EMITTER International Journal of Engineering Technology
Vol. 7, No. 1, June 2019
ISSN: 2443-1168

Content-Dependent Image Search System for Aggregation of Color, Shape and Texture Features

Arvita Agus Kurniasari, Ali Ridho Barakbah, Achmad Basuki

Politeknik Elektronika Negeri Surabaya
Kampus PENS, Jl. Raya ITS, Sukolilo Surabaya
E-mail: arvita.agus88@gmail.com, {ridho, basuki}@pens.ac.id

Abstract

The existing image search system often faces difficulty in finding an appropriate retrieved image corresponding to an image query. The difficulty is commonly caused by the users' intention for searching image is different with dominant information of the image collected from feature extraction. In this paper, we present a new approach for the content-dependent image search system. The system utilizes information of color distribution inside an image and detects a cloud of clustered colors as something - supposed as an object. We apply segmentation of an image as a content-dependent process before feature extraction in order to identify is there any object or not inside an image. The system extracts 3 features, which are color, shape, and texture features and aggregates these features for similarity measurement between an image query and image database. HSV histogram color is used to extract the color feature of the image. While the shape feature extraction used Connected Component Labeling (CCL) which is calculated the area value, equivalent diameter, extent, convex hull, solidity, eccentricity, and perimeter of each object. The texture feature extraction used Leung Malik (LM)'s approach with 15 kernels. For applicability of our proposed system, we applied the system with benchmark 1000 image SIMPLIcity dataset consisting of 10 categories namely Africans, beaches, buildings historians, buses, dinosaurs, elephants, roses, horses, mountains, and food. The experimental results performed 62% accuracy rate to detect objects by color feature, 71% by texture feature, 60% by shape feature, 72% by combined color-texture feature, 67% by combined color-shape feature, 72 % combined texture-shape features and 73% combined all features.

Keywords: Image feature extraction, Content-dependent image search, image search system, Feature aggregation.

1. INTRODUCTION

All technology is born to various purposes. For example, search engines are created to sort through massive amounts of data online. In each new technology, improvement is obtained by combining existing technology to create something new that is better than the technology used before. Digital information is increasing because of globalization which eliminates distance, space and time. That makes the level of information accuracy has an important role in the search process. The image search process on big data made a problem that is not easy to solve [1]. One solution that can be applied is Content-Based Image Retrieval (CBIR). CBIR increased the accuracy and efficiency of the image search system and managed large amounts of image data [2].

The term Content-Based Image Retrieval (CBIR) is purported experiments on automatic retrieval of images from a database by color and shape features. Although CBIR took features that could be either primitive or semantic, the feature extraction process must be able to identify dominant content and images [3]. IBM QBIC (Query by Image Content) is proposed methods to query large online image database used the content of the image as the basis of the queries [4]. The Photobook system is an interactive tool for browsing and searching image and image sequence that they made direct used of the image content rather than relying on text annotations [5]. SIMPLIcity is an image retrieval system used the semantic classification method with a wavelet-based approach for feature extraction and integrated region matching based on image segmentation [6].

CBIR aims to measure performance in getting images similar to search schemes. But the system cannot capture image content properly. Because the existing system focuses on the selection of extraction features and methods [7].

This paper developed a new approach about content-dependent image search system for aggregation of color, shape and structure features. The system introduced color feature extraction using RGB and HSL then it normalized the unique data. The main idea of this approach using a clustering technique that is Hierarchical K-means and Optimal K. And the results are used to detect objects. Then we extract the feature from a similar image with the desired image.

2. RELATED WORKS

Previous researcher has conducted research for image search [8][9][10][11][12][13][14]. In general, image search in its application involved feature extraction which includes a combination of color, shape and texture features [8][9][10][11][12]. Detecting objects on image search is used to approach content between user preferences and people in the database [13][14].

Ali Ridho Barakbah [8] has made a previous study of image search with feature extraction of colors, shapes, and textures with automatic weighting. Extract their color features using 3-Dimensional Color Vector Quantization. While feature extraction used eccentricity, area, same diameter, and convex area. Structural features are extracted using the Curvelet method. The key to this research [9] is the automatic weighting mechanism when selecting features based on a combination of color, shape and structure features. The researcher applies an automatic weighting mechanism to select features by analyzing the distribution of color information to determine representative features. The researcher extracts the color moment in the image and calculates the color distance for color weight and texture density for structural weights. The shape feature is measured to extract the shape area then the area is calculated to adjust the shape of the texture density to determine the shape weights.

Praheep Anantharatsamy [10] has proposed a content-based image retrieval based on three major types of visual information: color, shape, and texture. Comparison of distance calculations for all three space distance dimensions of retrieval. In the experimental results, they investigated several extraction feature methods and search algorithms from the content-based image retrieval. The results of the extraction feature selection obtained the highest accuracy based on the nearest 5-neighbor.

Naveena A K [11] has proposed a CBIR system based on color moments, wavelet and edge description to take desired images from the database. Extracting features used are color, texture, and shape. For color features used color moments, the texture feature used wavelet transform and the shape feature used edge histograms using canny edge detection. Comparison of the results of retrieval of each color, texture, and shape and comparison of features combined show that the combination of features provides better retrieval results than each feature.

K. Mala [12] has proposed a technique to produce image content descriptions with three features, there are: Color auto-Correlogram, Gabor Wavelet, and Wavelet Transform. The feature extraction process is based on the input request image of the IDB and features stored in the feature dataset. Manhattan distance is applied to users who are given query images and feature vectors calculated from database images to measure similarity. Another efficient feature such as Wavelet transformation for edge extraction, the shape is also used in conjunction with color and texture features to provide better results and can be used to obtain high-precision shooting. From the analysis conducted on the experimental results, it is illustrated that the proposed method achieves a better level of precision compared to other methods available. Thus the features taken from the proposed method achieved an average accuracy rate of 83% for the Corel database, while 88% for the Li database and 70% for the Caltech-101 database in the image search system.

Andrea Kutics [13] has described a new method for detecting objects in natural images with an unlimited domain. Which aims to capture important information in terms of assessment in terms of user semantics at the visual level for efficient shooting. The main obstacle to developing this method is the difficulty of accurately segmenting images into prominent areas. To overcome this difficulty, they developed a vector inhomogenous diffusion model that used several features.

Aamer Mohamed [14] has proposed an efficient content-based image retrieval with a scheme based on semantic object detection (SOD). The feature extraction process used a discrete cosine transform (DCT) blocks where retrieval results from query images use the calculation of a histogram-based approach. SOD is used with the aim of reducing the size of the database from the results of the drawing approach taken. By using SOD, it displays the improve retrieval results.

3. ORIGINALITY

The existing image search system often has difficulty finding images taken according to the image request. Difficulties are generally caused by the user's intention to look for different images with the dominant information of images collected from feature extraction. In this paper we present a new approach to image search systems that depend on content. This system uses color distribution information in images and detects color clouds that are grouped as something that is considered an object. We apply image segmentation as a process that depends on content before feature extraction to identify whether there are objects or not in the image. The system extracts 3 features, such as features of color, shape and texture and combines these features for measuring similarities between image requests and image databases. The HSV color histogram is used to extract image color features. While the form feature extraction uses Connected Component Labeling (CCL) which calculates area values, equivalent diameter, area, convex hull, solidity, eccentricity, and circumference of each object. The texture feature extraction uses the Leung Malik (LM)’s approach with 15 kernels.

4. SYSTEM DESIGN

Figure 1 shows the system design that we proposed research. There are 5 main functions in our system, there are: (1) preprocessing, (2) clustering, (3) object detection, (4) feature selection and similarity measurement. Each process is explained in section 4.1-4.5.

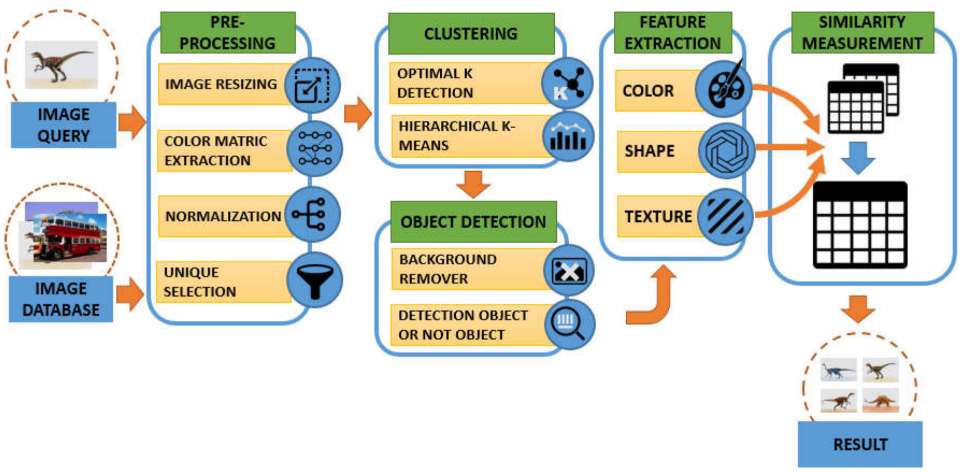


Figure 1. The System Design of our propose research

4.1 Pre-processing

Image preprocessing is the first stage of detection in order to improve the quality of images with color metric extraction and normalization. Color Metric Extraction reduced computational burdens and quantization of color can be used to represent images without significantly reducing image quality [15].

The color space included RGB and HSL. RGB is a color space which comprises the red, green, and blue spectral wavelength. The most frequent presentation of colors in image processing is RGB. Since RGB color space has some limitation in high-level processing, other color space representations have been developed [16]. HSL is known to improve the color system of HSV because it could present brightness better than saturation. Besides, the hue component in HSL color space is integrated with all chromatic information, it's stronger than the main color for image color segmentation [17].

The value between RGB and HSL is not the same, that makes this paper applied normalization. For the case of studies, used the softmax algorithm. The Softmax algorithm could achieve maximum and minimum values, but according to the specified limit. Transformations using Softmax are more or less linear in the middle range and have nonlinear fluency at both ends. The output range is between 0 and 1 [18].

$$newdata = \frac{1}{(1+e^{(-transfdata)})} \tag{1}$$

$$transfdata = \frac{(data - mean)}{(x * (\frac{std}{(2*3.14)}))} \tag{2}$$

Where:

x = linear response in standard deviation

4.2 Clustering

Object detection process involved the process of finding the best number of clusters from resizing images and clustering from original size images with getting optimal K and Hierarchical K-means. Optimal K Detection could find out the global optimal solution. Moving variance is defined as a variant in the cluster when determining calculated clusters and assesses that it has reached a globally optimal solution based on trends [19][20]. With experiments using widely available test data, comparing cluster results from this technique and existing non-hierarchical clustering techniques show the predominance of techniques. The ideal cluster has a minimum variance within the cluster (v_w) that represents internal homogeneity and maximum variance between clusters (v_b) which states external homogeneity.

$$v_w = \frac{1}{N-k}\sum_{i=1}^{k}(n_i - 1) \cdot v_i^2 \tag{3}$$

$$v_b = \frac{1}{k-1}\sum_{i=1}^{k} n_i(\overline{d_i} - \overline{d})^2 \tag{4}$$

$$v = \frac{v_w}{v_b} \tag{5}$$

Where:

N = Amount of all data
n_i = Number of cluster data *i*
v_i = Variant of cluster *i*

The best clusters that have been obtained are processed using Hierarchical K-means. In [21], they optimized the initial centroid for K-means. This utilizes all the results of grouping K-means at certain times, although some of them reach local optima. Then, the results are combined with the Hierarchical algorithm to determine the initial centroid for K-means. The algorithm is explained as follows:

Algorithm Hierarchical K-Means
1. Determine K as the initial cluster number. 2. Determine p as the amount of computing. 3. Determine $i = 1$ as the increase value. 4. Apply the K-Means algorithm. 5. Record the centroid from the cluster results as $C_i = [c_{ij} \| j = 1, \dots, K]$. 6. Add $i = i + 1$. 7. Repeat step five if $i < p$. 8. Assume $C = \{c_i \| i = 1, \dots, p\}$ as a data set, with the value K as the number of clusters to be formed. 9. Apply hierarchical clustering algorithms. 10. Save the centroid value from the results of clustering as $D = [d_i \| i = 1, \dots, K]$.

Then, $D = \{d_i | i = 1, \dots, K\}$as the initial initialization of the clusters mean value from K-means clustering.

4.3 Object Detection

The object detection in CBIR that made is one of the uniqueness of this research. In the object detection stage is involved the process of background remover and determining the object.

4.4.1 Background Remover

The background remover is one technique between the background images and image objects. This technique could be done by adjusting the object in the middle of the background image. The process is based on the results of the position normalization cluster values, RGB and HSL. The following is the removing technique that has 6 steps.

1. Initialize and specify tpmcluster value with value 1. It is intended to be used as a place to store the results of removing the cluster.

2. In Figure 2, the searching process is carried out by taking 25% of the image width calculated from the 0th column and 25% from the image height calculated from the 0th row. After that, changes in cluster value will be made with the value 0 based on the searching for the cluster value with the highest number.

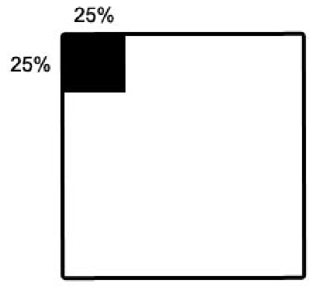


Figure 2. Second stage illustration

3. In Figure 3 we do the same step with the second stage which is different, such as the location of the searching for the cluster value with the highest number. The searching is done by taking 25% of the image width calculated from the last column and 25% and the image height is calculated from the 0th row.

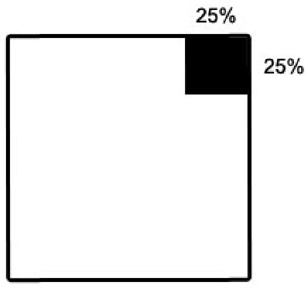


Figure 3. Third stage illustration

4. In Figure 4 we do the same step with the third stage which is different, such as the location of the searching for the cluster value with the highest number. The searching is done by taking 25% of the image width calculated from the 0th column and 25% and the image height calculated from the last row.

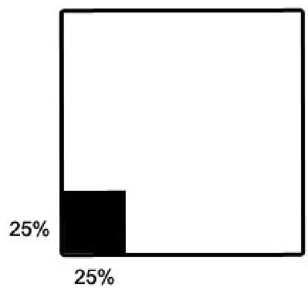


Figure 4. Fourth stage illustration

5. In Figure 5 we do the same step with the fourth stage which is different such as the location of the searching for the cluster value with the highest number. The searching is done by taking 25% of the image width calculated from the last column and 25% and the image height calculated from the last row.

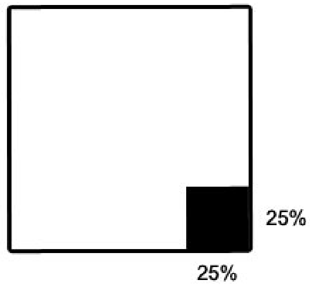


Figure 5. Fifth stage illustration

6. At the last stage, the technique is carried out which is to searching for cluster values that are not equal to 0 and then it find the cluster value with the highest number.

4.4.2 Determination of Objects

The last part of the object detection stage is the process of determining results, including objects or not objects. The determination is done by calculating the ratio of objects to the background.

It detects a cloud of clustered colors as something - supposed as an object if the number of object ratios is greater in the ratio than the number of outside ratios. The inner ratio is obtained from 25% -75% of the image and the outer ratio is obtained from the overall image minus the inner ratio. The simulation can be seen in Figure 6 below.

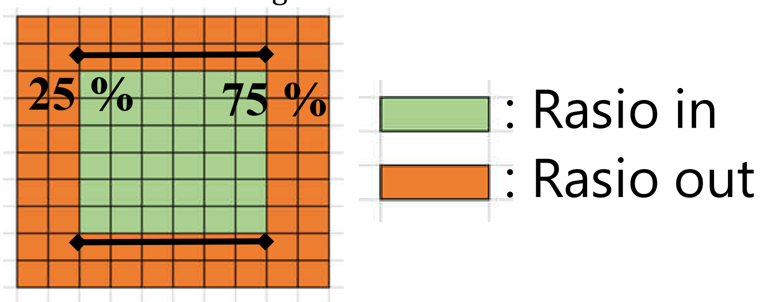


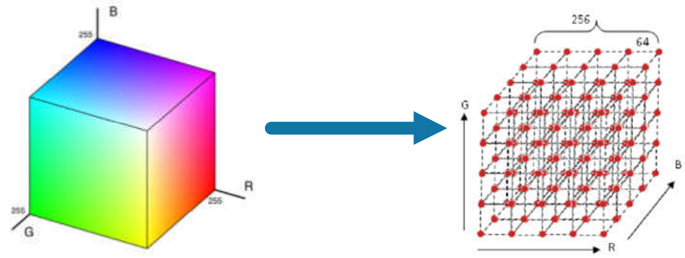
Figure 6. Object determination simulation

4.4 Feature Extraction

The process of taking an image by selected the appropriate feature extraction method and the measurement approach [2]. Extraction features that researchers used were color and texture.

4.4.1 Color Feature Extraction

The color feature extraction used 3D-Color Vector Quantization and HSV histogram. The main idea of 3D-Color Vector Quantization is that the system uniformly represents the color of the image at a certain position in the RGB vector color space. This means to reduce the complexity of the RGB color in the image and unite the color close to the vector space [8][22]. The size of the quantity is 5X5X5 so that it forms 125 positions in RGB.



Source: A. R. Barakbah, Y. Kiyoki, 2008, 3D-Color Vector Quantization for Image retrieval system, international database Sistem (IDB) 2008, Izaka, Japan [8]

Figure 7. Illustration 3D-Color Vector Quantization RGB [8]

The color feature extraction metadata that is formed is as follows:

$$MC_{x,i} = \{fc_{x,1}, fc_{x,2}, fc_{x,3}, \ldots, fc_{x,125}\} \quad (6)$$

Where:

$fc_{x,i}$ = feature extraction results from *i* the color histogram

4.4.2 Shape Feature Extraction

The feature extraction form of this paper using Connected Component Labeling [23] is applied based on the threshold of many models to identify all groups. Spatially connected groups where all pixels in connected components have pixel intensity values that are the same or connected to each other. That makes after all groups have been determined, each pixel is labeled according to the component specified. For extraction, the first step feature applies edge detection using canny detection[24] which aims to smooth the image and noise. This paper applied Connected Component Labeling (CCL) to get extracts of contours from the object. The contour is processed to obtain area values, equivalent diameters, extent, convex hull, solidity, eccentricity and perimeter of each object. The calculations applied to represent each pixel in the former's matrix use the mean (μ), median, standard deviation (σ), variance (σ^2), skewness (s), and kurtosis (k).

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} X_{ij} \quad (7)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}(X_{ij} - \mu_i)^2\right)^{\frac{1}{2}} \quad (8)$$

$$\sigma_i = \frac{1}{N}\sum_{j=1}^{N}(X_{ij} - \mu_i)^2 \tag{9}$$

$$s_i = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_{ij} - \mu_i}{\sigma_i}\right]^3 \tag{10}$$

$$k_i = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{X_{ij} - \mu_i}{\sigma_i}\right]^4 \tag{11}$$

Where:

X_{ij} is the value of the color component *i* in the pixel image *j*, and *N* is the number of pixel images.

The shape feature extraction metadata that is formed is as follows:

$$MS_{x,i} = \{fS_{x.area}, fS_{x.equv}, fS_{x.ext}, fS_{x.conv}, fS_{x.sol}, fS_{x.ecc}, fS_{x.per}\} \tag{12}$$

Where:

$fS_{x.area}$ = the result of feature extraction from the area in block *x*

$fS_{x.equv}$ = the result of feature extraction from the equivalent diameter in block *x*

$fS_{x.ext}$ = the result of feature extraction from the extent in block *x*

$fS_{x.conv}$ = the result of feature extraction from the convex hull in block *x*

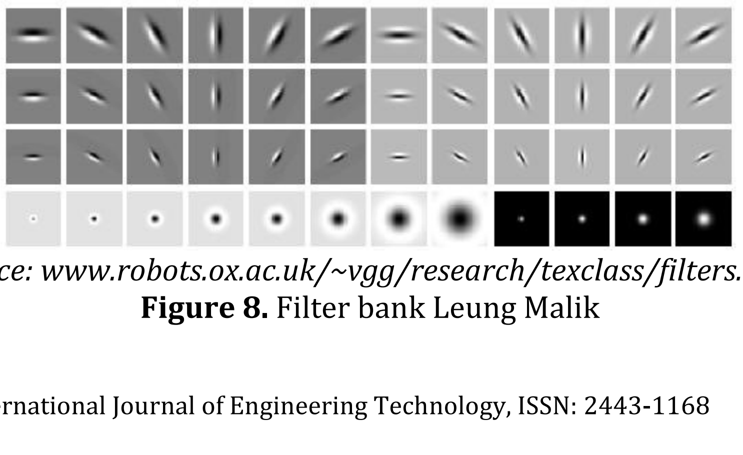
$fS_{x.sol}$ = the result of feature extraction from the solidity in block *x*

$fS_{x.ecc}$ = the result of feature extraction from the eccentricity in block *x*

$fS_{x.per}$ = the result of feature extraction from the perimeter in block *x*

4.4.3 Texture Feature Extraction

The texture feature extraction used Leung Malik (LM). The LM filter banks are rotationally variant. Therefore, the derivatives of the LM filter bank would change when subjected to different orientations. A set of images obtained under a known set of imaging conditions is considered as input data to the LM filters [25]. The LM filter set consists of 48 filters but in this paper used 15 filters.



Source: www.robots.ox.ac.uk/~vgg/research/texclass/filters.html

Figure 8. Filter bank Leung Malik

Leung Malik's approach implementation involves the process of segmenting canny detection edge detection[24]. Calculations applied for texture extraction researchers use calculations of the mean (μ), median, standard deviation (σ), variance (σ^2), skewness (s), and kurtosis (k).

The texture feature extraction metadata that is formed is as follows:

$$MT_{x,i} = \{fT_{x,1}, fT_{x,2}, fT_{x,3}, \ldots, fT_{x,15}\} \tag{13}$$

Where:

$fT_{x,1}$ = extracted texture features from i kernel.

4.5 Similarity Measurement

After features extraction, the metadata of color and structured are created. The metadata of images query is used to measure the similarity with proximity to the metadata of a image database. For the measurement approach, the researcher uses Normalize Canberra Distance where it can normalize distance data on each attribute [8]. Normalize Canberra Distance approaches data from image queries with image databases. The equation from Normalize Canberra Distance:

$$D_{(x,y)} = \frac{\sum_{i}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|}}{n} \tag{13}$$

Where:

N = number of attributes
x_i = metadata from image query
y_i = metadata from image database

5. EXPERIMENT AND ANALYSIS

For the testing phase in this paper used the benchmark 1000 image SIMPLIcity dataset from Wang et al. [26] which consists of 10 categories namely Africans, beaches, buildings historians, buses, dinosaurs, elephants, roses, horses, mountains, and food.

5.1 Background Remover and Object Detection

Based on the results of the background remover, it was found that there are images that cannot be supposed as image objects. For example the image of a mountain with a sky view, the system mostly suppose the image object is sky because the value of sky color extraction is more dominant than the color of the mountains. Except for dinosaur and flower images, the resulting color extraction is more dominant so it is supposed to be an object. Figure 9 illustrates some of the results of background remover.

Figure 9. Background Remover Results

In this paper, the researcher has attempted background remover by doing background replacement, the process of replacing background colors that have been removed with other colors. It aims to examine the background has been replaced by another color information whether or not to influence the similarity measurement. We attempted the results of the similarity to the detection of objects in each category using a method of calculation of the ratio of errors, accuracy, and scoring by determining the top 10 pictures from image search results for the query

$$Error = \sum_{i=1}^{10} err_i \begin{cases} err_i = 0 \leftarrow cr_i = cq \\ err_i = 1 \leftarrow otherwise \end{cases} \quad (14)$$

$$Score = \sum_{i=1}^{10} scr_i \begin{cases} scr_i = 10 - i + 1 \leftarrow cr_i = cq \\ scr_i = 0 \leftarrow otherwise \end{cases} \quad (15)$$

$$Precision = \sum_{i=1}^{10} pre_i \begin{cases} pre_i = 1 \leftarrow cr_i = cq \\ pre_i = 0 \leftarrow otherwise \end{cases} \quad (16)$$

Where:

cr_i = category of retrieval image

cq = category from image query

Tables 1,2 and 3 is displayed the results of errors, precision, and score values from the comparison between the results of the object removed by the object with the background color changed. Based on these results it could be seen that the background changed with other colors produces a high error, precision, and a low score compared to the background of the discarded object. The colors should indicate the object are not suitable because the background provides other information. However, if the information is discarded, it will provide better information. From the table above it can also be seen that the object detection technique is still not good if applied to images with backgrounds that are as varied as people, buildings, mountains, and buses. However, for the implementation of images with backgrounds that are not as diverse as dinosaurs, flowers, elephants, horses, food, and beaches, they have successfully detected objects well.

Table 1. Calculation Results of Average Error Ratio

Images	Object	Black	Yellow	Mean
Africa People	0.741	0.733	0.868	0.781
Beach	0.676	0.834	0.818	0.776
Building	0.750	0.824	0.940	0.838
Bus	0.722	0.901	0.923	0.849
Dinosaurs	0.140	0.636	0.839	0.539
Elephant	0.645	0.733	0.988	0.788
Flower	0.552	0.490	0.957	0.666
Food	0.578	0.659	0.934	0.724
Horse	0.455	0.732	0.865	0.684
Mountain	0.728	0.906	0.869	0.834
Average	**0.599**	**0.745**	**0.900**	**0.748**

Table 2. Calculation Results of Average Accuracy

Images	Object	Black	Yellow	Mean
Africa People	0.259	0.267	0.132	0.219
Beach	0.324	0.166	0.182	0.224
Building	0.250	0.176	0.060	0.162
Bus	0.278	0.099	0.077	0.151
Dinosaurs	0.860	0.364	0.161	0.461
Elephant	0.355	0.267	0.012	0.212
Flower	0.448	0.510	0.043	0.334
Food	0.408	0.341	0.066	0.272
Horse	0.545	0.268	0.135	0.316
Mountain	0.272	0.094	0.131	0.166
Average	**0.400**	**0.255**	**0.100**	**0.252**

Table 3. Calculation Results of Average Score

Images	Object	Black	Yellow	Mean
Africa People	18.29	15.65	0.13	5.94
Beach	21.47	10.06	0.18	9.34
Building	18.53	10.70	0.06	2.90
Bus	19.84	7.15	0.08	3.80
Dinosaurs	49.04	25.44	0.16	10.98
Elephant	23.87	18.22	0.01	0.71
Flower	29.29	32.38	0.04	2.10
Food	26.36	20.51	0.07	3.08
Horse	34.07	16.70	0.13	8.48
Mountain	19.46	6.60	0.13	6.37
Average	**26.02**	**16.34**	**0.10**	**5.37**

5.2 Selection of Image Similarity Techniques

The process of selecting the similarity measurement technique compared the Normalize Canberra Distance approach with cosine distance on the features of color shapes and textures. The following cosine distance equation is used as a comparison:

$$D_{(x,y)} = \frac{\sum_{i=1}^{N} x_i y_i}{\sqrt{\sum_{i=1}^{N} x_i^2 \sum_{i=1}^{N} y_i^2}} \tag{17}$$

Where:

N = number of attributes
x_i = metadata from image query
y_i = metadata from image database

Tables 4, 5, and 6 described error results, accuracy and score values from the comparison of approach methods. It could be seen that the method of normalizing Canberra distance is better to be applied to the color feature and texture feature extraction approach compared to the cosine distance method. Whereas for the shape features it is better to use the cosine distance method when compared to the Normalize Canberra Distance method.

Table 4. Calculation Results of Accuracy Values from Comparison of Similarity Measurement

Images	**Color**		**Shape**		**Texture**	
	Normalize Canberra	**Cosine**	**Normalize Canberra**	**Cosine**	**Normalize Canberra**	**Cosine**
Africa People	0.369	0.324	0.297	0.315	0.449	0.431
Beach	0.269	0.332	0.272	0.281	0.437	0.401
Building	0.443	0.256	0.318	0.345	0.426	0.440
Bus	0.355	0.287	0.390	0.386	0.767	0.773
Dinosaurs	0.853	0.865	0.672	0.643	0.753	0.744
Elephant	0.467	0.347	0.407	0.401	0.437	0.435
Flower	0.636	0.446	0.337	0.364	0.721	0.748
Food	0.479	0.451	0.294	0.259	0.509	0.490
Horse	0.656	0.625	0.417	0.394	0.563	0.568
Mountain	0.405	0.273	0.239	0.261	0.325	0.326
Average	**0.4932**	**0.4206**	**0.3643**	**0.3649**	**0.5387**	**0.5356**

Table 5 . Calculation Results of Error Values from Comparison of Similarity Measurement

Images	Color		Shape		Texture	
	Normalize Canberra	Cosine	Normalize Canberra	Cosine	Normalize Canberra	Cosine
Africa People	0.631	0.676	0.703	0.685	0.551	0.569
Beach	0.731	0.668	0.728	0.719	0.563	0.599
Building	0.557	0.744	0.682	0.655	0.574	0.560
Bus	0.645	0.713	0.610	0.614	0.233	0.227
Dinosaurs	0.147	0.135	0.328	0.357	0.247	0.256
Elephant	0.533	0.653	0.593	0.599	0.563	0.565
Flower	0.364	0.554	0.663	0.636	0.279	0.252
Food	0.521	0.549	0.706	0.741	0.491	0.510
Horse	0.344	0.375	0.583	0.606	0.437	0.432
Mountain	0.595	0.727	0.761	0.739	0.675	0.674
Average	**0.5068**	**0.5794**	**0.6357**	**0.6351**	**0.4613**	**0.4644**

Table 6. Calculation Results of Scoring Values from Comparison of Similarity Measurement

Images	Color		Shape		Texture	
	Normalize Canberra	Cosine	Normalize Canberra	Cosine	Normalize Canberra	Cosine
Africa People	24.78	21.74	19.78	21.18	28.41	27.68
Beach	18.65	21.93	18.61	19.82	27.47	25.73
Building	28.48	18.96	22.33	23.93	27.02	28.36
Bus	24.61	20.08	25.38	25.24	44.79	44.51
Dinosaurs	49.29	49.44	39.60	38.29	44.10	44.01
Elephant	29.37	23.56	26.57	26.49	28.21	27.58
Flower	38.72	29.23	22.60	24.05	42.05	42.65
Food	30.81	28.81	19.80	18.31	31.13	30.40
Horse	40.13	37.97	26.73	26.05	33.73	34.52
Mountain	26.62	19.32	17.18	18.71	22.12	22.39
Average	**31.146**	**27.104**	**23.858**	**24.207**	**32.903**	**32.783**

5.3 Result Retrieved

Some things that can be analyzed based on the results of the system and re-display the collection of database images that have similarities with image query to get the best solution in the form of images with the closest distance value to image query. Testing is done by comparing the features of color, shape, and texture through the calculation of the k-nearest neighbor's algorithm (k-NN) method with k = 3.

Table 7. The Results of Calculating Average Error Ratios, Accuracy, and Scoring

	Color	Texture	Shape	Color-Texture	Color-Shape	Texture-Shape	Color-Texture-Shape
Accuracy	62.02	71.21	60,99	72.03	67.18	72.65	73.48

Based on Table 7, it can be seen that the highest accuracy is 73.48% which is obtained by combining all the extraction features, the lowest accuracy is 60.99% which is obtained from the calculation of form features. From the results of the research, it can be concluded that combining all features is more optimal compared to the approach of each feature.

Tables 8, 9, 10 described from the calculation of error results, accuracy and score in each category. The dinosaur category generates the highest average accuracy and score on the comparison of each feature. Whereas for the beach category it generated the average accuracy value and the lowest score on the comparison of each feature. It can be concluded that the ocean category has diverse background information compared to the dinosaur category, so the system incorrectly supposed the object specified. This causes the results of the dinosaur category approach to be better than the ocean category because of the information to its appropriate so the results obtained are preferred or optimal.

Table 8. Comparison of Accuracy Calculation Results

Images	Color	Texture	Shape	Color-Texture	Color-Shape	Texture-Shape	Color-Texture-Shape
Africa People	0.369	0.449	0.297	0.412	0.494	0.520	0.496
Beach	0.269	0.437	0.272	0.414	0.347	0.358	0.358
Building	0.443	0.426	0.318	0.368	0.519	0.520	0.532
Bus	0.355	0.767	0.390	0.611	0.621	0.626	0.614
Dinosaurs	0.853	0.753	0.672	0.861	0.919	0.930	0.933
Elephant	0.467	0.437	0.407	0.397	0.540	0.565	0.554
Flower	0.636	0.721	0.337	0.736	0.767	0.769	0.770
Food	0.479	0.509	0.294	0.523	0.596	0.593	0.597
Horse	0.656	0.563	0.417	0.641	0.701	0.697	0.704
Mountain	0.405	0.325	0.239	0.334	0.430	0.435	0.426
Average	**0.493**	**0.539**	**0.364**	**0.530**	**0.593**	**0.601**	**0.598**

Table 9. Comparison of Error Calculation Results

Images	Color	Texture	Shape	Color-Texture	Color-Shape	Texture-Shape	Color-Texture-Shape
Africa People	0.631	0.551	0.703	0.506	0.606	0.480	0.504
Beach	0.731	0.563	0.728	0.653	0.730	0.642	0.642
Building	0.557	0.574	0.682	0.481	0.537	0.480	0.468
Bus	0.645	0.233	0.610	0.379	0.609	0.374	0.386
Dinosaurs	0.147	0.247	0.328	0.081	0.096	0.070	0.067
Elephant	0.533	0.563	0.593	0.460	0.506	0.435	0.446
Flower	0.364	0.279	0.663	0.233	0.347	0.231	0.230
Food	0.521	0.491	0.706	0.404	0.498	0.407	0.403
Horse	0.344	0.437	0.583	0.299	0.318	0.303	0.296
Mountain	0.595	0.675	0.761	0.570	0.599	0.565	0.574
Average	0.507	0.461	0.636	0.407	0.485	0.399	0.402

Table 10. Comparison of Scoring Calculation Results

Images	Color	Texture	Shape	Color-Texture	Color-Shape	Texture-Shape	Color-Texture-Shape
Africa People	24.78	28.41	19.78	31.51	26.02	32.90	31.82
Beach	18.65	27.47	18.61	22.97	18.81	23.37	23.39
Building	28.48	27.02	22.33	32.17	29.01	32.55	32.63
Bus	24.61	44.79	25.38	38.20	26.09	38.50	38.40
Dinosaurs	49.29	44.10	39.60	51.92	50.98	52.06	52.39
Elephant	29.37	28.21	26.57	33.13	30.54	34.29	33.93
Flower	38.72	42.05	22.60	44.94	39.94	45.27	44.92
Food	30.81	31.13	19.80	36.08	31.93	36.08	36.16
Horse	40.13	33.73	26.73	42.65	41.20	42.33	42.75
Mountain	26.62	22.12	17.18	27.49	26.07	27.89	27.42
Average	**31.15**	**32.90**	**23.86**	**36.11**	**32.06**	**36.52**	**36.38**

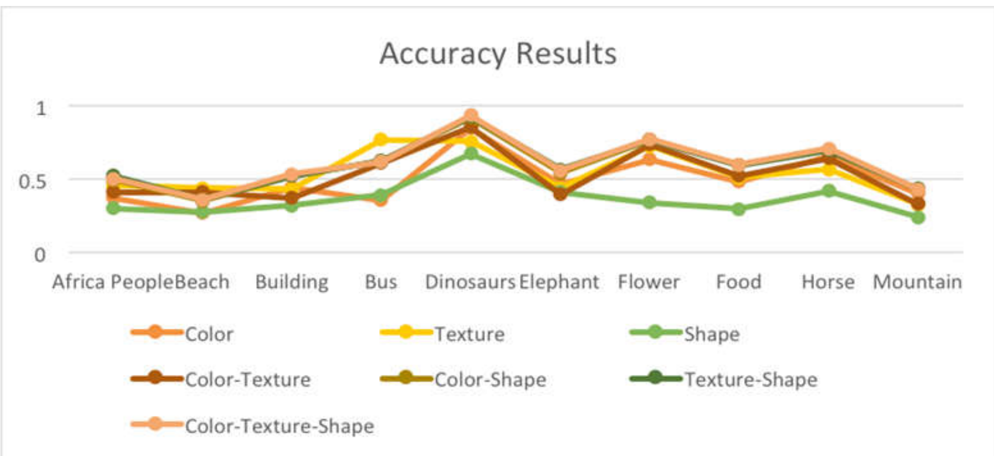


Figure 10. Results of Accuracy

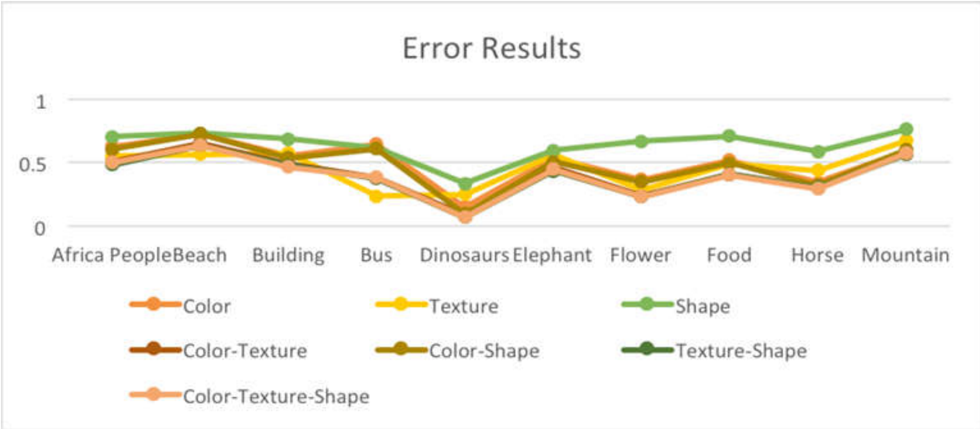


Figure 11. Results of Error

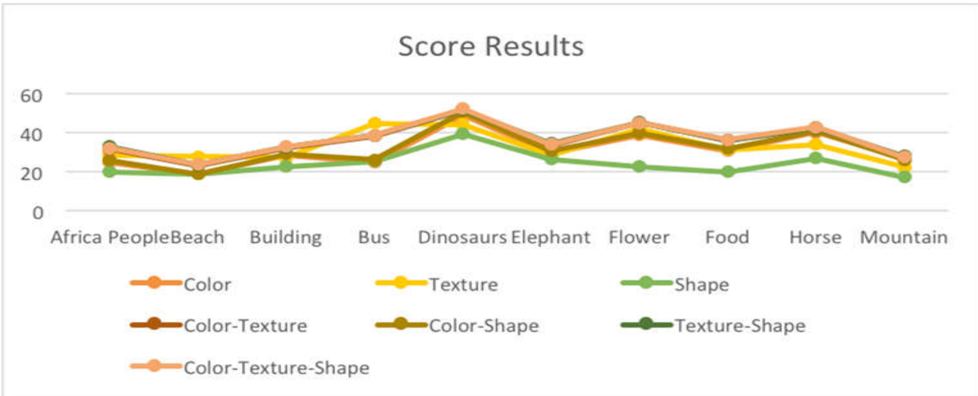


Figure 12. Results of Score

6. CONCLUSION

This paper presents a new approach to content-dependent image search system for aggregation of color, shape, and texture features. The system introduces color feature extraction using color extraction RGB and HSL then it normalization the unique data. The main idea of this approach is using clustering techniques that are Hierarchical K-means and Optimal K Selection. Our image search extracted three kinds of features which are color, shape, and texture. We used 3D-Color Vector Quantization color for extraction feature. The process of extraction features shape used Connected Component Labeling (CCL) which is calculated the area value, equivalent diameter, extent, convex hull, solidity, eccentricity, and perimeter of each object. The texture extraction feature was used by Leung Malik (LM) 's approach with 15 kernels.

In this research, the techniques that have been carried out are less able to find the desired object correctly if the images sought provide information that has a variety of backgrounds compared to information from the object.. For distance approach calculation techniques this paper have analyz that the calculation using the Canberra distance normalizes more high accuracy values than using cosine. The results are used for detecting the extraction accuracy rate of 62%, 71% by texture feature, 60% by shape feature, 72% by combined color-texture feature, 67% by combined color-shape feature, 72 % combined texture-shape features and 73% combined the feature. Analysis obtained from the experimental results, the system is more optimal if the feature extraction used is to combine all the features compared to just using each feature. For further work, we will apply an automatic weighting mechanism system to select this feature automatically.

REFERENCES

[1] A. K. Jain and A. Vailaya, **Image retrieval using color and shape,** *Pattern Recognit.*, vol. 29, no. 8, pp. 1233–1244, 1996.

[2] M. Bounthanh, K. Hamamoto, and B. Attachoo, **Content-Based Image Retrieval System Based on Combined and Weighted Multi-Features**, *Int. Symp. Commun. Inf. Technol.*, pp. 449–453, 2013.

[3] T. Kato, **Database architecture for content-based image retrieval**, *Proc. SPIE Conf. Image Storage Retr. Syst.*, vol. 1662, no. March, pp. 112–123, 1992.

[4] C. W. Niblack, **QBIC project: querying images by content, using color, texture, and shape**, *Proc. SPIE*, vol. 1908, no. 1, pp. 173–187, 1993.

[5] a. Pentland, R. W. Picard, and S. Sclaroff, **Photobook: Content-based manipulation of image databases**, *M.I.T. Media Laboratory Percaptual Computing Technical Report*, vol. 255, no. 3. pp. 233–254, 1993.

[6] G. W. Jia Li James Ze Wang, **SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries**, *IEEE Pami*, vol. 23, no. 9, pp. 947–963, 2001.

[7] M. Li and J. Xiuhua, ***An Improved Algorithm Based on Color Feature Extraction for Image Retrieval***. 2016.

[8] A. R. Barakbah and Y. Kiyoki, **3D-Color Vector Quantization for Image Retrieval Systems**, *Int. Database Forum 2008*, no. September 2008, pp. 13–18, 2008.

[9] A. R. Barakbah and Y. Kiyoki, **Image Search System with Automatic Weighting Mechanism for Selecting Features**, *6th Int. Conf. Inf. Commun. Technol. Syst.*, 2010.

[10] P. Anantharatnasamy, K. Sriskandaraja, V. Nandakumar, and S. Deegalla, **Fusion of colour, shape and texture features for content based image retrieval**, *Proc. 8th Int. Conf. Comput. Sci. Educ. ICCSE 2013*, no. Iccse, pp. 422–427, 2013.

[11] A. K. Naveena and N. K. Narayanan, **Image retrieval using combination of color, texture and shape descriptor**, *2016 Int. Conf. Next Gener. Intell. Syst. ICNGIS 2016*, 2017.

[12] K. Mala and S. Suganya, **Content Based Image Retrieval System based on Semantic Information Using Color, Texture and Shape Features**, *Comput. Technol. Intell. Data Eng.*, 2016.

[13] A. Kutics and A. Nakagawa, **Detecting prominent objects for image retrieval**, *Proc. - Int. Conf. Image Process. ICIP*, vol. 3, pp. 445–448, 2005.
[14] A. Mohamed, F. Khelifi, J. Jiang, and S. Ipson, **Efficient content based image retrieval based on Semantic Object Detection**, *10th Int. Conf. Inf. Sci. Signal Process. their Appl. ISSPA 2010*, no. Isspa, pp. 193–196, 2010.
[15] N. Suciati, D. Herumurti, and A. Y. Wijaya, **Fractal-based texture and HSV color features for fabric image retrieval**, *Proc. - 5th IEEE Int. Conf. Control Syst. Comput. Eng. ICCSCE 2015*, no. November, pp. 178–182, 2016.
[16] R. Garnavi, M. Aldeen, M. E. Celebi, A. Bhuiyan, C. Dolianitis, and G. Varigos, **Automatic Segmentation of Dermoscopy Images Using Histogram Thresholding on Optimal Color Channels**, vol. 1. 2010.
[17] and Y. P. Baogang Wei, Yonghuai Liu, Member, IEEE, **Using Hybrid Knowledge Engineering and Image Processing in Color Virtual Restoration of Ancient Murals**, *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 5, pp. 1338–1343, 2003.
[18] T. Adah, H. Ni, and B. Wang, **Partial Likelihood for Estimation of Multi-Class Posterior Probabilities**, *Neural Networks*, vol. 21250, pp. 1053–1056, 1999.
[19] A. R. Barakbah and K. Arai, **Identifying moving variance to make automatic clustering for normal data set**, *IECI Japan Work. 2004 2004*, no. May, pp. 26–30, 2004.
[20] A. R. Barakbah and K. Arai, **Determining Constraints of Moving Variance to Find Global Optimum and Make Automatic Clustering**, *Ind. Electron. Semin. 2004*, pp. 409–413, 2004.
[21] K. Arai and A. R. Barakbah, **Hierarchical K-means: an algorithm for centroids initialization for K-means**, *Rep. Fac. Sci. Engrg. Reports Fac. Sci. Eng. Saga Univ. Saga Univ.*, vol. 36, no. 1, pp. 36–125, 2007.
[22] A. R. Barakbah and Y. Kiyoki, **Image Retrieval Systems with 3D-Color Vector Quantization and Cluster based Shape and Structure Features**, *Inf. Model. Knowl. Bases XXI*, vol. 206, pp. 169–187, 2010.
[23] A. Rosenfeld and J. Pfaltz, *Pfalz, J.L,* **Sequential Operations in Digital Picture Processing**. *Journal of the ACM 13(4), 471-494*, vol. 13. 1966.
[24] Cai-Xia Deng, Gui-Bin Wang, and Xin-Rui Yang, **Image edge detection algorithm based on improved Canny operator**, *2013 Int. Conf. Wavelet Anal. Pattern Recognit.*, no. 1, pp. 168–172, 2013.
[25] T. Leung and J. Malik, **Representing and recognizing the visual appearance of materials using three-dimensional textons**, *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 29–44, 2001.
[26] J. Z. Wang, **Semantics-Sensitive Integrated Matching for Picture Libraries and Biomedical Image Databases**, 2000.