DSP GROUP
at home in your wireless world

White Paper

Version 1.1

Redefining Voice Quality and Intelligibility in Mobile Voice Communications

November 2012

Headquarters
USA
2580 North First Street
Suite 460
San Jose, CA 95131
Tel: (408)986-4300
Fax:(408)986-4323

Israel
5 Shenkar Street
Herzelia, Israel 46725
Tel: 972-9-952-9696
Fax: 972-9-954-1234

Europe
Germany
Nurnberg
Stromerstr. 7
90443 Nurnberg
Tel: +49-(911)-2001-1350

Scotland
Livingston
Geddes House
Kirkton North
Livingston EH54 6GU
Tel: +44-(0)1506-407-730

Switzerland
Binzstrasse
44 CH-8045 Zurich
Switzerland
Tel: +41-(44)-465-1900

APAC
China
11/F, 909 Cheung Sha Wan Rd.
Kowloon
Hong Kong
Tel: +852-(3965)-5888

DSPS Solutions
Rm. 24, 3/F, Blk. B,
Proficient Ind. Centre,
6 Wang Chiu Rd.,
Kowloon Bay,
Kowloon, Hong Kong
Tel: +852-(2750)-7325
Fax: +852-(2305)-0640

Ascend Technology Inc. Shenzhen
Room 1303, New World Center, No. 6009
Yitian Road, Futian District,
Shenzhen City, China
Tel: +86-755-82024598
Fax: +86-755-23982986

Ascend Technology Inc. Beijing
Rm 607, Hui Huang International Center,
1st Place, Shangdi 10th Road,
Haidian District, Beijing City, P.R. China

Japan
Tokyo
1-29-1 Nishi-Gotanda
Shinagawa-Ku Tokyo 141-0031
Tel:+81-(3)-3493-3050

Tomen Electronics
8-27, Kohnan 1 Chome,
Minato-ku, Tokyo
108-8510, Japan
Tel: +81-(3)-5462-9619
Fax: +81-(3)-5462-9686

Korea
Daesung Semiconductors
140-848, RM 401
Wonhyo BD. 46-1
Wonhyo-ro 3ka, Yongsan-gu,
Seoul, Korea
Tel: +82-(2)-3272-7300
Fax: +82-(2)-712-4632-3

Taiwan
Ascendtek Electronics, Inc
11F-7, No. 77, Sec. 1
Hsin Tai Wu Rd.,
Hsi Chih
Taipei Hsien,
Taiwan, R.O.C
Tel: 886-2-26988696
Fax: 886-2-86982138

India
Bangalore
Information Technology Park
Nagawara Village Kasaba Hobli
Bangalore 560045
Tel: +91 80 4024 8399

File: hdclear white paper v 1.1.doc saved 05/11/12 14:29

TABLE OF CONTENTS

1 INTRODUCTION 4

2 EMERGING MOBILE TECHNOLOGIES 5
2.1 Overview 5
2.2 Noise Cancellation 5
2.3 High-Definition (HD) Voice 6

3 THE HDCLEAR SOLUTION 7
3.1 Overview 7
3.2 Use Cases 7

4 HDCLEAR TECHNOLOGY 11

5 NOISE CANCELLATION EVALUATION 13

6 AUTOMATIC SPEECH RECOGNITION (ASR) 14

7 DSP GROUP'S VOICE PROCESSING OFFERING 16

8 SUMMARY 17

List of Figures 18

List of Tables 19

1 Introduction

Over the past few years, the craze for multimedia has eclipsed the essence of communicating: voice quality and intelligibility. While multimedia advances drove consumer demands to feed the industry's coffers, new voice technologies were largely neglected. As "What did you say? I can't hear you!" and "Speak slowly. You're coming across garbled." became painfully widespread grievances, consumers put voice quality and intelligibility back among their top requirements for mobile handsets.

Today, mobile handset suppliers and Mobile Network Operators (MNOs) have come to recognize that high-quality voice calls not only satisfy customer demands but also enable advanced services, such as Automatic Speech Recognition (ASR), and help optimize network usage by reducing inefficient traffic.

Voice enhancement requires a voice processor that provides advanced noise cancellation. Removing background noise significantly improves the accuracy rate of ASR and voice-controlled applications for smartphones and other Consumer Electronics (CE) devices. DSP Group's HDClear™ technology delivers unparalleled voice quality and intelligibility to enable consistently high performance, even at the harshest noise levels.

This paper describes the benefits of HDClear technology, and evaluates its performance against leading competitive solutions.

2 Emerging Mobile Technologies

2.1 Overview

Several years ago, GPS was an emerging technology in mobile devices. Nowadays, the availability of accurate and fast GPS is taken for granted. Similarly, other new technologies such as a MEMS gyroscope, an accelerometer and a built-in compass have emerged side by side with smartphone proliferation to provide users with a rich set of smartphone applications.

Just as GPS led to the development of better and easier social networking applications, voice enhancement through technologies such as noise cancellation and high-definition (HD) voice is driving the growth of new, friendlier applications, offering MNOs the opportunity to deliver new, advanced services. Among these services are speech-to-text applications activated by a single key, background noise replacement with sounds or music and even, in a few years' time, on-line universal translation.

2.2 Noise Cancellation

Noise cancellation delivers multiple value-added benefits to the near-end user (the caller), the far-end user (the person receiving the call), and to MNOs. Among these benefits are:

- A better voice experience for the far-end user when the near-end user is in noisy surroundings at home, in the office and on-the-go
- Privacy for the near-end user by eliminating noise that could potentially identify the caller's location
- Support for speakerphone mode
- Significantly higher accuracy rate for Automatic Speech Recognition (ASR) and other voice-controlled applications that are becoming a major human control interface for tablets, computers, Smart TVs and CE devices, and are being incorporated by industry giants like Google, Apple and Microsoft into their operating systems
- Higher network efficiency by maximizing its capacity through reducing inefficient traffic. This is achieved by more efficient codec operation (network bandwidth is not wasted on noise).

Traditional noise cancellation suffered from tradeoffs between the degree of noise reduction and voice quality: the higher the noise reduction levels, the greater the potential for voice distortion. Attempting to minimize the tradeoffs, engineers developed effective algorithms to reduce very specific types of background noise in limited settings, using standard handset modes of operation. Unfortunately this focus rendered them ineffective to achieve consistent performance in environments where other types of noise prevailed, and where the speaker chose to use other modes of operation, such as speakerphone mode.

Noise cancellation took a leap forward with the introduction of a second microphone in smartphones. Both microphones operate in similar manner to the human auditory system. However, this remarkable auditory intelligence capability does not provide sufficient noise cancellation to eliminate all background noise during voice calls. In noisy public venues, while driving or riding on public transportation, or even at home when, for instance, music is turned up loud, human hearing cannot satisfactorily block out background noise.

2.3 High-Definition (HD) Voice

HD voice is a standardized technology that enables MNOs to clearly differentiate their services with high-quality voice calls. Tier 1 operators began educating the mobile communications market on the benefits of HD voice when it was first introduced in 2009. It delivers much clearer, more natural sound, and removes the strain of listening and being understood. According to recent information from the Global Mobile Suppliers Association (GSA), 41 commercial mobile HD voice networks are active in 33 countries.[1]

HD voice technology is based on wideband codecs that use higher sampling rates than traditional narrow-band codecs. HD voice transmission requires both network support and an HD voice-enabled handset. With noise cancellation, HD voice technology operates more effectively.

[1] Global Mobile Suppliers Association (GSA) report: 41 commercial mobile HD Voice networks launched in 33 countries, May 21, 2012

3 The HDClear Solution

3.1 Overview

HDClear technology extends voice data acquisition from up to three sources and then applies DSP Group's 3D-Vocal algorithm on the acquired data. This unique combination delivers measurably higher and more consistent voice quality and intelligibility to mobile handset users compared to competing technologies.

The sources from which HDClear acquires voice signal are audio microphones, similar to competing technologies. After the combined input from the microphones is processed, the 3D-Vocal algorithm is applied to the acquired data to perform a multitude of voice processing tasks that include echo and background noise cancellation, loudness equalization and general voice enhancement.

The improvement in voice quality and intelligibility is consistently higher in HDClear-based systems, even when the mobile handset is used in noisy surroundings, either in handheld or speakerphone mode. The traditional tradeoff between voice intelligibility and noise reduction, known to distort speech so that it loses its human quality and becomes difficult to understand, is unnoticeable in HDClear-based systems. HDClear voice processing operates at low power, using a small footprint that can easily be integrated into any mobile handsets.

When compared with competing voice enhancement technologies, HDClear-based systems achieve measurably higher results (see section 5 for the results of acoustic laboratory) in a variety of use cases.

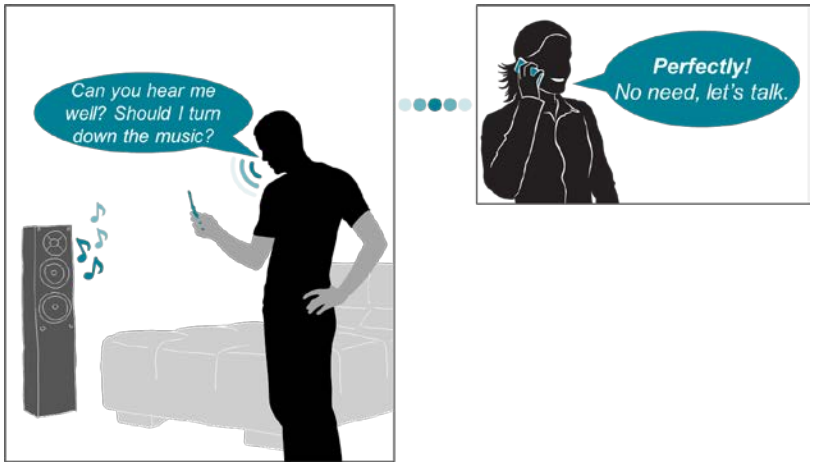
3.2 Use Cases

HDClear technology supports:

- Handheld and speakerphone mode
- Conference calls that can now be conducted even in noisy public venues
- High accuracy rate for ASR in all types of noisy environments and levels
- Background music and/or sounds during ongoing conversations, benefiting consumers and providing operators with a value-added edge
- Unique human interface with heightened sensitivity to support tapping control over the handset to enable a range of actions, such as activating mute or unmute, redirecting calls and activating ASR during a phone call

Value-added use cases are illustrated in Figure 3-1.

Speakerphone Mode



HDClear takes the strain out of hearing, being heard and being understood in any surroundings and operating mode.

Handheld Mode

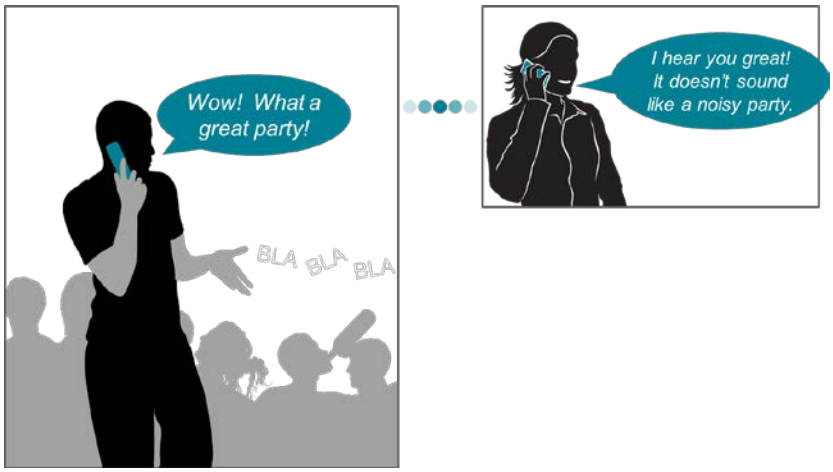
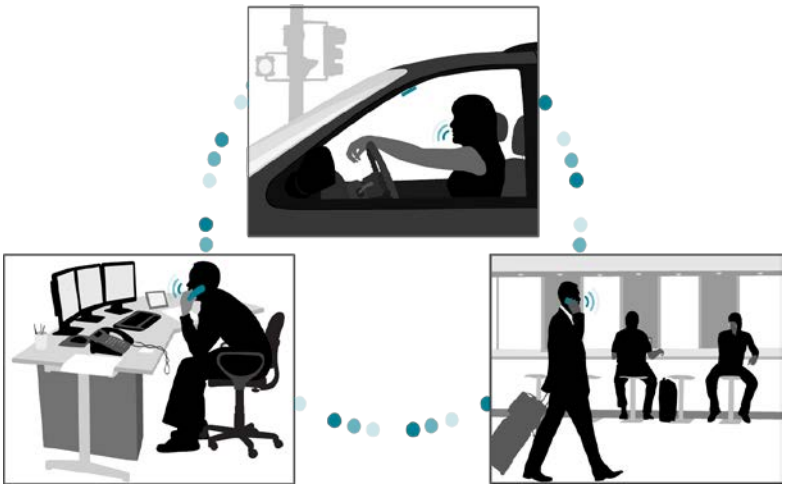


Figure 3-1: Value-Added HDClear Use Cases (Cont'd)

Conference Call

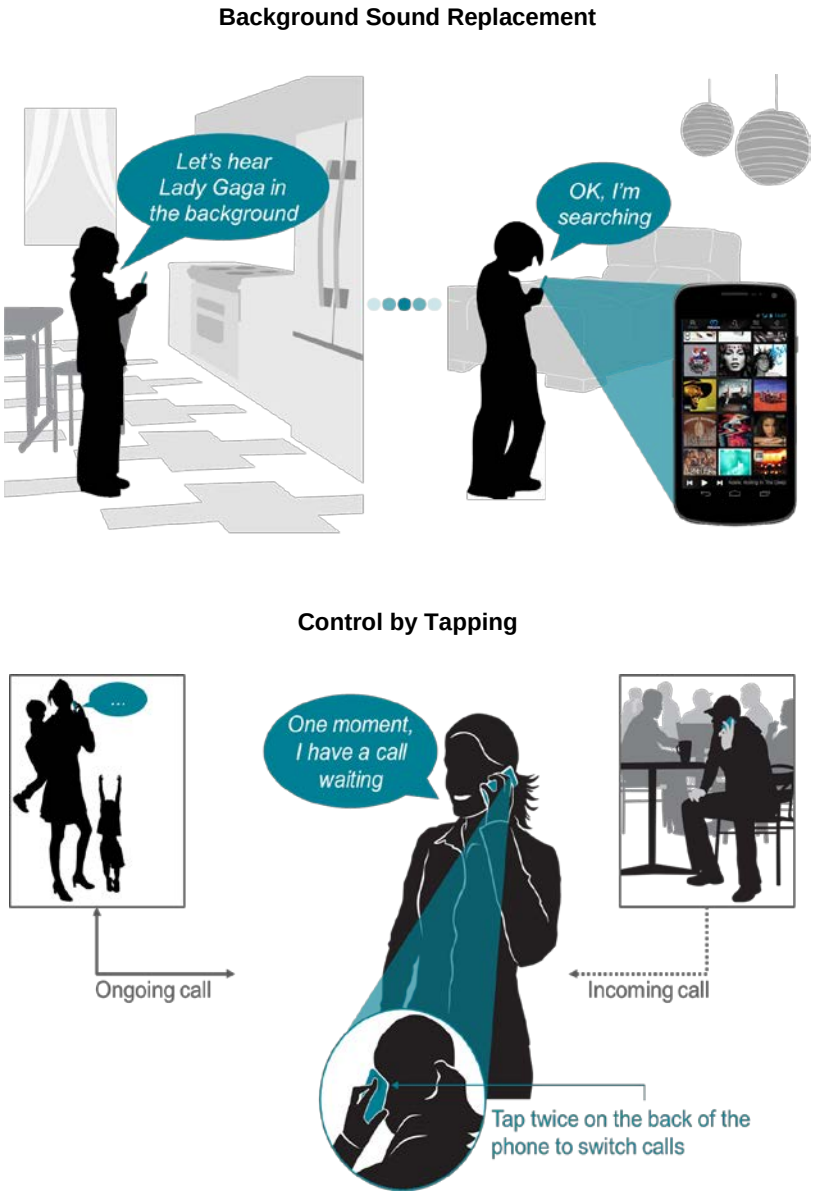


Conference calls can be taken on-the-go, in an office or in a noisy public venue without compromising voice quality or intelligibility.

ASR-Supported Call

ASR delivers higher accuracy rates to enable voice-activated applications, even when background noise is acute.

Figure 3-1: Value-Added HDClear Use Cases (Cont'd)



Background sound replacement, with music or other sounds, provides a potential revenue generator for operators.

HDClear-based phones offer customers an innovative and simple human interface that lets them tap on the back on their phones for control.

Figure 3-1: Value-Added HDClear Use Cases

4 HDClear Technology

HDClear automatically detects the various mobile handset use cases and optimizes performance accordingly. HDClear has been fine-tuned not only to optimize noise elimination and automatically adjust to the use case but also, no less important, to maximize cost/performance by implementing sophisticated software-based solutions.

The HDClear 3D-Vocal algorithm extracts features that characterize the speech source and distinguish between the sound components that belong to required speech vs. ambient noise. Thus, the performance level of HDClear remains very high regardless of the type and level of noise.

The 3D-Vocal algorithm is also used for ASR to preprocess the voice signal in noisy environments, significantly improving the accuracy rate.

HDClear has been rated for its ability to reduce far-end noise; i.e., noise transmitted to the receiving party. An example of how HDClear affects far-end noise is shown in Figure 4-1. The upper waveform illustrates the superposition of speech (S) and ambient noise (N), while the lower waveform shows the resulting speech signal after HDClear processing. Figure 4-2 shows a spectrogram, where the upper graph presents the superposition S+N, and the lower spectrogram shows the resulting speech signal after HDClear processing.

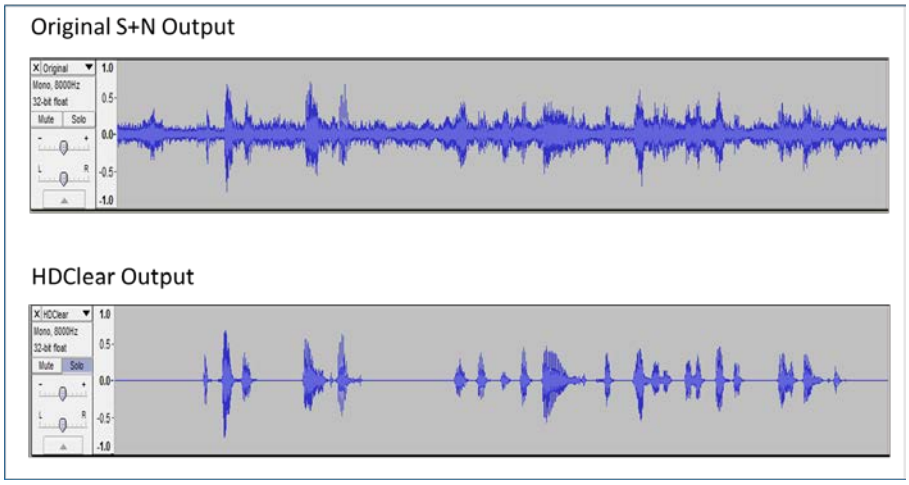


Figure 4-1: Typical HDClear Results on Speech and Ambient Noise

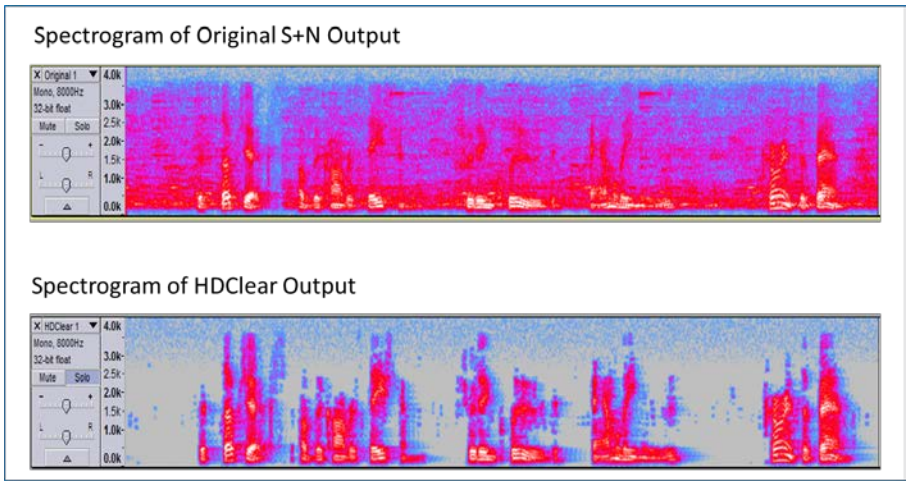


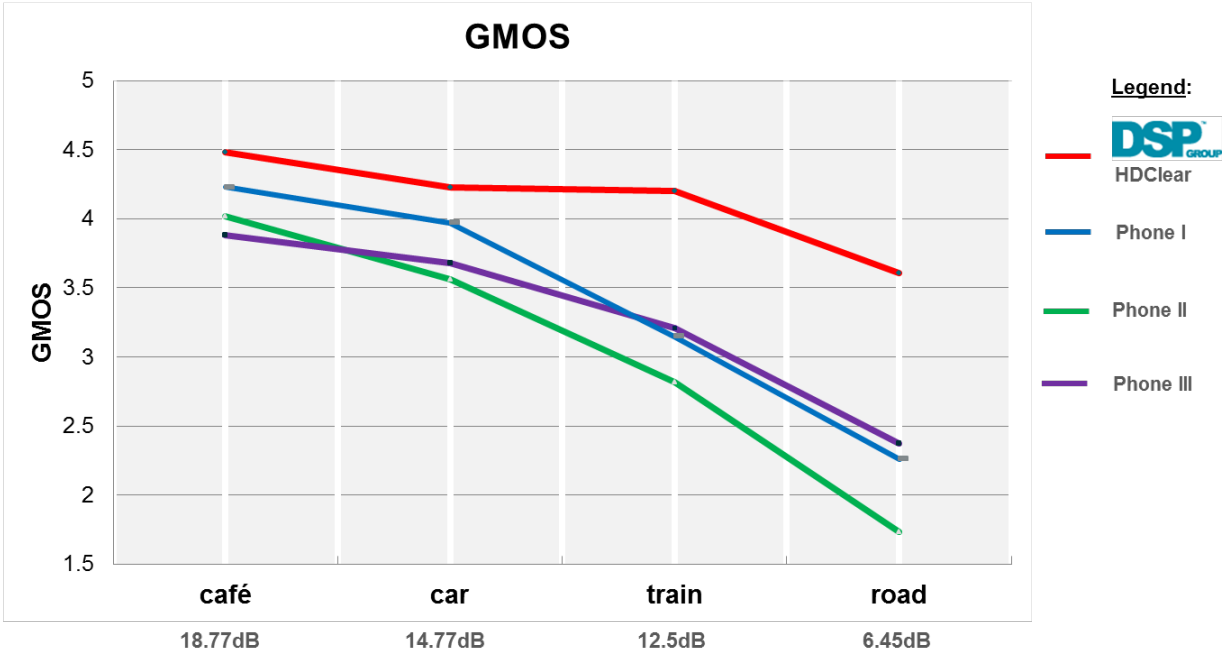
Figure 4-2: Spectrogram of Typical HDClear Results on Speech and Ambient Noise

5 Noise Cancellation Evaluation

This section provides some measurement results achieved with HDClear for objective test procedures. Several standards deal with quantitative evaluation of voice signals in noisy environments. ITU-T P.835 addresses subjective, mean opinion score (MOS) based, testing for noise reduction systems, and also includes an additional Appendix III for non-stationary noise signals. Objective test procedures for evaluation of noise reduction are given in ETSI EG 202 396-3 (3QUEST Tests) and ITU-T G.160 Appendix II. Table 5-1 explains general MOS (GMOS) ratings. Figure 5-1 shows how HDClear faired against other noise reduction systems in different scenarios when tested by 3QUEST methodology in GMOS conditions.

Table 5-1: General MOS Scores (GMOS)

SPEECH QUALITY	RATING
Excellent	5
Good	4
Fair	3
Poor	2
Bad	1



- Results of the 3QUEST Tests - Performed at HEAD Acoustics Lab, June & October 2012
- Score difference of 0.2 is a "significant difference"
- Phone I, II & III data refers to tests conducted on three smartphone brands (listed alphabetically): Apple iPhone 5, HTC One X and Samsung Galaxy S III

Figure 5-1: HDClear Performance vs. other Systems in GMOS Conditions

Remarkably, performance of HDClear remains above the threshold of acceptable user experience for all noise conditions, maintaining almost identical performance regardless of the environment. This achievement is a huge advantage over competing systems.

6 Automatic Speech Recognition (ASR)

Although ASR has been used in commercial applications for over two decades, only in recent years has it started to gain popularity, especially with the introduction of the "Siri" personal assistant on iPhone 4S devices, S-Voice for Samsung's Android smartphones and the Google Now service. This popularity can be attributed to a number of factors: advances in ASR, cloud computing that enables very high processing power with more robust and reliable algorithms, noise cancellation that reduces ambient noise, and language processing technologies. ASR is now considered mandatory in smartphones.

While ASR-enabled smartphone applications work well in quiet environments, their performance tends to degrade drastically in the presence of background noise. They have not been optimized for use in noisy cafés, on public transportation, or even when walking on a busy street.

According to Scientific American[1]:

> "Irregular background noise, wind and variable distance from mouth to microphone all make transcription perfection on a cell phone a towering task—and the results are much less accurate".

HDClear technology enables ASR to achieve a far better accuracy rate by suppressing background noise while preserving the natural voice of the speaker with only minor distortion, even in noisy environments. The degradation is so minor that the user experience is hardly affected when operating applications such as Siri or a text message dictation application in noisy public venues.

ASR is faced by challenges that may be somewhat different from "pure" noise reduction methodology implemented solely to improve the user experience. ASR algorithms usually have their own noise reduction sensitivity and may suffer from voice distortion/impairment that is a noise reduction tradeoff. Thus, it is important to evaluate ASR performance when noise reduction algorithms operate.

Figure 6-1 shows the results of the ASR accuracy score based on the following conditions:

- Vocabulary: Harvard sentences (http://www.cs.columbia.edu/~hgs/audio/harvard.html) played to the mouth of HATS
- Microphone location: reference microphone placed next to the primary microphone of the handset to measure the SNR Input level
- Noise Type: babble (party, café), using noise simulation system
- Speech recognition engine: Siri

While the performance is identical in a quiet environment, there is a dramatic difference in a noisy environment. HDClear-based performance remained consistent up to 7 dB with only minor deterioration (from ~85% to ~80). The iPhone 4S deteriorated in a linear fashion, and at 7dB, it was down to ~30%! The big added value of HDClear in this case is the fact that regardless of the noise level, performance degradation is minimal, enabling users to depend on ASR anywhere.

[1] Scientific American™, "Why Siri is Still the Future", David Pogue, July 31, 2012

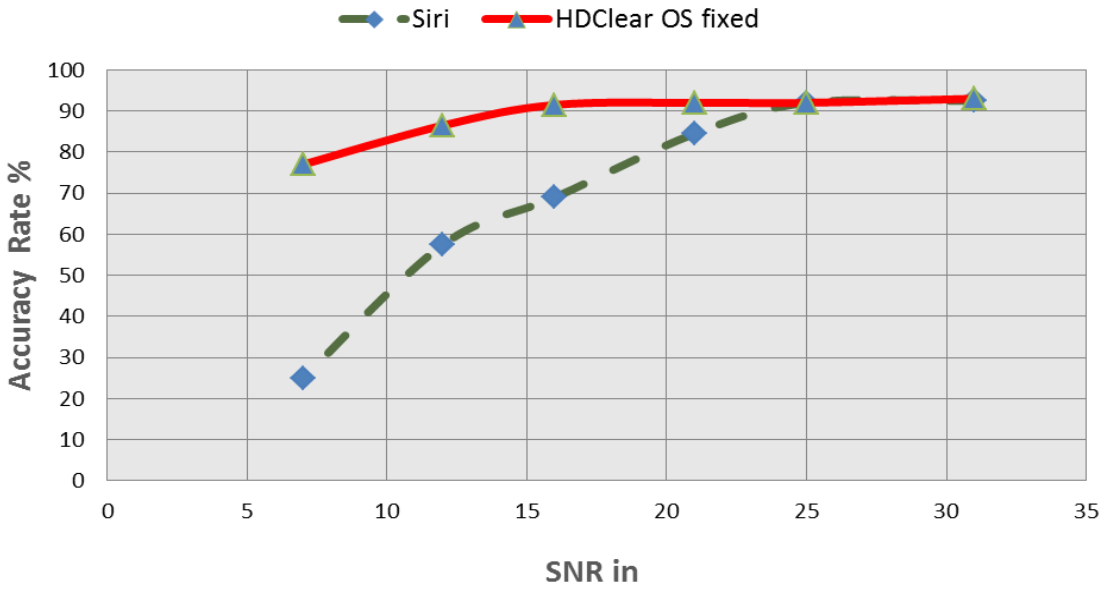


Figure 6-1: ASR Performance Comparison on iPhone 4S Using Siri with HDClear vs. Siri without HDClear

7 DSP Group's Voice Processing Offering

The DSP Group HDClear-based voice processing offering supports an extensive range of features:

- Narrowband and wideband voice
- Background noise cancellation and voice conditioning for the far-end user
  - Operating modes include: handheld (handset), hands-free (speakerphone)
  - Best-of-class speakerphone mode that operates in all venues, with active multiple microphones, a unique VSensor and AEC noise cancellation for the far-end user
- Ability to handle a broad spectrum of coherent/non-coherent and stationary/non-stationary noise sources, including:
  - Café
  - Car
  - Party
  - Road
- Improved accuracy rate for a variety of ASR-based applications
- FlexiSpeech – reduced/faster speech speed to ease listening and assist in voice message retrieval
- HD voice listening experience for the best, undistorted sound via speaker conditioning and user-customized frequency response

8 Summary

Noise cancellation is critical in today's mobile communications devices to enhance voice quality and intelligibility.

DSP Group's HDClear technology is implemented in an architecture optimized for mobile devices. It achieves measurably higher and consistently better results than competing technologies, even in conditions of acute background noise, when evaluated by using sophisticated laboratory equipment and industry-standard means of measurement. Traditional noise reduction techniques operate in similar fashion to the human auditory system, leveraging two inputs that are implemented as two microphones. However, even this remarkable auditory intelligence capability cannot completely remove background noise without compromising voice intelligibility.

HDClear employs DSP Group's 3D-Vocal algorithm that extends data acquisition to significantly improve background noise removal. HDClear provides clear and consistent voice quality, measurably higher than competing technologies.

LIST OF FIGURES

Figure 3-1: Value-Added HDClear Use Cases (Cont'd) 8

Figure 4-1: Typical HDClear Results on Speech and Ambient Noise 12

Figure 4-2: Spectrogram of Typical HDClear Results on Speech and Ambient Noise 12

Figure 5-1: HDClear Performance vs. other Systems in GMOS Conditions 13

Figure 6-1: ASR Performance Comparison on iPhone 4S Using Siri with HDClear vs. Siri without HDClear .. 15

LIST OF TABLES

Table 5-1: General MOS Scores (GMOS).....13