Gamut Intersection for Image Retrieval

Andrei Ouglov, Ali Alsam and Rune Hjelsvold
Dept of Computer Science and Media Technology
Gjøvik University College

Abstract

Colour histograms are the most dominate technique for image indexing based on image colour content. The colour histogram approach approximates a three-dimensional colour distribution of an image to a three-dimensional colour histogram. This paper describes image retrieval experiments using a novel Gamut Intersection approach. Gamut Intersection is an attempt of approximating an images' three-dimensional colour distribution by projecting it onto two orthogonal projection planes defined in the rgb-cube. This results in two 0 1 binary two-dimensional images which we use as our image descriptors. The method retains the advantages of colour histograms such as a simple computation, robustness to image rotation, image scaling and distribution of objects in the image. When comparing our method with a $16 \times 16 \times 16$ *histogram approach, we found that our new approach performs favorably, or equally well, to histograms for all the, 25, test images.*

Introduction and related work

The need for content-based image retrieval methods requires a selection of simple and effective image features for comparing images based on their overall appearance. Although numerous techniques have been proposed throughout the years, many of the content based image retrieval problems remain unsolved. Especially a problem relating to which features are sufficient to describe an image has been thoroughly investigated since the early 1990's. Nevertheless colour is agreed to be one of the most important and thus widely used features in image indexing and retrieval. The use of colour histogram has been and still is a common approach in colour image indexing. Histogram-based image indexing was originally described by Swain and Ballard.[1] Although a number of modifications to image indexing using both global and local histograms have been proposed[2–8] the actual technique is still very much the same. Advantages of colour histogram methods are its compactness and invariance to scaling and rotation. However, there are also some disadvantages such as the problems associated with quantization and bin misalignment.

Colour histograms are widely used alone or in combination with texture and shape information, for example by,[9 10] and.[11] The idea of colour-based image retrieval techniques, including histogram indexing, is to retrieve images that have perceptually similar colours to the user's query image.

Besides colour histogram, several other colour based methods have been developed. They include Colour Moments and Cumulative Colour,[3] Colour Sets,[6, 12] Colour Correlogram,[13] colour signatures and local colour regions,[14] colour coherent vectors,[15] and blobs.[16]

Image indexing by gamut intersection

Before we describe our new method let us consider colour distributions in the *rgb* space, of two arbitrary colour images. Examples of such distributions are shown in Figures 1:A and 1:B.

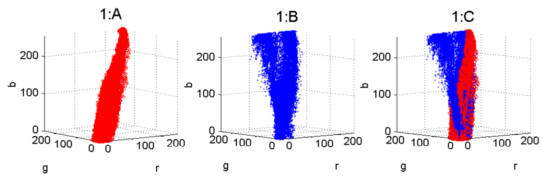

Figure 1. *1:A and 1:B show colour disributions of arbitrary images in the rgb space. 1:C shows the same colour distributions superimposed onto each other.*

Although, the *rgb* colour distributions depicted in the Figures 1:A and 1:B are specific to the actual images; they are similar in a number of ways. For example, we note that the data is elongated along the grey axis, i.e. the line segment from point $[0,0,0]$ to $[1,1,1]$ or black to white. Further, the data is stretched significantly less in the other directions. Having said that, we note that there is a clear difference between the colour distributions of the two images, which is especially evident when we superimpose the two images onto each other as is shown in the Figure 1:C.

As an outcome of the previous discussion, we propose a method based on the level of difference between the *rgb* distributions of the images. To achieve this we define a new image descriptor based on the image gamut as well as a metric to compare two descriptors.

A projection approach to image retrieval

From Figure 1:C we might wonder if it is possible to use the difference between the two *rgb* gamuts as a means for comparing the images. However, for such an approach to be successful it is necessary to store the whole image gamut as our image descriptor which means that we would lose one of the most significant advantages of colour histograms, namely, compactness. Instead of using the complete three-dimensional image gamut as our descriptor, for approximation, we propose to use the projection of the gamut onto two orthogonal planes; these are chosen such that the grey-axis is defined in the space of each plane. Prior to motivating our choice for the two projection planes, we start by giving a formal definition. The first plane which we call P_1 is defined as:

$$P_1 = \alpha[1\ 1\ 0]^T + \beta[0\ 0\ 1]^T \tag{1}$$

where α and β are scalars.

Similarly, the second projection plane is defined as:

$$P_2 = \alpha[1\ 0\ 0.5]^T + \beta[0\ 1\ 0.5]^T \tag{2}$$

Considering a unite cube the two planes are shown in Figure 2.

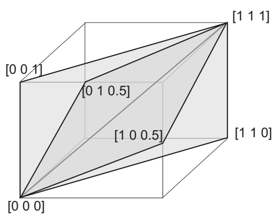

Figure 2. *The two orthogonal planes P_1 and P_2.*

The most important motivation for choosing the projection planes defined in Equations (1) and (2), is that the planes should be orthogonal to ensure that the projects are linearly independent and that both planes should contain the grey-axis since the colour distributions are elongated along this axis as illustrated in Figure 1. Having defined the planes, P_1 and P_2, we need to define the projection operators which take any point p in the *rgb* space to P_1 and P_2 respectively. For P_1 the projection operator is defined as:

$$PO_1 = P_1 \left(P_1^T P_1\right)^{-1} P_1^T \tag{3}$$

Similarly, for P_2 the projection operator is defined as:

$$PO_2 = P_2 \left(P_2^T P_2\right)^{-1} P_2^T \tag{4}$$

If we orthonormalize the vectors in both P_1 and P_2 such that:

$$P_1^T P_1 = P_2^T P_2 = I \tag{5}$$

where I is the identity matrix, then the projections operators one the respective planes can be rewritten as:

$$PO_1 = P_1 P_1^T \tag{6}$$

and

$$PO_2 = P_2 P_2^T \tag{7}$$

Now, if we project the *rgb* data of an image M onto P_1 and P_2 we will get two matrices which represent the three dimensional colour gamuts in the two-dimensional spaces of the respective planes. Let us define the first of these matrices as:

$$Q = MPO_1 \tag{8}$$

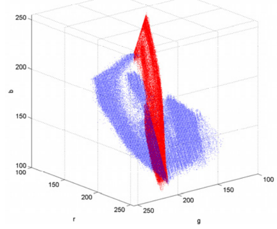

Figure 3. *The projection of an arbitrary image onto planes Q and W.*

and the second matrix as:

$$W = MPO_2 \tag{9}$$

Figure 3 shows the projection of an arbitrary image onto Q and W. We note that both Q and W are $n \times 3$ matrices where n is the number of pixels in the image M. We know, however, that any point in Q or W is defined in a two-dimensional plane. Thus, we are able to reduce the dimensionality of both Q and W to $n \times 2$. With reference to Figure 2 we can define the two axes for the first plane, P_1 as:

$$\widehat{q}_1 = \sqrt{q_1^2 + q_2^2} \tag{10}$$
$$\widehat{q}_2 = q_3$$

where q_1, q_2 and q_3 are coordinates of a point defined on the plane P_1 and are entries in the first, second and third columns of matrix Q. Similarly, by referring to Figure 2, we define the two dimensional coordinates of a point in plane P_2. A point in a plane P_2 is defined by matrix W and has following maximum values along x, y and z axis in the *rgb*-cube

$$w_{1max} = w_{2max} = w_{3max} = 255 \tag{11}$$

With reference to Figure 2 the maximum values the two dimensional coordinates of a point in P_2 defined by matrix $\widehat{W}$ are:

$$\widehat{w}_{1max} = \sqrt{w_{1max}^2 + \left(\frac{w_{3max}}{2}\right)^2} \tag{12}$$
$$\widehat{w}_{2max} = \sqrt{w_{2max}^2 + \left(\frac{w_{3max}}{2}\right)^2}$$

now two dimensional coordinates of a point in $\widehat{W}$ can be calculated as follows

$$\widehat{w}_1 = w_1 * \left(\frac{\widehat{w}_{1max}}{w_{1max}}\right) \tag{13}$$
$$\widehat{w}_2 = w_2 * \left(\frac{\widehat{w}_{2max}}{w_{2max}}\right)$$

Considering the new two-dimensional matrices $\widehat{Q}$ and $\widehat{W}$, as well as the representation of projection planes P_1 and P_2 plotted in Figure 2 and that of the example given in Figure 3; the size of $\widehat{Q}$ and $\widehat{W}$ is dependent solely on the dimensionality of the *rgb* colour distribution. Said differently, the size of the matrices $\widehat{Q}$ and $\widehat{W}$ is not related to the original size of the image M but rather to the size of the colour space. Given that the white point in the *rgb* is defined as point [255 255 255] then the dimensions of $\widehat{Q}$ will be $255 \times \sqrt{255^2+255^2}$ and the size of $\widehat{W}$

$\widehat{W}$ will be $\sqrt{255^2+\left(\frac{255}{2}\right)^2} \times \sqrt{255^2+\left(\frac{255}{2}\right)^2}$

or after factorization

$[255 \times 255\sqrt{2}]$ and $[\frac{255}{2}\sqrt{5} \times \frac{255}{2}\sqrt{5}]$.

Describing the gamut projection as a 0 1 binary image

We started our discussion by assuming that; it is possible to assess the similarity between two images based on the difference between their respective gamuts. Thus far, we have defined two two-dimensional matrices $\widehat{Q}$ and $\widehat{W}$ which serve to approximate the shape of the gamut in two dimensional spaces, however, both $\widehat{Q}$ and $\widehat{W}$ include some added information about the gamut, namely, the density or the number of points in the original three dimensional space which will project onto the same point in $\widehat{Q}$, i.e. $[\widehat{q}_{1j}\ \widehat{q}_{2j}]$ or equivalently the same point in $\widehat{W}$, i.e. $[\widehat{w}_{1j}\ \widehat{w}_{2j}]$, where j is an index. Using these matrices as our image descriptor is inefficient as it requires saving two additional matrices of a size that is equivalent to the original image. Further, we are only interested in the shape and distribution of the gamut rather than the density of the points in the projection planes. We thus propose to represent $\widehat{Q}$ and $\widehat{W}$ as 0 1 binary images where any point in the image can only have a value of one or zero. For $\widehat{Q}$, the transformation which results in the removal of the density information is defined as:

$$\widehat{Q}_b(x,y) = \left\{ \begin{array}{lllll} 1 & if & x=\widehat{q}_{1j} & and & y=\widehat{q}_{2j} \\ 0 & if & x\neq\widehat{q}_{1j} & or & y\neq\widehat{q}_{2j} \end{array} \right\} \tag{14}$$

where $\widehat{Q}_b$ is a matrix that has the same size as $\widehat{Q}$, i.e. $n \times 2$. Equally, for the second projection image $\widehat{W}$ the 0 1 binary image is defined as:

$$\widehat{W}_b(x,y) = \left\{ \begin{array}{lllll} 1 & if & x=\widehat{w}_{1j} & and & y=\widehat{w}_{2j} \\ 0 & if & x\neq\widehat{w}_{1j} & or & y\neq\widehat{w}_{2j} \end{array} \right\} \tag{15}$$

following the same procedure as before $\widehat{W}_b$ is a matrix which has the same size as $\widehat{W}$.

As an example we show a sample image form the **MPEG7** database Figure 4 together with its binary projections $\widehat{Q}_b$ and $\widehat{W}_b$.

The similarity of projection images

Now we need to define a similarity metric to compare projection images. Having calculated the projection images for all the images in a database we need a similarity metric; in this paper we propose using the difference between the projection images. For the projection images, $\widehat{Q}_b$ and $\widehat{W}_b$, we define the difference as:

Figure 4. *Projection of colour distribution onto the projection planes: right - source image, $\widehat{Q}_b$, left and $\widehat{W}_b$, middle.*

$$D\left(\widehat{Q}_b-\widehat{Q}_b^i\right) = \widehat{Q}_b \bigcup \widehat{Q}_b^i - \widehat{Q}_b \bigcap \widehat{Q}_b^i \tag{16}$$

$$D\left(\widehat{W}_b-\widehat{W}_b^i\right) = \widehat{W}_b \bigcup \widehat{W}_b^i - \widehat{W}_b \bigcap \widehat{W}_b^i \tag{17}$$

Said in words, the difference between two projection images is equal to the intersection of the white pixels subtracted from their union. Note that the index i in Equation (16) indicates the ith image in the database.

As an example, we show the projection images $\widehat{Q}_b$, left and middle images in Figure 5, for two similar images from the **MPEG7** database as well as their difference as defined in Equation (16).

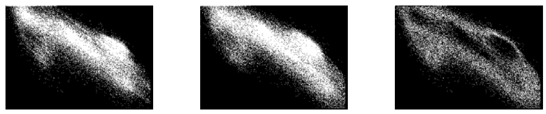

Figure 5. *Projection images similarity: the left most and the middle images show projection images of two similar images; right - their projection similarity.*

Finally, the total difference between two images M and M_i is defined as the sum:

$$D\left(M-M^i\right) = D\left(\widehat{Q}_b-\widehat{Q}_b^i\right) + D\left(\widehat{W}_b-\widehat{W}_b^i\right) \tag{18}$$

The first image in Figure 5 shows that some pixels are scattered or disconnected in a fashion which might result in a higher value of images difference than the one images actually should have. To address this problem we use the fillhole algorithm[17] on 4-connected neighbourhood and a median filter (on a 7-by-7 neighbourhood). As a result, small holes in the projection images are substituted with a filled region and the effect of the scattered pixels which lie outside the dense white region is reduced. The result of applying these processing steps is shown in the middle and most right images in Figure 6.

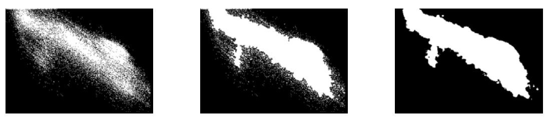

Figure 6. *Projection images similarity: the left most and the middle images show projection images of two similar images; right - their projection similarity.*

The similarity between the same pair of images as in Figure 5 after described pre-processing steps is shown in Figure 7.

Figure 7. Projection images similarity: 2 compared projection images, left and middle, right - their projection similarity.

Pre-processed projection images will be used as our image descriptors in the following section.

Experiment and results

Our experiment was performed on **MPEG7** image database. This database is commonly used for testing image retrieval algorithms; and is considered to be reasonably large and representative as it consists of 5466 images, covering a wide range of photographic images and sequences of video frames. Further, the database includes 50 ground-truth image sets each consisting of between 2 and 32 similar images. Thus, a successful indexing will return all images in a particular ground-truth set plus the query image itself. In our experiment 25 ground-truth query images were indexed and ranked using colour histogram with $16 \times 16 \times 16$ bins and our novel gamut intersection approach. Results from the two approaches were compared and showed that our new approach performs favorably to histograms for almost all the test images. The results are shown in Table 1.

Table 1: Image retrieval results using color histogram and gamut intersection approaches.

Query image	N	NH	NI
1-4	15	15/15	15/15
5	15	15/28	15/17
6	6	6/6	6/6
7-15	11	11/11	11/11
16	24	24/24	24/24
17	5	5/24	5/5
18	2	2/2	2/2
19	8	8/8	8/8
20-21	3	3/3	3/3
22	7	7/7	7/7
23	4	4/4	4/4
24	5	4/10	5/5
25	6	4/21	6/13

Where **N** is the size of the ground-truth set; **NH** is the ratio between retrieved images from the ground-truth set using colour histogram approach and the size of the ground-truth set; **NI** is the ratio between retrieved images from the ground-truth set using Gamut intersection and the size of the ground-truth set.

Conclusion

In this paper we presented a novel approach to image indexing based on colour distributions' gamut intersection. It approximates image colour distribution by projecting it onto two orthogonal projection planes which produces projection images to be used for image comparison. The gamut intersection approach matches colour histograms' advantages, such as a compact and simple computation, robustness to image rotation, image scaling and distribution of objects in the image. Our experiments performed on 25 images from the **MPEG7** database show that the gamut intersection approach performs favorably or equally well when compared with a $16 \times 16 \times 16$ bin histogram. However, it should be mentioned that achieved result is based on the 25 test images and an experiment based on different test images might come to a different conclusion.

References

[1] M.J. Swain and D.H Ballard. Color indexing. *Int Journal of Computer Vision*, 7(1):11–32, 1991.

[2] B.V. Funt and G.D. Finlayson. Color constant color indexing. Technical Report 91-09, School of Computing Science, Simon Fraser University, Vancouver, B.C., Canada, 1991.

[3] M. Stricker and M. Orengo. Simularity of color images. In *SPIE Storage and Retrieval for Image and Video Databases*, 1995.

[4] C.H. Chuan Y. Gong and G. Xiaoyi. Image indexing and retrieval using color histograms. *Multimedia Tools and Applications*, 2:133–156, 1996.

[5] T.S. Chua W. Hsu and H.K. Pung. An integrated color-spatial approach to content-based image retrieval. In *Proc. of the ACM Multimedia 95*, pages 305–313, 1995.

[6] J. Smith and S. Chang. Tools and techniques for color image retrieval. In *Proc. of the SPIE conference on the Storage and Retrieval for Image and Video Databases IV*, SPIE Proceedings Series, pages 426–437, San Jose, CA, USA, 1996.

[7] M.A. Stricker and A. Dimai. Color indexing with weak spatial constraints. In *Proc. of the SPIE conference on the Storage and Retrieval for Image and Video Databases IV*, SPIE Proceedings Series, pages 29–40, San Diego, CA, USA, Feb 1996.

[8] R. M. Rickman and T. J. Stonham. Content-based image retrieval using color tuple histograms. In *Storage and Retrieval for Image and Video Databases (SPIE)*, pages 2–7, 1996.

[9] V. Ogle and M. Stonebraker. Chabot retrieval from a relational database of images. *IEEE Computer*, 28(9):40–48, 1995.

[10] R. Picard A. Pentland and S. Sclaroff. Photobook content-based manipulation of image databases. *International Journal of Computer Vision*, 18(3):233–254, 1996.

[11] W. Niblack J. Ashley Q. Huang B. Dom M. Gorkani J. Hafner D. Lee D. Petkovic D. Steele M. Flickner, H. Sawhney and P. Yanker. Query by image and video content the qbic system. *IEEE Computer*, 28(9):23–32, 1995.

[12] J. Smith and S. Chang. Single color extraction and image query, 1995.

[13] M. Mitra W. Zhu J. Huang, S. Kumar and R. Zabih. Image indexing using color correlograms, 1997.

[14] L. Guibas Y. Rubner and C. Tomasi. The earth mover's distance, multi-dimensional scaling, and color-based image retrieval. In *Proc. ARPA Image Understanding Workshop*, 1997.

[15] G. Pass and R. Zabih. Histogram refinement for content-based image retrieval, 1996.

[16] H. Greenspan C. Carson, S. Belongie and J. Malik. Region-based image querying. In *Proceedings CVPR '97 Workshop on Content-Based Access of Image and Video Libraries*, 1997.

[17] P.Soille. *Morphological Image Analysis*. Springer-Verlag, 2nd edition, 2003.

