A time-varying-cell segmentation, tracking and measure scheme

Gabriela Maletti and Jesper Skjerning

Department of Informatics and Mathematical Modelling,
Technical University of Denmark,
DK-2800 Kgs. Lyngby, Denmark
gmm@imm.du.dk, jesperskjerning@hotmail.com
http://www.imm.dtu.dk/image

Abstract. A cell classification scheme in temporal sequences is proposed. This scheme assumes stationarity of the cells changes. The data set is composed by images of cells with a large variability in size, shape, color and texture. The cells are tracked between frames and cell descriptors within and between frames are computed and analyzed for real examples.

1 Introduction

Data is a set of 365 image frames, which joints to 15 sequences of moving and transforming cells. Each frame is a GIF-image of 640×480 pixels of size and 256 quantization levels. The sequences are grouped in touching cells and non-touching cells with and without transforming cells. Transforming cells correspond to the mitosis process. For each sequence, in general, the image background is the same. Its texture could be described as slightly rough. However, there is some variation in light. Some slight camera movements can also be seen. By visual appreciation, there is no common pattern of cell that can be found within and between sequences: to define an homogeneity criteria that is satisfied by all the cells using either spectral or textural information seems to be a difficult task. See some examples in Figure 1. Furthermore, much of the cell borders are covered by a bright area, probably caused by light sources from the side. There is also a large variation in size, shape, color and texture within cells. Regarding to the cell movements, some sequences contain slowly moving cells while others contain active moving cells. The cases for which cells interaction exists, some cells getting so close that they look like one can be appreciated. It has to be mentioned that there is no information about time between frames nor physical size of the pixels.

Based on previous comments, the approach is that the images could be segmented using the gradient of the original data. The present work is divided in three stages. In the first stage, for each sequence and frame, cells are segmented using the gradient. In the second stage corresponding objects between frames are found assuming minimal distance of center of mass locations. Third, object descriptors are computed and analyzed to describe the non moving cell change.

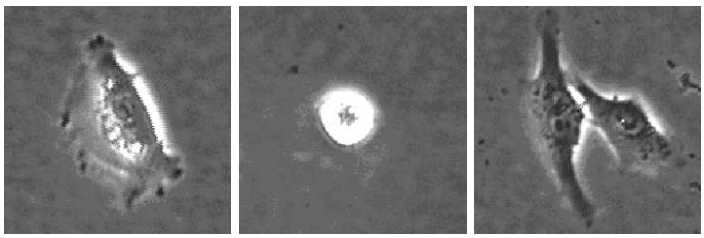

Fig. 1. Example cells. From left: No Mitosis, Mitosis and Touching).

2 Cells segmentation

In this section, some definitions are first given. Afterwards, details of the segmentation scheme proposed follow.

Definition 1. *Following [1], let a specific image with* $\mathsf{X}_t = \{x_t[r,c]\}$ *be defined over the given lattice:*

$$\mathsf{L} = \{[r,c] | 1 \leq r \leq max_r, 1 \leq c \leq max_c\} \tag{1}$$

Let each image X_t *be the realization of a two-dimensional random field* X_t [1]*;* X_t *is hierarchically defined in terms of the realization* $\mathsf{Z}_t = \{z_t[r,c]\}$ *of an underlying random field* Z_t *such that:* Z_t *represents the partition of the domain* L *in* K *regions of different types. Each* $z_t[r,c]$ *is a value of the set of labels* $B = \{b_1, b_2, ..., b_K\}$ *where* $z_t[r,c] = b_k$ *indicates that, at time* t*, the coordinate* $[r,c]$ *of a given pixel belongs to the region* k*.*

Definition 2. *Let the gradient of* X_t *in general be defined as:*

$$\nabla X_t = \sqrt{\frac{\partial X_t}{\partial t}^2 + \frac{\partial X_t}{\partial r}^2 + \frac{\partial X_t}{\partial c}^2} \tag{2}$$

The main assumption of the present scheme is stationarity in the change between frames. For this reason, it requires input data with a large variability between frames. Details of the approach follow.

For each sequence, a set of synthetical images composed by the gradients within and between frames is generated. Each synthetical image is assumed to be composed by two classes: large values corresponding to pixels belonging to the class cells and small values in the regions corresponding to the background.

[1] Notice that while X_t (sans serif) is a set of variables, X_t (italic) is a set of values of those variables.

In practice, the gradient is computed for each single pixel as follows:

$$\nabla x_t[r,c] = \sqrt{(x_t[r,c] - x_{t+1}[r,c])^2 + (x_t[r,c] - x_t[r+1,c])^2 + (x_t[r,c] - x_t[r,c+1])^2} \quad (3)$$

where $[r,c]$ is the pixel location in the lattice L (see Equation 1), given by the row r and column c. x_t is the pixel value at the time t.

The number of frames of this synthetical sequence is given by the number of frames in the original sequence, minus one frame. This is due to the synthetical sequences are based on the change from frame to frame. The values in each single frame of the synthetical images are normalized to 256 quantization levels. After applying a threshold, a new sequence of binary images is generated. Classified pixels are labelled as follows: pixels with value 1 indicate the class cells and pixels with value 0, the class background.

Wholes within objects are automatically filled. For each thematic map the three largest objects are selected. Information about the center of mass is also saved. No object correspondence between frames is given.

3 Object tracking

In this stage, the data produced by previous stage is analyzed and re-ordered. For each pair of consecutive frames in the sequence, objects with the minimal distance of their center of masses are considered corresponding.

Let $Y_t = \{Y_1, \ldots, Y_n\}$ be composed by the n-objects of largest size found, for the frame t, in the thematic map produced by the segmentation scheme described in Section 2. Let $L_s = \{[r_t[1], c_t[1]], \ldots, [r_t[n], c_t[n]]\}$ be the corresponding set of centers of masses of the objects in Y_t. Then, the distance between the i-th object in time $t1$ and the j-th object in time $t2$ is in general, defined as follows:

$$d(Y_{t1}[i], Y_{t2}[j]) = \sqrt{(r_{t1}[i] - r_{t2}[j])^2 + (c_{t1}[i] - c_{t2}[j])^2} \quad (4)$$

The object $Y_t[r]$ is corresponding with the object $Y_{t+1}[s]$ if their distance is minimal for all s:

$$Y_t[r] \longleftrightarrow Y_{t+1}[s] : d(Y_t[r], Y_{t+1}[s]) \; minimal \; \forall s \quad (5)$$

Given a sequence of n_f frames, the cell indicated by $Y_1[a]$ is considered completely tracked when the following sequence of corresponding objets was generated:

$$Y_1[a] \longleftrightarrow Y_2[b] \longleftrightarrow Y_3[c] \ldots Y_{n_f-1}[y] \longleftrightarrow Y_{n_f}[z] \quad (6)$$

4 Descriptors

For each tracked cell, the following descriptors were computed: Area, perimeter and Graylevel co-occurrence matrix (GLCM), with 32 graylevels, and 12 normalized features are extracted from these. The perimeter is based on 8-connectivity and GLCM looks on the 4 touching neighbors to the lower right (the pixel to the right and the 3 pixels below). The features from area and perimeter are Compacity, Fractal, Diameter, Circularity and Compactness. Most of these are different measures of the deviation from a circle.
The features from GLCM are Energy, Entropy, Maximum probability, Correlation, Contrast, Diagonal moment, Inverse different moment, Informational coefficient of correlation. There is no prior knowledge of which features could supply the wanted information hence so many are tried and later only a few will be selected.

5 Results

Segmentation
After normalization of the gradient in 256 quantization levels, two-peaks normalized histogram of the synthetical image formed by the gradient values was divided into two parts with a user-provided threshold of 20, empirically found. A structural element of 3×3 pixels, all with the value 1, was used to apply the morphological operation of dilation to thematic maps produced by the classification scheme. This was done so, after visual inspection of the results, in order to reduce the number of misclassified pixels. The gradients computed for pairs of consecutive frames in one of the sequences[2] can be appreciated in Figure 2. The overlay of the original images with the classification output using the gradients shown in Figure 2 can be seen in Figure 3.

Tracking
For the cell tracking three series have been chosen as examples and the first of these is worked through all the routines. The results in shown in Figure 4.
The tracking seems to work but for the third case the connected COGs (Center Of Gravity) looks wrong. This is due to cells leaving and entering the frame and because only the 3 largest cells are classified they are here forced to find a connection. In the colored trace for the Mitosis image (lower left image) the bottom cell is also yellow. This is due to cell split (one cell is classified as two) and treated as 2 separate cells (red and green together gives yellow).

Feature extraction
To demonstrate feature extraction only on the first series is shown. First the classified cells are shown (original image minus the background) then the enlarged cells (to observe shape changes and texture) and finally the extracted features.

[2] Codified as E362F0F2 and belonging to the group Mitosis.

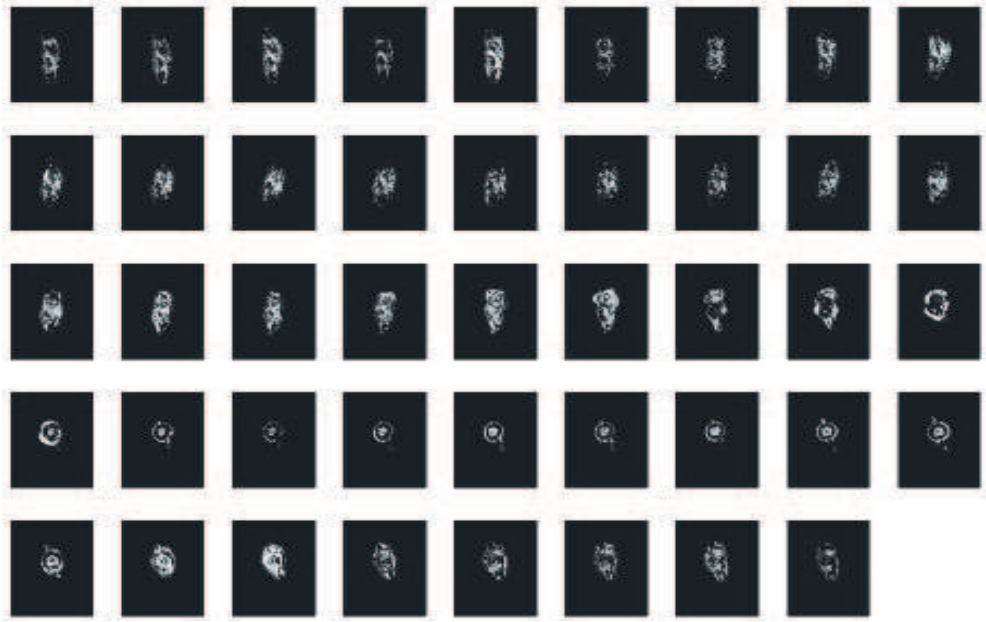

Fig. 2. Example of sequence of gradients.

This should give a good idea of the quality of the classification, transformation and features.

From the classified cells (Figure 5) it is seen that the movement here is small but the upper object (2 joint cells) change a lot. As for the rest of the figures the upper row starting with the left is the first images in the sequences.

From the first tracked cell it is seen (in Figure 6) that the cell transform is happening quite fast.

The next cell (Figure 7) hardly moves or change at all. It is seen that parts of the cell is missing in last frames. This was first seen in figure 4 where the missing part was yellow in a red cell.

Figure 8 show the third (and last) cell which is the fragments missing from the above cell. Not all the missing part is represented.

Figure 9 show the extracted features for each cell. The green curve is the small segments of the splitted cell and hence only 3 of frames have values larger than zero (results in peaks on the graphs).
The selection has not been carried out due to the needed improvement of the classification. As described above the features are extracted and tells about the cell stage but with the holes in the cells and maybe missing cell material outside the classified cell the features may change a lot during the improvement.

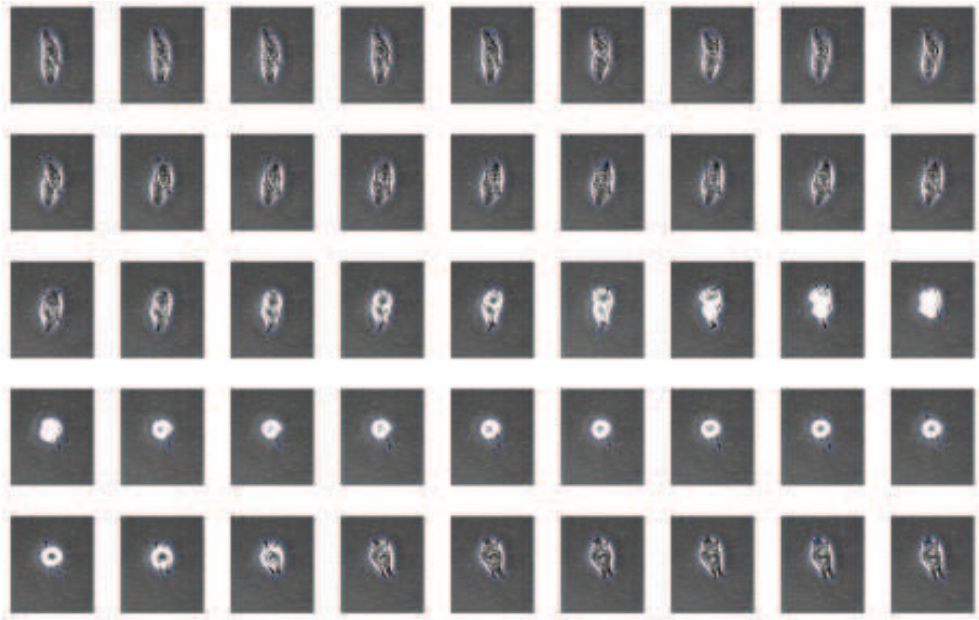

Fig. 3. Sequence of overlays (delineated in blue) of the original frames with the thematic maps produced using the gradients shown in Figure 2.

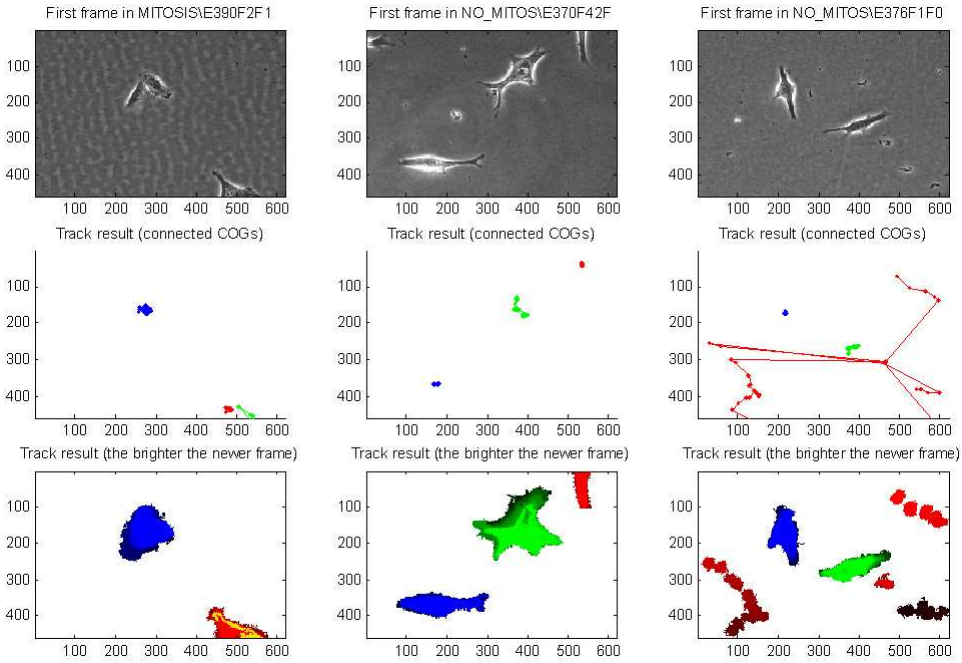

Fig. 4. Cell tracking results: Top row: The first frame in the 3 series. Middle row: The connected COGs. Bottom row: The cell area shown over time, the brighter the newer.

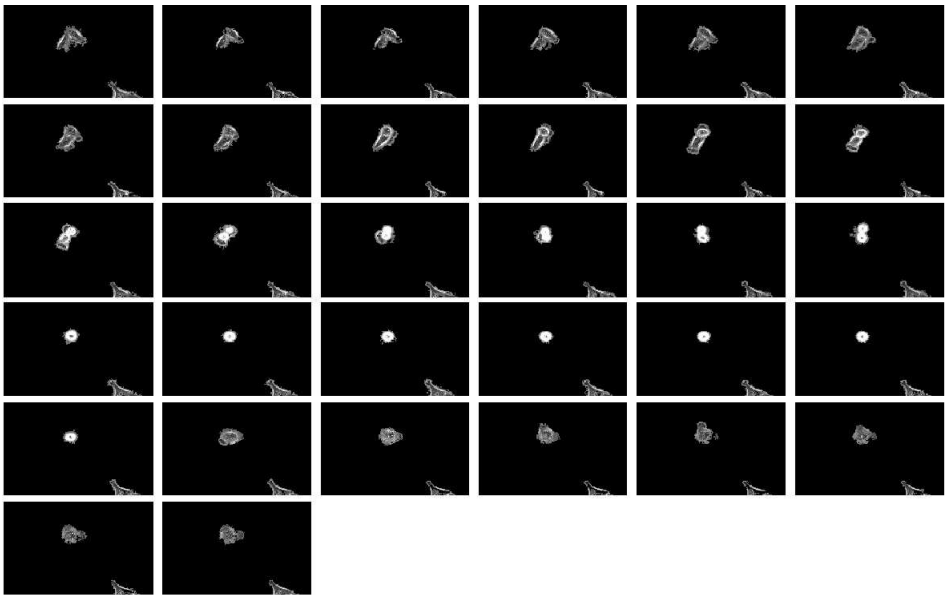

Fig. 5. All the frames from the series with the classified cells. It is not the original frames but only the classified cells that are shown. (There is only 2 cells in these frames).

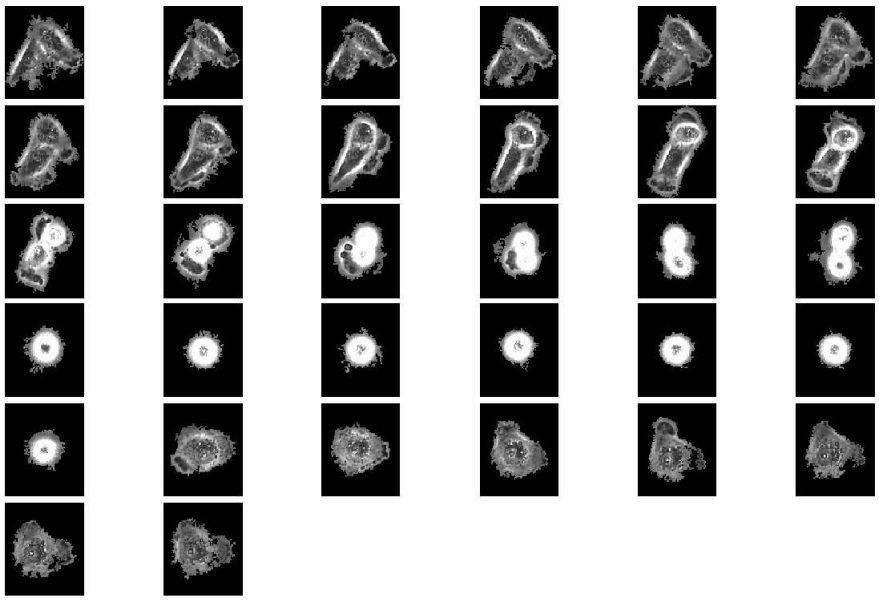

Fig. 6. The upper cell isolated and shown in equal scaling. Note the transformation speed and that it is actually 2 cells.

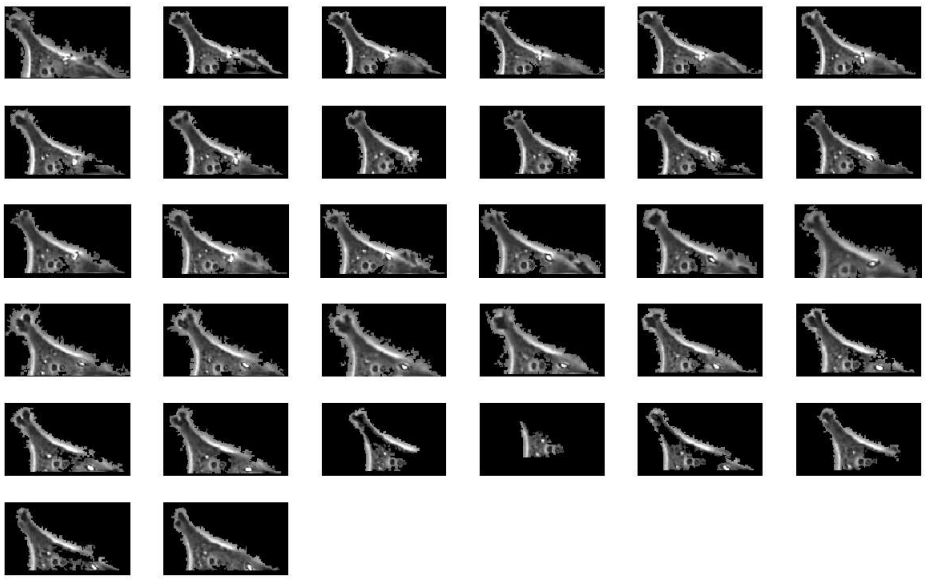

Fig. 7. The lower cell isolated and shown in equal scaling. Note the missing cell fragments in the 3 - 6 last frames.

Fig. 8. The fragments missing from the above cell.

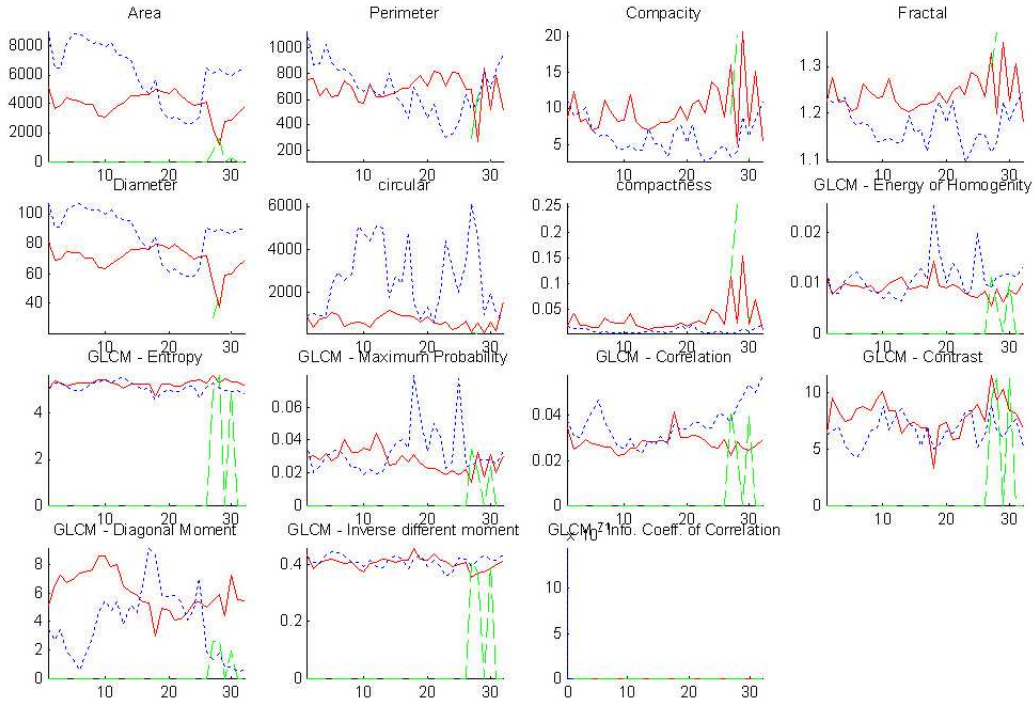

Fig. 9. The extracted features for each cell. The colors on the graphs correspond to the earlier cell colors hence the green part is the cell fragment separated from the rest of the cell and appear as peaks.

Observations

- Area is sometimes zero (not shown). This is probably due to a hole in the object and the identifying coordinate hits this hole / black spot.
- A cell can be classified as 2 individual cells.
- Classifying the 3 largest objects doesn't give the best result and should rather be all objects larger than ~ 200 pixels.
- Most cells moves in random directions, which could make prediction the position in the next frame fail.
- Some cells transforms fast which could give problems with tracking using features!
- Looking at changes from frame to frame can result in that the background is classified as a part of a cell. This can happen when the cell moves away from an area which results in a gradient.

6 Part 2

From the above observations and further experiments the second version of the segmentation, tracking and feature extraction consists of the following.

6.1 Improved segmentation

Now the frame is segmented with a object size threshold of 1000 pixels thus the possibility of a varying number of cells from frame to frame.
As with the first segmentation a closing with a circle of 5 pixels in diameter is applied followed by filling holes in the object and removing single pixels from the tematics.

Additional 3D closing is applied (the tematics is stacked to a volume) with a circle of 11 pixels in diameter and afterwards holes filing. This has been added to smoothen the border and to keep cells from splitting up or have cracks in them.

6.2 Improved tracking

For every cell in a frame the distance to all the cells in the next frame is calculated. The smallest distance always wins and other bids to the same cell is classified as a new cell from that frame forth. Now each cell gets a unique number / ID so a new cell wont get a leaving cells number. To correct cells missing due to size or leaving the area for a couple of frames, all entering cells are compared with cells left earlier and if their COGs are within a threshold they are merged.

6.3 Extra features

The get more information on the shape of the objects 2 new features are added.

Oblongness
Oblongness is the ratio of the length of longest direction of the cell divided with the length of shortest direction. This results in ~ 1 for a circle and grows with the oblongness. The length is measured in 4 directions; Vertical, horizontal and the 2 diagonals.

$$Oblongness = \frac{max(l)}{min(l)}, \tag{7}$$

where l i the length in the different directions. From [3].

Shape
Here the number of pixels in each connected row or column is placed in a matrix (one for each). They are used to generate 2 new matrixes, one with the largest

number of pixels in connected rows or columns for each pixel and another with the shortest (se Figure 10). The area is divided with the square of the total mean of these matrixes. The measure should be a more unique number describing the shape of the object, according to [1].

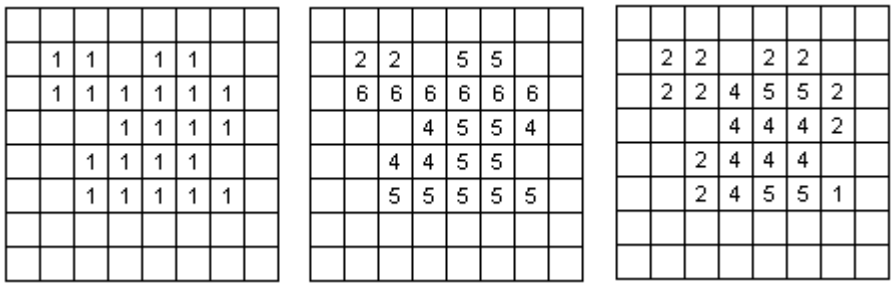

Fig. 10. Shape feature. From left: Original shape, Maximum(row,col) and Minimum(row,col).

$$Shape = \frac{A}{\left(\frac{\overline{max}+\overline{min}}{2}\right)} \tag{8}$$

7 Results of part 2

The improved routines are used on all the sequences and the results are described below.

7.1 Segmentation

This is a give and take situation. The cell split has been removed and holes and cracks as well. The drawback is that the background classified as cell has grown a bit, specially in cases of movement or cell change, where the 3D closing smoothen the cells and hence cover some background pixels. Due to 3D closing neighbor objects have the possibility to grow together, see Figure 11. This is only a problem in one sequence where there are a lot of small non-cells in the frames (which can be observed in the figure, between the two cells).

The touching cells are classified in joint groups (as expected) but the problem here with additional background is almost non-existing probably due to larger contrast between cells and background.

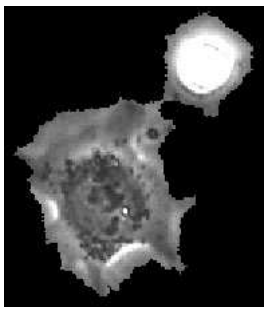

Fig. 11. Cells grown together. Note the small object between the cells.

7.2 Tracking

The improved tracking works without exceptions on No-mitosis and Mitosis with varying number of cells and temporary missing cells. Many of the cells doesn't move much but in Figure 12 some movement can be observed. The figures shows the connected COGs and colored trace respectively for the same sequence.

7.3 Descriptors

The 2 new features (Oblongness and Shape) seems to work and the problem in part 1 with holes in the objects etc. has been removed. All the features suffer

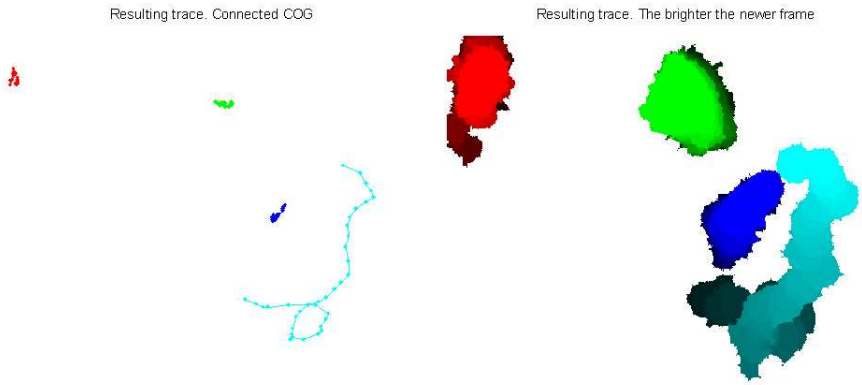

Fig. 12. No-mitosis sequence. Left: Connected COGs. Right: Colored trace.

from the same problem namely that there is some background in the cell class. This doesn't mean that the features can't be calculated but that the are that the values are smoothed from frame to frame. It can also result in peaks like in "GLCM - maximum probability" where a graylevel in the background becomes dominant.

Figure 13 (No-mitosis) shows the segmentation of one of the cells trough the sequence. In 2 of the frames the misclassified background is rather large. The corresponding features can be see in Ffigure 14. The cell transforms during the sequence which can be observed in the features Diagonal moment, Oblongness and Shape. Peaks are seen in Maximum probability and Homogeneity but this is due to the two cells with "extra background". This background consists of more uniform graylevels than the cell itself and therefore these descriptors are here not good measures of change.

The next example is a Mitosis sequence with cell transformation, se Figure 15. It starts as a normal cell (it actually consists of 2 cells, but thats besides the point.) and transform in to a bright almost circle shape and back again. The features extracted can be observed in Figure 16. The area of cause drops during the mitosis and Contrast and Shape as well. Maximum probability and Homogeneity rises. Oblongness gets a bit larger which at first looks strange when the shape turns to a circle but the point it that is not quite a circle due to a "tail" to the left of the circle making it oblong.

8 Selected features

From these few examples it is hard to tell which features is a good choice but looking through the sequences it sums to:
The best describing features are Contrast, Diagonal Moment and Shape. And

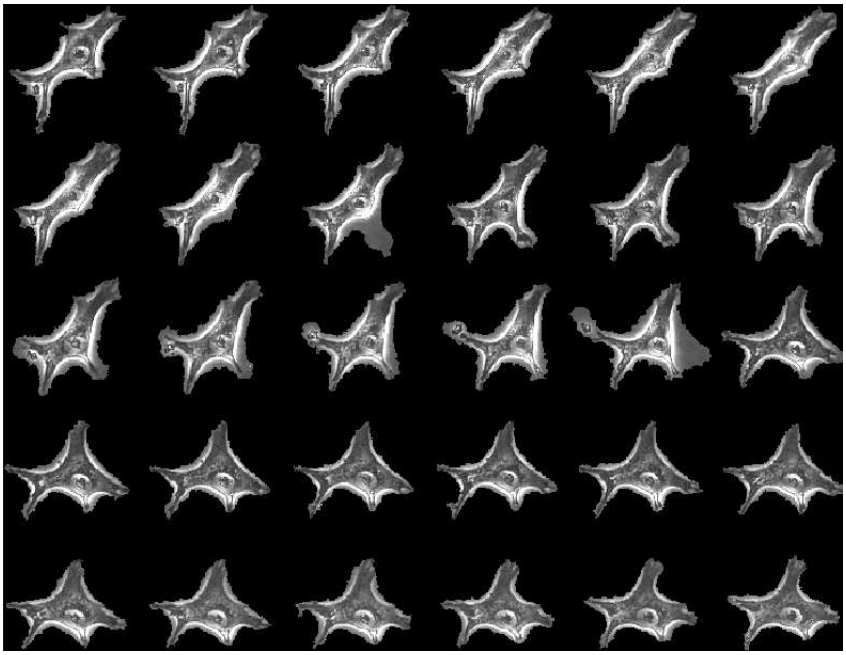

Fig. 13. Segmentation result of one of the cells trough a No-mitosis sequence. Note the 2 cells with large background attached.

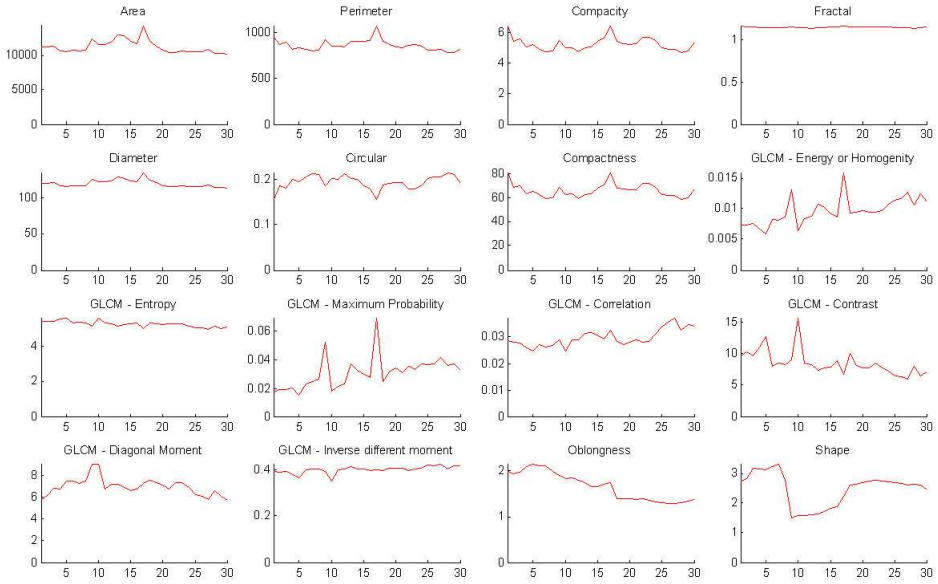

Fig. 14. Features calculated of the above shown cell.

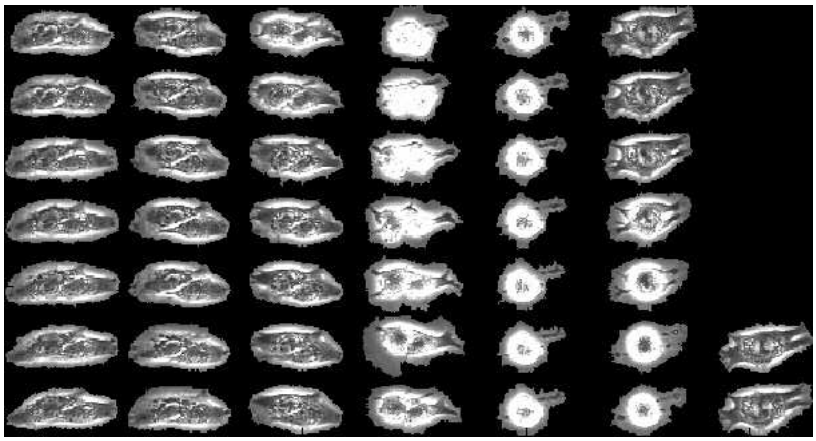

Fig. 15. Cell from a mitosis sequence. It is actually 2 joint cells. (The image is lying down which means that the left column starting at the bottom is the first cells.)

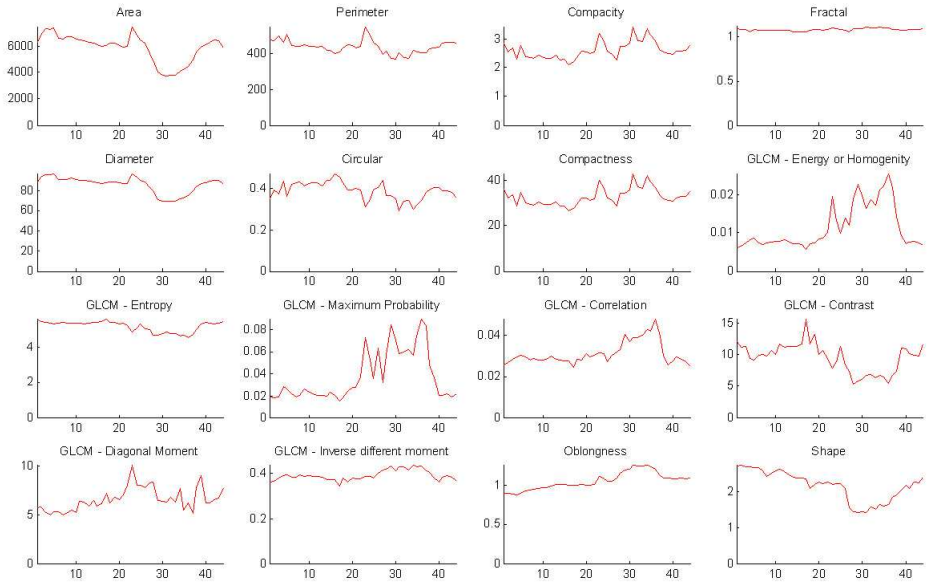

Fig. 16. Features calculated of the above shown cell.

Homogeneity and Maximum probability will also be good texture descriptors if there are less background in the cells.

9 Discussion

The main problem is the segmentation / classification of cells. Here is some suggestions of improvement:

- Look at the gradient again. As it is now the threshold is not to high. A higher threshold could result in less background classified as cell. A try could be with out the gradient between frames, or without normalization.

- For a better segmentation [2], [3] have theory about Edge detection, and [3] also describes Ultimate erosion and Morph segmentation which could be worth a try. [3] describes motion analysis and Kalman filters.

- Another way of segmentation is to describe the background and deviations from it would be cells. 3D FFT could be a way to look a the background noise and to design a filter to remove it.

- Calculate the mean of all the frames and subtract it for the frames. Needs a steady background and that the cells don't occupy the same place for more than approx. half the frames (which will be a problem in most of the sequences!)

- The shape of the cell could be used to predict the ROI in the next frame with a different threshold, but as described the threshold should not be lower that it is now.

To distinguish adjacent cells (Touching), erosion could be a solution. 2 or 3 joint cells are continually eroded and ends up with 2 - 3 points, one in each COG. Now the Euclidian distance is used to join / classify the rest of the pixels in the joint object to there respective cell represented by each COG.

Add a home made measure of change, by looking at the number of new / removed pixels and the change in each pixels graylevel.

10 Conclusions

There are errors (connected cells, background around a cell perceived as the cell) which can all be traced to the classification of the cell and closing (adds adjacent cells and result in "shadows" of moving cells). There simply is too much background in the classifications. Maximum probability (a GLCM feature) is a good measure of the error because the background is more uniform and hence more "dominating" graylevels.
Tracking the cells however is functioning perfect with the these data sets. If more than one cell is moving in large steps from frame to frame the tracking here would probably fail (could happen if another cell in the next frame is closer to the first cells COG in the earlier frame). Here direction and speed could predict a possible

region of the next position of each cell, but the direction is often random.
A couple of feature (Contrast, Diagonal moment and Shape) have been pointed out to exert the best description of non-moving cell change. However a better classification could probably improve the rest of the tested features.

Acknowledgments

The present work was carried out as a requirement for the Ph.D. course "Medical Image Analysis" given by Prof. Milan Sonka at the Technical University of Denmark on October, 2002. We thank Prof. Milan Sonka for providing the cell image sequences used and to the SITE Project funded by a grant from the Danish Technical Research Foundation (Project Number *STVF* 56-00-0123) for partially supporting the present work.

Appendix

References

1. J. Lira, *Introducción al Tratamiento Digital de Imágenes*, Instituto Politécnico Nacional, Universidad Nacional Autónoma de México, Fondo de Cultura Económica., September 2002, México D.F.
2. M. Sonka, J. Michael Fitzpatrick, *Medical Imaging. Volume 2. Medical Image Processing and Analysis*, 2000, United States of America.
3. M. Sonka, Vaclav Hlavac, Roger Boyle, *Image Processing, Analysis and Machine Vision*, 1999, United States of America.