PERSONALIZED GEO-TAG RECOMMENDATION FOR COMMUNITY CONTRIBUTED IMAGES

Prasanna Wadekar[1], Santosh Shinde[2]

[1]*PG Student, Department of Computer Engineering, VPCOE, Baramati, Maharashtra, India*
[2]*Assistant Professor, Department of Computer Engineering, VPCOE, Baramati, Maharashtra, India*

Abstract

Tagging is popularized by many social sharing websites, which allows us to add the description to object. Using tags users can organize their data so that it will be helpful for searching and browsing. Geotagging of a photo is the process in which a photo is marked with the geographical identification of the place it was taken. Geotagging can benefit users, to discover an Extensive Variation of exact Location related information. In personalized tag recommendation, tags that are relevant to the user's query are retrieved based upon the user's interest. The introduction of the Hypergraph learning is to find joint relevance between the visual and textual domains. Given a photo with Geolocation and without tags, System uses nearest neighbour search to obtain some user- predilected tags and geo-location predilected tags individually. It discovers the semantically and visually related images, and explores the idea of annotation-by-search to recommend tags for the untagged photo. In conclusion, the tags are recommended to the user.

Keywords: *Hypergraph Construction, Hypergraph learning, Personalization, Geo-tags, Preference learning*

---***---

1. INTRODUCTION

Advancement in web 2.0 technologies, multimedia creation, and sharing has become much easier than ever before. In communities there are many social sharing websites, which allows user to share photos, web links, songs, pictures etc. The photo sharing websites includes Flickr, Zoomer, Picasa, encourages users to create, annotate, share and comment on media data. A tag is a non-hierarchical keyword or term assigned to a piece of information (such as an Internet bookmark, digital image, or computer file) to describe that object[1]. Tagging allows user to find object when retrieving that object later. Tagging also increases accessibility of media object to the public as other users can find their relevant images.

Human can assign tags for photo but it requires a time. Tag recommendation inspires users to assign more tags in connecting the semantic gap betweenhuman concept and the features of media object, which provides a feasible solution for content based image retrieval. Many tag recommendation strategies have worked upon connection between tags and photos. Fig.1 Users have favor for photos while searching.
1) User can assign types for photos e.g. photos can be categorized such as architectural, natural, scientific etc.
2) Single photo can be tagged by two or more users with same or different tags.

Users like to create photo album with respect to the places they have visited and this task can be done with adding geo tags for photos. Geo tagging is the process of adding geographical information to various media objects in the form of metadata. Meta data for Geotagging contains longitude, latitude, city name etc. Same tags can be recommended to visually similar photos of user but if geo favor of user is considered then it will recommend photos that are relevant with location.

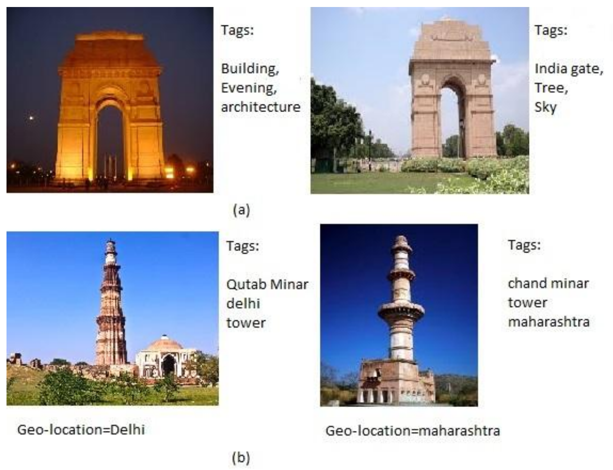


Fig.1 Tagging Behavior of user

Hypergraph learning will solve this tag recommendation by Homogeneous Hyperedge. Homogenous hyperedge is the edge between images and tags. Hypergraph will find joint relevance between image and tags by using visual content relationship hyperedge and textual relationship hyperedge uploaded by that particular user.

There exist two challenges:
1) To learn relevance of given tag to the visual content.
2) Image and Text are two different structures, to find common relation between these two structures is the task. To tackle these challenges, Personalized Geo-Tag Recommendation for Community Contributed Images is

used. It recommends tags based on users specific interest and geo specific interest by using Hypergraph learning.

The remainder of this paper is laid out as follows: Section 2 Describes Literature Survey. Section 3 Describes implementation of system. Section 4 Describes Experimental setup, Section 5 Describes Result and Dataset, Section 6 Describes Conclusion.

2. LITERATURE SURVEY

M. Ames et.al proposed in [2], it covered How to motivate users to provide tags to images. They have entitled all features that are provided by Flickr to the users. Authors have proposed a Zone tag application that has interface to provide title, tag, and description to photo that is going to be uploaded. However following are drawbacks of that method, Tag Suggestion becomes ambiguous and non-obvious. Irrelevant tag recommendation is encountered in this method. User preference is not considered with respect to visual content he has previously published.

A. Sun et.al proposed in [3], Tag recommendation is three stage process a) Tag relationship graph construction. b) Concept detection. c) Actual Tag recommendation. It enabled customized matching score computation of each user. It enhances scalability and efficiency of tag recommendation process. However, Geo specific information is not considered here.

B. Sigurbjörnsson et.al proposed in [4], Input to this system was photo along with user defined tags. System generated ordered list of m-candidate tags based on tag co-occurrence. List of candidate tags are used as input for tag aggregation and ranking. It has produced ranked list of n-recommended tags. It can handle growth of tags vocabulary. It can be used to recommend photos based on locations, artifacts, and objects. However it limits tags that are going to be provided by user. It is not personalized approach. System being less interactive as well as performance calculation is not accurate. System requires vital tuning parameters to be managed and this makes system expensive.

A. Silva and B. Martins proposed in [5], Georeferenced tag recommendation annotates geo referenced photos with descriptive tags. In this method supervised learning is used to rank methods by combining different estimators of tag relevance. Various ranking techniques are used. It does not consider content of image to improve to visual search as well as it ignores user preferences.

N.garg and I.Weber et al proposed in [6], A personalized tag recommendation idea that find out the tagging history by profile of users created by his sharing of photos and tags. System suggests tags dynamically based on it. It leads to conventional performance. It is less computationally complex than collective knowledge so that it can recommend tags dynamically. However, it ignores metadata of image and tags. Therefore, it leads to ignore preferences.

D. Rafailidis et al proposed in [7], a method that can handle very sparse data with Quadruple of user, image, tag and its associated weight. First, Construction of three order tensor then unfold matrix to create new three intermediate matrices. Use SVD on each unfolded matrix after that construction of core tensor s. Reconstruction of tensor then generation of item recommendation. It handles cold start and Sparsity problem. However it ignored geo specific information and time space computations were too much.

To tackle visual and textual tag recommendation along with geo tags Z. li et.al proposed in [8], subspace learning method in that unified latent space is find that combine visual, textual and Geo tagged relations.

Zhou et.al proposed in [9], Hypergraph is nothing but a generalization of the simple graph in which the edges, called hyperedges, are non-empty subsets of the vertex set. Therefore, the Hypergraph can be used to model entities and sparse relations.

Liu et.al proposed in [10], A transductive learning framework for image retrieval. It is based on a probabilistic Hypergraph. This method builds a Hypergraph by generating a hyperedge from each image and its adjacent neighbors after that ranking based on Hypergraph is then performed. Visual similarity matrix calculated from feature descriptors. Liu has chosen each image as a centroid vertex and created a hyperedge by a centroid and its k-nearest neighbors. It captured higher-order relationship.

Yue et.al proposed in [11], Intended an approach that simultaneously utilizes both visual and textual information for social image search.

In our proposed method, visual content and tags are used to generate the hyperedges of a Hypergraph. A relevance learning method is performed on the Hypergraph structure where a set of relevant samples are employed. Different from the conventional Hypergraph learning algorithms, our method learns not only the tag relevance based among images but also the Geo tags provided by that particular user. By using visual similarity edge and textual similarity edge we can recommend tags for that particular user.

3. SYSTEM IMPLEMENTATION

The proposed framework is organized into two stages Hypergraph Construction; Hypergraph based visual and tag learning, Hypergraph based Visual and Geo-tag learning and Tag Recommendation. The system architecture is as shown in fig.2.

3.1 Notations

Notations	Description
G=(V,E,W)	G indicates a Hypergraph where V and E indicate set of vertices and Hyperedges. Let w designates weight of hyperedge
u	User set

3.2 Mathematical Model

A Hypergraph G = {V, E, W} consists of the vertex set V, the hyperedge set E, and the hyperedge weight vector w. each edge e_i is assigned a weight $w(e_i)$ the Hypergraph G can be denoted by incident matrix H1.

$$h(v,e)=\begin{cases}1 \text{ if } v\in e\\ 0 \text{ if } v\notin e\end{cases}$$

Vertex degree of each vertex v ∈ V is:

$$d(v)=\sum_{e\in E} w(e)h(v,e)$$

For a hyperedge e ∈ E, hyperedge degree can be estimated by:

$$d(e)=\sum_{v\in V} h(v,e)$$

Let D_v and D_e denote diagonal matrices of vertex degree and hyperedge degrees, respectively, Let W denote the diagonal matrix of the hyperedge weights

$$W(i,j)=\begin{cases}1 \text{ if } i=j\\ 0 \text{ otherwise}\end{cases}$$

The regularize $\Omega(f)$ on the Hypergraph is defined as

$$\Omega(f)=\frac{1}{2}\sum_{e\in E}\sum_{u,v\in V}\frac{w(e)h(u,e)h(v,e)}{d(e)}\left(\frac{f(u)}{\sqrt{d(u)}}-\frac{f(v)}{\sqrt{d(v)}}\right)^2$$

$$\Omega(f)==f^T L.f \qquad (1)$$

3.3 Set Theory

Input Sets
I= {1, 2, 3,…, n} i.e. Set of images
T= {1, 2, 3,…,m} i.e. Set of tags
Processing sets {P, Q}
Output set T={T_1, T2, T3} ∈ T

Output: -Visual similarity matrix based on distance between them.

$$f(n)=\begin{cases}d \text{ if } v_i=v_j\\ 0 \;\; otherwise\end{cases}$$

Let d be distance between two images
1) P= *f(n)* be the function to construct visual content relationship.
Input: - set of images i.e. I= {1, 2, 3,…, n}
2) Q=f_1 *(n, m)* be the function to construct textual content relationship.
Input: - Set of images i.e. I= {1, 2, 3,…, n} And
Set of tags i.e. T= {1, 2, 3,…,m}

Output: -Visual textual similarity matrix based on whether particular tag is present or not. It is also called text representation matrix.

$$f_1(n)=\begin{cases}1 \text{ if } i_n \textit{ is tagged by tag } t_m\\ 0 \;\; otherwise\end{cases}$$

3) Z=f_3 *(P, Q)*
Input: - Processing sets i.e. output of above three phases P, Q.
Output: - for an image i_1 ∈ I set of tags will be recommended e.g. tag set for an image i_1= {T_1, T2, T3} ∈ T

3.4 System Architecture

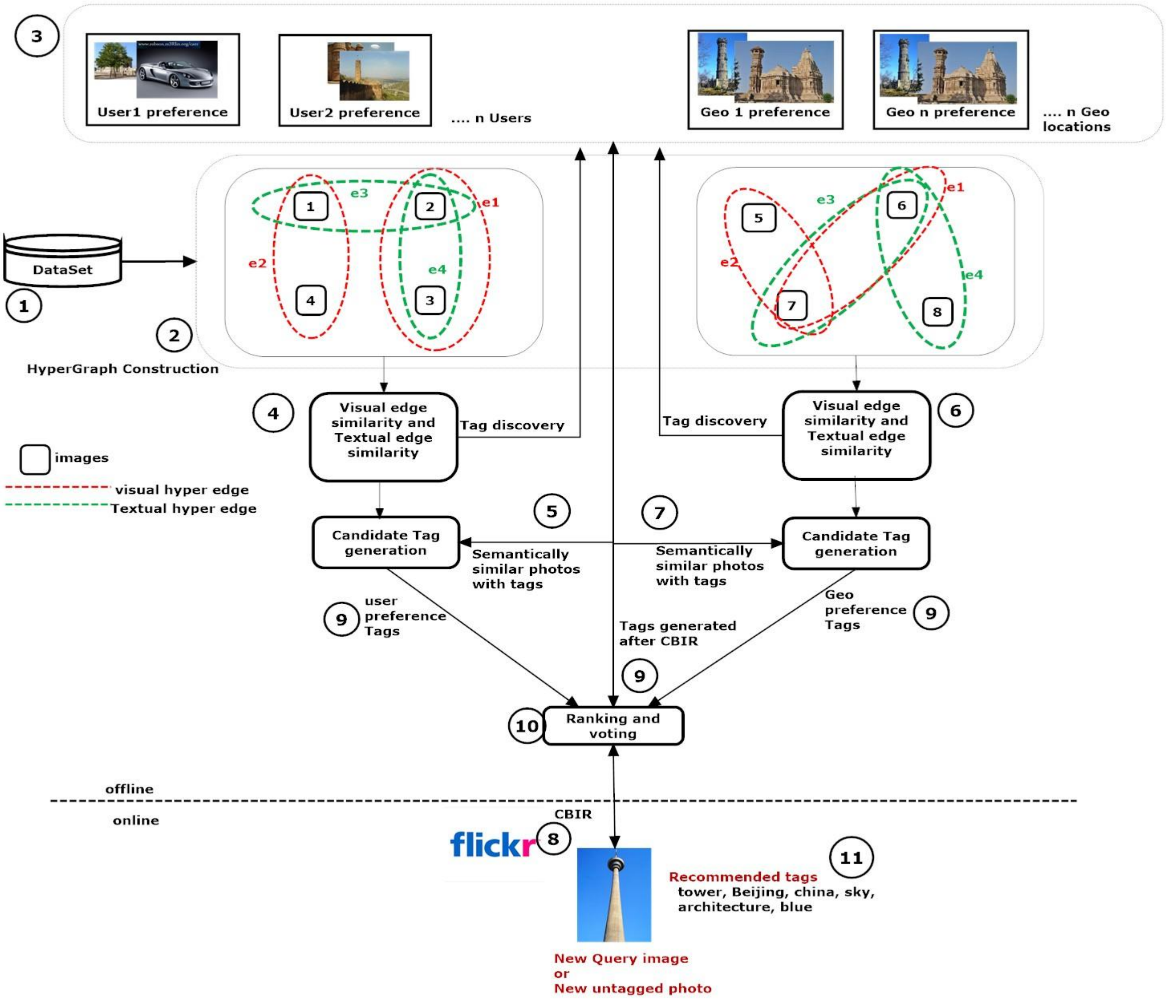


Fig. 2 System Architecture

4. EXPERIMENTAL SETUP

4.1 Offline Stage

4.1.1) Hypergraph Construction: - For Hypergraph construction, there are two types of vertices corresponding to the users and images, which constitute the vertex set denoted as V= {u, o}. The construction of the hyperedges is illustrated as follows:

Homogeneous hyperedges: - It is used to represent the visual-textual content relations among image vertices. There are two types of homogeneous Hyperedges. Visual content relation hyperedge ε_{visual} and textual content relation hyperedge $\varepsilon_{textual}$.

In our experiment, for each image 512-dimensional feature vector is extracted as the content representation. 512-dimensional GIST features. The visual content similarity Hyperedge's weight is set based on the visual similarity matrix Aij is calculated according to,

$$A_{ij} = \begin{cases} \exp\left(\dfrac{-\| x_i - x_j \|}{2\sigma^2}\right) & \text{if } j \in N_t(i) \text{ or } i \in N_t(j) \\ 0 & \text{otherwise} \end{cases}$$

Where N_t, denotes the index set for t nearest neighbors, x_i and x_j are feature vectors associated with images respectively, σ is a scaling parameter. Example from above equation.

To construct $\in_{textual}$, we build a tag vocabulary. Each tag is used to build a hyperedge, i.e., the images containing the same tag are connected by a hyperedge.

4.1.2) Calculation of Edge Similarity between visual and textual information: - We have to find relation between two edge similarities. If any column from visual incident matrix whose entries are having same values in textual similarity matrix then we say that there lies a common structure between image and text. Same procedure can be repeated for Geo tags.

4.1.3) Candidate tag generation: - When user uploads new image, system learns correlation between visual hyper edge and textual hyper edge of already uploaded images. It retrieves tags of images, which are correlated. System collects the tags of those visually and textually similar images and keeps the distinct tags. Candidate tags are generated from above step same procedure can be used for Geo tag.

4.2 Online Stage

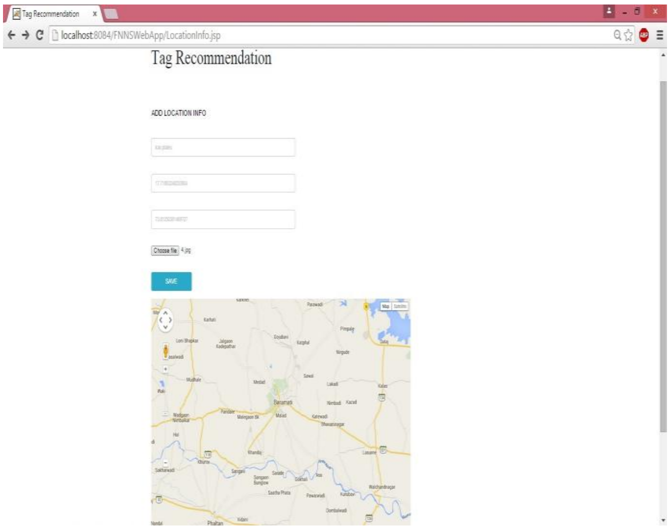
4.2.1) Tag Recommendation: - System will recommend tags based on learned correlation between visual and textual domains.

5. DATASET

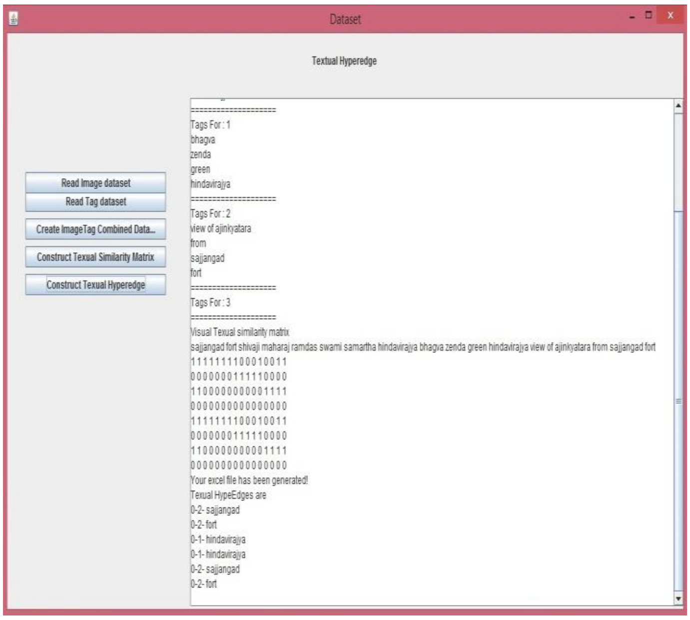
We have collected data from Flickr API to evaluate our approach. We have downloaded each user's information such as his Images, tags, geo tags (latitude and longitude). For geo tags we have extracted city level location names from Flickr. We have collected data such that each photo must have at least one tag, because photos without tag are not required to learning purpose. We divide user's photos into two sets one for training and one for evaluation. Evaluation photos are having tags but they are kept for ground truth purpose to check that whether our system recommends tags that are relevant to ground truth.

6. RESULTS

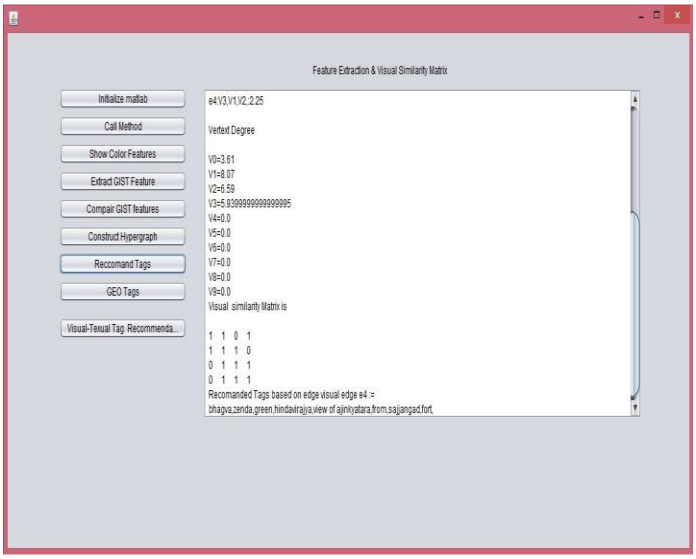
6.1 Upload Form



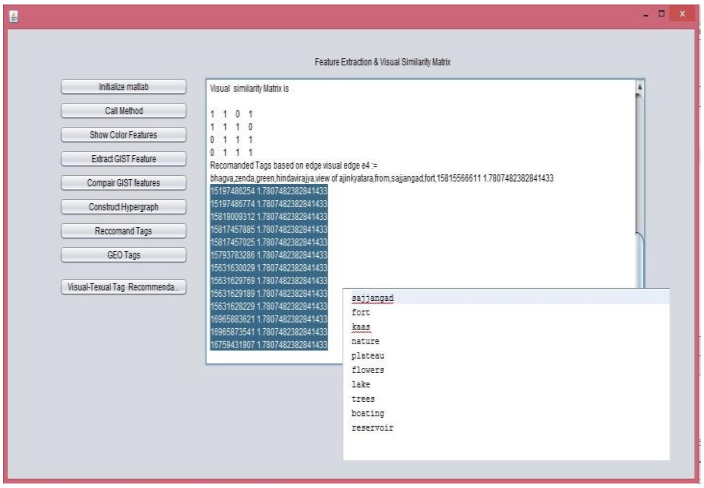
6.2 Textual Similarity and Textual Hyperedges



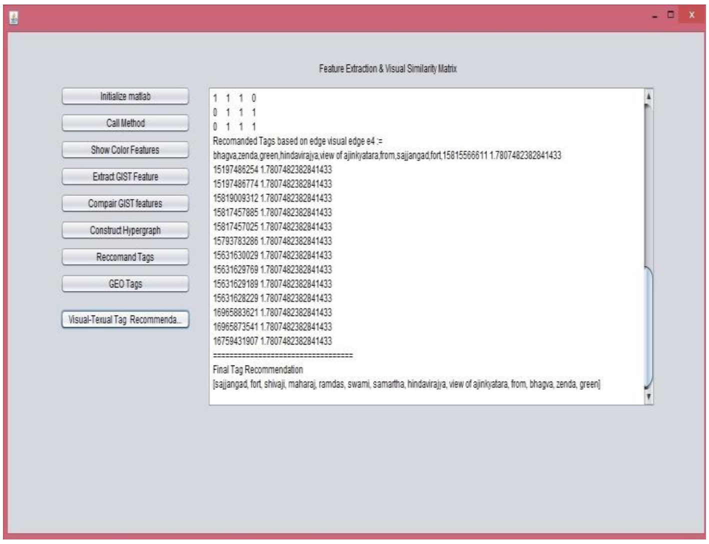
6.3 Visual Similarity and Visual Hyperedges



6.4 Geo Specific Tag Recommendation



6.5 Final Tag Recommendation



7. CONCLUSION

Personalized Geo-Tag recommendation for community contributed images is proposed to deal with the problem of learning joint relevance of tag to images. System bridges semantic gap between visual and textual features. It finds tags from visual homogenous hyperedge and they are used as input to find candidate tags from textual homogenous hyperedge. Finally, it recommends frequent but distinct tags.

ACKNOWLEDGEMENTS

I avail this opportunity to express my deep sense of gratitude and whole hearted thanks to my guide Prof. S. A. Shinde sir for giving his valuable guidance, inspiration and encouragement to embark this paper. Without his Coordination, guidance and reviewing, this task could not be completed alone. Prof. S. A. Shinde sir gave me all the freedom I needed for this project.

REFERENCES

[1] http://en.wikipedia.org/wiki/Tag_(metadata)

[2] M. Ames and M. Naaman, “why we tag: motivation for annotation in mobile and online media,” in *Proc.ACM CHI*, 2007.

[3] A. Sun, S. Bhowmick and J. Chong, “Social Image Tag Recommendation by Concept Matching,” in *Proc. ACM Multimedia*, 2011.

[4] B. Sigurbjörnsson and R. van Zwol, “Flickr tag recommendation based on collective knowledge,” in *Proc. ACM WWW*, 2008.

[5] A. Silva and B. Martins, “Tag recommendation for georeferenced photos,” in *Proc. ACM SIGSPATIAL Int. Workshop Location-Based Social Networks*, 2011.

[6] N. Garg and I. Weber, “Personalized, interactive tag recommendation for Flickr,” in *Proc. ACM Recommender Systems*, 2008.

[7] D. Rafailidis and P. Daras, “The TFC model: Tensor factorization and tag clustering for item recommendation in social tagging systems,” *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 43, no. 3, pp. 673–688, 2013.

[8] J. Liu, Z. Li, J. Tang, Y. Jiang, and H.Lu, “Personalized Geo-Specific Tag Recommendation for Photos on Social Websites,” *IEEE Trans. Multimedia*, vol. 16, no. 3, Apr. 2014.

[9] D. Zhou, J. Huang, and B. Schokopf, “Learning with hypergraphs: Clustering, classification, and embedding,” in *Proc. Adv. Neural Inf. Process. Syst.*, 2007, pp. 1–8.

[10] Y. Huang, Q. Liu, S. Zhang, and D. N. Metaxas, “Image retrieval via probabilistic hypergraph ranking,” in *Proc. CVPR*, 2010, pp. 3376–3383.

[11] Y. Gao, M. Wang, “Visual-Textual Joint Relevance Learning for Tag-Based Social Image Search,” in Proc. *ACM Multimedia*, 2011, pp. 15171520.

[12] http://en.wikipedia.org/wiki/Photo_sharing

BIOGRAPHIES

Prasanna Wadekar Has completed his BE degree in 2012 From Shivaji University. Currently he is pursuing his ME degree from VPCOE, Baramati, Savitribai Phule Pune University. He has attended various workshops organized by IITB and IITD with remote center. He has attended “Computer vision” by Dr. Sumantra Dutta Roy. He attended workshops of Latex and SciLab conducted by remote center of VPCOE under IITB. He has published Review paper on same topic in IJERGS. He has participated in cPGCON-2015 i.e. post Graduate Conference.

Prof. Santosh Shinde received his B.E. degree in computer engineering (First Class with Distinction) in the year 2003 from Pune University and M. E. Degree (First Class with Distinction) in Computer Engineering in 2010 from Pune University. He has eleven years of teaching experience at undergraduate and postgraduate level. He has attended various National and International Conferences, Seminars and Workshops on various subjects like Multimedia Techniques, Multicore Computing and Software Engineering. He is a life member of ISTE (Indian Society for Technical Education) and IACSIT (International Association of Computer Science and Information Technology). He has worked as a review committee member for the 1st International Conference on recent trends in Engineering and Technology (ICRTET’2012) held at SNJB’s COE, Nashik. He has worked as a Judge for the State level paper presentation (REBEL) held at SVPM’s COE, Malegaon. He has attended workshops on Robotics, Research paper writing and Latex conducted by IIT, Bombay remote center at VPCOE, Baramati. He has to his credit IBM RFT certification with 100% score. He has organized one day workshop on Software Engineering by Dr. S. A. Kelkar, adjunct Professor IIT, Powai.