

# Preface

With the widespread use of GPS-equipped handheld devices, location metadata (a.k.a. geotags) has rapidly become an integral part of photos and videos shared over the Web. This trend enabled location-based multimedia organization, search, and retrieval on many Internet services such as Google, Facebook, and Flickr. The main driving force behind these services is the creation of highly personalized user experiences, allowing for better recommendations and targeted advertisements. Even with this trend, it has been estimated that only about 5 % of the existing multimedia content on the Internet is actually geotagged. A significant amount of consumer-produced media content is still obtained using devices that do not have GPS functionality. Privacy concerns have motivated users to disable automatic geotagging of media. Furthermore, even GPS-enabled devices cannot provide accurate location information when the photo or video was captured in an indoor environment.

Nevertheless, the volume of high-quality geotagged videos and photos on the Web represents a quantity of training data for machine learning on an unprecedented scale, giving rise to the idea of creating an automated task that would try to locate non-geotagged media from the Web using models obtained through the geotagged subset. Put simply: Given a video and its associated textual metadata, can we infer the location where it was taken? This idea of “multimodal video location estimation” was proposed several years ago by the authors of the book.

Since then, the “Placing Task” of the MediaEval evaluation evaluated the task on a global scale and the United States Government has sponsored the research in their Finder program with a separate set of evaluations conducted by National Institute of Standards and Technology (NIST). As a result, the problem has been approached with diverse methods and ideas in the research community and significant improvements have been made. Multimodal Location Estimation has become a powerful tool and accuracies now come close to human capabilities.

The goal of the book is to present an overview of this field to software developers, engineers, and researchers and to bring together the different communities

working in the area. Apart from research interest, forensics experts, developers, and engineers for targeted advertising tools, as well as many people working in social media retrieval have become interested in the subject.

Berkeley, CA, USA, June 2014

Jaeyoung Choi  
Gerald Friedland



<http://www.springer.com/978-3-319-09860-9>

Multimodal Location Estimation of Videos and Images

Choi, J.; Friedland, G. (Eds.)

2015, XII, 191 p. 80 illus. in color., Hardcover

ISBN: 978-3-319-09860-9