Objective Evaluation of Video Segmentation Quality

Paulo Lobato Correia, Technical University of Lisbon, Portugal
Fernando Pereira, Technical University of Lisbon, Portugal

ABSTRACT

The evaluation of image and video segmentation results assumes a critical role for the selection of appropriate segmentation algorithms, as well as the adjustment of their parameters for optimal segmentation performance in the context of a given application. The current practice for the evaluation of video segmentation quality is based on subjective testing, which is an expensive and time-consuming process. Objective segmentation quality evaluation techniques can alternatively be used, once appropriate algorithms become available. Currently this is a field under development and this contribution proposes evaluation methodologies and objective segmentation quality metrics both for individual objects and for complete segmentation partitions. Standalone and relative evaluation metrics are proposed, to be used when a reference segmentation is missing, or available for comparison, respectively.

INTRODUCTION

The major objective of image and video segmentation algorithms is to produce appropriate results for the particular goals of the application addressed. Therefore, segmentation algorithms’ performance assessment assumes a crucial importance in evaluating the degree to which application targets are met.
The current practice for segmentation performance evaluation consists in subjective assessment by a representative group of human viewers (COST211, 2001). To be meaningful, such evaluations must follow precise methodologies, both in terms of test environment setup and grading techniques. Standard subjective evaluation methodologies have been established for video quality evaluation (ITU-R BT.500, 1995; ITU-T P.910, 1996; ITU-T P.930, 1996), targeting the evaluation of video degradation due to some types of processing, such as video coding, but not for segmentation quality evaluation.

Subjective evaluation of segmentation quality is usually considered as a “good enough” procedure, and often it is the only solution available. However, this type of evaluation has several drawbacks. Setting up the evaluation environment requires a considerable effort, besides requiring the presence of a significant number of evaluators to achieve statistically relevant results, becoming a time-consuming and expensive process. Also, the subjective nature of the evaluation may prevent the reproducibility of results, especially if precise and stable evaluation conditions and methodologies are not respected.

The alternative is to devise segmentation evaluation strategies that do not involve human evaluators. But, even if the development of segmentation algorithms is the topic of a large number of publications, the issue of their performance evaluation has not received comparable attention (Zhang & Gerbrands, 1994; Zhang, 1996; Rees & Grenway, 1999). This may be due to the difficulty in establishing a measure capable of adequately evaluating segmentation quality, except for very well constrained application scenarios.

Several non-subjective quality evaluation methodologies have been proposed since the 1970s, initially targeting the performance assessment of edge detectors. More recently, with the emergence of the MPEG-4 standard (ISO MPEG-4, 1999) and its ability to independently code arbitrarily shaped video objects, a new impetus to research on segmentation, as well as the development of segmentation quality evaluation methodologies, has been apparent (Erdem, Tekalp, & Sankur, 2001; Correia & Pereira, 2002; Mech & Marques, 2002; Correia & Pereira, 2003; Villegas & Marichal, 2004).

Some initial attempts tried to define analytical evaluation methods, assessing segmentation performance directly by examining the algorithms, but not requiring their implementation. The principles and properties of the algorithms’ components (e.g., complexity, efficiency, resolution of the results and processing approach) and their combination strategies were analyzed. These methods’ applicability to complex algorithms is limited, while requiring considerable human evaluation effort. Qualitative evaluation reports, focusing on the strengths and weaknesses of algorithms, or quantitative measures for particular techniques (Abdou & Pratt, 1979), could be produced as output.

Nowadays, several objective evaluation methods have been proposed to assess segmentation quality by using automatic measurement tools. Typically, objective evaluation operates on segmentation results produced by the algorithm being tested. As for subjective evaluation, two classes of techniques exist (Zhang, 1996):

- **Standalone objective evaluation** or “goodness” evaluation: Consists in evaluating a segmentation partition, understood as the set of objects that completely covers an image at a given instant without overlapping, on its own.
27 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:

[Link to full version]

Related Content

Detecting Corners for 3D Objects
[Link to article]

Enhancing Robustness in Speech Recognition using Visual Information
[Link to chapter]

A Technique to Approximate Digital Planar Curve with Polygon
[Link to chapter]

Shape Features of Overlapping Boundary for Classification of Moving Vehicles
[Link to article]

Copy Detection Using Graphical Model: HMM for Frame Fusion
Shikui Wei, Yao Zhao and Zhenfeng Zhu (2013). *Graph-Based Methods in Computer Vision: Developments and Applications* (pp. 261-279).
[Link to chapter]