



Audio Engineering Society Conference Paper 3

Presented at the International Conference on
Audio Education
2023 September 6–8, Hasselt, Belgium

This paper was peer-reviewed as a complete manuscript for presentation at this conference. This paper is available in the AES E-Library (<http://www.aes.org/e-lib>), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Hybrid teaching AV design implementation for music lectures in higher education

Iker Ormazabal¹, Rodrigo Sanchez-Pizani^{1,2}, and Christopher Robins¹

¹King's College London

²London Southbank University

Correspondence should be addressed to Iker Ormazabal (iker.ormazabal@kcl.ac.uk)

ABSTRACT

In this paper a scalable and adaptable solution for the AV design of a hybrid music teaching space is proposed. With the goal of making the AV not interfere with the delivery of lectures from a technical perspective, the solution provides a creative use of audio DSP processors in order to achieve this. The main factors that make the university's standard HyFlex AV design not suitable for music education are identified, and a design method based on separate microphone subsets is suggested. These subsets can be adapted for the appropriate speech and musical sources. In addition, automatic sound source detection and switching are introduced to achieve the desired technical unobtrusiveness. Finally the results are analysed and compared to those of a standard HyFlex system.

1 Introduction

In 2020, King's College London introduced HyFlex teaching as a means to supplement online and face-to-face teaching and in response to Covid-19 restrictions. This enabled teaching to a mixed cohort of students (both online and on campus) [1]. The technical implementation of this novel way of teaching, as well as its deployment in King's College London (KCL), was well described by Sanchez-Pizani et. al. [2]

Since then, there has been an increasing demand for rooms that include hybrid learning capabilities, and HyFlex has been in use by numerous departments at KCL. Some of the implementations have their own particularities. Adapting to these has occasionally required some deviations from the original design.

The specific case presented in this paper, is one where the technical implementation of the original approach

posed a few problems. Music lectures differ from traditional lectures as they require the capture of both music and speech.

1.1 Related works

Since the Coronavirus pandemic, the demand for remote teaching and learning has increased. In addition, there is an interest to keep an online alternative education model [3], which has also given rise to publications dealing with this topic.

There is little mention of AV design specifics regarding remote teaching implementations. Perhaps there is still not enough tradition of developing internal AV designs and consultancy within educational settings. However, there is relevant literature addressing student and academic experiences using these systems.

Lassoued et. al. [4] found among a sample of 100 University Professors, that 64% were concerned about the lack of preparation to deal with distance learning. While none of the 300 students reported this concern, both groups, however, felt that lack of capabilities (devices, internet, apps...) and training were two of the main obstacles. Molina et. al. [5] highlighted the importance of such resources and services and found that the link between technology and academic management practices needed to be strengthened. Similar discrepancy between the concerns and experience of the use of the technology were reported previously in a study by Dommett, et. al. [6] in 2019; in this case the study referenced lecture capture, however, the underlying technical experience is very similar.

Focusing on music lecturers from different countries (England, Turkey and Slovenia), Sonsel [3] and Birsa et. al. [7] found that the most popular platform between music teachers is Zoom. Sonsel found that the majority of lecturers were satisfied or highly satisfied with the quality of available materials and technology, however, they considered the remote teaching platform to be insufficient or inconvenient. Additionally, most lecturers in the study agreed that their digital literacy improved but were not willing to receive technology-related seminars.

Omur and Sonsel [8] conducted another study among piano teachers, and found that they were not satisfied with remote teaching over existing infrastructures (Zoom, MS Teams, etc.). They also thought that the lessons provided during emergency remote teaching were not enough to understand the music etudes and their musical specifications. Improvement of the infrastructure was seen as the second most relevant suggestion in order to improve the quality of remote piano lessons.

In addition to improvements in sound, factors such as audio-video synchronization should also be considered. Liu [9] studied the limitations of distance video teaching, and applied an audio-triggered system to obtain real-time synchronization. There are recent developments from platforms such as Elk Live [10], Jamulus [11], and Zoom to solve this issue, but there is still a compromise on audio-quality transmission when doing so.

1.2 Purpose

The aim of this research is to explain the particular implementation of HyFlex for music use, which retains

and builds upon the original features of the original design (initially implemented by KCL in 2020) and the AV Design principles described in the KCL AV Standard [12].

2 Methods

The research mixes several approaches: quantitative, observational and experimental. The main variables in the system are the input stage microphones, the video-conferencing platform and the output stage speakers/headphones. Additionally, there are two main sound sources at the initial stage of the system: speech and music.

2.1 Room Surveys

An omnidirectional dodecahedron speaker was utilized to evaluate the C50 metric in accordance with ISO 3382 standards. Impulse response measurements were performed and the data was processed using Arta software. The C50 values were calculated from the impulse responses. Measurements were repeated at different locations for reliability.

2.2 Room evaluation

The room was evaluated using 12 microphone-receiver combinations with a dodecahedron speaker, a calibrated Earthworks M50 microphone and the frequency sweep method in accordance with British Standards. [13] [14]

Arta software and an NTI sound level meter were used for acoustic measurement and analysis, together with Python programming language.

2.3 Baseline system

Latest efforts in Hybrid learning at KCL have been oriented towards capturing and transmitting student participation as well as lecturers. Therefore, there has been a preference towards beam-forming microphones with variable/ steerable polar patterns, also reducing the need for lapel microphones and the extra considerations needed when using the latter, such as: battery charging, microphone positioning, microphone sharing, etc.

Beamforming microphones can be used to reduce the impact of reverberation on speech signals [15]. Multi-channel algorithms, such as the generalized sidelobe canceller (GSC) [16], have been developed to further

enhance the performance of beamforming microphones in reverberant environments.

During the original implementation, the faculty staff received informal feedback about the hyflex implementation and the selected microphones. This was positive for most cases, except for music.

Selecting devices with noise cancelling algorithms when aiming for a design that supports both music and speech, adds an important problem. The desired signal has to be discernible from the unwanted sound or background noise such as office background noise. Unfortunately, some methods of detecting noise, such as the ones described by [17], will also regard most music as noise, since the combinations of organising sound in music are determined by a multiplicity of instruments. Each with its frequency spectrum, harmonic structure or even with additional modulation or time-based effects. Those instruments can create unlimited types of combinations.

Microphone arrays with adjustable beam widths and defined zones have proven the best results, however, distance from the sound source in fixed installations is still an issue when nuances of a musical piece need to be captured.

2.4 Input stage

The microphone chosen to capture music or speech signals is one of the essential differentiating aspects. Microphone techniques are beyond the scope of the manuscript, but detail on the subject is well documented in the literature [18, 19, 20]. Furthermore, the aim is to develop a design methodology that will allow combining different subsystems at the input stage without introducing phase cancellation or additional noise issues. As it will be seen later, commonly used microphone techniques have the potential to be part of the system.

The primary goals for the original system were:

1. Good noise cancellation for speech, this normally includes traffic noises, and other sounds external to the classroom.
2. Speech intelligibility, assuming normal conversational and raised voice levels ($63\text{dBA} - 70\text{dBA SPL}$).

3. Remove or limit the reverberation. While this is important for speech, some reverberation is usually desired from music performances to give a more natural feel.
4. It is normally useful to limit the speech frequency response to that of the human voice, while for music the requirement is to capture the full audible spectrum.
5. Scalability. In order to implement this solution in classrooms, it needs to be affordable and adaptable to a variety of sound sources and spaces.

In order to fulfil all the requirements, at least two microphones (or microphone arrays) are necessary. The primary microphone(s) to capture speech and the secondary microphone(s) to capture music. Additionally, the second subsystem alone can be modified to adapt the system to speech plus various other playback devices or acoustic and electronic instruments. As seen later in table: 5, the modulation transfer function (MTF) was used to compare the capacity of the several subsystems of reproducing the source information.

2.5 Combining inputs

When combining beamforming microphones and a stereo pair of microphones for music, several problems were encountered. These included phase cancellation and chorus like effects in music content. Music content was also incorrectly identified by the noise cancelling algorithm as 'unwanted noise' leading to its partial removal from the signal.

As a result, the decision was to not mix the signals for simultaneous playback. Instead, a system that could discern by automated methods which of the two signals should be sent to the conferencing software was developed.

2.6 Videoconferencing technology

Further problems can arise from the dedicated software noise-cancellation and audio compression bitrates of video conferencing software, which are normally optimized for speech intelligibility. There are currently platforms such as Audiomovers [21] and Sessionwire [22] that can overcome this, allowing for high bitrate and lossless codecs to be used in conjunction with what is presented here. However, not all of these platforms

capture and transmit audio and video simultaneously. Therefore, it needs to be considered that using these platforms could potentially introduce an extra step. At the time of writing, these music aimed platforms, alongside two of the most common found in higher education institutions, offer the bitrates and sample rates shown in table 1 at their higher quality settings.

Table 1: Audio codec bitrate per platform

Platform	Sample rate	Bitrate
Zoom	48 KHz	192 Kbps
MS Teams	32 KHz	128 Kbps
Audiomovers	96 KHz	6.15 Mbps
Sessionwire	48 KHz	400 Kbps

Given the institution’s preferences for MS Teams and Zoom, further research will focus in using these. Due to time constraints, this analysis is still not included in the research and it is planned for future research on the subject.

2.7 Output Stage

The output stage at the far end of the videoconferencing call cannot be controlled. Students are free to use the speakers or headphones that they have available. As such, there is a focus in reducing the number of variables at the input and transmission stage before further subjective evaluation is carried out.

3 Results

3.1 Room Surveys

Table: 2 shows the Rt30 as calculated for 35 classrooms at KCL with a Tmf of 0.7, and table: 3 shows the Rt30 as calculated for 10 listed classrooms at KCL with a Tmf of 0.9.

Figure: 1 shows the effect of augmenting the distance between sound source and microphone on the C50 clarity index. The graphic shows combined measurements for 10 rooms. It can be seen that despite the close proximity, in some cases the clarity index still provides poor measurements due to other factors such as noise or reverberation. Some rooms show late reflections (those arriving after 50ms) which are stronger than the early reflections (negative C50 values). It can also be

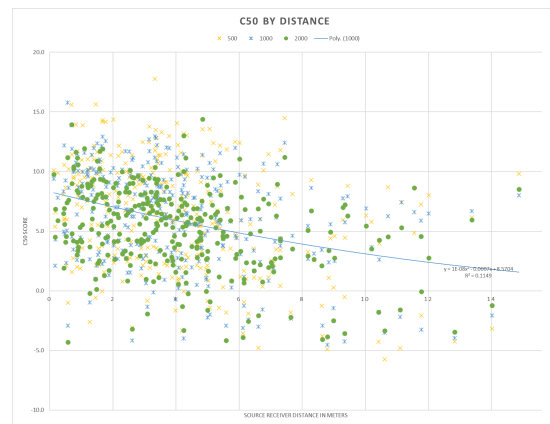


Fig. 1: C50 clarity index measurements in 10 listed rooms, several positions per room, 3 frequencies per measurement.

seen that generally, augmenting the distance from the source has a detrimental effect on the clarity.

Figure: 2 shows the early decay time (EDT) measurements in the same 10 rooms, and the effect that augmenting the distance between sound source and microphone has on the EDT measurements. Similar conclusions can be drawn from here: some long EDT times are measured despite the short distances from the source, and augmenting the distance from the source has a detrimental effect on the EDT measurements.

3.2 Room Characterisation

The room where this system has been implemented is normally used by the music department and 3 types of signal are usually present during teaching:

1. Speech. This can happen anywhere in the room, although most of it will originate at the lectern position.
2. Baby grand piano, harpsichord and other acoustic instruments. These are normally in fixed positions in the room.
3. Playback from Audio sources such as a CD Player, BluRay player... (using the front of house speakers.)

Table 2: Average Rt 30 per octave band for 35 rooms.

Freq. (Hz.)	63	125	250	500	1000	2000	4000	8000
Rt30 (s)	1.4	0.7	0.7	0.6	0.6	0.7	0.7	0.7

Table 3: Average Rt 30 per octave band for 10 listed rooms.

Freq. (Hz.)	63	125	250	500	1000	2000	4000	8000
Rt30 (s)	2.5	0.8	0.8	0.9	0.9	0.9	0.9	0.7

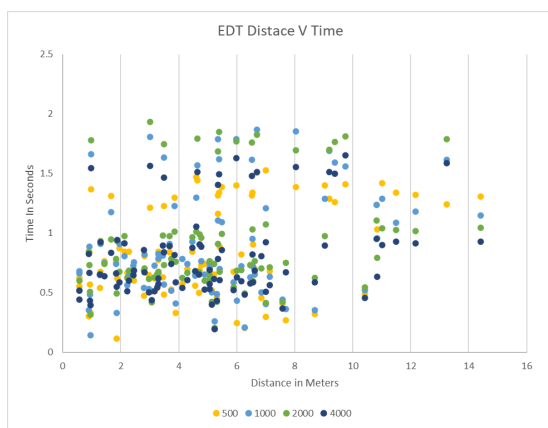
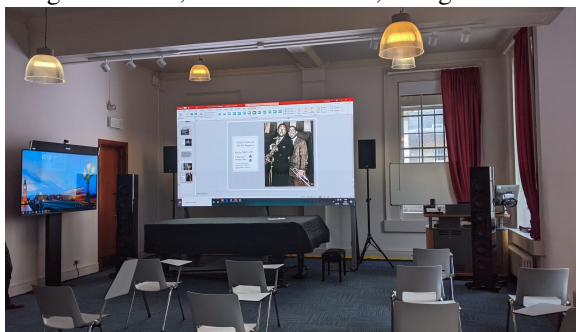


Fig. 2: Early Decay Time measurements in 10 listed rooms, several positions per room, three frequencies per measurement.

3.3 Room Dimensions

Length=11.46m, width= 7.62m, height= 4.26m.



3.4 Room Acoustics

The room’s acoustic evaluation can be observed in Table: 4 which displays the result of the parameters, measured per frequency as averages for all positions.

Table: 5 shows the MTF as calculated as the arithmetical average of octave bands for 500 Hz, 1KHz, 2KHz and 4KHz.

3.5 Frequency response and sound pressure level

Fig:3 shows: spectrogram (fig:3a) and waveform (fig:3b) for a condenser microphone capturing speech, followed by spectrogram (fig:3c) and waveform (fig:3d) for a beamforming microphone capturing speech, spectrogram (fig:3e) and waveform (fig:3f) for a condenser microphone capturing music and lastly, spectrogram (fig:3g) and waveform (fig:3h) for a beamforming microphone capturing music.

4 Discussion

4.1 Design

As mentioned earlier, the proposed solution is based on automatically alternating between two subsets of microphones. It uses a DSP design with a simplified structure such as the one in Fig. 4 and combines beamforming microphones for speech, a stereo condenser mic pair for musical examples, and automatic switching between both. This allows for the solution to be scalable and adaptable to multiple scenarios by changing or adding additional subsets with minimal modifications to the DSP workflow and signal path. It also includes inputs for program audio and can be used for electric and acoustic instruments.

Additionally, this solution aims to make the AV transparent to lecturers, and doesn’t require an operator to be present. Neither does it require that the lecturer interacts with the technology in a different way.

There are 4 microphone inputs which are part of 2 subsets.

Table 4: Acoustic parameters for the room.

Frequency (Hz)	Average	T30 (s)	T20 (s)	T10 (s)	EDT (s)	C80 (dB)	C50 (dB)	D50(%)
63	1.4	1.4	1.2	1.3	1.0	-1.6	-4.2	40.0
125	1.2	1.2	1.2	1.1	2.0	1.7	-1.3	86.0
250	1.0	1.1	1.0	1.0	3.0	2.7	-0.8	73.0
500	0.9	0.9	0.9	0.8	5.0	3.5	0.5	58.0
1000	0.8	0.8	0.8	0.8	5.0	5.0	1.9	55.0
2000	0.7	0.7	0.7	0.7	6.0	5.6	2.1	49.0
4000	0.6	0.6	0.6	0.6	8.0	6.5	2.9	39.0
8000	1.4	1.5	1.6	1.6	-2	8.5	4.7	426.0

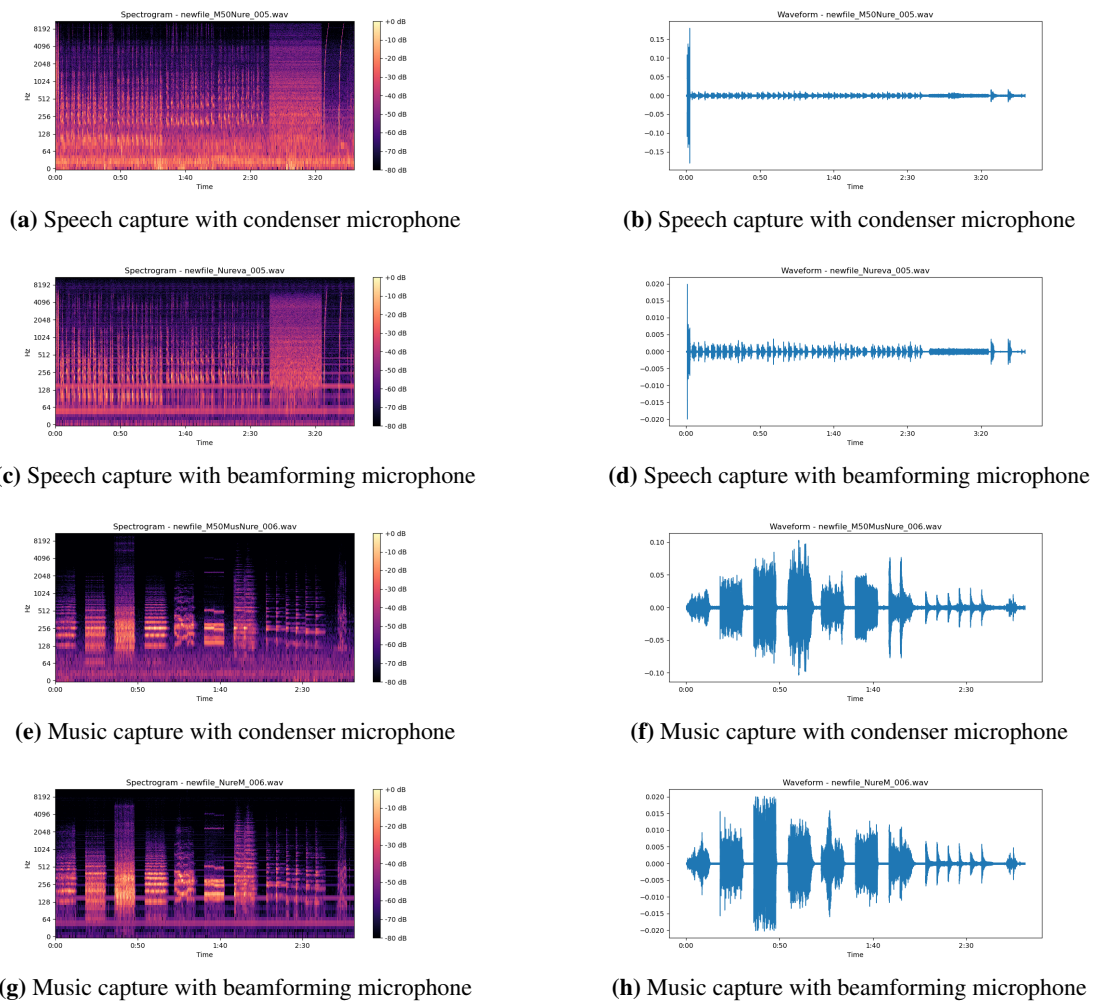


Fig. 3: Spectrograms and waveforms for the audio files.

1. A beamforming microphone with built in noise cancellation .
2. A stereo pair of condenser microphones dedicated to capturing music .

Table 5: Reverberation time and distance values for different microphone and loudspeaker combinations.

Combination	MTF (s)	Distance (m)
S1-R1	0.86	2.33
S2-R1	0.87	3.98
S1-R2	0.85	3.67
S2-R2	0.84	3.28
S1-R3	0.84	7.25
S2-R3	0.85	7.80
S1-R4	0.86	6.84
S2-R4	0.85	8.00
S1-R5	0.83	7.19
S2-R5	0.84	8.97
S1-R6	0.84	4.73
S2-R6	0.86	6.76

3. A piezoelectric transducer that will be used to trigger logic signals activated via a noise-gate .

Set number one is formed by the beamforming microphone array 1. Set number two is formed by the stereo pair 2 and the trigger 3.

The beamforming microphone is connected via USB to the DSP hardware. This could also have been an audio over IP microphone with noise cancelling technology. Its purpose is to capture the speech from both the lecturer and the students. It also improves speech intelligibility with built in speech recognition and noise cancellation algorithms.

The gate connected to the piezoelectric transducer, will in turn, control which microphones will be opened or closed via logic signals. This process is done in the audio DSP processor.

The piezoelectric transducer is placed on a resonant surface of the musical instrument that is being captured (a piano in this example). The transducer captures the sound waves travelling through a dense medium such as the wood of the piano, and produces a low voltage signal. It won't capture unwanted noise from the room, as it captures vibration from the surface almost exclusively. Sounds produced by playing the musical instrument are captured with enough signal to noise ratio by the piezoelectric transducer. To avoid having issues due to the very low level output from the

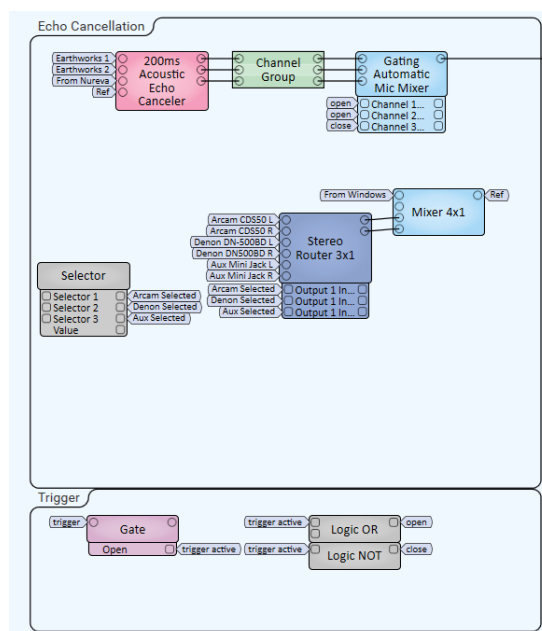


Fig. 4: DSP program. Trigger logic at the bottom, gating mixer receives logic at the top

transducer, a balanced wiring method or a DI interface should be used.

Before reaching the stage where the logic takes place, the microphone signals (except the piezoelectric transducer used as the trigger) as well as a reference signal, enter an acoustic echo canceller that will cancel echoes from the far end. Therefore, the far end signal being played back in the room will be discarded when it is captured by the room mics. Program audio from the near end shall also be part of the reference signal in the acoustic echo canceller, as suggested by the developer of the audio DSP processor [23].

After the echo cancelling stage, sound processing can be done (EQ, compression, etc.) to enhance the signal. The individual microphone signals will then reach the gating automatic mixer, which in this case is not activated by sound levels. Instead, channels will be allowed to open or close depending on logic signals that will be triggered by the piezoelectric transducer and a gate.

When the amplitude of the piezoelectric transducer reaches the desired threshold on the gate (Fig. 5), the “trigger active” signal will be set to a “high” state. Two

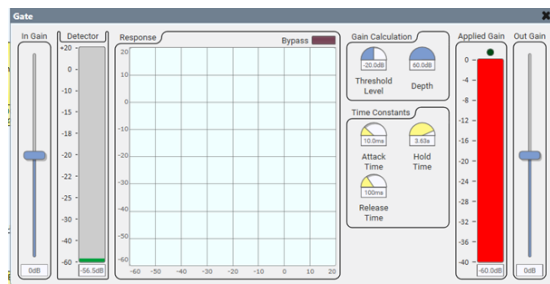


Fig. 5: DSP program gate interface.

subsequent logic gates will output the trigger output and its inversion. Two logic gates are used instead of just the “NOT” gate to ensure that both outputs will be produced at the same time by the audio DSP processor. When the signal goes below the threshold level, the logic signals will revert.

By routing these signals to the automatic gating mixer, the desired control can be obtained: the channels with the stereo microphone pair as an input will open when the logic signal is high. At the same time, the inverse of the logic signal will close the beamforming microphone channel.

4.1.1 Override

In this implementation of the design, buttons that allow users to manually choose the microphone subset were also installed. These are contact closure devices that connect to the audio DSP processor’s GPIO pins. By modifying the initial logic in charge of opening and closing the mics, auto and manual operation modes of the system can be included. Furthermore, these modes can be presented in the room control system UCI of choice by making the controls available to the room control system.

4.2 Analysis

A Python script with the Librosa library was used to load an audio file and compute its MFCC (Mel Frequency Cepstral Coefficients) features and the cosine similarity. The cosine similarity is a measure of the resemblance between two vectors, where a value of 1 indicates perfect similarity and a value of 0 indicates no similarity.

The results seemed to show discrepancy with the subjective evaluation done by the research team. A new

	Original	Omni	Beam
Original	-	0.994275	0.994463
Omni	0.994275	-	0.998502
Beam	0.994463	0.998502	-

Table 6: Similarity scores between audio files for music samples.

python script to compute the chroma feature to represent the pitch in the audio signal was used showing the results in Table: 7.

	Original	Omni	Beam
Original	-	0.994351	0.964761
Omni	0.994351	-	0.943037
Beam	0.964761	0.943037	-

Table 7: Cosine similarity of the chroma mean for music sample.

Looking at the results from performing the cosine similarity of the chroma mean, it can be seen that the beamforming microphone performs worse than the condenser microphone. Therefore it can be deduced that the beamforming microphone is less capable of capturing pitch information. It is worth noting the fact that after listening to the recordings, one would expect for the dissimilarity measurement between both to be greater. This underlines the impact on the accuracy of pitch and harmonic structure produced by the artifacts resulting from the noise cancelling processes. Despite not altering the fundamental structure of the musical content, these light variations in the harmonic content should be avoided in order to produce an accurate representation of the performed pieces.

Additionally, beamforming microphones used in this test (and in KCL in general), are installation microphones in fixed locations inside the rooms. One of the main reasons for this, is that the configuration on these microphones is dependant on where they are located in a room.

4.3 Signal to noise ratio

During the tests with a grand piano, the signal to noise ratio has not been a problem despite the high preamp gain needed in the piezoelectric transducer. Modifications to the frequency response and the sound pressure

level response of the piezoelectric transducer can provide different frequency responses by covering it with rubber, silicone or other materials either partially or entirely.

4.4 Further testing

Further tests with other beamforming microphones, have given different results. Focusing on available APIs, these are specially useful when they provide access to noise reduction, automatic gain compensation and automatic echo cancellation features. Allowing the trigger to activate and deactivate these features when music is being played, without introducing a second microphone subset.

4.5 Expansion upon the concept

The same design methodology can be applied to electric instruments. With the help of direct injection boxes, electric or electronic instruments can be routed to the DSP processor.

The logic in charge of opening and closing the automatic mixer channels can be expanded to accommodate more players and a preference system can be established by extending the DSP logic. There is a potential to use a wide range of microphone techniques and models without introducing additional noise or undesired artifacts.

4.6 Limitations

It would be necessary to modify this solution in order to accommodate the sung voice, for example when a lecturer is also a singer. In such a case, a simpler design is possible, and a dedicated microphone for voice would suffice for this particular application.

This approach doesn't tackle audio/video synchronization problems with current conferencing platforms, which is one of the main complains from user feedback.

Further research needs formal subjective-listening tests which are planned for the next stages of the research.

5 Summary

The article presents an AV design methodology for a hybrid music teaching space. The proposed solution uses a creative use of audio DSP processors to minimise technical obstructions and automatic sound source detection and switching. The solution is based on using two subsets of microphones, one for speech and the other for music. This allows for the solution to be scalable and adaptable to multiple scenarios by changing or adding additional subsets, with minimal modifications to the DSP workflow and signal path. Additionally, this solution aims to make the AV transparent to lecturers, and doesn't require an operator to be present or that the lecturer interacts with the technology in a different way. The proposed solution combines beamforming microphones for speech, a stereo condenser mic pair for musical examples, and automatic switching between both options. It also includes inputs for program audio and can be used for electric and acoustic instruments.

The paper shows that having two separate microphone subsystems such as the ones described, together with a detection method, provide the means to capturing disparate sources such as speech and music without one working in detriment of the other.

References

- [1] Detyna, M., Sanchez-Pizani, R., Giampietro, V., Dommett, E. J., and Dyer, K., "Hybrid flexible (HyFlex) teaching and learning: climbing the mountain of implementation challenges for synchronous online and face-to-face seminars during a pandemic," *Learning Environ. Res.*, 26(1), pp. 145–159, 2023, doi:10.1007/s10984-022-09408-y.
- [2] Sanchez-Pizani, R., Detyna, M., Dance, S., and Agustina, L. G., "Hybrid flexible (HyFlex) seminar delivery - A technical overview of the implementation," *Building and Environment*, 216, 2022, doi:10.1016/j.buildenv.2022.109001.
- [3] Sonsel, O. B., Kayaarslan, B., Güzeloğlu, G. B. T., and Onat, Y. A., "Emergency Remote Teaching During the COVID-19 Pandemic: The Case of Music Teachers in Turkey and England," *International Journal of Education & Literacy Studies*, 10(2), pp. 57–64, 2022, doi:10.7575/aiac.ijels.v.10n.2p.57.

- [4] Lassoued, Z., Alhendawi, M., and Bashitialshaaer, R., "An Exploratory Study of the Obstacles for Achieving Quality in Distance Learning during the COVID-19 Pandemic," in P. Edirisingha, editor, *Online and Distance Learning during Lock-down Times: COVID-19 Stories*, volume 2, chapter 1, pp. 1–14, MDPI, 2022.
- [5] Molina, A., Villegas, B., Ochoa, C. P., and Miranda, J., "Academic Continuity During the Covid-19 Global Health Emergency: Education 4.0 and the Flexible-Digital Model of Tecnologico de Monterrey University in Mexico Supporting Secondary Education," in F. J. M. F. M. Reimers, editor, *University and School Collaborations during a Pandemic Sustaining Educational Opportunity and Reinventing Education*, Knowledge Studies in Higher Education, chapter 10, pp. 149–164, Springer, 8 edition, 2022.
- [6] Dommett, E. J., Gardner, B., and van Tilburg, W., "Staff and student views of lecture capture: a qualitative study," *Int. J. Educ. Technol. High. Educ.*, 16(23), 2019, doi:<https://doi.org/10.1186/s41239-019-0153-2>.
- [7] Birsa, E., Kljun, M., and Kopačin, B., "ICT Usage for Cross-Curricular Connections in Music and Visual Arts during Emergency Remote Teaching in Slovenia," *Electronics*, 11(13), pp. 261–275, 2022, doi:[10.3390/electronics11132090](https://doi.org/10.3390/electronics11132090).
- [8] Omur, O. and Sonsel, O. B., "COVID-19 and Digital Literacy: Assessing Pre-service Music Teachers' Views on Piano Lessons Provided in Emergency Remote Teaching," *International Journal of Education & Literacy Studies*, 9(3), pp. 117–127, 2021, doi:[10.7575/aiac.ijels.v.9n.3p.117](https://doi.org/10.7575/aiac.ijels.v.9n.3p.117).
- [9] Liu, L., "Development and Application of Audio-triggered Courseware in Remote Music Teaching," *Int. J. Emerg. Technol. Learn.*, 10(6), pp. 37–41, 2015, doi:[10.3991/ijet.v10i6.4861](https://doi.org/10.3991/ijet.v10i6.4861).
- [10] ElkLive, "Elk Live," <https://help.elk.live/en/articles/7034218-elk-live-requirements> (Accessed 2023-06-27).
- [11] Jamulus, "Jamulus," <https://jamulus.io> (Accessed 2023-06-27).
- [12] Sanchez-Pizani, R., Ormazabal, I., Robins, C., and Grosvenor, D., "AV Standard," London: 2023.
- [13] BSI, "Acoustics — Measurement of Room Acoustic Parameters — BS EN ISO 3382-2-2008-," London: 2009.
- [14] BSI, "Acoustics: BS 18233 Application of New Measurement Methods in Building and Room Acoustics." London: 2006.
- [15] Hashemgeloogardi, S. and Braun, S., "Joint Beamforming and Reverberation Cancellation Using a Constrained Kalman Filter With Multichannel Linear Prediction," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 481–485, 2020, doi:[10.1109/ICASSP40776.2020.9053785](https://doi.org/10.1109/ICASSP40776.2020.9053785).
- [16] Priyanka, S. S. and Kumar, T. K., "Generalized Sidelobe Canceller Beamforming with Combined Postfilter and Sparse NMF for Speech Enhancement," *Fluct. Noise Lett.*, 20(02), p. 2150014, 2021, doi:[10.1142/S0219477521500140](https://doi.org/10.1142/S0219477521500140).
- [17] Hèansler, E. and Schmidt, G., *Acoustic echo and noise control : a practical approach*, Hoboken, New Jersey : Wiley-Interscience, 2004.
- [18] Davis, G. and Jones, R., *Sound Reinforcement Handbook*, Milwaukee: Hal Leonard, 1987.
- [19] Davis, D., Patronis, E., and brown, P., *Sound System Engineering*, Abingdon: Routledge, 4th ed edition, 2013.
- [20] McCarthy, B., *Sound Systems: Design and Optimization*, New York: Focal Press, 3 edition, 2016, ISBN 9780415731010.
- [21] Audiomovers, *ListenTo Quick Start User Guide*, Audiomovers LLC, 2022.
- [22] Sessionwire, "Sessionwire," <https://support.sessionwire.com/what-is-the-audio-quality-like-when-i-use-sessionwire> (Accessed 2023-06-17).
- [23] Lindemann, E., *Q-SYS Acoustic Echo Cancellation*, QSC LLC, 2016.