INTERNATIONAL STANDARD

ISO/IEC TR 15938-8

First edition
2002-12-15

AMENDMENT 3
2007-12-15

Information technology — Multimedia content description interface —

Part 8: Extraction and use of MPEG-7 descriptions

AMENDMENT 3: Technologies for digital photo management using MPEG-7 visual tools

Technologies de l'information — Interface de description du contenu multimédia —

Partie 8: Extraction et utilisation des descriptions MPEG-7

AMENDEMENT 3: Technologies pour la gestion des photos numériques à l'aide des outils visuels MPEG-7

Reference number
ISO/IEC TR 15938-8:2002/Amd 3:2007(E)

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

In exceptional circumstances, the joint technical committee may propose the publication of a Technical Report of one of the following types:

— type 1, when the required support cannot be obtained for the publication of an International Standard, despite repeated efforts;

— type 2, when the subject is still under technical development or where for any other reason there is the future but not immediate possibility of an agreement on an International Standard;

— type 3, when the joint technical committee has collected data of a different kind from that which is normally published as an International Standard ("state of the art", for example).

Technical Reports of types 1 and 2 are subject to review within three years of publication, to decide whether they can be transformed into International Standards. Technical Reports of type 3 do not necessarily have to be reviewed until the data they provide are considered to be no longer valid or useful.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 3 to ISO/IEC TR 15938-8:2002 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia Information*.

Information technology — Multimedia content description interface —

Part 8: Extraction and use of MPEG-7 descriptions

AMENDMENT 3: Technologies for digital photo management using MPEG-7 visual tools

Add after subclause 4.2.3.3:

4.2.3.4 Dominant Color Temperature

4.2.3.4.1 General

This subclause provides an advanced use scenario of the Dominant Color descriptor. The Dominant Color Temperature is a variation of Dominant Color, but suitable to implement perceptual similarity based retrieval. Images usually have one of a few dominant color temperatures perceived by users when they look at them. Dominant Color Temperatures enable users to search for images in scenarios such as query by example or query by value, and for image browsing regarding their color temperature. It can be useful for users who want to find images which look similar according to color temperature rather than to find images which have similar color regions.

4.2.3.4.2 Use scenario

Dominant Color Temperatures can be used in query by example and query by value search scenarios. Examples of such queries are depicted in Figure AMD3.1. In a query by example a user inputs an example image or draws a colored sketch (query by sketch) and the search application returns the most similar images regarding their color temperature. In a query by value a user chooses a temperature value, and the system retrieves images in which the appearance of color temperature is closest to the user choice.

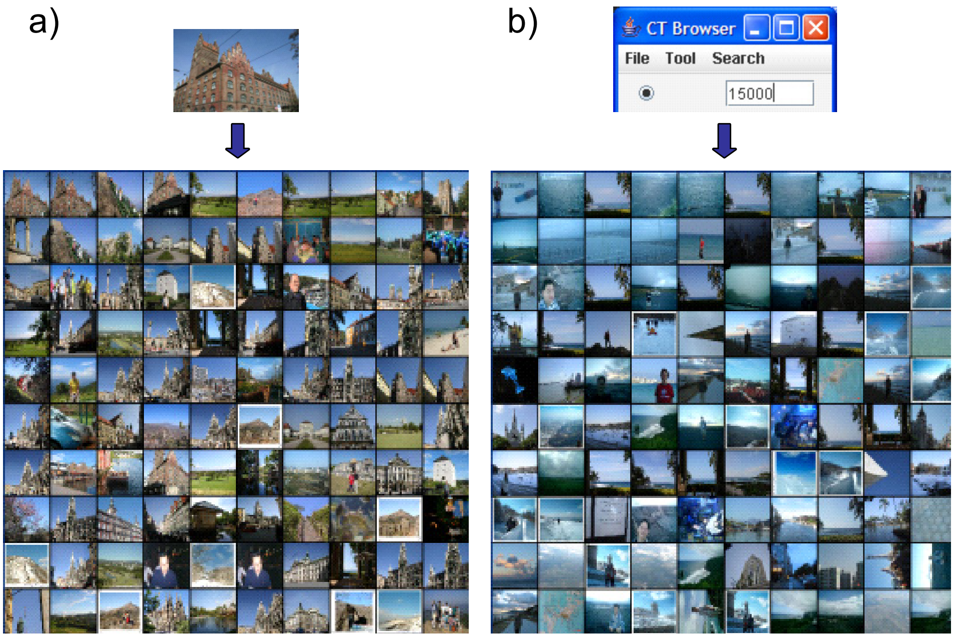

Figure AMD3.1 — Examples of image retrieval using Dominant Color Temperatures: a) query by example; b) query by color temperature value given in kelvins

4.2.3.4.3 Feature extraction

The Dominant Color Temperature, which consists of a maximum of eight pairs of color temperature and percentage, is obtained by the following steps.

1. Get RGB color values and percentages of dominant colors from a Dominant Color descriptor instance.

2. Convert each dominant color value from RGB to color temperature using the relevant method specified in the feature extraction method of Color Temperature descriptor [subclause 6.9.1.1]. The number of obtained color temperatures cannot, therefore, exceed the number of dominant colors in the Dominant Color descriptor instance. The colors that do not have significant color temperature (colors having luminance values below the luminance threshold specified in the extraction method of the Color Temperature descriptor) should be omitted.

3. Use the obtained color temperatures and their percentages given by the Dominant Color descriptor instance in queries: query by example, query by color temperature value, ranking search results, and others.

4.2.3.4.4 Similarity matching

The similarity is based on a distance function which is defined as an integral of absolute difference between two percentage distributions of dominant color temperature. The percentage distributions of dominant color temperature should be obtained first in the following steps:

1. Convert color temperature values T_i of Dominant Color Temperature description to Reciprocal Megakelvin scale RT_i [MK^{-1}] = 1000000/T_i [K].

2. Sort, in ascending order, the dominant color temperatures expressed in reciprocal scale.

3. Create the percentage distribution of dominant color temperature $D_i(RT_i)$ using the following equations:

$D(RT) = 0$ for $RT < RT_0$;

$D(RT) = p_0 + p_1 + ... + p_{i-1}$ for $RT_{i-1} \le RT < RT_i$, $1 \le i \le n-1$;

D(RT) = $p_0 + p_1 + \ldots + p_{n-1}$ for $RT \geq RT_{n-1}$;

Where:

n – number of dominant color temperatures;

RT_0 , RT_1 , ... , RT_{n-1} – sorted dominant color temperatures;

p_0 , p_1 , ... , p_{n-1} – percentages.

Figure AMD3.2 shows an example of a dominant color temperature distribution.

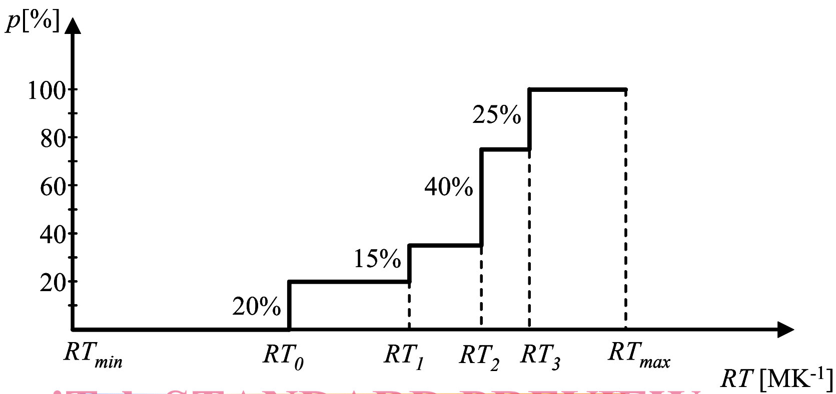

Figure AMD3.2 — Example of cumulative dominant color temperature distribution

The proposed distance function is given by the following equation, which is an integral of difference between two color temperature distributions.

$$dist = \int_{RT_{\min}}^{RT_{\max}} \left| D_1(RT) - D_2(RT) \right| dRT$$

This expression is equivalent to the geometrical area bounded by the two distributions. An example of distance calculation is depicted in Figure AMD3.3, where the distribution distances are shown graphically on distribution diagrams.

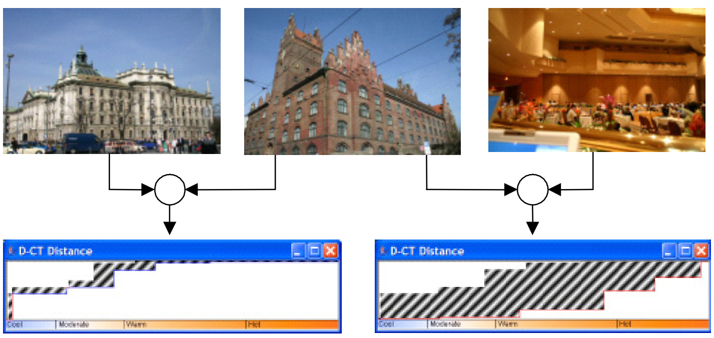

Figure AMD3.3 — Example of distance calculation

The distance function presented in the above equation can be efficiently implemented using the following steps:

1. Input: two percentage distributions of dominant color temperature:

RT1, *D1* – tables of temperature and percentage distribution for image 1,

RT2, *D2* – tables of temperature and percentage distribution for image 2;

2. Initialize: $dist$=0, $x_1 = RT_{min}$;

3. Take the next minimum temperature value t_{curr} from tables *RT1*, *RT2*, and let $x_2 = t_{curr}$;

4. Find in *D1*, *D2* the lower bound y_1 and the upper bound y_2 of the rectangle corresponding to the current x_1, x_2 coordinates;

5. $dist = dist + (x_2 – x_1)(y_2 – y_1)$;

6. $x_1 = x_2$;

7. If all values from tables *D1*, *D2* have been taken then return *dist* else go to step 3.

The tables used as an input to the algorithm above are obtained from the percentage distributions of dominant color temperature in the following way: $RTX[i] = RT_i$, $DX[i]=D(RT_i)$, for $0 \le i \le n$, where *X* stands for image 1 or 2.

In the case of query by color temperature value, the same distance function can be used, by assuming that the query value given by the user is a single dominant color temperature with a percentage of 100%. Although in this case, the distance function can be simplified to the following:

$$\Delta RT = \sum_{i=0}^{n-1} \left| RT_i - RT_{REF} \right| p_i$$

where RT_{REF} is the value of the query color temperature, RT_i are dominant color temperatures, p_i are percentages, and n is the number of dominant color temperatures in image.

4.2.3.4.5 Condition of usage

The same restrictions are applied as for the Dominant Color descriptor. Additionally, Dominant Color Temperatures cannot be used for very dark images of which all dominant colors have luminance values below the luminance threshold specified in the extraction method of Color Temperature descriptor.

Add after subclause 4.7:

4.8 High-level use scenarios

4.8.1 Content based Image retrieval

4.8.1.1 General

Content-based image retrieval gives an efficient and easy way of managing and retrieving digital images from enormous digital contents. In content-based image retrieval, there are two representative methods. One is a query by example, where a user selects a similar image to those expected for a query. The other is query by sketch, in which a user must draw a sketch and use it as a query. Since a seed picture is needed, some mechanism to assist users finding query image itself is required in the former scenario. One possible solution is to combine text-based image retrieval or query by sketch as a pre-processing step of query by example.

4.8.1.2 Query within region of interest (ROI)

4.8.1.2.1 General

This section provides a usage scenario to enable users to dynamically retrieve photographs with similar Region of Interest (such as background) in image space. Region-based image retrieval can be implemented by portioning an image into several small regions and assigning a StillRegionFeatureDS for each of them. However, in practice, such an approach is difficult as it requires prior segmentations that are often subjective and may depend on a particular query. The ROI-based photo retrieval gives users the benefit of defining ROI when making a query. Although query by example is very useful for image retrieval, one may want to retrieve photos with similar backgrounds. In other words, if the scenery is well known or quite beautiful, people tend to take pictures with the same background but different persons. For those photos, it will be more efficient to retrieve the photos by matching the background regions only. In this scenario, the user can select the region that he wants to retrieve in particular and send it to the system as a query.

4.8.1.2.2 Use Scenario

Figure AMD3.4 shows the flow of the proposed query method. The user first selects a query image. In the query image, the user selects a ROI by selecting local regions (shown in blue). The ROI is used as a query image for retrieval. Figure AMD3.5 shows the example of image retrieval within a ROI.

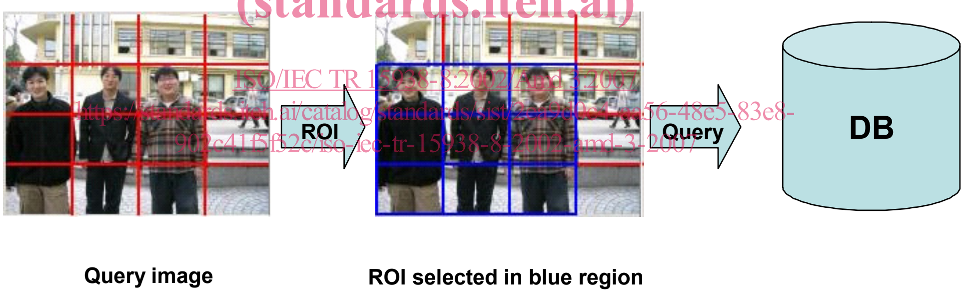

Figure AMD3.4 — Flow of query by ROI

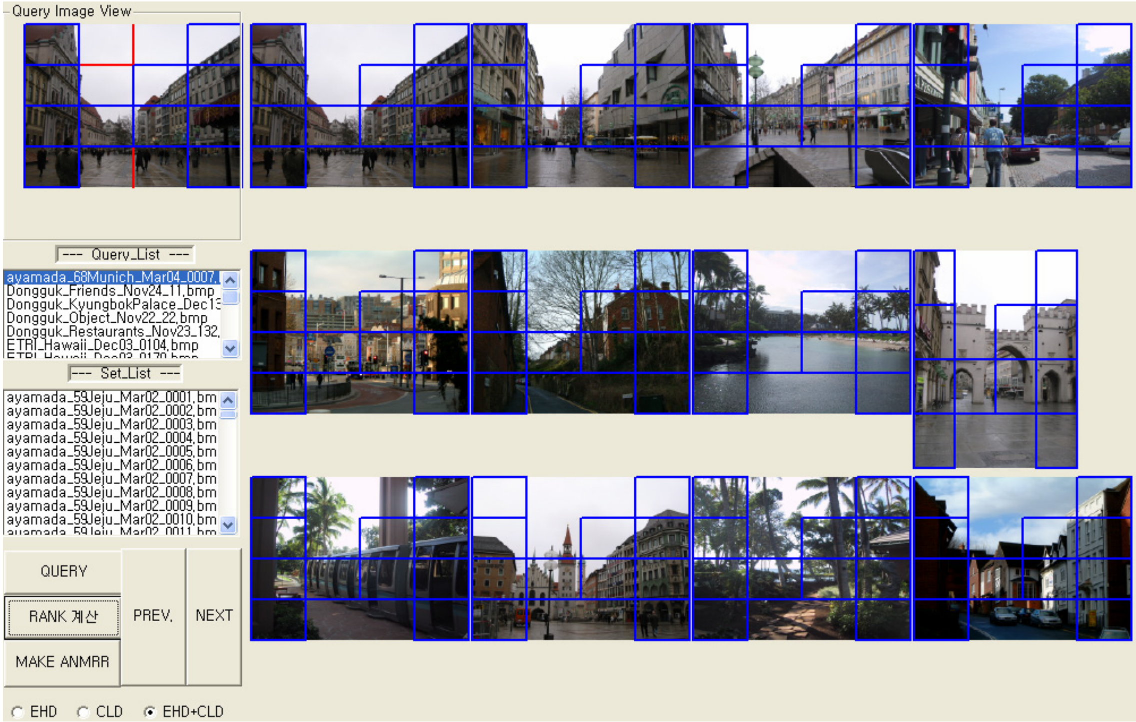

Figure AMD3.5 — Retrieval of image within ROI

4.8.1.2.3 Tools to be used

StillRegionFeatureDS or VideoSegmentFeatureDS is used for this scenario. Among the several elements included in these DSs, the Edge Histogram descriptor and the Color Layout descriptor should be instantiated to implement the functionality of ROI-based retrieval. For video retrieval, shots are extracted from the video sequence and for each shot, localized features from the specific region are extracted. Then, the ROI is used as a query for video retrieval.

4.8.1.2.4 Feature Extraction

ROI-based retrieval can be implemented by extracting localized features from the specified region. The extraction process of the localized feature from the instances of two mandatory description tools, Color Layout and Edge Histogram, is described in this subclause. Figure AMD3.6 illustrates this process. From the Edge Histogram Descriptor one can obtain a localized edge distribution in each 4 x 4 local rectangular region. From the Color Layout descriptor, one can obtain an 8 x 8 region-based DCT: by performing inverse quantization and taking the 8 x 8 inverse DCT (as described in subclause 4.2.5.2.3), we can obtain average color values for 8 x 8 local rectangular regions. Feature extraction of each descriptor is defined in ISO/IEC 15938-3, MPEG-7 Visual. As in Figure AMD3.6, a combination of the Edge Histogram Descriptor and Color Layout Descriptors can be used for the rectangular region-based query-by-ROI.

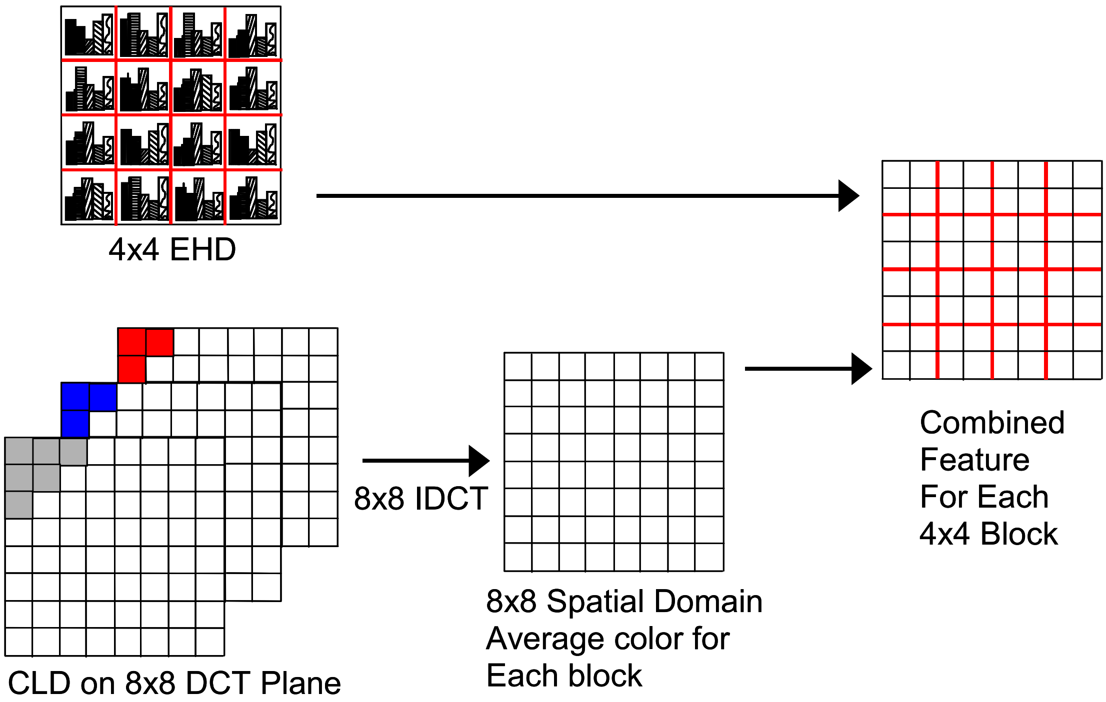

Figure AMD3.6 — 4x4 block-based “Query-by-ROI” with Edge Histogram Descriptor and Color Layout Descriptor

4.8.1.2.5 Similarity Matching

For the Color Layout descriptor, we can take an 8 x 8 inverse DCT for the quantized DCT coefficients of Y, Cr, and Cb. Then, we have representative color values for 8 x 8 blocks of the image. These block-wise color values are combined with the edge histogram bins for each 4 x 4 image region (see Figure AMD3.7). Thus, each rectangular image region of the (4 x 4) Edge Histogram descriptor blocks includes 4 (2 x 2) color blocks obtained by the inverse 8 x 8 DCT of the Color Layout descriptor. Now, a combination of the color and edge information in each of the (4 x 4) rectangular image regions will form a feature vector for the rectangular region-based similarity matching.

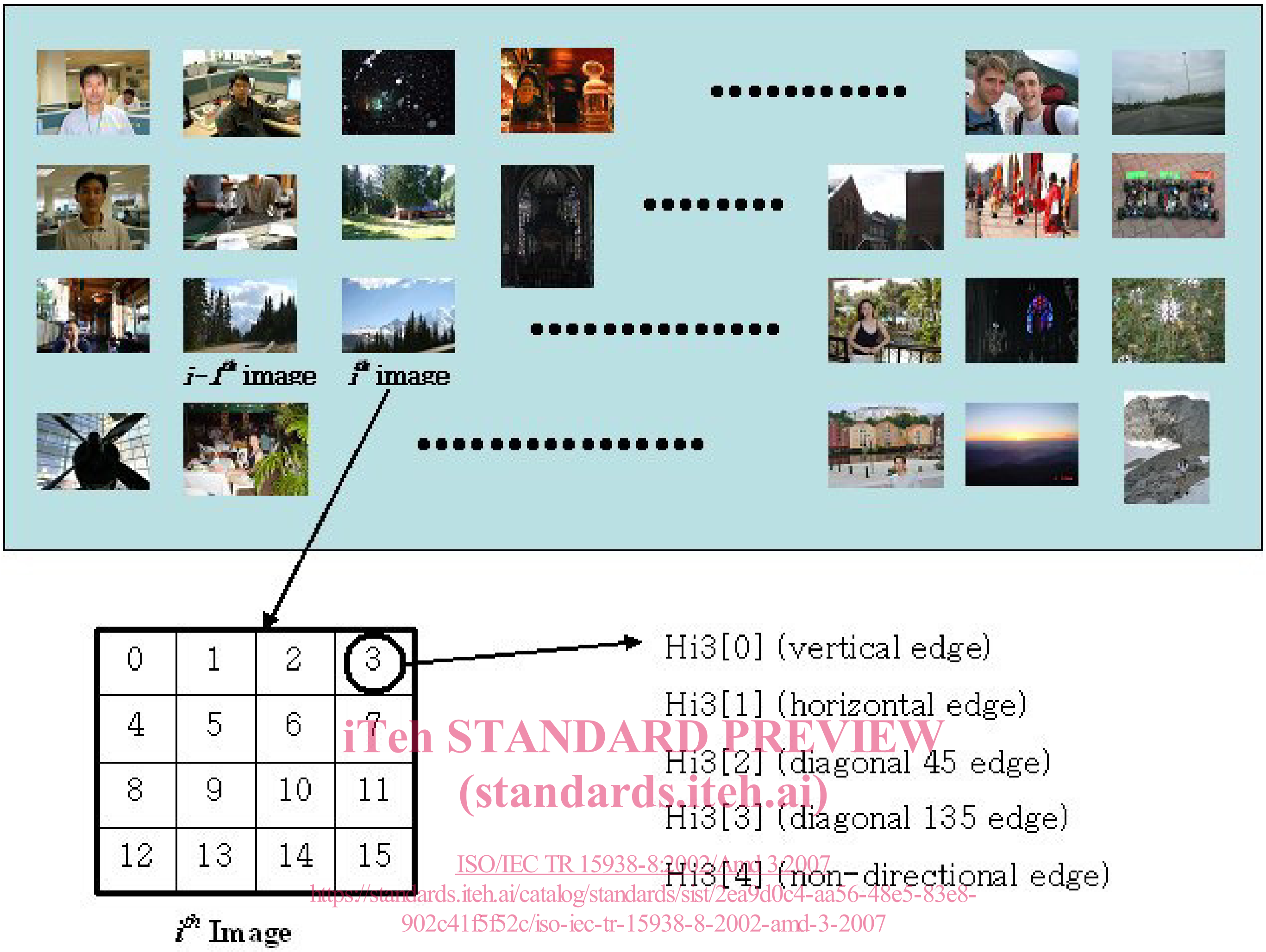

Figure AMD3.7 — Parameter value example of EHD

Figure AMD3.7 shows an example of parameter values when using the EHD for matching blocks. When the total number of images is N, the *j*th(*j*=0,1,2 ... 15) block of the *i*th(i=0,1,2,...N) image has five types(0°, 45°, 90°, 135°, non-directional) of edge value. If we represent these edge value as *k* (*k*=0,1,2,3,4), the parameter value $H_{ij}[k]$ is the *k*th edge value of the *j*th block of the *i*th image. For a query image Q, the edge value of the selected sub-image is $H^{Q}[k]$.The local distance of Edge Histogram $LD^{EHD}{}_{ij}$ is as follows.

$$\mathrm{LD}^{\mathrm{EHD}}{}_{\mathrm{ij}} = \sum_{k=0}^{4} | H^{Q}[k] - H_{ij}[k] | \qquad \text{(AMD1)}$$

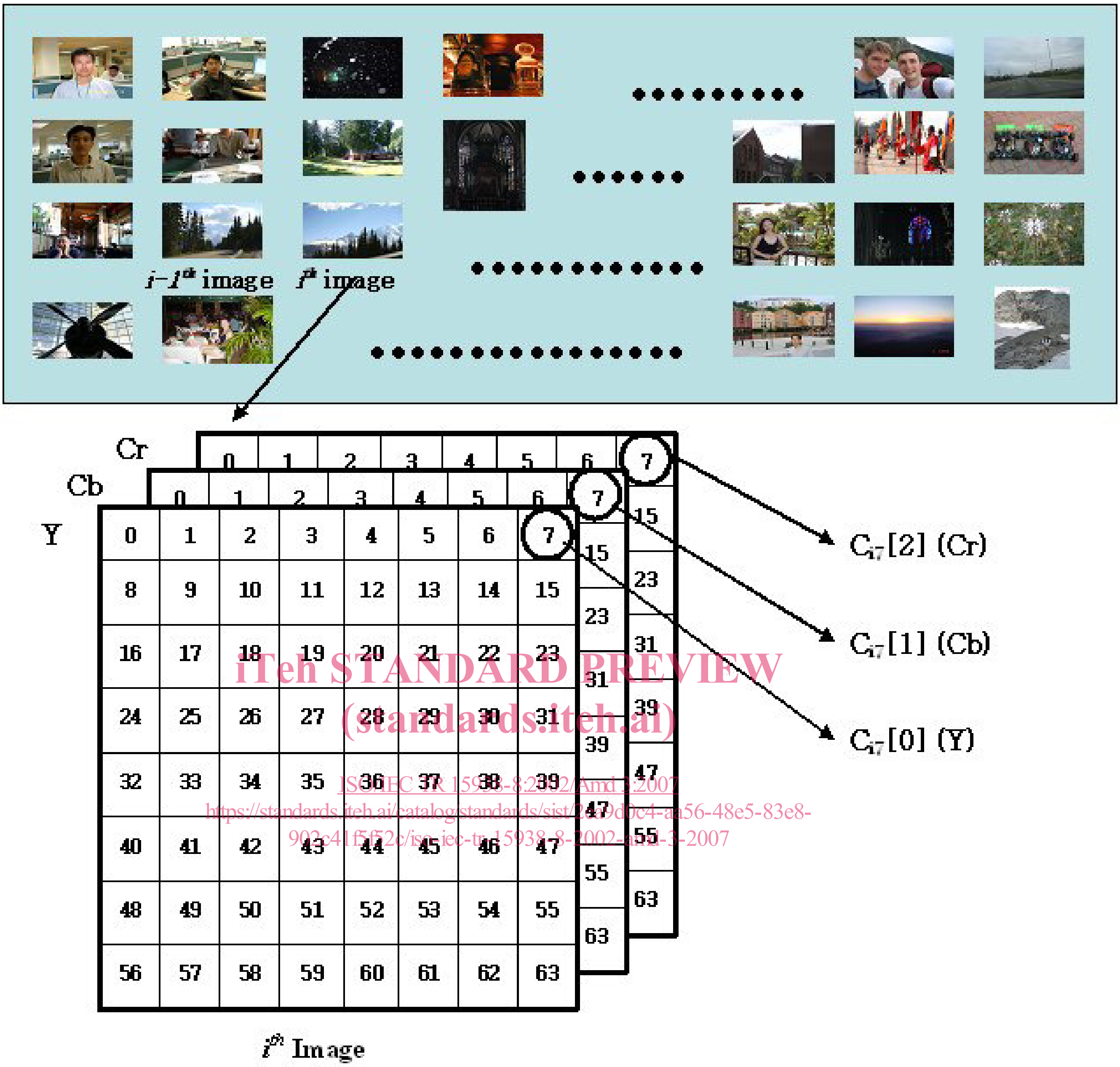

Figure AMD3.8 — Parameter of Inverse DCT Color Layout descriptor

Figure AMD3.8 shows the parameters of the inverse DCT Color Layout descriptor. The total number of images is N. We group 8 X 8 image blocks into 4 X 4 blocks. Newly grouped blocks can be labeled as β where each block consists of 4 blocks(β=0,1,2,3). Each Y, Cb, Cr is labeled as α(α=0,1,2) (α=0 for Y, α=1 for Cb, α=2 for Cr). Parameter value is $C_{ij\beta}[\alpha]$ for *j*th block(sub image) of *i*th image. $C_{ij\beta}[\alpha]$ represents color value α of β sub-block of *j*th block of *i*th image. $C^{Q}_{\beta}[\alpha]$ are parameter values of the query image *Q*. The local distance of Color Layout descriptor of *j*th sub image of *i*th image can be obtained as follows.

$$\mathrm{LD}^{\mathrm{CLD}}{}_{\mathrm{ij}} = \sum_{\alpha=0}^{2}\sum_{\beta=0}^{3}\frac{\left|C_{\beta}{}^{Q}[\alpha] - C_{ij\,\beta}[\alpha]\right|}{3} \qquad \text{(AMD2)}$$