

## Detection of mitoses in embryonic epithelia using motion field analysis

Parthipan Siva<sup>a1</sup> G. Wayne Brodland<sup>b\*</sup> and David Clausi<sup>a2</sup>

<sup>a</sup>Department of Systems Design Engineering, University of Waterloo, Waterloo, Ont., Canada; <sup>b</sup>Department of Civil Engineering, University of Waterloo, Waterloo, Ont., Canada

(Received 23 May 2008; final version received 13 October 2008)

Although computer simulations indicate that mitosis may be important to the mechanics of morphogenetic movements, algorithms to identify mitoses in bright field images of embryonic epithelia have not previously been available. Here, the authors present an algorithm that identifies mitoses and their orientations based on the motion field between successive images. Within this motion field, the algorithm seeks ‘mitosis motion field prototypes’ characterised by convergent motion in one direction and divergent motion in the orthogonal direction, the local motions produced by the division process. The algorithm uses image processing, vector field analyses and pattern recognition to identify occurrences of this prototype and to determine its orientation. When applied to time-lapse images of gastrulation and neurulation-stage amphibian (*Ambystoma mexicanum*) embryos, the algorithm achieves identification accuracies of 68 and 67%, respectively and angular accuracies of the order of 30°, values sufficient to assess the role of mitosis in these developmental processes.

**Keywords:** mitosis identification; motion field analysis; embryogenesis; epithelia

### 1. Introduction

Coordinated tissue motions are a crucial component of early embryo development and factors that are known to affect these motions include gene expression patterns, force generation by sub-cellular structural components and embryo geometry (Chen and Brodland 2008). Computational models suggest that mitosis may also be an important factor (Brodland and Veldhuis 2002). However, to assess its possible role, the locations, frequencies and orientations of mitoses must be quantified in live embryos, a task that is impractical to carry out manually.

The present study focuses on two critical processes common to all vertebrates; namely, gastrulation and neurulation. Embryos of the axolotl (*Ambystoma mexicanum*), a type of amphibian (Figure 1), were chosen for the study because a great variety of mechanical data – including kinematic data (Brodland et al. 1995), strain maps (Veldhuis et al. 2005), tensile properties (Wiebe and Brodland 2005) and *in vivo* stresses (Benko and Brodland 2007) – have already been obtained for it. Thus, a developmental mechanics context exists for interpretation of findings.

Currently available image processing algorithms can identify mitoses under certain conditions. For example, if cells remain isolated from each other, as they often do in culture, individual cells can be identified by image segmentation and the topological changes associated with the separation of the daughter cells (Bertuccio et al. 1998; Chen et al. 1999; Refai et al. 2003; Shimada et al. 2005) used to identify mitoses. Alternatively, if adequate contrast

exists between the cell boundary and the interiors of cells, then cell boundary algorithms can be used (Vincent and Masters 1992; Talukder and Casasent 1998; Puddister 2003; Phukpattaranont and Boonyaphiphat 2006; Iles et al. 2007). Another approach is to stain the cell nuclei and identify mitotic figures manually or automatically by geometric changes in the nucleus (Kaman et al. 1984; Simpson et al. 1992; van Diest et al. 1992; Kate et al. 1993; Cross and Start 1996; Gal et al. 2005). Unfortunately, none of these approaches are applicable to early-stage axolotls because the cells in the monolayer epithelium that form their surface are confluent, and suitable membrane and nuclear stains are not currently available for axolotls at such early stages of development.

The objective of the present study is to develop a computer algorithm that will allow mitoses to be identified in bright field images of the embryo surface (Brodland and Veldhuis 1998). These mitoses can be identified manually in such images because they produce local cell-level tissue motions (Figure 2). Manual observations of many such mitoses revealed that these motions are characterised by two components: inward motion along the plane of the cell division, as at the points marked + in Figure 2, and outward motion along an axis through the centers of the newly-formed cells, as at the points marked o and \*. Interkinetic migration (Baye and Link 2007), a characteristic apical–basal motion of the nucleus associated with mitosis, appears not to be a significant driver of the observed motions, since it would produce isotropic motions. Instead, these motions, which are consistent

\*Corresponding author. Email: brodland@uwaterloo.ca

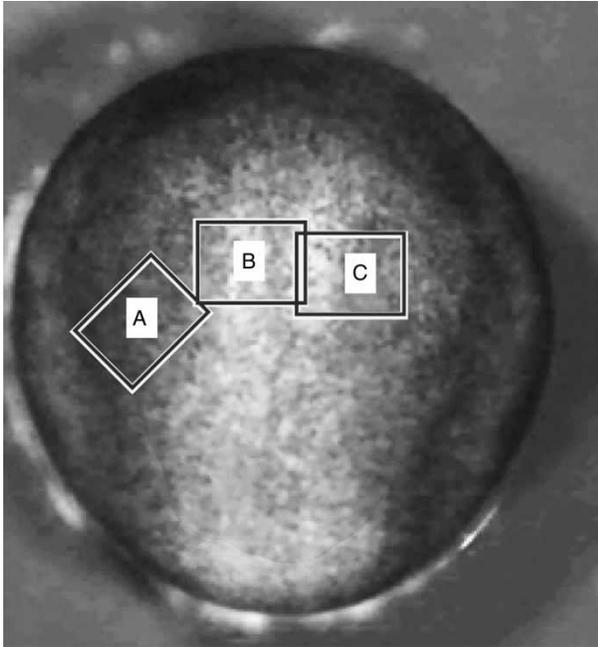


Figure 1. An early neurulation-stage axolotl embryo. In this dorsal view of a stage 13 embryo, the cephalic end of the embryo is toward the top of the image and the caudal end toward the bottom. The mid-sagittal plane lies along a vertical bisector of the image. Images were collected from the three-labelled regions.

with those predicted by 2D computer simulations (Brodland and Veldhuis 2002), are approximately normal to each other and of opposite sense. Mathematically, this pattern of motion (Figure 3) is known as a saddle point (Shu and Jain 1994) and herein this pattern of motion is referred to as the *mitosis motion field prototype* or simply *mitosis prototype*.

Automated detection of this pattern of motion in bright field images (Figure 2) is more difficult than one might expect and the solution we found required concepts from image processing, vector field analyses and pattern recognition (Siva 2007). It also required the use of mathematical function expansion and cut-off methods to improve discrimination of the algorithm. The resulting algorithm is able to determine the locations, frequencies and angular orientations of mitosis prototypes based on pairs of time lapse images. Such locations are deemed to be mitoses and the angle of the extensile component of the prototype to define the mitosis orientation.

The algorithms were tested on 15 time-lapse image sequences (Table 1) consisting of between 21 and 61 frames each. The first set consisted of 25,  $640 \times 480$  pixel images taken 5 min apart during early gastrulation, Stage 9–10 (Bordzilovskaya et al. 1989). The average cell area was 317 pixels, and no significant change in average cell size was seen in the image set. The other image sequences used here

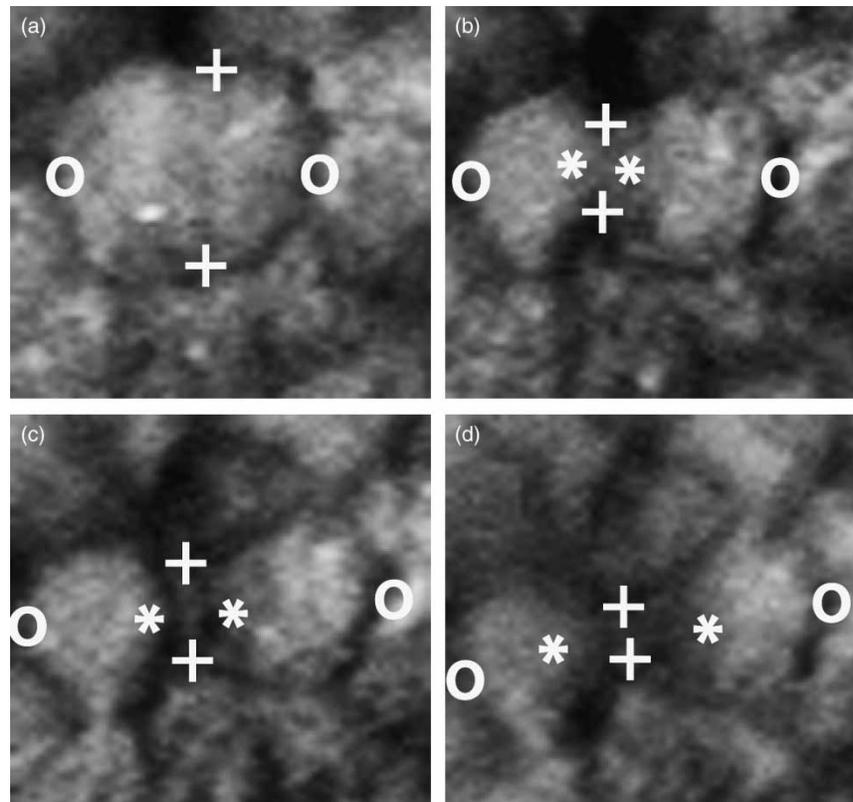


Figure 2. A high-magnification view of a single mitosis. Mitosis produces inward movement of the points labelled + and outward movement of the points labelled o and \*. A magnification of  $2.5 \mu\text{m}/\text{pixel}$  was used. (a)  $t = 0$  min, (b)  $t = 5$  min, (c)  $t = 15$  min and (d)  $t = 20$  min.

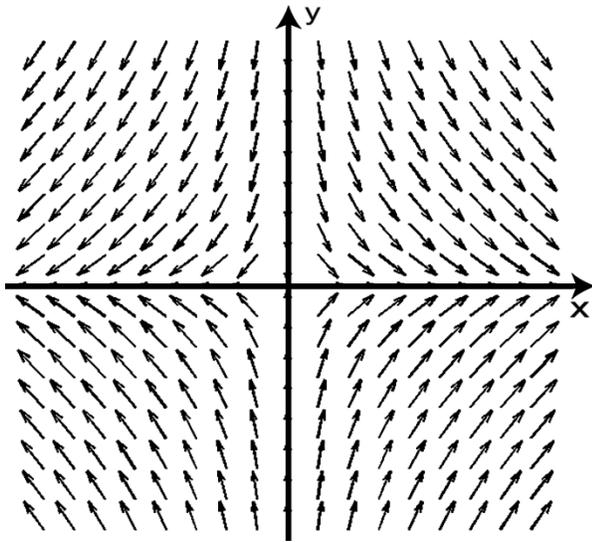


Figure 3. The mitosis motion field prototype. Outward movement occurs along the  $x$ -axis and inward movement along the  $y$ -axis. The field shown corresponds to Equation (4) with  $\mathbf{A}$  as defined in Equation (5).

were based on  $400 \times 400$  pixel sub-images extracted from one very long set that consisted of 389 images taken 1 min apart during early neurulation (Stages 12–13). A different camera system and higher lens magnification were used to collect these images and the average cell area was 2567 pixels. Amphibian embryos do not feed during early development and it is generally assumed that individual tissues do not change in mass. However, ongoing mitoses would tend to cause average cell volume to decrease slowly because the cell mass becomes more finely partitioned, and in-plane tissue expansion or contraction and their associated thickness changes could also affect cell planform area (Veldhuis et al. 2005). In the image sets used here, these effects were negligible. The algorithm

presented here would be applicable to more general situations, though, since the characteristic times of these effects would be much longer than those of the M phase of the cell cycle, the phase producing the motions the present algorithm detects. The algorithm achieved identification accuracies of 68 and 67% in the two respective image sets, and angular orientation accuracies were of the order of  $30^\circ$ . Where cell fabric analyses revealed cells to be elongated in a particular direction, mitoses orientations correlated with that direction, a finding consistent with current concepts of cell mechanics (Brodland and Veldhuis 2002).

## 2. The mitosis identification algorithm

Manual identification of mitoses in time lapse images is possible because of the remarkable processing capabilities of the human visual system. To replicate these capabilities using a computer is a challenging task. Here, we show that when a suitable combination of computer algorithms are used, mitoses locations and orientations can be identified.

In brief, image processing algorithms are used to calculate the optical flow of the image and to remove any bulk motions present in the image. Vector analysis and pattern recognition methods are then used to identify local patterns of motion and motion magnitudes consistent with mitosis. It was found that treating the patterns and magnitudes separately and then combining them using probability concepts reduced the dimension of the required searches and improved discrimination.

### 2.1 Calculation of optical flow

If the locations of all points in a source image are tracked to their new positions in a subsequent target image (Veldhuis and Brodland 1999), the resulting vector field of relative displacements is called a motion field. In embryos, these

Table 1. Summary of the image sets.

Stage	Frame	Sequence number	No. of frames	Frame interval (min)	Average cell area (pixels)	No. of mitoses
Early gastrulation	A	1	25	5	317	94
Early neurulation	B	2	21	1	2567	1
	B	3	51			2
	B	4	21			1
	B	5	21			1
	B	6	21			1
	B	7	21			1
	B	8	21			1
	B	9	21			2
	B	10	21			2
	B	11	21			1
	C	12	21			3
	C	13	21			2
	C	14	21			1
	C	15	61			5

fields tend to be relatively smooth and mathematically well behaved. Since embryo illumination is essentially uniform and constant, the motion field can be accurately approximated by its optical flow, 'the distribution of apparent velocities of movement of brightness patterns in an image' (Horn and Schunck 1981).

Optical flow can be extracted from image sequences by a variety of image processing methods including differential approaches, region matching, and energy and phase based techniques (Barron et al. 1994). Differential based techniques perform poorly in the present application because of image noise and the grainy pigmentation of the epithelium. Energy and phase based techniques, which are also more complex to implement, have limited usefulness here because of pigmentation (intensity) changes produced during nucleus division. A region matching method was ultimately chosen because it performs better than differential based techniques under noise and better than energy and phase methods in the presence of pigmentation changes.

A modified version of Singh's (1992) region matching algorithm was used here to determine the optical flow field  $\mathbf{U}(x,y)$ , where  $(x,y)$  are Cartesian image coordinates. The algorithm trial matches  $2a + 1$  by  $2b + 1$  pixel templates from a source image having pixel intensities  $I_1(x,y)$  with regions in a target image  $I_2(x,y)$  over a  $2A + 1$  by  $2B + 1$  search region centered on the location of the template in the source image. Here, the sum of squared difference used by Singh to assess the strength of each trial match was replaced by the more versatile normalised cross-correlation function

$$N(u, v) = \frac{\sum_{i=-a}^a \sum_{j=-b}^b [I_1(x-i, y-j) - \bar{I}_1(x,y)][I_2(x-u-i, y-v-j) - \bar{I}_2(x-u, y-v)]}{\sqrt{\sum_{i=-a}^a \sum_{j=-b}^b [I_1(x-i, y-j) - \bar{I}_1(x,y)]^2 \sum_{i=-a}^a \sum_{j=-b}^b [I_2(x-u-i, y-v-j) - \bar{I}_2(x-u, y-v)]^2}}. \quad (1)$$

For a pixel at  $(x,y)$  in the source image, the numerator calculates the amount by which each pixel in an  $(2a + 1)$  by  $(2b + 1)$  template centered on  $(x,y)$  differs from the mean  $\bar{I}_1$  of the points in that template. It then does the same for an  $(2a + 1)$  by  $(2b + 1)$  template in the target image centered on trial pixel positions  $(x + u, y + v)$  within the  $(2A + 1)$  by  $(2B + 1)$  search area. The products of corresponding normalised pixel values in the two templates are then summed. The sum is normalised by dividing it by the product of the root mean squared (RMS) values of the normalised pixels in each of the templates. For a particular pixel  $(x,y)$ , the function calculates the average correlation of pixels in the source template centered on that location with corresponding pixels offset by  $(u,v)$  in the target image. The ranges of  $u$  and  $v$  are  $-A \leq u \leq A$  and  $-B \leq v \leq B$ , respectively.

In many applications, the  $(u^*, v^*)$  pair giving the highest  $N(u,v)$  value is deemed to define the offset to the

new location of the pixel  $(x,y)$ . Here, however, we follow Singh's more robust approach of expanding the function  $N(u,v)$  mathematically to produce a new function

$$Q(u, v) = e^{kN(u,v)}. \quad (2)$$

A  $k$ -value of 7 was found to produce adequate selective expansion. The purpose of expanding the function is to emphasise the points of highest correlation, so that the mass center of the new function  $Q(u,v)$  over its  $(2A + 1)$  by  $(2B + 1)$  range can be taken as the best estimate of the offset or optical flow vector  $\mathbf{U}(x,y)$ . The vector field was then edited using a selective smoothing procedure (Singh 1992) so that any spurious motions could be eliminated. For the smoothing algorithm, a local template size of 1 cell diameter was used.

The template window size  $(2a + 1)$  by  $(2b + 1)$  was set to approximately  $2a + 1 = 2b + 1 = 0.5D$ , where  $D$  is a typical cell diameter and the search window size  $(2A + 1)$  by  $(2B + 1)$  was set to approximately  $2A + 1 = 2B + 1 = 1.5D$ . Average cell diameter can be obtained manually or using an automated algorithm (Iles et al. 2007). In terms of pixels, the search template was  $9 \times 9$  pixels in the gastrulation-stage images and  $21 \times 21$  pixels in the neurulation-stage images and the search region was  $29 \times 29$  pixels in the gastrulation-stage images and  $65 \times 65$  pixels in the neurulation-stage images.

While Singh's optical flow algorithm uses a central difference approach based on three successive frames, the motion field changed so much with time when mitoses occurred that more accurate results were obtained by using

pairs of images. Ideally, the images used should be separated by approximately 0.5 times the mitosis duration and the first image should show the area just before mitosis motions begin. Because mitoses are not synchronised with image collection, image frequency should be such that three to five images are collected during a typical mitosis. Although the algorithm is relatively insensitive to the spacing between the images used to calculate the motion field, this rate of image collection allows the timing criteria to be met without collection of unnecessary data. It also allows the user to choose a fixed spacing between the images used for calculating the motion field (the images are generally not consecutive) and parameters to be determined for successful mitosis detection within a given set of images. When analyzing gastrulation-stage embryos, motion fields were calculated using frames separated by 5 min, while for neurulation-stage embryos, frames were separated by 2 min.

## 2.2 Removal of bulk image motion

Dramatic bulk tissue movements occur during organogenesis, and these movements cause significant bulk translation, rotation, stretch and shear to be present in the images. Depending on the magnification of the collected images, particular components of this motion field will be more important than others.

In the present study, a  $1280 \times 1024$  pixel camera was used and magnifications were set high enough that the average cell diameter was at least 42 pixels. As a result, the length of the image diagonal was not more than 35 cell diameters. At relatively-high magnifications like this, bulk image translation dominates rotation and local tissue deformation. In some of the images sets used in this study, translations of as much as 3 cell diameters occurred during the course of a typical mitosis (20 min). Rotation and local deformation contributions were estimated in a few sample images and found to contribute less than 0.5 pixels to the total motion field during the course of a typical mitosis.

Based on these calculations, the total motion field  $\mathbf{U}(x,y)$  was deemed to consist of two components: bulk image translation and local deformation. In order that the mitosis-generated local motions not be obscured by bulk translations, the average bulk translation  $\mathbf{T}$  between two image frames had to be found and removed from the total motion field.

A variety of methods are available for estimating global translation (Maintz and Viergever 1998). The simplest of these methods is the normalised cross-correlation function, as defined in Equation (1), and it performs well when rotation, stretch and shear components are small and intensities between frames are constant, as they are in the present application. To estimate bulk motion, the template was chosen to be a rectangle one-third of the image height and width and it was centered on the source image. The search window was set to two-thirds of the image height and width, thus giving the largest search region possible.

A new motion field

$$\mathbf{L}(x,y) = \mathbf{U}(x,y) - \mathbf{T}, \quad (3)$$

(Figure 4) containing only local motions was then constructed and it was within this field that the mitosis prototypes were sought.

## 2.3 Defining the local mitosis motion prototype

Visual examination of many local motion fields  $\mathbf{L}(x,y)$ , led to the concept that the mitosis prototype could be described mathematically by the vector field equation (Shu and Jain 1994)

$$\mathbf{V}(x,y) = A \begin{bmatrix} x \\ y \end{bmatrix}. \quad (4)$$

In this equation,  $(x,y)$  is a point within the  $d$  by  $e$  range of the mitosis prototype. The values of  $d$  and  $e$  were chosen

to be nominally 0.5 times a typical cell diameter, which corresponds to approximately 9 pixels in the gastrulation-stage image set and 21 pixels in the neurulation-stage image set.

To represent a saddle point (Figure 3)  $A$  must be chosen such that its eigenvalues are real and the product of the eigenvalues is less than zero (Shu and Jain 1994). Observations of typical mitoses showed that the saddle point often involved inward and outward motion components of similar magnitude, a pattern produced by the matrix

$$A_1 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (5)$$

In some image sequences, the inward motion illustrated by  $+$  in Figure 2 is modest when compared to the other two motion patterns and a better choice for the  $A$  matrix is

$$A_2 = \begin{bmatrix} -1 & 0 \\ 0 & 100 \end{bmatrix}. \quad (6)$$

The two forms of  $A$  used in this study do not represent all possible saddle points but, because vector directions and magnitudes were treated separately and suitable cut-off parameters were used, are sufficient to capture the mitoses that occur.

## 2.4 Motion field analysis

The next task was to find all occurrences of the mitosis prototype in the local image motion field  $\mathbf{L}(x,y)$ . Two standard approaches to analysing vector fields are critical point analysis (Rao and Jain 1992; Shu and Jain 1994) and template matching (Heiberg 2001; Ebling and Scheuermann 2003). In critical point analysis, all zero velocity points in the local image motion field are identified because they are candidates for the zero velocity point at the center of the mitosis prototype. The region surrounding each such zero velocity point is then analysed mathematically to determine if a saddle point is present. The smoothing procedure applied to the optical flow field gives rise to a substantial number of points that have essentially zero velocity, making this approach impractical.

In the template matching approach (Heiberg 2001; Ebling and Scheuermann 2003), a user-given template motion field pattern, like the mitosis prototype, is matched against a larger motion field using only vector orientation information. It was found that in the present application, the template matching process is best done as two separate operations: one that matches motion patterns and another that ensures the motion amplitudes are of

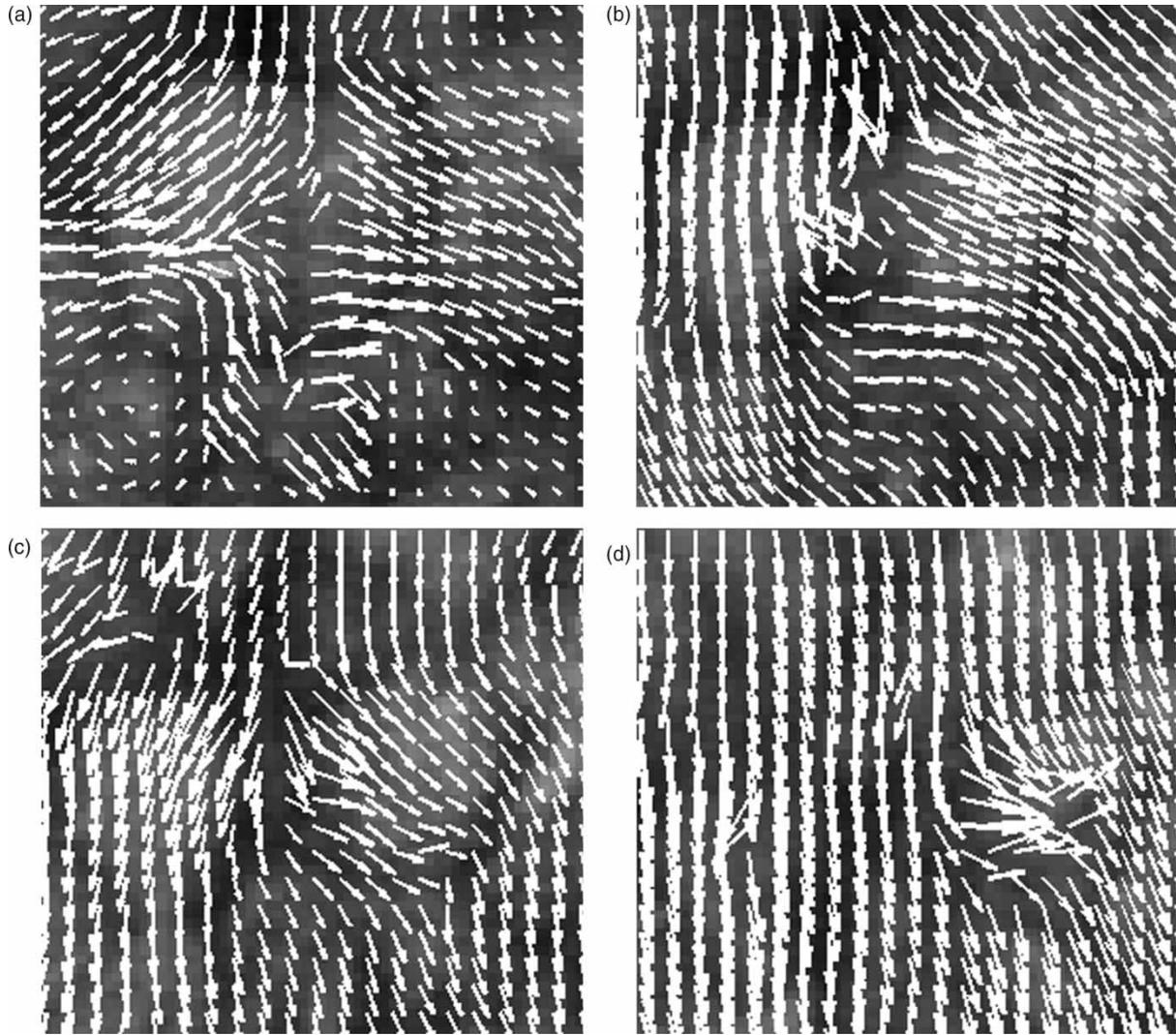


Figure 4. Typical local image motion fields  $\mathbf{L}$  associated with mitoses.

sufficient magnitude. For the vector matching operation, all of the vectors in the mitosis prototype field  $\mathbf{V}$  and the local image motion field  $\mathbf{L}$  are normalised to unit length. This approach works well because cases where the narrowing and elongation motions have different magnitudes map onto a single pattern, namely one that has unit vectors in the narrowing direction and unit vectors in the widening direction. The motion magnitude operation verifies that these motions are of cellular magnitude not just minor local undulations that coincidentally have local directional patterns like those produced by mitoses.

In the vector field matching step, the normalised prototype vector is translated so that it is centered on a location  $(i, j)$  in the local motion field. Each of the  $d$  by  $e$  points in the prototype field is then dotted with the motion field vector associated with the closest point in the motion

field, and

$$D(i, j) = \frac{1}{(2d+1)(2e+1)} \sum_{x=-d}^d \sum_{y=-e}^e |\mathbf{L}(i-x, j-y)^T \mathbf{V}(x, y)|, \quad (7)$$

the average of these dot products is calculated. The prototype is then rotated by  $5^\circ$  and a new correlation measure calculated. Whichever such product for total rotations up to  $\chi = 175^\circ$  yields the highest correlation value is deemed to be the best angular match for the prototype. In this step, both forms of  $A$  are used at each pixel location and the highest correlation value, denoted as  $D_{\max}(i, j)$ , is the one reported. Close parallels exist between the vector field correlation function given in Equation (7) and the template matching function in Equation (1).

When the vector template is aligned with the coordinate axes ( $\chi = 0$  or  $90^\circ$ ), the image field contains vectors that exactly underlie the template vectors. This alignment produces a bias compared to cases where the templates are rotated and the image field does not contain vectors that directly underlie the template vectors. To remove this bias, the correlations are multiplied by a factor

$$B(\chi) = \eta - (1 - \eta) \cos(4\chi). \quad (8)$$

Bias removal was not particularly sensitive to the value of  $\eta$  and a value of  $\eta = 0.55$  worked well for both image sets.

In order to accentuate matches, the resulting function  $D_{\max}(i, j)$ , which has a range between 0 and 1 is transformed using a sigmoid function (Duda et al. 2000) to yield a new function

$$P_O(i, j) = \frac{1}{1 + e^{-k_O(D_{\max}(i, j) - b_O)}}. \quad (9)$$

The values  $k_O = 5$  and  $b_O = 0.5$ , were found experimentally to provide good discrimination between matches and non-matches, and suitable values do not appear to depend on the particular image set being analysed. This sequence of operations completes the vector pattern portion of the template matching scheme and the function  $P_O(i, j)$  can be deemed to represent the probability of a mitosis prototype match at location  $(i, j)$  based on directional characteristics of the local motions in its vicinity.

To assess whether the magnitudes of the local motion were consistent with a mitosis, local motion field velocity magnitude (speed) information was calculated. The angle between the epithelium normal and the optic axis was small enough that foreshortening correction was unnecessary. The velocity measure used was the RMS of the image motion field over an  $n \times n$  pixel square centered on pixel  $(i, j)$ . The resulting field was then smoothed based on an area of 1 cell diameter, to produce a velocity measure  $\|\mathbf{L}(i, j)\|_{\text{smoothed}}$ . This function was then transformed using a sigmoid function to yield

$$P_M(i, j) = \frac{1}{1 + e^{-k_M(\|\mathbf{L}(i, j)\|_{\text{smoothed}} - b_M)}}. \quad (10)$$

This function can be considered a probability of mitosis measure based on local velocity magnitude. The purpose of the sigmoidal transformation is to produce a sharp transition at the velocities characteristic of mitosis (Figure 5). The mitosis velocity histogram is constructed by identifying 3–5 mitoses in the image sequence and plotting the  $\|\mathbf{L}(i, j)\|_{\text{smoothed}}$  values for those points. The value of  $b_M$ , the velocity for which the histogram has a value of 0.5, is chosen as the minimum velocity found on the histogram; this will ensure that most true mitosis regions are assigned a probability of match of at least 0.5. The value of  $k_M$ , which governs the steepness of the

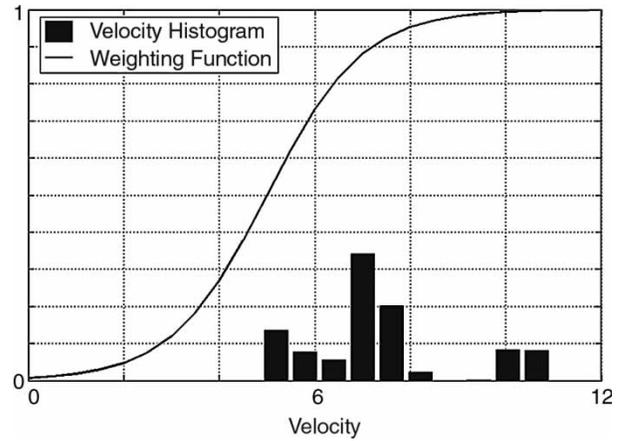


Figure 5. The sigmoid weighting function. To determine the parameters  $b_M$  and  $k_M$ , the a velocity magnitude histogram is drawn. The histogram shown is based on five manually selected mitoses. The sigmoid parameter  $b_M = 5$  is set to the smallest velocity in the histogram and  $k_M = 1$  is chosen, so that the sigmoid function reaches 1 by the time the maximum histogram value is reached.

sigmoid transition, is chosen such that  $P_M = 1$  at the maximum velocity found on the histogram; this approach will ensure that the maximum velocity of the mitosis region gets a probability of match of 1. For the data shown in Figure 5, taken from five manually identified mitoses in the early stage image set, the parameters are assigned values of  $k_M = 1$  and  $b_M = 5$ , giving the sigmoid curve shown. The functions  $P_M$  and  $P_O$  for the mitosis shown in Figure 2 are shown in Figure 6.

The motion pattern and magnitude match functions are then combined to produce a total probability of mitosis match function

$$P_{\text{mitosis}}(i, j) = \frac{P_O(i, j) + P_M(i, j)}{2}, \quad (11)$$

for each pixel in the source image.

## 2.5 Mitosis count and orientation

The final step in the algorithm is to determine those locations in the total probability of mitosis map where individual mitoses occur. Two magnitude thresholds  $T_{\text{low}}$ ,  $T_{\text{high}}$  and a spatial size threshold  $T_{\text{size}}$  are used in this process.

The threshold  $T_{\text{low}}$  is applied to  $P_{\text{mitosis}}$  to produce a binary image  $P'_{\text{mitosis}}$  (Figure 6(c)), where all potential mitosis locations are indicated by the high (white) pixels. Then, points associated with each other through an ‘8-connectivity rule’ (Gonzalez et al. 2004) – that is points that directly border each other horizontally, vertically or diagonally – are grouped together to form one of a series

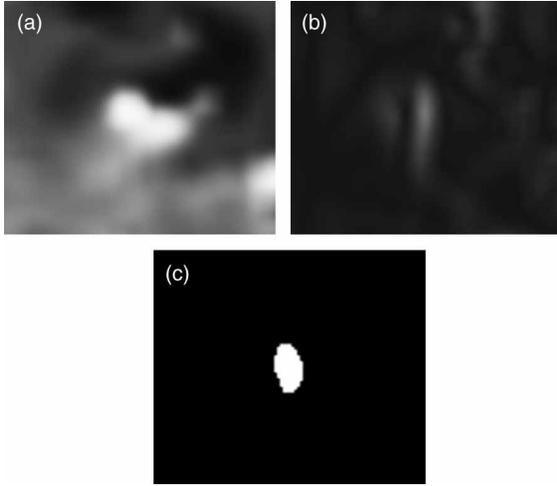


Figure 6. (a) Magnitude probability function. (b) Orientation probability function and (c) threshold of the sum of magnitude and orientation probability functions. Black corresponds to a function value of zero, while white corresponds to one.

of objects  $C_i$ . A morphological closing operator (Dougherty and Lotufo 2003) with a square structuring element of size  $N = 7$ , is performed on  $P''_{\text{mitosis}}$ , to connect any small gaps in  $P''_{\text{mitosis}}$  as a result of the thresholds. A larger value of structuring element size causes multiple adjacent mitoses regions to be erroneously joined. If the group is larger than  $T_{\text{size}}$ , a mitosis is considered to have occurred there. A mitosis is also considered to have occurred if one or more pixels in  $C_i$  has a value greater than  $T_{\text{high}}$ . Mathematically, the probability of a mitosis occurring at a particular location is deemed to be given by

$$P''_{\text{mitosis}}(C_i) = \begin{cases} 1 & \text{if size}(C_i) > T_{\text{size}} \text{ OR} \\ & \max(P_{\text{mitosis}} \in C_i) > T_{\text{high}} \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

To avoid multiple detection of the same mitosis in successive frames, any detected mitoses that are common between matching regions in consecutive frames are considered to be the same mitosis.

The mitosis orientation associated with the pixels in a particular mitosis region  $C_i$  will, in general, have different orientations  $\chi$ . The angle assigned for a particular mitosis region is the one most common within that region. Care must be exercised in this calculation because of the discontinuity produced as angle designations change suddenly from 0 to 360°.

Suitable values for the parameters  $T_{\text{low}}$  and  $T_{\text{high}}$  depend on the images and on the sigmoid function parameters  $k_O$ ,  $b_O$ ,  $k_M$  and  $b_M$ . The allowable ranges of these parameters are  $0.5 \leq T_{\text{low}} \leq 1$  and  $T_{\text{low}} < T_{\text{high}} \leq 1$ , and if the sigmoid function parameters are chosen to sufficiently separate the mitosis pixels from non-mitosis pixels, then  $T_{\text{low}} \approx 0.5$  and  $T_{\text{high}} \approx 1$  work well.

The threshold  $T_{\text{size}}$  is related to cell size and a value of approximately 0.5 times the number of pixels in a typical cell was found to work well.

## 2.6 Algorithm evaluation

Mitoses recognition is an event identification problem, a class of problem well known in the field of pattern recognition. Algorithm evaluation requires that mitoses occurrences, locations and orientation be identified by hand to establish the ‘ground truth’ for a set of test images. Then the algorithm is used to process the same image set. In terms of evaluating mitosis detection, a binary event, the data available are the number of correctly detected mitoses (TP, true positive), the number of incorrectly detected mitoses (FP, false positive) and the number of mitoses not detected (FN, false negative). If this were a classification problem, the true number of cells not undergoing mitosis (TN, true negative) would also be required, and if the total number of cells  $N$  in the images were known, it could be calculated as the difference between  $N$  and TP.

Two of the standard measures of the performance of an event identification algorithm are precision (Manning et al. 2008)

$$P = \frac{TP}{TP + FP}, \quad (13)$$

and recall (Manning et al. 2008)

$$R = \frac{TP}{TP + FN}. \quad (14)$$

Algorithm precision describes the fraction of the mitoses identified by the algorithm that actually are mitoses. Recall, also called the true positive rate, is the fraction of mitosis correctly detected by the algorithm out of the total mitosis present in the test data set. Ideally an algorithm would yield a precision of 100% and a recall of 100%.

The performance rate or score for a given algorithm is given by the *F-measure* (Manning et al. 2008)

$$F_\beta = \frac{(\beta^2 + 1)PR}{\beta^2 P + R}, \quad (15)$$

where the parameter  $\beta$  controls the importance of recall over precision. In this study both precision and recall are assumed to be equally important, and  $\beta$  is set to 1. The performance of an algorithm relative to a parameter in that algorithm can be studied using the precision-recall curve (Figure 7), a curve related to the receiver operator curve (ROC; Davis and Goadrich 2006) commonly used to evaluate classification algorithms.

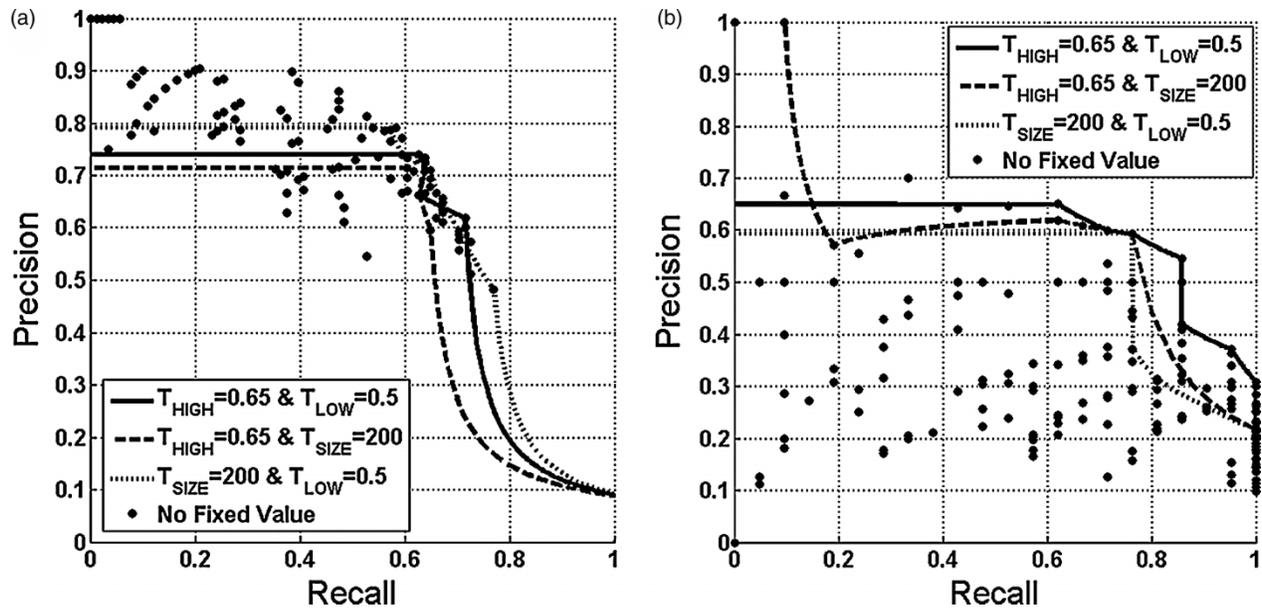


Figure 7. Precision-recall graphs. The graphs were constructed for various values of the parameters  $T_{low}$ ,  $T_{high}$  and  $T_{size}$ . Pairs of parameters were set to their optimal values while the third was varied. Large scale reshaping motion is present in the gastrulation-stage image set due to the use of low-magnification images, and high FN rates result. The neurulation-stage image sets use high-magnification images and the FN rates are lower. However, the high magnification gives rise to illumination changes and noise, thus producing high FP rates. (a) Gastrulation stage and (b) neurulation stage.

### 3. Results and conclusions

A set of 25 gastrulation-stage images (denoted as sequence 1 in Table 1) were collected at 5-minute intervals from the area denoted A in Figure 1. Using a different embryo and apparatus, images of neurulation-stage were collected at 1-minute intervals from the neural plate area. Sequences 2–11 were taken from area B and sequences 12–15 were from area C. In the first set, there are nominally 400 pixels per mm, while in the second there are 1400 pixels per mm. The lenses had sufficient numerical apertures that resolution was not diffraction limited.

Ground truth was determined by manual study of the image sets by an experienced observer. Image sequences were viewed using a custom Matlab interface that allowed images to be played forward or backward by single frames or in short sequences at user-defined rates. When single cells became visible as two cells, a mitosis was deemed to have occurred, and the time, location and orientation of the mitosis was registered by drawing a line along the newly formed cell boundary using the Matlab interface. The mitosis orientation was deemed to be normal to this line.

In order to determine suitable values for the analysis parameters  $T_{low}$ ,  $T_{high}$  and  $T_{size}$ , precision-recall curves (Figure 7) were constructed for candidate parameters. The points near the top of the graph correspond to parameter choices that produce high precision values, meaning that most of the detected mitoses are true mitoses. The points near the right of the graph have high recall, indicating that

most of the true mitoses are detected. Parameters that produce a point at the top right corner of the graph would be preferred, but this ideal cannot be achieved.

The missed detections (false negatives) in the late blastula stage can be attributed to large scale reshaping during the late blastula stage. This reshaping can alter the local image motion field  $\mathbf{L}$  sufficiently that the present mitosis prototype becomes less effective. In the early neurulation stage, large scale reshaping was not present, and high recall rates could be achieved. The low precision rates (high false positives) found in the neurulation image set can be attributed to illumination changes between frames and grainy additive noise in the images. Both of these factors cause errors in estimation of the local image motion field  $\mathbf{L}$ .

The manual study of the images in image sequence 1 revealed 94 mitoses. To tune the parameters, three of these mitoses were selected manually and used to determine that  $k_M = 1$  and  $b_M = 5$  are suitable parameter values. The parameters  $T_{low} = 0.5$ ,  $T_{high} = 0.65$  and  $T_{size} = 200$ , were chosen because a parametric search (Figure 7) showed that they maximised the  $F$ -measure,  $F_{\beta=1}$  over the range  $0.5 \leq T_{low} \leq 1$ ,  $T_{low} \leq T_{high} \leq 1$  and  $50 \leq T_{size} \leq 300$ .  $T_{low}$  is in fact, 0.5 as expected, and the relatively low value of  $T_{high}$  suggests that the values associated with  $P_{mitosis}$  do not span the entire range of 0–1.

When these parameters were used in the algorithm, it successfully identified 58 of the 91 mitoses that occurred.

Table 2. Detection rates.

	Predicted			
	Early gastrulation		Early neurulation	
	Positive	Negative	Positive	Negative
<i>Actual</i>				
Positive	58	33	16	5
Negative	21	n/a	11	n/a

While this ratio could have been increased to 70 out of 91 by a different choice of parameters, as shown in Figure 7, the number of false positives would then have been unacceptably high. Confusion matrices based on the parameters shown above are presented in Table 2, and the algorithm performance as determined using the *F-measure* is 68%. The detected locations were typically within a few pixels of the manually-identified mitosis locations.

The algorithm reported mitosis orientation histograms that were consistent with those found manually (Figure 8).

Differences between the manual and algorithmic angles reported for individual mitoses were often less than  $10^\circ$  (Figure 9), the width of the histogram class interval and usually less than  $30^\circ$ .

For the images collected every minute from the surface of a neurulation-stage embryo, a manual analysis identified a total of 24 mitoses (Table 1). The time lapse image sequences were then analysed by the algorithm using the same parameters as for the gastrulation stage images, except that  $b_M = 7$  and  $T_{low} = 0.5$ ,  $T_{high} = 0.95$  and  $T_{size} = 400$  based on a  $100 \leq T_{size} \leq 500$  search. The algorithm performance was found to be 67%, similar to the value obtained for the early blastula stage. The algorithm produced histograms of mitosis orientations consistent with the ground truth (Figure 8), and error histograms as shown in Figure 9. That the algorithm gives similar performance for image sets having different cell sizes and magnifications demonstrates that it is robust to these factors. If cell density or other relevant characteristics were to change significantly within a set of images, the parameters may have to be adjusted or made functions of time.

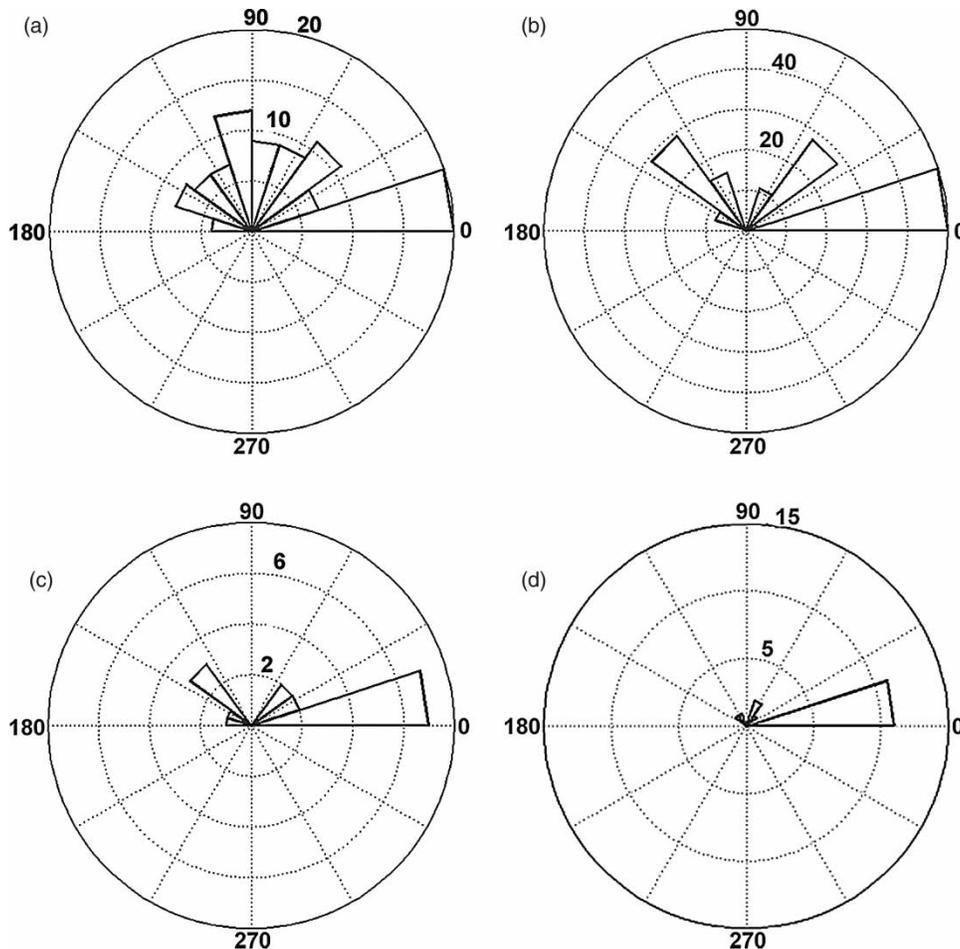


Figure 8. Ground truth and algorithmically-detected mitosis orientation histograms. (a) Gastrulation ground truth, (b) gastrulation algorithm detection, (c) neurulation ground truth and (d) neurulation algorithm detection.

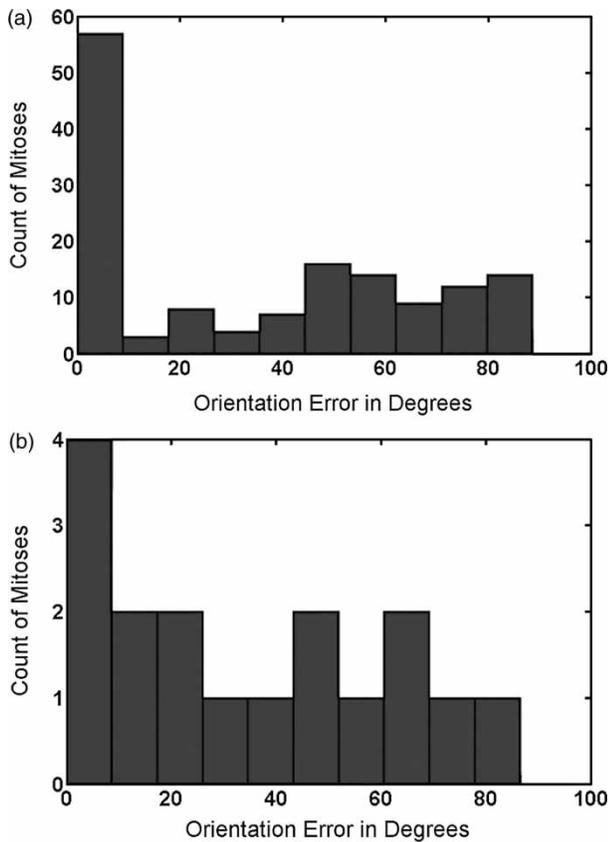


Figure 9. Orientation error histograms. The histograms show the angular error between ground truth and algorithmically-detected mitoses. (a) Gastrulation stage and (b) neurulation stage.

From a developmental biology perspective, the algorithm shows that the rate of cell division reduces by several orders of magnitude between early gastrulation and early neurulation. The algorithm also shows that mitosis orientation patterns change significantly. In the area studied during gastrulation, mitosis orientations were widely distributed, whereas during neurulation mitoses in the subject were quite strongly aligned with the medio-lateral axis. A number of biological studies have suggested that the long axes of cells (Jacobson and Gordon 1976) are key determinants of mitosis orientation (O'Connell and Wang 2000). Other relevant factors may include cell signalling, biochemical factors, mechanical stresses and tissue deformation (Brodland and Veldhuis 2002; Gong et al. 2004; Nelson et al. 2005). In turn, regularly oriented mitosis, regardless of its origin, has been implicated as a significant driver of tissue motions (Philip et al. 1992; Brodland and Veldhuis 2002; Gong et al. 2004).

A computer algorithm was used to determine the cellular fabric (Iles et al. 2007) in the images used here. It revealed that a significant average direction of cell elongation did not exist in the gastrulation-stage images, consistent with

broadly distributed mitosis orientations. In contrast, the fabric algorithm showed that cells in the neurulation-stage images had a significantly anisotropic fabric: they were elongated in the medio-lateral direction (normal to the embryo midline) by an average factor of  $\kappa = 1.1$  (Brodland et al. 2006), a degree of anisotropy not generally discernable by eye. Mitoses were found to be strongly aligned with this direction, suggesting that an average aspect ratio  $\kappa$  of 1.1 is sufficient to regulate mitosis orientation.

Computer simulations (Brodland 2006; Brodland and Veldhuis 2006) show that medio-laterally oriented cells would be expected in the neural plate because of the lamellipodium-driven convergent extension occurring there. Whether the mitosis orientations are a direct result of these cell geometries or are the consequence of morphogen gradients that trigger lamellipodium action (Keller et al. 2000; Wallingford and Harland 2002; Keller 2006) is not yet clear.

In terms of tissue reshaping, the monolayer epithelium that forms the surface of the embryo must expand dramatically during gastrulation so that tissue can stream through the blastopore, into the interior of the embryo, where it later forms various internal structures (Gilbert 2000). Whether the elevated cell division rates observed in the present study would facilitate the associated high strain rates is not clear. Strain rates are low in the cephalic region, where the neurulation stage images were collected (Veldhuis et al. 2005), suggesting that the strongly oriented but relatively low rate cell division observed there is not a primary driver of local tissue reshaping, a result consistent with computer simulations (Brodland and Veldhuis 2002).

The present study makes available regional data on mitosis frequency and orientation, factors that computer simulations have shown are mechanically associated with tissue deformation rates, cellular fabric and lamellipodium action (Brodland and Veldhuis 2002; 2006). To the kinematic data (Brodland et al. 1995), strain maps (Veldhuis et al. 2005), tensile properties (Wiebe and Brodland 2005) and *in vivo* stresses (Benko and Brodland 2007) now available for axolotl embryos, it adds mitosis data. By filling in one of the few remaining gaps it allows us to take another step toward construction of an integrated theory for the mechanics of early embryo development.

### Acknowledgements

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) through Discovery Grants to GWB and DC and a CGS-M Scholarship to PS. Animals were cared for in accordance with Canadian Council on Animal Care (CCAC) guidelines.

### Notes

1. Email: psiva@uwaterloo.ca
2. Email: dclausi@uwaterloo.ca

## References

- Barron JL, Fleet DJ, Beauchemin SS. 1994. Performance of optical flow techniques. *Int J Comput Vis.* 12(1):43–77.
- Baye MB, Link BA. 2007. Interkinetic nuclear migration and the selection of neurogenic cell divisions during vertebrate retinogenesis. *J Neurosci.* 27(38):10143–10152.
- Benko R, Brodland GW. 2007. Measurement of *in vivo* stress resultants in neurulation-stage amphibian embryos. *Ann Biomed Eng.* 35(4):672–681.
- Bertuccio L, Nunnari G, Randieri C, Rizza V, Sacco A. 1998. A cellular neural network based system for cell counting in culture of biological cells. Paper presented at: Proceedings of the IEEE International Conference on Control Applications; 1998 Sep 1–4; Trieste, Italy.
- Bordzilovskaya NP, Dettlaff TA, Duhon ST, Malacinski GM. 1989. Developmental-stage series of axolotl embryos. In: *Developmental biology of the axolotl*. New York (NY): Oxford University Press.
- Brodland GW. 2006. Do lamellipodia have the mechanical capacity to drive convergent extension? *Int J Dev Biol.* 50:151–155.
- Brodland GW, Veldhuis JH. 1998. 3D reconstruction of live embryos using robotic microscope images. *IEEE Trans Biomed Eng.* 45(9):1173–1181.
- Brodland GW, Veldhuis JH. 2002. Computer simulations of mitosis and interdependencies between mitosis, cell shape and epithelium reshaping. *J Biomech.* 35:673–681.
- Brodland GW, Veldhuis JH. 2006. Lamellipodium-driven tissue reshaping: a parametric study. *Comput Methods Biomech Biomed Eng.* 9:17–23.
- Brodland GW, Scott MJ, MacLean AF, Globus M, Globus SV, Gordon R, Veldhuis JH, Maestro RD. 1995. Morphogenetic movements during axolotl neural tube formation tracked by digital imaging. *Dev Genes Evol.* 205:311–318.
- Brodland GW, Chen DIL, Veldhuis JH. 2006. A cell-based constitutive model for embryonic epithelia and other planar aggregates of biological cells. *Int J Plast.* 22:965–995.
- Chen X, Brodland GW. 2008. Multi-scale finite element modeling allows the mechanics of amphibian neurulation to be elucidated. *Phys Biol.* 5(1):1–15.
- Chen YM, Biddell K, Sun AY, Relue AP, Johnson JD. 1999. An automatic cell counting method for optical images. Paper presented at: Proceedings of The First Joint BMES/EMBS Conference; 1999 October; Atlanta, GA, USA.
- Cross SS, Start RD. 1996. Estimating mitotic activity in tumours. *Histopathology.* 29:485–488.
- Davis J, Goadrich M. 2006. The relationship between Precision-Recall and ROC curves. Paper presented at: ICML 2006. Proceedings of the 23rd international Conference on Machine Learning; 2006 June 25–29; New York, NY, USA.
- Dougherty E, Lotufo R. 2003. *Hands-on morphological image processing*. Bellingham (WA): SPIE Press.
- Duda RO, Hart PE, Stork DG. 2000. *Pattern classification*. 2nd ed. New York: Wiley-Interscience.
- Ebling J, Scheuermann G. 2003. Clifford convolution and pattern matching on vector fields. *IEEE Vis.* 193–200.
- Gal R, Rath-Wolfson L, Rosenblatt Y, Halpern M, Schwartz A, Koren R. 2005. An improved technique for mitosis counting. *Int J Surg Pathol.* 13(2):161–165.
- Gilbert SF. 2000. *Developmental biology*. Sunderland (MA): Sinauer Associates.
- Gong Y, Mo C, Fraser SE. 2004. Planar cell polarity signalling controls cell division orientation during zebrafish gastrulation. *Nature.* 430.
- Gonzalez RC, Woods RE, Eddins SL. 2004. *Digital image processing using MATLAB*. Upper Saddle River (NJ): Prentice Hall.
- Heiberg EB. 2001. *Automated feature detection in multi-dimensional images – a unified tensor approach* [Licensiate]. Sweden: Linköping universitet.
- Horn BKP, Schunck BG. 1981. Determining optical flow. *Artif Intell.* 17(1-3):185–203.
- Iles PJW, Brodland GW, Clausi DA, Puddister SM. 2007. Estimation of cellular fabric in embryonic epithelia. *Comput Methods Biomech Biomed Eng.* 10(1):75–84.
- Jacobson AG, Gordon R. 1976. Changes in the shape of the developing vertebrate nervous system analyzed experimentally, mathematically and by computer simulation. *J Exp Zool.* 197:191–246.
- Kaman EJ, Smeulders AWM, Verbeek PW, Young IT, Baak JPA. 1984. Image processing for mitoses in sections of breast cancer: A feasibility study. *Cytometry.* 5(3):244–249.
- Kate TKT, Belien JAM, Smeulders AWM, Baak JPA. 1993. Method for counting mitoses by image processing in feulgen stained breast cancer sections. *Cytometry.* 14(3):241–250.
- Keller R. 2006. Mechanisms of elongation in embryogenesis. *Development.* 133:2291–2302.
- Keller R, Davidson L, Edlund A, Elul T, Ezin M, Shook D, Skoglund P. 2000. Mechanisms of convergence and extension by cell intercalation. *Philos Trans R Soc Land B.* 355:897–922.
- Maintz JBA, Viergever MA. 1998. A survey of medical image registration. *Med Image Anal.* 2(1):1–36.
- Manning CD, Raghavan P, Schütze H. 2008. *Introduction to information retrieval*. Cambridge: Cambridge University Press. [www-csli.stanford.edu/~schuetze/information-retrieval-book.html](http://www-csli.stanford.edu/~schuetze/information-retrieval-book.html) (In Press).
- Nelson CM, Jean RP, Tan JL, Liu WF, Sniadecki NJ, Spector AA, Chen CS. 2005. Emergent patterns of growth controlled by multicellular form and mechanics. *PNAS.* 102(33):11594–11599.
- O'Connell CB, Wang Y. 2000. Mammalian spindle orientation and position respond to changes in cell shape in a dynein-dependent fashion. *Mol Biol Cell.* 11(5):1765–1774.
- Philip L, Hertzler C, Wallis HC. 1992. Cleavage and gastrulation in shrimp *Sicyonia ingentis*: invagination is accompanied by oriented cell division. *Development.* 116:127–140.
- Phukpattaranont, P, Boonyaphiphat, P. 2006. An automatic cell counting method for a microscopic tissue image from breast cancer. Paper presented at: IFMBE. Proceedings of the 3rd Kuala Lumpur International Conference on Biomedical Engineering; 2006 December; Kuala Lumpur, Malaysia.
- Puddister S. 2003. *Estimating bulk geometrical properties of cellular structures [MASc]*. Waterloo (ON): University of Waterloo.
- Rao AR, Jain RC. 1992. Computerized flow field analysis: oriented texture fields. *IEEE Trans Pattern Anal Mach Intell.* 14(7):693–709.
- Refai HH, Li L, Teague TK, Naukam R. 2003. Automatic count of hepatocytes in microscopic images. Paper presented at: International Conference on Image Processing; Barcelona, Spain.
- Shimada T, Kato K, Kamikouchi T, Ito K. 2005. Analysis of the distribution of the brain cells on the fruit fly by an automatic cell counting algorithm. *Physica A.* 350:144–149.
- Shu CF, Jain RC. 1994. Vector field analysis for oriented patterns. *IEEE Trans Pattern Anal Mach Intell.* 16(9):946–950.

- Simpson JF, Dutt PL, Page DL. 1992. Expression of mitoses per thousand cells and cell density in breast carcinomas: a proposal. *Hum Pathol.* 23:608–611.
- Singh A. 1992. Optic flow computation: a unified perspective. IEEE Comput Soc Press.
- Siva P. 2007. Quantifying the frequency and orientation of mitoses in embryonic epithelia [MASC]. Waterloo (ON): University of Waterloo.
- Talukder A, Casasent DP. 1998. Multiscale gabor wavelet fusion for edge detection in microscopy images. *SPIE.* 3391(1): 336–347.
- Veldhuis JH, Brodland GW. 1999. A deformable block-matching algorithm for tracking epithelial cells. *Image Vis Comput.* 17:905–911.
- Veldhuis JH, Brodland GW, Wiebe CJ, Bootsma GJ. 2005. Multiview robotic microscope reveals the in-plane kinematics of amphibian neurulation. *Ann Biomed Eng.* 33(6):821–828.
- Vincent LM, Masters BR. 1992. Morphological image processing and network analysis of cornea endothelial cell images. *SPIE.* 1769:212–226.
- Wallingford JB, Harland RM. 2002. Neural tube closure requires dishevelled-dependent convergent extension of the midline. *Development.* 129:5815–5825.
- Wiebe CJ, Brodland GW. 2005. Tensile properties of embryonic epithelia measured using a novel instrument. *J Biomech.* 38:2087–2094.
- van Diest PJ, Baak JP, Matze-Cok P, Wisse-Brekemans EC, van Galen CM, Kurver PH, Bellot SM, Fijnheer J, van Gorp LH, Kwee WS, et al. 1992. Reproducibility of mitosis counting in 2,469 breast cancer specimens: results from multicenter morphometric mammary carcinoma project. *Hum Pathol.* 23:603–607.