

PRESS RELEASE

PRESS RELEASE

Sep 11th, 2023 || Page 1 | 3

Novelty at the fair: New tool for simultaneous face and speaker recognition enables fast people search in large media archives.

At the International Broadcasting Convention (IBC) trade fair in Amsterdam, the German Fraunhofer Institute for Digital Media Technology IDMT is presenting for the first time a solution that can automatically localize and identify people in large media archives based on their face and voice. For the "Audiovisual Identity Suite", the research institute has combined technologies for face and speaker recognition. With the help of Artificial Intelligence, vast amounts of media content are quickly analyzed for the presence of specific individuals.

The new combined face and speaker analysis offers program planners a comprehensive view of individual presences in TV broadcasts. For this purpose, the Audiovisual Identity Suite analyzes a large volume of data, i.e., any programs over many weeks, within a very short time. The results of the audiovisual recognition of specific persons are presented in an easy-to-understand and intuitive user interface and can be used for in-depth insights, trend analyses, and statistics.

If you want to detect the media presence of a specific individual within a program during a certain time, the tool shows in a so-called heatmap when and how often they were visible or audible on different TV channels. An important feature of the tool is that it also works reliably when the relevant person is speaking but is not shown in the picture. This is especially of interest in situations such as talk shows where reactions from the audience are captured, or other panelists are faded in while the person on the podium continues to speak.

This is possible by combining audio and video analysis methods. The institute has long-standing expertise in both research disciplines. Both analysis methods have already been successfully applied to various products and solutions.

Cross-modal combination of audio and video analysis methods

For the first time, the Audiovisual Identity Suite combines both methods into a cross-modal analysis tool. "This increases the validity and quality of the results significantly", explains Dr. Uwe Kühhirt, expert for video analysis at Fraunhofer IDMT and co-developer of the Audiovisual Identity Suite.

FRAUNHOFER INSTITUTE FOR DIGITAL MEDIA TECHNOLOGY IDMT

To identify people acoustically in programs, the institute relies on AI-based algorithms for recognizing speakers and classifying perceived gender. In addition, speech quality analysis enables the evaluation of entire programs or individual parts of programs regarding their acoustic intelligibility.

PRESS RELEASE

Sep 11th, 2023 || Page 2 | 3

Intelligent face recognition is used for the visual recognition of people in videos. In this process, facial features such as the visually perceived gender are extracted from the video data. In combination with the previously mentioned acoustic classification of perceived gender, very reliable statements can be made about how often men and women are seen or heard in the program. These findings can help, for example, in planning more gender-appropriate programming and for reporting.



Availability of the Audiovisual Identity Suite

Analyses and studies with the Audiovisual Identity Suite are initially carried out by Fraunhofer IDMT on behalf of the customer. The results of the analyses are then made available to the client in a customized user interface, prepared for his specific purposes.

*Picture 1:
The "Audiovisual Identity Suite" reliably identifies people in media archives by combining face and speaker recognition.
(Copyright Fraunhofer IDMT/istock.com/vm)*

In the future, the analysis tool should also be licensable for use at the customer's site.

Upcoming enhancements

The Audiovisual Identity Suite is set for further expansion. Upcoming features include age estimation based on visual analysis and audio advancements such as language recognition, speech-to-text conversion, and keyword analytics.

"Our planned enhancements will provide deeper opportunities for analysis. With the addition of text transcription, we can not only determine how often certain people appear but also which topics they are talking about," explains Christian Rollwage, expert for speaker recognition at the Fraunhofer IDMT Oldenburg Branch for Hearing, Speech and Audio Technology HSA.

FRAUNHOFER INSTITUTE FOR DIGITAL MEDIA TECHNOLOGY IDMT



*Picture 2:
The results of audiovisual
recognition of specific
individuals are presented in an
easy-to-understand and
intuitive dashboard and can be
used for trend analysis and
statistics.
(Copyright Fraunhofer IDMT).*

PRESS RELEASE

Sep 11th, 2023 || Page 3 | 3

Discover how the Audiovisual Identity Suite can simplify your daily work. Visit us from September 15 to 18, 2023 at the IBC trade show in Hall 8 at the Fraunhofer-Gesellschaft booth B.80 and let our experts show you the advantages of the new cross-modal analysis tool Audiovisual Identity Suite.

For pictures and more details about the Audiovisual Identity Suite visit <https://s.fhg.de/ibc2023>

About the Fraunhofer Institute for Digital Media Technology IDMT

The Fraunhofer Institute for Digital Media Technology IDMT is one of 76 institutes and research facilities of the Fraunhofer-Gesellschaft, the world's leading organization for application-oriented research.

At Fraunhofer IDMT's headquarters in Ilmenau (Thuringia, Germany), the experts work on the secure and efficient AI-based recognition and classification of audio and video data. Another focus is the development of audio technologies for virtual acoustic product experience as well as customized solutions for the production and reproduction of authentic and spatial sound experiences for the areas of professional audio, entertainment, and automotive.

Fraunhofer IDMT's Oldenburg Branch for Hearing, Speech and Audio Technology HSA stands for market-oriented research and development with a focus on speech and event recognition, sound quality, and speech intelligibility as well as mobile neurotechnology and systems for networked healthcare. With their in-house expertise in the development of hardware and software systems for audio system technology and signal enhancement, the employees at the Oldenburg site are responsible for transferring scientific findings into practical, customer-oriented solutions.