ROBUST FUNDAMENTAL FREQUENCY ESTIMATION IN THE PRESENCE OF INHARMONICITIES

N. R. Butt, S. I. Adalbjörnsson*, S. D. Somasundaram†, and A. Jakobsson**

*Centre for Mathematical Sciences, Lund University, SE-22100 Lund, Sweden.
† General Sonar Studies Group, Thales Underwater Systems, Stockport, Cheshire, SK3 0XB, U.K.
email: {naveed, stefan, aj}@maths.lth.se; sdsomasundaram@hotmail.com

ABSTRACT

We develop a general robust fundamental frequency estimator that allows for non-parametric inharmonicities in the observed signal. To this end, we incorporate the recently developed multi-dimensional covariance fitting approach by allowing the Fourier vector corresponding to each perturbed harmonic to lie within a small uncertainty hypersphere centered around its strictly harmonic counterpart. Within these hyperspheres, we find the best perturbed vectors fitting the covariance of the observed data. The proposed approach provides the estimate of the fundamental frequency in two steps, and, unlike other recent methods, involves only a single 1-D search over a range of candidate fundamental frequencies. The proposed algorithm is numerically shown to outperform the current competitors under a variety of practical conditions, including various degrees of inharmonicity and different levels of noise.

Index Terms— Fundamental frequency, inharmonicity, robust estimator, multi-dimensional covariance fitting.

1. INTRODUCTION

The estimation of the fundamental frequency, or *pitch*, of a set of harmonically related sinusoids is an integral part of many signal processing algorithms. While these algorithms most commonly find application in speech and audio signal processing, they can, in principle, be applied to harmonically related signals appearing in other fields, such as electrocardiography (ECG) [1]. Most developed estimators assume that the harmonics are exact integer multiples of the fundamental frequency (see, e.g., [1–3] and references therein). However, this is not always the case, and the deviation of the higher frequencies from exact integer multiples of the fundamental frequency, a phenomenon called *inharmonicity*, is often observed in real-world signals. For instance, it is well known that inharmonicity arises in piano tones due to the stiffness in the piano strings [4]. Inharmonicity has also been considered in the modelling and coding of speech signals, and several different models of inharmonicity have been developed [5,6], as, if not properly compensated for, the frequency deviations will lead to poor amplitude and pitch estimates [7]. To alleviate this problem, several *robust* fundamental frequency estimation algorithms have been proposed in the recent literature, allowing for inharmonicity in the observed signal. Most of these algorithms consider the scenario of stiff-stringed instruments where deviations from exact integer multiples of the fundamental frequency depend functionally on a single unknown stiffness parameter [8–11]. However, as discussed in [1, 7], and also elaborated upon below, a more general model that allows for random perturbations in the harmonics would lead to an estimator that covers a wider range of problems. Existing solutions, such as the maximum a posteriori (MAP) and subspace estimators presented in [1, 7], suffer from requiring exhaustive grid searches, such that the estimates are formed based on searches close to the expected unperturbed harmonics. Clearly, such combinatorial grid search approaches would increasingly become computationally inefficient with increasing number of harmonics, or for signals containing multiple sources. The main objective of this work is to develop a general robust fundamental frequency estimator that does not require searches over individual perturbed harmonics. In this regard, we incorporate the recently developed multi-dimensional covariance-fitting (MDCF) approach from the beamforming literature [12] into the robust pitch estimation problem by allowing the Fourier vector corresponding to each perturbed harmonic to lie within a small uncertainty hypersphere centered around its strictly harmonic counterpart. Within these hyperspheres, we find the best perturbed vectors fitting the covariance of the observed data. The proposed approach is more general than other recent robust methods such as [8–11] that deal only with simple parametric inharmonicity of the form in [4], and it avoids the exhaustive search approach of [1, 7]. Finally, we note that the proposed approach is different from several other robust pitch estimators [13–16] that are robust to different kinds of noise or to missing data. In contrast, our work focuses on robustness to inharmonicity.

This work was supported in part by the Swedish Research Council and Carl Trygger's Foundation, Sweden.

2. SIGNAL MODEL AND OTHER ESTIMATORS

Consider a harmonic signal with the fundamental frequency $\omega_0 > 0$, corrupted by an additive noise [1]

$$x(n) = \sum_{l=1}^{L} \alpha_l e^{in\omega_l} + e(n) \tag{1}$$

where $n = 0, \ldots, N-1$, $\alpha_l = |\alpha_l| e^{i\angle\alpha_l}$ denotes the complex amplitude of the lth harmonic, and $e(n)$ is a zero-mean white complex circularly symmetric Gaussian noise process with unknown variance σ_e^2. The harmonic frequencies, ω_l, are often formed as $\omega_l = \omega_0 l$, where ω_0 denotes the pitch frequency, but may, alternatively, using a model capturing the string stiffness also be modelled as

$$\omega_l(\omega_0, B) = l\omega_0\sqrt{1 + l^2 B} \tag{2}$$

where $B \ll 1$ is an unknown positive stiffness parameter. The main problem with such parametric models is that they are instrument dependent and one may have to consider many such models to develop an estimator that can be applicable to a wide range of pitch estimations problems. Additionally, in many audio signal processing problems, the inharmonicities may not be so well-behaved. To avoid such limitations, we will here consider the more general model used in [1], extending (1) to allow small independent perturbations in the harmonics, such that

$$x(n) = \sum_{l=1}^{L} \alpha_l e^{i\omega_l(\omega_0, \Delta_l)n} + e(n) \tag{3}$$

where $\omega_l(\omega_0, \Delta_l) = \omega_0 l + \Delta_l$, with Δ_l representing a perturbation of the l-th harmonic. We assume that the order L is known, or has been estimated using one of the order estimation algorithms available in literature [1]. We shall further assume, without loss of generality, that the perturbations are normally distributed zero-mean random variables with unknown but small variances, $\sigma_{\Delta_l}^2$. Among the pitch estimation algorithms available in literature, the maximum likelihood (ML) estimator offers a very powerful tool for estimating the fundamental frequency of a perfectly harmonic signal. It is known to be computationally efficient, and reduces to the optimal nonlinear least squares (NLS) estimator in case of white noise [1]. A robust version of the ML estimator, that allows for parametric inharmonicity of the form (2) has been presented in [1]. The algorithm is, however, computationally inefficient as it requires a 2-D search over ω_0 and B. Two of the relatively recent approaches that cover the general inharmonicity model in (3) are the MAP method of [1] and the subspace-based method of [7]. The MAP approach estimates the fundamental frequency and the perturbations by maximizing the posterior likelihood of observing the measured data under an assumed prior on the distribution of the perturbations. The subspace-based method [7], on the other hand, exploits a MUSIC-like approach to estimate the perturbed frequencies. However, both methods form the estimates based on searches over the parameters $(\omega_0, \{\Delta_l\})$.

3. PROPOSED ROBUST COVARIANCE-FITTING PITCH ESTIMATOR

Let

$$\mathbf{x}(n) = \begin{bmatrix} x(n) & x(n-1) & \ldots & x(n-M+1) \end{bmatrix}^T \tag{4}$$

$$\mathbf{A}_\Delta = \begin{bmatrix} \mathbf{a}_M(\omega_0 + \Delta_1) & \ldots & \mathbf{a}_M(\omega_0 L + \Delta_L) \end{bmatrix} \tag{5}$$

where $(\cdot)^T$ denotes the transpose, for $M < N$, with

$$\mathbf{a}_M(\omega) = \begin{bmatrix} 1 & e^{-i\omega} & \ldots & e^{-i\omega(M-1)} \end{bmatrix}^T \tag{6}$$

Note that $\mathbf{A}_\Delta$ is full-rank if $\omega_0 l + \Delta_l \neq \omega_0 m + \Delta_m, \forall\, l \neq m$. The covariance matrix of (3) can then be written as

$$\mathbf{R} = \mathrm{E}\{\mathbf{x}(n)\mathbf{x}^*(n)\} = \mathbf{A}_\Delta \mathbf{P} \mathbf{A}_\Delta^* + \sigma_e^2 \mathbf{I} \tag{7}$$

where $(\cdot)^*$ represents the Hermitian transpose, and

$$\mathbf{P} = \mathrm{diag}\{\begin{bmatrix} |\alpha_1|^2 & \ldots & |\alpha_L|^2 \end{bmatrix}\} \tag{8}$$

In order to utilize the powerful optimal filtering methods discussed in [1] without resorting to searches over the perturbations $\{\Delta_l\}$, we here propose to allow each perturbed Fourier vector $\mathbf{a}_M(\omega_0 l + \Delta_l)$ to lie within a small uncertainty hypersphere centered around its strictly harmonic counterpart $\mathbf{a}_M(\omega_0 l)$. Defining the nominal Fourier matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_M(\omega_0) & \ldots & \mathbf{a}_M(\omega_0 L) \end{bmatrix} \tag{9}$$

the set of constraints on the L Fourier vectors may be written compactly as

$$\|(\mathbf{A}_\Delta - \mathbf{A})\mathbf{e}_l\|_2 \leq \epsilon_l, \quad l = 1, \ldots, L \tag{10}$$

where ϵ_l is a user parameter reflecting on the expected level of inharmonicity, and where $\mathbf{e}_l$ is the l-th column vector of an $L \times L$ identity matrix. Further, noting from (7) that any selection of $\{\Delta_l\}$, $\boldsymbol{\alpha}$, and σ_e^2 must satisfy

$$\mathbf{A}_\Delta \mathbf{P} \mathbf{A}_\Delta^* + \sigma_e^2 \mathbf{I} \preceq \hat{\mathbf{R}} \tag{11}$$

where $\mathbf{A} \preceq \mathbf{B}$ denotes that $\mathbf{B} - \mathbf{A}$ is positive semidefinite, and where $\hat{\mathbf{R}} = \frac{1}{N-M+1} \sum_{n=0}^{N-M} \mathbf{x}(n)\mathbf{x}^*(n)$ is the sample covariance matrix, one may estimate the perturbed Fourier vectors as the ones that maximally explain the total observed signal power, $\log\det(\mathbf{R})$, while satisfying the constraints (10) and (11). To this end, we employ the MDCF concept [12], to recast the problem as

$$\begin{aligned} \max_{\mathbf{A}_\Delta, \mathbf{P}, \sigma_e^2 \geq 0} \quad & \log\det(\mathbf{A}_\Delta \mathbf{P} \mathbf{A}_\Delta^* + \sigma_e^2 \mathbf{I}) \\ \text{s.t.} \quad & \mathbf{A}_\Delta \mathbf{P} \mathbf{A}_\Delta^* + \sigma_e^2 \mathbf{I} \preceq \hat{\mathbf{R}} \\ & \|(\mathbf{A}_\Delta - \mathbf{A})\mathbf{e}_l\|_2 \leq \epsilon_l, \;\; l = 1, \ldots, L \\ & \mathbf{P} = \mathbf{P} \odot \mathbf{I}_L \succeq \mathbf{0} \end{aligned} \tag{12}$$

where $\odot$ is the element-wise matrix product, and the last constraint ensures that, in accordance with the definition (8), $\mathbf{P}$ is positive semidefinite and diagonal. As shown in [12], (12) may not be amenable to a standard numerical solution, and one may instead use semidefinite programming (SDP) to solve a local convex approximation of (12) as

$$\begin{aligned} \max_{\breve{\mathbf{A}}_\Delta, \sigma_e^2 \geq 0} \quad & 2\mathcal{R}\left\{ \text{tr}\{\breve{\mathbf{A}}^* \mathbf{R}_0^{-1} \breve{\mathbf{A}}_\Delta\} \right\} + \text{tr}\{\mathbf{R}_0^{-1}\}\sigma_e^2 \qquad (13) \\ \text{s.t.} \quad & \breve{\mathbf{A}}_\Delta \breve{\mathbf{A}}_\Delta^* + \sigma_e^2 \mathbf{I} \preceq \hat{\mathbf{R}} \\ & \mathcal{R}\{(\mathbf{A}\mathbf{e}_l)^* \breve{\mathbf{A}}_\Delta \mathbf{e}_l\} \geq \nu_l \|\breve{\mathbf{A}}\mathbf{e}_l\|_2; \ \ l = 1, \ldots, L \\ & \mathcal{I}\{(\mathbf{A}\mathbf{e}_l)^* \breve{\mathbf{A}}_\Delta \mathbf{e}_l\} = 0; \ \ l = 1, \ldots, L \end{aligned}$$

where $\mathcal{R}\{\cdot\}$ and $\mathcal{I}\{\cdot\}$ denote the real and imaginary parts of a complex number, respectively, $\nu_l = \sqrt{\|\mathbf{A}\mathbf{e}_l\|_2^2 - \epsilon_l^2}$, and $\mathbf{R}_0 = \breve{\mathbf{A}}\breve{\mathbf{A}}^* + \sigma_0^2 \mathbf{I}$, with

$$\breve{\mathbf{A}} = \mathbf{A}\mathbf{P}_0^{\frac{1}{2}} \qquad (14)$$

where $\mathbf{P}_0$ denotes an initial estimate of $\mathbf{P}$ obtained through any suitable spectral estimator, and σ_0^2 is formed by averaging the $M - L$ smallest eigenvalues of $\hat{\mathbf{R}}$. There are several reasons why this formulation cannot be directly used to estimate the perturbed frequencies $\{\omega_l\}$. Firstly, as can be seen from (6), the true Fourier vectors must satisfy $\overline{\mathbf{a}}_M(\omega)e^{-i\omega} = \underline{\mathbf{a}}_M(\omega)$, where $\overline{\mathbf{a}}_M(\omega)$ and $\underline{\mathbf{a}}_M(\omega)$ are formed by taking, respectively, the first $M-1$ and the last $M-1$ elements of the vector $\mathbf{a}_M(\omega)$. However, the formulation in (13) imposes no such constraint on the structure of $\breve{\mathbf{A}}_\Delta$. Secondly, we note that, by virtue of (14), an estimated $\breve{\mathbf{A}}_\Delta$ would include estimates of the amplitudes of the harmonics. Thus, the cost function of (13) is not suitable for a grid search over the fundamental frequency as it may wrongly compensate for the frequency perturbations by adjusting the estimates of the amplitudes and the noise variance σ_e^2. To address these issues for the robust pitch estimation problem, we propose the following two-step approach that can be applied over a very coarse grid of fundamental frequencies. We term the proposed two-step approach the robust covariance-fitting pitch (RCP) estimator.

3.1. Step one: coarse estimates

The main objective of the first step is to obtain an initial estimate of the perturbed matrix $\mathbf{A}_\Delta$. This estimate will then be used as the *assumed* matrix in the second step, and is formed using a single 1-D grid search over a range of fundamental frequencies. It is worth noting both that the search grid can be chosen to be rather coarse, and that the estimate may be formed without any search over the individual perturbations. The estimate is formed as:

(a). Form a grid of appropriate size, say K, over the expected range of fundamental frequencies, and choose a frequency point from the grid, say ω_0^k, and, assuming this to be the fundamental frequency, form the matrix $\mathbf{A}$ using (9), and $\mathbf{P}_0$ by computing the periodogram estimates at ω_0^k and its perfect harmonics. Using the evaluated $\mathbf{A}$ and $\mathbf{P}_0$, solve the SDP in (13) to get an initial estimate of $\breve{\mathbf{A}}_\Delta$.

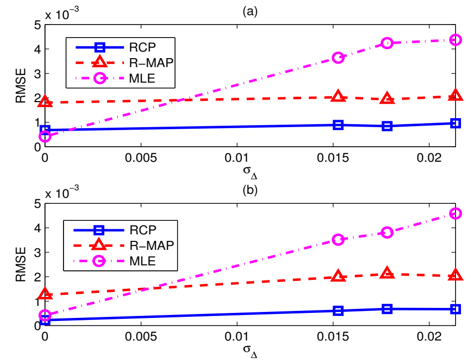

Fig. 1. RMSE of the fundamental frequency estimates against the level of inharmonicity, for $\omega_0 = 0.2137$, at SNR levels of (a) 5 dB and (b) 30 dB.

(b). In line with the discussion under (14), the perturbed harmonics are extracted from the estimated $\breve{\mathbf{A}}_\Delta$ by imposing the suggested structural constraint on its columns. More specifically, denoting the l-th column of the estimated $\breve{\mathbf{A}}_\Delta$ as $\mathbf{b}_l$, and noting that, to be a true Fourier vector for the l-th harmonic, it must satisfy $\overline{\mathbf{b}}_l\gamma_l = \underline{\mathbf{b}}_l$, where $\gamma_l = e^{-i\omega_l}$, and where $\overline{\mathbf{b}}_l$ and $\underline{\mathbf{b}}_l$ are defined similar to $\overline{\mathbf{a}}_M(\omega)$ and $\underline{\mathbf{a}}_M(\omega)$, respectively, form an estimate of the l-th harmonic frequency as $\hat{\omega}_l = -\mathcal{I}\{\ln(\hat{\gamma}_l)\}$, with

$$\hat{\gamma}_l = \frac{\overline{\mathbf{b}}_l^* \underline{\mathbf{b}}_l}{\|\overline{\mathbf{b}}_l\|_2^2} \qquad (15)$$

(c). Form an improved estimate of $\mathbf{A}_\Delta$, say $\hat{\mathbf{A}}_\Delta$, by substituting the estimates $\{\hat{\omega}_l\}$ in (5). With the estimate $\hat{\mathbf{A}}_\Delta$ now available, the problem reduces to a standard pitch estimation problem. Therefore, we propose to utilize the cost function

$$g_k \triangleq \text{tr}\left[\left(\hat{\mathbf{A}}_\Delta^* \hat{\mathbf{R}}^{-1} \hat{\mathbf{A}}_\Delta\right)^{-1}\right] \qquad (16)$$

which represents the total output power of a set of L Capon filters, and is maximized at the true perturbed frequencies (for details, see, e.g., [1, 17]).

(d). Repeat (a)-(c) for the K points in the grid, and choose $\{\hat{\omega}_l^{max}\}$ as the L estimates where $\{g_k\}_{k=1}^K$ is maximized.

3.2. Step two: refined estimates

While it is possible to use $\{\hat{\omega}_l^{max}\}$, obtained in the previous step, one may refine the estimates of the perturbed frequencies further by solving (13) with the following improved initializations. Firstly, in place of $\mathbf{A}$, $\hat{\mathbf{A}}_\Delta^{max}$ is used as the *assumed* Fourier matrix, where $\hat{\mathbf{A}}_\Delta^{max}$ is formed by substituting

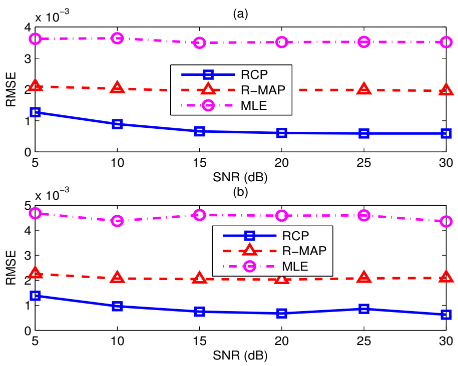

Fig. 2. RMSE of the fundamental frequency estimates against SNR, for $\omega_0 = 0.2137$ using σ_Δ equal to (a) $\omega_0/10$, and (b) $\omega_0/10$

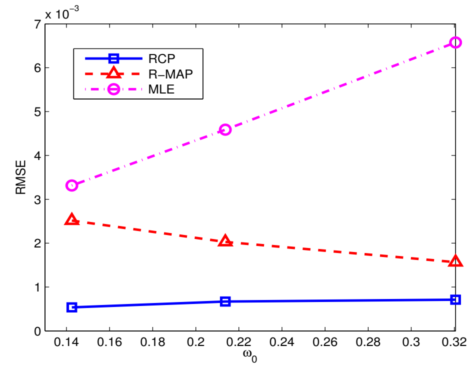

Fig. 3. RMSE at $\omega_0 = 0.1425, 0.2137, 0.3206$, for SNR = 30 dB, using $\sigma_\Delta = \omega_0/10$.

$\{\hat{\omega}_l^{max}\}$ in (5). Secondly, to give better initial estimates of the amplitudes of the harmonics, $\mathbf{P}_0$ should be formed by computing the periodogram amplitudes at $\{\hat{\omega}_l^{max}\}$. These two modifications together assure a better initialization for the SDP problem in (13), leading therefore to more accurate frequency estimates, which can be formed as in (b) of the first step.

3.3. Selection of ϵ_l

Noting that the left side of (10) may be written as the summation, $\sum_{m=1}^{M}\sqrt{2(1-\cos(\Delta_l m))}$, one may give a rough range for the selection of ϵ_l, such that it does not violate (10), as $0 \le \epsilon_l \le 2\sqrt{M}$. Practical experience shows that in order to restrict Δ_l to be very small (which is typically the case), one should choose $\epsilon_l \le \sqrt{M}/3$. Secondly, one should use a smaller ϵ_l in the second step of RCP as compared to the value used in the first step. This is because $\hat{\mathbf{A}}_\Delta^{max}$ is expected to be closer to the true value of $\mathbf{A}_\Delta$, as compared to $\mathbf{A}$ (which is used as the *assumed* matrix in the first step).

4. SIMULATIONS AND RESULTS

We proceed to numerically evaluate the performance of the proposed RCP estimator, comparing to the MLE [1] and the robust MAP (R-MAP) [7] estimators. The results are obtained through a number of experiments based on Monte Carlo simulations using synthetic signals. In each case, the synthetic signal was generated using (3), with $L = 4$ harmonics having unit amplitudes and uniformly distributed phases that are randomized in each Monte Carlo run. The experiments were repeated for several different fundamental frequencies, and for five different signal-to-noise ratio (SNR) levels from $5-30$ dB, where the SNR is defined as $10\log_{10}(\mathrm{tr}(\mathbf{P})/\sigma_e^2)$. All algorithms were tested at different levels of inharmonicity by increasing the standard deviation of the perturbations, σ_Δ, from 0 to $\omega_0/10$, where a variance of 0 indicates a perfectly harmonic signal. A total of $J = 150$ Monte Carlo simulations were used in each experiment to evaluate the root mean square error (RMSE), defined for the fundamental frequency estimates as, $RMSE = \sqrt{\frac{1}{J}\sum_{j=1}^{J}(\hat{\omega}_{0,j}-\omega_0)^2}$, where ω_0 and $\hat{\omega}_{0,j}$ represent the true fundamental frequency and the estimated fundamental frequency in the j-th Monte Carlo run, respectively. A data length of $N = 200$ samples was used, while the sub-vector length for RCP was set to $M = 50$, which is in accordance with the limit, $M \le N/2$, suggested in filtering literature (see, e.g., [1] and [17]). Typical results, comparing the proposed RCP estimator to the standard MLE and the R-MAP estimators, are shown in Figures 1-3. Following the guidelines in Section 3.3, all the results were obtained with the uncertainty parameter ϵ_l set to 4 for the first step and to 2 for the second step of RCP. The fundamental frequency search grid for MLE and R-MAP consisted of 300 equally-spaced points in the range [0.05, 0.5], whereas for the proposed RCP estimator, the grid consisted of only 30 equally-spaced points in the same range. Figure 1 shows the RMSE of the fundamental frequency estimates against the level of inharmonicity for $\omega_0 = 0.2137$ at SNR levels of 5 and 30 dB. As is clear from the figures, the proposed RCP estimator performs better at both low SNR and high SNR levels. As expected, the MLE method, not allowing for inharmonicity, suffers heavily with increase in inharmonicity. Figure 2 shows the RMSE against SNR for $\omega_0 = 0.2137$ for σ_Δ equal to $\omega_0/14$ and $\omega_0/10$. We see that while the performance of all the estimators degrades slightly at lower SNRs, the RCP estimator provides more accurate estimates at all levels. Finally, Figure 3 shows the RMSE at three different fundamental frequencies $\omega_0 = 0.1425, 0.2137$, and 0.3206, at SNR = 30 dB, and using $\sigma_\Delta = \omega_0/10$. We remark that the increase in the RMSE of MLE at the higher frequencies is because of a higher simulated inharmonicity at these frequencies. While both R-MAP and RCP show robustness to the inharmonicity, the proposed approach clearly provides more accurate estimates.

5. REFERENCES

[1] M. Christensen and A. Jakobsson, *Multi-Pitch Estimation*, Morgan & Claypool, 2009.

[2] H. Kameoka, *Statistical Approach to Multipitch Analysis*, Ph.D. thesis, University of Tokyo, 2007.

[3] W. Hess, "Pitch and voicing determination," *Advances in Speech Signal Processing*, pp. 3–48, 1992.

[4] H. Fletcher, "Normal vibration frequencies of stiff piano string," *Journal of the Acoustical Society of America*, vol. 36, no. 1, 1962.

[5] T. D. Rossing, *The Science of Sound*, Addison-Wesley Publishing Co., 1990.

[6] E. B. George and M. J. T. Smith, "Speech analysis/synthesis and modification using an analysis-by-synthesis/overlap-add sinusoidal model," *IEEE Transactions on Speech and Audio Processing*, vol. 5, no. 5, pp. 389 –406, Sep 1997.

[7] M. G. Christensen, P. Vera-Candeas, S. D. Somasundaram, and A. Jakobsson, "Robust Subspace-based Fundamental Frequency Estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Las Vegas, March 30-April 4, 2008.

[8] I. Barbancho, L. J. Tardon, S. Sammartino, and A. M. Barbancho, "Inharmonicity-based method for the automatic generation of guitar tablature," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 6, pp. 1857 –1868, Aug. 2012.

[9] E. Benetos and S. Dixon, "Joint multi-pitch detection using harmonic envelope estimation for polyphonic music transcription," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 6, pp. 1111 –1123, Oct. 2011.

[10] J. X. Zhang, M. G. Christensen, S. H. Jensen, and M. Moonen, "A robust and computationally efficient subspace-based fundamental frequency estimator," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 3, pp. 487 –497, March 2010.

[11] V. Emiya, R. Badeau, and B. David, "Multipitch estimation of piano sounds using a new probabilistic spectral smoothness principle," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 18, no. 6, pp. 1643–1654, August 2010.

[12] M. Rubsamen and A. B. Gershman, "Robust adaptive beamforming using multidimensional covariance fitting," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 740 –753, Feb. 2012.

[13] F. Huang and T. Lee, "Pitch estimation in noisy speech using accumulated peak spectrum and sparse estimation technique," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 1, pp. 99 –109, Jan. 2013.

[14] S. Gonzalez and M. Brookes, "A pitch estimation filter robust to high levels of noise (PEFAC)," in *19th European Signal Processing Conference (EUSIPCO)*, 2011.

[15] J. A. Morales-Cordovilla, Ning Ma, V. Sanchez, J. L. Carmona, A. M. Peinado, and J. Barker, "A pitch based noise estimation technique for robust speech recognition with missing data," in *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, May 2011, pp. 4808 –4811.

[16] J. O. Hong and P. J. Wolfe, "Robust and efficient pitch estimation using an iterative ARMA technique," in *INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association*, 2010.

[17] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, Upper Saddle River, N.J., 2005.