Protocol: Cell Segmentation and Linking.

Summary: This protocol describes the segmentation of the phase contrast image to determine regions corresponding to cells and the linking of these regions to determine cell boundaries.

1. Determination of putative cell boundaries.
   a. Identification of Microcolonies/Cell Clumps
      i. Images are smoothed and then thresholded to determine the boundaries of regions containing clumps of cells
   b. Segmentation of clump into individual cells
      i. Phase image is smoothed
      ii. The resulting Image is processed using the magic contrast filter
         1. Magic Contrast: At each pixel, the minimum intensity value in a region, centered on the pixel of interest, radius $r_m = 6$ pixels, is subtracted from the pixel of interest.
      iii. Next the image is thresholded upward
      iv. A watershed is then applied and masked by the cell clumps mask.
   v. All internal boundaries between putative cell regions are divided into segments.
   vi. Twenty characteristics of each segment are computed, including second derivative of the phase image over the segment, length, mean intensity, etc.
   vii. A probability of existence is assigned to each boundary segment using a Maximum Likelihood Model (MLM) evaluated on the segment characteristics.
   viii. All segments with existence probabilities above 99% are turned on all segment with existence probabilities below $1 \times 10^{-4}$ are turned off and the remaining segments are resolved in the next step.
   ix. The remaining boundaries are analyzed using a MLM which considers the shape of the resulting regions. Ten region characteristics are computed for each putative region. The combined Likelihood of the segments and regions are maximized together to determine to determine the cell boundaries.

2. Training of the Maximum Likelihood Model (MLM)
   a. To train the MLM, we assigned the existence of segments and regions by hand on a training data set, and then we optimized a parameterized model to predict the probability of a given boundary exists and the probability that a particular region is a cell.
   b. To get optimal performance from the algorithm, one typically has to train the algorithm for cell species, growth conditions, pixel size and magnification.

3. Region Linking
a. To generate cell trajectories, segmented regions must be linked between successive frames.
b. Regions with the most overlap between successive frames are linked.
c. After linking, Errors and inconsistencies are resolved. For instance, the splitting of one region into two is only allowed if this splitting persists for more than a frame. If the splitting is not persistent, the regions are merged to resolve the lining error.

4. **Keep only cells which are track without errors from division to division.**

<table>
<thead>
<tr>
<th>A.</th>
<th>B.</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Phase-Contrast Image" /></td>
<td><img src="image2.png" alt="Segmented Frame Image" /></td>
</tr>
</tbody>
</table>

**Figure 1:** (A) Phase-Contrast and (B) Segmented Frame Image. Cell regions are defined by region boundaries, shown in red, blue, and yellow in the segmented image. Red represents boundaries with a high probability of likelihood, blue represents boundaries with low probability of likelihood, and yellow represents a new boundary not present in the previous frame, indicating a cellular division event.