Shape Description for Automatically Structuring Graphical Data

Laura Keyes and Adam Winstanley

Dept. of Computer Science, National University of Ireland, Maynooth, Co. Kildare, Ireland
lkeyes@cs.may.ie, adam.winstanley@may.ie

Abstract. This work explores automatic object recognition and semantic capture in vector graphics through shape description. The low-level graphical content of graphical documents, such as a map or architectural drawing, are often captured manually and the encoding of the semantic content seen as an extension of this. The large quantity of new and archived graphical data available on paper makes automatic structuring of such graphical data desirable. Contour shape description techniques, such as Fourier descriptors, moment invariants play an important role in systems for object recognition and representation. However, most work carried out in this area has concentrated on categories of object boundaries representing very specific shapes (for example, a particular type of aircraft). Two classifiers were implemented and proved accurate in their automatic recognition of objects from drawings in different domains. Classical classifier combination techniques were used to improve performance. Further work will employ more complex fusion techniques and it is envisaged they will be used in combination with recognition based on object context using various modelling methods. A demonstration system has been constructed using all these techniques.

1 Introduction

Increased use of graphical information systems (such as GIS, CAD and multimedia systems) has motivated research in developing and applying graphical object recognition. A vast amount of data archived by organisations in the world is in graphical form (for example, diagrams, maps, technical drawings, and architectural plans). For this to be searched, analysed and synthesised automatically, it must be parsed and converted from simple graphics (points, line, symbols, polygons) to semantically rich graphical information ("circuit breaker", "building", spark-plug", "extractor fan"). Much work has been done in computer vision on the identification and classification of objects within images. However, less progress has been made on automating feature extraction and semantic capture in vector graphics. This is partly because the low-level graphical content of graphical documents has often been captured manually (on digitising tables) and the encoding of the semantic content seen as an extension of this. The successful automation of raster-vector conversion plus the large quantity of new and archived graphical data available on paper makes the automation of feature extraction and recognition of graphical data desirable.

Data capture for graphical information systems consists of two parts: the digitisation of the geometry and the addition of attributes indicating the object type

being depicted. Whereas the former can be automated using image processing and similar techniques, the latter is often a manual task. The manual structuring into composite objects and addition of labelling attributes is a labour-intensive, expensive and error prone process. One possible means of automatically structuring graphical data is through shape. Shape is an important part of the semantic content of an object within a graphical information system. Shape description methods used in image processing include Fourier Descriptors (FD) and Moment Invariants (MI).

Traditionally shape description methods are applied to the recognition of very specific shapes (for example, a particular make and model of aircraft). However, there are situations when, due to the nature of the data, we wish to classify previously unseen shapes into general classes representing categories of object. For example, in topographic object classification, a specific building shape may be unique. However, all building shapes have particular properties that distinguish them from objects in other classes. If descriptors can capture these properties, it would allow us to classify new examples. Other applications of general descriptors include the domain analysis of documents (for example as an electrical circuit diagram or a computer flow chart) and the segmentation of documents into different regions.

Previous work has applied this idea to topographic data from large-scale mapping [1]. As part of a project to speed up data capture and structuring of architectural and engineering drawings for a web-based multimedia Operations and Maintenance (O&M) management system, an initial attempt has been made to apply the same concept to this data domain. Standard shape description techniques are applied to object boundaries extracted from drawings represented as vector descriptions. The outputs obtained by the description methods provide a measurement of shape that characterises the object type. We then test the techniques' effectiveness at classifying shapes into general categories using standard data sets.

Different classifier schemes produce different classifications of the same data set. By analyzing these results against a standard data set for which ground-truth values are available, their effectiveness can be analysed and compared. It is usually seen that, although one technique yields the best performance, the set of features miss-classified by the different classifiers do not necessarily overlap. This suggests that different classifier techniques can offer complementary descriptions of the shapes to be classified, and a combination of classifiers would give optimal performance. A fusion methodology based on a Bayesian framework and using *max*, *min*, *sum* and *majority vote* rules [12], is applied to combine the results of the individual classifiers to derive an overall consensus decision. Results are presented here for the shape description task and matching module for identifying architectural features and symbols on drawings.

2   Automatically Structuring Graphical Data

Automating the structure of graphical data requires the recognition and representation of objects that are defined by a set of general shape properties. This involves the classification of a particular shape into a general class of similar object shapes. For example, the human visual system can recognise and identify a given chair from its shape properties, though the shape describing any particular chair may vary considerably. The shape properties of a given feature may vary while semantically they describe the same object. This semantic similarity must be considered when

attempting to classify graphical data. In this work, boundary shape description is investigated and evaluated on the problem of automatically recognising and interpreting graphical data on technical drawings for the development of an operation and maintenance information system for plans within buildings and other facilities.

2.1   Operation and Maintenance Information System

An Operation and Maintenance Information System holds centrally all relevant information pertaining to the operation and maintenance of plant and equipment within buildings and other facilities. This information is presented through a multimedia web interface and consists of drawings, data sheets, operating instructions, parts listings, suppliers, installers, manufacturers and other details of all the service utilities. The information on each component is comprehensively cross-referenced using links between corresponding items in drawings, data sheets, photographs and so on. The system can be implemented for all sizes of installations but comes particularly suited for the infrastructure management of large industrial or service sites. Current use includes a sports complex and large private dwelling.

The Operation and Maintenance Information System allows a user to select an example object (simple or composite) and the software finds similar objects in the same or other drawings. The tool generates data structures that can be used to build multimedia linkages between objects, drawings and related information. The information is accessed through a standard web browser interface including navigation through hot-links and key-word search facilities.

CAD drawings showing the location of utilities and services also act as browser navigational maps. In operation, the system's main use concerns day-to-day operation and maintenance tasks, for example retrieving plant operating and servicing instructions or keeping maintenance records. Systems commissioned so far have been constructed manually through digitising and structuring of this information appropriately. For the system to be economic, it is desirable to automate as much as possible of this compilation process. Automation possibilities include:

- Recognition and labelling objects/components on drawings through
  - shapes of objects within drawings
  - text labels in proximity to objects
  - meanings of standard symbols
  - context of objects, for example a water pump on a water pipe
- Generating links through string matching
- Compilation of databases of information from scanned text/drawings

Once recognised and classified, these objects can be assigned unique identifiers in the system. This allows their inclusion in the search and navigation functions. In this paper we evaluate the classification of objects and components on drawings based on their isolated shape. Other work is in progress to incorporate techniques based on context into the classification process.

2.2   Shape Description for Object Recognition and Representation

The recognition and description of objects plays a central role in automatic shape analysis for computer vision and it is one of the most familiar and fundamental problems in pattern recognition. Common examples are the reading of alphabetic characters in text and the automatic identification of aircraft. The shape description

techniques used include Fourier Descriptors (FD) Moment Invariants (MI) and Scalar Descriptors (SD) (area, elongation, number of corners etc.) []. Most applications using these for shape recognition deal with the classification of such definite shapes. However, this application (structuring components and symbols on technical drawings) is one of a number requiring the recognition and representation of object shapes that are defined by a set of general shape properties. Other examples include extracting buildings from topographic data and architectural drawings. In these cases, the semantic classes are reflected in the general properties of the objects shape as opposed to an exact template match. To identify graphical objects, each of the techniques needs to be extended to deal with general categories of shapes found in graphical documents, for example buildings.

The recognition of objects is largely based on the matching of numerical descriptions of shapes with a database of standard shapes. Fourier Descriptors, Moment Invariants and Scalar Descriptors are well understood when applied to images and can be normalised to describe shapes irrespective of position, scale and orientation [2]. They can also be easily applied to vector graphical shapes. Each technique is computed based on the shape boundary. The shapes can then be described using a small set of descriptor values (typically 7 to 10 real numbers). The results produced are used in the classification process. The recognition and classification is based on matching the descriptors of each shape to standard values representing typical shapes and choosing the closest match.

2.3   Classification

Shape description techniques generally characterise an object's shape as a set of real numbers. Classification of objects based on shape therefore consists of comparing these descriptors. In this work we are using supervised classification through Bayesian statistics [11,12].

Supervised classification involves two stages: a learning stage where criteria and methods are tried on the prototypes and recognition when the trained system is used to classify new data. Bayesian statistics uses the distribution of the values for each descriptor, for each class of object, in determining the probability that a particular object belongs to that class. Given a particular value for a descriptor, the likelihood of that value occurring in the distribution of values for a particular class can be determined. Applying Bayes theorem, the probability of the object belonging to that class is computed. Such a probability can be calculated for each class. The object then belongs to the class for which that descriptor gives the highest probability. The objective is to design classifiers that will classify an object in the most probable of the classes given.

2.4   Combining Classifiers

When setting out to design a shape recognition system the ultimate goal is to achieve the best possible classification performance. Attaining this goal involves the application of suitable classification schemes/techniques to the problem. Traditionally, choosing a classifier scheme as a final solution for the task at hand was based on an analysis of the results produced by each technique. However, although one technique would yield the best performance, the set of features miss-classified by the different classifiers would not necessarily overlap. This suggests that different

classifier techniques can offer complementary descriptions of the shapes to be classified, which leads to the combining of the classifiers for improved performance.

A central problem for the fusion algorithm is how to integrate several classifiers to produce a single final classification. The fusion technique employed here follows a methodology based on *max*, *min*, *sum* and *majority vote* strategies [12] and utilises a decision combination topology with a Bayesian approach. All three shape description methods (Fourier descriptors, scalar descriptors and moment invariants) are used for the decision making by combining their individual results to derive a consensus decision. Using the set of real descriptor values produced by each shape description method, a *likelihood* and *probability* of an input feature belonging to a particular class is calculated. Each strategy obtains a decision by computing the a-posteriori probabilities for each class (for example, the sum rule will compute the sum of the probabilities), taking the resulting maximum value as the combined result.

These fusion algorithms can be used in two configurations. When used to evaluate the effectiveness of these shape recognition techniques using a fully structured training set, it outputs measures of performance of individual classifiers as well as combinations of classifiers using different combination algorithms. Alternatively, when a particular configuration is known to be effective for a particular recognition problem, it can be parameterised to implement this combination and label the shapes according to the derived classifications.

3   Case Study: Shape Recognition in Architectural Drawings

Shape description techniques previously developed and implemented were applied to architectural data. The performance of all techniques is statistically analysed for the automatic recognition and labelling of features on the drawings. Figure 1 describes the architecture of the overall shape recognition system employed.
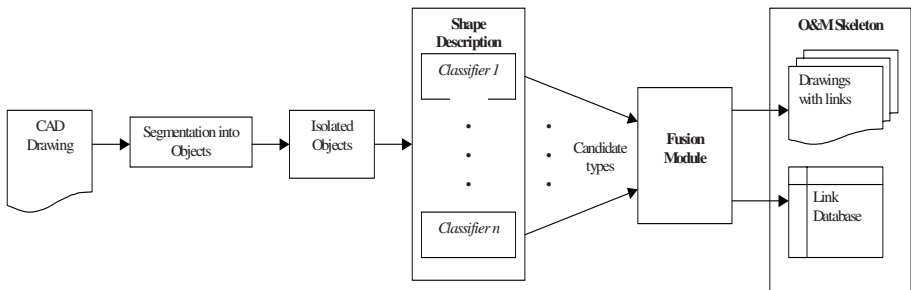

Fig. 1. Graphical Shape Recognition System configuration

3.1   Data Pre-processing

Before implementation of the shape techniques can be carried out there needs to be some pre-processing performed on the graphical data set. The architectural data available for construction of O&M systems is generally in a CAD vector format. Therefore, many of the problems of raster data and vectorisation do not apply. First,

the technical plan is segmented into its component objects for extraction. The data extracted must form separate/minimal isolated closed polygonal shapes and are stored in vector format for further processing. An interpolation method is applied to this data to sample the shape boundary, using a finite number (N) of equidistant samples representing the *x* and *y* co-ordinates of the objects shape. These points are stored in the appropriate format (complex valued data in the case of the FD method and *x-y* co-ordinates for the MI and SD methods) for processing with each shape description technique.

The data is extracted and stored as individual polygon shapes in a format suitable for processing with the shape descriptor techniques, Fourier Descriptors, Moment Invariants and Scalar Descriptors. These shape description techniques are calculated from the object boundary. The output from each shape description method can be used in subsequent stages of the overall system, that is, the component matching and later online database retrieval. Fourier Descriptors and Moment Invariants produce a set of real valued numerical descriptors to describe each object. So, for each component shape in the set we have twenty-one descriptor values, (the sixteen scalars, seven moments and fourteen Fourier descriptors, (FD(2) to FD(16)). FD(0) and FD(1) are redundant due to normalisation (translation, scale rotation) performed for each method.

3.2   Experimental Results

To evaluate the effectiveness of this method, an experiment was carried out on a corpus of technical drawing objects from the GREC 2003 Symbol Recognition Contest: architectural, electrical and symbols (figure 2). The aim of this experiment is not to recognise each individual shape but to classify each set of graphical objects into their respective domain. The sample datasets include an ideal set of all symbols used for training plus several sets containing various types of distortion and degradation. For this application, binary degradation is not a problem, therefore only vectoral distortion including scale and rotation were considered (figure 3).
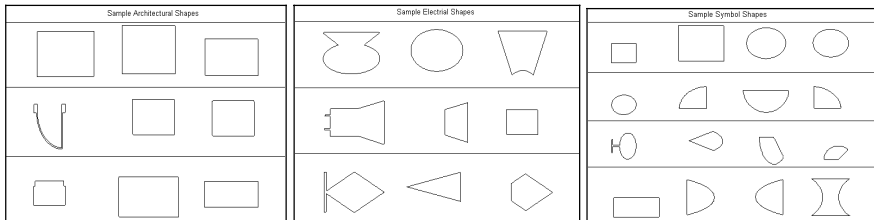

Fig. 2. Sample contour shapes from training sets

From each symbol image, closed contours were generated. Most symbols produce several contours. For each contour, Fourier and moment invariant descriptors were calculated. The ideal set was used to produce descriptor value distributions for each domain. The descriptors obtained from the testing data sets were used to classify the contours as belonging to one of the three classes through a maximum likelihood measure derived from the distributions. In this way, the data sets were used to test the effectiveness of each descriptor individually and combined using the implemented fusion algorithms.
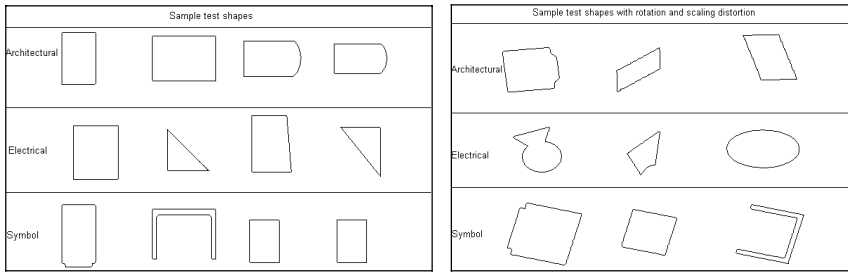
Fig. 3. Sample contour shapes from testing sets

To evaluate the performance of the graphical shape recognition system, standard precision and recall information retrieval metrics are employed. Precision, given here as a percentage, is a measure of selected shapes that the recognition system classified correctly (so-called true-positives as a fraction of all positives). Recall, also given as a percentage, is the measure of the proportion of the target class that the system detected. Tables 1, 2 and 3 show the results obtained for each shape description technique on a set of ideal tests and tests with scaling and rotation distortion, respectively. Each test set consists of Architectural, Electrical and Symbol data to be classified. Overall, the Fourier Descriptor method seemed to perform well when classifying Architectural data but poorly on Symbol data. Conversely, moment invariants performed well for classifying Symbol data but poorly for Architectural data. For Electrical, both techniques mis-classified the data as Symbols.

Table 1. Precision and recall of training set (59 objects, 212 contours)

Fourier Descriptors		Majority %	Max %	Min %	Median %	Sum %	Product %
Architectural	Precision	25.8	100	89.4	30.7	100	28
	Recall	96.0	76.0	84.0	92.0	92.0	92.0
Electrical	Precision	0	0	100	0	0	12.5
	Recall	0	0	7.1	0	0	3.57
Symbol	Precision	37.5	100	21.5	89.3	83.8	84.4
	Recall	15.7	35.8	26.1	18.7	23.1	20.9
Moment Invariants							
Architectural	Precision	0	100	2.5	100	100	11.1
	Recall	0	2	4	2	2	2
Electrical	Precision	0	0	100	0	0	1.3
	Recall	0	0	3.6	0	0	3.6
Symbol	Precision	42.3	100	94.5	99.3	63.8	98.4
	Recall	100	100	94.8	100	100	94.8

Using the ideal test data, results show that as descriptors for classifying Architectural data, Fourier descriptors prove best with 100% precision and 92% recall using max and sum rule combination strategies. Moment invariants proved best with 100% precision and 96% recall using the max rule and 98.1% precision and 99% recall using median rule. For the test set with scaling and rotation distortion, Fourier descriptors showed best results at 100% precision and 80.1% recall using the sum rule for Architectural data. On the other hand, moment invariants scored 100% and 99.8%

Table 2. Precision and recall for ideal test set (50 objects, 180 contours)

Fourier Descriptors		Majority %	Max %	Min %	Median %	Sum %	Product %
Architectural	Precision	29.8	100	89.7	33.1	100	32.2
	Recall	96.0	92.0	84.0	96.0	92.0	92.0
Electrical	Precision	0	0	100	0	0	22.2
	Recall	0	0	7.1	0	0	7.1
Symbol	Precision	38.9	100	20	84.2	80.8	66.7
	Recall	13.7	28.4	25.5	15.7	20.6	19.6
Moment Invariants							
Architectural	Precision	0	100	2.5	16.7	100	12.5
	Recall	0	2.0	4.0	2.0	2.0	2.0
Electrical	Precision	0	0	100	0	0	2.53
	Recall	0.0	0.0	3.6	0.0	0.0	7.1
Symbol	Precision	40.5	100	96	98.1	57.1	98
	Recall	100	96	94.1	99	98	94.1

Table 3. Precision and recall for deformed test set (scaling/rotation) (250 objects, 1147 contours)

Fourier Descriptors		Majority %	Max %	Min %	Median %	Sum %	Product %
Architectural	Precision	35.2	100	76.8	36	100	37
	Recall	81.2	78.6	76.6	81.2	80.1	78.1
Electrical	Precision	0	0	100	0	0	4
	Recall	0	0	4.2	0	0	2. 8
Symbol	Precision	30.9	100	20.7	53.8	83.2	45 .2
	Recall	24.4	27.9	23.9	24.5	25.8	18.6
Moment Invariants							
Architectural	Precision	26.1	100	5.7	14.6	100	18 .6
	Recall	1.5	2.3	8.1	1.5	1.5	4. 1
Electrical	Precision	0	0	100	0	0	0. 4
	Recall	0	0	1.4	0	0	1. 4
Symbol	Precision	37.6	100	91.1	99.8	54.2	98 .6
	Recall	97.7	96	90.9	97.9	96.7	93. 6

precision and 96% and 97.9% recall for the max and median rule, respectively. These results indicate the potential for using Fourier descriptors and moment invariants for the general classification of graphical domain data. Also the results show the potential of further combining each shape classifier to arrive at an improved consensus decision across all domains, as each provides complementary information, most notably on the Architectural and Symbol domains. Fusing the individual descriptors (14 FDs and 7 MIs) improves precision and recall of results for each test set.

4   Conclusions

Two types of contour/boundary shape description methods were investigated and their effectiveness evaluated on the problem of automatically recognising and interpreting graphical data on technical documents where generalized shape properties are

required. A demonstration system has been developed to assess the potential of this approach to automatically structuring graphical data for the development of an online operation and maintenance information system.

Automating the structure of graphical data requires the recognition and representation of objects that are defined by a set of general shape properties. This involves the classification of a particular shape into a general class of similar object shapes; for example, the shape properties representing an office on a building plan may vary while still describing the same object. Such semantic properties must be considered when attempting to classify graphical data.

Each device was classified by the individual and fused descriptors with an accompanying measure of certainty and confidence. Both techniques, Fourier Descriptors and Moment Invariants proved reasonably successful in certain domains but not in all. Some classical Bayesian fusion techniques were implemented to try to optimise recognition by combining classifiers with limited success. It is planned to implement a more sophisticated fusion approach based on neural networks in the future.

Further work currently being carried out is evaluating this approach on more complex, variable and larger sets of graphical data found on building plans. Also envisaged, as a natural extension of this is work, is the combination of recognition through shape with recognition based on object context using various context-modelling methods. The results presented here from the initial system based on shape indicate the potential for this approach.

References

1. Winstanley, A.C.: Structuring Vector Maps using Computer Vision Techniques, Conf. of the Association of Geographic Information, Birmingham, pp. 8.11.1–8.11.2, (1998).
2. Winstanley, A.C. and Keyes, L.: Applying Computer Vision Techniques to Topographic Objects, Int. Archives of Photogrammetry and Remote Sensing, 33 (B3): 480–487, (2000).
3. Keyes L. and Winstanley A.C.: Using Moment Invariants for Classifying Shapes on Large-scale Maps, Computers Environment and Urban Systems, 25, 119–130, (2001).
4. Keyes L. and Winstanley A.C.: Data Fusion for Topographic Object Classification, IEEE/ISPRS Workshop on Remote Sensing and Data Fusion, 8-9, Rome, Nov 2001.
5. Gonzalez, R.C., Wintz, P.: Digital Image Processing, Addison-Wesley, 1977.
6. Granlund, G.H.:   Fourier Pre-processing for Hand Print Character Recognition, IEEE Transactions on Computers, C-21: 195–201, 1972.
7. Wood, S. L.: Fourier Analysis of Object Boundaries From Two Dimensional Digitised Images, ICASS, 1986.
8. Hu, M. K.: Visual Pattern Recognition by Moment Invariants, IEEE Transactions on Information Theory, IT-8: 179–187, 1962.
9. Dudani, S. A., Breeding, K.J., McGhee, R. B., Aircraft Identification by Moment Invariants, IEEE Transactions on Computers, C-26 (1): 39–45, 1977.
10. Chaur-Chin Chen: Improved Moment Invariants for Shape Recognition, Pattern Recognition, 26(5): 683–686, 1993.
11. Costa, L.F., Cesar, R.M., Shape Analysis and Classification: Theory and Practice, CRC Press, 2001.
12. Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., On Combining Classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(3), 1998.