From Cells, to Images, to Numbers, Reliably

Erik Meijering

mRNA, vol. 2, no. 2, April 2017, pp. 18-20

Breakthroughs in science often go hand in hand with revolutions in technology. This is certainly true for biology. One branch of technology in particular – light microscopy – has consistently played a crucial role for over three centuries now in unraveling the miraculous workings of living cells. Our fellow Dutchman Antoni van Leeuwenhoek, using his own carefully handcrafted microscopes, was the first to observe microorganisms in action in a drop of rain water in the mid-1670s. And in the years to follow he continued perfecting his lenses and making one groundbreaking discovery after another. The growing skepticism about his incredible findings even made the Royal Society in London send a delegation of prominent British and Dutch clergymen to his home in Delft to see things for themselves.

Our reliance on visual evidence does not come as a surprise. Vision happens to be the most dominant of our five physical senses, involving more than half of our brain directly or indirectly. Thus it's only natural that we prefer to *see* things to get convinced. 'Seeing is believing', says the age-old adage. But how reliable is our visual system, really? It does not take many optical illusions to realize our eyes are easily deceived (**Figure 1**). And even if they are not, how accurate are we in estimating biophysical properties of the objects we observe in microscope images? If two persons are asked to manually measure, say, the average speed of certain fluorescently labeled proteins moving within a cell in a time-lapse recording, they likely get two different answers. Worse, if one person is asked to repeat the same measurement, the result

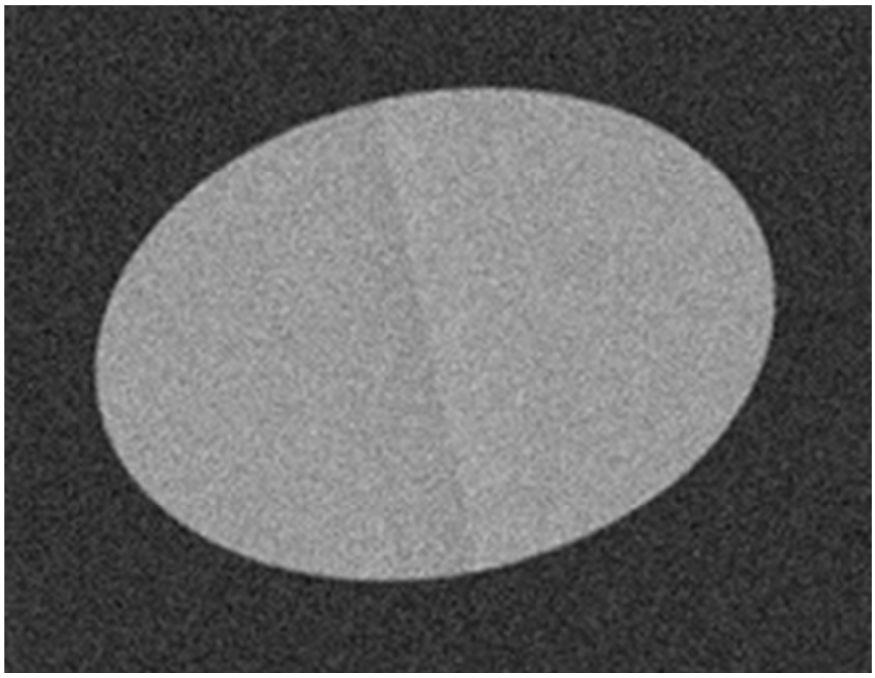

Figure 1 Which side of this 'cell' is brighter? Make a first guess, then put your finger across the centerline, and guess again.

will still be different each time. And if there are hundreds of objects to measure, in hundreds of images or movies, it does not take long before we get bored, tired, and sloppy. So we tend to make selections and ignore the rest of our data. But based on what criteria? And how can we be sure these do not bias our findings? Clearly, when it comes to measuring things objectively, reproducibly, and completely, human beings (even experts) quickly fail, and more rigorous means are needed.

Enter computers. They are relatively inexpensive, to buy as well as to feed, and they can process huge amounts of data and give the same answer when asked the same question, without ever getting bored or tired. One little problem: computers by themselves know absolutely nothing about microscope images (or anything, for that matter). They need to be 'educated', and this is where my research comes in. In my group we develop computer algorithms that help biologists get reliable numbers out of their images. This may involve many steps (**Figure 2**). The first is usually to preprocess the images to reduce noise, blurring, and other imaging artifacts. A next common step is to detect the presence of objects by using specific image filtering techniques. Often the most challenging step is segmentation, which aims to label and group image pixels as objects (relevant) or background (irrelevant). In the case of time-lapse images, a related problem is to determine the trajectories of moving objects by frame-to-frame association of image

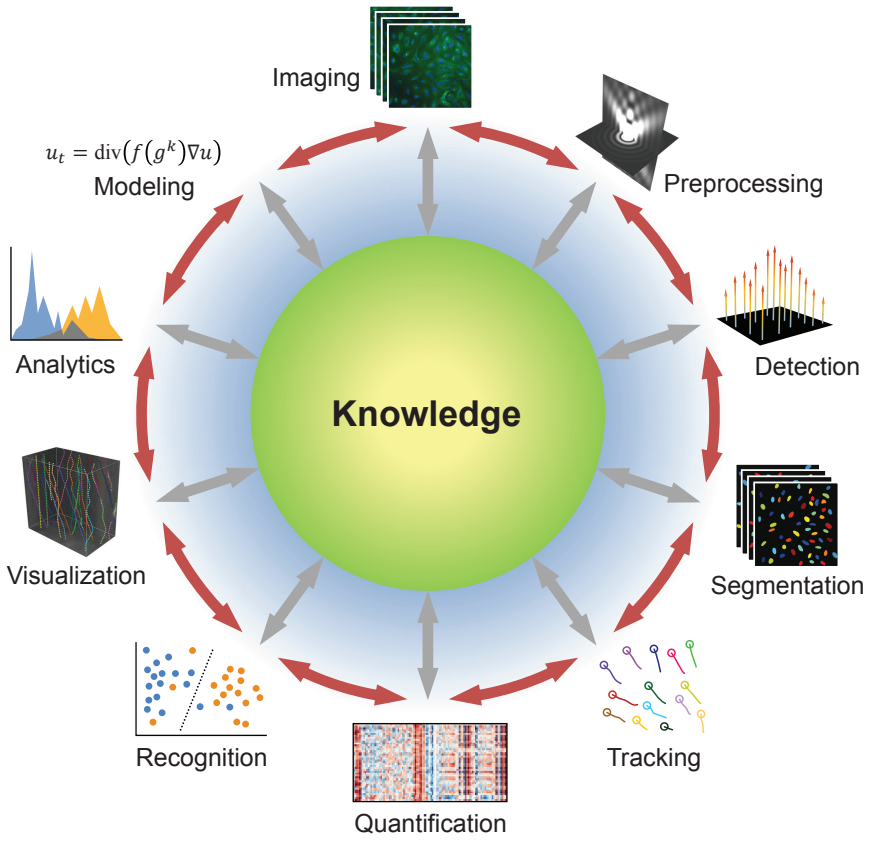

Figure 2 Common steps in computational image analysis. The ultimate goal is, of course, to increase our knowledge of biological processes. Depending on the experiment not all steps may always be needed (shortcuts are possible). Two-headed arrows are used everywhere to indicate the interrelation and possible feedback between steps.

information. Once this is done, it is relatively easy to compute many quantitative descriptors of object dynamics, shape, and texture. This enables the computer to characterize and distinguish objects using pattern recognition techniques. Various visualization techniques are available to show results to the user. And rigorous analytical techniques for statistical testing help in finally composing biophysical models of the (intra)cellular processes under investigation.

Well-designed computer algorithms for automated image analysis enable biologists to extract invaluable information from their data and make discoveries that would otherwise be impossible. For example, a few years ago, a consortium of European research groups successfully determined which of our ~21,000 human genes play a role in mitosis. In a series of RNA interference experiments they made ~190,000 time-lapse fluorescence microscopy images of in total ~19,000,000 cell divisions. The use of state-of-the-art techniques for segmenting the cell nuclei in all images, tracking them over time, extracting hundreds of quantitative features from them, and finally classifying them into different morphological classes related to the cell cycle, cell death, or other phenotypic changes, was absolutely crucial to the success of the study. In our own collaborations with biologists, too, we have developed a broad range of powerful image analysis techniques, in particular for particle detection, tracking, and statistical processing of their intensities in fluorescence microscopy movies. Example applications include the study of microtubule dynamics (**Figure 3**) and of proteins involved in DNA damage repair (discussed by Claire Wyman in mRNA 1.3).

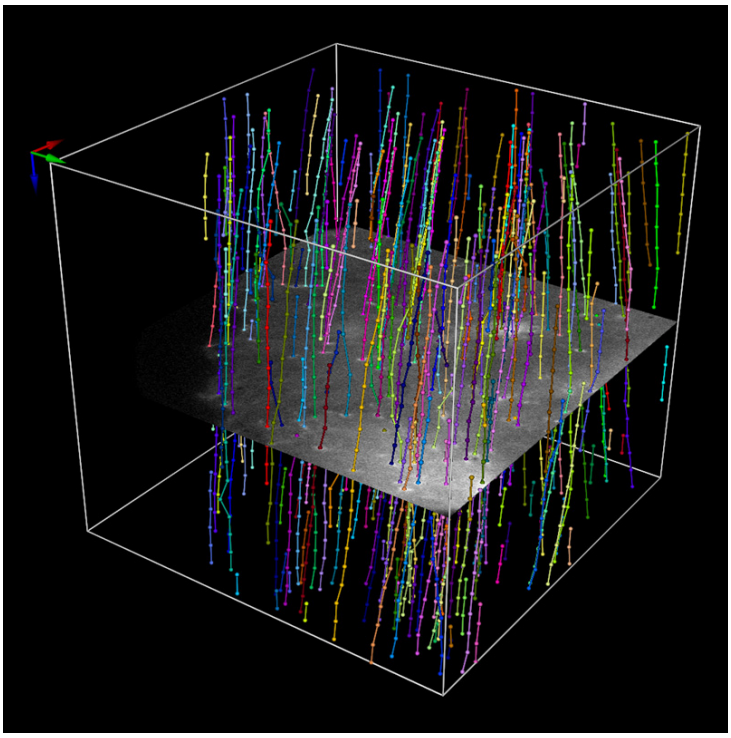

Figure 3 Spatiotemporal visualization (time runs along the vertical axis) of the trajectories of microtubule tips moving in the cytoplasm. The trajectories were extracted completely automatically using particle tracking techniques developed in our group. The visualization also shows one time frame of the original movie.

Where science meets technology, exciting things happen. Being part of an interdisciplinary educational program that fully recognizes this can be a challenge, but above all it is a great opportunity. No doubt you will see and contribute to major breakthroughs in your own career, improving the lives of many!