

Effective Multi-Query Expansions: Collaborative Deep Networks for Robust Landmark Retrieval

Abstract:

Given a query photo issued by a user (q-user), the landmark retrieval is to return a set of photos with their landmarks similar to those of the query, while the existing studies on the landmark retrieval focus on exploiting geometries of landmarks for similarity matches between candidate photos and a query photo. We observe that the same landmarks provided by different users over social media community may convey different geometry information depending on the viewpoints and/or angles, and may, subsequently, yield very different results. In fact, dealing with the landmarks with low quality shapes caused by the photography of q-users is often nontrivial and has seldom been studied. In this paper, we propose a novel framework, namely, multi-query expansions, to retrieve semantically robust landmarks by two steps. First, we identify the top-k photos regarding the latent topics of a query landmark to construct multi-query set so as to remedy its possible low quality shape. For this purpose, we significantly extend the techniques of Latent Dirichlet Allocation. Then, motivated by the typical collaborative filtering methods, we propose to learn a collaborative deep networks-based semantically, nonlinear, and high-level features over the latent factor for landmark photo as the training set, which is formed by matrix factorization over collaborative user-photo matrix regarding the multi-query set. The learned deep network is further applied to generate the features for all the other photos, meanwhile resulting into a compact multi-query set within such space. Then, the final ranking scores are calculated over the high-level feature space between the multi-query set and all other photos, which are ranked to serve as the final ranking list of landmark retrieval. Extensive experiments are conducted on real-world social media data with both landmark photos together with their user information to show the superior performance over the existing methods, especially our recently proposed multi-query based mid-level pattern representation method.