Metadata Management, Reuse, Inference and Propagation in a Collection-Oriented Metadata Framework for Digital Images

William Ku[1], Mohan S. Kankanhalli[1], and Joo-Hwee Lim[2]

[1] National University of Singapore, 3 Science 2, Singapore 117543
[2] Institute of Infocomm Research, 21 Heng Mui Keng Terrace, Singapore 119613
{kucheech,mohan}@comp.nus.edu.sg, joohwee@i2r.a-star.edu.sg

Abstract. Digital photography generates a lot more "shoeboxes" of photos than its conventional counterpart, resulting in image search and retrieval being more applicable. We briefly discuss some research challenges faced with the use of metadata in image search and retrieval. We then propose the structural use of metadata regularity of photos within collections (the *Group Effect*), in metadata management, reuse, inference and propagation. This application of the *Group Effect* is complemented by the *Social Networking Effect* whereby user interactions with image collections provide collaborative metadata. This is followed by our presentation of a set-theoretic approach to our framework (proposed in previous work [5,6]) and we then outline its application and utility.

Keywords: Image Metadata Management, Reuse, Inference, Propagation, Collection-Oriented Framework, Group Effect, Social Networking Effect.

1 Introduction

The high ease and low cost of digital photography empower people to take as many photos as they like. The "trigger-happy" user will find herself with tons of photos and while storage may not be an issue in terms of availability and costs, the user may have difficulties browsing, searching and retrieving from her massive collection, the exact photos that she may want to show to her family and friends. Image search and retrieval has become a needle-in-a-haystack issue. Nevertheless, digital photos have the capability to encode the "thousand words (and more) that they tell" in the form of metadata, which can be used in general to assist in search and retrieval. However, the key here is that the correct metadata and a proper search mechanism must be present in order for the search and retrieval in a digital photo collection to be workable. Unfortunately, the ideal type of annotation namely manual annotation is a tedious, inconsistent and erroneous process.

There has been significant research on automatic annotation. It is sufficient to state that at this point of time, there remain significant issues [1,10] that are yet to be fully addressed for efficient and accurate automatic annotation. Thus with large image collection, image search and retrieval has become an important research challenge. At a panel in ACM Multimedia 2005, the importance of image search and retrieval was

T.-J. Cham et al. (Eds.): MMM 2007, LNCS 4352, Part II, pp. 145 – 154, 2007.

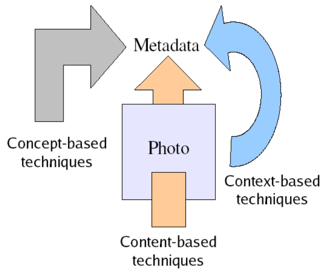

Fig. 1. The three main approaches to automatic metadata generation

emphasised by a panelist who said that "*Image retrieval may not be a killer app but not having it is an app killer*".

An alternative approach to enhance existing image search and retrieval techniques is to generate more metadata from the context of the image. The simultaneous use of content-based, concept-based and context-based techniques (Figure 1) should result in better image search and retrieval. We observe that digital images are usually part of a collection and that this association could provide some context on the nature of the photos themselves. For example, one would expect that a photo album with the title "My birthday party" would only contain photos pertaining to a birthday party. This will give rise to some interesting observations. Thus, photos belonging to this particular collection should have some identical information such as the location and date. More importantly, one should be able to obtain more contextual information from a group of photos than from the photos individually, such as the event-type. This group contextual information could be used for inference.

A second observation is that when users share their photos, the user interaction could provide collaborative annotations. Using the above example, one could group together the various photos (taken by different users) at the birthday party and examine their respective annotations to infer new contextual information such as the names of the participants.

Our contribution in this paper would be to incorporate the above two observations namely the *Group Effect* and the *Social Networking Effect* into our collection-oriented approach which would make use of metadata regularity at the collection level to extract contextual metadata for reuse, inference and propagation. The injection of collaborative metadata would further fuel reuse, inference and propagation.

The rest of this paper is organised as follows. We shall first touch on some related work on context-based techniques. This is followed by a discussion of our collection-oriented approach which is based on the *Group* and *Social Networking Effects*. We then present our set-theoretic approach to our collection-oriented framework. We concluded by applying our approach to a set of photos taken at ACM Multimedia 2005.

2 Related Work

In the domain of digital photos, there are some work on metadata sharing and reuse on a group basis. LOCALE [8] tags unlabeled photographs using shared information based on other photos taken in the same area. The social-temporal-social context of a group of cameraphone users would influence the metadata values of their images [3]. Snap2Tell [2, 7] matches a photo taken with cameraphone with a database, using content-based features and metadata. Event and location groupings can be used to suggest name labels [9]. *MyPhotos* [11] describes a prototype system for home photo management processing that replaces traditional folders with photo groups.

In summary, present work do not make structured use of photo metadata regularity in conjunction with collaborative metadata for personal photo collections while at the same time, provides for metadata conflict resolution, reuse, inference and propagation at the collection level. This is what our collection-oriented framework would attempt to achieve.

3 Collection-Oriented Approach

Our collection-oriented approach of handling digital images is based on two effects namely the Group Effect and the Social Networking Effect. The Group Effect is essentially the establishment of metadata regularity from the observation that images are usually part of a collection. The Social Networking Effect refers to the observation that one could share digital photos with his family and friends through the Internet and which provides a channel for collaborative annotation. A detailed discussion on the Group and Social Networking Effects can be found in [6].

4 Metadata Reuse, Propagation and Inference

In this section, we shall briefly discuss the relationship of metadata reuse, inference and propagation. Figure 2 gives an overview picture of this relationship. Initially, the metadata of the photos in a collection are in a state of *equilibrium*. That is to say, no further inference could be achieved on the existing set of metadata. However, when there is an interaction, the equilibrium would be disturbed and new metadata would be detected, inferred and propagated. Here, an interaction is an action defined by one

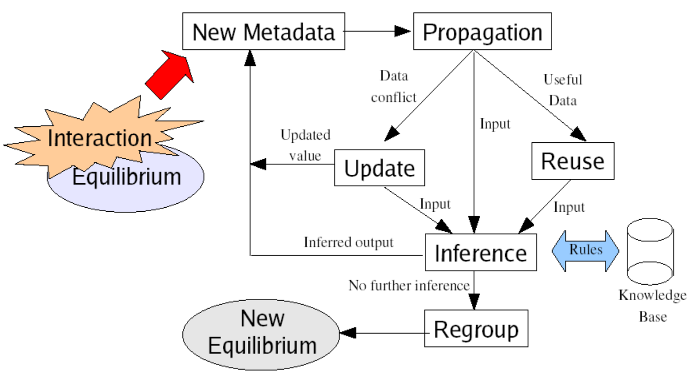

Fig. 2. Overview of the relationship between reuse, inference and propagation

of the following (but not limited to): Insert annotation(s), Create collection(s), Insert/Delete photo(s) into collection(s) and Merge/Partition collection(s).

Thus an interaction would trigger the propagation of new metadata within a photo itself and to other photos in the same collection. The new metadata could be an update to existing metadata or new useful information to be reused or as an input to an inference process. The updated or reused metadata could become an input to an inference process as well. The inference process may generate new metadata that would get propagated and this propagation cycle would go on until no further inference could take place. At this point of time, the existing metadata are examined to see if common metadata could be extracted and reused at the group level. When this is done and there is no further action, the photos are said to reach a new state of equilibrium. A further discussion on this issue together with the modeling of the relationship of metadata management, reuse, inference and propagation in some form of finite-state automata can be found in [6].

5 Collection-Oriented Framework for Digital Images

We had presented a collection-oriented framework for digital images in [5] and here we followed up with a set-theoretic approach to this framework.

5.1 Basic Definitions

Metadata Element definitions. The (*basic) metadata element E* is an *ordered pair (A,V)* consisting two members namely the *metadata attribute A* and the *metadata* value *V*. *E* can also be denoted as *<A,V>*. In set theory, a common definition of an ordered pair can be defined in the following manner: $(a,v)=\{\{a\},\{a,v\}\}$. *E* is *uninstantiated* when it does not have a metadata value. This is by defining *V* to be the empty set ($\emptyset$). In this case, we define *E* to be the *base metadata element* $\hat{E}$. This can be expressed as $E=\hat{E}\ \mathit{iff}\ V=\emptyset$.
All metadata elements are *similar* to their base metadata elements : $E\sim\hat{E}$.
E can always be reduced to $\hat{E}$: $E\rightarrow\hat{E}\Rightarrow V=\emptyset$.
Given that $E=(A,V)=\{\{a\},\{a,v\}\}$, $\hat{E}=\{\{a\}\}$. Thus, we observe that $\hat{E}$ is simply a set containing the singleton *A*. *E* is undefined when $A=\emptyset$.
We next define some fundamental concepts of the metadata element (see Figure 3).
E_1 is defined to be *identical (equal)* to E_2 when $E_1=E_2,\ \mathit{iff}\ A_1=A_2\wedge V_1=V_2$.
E_1 is defined to be *similar* to E_2 when $E_1\sim E_2,\ \mathit{iff}\ A_1=A_2\wedge V_1\neq V_2$.

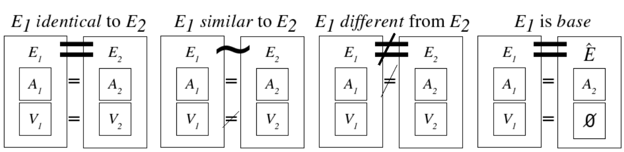

Fig. 3. Some fundamental element concepts

E_1 is defined to be *different* from E_2 when $E_1 \perp E_2, \textit{iff } A_1 \neq A_2$.
E_1 is defined to be *not equal (unique)* to E_2 when $E_1 \neq E_2, \textit{iff } E_1 \sim E_2 \vee E_1 \perp E_2$.

Metadata Schema definitions. The *metadata schema S* is a *set* of (*basic) metadata elements*. Hence, all elements in *S* are *unique*. Thus, $S = \{E_1, E_2, ..., E_{NE}\}$ where $NE = |S|$.

S is defined to be a *base schema (or schema template)* $\hat{S}$ when every of its elements is a *base element* (see Figure 4), that is $S = \hat{S}, \textit{iff } \forall E \in S, E = \hat{E}$. Thus, $\hat{S}$ is essentially a *set* of *base metadata elements*.

S can be reduced to $\hat{S}$ by reducing every of its members to its base form. $S \rightarrow \hat{S} \Rightarrow \forall E \in S \rightarrow \hat{E}$

We define some fundamental concepts of the metadata schema (see Figure 4).

S_1 is defined to be identical (equal) to S_2 when $S_1 = S_2, \textit{iff } S_1 \setminus S_2 = S_2 \setminus S_1 = \emptyset$.

S_1 is defined to be similar to S_2 when $S_1 \sim S_2, \textit{iff } S_1 \setminus S_2 \neq \emptyset \wedge \hat{S}_1 = \hat{S}_2$. S is always similar to $\hat{S}$.

S_1 is defined to be different from S_2 when $S_1 \perp S_2, \textit{iff } \hat{S}_1 \cap \hat{S}_2 \neq \emptyset$.
S_1 is defined to be not equal (unique) to S_2 when $S_1 \neq S_2, \textit{iff } S_1 \setminus S_2 \neq \emptyset \vee S_2 \setminus S_1 \neq \emptyset \vee S_1 \sim S_2 \vee S_1 \perp S_2$.

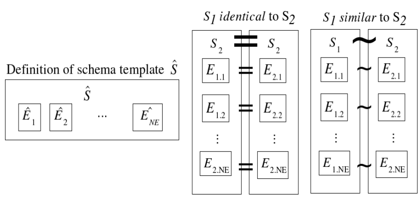

Fig. 4. Some fundamental schema concepts

Image Metadata Structure definitions. The *image metadata structure IM* of an image *I* is a *set* of *metadata schemas*. All elements in *IM* are *unique*. Thus, $IM = \{S_1, S_2, ..., S_{NS}\}$ where $NS = |IM|$.
IM_1 is defined to be identical (equal) to IM_2 when $IM_1 = IM_2, \textit{iff } IM_1 \setminus IM_2 = IM_2 \setminus IM_1 = \emptyset$.
IM is defined to be a *base image metadata structure (or image metadata template)* $\hat{IM}$ when every of its elements is a *base metadata schema*.
$IM = \hat{IM}, \textit{iff } \forall S \in IM, S = \hat{S}$
IM can be reduced to $\hat{IM}$ by reducing every of its members to its base form.
$IM \rightarrow \hat{IM} \Rightarrow \forall S \in IM \rightarrow \hat{S}$

IM_1 is defined to be similar to IM_2 when $IM_1 \sim IM_2, \textit{iff } IM_1 \setminus IM_2 \neq \emptyset \wedge \hat{IM}_1 = \hat{IM}_2$. IM is always similar to $\hat{IM}$.

IM_1 is defined to be different from IM_2 when $IM_1 \perp IM_2, iff\ \hat{IM}_1 \cap \hat{IM}_2 \neq \emptyset$.
IM_1 is defined to be not equal (unique) to IM_2 when

$$IM_1 \neq IM_2, iff\ IM_1 \setminus IM_2 \neq \emptyset \vee IM_2 \setminus IM_1 \neq \emptyset \vee IM_1 \sim IM_2 \vee IM_1 \perp IM_2$$

Image Collection definitions. An *image group G* is defined to be a *subset of a powerset of a set of unique images*. Its corresponding *group metadata structure GM* is defined to be a *subset of a powerset of a set of image metadata structures*. Thus, the definitions and operations defined for image metadata structures are applicable here.

5.2 Operations

In this section, we shall show some operations that the framework can provide. We would look at operations involving the metadata element and the metadata schema.

Element operations

1. Determining if an element E is a base element $\hat{E}$.
   $|E|=1 \Rightarrow E=\hat{E}$, otherwise $|E|=2 \wedge E \neq \hat{E}$
   Given $E=(a,v)=\{\{a\},\{a,v\}\}$, $\hat{E}=\{\{a\}\}$
   $E=\hat{E} \Rightarrow E=\{\{a\}\} \Rightarrow |E|=1$, otherwise $|E|=2 \wedge E \neq \hat{E}$

2. Determining the attribute a and the value v
   For an ordered pair *(a,v)*, $(a,v)=\{\{a\},\{a,v\}\}$
   Let $(a,v)=\{\{a\},\{a,v\}\}=\{X_{1,} X_2\}$ such that $|X_1|=1 \wedge |X_2|=2$
   Then $a=X_1 \cap X_2$ and $v=X_2 \setminus X_1$

3. Given that
   $E=(a,v)=\{\{a\},\{a,v\}\}$, $E_1=(a_{1,} v_1)=\{\{a_1\},\{a_{1,} v_1\}\}$ and
   $E_2=(a_{2,} v_2)=\{\{a_2\},\{a_{2,} v_2\}\}$
   A. $E_1 \cap E_2$
   A1. $E_1 \cap E_2 = E = E_1 = E_{2,} iff\ E_1 = E_2$
   $E_1 \cap E_2 = \{\{a_1\},\{a_{1,} v_1\}\} \cap \{\{a_2\},\{a_{2,} v_2\}\}$
   $E_1 = E_2 \Rightarrow a = a_1 = a_2 \wedge v = v_1 = v_2$
   $E_1 \cap E_2 = \{\{a\},\{a,v\}\} \cap \{\{a\},\{a,v\}\}$
   $E_1 \cap E_2 = \{\{a\},\{a,v\}\} = (a,v) = E$
   A2. $E_1 \cap E_2 = \hat{E}, iff\ E_1 \sim E_2$ (This operation may be used to derive $\hat{E}$)
   $E_1 \cap E_2 = \{\{a_1\},\{a_{1,} v_1\}\} \cap \{\{a_2\},\{a_{2,} v_2\}\}$
   $E_1 \sim E_2 \Rightarrow a = a_1 = a_2 \wedge v_1 \neq v_2$
   $E_1 \cap E_2 = \{\{a\},\{a,v_1\}\} \cap \{\{a\},\{a,v_2\}\}$
   $E_1 \cap E_2 = \{\{a\}\} = \hat{E}$

A3. $E_1 \cap E_2 = \emptyset, iff\ E_1 \perp E_2$

$E_1 \cap E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cap \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 \perp E_2 \Rightarrow a_1 \neq a_2$

$E_1 \cap E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cap \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 \cap E_2 = \emptyset$

B. $E_1 \cup E_2$

B1. $E_1 \cup E_2 = E = E_1 = E_{2,}\ iff\ E_1 = E_2$

$E_1 \cup E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cup \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 = E_2 \Rightarrow a = a_1 = a_2 \wedge v = v_1 = v_2$

$E_1 \cup E_2 = \{\{a\}, \{a, v\}\} \cup \{\{a\}, \{a, v\}\}$

$E_1 \cup E_2 = \{\{a\}, \{a, v\}\} = (a, v) = E$

B2. $E_1 \cup E_2 = ?, iff\ E_1 \sim E_2$

$E_1 \cup E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cup \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 \sim E_2 \Rightarrow a = a_1 = a_2 \wedge v_1 \neq v_2$

$E_1 \cup E_2 = \{\{a\}, \{a, v_1\}\} \cap \{\{a\}, \{a, v_2\}\}$

$E_1 \cup E_2 = \{\{a\}, \{a, v_1\}, \{a, v_2\}\} = ?$ => conflict resolution

B3. $E_1 \cup E_2 = \{E_{1,} E_2\}, iff\ E_1 \perp E_2$

$E_1 \cup E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cup \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 \perp E_2 \Rightarrow a_1 \neq a_2$

$E_1 \cup E_2 = \{\{a_1\}, \{a_{1,} v_1\}\} \cup \{\{a_2\}, \{a_{2,} v_2\}\}$

$E_1 \cup E_2 = \{\{a_1\}, \{a_{1,} v_1\}, \{\{a_2\}, \{a_{2,} v_2\}\} = ?$ => conflict resolution

Schema operations. The following are some example operations at the metadata schema level.

A. $S_1 \cap S_2$

A1. $S_1 \cap S_2 = S = S_1 = S_{2,}\ iff\ S_1 = S_2$

A2. $S_1 \cap S_2 = \emptyset\ iff\ S_1 \perp S_2$

A3. $S_1 \cap S_2 = S*, iff\ S_1 \sim S_2$

$S_1 \sim S_2 \Rightarrow \hat{S}_1 = \hat{S}_2 = \hat{S}$

Let $CS = S_1 \cap S_2$, $DS_1 = S_1 - CS$ and $DS_2 = S_2 - CS$

Then there exists $x_1 \in DS_1$ and $x_2 \in DS_2$ such that

$x_1 \sim x_2$ => conflict resolution

B. $S_1 \cup S_2$

B1. $S_1 \cup S_2 = S = S_1 = S_2, \; iff \; S_1 = S_2$

B2. $S_1 \cup S_2 = \{S_1, S_2\} \; iff \; S_1 \perp S_2$

B3. $S_1 \cup S_2 = S*, \; iff \; S_1 \sim S_2$ => conflict resolution

*S** refers to an updated *S*.

6 Experiments

In this section, we would like to demonstrate the utility of the Group and Social Networking Effects. Suppose we want to build a personal photo collection of photos taken at ACM Multimedia 2005. One way to do so is to search Flickr with the tag "multimedia2005" that the conference organisers have provided [12], which rendered a total of 70 photos taken by four users, all labeled with the "multimedia2005" tag.

Table 1 shows the breakdown of the number of photos taken by these four users and the least and most recent dates the photos were taken on. Thus, we have four collections and our role here is to merge these four collections into one and to make use of metadata regularity for metadata management, reuse, propagation and reuse.

We want to first select the collection with high metadata regularity. This would allow for high reuse. We do so by using a simple metric here.

$$metadata\ regularity = number\ of\ photos \times average\ tag\ frequency$$

where $average\ tag\ frequency = \dfrac{\sum tag\ frequency}{number\ of\ tags}$

Table 1. Breakdown of Flickr search results

User	*Number of photos*	*Least recent date*	*Most recent date*
A	26	07 Nov 05	09 Nov 05
B	28	07 Nov 05	11 Nov 05
C	13	07 Nov 05	10 Nov 05
D	3	08 Nov 05	09 Nov 05

Tag Frequency refers to the percentage of the photos in which a particular tag appears. Table 2 shows the breakdown of tag frequency while Table 3 tabulates the computation of this simple metric and we would want to choose the collection with the highest value for the metadata regularity metric, which is the collection by User C. We would reuse tags which have a tag frequency of 100% and exported it to the group level metadata, which also includes the number of photos and the least and most recent dates. At this time of point, our group metadata is as follows:

GM_C = {number_photo=13[number], least_recent_date="07 Nov 05"[date], most_recent_date="10 Nov 05"[date], common_tag="singapore"[text], common_tag="2005"[text]}.

Table 2. Breakdown of tag frequency

User	*Number of Tags*	*Tag Frequency (%)*
A	12	Singapore (100%), ACM (100%), Multimedia (100%), others (4% - 8%)
B	30	3.5% - 7%
C	2	Singapore (100%), 2005 (100%)
D	1	YRB (33%)

Table 3. Tabulation of metadata regularity

User	*Average Tag Frequency (%)*	*Metadata Regularity(no unit)*
A	29%	754
B	4%	111
C	100%	1300
D	33%	33

We would now merge collections in decreasing order of the metadata regularity metric, i.e. A followed by B and finally D. In combining collections, we would reuse tags with tag frequency of 100% provided that the metadata regularity is above 100, i.e. not tags of a collection containing just one photo.

Combining the collections by users C and A, we have

GM_{CA} = {number_photo=39[number], least_recent_date="07 Nov 05"[date], most_recent_date="10 Nov 05"[date],common_tag="singapore"[text], common_tag="2005"[text], common_tag="acm"[text], common_tag="multimedia"[text]}

Here, the metadata values of the group metadata would be adjusted and updated accordingly, subject to conflict resolution. For instance, there would not be a duplicate tag "singapore". The final collection group metadata would be

GM_{CABD} = {number_photo=70[number], least_recent_date="07 Nov 05"[date], most_recent_date="11 Nov 05"[date],common_tag="singapore"[text], common_tag="2005"[text], common_tag="acm"[text], common_tag="multimedia"[text]}

Thus any new photo that is inserted into this collection but does not have any metadata would be able to reuse the group metadata.

7 Conclusion and Future Work

We briefly covered some problems faced in digital image search and retrieval and discussed how we could make use of metadata regularity (the Group Effect) and collaborative annotations (the Social Networking Effect). We next presented our set-theoretic approach to our framework (proposed in a previous work [6]) and outlined its application and utility in an example image collection that made use of both the Group and Social Networking Effects.

In our experiments, we showed that it is possible to generate more metadata from the application of the Group Effect from the collaborative metadata made available from the Social Networking Effect. This generated metadata may be further applicable for reuse, inference and propagation.

In future work, we would be working on the EXIF metadata of the photos used in our experiments as well as that of the official ACM Multimedia photo collection [20] and to similarly apply our set-theoretic framework to determine the metadata reuse, inference and propagation.

References

1. Dick C.A. Bulterman. Is it Time for a Moratorium on Metadata? IEEE Multimedia, Vol. 11 Issue 4, pp. 11 – 17, 2004.
2. Jun Li, Joo-Hwee Lim and Qi Tian. Automatic Summarization for Personal Digital Photos, PCM 2003, vol. 3, pp. 1536-1540, 2003.
3. Marc Davis, Simon King, Nathan Good and Risto Sarvas. From context to content: leveraging context to infer media metadata, ACM Multimedia 2004, pp. 188 – 195.
4. Exchangeable Image File Format (EXIF), Version 2.2, JEITA.
5. William Ku, Mohan S. Kankanhalli and Joo-Hwee Lim. A Collection-Oriented Metadata Framework for Digital Images, ICME 2006, pp. 761 – 764.
6. William Ku. Exploiting "The World is Flat" syndrome in digital photo collections for contextual metadata, accepted for IEEE ISM 2006.
7. Joo-Hwee Lim, Jean-Pierre Chevallet and Sihem Nouarah Merah. SnapToTell: ubiquitous information access from camera. A picture-driven tourist information directory service, MobileHCI 2004, pp. 21-27.
8. Mor Naaman, Andreas Paepcke and Hector Garcia-Molina. From Where to What: Metadata Sharing for Digital Photographs with Geographic Coordinates, CoopIS 2003, pp. 196-217.
9. Mor Naaman, Ron B. Yeh, Hector Garcia-Molina and Andreas Paepcke. Leveraging Context to Resolve Identity in Photo Albums, JCDL 2005, pp. 178 – 187.
10. Frank Nack. All Content Counts: The Future in Digital Media Computing is Meta, IEEE MultiMedia, vol. 7, no 3, pp. 10-13, 2000.
11. Yanfeng Sun, Hongjiang Zhang, Lei Zhang and Mingjing Li. *MyPhotos*: a system for home photo management and processing, ACM Multimedia 2002, pp. 81 – 82 .
12. http://www.flickr.com/photos/tags/multimedia2005/
13. http://av.comp.nus.edu.sg/gallery/ACM-Multimedia-2005