

# Image Re-ranking using Information Gain and Relative Consistency through Multi-graph Learning

Poonam N. Borase  
Masters of Engineering  
MIT College, Aurangabad  
Maharashtra, India

Supriya A. Kinariwala  
Assistant Professor  
MIT College, Aurangabad  
Maharashtra, India

## ABSTRACT

After receiving a lot of attention towards text based searching for image retrieval, researchers have focused on content based image retrieval. Visual re-ranking is a method of image retrieval, which has been widely accepted to boost the accuracy of traditional text-based image retrieval. Current trend of this method is to combine the retrieval results from various visual features to boost the overall performance. The challenge in this trend of re-ranking is to exploit the complementary property of different features effectively. Our purpose basically comes under feature based image retrieval on three different modalities, so that retrieval re-ranking will be more accurate and effective. We deal with mainly two terms: information gain and relative ranking consistency among multiple modalities. Our submodular re-ranking framework can be easily used in re-ranking problems for real-time search engines.

## Keywords

Information gain, Relative Consistency, Graph Construction, Re-ranking.

## 1. INTRODUCTION

There is a huge research work focusing on the searching, retrieval and re-ranking of images[1] in the image dataset. The diverse and scattered work in this domain needs to improve in terms of accuracy. Images are similar in multiple facets, like having similar colors or textures, containing near-duplicate objects, or relative similar visual concepts. Different types of features characterize different aspects or views of images, while in many cases are also complementary to each other. Thus, to extract and fuse various kinds of visual features is a normal thought. The current trend in visual re-ranking is combining the re-ranking results from different feature channels to boost the performance in terms of precision and scalability. We present a submodular graph based technique for re-ranking images retrieved by multiple feature channels, which is fully unsupervised. Submodularity can be widely used in many optimization problems due to the availability of approximate optimization methods based on diminishing returns property - which states that as the incremental optimization algorithm proceeds, each item added to the evolving solution has less and less marginal value as the solution set grows. An information gain term and a relative ranking consistency term are two terms involved in submodularity objective function. We are using graph construction methods to achieve maximum information gain, where nodes are retrieved dataset images and edges represent similarities between two neighboring images. The graphs are converted into a transition matrix by random walk on a graph. Probability of walking from a node to its neighbors are then converted into edge weights. Pairwise relationships of retrieved images are taken into consideration while working with information gain. The inter-relationships among multiple

ranked lists obtained by different feature modalities is improved by relative ranking consistency. Both the relationships among retrieved images from a single feature channel and the relative ranks of image pairs across different feature channels are combined to achieve final submodular objective function. Recent

work combines multiple feature channels for re-ranking by multi-modal graph- based learning [2], query-specific graph fusion [3] or Co-Regularized Multi- Graph Learning [4], requires a large number of queries to compute relevance scores for initially retrieved images, which is only suitable for large sets of queries.

The contributions of our work is as follow :

- We are using greedy algorithm to maximize a submodular objective function to address the problem re-ranking natural images with different feature channels.
- We propose a Co-Regularization Multi-Graph Fusion for image re-ranking, in which multiple retrieval results from different modalities are fused based on a graph learning formulation. CR-MGF provides a novel perspective in graph based re-ranking by considering both the information gain and relative ranking consistency.

## 2. RELATED WORK

### 2.1 Visual Re-ranking

The prior work on visual re-ranking is divided into two types i.e. supervised and unsupervised, based on whether user's opinion is considered or not. The unsupervised re-ranking approach is also known as pseudo-supervised re-ranking. In supervised re-ranking requires user labeling in top ranks. Image re-ranking by example [5]: In this method a semi-supervised learning is used for image re-ranking.

### 2.2 Image Graphs

One major advantage enjoyed of graph based learning [4] is used to encode the data structure into the data similarity measurement. Visual re-ranking has introduced Graph learning. Wang *et al.* [2] developed a multi-graph learning approach to fuse multiple modalities, with the help of this approach semi-supervised learning is conducted for image re-ranking.

### 2.3 Feature Extraction

Feature extraction is a type of dimensionality reduction that efficiently represent interesting parts of an image as a feature vector. This approach is useful when image sizes are large and reduced feature representation is required to quickly complete tasks such as image matching and retrieval. The selected

features are expected to contain the relevant information from the input data, so that desired task can be performed using this reduced values instead of complete initial data[9].

### 2.4 Markov Process

A Markov process is a stochastic model [10] that has the Markov property. It can be used to model a random system that changes states according to a transition rule that only depends on the current state.

### 2.5 Information Gain

Information gain is used to determine which attribute in a given set of training feature vectors is most useful for discriminating between the classes to be learned. Information gain tells us how important a given attribute of the feature vectors is. We will use it to decide the ordering of attributes in the nodes of a decision tree.

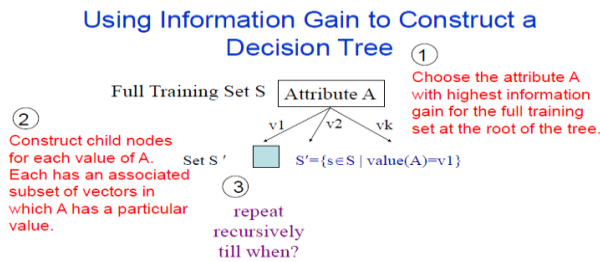


Fig. 1: Concept of Information Gain

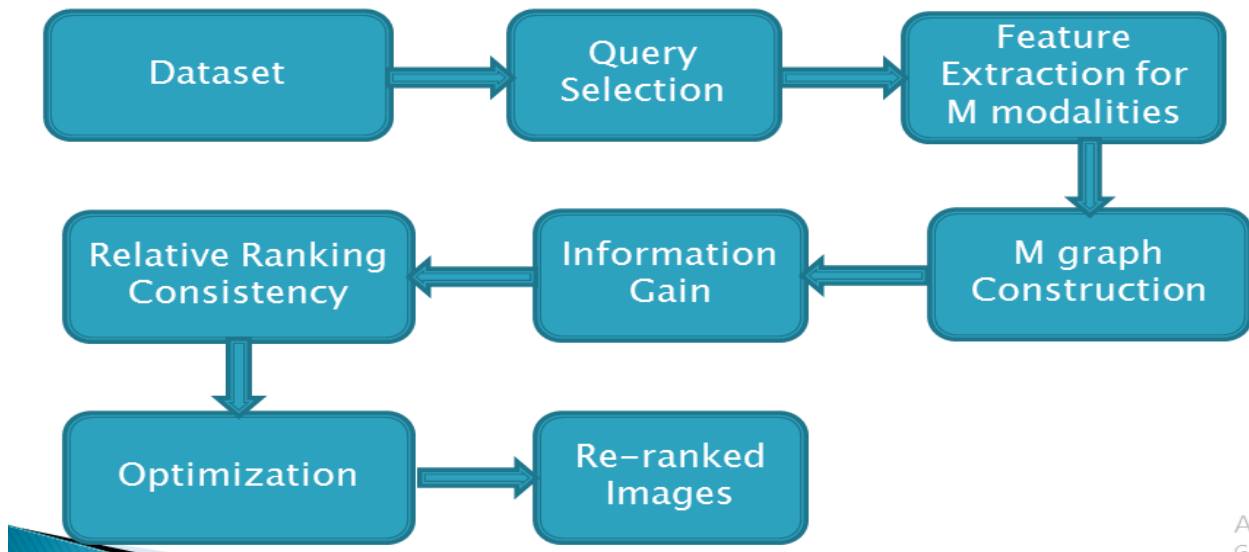


Fig 2: Architecture of Proposed Work

### 3.1 Features

We use the visual words from [6] to construct BoW vectors. 1M-dimension BoW vectors using single assignment and approximate k-means (AKM) [7]. Standard tf-idf weighting is used. For global representations, we use a 1192-dimension GIST feature and a 4000-dimension HSV color feature with 40 bins for H and 10 bins for S and V components[8].

### 3.2 Information Gain with Graph Construction

#### 3.2.1.1 Graph Construction

After selecting the three different modalities on which work has to be done, we work upon graph construction first then on information gain.

### 3. METHODOLOGY

Architecture diagram shows the relationship between different components of system. This diagram is very important to understand the overall concept of system. Architecture diagram is a diagram of a system, in which the principal parts or functions are represented by blocks connected by lines that show the relationships of the blocks. They are heavily used in the engineering world in hardware design, electronic design, software design, and process flow diagrams. For implementation of proposed system use of the dataset as INRIA Holidays, consisting of heavy images with extensions .jpg, .jpeg etc. The image dataset is used as an input to compute the information gain from the feature values of an image then the images are ranked to retrieve most relevant images to query.

Given, M feature channels, we obtain M-initial ranked list for each query image. k –retrieved images for each ranked list m. So, given initial ranked lists consisting of k-retrieval images, from modality m. We represent it as undirected graph  $G_m = (V_m, E_m)$ . Where nodes  $v_m \in V_m$  represents images from dataset.  $e_m(i, j) \in E_m$  is edge that contains  $v_m(i)$  and  $v_m(j)$ .  $A_m \in \mathbb{R}^{k \times k}$  is an affinity matrix used to represent graph with element  $a_m(i, j)$  corresponds to edge weight of  $e_m(i, j)$ , is a pairwise similarity between images  $v_m(i)$  and  $v_m(j)$ . After fusion of all graphs using information gain we will get a node set V,

$$V = V_1 \cup V_2 \cup V_3 \dots \dots \cup V_m \quad (1)$$

Now we will select a subset of nodes S from V, which are the most similar to query image, arrange them to obtain rerank

results.  $U$  is a subset of  $V$ , not selected images where  $U \cap S = \emptyset$  and  $V = U \cup S$ .

### 3.2.1.2 Information Gain

Given  $M$  graphs we find a method to combine them so that complementary modalities may help to discover images similar to query in a joint manner. For direct combination of graphs, pairwise similarity from different modalities are usually incomparable scales. To address this problem we will make a use of information gain theory which is based on a probabilistic model. A markov process is used to select a node from graph  $G_m$  randomly. A walker stays at node in a graph at time  $t$  and then randomly walks to one of its neighboring nodes at time  $t+1$  with some probability. Probability of walking between two nodes is refer as transition probability and is represented by  $P_m = D_m^{-1} A_m$ , is a row stochastic matrix indicating transition probabilities of random walk on a graph.  $P_m$  is used to convert edge weights to probabilities.

Information gain calculation for single graph,

$$F_m(S) = H(V_m \setminus S) - H(V_m \setminus S | S) \quad (2)$$

Where,  $S$ -subset selected from  $V$ ,

$V_m \setminus S$  -set with  $S$  removed,

$H(V_m \setminus S)$ -entropy of unselected nodes in graph  $G_m$ ,

$H(V_m \setminus S | S)$ - conditional entropy of remaining nodes on graph  $G_m$ .

$$H(V_m \setminus S) = -\sum_{v \in V_m \setminus S} pm(v) \log(pm(v)) \quad (3)$$

Where,  $pm(V)$ -marginal probability of  $V$  being similar to the query from modality  $m$ .

$$H(V_m \setminus S | S) = -\sum_{v \in V_m \setminus S, s \in S} pm(v, s) \log(pm(v, s)) \quad (4)$$

Where,  $pm(v, s)$ -transition probability of walking to a node  $V$  in  $G_m$  when walker is at  $S$ .

Similarity between top  $k$  images retrieved and query image from modality  $m$  as,

$$C_m = (C_{m,1}, C_{m,2}, \dots, C_{m,k}) \quad (5)$$

Finally to obtain total information gain, we simply do summing up the information gains of indivisual graphs.

$$I(S) = -\sum_m \left( \sum_{v \in V_m \setminus S} pm(v) \log(pm(v)) - \sum_{v \in V_m \setminus S, s \in S} pm(v, s) \log(pm(v, s)) \right) \quad (6)$$

Maximizing  $I(S)$  is equivalent to selecting group of images that are similar to query and closely related to each other.

## 3.3 Relative Ranking Consistency

Information gain does not utilize the inter relationships between modalities, it only considers pairwise similarities between images. To improve the re-ranking performance we use the initial ranks of retrieved images from different feature channels as an extra information. For example an image is similar to query and ranked lower by one of the modality may be having higher rank when seen from different view. So, only information gain is not sufficient to retrieve images with high performance.

We first define the relative ranking between a pair of images and then use it to measure the inter-rank among multiple ranked lists.

Formulation for Relative Ranking Consistency:

$r_m$ - position of top  $k$  images in initial ranked list by feature channel  $m$ .

$$r_m = (r_{m,1}, r_{m,2}, \dots, r_{m,k}) \quad (7)$$

where,  $r_{m,i}$  – position of image  $I_i$  in  $m^{\text{th}}$  ranked list.

Relative ranking between two images is calculated by

$$rr_m(v_i, v_j) = |r_{m,v_i} - r_{m,v_j}|, v_i, v_j \in V \quad (8)$$

where,  $v_i, v_j$  are images  $I_i$  and  $I_j$  in graph representation.

Relative ranking consistency measure across multiple ranked lists is given by,

$$C(v_i, v_j) = \frac{1}{Z} - \sum_{m, m' \in M, m \neq m'} \frac{\min(rr_m - rr_{m'})}{K} \quad (9)$$

Where,  $Z = M(M - 1)/2$ , number of all possible modality pairs.

Finally, Relative ranking consistency term is defines as,

$$R(S) = (1 - q) \sum_s q^s \cdot \frac{1}{S} \sum_{v_i, v_j \in S; r_{v_i} < r_{v_j}} C(v_i, v_j) \quad (10)$$

Where,  $\frac{1}{S} \sum_{v_i, v_j \in S; r_{v_i} < r_{v_j}} C(v_i, v_j)$ - allows to select image  $v_j$  with new rank  $s$  and compute the average relative ranking measure between  $v_j$  and all other  $S$  images with rank higher than  $v_j$ .

## 3.4 Optimization

Combining the information gain and relative ranking consistency terms, we obtain final objective function,

$$ObjFun(S) = I(S) + \alpha R(S) \quad (11)$$

Where  $\alpha$  is pre-defined weighting factor balancing two terms.

Direct optimization of objective function is hard, but it can be approximately optimized by a greedy algorithm

- Starting with an empty set= $\emptyset$
- Greedy algorithm iteratively adds a new element to  $S$  which provides the largest marginal gain at each iteration , until  $k_s$  elements selected.
- During each iteration, search for an image ,
  - a $\in V/S$  which gives largest combined marginal gain from information gain and relative ranking consistency terms.
- Add it to set  $S$  and obtain ranks.

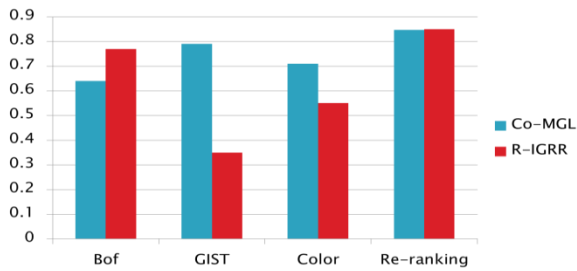


Fig. 3: Analysis of two systems at different level of implementation

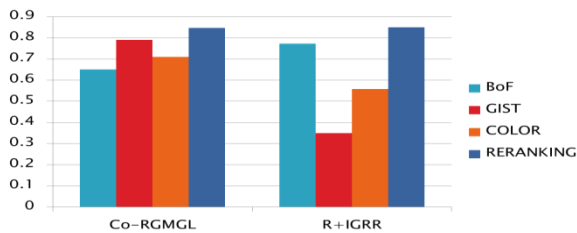


Fig. 4: Analysis of two systems at different level of implementation

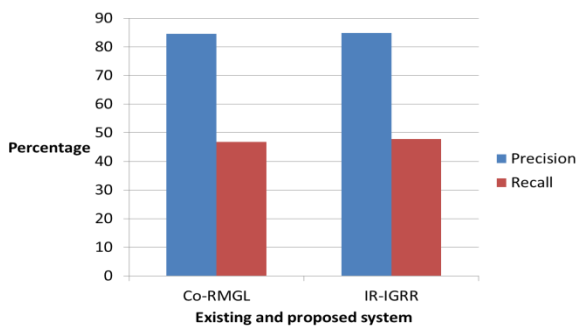


Fig. 5: Precision and Recall Percentage For Existing and Proposed System

## 4. RESULTS

Our primary focus is on improving re-ranking algorithm for multiple modalities that improves retrieval performance of multiple ranked lists modalities. Although our implementation depends only on pairwise similarities without spatial verification and query expansion, the performance by our submodular re-ranking is comparable to other state-of-the-art approaches using a single modality. mAP is improved by 0.01 % from previous method. Our algorithm outperform robustly for the INRIA Holidays dataset.

## 5. CONCLUSION

We proposed a method for image re-ranking by introducing a graph learning technology, where we worked on graphs with the help of two terms information gain and relative ranking consistency measure. Here, we have worked on three different feature channels to achieve better performance. So, by maximizing the information gain and final objective function we achieve efficient re-ranking results. Greedy algorithm is used to maximize the objective function. Experiments show that our re-ranking algorithm improves overall retrieval accuracy.

## 6. REFERENCES

- [1] V.Vinitha, Dr. J.Jagadeesan, R. Augustian Isaac, "Web Image Search Reranking Using CBIR," International Journal of Computer Science & Engineering Technology (IJCSSET), ISSN: 2229-3345 Vol. 5 No. 04 Apr 2014.
- [2] Wang, M., Li, H., Tao, D., Lu, K., Wu, X.: Multimodal graph-based re-ranking for web image search. IEEE Transactions on Image Processing 21 (2012) 4649-4661
- [3] Zhang, S., Yang, M., Cour, T., Yu, K., Metaxas, D.N.: Query speci\_c fusion for image retrieval. In: ECCV. (2012) 660-673
- [4] Deng, C., Ji, R., Liu, W., Tao, D., Gao, X.: Visual re-ranking through weakly supervised multi-graph learning. In: ICCV. (2013) 2600-2607
- [5] Zhang Wei, Xue Tianfan, "Image Reranking by Example:A Semi-supervised Learning Formulation," The Chinese University of Hong Kong.
- [6] Qin, D., Gammeter, S., Bossard, L., Quack, T., Gool, L.J.V.: Hello neighbor: Accurate object retrieval with k-reciprocal nearest neighbors. In: CVPR. (2011) 777-784
- [7] Philbin, J., Chum, O., Isard, M., Sivic, J., Zisserman, A.: Object retrieval with large vocabularies and fast spatial matching. In: CVPR. (2007) 1-8
- [8] <https://www.quora.com/How-does-features-extraction-on-images-work>
- [9] R. Gonzales, R. E. Woods, "Digital Image Processing," 2nd Ed.
- [10] Jagdish S. Rustagi , "Optimization techniques in statistics"