

# Improved Algorithm for Visual Summarization of Image Collections using Crowdsourcing and Infinite Push Ranking

<sup>1</sup>Rupali Waghmode, <sup>2</sup>Prajakta Shinde, <sup>3</sup>Deepali Lokare, <sup>4</sup>Ashvini Nikhade

Assistant Professor, ZCOER, Pune, India

## Abstract

New approach has been given for selection of images which are appropriate for part of the visual summaries. The novel method present crowdsourcing and Infinite Push method for performing Visual Summarization of image collections. Most of visual summarization algorithms are guided by notions of relevance, representativeness and diversity of visual content. Explicit information on human's perception regarding summarization process is not referred. This method is outlined on the base of how people generally thinking about generating summaries of image collections. Crowdsourcing experiment is performed for getting manually selected visual image summaries and method which help users for appropriate image selection. Based on insights of crowdsourcing experiment has purposed technique to select images automatically for generating visual summary. This technique utilizes the study various aspects of images like popularity of images, sentiment related to images, content and context of images. This technique defines images on the basis of their properties and semantic relation between images. It enables aesthetic features, sentiment analysis, and emotions associated with a specific group of images. Extraction of different features of images like luminance, colorfulness, sharpness, aspect ratio is done. These features are used to model image aesthetic appeal. The Infinite Push Ranking Method is used for training of crowdsourced dataset. This system is excellent in terms of performance as compared with existing RankSVM method.

**Keywords:** Crowdsourcing, Infinite Push Ranking, image aesthetic appeal

## 1. INTRODUCTION

Tremendous amount of digital data from digital media is accessible from content sharing and social networking websites and personal collection. Therefore there is need of powerful algorithms for representation, summarization and analysis of data for efficient retrieval and searching.

Summarization methods, provides a brief but comprehensive representation of multimedia data or text. Summaries may contain text, videos, segments, images. Existing visual summarization systems uses relevance, representativeness, diversity of visual content while developing summarization algorithm. The framework and details of the overview is depend on the goal it should fulfil, the exact evaluation of its quality can be evaluated with respect to its compatibility with the expectations of the individual users. Therefore the specific criteria containing the users perception of the summarization quality must be recognized and applied to guide the summarization algorithm. This approach present Crowd sourcing and Infinite Push method for performing Visual Summarization of image collections. User informed visual summarization is guided by notions of crowdsourcing study and Infinite Push pair wise ranking method. This improved algorithm is outlined on the basis of how people generally think while generating summarization of a set of images. For obtaining humans insights regarding summarization process crowdsourcing experiment is performed. This approach utilize the study various aspects of images like popularity of images, sentiment relates to images, content and context of images, image popularities. Here the extraction of different features of images like luminance, colourfulness, sharpness, aspect ratio is done. In crowdsourcing experiment first run a large scale crowdsourcing, try to acquire information in to how people perform visual summarization. Then utilize this information to determine on the correct functions, based on which images in the collection could be ranked. The position reflects the suitability of a image as a choice for addition in the summary, that's how probably a image could be picked for the overview by the users. The topic makes the following benefits:-

- Improved algorithm defines a new approach based on how human do selection of images for generation of visual summaries, which obtained with a inputs from crowd sourcing experiment.
- New approach uses properties of images such as semantic relations, aesthetic features and Infinite Push method. By using this approach generates images ranked with their suitability for inclusion in a visual summary.

## 2. RELATED WORK

### 2.1 Visual Summarization

In [1] present a new technique to summarization of visual information consists of images or video based on similarity technique. This technique depends on optimization of similarity method. Visual data is in the form of images, videos etc. Summary should capture essence of input data. It should not produce new visual artifacts. Bidirectional similarity method satisfies both the properties. Two signals are measured as visually similar only if all the patches of 1st image are present in 2nd image and vice versa. Use of those methods for solving problems like synthesis and completion of visual data, image collage, reshuffling of images and many. Generally generation of visual summaries plans to give concise details of a single video, set of video clips or a picture selection. In [2] they have done assume image analysis, tag data and images precise and implicit meta-data to draw out significant functions from community-contributed datasets. They utilize tags related with pictures by users and place meta-data to identify tags and place that signify milestone or geographical functions. Author performed visual research of pictures associated with discovered attractions to draw out associate sets of pictures for each milestone. Then they group the milestone pictures into creatively similar groups by applying different picture handling methods and produce links between those pictures that consist the similar visual things. In accordance with the k-means clustering method and on the produced link structure, they identify canonical opinions, and take the top associate pictures for each such view. This strategy helps for getting different and representative outcomes for landmarks queries. One of the demerits is concentrates on best opinions of the landmark only.

In [3] summarization is performed on user generated photos i.e images and travelogues and images on the websites. It provides structure of summarizing by using the rich textual and visual information from travelogue database. It selects location representative tags from travelogues representatives photos. The visualized tags and photos are lastly arranged appropriately to provide an informative summary in textual and visual format.

In [4] author had presented a strategy for automatically generation of visual summaries of geographical area. Topic discuss regarding a new retrieval strategy and learning structure for automated visual summarization of geographical place. While processing they yield geo-coordinates of specific place as feedback afterwards download pictures within set distance from Flickr website.

It takes meta-data, textual and visual methods of pictures. Algorithm signifies semantic interaction between pictures depending on user interaction. Here strategy takes Multimodal Image Context Graph (MICG) which brings together visual, textual and other methods together. The concept of random walking RWR over graph is applied to calculate diversity score, representative score.

### 3. IMPLEMENTATION DETAILS

The Visual Summarization system is having three important blocks as follows:-

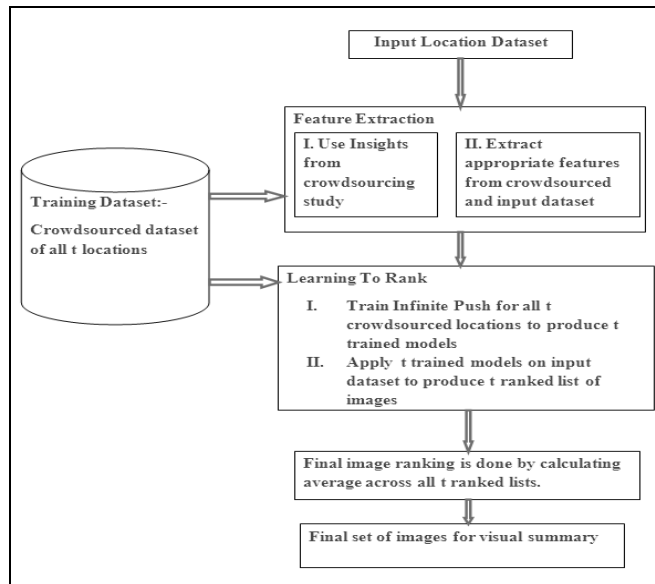


Figure1.1 System Architecture of Visual Summarization System

#### 3.1.Crowdsourcing Study

For 10 locations crowd sourcing study is done and reference summaries are generated. These manually generated reference summaries are as training dataset.

#### 3.2.Feature Extraction

On the basis of guidelines from crowd sourcing experiment feature extraction is done. Each image is defined based on its properties and the context of the other semantically related images. Each image is described with a feature vector. This system describe each image  $I$  with feature vector  $x_i$  based on its importance, popularity, aesthetic features. This is the preprocessing step. In this step the features such as luminance, sharpness, and aspect ratio are extracted from crowd sourced as well as input dataset. From the metadata of an image comments and view count is also extracted.

#### 3.3.Learning to Rank algorithm

Machine-learned ranking is typically supervised learning. MLR are important in the building of ranking models for information retrieval systems. In pair wise ranking training data consists of positive and negative samples within dataset. Training dataset is used by a learning algorithm to produce a ranking model which computes relevance of images or image collections. Here Infinite Push ranking method is used for training purpose. In Infinite Push ranking is a support vector machine based algorithm. It emphasizes on accuracy at the absolute top of the list. Infinite Push ranking method is a convex optimization problem. It uses a gradient and projection method. In the process of automatic selection of images for the summary we set a target to produce a ranked list of images per location, where the rank position of an image acts as a measure of its suitability for the visual summary.

Summarization algorithm works as follows:

- Human generated reference summaries from crowd sourcing study are used as training dataset for further process. This system takes the crowdsourced dataset of 10 locations as training dataset.
- By training the ranking function taking the features of images from crowd sourced dataset as the input. Here training with infinite push ranking is done for each location and  $t$  trained models are formed. Then approach start the training data selection by sorting the images per location based on collection subset selected for summaries. This approach selects a set of image preference pairs  $(i, j)$ , each containing of a top ranked and bottom ranked image.
- Finally in testing phase input location is taken and  $t$  ranked lists are formed by using infinite push trained model.
- At the end, a rank aggregation method is used to process the final set of ranked images. In rank aggregation final image ranking is done by calculating average across all  $t$  ranked lists. Based on the average value final rank position of image is computed.

**Table1.Results**

Location	Summary generation time using RankSVM	Summary generation time using Inifinite Push Ranking Algorithm
Chennai	4692	4359
Delhi	6392	5344
Kolkatta	3931	2510
Mumbai	3297	2390
Pune	2910	2219

Images in Input Dataset	Rank SVM Summary Generation Time in Miliseconds	Infinite Push Summary Generation Time in Miliseconds
10	1628	813
20	2347	1812
30	3394	2766

#### 4. CONCLUSION

New Visual Summarization approach select appropriate images for visual summary by taking insights from Crowdsourcing Study. Crowdsourcing study brings out features of images which are important for humans and provide training data. This approach uses Infinite Push Method to train crowdsourced dataset and generate list of images ranked by their compatability for inclusion in Visual Summary. The Visual Summarization System uses Crowd sourcing and Infinite Push ranking to enhance the performance. The Infinite push method is used to train crowdsourced dataset in order to minimize the summary generation time. The system compares the RankSVM Ranking Method and Infinite Push Ranking to improve the summary generation time. The result shows that training with Infinite Push ranking gives better performance over RankSVM Ranking algorithm. New approach shows a performance gain in summary generation time as compared to previous RankSVM algorithm. In future it is plan to investigate useful effective information from image title, tags and description generated by the uploader. It should be used in feature extraction phase.

#### References

- [1] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity", in Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. (2012), pp. 1-8.
- [2] L. S. Kennedy and M. Naaman, "Generating diverse and representative image search results for landmarks", in Proc. 17th Int. Conf. World Wide Web, ser. WWW08. New York, NY, USA: ACM, (2008), pp. 297-306.
- [3] Y. Pang, Q. Hao, Y. Yuan, T. Hu, R. Cai, and L. Zhang, "Summarizing tourist destinations by mining user-generated travelogues and photos", Comput. Vis. Image Understand., vol. 115, no. 3, (2011), pp. 352-363.
- [4] S. Rudinac, A. Hanjalic, and M. Larson, "Generating visual summaries of geographic areas using community contributed images", IEEE Trans. Multimedia, (2013)
- [5] G. Paolacci, J. Chandler, and P. G. Ipeirotis, "Running experiments on Amazon Mechanical Turk", Judgment and Decision Making, vol. 5, (2010), pp. 411-419.
- [6] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach", in Proc. 9th Eur. Conf. Computer Vision. Springer (2006), pp. 288-301
- [7] S. Siersdorfer, E. Minack, F. Deng, and J. Hare, "Analyzing and predicting sentiment of images on the social web," in Proc. Int. Conf. Multimedia, ser. MM10. New York, NY, USA: ACM, (2010), pp. 715-718.
- [8] D. R. Radev, E. Hovy, and K. McKeown, "Introduction to the special issue on summarization", Comput. Linguist., vol. 28, no. 4, (2002), pp. 399-408
- [9] Stevan Rudinac, Martha Larson, and Alan Hanjalic, "Learning Crowdsourced User Preferences for Visual Summarization of Image Collections", IEEE, (2013).
- [10] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "BLEU: a method for automatic evaluation of machine translation," in Proc. 40th Annu. Meeting on Association for Computational

- Linguistics, ser. ACL02. Stroudsburg, PA, USA: Association for Computational Linguistics, (2002), pp. 311-318.
- [11] C. Y. Lin, "ROUGE: A package for automatic evaluation of summaries," in Proc. ACL-04 Workshop Text Summarization Branches Out. Barcelona, Spain: Association for Computational Linguistics, (2004), pp. 74-81.
- [12] Y. Li and B. Merialdo, "VERT: automatic evaluation of video summaries," in Proc. Int. Conf. Multimedia, ser. MM 10. New York, NY, USA: ACM, (2010), pp. 851-854.
- [13] P. Over, A. F. Smeaton, and G. Awad, "The TRECVID 2008 BBC rushes summarization evaluation", in Proc. 2nd ACM TRECVID Video Summarization Workshop, TVS 08. New York, NY, USA: ACM, (2008), pp. 1-20.
- [14] G. Kazai, "In search of quality in crowdsourcing for search engine evaluation", in Advances in Information Retrieval, ser. Lecture Notes in Computer Science. Berlin, Germany: Springer, (2011), pp. 165-176.
- [15] T. Joachims, "Optimizing search engines using clickthrough data", in Proc. 8<sup>th</sup> ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, ser. KDD 02. New York, USA: ACM, (2002), pp. 133-142
- [16] O. Chapelle and S. S. Keerthi, "Efficient algorithms for ranking with SVMs", Inf. Retrieval, vol. 13, no. 3, (2010), pp. 201-215.
- [17] Agarwal, S., "The infinite push: A new support vector ranking algorithm that directly optimizes accuracy at the absolute top of the list". In Proceedings of the SIAM International Conference on Data Mining, (2011).
- [18] J. 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. and Develop. in Inform. Retrieval, ser. SIGIR 98. New York, NY, USA: ACM, (1998), pp. 335-336.
- [19] S. Siersdorfer, S. Chelaru, W. Nejdl, and J. S. Pedro, "How useful are your comments? analyzing and predicting YouTube comments and comment ratings", in Proc. 19th Int. Conf. World Wide Web, ser. WWW 10. New York, NY, USA: ACM, (2010), pp. 891-900.
- [20] C. Dwork, R. Kumar, M. Naor, and D. Sivakumar, "Rank aggregation methods for the web", in Proc. 10th Int. Conf. World Wide Web, ser. WWW 01. New York, NY, USA: ACM, (2001), pp. 613-622.
- [21] V. Pihur, S. Datta, and S. Datta, "Rankagg, an R package for weighted rank aggregation", BMC Bioinform., vol. 10, no. 1, (2009), pp. 62-62.