Adaptive QoE measurement on videostreaming IP services

Jose A. Lozano, Alfonso Castro, Beatriz Fuentes
Telefónica Investigación y Desarrollo, TID
Madrid, Spain
{jal, acast, fuentes}@tid.es

Juan Manuel González
Telefónica Global Technology
Madrid, Spain
juanmanuel.gonzalezmunoz@telefonica.ess

Álvaro Rodríguez
SoftTelecom
Madrid, Spain
arp@softtelecom.es

Abstract—**Nowadays broadband technologies have been introduced in millions of homes worldwide. They allow Services Providers to offer services with high bandwidth requirements, as IPTV. Essential for these services is the Quality of Experience (QoE) as it is perceived by the user. This paper presents a low cost device and a procedure to measure on-line the quality of a video streaming according to the end user perception. Furthermore, it describes a content aware and self-adjustable method to analyze the characteristics of specific video streaming flows.**

Keywords-component; IPTV services; video streaming monitoring; quality of experience (QoE)

I. INTRODUCTION

Video streaming is a key piece for an ever increasing number of telecommunication services as video conferencing, video on demand, live TV over Internet, etc. Although, some of these services are delivered using "best effort" quality scheme, being able to measure the video flows quality is an important issue for companies providing telecommunications managed services.

Although there are some methods and tools for estimating the quality of a video stream, most of them are based on the measurement of network parameters, which can only provide an estimation of the end users perception. Moreover, methods for estimating the quality of the video signals require expensive off-line resources with high processing capabilities. So, there is a lack of tools to effectively know the quality of the image that IPTV services end users are really receiving.

This paper presents a low cost device (Video Intelligent Probe - VIP) and a procedure to measure on-line the quality of a video streaming according to the end user perception, based on the quality of the images contained in the video streaming.

Furthermore, the solution allows this perceptual measuring of video flows to be correlated with information from technical network parameters measured extracted at the same point in time. In addition, it implements an interface for an operator to control the measurement process using high level commands. Another innovation is to make the algorithms context aware and self-adjustable to the characteristics and contents of specific video streaming flows. The solution presented makes the perceptual quality of video sensitive to both the viewer context and the operator interest.

The remainder of this paper is organized as follows: section II discuses existing solutions, section III presents in depth the proposed solution, section IV describes a proof of concept that validates the solution and, finally, section V provides some conclusions and proposes further work in this area.

II. EXISTING SOLUTIONS

The proposed solution about Internet Protocol television (IPTV) is related with different technologies: Videostreaming and codecs, network transmission parameters and video flow measurement.

Today's video codecs use a Group of Pictures (GoP) frame structure, which consists of an independently encoded reference frame (an "I" frame) followed by a sequence of "B" and "P" frames, in which only the motion changes from previous frame are encoded. When packet loss occurs, it can lead to decoding errors in one or more of these frame types on the receiving end of the video stream. An error occurring in an "I" or a "P" frame will propagate through all the remaining "B" and "P" frames, and thus be more likely to cause a visible impairment that may last up to several seconds. An error in a "B" frame, however, does not propagate to subsequent frames and may not even be noticeable to the viewer.

The quality perceived for users of IPTV services depends on the quality of the image they are receiving. Current QoS/QoE methods [1]-[3] measure the network parameters directly from the NE (Network Elements) or using some probes located at different points on the network. These probes can gather working parameters from the service protocol stack (IP, TCP, UDP, HTTP, etc.) as packet delay, packet loss, packet jitter, etc.

Regarding the analysis of the video flow itself there are two main references on such topic: the ITU-T recommendations, J.144 [4] and BT.1683 [5]. Recommendation ITU-T J.144 provides guidelines on the selection of appropriate objective perceptual video quality measurement equipment designed for use in digital cable television applications when the full reference video signal is available. In BT.1683, the reference and degraded video signal are compared to predict the video quality.

III. PROPOSED SOLUTION: VIP

The proposed solution, Video Intelligent Probe (VIP), integrates the analysis of video processing and network parameters; in order to adjust the thresholds of the video parameters to determine the customer perception.

Figure 1 presents the deployment its architecture, in the home environment. The VIP receives an IP video flow and by pass it to the Out interface, this makes this solution suitable for being deployed between the set top box and the customer router.

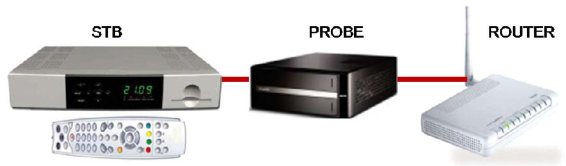

Figure 1. Adressing solution

Basically the solution presented in this section has three main components see Figure 2:

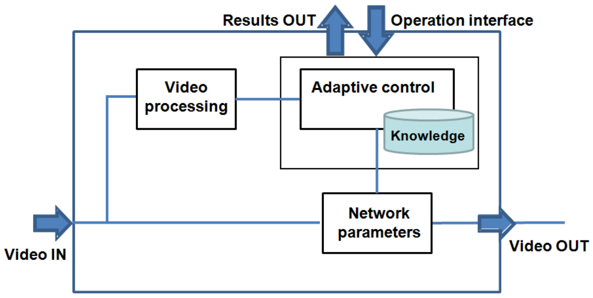

Figure 2. VIP architecture

- A zero-reference algorithm to analyze the video flow. The video processing block is in charge of analysing and detecting video artefacts that a customer can perceive. It is configured depending on a number of parameters such as the ratio of pixels with errors in a frame or the number of frames presenting artefacts in a certain amount of time.
- A network based measurement component enable the combination of direct quality measures by the video algorithm with perceived measures. This way the solution can anticipate problems in the video flow.
- The adaptive control component is in charge of giving the probe an autonomic behaviour. It contains a knowledge model which empowers it to decide which video quality profile will be applied to configure the video processing component. Furthermore it also correlates network information with the output of the video processing block.

A. Video processing

The video processing is implemented as a four stage algorithm: reception, conversion, analysis and result.

The reception stage extracts the frames from the video stream. Each frame is an input of the conversion stage where the frames are converted to YUV420p format. The YUV format is a colour space where Y stands for the luminance component and U and V are the chrominance components. To improve efficiency, the analysis stage only works with black and white frames so it only needs the Y component from the converted frame. The YUV 420p format allows obtaining a black and white frame easily by taking the Y component from the converted frame. That is the main reason to choose the YUV 420p format. Also in most of the IPTV broadcast systems, the YUV 420p is the emission format so in many cases the conversion stage can be omitted.

Previously to the development of the solution, an analysis of the customer claims during the period of one month was made. Following the total loss of service, the most important reported symptoms were freezing and pixilation, while the complaints related other parameters, like chroma, were not relevance. So, the analysis stage searches for two specific errors on the images of the video stream, frozen image and pixelation.

Figure 3 describes how these artifacts are characterized by a number of parameters in order to objectively express a perceived artifact.

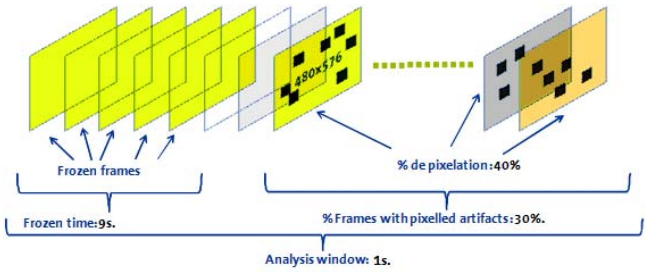

Figure 3. Stages of the algorithm

The video streaming analysis is based on the following parameters:

- Frozen time. It's the maximum time for one frame to remain unmodified before generating an event.
- Frame pixel ratio "pixelation": The number of pixels that show an error.
- Ratio of "frames" that show pixelled artifacts. It's the ratio between the number of frames shown pixelled artifacts above the stated pixilation value.
- Analysis window. Specification of the analysis duration.

Changing the threshold values of these parameters allows the Service Providers to fix different levels for perceived artifacts avoiding false events and improving QoE monitoring overall process.

For the frozen image analysis each frame is compared to the previous one obtaining the difference of movement between them. If there is no difference between two consecutive frames the image of the video stream is frozen.

The artefacts that compose a pixelation error have particular characteristics, they have a quadrangular shape and they have a similar texture. Taking into account these two characteristics it is possible to locate the pixelation errors in an image with the two analysis phases. The first one is an edge filtering and the other one is a Dirac delta analysis. The result of this stage is the image percentage that presents pixelation artefacts.

The edge filter is a Canny edge detector ([6] OpenCV Computer Vision with th OpenCV library Chapter 6). It detects the pixels of the image which are candidate for being an edge by using thresholding with hysteresis. It uses two thresholds, high and low. The pixels which have a higher gradient than the high threshold are marked as edge, the ones which have a gradient between the high and the low threshold are marked as possible candidates for being edge pixels and the ones which have a lower gradient than the low threshold are discarded to be edge pixels.

The Dirac Delta analysis detects the parts of the frame which have the same or a very similar Dirac Delta value. The Dirac Delta value of a specific zone of the frame represents it´s texture. The frame is divided into square components and for each component it´s Delta Dirac values is calculated. By comparing these values the algorithm can discover the image zones which have similar texture.

The result stage presents the percentage of movement and pixelation of the frame.

Finally, in Figure 4 it is depicted the whole process that is followed for pixelation and freeze image detection.

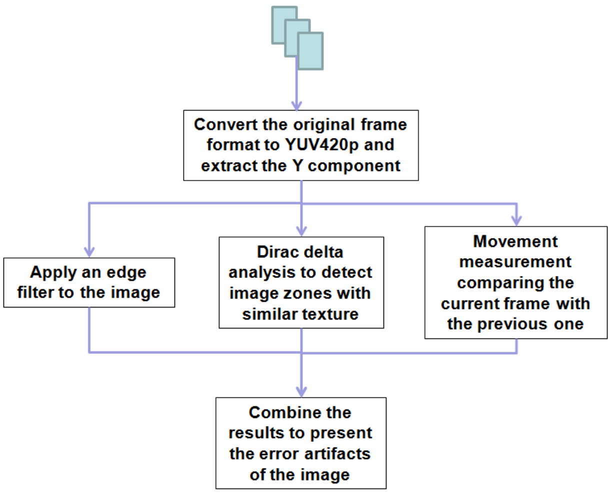

Figure 4. Algorithm

B. *Network measurements*

Network measurements are obtained by inspecting the network packets that compose the IPTV flows received by the user. The protocol used to transmit the IPTV flows is Transport Stream over RTP .

The probes provide three network measurements:

- Inter Arrival Time: time elapsed between the packet´s reception. This value is calculated comparing the packet´s reception timestamp.
- Packet Loss: number of packets lost. This value is calculated analyzing the sequence number of the RTP header.
- Transport Stream Discontinuities: number of discontinuities in the transport stream packets. Each packet of the Transport Stream protocol contained in the RTP payload has a sequence number; inspecting this sequence number the discontinuities of the Transport Stream can be arisen.

C. *Adaptation process*

The adaptation process is based in reasoning techniques that enable the component to match high level views of business and services from the operator into low level network metrics and video quality profiles.

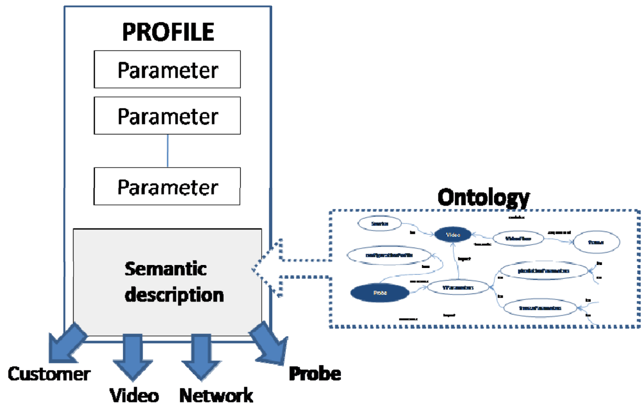

Figure 5. Adaptation process

The process is driven by the definition of a semantic profile, see Figure 5. A profile is the set of parameters that can be measured and characterize a specific domain. It also comprises a semantic description of what it means within the domain and how the parameters are related between different domains. The solution described in this paper uses an ontology to describe parameters and their relation according to four realms: Probe, Customer, Video and Network.

The relationships between the different realms are captured by means of an ontology language, namely OWL [7]. The use of an ontology language to define the management information offers advantages such the ability to use reasoning tools working on ontologies (e.g., inference engines used in artificial intelligence).

The second step is the implementation of the rules governing the system behavior, using SWRL [8](Semantic Web Rule Language) that extends the abstract syntax of OWL rules to include conditional rules into the ontology language.

The semantic model can be changed at any time and distributed to the probes without coding them again. It can be updated, extended or even shortened and then distributed again to the probes for them to work with the new domain description. This means that the presented solution allows the operator to introduce new concepts for the probe to manage them, and this can be done with minimum development effort.

IV. PROOF OF CONCEPT

A proof of concept consisting of the video analysis and network measurement phases detailed in the previous section has been implemented and tested.

As shown in Figure 6, the set of probes at the different customer premises conforms a distributed monitoring system. Each probe is measuring the quality of the image for a given customer. In order to allow the supervision of a number of customers, a central management system has been added, for storage of historical data and visualization purposes.

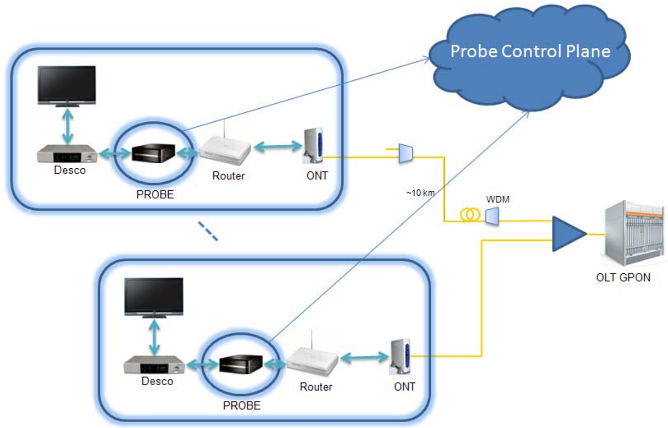

Figure 6. Adressing solution

The implementation of the video analysis algorithm in customer devices like STB or router was considered. Nevertheless, the characteristics of the existing home devices, with limited processing capabilities, are not enough to efficiently execute a video analysis algorithm at runtime. Currently, the hardware platform chosen for the probes is a barebon LEX LEO with an Intel Atom D510 1GB RAM 4GB Flash HD.

Each probe sends each minute to the control plane, via web services, a report with the results obtained by the analysis. This report contains: maximum IAT, packets lost, discontinuities in the transport-stream, number of pixelation errors and number of frozen image errors

The central management system implements a supervision mode. It receives the probe´s reports and saves the information in a database. If the data received belongs to a probe previously registered on the system it can be displayed graphically on a web browser.

Finally, it is worth mentioning that the management system allows the modification of the thresholds that determine the sensitivity of the probes in the analysis of the video images: the threshold for the detection of pixelation and the threshold to decide if the pixelation problem affects the QoE. If the thresholds are modified by a human operator, they are propagated to the registered probes via web services. The probes will take the new values in runtime, and further use them in the analysis of the images.

V. CONCLUSIONS AND FUTURE WORK

There are a number of advantages that this solution brings to the current state of art, namely:

- Real perceived quality versus network parameters-based estimation.
- Sensitivity can be adaptively adjusted in real time to match users' perception by means of specifying the type of contents, kind of users, etc.
- Correlation between perceived quality and network parameters for operational purposes.
- A semantic information model to provide a knowledge oriented interface for policy driven control and management.

As future work, we plan to enhance the current proof of concept with the development of the semantic-based self-adaptation process. Furthermore, a data mining process of the network measurements and the result of the video analysis algorithm will be researched, in order to find the root cause of the problem when a failure occurs.

REFERENCES

[1] Shu Tao, Apostolopoulos, J., Guerin, R.. "Real-Time Monitoring of Video Quality in IP Networks". IEEE/ACM Transactions on Networking. Oct 2008

[2] Amy R. Reibman, Subhabrata Sen, and Jacobus Van der Merwe."Network monitoring for Video Quality over IP". Dec 2007. AT&T Whitepaper.

[3] Jaime Lloret, Miguel García, Marcelo Atenas, Alejandro Cánovas. "A QoE management system to improve IPTV network". Int. J. Commun. Syst. 2011.

[4] Recommendation ITU-T J.144 "Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference", 2004.

[5] Recommendation ITU-R BT.1683 "Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference", 200410

[6] Gary Bradski Adrian Kaehler "Learning OpenCV: Computer Vision with the OpenCV Library", O'Reilly Media 6

[7] Ontology Web Language Reference, http://www.w3.org/TR/owl-features/

[8] SWRL: A Semantic Web Rule Language Combining OWL and RuleML. http://www.w3.org/Submission/SWRL/