

USER-ASSISTED IMAGE CLASSIFICATION ON PERSONAL PHOTO COLLECTIONS

Wei-Ta Chen, Ying-Hsiang Wen and Ming-Syan Chen

Department of Electrical Engineering
National Taiwan University

E-mail: {weita, winshung}@arbor.ee.ntu.edu.tw, mschen@cc.ee.ntu.edu.tw.

ABSTRACT

Image classification on personal photo collections can be extremely useful to various management tasks. However, there is little progress made forwarding it due to (1) lack of training data, and (2) subjectivity inherent in a user's photo-organizing behavior. In this paper, we propose a framework, User-assisted image Classification on Personal Photo Collections (UCP), to address this problem. The uniqueness of this framework is that it is user-centric and includes users in the loop. Our experimental results show that the techniques used in this framework are promising.

1. INTRODUCTION

With the widespread use of digital cameras and camera phones, users are amassing large collections of personal digital photographs. It becomes increasingly important to devise efficient and effective methods that help users to browse, search, retrieve or annotate photos on their personal collections [1], [2], [3]. Image classification has been recognized as a useful technique to support the various management tasks in a large general image repository [4], [5]. However, there is little progress made forwarding image classification on personal photo collections due to (1) lack of training data, and (2) subjectivity inherent in a user's photo-organizing behavior. Before further explaining these two difficulties, we give a few examples to explain why image classification is useful to management of personal photo collections.

First, consider an example of categorizing photos taken on a trip. It is common that a person takes lots of photos in a trip, and so do the rest persons in the same trip. Moreover, people usually exchange their photos after the trip. In this case, if the first user constructs an image classifier based on the photos he/she took, the subsequent photos he/she obtains from others can be categorized easily according to the classifier. Next, consider another example that a user constantly takes photos which belong to a set of categories. A classifier built based on the prior photos can help the user organize the subsequent new photos.

Classification is also referred to as supervised clustering. Compared with unsupervised clustering, classification aims

to obtain a classifier from a set of training data, which can be used to categorize subsequent new images as those mentioned in the above examples. Unsupervised clustering, due to lack of a classifier, cannot complete the task. In addition, as will be explained next, user feedback is extremely important to management of personal photo collections. User feedback about the structure of a classifier can make it easier for a classifier to get each individual's ground truth categorization. In contrast, there is no direct extension to equip an unsupervised clustering method with feedback mechanisms. Therefore, techniques of image classification is more applicable to these examples than those of image clustering [1], [2], [3].

We now explain the obstacles towards image classification on personal photo collections. First, there is no any set of training data that is able to cover various scenes of photos taken by users. In addition, even if we consider only a particular photo collection, class identities of photos in it are not known a priori. Therefore, the assumption of a traditional classification problem that we can restrict to a set of classes and that each photo in the training set has a known class label is not applicable to the case. Another major distinction in this case is that we are dealing with "*personal*" photo collection, which implies the subjectivity about the way an owner wants his collection to be categorized has to be taken into consideration. That is, a classification method suitable to the case must be capable of accepting user feedback so as to fulfill each individual's classification requirement.

We overcome these problems by devising a framework, User-assisted image Classification on Personal photo collections, abbreviated as UCP, in this paper. As shown in Figure 1, UCP includes users in the loop and uses both unsupervised and supervised clustering methods. In the framework, a user first selects a set of photos from his collection as a training set, and is asked to mark k photos in the set as seeds of representative photos in k classes. An unsupervised clustering method, PAM [6], is used to cluster photos into k clusters. We then assign different labels to the resultant clusters, which in turn makes the training data divide into k classes. Next, a modified decision tree approach [7] is used to construct a classifier. Compared with other classification methods, such as the SVM [4], or Bayesian classifier [5], a decision tree is easily understood by users. In addition, we visualize clas-

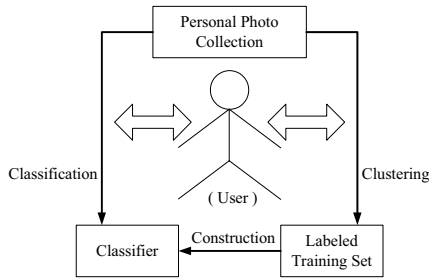


Fig. 1. The proposed UCP framework.

sification results by associating each tree node with a best representative image. The easily interpretable classifier structure and visualization of classification results enable users to give feedback to our decision tree classifier, which is uniquely equipped with feedback mechanisms.

The supervised and unsupervised clustering techniques in UCP can easily adapt to various image features as long as distance measures for such features are defined. It has been recognized that there is a semantic gap between the low-level visual features and the high-level semantics of images [8]. Personal photos, however, are usually associated with important metadata information about "when" a photo was taken, which has been shown to be a powerful memory cue that enables people to interact with their photo collections [9], [10]. Therefore, we use both metadata and visual features to represent a photo in our experiments.

For a literature review, we have seen an increasing amount of research in the management of personal photo collections due to the advance of acquisition technology in recent years. However, most proposed techniques focus on image clustering [1], [2], [3], especially through metadata information as is explained earlier. Although image classification in general image repositories has been extensively explored [4], [5], to our best knowledge, there are no prior works addressing the problem of image classification on personal photo collections despite its importance. This fact distinguishes this paper from others.

The remainder of this paper is organized as follows. In Section 2, we introduce the proposed framework. We present the experiment results in Section 3, followed by our conclusions and future work in Section 4.

2. DESIGN OF THE UCP FRAMEWORK

In Section 2.1, we describe the way UCP obtains class identities of photos in the training set through user assistance. We introduce the proposed decision tree classifier in Section 2.2, and explain the way we visualize classification results and accept user feedback in Section 2.3.

2.1. Clustering

In UCP, a user first selects a set of photos from his collection as a training set. Instead of asking a user to label each photo in the training set, which will be tedious when there are a lot of photos in it, UCP obtains class identities of photos by asking a user to mark k representative photos. These k representative photos are used as seeds for k individual classes, respectively. Then, the PAM [6] clustering method, which is robust in the presence of noise and outlier, is used to cluster photos into k groups. Given this preliminary result, a user can either adjust group members or mark a new set of k representatives to obtain a better clustering result according to his/she subjective preference. In the words, we use the unsupervised clustering technique to help a user label photos in a training set, and the whole process can be iterated until a clustering result that is well matched with a user's preference is obtained. UCP then assigns different labels to the resultant clusters, which in turn makes the training data divide into k classes.

2.2. Classification

We next introduce the proposed decision tree classifier in UCP. As mentioned earlier, a decision tree is easily understood by users, which makes it more suitable than any other classifiers to support user feedback. However, the traditional decision tree approach is only applicable to data whose attributes are all categorical. In our case, we represent photos by their visual content and metadata, which are mostly not categorical. Therefore, we use the PAM method to categorize features so as to cope with the decision tree growing process.

The growing algorithm of our modified decision tree is shown in Figure 2. The growing process is to recursively expand a tree node until a stop criterion is satisfied. Take the expansion at the root for example. Let n_0 denote the root and S_{n_0} denote the set of training images associating with n_0 . We expand n_0 by using the PAM algorithm to partition S_{n_0} against each image feature $d \in D$, where D is the set of features extracted. An attribute denoted by $d(n_0)$ that produces the most information gain is selected as the test attribute of n_0 . According to this test attribute $d(n_0)$, S_{n_0} is partitioned into S_1, \dots, S_q . New nodes n_1, \dots, n_q for S_1, \dots, S_q are created and added to the decision tree as the children of n_0 . After expanding the root, the same expanding procedure is recursively applied to each new node n_i unless n_i satisfies a stop criterion. Note that the PAM algorithm also generates a medoid image M_i for each partition S_i , where $1 \leq i \leq q$. Each new node n_i is also associated with a representative feature $r(n_i) = M_i(d(n_0))$, which is selected to stand for the representative feature of $d(n_0)$ of images in S_i . When we classify a new image, $r(n_i)$ is used to indicate how a new image traverses the tree to a leaf.

Algorithm DTG: DTG(S, D, f, ρ)
Input: a set S of training data, a set D of attributes, a representative feature f , and the purity threshold ρ ;
Output: a decision tree T ;

1. $n \leftarrow \text{MakeTreeNode}()$;
2. Let T denote the tree rooted by n ; $l(n)$ the class label of n ;
3. $S_n \leftarrow S$;
4. **if** ($f \neq \varphi$) **then** $r(n) \leftarrow f$;
5. **if** ($\text{purity}(S_n) \geq \rho$) **then**
6. $l(n) \leftarrow \text{MajorityLabel}(S_n)$;
7. **return** T ;
8. Let q be the number of branches which is equal to the number of classes covered by S_n ;
9. **for each** $d_i \in D$
10. $\text{Partitions}(d_i) \leftarrow \{S_1, \dots, S_q\} = \text{PAM}(S_n, d_i, q)$;
11. $g(d_i) \leftarrow \text{InformationGain}(\text{Partitions}(d_i))$;
12. $d(n) \leftarrow \text{argMax}_{d_i \in D} \{g(d_i)\}$;
13. **for each** $S_i \in \text{Partitions}(d(n))$
14. $I_M \leftarrow \text{Medoid of } S_i$;
15. $f_i \leftarrow I_M(d(n))$;
16. Insert a subtree $T_i = \text{DTG}(S, D, f_i, \rho)$ under n ;
17. **return** T ;

Fig. 2. The growing algorithm of the proposed decision tree.

2.3. Visualization and User feedback

The structure of our decision tree can be deemed as the skeleton that is capable of visualizing classification results in a tree hierarchy. We visualize the classification results by assigning a best representative image to every node but the root. The selection of best representative images are as follows: When a set U of new images are classified by the decision tree, each classification corresponds to traversing a path from the root to a leaf. Let $U_n \subseteq U$ denote the subset of new images falling into the subtree rooted by n . An image in U_n whose corresponding feature is the most similar to the representative feature $r(n)$ in n is selected as the best representative image of node n . The process of selecting best representative images can be incremental and completed along with each classification of a new image

In this way, a user can traverse the tree to view classification results. When a user visits an internal node n_i , UCP shows the user the collection of those best representative images in n_i 's child nodes. In contrast, when a user visits a leaf, UCP shows the user new images falling into this leaf. Moreover, we also show a user the splitting attribute or the class label associated with a node. Based on the information, users can easily interpret the classification results and give feedback to the classifier.

UCP is able to support the following user feedback: (1) subtree-pruning, and (2) subtree-growing. For the first operation, a user may give feedback to the classifier where there is no need to further expand a node n because doing so will either misclassify new images or over-classify the semantic concepts. The case of over-classification occurs when the subtree rooted by n divides images which should be simply grouped as one category into several sub-categories or when one of its branches has no associated best representative im-

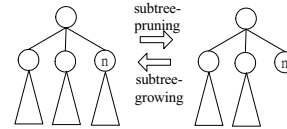


Fig. 3. The operations of subtree-pruning and subtree-growing.

Table 1. The Data Sets Used In Our Experiments

Collection	Number of Photos	Number of Classes
C ₁	694	23
C ₂	552	15
C ₃	520	18

age. We adjust the classifier to this feedback by pruning the corresponding subtree, i.e., mark n as a leaf. This new leaf n will then be associated with the class label of the majority in S_n . The upper arrow in Figure 3 illustrates the operation of subtree-pruning.

For the second operation, a user may give feedback to the classifier where further expanding a node n is needed because new images falling into n are misclassified or the semantic concept of images in n is too coarse. We adjust the classifier to this feedback by enhancing the purity threshold ρ of n to further expand n . That is, we apply Algorithm DTG with a higher purity threshold to grow a subtree for the training set S_n , with which the node n is associated. The lower-arrow in Figure 3 illustrates the operation of subtree-growing.

3. EXPERIMENT RESULTS

In this section, we evaluate the performance of UCP. As mentioned before, there were no prior works on image classification on personal photo collections. Therefore, we separate the evaluation of UCP into the clustering phase and classification phase. The experiments are conducted on a computer with Intel P4 2.8 GHz CPU and memory of 1GB.

Due to the lack of commonly adopted personal photo collections for performance evaluation, i.e., benchmark data sets, we compiled three personal photo collections from three contributors ourselves. Each collection contained hundreds of photos, all of which had timestamp information in the EXIF header. We asked each owners to manually divide his photos into a number of classes. Details of the collections are given in Table 1. We extracted both metadata and visual features to represent photos. The metadata of a photo's taken time was extracted. Color and texture extraction methods proposed in [11] were used to extract visual content.

3.1. Evaluation of the clustering phase

We compare the proposed user-assisted clustering technique with a trivial approach, in which a user labels every photo

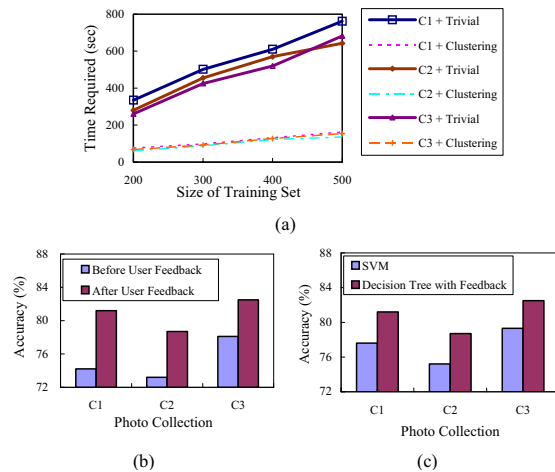


Fig. 4. Experiment results: (a) the time required to obtain class identities, and (b) classification accuracy.

manually. We randomly selected 200, 300, 400, and 500 photos from each collection and asked the owner to classify them by (1) the proposed approach, and (2) the trivial approach. Note that the proposed approach also allows users to manually adjust the clustering results and it terminates when users are satisfied with the results. Therefore, the qualities of the results obtained by the two methods are comparable in terms of a user's preference.

Figure 4 (a) shows the time required to label photos by the owners when the proposed or manual approach is used. The time required by the manual labeling approach is much longer than that by the proposed approach, especially when the size of a training set is large. Therefore, our semi-automatic approach can significantly reduce the time required while achieving comparable clustering results.

3.2. Evaluation of the classification phase

We compare the accuracy of our decision tree approach with that of the SVM-based approach [4]. First, in order to obtain the ground truth about class identities of photos, all photos on each collection were labeled in the clustering phase. Then, we randomly selected 80% of photos as the training set and the rest 20% as the test set. We first evaluate the performance of the feedback mechanism, with which our classifier is uniquely equipped, in Figure 4 (b). When the feedback mechanism is used, the accuracy is substantially enhanced. Therefore, our feedback mechanism is able to improve the performance. Then, the comparison with the SVM-based classifier is shown in Figure 4 (c). The result shows that the proposed classifier outperforms the SVM-based classifier because it is able to accept user feedback. However, we notice that regardless of the classifier used, the overall classification accuracy is not high. We reckon that this is due to the diversity of classes on

personal photo collections, and better representation methods about visual or metadata features have to be devised to alleviate this problem.

4. CONCLUSIONS

We have presented the proposed framework UCP, which aims to address the problem of image classification on personal photo collections. UCP overcomes the inherent difficulties in this problem by including users in the loop to label photos and to give feedback to the proposed decision tree classifier. Our experimental results showed that the techniques used in UCP are superior to their individual comparison methods.

Acknowledgement

The work was supported in part by the National Science Council of Taiwan, R.O.C., under Contracts NSC93-2752-E-002-006-PAE.

5. REFERENCES

- [1] M. Cooper, J. Foote, A. Girgensohn, and L. Wilcox, "Temporal event clustering for digital photo collections," in *ACM MM*, 2003, pp. 364–373.
- [2] M. Naaman, Y. J. Song, A. Paepcke, and H. G. Molina, "Automatic organization for digital photographs with geographic coordinates," in *JCDL*, 2004, pp. 53–62.
- [3] A. Pigeau and M. Gelgon, "Building and tracking hierarchical geographical & temporal partitions for image collection management on mobile devices," in *ACM MM*, 2005, pp. 141–150.
- [4] K. Goh, E. Chang, and K. Cheng, "SVM binary classifier ensembles for image classification," in *CIKM*, 2001, pp. 395–402.
- [5] A. Vailaya, M. A. T. Figueiredo, A. K. Jain, and H.-J. Zhang, "Image classification for content-based indexing," *IEEE Trans. on Image Processing*, vol. 10, no. 1, pp. 117–130, Jan. 2001.
- [6] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, New York: John Wiley & Sons, 1990.
- [7] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.
- [8] S. Antani, R. Kasturi, and R. Jain, "A survey on the use of pattern recognition methods for abstraction, indexing, and retrieval of images and video," *Pattern Recognition*, vol. 35, no. 4, pp. 945–965, 2002.
- [9] M. Naaman, S. Harada, Q.-Y. Wang, H. G. Molina, and A. Paepcke, "Context data in geo-referenced digital photo collections," in *ACM MM*, 2004, pp. 196–203.
- [10] K. Rodden and K. R. Wood, "How do people manage their digital photographs?," *ACM CHI*, vol. 4, pp. 409–416, Apr. 2003.
- [11] Y. Rubner, C. Tomasi, and L. J. Guibas, "A Metric for Distributions with Applications to Image Databases," in *ICCV*, 1998, p. 59.