

Social Photo Tagging Recommendation Using Community-Based Group Associations

Chien-Li Chou¹, Yee-Choy Chean¹, Yi-Cheng Chen¹, Hua-Tsung Chen² and Suh-Yin Lee¹

¹Department of Computer Science
National Chiao Tung University
Hsinchu, Taiwan

²Information & Communications Technology Lab
National Chiao Tung University
Hsinchu, Taiwan

e-mail: fallwind@cs.nctu.edu.tw, {ned.cs98g, ejen.cs95g}@nctu.edu.tw, {huatsung, sylee}@cs.nctu.edu.tw

Abstract—In the social network, living photos occupy a large portion of web contents. For sharing a photo with the people appearing in that, users have to manually tag the people with their names, and the social network system links the photo to the people immediately. However, tagging the photos manually is a time-consuming task while people take thousands of photos in their daily life. Therefore, more and more researchers put their eyes on how to recommend tags for a photo. In this paper, our goal is to recommend tags for a query photo with one tagged face. We fuse the results of face recognition and the user's relationships obtained from social contexts. In addition, the Community-Based Group Associations, called CBGA, is proposed to discover the group associations among users through the community detection. Finally, the experimental evaluations show that the performance of photo tagging recommendation is improved by combining the face recognition and social relationship. Furthermore, the proposed framework achieves the high quality for social photo tagging recommendation.

Keywords—social network; social context; face recognition; photo tagging recommendation

I. INTRODUCTION

With the affordable price and ubiquitous presence of digital capturing devices, more and more photos are captured in our daily life. People can take thousands of pictures anywhere through their digital cameras, cell phones and even tablet PCs. Nowadays, social network becomes a popular platform for people to interact with friends. Users upload photos onto the social network not only for keeping their own memory but also sharing their experiences. In addition, users can tag the people appearing in a photo with corresponding names so that the system can notify the tagged users and share photos with them. Photo management and search become much easier by utilizing the tags. However, tagging all uploaded photos is a time-consuming task for users. Hence, if tags can be recommended by the system, users can accomplish the tagging much quickly. For these reasons, many researchers have investigated in the field of social photo tagging recommendation. Fig. 1 shows the traditional framework of photo tagging recommendation.

For tagging recommendation, using face detection and recognition to identify the people in a photo is a

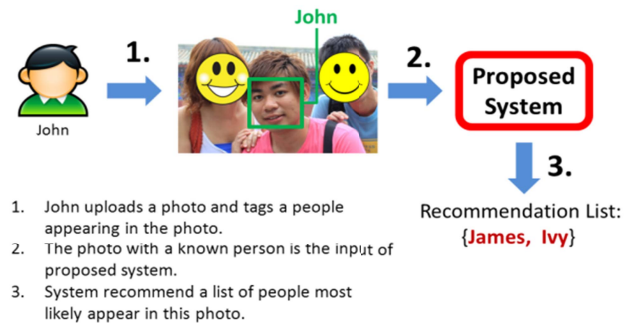


Figure. 1. Traditional framework of face tagging recommendation.

straightforward method. However, with the increasing number of people, the saliencies of faces are insufficient to distinguish the difference among the people in the photo. That is, the accuracy of face recognition decreases when the number of people increases. Besides, uncontrolled situations such as ambient illumination and capturing angle of photos are also the bottlenecks for practical face detection and recognition algorithms. Furthermore, many people use the graphic editing program, such as Photoshop, to make the photos look better before uploading. The post-producing process makes face recognition more difficult.

In general, to recognize the people in a photo, regarding many aspects including their demographic description, we can quickly identify them if they are familiar to us. Some questions related to the activities of, emotional states of, and relationships between people in a photo can be answered by humans. In other words, we draw conclusions based on not only what we see, but living experience and relationship with other people. In real world, it is difficult to quantize the relationship among people. Benefit from the rapid development of social network, the relationships between people can be retrieved from their interactions on the social network. For instance, if Alice and Bob usually appear in photos simultaneously, we can say that Bob is close with Alice. Afterward, when Bob uploads a new photo, the probability for Alice appearing in that photo is high. This kind of information becomes a complementation of low-level features when we recognize a face in a photo.

Based on the observations, we combine the face recognition technique and the relationship among people to

discover who is more possible to appear in the photo. To analyze the relationship, we collect two social contexts, photo co-occurrence and mutual friends. The co-occurrence relation and mutual friend relation are obtained from the social graph model constructed by social contexts. When we recommend the tags for the query faces in the query photo, we consider not only the relationship between the tagged face and the query face, but also that among all the query faces. With this aspect, the *Community-Based Group Associations*, called *CBGA*, is proposed to analyze the group relationship. According to the community detection algorithm, the users who interact more with each other are clustered together. The empirical evaluations on the collected real social data reveal that the proposed method is very promising to social face tagging recommendation in terms of precision.

The remaining of this paper is organized as follows. The related work is reviewed in Section 2. In section 3, we present the proposed approach in great detail. The experimental evaluations are described and discussed in Section 4. Finally, the conclusion and future work are stated in Section 5.

II. RELATED WORK

Social network is a social structure made up of people, which are connected by socially-meaningful relationships. For sharing the photo with each other on social network, users tag the faces appearing in the photo with their names. It can not only help the system recommend this photo to the users who we want to share with but also provide a good way to manage and search photo album easily afterwards [9]. Moreover, analyzing the photo tags is useful for both friend recommendation [10, 16] and photo tag recommendation techniques [1, 12, 14].

Some studies recommend the tags for the photos only by the image features. Face recognition techniques [18] are usually the basic approach for tag recommendation. Zhang et al. [17] adopted the Bayesian statistical model and the maximum a posteriori (MAP) estimation to annotate human faces in family albums. Choi et al. [3] proposed the situation clustering algorithm to cluster the photos under the same situation. In each situation cluster, the photos with the same subject are clustered with the similar face feature and clothing feature by subject clustering. Then, the face recognition based on face information and learning is performed to annotate faces in photos.

In consumer photos, Gallagher and Chen [5, 6] improve the face recognition by learning the prior probability of different individuals appearing together in a photo. In addition, analysis of the context in an album is used to improve the recognition as well [7]. The contexts include the information related to the photo of the scene surrounding the person, digital camera context such as location and photo capture time and the interactions among people. In social networking sites, social context can be analyzed to obtain the relationships among people. Rae et al. [8]

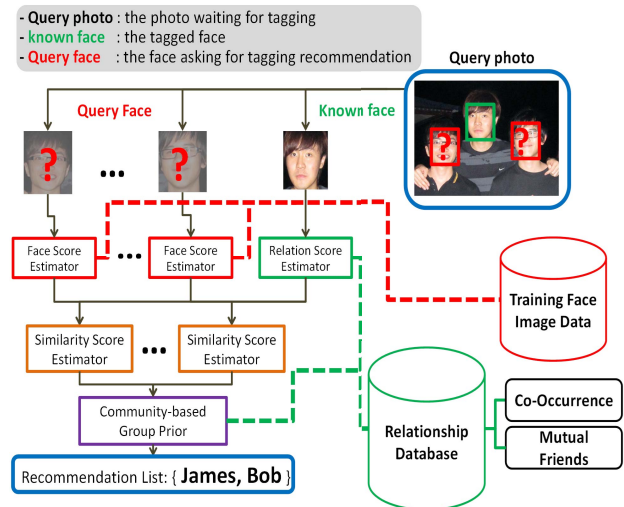


Figure 2. Framework of the proposed photo tagging recommendation method.

proposed the framework of partial tag recommendation through four kinds of context information. Choudhury et al. [4] demonstrated that the combination of low-level image feature and social community improve the performance of photo tag recommendation on Flickr.

Stone et al. [13] demonstrates a simple method to enhance face recognition with social network context. The goal of this work is to infer a joint labeling of face identities over all identities in social network by applying a pair-wise conditional random field (CRF). To reduce the computational cost and improve the performance of annotation, Choi et al. [2] proposed a collaborative face recognition framework for face annotation in social network. Multiple face recognition engines are constructed for each local network. Further, the social contexts in a local network are considered to select the suitable face recognition engine.

III. PROPOSED APPROACH

In this paper, we assume that there is one tagged face, called *known face*, in the query photo. The other detected faces are *query faces*. The framework of our proposed system is shown in Fig. 2. First, we estimate the relation score of known face from social contexts by Relation Score Estimator (RSE). The face scores of query faces are computed by Face Score Estimator (FSE). Then, the two scores are aggregated into a similarity score by Similarity Score Estimator (SSE). Finally, the Community-Based Group Associations (CBGA) is used to discover the group relationship.

A. Face Score Estimator (FSE)

We implement the eigenface method [15] to perform face recognition. The training faces in this paper are extracted from training photos. Thus, the number of training faces of each user may be different from each other. We use Euclidean distance to calculate the distance between two

eigenfaces. The face score f for a query face q and a user u_i can be defined as

$$f(q, u_i) = 1 - \left\| \min_{1 \leq j < n} d(q, e_i(j)) \right\|, \quad (1)$$

where $d(q, e)$ denotes the distance between the eigenface of q and the eigenface e , $e_i(j)$ is the j^{th} eigenface of user u_i , and n is the number of user u_i 's training faces. $\|\cdot\|$ is the normalization operation.

B. Relation Score Estimator (RSE)

In social network, the data used to obtain the relationships between users is called ‘‘social context’’. Two kinds of social contexts are used in this paper: the number of photo co-occurrence and the number of mutual friends. Given a social network S contains M users, $S = \{u_1, u_2, \dots, u_M\}$, and a photo collection set $P = \{p_1, p_2, \dots, p_N\}$, we construct the social graph model for each social context. The social graph model of G^t can be defined as $G^t = \{U^t, E^t, W^t\}$, where U^t is the set of nodes and each node is a user in social network S , E^t is the set of edges between any two nodes, and W^t is the set of weights on edges.

1) Co-Occurrence Relation (COR)

In general, the more co-occurrence between two persons, the closer the relationship between them. If a photo contains Alice and Bob, we say they co-occur in the photo. Therefore, we can construct the social graph model G^c from the co-occurrence relation with U^c , where U^c is the set of all users in the social network. The weight of the edge between u_i and u_j in G^c can be defined as

$$w^t(u_i, u_j) = \sum_{k=1}^N \delta_k(u_i, u_j), \quad (2)$$

where N is the number of query photos and

$$\delta_k(u_i, u_j) = \begin{cases} 1, & u_i \text{ and } u_j \text{ are in photo } p_k; \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

2) Mutual Friend Relation (MFR)

The number of mutual friends is another way to estimate the relationship among people. If two people have the more mutual friends, it is more possible for them to be in the same social circle of real life. They may study in the same school or work in the same company. Thus, they have more chance to take photo together. As co-occurrence relation, we can also construct a social graph model G^m from mutual friend relation with U^m , where U^m is the set of all users in the social network. The weight of the edge between u_i and u_j in G^m can be defined as

$$w^m(u_i, u_j) = FL_{u_i} \cap FL_{u_j}, \quad (4)$$

where FL_{u_k} is the set of the friends of user u_k .

3) Relation Score Combination

To calculate one relation score, we combine two social graph models into one model G^R since the nodes and edges are the same in them. When we combine the two models, the weights of edges should be combined as well. We perform the linear combination for COR and MFR. Thus, the weight of combined social graph can be defined as

$$w^R(u_i, u_j) = \alpha * \left\| w^c(u_i, u_j) \right\| + (1 - \alpha) * \left\| w^m(u_i, u_j) \right\|, \quad (5)$$

where α is the parameter to adjust the specific weight between COR and MFR, $0 < \alpha < 1$, and $\|w\|$ denotes the normalization of w . The normalization can be defined as

$$\left\| w^m(u_i, u_j) \right\| = \frac{w^t(u_i, u_j) - w_{\min}^t}{w_{\max}^t - w_{\min}^t}, \quad (6)$$

where w_{\min}^t and w_{\max}^t are the minimum and maximum value in W^t , respectively. The relation score r between two users, u_i and u_j can be defined as

$$r(u_i, u_j) = w^R(u_i, u_j). \quad (7)$$

C. Similarity Score Estimator (SSE)

In this section, we describe how to aggregate the face score and the relation score into similarity score (SS). For estimating the similarity score $S(q, u_i)$ for the query face q and the user u_i . We combine the face score f and relation score r by the formula as following.

$$S(q, u_i) = \beta * f(q, u_i) + (1 - \beta) * r(u_{\text{known}}, u_i), \quad (8)$$

where β is the parameter to adjust the specific weight between face score and relation score and $0 < \beta < 1$. Note that the importance of relation score becomes higher as β decreases. Let us take an example based on Fig. 3, 4, and 5. Assume that the social network contains 10 users with their training face images, photo collections and friend lists. Fig. 3 shows a query photo Q consisting of the known user, u_{known} , identified by the known face, and query faces, q_1 and q_2 . The goal is to recommend the candidate tags for the query faces, q_1 and q_2 . Let us take q_1 as an example.

In offline training phase, we train the eigenfaces from the training face images. In addition, we construct the social graph models from COR and MFR. Furthermore, we combine the two models into one model G^R to calculate the

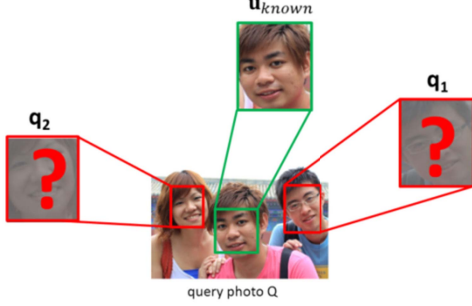


Figure 3. An example of query photo Q.

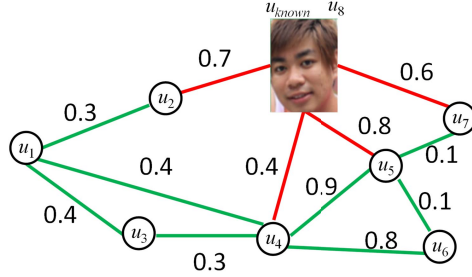


Figure 4. An example of social graph model G^R .

q_1	u_1	u_2	u_3	u_4	u_5	u_6	u_7	
Face score	0.8	0.4	0.2	0.1	0.8	0.3	0.9	
u_{known}	Relation Score	0	0.7	0	0.4	0.8	0	0.6
q_1	Similarity Score	0.4	0.55	0.1	0.25	0.8	0.15	0.75

Figure 5. Similarity score calculation by aggregating face score and relation score. (Assume that $\beta = 0.5$.)

relation score of each pair of users. Fig. 4 shows an example for social graph model G^R with assuming that there are 8 users, $\{u_1, u_2, \dots, u_8\}$, in the social network.

In online query, first, we adopt eigenface-based face recognition on q_1 . Hence, we can calculate the face score $f(q_1, u_j)$, $\{0.8, 0.4, 0.2, 0.1, 0.8, 0.3, 0.9\}$, for q_1 corresponding to the users $\{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$. Then, the known user u_{known} can be identified to u_8 from the known face. Since u_{known} and q_1 took the query photo together, there exist some relationships between them. The users who are close with u_{known} have higher possibility to appear in the query photo Q. For the known user u_{known} , we can retrieve the relation scores $r(u_{known}, u_j)$, $\{0, 0.7, 0, 0.4, 0.8, 0, 0.6\}$, corresponding to the users $\{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$ from social graph model G^R .

Next, similarity score $S(u_{known}, u_j)$ can be computed by aggregating the face score $f(q_1, u_j)$ and the relation score $r(u_{known}, u_j)$ for each user u_j . The higher the similarity score of u_j , the more possible that the u_j is the query face q_1 . An

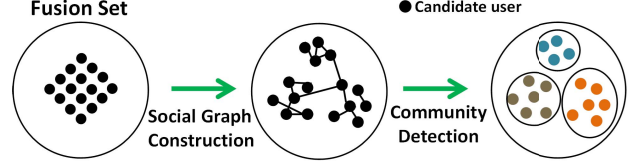


Figure 6. An example of social graph model construction and community detection.

example of similarity score calculation is shown in Fig. 5. Even if the face score of u_7 is the highest, the similarity score may not be the highest. It can reveal the importance of social relationships.

D. Community-Based Group Associations (CBGA)

In general, to recommend the tags for a query face, many studies utilize social contexts to discover who is most relevant to the known user and predict the query faces one by one. It only considers the relationship between the known user and the query face. However, most of people appearing together in a photo are caused by not only an event or activity, but some group associations and relationships existing among them. For example, user A, B and C usually take photo together, so do C and D. Given a query photo with known user C and two query faces, the set $\{A, B\}$ has higher possibility to be the one taking photo with C than the set $\{A, D\}$. For this purpose, we propose a *Community-Based Group Associations* method, called *CBGA*, to discover the group relationship and recommend a group of tags for a query photo. With CBGA, we consider the relationship among users to find a group of people who have the highest likelihood for the people appearing in the photo.

First, we select the top H candidate users with high similarity scores for each query face, and put the candidate users for all query faces into a *fusion set*. For example, there are three query faces in a query photo. If the top 10 candidate users for each query face are selected and put into the fusion set F , F contains 30 candidate users. To find a group of people with close relationship to each other in the fusion set, first we construct a social graph model G^F , where $U^F = F$. It means that each node in G^F is an individual in the fusion set F . Edges are the connections between individual, and weights are the relation score. Then, the community detection algorithm is used to detect the communities in G^F . For avoiding the sensitive threshold problem of density-based clustering, the parameter-free community detection algorithm SHRINK [8] is implemented to detect the communities. Fig. 6 illustrates an example of social graph model construction and community detection.

After community detection, we intend to choose a community which is most relevant to u_{known} . The relevant score $R(C_i)$ of a community C_i can be defined as

$$R(C_i) = \frac{\sum_{u_j \in C_i} r(u_{known}, u_j)}{|C_i|}, \quad (9)$$

where $|C_i|$ is the number of users in community C_i . According to the relevant score $R(C_i)$, all the communities are ranked. For each community, the candidate users are ranked by the relation score between u_{known} and themselves. For recommendation list construction, we select the community with higher rank and put the candidate user with higher rank in that community into the recommendation list, until the number of users in the recommendation list is equal to the number of query face.

IV. EXPERIMENT RESULT

We collect 1074 images and social contexts from Facebook.com with 94 volunteers. For training, 909 images are selected to be training image to train the eigenfaces and construct the social graph model for COR. The rest 165 images are testing images. Only the photo with more than two faces can be selected to be test data since we assume that there exists a known user in a query photo. To construct the groundtruths, we manually mark the location of the faces and tag the user's name onto them.

To evaluate the proposed method, two measurements, H-hit and precision, are used. H-hit is used to measure the robustness of the recommendation list for a query face, and defined as

$$H\text{-hit}(q_i) = \begin{cases} 1, & \text{if top } H \text{ tags contain the groundtruth of } q_i; \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

Precision is used to evaluate the robustness of CBGA, and defined as

$$Precision(Q) = \frac{|Predicted \cap Groundtruth|}{|Predicted|}, \quad (11)$$

where $Predicted$ is the set of predicted tags in the recommendation list, $Groundtruth$ is the set of groundtruths of query photo Q , and the $|\cdot|$ operation means the number of elements in the set.

In Fig. 7, we show the H-hit rate comparison with different score estimators. The H-hit rate with only face score is the lowest one since the face recognition is not accurate enough. The H-hit rates of similarity score with MFR and with COR are higher than face score. It says that COR and MFR are useful for relationship discovery. In addition, COR reflects the more real world relationship than MFR, since friends on the social network may not live in the same place or attend the same activity. Furthermore, the combination of COR and MFR has slightly improvement on H-hit rate because the MFR can discover the relationship when there is no COR between people.

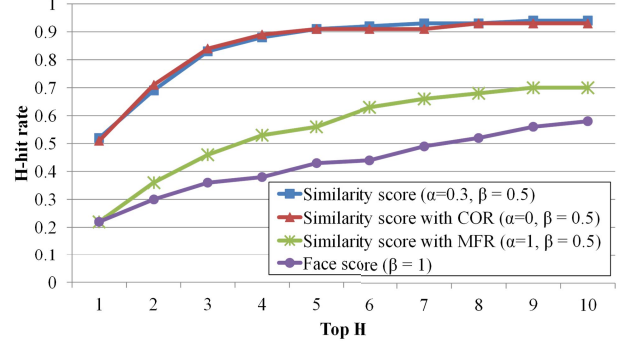


Figure 7. H-hit rate comparison with different score estimators.

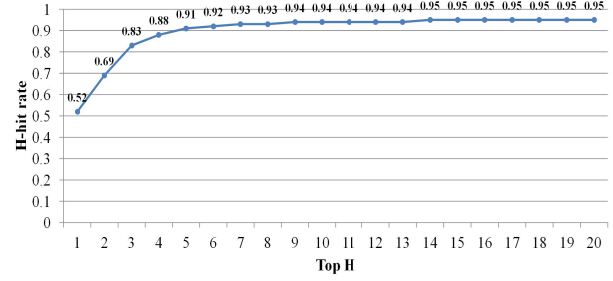


Figure 8. Top 20-hit rate for similarity score ($\alpha = 0.3$ and $\beta = 0.5$).

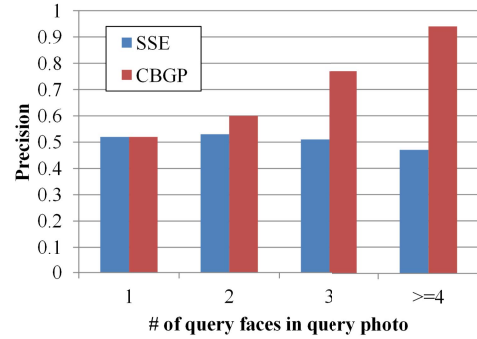


Figure 9. Precision comparison between SSE and CBGA.

For selecting the appreciate number of H , as shown in Fig. 8, we calculate the top 20 H-hit rate for the recommendation list with similarity score. We can observe that the hit rate increases slightly after $H = 10$. It means that 94% of the query face can be correctly identified in top 10 results. Thus, for efficiency consideration, we choose top 10 tags for each query faces as candidate users to construct the fusion set of CBGA.

For precision calculation, SSE selects the top 1 tag for each query face and put it into the recommendation list. The precision comparison between SSE and CBGA is shown in Fig. 9. The precision of SSE is about 50%, and there is no obvious change when the number of query faces increases. For CBGA, the more query faces, the higher the precision. That is, CBGA can discover the group relationship better if more query faces in a query photo. CBGA achieve 77% and

94% precision when the number of query faces is 3 and more than 4, respectively. The precisions of SSE and CBGA with one query face are the same since there is no group relationship for only one query face.

V. CONCLUSION

In this paper, we proposed a method for face tagging recommendation by community-based group associations (CBGA) on social network. The concept of CBGA is simple but effective for recommending a group of people to tag the query photo. Instead of only discovering the relation between the known user and the query face, CBGA can correctly tag the query faces via the relationships between query faces in the photo. From experimental results, CBGA can achieve 77% and 94% precision when the query faces are 3 and more than 4 in a query photo, respectively. To improve the performance and the robustness of the system, some enhancements can be done in the future:

1) Usage of text-based social context: in social websites, text-based social context contains much information that can be used for learning the relationship among network users. For instance, the profile page has information about users, such as: gender, name of high school, occupation, interest and so on. As a result, we can discover some relationship between two users by utilizing the text-based social context, such as they study in the same school.

2) Improvement of face recognition: although the using of relationship can improve the accuracy of face recommendation, face recognition still plays an important role in face tagging recommendation. An accurate face recognition approach makes the face tagging recommendation better.

ACKNOWLEDGMENT

This work is partially supported by "Aim for the Top University Plan" of the National Chiao Tung University and Ministry of Education, Taiwan, R.O.C., and partially by National Science Council of R.O.C. under the grant no. 98-2221-E-009-091-MY3 and 101-2218-E-009-004-.

REFERENCES

- [1] H.M. Chen, M.H. Chang, P.C. Chang, M.C. Tien, W. H. Hsu, and J.L. Wu, "SheepDog: group and tag recommendation for Flickr photos by automatic search-based learning," *Proc. of the 16th ACM International Conference on Multimedia*, Oct. 2008, pp. 737-740.
- [2] J. Y. Choi, W. D. Neve, K. N. Plataniotis, and Y. M. Ro, "Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks," *IEEE Transactions on Multimedia*, vol. 13, no. 1, Feb. 2011, pp. 14-28.
- [3] J. Y. Choi, W. D. Neve, Y. M. Ro, and K. N. Plataniotis, "Automatic face annotation in personal photo collections using context-based unsupervised clustering and face information fusion," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 10, Oct. 2010, pp. 1292-1309.
- [4] M. De Choudhury, H. Sundaram, Y.R. Lin, A. John, and D.D. Seligmann, "Connecting content to community in social media via image content, user tags and user communication," *Proc. of IEEE International Conference on Multimedia and Expo*, Jun. 2009, pp. 1238-1241.
- [5] A. C. Gallagher and T. Chen, "Using group prior to identify people in consumer images," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2007, pp. 1-8.
- [6] A. C. Gallagher and T. Chen, "Understanding images of groups of people," *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 256-263.
- [7] A. C. Gallagher and T. Chen, "Using context to recognize people in consumer images," *IPSP Transactions on Computer Vision and Applications*, vol. 1, Mar. 2009, pp. 115-126.
- [8] J. Huang, H. Sun, J. Han, H. Deng, Y. Sun, and Y. Liu, "SHRINK: a structural clustering algorithm for detecting hierarchical communities in networks," *Proc. of the 19th ACM International Conference on Information and Knowledge Management*, Oct. 2010, pp. 219-228.
- [9] H.N. Kim, A. El Saddik, K.S. Lee, Y.H. Lee, and G.S. Jo, "Photo search in a personal photo diary by drawing face position with people tagging," *Proc. of the 16th International Conference on Intelligent User Interfaces*, Feb. 2011, pp. 443-444.
- [10] M. Moricz, Y. Dosbayev, and M. Berlyant, "PYMK: friend recommendation at MySpace," In *Proc. of the International Conference on Management of Data - SIGMOD'10*, Jun. 2010, pp. 999-1002.
- [11] A. Rae, B. Sigubjörnsson, and R. van Zwol, "Improving tag recommendation using social networks," *Proc. of the 9th RIAO Conference on Adaptivity, Personalization and Fusion of Heterogeneous Information*, Apr. 2010, pp. 92-99.
- [12] B. Sigubjörnsson and R. van Zwol, "Flickr tag recommendation based on collective knowledge," *Proc. of the 17th International Conference on World Wide Web*, Apr. 2008, pp. 327-336.
- [13] Z. Stone, T. Zickler, and T. Darrell, "Autotagging Facebook: social network context improves photo annotation," *Proc. of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Jun. 2008, pp. 1-8.
- [14] Y. Song, L. Zhang, and C. L. Giles, "Automatic tag recommendation algorithms for social recommender systems," *ACM Transactions on the Web*, vol. 5, no. 1, Feb. 2011, pp. 4:1-31.
- [15] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, Jan. 1991, pp. 71-86.
- [16] Z. Wu, S. Jiang, and Q. Huang, "Friend recommendation according to appearances on photos," *Proc. of the 17th ACM International Conference on Multimedia*, Oct. 2009, pp. 987-988.
- [17] L. Zhang, L. Chen, M. Li, and H. Zhang, "Automated annotation of human faces in family albums," *Proc. of the 11th ACM International Conference on Multimedia*, Nov. 2003, pp. 355-358.
- [18] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, Dec. 2003, pp. 399-458.