# Fixed Representative Colors Feature Extraction Algorithm for Moving Picture Experts Group-7 Dominant Color Descriptor

[1]Ahmed Ibrahim, [2]Ala'a Al-Zou'bi, [3]Raed Sahawneh and [3]Maria Makhadmeh
[1]Department of Computer, Princess Nora Bint Abdul Rahman University, Riyadh, KSA
[2]Department of Computer Science, Irbid National University, Irbid, Jordan
[3]Department of Computer Information system, Irbid National University Irbid, Jordan

**Abstract: Problem statement:** MPEG-7 has defined a set of color descriptors for the multimedia contents; one of these descriptors is the Dominant Color Descriptor (DCD). The DCD adopted the Generalized Lloyd Algorithm (GLA) as a feature extraction technique. However, using this algorithm in image retrieval systems may produce perceptually irrelevant images in the ranked results. **Approach:** In this study, we proposed a new technique for the extraction of the DCD called Fixed Representative Colors Feature Extraction Algorithm (FRCFE). This technique is based on converting the original image colors to a set of predefined colors in the RGB color space. Also, we developed an image retrieval system to measure the effectives of this technique. **Results:** Experimental results revealed that FRCFE provides better accurate and perceptually relevant image retrieval for the MPEG-7 DCD than GLA in the chosen dataset. **Conclusion:** Mapping the original image colors to a set of predefined colors preserved the main features of the image and can produce more accurate results when used in image retrieval systems.

**Key words:** Moving picture experts group-7, image retrieval, dominant color descriptor, feature extraction

## INTRODUCTION

**Overview:** The term Content Based Image Retrieval (CBIR) describes the process of retrieving desired images from a large collection of images on the basis of features such as color, shape and texture that can be automatically extracted from the images themselves[1].

The ideal approach of querying an image database is using content semantics, which applies the human understanding about image[2]. It is like an information filter process and is expected to provide a high percentage of relevant images in response. In general, the image features tend to capture only some of the aspects of image. So, retrieval systems cannot be expected to find all correct images. Otherwise they select the most similar images to let the user choose the relevant images.

Among different types of features, color is the most straightforward information which can be easily retrieved from digital images with simple, while others require more processing and computational tasks[3].

While comparing image by color feature, two properties are usually considered:

- Area of matching-Count the area or number of pixels having same or similar colors. Larger matched area means more similar
- Color distance-Distance between colors, usually in a uniform color space such as CIELuv. Closer between matched colors means more similar

In a typical color similarity measure, area of matching is usually considered as the similarity. Color distance is used to control the matching between colors and to adjust the similarity.

**MPEG-7 dominant color descriptor:** An important achievement for CBIR is MPEG-7. MPEG-7 is an international standard for multimedia content description[4]. MPEG-7 has a collection of effective descriptors for images, videos, audios and other multimedia contents. In its visual part, several color descriptors are defined, in which Dominant Color Descriptor (DCD) is a compact and effective descriptor[5].

In DCD, Image feature is formed by a small number of representative colors. These colors are normally obtained by using clustering and color quantization. The descriptor consists of the

**Corresponding Author:** Ahmed Ibrahim, Department of Computer Princess Nora Bint Abdul Rahman University, Riyadh, KSA

representative colors, their percentages in a region, spatial coherency of the color and color variance[5]. The DCD is defined as (1):

$$F \{(c_i, p_i, v_i), s\}, (i = 1, 2, ..., N) \tag{1}$$

where, N is the number of dominant colors. Each dominant color value $c_i$ is a vector of corresponding color space (e.g., RGB color space). The percentage $p_i$ is the fraction of pixels in the image or image region corresponding to color $c_i$ and $\sum_i p_i = 1$. The optional color variance $v_i$ is the variation of the color values of the pixels in the corresponding representative color. The spatial coherency s is a single number that represent the overall spatial homogeneity of the dominant colors in the image.

The number of dominant colors N is suggested that a maximum of eight dominant colors is sufficient to represent an image or an image region. The dominant color values $c_i$ uses 1-12 bits per color component. It is controlled by the color space and color quantization level chosen[6]. Table 1 shows the binary syntax for each component.

**DCD extraction:** The extraction procedure for the dominant color uses the Generalized Lloyd Algorithm (GLA)[7] to cluster the pixel color values. It is recommended that the clustering be performed in a perceptually uniform color space such as CIELuv[8]. The distortion Di of i-th cluster is given as (2):

$$D_i = \sum_n h(n) \left\| x(n) - c_i \right\|^2, x(n) \in C_i \tag{2}$$

Where:
$c_i$         = The centroid of cluster
$C_i, x(n)$ = The color vector at pixel n
$h(n)$      = The perceptual weight for pixel n

Table 1: Binary syntax of dominant color descriptor

| Field | Suggested No. of bits | Meaning |
| --- | --- | --- |
| No. of colors | 3 | Specifies number of dominant colors (1-8) |
| Spatial coherency | 5 | Spatial coherency value (optional) |
| Percentage | 5 (per color) | Normalized percentage associated with each dominant color |
| Color Variance | 1 (per color component) | Color variance of each dominant color (optional) |
| Index | 1-12 (per color component) | Dominant color values |

The perceptual weights are calculated from local pixel statistics to account for the fact that human visual perception is more sensitive to changes in smooth regions than in textured regions.

The procedure is initialized with one cluster consisting of all pixels and one representative color computed as the centroid (center of mass) of the cluster. The algorithm then follows a sequence of centroid calculation and clustering steps until a stopping criterion (minimum distortion or maximum number of iterations) is met. The cluster with highest distortion is split by adding perturbation vectors to the centroid until the maximum distortion falls below a predefined threshold or the maximum number of clusters is generated. The percentage of pixels in each cluster of the image is then quantized to five bits. The color values are quantized according to the specifications of the color space and the associated color-quantization descriptors. Figure 1 shows the idea of DCD extraction.

**Fixed Representative Colors Feature Extraction Algorithm (FRCFE):** In FRCFE, we use a color quantization in RGB space using 38 perceptual colors to represent images based on dominant colors. Table 2 shows these colors selected from the RGB color space.
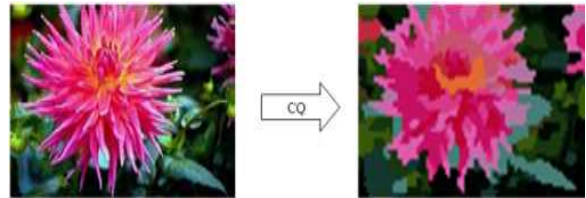


Fig. 1: DCD color quantization using GLA

Table 2: Perceptual colors to represent images based on dominant colors

| S/No. | Red | Green | Blue | S/No. | Red | Green | Blue |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 255 | 0 | 0 | 20 | 0 | 204 | 255 |
| 2 | 255 | 102 | 0 | 21 | 200 | 200 | 200 |
| 3 | 255 | 255 | 153 | 22 | 210 | 210 | 175 |
| 4 | 255 | 255 | 0 | 23 | 255 | 0 | 255 |
| 5 | 128 | 0 | 0 | 24 | 255 | 153 | 204 |
| 6 | 255 | 153 | 0 | 25 | 0 | 255 | 255 |
| 7 | 255 | 153 | 51 | 26 | 153 | 51 | 0 |
| 8 | 0 | 51 | 0 | 27 | 153 | 102 | 51 |
| 9 | 51 | 153 | 51 | 28 | 102 | 128 | 51 |
| 10 | 51 | 51 | 0 | 29 | 204 | 102 | 0 |
| 11 | 128 | 128 | 51 | 30 | 0 | 0 | 0 |
| 12 | 51 | 153 | 102 | 31 | 255 | 255 | 255 |
| 13 | 153 | 204 | 0 | 32 | 51 | 102 | 153 |
| 14 | 0 | 255 | 0 | 33 | 0 | 51 | 102 |
| 15 | 0 | 128 | 0 | 34 | 0 | 128 | 128 |
| 16 | 51 | 102 | 255 | 35 | 51 | 51 | 153 |
| 17 | 153 | 204 | 255 | 36 | 153 | 51 | 102 |
| 18 | 0 | 0 | 255 | 37 | 204 | 153 | 255 |
| 19 | 0 | 0 | 128 | 38 | 85 | 75 | 60 |
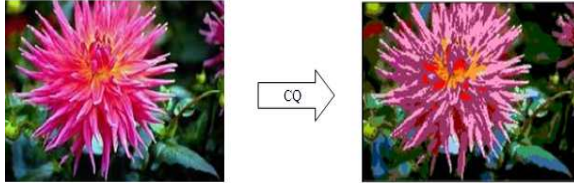
Fig. 2: DCD color quantization using FRCFE

After defining the colors, for each image implement the following steps:

- Read in the image and create an image array that contains the RGB components of each pixel in the image
- For each pixel in the image do:
  - Search color table for the nearest color by finding the distance between the pixel color I represented as $(P_r, P_g, P_b)$ and the color in the color table $C_i$ represented as $(C_{iR}, C_{iG}, C_{iB})$ using the distance formula (3):

$$C_d = \min\left( \sqrt{\frac{(P_r - C_{ir})^2 + (Pg - C_{iG})^2}{+ (Pb - C_{iB})^2}} \right), \tag{3}$$
$$i = 1, 2, ..., 38$$

  - Assign to the pixel the RGB entry in color table for which $C_d$ is the minimum
- Create a frequency table for each assigned color
- Sort the frequency table in descending order MPEG-7 DCD allows at most eight colors to be represented, therefore, the highest eight frequent colors are then selected with their percentages to create the description of the image. Figure 2 shows how an image will be converted

Since MPEG-7 uses XML as the textual representation for the descriptors[6], The DCD representation for the image in Fig. 2 is shown in Fig. 3 (without considering the optional color variance and spatial coherency fields).

**Experimental environment:**
**Similarity matching:** Given a query image, similarity retrieval involves searching the database for similar color distributions as the input query. Consider two DCDs F1 and F2 given by 4 and 5:

$$F_1\{(c_{1i}, p_{1i}, v_{1i}), s_1\}, (i = 1, 2, ..., N_1) \tag{4}$$



Fig. 3: The DCD representation for the image in Fig. 2

$$F_2\{(c_{2i}, p_{2i}, v_{2i}), s_2\}, (i = 1, 2, ..., N_2) \tag{5}$$

Without considering the optional color variance parameter and the spatial coherence, quadratic histogram distance measure is used to measure the dissimilarity $D^2 (F_1, F_2)$ between the two descriptors, which is defined as 6:

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} P_{1_i}^2 + \sum_{j=1}^{N_2} P_{2_{ji}}^2 - \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} 2a_{1i,2j} P_{1i} P_{2j} \tag{6}$$

where, $a_{1i}, 2_j$ is the similarity coefficient between two colors $c_{1i}$ and $c_{2j}$, it is identified as (7):

$$a_{1i,2j} = \begin{cases} 1 - \|c_{1i} - c_{2j}\| / aT_d, & \|c_{1i} - c_{2j}\| \leq T_d \\ 0, & \|c_{1i} - c_{2j}\| > T_d \end{cases} \tag{7}$$

Where:
$\|c_{1i} - c_{2j}\|$ = The Euclidean distance between two colors $c_{1i}$ and $c_{2j}$ in CIELuv color space
$T_d$ = The maximum distance for two colors to be considered as similar

In particular, this means that any two dominant colors from one single description are more than Td distance apart. A recommended value for Td is between 10 and 20 in the CIELuv color space and for α is between 1.0 and 1.5. The above dissimilarity measure is very similar to the quadratic distance measure that is commonly used in comparing two color histogram descriptors[3].

**Evaluation:** In MPEG-7 Experiments, datasets and evaluation metrics are defined[9]. For evaluation of retrieval performance, MPEG group have defined an evaluation metric Averaged Normalized Modified Retrieval Rate (ANMRR) to measure the performance of retrieval[3]. ANMRR is designed to measure performance on both number of correctly retrieved images and how high they are ranked.

Normalized Modified Retrieval Rate (NMRR) is used to measure the performance of each query. NMRR is defined by (8):

$$\text{NMRR}(q) = \frac{\left(\sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}\right) - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - 0.5 * NG(q)} \qquad (8)$$

Where:

NG (q) = The size of the ground truth set for a query image q

Rank(k) = The ranking of the ground truth images retrieved by the retrieval algorithm

K(q) = The "relevance rank" for each query

As the size of the ground truth set is normally unequal, a suitable K (q) is determined by (9):

$$K(q) = \min(4 * NG(q), 2 * GTM) \qquad (9)$$

where, GTM is the maximum of NG(q) for all queries. The NMRR is in the range of (0, 1) and smaller values represent a better retrieval performance. ANMRR is defined as the average NMRR over a range of queries, which is given by (10):

$$\text{ANMRR} = \frac{1}{NQ} \sum_{q=1}^{NQ} \text{NMRR}(q) \qquad (10)$$

where, NQ is number of query images.

**Corel_1k Dataset:** Corel_1k dataset is commonly used in image retrieval researches such as SIMPLIcity[10]. It consists of 1000 images. These images are divided into 10 categories based on semantic concepts. Although each category has its own semantic contents, the visual contents of images in one category could be very different. Ground truth sets of 20 sample query images from different categories are defined by[3].

### RESULTS

Experiments demonstrate that FRCFE has better average performance and it has better results than GLA in the chosen dataset. Table 3 listed the performance comparison of the dominant color descriptor in terms of ANMRR for the query images. Figure 4a and b demonstrate the visual differences for the retrieval results of query #12 "Yellow flower" of the two feature extraction methods. Figure 4 shows that the proposed method could achieve better perceptually relevant image retrieval.

Table 3: Experimental results using Corel_1k dataset

| S No. | Query image Corel_1k query) | NG(q) | K(q) | GLA | FRCFE |
|---|---|---|---|---|---|
| 1 | Beach 1 | 32 | 128 | 0.79300 | 0.3210 |
| 2 | Beach 2 | 29 | 116 | 0.65500 | 0.3970 |
| 3 | Rome | 25 | 100 | 0.32400 | 0.4470 |
| 4 | Statue | 20 | 80 | 0.37000 | 0.4350 |
| 5 | Blue bus | 8 | 32 | 0.48500 | 0.3510 |
| 6 | Yellow bus | 10 | 40 | 0.75200 | 0.2110 |
| 7 | Red bus | 29 | 116 | 0.19200 | 0.1230 |
| 8 | Dino painting 1 | 100 | 200 | 0.00100 | 0.1080 |
| 9 | Dino painting 2 | 100 | 200 | 0.02800 | 0.1110 |
| 10 | Elephant 1 | 19 | 76 | 0.53700 | 0.6360 |
| 11 | Elephant 2 | 20 | 80 | 0.83800 | 0.5720 |
| 12 | Yellow flower | 17 | 68 | 0.61300 | 0.2710 |
| 13 | White flower | 11 | 44 | 0.40600 | 0.1820 |
| 14 | Orange flower | 11 | 44 | 0.64600 | 0.3260 |
| 15 | Red flower | 36 | 144 | 0.64640 | 0.3550 |
| 16 | Purple flower | 26 | 104 | 0.89800 | 0.4680 |
| 17 | Brown horse | 82 | 200 | 0.34500 | 0.2620 |
| 18 | White horse | 47 | 188 | 0.39400 | 0.3290 |
| 19 | Ice mountain | 54 | 200 | 0.67700 | 0.5000 |
| 20 | Mountain and grassland | 26 | 104 | 0.80300 | 0.8590 |
| | ANMRR | | | 0.52017 | 0.3632 |



Fig. 4: The Corel_1k query "Yellow flower" from dataset

Fig. 4a: Retrieval result for the Corel_1k query #12 using DCD with GLA. 5 of 17 ground truth images found in first 20 retrievals



Fig. 4b: Retrieval result for the Corel_1k query #12 using DCD with FRCFE. 11 of 17 ground truth images found in first 20 retrievals

## CONCLUSION

MPEG-7 is considered an important standard for the description of multimedia content. In this study, a new color feature extraction algorithm for the MPEG-7 Dominant Color Descriptor (DCD) is presented. The new technique is called Fixed Representative Colors Feature Extraction Algorithm (FRCFE), this algorithm depends on converting color pixels of an image to its nearest color from a 38 fixed-colors set. After that the highest eight frequent colors with their percentages are selected to represent the DCD for that image. To evaluate the performance of the new technique, a CBIR system is implemented. Comparisons are made with the existing MPEG-7 feature extraction algorithm (the GLA) on the COREL_1k dataset. Experiments showed that the retrieval results using the new algorithm have performance "ANMRR = 0.3632" in most query images.

## REFERENCES

1. Rui, Y., T. Huang and S.F. Chang, 1999. Image retrieval: Current techniques, promising directions and open issues. J. Vis. Commun. Image Represent., 10: 39-62. DOI: 10.1006/JVCI.1999.0413
2. Flicker, M., H. Sawhney, W. Niblack, J. Ashley and Q. Huang *et al*., 1995. Query by image and video content: The QBIC system. IEEE. Comput., 28: 23-32.
   http://portal.acm.org/citation.cfm?id=620306
3. Po, L.M. and K.M. Wong, 2004. MPEG-7 dominant color descriptor using merged palette histogram similarity measure. Proceedings of the IEEE International Symposium on Intelligent Multimedia, Video and Speech Processing 2004, Oct. 20-22, IEEE Xplore Press, Hong Kong, pp: 651-654. DOI: 10.1109/ISIMP.2004.1434148
4. Chang, S.F., T. Sikora and A, Purl, 2001. Overview of the MPEG-7 standard. IEEE. Trans. Circ. Syst. Video Technol., 11: 688-695. DOI: 10.1109/76.927421
5. Sikora, T., 2001. The MPEG-7 visual standard for content description-an overview. IEEE. Trans. Circ. Syst. Video Technol., 11: 696-702. DOI: 10.1109/76.927422
6. Hunter, J., 2001. An overview of the MPEG-7 Description Definition Language (DDL). IEEE. Trans. Circ. Syst. Video Technol., 11: 765-772. DOI: 10.1109/76.927438
7. Gersho, A. and R.M. Grey, 1991. Vector Quantization and Signal Compression. Kluwer, Norwell, Massachusetts, ISBN: 0-7923-9181-0, pp: 732.
8. Wyszecki, G. and W.S. Stiles, 1982. Color Science: Concepts and Methods, Quantitative Data and Formulae. John Wiley and Sons, New York, ISBN: 025263196.
9. Manjunath, B.S., J.R. Ohm, V.V. Vasudevan and A. Yamada, 2001. MPEG-7 color and texture descriptors. IEEE. Trans. Circ. Syst. Video Technol., 11: 703-715.
10. Wang, J. and G. Wiederhold, 2001. SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. IEEE. Trans. Patt. Anal. Mach. Intell., 23: 947-963. DOI: 10.1109/34.955109