

Relevance- and Interface-driven Clustering for Visual Information Retrieval

Mohamed Reda Bouadjenek¹, Scott Sanner¹, Yihao Du¹

^a*School of Information Technology, Deakin University, Australia*

^b*Department of Mechanical and Industrial Engineering, The University of Toronto, Canada*

Abstract

Search results of spatio-temporal data are often displayed on a map, but when the number of matching search results is large, it can be time-consuming to individually examine all results, even when using methods such as filtered search to narrow the content focus. This suggests the need to aggregate results via a clustering method. However, standard unsupervised clustering algorithms like K -means (i) ignore relevance scores that can help with the extraction of highly relevant clusters, and (ii) do not necessarily optimize search results for purposes of visual presentation. In this article, we address both deficiencies by framing the clustering problem for search-driven user interfaces in a novel optimization framework that (i) aims to maximize the relevance of aggregated content according to cluster-based extensions of standard information retrieval metrics and (ii) defines clusters via constraints that naturally reflect interface-driven desiderata of spatial, temporal, and keyword coherence that do not require complex ad-hoc distance metric specifications as in K -means. After comparatively benchmarking algorithmic variants of our proposed approach – RadiCAL – in offline experiments, we undertake a user study with 24 subjects to evaluate whether RadiCAL improves human performance on visual search tasks in comparison to K -means clustering and a filtered search baseline. Our results show that (a) our binary partitioning search (BPS) variant of RadiCAL is fast, near-optimal, and extracts higher-relevance clusters than K -means, and (b) clusters optimized via RadiCAL result in faster search task completion with higher accuracy while requiring a minimum workload leading to high effectiveness, efficiency, and user satisfaction among alternatives.

Keywords: Visual Information Retrieval; Relevance-driven Clustering; Visual Search User Study; Clustering via Filter Optimization.

*This work has been primarily completed while the author was at the University of Toronto.