DSP Innovations

TWELP MR 300 bps ... 3600 bps Vocoder

(version 5.0) is multi-rate and operates in any of the 15 switchable modes:

- 3600 bps
- 2400 bps
- 1600 bps
- 1200 bps
- 700 bps
- 600 bps
- 480 bps
- 300 bps
- 1200/2400 bps (scalable)
- 700/1600 bps (scalable)
- 600/900 bps (scalable)
- 480/800 bps (scalable)
- 3600 NRT (variable bit rate, non-real-time)
- 1600 NRT (variable bit rate, non-real-time)
- 700 NRT (variable bit rate, non-real-time)

At each bit rate, the vocoder offers speech quality and intelligibility that were previously achievable only at nearly double, double, or higher bit rates.
Below are the results of speech quality and intelligibility testing, comparing the 2400 bps mode with the MELPe 2400 bps vocoder and the 3GPP AMR-NB 4750 bps vocoder, which operates at twice the bitrate.
For all other bit rates, you can find the necessary information on the corresponding web pages.

For Digital HF Radio and other markets.

TWELP Technology Features. The vocoder is based on newest technology of speech coding called "Tri-Wave Excited Linear Prediction" (TWELP) that was developed by experts of DSPINI.

TWELP technology is a new class of vocoders that differs from any other LPC-based vocoders by:

- advance reliable method of pitch estimation
- pitch-synchronous analysis
- advance tri-wave model of excitation
- newest quantization schemes
- pitch-synchronous synthesis

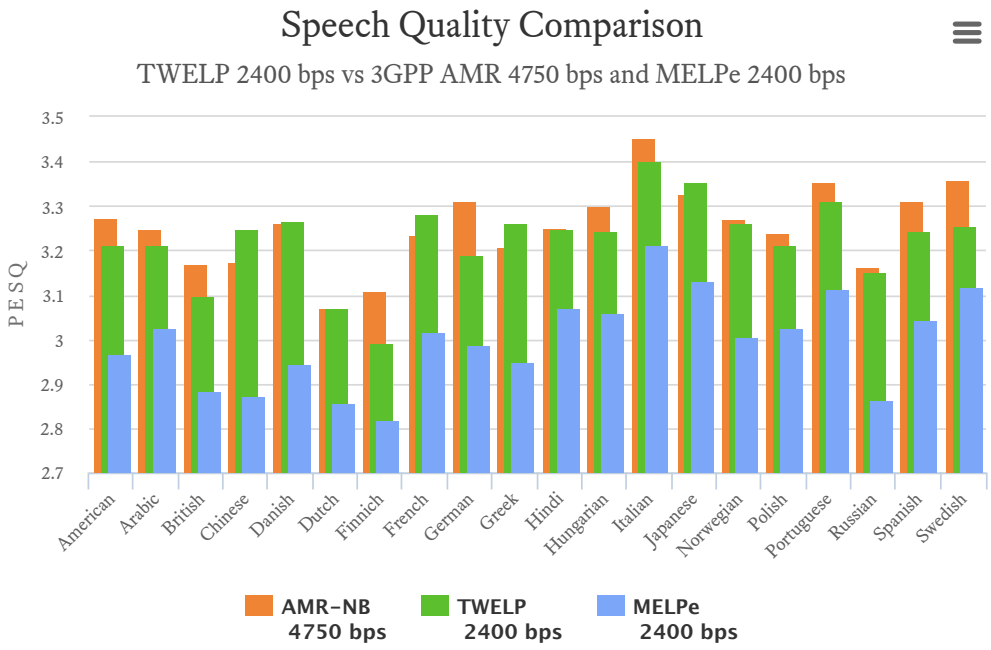
Thanks to these unique features, TWELP technology provides significantly better speech quality than other well-known technologies—including AMBE+2, MELPe, ACELP, and others—at equivalent bit rates ranging from 300 bps to 4800 bps and beyond.

Additionally, unlike other low-bitrate vocoders (such as MELPe, for example), TWELP delivers much higher quality for non-speech signals, including sirens, background music, and similar audio.

Speech Quality. This is a comparison with the 3GPP AMR-NB 4750 bps vocoder and MELPe 2400 bps vocoder.
The vocoders were tested using the ITU-T P.50 speech database in 20 different languages.

Note:
We have updated the speech database by minimizing inter-speech pauses to eliminate their impact on the evaluation results. Therefore, the numbers obtained from the quality measurements using this updated speech database differ from those previously obtained with the original speech database, where speech pauses were not removed.

The ITU-T P.862 tool was used to evaluate speech quality in terms of PESQ scores:



The diagram demonstrates that the speech quality of the TWELP 2400 bps vocoder is closer to that of the AMR-NB 4750 bps vocoder and significantly better than that of the MELPe 2400 bps vocoder. Exact numbers are shown in the table below.

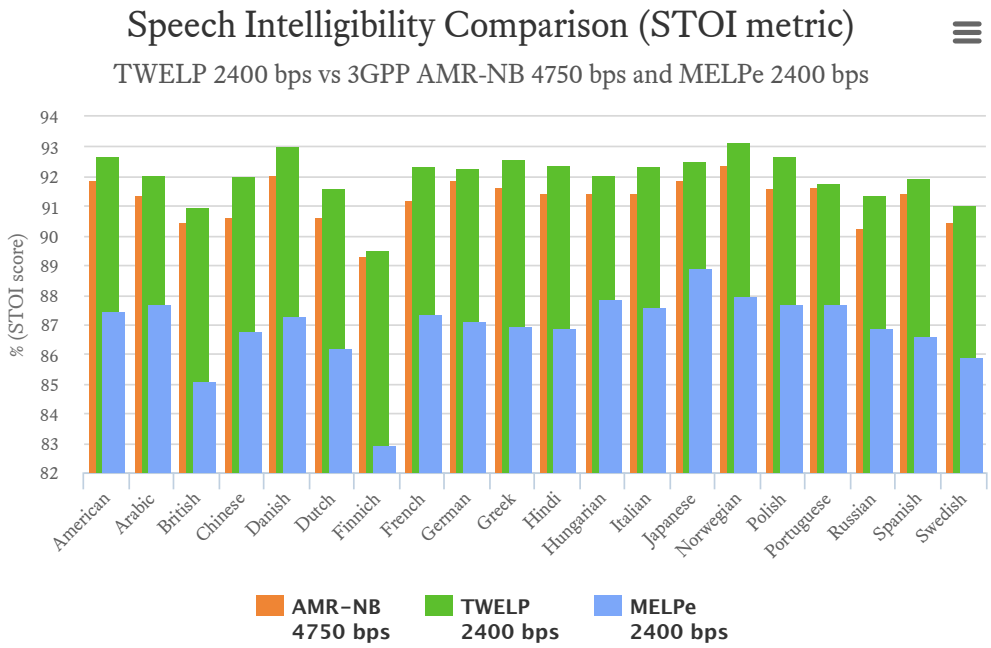
Language	AMR-NB 4750 bps	TWELP 2400 bps	MELPe 2400 bps
American	3.274	3.213	2.967
Arabic	3.246	3.214	3.024
British	3.170	3.098	2.885
Chinese	3.176	3.249	2.871
Danish	3.264	3.268	2.946
Dutch	3.072	3.070	2.856
Finnich	3.108	2.989	2.820
French	3.236	3.281	3.018
German	3.314	3.189	2.986
Greek	3.209	3.262	2.950
Hindi	3.250	3.247	3.070
Hungarian	3.300	3.245	3.060
Italian	3.454	3.399	3.213
Japanese	3.327	3.353	3.133

Norwegian	3.270	3.264	3.007
Polish	3.241	3.213	3.027
Portuguese	3.355	3.314	3.113
Russian	3.163	3.153	2.863
Spanish	3.312	3.244	3.046
Swedish	3.360	3.256	3.119
Average	**3.255**	**3.226**	**2.999**

The average difference between AMR-NB 4750 bps and TWELP 2400 bps is 0.029 PESQ, while the average difference between TWELP 2400 bps and MELPe 2400 bps is 0.227 PESQ.

Please visit the web pages with the corresponding bit rates to see speech quality estimations for other bit rates.

Speech Intelligibility. Here is a comparison with the 3GPP AMR-NB vocoder, which operates at twice the bitrate (4750 bps) and with the MELPe at the same bitrate (2400 bps) . The vocoders were tested using the ITU-T P.50 speech database, covering 20 different languages.
STOI (Short-Time Objective Intelligibility) and ESTOI (Extended Short-Time Objective Intelligibility) metrics were used to assess speech intelligibility:



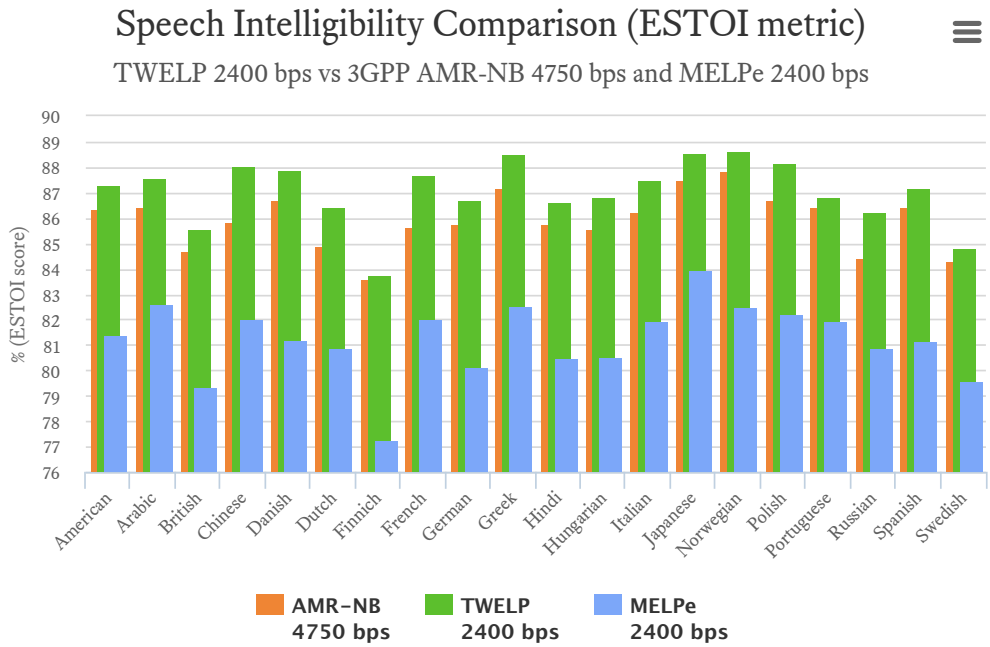
The diagram demonstrates the advantage of the TWELP 2400 bps vocoder over both the MELPe 2400 bps and AMR-NB 4750 bps vocoders, despite the latter operating at twice the bitrate. Exact values are provided in the table below:

Language	**AMR-NB 4750 bps**	**TWELP 2400 bps**	**MELPe 2400 bps**
American	91.89	92.69	87.43
Arabic	91.36	92.03	87.70
British	90.43	90.98	85.10

Chinese	90.63	92.00	86.78
Danish	92.06	93.00	87.30
Dutch	90.63	91.58	86.19
Finnich	89.29	89.54	82.94
French	91.16	92.33	87.35
German	91.86	92.28	87.10
Greek	91.64	92.54	86.96
Hindi	91.43	92.37	86.89
Hungarian	91.40	92.05	87.85
Italian	91.44	92.36	87.57
Japanese	91.87	92.53	88.88
Norwegian	92.39	93.13	87.97
Polish	91.56	92.68	87.66
Portuguese	91.64	91.77	87.68
Russian	90.28	91.38	86.89
Spanish	91.43	91.96	86.57
Swedish	90.46	91.02	85.92
Average	**91.24**	**92.01**	**86.94**

The average difference between the TWELP 2400 bps and the AMR-NB 4750 bps vocoders is 0.77%, while the average difference between the TWELP 2400 bps and the MELPe 2400 bps vocoders is 5.07%.

Considering that a low-bitrate vocoder is a nonlinear device that significantly distorts the spectrum of the original speech signal, the ESTOI metric provides more accurate assessments of speech intelligibility after vocoding:



The diagram demonstrates the advantage of the TWELP 2400 bps vocoder over both the MELPe 2400

bps and AMR-NB 4750 bps vocoders, despite the latter operating at twice the bitrate. Exact numbers are shown in the table below.

Language	AMR-NB 4750 bps	TWELP 2400 bps	MELPe 2400 bps
American	86.39	87.35	81.43
Arabic	86.48	87.59	82.61
British	84.71	85.55	79.35
Chinese	85.83	88.03	82.03
Danish	86.75	87.91	81.22
Dutch	84.90	86.47	80.88
Finnich	83.66	83.80	77.26
French	85.63	87.69	82.03
German	85.75	86.69	80.14
Greek	87.19	88.52	82.55
Hindi	85.81	86.62	80.48
Hungarian	85.55	86.83	80.57
Italian	86.22	87.52	81.95
Japanese	87.49	88.62	83.96
Norwegian	87.86	88.65	82.53
Polish	86.71	88.20	82.26
Portuguese	86.46	86.87	81.94
Russian	84.47	86.25	80.86
Spanish	86.47	87.22	81.18
Swedish	84.33	84.82	79.60
Average	**85.93**	**87.06**	**81.24**

The average difference between the TWELP 2400 bps and the AMR-NB 4750 bps vocoders is 1.13%, while the average difference between the TWELP 2400 bps and the MELPe 2400 bps vocoders is 5.82%.

You can download the P.862 and STOI/ESTOI utilities, along with all speech samples, by using the links in the 'Downloads' section at the bottom of the page, and then check all the numbers presented above.

Please visit the web pages with the corresponding bit rates to see speech intelligibility estimations for other bit rates.

Speech Samples (WAV-files).

A few independent experts compared the TWELP 2400 bps vocoder with the AMR-NB 4750 bps and MELPe 2400 bps vocoders using the preference method.

Although opinions were almost evenly split, the majority of listeners preferred the TWELP vocoder over the AMR-NB vocoder, noting that the speech sounded clearer and more natural. All listeners preferred the TWELP vocoder over the MELPe vocoder, noting that the speech sounded more natural and less synthetic.

You can play and listen to short samples of the source speech, as well as the speech processed by the AMR-NB 4750 bps vocoder, MELPe 2400 bps vocoder and the TWELP 2400 bps vocoder, using the links in the table below.

For the best listening experience, we recommend using high-quality headphones or premium audio equipment to hear the nuances and differences in the vocoder sound more clearly.

You can also download the complete set of P.50 samples as zip files for all languages simultaneously by using the links in the 'Downloads' section at the bottom of the page.

Language	Source speech	AMR-NB 4750 bps	MELPe 2400 bps	TWELP 2400 bps
American	▶	▶	▶	▶
Arabic	▶	▶	▶	▶
British	▶	▶	▶	▶
Chinese	▶	▶	▶	▶
Danish	▶	▶	▶	▶
Dutch	▶	▶	▶	▶
Finnich	▶	▶	▶	▶
French	▶	▶	▶	▶
German	▶	▶	▶	▶
Greek	▶	▶	▶	▶
Hindi	▶	▶	▶	▶
Hungarian	▶	▶	▶	▶
Italian	▶	▶	▶	▶
Japanese	▶	▶	▶	▶
Norwegian	▶	▶	▶	▶
Polish	▶	▶	▶	▶
Portuguese	▶	▶	▶	▶
Russian	▶	▶	▶	▶
Spanish	▶	▶	▶	▶
Swedish	▶	▶	▶	▶

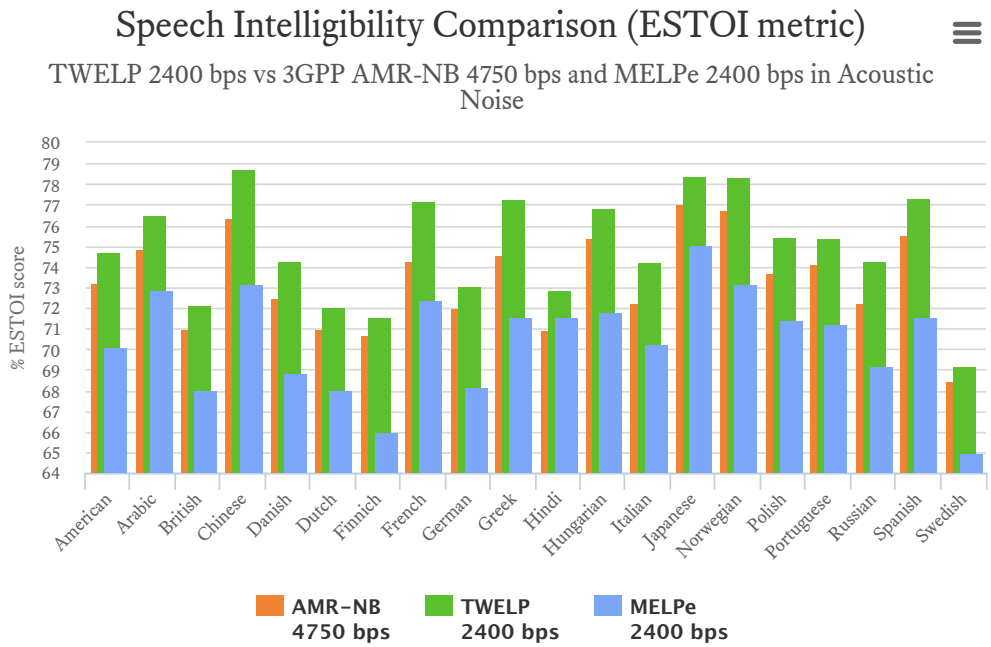
Please visit the web pages with the corresponding bit rates to hear speech samples for other bit rates.

Superiority In Quality Of The Non-speech Signals. In contrast to other LBR vocoders (MELPe, AMBE+2, etc.), TWELP vocoders provide high quality of non-speech signals, including police, ambulance, fire sirens, etc. This feature in conjunction with high quality natural human-sounding of voice makes TWELP vocoders well suitable for replacement of analog radio by digital radio and also for other applications where high quality transmitting of non-speech signals is relevant along with high quality transmitting of speech signals.

Source signal	AMR-NB 4750 bps	MELPe 2400 bps	TWELP 2400 bps
▶	▶	▶	▶

You can hear that the TWELP vocoder processes non-speech signals just as well, if not more accurately, than the waveform AMR-NB codec.

High Robustness To Acoustic Noise. In contrast to other LBR vocoders, TWELP vocoders are highly robust to acoustic noise due to a reliable pitch estimation method and other features of TWELP technology.



The diagram demonstrates the advantage of the TWELP 2400 bps vocoder over both the MELPe 2400 bps and AMR-NB 4750 bps vocoders, despite the latter operating at twice the bitrate. Exact numbers are shown in the table below.

Language	AMR-NB 4750 bps	TWELP 2400 bps	MELPe 2400 bps
American	73.26	74.70	70.09
Arabic	74.92	76.56	72.88
British	71.00	72.18	68.02
Chinese	76.39	78.73	73.16
Danish	72.52	74.29	68.87
Dutch	70.99	72.06	68.01
Finnich	70.64	71.56	65.96
French	74.32	77.23	72.39
German	71.96	73.04	68.23
Greek	74.55	77.26	71.61
Hindi	70.92	72.90	71.61
Hungarian	75.41	76.83	71.84
Italian	72.22	74.19	70.28
Japanese	76.99	78.40	75.04
Norwegian	76.76	78.31	73.14
Polish	73.70	75.47	71.40
Portuguese	74.12	75.40	71.25
Russian	72.22	74.27	69.19

Spanish	75.51	77.32	71.55
Swedish	68.48	69.17	65.02
Average	**73.34**	**74.95**	**70.33**

The average difference between the TWELP 2400 bps and the AMR-NB 4750 bps vocoders is 1.61%, while the average difference between the TWELP 2400 bps and the MELPe 2400 bps vocoders is 4.62%.

Please visit the web pages with the corresponding bit rates to see the numbers for other bit rates.

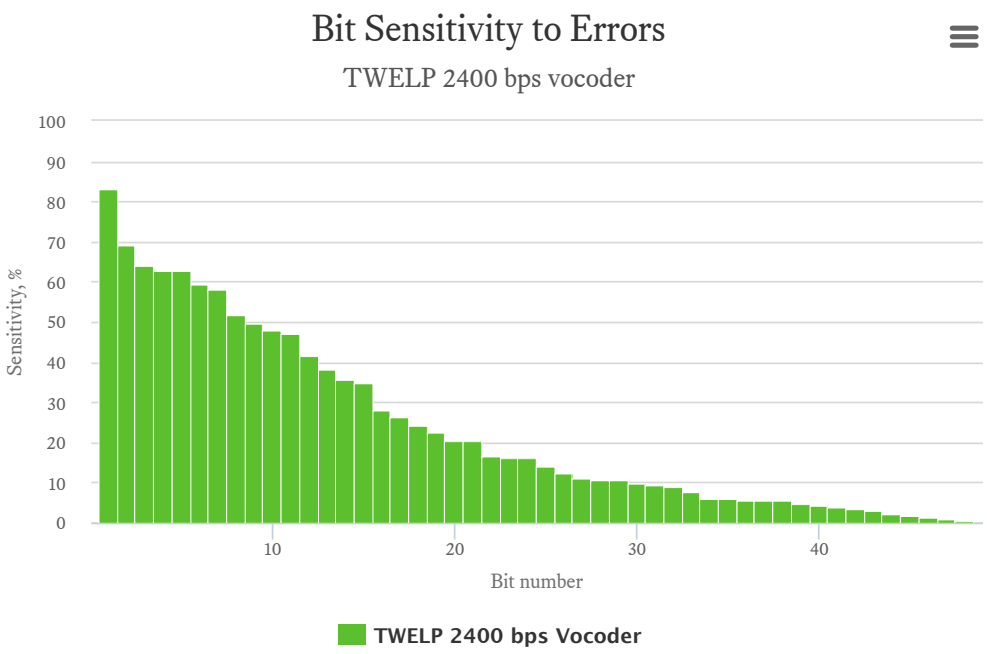
Additionally, the TWELP vocoder features an NCSE (Noise Cancellation Speech Enhancement) preprocessor, which cleans the input speech signal from noise and enhances speech quality.
Below, you can listen to a short fragment of heavily noisy English speech after passing through MELPe and TWELP vocoders, with NPP (Noise Pre-Processor) and NCSE disabled and enabled, respectively.

NPP NCSE	**Input speech (SNR=10dB)**	**AMR-NB 4750 bps**	**MELPe 2400 bps**	**TWELP 2400 bps**
Disabled	▶	▶	▶	▶
Enabled	▶		▶	▶

The NCSE integrated into the TWELP vocoder is described in more detail on the webpage for our standalone product, 'NCSE-AGC Preprocessor'.

High Robustness To The Channel Errors.
The TWELP technology offers highly efficient speech compression by eliminating redundancy while preserving excellent quality and intelligibility. To enhance robustness against transmission errors, we provide specialized versions called **TWELP Robust**.
These vocoders are based on an effective Joint Source-Channel Coding approach. Each vocoder is equipped with a custom-designed FEC, tailored to its specific characteristics and operational conditions.
TWELP Robust vocoders provide high speech quality simultaneously in noisy channel as well as in noiseless channel. FEC can operate with "soft decisions" as well as with "hard decisions" from a modem. "Soft decisions" mode provides much better robustness in comparison with the "hard decisions" mode.
For all users of our non-robust vocoder versions, we offer the following recommendations.

The diagram below illustrates the sensitivity of bits at the output of the vocoder to communication channel errors.
Essentially, the diagram shows by what percentage speech quality is reduced when a specific bit is distorted. The first bits in order cause catastrophic distortions, while the latter bits have significantly less impact on quality.



We strongly recommend using FEC (Forward Error Correction) with unequal protection of the bits in strong accordance with their sensitivity to errors and utilizing 'Soft Decisions' decoding. This will provide the highest robustness of the vocoder against errors in the channel.

Please visit the web pages with the corresponding bit rates to see the diagram for other bit rates.

Additional Functionalities. The following additional functionalities are developed by DSPINI and integrated into TWELP vocoders:

- Noise Cancellation Speech Enhancement (NCSE)
- Automatic Gain Control (AGC),
- Voice Activity Detector (VAD),
- Discontinuous Transmission (DTX),
- Tone Detection/Generation (Single tones and Dual tones). The tones are transmitted by the vocoder facilities.

Note:
The Tone Detector/Generator functionalities are not integrated into the code by default but can be added free of charge upon request.

Each functionality has unique features, performance and characteristics, providing significant superiority over any well-known implementations on the market.

Technical Characteristics And Resource Requirements:

Technical characteristics

Bit Rate (bps)	Algorithm	Frame size (ms)	Algorithmic delay (including frame size) (ms)	Sampling rate (kHz)	Signal format	Bit stream format (bits)
3600	TWELP™	20	40	8	Linear 16-bit PCM	72
2400		20	40			48
1600		40	60			64
1200		40	60			48
700		80	100			56
600		80	100			48
480		100	120			48
300		100	120			30
1200/2400 (Scalable)		40	60			48+48=96 (DTX is disabled) 11...97 (DTX is enabled)
700/1600 (Scalable)		80	100			56+72=128 (DTX is disabled) 11...129 (DTX is enabled)
600/900 (Scalable)		80	100			48+24=72 (DTX is disabled) 11...73 (DTX is enabled)
480/800 (Scalable)		100	120			48+32=80 (DTX is disabled) 11...81 (DTX is enabled)
3600 NRT (VBR)		20xN	20xN+20			11xN...73xN
1600 NRT (VBR)		40xN	40xN+20			11xN...65xN
700 NRT (VBR)		80xN	80xN+20			11xN...57xN

Additional functionalities

Name	Functionality	Technical characteristics	
		Name	Value
AGC	Automatic Gain Control	Control range:	0 ... +42 dB
NCSE	Noise Canceller - Speech Enhancer	SNR increasing	> 6 dB
		Speech quality improvement	> 0.1 PESQ
Tone Detector	Single/Dual tones detection	In accordance with international standards	
Tone Generator	Single/Dual tones generation	Special generator, kept continuity of signal (phase and amplitude of signal of previous frame)	
VAD	Voice Activity Detection	Reliable detection speech in background noise	
CNG	Comfort Noise Generation	Type of noise	"white"
		Level	- 60 dB

Resources for TI's DSP C64+ platform

Bitrate	Module	MIPS* peak	**Memory** (KBytes)				
			Program	Data			
				Constants	Channel	Heap	Stack
3600	Voice Encoder	20.0					
	NCSE	3.0					
	AGC	0.3					
	Voice Decoder	5.6					
	Voice Encoder + Voice Decoder	25.6					
	Total	28.9					
	Voice Encoder	18.7					
	NCSE	2.9					
	AGC	0.3					

2400	Voice Decoder	5.4
	Voice Encoder + Voice Decoder	24.1
	Total	27.3
1600	Voice Encoder	38.9
	NCSE	2.8
	AGC	0.3
	Voice Decoder	5.4
	Voice Encoder + Voice Decoder	44.3
	Total	47.4
1200	Voice Encoder	35.0
	NCSE	3.1
	AGC	0.3
	Voice Decoder	5.4
	Voice Encoder + Voice Decoder	40.4
	Total	43.8
700	Voice Encoder	30.0
	NCSE	2.6
	AGC	0.3
	Voice Decoder	5.3
	Voice Encoder + Voice Decoder	35.3
	Total	38.2
600	Voice Encoder	19.0
	NCSE	2.6
	AGC	0.3
	Voice Decoder	5.3
	Voice Encoder + Voice Decoder	24.3
	Total	27.2

480	Voice Encoder	17.6	295	2850	5.0	6.7	2.1
	NCSE	2.6					
	AGC	0.3					
	Voice Decoder	5.3					
	Voice Encoder + Voice Decoder	22.9					
	Total	25.8					
300	Voice Encoder	25.9					
	NCSE	2.8					
	AGC	0.1					
	Voice Decoder	3.8					
	Voice Encoder + Voice Decoder	29.7					
	Total	32.6					
1200/2400 scalable	Voice Encoder	51.9					
	NCSE	2.7					
	AGC	0.3					
	Voice Decoder	5.5					
	Voice Encoder + Voice Decoder	57.4					
	Total	60.4					
700/1600 scalable	Voice Encoder	30.3					
	NCSE	2.4					
	AGC	0.3					
	Voice Decoder	5.4					
	Voice Encoder + Voice Decoder	35.7					
	Total	38.4					
	Voice Encoder	23.2					
	NCSE	3.0					
	AGC	0.3					

600/900 scalable	Voice Decoder	5.3
	Voice Encoder + Voice Decoder	28.5
	Total	31.8
480/800 scalable	Voice Encoder	17.8
	NCSE	3.0
	AGC	0.3
	Voice Decoder	5.3
	Voice Encoder + Voice Decoder	23.1
	Total	26.4
3600 NRT (VBR)	Voice Encoder	20.0
	NCSE	3.0
	AGC	0.3
	Voice Decoder	5.6
	Voice Encoder + Voice Decoder	25.6
	Total	28.9
1600 NRT (VBR)	Voice Encoder	38.9
	NCSE	2.8
	AGC	0.3
	Voice Decoder	5.4
	Voice Encoder + Voice Decoder	44.3
	Total	47.4
700 NRT (VBR)	Voice Encoder	29.2
	NCSE	2.4
	AGC	0.3
	Voice Decoder	5.3
	Voice Encoder + Voice Decoder	34.5

Total	37.2					

Resources for ARM Cortex M4 platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder						
NCSE						
AGC						
Voice Decoder						
Voice Encoder + Voice Decoder						
Total						

Resources for ARM Cortex M4 platform are available on request.
Approximately, you can estimate the resources based on the figures provided on the corresponding webpages for different bit rates.

Resources for TI′s C55 DSP platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder						
NCSE						
AGC						
Voice Decoder						
Voice Encoder + Voice Decoder						
Total						

Resources for TI's DSP C55 platform are available on request.
Approximately, you can estimate the resources based on the figures provided on the corresponding webpages for different bit rates.

* DSPINI continues optimization of the TWELP algorithm and code in order to minimize computational complexity of the vocoder.

Software Integrity and Security. DSPINI guarantees the ABSOLUTE integrity of its software, free from any undocumented features, undeclared capabilities, or hidden functions. Our customers can be assured that none of our software/code contains any secret features or functionalities concealed from the user. If necessary, we are ready to provide the source code of our software products for appropriate certification. Moreover, our software is available in source code form—you simply need to purchase the appropriate license to use it.

Guarantee And Support. DSPINI guarantees a quality and accordance of all technical characteristics of the product to requirement of current specifications. Testing and other method of quality control are used for guarantee support.

Any Platforms. DSPINI can port this vocoder software into any other DSP, RISC or general- purposes platform inshort time: 1-2 months.

Licensing Terms. To use the vocoder, customer should obtain a license from DSPINI only.

Customization. The vocoder can be customized under any specific requirements- other bit rate, frame size, any other robustness to channel errors, etc. Please contact with us for details.

Prospects. DSPINI is impoving and developing continuously a set of new vocoders with range from 300 bps up to 9600 bps, based on TWELP technology.

Related Software. This vocoder may be effectively used in a bundle with other DSPINI's products:

- Linear and acoustic echo cancellers,
- Multichannel noise cancellers (including two-microphone adaptive array),
- Wired or radiomodems for any types of channels and bitrates,
- Other products.

Downloads:

- Datasheet (pdf)
- ITU-T P.50 source speech samples (zip)
- AMR-NB 4750 bps speech samples (zip)
- MELPe 2400 bps speech samples (zip)
- TWELP 2400 bps speech samples (zip)
- P.862 and STOI/ESTOI utilities
- PC-evaluation package (zip) — on request
- User's Guide document (pdf) — on request

Please visit the web pages with the corresponding bit rates to download speech samples for other bit rates.

Send us an e-mail
request@dspini.com
or call
+33 9 70 40 33 99