

ANALISIS SELEKSI CITRA MIRIP DENGAN MEMANFAATKAN KONSEP CBIR DAN ALGORITMA THRESHOLD

A. Haris Rangkuti

Computer Science Department, School of Computer Science Binus University
Jl. K.H. Syahdan No. 9, Palmerah, Jakarta Barat 11480
rangku2000@yahoo.com

ABSTRACT

Content base image retrieval (CBIR) is the concept of image retrieval by comparing the existing image on the sample to that of the database (query by example). CBIR process based on color is carried out using adaptive color histogram concept, while one based on shape is performed using moment concept. Following up the process results, a sorting process is done based on a threshold value of the sample image through the utilization threshold algorithm. The image displayed is be sorted from the one that is nearly similar to the query image (example) to the resemblance of the lowest (aggregation value). The threshold value of the query image used as reference is compared with the aggregation value of the image database. If the comparison in the search for similarities by using the concept of fuzzy logic approaches 1, the comparison between the threshold value and the aggregation value is almost the same. Otherwise, if it reaches 0, the comparison results in a lot of differences.

Keywords: CBIR, database, threshold algoritma, similarity, query by example, adaptive histogram, invariant moment, threshold value, aggregation value, logika fuzzy

ABSTRAK

Content base image retrieval (CBIR) merupakan konsep temu kembali citra dengan membandingkan citra yang ada pada contoh dengan yang ada pada basis data (query by example). Proses CBIR berdasarkan warna dilakukan menggunakan konsep adaptive histogram; sedangkan yang berdasarkan bentuk dilakukan menggunakan konsep invariant moment. Dari hasil proses pencarian berdasarkan bentuk dan warna, dilakukan proses pengurutan berdasarkan nilai ambang (threshold) dari contoh citra melalui pemanfaatan algoritma threshold. Citra yang ditampilkan diurutkan dari yang hampir sama atau serupa dengan citra query (contoh) hingga kemiripan yang terendah (nilai agregasi). Nilai ambang citra query yang digunakan sebagai acuan dibandingkan dengan nilai agregasi dari basis data citra. Jika proses perbandingan dalam mencari kemiripan dengan menggunakan konsep logika fuzzy mendekati angka 1, perbandingan antara nilai ambang dan nilai agregasi hampir sama. Sebaliknya, jika mendekati angka 0, perbandingan menjadi sangat jauh berbeda.

Kata kunci: CBIR, basis data, algoritma threshold, kemiripan, query by example, adaptive histogram, invariant moment, nilai ambang, nilai agregasi, logika fuzzy

PENDAHULUAN

Secara umum gambaran tentang sistem temu kembali citra yang tersimpan di basis data adalah dengan melakukan *query* ke dalam atribut *file*. Dengan beberapa informasi tambahan dalam atribut *file* bervariasi seperti tanggal pembuatan, judul, subyek, komentar dan kata kunci yang berbasis teks, tergantung sistem operasi yang digunakan. Namun, proses *query* akan menjadi tidak mudah, jika proses *query* tidak terhadap informasi tambahan dari image. Terdapat beberapa metode lain yang digunakan, di antaranya proses *query* dengan menggunakan citra berbasis ciri (*image feature*). Beberapa penelitian yang membandingkan antara ciri citra *query* dengan ciri citra di basis data, proses kemiripan yang terjadi dapat juga dilakukan dengan menggunakan logika *fuzzy*. Sehingga secara visualisasi akan terlihat beberapa citra yang mempunyai tingkat kemiripan yang tinggi. Disini peranan algoritma *threshold* yang merupakan turunan dari algoritma Fagin untuk membuat peringkat kemiripan, mulai dari yang hampir sama hingga kurang sama. Ini dilakukan bertujuan untuk meningkatkan kinerja sistem temu kembali.

Pada dasarnya penelitian seleksi citra dengan model CBIR ini terdiri dari tiga tahap yaitu ekstraksi ciri, pembuatan basis nilai jarak (*fuzzy*) dan pengukuran kinerja sistem. Tahap ekstraksi ciri merupakan proses segmentasi citra berdasarkan warna dan bentuk atau pengindeksan citra. Metode yang digunakan untuk segmentasi warna adalah histogram dengan bin warna yang sudah didefinisikan (warna referensi) dan metode untuk segmentasi bentuk adalah *binary* citra. Dalam penggunaan metode ini terlihat beberapa kendala karena untuk segmentasi bentuk, harus dipisahkan bentuk yang rumit atau yang sederhana. Dalam penelitian ini memfokuskan kepada bentuk yang sederhana dari buah. Oleh karena itu setiap gambar / citra yang disiapkan baik pada citra *query* dan citra basis data harus terlihat dengan jelas bentuknya dan tidak rumit. Tahap pembuatan basis nilai jarak (*fuzzy*) adalah tahapan untuk mengimplementasikan nilai jarak citra *query* dan citra basis data dalam sistem temukembali. Nilai jarak yang terjadi antara citra *query* dan basis data akan diurutkan berdasarkan nilai ambang yang didapat dari citra *query*.

Dalam menggunakan konsep seleksi citra *query*, nilai peringkat yang mendekati dengan citra *query* akan ditampilkan secara berurutan. Citra yang akan ditampilkan mulai dari peringkat tertinggi atau yang menyerupai dengan citra *query*, hingga peringkat yang terendah. Untuk proses menampilkan citra yang sudah terseleksi sesuai dengan peringkat tersebut, algoritma yang digunakan adalah algoritma *threshold* yang merupakan turunan dari algoritma Fagin. Dimana dengan algoritma tersebut, maka fokus pencarian atau pemanggilan data citra hanya yang menyerupai, atau hampir sama dengan data citra yang tersimpan pada basis data. Fokus penelitian ini mencakup dua hal yaitu keakuratan dan efisiensi waktu sistem temu kembali citra.

Dengan melakukan proses temu kembali citra salah satunya dapat dilakukan dengan memanfaatkan kemiripan ciri citra yaitu warna, tekstur dan bentuk. Namun metode ini masih memiliki masalah yang harus terus dilakukan penelitian yaitu: (1) terdapat perbedaan persepsi pengguna terhadap citra; (2) pemahaman secara subyektif terhadap pengguna sistem; (3) melakukan pembobotan pada masing-masing ciri citra sehingga sistem tidak bisa merefleksikan pola pikir manusia; (4) kapasitas cahaya yang berbedanya pada tiap citra; (5) warna *background* pada obyek yang tidak jelas; (6) terjadinya perbedaan skala beberapa obyek; (7) *edge detection* dari obyek yang kurang jelas (Xiaoling dan Kanglin, 2005).

Penelitian ini bertujuan untuk: (1) mengembangkan konsep CBIR yang berdasarkan warna dan bentuk dengan menggunakan algoritma *threshold* dan metode logika *fuzzy*; (2) analisis akurasi dan efisiensi sistem temu kembali citra dengan menggunakan algoritma *threshold* dengan menjalankan konsep CBIR; (3) untuk menghasilkan citra yang mirip dengan citra *query* dengan tingkat kebenaran sama dengan citra yang mirip menurut penglihatan mata manusia; (4) menjadi bahan referensi dalam melakukan pencarian data Image yang dapat dikembangkan pada kasus-kasus lainnya.

Sedangkan urgensi (keutamaan) penelitian ini adalah sebagai berikut: (1) dalam proses pencarian citra yang tersimpan dalam basis data, tingkat akurasi yang diharapkan dapat sesuai dengan kenyataan. Jumlah citra di basis data selalu ditambah tapi tidak merubah akurasi; (2) meningkatkan kinerja sistem temu kembali dalam melakukan proses pencarian citra dengan model CBIR; (3) menjadi bahan referensi dalam melakukan pencarian citra yang dapat dikembangkan pada kasus-kasus lain seperti kasus pada bidang kesehatan, ekonomi dan kebudayaan; (4) dapat menghasilkan sistem prototipe temu kembali citra menggunakan algoritma Fagin dan *threshold*; (5) sebagai salah satu metode untuk pengembangan aplikasi sistem temu kembali citra dalam berbagai bidang misalnya perpustakaan digital, sortasi dan klasifikasi.

Penelitian ini dibatasi pada ruang lingkup adalah sebagai berikut: (1) obyek penelitian adalah citra buah dalam format citra *JPG*, *PNG*, *BMP*; (2) segmentasi citra dilakukan pada warna dan bentuk; (3) analisis kemiripan secara visual menggunakan algoritma *threshold*; (4) analisis warna citra menggunakan metode *adaptive histogram* dan berdasarkan bentuk dengan menggunakan citra *binary*; (5) model warna yang digunakan *RGB,HSV*; (6) warna latar belakang polos atau tidak berwarna; (7) bentuk citra dipengaruhi oleh posisi dari citra, sehingga diharapkan posisi citra dalam keadaan tegak; (8) tidak dipengaruhi oleh *noise* citra.

Landasan Teori

Content Base Image Retrieval (CBIR)

Sistem temu kembali citra merupakan bidang yang mengintegrasikan berbagai disiplin ilmu antara lain: temu kembali informasi (*information retrieval*), pemrosesan citra (*digital image processing*) dan basis data (*database*). Penelitian dalam bidang sistem temu kembali citra ini dipicu oleh semakin pesatnya perkembangan internet dan *website* serta semakin mudahnya teknologi peralatan akuisisi citra (*imaging*), aplikasi perpustakaan digital (*digital libraries*), pengarsipan citra (*image archive*) dan *video-on-demand*.

Menurut Smeulders (2000), tantangan terbesar adalah pemanfaatan hasil-hasil penelitian dalam bidang ini tidak hanya terbatas untuk keperluan *computer vision* dan basis data, tetapi harus juga bisa digunakan untuk aplikasi di bidang lain seperti permodelan biologi, peramalan cuaca, citra medis, citra satelit dan lain-lain dalam bentuk *interactive image understanding*. Masalah utama dalam pengembangan sistem temu kembali citra adalah bagaimana mengintegrasikan antara konsep representasi visual *low-level* yang tidak ada korelasinya dengan konsep semantik *high-level* untuk mendapatkan informasi dari basisdata citra.

Menurut Smeulder et al (2000), terdapat tiga jenis pendekatan spesifikasi *query* citra. Pertama, *query* dengan bentuk spasial yang menghasilkan bentuk spasial yang berhubungan dengan nilai citra secara literal; (2) *query* dengan citra contoh yang memberikan masukan pada sistem berupa *array of pixel* dan melakukan *query* pada basis data sehingga menghasilkan urutan nominasi citra; (3) *query* dengan grup citra contoh yaitu melakukan seleksi citra sesuai dengan spesifikasi yang didefinisikan secara semantik. Hasil dari pendekatan ini adalah kumpulan citra dengan klasifikasi yang sama.

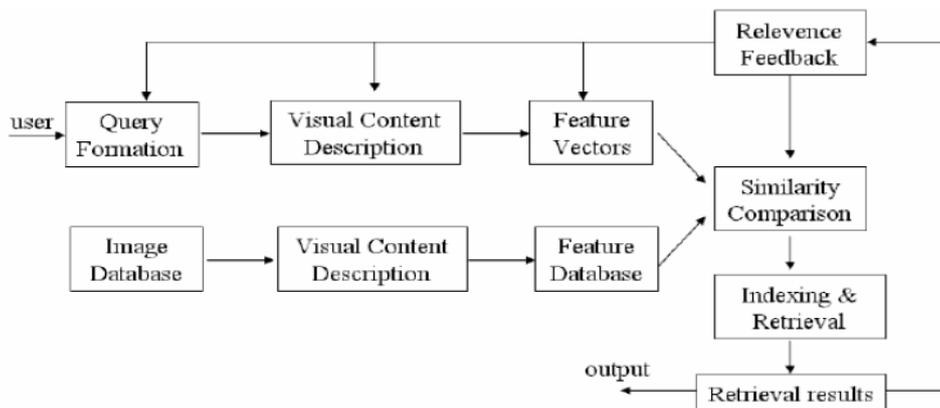
Perkembangan pesat penelitian dalam bidang ini pada umumnya lebih difokuskan pada metode untuk merepresentasikan ciri citra yaitu warna, tekstur dan bentuk. Ditambah lagi dengan adopsi logika *fuzzy* dalam sistem temu kembali citra yang bertujuan untuk meningkatkan kinerja sistem temu kembali yaitu keakuratan sistem temu kembali citra. Rancangan sistem temu kembali citra (Gambar 1) dibagi menjadi dua proses utama, yaitu pengindeksan dan temu kembali. Proses pengindeksan adalah proses melakukan segmentasi berdasarkan ciri citra yaitu ciri warna menggunakan metode *adaptive histogram* dan ciri bentuk menggunakan metode citra *binary*.

Sedangkan untuk mengurutkan ciri citra, proses pengimplementasian algoritma Fagin dan *threshold* terdiri dari tahapan *fuzzy*-fikasi nilai *crisp* ciri citra, agregasi serta visualisasi nominasi citra yang relevan.



Gambar 1. Seleksi citra di basis data menggunakan algoritma *threshold*.

CBIR merupakan konsep temu kembali citra dengan membandingkan citra yang ada pada contoh citra dengan yang ada pada basis data citra (*query by example*). CBIR dilakukan dengan membandingkan nilai jarak citra *query* dengan citra pada basis data (*image distance measure*). Pengukuran nilai jarak citra dapat dikelompokkan dengan tiga kelas utama, yaitu: kemiripan warna, kemiripan bentuk dan kemiripan tekstur. Secara umum diagram sistem CBIR diilustrasikan pada Gambar 2. Gambar tersebut menjelaskan bahwa *feature* atau ciri citra menjadi sangat menentukan dalam proses pencarian citra yang berdasarkan kemiripan yang berdasarkan warna, bentuk dan tekstur.



Gambar 2. Diagram umum konsep CBIR.

Warna – Model warna RGB (*red, green, blue*) mendeskripsikan warna sebagai kombinasi positif dari tiga warna, yaitu merah, hijau, dan biru. Ciri warna suatu gambar dapat dinyatakan dalam bentuk histogram dari Gambar 2. Histogram menghitung jumlah piksel pada masing-masing jenis warna dengan membaca masing-masing piksel citra hanya sekali dan menambah jumlahnya pada tempat penyimpanan yang tepat di histogram.

Bentuk – Ciri bentuk suatu gambar dapat ditentukan oleh tepi (sketsa), atau besaran moment dari suatu gambar. Pemakaian besaran moment pada ciri bentuk ini banyak digunakan dengan memanfaatkan nilai-nilai transformasi *fourier* dari gambar. Proses yang dapat digunakan untuk menentukan ciri bentuk adalah deteksi tepi, *threshold*, segmentasi dan perhitungan moment seperti (*mean*, *median* dan standar deviasi dari setiap lokal gambar).

Sedangkan tekstur merupakan karakteristik yang sangat penting untuk analisa permukaan berbagai jenis obyek. Istilah tekstur secara umum mengacu kepada pengulangan elemen–elemen dasar tekstur yang disebut *texel* yang tersebar secara periodik, kuasi periodik atau secara acak. Dari berbagai penelitian tentang penglihatan manusia diperoleh kesimpulan bahwa analisa ruang frekuensi atau multiskala lebih

Algoritma Threshold

Pada dasarnya algoritma Fagin sudah cukup optimal pada situasi tertentu. Namun pada situasi lain algoritma Fagin kurang menunjukkan performa yang kurang baik. Sebagai contoh bahwa tidak ada jaminan fungsi agregasi tidak tetap karena pada beberapa kasus jelas terlihat fungsi agregasi adalah konstant.

Berikut gambaran algoritma *threshold* menurut Fagin R (1998). Kerjakan akses pengurutan secara paralel untuk setiap pada m pada list urut $L1$. Sebagai objek R dilihat dibawah akses urutan pada beberapa list. kerjakan urutan secara random untuk list yang lain untuk mendapatkan peringkat $x1$ pada objek R dalam setiap list $L1$. Kemudian hitung peringkat $t(R)=t(x1..xm)$ pada objek R . Jika peringkat k ini adalah salah satu dari nilai tertinggi yang sudah dilihat, ingat objek R dan pada peringkat $t(R)$. Untuk setiap list $L1$, biarkan $x1$ menjadi tingkatan pada objek terakhir yang dilihat pada proses akses pengurutan. Definisikan nilai threshold τ menjadi $t(x1..xm)$. Sesegera sekurang kurangnya pada objek k yang sudah dilihat dimana peringkat adalah sekurang-kurangnya sama dengan τ kemudian keluar. Biarkan Y menjadi satu set yang berisi objek k yang sudah dilihat pada peringkat yang tertinggi, *ouput*-nya adalah set peringkat $\{(R,t(R))|R \in Y\}$. Threshold Algorithm adalah benar untuk setiap fungsi *t aggregation* yang *monoton*.

Pengukuran Kinerja Sistem

Dalam pemilihan citra buah yang beragam, didasari karena banyak sekali faktor-faktor yang mempengaruhi warna dan bentuk citra saat perekaman oleh kamera adalah sudut pandang dan pencahayaan. Perbedaan sudut pandang dan pencahayaan menyebabkan citra yang dihasilkan bervariasi warna dan bentuknya untuk setiap tema citra. Basis data citra yang digunakan dalam penelitian ini adalah citra dengan tema yang beragam. Hal ini dimaksudkan untuk menguji secara lebih mendetil keakuratan dan efisiensi algoritma Fagin dan *threshold*. Pemilihan tema citra yang beragam ini didasari karena banyak sekali faktor-faktor yang mempengaruhi warna dan bentuk citra, pada saat perekaman oleh kamera, atau sudut pandang dan pencahayaan. Perbedaan sudut pandang dan pencahayaan menyebabkan citra yang dihasilkan bervariasi warna dan bentuknya untuk setiap tema citra. Dalam penelitian ini penulis memfokuskan kepada citra buah terhadap faktor warna dan bentuk yang dimiliki citra buah.

Dalam menentukan kinerja sistem temu kembali pada citra yang berdasarkan bentuk dan warna dengan menggunakan algoritma Fagin dan *threshold*, akan menjadi benar dan akurat jika jumlah citra yang berhasil diidentifikasi dengan benar oleh sistem. Pengukuran kinerja sistem ini dapat menggunakan persamaan sebagai berikut:

Citra buah yang relevant (X)

$$X = \frac{\sum_y^x \text{citrayangrelevanhasilquery}}{\sum \text{citrayangrelevandalamDatabase}}$$

Citra buah yang presisi (Y)

$$Y = \frac{\sum_y^x \text{citrayangrelevanhasilquery}}{\sum \text{citradalamDatabase}}$$

Pengukuran Kedekatan Citra

Dalam melakukan pengukuran akan kedekatan citra, dapat dilakukan berdasarkan fungsi jarak antara dua ciri citra. Jarak digunakan untuk mengukur ke(tidak)miripan antara dua obyek data. Kemiripan merupakan salah satu landasan dari proses pencarian sebuah citra. Dimana dalam penelitian ini citranya adalah citra buah-buahan. Dengan ketentuan bahwa sudah disiapkan data buah-buahan ke dalam bentuk basis data.

Pada dasarnya terdapat banyak cara untuk menghitung jarak, namun pada tesis ini hanya akan dibahas tiga jarak yang paling banyak digunakan dan di antara ketiga jarak tersebut, yang paling populer adalah jarak Euclid.

Jarak Minkowski didefinisikan sebagai:

$$d(i, j) = \sqrt[q]{\left(|x_{i1} - x_{j1}|^q + |x_{i2} - x_{j2}|^q + \dots + |x_{ip} - x_{jp}|^q\right)}$$

dengan $d(i, j)$ = jarak Minkowski antara data ke- i dan data ke- j ,

x = obyek data,

p = banyaknya atribut data, dan

q adalah bilangan bulat positif.

Jarak Manhattan merupakan kasus khusus ($q=1$) dari Jarak Minkowski.

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|,$$

Jarak Euclid sama seperti Jarak Manhattan, jarak Euclid merupakan kasus khusus dari jarak Minkowski dengan $q=2$

$$d(i, j) = \sqrt{\left(|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2\right)}$$

METODE

Kerangka Pemikiran

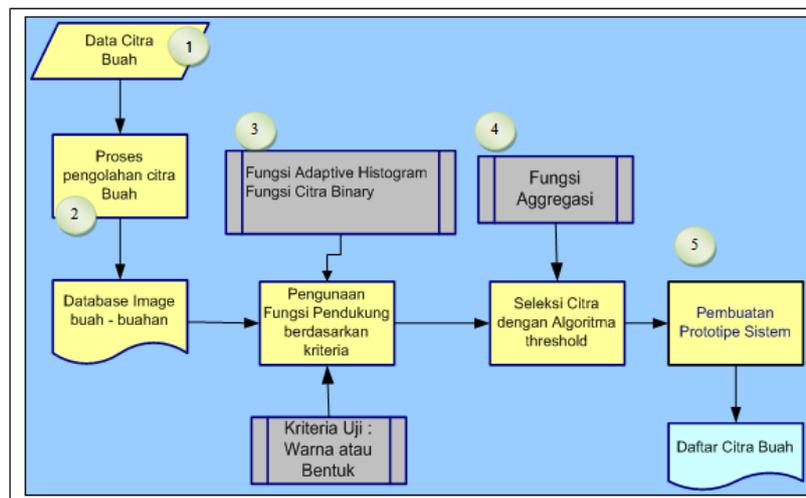
Untuk melakukan penelitian terhadap seleksi citra dalam pencarian citra berbasis ciri di basis data, maka ada lima tahap yang akan dilakukan yaitu: (1) persiapan (pengumpulan dan pengolahan data citra); (2) pemasukan basis data citra; (3) penggunaan fungsi-fungsi pendukung; (4) seleksi citra dengan algoritma *threshold*; (5) pembuatan prototipe sistem.

Dengan seleksi citra dengan menggunakan konsep CBIR, urutan untuk menampilkan kemiripan citra akan dilakukan oleh algoritma *threshold*. Pada dasarnya algoritma ini menggunakan model

logika *fuzzy* dalam menjalankan tahapan prosesnya. Dengan model *fuzzy* ini citra yang dipanggil berdasarkan kemiripan warna dan bentuk terhadap data citra buah di dalam basis data karena untuk kalimat menyerupai atau mirip (*similar*) merupakan kalimat yang bersifat *fuzzy*. Namun jika ingin mengetahui dan mencari citra tertentu dapat menggunakan beberapa fungsi bantu lainnya seperti fungsi citra *binary* dan *adaptive histogram*.

Fungsi algoritma *threshold* yang merupakan turunan dari algoritma Fagin akan menampilkan beberapa citra sekaligus berdasarkan kemiripan dari citra tersebut. Dimulai dari peringkat yang mendekati angka 1 (satu) dengan maksud sangat mirip sekali dan mendekati angka 0 (nol) dengan maksud tidak mirip sama sekali. Baik dari sisi warna dan bentuk. Untuk menampilkan urutan citra menggunakan dua model akses: model akses secara berurut (*sequential*) dan model akses dengan akses acak (*random*). Diasumsikan bahwa akses berurut diwakilkan dengan variable *s* dan akses acak diwakilkan dengan variable *r* sehingga untuk menghasilkan proses *middleware* yang optimal adalah $s + r$.

Untuk itu kerangka pikir dari penelitian ini dapat dilihat pada Gambar 3.

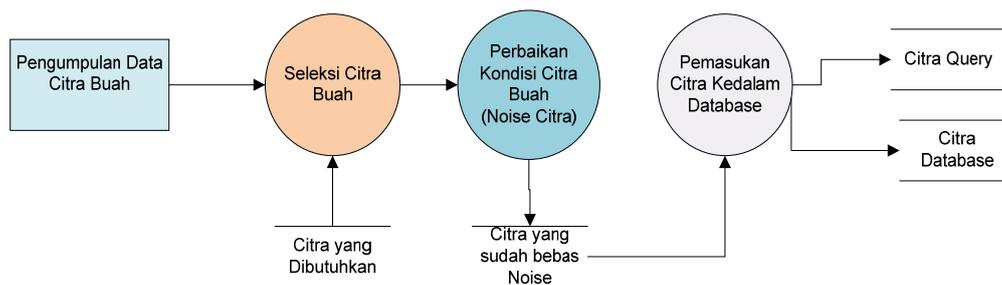


Gambar 3. Kerangka pikir seleksi citra buah dengan konsep CBIR dan algoritma *threshold*.

Tahap Persiapan

Dalam pengumpulan data citra buah, dilakukan melalui beberapa cara seperti: melalui internet, pengambilan dengan foto digital ataupun melalui kumpulan cd citra buah-buahan. Setelah dilakukan pengumpulan data, semua citra harus mempunyai ekstension JPG atau BMP (*.jpg, *.png dan *.bmp). Setelah itu dilakukan seleksi terhadap citra buah yang dapat digunakan sebagai bahan penelitian. Kondisi citra buah yang diharapkan dan sesuai hanya tampak satu-satu bukan kumpulan buah, termasuk juga citra juga jelas bentuk dan warnanya.

Sedangkan secara lebih detail gambaran proses tata laksana untuk menghasilkan citra buah yang siap untuk dimasukkan kedalam basis data citra buah dapat dilihat pada Gambar 4.



Gambar 4. Teknik pengumpulan citra dalam basis data citra.

Pada Gambar 4 terlihat juga bahwa dalam melakukan penelitian ini penulis melakukan beberapa tahapan dalam rangka mengolah citra buah, untuk kemudia dipisah semua data citra buah yang sudah dikumpulkan. Sebahagian citra diletakkan sebagai citra *query* sedangkan yang lainnya diletakkan ke dalam basis data. Pengolahan dimulai dari saat pengumpulan citra buah, seleksi citra hingga perbaikan terhadap citra buah. Ini dilakukan agar data citra yang dimasukkan ke dalam basis data adalah data yang benar dan sesuai dengan kebutuhan prototipe sistem yang akan dikembangkan.

Tahapan Pemasukan Data Citra Ke Basis data

Setelah semua citra diperbaiki dan dibersihkan dari noise citra maka semua citra tersebut dimasukkan ke dalam basis data citra buah. Dengan demikian semua data citra buah yang dimasukkan ke dalam basis data sudah dilakukan perbaikan/modifikasi terhadap kondisi citra, baik dari kejelasan warna, ketepatan bentuk dan penghapusan warna latar. Untuk jumlah data citra buah yang sudah dimasukkan kedalam basis data sebanyak 180 citra buah.

Tahapan Penggunaan Fungsi

Untuk melakukan komparasi antara citra *query* dengan citra di basis data dalam menjalankan konsep CBIR, maka diperlukan beberapa fungsi pendukung. Untuk memanggil kembali citra buah berdasarkan warna, terlebih dahulu harus menggunakan fungsi adaptive histogram, setelah itu dilanjutkan dengan proses euclid yang berguna untuk menentukan jarak warna citra. Kemudian dilanjutkan dengan fungsi sigmoid yang berguna untuk memproses nilai terhadap citra berdasarkan perbandingan dengan citra *query*. Sementara untuk memanggil kembali data citra berdasarkan bentuk, terlebih dahulu harus menggunakan fungsi citra binary, setelah itu dilanjutkan dengan proses euclid yang berguna untuk menentukan jarak bentuk citra. Kemudian dilanjutkan dengan fungsi sigmoid yang berguna untuk memproses nilai terhadap citra berdasarkan perbandingan dengan citra *query*. Dengan nilai sigmoid ini, jika citra bentuk yang ada di dalam basis data hampir sama dengan citra bentuk di *query*.

Tahapan Pemanfaatan Algoritma Threshold

Untuk menghasilkan urutan dari yang mirip sekali hingga yang kurang mirip, penulis menggunakan algoritma threshold yang merupakan turunan dari algoritma Fagin. Nilai threshold (ambang) yang diambil dari citra *query* akan dibandingkan dengan nilai ambang pada semua citra yang ada di basis data citra melalui penggunaan fungsi *euclid*. Dengan model nilai ambang ini, citra yang memiliki nilai ambang mirip akan ditampilkan secara berurutan (*sorting*). Tampilan citra mulai dari yang seperti mirip yang asli hingga yang agak kurang mirip. Proses perbandingan nilai ambang berdasarkan warna dan bentuk akan dilakukan proses *fuzzy logic* dengan hitungan

$$\mu_{A \cap B} = \min(\mu_A[x], \mu_B[y])$$

dimana a = warna b = bentuk.

Pada dasarnya konsep algoritma Fagin dan *threshold* merupakan bagian dari sistem pakar, di mana algoritma tersebut dapat melakukan pengurutan terhadap citra buah yang sudah tersimpan di dalam Basis data. Di dalam penggunaan algoritma ini, berdasarkan perkembangan penelitian ada tiga jenis yang dapat dilakukan yaitu berdasarkan bentuk, warna dan tekstur citra. Dalam penelitian ini, penulis melakukan terhadap pencarian citra buah yang memfokuskan kepada dua jenis, yaitu bentuk dan warna. Dalam melakukan proses pencarian citra buah dengan algoritma ada tiga proses utama, yaitu: (1) proses pengindeksan yang terdiri dari proses segmentasi; (2) proses temu kembali, yaitu “*gradeness*” warna dan “*gradeness*” bentuk, termasuk penentuan kemiripan dengan fungsi agregasi; (3) proses pengurutan yaitu proses algoritma *threshold*. Dalam melakukan penelitian yang berhubungan dengan pencarian citra buah dengan menggunakan konsep logika *fuzzy* berdasarkan Threshold penulis merancang dan menguji dengan perangkat lunak Matlab.

HASIL DAN PEMBAHASAN

Pemanfaatan Parameter Kurva Pusat dan Pemulus untuk Mencari Presisi yang Optimal

Dalam penelitian untuk teknik Pencarian citra ini , dilakukan untuk menentukan nilai presisi yang optimal dari hasil uji coba sebanyak 240 kali menggunakan 2 parameter untuk kurva pusat dan 5 parameter untuk kurva pemulusnya, di mana nilainya untuk kurva pusat adalah -10, dan -8. Sedangkan untuk kurva pemulus mempunyai nilai yaitu 0.2, 0.5, 0.7, 0.8, 0.9. Diharapkan dengan menentukan nilai kurva pusat sebanyak 2 nilai dan nilai pemulus sebanyak 5 nilai akan menghasilkan nilai presisi yang optimal. Dengan perhitungan untuk total percobaan yang dilakukan adalah 2 kurva pusat x 5 kurva pemulus x 16 jumlah jenis citrabuah x 2 jenis algoritma = 320 kali percobaan.

Dengan demikian sebanyak 320 proses yang harus dilakukan oleh penulis dalam mencari citra buah yang ada di dalam basis data sebanyak 180 citra buah.

EFISIENSI SISTEM

Untuk penelitian ini terdapat 180 citra dimana dalam proses CBIR menggunakan 16 citra buah sebagai citra *query*. Biasanya indeks basisdata yang dihasilkan pada proses ekstraksi ciri besarnya meningkat dalam ukuran *kilobyte* dibandingkan dengan basis data yang meningkat dalam ukuran *megabyte*. Waktu rata-rata yang dibutuhkan untuk melakukan CBIR berdasarkan kemiripan ciri citra, dengan kondisi terhadap warna dan bentuk citra adalah sebesar 0,22 detik dan untuk pencarian secara sekuensial dalam *file* indeks adalah sebesar 0,16 detik. Waktu yang dibutuhkan untuk melakukan CBIR dalam pencarian citra ini berbanding lurus dengan peningkatan jumlah citra dalam basis data. Gambaran untuk seleksi efisiensi sistem temu kembali dapat dilihat pada Tabel 1.

Tabel 1
Efisiensi Sistem Temu Kembali Citra Buah Berbasis CBIR

No	Nama File	Ukuran Data		Rasio	Ekstraksi	Proses Temukembali			
		Jumlah citra	Basis data (byte)	Indeks (byte)	Idx/ db (%)	Basis data (detik)	Pencarian (detik)	Visualisasi (detik)	Jumlah (detik)
1	Idxfagin	180	8.560.160	5.448	6,67%	56,273	0,721	0,759	1,480
2	Idxthreshold	180	8.560.160	4.648	4,26%	44,285	0,503	0,516	1,019

Keterangan:

Idx: nama *file* untuk algoritma Fagin

File idx: nama *file* untuk algoritma *threshold*

Kompleksitas Sistem

Kompleksitas sistem ini ditentukan untuk mengetahui kecepatan proses temu kembali citra yang tersimpan di dalam basis data berdasarkan citra contohnya. Untuk itu gambaran untuk kompleksitas waktu pada seleksi citra melalui konsep CBIR dengan algoritma Fagin adalah

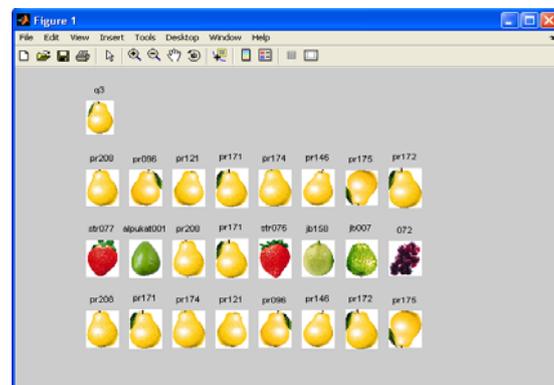
$$T(n) = (n * n + m)$$

atau $T(n) \in O(n^2 + m)$ (n = total proses pengurutan yang harus dilakukan oleh algoritma Fagin yaitu pengurutan secara berurutan dan Acak).

Sehingga Secara keseluruhan, sistem memiliki kompleksitas $O(n^2 + m)$. Sedangkan untuk algoritma *threshold* untuk kompleksitas waktu adalah $T(n) = n + m$ atau proses CBIR dengan menggunakan algoritma *threshold* sebesar $O(n + m)$ (n = total proses pengurutan yang harus dilakukan oleh algoritma *threshold* yaitu pengurutan secara berurutan dan acak dengan nilai *threshold* yang menjadi acuan. Hasil CBIR dalam menseleksi citra buah jambu dan pir menggunakan algoritma *threshold* (sistem prototipe) dapat dilihat pada Gambar 4 dan 5.



Gambar 4. Hasil CBIR dalam seleksi citra buah jambu.



Gambar 5. Hasil CBIR dalam seleksi citra buah pir.

KESIMPULAN DAN SARAN

Kesimpulan

Berikut beberapa kesimpulan yang dapat ditarik dari hasil penelitian ini. Pertama, seleksi citra yang berdasarkan bentuk dan warna dapat diimplementasikan menggunakan konsep CBIR yang memanfaatkan algoritma threshold dalam mengurutkan citra buah. Mengurutkan citra berdasarkan kemiripan citra yang menggunakan algoritma *threshold* dan Fagin dapat dilakukan menggunakan konsep *fuzzy*. Kedua, untuk menghasilkan kinerja sistem temu kembali terhadap citra yang terseleksi menjadi lebih baik, konsep Adaptive histogram dapat digunakan untuk warna citra, dan citra *binary* untuk bentuk citra. Sedangkan mengukur kemiripan citra *query* dan citra basis data dapat dilakukan dengan fungsi *euclid* dan fungsi *sigmoid*. Untuk menghasilkan pengurutan citra yang mirip dapat dilakukan menggunakan konsep peringkat, di mana fungsi yang digunakan adalah fungsi agregasi.

Ketiga, ketika menghitung nilai presisi menggunakan parameter kurva pusat dengan nilai -8 dan -10 dan parameter kurva pemulus dengan nilai 0.2, 0.5, 0.7, 0.8, 0.9, dihasilkan nilai presisi yang optimal sebesar 85% untuk algoritma Fagin dan 90% untuk algoritma *threshold*. Keempat, algoritma Fagin membutuhkan waktu untuk mengeksekusi rata-rata 1,48 detik untuk mengetahui kecepatan proses eksekusi yang dilakukan terhadap 180 citra buah yang ada di dalam Basis data dan 16 citra buah yang ada. Sedangkan algoritma *threshold* membutuhkan waktu eksekusi rata-rata 1.019 detik. Kelima, proses CBIR yang menggunakan algoritma Fagin memiliki kompleksitas waktu sebesar $O(n^2m)$ (n = total proses pengurutan yang harus dilakukan oleh algoritma Fagin yaitu pengurutan secara berurutan dan Acak). Dengan demikian, algoritma memiliki kompleksitas $O(n^2+m)$. Sedangkan algoritma *threshold* untuk kompleksitas waktu adalah $T(n) = n + m$ atau kompleksitas adalah $O(n + m)$ (n = total proses pengurutan yang harus dilakukan oleh algoritma *threshold* yaitu pengurutan secara berurutan dan acak).

Saran

Berikut beberapa saran untuk pengembangan penelitian di masa mendatang: (1) seleksi pencarian citra ini dapat menjadi awal dalam penelitian lanjutan yang berhubungan dengan teknik temu kembali lebih spesifik dan berguna untuk aplikasi lainnya; (2) seleksi terhadap pencarian citra yang terdekat kemiripannya antara citra *query* dan citra basis data, dilakukan dengan menggunakan algoritma *threshold*. Teknik ini dapat dikembangkan kembali dengan berdasarkan algoritma lainnya dalam rangka penyempurnaan; (3) sistem prototipe temu kembali citra dapat diimplementasikan dalam temu kembali informasi untuk bibit buah atau tanaman lainnya. Akurasi hasil temu kembali informasi pertanian dapat ditingkatkan dengan bantuan tenaga ahli pertanian.

DAFTAR PUSTAKA

- Fagin, R. (1998). Fuzzy Queries in Multimedia Database System. *17th ACM Symposium on Principle of Database System, Seattle, 1998, pp 1-10.*
- Xiaoling W, Kanglin, X. (2005). *Application of the Fuzzy Logic in Content-based Image Retrieval JCS & T, 5 (1), 19-24.*
- Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R. Content-based image retrieval at the end of the early years. *IEEE PAMI, 22(12), 1349-1380.*