

Automatic Annotation to Rank the Images

Prof. Sunil B. Hebbale

Associate Professor

*Dept. of CSE, KLE College of Engg. & Tech,
Chikodi,, Belagavi, India*

Prof. Ashwini V. Gavali

Assistant Professor

*Dept. of CSE, KLE College of Engg. & Tech,
Chikodi, Belagavi, India*

Abstract- Due to the popularity of image capturing digital devices and the ease of social network/photo sharing services (e.g., Facebook, Twitter, Flickr), image annotation came into limelight due to its application in image matching and retrieval. Previously image annotation was casted into a multilabel classification problem which had a drawback that it required a large number of training images with clean and complete annotations in order to learn a reliable model for tag prediction. To overcome this limitation we develop a novel approach that combines the strength of tag ranking with the power of matrix recovery. In this work tags are ranked in the descending order of their relevance to the given image, thus simplifying the problem. The proposed method also aggregates the prediction models into a matrix, and casts tag ranking into a matrix recovery problem which introduces the matrix trace norm to explicitly control the model complexity so that tag ranking can be done even when the tag space is large and the number of training images is limited. Experiments on various image data sets show the effectiveness of the proposed framework for tag ranking compared with the previous approaches for image annotation and tag ranking.

Keywords – Automatic image annotation, tag ranking, matrix recovery, low-rank, trace norm.

I. INTRODUCTION

Due to the popularity of digital devices which led to an explosive growth of digital images that are available over Internet, image annotation came into existence. Among enormous collection of digital photos, retrieving images accurately has become an important issue. Content-Based Image Retrieval (CBIR) addresses this issue by identifying the matched images based on their visual similarity to a query image. In CBIR technique there was a performance issue due to the semantic gap between the low-level visual features and the high-level semantic tags. To overcome this limitation of CBIR many algorithms have been developed for Tag Based Image Retrieval (TBIR) that represents images by manually assigned keywords/tags. It allows a user to present his/her information needs by textual information and find the relevant images based on the match between the textual query and the assigned image tags.

Since it is time-consuming to manually label images, various algorithms have been developed for automatic image annotation. As image annotation is viewed as a multi-label classification problem where binary classification model is built for each tag which requires a large number of training images with clean and complete annotations which would be a overhead. Thus in our work 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 which leads to a better performance than the traditional classification based approaches for image annotation. As tag ranking approaches are more robust to noisy and missing tags than the classification approaches we address this limitation by casting tag ranking in the matrix recovery problem. The key idea is to aggregate the prediction models for different tags into a matrix. We also introduce trace norm regularization to capture the dependence among different tags and to control the model complexity by this a reliable prediction model can be learned for tag ranking even when the tag space is large and the number of training images is limited.

II. RELATED WORK

Under this particular section we present the summarized survey study on the previous existing system related to automatic image annotation and tag ranking. As there is a rich documentation available on both the subjects, we portray the studies that are closely related to this work.

Automatic image annotation aims to find a subset of keywords/tags that describes the visual content of an image. It plays an important 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
 - (iii) Search based approach include CBIR and TBIR.
- ❖ In one of the existing system, a Gaussian mixture model is used to model the dependence between keywords and visual features.
- ❖ In another system, kernel density estimation is applied to model the distribution of visual features and to estimate the conditional probability of keyword assignments given the visual features. Topic models annotate images as samples from a specific mixture of topics, which each topic is a joint distribution between image features and annotation keywords.

DISADVANTAGES OF EXISTING SYSTEM:

- ❖ Although multiple algorithms have been developed for tag ranking, they tend to perform poorly when the number of training images is limited compared to the number of tags, a scenario often encountered in real world applications.
- ❖ Another limitation of these approaches is that they are unable to capture the correlation among classes, which is known to be important in multi-label learning.
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 limited.

III. PROPOSED SYSTEM

- ❖ In this work, we have proposed a novel tag ranking scheme for automatic image annotation.
- ❖ We first present the proposed framework for tag ranking that is explicitly designed for a large tag space with a limited number of training images.
- ❖ The proposed scheme casts the tag ranking problem into a matrix recovery problem and introduces trace norm regularization to control the model complexity. 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.

ADVANTAGES OF PROPOSED SYSTEM:

- ❖ The proposed scheme casts the tag ranking problem into a matrix recovery problem and introduces trace norm regularization to control the model complexity.
- ❖ 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.

IV. METHODOLOGY

A. A Regularization Framework for Tag Ranking

The above Figure illustrates the basic idea of the proposed framework for tag ranking. A straightforward 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 overfitting of training data when the number of training images is relatively small and the number of unique tags is large. In order to effectively capture the correlation among different tags, we follow and assume that the linear prediction functions in W are linearly dependent and consequentially W is a low rank matrix, leading to the following optimization problem

$$\min_{W \in \Omega} f(W)$$

where the domain Ω is defined as

$$\Omega = \{W \in \mathbb{R}^{d \times m}, \text{rank}(W) \leq r, \|W\|_2 \leq s\}$$

In optimization problem we restrict the solution W to domain Ω in order to control the complexity of the prediction model.

B. Optimization

Since both the loss function $f(W)$ and the trace norm $\|W\|_*$ are convex, one popular approach for solving the optimization problem is gradient descent. 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 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. Hence, we adopt the accelerated proximal gradient (APG) method for solving the optimization problem.

Algorithm Solving Problem by Accelerated Gradient Algorithm
<p>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.</p> <p>Initialize: $\eta_0 = 1, \gamma = 2, \alpha_1 = 1, W_0 = Z_0 = Z_1 \in \mathbb{R}^{d \times m}$</p> <p>while not converged do</p> <ol style="list-style-type: none"> 1. Set $\bar{\eta} = \eta_{k-1}$ 2. While $F(p_{\bar{\eta}}(Z_{k-1})) > Q_{\bar{\eta}}(p_{\bar{\eta}}(Z_{k-1}), Z_{k-1})$, set $\bar{\eta} := \frac{\bar{\eta}}{\gamma}$ 3. Set $\eta_k = \bar{\eta}$ and update $W_k = p_{\eta_k}(Z_k),$ $\alpha_{k+1} = \frac{1 + \sqrt{1 + 4\alpha_k^2}}{2},$ $Z_{k+1} = W_k + \left(\frac{\alpha_k - 1}{\alpha_{k+1}}\right)(W_k - W_{k-1}).$ <p>end while</p> <p>Output: The optimal solution W_*.</p>

C. Automatic Image Annotation and Tag Ranking

Given the learned matrix and a test image, we compute scores for different tags 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.

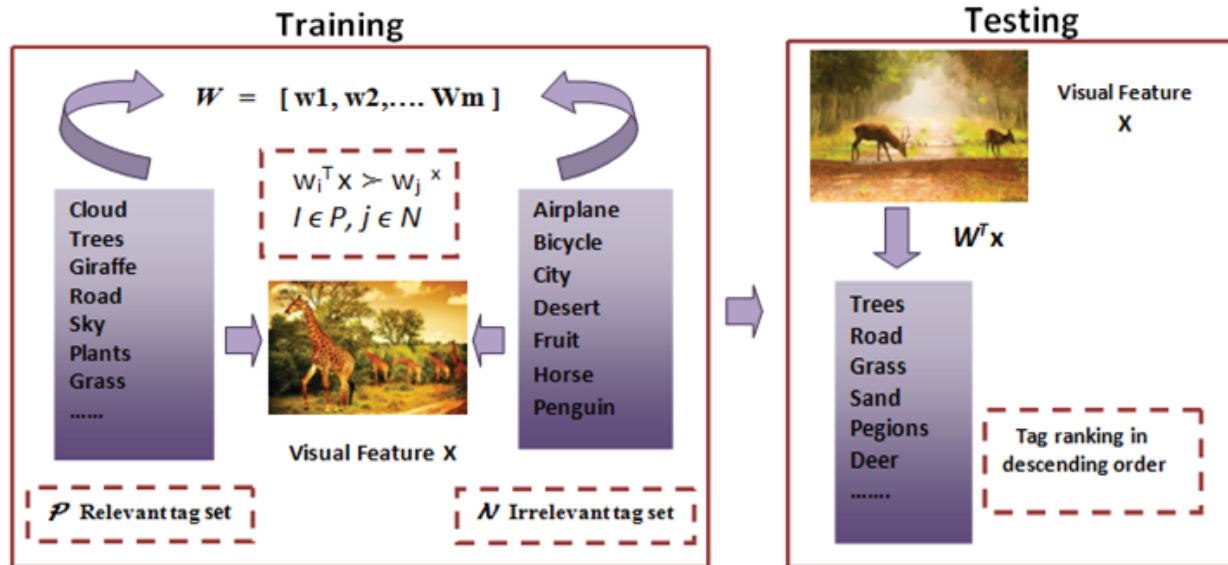


Fig. Schematic illustration of the Proposed Method.

STATISTICS FOR THE DATASETS USED IN THE PREVIOUS EXPERIMENTS.

	Corel5K	ESPGame	IAPRTC-12	Pascal VOC2007	SUNAttribute
No.of images	4,999	20,770	19,627	9,963	14,340
Vocabulary size	260	268	291	399	102
Tags per image	3.4/5	4.69/15	5.72/23	4.2/35	15.5/37
Image per tag	58.6/1,004	363/5,059	386/5,534	53/2,095	2,183/11,878

V. PREVIOUS EXPERIMENTAL RESULTS

In our proposed model, we describe the previous experimental setup including Image Datasets, Feature Extraction, and Evaluation Measures. To verify the effectiveness of the tag ranking approach, we here present three sets of experiments they are as follows,

The first experiment evaluates the performance of image annotation with limited training examples.

The second experiment evaluates the performance of image annotation using training images with missing tags.

The last experiment examines the performance of the proposed algorithm for tag ranking.

Finally we evaluate the sensitivity of the proposed algorithm to parameter λ .

A. Image Datasets

For evaluating the proposed algorithm to image tagging, we conduct extensive experiments on five benchmark datasets for image annotation/tagging, including Corel5K, ESPGame, IAPRTC-12, Pascal VOC2007 and SUN Attribute. In these the first three image datasets are used to evaluate the performance of automatic image annotation, and the last two image datasets are used to evaluate tag ranking and hence a relevance score is provided for every assigned tag.

B. Evaluation Measures

At first, to evaluate the performance of automatic image annotation, we adopt the Average Precision ($AP@K$) and Average Recall ($AR@K$) as the evaluation metrics. These are defined as shown below:

$$AP@K = \frac{1}{nt} \sum_{i=1}^{nt} \frac{Nc(i)}{K}$$

$$AR@K = \frac{1}{nt} \sum_{i=1}^{nt} \frac{Nc(i)}{Ng(i)}$$

where K is the number of truncated tags,

nt is the number of test images,

$Nc(i)$ is the number of correctly annotated tags for the i th test image,

$Ng(i)$ is the number of tags assigned to the i th image. Both average precision and recall compares the automatically annotated image tags to the manually assigned ones.

C. Experimental (I): Automatic Image Annotation with Limited Number of Training Images

In this experiment, the annotation performance of the proposed image tagging method with limited training images is evaluated. At the end, we sample only 10% of images for training and use the remaining 90% for testing. Each experiment is repeated 10 times, with a different splitting of training and testing data. The result is reported based on the average over the trials. The previous approaches of image annotations are used as the baseline approaches in our evaluation. The baseline approaches used are as follows:

- *Joint Equal Contribution Method (JEC)*- It finds appropriate annotation words for a test image based on a k nearest neighbour classifier.
- *Tag Propagation Method (TagProp)*- It propagates the tag information from the labeled images to the unlabeled ones via a weighted nearest neighbour graph.
- *Multi-Class SVM Method (SVM)*- It simply implements One-versus-All (OvA) SVM classifier for each tag, and ranks the tags based on the output probability values.
- *Fast Image Tagging Method (FastTag)*- It explores multi-view learning technique for multi-label learning and introduces a co-regularizer in the objective function to enforce that the predictions based on different views are consistent for most training examples.
- *Efficient Multi-Label Ranking Method (MLR)* -This approach explores the group lasso technique in multilabel ranking to effectively handle the missing class labels.

							
Ground Truth	building door frame sky street window	adult child front house square woman	fog mountain roof stripe train	grandstand lawn roof round spectator stadium	hill landscape mountain rock woman	bike car cycling sky cyclist frame helmet jersey rack landscape roof short	boat green hill jacket lake orange mountain people range shore life sky summit
JEC	<i>table front man house roof wall woman boy building chair</i>	<i>tourist woman child front wall classroom people room table building</i>	<i>front child classroom man sky table board car mountain wall</i>	<i>grass bush hill house lawn man people player short slope</i>	<i>sky landscape hill rock tree lake man mountain bay cliff</i>	<i>cyclist jersey helmet sky side road short bike sand cycling</i>	<i>mountain sea sky shore cloud man house lake rock tourist</i>
TagProp	<i>front wall table man rail house woman level building child</i>	<i>people wall tourist front woman table man side round classroom</i>	<i>man sky mountain front classroom wall child table tourist house</i>	<i>man short sky tree lawn house grass bush rock people</i>	<i>sky rock man landscape mountain jeep sea hill tree palm</i>	<i>cyclist sky helmet jersey short highway road side meadow bike</i>	<i>sea shore sky mountain lake cloud man house boat woman</i>
SVM	<i>rail level front house building table roof wall man sky</i>	<i>people side front wall tree woman tourist building sky house</i>	<i>mountain sky front man cloud wall child classroom table car</i>	<i>man short woman sky tree house grass wall bush people</i>	<i>jeep sky landscape mountain rock lake road hill man tree</i>	<i>sky side landscape highway car tree short road bike people</i>	<i>mountain lake sea sky cloud boat fountain hill house man</i>
FastTag	<i>sky front man stripe level tree people house rack pond</i>	<i>sky front man people rail house frame tree bedside centre</i>	<i>sky gate front man people train tree penguin portrait carpet</i>	<i>sky people front man tree lawn dirt short house tussock</i>	<i>sky front man tree house tussock kid people mountain formation</i>	<i>sky front tree man highway people short rack trouser pinnacle</i>	<i>sky man front lagoon tree shore fountain garden mountain cloth</i>
MLR	<i>house wall room cobblestone door front palm lamp tower bed</i>	<i>house tree landscape people jacket square building sweater shelf dog</i>	<i>mountain house palm flower landscape wall gate orange sky train</i>	<i>people house lawn round view green field building stand stadium</i>	<i>mountain grass landscape house shrub rock wall tussock slope snow</i>	<i>cyclist sky car tree jersey helmet cycling short sign sand</i>	<i>range shore field sky lagoon bay lake river tourist road</i>
Proposed	wall building front window house door street room column balcony	front man house wall people woman child tourist room building	mountain sky train front cloud tourist door roof window wall	stadium lawn slope house grandstand road field tree player people	mountain landscape sky rock middle hill desert lake man cliff	cyclist jersey sky short road cycling helmet bike pole car	mountain sky lake range cloud shore summit sea stone hill

							
Ground Truth	front group meadow stone tourist wall	boat man ocean sea water wave	balcony building sky fountain lamp palm square street	cloud leave palm plant sea sky	bed bedcover bedside curtain lamp room table wall	bush cactus man slope tree woman	face girl hair smile white woman
JEC	<i>slope jersey man tree cyclist helmet cycling wall sky front</i>	<i>man green people grass picture photo family blue woman black</i>	<i>sky front table child grey tree wall classroom house man</i>	<i>man round tee shirt wall woman building child classroom jacket</i>	<i>front room fence gravel child cloud curtain door floor man</i>	<i>man forest grass bush cliff middle adult fence hill leave</i>	<i>website red woman black art box brown building church colors</i>
TagProp	<i>man jersey slope front tree wall mountain bush meadow people</i>	<i>airplane arrow red man blue sky water maga the word ocean</i>	<i>front sky group people table man wall tree room house</i>	<i>man bench blanket wall table woman front child classroom sky</i>	<i>front child people wall room table sky fence man house</i>	<i>man hill grass forest tree pant sky bush rock middle</i>	<i>man cd red white black circle green logo face people</i>
SVM	<i>man front meadow wall cycling cyclist forest helmet tree bike</i>	<i>water man sea sand tree word blue car gray mountain</i>	<i>front group sky wall tree people table man house building</i>	<i>bench sky man wall people woman sea tree building tourist</i>	<i>front child wall room man people tourist fence gravel table</i>	<i>hill pant man sky tree grass bush forest meadow cliff</i>	<i>man logo red black cd face white circle green pink</i>
FastTag	<i>sky man tree bone people front shelter building pant house</i>	<i>water red sky ocean tree smile beak nose green cloud</i>	<i>sky front tree harbour man people house building pot fountain</i>	<i>sky jetty front man tree house corridor edge lagoon flagpole</i>	<i>sky tree table cloth flagpole wall man ravine room neck</i>	<i>sky man front tree people house grass bicycle lawn leave</i>	<i>hair face man smile nose woman girl circle glasses teeth</i>
MLR	<i>building trunk dog man tourist photo sky meadow sand paving</i>	<i>water ocean tree nose blue fire white roof cloud wing</i>	<i>building sky house tree tower palm street people sign man</i>	<i>sky woman house tree sea cloud salt lake orange tower</i>	<i>room bed wall tree house curtain night side painting bedside</i>	<i>grass tree leave sand rock dog bush house cliff trail</i>	<i>white pink square purple face bald nose man word blue</i>
Proposed	front man wall people tourist helmet group woman photo meadow	ocean sea sky water sand blue red boat man tree	building sky lamp tree street front people tower palm grey	sky cloud man sea palm tree lake grass house shore	wall room curtain front bed painting bedside fence blanket table	tree bush man grass forest slope middle rock jungle path	white man hair black red face blue woman girl hat

Fig. Examples of test images from both the ESPGame and IAPRTC-12 datasets with top 10 annotations generated by different methods. The correct tags are highlighted by bold font whereas the incorrect ones are highlighted by italic font.

VI. CONCLUSION AND FUTURE WORK

In this work, we have proposed a novel tag ranking scheme for automatic image annotation. The tag ranked in descending order of their relevant scores and only the tags ranked at the top will be used to annotate the test image. Besides image annotation, the proposed model can also be used when a subset of tags is provided to test image and needs to be re-ranked in order to remove noisy tags. 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. The proposed framework can be applied 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," vol. 24, no. 4, 2015.
- [2] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, no. 2, 2008, Art. ID 5.
- [3] A. Makadia, V. Pavlovic, and S. Kumar, "Baselines for image annotation," *Int. J. Comput. Vis.*, vol. 90, no. 1, pp. 88–105, 2010.
- [4] S. Zhang, J. Huang, Y. Huang, Y. Yu, H. Li, and D. N. Metaxas, "Automatic image annotation using group sparsity," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3312–3319.
- [5] Y. Verma and C. V. Jawahar, "Image annotation using metric learning in semantic neighbourhoods," in *Proc. 12th Eur. Conf. Comput. Vis.*, 2012, pp. 836–849.
- [6] C. Wang, S. Yan, L. Zhang, and H.-J. Zhang, "Multi-label sparse coding for automatic image annotation," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1643–1650.
- [7] H. Wang, H. Huang, and C. Ding, "Image annotation using multi-label correlated Green's function," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 2029–2034.
- [8] D. Liu, X.-S. Hua, L. Yang, M. Wang, and H.-J. Zhang, "Tag ranking," in *Proc. 18th Int. Conf. WWW*, 2009, pp. 351–360.
- [9] Z. Wang, J. Feng, C. Zhang, and S. Yan, "Learning to rank tags," in *Proc. ACM Int. Conf. CIVR*, 2010, pp. 42–49.
- [10] J. Zhuang and S. C. H. Hoi, "A two-view learning approach for image tag ranking," in *Proc. 4th ACM Int. Conf. WSDM*, 2011, pp. 625–634.