

Image retrieval based on distance and the Pareto optimization problem

Nguyen Thu Hang

Abstract: This thesis investigates the integration of distance-based similarity measures with Pareto optimization for content-based image retrieval. Traditional retrieval systems often rely on a single aggregated similarity metric, which can obscure the complex, multi-faceted nature of visual content. To address this limitation, we propose a framework that models image retrieval as a multi-objective optimization problem. Each image is represented by a vector of features, and the retrieval task is formulated as finding images that are non-dominated with respect to multiple distance criteria, corresponding to different visual attributes such as color, texture, and shape. The Pareto-optimal set provides a set of trade-off solutions, offering users a richer and more balanced set of results than a single ranked list. We explore efficient algorithms for computing the Pareto front in high-dimensional feature spaces and evaluate the proposed method against conventional distance-based approaches. Experimental results on standard benchmark datasets demonstrate that our Pareto-based method significantly improves retrieval precision and user satisfaction by effectively managing conflicting objectives and reducing the dominance of any single feature.

Keywords: image retrieval, distance metric, Pareto optimization, multi-objective optimization