

No more “Whaaaat?!?”

Telos integrates the MPEG-H Production Technology Dialog+, enabling broadcasters to automatically enhance the dialogue of their content.

It's movie night and you are deeply involved in the adventures on your TV screen. The crucial scene unfolds and ... WHAT did they just say?! Why does that silly person have to start mumbling? Or was it the music that suddenly got louder? Who cares – the thread is lost, and you have to hit the rewind button.

Everybody has had this experience at least once in their life. Very annoying and not only for people who suffer from hearing loss. But it would not be fair to accuse the film's sound designers of not having done their job properly. Studies have long since established that the optimum loudness difference between dialogue and background sounds is a very personal matter and differs widely between individuals. This means that it is next to impossible to please everyone with one mix, regardless of how hard they try. But this does not mean that we have to endure mumbling actors until the end of time. In modern film production, creators and broadcasters can choose to implement audio systems such as MPEG-H Audio that enable listeners to adapt the loudness of individual audio objects to their personal taste.

And for our beloved old(-ish) movies, the scientists at Fraunhofer IIS developed MPEG-H Dialog+, a file-based dialogue separation technology. The AI-based MPEG-H production technology uses Deep Neural Networks to separate the dialogue from the background of an existing broadcast mix and outputs a new, easier-to-understand remix. The technology has just been implemented into the Telos Alliance Minnetonka AudioTools Server and released as their Dialog+ module.

In preparation of the release, Fraunhofer IIS and Telos teamed up with the German broadcaster WDR to create the best training environment for the AI. In a first step, the IIS conducted field tests over DVB and the VoD platform “ARD Mediathek” to refine the requirements and production workflows that should be implemented to achieve dialogue mixing with enhanced speech intelligibility. The results were then fed into the Telos product development. WDR as well as other ARD broadcasters provided suitable training material for the Deep Neural Network and led the workflow design. WDR also launched the “Klare Sprache” (Clear Speech) service in the “ARD Mediathek”. As the result of this close cooperation, the software has now been implemented as part of an automatic workflow – from archive to transcoding farm – in the WDR production infrastructure.

The integration of MPEG-H Dialog+ into the Telos Alliance product range enables broadcasters to rework their legacy content with a state-of-the-art dialogue separation algorithm in order to provide customizable audio mixes to their audience. The workflow is automated, which makes it scalable and cost-efficient. Thanks to a set of presets customized for different use cases, content providers can apply processing optimized, for example, for documentaries, music films, and sports content. MPEG-H Dialog+ can output two formats: A stereo mix for legacy workflows, plus an ADM file that supports all the groundbreaking features offered by Next Generation Audio, like user personalization and universal delivery for all playback devices.

And for audiences entering the world of their favorite characters this means that, finally!, there is no more “Whaaat?!” and they can lose themselves in the adventure completely and without interruptions.