

## Abstract

The images can be considered as a union of many parts or a composition of multiple segments. Some of the examples of such parts/segments can be different objects present in the image, foreground and background regions, or the textual (regions containing text) and non-textual regions in the image. In order to get the inherent semantics and higher level knowledge associated with the images, getting insight of such parts/segments is essential. In this thesis, we introduce the notion of “decomposition” in images. The decomposition refers to the phenomenon of break down of images into its constituents meaningful parts/segments. Those meaningful parts depends on the task we are interested in. In case of foreground and background decomposition, it can be a pixel accurate segmentation of a foreground or the tight rectangular box enclosing the foreground segment. In this work, we have discussed the problem of decomposition in two kinds of images – natural images and document images. We also show how popular computer vision tasks such as object detection, semantic segmentation, document layout analysis, word spotting, etc. can be perceived as a decomposition task.

In this thesis, we solve two decomposition problems in a linear optimization framework. Firstly, decomposing a global histogram of a natural image into histograms of its associated objects and regions and secondly, decomposing a questioned document image into regions containing copied and non-copied/original contents in a recognition free setting.

The decomposition of a global histogram representation of an image into histograms of its associated objects and regions is formulated as an optimization problem, given a set of linear classifiers, which can effectively discriminate the object categories present in the image. This decomposition bypasses harder problems associated with localization and the explicit pixel-level segmentation. Our decomposition framework is also applicable for separating histograms of object and background in an image. Our solution is computationally efficient and we demonstrate its utility in multiple situations. We evaluate our method on a wide variety of composite histograms and also compare it with MRF-based solutions. We decompose histograms at an average accuracy of 86.4% on a Caltech-256 based dataset. In addition to merely measuring the accuracy of decomposition, we also show the utility of the estimated object and background histograms for the task of image classification on PASCAL VOC 2007 dataset.

To solve the problem of decomposition in questioned document images, we detect documents from the database which have exact or similar text to a given query document or region image. Exact duplicate is the direct cut and paste of content from multiple documents in the database whereas near duplicate document segments (similar text) could arise due to various document manipulations like summariza-

tion, copying, rewriting, editing, formatting, cut-and-paste, etc. We refer to the corresponding problems as retrieval of exact and near duplicate document images. We formulate the problem as a document retrieval task, and solve it in a recognition-free setting. We propose two approaches which are capable of detecting regions generated by these operations accurately without depending on a reliable OCR. First approach is based on modelling the solution as finding a mixture of homographies, and designing a linear programming (LP) based solution to compute the same while the second approach is based on learning a discriminative classifier for a questioned document region to retrieve duplicate documents. Using both the approaches, we get encouraging results.