Contour Co-occurrence Matrix – A Novel Statistical Shape Descriptor

Rami Rautkorpi and Jukka Iivarinen

Helsinki University of Technology,
Laboratory of Computer and Information Science,
P.O. Box 5400, FIN-02015 HUT, Finland
{rami.rautkorpi, jukka.iivarinen}@hut.fi

Abstract. In this paper a novel statistical shape feature called the Contour Co-occurrence Matrix (CCM) is proposed for image classification and retrieval. The CCM indicates the joint probability of contour directions in a chain code representation of an object's contour. Comparisons are conducted between different versions of the CCM and several other shape descriptors from e.g. the MPEG-7 standard. Experiments are run with two defect image databases. The results show that the CCM can efficiently represent and classify the difficult, irregular shapes that different defects possess.

1 Introduction

There are lots of different features available that are used in image classification and retrieval. The most common ones are color, texture and shape features [1]. Shape features can be divided into two main categories [2]: syntactical, using structural descriptions suitable for regular shapes such as man-made objects, and statistical, which is more suitable for irregular, naturally occuring shapes. Statistical features can be extracted efficiently using histogram techniques, which are popular due to their simplicity as well as their good performance.

In this paper a novel statistical shape feature called the Contour Co-occurrence Matrix (CCM) is proposed. The CCM indicates the joint probability of contour directions in a chain code representation of an object's contour. Different versions of the CCM are experimented with and comparisons are made between them and several other shape descriptors from e.g. the MPEG-7 standard. The classification performance is tested with two defect image databases. Some earlier work with these databases and the MPEG-7 features are found e.g. in [3,4].

2 Contour Co-occurrence Matrix (CCM)

The Contour Co-occurrence Matrix (CCM) contains second-order statistics on the directionality of the contours of objects in an image. It resembles the gray level co-occurrence matrix (GLCM) [5], but instead of a two-dimensional image, the co-occurrence information is calculated from the Freeman chain code [6] of the contour of an object. In this regard, it is related to the Chain Code Histogram (CCH) [7] which is the first-order counterpart of CCM.

F. Roli and S. Vitulano (Eds.): ICIAP 2005, LNCS 3617, pp. 253–260, 2005.

2.1 Feature Extraction

The first step in calculating the CCM of an object is to generate the chain code of its contour. The starting point of the contour is not stored, so the resulting feature descriptor is translation invariant.

The co-occurrence matrix is then formed from the pairs of links separated by a given displacement. Let A be a chain of length n and let d be a displacement, i.e. the difference between two chain link indices (not the distance, i.e. the absolute value of the difference). Then the contour co-occurrence matrix $\mathbf{H}^{CCM}$ is defined as a matrix, where the (i, j)th element is the number of instances of a link with value i separated from a link with value j by the displacement d,

$$H_{ij}^{CCM} = \#\{k \mid a_k = i,\ a_{k+d\,(mod\ n)} = j\}, \qquad (1)$$

where # is the number of elements in the set and k runs through the values $0, 1, \ldots n-1$. Because the chain is derived from a closed contour, the index k and displacement d are summed modulo n, so that the displacement wraps around the chain's arbitrary starting point. Since the chain code is octal, the size of the CCM is 8×8.

Implementing rotation invariance is problematic. The contour direction is quantized into eight values, and certain rotation angles result in predictable transformations of the CCM. A rotation of 90° shifts the elements of the matrix by two steps. Rotations of 45° result in a similar shifting effect, but due to the rectangular grid, there can be a significant change in the distribution of edge pixels, an effect similar to quantization noise. Invariance with respect to these rotations can be achieved by matching shifted versions of the matrix, but the differing lengths of chain links in different directions have to be taken into account by normalizing the elements of the CCM to the lengths of the respective link directions.

Multiple displacements can be used in order to obtain information about the contour at different scales, thereby improving the performance of the descriptor. The resulting matrices can be either summed or concatenated to produce the final feature descriptor.

Two basic variations of the CCM may be considered, based on whether the displacement d is constant over all examined contours (let this be called the CCM1), or dependent on the length of the chain, i.e. $d = cn$, where c is a real number in the range $[0, 1[$ (let this be called the CCM2). If the sum of the CCM's elements is normalized to unity, the matrix represents the joint probability of the link values i and j occurring at link indices with the difference d. Thus normalized, the CCM2 becomes scale invariant.

A matrix that has been normalized to unit sum can be interpreted as a probability distribution, from which a set of features can be calculated. These features were originally proposed by Haralick [5] for use with the GLCM, and additional features were introduced by Conners and Harlow [8]. This way the dimensionality of the feature vector can be reduced, which makes computations such as distance calculations more efficient.

2.2 Example CCMs

Some examples of the CCM, calculated as described later in Section 4.1, are presented in Figure 1. The matrices are presented as bitmaps, with intensity representing bin values. For clarity, the values have been normalized so that the highest bin value in a matrix is shown as white. These images show how the CCM captures some general properties of an object. The indices 2 and 6 represent the vertical directions. For a vertical stripe, the matrix contents are concentrated on the intersections of these indices, representing the relationships of points on the same side of the contour, at $(2, 2)$ and $(6, 6)$, and points on opposite sides, at $(2, 6)$ and $(6, 2)$. For a cluster of somewhat round spots with some distinctly vertical features, the contents are spread more loosely around the same elements. For a slightly vertically elongated, irregular spot, some of the same structure is visible, but the contents are more spread out and slightly shifted away from the indices that would represent a regular, vertical shape.

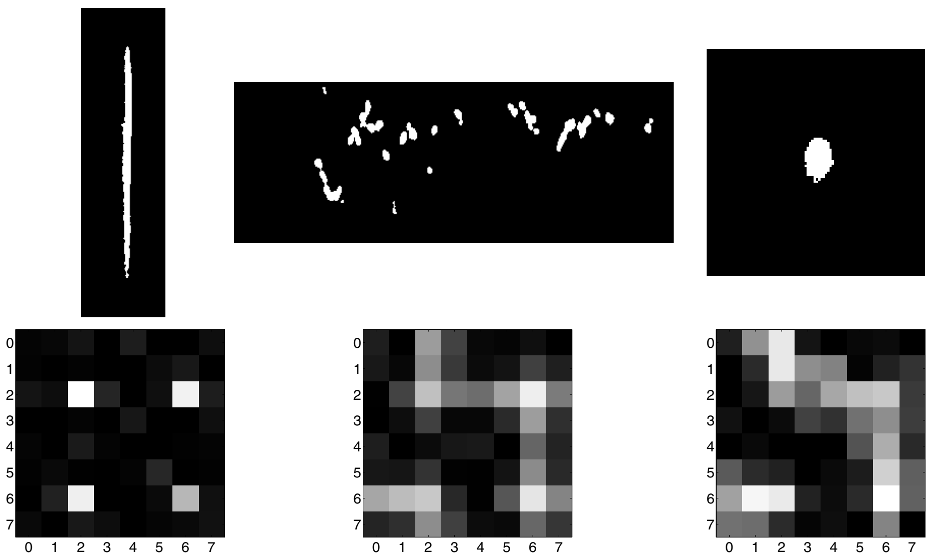


Fig. 1. Contour Co-occurrence Matrices for three images from the metal database

3 Other Shape Descriptors

Other shape descriptors considered in this paper are taken from the MPEG-7 standard, formally named "Multimedia Content Description Interface" [9]. These descriptors were selected for several reasons. They are well standardized descriptors that are used in searching, identifying, filtering and browsing images or video in various applications. In addition to MPEG-7 shape features we also tested three other shape descriptors that we have used previously for defect image classification and retrieval.

Edge Histogram (EH). (MPEG-7) calculates the amount of vertical, horizontal, 45 degree, 135 degree and non-directional edges in 16 sub-images of the picture, resulting in a total of 80 histogram bins.

Simple Edge Histogram (SEH). is similar to its MPEG-7 counterpart, but instead of dividing an image into several sub-images, it is calculated for the whole image.

Contour-based Shape (CBS). (MPEG-7) consists of a set of peak coordinates derived from a Curvature Scale Space (CSS) representation of a contour, and the eccentricities and circularities of the contour and its convex prototype, which is created by repeatedly low-pass filtering the contour.

Region-based Shape (RBS). (MPEG-7) utilizes a set of 35 Angular Radial Transform (ART) coefficients that are calculated within a disk centered at the center of the image's Y channel.

Simple Shape Descriptor (SSD). [10] consists of several simple descriptors calculated from an object's contour. The descriptors are convexity, principal axis ratio, compactness, circular variance, elliptic variance, and angle.

Chain Code Histogram (CCH). [7] is an 8-dimensional histogram calculated from the Freeman chain code of a contour. It is the first-order equivalent of the CCM.

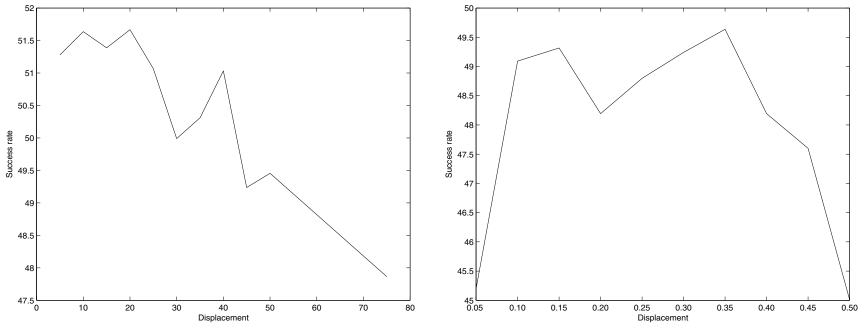
4 Experiments

Experiments were carried out with two image databases containing defect images, one from a metal web inspection system and the other from a paper web inspection system. All images are grayscale images, supplied with binary mask images containing segmentation information, from which the contours of the objects were extracted. The images have different kinds of defects and their sizes vary according to the size of a defect. Classification of defects is based on the cause and type of a defect, and different classes can therefore contain images that are visually dissimilar in many aspects. The paper defect database has 1204 images. They are preclassified into 14 different classes with between 63 and 103 images in all of the classes but one which has only 27 images. The metal defect database has 2004 images. They are preclassified into 14 different classes, with each class containing from 101 up to 165 images. The databases were provided by ABB Oy. More information on these databases can be found e.g. in [3,4].

Classification performance was tested with K-Nearest Neighbor leave-one-out cross-validation, using $K = 5$ and the Euclidean distance measure.

4.1 Displacement Selection

Some initial experiments were made to determine good displacement values. For the CCM1, the range of possible displacements is clearly too large to be exhaustively searched. For a given contour, the CCM1 is periodical with respect



(a) CCM1, metal database (b) CCM2, metal database

Fig. 2. Classification success rates (%) using different displacements

to a displacement. While the period for a single contour is its length n, the period for a set of contours is the least common multiple of the lengths of all the contours. Figure 2(a) shows that for the metal database good results are obtained with displacements in the range from 10 to 40, from which 10 and 20 were chosen. For the paper database a good range is from 20 to 80, and the displacements 20 and 40 were chosen. These values depend on the dataset in a very complex way. In contours of different lengths, a specific displacement value represents a very different relationship between points. The good values for displacement emerge from the entirety of the dataset.

With the CCM2 it is easier to select displacements that can be expected to give sufficiently good results, since the displacement parameter has the range [0 1[. Disregarding rounding effects, the CCM calculated using a relative displacement of c is the transpose of the CCM calculated using a relative displacement of $1 - c$, and therefore only the range $[0, 0.5]$ needs to be examined. Figure 2(b) shows the classification success rates using relative displacements from 0.05 to 0.50 at intervals of 0.05. Based on these results, the relative displacements 0.10, 0.20, 0.30, and 0.40 were chosen, and the matrices were summed together to form the feature vector. Since the displacement is relative to chain length, these choices can be expected to give good results in other databases as well.

4.2 Comparison with Different CCMs

Classification results using the descriptors CCM1 and CCM2 as developed in Section 2.1 are presented in Table 4.2. Although the CCM1 performed better, it also required more care in selecting the displacements. If optimizing the selection of displacements is not possible, e.g. the database grows during use, and the initial set is not representative of the actual pool of data being used, then the CCM2 is probably more reliable, due to the use of relative displacements. Here we assume that the training set is representative of the data, so using the CCM1 with optimized displacements gives a slight advantage. In the remaining experiments only the CCM1 will be used, and will be referred to simply as the CCM.

Table 1. Comparison between the CCM1 and the CCM2

	Classification success rates (%)			
	CCM1 unnorm.	CCM1 norm.	CCM2 unnorm.	CCM2 norm.
Metal	53	49	51	47
Paper	56	58	55	54

4.3 Comparison with Other Shape Descriptors

A comparison was made between the other shape descriptors and the unnormalized CCM, the normalized version and the extracted Haralick features. The 14 features suggested by Haralick in [5] and cluster shade and cluster prominence, added by Conners and Harlow in [8], were considered. The set of features was pruned with a greedy algorithm that eliminated on each iteration the feature that would result in the smallest decrease in the classification rate. The selected features are listed in Table 2. Since the values of the features have very different ranges, the feature vector was whitened using PCA, resulting in a vector where each component has zero mean and unit variance.

Table 2. The sets of features calculated from the CCM and used in classification

Metal	Paper
Difference entropy	Inverse difference moment
Information measures of correlation 1	Entropy
Cluster shade	Information measures of correlation 2
Cluster prominence	Cluster prominence

The CCM results are compared with those obtained with six other shape descriptors: the Edge Histogram (EH), the Simple Edge Histogram (SEH), the Contour-based Shape (CBS), the Region-based Shape (RBS), the Simple Shape Descriptors (SSD), and the Chain Code Histogram (CCH). The results are shown in Tables 3 and 4.

In the metal database the best descriptor was the unnormalized CCM, which scored 4% better than both the normalized CCM and the EH. However, the advantage over the SEH was 9%. Haralick features scored 11% lower than the unnormalized CCM. The CBS and the RBS were clearly the weakest, with the CCH and the SSD falling in-between.

In the paper database the best descriptor was the EH, scoring 2% better than the normalized CCM, which in turn was 2% better than the normalized CCM. However, the SEH was considerably worse, scoring 18% lower than the normalized CCM. This shows that in the paper database dividing the images into segments gives a great advantage. In the metal database the difference was much smaller. Haralick features scored 6% lower than the normalized CCM, the same as the SSD, slightly better than the CBS and the RBS. The CCH was the worst one.

Table 3. KNN classification results of the metal database

	Classification success rates (%)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	avg
CCM unnorm.	25	51	54	41	59	69	56	33	46	50	82	18	52	99	53
CCM norm.	32	25	59	21	27	56	46	48	67	38	91	21	56	99	49
CCM Haralick	25	32	15	32	41	60	47	38	37	30	73	12	46	96	42
EH	37	45	27	30	62	63	54	26	55	61	61	21	33	91	49
SEH	30	15	15	22	29	71	59	14	60	24	87	33	42	97	44
SSD	29	45	8	33	42	59	55	40	25	46	44	16	43	96	42
CCH	32	23	9	22	21	46	41	36	34	20	63	18	36	95	36
CBS	15	38	14	29	42	38	32	11	13	51	42	7	26	65	31
RBS	16	13	13	9	30	37	25	8	17	14	16	7	11	65	20

Table 4. KNN classification results of the paper database

	Classification success rates (%)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	avg
CCM norm.	45	68	43	95	54	28	31	71	82	67	42	11	46	94	58
CCM unnorm.	31	58	44	93	55	19	38	80	82	65	39	19	50	92	56
CCM Haralick	39	55	38	85	48	23	33	59	65	71	33	10	53	89	52
EH	41	23	72	93	58	31	50	74	54	86	30	8	71	93	60
SSD	53	41	40	84	35	23	42	77	70	74	19	6	33	89	52
CBS	48	46	31	69	28	28	46	76	71	74	10	4	19	89	49
RBS	57	26	50	79	17	33	44	43	27	71	17	1	37	86	46
CCH	33	30	41	69	29	16	29	56	49	57	26	1	18	68	40
SEH	31	28	31	93	25	31	30	34	16	77	11	0	25	75	40

5 Discussion

The CCM was developed for use as a part of a feature set in a surface inspection application. The feature set is used in PicSOM [11], a content-based image retrieval system developed in the Laboratory of Computer and Information Science at Helsinki University of Technology. The set contains three MPEG-7 feature descriptors: the Color Structure (CS) for color, the Homogeneous Texture (HT) for texture, and the Edge Histogram (EH) for shape description. The Simple Shape Descriptors (SSD), representing a different approach to shape description, is also included in the set. The dominant feature in this set is the texture feature, while the shape features contribute the least to the retrieval performance. Nevertheless, the CCM has been found to work as well or slightly better than the other shape descriptors in experiments with different feature sets in both KNN experiments and the PicSOM system.

6 Conclusions

In this paper a novel statistical shape feature called the Contour Co-occurrence Matrix (CCM) was presented. The classification performance was tested and compared with several other shape features using two defect image databases. The results in all cases show the CCM to be quite efficient.

The length of the CCM feature vector can be decreased by calculating a set of Haralick features from the matrix. It is possible to keep the decrease in performance quite low by selecting the features individually for each dataset.

Acknowledgments. The financial supports of the Technology Development Centre of Finland (TEKES's grant 40102/04) and our industrial partner ABB Oy (J. Rauhamaa) are gratefully acknowledged.

References

1. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis and Machine Vision. d Chapman & Hall Computing, London (1993)
2. Marshall, S.: Review of shape coding techniques. Image and Vision Computing **7** (1989) 281–294
3. Pakkanen, J., Ilvesmki, A., Iivarinen, J.: Defect image classification and retrieval with MPEG-7 descriptors. In Bigun, J., Gustavsson, T., eds.: Proceedings of the 13th Scandinavian Conference on Image Analysis. LNCS 2749, Gteborg, Sweden, Springer-Verlag (2003) 349–355
4. Iivarinen, J., Rautkorpi, R., Pakkanen, J., Rauhamaa, J.: Content-based retrieval of surface defect images with PicSOM. International Journal of Fuzzy Systems **6** (2004) 160–167
5. Haralick, R., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics **SMC-3** (1973) 610–621
6. Freeman, H.: Computer processing of line-drawing images. Computing Surveys **6** (1974) 57–97
7. Iivarinen, J., Visa, A.: Shape recognition of irregular objects. In Casasent, D.P., ed.: Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling. Proc. SPIE 2904 (1996) 25–32
8. Conners, R., Harlow, C.: Toward a structural textural analyser based on statistical methods. Computer Graphics and Image Processing **12** (1980) 224–256
9. Manjunath, B.S., Salembier, P., Sikora, T., eds.: Introduction to MPEG-7: Multimedia Content Description Interface. John Wiley & Sons Ltd. (2002)
10. Iivarinen, J., Visa, A.: An adaptive texture and shape based defect classification. In: Proceedings of the 14th International Conference on Pattern Recognition. Volume I., Brisbane, Australia (1998) 117–122
11. Laaksonen, J., Koskela, M., Laakso, S., Oja, E.: Self-organising maps as a relevance feedback technique in content-based image retrieval. Pattern Analysis and Applications **4** (2001) 140–152