Defence Research and Development Canada | Recherche et développement pour la défense Canada

DEFENCE | DÉFENSE

Improving Situational Awareness in Noisy Environments

A helmet-based system for speech enhancement, hearing protection, and shock localization

Karl Wiklund
Simon Haykin
McMaster University

Andreas Freibert
Fan Zhang
Advanced Ultrasound Technologies Inc.

Prepared By:
Karl Wiklund
McMaster University
1280 Main Street West
Gilmour Hall - 306
Hamilton, ON
L8S 4L8
Department of Electrical and Computer Engineering, McMaster University
:

PWGSC Contract Number: PWGSC Contract Number: W7711-088143/A

CSA: Craig Burrell, Defence Scientist, (416) 635-2000 ext. 3051

Defence R&D Canada

Contract Report

DRDC Toronto CR 2010-048

March 2010

Canada

Improving Situational Awareness in Noisy Environments

A helmet-based system for speech enhancement, hearing protection, and shock localization

Karl Wiklund
Simon Haykin
McMaster University

Andreas Freibert
Fan Zhang
Advanced Ultrasound Technologies Inc.

Prepared By:
Karl Wiklund
McMaster University
1280 Main Street West
Gilmour Hall - 306
Hamilton, ON
L8S 4L8
Department of Electrical and Computer Engineering, McMaster University

PWGSC Contract Number: PWGSC Contract Number: W7711-088143/A

CSA: Craig Burrell, Defence Scientist, (416) 635-2000 ext. 3051

Defence R&D Canada – Toronto

Contract Report
DRDC Toronto CR 2010-048
March 2010

Principal Author

Original signed by Karl Wiklund

Karl Wiklund

Research Engineer

Approved by

Original signed by Craig Burrell

Craig Burrell

Defence Scientist, Individual Behaviour and Performance Section

Approved for release by

Original signed by Dr. Stergios Stergiopoulos

Dr. Stergios Stergiopoulos

Acting Chair, Knowledge and Information Management Committee

Acting Chief Scientisty

Abstract

Soldiers in the field work in an acoustically complex environment. Spoken words of particular interest may be embedded in other less important speech or environmental noise. Loud engine noise and impulsive sounds may also threaten not only the soldier's awareness of their surroundings, but their hearing as well. Just as importantly, some sounds (including impulsive events), are critical for situational awareness, which the soldier must hear or be made aware of. To address these competing problems, we have introduced the Smart Helmet system which combines hearing protection, speech enhancement and source localization. This report outlines the overall system design, presents lab-testable prototypes, and describes more integrated software models that may be used for further investigation. Areas needing further research are also described.

Résumé

Soldats dans le champ de travail dans un environnement acoustique complexe. Paroles d'un intérêt particulier peuvent être incorporés dans d'autres discours moins importantes ou le bruit dans l'environnement. Bruit de moteur fort et sons impulsifs peuvent menacer aussi non seulement le soldat conscience de leur environnement, mais aussi leur audience. Tout aussi important, certains sons (y compris les événements impulsifs), sont essentiels pour la conscience, qui le soldat doit entendre ou être mis au courant des. Pour résoudre ces problèmes de concurrents, nous avons introduit le système de casque dynamique qui associe la protection auditive, amélioration des discours et de la localisation de la source. Ce rapport décrit la conception globale du système, présente des prototypes de laboratoires-tests et décrit les modèles de logiciel plus intégrés qui peuvent être utilisées pour l'enquête supplémentaire. Les domaines nécessitant davantage de recherche sont également décrites.

This page intentionally left blank.

Executive summary

Improving Situational Awareness in Noisy Environments: A helmet-based system for speech enhancement, hearing protection, and shock localization

Karl Wiklund; Simon Haykin; Andreas Freibert; Fan Zhang; DRDC Toronto CR 2010-048; Defence R&D Canada – Toronto; March 2010.

Background: Soldiers in the field are often subjected to very loud noises which can both interfere with spoken communication, as well as cause long-term hearing loss. Hearing protection can create an additional hazard because soldiers lose a certain amount of situational awareness, including awareness of important environmental events such as gunshots or explosions, thereby placing them at further risk.

To combat these problems, we have proposed the Smart Helmet system, which combines three complementary technologies: a speech enhancement for improved verbal communication in the presence of noise, a microphone array for detection and localization of acoustic shocks such as those produced by gunfire, and hearing protection.

The Smart Helmet is designed to recognize the nature of the acoustic environment and switch between different processing modes to select the one that is most appropriate for a given environment. These processing modes include: active noise control, speech enhancement, and a special mode for acoustic shock suppression. In addition, the Smart Helmet includes an independent direction-finding system for rapidly localizing impulsive acoustic sources (such as gunshots).

Results: A demonstration system is provided in three parts. A hardware implementation of the core speech enhancement system has been built, along with the head-mounted system comprised of the microphones and amplifiers. In addition, a second hardware component implements the direction-finding system, which includes the helmet-mounted microphone array. The third component is a full software model that incorporates the main speech processing features of the Smart Helmet. This model is implemented in MATLAB and combines the basic processing modes and the mode-switching mechanisms.

Significance: Sample results show that the system is capable of performing its major functions, including mode-switching under a time-varying set of acoustic inputs. Noise suppression results show an SNR improvement of up to 10 dB when using the speech enhancement mode, and a suppression level of about 17 dB when in Active Noise Control mode (simulated only). The acoustic shock localization system is demonstrated on a hardware implementation consisting of a helmet-mounted array, an associated DSP, and a linear array of LEDs, used to indicate the source direction. Preliminary results are discussed, along with an assessment of future areas for improvement.

Future plans: Preliminary results indicate that the prototype system resulting from this contract vindicates the original concept. More thorough tests are required to assess the strengths and

weaknesses of the prototype system. Numerous technical objectives which could improve the performance of the various components of the Smart Helmet are presented to guide further development.

Sommaire

Improving Situational Awareness in Noisy Environments: A helmet-based system for speech enhancement, hearing protection, and shock localization

Dr. Karl Wiklund, Dr. Simon Haykin; Andreas Freibert; Fan Zhang; DRDC Toronto CR 2010-048; Defence R&D Canada – Toronto; March 2010.

Introduction ou contexte:

Soldats dans le champ sont souvent soumis à des bruits très forts qui peuvent interférer avec la communication parlée, aussi bien que causer la perte d'audition à long terme. Le niveau de bruit peut également créer un danger supplémentaire que soldats travaillant dans ces environnements sont également susceptibles de perdre un certain degré de conscience, ce en les plaçant à davantage de risques. Pour lutter contre ce problème, nous avons proposé le système casque dynamique, qui combine la protection auditive avec amélioration de discours avec logiciel capable de prendre des décisions de traitement intelligent de face à un environnement complexe et dynamique.

Ce système est conçu pour reconnaître la nature de l'environnement acoustique et basculer entre les modes de traitement différent pour sélectionner celui qui est le plus approprié pour un environnement donné. Ces modes de traitement comprennent : lutte active contre le bruit, amélioration du discours, ainsi qu'un mode spécial pour la suppression de choc acoustique. En outre, le casque dynamique est conçu pour dotés d'un système de direction-recherche pour localiser rapidement impulsifs sources acoustiques (comme gunshots) et donc agir comme une aide à la conscience du soldat.

Un système de démonstration est fourni en trois parties. Une mise en œuvre de matériels du système principaux discours amélioration a été construit, ainsi qu'avec le système de tête-monté composé des microphones et amplificateurs. En outre, un second composant matériel met en œuvre le système de direction-recherche, ce qui inclut la baie montés casque microphone. Le troisième volet est un modèle de logiciel complet qui intègre les fonctions de traitement de discours principal de la casque dynamique. Ce modèle est mis en œuvre dans MATLAB et combine les modes de traitement de base et les mécanismes de basculement de mode.

Résultats d'exemples montrent que le système est capable d'exécuter ses fonctions principales, y compris le basculement de mode dans le cadre d'un jeu variant dans le temps des intrants acoustiques. Résultats de la suppression du bruit montrent une amélioration du rapport signal/bruit de jusqu'à 10 dB lorsque vous utilisez le mode d'amélioration de la parole et un niveau de suppression d'environ 17 dB en mode ANC (simulée uniquement).

En plus de la protection auditive, une exigence fondamentale est la conscience. Au-delà de l'aspect du discours de sonore amélioration mentionnés précédemment, cela comprend la prise de conscience des événements environnement importants, tels que gunshots ou d'explosions. Dans un milieu bruyant et potentiellement chaotique, un soldat peut ne pas toujours être

immédiatement au courant d'un tel événement, ou au courant de sa direction. Pour aider à cela, un second système est ajouté à la casque dynamique, qui se compose d'une baie de tête-monté le microphone circulaire. L'objectif de ce système est de localiser immédiatement des sources sonores impulsifs et fournir le porteur avec une indication visuelle de son orientation d'origine. Ce système est démontré sur une mise en œuvre de matériels consistant en une baie montés casque, un DSP associé et une baie linéaire de témoins, utilisée pour indiquer la direction de la source. Résultats exemplaires sont abordées, avec une évaluation de futurs domaines.

Table of contents

List of Figures

List of Tables

This page intentionally left blank.

1 Problem Statement

Soldiers in the field may be required to work in noise-intensive environments in which hearing-protection apparatus must be worn. The combination of high noise levels and the obscuring effect of the hearing protection itself serve not only to make verbal communication between the soldiers difficult, but also to reduce a soldier's awareness of the environment. Commands and other verbal information may end up going unheard or misunderstood, and reaction to other vital information may be similarly impaired.

While Active Noise Control (ANC) devices such as NACRE [1] are useful for reducing the effects of extremely high noise levels, and thus protecting the wearer's hearing, there are limitations. In noisy environments below NACRE's threshold of operation, even when there is no risk to the wearer's hearing, the noise levels may still be sufficient to impair speech intelligibility [2]. In addition, highly dynamic noise sources such as speech and other environmental noise are difficult to filter out in any case [3]. As a result, there is a need to develop a new form of active noise-control, which is suitable for people working in dynamic, noise-intensive environments. Such a system should be designed to reduce potentially harmful noise levels and improve the efficacy of communication. In addition, it is also necessary to enhance or at least preserve the soldier's situational awareness in response to non-verbal acoustic events.

2 System Overview

2.1 Processing Overview

To address these issues, we have proposed the Smart Helmet system, which combines hearing protection with speech enhancement along with software capable of making intelligent processing decisions in the face of a complex and dynamic environment. Because this system is expected to operate in a wide variety of conditions, it must also be capable of employing different processing methods, suitable for the expected acoustic circumstances or environments. The nature of these environments mean that processing options must include speech enhancement, noise control, acoustic shock detection and protection, as well as the option to simply transmit the sound with no processing at all. As a result, not all of these processing modes can be engaged at once, and it is the job of the software to determine which processing mode should be employed at any given time.

The qualitative criteria for each processing mode can be summarized as follows:

1) Quiet: In this case, any processing beyond correcting for attenuation caused by the passive headgear may result in loss of situational awareness. As a result no processing other than basic amplification should be performed.

2) Noisy (but not harmful): The noise level is such that there is a loss of intelligibility in spoken communication as well as a loss in situational awareness. Of particular interest is the case when the noise is non-stationary, and cannot be filtered using conventional means (e.g. babble noise from crowds or other acoustically complex scenarios). In this case, some signal processing should be performed to enhance speech intelligibility.

3) Very Noisy: In this case, a loud sustained noise source is present. This both obscures spoken communication, as well as threatens to cause hearing damage. In this case, systems such as Active Noise Control (ANC) should be engaged, along side other methods for noise reduction (where appropriate).

4) Acoustic Shocks: Very loud, impulsive sounds cannot be meaningfully filtered out, and also threaten to cause hearing damage. In this case, the only thing to do is rely completely on the passive hearing protection by shutting down all audio transmission to the ears.

The quantitative aspects of these decisions are more difficult and depend directly on both the particular acoustic properties of the headset, as well as the safe thresholds as determined by audiologists. In the absence of a suitable acoustics lab, a MATLAB simulation of this system makes use of a generic parameter set that may be tuned as better information becomes available.

2.2 Smart Helmet Processing Model

The form of the Smart Helmet that was originally proposed was to be an all-hardware model that switched between NACRE's ANC headset, and another unit, the Fuzzy Cocktail Party Processor (FCPP) [3], depending on the current noise environment. The purpose of the ANC mode, of course, being to eliminate very loud, steady noises, while the FCPP unit was engaged in moderately noisy environments for the purpose of eliminating noise in general, and non-stationary noises (babble) in particular. However, owing to unforeseen legal issues, this implementation was not possible. In its place, a more tightly integrated system was proposed that would incorporate the various processing modes in a single device. Ultimately, this approach allowed for a greater degree of design flexibility, since the difficulty of switching between two dissimilar hardware units was eliminated. However, the time constraints of the current project deadline meant that a full hardware implementation was no longer feasible. In its place, a MATLAB model of the proposed system has been developed that includes the basic functionality of the Smart Helmet system.

The overall operation of the system is depicted in the flowchart of Figure 1. As can be seen, this system incorporates the main processing modes described in Section 2.1, as well as the specific decision points that relate to mode switching. For safety reasons, acoustic shock detection has been given priority over all other decision points.

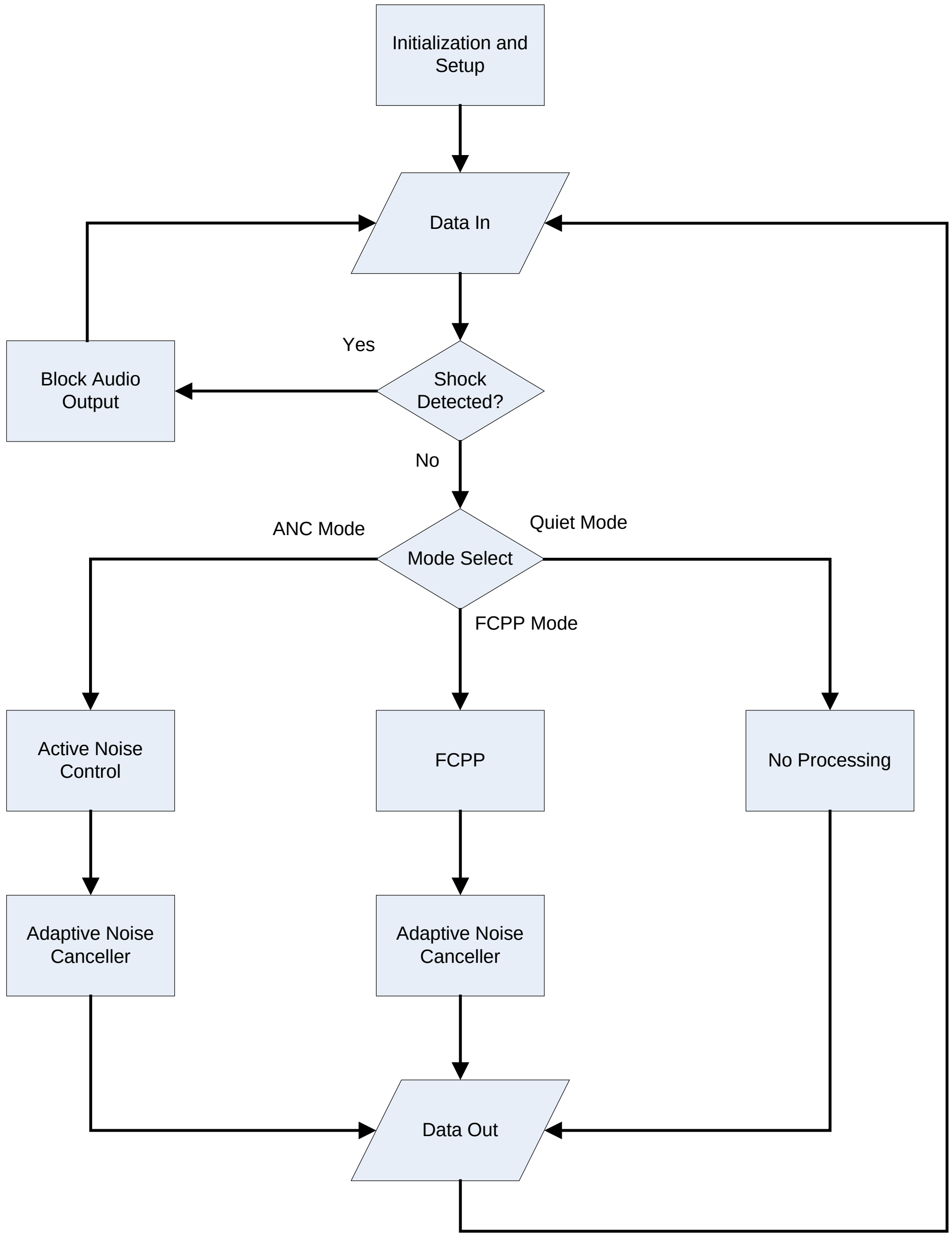

Figure 1 The basic execution path for the Smart Helmet system.

2.3 Physical Model

For the current Smart Helmet implementation, a six-microphone model is used, with three microphones on each side of the head. Of these microphones, two are external, and are arranged so that one is forward-facing and the other rearward-facing. The configuration is necessary to implement both the FCPP and the adaptive noise canceller. One of the microphones is also used as the reference sensor for the active noise controller. The remaining internal microphone is used as the error sensor, which is also part of the ANC system.

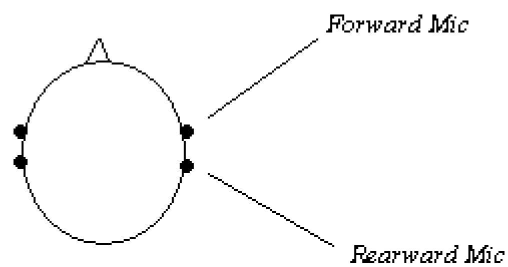

Figure 2 Placement of the external FCPP microphones (top view).

A simple, physical acoustic model of the headphone unit itself was also developed, which incorporates a primary acoustic path, P(z), as the transfer function between the outside surface and the wearer's ear, as well as a secondary path, S(z) [4], that models the path between the internal loudspeaker and the error microphone (placed close to the wearer's ear). This model is necessary for developing ANC algorithms, as well as for modelling the effects of acoustic leakage in different situations. However, since accurate physical information was not available, the primary path was modeled as a simple lowpass filter with a low delay and is shown below as Equation (1)

$$h(t) = \frac{1}{60}\left[\exp(-\frac{1}{65}(t-2))\right]\cdot u(t-2) \tag{1}$$

where $h(t)$ is the impulse response of the primary path, and $u(t)$ is the unit step function. The secondary path is modeled as a rapidly decaying impulse function, with some random noise added (see Figure 3). The noise is randomized for each simulation to ensure greater robustness, as the secondary path estimate is prone to modeling errors as well as some variation from user to user. In addition, while these models are broadly correct, they should be assessed using more accurate data.

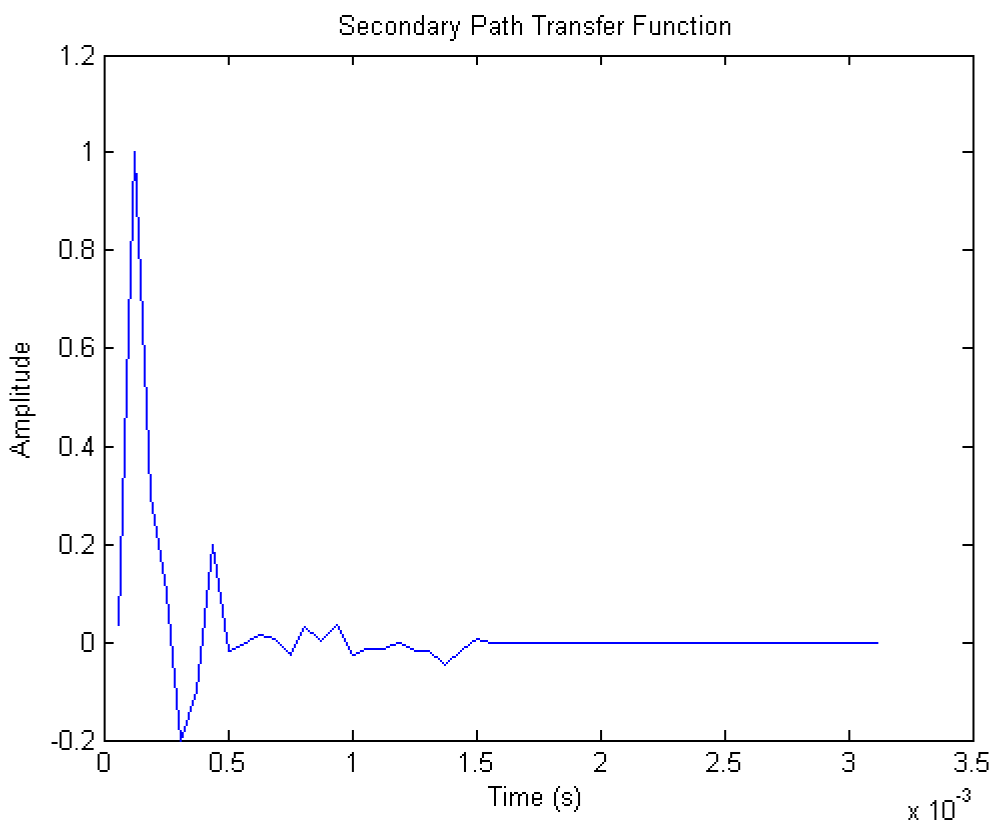

Figure 3 The secondary path transfer function reflects the fact that the loudspeaker and error microphone are placed close together. Some random variation from trial to trial is also included.

2.4 Shock Detection

Owing to the fact that in the field, loud acoustic shocks can happen at any time, and while the helmet is in any processing mode, a shock detection system that is separate from the mode switching block is necessary. The job of the detector is to detect sudden, loud shocks, and determine whether or not these shocks constitute a threat to the wearer's hearing. The current system operates independently of the direction-of-arrival estimator that has been developed by Advanced Ultrasound Technologies as an additional part of this contract (see Annexes D-F). This separation owes more to practical development issues than it does to any sort of actual necessity. Any future system should seek to integrate these functions.

The system we propose is similar to the speech onset detection system used in [3], except that the loudness comparisons are intra-frame instead of inter-frame. Basically speaking, the system looks for events such that for a frame *S* of length *M*, and for *L*<*M*,

$$(\max(|S(1:L)|) > \theta_1 \cdot \max(|S(L+1:M)|)) \wedge (\max(|S(1:L)|) > \theta_2) \tag{2}$$

Equation (2) means that if the maximum value of the *L* most recent samples is greater by some factor θ_1 than the preceding *M*-*L* samples, and the basic loudness is above some pre-defined threshold θ_2, then a shock event has been detected. In the current implementation, *L* is set equal to 60 samples, and the sampling rate is 20 kHz, meaning that the helmet can react to an impulsive source in about 3 ms. This is a tuneable parameter however. In addition, the system automatically reverts to its previous mode after 128 samples. To reduce the likelihood of transient effects, this comparison is carried out across all frequency bands; shock suppression mode triggers if the number of frequency bands satisfying Equation (2) exceeds a tuneable threshold value.

While in suppression mode, no sound is transmitted electronically, and any adaptive algorithms are suspended. The only sound that will be heard by the wearer is that which

is physically transmitted through the headphones. Figure 4 and Figure 5 for example, illustrate the level of suppression that is achievable using this method. Alternative approaches to the currently implemented practice of total suppression are also possible, and may include loudness compression or some other form of gain control.

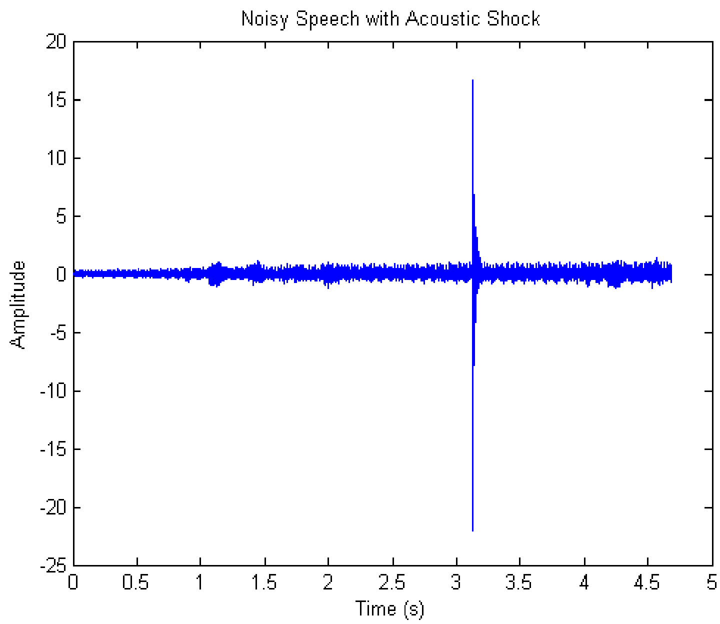

Figure 4 The speech signal shown here is corrupted by several noise sources, as well as a sudden impulsive event occurring at about 3 s

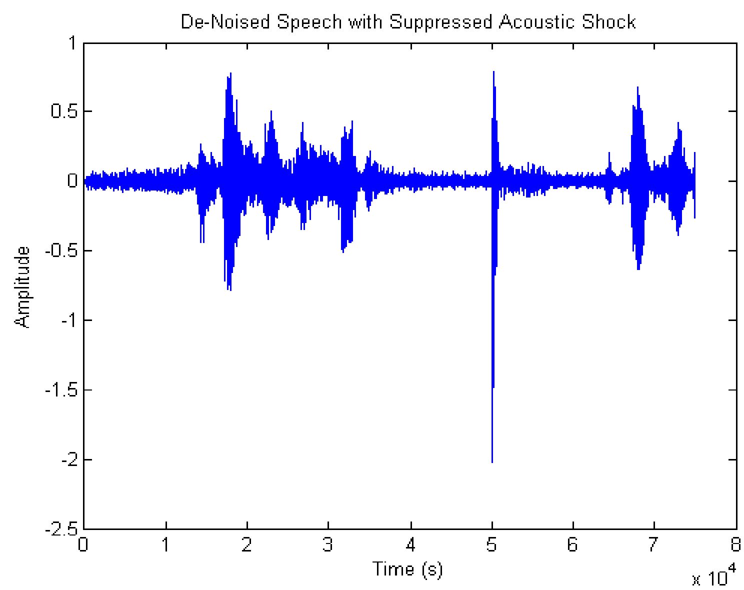

Figure 5 The loud impulsive source is partially suppressed, but still audible without affecting the de-noising capabilities of the system.

2.5 Mode Switching

In some possible scenarios, it may be that the wearer experiences an acoustic environment that drives the system to operate close to one of the mode decision boundaries. In this case, it is undesirable to have the system flip back and forth between processing modes due to small changes in volume, head position, and so on. At the same time however, it is not practical to run two modes in parallel and then sum them using some combination rule. This is especially true when switching between the FCPP and ANC modes, because of the fact that the ANC mode incorporates a feedback loop.

As a result, the system is run in a single mode at a time, and to mitigate the potential for mode flipping, a hysteresis-based switching mechanism is used. This ensures that when the system moves into a given processing mode, the input must move further back than the initial switch point if the previous mode is to be restored (see Figure 6). This system is applied to all

modes except for the shock suppression mode as it does not result from any sustained source, and must be entered and exited very rapidly.

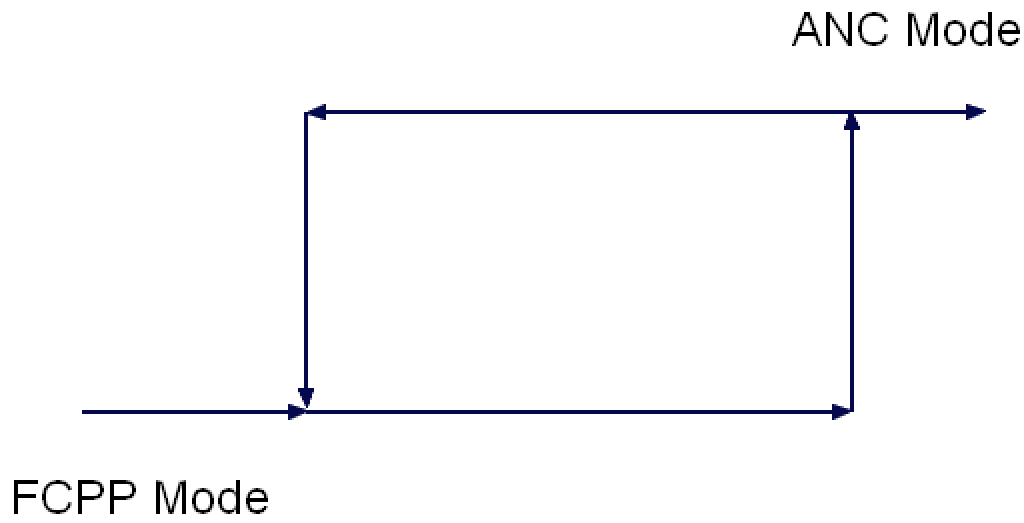

Figure 6 Mode switching using a hysteresis pattern. This approach is applied to the Quiet, FCPP, and ANC modes, with the Shock mode being excepted. Currently, the switching boundaries should be viewed as somewhat arbitrary pending accurate acoustic and physiological data. The switching decision criterion is based on the ambient noise power level.

In the current implementation, the decision boundaries are more or less arbitrary, as they are based on informal perceptual criteria. A more involved model should incorporate decision boundaries derived from both physical acoustics and physiological data.

2.6 Fuzzy Cocktail Party Processor

The fuzzy cocktail party processor is designed to be a noise reduction system suitable for coping with moderate levels of noise in a highly non-stationary acoustic environment. This system for example, is able to reduce the level of acoustic interference in multi-talker or “babble” environments, and thus enhance the intelligibility of face-to-face communication in otherwise difficult situations. When in the FCPP mode the system executes essentially the same algorithm described in [3], the outline of which is summarized in Figure 7.

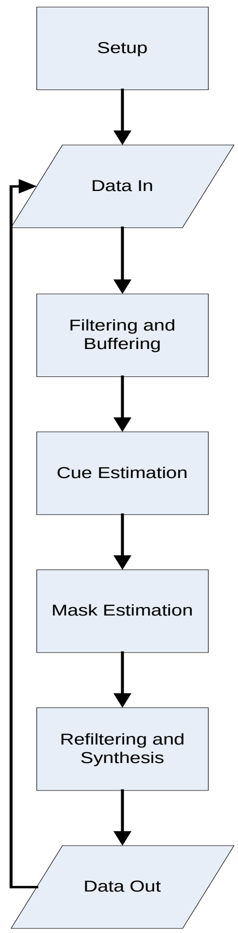

Figure 7 The execution flowchart for the basic FCPP algorithm.

In order to cope with the highly non-stationary acoustic scenarios that characterize many real acoustic environments, the FCPP is based a set of principles generally referred to as "Computational Auditory Scene Analysis" (CASA). In short, CASA-based systems attempt to make use of one or more of the methods by which the human auditory system is believed to separate and group various acoustic streams. To accomplish this task, the FCPP calculates approximations to several auditory cues using the windowed outputs of a set of fixed filterbanks. The cues are known to be used by the human auditory system, and can also be readily

implemented on a standard DSP. From these cues, the FCPP is then able to classify different time-frequency blocks as being either signal or interference, and thus filter them out.

For the FCPP, the signal / noise classification is not a binary choice; rather, the set of time-frequency gains are calculated based on the level of confidence the system has in the correct classification. The gains therefore, are real numbers taken from the range [0,1], and are assigned based on the truth values of a set of fuzzy logic operations whose inputs are the auditory cues described previously. Following the application of the gains, the signal is re-synthesized in the usual way.

The only significant change to the original FCPP algorithm described in [3], is the use of a bank of Princen-Bradley FIR [5] filters to perform the analysis and synthesis functions. This was necessary because of the fact that the phase delay properties of the original cochlear filter bank [6] proved to be incompatible with good performance for the active noise canceller. At the same time, the use, and swapping of different filterbanks was judged to be undesirable based on the computational complexity of such an arrangement. Therefore, a single filterbank structure was decided on that could meet the needs of both the FCPP and the ANC units.

For this purpose, the basic Princen-Bradley structure was altered so that it could approximate the frequency distribution of the original cochlear filters. To do this, the higher frequency bands of the system were simply summed so as to increase the bandwidth of the filters at the specific centre frequency. Unfortunately, this is not an ideal solution, since the use of FIR filters also increases the system's computational complexity. Preferably, some other filterbank structure should be sought out that provides an adequate frequency decomposition, while also possessing the low group-delay needed for the active noise control program.

The FCPP is also supplemented with an adaptive noise canceller configured as a post-filter. This allows for additional noise reduction in the case of engine sounds, a case for which

the FCPP is not well-optimized. The combination of the two algorithms, which combine stationary and non-stationary noise reduction seems to be a good one, as neither seems to compromise the other. The FCPP responds to fast changes, and can eliminate segments that are clearly dominated by noise, but it ultimately lacks the kind of frequency resolution possible with a linear filter. At the same time, the adaptive noise canceller cannot respond to quick-changing sounds like speech, and so only adapts to remove long-term, semi-stationary noise sources.

2.7 Active Noise Control and Active Speech Transmission

In very loud environments, some form of hearing protection must be worn. Of particular interest is the case when the noise source is dominated by low-frequency components. In such cases, the protection afforded by simply passive systems is inadequate and active noise control (ANC) must be used. For this study, it was decided that a hybrid feed-forward, feedback version of the FXLMS filter [4] would be used. The model chosen follows that used in [7] by Ray et al. As in that study, a fixed feedback filter is used in conjunction with an adaptive feed-forward section. The use of the fixed feedback filter is a compromise that offers improvements in terms of performance and stability over the basic feed-forward FXLMS, while not incurring the cost of using a fully adaptive hybrid system.

Because the primary purpose of the active noise control system is the elimination of low-frequency (below about 1000 Hz) noise, it is possible to design the system so that only noise components below this frequency limit are suppressed. Frequency bands above this limit can be processed separately, using methods that are not subject to the same constraints as active noise control algorithms (see Figure 8). Such processing can be as simple as straight-through transmission with automatic gain control (termed “Active Speech Transmission” (AST) by Oinonen [8]), or other noise reduction algorithms. Situational awareness is therefore retained,

since the high-frequency components of environmental sounds can still be transmitted to the wearer. Spoken communication is also enhanced, as the speech signal still retains significant intelligibility despite the loss of the low-frequency components.

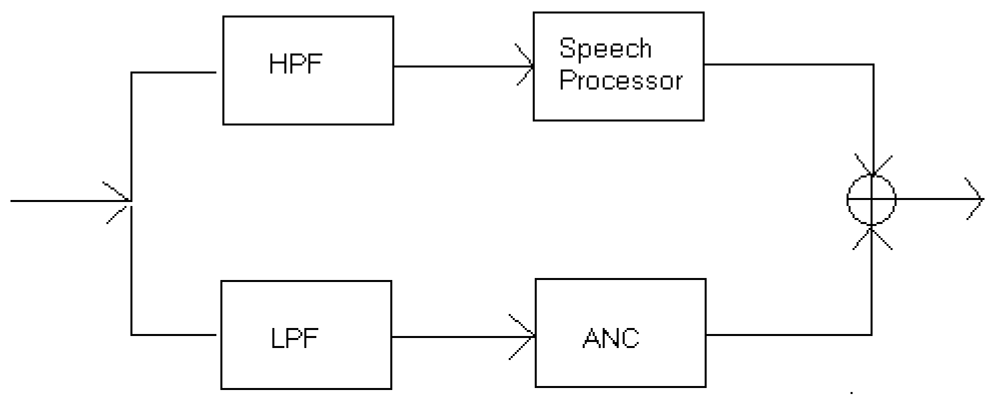

Figure 8 Hearing protector design using separate processing for the high and low frequency bands.

In spite of the potential design flexibility of the AST section, there are still some important constraints that must be met. First and foremost is the need to keep the computational complexity as low as possible. Obviously, one approach to meeting this goal is to use a processing algorithm that is also computationally efficient. Another is to ensure that the proposed algorithm is structured so as to be readily compatible both with the ANC system itself, as well as with the other processing modes. In particular, this means avoiding some overhead and scheduling issues by restricting the AST algorithm to one that can operate on a sample-by-sample basis, as opposed to one that is block-based. In addition, it would also be preferable to be able to use the same filter structure throughout. That is, no significant reconfiguration of the analysis or synthesis filters should be required.

To this end, it was decided that the basic adaptive noise cancellation algorithm [9] should be the approach of choice for the purposes of preliminary testing, as it meets all of the criteria described above. In addition, this algorithm makes good use of the proposed microphone

structure used by the FCPP, as the backward-facing or forwards-facing microphones can be used as the reference and error microphones respectively.

The FCPP was also considered, but its computational complexity, and most importantly, its block-based nature make it difficult to directly incorporate. Aside from the overhead associated with the many context switches required, some routines which have been highly optimized for a specific processor, may not be safely interruptible (the FFT routine for Texas Instrument's C6713 processor is an example of this [10]). However, with the increasing presence of multicore DSPs on the market, it is likely that these issues may in time be eliminated or at least mitigated.

2.8 Sample Results

Using the FCPP mode alone, and comparing the Smart Helmet implementation against the original, we find that for a simple acoustic scenario with random spatially distributed interferers and an input SNR of 0 dB, the following average results are obtained after twenty randomized trials:

Table 1 Sample Output SNRs for Speech Scenarios

Type	Output SNR	Variance
FCPP	9.45	0.086
Smart Helmet FCPP	9.78	0.062

In Table 1, it should be noted that although the original FCPP algorithm seems to perform marginally poorer than the revised version, informal listening tests indicate that it is still the

preferred algorithm when it is practical to use. Owing to problems with the testing facilities at the DRDC acoustics lab, no similar tests have yet been conducted for the hardware version.

With respect to the system's ANC capabilities, preliminary work has so far verified both the mode-switching and de-noising capabilities of the proposed system. The tests shown below, for example, were carried out using a speech signal contaminated by other speech signals, as well as a loud engine noise of time-varying intensity (linearly increasing, flat, then decreasing). In these tests, the system begins in the quiet mode, rapidly switches to the FCPP mode, and as the power of engine noise increases, switches to the ANC/AST mode. At the loudest point, the input SNR is -10 dB. For this case, the output SNR can only be measured for frequencies above 1 kHz, due to the wideband suppression by the ANC unit. For this set of measurements, the average output SNR was measured to be 8 dB over the range of applicable frequencies.

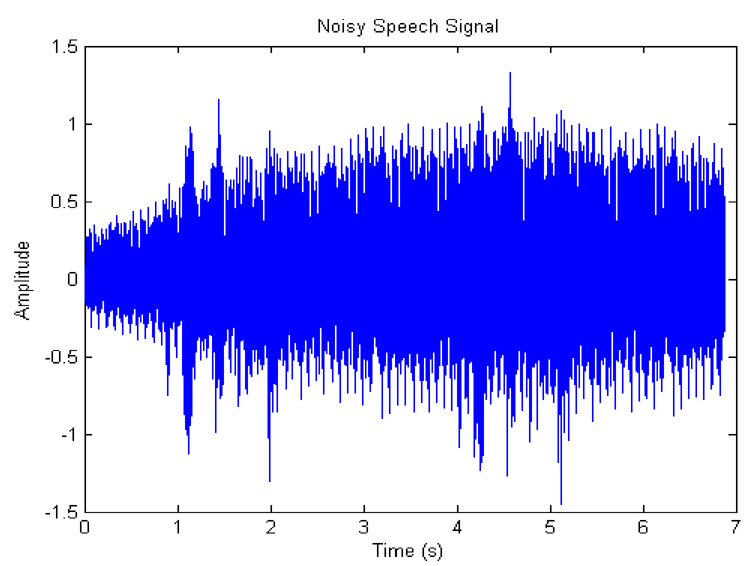

Figure 9 A sample window of the first 7 seconds of the noisy speech signal. The input SNR is -8 dB.

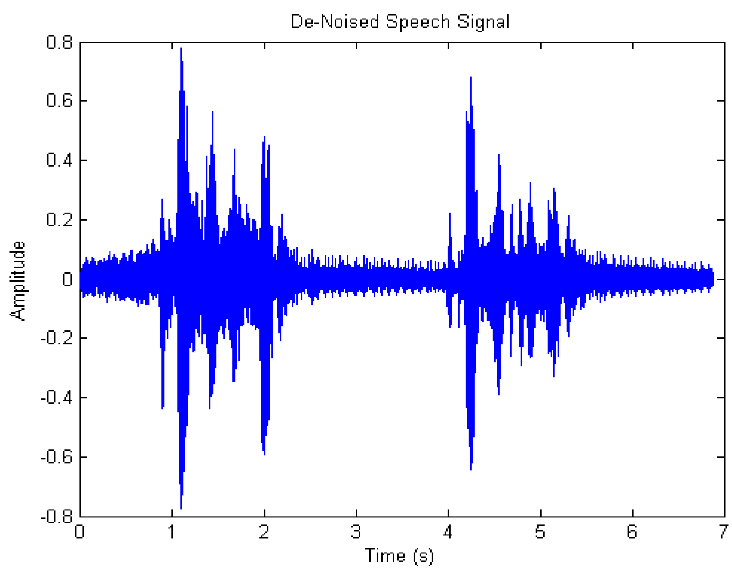

Figure 10 The de-noised speech signal. The average segmental output SNR for the ANC section is 8.2 dB, although that is only for frequencies above 1000 Hz.

3 Hardware Implementation of the FCPP

For the cocktail party processor to be a credible technology, and thus a practical element of the Smart Helmet system, it was essential that it be shown to work on a real-time DSP platform similar to what might eventually be used in the field. Only by doing this, could essential information about the timing constraints and hardware issues be fully understood.

3.1 Hardware Platform

The platform chosen for demonstrating the real-time capabilities of the FCPP was the SHARC ADSP-21369 EZ-KIT LITE demonstration board, set to run at 392 MHz. The DSP itself has an on-board SRAM of 2 Mbits, large enough to keep the entire program and most of the data in cache. The board also features an external flash memory, which is used for the stand-alone boot mode, and an external complement of 4 banks of 4 M x 32 bit SDRAM chips. This board was chosen primarily for the speed of the processor, although it also allowed for the possibility of creating a simplified implementation by removing the need to build the hardware portions of the I/O interface. At the same time, the board is also extendable in that additional peripherals may be added in order to create more involved implementations of the system. However, only the basic two-microphone implementation is available at this time. Setup instructions for both of these options are described in Annex A.

3.2 Timing and Computational Complexity

Both the two- and the four-microphone version of the FCPP can be run in real-time on the chosen DSP. The present implementation can allow for either 32 audio channels for the former system, and 31 channels for the latter. In both cases, the CPU is operating close to its

maximum capacity, and cannot accommodate any other processes of significant complexity. However, the possibility of using a multi-rate approach to cue estimation has been identified as a potential area for significant computational savings [11]. This would either allow for adding additional processing routines to be used in parallel with the FCPP, or improving the audio output quality by increasing the number of channels used for mask estimation. Further system improvements may also be possible by migrating the implementation to a fixed-point processor, as these DSPs can typically run critical routines faster than floating-point processors like the SHARC.

3.3 Mixed-Signal Interface

Of critical importance was the fact that the on-board analog I/O was limited to a stereo input only. This permitted only a basic, 2-microphone version of the cocktail party processor to be implemented directly on the board. However, as this system only required the construction of a single combined microphone and amplifier unit for each ear, it greatly simplified the design and implementation of the system.

An additional version making use of the full four microphones required in the original FCPP design demanded the further construction of an interface board to handle the extra inputs. This was accomplished first through bread-boarding and later through the use of a custom-made printed circuit board (see Annex C for the relevant schematics). Unfortunately, both of these implementations exhibit an unaccounted-for attenuation at low to mid-level audio frequencies. While this can be readily fixed using a suitable equalization stage, the time constraints associated with the present project mean that this solution must be left as a requirement for any future work.

3.4 Audio Input

The audio input units consist of a pair of microphones, one located on each ear, similar to what was shown in Figure 2 of this report. The microphone signal is amplified directly on the board using the circuit shown in Annex C before being fed to the relevant audio input on either the DSP board or the external codec board. The four-microphone version of this circuit is identical, except that instead of a single forward-facing microphone, each board contains one forward-facing, and one backward-facing microphone. Future improvements for this unit should include a secondary gain stage for the purposes of frequency equalization.

4 Recommendations for Further Work

Smart Helmet System

1) Add in the component for radio communication.

2) Tuning of system parameters to reflect correct physiological / physical data.

3) Build better decision-making capability into the mode-switching system.

4) Replace the ad-hoc acoustic models with physically measured ones.

5) Implement the combined system in hardware.

Fuzzy Cocktail Party Processor

1) Reduce the algorithm's computational complexity through sub-sampling.

2) Explore alternate filterbank structures in order to improve performance and reduce computational complexity.

3) Implement a version of Coherent Independent Components Analysis (cICA) for use with this system.

4) Implement the adaptive noise canceller as a pre-filter instead of a post-filter.

Active Noise Control / Adaptive Noise Cancellation

1) Replace current adaptive FIR algorithms with IIR-lattice type implementations.

2) Implement additions for a radio communications hook-up.

Hardware

1) Replace current codec set with a more flexible design.

2) Create a second amplifier stage for gain equalization.

3) Use balanced audio input lines instead of the single-ended input currently in place.

5 Conclusions

The goal of this project was to develop the design for an integrated hearing protection system that could provide both speech enhancement and a dynamic response to a changing acoustic environment. To this end, the result of this work has been this report's proposed "Smart Helmet" design. This design is based on a combined ANC and Fuzzy Cocktail Party Processor system, which provides both high-quality noise suppression in extremely noisy situations, but also allows for the elimination of highly non-stationary noise in more moderate environments.

The overall system design has been effectively demonstrated in MATLAB, which combines all of the primary features of the helmet, and allows for further study of both the individual components as well as their interactions. In addition, a hardware implementation of the Cocktail Party Processor has also been constructed. This system allows for a demonstration of both the practicality of this technology, but is also flexible enough that the primary algorithm can be studied using both direct PC audio connections, as well as live acoustic testing. It is maintained therefore that this project has been successful in its key aims, not only showcasing the value of the technology, but also providing the basis for further development that will lead to a field-testable prototype.

Annex A Hardware Details

A.1 Setup Instructions for the MATLAB Smart Helmet Model

The MATLAB SmartHelmet program can be invoked in the following way:

```
output = SmartHelmet ( front_pair, rear_pair, P)
```

where the variable `front_pair` refers to the vector containing the front-facing microphones, `rear_pair` refers to the rear-facing microphones, and `P` is the headphone's primary acoustic path transfer function. Various speech-in-noise scenarios can be tested this way, provided they can be created by the user. Several sample scenarios are also available in the file "SmartHelmet.mat".

A.2 Setup Instructions for the DSP Board

The basic development platform for this project was the ADSP-21369 EZ-KIT [12]. This is the development/demonstration board for Analog Devices' SHARC ADSP-21369 floating point DSP chip. This board allows for a single stereo input, as well as multiple stereo outputs. A different arrangement of inputs and outputs is possible through the use of connectable peripherals.

A.2.1 Board Settings and Connectors: Two-Microphone Case

The FCPP algorithm can work using either a two-microphone or a four-microphone arrangement. If the two-microphone implementation is chosen, the audio input and outputs connected directly to the development board using the board's default settings. The following summarizes the system's hardware settings and connectors.

Switch Settings:	All switches should be set to the default positions.
Audio Input:	Use connector J10 (stereo RCA connectors)
Audio Output:	Use connector J9, also denoted as CH4_HP. This connector is a standard 3.5 mm stereo jack.

Volume Control: Pushbutton 1 and Pushbutton 2.

Using this arrangement allows for simpler implementation in that both signal conversion, and amplification can be handled using the EZ-KIT. However, as noted, earlier, this limits the capability of the FCPP since it only permits the use of two input microphones. For best results, a 3-4 microphone system should be implemented.

A.2.2 Board Settings and Connectors: Four Microphone Case

If a four-microphone implementation is desired, the board should be reconfigured as follows:

Switch Settings: Change SW3.4 to the "off" position. This disconnects the on-board codec from the DSP.

Audio Input: Connect the right audio inputs to jack X1 on the peripheral device, and the left audio inputs to jack X2.

Audio Output: The audio output should be connected to jack X3 on the peripheral device.

Peripheral: The current implementation connects to the DAI header, labelled on the EZ-KIT as P4.

There is no volume control available for this version, owing to the fact that the AD73322 codec does not support a programming interface separate from the serial data interface. It should also be noted that the programs for the two and four microphone cases are not currently interchangeable. The reason for this is the different setup procedures needed by the two versions.

A.2.3 Microphone Mountings

Circuit boards for mounting the microphones and for performing initial amplification have been constructed for both the two- and four-microphone systems. All boards use Knowles Electronics FG-23629-C36 electret microphones, which are the same type currently used in some

common hearing-aid systems. Also common to all versions is the choice of amplifier, the LME49721 from National Semiconductor. Readers are referred to Annex C for the relevant schematics. The current implementation of the amplifier boards uses a 5 V line to power the amplifiers, and an additional 1.1 V button cell to power the microphones. The battery is an on-board device, while the 5 V line connects to an external power source. Refinement of this design should include replacing this set-up with a single unified power supply.

A.2.4 Physical Mounting

For a given version of the software, it is not necessary to run the DSP board from the manufacturer's IDE. Instead, the code can be loaded into flash memory, which will run on boot-up. A system reset may be necessary though if the IDE was previously in use.

When operating the combined system for testing, the following options are available:

1) Direct PC hook-up.

2) Full set up in an acoustics laboratory.

These options are explained below

PC Hookup

Using the two-microphone version of the software, the system can be run directly from the PC's sound card. All that is required is that the PC's audio line-out be connected to the EZ-KIT's line-in (J10). This eliminates the need for live acoustic tests during much of the development process. Any set of binaural recordings will suffice for testing purposes. For the most realistic results, however, these recordings should be made from an appropriately configured acoustics mannequin. Use of R-HINT-E or similar software for creating test scenarios is recommended [13].

Full Set-up

The full set-up is necessary for live acoustic tests, or other cases where some aspects of the peripherals need to be tested. In this case, the following steps must be taken:

1) Change the EZ-KIT's board switches to the correct settings.

2) If appropriate connect the external codec board via the ribbon cable.

3) Mount the microphone boards on the helmet according to the manner shown in Figure y.

4) Hook up the external power cable to the microphone board's 5 V connectors.

5) Connect the audio cables to the microphone board and if appropriate, the external peripheral. In the four-microphone version, the order of connection should be such that the red cable connects to the red plug. This matters since it is the difference between the front and back microphones.

6) On the two-microphone version the audio output should be connected to J9. On the codec board, the audio output should be connected to the audio jack provided. Both output connectors are standard 3.5 mm jacks.

7) Make sure the batteries are properly connected.

8) Power up the external power supply, and the EZ-KIT. The code should run on boot-up.

A.2.5 Audio Recording

For the purposes of testing, it is necessary to be able to record to the PC from FCPP system. The procedure is somewhat different for each of the two configurations, but still fairly simple:

1)From the EZ-KIT: Connect the PC audio input to connector J9 on the EZ-KIT.

2)From the codec board: Connect the PC audio input to connector X3 on the board.

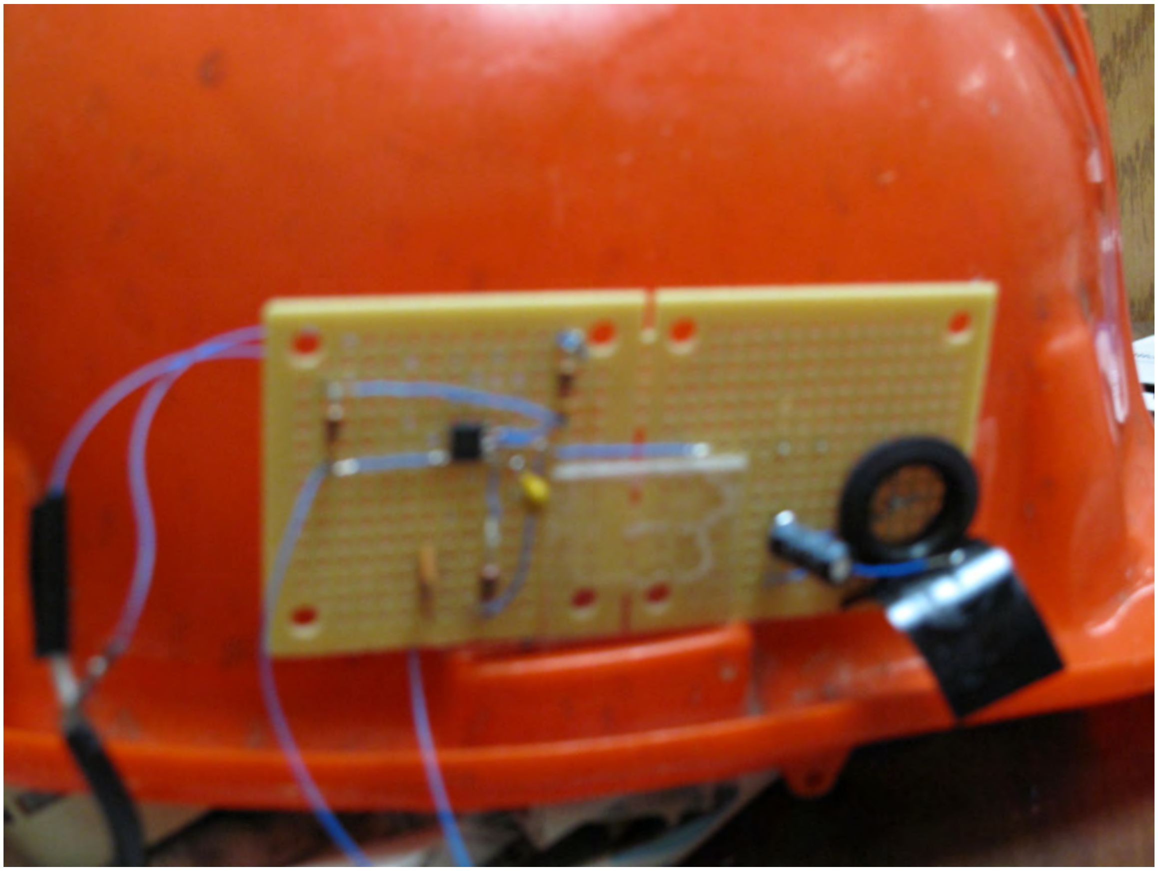

Figure A.1 Physical mounting for the microphone / amplifier boards. For the single microphone system the microphone should face forwards. The rubber washer is meant to hold a standard 1.1 V button cell battery.

This page intentionally left blank.

Annex B Software Documentation

B.1 MATLAB Software Modules

Table B.1 The main SmartHelmet modules

Task	**Function**	**Description**
Initialize system variables	SmartHelmet.m	This sets up various constants needed elsewhere in the program, such as sampling length, filterbanks, etc.
Create headphone model	HeadphoneCreate.m	This creates a structure containing the model information for the headphones such as the primary and secondary path models.
Create ANC model	ANCCreate.m	This creates a structure containing the basic variables, filter vectors and buffers used by the ANC model.
Perform ANC filtering and adaptation.	ANCFilter.m	Calculates the current output signal and updates the ANC adaptive filter.
Create Adaptive Noise Cancellation model	AdNCCreate.m	As a above, but for the adaptive noise canceller.
Perform Adaptive Noise Cancellation.	AdNC_Update.m	Calculates the filtered output and updates the adaptive filter.

B.2 FCPP Modules for the ADSP-21369 EZ-KIT

Table B.2 Internal Configuration

Task	Modules	Description
CPU Setup	InitPLL_SDRAM.c	This configures the CPU and internal clocks.
SRU Setup	InitSRU.c InitSRU_4mic	For the Signal Routing Unit (SRU). This configures the routing of I/O signals within the DSP.
I/O Interface	InitSPORT.c InitSPORT_4mic.c	The configuration here is concerned with the correct setup of the direct memory access (DMA) interfaces. These are responsible for the reading and writing of external serial data, including the correct buffering, formatting, and data rates.
Interrupts	IRQProcess.c	The interrupt handling is responsible for activating the processing loop upon receipt of a full data frame. In addition, for the two-microphone version, two other interrupt handlers are configured to control the loudness of the audio output.

Table B.3 Configuration of External Devices

Task	Modules	Description
SDRAM	InitPLL_SDRAM.c	Owing to the size of the data sets that must be manipulated, it is necessary to access the slower, external RAM. Configuration involves setting up both the hardware-level access information as well installing the memory heap
2-mic Codec	Init1835viaSPI.c	The on-board 1835A codec chip must be correctly configured for the desired sampling and data-transfer rates.
4-mic Codec	InitSPORT_4mic.c	The AD73322L codecs must be programmed differently than the onboard codec. This also requires a slightly different internal setup as well with respect to buffers, data width and so on.

Annex C   Circuit Schematics

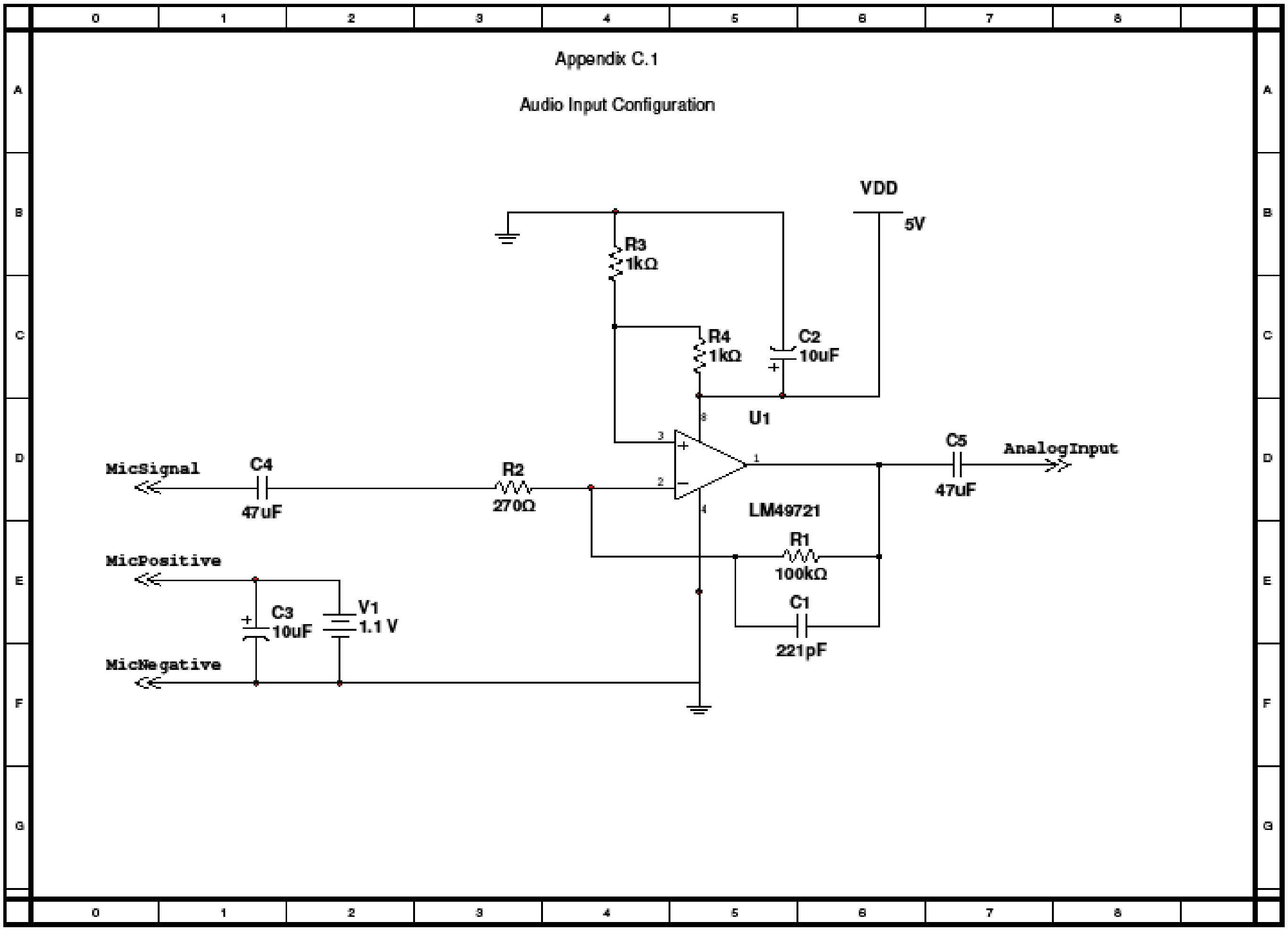

Figure C.1 The circuit diagram for the audio input stage. This forms the combined microphone and amplifier section, which is to be mounted on the wearer's head. This circuit is repeated for all microphones.

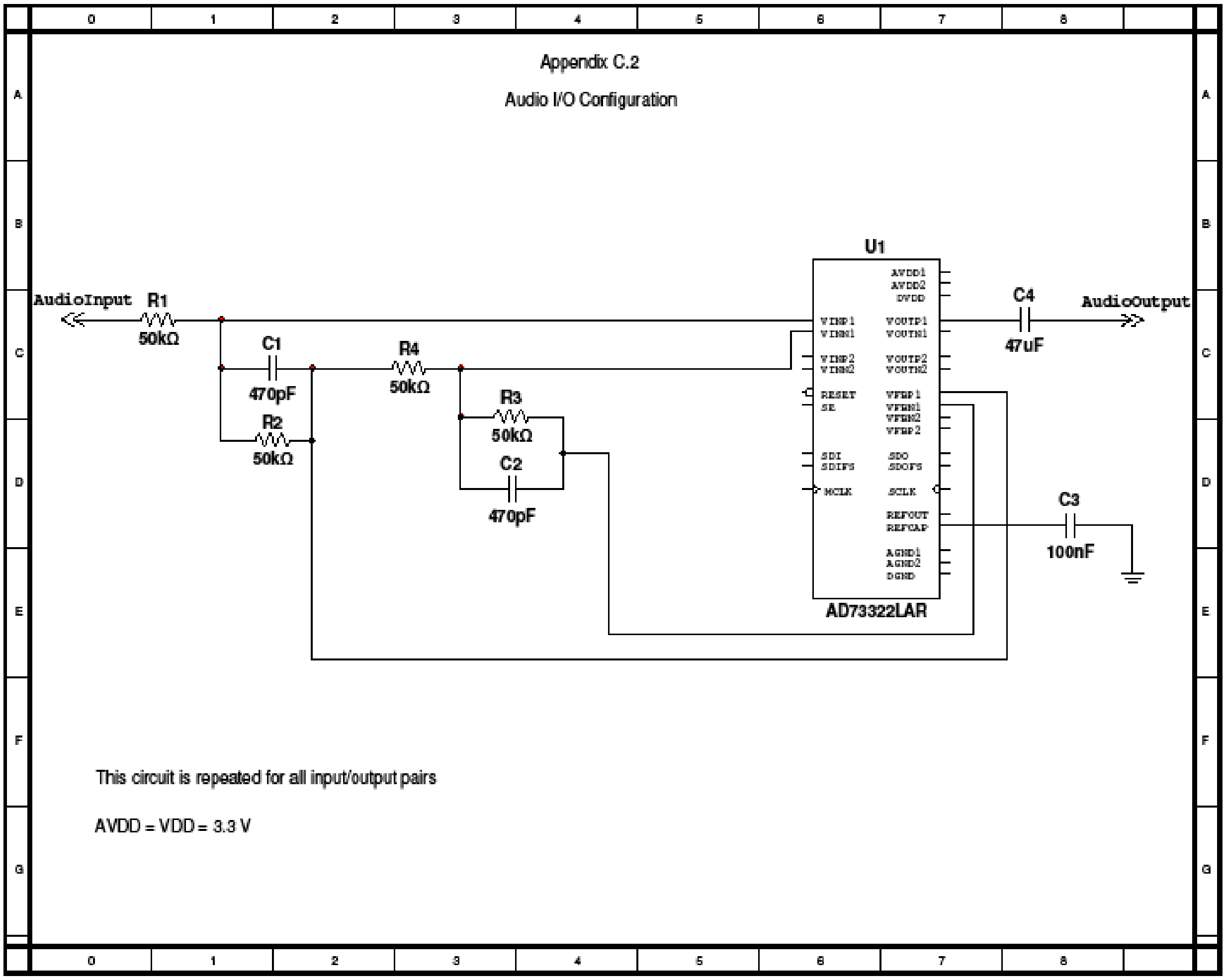

Figure C.2 The analog I/O connections for the codec board. The relevant schematics for this section were taken from the suggested design in [14].

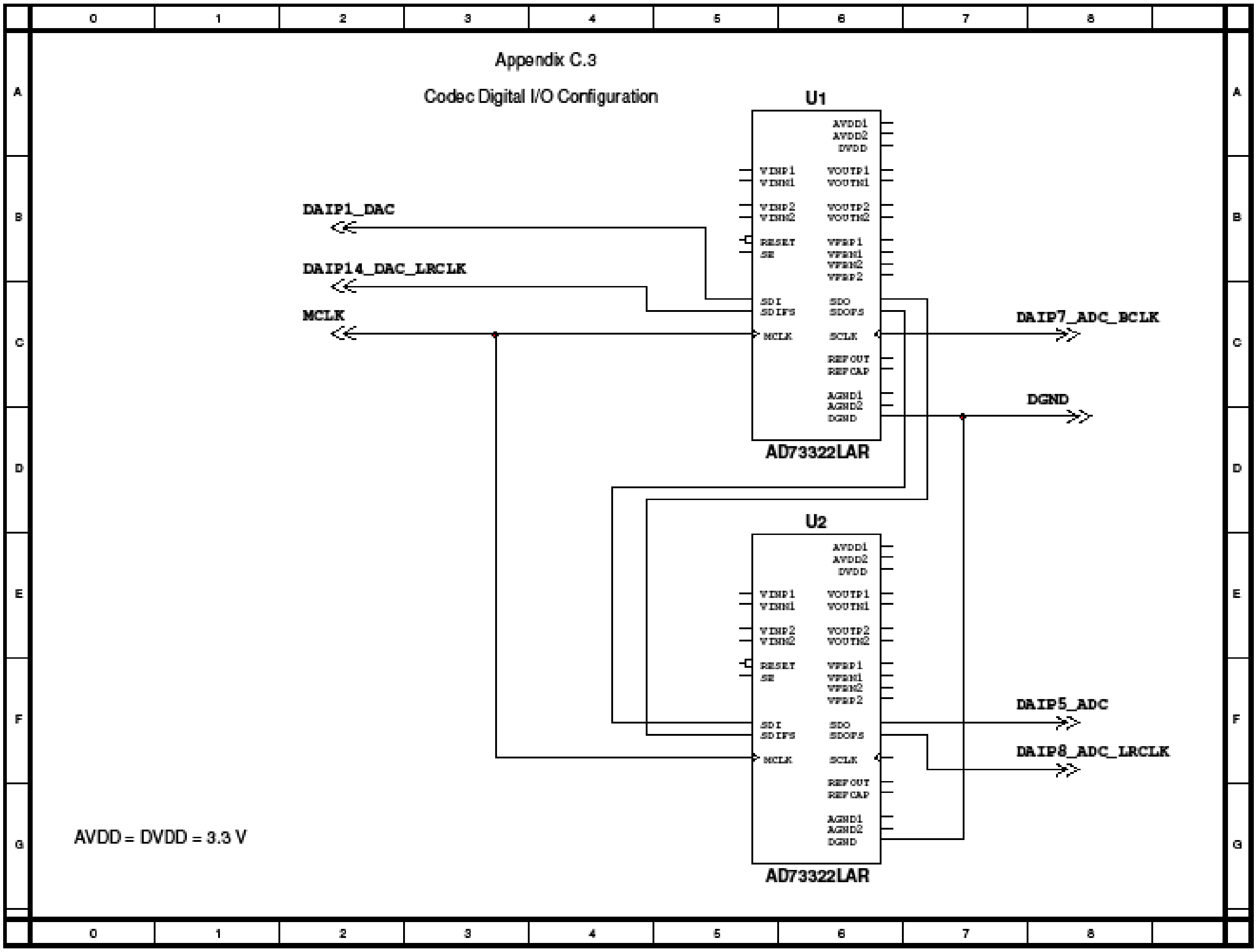

Figure C.3 The digital I/O section connects the codecs to the processor.

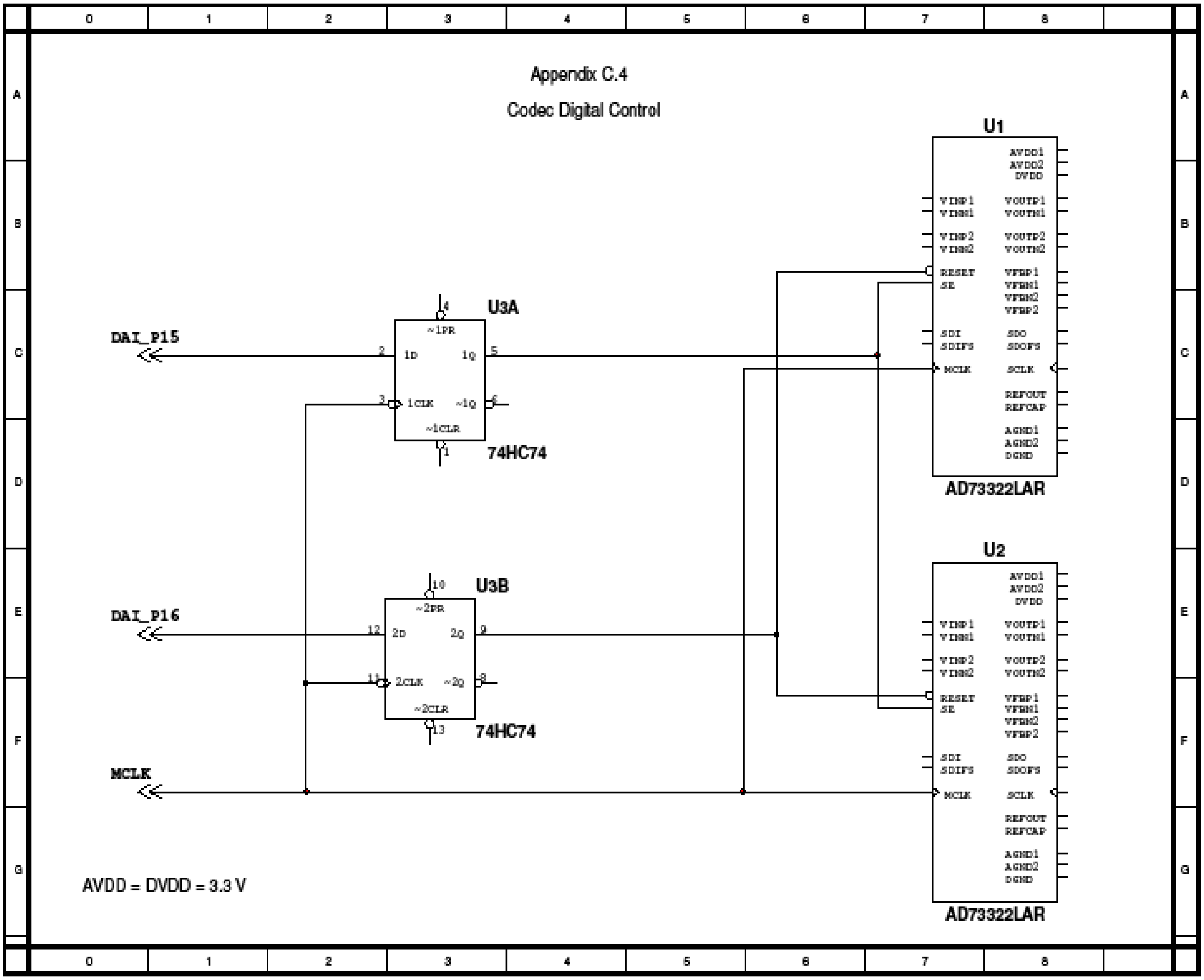

Figure C.4 The synchronization control for the twin codecs. Providing synchronized resets ensures that the codecs transmit and receive in order and at the right times.

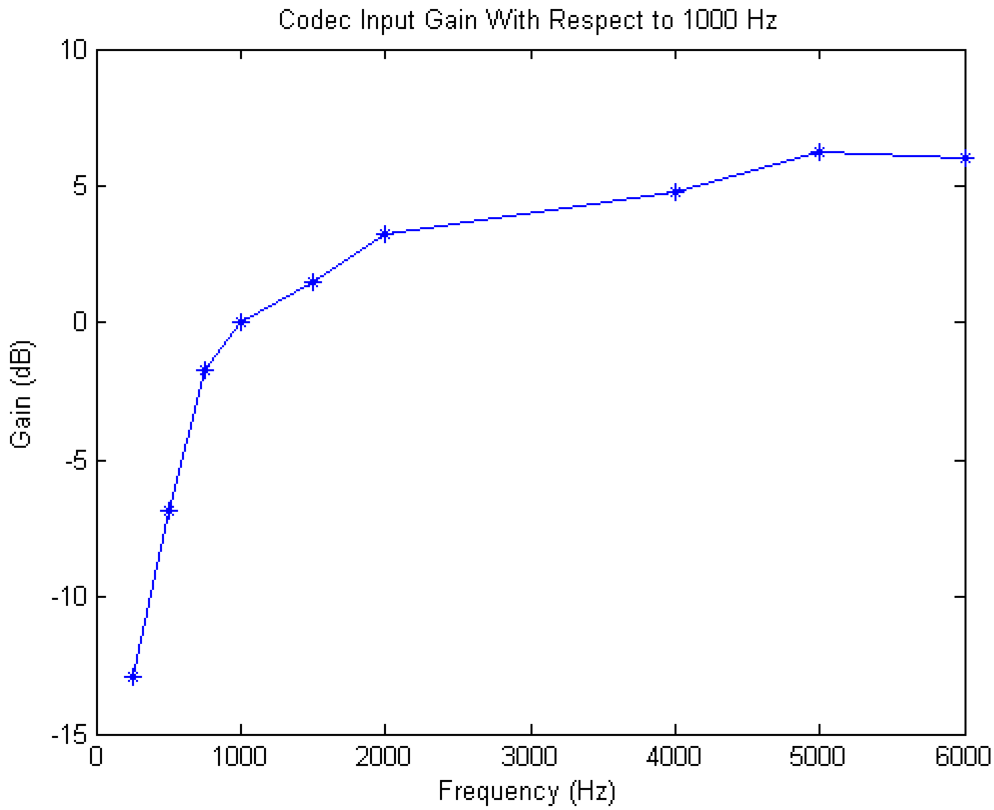

Figure C.5 The frequency response of the PCB-mounted audio codec. This situation may be remedied by the addition of a secondary gain stage, to amplify the low-frequency portions of the signal. Alternatively, the reason for this effect should also be sought out

.

This page intentionally left blank.

Annex D Smart Helmet Array Processor

Smart Helmet:

ARRAY PROCESSOR

Built for Speech Communication in Noise Environment Project

PWGSC Solicitation No.: W7711-088143/A

15 March, 2010

Prepared by

Advanced Ultrasound Technologies Inc.

122 James Gray Dr., North York, Ontario, M2H 1P1

Tel: (647) 818 1861 Email: fan.zhang@utoronto.ca

D.1 Summary

The main objective of this project was the development of a system, herein called Array Processor, that can detect and pinpoint the occurrence of Impulsive Events in the presence of Environmental Noise. An example of such an environment is given in Figure D.1.

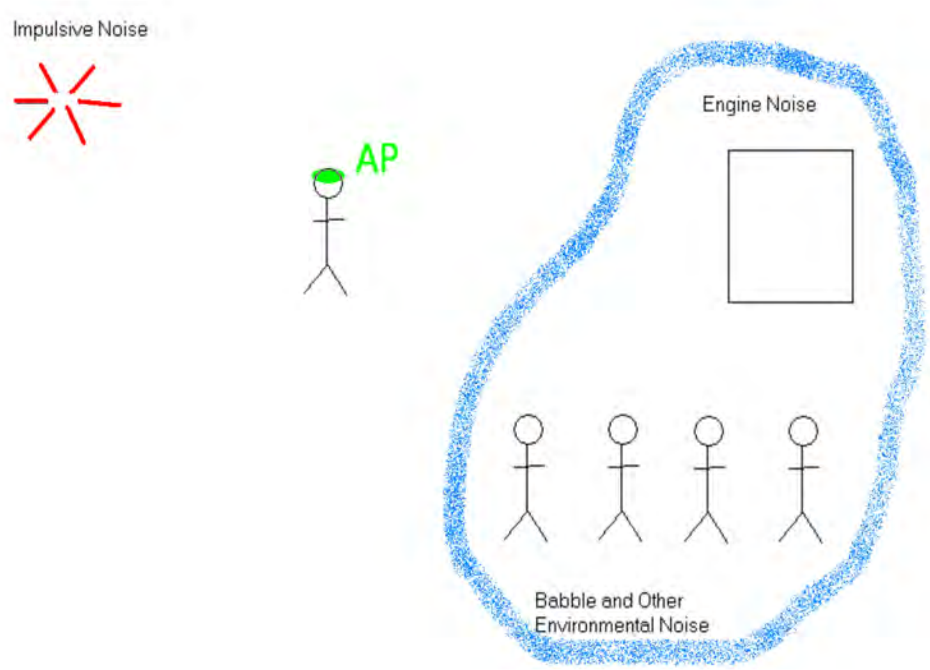

Figure D.1 In the presence of environmental noise, the array processor (AP) may also be exposed to impulsive noise events.

The Array Processor's acquisition unit was designed to be able to adapt to changing noise levels allowing operation in both quiet and very noisy environments. Several parameters in the algorithm of the Processing Unit can be configured to only evaluate impulsive events that match a certain pattern.

During its early stages of development the Array Processor was successfully tested with synthetic data externally generated and uploaded into the Processing Unit of the Array Processor.

At a later stage, the hardware's performance and the accuracy of the algorithm were tested with real acoustic signals recorded with the system's microphones.

D.2 System Overview

The current version of the Array Processor can be divided into the Acquisition and Processing Units. Each of the two units is designed on an individual Printed Circuit Board (PCB) that allows various configurations to be tested. In its present configuration the Acquisition Board contains sixteen (16) identical analog channels and two (2) Analog to Digital Converters (ADC). The Processing Board houses the various power supplies required by the different units as well as the Processing Unit itself. The assembly of the Array Processor's Acquisition and Processing Unit is shown in Figure D.2 .

Figure D.2 The acquisition and processing unit.

D.3 Components of the Array Processor

D.3.1 Acquisition Unit

The Acquisition Unit of the Array Processor contains sixteen channels to allow the connection of sixteen individual microphones. Each of these channels is connected to an Analog-to-Digital Converter that samples and digitizes the analog signal. The output of the ADC is forwarded to the Processing Unit. The Acquisition Unit also generates the required bias voltages for the attached microphones.

To reduce the footprint of the system the design utilizes devices with more than one analog channel. In particular, of the total sixteen channels, each pair of analog channels is combined into a single unit. Four of these dual channel units are connected to a single octal Analog-to-Digital Converter. Both Analog-to-Digital Converters are connected to a single clock signal and synchronized. This allows all sixteen channels to be sampled at the exact same time. Figure D.3 shows these eight pairs of analog channels and the two octal Analog-to-Digital Converters.

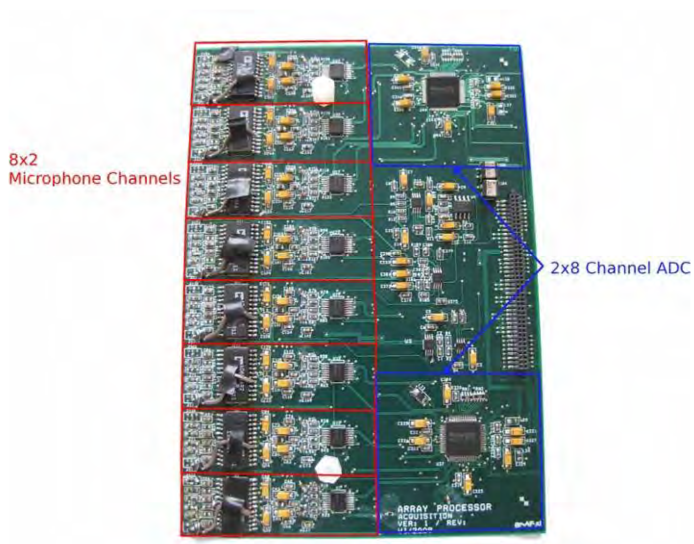

Figure D.3 Acquisition unit showing eight dual channels and two eight-channel ADCs.

D.3.2 Processing Unit

The two Analog to Digital Converters produce a data volume in the approximate rate of 6 MB/s. The Processing Unit has to be fast enough to process this data in real time. There are basically two devices that could be used for this application: The highly versatile and fast FPGA or a DSP. An FPGA implementation offers enormous flexibility for building high-speed parallel processing structures. However, the disadvantage of the FPGA is that standard interfaces like SDRAM, SPI, High-Speed Serial Ports, etc. have to be programmed into the device and thus occupy a major portion of its programmable gate arrays. A DSP implementation on the other hand provides flexible hardware interfaces for a wide variety of memory standards as well as a variety of external devices. Some of these interfaces can even function without straining the Central Processing Unit (CPU). The use of either a fast or a multi-core DSP offers similar performances and already includes all interfaces in hardware. Figure D.4 depicts the current design of the Processing Unit on a six layer Printed Circuit Board.

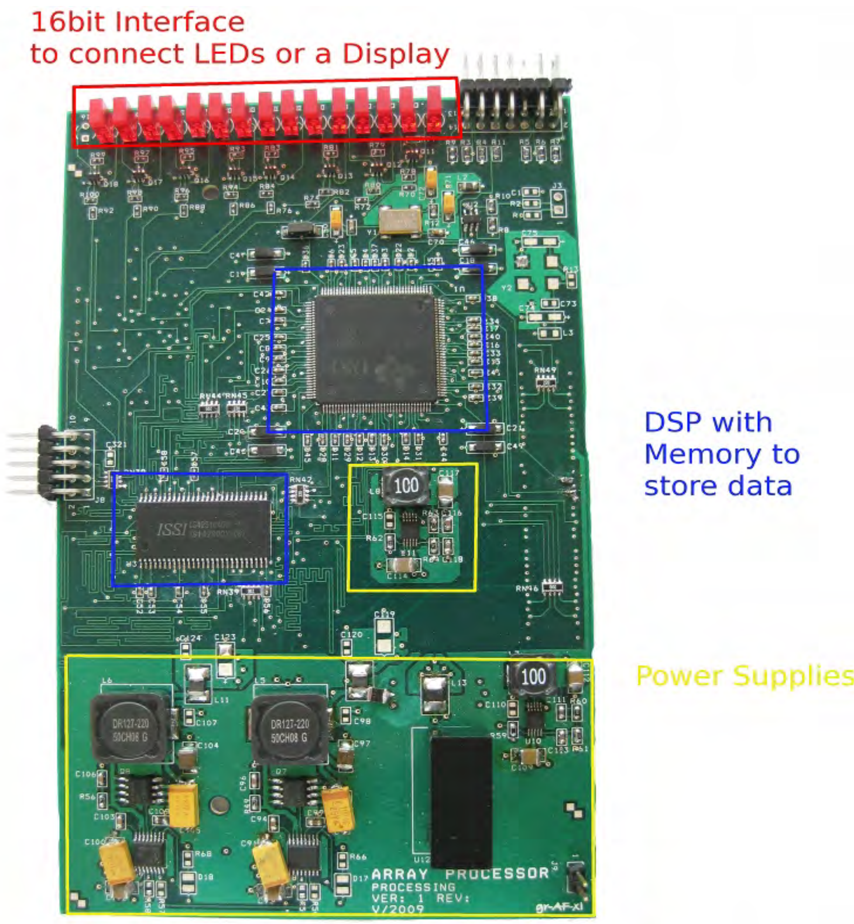

Figure D.4 The processing unit.

D.4 System Processing

The processing evaluated in this report consists of two major parts:

- Data Acquisition
- Signal Processing

D.4.1 Data Acquisition

The data acquisition is fully automated utilizing the DSP's Multichannel Serial Audio Port (McSAP)[1] and dMAX[2] hardware. The McSAP is configured to produce a Frame sync signal at a rate of 97,646.25 Hz. The Frame sync signal is fed to the ADC that in turn outputs the data on the rising slope of the frame sync signal. Each frame is divided into two slots. The first slot (slot0, HIGH Frame sync) contains the data from the ADC while the second (slot1) remains empty allowing the frame sync signal to go low and become ready for the next frame. A graphical representation is shown in Figure D.5.

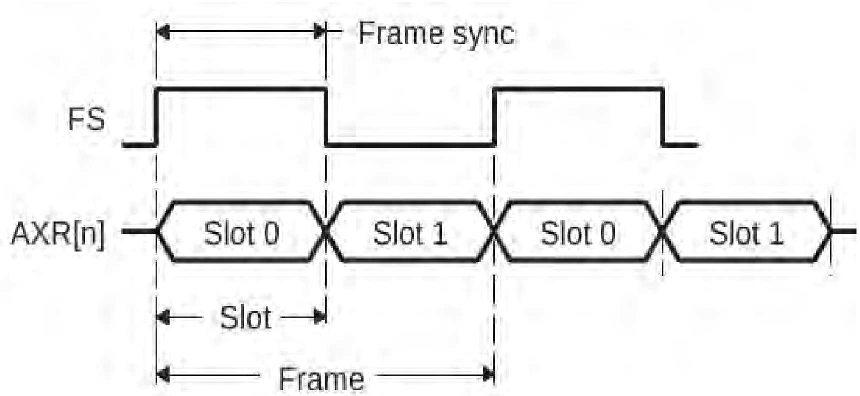

Figure D.5 Definition of frame and slot.

The McSAP triggers the dMAX at the end of the first slot. The dMAX reads the data from input buffers of the McSAP and stores it in one of two small buffers allocated in the internal memory. The two buffers are used in a “PING”-”PONG” manner. While new data is written into the PING buffer, the previous data can be read from the PONG buffer and vice versa. Each of the buffers can currently store 128 samples. When one of the buffers is full the buffers are switched and the DSP is interrupted to start reading and pre-processing the new data.

---

1 McSAP: Multi channel Serial Audio Port allows the parallel interface of up to 16 serial channels of audio data

2 dMAX : Data Movement Accelerator allows the fast transfer of data from one memory to another

D.4.2 Data Processing

High Priority Tasks

The data processing is divided into tasks with different priorities. A semaphore is raised when the dMAX triggers the DSP and new data has become available. A high priority task is called that starts reading the new data from the internal buffer. This task writes the new data into circular buffers that are allocated in the external memory. There is one circular buffer for each channel. To reduce the amount of data being processed, this task also creates a down sampled version of the original signal. At the end of this task the DSP checks the signal level and adjusts the gain of the Analog Front End if necessary.

Low Priority Tasks

Upon completion of the high priority task the DSP switches to lower priority tasks. These tasks analyze the down-sampled signal for possible impulsive events using the gradient of the signal's envelope. If a strong gradient signals the possibility of an impulsive event in one channel, the Peak Detection task is called to verify the presence of an actual impulsive event.

Peak Detection

The verification of such event includes the analysis of the pulse's duration and a spectral analysis. The spectral analysis is based upon a 2 k FFT of the original data sampled at the higher frequency. Since reliable real data is currently not available the result of this analysis is ignored. Yet, the code is active to test the real time performance of the system.

Direction Finding

Once a signal is considered as an impulsive event, the direction finding task tries to define the direction of the source. To get accurate information about the recorded signal, the data recorded at the original higher speed is processed.

Due to tolerances of the individual components or external influences (dirt on microphones) the channel that triggered the algorithm is not necessarily the channel that carried the signal first, but it is assumed that this channel is very close to the actual direction. To further assure the detected channel is the right one, a cross correlation for the next two adjacent channels is calculated. The replica for the correlation is derived from the channel that triggered the direction finding algorithm. The channel that has its correlation peak first will determine the direction of the signal.

The current hardware cannot calculate the cross correlation for all sixteen channels while satisfying the real time criteria. As a matter of fact, the correlation information of channels from the back side of the helmet might reduce the accuracy of the system. However, it can be practical for a reduced set of microphones. Shown below in Figure D.6, for example, are the recordings for a set of three microphones placed 2.5 cm apart on a circular aperture.

In the left subplot of the figure, the source was closer to the left channel (denoted by '*'), and so it is seen peaking first. In the right subplot, the placement of the source is reversed, and it is closest to the right channel (denoted by the '∀'). In both cases, correctly- the middle channel peaks between these two extremes

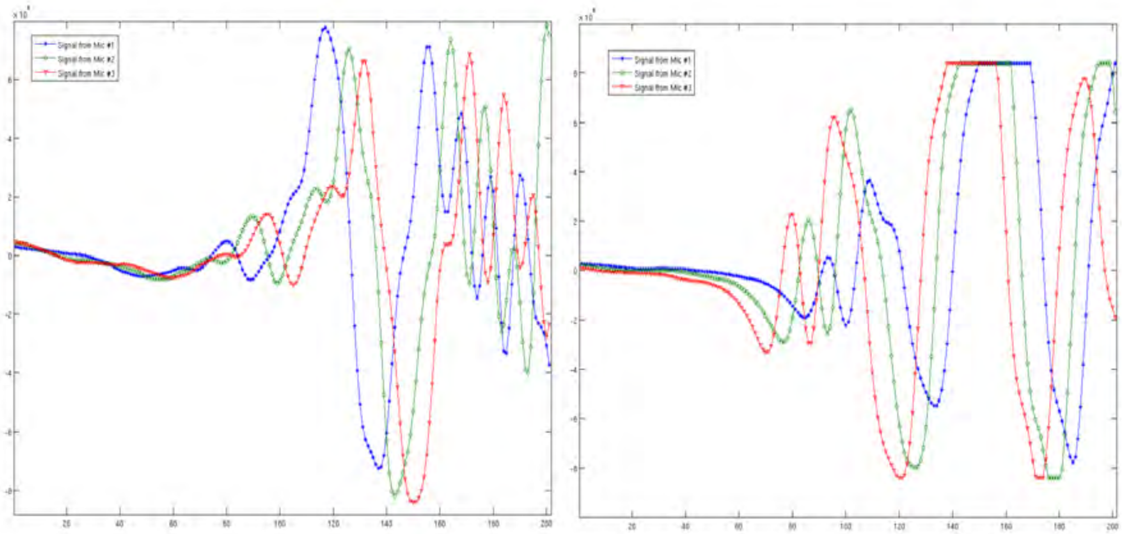

Figure D.6 Comparison of data recorded from two different directions, using three different microphones.

The difference between two adjacent channels is approximately 6 sample points. This equals at the sample rate of the system to

$$\Delta t = \frac{6}{97{,}656.25Hz} = 61.4\mu s$$

Given that the speed of sound is approximately 340 m/s the two microphones are

$\Delta d = 61.4\mu s \cdot 340\text{m}/s = 2.1cm$ apart. The difference between the theoretical and measured value results from the position of the source not being on the same plane as the microphones and the initial coarse estimate using sample points.

D.5 Microphone Mounting

To test the performance of the Array Processor the sixteen microphones had to be mounted on the helmet. In the current setup, equally spaced Velcro patches were affixed to the circumference of the helmet and on the backside of each microphone. This approach allows the test of different numbers of microphones and locations on the helmet. Furthermore, it is possible to evaluate different models of microphones. The following figure depicts the Array Processor and the mounting of the microphones on the helmet.

Figure D.7 The Smart Helmet Processing array.

D.6 Preliminary Testing

In the following the Array Processor is tested with external acoustic signals. To allow signals with high amplitude, small, air filled bags were exploded in the proximity of the system. The relative location of the acoustic signal is changed to test the performance of all available channels. In addition, to evaluate the effect of different signal amplitudes the sensitivity of some channels was altered.

During normal operation the Array Processor collects data from all sixteen channels and stores is into a circular buffer. This buffer is big enough to store one second of data. Once this buffer is full previous recorded data is normally overwritten. To better understand the results of the direction finding algorithm the data acquisition is stopped as soon as the Array Processor detects a possible impulsive signal. The data from all sixteen channels is then extracted from the Array Processor to allow the display and off-line processing. This last procedure is done only for the purposes of data collection for this report.

D.6.1 Normal Condition

In this test all channels are working properly and there are no differences between the analog channels other than the ones caused by manufacturing tolerances of the individual devices. The source of the acoustic signal is located five meters away from the microphone attached to channel '1' and is approximately 4 ms long. The following figure shows the data for the channels 0 through 15. The time period within the one second buffer is shown on the *x* axis while the *y* axis depicts the amplitude of all sixteen channels.

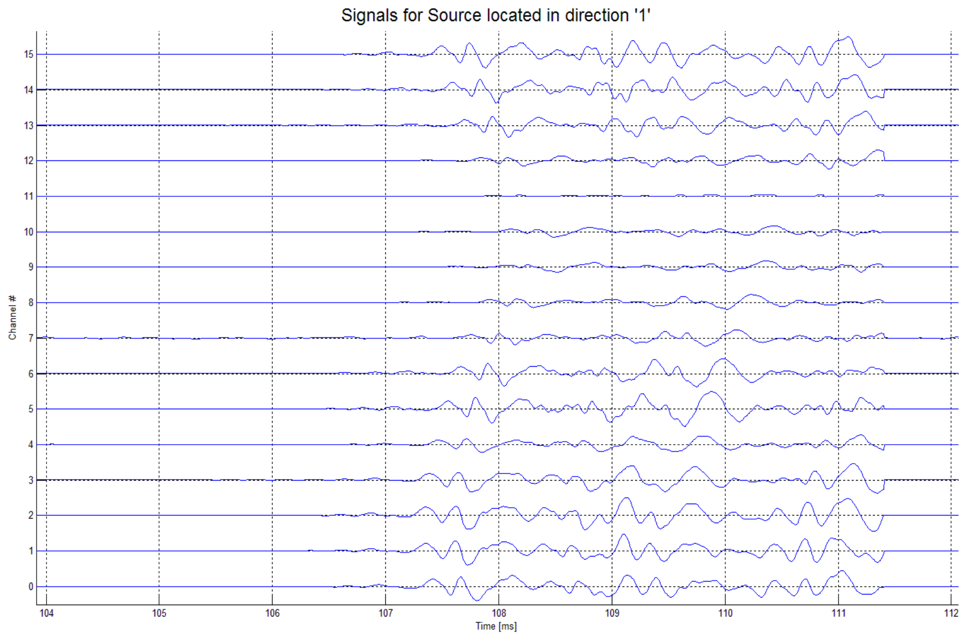

Figure D.8 Data recorded on the sixteen channels for a 4 ms acoustic input signal.

A look at the signal amplitude shows that the acoustic signal was recorded on all sixteen channels. The amplitude gradually decreases when approaching the signal of channel 11 from either side, but it is hard to tell which channel carries the strongest signals. To find out where the source is located, the amplitude cannot provide sufficient information. For that reason, the signal of the channel which triggered the processing, is used as a replica and correlated with the data of the other channels. The outputs of all calculated correlations are then analyzed and the channel that received the signal first is calculated. This is shown in Figure D.9 below, where the channel corresponding to the source direction is indicated by a solid square. The result for this test is channel '1', which agrees with the actual source direction for the trial.

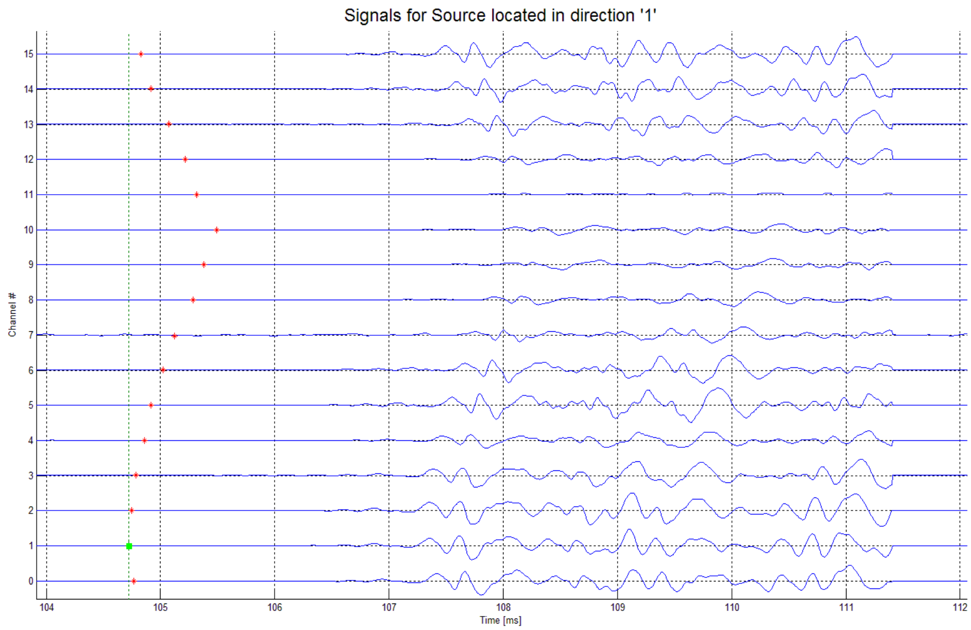

Figure D.9 Results after correlation (denoted by ''), and the selected channel, denoted by '□'.*

D.6.2 Sensitivity

For this test all channels are working properly. Some (3) channels have a higher gain than the remaining. This test is to simulate the compromised functionality of some microphones due to external influences (dirt or leaves on microphone), resulting in a reduced sensitivity. The source of the acoustic signal is approximately 4 ms long and again located five meters away from the closest microphone in a direction close to the altered channels. Figure D.10 shows again the data for the channels 0 to 15. The channels 0 through 2 are the ones with the higher gain. As in the previous figures the time is plotted on the x - axis while the y - axis lists the amplitudes of the sixteen channels. Also as in the previous example, the results of a cross correlation are compared and the channel that carries the signal first is chosen. For these tests, it is seen that the direction finding system is insensitive to gain differences between the various microphones.

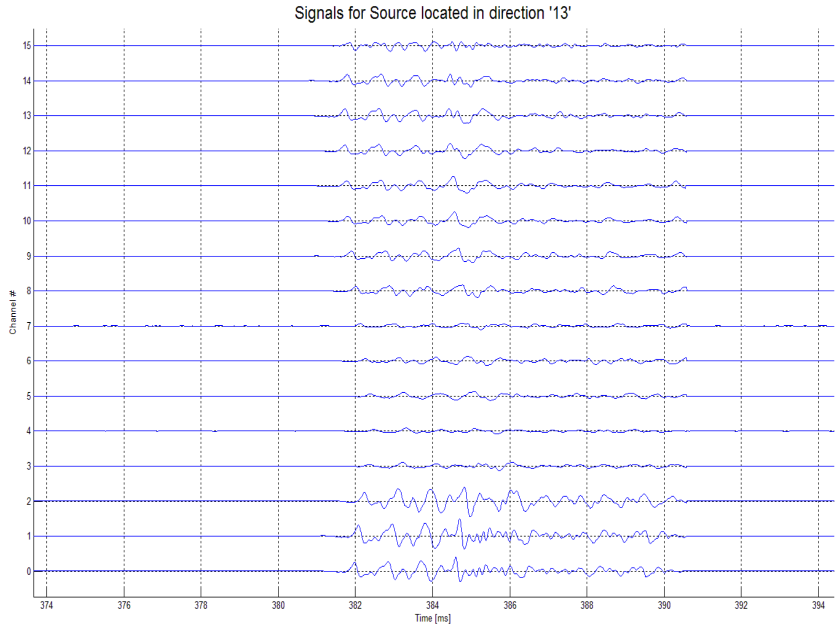

Figure D.10 Recorded data for a source close to channel 13. Channels 0 through 2 are configured with a higher gain.

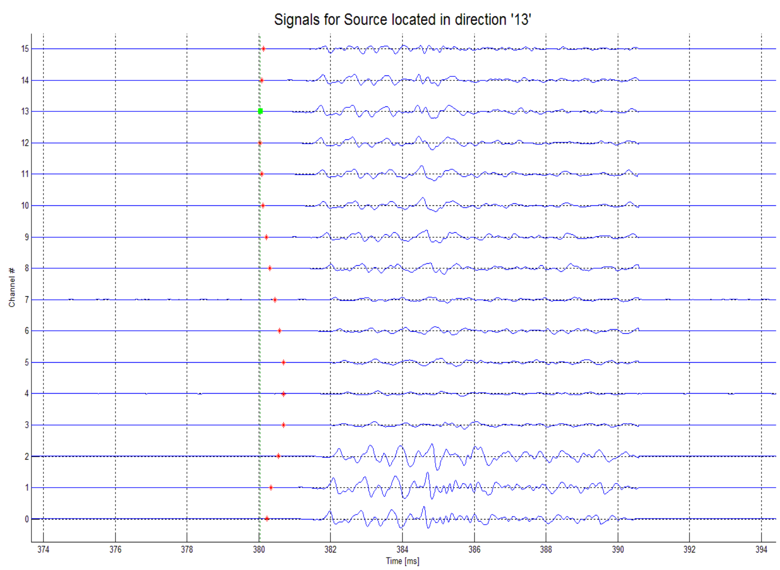

Figure D.11 Results after correlation. The selected channel is denoted by the '□' symbol.

D.6.3 Distortion

The Array Processor's digital automatic gain control (AGC) requires some time to adjust from a quiet environment to very loud impulsive events (increase of more than +54 dB). This can cause the signal to be distorted. To test the system's behaviour, the AGC is set to a fixed high gain. Due to the limited voltage range in the Analog Front End (Amplifiers, Filters, ADC) the amplitude of some signals will be distorted (clipped). As it can be seen in Figure D.12, the amplitude for almost all recorded signals suffers distortion. Compared to the previous data sets, the signals above will contain a variety of harmonics that artificially increases the bandwidth of the recorded signal. The extent to which the harmonics were created depends on the amount of signal that was clipped off and varies for each channel. Rather than trying to restore the signal by filtering and extrapolation, the Array Processor uses again the first detected signal as replica.

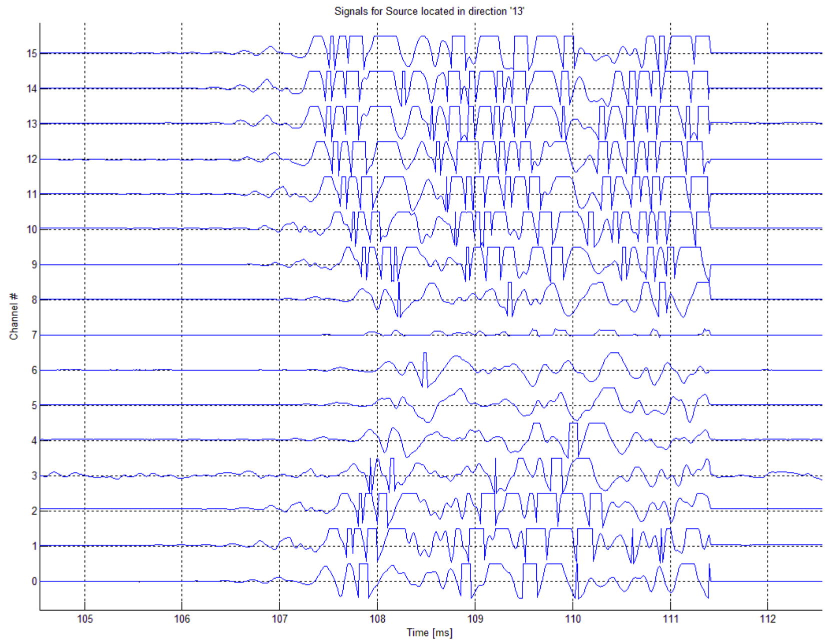

Figure D.12 Distorted input signals for a very loud signal.

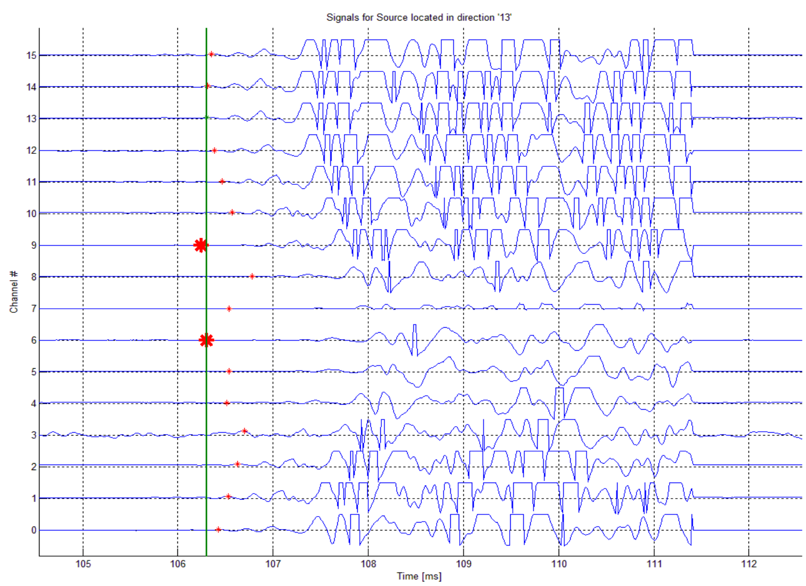

Figure D.13 Results of the direction-finding algorithm.The bold stars would create wrong results.

.

As expected, the result is not as good as for undistorted signals. The results of the cross correlations where the signal was less or not distorted are not usable. The bold red stars in Figure D.13 are not only incorrect but would also cause the algorithm to fail, if considered by the algorithm. As described previously, the actual algorithm on the Array Processor does not compute the cross correlation for all sixteen channels but rather for the channel that triggered the algorithm and the two adjacent channels. This allows the Array Processor to meet the real time criteria and does also eliminate the results for channels that did not receive the signal with its full dynamic range.

The Processing Unit of the Array Processor currently processes data in segments of 128 sample points. The information from each of these individual segments is used by the Automated Gain Control unit to adjust the amplitude. At the given sampling frequency of 97.656 kHz new information is available every 1.3 ms. To further overcome potential problems due to clipping, this response time can be reduced by analyzing smaller segments of data. Smaller segments on the other hand increase the frequency at which the DSP receives interrupts, and thus the rate at which it must analyze new data. Due to the fact that it takes some time for the DSP to stop its current processing in order to service an interrupt, the over all performance will decrease with an increasing frequency of interrupts.

D.7 Areas For Further Research

D.7.1 Testing Under Real Conditions

The most important task is testing the system in a controlled environment as close as possible to its final application. This step is necessary to define the system's capabilities and limitations.

D.7.2 Power Reduction

The system currently requires approximately 10 W of power. Most of it is used by the sixteen receiving channels in the Analog Front End. In order to make this device portable and to run it from a battery of manageable size the power consumption has to be reduced. A Standby-Mode could be one solution. Most active devices feature a standby mode where they require only a fraction of their regular power. The downside is that all these active components require some time to wake-up from their standby-mode. The impulsive event might be over by the time the system is then fully operational again.

D.7.3 Channel Multiplexing

Another way to accomplish a reduction of power is by reducing the number of parts in the analog front end. Instead of providing an individual analog channel for each microphone, all connected microphones could be multiplexed into one analog channel that requires only one Analog to Digital Converter. The rate at which the data can be acquired is mostly defined by the speed of the analog switch. With transition speeds of around 200 ns the effective sampling speed for sixteen channels would then be around 300 kHz (currently 100 kHz). The only disadvantage is that the algorithm has to consider the constant offset between the data from different channels.

D.7.4 Lower Microphone Count

To further reduce the power, the overall number of microphones can be reduced. Instead of using sixteen microphones the system should be able to operate with the same or even better resolution and only eight microphones.

D.7.5 Improved Microphone Mounting

Once the total amount and type of microphones is defined, clips made out of metal or plastic can be designed to affix the microphones onto the helmet. These clips can be made to slide onto the rim of the helmet just between the rubber protection and the helmet without compromising the integrity of the helmet. Furthermore, the clips can a have small fixture on the inside to run the cabling and protect it from being torn off of the helmet.

Annex E Array Processor I/O

E.1 Microphones

The Sound Pressure Level (SPL) of acoustic signals covers a wide range. Table E.1 shows a few examples with their corresponding SPLs.

Table E.1 Sound sources

Source of sound	Sound pressure	Sound pressure level
	pascal	dB re 20 µPa
Krakatoa explosion at 100 miles (160 km) in air[dubious - discuss]	20,000 Pa	180 dB
Simple open-ended thermoacoustic device [6]	12,619 Pa	176 dB
.30-06 carbine 1 m to shooter's left side	7,265 Pa	171 dB (peak)
M1 Garand being fired at 1 m	5,023 Pa	168 dB
Jet engine at 30 m	632 Pa	150 dB
Threshold of pain	63.2 Pa	130 dB
Hearing damage (due to short-term exposure)	20 Pa	approx. 115 dB
Jet at 100 m	6.32 - 200 Pa	110 - 140 dB
Jack hammer at 1 m	2 Pa	approx. 100 dB
Hearing damage (due to long-term exposure)	0.356 Pa	78 dB
Major road at 10 m	2×10^{-1} - 6.32×10^{-1} Pa	80 - 90 dB
Passenger car at 10 m	2×10^{-2} - 2×10^{-1} Pa	60 - 80 dB
TV (set at home level) at 1 m	2×10^{-2} Pa	approx. 60 dB
Normal talking at 1 m	2×10^{-3} - 2×10^{-2} Pa	40 - 60 dB
Very calm room	2×10^{-4} - 6.32×10^{-4} Pa	20 - 30 dB
Leaves rustling, calm breathing	6.32×10^{-5} Pa	10 dB
Auditory threshold at 1 kHz	2×10^{-5} Pa	0 dB

To cope with high SPL signals like gunshots or explosions the recording microphone cannot be too sensitive. Therefore, at this stage of development the waterproof microphone VEK-F-30350-000 (Knowles Electronic) is evaluated. This microphone has a linear sensitivity of -55 dB in the frequency range from 100 Hz to about 4 kHz. The sensitivity is shown in Figure E.1.

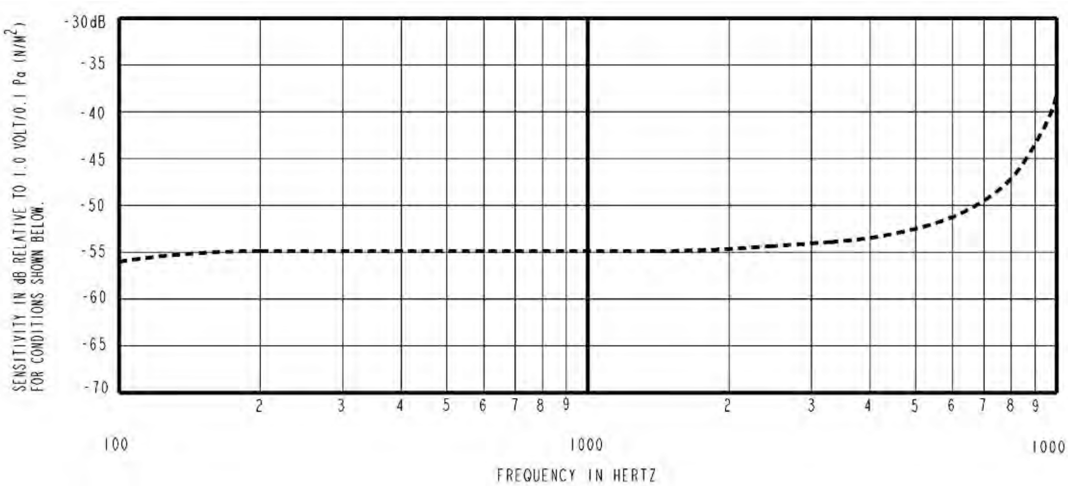

Figure E.1 Microphone sensitivity.

The sensitivity S of -55 dB converts with

$$S = 20 \cdot \log\left(T_F\right)$$

to a Transfer function of:

$$T_F = 1.778\, mV / Pa\ ;$$

The maximum output of the microphone is according to the data sheet

$$V_{mic_{out}} = 400\text{mV}\ ;$$

Thus, the microphone can theoretically record Sound Pressure Levels up to 280 Pa or 140 dB.

E.2 Analog Front End

The output voltage generated by the microphone is too weak to be directly connected to an Analog to Digital Converter; an amplifier is required to increase the signal level. Due to the expected high dynamic range of the microphone signal the input stage is equipped with a Variable Gain Amplifier.

E.2.1 Variable Gain Amplifier

Analog Device's ultra-low noise dual-channel amplifier AD604 provides an adjustable gain of 48 dB. The built-in preamplifier can be adjusted to provide an additional fixed gain of up to 6 dB. This allows SPLs to range from the above 140 dB down to almost 90 dB. The AD604 also features a single-ended analog unipolar gain control with 1 dB gain resolution, which provides a simple method to control all 16 channels in a centralized and accurate way. Figure E.2 depicts the a detail schematic of the AD604 showing the preamplifier (PA) on the top and the Variable Gain Amplifier (VGA) on the bottom. The resistor R (here R23) defines the preamplifier's gain.

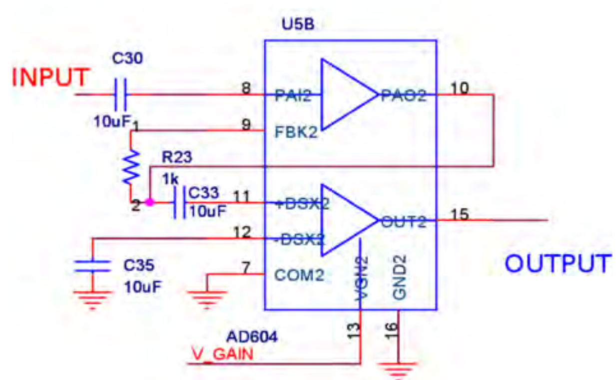

Figure E.2 Schematic for the AD604 amplifier.

E.2.2 Low Pass Filter: LTC 1069

The amplified signal is passed to an adjustable Low-Pass Filter built around Linear Technologies' LTC1069-1. The adjustable cut-off frequency can be clock-tuned up to 10 kHz which is beyond the frequency where the microphone is expected to be used. Due to the lack of good sample files and to avoid removal of important information, the cut-off frequency is currently fixed to its maximum value but can later be controlled adaptively by the Processing Unit. The ability to change the cut-off frequency allows the Array Processor to adapt to the presence of external interferences and helps improving its performance. Regardless of its setting, the LTC1069-1 acts as an Anti-Alias filter for the ADC. The schematic for the LTC1069-1 is shown in Figure E.3.

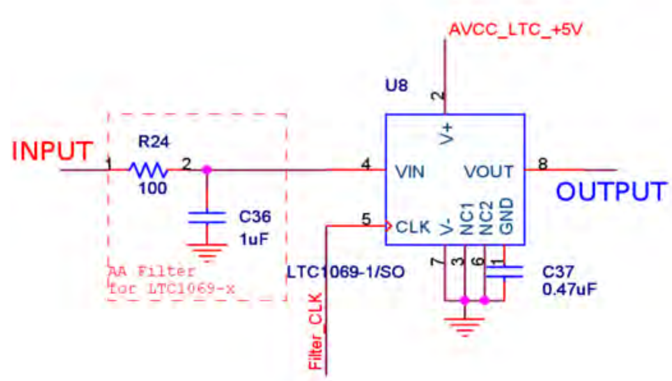

Figure E.3 Schematic for the anti-aliasing filter.

E.2.3 Signal Conditioning

The filtered single ended signal needs to be conditioned to be connected to the differential analog input of the ADC. Texas Instruments' THS4522 is used to convert the single ended signal into a differential one and to adjust to the Common Voltage input of the ADC. It

further adds a small gain to the output voltage of the AD604 - which is around 2 Vpp – to reach the full scale input (2.5 V) of the ADC. The figure below shows a detail of this circuit.

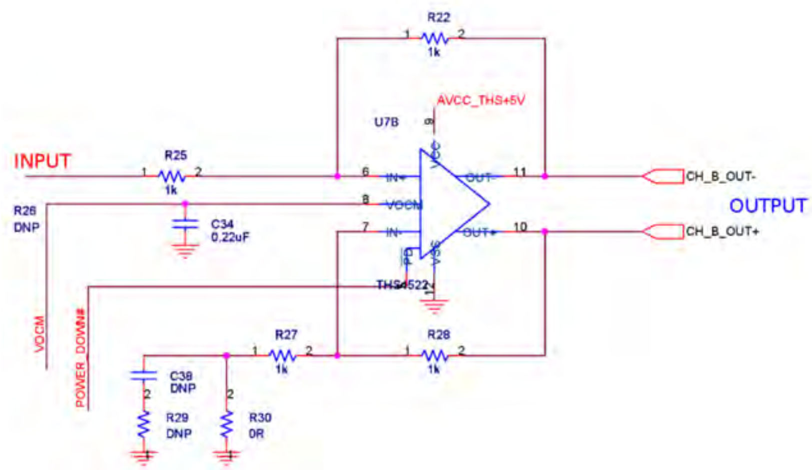

Figure E.4 Single-ended to differential converter.

E.2.4 Variable Gain Control

The current design controls the signal level in the Processing Unit. The analog audio signals of all sixteen channels are recorded and analyzed in the Processing Unit. In case the ambient noise level changes, the Processing Unit adjusts the gain of all sixteen channels using a single point of control. To interface the Digital Processing Unit with the analog front end, Analog Device's AD5160 digital potentiometer is used. The AD5160 is a 256-Position SPI-Compatible Digital Potentiometer. Its 8 bit resolution theoretically allows the gain to be adjusted with a precision of 0.2 dB. Using this control the system can adapt to the change of noise levels.

E.3 Analog / Digital Conversion

Special requirements are necessary for the Analog to Digital Converter. To provide accurate digital signals for the Processing Unit all analog channels have to be sampled at the exact same time. Most multichannel ADCs use a multiplexer at the input and scan sequentially the analog channels. An ADC with this configuration cannot be used in this application.

E.3.1 Analog to Digital Conversion

Texas Instruments' ADS1278 is an octal, simultaneous sampling, 24-Bit delta-sigma Analog-to-Digital Converter. When used in high speed mode the maximum sample rate can be as high as 128 kHz. In the current configuration the sampling rate is set to 97.656 kHz. This sample frequency is sufficient to provide a good spatial resolution. Under the assumption that the speed of sound is 343 m/s the spatial resolution for a planar wave is:

$$R_{spatial} = \frac{343\frac{m}{s}}{97{,}656\frac{1}{s}} = 3.5\text{mm}$$

The spatial difference between two of the sixteen microphones mounted on the circumference of a helmet is between 8 mm and 30 mm. Thus, the signals are sampled with delays between 2 and 8 sample points. This is sufficient to determine the direction of the sound source.

To synchronize the two ADCs, that are required to sample the sixteen channels, the ADS1278 also provides a special SYNC signal. The SYNC pin can be used to synchronize multiple devices within the same clock cycle. The digital output format of the ADS1278 is highly variable and can be configured to different protocols. In the current design the serial data of all sixteen channels is read simultaneously. A detailed schematic of the ADS1278 is given in Figure E.5.

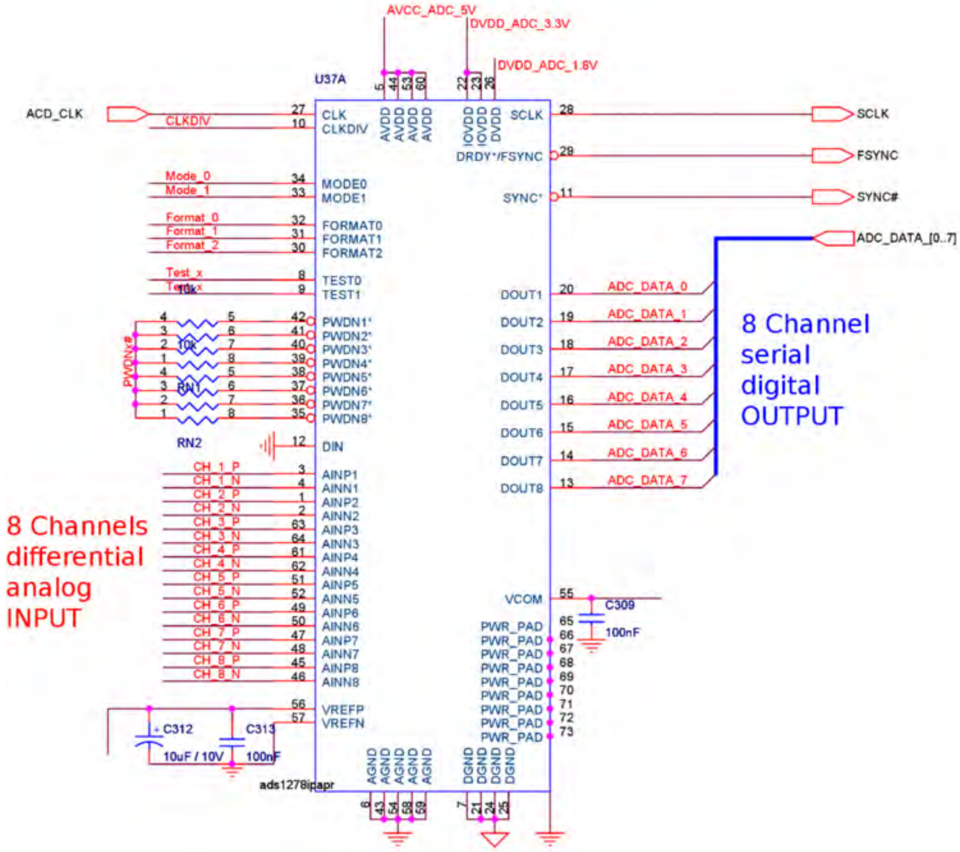

Figure E.5 Schematic for the signal conversion unit.

This page intentionally left blank.

Annex F Digital Signal Processor

F.1 DSP

In the current design Texas Instruments' TMS320C6726B DSP is used. This DSP features a Multichannel Serial Audio Port (McSAP) that can be configured to be sixteen channels wide. Furthermore, it has a dedicated McSAP DMA Bus to provide a fast and reliable transfer of the received data within the device. With the new built-in audio-specific functions for faster calculations of biquad filters and faster FFT instructions this DSP can further speed up the processing.

Attached to the DSP is a sixteen bit wide 64 Mbit synchronous dynamic RAM (SDRAM). The major part of this memory is configured as a circular buffer to concurrently store one second of data for all the sixteen channels. In addition to the SDRAM an external flash based Read-Only Memory (ROM) is attached to the DSP that provides the program and additional configuration instructions during BOOT time. The algorithm that runs on the DSP is explained in detail in Annex D.

F.2 User Interface

The user interface is designed with flexibility in mind. It consists of a programmable device (CPLD) that interfaces on one side with the 16 bit interface bus of the DSP and on the other side with a 16 bit bus. This bus can connect to an external graphical display or to a series of buffers that supply enough current to drive an attached LED. The current configuration uses sixteen LED's. The channel number where the signal was detected is turned on for two seconds. Should other events occur within this period, additional LEDs are turned on for two seconds.

F.3 Power Supply

The system can be powered with a single 6 V to 20 V supply. All necessary lower voltages are created on the Processing Board. The critical voltages for the DSP's core and digital I/Os are sequenced. The core voltage turns on only when the digital I/O voltage is present. The DSP remains in reset mode until both voltages have reached 90% of their configured value

References

[1] NACRE AS. QuietPro User's Manual. Available at: http://www.nacre.no/html/instructions.asp

[2] Winberg, Mathias. Speech Enhancement and Noise Combat. Ph.D. Thesis, Karlskrona/Ronneby, 1998.

[3] Wiklund, Karl. The Cocktail Party Problem: Solutions and Applications. Ph.D. Thesis, McMaster University, 2009.

[4] Kuo, Sen M. and Dennis R. Morgan. "Active Noise Control: A Tutorial Review". Proceedings of the IEEE. 87 (1999): 943-973.

[5] Akansu, Ali N. Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets. Boston, MA: 1992.

[6] Slaney, Malcom. An Efficient Implementation of the Patterson-Holdsworth Auditory Filter. Technical Report 45, Apple Computer.

[7] Ray, Laura R. et al. "Hybrid Feedforward-Feedback Active Noise Reduction for Hearing Protection and Communication". Journal of the Acoustic Society of America. 120 (2006): 2026-2036.

[8] Oinonen, Mika K. Development of Active Noise Control and Voice Communication Systems For Personal Hearing Protectors. D.Tech Thesis. Tampere University of Technology, 2006.

[9] Widrow, B. and Goodlin, R.C. "Adaptive Noise Cancelling: Principles and Applications". Proceedings of the IEEE. 63 (1975): 1692-1716.

[10] Chassaing, Rulph. Digital Signal Processing and Applications with the C6713 and C6416 DSK. Hoboken, NJ: Wiley, 2005.

[11] Kit, Wong Chun. A Decimated Electronic Cochlea on a Reconfigurable Platform. Master's Thesis, Chinese University of Hong Kong, 2006.

[12] Analog Devices. ADSP-21369 EZ-KIT LITE Evaluation System Manual. Norwood, MA: Analog Devices, 2006.

[13] Wiklund, Karl. R-HINT-E: A Realistic Hearing in Noise Test Environment. M.A.Sc. Thesis, McMaster University, 2003.

[14] Analog Devices. AD73322L Datasheet. Norwood, MA: Analog Devices, 2.

This page intentionally left blank.

List of symbols/abbreviations/acronyms/initialisms

ADC	Analog-to-Digital Converter
AGC	Automatic Gain Control
ANC	Active Noise Control
AST	Active Speech Transmission
CASA	Computational Auditory Scene Analysis
cICA	Coherent Independent Components Analysis
CPLD	Complex Programming Logic Device
CPU	Central Processing Unit
DAC	Digital-to-Analog Converter
DMA	Direct Memory Access
dMAX	Data Movement Accelerator
DRDC	Defence Research & Development Canada
DSP	Digital Signal Processor
FCPP	Fuzzy Cocktail Party Processor
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FPGA	Field Programmable Gate Array
FXLMS	Filtered-X Least Mean Square
LED	Light Emitting Diode
I/O	Input/Output
IDE	Integrated Development Environment
IIR	Infinite Impulse Response
McSAP	Multichannel Serial Audio Port
PCB	Printed Circuit Board
SNR	Signal-to-Noise Ratio
SPI	Serial Peripheral Interface
SPL	Sound Pressure Level
ROM	Read-Only Memory
VGA	Variable Gain Amplifier

This page intentionally left blank.

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's document, or tasking agency, are entered in section 8.) Publishing: DRDC Toronto Performing: Simon Haykin, McMaster University, 1280 Main Street West, Gilmour Hall − 306, Hamilton, ON, L8S 4L8 Monitoring: Contracting: DRDC Toronto Craig Burrell, Scientific Authority	2. SECURITY CLASSIFICATION (Overall security classification of the document including special warning terms if applicable.) UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title)

Improving Situational Awareness in Noisy Environments: A helmet−based system for speech enhancement, hearing protection, and shock localization (U)
(U)

4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.)

Karl Wiklund; Simon Haykin; Andreas Freibert; Fan Zhang

5. DATE OF PUBLICATION (Month and year of publication of document.) March 2010	6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 88	6b. NO. OF REFS (Total cited in document.) 14

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Contract Report

8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development − include address.)

Sponsoring:

Tasking:

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.) 14dc04	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.) W7711−088143/A
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document) DRDC Toronto CR 2010−048	10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on the dissemination of the document, other than those imposed by security classification.)

Defence departments in approved countries TTCP and NATO countries

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Canadian and foreign defence departments

DOCUMENT CONTROL DATA

(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) Soldiers in the field work in an acoustically complex environment. Spoken words of particular interest may be embedded in other less important speech or environmental noise. Loud engine noise and impulsive sounds may also threaten not only the soldier’s awareness of their surroundings, but their hearing as well. Just as importantly, some sounds (including impulsive events), are critical for situational awareness, which the soldier must hear or be made aware of. To address these competing problems, we have introduced the Smart Helmet system which combines hearing protection, speech enhancement and source localization. This report outlines the overall system design, presents lab−testable prototypes, and describes more integrated software models that may be used for further investigation. Areas needing further research are also described.

(U) Soldats dans le champ de travail dans un environnement acoustique complexe. Paroles d'un intérêt particulier peuvent être incorporés dans d'autres discours moins importantes ou le bruit dans l'environnement. Bruit de moteur fort et sons impulsifs peuvent menacer aussi non seulement le soldat conscience de leur environnement, mais aussi leur audience. Tout aussi important, certains sons (y compris les événements impulsifs), sont essentiels pour la conscience, qui le soldat doit entendre ou être mis au courant des. Pour résoudre ces problèmes de concurrents, nous avons introduit le système de casque dynamique qui associe la protection auditive, amélioration des discours et de la localisation de la source. Ce rapport décrit la conception globale du système, présente des prototypes de laboratoires−tests et décrit les modèles de logiciel plus intégrés qui peuvent être utilisées pour l'enquête supplémentaire. Les domaines nécessitant davantage de recherche sont également décrites.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

(U) Communication; Signal Processing; Speech Enhancement; Hearing Protection; Situational Awareness; Cocktail Party Problem; Computational Auditory Scene Analysis; Active Noise Control; Acoustic Shock Detection; Acoustic Shock Localization

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca