Application

2015-03-31

2015-05207	Schüldt, Christian	NT-14

Information about applicant

Name: Christian Schüldt

Doctorial degree: 2012-12-11

Birthdate: 19780816

Academic title: Doktor

Gender: Male

Employer: Limes Audio

Administrating organisation: Kungliga Tekniska högskolan

Project site: Avdelningen för Signalbehandling

Information about application

Call name: Forskningsbidrag Stora utlysningen 2015 (Naturvetenskap och teknikvetenskap)

Type of grant: Projektbidrag

Focus: Unga forskare

Subject area:

Project title (english): Open problems in multimedia signal processing related to reverberation

Project start: 2016-01-01

Project end: 2019-12-31

Review panel applied for: NT-14

Classification code: 20205. Signalbehandling, 20204. Telekommunikation

Keywords: Blind estimation, Reverberation time, Dereverberation

Funds applied for

Year:	2016	2017	2018	2019
Amount:	1,128,575	1,162,684	1,227,108	1,291,535

Descriptive data

Project info

Project title (Swedish)*

Öppna problem i multimedia-signalbehandling relaterat till reverberation

Project title (English)*

Open problems in multimedia signal processing related to reverberation

Abstract (English)*

Speech is perhaps the most fundamental mean of human communication. However, in many situations, especially in modern telecommunication, speech communication can be significantly impaired by noise and reverberation (i.e., the persistence of sound in an enclosure after the sound has been produced). Although extensive research progress have been made in the area of signal processing for speech enhancement, noise and reverberation still constitute challenging problems in modern communication.

A signal processing method to mitigate reverberation in speech signals is de-reverberation. Many de-reverberation methods rely, either directly or indirectly, on "blind" estimates of the reverberation time, i.e., estimates of the reverberation time directly from a recorded signal (e.g., speech) in an enclosure where the excitation signal is unknown, although it should be noted that blind reverberation time estimation also can be used for other applications such audio forensics, as well as non-intrusive acoustical measurements in ordinary rooms, offices and concert halls. However, blind reverbereration time estimation is a non-trivial problem; a recent comparison between three state-of-the-art algorithms show, e.g, that all algorithms suffer from serious performance degradation in case of noise, even under fairly moderate noise levels.

Two challenges have recently been presented to researchers in the areas of de-reverberation and blind reverberation time estimation. Results from the de-reverberation challenge shows that while most submissions manage to improve the objective measures (direct signal-to-reverberation ratio, signal-to-noise ratio, etc.), several submissions simultaneously reduce the subjective audio quality through audible artifacts, especially the algorithms that do only use one microphone. This is also the reason why speech de-reverberation seldom is used in current commercial telecommunication products. The blind reverberation time estimation challenge submission deadline has not yet passed, and thus the results from this challenge have not been made available. Once the results have been published it is crucial to investigate each algorithm in detail to fully understand advantages and disadvantages of the most promising algorithms. This will be the immediate initial focus of the project.

In a study preceding this project proposal, the applicant has conducted basic research on decay rate estimation for blind and non-blind reverberation time estimation. The results have reached interest from the research community and have recently been published in the most prestigious journals and conferences in the area. The purpose of the herein described project is to continue this line of research; to take the investigation of blind reverberation time estimation further, and to study its application for de-reverberation, with the goal to solve some of the open research problems related to both blind reverberation time estimation and de-reverberation. A significant amount of work in the project will be carried out by a PhD student, under supervision of the applicant.

Popular scientific description (Swedish)*

Reverberation (även efterklang eller eko) kan löst beskrivas som det ljud som finns kvar i ett utrymme efter att källan har tystnat. Ljud som utsatts för reverberation uppfattas vanligtvis som spatiöst och med djup. Dock lider tal som utsatts för reverberation både av nedsatt uppfattbarhet samt nedsatt kvalité, jämfört med icke-reverberant tal.

Ett sätt att minska mängden reverberation i inspelade signaler är genom de-reverberation. Nyligen genomfördes en forskartävling för de-reverberation av talsignaler, och resultatet visade att trots att de allra flesta metoderna för de-reverberation lyckades reducera ett antal objektiva reverberationsmått, så reducerade även många den subjektiva talkvalitén genom hörbara artefakter, särskilt de som endast använder en mikrofon. Detta visar tydligt på att mycket arbete återstår inom forskningsområdet, samt ger en fingervisning om varför de-reverberation idag sällan används i kommersiella telekommunikationsprodukter.

Många de-reverberationsalgoritmer förutsätter tillgängliga mått på efterklangstiden, dvs. den tid det tar för efterklangen/reverberationen att dö ut. Att skatta efterklangstiden utifrån en mikrofonsignal utan att känna till excitationssignalen kallas blind efterklangstidsskattning. Att skatta efterklangstiden på detta sätt är dock inte helt lätt, vilket bland annat tydliggjorts av en relativt nyligen genomförd jämförelse av tre olika metoder. Jämförelsen visar att samtliga metoder lider av överskattning i brusiga miljöer, även för måttliga brusnivåer. Även inom detta område har en forskartävling nyligen utlysts, men resultaten har ej ännu offentliggjorts.

Båda forskartävlingarna inom de-reverberation och blind skattning av efterklangstid visar på stor forskningsaktivitet inom området, och problemen med dagens lösningar (beskrivna ovan) understryker betydelsen av att framsteg görs.

Målet med detta forskningsprojekt är att försöka lösa några av de öppna forskningsproblemen som finns inom områdena de-reverberation och skattning av efterklangstid. Arbetet kommer initialt att fokusera på noggrann utvärdering av resultatet från tävlingen för blind skattning av efterklangstid, för att sedan mer i detalj undersöka samt samspelet mellan blind skattning av efterklangstid och de-reverberation.

Project period

Number of project years*

4

Calculated project time*

2016-01-01 - 2019-12-31

Deductible time

Deductible time

Cause	Months

Career age: 27

Career age is a description of the time from your first doctoral degree until the last day of the call. Your career age change if you have deductible time. Your career age is shown in months. For some calls there are restrictions in the career age.

Classifications

Select a minimum of one and a maximum of three SCB-codes in order of priority.

Select the SCB-code in three levels and then click the lower plus-button to save your selection.

SCB-codes*

2. Teknik > 202. Elektroteknik och elektronik > 20205. Signalbehandling

2. Teknik > 202. Elektroteknik och elektronik > 20204. Telekommunikation

---

Enter a minimum of three, and up to five, short keywords that describe your project.

Keyword 1*

Blind estimation

Keyword 2*

Reverberation time

Keyword 3*

Dereverberation

Keyword 4

Keyword 5

Ethical considerations

Specify any ethical issues that the project (or equivalent) raises, and describe how they will be addressed in your research. Also indicate the specific considerations that might be relevant to your application.

Reporting of ethical considerations*

Inga etiska frågor är aktuella.

The project includes handling of personal data

No

The project includes animal experiments

No

Account of experiments on humans

No

Research plan

Executive summary

Speech is perhaps the most fundamental mean of human communication. However, in many situations, especially in modern telecommunication, speech communication can be significantly impaired by noise and reverberation (i.e., the persistence of sound in an enclosure after the sound has been produced). Although extensive research progress have been made in the area of signal processing for speech enhancement, noise and reverberation still constitute challenging problems in modern communication. For example, the results from the recent REVERB challenge [23], which is a challenge for assessing and comparing current state-of-the-art methods for de-reverberation, shows that while most submissions manage to improve the objective measures (direct signal-to-reverberation ratio, signal-to-noise ratio, etc.), several submissions simultaneously reduce the subjective audio quality through audible artifacts, especially the algorithms that do only use one microphone. This is also the reason why speech de-reverberation seldom is used in current commercial telecommunication products.

Many de-reverberation methods rely, either directly or indirectly, on “blind” estimates of the reverberation time, i.e., estimates of the reverberation time from a recorded signal (e.g., speech) in an enclosure where the excitation signal is unknown, although it should be noted that blind reverberation time estimation also can be used for other applications such audio forensics, as well as non-intrusive acoustical measurements in ordinary rooms, offices and concert halls. However, blind reverberation time estimation is a non-trivial problem; a recent comparison between three state-of-the-art algorithms [3] show, e.g, that all algorithms suffer from serious performance degradation in case of noise, even under fairly moderate noise levels.

There has recently been an increasing research interest in the area of blind reverberation time estimation, illustrated by the ongoing ACE challenge [4] for comparison of current blind reverberation time estimation algorithms. At this time, the challenge submission deadline has not yet passed, and thus the results have not been made available. Once the results have been published it is crucial to investigate each algorithm in detail to fully understand advantages and disadvantages of the most promising algorithms. It should also be noted that there currently are a substantial number of open research problems related to both de-reverberation and blind reverberation time estimation: How does the reverberation time estimation performance affect de-reverberation? How is decay rate estimation performance affected when the decaying signal does not comply with the Polack model? How do automatic speech segmentation affect the estimation accuracy of different kinds of decay rate estimators? Is it feasible to combine blind beamforming and de-reverberation? Can spatial cues from video be used to aid blind reverberation time estimation and de-reverberation? (See also section Survey of the field below.)

This clearly indicates that there is still important work to be done in the field.

In a study preceding this project proposal, the applicant has conducted basic research on decay rate estimation for blind and non-blind reverberation time estimation. The theoretical results show that a simple linear regression-type estimator consistently performs about 4 dB worse compared to a maximum-likelihood estimator (which attains the Cramér Rao bound (CRB)), and that both estimators suffer from similar bias problems in case of noise. It has also been shown that the use of the backward integration approach for decay rate estimation can be problematic, especially in situations where the signal-to-noise ratio is poor, which can be the case for blind reverberation time estimation. The latter result is especially interesting, as there indeed have been some recent approaches using backward integration for blind reverberation time estimation.

The applicant is a young promising researcher who combines (1) basic research in academia, holding a post-doc position at KTH, and (2) applied research and development in a high-tech company, motivated by the increasingly faster turnaround from basic research to product development in the area of signals and systems. Although a young researcher in an early stage of the career, the applicant has already made a significant scientific impact, manifested by far more than 2,000 citations counted in Google Scholar. The 2004 conference paper on support vector machines has become a "Citation Classic" with some 1,800 citations alone.

Purpose and aims

The purpose of the project is to try and solve some of the current open problems in multimedia signal processing related to blind reverberation time estimation and speech de-reverberation. This will effectively give a better understanding of how to accurately measure and characterize the reverberation in reverberant audio signals, as well as give deepened insights into methods for reducing the amount of reverberation in reverberant (and possibly noisy) speech signals without reducing the speech intelligibility, without introducing artifacts, or otherwise compromising the speech quality. As well as broadening the general knowledge of the field of multimedia signal processing, the results will be of importance for acoustical measurements in ordinary rooms, offices and concert halls by providing means of increased accuracy. Moreover, the results will also be beneficial for other applications such as audio forensics, and help to improve the quality of speech communication systems (mobile phones, conference phones, video conferencing, etc.).

The aims can be concertized as

- To gain deepened insight into what algorithm sub-blocks are important for robust and accurate blind reverberation time estimation and how the sub-blocks interact. It should be noted that this currently is something that is practically unexplored in the research literature.
- An increased understanding of decay rate estimation accuracy when the signal does not comply with the Polack model (see the section Survey of the field below). This will be of importance for predicting estimation accuracy in realistic situations, as well as aid in the process of constructing a beyond state-of-the art blind reverberation time estimator.
- To study and to quantify the importance of blind reverberation time estimation accuracy for blind speech de-reverberation. The obtained results will of interest for deciding future research directions.
- Possible exploration of the use of visual cues from video to aid de-reverberation and/or reverberation time estimation.

Survey of the field

When a sound is produced in an enclosure, such as, e.g., when a person is speaking in a room, the sound waves are reflected (and damped) by the walls, ceiling and floor, and the interior. A microphone located in the room will thus pick up this reverberation, in addition to the (more or less) direct sound waves. The reverberation will significantly affect the speech through temporal smearing, which in turn reduces the intelligibility, as well as the quality. In case of strong reverberation, the speech will smear out so that reverberation from one syllable is still heard in the room when the next syllable is spoken, significantly reducing the intelligibility.

Perhaps the most fundamental measurable entity associated with the reverberation is the reverberation time T_{60}, which is defined as the time it takes for a sound to decay 60 dB. The reverberation time has been measured since the late 19 century, when W. C. Sabine started investigating the relationship between sound absorption and reverberation time, using organ pipes, a stop-watch and his ears. In 1965, M. R. Schröder showed that the ensemble average of squared white noise decay is identical to a certain integral over the squared room impulse response, implying that the reverberation time could be obtained directly from the impulse response, rather than from measuring the decay of multiple noise bursts [1]. The technique proposed by Schröder is commonly denoted backward integration and is often used for estimation of the reverberation time in many situations, also being part of the ISO 3382 standard [2].

Unfortunately, in many cases it is not practical, or even possible, to perform measurements using a controlled excitation signal. In such cases, the only available information is typically a reverberant signal, containing speech and/or other sounds, and a ”blind” approach is required for estimating the reverberation time. A number of different blind estimators have been proposed in the literature, and a fairly recent comparison have been made between three methods [3], showing that all three methods are able to estimate the reverberation time within ±0.2 s for $T_{60} \leq 0.8$ s and signal-to-noise-ratio ≥ 30 dB, while increasing the noise level causes overestimation, which illustrates that there indeed is room for improvement within the field. Also, a recent challenge named the Acoustic Characterisation of Environments (ACE) challenge [4] has been announced, which will allow comparison of state-of-the-art algorithms and the gaining of new insights regarding potential future research directions. As blind estimators typically comprise multiple algorithm components, such as e.g., segmentation mechanisms for finding signal sections where the reverberant energy is dominant, and decay rate estimators for estimating the rate of decay of the reverberant energy, analyzing the results from the ACE challenge and comparing the competing algorithms will give an indication of which algorithm components are most important for the overall performance in a number of real scenarios.

The applicant Christian Schüldt has, together with professor Peter Händel at KTH Department of Signal Processing, recently carried out research on theoretical bounds and performance of different types of decay rate estimators, both in the context of blind as well as non-blind reverberation time estimation [5] [6] [7] [8]. The results show that a simple linear regression-type estimator, such as used in, e.g., [9] [10] [11] [12], consistently performs about 4 dB worse compared to a maximum-likelihood estimator (which attains the Cramér Rao bound (CRB)), used in, e.g., [13] [14] [15] [16], and that both estimators suffer from similar bias problems in case of noise [5]. In [6], the well-known Schröder backward integration method [1], which essentially can be seen as a linear regression pre-processing step, is analyzed. It is shown that the backward integration process has the potential to improve the linear-regression performance down to the CRB, although in some cases the estimation performance can decrease significantly, and that controlling, or even predicting, this effect can be non-trivial, especially in situations with strong noise. This means that if using backward integration in blind reverberation time estimation applications, as in, e.g., [17] and [18], special care must be taken to avoid degradation of the estimation accuracy.

The theoretical analysis carried out in [5] [6] [7] and [8] is based on a common model named the Polack model [19], which is simplistic but captures the most important properties associated with reverberation, given a number of assumptions. This model is well established and widely used by researchers in the field, see e.g., [20] and the references therein. However, it has not been quantified how well different types of decay rate estimators perform when the decaying signal does not comply with the Polack model, such as e.g., when the source (speaker)-to-microphone distance is less than the critical distance and/or when the sound field is non-diffuse [21]. This is an important challenge, as increased knowledge in this regard would mean that more general bounds on estimation performance could be derived. This would in turn increase the knowledge of what T_{60} estimation performance that can be expected in an increased number of real situations.

In blind reverberation time estimation, the reverberant signal is often automatically segmented into different sections corresponding to diffuse reverberation decay, and decay rate estimation is applied to each segment [9] [14] [15]. Thus, in such situations it is essentially up to the segmentation algorithm to define if the decay rate should be estimated from a signal section or not, and hence the segmentation performance will significantly affect the overall blind reverberation time estimation performance. It should however also be noted that while some estimators rely on this type of automatic segmentation in one for or another, numerous alternative approaches have been presented where the decay rate is estimated continuously over the whole recorded signal [10] [11] [12] [13] [16], hence not utilizing any explicit segmentation and instead relying on statistical analysis such as, e.g., histogram-based post-processing. Nevertheless, no comprehensive analysis of different automatic segmentation approaches for blind reverberation time estimation has been carried out, and it has not been shown whether or not this segmentation actually is beneficial for the estimation performance, compared to, e.g., the histogram based post-processing (although it is fairly intuitive that the segmentation tends to contribute to a reduction of the overall computational complexity, and the method in [14] is shown to be more suitable to track time-varying reverberation times than related approaches). Furthermore, in audio forensic applications, it is common to manually, or semi-manually, identify sections of decaying reverberation [15]. It has been shown that this manual interaction aids the T_{60} estimation performance [15], indicating that there indeed is room for potential improvement of fully automatic segmentation.

A common application for blind reverberation time estimation algorithms is blind speech de-reverberation, where a speech signal is processed in order to increase the speech-to-reverberation ratio. Many such methods are based on a spectral subtraction approach, where the signal is essentially filtered by an adaptive Wiener filter, which is continuously adapted according to the estimated speech-to-reverberation ratio [20] [22]. Such an approach is typically performed in frequency sub-bands, and ideally an estimate of the reverberation time in each sub-band is used in the de-reverberation process. Interesting in this regard is that the relationship between the sub-band and full-band T_{60} estimates is quite difficult to model and constitutes an open problem in the associated literature [17] (see also the references therein).

Recently, a challenge for blind speech de-reverberation, called the REverberant Voice Enhancement and Recognition Benchmark (REVERB) challenge [23], was organized and carried out. The results from the challenge show that most submissions manage to improve the objective measures (direct signal-to-reverberation ratio, signal-to-noise ratio, etc.), but several submissions simultaneously reduce the subjective audio quality through audible artifacts, especially the algorithms that do only use one microphone. The conclusion is thus that de-reverberation is still a challenging task, with much room for improvements. It can also be noted that the challenge did not evaluate speech intelligibility, which is a crucial measure for speech communication applications. Another open research problem is how the accuracy of reverberation time estimation affects the de-reverberation process. Moreover, the use of microphone beamforming [24] in this context is interesting; one of the best performing algorithms of the REVERB challenge used microphone beamforming as a pre-processing step to the actual de-reverberation [22], and the beamforming approach has also recently been used for reverberation time estimation [25]. Hence, in a multi-microphone situation, it is clear that a beamforming approach could facilitate both reverberation time estimation and de-reverberation. Although some research has been made in this area (see e.g., [20] [22] [25] and the references therein), there are many aspects that are still unexplored (for example, the combination of blind beamforming[1] and de-reverberation has, to the best of our knowledge, not been considered in the literature), with great potential for new insights and improvements.

Another interesting approach is the fusing of audio and video information to perform both reverberation time estimation and de-reverberation. For applications such as video conferencing, there has been a number of methods for speakers tracking and the use of this information for adaptive microphone beamforming, i.e., essentially directing a beam towards the speaker [26] [27]. This basic principle of using spatial video cues for audio processing could likely also be used for, e.g., first estimating how far away from the camera the current speaker is, using either some stereoscopic camera rig and/or simply measuring the size of the speakers head in individual video frames, and then use this distance information as an aid in the estimation of the direct speech-to-reverberation ratio (naturally, the direct speech-to-reverberation ratio reduces with increased distance between speaker and microphone) for de-reverberation. Also, additional spatial

1 Where the spatial locations of the microphones in the array are unknown.

video clues such as corners, walls, etc., could potentially be used when estimating the reverberation time.

Project description

The results from the ACE challenge will give a solid indication of what type of algorithms are accurate and robust in practical scenarios. However, to fully understand advantages and disadvantages of the most promising algorithms, it is crucial to investigate each algorithm in detail, to perform extended controlled tests and to modify parameters in an iterative manner. Thus, initially the goal of the project will focus on on-depth analysis of the results from the ACE challenge and its algorithms (**phase #1**). Crucial in this phase is the identification of why certain algorithms work well (or not) in certain situations. Also, the goal is to try and solve the automatic segmentation problem (discussed in the previous section Survey of the field) for finding signals sections where the decaying reverberation is dominant, in conjunction with the applicant's previous work on decay rate estimation in [5] [6] [7] [8]. It is estimated that this phase will take place over about 12 months, but it will naturally depend on a number of factors such as the actual results of the challenge and if the source code of the algorithms of interest are publicly available or not.

In the next phase (**phase #2**) of the project, the goal is to study decay rate estimation in cases when the decaying (reverberated) signal does not comply with the Polack model. This will provide increased knowledge of decay rate estimation for more general cases, which would be useful for predicting estimation accuracy in realistic situations, as well as aid the process of designing a beyond state-of-the art blind reverberation time estimator – something which is a partial goal in this phase. The premises of the estimator, e.g., if it utilizes one or multiple microphones, if the spatial microphone array structure must be known (in case of multiple microphones), is undecided at this point, but will be determined after investigating the results of the ACE challenge (i.e., during/after the first project phase).

Since many of the successful algorithms in the REVERB challenge rely upon available (either direct of indirect) estimates of the reverberation time, the following phase (**phase #3**) will be dedicated to investigating the effect of blind reverberation time estimation on blind speech de-reverberation. This effect has not been studied previously in the literature, as described in the previous section Survey of the field, but is essential for determining, e.g., what actual de-reverberation improvement can be achieved with more accurate reverberation time estimation. An important part of this project phase will be to investigate and quantify this matter.

The final project phase (**phase #4**) will be dedicated to further improvement of blind reverberation time estimation and/or de-reverberation. The exact focus of this phase is not determined at the moment, but will be decided during the course of the project, depending on the findings and results obtained in previous phases. One idea is to further improve the reverberation time estimation accuracy using spatial cues from a video camera. However, if it is shown in the preceding project phase that the accuracy of the reverberation time estimates are not critical, or that the accuracy of the beyond state-of the art method (or even the current state-of-the-art methods) is in some sense ”good enough” for the de-reverberation process, this project phase will of course focus on other matters. Perhaps the distance information obtained from the spatial cues can be used in some other way (i.e., apart from explicitly estimating the reverberation time) to aid the de-reverberation process?

The different phases of the project are roughly scheduled as

- Phase #1 – Jan 2016 to Dec 2016
- Phase #2 – Jul 2016 to Dec 2017
- Phase #3 – Jan 2017 to Jun 2018
- Phase #4 – Jan 2018 to Dec 2019

Note that there is some overlap of the phases.

A significant amount of work in the project will be carried out by a PhD student who will be admitted at KTH, under supervision of the applicant, Christian Schüldt, together with professor Peter Händel. The PhD student will be involved full time (100%) in the project while Christian will be working 25%. Peter Händel will function as main advisor to the PhD student and will be working 5% in the project.

A strong candidate for the PhD student position is Robin Lundberg, who is currently employed as a developer at Limes Audio and has shown excellence in both his work at Limes Audio, as well as in his preceding studies.

Significance

Multimedia (audio and video) signal processing have traditionally been the main signal processing applications, and are still today - illustrated, e.g., by the fact that the IEEE Signal Processing Society's two “flagship” conferences are the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) and the International Conference on Image Processing (ICIP). At KTH, this field of research has since 1998, until the end of 2012, been carried out at the Sound and Image Processing Lab headed by professor Bastiaan Kleijn, now at the Victoria University of Wellington, New Zealand, and professor Arne Leijon, recently retired. After 2012 the Sound and Image Processing Lab was absorbed by the Communication Theory group and efforts have been made to secure the competence for the future through, e.g., employing an tenure track assistant professor in multimedia communication first in 2012, without finding any suitable applicant, and now again in 2015. It is of national strategic importance that the research competence at KTH - Sweden's largest technical university, within the field of multimedia signal processing is nurtured and evolved, something that will be greatly aided by a granted research project as described herein.

The ACE and REVERB challenges [4] [23] clearly illustrate that the areas of blind room acoustic parameter estimation and speech de-reverberation recently have attracted significant research interest. As mentioned previously, the challenges allow direct performance comparison of state-of-the-art algorithms in each corresponding area, although there is also a need for more analytical comparison of the algorithms and their respective components, to give an understanding of *why* an algorithm is performing well or not. A successful project will provide a deepened understanding in this regard. Moreover, a successful project will also lead to better performing algorithms for both (blind) reverberation time estimation and de-reverberation, helping to improve various different applications such as audio forensics, where a single recording can help characterizing a room, giving an indication of *where* the recording was made, speech communication systems (mobile phones, conference phones, video conferencing, etc.) and acoustical measurements in general. The use of microphone beamforming to increase audio pickup quality by reducing noise and reverberation is gaining popularity in the telecommunication industry, exemplified by audio conferencing products with beamforming functionality from companies such as ClearOne, Mitel and Lifesize (acquired by Logitech in 2009, but still a separate brand). Speech de-reverberation, on the other hand, is seldom used due to the fact that current state-of-the-art algorithms typically are not robust enough; e.g., although the actual reverberation is reduced, the subjective audio quality can in some cases be compromised through audible artifacts (as mentioned in the previous section Survey of the field). Improving the performance and robustness of speech de-reverberation approaches in real practical situations, especially for single microphones, or e.g., used in conjunction with beamforming, could thus be of great benefit to audio conferencing.

Preliminary results

In a study preceding this project proposal, the applicant Christian Schüldt, together with Peter Händel, has conducted basic research on decay rate estimation for blind and non-blind reverberation time estimation. The results have reached interest from the research community and have recently been published in the most prestigious journals and conferences in the area [5] [6] [7] [8]. The theoretical results show that a simple linear regression-type estimator consistently performs about 4 dB worse compared to a maximum-likelihood estimator (which attains the Cramér Rao bound (CRB)), and that both estimators suffer from similar bias problems in case of noise [5]. It has also been shown that the use of the backward integration approach for decay rate estimation can be problematic, especially in situations where the signal-to-noise ratio is poor, which indeed can be the case for blind reverberation time estimation [6]. Two novel low complexity decay rate estimators exhibiting noise robustness have also been proposed [5] [7], which is promising with the noise robustness issues of current blind reverberation time algorithms [3] in mind. The work in [5] [6] [7] [8] should be seen as a preceding step to the described project. For example, decay rate estimation could be combined with some type of automatic segmentation approach for finding sections in a speech signal with dominant decaying reverberation, and/or possibly other type of processing such as pre-whitening [16] [28] and histogram-based techniques [13] [14].

Independent line of research

Christian Schüldt finished his PhD studies in December 2012, with professor Ingvar Claesson at Blekinge Institute of Technology (BTH) as main supervisor. During his studies he was also admitted to the Graduate School of Telecommunications – GST, which was a research and educational collaboration within the field of mobile and wireless communications between The Royal Institute of Technology (KTH), Blekinge Institute of Technology (BTH), Mid Sweden University (MU), and University of Gävle (HiG). During the final years before presenting his dissertation, Christian has been working increasingly independently (illustrated, e.g., by the fact that he effectively was pursuing his studies at a distance from BTH, Karlskrona, instead being located in Stockholm). Having left BTH, he joined the Department of Signal Processing at KTH in 2013 as a part time post-doc researcher under the supervision of professor Peter Händel, and initially shared his time equally between KTH and Limes Audio (of which he is one of the founders and currently CTO), i.e., 50% at KTH and 50% at Limes Audio. The following year (2014), he was working 25% at KTH and 75% at Limes Audio, and currently during 2015 he is working 5% at KTH and 95% at Limes Audio. The main reason for the decreasing amount of time spent with KTH, even though Limes Audio in some sense has required his increasing attention, is the lack of research funding. A granted research project will give Christian opportunity to focus on his basic research and also provide an opportunity to grow and to develop as and independent researcher through the supervision of the PhD student. Being involved with both KTH and Limes Audio has allowed the combination of basic research in academia and applied research and development in a young hot fast growing technological company – Limes Audio secured a place on the Swedish “33-listan” (Ny Teknik and Affärsvärlden's prestigious list of Sweden's hottest growing tech firms) in both 2013 and 2014, and was also listed as number 21 on Computer Sweden's HIT list of the hottest IT companies in Sweden during 2014. It is important to note that all ideas outlined in this project proposal originates from and are developed by Christian, and the intention is that the PhD student will assist in carrying out the research. Through this process, the PhD student will learn research methodology and Christian will gain graduate student level supervision experience. Formally, professor Peter Händel will function as the PhD student's main supervisor and Christian will be assistant supervisor.

Employment status

Christian is currently working 95% as Chief Technical Officer (CTO) at Limes Audio AB and 5% as a researcher at KTH Department of Signal Processing. During the herein described project, Christian will be part-time employed (25%) by KTH, funded by the project throughout its entire length, as well as 75% at Limes Audio. The PhD student will be employed by KTH.

Equipment

Limes Audio will provide resources in form of real acoustic environments for testing. The company has a quite unique and well-equipped lab and is conducting acoustic tests on with various custom-built loudspeaker- and microphone configurations on a daily basis. Limes Audio's research and development (R&D) team consists of around a dozen engineers specialized in audio signal processing and related areas, such as audio measurements as well as mechanical and electrical design with focus on audio applications. The team collaborates, and are engaged in development projects, with companies like Avaya, Cisco, FoxConn, Google, Alcatel-Lucent and Samsung, and can provide relevant input in terms of where the industry and market is heading and the demands and desires of the customers.

References

[1] M. R. Schroeder, "New method of measuring reverberation time," *Journal of the Acoustical Society of America*, vol. 37, no. 3, pp. 409–412, 1965.

[2] ISO 3382-2:2008, *Acoustics - Measurement of room acoustic parameters - Part 2: Reverberation time in ordinary rooms*, International Standards Organization, 2008.

[3] N. D. Gaubitch, H. W. Löllmann, M. Jeub, T. H. Falk, P. A. Naylor, P. Vary, and M. Brookes, "Performance comparison of algorithms for blind reverberation time estimation from speech," in *Proceedings of the International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept. 2012.

[4] The Acoustic Characterisation of Environments (ACE) Challenge http://www.ee.ic.ac.uk/naylor/ACEweb/index.html

[5] C. Schüldt and P. Händel, "Decay rate estimators and their performance for blind reverberation time estimation," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 8, pp. 1274-1284, Aug. 2014.

[6] C. Schüldt and P. Händel, "On implications of the ISO 3382 backward integration method for decay rate estimation," *Journal of the Audio Engineering Society*, Accepted for publication 2015.

[7] C. Schüldt and P. Händel, "Noise robust integration for blind and non-blind reverberation time estimation," to appear in the *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015.

[8] C. Schüldt and P. Händel, "Blind low-complexity estimation of reverberation time," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, Oct. 2013.

[9] K. Lebart, J. M. Boucher, and P. N. Denbigh, "A new method based on spectral subtraction for speech dereverberation," *Acta Acustica united with Acustica*, vol. 87, no. 3, pp. 359–366, May/June 2001.

[10] J. Y. Wen, E. A. Habets, and P. A. Naylor, “Blind estimation of reverberation time based on the distribution of signal decay rates,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 329–332, March 2008.

[11] R. Talmon and E. A. P. Habets, “Blind reverberation time estimation by intrinsic modeling of reverberant speech,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 156– 160, May 2013.

[12] B. Dumortier and E. Vincent, “Blind RT60 estimation robust across room sizes and source distances,” in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2014.

[13] R. Ratnam, D. L. Jones, B. C. Wheeler, W. D. O’Brien, C. R. Lansing, and A. S. Feng, “Blind estimation of reverberation time,” *The Journal of the Acoustical Society of America*, vol. 114, no. 5, pp. 2877–2892, Nov. 2003.

[14] H. W. Löllmann, E. Yilmaz, M. Jeub, and P. Vary, “An improved algorithm for blind reverberation time estimation,” in *Proceedings of the International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Aug. 2010.

[15] H. Malik, ”Acoustic Environment Identification and Its Applications to Audio Forensics ,” *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 11, pp. 1827–1837, Sept. 2013.

[16] A. Keshavarz, S. Mosayyebpour, M. Biguesh, T.A. Gulliver, and M. Esmaeili, “Speech-model based accurate blind reverberation time estimation using an LPC filter,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 6, pp. 1884–1893, 2012.

[17] T. d. M. Prego, A. A. de Lima, S. L. Netto, B. Lee, A. Said, R. W. Schafer, and T. Kalker, “A blind algorithm for reverberation-time estimation using subband decomposition of speech signals,” *The Journal of the Acoustical Society of America*, vol. 131, pp. 2811–2816, 2012.

[18] S. Vesa and A. Harma, “Automatic estimation of reverberation time from binaural signals,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2005, vol. 3, pp. 281–284.

[19] J-D. Polack, *La transmission de l’energie sonore dans les salles, Ph.D. thesis*, Le Mans, 1988.

[20] E. A. P. Habets, *Single- and Multi-Microphone Speech Dereverberation using Spectral Enhancement*, PhD thesis, Jun 2007.

[21] H. Kuttruff, *Room Acoustics*, 4th ed. Taylor & Francis, Oct. 2000.

[22] B. Cauchi, I. Kodrasi, R. Rehr, S. Gerlach, A. Jukic, T. Gerkmann, S. Doclo, and S. Goetze, “Joint dereverberation and noise reduction using beamforming and a single-channel speech enhancement scheme,” in *Proceedings of The REVERB challenge*, Florence, Italy, May 2014.

[23] The REverberant Voice Enhancement and Recognition Benchmark (REVERB) challenge http://reverb2014.dereverberation.com

[24] J. Benesty, J. Chen, and Y. Huang, *Microphone array signal processing,* Springer, vol. 1, 2008.

[25] J. Eaton, A. Moore, P. A. Naylor, and J. Skoglund, ”Direct-to-reverberant ratio estimation using a null-steered beamformer,” to appear in the *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2015.

[26] J-M Valin, F. Michaud, and J. Rouat, "Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering," *Robotics and Autonomous Systems,* vol. 55, no. 5, pp. 216–228, March 2007.

[27] N. Strobel, S. Spors, and R. Rabenstein, "Joint audio-video object localization and tracking," IEEE *Signal Processing Magazine,* vol. 18, no. 1, pp. 22–31, Jan. 2001.

[28] N. López, Y. Grenier, and I. Bourmeyster, "Low variance blind estimation of the reverberation time," in *Proceedings of the International Workshop for Acoustic Echo Cancellation and Noise Control (IWAENC)*, Sept. 2012.

Interdisciplinarity

My application is interdisciplinary

☐

An interdisciplinary research project is defined in this call for proposals as a project that can not be completed without knowledge, methods, terminology, data and researchers from more than one of the Swedish Research Councils subject areas; Medicine and health, Natural and engineering sciences, Humanities and social sciences and Educational sciences. If your research project is interdisciplinary according to this definition, you indicate and explain this here.

Click here for more information

Scientific report

Scientific report/Account for scientific activities of previous project

Budget and research resources

Project staff

Describe the staff that will be working in the project and the salary that is applied for in the project budget. Enter the full amount, not in thousands SEK.

Participating researchers that accept an invitation to participate in the application will be displayed automatically under Dedicated time for this project. Note that it will take a few minutes before the information is updated, and that it might be necessary for the project leader to close and reopen the form.

Dedicated time for this project*

	Role in the project	Name	Percent of full time
1	PhD Student	Robin Lundberg	100
2	Applicant	Christian Schüldt	25

Salaries including social fees

	Role in the project	Name	Percent of salary	2016	2017	2018	2019	Total
1	PhD Student	Robin Lundberg	100	491,040	509,640	546,840	584,040	2,131,560
2	Applicant	Christian Schüldt	25	186,000	188,325	190,650	192,975	757,950
	Total			677,040	697,965	737,490	777,015	2,889,510

Other costs

Describe the other project costs for which you apply from the Swedish Research Council. Enter the full amount, not in thousands SEK.

Premises

	Type of premises	2016	2017	2018	2019	Total
1	Kontor/labb	67,704	69,797	73,749	77,702	288,952
	Total	67,704	69,797	73,749	77,702	288,952

Running Costs

	Running Cost	Description	2016	2017	2018	2019	Total
1	Publiceringskostnader		25,000	25,000	25,000	25,000	100,000
2	IT/infrastruktur		54,163	55,837	58,999	62,161	231,160
3	Overhead KTH		236,964	244,288	258,121	271,955	1,011,328
4	Resor		67,704	69,797	73,749	77,702	288,952
	Total		383,831	394,922	415,869	436,818	1,631,440

Depreciation costs

Depreciation cost	Description	2016	2017	2018	2019

Total project cost

Below you can see a summary of the costs in your budget, which are the costs that you apply for from the Swedish Research Council. Indirect costs are entered separately into the table.

Under Other costs you can enter which costs, aside from the ones you apply for from the Swedish Research Council, that the project includes. Add the full amounts, not in thousands of SEK.

The subtotal plus indirect costs are the total per year that you apply for.

Total budget

Specified costs	2016	2017	2018	2019	Total, applied	Other costs	Total cost
Salaries including social fees	677,040	697,965	737,490	777,015	2,889,510		2,889,510
Running costs	383,831	394,922	415,869	436,818	1,631,440		1,631,440
Depreciation costs					0		0
Premises	67,704	69,797	73,749	77,702	288,952		288,952
Subtotal	1,128,575	1,162,684	1,227,108	1,291,535	4,809,902	0	4,809,902
Indirect costs					0		0
Total project cost	1,128,575	1,162,684	1,227,108	1,291,535	4,809,902	0	4,809,902

Explanation of the proposed budget

Briefly justify each proposed cost in the stated budget.

Explanation of the proposed budget*

The PhD student will be employed by KTH and the salary will be according to the KTH collective bargaining agreement for PhD students according to RALS 2013, dnr V-2014-0249, where the salary is based on a staircase model, increasing each year along with the student's progress. The salaries for employees hired between 2013-10-01 and 2014-09-30 are 26,400, 27,400, 29,400 and 31,400 SEK/month, for each of the 4 years, respectively. Based on these figures, and with addition of 55,5% social charges, the total budgeted salary cost for the PhD student is 2,131,560 SEK.

The project manager will be employed by KTH 25%. A full-time salary of 40,000 SEK/month for the first year, 40,500 SEK/month the second year, 41,000 SEK/month the third year and 41,500 SEK/month the fourth year gives a total of 757,950 SEK for the whole project.

Cost of the premises, as well as travel costs are each budgeted at 10% of the salary costs.

Publication costs for scientific journals and conferences are budgeted at 25,000 SEK/year, while IT/infrastructure and KTH overhead costs are budgeted at 8% and 35%, respectively, of the salary costs.

Other funding

Describe your other project funding for the project period (applied for or granted) aside from that which you apply for from the Swedish Research Council. Write the whole sum, not thousands of SEK.

Other funding for this project

Funder	Applicant/project leader	Type of grant	Reg no or equiv.	2016	2017	2018	2019

CV and publications

CV

Curriculum Vitae – Christian Schüldt (born: 1978)

1. **Högskoleexamen (Higher education qualification(s)):**
   2001, Högskoleingenjör elektroteknik, KTH Kista
   2004, Civilingenjör elektroteknik, KTH
2. **Doktorsexamen (Doctoral degree):**
   2012, Telekommunikation, BTH (GST - Graduate School of Telecommunications), Low-Complexity Algorithms for Echo Cancellation in Audio Conferencing Systems, Ingvar Claesson
3. **Postdoktorsvistelser (Postdoctoral positions):**
   2013 KTH (50%)
   2014 KTH (25%)
4. **Docentkompetens (Qualification required for appointments as a docent):**
5. **Nuvarande anställning (Current position):**
   CTO Limes Audio (95% of full-time) and researcher at KTH (5%)
6. **Tidigare anställningar (Previous positions and periods of appointment):**

- Embedded systems & DSP consultant, self employed, April 2007 – Jan 2008
- Software developer, Konftel, June 2006 – April 2007
- Research engineer, BTH, June 2004 – June 2006
- Software developer (part-time), Zenitel RadioTeknik AB, May 1999 – August 2003
- Software developer, TicketAnywhere, March 2001 – July 2001

7. **Uppehåll i forskningen (Interruption in research):**
8. **Handledning (Supervision):**
9. **Eventuell övrig information av betydelse för ansökan (Other information of relevance to the application)**

- Audio & Video Systems Tack Chair, IEEE International Conference on Consumer Electronics, 2013 –
- Audio Track Chair, IEEE International Conference on Consumer Electronics, 2008 – 2013.

Publication list

List of publications (last 8 years) – Christian Schüldt

The five publications considered most important for the project are marked with an asterisk (*).

1. **Fackgranskade originalartiklar (Peer-reviewed original articles)**

* C. Schüldt and P. Händel, "On implications of the ISO 3382 backward integration method for automated decay rate estimation," *Journal of the Audio Engineering Society*, vol. 63, no. 3, pp. 161–173, March 2015.

* C. Schüldt and P. Händel, "Decay Rate Estimators and Their Performance for Blind Reverberation Time Estimation," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, ISSN 2329-9290, vol. 22, nr 8, 1274–1284, 2014.

C. Schüldt, F. Lindstrom, and I. Claesson, "A Delay-Based Double-Talk Detector," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 6, pp. 1725–1733, 2012.

M. Borgh, M. Berggren, C. Schüldt. F. Lindstrom, and I. Claesson, "An Improved Adaptive Gain Equalizer for Noise Reduction with Low Speech Distortion," *EURASIP Journal on Audio, Speech, and Music Processing*, 2011:7, doi: 0.1186/1687-4722-2011-7, 2011.

M. Berggren, M. Borgh, C. Schüldt, F Lindstrom, I Claesson, ”Low-complexity network echo cancellation approach for systems equipped with external memory," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 8, pp. 2506–2515, 2010.

C. Schüldt, F Lindstrom, H Li, I Claesson, ”Adaptive filter length selection for acoustic echo cancellation," *Signal Processing*, vol. 89, no. 6, pp. 31185–1194, 2009.

C. Schüldt, F Lindstrom, I Claesson, "A low-complexity delayless selective subband adaptive filtering algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 5840–5850, 2008.

F. Lindstrom, C. Schüldt, I. Claesson, ”Efficient multichannel NLMS implementation for acoustic echo cancellation," *EURASIP Journal on Audio, Speech, and Music Processing*, 2007.

* I. Laptev, B. Caputo, C. Schüldt, and T. Lindeberg, ”Local velocity-adapted motion events for spatio-temporal recognition," *Computer Vision and Image Understanding*, vol. 108, no. 3, pp. 207–229, 2007.

F. Lindstrom, C. Schüldt, M Långström, and I Claesson, ”A method for reduced finite precision effects in parallel filtering echo cancellation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 9, pp. 2011–2018, 2007.

F. Lindstrom, C. Schüldt, and I. Claesson, "A Hybrid Acoustic Echo Canceller and Suppressor," *Signal Processing*, vol. 87, pp. 739–749, 2007.

2. **Fackgranskade konferensbidrag (Peer-reviewed conference contributions)**, vars resultat inte finns i andra publikationer.

   * C. Schüldt and P. Händel, "Noise robust integration for blind and non-blind reverberation time estimation," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Accepted 2015.

   * C. Schüldt and P. Händel, "Blind low-complexity estimation of reverberation time," in *Proceeding of IEEE Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, NY, USA, October 2013.

   J.O. Nilsson, C. Schüldt, and P. Händel, "Voice radio communication, pedestrian localization, and the tactical use of 3D audio," in *Proceedings of International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Montbeliard-Belfort, France, October 2013.

   M. Borgh, C. Schüldt, and I. Claesson, "Efficient asynchronous re-sampling implementation on a low-power fixed-point DSP," in *Proceedings of IEEE International Conference on Consumer Electronics (ICCE)*, pp. 378–379, Las Vegas, NV, USA, January 2013.

   C. Schüldt, F. Lindstrom, and I. Claesson, "Robust Low-Complexity Transfer Logic for Two-Path Echo Cancellation," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 173–176, Kyoto, Japan, March 2012.

   C. Schüldt, F. Lindstrom, and I. Claesson, "Evaluation of an Improved Deviation Measure for Two-Path Echo Cancellation," in *Proceedings of the 12th International Workshop on Acoustic Echo and Noise control (IWAENC)*, Tel Aviv, Israel, September 2010.

   C. Schüldt, F. Lindstrom, and I. Claesson, "An improved deviation measure for two-path echo cancellation," in *Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp. 305–308, Dallas, TX, USA, March 2010.

   C. Schüldt, F. Lindstrom, and I. Claesson "A Distortion Reducing Subband Limiter Implementation for Conference Phones," in *Proceedings of IEEE International Conference on Consumer Electronics*, Las Vegas, NV, USA, January 2008.

   C. Schüldt, F Lindstrom, and I. Claesson, "A Combined Implementation of Echo Suppression, Noise Reduction and Comfort Noise in a Speaker Phone Application," in *Proceedings of IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, USA, January 2007.

3. **Monografier (Monographs)**
4. **Forskningsöversiktsartiklar (Research review articles)**
5. **Böcker och bokkapitel (Books and book chapters)**
6. **Patent (Patents)**

   C. Schüldt and F. Lindstrom, "Method and device for microphone selection," Sweden Application Serial No. 1150031-1, filed on January 19, 2011. Patent pending.

   F. Lindstrom, C. Schüldt and I. Claesson, "Device and method for controlling damping of residual echo," Sweden Application Serial No. 0901012-5, filed on July 20, 2009. PCT Application Serial No. PCT/SE2010/050676, filed on June 17, 2010. U.S. Application Serial No. 13/384554, filed on January 17, 2012.

7. **Egenutvecklade allmänt tillgängliga datorprogram eller databaser (Open access computer programs or databases you have developed)**
8. **Populärvetenskapliga artiklar/presentationer (Popular science articles/presentations)**

CV

Name: Christian Schüldt	**Doctorial degree:** 2012-12-11
Birthdate: 19780816	**Academic title:** Doktor
Gender: Male	**Employer:** Limes Audio

Research education

Dissertation title (swe)

Algoritmer med låg beräkningskomplexitet för ekosläckning i konferenssystem

Dissertation title (en)

Low-Complexity Algorithms for Echo Cancellation in Audio Conferencing Systems

Organisation	Unit	Supervisor
Blekinge Tekniska Högskola, Sweden Sweden - Higher education Institutes		Ingvar Claesson

Subject doctors degree	ISSN/ISBN-number	Date doctoral exam
20205. Signalbehandling	978-91-7295-242-3	2012-12-11

Publications

Name: Christian Schüldt	**Doctorial degree:** 2012-12-11
Birthdate: 19780816	**Academic title:** Doktor
Gender: Male	**Employer:** Limes Audio

Schüldt, Christian has not added any publications to the application.

Register

Terms and conditions

The application must be signed by the applicant as well as the authorised representative of the administrating organisation. The representative is normally the department head of the institution where the research is to be conducted, but may in some instances be e.g. the vice-chancellor. This is specified in the call for proposals.

The signature*from the applicant*confirms that:

- the information in the application is correct and according to the instructions form the Swedish Research Council
- any additional professional activities or commercial ties have been reported to the administrating organisation, and that no conflicts have arisen that would conflict with good research practice
- that the necessary permits and approvals are in place at the start of the project e.g. regarding ethical review.

The signature *from the administrating organisation* confirms that:

- the research, employment and equipment indicated will be accommodated in the institution during the time, and to the extent, described in the application
- the institution approves the cost-estimate in the application
- the research is conducted according to Swedish legislation.

The above-mentioned points must have been discussed between the parties before the representative of the administrating organisation approves and signs the application.

Project out lines are not signed by the administrating organisation. The administrating organisation only sign the application if the project outline is accepted for step two.

Applications with an organisation as applicant is automatically signed when the application is registered.