Chapter 10
A Statistical and Biophysical Toolbox to Elucidate Structure and Formation of Stress Fibers

Benjamin Eltzner, Lara Hauke, Stephan Huckemann, Florian Rehfeldt and Carina Wollnik

Abstract We are concerned with statistically validated early mechanically guided differentiation of human mesenchymal stem cells (hMSCs). This chapter reviews and extends methods of fixed and live imaging of hMSCs, automated reliable and unbiased near real-time filament extraction and digitization for massive data via the FilamentSensor, suitable aggregation of simple (area, mean orientation, aspect ratio and order parameter) and advanced (orientation mode persistence and orientation fields) data descriptors and methods of their non-euclidean inferential statistics. Exemplary, we study the morphology of stress fibers in fixed and live hMSCs within 24 h post seeding on elastic matrices exhibiting Young's moduli of 1 kPa (soft, brain-like elasticity), 11 kPa (intermediate, muscle-like stiffness) and 30 kPa (hard, pre-calcified bone rigidity). The combination of these methods constitutes a novel integrated toolbox, where for instance, statistical insight may be used to guide experimental design.

B. Eltzner (✉) · S. Huckemann
Felix-Bernstein-Institute for Mathematical Statistics in the Biosciences, Universität Göttingen, Göttingen, Germany
e-mail: benjamin.eltzner@mathematik.uni-goettingen.de

S. Huckemann
e-mail: stephan.huckemann@mathematik.uni-goettingen.de

L. Hauke · F. Rehfeldt · C. Wollnik
Third Institut of Physics - Biophysics, Universität Göttingen, Göttingen, Germany
e-mail: lara.hauke@phys.uni-goettingen.de

F. Rehfeldt
e-mail: rehfeldt@physik3.gwdg.de

C. Wollnik
e-mail: carina.wollnik@gmx.de

10.1 Introduction

During the last two decades it has become evident that the mechanical properties of the cellular micro-environment are as important for cellular behavior and homeostasis as traditionally investigated biochemical cues [1, 2]. Especially striking was the finding that differentiation of human mesenchymal stem cells (hMSCs) can be mechanically induced by culturing them on elastic substrates of different Young's moduli E [3]. While upregulation of specific differentiation markers is typically observed after five or more days, fundamental mechanical interactions between cells and the substrates take place immediately after adhesion on the substrate. Interestingly, during this early stage (within the first 24 h) of this mechano-guided differentiation process in hMSCs, the structure and polarization of actin-myosin *stress fibers* as quantified by an order parameter S depend critically on Young's modulus E [4].

Stress fibers are contractile structures mainly composed of actin *filaments*, myosin motor mini filaments (in particular non-muscle myosin II isoforms) and distinct types of actin cross-linking proteins (e.g. α-actinin, fascin, etc.). They play the role of 'cellular muscles' generating contractile forces and connecting to the *extracellular matrix* (ECM) via focal adhesions, thereby also transmitting forces to the ECM [5]. Acto-myosin filaments are also considered to be the principal force sensors of the cell that translate mechanical cues from the surroundings into biochemical signaling, eventually leading to cell differentiation [2, 6]. Previous experiments with fixed cells revealed the important role of acto-myosin cytoskeleton structure formation for the mechanically induced differentiation of hMSCs [4, 7].

Building statistically validated models and theories linking substrate elasticity to early hMSCs differentiation, the filament structure has to be visualized over time, binarized in an unbiased fashion, aggregated into descriptors and analyzed, possibly within a feedback loop. Due to high biological diversity, large amounts of data are required for statistical power. In turn, such massive data require near real-time processing.

In order to visualize these filaments selectively, fluorescence microscopy proves useful and typical images of acto-myosin stress fibers of different quality in fixed cells are displayed in Fig. 10.1. One of the main differences, however, of experimental visualization of the acto-myosin cytoskeleton in fixed cells at particular time points and life cell imaging is given by the signal to noise ratio of the microscope images. Fixed cells can be stained with many different methods and allow for saturation with fluorescent dyes that typically lead to nice and crisp images (see top row of Fig. 10.1). In contrast, life cell imaging, as detailed in Sect. 10.2, relies on transfection with fluorescent fusion proteins that need to be expressed in the cell and typically leads to worse signal to noise ratios (see bottom row of Fig. 10.1), that are challenging for subsequent image processing.

Once these images are obtained, the challenge consists in extracting the underlying filament structure in near real-time in an automated and unbiased fashion. To this end, we review the *FilamentSensor* (FS) from [8] in Sect. 10.3. It integrates general and specifically tailored preprocessing with an elaborate binarization routine, to identify
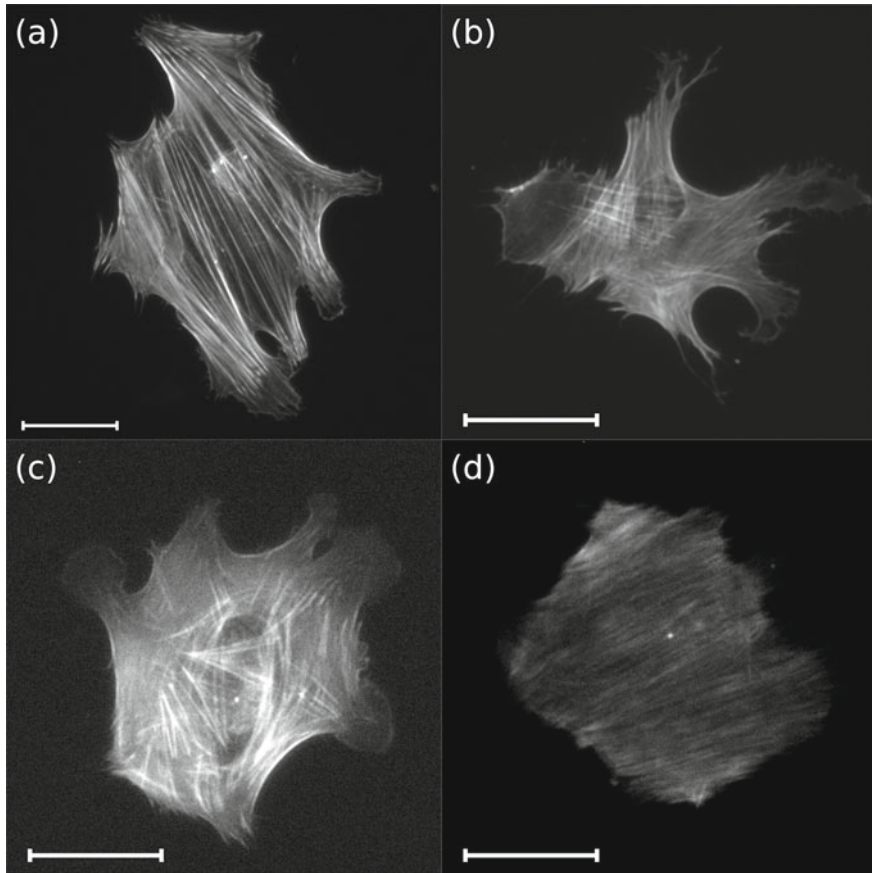
Fig. 10.1 Varying quality (top fixed, bottom live) filament expressions of fluorescence micrographs of human mesenchymal stem cells with scale bars at 50 μm. Subfigures: **a** *good quality* image of a fluorescently stained fixed cell of large size with clearly visible stress fibers on a substrate with a Young's modulus $E = 10$ kPa; **b** *medium quality* image of a fluorescently stained fixed cell of moderate size with inhomogeneous brightness and slight blur on glass; **c** *poor quality* image of a live cell of moderate size with considerable noise and excessive brightness due to overexposure on glass; **d** *very poor quality* image of a live cell of moderate size with very low contrast due to bleaching, considerable blur and hardly discernible stress fibers

filaments of varying widths, lengths and angles. We describe a novel algorithm able to detect slightly bent filaments. Since the FS is modular and open source, it can be easily extended to suit related image analysis tasks. Guaranteeing unbiasedness, tunable parameters can be learned, among others, from the *benchmark data set* (BDS) in [8] that has been manually labeled by specialists.

From the binarized filament structure various morphological descriptors can be extracted. Simple summary statistics are (weighted) mean orientation, area, aspect ratios of principal components (PCs). More subtle is the *order parameter* derived

from the angle between extrinsic mean orientation and first PC, quantifying the anisotropy of the acto-myosin cytoskeleton, and [4] linked statistically significant substrate elasticity to stem cell differentiation. A more sophisticated analysis [9] links number and persistence (under smoothing) of modes (i.e. dominating directions) of the distribution of weighted filament orientations to substrate elasticity, and, in particular, this requires development of a causal circular scale space theory in [9].

The above descriptors coarsely describe common filament orientations. In a finer approach, the concept of different single *orientation fields* (OFs) described in Sect. 10.4, pays credit to the tendency of filament orientations to change in a spatially smooth way. In order to simultaneously analyze statistically different moments of their distribution (e.g. jointly the intrinsic mean on a first geodesics PC together with the first geodesic PC), *backward nested descriptor* (BND) analysis is applied in Sect. 10.4.2 along with its asymptotic theory from [10]. This allows to elucidate fundamental differences between fixed and live cell analysis, with consequences for experimental design.

In Sect. 10.5 we conclude with an outlook how our biostatistical toolbox can be used, in various combinations, to tackle problems that have arisen through this research and problems currently high interest, for example tracking of individual filament dynamics and defining and analyzing corresponding descriptors such as filament life times.

10.2   Live Cell Imaging-Opportunities and Challenges

As mentioned above, novel insights into mechanisms of the complex mechanical interplay between cells and the extracellular matrix require the analysis of the dynamics of stress fiber formation and arrangement. Such an experimental approach differs fundamentally from fluorescence microscopy of fixed cells. In experiments using chemically fixed cells, these can be stained with a variety of fluorescent dyes using either antibodies or other small molecules that selectively bind to the protein of choice. Living cells however need to be modified genetically to express a fluorescent protein fused to the protein of interest or a respective binding partner. While both methods allow for fluorescently labeled cellular structures a significant difference is the homogeneity of the fluorescent intensity: fixed cells will overall show similar intensities that depend on the staining method and the cellular concentration of the protein of choice; living cells that are transiently expressing the fluorescent marker will show a broad distribution of intensities (and even dark cells that do not express any fluorescent marker) due to the stochastic nature of the transfection process. In addition, it is essential to monitor many ($N > 30$) cells in parallel leading to an enormous amount of microscopic images. These issues pose distinct challenges to the image segmentation and analysis algorithm that could not be resolved with our traditional approach of a pixel-based eLoG orientation analysis [4] mainly due the varying image quality (as illustrated e.g. in Fig. 10.1) and also computational time.

In the case of fixed cells, we are using a rigorous protocol to ensure the unbiased microscopic analysis of single cells. Due to the intrinsic variation of cellular morphologies it is of paramount importance to exclude any human bias. First, cells are searched in the fluorescence channel of the nucleus dye to find an isolated cells, whose nucleus looks normal and has no direct neighbors. Here, it is critical to exclude deformed nuclei, nuclei with a doubled set of DNA (high intensity), or any nuclei that are within the division process. Next, the fluorescent channel of interest (e.g. actin, non-muscle myosin IIa) is recorded regardless of the morphology, except for cases where now neighboring cells are observed that might interfere.

In contrast to the above described protocol for fixed cells, the situation is more subtle for parallel live cell imaging. Firstly, cells need to be transfected with a fluorescent marker that tags the protein of choice (in our case actin). This can be done using a fluorescent fusion protein (e.g. GFP-actin) with the immediate drawback of over-expression of that protein, differences in assembly kinetics, and potential problems with incorporation in distinct actin structures [11]. Most of these issues can be avoided using LifeAct [12], a short amino-acid sequence that binds to actin and is fused with a fluorescent protein. However, direct comparison also leads to minor differences between this visualization and staining later fixed cells with a phalloidin dye [11].

To avoid influence from neighboring untransfected cells, it is advisable to also record always an image in phase contrast or brightfield mode. That way any unwanted additional interactions can be ruled out. However, during time lapse microscopy recordings several incidents can occur that might affect the statistical analysis of the cell population. Cells might migrate out of the field-of-view, an aspect that we will address by smart repositioning the sample with real-time analysis of the microscopic pictures. Therefore, the full recording will lack a subset of very motile cells. Cells might undergo apoptosis and exclude themselves from further analysis reducing the number of samples. Cells might divide and therefore make the analysis of their cytoskeletal dynamics very complicated, even precluding it in case of not thoroughly separating cells. Cells might interact during the time period with neighboring cells that will also affect their cell-matrix mechanical interplay and acto-myosin structure. Altogether, it becomes clear that the population subsets of microscopy analysis of fixed cells and living cells can differ significantly and appropriate measures and controls need to be developed to fully understand its impact on the statistical analysis.

10.3   Automated Unbiased Binarization of Filament Structure

The present section is heavily based on the authors' previous publication [8], which is published under an open license (CC BY 4.0). The text and contents from said publication is reproduced here to achieve a self-contained description of the FilamentSensor.

10.3.1   Related Work

There is an impressive body of techniques for image processing and in particular, for line detection (for an overview e.g. [13], Chap. 4). Previous to our development of the FS, however, methods for the detection of filaments in cell images were often ad hoc, required manual processing to a considerable extent and were computationally rather time consuming, e.g. [4, 14, 15]. The latter two issues are particularly unfavorable considering the large number of images to be processed in live cell imaging.

Moreover, there is a large number of algorithms focusing on analyzing networks of strongly curved microtubuli (this property is not shared by single filaments), such as line thinning by [16], active contours by [17] and the constrained inverse diffusion (CID) method by [14]. These methods, however, detect only a skeletal filament network structure, they leave out filament orientation, length, and width. They aim at identifying thin microfilaments and not wide stress fibers as we are interested in.

There are methods which aim at extraction not only of filament pixel position but also of local orientation such as the FiberScore algorithm by [18], elongated Laplacians of Gaussians (eLoGs) by [4] and gradient based methods, e.g. [15, 19].

The eLoG method, like the gradient method is geared towards the detection not only of filament pixels but also of their orientations. Although filament width and length are not extracted by these methods, counting the number of pixels per orientation, they yield histograms of cumulated filament length per orientation angle and these histograms are then further analyzed [4, 15].

Local orientation and centerline images are produced by the FiberScore program [18] which provides global information on accumulated line length and average width. Line objects are not produced, however. For our cell images, we tried out the methods applied in FiberScore, but did not yield optimal results [8]. A fundamental drawback for applying FiberScore, however, is that neither the program's nor its framework's source codes are freely available. Even though the original developer has been very helpful and supportive to make the program run, FiberScore could not be tailored to our needs.

With the FS we have developed an image processing tool that returns stress fiber structures from live cell images, as well as from fixed cells images applicable to the use case where images vary widely in brightness, contrast, sharpness and homogeneity of fluorescence, cf. Fig. 10.1. Typically, in our setup of live cell imaging, we observe 30 cells over a period of 24 h, taking an image every 10 min. As we aim at real time processing, this leaves about 20 s process time per image.

The FS thus developed can be used to binarize filament structures for any (sets of) images containing fiber features. Applications in a wide range of use cases come to mind, in particular in the context of actin fiber structures, e.g. [20, 21], but also for more general contexts in medical imaging, biology, and in the material science. As the FS is modular and easily extensible, several authors, including [22, 23], have built on it after it was first published.

Notably, there is also a rising demand to address the task of tracing and tracking stress fibers, both over space and over time. We mention studies on migrating cells

which display a variety of stress fiber types (dorsal, ventral, arc) that appear at different loci inside a migrating cell [21, 24–28]. Their exact cellular function is still in the dark, it could be clarified, however, using live cell filament digitization. Indeed, when stress fiber dynamics are followed over time, this may give further insight into formation and function of filament structure. A novel method to analyze traction force microscopy data, so called model-based traction force microscopy has been recently described by [29]. In this context, it is necessary to detect and mark the stress fibers in a cell in order to link forces to fiber location and develop deeper insight into cellular force generation and transmission to the substrate. As mentioned before, ideally, corresponding live cell experiments are performed simultaneously for many cells in order to arrive at statistics that are sufficiently significant. This requires that algorithms for fiber analysis perform tracing and tracking in (nearly) real-time, ideally.

10.3.2   The FilamentSensor and the Benchmark Dataset

To obtain the full information of the stress fibers in cells, namely location, length, width, and orientation, from repeated observations of living cells under widely varying conditions in near real time the FS has to extract

  (I)   fast and unsupervised
 (II)   robustly
(III)   all filament features: location, length, width and orientation;

where (II) implies dealing with several specific problems illustrated in Fig. 10.2

(IIa)   detecting darker lines crossing bright lines,
(IIb)   dealing with image inhomogeneities and
(IIc)   dealing with image blur and noise.

The FS is specifically designed to meet these challenges. Dealing with image inhomogeneity calls for the application of local image processing tools. Blurring effects will be mitigated by line enhancement through direction sensitive methods. Crossings of lines of varying intensities can be successfully detected by what we call *line Gaussians* which utilize oriented thin masks. After local binarization, an adaption of the semilocal line sensor approach to fingerprint analysis [30] is applied to extract all filament features. As the FS is modularized, employs local and orientation dependent image analysis methods and outputs the entire filament data, expert knowledge such as detecting fewer filaments in specific low variance areas, say, can be easily incorporated.
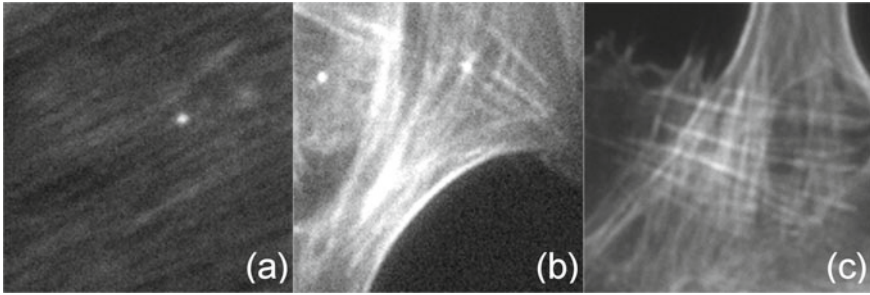
Fig. 10.2 Challenges for filament extraction. a blur (detail from Fig. 10.1d). The overall contrast of the cell body is very low and lines are hardly discernible. **b** overexposure and noise (Fig. 10.1c). The extensive regions of maximal brightness hide any structure that may be present in those regions. Salt and pepper noise is clearly visible as dark spots in bright areas and bright spots in dark areas. **c** filament crossings (Fig. 10.1b). A bundle of roughly vertical filaments of varying brightness crosses a bundle of roughly horizontal filaments with varying brightness

10.3.3 Detecting Slightly Bent Filaments

After preprocessing and binarization, as described in [8], filament data is extracted from the white pixels of the binarized image. Visual inspection of fluorescence microscopic images reveals that actin stress fibers can be slightly bent. To take this into account, we have adjusted the FS to follow slightly curved lines on a piecewise linear path. Line detection is performed by the following algorithm.

1. Every white pixel (x, y) is assigned a *width*, $W(x, y)$. This is done by taking circular pixel neighborhoods of the pixel (cf. Fig. 10.3) with increasing diameter. A diameter is accepted, if the ratio of white pixels of the binary image is above an adjustable tolerance (default 95%). If a diameter was accepted, the next larger diameter is proposed until a diameter is not accepted. The width $W(x, y)$ at the pixel is then given by the largest accepted diameter at the pixel. In particular, this gives a range of widths $1 \leq w_1 < \ldots < w_k = \max W(x, y)$ attained by pixels.
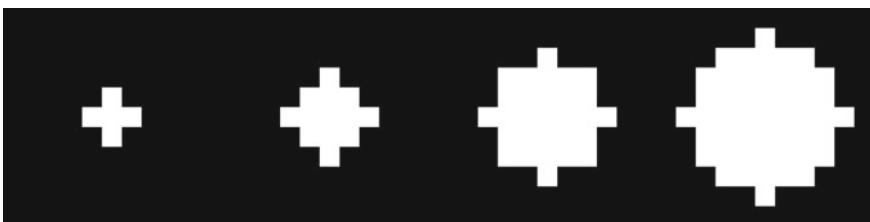

Fig. 10.3 Some circle masks. These are examples of the circular masks used by the segment sensor algorithm to determine line width. The circles displayed here correspond to diameters of 2, 4, 6 and 8 pixels. The masks are squares with an odd number of pixels as they are centered at a unique pixel

A temporary List L, the *filament data set* $\mathcal{F}$, and the *orientation field* $\mathcal{O}$ are each initialized by the empty set.

2. For every white pixel, starting with the highest width value and continuing with decreasing width value, apply the *CurveTracer (CT)*. The curves are represented piecewise linearly. The user specifies four parameters, namely the length of linear pieces l_{lin}, the direction step size in degrees ϕ_{step}, the maximal number of angle steps between two adjacent linear pieces n_{step}, leading to maximal angle of $\phi_{\text{step}} \cdot n_{\text{step}}$ between adjacent line pieces, and a minimal line length l_{min}.

   a. For each white pixel (x, y) the CT probes into a number of directions (by default $\phi_{\text{step}} = 3°$; this corresponds to 60 orientations $3°, 6°, \ldots, 360°$). For each direction the CT follows a straight line from (x, y) for a number of $2l_{\text{lin}}$ pixels. For each of these directions, the average width value is calculated and two almost opposite directions (with a relative angle in $[180° - n_{\text{step}}\phi_{\text{step}}, 180° + n_{\text{step}}\phi_{\text{step}}]$) with the largest combined average width are selected. For each of these two directions the CT now proceeds separately as follows, using a point list P containing the starting point.

      i. Move l_{lin} pixels in the current direction ϕ_c and add the end point p to P. If the average width along this line is below 1, remove all pixels with width 0 from the end of the line and then proceed removing pixels (possibly with width greater than 0) from the end until the average width is at least 1. Then add the new final point p of the line to P and stop.

      ii. From p probe for $2l_{\text{lin}}$ pixels into the $2n_{\text{step}} + 1$ directions $\{\phi_c - n_{\text{step}}\phi_{\text{step}}, \phi_c - (n_{\text{step}} - 1)\phi_{\text{step}}, \ldots, \phi_c + n_{\text{step}}\phi_{\text{step}}\}$ and calculate average width values for every direction. Set the new ϕ_c to the direction of highest average width.

      iii. Return to step 1.

   b. When the CT searches in both directions have reached their end points, the combined length of the line pieces is determined and if it is larger than l_{min} the list of points from both pieces is stored to L. The CT is illustrated in Fig. 10.4.

   c. In the next step, segments in L are called in the order of their length, long segments first. For every segment, the orientation field $\mathcal{O}$ (which is empty when first called) is looked up for every pixel on the segment. If less than 30 % of the segment's pixels have a *conflicting orientation* entry in $\mathcal{O}$,—i.e. the entry in $\mathcal{O}$ differs by less than an adjustable tolerance angle (per default 20°) from the segment's orientation—the segment is accepted as valid. For every pixel within a circular neighborhood with diameter $w_j + 2$ pixels (in order to avoid duplicate lines in case the CT does not perfectly follow the path of maximal width) of a segment pixel, the segment's orientation is stored to $\mathcal{O}$ overwriting possible previous entries. The segment is then also added to $\mathcal{F}$. If at least 30 % of the pixels on a segment have a conflicting orientation, we have the following cases.

      i. If $\mathcal{O}$ does not carry a conflicting orientation for any of the endpoints, the segment is discarded.
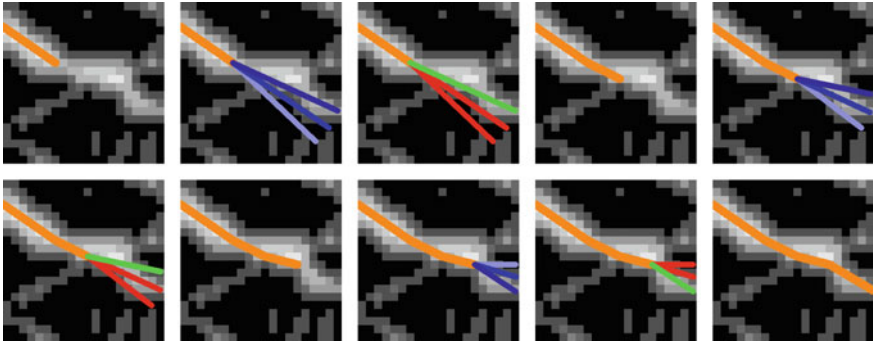
Fig. 10.4 Illustration of the CT algorithm. The probing distance is twice as far as the step length of the CT. Per default, the CT probes three directions at each step where the number of directions as well as the angle between them can be adjusted by the user. The probing directions are visualized in three shades of blue. The chosen direction is then marked green and the others are marked red

ii. Otherwise, the endpoints with conflicting orientations are iteratively removed from the segment until the remaining segment's endpoints no longer have a conflicting orientation. If the resultant segment length is above the threshold of minimal filament length, this new segment is added back to L and the original one is removed. The new segment is revisited when its length is called.

As lines are blurred due to scattering and as the preprocessing usually enhances line width, the FS tends to find greater line width than a human expert (cf. benchmark data set in [8]).

10.4 Orientation Fields

We make use of the individual line data provided by the FS to identify local orientation fields of fibers. These provide a more detailed picture of the cytoskeleton than simple summary statistics or orientation histograms, since they take local features into account. Orientation fields are contiguous regions in the cell filled with filaments of similar orientation. The local orientation of the field may change slowly over the cell, so as to encompass the case of a curved cell, where stress fibers follow the cell shape. As input we use an *orientation map* denoting the line orientations determined by the filament sensor at each pixel. Our algorithm uses relaxation labeling, first described by [31].

As a first step, the image is covered with a rectangular equidistant grid of pixels with spacing a. The grid points define *blocks* in the following. The minimal spacing of the grid is 5 pixels. While the number of blocks b covering the cell area in the image exceeds 500, the spacing is successively increased by steps of two pixels. As

a result, the maximal number of non-trivial blocks is 500. In this way, the algorithm can deal with images of different magnification and cells of different size.

For every block, we use a square isotropic Gaussian mask with $\sigma = \max(15, \lfloor 1.2a \rfloor)$ centered at the grid point to assign weights to nearby pixels. The length of the mask from the center is $l = \lfloor 2.5\sigma \rfloor$ pixels in every direction, after which it is truncated. Using the orientation map and these weights, we get an orientation histogram for each block. This orientation histogram is then smoothed with a wrapped Gaussian kernel with $\sigma = 6°$. Of the smoothed histogram, all maxima are stored as local orientations of the block. If several neighboring bins of the smoothed histogram have the same, locally maximal, value, the leftmost bin, corresponding to the smallest angle, is used.

To perform relaxation labeling, it is necessary to have a number of blocks with fixed orientations that can be used as a seed. In order to achieve reliable results, it is desirable to have as many seed blocks as possible. Therefore, we try to determine at least $s = 0.05b$ seed blocks. For this we collect all blocks with only one orientation and apply the following *cleanup procedure*:

1. Keep only the largest contiguous region.
2. Make a histogram of the block orientation smoothed with a Gaussian kernel with $\sigma = 6°$. And determine the global maximum $\phi_{\max}$.
3. Starting with $k = 0$, we determine the largest contiguous region of blocks whose orientation ϕ satisfies $\phi - \phi_{\max} \leq k$. We then increase k by steps of 1, until we reach $k = 6$ or until the largest contiguous region of blocks reaches or exceeds the number of $s = 0.05b$ blocks.

If the number of seed blocks is smaller than s, we repeat the cleanup procedure for all blocks using all orientations of every block in step 2.

Once we have a set of seed blocks with seed orientations, we fix these orientations and perform a relaxation labeling over all orientations of all non-seed blocks. For the relaxation labeling, we use a von Mises type *compatibility function*

$$f(\phi) = C + B \exp(A \cos(2\phi))$$
$$\text{with} \quad f(0) = 1, \quad f(90) = -1, \quad f(\sigma) = 0$$

where we start with $\sigma = 15°$. If the largest field contains less than 2/3 of line pixels or less than 85% of the blocks, we repeat the relaxation labeling, increasing σ by steps of 5° to a maximum of $\sigma = 25°$.

To preclude too large changes of orientation at medium range, we do not use only nearest neighbors for the relaxation, but every block reacts with isotropic Gaussian weights with surrounding blocks, where the standard deviation of the Gaussian is $\Sigma = 5$ blocks. Every block also has a dummy orientation, whose probability will slowly grow, if none of the block orientations match their neighborhood. Blocks neighboring on seed blocks and having reached a probability of 0.999 on one of their orientations will be turned into seed blocks, so the field will gradually "freeze" to ensure convergence.
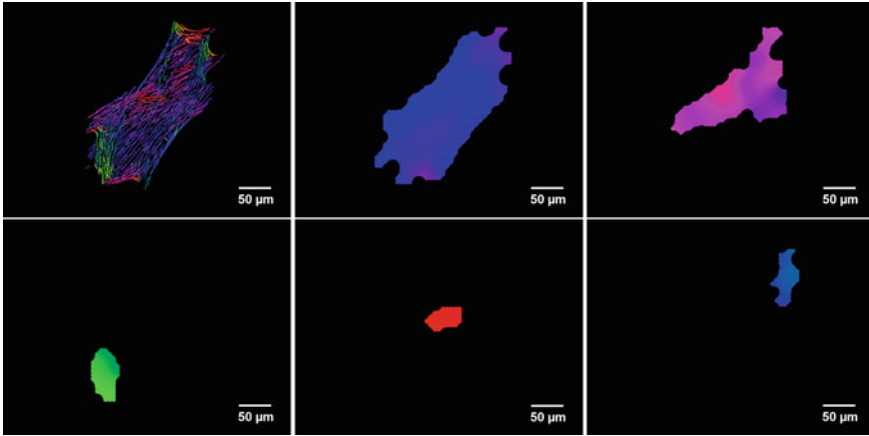
Fig. 10.5 Orientation field detection. Lines segments extracted by the FS (upper left panel) colored by local orientation, giving five dominating orientation fields (their range and local orientations displayed in the other panels). Indeed, colors of the fields, indicating their local orientation, vary only slightly

When the relaxation has converged, the resulting field is saved and its corresponding orientations are removed from the blocks. If the remaining nontrivial blocks do not form a contiguous region, all regions below minimal size s are removed. The procedure is then repeated until no non-trivial blocks remain. Finally, every filament is sorted into the orientation field whose local orientation best matches its own orientation (Fig. 10.5). If the filament orientation diverges by more than 15° from all local orientation fields, it is not associated to any field.

10.4.1 Orientation Field Evolution

In most cells on substrates with stiffness 10 and 30 kPa a single orientation field emerges over 24 h, which contains more than 80% of stress fiber length. In order to illustrate the evolution of orientation fields in time, we represent the orientation field at each time point by a gray circle, whose gray level displays the relative amount of fiber length represented by that field, such that the circle is black when all fibers are included in the field. The standard deviation of fiber orientations in the field are displayed by error bars for each circle. The evolution of orientation fields for a typical cell on an intermediate or stiff gel is displayed in Fig. 10.6.

In Fig. 10.7 we show a typical orientation field evolution for a cell on a soft substrate with stiffness 1 kPa. In cells on such soft substrates the cytoskeleton is much less ordered which is reflected by a large number of small orientation fields which are found over time. In a cell where fibers are not ordered, orientation fields
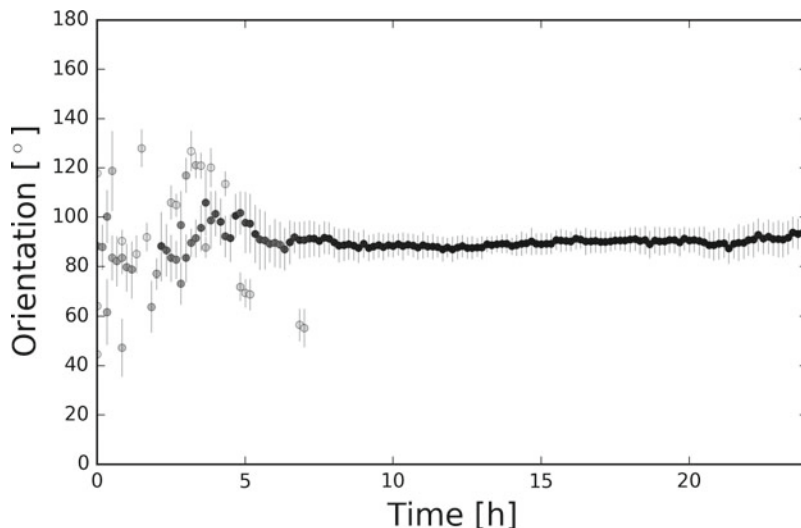
Fig. 10.6 **Orientation field time series for a typical cell on a substrate with stiffness** ≥ 10 kPa, at each time point represented by a gray circle, whose gray level displays the relative amount of fiber length represented by that field (the circle is black when all fibers are included in the field). The standard deviation of fiber orientations in the field are displayed by error bars for each circle. As is typical for a cell on a stiff substrate, a single main orientation field emerges after a few hours and remains stable throughout imaging time

often appear only for few images. While such fields can be considered spurious, they still serve to illustrate the disorder of the actin cytoskeleton.

In many cells on all gels the cytoskeleton is not fully described by just one orientation field but is partially ordered. A frequently observed evolution starts out with an almost chaotic cytoskeleton where the short lived small orientation field converge into a main orientation field over time, as in Figs. 10.6 and 10.7. This process can take between 4 and 20 h and can even be unfinished at the end of the 24 h observation span. However, there are also cases, where a main orientation field, which has remained stable for many hours suddenly disperses as the cell starts to move (corresponding to Figs. 10.6 and 10.7 with inverted time axis).

A behavior, which is observed in less than 5 % of cells, is illustrated in Fig. 10.8. In this case, a stable main orientation field exists, when at some point in time a new orientation field begins to form and the original main field starts to dissolve. This behavior requires more thorough investigation, both into the underlying cell dynamics and into adequate statistical representation of cytoskeleton order. Elucidating this behavior is left for future research.
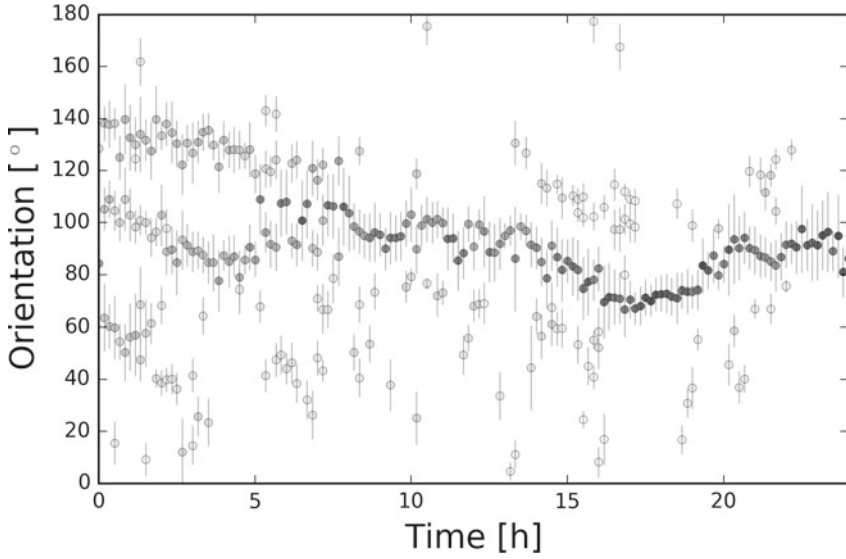
Fig. 10.7 Orientation field time series for a typical cell on a substrate with stiffness 1 kPa (notation as in Fig. 10.6) with several local orientation fields indicating a less ordered cytoskeleton. While over time a dominating orientation field emerges, smaller orientation fields pop out until the end
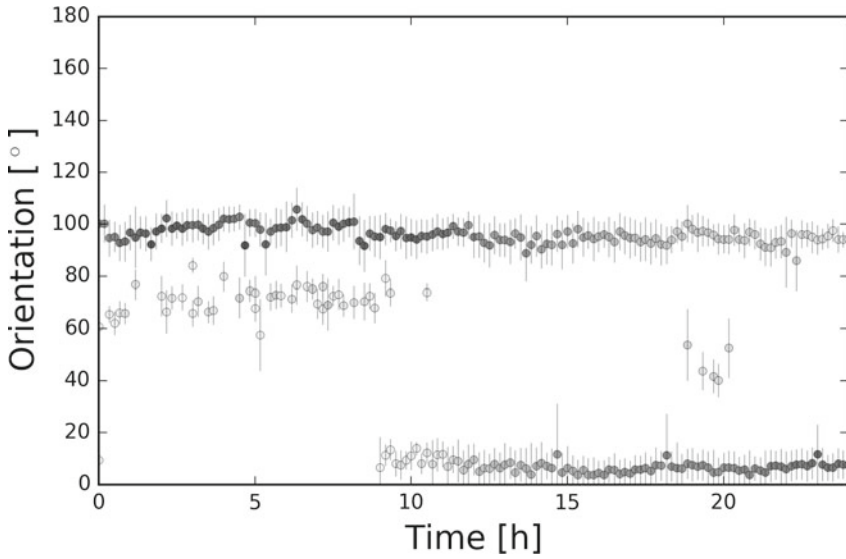

Fig. 10.8 Orientation field time series for a cell on a substrate with stiffness 10 kPa (notation as in Fig. 10.6). The initial main orientation field starts to decay after 10 h and is superseded by a new orientation field which is almost perpendicular to it

10.4.2   Backward Nested Descriptor Analysis

In order to use orientation fields for quantitative analysis, we devised the simple low dimensional orientation fields representation (10.1) in [10]. Denote by M the number of all filament pixels in a cell image, m_1 denotes the number of filament pixels in the largest orientation field and m_2 the number of pixels in all smaller fields combined. $M - m_1 - m_2$ then enumerates the pixels from "rogue" filaments which do not fit into any of the orientation fields due to strongly deviating orientation. In order to compare relative diameters rather than relative areas we observe the quantity

$$x = (x_1, x_2, x_3) := (\sqrt{m_1/M}, \sqrt{m_2/M}, \sqrt{1 - (m_1 + m_2)/M})^T \in S^2 \tag{10.1}$$

on a two-sphere. In this representation, the spherical data lie in the first octant. We observe that points tend to accumulate close to the $x_2 = 0$ plane, representing cells with one single orientation field, cf. Fig. 10.9.

In order to interpret live cell observations, we compare these observations to time series of cells which were fixed after different times on a gel. We have taken images of fixed cells for each gel rigidity in intervals of 4 h of time on the gel. The sample sizes are displayed in Table 10.1. These cells were stained with phalloidin as opposed to the live cells, which were transfected with liveAct. To compare the live cell experiment to the fixed cell experiment, we only consider images from live cell movies corresponding to the fixation times. Since we have between 50 and 60 movies on each gel, we can expect a higher data variance for the live cells compared to fixed cells. Some of the investigated samples are displayed in Fig. 10.9.

We analyze the samples on S^2 by applying dimension reduction via *principal nested great spheres* from [32]. This means we first identify the great circle which fits the data best (in terms of accumulated squared spherical distance), then orthogonally (along great circles) project data to this great circle and determine their Fréchet mean on this great circle, called the *nested mean*. Jointly, the two give our *backward nested data descriptor*. To estimate its variance, i.e. the variance of the great circle and the nested mean, we use $B = 1000$ bootstrap replicates from the data. Figure 10.10 displays the backward nested descriptor and bootstrapped means illustrating the spread of the nested mean estimator.

The lower sample size leads to considerably higher variance for live cells, as expected. However, fixed and live cells on the same substrates are strikingly different from each other.

While live cells on the soft ($E = 1$ kPa) gels show little temporal evolution, the fixed cells exhibit a development towards a dominant main field with few smaller fields and eventually less rogue filaments. For the cells on stiffer gels, while the fixed cells strengthen their main field, mainly at the expense of rogue filaments, for live cells this effect is stronger, leading to fewer rogue filaments over time as well. Remarkably, for the fixed cells we observe an increased number of rogue filaments after 16 and 20 h, which does not occur for live cells. A T^2 hypothesis test developed in [10] for backward nested descriptors verifies that this effect is significant.
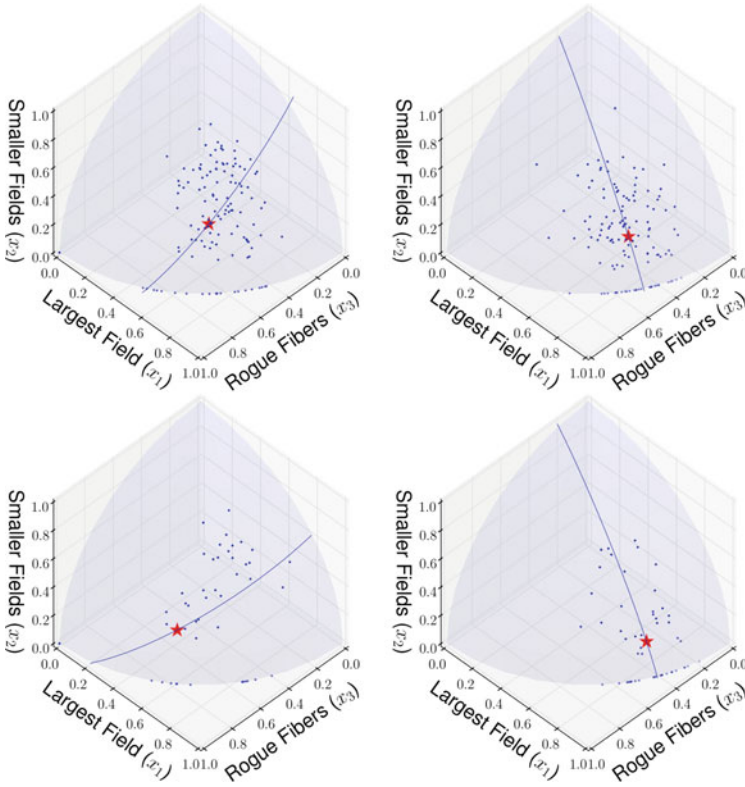
Fig. 10.9 Visualization of the orientation fields representatives on S^2 *(blue points) from* (10.1) with backward nested data descriptor given by the best approximating great circle (blue line) with nested mean (red star). All images correspond to 16 h on the gel. Upper left: 135 fixed cells on a 1 kPa gel; upper right: 127 fixed cells on a 10 kPa gel; lower left: 59 live cells on a 1 kPa gel; lover right: 53 live cells on a 30 kPa gel

Table 10.1 Sample sizes of hMSC skeleton images over varying Young's moduli and cultivation time, left for fixed cells, right for live cells

Time	1 kPa	10 kPa	30 kPa
4h	159	168	153
8h	163	164	153
12h	176	171	173
16h	135	127	147
20h	138	126	127
24h	166	152	152

1 kPa	10 kPa	30 kPa
59	54	53
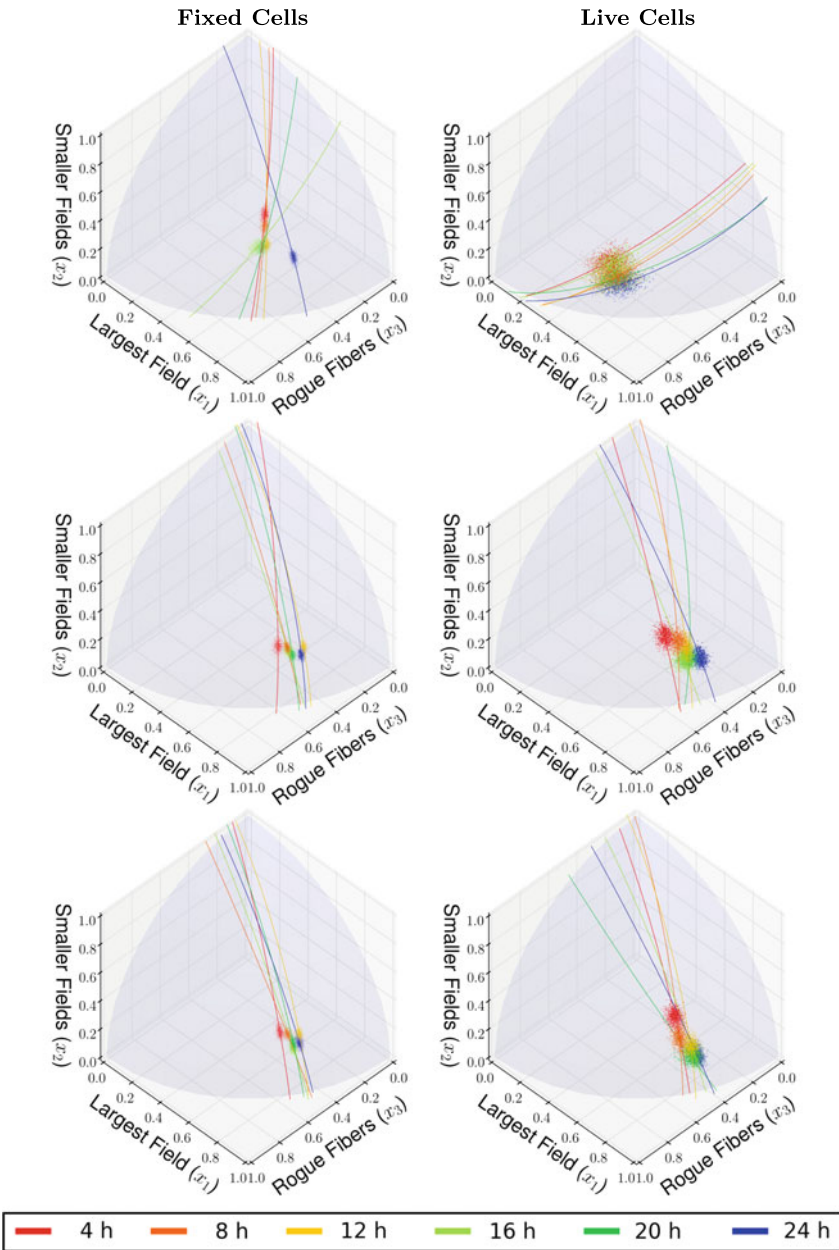
Fig. 10.10   Best fitting great circles with nested means. For every time and sample, nested means of $B = 1000$ bootstrap samples are displayed to illustrate variance of the mean. Rows from top to bottom: 1 kPa gel, 10 kPa gel, 30 kPa gel

Upon closer inspection, as noted in [10], describing cells by their orientation field decomposition, the temporal evolution of fixed cells comes to a stop and at roughly 12 h, often reversed then, hinting to an increased rate of cell division, albeit cells near the division process have been singled out previously, cf. Sect. 10.2. Although these cells are not intentionally synchronized, an increased rate of cell division after a particular time after seeding is not surprising as trypsinization and re-seeding slightly decreases the isotropic temporal distribution on the cell cycle. For live cells, such an effect is not observed, since all movies during which the cells divide are left out. The statistical analysis confirms our initial hypothesis that a direct comparison of results from live and fixed cells is complicated due to the different population subsets and points out the importance of careful experiment design including proper controls for future studies.

10.5   Outlook

In this chapter we have briefly illustrated the statistical biophysical toolbox developed over 8 years in project B8 of the SFB 755, the support of which we gratefully acknowledge, and we have applied a typical set of descriptors (mean great circle and nested mean on it), exemplary to highlight differences in the actin-myosin cytoskeleton structure of live and fixed cells. Future application include tracing and tracking stress fibers over space and time and usage in many other demanding research areas. For example, studies on migrating cells indicate various stress fiber types classically described as (dorsal, ventral, and arcs) appearing at different locations inside a migrating cell [21, 24–28]. Following the filament dynamics over time will give further insight into the formation and function of stress fibers. Using our toolbox applied to live cell imaging, it seems promising that we can come to an unbiased statistical classification of the cytoskeleton that relates temporal and spatial persistence to function. Recently, [29] described a novel method to analyze traction force microscopy data, so called model-based traction force microscopy. Here, it is imperative to detect and mark the stress fibers of a cell in order to link forces to fiber location and gain more insight into cellular force generation and transmission to the substrate.

References

1. Discher, D.E., Janmey, P., Wang, Y.L.: Tissue cells feel and respond to the stiffness of their substrate. Science **310**(5751), 1139–1143 (2005)
2. Rehfeldt, F., Engler, A.J., Eckhardt, A., Ahmed, F., Discher, D.E.: Cell responses to the mechanochemical microenvironment-implications for regenerative medicine and drug delivery. Adv. Drug Deliv. Rev. **59**(13), 1329–1339 (2007)
3. Engler, A.J., Sen, S., Sweeney, H.L., Discher, D.E.: Matrix elasticity directs stem cell lineage specification. Cell **126**(4), 677–689 (2006)
4. Zemel, A., Rehfeldt, F., Brown, A.E.X., Discher, D.E., Safran, S.A.: Optimal matrix rigidity for stress-fibre polarization in stem cells. Nat Phys **6**(6), 468–473 (2010)
5. Wen, J.H., Vincent, L.G., Fuhrmann, A., Choi, Y.S., Hribar, K.C., Taylor-Weiner, H., Chen, S., Engler, A.J.: Interplay of matrix stiffness and protein tethering in stem cell differentiation. Nat. Mater. **13**(10), 979–987 (2014)
6. Swift, J., Ivanovska, I.L., Buxboim, A., Harada, T., Dingal, P.C.D.P., Pinter, J., Pajerowski, J.D., Spinler, K.R., Shin, J.W., Tewari, M., Rehfeldt, F., Speicher, D.W., Discher, D.E.: Nuclear lamin-a scales with tissue stiffness and enhances matrix-directed differentiation. Science **341**(6149) (2013)
7. Zemel, A., Rehfeldt, F., Brown, A.E.X., Discher, D.E., Safran, S.A.: Cell shape, spreading symmetry, and the polarization of stress-fibers in cells. J. Phys. Condens. Matter **22**(19), 194,110 (2010)
8. Eltzner, B., Wollnik, C., Gottschlich, C., Huckemann, S., Rehfeldt, F.: The filament sensor for near real-time detection of cytoskeletal fiber structures. PloS one **10**(5), e0126,346 (2015)
9. Huckemann, S., Kim, K.R., Munk, A., Rehfeldt, F., Sommerfeld, M., Weickert, J., Wollnik, C.: The circular sizer, inferred persistence of shape parameters and application to early stem cell differentiation. Bernoulli **22**(4), 2113–2142 (2016)
10. Huckemann, S.F., Eltzner, B.: Backward nested descriptors asymptotics with inference on stem cell differentiation. Ann. Stat. **46**(5), 1994–2019 (2018). ArXiv:1609.00814
11. Belin, B.J., Goins, L.M., Mullins, R.D.: Comparative analysis of tools for live cell imaging of actin network architecture. BioArchitecture **4**(6), 189–202 (2014)
12. Riedl, J., Crevenna, A.H., Kessenbrock, K., Yu, J.H., Neukirchen, D., Bista, M., Bradke, F., Jenne, D., Holak, T.A., Werb, Z., Sixt, M., Wedlich-Soldner, R.: Lifeact: a versatile marker to visualize f-actin. Nat. Methods **5**, 605 (2008)
13. Szeliski, R.: Computer vision: algorithms and applications. Springer (2010)
14. Basu, S., Dahl, K., Rohde, G.: Localizing and extracting filament distributions from microscopy images. J. Microsc. **250**, 57–67 (2013). Retracted due to image copyright issues
15. Faust, U., Hampe, N., Rubner, W., Kirchgeßner, N., Safran, S., Hoffmann, B., Merkel, R.: Cyclic stress at mHz frequencies aligns fibroblasts in direction of zero strain. PLoS ONE **6**(12), e28963 (2011)
16. Chang, S., Kulikowski, C., Dunn, S., Levy, S.: Biomedical image skeletonization: a novel method applied to fibrin network structures. Medinfo **84**(2), 901–906 (2001)
17. Dormann, D., Libotte, T., Weijer, C.J., Bretschneider, T.: Simultaneous quantification of cell motility and protein-membrane-association using active contours. Cell Motil Cytoskelet. **52**(4), 221–230 (2002)
18. Lichtenstein, N., Geiger, B., Kam, Z.: Quantitative analysis of cytoskeletal organization by digital fluorescent microscopy. Cytometry A **54**(1), 8–18 (2003)
19. Herberich, G., Würflinger, T., Sechi, A., Windoffer, R., Leube, R., Aach, T.: Fluorescence microscopic imaging and image analysis of the cytoskeleton. In: Conference record of the forty fourth Asilomar conference on signals, systems and computers (ASILOMAR), pp. 1359–1363 (2010)
20. Sanchez, T., Chen, D.T.N., DeCamp, S.J., Heymann, M., Dogic, Z.: Spontaneous motion in hierarchically assembled active matter. Nature **491**(7424), 431–434 (2012)
21. Tojkander, S., Gateva, G., Lappalainen, P.: Actin stress fibers-assembly, dynamics and biological roles. J. Cell Sci. **125**(8), 1855–1864 (2012)

22. Alioscha-Perez, M., Benadiba, C., Goossens, K., Kasas, S., Dietler, G., Willaert, R., Sahli, H.: A robust actin filaments image analysis framework. PLOS Computational Biol. **12**(8), 1–23 (2016)
23. Rogge, H., Artelt, N., Endlich, N., Endlich, K.: Automated segmentation and quantification of actin stress fibres undergoing experimentally induced changes. J. Microsc. **268**(2), 129–140 (2017)
24. Ciobanasu, C., Faivre, B., Le Clainche, C.: Actin dynamics associated with focal adhesions. Int. J. Cell Biol. **2012**, 941,292 (2012)
25. Hotulainen, P., Lappalainen, P.: Stress fibers are generated by two distinct actin assembly mechanisms in motile cells. J. Cell Biol. **173**(3), 383–394 (2006)
26. Naumanen, P., Lappalainen, P., Hotulainen, P.: Mechanisms of actin stress fibre assembly. J. Microsc. **231**, 446–454 (2008)
27. Pellegrin, S., Mellor, H.: Actin stress fibres. J. Cell Sci. **120**, 3491–3499 (2007)
28. Vallenius, T.: Actin stress fibre subtypes in mesenchymal-migrating cells. Open Biol. **3**, 13,001 (2013)
29. Soiné, J.R.D., Brand, C.A., Stricker, J., Oakes, P.W., Gardel, M.L., Schwarz, U.S., Vikram, D.: Model-based traction force microscopy reveals differential tension in cellular actin bundles. PLoS Comput Biol **11**(3), e1004,076 (2015)
30. Gottschlich, C., Mihailescu, P., Munk, A.: Robust orientation field estimation and extrapolation using semilocal line sensors. IEEE Trans. Inf. Forensics Secur. **4**(4), 802–811 (2009). https://doi.org/10.1109/TIFS.2009.2033219
31. Rosenfeld, A., Hummel, R.A., Zucker, S.W.: Scene labeling by relaxation operations. IEEE Trans. Syst. Man Cybern. **6**(6), 420–433 (1976)
32. Jung, S., Dryden, I.L., Marron, J.S.: Analysis of principal nested spheres. Biometrika **99**(3), 551–568 (2012)