Summary: In this paper, we address a ranking problem in web image retrieval. Due to the growing availability of web images, comprehensive retrieval of web images has been expected. Conventional systems for web image retrieval are based on keyword-based retrieval. However, we often find undesirable retrieval results from the keyword-based web image retrieval system since the system uses the limited and inaccurate text information of web images; a typical system uses text information such as surrounding texts and/or image filenames, etc. To alleviate this situation, we propose a new ranking approach which is the integration of results of text and image content via analyzing the retrieved results. We define four ranking methods based on the image contents analysis of the retrieved images; (1) majority-first method, (2) centroid-of-all method, (3) centroid-of-top $K$ method, and (4) centroid-of-largest-cluster method. We evaluate the retrieval performance of our methods and conventional one using precision and recall graphs. The experimental results show that the proposed methods are more effective than conventional keyword-based retrieval methods.

For the entire collection see [Zbl 1026.68831].

MSC:

- 68U99 Computing methodologies and applications
- 68P20 Information storage and retrieval of data
- 68T45 Machine vision and scene understanding
- 68U35 Computing methodologies for information systems (hypertext navigation, interfaces, decision support, etc.)

Software:

- SIMPLIcity
- NeTra
- WebSeer

Full Text: Link