

# Using Colour Pair Patches for Image Retrieval

*Mike Westmacott, Paul Lewis, and Kirk Martinez*

*Department of Electronics and Computer Science, University of Southampton,  
Southampton, United Kingdom*

## Abstract

An algorithm for image and sub-image retrieval is presented which is based on colour pairs occurring in inhomogeneous patches in the query image. The algorithm is a modification of the Multi-modal Neighbourhood Signature (MNS) method developed at the University of Surrey. By quantising into cells the space used to represent the colour pairs and then indexing the cells in a database, significantly faster retrieval times are achieved. Retrieval performance is enhanced and this is also demonstrated using recall-precision graphs for a test database of images.

## 1. Introduction

Image matching and retrieval based on colour signatures has been a popular approach to content based image handling in image databases.<sup>1</sup> Early work on matching the colour histogram provided a mechanism for finding similar images, at least in terms of global colour content and more elaborate schemes based on hierarchies of histograms provided more spatially related colour information in the signatures. The colour coherence vector approach captures more local information, small regions of colour contributing to the incoherent histogram and large regions contributing to the coherent histogram.<sup>2,4</sup>

Relatively recently, a novel scheme for colour based image (and video frame) retrieval has been developed by Matas and Koubaroulis. Their approach is called the multimodal neighbourhood signature (MNS) algorithm<sup>3</sup> and in their implemented version the signature records colour pairs occurring in bimodal patches in the image. Good retrieval results are reported.

This paper describes variants of a colour pair retrieval (CPR) algorithm which we have developed and tested and which have their roots in the MNS algorithm. The motivation for these developments was to provide a robust signature and indexing strategy allowing rapid content based similarity matching for use with large collections of high resolution images. We begin with a brief description of the MNS algorithm on which our approach is based. This is followed by a description the three variations of our colour pair retrieval algorithm which we will refer to as CPR1, CPR2 and CPR3. We then describe experiments to compare the three CPR algorithms, the standard MNS algorithm and a basic colour histogram algorithm. Finally, our conclusions and plans for further work are described.

## 2. The MNS Algorithm

Briefly, the published MNS algorithm<sup>3</sup> divides the image into a grid of rectangular patches. To avoid aliasing, patches are perturbed with small random displacements in the x and y directions. For each patch, the modes of the colour distribution of its pixels are identified using the mean shift algorithm<sup>5</sup> Only modes corresponding to significant clusters in colour space are considered and, so far, the authors, Matas et. al., have considered bimodal patches. Homogeneous (uni-modal) regions as well as regions with more than two significant modes are discarded.

The two modes for bi-modal patches are represented by a single vector in a six dimensional joint colour space which they denote by *RGB*. A cluster analysis of the vectors for all the bimodal patches in an image is then executed in the *RGB*-space, again using the mean shift algorithm. The modes of the resulting clusters in the *RGB*-space are used to represent the colour pairs present in the image. They constitute the features for the MNS image signature, either directly or transformed in some way to provide appropriate colour invariance properties according to their chosen colour change model.

For image retrieval, signatures are matched by calculating all pairwise distances between the query and database image features. Then for each query image feature in turn, the nearest database image feature is chosen (providing it is less than some threshold), and its distance is used to contribute to the dissimilarity measure and eliminated from further consideration in that match. Some of the reported benefits of the algorithm are that it is fast to compute, provides concise and informative colour content representation and it can be used to reliably locate query objects in database images.

## 3. The CPR Algorithms

In our application, we were seeking fast and reliable similarity matching strategies for content based image retrieval from databases of large numbers of high resolution images, mainly of paintings or other museum artefacts. In some instances, our queries are sub-images and match only a part of one of the images in the database. The development of our colour pair matching began with an implementation of the MNS algorithm. This was partly because of the direct way in which MNS facilitates sub-image matching and partly because of the effective way it makes use of local colour information.

However, the potential size of our database suggested that an indexing strategy for the signatures would be desirable as the complex distance calculations of MNS

were computationally intensive when applied to every database image.

Our first algorithm, CPR1, begins like the MNS algorithm by finding colour pair patches using clustering in the colour space for each patch. Also like MNS, the colour pairs are represented by vectors in *RGB*-space. However, instead of clustering again, we quantise the *RGB*-space and store the information about the colour pairs as a table of cell numbers and frequencies for occupied cells. The frequency of a cell is the number of patches, in that image, for which the colour pair falls in that cell. We also analysed the occupancy of cells across the database and eliminated from consideration those cells which were occupied for a significant fraction of the images in the database. Such cells have poor discriminating power and decrease the performance of the algorithms.

CPR1, CPR2 and CPR3 only differ in the way in which the similarity match is performed. In CPR1 the frequency of pairs in a cell is not used in the match, only whether a cell is occupied or empty. The similarity measure is the number of occupied cells for the query image which are also occupied in the database image. The cell numbers are indexed so that for each occupied cell from the query we retrieve directly the image identifiers of all images containing at least one colour pair in the same cell. We then order the images according to their frequency of occurrence in the total set of identifiers returned. This is essentially a crude quantised version of the MNS algorithm but allowing fast indexed retrieval.

An obvious disadvantage of CPR1 is that occupied cells in *RGB*-space all have the same weight regardless of how many colour pair patches they represent. (MNS has a similar disadvantage in that cluster sizes in *RGB*-space are not taken into account). In our second algorithm, CPR2, for an occupied cell from the query which is also occupied for the database image, the contribution to the similarity index is weighted by the frequency of points in that cell for the query image. This is a valid approach for both whole image matching and also when the query is a sub-image.

For whole image matching, (but not sub-image matching) the closeness of the normalised cell counts for the query and database images can be used as a similarity indicator and this forms the basis of the matching in the third colour pair matching algorithm, CPR3. If  $N F_q$  is the normalized frequency of colour pairs in a cell in the *RGB*<sup>2</sup> space and  $N F_d$  is the normalised frequency for the same cell in a database image, then in CPR3, for each cell in *RGB*<sup>2</sup> space having a nonzero entry for the query, the similarity is incremented by

$$N F_q \left( 1 - \left( N F_q - N F_d \right) / \left( \max(N F_q, N F_d) \right) \right)$$

#### 4. Experimental Results

Each of the three algorithms, CPR1, CPR2 and CPR3 was tested against our implementation of MNS and also against an RGB Histogram algorithm for which the colour space was divided into 64 bins. The algorithms were all written in C++ and a test harness was used to execute multiple

queries and calculate precision and recall metrics for a given test set. The test environment we used consisted of an HP 6-CPU Pentium 600Mhz running Linux, although only one processor was used in retrievals. The relational database, MySQL, was used to store the signatures. In the MNS algorithm the patch size used was 8 x 8 pixels and in the CPR algorithms the *RGB*<sup>2</sup> space was quantised into 4096 cells. The threshold for deleting cells occupied by many images was set at 50% so that cells were deleted if they were occupied for more than half of the images in the database.

The suite of images that was used was a collection of 769 photographic images of varying subjects ranging from scenes of animals to still life. Each image was 384 by 256 pixels in landscape or 256 by 384 pixels in portrait. The collection contains a number of similar images which are grouped together, and these groups were taken to form the basis of categories for our testing. Five groups were identified and group membership was determined manually both by the image content and also the colour layout. The numbers of images in the sets varied between 6 and 47.

A query was run for each image in each category, and for each algorithm. The results were used to calculate precision and recall data for each query together with the speed of execution for the retrievals. Both recall and precision metrics are defined in the normal way:

$$\text{Recall} = \frac{\text{Number of relevant hits in hit list}}{\text{Total number of relevant documents}}$$

$$\text{Precision} = \frac{\text{Number of relevant hits in hit list}}{\text{Total hits in the hit list}}$$

An overall metric that takes into account the quality of the ranking of retrieved documents is average precision. This is calculated as the sum of the precision at each relevant hit in the hit list divided by the total number of documents.

The precision and recall data was calculated for each query, and averaged across each category, and algorithm. For each query the top 100 retrieved images were taken into consideration.

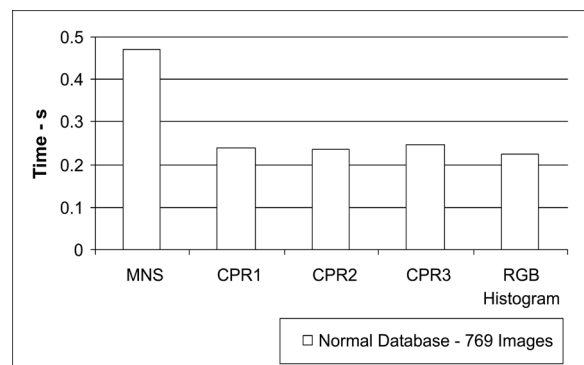


Figure 1. Comparison of retrieval Times

The first aim of our implementation was to develop an algorithm which allowed indexed retrieval so that it would scale sensibly for very large image databases. In figure 1 we show a comparison of the average retrieval times for the five algorithms. It can be seen that the CPR algorithms are significantly faster than MNS and of the same order of magnitude as the times for the RGB histogram retrieval for this size of database.

Algorithm	Average Precision
MNS	0.356
QMNS CPR1	0.461
QMNS CPR2	0.45
QMNS CPR3	0.464
RGB Histogram	0.392

Figure 2. Average Precision Values

In order to compare retrieval quality we use the recall and precision metrics. In figure 2 we show the average precision, as defined earlier, for each of the five algorithms. It can be seen overall the three CPR algorithms do best with CPR3 having the highest average precision.

In figure 3 we show the recall-precision curves for the five algorithms. It can be seen that, although the MNS and histogram curves cross the CPR curves, the CPR algorithms do consistently better at the higher recall values.

The area under the graphs corresponds to the average precision values given in figure 2.

In figure 4 we show examples of the ten best matches for a whole image retrieval using MNS and using CPR3 and in figure 5 we show an example of sub-image retrieval using MNS and CPR2. Note that CPR3 can not be used reliably for sub-image retrieval because it uses frequencies from the database image in the similarity calculation.

## 5. Conclusions and Future Work

Signatures for content based retrieval using colour pairs have been developed and algorithms for comparing images using the signatures have been described. The CPR approach provides fast indexed image and sub-image retrieval with the additional advantages of local colour characterisation. The increased speed is achieved whilst also improving on the retrieval performance of earlier approaches.

Ways of providing selection mechanisms for the patches to be analysed rather than processing all patches in the initial grids are being considered. The CPR approach suffers from the problems inherent in quantisation approaches when points fall near to a cell boundary. To overcome this, indexing the cluster mode vectors in RGB-space directly (rather than cell numbers) using a multidimensional indexing strategy is also under consideration. The number of points in each cluster would then replace the cell counts in the CPR2 and CPR3 algorithms.

## 6. Acknowledgements

The first author wishes to acknowledge the support of an EPSRC studentship and the other two authors are grateful to the European Commission for their support through grant IST-1999-11978.

## 7. References

1. A. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-Based Retrieval at the End of the Early years, *Trans. Pattern Analysis and Machine Intelligence*, vol 22, 1349-1380, 2000.
2. G. Pass, R. Zabih, J. Miller, Comparing Images Using Colour Coherence vectors, *ACM Conference on Multimedia*, 65- 73, 1996.
3. J. Matas, D. Koubaroulis, J. Kittler, Colour Image Retrieval and Object Recognition Using Multi-modal Neighbourhood Signature, *Proc. European Conf. On Computer Vision*, 48- 64, 2000
4. S. Chan, K. Martinez, P. Lewis, C. Lahanier, J. Stevenson, Handling Sub-image Queries in Content-Based Retrieval of High Resolution Art Images, *International Cultural Heritage Informatics Conference*, 157-163, 2001
5. K. Fukunaga, L. Hostetler, The Estimation of the Gradient of a Density Function with Applications in Pattern Recognition, *IEEE Trans. in Information Theory*, 32-40, 1975.

## 8. Biography

Michael Westmacott graduated in Computer Science in 2000 at the University of Southampton, Department of Electronics and Computer Science. He is now a graduate student doing research in the area of content based image retrieval in the same department under the supervision of Dr Paul Lewis.

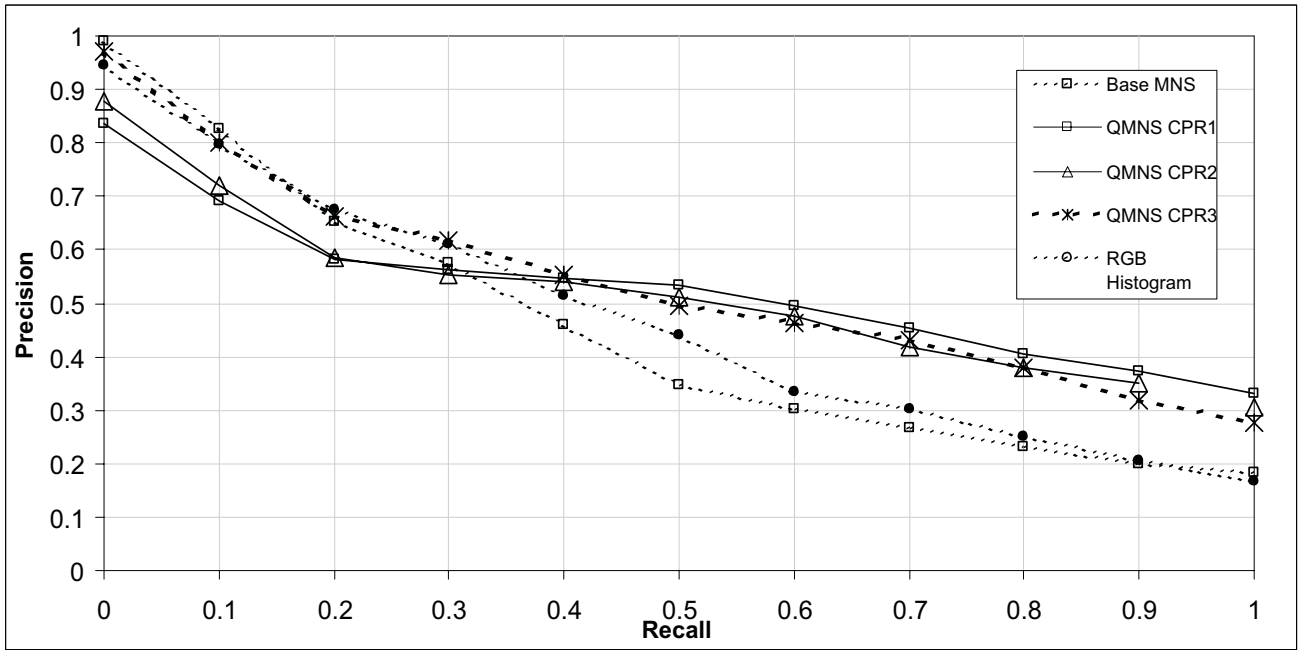


Figure 3: Recall-Precision Curves for the Algorithms

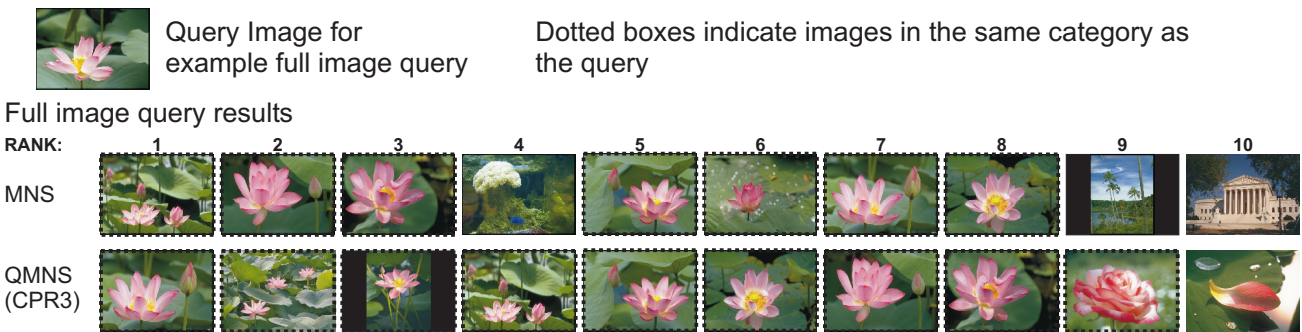


Figure 4: Example of Full Image Retrieval

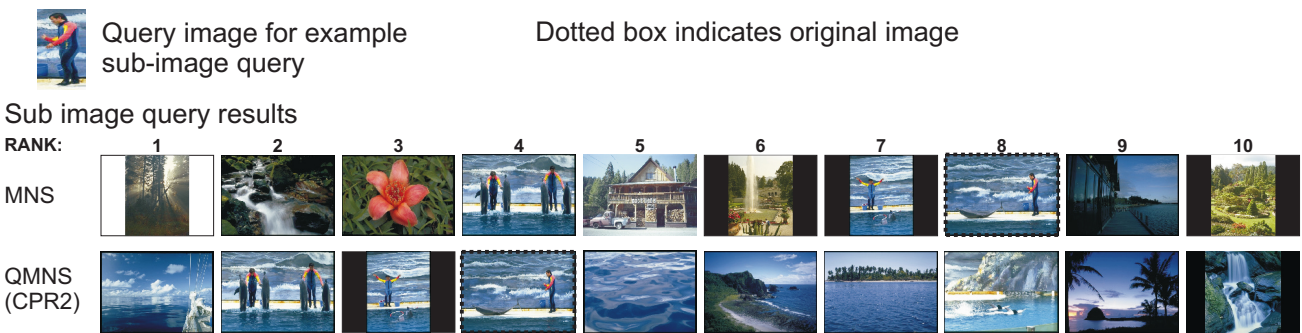


Figure 5: Example of Sub-image Retrieval