Full reference testing helps ensure quality audio delivery

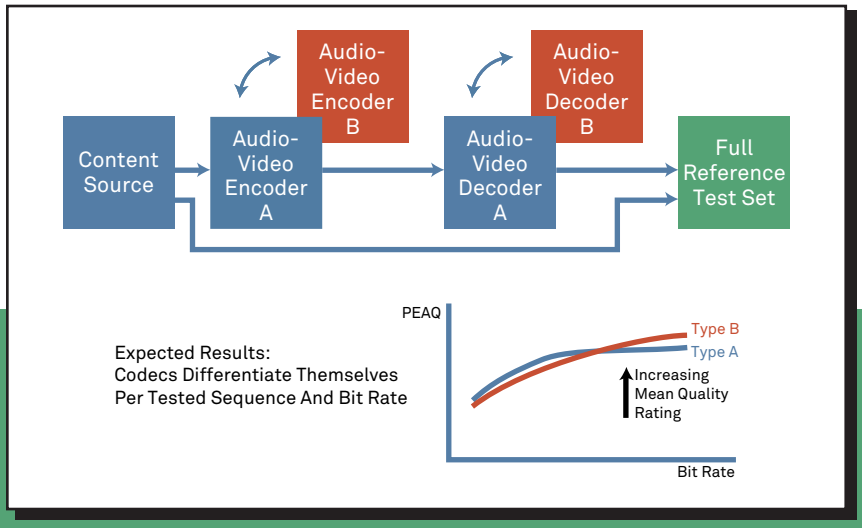

Figure 1: Target Bit-Rate Designation

by Adam Schadle, Vice President of Video Clarity

With the ever-increasing amount and variety of programming, devices, screen sizes, quality levels, and delivery methods today, broadcasters and delivery networks must be able to measure the quality of audiovisual content — and their processes for delivering it — to ensure it will meet standards that satisfy audiences, subscribers, advertisers, and regulators. One of the most accurate methods for assessing changes in quality is full-reference testing, which compares original content to processed content. Full-reference testing produces very accurate and highly repeatable results at a much lower cost than subjective testing with human listeners. Astute content distributors recognize that fact and are using full-reference testing to help them determine the best combination of equipment and delivery schemes to use in their operations.

In terms of selecting an appropriate delivery bit rate, the workflow is fairly straightforward (Figure 1): An A/V source player feeds an uncompressed signal to the encoder's input. The compressed output of the encoder can be fed directly to the decoder's input and captured in real time to form one element in a set of full-reference test sequences. Additional outputs are then captured and added to the sequence set by making small-step changes in the encoder bit rate. You can create more full-reference test sequence sets by using different source content and repeating the process of capturing and storing the decoder output sequences. (This workflow can also be performed as a file process in place of real-time play/record functions.)

Once the full-reference comparisons are made for various bit rates, differences in audio quality caused by each bit-rate change can be measured in each sequence set using objective perceptual techniques such as PEAQ (Perceptual Evaluation of Audio Quality). The PEAQ model, an international standard for audio quality measurement, was calibrated to deliver scores that correlate with the mean opinion scores of a panel of expert listeners who took part in a formal subjective test. In this sense, PEAQ can be viewed as an "average expert listener" score. By plotting the scores for each set of sequences on a graph, it's easy to see the quality trend for each bit rate under comparison.

PEAQ is optimized to measure the perceptual quality of wideband (20 kHz) audio codecs and has been successfully used to measure the quality of other wideband audio devices such as analog-to-digital converters, digital-to-analog converters, and sample-rate converters.

The mean quality rating is either scored as a 0 to -4 or 1 to 5. The score is an objective differential grade (ODG), which represents the overall severity of the impairments in the test signal. Its meaning is defined on the rating scales of Figure 2.

This objective, full-reference test measurement will provide a quality value for the corresponding audio in each frame of video. These frame-by-frame quality values can be compared directly to the quality values from other elements

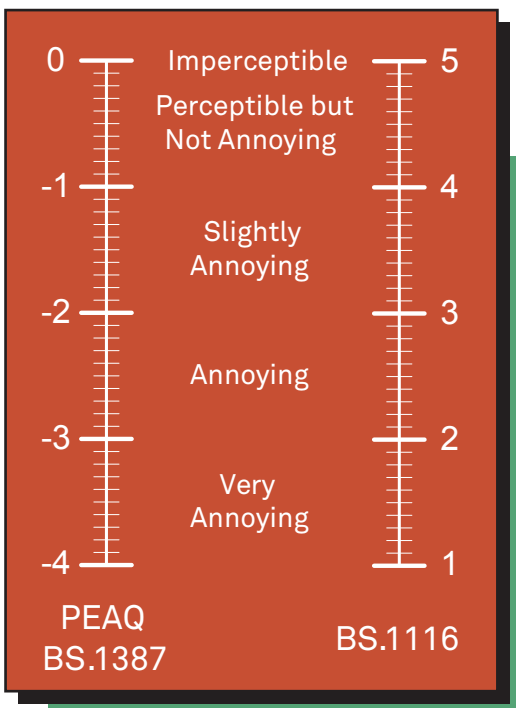

Figure 2: PEAQ Impairment Scales

of the same test sequence set (since they all used the same source sequence), or they can be compared to other sets that have been generated using the same source sequence at another bit rate. If objective measurements vary greatly from one bit rate to another in portions of the output sequences, then you can examine them more extensively using subjective techniques.

The full-reference test and measurement process for evaluating codec technologies is very similar to the one for choosing a bit rate, except that you change the codec engine for each test instead of the bit rate (Figure 3). After evaluating the first codec (A), simply insert a second codec (B) into the test workflow and start again. From there PEAQ measurements can be made to identify trends in the average sequence scores as a result of the different codecs.

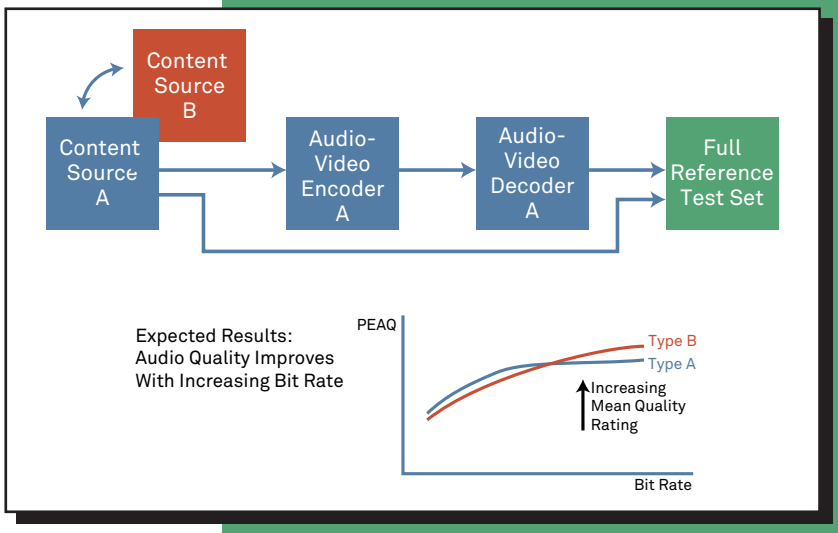

Figure 3: Codec Technology Selection

By using full-reference testing and the PEAQ audio quality scoring model, content originators can get a clear idea of how close a processed audio signal is to the original, and how it is affected by changes in different delivery variables. The results of these tests help them to ensure a highly accurate measurement of quality for audio delivery.