4. Image Retrieval using Transformed Image Content

The desire of better and faster retrieval techniques has always fuelled to the research in content based image retrieval (CBIR). A class of unitary matrices used for representing image can be referred as image transforms. Transforms change the representation of image by projecting it into a set of basis functions (commonly referred as basis images). The image transforms do not change the information content present in the image, they only allow the image representation to move from one domain (time domain) to another domain (frequency domain). This image transformation from one representation to another is mainly advantageous in two aspects. First, the transformation isolates critical components of image patterns which then can directly be accessed for analysis. Second, transformation places image data in more compact form that can be stored and transmitted efficiently. These two advantages of image transforms make them inevitable choice for feature vector size reduction in image retrieval techniques. These aspects of image transforms are exploited in the proposed CBIR techniques using fractional energy, row mean of column transformed image, energy compaction and Principle component analysis discussed in this chapter.

4.1 Image Transforms

If applied on image, the orthogonal image transforms have the property of compaction of energy towards the low frequency region. The benefit of energy compaction of transforms in higher coefficients is taken to reduce the feature vector size per image in image retrieval techniques proposed in this chapter. Smaller feature vector size results in less time for comparison of feature vectors
resulting in faster retrieval of images. Here the fairly large number of popular image transforms namely Cosine transform, Walsh transform, Haar transform, Sine transform, Slant transform, Hartley transform and newly introduced Kekre transform are considered. The section takes quick review of these image transforms.

4.1.1 Discrete Cosine Transform (DCT)

The discrete cosine transform (DCT) [65,227,228,229] is closely related to the discrete Fourier transform. The NxN cosine transform matrix C={c(k,n)}, also called the discrete cosine transform (DCT), is defined as equation 4.1.

\[
c(k,n) = \begin{cases} 
\frac{1}{\sqrt{N}}, & k = 0, \quad 0 \leq n \leq N - 1 \\
\sqrt{\frac{2}{N}} \cos \left( \frac{\pi(2n+1)k}{2N} \right), & 1 \leq k \leq N - 1, \quad 0 \leq n \leq N - 1 
\end{cases}
\]  

(4.1)

DCT of a sequence \(\{u(n), 0 \leq n \leq N-1\}\) is defined as equation 4.2.

\[
v(k) = \alpha(k) \sum_{n=1}^{N-1} u(n) \cos \left( \frac{\pi(2n+1)k}{2N} \right), \quad 0 \leq k \leq N - 1
\]

(4.2)

where,

\[
\alpha(0) \approx \frac{1}{\sqrt{N}}, \quad \alpha(k) \approx \frac{2}{\sqrt{N}} \quad \text{for} \ 1 \leq k \leq N - 1
\]

(4.3)

It is a separable linear transformation, which means the two-dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension. The DCT tends to concentrate information, making it useful for image compression applications and also helping in minimizing feature vector size in CBIR [150]. For full 2-Dimensional DCT for an NxN image the number of multiplications required are \(N^2(2N)\) and number of additions required are \(N^2(2N-2)\).
4.1.2 Walsh Transform

Walsh transform matrix \([230, 231, 232]\) is defined as a set of \(N\) rows, denoted as \(W_j\), for \(j = 0, 1, ..., N - 1\), which have the following properties.

- \(W_j\) takes on the values +1 and -1.
- \(W_j[0] = 1\) for all \(j\).
- \(W_j \times W_k^T = 0\) for \(j \neq k\) and \(W_j \times W_k^T = N\), for \(j=k\).
- \(W_j\) has exactly \(j\) zero crossings, for \(j = 0, 1, ..., N-1\).
- Each row \(W_j\) is even or odd with respect to its midpoint.

Walsh transform matrix is defined using a Hadamard matrix \([233]\) of order \(N\). The Walsh transform matrix row is the row of the Hadamard matrix specified by the Walsh code index, which must be an integer in the range \([0, ..., N - 1]\). For the Walsh code index equal to an integer \(j\), the respective Hadamard output code has exactly \(j\) changes of signs (‘-1 to +1’ or ‘+1 to -1’) zero crossings, for \(j = 0, 1, ..., N-1\). For the full 2-Dimensional Walsh transform applied to image of size \(N \times N\), the number of additions required are \(2N^2(N-1)\) and absolutely no multiplications are needed in Walsh transform.

4.1.3 Haar Transform

This sequence was proposed in 1909 by Alfréd Haar \([234]\). Haar used these functions to give an example of a countable orthonormal system for the space of square-integrable functions on the real line. The study of wavelets, and even the term "wavelet", did not come until much later \([235, 277]\). The Haar wavelet is also the simplest possible wavelet. The technical disadvantage of the Haar wavelet is that it is not continuous, and therefore not differentiable. This property can, however, be an advantage for the analysis of signals with sudden transitions, such as monitoring of tool failure in
machines. For full 2-Dimensional Haar Transform for an \( \text{N} \times \text{N} \) image the number of additions required are \( 2\text{N}^2 \log_2(\text{N}) \) and no multiplications needed. The Haar wavelet's mother wavelet function \( \psi(t) \) is described as equation 4.4.

\[
\psi(t) = \begin{cases} 
1 & 0 \leq t < \frac{1}{2} \\
-1 & \frac{1}{2} \leq t < 1 \\
0 & \text{otherwise}
\end{cases} \quad (4.4)
\]

And its scaling function \( \varphi(t) \) can be described as equation 4.5.

\[
\varphi(t) = \begin{cases} 
1 & 0 \leq t < 1 \\
0 & \text{otherwise}
\end{cases} \quad (4.5)
\]

### 4.1.4 Kekre Transform

Newly introduced Kekre transform matrix [154,159] is the generic version of Kekre’s LUV colour space matrix. Kekre transform matrix can be of any size \( \text{N} \times \text{N} \), which need not have to be in powers of 2 (as is the case with most of other transforms). All upper diagonal and diagonal values of Kekre transform matrix are one, while the lower diagonal part except the values just below diagonal is zero. Generalized \( \text{N} \times \text{N} \) Kekre transform matrix is given in equation 4.6.

\[
K_{\text{N,N}} = \begin{bmatrix}
1 & 1 & 1 & \ldots & 1 & 1 \\
-\text{N}+1 & 1 & 1 & \ldots & 1 & 1 \\
0 & -\text{N}+2 & 1 & \ldots & 1 & 1 \\
0 & 0 & 0 & \ldots & 1 & 1 \\
0 & 0 & 0 & \ldots & -\text{N}+(\text{N}+1) & 1
\end{bmatrix} \quad (4.6)
\]

The formula for generating the term \( K(x,y) \) of Kekre transform matrix is given in equation 4.7.

\[
K(x,y) = \begin{cases} 
1 & x \leq y \\
-\text{N}+(x-1) & x = y + 1 \\
0 & x > y + 1
\end{cases} \quad (4.7)
\]
For taking Kekre’s transform of an NxN image, the number of required multiplications are $2N(N-2)$ and number of additions required are $N(N^2+N-2)$.

### 4.1.5 Discrete Sine Transform (DST)

In mathematics, the discrete sine transform [236,237,238] is a Fourier-related transform similar to the discrete Fourier transform (DFT), but using a purely real matrix. It is equivalent to the half of the imaginary parts of a DFT, where the input and/or output data are shifted by half a sample. Formally, the discrete sine transform [236,237] is a linear, invertible function. The $NxN$ sine transform matrix $S=\{s(k,n)\}$, also called discrete sine transform (DST) is defined as equation 4.8.

$$s(k,n) = \sqrt{\frac{2}{N+1}} \sin \frac{\pi (k+1)(n+1)}{N+1}, 0 \leq k, n \leq N-1 \quad (4.8)$$

The sine transform pair of one-dimensional sequences is defined as equations 4.9 and 4.10.

$$v(k) = \sqrt{\frac{2}{N+1}} \sum_{n=0}^{N-1} u(n) \sin \frac{\pi (k+1)(n+1)}{N+1}, 0 \leq k \leq N-1 \quad (4.9)$$

$$u(n) = \sqrt{\frac{2}{N+1}} \sum_{k=0}^{N-1} v(k) \sin \frac{\pi (k+1)(n+1)}{N+1}, 0 \leq n \leq N-1 \quad (4.10)$$

The DST matrix is orthogonal. The elements of first row of DST matrix are not ‘1’ as the case with other transform matrices. The direct application of the DST for $NxN$ image would require $2N^2(N-1)$ additions and $N^2(2N)$ multiplications.
4.1.6 Slant Transform

The concept of an orthogonal transformation containing saw tooth waveforms or slant basis vectors was introduced by Enomoto and Shibata [239]. The slant transform has its first basis function as constant and second basis function as linear. The Slant vector is a discrete saw tooth waveform decreasing in uniform steps over its length [240,241]. It has been seen that Slant vectors are suitable for efficiently representing gradual brightness change in an image line as in the case with television signal and many other images. Slant Matrix Construction is explained below.

Let \( S(n) \) denotes the \( N \times N \) Slant matrix (\( N=2^n \)), then

\[
S(1) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (4.11)
\]

The Slant matrix for \( N=4 \) can be written as equation 4.12.

\[
S(2) = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ a+b & a-b & -a+b & -a-b \\ 1 & -1 & -1 & 1 \\ a-b & -a-b & a+b & -a+b \end{bmatrix} \quad (4.12)
\]

Where \( a \) and \( b \) are real constants to be determined subject to the conditions that the step size must be uniform and \( S(2) \) must be orthogonal, which gives

\[
a = \frac{2}{\sqrt{5}}, \quad b = -\frac{1}{\sqrt{5}}
\]

So normalized matrix \( S(2) \) becomes as shown in equation 4.13.

\[
S(2) = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ \frac{3}{\sqrt{5}} & \frac{1}{\sqrt{5}} & -\frac{1}{\sqrt{5}} & -\frac{3}{\sqrt{5}} \\ 1 & -1 & -1 & 1 \\ \frac{1}{\sqrt{5}} & -\frac{3}{\sqrt{5}} & \frac{3}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \end{bmatrix}
\]

Number of sign changes

0
1
2
3

C4-6
It is observed that S(2) possess the sequency property like Walsh Hadamard transform. The direct application of Slant transform on NxN image would require 2N²(N-1) additions and N²(N) multiplications.

4.1.7 Discrete Hartley Transform (DHT)

A discrete Hartley transform (DHT) [242,243,244,245] is a Fourier-related transform of discrete, periodic data similar to the discrete Fourier transform (DFT), with analogous applications in signal processing and related fields. Its main distinction from the DFT is that it transforms real inputs to real outputs, with no intrinsic involvement of complex numbers. Just as the DFT is the discrete analogue of the continuous Fourier transform, the DHT is the discrete analogue of the continuous Hartley transform, introduced by R. V. L. Hartley [246] in 1942. Because there are fast algorithms for the DHT analogous to the fast Fourier transform (FFT) [247], the DHT was originally proposed by R. N. Bracewell [242,243,244,245] in 1983 as a more efficient computational tool in the common case where the data are purely real. It was subsequently argued, however, that specialized FFT algorithms for real inputs or outputs can ordinarily be found with slightly fewer operations than any corresponding algorithm for the DHT. Formally, the discrete Hartley transform is a linear, invertible function \( H : \mathbb{R}^n \rightarrow \mathbb{R}^n \) (where \( \mathbb{R} \) denotes the set of real numbers). The \( n \) real numbers \( x_0, ..., x_{n-1} \) are transformed into the \( n \) real numbers \( h_0, ..., h_{n-1} \) according to the equation 4.14.

\[
H_k = \sum_{n=0}^{N-1} x_n \left[ \cos \left( \frac{2\pi n k}{N} \right) + \sin \left( \frac{2\pi n k}{N} \right) \right], \quad k = 0, ..., N - 1 \tag{4.14}
\]

The transform can be interpreted as the multiplication of the vector \((x_0, ..., x_{N-1})\) by an \( N \times N \) matrix; therefore, the discrete Hartley transform is a linear operator. The matrix is invertible; the inverse
transformation, which allows one to recover the $x_n$ from the $H_k$, is simply the DHT of $H_k$ multiplied by $1/N$. That is, the DHT is its own inverse, up to an overall scale factor. For taking Hartley transform of an $N \times N$ image, the number of required multiplications are $2N(N-2)$ and number of additions required are $N(N^2+N-2)$.

### 4.2 Image Retrieval using Fractional Energy of transformed Image Content

The innovative content based image retrieval (CBIR) techniques based on feature vectors as fractional coefficients of transformed images using various orthogonal transforms are presented in this section. The benefit of energy compaction of transforms in higher energy coefficients is taken to reduce the feature vector size per image by taking fractional coefficients of transformed image. Smaller feature vector size results in less time for comparison of feature vectors resulting in faster retrieval of images. Instead of using all coefficients of transformed images as feature vector for image retrieval, the fourteen coefficients sets (as shown in figure 4.1) with reduced size are used, resulting into better performance of image retrieval with lower computations.

The proposed CBIR techniques are tested on image database. For each proposed CBIR technique 55 queries (randomly selected 5 per category) are fired on the database and net average precision and recall are computed for all feature sets per transform. The results have shown performance improvement (higher precision and recall values) with fractional coefficients compared to complete transform of image at reduced computations resulting in faster retrieval. Finally newly introduced Kekre transform surpasses all other discussed transforms in performance with highest precision and recall values for fractional coefficients (6.25% and 3.125% of all
coefficients) and computation are lowered by 94.08% as compared to DCT or DST or Hartley transform.

4.2.1 Fractional Energy of Transformed Image

Figure 4.1 gives the feature sets extraction method for proposed CBIR techniques using fractional coefficients of transformed images [150,151]. The feature vectors are extracted in fourteen different ways from the transformed image, with the first being considering all the coefficients of transformed image (100%) and then fourteen fractional coefficients sets (as 50%, 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% of complete transformed image) are considered as feature vectors. This reduces the feature vector size from 256x256 to 2x2 per colour plane.

Figure 4.1 Feature Extraction for CBIR using Fractional Energy of Transformed Image
4.2.2 CBIR using Fractional Energy

The seven transforms are applied on the colour components of images to extract RGB feature sets as fourteen fractional coefficient sets. These fractional coefficient sets are then used for image retrieval. Table 4.1 gives comparative analysis of computational complexity of applying image transform to image of size NxN for all considered image transforms.

Table 4.1 Computational Complexity for applying transforms to image of size NxN

<table>
<thead>
<tr>
<th>Transform</th>
<th>Number of Additions For NxN image</th>
<th>Number of Multiplications</th>
<th>Total Additions for transform of 256x256 image</th>
<th>Computations Comparison (For 256x256 image)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCT</td>
<td>$2N^2(N-1)$</td>
<td>$N^2(2N)$</td>
<td>301858816</td>
<td>100</td>
</tr>
<tr>
<td>Walsh</td>
<td>$2N^2(N-1)$</td>
<td>0</td>
<td>33423360</td>
<td>11.07</td>
</tr>
<tr>
<td>Haar</td>
<td>$2N^2\log_2(N)$</td>
<td>0</td>
<td>1048576</td>
<td>0.35</td>
</tr>
<tr>
<td>Kekre</td>
<td>$N[N(N+1)-2]$</td>
<td>$2N(N-2)$</td>
<td>17882624</td>
<td>5.92</td>
</tr>
<tr>
<td>DST</td>
<td>$2N^2(N-1)$</td>
<td>$N^2(2N)$</td>
<td>301858816</td>
<td>100</td>
</tr>
<tr>
<td>Slant</td>
<td>$2N^2(N-1)$</td>
<td>$N^2(N)$</td>
<td>167641088</td>
<td>55.54</td>
</tr>
<tr>
<td>Hartley</td>
<td>$2N^2(N-1)$</td>
<td>$N^2(2N)$</td>
<td>301858816</td>
<td>100</td>
</tr>
</tbody>
</table>

[Observation: Here one multiplication is considered as eight additions for second last row computations and DCT computations are considered to be 100% for comparison in last row]

4.2.2.1 Feature Extraction for feature vector ‘T-RGB’

Here the feature vector space of the image of size NxNx3 has $NxNx3$ number of elements. This is obtained using following steps to get ‘T-RGB’

i. Extract Red, Green and Blue components of the color image.

ii. Apply the Transform ‘T’ on individual color planes of image to extract feature vector.

iii. The result is stored as the complete feature vector ‘T-RGB’ for the respective image.

Thus the feature vector database for DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform is generated as DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB respectively. Here the size of feature database is NxNx3.

C4-10
4.2.2.2 Feature Vector Database ‘Fractional T-RGB’

The fractional coefficients of transformed image as shown in figure 4.1, are considered to form ‘fractional T-RGB’ feature vector databases. Here first 50% of coefficients from upper triangular part of feature vector ‘T-RGB’ are considered to prepare the feature vector database ‘50%-T-RGB’ for every image as shown in figure 4.1. Thus DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB feature databases are used to obtain new feature vector databases as 50%-DCT-RGB, 50%-Walsh-RGB, 50%-Haar-RGB, 50%-Kekre-RGB, 50%-DST-RGB, 50%-Slant-RGB, 50%-DHT-RGB respectively. Then per image first 25% number of coefficients (as shown in figure 4.1) from the feature vectors database DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB are stored separately as feature vector databases as 25%-DCT-RGB, 25%-Walsh-RGB, 25%-Haar-RGB, 25%-Kekre-RGB, 25%-DST-RGB, 25%-Slant-RGB, 25%-DHT-RGB respectively. Then for each image in the database fractional feature vector set for DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB using 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% of total coefficients are formed. The comparison of sizes of these feature vectors for 256x256 image are given in table 4.2.

<table>
<thead>
<tr>
<th>Fractional Coefficients (in %)</th>
<th>Number of Elements in Feature Vector</th>
<th>Fractional Coefficients (in %)</th>
<th>Number of Elements in Feature Vector</th>
<th>Fractional Coefficients (in %)</th>
<th>Number of Elements in Feature Vector</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>65536</td>
<td>3.125</td>
<td>2048</td>
<td>0.097</td>
<td>64</td>
</tr>
<tr>
<td>50</td>
<td>32768</td>
<td>1.5625</td>
<td>1024</td>
<td>0.048</td>
<td>32</td>
</tr>
<tr>
<td>25</td>
<td>16384</td>
<td>0.7813</td>
<td>512</td>
<td>0.024</td>
<td>16</td>
</tr>
<tr>
<td>12.5</td>
<td>8192</td>
<td>0.39</td>
<td>256</td>
<td>0.012</td>
<td>8</td>
</tr>
<tr>
<td>6.25</td>
<td>4096</td>
<td>0.195</td>
<td>128</td>
<td>0.006</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 4.2 Feature vector size in proposed image retrieval using fractional coefficients of transformed image for image of size 256x256
4.2.2.3 Query Execution for ‘T-RGB’ CBIR

Here the feature set of size NxNx3 for the query image is extracted using transform ‘T’. This feature set is compared with each entry from the feature database using Euclidian distance as similarity measure.

Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform based feature sets are extracted from query image and are compared respectively with DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB feature sets using average Euclidian distance to find the best match in the database for respective transform based image retrieval method.

4.2.2.4 Query Execution for ‘Fractional T-RGB’

For 50%-T-RGB query execution, only 50% number of coefficients of upper triangular part of ‘T’ transformed query image (with NxNx3 coefficients) are considered for the CBIR and are compared with ‘50%-T-RGB’ database feature set for Euclidian distance computations.

Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform based feature sets are extracted from the query image and are compared respectively with 50%-DCT-RGB, 50%-Walsh-RGB, 50%-Haar-RGB, 50%-Kekre-RGB, 50%-DST-RGB, 50%-Slant-RGB, 50%-DHT-RGB feature sets to find average Euclidian distances. For 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% T-RGB based query execution, the feature set of the respective percentages are considered from the ‘T’ transformed NxNx3 image as shown in figure 4.1, to be compared with the respective percentage T-RTB feature set database to find average Euclidian distance.
4.2.3 Results of CBIR using Fractional Energy

For testing the performance of each proposed CBIR technique, per technique 55 queries (randomly selected 5 from each category) are fired on the image database. The query and database image matching is done using average Euclidian distance. The average precision and average recall are computed by grouping the number of retrieved images sorted according to ascending average Euclidian distances with the query image.

In all transforms, the average precision and average recall values for CBIR using fractional coefficients are higher than CBIR using full set of coefficients. The crossover point of precision and recall of the CBIR techniques acts as one of the important parameters to judge their performance.

Figure 4.2 shows the precision-recall crossover points plotted against number of retrieved images for proposed image retrieval techniques using DCT. Here DCT features 0.012% fractional feature set (1/8192th of total coefficients) based image retrieval gives highest precision and recall values.

Figure 4.3 shows the graphs of precision/recall values plotted against number of retrieved images for Walsh transform based image retrieval techniques. Here 1/4096th fractional coefficients (0.024% of total Walsh transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance using Walsh transform.
**Figure 4.2 Crossover Point of Precision and Recall for DCT Fractional Coefficients based CBIR.**

[Observation]: Here the highest precision-recall crossover point is given by 0.012% fractional coefficients (1/8192th of total coefficients) of DCT transformed image content, followed by 0.024% fractional coefficients and 0.006% fractional coefficients. All fractional coefficients based precision-recall crossover points are higher than all coefficients (All-Coeff) of DCT transformed image content.

**Figure 4.3 Crossover Point of Precision and Recall for Walsh transform Fractional Coefficients based CBIR.**

[Observation]: Here the highest precision-recall crossover point is given by 0.024% fractional coefficients (1/4096th of total coefficients) of Walsh transformed image content, followed by 0.012% fractional coefficients. The precision-recall crossover points in all fractional coefficients are higher than all coefficients (All-Coeff) of Walsh transformed image content.
Figure 4.4 shows the precision-recall crossover points plotted against number of retrieved images for proposed image retrieval techniques using Haar transform. Here 0.024% fractional feature set (1/4096th of total coefficients) based image retrieval gives highest precision and recall values.

![Figure 4.4 Crossover Point of Precision and Recall for Haar transform Fractional Coefficients based CBIR.](image)

**Observation**: Here the highest precision-recall crossover point is given by 0.024% fractional coefficients (1/4096th of total coefficients) of Haar transformed image content, followed by 0.78% fractional coefficients and 0.195% fractional coefficients. The precision-recall crossover points in all fractional coefficients are higher than all coefficients (All-Coeff) of Haar transformed image content.

Figure 4.5 gives average precision/recall values plotted against number of retrieved images for all Kekre-RGB image retrieval techniques. Here 1/32th fractional coefficients (3.125% of total Kekre transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Kekre transform.
Figure 4.5 Crossover Point of Precision and Recall for Kekre transform Fractional Coefficients based CBIR

Observation: Here the highest precision-recall crossover point is given by 3.125% fractional coefficients (1/32\textsuperscript{th} of total coefficients) of Kekre transformed image content, followed by 0.78% fractional coefficients and 12.5% fractional coefficients. All fractional coefficients based precision-recall crossover points are higher than all coefficients (All-Coeff) of Kekre transformed image content.

Figure 4.6 Crossover Point of Precision and Recall for DST Fractional Coefficients based CBIR

Observation: Here the highest precision-recall crossover point is given by 0.048% fractional coefficients (1/2048\textsuperscript{th} of total coefficients) of DST transformed image content, followed by 0.012% fractional coefficients. The precision-recall crossover points in all fractional coefficients are higher than all coefficients (All-Coeff) of DST transformed image content.
Precision/recall values for DST-RGB image retrieval techniques are plotted in figure 4.6. Here $1/2048^{th}$ fractional coefficients (0.048% of total DST transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Discrete Sine Transform.

Figure 4.7 gives average precision/recall values plotted against number of retrieved images for all Slant-RGB image retrieval techniques. Here $1/8^{th}$ fractional coefficients (12.5% of total Slant transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Slant Transform.

![Figure 4.7 Crossover Point of Precision and Recall for Slant transform Fractional Coefficients based CBIR](image)

[Observation : Here the highest precision-recall crossover point is given by 12.5% fractional coefficients ($1/8^{th}$ of total coefficients ) of Slant transformed image content, followed by 1.56% fractional coefficients and 0.195 % fractional coefficients. The energy compaction in Slant transform is not as good as in other image transforms.]

Precision/recall values for DHT image retrieval techniques are plotted in figure 4.8. Here $1/2$ fractional coefficients (0.50% of total DHT transformed coefficients) based image retrieval gives the
highest precision/recall crossover values specifying the best performance for using Discrete Hartley Transform.

![Figure 4.8 Crossover Point of Precision and Recall for Hartley Transform Fractional Coefficients based CBIR](image)

[Observation : Here the highest precision-recall crossover point is given by 50% fractional coefficients (1/2 of total coefficients) of Hartley transformed image content, followed by 6.25% fractional coefficients. The precision-recall crossover points in all fractional coefficients are higher than all coefficients (All-Coeff) of Hartley transformed image content]

Figure 4.9 shows the performance comparison of all the seven transforms for proposed CBIR techniques (DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, and DHT-RGB CBIR with different percentage of fractional coefficients). Here Slant-RGB CBIR outperforms all other transforms till 12.5% of coefficients as feature vector then Kekre-RGB CBIR outperforms all other transforms till 0.097% of coefficients as feature vector then Walsh-RGB CBIR takes over till 0.024% then DCT-RGB performs best for 0.012% of coefficients. In Walsh-RGB and Haar-RGB CBIR the feature vector with 0.024% of coefficients gives best performance, in DCT-RGB CBIR 0.012% of coefficients shows highest crossover value of average precision and average recall and Kekre transform gives the best performance when 6.25% of coefficients are
considered. In all, CBIR using Kekre transform with 6.25 % of fractional coefficients gives the best performance for among fractional energy based CBIR techniques discussed here.

![Figure 4.9 Performance Comparison of CBIR methods using Fractional Coefficients across all discussed Transforms](image)

[Observation]: Here for all image transforms the image retrieval methods considering fractional coefficients as feature vectors have given higher precision-recall crossover points as compared to consideration of all coefficients of transformed images as feature vector, proving that the discrimination capability of fractional coefficients of transformed image content is better than all coefficients of transformed image. In all considered image retrieval methods using fractional coefficients of transformed image content Kekre transform at 6.25% of fractional coefficients has given best performance indicated by highest crossover point value (0.42)]

Computational complexity and retrieval efficiency are the key objectives in the image retrieval system. Nevertheless it is very difficult to reduce the computations and improve the performance of image retrieval technique.

Here the performance of image retrieval is improved using fractional coefficients of transformed images at reduced computational complexity. Fairly large numbers of popular image transforms are
considered here with newly introduced Kekre transform. In all transforms (DCT, Walsh, Haar, Kekre, DST, Slant and DHT), the average precision and average recall values for CBIR using fractional coefficients are higher than CBIR using full set of coefficients. Hence the feature vector size for image retrieval could be greatly reduced, which ultimately will result in faster query execution in CBIR with better performance. In all Kekre transform with fractional coefficients (6.25%) gives best performance with highest crossover points of average precision and average recall. Feature extraction using Kekre transform is also computationally lighter as compared to DCT or Walsh transform. Thus feature extraction in lesser time is possible with increased performance.

The fractional coefficients gives better discrimination capability in CBIR than the complete set of transformed coefficients and image retrieval with better performance at much faster rate.

Till the feature vector size is quite large and the limitation of image retrieval using fractional energy of transformed image content is the need of all database images to be of same size (same and square dimensions i.e. number of rows and number of columns must be same) to be able to apply image transform on the image data for feature extraction. This limitation is partially taken out in section 4.3 where the feature vector dimension is further reduced using row mean of column transformed image, where all images in the database must have to have same number of rows but number of columns may vary.
4.3 Image Retrieval using Row Mean of Column Transformed Image Content

To take advantage of energy compaction property of image transforms, the earlier section 4.2 uses images in transformed domain for feature extraction in CBIR. But taking transform of image is time consuming. Reducing the size of feature vector by applying transform on columns of the image and finally taking row mean of transformed columns and till getting the improvement in performance of image retrieval is the theme of the work presented in form of innovative image retrieval techniques in this section as image retrieval using row mean of column transformed image content [152,153,248].

The seven assorted image transforms discussed in section 4.1 are considered to obtain seven variants of proposed image retrieval method. The techniques are compared with CBIR using full transformed image as feature vector. The results have shown the performance improvement (higher precision and recall values) with proposed methods compared to complete data of transformed image at reduced computations resulting in faster retrieval. The variations of considering DC component of transformed columns as part of feature vector and excluding it are also tested and it is found that presence of DC component in feature vector improvises the results in image retrieval. The ranking of image transforms for performance in proposed CBIR techniques with DC component starting from best can be listed as discrete Cosine transform, Haar transform, Walsh transform, Slant transform, discrete Sine transform, Hartley transform and Kekre transform.

4.3.1 Row Mean of Column Transformed Image

Figure 4.10 shows the Feature Extraction in Proposed content based image retrieval (CBIR) Technique with Row Mean of Transformed
Image Columns. Here image transform is applied on each column of image. Then row mean of the transformed columns is used as feature vector. The obtained feature vector is used in two different ways (with and without DC component) to see the variations in retrieval accuracy. As indicated by experimental results, image retrieval using DC component value proves to be better than retrieval excluding it.

![Feature Extraction in Proposed CBIR Technique with Row Mean of Transformed Image Columns](image)

**Figure 4.10 Feature Extraction in Proposed CBIR Technique with Row Mean of Transformed Image Columns**

### 4.3.2 CBIR using Row Mean of Column Transformed Image

The steps to be followed for image retrieval using the proposed image retrieval techniques can be given as

i. Apply transform $T$ on the column of image of size $N \times N$ ($I_{N \times N}$) to get column transformed image of the same size ($cI_{N \times N}$)

$$cI_{N \times N} = [T_{N \times N}] [I_{N \times N}]$$

ii. Calculate row mean of column transformed image to get feature vector of size $N$ (instead of $N^2$)

iii. The feature vector is considered with and without DC component to see variations in results. Then Euclidean Distance is applied to obtain precision and recall.

Applying transform on image columns instead of applying transform on the whole image, saves 50% of computations required resulting in faster retrieval. Again row mean of column transformed image is
taken as feature vector which further reduces the required number of comparisons among feature vectors resulting in faster retrieval.

The performance of proposed image retrieval techniques with DC component (referred as ‘Transform-RM-DC’) and without DC component (referred as ‘Transform-RM’) for each transform is compared with content based image retrieval (CBIR) using complete transformed image as feature vector (referred as ‘Full’), spatial row mean vector of image as feature vector (referred as ‘RM’, discussed in sections 3.1.1 and 3.1.3). For all image transforms the proposed image retrieval methods are performing better than the compared counterparts.

4.3.3 Results of CBIR using Colour Row Mean of Column Transformed Image

The performance comparison of content based image retrieval (CBIR) using transformed full image as feature vector (referred as ‘Full’), content based image retrieval (CBIR) using simple row mean feature vector of image (referred as ‘RM’) and the proposed content based image retrieval (CBIR) techniques with DC coefficient (referred as ‘Transform-RM-DC’) and without DC component (referred as ‘Transform-RM’) is given in Figure 4.11 to figure 4.17 in the form of crossover points of precision and recall obtained by applying all these image retrieval techniques on image database.

In all transforms it is observed that consideration of DC coefficient in feature vector improves the content based image retrieval (CBIR) performance (as indicated by higher precision and recall crossover point values).
Figure 4.11 Crossover Point of Precision and Recall v/s Number of Retrieved Images using Row mean of Column transformed image with DCT

[Observation]: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using DCT (DCT-RM-DC-P/R), which is higher than CBIR with feature vector as DCT applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance.

Figure 4.12 Crossover Point of Precision and Recall v/s Number of Retrieved Images using Row mean of Column transformed image with DST

[Observation]: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using DST (DST-RM-DC-P/R), which is higher than CBIR with feature vector as DST applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance.

C4-24
Observation: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using Walsh transform (WALSH-RM-DC-P/R), which is higher than CBIR with feature vector as Walsh transform applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance.

Observation: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using Haar transform (HAAR-RM-DC-P/R), which is higher than CBIR with feature vector as Haar transform applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance.
Figure 4.15 Crossover Point of Precision and Recall v/s Number of Retrieved Images using Row mean of Column transformed image with Slant Transform

[Observation]: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using Slant transform (SLANT-RM-DC-P/R), which is higher than CBIR with feature vector as Slant transform applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance

Figure 4.16 Crossover Point of Precision and Recall v/s Number of Retrieved Images using Row mean of Column transformed image with Hartley Transform

[Observation]: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using Hartley transform (HARTLEY-RM-DC-P/R), which is higher than CBIR with feature vector as Hartley transform applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance
Observation: Here the highest precision and recall crossover point value is obtained in proposed CBIR method with DC coefficient using Kekre transform (KEKRE-RM-DC-P/R), which is higher than CBIR with feature vector as Kekre transform applied to full image (Full) and row mean of the image (RM). Consideration of DC coefficient in feature vector helps in enhancing the image retrieval performance.

**Figure 4.17** Crossover Point of Precision and Recall v/s Number of Retrieved Images using Row mean of Column transformed image with Kekre Transform

Observation: Here in all considered image transforms the precision and recall crossover point value of CBIR using row mean of column transformed image are higher than that of CBIR using all transformed image coefficients as feature vector. This proves that the proposed techniques are better and faster. In all DCT gives best performance followed by Haar and Walsh transforms. However the performance variation across the transforms is marginal.

**Figure 4.18** Performance comparison of image retrieval methods using row mean of column transformed image with & without DC component for all transforms
To decide which image transform proves to be the best for proposed CBIR techniques, the crossover points of proposed image retrieval methods with and without DC coefficient are shown in figure 4.18. Here it is observed that the proposed CBIR techniques for all transforms are giving better performance in DC component consideration than neglecting it. Also in all transforms proposed colour content based image retrieval (CBIR) method with DC component outperforms the complete transform based CBIR technique. Here the best results are obtained using DCT-RM-DC followed by HAAR-RM-DC. The ranking of transforms for performance in proposed image retrieval techniques with DC component consideration can be given as discrete cosine transform (DCT), Haar transform, Walsh transform, Slant transform, discrete sine transform (DST), Hartley transform and Kekre transform. All transforms with proposed image retrieval technique are showing improvement in performance as compared to image retrieval based on complete transformed image as feature vector at great reduction in computational complexity, proving the proposed CBIR technique better and faster.

The herculean task of improving the performance of image retrieval and simultaneously reducing the computational complexity is achieved by proposed image retrieval technique using row mean of transformed column image. The performance of proposed techniques is compared with content based image retrieval (CBIR) using complete transformed image as feature vector and row mean of image as feature vector. Total seven image transforms like discrete cosine transform (DCT), discrete sine transform (DST), Haar transform, Hartley transform, Kekre transform, Walsh transform and Slant transform are considered. Experimental results show that in all transforms proposed content based image retrieval
(CBIR) technique with DC component outperforms other methods with great reduction in computation time. Consideration of DC component in proposed image retrieval techniques gives higher performance as compared to neglecting it.

The approach used for feature vector dimension reduction in image retrieval using fractional energy of transformed image is impudent. More sophisticated feature vector dimension reduction in image retrieval using energy compaction in transform domain is presented in section 4.4.

### 4.4 Image Retrieval using Energy Compaction in Transform Domain

This section introduces proposed image retrieval methods based on energy compaction in transform domain. All the transforms have characteristic to compress most of the signal energy in low frequency region in transform domain. In transform domain very few high frequency coefficients do contain most of the energy, these small number of coefficients if considered as feature vector gives performance improvement in image retrieval as compared to consideration of all coefficients as feature vector with tremendously reduced complexity of query execution. Reducing the feature vector size in image retrieval using these low frequency (high energy) components with retention of signal energy to 98%, 96% and 94% is the basis of the work elaborated in this section. Instead of randomly selecting few starting transform domain coefficients as feature vector, average energy and cumulative energy vectors do statistically help in finding the number of coefficients for retaining some percentage of energy. Here image retrieval techniques are proposed by considering the row mean, column mean, diagonal means elaborated in section 3.1.1 in transform domain with energy compaction using seven assorted image transforms [154, 156, 157].