Chapter X
Content Analysis from User's Relevance Feedback for Content-Based Image Retrieval

Chia-Hung Wei
Ching Yun University, Taiwan

Chang-Tsun Li
University of Warwick, UK

ABSTRACT

An image is a symbolic representation; people interpret an image and associate semantics with it based on their subjective perceptions, which involves the user's knowledge, cultural background, personal feelings and so on. Content-based image retrieval (CBIR) systems must be able to interact with users and discover the current user's information needs. An interactive search paradigm that has been developed for image retrieval is machine learning with a user-in-the-loop, guided by relevance feedback, which refers to the notion of relevance of the individual image based on the current user's subjective judgment. Relevance feedback serves as an information carrier to convey the user's information needs / preferences to the retrieval system. This chapter not only provides the fundamentals of CBIR systems and relevance feedback for understanding and incorporating relevance feedback into CBIR systems, but also discusses several approaches to analyzing and learning relevance feedback.

1. INTRODUCTION

The rapid growth in the amount of digital images has highlighted the importance of effective retrieval approaches in order to facilitate the searching and browsing of large image databases. Although the design of content-based image retrieval (CBIR) systems is based on the nature of the underlying images

and the system's purposes, one of the common purposes of all image retrieval systems is to satisfy human information needs and support human activities in an efficient and effective way. The development of an image retrieval system has to take human factors into account. Among human factors, subjective perception is one of the most challenging issues. An image is a symbolic representation; people interpret an image and associate semantics with it based on their subjective perceptions, which involves the user's knowledge, cultural background, personal feelings and so on (Jaimes, 2006b).

An important assumption in image retrieval is that each user's information need is different and time varying (Zhou & Huang, 2003). This assumption indicates that humans exhibiting subjective perceptions when interpreting images can be classified as *different information seekers* or *same information seekers* (Jaimes, 2006a). Different information seekers normally have different interpretations for the same image based on their individual perceptions. As a result, when different information seekers provide the same query example, they will have different satisfactory degrees for the same search results; even the same information seekers have different subjective perceptions as time evolves.

Another challenging issue arises from the difference between two descriptions of an object by high-level semantics and representations of low-level pixel data (Liu, Zhang, Lu, & Ma, 2007; Vasconcelos, 2007). The difference, known as the semantic gap, exists because low-level features are more easily computed in the system design process, but high-level queries are used as the starting point of the retrieval process. The semantic gap involves not only the conversion between low-level features and high-level semantics, but also the understanding of contextual meaning of the query involving human knowledge and emotion. Figure 1 shows that visual similarity mismatches human similarity judgments, resulting in a semantic gap between the user and the CBIR system. The "riding bicycle" query contains color gradients and two circular shapes that might guide a CBIR system, which utilizes shape and color as features for discriminating images, to associate it with objects, such as earphones, glasses, binoculars and two coins, with similar low-level features. However, the user actually looks for "riding bicycle" images, correlating the query example with high-level semantic context.

To compensate for the vagueness of subjective human perceptions, image retrieval systems must be able to interact with users and discover the current user's information needs. Due to the vagueness of subjective human perceptions, a retrieval system cannot adopt a fixed clustering structure to deal with the problem. An interactive search paradigm that has been developed for image retrieval is machine learning with a user-in-the-loop, i.e. in the search process users will be required to interact with the

Figure 1. Visual similarity does not match with human similarity judgments, resulting in a semantic gap. (a) A query example. (b) Visual feature matches. (c) Semantic matches.
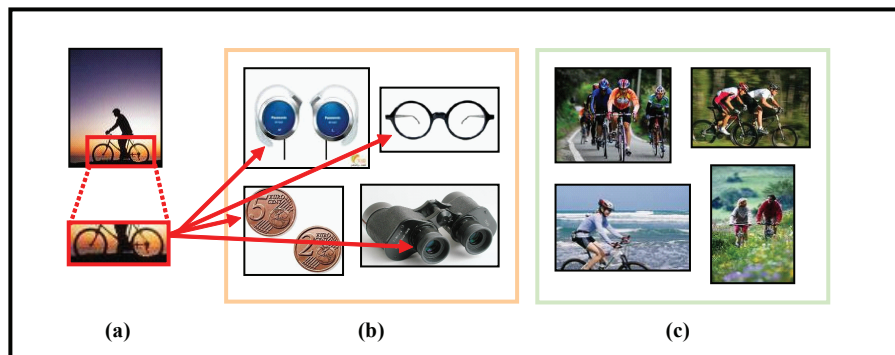
Related Content

Wireless Body Area Networks Combined with Mobile Cloud Computing in Healthcare: A Survey
Srijan Goswami, Payel Roy, Nilanjan Deyand Sayan Chakraborty (2016). *Classification and Clustering in Biomedical Signal Processing (pp. 388-402).*
www.irma-international.org/chapter/wireless-body-area-networks-combined-with-mobile-cloud-computing-in-healthcare/149400

Human Skin Detection in Color Images Using Deep Learning
Mohammadreza Hajiarbabiand Arvin Agah (2015). *International Journal of Computer Vision and Image Processing (pp. 1-13).*
www.irma-international.org/article/human-skin-detection-in-color-images-using-deep-learning/159804

A Hybrid Moth-Flame Optimization Technique for Feature Selection in Brain Image Classification and Image Denoising by Improved Log Gabor Filter
P. M. Diazand M. Julie Emerald Jiju (2022). *International Journal of Computer Vision and Image Processing (pp. 1-20).*
www.irma-international.org/article/a-hybrid-moth-flame-optimization-technique-for-feature-selection-in-brain-image-classification-and-image-denoising-by-improved-log-gabor-filter/296585

Fingerprint Matching Using Rotational Invariant Orientation Local Binary Pattern Descriptor and Machine Learning Techniques
Ravinder Kumar (2017). *International Journal of Computer Vision and Image Processing (pp. 51-67).*
www.irma-international.org/article/fingerprint-matching-using-rotational-invariant-orientation-local-binary-pattern-descriptor-and-machine-learning-techniques/195010

Cross-Domain Usage in Real-Time Video-Based Tracking
Satbir Singh, Rajiv Kapoorand Arun Khosla (2018). *Handbook of Research on Advanced Concepts in Real-Time Image and Video Processing (pp. 105-129).*
www.irma-international.org/chapter/cross-domain-usage-in-real-time-video-based-tracking/186275