grass valley
A BELDEN BRAND

Streamlining HD Monitoring with Media Fingerprinting

Marco Lopez, Senior Vice-President – Infrastructure & Monitoring

WHITE PAPER

Lowering the operating costs of television playout has become something of a mantra across the television industry, and it has been especially evident in playout monitoring. Since monitoring has traditionally been labor-intensive and also prone to human error, it has been a natural focus for many streamlining and automation projects.

The widespread adoption of multiviewers over recent years, as well as IP-based facility monitoring with signal probing, has opened up new, more efficient workflows. A single operator can now manage and monitor complex systems efficiently, with a consequent large improvement in productivity.

Fingerprint Generation

Naturally, the value and strength of a media fingerprint system are dependent on the underlying creation algorithm. Currently, there are multiple proprietary technologies used for the generation of fingerprints, and these are usually level and/or motion based. Some use luminance characteristics, while others use transitions, edges, peaks, frequency and color characteristics.

A typical media fingerprint generator is shown in Figure 1, and this creates separate fingerprint data for the video and the individual audio channels to allow effective multi channel monitoring. This data is multiplexed, and can be streamed for storage or live comparison. This fingerprint generation requires minimal hardware, and can be implemented on a simple monitoring-grade DA module or more complex interface cards. The fingerprint data stream generated is very small, and represents just 0.0004 percent of a 1080i60 HD signal. This means that the fingerprint data from many HD channels can be transferred quite easily over standard IP networks.

This type of media fingerprint is highly resilient to normal television playout processes, with insensitivity to typical video level adjustments, as well as up/downconversion and video compression. Similarly, the process is insensitive to audio bit rate reduction (compression), audio loudness control performed by gain or dynamic range adjustment, and sample rate conversion.

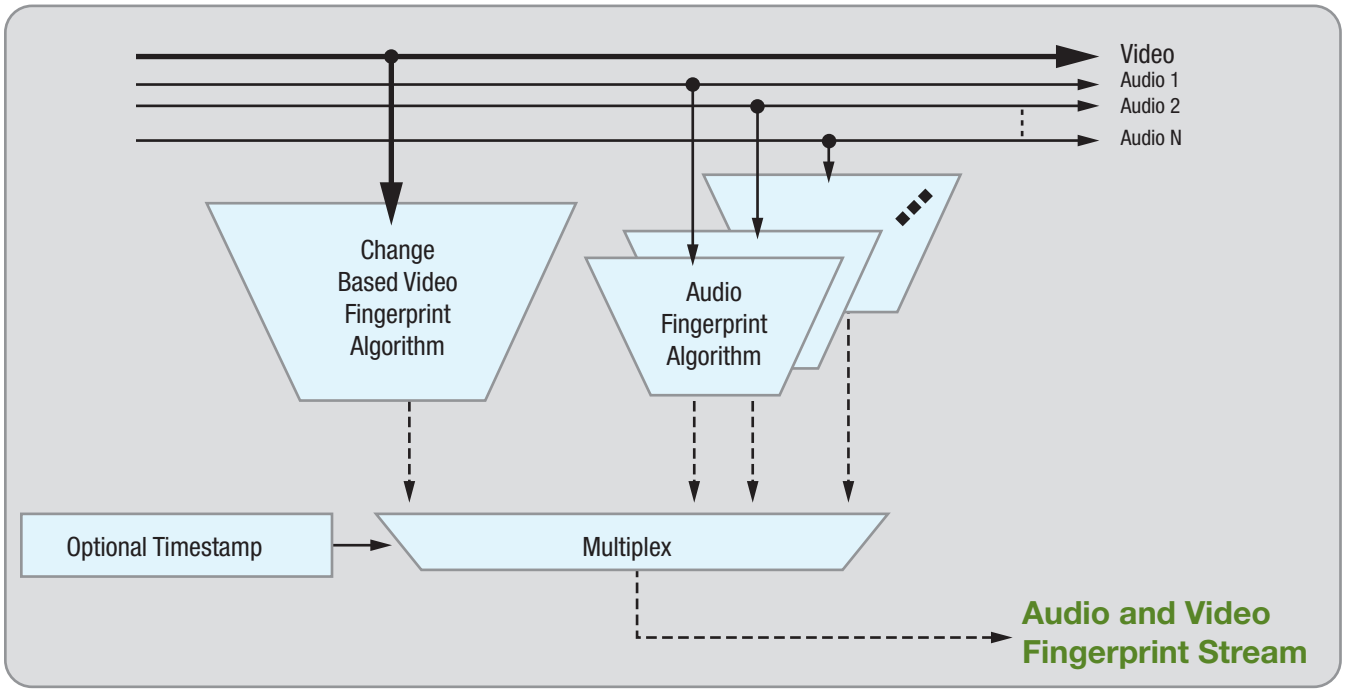


Figure 1: Typical media fingerprint generator functionality with separate video and audio algorithms

Media fingerprints are weakened, but not disabled, by aspect ratio changes, and by the insertion of small graphics. The video processes that can cause problems are standards conversion like going from 50 Hz to 60 Hz, and prolonged periods of frozen video. On the audio side, an obvious problem is when new audio content is mixed into the original content, such as voice-overs and stings.

During fingerprint monitoring, a simple convolution engine (see Figure 2) is used to look for matching patterns between two media fingerprint streams at different points in the playout process. At the simplest level, this comparison process can be performed by a single interfacing module. However, the process is entirely scalable, and multiple streams of fingerprint data can be correlated and analyzed using a standard PC based platform to allow end-to-end deployment in a large facility (see Figure 3). It can even be used across multiple remote facilities, as the fingerprint data is so small that transfers over a WAN are possible.

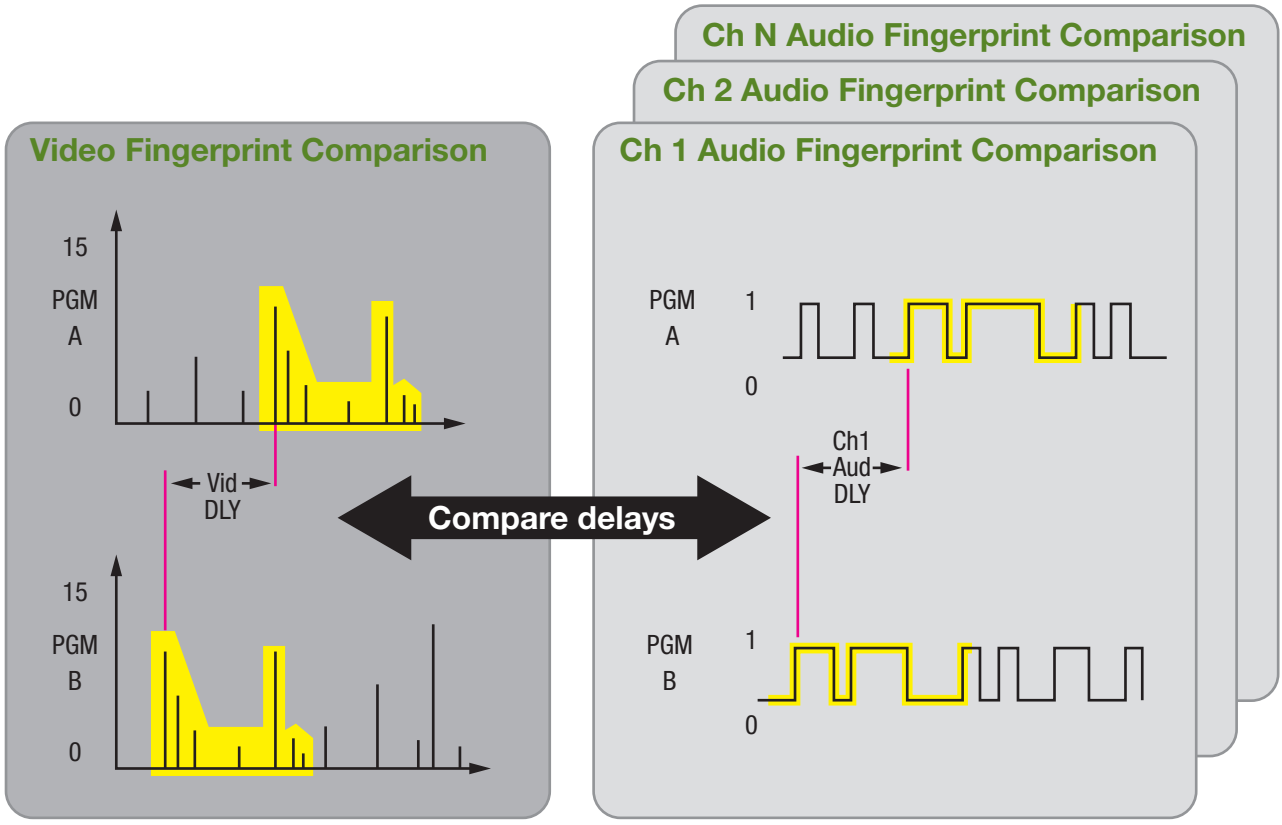


Figure 2: A simple convolution is used to look for matching patterns between two media fingerprint streams at different points in the playout process.

Once a match is found between the fingerprint streams, two timing measurements can be performed. The first is to calculate the program to program delay. This becomes the regular delay that is found within the signal distribution path. The second measurement is to calculate the difference between the video and audio delay. If the difference is zero, then there is no problem. However, if the difference is not zero, then we have a drift in audio with reference to the video. This drift is also known as a lip sync error. This comparison process allows any video or audio differences between two or more signals to be quickly identified. The delay between two streams can be measured with a resolution of just ±1 ms.

Fingerprint data streams and alarms can be analyzed by the latest generation of facility monitoring systems, and operators can be immediately alerted whenever problems emerge to promote rapid fault resolution.

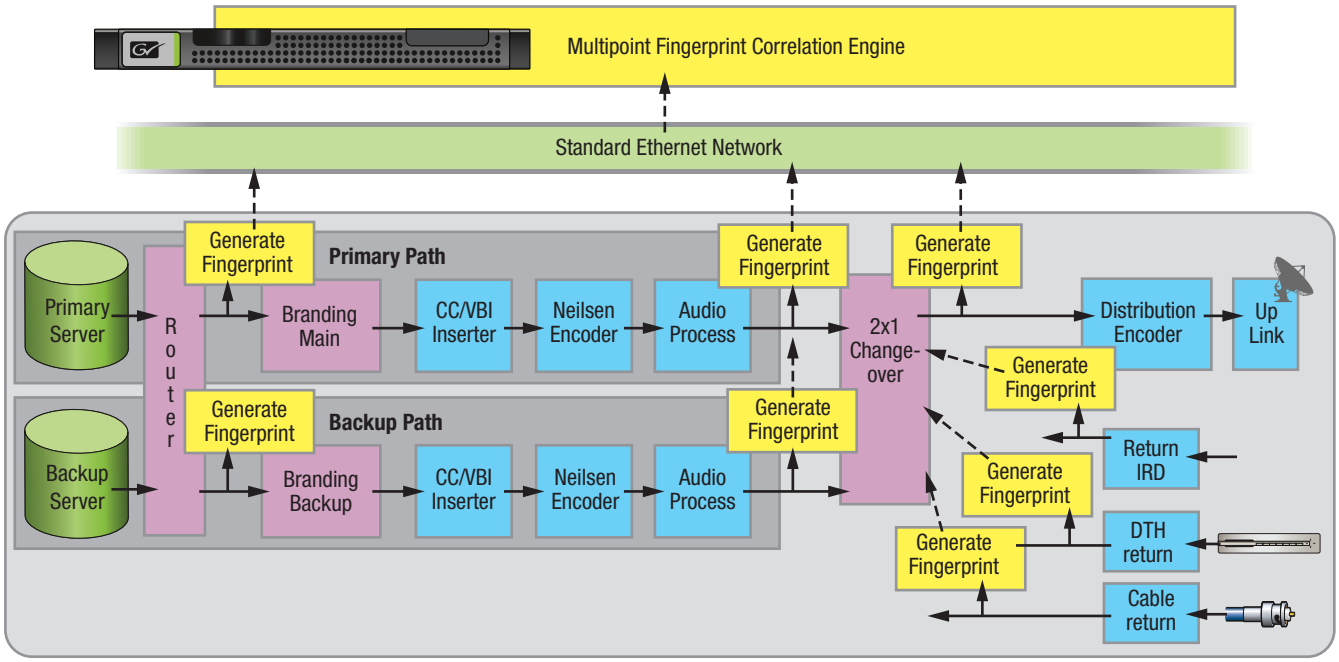


Figure 3: Multipoint media fingerprint monitoring across a facility to detect playout errors

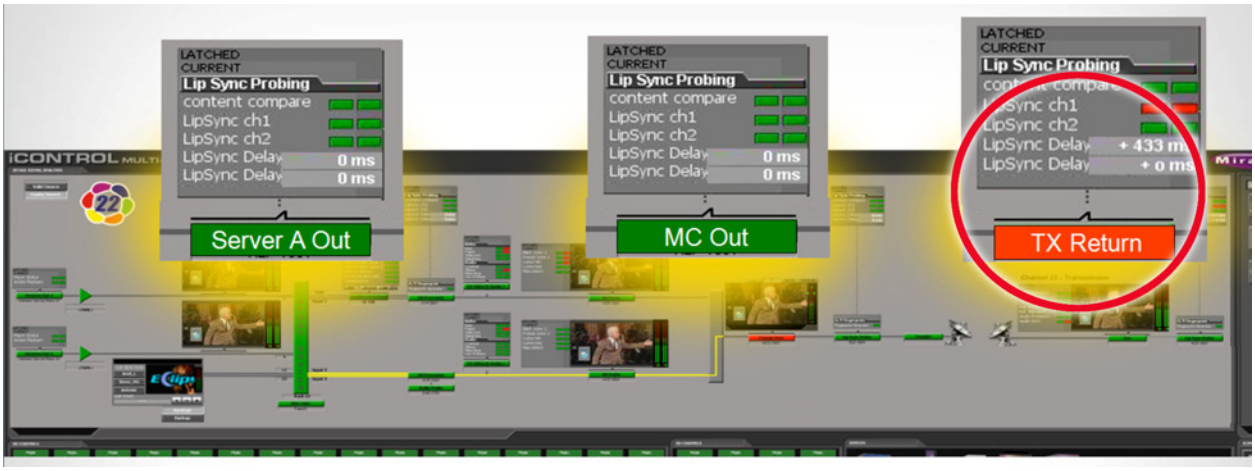
Content Verification and Lip Sync Detection

One of the key applications of fingerprinting for television playout is automated content verification while a channel is played out across its distribution path, from the server to the uplink, and back via the return feeds.

Fingerprint technology is well-suited to addressing many typical playout issues, such as a back-up channel not having exactly the same content as the primary channel. This kind of problem can be intercepted quickly and accurately by using fingerprint capture and monitoring along the primary and back-up playout paths. For instance, by using fingerprint detection at the server and back-up change-over, the facility monitoring system can quickly identify any content errors. Similarly, fingerprint detection at the distribution encoder and cable return can spot problems that have been encountered at a cable operator. Even subtle differences in content can be identified, such as missing branding graphics on the main playout chain, when they are present on the back-up channel.

Another key monitoring problem that can be detected with fingerprinting is lip sync errors. This has traditionally been one of the tougher monitoring challenges in a facility, as it is not immediately evident with traditional monitoring using a multiviewer. Naturally, this situation has become more demanding as channel counts have risen ever higher. Operators of a multichannel playout facility cannot be listening to the audio of each channel; it is just too confusing. Today, many operators rely on level meters to determine that audio is present. Unfortunately, this approach does not allow for the detection of a lip sync problem.

In contrast, fingerprinting can quickly identify lip sync errors anywhere along the playout chain, on any individual audio channel. These errors can trigger alarms to instantly direct the attention of the operator to the source of the problem (see Picture 1) so that the necessary corrective measures can be taken.



Picture 1: Using media fingerprinting to monitor lip sync issues using IP-based facility monitoring

Looking Ahead

Media fingerprint-based error detection is already a practical proposition for more streamlined television monitoring, when used alongside the latest generation of facility monitoring systems. However, the technology would really benefit from standardization and full interoperability, in the way that SNMP monitoring has become a benchmark for reporting across all types of television and telco equipment. This would remove the need for additional equipment to capture fingerprints across a facility, as the fingerprints could be generated via the equipment that is already part of the standard infrastructure.

Obviously, much wider deployment of media fingerprinting across many vendors is still some way off, but there are encouraging signs towards standardization of video and audio fingerprint algorithms. The 22TV Lip Sync Ad-Hoc Group recently met at the SMPTE Technology Committees in Montreal, and the media fingerprint technology from multiple vendors was reviewed as part of an ongoing process of developing a new standard. It is anticipated that this process will ultimately yield new SMPTE standards that will take this formative technology to the next level of adoption.