

ABSTRAK

Perkembangan teknologi yang semakin canggih memicu makin berkembangnya data di berbagai bidang seperti hiburan, perdagangan, pendidikan, *biomedicine*, dan lain-lain. Data citra dari berbagai bidang tersebut semakin bertambah dengan cepat pula. Data citra membutuhkan ruang simpanan yang besar, dan kemungkinan dalam suatu direktori penyimpanan pada *harddisk* terdapat banyak citra. Sehingga sistem penyimpanan dari berbagai macam informasi digital tersebut semakin meningkat dan membuat masalah dalam pencarian kembali dan pengolahannya.

Teknik pencarian yang sudah ada saat ini merupakan pencarian dengan *inputan* berupa teks namun menghadapi berbagai masalah diantaranya kurang praktis dan hasilnya terkadang tidak sesuai dengan yang dicari. Maka dibangunlah *Content Based Image Retrieval (CBIR)*. Metode *color histogram* digunakan untuk mengekstraksi fitur warna dan juga sebagai nilai pembanding untuk menemukan kembali citra. Citra *query* maupun citra acuan diekstrak *color histogram*nya kemudian dihitung selisihnya. Selain itu digunakan pula teknik *clustering* menggunakan algoritma *K-Means*. Penggunaan centroid hasil pengelompokan dataset yang berasal dari *color histogram* dari kumpulan citra dapat digunakan sebagai acuan untuk melakukan pencarian. Sistem ini dibangun dengan bahasa pemrograman Java. Metode pengembangan perangkat lunak menggunakan metode RUP (*Rational Unified Process*). Hasil percobaan dari pencarian dengan *color histogram* dan teknik *clustering* ini ternyata berhasil menemukan citra yang mirip dari segi warna dengan akurasi cukup tinggi dan waktu pencarian yang cepat.

Kata Kunci : *Content Based image retrieval (CBIR), color histogram, K-Means clustering*

ABSTRACT

The development of increasingly sophisticated technology triggers the growing data in various fields such as entertainment, commerce, education, biomedicine, and others. Image data from a variety of fields growing quickly as well, where it is based on human needs tend to be constantly increasing. Image data requires a large storage space, and possibly a storage directory on the hard drive there is a lot of imagery. So the storage systems of various kinds of digital information is increasing and causing trouble in the retrieval and processing.

Existing search techniques now is a search with text input, but face many problems such as lack of practical and sometimes the results are not in accordance with the search. So a Content Based Image Retrieval (CBIR) was built. Color histogram method is used to extract color features as well as a comparison value to reinvent the image. Color histogram of query image and the reference image are extracted and then calculated the difference. In addition it is also used clustering techniques using K-Means algorithm. Using the results of clustering centroid dataset derived from the color histogram of the image collection can be used as reference to search. The system is built with the Java programming language. Software development method using the RUP (Rational Unified Process). The experimental results of the search by color histogram and clustering techniques proved able to find images that are similar in terms of color with a high accuracy and a fast search time.

Keywords : *Content Based image retrieval (CBIR), color histogram, K-Means clustering*