ISSN:2231-0711

K.Lahari et al | IJCSET(www.ijcset.net) | November 2016 | Vol 6, Issue 11, 375-379

Reinforced OMDML with Ambiguity Resolving Problem in Ranking based CBIR System

K.Lahari
Student of M.Tech(CSE)
Department of Computer Science Engineering
Avanthi Institute of Engineering & Technology

Ch.Kodandaramu **M.sc , M.Tech , (Ph.D)**
Assistant Professor,
Department of Computer Science and Engineering
Avanthi Institute of Engineering & Technology

Abstract-The paper proposes a new initiate to enhance the existing online multi-modal distance metric learning (OMDML) with a new feature of extension to solve the Image ambiguity issue using Conditional Random Field (CRF) Algorithm. The main intention of proposing this model of system is to annotate/tag the images with some manually defined concepts for learning an inherent space, using visual and contextual features. More particularly, by making the system to feed the latent vectors into existing classification representations, it can be enforce for application of image annotation, which is considered as the needed issue in image retrieval. As an extension to the available model, we recommend and add the entity feature of the problem of solving the ambiguity. The Conditional Random Filed Algorithm model is used for training the system and results of reinforced online multi-modal distance metric learning system provides an better outcome of content based image retrieval model. This solution is the future enhancement where the contribution of providing more accuracy to the proposed system by enhancing using ambiguity resolving problem.

I. INTRODUCTION

What Is Image Processing?

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Usually Image Processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

It is among rapidly growing technologies today, with its applications in various aspects of a business. Image Processing forms core research area within engineering and computer science disciplines too.

Image processing basically includes the following three steps:

- ❖ Importing the image with optical scanner or by digital photography.
- ❖ Analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs.
- ❖ Output is the last stage in which result can be altered image or report that is based on image analysis.

Purpose of Image processing:

The purpose of image processing is divided into 5 groups.

They are:

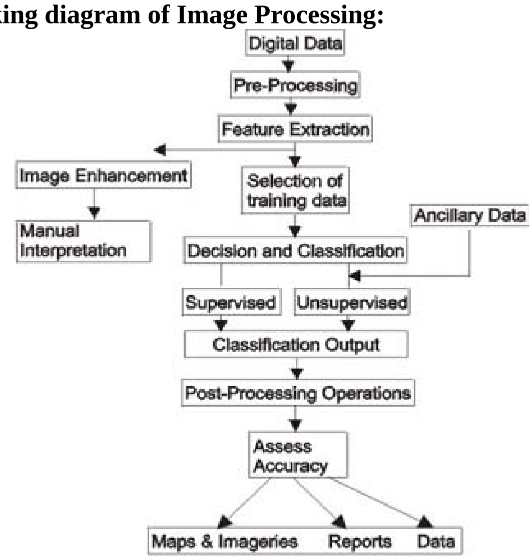
1. Visualization - Observe the objects that are not visible.
2. Image sharpening and restoration - To create a better image
3. Image retrieval - Seek for the image of interest.
4. Measurement of pattern – Measures various objects in an image.
5. Image Recognition – Distinguish the objects in an image.

Types of Image Processing:

The two types of methods used for Image Processing are Analog and Digital Image Processing. Analog or visual techniques of image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. The image processing is not just confined to area that has to be studied but on knowledge of analyst. Association is another important tool in image processing through visual techniques. So analysts apply a combination of personal knowledge and collateral data to image processing.

Digital Processing techniques help in manipulation of the digital images by using computers. As raw data from imaging sensors from satellite platform contains deficiencies. To get over such flaws and to get originality of information, it has to undergo various phases of processing. The three general phases that all types of data have to undergo while using digital technique are Pre-processing, enhancement and display, information extraction.

Working diagram of Image Processing:



Characteristics of Image Processing:
Before going to processing an image, it is converted into a digital form. Digitization includes sampling of image and quantization of sampled values. After converting the image into bit information, processing is performed. This processing technique may be, Image enhancement, Image restoration, and Image compression.

Image enhancement:
It refers to accentuation, or sharpening, of image features such as boundaries, or contrast to make a graphic display more useful for display & analysis. This process does not increase the inherent information content in data. It includes gray level & contrast manipulation, noise reduction, edge crispening and sharpening, filtering, interpolation and magnification, pseudo coloring, and so on.

Image restoration:
It is concerned with filtering the observed image to minimize the effect of degradations. Effectiveness of image restoration depends on the extent and accuracy of the knowledge of degradation process as well as on filter design. Image restoration differs from image enhancement in that the latter is concerned with more extraction or accentuation of image features.

Image compression:
It is concerned with minimizing the number of bits required to represent an image. Application of compression are in broadcast TV, remote sensing via satellite, military communication via aircraft, radar, teleconferencing, facsimile transmission, for educational & business documents, medical images that arise in computer tomography, magnetic resonance imaging and digital radiology, motion, pictures, satellite images, weather maps, geological surveys and so on.

- Text compression – CCITT GROUP3 & GROUP4
- Still image compression – JPEG
- Video image compression – MPEG

Advantages of Image Processing:

- The processing of images is faster and more cost-effective. One needs less time for processing, as well as less film and other photographing equipment.
- It is more ecological to process images. No processing or fixing chemicals are needed to take and process digital images. However, printing inks are essential when printing digital images.
- When shooting a digital image, one can immediately see if the image is good or not.
- Copying a digital image is easy, and the quality of the image stays good unless it is compressed. For instance, saving an image as jpg format compresses the image. By resaving the image as jpg format, the compressed image will be recompressed, and the quality of the image will get worse with every saving.
- Fixing and retouching of images has become easier. In new Photoshop 7, it is possible to smoother face wrinkles with a new Healing Brush Tool in a couple of seconds.
- The expensive reproduction (compared with rastering the image with a repro camera) is faster and cheaper.
- By changing the image format and resolution, the image can be used in a number of media.

II. Related Work

Our work is related to three major groups of research: content-based image retrieval, distance metric learning, and online learning. In the following, we briefly review the closely related representative works in each group.

A. Content-Based Image Retrieval

With the rapid growth of digital cameras and photo sharing websites, image retrieval has become one of the most important research topics in the past decades, among which content-based image retrieval is one of key challenging problems [1], [2], [3]. The objective of CBIR is to search images by analyzing the actual contents of the image as opposed to analyzing metadata like keywords, title and author, such that extensive efforts have been done for investigating various low-level feature descriptors for image representation [14]. For example, researchers have spent many years in studying various global features for image representation, such as color features [14], edge features [14], and texture features [15]. Recent years also witness the surge of research on local feature based representation, such as the bag-of-words models [16], [17] using local feature descriptors (e.g., SIFT [18]). Conventional CBIR approaches usually choose rigid distance functions on some extracted low-level features for multimedia similarity search, such as the classical Euclideann distance or cosine similarity. However, there exists one key limitation that the fixed rigid similarity/distance function may not be always optimal because of the complexity of visual image representation and the main challenge of the semantic gap between the low-level visual features extracted by computers and high-level human perception and interpretation. Hence, recent years have witnesses a surge of active research efforts in design of various distance/similarity measures on some low-level features by exploiting machine learning techniques [19], [20], [21], among which some works focus on learning to hash for compact codes [19], [22], [23], [24], [25], and some others can be categorized into distance metric learning that will be introduced in the next section. Our work is also related to multimodal/multiview studies, which have been widely studied on image classification and object recognition fields [26], [27], [28], [29]. However, it is usually hard to exploit these techniques directly on CBIR because (i) in general, image classes will not be given explicitly on CBIR tasks, (ii) even if classes are given, the number will be very large, (iii) image datasets tend to be much larger on CBIR than on classification tasks. We thus exclude the direct comparisons to such existing works in this paper. There are still some other open issues in CBIR studies, such as the efficiency and scalability of the retrieval process that often requires an effective indexing scheme, which are out of this paper's scope.

B. Distance Metric Learning

Distance metric learning has been extensively studied in both machine learning and multimedia retrieval communities [7], [30], [31], [32], [33], [34], [35], [36]. The essential idea is to learn an optimal metric which minimizes the distance between similar/related images and simultaneously maximizes the distance between dissimilar/unrelated images. Existing DML studies can be grouped into different categories according to different learning settings and principles. For example, in terms of different types of constraint settings, DML techniques are typically categorized into two groups: _ Global supervised approaches [7], [30]: to learn a metric on a global setting, e.g., all constraints will be satisfied simultaneously; _ Local supervised approaches [32], [33]: to learn a metric in the local sense, e.g., the given local constraints from neighboring information will be satisfied. Moreover, according to different training data forms, DML studies in machine learning typically learn metrics directly from explicit class labels [32], while DML studies in multimedia mainly learn metrics from side information, which usually can be obtained in the following two forms: _ Pairwise constraints [7], [9]: A must-link constraint set S and a cannot-link constraint set D are given, where a pair of images ðpi; pjÞ 2 S if pi is related/ similar to pj, otherwise ðpi; pjÞ 2 D. Some literature uses the term equivalent/positive constraint in place of “must link”, and the term inequivalent/negative constraint in place of “cannot-link”. _ Triple constraints [20]: A triplet set P is given, where P ¼ fðpt; pþt ; p_t Þjðpt; pþt
Þ 2 S; ðpt; p_t Þ 2 D; t ¼ 1; . . . ; Tg, S contains related pairs and D contains unrelated pairs, i.e., p is related/similar to pþ and p is unrelated/ dissimilar to p_. T denotes the cardinality of entire triplet set. When only explicit class labels are provided, one can also construct side information by simply considering relationships of instances in same class as related, and relationships of instances belonging to different classes as unrelated. In our works, we focus on triple constraints. Finally, in terms of learning methodology, most existing DML studies generally employ batch learning methods which often assume the whole collection of training data must be given before the learning task and train a model from scratch, except for a few recent DML studies which begin to explore online learning techniques [37], [38]. All these works generally address single-modal DML, which is different from our focus on multi-modal DML.

C. Online Learning

Our work generally falls in the category of online learning methodology, which has been extensively studied in machine learning [41], [42]. Unlike batch learning methods that usually suffer from expensive re-training cost when new training data arrive, online learning sequentially makes a highly efficient (typically constant) update for each new training data, making it highly scalable for large-scale applications. In general, online learning operates on a sequence of data instances with time stamps. At each time step, an online learning algorithm processes an incoming example by first predicting its class label; after the prediction, it receives the true class label which is then used to measure the suffered loss between the predicted label and the true label; at the end of each time step, the model is updated with the loss whenever it is nonzero. The overall objective of an online learning task is to minimize the cumulative loss over the entire sequence of received instances. In literature, a variety of algorithms have been proposed for online learning [43], [44], [45], [46], [47]. Some wellknown examples include the Hedge algorithm for online prediction with expert advice [48], the Perceptron algorithm [43], the family of passive-Aggressive (PA) learning algorithms [44], and the online gradient descent (OGD) algorithms [49]. There is also some study that attempts to improve the scalability of online kernel methods, such as [50] which proposed a bounded online gradient descent for addressing online kernel-based classification tasks. In this work, we apply online learning techniques, i.e., the Hedge, PA, and online gradient descent algorithms, to tackle the multi-modal distance metric learning task for contentbased image retrieval. Besides, we note that this work was partially inspired by the recent study of online multiple kernel learning which aims to address online classification tasks using multiple kernels [51].

III. MODULES DESCRIPTION

A. Content-Based Image Retrieval

Content-based image retrieval is one of key challenging problems The objective of CBIR is to search images by analyzing the actual contents of the image as opposed to analyzing metadata like keywords, title and author, such that extensive efforts have been done for investigating various low-level feature descriptors for image representation. Conventional CBIR approaches usually choose rigid distance functions on some extracted low-level features for multimedia similarity search, such as the classical Euclidean distance or cosine similarity.

B. Utilization of Text-Based Search

Since in reranking, the text-based search provides original ranking lists instead of quantized scores, a necessary step is to turn the ranking positions into scores. In this work, we investigate the relationship between y_i and the position i with a large number of queries. We elaborate the proposed attribute-assisted image search re ranking framework. An attribute-assisted hypergraph learning method to reorder the ranked images which returned from search engine based on textual query. The attribute-assisted re ranking method, we compare the following approaches for performance evaluation.

C. Keyword Model Update

To collect training data for the updating process, each of the 50 keywords is used as the query once. In this case, users’ feedback processes are simulated as follows. For a query image, 5 iterations of user-and-system interaction were carried out. At each iteration, 5 most positive images are labeled by the user. Both positive and negative examples are considered. The initial retrieval result after updating the keyword model is presented, together with that without the updating procedure as a reference. The effect of updating process on subsequent relevance feedback sessions is also evaluated and the retrieval results.

D. Ambiguity search

The main intention of proposing this model of system is to annotate/tag the images with some manually defined concepts for learning an inherent space, using visual and contextual features. More particularly, by making the system to feed the latent vectors into existing classification representations, it can be enforce for application of image annotation, which is considered as the needed issue in image retrieval. As an extension to the available model, we recommend and add the entity feature of the problem of solving the ambiguity. The Conditional Random Filed Algorithm model is used for training the system and results of reinforced online multi-modal distance metric learning system provides an better outcome of content based image retrieval model. This solution is the future enhancement where the contribution of providing more accuracy to the proposed system by enhancing using ambiguity resolving problem.

IV. Proposed CRF Algorithm

The proposed works in two phases namely training and testing phase. In training phase, the input image is taken as user query. For that input image, the ambiguity is set. Then the image is annotated and trained. In testing phase, when the user submits the query, the query will be treated into context and content based systems. The ambiguity will be analyzed. Let us consider that the set of data $X = \{x_1, x_2 .... x_n\} \subset \mathrm{R}^m$. With these set of coordinate, it build the graph using traditional models. $W \in R^{n*n}$ represents the adjacency matrix of an item w_{ij} whose weights estimates the edges between point i and j. The weight can be calculated as:

$$w_{ij} = \exp[-d^2(x_i, x_j)/2\sigma^2)]$$

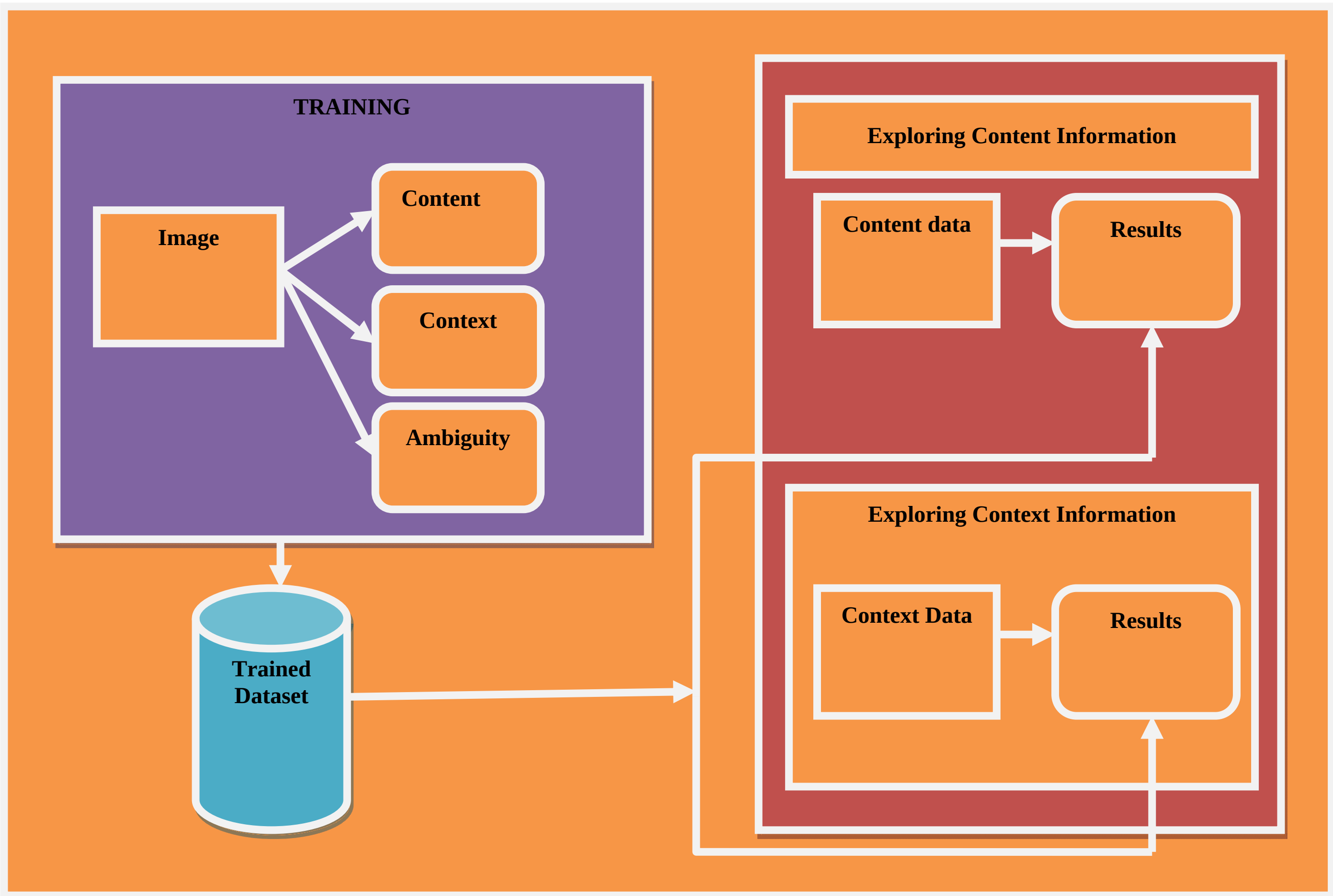
The distance metric of x_i and x_j is given as the function of d (x_i, xj). The ranking matrix can be assigned to each point x_i and its ranking score as r_i. Then the initial vector assigned as

$y= [y1…y_n]^T$

Here, y_i =1 if x_i is a query otherwise y_i=0.

V. System Design

Our Proposed System Architecture



VI. System Analysis

A. Existing System

- In recent years, one promising direction to address this challenge is to explore distance metric learning (DML) by applying machine learning techniques to optimize distance metrics from training data or side information, such as historical logs of user relevance feedback in content-based image retrieval (CBIR) systems.
- As a classical well-known online learning technique, the Perceptron algorithm simply updates the model by adding an incoming instance with a constant weight whenever it is misclassified.
- Recent years have witnessed a variety of algorithms proposed to improve Perceptron, which usually follow the principle of maximum margin learning in order to maximize the margin of the classifier.
- Among them, one of the most notable approaches is the family of Passive-Aggressive learning algorithms, which updates the model whenever the classifier fails to produce a large margin on the incoming instance.

B. Disadvantages Of Existing System

- Although various DML algorithms have been proposed in literature, most existing DML methods in general belong to single-modal DML in that they learn a distance metric either on a single type of feature or on a combined feature space by simply concatenating multiple types of diverse features together.
- In a real-world application, such approaches may suffer from some practical limitations:
- Some types of features may significantly dominate the others in the DML task, weakening the ability to exploit the potential of all features; and
- The naïve concatenation approach may result in a combined high dimensional feature space, making the subsequent DML task computationally intensive.

C. Proposed System

- This paper investigates a novel framework of Online Multi-modal Distance Metric Learning (OMDML), which learns distance metrics from multi-modal data or multiple types of features via an efficient and scalable online learning scheme.
- The key ideas of OMDML are twofold:
- It learns to optimize a separate distance metric for each individual modality (i.e., each type of feature space), and
- It learns to find an optimal combination of diverse distance metrics on multiple modalities.
- We present a novel framework of Online Multimodal Distance Metric Learning, which simultaneously learns optimal metrics on each individual modality and the optimal combination of the metrics from multiple modalities via efficient and scalable online learning
- We further propose a low-rank OMDML algorithm which by significantly reducing computational costs for high-dimensional data without PSD projection.
- We offer theoretical analysis of the OMDML method
- We conduct an extensive set of experiments to evaluate the performance of the proposed techniques for CBIR tasks using multiple types of features.

D. Advantages Of Proposed System

- OMDML takes advantages of online learning techniques for high efficiency and scalability towards large-scale learning tasks.
- To further reduce the computational cost, we also propose a Low-rank Online Multi-modal DML (LOMDML) algorithm, which avoids the need of doing intensive positive semi-definite (PSD) projections and thus saves a significant amount of computational cost for DML on high-dimensional data.

VII. Conclusion

In this work we implemented a new system to make enhancement on the available existing online multi-modal distance metric learning (OMDML) with a new feature of extension to solve the Image ambiguity issue using Conditional Random Field (CRF) Algorithm. The implementation results show that the proposed model is very efficient in providing the solution for the problem of ambiguity in the Content based Image Retrieval System. CRF model works well than the available existing model and results proved it too.

Reference:

Pengcheng Wu, Steven C. H. Hoi, Peilin Zhao, Chunyan Miao, and Zhi-Yong Liu, “Online Multi-Modal Distance Metric Learning with Application to Image Retrieval”, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 28, NO. 2, FEBRUARY 2016.