DSP Innovations

SPR 600 bps Vocoder

Provides high speech quality and robustness to the channel errors - better than MELPe 600 bps vocoder, at very low consumption of the computing resources and memory (~2.5 times less than MELPe).

For low cost Digital HF Radio, Digital Mobile Radio (DMR), Voice Storage devices and other markets.

SPR Technology Features. The SPR 600 bps Vocoder is based on Sinusoidal-Pulsed Representation (SPR) model developed by DSP Innovations.

SPR™ model is based simultaneously on two well-known models: Sinusoidal (Harmonic) Coding (SHC) and Linear Predictive Coding (LPC), where system function is presented by LPC-filter and excitation function is formed by Sinusoidal Harmonic model.

However, SPR model uses more complex excitation of the synthesize LPC-filter that contains three components:

- voiced (sinusoidal harmonics of the "fundamental frequency"),
- unvoiced (noise),
- pulsed (aperiodic pulses)

As rule, the first two components are used in Sinusoidal Coding, and third component is used in Multi-Pulse Excitation coding on high bit rates.

We combined advantages of these models to achieve high speech quality at low bit rate and low resources costs.

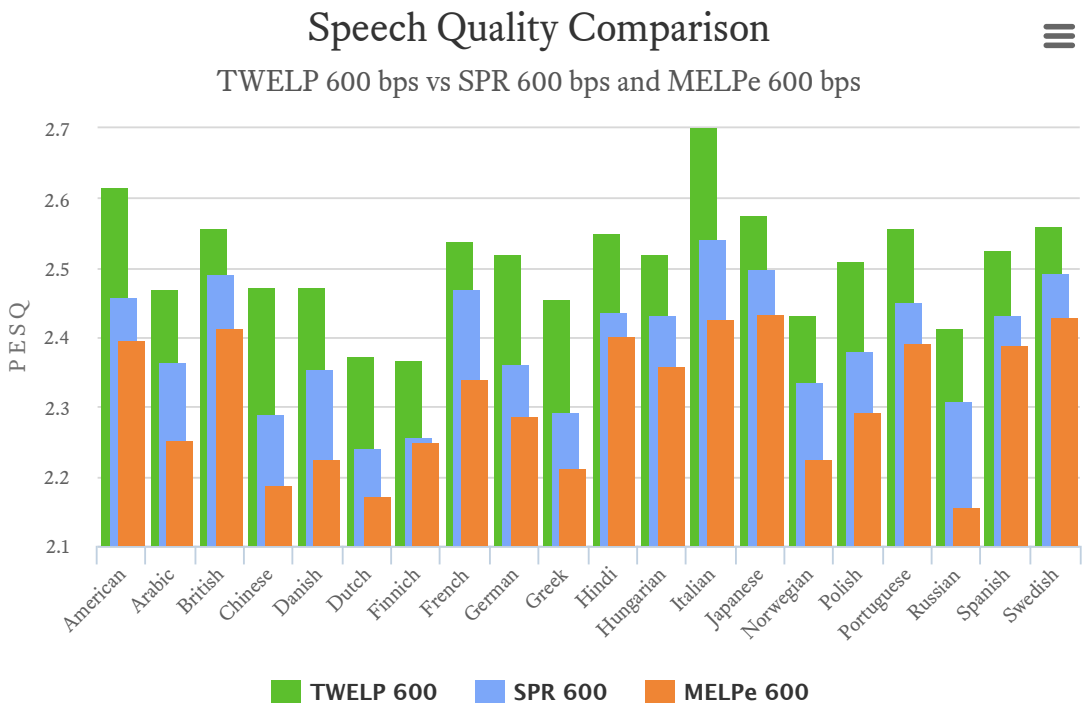
So, for example, SHC model can't represent complex speech intervals with aperiodic fluctuations of vocal chords. However, such intervals can be represented by pulse excitation very well. And vise versa, MPE-model can't provide high speech quality if number of pulses is not enough, therefore this model is used on relatively high bit rates (more 4500 bps).

The combining of these models to form an excitation signal for LPC- model allows achieving unique characteristics of vocoders based on the SPR model.

Therefore, bit stream on output of the Voice Encoder contains the following quantized parameters:

- LPC-parameters (as LSF),
- "Fundamental frequency" (Pitch),
- Frequency-depended Voiced/Unvoiced decisions,
- Pulse parameters,
- Gains

Superiority In Speech Quality. Here is the comparison with MELPe 600 bps vocoder in noiseless channel. SPR 600 bps vocoder, MELPe 600 bps vocoder and TWELP 600 bps vocoder were tested, using ITU-T P.50 speech base for 20 different languages. ITU-T P.862 utility was used for estimation of the speech quality in PESQ terms:



A diagram demonstrates superiority of the SPR 600 bps over MELPe 600 bps in speech quality in clear channel. TWELP 600 bps vocoder provides the best speech quality, but consumes much more (~3 times more) computing resources and memory in comparison with SPR 600 bps vocoder. Exact numbers of the speech quality are presented in the table below.

Language	TWELP 600	SPR 600	MELPe 600
American	2.617	2.460	2.397
Arabic	2.470	2.366	2.254
British	2.559	2.491	2.414
Chinese	2.471	2.289	2.189
Danish	2.471	2.354	2.227
Dutch	2.374	2.241	2.173
Finnich	2.367	2.257	2.249
French	2.539	2.470	2.34
German	2.521	2.362	2.287
Greek	2.457	2.294	2.212
Hindi	2.549	2.437	2.403
Hungarian	2.520	2.431	2.359
Italian	2.705	2.543	2.428
Japanese	2.576	2.499	2.436
Norwegian	2.433	2.335	2.225

Polish	2.511	2.382	2.294
Portuguese	2.558	2.451	2.393
Russian	2.414	2.310	2.156
Spanish	2.525	2.431	2.389
Swedish	2.562	2.495	2.43
Average	**2.50995**	**2.3949**	**2.31275**

Superiority of the SPR 600 bps vocoder over MELPe 600 bps vocoder is on average 0.082 PESQ

Speech Samples (WAV-files). A few independent experts listened SPR 600 bps vocoder in comparison with MELPe 600 bps vocoder and TWELP 600 bps vocoder, using method of preferences. Majority of experts preferred SPR to MELPe and TWELP to SPR, having noted much better intellegibility of speech in SPR in comparison with MELPe and much more natural human-sounding of voice in the TWELP vocoder. You can play and listen short samples of the source speech as well as the speech processed by these vocoders for any of 20 languages, using links in the table below.
Also, you can download full set of the P.50 samples as zip-files for all languages simultaneously, using the links in the "Downloads" para in a bottom of the page.

Language	Source speech	MELPe 600 bps	SPR 600 bps	TWELP 600 bps
American	▶	▶	▶	▶
Arabic	▶	▶	▶	▶
British	▶	▶	▶	▶
Chinese	▶	▶	▶	▶
Danish	▶	▶	▶	▶
Dutch	▶	▶	▶	▶
Finnich	▶	▶	▶	▶
French	▶	▶	▶	▶
German	▶	▶	▶	▶
Greek	▶	▶	▶	▶
Hindi	▶	▶	▶	▶
Hungarian	▶	▶	▶	▶
Italian	▶	▶	▶	▶
Japanese	▶	▶	▶	▶
Norwegian	▶	▶	▶	▶
Polish	▶	▶	▶	▶
Portuguese	▶	▶	▶	▶
Russian	▶	▶	▶	▶
Spanish	▶	▶	▶	▶
Swedish	▶	▶	▶	▶

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 type	Source signal	MELPe 600 bps	SPR 600 bps	TWELP 600 bps
Siren only	▶	▶	▶	▶
With voice	▶	▶	▶	▶

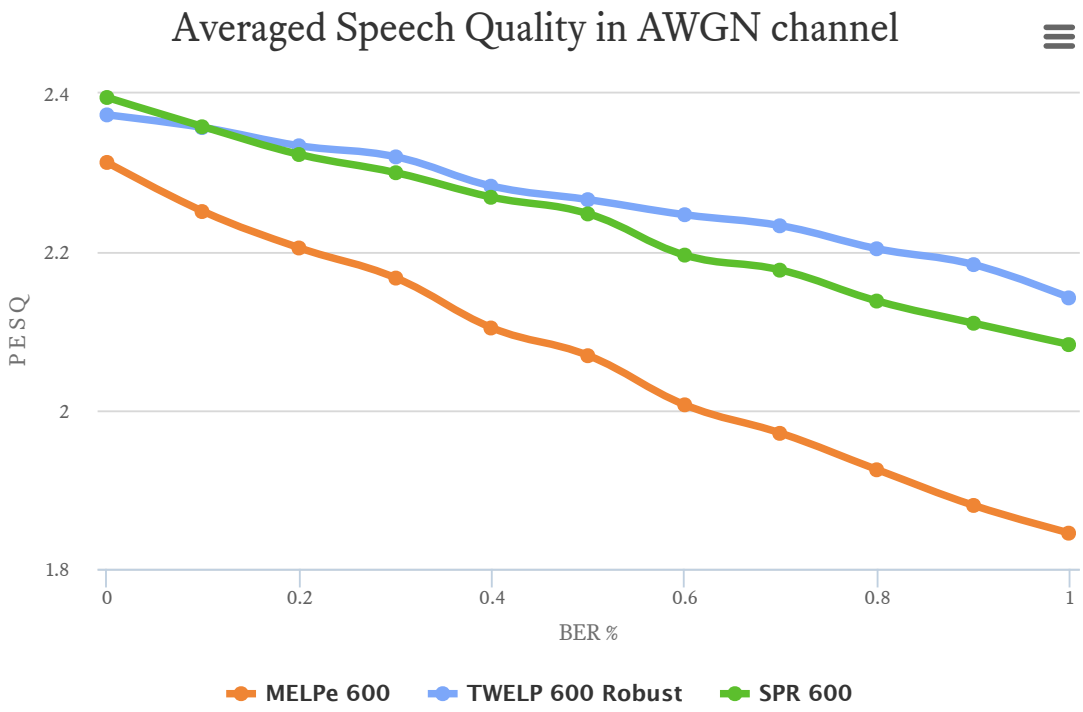
High Robustness To Acoustic Noise. In contrast to other LBR vocoders, TWELP vocoders are well robust to acoustic noise thanks to robust reliable method of pitch estimation and other features of TWELP technology.

Moreover, vocoder includes in-built Noise Cancellation—Speech Enhancement (NCSE) functionality that improves speech quality in noisy acoustic environment.

NCSE Mode	Source signal	MELPe 600 bps	SPR 600 bps	TWELP 600 bps
Disabled	▶	▶	▶	▶
Enabled	▶	▶	▶	▶

High Robustness To The Channel Errors. The diagram and table below show a dependence of the averaged speech quality for AWGN-noisy channel for different BER in comparison with other vocoders.

SPR vocoder is robust already to the channel errors. Also, special "robust" versions of the SPR and TWELP vocoders include FEC that are integrated with vocoder on base of "joint source-channel coding" approach that provides 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.



BER %	MELPe 600	TWELP 600 Robust	SPR 600
0.00	2.313	2.373	2.395
0.10	2.251	2.357	2.358
0.20	2.205	2.334	2.323
0.30	2.167	2.32	2.300
0.40	2.104	2.283	2.269
0.50	2.069	2.266	2.248
0.60	2.007	2.247	2.196
0.70	1.971	2.233	2.177
0.80	1.925	2.204	2.138
0.90	1.880	2.184	2.110
1.00	1.845	2.142	2.083

Additional Functionalities. The following additional functionalities are developed by DSPINI and can be integrated (on a request) into vocoders:

- Automatic Gain Control (AGC),
- Noise Cancellation for Speech Enhancement (NCSE)
- Voice Activity Detector (VAD),

- Tone Detection/Generation (Single tones and Dual tones). The tones are transmitted by the vocoder facilities.

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**
600	SPR™	80	100	8	Linear 16-bit PCM	48

Additional functionalities

Name	**Functionality**	**Technical characteristics**	
		Name	**Value**
AGC	Automatic Gain Control	Control range:	0 ... +20 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 ARM Cortex-M4 platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder	20.8	29.8	9.5	3.6	2.4	0.5
Voice Decoder	10.4					
Total	31.2					

Resources for TI's C64 DSP platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder	6.0	39.5	9.5	3.6	2.4	0.5
Voice Decoder	2.7					
Total	8.7					

Resources for TI's C55 DSP platform

Module	**MIPS* peak**	**Memory** (KBytes)				
		Program	**Data**			
			Constants	**Channel**	**Heap**	**Stack**
Voice Encoder	9.0	17.2	9.5	3.6	2.4	0.5
Voice Decoder	4.8					
Total	13.8					

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

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 SPR and TWELP technologies.

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)
- MELPe 600 bps speech samples (zip)
- SPR 600 bps speech samples (zip)
- TWELP 600 bps speech samples (zip)
- PC-evaluation package (zip) — on request
- User′s Guide document (pdf) — on request

Please contact us by e-mail:
request@dspini.com
or by phone:
+44 20 81 33 00 44
+33 9 70 40 33 99