2013 IEEE 10th International Symposium on Biomedical Imaging:
From Nano to Macro
San Francisco, CA, USA, April 7-11, 2013

3D HAAR-LIKE ELLIPTICAL FEATURES FOR OBJECT CLASSIFICATION IN MICROSCOPY

Fernando Amat, Philipp J. Keller

Howard Hughes Medical Institute, Janelia Farm Research Campus, Ashburn, Virginia, USA

ABSTRACT

Object detection and classification are key tasks in computer vision that can facilitate high-throughput image analysis of microscopy data. We present a set of local image descriptors for three-dimensional (3D) microscopy datasets inspired by the well-known Haar wavelet framework. We add orientation, illumination and scale information by assuming that the neighborhood surrounding points of interests in the image can be described with ellipsoids, and we increase discriminative power by incorporating edge and shape information into the features. The calculation of the local image descriptors is implemented in a Graphics Processing Unit (GPU) in order to reduce computation time to 1 millisecond per object of interest. We present results for cell division detection in 3D time-lapse fluorescence microscopy with 97.6% accuracy.

Index Terms — light microscopy, object classification, local descriptors, Haar features, cell division

1. INTRODUCTION

The amount of image data in the life sciences is growing exponentially and in many experiments it is impossible to extract useful biological insights without the help of efficient statistical image processing algorithms [1]. Automatic or semi-automatic approaches are required in order to screen information efficiently and design high-throughput experiments to extract quantitative information from voxel intensities. In many cases, object classification is one of the key image processing tasks in such high-throughput pipelines.

Cell lineage reconstruction from three-dimensional (3D) time-lapse fluorescence light microscopy is an example of such a scenario. The ultimate goal is to reconstruct the location of each individual cell in space and time for complex organisms such as entire *Drosophila* and zebrafish embryos. Aside from segmentation and tracking, cell division detection is a key task required for accurate cell lineage reconstruction, since misclassification of dividing cells immediately disrupts the topology of the cell lineage tree [2].

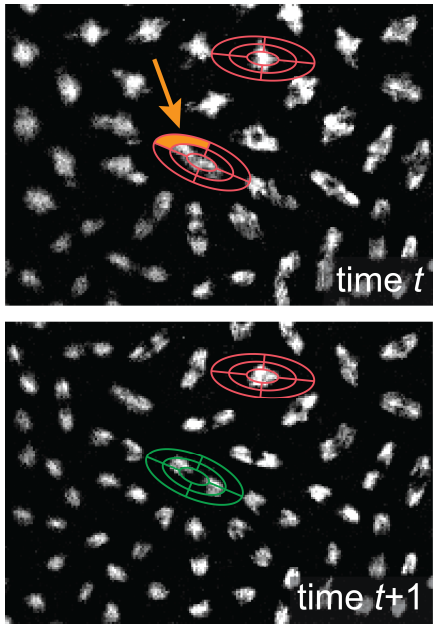

Figure 1. Maximum intensity projections of two time points in an *in vivo* recording of *Drosophila* embryogenesis, obtained with SiMView light-sheet microscopy. Superimposed ellipsoids indicate two nuclei tracked through time using a Gaussian parametric shape model for each nucleus. The concentric ellipsoids indicate the division of the space to calculate each Haar-like feature. Orange arrow points to one specific radial and angular sector. Red color indicates non-dividing cells and green color indicates dividing cells. This figure is just for illustration purposes, since data and features discussed in this paper are 3D.

2. RELATED WORK

Object detection and classification are key tasks in computer vision. Many local image descriptors, such as SIFT [3] and Histograms of Oriented Gradients (HOG) [4], have been derived to facilitate these tasks by incorporating relevant local image information in an array of features related to points of interest. However, extension to 3D does not always produce the same successful results [5], since the

number of degrees of freedom for object appearance in a local neighborhood increases, making it harder to summarize all local information in a small set of features.

One of the most popular approaches for object detection is the set of Haar-like features popularized by the seminal paper of Viola and Jones on face detection [6]. However, its popularity can be attributed more to its fast implementation through the integral image rather than its discriminative power [7]. Some improvements have been suggested to increase the expressiveness of Haar-like features, such as 45 degrees rotation in order to capture tilted objects [8] and addition of edge information in the image in order to incorporate shape information [9]. These ideas have also been successfully extended to 3D biomedical imaging applications [10]. Inspired by all these different contributions, we present an efficient implementation of 3D Haar-like elliptical local image descriptors computed in a sparse set of points of interest in the image. Our method uses ellipsoidal descriptions of objects and Difference of Gaussians (DoG) filters in order to incorporate arbitrary rotation, scale and shape information into a local image descriptor.

Skibbe *et al.* [5] have proposed multiple local descriptors specifically designed for 3D microscopy data based on the theory of spherical harmonics and spherical Gabor basis functions. Their results in different bioimaging pattern recognition problems, such as cell division detection, show that good local descriptors are a key component of the solution. However, they have focused on approaches to efficiently extract dense local features, which can be prohibitive in terms of memory and time consumption for the large 3D datasets generated in light microscopy experiments.

In our experience, most fluorescence light microscopy image stacks are sparse (over 90% of the volume can be background) because only a fraction of the specimen volume contains the fluorescent probes that mark objects of interest. Moreover, the signal-to-noise ratio (SNR) is often sufficiently high to enable the use of simple techniques, such as image thresholding, to reduce the number of points of interest to a manageable number.

3. METHOD

3.1-Angular and radial sectors

In our approach, we assume that the local neighborhoods for objects of interests can be described as 3D ellipsoids (or blobs), which can be parameterized by 9 coefficients (3 for the center and 6 for the covariance matrix describing the quadratic surface). This kind of representation is very common in 2D computer vision approaches, since ellipses are a compact representation of object position, scale and orientation and are directly linked to Gaussian distributions. This statement is also true in the case of 3D microscopy images, especially since many biological features are below the diffraction limit and appear in the image as blurred elliptical spots due to the point-spread function (PSF) of the microscope.

Using a sphere, we can decompose the local image space into different sectors based on angular and radial information (Fig. 1). In our case, we first generate R different shells at different radial components. Shell R_i contains voxels at a radial distance between i and i+1 standard deviations of the multivariate Gaussian. Second, we partition each radial shell into A angular sectors using the Hierarchical Equal Area isoLatitude Pixelization (HEALPix) of a sphere [11], in order to obtain sectors of equal volume within each of the radial shells. The only exception is the central radial sector (i = 0), which we do not split into angular sectors due to lack of angular resolution in an image grid. The ellipsoid defining the point of interest allows us to easily map a partition of the inside volume of a sphere into a local neighborhood of the point of interest using the eigenvector decomposition. The total number of sectors is 1 + A (R-1).

3.3-Haar-like features

Once the basic sectors are defined, we can use these to define Haar-like features. For each sector we can calculate scalar values, such as average image intensity, in order to obtain local image features. At the same time, we can combine all these sector features by addition or subtraction in order to obtain an exponentially growing number of higher order features. In our case, we select the following Haar-like combinations (Fig. 2):

- *Type 1:* Average intensity in each radial shell. This type adds R features.
- *Type 2:* Average intensity in each sector. This type adds 1 + A (R-1) features.
- *Type 3:* Subtraction of any possible pair of intensities described in type 3. This type adds 1 + A (R-1) choose 2 features.
- *Type 4:* Generation of intermediate features by addition of any possible pair of intensities described in type 3. Then, subtraction of the results of each possible pair of intermediate features. This type generates the largest set of features, adding 1 + A (R-1) choose 2 choose 2 new features.
- *Type 5:* The ratio of the length of each possible pair of semi-axis of the ellipsoid as a measure of eccentricity. This type adds 3 features.

By joining all features into a single 1D vector and defining H = 1 + A (R-1) choose 2, we generate an array of length [R + 1 + A (R-1) + H + H choose 2 + 3]. Any method

for feature selection can be used later on to decide which of these features are most relevant to the task at hand.

One common objection to Haar-like features is that they are not as discriminative as other techniques. The ellipsoid orientation and scale information already improves this aspect. We also use the scale information in order to improve the discrimination power of our set of local image features by calculating Haar features over the gradient image, which effectively includes shape information from the object and dramatically improves the expressivity of the features. Thus, we double the length of our array of features by calculating the same set of Haar-like features over the image convolved with a Difference of Gaussians (DoG) filter. The scale of the filter is defined by the size of the ellipsoid in each principal direction.

3.2-Implementation details

The calculation of the scalar values for the $1 + A\,(R-1)$ angular and radial sectors can be efficiently calculated using a Graphical Processor Unit (GPU). In particular, the interpolations required for ellipsoidal sectors and the separable convolution performed for the DoG are highly accelerated by the GPU architecture.

For all results presented in section 4, we use trilinear interpolation, A equals 12 (level 0 of HEALPix) and R equals 3, which results in a total of 76,529 3D elliptical Haar-like features. We also normalize intensity levels in a box around each ellipsoid to obtain an average of zero and a standard deviation of one in order to ensure robustness against intensity changes, which are very common in microscopy images of samples recorded in different experiments.

4. RESULTS

We evaluated our approach in scanned light-sheet microscopy datasets for the purpose of detecting cell divisions (Fig. 1). Light-sheet microscopy provides exceptionally high imaging speeds while minimizing the energy load on the biological specimen. This combination of capabilities is invaluable for live imaging applications and enables quantitative imaging of cellular dynamics throughout the development of complex organisms such as entire *Drosophila* embryos. Cell divisions are of central importance throughout the development of multi-cellular organisms and their automatic detection through image processing is crucial in a wide spectrum of biological applications, such as complete cell lineage reconstructions.

All results presented in this section were obtained using previously published microscopy datasets of *Drosophila* embryos [12]. In total, we annotated 5,764 blobs for different time points, indicating if these blobs contained a nucleus that was dividing or not. 720 of these annotations (12%) were positive samples and the rest were negative. This difference in positive versus negative examples is expected, since cell divisions are typically rare events. We split the data into two sets, comprising roughly one-third of the volumes for testing and two-thirds for training.

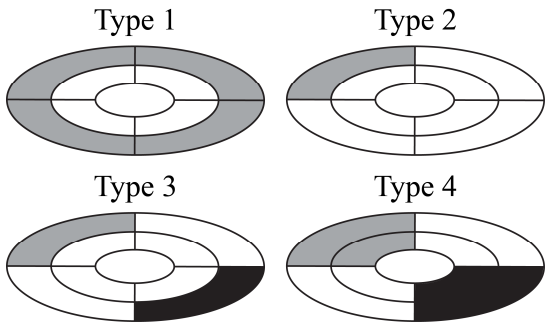

Figure 2. Schematic illustration in 2D of our elliptical Haar-like features. The actual division in radial and angular sectors is performed in 3D using HEALPix [11]. Gray and black shadings indicate positive and negative signs, respectively. See main text for a full description of each feature type.

Fig. 3 shows precision-recall curves on the test data for three different methods: first, our set of 3D elliptical Haar-like features applied only to the image intensity. Second, twice as many features by calculating our 3D elliptical Haar-like features on the image intensity and the output of the DoG filter. In both cases we train a classifier using GentleBoost [13] with 400 weak learners. Each weak learner is a tree with three levels, resulting in 426 features selected for classification out of the 76,529. Finally, for comparison purposes, we show results with the method described in [14] specifically targeting cell divisions, where the Kullback-Leibler divergence is used to decide if the blob is better represented with a single Gaussian or a mixture of two Gaussians.

Fig. 3 shows the advantage of using our image features in order to detect cell division. In particular, the shape information encoded in the DoG filter helps improving accuracy. Using our GPU implementation on a Tesla C2070 (Nvidia), we calculated all 76,529 features in 1 millisecond per blob of interest, including data transfer of a cropped image box of size 50 x 50 x 50 voxels around the point of interest. The array of 76,529 image features in single-point precision has a size of 0.3 MB. Thus, even when locating on the order of 1,000 objects of interest in a microscopy image, we can calculate all necessary image features in 1 second and store these in an array with total size of 300 MB. Given that most fluorescent microscopy images contain large amounts of background and that the SNR is sufficient to select points of interest using fast image processing techniques such as detection of local intensity maxima, we tend to favor descriptors on a sparse set of blobs versus dense descriptors when analyzing 3D microscopy volumes.

An earlier version of the methodology presented in this paper was part of the tracking and segmentation pipeline used in [12] to reconstruct individual cell lineages of complex multicellular organisms such as *Drosophila.* The approach presented in [12] considered only features of type 1, 2 and 5 and was not implemented in a GPU, which increased the computational time 30-fold. This performance difference is crucial, since it represents the difference between being able to identify cell divisions on the whole-embryo level in real-time or not. Moreover, the accuracy of 97.6% for cell division detection reported here is substantially improved over the accuracy of 93.8% in [12].

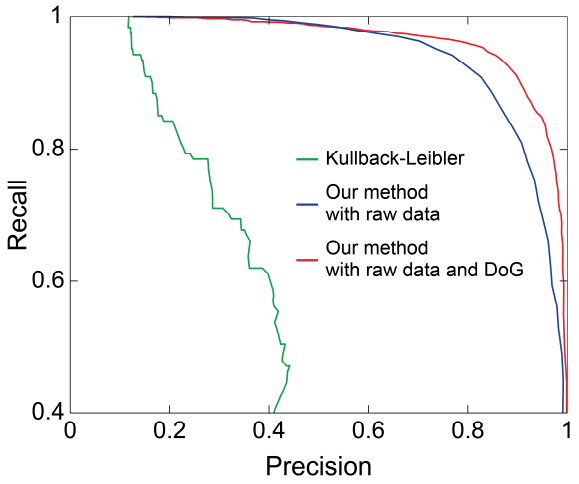

Figure 3. Precision-recall curves for cell division detection in the light-sheet microscopy test set. Training set contains 3,843 samples. Test set contains 1,921 samples. Curves are shown for our features in raw image data (blue), our features in raw image data and DoG enhanced data (red), and for Kullback-Leibler measure [14] (green).

5. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a systematic construction of 3D Haar-like features adapted to points of interest with local neighborhoods described by ellipsoids, which contain position, scale and orientation information. We also showed that adding edge information from DoG filtering to the features helps in its discrimination power. Finally, a fast GPU implementation was discussed in order to facilitate the application of these features to large 3D microscopy datasets. In the future, these features can be extended to 3D + time by tracking the point of interest as a function of time or, in the case of cell division, they can be incorporated in a temporal graphical model framework, such as the one presented by Huh *et al.* [2], in order to improve detection accuracy.

6. REFERENCES

[1] A. Cardona and P. Tomancak, "Current challenges in open-source bioimage informatics," *Nature Methods*, vol. 9, no. 7, pp. 661–665, 2012.

[2] Seungil Huh, Sungeun Eom, R. Bise, Zhaozheng Yin, and T. Kanade, "Mitosis detection for stem cell tracking in phase-contrast microscopy images," in *ISBI*, pp. 2121–2127, 2011.

[3] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *IJCV* , vol. 60, no. 2, pp. 91–110, 2004.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *CVPR*, vol. 1, pp. 886–893, 2005.

[5] H. Skibbe, M. Reisert, T. Schmidt, T. Brox, O. Ronneberger, and H. Burkhardt, "Fast Rotation Invariant 3D Feature Computation utilizing Efficient Local Neighborhood Operators," *TPAMI*, vol. 34, no. 8, pp. 1563–1575, 2012.

[6] P. Viola and M. J. Jones, "Robust Real-Time Face Detection," *IJCV*, vol. 57, pp. 137–154, 2004.

[7] A. Coates, P. Baumstarck, Q. Le, and A. Y. Ng, "Scalable learning for object detection with GPU hardware," in *ICIRS*, pp. 4287–4293, 2009.

[8] R. Lienhart and J. Maydt, "An extended set of Haar-like features for rapid object detection," in *ICIP*, p. 900–903, 2002.

[9] K. Levi and Y. Weiss, "Learning object detection from a small number of examples: the importance of good features," in *CVPR*, vol. 2, 53–60, 2004.

[10] G. Carneiro, F. Amat, B. Georgescu, S. Good, and D. Comaniciu, "Semantic-based indexing of fetal anatomies from 3-D ultrasound data using global/semi-local context and sequential sampling," in *CVPR*, pp. 1–8, 2008.

[11] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere," *The Astrophysical Journal*, vol. 622, pp. 759–771, 2005.

[12] R. Tomer, K. Khairy, F. Amat, and P. J. Keller, "Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy," *Nature Methods*, vol. 9, no. 7, pp. 755–763, 2012.

[13] J. Friedman, T. Hastie, and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting," *Annals of Statistics*, vol. 28, p. 2000, 1998.

[14] G. Xiong, C. Feng, and L. Ji, "Dynamical Gaussian mixture model for tracking elliptical living objects," *Pattern Recognition Letters*, vol. 27, no. 7, pp. 838–842, 2006.