International Journal of Scientific and Research Publications, Volume 3, Issue 4, April 2013
ISSN 2250-3153

Introducing Feature Extraction & Least square regression in Automatic Image Annotation

Shalini K.Kharkate*, Prof. Nitin J. Janwe**

* Rajiv Gandhi college of Engg., Research and Technology, Chandrapur.(M.S.),India
** Computer Department, Rajiv Gandhi college of Engg., Research and Technology , Chandrapur., (M.S.),India

Abstract- Automatic image annotation is the process of assigning keywords to digital images depending on the content information. Automatically assigning keywords to images is of great interest as it allows one to index, retrieve, and understand large collections of image data. Selection & implementation of proper feature extraction and weight calculation method plays important role in image annotation. Many techniques have been proposed for image annotation in the last decade that gives reasonable performance on standard datasets. This paper presents detailed literature about the Automatic Image Annotation and explains the feature extraction as well as weight calculation method.

Index Terms- Automatic image annotation, Literature, feature vector, concept modeling,Saliency.

I. Introduction

With the prevalence of digital imaging devices such as digital cameras and mobile phones, a large number of images are produced everyday. An emerging issue is how to efficiently and effectively search required image items from a huge image database. A promising approach is keyword-based image retrieval that allows the user to search for image items using keywords, but traditional manual image annotation is both laborious and subjective. Hence automatic image annotation has attracted considerable attention. Automatic Image annotation has been a topic of on-going research and several techniques have been proposed. [1,2,3,4,7,8,9,10]

Automatic image annotation consists of analyzing the visual content (colors, textures, shapes) of the images (or the objects existing in the images) in order to transform it into meaningful symbolic (or textual) information.. It is used in many applications domains such as entertainment, commerce education, biomedicine, military, and web image classification also saves lot of manual work of human. But AIA is difficult task because visual content to be analyzed depend on many factors such as shooting conditions, instances of objects, lighting condition, resolution of camera , and the background clutters.

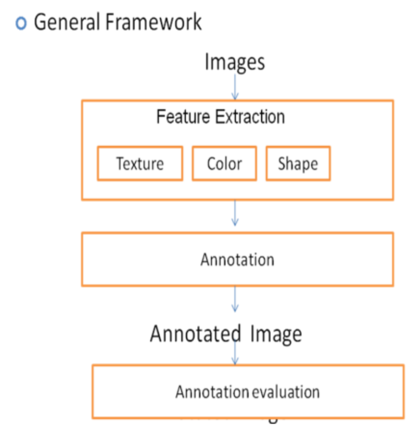

Fig. 1 General Framework Automatic Image annotation

State of the art automatic image annotation systems can be analyzed and grouped from various point of views. The remainder of paper is organized as follows. Review of the existing approaches is explain in section II. Proposed method is explain in III and we conclude our work in section IV.

II. RELATED WORK

A. Image Annotation

In Automatic Image Annotation three main groups are identified according to learning techniques and the application domain. classification model, probabilistic model and graph based models.

a) Classification model: Use discriminative model to classify images. Annotating images treats as classifying it to number of predefined groups called concepts or keyword. Since image is annotated by more than two words . So AIA problem is also considered as multi classification problem Classifier calculates the similarity of all trained classes and models unobserved instance to the class with highest similarity[6,7,9,12]. Both supervised[21,28] and unsupervised learning methods are used for classification. In supervised classification pre-trained classes are used whereas in unsupervised clustering used for classification.

b) Probabilistic model: Generative models are used to specify joint probability of image region feature and set of words [13,21,22] word to image relation and word to word relation are estimated. After estimating words are ranked according to probabilities. Probabilistic model are more flexible in learning process but computational cost of parameter estimation is more.

c) Graph based model : Various graphical models are used for finding relation between word to image or word to word relation.[3,11,27] However time complexity and space complexity makes it difficult to apply on real word annotation.

Authors Reviewed many supervised learning models such as probabilistic classifier, artificial neural network, support vector machine, decision trees ,k-nearest neighbor, majority voting, bagging, boosting, stacked generalization. Also reviewed hybrid learning models which based on both clustering and classification[15].Most of these treat annotation as translation from image instances to keywords. The translation paradigm is typically based on some model of image and text co-occurrences.

B. Features

The essential requirement in image retrieval, indexing, classification is extracting efficient features from images. The color, texture and shape features are the most widely used visual features. There are many color spaces used RGB,HSV, LUV etc. Color feature is extracted from images or regions. Many color feature can be used such as color moments such as mean, standard deviation and skewness. Usually they are calculated for each color channels (components) separately .But all color information in image is not represented by moments.

The color histogram describes the color distribution of an image. The color coherence vector (CCV) incorporates spatial information into the basic color histogram. CCV is preferable than histrogram.

MPEG-7 also standardizes a number of color features including dominant color descriptor (DCD), color layout descriptor (CLD), color structure descriptor (CSD), and scalable color descriptor (SCD).

Texture is another important image feature.

Many methods have been proposed. In spatial approach, texture features are extracted by finding the local pixel structures in original image domain. This method is easy to understand but sensitive to noise.

In spectral texture feature extraction techniques, an image is transformed into frequency domain and then feature is calculated from the transformed image. For images or regions with sufficient size, spectral texture features best choice. However, for small images or regions, especially when the regions are irregular, spatial features desirable.

Shape feature is useful for extracting feature in many applications. Many region shape descriptors are commonly used in color image retrieval, including, area, moments, circularity, and eccentricity. In this work salient regions are consider for color feature extraction.

C. Saliency

Human eye is perceptually more sensitive to certain colors and intensities and objects with such features are considered more salient. Salient regions are most important point in image which attracts greater attention by visual system than other part of the image. These regions has distinctive features when compared with others in image. Eg. a polar bear is salient on dark rocks, but almost invisible in snow.

Recently, several saliency approaches came up that are based on computational and mathematical ideas and usually less biologically motivated. These approaches range from the computation of entropy[28],[29] over determining features that best discriminate between a target and a null hypothesis to learning the optimal feature combination with machine learning techniques.

In[30] proposed saliency detection method spectral residual. Spectral residual is the difference between original log spectrum and its mean-filtered version. The saliency map is obtained by applying inverse Fourier transform to the spectral residual. We compute the color histogram of saliency regions for the color space RGB..

III. PROPOSED WORK

A. Automatic Image Annotation

AIA is a process of assigning keywords to digital images automatically depending on the content information.

More formally, an image I with set of visual features Vi={V1,V2….Vn} and set of keywords W={w1,w2….wm},find automatically the keyword subset Wi which describe image I.

B. Concept

A training set S consisting of N images with n feature vectors. n feature vectors forms a feature matrix and A pair of similar and dissimilar images(L). The main purpose of this paper is to investigate the feature selection properties in the image annotation task. This image pair setting helps us to create a feature matrix that contains the same groups of features. Thus, we can directly do feature analysis on this matrix within the same framework.

Calculate the weight assign to each feature vector by using feature matrix and L, which is final step of training stage. Weight vector is used to find relevancy of keyword to the image. Sufficient training is necessary to have correct annotations to the image in testing stage.

Feature vector of input image is compared with feature matrix of the training images. Based on the weights calculated, most similar images are find out from L. Keywords from L are getting assigned to test image which are annotations.

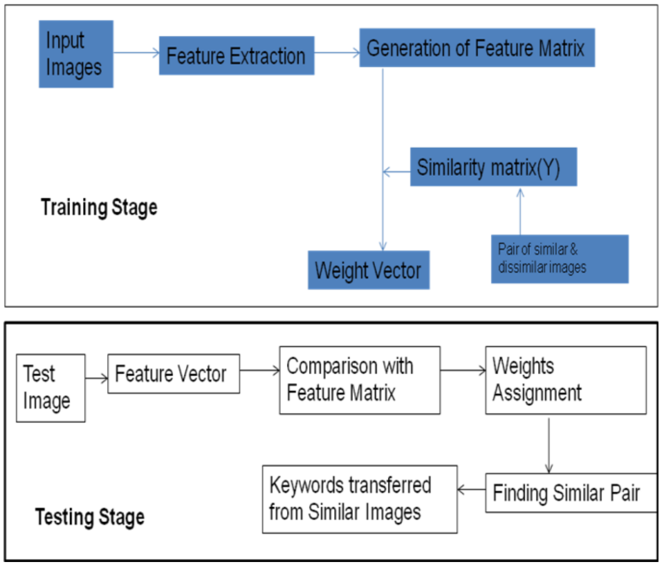

Fig 2.Proposed Architecture

C. *Feature Extraction*

[30] Paper argues that the spectrum residuals corresponds to image saliency. Given an input image, the log spectrum L(*f*)is computed from the down-sampled image with height (or width) equals 64 px. The selection of the input size is related to visual scale.

If the information contained in the L(*f*) is obtained previously,the information required to be processed is:

$$H(\mathrm{R}\ (\mathrm{f})) = H(\mathrm{L}(f))|\mathrm{A}(\mathrm{f})) \quad (1)$$

Where A(f) denotes the general shape of log spectra, which is given as prior information. R (f) denotes the statistical singularities that is particular to the input image, define R (f) as the spectral residual of an image. The average spectrum A(f) can be approximated by convoluting image

$$\mathrm{A}(\mathrm{f})= h_n(f) * \mathrm{L}(f) \quad (2)$$

Therefore the *spectral residual*R(f) can be obtained by:

$$\mathrm{R}(\mathrm{f}) = \mathrm{L}(\mathrm{f}) - h_n(f) * \mathrm{L}(f) \quad (3)$$

And saliency is obtained by

$$\mathrm{S}(\mathrm{x})=\mathrm{F}^{-1}[\exp(\mathrm{R}(\mathrm{f}) +\mathrm{P}(f))]$$

In order to obtain a better visual display the final saliency map was actually presented as

$$\mathrm{S}(\mathrm{x}) = g(\mathrm{x}) * \mathrm{F}^{-1}[\exp (\mathrm{R}(\mathrm{f}) + \mathrm{P}(\mathrm{f})]^2 \quad (4)$$

Where F and F^{-1} denote Fourier transform and inverse Fourier transform respectively; *g and* h_n are low pass filter. P(f) denotes phase spectrum of image which is assumed to be preserved during process.

D. *Weight Vector calculation*

Here weight of each feature vector is calculated by using feature matrix obtained using SR method and set of similar and dissimilar image pair. In this setting we consider any pair of images that share enough keywords to be positive training samples and any pair with no keywords in common to be negative example. In this work we obtained training samples from the designated training set of the Corel5K dataset. Image paires that had at least four common keywords were treated as positive sample for training and those with no common keywords were used as negative samples[31].

Weighted least square is an efficient method that makes good use of small data sets.It also shares the ability to provide different types of easily interpretable statistical intervals for estimation , prediction

The most popular loss function to calculate w in this regression problem is the least square estimate, which is also named as the minimizer of the residual sum of squared errors and is given as

$$w = \arg\min_{w \in \mathbb{R}p} \|Xw - Y\|^2_2$$

This weight is used for testing stage in image annotation task.

Here we are using image pairs instead of one to-the-others mode, In this image annotation task, we have hundreds of keywords to model. Each image may have three to five keywords. Furthermore, different data set can have different sets of keywords. Thus, it is complicated to model exhaustively such relations using one-to-the-others mode. Image pair setting does not need to concern modelling these keywords directly. Instead, it just models the keyword similarity among images. Thus, it can easily assign keywords to testing images using existing annotations from training data.

IV. CONCLUSION

In this paper, detailed literature about various image annotation algorithms is mentioned. Proper feature extraction and weight calculation method plays important role in image annotation. This paper explains the method for the same. These methods are used as training stage in image annotation task. Future work is to work for testing stage and to implement this algorithm and to compare with various algorithms available.

REFERENCES

[1] Tianxia Gong, Shimiao Li, Chew Lim Tan,"A Semantic Similarity Language Model to Improve Automatic image annotation" 22nd International Conference on Tools with Artificial Intelligence, 2010

[2] Rami albatal, Philippe mulhem, Yues chiaramella, "A new ROI grouping schema for automatic image annotation"

[3] Jing Liua,, Mingjing Lib, Qingshan Liua, Hanqing Lua, Songde Ma "Image annotation via graph learning", Pattern Recognition 42 (2009) 218 -- 228

[4] Yunhee Shin, Youngrae Kim, Eun Yi Kim," Automatic textile image annotation by predicting emotional conceptsfrom visual features". Image and Vision Computing 28 (2010)

[5] Y. Alp Aslandogan and Clement T. Yu, Senior Member, IEEE, Techniques and Systems for Image and Video Retrieval VOL. 11, NO. 1, JANUARY/FEBRUARY 1999

[6] Lei Ye, Philip Ogunbona and Jianqiang Wang" Image Content Annotation Based on Visual Features" Proceedings of the Eighth IEEE International Symposium on Multimedia (ISM'06).

[7] Dongjian He & Yu Zheng, Shirui Pan, Jinglei Tang-2010 3rd International Congress on Image and Signal Processing, "Ensemble Of Multiple Descriptors For Automatic Image Annotation"

[8] Hamid ansari, Mansour Jamzad "Large-Scale Image Annotation using Prototype-based Models"7th International Symposium on Image and Signal Processing and Analysis (ISPA 2011) September 4-6, 2011, Dubrovnik, Croatia

[9] Golnaz Abdollahian, Murat Birinci †, Fernando Diaz-de-Maria ‡, Moncef Gabbouj †, Edward J. Delp "A region-dependent image matching methodfor image and video annotation" 2011 IEEE

[10] Yinjie Lei Wilson Wong Mohammed Bennamoun Wei Liu "Integrating Visual Classifier Ensemble with TermExtraction for Automatic Image Annotation",2011

[11] Hua Wang, Heng Huang and Chris Ding, pages 126–139, 2011, "Image Annotation Using Bi-Relational Graph of Images and Semantic Labels"

[12] Ran Li, YaFei Zhang, Zining Lu, Jianjiang Lu, Yulong Tian-2010 Second International Conference on MultiMedia and Information Technology, "Technique of Image Retrieval based on Multi-label Image Annotation"

[13] J. Liu, B. Wang, M. Li, Z. Li, W. Y. Ma, H. Lu and S. Ma, "Dual Cross-Media Relevance Model for Image Annotation," in Proceedings of the 15th International Conference on Multimedia 2007, p. 605 – 614.

[14] T. Jiayu, "Automatic Image Annotation and Object Detection," PhD thesis, University of Southampton, United Kingdom, May 2008.

[15] C. F. Tsai and C. Hung, "Automatically Annotating Images with Keywords: A Review of Image Annotation Systems," Recent Patents on Computer Science, vol 1, pp 55-68, Jan., 2008.

[16] R. Datta, D. Joshi, J. Li and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," ACM Computing Surveys (CSUR), vol 40, Apr. 2008.

[17] Chung, K.P. and Fung, C.C. Relevance feedback and intelligent technologies in content-based image retrieval system for medical applications. Australian Journal of Intelligent Information Processing Systems, 8 (3). pp. 113-122, 2004.

[18] Ning Yu ⇑, Kien A. Hua, Hao Cheng." A Multi-Directional Search technique for image annotation propagation" J. Vis. Commun. Image R. 23 (2012) 237–244

[19] E. Y. Chang, K. Goh, G. Sychay, and G. Wu. CBSA: content-based soft annotation for multimodal image retrieval using bayes point machines. IEEE Trans. Circuits Syst. Video Techn, 13(1):26–38, 2003.

[20] E. Izquierdo and A. Dorado. Climbing the semantic ladder: Towards semantic semi-automatic image annotation using MPEG-7 descriptor schemas. In Proc. IEEE Int'l Workshop on Computer Architecture for Machine Perception, pages257–262, Italy, Jul 2005.

[21] J. Jeon, V. Lavrenko, and R. Manmatha, "Automatic image annotation and retrieval using cross-media relevance models," in SIGIR '03: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, (New York, NY, USA), pp. 119–126, ACM Press, 2003.

[22] V. Lavrenko, R. Manmatha, and J. Jeon, "A model for learning the semantics of pictures," in Advances in Neural Information Processing Systems 16 (S. Thrun, L. Saul, and B. Sch¨olkopf, eds.), Cambridge, MA: MIT Press, 2004.

[23] Fei Shi, Fangfang Yang, Jiajun Wang," Supervised Semantic Image Annotation Using Region Relevance 2012 International Conference on Medical Physics and Biomedical Engineering

[24] Hichem Bannour and Céline Hudelot Applied Mathematics and Systems Department, Ecole Centrale Paris, "Towards Ontologies for Image Interpretation and Annotation" 2011 IEEE

[25] Feichao Wang "A survey on automatic image annotation and trends of the new age. 2011,Elsevier

[26] Hao Ma, Jianke Zhu, Member, IEEE, Michael Rung-Tsong Lyu, Fellow, IEEE, and Irwin King, Senior Member, IEEE "Briding the semantic gap between I,age content and bags"

[27] O zge O ztimur Karadag, Akdeniz University, Fatos¸ T. Yarman Vural, Member IEEE Middle East Technical University Ankara, Turkey.,IEEE 09, "HANOLISTIC: A Hierarchical Automatic Image Annotation System Using Holistic Approach",

[28] S. Frintrop, E. Rome, and H. I. Christensen. "Computational visual attention systems and their cognitive foundation: A survey." ACM Trans. on Applied Perception, 7(1), 2010

[29] T. Kadir, A. Zisserman, and M. Brady. "An affine invariant salient region detector." In Proc. of ECCV, 2004

[30] Xiaodi Hou and Liqing Zhang "Saliency Detection: A Spectral Residual Approach"

[31] Ameesh Makadia1, Vladimir Pavlovic2, and Sanjiv Kumar "A New Baseline for Image Annotation"

AUTHORS

First Author – Shalini K.Kharkate ,Rajiv Gandhi college of Engg., Research and Technology , Chandrapur.(M.S.),India shalin_k1@rediffmail.com

Second Author – Prof. Nitin J. Janwe, Computer Department Rajiv Gandhi college of Engg., Research and Technology , Chandrapur., (M.S.),India