

( 続紙 1 )

京都大学	博士 (情報学)	氏名	李 吉屹 (LI JIYI)
論文題目	A Study on Image Retrieval in Social Image Hosting Websites		
(論文内容の要旨)			
<p>The websites providing social networking services, such as Flickr, Twitter, Facebook and so on, have been important platforms for people to share information and communicate with others. Social images refer to huge image collections that users share on social networking service websites with additional user generated textual information. Social image hosting websites, such as Flickr, is a kind of social networking service websites that are specially constructed for sharing social images and have been rapidly growing recently. On these websites, users can upload, tag and share their images; and other users can search and browse images using these tags. Keyword annotation in traditional image retrieval systems, which asks for several experts to annotate images with some fixed keywords and costs too much time and labor for a large database. By contrast, social image hosting websites allow large number of users tag his/her own images. Therefore, it is becoming increasingly easy to construct a large database of tagged images. Social tags on images shed a new light on promoting image search for practical applications.</p> <p>The research on social image retrieval utilizing tags can be addressed into two dimensions: text-based and content-based. Social tags are naturally effective for text-based social image retrieval. Flickr has provided the service of tag-based image search, and there are also some existing works that can be used for improving text-based social image retrieval.</p> <p>However, whether and how social tags, which have an open vocabulary and lots of noise, can be used for improving content-based social image retrieval still has not been well investigated in existing work. This dissertation focuses on this topic and study three sub-topics.</p> <p>There are two problems in this overall topic. One problem is the semantic gap problem between low level visual features and high level image semantics. To handle this problem, this dissertation constructs a unified image-tag relationship graph model to analyze relationships between social images and tags. The approaches proposed in each sub-topic fuse both visual and textual information. The other problem is low quality tags describing image semantic content. This dissertation summarizes different kinds of low quality tags. Each sub-topic solves this problem with different manners respectively. The three sub-topics are as follows.</p> <p>The first sub-topic is concerned with unsupervised ranking utilizing social tags for content-based social image retrieval. The purpose of this sub-topic is to auto-</p>			

matically rank the images in initial content-based image search results for a given query image utilizing social tags. This thesis proposes an optimized mutual reinforcement process with an image-tag relationship graph model. The approach extracts and mutually propagates textual and visual information through the graph links to fuse these two kinds of information. The experimental results show that this approach outperforms existing ones which use visual information only or textual information only; existing approach which uses linear combination on these two kinds of information; and naive mutual reinforcement approach. It also has a better performance on noisy tag resistance.

The second sub-topic is supervised re-ranking utilizing social tags with multi modal relevance feedback for content-based social image retrieval. The task of this work is to boost the performance of image search results with only a few user interactions. This thesis proposes a novel multi modal relevance feedback scheme which allows users to label any social images and tags as positive and negative relevance feedback instances because information from various modalities is available in the scenario of social image retrieval. This thesis proposes a new approach that can effectively utilize MMRF information to improve content-based social image search results. This approach propagates MMRF information and fuses textual and visual information on an image-tag relationship graph model. It has better performance than the adapted approaches that are originally proposed for single modal relevance feedback. This work also analyzes various cases of MMRF selections and the potential user semantics contained in these selections. The original re-ranking approach is able to handle the potential user semantics related to the relationships among MMRF instances, with its graph model, computation rules and propagation-based iterations. In addition, the experimental results also show that MMRF scheme can significantly improve performance in contrast to traditional single modal relevance feedback scheme.

The third sub-topic is concerned with improving folksonomy tag quality in social image hosting websites: This dissertation does not only improve social image search results by proposing optimized image ranking and re-ranking approaches, it also provides another solution by improving data quality in the social image database. The target of this work is to add new tags and to rank the new tags as well as the raw tags for an input social image which can be tagged image or untagged image. After the tag quality of all images in the database is improved, it can generate an improved social image database. This work proposes a unified framework which has two stages to solve the problems of various cases of low quality tags. The first stage is to collect the candidate images with visual information and candidate tags with textual information, and can handle the problem of missing tags. The second stage uses such visual and textual information to analyze the relevance relationships between social images and tags to evaluate the relevance of the tags to the

input image, and can handle the problem of imprecise tags, meaningless tags and unranked tags. This work provides a series of approaches as the solutions of the second stage including three approaches that refer the approaches in the existing work and an original approach that has a mutual reinforcement process and fuse both visual and textual information iteratively. Because of the unified framework, all these approaches do not have the disadvantage of too depending on the initial tag set. They improves tag quality based on user generated folksonomy tags directly, which have a large concept space, without constructing and using manual training set. This work conducts computational experiments based on datasets constructed from Flickr. The experimental results show that all these approaches can improve tag quality prominently. In contrast to the approaches adapted from existing work, the original approach proposed in this work has better performance on NDCG metric and time complexity.

注)論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 words で作成し  
審査結果の要旨は日本語500～2,000字程度で作成すること。

(論文審査の結果の要旨)

本論文は、近年広く利用されるようになったソーシャル画像ホスティングウェブサイトにおける画像検索に関する研究に取り組んでいる。

内容に基づくソーシャル画像検索のためのソーシャルタグを用いたリランキングでは、画像とタグから成る二部グラフ上の枝に沿ってテキスト情報及び画像情報を伝搬する最適化された相互強化過程を実行することによりこれら二種類の情報を融合し、タグの情報を考慮したソーシャル画像の内容検索の精度向上を実現している。実験でそれを確認しており提案手法の有効性が認められる。

ソーシャル画像検索のための教師付再ランキングでは、内容に基づく画像検索のための新たなマルチモーダル適合性フィードバック手法として、利用者が画像、タグ、またはその両者に正例、負例のラベルを付与し、またそれらのフィードバック情報をデータベース内の画像、タグ情報から成るグラフ構造上で伝搬することによりすべて融合する手法を提案している。提案手法は、従来の単一モーダル適合性フィードバックスキームに比べて性能を大幅に改善できることを示し、新規性及び有用性の点で評価できる。

ソーシャル画像に付与されたフォークソノミータグは、各利用者が統制語彙に従わず自由に付与するため、不正確なタグ、意味のないタグ、付与されるべきタグの欠落など低品質なタグ付けが多く存在する。ソーシャルタグの改善では、全体のプロセスを、対象とする画像に関連する他の画像集合およびタグ集合を抽出する段階と抽出された画像およびタグ集合を基に対象画像のタグを改善する段階に分割する統一的な枠組みを提案した。第二の段階には既存のタグ推薦方法も利用できるため、実際に既存の三種類の手法を適用した結果、既存手法を単独で実行した場合に比較し性能が改善されることを確認した。また、第二段階として画像及びタグの情報を融合した新たな手法も提案し、第二段階に既存手法を使用した場合に比べて性能が改善されることを示した。このように新たな枠組みとタグ改善手法を提案し性能を改善した点が新規性と有用性の点で評価できる。

このように、本論文はソーシャル画像の内容検索のために画像検索ランキングの改善とタグ改善の問題に取り組んだものであり、その新規性と有用性が高く評価される。

よって、本論文は博士（情報学）の学位論文として価値あるものと認める。また、平成25年8月19日に実施した論文内容とそれに関連した試問の結果合格と認めた。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。更に、試問の結果の要旨（例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」）を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。  
要旨公開可能日： 年 月 日以降