

## Abstract

Recent years has seen the emergence of thousands of photo sharing websites on the Internet where billions of photos are being uploaded every day. All this visual content boasts a large amount of information about people, objects and events all around the globe. It is a treasure trove of useful information and is readily available at the click of a button. At the same time, significant effort has been made in the field of text mining, giving birth to powerful algorithms to extract meaningful information scalable to large datasets. This thesis leverages the strengths of text mining methods to solve real world computer vision problems. Leveraging such techniques to interpret images is accompanied by its own set of challenges. The variability in feature representations of images makes it difficult to match images of the same object. Also, there is no prior knowledge about the position or scale of the objects that have to be mined from the image. Hence, there are infinite candidate windows which have to be searched.

The work at hand tackles these challenges in three real world settings. We first present a method to identify the owner of a photo album taken off a social networking site. We consider this as a problem of prominent person mining. We introduce a new notion of prominent persons, where information about the location, appearance and social context is incorporated into the mining algorithm to be effectively able to mine the most prominent person. A greedy solution based on an eigenface representation is proposed and We mine prominent persons in a subset of dimensions in the eigenface space. We present excellent results on multiple datasets - both synthetic as well as real world datasets downloaded from the Internet.

We next explore the challenging problem of mining patterns from architectural categories. Our mining method avoids the large numbers of pair-wise comparisons by recasting the mining in a retrieval setting. Instance retrieval has emerged as a promising research area with buildings as the popular test subject. Given a query image or region, the objective is to find images in the database containing the same object or scene. There has been a recent surge in efforts in finding instances of the same building in challenging datasets such as the Oxford 5k dataset[46], Oxford 100k dataset and the Paris dataset[48]. We leverage the instance retrieval pipeline to solve multiple problems in computer vision. Firstly, we ascend one level higher from instance retrieval and pose the question: *Are Buildings Only Instances?* Buildings located in the same geographical region or constructed in a certain time period in history often follow a specific method of construction. These architectural styles are characterized by certain features which distinguish them from other styles of architecture. We explore, beyond the idea of buildings as instances, the possibility that buildings can be categorized based on the architectural style. Certain char-

acteristic features distinguish an architectural style from others. We perform experiments to evaluate how characteristic information obtained from low-level feature configurations can help in classification of buildings into architectural style categories. Encouraged by our observations, we mine characteristic features with semantic utility for different architectural styles from our dataset of European monuments. These mined features are of various scales, and provide an insight into what makes a particular architectural style category distinct. The utility of the mined characteristics is verified from Wikipedia.

We finally generalize the mining framework into an efficient mining scheme applicable to a wider varieties of object categories. Often the location and spatial extent of an object in an image is unknown. The matches between objects belonging to the same category are also approximate. Mining objects in such a setting is hard. Recent methods [55] model this problem as learning a separate classifier for each category. This is computationally expensive since a large number of classifiers are required to be trained and evaluated, before one can mine a concise set of meaningful objects. On the other hand, fast and efficient solutions have been proposed for the retrieval of instances (same object) from large databases. We borrow, from instance retrieval pipeline, its strengths and adapt it to speed up category mining. For this, we explore objects which are “near-instances”. We mine several near-instance object categories from images obtained from Google Street View. Using an instance retrieval based solution, we are able to mine certain categories of near-instance objects much faster than an Exemplar SVM based solution.