RAPID LOCATION OF CONVEX OBJECTS IN DIGITAL IMAGES

E. R. Davies
Machine Vision Group, Department of Physics
Royal Holloway, University of London
Egham Hill, Egham, Surrey, TW20 0EX, UK
Tel: +44 1784 443497; fax: +44 1784 472794
e-mail: e.r.davies@rhbnc.ac.uk

ABSTRACT

This paper studies sampling strategies for the rapid location of objects in digital images, and shows how point sampling can be used to minimise computational effort. The process can be extremely efficient, especially when the image space is sparsely populated and large convex objects are being detected. In the case of ellipses, exact location is considerably aided by the new 'triple bisection' algorithm. The approach has been applied successfully to the location of well separated nearly elliptical cereal grains which are to be scrutinised for damage and varietal purity.

1 INTRODUCTION

Machine vision is a task orientated subject which involves detailed image analysis. In general, objects and features have to be located in an image before sufficient understanding of its content can be achieved for specific task orientated decisions to be made. However, the low and intermediate level problems of object location often take a substantial proportion of the overall processing effort, as they involve unconstrained search over the image data: nowhere is this felt more than in industrial inspection where real-time implementation is mandatory. This paper is concerned with such real-time applications, and in particular with the problem of extremely rapid location of objects in 2-D images, a topic on which relatively little systematic work has been carried out (but see [1–5]).

We shall concentrate on images which have non-complex intensity patterns so that edge detection and thresholding techniques are reasonably effective, and Hough transforms provide an obvious path to object recognition [4].

2 EXTENDING THE SAMPLING APPROACH

In an earlier project, the author had the problem of finding the centres of circular objects such as coins and biscuits significantly more rapidly than for a conventional Hough transform, while retaining as far as possible the robustness of that approach [2]. The best solution appeared to be to scan along a limited number of horizontal lines in the image recording and averaging the x-coordinates of mid-points of chords of any objects, and repeating the process in the vertical direction to complete the process of centre location. The method was successful and led to speedup factors as high as 25 in practical situations.

In the present project, extreme robustness was not necessary, and it seemed worth finding how much faster the scanning concept could be taken. It was envisaged that significant improvement might be achieved by taking a minimum number of sampling points in the image rather than by scanning along whole lines.

Suppose that we are looking for an object such as that shown in Figure 1(a), whose shape is defined relative to a reference point R as the set of pixels $A = \{\mathbf{r}_i: i = 1 \text{ to } n\}$, n being the number of pixels within the object. If the position of R is $\mathbf{x}_R$, pixel i will appear at $\mathbf{x}_i = \mathbf{x}_R + \mathbf{r}_i$. This means that when a sampling point $\mathbf{x}_s$ gives a positive indication of an object, the location of its reference point R will be $\mathbf{x}_R = \mathbf{x}_s - \mathbf{r}_i$. Thus the reference point of the object is known to lie at one of the set of points $U_R = \cup_i (\mathbf{x}_s - \mathbf{r}_i)$, so knowledge of its location is naturally incomplete. Indeed, the map of possible reference point locations has the same shape as the original object, but rotated through 180° – because of the minus sign in front of $\mathbf{r}_i$. Furthermore, the fact that positions of the reference point are only determined within n pixels means that many sampling points will be needed, the minimum number required to cover the whole image clearly being N/n, if there are N pixels in the image. This means that the optimum speedup factor will be N/(N/n) = n, as the number of pixels visited in the image is N/n rather than N.

Unfortunately, it will not be possible to find a set of sampling point locations such that the 'tiling' produced by the resulting maps of possible reference point positions covers the whole image without overlap. Thus there will normally be some overlap (and thus loss of efficiency in locating objects) or some gaps (and thus loss of effectiveness in locating objects). Clearly, the set of tiling squares shown in Figure 1(b) will only be fully effective if square objects are to be located.

However, a more serious problem arises because objects may appear in any orientation. This prevents an ideal tiling from being found. It appears that the best that can be achieved is to search the image for a maximal *rotationally invariant* subset of the shape, which must be a circle, as indicated in Figure 2. Furthermore, as no perfect tiling for circles exists, the tiling that must be chosen is either a set of hexagons or, more practically, a set of squares. This means that the speedup factor for object location will be significantly less than n, though it will still be substantial.

3 APPLICATION TO GRAIN INSPECTION

The application we had in mind when approaching this problem was that of fast location of grains on a conveyor in order to scrutinise them for damage and varietal purity. Under these circumstances it is best to examine each grain in isolation: specifically, touching or overlapping grains would be difficult to cope with. Thus we envisaged the grains as being spread out with at most 25 grains being visible in any 256×256 image. Under these circumstances there would be an intensive search problem, with far more pixels having to be considered than would otherwise be the case. Hence a very fast object location algorithm would be of especial value.

We have found that wheat grains are well approximated by ellipses in which the ratio of semi-major (a) to semi-minor (b) axes is almost exactly two. The deviation is normally less than 10%, in spite of some quite large apparent differences between the intensity patterns for different grains. Hence it seemed worth using this model as an algorithm optimisation target. First, the (non-ideal) $L \times L$ square tiles would appear to have to fit inside the circular maximal rotationally invariant subset of the ellipse, so that $\sqrt{2}L = 2b$, i.e. $L = \sqrt{2}b$. This value should be compared with the larger value $L_0 = (4/\sqrt{5})b$ which could be used if the grains were constrained to lie parallel to the image x-axis – see Figure 3 (here we are ignoring the dimensions $2\sqrt{2}b \times \sqrt{2}b$ for optimal *rectangular* sampling tiles).

Another consequence of the difference in shape of the objects being detected (here ellipses) and the tile shape (square) is that the objects may be detected at several sample locations, thereby wasting computation (see Section 2). A further consequence of this is that we cannot merely count the samples if we wish to count the objects: instead we must relate the multiple object counts together and find the centre locations of the objects. This also applies if the main goal is to locate the objects for inspection and scrutiny. In our case, the objects are convex, so we only have to look along the line joining any pair of sampling points to determine whether there is a break and thus whether they correspond to more than one object. We shall return later to the problem of systematic location of object centres.

For ellipses, it is relevant to know how many sample points could give positive indications for any one object. Now the maximum distance between one sampling point and another on an ellipse is 2a, and for the given eccentricity this is equal to 4b which in turn is equal to $2\sqrt{2}L$. Thus an ellipse of this eccentricity could overlap three sample points along the x-axis direction if it were aligned along this direction; alternatively, it could overlap just two sample points along the 45° direction if it were aligned along this direction, though it could in that case also overlap just one laterally placed sample point. In an intermediate direction (e.g. at an angle arctan 0.5 to the image x-axis), the ellipse could overlap four points. Similarly, it is easy to see that the minimum number of positive sample points per ellipse is two. The possible arrangements of positive sample points are presented in Figure 4.

Fortunately, the above approach to sampling is over-rigorous. Specifically, we have insisted upon the sampling tile being contained within the ideal (circular) maximal rotationally invariant subset of the shape. However, what is required is that the sampling tile must be of such a size that all possible orientations of the shape are allowed for. In the present example the limiting case that must be allowed for occurs when the ellipse is orientated parallel to the x-axis, and it must be arranged that it can just pass through four sampling points at the corners of a square, so that on any infinitesimal displacement, at least one sampling point is contained within it. For this to be possible it can be shown that $L = (4/\sqrt{5})b$, the same situation as already depicted in Figure 3. This leads to the possible arrangements of positive sampling points shown in Figure 5 – a distinct reduction in the average number of positive sampling points, which leads to useful savings in computation (the average number of positive sampling points per ellipse is reduced from ~3 to ~2).

Object location normally takes considerable computation because it involves an unconstrained search over the whole image space, and in addition there is normally (as in the ellipse location task) the problem that the orientation is unknown. This contrasts with a certain other aspect of inspection, that of object scrutiny and measurement, in that relatively few pixels have to be examined in detail, so relatively little computation is involved in this aspect. Clearly, the sampling approach outlined above largely eliminates the search aspect of object location, since it quickly eliminates any large tracts of blank background. Nevertheless, there is still the problem of refining the object location phase. One way of approaching this problem is to expand the positive samples into fuller regions of interest and then perform a restricted search over these regions. For this purpose we could use the same search tools that we might use over the whole image if sampling were not being performed. However, the preliminary sampling technique is so fast that this approach would not take full advantage of its speed. Instead we could use the following procedure.

For each positive sample, draw a horizontal chord to the boundary of the object, and find the local boundary tangents. Then use the chord-tangent technique (join of tangent intersection to mid-point of chord [4]) to determine one line on which the centre of an ellipse must lie. Repeat this for the all positive samples, and obtain all possible lines on which ellipse centres must lie. Finally, deduce what the possible ellipse centre locations are, and check each of them in detail in case some correspond to false alarms arising from objects which are close together rather than from genuine self-consistent ellipses. Note that in cases where there is a single positive sampling point, another positive sampling point has to be found (say L/2 away from the first).

We next propose an even faster approach, which we may call the *triple bisection* algorithm. Draw horizontal (or vertical) chords through adjacent vertically (or horizontally) separated pairs of positive samples, bisect them, join and extend the bisector lines, and finally find the mid-points of these bisectors (Figure 6). (In cases where there is a single positive sampling point, another positive sampling point has to be found, say L/2 away from the first.) This is the approach we have adopted in our studies on grain. It has the additional advantage of not requiring estimates of tangent directions to be made at the ends of chords, which can prove inaccurate when objects are somewhat fuzzy, as in our grain images. The result of applying this technique to an image containing mostly well-separated grains is shown in Figure 7: this illustrates that the whole procedure for locating grains by modelling them as ellipses and searching for them by sampling and chord bisection approaches is a viable one. In addition, the procedure is very fast, as the number of pixels that are visited is a small proportion of the total number in each image.

Finally, we show why the triple bisection algorithm presented above is appropriate. First note that it is correct for a circle, for reasons of symmetry. Second, note that in orthographic projection, circles become ellipses, straight lines become straight lines, parallel lines become parallel lines, chords become chords, and midpoints become midpoints. Hence choosing the right orthogonal projection to transform the circle into a correctly orientated ellipse of appropriate eccentricity, the midpoints and centre location shown in the diagram of Figure 6 must be validly marked. This proves the algorithm.

4 CONCLUDING REMARKS

This paper has studied sampling strategies for the rapid location of objects in digital images. Motivated by the success of an earlier line-based sampling strategy [2], it has shown that point samples lead to the minimum computational effort when the 180°-rotated object shapes form a perfect tiling of the image space. In practice imperfect tilings have to be used, but these can be extremely efficient, especially when the image intensity patterns permit thresholding, the images are sparsely populated with objects, and the latter are convex in shape. In important feature of the approach is that detection speed is *improved* for larger objects, though naturally exact location involves some additional effort. In the case of ellipses, the latter process is considerably aided by the new triple bisection algorithm.

We have applied the approach successfully to the location of well separated cereal grains, which can be modelled as ellipses with 2:1 aspect ratio, prior to scrutiny for damage and varietal purity.

Acknowledgement

Thanks are due to J. Chambers and C. Ridgway of CSL, York, UK for providing the original grain images used for testing the algorithm.

REFERENCES

1. VanderBrug, G.J. and Rosenfeld, A. (1977) "Two-stage template matching", IEEE Trans. Comput., **26**, pp. 384–393
2. Davies, E.R. (1987) "A high speed algorithm for circular object location", Pattern Recogn. Lett., **6**, no. 5, pp. 323–333
3. Davies, E.R. (1987) "Lateral histograms for efficient object location: speed versus ambiguity", Pattern Recogn. Lett., **6**, no. 3, pp. 189–198
4. Davies, E.R. (1997) *Machine Vision: Theory, Algorithms, Practicalities*, Academic Press (2nd edition), pp. xxxi + 750
5. Davies, E.R. (1992) "A skimming technique for fast accurate edge detection", Signal Process., **26**, no. 1, pp. 1–16

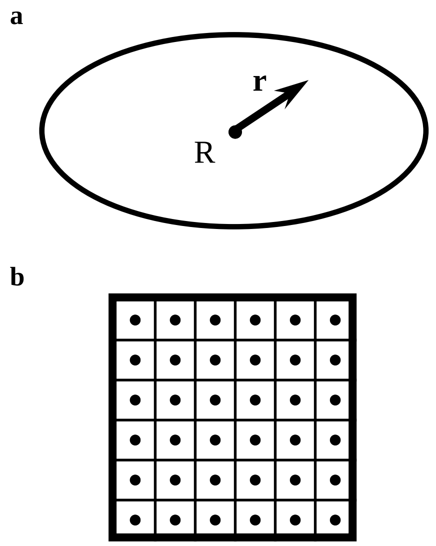

Figure 1 Object shape and method of sampling. (a) object shape, showing reference point R and vector **r** pointing to a general location $\mathbf{x}_R + \mathbf{r}$. (b) Image and sampling points, with associated tiling squares.

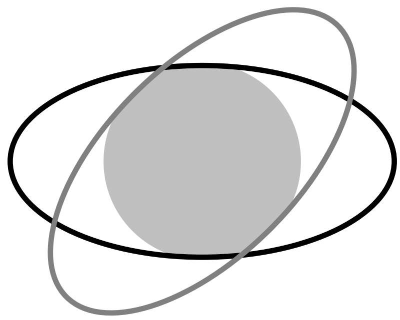

Figure 2 Ellipse in two orientations and maximal rotationally invariant subset.

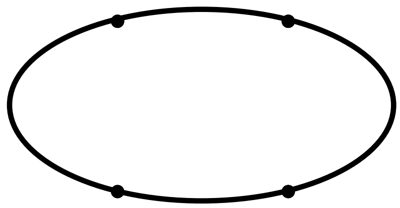

Figure 3 Horizontal ellipse and geometry showing size relative to largest permitted spacing of sampling points.

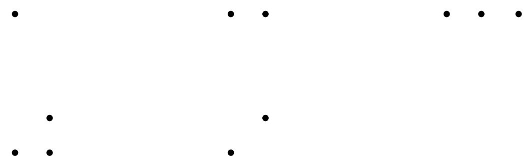

Figure 5 Possible arrangements of positive sampling points for ellipse, taking L = (4/√5)b.

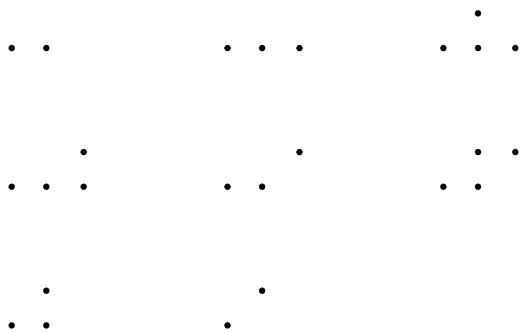

Figure 4 Possible arrangements of positive sampling points for ellipse, taking L = √2b.

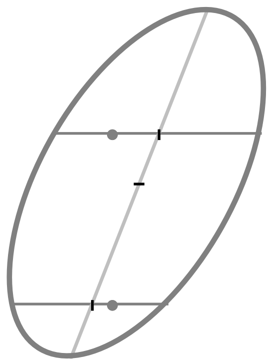

Figure 6 Illustration of triple bisection algorithm. The round spots are the sampling points, and the short bars are the midpoints of the three chords, the short horizontal bar being at the centre of the ellipse.

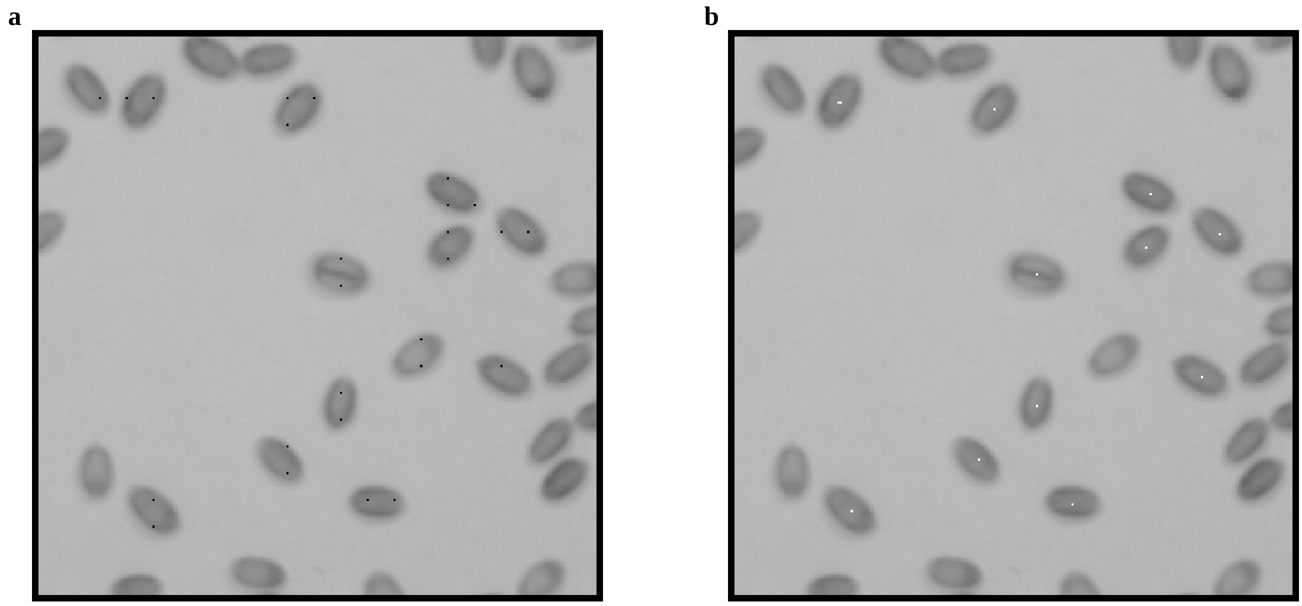

Figure 7 Image showing grain location using new sampling approach. (a) sampling points. (b) final centre locations.