CHAPTER

Inspection of Cereal Grains

21

Inspection of cereal grains is an exceptionally mundane and repetitive task, and to some extent contrasts with many of the examples given in Chapter 20 in that the emphasis is on detecting contaminants rather than finding manufacturing faults. This chapter presents three case studies that cover the topic and at the same time air and solve relevant theoretical questions.

Look out for:

- the limitations of the immediately obvious technique—global thresholding.
- the value of morphological filtering.
- the relatively unusual need for median filtering as a morphological operation.
- the use of bar (linear feature) detectors for locating insects.
- how the vectorial bar detector operator is optimized.
- how sampling can be used to speed up object location.
- how the outputs of oblique template masks can optimally be combined.

This chapter repeatedly alludes to the problems of achieving real-time operation and largely solves these problems for moderately demanding situations (flow rates up to ∼300 items per second): for more demanding situations, dedicated hardware accelerators are needed, as discussed in Chapter 26. At a more detailed level, the balance between false positives (false alarms) and false negatives needs to be optimized, as discussed in Section 24.7.

21.1 INTRODUCTION

Cereal grains are among the most important of the foods we grow. A large proportion of cereal grains is milled and marketed as flour: this is then used to produce bread, cakes, biscuits and many other commodities. Cereal grains can

also be processed in a number of other ways (such as crushing), and thus they form the basis of many breakfast cereals. In addition, there are some cereal grains, or cereal kernels, which are eaten with a minimum of processing: rice is obviously in this category—although whole wheat and oat grains are also consumed as decorative additives to "granary" loaves.

Wheat, rice, and other cereal grains are shipped and stored in vast quantities measured in millions of tons, and a large international trade exists to market these commodities. Transport also has to be arranged between farmers, millers, crushers, and the major bakers and marketers. Typically, transit by road or rail is in relatively small loads of up to 20 tons, and grain depots, warehouses, and ports are not unlikely to receive lorries containing such consignments at intervals ranging from 20 min to as little as 3 min. All the necessary transportation and storage result in grains being subject to degradation of various sorts: damage, molds, sprouting, insect infestation, and so on. In addition, the fact that grains are grown on the land and threshed implies the possibility of contamination by rodent droppings, stones, soil, chaff, and foreign grains. Finally, the quality of grain from various sources will not be uniformly high, and varietal purity is an important concern.

These factors mean that ideally, the grain that arrives at any depot should be inspected for a good many possible causes of degradation. This chapter is concerned with grain inspection. In the space of one chapter, we shall not aim to cover all possible methods and modes of inspection. Indeed, this would be impossible as the subject is moving ahead quite fast: not only are inspection methods evolving rapidly, but the standards against which inspection is carried out are also evolving fairly quickly. To some extent, the improvement of automatic inspection methods and the means of implementing them efficiently in hardware are helping to drive the process onward. We shall explore the situation with the aid of three main case studies, and then we shall look at the overall situation.

The first case study involves the examination of grains to locate rodent droppings and molds, such as ergot. The second case study considers how grains may be scrutinized for insects, such as the saw-toothed grain beetle. The third case study is concerned with inspection of the grains themselves. However, this case study is more general, and is less involved with the scrutiny of individual grains than with how efficiently they can be located: this is an important factor when lorry loads of grain are arriving at depots every few minutes—leading to the need to sample of the order of 300 grains per second. As remarked in Chapter 20, object location can involve considerably more computation than object scrutiny and measurement.

21.2 CASE STUDY: LOCATION OF DARK CONTAMINANTS IN CEREALS

As noted above, there is a demand for grain quality to be monitored before processing to produce flour, breakfast cereals, and a myriad of other derived products. Early work in this area was applied mainly to the determination of grain quality

(Ridgway and Chambers, 1996), with concentration on the determination of varietal purity (Zayas and Steele, 1990; Keefe, 1992) and the location of damaged grains (Liao et al., 1994). In fact, there is also the need to detect insect infestations and other important contaminants in grain—while not being confused by up to $\sim 2\%$ "permitted admixture" such as chaff and dust. The inspection work described in this section (Davies et al., 1998a) pays particular attention to the detection of noninsect contaminants. Relevant contaminants in this category include rodent (especially rat and mouse) droppings, molds such as ergot, and foreign seeds. (Note that ergot is poisonous to humans, so locating any instances of it is of especial importance.) In this case study the substrate grain is wheat, and foreign seeds such as rape would be problematic if present in too great a concentration.

Many of the potential contaminants for wheat grains (and for grains of similar general appearance such as barley or oats) are quite dark in color. This means that thresholding is the most obvious approach for locating them. However, there are a number of problems, in that false alarms arise from shadows between grains, dark patches on the grains, rapeseeds, chaff and other admixture components: this means that further recognition procedures have to be invoked to distinguish between the various possibilities. As a result, the thresholding approach is not eventually as attractive as might *a priori* have been thought (Fig. 21.1(a) and (b)). This problem is exacerbated by the extreme speeds of processing required in real applications. For example, a successful device for monitoring lorry loads of grain might well have to analyze a 3-kg sample of grain in 3 min (the typical time between arrival of lorries at a grain terminal), and this would correspond to some 60,000 grains having to be monitored for contaminants in that time. This places a distinct premium on rapid, accurate image analysis.

This case study concentrates on monitoring grain for rodent droppings. As indicated above, these types of contaminant are generally darker than the grain background, but cannot simply be detected by thresholding since there are significant shadows between the grains, which themselves often have dark patches. In addition, the contaminants are speckled because of their inhomogeneous constitution and because of lighting and shadow effects. In spite of these problems, the contaminants are identifiable by human operators because they are relatively large and their shape follows a distinct pattern (e.g., an aspect ratio of three or four to one). Thus, it is the combination of size, shape, relative darkness, and speckle that characterizes the contaminants and differentiates them from the grain substrate.

Designing efficient and rapidly operating algorithms to identify these contaminants is something of a challenge, but an obvious way of tackling it is *via* mathematical morphology—as we shall see in Section 21.2.1.

21.2.1 Application of Morphological and Nonlinear Filters to Locate Rodent Droppings

As indicated above, the obvious approach to the location of rodent droppings is to process thresholded images by erosion and dilation. In this way, shadows between

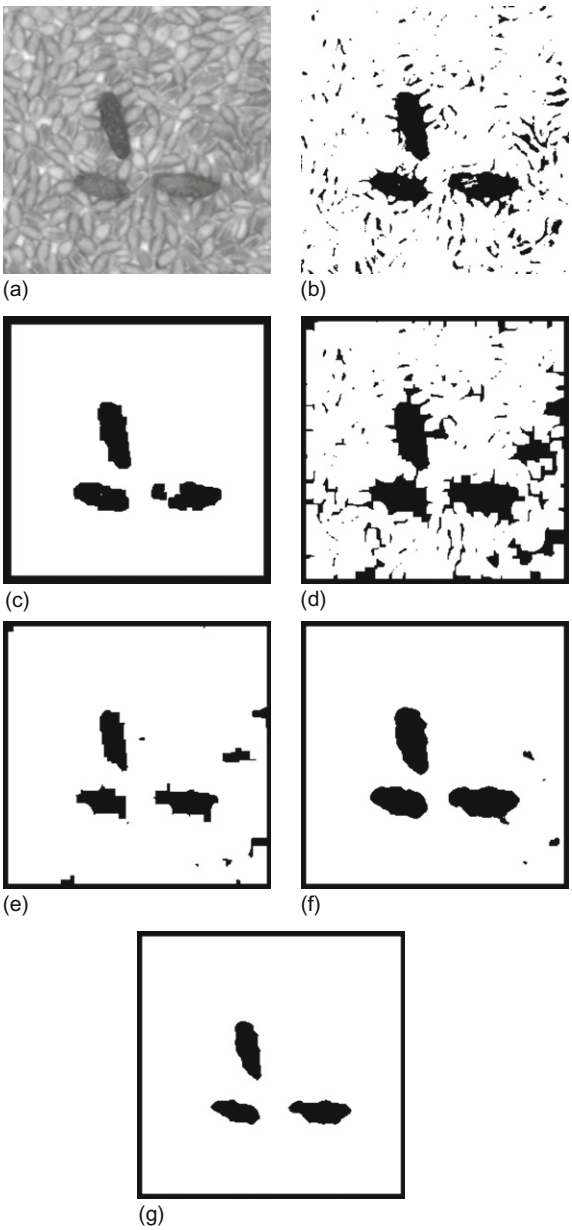


FIGURE 21.1

Effects of various operations and filters on a grain image. (a) Grain image containing several contaminants (rodent droppings). (b) Thresholded version of (a). (c) Result of erosion and dilation on (b). (d) Result of dilation and erosion on (b). (e) Result of erosion on (d). (f) Result of applying 11×11 median filter to (b). (g) Result of erosion on (f). In all cases, "erosion" means three applications of the basic 3×3 erosion operator, and similarly for "dilation."

grains, and discoloration of grains would be eliminated by the erosions, and the shapes and sizes of the contaminants restored by the subsequent dilations. The effect of this procedure is shown in Fig. 21.1(c). Note that the method has been successful in eliminating the shadows between the grains, but has been decidedly weak in coping with light regions on the contaminants. Remembering that while considerable uniformity might be expected between grains, the same cannot be said about rodent droppings, whose size, shape, and color vary quite markedly. Hence, the erosion−dilation schema has limited value, although it would probably be successful in most instances. Accordingly other methods of analysis were sought (Davies et al., 1998a).

The second approach is to overcome the problem of fragmentation of the contaminants which arose with the previous approach: this was attempted by first consolidating the contaminants by applying dilation before erosion. The effect of this approach is shown in Fig. 21.1(d). Notice that the result is to consolidate the shadows between grains even more than the shapes of the contaminants. Even when an additional few erosions are applied (Fig. 21.1(e)), the consolidated shadows do not disappear, and are of comparable sizes to the contaminants. Hence, the approach is not viable, and creates more problems than it solves. One possibility is to use the results of the earlier erosion−dilation schema as "seeds" to validate a subset of the dilation−erosion schema. However, this would be far more computation intensive and the results would clearly not be especially impressive (see Fig. 21.1(c) and (e)). Instead, a totally different approach was adopted.

The new approach was to apply a large median filter to the thresholded image, as shown in Fig. 21.1(f). This gives good segmentation of the contaminants, retaining their intrinsic shape to a reasonable degree, and suppresses the shadows between grains quite well. In fact, the shadows immediately around the contaminants enhance the sizes of the latter in the median filtered image, while some shadows further away are consolidated and retained by the median filtering. It was found useful to perform a final erosion operation (Fig. 21.1(g)): this eliminates the extraneous shadows and brings the contaminants back to something like their proper size and shape: although the lengths are slightly curtailed, this is not a severe disadvantage. Overall, the median filtering−erosion schema gave easily the greatest fidelity to the original contaminants, while being particularly successful at eliminating other artifacts (Davies et al., 1998a). In this case, it seems that the median filter is acting as an analytical device that carefully meditates and obtains the final result in a single rigorous stage—thus avoiding the error-propagation inherent in a two-stage process.

Finally, although median filtering is intrinsically more computation intensive than erosion and dilation operations, many methods have been devised for accelerating median operations (see Chapter 3), and indeed the speed aspect was found to be easily soluble within the specification set out above.

21.2.2 Problems with Closing

The above case study was found to require the use of a median filter coupled with a final erosion operation to eliminate artifacts in the background. An earlier test similarly involved a closing operation (a dilation followed by an erosion), followed by a final erosion. In other applications, grains or other small particles are often grouped by applying a closing operation to locate the regions where the particles are situated (Fig. 7.6). It is interesting to speculate whether, in the latter type of approach, closing should occasionally be followed by erosion, and also whether the final erosions used in the tests were no more than *ad hoc* procedures or whether they were vital to achieve the defined goal.

The situation was analyzed by Davies (2000c) and is summarized in Sections 7.5.1 and 7.5.2. The analysis starts by considering two regions containing small particles with occurrence densities ρ_1, ρ_2, where $\rho_1 > \rho_2$ (Figs. 7.7 and 7.8). It finds that a final erosion is indeed required to eliminate a shift in the region boundary, the estimated shift δ being:[1]

$$\delta = 2ab\rho_2(a + b) \tag{21.1}$$

where a is the radius of the dilation kernel and b is the width of the particles in Region 2.

It is important to notice that if $b = 0$, no shift will occur, but for particles of measurable size this is not so. Clearly, if b is comparable to a or if a is much greater than 1 pixel, a substantial final erosion may be required. On the other hand, if b is small, it is possible that the two-dimensional shift will be less than 1 pixel. In that case, it will not be correctable by a subsequent erosion, but due allowance for the shift can be made during subsequent analysis. While in this work, the background artifacts were induced mainly by shadows around and between grains, in other cases impulse noise or light background texture could give similar effects, and care must be taken to select a morphological process that limits any overall shifts in region boundaries. In addition, notice that the whole process can be modeled and the extent of any necessary final erosion estimated accurately. More particularly, the final erosion is a necessary measure, and is not merely an *ad hoc* procedure (Davies, 2000c).

21.2.3 Ergot Detection Using the Global Valley Method

While the early parts of Section 21.2 emphasized the difficulty of locating dark contaminants in cereals, and concluded that, in general, morphological methods would be needed for this purpose, it turns out that ergot can be isolated by thresholding. However, there is an intrinsic difficulty in finding a suitable threshold value, and appealing to the intensity histogram shows that it usually approximates closely to a unimodal distribution with no real indication of where the best

[1] A full derivation of this result appears in Section 7.5.1.

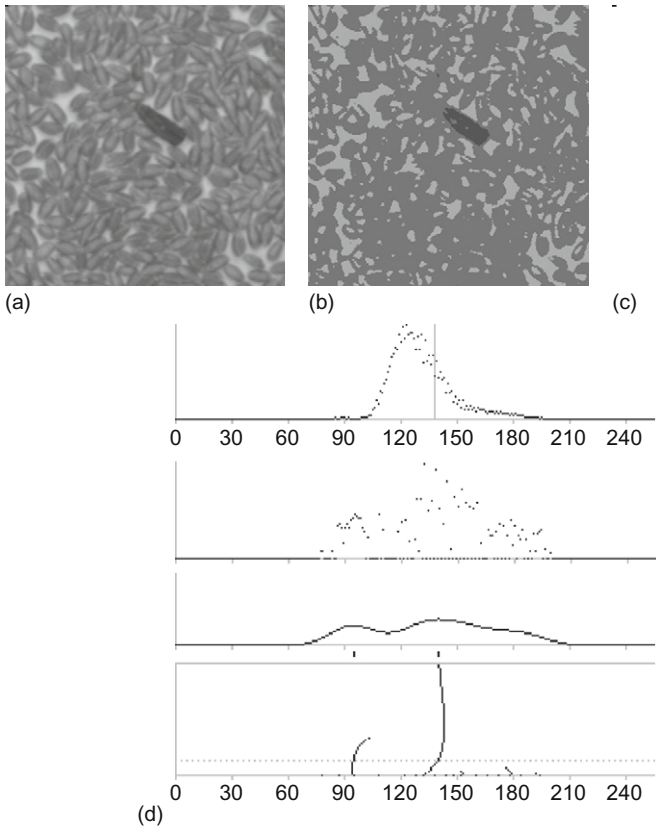


FIGURE 21.2

Location of ergot among wheat grains. (a) Original image. (b) Doubly thresholded image. (c) Result of only applying the lower threshold. (d) Top to bottom: Intensity histogram of (a). Result of applying the global valley transformation. Result of smoothing the global valley transform. The two thresholds used in (b) being located automatically at the dotted line. For further details, see text in Chapter 4.

Source: © *IET 2008*

thresholding point would be. However, the global valley transformation described in Chapter 4 is highly effective in this situation and leads to a global valley transform which, when smoothed, provides a virtually perfect thresholding value (Fig. 21.2). In fact, the method yields two threshold values, one which is useful for locating ergot and the other which is suited to separating wheat grains from a light conveyor background: clearly, there is no problem ignoring the latter threshold if it is not needed.

The reason why ergot can be detected in this way is that it is generally somewhat darker than rodent droppings. Ergot also differs from rat droppings in that the pieces are rather smaller, although in this respect they are not dissimilar to

mouse droppings. However, its most important characteristic is that it is poisonous to humans (rodent droppings are unlikely to be poisonous, even though they are clearly highly undesirable). Hence, a further method that can be employed for detecting this type of contaminant is of some value.

21.3 CASE STUDY: LOCATION OF INSECTS

The case study described in this section pays particular attention to the need to detect insects. Note that insects present an especially serious threat, because they can multiply alarmingly in a short span of time, so greater certainty of detection is vital in this case. This means that a highly discriminating method is required for locating adult insects (Davies et al., 1998b).

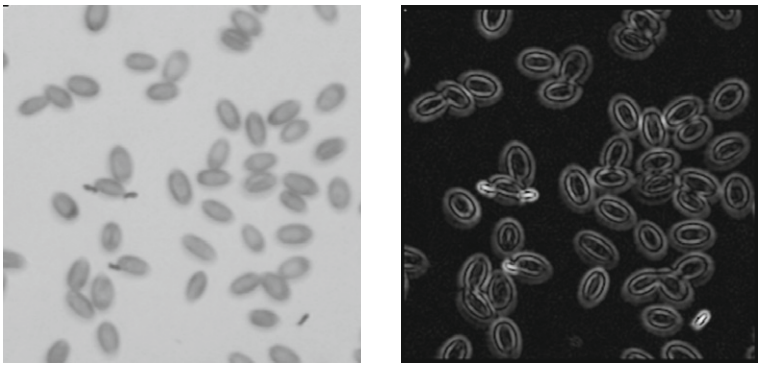
Not surprisingly, thresholding initially seemed to be the approach offering the most promise for locating insects, which appear dark relative to the light brown color of the grain. However, early tests on these lines showed that a good many false alarms would result from chaff and other permitted admixture, from less serious contaminants such as rapeseeds, and even from shadows between, and discolorations on, the grains themselves (Fig. 21.3). These considerations led to use of the linear feature detection approach for detecting small adult insects. This proved possible because these insects appear as dark objects with a linear (essentially bar-shaped) structure; hence attempts to detect them by applying bar-shaped masks led ultimately to a linear feature detector that had good sensitivity for a reasonable range of insect sizes. Before proceeding further, we consider the problems of designing linear feature detector masks.

21.3.1 The Vectorial Strategy for Linear Feature Detection

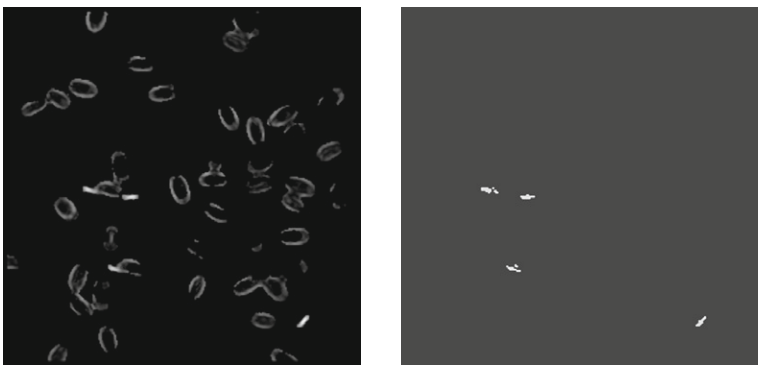
In earlier chapters we found that location of features in 3×3 windows typically required eight template masks to be applied in order to cope with arbitrary orientation. In particular, to detect corners, eight masks were required, while only one is needed to locate small holes because of the high degree of symmetry present in the latter case (Chapter 6). To detect edges by this means, four masks were required because 180° rotations correspond to a sign change that effectively eliminates the need for half of the masks: however, edge detection is a special case, as edges are vector quantities with magnitude and direction, and hence they are fully definable using just two component masks. In contrast, a typical line segment detection mask would have the form:

$$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix}$$

This indicates that four masks are sufficient, the total number being cut from eight to four by the particular rotation symmetry of a line segment. Curiously,



(a) (b)



(c) (d)

(e)

FIGURE 21.3

Insects located by line segment detector. (a) Original image. (b) Result of applying vectorial operator. (c) Result of masking (b) using an intensity threshold at a standard high level. (d) Result of thresholding (c). (e) Result of applying TM operation within (d).

while only edges qualify as strict vectors, requiring just two component masks, the same can be made to apply for line segments. To achieve this, a rather unusual set of masks has to be employed:[2]

$$L_0 = A\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \quad L_{45} = B\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{21.2}$$

The two masks are given different weights, so that some account can be taken of the fact that the nonzero coefficients are respectively 1 and $\sqrt{2}$ pixels from the center pixel in the window. These masks give responses g_0, g_{45} leading to:

$$g = (g_0^2 + g_{45}^2)^{1/2} \tag{21.3}$$

$$\theta = \frac{1}{2}\arctan\left(\frac{g_{45}}{g_0}\right) \tag{21.4}$$

where the additional factor of one half in the orientation arises as the two masks correspond to basic orientations of 0° and 45° rather than 0° and 90° (Davies, 1997c). Nevertheless, these masks can cope with the full complement of orientations as line segments have 180° rotational symmetry and the final orientation should only be determined within the range 0−180°. In Section 21.5, we comment further on the obvious similarities and differences between Eqs. (21.3), (21.4) and those pertaining to edge detection.

Next, we have to select appropriate values of the coefficients A and B. Applying the above masks to a window with the intensity pattern:

a	b	c
d	e	f
g	h	i

leads to the following responses at 0° and 45°:

$$g_0 = A(d + f - b - h) \tag{21.5}$$

$$g_{45} = B(c + g - a - i) \tag{21.6}$$

Using this model, theory and simulations were carried out (Davies, 1997c) for the case of a line segment of width w passing through the center of the 3×3 window, pixel responses being taken in proportion to the area of the line falling within each pixel. These showed that high orientation accuracy occurred for $w = 1.4$, when $B/A = 0.86$, giving a surprisingly low maximum error of just 0.4°. However, in this application, high orientation accuracy was not required, so

[2]The proof of the concepts described here involves describing the signal in a circular path around the current position and modeling it as a sinusoidal variation, which is subsequently interrogated by two quadrature sinusoidal basis functions set at 0° and 45° in real space (but by definition at 90° in the quadrature space). The mapping between orientation θ in the quadrature space and orientation φ in real space leads to the additional factor of one half in Eq. (21.4).

Eq. (21.3) was used on its own to obtain an accurate estimate of line contrast using the two masks. In this way, high sensitivity could be attained in the detection of image features corresponding to insect contaminants.

21.3.2 Designing Linear Feature Detection Masks for Larger Windows

When larger windows are used, it is possible to design the masks more ideally to match the image features that have to be detected, because of the greater resolution then permitted. However, there are many more degrees of freedom in the design, and there is some uncertainty as to how to proceed. The basic principle (Davies et al., 1998b) is to use masks whose profile matches the intensity profile of a linear feature around a ring of radius R centered on the feature, and at the same time follows a particular mathematical model—namely an approximately sinusoidal amplitude variation. For a given linear feature of width w, the sinusoidal model will achieve the best match for a thin ring of radius R_0 for which the two arc lengths within the feature are each one quarter of the circumference $2\pi R_0$ of the ring. Simple geometry (Fig. 21.4) shows that this occurs when:

$$2R_0 \sin\left(\frac{\pi}{4}\right) = w \tag{21.7}$$

$$\therefore \quad R_0 = \frac{w}{\sqrt{2}} \tag{21.8}$$

The width ΔR of the ring should in principle be infinitesimal, but in practice, considering noise and other variations in the dataset, ΔR can validly be up to about 40% of R_0. The other relevant factor is the intensity profile of the ring, and how accurately this has to map to the intensity profile of the linear features to be

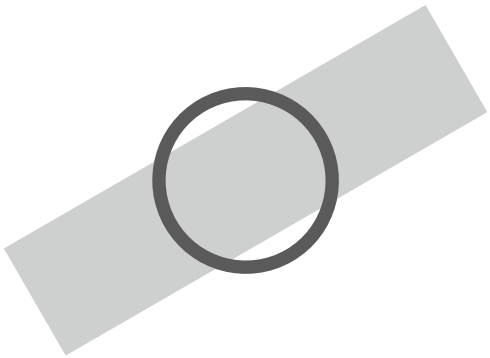


FIGURE 21.4

Geometry for application of a thin ring mask. Here two quarters of the ring mask lie within and two outside a rectangular bar feature.

Source: © World Scientific 2000

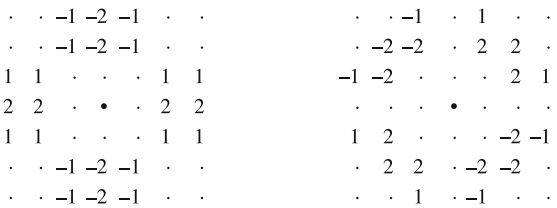


FIGURE 21.5

Typical (7 × 7) linear feature detection masks.

Source: © EURASIP 1998

located in the image. In many applications, such linear features will not have sharp edges, but will be slightly fuzzy and the sides will have significant and varying intensity gradient. Thus, the actual intensity profile is quite likely to correspond reasonably closely to a true sinusoidal variation. Masks designed on this basis proved close to optimal when experimental tests were made (Davies et al., 1998b). Figure 21.5 shows masks that resulted from this type of design process for one specific value of R (2.5 pixels).

21.3.3 Application to Cereal Inspection

The main class of insect that was targeted in this study (Davies et al., 1998b) was *Oryzaephilus surinamensis* (saw-toothed grain beetle): insects in this class approximate to rectangular bars of about 10 × 3 pixels, and masks of size 7 × 7 (Fig. 21.5) proved to be appropriate for identifying the pixels on the centrelines of these insects. In addition, the insects could appear in any orientation, and thus the vectorial approach to template matching was used, employing two masks as outlined above.

Preliminary decisions on the presence of the insects were made by thresholding the output of the vectorial operator. However, the vectorial operator responses for insect-like bars have a low-intensity surround, which is joined to the central response at each end, and between the low-intensity surround and the central response the signal drops close to zero except near the ends (Fig. 21.3(b)). This response pattern is readily understandable because symmetry demands that the signal be zero when the centers of the masks are coincident with the edge of the bar. To avoid problems from the parts of the response pattern outside the object, intensity thresholding (applied separately to the original image) was used to set the output to zero where it indicates no insect could be present (Fig. 21.3(c)).

21.3.4 Experimental Results

In laboratory tests, the vectorial operator was applied to 60 grain images containing a total of 150 insects. The output of the basic enhancement operator

```
 2  2  2  2  2  2  2      2  2  3  2  1  0  0      4  2  2  0  0 -1 -2
 0  0  0  0  0  0  0      2  2  0  0  0 -1 -1      2  2  0  0 -1 -2 -1
-1 -1 -1 -1 -1 -1 -1      0  0  0 -2 -2 -2 -1      2  0  0 -1 -3 -1  0
-2 -2 -2 -2 -2 -2 -2      0 -2 -2 -2 -2 -2  0      0  0 -1 -2 -1  0  0
-1 -1 -1 -1 -1 -1 -1     -1 -2 -2 -2  0  0  0      0 -1 -3 -1  0  0  2
 0  0  0  0  0  0  0     -1 -1  0  0  0  2  2     -1 -2 -1  0  0  2  2
 2  2  2  2  2  2  2      0  0  1  2  3  2  2     -2 -1  0  0  2  2  4
```

FIGURE 21.6

7×7 template matching masks. There are 8 masks in all, orientated at multiples of 22.5° relative to the *x*-axis: only three are shown as the others are straightforwardly generated by 90° rotation and reflection (symmetry) operations.

(described in Section 21.3.1) was thresholded to perform the actual detection and then passed to an area discrimination procedure: objects having fewer than 6 pixels were eliminated, as these normally correspond to dark shadows or discolorations on the grains.

The detection threshold was adjusted to minimize the total number of errors. Minor failures were typically due to two insects in contact being interpreted as one, insects being masked by lying along the edges of grains, and the dark edge of a grain appearing similar to an insect.

While adjustment of the detection threshold was found to eliminate either the false positive or two of the false negatives (but not the one in which two insects were in contact), the ultimate reason why the total number of errors could not be reduced further lay in the limited sensitivity of the vectorial operator masks, which contain relatively few coefficients. This was verified by carrying out a test with a set of conventional template matching (TM) masks (Fig. 21.6), which led to a single false negative (due to the two insects being in contact) and no false positives—and a much increased computational load (Davies et al., 2003a). It was accepted that the case of two touching insects would remain a problem that could only be eliminated by introducing a greater level of image understanding—an aspect beyond the scope of this work (although this particular problem could easily be eliminated by noting the excessive length and/or area of this one object).

Next, it was desired to increase the interpretation accuracy of the vectorial operator in order to limit computation, and tests were made of a two-stage system, aimed at combining the speed of the vectorial operator with the accuracy of TM. Accordingly, the vectorial operator was used to create regions of interest for the TM operator. To obtain optimum performance, it was found necessary to decrease the vectorial operator threshold level to reduce the number of false negatives, leaving a slightly increased number of false positives that would subsequently be eliminated by the TM operator. The final result was an error rate exactly matching that of the TM operator, but a considerably reduced overall execution time composed of the adjusted vectorial operator execution time plus the TM execution time for the set of regions of interest. Further speed improvements, and no loss in

accuracy, resulted from employing a dilation operation within the vectorial operator to recombine fragmented insects before applying the final TM operator: in this case, the vectorial operator was used without reducing its threshold value. At this stage, the speedup factor on the original TM operator was around seven (Davies et al., 2003a).

Finally, intensity thresholding was used as a preliminary skimming operation on the vectorial operator—where it produced a speedup by a factor of about 20: this brought the overall speedup factor on the original TM algorithm into the range 50−100, being close to 100 where insects are rare. Perhaps more important, the combined algorithm was in line with the target to inspect 3 kg of grain for a variety of contaminants within 3 min without additional fast hardware.

21.4 CASE STUDY: HIGH-SPEED GRAIN LOCATION

It has already been mentioned several times in this book that object location often requires considerable computation, as it involves unconstrained search over the image data: as a result the whole process of automated inspection can be slowed down significantly, and this can be of crucial importance in the design of real-time inspection systems. Indeed, if the scrutiny of particular types of object requires quite simple dimensional measurements to be made, the object location routine can be the bottleneck. This case study is concerned with the high-speed location of objects in 2D images, a topic on which relatively little systematic work has been carried out—at least on the software side—although many studies have been made on the design of hardware accelerators for image processing. As hardware accelerators represent the more expensive option, this case study concentrates on software solutions to the problem, and then specializes it to the case of cereal grains in images.

21.4.1 Extending an Earlier Sampling Approach

Following on from the earlier sampling approach of Section 12.5, in which circular objects were rapidly located by bisecting a limited number of horizontal and vertical chords, a new approach was tried. This method aimed at even greater speed by taking a minimum number of sampling points in the image rather than by scanning along whole lines (Davies, 1998b).

Suppose that we are looking for an object such as that shown in Fig. 21.7(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

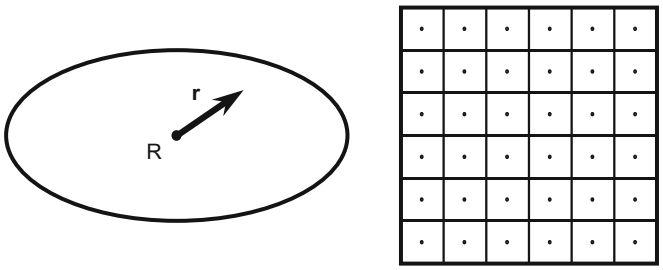


FIGURE 21.7

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.

shape as the original object, but rotated through 180°—because of the minus sign in front of $\mathbf{r}_i$. Furthermore, the fact that reference point positions 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 (Davies, 1997d).

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 Fig. 21.7(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. Thus, 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 Fig. 21.8(a). 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, although it will still be substantial.

21.4.2 Application to Grain Inspection

A prime application for this technique is that of fast location of grains on a conveyor in order to scrutinize them for damage, varietal purity, sprouting, molds, etc. Under these circumstances it is best to examine each grain in isolation: specifically, touching or overlapping grains would be more difficult to cope with.

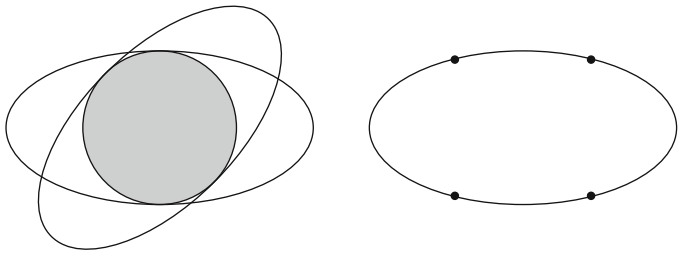


FIGURE 21.8

Geometry for location of ellipses by sampling. (a) Ellipse in two orientations and maximal rotationally invariant subset (shaded). (b) Horizontal ellipse and geometry showing size relative to largest permitted spacing of sampling points.

Source: © EURASIP 1998

Thus, the grains would need to be spread out with at most 25 grains being visible in any 256×256 image. With so much free image space 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.

Wheat grains are well approximated by ellipses in which the ratio of semi-major (a) to semi-minor (b) axes is almost exactly 2. The deviation is normally less than 20%, although there may also be quite large apparent differences between the intensity patterns for different grains. Hence this model was used as an algorithm optimization target. First, the (nonideal) $L \times L$ 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 $L = (4/\sqrt{5})b$ that could be used if the grains were constrained to lie parallel to the image x-axis—see Fig. 21.8(b) (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 21.4.1). 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 centers of the objects. This also applies if the main goal is to locate the objects for inspection and scrutiny. In the present 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 centers.

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

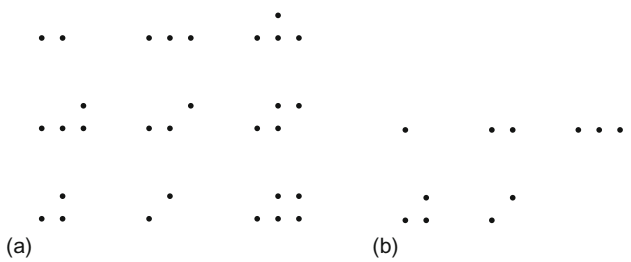


FIGURE 21.9

Possible arrangements of positive sampling points for ellipse, (a) with $L = \sqrt{2}b$ and (b) with $L = (4/\sqrt{5})b$.

Source: © EURASIP 1998

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, although it could in that case also overlap just one laterally placed sample point. In an intermediate direction (e.g., at an angle of around arctan 0.5 to the image x-axis), the ellipse could overlap up to five 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 Fig. 21.9(a).

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 Fig. 21.8(b). This leads to the possible arrangements of positive sampling points shown in Fig. 21.9(b)—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 the other crucial aspect of inspection, that of object scrutiny and measurement, in that relatively few pixels have to be examined in detail, requiring relatively little computation. 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

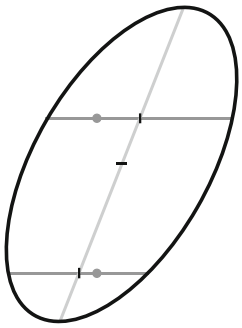


FIGURE 21.10

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 center of the ellipse.

Source: © EURASIP 1998

positive samples into fuller regions of interest and then to 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 midpoint of chord: see Chapter 12) to determine one line on which the center of an ellipse must lie. Repeat this for the all positive samples, and obtain all possible lines on which ellipse centers must lie. Finally, deduce the possible ellipse center locations, and check each of them in detail in case some correspond to false alarms arising from objects that 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).

In fact, a significantly faster approach called the *triple bisection* algorithm was developed (Davies, 1998). 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 midpoints of these bisectors (Fig. 21.10). (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.) The triple bisection algorithm 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 many grain images. The result of applying this technique to an image containing mostly well separated grains is shown in Figure 21.11: this illustrates that the whole procedure for locating grains by modeling them as ellipses and searching for them by sampling and chord bisection

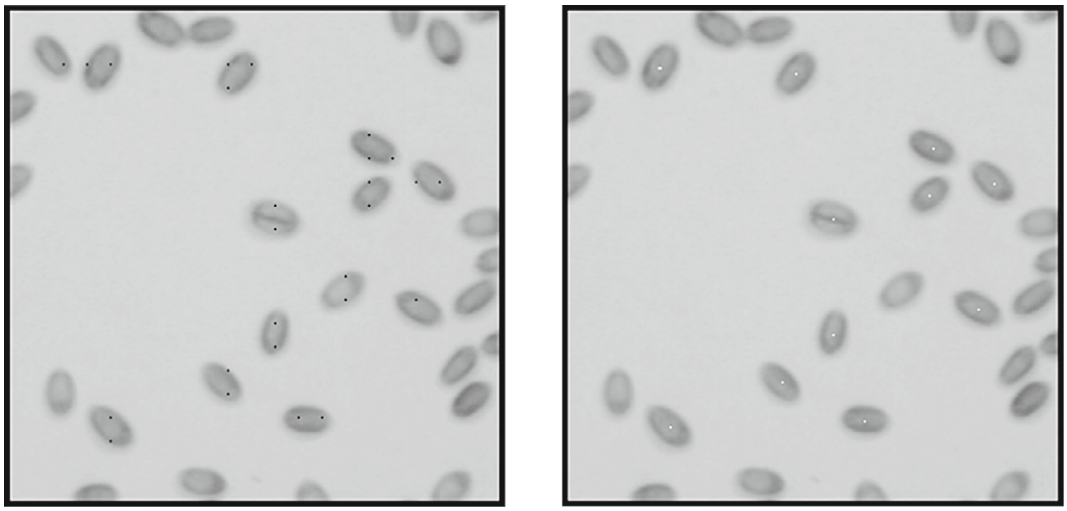


FIGURE 21.11

Image showing grain location using the sampling approach. (a) Sampling points. (b) Final center locations.

Source: © EURASIP 1998

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 remain straight lines, parallel lines remain parallel lines, chords remain chords, and midpoints remain midpoints. Hence, choosing the right orthogonal projection to transform the circle into a correctly orientated ellipse of appropriate eccentricity, the midpoints and center location shown in the diagram of Figure 21.10 must be validly marked. This proves the algorithm. (For a rigorous algebraic proof, see Davies, 1999b.)

21.4.3 Summary

This case study has studied sampling strategies for the rapid location of objects in digital images. Motivated by the success of an earlier line-based sampling strategy (Davies, 1987f), 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. An important feature of the approach is that detection speed is *improved* for larger objects, although exact location involves some additional effort. In the case of ellipses, the latter process is considerably aided by the triple bisection algorithm.

The method has been applied successfully to the location of well separated cereal grains, which can be modeled as ellipses with 2:1 aspect ratio, ready for scrutiny to assess damage, varietal purity or other relevant parameters.

In a more recent development (Davies, 2007c, 2008a), the line- and point-based sampling approaches have been analyzed to determine how much can be learnt by making the particular sampling tests. It turns out that the greatest gain in certainty (amount that can be learnt) is obtained by employing a logarithmic estimate of information and seeking a maximum of the entropy function:

$$E = -[P\log P + (1-P)\log(1-P)] \tag{21.9}$$

where P is the chance of hitting an object at any point. The result of this theory is that a point-based solution will be optimal when object positions are totally unknown, whereas when their positions are already fairly well known—and *a fortiori* when knowledge of their positions is being refined—line-based solutions will be optimal.

21.5 OPTIMIZING THE OUTPUT FOR SETS OF DIRECTIONAL TEMPLATE MASKS

Several earlier sections of this chapter highlighted the value of low-level operations based on template masks, but did not enter into all the details of their design. In particular, they did not tackle the problem of how to calculate the signal from a set of n masks of various orientations—a situation exemplified by the case of corner detection, for which eight masks are frequently used in a 3×3 window (see Section 21.2). The standard approach is to take the maximum of the n responses, although clearly this represents a lower bound on the signal magnitude λ and gives only a crude indication of the orientation of the feature. Applying the geometry shown in Fig. 21.12, the true value clearly has components:

$$c_1 = \lambda\cos\alpha \tag{21.10}$$

$$c_2 = \lambda\cos\beta \tag{21.11}$$

Finding the ratio between the components yields the first important result:

$$\begin{aligned}\frac{c_2}{c_1} = \frac{\cos\beta}{\cos\alpha} &= \cos(\gamma-\alpha)\sec\alpha \\ &= \cos\gamma + \sin\gamma\tan\alpha\end{aligned} \tag{21.12}$$

and leads to a formula for tan α:

$$\tan\alpha = \frac{(c_2/c_1 - \cos\gamma)}{\sin\gamma} \tag{21.13}$$

Hence:

$$\alpha = \arctan\left[\left(\frac{c_2}{c_1}\right)\operatorname{cosec}\gamma - \cot\gamma\right] \tag{21.14}$$

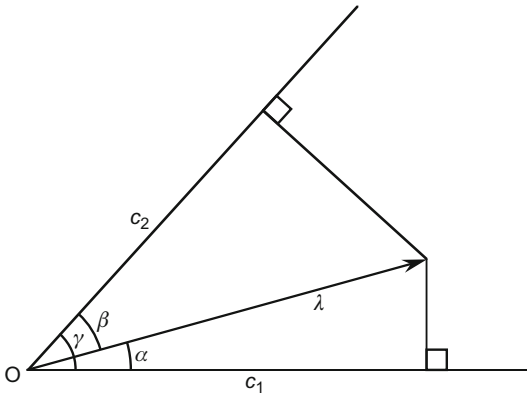


FIGURE 21.12

Geometry for vector calculation.

Source: © IEE 2000

Next, we find from Eq. (21.10):

$$\lambda = c_1 \sec \alpha \tag{21.15}$$

and elimination of α gives:

$$\lambda = c_1 \left[1 + \left(\frac{c_2/c_1 - \cos\gamma}{\sin\gamma} \right)^2 \right]^{1/2} \tag{21.16}$$

On expansion, we finally obtain the required symmetric formula:

$$\lambda = (c_1^2 + c_2^2 - 2c_1 c_2 \cos\gamma)^{1/2} \operatorname{cosec} \gamma \tag{21.17}$$

The closeness to the form of the cosine rule is striking and can be explained by simple geometry (Davies, 2000b).

Interestingly, Eqs. (21.17) and (21.14) are generalizations of the standard vector results for edge detection:

$$\lambda = (c_1^2 + c_2^2)^{1/2} \tag{21.18}$$

$$\alpha = \arctan\left(\frac{c_2}{c_1}\right) \tag{21.19}$$

which apply when $\gamma = 90^\circ$. Specifically, the results apply for oblique axes where the relative orientation is γ.

21.5.1 Application of the Formulae

To proceed further, we define features that have $2\pi/m$ rotation invariance as m-vectors. It follows that edges are 1-vectors, and line segments (following the

discussion in Section 21.3.1) are 2-vectors. In addition, it would appear that m-vectors with $m > 2$ are rarely important in practice.

So far we have established that the formulae given in Section 21.5 apply for 1-vectors such as edges. However, to apply the results to 2-vectors such as line segments, the 1: 2 relation between the angles in real and quadrature manifolds has to be acknowledged. Thus, the proper formulae are obtained by replacing α, β, γ, respectively by 2α, 2β, 2γ in the earlier equations. This leads to the components:

$$c_1 = \lambda \cos 2\alpha \tag{21.20}$$

$$c_2 = \lambda \cos 2\beta \tag{21.21}$$

with $2\gamma = 90^\circ$ replacing $\gamma = 45^\circ$ in the formula giving the final orientation:

$$\begin{aligned} 2\alpha &= \arctan\left[\left(\frac{\cos 2\beta}{\cos 2\alpha}\right)\operatorname{cosec} 90^\circ - \cot 90^\circ\right] \\ &= \arctan\left[\left(\frac{\cos 2\beta}{\cos 2\alpha}\right)\right] \end{aligned} \tag{21.22}$$

Hence:

$$\alpha = \frac{1}{2}\arctan\left(\frac{c_2}{c_1}\right) \tag{21.23}$$

as found in Section 21.3.1.

Finally, we apply the methods described above to obtain interpolation formulae for sets of 8 masks for both 1-vectors and 2-vectors. For 1-vectors (such as edges), 8-mask sets have $\gamma = 45^\circ$, and Eqs. (21.17) and (21.14) take the form:

$$\lambda = \sqrt{2}(c_1^2 + c_2^2 - \sqrt{2}c_1c_2)^{1/2} \tag{21.24}$$

$$\alpha = \arctan\left[\sqrt{2}\left(\frac{c_2}{c_1}\right) - 1\right] \tag{21.25}$$

For 2-vectors (such as line segments), 8-mask sets have $\gamma = 22.5^\circ$, so $2\gamma = 45^\circ$, and the relevant equations are:

$$\lambda = \sqrt{2}(c_1^2 + c_2^2 - \sqrt{2}c_1c_2)^{1/2} \tag{21.26}$$

$$\alpha = \frac{1}{2}\arctan\left[\sqrt{2}\left(\frac{c_2}{c_1}\right) - 1\right] \tag{21.27}$$

21.5.2 Discussion

This section has examined how the responses to directional template matching masks can be interpolated to give optimum signals. It has arrived at formulae that

cover the cases of 1-vector and 2-vector features, typified respectively by edge and line segments. While it appears that there are few likely instances of m-vectors with $m > 2$, features can easily be tested for rotation invariance, and it turns out that corners and most other features should be classed as 1-vectors. An exception to this is a symmetrical S-shape, which should be classed as a 2-vector. Support for classifying corners as 1-vectors arises as they form a subset of the edges in an image. These considerations mean that the solutions arrived at for 1-vectors and 2-vectors will solve all the foreseeable problems of signal interpolation between responses for masks of different orientations. Thus the main aim of this section is accomplished, to find a useful nonarbitrary solution to determining how to calculate combined responses for n-mask operators.

Oddly, these useful results (Davies, 2000b) do not seem to have been reported in the earlier literature of the subject.

21.6 CONCLUDING REMARKS

This chapter has covered three main case studies relating to cereal grain inspection. This topic is quite specialized: not only is it just one aspect of automated visual inspection, but also it covers just one sector of food processing. Yet its study has taken these topics to the limits of the capabilities of a number of algorithms and some useful theory has been developed and applied: this has made the chapter considerably more generic than might *a priori* have been thought. Indeed, this is a major reason for the rather strange inclusion of a section on template matching using sets of directional masks later on in the chapter: it happened to be needed to take the subject onward. This also emphasizes that there is often the need for more theory than is actually available in the literature. Another aspect of the situation is that there is a tendency for inspection and other algorithms to be developed *ad hoc*, whereas there is a need for solid theory to underpin this sort of work and above all to ensure that any techniques that are used are effective, robust, and close to optimal.

Food inspection is an important aspect of automated visual inspection. This chapter has shown how bulk cereals may be scrutinized for insects, rodent droppings, and other contaminants. The application is subject to highly challenging speed constraints and special sampling techniques had to be developed to attain the ultimate speeds of object location.

21.7 BIBLIOGRAPHICAL AND HISTORICAL NOTES

Food inspection is beset with problems of variability, the expression “Like as two peas in a pod” giving a totally erroneous indication of the situation. The author summarized the position in his book (Davies, 2000a) and later reviews (Davies,

2001a, 2003b, 2009). Graves and Batchelor (2003) edited a volume that collects much further data on the problems of variability.

The work on insect detection described in this chapter concentrates on the linear feature detector approach, which was developed in a series of papers already cited but founded on theory described in Davies (1997e, 1999e) and Davies et al. (2002, 2003a, 2003b). (For other relevant work on linear feature detection but not motivated by insect detection, see Chauduri et al. (1989), Spann et al. (1989), Koller et al. (1995), and Jang and Hong (1998).) Zayas and Flinn (1998) offer an alternative strategy based on discriminant analysis, but this targets the lesser grain borer beetle, which has a different shape, and the data and methods are not really comparable. That different methods are needed to detect different types and shapes of insect is obvious, and indeed, the author's own work shows that morphological methods are more useful for detecting large insects, as well as rodent droppings and ergot (Davies et al., 2003b; Ridgway et al., 2002). In fact, the issue is partly one of sensitivity of detection for the smaller insects *versus* functionality for the larger ones, coupled with speed of processing—optimization issues typical of those discussed in Chapter 27.

Work in this area also covers NIR detection of insect larvae growing within wheat kernels (Davies et al., 2003c), although the methodology is totally different, being centered on the location of bright (at NIR wavelengths) patches on the surfaces of the grains.

The fast processing issue arises again in respect of the inspection of wheat grains, not least because the 30 tonnes of wheat that constitute a typical lorry load contain some 6000 million grains, and the turnaround time for each lorry can be as little as 3 min. The theory for fast processing by image sampling and the subsequent rapid centering of elliptical shapes is developed in Davies (1997d, 1999b, 1999d, 2001b): see Davies (2007c, 2008a) for an integrated analysis in which entropy is used to optimize sampling when object locations vary from totally unknown to approximately known.

Sensitive feature matching is another aspect of the work described in this chapter. Relevant theory was developed by Davies (2000b), but there are other issues such as the "equal area rule" for designing template masks (Davies, 1999a) and the effect of foreground and background occlusion on feature matching (Davies, 1999c): these theories should all have been developed far earlier in the history of the subject and merely serve to show how little is still known about the basic design rules for image processing and analysis. A summary of much of this work appears in Davies (2000d).

21.7.1 More Recent Developments

Recent work in the cereals and grain industry has included investigations aimed at separating touching grain kernels (Zhang et al., 2005), classifying cereal grains (Choudhary et al., 2008), differentiating wheat classes (Mahesh et al., 2008), detecting sprouted wheat grains (Neethirajan et al., 2007), sorting grains by color

(Bayram and Öner, 2006), detecting the creases in wheat grains (Sun et al., 2007), and detecting insects inside wheat kernels (Manickavasagan et al., 2008). The different approaches have involved near infrared, thermal imaging, X-ray, and hyperspectral modalities. While the methods have employed a number of relatively standard image analysis techniques—ellipse fitting, wavelets, morphology, color analysis, and stereo vision—the work is far from trivial, because the various varieties of grain and their defects and contaminants involve high levels of variation: this means that getting the techniques to work well is anything but straightforward.