iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 3 **Issue:** VI **Month of publication:** June 2015

DOI:

www.ijraset.com Volume 3 Issue VI, June 2015
IC Value: 13.98 ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Re-Ranking Of Web Images Using Semantic Signature and Parallel SVM

Nisha Abdul[1], Kiran V K[2]
[1,2]Dept. of Computer Science & Engineering,
NSS College of Engineering, Palakkad, Kerala

Abstract*— Image re-ranking has been adopted by current commercial search engines to improve the results of web-based image search. In image re-ranking, for a given a query keyword, first a pool of images are retrieved by the search engine and the user is asked to select a query image from the pool. The remaining images are re-ranked based on their visual similarities with the query image. But there are difficulties in understanding users' search intention and in learning a universal visual semantic space to characterize the web images. An image re-ranking framework is proposed which automatically offline learns different visual semantic spaces for different query keywords through keyword expansions. The visual features of images are projected into their related visual semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the visual semantic space specified by the query keyword. An image can have a single semantic signature or multiple semantic signatures for a specific keyword. Computing multiple semantic signatures for an image has higher precision and re-ranking accuracy than using single semantic signature. But the lengthy combined semantic signatures results in more time consumption. Hence a number of SVMs (Support Vector Machines) can be used in parallel at the online stage which compares multiple semantic signatures at the same time and results are merged at the later stages. Thus images are re-ranked in faster manner while preserving the accuracy.

***Keywords*— image re-ranking, SVM, semantic signature, keyword expansion, reference class**

I. INTRODUCTION

Nowadays the users all around the world are depending on the web images for varied purposes. This has led to the development of image search engines such as Google Image Search and Microsoft Live Image Search. Most of the search engines use keywords as queries and rely on surrounding text of images to search for them. Such keyword based searching is suffering from the ambiguity of keywords, because it is hard for users to accurately describe the visual content of images only using keywords. This results in an inefficient and imprecise output. For example, if the query keyword is "apple", the search engine will retrieve images belonging to different categories such as "apple tree", "red apple", and "apple iPhone" [1].

One way to improve the keyword based search is known as image re-ranking. The web-based image search used by Bing and Google uses image re-ranking [2]. Image re-ranking is used widely since image based features are more important than text based features. In image re-ranking, first the user has to give a query keyword. A pool of images which are relevant to the given keyword will be retrieved according to a stored word-image index file. But the image pool will contain images from different categories. Secondly the user is asked to select a query image from the pool and the remaining images in the pool are re-ranked based on their visual similarities with the query image [1]. Thus a reordering of images is taking place. The query image reflects the user's search intention more deeply than query keyword. Both query keyword and query image take part in the re-ranking of images. The word-image index file and visual features of images are precomputed at the offline stage and stored by the search engine. This is what happens in conventional image re-ranking framework. A major challenge of image re-ranking is that even if the visual features of images are similar, they may not be semantically similar. The main online computational cost of image re-ranking is on comparing visual features. But comparing the visual features at the online stage is a tedious task and it increases the computational cost. One solution is to map visual features to a set of concepts as semantic signature. But it is difficult to design a huge concept dictionary which considers the diverse images on web [1]. A framework can used for web image re-ranking which automatically learns different semantic spaces for different query keywords at the offline stage, instead of manually designing a universal concept dictionary. Each keyword is associated with a semantic space of related concepts .Each image is associated with multiple keywords and thus to multiple semantic spaces. The visual features of images are projected into their related semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures for corresponding query keyword. The semantic correlation between concepts is explored when computing the similarity of semantic signatures. This query-specific semantic signatures can significantly improve both the accuracy and efficiency of image re-ranking. Semantic space of a query keyword can be described by just 25 concepts on average. Therefore the semantic signatures are very short and online image re-ranking becomes extremely efficient. These concepts that constitute the semantic space are also called as reference classes. The semantic spaces of query keywords are automatically learned through keyword expansion because of the large number of keywords and the dynamic variations of the web [1].

II. RELATED WORKS

Image retrieval has become an important requirement for internet users. A number of image retrieval systems have been developed over the past decades. They include text-based image retrieval (TBIR), content-based image retrieval (CBIR), user's relevance feedback and adaptive similarity.

A. Text-Based Image Retrieval

The text-based image retrieval has been used widely by popular search engines. In such retrieval, the user has to input a keyword as a textual query to the retrieval system. Then the system returns the images ranked in the order of surrounding texts containing the given query keyword The ranking score is obtained according to some similarity measurements such as cosine distance between the query keyword and the textual features of relevant images [3]. The TBIR system may retrieve irrelevant and disorganized images and it does not consider user's search intention.

B. Content-Based Image Retrieval

Content-based image retrieval (CBIR) is also known as query by image content and content-based visual information retrieval. Google is one of the search engines that work on content based Image re-ranking. CBIR systems overcomes the difficulties of text based approach since it uses the visual content of the images such as color, texture and shape features for image re-ranking [4][5]. In CBIR, the user has to input a query image to the system and then it will return images in the order of similarity of contents with the query image. Content-based image search analyses the contents of the image rather than the textual data such as keywords, tags, or descriptions associated with the image. It is not at all possible for human to enter the keywords for images manually. So a method which compares the visual content is more desirable than identifying the keywords associated with the images [3][6]. In content-based image retrieval, the fundamental problem of finding semantically similar images remains largely unsolved since it considers the visual similarities only. CBIR systems do not consider user's search intention [4].

C. User's Relevance Feedback

A relevance feedback based interactive retrieval approach requires the user to select multiple relevant and irrelevant image examples. CBIR systems can make use of relevance feedback, where the user progressively refines the search results by marking images in the results as "relevant", "not relevant", or "neutral" to the search query, then repeating the search with the new information. This helps CBIR systems to be aware about the user's search intention more deeply. Through the submission of desired images or visual content-based queries, the re-ranking for image search results can achieve significant performance improvement [1][6]. In the relevance feedback based approach, the retrieval process is interactive between the computers and human. User's relevance feedback considers user's search intention but it requires more user effort and online training is needed. For web-scale commercial systems, users' feedback has to be limited to the minimum without online training. The most challenging problem in this framework is in defining the similarity.

D. Adaptive Similarity

Adaptive similarity is an image re-ranking approach which limited users' effort to just one-click feedback. In adaptive similarity, a query image is first categorized into one of the predefined intention categories, and a specific similarity measure is used inside each category to combine image features for re-ranking based on the query image. In order to characterize images from different perspectives, such as color, shape, and texture, adaptive similarity adopt and design a set of features that are both effective in describing the content of the images. The main features include Attention Guided Color Signature, Color Spatialet, SIFT (Scale-invariant Feature Transform), Histogram of Gradient, Facial Feature. In adaptive similarity no online training is required. Just one-click is enough for the simple user interface. But the query image may be classified into wrong category and the available categories can't cover all web images [7].

III.EXISTING IMAGE RE-RANKING SYSTEM

The major challenge of image re-ranking is that similarities in visual features do not indicate the semantic similarity. It is impractical to compare the visual features of thousands of dimension at the online stage. One solution is to map visual features to a set of concepts as semantic signature. But it is difficult to design a huge concept dictionary which considers the diverse images on web. Web image re-ranking using query specific semantic signature is an existing system which learns different visual semantic spaces for different query keywords individually and automatically. Thus avoids the requirement of designing a universal concept dictionary which covers the diverse images on the web. The system stores the semantic signatures and classifiers of reference classes instead of the visual features. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the visual semantic space specified by the query keyword. Thus the original visual features of thousands of dimensions can be projected to shorter semantic signatures. Each keyword is associated with a semantic space which contains the related concepts of that keyword. Because of the large number of keywords and the dynamic variations of the web, the semantic spaces of query keywords are individually and automatically learned through keyword expansion. The query-specific semantic spaces can help to accurately re-rank the images, since they have excluded the irrelevant concepts. The visual features of images are then projected into their related semantic spaces to get semantic signatures. The framework has offline and online parts. At the offline stage, the reference classes (which represent different concepts) related to query keywords are automatically discovered and their training images are automatically collected. For each keyword, its reference classes are defined by finding a set of keyword expansions that are most relevant to it. To achieve this, a set of images are retrieved by the search engine using the keyword as query. Keyword expansions are found from the words extracted from the images in the set. A keyword expansion is expected to frequently appear in that set. For each image I in the set, all the images are re-ranked according to their visual similarities to I. A number of most frequent words among top re-ranked images are found. If a word w is among the top ranked image, it has a ranking score rI (w) according to its ranking order; otherwise rI (w) = 0. The overall score of a word w is its accumulated ranking scores over all the images. Words with highest scores are selected and combined with the original keyword to form keyword expansions, which define the reference classes. Keyword expansions are used to retrieve images belonging to the reference classes. In order to improve the efficiency of online image re-ranking, redundant reference classes are removed. To compute similarity between two reference classes, we use half of the data in both classes to train a binary SVM classifier to classify the other half data of the two classes. If they can be easily separated, then the two classes are considered not similar. Finally a set of reference classes are selected from the candidates.

For each query keyword, its reference classes forms the basis of its semantic space. Given a set of N reference classes for keyword q, a multi-class classifier on the visual features of images is trained and stored offline. The multi-class classifier outputs an N-dimensional vector p, indicating the probabilities of a new image I belonging to different reference classes. The semantic signature of an image is extracted by computing the similarities between the image and the reference classes of the query keyword using the trained multiclass classifier. Then p is used as semantic signature of I. The distance between two images I^x and I^y are measured as the distance between their semantic signature p^x and p^y. An image may be associated with multiple query keywords and thus to multiple semantic spaces. Therefore, it may have different semantic signatures. The query keyword input by the user decides which semantic signature to choose. Suppose an image is associated with three keywords “apple,” “mac” and “computer”. When using any of the three keywords as query, this image will be retrieved and re-ranked. However, under different query keywords, different semantic spaces are used. Therefore an image could have several semantic signatures obtained in different semantic spaces. They all need to be computed and stored offline. At the online stage, a pool of images is retrieved by the search engine according to the query keyword. Since all the images in the pool are associated with the query keyword according to the word-image index file, they all have pre-computed semantic signatures in the same semantic space specified by the query keyword. Once the user chooses a query image, its semantic signature for that keyword is compared with semantic signatures of remaining images for the same keyword. Images will be re-ranked according to the similarity of its semantic signature with that of query image. These semantic signatures are used to compute image similarities for re-ranking.

There are two different ways of computing semantic signatures.

Query-specific visual semantic space using single signatures (QSVSS Single). For an image, a single semantic signature is computed from one SVM (Support Vector Machines classifier trained by combining all types of visual features.

Query-specific visual semantic space using multiple signatures (QSVSS Multiple). For an image, multiple semantic signatures are computed from multiple SVM classifiers, each of which is trained for one type of visual features.

A simple idea is to combine all types of visual features to train a single powerful SVM classifier which better distinguish different reference classes. If there are K types of visual features, such as color, texture, and shape, first approach combine them together to train a single classifier, which generates one semantic signature for an image. It is also possible to train a separate classifier for each type of features. Then, the K classifiers based on different types of features extract K semantic signatures, which are combined at the later stage of image matching. But second approach can increase the image re-ranking accuracy. If only single classifier is trained combining all types of visual features the semantic signatures are of 25 dimensions on average. If separate classifiers are trained for different types of visual features, the semantic signatures are of 100-200 dimensions based on number of features considered. Computing multiple semantic signatures has higher precision and re-ranking accuracy than using single semantic signature. But the disadvantage is that length of combined semantic signatures is more and thus comparison is

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

time consuming [1].

IV. PROPOSED SYSTEM

The new framework is an extension of web image re-ranking system that uses semantic signature. The proposed system automatically offline learns different semantic spaces and considers users' search intention such that less user effort and no online training is required. It does not have a set of predefined categories so that image may be classified into wrong category. The accuracy of image re-ranking is improved with the use of multiple semantic signatures and produces the result in a faster manner using parallel SVMs at the online stage. In web image search engines, semantic signature can be computed in two ways. First using single SVM classifier which is trained by combining all of the visual features. Second using separate SVM classifier for each type of visual feature. The second method has high precision and re-ranking accuracy than computing a single semantic signature. The new approach is aimed at using multiple semantic signatures by overcoming the disadvantage of time consumption while comparison. Instead of comparing the lengthy combined multiple semantic signatures, individual semantic signatures are compared in parallel using parallel SVM and the results are merged at later stages. Merging provides images which are highly correlated with query image. This approach produces faster result with accuracy and precision. It avoids the time consumed during comparing the lengthy semantic signature. Thus accurate result with almost same time as using single semantic signature can be obtained. Support Vector Machines are powerful classification and regression tools, but their compute and storage requirements increase rapidly with the number of training vectors [8].

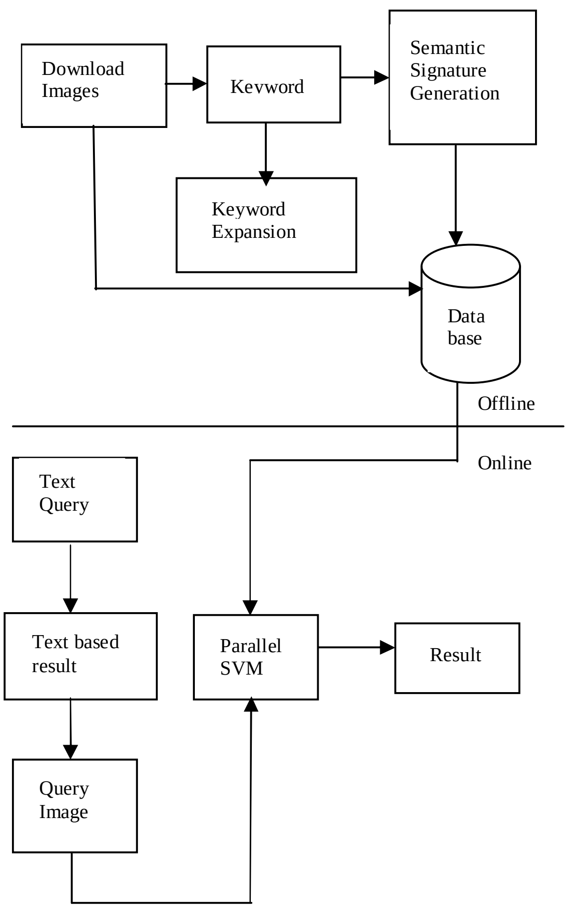


Fig.1.System Architecture

At the online stage, once a query image is selected, its semantic signatures are compared with semantic signatures of the remaining images in the pool for each visual feature parallely. One SVM classifier is used for each visual feature. The SVM classifier will return images whose semantic signature is highly correlated with that semantic signature of the query image for that visual feature. The partial results from parallel SVMs are merged and filtered at later stages in cascade of SVMs. This will rank images in the order of correlation with query image. Fig. 1 shows the system architecture of new approach. At the offline stage, images are downloaded using their URLs provided in the dataset. Keywords are expanded to find reference classes. If new reference classes are found, they are added to database. Otherwise image is included in already existing reference classes.

For a downloaded image, multiple semantic signatures are generated for multiple visual features and this is repeated for each keyword of that image. These semantic signatures and multiclass classifiers are stored in the database. At the online stage, a query image is selected and its all semantic signatures are compared with semantic signatures of other images in the pool parallel using parallel SVM. The results are merged through cascade of SVMs. Finally images which are highly correlated with the query image both semantically and visually can be obtained. Thus the time for comparing lengthy semantic signatures is saved. The time taken by proposed system includes time for comparing as if a single semantic signature exists with an additional time of merging. Thus the result is obtained with same re-ranking accuracy as that of using multiple semantic signatures but in a faster manner. The different visual features of images used include Attention Guided Color Signature [9], Color Spatialet, SIFT [10], Multi-Layer Rotation Invariant EOH, SURF.

V. CONCLUSION

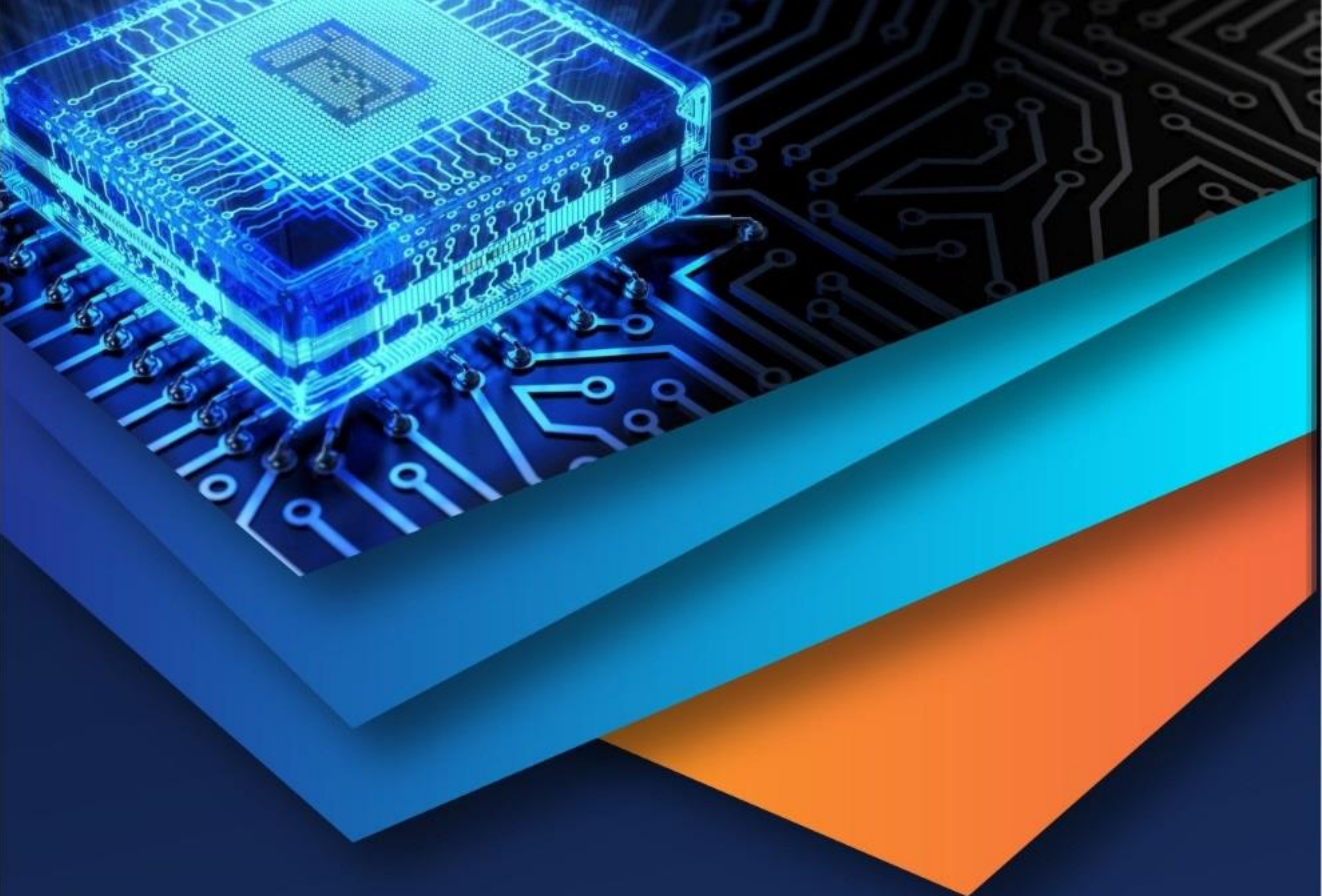
Web image re-ranking system using semantic signature and parallel SVM overcomes the major challenges of image re-ranking. It automatically offline learns different visual semantic spaces for different query keywords through keyword expansions. The visual features of images are projected into their related visual semantic spaces to get semantic signatures. At the online stage, images are re-ranked by comparing their semantic signatures obtained from the visual semantic space specified by the query keyword. An image can have a single semantic signature or multiple semantic signatures for a specific keyword. The system is having same re-ranking accuracy and precision as that of using multiple semantic signatures and takes computational time of using the single semantic signature plus the time for merging. It is faster than re-ranking using combined multiple semantic signatures. The system allows images to be re-ranked in faster manner while preserving the accuracy.

VI. ACKNOWLEDGMENT

The authors would like to thank all the staff members of Computer Science and Engineering Department of NSS College of Engineering for their guidance and support.

REFERENCES

[1] Xiao gang Wang, Shi Qiu, Ke Liu, and Xiaoou Tang, "Web Image ReRanking Using Query-Specific Semantic Signatures," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.36, no. 4,pp.810-823, April 2014.
[2] Harshil Jain, Riya Lodha , and Khushali Deulkar," Web-based Image Search using Image Re-Ranking Technique: A Review Paper," International Journal of Current Engineering and Technology , Vol.4, no.5 ,pp.3300-3303, Oct. 2014.
[3] Kirti Yadav, and Sudhir Singh," Improving Web Image Search ReRanking Using Hybrid Approach," International Journal of Advanced Research in Computer Science and Software Engineering, Vol.4, no.6, pp.629-639, June 2014.
[4] Shi Qiu, Xiaogang Wang, and Xiaoou Tang," Visual Semantic Complex Network for Web Images," Proc. IEEE Conf. Computer Vision (ICCV), 2013.
[5] Y. Rui, T.S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Trans. Circuits and Systems for Video Technology, vol. 8, no. 5, pp. 644-655, Sept. 1998.
[6] Content-based image retrieval,http://en.wikipedia. org/wiki/Content-based_image_retrieval
[7] J. Cui, F. Wen, and X. Tang, "Real Time Google and Live Image Search Re-Ranking," Proc. 16th ACM Int'l Conf. Multimedia, 2008.
[8] H.P. Graf, E. Cosatto, L. Bottou, I. Durdanovic, and V. Vapnik, "Parallel Support Vector Machines: The Cascade SVM", ;in Proc. NIPS, 2004.
[9] Y. Rubner, L. J. Guibas, and C. Tomasi,"The earth mover's distance, multi-dimensional scaling, and color-based image retrieval," Proc. ARPA Image Understanding Workshop, 1997.
[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision, vol.60, no.2 , pp.91–110, 2004.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY