Adaptive H-Extrema For Automatic Immunogold Particle Detection

Guillaume Thibault, Kristiina Iljin, Christopher Arthur, Izhak Shafran, and Joe Gray

Oregon Health & Science University (OHSU),
3181 SW Sam Jackson Park Rd Portland, OR 97239, USA

Abstract. Quantifying concentrations of target molecules near cellular structures, within cells or tissues, requires identifying the gold particles in immunogold labelled images. In this paper, we address the problem of automatically detect them accurately and reliably across multiple scales and in noisy conditions. For this purpose, we introduce a new contrast filter, based on an adaptive version of the H-extrema algorithm. The filtered images are simplified with a geodesic reconstruction to precisely segment the candidates. Once the images are segmented, we extract classical features and then classify using the majority vote of multiple classifiers. We characterize our algorithm on a pilot data and present results that demonstrate its effectiveness.

Keywords: Adaptive H-extrema, Mathematical morphology, Immunogold particle detection, Pattern recognition.

1 Introduction

Immunogold staining (IGS) is a technique used in electron microscopy (EM) to localize a molecule of interest – target molecule. This often achieved by attaching a primary antibody to the molecule of interest, which is then linked to the immunogold particle through a secondary antibody. After the gold particles are attached to the target molecules in this manner, the specimen is imaged using an electron microscope where the gold particles appear as "dark spots" due to the high electron density (see image 1). The IGS allows indirect visualization of target proteins and their approximate locations (the distance between primary antibody and immunogold is in range 15 to $30nm$). The immunogold particles are extremely small and so the IGS is typically employed in studies where cells or tissues are imaged at high resolution. The high resolution images in such studies are manually tagged, which is a time consuming process.

In this paper, we describe a new scheme to automatically detect the immunogold particles in high resolution images. We first explain the challenges in this problem (section 2), and then describe a new adaptive version of $H-extrema$, a mathematical morphology algorithm (section 3.1) for accurately detecting the particles in all conditions. In Section 4, we evaluate our method empirically to understand its capabilities and limitations and report results.

2   Problem Description

Immunogold particles appear as dark spots under good imaging conditions. However, the acquired EM images vary significantly depending on conditions of image acquisition, which are difficult to control precisely. These variations makes it difficult to detect or locate the immunogold particles in the image. The most common variation is the magnification of the EM images; figure 1 (left and middle) shows the same group of golds acquired with two different magnifications. The change in magnification, not only impacts the scale of the objects in the view, but also the shape and the intensity profile of the gold. Moreover, as shown in Figure 1, the intensity of the image is effected by the presence of relatively larger (dark) structures in the close neighborhood. Another factor that influences the quality of the image relates to noise or fuzziness, as shown in Figure 1 (right), arising from variations in specimen preparation, image acquisition, clustering of particles in the same location and the nature of organic tissue. The above mentioned variations can substantially affect the appearance of the gold particles, making the task of automatic detection challenging.
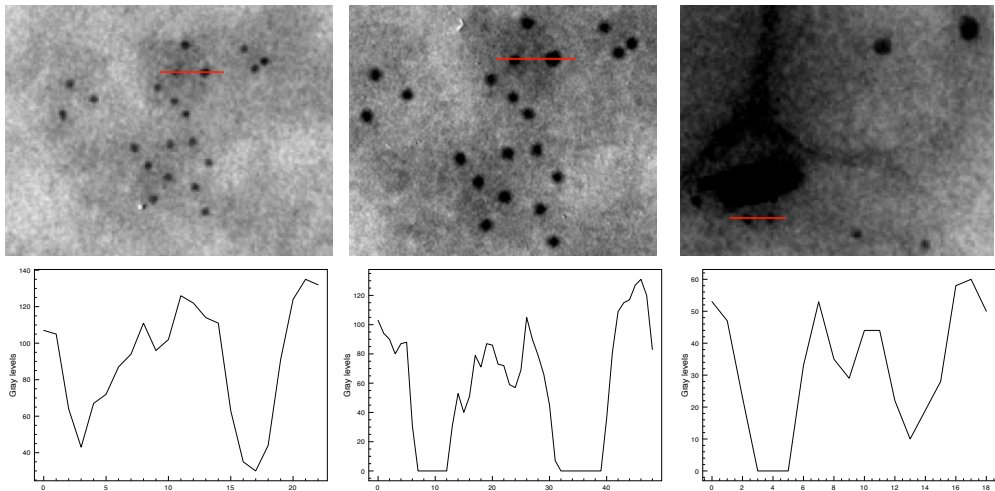
Fig. 1. Example of set of golds acquired at different magnifications (top left and middle), with different contrasts in the same original image (top middle and right), and their corresponding intensity profiles along the red segments.

3   Our Approach

Let $f : \begin{cases} E \to \mathcal{T} \\ \mathbf{x} \mapsto f(\mathbf{x}) \end{cases}$ be a gray-levels image, where $E \subset \mathbf{Z}^2$ is the support space of pixels and the image intensities are discrete values which range in a closed

set $\mathcal{T} = \{t_1, t_2, ..., t_N\}$, $\Delta t = t_{i+1} - t_i$, e.g., for a 8 bits image we have $t_1 = 1$, $N = 256$ and $\Delta t = 1$. Let us assume also that image f is segmented into its J flat zones $R_j[f]$ (i.e., connected regions of constant value): $E = \cup_{j=1}^{J} R_j[f]$, $\cap_{j=1}^{J} R_j[f] = \emptyset$. The size (surface area) of each region is $s(j) = |R_j[f]|$ ($|.|$ the cardinal). Hence, we consider that each zone $R_j[f]$ has associated a constant gray-level intensity $g(j)$.

3.1   Review: H-extrema

H-extrema, a mathematical morphology algorithm [1, 2], is a powerful non-linear filter to detect structures with certain intensity profile. The algorithm is comprised of two distinct algorithms – the h-minima and its dual operation the h-maxima. The h-minima (resp. maxima) detects dark (resp. bright) patterns with a intensity range of at least h. A constant h is added (resp. subtracted) to the original image f. The new image with $f + h$ (resp. $f - h$) is used in an over (resp. under) geodesic reconstruction, $OverRec(f, f + h)$ (resp. $UnderRec(f, f - h)$). In effect, the algorithm erases all dark (resp. bright) patterns with an intensity range lower than h, retaining all other structures (flat zone with a lower/higher gray level value than its neighborhood), as illustrated in Figure 2(b). So each local extremum in the resulting image corresponds to a local extremum in the original image with at least a dynamic range of h.

The main inconvenience of h-extrema is the fixed value of h that is added to or subtracted from the entire image. The fixed value doesn't take into account any local information, and hence it is not optimal for our task of gold detection. Figures 2 (b and c) illustrate this weakness. Often, gold particles close to dark areas are merged with the neighborhood, and thus erased.

3.2   Adaptive H-Extrema

We introduced a simple new adaptive version of h-extrema, which we refer to as *A-Extrema*, where for each pixel we adapt the value of the additive parameter h according to its neighborhood.

First a filter $\mathcal{F}$ is applied on the original image f, in order to get a new simplified image G_f, smoothed and containing only (preferably) global variations. Next, in the case of a-minima, for each pixel $\mathbf{x}$ of f, the value added is computed according to a function $\mathcal{A}$ and the corresponding value of $\mathbf{x}$ in G_f, $\mathcal{A}(G_f(\mathbf{x}))$. Finally the same over reconstruction is performed, as in h-minima. The algorithm 1 enumerates all the steps:

> **Data**: Image f, filter $\mathcal{F}$, function $\mathcal{A}$
> **Result**: Result image $Amin_f$
> **begin**
> $G_f \leftarrow \mathcal{F}(f)$;
> $Add_f \leftarrow f + \mathcal{A}(G_f)$; $[\Leftrightarrow \forall \mathbf{x}, Add_f(\mathbf{x}) \leftarrow f(\mathbf{x}) + \mathcal{A}(G_f(\mathbf{x}))]$;
> $Amin_f \leftarrow OverRec(f, Add_f)$;
> **end**

Algorithm 1: A-minima algorithm.

In our application, we chose the filter $\mathcal{F}$ to be an alternate sequential filter (ASF, alternation of openings and closings with structuring elements of increasing sizes [1]) because it simplifies the image without being affected by small local patterns. Moreover, ASF is known for its insensitivity to noise, which is useful in our application. For the ASF, we use disk type as our structuring elements, whose maximum radius equal to gold particles sizes so that we can erase them while computing the global variations.

For the function $\mathcal{A}$, we compute a percentage. Thus, in our adaptive algorithm h is computed as a proportion of the global variation computed from the neighborhood of each pixel according to $\mathcal{F}$. Thus by design, in dark areas a low h value is employed, whereas a high h value is employed in bright areas.

(a)   (b)   (c)

(d)   (e)   (f)
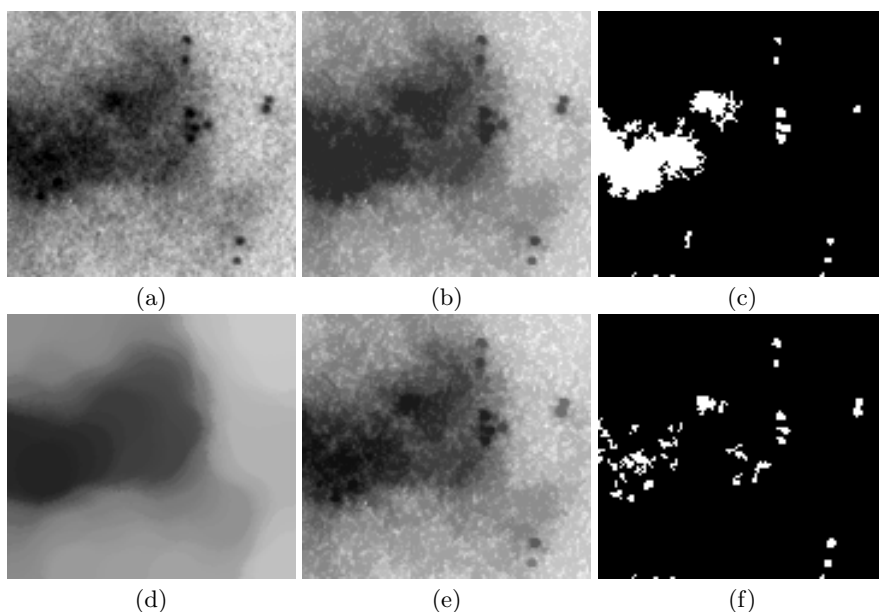
Fig. 2. The original image (a) and the various processes: the 43-minima result (b) and its local minima in white (c), the ASF result (d), the a-minima result for 43% of the ASF (e) and its local minima in white (f).

The image 2 e shows the A-minima result. The percentage for the function $\mathcal{A}$ needs to be computed empirically. In our example, we collected the statistics of contrasts in gold particles and found that the dynamic range was at least 43%. This is compared with h-minima results using the optimal value, $h = 43$. The figure illustrates the advantage of our $A - minima$ algorithm, which preserves gold particles with higher fidelity in both bright areas with high contrast and dark areas with low contrast. Thus, the combination of the filter $\mathcal{F}$ and the function $\mathcal{A}$ is effective in preserving the gold particles under different image contrasts.

Remarks:

- We preferred the ASF to classical mean or Gaussian filters, because it is not affected by noise and small artifacts.
- This new adaptive method can be adapted for the dual operation, A-maxima, to detect peaks.
- The function $\mathcal{A}$ can be generalized to larger class of functions. For example, in case of the detection of dark patterns in sub-exposed images, $\mathcal{A}$ can be modified as $Add_f \leftarrow f + Invert(G_f) \times p$ to be more effective.

3.3 Detection of Gold Particles

The A-minima preserves the sufficiently contrasted dark patterns and removes all other patterns with a lower contrast, thus provides a simplified image $Amin_f$. On the image $Amin_f$, we apply a new ASF in order to estimate the new global variations and then we compute the difference between the ASF result and the simplified image in order to extract all candidates: $C_f \leftarrow ASF(Amin_f) - Amin_f$. Each candidate is then isolated and characterized with 17 features:

- Geometrical Features: the surface and 3 radii (maximum, minimum and average). All these physical measures are "real" values estimated according to measurements on the image and the magnification.
- Texture:
    - Intensity Features: the average, the median and the range of intensity. Note, since the candidates are identified after applying $A-minima$, their dynamic range is guaranteed to be more than the minimum specified in the algorithm. The average and median provides more information about the intensity and shape.
    - Basic moments.
    - Contrasts.
- Shape (indexes) [3]:
    - Circularity: according to radii and the inscribed disk. Gold particles are expected to have a circular shape.
    - Besicovitch symmetry: Even though the image of gold particles may be merged, they still respect symmetry along a certain axis. Candidates with no symmetry are unlikely to be gold particles.
    - Gaussian: sigma of the best fitted circular $2D$ Gaussian and the residual error of the fit. As illustrated in Figure 1, gold particles can be easily approximated using Gaussians.

For each candidate, we compute the above features and experimented with three different types of classifiers for identifying the gold particles.

- Logistic Regression [5] (RL) is a linear regression function particularly well-suited to binary classification problems, allowing a variety of complex features.
- Random Forests [6] (RF) is a powerful, state-of-the-art classifier, consisting of an ensemble of trees.

– Neural network [7] (NN) is non-linear classifier, whose parameters are learned using back-propagation to minimize the cost function such as average squared loss.

These three classifiers have very different strengths, thus we expect different types of errors. In our approach, we combine the results from all the three classifiers using majority vote.

4   Experiments and Results

4.1   EM Images

The EM images were acquired from a realistic biological experiment. SKBR3 breast cancer cells were prepared for immuno-electronmicroscopy using Tokuyasu's method as previously described in [4]. Briefly, cells were chemically fixed in 4% paraformaldehyde in PHEM buffer, washed and embedded in 12% gelatin. After solidification, cell pellets were cut in small blocks and infiltrated in 2.3M sucrose. Blocks were mounted on specimen pins, frozen in liquid nitrogen and ultra-thin $80nM$ sections were cut with Leica cryoultramicrotome. Primary antibody recognizing protein disulfide isomerase localized in endoplasmic reticulum was selected since it has been previously shown to work in immunogold labeling for TEM (e.g. shown in [8, 9]) and includes incubation with bridging antibody (rabbit-anti-mouse IgG) and 5 nm protein A gold particles (from Dr. George Posthuma, UMC-Utrecht, the Netherlands), followed by contrasting in uranyloxalate and uranylasetate-methylcellulose. Imaging was performed using iCorr microscope (FEI).

4.2   Results

We evaluate our method using a data set of images where all the immunogold particles are manually annotated by experts. The data set consisted of 14 images, containing approximately 8500 gold particles. The evaluation was performed using a leave one out cross validation: an image is discarded from the data set, all golds from the remaining images are used to train the classifier, the discarded image is then processed and the result evaluated. The same process is performed for all the images, thus for a data set of N images, each image is used 1 time for validation and $N - 1$ times for training. The figure 3 *left* show the results for all the magnifications available. We can observe that our method provides particularly good sensitivity (nearly 100%) for magnifications from 1 to $\sim 3nm/pixel$, with a comparable specificity, which indicates that false alarms are minimal, and only gold particles are detected. Moreover, the area under ROC is equal to 0.9797, which demonstrates that our algorithm is effective in this task. For magnifications greater than $4nm/pixel$, the performances decrease rapidly. This is because at this magnification, a gold is represented by approximately 4 pixels, too few pixels to extract the relevant information accurately and robustly. At this resolution, they could easily be confused with noise in the image.
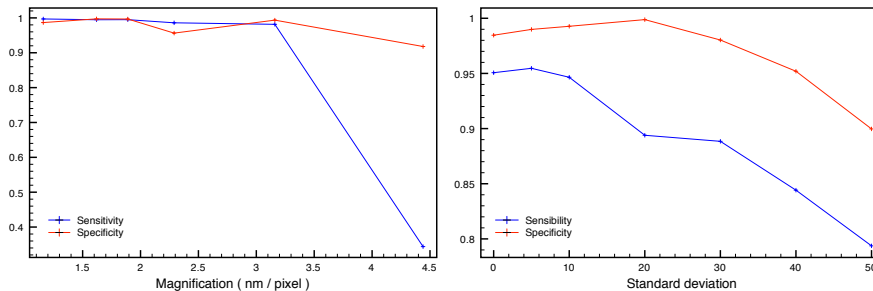
Fig. 3. The sensitivity/specificity obtained by the classifiers: without noise and according to the magnification (left); with artificially added noise and for a fixed magnification of $3.16nm/pixel$ (right).

But in EM imaging, it is particularly frequent to acquire images altered with Gaussian noise. In order to evaluate the algorithm's noise sensitivity, we artificially added Gaussian noise at different levels with respect to its standard deviation. These experiments were performed on images with a magnification of $3.16nm/pixel$, the current limit of accurate detections of gold particles. The performance of our algorithm is not significantly altered until the noise reaches a standard deviation of 10. Figure 3 (*right*) illustrates that from 10 to 20 the sensitivity decrease but not the specificity in spite of the fact that images are already extremely degraded[1] (see 4). At higher standard deviations, the specificity starts to decrease and the number of false positive increases.

5   Conclusions

In this paper, we developed a complete pipeline for automatically detecting immunogold particles. First we introduced a new adaptive version of h-extrema which filters contrasted patterns according to their dynamic and neighborhood: it preserves contrasted patterns even in really low contrasted neighborhood. Then this new method was successfully applied to simplify the images and find all potential candidates. Each candidate was then classified using machine learning algorithms. The results on a data set of annotated images show that our method detects immunogold particles in EM images with high accuracy, both high sensitivity and specificity, even in highly noisy images.

---

[1] Experts do not analyze images with such quantity of noise

(a) No noise   (b) $SD = 10$ (c) $SD = 20$ (d) $SD = 40$
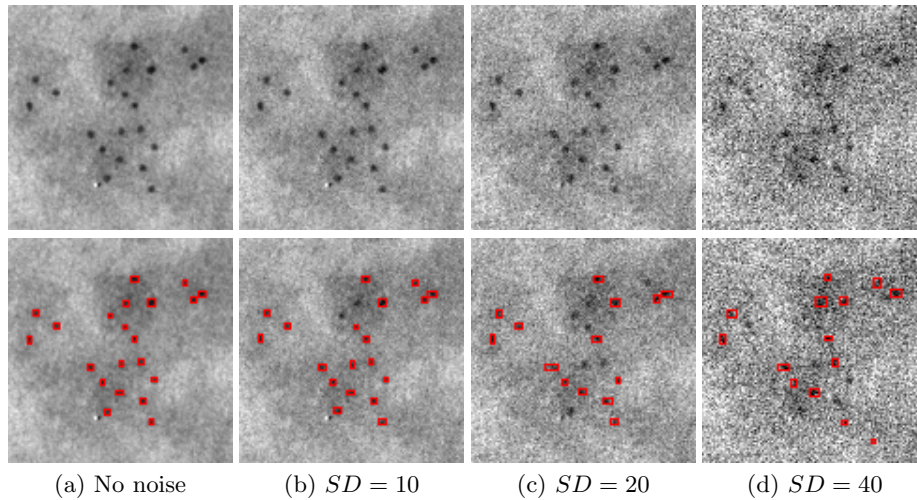
Fig. 4. Examples of detections with additive Gaussian noise at different standard deviations.

References

1. Jean Serra: Image Analysis and Mathematical Morphology. Academic Press, London (1982).
2. Michel Grimaud: New measure of contrast: the dynamics. In Proc. SPIE Vol. 1769, p. 292-305, Image Algebra and Morphological Image Processing III, Paul D. Gader; Edward R. Dougherty; Jean C. Serra; Eds., (1992).
3. Guillaume Thibault, Bernard Fertil, Claire Navarro, Sandrine Pereira, Pierre Cau, Nicolas Levy, Jean Sequeira, Jean-Luc Mari: Shape and Texture Indexes: Application to Cell Nuclei Classification. International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 27(1), (2013).
4. Jan W. Slot, Hans J. Geuze: Cryosectioning and immunolabeling. Nature, 10, pp. 2480–2491, (2007).
5. J. Berkson: Application of the Logistic Function to Bio-Assay. Journal of the American Statistical Association, 39, (1944).
6. Leo Breiman: Random forests. Machine Learning, 45(1), pp. 5–32, (2001).
7. Warren Sturgis McCulloch, Walter Pitts: A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, pp. 115–133, (1944).
8. Matthia A. Karreman, Elly G. van Donselaar, Hans C. Gerritsen, C. Theo Verrips, Arie J. Verkleij: VIS2FIX: a high-speed fixation method for immuno-electron microscopy. Traffic, 12(7), pp. 806–814, (2011).
9. Matthia A. Karreman and A.V. Agronskaia and Elly G. van Donselaar and K. Vocking and F. Fereidouni and B.M. Humbel and C. Theo Verrips and Arie J. Verkleij and H.C. Gerritsen: Optimizing immuno-labeling for correlative fluorescence and electron microscopy on a single specimen. Journal of structural biology, (2012).