2010 2nd International Conference on Information and Multimedia Technology (ICIMT 2010)
IPCSIT vol. 42 (2012) © (2012) IACSIT Press, Singapore
DOI: 10.7763/IPCSIT.2012.V42.16

Picture Browser with Face Clustering on Mobile Devices

[1]Zhijie Li, [1]Wendong Wang, [2] Xiaohui Xie, [2]Xiao Zeng and [2]Kongqiao Wang

[1] State Key Lab of Switching & Networking Technology, Beijing University of Posts and Telecommunications, Beijing, China

[2]Nokia Research Center, Nokia(China) Investment CO. LTD, Beijing, China

xiaolanhai228@gmail.com, wdwang@bupt.edu.cn, {xiaohui-kevin.xie, ext-xiao.2.zeng, kongqiao.wang}@nokia.com

Abstract—Classified users' photo album on mobile devices could increase the browsing efficiency, and effective grouping method could greatly improve user experience. In this paper, a novel picture organizing method based on automatic face detection and face clustering is studied and analyzed, a picture grouping method which contains initializing, updating and re-clustering stages is also proposed to increase efficiency and reliability for mobile devices. A sorting method based on correlation is applied to improve user experience of displaying clustering results. User study results show the effectiveness of grouping method and sorting method, which not only increase browsing efficiency but also improve user experience.

Keywords- picture browsers; mobile devices; face detection; face clustering; user study

1. Introduction

Mobile devices with built-in camera bring convenience to capturing pictures. How to group and manage pictures efficiently to enhance browsing is a considerably important research problem. Existing grouping methods tend to be based on time stamp and GPS message of pictures[1~3]. If those metadata cannot be retrieved from picture archives, those methods will be not validated for grouping pictures. Grouping methods based on content or context generally extract features or textures of objects from pictures, such as beach or sunset, and group pictures by those features or textures[4].

However, on mobile devices, portraits take up a high proportion of users' digital album. So researching on face-based grouping method is of significant importance. On desktop PCs, there are some applications introducing face-based grouping technologies, such as Google's Picasa[5] and Apple's iPhoto[6]. All those photo browsers used a semi-automatic method and introduced in manual process. That is not suitable on mobile devices due to its complicated interaction. Face Bubble[7,8] used a two-stepped, automatic grouping method, combined with semi-supervised learning. Despite that this method achieved acceptable accuracy, its complicated steps and computing complexity make it difficult to implement on mobile devices. In this paper, a more simple face-based grouping method is proposed to balance efficiency and accuracy. The method can be described in three stages: initializing, updating and re-clustering.

Meanwhile, it is a valuable question to research on that how to display the results of face detection and face clustering. Down sampling is a traditional thumbnail creation method, which may cause loss of important visual messages. On mobile device, due to its limited display size, the situation is rather worse. Traditional method of displaying grouping results is to display by groups, that is to say each group is represented by a selected picture or face. There are some problems in represented picture's selection. And it may cause some confusion because of disturbing user's browsing. In this paper, a face-based thumbnail creation method is used to emphasize facial area, and to make sure that users can realize facial messages. A method of displaying clustering results by the sorting of correlation is used to highlight the most concerned picture by users.

The rest of the paper is consisted as follow. In section Ⅱ, the paper describes the method of face grouping, including initializing, updating and re-clustering stages. In section Ⅲ a prototype named FaceLink is designed and implemented to show clustering results. The sorting method of correlation is emphasized in this section. In section Ⅳ, a comparable experiment is carried out to prove the effectiveness of the clustering method and results displaying method. And users' browsing interests on mobile devices with face clustering are analyzed and discussed in this section. In section Ⅴ, we summarize our main work done in this paper.

2. Face Clustering Mechanism

From the view point of time selection of grouping process' execution, there are two types of face clustering mechanisms. One is called on-line mechanism, the other is off-line mechanism. The former one employs a method that users browse pictures at the same time as the running of clustering process. The updating of clustering results for display is real-time. The advantage of this mechanism is that it does not affect users' browsing. Whereas the disadvantage is that it does not take full use of computing resource of computer, and also the UI may be more complicated. Desktop Picasa[5] used this mechanism. In contrast, the off-line mechanism applies another method that users cannot access application to browse until the clustering process ends. The disadvantage is that users have to wait for the process, but it can take full use of computing resource to complete the clustering algorithm.

On one hand, mobile devices have limited computing power, so we must take full use of it. On the other hand, popular operating systems of mobile devices provide multi-task supporting, so blocking a single application does not affect other applications. Additionally, if on-line mechanism is applied, due to mobile devices' limited computing power, it may affect users' browsing operation as well as clustering process, and also real-time updating of displaying grouping results may increase the complexity of UI design. So we apply off-line mechanism.

Meanwhile, there are situations that users may add or delete pictures dynamically, or the number of users' pictures in album is not large enough. So some processes like updating and re-clustering are applied to balance efficiency, effectiveness and accuracy. Moreover, the false positive situation is not considered, because it is rare and users can accept a few errors. In additional, the error handling will increase the complexity of UI design which may outweigh the benefits.

2.1 Initializing

The initializing process of automatic grouping method can be divided into there steps(as shown in Figure 1).

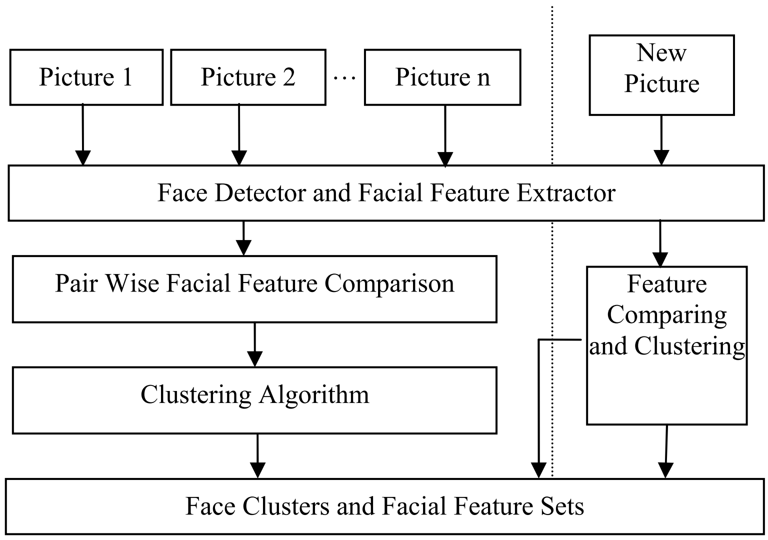

Figure 1. Face clustering framework.

The first step is face detection and facial feature extraction from all users' pictures. An Adaboost-based face classifier[9] is applied to detect faces. Local Binary Pattern(LBP) features[10] are extracted from all detected faces. Every face is described as a feature vector.

The second step is computing of feature similarity. Cosine value of every feature vector pair is calculated to represent their similarity. So a similarity matrix is retrieved.

Finally, a hierarchical clustering method is applied to grouping face clusters. Firstly, every facial feature forms a cluster. Then, the most similar two clusters are selected and merged into one cluster. The merging operation repeats until the similarity value between the two merging clusters falls below a stopping threshold.

2.2 Updating

Once newly added pictures are detected in users' album, updating process is executed. Firstly, all facial features in newly added picture are extracted. Each of the extracted features is compared with all existed facial features. If the similarity value between them exceeds such a threshold, the facial feature is assigned to the existed cluster. If the facial feature is assigned to more than one clusters, all the clusters will be merged together.

The reason for merging clusters is the fact that there will always be some clusters for one person after initializing stage, and for every cluster there are only a few pictures.

It is notable that when users delete a picture, the facial features of this picture is not removed from feature database, and the clustering results will not be updated either. For an obvious reason that more facial feature messages will benefit for clustering accuracy.

2.3 Re-clustering

Certainly, if there are only a few facial features contained in users' album, the initial results of clustering may not be so acceptable. As users' adding pictures and the updating process's execution, the number of facial feature increases, but the final clustering result may not be acceptable either. When the number of facial features reached a certain threshold, a re-clustering process executes. Whether re-clustering process executes or not can be decided by users.

The re-clustering process re-calculates each pair of facial features, repeats the hierarchical clustering method mentioned in initializing stage, and generates new reasonable clustering results.

3. Graphical User Interface

3.1 Design

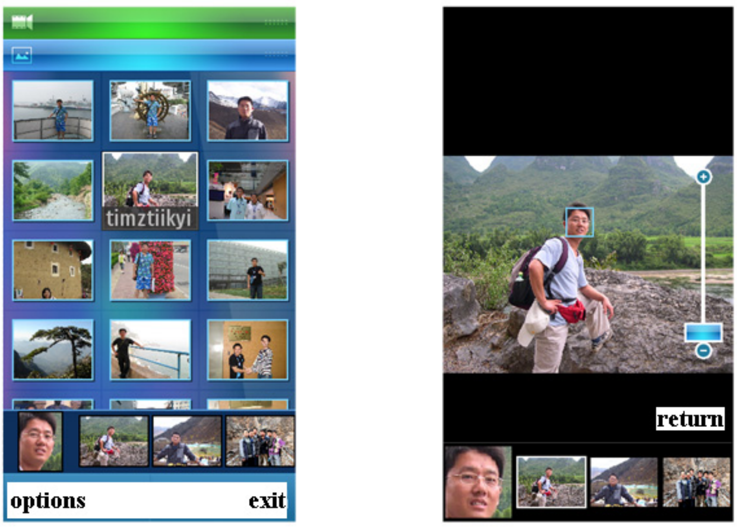

Figure 2. Thumbnails view and detail view.

We designed a prototype named FaceLink to test aforementioned face grouping mechanism and to display clustering results. FaceLink applies a traditional 2D grid thumbnails view to allow users' quick browsing and positioning pictures, and another full-screen view to display the detail of a certain picture. Researchers in [11] proved that 2D grid thumbnails view was simple but could effectively reduce users' mal-operation. Figure 2 shows the thumbnails view and the detail views.

Face-based thumbnails creation method is adopted in this prototype, which was proved to enhance the browsing efficiency in [12]. To accelerate users' searching speed in lots of pictures, a menu entry of filtering portraits from other pictures is provided in options menu.

When a certain picture is focused, all faces appeared in this picture and other pictures relative with the current picture are listed in the bottom. The detail of sorting method of correlation will be demonstrated in the next sub-section.

3.2 Order of Correlation

According to the scope of users' concern, there are two types of situations. The first situation is that users concern about the whole picture. In this situation, all persons appeared in this picture will be assigned the same importance. Other pictures containing any persons same as in the current picture will be listed. The priority of related pictures can be represented by the number of persons appeared in both related picture and current picture. The other situation is that users focus on a certain person in a selected picture. In this situation, the focus person will be assigned high priority. Any other pictures containing this person will be assigned high priority correspondingly.

To make the order of correlation distinct, the priorities of two relation types are describes as follow respectively.

In the first situation, the priority value of relationship between the focused picture P_i and other picture P_j is defined as follow:

$$R(P_i, P_j) = \frac{\#\{F_s \mid F_s \in P_i, F_s \in P_j\}}{\#\{F_s \mid F_s \in P_j\}} \tag{1}$$

where $F_s \in P_i$ means that face F_s appeared in picture P_j.The picture P_j with the larger R value will be arranged at the more front position.

In the other situation, the priority value of relationship between the focused face F_s in picture P_i and other picture P_j is defined as follow:

$$RF(<P_i, F_s>, P_j) = \frac{Include(P_j, F_s)}{\#\{F_k \mid F_k \in P_j\}} * Inf + R(P_i, P_j) \tag{2}$$

where function $Include$ is defined as this:

$$Include(P_i, F_s) = \begin{cases} 1 & F_s \in P_i \\ 0 & Other \end{cases} \tag{3}$$

where Inf is a large enough value to make sure that any picture P_j contains the focus face F_s will have the larger RF value. The order of related pictures in this situation is decided by RF value of other pictures. The picture P_j with the larger RF value will be arranged at the more front position.

4. Usability study

A controlled experiment was set up to test if the proposed clustering method and the displaying method can improve browsing efficiency and increase users' satisfaction. The contrasted browsers are Nokia N97's system gallery application and facial enhanced browser FaceLink. Users' reactions in using such a facial enhanced browser were recorded in the experiment.

4.1 Experiment

The experiment contained two different searching tasks to measure browsing efficiency. After the experiment, the subjective scores of participants were used to represent their satisfaction. The experiment employed a 4×2 within-subjects factorial design, with picture collections and different applications as independent variables. We measured the task completion time as a dependent variable.

We assume that the main types of pictures in users' album are portraits and scenery images. But in different users' album, the proportion of different types and the number of pictures are different. Therefore, four picture sets with different proportions and different numbers were selected from real world mobile phone album or downloaded from the Internet. To make sure that all participants were familiar with all persons appeared in the experiment, all pictures are selected carefully. The detail description of picture sets and the face clustering results are as shown in Figure 3.

We tested ten participants(all males, age: mean 27.3 +- SD 4.2). All participants were relatively experienced computer and touch-screen mobile phone users. Two search tasks were designed for all participants. For the first task, participants were asked to find two scenery pictures for each browsing condition, the target pictures were randomly chosen and were provided by printed papers. For the second task, participants were asked to find out most(around 70%) pictures of a certain person for each browsing condition, the target person was randomly chosen and was described by names. Before the experiment, trainings were applied to make sure all participants were familiar with both applications. They were asked to complete some browsing tasks for the training, and after that give their comments on each picture browser. We tested participants one by one. To minimize the duplicate exposure of pictures, the order of

Set Id	Total Pictures	People Pictures	Detect False Posiive	Detect False Negative	Cluster False Positive	Cluster False Negative
1	40	12	0	0	0	4
2	200	60	5	4	2	6
3	40	32	1	0	2	11
4	200	160	1	12	6	40

Figure 3. Picture sets' description and clustering result.

applied browsers was alternate.Every participant took about 50 minutes. At the end of the experiment, they were asked to score both picture browsers from their subjective satisfaction.

4.2 Quantitative Results

The left and right side of Figure 4 shows the mean task completion time for task 1 and task 2 respectively by comparison between FaceLink and N97's gallery application. As we had expected, for N97's gallery, the task completion times are relative with the number of pictures and number of the same type pictures(portraits or scenery picture), for the fact of large visual differences between different types. But for FaceLink, this phenomenon is not so obvious.

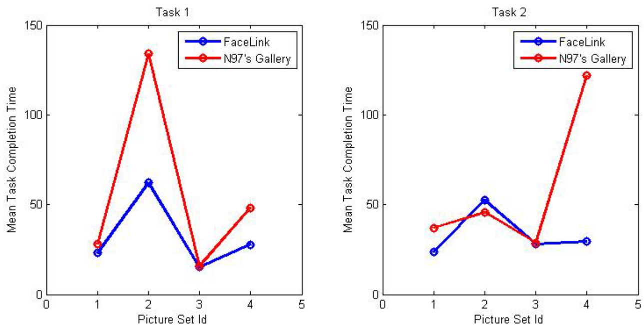

Figure 4. Mean completion time of searching tasks.

To compare the effectiveness of search task in two browsers more precisely, we applied paired t-test for each task and picture set. We observed a statistically significant less time consumption of FaceLink compared with N97's gallery in two browsing conditions for each task. Figure 5 shows the analysis result of visual search tasks, and each subfigure in Figure 6 represents one browsing conditions under which two browsers have significant differences.

In task 1, the task completion times for both picture browsers are shown to be roughly proportional with the picture counts. For the reason that FaceLink filtered portraits from other pictures, in large picture sets(set 2 and set 4), FaceLink was significantly more effective than N97's gallery. In task 2, the situation changed a little bit. The task completion time of N97's gallery seemed to be similar with that in task 1 while FaceLink's effectiveness was very different. Even, with picture sets 2, the mean time consumption of FaceLink is a little more than that of N97's gallery, though it is not statistically significant. We believe it was the difference of face clustering result that impacted FaceLink's browsing efficiency.

Condition	t-Value	p-Value
Task 1 and picture set 2	-3.02	0.01
Task 1 and picture set 4	-2.67	0.03
Task 2 and picture set 1	-2.41	0.04
Task 2 and picture set 4	-3.83	<0.01

Figure 5. Analysis results of visual search tasks(paired t-test).

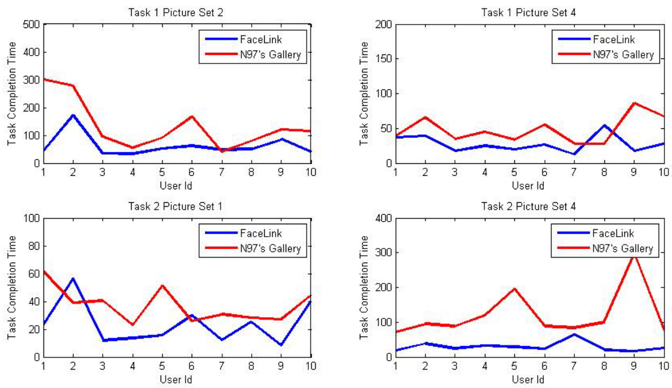

Figure 6. Results of visual search tasks.

4.3 Subjective Satisfaction Results

On a scale of 1-7, FaceLink scored an average 5.4, while N97's gallery scored 3.5. Figure 8 shows the detail of users' scores of both picture browsers. There is only one participant who thought the two browsers satisfied him evenly, while others thought FaceLink should be scored higher. They comment that though some problems existed for FaceLink, such as the loading speed and scrolling speed, it gave them a more convenient and interesting browsing interfac than N97's gallery. Some participants thought they could complete the tasks more quickly if they were more familiar with the interface of FaceLink.

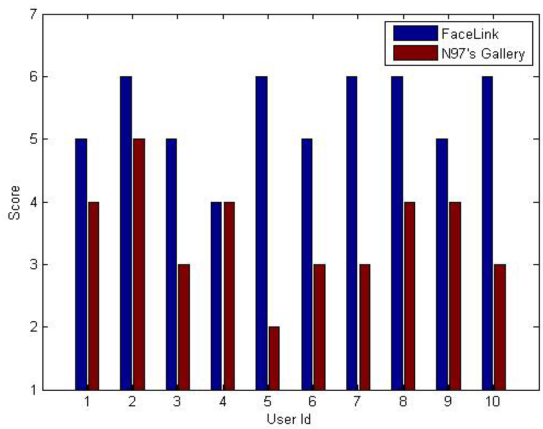

Figure 7. Subjective score of two picture browsers.

4.4 Discussion

We observed some positive phenomenon on using FaceLink to browse during the experiment and the training. They were record as follow:

- All participants tried to use the function of clicking on the face to directly zoom when browsing by FaceLink, which demonstrated its convenience.
- All participants were experienced with the situation that they could not recognize the person by thumbnails in N97's gallery and had to access the archives when asked for searching, but this never happened in FaceLink for the helpful face-based thumbnails creation technology.
- 8 of 10 participants tried and succeeded to use the function of filtering people pictures from others when coping with the searching task, which demonstrated their confidence of face detection technology.
- 7 of 10 participants tried to use the function of link by face to complete the searching task of a certain person.
- When browsing, as some users found a false positive picture in popular stars' cluster, they would glad to say: "Oh, he looks like a super star!" Obviously, they had found the fun of FaceLink.

At the same time, we observed one negative phenomenon. When finding pictures of a certain person by using the face link function, if failed once or twice, they began to complaint the bad performance of face clustering and gave up using the function to search further. So, a more accurate face clustering method should satisfy users better.

In addition, four users suggested a simple interface for navigation by clusters and merging clusters, since they found that lots of pictures of the same person were clustered in different groups. They believed it will be better if they can manually merge clusters with a few operations though it is not a must. They thought they would be patient enough for the pre-computing process before they could access the photo album fluently, because they could just leave it and do something else. They also thought they can bear a few errors and so that error correcting operations are not that important.

From the study, we were glad to see users' confidence with face detection technology, while there should be much work to do with face clustering, both on algorithm and on user interface designing.

5. Conclusion

In this paper, we proposed a face clustering method and a sorting method based on correlation to display clustering results. A prototype named FaceLink was designed and implemented to test the efficiency and usability of proposed methods. Both objective and subjective results of the user study demonstrate the effectiveness and fun of FaceLink.

Since the face clustering results are not so satisfactory and this situation cannot be greatly changed in short time on mobile devices, our future work will focus on UI design to improve users' satisfaction. And we will also attempt to add a view by clusters and evaluate its usability. Some experiments will be celebrated to find the relationship between face clustering results and users' browsing efficiency and satisfaction.

6. Acknowledgment

The work described in this paper is supported by BUPT－NOKIA Innovation UI Design cooperation project, the National Hi-Tech Research and Development 863 Program of China under grant No. 2009AA01Z210 and the EU FP7 Project EFIPSANS (INFSO-ICT-215549). The authors would like to thank all developing team members of FaceLink: Yingfei Liu, Hao Wang and Like Zhu for their hard work and good suggestions. They would also like to give the special thanks to Jing Liu for her planning and organizing of the usability study.

7. References

[1] J.C.Platt, M.Czerwinski, and B.A.Field. PhotoTOC: Auto-matic Clustering for Browsing Personal Photographs, Tech-nical Report MSR-TR-2002-17, Microsoft Research, 2002.

[2] M. Coorper, J.Foote, A.Girgensohn. Temporal Event Clustering for Digital Photo Collections. ACM Transactions on Multimedia Computing, Communications and Applications, Vol.1, No.3, Aug. 2005, New, York, USA.

[3] Y.Orii, T.Nozawa, T.Kondo. User Study of Automatic Photo Classifier by Color and Timestamp. IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology. Sep. 2009.

[4] R. Datta, D.Joshi, J. Li, and J.Z. Wang. Image retrieval: Ideas, influences, and trends of the new age, ACM Computing Surveys(CSUR), Apr. 2008, 1-60.

[5] Google Picasa. http://picasa.google.com.

[6] Apple iPhoto. http://www.apple.com/life/iphoto.

[7] J. Xiao, T. Zhang. Face Bubble: Photo Browsing by Faces. ACM AVI'08, May. 2008, Napoli, Italy.

[8] T. Zhang, J. Xiao, D. Wen and X.Q. Ding. Face Based Image Navigation and Search. ACM MM'09, Oct.2009, Beijing, China.

[9] Jones, M. , Viola,P. Fast Multi-view Face Detection. TR2003-96, Aug. 2003.

[10] T. Ahonen, A.Hadid and Matti,Pietikäinen. Face recognition with local binary patterns. in Proc. 8th European Conference on Computer Vision, ser.Lecture Notes in Computer Science, vol. 3021. Springer, 2004, pp. 469-481.

[11] T.T.A.Combs and B.B.Bederson. Dose Zooming Improve Image Browsing?, ACM DL99, Aug.1999, Berkeley, CA USA.

[12] B.Suh, H.B. Ling, B.B. Bederson, and D.W. Jacobs. Automatic Thumbnail Cropping and its Effectiveness. ACM UIST'03, Nov.2003, Vancouver, BC, Canada.