International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970)
Volume-4 Number-2 Issue-15 June-2014

Representative and Diverse Image Set Gathering for Geographic Area and its Surrounding Region

Sonali B. Ulape[1], Sheetal. A. Takale[2]

Abstract

Geo-tagging is becoming famous in today's era of web 2.0 technologies. Geo-tagging is a concept in which we are adding information about location to the image. In proposed approach such geo-tagged images are used for making visual summary. Generally, visual summary of geographic location provides images of dominate location and ignore surrounding areas of that location which are potentially valuable. Multiple images of the same location suppress other places close to the dominant location and can be visited. In this proposed approach automatic visual summary of geographic area and its surrounding location is generated. For this community contributed images and their associated metadata are extracted from popular social websites such as Flickr [11]. Proposed approach generate visual summary which contains representative and diverse image set of given geographic location and its surrounding areas. For automatic generation of visual summary four layer graph of images and its associated metadata is formed. Proposed edge weight calculation method in this system is novel. With the help of random walk with restarts representative and diverse image set is selected which is visual summary of the given geographic location. Such visual summary can motivate a person to visit nearby areas of given geographic location.

Keywords

Geo-coordinates, Representative Image, Diverse Images, Visual Features, Visual Summary, Adjacency Matrix, Random Walk with Restarts (RWR), Algorithms, Multimedia.

This work was supported in part by the Department of Computer, Vidya Pratishthan's College of Engineering, Baramati, Pune, India.

Manuscript received May 21, 2014.

Sonali Babasaheb Ulape, Computer Department, Pune University, Vidya Pratishthan's College of Engineering, Baramati, Pune, India.

Sheetal. A. Takale, Information Technology Department, Pune University, Vidya Pratishthan's College of Engineering, Baramati, Pune, India.

1. Introduction

Concept of geo-tagging is introduced by flickr. Geo-tagging means attaching information of location to the image to know where this image has taken in the world. Location information is store by geo-coordinates i.e. latitude and longitude of the location. Obtaining geo-coordinates of location is possible with the help of GPS (Global Positioning System). Now a day's mobile, cameras are available from which person can take images with their geo-tagged information. Geo-tagging allows users to share images of locations they have visited with their friends with the help of social websites. It also helps to keep documentation of holidays by keeping together geo-tagged images of location. Flickr allow users to upload images with their geo-tags and place images on the world map. It also allow user to give title, description and tags to uploaded images and other people to comment their views on those images. Like general summary which is conveying large information shortly, proposed visual summary is giving information about circular geographic area shortly in the form of images. In this approach visual summary of geographic area and its surrounding locations is automatically formed. Generally when user is trying to search images of any location, many images of dominant location are extracted in result set. Such result set cannot give idea about surrounding locations of the given geographic area. Normally, there can be many places near single location which are less popular but potentially valuable e.g. good shop to buy famous products from that place, good museum to visit, good hotels to stay, libraries. Such nearby places are suppressing under most popular locations. For achieving this motive, this proposed approach is automatically generating visual summary of given geographic area and its surrounding locations. Due to such auto generated visual summary of geographic area people may be interested to visit unpopular nearby locations along with specified location. Such visual summary contains information about mainstream location i.e. given location in which user is interested along with non-mainstream location i.e. surrounding area of given location in which small group of users are interested. Such short visual summary will increase

time efficiency for selecting locations to visit along with its surrounding area.

This proposed approach is using community contributed images of users on social website like flickr and its associated metadata. Explicit metadata i.e. title, description, tags on the images and implicate metadata like uploader/commentator of image is extracted for each extracted image from flickr. Proposed system is graph based in which novel method for calculation of weight on links is proposed. This automatic weighting method helps to find out contribution of each modality i.e. text, users, visual feature of images to the performance of proposed visual summarizer system. With the help of random walk with restarts concept representative image of the given location and diverse set of images of surrounding areas are extracted in result set. Input to the proposed system is location which is specified by geo-coordinate of that location (i.e. latitude and longitude of the location). This geo-coordinate uniquely identifies that location on the world map. Output from the proposed system is formation of visual summary of given geographic area and its surrounding location. This paper is organized as: Section II introduced literature survey of previous work on visual summary. Section III represents proposed approach. Section IV contains experimental results. Finally Section V concludes the proposed system.

2. Literature Review

In [1] graph-based Mixed Media Graph (MMG) is constructed. In multimedia collection proposed MMG graph finds out the cross-media correlations between multimedia objects and their attributes both represented by nodes of the graph. Random walk with restarts (RWR) concept is used for estimating the similarities between objects in the graph.

In [2] for information retrieval, this approach proposes a retrieval model and learning framework. Framework builds based on semantically related multimedia content which are integrated with textual and visual feature. This approach constructs Image-Context Graph (ICG) for showing relation between features, images and terms.

In [3] Maximal marginal relevancy (MMR) ranking method is proposed. This method fetches relevant results to user query with minimum redundancy. Based on originality of information and query relevance, results from the ranked list are selected.
In [4] representative and diverse image set is generated with the help of image tags, landmark specific tags, image metadata and visual features of images. Landmark tags are extracted from various geographic areas. Global colour and texture features of images are extracted using local color moment, Gabor filter respectively. Local features are extracted using scale-invariant feature transform (SIFT). K-means clustering is applied and based on variety of views of landmark, clusters are ranked. . From these ranked clusters, images from each of these clusters are ranked.

In [5] weights of visual features i.e. colour, shape and texture of image are calculated for extracted images for the given user query. These weights are required for ranking function of clustering technique which returns final image set for the given search query. In this approach three clustering algorithms are proposed namely folding, Maxmin and reciprocal election. The folding algorithm makes use of already ranked search result for clustering. It selects representative cluster based on higher possibility of higher ranked images. Maxmin algorithm provides various representative images which are visually different from one another. In reciprocal election each image votes for the other image which it considers as best representative of itself. So the chances of being a representative is determine by all the obtained votes for images. This voting process is estimated by reciprocal ranks of ranked images.

In [6] enriched topic coverage retrieval results are obtained by making use of re-ranking method which is based on topic richness analysis. To evaluate topical convergence of the retrieved images diversity score method is proposed. These two methods help to improve the retrieved result set of images.

In [7] based on locations, Geo-tagged photos are clusters. For each cluster, representative images and tags are selected. For user search query (keyword of the interested place to visit or image of that place) this approach compares query against representative images or tags and provides final result.

In [8] method for overview generation of the location is proposed. On the web many travellers share their experiences of travelling. Proposed method uses this information to extract representative tags of location. These extracted tags will help to retrieve images from

web. Finally through user interface, extracted images and tags are return to users as location overview.
In [9]: various clusters of complex image query are discovered with the content-based image retrieval method. In this approach inputted image is parsed to get query point in the feature space which is actual query to the system. With the help of distance function, query point from image query is compared with images in dataset and user desired no of images return to user. These return images get relevance score. This paper proposes classifying process which keeps each element from relevance set to existing or new cluster. Another proposed cluster-merging process merges clusters for decreasing query points for the next iteration. Obtained cluster representative forms the new query point which is given as input to the next iteration. After some iteration final result set of clusters is obtained to the inputted query image.

3. Dataset

We have created our own dataset of geo-tagged images with the help of GPS. For this vidya pratishthan college, vidyanagari, baramati, pune location is selected. We have taken images of this college and its surrounding region by keeping GPS on. These images are uploaded with their co-ordinates to the flickr websites with deferent user name. Co-ordinates of images are representing latitude and longitude of the location where images are taken. To these images title, description, tags are given by uploader and comments are given by commentators. Commentators can also give tags to the images. This tag count can differs from zero to ten. Fig. 1 shows distribution of images on world map for GPS locations. These geo-tagged images are distributed on the world map as per latitude and longitude of the location.

Fig. 1. Distribution of location's images on world map.

4. Proposed System

In the proposed approach input is geographic location. Geo-coordinate of the location is given to the flickr. From defined geo-coordinates, images and its associated metadata of the given location and its surrounding circular region is extracted from flickr. Surrounding circular region is limited for the radius of 1 km from given geographic location. System architecture of the proposed approach is shown in Fig. 2.

4.1 Image-Attribute Graph Construction

Proposed visual summarizer system constructs four-layer graph in which node are retrieved images from flickr and its attributes. The edges represent weight connecting one node to another. This graph will help to calculate adjacency matrix A with is combination of 16 sub-matrices. As relation between nodes in the graph is symmetric undirected graph G= (V, E) is constructed in which V is set of nodes and E is set of edges. The graph is having 4 types of nodes and 2 types of edges.

Four types of edges in the graph are Image Node I= $\{i_1, i_2,\ldots, i_N\}$ - for each extracted image of the location, image node is assign. Visual feature Node F=$\{ f_1, f_2,\ldots f_N\}$ – for visual features of each image, visual feature node is assign. In proposed system two visual features of image are considered i.e. colour, texture features of image. Term Node T= $\{t_1, t_2,\ldots t_N\}$ - metadata of the image i.e. title, description, tags are gathered in the single term node. User Node U=$\{ u_1, u_2,\ldots u_N\}$ – for the user which have uploaded image or has given comments to somebody else's uploaded image, user node is assign. Fig. 3 shows a four-layer graph structure. So the final set of nodes for graph G will be: $V = I \cup T \cup F \cup U$

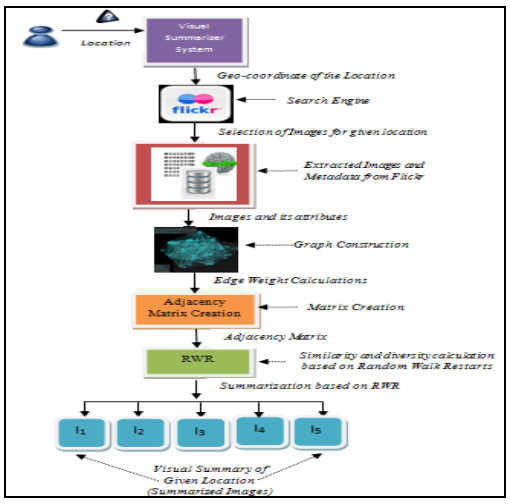

Fig. 2. System Architecture.

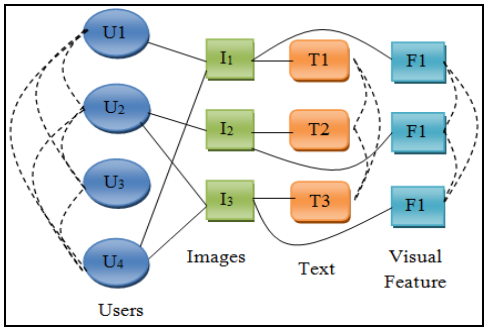

Fig. 3. Four-layer graph.

Two types of edges in the graph are similarity edges which are connecting nodes of the same types. These edges are represented by dashes line in Fig. 3. And attribute edges which are the edges connecting image nodes with their attributes i.e. image with text, visual features and users. These edges are represented by solid lines in Fig. 3.

4.2 Edge Weight Calculation

-Weight of edges connecting visual feature nodes of images is based on a visual feature similarity score of linked feature nodes and is calculated as [10]:

$$W_f(l,j) = sim(f_l, f_j) = \exp(-\frac{\|f_l - f_j\|^2}{2\sigma^2})$$

Where, W_f is $N \times N$ weight matrix. $N \times N$

-Weight of edges connecting user nodes of images is based on similarity of user. Similarity of the user is based on how many images are related to linked users. E.g. one user uploaded the image and another commented on it or both the users commented on the same image. So the user similarity measure is calculated as [10]:

$$W_u(l,j) = sim(u_l, u_j) = \frac{|I_l^{'} \cap I_j^{'}|}{|I_l^{'} \cup I_j^{'}|}$$

Where, W_u is the matrix $N_u \times N_u$ of user similarities and $I_l^{'}, I_j^{'} \subset I$ are the sets of images uploaded/ commented by the users.

-Weight of edges connecting text node is calculated using cosine similarity between vectors of TF-IDF weights [10]:

$$W_t(l,j) = sim(t_l, t_j) = \frac{t_l \times t_j}{\|t_l\|\|t_j\|}$$

Where, W_t is $N \times N$ weight matrix.

Adjacency matrix A of the graph is shown in Fig. 4. II, FT, FU, TU are matrices for which affinity between nodes is not known so these matrices are filled with zeros. FF, TT, UU are calculated using

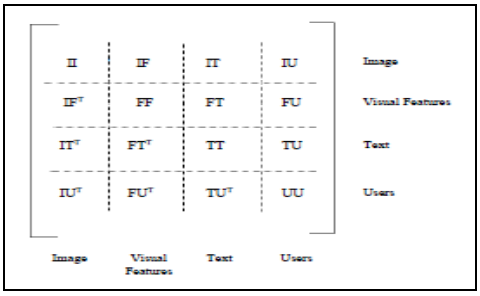

Fig. 4. Adjacency matrix A of the graph G.

formula $W_m^{'} = \beta_m W_m, m \in \{f, t, u\}$. This is giving weight on edges connecting features-features, text-text, and user-user node. IF, IT are calculated using formula $\beta_m I_N, m \in \{f, t\}$. This is giving weight on edges connecting image-feature and image-text node. IU is β_u if user uploaded or commented on the images otherwise it is assign to zero. All other matrices are transpose of their respective matrices. The adjacency matrix A is column-normalized i.e. the values in each column sum to 1.

4.3 Selection of representative and diverse images

For a selection of representation images Random walk with restarts (RWR) concept is used. Let R is random vector of dimension $R \times 1$. A is the adjacency matrix of dimension $R \times R$ of graph G. Random walker starts from any image node and repeats for all image nodes in the graph. Algorithm for selection of representative image is as:

Algorithm: Selection of Representative Images

Input- Adjacency matrix A
Output-Representative image

1. Let $R = 0$, for all N image nodes, except 1 for the starting image nodes.
2. Calculate steady state probability: $p = (1-c)Ap + cR, c = 0.5$
3. Repeat process 1 and 2 for all image nodes.
4. Store first N values of p for all image nodes in matrix $S_{N \times N}$

5. For arbitrary image i_l, compute the sum of its similarity to all other images in graph excluding (i_l) as: $x_l = \sum_{j=1:N, j \neq l} S_{lj}$
6. Store similarity score for all images in $X = [x_1, x_2, ..., x_N]$
7. Sort images by increasing order of X.
8.Positions of the image will be considered as representative score(RS)
9. Select the image with the highest RS in the target image set (optimal set). This is representative image.

After selection of representative image, diverse set of images are selected iteratively. These next selected images must be dissimilar from earlier selected images but at the same time they should have high representative scores. Algorithm for selection of diverse images is as:

Algorithm: Maximization of Set Diversity

Input: Adjacency matrix A, RS of images, previously selected image in OS
Output: Diverse set of images
1. Let $v = 1/|OS|$, for all previously selected image nodes, 0 for all other image nodes.
2. For all image nodes calculate steady state probability: $p = (1-c)Ap + cR, c = 0.5$
3.Store values of p for all images in $p^{(OS)} = [p_1^{(OS)} p_2^{(OS)} ... p_N^{(OS)}]$
4. Sort images by decreasing order of $p_j^{(OS)}$.
5.Positions of the images will be considered as diversity score(DS)
6. Select the image with the highest RS*DS and earlier not selected. This is image diverse image.
7. Repeat the process until desired no of summarized images are not selected.

5. Experimental Results

To find out automatic visual summary of inputted geographic location- VP College, Vidya Nagari Baramati, Pune India, geo-coordinate of the location need to find out. Below Fig. 5 shows name of location, its latitude, longitude, no of images need to extract from flickr and no of images required to summarize. Latitude, longitude is given to the flickr web site.

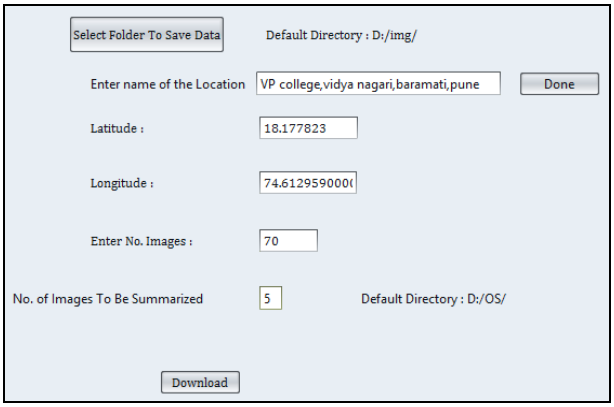

Fig. 5. Input to the visual summarizer system.

After doing all calculation parts i.e. image-attribute graph construction, edge weight calculations, adjacency matrix creation, selection of representative and diverse images with the help of random walk with restarts, final output is obtained.
For testing we have compared our proposed approach with three baseline methods:
Random: Random selection of visual summarized images from image collection of selected location.
View Count: Selection of visual summarized images with highest view count is selected.
Comment Count: Selection of visual summarized images with the largest number of comments on is selected.

Fig. 6. Comparison of several visual summaries for the VP College, Vidya Nagari, Baramati, Pune, India. Summary (a) illustrates the output by Random Method. Summary (b) illustrates the output by View Count Method. Summary (c) illustrates the output by Comment Count Method. Summary (d) illustrates the output by Proposed Method.

6. Conclusion

By using proposed approach a visual summary of given geographic area and its surrounding circular region is generated automatically. Such auto generated visual summary may increase interest of the people to visit unpopular nearby locations along with specified popular location. For this visual summarization of a geographic region, shared images by users and metadata associated with them is used. This approach proposes new technique for edge weight calculation which is novel. It provides informative, short summary of geographical region, which will experience good time efficiency of user for image retrieval.

References

[1] J.-Y. Pan, H.-J. Yang, C. Faloutsos, and P. Duygulu, "Automatic multimedia cross-modal correlation discovery," in Proc. 10th ACMSIGKDD Int. Conf. Knowl. Discovery Data Mining, 2004, pp. 653–658.

[2] J. Urban and J.M. Jose, "Adaptive image retrieval using a graph model for semantic feature integration," in Proc. 8th ACM Int.WorkshopMultimedia Inf. Retrieval, 2006, pp. 117–126.

[3] Carbonell and J. Goldstein, "The use of MMR, diversity-based reranking for reordering documents and producing summaries," in Proc. 21st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf., Retrieval, 1998, pp. 335–336.

[4] L. S. Kennedy andM. Naaman, "Generating diverse and representative image search results for landmarks," in Proc. 17th Int. Conf. World Wide Web, 2008, pp. 297–306.

[5] R. H. van Leuken, L. Garcia, X. Olivares, and R. Zwol, "Visual diversification of image search results," in Proc. 18th Int. Conf. World Wide Web, 2009, pp. 341–350.

[6] K. Song, Y. Tian, W. Gao, and T. Huang, "Diversifying the image retrieval results," in Proc. 14th Annu. ACM Int. Conf. Multimedia, 2006, pp. 707–710.

[7] L. Cao, J. Luo, A. Gallagher, X. Jin, J. Han, and T. Huang, "A worldwide tourism recommendation system based on geotagged web images," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., Mar. 2010, pp. 2274–2277.

[8] Q. Hao, R. Cai, X.-J.Wang, J.-M. Yang, Y. Pang, and L. Zhang, "Generating location overviews with images and tags by mining user-generated travelogues," in Proc. 17th ACM Int. Conf. Multimedia, 2009, pp. 801–804.

[9] Deok-Hwan Kim,ChinWan Chung, " Qcluster: Relevance Feedback Using Adaptive Clustering for Content Based Image Retrieval " SIGMOD 2003, June 912, 2003, San Diego.

[10] S. Rudinac, A. Hanjalic, and M. Larson "Generating Visual Summaries of Geographic Areas Using Community-Contributed Images" in IEEE Transaction on Multimedia 15(4), June 2013, pp. 921-932.

[11] Flickr.http://www.flickr.com/

Sonali Babasaheb Ulape has obtained her B.E. in 2010 and she is now persuing her M.E. in Computer Engineering from Vidya Pratishthan's College Of Engineering, Baramati.

Sheetal Ajay Takale has obtained her B.E. in 1998 and M.E. in 2005 in Computer Science and Engineering from Walchand College of Engineering, Sangli. She also qualified GATE 2011. She is now pursuing her Ph.D. in Computer Science and Engineering from Shivaji University Kolhapur. Sheetal has worked as a lecturer in Computer Engineering from 1998 to 2002 at M.I. T. Pune and from 2002 to 2007 at VPCOE, Baramati. Since 2007 to till date Sheetal has been working as an Assistant Professor and HOD at Vidya Pratishthan's college of Engineering , Baramati.