DMP title

Project Name C1 Internal Funds DMP - DMP title

Project Identifier C14/21/075

Grant Title C14/21/075

Principal Investigator / Researcher Marc Moonen / Toon van Waterschoot

Project Data Contact Aldona Niemiro-Sznajder, +32 16 329607, aldona.niemiro-sznajder@kuleuven.be

Description The project is targeted to speech and audio signal processing, such as in speech communication devices (e.g. smartphones), teleconferencing systems (e.g. VoIP in laptops), audio recording devices (e.g. digital video cameras), hearing instruments (e.g. hearing aids), voice controlled systems (e.g. smart speakers), sound reinforcement systems (e.g. public address systems, interactive virtual acoustics systems), and noise control systems (e.g. noise-cancelling headsets). Present-day devices often have advanced digital signal procesing (DSP) functions on board for improved sound capture and playback, e.g. for noise reduction and signal enhancement, dereverberation, acoustic echo and feedback cancellation, active noise control. These individual DSP functions, however, are still developed and analysed in isolation, so that whenever such DSP functions are cascaded in a complete system, their interaction -even counteraction- is rarely taken into account. To address this problem, a generic holistic approach will be developed in this project, integrating individual DSP functions in a common acoustic control problem statement and solution strategy, from which suitable integrated and distributed algorithms can then be designed. Alongside this generic approach, two representative and relevant use cases will be developed, focusing on i) hearing instruments technology and ii) public address and interactive virtual acoustics technology, where eventually the performance of the integrated and distributed algorithms can be compared with the current state-of-the-art. Data will be recorded in controlled lab environments and used for off-line and real-time simulations to asses the performance of all developed integrated algorithms against the performance of state-of-the-art algorithms.

Institution KU Leuven

1. General Information

Name of the project lead (PI)

Marc Moonen / Toon van waterschoot

Internal Funds Project number & title

C14/21/075 - "A holistic approach to the design of integrated and distributed digital signal processing algorithms for audio and speech communication devices"

2. Data description

2.1. Will you generate/collect new data and/or make use of existing data?

- Generate new data

2.2. What data will you collect, generate or reuse? Describe the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a numbered list or table and per objective of the project.

All data used to validate research results and/or to support research findings will be generated by the PIs and co-workers.

1) Experiment input data: Experiments will be conducted which will be based on input data either generated by computer simulations of specific acoustic scenarios, or recorded (collected) in specific acoustic scenarios (i.e. laboratory conditions). These data will be exported and stored in data formats that are compatible with the software (e.g. Matlab/Simulink(TM)) used for the experiments (e.g. mat and ascii files). Metadata about the data sets with specific information on the acoustic scenario and simulation/recording conditions will be stored in associated text files.

2) Experiments: Experiments (i.e. numerical simulations) will be conducted with standard scientific computing software packages (e.g. Matlab/Simulink), either running off-line on a laptop/desktop computer, or running in real-time on a real-time simulation and test platform (e.g. Speedgoat(TM)). Numerical simulation codes will be developed and stored in software specific file formats (e.g. Matlab m-files, and C code).

3) Experiment output data: Results from numerical simulations will exported and stored in data formats that can be imported in (other) software packages (e.g. Matlab/Simulink) for post-

processing and interpretation (e.g. mat and ascii files). Metadata about the data sets with specific information on the numerical simulations will be stored in associated text files.

3. Ethical and legal issues

3.1. Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

No.

3.2. Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s).

No.

3.3. Does your research possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

No.

The project aims to make fundamental contributions and publish its research findings in the open literature, possibly also accompanied by computer code that is then made publicly available. In a later stage, the project may be followed by application and valorization projects, possibly with third parties.

3.4. Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions regarding reuse and sharing are in place?

No.

4. Documentation and metadata

4.1. What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

The deliverables of the project include algorithms and software. Algorithms will be described/documented in internal reports and/or published in conference proceedings and scholarly journals. Software code will include documentation to allow re-use, e.g. details related to mathematical concepts and operations used in the software.

4.2. Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.

No.

Metadata with specific information on experiments carried out with the developed algorithms and software will be stored in associated text files.

5. Data storage and backup during the project

5.1. Where will the data be stored?

The data generated in the course of the project will be stored on storage facilities of the research unit.

5.2. How will the data be backed up?

The storage used by the research unit is backed up daily. The backup process is managed by the IT division of the department (ESAT).

5.3. Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

Yes.

The data generated in the project will be stored on storage facilities of the research unit, with a total size of 8TB, of which 4 TB is still available.

5.4. What are the expected costs for data storage and backup during the project? How will these costs be covered?

The storage facilities of the research unit are currently available at no cost.

5.5. Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

For all data stored in the department's servers, access is regulated by an access control list (ACL) that grants read-write access to the project owner and read-only access to specific users. The ACL is managed by the project owner. Client computers can access the data using: SMB2 (or higher) from specific IP ranges NFSv4 from specific (IT managed) systems.

6. Data preservation after the end of the project

6.1. Which data will be retained for the expected 10 year period after the end of the project? If only a selection of the data can/will be preserved, clearly state why this is the case (legal or contractual restrictions, physical preservation issues, ...).

Any relevant data produced in the course of the project (cfr supra), will be kept for a period of 10 years after the project has ended.

6.2. Where will these data be archived (= stored for the long term)?

After the period of 10 years, any data generated in the course of the project that is not longer in use will be removed from the department's servers.

6.3. What are the expected costs for data preservation during these 10 years? How will the costs be covered?

The storage facilities of the research unit are available for the researchers for free.

7. Data sharing and re-use

7.1. Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions or because of IP potential)?

No.

7.2. Which data will be made available after the end of the project?

The deliverables of the project include algorithms and software. Algorithms will be described/documented in internal reports (publicly available in the research group's publication repository) and/or published in conference proceedings and scholarly journals. Whenever relevant, software code will be made available together with the reports/publications, e.g. in venues provided by the publisher (such as IEEE's CodeOcean).

7.3. Where/how will the data be made available for reuse?

- In an Open Access repository
- In a restricted access repository

See previous question, open or restricted access depending on the selected publication venue.

7.4. When will the data be made available?

- Upon publication of the research results

7.5. Who will be able to access the data and under what conditions?

See previous questions (subscribers if published under restricted (toll) access, general public if published under open access).

7.6. What are the expected costs for data sharing? How will these costs be covered?

There are no additional costs of data sharing foreseen.

8. Responsibilities

8.1. Who will be responsible for the data documentation & metadata?

The PIs (Marc Moonen and Toon van Waterschoot) will be responsible for data documentation and metadata.

8.2. Who will be responsible for data storage & back up during the project?

The IT division of the department (ESAT) will be responsible for the data storage and back-up during the project.

8.3. Who will be responsible for ensuring data preservation and sharing?

The PIs (Marc Moonen and Toon van Waterschoot) will be responsible for the data preservation and reuse.

8.4. Who bears the end responsibility for updating & implementing this DMP?

The PIs (Marc Moonen and Toon van Waterschoot) bear the end responsibility of updating and implementing this DMP.