Acquaintances Clustering for Social Relationship-Based Indexing of Digital Photos

Jonghak Kim, Taekwon Jang, Joonhyuk Yang and Jung-hee Ryu,

Graduate School of Culture Technology,
Korea Advanced Institute of Science and Technology (KAIST),
373-1 Guseong-dong, Yuseong-gu, Daejeon, Republic of Korea
{airjonghak, taekwon, joony.yang, junghee.ryu}@kaist.ac.kr

Abstract. One of the effective ways to manage large collections of digital photos is to tag names of people appearing in those photos. However, the number of people appearing in photo collections may range in the hundreds and the names of tagged people are usually presented in alphabetical order or on a first-tag first-place basis. As a result, it is difficult to quickly search for name tags that a user wishes to find. In order to solve this problem, we developed a digital photo management system that automatically groups the name tags based on their social relationships. This system was tested on users' own photos against three other comparison interfaces. The average searching time for name tags was significantly faster with our system. Also, the user satisfaction was higher than the others.

Keywords: Acquaintances clustering, episodic memory, group of people, name tag, photo indexing, photo management, tag clustering.

1   Introduction

As the use of digital cameras became popular, personal digital photo collections quickly increased in number and size. The increasing number of photos makes it difficult for users to manage and retrieve their collections. As a result, there is a growing demand for tools to help managing, organizing, and browsing those large photo collections.

Various researches were conducted to develop an effective photo management system. Many of them focused on human episodic memory because it is important for people to recall past experiences [7, 32]. Therefore, researchers tried to develop a photo management system that fits with contents and structures of episodic memory. Episodic memory means that memories may be organized by events (episodes). The events include information such as the location of an event, the persons who were present, and what occurred before, during, and after the event. Among this information, the core information of episodic memory is a person [15, 28]. When

people described about their photos, names of people in the photos were the most frequently mentioned information [23]. Also, people's ability to recall the participants of an event decreases quite gradually compared to other information [31]. In addition, a person enabled faster access to past experiences and helped to recall a greater number of total events than the others [3, 25, 26]. Finally, the most frequent content type of photos was a person [13]. Therefore, managing personal photo collections by person information is very important to develop an effective photo management system.

1.1 Related Work

There were many approaches to use person information for managing digital photos. One of the ways to use person information was to tag names of people appearing in photos. Name tagging could significantly improve the usefulness of photo collections [18, 19]. The most commonly-cited benefits were to help recall and support search. However, people have mostly avoided tagging, because it was laborious and the future benefits were unanticipated. In order to support the creation of name tags, various face detection and recognition technologies were used [1, 2, 9, 11, 14]. Consequently, name tagging activities became much easier and faster than before.

In many of the previous works, the name tags were usually presented in alphabetical order [1, 9, 11, 14] or on a first-tag first-place basis [2]. However, the number of people appearing in users' photo collections could range in the hundreds and the span of human memory imposed severe limitations [22]. As a result, it was difficult to manage and search the name tags that a user wished to find. Even worse, there were no relationships between closely placed name tags. Whenever users attempted to search for the name tags of related people, they needed to search the whole list repeatedly. As the human brain operates by association [5], the searching for the related name tags could occur frequently and become one of the reasons to increase the searching time.

In order to solve this problem, we focused on the cognitive strategies of people to reduce the complexity of memory and searching tasks. These are 'chunking' and 'narrowing down'. First, the term chunking means the formation of stimulus elements into subgroups in order to facilitate the assimilation of total information [22]. It is required because the memory span of human has a fixed number of chunks (7 ± 2). In order to break this information bottleneck, people are able to increase the number of information that it contains by building larger chunks. Second, the narrowing down is to use special landmarks or anchors for guiding recall and search [4, 6, 7, 12, 16, 17, 20, 21, 27, 30]. This was frequently observed when people wanted to search for a particular photo. For example, they would first think of an event that was relevant and use it as a guideline to dip into the collection and then move backwards or forwards. Through this process, the actual search-space or browse-space could be relatively limited. By using these two strategies, people could reduce the time for searching and retrieving information.

1.2   Research Problem

Based on our related works, we hypothesized that if a system provides the chunking and narrowing down, it could reduce the time for searching name tags. In order to achieve these, we focused on the intuition that groups of people exist and they are usually exclusive to each other. A group is defined as two or more people who for longer than a few moments interact with and influence one another [29]. We expected that the group could be an effective chunking and narrowing down criteria. We established two main objectives. The first was to develop a name tag clustering method that classifies each name tag into chunks that only hold socially related name tags. The other was to verify whether arranging name tags by their groups could facilitate the task for searching name tags and increase the user satisfaction.

2   System Development

We considered two or more people who meet and interact together as groups based on the Shaw's definition. Our input data were time-stamped photos with name tags of people appearing in them. By using this data, we developed the name tag clustering method that automatically groups the name tags on the basis of their co-attendance at events.

2.1   Step 1: Name Tag Clustering by Co-occurrence

First, we counted the number of co-occurrences (name tags that appear in the same photo) of every pair of name tags. After that, we classified each name tag into groups according to the rule that the top three name tags having the highest co-occurrence frequency be included in the same group. We repeated this process until every name tags satisfied this condition. We could then extract the primary groups.

After the above process, originally two or more groups could be merged into one group because a name tag was able to co-occur with members of several other groups. For example, a user's girl friend could appear in the same photo with his family or co-workers. In this case, his family, co-workers and his girl friend could be merged into a group as she took photos with them. In order to divide this heterogeneous group, we extracted the key name tag which co-occurred most frequently with its group members. Then, we made a list of name tags co-occurring with the key name tag. Next, we computed co-occurrences of each name tag in the list with the heterogeneous group members. Then, if more than 49% of the name tags in the list co-occurred with less than 13% of their group members (the percentages were chosen empirically), we grouped the name tags again by the same rule without the key name tag. If the heterogeneous group was divided into two or more groups, the key name tag was inserted into the group that held a name tag which co-occurred most frequently with the key name tag. Pseudo code was described in Fig 1.

```
Procedure Group classification (total name list, total photo list) {
   Co-table ← Build Co-occurrence table (total name list, total photo list)
   Group list ← Grouping (total name list, Co-table)
   Final Group list ← Regrouping (Group list, total photo list)
}

Sub-Procedure Build Co-occurrence table (total name list, total photo list) {
  For each name from total name list
   Choose photos which contain current name from given total photo list

  For each chosen photo
   Extract tags (=names) without duplication from photo

  For each extracted tag
   Update co-occurrence frequency of [name, tag] pair
}

Sub-Procedure Grouping (total name list, co-occurrence table) {
  For each name from total name list
   Choose the top three names having the highest co-occurrence frequency with name

  For each chosen candidate
  Make group (name, candidate)
}

Sub-Procedure Make group (name, candidate) {
  If name and candidate are not group members then
   Make new group then add name and candidate as member

  If name and candidate are belong to different groups then
   Move one of them to bigger group

  If one of (name and candidate) is belong to certain group then
   Move non-member to certain group
}

Sub-Procedure Regrouping (total group list, total photo list) {
  For each group from total group list
   Choose group-related photo list from total photo list
  Build Co-occurrence table(group members, chosen photo list) to make local table

  If group is separable then
   Remove key name from group who has the highest co-occurrence with its members

  Grouping (current group members, local table)
   Insert 'key name' into the group that includes the highest number of names co-
   occurred with it
}
```

Fig. 1. Pseudo code of our name tag clustering method (Step 1).

2.2   Step 2: Name Tag Clustering by Photo Creation Time

After the group refinement, there could be name tags that do not belong to any group, because they did not have co-occurrence with other name tags. In order to classify these name tags into groups, time-based clustering was conducted. Time clustering detects noticeable time gaps in the creation time of digital photos for identification of events [8]. If a gap is much longer than the local average gap, it is considered a change of event. When an independent name tag appeared in a time cluster containing other name tags, we assumed that the independent name tag attended an event

together with them. We then inserted the independent name tag into the group that held the highest number of name tags appearing in the same time cluster. As time gaps had a very wide range, we used Platt et al. [10, 24]'s time clustering algorithm, because this algorithm adaptively determines the gap between events. After time clustering, there could still be name tags that weren't belonged to any group. It was because they didn't have co-occurrence and didn't appear in a time cluster containing other name tags. These name tags were classified into groups alone.

2.3   Step 3: Arrangement of Groups and Its Members

After name tag clustering, we computed the time that the most recent photo was taken for each group member. We then arranged the most recently taken group at the first place because the most important use of personal photos is looking at recently taken photos and sharing those with friends and family [6]. In each group, members were placed in alphabetical order. This was provided as another means of narrowing down to limit the actual search area.
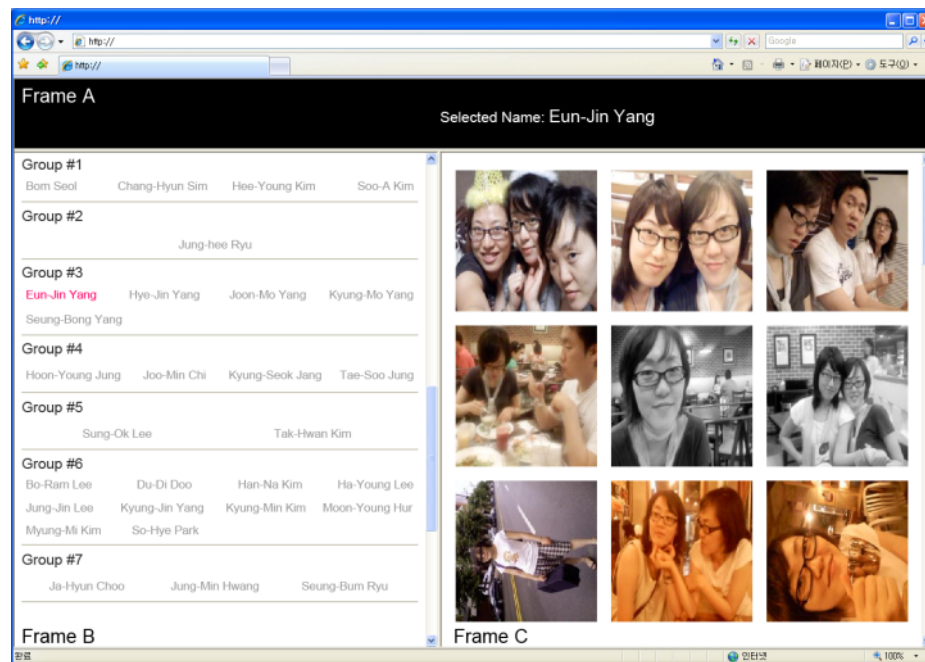

Fig. 2. Prototype photo management system applied our clustering method.

2.4 Prototype Photo Management System

We developed a prototype photo management system applied our name tag clustering, as shown in Fig 2. The system was implemented by using C, HTML and PHP languages. Our user interface consisted of 3 frames. In frame B, the name tags were arranged according to their groups. Groups were distinguished by horizontal lines. Users could navigate the arranged name tags by using the scroll bar. When users selected a name tag, its color was changed from gray to red and it was displayed in frame A. Then, photo thumbnails related to the selected name tag were displayed in frame C. After that, users could browse the extracted photo collections. If users clicked a photo, it was presented by its original size.

3 User Study

We conducted user studies to evaluate the performance of our name tag clustering method and the value of arranging the name tags according to their social relationships (groups of people).

3.1 Subjects

We recruited fourteen subjects who had their own digital cameras and had taken photos for more than four years. These included 6 men and 8 women. Their ages ranged between 25 and 54, with an average age of 30.5 (SD = 7.3). All subjects had normal or corrected-to-normal vision, and had experiences in managing their own digital photo collections.

3.2 Data Acquisition

Before the experiments, each subject was asked to provide their personal digital photos taken during the period from January 2005 to September 2008. Among the vast amount of photo collections, there were some photos that didn't have time stamps. We excluded them from our data set. After that, the number of time-stamped photos ranged from 247 to 5468 photos, the average size being 1739. The names of people appearing in each photo were manually tagged by each subject. In order to help the creation of name tags, we provided a web site. Log-in IDs were allotted to each subject. When they logged-in, their personal photos were presented in chronological order. The creation time of each photo was presented together. In each photo, it had a button for inputting name tags. The number of people appearing in their photo collections ranged from 9 to 132, the average being 55.

3.3 Measures

According to ISO 9241-11[1], usability is defined as effectiveness, efficiency and satisfaction. Effectiveness can be evaluated by accuracy and efficiency by time. At the beginning, in order to evaluate our clustering method by its accuracy, we conducted a questionnaire about how subjects classify their name tags into groups. However, their grouping criteria or levels were very diverse. For example, some subjects classified their father, mother and relatives into one group, 'family'. On the other hand, others classified them into two or more groups, 'family', 'relatives', etc. Therefore, it was hard to evaluate our system with the criterion of accuracy. As a result, we used time and user satisfaction to evaluate our system.
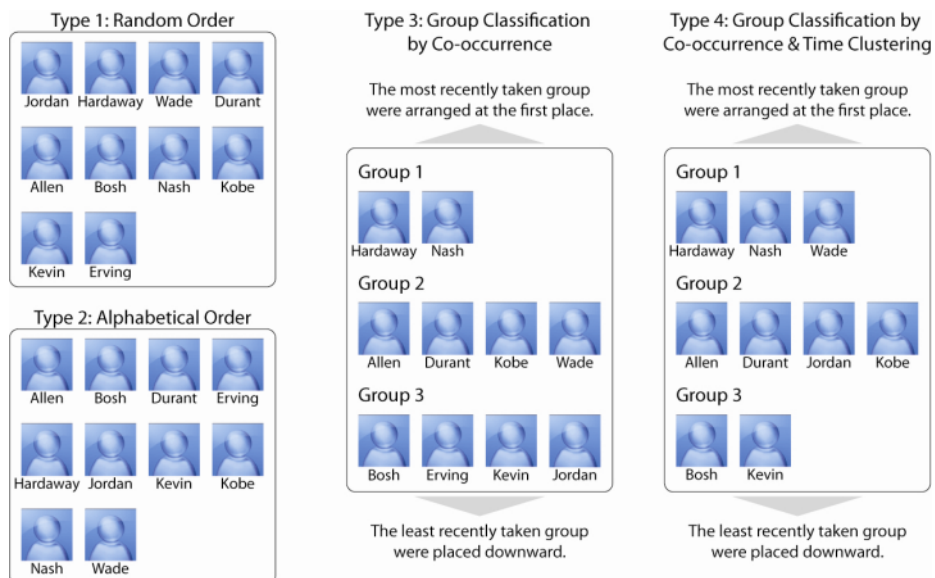

Fig. 3. Differences of each comparison interface.

3.4 Comparison Interfaces

We made three comparison interfaces to verify the performance of our system by measuring time for searching each name tag, as shown in Fig 3. In Type 1 interface, name tags were arranged in a random basis without name tag clustering. Type 2 employed alphabetical order, also without name tag clustering. Type 3 employed a name tag clustering method based only on the co-occurrence of each name tag. For example, if name tags A, B appeared in photo 1 and B, C in photo 2, this classified A,

---

[1] ISO 9241 is a multi-part standard covering a number of aspects for people working with computers.

B, and C as belonging to the same group. Type 4 employed our proposed name tag clustering method (Co-occurrence + Time Clustering).

There were major differences between Type 3 and 4. Type 4 classified name tags into groups by using both co-occurrence and time clustering. By time clustering, therefore, Type 4 was able to classify name tags that always occurred alone in photos into groups. In addition, Type 4 also considered the frequency of co-occurrence. As a result, if name tags of other groups co-occurred in a photo, this could classify them into different groups. On the other hand, Type 3 classified name tags by using only co-occurrence without considering frequencies. Therefore, Type 3 couldn't classify the independent name tags into groups and distinguish other group members co-occurring in a photo. We made Type 3 to compare with our proposed method (Type 4). In both Types 3 and 4, the most recently taken group was placed in the first place and the members of each group were arranged in alphabetical order.

3.5 Experiments

Each subject completed a series of name tag searching tasks on four different interfaces. When one of their name tags was randomly presented on the screen, the subjects navigated the interface, found the indicated name tag and clicked on it. Then, the searching time was recorded and the next name tag was automatically presented. This task was repeated ten times in each interface. The conditions were counter-balanced to avoid learning effects, so that half of the subjects experienced the non-group condition (Type 1, 2) before the grouped condition (Type 3, 4), and the other half experienced the conditions in the reverse order. In order to avoid ordering effects, the sequence of name tags was randomly changed for every pair of subjects.

After completing the task of searching name tags, the subjects were asked to use each interface for some time. Then, we conducted a questionnaire and the subjects rated their degree of satisfaction with each interface on a 7-point Likert scale ranging from 1, very dissatisfied, to 7, extremely satisfied.

4   Result

4.1 Searching Time

The average times for searching name tags were 7.54 sec (Type 1: random order), 4.12 sec (Type 2: alphabetical order), 4.20 sec (Type 3: grouped by only co-occurrence) and 3.37 sec (Type 4: grouped by our proposed method, using co-occurrence and time clustering), as shown in Fig 4. We used an RM-ANOVA on the interfaces and then performed post-hoc analysis using least-significant difference (LSD) to see the difference between each interface. Our system (Type 4) was significantly different with Type 1 ($p < 0.005$) and Type 3 ($p < 0.028$). However, it slightly failed to reach statistical significance compared to Type 2 ($p < 0.065$).
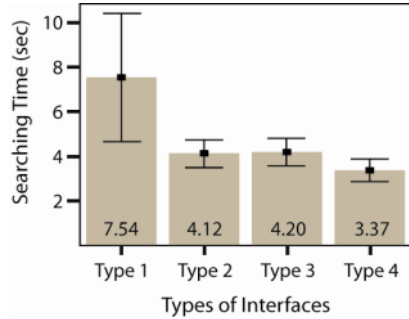
Fig. 4. The average times for searching name tags of each interface. Error bars represent standard errors.

We anticipated that our system would be more effective when the number of name-tagged people was increased. For further investigation, we classified the subjects into two groups (A and B). Group A consisted of the subjects whose collections contained more than 55 people (the average number of people appearing in subject's photo collections). Group B was less than 55. As shown in the Fig 5, when the number of name-tagged people was below 55, there were no significant time differences for searching name tags between each interface. However, when the number was more than 55, the average time was significantly increased in Type 1, 2 and 3. In contrast, there was no significant time increasing in Type 4. The average times for searching name tags were 11.35 sec (Type 1), 5.33 sec (Type 2), 5.35 sec (Type 3) and 3.31 sec (Type 4).
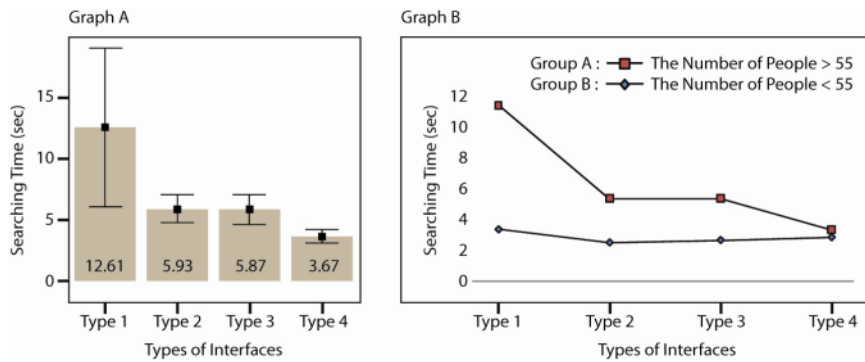
Fig. 5. Graph A - The average times for searching name tags of Group A. Graph B - Time differences for searching name tags between Group A and B.

A repeated measured analysis of variance (RM-ANOVA) with one within-subject factor (Interfaces) and one between-subject factor (Number-of-People) indicated a significant difference between group A and B (Wilks' lambda = 0.853, F=7.011, df=3.000, p-value < 0.05). The RM-ANOVA and LSD was performed only with the data of Group A. In this condition, our system (Type 4) showed significant differences with the other interfaces (Type 1: p<0.007, Type 2: p<0.001, Type 3: p<0.002). As a result, we could find that users were able to find their name tags

significantly faster with our system. Also, it showed a stable searching time regardless of the number of people appearing in user's photo collections.

4.2 User Satisfaction

One-way ANOVA analysis was performed on the satisfaction for each interface. The result was statistically significant with F-value=12.52, $p<0.0001$. The average user satisfaction was 2.64 (Type 1), 4.43 (Type 2), 5.71 (Type3) and 5.86 (Type 4), as shown in Fig 6. The degree of user satisfaction with Type 2, 3 and 4 was significantly higher than Type 1. Especially, it was much higher in Type 3 and 4.
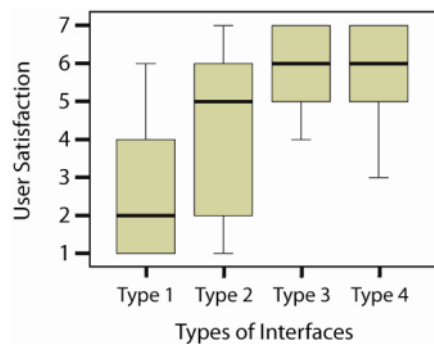

Fig. 6. The degree of user satisfaction with each interface on a 7-point Likert scale.

In addition, our subjects were satisfied with the clustering results. Fig 2 showed the name tag clustering results of our subject 3 (female and 28 years old). For example, her co-workers were classified into group 1. Group 2 was her graduate school professor. Group 3 was her family and relatives. Group 4 was her juniors in design school. Group 5 was administrative staffs of her graduate school. Group 6 was friends of her university. Group 7 was Korean movie stars. She was very satisfied with this clustering result.

5 Conclusion

In order to reduce the complexity of tasks for searching name tags in photo management system, we developed the name tag clustering method that automatically groups those tags based on their social relationships. We also implemented the prototype photo management system applied our clustering method.

After the implementation, we conducted user studies to evaluate our clustering method and the effect of arranging the name tags according to their social relationships. Our system was tested on subjects' own photo collections against three other comparison interfaces. The results showed that subjects could find their name tags significantly faster with our system. Especially, it showed a stable searching time

regardless of the number of people appearing in user's photo collections. In addition, the user satisfaction was higher than the others.

6   Future Work

Events include information such as what, where, when and who. Among them, we firstly used the 'who' information because it was considered the most effective information in photo management. In the future, we are going to improve our system by applying other information (e.g. categorized by location or time) and develop a more refined photo management system.

7   Acknowledgement

References

1. A. Girgensohn, J. Adcock, and L. Wilcox.: Leveraging face recognition technology to find and organize photos. In MIR '04: 6th ACM SIGMM international workshop on Multimedia information retrieval (2004)
2. Apple iPhoto, http://www.apple.com/ilife/iphoto/
3. Barsalou, L. W.: The content and organization of autobiographical memories. In U. Neisser & E. Winograd (Eds.), Remembering reconsidered: Ecological and traditional approaches to the study of memory (pp. 193-243). New York: Cambridge University Press (1988)
4. Brown, N.R., Shevell, S.K., & Rips, L.J.: Public memories and their personal context. In D.C. Rubin (Ed.), Autobiographical memory. Cambridge, UK: Cambridge University Press (1986)
5. Bush, V.: As we may think. Atlantic Mon. 176, 1 (July), 101–108 (1945)
6. David Kirk, Abigail Sellen, Carsten Rother, Ken Wood.: Understanding photowork, Proceedings of the SIGCHI conference on Human Factors in computing systems, April 22-27, 2006, Montréal, Québec, Canada (2006)
7. E. Horvitz, S. Dumais, and P. Koch.: Learning predictive models of memory landmarks, CogSci 2004: 26th Annual Meeting of the Cognitive Science Society, pp. 1-6 (2004)
8. Graham, A., Garcia-Molina, H., Paepcke, A., and Winograd, T.: Time as essence for photo browsing through personal digital libraries. In Proceedings of JCDL 2002, 326–335. ACM (2002)
9. J. Cui, F. Wen, R. Xiao, Y. Tian, and X. Tang.: EasyAlbum: An interactive photo annotation system based on face clustering and re-ranking. Proc. CHI 2007. ACM Press (2007)
10. J. Platt.: Autoalbum: Clustering digital photographs using probabilistic model merging, (2000)
11. Kang, H., & Shneiderman, B.: Visualization Methods for Personal Photo Collections Browsing and Searching in the PhotoFinder. In Proceedings of IEEE International Conference on Multimedia and Expo (1CME2000) New York: IEEE, pp. 1539-1542 (2000)

12. Kemp, S.: An associative theory of estimating past dates and past prices. Psychonomic Bulletin and Review, 6, 41-56 (1999)
13. Kindberg, T., Spasojevic, M., Fleck, R. and Sellen, A.: The Ubiquitous Camera: An In-Depth Study of Camera Phone Use. In IEEE Pervasive Computing 4, 2, 42-50 (2005)
14. Kuchinsky, A., Creech, M., Freeze D., Gwizdka J., Pering C., & Serra B.: FotoFile: A consumer multimedia organization and retrieval system. Proceedings of CHI'99:496-503. New York: ACM Press (1999)
15. Lancaster, J.S., & Barsalou, L.W.: Multiple organisations of events in memory. Memory, 5, 569-599 (1997)
16. Lansdale, M., and Edmonds, E.: Using memory for events in the design of personal filing systems. International Journal of Man-Machine Studies, 36(1):97-126 (1992)
17. Loftus, E.F., & Marburger, W.: Since the eruption of Mount St. Helens has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. Memor and Congnition, 11, 114-120 (1983)
18. M. Ames and M. Naaman.: Why We Tag: Motivations for Annotation in Mobile and Online Media. In Proc. ACM CHI, San Jose, USA, May (2007)
19. Marlow, C., Naaman, M., boyd, d., and Davis, M.: Tagging Paper, Taxonomy, Flickr, Academic Article, ToRead. In Proc. Hypertext '06, ACM Press (2006)
20. Means, B., Mingay, D.J., Nigam, A., & Zarrow, M.: A cognitive approach to enhancing health survey reports of medical visits. In M.M. Gruneberg, P.E. Morris, & R.N. Sykes (Eds.), Practival aspects of memory: Current research and issues, Vol. 1. Chichester, UK: Wiley (1988)
21. Merriam-Webster, Merriam-Webster's Online Dictionary, Merriam-Webster, Inc. (2008)
22. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information, The Psychological Review 63, 81-97 (1956)
23. Naaman, M, et al.: Context Data in Geo-Referenced Digital Photo Collections. In proceedings, Twelfth ACM International Conference on Multimedia (2004)
24. Platt, J.C., Czerwinski, M., and Field, B.A.: PhotoTOC: Automatic clustering for browsing personal photographs. Technical Report MSR-TR-2002-17, Microsoft Research (2002)
25. Reiser, B.J., Black, J.B., & Abelson, R.P.: Knowledge structures in the organisation and retrieval of autobiographical memories. Cognitive Psychology, 17, 89-137 (1985)
26. Ringel, M., Cutrell, E., Dumais, S. and Horvitz, E.: Milestones in time: The value of landmarks in retrieving information from personal stores. To appear in the Proceedings of Interact 2003 (2003)
27. Rodden, K. and Wood, K.: How do people manage their digital photographs? In Proceedings of the ACM Conference on Human Factors in Computing Systems. ACM Press, New York, NY, 409-416 (2003)
28. S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, R. Sarin and D. Robbins.: Stuff I've seen: a system for personal information retrieval and re-use, Proceedings of SIGIR'03 Toronto, ACM Press (2003)
29. Shaw, M.E.: Group dynamics: The psychology of small group behavior. New York: McGraw-Hill (1976)
30. Shum, M.S.: The role of temporal landmarks in autobiographical memory process. Psychol. Bull. 124, 423–442 (1998).
31. Thompson, C. P., Skowronski, J. J., Larsen, S. F., & Betz, A.: Autobiographical memory: Remembering what and remembering when. Mahwah, NJ: Erlbaum (1996)
32. Tulving E.: Episodic memory: from mind to brain. Annu. Rev. Psychol. 53:1–25 (2002)