

Content Based Image Retrieval

STBI Multimedia

- Tujuan: untuk memberikan jawaban terbaik yang sesuai dengan kebutuhan pengguna
- STBI multimedia : sistem untuk manajemen (penyimpanan, pengambilan, dan manipulasi) data beberapa media seperti kombinasi data tabulasi/administrasi, dokumen teks, gambar, spasial, sejarah, audio, dan data video (joan, 2008).

Komponen STBI Multimedia

- Dokumen multimedia: sistem mampu menyimpan dokumen multimedia.
 - Dokumen berupa data multimedia (teks, gambar, audio, video, dll)
 - Disimpan semi terstruktur: data terstruktur dan tidak terstruktur
- Model temu kembali
 - Menyajikan daftar dokumen yang diurutkan dalam peringkat

Komponen STBI Multimedia

- Analisis dokumen: dokumen diproses untuk mengumpulkan informasi statistik
- Teknik pencarian interaktif: sistem mendukung interaksi dengan pengguna untuk meningkatkan kemungkinan keberhasilan pencarian
 - Pengguna dapat memberikan informasi umpan balik kepada sistem

Jenis-jenis Sistem Temu Kembali

- Temu Kembali Audio berbasis Konten (Content Based Audio Retrieval)
- Temu Kembali video berbasis Konten (Content Based Video Retrieval)
- Temu Kembali Citra Berbasis Konten (Content Based Image Retrieval)
- Temu Kembali Teks Berbasis Konten (Content Based Text Retrieval)

Sistem temu kembali citra berbasis konten

- suatu teknik pencarian kembali gambar yang mempunyai kemiripan karakteristik atau content dari sekumpulan gambar.
- Proses umum dari CBIR adalah gambar yang menjadi query dan penyimpanan dalam database dilakukan proses ekstraksi fitur
- Fitur gambar yang dapat digunakan untuk retrieval pada system ini misalnya histogram, susunan warna, tekstur, dan shape, tipe spesifik dari obyek, tipe event tertentu, lokasi, emosi.

Pengukuran Jarak Antar Dua Histogram

- Fitur warna merupakan fitur yang paling banyak digunakan pada sistem CBIR.
- Banyak diantaranya menggunakan image color histogram.
- Color histogram antara dua gambar tadi kemudian dihitung jaraknya.
- Gambar yang memiliki jarak paling kecil, merupakan solusinya.

Pengukuran Jarak Antar Dua Histogram

- dimisalkan ada dua gambar dengan histogram 4 warna yang sudah terkuantisasi sebagai berikut:

$$H^A = \{20\%, 30\%, 10\%, 40\%\}$$

$$H^B = \{10\%, 10\%, 50\%, 30\%\}$$

- Rumus menghitung jarak:

$$d(A, B) = \sum_{j=1}^n |H_j^A - H_j^B|$$

- Sehingga:

$$d(A, B) = |0.2 - 0.1| + |0.3 - 0.1| + |0.1 - 0.5| + |0.4 - 0.3| = 0.8$$

Pengukuran Jarak Antar Dua Histogram

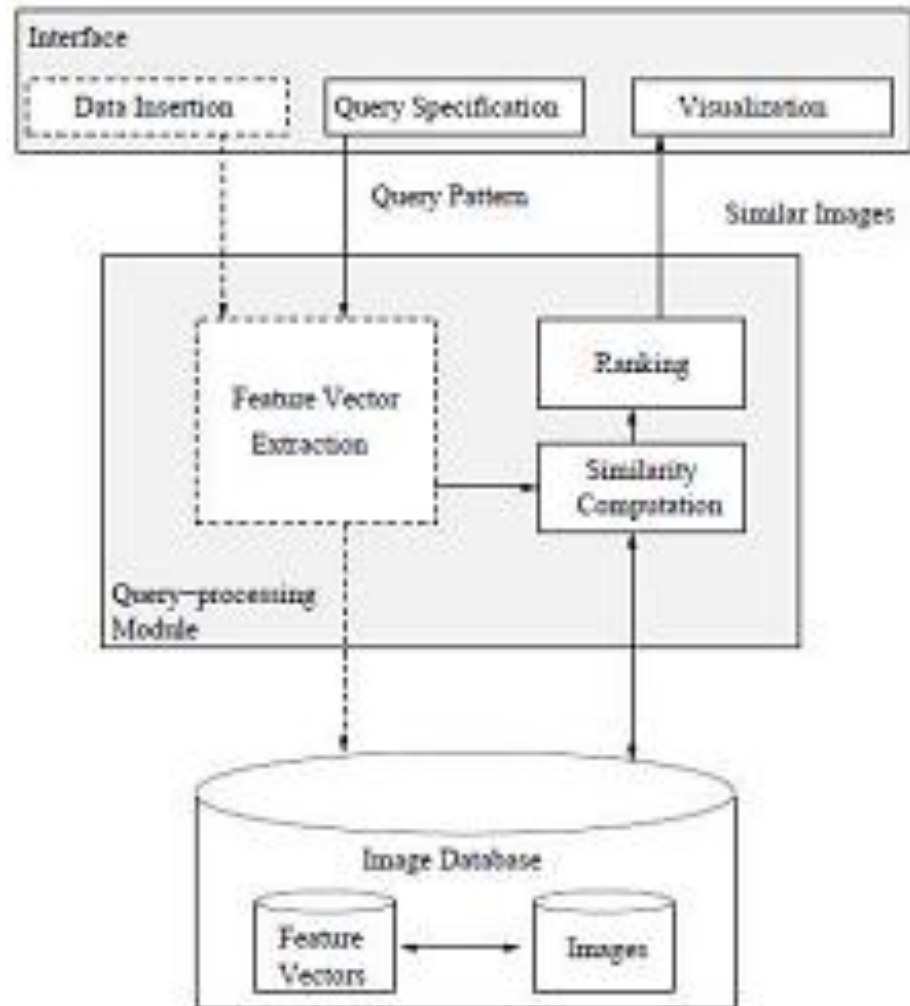
- Cara lain dg rumus Euclidan:

$$d(A, B) = \sqrt{\sum_{j=1}^n (H_j^A - H_j^B)^2}$$

- Sehingga:

$$d(A, B) = \sqrt{\sum_{j=1}^n (0.2+0.1)^2 + (0.3+0.1)^2 + (0.1-0.5)^2 + (0.4-0.3)^2} = 0.47$$

Arsitektur CBIR



Bagian Arsitektur CBIR

1. Interface : bagian untuk interaksi antara pengguna dengan sistem CBIR melalui aplikasi GUI (Graphical User Interface). Terdiri dari:
 - Data Insertion : untuk memasukkan citra yang akan di ekstraksi.
 - Query Specification untuk menentukan citra yang akan dijadikan citra query serta menentukan metode ekstraksi fitur.
 - Visualization untuk menampilkan citra query dan citra hasil pencarian.

Bagian Arsitektur CBIR

2. Query-processing module, terdiri dari:
 - Feature vector extraction untuk mengekstraksi baik citra yang ada didatabase citra maupun citra query.
 - Similarity Computation digunakan untuk menghitung kesamaan fitur citra.
 - Ranking digunakan untuk mengurutkan citra yang memiliki tingkat kemiripan dengan citra query.

Bagian Arsitektur CBIR

3. Image Database, terdiri dari

- Feature Vectors yang digunakan untuk menyimpan hasil ekstraksi fitur citra. Dan disimpan didalam database derby.
- Images merupakan database citra yang secara fisik berupa folder yang didalamnya terdapat kumpulan citra