

A Fuzzy Rank-Based Late Fusion Method for Image Retrieval

Savvas A. Chatzichristofis^{1,2}, Konstantinos Zagoris¹,
Yiannis Boutalis^{1,3}, and Avi Arampatzis¹

¹ Department of Electrical and Computer Engineering
Democritus University of Thrace, Xanthi 67100, Greece

² Informatics and Telematics Institute (ITI)

Centre for Research and Technology Hellas (CE.R.T.H.), Greece

³ Department of Electrical, Electronic and Communication Engineering
Chair of Automatic Control, Friedrich-Alexander University of Erlangen-Nuremberg
Erlangen 91058, Germany
{schatzic,kzagoris,ybout,avi}@ee.duth.gr

Abstract. Rank-based fusion is indispensable in multiple search setups in lack of item retrieval scores, such as in meta-search with non-cooperative engines. We introduce a novel, simple, and efficient method for rank-based late fusion of retrieval result-lists. The approach taken is rule-based, employs a fuzzy system, and does not require training data. We evaluate on an image database by fusing results retrieved by three MPEG-7 descriptors, and find statistically significant improvements in effectiveness over other widely used rank-based fusion methods.

Keywords: Image Retrieval, Rank-Based Late Fusion, Fuzzy Systems, Heterogeneous Databases.

1 Introduction

Fusion in image retrieval is critical for the future of image retrieval research [5] and is not trivial [15]. Two main approaches to fusion have been taken: *early fusion*, where multiple image descriptors are composed to form a new one before indexing, and *late fusion*, where result rankings from individual descriptors are fused during query time. In general, late fusion approaches concern every technique for combining outputs of distinct systems [12] and can be accomplished either as a function of retrieval scores, or as a function of the position in which the results appear in each rank-list. In most cases, score-based late fusion is a better performer [2], but since in some practical situations scores are unknown, the use of rank-based fusion is necessary. A typical need for rank-based fusion arises in meta-search setups with non-cooperative search engines. Additionally, the score-based strategies, require a normalization among all systems in order to balance the importance of each of them, which is not the case of the rank-based strategies [12].

A commonly used method for rank-based fusion is Borda Count (BC), which originates from social theory in voting back in 1770. The image with the highest rank on each rank-list gets n votes, where n is the collection size. Each subsequent rank gets