Robust Pitch Estimation Using an Optimal Filter on Frequency Estimates

Sep. 4, 2014

Sam Karimian-Azari, Jesper Rindom Jensen, and Mads Græsbøll Christensen
{ska, jrj, mgc}@create.aau.dk

Audio Analysis Lab, AD:MT,
Aalborg University, Denmark

Funded by the Villum foundation and DFF–1337-00084
Agenda

Introduction
Harmonic Signal Model
Pitch Estimation

The WLS Pitch Estimator

Proposed Method
Problem Formulation
Filtering UFEs
Cramér-Rao Lower Bound

Experimental Results

Conclusion
Introduction

In speech and audio signal processing applications such as:

- hearing-aids
- teleconference systems

we can design filters based on parametric models.

**Pitch** estimate is often necessary in some applications such as:

- Coding
- Enhancement
- Separation
- Compression
Harmonic Signal Model:

\[ s(n, \theta) = \sum_{l=1}^{L} \alpha_l e^{j(\omega_l n + \varphi_l)}, \]  

\[ \theta = [\alpha_1, \varphi_1, \omega_1, \ldots, \alpha_L, \varphi_L, \omega_L]^T, \]  

where \( \omega_l = l\omega_0 \) (for \( l = 1, \ldots, L \)),

- \( L \): number of sinusoids
- \( \alpha_l \): real magnitudes
- \( \omega_0 \): fundamental frequency
- \( \varphi_l \): phases of harmonics
Existing approaches*:

- **Filtering Methods**: Based on filtering the observed signal that pass power undistorted frequency of harmonics, e.g., optimal adaptive designs.

- **Subspace Methods**: Based on the principles of subspace orthogonality, e.g., MUSIC method

- **Statistical Methods**: Based on maximizing the likelihood function of pitch, e.g., the non-linear least-squares (NLS), and weighted least squares (WLS) [H. Li, P. Stoica, and J. Li 2000].

* [M.G.Christensen and A.Jakobsson 2009]
The Weighted Least Squares (WLS) Parameter Estimator
[H. Li, P. Stoica, and J. Li 2000]

The observed signal is contaminated by Gaussian noise \( v(n) \) with variance of \( \sigma^2 \)

\[
x(n) = \sum_{l=1}^{L} \alpha_l e^{j(\omega_l n + \varphi_l)} + v(n). \tag{2}
\]

With the assumption of white Gaussian noise, the WLS pitch estimator yields

\[
\hat{\omega}_0 = \frac{1}{\sum_{l=1}^{L} (l \hat{\alpha}_l)^2} \sum_{l=1}^{L} l \hat{\alpha}_l^2 \hat{\omega}_l, \tag{3}
\]

where \( \hat{\omega}_l \) and \( \hat{\alpha}_l \) can be estimated using different methods, e.g., the DFT, MUSIC, ESPRIT, NLS, Capon, and APES.
The WLS Pitch Estimator

Drawbacks

The WLS pitch estimator is **computationally efficient** with good **statistical performance**,

BUT! It is not optimal for cases:

- nonidentical noise variances among harmonics, e.g., colored noise.
- spurious frequency estimates.
- missing harmonics.
Proposed Method

Assumption

At a high narrowband SNR, i.e.,

\[
\text{SNR}(\omega_l) = \frac{\alpha_l^2}{\sigma^2} \gg 1, \quad (4)
\]

the harmonic frequency \(\omega_l\) is perturbed with a real angular noise [S. Tretter 1985]

\[
\Delta \omega_l(n) = \frac{v(n)}{\alpha_l} \sin(l\omega_0 n + \varphi_l), \quad (5)
\]

which has a normal distribution with zero mean and variance

\[
E\{(\Delta \omega_l)^2\} = \frac{1}{2 \text{SNR}(\omega_l)}, \quad (6)
\]

where the white Gaussian noise has a homogeneous power spectrum \(\Phi(\omega) = \sigma^2\) across frequencies \(\omega \in [0, \pi]\).

In general,

\[
E\{(\Delta \omega_l)^2\} = \frac{\Phi(\omega_l)}{2 \alpha_l^2}. \quad (7)
\]
Proposed Method

Problem Formulation

We can approximate \( x(n) = \sum_{l=1}^{L} \alpha_l e^{j(\omega_l n + \varphi_l)} + v(n) \) like,

\[
x(n) \approx \sum_{l=1}^{L} \alpha_l e^{j(\omega_l n + \Delta \omega_l(n) + \varphi_l)}.
\]

Harmonic frequencies:

\[
\Omega = [\omega_1, \omega_2, \ldots, \omega_L]^T = d_L \omega_0,
\]

where \( d_L = [1, 2, \ldots, L]^T \).

Unconstrained frequency estimates (UFEs):

\[
\hat{\Omega} = [\hat{\omega}_1, \hat{\omega}_2, \ldots, \hat{\omega}_L]^T = \Omega + \Delta \Omega = d_L \omega_0 + \Delta \Omega,
\]

where \( \Delta \Omega = [\Delta \omega_1, \Delta \omega_2, \ldots, \Delta \omega_L]^T \).
Proposed Method
Filtering UFEs

Apply a filter \( h \in \mathbb{R}^L \)

\[
\hat{\omega}_0 = h^T \hat{\Omega} = h^T d_L \omega_0 + h^T \Delta \Omega. \tag{11}
\]

With the distortionless constraint that \( h^T d_L = 1 \), the mean squared error (MSE) of the unbiased estimator is given by

\[
\text{MSE}[\hat{\omega}_0] = E\{(\hat{\omega}_0 - \omega_0)^2\} = E\{(h^T \Delta \Omega)(\Delta \Omega^T h)\} = h^T \Phi_{\Delta \Omega} h. \tag{12}
\]
Proposed Method
Formulation

Independent UFEs would be implicitly uncorrelated, i.e., \( E\{\Delta \omega_i \Delta \omega_k\} = 0 \) for \( i \neq k \), when harmonics are well-separated and the narrowband SNRs are high enough;

\[
\Phi_{\Delta \Omega} = E\{\Delta \Omega \Delta \Omega^T\}
\]

\[
= \text{diag}\left\{ \frac{\Phi(\omega_1)}{2 \alpha_1^2}, \frac{\Phi(\omega_2)}{2 \alpha_2^2}, \ldots, \frac{\Phi(\omega_L)}{2 \alpha_L^2} \right\}. \quad (13)
\]

In practice:

\[
\Phi_{\Delta \Omega} = E\{(\hat{\Omega} - E\{\hat{\Omega}\})(\hat{\Omega} - E\{\hat{\Omega}\})^T\}, \quad (14)
\]

where \( E\{\cdot\} \) is the mathematical expectation, and

\[
E\{\hat{\Omega}(n)\} = \frac{1}{M} \sum_{m=0}^{M-1} \hat{\Omega}(n-m). \quad (15)
\]
Optimal filter is given as the solution to

\[
\min_h \ h^T \Phi_{\Delta \Omega} h
\]

subject to

\[
h^T d_L = 1
\]

Minimum variance distortionless response (MVDR):

\[
h_{\text{MVDR}} = \Phi_{\Delta \Omega}^{-1} d_L (d_L^T \Phi_{\Delta \Omega}^{-1} d_L)^{-1}.
\]

Maximum likelihood (ML):

Assuming white Gaussian noise with the variance \( \sigma^2 \)

\[
\Phi_{\Delta \Omega}^{-1} = \frac{2}{\sigma^2} \text{diag}\left\{ [\alpha_1^2, \alpha_2^2, \ldots, \alpha_L^2] \right\},
\]

\[
h_{\text{ML}} = \frac{1}{\sum_{l=1}^L (l\alpha_l)^2} [\alpha_1^2, 2\alpha_2^2, \ldots, L\alpha_L^2]^T.
\]
Asymptotic phase estimates of harmonics at \( n \in [0, N - 1] \):

\[
\hat{\Psi}(n) = d_L \omega_0 n + \Xi_L + \Delta \Omega(n)
\]  

(19)

Likelihood Function:

\[
P(\hat{\Psi}, \xi) \sim \mathcal{N}(d_L \omega_0 + \Xi_L, \Phi_{\Delta \Omega})
\]  

(20)

where \( \xi = [\omega_0, \Xi_L]^T \) and \( \Xi_L = [\varphi_1, \varphi_2, \ldots, \varphi_L]^T \).

CRLB of pitch estimate:

\[
[I(\xi)^{-1}]_{1,1} = \frac{12}{N(N^2 - 1)} (d_L^T \Phi_{\Delta \Omega}^{-1} d_L)^{-1},
\]  

(21)

where \( I(\xi) = -E\left\{ \frac{\partial^2 \ln p(\Psi, \xi)}{\partial \xi \partial \xi^T} \right\} \) is the Fisher information matrix.
Experimental Results
Experiment 1 (synthetic signal)

A synthetic harmonic signal $\omega_0 = 0.15\pi$, $L = 5$ sinusoids with identical amplitudes. We estimated UFEs from $N = 60$ samples using the subspace orthogonality method based on the MUSIC method.

- Additive colored Gaussian noise, different SNR:
Experimental Results

Experiment 2 (synthetic signal)

- White Gaussian noise in SNR = 5 dB, different $\omega_0$: 

![Graph showing MSE vs. $\omega_0/2\pi$ for different pitch estimators.]

- Harmonic Signal Model
- Pitch Estimation
- The WLS Pitch Estimator
- Proposed Method
- Problem Formulation
- Filtering UFEs
- Cramér-Rao Lower Bound
- Experimental Results
- Conclusion
Experimental Results

Experiment 3 (real signal)

- A trumpet signal contaminated by colored Gaussian noise, \( f_s = 8.0 \text{ kHz}, \; N = 128 \)
- We estimated \( L \) using the subspace orthogonality method [M.G.Christensen et al. 2007]
Summery and Conclusion

- We estimated pitch from unconstrained frequency estimates (UFEs).
- We designed an optimal filter based on noise statistics which is estimated from UFEs.
- The optimal filter design is a general method that we can derive the WLS pitch estimator.
- Pitch estimation using the optimal filter is robust against different noise situations, e.g., colored noise.
- Pitch estimation using the optimal filter is robust against missing harmonics in some time-frames and outperforms the WLS pitch estimator.
Thank you!