

Visual Tagging Through Social Collaboration: a concept paper

Andrea Bellucci¹, Stefano Levialdi¹ and Alessio Malizia²

¹Department of Computer Science
University Sapienza of Rome
Via Salaria 113, 00100, Rome, Italy
{bellucci,levialdi}@di.uniroma1.it

²DEI Laboratory, Computer Science Department
Universidad Carlos III de Madrid
Avda. de la Universidad, 30. 28911-Leganés, Madrid (Spain).
alessio.malizia@uc3m.es

Abstract. Collaborative tagging has grown on the Internet as a new paradigm for web information discovering, filtering and retrieval. In the physical world, we use visual tags: labels readable by smartphones with cameras. While visual tags are usually related to a web site address, collaborative tagging, instead, provides updated, recommended information contributed and shared by users. In this paper we investigate the combination of collaborative tagging systems with visual tags. We present a prototype of a semiautomatic system generating visual tags which gather information from collaborative tagging. The user can interact with a list of relevant tags (built by clustering closely related tags) that can be further encoded in a visual tag, according to user's preferences. The user experience is enriched by retrieving multimedia content linked to the selected tags, present on the web. We finally show a case study illustrating our approach.

1 Introduction

Collaborative tagging[1,2] is the process by which users add metadata, by means of descriptive terms (tags), to community-shared content in order to classify or organize documents for further navigation, filtering or search. These tags are usually appended informally and personally by the author/creator of a document. Based on social structures, tagging is an adaptable process; it takes the form best supported by the content, letting users decide the categorization, rather than imposing a rigid structure on it. Nowadays there are a number of eminent web sites that rely on collaborative tagging: these systems allow users both to add tags and share content, so they can browse content categorized by other users as well as categorize information for their own purposes. An example of such sites is Flickr (<http://www.flickr.com>), a service giving to users the possibility to tag photographs they own. In the real world we assist to another tag proliferation: the one of visual tags[3]. Barcodes are the first example of visual tags; in this

sense tagging simply places a machine-readable label on a physical object. The main advantage of visual tagging is that the actual modification on the object is minimal, as its cost. Visual codes can be easily incorporated when designing the package or printed on labels, but they must be scanned using a camera, or a code reader, and they need to display the retrieved information. Smartphones seem to embody the natural solution: small mobile devices that incorporate a tag reader (the mobile camera), a display and an online connection into a single device. The prevalent use of visual tags is to associate physical objects to digital information. Systems like the Japanese QR Code[4] provide visual tags, easy to generate and print, that are readable with a mobile camera. QR Codes appear in the most unpredicted places in Japan: for example one common use is placing codes, related to a web site address, on magazines or advertisements. You can also find QR Codes on the timetable at a bus stop, linked to real-time data on traffic and bus schedules. It follows that the real power of mobile visual tagging is in providing complementary information when and where a user actually need it. In this paper we investigate the possibilities offered by collaborative tagging activities on the web combined with visual tags.

2 Motivation

In practice web sites have a static meaning since the content is only organized by the web master. Collaborative tagging systems, instead, usually contain information shared by users, frequently updated and providing hints (e.g. tags suggesting whether or not something was appreciated by the "tagger", for example recent photos of a vacation trip, and so on). Such systems also contain individuals' structural knowledge about documents. Structural knowledge has been defined as the knowledge of how concepts in a domain are interrelated (Diekhoff and Diekhoff 1982)[5]. Therefore, in a collaborative tagging system, tags encode the knowledge of relationships among documents and concepts represented by the tags. Moreover, as users with similar interests tend to have a shared vocabulary, tags created by one user turn useful to others, particularly those with similar interests as the tagger's.

However while tagging systems claim to have many advantages over controlled vocabularies or taxonomies, they still suffer from some limitations, such as the lack of lexical relations[6] or information overload[7]. Searching and retrieving may turn very limited if we ignore lexical relations, especially in a system where many users with different backgrounds have added different tags. A solution to these handicaps can be found by clustering techniques within the tag space: starting from a seed word, we identify 'families' of tags having a strongly related meaning.

We suggest that if we use individuals' knowledge of a given concept (i.e. about *Rome*), revealed through the set of personal tags, in conjunction with clustering sets of tags related to *Rome*, we may then identify a core group of *key tags*: such *key tags* make up a family of potentially interesting tags for the user. Such tags best describe an associated concept that the system user required: for

instance Rome can be followed by *Tiber* and *bridge*, representing different Roman bridges whilst followed by *Vatican* and *church* will depict Roman churches. Once identified, these *key tags*, can be encoded into a visual tag. The visual tag can be placed in a variety of physical objects (books, ads, historical ruins) and the contained information allows retrieval of online multimedia content in order to enrich the user experience. In this way, we have firstly established a bridge between the information contents (digital and online) provided by a group of people and the final user but also the possibility of sharing such information between a wide range of users so increasing the expressive power of such visual tag. The user may always retrieve the updated contents produced by other users in the world.

3 The system

We distinguish two classes of users: the editor of visual tags and the final user interested in retrieving multimedia content (images in our case) corresponding to her interests through the visual tag. The editor will submit a seed word and the underlying collaborative system (Flickr in our case) will firstly provide a list of strongly related tags established by clustering performed directly from the Flickr web site. The editor will choose those tags representative of her intended aim. For example if the seed word is *Rome* and the editor wants to provide pictures representing bridges on the Tiber river, she will select the tag *bridge* and the tag *Tiber* from the list (see Fig.1).

The editor will test her new tags by looking at corresponding images and either leave the chosen ones or add/delete other tags after checking their correspondence to her target concept. We next encode such tags into a Quick Response Code (QR Code) such as that shown on Fig. 2 (step 4). On the final user side, she will read the visual tag (on the site, totem or physical object) by means of a camera phone enabling her to retrieve updated information from Flickr.



Fig. 1. List of tags for the seed word *Rome*

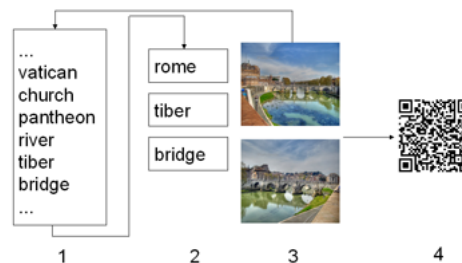


Fig. 2. The encoding process: input list (1), selected tags (2), corresponding images (3) and final, visual tag (4)

4 Case study

Our test example is in the area of sightseeing and vacation guides. Such booklets provide printed indications about places to see and their related history. This information is traditionally static (printed) and contains few (expensive) images of the site: such drawbacks may be overcome by using our visual tags that may be printed on a page of such travel book. The tourist with her camera phone may read the visual tag so obtaining recent (daily updated) images, related to the site she will be visiting, taken by other visitors of the same site.

5 Conclusion and future work

In this paper we suggest to combine online collaborative tagging with visual tags. We have presented a semiautomatic system for generating visual tags starting from a list of semantically related tags of a given seed word. We have briefly described a case example to illustrate our approach, using Flickr. Our system may overcome the digital divide by providing information obtained from other countries in the world (travelers in our case) so expanding the knowledge of the final user to a limit similar to the one of the rest of the world. We are currently investigating the use of other collaborative tagging systems, such as Youtube (<http://www.youtube.com>) in order to extend the retrieval process to a wide range of multimedia contents; at the same time we will explore different clustering techniques so as to maximize the semantic closeness of our tags.

References

- [1] Mates A., Folksonomies - Cooperative Classification and Communication Through Shared Metadata, Computer Mediated Communication, LIS590CMC, 2004.
- [2] Hammond T., Hannay T., Lund B., Scott J., Social Bookmarking Tools (I): A General Review, D-Lib Magazine 11(4), 2005.
- [3] Holmquist L.E.: Tagging the World. ACM Interactions 13(4), ACM Press, New York, pp 51-ff, 2006.
- [4] DENSO WAVE: <http://www.denso-wave.com/qr/code/index-e.html>, 2004.
- [5] Diekhoff G.M., Diekhoff K.B., Cognitive maps as a tool in communicating structural knowledge. Educational Technology, 22(4), pp. 28-30, 1982.
- [6] Begelman G., Keller P., Smadja F.: Automated Tag Clustering: Improving search and exploration in the tag space. WWW2006, May 22-26, Edinburgh, 2006.
- [7] Wu H., Zubair M., Maly K.: Harvesting Social Knowledge from Folksonomies. HYPERTEXT '06: Proceedings of the seventeenth conference on Hypertext and hypermedia, ACM Press, New York, pp 111-114, 2006.