Summary on

A survey on automatic image annotation and trends of the new age

Introduction:

This paper is based on automatic image annotation problem. Based on this researching problem they proposed some model which give idea how we can annotate image automatically by using these model. Here they have discussed about three type of model which have different method for solving image annotation problem.

Generative Model:

Generative Model is observing data randomly from some hidden parameters. It is used in machine learning for modeling data directly or also use as an intermediate step to forming a conditional probability density function. In this model, recognition is a process of annotating image regions with words. Firstly, images are segmented into regions. Then image are classified into region using a variety of features. It is mapping with the image between region types and keywords. That region can be describes using a small vocabulary of blobs. That blobs are generated from image features using clustering. From training set of image with annotation that probabilistic models can predict the probability of generating a word given the blobs in an image. By learning the semantics of images it automatically annotate an image with keywords and to retrieve images based on text queries.

Another kind of generative model is topic model, which is also widely used in automatic image annotation. Topic model could discover the relationship between image visual feature and annotation. However, there are some parameters in topic model to be estimated and optimized latent topic number is not easy to obtain. On the other hand, topic model based automatic annotation method is more suitable for small-scale image dataset.

Discriminative Model:

Discriminative models are a class of models used in machine learning for modeling the dependence of an unobserved variable y on an observed variable x . Within a statistical framework, this is done by modeling the conditional probability distribution $P(y|x)$, which can be used for predicting y from x. Discriminative model do not allow to generate samples from the joint distribution of x and y. the main idea of this model based on automatic image annotation problem to classification problem. In this paper they proposed some frame for classification image that are-

SML In this work, a probabilistic formulation for semantic image annotation and retrieval is proposed. Annotation and retrieval are posed as classification problems where each class is defined as the group of database images labeled with a common semantic label.

HSVM-MIL Heuristic Support Vector Machine-based MIL algorithm to learn the correspondence between image regions and keywords under Multiple Instance Learning (MIL) framework. It tries to change the class label of only one instance to minimize the classification risk

Hierarchical classification framework for bridging the semantic gap effectively and achieving multi-level image annotation automatically. In this work, the semantic gap between the low-level computable visual features and the users' real information needs is partitioned into four smaller gaps, and multiple approaches are proposed to bridge these smaller gaps more effectively.

Graph Model:

In this paper they proposed three types of graph model and these graph has two basic process. First annotation process which is the image-based graph learning is utilized to obtain the candidate annotations and the other is the annotation refinement process, the word based graph learning is used to refine those candidate annotations from the prior process.

First model is **BGRM**, bipartite graph reinforcement model is proposed for web image annotation. Given a web image, a set of candidate annotations is extracted from its surrounding text and other textual information in the hosting web page. As this set is often incomplete, it is extended to include more potentially relevant annotations by searching and mining a large-scale image database. All candidates are modeled as a bipartite graph. Then a reinforcement algorithm is performed on the bipartite graph to re rank the candidates. Only those with the highest ranking scores are reserved as the final annotations.

Second is an automatic image annotation approach based on **Automatic Multimedia Cross-modal Correlation Discovery**. The main idea of this work is to represent all the objects, as well as their attributes as nodes in a graph. For multimedia objects with m attributes, we obtain an (m+1)-layer graph. There are m types of nodes and one more type of nodes for the objects.

Third is a graph learning framework for image annotation. In this work, the image-based graph learning is performed to obtain the candidate annotations for each image. In order to capture the complex distribution of image data, the authors proposed a Nearest Spanning Chain (NSC) method to construct the image-based graph, whose edge-weights are derived from the chain-wise statistical information instead of the traditional pairwise similarities. Moreover, the word-based graph learning is developed to refine the relationships between images and words to get final annotations for each image.

Conclusion: In this paper we came to know about three types of model for image annotation. All three models are for automatic image annotation. Each model are differ from another, generative models are typically more flexible than discriminative models in expressing dependencies in complex learning tasks and discriminative models differ from generative models in that they do not allow one to generate samples from the joint distribution of x and y . Lastly the graph model time complexity and space complexity are always high, and it is difficult to apply it directly in real world image annotation but we can try to graph model for parallel processing to increase computing speed.