International Journal of Multimedia Data Engineering and Management
Volume 10 • Issue 1 • January-March 2019

A No-Reference Image Quality Model for Object Detection on Embedded Cameras

Lingchao Kong, University of Cincinnati, Cincinnati, USA
Ademola Ikusan, University of Cincinnati, Cincinnati, USA
Rui Dai, University of Cincinnati, Cincinnati, USA
Jingyi Zhu, University of Cincinnati, Cincinnati, USA
Dara Ros, University of Cincinnati, Cincinnati, USA

ABSTRACT

Automatic video analysis tools are an indispensable component in imaging applications. Object detection, the first and the most important step for automatic video analysis, is implemented in many embedded cameras. The accuracy of object detection relies on the quality of images that are processed. This paper proposes a new image quality model for predicting the performance of object detection on embedded cameras. A video data set is constructed that considers different factors for quality degradation in the imaging process, such as reduced resolution, noise, and blur. The performances of commonly used low-complexity object detection algorithms are obtained for the data set. A no-reference regression model based on a bagging ensemble of regression trees is built to predict the accuracy of object detection using observable features in an image. Experimental results show that the proposed model provides more accurate predictions of image quality for object detection than commonly known image quality measures.

KEYWORDS

Automatic Video Analysis, Image Distortion, Image Quality Assessment, No-Reference, Object Detection

DOI: 10.4018/IJMDEM.2019010102

Copyright © 2019, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

INTRODUCTION[1]

Wireless embedded camera sensors have become ubiquitous components in various imaging applications, such as public safety and security systems, smart building operations, intelligent transportation, and remote health care. Rather than merely presenting raw data collected by camera sensors to the user, an application usually aims to automatically discover and extract meaningful information from the camera sensors and to achieve as much autonomy as possible in the physical system. Automatic video data analysis tools, which could detect, recognize, track objects of interest, and understand their behaviors, have become indispensable components in today's imaging applications.

The performance of automatic analysis methods relies on the quality of images that are processed. It is therefore essential to introduce objective metrics for predicting the quality of images evaluated by automatic analysis algorithms. In the field of image quality assessment (IQA), a diverse range of image quality models, ranging from full-reference to reduced-reference and no-reference ones, were designed for predicting the perceptual quality evaluated by human subjects (Ma et al., 2018, pp. 1202-1213; Wang et al., 2018, pp. 1-14; Wang, Bovik, Sheikh, & Simoncelli, 2004, pp. 600-612).

The quality of a video sequence judged by an automatic analysis algorithm, however, is not necessarily sensitive to the same factors that drive human perceptions. The perceptual image quality assessments usually try to emulate known characteristics of the human visual system (HVS), such as the contrast sensitivity and the visual attention mechanisms. The contrast sensitivity mechanism means that the HVS is sensitive to the relative luminance change rather than the absolute luminance change (Wang et al., 2004, pp. 600-612). The visual attention mechanism is that only a local area in the image can be perceived with high resolution by the human observer at one time instance at typical viewing distances, due to the foveation feature of the HVS (Yang et al., 2016, pp. 3475-3488). On the other hand, automatic analysis methods run by machines can "perceive" the absolute luminance change precisely and have a better global "view". For example, the problem of evaluating motion imagery quality for tracking in airborne reconnaissance systems was studied in Irvine and Wood's research (2013, p. 87130Z). It was found that automated target detection algorithms are less sensitive to spatial resolution than humans, but factors such as jitter in the temporal domain, texture complexity, edge sharpness, and level of noise have a strong effect on the performance of target detection. In our recent work (Kong, Dai, & Zhang, 2016, pp. 3797-3801), we found that unlike human beings who can easily extract and focus on a moving object from a blurred background, the performance of object detection algorithms can be affected by the quality of the background. These results suggest that new models are needed for evaluating the quality of images from the perspective of automatic analysis algorithms.

In a wireless imaging system, automatic analysis could be deployed using two strategies: in the central server on compressed videos; or at the local cameras on uncompressed videos as a preprocessing step. The impact of video compression on the accuracy of analysis algorithms has been studied in some recent works (Tahboub, Reibman, & Delp, 2017, pp. 4192-4196; Zhong, & Reibman, 2018, pp. 1-6), which aim at finding the optimal compression rates under a quality requirement. Apart from the distortion introduced by compression, the quality of an image or a video could be degraded during the data acquisition or sensing process, e.g., distortion caused by noise or motion blur, or reduced image resolution due to storage or bandwidth constraints on embedded cameras. These factors should also be taken into consideration to evaluate the quality of an image.

Object detection is the first and the most important step in the process of automatic analysis, because the detected objects provide a focus of attention for the following tasks such as tracking and recognition. In this paper, we propose a blind regression model based on a bagging ensemble of trees to predict the performance of object detection on an image. The model utilizes local features in an image such as edge and oriented gradient and global features including image gradient and estimated object size, which could be easily extracted from an image. The model is trained using a large number of images with different scene characteristics and four types of distortions including

16 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/a-no-reference-image-quality-model-for-object-detection-on-embedded-cameras/232180

Related Content

Parallel and Distributed Multimedia Databases
S. Geislerand O. Kao (2008). *Multimedia Technologies: Concepts, Methodologies, Tools, and Applications (pp. 223-232).*
www.irma-international.org/chapter/parallel-distributed-multimedia-databases/27084

Ontology Instance Matching based MPEG-7 Resource Integration
Hanif Seddiquiand Masaki Aono (2010). *International Journal of Multimedia Data Engineering and Management (pp. 18-33).*
www.irma-international.org/article/ontology-instance-matching-based-mpeg/43746

Low Light Face Detection System
G. Ananthi, T. Balaji, M. Pugalenthiand M. Rajkumar (2023). *Recent Advancements in Multimedia Data Processing and Security: Issues, Challenges, and Techniques (pp. 166-181).*
www.irma-international.org/chapter/low-light-face-detection-system/331441

A Novel Research in Low Altitude Acoustic Target Recognition Based on HMM
Hui Liu, Wei Wangand Chuang Wen Wang (2021). *International Journal of Multimedia Data Engineering and Management (pp. 19-30).*
www.irma-international.org/article/a-novel-research-in-low-altitude-acoustic-target-recognition-based-on-hmm/276398

Communication and Media: Types, Functions, and Key Concepts
Bradley Freeman (2018). *Handbook of Research on Media Literacy in Higher Education Environments (pp. 25-40).*
www.irma-international.org/chapter/communication-and-media/203989