Neapolis University

HEPHAESTUS Repository **http://hephaestus.nup.ac.cy**

School of Information Sciences Articles

2011

Fast Retrieval from Image Databases via Binary Haar Wavelet Transform on the Color and Edge Directivity Descriptor

Chatzichristofis, Savvas A.

IARIA

http://hdl.handle.net/11728/10171

Downloaded from HEPHAESTUS Repository, Neapolis University institutional repository

Fast Retrieval from Image Databases via Binary Haar Wavelet Transform on the Color and Edge Directivity Descriptor

Savvas A. Chatzichristofis[1] Yiannis S. Boutalis[1,2] Avi Arampatzis[1]
[1]Department of Electrical & Computer Engineering,
Democritus University of Thrace, Xanthi, Greece
[2]Department of Electrical, Electronic and Communication Engineering,
Chair of Automatic Control, University of Erlangen-Nuremberg, Germany
{schatzic,ybout,avi}@ee.duth.gr

***Abstract*—In this paper, we are evaluating several accelerating techniques for content-based image retrieval, suitable for the Color and Edge Directivity Descriptor (CEDD). To date, the experimental results presented in the literature have shown that the CEDD achieves high rates of successful retrieval in benchmark image databases. Although its storage requirements are minimal, only 54 bytes per image, the time required for retrieval may be practically too long when searching on large databases. The proposed technique utilizes the Binary Haar Wavelet Transform in order to extract from the CEDD a smaller and more efficient descriptor, with a size of less than 2 bytes per image, speeding up retrieval from large image databases. This descriptor describes the CEDD, but not necessarily the image from which it is extracted. The effectiveness of the proposed method is demonstrated through experiments performed on several known benchmarking databases.**

***Keywords*-CEDD; Binary Haar Wavelet Transform; Content-Based Image Retrieval;**

I. Introduction

As the use of computers, internet and cameras is getting more popular, efficient content-based image retrieval is more essential than ever. Any technology that, in principle, helps to organize digital image archives by their visual content is defined as content-based image retrieval (CBIR). By this definition, anything ranging from an image similarity function to a robust image annotation engine falls under the purview of CBIR [1].

Online image repositories such as Flickr contain hundreds of millions of images and are growing quickly [2]. The requirements of modern retrieval systems are not limited to providing good retrieval results, but extend to their ability for quick results. The majority of internet users would compromise the partial reduction of result accuracy in order to save time from searching.

Image retrieval, as well as text retrieval, may be described by the similarity search paradigm [3]. Efficient approaches that allow application on generic similarity search problems still need to be investigated [4]. A promising direction to address this issue is the *approximate* similarity search paradigm [5], [6], [7], [8]. Approximate similarity search provides an improvement in search performance at the price of some imprecision in the results. An interesting approach of approximate similarity search was proposed in [4]. The idea at the basis of this technique is that when two objects are very close to each other they 'see' the world around them in the same way.

In order to achieve image retrieval from large databases, the representation of images by Latent Dirichlet Allocation (LDA) [9] models is studied in [2]. Image representations are learned in an unsupervised fashion, and each image is modeled as a mixture of its depicted topics or object parts.

The present paper proposes a different approach for searching in large databases. First of all, in order to ensure quality of the results, the Color and Edge Directivity Descriptor (CEDD), proposed in [10], [11], is utilized. The size of this descriptor is 54 bytes/image. Subsequently, the Binary Haar Wavelet Transform [12] is used for the extraction of a second descriptor, we call Binary CEDD (B-CEDD). This second descriptor is employed during the retrieval procedure instead of the image. In this way, reduced retrieval times are achieved. A preliminary version of this work has been presented in [13]. Details concerning the CEDD and the Binary Haar Wavelet Transform are given in Sections II and III, respectively.

In order to shape B-CEDD, we follow and evaluate three different approaches. First, we consider CEDD as a single vector and apply the Binary Haar Wavelet Transform up to 15 coefficients. Second, we consider CEDD as a result of early fusion of two independent vectors, one capturing color and the other texture information. Also in this case, the resulting compact descriptor consists of 15 coefficients. In final third approach, we again consider CEDD as a result of early fusion of 6 independent vectors and apply Binary Haar Wavelet Transform to each of the vectors separately. The length of the resulting descriptor is now 18 coefficients. In Section IV, we will describe these three approaches in detail.

During the search process, an image query is entered and the system returns images with a similar content. Initially, the similarity/distance between the query and each image in the database is calculated with the proposed descriptor, and

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

only if the distance is smaller than a predefined threshold, the comparison of their CEDDs is performed. The entire retrieval procedure is described in Section V. In order to estimate the appropriate threshold value, efficient techniques, described in Section VI, were used. The experimental results are presented in Section VII, and the conclusions of this study are drawn in Section VIII.

II. The Color and Edge Directivity Descriptor

The descriptors, which include more than one features in a compact histogram, can be regarded that they belong to the family of Compact Composite Descriptors. A typical example of CCD is the CEDD descriptor. The structure of CEDD consists of 6 texture areas. In particular, each texture area is separated into 24 sub regions, with each sub region describing a color. CEDD's color information results from 2 fuzzy systems that map the colors of the image in a 24-color custom palette. To extract texture information, CEDD uses a fuzzy version of the five digital filters proposed by the MPEG-7 EHD [14], [15]. The CEDD extraction procedure is outlined as follows: when an image block (rectangular part of the image) interacts with the system that extracts a CCD, this section of the image simultaneously goes across 2 units. The first unit, the color unit, classifies the image block into one of the 24 shades used by the system. Let the classification be in the color $m, m \in [0, 23]$. The second unit, the texture unit, classifies this section of the image in the texture area $a, a \in [0, 5]$. The image block is classified in the bin $a \times 24 + m$. The process is repeated for all the image blocks of the image. On the completion of the process, the histogram is normalized within the interval [0,1] and quantized for binary representation in a three bits per bin quantization.

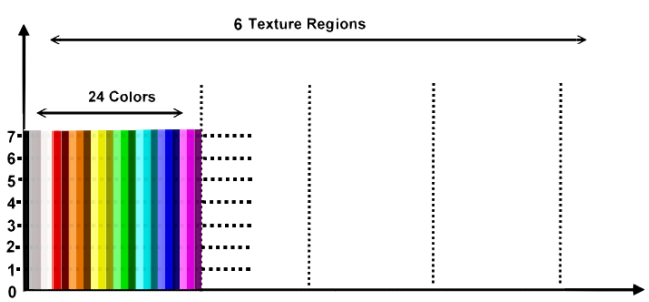

Figure 1. The structure of the CEDD.

The most important attribute of CEDDs is the achievement of very good results that they bring up in various known benchmarking image databases. Table 1 shows the ANMRR [14] results in 3 image databases, along with those obtained by MPEG-7 descriptors. The ANMRR ranges from '0' to '1', and the smaller the value of this measure is, the better the matching quality of the query. ANMRR is the evaluation criterion used in all of the MPEG-7 color core experiments. Evidence shows that the ANMRR measure approximately coincides linearly with the results of subjective evaluation of search engine retrieval accuracy. More details on the ANMRR are given in section VII. The ANMRR values for the MPEG-7 descriptors in WANG's [16] database as well as the ground truths that were used are available at [17]. Since MPEG-7 descriptor results are not available for the UCID [18] and NISTER [19] databases, an implementation of CLD, SCD and EHD in img(Rummager) [20] and LIRe Demo [21] retrieval systems is used. Details regarding the experimental results, the implementation of the MPEG-7 descriptors, as well as the ground truths that were used, are available online.

Another important attribute of CEDD is its small size requirements for indexing images. The CEDD length is 54 bytes per image.

Table I
ANMRR results in three benchmark image databases.

Descriptor	WANG	UCID	NISTER
CCD			
CEDD	**0.25283**	**0.28234**	**0.11297**
MPEG-7			
DCD MPHSM	0.39460	-	-
DCD QHDM	0.54680	-	-
SCD	0.35520	0.46665	0.36365
CLD	0.40000	0.43216	0.2292
CSD	0.32460	-	-
EHD	0.50890	0.46061	0.3332
HTD	0.70540	-	-

The img(Rummager) and LIRe Demo retrieval systems use these descriptors to create index files from which they carry out the search. Img(Rummager) makes XML-type index files, while LIRe utilizes a Lucene index to store the descriptors.

III. Binary Haar Wavelet Transform

The Binary Haar Wavelet Transform coefficients of the histogram are calculated with the use of following Haar Wavelet Transform [12]:

$$\psi(x) = \begin{cases} 1 & 0 \leq x < 0.5 \\ -1 & 0.5 \leq x < 1 \\ 0 & else \end{cases} \quad (1)$$

Figure 2 shows the four basis functions of the Haar wavelet of length eight. The Haar wavelet coefficients are obtained by taking the inner product of the basis functions with the given histogram. This transformation is very fast as it does not involve multiplications.

The Haar coefficients capture the qualitative aspects of the histogram [22]. For example, the second coefficient (from the basis function 2 in Figure 2) is positive if the sum of the left half of the histogram bins is greater than the right half and negative otherwise. Similarly, the third coefficient is

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

positive if the sum of the first quarter of the histogram bins is greater than the second quarter and negative otherwise. In the Binary Haar descriptor, each of these coefficients is quantized to '0' or '1', depending on whether its value is negative or positive, hence a binary representation is obtained.

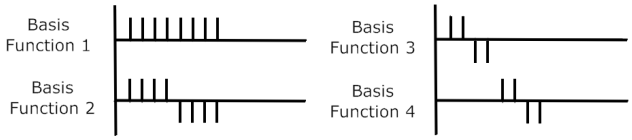

Figure 2. Four basis functions of the Haar wavelet of length eight.

At the first level, the k bins of the histogram are divided into two halves. If the sum of the histogram values in the left half is greater than the sum of the histogram values in the right half then the second bit of descriptor is '1', while is '0' otherwise. Note that the first coefficient corresponds to the sum of all probabilities in a histogram and it is always positive. Therefore is quantized to 1 and is not used in similarity matching.

At the second level, the $k/2$ bins of each half of the histogram are divided into two halves. If the sum of the histogram values in the first half is greater than the sum of the histogram values in the second half then the second bit of descriptor is '1' else it is '0'. Similar, if the sum of the third half is greater than the sum in the fourth half, then the third bit of descriptor is '1' else it is '0'. This is repeated recursively for the third and the fourth level.

IV. Binary CEDD (B-CEDD)

In order to describe the contents of CEDD with another descriptor significantly smaller in storage needs, we followed 3 different approaches with the use the Binary Haar Wavelet Transform.

A. Approach 1: B-CEDD$_1$

The first approach considers CEDD as a 144-dimensional vector, in which we apply 4-level Binary Haar Wavelet Transform. In the first level, the 144 elements are split in two groups of 72 elements. If the sum of the coefficients of the first group is greater or equal to the corresponding sum of the second group, then the first output coefficient of the transform is 1; otherwise, it is 0.

In the second level, the 144 elements are split in groups of 36 elements, and compared in consequent pairs using the same method as in the first level. The output of the transform consists of two coefficients. The third level of the transform compares consequent pairs of 18 elements, producing another 4 output coefficients.

Finally, the fourth level compares consequent pairs of 9 elements, producing another 8 coefficients. Overall, the output vector of the transform consists of 15 coefficients. We chose to apply the transform four times in order to produce an output vector of 2 bytes. In the rest of this paper, the result obtained from the application of the Binary Haar Transform on the CEDD descriptor will be referred as Binary CEDD with index 1 (B-CEDD$_1$).

B. Approach 2: B-CEDD$_2$

The second approach considers CEDD as the results of early fusion of two independent vectors; a vector capturing the color information of the image described by the CEDD, and a vector describing the texture information. To apply the Binary Haar Wavet Transform, we work as follows: First, we isolate the color and texture information from the descriptor in two independent vectors, the CEDD_Color vector consisting of 24 elements and the CEDD_Texture vector consisting of 6 elements. The separation is straightforward since each information item is distinctively placed in the descriptor. The separation pseudocode is the following:

```
for (int i = 0; i < 6; i++)
  {
   for (int j = 0; j < 24; j++)
      {
          CEDD_Color [j] += CEDD[24 * i + j];
      }
   }
for (int i = 0; i < 6; i++)
  {
   for (int j = 0; j < 24; j++)
    {
     CEDD_Texture [i] += CEDD[24 * i + j];
    }
   }
```

The Binary Haar Transform is applied on the CEDD_Color vector up to the third level resulting in 7 coefficients (1 coefficient from the first transformation level, 2 coefficients from the second transformation level, and 3 from the third level).

Regarding the CEDD_Texture vector, given that the resulting 2 halves include 3 elements, the problem arising is that the transform may be applied only once. In order to overcome this constraint, whenever this is met during the transform application, we propose the following solution: During the first transform level, the 6 elements are divided in 2 triads. To apply the second level, the middle element of each triad is cloned. The 2 identical elements replace the original element from which they came from. In this way, each triad is replaced by a quartet of elements, which now the transform may be applied on. In the third level of the Binary Haar Transform the cloned elements are removed and the transform is applied directly on the vector, comparing this time the elements per pair. On the whole, 8 elements are created (1 from the first transform level, 4 from the second level and 3 from the third). The complete extraction procedure of the Modified Binary Haar Wavelet Transform from CEDD_Texture is illustrated in Figure 3.

At the end of the procedure, the 2 resulting vectors are placed consecutively. The size of the produced descriptor is

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

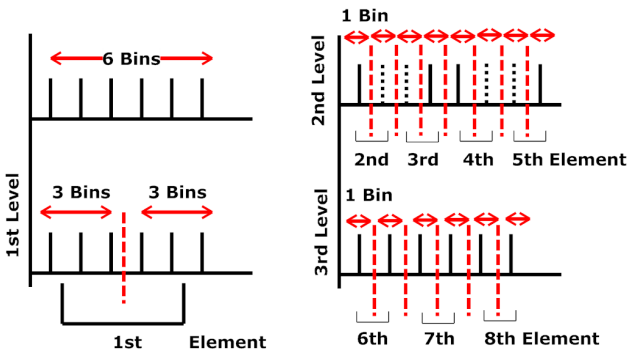

Figure 3. Extraction procedure of the Modified Binary Haar Wavelet Transform from CEDD_Texture.

limited to 15 binary bins and its storage requirements are much smaller than 2 bytes per image (15 bits). In the rest of this paper, this descriptor will be referred to as Binary CEDD with index 2 (B-CEDD$_2$).

C. Approach 3: B-CEDD$_3$

The third approach splits CEDD to 6 independent vectors in 24 dimensions. Each vector corresponds to one of the 6 texture areas used by the descriptor. The transform is applied twice to each of the vectors, splitting the vectors to groups of 12 and then to 6. From each vector, 3 coefficients are produced. We produce B-CEDD$_3$ by taking consequently the 18 output coefficients ($3 \times 6 = 18$). This approach results to the largest descriptor size and is closer to the essence of Compact Composite Descriptors where CEDD belongs to, but as we will see in Section VII it does not lead to the best end-results.

V. System Overview

One of the most important attributes of the Binary CEDD is that it is extracted directly from the CEDD with no intervention of the described image. This results in its immediate extraction from the already existent index files.

The search procedure based on the use of the 2 descriptors, CEDD and B-CEDD, is illustrated in Figure 4. The user enters a query image in the system. From this image, both the CEDD and the B-CEDD descriptors are extracted. The system uses an image database in which the indices are described by both descriptors. During data retrieval from databases, the length of the retrieved information is of great significance [23]. For this reason, in a first phase the system retrieves only the B-CEDD descriptor, which, due to its small storage requirements as well as its small length, is retrieved practically in an instance.

For each database image, the distance between the B-CEDD descriptor with the corresponding B-CEDD descriptor of the query image is calculated by a simple X-OR gate. In the case of 2 identical descriptors, the X-OR output is equal to 0, while in the worst scenario the obtained distance is 15 (equal to the B-CEDD length). The logic gate X-OR requires much less computing resources than the Tanimoto coefficient. The Tanimoto coefficient is used to calculate the distance between CEDD descriptors.

If the distance of the 2 descriptors is found to be smaller than a threshold T, then the CEDD descriptor is retrieved from the database and its distance from the corresponding CEDD descriptor of the query image is calculated. The procedure is repeated for all the database images.

After the completion of the procedure, the classified results are shown to the user in ascending order of the distance obtained during the CEDD descriptors comparison.

The most important issue that should be taken into consideration during the aforementioned procedure is the determination of the threshold T. This threshold defines whether an image is potentially similar to the query image. If it is, then the retrieval of the image's CEDD is requested and the process for its comparison with the corresponding CEDD of the query image is activated. The T value investigation is described in detail in the following section.

VI. Investigating the T Value

In order to determine an appropriate T value, we work as follows: We choose 35 images from the Wang database and regard them as historical queries. The historical queries idea comes from the text retrieval area and has been used to normalize retrieval scores of documents in cases of fusion and distributed information retrieval [24], [25].

For each one of the historical queries, searching is performed in the database from which they originate. In particular, the distances between the B-CEDD descriptors of each historical query and each image of the database are calculated and a ranking list, in which the database images are ordered according to their distance from the historical query, is obtained. The procedure is repeated for all the historical queries. At the end of the process, 35 ranking lists emerge. Since the Wang database includes 1000 images, 35000 values (35 ranking list $\times$ 1000 images) are finally obtained. By plotting these values the distance distribution is obtained. As depicted in Figure 5, which shows score distributions from the three B-CEDD approaches, the first two approaches produce similar distributions.

Subsequently, the set of these 35000 values for each approach, enters in a Gustafson Kessel fuzzy classifier [26]. The Gustafson Kessel is an extension of the Fuzzy C-Means algorithm. The Gustafson Kessel parameters are selected as: Clusters=4, Repetitions=3000, $e = 0.001$ and $m = 2$.

The four resulting classes were used to form a single input fuzzy system for each approach. The fuzzy system includes four membership functions which are labeled as: 'Low', 'Medium', 'High' and 'Highest' The centers of the classes, as they result from Gustafson Kessel classifier, correspond to the tops of the membership functions. Given that the score distribution of the B-CEDD$_1$ and B-CEDD$_2$ are similar to

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

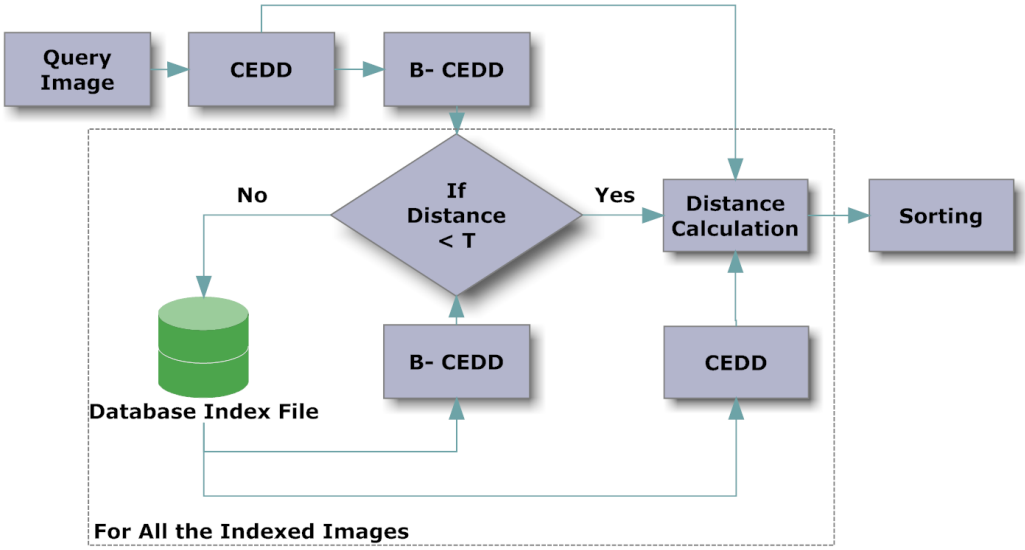

Figure 4. Searching using the 2 descriptors.

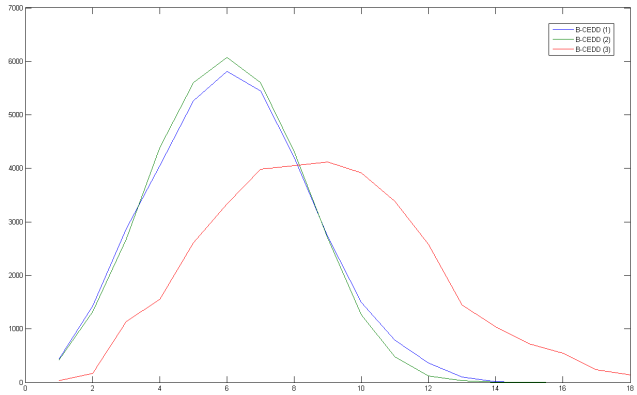

Figure 5. Distribution of the B-CEDD distances for 35 historical queries.

each other, the fuzzy systems that come from the Gustafson Kessel classifier output are identical.

The fuzzy system that was shaped operates as follows: The system gets as input the distance between the B-CEDD descriptors of the query image and any other image. The vertical axis shows the distance that may be obtained during the comparison of the 2 B-CEDD descriptors, while the horizontal axis shows the activation degree for the membership function of each class. Consider, for instance, that the distance between 2 B-CEDD descriptors was found to be equal to 4 (see Figure 6). This value activates both the first and the second membership function by 0.5.

For the simplest scenario of the model usage we should specify that if the 'Low' membership function is activated with a value greater than the activation value of any other membership function, then the image under study is likely to be visually similar to the query image. Thus, the CEDD descriptor should be also retrieved in order to perform the comparison.

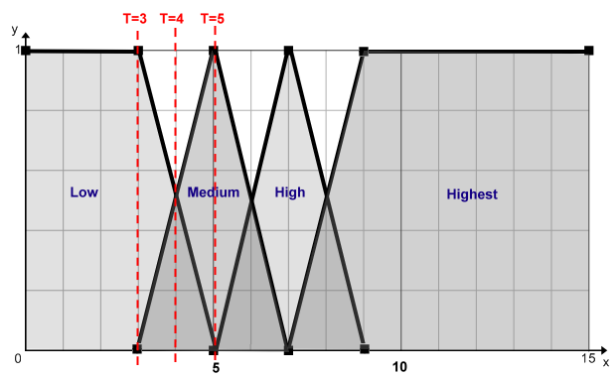

Figure 6. Outcome of the historical queries distances distribution classification in 4 classes.

Similarly, if the 'Low' activation degree is smaller than that of another membership function, the image is discarded. In the next section, the experimental results of a threshold tuning attempt are presented.

VII. Experimental Results

To the best of our knowledge, there is no large scale image database with ground truths data available that could be used for the performance evaluation of the proposed method. For this reason, experiments have been performed on two known small scale benchmarking databases.

For the performance evaluation the following measures were used:

1) Recall at n, where n is the number of the retrieved through the proposed method images. This measure

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

presents the number of relevant documents retrieved by a search divided by the total number of existing relevant documents.

2) ANMRR at n. The ANMRR ranges from 0 to 1. The smaller the value of this measure is, the better the retrieval quality. This measure captures both precision and recall in one value.

The ANMRR computation requires the average rank computation. The average rank AVR(q) for query q is:

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \tag{2}$$

where

- $NG(q)$ is the number of ground truth images for the query q.
- K is the top ranked retrievals examined where:

$$K = min(X \times NG(q), 2 \times GMT) \tag{3}$$

$$GMT = max\,(NG\,(q)) \tag{4}$$

If $NG(q) > 50$ then $X = 2$ else $X = 4$.

$Rank(k)$ is the retrieval rank of the ground truth image. For a query q, suppose that as a retrieval result the kth ground truth image is found at a position R. If this image is in the first K results then $Rank(k) = R$ else $Rank(k) = K + 1$.

The modified retrieval rank is:

$$MRR(q) = AVR(q) - 0.5 - 0.5 \times N(q) \tag{5}$$

The normalized modified retrieval rank is computed as follows:

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 \times N(q)} \tag{6}$$

The average of NMRR over all queries is defined as:

$$ANMRR(q) = \frac{1}{Q} \sum_{q=1}^{Q} NMRR(q) \tag{7}$$

The proposed method is implemented in the retrieval system img(Rummager) [20]. For better time measurements, each experiment was repeated 10 times. All the experiments were performed on an Intel Core 2 Quad Q9550 @2.83GHz PC with 3GB of RAM.

As elaborated in the previous section, the necessity for accurate estimation of the T threshold value which defines whether the CEDD descriptor of an image should be retrieved, is imperative. In order to determine the appropriate threshold, we experiment with the following scenarios: Initially, for each of the approaches of B-CEDD, we consider that the two images may be potentially similar (and calculate the distance between the two CEDD descriptors) if the distance of their B-CEDD descriptors activates the membership function 'Low' with value equal to 1. This threshold is defined as $T1$.

According to the second scenario, two images may be potentially similar if the distance of their B-CEDD descriptors activates the membership function 'Low' with value greater than 0.5. This threshold is defined as $T2$. At the end, according to the third scenario, two images may be potentially similar if the distance of their B-CEDD descriptors activates the membership function 'Low' with value greater than 0. This threshold is defined as $T3$.

The values of T for all the approaches are given in Table II, while in the case of B-CEDD$_1$ and B-CEDD$_2$ the threshold values are depicted in Figure 6 with a dashed line.

Table II
THRESHOLD VALUES

	$T1$	$T2$	$T3$
B-CEDD$_1$ Thresholds	3	4	5
B-CEDD$_2$ Thresholds	3	4	5
B-CEDD$_3$ Thresholds	5	7	8

Initially, our experiments were performed using Wang database. The Wang database is a subset of 1000 manually-selected images from the Corel stock photo database and forms 10 classes of 100 images each. In particular, the queries and ground-truth proposed by the MIRROR[17] image retrieval system are used. MIRROR separates the WANG database into 20 queries.

For each of these queries, the time of the retrieval through the proposed system is measured, as well as the time required when only the CEDD descriptor is used. Table III illustrates the mean results for the 20 queries of the Wang database for all the approaches and all the T values. The Recall index represents the ratio of the correct image retrievals for each query (images that belong to the ground truth of the query) to the size of the ground truth. Therefore, the Recall @ n describes the percentage of the correct images that were retrieved for all the queries. On the other hand, the ANMRR index evaluates the order in which the results were placed after the completion of the procedure. Thus, in order to assess the systems effectiveness properly both measures should be taken into account.

Considering these results, it can be readily observed that the proposed method improves the searching procedure times significantly. For $T1$, all approaches are capable of retrieving almost 112,000 images per second. But in all three approaches both Recall @ n and ANMRR have a very small value. This means that a lot of images from the ground truth were absorbed during the retrieval procedure. By comparing the three approaches, B-CEDD$_1$ seems to perform better.

The threshold $T2$, which retrieves almost 83,333 images

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

Table III
EXPERIMENTS ON THE WANG DATABASE

CEDD Time	45.2ms		
	$T1$	$T2$	$T3$
B-CEDD$_1$ Time	10.0ms	13.1ms	19.0ms
B-CEDD$_2$ Time	9.1ms	12.2ms	19.0ms
B-CEDD$_3$ Time	9.1ms	13.7ms	19.1ms
B-CEDD$_1$ Retrieved Im.	242.95	404.35	562.20
B-CEDD$_2$ Retrieved Im.	218.40	381.45	563.15
B-CEDD$_3$ Retrieved Im.	212.20	445.75	599.65
B-CEDD$_1$ Recall @ n	**0.6476**	**0.8188**	**0.9234**
B-CEDD$_2$ Recall @ n	0.5441	0.7074	0.8675
B-CEDD$_3$ Recall @ n	0.4971	0.7715	0.8704
CEDD ANMRR	0.2528		
B-CEDD$_1$ ANMRR	**0.4011**	**0.3100**	**0.2636**
B-CEDD$_2$ ANMRR	0.4836	0.3747	0.2902
B-CEDD$_3$ ANMRR	0.5322	0.3397	0.2824

Table IV
EXPERIMENTS ON THE UCID DATABASE

CEDD Time	60.3ms		
	$T1$	$T2$	$T3$
B-CEDD$_1$ Time	22.7ms	25.2ms	34.0ms
B-CEDD$_2$ Time	19.5ms	25.9ms	33.3ms
B-CEDD$_3$ Time	16.0ms	18ms	27.4ms
B-CEDD$_1$ Retrieved Im.	550.13	785.68	1007.61
B-CEDD$_2$ Retrieved Im.	472.08	825.04	987.37
B-CEDD$_3$ Retrieved Im.	379.14	618.90	866.44
B-CEDD$_1$ Recall @ n	**0.8920**	**0.9457**	**0.9734**
B-CEDD$_2$ Recall @ n	0.7990	0.8926	0.9478
B-CEDD$_3$ Recall @ n	0.8349	0.9424	0.9653
CEDD ANMRR	0.2823		
B-CEDD$_1$ ANMRR	**0.2971**	**0.2833**	**0.2823**
B-CEDD$_2$ ANMRR	0.3258	0.2905	0.2831
B-CEDD$_3$ ANMRR	0.3140	0.2874	0.2848

per second and its performance is found satisfactory for both the Recall @ n and the ANMRR, could be considered as the 'golden-mean' solution. Compared to the CEDD method, the proposed method with $T2$ retrieves almost four times more images per second. Also in this case, B-CEDD$_1$ seems to advance, although it requires slightly more time than the B-CEDD$_2$ (8.33% more time / 1 ms), it improves the Recall @ n by 16.6% compared to the performance of B-CEDD$_2$. Better performance is observed also in ANMRR where B-CEDD$_1$, having the value of 0.31, is better than B-CEDD$_3$ by 8.74%.

With $T3$, good results were achieved from all the approaches for both the Recall @ n and the ANMRR but the searching time was over doubled in comparison to the corresponding time for $T1$. Also in this case B-CEDD$_1$ performs better than the other two approaches, with its value of Recall @ $n = 0.9234$ approaching the CEDD performance, consider that the value of ANMRR is smaller by 4% from the corresponding value that the CEDD descriptor presents.

In the sequel, experiments were performed using the UCID Database. The UCID database was created as a benchmark database for CBIR and image compression applications. This database currently consists of 1338 uncompressed TIFF images on a variety of topics including natural scenes and man-made objects, both indoors and outdoors. All the UCID images were subjected to manual relevance assessments against 262 selected images, creating 262 ground truth image sets for performance evaluation. The results for all the three B-CEDD extraction approaches, for all the T values are illustrated in Table IV.

In this database we observe that even the smallest values of T retrieve more images compared to the images that were retrieved for the corresponding values of T on Wang database. We also observe that even for very small T, as for $T1$, there are good results for Recall @ n from all the three approaches that we used. Also in this case, B-CEDD$_1$ performs way better than the other two approaches, showing improvement from the B-CEDD$_3$ by 6.83%, while the improvement from B-CEDD$_2$ equals to 11.64%. B-CEDD $_1$ ANMRR value is by 5.24% smaller than the corresponding value of the CEDD descriptor, but the system retrieves twice as many images in the same time.

For $T2$, B-CEDD$_1$ performance is approaching the performance of CEDD. With Recall @ n value at 0.9457, ANMRR presents deviation from the corresponding CEDD value by just 0.35%. Significant improvements also presented by the other two approaches.

Finally, for $T3$, the performance of B-CEDD$_1$ is identical to the CEDD performance. At the same time, the other two methods also improve their performance, with B-CEDD$_2$ to behave slightly better than B-CEDD$_3$.

Observing the results in the two databases, is obvious that the B-CEDD$_1$ approach performs better than the other two approaches. Considering the threshold value, is obvious that by $T2$, a satisfactory trade-off between the acceleration rate of the retrieval procedure and the performance rate of the system is ensured. But how important is the reduction in the results? In order to quantize the reduction that is observed, we use a significance test. Significance tests tell us whether an observed effect, such as a difference between two means or a correlation between two variables, could reasonably occur 'just by chance' in selecting a random sample [27]. We used a bootstrap test, one-tailed, at significance levels 0.05(*), 0.01 (**), and 0.001 (***), against the CEDD results baseline in UCID database . The significance test was applied on Mean Average Precision (MAP):

$$MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q) \tag{8}$$

where Q is the set of queries q.

$$AP(q) = \frac{1}{N_R} \sum_{n=1}^{N_R} P_Q(R_n) \tag{9}$$

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

where R_n is the recall after the nth relevant image was retrieved. N_R is the total number of relevant documents for the query.

Table V
SIGNIFICANCE TEST RESULTS

CEDD MAP	0.6748		
	$T1$	$T2$	$T3$
B-CEDD$_1$ MAP	**0.6598****	**0.6720-**	**0.6748-**
B-CEDD$_2$ MAP	0.6324***	0.6639**	0.6730-
B-CEDD$_3$ MAP	0.6747***	0.6685**	0.6720-

Observing the results of Table V, we conclude that all three approaches for $T3$ have non-significant reductions in their results. On the other hand, for $T2$ B-CEDD$_1$ is the only approach which has a non-significant reduction. The significance test results support the conclusion that the best method is B-CEDD$_1$ for $T2$.

Finally, given that the proposed descriptor is an MPEG-7 like descriptor, the schema of the B-CEDD as an MPEG-7 extension is described as follows:

```
<?xml version ="1.0" encoding="UTF-8"?>
<schema xmlns="http://www.w3.org/2001/XMLSchema"
       xmlns:mpeg7="urn:mpeg:mpeg7:schema :2004"
        xmlns:SpCDNS="B-CEDDNS"
                    targetNamespace="B-CEDDNS">
  <import namespace="urn:mpeg:mpeg7:schema:2004"
               schemaLocation="Mpeg7-2004.xsd"/>
  <complexType name="B-CEDDType" final="#all">
    <complexContent>
      <extension base="mpeg7:VisualDType">
       <sequence>
         <element name="value">
          <simpleType>
           <restriction>
            <simpleType>
              <listitemType="mpeg7:Binary"/>
            </simpleType>
             <length value="15"/>
           </restriction>
          </simpleType>
         </element>
       </sequence>
      </extension>
    </complexContent>
  </complexType>
</schema>
```

VIII. CONCLUSIONS

We proposed an extension of the Color and Edge Directivity Descriptor which improves the speed efficiency of the CEDD. Through the application of the Modified Binary Haar Wavelet Transform on the CEDD, the proposed method achieves the extraction of a second, smaller (15 bits length), descriptor. Essentially, each CEDD descriptor is described by another compact binary descriptor. During the image searching process, the compact versions of the descriptors are employed, and only when their distance is smaller than a given threshold the searching continues with the CEDD. The distance between the B-CEDD descriptors is calculated by using a simple X-OR gate. The logic gate X-OR has much less computational cost than the Tanimoto coefficient. One of the most important attributes of the Binary CEDD (B-CEDD) is that it is extracted directly from the CEDD, without the need of the described image. This enables its immediate extraction from pre-existing index files. The effectiveness of the proposed method was demonstrated through experiments. Finally, it is worth noting that the proposed method can be applied to all Compact Composite Descriptors [23], [28], [29].

REFERENCES

[1] R. Datta, D. Joshi, J. Li, and J. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40(2), pp. 1–60, 2008.

[2] E. Hörster, R. Lienhart, and M. Slaney, "Image retrieval on large-scale image databases," in *Proceedings of the 6th ACM international conference on Image and video retrieval*. ACM, 2007, pp. 17–24.

[3] H. Jagadish, A. Mendelzon, and T. Milo, "Similarity-based queries," in *Proceedings of the fourteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*. ACM, 1995, pp. 36–45.

[4] G. Amato and P. Savino, "Approximate similarity search in metric spaces using inverted files," in *Proceedings of the 3rd international conference on Scalable information systems*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2008, pp. 1–10.

[5] P. Zezula, P. Savino, G. Amato, and F. Rabitti, "Approximate similarity retrieval with m-trees," *The VLDB Journal*, vol. 7, no. 4, pp. 275–293, 1998.

[6] P. Ciaccia and M. Patella, "Pac nearest neighbor queries: Approximate and controlled search in high-dimensional and metric spaces," in *ICDE*. Published by the IEEE Computer Society, 2000, p. 244.

[7] G. Amato, F. Rabitti, P. Savino, and P. Zezula, "Region proximity in metric spaces and its use for approximate similarity search," *ACM Transactions on Information Systems (TOIS)*, vol. 21, no. 2, pp. 192–227, 2003.

[8] H. Ferhatosmanoglu, E. Tuncel, D. Agrawal, and A. El Abbadi, "Approximate nearest neighbor searching in multimedia databases," in *Proceedings Of The International Conference On Data Engineering*. Citeseer, 2001, pp. 503–514.

[9] D. Blei, A. Ng, and M. Jordan, "Latent dirichlet allocation," *The Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[10] S. Chatzichristofis and Y. Boutalis, "Cedd: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval," *LNCS, Computer Vision Systems*, pp. 312–322, 2008.

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org

[11] S. A. Chatzichristofis, K Zagoris, Y. S. Boutalis and N. Papamarkos, "Accurate image retrieval based on compact composite descriptors and relevance feedback information," *International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI)*, vol. 24 (2), pp. 207–244, 2010.

[12] Z. Struzik and A. Siebes, "The haar wavelet transform in the time series similarity paradigm," *Principles of Data Mining and Knowledge Discovery*, pp. 12–22, 1999.

[13] S. A. Chatzichristofis, Y. Boutalis, and A. Arampatzis, "Accelerating image retrieval using binary haar wavelet transform on the color and edge directivity descriptor," in *The Fifth International Multi-Conference on Computing in the Global Information Technology (ICCGI 2010)*, 2010, pp. 41–47.

[14] B. Manjunath, J. Ohm, V. Vasudevan, A. Yamada *et al.*, "Color and texture descriptors," *IEEE Transactions on circuits and systems for video technology*, vol. 11, no. 6, pp. 703–715, 2001.

[15] C. Won, D. Park, and S. Park, "Efficient use of mpeg-7 edge histogram descriptor," *Etri Journal*, vol. 24, no. 1, pp. 23–30, 2002.

[16] J. Wang, J. Li, and G. Wiederholdy, "Simplicity: Semantics-sensitive integrated matching for picture libraries," *Advances in Visual Information Systems*, vol. 1929/2000, pp. 171–193, 2000.

[17] K. Wong, K. Cheung, and L. Po, "Mirror: an interactive content based image retrieval system," in *IEEE International Symposium on Circuits and Systems*, vol. 2. IEEE; 1999, 2005, pp. 1541–1544.

[18] G. Schaefer and M. Stich, "Ucid-an uncompressed colour image database," *Storage and Retrieval Methods and Applications for Multimedia 2004*, vol. 5307, pp. 472–480, 2004.

[19] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2. Citeseer, 2006, pp. 2161–2168.

[20] S. A. Chatzichristofis, Y. S. Boutalis and M. Lux, "Img(rummager): An interactive content based image retrieval system," in *2nd International Workshop on Similarity Search and Applications (SISAP)*, 2009, pp. 151–153.

[21] M. Lux and S. Chatzichristofis, "Lire: lucene image retrieval: an extensible java cbir library," in *Proceeding of the 16th ACM international conference on Multimedia*. ACM, 2008, pp. 1085–1088.

[22] S. Krishnamachari and M. Abdel-Mottaleb, "Compact color descriptor for fast image and video segment retrieval," in *IS&T/SPIE Conference on Storage and Retrieval of Media Databases*, 2000, pp. 581–589.

[23] S. A. Chatzichristofis, Y. S. Boutalis and M. Lux, "Spcd - spatial color distribution descriptor. a fuzzy rule based compact composite descriptor appropriate for hand drawn color sketches retrieval," in *2nd International Conference on Agents and Artificial Intelligence (ICAART)*, 2010, pp. 58–63.

[24] A. Arampatzis and J. Kamps, "A signal-to-noise approach to score normalization," in *Proceeding of the 18th ACM conference on Information and knowledge management*. ACM, 2009, pp. 797–806.

[25] M. Fernández, D. Vallet, and P. Castells, "Using historical data to enhance rank aggregation," in *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2006, pp. 643–644.

[26] D. Gustafson and W. Kessel, "Fuzzy clustering with a fuzzy covariance matrix," in *1978 IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes*, vol. 17, 1978, pp. 761–766.

[27] D. Moore and L. Sorenson, *Introduction to the Practice of Statistics SPSS Manual*. WH Freeman, 2005.

[28] S. Chatzichristofis and Y. Boutalis, "Fcth: Fuzzy color and texture histogram-a low level feature for accurate image retrieval," in *Image Analysis for Multimedia Interactive Services, 2008. WIAMIS'08. Ninth International Workshop on*, 2008, pp. 191–196.

[29] Chatzichristofis, S.A. and Boutalis, Y.S., "Content based radiology image retrieval using a fuzzy rule based scalable composite descriptor," *Multimedia Tools and Applications*, vol. 46, pp. 493–519, 2010.

2011, © Copyright by authors, Published under agreement with IARIA - www.iaria.org