**USULAN TUGAS AKHIR**

1. **IDENTITAS PENGUSUL**

Nama : Muhammad Arizona Fauzie

NRP : 5108100021

Dosen Wali : Yudhi Purwananto, S.Kom., M.Kom.

1. **JUDUL TUGAS AKHIR**
   1. **“Pemilihan Fitur Citra pada Algoritma Genetika untuk Umpan Balik Sistem Temu Kembali Citra”**
   2. **“Image Feature Selection on Genetic Algorithm for Relevance Feedback on Image Retrieval System”**
2. **LATAR BELAKANG**

Sistem temu kembali citra merupakan penerapan teknik visi komputer untuk mencari citra digital pada *database* citra dalam skala besar. Sistem temu kembali citra berbasis konten berarti pencarian akan dilakukan dengan menganalisa konten sebenarnya dari sebuah citra daripada menggunakan metadata seperti kata kunci, label, maupun diskripsi yang berhubungan dengan citra. Konten dari citra tersebut dapat berupa warna, tekstur, bentuk, ataupun informasi yang lain yang dapat diekstraksi dari sebuah citra. Sistem temu kembali citra memiliki dua tahap penting yaitu ekstraksi dan penyimpanan fitur – fitur citra pada sebuah *database* serta proses temu kembali citra.

Pada proses temu kembali citra masih dimungkinkan terdapat hasil yang berbeda dengan yang dibutuhkan oleh pengguna, untuk itu diperlukan pengembangan untuk meningkatkan relevansi hasil yaitu dengan menggunakan umpan balik. Umpan balik memungkinkan pengguna untuk mengevaluasi hasil temu kembali sehingga perbedaan informasi yang dibutuhkan oleh pengguna dengan hasil temu kembali dapat diminimalisir.

Salah satu metode umpan balik yang ada adalah menggunakan Algoritma Genetika. Algoritma Genetika merupakan algoritma yang efisien untuk menyelidiki daerah pada lingkup pencarian yang belum tereksplorasi dan mendapatkan informasi dari solusi yang memiliki kualitas lebih tinggi [2]. Pada umumnya Algoritma Genetika memiliki populasi dari solusi yang mungkin. Setiap individu pada populasi tersebut

direpresentasikan dengan kromosom. Setiap generasi populasi baru dibuat melalui tiga tahap yaitu evaluasi nilai *fitness* tiap kromosom, pemilihan individu untuk proses generasi berikutnya dan perkawinan individu menggunakan operator genetika yaitu kawin silang dan mutasi untuk memciptakan generasi baru.

Pada sistem temu kembali citra berbasis konten, pemilihan konten citra sebagai representasi kromosom dapat menentukan tingkat relevansi hasil temu kembali. Apabila pada proses temu kembali citra hanya digunakan warna sebagai representasi kromosom maka kemungkinan besar citra yang didapat akan memiliki tingkat relevansi yang rendah dan harus melakukan generasi lebih banyak untuk meningkatkan hasi temu kembali. Sebaliknya apabila menggunakan lebih banyak konten dari citra sebagai representasi kromosom relevansi yang diperoleh bisa lebih tinggi pada tingkat generasi yang lebih rendah.

1. **TUJUAN TUGAS AKHIR**

Tujuan dari Tugas Akhir ini adalah mengimplementasikan Algoritma Genetika Interaktif sebagai umpan balik agar dapat meningkatkan relevansi hasil pada sistem temu kembali citra berbasis konten.

1. **MANFAAT TUGAS AKHIR**

Tugas Akhir ini diharapkan dapat memberikan kemudahan bagi pengguna untuk mencari citra yang relevan dengan yang dibutuhkan pengguna.

1. **RUMUSAN MASALAH**

Rumusan masalah yang diangkat umtuk mengimplementasikan Algoritma Genetika Interaktif pada sistem temu kembali citra berbasis konten ini adalah :

1. Bagaimana menentukan elemen kromosom pada Algoritma Genetika?
2. Bagaimana menentukan operator genetika yang paling sesuai?
3. Bagaimana melakukan analisa fitur algoritma genetika sehingga hasil yang didapat memiliki relevansi yang tinggi?
4. **BATASAN MASALAH**

Permasalahan yang dibahas dalam Tugas Akhir ini memiliki batasan sebagai berikut :

1. Data set yang digunakan berupa citra berwarna yang berasal dari http://wang.ist.psu.edu/docs/related/ yang terdiri dari 10 kategori dan tiap kategori terdiri dari 100 sampel data.
2. Fitur yang diekstraksi pada citra adalah :
   1. Warna direpresentasikan *Hue Saturation Value* (HSV).
   2. Tekstur menggunakan *Gray Level Co-ocurrence Matrix* (GLCM).
   3. Edge menggunakan *Edge Histogram Descriptor* (EHD).
3. Metode seleksi yang digunakan pada Algoritma Genetika adalah :
   1. *Roulette wheel selection.*
   2. *Tournament selection.*
   3. *Rank selection.*
4. Metode persilangan yang akan digunakan pada Algoritma Genetika adalah :
   1. Persilangan satu titik.
   2. Persilangan dua titik.
5. **RINGKASAN TUGAS AKHIR**

Sebuah sistem temu kembali memiliki antar muka agar pengguna dapat berinteraksi dengan sistem. Antar muka tersebut akan menerima *query* berupa citra dan kemudian akan menampilkan hasil pencariannya kepada pengguna. Meskipun begitu, selama proses temu kembali citra masih menggunakan fitur tingkat rendah pasti akan ada hasil citra yang tidak sesuai dengan *query*. Untuk itu disediakanlah fitur umpan balik yang memungkinkan interaksi antara pengguna dan computer sehingga relevansi hasil temu kembali citra dapat ditingkatkan..

Berdasarkan konsep diatas akan didesain sistem temu kembali citra berdasarkan Algoritma Genetika. Sistem ini memiliki empat tahap :

1. *Querying* : Pada tahap ini pengguna akan memasukkan input berupa citra sebagai *query* dari sistem temu kembali citra.
2. Penghitungan similaritas : Pada tahap ini sistem akan menghitung similaritas antara citra pada *query* dengan koleksi citra pada database. Perhitungan similaritas ini didasarkan pada nilai – nilai fitur citra yang telah diekstraksi kedalam database yaitu :
   1. Warna

Warna direpresentasikan dengan *Hue Saturation Value* (HSV). *Hue* digunakan untuk membedakan warna. *Saturation* digunakan untuk mengetahui kadar cahaya putih yang ditambahkan dalam warna asli. *Value* digunakan untuk megetahui kadar cahaya [6]. Untuk mendapatkan properti global citra maka dibuat histogram warna HSV untuk mengetahui distribusi warna dari citra. Dari histogram tersebut diperoleh nilai berupa rataan dan standar deviasi elemen warna HSV dari citra. Sedangkan properti lokal citra diperoleh dengan menggunakan metode *image binary bitmap*, yang memiliki 3 tahap yaitu :

* + - 1. Membagi citra kedalam beberapa blok yang tidak saling tumpang tindih.
      2. Menghitung nilai rataan elemen warna HSV dari setiap blok.
      3. Membandingkan tiap nilai rataan elemen HSV pada blok dengan nilai rataan global HSV pada citra. Jika nilai rataan HSV pada blok lebih besar atau sama dengan rataan nilai HSV global maka nilainya 1 jika tidak 0.
  1. Tekstur

Representasi tekstur dihitung dengan metode *gray level co-ocurrence matrix* (GLCM). GLCM menunjukkan probabilitas kemunculan pasangan piksel yang memiliki tingkat keabuan pada jarak dan arah tertentu. Setelah matriks terbentuk maka dapat diukur informasi tekstural dari suatu citra [3]. GLCM memiliki tahap sebagai berikut :

* + - 1. Inisialisasi matriks yang memiliki ukuran sesuai dengan skala keabuan maksimal.
      2. Tentukan arah dan jarak dua piksel yang akan dicari kemunculan berulangnya.
      3. Hitung kemunculan pasangan piksel dan isi matriks dengan nilai tersebut.
      4. Hitung informasi tekstural citra.

Untuk meneyederhanakan hitungan maka pada Tugas Akhir ini akan digunakan pasangan piksel pada jarak 1 (d = 1), arah 0o, dan informasi tekstural yang akan digunakan adalah *Entropy*.

* 1. Edge

*Edge* *Histogram* digunakan untuk mendeskripsikan distribusi dari *edge* berdasarkan histogram sebuah citra [4]. Proses *Edge* *Histogram Descriptor* (EHD) terdiri dari 5 tahap yaitu :

1. Suatu citra dibagi kedalam 4 x 4 sub citra.
2. Setiap sub citra dipartisi kembali ke dalam blok citra yang lebih kecil.
3. *Edge* dari blok dikategorikan ke dalam 5 kategori edge yaitu : vertikal, horizontal, diagonal 45o, diagonal 135o, dan tak berarah.
4. Dengan demikian histogram dari setiap blok citra menunjukkan frekuensi kemunculan dari lima tipe edge pada sub citra yang dimaksud.
5. Setelah memeriksa seluruh blok citra pada sub citra, lima nilai tersebut kemudian dinormalisasi dengan jumlah blok pada sub citra. Terakhir nilai yang telah dinormalisasi tersebut dikuantisasi untuk representasi biner. Nilai inilah yang merupakan nilai *EHD.*
6. Penerimaan hasil pencarian : Sistem akan menerima dan menampilkan citra sesuai dengan urutan rangking tingkat similaritas. Dengan demikian pengguna akan dimungkinkan untuk menemukan citra yang paling sesuai dengan kebutuhan dengan mengambil citra yang memiliki peringkat tertinggi.
7. Perbaikan hasil pencarian : Setelah memperoleh beberapa image yang relevan, sistem akan menyediakan mekanisme umpan balikdengan menggunakan Algoritma Genetika yang memungkinkan pengguna untuk mengevaluasi citra yang telah diterima tersebut. Evaluasi diberikan dengan memberikan tingkat relevansi dari citra yang telah diperoleh sebagai bobot untuk meningkatkan hasil pencarian selanjutnya dengan Algoritma Genetika. Pada tahap ini pengguna dapat memilih fitur citra yang akan digunakan untuk representasi kromosom serta operator genetika yang akan digunakan untuk proses generasi pada Algoritma Genetika. Proses temu kembali dapat dilanjutkan sampai hasil yang didapatkan memberikan kepuasan pada pengguna.

Untuk mengaplikasikan Algoritma Genetika Interaktif terlebih dahulu harus didefinisikan representasi solusi dari permasalahan tersebut, populasi awal, fungsi fitness, dan operator genetika.

1. Representasi solusi : Solusi direpresentasikan dengan kromosom yang berisi nilai fitur citra yaitu warna, tekstur, dan edge.
   * Representasi kromosom citra.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **µH** | **µS** | **µV** | **σH** | **σS** | **σV** | **BM** | **E** | **EHD** |

Keterangan :

**µH =** nilai rataan *Hue*.

**µS =** nilai rataan *Saturation*.

**µV =** nilai rataan *Value*.

**σH =** nilai standar deviasi *Hue*.

**σS =** nilai standar deviasi *Saturation*.

**σV =** nilai standar deviasi *Value*.

**BM =** nilai binary bitmap dari citra.

**E =** nilai Entropy citra.

**EHD =** nilai *Edge Histogram Descriptor* citra.

1. Populasi awal : Populasi awal berasal dari citra yang diterima pada proses pencarian pertama kali.
2. Fungsi *fitness* : Fungsi *fitness* digunakan untuk mengevaluasi kualitas kromosom pada populasi. Fungsi *fitness* didapatkan dari pengukuran similaritas citra dengan koleksi citra pada database ditambah tingkat relevansi citra yang diberikan oleh pengguna setelah menerima hasil pencarian. Fungsi fitness ditunjukkan pada persamaan berikut.

*F*(*q,C*) = *w*1 *・* sim(*q,C*) + *w*2 *・ δ*

sim(q,C) merupakan nilai similaritas antara citra pada *query* dengan citra paa koleksi database, sedangkan *δ* merupakan tingkat relevansicitra yang diberikan oleh user. Similaritas tersebut diukur sesuai dengan fitur yang dipilih. Apabila fitur citra yang dipilih pada saat proses umpan balik adalah warna, tekstur ( menggunakan entropy ), dan edge maka persamaannya adalah sebagai berikut :

Keterangan :

q = citra *query*.

C = koleksi citra pada database

= nilai rataan

= standar deviasi

= nilai *binary bitmap* citra

= jumlah blok

= nilai *Entropy*

= nilai *Edge Histogram Descriptor* untuk 4 x 4 sub citra dan 5 arah edge.

merupakan *hamming distance* dari citra *query* dan citra pada database. Hamming distance diukur dengan persamaan di bawah ini.

1. Operator genetika : Digunakan untuk mendapatkan kromosom baru dari dua kromosom induk. Pada sistem ini akan digunakan beberapa metode seleksi untuk memilih dua kromosom induk yaitu :
   1. *Roulette wheel selection.*

Induk dipilih berdasarkan nilai *fitness*. Semakin bagus nilai fitness maka semakain besar peluang terpilihnya kromosom. *Roulette wheel selection* memilikitahap sebagai berikut :

* + 1. Hitung total nilai fitness pada populasi kromosom.
    2. Bangkitkan bilangan acak dari interval (0, total *fitness*) .
    3. Lakukan perulangan pada populasi dan hitung nilai fitness. Ketika nilai total fitness tersebut lebih besar dari bilangan acak yang dibangkitkan hentikan perulangan dan pilih kromosom terakhir yang dimasukkan dalam perhitungan sebagai nduk.

* 1. *Tournament selection.*

Metode *tournament selection* akan melakukan turnamen antara individu yang dipilih acak dari populasi. Pemenang dari tiap turnamen dijadikan induk pada persilangan untuk menghasilkan individu baru. *Tournament selection* memiliki tahap sebagai berikut :

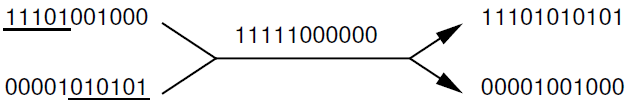
* + 1. Tetapkan ukuran turnamen.
    2. Pilih individu dari populasi secara acak sesuai ukuran turnamen.
    3. Pilih individu terbaik pada turnamen.
    4. Individu yang telah dipilih dapat dihilangkan dari populasi atau dipilih kembali.
    5. Ulangi langkah 1 – 3 sampai diperoleh 2 individu untuk disilangkan.
  1. *Rank selection.*

Metode *rank selection* digunakan untuk mengatasi kromosom yang memiliki nilai *fitness* sangat berbeda dengan kromosom lainnya. Metode seleksi ini akan memberikan nilai *fitness* berdasarkan peringkat. Kromosom terbaik akan memiliki nilai *fitness* 1 sedangkan yang terendah mendapatkan nilai *fitness* sesuai dengan jumlah kromosom dalam populasi. Pemeringkatan kromosom didasarkan pada fungsi fitness awal.

Operator genetikauntuk mendapatkan kromosom baru menggunakan persilangan yaitu :

* 1. Persilangan satu titik

Persilangan satu titik hanya menunjuk satu titik pada kromosom. Semua data di belakang titik tersebut akan ditukar dengan induk yang lainnya sehingga menghasilkan individu baru.

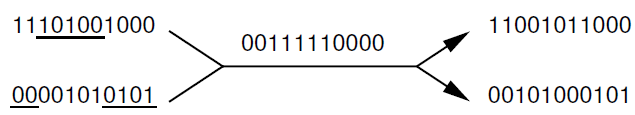


Crossover mask

Kromosom induk Kromosom anak

* 1. Persilangan dua titik

Persilangan dua titik menunjukkan dua titik pada kromosom. Semua data diantara dua titik tersebut akan ditukar dengan induk yang lainnya sehingga menghasilkan individu baru.



Crossover mask

Kromosom induk Kromosom anak

Secara umum sistem temu kembali citra menggunakan Algoritma Genetika Interaktif dapat digambarkan dalam bagan di bawah ini :

 Gamabar 1. *Flowchart* Sistem Temu Kembali Citra Menggunakan Algoritma Genetika Interaktif.

Eevaluasi hasil temu kembali citra dengan pemilihan fitur citra dan operator genetika yang berbeda dilakukan dengan mengukur relevansi berdasarkan *precision* dan *recall*. Evaluasi dilakukan pada tiap generasi dan generasi dilakukan sebanayak 5 kali. *Precision* dan *recall* dihitung dari 50 citra teratas yang diterima. *Precision* diukur dengan membagi jumlah citra hasil temu kembali yang relevan dengan jumlah seluruh citra yang diterima. *Recall* didapat dengan membagi jumlah citra hasil temu kembali yang relevan dengan jumlah citra yang relevan pada database.

*Precision* =

*Recall* =

Evaluasi juga dilakukan pada waktu eksekusi tiap generasi. Hal ini dikarenakan suatu sistem temu kembali harus dapat bekerja secara cepat.

1. **METODOLOGI**

Metodologi yang akan dilakukan dalam Tugas Akhir ini memiliki beberapa tahapan, antara lain sebagai berikut :

1. Penyusunan Proposal Tugas Akhir

Tahap awal untuk memulai pengerjaan Tugas Akhir adalh dengan menyusun Proposal Tugas Akhir. Pada proposal ini penulis mengajukan gagasan pembuatan sistem temu kembali citra berbasis konten dengan Algoritma Genetika Interaktif

1. Studi Literatur

Pada tahap ini dilakukan pencarian, pengumpulan, penyaringan, pembelajaran dan pemahaman literatur yang berhubungan dengan proses temu kembali citra berbasis konten khususnya yang meliputi permasalahan ekstraksi fitur dari suatu citra dan Algoritma Genetika. Literatur yang digunakan dalam pengerjaan Tugas Akhir ini sebagian besar berasal dari internet berupa artikel ilmiah, makalah, materi kuliah serta buku – buku referensi

1. Implementasi

Implementasi merupakan tahap membangun sistem temu kembali citra berbasis konten.

1. Pengujian dan Evaluasi

Pada tahap ini akan dilakukan uji coba terhadap sistem yang telah dibuat meliputi pengamatan kinerja sistem dan identifikasi kendala yang mungkin timbul.

1. Penyusunan Buku Tugas Akhir

Tahap terakhir merupakan penyusunan laporan yang memuat dokumentasi mengenai pembuatan sistem serta hasil dari implementasi sistem yang telah dibuat.

Bagian inti dari buku Tugas Akhhir ini terdiri dari :

1. Pendahuluan
   1. Latar belakang masalah.
   2. Rumusan Permasalahan.
   3. Batasan masalah.
   4. Tujuan.
   5. Manfaat.
2. Tinjauan Pustaka
   1. Hasil penelitian terdahulu yang relevan.
   2. Teori yang digunakan untuk menyelesaikan permasalahan.
3. Metodologi
   1. Metode yang digunakan.
   2. Bahan dan peralatan yang digunakan.
   3. Urutan pelaksanaan percobaan.
4. Hasil dan pembahasan
   1. Data hasil percobaan atau pengukuran.
   2. Pembahasan (analisa, sintesis dan evaluasi).
5. Kesimpulan dan Saran
6. **JADWAL PEMBUATAN TUGAS AKHIR**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No** | **Tahapan** | **Bulan** | | | | | | | | | | | | | | | | | | | |
| **Maret** | | | | **April** | | | | **Mei** | | | | **Juni** | | | | **Juli** | | | |
| **1** | Pembuatan Proposal |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **2** | Studi Literatur |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **3** | Implementasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **4** | Pengujian dan Evaluasi |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| **5** | Penyusunan Buku Tugas Akhir |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. **DAFTAR PUSTAKA**
   1. Lai, Chih-Chin dan Ying Chuan Chen, “*A User-Oriented Image Retrieval System Based on Interactive Genetic Alghorithm*,”*IEEE Transaction on Instrumentation and Measurement*, vol. 60,no. 10, pp. 3318–3325, Oct. 2011.
   2. S. F. da Silva, M. A. Batista, and C. A. Z. Barcelos, “*Adaptive image retrieval through the use of a genetic algorithm*,” in *Proc. 19th IEEE Int.Conf. Tools With Artif. Intell.*, 2007, pp. 557–564.
   3. Myriax Software , “About the GLCM and texture”. http://support.echoview.com /WebHelp/Reference/Algorithms/Operators/GLCM\_textures/About\_the\_GLCM\_and\_textures.htm (diakses tanggal 3 Maret 2012).
   4. Park, Dong Kwon, Yoon Seok Jeon, dan Chee Sun Won,”*Efficient Use of Local Edge Histogram Descriptor”*, *Dep. of Electronic Engineering,Dongguk Universiyt*
   5. Wikipedia, “Crossover (Genetic Algorithm)”. http://en.wikipedia.org/wiki/Crossover\_(genetic\_algorithm)*.* (diakses tanggal 8 Maret 2012).
   6. Wikipedia, “HSV and HSL”. http://en.wikipedia.org/wiki/HSL\_and\_HSV (diakses tanggal 8 Maret 2012).

**LEMBAR PENGESAHAN**

Surabaya, 15 Maret 2012

Menyetujui,

|  |  |
| --- | --- |
| Dosen Pembimbing I,  **( Diana Purwitasari, S,Kom, M.Sc )**  **(NIP. 197804102003122001)** | Dosen Pembimbing II,  **(Anny Yuniarti, S.Kom.,M.Com.Sc.)**  **(NIP. 198106222005012002)** |