

## **ANALISIS EKSTRAKSI FITUR MENGGUNAKAN COLOR HISTOGRAM, MOMENT, GRAY LEVEL DIFFERENCE VECTOR**

### **ANALYSIS FEATURE EXTRACTION USING COLOR HISTOGRAM, MOMENT, GRAY LEVEL DIFFERENCE VECTOR**

**Jimmy Eduard Rizky<sup>1</sup>, Tjokorda Agung Budi Wirayuda<sup>2</sup>, Retno Novi Dayawati<sup>3</sup>**

<sup>1</sup>Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

---

#### **Abstrak**

Image Retrieval adalah salah satu bidang dimana suatu informasi dari suatu image diambil berdasarkan fitur-fitur yang terdapat pada suatu image. Setiap image mengandung suatu informasi, seperti halnya dokumen teks. Teknik yang dipakai untuk menggali informasi bisa menggunakan teks, content, ataupun gabungan dari keduanya. Image retrieval dimana teknik pencarian citra yang didasarkan atas informasi dari isi citra tersebut disebut Content Based Image Retrieval.

Pada Tugas Akhir ini, CBIR yang dikembangkan berdasarkan fitur warna, bentuk, dan tekstur. Fitur dari warna diekstraksi dengan menggunakan Color Histogram, dimana metode ini memanfaatkan nilai kemunculan dari setiap warna pada citra. Untuk ekstraksi bentuk, digunakan metode Moment Invariant. Moment invariant ini menggunakan tujuh nilai vektor yang digunakan sebagai vektor ciri yang konstan terhadap perubahan geometri. Dan untuk fitur tekstur diekstraksi dengan menggunakan metode Gray Level Difference Vector, dimana metode ini merupakan lanjutan dari metode Gray Level Co-occurrence Matrix.

Hasil penelitian menunjukkan bahwa setiap metode ekstraksi fitur ini mempunyai performansi yang berbeda untuk setiap jenis citra. Pada hasil penelitian didapatkan hasil bahwa ekstraksi menggunakan fitur tekstur, mempunyai performansi yang merata terhadap semua jenis citra. Dan penggabungan esktraktor ciri yang dilakukan juga belum bisa memberikan nilai performansi yang lebih bila digunakan ciri fitur secara individu.

**Kata Kunci : content based image retrieval, color histogram, moment invariant, gray level difference vector**

---

#### **Abstract**

Image Retrieval is one area where some information from an image taken based on the features contained in an image. Each image contains some information, such as text documents. The technique can be used to dig information using text, content, or combination of both. Image retrieval based on information from the contents of that image is called Content-Based Image Retrieval.

In this final project, which was developed CBIR based on color features, shape, and texture. Color feature is extracted using Color Histogram method which takes advantage of appearance of each color in the images. For the shape feature, is extracted using Moment Invariant. It uses seven invariant moment vector value that is used as a feature vector which is constant with changes in geometry. And for the extracted texture features using Gray Level Difference Vector method, which is a continuation of the method of Gray Level Co-occurrence Matrix.

The results showed that each of these feature extraction methods have different performance for each type of image. In research results showed that extraction using texture features, has a uniform performance for all types image. And the merger of these feature do not give better performance value than use individual feature.

**Keywords : content based image retrieval, color histogram, moment invariant, gray level difference vector**

# 1. Pendahuluan

## 1.1 Latar belakang masalah

*Image Retrieval* adalah salah satu bidang dimana suatu informasi dari suatu *image* diambil berdasarkan fitur-fitur yang terdapat pada suatu *image*. Setiap *image* mengandung suatu informasi, seperti halnya dokumen teks. Jika penggalian informasi tersebut didasarkan atas mata manusia, maka manusia dengan mudah melakukan pembacaan informasi, melakukan klasifikasi, dan melakukan pencocokan dari *image query* terhadap koleksi *image* yang ada. Akan tetapi jika hal tersebut dilakukan oleh komputer, maka hal itu sangat sulit dilakukan karena kemampuan dari komputer sangat terbatas.

Dalam melakukan penggalian informasi dari suatu *image*, biasanya dilakukan ekstraksi fitur-fitur yang ada pada *image* tersebut. Salah satu fitur dari *image* yang populer adalah *color*. *Color* dapat digunakan untuk memberikan deskripsi tentang informasi apa yang terkandung dalam *image*. Akan tetapi *color feature* ini hanya memberikan kesamaan dua buah *image* didasarkan atas distribusi dari warna saja. Sehingga bisa terjadi suatu *image* dikatakan sama atas dasar distribusi warna saja, padahal informasi yang terkandung bisa berbeda jika dilihat dari *content* yang ada pada *image*. Fitur *image* lainnya yang bisa digunakan untuk ekstraksi adalah bentuk (*shape*). Bentuk bisa digunakan untuk mendeskripsikan objek yang terdapat pada *image*. Suatu ekstraktor fitur bentuk harus bisa mengenali objek pada citra yang sudah mengalami perubahan geometri seperti translasi, rotasi, dan penskalaan. Fitur lainnya yang bisa digunakan untuk mendeskripsikan informasi pada citra adalah dengan menggunakan fitur tekstur (*texture*). Tekstur mendeskripsikan informasi yang dimiliki pada suatu *image* dengan menggambarkan pengulangan pola yang terdapat pada citra. Setiap fitur mempunyai kelebihan dan kekurangannya bila digunakan secara individual untuk mendeskripsikan informasi yang terkandung pada *image*.

Untuk menutupi kekurangan yang terdapat pada suatu fitur dapat diatasi dengan melakukan kombinasi dari *color feature*, *shape feature*, dan *texture feature*. Hasil yang diharapkan dari penggunaan metode ini adalah mengurangi ketidakrelevanan suatu *image* akan ter-*retrieve*. Sehingga bisa memberikan nilai keakuratan tersendiri. Karena dengan menggunakan beberapa fitur dasar dari *image*, bisa mengurangi kekurangan dari ekstraksi ciri hanya didasarkan atas satu fitur saja yang ada pada *image*.

## 1.2 Perumusan masalah

Rumusan masalah yang ada pada Tugas Akhir ini yaitu sebagai berikut :

1. Bagaimana menerapkan pemrosesan *image* pada komputer, dari melakukan ekstraksi informasi dari suatu *image* dan melakukan pencocokan *image* yang terdapat dalam *database image*.
2. Bagaimana mengimplementasikan pemakaian *color histogram*, *moment invariants*, dan *gray-level difference vector* (GLDV) dalam mengekstraksi fitur citra.
3. Bagaimana pengaruh pengimplementasian dari metode ekstraksi fitur ciri yang ada, bila diimplementasikan secara individual ataupun digabungkan untuk melihat keakuratan sistem *image retrieval*.

4. Bagaimana mengimplementasikan metode *canberra distance* untuk mengetahui jarak dua buah vektor hasil ekstraksi *image*.

Adapun batasan masalah dari Tugas Akhir ini adalah :

1. Pada implementasi untuk pengujian dari hasil analisis digunakan beberapa *library* pendukung guna memudahkan dalam proses pengembangan sistem CBIR (bukan *library* dari metode yang akan digunakan). Adapun *library* yang digunakan adalah sebagai berikut *imadjust*, *rgb2ind*, *rgb2gray*, *im2bw*, *bwmorph*, dan *graycomatrix*.
2. Sistem yang dibangun digunakan sebagai implementasi dari analisis yang dilakukan dan digunakan sebagai pengujian.
3. *Query* yang dimasukan adalah *query* berjenis *image*, bukan teks.
4. Dalam Tugas Akhir ini yang dibahas bagaimana cara mengekstraksi suatu fitur yang terdapat pada *image*. Untuk proses pencarian yang terbaik tidak dibahas pada penelitian ini.
5. *Moment invariants* yang dipakai untuk ekstraksi ciri berdasarkan *shape* menggunakan *moment invariants* ordo yang kedua dan ketiga.
6. Untuk menghitung nilai kemiripan dua buah vektor digunakan *canberra distance*.

### 1.3 Tujuan

Adapun tujuan yang ada pada Tugas Akhir ini adalah sebagai berikut :

1. Menerapkan pemrosesan *image* pada komputer, dari melakukan ekstraksi ciri sampai pencocokan tingkat kemiripan antar citra.
2. Menganalisa bagaimana kinerja *image retrieval* dengan menggunakan *precision* dan *recall* jika :
  - 1) Menggunakan *color feature* saja.
  - 2) Menggunakan *shape feature* saja.
  - 3) Menggunakan *texture feature* saja.
  - 4) Menggabungkan *feature-feature* yang terdapat pada *image*.

### 1.4 Metodologi penyelesaian masalah

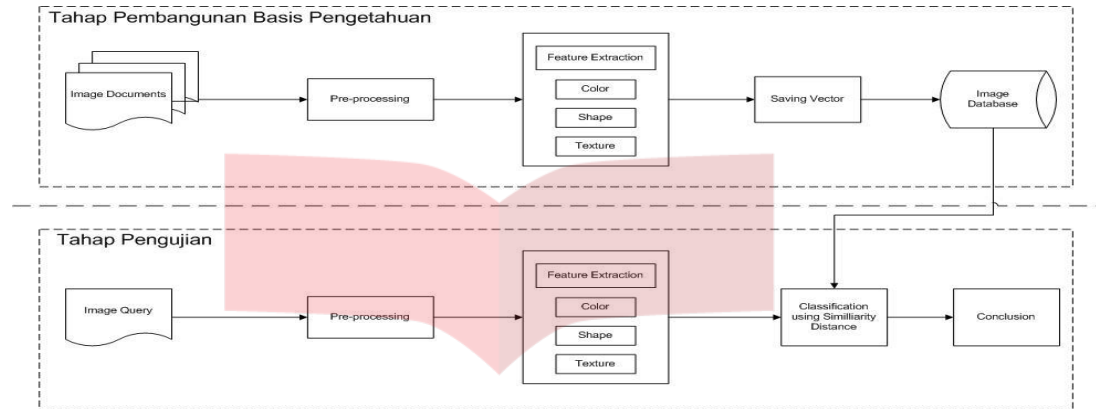
Metodologi yang digunakan dalam menyelesaikan Tugas Akhir ini adalah sebagai berikut :

1. Studi Literatur

Pada tahap ini dilakukan pengumpulan data untuk referensi dalam pembuatan Tugas Akhir ini. Kemudian mempelajari literatur-literatur yang berkaitan dengan konsep *image processing*, *image retrieval*, deskripsi dan teori tentang *color histogram*, *moment invariants*, *gray-level difference vector* (GLDV), *canberra distance* dan cara pemakaian *software* Matlab untuk membangun perangkat lunak yang digunakan untuk pengujian.

2. Analisis dan Perancangan

Tahapan ini meliputi analisa kebutuhan, pengumpulan data yang diperlukan, dan melakukan perancangan untuk tahap implementasi. Adapun perancangan dari sistem yang akan dibangun dapat dilihat pada Gambar 1-1:



Gambar 1-1 : Arsitektur sistem yang dibangun

3. Implementasi

Tahap ini melakukan implementasi pembangunan sistem CBIR dengan menggunakan Matlab dan mengimplementasikan metode *color histogram*, *moment invariants*, *gray-level difference vector (GLDV)*, *canberra distance*.

4. Pengujian dan Analisis Hasil Implementasi

Pada tahap ini dilakukan pengujian terhadap sistem CBIR yang telah dibangun dan melakukan analisa terhadap output yang dihasilkan oleh sistem. Pada pengujian yang dilakukan ada beberapa skenario yang akan dilakukan yaitu :

- Jenis pengujian yang dilakukan yaitu berdasarkan *color feature* saja, *shape feature* saja, *texture feature* saja, dan gabungan dari semua *feature* tersebut.
- Jenis *image* yang akan diuji memiliki kualitas yang berbeda-beda, misalnya ada *image* yang kualitasnya bagus, lalu ada *image* yang mempunyai *noise*.
- Untuk menguji hasil dari fitur warna akan diuji tipe gambar yang mempunyai konten dan distribusi warna yang beragam.
- Untuk pengujian pada fitur bentuk akan diuji tipe gambar yang mempunyai bentuk yang sama akan tetapi sebelumnya dilakukan operasi geometri.
- Untuk pengujian pada fitur tekstur akan digunakan tipe gambar yang mempunyai tingkat kompleksitas yang beragam. Hal ini dikarenakan pola yang dihasilkan juga dipengaruhi oleh faktor kompleksitas dari gambar tersebut.

Kemudian dari hasil pengujian ini akan dianalisa. Analisa ini dilihat dari keakuratan *image* yang di-*retrieve* (berdasarkan nilai *precision* dan *recall*).

5. Pembuatan Laporan

Mendokumentasikan tahap-tahap yang telah dilakukan mulai dari studi literatur sampai analisis hasil testing, dan menarik kesimpulan.

## 1.5 Sistematika penulisan

Sistematika penulisan pada tugas akhir ini seperti yang tercantum berikut :

### BAB 1 : PENDAHULUAN

Dalam bab ini dijelaskan tentang latar belakang, perumusan masalah, tujuan, metodologi penyelesaian masalah, dan sistematika penulisan.

### BAB 2 : LANDASAN TEORI

Bab ini menjelaskan tentang landasan teori yang akan menunjang dalam pembuatan tugas akhir ini.

### BAB 3 : ANALISIS SISTEM

Di bab ini dijelaskan tentang analisis kebutuhan dari sistem yang dibangun, beserta tentang perancangan dari sistem yang dibangun.

### BAB 4 : PENGUJIAN SISTEM DAN ANALISIS

Pada bab ini dijelaskan tentang pengujian terhadap setiap ekstraksi fiturnya dan dilakukan analisa terhadap hasil pengujiannya.

### BAB 5 : PENUTUP

Dalam bab ini merupakan penutup, berisi tentang kesimpulan yang dapat diambil berdasarkan data yang ada dan saran yang selayaknya dilakukan bila tugas akhir ini akan dilanjutkan kembali.

## 5. Penutup

### 5.1 Kesimpulan

Dari hasil penelitian yang telah dilakukan pada ekstraksi dengan menggunakan *color histogram*, *moment invariant*, dan *gray level difference vector* dapat diambil kesimpulan sebagai berikut :

1. Penggunaan *color histogram* sebagai ekstraktor ciri warna, mempunyai performansi yang bagus pada data citra yang mempunyai distribusi warna yang dominan seperti pada kelas citra *seaside*, *bus*, *horse*, dan *flower*. Dan kualitas dari vektor ciri yang dihasilkan juga dipengaruhi oleh penggunaan jumlah warna yang digunakan untuk proses kuantisasinya.
2. Pada penggunaan kuantisasi 64 warna dengan 128 warna, ternyata performansi kuantisasi 64 warna memberikan nilai yang lebih bagus. Hal ini dengan melihat aspek nilai *precision* dan *recall* yang diberikan dengan waktu komputasi yang diperlukan.
3. Penggunaan *moment invariant* sebagai ekstraktor ciri bentuk, mempunyai performansi yang bagus pada data citra yang gambar *background*-nya tidak kompleks, sehingga citra dapat tersegmentasi dengan baik (*foreground* dan *background* dapat terpisah) seperti pada citra *dinosaur*, *circle*, dan *triangle*. Untuk kualitas dari vektor ciri yang dihasilkan sangat dipengaruhi dari penggunaan metode segmentasi yang bagus. Hal ini dikarenakan pada proses segmentasi terjadi pemisahan antara *foreground* dan *background*-nya. Jika pada proses segmentasinya tidak bisa melakukan pemisahan antara *foreground* dan *background*-nya secara sempurna, maka akan berakibat pada hasil vektor ciri yang dihasilkan tidak bagus.
4. Penggunaan *gray level difference vector* sebagai ekstraktor ciri tekstur dapat diaplikasikan pada citra yang beragam, baik itu *background*-nya kompleks, ataupun sederhana. Dan dikarenakan metode GLDV merupakan lanjutan dari metode GLCM, maka kualitas dari vektor ciri yang dihasilkan juga dipengaruhi oleh besarnya sudut pembacaan yang digunakan.
5. Sudut pembacaan GLCM yang bagus terdapat pada penggabungan penggunaan sudut  $0^0$ ,  $45^0$ ,  $90^0$ , dan  $135^0$ .
6. Penggabungan ekstraktor ciri belum tentu menghasilkan performansi yang lebih bagus bila dibandingkan dengan penggunaan ekstraktor ciri secara individual. Untuk mendapatkan performansi yang bagus dari penggabungan ekstraktor ciri, setiap ekstraktor ciri harus mempunyai performansi yang bagus jika digunakan secara individu, dan dapat diaplikasikan untuk semua jenis citra.
7. Bila dilihat dari nilai ambang batas (*threshold*) yang diberikan untuk digunakan sebagai prosentasi tingkat kemiripannya, kombinasi penggabungan ekstraktor ciri mempunyai perubahan yang relatif lebih stabil, jika dibandingkan dengan penggunaan secara individu atau mandiri. Hal ini disebabkan karena semakin kecilnya data citra yang tidak relevan yang terambil pada proses pencariannya.

8. Tahapan praproses merupakan tahapan yang paling penting sebelum dilakukan ekstraksi fitur dan dapat mempengaruhi kualitas dari vektor ciri yang dihasilkan.

## 5.2 Saran

Saran-saran yang dapat diambil berdasarkan hasil penelitian yang dapat digunakan untuk penelitian selanjutnya antara lain :

1. Untuk penelitian selanjutnya sebaiknya parameter pemilihan praproses juga ikut disertakan sebagai parameter pengujian. Sehingga diharapkan dapat mendapatkan kombinasi praproses yang dapat menghasilkan performansi yang lebih bagus.
2. Pada penelitian selanjutnya, sebaiknya perlu ditambahkan modul untuk melakukan *relevance feedback*, guna meningkatkan hasil performansi dari data yang terambil dan relevan.



## Daftar Pustaka

- [1] Acharya, Thinku, dan Ajoy K. Ray. "*Image Processing: Principle and Applications*". New Jersey : John Willey and Sons, Inc. 2005.
- [2] Anonim. "Binari Image". [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/OWENS/L\\_ECT2/node3.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/OWENS/L_ECT2/node3.html). Diakses tanggal 29 Juni 2010
- [3] Away, Gunaidi Abdia. "*The Shortcut of Matlab Programming*". Bandung : Penerbit Informatika. 2006.
- [4] Beyer, Mryka Hall. <http://www.fp.ucalgary.ca/mhallbey/tutorial.htm>. Diakses tanggal 29 Juni 2010.
- [5] Gonzales, Rafael C, dan Richard E. Woods. "*Digital Image Processing*". 2<sup>nd</sup> ed. New Jersey : Prentice Hall. 2002.
- [6] Hunt, Brian R., Ronald L. Lipsman, dan Jonathan M. Rosenberg. "*A Guide to Matlab for Beginners and Experienced Users*". New York : Cambridge University Press. 2001.
- [7] Jain, Suyog Dutt. "*Content Based Image Retrieval*". Slide presentasi Manipal Institute of Technology.
- [8] Jeong, Sangoh. <http://scien.stanford.edu/pages/labsite/2002/psych221/projects/02/sojeong/>. Diakses tanggal 10 Maret 2010.
- [9] Nixon, Mark, dan Alberto Aguado. "*Feature Extraction and Image Processing*". 2nd ed. London : Elsevier. 2008.
- [10] Wijaya, Marvin Ch, dan Agus Prijono. "*Pengolahan Citra Digital Menggunakan Matlab*". Bandung : Penerbit Informatika. 2007.
- [11] Pratt, William K. "*Digital Image Processing*". 3<sup>rd</sup> ed. New York : John Willey and Sons, Inc. 2001.
- [12] Putra, Darma. "*Sistem Biometrika : Konsep Dasar, Teknik Analisis Citra, dan Tahapan Membangun Aplikasi Sistem Biometrika*". Yogyakarta: Penerbit Andi. 2009.
- [13] Register, Andi H. "*A Guide To Matlab Object Oriented Programming*". New York : Scitech Publishing. 2007.
- [14] Shutler, James. "Statistical Moment". [http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/SHUTLER3/CVonline\\_moments.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/SHUTLER3/CVonline_moments.html). Diakses tanggal 29 Juni 2010.
- [15] Sutoyo, T, Edi Mulyanto, Vincent Suhartono, Oky Dwi Nurhayati, dan Wijanarto. "*Teori Pengolahan Citra Digital*". Yogyakarta : ANDI. 2009
- [16] W, Tjokorda Agung Budi. "*Pengantar Pengolahan Citra*". Bandung. 2008.