Color Content Descriptors of Images by Vector Quantization

Jan MIHALIK, Iveta GLADISOVA

Laboratory of Digital Image Processing and Videocommunications, Department of Electronics and Multimedia Communications, Faculty of Electrical Engineering and Informatics, Technical University of Kosice, Park Komenskeho 13, 041 20 Kosice, Slovak Republic

jan.mihalik@tuke.sk, iveta.gladisova@tuke.sk

DOI: 10.15598/aeee.v18i4.3799

Abstract. *In the paper, we propose color content descriptors of images by vector quantization in RGB color space compared to scalar ones. In order to obtain a much more accurate discrete representation of this space we use our algorithm for optimization of vector quantizers. We introduce several modifications of these descriptors such as global, structural as well as of dominant color. We consider a different number of bins and evaluated the similarity of the color content of images using mean square error of color histograms of a reference image and a searched image. Then, the color content of the image with the minimum error was the most similar to the reference image. We also used a parameter of variance, if the color content of several searched images was very similar based on the mean quadratic error for the structural descriptor. By vector quantization of RGB space we can achieve 2–3 times decreasing the number of bins at the same accuracy.*

Keywords

Color content, color histogram, descriptors, images, vector quantization.

1. Introduction

Nowadays, with ever-increasing bit rates of the Internet, the world is awash with digital photos, images, graphics, audio and video files. This large amount of various data requires a different technique for searching and sorting them compared to the text. It is often necessary to perform various operations, which can be associated either with different attribute features (size, resolution, etc.) or content features (color, texture, motion, etc.) [1]. In most cases, the second group of features requires the most storage space. The solution to this problem is MPEG-7 [2], which consists of standard tools for the description of multimedia content [3]. The main advantage of MPEG-7 lies in the so-called descriptors used to describe individual content characteristics of the medium to search for information or compare similarity in a large amount of multimedia data, such as still pictures, 2D or 3D graphics, audio, video, but also speech or face [4].

In general, the description of multimedia content consists of several descriptors. Descriptors include applicability to large amounts of multimedia data and, therefore, are divided into:

- graphical descriptors (color, texture and shape),
- video descriptors,
- text descriptors,
- other descriptors (e.g. face recognition).

This paper presents mainly the descriptors of the color content of images [5] and [6].

2. Color Content Descriptors of Images

Color is one of the most important visual attributes of an image. The basic structure of a color descriptor representing the color content of an image consists of a color space definition, quantization of that space and displaying of colors [7]. Definition and quantization of the color space is mainly used in combination with displaying the color by color histograms [8]. RGB space is the primary color space for displaying the color content of an image. The representation of the color content in RGB space can be easily converted into other color spaces such as YUV, HSV, HSI and CIE using well-known transformation relations [9].

Different quantization methods can be applied to respective color components in any color space. The number of color components and their value ranges are known from the description of individual color spaces. The quantizer covers the entire range of values of each component, while the quantized values are considered to be normalized to their range width, as e.g. for R, G, B components or for the radial range 2π for component H (Hue). While the identifier determines the type of quantizer, the remaining quantized values at its output are generally different for each type of quantizer. The relevant types of quantizers include uniform and nonuniform scalar quantizer and vector quantizer.

Color histogram is one of the most basic visual characteristics. It provides graphic information about the number of separate colors in an image. The entire RGB color space is divided into specific areas (bins) with subsequent assignment of individual color pixels to these areas [10]. Then, the color histogram can be simply described as color distribution of the image in RGB space and mathematically is expressed as a discrete function:

$$h\left(f_k\right) = n_k, \tag{1}$$

where f_k is k-th color $(k = 0, 1, \ldots, N)$ and n_k is number of color pixels with color f_k in the image. In practical terms, it is useful to recalculate this histogram by dividing the number of color pixels by the total number of pixels L in the image and, thus, eliminate its dependence on the size (raster) of the image. This histogram is dependent on quantization vectors or their corresponding bins [11], since color pixels for the description of the color content of an image using the vector quantizer are determined by its quantization vectors. This image is then normalized with the maximum value of the range of its color components before its calculation in order to eliminate the dependence of this histogram on the light intensity of the image. This way, the color content of each image is generally represented in one normalized RGB space, and the histogram that is calculated subsequently for this space determines the probability of the occurrence of individual colors specified by the vector quantizer used.

The dominant color is an important feature in color segmentation but it can also be used for the description of the color content of the whole image [12]. In principle, this dominant color description is a special case of using a complete histogram of a quantized color image, where only one bin is given. In this case, the dominant color descriptor requires only the bin index of the dominant color. There are many kinds of such descriptors, e.g. a descriptor of the first K dominant colors, a descriptor determining the variance of the dominant color (variance around its quantization vector in the color space). The dominant color descriptor is simpler, but less accurate than the color content descriptor using a full-color histogram.

The global color content descriptor of the image uses its full-color histogram [13]. Then, this description may give the same results for color images with the same color content (with full histogram) but with different spatial color distribution. This disadvantage is eliminated by the structural descriptor of the color content, where the color image is first divided into segments, e.g. rectangular blocks, whose color content is described using separate histograms. As a result, the color content is in general different for each segment of the color image with the same outcome using the global descriptor. It means that the structural descriptor of the image color content enables to distinguish color images more accurately, since it takes into account also spatial distribution of colors. Similarly, it would be possible to create structural descriptors of dominant colors for individual segments.

3. Quantizers

The Uniform Scalar Quantizer (USQ) [14] is the simplest one and requires only one parameter indicating the number of quantization cells for each color component. Assuming a uniform scalar quantization of individual color components but with a different number of quantization cells, this number would be three times greater, because it would have to be entered for each component separately. The design of the USQ itself is relatively simple and lies in calculation of the quantization step, which then determines all its quantization levels.

The separable Non-Uniform Scalar Quantizer (NUSQ) [14] means that the quantization of the color component is independent of the current state of other components. On the other hand, an inseparable NUSQ is suitable if, for specific values of one component, other components are differently quantized (e.g. in HSV color space, for low intensity, it is useless to distinguish between saturation and hue, and for low saturation, various hues are not necessary). Then, the number of quantization levels required to specify NUSQ depends on its separability. The number of quantization levels should be specified for at least each color component, which has to be quantized non-uniform, while for an inseparable NUSQ, the number of quantization levels may be even higher. The design of NUSQ is more complicated than the design of USQ because it is not based on calculation of only one parameter but of all generally different quantization levels.

Vector Quantizer (VQ) is the result of vector generalization of NUSQ [15]. In general, its input is based on the sequence of random vectors. Depending on the region $(O_1, \ldots, O_N)$ of the color space R_3 with the dimension $v = 3$, where the input vector is located, it is assigned one of N possible quantization vectors $\vec{b}_1$ to $\vec{b}_n$.

The mean square value of the quantization distortion per dimension at the VQ output is:

$$\sigma_q^2 = \frac{1}{\nu} \sum_{i=1}^{N} \int_{O_i} \left\| \vec{x} - \vec{b}_i \right\|^2 f(\vec{x}) d\vec{x}, \tag{2}$$

where ν is dimension of the input vector, N is the number of quantization vectors, $f(\vec{x})$ is the joint PDF of random vector $\vec{X}$, where $\vec{x}$ denotes its possible results. Equation (2) shows that σ_q^2 depends on variables: ν, N, $f(\vec{x})$, O_i and $\vec{b}_i$. The mean square value σ_q^2 depends on O_i and $\vec{b}_i$ for the selected ν, N and $f(\vec{x})$. Optimization of VQ means finding such a division of the vector space R_V into regions O_i and quantization vectors $\vec{b}_i$ as to minimize the mean square error σ_q^2 per dimension. It means that the following must be valid:

$$\vec{b}_i = \frac{\int\limits_{O_i} \vec{x} f(\vec{x}) d\vec{x}}{\int\limits_{O_i} f(\vec{x}) d\vec{x}} = E\left(\vec{X} \mid \vec{X} \in O_i\right), \tag{3}$$

where E is the statistical mean value operator. From Eq. (3) it is obvious that $\vec{b}_i$ is conditional mean value of $\vec{x}$ on the condition it is located in the region O_i. If $\vec{x} \in O_i$ the following must be applied:

$$\left\| \vec{x} - \vec{b}_i \right\| < \left\| \vec{x} - \vec{b}_j \right\| \quad \text{for all} \quad j \neq i. \tag{4}$$

By the inequality in Eq. (4), optimum division of the vector space R_V is determined, which is also known as Voronoi (Dirichlet) division.

Then, the VQ optimization algorithms [16] must find such a division of the vector space R_3 into O_i regions and quantization vectors $\vec{b}_i$, so that σ_q^2 is minimum. Subsequently, the optimal vector quantizer is described by the vector function $\vec{Q}\left(\vec{X}\right)$ which is determined by the optimum division of the regions $O_o = \{O_i, i = 1, \ldots, N\}$ of the color space R_3 and by the optimum reproduction alphabet $B_o = \left\{\vec{b}_i, i = 1, \ldots, N\right\}$ (a set of quantization vectors). The multidimensional distribution function of the ergodic sequence of random input vectors is often not known. In this case, VQ optimization is based on the knowledge of a sufficiently long training sequence of the input vectors $\{\vec{x}_t\}$, $t = 1, \ldots, T$, where T is the length of the sequence. Thus, we derive VQ optimization algorithms for an unknown distribution function from VQ optimization algorithms for a known distribution function. The computation becomes simpler and gives results that are very close to those derived from the explicit expression of the distribution function if the training sequence of input vectors is long enough. The above-mentioned VQ optimization algorithms are universally used for VQ with any dimension of vector and number of quantization vectors, where optimal VQ has the highest quantization efficiency. Therefore, it shows that also optimal NUSQ is a special case of optimal VQ with dimension equal to one.

4. Quantization of RGB Color Space

Quantization of a discrete color image is an irreversible operation that transforms a sequence of its pixels with a continuously varying value into a sequence of pixels with a discretely varying value. Generally, the input pixels are random (continuous random variables), then the quantized pixels are also of a random character and represent discrete random variables [17].

When designing and evaluating color content descriptors in RGB space, we used a test set of 19 color images of countries (Fig. 3), which had different sizes, but their dimensions were always divisible by 4. Figure 1(a) shows a point correlation diagram that gives the total number of all pixels from the test set, i.e. 2,802,000 pixels. Figure 1(a) shows that the pixels of the test set of images of the countries are mostly distributed on the diagonal of RGB space.

Each component of RGB space is quantized separately by the uniform scalar quantization of color images. Figure 1(b) shows 64 quantization points of USQ obtained after uniform dividing the ranges of the color components of the normalized RGB space. From comparison of the distribution of these quantization points of USQ in Fig. 1(b) with the point correlation diagram in Fig. 1(a), it is obvious that this distribution is not very well adapted, especially at the edges of RGB coordinates.

The distribution of quantization points of NUSQ in Fig. 1(c) is better adapted to this correlation diagram. This distribution was obtained using our designed and implemented algorithm of optimization of vector quantizers for the unit dimension.

The scalar quantization efficiency of RGB space can be significantly increased by its vector quantization. Due to vector quantization of the discrete color image, it is required to convert this image into a sequence of vectors by grouping three pixels with the same spatial coordinates in RGB images, as shown in Fig. 2. In vector quantization of color images, color pixels in

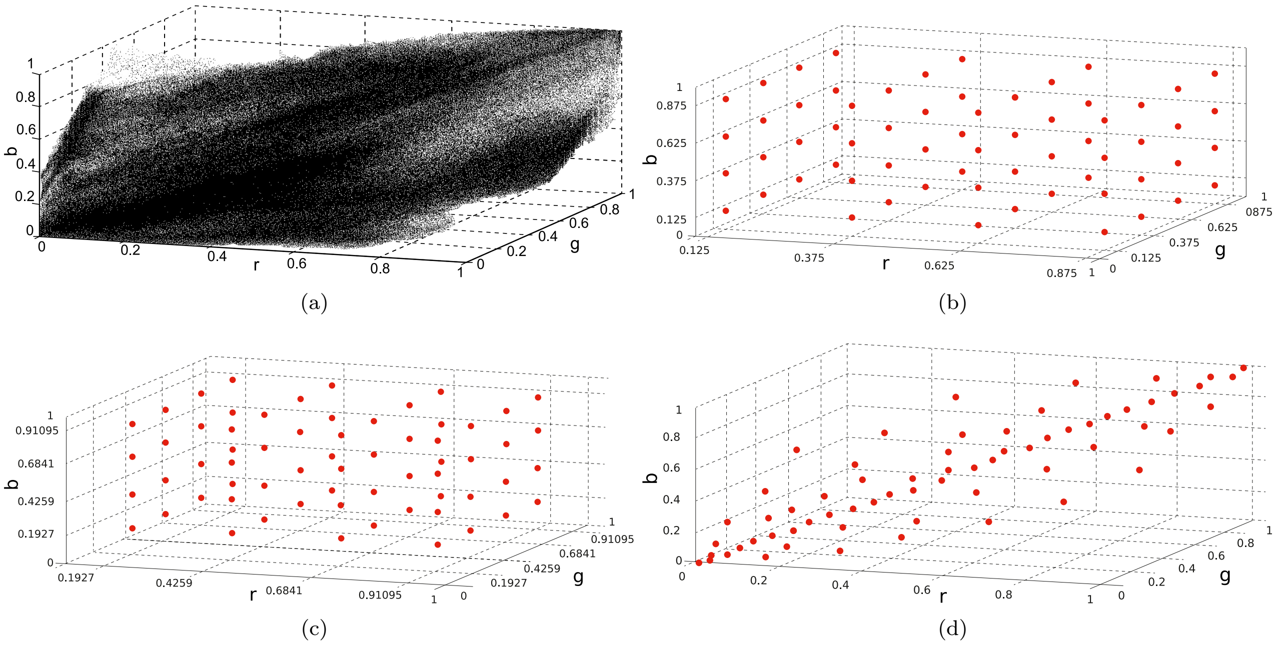


Fig. 1: (a) The point correlation diagram for the test set of 19 color images of countries. Mapping of 64 quantization points in a normalized RGB space for: (b) USQ, (c) NUSQ, (d) VQ.

RGB space are quantized all at the same time and not by components as it was in scalar quantization.

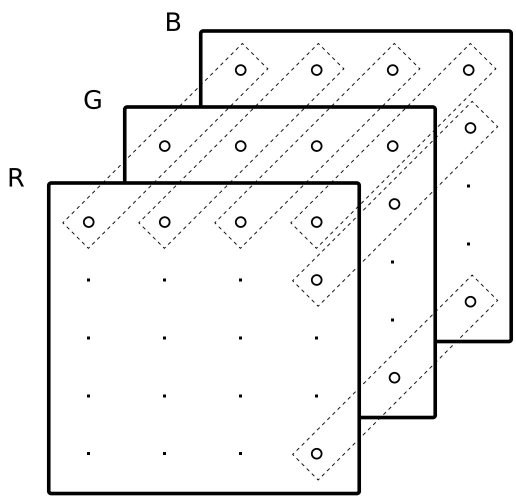


Fig. 2: Process of generating a sequence of three-dimensional vectors of RGB images.

The optimum distribution of 64 quantization points obtained by our VQ optimization algorithm with dimension equal to 3 is given in Fig. 1(d). The figure shows that these quantization points are located in RGB space in places of the highest occurrence of color pixels of the test set of images. In this case, this distribution of quantization points of VQ is best adapted to the point correlation diagram in Fig. 1(a) and represents its most accurate discrete representation in RGB space.

5. Implementation of Color Content Descriptors of Images by Vector Quantization

The testing was performed for the RGB color space using a test set of 19 color images of countries in Fig. 3. The statistical mean square value of quantization noise σ_q^2 was used to objectively evaluate the results obtained for the given types of quantization [18]:

$$\sigma_q^2 = \frac{\sigma_q^2(r) + \sigma_q^2(g) + \sigma_q^2(b)}{3}, \tag{5}$$

and

$$\sigma_q^2(r) = \frac{\sum (r - \tilde{r})^2}{N_r}, \quad \sigma_q^2(g) = \frac{\sum (g - \tilde{g})^2}{N_g}, \quad \sigma_q^2(b) = \frac{\sum \left(b - \tilde{b}\right)^2}{N_b}, \tag{6}$$

where $(r - \tilde{r})$ expresses the difference between the value r of component R and its quantized value of the color pixel of the original image. The same applies to the differences $(g - \tilde{g})$ and $\left(b - \tilde{b}\right)$ of the G and B components of the same color pixel. Variables N_r, N_g and N_b indicate the total number of these components and are the same. Then, the Signal to quantization Noise Ratio (SNR) in decibels is

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{1}{\sigma_q^2} \right). \tag{7}$$

SNR ratios and σ_q^2 values achieved for the test set of 19 color images are given in Tab. 1 and show that the distribution of 64 quantization points in Fig. 1(d) is best for vector quantization compared to uniform and nonuniform scalar quantization of these color images.

Tab. 1: Values σ_q^2 and SNR for the test set of color images.

Type of quantization	σ_q^2	SNR (dB)
vector quantization	0.00001	49.9646
uniform scalar quantization	0.03033	15.1813
non-uniform scalar quantization	0.01173	19.3070

5.1. Global Descriptor

The global descriptor describes the color content of an image as a whole. The image *nature*-1 was considered as a reference image in the test set of 19 color images (Fig. 3). Images with the most similar color content to the reference image were searched using their calculated histograms with a total number of 64 bins as well as for the reduced number of 20 and 10 bins.

The reduction in the number of bins is made not according to their original order, but by the extent of the probability of their occurrence for the reference image *nature*-1. Their original order corresponds to the total number of 64 bins arranged according to the size of the norms of their quantization points (vectors). As the norm increases, the serial number of the bin on the histograms of individual images also increases. The Mean Square Error (MSE) of the histogram of the reference image *nature*-1 and the histogram of the searched image that has similar color content is calculated as follows:

$$\mathrm{MSE} = \frac{1}{64}\left(\mathrm{hist}(A) - \mathrm{hist}(B)\right)^2, \tag{8}$$

where $\mathrm{hist}(A)$ indicates values of the histogram of the reference image *nature*-1 and $\mathrm{hist}(B)$ indicates values of the histogram of the searched image.

As an example, Fig. 4 shows two images *nature*-1 and *nature*-5 with their histograms after vector quantization using 64 bins. The Figure gives a significant deviation in histogram values, which is also reflected in the MSE value for the image *nature*-5 (Tab. 2).

Table 2 gives calculated MSE values for individual test images for the number of bins equal to 64, 20, and 10. It can be stated that the least similar image from the entire test set compared to the reference image *nature*-1 (Fig. 4(a)) is the image *nature*-5 (Fig. 4(b)) with the highest MSE achieved in all three cases of the number of bins. The most similar image with the smallest MSE to the reference image is the image *nature*-11 for all considered numbers of bins.

Table 2 shows that, when calculating values of the mean square error for the reduced number of bins (20 and 10), the MSE values for individual images are higher. These bins have the highest occurrence for the image *nature*-1, i.e. the highest countability. The countability of the selected 20 or 10 bins can be different for other test images. Even the selected bin can have a zero countability, because the same bins selected for the image *nature*-1 are compared. This is also proved by the increased values of the MSE.

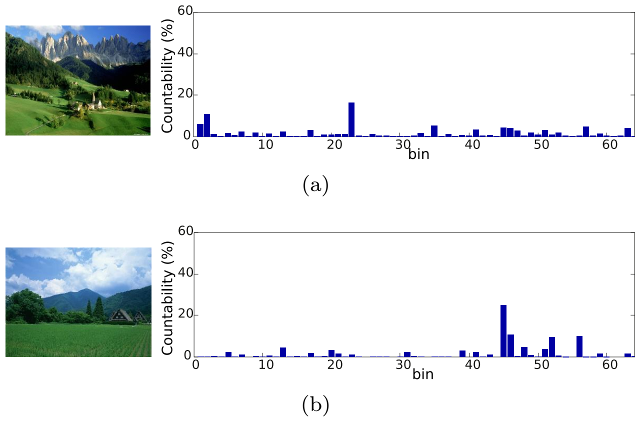


Fig. 4: *Nature*-1 (a) and *nature*-5 (b) images with their histograms after vector quantization.

Tab. 2: MSE of the global descriptor with 64, 20 and 10 bins for the test set of color images.

Image	MSE		
	Number of bins		
	64	20	10
nature-2	4.7295	12.2932	23.4074
nature-3	5.5522	9.1971	15.8606
nature-4	10.2684	24.7529	45.8910
nature-5	**19.1771**	**49.3535**	**96.4084**
nature-6	8.1308	16.0253	31.2921
nature-7	5.4103	16.3453	30.7348
nature-8	7.6334	23.0139	41.6209
nature-9	8.9155	25.7036	49.0282
nature-10	11.0275	39.9112	68.8119
nature-11	**3.4088**	**9.1877**	**14.7189**
nature-12	6.3227	16.3749	29.1901
nature-13	7.5692	21.0389	31.3449
nature-14	4.3369	10.8426	19.4242
nature-15	7.9470	15.7750	27.6932
nature-16	8.5688	24.2469	41.7930
nature-17	7.9559	12.2932	21.3222
nature-18	3.9305	9.1971	20.3473
nature-19	7.2550	24.7529	26.5308

5.2. Structural Descriptor

The description of the color content of images using a structural descriptor is made in blocks. Each image is at first divided into 16 (4×4) blocks of the same size, while their size for individual images may be different depending on their horizontal and vertical dimensions, which must be divisible by 4. For example, for an image of 600×800 pixels when divided into 4×4 blocks, the

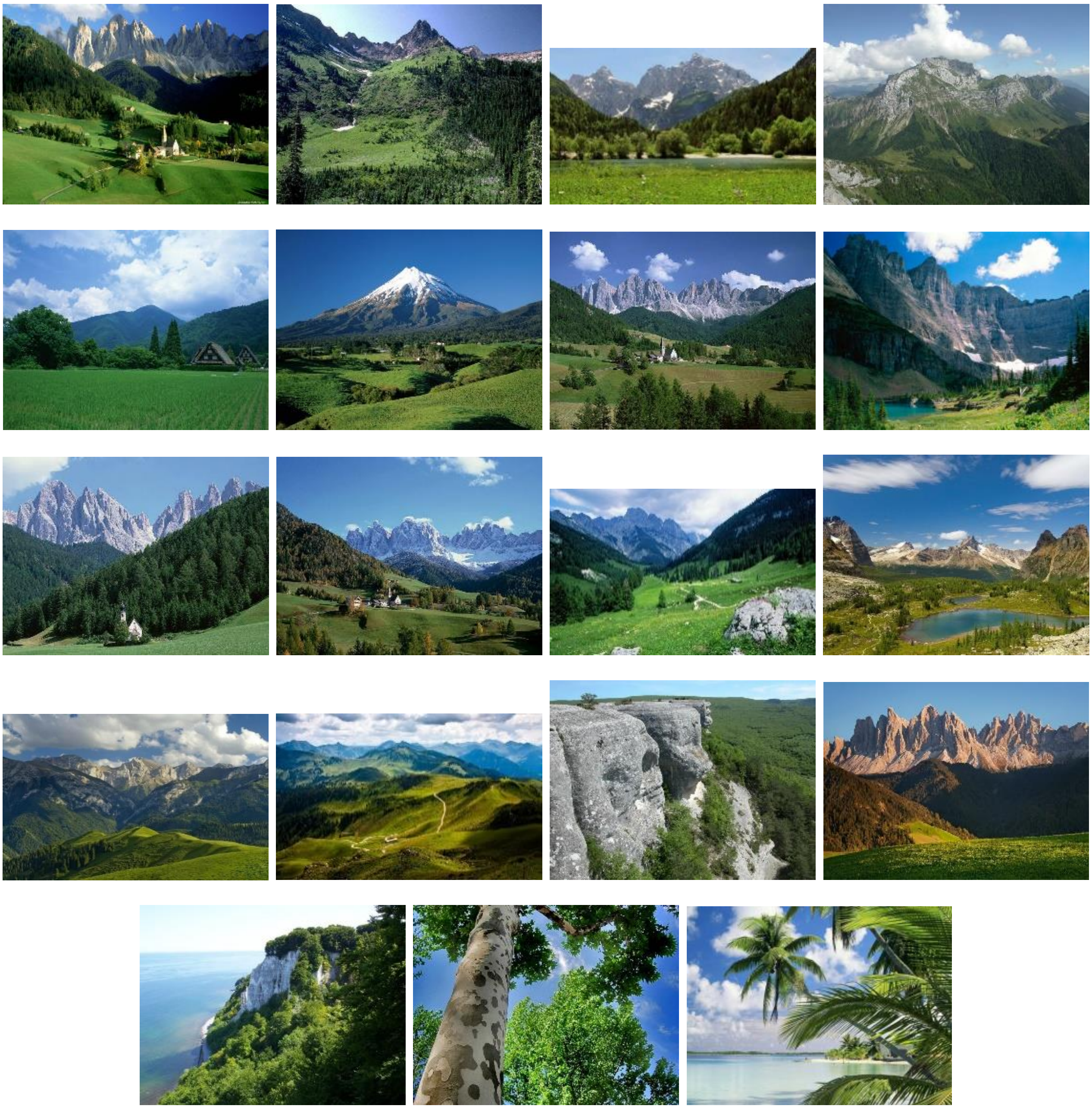


Fig. 3: The test set of 19 color images: *nature-x*, where x is changing from 1 to 19 from left to right and from top to bottom.

block size is 150×200 pixels. The structural descriptor describes their color content using corresponding histograms with a total of 64 bins or a reduced number of 20 and 10 bins, similar to the global descriptor. The criterion of similarity of color content of the corresponding blocks of two different images is MSE of histograms with different numbers of bins for these blocks. This allows to distinguish two images with the same global color content but different color layout. The average Mean Square Error (MSEba) in Tab. 3 is calculated as the average from MSEbi ($i = 1, 2, \ldots, 16$) of each block in the image. The reference image compared to the remaining 18 images in the test set of images is *nature*-1, the same as in the description of the global descriptor. Firstly, the description was made using a structural descriptor with 64 bins, where the smallest MSEba $= 28.1748$ was for the image *nature*-11 and the largest MSEba $= 63.5328$ for the image *nature*-5, i.e. it achieved the lowest degree of similarity with the reference image, which was the same result as in the case of the global descriptor.

A description using the structural descriptor with a reduced number of bins was also performed for the test set of images. For the reference image (*nature*-1), 20 bins with the highest occurrence for all 16 blocks of this image were at first selected and then compared with the corresponding blocks of other tested images. Similar to the description made using the global descriptor, the *nature*-5 image had the largest average MSEba = 149.8374, i.e. the worst-matched color content. The difference was in evaluating the smallest MSEba, which did not correspond to the image *nature*-11 (as was the case of the description with the global descriptor as well as the structural descriptor with 64 bins), but to the image *nature*-12 with MSEba = 55.9513. The same evaluation was achieved using 10 bins, so the largest MSEba = 223.8084 was for the image *nature*-5 and the smallest MSEba = 99.4371 for *nature*-12. Therefore, it can be stated that, when using the structural descriptor with a considerably smaller number of bins, the results of average MSEba were not the same as with the global descriptor despite the fact that with the reduced number of bins, the ones with the highest occurrence in individual blocks of the reference image were selected and compared.

Tab. 3: MSEba of the structural descriptor with 64, 20 and 10 bins for the test set of color images.

Image	MSEba		
	Number of bins		
	64	20	10
nature-2	28.2864	68.7473	117.2718
nature-3	31.6606	77.5587	129.1234
nature-4	41.5519	96.5063	142.7807
nature-5	**63.5328**	**149.8374**	**223.8084**
nature-6	43.3579	66.3779	110.6587
nature-7	32.8773	81.7118	108.3387
nature-8	41.8205	85.8969	146.8312
nature-9	43.4990	108.3950	188.2353
nature-10	49.8167	94.7748	116.9942
nature-11	**28.1748**	69.7911	102.0597
nature-12	35.9202	**55.9513**	**99.4371**
nature-13	42.7183	96.5258	139.7433
nature-14	29.5948	71.7796	114.6542
nature-15	34.4607	84.0031	131.9240
nature-16	46.3036	106.2027	149.0051
nature-17	51.9989	99.1175	135.3881
nature-18	29.6607	106.0526	110.6094
nature-19	40.1305	87.2709	130.2606

The achieved MSE values of individual blocks labeled MSEb1, MSEb2, ..., MSEb16 from the entire test set of color images are given in Tab. 4 for 64 bins. Table 4 shows large differences in the color content of the blocks in the least similar image (*nature*-5) compared to the most similar image (*nature*-11). For example, only in blocks b5 and b16 of the least similar image *nature*-5, MSEbis are lesser than those in the same blocks of the most similar image *nature*-11 and are significantly higher in all others.

Variance σ^2 is an additional quantity of the structural descriptor for evaluation of the similarity of the color content of images:

$$\sigma^2 = \frac{1}{16}\sum_{i=1}^{16}\left(\mathrm{MSEb}i - \mathrm{MSEba}\right)^2, \tag{9}$$

After calculating the values of variance according to Eq. (9), we compared them in the best-matched images with the reference image *nature*-1 in terms of color content and the second best-matched images with different numbers of bins. For 64 bins, the best result with the smallest MSEba = 28.1748 was obtained for the image *nature*-11, whose variance was 241.124. The second was the image *nature*-2 with MSEba = 28.2864 and variance equal to 247.857. Visual evaluation of these results also showed that the color pattern of these two images was similar to the reference one. For 20 bins, the best-matched image to the reference one was the image *nature*-12 with MSEba = 55.9513 and variance equal to 1757.636, the second was *nature*-6 with MSEba = 66.3779 and variance equal to 1831.004. For 10 bins, the best-matched image was also the image *nature*-12 with MSEba = 99.4371 and variance equal to 10030.61, but the second best-matched image was *nature*-11 with MSEba = 102.0597 and variance equal to 6612.508.

The results of the first two comparisons of similarity of the color content for 64 and 20 bins, were also confirmed by the obtained values of variance according to the MSEba. In the last case of comparison for 10 bins, the variance of the second most similar image (*nature*-11) was significantly lower than that of the image of *nature*-12, which, however, had smaller MSEba. From the overall evaluation of the comparison of the color content of images in the test set to the reference image for the structural descriptor, it can be concluded that with the decreasing number of bins, the inaccuracy in finding the most color content-matched image increases.

5.3. Dominant Color Descriptor

The greatest occurrence of the dominant color is in the image with the corresponding histogram of the quantized color image with one most probable bin. The dominant color can be used to describe the entire color image (global description) or for smaller parts of this image - blocks (structural description). For example, for global description, if we compare the dominant colors of the reference image *nature*-1 (Fig. 5(a)) to the image with the most similar color content - *nature*-11 (Fig. 5(b)) and the image with the least similar color content - *nature*-5 (Fig. 5(c)), it is obvious that when using also the dominant color descriptor of these images, the dominant color of *nature*-5 (Fig. 5(c), bottom) is less similar to the dominant color of the reference image *nature*-1 (Fig. 5(a), bottom) than the dominant color of the image *nature*-11 (Fig. 5(b), bottom).

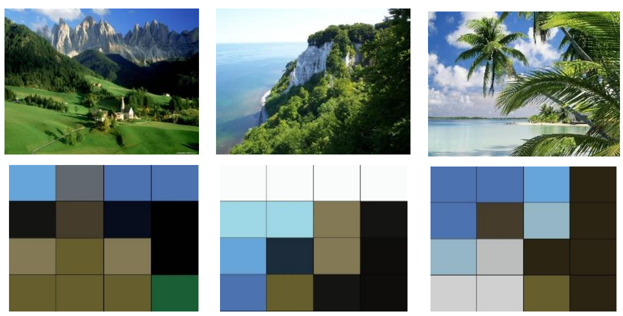
Tab. 4: MSEbi of the individual blocks ($i = 1$ to 16) of structural descriptor with 64 bins for test set of color images.

Block	MSEbi of images *nature*-x (for $x = 2$ until 10)								
	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$	$x = 7$	$x = 8$	$x = 9$	$x = 10$
b1	35.12	66.64	15.47	46.68	153.13	82.07	44.71	10.34	130.18
b2	22.66	14.36	23.82	31.61	65.63	14.54	17.83	31.45	18.84
b3	19.16	24.37	60.17	63.47	40.44	10.43	56.86	68.99	22.88
b4	59.89	1.78	65.262	189.936	43.629	23.056	75.283	70.953	40.712
b5	41.046	70.312	14.509	34.989	86.975	89.953	14.019	62.772	64.111
b6	6.304	12.115	15.496	31.524	18.822	13.371	9.823	16.064	55.833
b7	10.965	4.393	36.243	39.618	6.122	3.400	42.008	25.543	9.555
b8	33.342	2.54	61.179	123.377	7.291	45.164	37.192	50.435	6.09
b9	26.665	49.732	13.94	70.958	44.985	51.708	49.618	76.554	74.93
b10	8.7	22.321	18.718	31.901	19.454	18.471	24.062	15.297	48.763
b11	8.252	4.265	27.713	24.661	3.808	5.978	23.887	18.044	12.562
b12	53.499	10.775	85.969	118.615	34.94	40.051	48.726	62.584	19.52
b13	43.049	67.95	57.471	97.632	38.811	21.009	104.994	74.013	132.146
b14	19.816	54.792	41.443	44.07	55.475	31.016	47.101	35.6	78.702
b15	25.262	25.945	49.219	48.482	25.836	29.552	41.825	33.085	41.787
b16	38.849	74.272	78.203	18.998	48.374	46.259	31.186	44.248	40.448
Block	**MSEbi of images *nature*-x (for $x = 11$ until 19)**								
	$x = 11$	$x = 12$	$x = 13$	$x = 14$	$x = 15$	$x = 16$	$x = 17$	$x = 18$	$x = 19$
b1	14.46	46.51	43.4	19.07	30.18	37.83	114.87	18.08	48.05
b2	26.2	50.16	49.96	27.63	25.81	17.17	65.25	39.8	34.98
b3	53.6	20.43	74.62	46.39	51.87	46.18	77.3	26.81	65.7
b4	10.905	18.281	84.858	51.32	73.95	82.362	60.392	49.747	112.126
b5	41.432	42.696	21.087	17.003	21.712	36.074	137.693	14.367	15.932
b6	25.108	62.808	11.72	2.88	9.794	4.529	26.015	12.993	4.353
b7	12.728	9.651	30.345	11.995	20.04	14.59	13.652	10.272	24.583
b8	27.244	12.143	33.343	23.807	30.175	30.735	18.02	36.824	87.294
b9	53.515	89.942	36.112	28.457	40.31	45.530	65.08	30.024	18.959
b10	7.726	41.48	14.573	20.404	20.223	18.248	14.778	22.955	18.841
b11	15.247	8.949	34.429	39.071	6.556	83.02	3.91	8.422	6.668
b12	29.498	14.205	26.438	19.447	32.1	33.566	58.613	47.608	37.717
b13	59.28	23.387	23.554	33.998	58.372	120.753	46.105	39.059	55.116
b14	22.27	46.198	49.892	46.818	47.325	31.644	38.485	38.739	40.823
b15	30.975	42.964	74.636	22.543	35.678	91.52	34.111	46.086	30.936
b16	20.605	44.92	74.516	62.685	47.267	47.098	57.70	32.787	40.009

(a) (b) (c)

Fig. 5: Images (a) *nature*-1, (b) *nature*-11, (c) *nature*-5 with their dominant colors.

Another comparison was made by detecting block dominant colors using the structural descriptor, as it can be seen in Fig. 6. For this comparison, we selected images with overall image structure clearly different from the reference image *nature*-1 (Fig. 6(a)). Hence, MSE values of the images *nature*-17 (Fig. 6(b)) and *nature*-19 (Fig. 6(c)) had also higher values among all the tested images, whether it was a description with the global or structural descriptor with 64 bins (Tab. 2 and Tab. 3). It is also obvious from the comparison of the block dominant colors of the structural descriptor given below for each image. More thorough examination of the color content by the dominant color can be achieved using the structural descriptor with a larger number of blocks, into which the input color image has to be divided.



(a) (b) (c)

Fig. 6: Images (a) *nature*-1, (b) *nature*-17, (c) *nature*-19 with their block dominant colors using the structural descriptor with 64 bins.

6. Conclusion

In the paper, we first discussed in general terms the color content descriptors of an image for multimedia standard MPEG-7 based on color space quantization and, after that, considered RGB space. There are classical methods of quantization in this space using a uniform and non-uniform scalar quantizer. To increase the accuracy of its discrete representation we applied vector quantizer for its quantization. For this purpose, we introduced a method of generating the input sequence of vectors from the described color image. Then, we designed an algorithm for its optimization, which allowed us to obtain a discrete representation of RGB space with the highest SNR.

Based on the experimental results, we showed that 64 bins were enough for sufficiently accurate description of the color content of an image, which was 2–3 times less than when using scalar quantizers. The global descriptor, being one of the proposed modifications of color content descriptors of an image using vector quantization, describes this content without the requirement for dividing the image region. If it is necessary to do so, then it is better to use the structural descriptor. The simplest is either global or structural descriptor of the dominant color, whose role is only to decide on the dominant color in the image as a whole or in its individual parts. In general, decreasing the number of bins worsens the description accuracy of the color content of an image, and a minimum of 20 bins is required for the descriptors with vector quantization.

References

[1] OHM, J.-R., F. BUNJAMIN, W. LIEBSCH, B. MAKAI, K. MULLER, A. SMOLIC and D. ZIER. A set of visual feature descriptors and their combination in a low-level description scheme. *Signal Processing: Image Communication*. 2000, vol. 16, iss. 1, pp. 157–179. ISSN 0923-5965. DOI: 10.1016/S0923-5965(00)00023-0.

[2] SIKORA, T. The MPEG-7 visual standard for content description-an overview. *IEEE Transactions on Circuits and Systems for Video Technology*. 2001, vol. 11, iss. 6, pp. 696–702. ISSN 1558-2205. DOI: 10.1109/76.927422.

[3] PUJARI, J., S. N. PUSHPALATHA and D. PADMASHREE. Content-Based Image Retrieval using color and shape descriptors. In: *2010 International Conference on Signal and Image Processing*. Chennai: IEEE, 2010, pp. 239–242. ISBN 978-1-4244-8594-9. DOI: 10.1109/ICSIP.2010.5697476.

[4] OVSENIK, L., J. TURAN, T. HUSZANIK, J. ORAVEC, O. KOVAC and M. ORAVEC. Image Encryption Algorithm with Plaintext Related Chaining. *Computing and Informatics*. 2019, vol. 38, iss. 3, pp. 647–678. ISSN 2585-8807. DOI: 10.31577/cai_2019_3_647.

[5] SINGHA, M. and K. HEMACHANDRAN. Content Based Image Retrieval using Color and Texture. *Signal & Image Processing : An International Journal*. 2012, vol. 3, iss. 1, pp. 39–57. ISSN 0976-710X. DOI: 10.5121/sipij.2012.3104.

[6] YASMIN, M., M. SHARIF and M. A. SHAHID. Content Based Image Retrieval Based on Color: A Survey. *International Journal of Advanced Networking and Applications*. 2015, vol. 7, iss. 3, pp. 2724–2735. ISSN 0975-0290.

[7] CIEPLINSKI, L. MPEG-7 Color Descriptors and Their Applications. In: *International Conference on Computer Analysis of Images and Patterns (CAIP)*. Warsaw: Springer, 2001, pp. 11–20. ISBN 978-3-540-44692-7. DOI: 10.1007/3-540-44692-3_3.

[8] BRUNELLI, R. and O. MICH. Histograms analysis for image retrieval. *Pattern Recognition*. 2001, vol. 34, iss. 8, pp. 1625–1637. ISSN 0031-3203. DOI: 10.1016/S0031-3203(00)00054-6.

[9] GEVERS, T., M. A. GIJSENIJ, J. WEIJER and J. M. GEUSEBROEK. *Color in Computer Vision: Fundamentals and Applications*. 1st ed. Hoboken: John Wiley & Sons, 2012. ISBN 978-0-470-89084-4.

[10] SHARMA, N., P. RAWAT and J. SINGH. Efficient CBIR Using Color Histogram Processing. *Signal & Image Processing : An International Journal*. 2011, vol. 2, iss. 1, pp. 94–112. ISSN 0976-710X. DOI: 10.5121/sipij.2011.2108.

[11] JEONG, S., C. S. WON and R. M. GRAY. Image retrieval using color histograms generated by Gauss mixture vector quantization. *Computer Vision and Image Understanding*. 2004, vol. 94, iss. 1–3, pp. 44–66. ISSN 1077-3142. DOI: 10.1016/j.cviu.2003.10.015.

[12] ISLAM, M., D. ZHANG and G. LU. Automatic Categorization of Image Regions Using Dominant Color Based Vector Quantization. In: *2008 Digital Image Computing: Techniques and Applications*. Canberra: IEEE, 2008, pp. 191–198. ISBN 978-0-7695-3456-5. DOI: 10.1109/DICTA.2008.17.

[13] MUSTIKASARI, M., S. MADENDA, E. PRASETYO, D. KERAMI and S. HARMANTO. Content Based Image Retrieval Using Local Color Histogram. *International Journal of Engineering Research*. 2014, vol. 3, iss. 8, pp. 507–511. ISSN 2319-6890. DOI: 10.17950/ijer/v3s8/807.

[14] MIHALIK, J. and R. STEFANISIN. Quantization Algorithms of Standard Videocodecs. *Acta Electrotechnica et Informatica*. 2004, vol. 4, iss. 1, pp. 47–54. ISSN 1335-8243.

[15] GERSHO, A. and R. M. GRAY. *Vector Quantization and Signal Compression*. 1st ed. Norwell: Kluwer Academic Publishers, 1992. ISBN 978-0-7923-9181-4.

[16] MIHALIK, J. Neural Network Clustering Vector Quantizer Design. *Neural Network World*. 1993, vol. 2, iss. 5, pp. 197-208. ISSN 1210-0552.

[17] BARAKBAH, A. R. and Y. KIYOKI. 3D-Color Vector Quantization for Image Retrieval Systems. In: *International Database Forum (iDB)*. Izaka: IPSJ, 2008, pp. 13–18.

[18] STRICKER, M. A. and M. ORENGO. Similarity of color images. In: *IS&T/SPIE's Symposium on Electronic Imaging: Science and Technology*. San Jose: SPIE, 1995, pp. 381–392. ISBN 0-8194-1767-X. DOI: 10.1117/12.205308.

About Authors

Jan MIHALIK graduated from the Technical University in Bratislava in 1976. In 1979, he joined the Faculty of Electrical Engineering and Informatics of Technical University of Kosice, where he received his Ph.D. degree in Radioelectronics in 1985. Currently, he is full professor of electronics and telecommunications and the head of the Laboratory of Digital Image Processing and Videocommunications at the Department of Electronics and Multimedia Telecommunications. His research interests include information theory, image and video coding, digital image and video processing and multimedia videocommunications.

Iveta GLADISOVA received M.Sc. and Ph.D. degrees from the Department of Electronics and Multimedia Telecommunications, Technical University of Kosice in 1984 and 1997. She works at same university currently as Assistant Professor. Her research interests include entropy coding, image segmentation, vector quantization and digital image processing.