real imagery, via the k-nearest neighbour rule, the results suggested that co-occurrence offered the best overall performance. Wavelets (Porter and Canagarajah, 1996), Gabor wavelets and Gaussian Markov random fields have been compared (on a limited subset of the Brodatz database) to show that the wavelet-based approach had the best overall classification performance (in noise as well), together with the smallest computational demand.

8.4.2 Other classification approaches

Classification is the process by which we attribute a class label to a set of measurements. Essentially, this is the heart of pattern recognition: intuitively, there must be many approaches. These include *statistical* and *structural* approaches; a review can be found in Shalkoff (1992) and a more modern view in Cherkassky and Mulier (1998). One major approach is to use a *neural network*, which is a common alternative to using a classification rule. Essentially, modern approaches centre around using *multilayer perceptrons* with *artificial neural networks* in which the computing elements aim to mimic properties of neurons in the human brain. These networks require *training*, typically by error back-propagation, aimed to minimize classification error on the training data. At this point, the network should have learnt how to recognize the *test* data (they aim to learn its structure): the output of a neural network can be arranged to be class labels. Approaches using neural nets (Muhamad and Deravi, 1994) show how texture metrics can be used with neural nets as classifiers, while another uses cascaded neural nets for texture extraction (Shang and Brown, 1994). Neural networks lie within a research field that has shown immense growth in the past two decades. Further details may be found in Michie et al. (1994) and Bishop (1996; often a student favourite), and information more targeted at vision in Zhou and Chellappa (1992). *Support vector machines* (SVMs) (Vapnik, 1995) are one of the more popular approaches to data modelling and classification, more recently subsumed within *kernel methods* (Shawe-Taylor and Cristianini, 2004). Their advantages include excellent *generalization* capability, which concerns the ability to classify correctly samples that are not within feature space used for training. SVMs have found application in texture classification (Kim et al., 2002). Recently, interest in biometrics has focused on combining different classifiers, such as face and speech, and there are promising new approaches to accommodate this (Kittler, 1998; Kittler et al., 1998).

There are also methods aimed to improve classification capability by pruning the data to remove that which does not contribute to the classification decision. Guided ways that investigate the potency of measures for analysis are known as *feature (subset) selection*. *Principal components analysis* (Appendix 4) can reduce dimensionality, orthogonalize and remove redundant data. There is also *linear discriminant analysis* (also called *canonical analysis*) to improve class separability, while concurrently reducing cluster size (it is formulated concurrently to minimize the within-class distance and maximize the between-class distance). There are also algorithms aimed at choosing a reduced set of features for classification: feature selection for improved discriminatory ability; a comparison can be found in Jain and Zongker (1997). Alternatively, the basis functionals can be chosen in such a way as to improve classification capability.

8.5 Segmentation

To *segment* an image according to its texture, we can measure the texture in a chosen region and then classify it. This is equivalent to *template convolution*, but where the result applied to pixels is the class to which they belong, as opposed to the usual result of template convolution.

Here, we shall use a 7×7 template size: the texture measures will be derived from the 49 points within the template. First, though, we need data from which we can make a classification decision, the training data. This depends on a chosen application. Here, we shall consider the problem of segmenting the eye image into regions of *hair* and *skin*.

This is a two-class problem for which we need samples of each class, samples of skin and hair. We will take samples of each of the two classes; in this way, the classification decision is as illustrated in Figure 8.5. The texture measures are the energy, entropy and inertia of the co-occurrence matrix of the 7×7 region, so the feature space is three-dimensional. The training data is derived from regions of hair and from regions of skin, as shown in Figure 8.6(a) and (b), respectively. The first half of this data is the samples of hair, the other half is samples of the skin, as required for the k-nearest neighbour classifier of Code 8.5.

We can then segment the image by classifying each pixel according to the description obtained from its 7×7 region. Clearly, the training samples of each class should be classified correctly. The result is shown in Figure 8.7(a). Here, the top left corner is first (correctly) classified as hair, and the top row of the image is classified as hair until the skin commences (note that

Figure 8.6 Training regions for classification

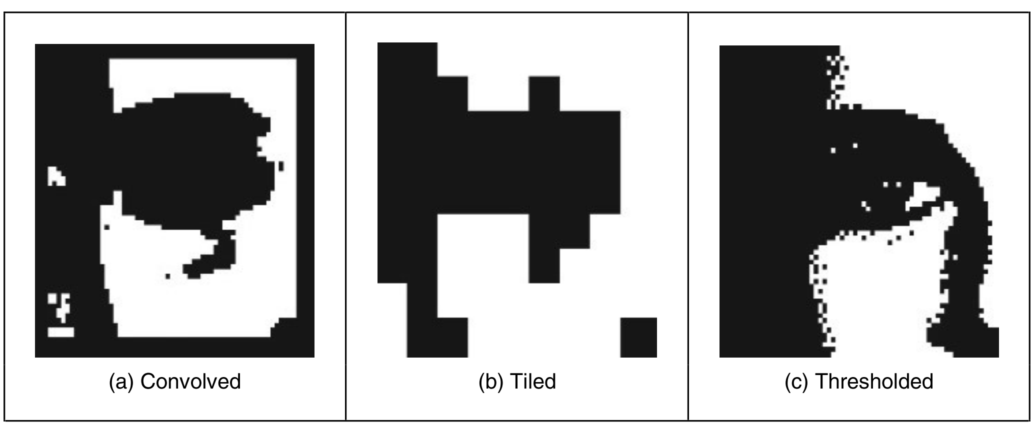

Figure 8.7 Segmenting the eye image into two classes

the border inherent in template convolution reappears). In fact, much of the image appears to be classified as expected. The eye region is classified as hair, but this is a somewhat arbitrary decision; it is simply that hair is the closest texture feature. Some of the darker regions of skin are classified as hair, perhaps the result of training on regions of brighter skin.

This is a computationally demanding process. An alternative approach is simply to classify regions as opposed to pixels. This is the tiled approach, with the result shown in Figure 8.7(b). The resolution is very poor: the image has effectively been reduced to a set of 7×7 regions, but it is much *faster*, requiring only 2% of the computation of the convolution approach.

A comparison with the result achieved by uniform thresholding is given, for comparison, in Figure 8.7(c). This is equivalent to pixel segmentation by brightness alone. There are no regions where the hair and skin are mixed and in some ways the result appears superior. This is in part due to the simplicity in implementation of texture segmentation. However, the result of thresholding depends on *illumination* level and on appropriate choice of the threshold value. The texture segmentation method is completely *automatic* and the measures are known to have *invariance* properties to illumination, as well as other factors. In addition, in uniform thresholding there is no extension possible to separate *more* classes (except perhaps to threshold at differing brightness levels).

8.6 Further reading

There is much further reading in the area of texture description, classification and segmentation, as evidenced by the volume of published work in this area. The best place to start is Maria Petrou's book (Petrou and Sevilla, 2006) (the same author as in edge detection). There is one fairly comprehensive, but dated, survey (Reed and du Buf, 1993). An updated review has a wide bibliography (Tuceryan and Jain, 1998). Another (Zhang and Tan, 2002) offers a review of the approaches that are invariant to rotation, translation, and affine or projective transforms, but texture is a large field of work to survey with many applications. Even though it is a large body of work, it is still only a subset of the field of pattern recognition. In fact, reviews of pattern recognition give many pointers to this fascinating and extensive field (e.g. Jain et al., 2000). In this text, the general paradigm is to extract features that describe the target and then to classify them for purposes of recognition. In vision-based systems such approaches are used in *biometrics*: ways of recognizing a person's identity by some innate human properties. The biometrics of major recent interest are *signatures, speech, irises* and *faces*, although there is work in other areas including hand geometry (as used in US immigration) and gait. The first text on biometrics is not very old (Jain et al., 1999) and surveys all major biometric approaches. (It has just been updated.) There is much interest in automatic target recognition in both military and commercial applications. This translates to medical studies, where the interest is in either diagnosis or therapy. Here, researchers seek to be able to identify and recognize normal or abnormal features within one of the many medical imaging modalities, for surgical purposes. This is the world of image processing and computer vision. But all these operations depend on *feature extraction*, which is why this text has concentrated on these basic methods, for no practical vision-based system yet exists without them. We finish here; we hope you enjoyed the book and will find it useful in your career or study. Certainly have a look at our website, http://www.ecs.soton.ac.uk/~msn/book/, as you will find more material there. Don't hesitate to send us your comments or any suggestions. À bientôt!