Sensitivity based Image Filtering for Multi-Hashing

Daniel Yeung

Abstract

Hashing is an effective method to retrieve similar images from a large scale database. However, a single hash table requires searching an exponentially increasing number of hash buckets with large Hamming distance for a better recall rate which is time consuming. The union of results from multiple hash tables (multi-hashing) yields a high recall but low precision rate with exact hash code matching. Methods using image filtering to reduce dissimilar images rely on Hamming distance or hash code difference between query and candidate images. However, they treat all hash buckets to be equally important which is generally not true. Different buckets may return different number of images and yield different importance to the hashing results. We propose two descriptors to score both the hash bucket and the candidate images that it contains using a location based sensitivity measure. A neural network is trained to filter dissimilar images based on the Hamming distance and the two proposed descriptors. Both the neural network and the two descriptors could be computed offline when hash tables become available. Hence the proposed Sensitivity based Image Filtering method is efficient during retrieval. Experimental results using four large scale databases show that the proposed method improves precision at the expense of a small drop in the recall rate for both data-dependent and data-independent multi-hashing methods as well as multi-hashing combining both types. When data is distributed as a hyper sphere and hashed by Locality Sensitive Hashing, the Sensitivity based Image Filtering will perform no better than filtering using the count of hash code differences.