Client-Side Relevance Feedback Approach for Image Retrieval in Mobile Environment

Ning Yu, University of Central Florida, USA
Kien A. Hua, University of Central Florida, USA
Danzhou Liu, Symantec Corporation, USA

ABSTRACT

During the last decade, high quality (i.e., over 1 megapixel) built-in cameras have become standard features of handheld devices. Users can take high-resolution pictures and share with friends via the internet. At the same time, the demand of multimedia information retrieval using those pictures on mobile devices has become an urgent problem to solve, and therefore attracts attention. A relevance feedback information retrieval process includes several rounds of query refinement, which incurs exchange of images between the mobile device and the server. With limited wireless bandwidth, this process can incur substantial delay, making the system unfriendly to use. This issue is addressed by considering a Client-side Relevance Feedback (CRF) technique. In the CRF system, Relevance Feedback (RF) is done on client side along. Mobile devices’ battery power is saved from exchanging images between server and client and system response is instantaneous, which significantly enhances system usability. Furthermore, because the server is not involved in RF processing, it is able to support more users simultaneously. The experiment indicates that the system outperforms the traditional server-client relevance feedback systems on the aspects of system response time, mobile battery power saving, and retrieval result.

Keywords: Client-Side Relevance Feedback, Image Retrieval, Mobile Information Retrieval, Mobile Platform, Query by Photo

1. INTRODUCTION

Mobile devices with high resolution built-in camera are becoming ubiquitous. It is desirable to support multimedia information retrieval using the images taken from the camera. As examples, a student can use pictures of a plant to search for information of similar species in a remote digital library. Providing this capability calls for efficient techniques to facilitate Content-Based Image Retrieval (CBIR) in mobile environment. In a CBIR system, images are characterized by their low level features such as color, texture, and shape. Since it is difficult to “describe” a query through those features, RF is widely used and plays an important role (Binderberger & Mehrotra, 2004; Chen et al., 2009; Fu et al., 2008; Goh et al., 2004; Gevers, 2008).
& Smeulders, 2004; Hoi et al., 2006, Kim & Chung, 2003). It helps the system to understand the user’s intention. In such a system, the user interacts with the system as follows: In each round, he helps by identifying the relevant images within the returned set; the system then utilizes this feedback to modify the query and thus to improve its retrieval results in the next round. This process can be repeated until the user is satisfied with the results.

Lots of information retrieval systems in mobile environment have been introduced in recent years (Bezerra et al., 2005; Jia et al., 2006; Kim et al., 2006; Lee & Jayant, 2006; Shyu et al., 2006; Sonobe et al., 2004; Sarvas et al., 2004; Tollmar et al., 2004). In Bezerra et al. (2005), the query is composed by keywords and visual sub-queries. The keyword part is submitted to an existing search engine and the resulting images are passed to a query refinement agent that processes the visual part of the query in the client side. In Jia et al. (2006) and Kim et al. (2006), visual features are used to search in a large database and the result is returned to the user for relevance feedback. In summary, in those systems, the user needs to interact with the server for result. The mobile device works as an interface. However, this computation model is not suitable for RF because the wireless bandwidth is limited for exchanging images in each round of user RF. The long delay in feedback would affect the usability of the mobile system. Moreover, communication is generally hundreds of times more demanding on mobile power than computation is (Pottie & Kaiser, 2000; Stemm & Katz, 1997). Constantly sending and receiving queries and results can quickly exhaust mobile device’s battery power. In practice, power required by CPU is minimal compared to sending data over the wireless radio. For example, the energy cost of transmitting 1Kb message over a distance of 100 meters is approximately 3 joules. By contrast, a general-purpose processor with 100 MIPS/W power could efficiently execute 3 million instructions for the same amount of energy (Pottie & Kaiser, 2000). Also the storage power for mobile devices is always some hundred mW (Zheng et al., 2003), which is 20-30% of the communication power consumption. Therefore, saving the communication power will be more urgent in mobile RF system.

Another issue in the existing image retrieval techniques is the semantic gap between the high level semantic concept and the low level visual features. In most of the current systems (Ishikawa et al., 1998; Kim & Chung, 2003), the searching range is always confined in a single neighborhood. This is not consistent with the reality: a semantic concept can have very different representation (for example, side view of a sedan and front view of a sedan can be very different). Trying to bridge the gap, some mobile information management systems use tags generated at the point of picture was taken: In Davis et al. (2004), Liu et al. (2005), and Lahti et al. (2005), location, temporal, and sometimes social contextual metadata are used as image features instead of the visual features. However, a picture is worth a thousand words. It is hard and tedious to fully describe an image using concise verbal terms. The most convenient and common way for a mobile user to express himself is to use the picture he took as the query (Bentley et al., 2006).

To address the aforementioned issues, we investigate a Client-side Relevance Feedback (CRF) technique in this paper. In this model, all RF except the last round is processed in the mobile device to avoid the communication cost. To achieve this, we adapt the Relevance Feedback Support (RFS) structure (Hua et al., 2006; Yu et al., 2007) to address the storage and computation limitation of mobile devices. In the CRF system, the initial user query might be decomposed to cover more semantically relevant images, thus to bridge the semantic gap. We list the contributions as below:

1. The CRF technique is detailed described and investigated.
2. We take the interface design into consideration. We develop different interfaces for different mobile users.
Related Content

Improving Gender Classification Using an Extended Set of Local Binary Patterns
Abbas Roayaei Ardakany, Mircea Nicolescu and Monica Nicolescu (2014). International Journal of Multimedia Data Engineering and Management (pp. 47-66).
www.igi-global.com/article/improving-gender-classification-using-an-extended-set-of-local-binary-patterns/117893?camid=4v1a

Learning Object Model for Online Laboratories
www.igi-global.com/chapter/learning-object-model-online-laboratories/49402?camid=4v1a
FaceTimeMap: Multi-Level Bitmap Index for Temporal Querying of Faces in Videos
www.igi-global.com/article/facetimemap/233863?camid=4v1a

Situated Multimedia for Mobile Communications
www.igi-global.com/chapter/situated-multimedia-mobile-communications/27079?camid=4v1a