

Preface

Digital images and videos are proliferating at an amazing speed in science, engineering and technology, media, and entertainment. With the huge accumulation of such data, the keyword search or manual annotation scheme may no longer meet the practical demand for retrieving the relevant contents in images and videos. Intelligent image search and video retrieval thus have broad applications in, for example, smart cities such as intelligent traffic monitoring systems, surveillance systems and security, and Internet of things such as face search in social media and on web portals.

This book first reviews the major feature representation and extraction methods and the effective learning and recognition approaches that have broad applications in intelligent image search and video retrieval. It then presents novel methods, such as the improved soft-assignment coding, the inheritable color space (InCS) and the generalized InCS framework, the sparse kernel manifold learner method, the efficient support vector machine (eSVM), and the SIFT features in multiple color spaces. This book finally presents clothing analysis for subject identification and retrieval, and performance evaluation of video analytics for traffic monitoring.

Specifically, this book includes the following chapters. Chapter 1 reviews the following representative feature representation and extraction methods: the spatial pyramid matching (SPM), the soft-assignment coding (SAC), the Fisher vector coding, the sparse coding and its variants, the local binary pattern (LBP), the feature LBP (FLBP), the local quaternary patterns (LQP), the feature LQP (FLQP), the scale-invariant feature transform (SIFT), and the SIFT variants. These methods have broad applications in intelligent image search and video retrieval. Chapter 2 first discusses some popular deep learning methods, such as the feedforward deep neural networks, the deep autoencoders, the convolutional neural networks, and the Deep Boltzmann Machine (DBM). It then reviews one of the popular machine learning methods, namely the support vector machine (SVM): the linear SVM, the soft-margin SVM, the nonlinear SVM, the simplified SVM, the efficient SVM (eSVM), as well as the applications of SVM to image search and video retrieval. Finally, it briefly addresses other popular kernel methods and new similarity measures. Chapter 3 first analyzes the soft-assignment coding or SAC method from

the perspective of kernel density estimation and then presents an improved SAC (ISAC) method by introducing two enhancements—the thresholding normalized visual word plausibility and the power transformation method. The ISAC method is able to enhance the image classification performance when compared to the SAC method while keeping its simplicity and efficiency. Chapter 4 proposes a novel inheritable color space or InCS and its two important properties: the decorrelation property and the robustness to illumination variations property. It further presents a new generalized InCS framework to extend the InCS from the pixel level to the feature level for improving the verification performance as well as the robustness to illumination variations. Chapter 5 introduces a new sparse kernel manifold learner (SKML) method for learning a discriminative sparse representation by considering the local manifold structure and the label information based on the marginal Fisher criterion. The objective of the SKML method is to minimize the intraclass scatter and maximize the interclass separability that are defined based on the sparse criterion. This chapter further presents various novel features, such as a new DAISY Fisher vector (D-FV) feature, a WLD-SIFT Fisher vector (WS-FV) feature, an innovative fused Fisher vector (FFV) feature, and a novel fused color Fisher vector (FCFV) feature. Chapter 6 contributes a new efficient support vector machine or eSVM for image search and video retrieval in general and accurate and efficient eye search in particular. The eSVM is able to significantly reduce the number of support vectors and improve the computational efficiency without sacrificing the generalization performance. The eSVM method is an efficient and general learning and recognition method, and hence, it can be broadly applied to various tasks in intelligent image search and video retrieval. Chapter 7 presents new descriptors—the Color SIFT Fusion (CSF), the Color Grayscale SIFT Fusion (CGSF), and the CGSF+PHOG descriptors—by integrating the oRGB-SIFT descriptor with other color SIFT features. These new descriptors are effective for image classification with special applications to image search and video retrieval. Chapter 8 derives a perceptive guide for feature subset selection and enhanced performance through the extended analyses of the soft clothing attributes and the studies on the clothing feature space via detailed analysis and empirical investigation of the capabilities of soft biometrics using clothing attributes in human identification and retrieval. This chapter also offers a methodology framework for soft clothing biometrics derivation and their performance evaluation. Chapter 9 discusses a pilot study that is designed and conducted to evaluate the accuracy of a video analytic product by integrating it with the closed-circuit television (CCTV) cameras deployed on highways. This pilot study is designed to evaluate the accuracy of video analytics in detecting traffic incidents and collecting traffic counts.

Recent Advances in Intelligent Image Search and Video
Retrieval

Liu, C. (Ed.)

2017, XVII, 235 p. 88 illus., 85 illus. in color., Hardcover

ISBN: 978-3-319-52080-3