

# Abstract

This thesis studies and develops different image feature representation and learning methods for visual search applications. We will develop image feature representation and learning methods using both handcrafted features as well as deep learning features. Handcrafted features have shown good performance in visual search problems such as instance (specific) object recognition. The handcrafted-based methods are generally more light-weight, easier to train, and do not require large training data. However, they have been outperformed by deep learning methods in various visual search problems in recent years. Deep learning methods and features have demonstrated strong performance in visual classification and retrieval. Nevertheless, deep learning methods generally require significant data and computational power during training. Further, they still experience limitations in certain issues such as inadequate handling of geometric consistency when performing image matching.

In view of this, this thesis will study both handcrafted methods as well as deep learning methods for two selected visual search applications. In the first application, we will develop handcrafted-based methods for image landmark recognition. Landmark recognition is a challenging problem in instance object recognition as it involves 3-D objects under different imaging conditions such as viewpoint and scale variation, occlusion, background clutters, etc. In addition, it also encounters the issue of repetitive features (e.g. windows, tiles, pillar, etc.) that poses additional difficulty in landmark recognition. To address these problems, we will propose feature detection and analysis using handcrafted features to perform landmark recognition. In the second application, we will develop deep learning methods for visual fashion clothing search and recommendation. Fashion clothing search is an emerging research problem with huge commercial opportunities. Fashion search is a challenging problem as clothing is deformable, contains multiple attributes, and experiences various imaging conditions such as pose change, background clutters, etc. To address this problem, we will develop deep metric learning and feature representation to perform fashion search. The following paragraphs will explain the contributions of the thesis in greater details.

The first work studies the effect of repetitive patterns in images for visual search. Repetitive patterns are ubiquitous in many cityscape images. These patterns will cause visual burstiness, impact the image representation and negatively affect the search performance. In order to alleviate this issue, we present a new lattice-support repetitive local feature detection (LS-RLF) technique to effectively detect and represent the repetitive patterns in images. We also present a new feature repetitiveness similarity (FRS) metric that uses the information from repetitive patterns to enhance the search performance. Experiments conducted on three benchmark datasets namely, Oxford, Paris and Inria Holidays datasets show the effectiveness of the proposed methods.

The second work studies deep learning methods for fashion clothing visual search. Convolutional Neural Networks (CNN) are used for feature learning and image representation. In this work, we also develop a brand-aware fashion search which takes user brand preference into account during the search. To this end, we construct a new clothing dataset called the NTUFashionBrand dataset with 10K fashion images which are annotated with rich attribute information. We develop two different methods for fashion search. In the first method, we propose a deep feature encoding known as Principal Component Maximum Activation of Convolutions (PMAC) which leverages on hierarchies of CNN activations to extract rich visual representation from clothing images. In the second method, we propose a new deep metric learning framework which incorporates image attribute information to supervise the triplet network training. This serves two purposes: (i) mining of informative triplets, especially when the exact positive-pairs annotation is not available, and (ii) treating the triplets in a soft-manner based on their importance, which helps in capturing similarity at different levels. Experiments conducted on the NTUFashionBrand and Deep Fashion datasets show the effectiveness of the proposed method for visual search and recommendation.

The third work studies methods for fashion clothing popularity prediction based on visual analysis of clothing images. The clothing popularity prediction can help e-commerce companies to manage their production, supply and inventory. In this work, we develop a new framework for fashion popularity prediction based on click rate estimation. A high click rate of a clothing item displayed on an online portal indicates that the item is attractive and popular to the user. We use CNNs to extract useful visual features from clothing images. In addition, the developed framework also incorporates several exogenous factors such as price, discount and brand information. Experiments conducted on datasets collected from an online retail shop shows satisfactory results for fashion clothing popularity prediction.