SEARCHING SIMILAR OF PAINT COLOR BASED ON RGB AND HSV SPECTRUM USING K-MEANS METHOD

1RIAN TRI CAHYONO, 2EKO PRASETYO, 3RIFKI FAHRIAL
1,2,3Informatics Department, Faculty of Engineering, Bhayangkara Surabaya University
Jl. Ahmad Yani No 114 Surabaya
E-mail: 1riant57@gmail.com, 2eko@ubhara.ac.id, 3rifki@ubhara.ac.id

ABSTRACT

Currently the community in determining the desired paint color choice still manually searches by selecting the drawings one by one in the manual paint catalog, the more color choices available in the catalog take a long time to determine the desired color choice. This research replaces the search manually by looking at the catalog with a computer based system. Similar paint color search uses image processing, ie reading the intensity of the Red, Green, Blue (RGB) value of the image, the value of the RGB is then calculated to get the HSV value, so as to get the RGB and HSV features as the input calculation method kmeans clustering, the data used as training data obtained from dulux wall paint catalog of 317 data. The results in this study average data can be grouped into 11 groups and 15 iterations, also calculated by a method to check the cluster validity of Davies Bouldin Index (DBI) sebanya 3 times the experiment got an average value of 0879.

Keywords: Image, RGB, HSV, K-means, color, paint

1. INTRODUCTION

Paint is a liquid consisting of a variety of colors used to coat the surface of a material with the aim of beautifying, strengthening or protecting the material, paint color is one important element that determines the beauty, one of the use of paint is the paint of the wall.

The color of the paint in this study is represented in digital image. Characteristic of human eye absorption, color is considered as a combination of variables called primary color ie red (R), green (G), and blue (B). Common characteristics used to distinguish one color with another color is brightness, hue, and saturation.

Utilizing digital image processing combined with k-means clustering method, using 317 kinds of color from dulux wall paint catalogs as training data, it is hoped this research can help people in searching paint color similar in a more efficient way.

2. THEORETICAL BASIS

2.1. Color Processing on Digital Image

Digital imagery is an image that can be processed by a computer [10]. A digital image can be defined as a two-dimensional function \( f(x, y) \), where \( x \) and \( y \) are spatial coordinates, and the amplitude of \( f \) in any coordinate pair \( (x, y) \) is called the light intensity (gray level) at that point. Image colors are formed by a combination of individual 2-D imagery.

2.1.1. RGB Color Model

The image represented in the RGB color model consists of three image components, namely Red, Green and Blue. when displayed on an RGB monitor, these three image combinations are on the phosphor screen to produce composite color images. The number of bits used to represent each pixel into an RGB space is called a pixel depth. In digital image processing, RGB is represented as an 8-bit image. In the condition of each color of the RGB pixel [then, the triplet of values (R, G, B)] has a 24-bit depth (3 layers of image with bits per layer) [8].

405
2.1.2. HSV Color Model

HSV is one of the color model of RGB value conversion, HSV has 3 attributes, Hue, Saturation, and Value with description as follows:

a. Hue
   - Express the true colors, such as red, violet, and yellow. Hue is used to distinguish colors and determine redness, greenness, etc. from light.

b. Saturation
   - Express the level of light color purity, ie indicating how much white color is applied to the color.

c. Value
   - Attribute that states the amount of light received by the eye regardless of color. The range of values is between dark (black) and light (white) [6].

The RGB to HSV conversion formula [14]:

The RGB value is divided by 255 to change the distance from 0-255 to 0-1

\[ R' = \frac{R}{255} \quad C_{max} = \max (R', G', B') \]
\[ G' = \frac{G}{255} \quad C_{min} = \min (R', G', B') \]
\[ B' = \frac{B}{255} \quad \Delta = C_{max} - C_{min} \]

a. Hue value calculation

The hue value is calculated with the highest RGB intensity of each image, if in an image of the highest intensity Red value then use the formula \( C_{max} = R' \), as well as for Green and Blue intensity, if \( \Delta \), the calculation of \( C_{max}-C_{min} \) is = 0 then the value hue is represented in 0°

\[
H = \begin{cases} 
0^\circ, & \Delta = 0 \\
60^\circ \times \frac{(G' - B')}{\Delta} + 6, & C_{max} = R' \\
60^\circ \times \frac{(B' - R')}{\Delta} + 2, & C_{max} = G' \\
60^\circ \times \frac{(R' - G')}{\Delta} + 4, & C_{max} = B' 
\end{cases}
\]

b. Saturation Value Calculation

If the value of \( C_{max} = 0 \) then the value of saturation is 0, and if the value of saturation \( \neq 0 \) then the hue value is calculated using the formula \( \Delta \) divided \( C_{max} \)

\[
S = \begin{cases} 
0, & C_{max} = 0 \\
\frac{\Delta}{C_{max}}, & C_{max} \neq 0 
\end{cases}
\]

c. Value calculation

Calculation of value obtained from calculation of \( C_{max} \) value

\[
V = C_{max}
\]

2.2. Data Grouping Using K-means Clustering Method

K-Means is one method of grouping nonhierarchy data (sekatan) that try to partition the existing data into two or more groups. This method partitions the data into groups so that the same characteristic data is entered into the same group and the different characteristic data are grouped into another group [7].
General data grouping algorithm is K-means method:

a. Determine the number of groups.
b. Allocate data into the randomness of the guesthouse.
c. Calculate the center of the group (centroid or average) of the data in each group. Calculation of centroid value using the formula:

\[ C_i = \frac{1}{M} \sum_{j=1}^{M} x_j \]

d. Allocate each data to centroid or nearest average Distance data to centroid using Euclidean parameter:

Data allocation is calculated based on data distance to the centroid

\[ D(x, y) = \left\| x - y \right\| = \sum_{j=1}^{N} |x_j - y_j| \]

The data is reallocated explicitly to the group that has the centroid with the closest distance from the data. This allocation can be formulated as follows (MacQueen, 1967):

\[ a_{ij} = \begin{cases} 1 & d = \min \{(x_i, C_j)\} \\ 0 & \text{lainnya} \end{cases} \]

e. Return to step 3, if there is still data moving group, or if there is a change of centroid value above the specified threshold value, or if the value change in the objective function used is still above the specified threshold value. The formula for calculating the value of objective function is as follows:

\[ J = \sum_{i=1}^{N} \sum_{k=1}^{K} a_{ik} D \left( x_i, C_k \right) \]

Iteration is stopped if the value of objective function = 0, as a representation that there is no data transfer in the last iteration.

3. SYSTEM DESIGN

1. Read the RGB image
   Each pixel of an image read in the RGB value using a function of the programming language, in order to obtain the RGB value of each pixel from a whole image then performed the calculation of the average by doing the sum of each pixel divided by the number of pixels.
   From the average RGB value then converted to HSV color model, then each RGB and HSV value is used as input feature in k-means method calculation.

2. K-means clustering
   The intensity of RGB and HSV values which are representations of the images are grouped using the k-means method. Calculation of data on the centroid and the objective function using Euclidean distance parameter. The iteration is done up to the value of the objective function = 0 as an indication that in the last iteration there is no data moving group.

3. Search color recommendation
   After iteration stops the test data is recalculated using the Euclidean distance parameter against each training data that is in the group with the test data. The smallest value of the calculation indicates that the data has characteristics similar to the test data.
4. TESTING THE SYSTEM

Cluster testing in this study using Davies Bouldin Index (DBI). DBI is a cluster validation method of clustering results. The smaller the DBI value indicates the optimal optimal cluster scheme [13]. The first test using sample 9 data is divided into 3 clusters and subsequent tests using 317 data from dulux cat catalogs are divided into 11 clusters.

<table>
<thead>
<tr>
<th>Number of Testing</th>
<th>Picture test iteration</th>
<th>clusters</th>
<th>centroid</th>
<th>SSW</th>
<th>DBI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>1</td>
<td>blue</td>
<td>76.2</td>
<td>0.479</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>red</td>
<td>53.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>green</td>
<td>45.73</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>1</td>
<td>blue</td>
<td>61.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2</td>
<td>white</td>
<td>31.34</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>yellow</td>
<td>44.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>brown</td>
<td>56.56</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>light blue</td>
<td>25.50</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>light pink</td>
<td>63.71</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>7</td>
<td>light blue</td>
<td>53.75</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8</td>
<td>light green</td>
<td>33.78</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td>light brown</td>
<td>50.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>light yellow</td>
<td>26.57</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>11</td>
<td>light pink</td>
<td>33.93</td>
<td></td>
</tr>
</tbody>
</table>

5. CONCLUSION

1. Similar paint color search system based on RGB and HSV spectrum using k-means clustering method to obtain cluster validity value calculated using DBI algorithm with 3 times testing in get the average value of 0.879
2. Each value of the Red Green Blue (RGB) and Hue Saturation Value (HSV) features is quite representative of the characteristics of each data that is the input of the K-means algorithm calculation
3. To determine the best number of clusters in k-means clustering method it is necessary to add cluster optimization method
4. Image detection can be developed using a color sensor so as not to manually input the image.
6. REFERENCES