

ANALISIS DAN IMPLEMENTASI CONTENT BASED IMAGE RETRIEVAL BERDASARKAN CIRI WARNA, TEKSTUR DAN BENTUK

Abdul Qadir¹, Tjokorda Agung Budi Wirayuda², Retno Novi Dayawati³

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Kebutuhan akan sistem yang mampu melakukan pencarian image sesuai keinginan user dan seiring dengan terus berkembangnya ukuran basis data image membuat metode pencarian image dengan kata kunci berupa teks tidak lagi bisa diandalkan karena kata kunci yang diinputkan mungkin tidak sesuai dengan image yang diharapkan sehingga dengan memberikan kata kunci saja tidak cukup. Hal ini disebabkan pemberian nama image bisa bersifat tidak objektif.

Dari permasalahan tersebut dibangunlah sistem yaitu Content Based Image Retrieval (CBIR) dengan menerapkan ekstraksi ciri warna (Color Moments), ciri tekstur (Haar Wavelet) dan ciri bentuk (Centroid Contour Distance) untuk mendapatkan image yang sesuai dengan image yang dicari. Ekstraksi fitur warna menggunakan Color Moments memanfaatkan distribusi probabilitas warna sebuah image yang terdiri dari 3 moments yaitu : mean, standard deviation, dan skewness. Sedangkan untuk ekstraksi tekstur dengan melakukan dekomposisi Haar Wavelet dengan pendekatan pyramid-structured wavelet transform (PWT) yang menghasilkan 12 fitur ciri. Untuk ekstraksi bentuk menggunakan Centroid Contour Distance dengan menghitung jarak dari centroid ke tepi objek dengan menggunakan perhitungan sudut 50 dan menghasilkan 72 fitur ciri.

Hasil penelitian menunjukkan bahwa penggabungan ketiga metode secara paralel menghasilkan akurasi dan performansi yang lebih baik serta menghasilkan peningkatan nilai F-Measure dari pada penggabungan secara serial maupun individu. Akurasi yang diperoleh 96.39% dan performansi 43.77% serta selisih nilai F-Measure 19.65%.

Kata Kunci : content based image retrieval, color moments, haar wavelet, centroid contour distance.

Abstract

Along with the continued development of image database making searching method a image based on keyword is not enough. Because, name of image can be given not objective.

These problems built Content Based Image Retrieval based on color features extraction (Color Moments), texture features (Haar Wavelet) and shape features (Centroid Contour Distance) to obtained an appropriate with image query. Extraction of color features with Color Moments which is use a color probability distributions image that consist of 3 moments : mean, standar deviation and skewness. To extracting texture used decomposition Haar Wavelet approach pyramid-structured wavelet transform (PWT) which produces 12 feature texture. For shape extraction with Centroid Contour Distance calculating the distance from centroid to the edge of object. Calculating used angle 50 dan produces 72 shape feature.

The result showed that paralel combination of 3 method has the accuracy and performance result better also increased value of F-Measures than serial combination or without combination. Accuracy obtained 96.39% and performance 43.77% with difference value of F-Measures 19.65%.

Keywords : content based image retrieval, color moments, haar wavelet, centroid contour distance.

1. Pendahuluan

1.1 Latar belakang

Kebutuhan akan sistem yang mampu melakukan pencarian *image* sesuai keinginan user, seiring dengan terus berkembangnya ukuran basis data *image* membuat metode pencarian *image* dengan kata kunci berupa teks tidak lagi bisa diandalkan karena kata kunci yang diinputkan mungkin tidak sesuai dengan *image* yang diharapkan sehingga dengan memberikan kata kunci saja tidak cukup. Hal ini disebabkan, pemberian nama terhadap sebuah *image* mungkin tidak bersifat objektif dan mendeskripsikan sebuah *image* hanya dengan beberapa kata mungkin tidak terlalu tepat untuk mewakili semua informasi yang terdapat pada *image* tersebut, karena *image* mengandung beragam makna [2]. Oleh karena itu, dibutuhkan sebuah sistem untuk mengatasi permasalahan tersebut yaitu *Content Based Image Retrieval* (CBIR).

CBIR adalah salah satu metodologi untuk pemanggilan kembali data *images* berdasarkan *content* sebuah *image* [1]. CBIR mengekstrak *feature vector* yang merepresentasikan *images properties* seperti warna (*color*), tekstur (*texture*) dan bentuk (*shape*) Sehingga memungkinkan untuk *re-trieve images* yang mempunyai kemiripan dengan *images* yang diinginkan *user* (*query-by-example*) [2]. Fitur warna merupakan fitur visual yang paling banyak digunakan dalam CBIR. Warna sudah banyak diaplikasikan dalam pencarian *image* karena memiliki hubungan yang kuat dengan obyek dalam citra. Selain itu fitur warna *robust* dalam komplikasi *background*, perskalaan, orientasi, perspektif dan ukuran citra [4]. Salah satu metode yang dapat digunakan untuk ekstraksi warna adalah *Color Moments*. Metode ini mengasumsikan distribusi warna dari sebuah *image* sebagai distribusi probabilitas. Dalam merepresentasikan penyebaran warna dari sebuah *image* digunakan tiga moment, yaitu : *mean*, *deviation*, *skewness* [3].

Tekstur merupakan karakteristik yang sangat penting untuk menganalisa permukaan suatu objek. Fitur tekstur (*texture*) merupakan salah satu ciri yang menggambarkan pola-pola dari permukaan objek dalam suatu citra [6]. Metode *Haar Wavelet* mengekstrak ciri-ciri tekstur dengan melakukan proses *averaging* untuk mendapatkan bagian dari gambar yang berfrekuensi rendah (*Low Frequency*) dan melakukan proses *differencing* untuk mendapatkan bagian dari gambar yang berfrekuensi tinggi (*High Frequency*)[8]. Kemudian dilakukan transformasi *Haar Wavelet* pada setiap baris dan kolom sehingga menghasilkan bagian LL (*Low Low*), LH (*Low High*), HL (*High Low*), HH (*High High*). Selanjutnya transformasi *Haar Wavelet* dilakukan pada bagian LL sehingga dihasilkan 4 bagian lagi. Setiap bagian ini akan dihitung nilai *mean absolute* dan *variansi*. Nilai *mean absolute* dan *variansi* ini dianggap sebagai ciri tekstur pada *image*.

Fitur bentuk (*shape*) merupakan fitur pokok sebuah *image* yang dapat merepresentasikan bentuk sebuah *image*. Salah satu metode untuk mendapatkan fitur bentuk dari sebuah *image* adalah *Centroid Contour Distance*. Metode ini dapat merefleksikan karakter umum dari suatu bentuk dengan menghitung jarak dari *centroid* suatu bentuk dengan batasannya berdasarkan kelipatan sudut [5].

Setiap jarak antara sisi tepi objek dengan *centroid* suatu bentuk yang diperoleh akan dianggap sebagai ciri dari bentuk tersebut.

Pada tugas akhir ini, proses *retrieval* akan dilakukan dengan menggunakan kombinasi fitur warna, tekstur dan bentuk. Dengan kombinasi tersebut diharapkan dapat meningkatkan performansi *image retrieval* karena *image* yang tidak relevan difilter dengan tiga kombinasi (warna, tekstur dan bentuk) sehingga dapat menghasilkan citra akhir dengan kemiripan dan akurasi yang lebih tinggi dengan citra *query input*.

1.2 Perumusan masalah

Dalam tugas akhir ini dirumuskan beberapa masalah sebagai berikut :

1. Bagaimana mengimplementasikan metode *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance* dalam mengekstrak fitur citra.
2. Bagaimana mendapatkan teknik CBIR untuk mendapatkan citra yang sesuai dengan kelas *image query*.
3. Bagaimana mengukur kemiripan antara *image query* dengan *image* yang ada pada database serta mengukur keakuratan sistem dalam menghasilkan citra relevan.
4. Bagaimana pengaruh kombinasi warna, tekstur dan bentuk pada *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance* dengan keakuratan sistem *image retrieval*.

Adapun batasan masalah dalam tugas akhir ini adalah sebagai berikut :

1. Data berupa *image* dengan tipe bitmap dengan ukuran gambar 256x256 piksel.
2. *Image* yang digunakan adalah memiliki kategori seperti : bunga, brick, binatang dan daun. Bunga, dinosaurus dan daun digunakan untuk pengujian warna, bentuk dan tekstur. Sedangkan brick hanya digunakan untuk pengujian tekstur.
3. Pengukuran tingkat kemiripan antara *image* yang diuji dengan citra pada database diukur menggunakan perhitungan jarak dari vektor tiap-tiap fitur.
4. Keakuratan sistem dalam menghasilkan sebuah *image* yang relevan diukur dengan *Precision* dan *Recall*.
5. Pengujian dilakukan dengan ciri warna saja, tekstur saja, bentuk saja dan mengkombinasikan ketiga ciri tersebut secara serial, yaitu warna dan bentuk, warna dan tekstur, bentuk dan tekstur serta gabungan warna, tekstur dan bentuk secara paralel.

1.3 Tujuan

Tujuan yang ingin dicapai dalam penelitian ini adalah :

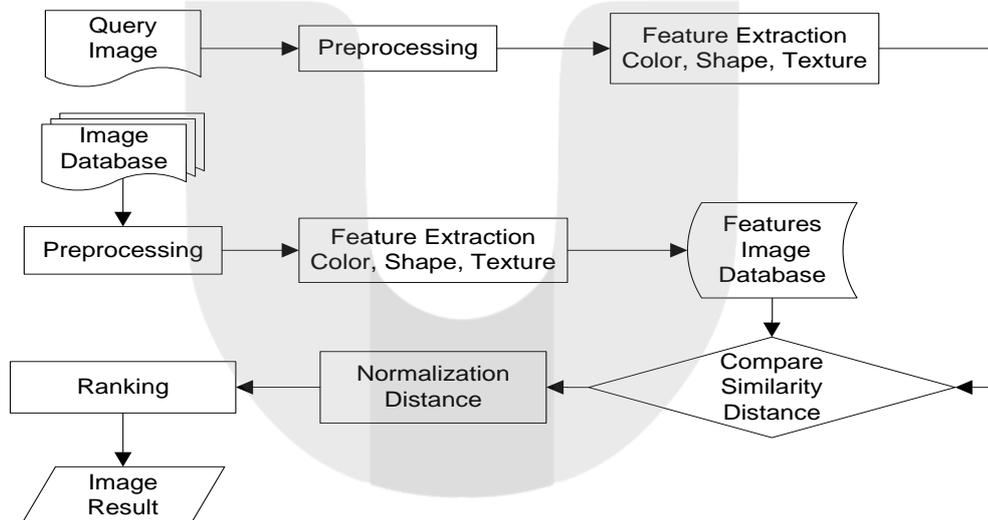
1. Mengimplementasikan *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance* dalam mengekstraksi fitur warna, tekstur dan bentuk.
2. Mengukur *image* relevan yang ter-*retrieved* menggunakan *precision* dan *recall*.
3. Menganalisis pengaruh kombinasi fitur warna, bentuk dan tekstur dalam peningkatan keakuratan sistem *image retrieval*.

4. Menganalisis keakuratan sistem *image retrieval* jika *image query* mengalami *blur*.

1.4 Metodologi penyelesaian masalah

Secara umum metode penelitian dilakukan adalah sebagai berikut :

1. Studi literatur
 Pada tahap ini, dilakukan pengumpulan data referensi untuk pembuatan Tugas Akhir. Kemudian mempelajari serta memahami konsep CBIR, *image processing*, *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance* melalui literatur-literatur yang relevan dengan permasalahan diatas. Selanjutnya mempelajari pemakaian *software* Matlab untuk membangun sistem untuk pengujian Tugas Akhir.
2. Analisis dan Desain
 Menganalisis kebutuhan fungsionalitas dari sistem *image retrieval*, pengumpulan data dan melakukan perancangan untuk tahap implementasi. Adapun skema proses pada sistem ini adalah sebagai berikut :



Gambar 1-1 : Desain Skema Proses

3. Implementasi
 Tahap ini meliputi implementasi pembangunan perangkat lunak menggunakan bahasa pemrograman Matlab dengan metode *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance*.
4. Pengujian dan Analisis Hasil
 Pada tahap ini dilakukan pengujian terhadap perangkat lunak dengan beberapa skenario dan menganalisis output yang dihasilkan oleh sistem. Adapun skenario pengujian adalah sebagai berikut :
 - Melakukan pengujian berdasarkan *feature* warna, bentuk, dan tekstur serta kombinasi dari ketiga *feature-feature* tersebut.
 - Melakukan pengujian untuk *image* yang mengandung *blur*.
 Kemudian hasil dari *image* yang ter-*retrieve* akan dianalisa berdasarkan *precision* dan *recall*.

5. Penyusunan Laporan

Pada tahap ini dilakukan penyusunan laporan hasil penelitian sesuai tahap-tahap dari studi literatur sampai analisis hasil *testing* dan mengambil kesimpulan, kemudian didokumentasikan ke dalam sebuah buku laporan Tugas Akhir.

Laporan yang disusun terdiri-dari :

1. Pendahuluan

Bab ini menjelaskan latar belakang penelitian, permasalahan yang diteliti, tujuan penelitian, dan metodologi penyelesaian masalah.

2. Landasan Teori

Landasan teori memberikan dasar-dasar pengetahuan yang dapat membantu penelitian.

3. Analisis Sistem

Bab ini memberikan gambaran dan penjelasan mengenai tahap-tahap pengerjaan sistem CBIR menggunakan metode *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance*.

4. Pengujian dan Analisis

Bab ini berisi hasil-hasil penelitian dan pengujian sistem yang telah dibangun. Kemudian dijelaskan juga hal-hal yang dianalisis serta bagaimana hasil analisis dari pengujian yang dilakukan.

5. Penutup

Bab terakhir ini memberikan kesimpulan yang didapatkan berdasarkan penelitian yang telah dikerjakan. Selain itu diberikan saran-saran yang dapat membantu pelaksanaan penelitian berikutnya yang dapat dilakukan.

5. Kesimpulan dan Saran

5.1 Kesimpulan

Dari hasil analisis terhadap penelitian yang dilakukan pada *Content Based Image Retrieval* berdasarkan ciri warna, tekstur dan bentuk dengan menggunakan metode *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance*, maka dapat diambil kesimpulan sebagai berikut :

1. Penggunaan metode *Color Moments* dalam mengekstrak ciri warna mempunyai keakuratan lebih baik dengan pemberian bobot yang berbeda pada setiap *moment* pada masing-masing kelas sangat mempengaruhi dalam menghasilkan *image* ter-retrieved yang relevan. Seperti pada *image* bunga.
2. Penggunaan metode *Haar Wavelet* mempunyai akurasi yang baik dalam mengekstrak tekstur pada *image* normal. Namun, metode ini mempunyai kelemahan ketika *image query* mengalami *blur* dimana terjadinya pengurangan nilai frekuensi tinggi pada *image*. Penggunaan frekuensi tinggi untuk mendapatkan ciri tekstur sangat bagus pada *image* seperti bunga, dinosaurus, daun dan brick.
3. Penggunaan metode *Centroid Contour Distance* mempunyai akurasi yang baik dalam mengekstrak ciri bentuk pada *image* normal. Namun, pada *image query* yang mengalami *blur* metode ini kurang bagus disebabkan proses segmentasi yang tidak berjalan dengan sempurna. Sehingga mempengaruhi dalam menentukan titik *centroid* dan jarak dari titik *centroid* ke sisi tepi. Pada metode ini kelas dinosaurus dan daun mempunyai akurasi yang bagus karena proses segmentasi yang berjalan dengan baik.
4. Penggabungan metode secara serial sangat dipengaruhi oleh metode ekstraksi terakhir, karena semakin bagus akurasi pada metode terakhir akan sangat menentukan *image* ter-retrieved yang benar-benar relevan.
5. Penggabungan metode secara paralel sangat dipengaruhi juga oleh tingginya nilai akurasi pada masing-masing metode. Kombinasi secara paralel memiliki akurasi dan performansi yang stabil dibandingkan dengan ekstraksi secara mandiri maupun penggabungan secara serial.
6. Akurasi relatif stabil pada penggunaan *threshold* dalam penggabungan ciri secara serial maupun paralel dibandingkan dengan penggunaan secara mandiri.

5.2 Saran

Hasil analisis terhadap penelitian yang telah dilakukan pada *Content Based Image Retrieval* berdasarkan ciri warna, tekstur dan bentuk dengan menggunakan metode *Color Moments*, *Haar Wavelet* dan *Centroid Contour Distance*. Berikut beberapa saran yang dapat digunakan untuk mengembangkan sistem ini.

1. Bandingkan *Color Moments* jika menggunakan model warna lainnya seperti *CIE LAB Color Space*.

2. Bandingkan *Haar wavelet* jika menggunakan mendekomposisi sebuah imager dengan *Tree-Structured Wavelet Transform* dengan dekomposisi penuh pada semua bagian *image* pada LL,LH,HL dan HH.
3. Bandingkan kombinasi fitur warna, tekstur dan bentuk ini dengan fitur ekstraksi lainnya.
4. Penggunaan *image* database yang lebih besar untuk mendapatkan hasil yang lebih akurat.



Daftar Pustaka

- [1] Wirawan, Setia. 2004. *Jurnal: Content based Image Information Retrieval*. Universitas Gunadarma. <http://ejournal.gunadarma.ac.id/files/B10.pdf> diakses tanggal 5 Desember 2010.
- [2] Torres, Ricardo da Silva., and Falcao, Alexandre Xavier. 2006. *Content-Based Image Retrieval: Theory and Applications*, <http://www.dcc.unicamp.br/~rtorres/artigos/journal/torres06rita.pdf>, diakses tanggal 10 Desember 2010
- [3] Noah Keen. *Color Moments*. 2005. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0405/KEEN/av_as2_nkeen.pdf. Diakses tanggal 10 Desember 2010
- [4] Acharya, Tinku., Ray, Ajoy K.(2005). *Image Processing Principles and Applications*. Wiley Interscience. Canada.
- [5] Levner, Ilya. *Shape Detection, Analysis, and Recognition*, http://www.cs.ualberta.ca/~ilya/courses/c615-PerceptualSystems/c615_Final_paper.pdf.gz, diakses terakhir tanggal 5 Januari 2011.
- [6] Kondekar V. H., Kolkure V. S., Prof.Kore S.N. 2010. <http://arxiv.org/ftp/arxiv/papers/1002/1002.1951.pdf> diakses terakhir tanggal 17 Maret 2011
- [7] Long, Lee Kuen and Hwei, Chen Ling. A New Method For Extracting Primitives Of Regular Textures Based On Wavelet Transform <http://debut.cis.nctu.edu.tw/Publications/pdfs/J37.pdf> , diakses terakhir tanggal 6 Januari 2011.
- [8] Sani Muhammad Isa, Elsa Juwita. 2007. <http://yudiagusta.files.wordpress.com/2009/11/221-226-snsi07-039-aplikasi-image-retrieval-berdasarkan-tekstur-dengan-menggunakan-transformasi-haar-wavelet.pdf> diakses terakhir tanggal 5 januari 2011
- [9] Setiawan Hadi, 2005, *Pengembangan Metode Pra-pemrosesan Bagi Model Generatif Pengenalan Wajah Berbasis Model 3D Morphables*, Institut Teknologi, Bandung, http://students.itb.ac.id/~setiawan.hadi/webS3/33503505/Laporan/LP_IF80ZA.pdf, diakses tanggal 11 Januari 2011
- [10] Wen Chen¹, Yun Q.Shi¹, Guorong Xuan². *Identifying Computer Graphs Using HSV Color Model and Statistical Moments of Characteristic Functions*.¹New Jersey Institute of Technology, USA. ²Tongji University, China, <http://web.njit.edu/~shi/PaperDownload/forensics/ICME07-CG.pdf>, diakses tanggal 10 Desember 2010
- [11] Tjokorda Agung BW, ST, MT. 2009. CS3204 Slide Perkuliahan, *Image Segmentation*. Departemen Informatika IT Telkom, Bandung.
- [12] <http://yudiagusta.files.wordpress.com/2009/11/321-327-knsi08-057-flower-image-retrieval-berdasarkan-color-moments-centroid-contour-distance-dan-angle-code-histogram.pdf> diakses terakhir tanggal 10 Desember 2010
- [13] Putra, Darma, 2010. "Pengolahan Citra Digital". Andi. Yogyakarta.
- [14] Tjokorda Agung, BW. Mela Rahmawati. Retno Novi Dayawati. Ade Romadhony. 2009. *Proses Up-S calling Citra Digital pada Domain Frekuensi dengan menggunakan Metode Discrete Wavelet Transform*. Makalah KNSI. Bali
- [15] Hong, An-xiang., Chen, Gang., Li, Jun-li; Chi, Zhe-ru., and Zhang. A flower image retrieval method based on ROI feature. <http://www.zju.edu.cn/jzus/2004/0407/040704.pdf> , diakses terakhir tanggal 20 Desember 2010.
- [16] Image Database http://www.vision.caltech.edu/Image_Datasets/Caltech101/101_Object_Categories.tar.gz diakses 20 Februari 2011
- [17] Sutoyo, T, Edi Mulyanto, Vincent Suhartono, Oky Dwi Nurhayati, dan Wijanarto. *Teori Pengolahan Citra Digital*. Yogyakarta : ANDI. 2009
- [18] Wijaya Marvin, CH, Prijono Agus, 2007, *Pengolahan Citra Digital Menggunakan Matlab*. Informatika. Bandung.