

ETSI TS 126 243 V14.0.0 (2017-04)



**Digital cellular telecommunication system (Phase 2+) (GSM);
Universal Mobile Telecommunications System (UMTS);
LTE;
ANSI-C code for the fixed-point distributed speech recognition
extended advanced front-end
(3GPP TS 26.243 version 14.0.0 Release 14)**



Reference

RTS/TSGS-0426243ve00

Keywords

GSM,LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
<http://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:
<https://portal.etsi.org/People/CommiteeSupportStaff.aspx>

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.
The copyright and the foregoing restriction extend to reproduction in all media.

© European Telecommunications Standards Institute 2017.
All rights reserved.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are Trade Marks of ETSI registered for the benefit of its Members.
3GPP™ and **LTE™** are Trade Marks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners.
GSM® and the GSM logo are Trade Marks registered and owned by the GSM Association.

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org>).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under <http://webapp.etsi.org/key/queryform.asp>.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights	2
Foreword.....	2
Modal verbs terminology.....	2
Foreword.....	4
1 Scope	5
2 References	5
3 Definitions and abbreviations.....	5
3.1 Definitions	5
3.2 Abbreviations	5
4 C code structure.....	5
4.1 Contents of the C source code	5
4.2 Program execution.....	6
4.3 Code hierarchy	7
4.5 Variables, constants and tables.....	11
4.5.1 Description of constants used in the C-code.....	12
4.5.2 Description of fixed tables used in the C-code.....	15
4.5.3 Static variables used in the C-code	16
5 File formats	19
5.1 Speech file	19
Annex A (informative): Change history.....	20
History	21

Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
 - 1 presented to TSG for information;
 - 2 presented to TSG for approval;
 - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

ITeH STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/7fce-2016-4220-4061-a33e-724d4f1e987d/etsi-ts-126-243-v14.0.0>
2017-04

1 Scope

The present document contains an electronic copy of the ANSI-C code for DSR Extended Advanced Front-end. The ANSI-C code is necessary for a bit exact implementation of DSR Extended Advanced Front-end.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- [1] ETSI ES 202 050 (2007-01) V1.1.5: "Distributed Speech Recognition; Advanced Front-end Feature Extraction Algorithm; Compression Algorithm".
 - [2] ETSI ES 202 212 (2005-11) V1.1.2: "Distributed Speech Recognition; Extended Advanced Front-end Feature Extraction Algorithm; Compression Algorithm, Back-end Speech Reconstruction Algorithm".
 - [3] 3GPP TS 26.177: "Speech Enabled Services (SES); Distributed Speech Recognition (DSR) extended advanced front-end test sequences".
-

3 Definitions and abbreviations

3.1 Definitions

Definition of terms used in the present document, can be found in [1], [2]

3.2 Abbreviations

For the purpose of the present document, the following abbreviations apply:

ANSI	American National Standards Institute
I/O	Input/Output
RAM	Random Access Memory
ROM	Read Only Memory
AFE	Advanced Front-end
X-AFE	eXtended Advanced Front-end
DSR	Distributed Speech Recognition

4 C code structure

This clause gives an overview of the structure of the bit-exact C code and provides an overview of the contents and organization of the C code attached to this document.

The C code has been verified on the following systems:

- Sun Microsystems workstations and GNU gcc compiler
- IBM PC compatible computers with Linux operating system and GNU gcc compiler.

ANSI-C was selected as the programming language because portability was desirable.

4.1 Contents of the C source code

The distributed files with suffix "c" contain the source code and the files with suffix "h" are the header files.

Makefiles are provided for the platforms in which the C code has been verified (listed above).

4.2 Program execution

There are separate executables for the FrontEnd and Vector Quantization, with and without Extensions. The command line options are described below.

⟨⟩ - indicates parameters for the given option for running the executable
 () – indicates default parameter.

FrontEnd w/ Extension:

USAGE: bin/ExtAdvFrontEnd infile HTK_outfile pitch_outfile class_outfile [options]

OPTIONS:

-q Quiet Mode (FALSE)
 -F format Input file format <NIST,HTK,RAW> (NIST)
 -fs freq Sampling frequency in kHz <8,16> (8)
 -swap Change input byte ordering (Native)
 -noh No HTK header to output file (FALSE)
 -noc0 No c0 coefficient to output feature vector (FALSE)
 -nologE No logE component to output feature vector (FALSE)
 -skip_header_bytes n - Skip header, first n bytes (Only for -F RAW)

-noh, -noc0, -nologE and -skip_header_bytes are not used and should not be changed.

FrontEnd w/o Extension:

USAGE: bin/AdvFrontEnd infile HTK_outfile [options]

OPTIONS: - Same as FrontEnd w/ Extension

Vector Quantization w/ Extension:

Usage: extcoder htk_file_in pitch_file_in class_file_in bitstream_file_out pitch_file_out txt_file_out -freq x -

VAD/No_VAD
 htk_file_in Input mel-frequency cepstral coefficient file in HTK MFCC format.
 pitch_file_in Input pitch period file.
 class_file_in Input classification file.
 bit_file_out Output binary bitstream.
 pitch_file_out Output quantised pitch period file.
 txt_file_out Vector quantiser output in text format.
 -freq x Sampling frequency in kHz (8 or 16).
 -VAD Use voice activity detector data. Voice activity input file must have same name as htk_file, but extension .vad
 -No_VAD Do not incorporate voice activity detector information in output bitstream.

Vector Quantization w/o Extension:

Usage: coder htk_file_in bitstream_file_out txt_file_out -freq x -VAD/No_VAD

htk_file_in Input mel-frequency cepstral coefficient file in HTK MFCC format.
 bit_file_out Binary output bitstream.
 txt_file_out Vector quantiser output in text format.
 -freq x Sampling frequency in kHz (8 or 16).
 -VAD Use voice activity detector data. Voice activity input file must have same name as htk_file, but extension .vad
 -No_VAD Do not incorporate voice activity detector information in output bitstream.

File extension descriptions as generated by the sample script:

.cep – Binary file containing cepstral features in HTK format. Output from the FrontEnd, input to the vector quantizer.
 .pitch – Binary file containing pitch information. Output from the FrontEnd, input to the vector quantizer. Only used for Extension.
 .class – Ascii file containing class information. Output from the FrontEnd, input to the vector quantizer. Only used for Extension.
 .bs – Binary file containing the bitstream. Output from the vector quantizer.
 .log – Log files from the different executables.

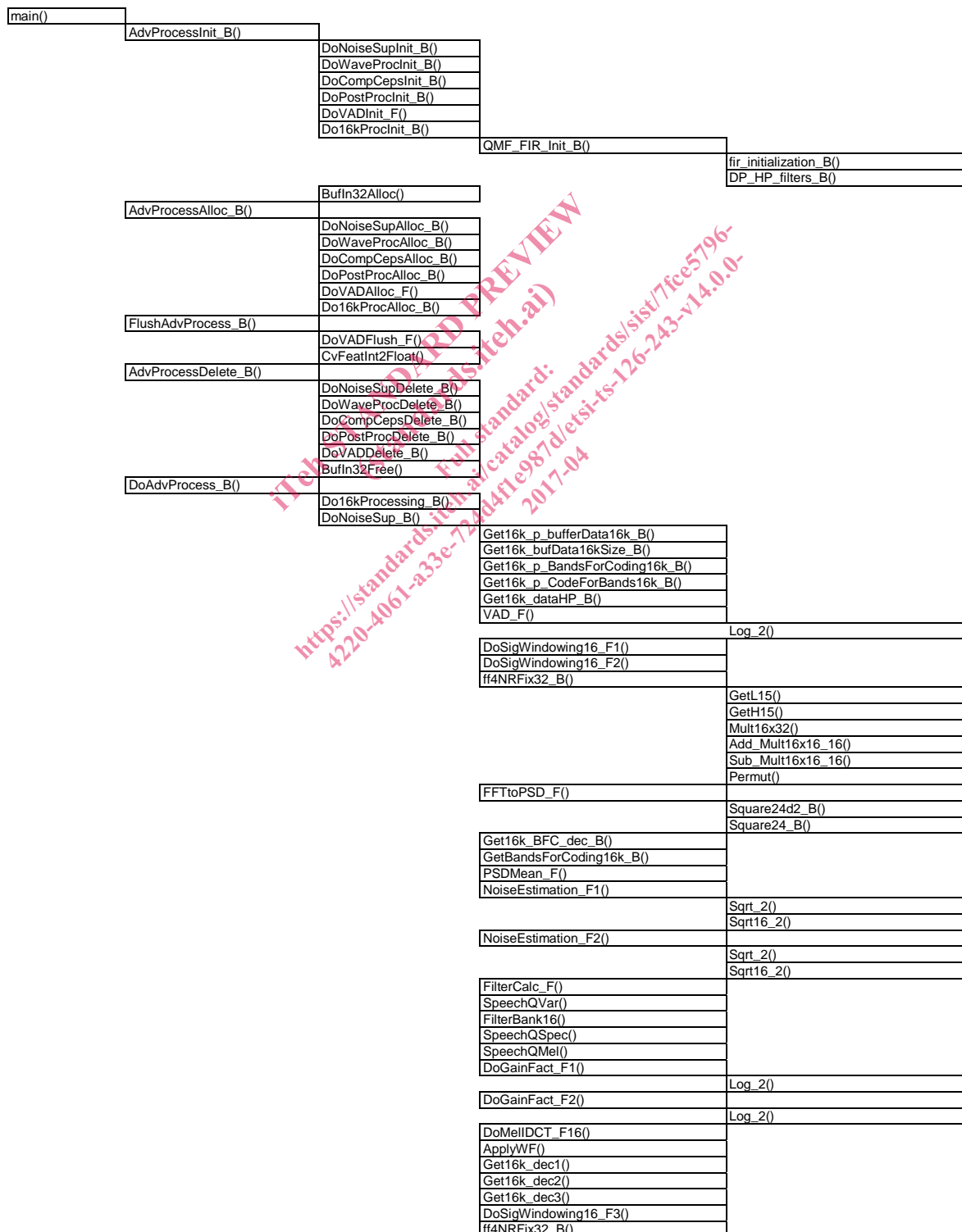
4.3 Code hierarchy

Tables 1 to 3 are call graphs that show the functions used for AFE (table 1), VQ (table 2), and Extension (table 3).

Each column represents a call level and each cell a function. The functions contain calls to the functions in rightwards neighboring cells. The time order in the call graphs is from the top downwards as the processing of a frame advances. All standard C functions: printf(), fwrite(), etc. have been omitted. Also, no basic operations (add(), L_add(), mac(), etc.) or double precision extended operations (e.g. L_Extract()) appear in the graphs.

The basic operations are not counted as extending the depth, therefore the deepest level in this software is level 7.

Table 1: AFE call structure



	GetL15()
	GetH15()
	Mult16x32()
	Add_Mult16x16_16()
	Sub_Mult16x16_16()
	Permut()
FFTtoPSD_F()	
	Square24d2_B()
	Square24_B()
DoMelFB_B()	
CodeBands16k_B()	
DoSpecSub16k_B()	
	Log_2()
UpDateDecal()	
ApplyDecal()	
DCOffsetFil_F()	
Get16k_hpBandsSize_B()	
Get16k_p_hpBands_B()	
Get16k_p_bufferCodeForBands16k_B()	
Get16k_p_CodeForBands16k_B()	
Get16k_p_bufferCodeWeights_B()	
Get16k_p_codeWeights_B()	
Set16k_hpBands_dec_B()	
DoWaveProc_B()	
	TeagerEng()
	GetTeagerFilter()
DoCompCeps_B()	GetMaximaPositions()
	CepsCompute()
	Get16k_p_bufferCodeWeights_B()
	Get16k_p_bufferCodeForBands16k_B()
	PreEmphHamm()
	ff4NB16_B()
	GetBandsForDecoding16k_B()
	DecodeBands16k_B()
	FilterBank()
	Get16k_hpBands_dec_B()
	Get16k_p_hpBands_B()
	MergeSSandCoded_B()
	CorrectEnergy_B()
	CosInv16Khz()
	cosInv() (only for 8kHz)
DoPostProc_B()	
DoVADProc_F()	
	focalpoint()

Table 2: VQ call structure

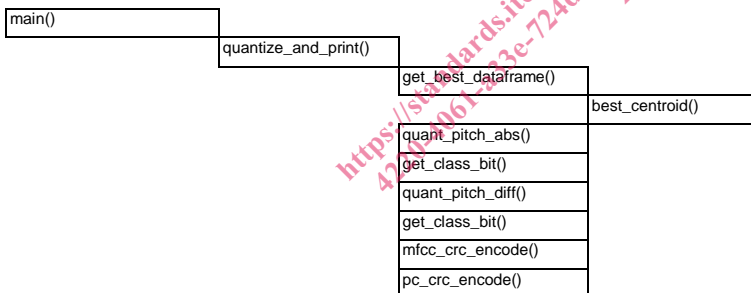
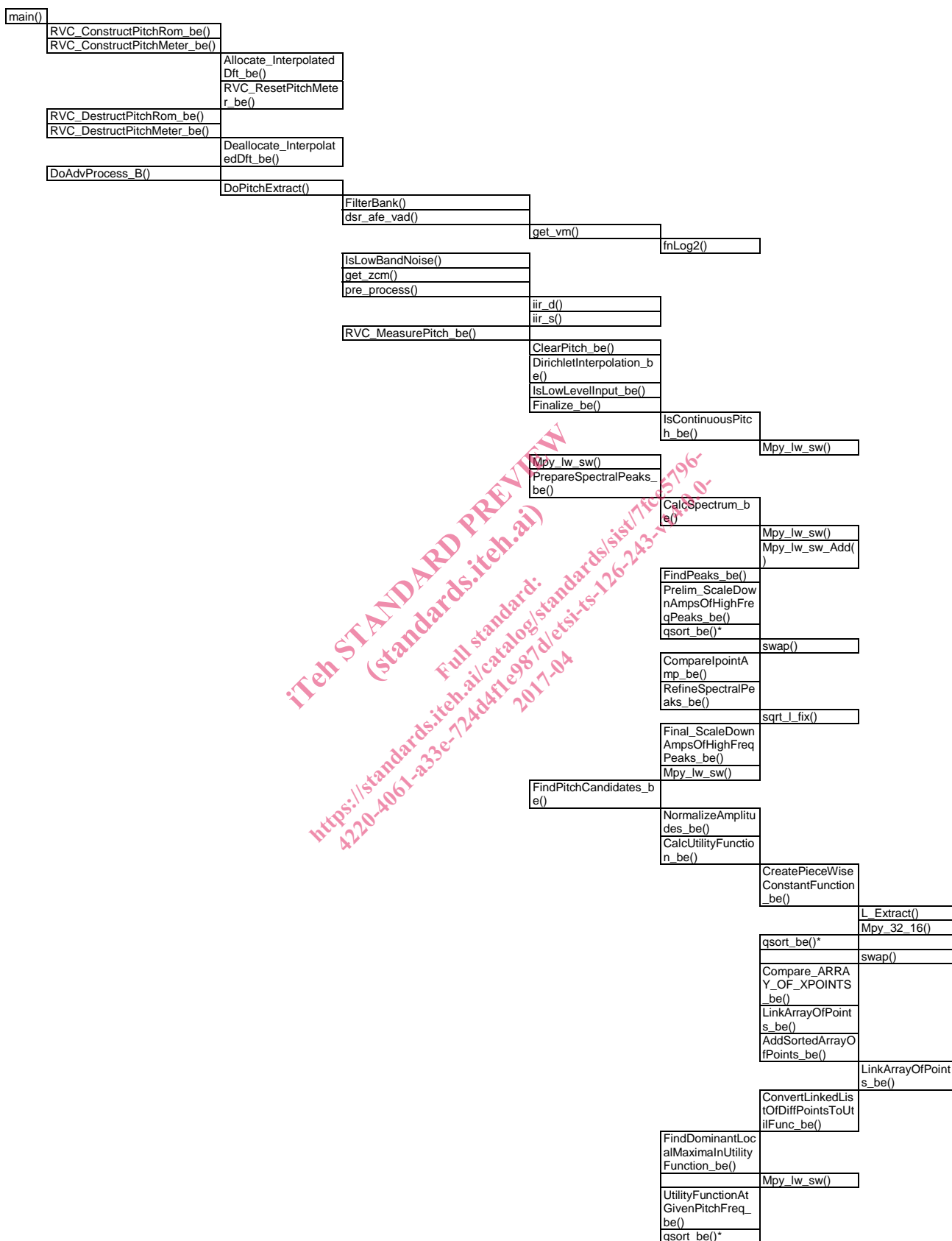


Table 3: Extension call structure



iTeh STANDARD PREVIEW
 (standards.iteh.ai)
 Full standard:
<https://standards.iteh.ai/catalog/standards/sist/71655796-4220-4061-a33e-724d4f1e987d/etsi-ts-126-243-v14-0-2017-04>