MPEG 7 Color Structure Descriptor
for visual information retrieval project VizIR[1]

Adis Buturovic

Institute for Software Technology and Interactive Systems
Technical University Vienna
Favoritenstrasse 9-11/188/2, A-1040 Vienna, Austria

e0125423@student.tuwien.ac.at

ABSTRACT

The goal of the project is to implement the Color Structure Descriptor (further CSD) for MPEG 7 visual information retrieval project VizIR. Main functionality of this descriptor is image-to-image matching. It expresses local color structure in an image by use of a structuring element that is comprised of several image samples.

1. INTRODUCTION

The global integration of information systems with the ability of easy creation and digitization of visual content have risen the problem of how these vast amounts of data in collections or databases are managed. One of the crucial success factors of all approaches to solve this problem is apparently the implementation of effective but still easy to handle retrieval methods. Visual image retrieval is a rather new approach to overcome the problems by deriving features (or: descriptors; like color histograms, etc.) from the visual content and comparing visual objects by measuring the distance of features with distance functions [1]. In this part of the VizIR[1] project one of such descriptors – the *Color Structure Descriptor* is implemented in Java. The main objective of this document is to give an overview of *CSD* and describe the implemented *Color Structure Descriptor* as part of VizIR. For a complete technical description of MPEG 7 descriptors, the interested reader is referred to [3],[4].

2. COLOR STRUCTURE DESCRIPTOR

Color Structure Descriptor is mostly used for still image retrieval. It expresses local color structure in an image by use of a structuring element as shown on Figure 1. The CSD is computed by visiting all location in the image, retrieving colors C{0-7} of all pixels contained in the 8x8 pixel structure element overlaid on each location, and incrementing the CSD bins assigned to color *Cm*.

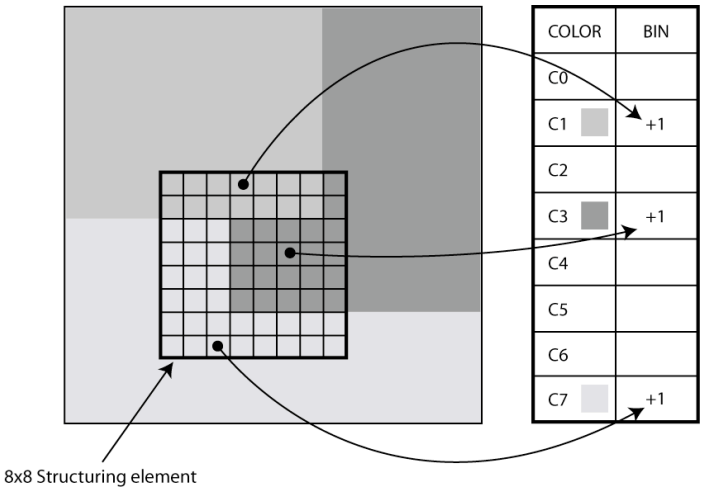

Figure 1. – CSD structuring element

The number of structuring elements is always 64 and the distance between the structuring points increases with the image size, as shown on Figure 2. This method is equivalent to subsampling the image by the power of 2 and then using a structuring element of 8x8 pixels. Due to referenced paper [2] the following simple rule determines the spatial extent of the structuring element given the image size:

$$p = \max\{0, \text{round}(0.5 * \log_2(\text{height} * \text{width}) - 8)\}$$
$$K = 2^p,\ E = 8 * K$$

where:

height, width	image height and width
E × E	spatial extent of the structuring element
K	subsampling factor.

[1] VizIR Project – http://vizir.ims.tuwien.ac.at (last visited on 2005-01-12)

As shown in the Figure 2. for images smaller then 320 x 240 pixels an 8x8 structure element with no subsampling is used, and for image size 640 x 480 (p = 1, K = 2, and E = 16) structuring element is 16x16 and subsampling is 2x2. The structuring element of size 8x8 is applied to a subsampled image.

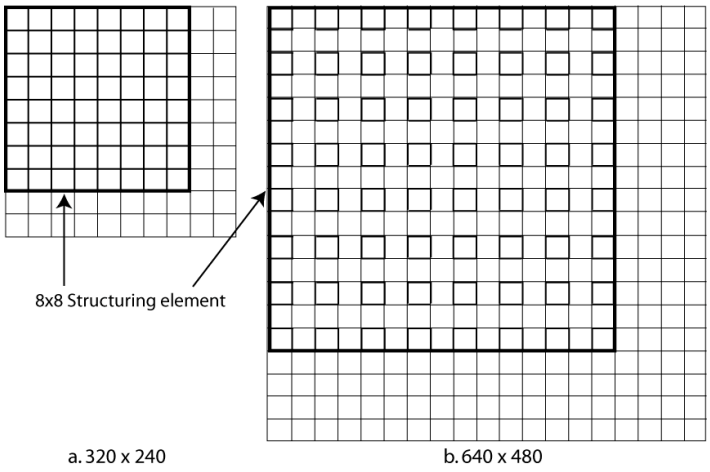

Figure 2. Structuring elements for images with different size

This descriptor is using HMMD color space, and is defined using four color space quantization operating points: 256, 128, 64 and 32 bins. The quantization, shown on Figure 3. is doing in following steps. First, the HMMD color space is divided into five subspaces (0, 1, 2, 3, and 4). The subspace division is preformed on the HMMD diff value by following intervals: [0, 6], [6, 20], [20, 60], [60,110], and [110, 255].

Component	Subspace	Number of quantization levels for different numbers of histogramm bins			
		256	128	64	32
Hue	0	1	1	1	1
	1	4	4	4	4
	2	16	8		
	3	16	8	8	4
	4				
Sum	0	32	16	8	8
	1	8	4	4	4
	2	4			
	3	4	4	2	1
	4			1	

Figure 3. – HMMD color space quantization for CSD

Second each color subspace is uniformly quantized along the Hue and Sum axes where the number of quantization levels is shown on Figure 3 (for both hue and sum parameter).
In the final step the values that contains the Color Structure Histogram are normalized to the 8-bit code value, so that the range of the histogram is converted from [0, 1] values to 0, 255.

3. Class Description and Usage

org.vizir.ColorStructureDescriptor.java - Contains all necessery methods for feature extraction from media content using color structure descriptor:

- extractFeature()
  loads the media content, convert it to HMMD color space and executes the CSD extraction and quantization.
- HMMDColorStuctureExtraction()
  builds the Color Structure Histogram trough described steps in the CSD chapter.
- reQuantization()
  Non-uniform quantization to 8-bit code values.
- RGB2HMMD()
  converts the RGB pixel into HMMD color space.
- setDescriptionFromString()
  verifies if the input String is made by valid Color Structure Descriptor.
- getDescriptionAsString()
  creates a XML string with the results created by feature extraction. The output can also be done by saving the result to a XML file with name specified with setXMLfileName().
- setQuantizationLevels()
  choose between four quantisation bins: 32, 64, 128 and 256
- setXMLfileName()
  if output to XML file is activated, you can set the filename of the output file.

src.test.org.vizir.CSDtest.java – This test class runs the Color Structure Descriptor and saves results to XML iso-8859-1 string and / or XML file.

- testExtractFeature()
  verify the result created by extractFeature().
- testDescriptionAsString()
  verify the result created by getDescriptionAsString().

Known bugs: in MediaFrame: MediaFrame.getPixelAt if source image width is larger then image height.

Usage example:

```
URL url = new URL("testimage.jpg");
MediaContent newMedia = new MediaContent(url);
ColorStructureDescriptor.setQuantizationLevels(64);
ColorStructureDescriptor CSD = new ColorStructureDescriptor();
CSD.extractFeature(newMedia);
String result = CSD.getDescriptionAsString();
```

After chosing the image, and number of quantization bins you can run extractFeature() methode. The result will be stored into a string.

4. References

[1] H. Eidenberger, C. Breiteneder, "VizIR – a framework for visual information retrieval "– jvlc2003.pdf

[2] B. S. Manjunath, "Color and Texture Descriptors" IEEE Transactions on circuits and systems for video technology, vol. 11, no.6, June 2001

[3] Text of ISO/IEC 15938-3 Multimedia Content Description Interface - Part 3: Visual. Final Committee Draft, ISO/IEC/JTC/SC29/WG11, Doc. N4062, Mar. 2001.

[4] MPEG-7 Visual Experimentation Model (XM), Version 10.0, ISO/IEC/JTC1/SC29/WG11, Doc. N4063, Mar. 2001.