IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

A Review On Rank Image Tags Using Image Annotation

Shilpa Shree. B[1], Dr.Chethan H.K[2]

PG Scholar, Department of Computer Science and Engineering, Maharaja Institute of Technology (MIT), Mysore, Karnataka, India[1]

Professor, Department of Computer Science and Engineering, Maharaja Institute of Technology (MIT), Mysore, Karnataka, India[2]

Abstract: With the raising in the amount of images in image matching and retrieval in social Media , Image annotation has egressed as an significant research issues due to its practical application .Many social image search engines are based on keyword/tag matching . In this paper , the proposed method combines the prediction models for different tags into a matrix, and casts tag ranking into a matrix recovery problem. It acquaints the matrix trace norm to explicitly control the model complexity, so that a reliable prediction model can be acquired for tag ranking even when the tag space is big.

Keywords: Automatic Image Annotation; Tag Ranking; User Tags; Image Retrieval.

I. INTRODUCTION

Growth in Digital imaging and Internet technologies lead to an volatile increase of Digital images that are available over the Internet . It is very significant to efficiently store and retrieve images for different application such as fashion design, crime prevention, medicine, architecture, etc. For this purpose, many general purpose image retrieval systems have been developed.

Retrieve images from ginormous collections of digital system has become an crucial research topic. Content-based image retrieval (CBIR) addresses this challenge by retrieving the matched images based on their visual similarity to a query image. Limitation of CBIR can be addressed by Tag based image retrieval (TBIR) images which use manually assigned keywords/tags. It allows a user to provide his/her data by textual information and find the relevant images based on the match between the textual query and the assigned image tags.

TBIR is usually more effective than CBIR in identifying the relevant images since it is time-consuming to manually label images, many algorithms have been developed for automatic image annotation , we focus on the tag ranking approach for automatic image annotation.

Instead of having to decide, for each tag, if it should be assigned to a given image, the tag ranking approach ranks tags in the descending order of their relevance to the given image. By avoiding making binary decision for each tag, the tag ranking approach significantly simplifies the problem, leading to a better performance than the traditional classification based approaches for image annotation.

In addition, studies have shown that tag ranking approaches are more robust to noisy and missing tags than the classification approaches

II. LITERATURE SURVEY

A. Content based Image Retrieval (CBIR)
CBIR is used for automatic indexing and retrieval of images depending upon contents of images known as features. The features may be low level or High level. The low level features include colour, texture and shape. The high level feature describes the concept of human brain. The difference between low level features extracted from images and the high level information need of the user known as semantic gap. Semantic gap between Low level visual features and High Level Semantic concept made retrieval performance of CBIR not Satisfactory.

B. Tag based Image Retrieval(TBIR)
TBIR is the application of computer vision techniques to the image retrieval TBIR faces the problem searching for digital images in large databases. TBIR seeks for the tag that the user entered as a search query in the browser of any system in the world. It looks the similar tag that has been attached with the image and retrieves the image to the user.

It didn't check the content of the image; it only checks the tag in the image. TBIR is the most efficient technique in image retrieval but it is dependent the tags. The tags are added manually by the users during the time of uploading. TBIR is not only efficient but also effective.

The performance of TBIR is highly dependent on the availability and quality of manual tags. Analyses have shown that manual tags are often unreliable and inconsistent. Many users tend to choose general and ambiguous tags in order to minimize their efforts in choosing appropriate words, tags that are specific to the visual content of images tend to be missing or noisy, leading to a limited performance of TBIR

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

III. PROJECT WORK

In this Paper, we study the problem of tag completion, where the goal is to automatically fill in the missing tags as well as correct noisy tags for given images. We represent the image-tag relation by a tag matrix, and search for the optimal tag matrix consistent with both the observed tags and the visual similarity .In this section we review the work on automatic image annotation and tag ranking

A. Automatic Image Annotation

Automatic image annotation aims to find a subset of keywords/tags that describes the visual content of an image. It plays an significant role in bridging the semantic gap between low-level features and high-level semantic content of images. Most automatic image annotation algorithms can be classified into three categories (i) generative models that model the joint distribution between tags and visual features, (ii) discriminative models that view image annotation as a classification problem, and (iii) search based approaches. Both mixture models and topic models, two well-known approaches in generative model, have been successfully applied to automatic image annotation. In, a Gaussian mixture model is used to model the dependence between keywords and visual features.

Since a large number of training examples are required for estimating the joint probability distribution over both features and keywords, the generative models are unable to handle the challenge of large tag space with limited number of training images. Discriminative models, views image annotation as a multi-class classification problem, and learn one binary classification model for either one or multiple tags. A structured max-margin algorithm is developed in to exploit the dependence among tags. One problem with discriminative approaches for image annotation is imbalanced data distribution because each binary classifier is designed to distinguish image of one class from images of the other classes. It becomes more severe when the number of classes/tags is large

B. Tag Ranking

Tag ranking aims to learn a ranking function that puts relevant tags in front of the irrelevant ones. In the simplest form, it learns a scoring function that assigns larger values to the relevant tags than to those irrelevant ones. Classification framework for tag ranking that computes tag scores for a test image based on the neighbour voting. It was extended in to the case where each image is represented by multiple sets of visual features. Kernel Density Estimation (KDE) is used to calculate relevance scores for different tags, and performs a random walk to further improve the performance of tag ranking by exploring the correlation between tags. Tang et al. proposed a two-stage graph-based relevance propagation approach. In a two-view tag weighting method is proposed to effectively exploit both the correlation among tags and the dependence between visual features and tags. In a max-margin riffled independence model is developed for tag ranking. As mentioned in the introduction section, most of the existing algorithms for tag ranking tend to perform poorly when the tag space is large and the number of training images is fixed.

C. Low-rank

In mathematics, low-rank approximation is a minimization problem, in which the cost function measures the fit between a given matrix (the data) and an approximating matrix (the optimization variable), subject to a constraint that the approximating matrix has reduced rank. The problem is used for mathematical modelling and data compression. The rank constraint is related to a constraint on the complexity of a model that fits the data. In applications, often there are other constraints on the approximating matrix apart from the rank constraint, e.g., non-negativity and Henkel.

We study the rank, trace-norm and max-norm as complexity measures of matrices, focusing on the problem of fitting a matrix with matrices having low complexity. We present generalization error bounds for predicting unobserved entries that are based on these measures. We also consider the possible relations between these measures. We show gaps between them, and bounds on the extent of such gaps.

D. Matrix recovery

A common modelling assumption in many engineering applications is that the underlying data lies (approximately) on a low-dimensional linear subspace. This property has been widely exploited by classical Principal Component Analysis (PCA) to achieve dimensionality reduction. However, real-life data is often corrupted with large errors or can even be incomplete. Although classical PCA is effective against the presence of small Gaussian noise in the data, it is highly sensitive to even sparse errors of very high magnitude. Paper propose powerful tools that exactly and efficiently correct large errors in such structured data. The basic idea is to formulate the problem as a matrix rank minimization problem and solve it efficiently by nuclear-norm minimization.

IV. IMPLEMENTATION

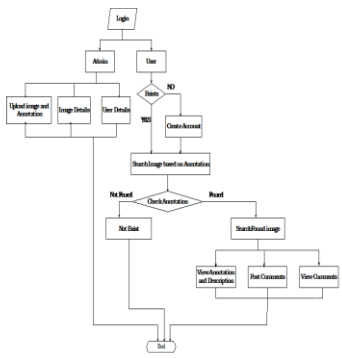

Fig 1 . Flowchart of the proposed framework .

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

Module in the proposed system contain

- Regularization Framework for Tag Ranking
- Optimization
- Image Annotation and Tag Ranking

A. Regularization Framework For Tag Ranking:
The flowchart of the framework is given in Fig 1 .The framework contain two module admin module and User module .In admin module , admin can upload the Image along with the tag ,user details and Image details are monitored . In user module tag based search for the image is done . View and post comments are done .

Consider collection of training images be denoted by I = $\{x_1, x_2,..., x_n\}$, each image $x_i \in R^d$ is a vector of d dimensions and n is the number of training examples. Let $\mathcal{T} = \{t_1,t_2,... t_m\}$ be the set of tags used to annotate images. Let $Y = (y_1,..., y_n) \in \{0, 1\}^{m\times n}$ represent tag assignments for training images, where $y_i \in \{0, 1\}^m$ represents the tag assignment for the ith image. Assigned tag is indicated by $y_{ji} = 1$ to image xi and zero.

Ranking function that assign a higher score to tag t_j than to a tag t_k for image x_i if $y_{ji} = 1$ and $y_{ki} = 0$. More specifically, let $f_i(x)$ be the prediction function for the ith tag, and let (z) be a loss function. Let $\varepsilon_{j,k}(x, y)$ measure the error in ranking tag t_j and t_k for image x with respect to the true tag assignments y . It is defined as follows:

$\varepsilon_{j,k}(x, y) = I(y_j = y_k)$ $l((y_j - y_k)(f_j(x) - f_k(x)))$

where I(z) is an indicator function that outputs 1 when z is true and 0, otherwise. Using the ranking error $\varepsilon_{j,k}(x, y)$, we can now define the ranking error for an individual image x as

$$\varepsilon(x,y) = \sum_{i=1}^{n} \varepsilon_{j,k}(x,y)$$

The overall ranking error for all the training images in collection I as $\sum_{i=n}{}^{n}\varepsilon(x_i, y_i)$. For the simplicity of computation, we restrict the prediction functions $\{fi\}^m_{i=1}$ to linear functions, i.e. $fi(x) = w_i^T x$. Define W = $[w_1,..., w_m] \in R^{d\times m}$ and the overall loss f (W) as

$$f(W) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{m} \varepsilon_{j,k}(\mathbf{x}_i, \mathbf{y}_i)$$
$$= \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{m} I(y_{ji} \neq y_{ki})\ell\big((y_{ji}-y_{ki})(\mathbf{w}_j^\top \mathbf{x}_i - \mathbf{w}_k^\top \mathbf{x}_i)\big)$$

Straight forward approach for tag ranking is to search for a matrix W that minimizes the ranking error f (W). This simple approach is problematic and could lead to the over fitting of training data when the number of training images is relatively small and the number of unique tags is largest\

B. Optimization:
In this module we develop the Optimization process. The main computational challenge in implementing the gradient descent approach for optimizing arise from the high cost in computing the singular value decomposition of Wt . It is known that when the objective function is smooth, the gradient method can be accelerated to achieve the optimal convergence rate. It was shown recently that a similar scheme can be applied to accelerate optimization problems where the objective function consists of a smooth part and a trace norm regularization. In this work, we adopt the accelerated proximal gradient (APG) method for solving the optimization problem. Specifically, we update the solution Wt by solving the following optimization problem. The final component of the accelerated algorithm is to determine the step size ηt, which could have a significant impact on the convergence of the accelerated algorithm.

Trace-norm and max-norm as complexity measures of matrices, focusing on the problem of fitting a matrix with matrices having low complexity. We present generalization error bounds for predicting unobserved entries that are based on these measures. We also consider the possible relations between these measures.

Gradient descent is a first-order optimization algorithm. To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or of the approximate gradient) of the function at the current point. If instead one takes steps proportional to the positive of the gradient, one approaches a local maximum of that function; the procedure is then known as gradient ascent.

Gradient descent is also known as steepest descent, or the method of steepest descent. When known as the latter, gradient descent should not be confused with the method of steepest descent for approximating integrals.

ALGORITHM :

Input: Training image collection $\mathcal{I} = \{\mathbf{x}_i \in \mathbb{R}^d\}_{i=1}^n$, tag assignments for training images $\mathcal{Y} = \{\mathbf{y}_j \in \{0,1\}^m\}_{j=1}^n$, parameter λ.

Initialize:
$\eta_0 = 1,\ \gamma = 2,\ \alpha_1 = 1,\ W_0 = Z_0 = Z_1 \in \mathbb{R}^{d\times m}$
While not converged do
1. Set $\bar{\eta} = \eta_{k-1}$
2. While
$F\left(p_{\bar{\eta}}\left(Z_{k-1}\right)\right) > Q_{\bar{\eta}}\left(p_{\bar{\eta}}\left(Z_{k-1}\right), Z_{k-1}\right)$, set
$\bar{\eta} := \frac{\bar{\eta}}{\gamma}$
3. Set $\eta_k = \bar{\eta}$ and update
$$W_k = p_{\eta_k}(Z_k),$$
$$\alpha_{k+1} = \frac{1+\sqrt{1+4\alpha_k^2}}{2},$$
$$Z_{k+1} = W_k + \left(\frac{\alpha_k - 1}{\alpha_{k+1}}\right)(W_k - W_{k-1})$$

end while
Output: The optimal solution W_* .

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 5, May 2016

Loss function f (W) and the trace norm $\|W\|_*$ are convex, to solve the optimization problem in is gradient descent. For the sake of clarity, we set the loss function in (1) to be a logistic loss. i.e., $\ell(z) = \log\left(1 + e^{-z}\right)$.At each iteration t, given the current solution Wt for , we first compute a sub gradient of the objective function F (W) at W = Wt, denoted by ∇F (Wt), and then update the solution by

$$W_{t+1} = W_t - \eta_t \nabla F(W_t)$$

where ηt > 0 is a step size at the t-th iteration. Let $W_t = U_t \sum_t V_t^T$ be the singular value decomposition of Wt . Since $U_t V_t^T$ is a sub gradient of ||W||* at W = Wt, we have

$$\nabla F(W_t) = \lambda U_t V_t^\top + \frac{1}{n}\sum_{i=1}^{n}\sum_{j,k=1}^{m} \alpha^i_{jk}\mathbf{x}_i \left(\mathbf{e}^m_j - \mathbf{e}^m_k\right)^\top$$

Where

$$\alpha^i_{jk} = I(y_{ji} \neq y_{ki})\ell'((y_{ji} - y_{ki})\mathbf{x}_i^\top(\mathbf{w}_j - \mathbf{w}_k))$$

and e_j^m is a vector of m dimensions with all the elements being zero except that its jth entry is 1.

The main computational challenge in implementing the gradient descent approach for optimizing above equation aries from the high cost in computing the singular value decomposition of Wt. It is known that when the objective function is smooth, the gradient method can be accelerated to achieve the optimal convergence rate of .similar scheme can be applied to accelerate optimization problems where the objective function consists of a smooth part and a trace norm regularization. we adopt the accelerated proximal gradient (APG) method] for solving the optimization problem in. Specifically, according to ,we update the solution Wt by solving the following optimization problem

$$W_t = \arg\min_W \frac{1}{2\eta_t}\|W - W'_t\|_F^2 + \lambda\|W\|_*$$

Where

$$W'_t = W_{t-1} - \eta_t \nabla f(W_{t-1})$$

The optimal solution ,is obtained by first computing the singular value decomposition (SVD) of w_t' and then applying soft-thresholding to singular values of w_t'. More specifically, the optimal solution is given as

$$W_t = U\Sigma_{\lambda\eta_t}V^\top,$$

Where $W_t' = U\sum V^T$ is the SVD of w_t' and $\Sigma_{\lambda\eta_t}$ is a diagonal matrix with its diagonal elements computed as $(\Sigma_{\lambda\eta_t})_{ii} = \max\{0, \Sigma_{ii} - \lambda\eta_t\}$.

The final component of the accelerated algorithm is to determine the step size ηt, which could have a significant impact on the convergence of the accelerated algorithm.

C. Image Annotation And Tag Ranking

In this module we develop the Image annotation. Given the learned matrix W∗ and a test image represented by vector xt , we compute scores for different tags by yt = W_ ∗ xt that indicate the relevance of each tag to the visual content of the test image. The tags are then ranked in the descending order of the relevant scores and only the tags ranked at the top will be used to annotate the test image. Besides image annotation, the learned model can also be used when a subset of tags is provided to the test image and needs to be re-ranked in order to remove the noisy tags.

V. CONCLUSION

The proposed scheme casts the tag ranking problem into a matrix recovery problem and introduces trace norm regularization to control the model complexity. A tag matrix completion method for image tagging and image retrieval. We consider the image-tag relation as a tag matrix, and aim to optimize the tag matrix by minimizing the difference between tag based similarity and visual content based similarity. The proposed method falls into the category of semi-supervised learning in that both tagged images and untagged images are exploited to find the optimal tag matrix. Extensive experiments on image annotation and tag ranking have demonstrated that the proposed method significantly outperforms several state-of-the-art methods for image annotation especially when the number of training images is limited and when many of the assigned image tags are missing. In the future, we plan to apply the proposed framework to the image annotation problem when image tags are acquired by crowd sourcing that tend to be noisy and incomplete.

REFERENCES

[1]. Songhe Feng, Zheyun Feng, and Rong Jin, “Learning to Rank Image Tags With Limited Training Examples”, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 24, NO. 4, APRIL 2015.
[2]. R. Datta, D. Joshi, J. Li, and J. Wang, “Image retrieval: ideas, influences, and trends of the new age,” ACM Computing Surveys, vol. 40, no. 2, 2008.
[3]. L. Wu, R. Jin, and A. Jain, “Tag completion for image retrieval,” IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 35, no. 3, pp.716–727, 2013.
[4]. A. Makadia, V. Pavlovic, and S. Kumar, “Baselines for image annotation,” International Journal of Computer Vision, vol. 90, no. 1, pp.88–105, 2010.
[5]. M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, “Tagprop: discriminative metric learning in nearest neighbor models for image auto-annotation,” in IEEE Int. Conf. on Computer Vision, 2009, pp. 309–316.
[6]. X. Li, C. Snoek, and M. Worring, “Learning social tag relevance by neighbor voting,” IEEE Trans. on Multimedia, vol. 20, no. 11, pp. 1254–1259, 2009.
[7]. J. Zhuang and S. Hoi, “A two-view learning approach for image tag ranking,” in ACM Int. Conf. on WSDM, 2011.
[8]. Karthikram1, Mailaivasan, Parthiban3, Ganesh Assistant Professor, Student, IT Department, Sri Manakula Vinayagar Engineering College, Pondicherry, India ,International Journal of Research in Engineering and Technology on "TAG BASED IMAGE RETRIEVAL (TBIR) USING AUTOMATIC IMAGE ANNOTATION "
[9]. Arif Ur Rahman, Muhammad Muzammal, Humayun Zaheer Ahmad Department of Computer Science, Bahria University, Islamabad, Pakistan Awais Majeed Department of Software Engineering, Bahria University, Islamabad, Pakistan Zahoor Jan Department of Computer Science, Islamia College University, Peshawar, Pakistan . International Journal of Advanced Computer Science and Applications, Vol. 6, No. 12, 2015 on "A Novel Approach for Ranking Images Using User and Content Tags ".
[10]. Miao Fan, Qiang Zhou and Thomas Fang Zheng Center for Speech and Language Technologies, Division of Technical Innovation and Development, Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University Beijing 10084, China . 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology on "Content-based Semantic Tag Ranking for Recommendation.