

# Tools for image annotation Using context-awareness, NFC and image clustering

Randi Karlsen and Anders Andersen

Department of Computer Science

UiT The Arctic University of Norway, Tromsø, Norway

randi.karlsen@uit.no, anders.andersen@uit.no

## Abstract

Annotation of images is crucial for enabling keyword based image search. However, the enormous amount of available digital photos makes manual annotation impractical, and requires methods for automatic image annotation. This paper describes two complementary approaches to automatic annotation of images depicting some public attraction. The LoTagr system provides annotation information for already captured, geo-positioned images, by selecting nearby, previously tagged images from a source image collection, and subsequently collect the most frequently used tags from these images. The NfcAnnotate system enables annotation at image capture time, by using NFC (Near Field Communication) and NFC information tags provided at the site of the attraction. NfcAnnotate enables clustering of topically related images, which makes it possible to annotate a set of images in one annotation operation. In cases when NFC information tags are not available, NfcAnnotate image clustering can be combined with LoTagr to conveniently annotate every image in the cluster in a single operation.

## 1 Introduction

Keyword-based search is today the most common technique for searching images [18, 15]. Text-based image retrieval (TBIR) allows users to formulate high-level semantic queries, and are often more accurate and efficient in identifying relevant images compared to content-based image retrieval (CBIR) [3, 22, 23]. However, to be effective, a necessary basis for TBIR is the availability of annotations that reflect image content. Manual annotations are usually subjective and the annotation task is time consuming so that users are not likely to devote enough time to tag all their images [21, 23]. Another aspect is that people do not necessarily know or remember the names of all depicted objects (for instance attractions visited during a holiday).

Many tools for automatic and semi-automatic image annotation use CBIR techniques for linking visual features to keywords [20, 23]. Despite the achievements

---

*This paper was presented at the NIK-2014 conference; see <http://www.nik.no/>.*

in CBIR, bridging the semantic gap between low-level visual features and high-level semantic concepts is still a challenging task [3]. In addition, CBIR techniques often suffer from low efficiency and scalability caused by the high dimensionality of visual features [23]. To leverage these problems, some recent approaches for automatic image annotation combine content analysis techniques with the use of image metadata, where for instance GPS position of an image is used to focus tag collection to a relevant area [5, 6, 18, 11].

In our work, we have explored different techniques for automatic annotation of images. We first investigated the ability to automatically tag images based only on image metadata. We have designed a system, named LoTagr, which automatically annotates images based on the images' GPS coordinates. The coordinates are used to first select related, previously tagged images from a source image collection, and subsequently collect the most frequently used tags from these images. Important to LoTagr are some dynamic techniques that adapt image and tag selection to the availability of relevant images and tags in the source image collection.

In the second approach, we use NFC (Near Field Communication) and context-awareness as tools to collect accurate annotations at image capture time. We have designed a photo capturing application (for an Android based mobile phone) that clusters images and assigns annotations available through NFC tags. When taking images of an attraction, the user obtains image annotations by simply touching an NFC tag related to the attraction. An important part of this system, is our solution to the challenge of associating images to the correct NFC tag.

This paper describes both LoTagr and an NFC-based image annotation system named NfcAnnotator. We also discuss the usefulness of the image clustering functionality of NfcAnnotator, and how NfcAnnotator and LoTagr can be combined.

## 2 Related work

Research on automatic image annotation has resulted in different approaches for selecting textual terms from previously annotated images. A number of systems annotate a query image by selecting terms from related images gathered from online image collections (such as Flickr<sup>1</sup> and Panoramio<sup>2</sup>) based on a combination of geographic position and visual similarity [5, 6, 10, 11]. The general technique is to first collect a set of images within a certain radius of the query image, narrow down the set by using visual similarity techniques, and finally collect terms from the remaining images.

Expansion of user provided image keywords is described in the work of [22, 4, 9, 17]. New terms, recommended as expansions, are selected from relevant images based on visual similarity and/or co-occurrence analysis of tags. Automatic image tagging in mobile phone applications is described in for example [12, 16], where people can be identified using sensors detecting movement and direction, and location tags are suggested based on information such as location, previously used tags, tags from social contacts and temporal information.

Our work contrasts the referred work in that we do not include analysis of visual features in images. In the LoTagr system, we rather provide automatic image tagging based solely on image locations metadata. We chose this approach in

---

<sup>1</sup><http://www.flickr.com/>

<sup>2</sup><http://www.panoramio.com/>

order to avoid the semantic gap problem, and investigate the effectiveness of using image GPS coordinates as basis for the image tagging process. In the NfcAnnotate system, we offer accurate image annotation through the use of NFC technology. This is a novel approach, where the upcoming NFC technology for tagging our environment, conveniently can be used to add exact annotations to images, and thus avoid interpretations and possible selection of wrongful annotations.

### 3 Location-based annotation of images

Our approach to location-based image annotation is illustrated in Figure 1. A query image is submitted to the *Context-based tagger*, where the *Image Selector* uses the image's GPS coordinates to search Flickr for geographically close images. A list of photos are returned from Flickr, and Image Selector executes an algorithm for *dynamic image selection*. This algorithm may require repeated Flickr photo searches before a set of images has been selected as the source for tagging the query image. Using a *dynamic tag selection* algorithm, the *Tag Selector* processes the selected source images, collects the most frequently used tags and assigns them to the query image.

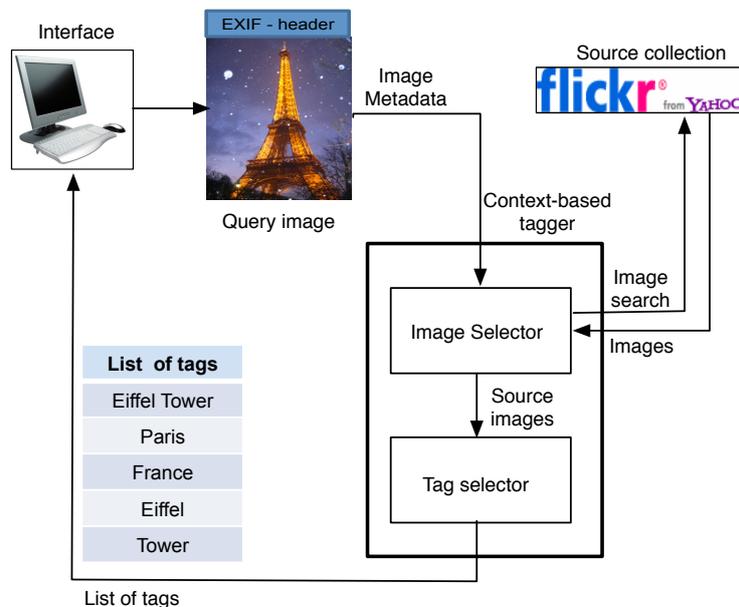


Figure 1: Location-based image annotation

Based on this architecture we have implemented a system, named LoTagr, that searches Flickr using the flickr.photo.search<sup>3</sup> API function with the image's GPS coordinates and a radius as parameters. The image itself is not part of the Flickr query. A description of LoTagr, together with other systems, is also found in [8].

#### Dynamic selection techniques

To select relevant tags for image annotation, we use two dynamic selection techniques. One technique selects an appropriate number of geographically nearby source images, while a second technique selects tags from the source image set.

<sup>3</sup><https://www.flickr.com/services/api/flickr.photos.search.html>

### *Dynamic image selection*

We use a *dynamic image selection* method that selects many Flickr images if there are many images close to the query image, and fewer images if there are few nearby images. The technique initially searches Flickr with a very small radius,  $r$ , and a fairly high amount of required images,  $ReqIm$ . The set of selected images,  $S$ , is large enough if  $ReqIm$  or more images are selected. If a sufficient number of images is selected, tags can be collected, otherwise the radius is doubled and the number of required images is reduced with one tenth of its last value, as shown in (1).

$$\begin{aligned} &\text{While } |S| < ReqIm \text{ AND } r < 32km \\ &\quad r = r * 2 \\ &\quad ReqIm = ReqIm - (ReqIm/10) \\ &\quad \text{Get new source set } S \text{ from Flickr} \end{aligned} \quad (1)$$

The image search to Flickr continues until  $S$  has reached the required number of images. The maximum allowed radius in flickr.photo.search is  $32km$ , and the search process therefore ends when the radius reaches  $32km$ , even if no images are found. Since images with very few tags provide little valuable information, images are selected so that  $S$  does not include images with less than 2 tags. LoTagr were tested using  $r = 0.001km$  and  $ReqIm = 50$  as initial values.

### *Dynamic tag selection*

The Tag Selector collects the most frequently used tags from images in the source set  $S$ , and assigns them to the query image. Formula (2) calculates Tag Frequency for term  $t$  in  $S$ . There  $|S(t)|$  represents the number of images in  $S$ , from unique owners, that use term  $t$ , while  $|S|$  represents the number of images in  $S$ .

$$TagFreq(t, S) = \frac{|S(t)|}{|S|} \quad (2)$$

The tag frequency function counts each unique tag once per image owner. This is done to avoid very user specific tags, in cases when users upload many images with the same set of tags.

The set of selected tags for query image  $Q$ , is specified in formula (3). Our approach is to discard tags with low frequencies (determined by the value of  $TagThreshold$ ). If the number of image in  $S$  is low (i.e. less than the value of  $NumLow$ ), the minimum times a tag must appear is set to  $TagMin$  (in stead of using the threshold value). This is done to prevent noisy tags from appearing.

$$\begin{aligned} SelectedTags(Q) = \{ &t | (t \notin FilterList) \wedge \\ &((TagFreq(t, S) \geq TagThreshold \wedge |S| \geq NumLow) \vee \\ &(|S| < NumLow \wedge |S(t)| \geq TagMin)) \} \end{aligned} \quad (3)$$

In the testing, we used the values  $TagThreshold = 0.2$ ,  $NumLow = 15$  and  $TagMin = 3$ . During testing we found these values to be effective for discarding noisy tags. Tags in the set  $FilterList$  are of no interest to any query image (examples include “geotagged”, “latitude” and “flickr”) and are thus filtered out.

Testing shows that when the number of images in the source set is very low, precision of the collected tags is also low. To tag a query image, we therefore

require a minimum number of images ( $ReqImMin$ ) in  $S$  (see Formula (4)). In the current testing,  $ReqImMin$  is set to 5.

$$\text{If } |S| < ReqImMin \text{ then } SelectedTags(Q) = \phi \quad (4)$$

## Testing and results

We tested LoTagr using 110 query images, with 78 images depicting an object and 32 an event. The content varied from very famous objects and events, such as the Eiffel Tower and Carnival in Rio, to less famous objects/events.

To test the effectiveness of using the dynamic selection techniques, we implemented and compared against *SimpleTagr*, which is a simplified version of LoTagr where no dynamic techniques are used. SimpleTagr searches Flickr based on GPS coordinates, selects the up to 250 images that are closest to the query image and returns all tags that appear in at least 20% of the selected images.

### Collected tags

For each image, we manually examined the automatically collected tags and distinguished between tags that were *content relevant*, *location names*, *unsure* and *irrelevant*. Examples of the different types of tags are shown in Figure 2, where we see one of the query images together with tags collected by LoTagr.

*Content relevant* tags (displayed in blue) are relevant with respect to what is seen in the image, while *location* tags (green) give correct location names. Incorrect tags and tags not relevant to image content or location are classified as *irrelevant* (violet), while tags that may or may not be considered relevant are classified as *unsure* (red). Typical examples of unsure tags are tags relevant to the location of the image but not visible on the image. Other examples are the two tags “hancock” and “city” in Figure 2.



Figure 2: Query image (of John Hancock Centre in Chicago) and collected tags

The average number of tags collected by LoTagr and SimpleTagr is shown in Table 1, while Figure 3 shows the relative proportion of the collected tag types for each system. We see that LoTagr collected the highest number of Content Relevant tags for both object and event images, and the fewest irrelevant tags. The number of location names and unsure tags are approximately the same in the two systems. The most noticeable difference between LoTagr and SimpleTagr is seen for object images, where the dynamic selection techniques increased the amount of content relevant tags with 63% (from 1.6 to 2.6) and decreased the amount of irrelevant tags with 73%.

### Tag relevancy and accuracy

The relevancy of the collected tags is evaluated by calculating precision scores for each image. As exact location names easily can be obtained for geo-positioned

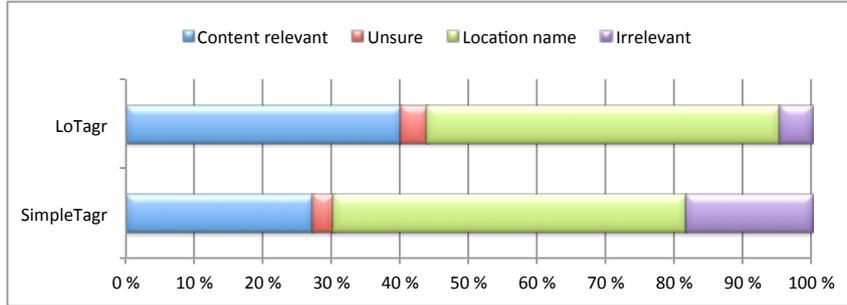
	CRelevant	Unsure	LocName	Irrelevant	Total
LoTagr	2.6	0.2	3.3	0.3	6.4
SimpleTagr	1.6	0.2	3.0	1.1	5.9

(a) Object images

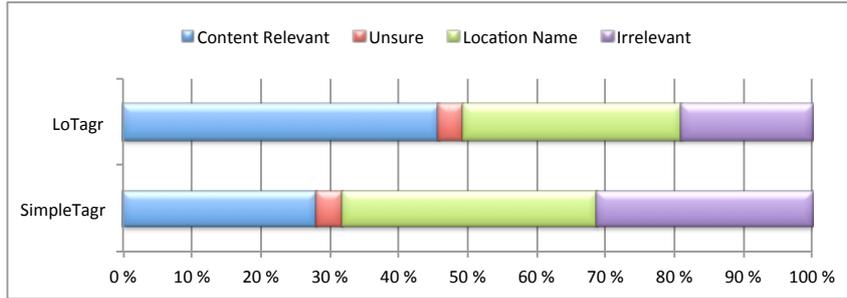
	CRelevant	Unsure	LocName	Irrelevant	Total
LoTagr	3.8	0.3	2.6	1.5	8.2
SimpleTagr	2.6	0.4	3.4	2.9	9.3

(b) Event images

Table 1: Average number of tags for event images



(a) Object images



(b) Event images

Figure 3: Relative proportion of different tag types

images, through for instance the Flickr function `flickr.places.findByLatLon`<sup>4</sup>, our main interest here is to evaluate the ability to select content relevant tags. All location names are therefore disregarded in the calculation of precision, even though they are relevant to the image.

Precision is calculated as shown in Formula 5, where  $CRelev(Q)$ ,  $LName(Q)$  and  $Unsure(Q)$  represent the sets of content relevant, location name and unsure tags, respectively, for image  $Q$ , while  $SelectedTags(Q)$  represents the set of all collected tags for  $Q$ . When comparing average precision scores for LoTagr and SimpleTagr, we found that the dynamic techniques clearly improves the relevance of the collected tags (see Table 2). Calculated over all images, the average precision score improved from 0.51 to 0.82 using dynamic techniques. This is a statistically significant improvement.

$$prec(Q) = \frac{|CRelev(Q)| + |Unsure(Q)|}{|SelectedTags(Q)| - |LName(Q)|} \quad (5)$$

<sup>4</sup><http://www.flickr.com/services/api/flickr.places.findByLatLon.html>

	<b>Object images</b>	<b>Event images</b>	<b>All images</b>
LoTagr	0.82	0.80	0.82
SimpleTagr	0.48	0.59	0.51

Table 2: Average precision of content relevant tags

To evaluate the accuracy of the collected tags, we identified the number of test images that received a tag that named the depicted object or event. The results are presented in Table 3, where we see that LoTagr collected the highest number of naming tags for both object and event images.

	Object images	Event images	All images
LoTagr	65%	44%	59%
SimpleTagr	40%	34%	37%

Table 3: Amount of images with a tag naming the depicted object/event

## 4 Image clustering and NFC-provided annotation

To enable accurate image annotation, we have developed NfcAnnotator, a photo capturing application for an Android based cellular phone that automatically adds NFC-provided information as annotation to images. When taking images of an attraction or event, the user collects image annotations by simply touching an NFC tag related to the attraction/event.

### NFC technology

Near field communication (NFC) is a set of short range wireless technologies in family with RFID [14], limited to distances below 10 cm (typically 4cm). The short range is a feature of NFC (not a limitation) that provides context-awareness. NFC communication involves an initiator and a target. The initiator generates an RF field that power a passive target, which consequently can take a simple form factor (such as a sticker, key fob or card).

Three NFC operating modes have been defined; card emulation, peer-to-peer and reader/writer mode. In *card emulation* mode credit cards, key cards, tickets, loyalty cards or similar are emulated, while in *peer-to-peer* mode two NFC devices are communicating as peers. The *reader/writer* mode, where the initiator can read and write data to NFC tags, is used in this work to annotate images.

There are many NFC-based information services using the reader/writer mode. In its most basic usage, data is collected from the NFC tag and displayed on the screen of the mobile device, while in more advanced applications the receiving of information triggers additional processing or delivery of user provided information [13]. Examples of some information-providing applications include smart poster applications [2], location-based wikis [19] and mobile museum guides [1].

## NfcAnnotate architecture

The architecture for NFC-based image tagging, shown in Figure 4, includes two parts, where the upper right part displays the activity of taking images and annotating by touching NFC tags, while the lower left part displays the process of storing image annotation information on the NFC tag. Annotation information can either be stored directly on the NFC tag or in a backend system, which responds with a URI that is written to the tag. NFC tags are then placed at/or their respective attraction.

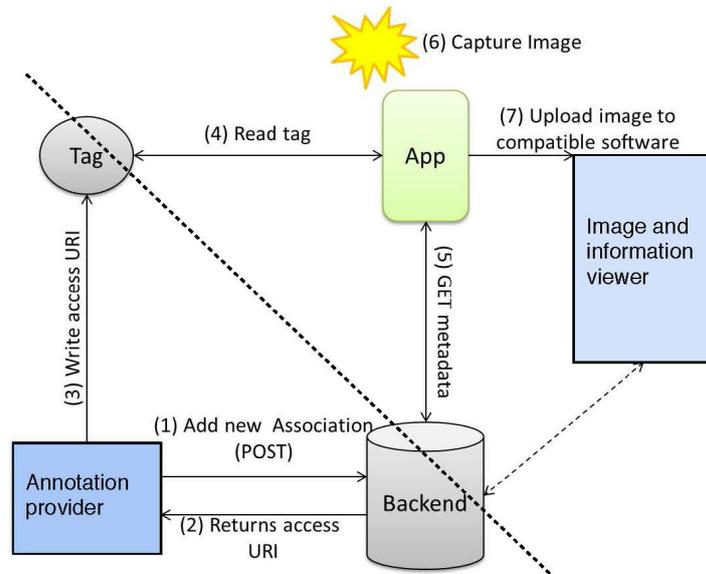


Figure 4: Architecture of NFC-based image annotation

When the NfcAnnotate application reads a URI on the NFC tag, it collects annotation information from the backend system and stores it as metadata on the image. If a connection to the backend system is not available, NfcAnnotate stores the URI on the image, making it possible to later download the annotations.

Images can also be uploaded to some compatible computer software (via wifi or Bluetooth), where they can be displayed, managed and more information about the image can be collected. This is described in [7] where also the software for registering annotation information is described.

## Associating tags and images

A main challenge for NFC-based image annotation is to associate an image with the correct NFC tag. A number of images can be taken of the same attraction, while we only want to touch the corresponding NFC tag once. One touch and multiple tag-image associations must thus be possible. Also, the tag might be scanned before or after image capture, and there might be a significant time gap between touch and capture. Finally, as there will never be an NFC tag for every image motif, some images will not have an image-tag association.

We solve these challenges by clustering related images. An image cluster  $O$  is in our system a set of continuously captured images related to the same

attraction/event. Cluster  $O$  can be associated to an NFC tag  $t$ , meaning that all images in  $O$  will be given the same set of annotation terms, obtained from  $t$ .

Table 4: Image group - tag association when starting and ending image groups

Activity	Situation	Image grouping	Association
Touch tag $t$	no active group	start group $O$	$\{O, t\}$
	active group, $\{O_1, t'\}$	end group $O_1$ , start group $O_2$	$\{O_1, t'\}$ $\{O_2, t\}$
	active group, $\{O, -\}$	continue group $O_1$	$\{O, t\}$
Push Change	no active group	start group $O$	$\{O, -\}$
	active group, $\{O_1, t'\}$	end group $O_1$ , start group $O_2$	$\{O_1, t'\}$ $\{O_2, -\}$
	active group, $\{O, -\}$	end group $O_1$ , start group $O_2$	$\{O_1, N\}$ $\{O_2, -\}$
Push Stop	no active group	nothing happens	-
	active group, $\{O, t\}$	end group $O$	$\{O, t\}$
	active group, $\{O, -\}$	end group $O$	$\{O, N\}$

Table 4 shows the cluster start/end activities and the resulting association between a cluster  $O$  and NFC tag  $t$ , denoted  $\{O, t\}$ . The *undefined association*  $\{O, -\}$  is used when the photographer is still capturing images in  $O$ , but an NFC tag has not yet been scanned, while the *none association*  $\{O, N\}$  is used when a cluster has ended without a tag scan.

If  $O$  is started with a touch of tag  $t$ , the association  $\{O, t\}$  is immediately stored. The cluster remains active, meaning that new images are added to the cluster, until a new tag is touched or the Change or Stop button is pushed. Touching a tag  $t$  having an active cluster  $O_1$  associated with tag  $t'$ , i.e.  $\{O_1, t'\}$ , results in the end of  $O_1$  and start of cluster  $O_2$  with association  $\{O_2, t\}$ .

To correctly cluster images taken before the relevant tag is scanned, NfcAnnotate gives the user the possibility to manually start/end clusters by using the *Change* and *Stop* buttons available in the applications' main menu. Pushing the Change button, ends an active cluster and starts a new cluster with an undefined association, i.e.  $\{O, -\}$ . A following NFC tag touch, does not start a new group, but identifies the association  $\{O, t\}$  and continue the current cluster. If the Change button is pushed without a touch of a tag, the association is changed to *none*,  $\{O, N\}$ . The relation between images in  $O$  are kept, as they belong to the same (undefined) topic. The user can later assign annotations to the set of images.

The Stop button ends a group without starting a new one. The following images will not be clustered and there will be no association to any tag. This is useful when the next image(s) will not be related to a common topic.

## Usefulness of image clustering

An effect of using the NfcAnnotate application is that images are clustered into groups that depict a common topic (i.e an attraction or event). Image clustering is known as a useful tool for image annotation, for example to ease the burden of manual annotation by allowing users to annotate a set of images in one operation [20].

In contrast to automatic image clustering techniques, based on visual features or date/time information [3], our approach provides topic-based clustering and has the advantage of allowing the user to cluster images as the user finds best. We expect visually dissimilar and locationally (somewhat) spread images to be grouped

together, as an attraction can be photographed from different angles, distances, outside and inside. Automatic clustering of images based on content analysis or closeness in location and time [3], is thus not suited for our purpose. A limitation to our approach is the need for user interaction that may complicate the process of taking images, in that the user may need to actively start and/or end image clusters.

Image clustering of NfcAnnotate and automatic annotation provided by LoTagr, can be combined to easily annotate a cluster of images. By using LoTagr to annotate one image in a cluster, all images in the cluster can be annotated with the same set of selected terms. Thus, given an image cluster  $O$  and an image  $Q$  such that  $Q \in O$ , the *SelectedTags*( $Q$ ) from LoTagr can be used for annotating every image in  $O$ .

When choosing an image  $Q \in O$  for LoTagr annotation, one should use an image that is close to the object of interest, and not an image taken from a distance. This is to use the most accurate GPS coordinate for the attraction, which also gives the most accurate LoTagr annotations.

NFC-based annotation provides accurate image annotation, as the photographer explicitly chooses the context of the images (i.e. an NFC tag) and the owner of the attraction explicitly adds correct annotation terms to the NFC tag. In cases when NFC tags are unavailable, and a combination of image clustering and LoTagr annotations are used, one can expect annotation precision and accuracy as described in Section 3.

Clustering of successively captured images may well result in a number of clusters covering the same topic. This may happen if the user switches between topics. NfcAnnotate supports storing of annotation information. Thus, once a set of annotation terms is provided (either through NFC annotation, LoTagr or by manually annotation), it is stored by the NfcAnnotator. The set of annotations can later be manually associated to new image clusters depicting the specific topic.

## 5 Conclusions

We have described two approaches to automatic image annotation. LoTagr is a system for automatically annotating geo-positioned images based on the images' GPS latitude and longitude values. A GPS coordinate and a radius is used for selecting a set of tagged, locationally relevant images from Flickr, from which the most frequently used tags are selected and used as annotation the query image. Important to LoTagr are some dynamic techniques that adapt image and tag selection to the availability of relevant images and tags in the source image collection. Testing shows that the dynamic techniques are effective for improving annotation precision.

NfcAnnotator is a system for annotating at image capture time using NFC. Images are clustered and assigned annotations available through NFC tags by simply touching the NFC tag related to the depicted object or event. We have in the paper described a technique for associating image clusters to the correct NFC-tag.

We have also describes how the two systems can be combined, as a set of topically clustered images from the NfcAnnotate application can be collectively annotated by simply providing LoTagr annotations for one of the images in the cluster. Sets of annotation terms, obtained either from LoTagr or NfcAnnotator, can also be stored and reused by manually relating the annotation sets to other, topically equivalent clusters.

## 6 Acknowledgments

The authors appreciate support from the Norwegian Research Council, NFR, through project nr. 176858, Context-aware Image Management, CAIM.

## References

- [1] Blockner, M., et. al., Please touch the exhibits!: using NFC-based interaction for exploring a museum. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*, ACM, New York, NY, USA, 2009.
- [2] Borrego-Jaraba, F., Luque Ruiz, I., Gmez-Nieto, M., A NFC-based pervasive solution for city touristic surfing, In *Personal Ubiquitous Comput*, Vol. 15, Issue 7, Springer-Verlag, London, UK, 2011.
- [3] Datta, R., Joshi, D., Li, J. and Wang, J. (2008) Image retrieval: Ideas, influences, and trends of the new age, *Computing Surveys (CSUR)*, Vol. 40, No. 2.
- [4] Gao, S., et. al., Automatic image tagging via category label and web data. In *Proceedings of the international conference on Multimedia (MM '10)*, ACM, New York, NY, USA, 2010.
- [5] Jones, G.J.F., et. al., Automated Annotation of Landmark Images using Community Contributed Datasets and Web Resources. In *Proceedings of the 5th international conference on Semantic and digital media technologies, SAMT'10*, Springer-Verlag, Berlin, Heidelberg, 2010.
- [6] Joshi, D., et. al., Using Geotags to Derive Rich Tag-Clouds for Image Annotation. In *Social Media Modeling and Computing*, Springer London, 2011
- [7] Karlsen, R., Andersen, A., NFC-Based Image Annotation. In *Trends in Mobile Web Information Systems*, Springer International Publishing, 2013
- [8] Karlsen, R., Evertsen, M.H., Elahi, N., Metadatabased automatic image tagging. In *International Journal of Metadata, Semantics and Ontologies*, Vol. 8, No. 4, 2013
- [9] Kucuktunc, O., et. al., Tag suggestr: Automatic photo tag expansion using visual information for photo sharing websites. In *SAMT '08: Proceedings of the 3rd International Conference on Semantic and Digital Media Technologies*, Berlin, Heidelberg, 2008. Springer-Verlag.
- [10] Mousselly, H., et. al., Geo-based Automatic Image Annotation. In *Proceeding of the ACM International Conference on Multimedia Retrieval (ICMR'12)*, Hong Kong, China, June, 2012.
- [11] Moxley, E., Kleban, J. and Manjunath, B., Spirittagger: A Geo-Aware Tag Suggestion Tool Mined from Flickr. *MIR '08: Proceeding of the 1st ACM international conference on Multimedia information retrieval*, Vancouver, Canada, Oct 2008.

- [12] Naaman, M., Nair, R., ZoneTag's Collaborative Tag Suggestions: What is This Person Doing in My Phone?. *IEEE MultiMedia* 15, 3 (July 2008), 34-40. IEEE Computer Society Press, Los Alamitos, CA, USA, 2008
- [13] Ok, K., Coskun, V., Aydin, M.N., Ozdenizci, B., Current Benefits and Future Directions of NFC Services, In *Proceedings of the International Conference on Education and Management Technology (ICEMT 2010)*, Cairo, Egypt, Nov. 2010.
- [14] Ortiz, C.E., *An introduction to near-field communication and the contactless communication API*, Tech. rep., Oracle Sun (2008)
- [15] Pagare, R., Shinde, A. (2012) A study on image annotation techniques, *International Journal of Computer Applications*, Vol. 37, No. 6, pp.4245.
- [16] Qin, C., et. al., TagSense: A Smartphone-based Approach to Automatic Image Tagging. In *MobiSys'11: Proceedings of the 9th International Conference on Mobile Systems, Applications and Services*, Bethesda, Maryland, USA, June 28-July 1, 2011.
- [17] Rae, A., Sigurbjörnsson, B., van Zwol, R., Improving tag recommendation using social networks. In *Proceedings of the 9th International Conference on Adaptivity, Personalization and Fusion of Heterogeneous Information (RIA0 '10)*, Paris, France, 2010.
- [18] Sergieh, H.M., Gianini, G., Dller, M., Kosch, H., Egyed-Zsigmond, E. and Pinon, J-M., Geo-based automatic image annotation, In *Proceedings of the 2nd ACM International Conference on Multimedia Retrieval in ICMR12*, ACM, New York, USA, 2012.
- [19] Siira, E., Tuikka, T., Tormanen, V., Location-Based Mobile Wiki Using NFC Tag Infrastructure. In *Proceedings of the 2009 First International Workshop on Near Field Communication (NFC '09)*, IEEE Computer Society, Washington, DC, USA, 2009.
- [20] Wang, M., et. al., Assistive Tagging: A Survey of Multimedia Tagging with Human-Computer Joint Exploration. In *ACM Computing Surveys*, Vol. 44, No. 4, ACM, New York, NY, USA, 2012.
- [21] Whittaker, S., Bergman, O. and Clough, P., Easy on that trigger dad: a study of long term family photo retrieval, In *Personal Ubiquitous Computing*, Vol. 14, No. 1, 2010, pp.3143.
- [22] Wu, L., Jin, R., Jain, A., Tag Completion for Image Retrieval, in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, May 2012.
- [23] Zhang, D., Islam, M., Lu, G., A review on automatic image annotation techniques, In *Pattern Recognition*, Vol. 45, No. 1, Elsevier Science Inc., 2012.