PARAMETER OPTIMIZATION METHODS FOR THE EDS MODEL

Stijn Lamens, Wim D'haes

University of Antwerp
Vision Lab
Groenenborgerlaan 171 - 2020 Antwerp - Belgium

ABSTRACT

Various methods have been proposed to estimate the parameters of the exponentially damped sinusoidal (EDS) model. As most estimation methods do not yield optimal parameter values, it is often useful to apply further optimization. In this work, various optimization algorithms which can be used for both frequency and damping factor parameters are developed and compared. The optimization methods include Newton and Gauss-Newton methods, which were improved by applying a regularization factor resulting in the Levenberg-Marquardt method. Estimation and optimization algorithms were implemented and tested in two different methodologies: a method where optimization is performed on all partials simultaneously, and a pseudo-simultaneous method where optimization is performed on individual partials in an iterative manner.

1. INTRODUCTION

Sinusoidal modeling of audio signals is a powerful and flexible method, because of the many applications it allows, like low-bitrate coding and time/pitch-scale modifications. The original model [1] is limited to sinusoidal components with a constant amplitude and frequency, making it unsuitable for fast varying signals unless small time-domain windows are used for analysis and synthesis. By using exponentially damped sinusoids [2], the model is made more suitable for exponentially decaying sounds, which are common in music (e.g. piano and percussive instruments). It consists of K damped sinusoids with frequencies ω_k, complex amplitudes $A_k = A_k^r + iA_k^i$ and damping factors ρ_k, for k = 0,...,K-1, and can be expressed as follows:

$$y_n = \sum_{k=0}^{K-1} e^{(\rho_k \frac{n}{N})} \left[A_k^r \cos(2\pi\omega_k \frac{n}{N}) - A_k^i \sin(2\pi\omega_k \frac{n}{N}) \right]$$

Various parameter estimation methods have been proposed in the literature, like spectral peak-picking [1], matching pursuits [3] and subspace-based methods [2, 4]. The first two methods can be implemented in a computationally cheap manner, but lead to sub-optimal results. Subspace-methods result in a considerably higher modeling accuracy, but have a high computational complexity, namely $\mathcal{O}(N^3)$. A way to improve the results of spectral peak picking and matching pursuit methods, is to apply an optimization method (e.g. Newton) to the estimated parameters. However, this method isn't used often because of it's high computational complexity (typically $\mathcal{O}(NK^2)$ when all frequency parameters are optimized simultaneously). Recently, it was shown that the computational complexity of the optimization can be reduced to $\mathcal{O}(N \log(N))$ in the case of stationary sinusoids and if analysis windows with band-limited frequency responses are used [5]. In this work, optimization methods for the exponentially damped model will be discussed.

2. COMPLEX AMPLITUDE CALCULATION

If the frequencies and damping factors of the sinusoidal components are known, their complex amplitudes can be calculated by using a least squares method. The partial derivatives, with respect to each A_l^r and A_l^i, for l = 1,...,K, of the squared difference between the input signal x_n and the model y_n are put to zero:

$$\frac{\partial}{\partial A_l^i} \sum_{n=0}^{N-1} (x_n - y_n)^2 = 0, \frac{\partial}{\partial A_l^r} \sum_{n=0}^{N-1} (x_n - y_n)^2 = 0$$

These expressions are lineair in A_l^i and A_l^i respectively, which means that all amplitudes can be calculated simultaneously by solving a 2K-dimensional linear system. An alternative method is to calculate the amplitudes in a sequential manner, in which case the least squares difference between $x_n^{(l-1)}$ (which is the original signal where the l-1 previously analysed partials have been subtracted from) and the l-th partial is considered and the above system reduces to a 2-dimensional system. The computational complexity of the sequential method is $\mathcal{O}(NK)$, where the simultaneous method scales $\mathcal{O}(NK^2)$. However, the sequential method results in incorrect values if the sinusoidal components have an overlapping frequency respons, which is especially the case with exponentially damped sinusoids. This will be discussed further in section 5.1.

3. FREQUENCY OPTIMIZATION

The partial derivatives of the error function which respect to ω_l are given by

$$\frac{\partial}{\partial\omega_l}\sum_{n=n_0}^{N-1}(x_n-y_n)^2=\sum_n 2(x_n-y_n)e^{(\rho_l\frac{n}{N})}(2\pi\frac{n}{N})\left[A_l^r\sin(2\pi\omega_l\frac{n}{N})+A_l^i\cos(2\pi\omega_l\frac{n}{N})\right] \quad (1)$$

which are non-linear in ω_l. Therefore, they will be estimated by spectral peak picking and iteratively optimized using Newton or Gauss-Newton methods, resulting in a non-linear least squares method.

3.1. Newton optimization

A Newton step for frequency optimization is based on a quadratic approximation of the error function and typically looks as follows:

$$\omega=\hat{\omega}-\mathbf{H}^{-1}\mathbf{h}$$

meaning a new frequency value ω is estimated based on a previous estimation $\hat{\omega}$ at each optimization step. The elements of the gradient $\mathbf{h}$ are defined as the first order partial derivatives of the error function, and are therefore equal to equation (1). In case of the simultaneous method, the hessian $\mathbf{H}$ is a K x K matrix for which each element $\mathbf{H_{l,k}}$ is defined as $\frac{\partial}{\partial\omega_l}\frac{\partial}{\partial\omega_k}(x_n-y_n)^2$. In case of the sequential and pseudo-simultaneous method, $\mathbf{H}$ and $\mathbf{h}$ are scalars.

3.2. Gauss-Newton optimization

As an alternative method, a Gauss-Newton optimization step can be performed. This is done by making a linear approximation of the model, by using a first order Taylor approximation around point $\hat{\omega}$, which results in

$$\sum_{k=0}^{K-1}e^{(\rho_k\frac{n}{N})}\left[A_k^r\cos(2\pi\hat{\omega_k}\frac{n}{N})-A_k^i\sin(2\pi\hat{\omega_k}\frac{n}{N})\right]+(2\pi\frac{n}{N})e^{(\rho_k\frac{n}{N})}\left[-A_k^r\sin(2\pi\hat{\omega}_k\frac{n}{N})-A_k^i\cos(2\pi\hat{\omega}_k\frac{n}{N})\right](\omega_k-\hat{\omega}_k)$$

denoted as $\tilde{y}_n$. The partial derivatives of the error function with respect to each $(\omega_l-\hat{\omega}_l)$ are linear so we can find it's minimum by putting them to zero:

$$\frac{\partial\sum_n(x_n-\tilde{y}_n)^2}{\partial(\omega_l-\hat{\omega}_l)}=0$$

resulting in the linear system

$$\mathbf{GA}=\mathbf{g}$$

with

$$\mathbf{A}_l=\omega_l-\hat{\omega}_l$$

yielding

$$\omega_l=\hat{\omega}_l+\mathbf{G^{-1}g}$$

It can easily be shown that $\mathbf{g}$ equals $-\mathbf{h}$ and that $\mathbf{G}$ and $\mathbf{H}$ only differ in the elements on their diagonal.

3.3. Regularized variants

It is commonly known that Newton and Gauss-Newton methods have some shortcomings. One of them is that the direction of the optimization step may be incorrect if the system matrix is not positive definite. There is no method to verify this condition, but in practice one can check if the optimization step doesn't point in the opposite direction of the gradient and the value of the error function decreases after each optimization step. If these conditions are not satisfied, the direction of the optimization step is not correct or the step size is too large. Both problems can be dealt with by adding a regularization term λI, with I being the unit matrix. For large values of λ, the optimization nearly equals the gradient descent method, with a learning speed of $\frac{1}{\lambda}$. For small values of λ, the effect of this addition is small, hence the optimization step will be close the Newton (Gauss-Newton) optimization step. A commonly used strategy to determine the optimal value for λ at each step is by initializing λ with a small value, increasing it with a factor 10 each time an optimization step fails (in which case ω is not updated), and decrease it otherwise. The Gauss-Newton method with a correction factor taken into account is also known as the Levenberg-Marquardt method and is known to be one of the most powerful and robust optimization techniques [6].

4. DAMPING FACTOR ESTIMATION AND OPTIMIZATION

The damping factors can be optimized in a similar way after an initial estimation has been made. Both the Newton and Gauss-Newton optimization methods were implemented in this context, where the same considerations apply as in the case of frequency optimization. The initial estimation is made by considering the frequency responses of ω_k within the original analysis window and a window that has been shifted by p samples. It can be shown that

$$\rho_k=\frac{N-p}{p}\ln\frac{\sum_{n=p}^{N-1}y_ne^{(2\pi i\omega_k\frac{n}{N})}}{\sum_{n=0}^{N-p-1}y_ne^{(2\pi i\omega_k\frac{n}{N})}} \quad (2)$$

If the input signal x_n is conform to the assumptions of the model, ρ_k can be estimated by substituting x_n for y_n in (2).

5. COMPLETE ANALYSIS OF A GIVEN INPUT SIGNAL

5.1. Sequential, pseudo-simultaneous and simultaneous optimization methods

To do a complete analysis of an input sound, various strategies can be applied to obtain the parameters for each sinusoidal component. The simplest strategy is to analyse all partials in a sequential order. This consists of a repetitive process of estimating and optimizing the parameters of the strongest component in the signal and subtracting that component from the signal, until a certain stopping criterium is reached. A possible stopping criterium could be that a maximum number of components has been analysed, or that the power of the residual signal has decreased below a certain treshold. This method is computationally cheap: it has a complexity of $\mathcal{O}(N)$ per optimization step for one partial. A major drawback of this method is that the estimated parameters will not be accurate if the frequency responses of some components overlap. The simultaneous method is able to determine the exact parameters but the computational complexity of one simultaneous optimization step is $\mathcal{O}(K^2N)$. The sequential method can be improved by iterating a number of times over all the partials and reanalysing them using a signal where all other partials have been subtracted from. We call this the pseudo-simultaneous method.

5.2. Obtaining the various parameters of a partial

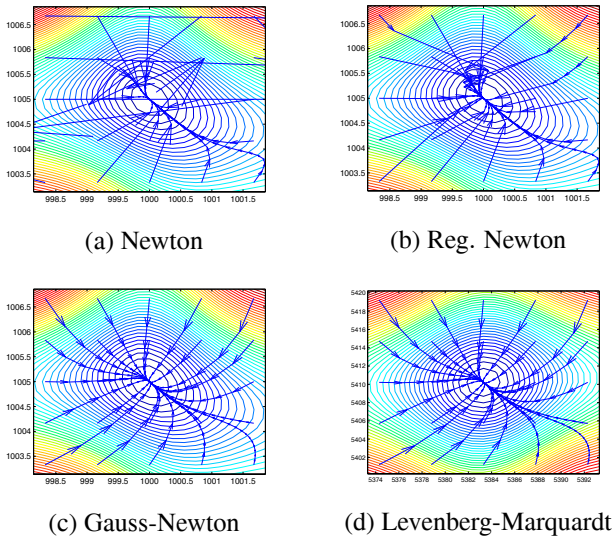
For each partial that is analysed, three different parameters have to be obtained. The main problem in this context is that each calculation or optimization step implicitely assumes that the other parameters are known. In practice, an estimation of the parameters is often known and the optimization steps can be applied on the estimated parameters. This leads to suboptimal results, but when the analysis algorithm performs the several calculation and optimization steps in a well-chosen order, all parameters will converge to their optimal value. A possible method to obtain the parameters of one or more partials is given in algorithm 1, where we chose to optimize ω before ρ because the ω can be optimized quite accurate when ρ is incorrect. When, on the other hand, the estimated frequency value of a partial is far from it's optimum, damping optimization will fail because the error function will be at it's minimum for $\rho_l = -\infty$. The stop criterium can be chosen depending on the needs of the situation.

Algorithm 1 AnalyseNPartials(x_n, N)

```
ω = estimFreq(x_n, N);
ρ = estimDamping(x_n, ω, N);
repeat
    for j = 0 to nrOfFreqOptimSteps do
        A = calcAmp(x_n, ω, ρ, N);
        ω = freqOptim(x_n, A, ω, ρ, N);
    end for
    for j = 0 to nrOfDampingOptimSteps do
        A = calcAmp(x_n, ω, ρ, N);
        ρ = dampingOptim(x_n, A, ω, ρ, N);
    end for
until stop criterium reached
```

6. TEST RESULTS

A comparison between Newton, Gauss-Newton and their regularized variants was made. A synthetic test signal consisting of two partials with a strongly overlapping frequency response was used. It's parameters are given by ω = $[1000, 1005]$, $\mathbf{A} = [3 + 2i, 7 - 4i]$, $\rho = [-5, -9]$ and a fixed length of N = 8192 samples. In figure 1, the optimization trails of the tested methods are shown for various sets of initial frequency values, from which one can conclude that the Newton method behaves very bad at the edges of the frequency range considered, so that the mimimum is not reached. Regularized Newton performs considerably better and is able to find the optimum for each tested set of initial values. The Gauss-Newton method leads to very good results in this test, as there is no regularization needed for any of the initial values. These experiments were also executed for damping factor optimization, yielding similar results. The pseudo-simultaneous and simultaneous Newton and Gauss-Newton methods were compared in terms of error decrease and computational time per optimization sequence. For this test, we used only the regularized variants of the algorithms. An optimization sequence consists of 4 frequency optimization steps, followed



(a) Newton (b) Reg. Newton

(c) Gauss-Newton (d) Levenberg-Marquardt

Fig. 1. Different simultaneous frequency optimization algorithms applied to a signal consisting of two partials

by 3 damping factor optimization steps. Each optimization step was preceded by a complex amplitude calculation. Various synthetic test signals with a different number of partials were used. Their parameters were chosen so that all partials were closely spaced in the frequency domain and had an overlapping frequency response. Initial frequency values were obtained by pertubating the frequency values of the testsignal, and damping factor estimations were obtained by the method described in section 4. The error decrease during the first 30 optimization steps is shown in figure 3, from which we can conclude that the simultaneous methods need less steps than pseudo-simultaneous method. For large values of K however, the pseudo-simultaneous method will be faster in terms of computational time needed. This is shown in figure 2, where the computational time of one optimization sequence of each method is shown with respect to the number of partials. This shows that the complexity of the pseudo-simultaneous methods is only slightly larger than $\mathcal{O}(NK)$, while simultaneous methods scale about $\mathcal{O}(NK^2)$.

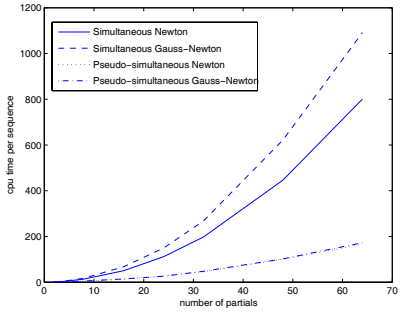
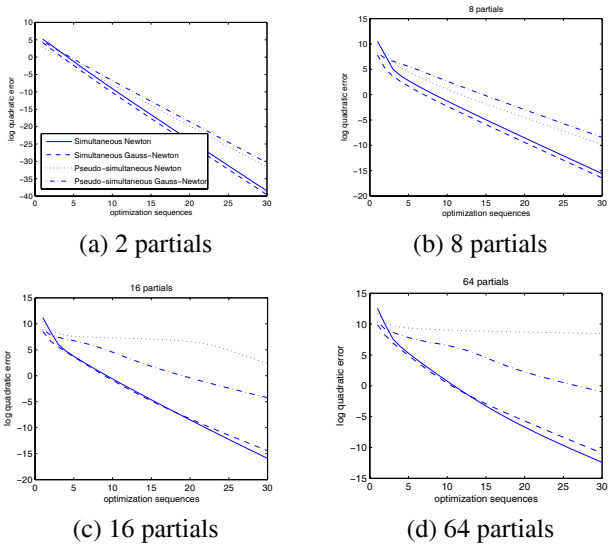


Fig. 2. Computational time per optimization sequence

7. CONCLUSIONS AND FUTURE WORK

The general conclusion one can make is that the Levenberg-Marquardt method, performed in a pseudo-simultaneous manner, is an efficient method for the optimization of frequency and damping factor values of the EDS model. One of the remaining problems of the implemented optimization methods is that the optimization steps for each parameter are not optimal because they depend on other unknown parameters. The frequency and damping factor could be optimized simultaneously, but that isn't worth the computational complexity because the frequency depends only weakly to the damping factor. Generally speaking, optimizing multiple parameters simultaneously is efficient when the mutual dependence of the parameters is strong enough to justify the additional computational cost. In the case of the EDS model, it is clear that the amplitudes of overlapping partials depend on each other, so it makes sense to calculate them in a simultaneous way. There is also a strong dependence between the complex amplitude and



(a) 2 partials (b) 8 partials

(c) 16 partials (d) 64 partials

Fig. 3. Comparing the logarithmic error decrease of the tested optimization methods

the frequency and between the amplitude and damping factor of each partial, forcing us to execute a amplitude computation step before each optimization step, still yielding non-optimal optimization steps. If this dependency could be taken into account, the required number of optimization steps might be reduced further.

8. REFERENCES

[1] R.J. McAulay and T.F. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," in *IEEE Transactions on Acoustics, Speech and Signal Processing 34(4): 744–754*, 1987.

[2] Kris Hermus, *Signal Subspace Decompositions for Perceptual Speech and Audio Processing*, Ph.D. thesis, K.U. Leuven, December 2004.

[3] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, 1993.

[4] Rémy Boyer and Karim Abed-Meraim, "Audio modeling based on delayed sinusoids," in *IEEE Transactions on speech and audio processing, vol. 12, no. 2*, March 2004.

[5] W. D'haes, "A highly optimized nonlinear least squares technique for sinusoidal analysis: From O(K^2N) to O(N log(N))," in *Proceedings of the 116th Audio Engineering Society Convention (AES)*, Berlin, Germany, May 2004, Audio Engineering Society (AES).

[6] Christopher M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1996.