

# 1. Pendahuluan

## 1.1 Latar belakang masalah

Content Based Image Retrieval(CBIR) adalah teknik pencarian citra dengan masukan berupa citra. Dengan CBIR hasil pencarian akan lebih relevan karena yang dicari bukan informasi tekstual (tag atau metadata) yang ada pada citra melainkan isi dari citra itu sendiri.

Inti permasalahan dari CBIR adalah membuat model matematis dari citra dan mencari kesamaan dari model yang telah dibuat[6]. Pembuatan model matematis, atau signature, diperlukan karena representasi dari citra yang berupa matrix berisi nilai pixel ternyata tidak mencukupi. Pemodelan ini umumnya dilakukan dengan cara melakukan ekstraksi fitur baik itu fitur warna, texture atau bentuk. Untuk mencari kesamaan model perlu dicari “selisih” antara fitur yang dicari dan didapat. Beberapa metode yang digunakan pada paper yang membahas CBIR adalah euclidian distance[11], hamming distance[9], affine normalisation[8].

Pada tugas akhir ini CBIR yang akan digunakan adalah berbasis hash. Fungsi hash dipilih karena dengan mengubah signature ke nilai hash maka mencari selisihnya akan lebih cepat[8] karena hanya dicari bit-bit yang berbeda (*hamming distance*). Namun kelemahan fungsi hash adalah ketika citra masukan mengalami sedikit perubahan maka nilai hashnya akan berbeda sehingga berdampak nilai hamming distance tinggi dan nilai yang tinggi tersebut mempengaruhi prosentase kesamaan pada citra yang dicari.

Untuk mengatasi hal tersebut dua solusi ditawarkan. Pertama melakukan hashing pada fitur-fitur yang ada pada citra bukan keseluruhan citra. Dengan begitu meskipun secara keseluruhan citra banyak mengalami perubahan tetap bisa dicari kesamaannya. Metoda feature extraction yang akan digunakan adalah Maximally stable extremal regions (MSER). MSER dipilih karena berdasarkan pengujian yang dilakukan oleh Mikolajczyk [4] metoda ini mempunyai waktu komputasi paling kecil dan matching score paling tinggi dibandingkan dengan metoda lain. Kedua, menggunakan fungsi hash yang tidak terlalu sensitif terhadap perubahan kecil (minor).Hal ini dilakukan untuk jaga-jaga apabila fitur yang diekstrak bentuknya mengalami sedikit perubahan. Untuk saat ini fungsi hash yang bisa melakukan hal tersebut adalah Perceptual Hash (pHash)[9].

## 1.2 Perumusan masalah

Permasalahan pada tugas akhir ini adalah:

1. Bagaimana mengimplementasikan MSER dan perceptual hash pada CBIR?
2. Apakah penggabungan MSER dan perceptual hash dapat meningkatkan akurasi?

## 1.3 Batasan Masalah

Dalam implementasi tugas akhir ini dibatasi oleh beberapa hal, sebagai berikut:

1. Data masukan diambil dari <http://www.robots.ox.ac.uk/~vgg/research/affine/> dan dari hasil pencarian dengan CBIR di Internet dengan alamat [www.tineye.com](http://www.tineye.com)
2. Implementasi MSER menggunakan *library* dari openCV.
3. Pengujian hanya melibatkan sistem yang menggunakan pHash dan pHash+MSER.

## 1.4 Tujuan

Tujuan yang akan dicapai dalam penelitian ini adalah:

1. Membangun sistem *ContentBased Image Rretrieval* dengan menggabungkan metoda MSER dan perceptual hash agar hasil pencarian bisa lebih akurat.
2. Menguji dan menganalisis kinerja sistem tersebut

## Hipotesa:

Hasil pengujian yang diperoleh Mikolajczyk [4] adalah MSER mampu mendeteksi fitur-fitur pada citra meskipun citra tersebut diambil dari sudut pandang yang berbeda, waktu pengambilan yang berbeda, dan pada citra yang telah mengalami blurring. Artinya metoda ini mampu mengekstrak fitur meskipun citra tersebut telah mengalami berbagai macam perubahan. Dari segi performansi hasil pengujian pada [5] adalah 0.66 detik, rata-rata jumlah region yang bisa diekstrak adalah 533 dan prosentase correct match tertinggi, 70-80 %. Pada paper [8] MSER diujikan sebagai CBIR dan hasilnya average precision sebesar 93,5% dan average recall sebesar 83,3%

pHash mampu mendeteksi citra sejenis yang pencahayaannya berbeda karena ketika membuat nilai hash terdapat proses pengurangan warna. Selain itu berdasarkan hasil pengujian pHash mampu membuat hash 140 citra yang beresolusi diatas 2000x2000 selama 58 s

Dengan penggabungan akurasi dari MSER dan kecepatan pHash, sistem ini mampu membuat hasil pencarian yang lebih akurat.

## 1.5 Metodologi penyelesaian masalah

1. Studi literatur.  
Pencarian materi-materi dan referensi yang berkaitan dengan permasalahan yang dibahas, seperti materi tentang MSER, perceptual hash, dan materi pendukung lainnya.
2. Analisis dan perancangan kebutuhan sistem.  
Memaparkan paramter-parameter dan dataset yang akan digunakan ketika pengujian serta perancangan perangkat lunak untuk CBIR
3. Implementasi sistem.  
Melakukan *coding* menggunakan *python* untuk membangun sistem sesuai dengan rancangan pada tahap sebelumnya.
4. Pengujian sistem.  
Pada tahap ini, dilakukan pengujian terhadap sistem yang telah dibangun. Hal yang diujikan ialah seperti yang telah dipaparkan pada tahap perancangan.
5. Analisis hasil pengujian.  
Dari tahap pengujian sistem yang dilakukan sebelumnya, dilakukan analisis terhadap pengaruh MSER terhadap kinerja.
6. Penyusunan laporan Tugas Akhir.  
Pada tahap ini, dilakukan penyusunan laporan akhir dan pengumpulan dokumentasi yang diperlukan, format laporan mengikuti kaidah penulisan yang benar dan yang sesuai dengan ketentuan-ketentuan yang telah ditetapkan oleh institusi.

## 1.6 Sistematika Penulisan

Tugas akhir ini disusun dengan sistematika penulisan sebagai berikut:

### 1 Pendahuluan

Bab ini menguraikan tugas akhir ini secara umum, meliputi latar belakang masalah, perumusan masalah, batasan masalah, tujuan, hipotesa awal, dan metode penyelesaian masalah.

### 2 Landasan Teori

Bab ini membahas mengenai uraian teori yang berhubungan dengan *Content Based Image Retrieval*, MSER, dan perceptual hash

### 3 Analisis Perancangan dan Implementasi

Bab ini berisi analisis kebutuhan dari system dan masalah-masalah yang ada di dalamnya. Hasil analisis ini dituangkan ke dalam suatu system pemodelan secara terstruktur. Dari tahap analisis kemudian dilanjutkan ke tahap perancangan dan implementasi.

#### **4 Pengujian dan Analisis Hasil Percobaan**

Bab ini membahas mengenai pengujian hasil implementasi yang telah dilakukan pada bab sebelumnya. Pengujian dilakukan dengan membandingkan hasil implementasi dengan data aslinya. Tahap pengujian dilanjutkan dengan tahap analisis hasil pengujian.

#### **5 Kesimpulan dan Saran**

Berisi kesimpulan dari penulisan Tugas Akhir ini dan saran-saran yang diperlukan untuk pengembangan lebih lanjut.