Signal Processing 135 (2017) 188–197

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Fast fundamental frequency estimation: Making a statistically efficient estimator computationally efficient

Jesper Kjær Nielsen[a,b,*], Tobias Lindstrøm Jensen[c,1], Jesper Rindom Jensen[a,2], Mads Græsbøll Christensen[a], Søren Holdt Jensen[c]

[a] Aalborg University, Audio Analysis Lab, AD:MT, Rendsburggade 14, DK-9000 Aalborg, Denmark
[b] Bang & Olufsen A/S, Acoustic Research, Peter Bangs Vej 15, DK-7600 Struer, Denmark
[c] Aalborg University, Dept. Electronic Systems, Fredrik Bajers Vej 7B, DK-9220 Aalborg, Denmark

ARTICLE INFO

Keywords:
Fundamental frequency estimation
Toeplitz
Hankel
Fast algorithms
Pitch

ABSTRACT

Modelling signals as being periodic is common in many applications. Such periodic signals can be represented by a weighted sum of sinusoids with frequencies being an integer multiple of the fundamental frequency. Due to its widespread use, numerous methods have been proposed to estimate the fundamental frequency, and the maximum likelihood (ML) estimator is the most accurate estimator in statistical terms. When the noise is assumed to be white and Gaussian, the ML estimator is identical to the non-linear least squares (NLS) estimator. Despite being optimal in a statistical sense, the NLS estimator has a high computational complexity. In this paper, we propose an algorithm for lowering this complexity significantly by showing that the NLS estimator can be computed efficiently by solving two Toeplitz-plus-Hankel systems of equations and by exploiting the recursive-in-order matrix structures of these systems. Specifically, the proposed algorithm reduces the time complexity to the same order as that of the popular harmonic summation method which is an approximate NLS estimator. The performance of the proposed algorithm is assessed via Monte Carlo and timing studies. These show that the proposed algorithm speeds up the evaluation of the NLS estimator by a factor of 50–100 for typical scenarios.

1. Introduction

Periodic signals are encountered in many real-world applications such as music processing [1,2], speech processing [3,4], sonar [5], order analysis [6], and electrocardiography (ECG) [7]. Such signals can be modelled as a weighted sum of sinusoids whose frequencies are integer multiples of a common fundamental frequency which in audio and speech applications is often referred to as the pitch [2]. Therefore, an important and fundamental problem in the above mentioned applications is to estimate this fundamental frequency from an observed data set. Multiple estimation methods have been proposed in the scientific literature ranging from simple correlation-based methods [8] to parametric methods [2]. Although the parametric methods in general are much more accurate than the correlation-based methods, they suffer from a high computational complexity. Consequently, the correlation-based methods remain very popular despite that they require all sorts of heuristic post-processing to give a satisfactory performance [9–13]. Since many applications require real-time processing, the computational complexity of the parametric methods must be reduced to make them a viable alternative to the correlation-based methods, and the contribution presented in this paper should be seen in this context.

The main difficulty in estimating the fundamental frequency is that a non-linear optimisation problem has to be solved. No closed-form solution is available, and we, therefore, have to search for the global optimiser of an often very oscillatory cost function such as the two examples shown in Fig. 1. This search for the optimiser is often performed using the following steps.

1. The cost function is evaluated on a grid and one or several candidate optimisers are selected on this grid. Often, the grid is uniform since a part of the cost function can then be evaluated efficiently using an FFT algorithm.
2. The candidate optimisers are refined using, e.g., interpolation

* Corresponding author at: Aalborg University, Dept. Electronic Systems, Fredrik Bajers Vej 7B, DK-9220 Aalborg, Denmark.
E-mail addresses: jkn@create.aau.dk (J.K. Nielsen), tlj@es.aau.dk (T.L. Jensen), jrj@create.aau.dk (J.R. Jensen), mgc@create.aau.dk (M.G. Christensen), shj@es.aau.dk (S.H. Jensen).
[1] The work of T.L. Jensen was supported by the Independent Research Council for Technology and Production 4005-00122.
[2] The work of J.R. Jensen was supported by the Independent Research Council for Technology and Production 1337-00084.

http://dx.doi.org/10.1016/j.sigpro.2017.01.011
Received 4 October 2016; Received in revised form 5 January 2017; Accepted 10 January 2017
Available online 16 January 2017
0165-1684/ © 2017 Elsevier B.V. All rights reserved.

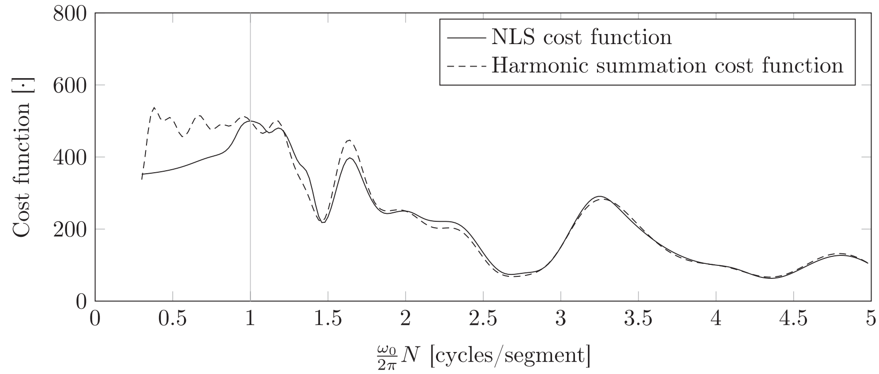

Fig. 1. Example of the exact NLS and harmonic summation cost functions. N=100, L=10, $\omega_0 = 2\pi/N$, and constant amplitude $\sqrt{a_i^2 + b_i^2} = 1,\ i = 1,\ldots,L$.

methods, line searches, or derivative-based methods.

3. For the parametric methods, model order estimation has to be performed when the model is unknown to reduce the risk of estimating an integer multiple or division of the true fundamental frequency. Often, this problem is also referred to as pitch halving/doubling or as octave errors. Estimating an unknown model often means that we have to repeat the first two steps above for every candidate model, thus increasing the computational complexity significantly.

In the correlation-based methods, the cost function is the autocorrelation function (or some variation thereof) which can typically be computed very efficiently. Adding to this, the correlation-based methods are not model based so it is not necessary to do model comparison to determine the number of harmonic components in the signal. From a computational perspective, the correlation-based methods are, therefore, very attractive. Unfortunately, they have a suboptimal estimation performance, are not very robust to noise (see, e.g., Fig. 3), and do not work for low fundamental frequencies [14]. Here, a low frequency means the number of cycles in a segment of data rather than the frequency measured in, e.g., Hz or radians/s, and this also explains the somewhat non-standard value on the x-axis in Fig. 1. The poor performance for low fundamental frequencies is hardly surprising since fewer and fewer data points are used in the computation of the autocorrelation function as the candidate fundamental frequency decreases. As exemplified by the very popular YIN method [11], this is often solved by using data from the previous data segment, but this trick corresponds to doubling the segment length and using a 50% overlap. Thus, the correlation-based methods cannot provide the same time-frequency resolution as those parametric methods which also work for a low fundamental frequency.

The poor noise robustness and time-frequency resolution seem to be fundamental flaws of the correlation-based methods and the main reason for considering alternative estimators based on a parametric model. Unfortunately, the evaluation of the cost function in the parametric methods is often quite numerically costly since they can involve eigenvalue decompositions of covariance matrices [15,16] or matrix inversions [17,18]. A notable exception, though, is the harmonic summation method (HS) [19,20] which is an approximate non-linear least-squares (NLS) estimator and can be implemented efficiently using a single FFT [2]. The HS summation estimator is statistically efficient and robust to noise and is, therefore, a very attractive alternative to the correlation-based methods. Unfortunately, the HS method also fails for low fundamental frequencies and, therefore, suffers from a suboptimal time-frequency resolution. This is in contrast to the NLS estimator [17] which has a much better performance for low fundamental frequencies and, consequently, a better time-frequency resolution [21], although at a much higher computational complexity. In Fig. 1, the differences and similarities between the HS and NLS cost functions are illustrated. When more than approximately two periods or more are assumed to be in a segment, the two cost functions are nearly the same whereas they become more and more distinct for a decreasing fundamental frequency. When the fundamental frequency is low and the data are real-valued, an error is also made if estimators based on the complex-valued harmonic model are used instead of estimators based on the real-valued harmonic model. The error is introduced when the real-valued signal is converted into an analytic signal (complex-valued) by use of the Hilbert transform which ignores the interaction effects that occur between positive and negative frequency components when the fundamental frequency is low. As demonstrated in [21], a much better estimation accuracy is obtained for low fundamental frequencies if the NLS estimators for the real-valued signal model is used instead of the NLS estimator for the complex-valued one.

Although the NLS estimator has been known for at least 25 yr and has some very attractive properties, which have been investigated thoroughly in, e.g., [17,22,23,2,21,24], we are not aware of any fast implementations of it. In fact, we believe that one of the main reasons for the popularity of the HS method is that it shares many desirable properties with the NLS method, but is much more computationally efficient. In this paper, however, we show that the evaluation of the NLS cost function can be reduced to the same order of time complexity as that of the HS method. More precisely, we show that we can reduce the cost of evaluating the NLS cost function on an F-point grid for all candidate model orders $l = 1,\ldots,L$ from $O(F\log F) + O(FL^3)$ to $O(F\log F) + O(FL)$ which is the same as that for the HS method. In addition to making each cost function evaluation as cheap as possible, we also derive how the number of grid points F depends on the segment length N and the maximum candidate model order L. This result is important to ensure that we neither over- nor undersample the cost function and can also be used to make the HS method faster.

The rest of this paper is organised as follows. In Section 2, we first introduce the signal model, the ML estimator, the NLS cost function, and the HS method. We also show how they are related to each other. Then, the standard way of computing the NLS and HS cost functions are described in Section 3. To speed up the computation of the NLS cost function, we describe how the number of cost function evaluations are minimised in Section 4 and how each cost function evaluation can be made efficiently in Section 5. Finally, we investigate how fine the cost function grid should be and quantify the computational savings using Monte Carlo simulations in Section 6.

2. Signal model and NLS cost function

The real-valued signal model for a uniformly sampled and periodic signal in additive noise $e(n)$ is given by

$$x(n) = \sum_{i=1}^{l} [a_i \cos(i\omega_0 n) - b_i \sin(i\omega_0 n)] + e(n) \tag{1}$$

where a_i and b_i are the linear weights of the i'th harmonic component and ω_0 is the fundamental frequency in radians per sample. If an N-

Download English Version:

https://daneshyari.com/en/article/4977576

Download Persian Version:

https://daneshyari.com/article/4977576

Daneshyari.com