Mirror Symmetry Histograms for Capturing Geometric Properties in Images

Marcelo Cicconet  
NYU  
mc4527@nyu.edu

Davi Geiger  
NYU  
geiger@cims.nyu.edu

Kristin C Gunsalus  
NYU  
kcgl@nyu.edu

Michael Werman  
The Hebrew University  
werman@cs.huji.ac.il

Abstract

We propose a data structure that captures global geometric properties in images: Histogram of Mirror Symmetry Coefficients. We compute such a coefficient for every pair of pixels, and group them in a 6-dimensional histogram. By marginalizing the HMSC in various ways, we develop algorithms for a range of applications: detection of nearly-circular cells; location of the main axis of reflection symmetry; detection of cell-division in movies of developing embryos; detection of worm-tips and indirect cell-counting via supervised classification. Our approach generalizes a series of histogram-related methods, and the proposed algorithms perform with state-of-the-art accuracy.

1. Introduction

This paper is a study on capturing the quantitative mirror symmetry between pairs of pixels, and on the use of this pairwise measure as a tool for finding various geometric entities, as well as for constructing shape representations.

A classical method for geometry representations is the Hough Transform [7], where the evidence for a certain shape comes from pixels with high edge intensity. Recently, there has been great advance in building computational representations of more complex objects via so-called Deep Learning methods [3] [10] [13], where the basic idea is to perform convolutions with filters and pooling coefficients from the convolution outputs.

The present work also explores the convolution and pooling techniques—we convolve the input image with a bank of Morlet wavelets, and pool the magnitude and orientation of the wavelet of maximum output at every pixel—but instead of considering global relationships between pixels via recursive convolution and pooling, we approach them directly by computing a coefficient of mirror symmetry for every pair of pixels. Considering pairs, as opposed to individual pixels, is also what sets us apart from the Hough Transform methods.

2. Pairwise Mirror Symmetry

Let I be a grayscale image, p ∈ I a pixel, S¹ the unitary circle in R², and τp ∈ S¹ a tangent vector that captures the direction perpendicular to the one in which the contrast is maximum around pixel p. Now, for v ∈ S¹ and θv ∈ [0, 2π) its associated angle, the reflection matrix with respect v is

\[ S(v) = \begin{pmatrix} \cos 2\theta_v & \sin 2\theta_v \\ \sin 2\theta_v & -\cos 2\theta_v \end{pmatrix}. \]

Let p and q be distinct points in the plane, with associated unit tangents τp and τq, and Tpq = \( \frac{q-p}{\|q-p\|} \). The pair ((p, τp), (q, τq)) is mirror symmetric if τq = ±S(Tpq)τp. For simplicity, sometimes we will just say that p and q are mirror symmetric, without mentioning the associated tangent vectors.

We construct a mirror symmetry measure (a mirror symmetry coefficient) between two distinct points p and q as

\[ f(p, τ_p, q, τ_q) = |τ_q S(T_{pq}) τ_p|. \]

Notice how f is invariant to a change of sign in τp or τq. We can simplify the notation and write f(p, q) instead of
\[ f(p, \tau_p, q, \tau_q). \] Observe that \( f(p, q) \in [0, 1] \), and that the larger \( f(p, q) \), the more mirror symmetric \( p \) and \( q \) are.

For each pixel \( p \) in \( I \), convolutions with wavelets of \( n \) different orientations (equally distributed in the interval \([0, \pi)\)) are performed. After this we have a set of \( n \) convolved images, \( \{ J_k = I \ast \psi_{\lambda_k} : k = 1 \ldots n \} \), where \( \lambda_k = (\sigma, \tau_k, m, \gamma) \) are the wavelet parameters (scale, orientation, number of visible peaks, and elongation).

We modify the mirror symmetry coefficient to include the strengths of the wavelets of maximum response, \( J_k(p) = \max_k J_k(p) \) and \( J_k(q) = \max_k J_k(q) \):

\[ f_k(p, q) = f(p, q) J_k(q) J_k(p). \] (3)

3. HMSC and its Marginals

The pairwise mirror symmetry coefficients are grouped in a 6-dimensional histogram \( H \) defined by

\[ H(c_x, c_y, \theta, \phi_p, \phi_q) = \sum_{p, q \neq p} f_k(p, q) \chi_{c_x, c_y, \theta, \phi_p, \phi_q} \] (4)

where

\[ \chi_{c_x, c_y, \theta, \phi_p, \phi_q} = \delta_{c - \frac{\tau_p}{2}} \cdot \delta_{d - \frac{\tau_q}{2}} \cdot \delta_{\theta - \phi_p} \cdot \delta_{\phi_q}. \] (5)

c = (c_x, c_y), \( \theta_{pq} \) is the angle associated with \( T_{pq} \) (the unit vector perpendicular to \( q - p \)), \( \phi_p \) and \( \phi_q \) are the angles associated with \( \tau_p \) and \( \tau_q \), and \( \delta_2 \) is the Kronecker delta.

The main idea of this paper is that we can detect geometric properties and shapes in images via histogram marginalizations in \( H \): we perform sums in \( H \) over the variables we do not focus on, or on a subset of the domain of \( H \) where the variables satisfy certain restrictions; this reduces the dimensionality of the histogram, leaving only the variables needed for the specific problem.

In general, marginalization assumes the form

\[ H(\beta) = \sum_{\alpha_1, \ldots, \alpha_n} H(\alpha) \prod_{i=1}^{k} \delta_{\beta_i - f_i(\alpha)}, \] (6)

where \( \beta = (\beta_1, \ldots, \beta_k) \), with \( k \leq 6 \), is the new set of parameters, and the set of functions \( f_i(\alpha_1, \ldots, \alpha_n) \) depend on the variable restrictions related to the particular application.

The HMSC is a theoretical framework for describing shapes and developing algorithms. In the applications, we compute the marginal histograms directly. Computational complexity is linear in the size of each marginal histogram, and running times are of the order of a few seconds for the applications described in section 4.

**Circularity** Setting \( k = 3, \beta_1 = C_x, \beta_2 = C_y, \beta_3 = r \), and \( f = (f_1, f_2, f_3) \) as the function representing the parameters (center coordinates and radius) of the best circle fitting \( p, \tau_p, q, \tau_q \), equation 6 takes the form

\[ H_C(C, r) = \sum_{\alpha} H(\alpha) \delta_{C - (f_1(\alpha), f_2(\alpha))} \delta_{r - f_3(\alpha)}. \] (7)

where \( \alpha = (c_x, c_y, \theta, \phi_p, \phi_q) \). We define the best circle fitting \( p, \tau_p, q, \tau_q \) as the one centered at the line with inclination \( T_{pq} \) passing through \( \frac{\tau_p + \tau_q}{2} \), whose boundary passes through \( p \) with tangent parallel to \( \tau_p \). In our applications we marginalize \( H_C(C, r) \) further. A centers histogram \( H_C(c) \) is created by summing over the radius variable:

\[ H_C(C) = \sum_{r} H_C(C, r). \] (8)

A radii histogram, \( H_r(r) \), is obtained by summing over the circle centers:

\[ H_r(r) = \sum_{C} H_C(C, r). \] (9)

Applications of centers and radii histograms are reported in the three first experiments of section 4.

**Reflection Symmetry** Suppose our goal is to extract the best symmetry axis (line) from an image. A line can be described by two parameters, \( (\theta, \rho_c) \), where \( \rho_c \) is the distance between the origin (top-left corner of the image) and the line, while the line inclination is given by \( \theta \). Then, given \( (c_x, c_y, \theta) \) in the symmetry histogram space, we have \( \rho_c = \sqrt{c_x^2 \sin^2 \theta + c_y^2 \cos^2 \theta} \), so we define the restriction \( f(c_x, c_y, \theta) = \rho_c \). We then marginalize \( H(\alpha) \) into \( H_S(\theta, \rho) \) as follows:

\[ H_S(\theta, \rho) = \sum_{c_x, c_y} H(\alpha) \delta_{\rho - \sqrt{c_x^2 \sin^2 \theta + c_y^2 \cos^2 \theta}}, \] (10)

where \( \alpha \) is as in Equation 7. Local maxima in \( H_S(\theta, \rho) \) correspond to axis of symmetry in images. Examples are shown in Detecting the Axis of Symmetry, section 4.

**Shape Description** The set of points and directions of wavelets of maximum response, \( \{(p, J_{kp}(p)) : p \in I\} \), can be thought of as a vector field, and statistics of the field can be used as a descriptor, in a supervised classification context, for shape detection. In particular, we choose a pivot point, say \( \bar{p} \), and look at the bi-dimensional distribution where one dimension (variable \( x \)) is given by the distances between \( \bar{p} \) and \( q \) (which in our notation corresponds to \( 2d \)), and the other (variable \( y \)) by the norm of the dot product between \( T_{pq} \) and \( \tau_{\bar{p}} \). Thus, setting \( f(\alpha) = (p, 2d, |(T_{pq}, \tau_{\bar{p}})|) \), the marginalization we are referring to is
Figure 1. Circle detection. Left: wavelets used to compute mirror symmetry coefficients. Right: centers histogram (in red); center (yellow) and boundary (green) of the fit circle.

\[ H_D(x, y) = \sum_{\alpha} H(\alpha) \delta(p, x, y) - f(\alpha) , \]  

(11)

where again \( \alpha = (c_x, c_y, d, \theta, \phi_p, \phi_q) \). Notice that the restriction \( f(\alpha) = \bar{p} \) implies that, differently from previous cases, the value in a domain point of the marginalized histogram doesn’t come from all possible pairs of distinct points in the image, but rather from pairs where one of the points is the pivot (\( \bar{p} \), in this case). \( H_D \) is used in Detecting Tips of C. Elegans via supervised classification, section 4.

4. Experiments

**Counting Cells** Local maxima in \( H_c(c) \) correspond to the centers of circles in the image. If we are looking for circles of a known radius, or range of radii, say \([r_1, r_2] \), we modify the normalization of equation 8 to reflect this:

\[ H_c(C) = \sum_{r \in [r_1, r_2]} H_0(C, r) . \]

(12)

Since \( H_c \) is usually sparse, a smoothing operation is performed before searching for local maxima. Figure 1 shows intermediate and final steps of this algorithm on an image containing a cell.

We conducted a series of experiments to evaluate the performance of this circle detection method. First, we tested the algorithm for finding the number of embryos in a well (where embryos are in the 1-cell stage). Figure 2 shows a sample image. We manually located the centers of cells in 112 images, totaling 1442 embryos, where cell radii are of about 30 pixels. Detection is considered correct if the located center is at most \( \Gamma = 15 \) pixels away from the ground-truth center. If \( n > 2 \) centers are located at a distance < \( \Gamma \) from a ground truth center, \( n - 1 \) are considered false positives. False negatives correspond to lack of detection at a distance < \( \Gamma \) from a true center, and to centers located more than \( \Gamma \) pixels away from any true center.

Figure 3 shows the obtained results in comparison with the CHT method. The parameter varying in the CHT method is the minimum gradient magnitude from which edges are considered for voting (from 0.1 to 0.2 of the maximum value). In the HMSC method, we vary the minimum wavelet magnitude from which a point is considered (from 0.16 to 0.24 of the maximum value).

The first experiment doesn’t contain images where cells overlap. On a second experiment we evaluated the circle detection algorithm on two datasets of overlapping cells. The datasets were obtained from microscopic images of mouse embryos at the 4-cell stage (as in Figure 2, right), where the locations of the centers were annotated with the help of tracking software \[4\]. Table 1 summarizes the obtained performance.

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHT</td>
<td>99.86%</td>
<td>99.51%</td>
</tr>
<tr>
<td>HMSC</td>
<td>99.75%</td>
<td>99.45%</td>
</tr>
</tbody>
</table>

Our results are similar to those of state-of-the-art algorithms for cell detection. \[23\], for instance, reports 95% precision / 80% recall for cytoplasm detection on simulated cell culture images, and 82% / 90% on a Drosophila Kc167 dataset; \[1\] claims an accuracy of > 94% on nuclei detection in histological images; and \[22\] publishes pre-
cision/recall rates of 90% and 78% for cell segmentation in Hematoxylin-stained breast TMA specimen images. We were able to run the method in [23] on our dataset of 112 hand-labeled images totaling 1442 embryos, obtaining precision/recall (in %) of 99.86/90.17, which is slightly worse than ours (99.86/99.51).

Detecting Cell Division By monitoring the centers histogram along a sequence of frames, cell division detection can be performed. Here we describe a method for detecting the first division in a time-lapse movie of a mouse embryo. The movie starts with one cell, of radius $r_1$, which eventually divides in two cells of radii $r_2 < r_1$. The movie ends at the 2-cell stage.

For each frame $t$, we consider the marginal histogram

$$H_t^r(C) = \sum_{r \in [r_1 - \epsilon, r_1 + \epsilon]} H_t^r(C, r),$$

where the superscript $t$ denotes histogram computation at the frame of index $t$ in the video sequence, and $\epsilon$ is about 5 pixels (the range $[r_1 - \epsilon, r_1 + \epsilon]$ allows for deviations of the cell shape from perfect circularity).

Thus $H_t^r(C)$ accumulates the evidence that there is a circle of radius in the range $[r_1 - \epsilon, r_1 + \epsilon]$, centered at point $C$, in frame $t$. Therefore, the value $L_t = \sum_C H_t^r(C)$ represents the likelihood that there is a circle of radius in such range in the entire image.

If $\epsilon$ is small enough so that $r_2 < r_1 - \epsilon$, the likelihood $L_t$ will drop significantly when the first cell divides. The frame of division is the global minimum of a cost function $g(t)$, as in the experiment Counting Cells (see main text), with $\Gamma$ = 30 pixels.

Table 1. Performance of our circle detection method on largely overlapping cells of radius $r \approx 60$ pixels. Each image contains 4 cells, where the minimum distance between a pair of cells is larger than $\Gamma_1 \cdot r$ and smaller than $\Gamma_2 \cdot r$. Notice that $\Gamma_2 = 2$ corresponds to cells that intersect at least one point (their boundaries touch), and $\Gamma_1 = 1$ implies that cells can intersect as much as to share a radius. The criterion for correct detection is as in the experiment Counting Cells (see main text), with $\Gamma = 30$ pixels. Legend: NI: number of images; NC: number of cells.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Acc. (1/2/4 cells)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HoG</td>
<td>0.4229</td>
</tr>
<tr>
<td>HTRH</td>
<td>0.4788</td>
</tr>
<tr>
<td>RH</td>
<td>0.9463</td>
</tr>
</tbody>
</table>

Table 2. Accuracy of the kNN ($k = 5$) method with different descriptors in the separation of 2 (one cell or two cells) or 3 (one, two, or four cells) classes. Legend: HoG = Histogram of Oriented Gradients; HTRH = Hough Transform Radii Histogram; RH = Radii Histogram (our descriptor).

Counting Cells via Supervised Classification In mouse-embryo images, the radii of cells are roughly constant within generations, and due to the the fact that the total volume of the cells in the early stage of development is almost invariant, the radii decrease with the passing of generations. Therefore, we expect the radii histogram (equation [10]) to reflect at which development stage the embryo is.

Let $r_1$, $r_2$, and $r_3$ be the estimated radii of first, second, and third generation cells. Let $\epsilon_i$, be range parameters (for $i = 1, 2, 3$), big enough to allow for small variations in expected radii, and small enough so that the intervals $[r_1 - \epsilon_i, r_1 + \epsilon_i]$ don’t intersect. Restricting the radii histogram $H_t^r(C)$ to the mentioned intervals, and discretizing it so as to have an integer number of bins, we build a descriptor that can be used to count cells in a supervised classification context.

We did an experiment where we consider the ranges $\{55, 56, ..., 60\}$, $\{45, 46, ..., 50\}$, and $\{30, 31, ..., 40\}$ (pixels), so that our descriptor has dimension 21. We use the kNN algorithm to separate images containing one, two, or four embryonic cells (examples are shown in Figure 2 from the second to the last). Training and test sets have 200 and 800 samples per class, with images taken from different embryos and different experiments.

We compared our descriptor with the Histogram of Oriented Gradients [5]. and an equivalent (i.e., same intervals, same discretization) radii histogram coming from the standard CHT algorithm [7], which considers votes point-by-point, as opposed to our pairwise approach. Table 2 shows the best obtained results for each of the methods.

Detecting the Axis of Symmetry Local maxima in $H_S$ (equation [10]) correspond to the axis of symmetry in an image. Figure 4 shows an example of the method (histogram and the located symmetry axis). The picture was taken
Figure 4. $H_S$ and symmetry detection. Left: $H_S$, where the red dot is the point of maximum. Right: white dots represent pairs with high mirror symmetry coefficient, white big dots are the vertices of the convex hull of those pairs, and the line is the detected symmetry axis; the white part of the line is the intersection of the symmetry axis with the convex hull of highly symmetric pairs.

Figure 5. (a) Positive (tips) labels: for each tip location (labeled manually), a neighborhood of 5 pixels is used. (b) Negative labels, distributed roughly evenly between in-worm and background pixels. (c,d) Examples of test images, where highlighted pixels correspond to pixels our method classified as tips.

from the database of the Symmetry Detection Competition of CVPR 2011.

Out of the 14 images with single symmetry axis in the database (6 real world images and 8 synthetic), global maxima in $H_S$ provided correct answers in all but one image (a real world example). This corresponds to an accuracy of 93%, which is the same obtained by the best performing method for such a subset as reported in [24].

Detecting Tips of C. Elegans In this application, rather than computing a histogram globally (in the entire image), we use it as in a “filtering approach,” that is, we scan the image pixel by pixel and compute the histogram in a local window. For the particular experiment detailed here, we use windows of size $11 \times 11$ pixels.

We plugged our shape histogram (equation (11)) into a SVM classifier (using a Radial Basis Function kernel), and tested its accuracy for detecting tips of Caenorhabditis Elegans (a model-organism worm often used in Biology experiments). Figure 5 shows the images used for training, and two examples of images used for testing.

Keeping the SVM classifier, we compared our descriptor with HoG [5] (using the same window size) and SIFT [15] (computed at the center pixel location). We trained with 44 tip pixels, 73 non-tip worm pixels, and 77 non-tip background pixels. For test we used 42 images containing a total of 111 tips (whose locations were hand labeled). Blobs of pixels classified as tips were collapsed into one tip location only. Accuracy was measured as in the cell-counting experiment, with $\Gamma = 5$. Table 3 shows the best obtained results for each of the methods.

Table 3. Accuracy of the SVM method (using a RBF kernel), with different descriptors in the detection of tips of C. Elegans.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Precision</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histogram of Oriented Gradients</td>
<td>0.4685</td>
<td>0.4685</td>
</tr>
<tr>
<td>Scale Invariant Feature Transform</td>
<td>0.7500</td>
<td>0.4865</td>
</tr>
<tr>
<td>Shape Histogram (our descriptor)</td>
<td>0.8272</td>
<td>0.6036</td>
</tr>
</tbody>
</table>

5. Conclusion

We introduced a framework for shape analysis that consists of different marginalizations of a 6-dimensional Histogram of Mirror Symmetry Coefficients (HMSC) for the geometric study of objects in images. Properties such as circularity, reflection symmetry, and shape description are considered.

A number of marginalizations are described in detail, and computational experiments for five applications are reported. Our approach performs with state-of-the-art accuracy on circle detection (in the context of counting cells), as well as on symmetry axis detection; and our histogram descriptors outperform classic descriptors (such as HoG and SIFT) in association with classification techniques for counting cells and detecting tips of worms.

We are very excited with the generality and range of applications of the method, as well as with the avenues of research that it opens. One possibility, for instance, is to build a system that “learns” the proper marginalization for shape recognition based on the input images.

Sample code (in Matlab and C++) and related material is available at [marceloc.net/science/hmsc](http://marceloc.net/science/hmsc).

6. Related Research

Mirror symmetry has already been shown to lead to the extraction of skeletons [8], and is the basis for more general methods for detecting the global axes of symmetry [12, 11, 20]. A recent review of methods specially designed for reflection symmetry detection can be found in [14], and a more recent work is [21]. Detection of mirror symmetry also appeared as a particular case in [26]. Although one of our applications is the detection of symmetry axis, we did not focus on this problem, but rather on pairwise symmetry as a general framework for geometric image analysis.

Due to the fact that marginalizing histograms is a “voting scheme,” and considering that we also deal with circle detection, our work naturally resembles the Circular Hough Transform (CHT) [7]. We provided a more detailed comparison with the CHT in section 4. Similarly, and also related to the problem of detecting the axis of symmetry,
our method resembles [16], in the sense that pairwise information is accumulated in a “Hough space.” Differently, though, our symmetry measurement is based on the outputs of wavelets, instead of the more complex SIFT descriptor [15] on selected feature points.

Methods based on symmetry between pairs of pixels with tangents have been proposed in the past, e.g. in [28] [19] [8] [27]. [18] provides a good application for 3D object recognition, but the work does not develop the histogram framework, nor is it based on wavelet filters. [11] is of interest as well, as they similarly define symmetry via reflection matrices. However, they also do not have a histogram formulation.

The literature for shape detection is likewise vast, including RANSAC methods [25], Multiresolution Histograms [9], and Shape Contexts, [2], to cite a few. Our approach is similar to Multiresolution Histograms, but we accumulate pairwise information of magnitudes and angles, instead of pixel luminances for different image resolutions. Also, our method for worm-tips detection has a connection to Shape Contexts, and the method for computing similarity between vector fields described in [6], but we use a different technique to compute the image gradients (wavelets), and those methods do not build histograms using pairwise symmetry measurements, which is the focus of our work.

References