

# Joint Hypergraph Learning for Tag-based Image Retrieval

## Peddakottam Anusha

Department of Computer Science, Sri Venkateswara University, Tirupati

**Abstract**— Compared with content-based image retrieval (TBIR), tag-based image search is more commonly used in social media. It is one of the important ways to find images contributed by social users. In this research field, tag information and diverse visual features have been investigated. A pseudo-relevance feedback mechanism to obtain the pseudo-positive images. Finally, with the hypergraph and pseudo relevance feedback. Hypergraph not only takes pairwise relationship into consideration, but also models the higher order relationship among three or more vertices containing grouping information. Hypergraph method is widely used in data mining and information retrieval tasks [6, 7]. Cai et al. [9] first train attribute ion classifiers, then construct hypergraph based on these classifiers, finally they obtain the relevance score by hypergraph learning. Finally, with the hypergraph and pseudo relevance feedback, we adopt the hypergraph learning algorithm to calculate the relevance score of each image to the query. Experimental results demonstrate the effectiveness of the proposed approach.

**Keywords:** Tag-based Image Retrieval. Hypergraph. Feature fusion. Visual Feature. Pseudo Relevance Feedback.

## I. INTRODUCTION

The large-scale web images demand the researchers to develop efficient algorithms for more accurate indexing and retrieval. Compared with content-based image retrieval (TBIR), tag-based image search is more commonly used in social media [32, 51]. In the last few decades, extensive efforts have been dedicated to image relevance retrieval. However, many algorithms can't achieve satisfactory results for tag mismatch, noisy tags and query ambiguity problems [51]. Thus, more and more researchers attempt to utilize visual features and user relevance feedback to improve the retrieval accuracy. There are several visual features designed to express images such as colour feature [29], shape feature [36], textural feature [37], edge feature [1], SIFT [16] and deep feature [58, 62]. Different visual features describe different aspects of an image. Therefore, some algorithms try to fuse multiple visual features to improve the image retrieval precision [2, 4, 5, 33]. However, most existing methods usually explore multiple. visual features separately. For example, Yang et al. [2] first construct a graph for every feature. Then they apply random walk model to get a relevance score according to each constructed graph. Finally, they re-rank the images by the linear combination of the relevance scores of different features. Zhang et al. [4] first select training samples, then they apply multiple visual features by simple to train the classification function for image ranking. In [5], Yang et al. learn the Mahala Nobis matrix for different visual features and calculate the distance of images by the Mahala Nobis distance of corresponding visual feature. Yu et al. [33] construct five hypergraphs for five visual features, and integrate the visual consistency constrains of these hypergraphs to learn a linear model for ranking. Gao et al. [34] construct hypergraph by local visual feature only and abandon the global visual information of images. However, different visual features have district emphasis on describing the content of an image, therefore, separately or sequentially using this information is suboptimal for social image retrieval.

### 1.1 What is an image?

An image is defined as a two-dimensional function  $(x,y)$ , where  $x$  and  $y$  are spatial coordinates, and the amplitude of  $F$  at any pair of coordinates  $(x,y)$  is called the intensity of that image at that point. When  $x,y$ , and amplitude values of  $F$  are finite, we call it a digital image.

In other words, an image can be defined by a two-dimensional array specifically arranged in rows and columns.

Digital Image is composed of a finite number of elements, each of which elements have a particular value at a particular location. These elements are referred to as picture elements, image elements, and pixels. A Pixel is most widely used to denote the elements of a Digital Image.

### 1.2 Types of an image

**BINARY IMAGE**– The binary image as its name suggests, contain only two-pixel elements i.e 0 & 1, where 0 refers to black and 1 refers to white. This image is also known as Monochrome.

**BLACK AND WHITE IMAGE**– The image which consist of only black and white color is called **BLACK AND WHITE IMAGE**.

**8 bit COLOR FORMAT**– It is the most famous image format. It has 256 different shades of colors in it and commonly known as Grayscale Image. In this format, 0 stands for Black, and 255 stands for white, and 127 stands for gray.

**16 bit COLOR FORMAT**– It is a color image format. It has 65,536 different colors in it. It is also known as High Color Format. In this format the distribution of color is not as same as Grayscale image.

A 16 bit format is actually divided into three further formats which are Red, Green and Blue. That famous RGB format.

### 1.3 Phases of Image Processing:

1.ACQUISITION– It could be as simple as being given an image which is in digital form. The main work involves:

a) Scaling

b) Color conversion (RGB to Gray or vice-versa)

2.IMAGE ENHANCEMENT– It is amongst the simplest and most appealing in areas of Image Processing it is also used to extract some hidden details from an image and is subjective.

3.IMAGE RESTORATION– It also deals with appealing of an image but it is objective (Restoration is based on mathematical or probabilistic model or image degradation).

4.COLOR IMAGE PROCESSING– It deals with pseudo color and full color image processing color models are applicable to digital image processing.

5.WAVELETS AND MULTI-RESOLUTION PROCESSING– It is foundation of representing images in various degrees.

6.IMAGE COMPRESSION-It involves in developing some functions to perform this operation. It mainly deals with image size or resolution.

7.MORPHOLOGICAL PROCESSING-It deals with tools for extracting image components that are useful in the representation & description of shape.

8.SEGMENTATION PROCEDURE-It includes partitioning an image into its constituent parts or objects. Autonomous segmentation is the most difficult task in Image Processing.

9.REPRESENTATION & DESCRIPTION-It follows output of segmentation stage, choosing a representation is only the part of solution for transforming raw data into processed data.

0.OBJECT DETECTION AND RECOGNITION-It is a process that assigns a label to an object based on its descriptor.

### 1.4 Image compression

An important development in digital image compression technology was the discrete cosine transform (DCT), a lossy compression technique first proposed by Nasir Ahmed in 1972.<sup>[14]</sup> DCT compression became the basis for JPEG, which was introduced by the Joint Photographic Experts Group in 1992.<sup>[15]</sup> JPEG compresses images down to much smaller file sizes, and has become the most widely used image file format on the Internet.<sup>[16]</sup> Its highly efficient DCT compression algorithm was largely responsible for the wide proliferation of digital images and digital photos,<sup>[17]</sup> with several billion JPEG images produced every day as of 2015.<sup>[18]</sup>

### 1.5 Digital signal processor (DSP)

Electronic signal processing was revolutionized by the wide adoption of MOS technology in the 1970s.<sup>[19]</sup> MOS integrated circuit technology was the basis for the first single-chip microprocessors and microcontrollers in the early 1970s,<sup>[20]</sup> and then the first single-chip digital signal processor (DSP) chips in the late 1970s.<sup>[21][22]</sup> DSP chips have since been widely used in digital image processing.<sup>[21]</sup>

The discrete cosine transform (DCT) image compression algorithm has been widely implemented in DSP chips, with many companies developing DSP chips based on DCT technology. DCTs are widely used for encoding, decoding, video coding, audio coding, multiplexing, control signals, signaling, analog-to-digital conversion, formatting luminance and color differences, and color formats such as YUV444 and YUV411. DCTs are also used for encoding operations such

as motion estimation, motion compensation, inter-frame prediction, quantization, perceptual weighting, entropy encoding, variable encoding, and motion vectors, and decoding operations such as the inverse operation between different color formats (YIQ, YUV and RGB) for display purposes. DCTs are also commonly used for high-definition television (HDTV) encoder/decoder chips.<sup>[23]</sup>

### 1.6 Medical imaging

In 1972, the engineer from British company EMI Housfield invented the X-ray computed tomography device for head diagnosis, which is what is usually called CT (computer tomography). The CT nucleus method is based on the projection of the human head section and is processed by computer to reconstruct the cross-sectional image, which is called image reconstruction. In 1975, EMI successfully developed a CT device for the whole body, which obtained a clear tomographic image of various parts of the human body. In 1979, this diagnostic technique won the Nobel Prize.<sup>[4]</sup> Digital image processing technology for medical applications was inducted into the Space Foundation Space Technology Hall of Fame in 1994.<sup>[24]</sup>

Digital image processing allows the use of much more complex algorithms, and hence, can offer both more sophisticated performance at simple tasks, and the implementation of methods which would be impossible by analogue means.

In particular, digital image processing is a concrete application of, and a practical technology based on:

- Classification
- Feature extraction
- Multi-scale signal analysis
- Pattern recognition
- Projection

Some techniques which are used in digital image processing include:

- Anisotropic diffusion
- Hidden Markov models
- Image editing
- Image restoration
- Independent component analysis
- Linear filtering
- Neural networks
- Partial differential equations
- Pixelation
- Point feature matching
- Principal components analysis
- Self-organizing maps
- Wavelets

## II. LITERATURE SURVEY

**Title:** An image registration method based on similarity of edge information

**Author:** Cai Xi; Han Guang; Xiao Shunli

**Abstract:** Considering that spatial structural information of images, taken from the same scene even with diverse illustration or in different wavebands, has much in common with each other, we propose an image registration method utilizing similarity

of edge information as a new similarity measure. Similarity of edge information can be achieved from gradient information of detected edge pixels in two images, so it reflects well the relevance of spatial structures between two corresponding scenes. Experimental results indicate that, no matter when lighting changes or for images sensed in different wavebands, the proposed similarity measure exhibits a perfect single peak response with respect to transformation parameters and the proposed method achieves more accurate alignment than the one based on mutual information.

**Title:** Click-boosting multi-modality graph-based re-ranking for image search

**Author:** Yongdong Zhang

**Abstract:** Image reranking is an effective way for improving the retrieval performance of keyword-based image search engines. A fundamental issue underlying the success of existing image reranking approaches is the ability in identifying potentially useful recurrent patterns from the initial search results. Ideally, these patterns can be leveraged to upgrade the ranks of visually similar images, which are also likely to be relevant. The challenge, nevertheless, originates from the fact that keyword-based queries are used to be ambiguous, resulting in difficulty in predicting the search intention. Mining useful patterns without understanding query is risky, and may lead to incorrect judgment in reranking. This paper explores the use of click-through data, which can be viewed as the footprints of user searching behavior, as an effective means of understanding query, for providing the basis on identifying the recurrent patterns that are potentially helpful for reranking. A new reranking algorithm, named click-boosting multi-modality graph-based reranking, is proposed. The algorithm leverages clicked images to locate similar images that are not clicked, and reranks them in a multi-modality graph-based learning scheme. Encouraging results are reported for image reranking on a real-world image dataset collected from a commercial search engine with click-through data.

**Title:** Learning with Hypergraphs: Clustering, Classification, and Embedding

**Author:** John Platt; Thomas Hofmann

**Abstract:** We usually endow the investigated objects with pairwise relationships, which can be illustrated as graphs. In many real-world problems, however, relationships among the objects of our interest are more complex than pairwise. Naively squeezing the complex relationships into pairwise ones will inevitably lead to loss of information which can be expected valuable for our learning tasks however. Therefore, we consider using hypergraphs instead to completely represent complex relationships among the objects of our interest, and thus the problem of learning with hypergraphs arises. Our main contribution in this paper is to generalize the powerful methodology of spectral clustering which originally operates on undirected graphs to hypergraphs, and further develop algorithms for hypergraph embedding and transductive classification on the basis of the spectral hypergraph clustering approach. Our experiments on a number of benchmarks showed the advantages of hypergraphs over usual graphs.

**Title:** Image Search Reranking With Query-Dependent Click-Based Relevance Feedback

**Author:** Xiaopeng Yang

**Abstract:** Our goal is to boost text-based image search results via image reranking. There are diverse modalities (features) of images that we can leverage for reranking, however, the effects of different modalities are query-dependent. The primary challenge we face is how to fuse multiple modalities adaptively for different queries, which has often been overlooked in previous reranking research. Moreover, multimodality fusion without an understanding of the query is risky, and may lead to incorrect judgment in reranking. Therefore, to obtain the best fusion weights for the query, in this paper, we leverage click-through data, which can be viewed as an “implicit” user feedback and an effective means of understanding the query. A novel reranking algorithm, called click-based relevance feedback, is proposed. This algorithm emphasizes the successful use of click-through data for identifying user search intention, while leveraging multiple kernel learning algorithm to adaptively learn the query-dependent fusion weights for multiple modalities. We conduct experiments on a real-world data set collected from a commercial search engine with click-through data. Encouraging experimental results demonstrate that our proposed reranking approach can significantly improve the NDCG@10 of the initial search results by 11.62%, and can outperform several existing approaches for most kinds of queries, such as tail, middle, and top queries.

**Title:** Web Image Search Re-Ranking With Click-Based Similarity and Typicality

**Author:** Jie Liu; Shin'ichi Satoh

### III. EXISTING SYSTEM:

A new pseudo relevance feedback mechanism for tag-based image retrieval. First, we conduct clustering on the co-occurrence tags. The introducing of cluster relevance is an assistant for calculating image initial relevance score, which is superior to using image relevance only. In online retrieval, we only match the query tag to get the retrieval results, thus the online search is very efficient. Hypergraph not only takes pairwise relationship into consideration, but also models the higher order relationship among three or more vertices containing grouping information. Hypergraph method is widely used in data mining and information retrieval tasks [6, 7]. Cai et al. [9] first train attribute ion classifiers, then construct hypergraph based on these classifiers, finally they obtain the relevance score by hypergraph learning. Hypergraph has shown its effectiveness of higher-order sample relationship modelling in many mechanism learning tasks such as data mining and information retrieval. For image retrieval, many algorithms are designed based on hypergraph technique. Liu et al. [7] propose a soft hypergraph, which assigns each vertex to a hyperedge in a soft way. The image retrieval task is formulated as the problem of hypergraph ranking. Zhu et al. learn hash codes for mobile image retrieval based on hypergraph [68], they utilize hypergraph to model the high-order semantics of images.

#### 3.1 Disadvantage:

The online retrieval can be conducted by key words match. However, the retrieval results are not satisfactory for the unreliable tags

### IV. PROPOSED SYSTEM

- A new pseudo relevance feedback mechanism for tag-based image retrieval. First, we conduct clustering on the co-occurrence tags.
- The introducing of cluster relevance is an assistant for calculating image initial relevance score, which is superior to using image relevance only.
- In online retrieval, we only match the query tag to get the retrieval results, thus the online search is very efficient.
- We present a novel joint learning approach for tag-based web image retrieval (JHR), which utilizes the global, local visual features and textual feature simultaneously. Compared to using global, local visual feature or textual feature alone or separately, the joint hypergraph learning approach can capture more reliable relationships between images.
- We propose a new pseudo relevance feedback mechanism for tag-based image retrieval. First, we conduct clustering on the co-occurrence tags. Then we assign images to clusters. Finally, we estimate the relevance between image and query by fusing the image cluster relevance and image relevance to query.
- The introducing of cluster relevance is an assistant for calculating image initial relevance score, which is superior to using image relevance only. 3) We build the inverted file system for tags in offline part. All steps of our algorithm are conducted offline. In online retrieval, we only match the query tag to get the retrieval results, thus the online search is very efficient.

#### 3.1 Advantages:

- High security and more effective.
- This system gives more accuracy compared to another systems.
- No illegal key share doesn't happen, because it should be traced.
- It supports public traceability, key abuse, and key escrow.

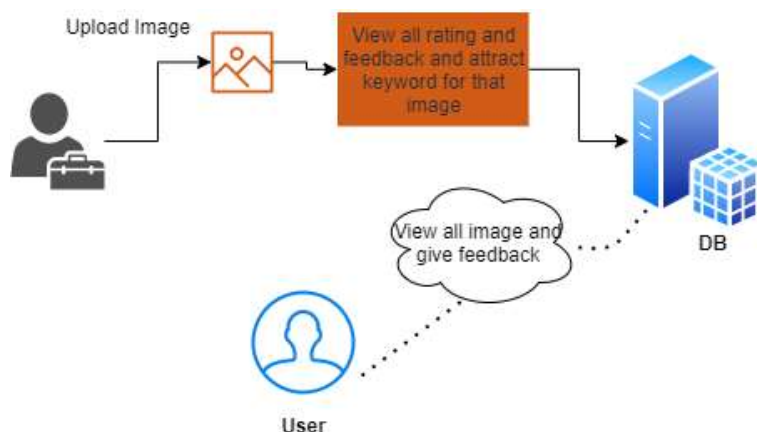
#### 3.2 Algorithm:

**Ciphertext:** Ciphertext is also known as encrypted or encoded information because it contains a form of the original plaintext that is unreadable by a human or computer without the proper cipher to decrypt it. Decryption, the inverse of encryption, is the process of turning ciphertext into readable plaintext.

**AES:** The algorithm described by AES is a symmetric-key algorithm, meaning the same key is used for both encrypting and decrypting the data. The Advanced Encryption Standard (AES) is a symmetric block cipher chosen by the U.S. government to

protect classified information. AES is implemented in software and hardware throughout the world to encrypt sensitive data. It is essential for government computer security, cybersecurity and electronic data protection. Symmetric, also known as secret key, ciphers use the same key for encrypting and decrypting, so the sender and the receiver must both know -- and use -- the same secret key. The government classifies information in three categories: Confidential, Secret or Top Secret. All key lengths can be used to protect the Confidential and Secret level. Top Secret information requires either 192- or 256-bit key lengths.

## V. SYSTEM ARCHITECTURE



### Applications:

In this method can be implemented in all other websites, for example who wants more security for Image. We use this is college, government, company's etc.

## VI. FUTURE ENHANCEMENT

Joint re-ranking method for social image retrieval, in which we simultaneously utilize global, local visual features and textual feature to improve the retrieval accuracy the global and local visual features is much better than using any of them alone and also more efficient than the comparison methods. The discussions in experiment show that our method has lighter dependence on the learning parameters, clustering methods and the metric

## VII. CONCLUSION

A new joint re-ranking method for social image retrieval, in which we simultaneously utilize global, local visual features and textual feature to improve the retrieval accuracy. Experiment results on NUS-Wide dataset show that combing the global and local visual features is much better than using any of them alone and also more efficient than the comparison methods. The discussions in experiment show that our method has lighter dependence on the learning parameters, clustering methods and the metric methods we apply. However, in our method, we only consider the relevance of result and ignore the diversity. In our future work, we will investigate the diversity by multiple visual features.

## REFERENCES

- [1] X. Cai, G. Han, and S. Xiao, "An image registration method based on similarity of edge information", IEEE International Symposium on Industrial Electronics. IEEE Press, 2012, pp. 217-224.
- [2] X. Yang, Y. Zhang., T. Yao, C. Ngo, and T. Mei, "Click-boosting multi-modality graph-based re-ranking for image search", Multimedia Systems, vol.21, issue 2, pp.217-227, 2015.
- [3] D. Zhou, J. Huang, and B. Scholkopf, "Learning with hypergraphs: Clustering, classification, and embedding", NIPS, vol. 19, 2006.
- [4] Y. Zhang, X. Yang, and T. Mei, "Image Search Re-ranking With Query-Dependent Click-Based Relevance Feedback". IEEE Transaction on Image Processing, vol.23, no.10, pp.2310-4448, 2014.
- [5] X. Yang, T. Mei, Y. Zhang, and J. Liu, "Web Image Search Re-Ranking with Click-Based Similarity and Typicality", IEEE Transaction on Image Processing, vol.25, no.10, pp.4617-4630, 2016.
- [6] Y. Huang, Q. Liu, S. Zhang, and D. Metaxas, "Image retrieval via probabilistic hypergraph ranking", CVPR. IEEE, pp. 3376– 3383, 2010