Multimedia content classification metrics for content adaptation

Rui Fernandes[1], M. T. Andrade[2]
[1]Faculty of Engineering, University of Porto, Portugal (mpt09015@fe.up.pt); ESTIG-IPB, Bragança, Portugal; [2]INESC TEC, Faculty of Engineering, University of Porto, Portugal (mandrade@fe.up.pt)

Abstract
Multimedia content consumption is very popular nowadays. However, not every content can be consumed in its original format: the combination of content, transport and access networks, consumption device and usage environment characteristics may all pose restrictions to that purpose. One way to provide the best possible quality to the user is to adapt the content according to these restrictions as well as user preferences. This adaptation stage can be best executed if knowledge about the content is known *à-priori*. In order to provide this knowledge we classify the content based on metrics to define its temporal and spatial complexity. The temporal complexity classification is based on the Motion Vectors of the predictive frames of the content, and the spatial complexity classification is based on an edge detection algorithm and an image activity measure.

Author Keywords. Multimedia classification, temporal complexity, spatial complexity, multimedia adaptation.

1. Introduction

Today, there is a wide array of possibilities for consuming Multimedia content, from TV displays at home to portable devices on the go, with different types of transport and access networks. Furthermore, there is already a great diversity of high definition (HD) content available and these consumption scenarios don't always meet the minimum requirements to enable an optimal HD content consumption. One way to overcome this limitation is to adapt the multimedia content, and content classification can provide insightful information to execute this task.

We present the implementation of different metrics to classify multimedia content, based on temporal and spatial complexities. This kind of classification can be very useful to decide the type of adaptation that should be performed on the content, when there are restrictions imposed by the distribution chain and the consumption context.

2. Temporal Complexity Metrics

One way to classify the temporal complexity of a sequence is to look into the magnitude of the motion vectors (MVs) present on the frames of the sequence. In (Amel and Abdessalem and Abdellatif, 2010), the authors present an algorithm to perform this type of classification per frame, based on the intensity of motion, using a five level classification scale.

We implemented this metric to establish the original (1) frame-by-frame classification, (2) GOP-by-GOP or group of GOPs-by-group of GOPs classification, (3) scene-by-scene classification and also (4) multimedia clip classification. Both (2) and (3) approaches are defined based on the mean average value of the

intensity of motion, of the GOP/scene frames, and are to be used together with the definition of points in time where an adaptation should take place. For the implementation of (4), different approaches were used, namely: (4.1) the average intensity of motion, of all frames under consideration, was determined and used along with the original classification scale to define the clip's classification; (4.2) the statistical mode of the frames classifications was determined and attributed as the clip classification. (4.3) A set of complementary characteristics (mean average and standard deviation of the intensity of motion, considering all frames from the clip) was determined to produce more knowledge about the multimedia clip in terms of motion complexity degree and its variation throughout the sequence.

3. Spatial Complexity Metrics

Spatial complexity is usually established through the use of edge detection algorithms over the luminance component of the sequence. A variation of the Sobel edge detection algorithm was selected and computed according to the definition established in (*ITU-T Study Group 12, 1999*). Consequently, one of the methodologies defined in (Yu and Winkler, 2013) was employed to obtain a measure of the spatial complexity of a frame. This method was chosen because it provides good results at low complexity for its implementation.

To assess the quality of the obtained results, a second metric was used, based on both local vertical and horizontal gradients calculated over the luminance component of the frames. Our results confirmed that this metric, presented in (*Engelke et al, 2009*), quantifies image activity very accurately, delivering higher values as the spatial complexity increases.

4. Conclusions and Results

The implemented metrics proved to be adequate to classify multimedia content by measuring with a good level of accuracy its temporal and spatial complexity on a frame, GOP, scene or complete clip basis. This classification is crucial for an adaptation decision engine to be able to select the best adaptation among several possible ones aiming at offering the best possible Quality of Experience to the user.

By comparing the automatic classification of the content using the temporal complexity metrics with a visual inspection of the content, it is possible to conclude that it delivers reasonably accurate results. Still, the strategies devised for groups of frames classification require a more thorough analysis to establish which method provides more reliable results. Furthermore, an analysis using derivatives is being conducted to assess whether it is possible to perform the classification without the need to process all frames. The results for the spatial complexity metrics yields similar classifications with an almost linear relation between the metrics' values in which the classification is based.

References

Amel, Abdelati Malek and Abdessalem, Ben Abdelali and Abdellatif, Mtibaa. 2010. "Video shot boundary detection using motion activity descriptor". Journal of Telecommunications, 2 (1):54–59.

ITU-T Study Group 12. 1999. "Subjective video quality assessment methods for multimedia applications". ITU-T Recommendation P.910, International Telecommunication Union.

Yu, Honghai and Winkler, Stefan. 2013. "Image complexity and spatial information". In Quality of Multimedia Experience (QoMEX), 2013 Fifth International Workshop on, 12-17.

Engelke, Ulrich and Kusuma, Maulana and Zepernick, Hans-Jürgen and Caldera, Manora. 2009. "Reduced-reference metric design for objective perceptual quality assessment in wireless imaging". In Image Commun., 24 (7):525-547.