
ATC-10B Vocoder - Verifying Equivalence of Floating and Fixed Point Implementations

CONTENTS

- INTRODUCTION
- TEST APPROACH
 - Interoperability
 - Source Code Review
 - Comparative Listening Tests
 - Performance: Background Noise, Channel Bit Errors, and Truncation Mode
- TEST CONDUCT

This white paper outlines a test approach for verifying the functional equivalence of the ATC-10B vocoder floating and fixed point implementations.

INTRODUCTION

The ATC-10B voice compression algorithm developed by Digital Voice Systems, Inc (Westford, MA) is the approved voice compression algorithm specified by DO-224A for use in air traffic voice communication. The Federal Aviation is leading the effort to specify and standardize new communication equipment under an effort known as the Next Generation Communication (NEXCOM) program. The ATC-10B vocoder is one of the components of the NEXCOM system. The FAA has also been working with the International Civil Aviation Organization (ICAO) to coordinate international acceptance of the new standards.

The ATC-10B algorithm was originally evaluated and selected using software written and compiled for a floating point processor. A fixed point implementation has subsequently been developed to provide a bit exact version to facilitate testing, validation and DO-178B certification efforts.

As part of the DO-178B certification effort, the equivalence of the fixed and floating point implementations must be supported. This white paper outlines a proposed test and evaluation approach to support equivalence of the two implementations.

TEST APPROACH SUMMARY

The test approach will compare the floating and fixed point vocoder implementations in the following areas:

- Interoperability
- Source Code Review
- Voice Quality
- Performance

Interoperability

Interoperability testing ensures that voice that has been encoded by one implementation can be decoded by the other implementation. It verifies interface equivalence.

The compressed voice output stream from the floating point implementation is provided as an input stream to the fixed point implementation, and vice versa. The decoded voice output is compared with the speech input to insure the voice was faithfully reproduced, i.e., sounds the same. The intent is to validate interoperability and not voice quality. Separate voice quality tests will be conducted (see below).

Fortunately, this type of testing has already been conducted on a one or two speech segments with positive results. Additional testing will be done on at least 50 more speech segments. These voice segments will be randomly selected from the pool of Air Traffic voice samples originally used to evaluate the floating point implementation.

Source Code Review

A general review of the source code for both floating-point and fixed-point implementations ensures the code implementations are comparable. The objective of this review is to increase the level of confidence that the two implementations are equivalent, i.e., material differences in the two implementations are not evident.

Voice Quality

Voice quality is the most important attribute when comparing the floating and fixed point implementations. A blind AB comparative listening test will be performed using a sufficient number listeners and speech samples.

The comparative listening test will use at least 8 listeners and 200 speech sample pairs. Each listener will be presented the speech sample pairs and will be asked to select the speech sample from the pair that sounds better. Both speech samples in the pair will be derived from a single original speech segment and should sound the same; however, one or both of speech samples may be processed with the floating point or the fixed point algorithm. To develop a processed sample, the speech will be encoded and decoded using one of the ATC-10B algorithm implementations. The listener will not know anything about the type of processing, if any, that was performed on the speech. Since the sample pairs will be constructed with the objective of comparing floating and fixed point implementations, a majority of the sample pairs

will consist of a floating-point processed sample and fixed-point processed sample. The speech sample pairs will be derived from Air Traffic speech samples used for the original floating point algorithm evaluation.

Performance

Performance comparisons ensure that the floating and fixed point implementations function equivalently in suboptimal environments. The tests will compare performance in the following three areas: background noise performance, channel error performance, and truncated mode performance. The intent of these performance tests is to compare basic performance, i.e., there are no obvious/ differences in performance between the two implementations. The intent of the tests is to compare performance for reasonable impairments that are likely to be encountered during use. Again, these tests will be conducted using the ATC speech samples described above.

TEST CONDUCT

The voice quality testing will be conducted using FAA NEXCOM Program Office designated listeners. The voice samples and test instructions will be provided on at least two CDs: an audio CD with speech sample pairs and a data CD containing test instructions and test data sheet. The preferred method for completing the test data sheet will be electronic (e.g. Excel spreadsheet) to facilitate the compilation of test results and to minimize clerical data entry. Paper versions of the test instructions and test data sheet will also be available for listeners that are not comfortable with the electronic entry. Statistics will be developed from the raw data.

The remaining tests will be conducted by engineering representatives in a laboratory environment. A report will be generated including raw data, test procedures, and test results.