

***K-Means Clustering Untuk Pengenalan Buah
Berdasarkan Karakteristik Warna Citra***

Yustina Retno Wahyu Utami ²⁾

Abstract

Research on pattern recognition in images has been developed. Objects can be recognized by color, texture, or shape. This study took four types of fruit as its object of apple, pear, bark, and orange. Feature extraction image taken based on the color characteristics of objects in the image of Hue Saturation Value. Characteristic images obtained were classified using k-means clustering. This experiment using 20 data has resulted in 70% accuracy.

Keyword: *feature extraction, Hue Saturation Value, k-means clustering, recognized of fruit*

1. PENDAHULUAN

Warna suatu obyek dengan obyek lain berbeda. Untuk mengenali suatu obyek apakah berbeda dengan obyek lain dapat dilakukan dari perbedaan warnanya. Demikian pula pada citra, obyek yang berbeda memiliki warna yang berbeda pula. Dalam penelitian ini, obyek yang akan digunakan sebagai sampel adalah obyek buah.

Pada akuisisi citra, citra yang dihasilkan berupa citra berwarna 24 bit. Citra berwarna ini memiliki komponen red, green dan blue (RGB). Citra RGB ditransformasikan ke model warna HSV (*Hue Saturation Value*) untuk memperoleh ciri citra. Dalam penelitian ini, akan dikenali beberapa jenis buah berdasarkan karakteristik warna (hue dan saturasi). Untuk mengklasifikasikan ciri, digunakan metode *k-means clustering*. Sedangkan untuk mengelompokkan citra uji ke dalam kelas yang ada, akan digunakan metode jarak terdekat dengan pusat kluster.

²⁾ Staf Pengajar STMIK Sinar Nusantara Surakarta

II. METODE PENELITIAN

Dalam penulisan ini metode penelitian yang digunakan adalah:

1. Metode Pengamatan
Dengan cara mengamati citra berwarna 24 bit dengan obyek buah, baik pada model RGB maupun HSV.
2. Metode Percobaan
Citra dipotong, diambil pada bagian obyeknya kemudian ditransformasikan ke model warna HSV. Ciri rata-rata dan standard deviasi dari citra H dan S diekstrak kemudian disimpan dalam database.
3. Studi Literatur
Untuk memperkaya wawasan dan bahan kajian dalam menyimpulkan bahan kajian, diperlukan literature dari perpustakaan ataupun internet.

III. DASAR TEORI

A. Model Warna

Model Warna Red Green Blue (RGB):

Model warna RGB dipakai untuk menampilkan warna pada monitor. Model ini berdasar pada tri-stimulus *vision theory* yang mengatakan bahwa manusia melihat warna dengan cara membandingkan cahaya yang datang dengan sensor-sensor peka cahaya pada retina (yang berbentuk kerucut). Sensor-sensor tersebut paling peka terhadap cahaya dengan panjang gelombang 630 nm (merah), 530 nm (hijau) dan 450 nm (biru).

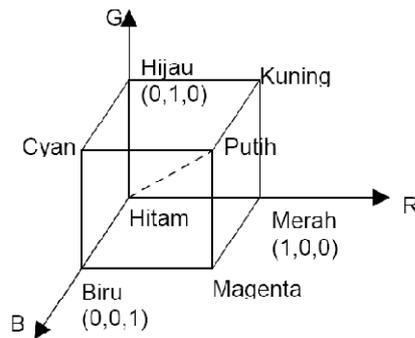
Model Gambar 1 dapat digambarkan dengan kubus dengan sumbu-sumbu R, G dan B, yang mana warna sudut kubus pada sumbu utama menyatakan warna primer. Sedangkan warna sudut kubus diluar sumbu utama menyatakan warna komplementer (merah dengan cyan, hijau dengan magenta, biru dengan kuning). Warna gray dinyatakan sepanjang diagonal hitam-putih.

Model Warna Hue Saturation Value (HSV)

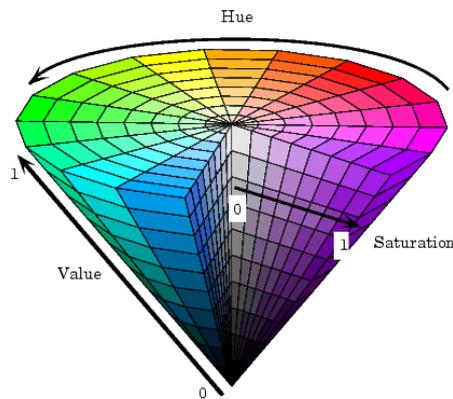
Model HSV diturunkan dari model warna RGB (dengan cara melihat sepanjang diagonal hitam putih).

- Hue: sudut putar seputar sumbu tegak dimana warna merah pada 0° dan yellow pada 60° dst
- Saturation: bernilai antara 0 dan 1. Dinyatakan sebagai ratio antara kemurnian warna terpilih dengan kemurnian warna maksimumnya ($S = 1$). $S = 0$ berarti gray, $S = 1$ berarti murni.
- Value bernilai antara 0 (hitam) dan 1 (putih)

Model HSV dapat digambarkan Gambar 2.



Gambar 1. Model Warna RGB



Gambar 2. Model Warna HSV

B. Ekstraksi Ciri

Ciri citra yang digunakan adalah rata-rata piksel dan standard deviasi. Rata-rata piksel dalam citra hue merepresentasikan rata-rata kemurnian warna obyek. Sedangkan standard deviasi merepresentasikan seberapa jauh simpangan nilai warna terhadap nilai rata-ratanya. Bila standard deviasi besar, menggambarkan bahwa terdapat nilai piksel yang jauh berbeda, sebaliknya bila standard deviasi kecil, artinya nilai piksel seragam atau hampir sama.

C. Clustering

Dalam penelitian ini, digunakan metode *k-means clustering*, dimana jumlah kluster telah diketahui.

Algoritma *k-means clustering* adalah sebagai berikut:

1. Set inisial kluster sebanyak k , k adalah jumlah kluster. Inisial kluster sebagai pusat kluster.
2. Tentukan anggota kluster berdasarkan jarak terdekat ke pusat kluster.
3. Hitung kembali pusat kluster baru dari rata-rata anggota kluster.
4. Ulangi langkah 2 dan 3 sampai pusat kluster stabil/tidak berubah.

Untuk membandingkan antara ciri citra uji dengan ciri citra yang disimpan digunakan jarak euclidian. Ciri citra uji diklasifikasikan dengan menghitung jarak terhadap setiap rata-rata kelas, dan memasukkan citra uji tersebut ke dalam kelas yang memiliki jarak terdekat. Bila terdapat suatu obyek baru dengan ciri x , penghitungan jarak terdekat dengan jarak Euclidian adalah sebagai berikut:

$$D_j(x) = \|x - m_j\|$$

dimana:

$j = 1, 2, 3, \dots, n$

$m =$ pusat kluster

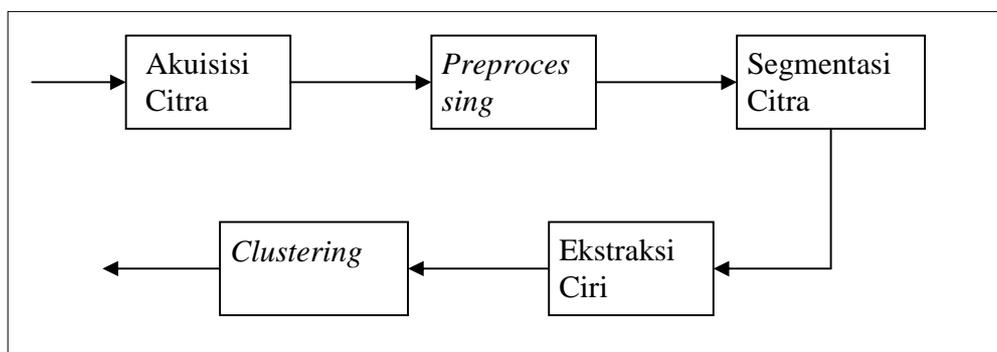
IV. LANGKAH-LANGKAH PENELITIAN

Langkah-Langkah yang dilakukan dalam penelitian ini seperti Gambar 3.

a) Akuisisi Citra

Citra yang digunakan dalam penelitian ini diambil dengan kamera digital. Citra yang dihasilkan berupa citra berwarna dengan resolusi 1728x2304 piksel. Obyek diamati dengan jarak pengamatan (80 cm) dan

zoom tetap. Untuk keperluan pengolahan, dipilih region of interest (ROI) dengan memotong citra sehingga citra memiliki ukuran 60 x 60 piksel.



Gambar 3. Langkah Penelitian

b) *Preprocessing*

Dari model warna RGB dan HSV, dapat dipelajari bahwa citra RGB dapat diturunkan ke model HSV. Ciri yang akan dipakai adalah rata-rata dan standard deviasi Hue dan Saturasi. Alasan pemilihan ciri ini adalah:

- pada model RGB, harus memperhatikan seluruh komponen warna yakni R, G, dan B
- pada model HSV, hanya memperhatikan komponen warna hue dan saturasi saja dimana hue menunjukkan warna obyek sedangkan saturasi menunjukkan kemurnian warna obyek.

c) *Ekstraksi Ciri*

Setelah citra ditransformasikan ke model HSV, hitung rata-rata dan standard deviasi dari hue dan saturasi. Dari sampel yang diambil, ciri citra jenis apel, pear, jeruk, dan salak adalah terlihat pada Tabel 1.

d) *Clustering*

Metode yang digunakan untuk klasifikasi obyek adalah *k-means clustering*. Perhitungan jarak menggunakan jarak Euclidean. Dalam penelitian ini terdapat 4 kelas obyek yakni:

- Kelas A : Apel
- Kelas B : Pear
- Kelas C : Jeruk

Kelas D : Salak

Tabel 1. Ciri citra

| Kelas Obyek | Rata-Rata | | Standard Deviasi | |
|-------------|-----------|----------|------------------|----------|
| | H | S | H | S |
| Apel | 0.0561 | 0.689265 | 0.017595 | 0.051765 |
| | 0.0765 | 0.635715 | 0.0255 | 0.068085 |
| | 0.0816 | 0.69309 | 0.024735 | 0.06426 |
| Pear | 0.09435 | 0.601035 | 0.005865 | 0.04029 |
| | 0.0918 | 0.581145 | 0.006375 | 0.033915 |
| | 0.09435 | 0.59007 | 0.007905 | 0.05253 |
| Jeruk | 0.12495 | 0.90984 | 0.008925 | 0.086445 |
| | 0.12495 | 0.84915 | 0.00969 | 0.118575 |
| | 0.11985 | 0.84609 | 0.006885 | 0.139485 |
| Salak | 0.30345 | 0.6171 | 0.441405 | 0.114495 |
| | 0.6528 | 0.44829 | 0.439365 | 0.165495 |
| | 0.2295 | 0.458235 | 0.393975 | 0.160905 |

V. IMPLEMENTASI DAN PEMBAHASAN

A. Implementasi

Aplikasi pengenalan buah yang dibuat menggunakan MATLAB 6.5, meliputi proses :

- konversi ke model warna HSV
- ekstraksi ciri
- klustering
- menyimpan ciri dalam database (berupa file berformat dat)

Transformasi Citra RGB ke model HSV, diimplementasikan dengan kode berikut

```
J=rgb2hsv(I);
```

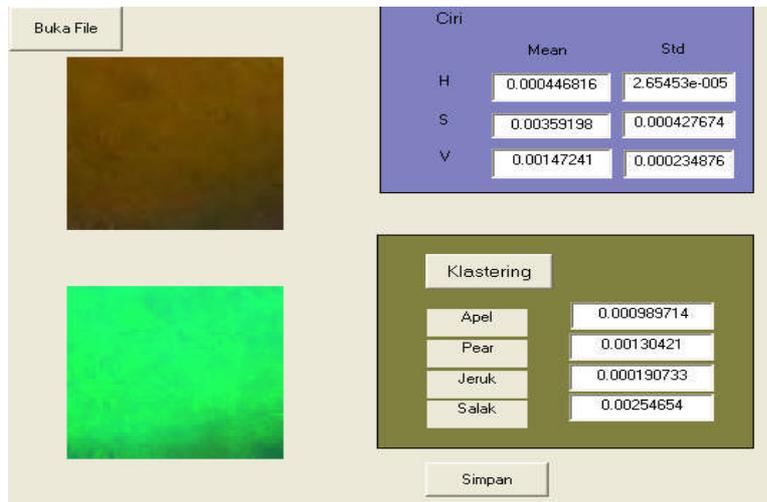
di mana I adalah citra RGB dan J adalah citra HSV.

Ekstraksi ciri yang diambil berupa rata-rata/mean dan standard deviasi citra hue dan saturasi. Kode untuk ekstraksi ciri seperti di bawah ini.

```
mh=mean(h(:)); ms=mean(s(:)); mv=mean(v(:));  
stdh=std2(h); stds=std2(s); stdv=std2(v);
```

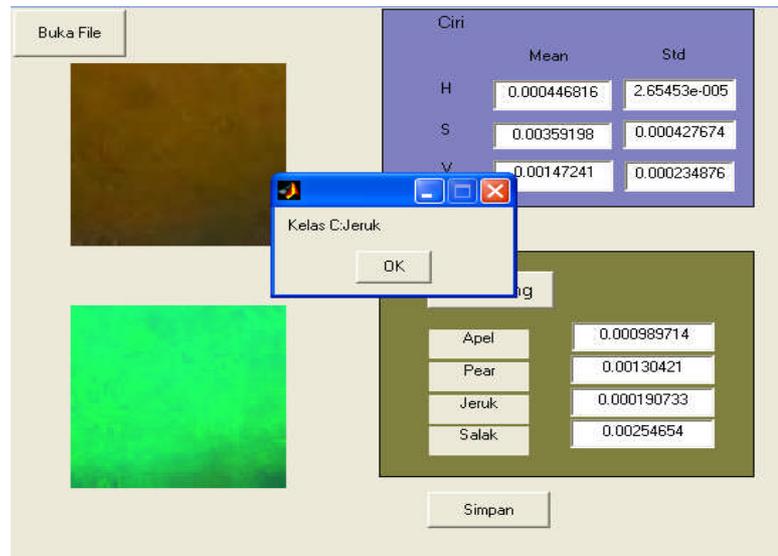
Proses *clustering* menggunakan *k-mean* diimplementasikan dengan memanggil fungsi $[IDX, C] = kmean(data, 4);$

Antarmuka aplikasi untuk pengenalan buah ini seperti tampak pada Gambar 4.



Gambar 4. Antarmuka Pemakai

Citra obyek ditranformasikan ke model HSV. Dari citra hue dan saturasi, diperoleh ciri rata-rata dan standard deviasi dari hue dan saturasi. Ciri tersebut dikelompokkan menggunakan metode *k-means clustering* (Gambar 5). Citra akan dikelompokkan dalam kelas berdasarkan jarak terdekat dengan pusat klaster. Bila diinginkan ciri citra uji dapat pula disimpan.



Gambar 5. Pengelompokan Obyek dengan *K-Means Clustering*

B. Pembahasan

Uji terhadap penelitian ini dilakukan terhadap 20 data uji menggunakan data sampel pada tabel 1. Hasil pengujian disajikan pada tabel 2. Berdasarkan hasil pengujian pada tabel 2, terdapat 6 data yang dikelompokkan ke kluster yang salah dan 14 data yang dikelompokkan dengan benar. Hasil akurasi sistem adalah $14/20 \times 100\% = 70\%$.

VI. KESIMPULAN SARAN

Kesimpulan yang diperoleh dari penelitian ini adalah:

1. Karakteristik warna hue dan saturasi dapat digunakan untuk membedakan beberapa obyek buah namun masih terdapat kesalahan dalam klasifikasi.
2. Diperlukan ciri lain untuk membedakan obyek seperti ciri tekstur dan bentuk dari obyek buah tersebut.
3. K-means dapat digunakan untuk membedakan jenis buah berdasarkan karakteristik warna citra.
4. Pengujian sistem terhadap 20 data uji diperoleh akurasi sebesar 70%.

Saran pengembangan:

1. Karena sampel citra masih terlalu sedikit, perlu untuk mencoba dengan sampel lebih banyak
2. Perlu dikembangkan untuk jenis buah yang lain.
3. Ciri yang diambil berupa karakteristik warna citra dalam domain spasial. Perlu dikembangkan penelitian pengenalan berdasarkan karakteristik citra dalam domain frekuensi.

Tabel 2. Hasil Pengujian

| No | Sampel | Hasil Sistem | Status |
|----|---------|--------------|--------|
| 1 | Jeruk 1 | Jeruk | TRUE |
| 2 | Jeruk 2 | Jeruk | TRUE |
| 3 | Jeruk 3 | Apel | FALSE |
| 4 | Jeruk 4 | Apel | FALSE |
| 5 | Jeruk 5 | Jeruk | TRUE |
| 6 | Apel 1 | Apel | TRUE |
| 7 | Apel 2 | Apel | TRUE |
| 8 | Apel 3 | Jeruk | FALSE |
| 9 | Apel 4 | Apel | TRUE |
| 10 | Apel 5 | Apel | TRUE |
| 11 | Pear 1 | Pear | TRUE |
| 12 | Pear 2 | Pear | TRUE |
| 13 | Pear 3 | Apel | FALSE |
| 14 | Pear 4 | Apel | FALSE |
| 15 | Pear 5 | Pear | TRUE |
| 16 | Salak 1 | Salak | TRUE |
| 17 | Salak 2 | Salak | TRUE |
| 18 | Salak 3 | Salak | TRUE |
| 19 | Salak 4 | Salak | TRUE |
| 20 | Salak 5 | Pear | FALSE |

VII. DAFTAR PUSTAKA

1. Gonzalez. R.C., dan Woods. R. E., 2008, *Digital Image Processing*, Prentice Hall.
2. Marvin C. W. & Projono. A., *Pengolahan Citra Digital Menggunakan MatLAB*, Informatika, 2007
3. MATLAB Reference.
4. Schalkoff, R., *Pattern Recognition: Statistical, Syntactic and Neural Approaches*, John Wiley & Sons, 1992