CBIR using Combined Feature Vectors of Column-Wise and Row-Wise DCT Transformed Plane Sectorization.

H.B.Kekre
Sr. professor, Computer Engg.
MPSTME,SVKM's NMIMS
Vile Parle West Mumbai-India

Dhirendra Mishra
Associate professor, Computer Engg.
MPSTME,SVKM's NMIMS
Vile Parle West Mumbai-India

Rohan Shah, Chirag Thakkar, Shikha Shah
B.Tech Students
MPSTME,SVKM's NMIMS
Vile Parle West Mumbai-India

ABSTRACT
Content Based Image Retrieval is a way of computer viewing technique used to retrieve digital images from a huge database. In this paper we have first calculated the feature vector column-wise and row-wise separately. After this we have concatenated the feature vectors of column-wise and row-wise. To evaluate the performance of the proposed method we have used Precision-Recall crossover point, LIRS, LSRR and LSRI. Sum of Absolute Distance and Euclidean Distance are the two similarity measures used. The column-wise row-wise DCT transformed image is sectorized on the basis of even-odd column components of transformed image with augmentation of zero and highest row components. The proposed algorithm is applied to a database of thousand images. These thousand images are grouped in ten different classes. Performance is evaluated and compared for 4, 8, 12, 16 DCT sectors.

General Terms
The general terms used are CBIR (Content Based Image retrieval), LSR, LIRS, Absolute distance, LSRI (Longest string of relevant Retrieved images).

1. INTRODUCTION
Content Based Image Retrieval (CBIR) algorithms are a means to access a digital image from a database [2]. In CBIR the actual contents of the image are used to describe as well as analyse an image. The contents of an image are used to form a set of feature vectors. Various techniques are used for extraction of the feature vectors for an image. Some of these techniques involve using attributes like shape [1], color [13], textures [13] and edge density [13] of an image to extract the feature vector. Different CBIR systems have adopted different techniques. Feature vectors are unique identities of images [1]. They are used to differentiate among images. They are used to facilitate better retrieval of relevant images. The feature vector of the query image is compared to that of the feature vectors of the images in the database using the similarity measures such as Euclidean and Sum of Absolute distance. This results in retrieval of the images with feature vectors near to that of the query images. CBIR has attracted research interests of people from various fields such as Artificial Intelligence, Data Mining, Web Development, Information Theory, Statistics, etc. [4]. Applications of CBIR are fingerprint recognition [3], iris recognition [5], face recognition [6], etc. This paper proposes the use of sectorization of column-row-wise DCT transformed images on the basis of even-odd column components of Column-Row DCT transformed image. This is done with and without augmentation of zero and highest row components for feature vector extraction. In this paper first we have done the sectorization column-wise then row-wise and then we have concatenated both the output. A Wang database [10, 11] of thousand images is taken. CBIR has attracted research interests of people from various fields such as Artificial Intelligence, Data Mining, Web Development, Information Theory, Statistics, etc. [7]. There are many applications of CBIR. These include bit truncation coding [15, 16], Gaussian Mixtures [14], feature vector extraction using color histograms [17, 18], Biometrics [19], feature vector extraction using color histograms [17, 20].

2. DISCRETE COSINE TRANSFORM
DCT is made up of cosine functions taken over half the interval and dividing this half interval into N equal parts and sampling each function at the center of these parts. The discrete cosine transform matrix is formed by keeping these sequences [7]. The most common DCT definition of 1D Sequence of Length N is

\[ C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos \left( \frac{\pi (2x+1)u}{2N} \right) \]  

For \( u = 0, 1, 2, \ldots, N-1 \).

Similarly inverse transform is given as:

\[ f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos \left( \frac{\pi (2x+1)u}{2N} \right) \]  

For \( x = 0, 1, 2, \ldots, N-1 \) and \( \alpha(u) \) for both the above equations is given as:

\[ \alpha(u) = \begin{cases} \frac{1}{\sqrt{N}} & \text{for } u = 0 \\ \frac{2}{\sqrt{N}} & \text{for } u \neq 0 \end{cases} \]  

3. FEATURE VECTOR GENERATION:
In our proposed method, we apply DCT transform column-wise and row-wise separately to the image and then we concatenate the result of both. We have first calculated the feature vector by column-wise and row-wise separately. After this we have concatenated the feature vectors of column-wise and row-wise. The proposed algorithm uses a combination of even-odd co-efficient pair to sectorize the DCT transformed image putting odd co-efficient on Y-axis and even co-efficient on the X-axis. Considering these components as co-ordinates, we get a point in X-Y plane as shown in Figure 1. We will be sectorizing the transformed images in 4, 8, 12, 16 sectors.
based on values of co-efficient. This proposed even-odd plane is used for feature vector generation.

**Fig. 1: DCT even-odd plane used**

We have proposed the use of three different types feature vectors per plane. Mean value of all the co-efficient in a sector with augmentation of two extra components for each color plane i.e. R, G and B: Here the average of all the co-efficient placed in a sector for every plane is taken into consideration. The zeroth and highest row components are augmented to the feature vector for every plane. For 4 DCT sectors, there are four components in the feature vector per plane, 1 for each sector and 2 components for the augmented rows. Thus for 4 DCT sectors, the size of the feature vector is 6*3=18. Similarly, for 8, 12, 16 DCT sectors, the sizes of the database are 30, 42 and 54 respectively. The results for a given query image are computed for each of these feature vectors and using the similarity measures Sum of Absolute Distance [3,7,8,14] and Euclidean Distance[4,3,8,14].

### 3.1 4 DCT sectors

DCT of the image is measured in all three planes namely red, blue and green planes. The even row/columns and odd row/columns components are checked for negative and positive signs. It can be better understood we the following chart

**Table 1:** Formation of Four DCT sectors

<table>
<thead>
<tr>
<th>Sign of even row/column</th>
<th>Sign of odd row/column</th>
<th>Quadrant assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>I (0 – 90°)</td>
</tr>
<tr>
<td>+</td>
<td>–</td>
<td>II (90 – 180°)</td>
</tr>
<tr>
<td>–</td>
<td>–</td>
<td>III (180 – 270°)</td>
</tr>
<tr>
<td>–</td>
<td>+</td>
<td>IV (270 – 360°)</td>
</tr>
</tbody>
</table>

### 3.2 8 DCT sectors

Each sector obtained in the previous section is divided into 2 equal sectors, each of 45°. In all we have 8 sectors for each plane which is shown in our previous work [7, 12].

### 3.3 12 DCT Sectors

Each sector obtained in the previous section of 4 sectors is divided into 3 equal sectors, each of 30°. In all we have 12 sectors for each plane. For one plane the sectors are divided as shown in our previous work [7, 12].

### 3.4 16 DCT sectors

Sixteen sectors are obtained by dividing each one of eight sectors into two equal parts each of 22.5°.

### 4. RESULTS AND DISCUSSIONS

The database used to calculate the performance of our proposed algorithm consisted of 1000 images of ten different classes. Classes include images of Tribal, Beaches, Monuments, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food Dishes. The figure below shows the sample images from the database.

The images in above figure are displayed class-wise i.e. Class 1 has images of tribal and so on. Every class has 100 images. For all the images feature vector is generated and feature database is formed. From every class five random images are selected to evaluate the performance. Feature vector of the query images are compared with the feature vector in the feature vector database using Similarity measures Euclidean Distance and Absolute Distance. The smaller the similarity measure better is the match with the respective image in the database.

**Fig. 2:** Sample images from the Database.

To check the effectiveness we have made use of precision and recall using the equations given below:

\[
\text{Precision} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Images Retrieved}} \quad \ldots (4)
\]

\[
\text{Recall} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Relevant Images in Database}} \quad \ldots (5)
\]

Also we have made use of two more performance evaluation parameters: Length of Initial Relevant retrieval Strings (LIRS) and Length of String to Recover all Relevant images in database (LSRR) given by the equations below:

\[
\text{LIRS} = \frac{\text{Length of Initial Relevant String of Images}}{\text{Total Number of Relevant Images Retrieved}} \quad \ldots (6)
\]

\[
\text{LSRR} = \frac{\text{Length of String to Recover all Relevant Images}}{\text{Total Number of Images in the Database}} \quad \ldots (7)
\]

Higher the precision-recall crossover point better is the performance. Similarly higher the value of LIRS better is the performance of the proposed algorithm, whereas lower the value of LSRR better is the performance.

We have introduced a new parameter: LSRI [12] i.e. length of the longest string of images given by the following equation.
This is the measure of length of the longest string of relevant images retrieved at any point during the process of retrieval.

\[ LSRI = \frac{\text{Longest string of relevant images retrieved}}{\text{Total Number of relevant Images in the Database}} \quad (8) \]

All these performance evaluation parameter values are in the range of 0-1 and these can be represented in the percentage.

Figure 3 and 4 depicts the average crossover point of Precision-Recall DCT Transform (column-row wise) sectorization using 4, 8, 12 and 16 sectors.

The performance of the evaluation parameter Precision-Recall crossover point is shown in figure 5 and figure 6 for both the similarity measure using Euclidean Distance and Sum of Absolute Distance.

The performance with respect to retrieval rate is 40% with Euclidean Distance as similarity measure and 40% with Sum of Absolute Distance as similarity measure for 4 sectors. The performance with respect to retrieval rate is 40% with Euclidean Distance as similarity measure and 43% with Sum of Absolute Distance as similarity measure for 8 sectors. The performance with respect to retrieval rate is 40% with Euclidean Distance as similarity measure and 40% with Sum of Absolute Distance as similarity measure for 12 sectors. The performance with respect to retrieval rate is 40% with Euclidean Distance as similarity measure and 43% with Sum of Absolute Distance as similarity measure for 16 sectors. We infer that the proposed algorithm performs better when 16 sectors are used to sectorize the image and when using Sum of Absolute Distance as similarity measure. From Figure 6 and Figure 7 we conclude that the performance of the proposed algorithm varies based on the use the similarity measure. Also, classes 1 and 7 yield better results when Euclidean Distance is used as the similarity measure. Also, when Sum of Absolute Distance is used as the similarity measure, classes 9 have a decline in the retrieval rate with the increase in the number of sectors which again opposes the conclusion which depends on the overall performance of the algorithm.

The best result class-wise with respect to Precision – Recall Crossover point using Euclidean and Sum of Absolute Distance is shown in the following Table 2 and Table 3. The worst performance is seen in class 3 with respect to Precision – Recall i.e. images of monuments with an average of 22% using Euclidean Distance and Sum of Absolute distance as a similarity measure each. The best performance is marked with yellow colour in the following table.

![Figure 4: Plot for average precision recall crossover point using sum of absolute distance](image)

Figure 4: Plot for average precision recall crossover point using sum of absolute distance

**Table 2: Top 3 class-wise performance using Euclidean Distance w.r.t. Precision-Recall**

<table>
<thead>
<tr>
<th>Class</th>
<th>Average Precision and recall crossover point plot for sector 4, 8, 12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>98%</td>
</tr>
<tr>
<td>7</td>
<td>39%</td>
</tr>
<tr>
<td>8</td>
<td>48%</td>
</tr>
</tbody>
</table>

**Table 3: Top 3 class-wise performance using Sum of Absolute Distance w.r.t. Precision-Recall**

<table>
<thead>
<tr>
<th>Class</th>
<th>Average Precision and recall crossover point plot for sector 4, 8, 12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>98%</td>
</tr>
<tr>
<td>7</td>
<td>55%</td>
</tr>
<tr>
<td>8</td>
<td>40%</td>
</tr>
</tbody>
</table>

The performance of the evaluation parameter LIRS is shown in figure 5 and figure 6 for both the similarity measure using Euclidean Distance and Sum of Absolute Distance. For all 4, 8, 12 & 16 sectors the performance of LIRS is compared.

It is evident that higher the value of LIRS higher is the performance of the algorithm from equation 6. The performance of algorithm with respect to LIRS is 13% with Euclidean Distance and 12% with respect to Sum of Absolute distance in 4 sectors. The performance of algorithm with respect to LIRS is 12% with Euclidean Distance and 13% with respect to Sum of Absolute distance in 8 sectors. The performance of algorithm with respect to LIRS is 13% with Euclidean Distance and 14% with respect to Sum of Absolute distance in 12 sectors. The performance of algorithm with respect to LIRS is 14% with Euclidean Distance and 14% with respect to Sum of Absolute distance in 16 sectors. Classes 3 and 5 yield better results when Euclidean Distance is used as the similarity measure. Also, when Sum of Absolute Distance is used as the similarity measure, classes 5 have a decline in the retrieval rate with the increase in the number of sectors which again opposes the conclusion which depends on the overall performance of the algorithm.
Figure 5: Plot for LIRS using Euclidean Distance as similarity measure.

Figure 6: Plot for LIRS using Sum of Absolute Distance as similarity measure.

The worst performance is seen in class 4 with respect to LIRS i.e. images of buses with an average of 1% and 5% using Euclidean Distance and Sum of Absolute distance respectively. The best result class-wise with respect to LIRS using Euclidean and Sum of Absolute Distance is shown in the following Table 4 and Table 5. The best performance is marked in yellow color in the following table.

Table 4: Top 3 class-wise performance using Euclidean Distance w.r.t. LIRS

<table>
<thead>
<tr>
<th>Class</th>
<th>Average LIRS plot for sector 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>94%</td>
</tr>
<tr>
<td>8</td>
<td>9%</td>
</tr>
<tr>
<td>7</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 5: Top 3 class-wise performance using Sum of Absolute Distance w.r.t. LIRS

<table>
<thead>
<tr>
<th>Class</th>
<th>Average Precision and recall crossover point plot for sectors 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>92%</td>
</tr>
<tr>
<td>8</td>
<td>11%</td>
</tr>
<tr>
<td>7</td>
<td>10%</td>
</tr>
</tbody>
</table>

The performance of the evaluation parameter LSRR is shown in figure 7 and figure 8 for both the similarity measure using Euclidean Distance and Sum of Absolute Distance. For all 4,8,12 & 16 sectors the performance of LSRR is compared.

It is evident that lower the value of LSRR higher is the performance of the algorithm from equation 6. The performance of algorithm with respect to LSRR is 75% with Euclidean Distance and 74% with respect to Sum of Absolute distance in 4 sectors. The performance of algorithm with respect to LSRR is 75% with Euclidean Distance and 73% with respect to Sum of Absolute distance in 8 sectors. The performance of algorithm with respect to LSRR is 74% with Euclidean Distance and 73% with respect to Sum of Absolute distance in 12 sectors. The performance of algorithm with respect to LSRR is 75% with respect to Euclidean Distance and 72% with respect to Sum of Absolute distance in 16 sectors. Classes 1, 6 and 7 yield better results when Euclidean Distance is used as the similarity measure. Also, when Sum of Absolute Distance is used as the similarity measure, classes 4, 6, 7 and 8 have a decline in the retrieval rate with the increase in the number of sectors which opposes the conclusion which depends on the overall performance of the algorithm.

Table 6: Top 3 class-wise performance using Euclidean Distance w.r.t. LSRR

<table>
<thead>
<tr>
<th>Class</th>
<th>Average LSRR plot for sector 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>92%</td>
</tr>
<tr>
<td>8</td>
<td>11%</td>
</tr>
<tr>
<td>7</td>
<td>10%</td>
</tr>
</tbody>
</table>

Table 7: Top 3 class-wise performance using Sum of Absolute Distance w.r.t. LSRR

<table>
<thead>
<tr>
<th>Class</th>
<th>Average Precision and recall crossover point plot for sectors 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>92%</td>
</tr>
<tr>
<td>8</td>
<td>11%</td>
</tr>
<tr>
<td>7</td>
<td>10%</td>
</tr>
</tbody>
</table>

The overall performance classes 1, 6 and 7 yield better results when Euclidean Distance is used as the similarity measure. Also, when Sum of Absolute Distance is used as the similarity measure, classes 4, 6, 7 and 8 have a decline in the retrieval rate with the increase in the number of sectors which opposes the conclusion which depends on the overall performance of the algorithm.

The worst performance is seen in class 3 with respect to LSRR i.e. images of monuments with an average of 92% using Euclidean Distance and Sum of Absolute Distance both. The best result class-wise with respect to LIRS using Euclidean and Sum of Absolute Distance is shown in the following Table 6 and Table 7. The best performance is marked in yellow color in the following table.
Table 6: Top 3 class wise performance w.r.t. LSRR using Euclidean distance

<table>
<thead>
<tr>
<th>Class</th>
<th>Average LSRR plot for sector 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>70%</td>
</tr>
<tr>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>7</td>
<td>62%</td>
</tr>
</tbody>
</table>

Table 7: Top 3 class wise performance w.r.t. LSRR using Sum of Absolute Distance

<table>
<thead>
<tr>
<th>Class</th>
<th>Average LSRR plot for sector 4,8,12 &amp; 16</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>10%</td>
</tr>
<tr>
<td>6</td>
<td>68%</td>
</tr>
<tr>
<td>7</td>
<td>50%</td>
</tr>
</tbody>
</table>

More the value of the above parameter better is the performance in LSRI. The performance of the evaluation parameter LSRI is shown in figure 9 and figure 10 for both the similarity measure using Euclidean Distance and Sum of Absolute Distance. The performance of algorithm with respect to LSRI (Longest String of Relevant Retrieved Images) is 14% with respect to Euclidean Distance and 14% with respect to Sum of Absolute distance in 4 sectors. The performance of algorithm with respect to LSRI is 14% with Euclidean Distance and 15% with respect to Sum of Absolute distance in 8 sectors. The performance of algorithm with respect to LSRI is 15% with Euclidean Distance and 15% with respect to Sum of Absolute distance in 12 sectors. The performance of algorithm with respect to LSRR is 14% with Euclidean Distance and 15% with respect to Sum of Absolute distance in 16 sectors.

Classes 3, 5 yield better results when Euclidean Distance is used as the similarity measure. Also, when Sum of Absolute Distance is used as the similarity measure, classes 7 have a decline in the retrieval rate with the increase in the number of sectors which opposes the conclusion which depends on the overall performance of the algorithm. All the query images used by us have the value of LSRI is Always bigger than 1. Thus we come to a conclusion that this algorithm can retrieval relevant images in a sequence at least once during the process of retrieval. The worst performance is seen in class 9 with respect to LSRI with 3% using Euclidean distance and class 3 with respect to LSRI with 3% using Sum of Absolute distance. The best result class-wise with respect to LIRS using Euclidean and Sum of Absolute Distance is shown in the following Table 8 and Table 9. The best performance is marked in yellow color in the following table.
5. CONCLUSION
An algorithm of sectorization of combined column-row wise DCT transformed planes of images is proposed. The plane has been sectorized in 4, 8, 12 and 16 sectors to generate the feature vector and also performance parameter are used such as Precision-Recall crossover point, LIRS, LSRR and LSRI (Longest string of relevant images retrieved) are used for similarity measures such as Euclidean and Sum of Absolute distance. This algorithm is checked with even-odd row DCT component of column-wise DCT transformed image and even-odd column DCT component of row-wise DCT transformed image. The average value of zeroth value is taken along with last row and then augmented into feature vector. Performance is evaluated for mentioned parameters and then compared with Euclidean distance and Sum of Absolute distance. Thus we come to a conclusion that when Euclidean Distance is 40% and 42% when Sum of Absolute Distance is used as the similarity measure. With respect to LIRS, the performance of the algorithm when Euclidean Distance is 12% and 13% when Sum of Absolute Distance is used as the similarity measure. With respect to LSRR, the performance of the algorithm when Euclidean Distance is 75% and 73% when Sum of Absolute Distance is used as the similarity measure. Thus we can say that the algorithm produces better results when Sum of Absolute distance is used. From the Precision – Recall crossover point plots for all the query images we can say that the crossover point is always generated when the number of retrieved images is equal to the number of images in the relevant class which in our case is 100 for all classes. From table 9 we come to a conclusion that when Euclidean Distance is used as a similarity measure Precision-Recall, LSRR give better result in column-wise feature extraction while LSRI, LIRS yield better result in column-wise feature extraction. When Sum of Absolute Distance is used as similarity measure Precision - Recall yields better result in column-wise feature extraction. While LIRS, LSRI gives better results in column-row wise feature extraction. But when overall review is done it is found that column wise feature extraction gives a better retrieval rate in comparison with that of column-row wise.

6. REFERENCES


