ISSN 2278-3091
Volume 8, No.1.2, 2019
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse2381.22019.pdf
https://doi.org/10.30534/ijatcse/2019/2381.22019

Semantically enriched Tag clustering and image feature based image retrieval system

Saravana Balaji B[1], Krishna Kumar V[2], Ahmed Najat Ahmed[3]
[1] Assistant Professor, Dept of Information Technology, Lebanese French University, Erbil, KR-Iraq
saravanabalaji.b@lfu.edu.krd
[2]Assistant Professor, Department of Computer Science & Eng., Sri Ramakrishna Engg College, Coimbatore, India
krish232007@gmail.com
[3] Assistant Lecturer, Department of Computer Network, Lebanese French University, Erbil, KR-Iraq
a.afandy@lfu.edu.krd

ABSTRACT

Image retrieval systems are always a key research area in the academia and in industry. Image processing systems are playing a pivotal role in many domains like medical etc. The WWW is a core collection of trillions of images which are searched by many users for usage. But the traditional search engines cannot provide relevant images according to the user queries as they are having their own issues. To overcome this, in this paper, an image retrieval system based on tag clustering and image features is explained. This systems clusters the images based on the tags associated with the images and uses the image features like color, texture and shape. Then the user query is matched for its membership value in the tag clusters and based on the membership value the imagers are retrieved. The performance measure of the system provides consistent results for different queries.

Key words: CBIR, Image Retrieval, Tag Clustering, Visual Features.

1. INTRODUCTION

A image retrieval structure is a system for scrutinizing, looking and recuperating pictures from a broad database of relevant pictures. Most standard and essential strategies for picture recuperation utilize some procedure for including metadata, for instance, engraving, watchwords, or depictions to the photos with the objective that recuperation can be performed over the clarification words. The extension in social web applications and the semantic web have roused the progression of a couple of electronic picture remark contraptions. The image look techniques are of two sorts. Meta data based picture interest and substance based picture recuperation. Metadata based picture look for uses names, catchphrases, customer content and other information that are connected with the image in the database.The substance based picture recovery technique depends on likenesses in their substance (surfaces, hues, shapes and so forth.) to a client provided question picture or client indicated picture highlights. Free of which seek conspire is sent, a picture web index for the most part works in two fundamental advances: the disconnected and the online advance. For some inquiry keywor,ds, the picture recovery execution is great, however the exactness of the returned outcomes is still generally low. They experience the ill effects of the uncertainty of question catchphrases since it is troublesome for clients to precisely depict the visual substance of target pictures just utilizing inquiry watchwords. One of the real difficulties is the contention between the substance of the picture and the site page printed data. This paper endeavors to determine this bottleneck by relying upon both the literary data and visual data. Another significant test in the current frameworks is that its similitudes of low dimension visual highlights may not correspond with the h the picture's abnormal state semantic implications. To diminish this semantic hole, visual highlights are mapped to predefined qualities known as visual semantic descriptors. The greater part of the current frameworks for web picture recovery have a question situated point of view and are not client arranged. A definitive objective of any recovery framework must be based and coordinated according to the client's decision, in this way fulfilling the client's requirement for pictures. Certain current frameworks which catch client's inclinations still don't make an imprint as they disregard the point of view of the client to give the best outcomes. This improves the clamor of web picture inquiry and expands the superfluity of pictures recovered with regards to web picture look through that should be survived. In CBIR frameworks, highlights are consequently removed from picture pixels and utilized inside explicitly chosen similitude measures so as to recover comparable pictures. Given the component extraction level, we recognize two principle CBIR approaches: worldwide and nearby methodologies. The worldwide methodology depends on visual highlights registered overall picture, though, the neighborhood approach depends on depictions of each picture districts (objects) of the picture [1]. Hence in this paper, we have proposed an image retrieval system which is based on tag clustering of images and image features. This system will first cluster the images based on the tag clustering technique so that the search time will be reduced. Then it will check for image features of the cluster by having the SVM based feature extraction which will fine grain the image cluster with related images only. Section II discusses some related work regarding the problem. Section III describes the system operation, and section IV describes the performance evaluation of the system compared with a few other systems. Section V concludes the paper.

2. RELATED WORK

Various investigations have been carried on semantic picture grouping. Some of spotlights have been on ordering pictures, explanations, significance criticism, and cosmology. Recovery frameworks are created by fusing expansive visual codebooks and spatial data [16-18], and incorporate inquiry based [19, 20] and agent approaches [15, 21], among others [22, 23]. Rasiwasia et al. [24] examined two critical research questions: 1) in the case of demonstrating relationships viably between printed segments and visual substance parts can help by and large pursuit execution and 2) regardless of whether the displaying can be progressively compelling in highlight spaces with semantic theoretical. The exact outcomes demonstrate that the look frameworks representing cross-methodology relationships and abnormal state reflection show a promising exhibition. Despite the fact that this technique understands the objective of questioning picture content through printed watchwords, the scholarly factual model just gets the inactive connection among content and visual segments. The semantic gap between picture content and semantic ideas stays unsolved.

3. PROPOSED SYSTEM

The proposed systems consist of three parts where the first component is to cluster the given dataset of the image based on tags corresponding with them. The second component is for fine grain the cluster of images by comparing the image features with the corresponding test image. The third component is for retrieving the images from the clusters by analyzing the user query which is given in natural language.

3.1Tag Clustering

This stage starts with a realizing, when an effectively named preparing dataset is utilized by AI procedures to make a measurable model for every idea from the objective vocabulary. The models are then utilized amid the genuine comment stage to choose the importance of individual ideas as for a given question picture. As reviewed for example in [2], various learning systems have been examined in setting of picture comment. As of late, great outcomes have been gotten by profound convolutional neural system classifiers [3, 4]. In the subsequent stage, the catchphrases that showed up in the explanations of pictures in Simq are collected. We will allude to them as the underlying competitor watchwords Kw. These applicant watchwords are related with introductory probabilities, which are normally gotten from certain properties of Sim_q. Extra content mining systems would then be able to be connected to grow and refine the arrangement of applicant watchwords and to recompute the enrollment of individual competitors. Toward the second's end stage, we acquire Kw_{Final} – a last rundown of catchphrases and their membership esteems.

3.2 Image Feature Extraction

The image features normally considered for extraction by any image retrieval systems are color, shape, texture etc. These three features are extracted as follows:

1) Color: It is a standout amongst the most broad vision qualities because of its nearby connection with picture items, forefronts, and foundations. The shading is likewise a vigorous visual component as it doesn't rely upon the condition of picture substance, for example, the heading, size and edge. The prominent shading portrayals that have mostly been utilized are shading histogram, shading minutes [5], shading correlogram [6], and shading co-event network [7].

2) Texture: In PC vision, there is no exact meaning of picture surface, however it very well may be characterized as all what is left in the wake of thinking about hues and shapes, or as a depiction of picture structure, haphazardness, granulation, linearity, harshness, and homogeneity. Picture surface is a vital picture include for depicting intrinsic surface properties of a specific article and its association with the encompassing locales [8]. Some ordinarily utilized calculations as surface descriptor are Gabor channels, Wavelet changes, dark dimension co-event framework (GLCM) [9], Markov arbitrary field (MRF) [10], edge histogram descriptor (EHD) [11], steerable pyramid deterioration (SPD) [12], and Tamura highlights [13]. The technique adjusted for the advancement of this framework is [14] have presented a probabilistic surface recovery approach. It depends on the picture portrayal in the unpredictable wavelet area and a few factual models for the extent of the complex change coefficients. Moreover, this methodology incorporates shut structure articulations for the KL-divergences between the proposed factual models which permit consistent intricacy closeness estimations.

3) Shape: Image shape highlight essentially conveys semantic data and can be extensively sorted as limit based and locale based. The limit put together strategy removes highlights based with respect to the external limit of the locale while the area put together concentrates highlights based with respect to the whole district. By and large, shape-based recovery strategies experience the ill effects of issues related with the interpretation, scaling, revolution invariances and the soundness with slight changes fit as a fiddle. In ensuing, shape descriptors are generally extricated and utilized with different highlights, for example, shading and surface and will in general be proficient in explicit applications, for example, man-made articles. Shape descriptor can be spoken to utilizing numerous basic techniques, for example,

polygonal estimate, Fourier descriptors, invariant minutes, deformable formats, B-splines, curvature scale space (CSS), viewpoint proportion, circularity, and back to back limit portions.

Clustering of Images

The KMeans algorithm used in this clustering approach. The steps are given below

1. Number of clusters required X.
2. Make X clusters randomly with tags and calculate center point of each cluster, or randomly choose Y points as initial center point.
3. Based on the nearest center point, assignment each point to the corresponding cluster.
4. Re-iterate for the center point Y of the cluster.
5. Repeat step number 3 and 4 until the convergence criterion achieved (there is no membership change in each cluster or until some number of iteration)

For the above steps the features mentioned above (color, shape and texture) are used and based on this the images are clustered.

3.3 Image Retrieval based on User Query

The semantic closeness estimation takes puts between homologous parts that are material to the request similarly as the thought dimension substances of the gathering. Further, the semantic comparability is prepared between the inquiry parts and the genuine thoughts isolated from the homologous thoughts by using semantic proportionality organizing. An InfoPath is figured using the semantic likeness measures and the depiction methods of reasoning of the name substances. The InfoPath molded is a dynamic change of the planning marks subject to the semantic similarity regards and the image delineations. A dimension mapping of the gathering parts in the InfoPath and the photos subject to picture marks, picture metadata or the watchwords of the image content page is cultivated. The level mapping is cultivated by the semantic likeness figuring.

The contextual information of Kwi and Kwj, denoted by $cont(Kwi,Kwj)$, is defined by:

$$cont(Kwi,Kwj) = log\ P(Kwi,Kwj) \qquad (1)$$

$P(Kwi)$ represents the appearance membership of the keyword Kwi in the database image. $P(Kwi,Kwj)$ represents the joint membership of the two keywords Kwi and Kwj together

4. EXPERIMENTAL EVALUATION

The informational pictures for experimentation are gathered from the consequences of various picture web search tools. The experimentation was done for 2314 pictures; out of which 1814 pictures were consequently crept utilizing a tweaked picture crawler and the rest of the pictures were physically gone into the database. Every one of the pictures were gathered with their marks or labels. The precedent groups are Apple, Tank, and Horse separately. The performance of the system is calculated using the well-known formulas given below.

Precision = No of Image Retrieved/ Total Image Retrieved (2)

$$Recall = \frac{No\ of\ Relevant\ Images\ Retrieved}{Total\ No\ of\ Relevant\ Images} \qquad (3)$$

$$Accuracy = \frac{Precision + Recall}{2} \qquad (4)$$

The above mentioned values are given in the table below. From the table it is evident that the proposed system is performing well in complex queries.

Table 1: Performance Evaluation

S.No	Query	Precision	Recall	Accuracy
1	Apple	96.2	97.4	96.8
2	Tank	95.4	96.7	96.05
3	Horse	97.5	98.1	97.8

5. CONCLUSION

The image searching in the WWW is always an open problem. There are many image searching techniques like content based image retrieval, keyword based image retrieval are proposed. All of them have their own issues. In this proposed system, a combination of image visual features and their corresponding tags are clustered. Then the user query is matched for the membership of the cluster and based on that the images are retrieved as result. The performance evaluation of the systems shows that it provides consistent results for different types of user queries given in natural language.

REFERENCES

1. Liu Y, Zhang D, Lu G, Ma W-Y (2007) **A survey of content-based image retrieval with high-level semantics**. Pattern Recogn 40(1):262–282 https://doi.org/10.1016/j.patcog.2006.04.045

2. Zhang D, Islam MM, Lu G (2012) **A review on automatic image annotation techniques**. Pattern Recogn 45(1):346–362 https://doi.org/10.1016/j.patcog.2011.05.013

3. Krizhevsky A, Sutskever I, Hinton GE (2012) **ImageNet classification with deep convolutional neural networks**. In: Advances in neural information processing systems (NIPS 2012), pp 1106–1114

4. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) **Going deeper with convolutions**. In: IEEE conference on computer vision and pattern recognition (CVPR 2015), pp 1–9 https://doi.org/10.1109/CVPR.2015.7298594

5. D. Xiaoyin, **Image retrieval using color moment invariant**, The Seventh International Conference on

Information Technology: New Generations (ITNG), Las Vegas, NV, 12–14 (2010), pp. 200–203.
6. H. Jing, SR Kumar, M. Mitra, WJ. Zhu, R. Zabih, **Image indexing using color correlograms**, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Proceedings (1997), pp.762-768.
7. G.P. Qiu, Color image indexing using BTC, IEEE Transactions on Image Processing, 12 (1) (2003), pp. 93–101 https://doi.org/10.1109/TIP.2002.807356.
8. N. Shrivastava, V. Tyagi, **Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching**, Information Sciences 259 (2014), pp. 212-224.
https://doi.org/10.1016/j.ins.2013.08.043
9. R. Haralick, **Statistical and structural approaches to texture**, Proceedings of the IEEE 67 (1979), pp. 786–804.
https://doi.org/10.1109/PROC.1979.11328
10. G. Cross, A. Jain, **Markov random field texture models**, IEEE Transactions on Pattern Analysis and Machine Intelligence, 5 (1) (1983), pp. 25–39.
https://doi.org/10.1109/TPAMI.1983.4767341
11. DK. Park, YS. Jeon, CS. Won, **Efficient use of local edge histogram descriptor**, In Proceedings of the ACM workshops on Multimedia, ACM (2000), pp. 51-54
https://doi.org/10.1145/357744.357758
12. EP. Simoncelli, WT. Freeman, The steerable pyramid: **A flexible architecture for multi-scale derivative computation**, In International Conference on Image Processing, IEEE Computer Society, vol. 3 (1995), pp. 3444-3444.
13. H. Tamura, S. Mori, T. Yamawaki, **Textural features corresponding to visual perception**, Systems Man and Cybernetics, IEEE Transactions on 8, no. 6 (1978), pp. 460-473.
https://doi.org/10.1109/TSMC.1978.4309999
14. R. Kwitt, A. Uhl, **Lightweight probabilistic texture retrieval**, IEEE Transactions on Image Processing, 19(1) (2010), pp. 241-253
https://doi.org/10.1109/TIP.2009.2032313
15. Jegou H, Perronnin F, Douze M, Sanchez J, Perez P, Schmid C (2012) **Aggregating local descriptors into compact codes**. In: PAMI
https://doi.org/10.1109/TPAMI.2011.235
16. Philbin J, Chum O, Isard M, Sivic J, Zisserman A (2007) **Object retrieval with large vocabularies and fast spatial matching**. In: CVPR
https://doi.org/10.1109/CVPR.2007.383172
17. Avrithis Y, Kalantidis Y (2012) **Approximate gaussian mixtures for large scale vocabularies**. In: ECCV
https://doi.org/10.1007/978-3-642-33712-3_2
18. Hinami R, Matsui Y, Satoh S (2017) **Region-based image retrieval revisited.** https://doi.org/10.1145/3123266.3123312
19. Chum O, Mikulik A, Perdoch M, Matas J (2011) **Total recall II: query expansion revisited**. In: CVPR
https://doi.org/10.1109/CVPR.2011.5995601
20. Tolias G, Jegou H (2014) **Visual query expansion with or without geometry: refining local descriptors by feature aggregation.** Pattern Recognit
https://doi.org/10.1016/j.patcog.2014.04.007
21. Perronnin F, Liu Y, Sanchez J, Poirier H (2010) **Large-scale image retrieval with compressed fisher vectors**. In: CVPR
https://doi.org/10.1109/CVPR.2010.5540009
22. Radenovi F, Jegou H, Chum O (2015), **Multiple measurements and joint dimensionality reduction for large scale search with short vectors,** In: ICMR
https://doi.org/10.1145/2671188.2749366
23. Tolias G, Furon T, Jegou H (2014) **Orientation covariant aggregation of local descriptors with embeddings**. In: ECCV
https://doi.org/10.1007/978-3-319-10599-4_25
24. N. Rasiwasia *et al.*, **A new approach to cross-modal multimedia retrieval**, in *Proc. Int. Conf. Multimedia*, 2010, pp. 251–260
https://doi.org/10.1145/1873951.1873987
25. P. Ashok Babu, **Super Resolution Image Reconstruction for single image using Approximate BPTSRIRTD Algorithm**, International Journal of Advanced Trends in Computer Science and Engineering, Volume 7, No 6, 2018, pp. 144-148.