from Center for Muscle Biology (CMB) at University of Kentucky. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health. This project is also funded from MYOAGE (no. 223576) through the European Commission under FP7, and Nordea Foundation (Healthy Aging grant). The authors want to thank for the detailed comments of the reviewers, which greatly improve the quality of this manuscript.

References

- Chan TF, Vese LA. Active contours without edges. *IEEE Trans. Image Process.* 2001; 10:266–277. [PubMed: 18249617]
- Chen C, Li H, Zhou X, Wong S. Constraint factor graph cut-based active contour method for automated cellular image segmentation in RNAi screening. *J. Microsc.* 2008; 230:177–191. [PubMed: 18445146]
- Cohen LD. On active contour models and balloons. *Comput. Vis. Graph. Image Process.* 1991; 53:211–218.
- Dall'Ago P, Chiappa GR, Guths H, Stein R, Ribeiro JP. Inspiratory muscle training in patients with heart failure and inspiratory muscle weakness: a randomized trial. *J. Am. Coll. Cardiol.* 2006; 21:757–764. [PubMed: 16487841]
- Di Zeno S. A note on the gradient of a multi-image. *Comput. Vis. Graph. Image Process.* 1986; 33:116–125.
- Díaz, G.; Gonzalez, F.; Romero, E. *Progress in Pattern Recognition, Image Analysis and Applications.* Springer; New York, NY: 2007. Automatic clump splitting for cell quantification in microscopical images; p. 763-772.
- Elter M, Daum V, Wittenberg T. Maximum-intensity-linking for segmentation of fluorescence-stained cells. *Proceedings of Microscopic Image Analysis with Applications in Biology.* 2006:46–50.
- Faustino, GM.; Gattass, M.; Rehen, S.; de Lucena, C. Automatic embryonic stem cells detection and counting method in fluorescence microscopy images. *IEEE International Symposium on Biomedical Imaging*; 2009. p. 799-802.
- Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. *Ann. Statist.* 2000; 38:337–374.
- Gevers T. Adaptive image segmentation by combining photometric invariant region and edge information. *IEEE Trans. Pattern Anal. Mach. Intell.* 2002; 24:848–852.
- Jones, TR.; Carpenter, A.; Golland, P. *Computer Vision for Biomedical Image Applications.* Springer; New York, NY: 2005. Voronoi-based segmentation of cells on image manifolds; p. 535-543.
- Julesz B. Textons, the elements of texture perception, and their interactions. *Nature.* 1981; 290:91–97. [PubMed: 7207603]
- Kothari, S.; Chaudry, Q.; Wang, MD. Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques. *IEEE International Symposium on Biomedical Imaging*; 2009. p. 795-798.
- Kreyszig, E. *Differential Geometry.* 1st edn. Dover Publications; Mineola, New York: 1991.
- Lexell J, Taylor CC, Sjostrom M. What is the cause of the ageing atrophy? Total number, size and proportion of different fiber types studied in whole vastus lateralis muscle from 15- to 83-year-old men. *J. Neurol. Sci.* 1988; 84:275–294. [PubMed: 3379447]

- Lin G, Chawla MK, Olson K, Guzowski JF, Barnes CA, Roysam B. Hierarchical, model-based merging of multiple fragments for improved three-dimensional segmentation of nuclei. *Cytometry A*. 2005; 63:20–33. [PubMed: 15584021]
- Lindboe CF, Platou CS. Disuse atrophy of human skeletal muscle. *Acta. Neuropathol.* 1982; 56:241–244. [PubMed: 6211887]
- Malladi R, Sethian JA, Vemuri BC. Shape modeling with front propagation: a level set approach. *IEEE Trans. Pattern Anal. Mach. Intell.* 1995; 17:158–175.
- Maughan RJ, Watson JS, Weir J. Strength and cross-sectional area of human skeletal muscle. *J. Physiol.* 1983; 338:37–49. [PubMed: 6875963]
- Nasr-Isfahani, S.; Mirsafian, A.; Masoudi-Nejad, A. A new approach for touching cells segmentation. *International Conference on BioMedical Engineering and Informatics*; 2008. p. 816-820.
- Osher S, Sethian JA. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Comp. Phys.* 1988; 79:12–49.
- Pernus F, Erzen I. Fibre size, atrophy, and hypertrophy factors in vastus lateralis muscle from 18- to 29-year-old men. *J. Neurol. Sci.* 1994; 121:194–202. [PubMed: 8158215]
- Peterson CM, Johannsen DL, Ravussin E. Skeletal muscle mitochondria and aging: a review. *J. Aging Res.* 2012; 2012:1–21.
- Porikli, F. Integral histogram: a fast way to extract histograms in Cartesian spaces. *IEEE International Conference on Computer Vision and Pattern Recognition*; 2005. p. 829-836.
- Roerdink JBTM, Meijster A. The watershed transform: definitions, algorithms and parallelization strategies. *Fundam. Inform.* 2001; 41:187–228.
- Sapiro G, Ringach DL. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Trans. Image Process.* 1996; 5:1582–1586. [PubMed: 18290076]
- Schantz P, Randall-Fox E, Hutchison W, Tyden A, Astrand PO. Muscle fibre type distribution, muscle cross-sectional area and maximal voluntary strength in humans. *Acta Physiol. Scand.* 1983; 117:219–226. [PubMed: 6223509]
- Schmid, C. Constructing models for content-based image retrieval. *IEEE International Conference on Computer Vision and Pattern Recognition*; 2001. p. 39-45.
- Viola P, Jones M. Fast and robust classification using asymmetric adaboost and a detector cascade. *Advances in Neural Information Processing System*. 2002; Vol. 14:1311–1318.
- Wählby C, Lindblad J, Vondrus M, Bengtsson E, Björkstén L. Algorithms for cytoplasm segmentation of fluorescence labelled cells. *Anal. Cell. Pathol.* 2002; 24:101–111. [PubMed: 12446959]
- Yan P, Zhou X, Shah M, Wong ST. Automatic segmentation of high-throughput RNAi fluorescent cellular images. *IEEE Trans. Inf. Technol. Biomed.* 2008; 12:109–117. [PubMed: 18270043]
- Yu W, Lee HK, Hariharan S, Bu W, Ahmed S. Quantitative neurite outgrowth measurement based on image segmentation with topological dependence. *Cytometry A*. 2009; 75:289–297. [PubMed: 18951464]
- Zhang, B.; Zimmer, C.; Olivo-Marin, J-C. Tracking fluorescent cells with coupled geometric active contours. *IEEE International Symposium on Biomedical Imaging*; 2004. p. 476-479.
- Zhao H-K, Chan T, Merriman B, Osher S. A variational level set approach to multiphase motion. *J. Comp. Phys.* 1996; 127:179–195.
- Zhou, X.; Liu, K-Y.; Bradley, P.; Perrimon, N.; Wong, ST. Towards automated cellular image segmentation for RNAi genome-wide screening. *International Conference on Medical Image Computing and Computer- Assisted Intervention*; New York, NY: Springer; 2005. p. 885-892.

Fig.1.

The challenges of automatic image segmentation. The H&E stained muscle images 1, 2, 3 and 4 correspond to: (1) weak cell boundaries of the muscle fibres; (2) nonmuscle fibres; (3) sample preparation artefacts due to staining and folding of muscle sections and (4) freeze artefacts.

Fig. 2.

The entire automatic muscle fibre segmentation framework. The algorithm flowchart of our proposed muscle fibre segmentation framework.

Fig. 3.

The image filter bank used to generate the texton histogram. The Schmid filter bank used to generate the texton histogram for learning-based seed detection of muscle fibres. The Schmid filter bank contains 13 rotationally invariant filters. As one can tell from the illustration of these filters, all the image filters in Schmid filter bank is rotational symmetry.

Fig. 4.

The learning-based automatic detection of the geometric centres of the muscle fibres. The entire robust muscle fibre seed detection procedure using the asymmetric online boosting. The positive and negative training sample image patches are extracted from the digitized muscle specimens. An asymmetric Adaboost classifier is trained to separate positive and negative samples. The classifier is online updated to adopt to different staining variations. A dynamic library is also online updated. The final online classifier is used to separate the centre region of the muscle fibres (seeds) and the cell boundaries among different muscle fibres.

Fig. 5.

The colour gradient-based deformable model used to find the boundary of the muscle fibres. The 1st, 10th, 15th and 50th iteration of the contour evolution using the colour gradient repulsive balloon snake (RBS) model initiated with the borders of detected muscle seeds. Note that from left to right, we show how the colour gradient RBS gradually converge to the muscle cell boundaries.

Fig. 6.

The comparative experimental results. An example of segmentation results using the proposed method. (A) The original image; (B) the green lines represent the manually delineated contours of muscle fibres, and the white numbers denote the CSA (in μm^2) of each individual muscle fibre and (C) the automatic segmentation results using our automatic algorithm.

Fig. 7.

Two representative image patches that illustrate the automatic segmentation results of the challenging cases: (A) A challenging case due to the weak boundaries between muscle fibres. (B) Some freeze artefacts and sample defects during muscle specimen preparation. (C) The automatic segmentation results of (A). Some zoom-in image patches are presented below (C) for illustration purposes. (D) The automatic segmentation results of (B). Some zoom-in image patches are presented below (D) for illustration purposes.

Fig. 8.

Automatic segmentation results for 12 H&E stained image patches. The automatic delineated fibre boundaries are drawn with green contours. The CSA is labelled on each segmented muscle fibre with white fonts. The automatic segmentation result of a bigger H&E stained muscle cross-section that contains hundreds of muscle fibres is shown in the bottom of the figure.

Table 1

The quantitative comparison of cross-sectional area (CSA) calculated using the proposed automatic segmentation method and the manual annotation of 30 randomly selected image patches.

Image ID	Number of Fibres	C S A (manual)	C S A (auto)	Difference (%)
1	87	1124	1156	2.86
2	78	1194	1224	2.53
3	82	1253	1288	2.8
4	51	1834	1878	2.41
5	69	1252	1280	2.27
6	78	1249	1280	2.48
7	72	1346	1381	2.59
8	78	1269	1299	2.4
9	78	1309	1346	2.86
10	79	1287	1317	2.3
11	52	1789	1800	0.63
12	72	1338	1373	2.61
13	63	1642	1678	2.17
14	42	1903	1940	1.93
15	73	1362	1402	2.92
16	65	1535	1578	2.76
17	75	1361	1396	2.6
18	51	1729	1739	0.52
19	77	1302	1339	2.86
20	47	1819	1830	0.58
21	46	1971	2017	2.31
22	78	1208	1211	0.23
23	80	1222	1238	1.3
24	37	1932	1899	1.71
25	67	1435	1465	2.07
26	65	1371	1402	2.23
27	72	1435	1467	2.25
28	57	1604	1640	2.25
29	57	1530	1560	1.91
30	83	1073	1075	0.14
Mean	67	1456	1483	2.05
Std	14	259	258	0.83
Min	37	1073	1075	0.14
Max	87	1971	2017	2.92
Median	72	1362	1399	2.29

Algorithm 1**Asymmetric online boosting**

Define: Let $\{x_i, y_i\}_{i=1}^N$ represent the data set $x_i = \{x_{i1}, \dots, x_{id}\}$, where d is the dimension of the feature vector, and $y_i \in \{-1, 1\}$ represents its label. The γ is the learning rate. $\omega = \{\omega_1, \omega_2, \dots, \omega_N\}$ is the samples' weight. The $H_t(x)$ is the strong classifier in the t th iteration with confidence function $f_t(x)$, the weighted accuracy and error are denoted as λ_m^{cor} and λ_m^{wro} , $m=1, \dots, M$, where M is number of weak classifiers in the pool. The I is the indicator function.

```

1   $\omega_i = 1, i = 1, \dots, N$ 
2  for  $t = 1, 2, \dots, T$  do
3       $\omega_i = \omega_i \exp\left(\frac{y_i \sqrt{k}}{T}\right), \omega_i = \frac{\omega_i}{\sum_{j=1}^N \omega_j}$ 
4      for  $m = 1, 2, \dots, M$  do
5           $h_m(x) = \text{update}(h_m(x), \{x_i, y_i\}_{i=1}^N, \omega), \lambda_m^{cor} = \lambda_m^{cor}(1 - \gamma) + \gamma \sum_{i=1}^N \omega_i * I(h_m(x_i) = y_i)$ 
6           $\lambda_m^{wro} = \lambda_m^{wro}(1 - \gamma) + \gamma \sum_{i=1}^N \omega_i * I(h_m(x_i) \neq y_i), \epsilon_m = \frac{\lambda_m^{wro}}{(\lambda_m^{wro} + \lambda_m^{cor})}$ 
7      end for
8       $m^* = \text{argmin}_m \epsilon_m, \alpha_t = \frac{1}{2} \log\left(\frac{1 - \epsilon_{m^*}}{\epsilon_{m^*}}\right), \omega_i = \omega_i \exp\left(-y_i * \alpha_t * h_{m^*}(x_i)\right)$ 
9       $f_t(x) = f_{t-1}(x) + \alpha_t h_{m^*}(x), H_t(x) = \text{sign}(f_t(x))$ 
10 end for
```