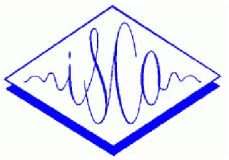

INTERSPEECH 2012
ISCA's 13th Annual Conference
Portland, OR, USA
September 9-13, 2012

Demonstration of Advanced Multi-Modal, Network-Centric Communication Management Suite

Victor Finomore, Jr.[1], John Stewart[2], Rita Singh[3], Bhiksha Raj[3], & Ron Dallman[4]

[1]Battlespace Acoustics Branch, Air Force Research Laboratory, Wright-Patterson AFB, OH, USA
[2]Independent Contractor, Dayton OH, USA
[3]Language Technologies Institute, Carnegie Mellon University, Pittsburgh, PA, USA
[4]Ball Aerospace, Wright-Patterson AFB, OH, USA

victor.finomore@wpafb.af.mil,stewartjm@gmail.com,rsingh@cs.cmu.edu,bhiksha@cs.cmu.edu,
ron.dallman.ctr@wpafb.af.mil

Abstract

The Multi Modal Communication tool was developed to improve mission effectiveness for Command and Control operators. This communication management suite combines mature communication technology to provide the operator with an intuitive display that allows on-demand access to mission critical communication data.

Index Terms: Multi-Modal Display, Communication Interface Automatic Speech Recognition,

1. Introduction

The success of a mission relies heavily on the ability of Command and Control (C2) operators to effectively communicate with their assets and one another. C2 operators monitor communications which allows them to construct a real-time picture of the dynamic battlefield in order to coordinate and direct assets based on its evolution. Traditionally this communication between operators has been accomplished primarily through voice communication. However, this means of communication imposes a high degree of workload on the operator since they can be overloaded with multiple assets speaking at the same time thus reducing message intelligibility [1]. An additional problem is the transient nature of radio transmissions, in that the recipient has a fleeting opportunity to comprehend information being transmitted before it is missed altogether.

Due to these issues with radio communication and the growing need to communicate to a growing number of distributed operators, the use of real-time text-based communication (chat) has proliferated C2 centers. Warfighters report that chat is an essential tool to accomplish their current mission; however, there are potential drawbacks to this collaborative technology. First, chat displays are not well integrated into legacy systems thus forcing operators to divert visual attention away from the tactical display. Additionally, monitoring multiple chat rooms can be a challenge since the operators have to divide attention between multiple conversations, which can result in missed or misunderstood information [2].

In an attempt to better equip C2 operators with collaborative tools that could aid in their mission, researchers and engineers at the Battlespace Acoustics Branch in the Air Force Research Laboratory developed a network-centric communication management suite to help improve communication performance. Extensive interviews with C2 operators from a wide range of mission sets were conducted to understand difficulties. A survey of C2 operators was also given to assess the perceived usefulness of commercial-off-the-shelf collaborative tools [3]. The survey found the two highest ratings for potential technologies were data capture/replay tools and chat/messaging tools. This information led to the development of a program called Multi-Modal Communication (MMC), which combines mature communication technologies into an integrated communication management system to provide C2 operators with the tools to be more effective in their mission.

2. Description of MMC

MMC provides operators with the ability to manage communication from voice and text-based systems in a single intuitive, dynamic display. It captures, records, and displays radio (via automatic speech recognition) and chat communications to the operator so that they have on-demand access and control over all current and past information. In addition to aiding in the retrieval of information, speech intelligibility over the radio channels is increased by spatially separating each of the radio channels to virtual locations around the operator via their headphones.

2.1 Radio

2.1.1 Intelligibility

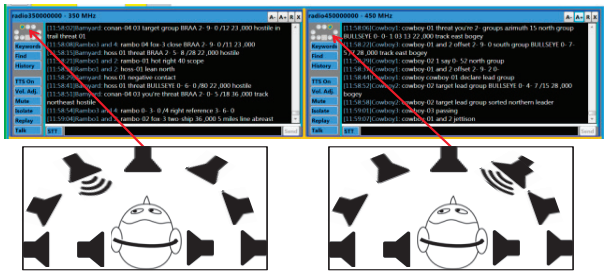

Figure 1. Graphic showing spatial audio display

A Spatial Audio display was implemented into MMC to improve radio intelligibility and reduce the workload associated with monitoring multiple channels. Brungart and Simpson [4] systematically evaluated different configurations of digital filters, called head-related transfer functions (HRTS), which take advantage of binaural cues used in sound localization and replicate them over headphones. The study found an optimal set of HRTS that

maximizes the number of talkers while preserving intelligibility. These filters are included into MMC to allow users in real-time to place radio channels into one of nine virtual locations for which sounds will be rendered. Figure 1 depicts two MMC radio windows with pop-out graphics showing the sounds being rendered in virtual locations.

2.1.2 Capture and Display

As previously mentioned an issue with verbal communication is that it is perishable, there is only a brief chance to hear the message before it is gone. In addition, there are a number of competing factors that reduce the likelihood that speech communication will be accurately heard; such as ambient noise, loss of attentional focus of the listener, or masking by competing speech signals. To combat this issue, MMC employs near real-time speech-to-text transcription for verbal communication and displays the transcribed speech as text in a chat-like window thus preserving the message for later recall. These messages are time-stamped as well as formatted to operators' preference for reading messages in order to aid in retrieval and comprehension of information.

In addition to transcribing the message, MMC also captures and stores the original audio transmission and links it to the text. Operators simply need to select a line or lines of text to re-listen to the message. During playback mode, operators have the option to increase or decrease the speed of the audio messages to twice as fast as real-time or slowed down to half speed without disrupting the integrity of the message. In addition to changing the speed of the message, all silences in between messages are removed so that re-play can occur in faster than real time. The ability to read verbal messages or hear them again reduces the need to request a message to be repeated, thus saving time and communication bandwidth but more importantly providing the operator with quick and easy access to find and verify critical information. Figure 2 shows the transcriptions of audio messages and the playback feature.

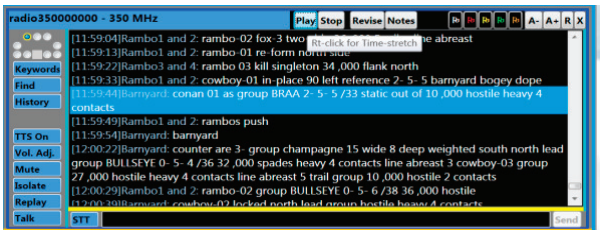

Figure 2. Radio window with automatic speech recognition and playback feature

2.2 Chat

Instead of a separate display added to an already busy workstation to present chat rooms, MMC integrates text-bases messaging into a common interface. The chat windows as seen in Figure 3, allows operators to type or speak their message and transmit them as text. Speech-to-Text was included into the chat window to combat the issue people are having with the time it takes to type out messages [5] or the time and effort in abbreviating phrases or the translation of them [6]. Text-to-Speech was also included into the chat window to allow operators to hear messages when their visual attention is need elsewhere. This feature also allows chat windows to be minimized to save space on the display.

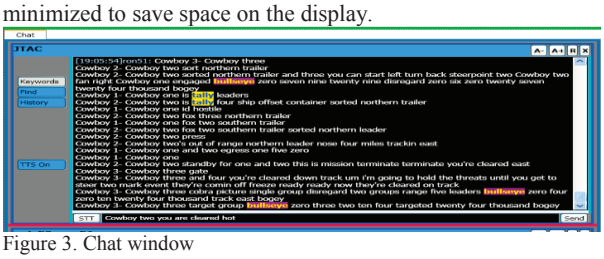

Figure 3. Chat window

2.3 Accessing Data

A number of features are included in MMC to assist operators in finding the information they need. These features are found in both the voice and chat windows and can be utilized locally for the specific channel or globally across channels. The Find function allows operators to search for specific words or phrases to allow operators to quickly identify critical information. The Keyword feature allows operators to define words to be automatically highlighted as they are transcribed to aid operators in finding information based on time. The History feature is used to go to a specific time or to go back from the current time.

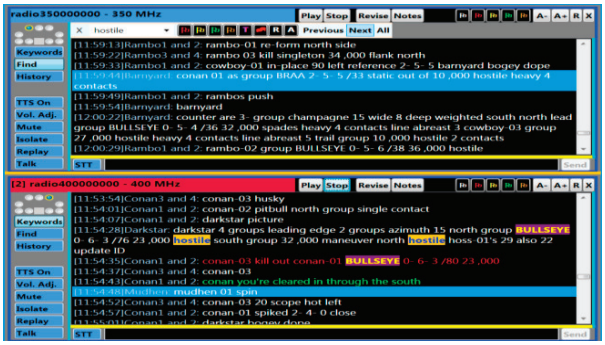

Figure 4. Radio windows demonstrating tools to access data

3. Conclusion

The combination of these features provides operators with the tools necessary to monitor multiple communication channels for critical information and make quick and accurate decisions, affording operators greater situation awareness while also reducing their perceived mental workload.

4. References

[1] Bolia, R. S. (2003). Effects of spatial intercoms and active noise reduction headsets on speech intelligibility in an AWACS environment. *Proceedings of the Human Factors and Ergonomics Society, 47,* 100-104.

[2] Ramachandran, S., Jansen, R., Bascara, O., Carpenter, T., Denning, T., Sucillon, S. (2009). After action review tools for team training with chat communications. *Interservice/Industry Training, Simulation, and Education Conference*.

[3] Nelson, W. T., Boila, R. S., Vidulich, M. A., & Langhorne, A. L. (2004). User-centered evaluation of multi-national communication and collaborative technologies in a network-centric air battle management environment. *Proceedings of the Human Factors and Ergonomics Society, 48,* 731-735.

[4] Brungart, D. S., & Simpson, B. D. (2005). Optimizing the spatial configuration of a seven-talker speech display. *ACM Transactions on Applied Perception, 4,* 430-436.

[5] Baltes, B. B., Dickson, M. W., Sherman, M. P., Bauer, C. C., LaGanke, J. S. (2002). Computer-mediated communication and group decision making: a meta analysis. *Organizational Behavior and Human Decision Processes, 87,* 156-179.

[6] Knott, B. A., Nelson, W. T., Brown, R. D., Dukes, A. W., Boila, R. S. (2007). Assessment of a novel chat messaging interface for rapid communication in tactical command and control. *Proceedings of the Human Factors and Ergonomics Society, 51,* 1163-167.