GLOBAL AND LOCAL DESCRIPTOR FOR CBIR AND IMAGE ENHANCEMENT USING MULTI-FEATURE FUSION METHOD

Devrat Arya *1, Jaimala Jha 2

*1 Department of CSE/IT, MITS, Gwalior, INDIA
2 Professor, Department of CSE/IT, MITS, Gwalior, INDIA

DOI: https://doi.org/10.29121/granthaalayah.v4.i6.2016.2651

ABSTRACT

The research is ongoing in CBIR it is getting much popular. In this retrieval of image is done using a technique that searches the necessary features of image. The main work of CBIR is to get retrieve efficient, perfect and fast results. In this algorithm, fused multi-feature for color, texture and figure features. A global and local descriptor (GLD) is proposed in this paper, called Global Correlation Descriptor (GCD) and Discrete Wavelet Transform (DWT), to excerpt color and surface feature respectively so that these features have the same effect in CBIR. In addition, Global Correlation Vector (GCV) and Directional Global Correlation Vector (DGCV) is proposed in this paper which can integrate the advantages of histogram statistics and Color Structure Descriptor (CSD) to characterize color and consistency features respectively. Also, this paper is implemented by Hu moment (HM) for shape feature, it extract 8 moments for image. For the classification process, apply kernel Support vector machine (SVM). The experimental result has computed precision, recall, f_measure and execution time. Also, worked on two datasets: Corel-1000 and Soccer-280.

Keywords:
Image retrieval, HSV color space, Global Correlation Vector, DWT, DGVC, SVM.


1. INTRODUCTION

CBIR is a method to search and also index images in the massive set of database based on its visual contents, for example the basis of textures, colors, shapes or spatial layouts in its place of applying tags or various descriptive metadata keywords that might associate with images in database [1]. In traditional CBIR systems work is done through extracting one or extra multi-dimensional vectors from all image in database, this procedure is complete in a posterior step to start retrieving. At query time, the few vectors are typically extracted from query image and a similarity based function is used then to measure the quantity of variance between different
images and query image vector in the database. Those images have a similarity vector to query vectors are the set of final retrieved result.

CBIR finds its applications in numerous domains for example medical diagnostics, GIS and military applications, pattern recognition, computer vision and numerous others [2]. However, in most applications CBIR systems are essentially depending on extracting few features, i.e. characteristics that can capture certain visual properties of an image either internationally for the complete image or nearby for its regions [3] [4] and significant features that can efficiently discriminate images and help in matching the most similar ones is the most challenging issue in CBIR systems. Color features are extensively used in CBIR systems as they are independent of image size and orientation [5]. They are typically extracted from various color spaces, e.g. RGB, HSV, YCbCr, through computing the color histogram, color moments or dominant colors.

2. METHODS USED IN THIS APPROACH

2.1. IMAGE PYRAMID

The image pyramid is designed on data structure to support effective scaled convolution by decrease image representation. It contains different original copies of an image in sequence with both decrease sample density and resolution in regular steps [6]. Stages of the lower resolution pyramid are achieved themselves by using the highly effective iterative algorithm. The zero or lower level of pyramid G0 is equal to original image.

2.2. COLOR SPACE AND COLOR QUANTIZATION

2.2.1. COLOR SPACE

Color is natural component and perform necessary role in CBIR systems. The Hue, Saturation and Value (HSV) give the perception representation value according to the human visual feature. The HSV mannequin defines a colour area in terms of range i.e for Hue stages from zero to 360, saturation levels from 0 to a 100% with vibrancy of colour and it's known as purity. Worth ranges from zero to one 100% with color brightness[7].

2.2.2. COLOR QUANTIZATION

Color quantization is a procedure that decreases various distinct colors used in an image. In this a novel image much visually similar as probable to the original image. For a true color image, the different colors are up to $2^{24} = 16777216$, so the color feature extraction from the true color will prompt to a large computation. To reduce the computation without affecting image quality, some color is extracted to represent the image, by this processing speed, improve and reduces the storage space [8]. Several authors reported the effect of color quantization on image retrieval performance [9], with different quantization schemes, like RGB (8X8X8), Lab (4X8X8), HSV (16X4X4), Lu*v* (4X8X8).
2.3. GLOBAL CORRELATION VECTOR

Color histogram is a general method to extract color feature of an image and has great performance when applied to image recognition. With this technique, it is to calculate frequency of occurrence of each color[10]. GCV is proposing to extract color feature in image pyramid. The distributions of colors and the spatial correlation between pair of colors can be extracted simultaneously.

2.4. DISCRETE WAVELET TRANSFORM

The DWT is based on coding of sub-band, is to create to yield a Wavelet Transform fast computation. It is simply to implement and decreases the computation time and resources need. The 2-D DWT of size N1 x N2 of image function s(n1,n2) may be expressed as [11]

\[
W_{\varphi}(j_0,k_1,k_2) = \frac{1}{\sqrt{N_1N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} s(n_1,n_2) \varphi_{j_0,k_1,k_2}(n_1,n_1)
\]

\[
W_{i}(j_0,k_1,k_2) = \frac{1}{\sqrt{N_1N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} s(n_1,n_2) \psi_{i,j_0,k_1,k_2}(n_1,n_1)
\]

Where \( i = \{H, V, D\} \) indicate the wavelet function direction index. As in 1-D case, \( j_0 \) shows any starting scale, which may be treated as \( j_0=0 \). Above two different equations are 2-D DWT.

\[\text{Figure 1: Discrete wavelet sub-band decomposition.}\]

The DWT [6] comparing to other transforms is characteristics of frequency or time used in various application fields. The DWT Sub band flow chart for digital image is present in figure 1, where L represents to component of low frequency, H represents to high frequency and the number 1 and 2 represents to decomposition level of the DWT. The sub image LL is element of low frequency, it is the approximate usual image sub image; sub image HL is the component of low frequency in horizontal course and the high in vertical direction, it manifests the long-established image horizontal side. LH sub image is the factor of high frequency in horizontal course and in vertical direction, it manifests the original image vertical edge; sub image HH is
the component of high frequency, it manifests the original image oblique edge. It is present that original image energy is contained in the LL2 low frequency region.

2.5. DGCV

In GCV technique the texture features of an image information can be lost directly, so a method called DGVC is proposed to extract texture feature of filtered images. In this technique the two neighbors’ text in filtered orientation are considered to compare with the central pixels.

2.6. SVM

SVM is a supervised learning approach used in organization and reversion analysis. An SVM training algorithm constructs a model that can classify whether a new sample go down into that category in which the system is trained for the database.

3. LITERATURE REVIEW

Xiang-Yang[13] In this paper, a CBIR method has been proposed which uses the combination of Zernike chromaticity distribution moments and rotation-scale invariant Contourlet texture descriptor. The proposed approach outcome yielded bigger retrieval accuracy than different approaches and not using a better function vector dimension. Also the proposed method shows better performance gain in average normal precision, average normal recall, and average retrieval time over the other methods.

Su dipta.et.al[14] A novel approach is proposed for texture image retrieval. This approach can be used to overcome the bottleneck of simple distance based image retrieval. The approach is tested using three different databases of varying size, orientation, complexity and number of texture class. Performance of this approach is compared with other promising distance based as well as classifier based retrieval methods.

D. Feng et al.[15] In this paper proposed an efficient indexing technique for CBIR. The proposed technique introduces the ordered quantization to growth distinction among descriptors of quantized feature. Thus, feature point correspondences can be defined through quantized feature descriptors, and they are used to similarity measure between query image and database image.

4. PROPOSED METHODOLOGY

1) Consider color query image ‘I’ as an N X M size of an image and extract the Red, Green, and Blue Components from an image.
2) Apply adaptive histogram on query image for enhancing the image.
3) Apply image pyramid reduction, initialize kernel center weight to 0.375 for decreasing the size of an image.

\[ Img_{filtered} = \sum_{i=1}^{3} Blur\_img(I, kernel) \]
Where $\text{img}_{\text{filtered}}$ is pyramid reduced image, Blur_img is blurred image, kernel is kernel center weight.

4) Apply GCV by assigning 8 levels each to hue, saturation and value provide a quantized HSV space with 8x8x8=512 histogram bins. And also find color structure descriptor.

5) In the feature extraction process, extract features of an image using DGCV (entropy, energy, homogeneity and contrast).

$$\text{hom} = \sum \frac{p}{\text{greydiff}^2 + 1}$$
where greydiff=[0:1], P is probability function and hom denotes to homogeneity.

$$\text{con} = \sum (P \times \text{greydiff}^2 + 1)$$
Where con is contrast of an image.

$$\text{eng} = \sum (P)^2$$
Where eng denotes energy of an image.

$$\text{ent} = -\sum P \times \text{log}(P + \text{eps})$$
Where and denotes the entropy of an image.

$$B1 = \sum P \times \text{greydiff}$$
Where B1 is mean value of image.

6) Apply Discrete Wavelet Transform at 1st level to get estimated coefficient and vertical, horizontal and diagonal detail coefficients.

7) Calculate moment $(p,q)$ of an image $f(x,y)$ of size $M*N$ is defined using this equation:

$$m_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) x^p y^q$$
where $p$ is the order of $x$ and $q$ is the order of $y$.

8) A central moment is basically the same as the moments just described except that the values of $x$ and $y$ used in the formulas are displaced by the mean values

$$\beta_{p,q} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) (x - x_{\text{avg}})^p (y - y_{\text{avg}})^q$$
Where $x_{\text{avg}} = \frac{m_{10}}{m_{00}}$ and $y_{\text{avg}} = \frac{m_{01}}{m_{00}}$.

9) The normalized moments $\gamma_{p,q}$ are the same as the central moments except that they are all divided by an appropriate power of $m_{00}$. 
10) Hu invariant moments are linear combinations of the central moments and here is how are defined eight Hu moments:

\[ I_1 = \gamma_20 + \gamma_02 \]
\[ I_2 = (\gamma_20 - \gamma_02)^2 + 4 \times \gamma_{11} \]
\[ I_3 = (\gamma_30 - 3 \times \gamma_{12})^2 + (\gamma_{03} - 3 \times \gamma_{21})^2 \]
\[ I_4 = (\gamma_30 + \gamma_{12})^2 + (\gamma_{03} + \gamma_{21})^2 \]
\[ I_5 = (\gamma_30 - 3 \times \gamma_{12}) \times (\gamma_30 + \gamma_{12}) \times ((\gamma_30 + \gamma_{12})^2 - 3 \times (\gamma_{21} + \gamma_{03})^2) + (3 \times \gamma_{21} - \gamma_{03}) \times (\gamma_{21} + \gamma_{03}) \times (3 \times (\gamma_30 + \gamma_{12})^2 - (\gamma_30 + \gamma_{12})^2) \]
\[ I_6 = (\gamma_20 - \gamma_02) \times ((\gamma_30 + \gamma_{12})^2 - (\gamma_{21} + \gamma_{03})^2) + 4 \times (\gamma_30 + \gamma_{12}) \times (\gamma_{21} + \gamma_{03}) \]
\[ I_7 = (3 \times \gamma_{21} - \gamma_{03}) \times (\gamma_30 + \gamma_{12}) \times ((\gamma_30 + \gamma_{12})^2 - 3 \times (\gamma_{21} + \gamma_{03})^2) + (\gamma_{03} - 3 \times \gamma_{21} \times (\gamma_{21} + \gamma_{03}) \times (3 \times (\gamma_30 + \gamma_{12})^2 - (\gamma_30 + \gamma_{12})^2) \]
\[ I_8 = (\gamma_{11} \times (\gamma_30 + \gamma_{12})^2 - (\gamma_30 + \gamma_{12})^2 - (\gamma_20 - \gamma_02) \times (\gamma_30 + \gamma_{12}) \times (\gamma_{21} + \gamma_{03}) \]

11) Repeat step 2 to step 9 on a query image within the database.

12) Determine the similarity matrix of query image and image database using Euclidean distance, L1 distance and Weighted L1 distance. We have used Euclidean distance which is the most predictable metric for calculating the lack of involvement between two vectors. Given two vectors Q and T, where

\[ d(Q, T) = \sum_m \sum_n d_{mn}(Q, T) \]

Where

\[ d_{mn} = \frac{|(\mu_{mn}^Q - \mu_{mn}^T)|}{|\mu_{mn}^Q| + |\mu_{mn}^T|} + \frac{|(\mu_{mn}^Q - \mu_{mn}^T)|}{|\mu_{mn}^T| + |\mu_{mn}^Q|} \]

13) Repeat the steps from 2 to 9 for all the images in the database.

14) Classify the images using SVM classifier and combine global and local features.

\[ f_{QUERY} = (f_G, f_L) \]

15) Calculate precision, f-measure, execution time and recall of retrieved images.

\[ Precision = \frac{\text{No. of relevant image retrieved}}{\text{Total number of image retrieved}} \]

\[ Recall = \frac{\text{No. of relevant image retrieved}}{\text{number of image in the database}} \]

\[ F_{\text{measure}} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \]
5. RESULT ANALYSIS

We choose the other two datasets, namely Corel-1000 and Soccer-280, to compare GLD with these image feature descriptors. There are 8 images chosen randomly for each category as query images in the Corel dataset and 7 images from Soccer dataset. Graph1 show the experimental results of comparison among GLD and the other descriptor in these datasets. It can be seen that the GLD algorithm have better performance than previous algorithm.
Figure 3: Corel -1000 Dataset

Figure 4: Soccer-280 Dataset

Figure 5: Retrieval result on African Dataset
The performance of retrieval system can be measured in terms of its recall and precision. Don't forget measure the method capacity to retrieve every the units which can be central, whilst precision measures the approach potential to retrieve only models which can be valuable.

**Table 1:** Shows base and proposed result on precision and recall for different images

<table>
<thead>
<tr>
<th>Category</th>
<th>Base Precision (%)</th>
<th>Proposed Precision (%)</th>
<th>Base Recall (%)</th>
<th>Proposed Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>African 1.jpg</td>
<td>82.05</td>
<td>91.17</td>
<td>64</td>
<td>62</td>
</tr>
<tr>
<td>Beach 123.jpg</td>
<td>71.05</td>
<td>83.33</td>
<td>54</td>
<td>70</td>
</tr>
<tr>
<td>Building 222.jpg</td>
<td>72.97</td>
<td>77.41</td>
<td>54</td>
<td>48</td>
</tr>
<tr>
<td>Bus 343.jpg</td>
<td>73.33</td>
<td>82.85</td>
<td>66</td>
<td>58</td>
</tr>
<tr>
<td>Elephant 511.jpg</td>
<td>80</td>
<td>85.71</td>
<td>56</td>
<td>60</td>
</tr>
<tr>
<td>Flower 634.jpg</td>
<td>78.95</td>
<td>86.48</td>
<td>60</td>
<td>64</td>
</tr>
<tr>
<td>Mountain 804.jpg</td>
<td>78.26</td>
<td>87.50</td>
<td>72</td>
<td>70</td>
</tr>
<tr>
<td>Food 991.jpg</td>
<td>77.42</td>
<td>80</td>
<td>48</td>
<td>64</td>
</tr>
</tbody>
</table>

In Table 1 compare the precision and Recall of Base and Proposed work for Corel-1000 dataset. The Precision of Base has reached upto 82.05% whereas the proposed is 91.17%.
Table 2: Shows base [10] and proposed result on F-measure and Time for different images

<table>
<thead>
<tr>
<th>Category</th>
<th>Base F-measure</th>
<th>Proposed F-measure</th>
<th>Base Time</th>
<th>Proposed Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>African 1.jpg</td>
<td>0.26</td>
<td>0.28</td>
<td>3.58</td>
<td>2.92</td>
</tr>
<tr>
<td>Beach</td>
<td>0.21</td>
<td>0.26</td>
<td>2.78</td>
<td>2.67</td>
</tr>
<tr>
<td>Building</td>
<td>0.20</td>
<td>0.27</td>
<td>2.92</td>
<td>2.61</td>
</tr>
<tr>
<td>Bus</td>
<td>0.25</td>
<td>0.26</td>
<td>2.78</td>
<td>2.60</td>
</tr>
<tr>
<td>Elephant</td>
<td>0.21</td>
<td>0.22</td>
<td>2.82</td>
<td>2.61</td>
</tr>
<tr>
<td>Flower</td>
<td>0.241</td>
<td>0.24</td>
<td>2.84</td>
<td>2.78</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.26</td>
<td>0.27</td>
<td>2.75</td>
<td>2.61</td>
</tr>
<tr>
<td>Food</td>
<td>0.18</td>
<td>0.25</td>
<td>2.82</td>
<td>2.61</td>
</tr>
</tbody>
</table>

In Table 2 compare the F_measure and Time of Base and Proposed work for Corel-1000 dataset. The F_measure of Base has reached upto 0.26 whereas the proposed is 0.27. Time is also decreased as compared to previous algorithm.

Table 3: Shows base [10] Precision result on Similarity measure

<table>
<thead>
<tr>
<th>Similarity Measure</th>
<th>Base Precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>L1</td>
<td>73.17</td>
</tr>
<tr>
<td>Euclidean</td>
<td>88.24</td>
</tr>
<tr>
<td>Weighted L1</td>
<td>70.73</td>
</tr>
</tbody>
</table>

Table 3 to Table 6 compare the precision and recall between Base and Proposed system, the total number of the retrieved images are fix from 10 to 30 in the experiments. It can be seen that the L1 distance performs better than other similarity measures and it is much more computationally efficient. Euclidean distance is typically used similarity measures, but now not normally the quality one considering the fact that the distances put too much emphasis on features which are generally distinctive. Weighted L1 distance can be considered as a weighted L1 distance with different weights.

Table 4: Shows Proposed Precision result on Similarity measure

<table>
<thead>
<tr>
<th>Similarity Measure</th>
<th>Proposed Precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
</tr>
<tr>
<td>L1</td>
<td>81.57</td>
</tr>
<tr>
<td>Euclidean</td>
<td>92.59</td>
</tr>
<tr>
<td>Weighted L1</td>
<td>79.06</td>
</tr>
</tbody>
</table>
Table 5: Shows Base Recall result on Similarity measure

<table>
<thead>
<tr>
<th>Similarity Measure</th>
<th>Base Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>60 46 66 68 68</td>
</tr>
<tr>
<td>Euclidean</td>
<td>60 68 60 62 60</td>
</tr>
<tr>
<td>Weighted L1</td>
<td>58 54 72 66 60</td>
</tr>
</tbody>
</table>

Table 6: Shows Proposed Recall result on Similarity measure

<table>
<thead>
<tr>
<th>Similarity Measure</th>
<th>Proposed Recall (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>62 54 52 58 68</td>
</tr>
<tr>
<td>Euclidean</td>
<td>50 66 70 54 56</td>
</tr>
<tr>
<td>Weighted L1</td>
<td>68 52 50 62 58</td>
</tr>
</tbody>
</table>

Table 7: Shows Proposed result on F-measure and Time for Soccer dataset on each category

<table>
<thead>
<tr>
<th>Category</th>
<th>Proposed Time</th>
<th>Proposed Precision (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acemilan</td>
<td>0.42</td>
<td>20.83</td>
</tr>
<tr>
<td>Barcelona</td>
<td>0.45</td>
<td>33.33</td>
</tr>
<tr>
<td>Chelsea</td>
<td>0.44</td>
<td>23.33</td>
</tr>
<tr>
<td>Juventus</td>
<td>0.447</td>
<td>50</td>
</tr>
<tr>
<td>Liverpool</td>
<td>0.455</td>
<td>40</td>
</tr>
<tr>
<td>Madrid</td>
<td>0.445</td>
<td>38.88</td>
</tr>
<tr>
<td>Psv</td>
<td>0.47</td>
<td>36.36</td>
</tr>
</tbody>
</table>

Graph 1: Comparison of Precision between base and Proposed Approach
6. CONCLUSION

On the basis of previous researches, the paper explored low-level features of shape, color and texture extraction of CBIR. After matching the CBIR founded on the shape, color and texture categories with that of the texture and color fused features, it is observed results of shape, color and texture fused categories are more robust than the texture and color features based image recovery. The investigational results have demonstrated that the DGCV algorithm is much more robust and discriminative than other image descriptors in CBIR. The investigational results show good precision up to 91.17% as matched to previous techniques. Furthermore, work on Soccer dataset for improving precision and recall, increase dataset images for enhance effectiveness of the system.

7. REFERENCES


