IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-9 E-ISSN: 2347-2693

How Flickr Helps to Know the Place: Visual and Textual Summarization of Geo-location

S. A. Takale[1*], P. J. Kulkarni[2]

[1*] Dept. of Information Technology, VPKBIET-SPPU, Pune, India
[2] Dept. of Computer Science and Engineering WCE, Sangli, India

Corresponding Author: sheetaltakale@gmail.com Mobile No: +91-9823149304

Available online at: www.ijcseonline.org

Received: 22/Aug/2017, Revised: 04s/Sep/2017, Accepted: 19/Sep/2017, Published: 30/Sep/2017

***Abstract*—**The work proposed here addresses the task of organizing the geo-referenced media on Flickr to generate Visual and Thematic Summarization of specified geo-location. The major challenge addressed in this work is: how to use the unstructured and unrestricted community contributed media and annotations to generate knowledge? The social tags associated with social images suffer from various problems such as, "free nature of tags", "tag relevance", and "cold start". To deal with these problems, we consider ternary interrelations and multiple intra-relations among user, image and tag and model the relations using HOSVD, a Tensor Reduction Technique. As a result, context information for geo-location images is generated using user's potential annotations. Content and context information for geo-location images and probabilistic generative model is utilized for location visualization. The novel visualization scheme proposed here summarizes the geo-location with a rich display landscape and provides location description using location representative tags. Experiments are performed on geo-tagged Flickr images for various geo-locations. The experimental results have validated the proposed method.

Keywords— Tag Refinement, Tensor Factorization, Geo-Referenced Photos, Summarization, Clustering, Image Search, Collection Visualization.

I. INTRODUCTION

Due to the advent of geo-referenced digital photography photos are now associated with the metadata describing the geographic location in which they were taken. The rapid growth of content sharing websites has resulted in sheer volume of location data stored on social media websites. The largest repositories of user location histories are in fact photo sharing web sites like Flickr. Viewing and interacting with such collections has a broad social and practical importance[5]. However, these collections are inherently difficult to organize, browse and search due to their size and the inability of computers to understand the content of the images. This is leading to interesting tasks like geographically organizing photos and location visualization[2,4,7,8].

The work proposed here addresses an interesting task of organizing these geo-referenced media on Flickr to generate Visual and Thematic Summarization of specified geo-location. The fact is that the original annotations available are not enough to describe the context of image. Therefore, we transfer the problem of location summarization to users' annotation prediction [3,6]. Sub part of our work is to improve the underlying associations between the images and tags provided with the raw tagging data from photo sharing websites. As a result, context information for geo-location images is generated using user's potential annotations. Content and context information for geo-location images and probabilistic generative model based on PLSA [11] is used for location visualization and theme generation.

User affinity, image affinity and tag affinity are modeled in the form of graph to represent intra-relations among them. Multiple intra-relations and interrelations among User, Image and Tag are incorporated into 3-order tensor. We utilize tensor factorization framework [3,6,10] for tag refinement . Major contributions of the paper include:

- We introduce the problem of Location visualization using context and content information associated with the geo-referenced social media.
- The unstructured and unrestricted community contributed images and annotations are utilized to generate the knowledge.
- Annotation refinement or recommendation of tags for un-tagged geo-referenced images is achieved.
- We propose a probabilistic generative model to discover the location theme by exploiting and combining textual and visual content of images.

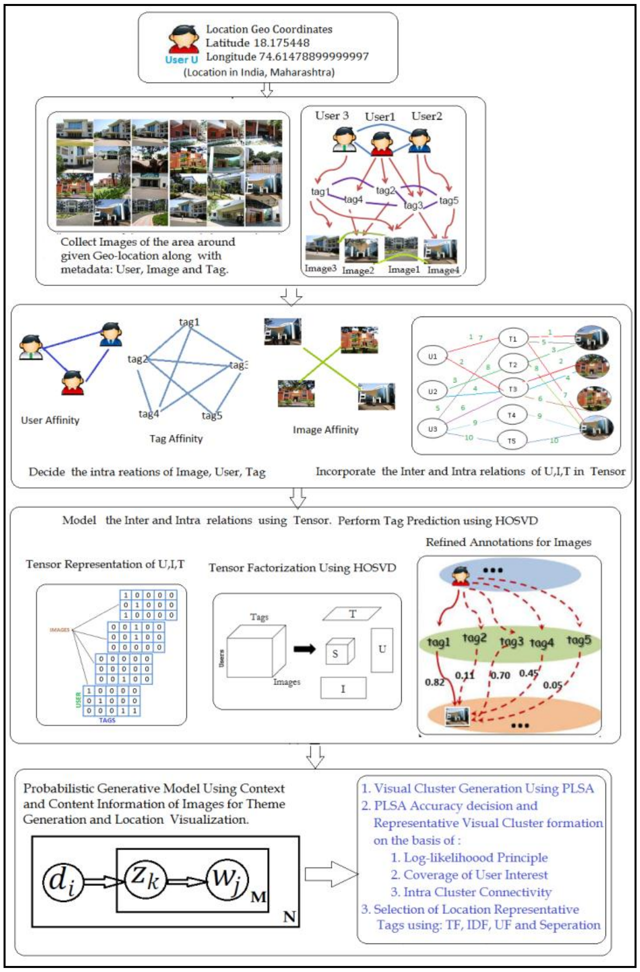


Figure 1: System Architecture

II. RELATED WORK

Stevan Rudinac et. al. [9] developed an approach for automatic visual summarization of a geographic area that exploits user contributed images and related explicit and implicit metadata. It is based on the random walk with restarts over a graph that models relations between images, visual features extracted from them, associated text, as well as the information on the uploader and commentators.

Quan Fang et.al [8] investigated the problem of location visualization from multiple theme. They identify the highly photographed place (POI) and discover their distributed themes.

Qiang Hao et.al [7] proposed an approach for location overview generation approach, which first mines location representative tags from travelogue and then uses these tags to retrieve images from web.

Lyndon Kennedy et.al [4] proposed an approach to make sense of the world using Flickr images. The two step method extracts the representative tags first, which are further utilized to automatically find place and event semantics.

III. METHODOLOGY

Figure 1 outlines the framework of the system for visual summarization of location using community contributed images and UGC.

Table 1. : List of Key Notations

Symbol	*Description*
$\boldsymbol{Y}$	*User-Image-Tag Ternary Interrelations*
$\boldsymbol{U}, \boldsymbol{I}, \boldsymbol{T}$	*User, Image and Tag Matrix*
$\boldsymbol{u}, \boldsymbol{i}, \boldsymbol{t}$	*Represent user, image and tag index.*
$\boldsymbol{C}$	*Core Tensor*
$\widehat{\boldsymbol{Y}}$	*Refined User-Image-Tag Ternary Interrelations*
$\boldsymbol{r_u}, \boldsymbol{r_i}, \boldsymbol{r_t}$	*Rank of matrix U, I, T respectively.*

A. Image Tag Refinement using Tensor Factorization

Tensor is a mathematical representation of a multi-way array. The order of a tensor is the number of modes. A second-order tensor is a matrix, and a higher-order tensor has three or more modes. The most important tensor operation is tensor factorization. Many tensor factorization methods have been proposed, among which, CANDECOMP/PARAFAC (CP) and Tucker decomposition are the most popular ones. Tensor decomposition methods [3,6,10] have been exploited for personalized tag recommendations.

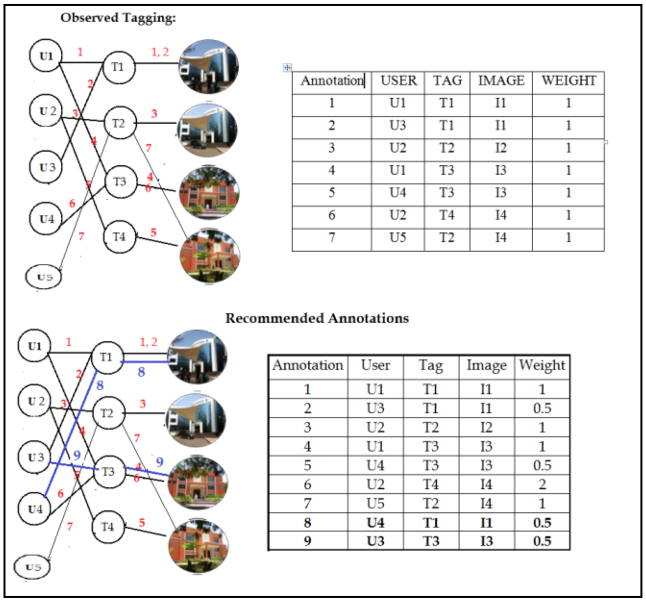
The raw tagging information obtained from community contributed images on the web has three types of interrelated entities, i.e. user, image and tag. The relations among user, image and tag are complicated. There exist intra relations among objects of the same type as well as 3-order interrelations among objects of different type. The goal of our work is to improve underlying associations between images and tags. We perform 3-dimensional analysis of the tagging data in order to discover the associations among these multi-type objects consequently the tags can be recommended according to captured associations. We use 3-order tensor to represent the ternary relations among user, image and tag. We apply the dimensionality reduction in 3-order tensor to reveal the latent semantic association among user, image and tag.

The tagging data can be viewed as a set of triplets. Let U, I, T denote the sets of users, images, tags. The set of observed tagging data is denoted as, $O \subset U \times I \times T$. The ternary interrelations can then constitute a three dimensional tensor, $Y \in R^{|U| \times |I| \times |T|}$ .To jointly model the three factors of *user*, *image*, and *tag*, we employ the general tensor factorization model, Tucker decomposition. In Tucker Decomposition [3,6,10], the tagging data Y are estimated by three low rank matrices and one core tensor. $\hat{Y} = C \times_u U \times_i I \times_t T$, where, $\times_n$ is the tensor product of multiplying a matrix on mode *n*.

The Tensor decomposition problem is reduced to minimizing a point-wise loss on $\hat{Y}$, defined as :

$$\min_{U,I,T,C} \sum_{(u,i,t)\in|U|\times|I|\times|T|} \left(\hat{Y}_{u,i,t} - Y_{u,i,t}\right)^2 \quad (1)$$

where, Y is the original observed tagging data and $\hat{Y}$ is the result of tensor factorization.



Annotation	USER	TAG	IMAGE	WEIGHT
1	U1	T1	I1	1
2	U3	T1	I1	1
3	U2	T2	I2	1
4	U1	T3	I3	1
5	U4	T3	I3	1
6	U2	T4	I4	1
7	U5	T2	I4	1

Annotation	User	Tag	Image	Weight
1	U1	T1	I1	1
2	U3	T1	I1	0.5
3	U2	T2	I2	1
4	U1	T3	I3	1
5	U4	T3	I3	0.5
6	U2	T4	I4	2
7	U5	T2	I4	1
8	**U4**	**T1**	**I1**	**0.5**
9	**U3**	**T3**	**I3**	**0.5**

Figure 2: Tag Recommendation Using HOSVD

B. HOSVD Algorithm Steps :

1. Construct the initial tensor:

The ternary relation between user, image and tag is mapped to a three dimensional tensor $Y \in R^{|U|\times|I|\times|T|}$ which is defined as: $Y_{u,i,t} = \begin{cases} 1 & if(u,i,t) \in O \\ 0 & otherwise \end{cases}$

2. Matrix Unfolding :

The initial tensor Y is matricized in three modes i.e. User, Image and Tag and we create 3 new matrices:

$$Y_1 \in R^{|U|\times|I||T|},\ Y_2 \in R^{|I|\times|U||T|}, Y_3 \in R^{|T|\times|U||I|}$$

3. Application of SVD on Each Mode:

SVD is employed to generate the three matrices A1, A2, A3.

$$A_1 = Y_1.Y_1^T, A_2 = Y_2.Y_2^T, A_3 = Y_3.Y_3^T$$

$$A_1 = U.S_u.V_u^T \quad A_2 = I.S_i.V_i^T \quad A_3 = T.S_t.V_t^T$$

4. Computing Low Rank Approximation

Compute the numbers c_1, c_2 and c_3 for approximation of tensor, where c_i is the numner of dimensions maintained for i-mode${U_{c1}^{(1)}}^T, {U_{c2}^{(2)}}^T, {U_{c3}^{(3)}}^T$.

5. Core Tensor C Construction:

The core tensor C governs the interactions among user, item and tag entities. $C = Y \times_1 {U_{c1}^{(1)}}^T \times_2 {U_{c2}^{(2)}}^T \times_3 {U_{c3}^{(3)}}^T$

6. The tensor $\hat{Y}$ Construction

Finally, tensor $\hat{Y}$ is build by the product of the core tensor C and the mode products of the three matrices$U^{(1)}, U^{(2)}, U^{(3)}$, as follows: $\hat{Y} = C \times_1 U_{c1}^{(1)} \times_2 U_{c2}^{(2)} \times_3 U_{c3}^{(3)}$

C. Multiple Intra Relations of User, Image and Tag

To handle the problem of sparsity and to increase the quality of recommendations, multiple Intra relations [3] between $U, I,$ and T are incorporated into the Tensor Factorization. The multiple intra relations between $U, I,$ and T are modeled and represented in the form of graphs. Graph based clustering algorithm is further utilized. Tag refinement problem addressed here is divided into two steps: incorporation of intra relations of $U, I,$ and T into tensor and, tensor Factorization using Tucker Decomposition.

1) *User Affinity :*

Each image is owned, shared, annotated or commented by user. For example if the image owned by user U_1 is commented by the user U_2, then the user U_1 and user U_2 are associated. Therefore, we measure the affinity relationship between two users using the number of images shared, annotated, commented and liked or marked favorite by them. Assume that u_i is user i and u_j is user j. $n(u_i)$ and $n(u_j)$ represent the number of images shared, annotated, commented, liked or marked favorite by user i and user j, respectively. $n(u_i, u_j)$ is the number of images shared, annotated, commented, liked or marked favorite by both the user i and user j.

$$Sim_{i,j}^U = \frac{n(u_i, u_j)}{n(u_i) + n(u_j)}$$

2) *Image Affinity :*

For each image we extract three types of features to capture 81- dimensional global color and 120-dimensional texture content and 512-dimensional GIST features of image [8]. Grid color moment features are used to represent the spatial color distributions in the images and Gabor textures to represent the texture. We concatenate these three feature sets together to produce a single feature vector each image in the data set. Visual similarity between images is defined by using RBF kernel.

$$Sim_{i,j}^I = exp\left(-\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right)$$

where $\left\|x_i - x_j\right\|^2$ is squared Euclidean distance between the two feature vectors and σ is a scaling parameter which is set to median values of Sim^I.

3) ***Tag Affinity:***

Tag affinity graph is constructed based on the tag context and semantic relevance [3]. The context relevance of two tags is simply encoded by their weighted co-occurrence in the image collection. Tag semantic relevance information is obtained using WordNet. Contextual Similarity Between two tags t_i is tag i and t_j is tag j is computed as:

$$CSim_{i,j}^T = \frac{n(t_i, t_j)}{n(t_i) + n(t_j)}$$

Semantic similarity between two tags t_i is tag i and t_j is tag j represented as $SSim_{i,j}^{T}$ is computed using WordNet. If two tags share the semantic relations such as synonym, hypernym, hyponym, meronym, and holonym, the words are considered as “related”. We assume λ_c and λ_s are the weights of context relevance and semantic relevance $Sim_{m,n}^{T}$ represents similarity between two tags t_m and t_n.

$$Sim_{i,j}^{T} = \lambda_c\, CSim_{i,j}^{T} + \lambda_s\, SSim_{i,j}^{T}$$

4) *Regeneration of Tensor by Incorporating Multiple Intra Relations:*

The multiple intra relations of user, image and tag are incorporated into Tensor Factorization framework [1]. The rank of tensor in example is $R^{5\times4\times4}$ as there are 5 Users, 4 images and 4 tags. Based on Tag affinity definition, clustering result is: $\{T_1, T_2, T_3\}$ and $\{T_4\}$. So we regenerate initial triplets in the form of tag clusters: $CT_1 - \{T_1, T_2, T_3\}$ $CT_2 - \{T_4\}$. Based on regenerated triplets, the initial tensor is reconstructed as, $Y \in R^{|U|\times|CT|\times|I|}$. Similarly, based on image affinity and user affinity, the initial tensor is reconstructed as, $Y \in R^{|U|\times|T|\times|CI|}$ and $Y \in R^{|CU|\times|T|\times|I|}$.

ID	USER	TAG	IMAGE	WEIGHT
1	U1	T1	I1	1
2	U3	T1	I1	1
3	U2	T2	I2	1
4	U1	T3	I3	1
5	U4	T3	I3	1
6	U2	T4	I4	1
7	U5	T2	I4	1

(a)

ID	USER	TAG	IMAGE	WEIGHT
1	U1	CT1	I1	1
2	U3	CT1	I1	1
3	U2	CT1	I2	1
4	U1	CT1	I3	1
5	U4	CT1	I3	1
6	U2	CT2	I4	1
7	U5	CT2	I4	1

(b)

Figure 3 : (a) Triplets in the Initial Form (b) Representation of Triplets using Tag Clustering

D. Visual Cluster Generation using Extended PLSA:

For document clustering and theme generation, we propose the use of probabilistic generative model which extends standard PLSA. Assume, for a given geo-location, a set of N images $D = \{d_1, d_2, \dots\dots\dots, d_N\}$ is retrieved. Each image d is represented as a vector of word occurrences, $W = \{w_1, w_2, \dots\dots\dots, w_M\}$, which are collected from associated tags. By considering image as virtual document and tags as terms, we obtain Document-Term matrix. We apply PLSA to model the generation of location images and tag occurrences. The Document-Term Matrix is constructed as follows:

The rank of matrix is $DT \in R^{|I|\times|T|}$, where,

$$DT_{i,j} = \begin{cases} 1, if\ term\ (tag) j\ belongs\ to\ document\ (image)\ i \\ 0, otherwise \end{cases}$$

PLSA associates an unobserved class variable $z \in Z = \{z_1, z_2, \dots\dots, z_k\}$ with each occurrence of $w \in W = \{w_1, w_2, \dots\dots\dots, w_M\}$ in a document $D = \{d_1, d_2, \dots\dots\dots, d_N\}$. In terms of probabilistic generative model it can be defined as:

- (d_i, w_j) is an observed pair, tag w_j is associated with document d_i.
- select an image d_i with probability $P(d_i)$
- pick a latent class/topic z_k with probability $P(z_k|d_i)$
- Generate a word w with probability $P(w_j|z_k)$

As a result one obtains an observed pair (d_i, w_j), when the latent class variable z_k is discarded.

$$P(d_i, w_j) = P(d_i)P(w_j|d_i) \text{ where,}$$

$$P(w_j|d_i) = \sum_{k=1}^{K} P(w_j|z_k)P(z_k|d_i)$$

As discussed in [12], the standard Expectation-Maximization approach is used to compute the model parameters, $P(z_k|d_i)$ and $P(w_j|z_k)$.

Standard PLSA results in k number of image clusters. The likelihood function considers only the textual associations among the images and tags and does not include the content level information for images. *Fact is, images with similar content should share common topics*. Therefore, we can use the image similarity as a constraint over images to learn the latent topics of interest more accurately. Similarly, the image cluster generated by the standard PLSA may have all the images from same user. Ideally, the image cluster for specified geo-location should be covering a broad interest. As discussed in [12], the extended PLSA proposed here aims at maximization of following two conditions in addition to the log-likelihood principle.

- Coverage of User’s Interest: It is the most important factor for clustering of images, because a cluster must be covering broad interest. *Aim is maximization of number of users $|U_k|$ that are represented in photos from cluster z_k.*
- Intra Cluster Connectivity: If cluster’s photos are linked to many other photos in the same cluster, then the cluster is more likely to be representative. The links between photos represent that the photos are visually similar and share same context (set of tags). *Aim is maximization of average number of links per photo in the cluster.* Visual similarity between photos is decided by following the method described in Image Affinity. Image context information will be decided as follows: Assume T_u and T_v are the set of tags assigned to the two images D_u and D_v, respectively. Using Jaccard similarity measure:

$$ConxtSim = \frac{n(T_u \cap T_v)}{n(T_u) \cup n(T_v)}$$

where, $n(T_u)$ and $n(T_v)$ represent the number of tags assigned to image u and v respectively and $n(T_u \cap T_v)$ represents the number of tags common for image u and v.

Ranking of Cluster:
The score of the cluster is decided for ranking of results. The score is computed as [12]:

$$score(z_k) = \sum_{i=1}^{|U|} \log(N_{images}(u_i) + 1)$$

where, $N_{images}(u_i)$ is the number of images in the cluster z_k which are of user's interest and which satisfy cluster coherance. From each ranked cluster we select representative images

5) ***Location Representative Tag Generation:***

Tag Cleaning:
Once the clusters have been determined, we perform tag cleaning on each cluster. A lot of tags are collected from all the images in the cluster. Many of these tags are noisy, irrelevant, general frequent tags or very rarely occurring tags. These tags need cleaning.

Tag Score:
The system computes scores for each cluster's tags to extract representative tags. In other words, we consider each cluster z_k, and the set of tags T_k that appear with photos from the cluster.

We assign a score to each tag $t \in T_k$ according to the three factors: *Term Frequency (TF)-* $tf(z_k, t)$, *Inverse Document Frequency (IDF)-* $idf(t)$, *User Frequency(UF) -* $uf(t)$.

One of the factors we use is TF-IDF (term frequency, Inverse Document Frequency). This metric assigns a higher score to tags that have a larger frequency within a cluster compared to the rest of the area under consideration. The assumption is that the more unique a tag is for a specific cluster, the more representative the tag is for that cluster. To avoid tags that appear only a few times in the cluster, the term frequency element prefers popular tags. While the tag weight is a valuable measure of the popularity of the tag, it can often be affected by a single user who accesses the image a large number of times. To guard against this scenario, we include a user element in our scoring, that also reflects the heuristic that a tag is more valuable if a number of different users use it. In particular, we factor in the percentage of users in the cluster z_k that use the tag t.

The final score for tag t in cluster z_k is computed by

$$score(z_k, t) = tf(z_k, t) \cdot idf(t) \cdot uf(\mathrm{t})$$

The higher the $TF\text{-}IDF$ score and the UF score, the more distinctive the tag is within a cluster. For each cluster, we retain only the tags that score above a certain threshold. The threshold is needed to ensure that the selected tags are meaningful and valuable for the aggregate representation. We use an absolute threshold for all computed clusters to ensure that tags that are picked are representative of the cluster.

Tag Selection:
The goal of tag selection algorithm is that a) important textual concepts that are related to specific location are selected and b) unimportant or highly personal tags are demoted.

For user specified geo-location, given a set of images, $D = \{d_1, d_2, \dots\dots, d_N\}$. $D_k = \{d_l\}_{l=1}^{N} \in D$ is a set of images for topic k. We aim to extract representative tags for this topic from the complete set of tags W_k associated with the topic k. We follow the following two conditions for extraction of representative tags:

Concept Representative Tag: If a tag t is representative tag for theme/topic k, then the probability of observing the tag t among images D_k of theme k is larger than the probability of observing it among all images in D.

Visual Representative Tag: A tag t is a visually representative tag if its annotated images are visually similar to each other.

*User's Interest Representative Tag:*A tag t is representative as per user's interest if same tag is utilized by maximum number of users.

Condition 2 and 3 are already covered according ranking of cluster's. Occurrence probability of tag t_i in the set of images D_k of theme k is computed as:

$$p(t_i|D_k) = \frac{n(t_i \cap D_k)}{n(D_k)}$$

In the same way, occurrence probability of tag t_i in the complete set of images D is computed as:

$$p(t_i|D) = \frac{n(t_i \cap D)}{n(D)}$$

Rank of tag t_i for topic k is decided by the condition:

$$p(t_i|D_k) - p(t_i|D) > 0$$

IV. Results and Discussion

A. Data Set

We have conducted extensive experiments to evaluate the effectiveness and usefulness of the proposed framework. Our dataset was collected by crawling images and photo metadata using Flickr API. To test our approach we selected different locations in India which are well known. Only those locations are selected for which we can retrieve at least 100 CC-licensed Flickr images. Flickr images within the radius of 1 km from each location are selected. Together with image, the accompanying metadata, i.e. tags and user information are also collected. Tags are preprocessed: stop words, meaningless words, time and number related words and camera related words, are removed from the tag vocabulary. some general frequent tags are removed by following Luhn's idea. Suffix removal is done using Porter's Algorithm.

B. User Survey

We conduct a small-scale user study to evaluate the effectiveness of the proposed method and the user experience of the novel visualization form. The experiment was conducted with two different locations which were known to the user. For each location, images are grouped into four different themes as shown in figure 4. Four criteria are considered: *Representativenes*- the total number of images in the cluster and number of representative images in the cluster. *Coverage*- the extent that the mined visual themes and representative tags provide sufficient information about the location. *Uniqueness*: - the number of representative images minus redundant photos in the cluster.(0: Not Unique, 10: Unique). *Satisfaction*- how satisfactory are the aggregated multiple themes for location visualization. We invited 20 participants for the user study experiment. The eight themes are selected for evaluation. Those 20 participants are selected for the evaluation, who are well known to the selected geo-location. The results are averaged over all participants for each theme and shown in Figure 4. The participants gave positive feedback to the novel location visualization scheme.

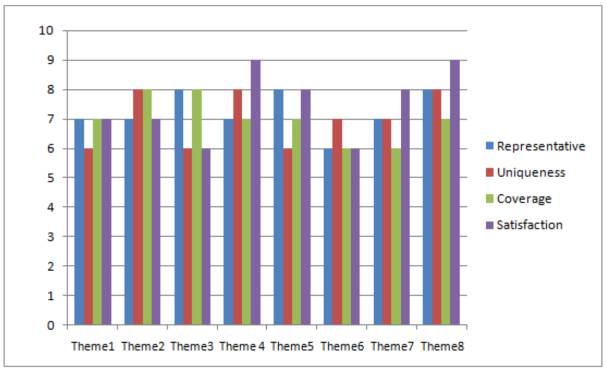


Figure 4: User Survey Results

V. Conclusion

The phenomenal growth of personal and shared digital photo collections presents considerable challenges in building navigation and summarization applications. By utilizing our method for location visualization, we enable users to view the most relevant samples from large-scale photo collections. We have presented a novel location visualization scheme using extended PLSA for geographically and thematically organizing photos into multiple themes. The proposed tensor factorization HOSVD using intra-relations between use, image and tag helps deal with the tagging problems of, “visual textual relevance”, “cold start”, “ternary relationship among user, image and tag”. Experiments on Flickr datasets for various known locations show that the proposed framework greatly outperforms the baseline and also shown its advantage in deriving compact location visualization and themes for improving user experiences.

References

[1] Dimitrios Rafailidis, Petros Daras, (2013), “*The TFC Model: Tensor Factorization and Tag Clustering for Item Recommendation in Social Tagging Systems*”, IEEE Transactions On Systems, Man, And Cybernetics: Systems, Vol. 43, No. 3, May 2013

[2] Jaffe A, Naaman M, Tassa T, Davis M, (2006), “*Generating Summaries And Visualization For large Collections Of Geo-Referenced Photographs”.* In: Proceedings of the 8th ACM international workshop on multimedia information retrieval. ACM, New York, pp 89–98.

[3] Jitao Sang, Changsheng Xu, Dongyuan Lu,(2012), “*User-Aware Image Tag Refinement via Ternary Semantic Analysis”*, IEEE Transactions On Multimedia, Vol. 14, No. 3, June 2012.

[4] Kennedy L, Naaman M (2008) “*Generating Diverse and Representative Image Search Results for Landmarks”.* In: Proceeding of the 17th international conference on World Wide Web. ACM, New York, pp 297–306.

[5] Yan-Tao Zheng · Zheng-Jun Zha · Tat-Seng Chua (2011), “*Research and applications on georeferenced multimedia: a survey*," Multimed Tools Appl 51:77–98 DOI 10.1007/s11042-010-0630-z.

[6] Panagiotis Symeonidis, Alexandros Nanopoulos, and Yannis Manolopoulos, (2010), “*A Unified Framework for Providing Recommendations in Social Tagging Systems Based on Ternary Semantic Analysis*”, IEEE Tra nsactions On Knowledge And Data Engineering, Vol. 22, No. X, 2010 .

[7] Qiang Hao, Rui Cai, Xin-Jing Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang, 2009, “*Generating Location Overviews with Images and Tags by Mining User-Generated Travelogues*”, MM’09, October 19–24, 2009, Beijing, China.

[8] Quan Fang, Jitao Sang, Changsheng Xu, and Ke Lu, 2013, “*Paint the City Colorfully: Location Visualization from Multiple Themes*”, MMM 2013, Part I, LNCS 7732, pp. 92–105, 2013.

[9] Stevan Rudinac, Alan Hanjalic, Martha Larson, (2013) “*Generating Visual Summaries of Geographic Areas Using Community-Contributed Images*”, IEEE Transactions On Multimedia, Vol. 15, No. 4, June 2013.

[10] T. G. Kolda and B. W. Bader, (2009), “*Tensor decompositions and applications*,” SIAM Rev., vol. 51, no. 3, pp. 455–500, 2009.

[11] Thomas Hofmann, (1999), “*Probabilistic Latent Semantic Indexing*”. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR ’99, pages 50–57, New York, NY, USA, 1999. ACM.

[12] S.A. Takale, P.J. Kulkarni, “*Extract Knowledge About Geo-Location Using Context and Content Information of Geo-Tagged Social Media*”, LNCS Proceedings, Part I, of the 16th International Conference on Web Information Systems Engineering - WISE 2015 - Volume 9418 Pages 601-615.

