LA-UR-14-23579
Approved for public release; distribution is unlimited.

Title: MAMA Software Features: Quantified Attributes

Author(s): Ruggiero, Christy E.
Porter, Reid B.

Intended for: Web

Issued: 2014-05-20

MAMA Software Features:

Quantified Attributes

Abstract:

This document reviews the size and shape attributes that are quantified by the MAMA software, shows how they are calculated, and then gives some visual example shapes.

Contents

Classification Review: UNCLASSIFIED

☒ **Unclassified**

Derivative Classifier Name, Org., Signature, Date:

Reviewed using CG-TNF-1, 4/1/14

NAME: Diana Hollis, Classification Office, LANL

(Reviewed as part of larger document):

LA-UR: PENDING

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

1.0 Mama Attribute Summary

The table at the end describes how the attributes are mathematically calculated so that you can fully understand the values the software produces. We have determined uncertainty in the calculation for most of the basic attributes (it's a function of size and shape), but do not yet have final numbers for this yet (coming soon!). No single set of attributes can accurately define the shapes of the large variety of particles possible, so multiple attributes are calculated to help distinguish different shape—the ones that will be the most useful for your materials will depend on the materials. If your material contains very odd shapes, such as spirals or threads, or features (specific holes/layers/structures) different attributes may be needed to better quantify the materials attributes.

1.1 Why these attributes?

1.1.1 Area and Pixel Count:

Area and pixel count are obvious primary size attributes calculated that are commonly reported.
ECD is a size attribute derived from the area, and it is included since it is very commonly reported in the literature.

1.1.2 Convex Hull Area:

Convex Hull Area is not as commonly used by itself, but is useful for convexity calculations. It has also proven to a very robust attribute. By smoothing the object perimeter, the convex hull calculation eliminates many segmentation variations that can occur in different images/among users, so give a more consistent result that is, in validation testing, averages less than 0.5% different than the directly measured attribute. It may prove to be equally or more useful attribute in a large database of samples than area.

1.1.3 Perimeter and Convex hull perimeter:

Perimeter and Convex hull perimeter are probably y not useful as direct numbers for quantifying attributes, but they are the basis of calculated **Circularity/Roundness and Perimeter convexity**. The perimeters are displayed so that if there are any anomalies in these calculated secondary attributes, users can see the data that went into the attribute calculation. To verify and validate the derived attributes, these numbers must be used.

1.1.4 Aspect Ratios, Ellipse and Chordal :

Aspect Ratios, Ellipse and Chordal: After a size attribute, aspect ratio (or elongation) is probably the second most commonly reported on particles. After testing a number of aspect ratio calculations methods (require orthogonal or not, through centroid or not, ferret/caliper, bounding box, etc.), we have settled on these two aspect ratios because they will give different values from each other depending on particle features. Ellipse aspect ratio, since by a best-fit ellipse, is essentially the aspect ratio of the smoothed particle. It is a somewhat better, but very similar, to doing the aspect ratio of the convex hull. It is a very good attribute for how generally elongated a particle is. However, it has two major caveats: It cannot be directly vertified by a user-performed measurement on the object, except in limited cases, and it is very poor at distinguishing all convexity feature s, just as convex hull calculations

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

would be.   The Chordal aspect ratio on the other hand will accentuate convex and irregular edge features. It can be directly measured for validation, but is also more sensitive to minor differences in segmentation.

1.1.5 Ellipse perimeter, Ellipse Min & Max :

Ellipse perimeter, Ellipse Min & Max are probably not typically as direct numbers of quantifying attributes and features in objects, but are needed for any verification of the Ellipse aspect ratio.   The max axis of the ellipse (or the max chordal length, which is not displayed anymore) may be helpful in differentiating particles since either is a direct measurement ratio than a ratio. Any of these ellipse attributes can also be used in secondary calculations, but such calculations are can also be represented by more typical circularity/convexity numbers.

1.1.6 Area and Perimeter Convexity, Circularity and Roundness:

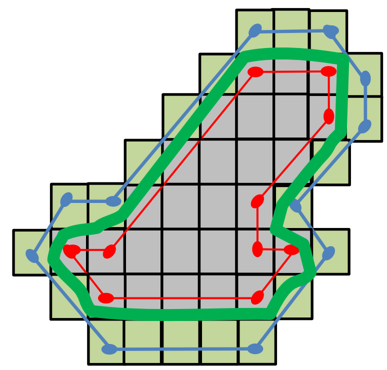
Area and Perimeter Convexity are fantastic attributes in their robustness in calculations, and complete a basis set (size, shape, convexity) to describe basic particle attribute. Some ratio of perimeter to area or ratio or an attribute to its convex hull equivalent attribute is needed in order to distinguish objects (It is easy to generate a set of object that are identical in area and ellipse aspect ratio, but that look totally different and vary in convexity/circularity measurements.   Circularity, and its reciprocal roundness, are more commonly used, but are less robust to object representation variations (how the object aligns in pixel shape) since they magnify variations in perimeter. They also both vary with convexity and with aspect ratio. However, they are a very commonly used (albeit in many variations and names). Since all are derived calculations from fundamental attributes (Areas and Perimeters), they are easy to calculate and validate. Most importantly, they are essential for highly convoluted shapes. As vectorized calculations, unlike theboth aspect ratios, the will give large difference in values for convoluted shapes (e.g. a spiral) for which aspect ratios would be nearly identical. All are included to increase comparability for samples not analyzed within this software.

1.1.7 The Hu's:

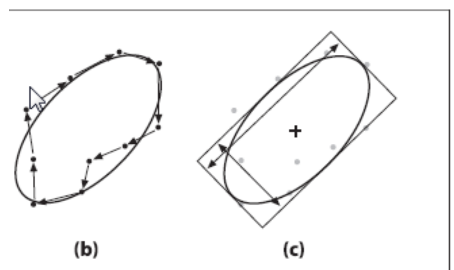
You are not expected to understand or have an intuitive sense of Hu moments. They are one of a handful of derived set of shape attributes sets that are commonly used in machine learning to classify and categorize objects by shapes. They often do far better at distinguishing shapes than traditional size/shape attributes that are more intuitive to humans in computer-based learning and classifying efforts. They are implemented for when we generate data sets for which we want to classify groups of particles by shape, and will likely help when standard measures fail. For now, they are a great tool for which to explore the particle shapes. They have been tested, but not fully validated yet.
Grayscale and Gradient: Although listed as texture, these four attributes are simple here for triaging and labeling your data. They are not robust to image variations in any way since they are not normalized. They should not be thought of as reportable and comparable values.

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

2.0 Calculation Details

Term	Definition
Area and Perimeter	
Pixel count	This is the area of the object, not including holes. This is calculated by counting the total number of pixels within the object boundary (segmentation boundary), and multiplying by the micrometer/pixel scale.
Area	This is calculated using the contour representation of each object. A contour vector is calculated using the openCV function cvfindContour. This vector is an ordered sequence of points on the object's external perimeter. This calculation does not remove holes before the area calculation is performed. Because this vector is calcualated from the center of the boundary pixels (red line), it represents a shape slightly smaller than the actual object. Therefore, the object is dialated (expanded) by one pixel (green pixels) and the cvFind Contour vector is calculated a second time (blue line). The area of each contours is calculated using cvContourArea using Green's theorem, and then the two areas are averaged to give the calculated area of the shape. NOTE: this area estimate is typically very close to the pixel **Area**.  Example of a irregular shape (bright green outline) in pixels. Pixel Area (Gray) = 27 Vector Area, internal(red lines) = 17 Vector Area, external (blue lines)= 36.5 Average: 26.75 pixels.
Convex Hull Area	This is the area of the convex hull (convex polygon outline) of the object. This is calculated from the vectors determined using the cvFindContours algorithm (as above), and then the external points of this set of points is calculated using cvConvexHull (which using Sklansky's algorithm) to give a reduced vector (reduced set of points). These vectors are then used in the cvContourArea algorithm to calculate the convex areas, which are averaged as above. Note: If the input contour has points of self-intersection, the region area within the contour may be calculated incorrectly.
Perimeter	Perimeter is calculated using the object contour vector determined by cvFindContours. The function cvArcLength calculates the length of the perimeter as the sum of segments between subsequent points. Like the area calculation, the cvFindContour vectors from the shape and the dilated shape are both used, and the resulting perimeters are averaged. Because we are representing objects that may be round with a finite set of pixels (points) for the boundary, and calculating the boundary as steps/diagonals around the perimeter, this perimeter will typically be slightly larger (a few percent) than the theoretical perimeter of the shape. (For a circle 100 pixels in diameter, the vector perimeter is ~ 6% larger than the mathematically calculated perimeter. For a hemi circular arc that is 100 pixels in diameter with an internal arc of 50 pixels in diameter, the perimeter is ~3% larger than the mathematically calculated perimeter.) For example, in the shape above, the green perimeter is roughly 21 pixel in length. The full pixel count perimeter (step wise along the edge of each gray pixel) gives 30 pixel perimeter. Estimating using stepwise and diagonals gives a better estimate of the perimeter at ~25 pixels The vectors give 21.3 and 26.3, giving and average of 24 pixels, which is in between the actual shape of the object (which we can never measure since we start with a pixelated image) and the best estimate from is representation is pixel space

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

Convex Hull Perimeter	This is the perimeter of the convex hull of the object. This is calculated as above, only using the convex haul vector points. Because the convex hull eliminated the perimeter variations, either real or resulting from segmentation variations, this estimate of perimeter is very close to the mathematical calculated perimeter for the convex hull of a regular shape. (For a circle 100 pixels in diameter, the convex hull perimeter is <1% different than the than the mathematically calculated perimeter. For a hemi circular arc that is 100 pixels in diameter with an internal arc of 50 pixels in diameter, the convex hull perimeter, which would be the perimeter of a hemicircle, is < 1% differentr than the mathematically calculated perimeter.)
Ellipse Perimeter	The perimeter of the best-fit ellipse is estimated using the values obtained from the calculations of the major and minor axes of the best-fit ellipse , with the following equation (where *C* is the estimated perimeter of the ellipse): a = length of major radius of ellipse (i.e. half the length of the major axis) b = length of minor radius of ellipse (i.e. half the length of the minor axis) $C \approx \pi (a+b) \left(1 + \frac{3\left(\frac{a-b}{a+b}\right)^2}{10 + \sqrt{4 - 3\left(\frac{a-b}{a+b}\right)^2}} \right);$ The equivalent ellipse is calculated using cvFitEllipse: All points (pixels) in the object perimeter and the points (pixels) in the dilated object (object + 1 pixel) are used for fitting this ellipse. The cvFitEllipse uses a fitting function and returns the ellipse that is the best approximation to these perimeter points. This means that not all points in the contour will be enclosed in the ellipse. The fitting is done using a least-squares fitness function. *This means that the ellipse that is found that minimizes the geometric distance of the boundary points on the particle from the fitted ellipse. The ellipse is not constrained to be the same area , centroid, or any actual diameter as in the particle. This would minimize odd narrow tails, fringes, or extensions on the particle—in essence; the ellipse is a smoothed version of the particle. By using a 2 pixel perimeter, for a regular elliptical shape, the fitted ellipse will be the exact bounding ellipse.* The results of the fit are returned in box structure that exactly encloses the ellipse. From this the major and minor axis of the ellipse are easily extracted. See Figure below (From OReilly Learning OpenCV : www.cse.iitk.ac.in/users/users/vision/dipakmj/papers/**OReilly**%20Learning%20**OpenCV**.pdf *(detailed math on the fitting can be found at* www.cs.unc.edu/Research/stc/FAQs/**OpenCV/OpenCV**Reference**Manual**.pdf  Equivalent ellipse area could also be implemented, but is not commonly used.
Axis Diameters	
Equivalent Circle Diameter	This is the diameter of the equivalent circle (diameter of a circle with an equivalent area). Area used is vector based area, and diameter = $(4A/\pi)^{½}$. The two vector based areas are used to each calculate the ECD, and the resulting ECDs are averaged.
Major & Minor Ellipse	These diameters are calculated from the equivalent ellipsoid (same ellipse as calculated in perimeter).
Aspect Ratios	
Ellipse Aspect Ratio	This is the ratio of the Major Axis Length to the Minor Axis Length. (1 = sphere, > 1 longer than sphere. Note that this is a fitted ellipse. It will give an aspect ratio that may be different than you

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

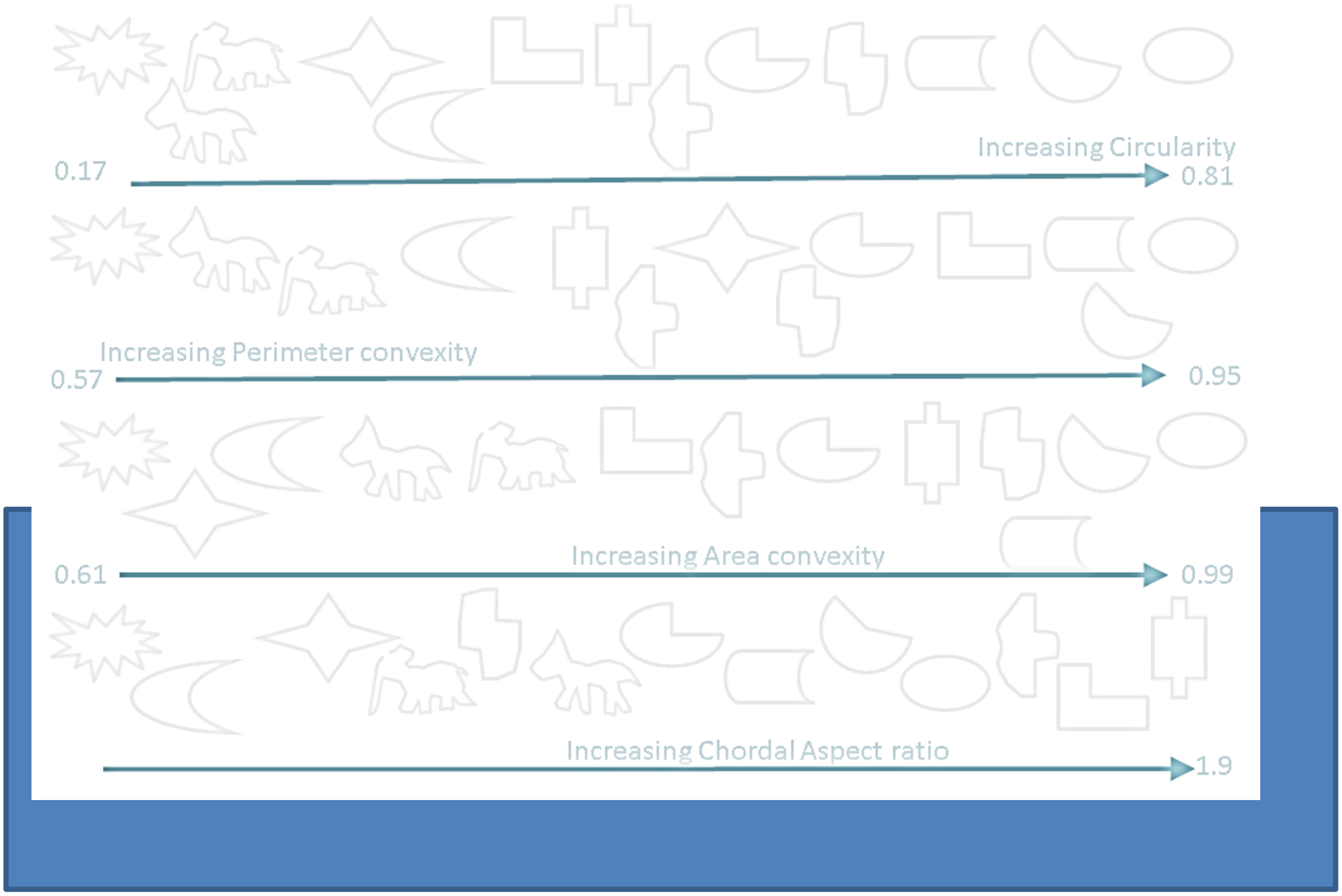
	expect if your object is convex, contains numerous holes, or is a spiral/curved shape. Although robust to image variations, it is insensitive to many object differences.
Maximum Chordal Diameter Aspect Ratio	The perimeter (boundary) pixels are used as the set of points. Next, the distance between all pairs of points in the boundary is calculated, along with the angle that the chord between the two points makes with the image horizontal axis. The pair of points that has the greatest distance between them is noted (i.e. the maximum-distance chord), together with this distance measure (chord length) and the angle that this maximum-distance chord makes with the axis. The list of point pairs is then searched to find the pair that has the largest distance between them, but with the constraint that they must be within a small angle (angle tolerance is 2 degrees) of being orthogonal to the maximum-distance chord (the longest orthogonal chord). The maximum diameter aspect ratio is then calculated as the ratio of the length of the maximum-distance chord to the distance of the longest orthogonal chord. Note that neither of these are constrained by going through the centroid; an equilateral triangle would have an aspect ratio of 1 by this calculation. (It is very similar to ferret aspect ratios, but without a constraint that it must go through the centroid of the object.) Also note that for small objects, the angle step between adjacent pixels can be large than 5 degrees, limiting the accuracy of this calculation. The raw numbers used for this calculation (maximum chordal diameter, and chordal diameter ~ perpendicular to that) could be added to the output data if users need it. NOTE of CAUTION: The current implementation looks for the maxium chordal length, and then calculates the orthroganal chord. It does not check for duplicate maximum chords. This can lead to differences in aspect ratio than a user would expect. This will be corrected in a pending update.
Shape	
Circularity	This is calculated from 4 π Area /(Perimeter^2) using the vector based area and perimeter. This is the inverse of the object roundness, and can be useful since it must be between 0 and 1, where circle =1)
Roundness	This is calculated from (Perimeter^2) **/** 4 πArea using vector based area and perimeter. This will be greater than 1 if not a circle, and 1 for a perfect circle. (The square root of this sometimes called Heywood circularity factor)
Perimeter Convexity	This is the ratio of the convex hull perimeter to the object vector perimeter. (Must be between 0-1)
Area Convexity or Compactness	This is the ratio of the vector area of the object to the convex hull area (Must be between 0-1)
Invariant Moments	
Hu 1 to 6	In image processing an image moment is a weighted average (moment) of an images pixels' intensities by a selected function. The image used is a binary mask of a segmented section of the original image. (i.e. a black/white or 0/1 box around a segmented image section; in this case, a black box with a white particle inside it), so only contains information about the image (particle) shape, size, and orientation. The moments are calculated using openCVs function cvMoments, which calculates spatial and central moments up to the third order. Moments can be used to calculate the center of gravity, area, main axes, inertia, and other shape characteristics. In this case, the moments are used to calculate Hu moments (cvGetHuMoments). Hu moments are linear combinations of the normalized central moments which are nominally invariant under translation, changes in scale, and rotation. They are scalar quantities. Hu moments are used in image processing for image/shape matching (such as letter recognition), since they are invariant to scale/rotation/translation. 6 Hu moments are calculated, but only 2 are currently displayed.

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

Simple Texture Measures	
Grayscale mean	When calculating grayscale statistics for an object, a mask which defines the set of pixel locations belonging to an object is used. At every pixel location defined by this mask, the corresponding grayscale pixel intensity value is extracted and used in calculations. Mean (m) is the sum of the pixel intensity values divided by the number of pixels in the particle. $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i.$ The measurement should not be used to compare different samples or images, since it will vary greatly by image acquisition conditions. It can be used to select and label particles/sub particles in a single image.
Grayscale variance	Grayscale variance (s^2) is the mean of the squared deviation (difference) of the pixel values from the grayscale mean. (this is the square of standard deviation) $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mu)^2.$ The measurement should not be used to compare different samples or images, since it will vary greatly by image acquisition conditions. It can be used to select and label particles/sub particles in a single image.
Gradient Mean	This uses cvSobel (Sobel operator, a differentiation operator) to calculate an approximation of the gradient of the objects pixel intensity values in the x and y directions, and then combines them to give a single gradient magnitude value for each pixel location. The mean of these values are then calculated. The measurement should not be used to compare different samples or images, since it will vary greatly by image acquisition conditions. It can be used to select and label particles/sub particles in a single image.
Gradient Variance	This is the variance in the gradient magnitude values. The measurement should not be used to compare different samples or images, since it will vary greatly by image acquisition conditions. It can be used to select and label particles/sub particles in a single image.

DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory

3. 0 Visual Examples:



DHS Funding Statement: The Department of Homeland Security sponsored the production of this material under the Department of Energy contract for the management and operation of Los Alamos National Laboratory