1 Introduction

- Linear prediction of speech based on 1-norm minimization has already proved to be an interesting alternative to 2-norm minimization.
- Choosing the 1-norm as a convex relaxation of the 0-norm, the corresponding linear prediction model offers a sparser residual better suited for coding applications.
- Purpose of this paper is to overcome the mismatch between 0-norm minimization and 1-norm minimization while keeping the problem solvable with convex optimization tools.
- Experimental results prove the effectiveness of the reweighted 1-norm minimization, offering better coding properties compared to 1-norm minimization.

2 Sparse Linear Prediction

- The problem considered in this paper is based on the following Auto-Regressive (AR) speech production model, where a sample of speech $x(n)$ is written as a linear combination of $K$ past samples:

$$x(n) = \sum_{k=0}^{K} a_k x(n-k) + r(n), \quad 0 < n \leq N,$$

where $\{a_k\}$ are the prediction coefficients and $r(n)$ is the driving noise process (commonly referred to as the prediction residual).
- The prediction coefficient vector $a$ (order $K$) is found by minimizing the $l_1$-norm of the residual $r$:

$$\hat{a} = \arg \min_{a} \| r \|_1 \quad \text{s.t.} \quad r = x - Xa.$$

An interesting alternative problem formulation is obtained when sparsity is also imposed on the predictor:

$$\hat{a} = \arg \min_{a} \| r_{+} \|_1 + \gamma \|a\|_1 \quad \text{s.t.} \quad r = x - Xa,$$

where $r_{+}$ is the positive part of $r$.
- In this case the sparse structure of the predictor (in this case high order) allows a joint estimation of a short-term and a long-term predictor.

3 Reweighted 1-norm

- Iteratively Reweighted 1-norm Minimization of the Residual

Inputs: speech segment $x$

Outputs: predictor $\hat{a}$ and residual $\hat{r}$

1. $\hat{a} = \arg \min_{a} \| W r \|_1$
2. $W = D^{-\gamma}$
3. $\hat{r} = x - X\hat{a}$

- Iteratively Reweighted 1-norm Minimization of Residual and Predictor

Inputs: speech segment $x$

Outputs: predictor $\hat{a}$ and residual $\hat{r}$

1. $\hat{a} = \arg \min_{a} \| W r \|_1$
2. $W = D^{-\gamma}$
3. $\hat{r} = x - X\hat{a}$

4 Experiments

- The reweighting process will balance the dependence on the magnitude of the 1-norm, changing the cost function and moving the problem towards the 0-norm minimization.
- The weights are chosen as the inverse of the magnitude of the residual.
- $\epsilon > 0$ is used to provide stability when a component of $\hat{r}$ goes to zero.
- $\| \hat{r} \|_1 \leq \| \hat{r} \|_1$.
- The halting criterion can be chosen as either a maximum number of iterations or as a convergence criterion.

5 Validation

- The residual sequence is coded with sparse multipulse techniques ($M = 20$) after the reweighted scheme. The frame length is $N = 160$ (20 ms). The prediction order is $K = 10$.

In MPE it is not necessary to calculate the positions of the nonzero pulses are located (as it is usually done in MPE coding), we simply exploit the information coming out of the predictive analysis. We are then clearly moving our problem towards a more synergistic way to code a signal.

Comparison between the MPE residual estimation methods in terms of Segmental SNR and Mean Opinion Score (PESQ evaluation). A 95% confidence intervals is given for each value. Each signal is coded with 9500 bits.

6 Discussion

- With just few iterations, we were able to move the error minimization criterion toward the 0-norm solution, showing general improvements over conventional 1-norm minimization in coding purposes.
- We have no prior knowledge of where the residual should be nonzero. This brings the bit allocated to describe the position of few samples to significantly increase the rate.
- An interesting case would be to structure the reweighting process by imposing where we would like to have the nonzero pulses located, as in RPE encoding scheme.

References