SCELP: Low Delay Audio Coding with Noise Shaping Based on Spherical Vector Quantization

Coding of Audiovisual Contents

Miquel Oller Oliveras & Alvaro Scherk Fontanals

Barcelona, November 7, 2017

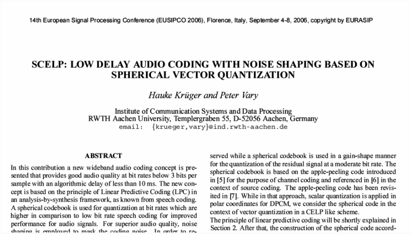
UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona

Table of Contents:

SCELP: An overview of the proposal

Abstract[1]



14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

SCELP: LOW DELAY AUDIO CODING WITH NOISE SHAPING BASED ON SPHERICAL VECTOR QUANTIZATION

Hauke Krüger and Peter Vary

Institute of Communication Systems and Data Processing
RWTH Aachen University, Templergraben 55, D-52056 Aachen, Germany
email: {krueger,vary}@ind.rwth-aachen.de

ABSTRACT

In this contribution a new wideband audio coding concept is presented that provides good audio quality at bit rates below 3 bits per sample with an algorithmic delay of less than 10 ms. The new concept is based on the principle of Linear Predictive Coding (LPC) in an analysis-by-synthesis framework, as known from speech coding. A spherical codebook is used for quantization at bit rates which are higher in comparison to low bit rate speech coding for improved performance for audio signals. For superior audio quality, noise shaping is employed to mask the coding noise. In order to re-

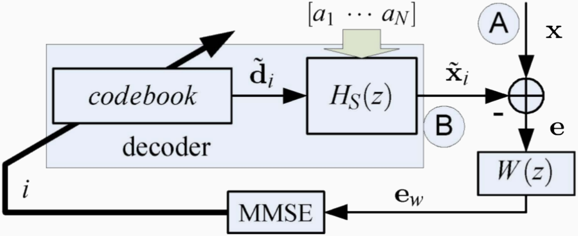
served while a spherical codebook is used in a gain-shape manner for the quantization of the residual signal at a moderate bit rate. The spherical codebook is based on the apple-peeling code introduced in [5] for the purpose of channel coding and referenced in [6] in the context of source coding. The apple-peeling code has been revisited in [7]. While in that approach, scalar quantization is applied in polar coordinates for DPCM, we consider the spherical code in the context of vector quantization in a CELP like scheme.
The principle of linear predictive coding will be shortly explained in Section 2. After that, the construction of the spherical code accord-

- Good audio quality at bit rates below 3 bits per sample
- Based in LPC with analysis-by-synthesis framework
- Spherical codebook for vectorial quatization
- Masked coding noise
- All-pole filter models spectral envelope of input signal

[1] Proposal by Haunke Krüger and Peter Vary, both from the Aachen University, Germany. Published in 2006

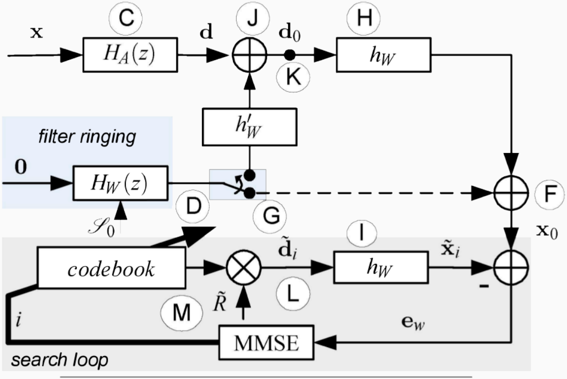
Adaptative Linear Prediction: Closed Loop Quantization

- Windowed segment of length L_{lpc} to obtain coefficients LPC
- Coefficients must be sent together with the cuantized residual LP
- Vector quantization instead of scalar



Optimized Excitation Search Modified

Complexity Reduction scheme:



$W(z) = H_A(z) * H_s(z) * W(z) = H_A(z) * H_W(z)$

Conclusion and Results

- Sample rate 16kHz
- Lv = 11
- Outperforms G.722
- Encoder: 20-25 WMOPS
- Decoder: 1-2 WMOPS
- Bitrates below 48 Kbps

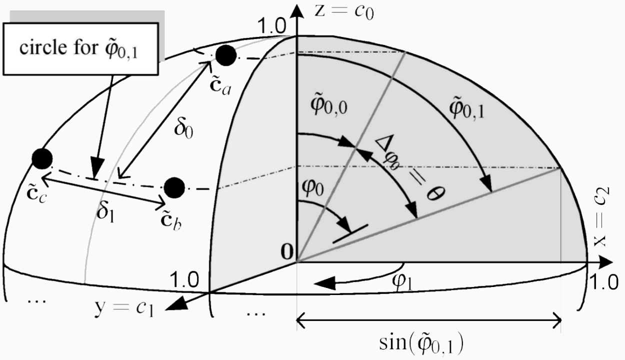
Spherical Codebook

Spherical Codebook

- Codewords composed by:
  - Gain (scalar)*
  - Shape (vector)

*We quantify it using a log quanifier

- Apple-peeling rule



Spherical Codebook Construction: 3D model

All layers expcept the last:

- Separation between layers:

  $\theta(N_{sp}) = \pi/N_{sp}$ where Nsp denotes the number of sublayers.

- Angle for the next layer:

$$\tilde{\phi}_{0,i_0} = (i_0 + 1/2) \cdot \theta(N_{sp})$$

Last layer:

- Number of centroids:

$$N_{sp,1}(\tilde{\phi}_{0,i_0}) = \lfloor \frac{2\pi}{\theta(N_{sp})} \cdot sin(\tilde{\phi}_{0,i_0}) \rfloor$$

- Centroid angle:

$$\tilde{\phi}_{1,i_1}(\tilde{\phi}_{0,i_0}) = (i_1 + 1/2) \cdot \frac{2\pi}{N_{sp,1}(\tilde{\phi}_{0,i_0})}$$

Spherical Codebook: Cartesian Coordinates of centroids

- Cartesian coordinates:[2]

$$x_1 = r \cdot cos(\phi_1)$$

$$x_2 = r \cdot sin(\phi_1) \cdot cos(\phi_2)$$

$$x_3 = r \cdot sin(\phi_1) \cdot sin(\phi_2) \cdot cos(\phi_3)$$

$$\vdots$$

$$x_{n-1} = r \cdot sin(\phi_1) \cdot \ldots \cdot sin(\phi_{n-2}) \cdot cos(\phi_{n-1})$$

$$x_n = r \cdot sin(\phi_1) \cdot \ldots \cdot sin(\phi_{n-2}) \cdot sin(\phi_{n-1})$$

[2] Font: Wikipedia: https://en.wikipedia.org/wiki/N-sphere

Spherical Codebook: N-Spherical Coordinates of Centroids

- Spherical coordinates:[3]

$$r = \sqrt{x_n^2 + x_{n-1}^2 + \ldots + x_2^2 + x_1^2}$$

$$\phi_1 = arccot \frac{x_1}{\sqrt{x_n^2 + x_{n-1}^2 + \ldots + x_2^2}} = arccos \frac{x_1}{\sqrt{x_n^2 + x_{n-1}^2 + \ldots + x_2^2 + x_1^2}}$$

$$\phi_2 = arccot \frac{x_2}{\sqrt{x_n^2 + x_{n-1}^2 + \ldots + x_3^2}} = arccos \frac{x_2}{\sqrt{x_n^2 + x_{n-1}^2 + \ldots + x_2^2}}$$

$$\vdots$$

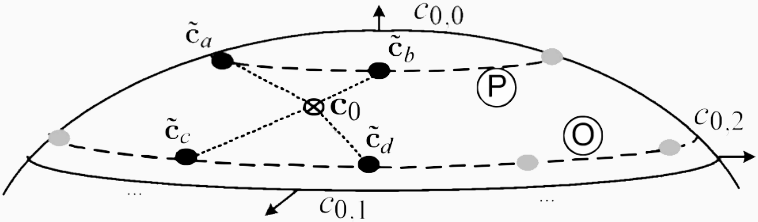
$$\phi_{n-2} = arccot \frac{x_{n-2}}{\sqrt{x_n^2 + x_{n-1}^2}} = arccos \frac{x_{n-2}}{\sqrt{x_n^2 + x_{n-1}^2 + x_{n-2}^2}}$$

$$\phi_{n-1} = 2 \cdot arccot \frac{x_{n-1} + \sqrt{x_n^2 + x_{n-1}^2}}{x_n}$$

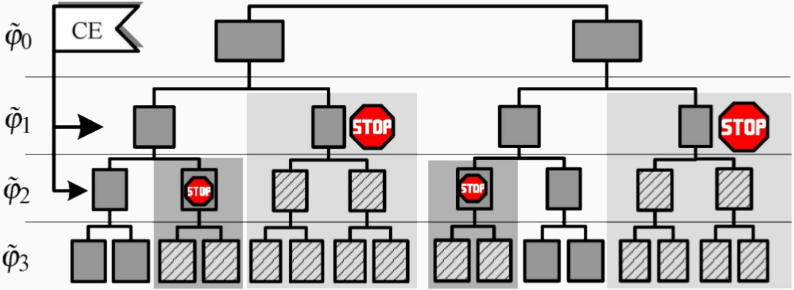
[3] Font: Wikipedia: https://en.wikipedia.org/wiki/N-sphere

Preselection

- Select neighbour centroids surrounding the unquantized signal
- 2 new layer candidates for each layer. Select the upper and lower circumferences for each ϕ_i
- 2^{L_v-1} total neighbours which are the candidates



Candidate-exclusion



Minimize the partial distorsion[4]:

- Reduce the computation complexity
- Loss in quatization SNR

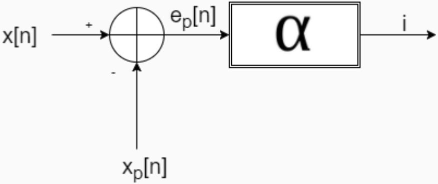
[4] Partial Distortion: $D_i = \sum_{j=0}^{i_0} (x_{0,j} - \tilde{\mathbf{x}}_{i,j}|_{[0\ldots i_0]})^2$

Our Implementation

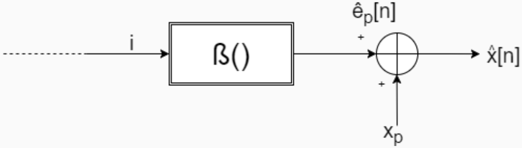
Our Implementation: Open-Loop (Basic Idea)

Predict with the unquantized values

- Encoder:



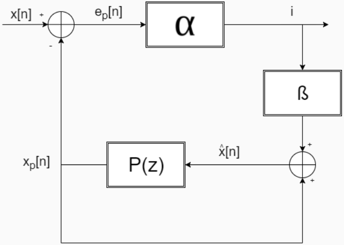
- Decoder:



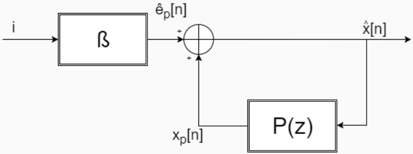
Our Implementation: Open-Loop (Advanced Scheme)

Use the quantized values in the predictor

- Encoder:



- Decoder:



Our Simulations & Results

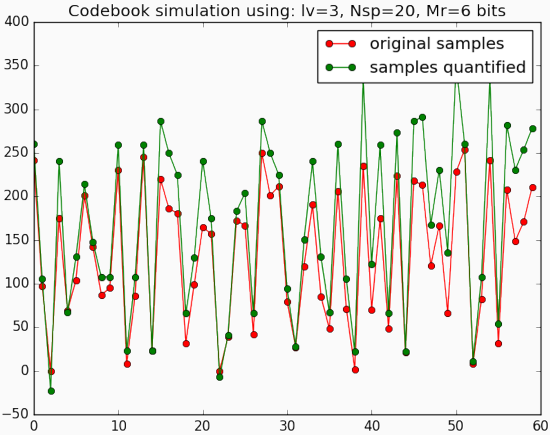
Route-Map

- Test the codebook
- Test the open-loop scheme (basic)
- Compare with uniform quanitfier

Metrics:

$$SNR = 10log(\frac{P_{x_p}}{P_{eq}})$$
$$P_{eq} = \frac{1}{N}\sum e_q^2$$
$$P_{xp} = \frac{1}{N}\sum x_p^2$$
$$\text{Prediction error: } e_q = x_p - \hat{x}_p$$
$$\text{Original signal: } x_p$$
$$\text{Quantified signal: } \hat{x}_p$$

Spherical Codebook Testing: (l_v, N_{sp}, M_r)=(3,20,6)

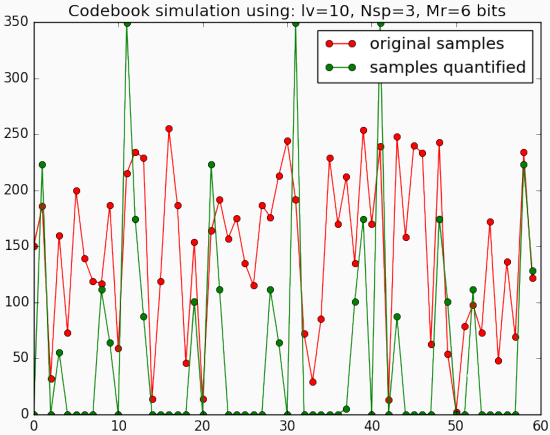


Number of centroids: 429 $\rightarrow$ 9 bits

$SNR = 8.68dB$

Reduction: 24/15 (38%)

Spherical Codebook Testing: (l_v, N_{sp}, M_r)=(10,3,6)

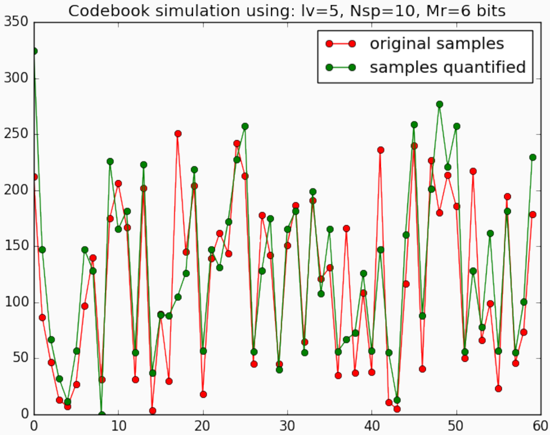


Number of centroids: 26244 $\rightarrow$ 15 bits

$SNR = 1.73dB$

Reduction: 80/21 (73.7%)

Spherical Codebook Testing: (l_v, N_{sp}, M_r)=(5,10,6)



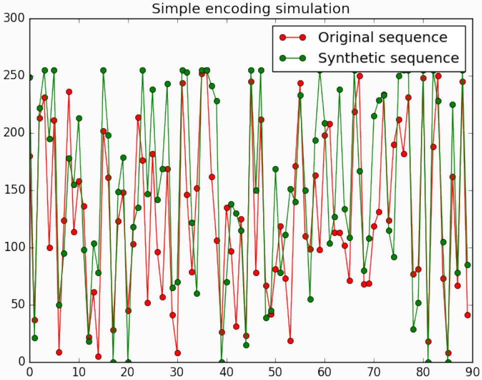
Number of centroids: 12400 $\rightarrow$ 14 bits

$SNR = 7.85dB$

Reduction: 40/20 (50%)

General Prediction: Using the Basic Scheme

Using the predictor based on the unquanitzed values:

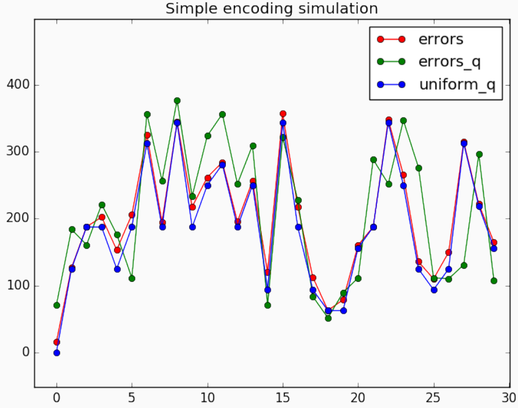


Codebook definiton: (l_v, N_{sp}, M_r)=(5,10,6) $\rightarrow$ 20 bits/ 5 samples

$SNR = 8.05dB$

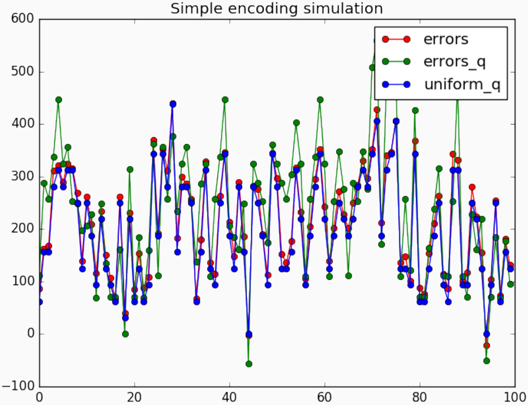
Reduction: 40/20 (50%)

Error Prediction Test



Codebook definiton: $(l_v,\ N_{sp},\ M_r)$=(5,10,6) $\rightarrow$ 20 bits / 5 samples
Uniform quantifier: 4 bits/sample
$SNR_{spherical} = 10.8\ dB$
$SNR_{uniform} = 22.4\ dB$

Error Prediction Test



Codebook definiton: $(l_v,\ N_{sp},\ M_r) = \rightarrow$ 20 bits / 5 samples

Uniform quantifier: 4 bits/sample

$SNR_{spherical} = 10.8\ dB$

$SNR_{uniform} = 22.4\ dB$

Conclusions

Conclusions

- Low bitrate coding acomplish
- Too complex model scheme
- Open-loop configuration can behave unexpectedly
- Big dimentional CB $\rightarrow$ great time-consuming impact
- Further improves are needed

SCELP: Low Delay Audio Coding with Noise Shaping Based on Spherical Vector Quantization

Coding of Audiovisual Contents

Miquel Oller Oliveras & Alvaro Scherk Fontanals

Barcelona, November 7, 2017

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Tècnica Superior d'Enginyeria de Telecomunicació de Barcelona