

## A HYBRID RANKING OF LINK AND POPULARITY FOR NOVEL SEARCH ENGINE

HIROMITSU HAMA, THI THI ZIN AND PYKE TIN

Graduate School of Engineering  
Osaka City University  
Sugimoto 3-3-138, Sumiyoshi-ku, Osaka 558-8585, Japan  
hama@info.eng.osaka-cu.ac.jp; { thithi; pyketin }@sys.info.eng.osaka-cu.ac.jp

Received September 2008; revised February 2009

**ABSTRACT.** *In this paper, we explore a new paradigm to enable Web search at image level reflecting the most relevant results to the users. So we introduce a new concept of Popularity-based Rank (PR), a content-level ranking and searching model for image retrieval. Specifically, we establish a PR Operation which is combined with new link structure analysis. The strategy to decide searching order by taking similarity into consideration is also proposed and proved to be effective and efficient. Experimental results show that the combined analysis can achieve significantly better ranking results than naively applying page-level ranking on the image model which is usually used in the current search engines.*

**Keywords:** Hybrid ranking, Image rank, Popularity rank, Search engine, Searching order

1. **Introduction.** The emergence of World Wide Web has created many new opportunities but also challenges for organizing and searching a large volume of images available publicly [1-5]. In this paper, we explore and investigate the problem of searching images on the Web by introducing new concepts of Popularity-based rank, image link structure and image similarities. Most of previous Web-based applications [6-10] regard Web pages as atomic. However, it is often the case that a single Web page contains multiple topics. Here after, we will use the term HP to stand for Web page. Generally, a single HP is made up of hundreds of basic elements. The functional role of each element is different. For example, an image can be the banner of a Web site, an advertisement, or a picture for a news article [11-13]. Thus, a HP should not be the smallest unit. Naturally, it is more reasonable to regard the images and its contents as the smallest units of information. It is also worthwhile to note that there are various kinds of objects embedded in HPs. Each HP can be divided into a number of classes which contain semantically related information including images. Like HPs that are connected among each other to form a Page-to-Page graph, various classes and images also form Class-to-Class graph and Image-to-Image graph. In the Page-to-Page graph, different pages have different importance scores according to the number of in-links. Most of previous link analysis algorithms such as HITS [7] and Page Rank [6] use link structure to construct a Page-to-Page graph. Since then, a number of extensions to these algorithms have been also proposed [10,14-17]. All these algorithms are based only on a Page-to-Page graph. In this work, we consider three kinds of relationships: Page-to-Class, Class-to-Page and Class-to-Image, which ultimately lead Page-to-Page, Class-to-Class and Image-to-Image models.

In addition to these models, we adopt a new concept of popularity ranking which will play an important role in development of search engines. It is obvious that objects