# Chapter 7

# Conclusions and future work

## 7.1 Summary and closing discussion

In summary, this thesis examines several image decomposition problems using the framework of the UpWrite. As outlined in Chapter 1, the principal objectives and desirable properties are to produce a mechanism that is generic, automatically extracts structure, performs classification and recognition, allows image reconstruction, is hierarchical and syntactic, and is computationally viable. These characteristics can now be discussed with respect to the detection of anomalies, segmentation of newspaper images and analysis of outdoor scenes covered in the previous three chapters.

The mechanisms used are clearly generic. The same feature extraction process is applied to each problem with appropriate modifications to the choice of window measurements. Similar modelling techniques, making use of local, global and region based geometric moments, are used throughout. This is most clearly apparent in the partial DownWrite of painted steel images and colour scenes. Furthermore, there are few free parameters, and those that exist relate to general image properties such as image resolution.

In terms of automation, following parameter choice the processes used throughout are unsupervised, requiring only labelling and training to perform classification. Feature vectors tend to be normalised based on the attributes of the image being analysed or from the image class. For example, the $\eta - \mathbf{d}$ parameterisation, used in Chapter 5, normalises feature vectors with respect to the value of the current image assuming only that the image contains some text regions. The basis for the entire anomaly detection algorithm is another clear example.

The ability to perform classification is demonstrated in each of the three restricted problems. The results for anomaly detection are very good and the classification of Corel images is very encouraging given the scope of the problem and the simplicity of the final feature vectors used.

Although the partial DownWrite performs image reconstruction, the accumulated limitations of the representation and reconstruction schemes make it inappropriate for image compression. However, as intended, the reconstructed image provides valuable insight into the effectiveness of the underlying algorithms.

The algorithms used are clearly hierarchical and syntactic, combining low level primitives into high level structure. In Chapter 4 the feature vectors are combined to form the background and anomalous regions. The anomaly is in turn examined for structure. In Chapter 6 similar feature vectors form multiple image regions which are in turn amalgamated into more complete objects.

The final desirable property is computational viability. The inherent nature of the adopted framework caters well for this. The early local analysis and amalgamation of many thousand pixels into relatively few objects is computationally attractive. The final representations are typically compact and concise. For example, although the segmentation of a fairly large colour image takes approximately 30 seconds, the amalgamation based on mutual prediction takes less than a second and the comparison between regions is in the order of microseconds. Note that the first two of these three operations need only be done once. Furthermore, the recent speed increases of cheap desktop machines compared to traditionally more expensive ones, as apparent from Section 6.5.6, provides an optimistic prospect for algorithms within this framework. Consequently, a cheap real-time implementation of the anomaly detection and classification algorithms should soon be feasible.

In view of the above objectives and desirable properties, the UpWrite framework has successfully been applied to three very different restricted problems. An extremely brief summary of each follows. Refer to individual chapter summaries for more detail (Sections 4.10, 5.10.2 and 6.10).

For anomaly detection restricted to images with homogeneous backgrounds, a novel generic feature extraction process is applied using a simple average intensity window measurement. The entire feature space is modelled using a multivariate Gaussian distribution, with the background forming a dense cluster. Anomalous regions are identified as those which depart from this model. They are examined using a local Karhunen-Loeve Transformation to extract topological structure in the feature space. Using the structure obtained from images of three defect classes, classification results yield 49 correct, 8 unsure and 1 misclassification, or alternatively 51 correct and 7 misclassifications. Anomaly removal is presented to gain an insight into the background modelling.

The segmentation of newspaper images uses the same feature extraction method, but using the standard deviation of intensity as a window measurement. The feature vectors are transformed into the $\eta - \mathbf{d}$ parameterisation for Gaussian mixture modelling and subsequent initial region identification. Spatial neighbourhood information provides the final classifications after which post-processing forms rectangular regions corresponding to pictures and text. A comparison with an optimal Gabor filter technique suggests this method is superior, more general, is far more efficient and possesses numerous invariances. It is also simple, effective and easily adaptable to image resolution.

The more general colour segmentation also uses the same feature extraction method with measurements for intensity, standard deviation of intensity and colour. A variation of the hue-saturation plane is used for colour with all measurements normalised. Through identifying regions with little local change, a region growing type segmentation algorithm is applied. Regions are grown with respect to an incremental model with a dynamic threshold to avoid region leaking. Amalgamation of regions into more complete objects, a partial DownWrite and classification are also presented. Classification of 1000 images from 11 categories yields 55.1% accuracy.

Overall, the adopted framework has successfully been applied to the detection and classification of anomalies and the segmentation of mixed-mode images. These two restricted and relatively well defined problems both suit and demonstrate the effectiveness of the approach. However, the segmentation and classification of coloured natural images stretches its bounds. This third problem is not well defined and in general is very ambitious. The ability to sufficiently represent and compare shape as well as the ability to form a general description which combines multiple objects (as a point in Euclidean space), outlines the infancy of the UpWrite approach.

In all, the UpWrite provides a good framework which works reasonably well over numerous different problems. In combination with the associated algorithms proposed in this thesis, it provides numerous tools for analysing images which can be used to provide a quick insight into many different specific problems. It is however inherently generic and hence may not be adequate for specific problems requiring expert knowledge, although specific analysis may obviously be applied as required.

## 7.2 Original contribution

The material presented in Chapters 4, 5 and 6 represents original work. As outlined in Section 1.1 and 2.8, the UpWrite has previously been used for analysing the structure of English text and strokes in binary images, both which deal with ordered data. In this thesis, the application of the framework deals with low level primitives which cannot be ordered since the problems are two dimensional in nature. Hence, this work also extends the UpWrite to a new class of problems.

More specifically, the original contribution includes the following.

- The feature extraction process detailed in Section 4.3. It includes parameters for *blocksize* $\mathbf{k}$, *windowsize* $\mathbf{n}$, *shiftsize* $\mathbf{s}$, shifts $\delta_x$ and $\delta_y$, and local window measurements. This generic low level feature extraction method is adaptable to image resolution ($\mathbf{n}$), neighbourhood complexity ($\mathbf{k}$), localisation ($\mathbf{s}$), memory restrictions ($\delta_x$, $\delta_y$) and appropriate local image characteristics.

- The use of feature vectors to identify anomalies as outlined in Section 4.5. Through modelling the entire image using a standard multivariate Gaussian model, regions which disagree with the uniform background are identified.

- A mechanism for refining unimodal histograms, presented in Section 4.5.4. By recursively modelling the distribution of the Mahalanobis distance of points from a Gaussian model, the distinction between points from the distribution and anomalies can be distinguished more easily.

- The application of the local Karhunen-Loeve Transformation for locating surfaces and other manifolds from a dataset, with the intention of using identified structure to classify images. The method used to approximate local dimension from an arbitrary set of eigenvalues (obtained from a matrix of positive semi-definite form) is also original. These techniques appear in Section 4.7.

- The reconstruction of image pixels through restricting a multivariate Gaussian model using the prediction matrix $\Psi$ derived in Section 4.9. This technique is applied to visually remove anomalies, thus providing an insight into the modelling. It is used similarly in Section 6.9 to reconstruct the local characteristics of colour images.

- The $\eta - \mathbf{d}$ parameterisation outlined in Section 5.4.4. This provides a novel set of measurements to concisely describe a comparable set of measurements containing an arbitrary number of elements, or equivalently a feature vector of an arbitrary dimension (for $\mathbf{N} > 1$).

- The tradeoff between the HSI cylindrical and conical colour systems, as presented in Section 6.3. The resulting colour space is also decorrelated to form measurements similar to *Ohta's colour features* [228] or those obtained by principal components analysis [229]. These are an extension to the material presented in Section 2.6.1 and illustrated in Tables 2.6 and 2.7.

- The region growing segmentation algorithm detailed in Section 6.5. Growing regions based on a multivariate Gaussian describing local characteristics and inhibiting growth as the distribution of the model increases to significantly occupy the bounded feature space is an original idea.

- The use of the prediction matrix $\Psi$ to amalgamate spatially disjoint regions with mutually predicting local characteristics, thus tackling the problem of occlusion. This is outlined in Section 6.7.

## 7.3 Future research

This, the final section of this thesis, briefly outlines the future research tasks. The first two relate to Chapter 4, the next to Chapter 5, the following five to Chapter 6 and the final two are applicable across all three chapters.

- **Manifold modelling**. The problem of identifying and representing manifolds embedded in a higher dimensional space requires thorough examination. Even for artificial data points, produced from embedding a mathematical function at discrete intervals with limited noise, it is non-trivial to identify and amalgamate the structure. For example, consider a discrete dataset representing a solid sphere in $\mathbb{R}^3$ embedded in $\mathbb{R}^4$. Given no noise and a dense enough set, the local dimension, i.e. 3, can easily be ascertained. However, given a combination of a sparse dataset and additive noise, the structure is quickly lost. Thus, a statistical problem exists to determine the relationship between the density of points, quantity of noise, and the dimensions of the space and embedded manifold. Such a relationship would provide the information necessary to ascertain whether it is feasible to attempt to locate manifolds of a given density in a particular dimension. Given that it is, these relationships would also provide an indication of the tolerance to noise and the chance of identifying an embedded manifold with ideal structure. The problem of representing such a structure is even more challenging.

- **The hardware implementation of a real-time anomaly detection and classification system**. Such a scheme would require access to vast amounts of training data, further clarification of adequate classification results and preferably a controlled environment.

- **The use of newspaper segmentation results to read the paper**. Additional research might make use of the segmentation to aid skew analysis and optical character recognition.

- **The examination of alternative local measurements**. The segmentation algorithm presented in Chapter 6 uses simple averages and standard deviations for local window measurements. The investigation of other simple measurements, such as local autocorrelation, as well as more complicated ones, is applicable.

- **The incorporation of mixture modelling into the colour segmentation algorithm**. As indicated in Section 6.5.7, it may be beneficial to model regions containing a mixture of different local regions, or multiple objects which occupy the same image location. The incorporation of this into the segmentation algorithm, or the independent examination of this phenomenon can be investigated.

- **Shape representation**. An appropriate shape representation should allow regions to be appropriately amalgamated, compared to one another, and ideally support clustering of *similar* shapes. This is a significant fundamental problem of image recognition.

- **Analysis of non-homogeneous regions**. The colour segmentation algorithm detects spatially disjoint regions with homogeneous local characteristics. However, much of the information exists at the edges of these regions where relatively anomalous characteristics are present. One possibility is to combine the techniques presented in this thesis with those used for detecting and analysing strokes in binary images (Section 2.8.3) to ascertain the nature of these missing regions. Some will obviously correspond to boundaries between the regions, whereas others will contain important information.

- **User query driven CBIR**. Given the description of disjoint image regions, a user should be able to offer a prototype region and receive images containing similar regions. This requires a user interface, a mechanism to define a prototype, and the flexibility to select desired invariances by changing comparison measures.

- **Multi-scale analysis**. This is applicable to all methods presented in Chapters 4, 5 and 6. Processing images over a range of scales ensures that errors are not produced by artifacts at one particular scale. The ability to identify the same structure at multiple scales provides reassurance that it actually is there.

- **Algorithm success measures**. The algorithms presented do not attempt to determine the strength of the identified structure or to dismiss images containing inappropriate information. Perhaps the worse case is the mixed-mode segmentation algorithm which silently fails when presented with a picture that violates the imposed restrictions. The development of such algorithm measures is left for future work.