CIRCULARITY OF OBJECTS IN IMAGES

Murk J. Bottema

School of Information Science and Engineering
Flinders University of South Australia
PO Box 2100, Adelaide, SA 5001, Australia
The Cooperative Research Centre for Sensor Signal and Information Processing
Warrendi Road, Mawson Lakes, SA 5095, Australia
email: murkb@ist.flinders.edu.au

ABSTRACT

The most commonly used measure of circularity of objects in images is shown to give incorrect results. An alternative measure of circularity based on the distance between a set and a discrete disk is described. The alternative measure gives circularity zero (distance zero) for discrete disks and values in the range (0,1] for discrete sets which are not disks.

1. INTRODUCTION

In order to detect or classify objects in images, features are extracted which are thought to provide discriminatory information. For example, to detect the presence of lobular calcifications in a digital mammogram, knowledge that such calcifications are generally spherical in shape and of higher density than surrounding tissue, motivates searching for peaks in the intensity surface with circular cross section [2]. This settles the choice of feature, but a method for extracting the feature, in this case a measure of the circularity of the region, must be selected also.

Ultimately, the feature measured is defined by the algorithm chosen for extraction, regardless of the original intention. The value of the feature in terms of discriminatory power and its accuracy in representing the intented feature are really separate issues. In most cases, however, it is highly desirable to chose methods of feature extraction which measure the intended feature faithfully.

In the case of measuring the circularity of an object, a popular technique is to measure the perimeter and the area of set of pixels representing the object and to compute

$$\frac{\text{perimeter}^2}{4\pi \text{ area}} \quad (1)$$

The author would like to thank Stephen Weissenhofer for many discussions and for reading a draft of the paper.

The purpose of this paper is first to show that this measure does not represent the circularity of the object: all disks are not equally circular and squares are more circular than disks. The second goal is to show that an alternative method exists for measuring circularity based on the "distance" between a set and ideal circular set which is easy to compute and which does not suffer from the inadequacies of (1).

1.1. Notation and conventions

A digital image is viewed as a function on $\mathbb{R}^2$ which is constant on pixels. Normally pixels are squares of unit side length centered at integer valued coordinates in the plane and are identified by their centers. Thus a pixel also may be viewed as a point with integer coordinates, that is, an element of the discrete array $\mathbb{Z}^2$. The two views will be allowed to coexist, with context dictating which is preferred in any instance. In a few cases, pixels of side length $s \neq 1$ will be considered with the obvious adjustments taken as understood.

It will be important to distinguish between sets in $\mathbb{R}^2$ which consist of unions of pixels (called discrete sets) and general sets in $\mathbb{R}^2$. To do so, calligraphic letters, $\mathcal{A}, \mathcal{B}, \ldots$, will be used to indicate general sets and regular capital letters, $A, B, \ldots$, will be used for sets which are the unions of pixels. The discretization of $\mathcal{A}$, is the set of pixels whose centers lie in $\mathcal{A}$. Although the precise statements which follow depend on this choice of discretization, the essence of the discussion is valid for other reasonable choices.

The discrete disk $D(c,r)$ is the discretization of the set $\mathcal{D}(c,r) = \{x \in \mathbb{R}^2 : \|x - c\| \leq r\}$, (Fig. 1). Note that by this convention, it makes sense to discuss discrete disks with non-integer valued centers and radii and that it may happen that $D(c_1, r_1) = D(c_2, r_2)$ even when $c_1 \neq c_2$ or $r_1 \neq r_2$ or both. Also, note that the geometric center of $D(c,r)$ need not be c.

0-7803-6293-4/00/$10.00 ©2000 IEEE.

The area of a set A will be denoted by $|A|$ or $|\mathcal{A}|$. For discrete sets, $|A|$ is just the number of pixels of A.

Figure 1. Examples of discrete disks. On the left is $D((0,0),3.3)$, the middle is $D((.5,0),0.6)$, and on the right is $D((.4,.2),3.3)$

2. SHORTCOMINGS OF $\frac{\text{PERIMETER}^2}{4\pi\,\text{AREA}}$

The formula in (1) is appealing because for general sets in $\mathbb{R}^2$, the shape which maximizes the area for a given perimeter is the disk. Hence the formula gives value 1 for disks and is strictly larger for other shapes.

The fact that this property does not carry over to the discrete setting has been noted previously [10]. Since the formula in (1) continues to appear, albeit under aliases "roundness" [6], "shape factor" [1], "compactness" [4, 6, 8], "thinness ratio" [9], as well as circularity [5, 1, 9, 8], it is worthwhile to delineate it's shortcomings.

The culprit in (1) is perimeter. This is not a quantity which can be approximated well in digital images. There are theoretic justifications for this statement, but here this point will be illustrated by showing the consequence of various popular methods for computing perimeter on circularity according to (1).

2.1. Perimeter by counting pixel edges

Edges are taken to be the usual vertical and horizontal line segments which define the boundaries between pixels. If perimeter is measured by counting the number of pixel edges which separate pixels inside the set from pixels outside the set, then the discrete disk in Fig. 2 has circularity ≈ 1.8093. The square in the Fig. 2 has much larger area than the disk, but the same perimeter. It has circularity ≈ 1.2732. Not only is the circularity of the disk far from 1, the disk is not the shape having the maximum area for a given perimeter. The formula in (1) measures "squareness" rather than circularity.

The situation is not improved if the resolution is increased. If the the disk $\mathcal{D}((0,0),4.2)$ and the square in Fig. 2 are discretized using pixels of sidelength s, then for the disk

$$\lim_{s\to 0} \frac{P(s)^2}{4\pi A(s)} = \frac{33.6^2}{4\pi^2(4.2)^2} \approx 1.6211$$

and for the square

$$\lim_{s\to 0} \frac{P(s)^2}{4\pi A(s)} = \frac{36^2}{4\pi 81} \approx 1.2732.$$

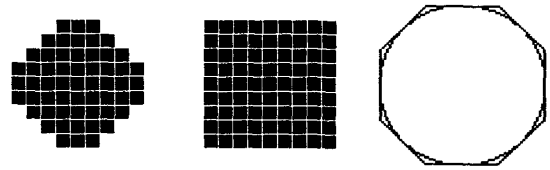

Figure 2. On the left is the discrete disk $D((0,0),4.2)$. The perimeter of this disk, both in terms of number of the number of pixel edges or in terms of number of 4-connected boundary pixels is the same as the perimeter of the 9×9 square in the middle. But the square has greater area and so is more circular than the disk by the formula in (1). On the right are the 8-connected boundary pixels for the discrete disk $D((0,0),37)$ and for an octagon. If the perimeter is measured by adding the lengths of line segments connecting the centers of 8-connected boundary pixels, these figures have equal perimeter, but the area of the octagon is larger than the disk.

2.2. Perimeter by 4-connected boundary pixels

In this case the disk and the square in Fig. 2 both have perimeter 32. The circularity is ≈ 1.4296 for the disk and ≈ 1.0060 for the square. The limiting values or circularity for increasing resolution are the same as the pixel edge counting method presented above.

2.3. Perimeter by allowing diagonal segments

The perimeter is sometimes defined as the sum of the lengths of line segments connecting the centers of the adjacent pixels in the 8-connected boundary of the set. This allows diagonal line segments of length $\sqrt{2}$ in addition to vertical and horizontal segments of unit length. With this increased flexibility, the circularity defined in (1) seems to improve. The circularity of the discrete disk in Fig. 2 is now ≈ 1.0787 and the circularity of the square is ≈ 1.2732. At least the disk is more circular than the square. But this particular disk only has circularity near 1 because it is very close to the octagon of the same area. The right most example in Fig. 2 shows another discrete disk and the regular octagon of the same perimeter as measured by the technique of this paragraph. The octagon has greater area and thus has lower circularity than the disk. By allowing diagonal line segments in computing the perimeter, the formula in (1) apparently measures "octagonality".

3. CIRCULARITY BASED ON A METRIC

To define the circularity of a discrete set without relying on the perimeter, a metric is introduced which measures the "distance" between the set and a discrete disk.

For discrete sets A and B, the function d defined by

$$d(A,B) = \frac{|A \setminus (A \cap B)| + |B \setminus (A \cap B|}{|A| + |B|} \quad (2)$$

has the following properties:

(a) $d(A,B) \geq 0$,
(b) $d(A,B) = 0$ if and only if A = B,
(c) $d(A,B) = d(B,A)$.

However, this function is not a metric on the set of non-empty discrete sets in the plane, because the triangle inequality fails.

If the class of discrete sets is limited to sets of a fixed area $M > 0$, then

$$d(A,B) + d(B,C) - d(A,C) = \frac{|B \setminus (A \cup C)| + |(A \cap C) \setminus (A \cap B \cap C)|}{M} \geq 0$$

and so d in (2) is a metric on this restricted collection of discrete sets.

If A is any discrete set, and D is a discrete disk having the same area as A, then

$$d(A,D) = \frac{|A \setminus (A \cap D)|}{|A|} \quad (3)$$

Loosely speaking, the definition of circularity of a discrete set A will be $d(A,D)$ where D is the discrete disk having the same area as A and the same center of mass. The reason for the "loosely speaking" caveat is that such a discrete disk does not always exist. For example, the set A in Fig. 3 has center $(0,0)$ and area 3 but the discrete disk $D((0,0),r)$ has area 1 for $r < 1$ and area ≥ 5 for $r \geq 1$.

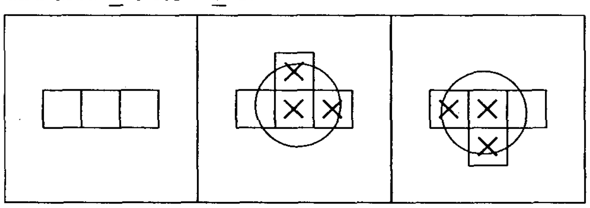

Figure 3. The left panel shows a three pixel set A for which there does not exist a discrete disk having the same center and area. The center and right panel each show an example of a discrete disk D (pixels marked $\times$) having the same area as A. In each case the circle indicates the boundary of a disk $\mathcal{D}$ for which D is the discretization.

However, the discrete disk $D((1/6,1/6),7/6)$ does have area 3 and has center not very far from $(0,0)$. More importantly, the same holds for any discrete disk $D((\epsilon,\epsilon),1+\epsilon)$ where $0 < \epsilon < 1/6$ and in addition

$$D((\epsilon,\epsilon),1+\epsilon) = D((1/6,1/6),7/6)$$

for $0 < \epsilon < 1/6$. In summary, a discrete disk has been found which has area equal to the area of A and which is the discretization of a disk with center arbitrarily close to the center of A. When this disk is used in formula (3), the result is $d(A,D) = 1/3$ which will be the circularity of A as formally defined below. The discrete disk found in this example is not unique as is apparent from Fig. 3. The following definition provides a unique discrete disk for use in formula (3).

Definition 3.1

Given a point c in the plane, $D_n = D_n(c)$ is the set of "n closest pixels" to c. More specifically, p_1 is the pixel such that $||c - p_1|| \leq ||c - q||$ for all pixels q. If there is more than one such pixel, choose the one for which the vector $p_1 - c$ makes the smallest positive angle with the positive x-axis. The sets D_n are defined recursively by $D_1 = p_1$ and $D_n = D_{n-1} \cup p_n$ where p_n is the pixel such that $p_n \notin D_{n-1}$ and $||c-p_n|| \leq ||c-q||$ for all pixels $q \notin D_{n-1}$. If there is more than one such pixel, choose the one for which the vector $p_n - c$ makes the smallest positive angle with the positive x-axis.

Theorem 3.2

Let c be a point in the plane. For each positive integer n, the set D_n defined above has area n and either there exists a disk $\mathcal{D}$ with center c such that D_n is the discretization of $\mathcal{D}$ or there exists a disk $\mathcal{D}$ with center arbitrarily close to c such that D_n is the discretization of $\mathcal{D}$.

A proof appears in the appendix.

Definition 3.3

For a non-empty discrete set A with area $|A| = n$, the circularity of A is defined to be $d(A, D_n(c))$, where d is as in (3), c is the geometric center of A, and $D_n(c)$ is as in Definition 3.1.

4. EXAMPLES

In this section, circularity will refer to Definition 3.3.

1. Every discrete disk has circularity 0 and every discrete set which is not a discrete disk has circularity > 0.

2. The square in Fig. 2 has circularity ≈ 0.0494. In the limit of infinite resolution, the circularity of

any square is $(4/\pi)\tan^{-1}(\sqrt{(4/\pi)-1})-\sqrt{(4/\pi)-1}\approx 0.0905$.

3. The triangle in Fig. 4 has circularity ≈ 0.1872. In the limit of infinite resolution, the circularity of any equilateral triangle is $(3/\pi)\tan^{-1}(w)-(w/\sqrt{3})\approx 0.1825$, where $w=\sqrt{(3\sqrt{3}/\pi)-1}$.

4. The circularity of a regular n-sided polygon in the limit of infinite resolution is

$$\frac{n}{\pi}\left(\tan^{-1}\left(\sqrt{w(n)-1}\right)-\frac{\sqrt{w(n)-1}}{w(n)}\right)$$

where $w(n)=(n/\pi)(\tan(\pi/n))$.

5. The snake in Fig. 4 has circularity ≈ 0.5763.

6. The "c-shaped" set in Fig. 4 has circularity 1.

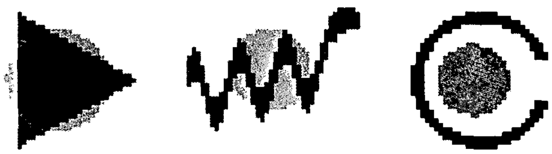

Figure 4. Three sets (dark) with their discrete disks (gray) as defined in Definitions 3.1 and 3.3.

5. DISCUSSION

The definition of circularity proposed in this paper is very similar to the definition of circularity given by

$$\text{circularity}(A)=\frac{|A\cap D|}{|A|}, \tag{4}$$

where D is the disk centered at the geometric center of A and having radius $r=\sqrt{A/\pi}$. This definition was introduced in [3] and has been used elsewhere [7, 2]. The connection to the definition introduced here is clear if the formula in (3) is written as

$$d(A,D)=\frac{|A\setminus(A\cap D)|}{|A|}=1-\frac{|A\cap D|}{|A|} \tag{5}$$

However, the definition in (4) does not address the issue of the existence of a discrete disk D with the desired properties. For the set in Fig. 3, for example, the disk to use for computing circularity via (4) is $\mathcal{D}((0,0),\sqrt{3/\pi})$, which when discretized, is a disk of area 1 instead of area 3.

6. APPENDIX: PROOF OF THE THEOREM.

Let $p_1,p_2,...,p_n$ be the pixels chosen according to Definition 3.1. Let $r=||c-p_n||$. If $||c-q||>r$ for all pixels $q\notin D_n$, then $D_n=D(c,r)$. The other possibility is that there exist $q_1,...,q_m\notin D_n$ such that $||c-q_j||=r$ and $||c-q||>r$ for all $q\notin D_n\cup_{j=1}^m q_j$. In this case, set $r_q=\min\{||c-q||:||c-q||>r\}$. Let k be the smallest integer such that $||c-p_i||=r$ for $i=k,k+1,...,n$. If $k>1$, set $r_p=\max\{||c-p_i||:i=1,2,...,k-1\}$ and if $k=1$, set $r_p=0$. Note that $r_p<r<r_q$.

For any $s\in\mathbb{R}^2$, let $\phi(s)$ be the positive angle between the x-axis and the vector $s-c$. Let w be the point such that $\phi(w)=(1/2)(\phi(p_k)+\phi(p_n))$ and $||w||=1$. Set $\lambda_0=\frac{1}{4}\min\{r-r_p,r_q-r\}$ and for $\lambda\in(0,\lambda_0]$ define $c_\lambda=c+\lambda w$. It is easy to check that for $r_\lambda=||c_\lambda-p_n||$, the disk $D(c_\lambda,r_\lambda)$ contains p_i, $i=1,2,...,k-1$ and does not contain any pixel q with $||c-q||>r$.

It remains to show that $p_i\in D(c_\lambda,r_\lambda)$ for $i=k,...,n$ and $q_j\notin D(c_\lambda,r_\lambda)$ for $j=1,...,m$. For any point $s\neq c$ in the plane, define $\theta_s=\angle(c_\lambda cs)\in[-\pi,\pi)$. By the way in which the pixels p_i were chosen, $|\theta_{p_i}|\leq|\theta_{p_n}|$ for $i=k,...,n$ and $|\theta_{q_j}|>|\theta_{p_n}|$ for $j=1,...,m$. Since $||q_j-c_\lambda||^2=\lambda^2+r^2-2\lambda r\cos(\theta_{q_j})$ and $||p_i-c_\lambda||^2=\lambda^2+r^2-2\lambda r\cos(\theta_{p_i})$, $||p_i-c_\lambda||\leq||p_n-c_\lambda||<||q_j-c_\lambda||$ for $i=k,...,n$ and $j=1,...,m$.

Thus $D_n=D(c_\lambda,r_\lambda)$ for any $\lambda\in(0,\lambda_0]$.

7. REFERENCES

[1] G. J. Awcock and R. Thomas. *Applied Image Processing.* Macmillan, London, 1995.

[2] M. J. Bottema and J. S. Slavotinek. Detection of subtle microcalcifications in digital mammograms. In N. Karssemeijer, M. Thijssen, J. Hendriks, and L. van Erning, editors, *Digital Mammography,* 1998.

[3] M. Giger, K. Doi, and H. MacMahon. Image feature analysis and computer-aided diagnosis in digital radiography. 3. automated detection of nodules in peripheral lung fields. *Medical Physics,* 15(2):158–166, 1988.

[4] R. C. Gonzales and R. E. Woods. *Digital Image Processing.* Addison-Wesley, Readiong, Mass., 1993.

[5] B. Jahne. *Practical Handbook of Image Processing for Scientific Applications.* CRC Press, Boca Raton, 1997.

[6] A. K. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, London, 1989.

[7] N. Petrick, H.-P. Chan, B. Sahiner, and D. Wei. An adaptive density-weighted contrast enhancement filter for mammographic breast mass detection. *IEEE Transactions on Medical imaging,* 15(1):59–67, 1996.

[8] I. Pitas. *Digital Image Processing Algorithms.* Prentice Hall, New York, 1993.

[9] W. Pratt. *Digital Image Processing.* Wiley, New York, 1991.

[10] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, London, 1982.