

Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation

Abstract:

This paper investigates a framework of search-based face annotation (SBFA) by mining weakly labeled facial images that are freely available on the World Wide Web (WWW). One challenging problem for search-based face annotation scheme is how to effectively perform annotation by exploiting the list of most similar facial images and their weak labels that are often noisy and incomplete. To tackle this problem, we propose an effective unsupervised label refinement (ULR) approach for refining the labels of web facial images using machine learning techniques. We formulate the learning problem as a convex optimization and develop effective optimization algorithms to solve the large-scale learning task efficiently. To further speed up the proposed scheme, we also propose a clustering-based approximation algorithm which can improve the scalability considerably. We have conducted an extensive set of empirical studies on a large-scale web facial image testbed, in which encouraging results showed that the proposed ULR algorithms can significantly boost the performance of the promising SBFA scheme.