KALMAN FILTER ANALYSIS FOR QUASI-PERIODIC SIGNALS

Kazuki Nishi

University of Electro-Communications
1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, JAPAN (E-mail: *nishi@ice.uec.ac.jp*)

ABSTRACT

An Optimal filter for extracting quasi-periodic signals such as a voiced speech or instrumental sound from the noise-corrupted observation are proposed. They are derived through the Kalman-Bucy filter analysis in which the dynamics of amplitude and pitch fluctuations are described through Itô stochastic differential equations. The Laplace analysis to the filter equation leads to three types of comb filters, i.e., constant-BW (-bandwidth) type, constant-Q type and those mixture type that have robustness to the amplitude fluctuation, pitch fluctuation and both of them, respectively. All-pole digital filters can be also realized for real-time processing. Examples of filter design are presented, and the performance of harmonics extraction is examined by comparison between the constant-BW type and constant-Q type.

1. INTRODUCTION

Quasi-periodic signals (or time-varying harmonics) such as a voiced speech, instrumental sound or mechanical vibration play an essential role in our daily communication. It is, therefore, one of the essential topics in the signal processing field to extract such harmonics from the noise-corrupted observation. Much effort has been spent to solve this problem, but the effective method has not been established yet.

The comb filter is well known as the most typical method for harmonics extraction. It works so as to enhance the periodicity of harmonics by averaging over every point spacing at a pitch interval[1] and thus emphasize a discrete spectral structure in harmonics[2]. The Kalman filtering approach has been proposed for estimation of time-varying harmonic components. The state-space signal model allows the introduction of the time-varying phenomena[3], especially the effect of harmonic amplitude fluctuation has been investigated[4]. Even a slight indistinctness of the pitch information, however, causes a significant distortion to the filter output, because any conventional methods do not take into account the effect of the inevitable pitch estimation error and fluctuation. In our previous work[5], an optimal comb filter has been obtained by the least mean square estimation under the strict assumption that the pitch estimate includes the uniform bias error.

In this paper, we propose the generalized comb filters which have a robustness to both of amplitude and pitch fluctuations. They are derived by the Kalman-Bucy filter analysis in which those fluctuation dynamics are described through Itô stochastic differential equations. The Laplace analysis to the filter equation provides the noteworthy following results: (1)When only the amplitude fluctuation exists, the constant-BW (-bandwidth) comb filter of which every passband has a constant bandwidth on each harmonic frequency is derived. (2)When only the pitch fluctuation exists, the constant-Q comb filter of which each bandwidth is homogeneously dilated in proportion to the harmonic frequency is obtained. (3)When both amplitude and pitch fluctuations exist, the filter works as the constant-BW type in lower frequency, the constant-Q type in higher frequency, and those mixture appears in the intermediate frequency. Furthermore, the digital filter formulation can be also derived through the s-z transform and the discretization of the filter equation.

2. KALMAN-BUCY FILTER EQUATION

The basic behavior of quasi-periodic signals can be described by the following Itô stochastic differential equation:

$$d\boldsymbol{x}(t) = \Omega(t)\boldsymbol{x}(t)dt + dB(t)\boldsymbol{x}(t), \tag{1}$$

where

$$\boldsymbol{x}(t) = [x_1(t), x_2(t), \cdots, x_N(t)]^T,$$

$$\Omega(t) = \mathrm{diag}\{-\frac{1}{2}(\sigma_{a1}^2(t) + \sigma_\omega^2(t)) + j\omega_0(t), \cdots, -\frac{1}{2}(\sigma_{aN}^2(t) + N^2\sigma_\omega^2(t)) + jN\omega_0(t)\},$$

$$dB(t) = \mathrm{diag}\{\sigma_{a1}(t)d\beta_{a1}(t) + j\sigma_\omega(t)d\beta_\omega(t), \cdots, \sigma_{aN}(t)d\beta_{aN}(t) + jN\sigma_\omega(t)d\beta_\omega(t)\}.$$

$x_n(t)$ is each harmonic component of the quasi-periodic signal and $\omega_0(t)$ denotes a slowly time-varying pitch frequency estimated in advance. $\beta_{an}(t)$ and $\beta_{\omega}(t)$ mean normal Brownian motion processes to describe randomness of amplitude changes and pitch fluctuations with the variance $\sigma_{an}^2(t)$ and $\sigma_{\omega}^2(t)$, respectively.

The actually observed signal is given by summation of harmonic components and a zero-mean white Gaussian measurement noise $v(t)$ with variance $\sigma_v^2(t)$, i.e.,

$$y(t) = \mathbf{1}^T \boldsymbol{x}(t) + v(t), \quad (2)$$

where $\mathbf{1} = [1 \cdots 1]^T$, superscripts T means the matrix transpose. $v(t)$,$\Delta\omega(t)$ and $\{x_1(0), \cdots, x_N(0)\}$ are assumed to be independent each other.

Our goal is to compute an optimum estimate $\hat{\boldsymbol{x}}(t)$ to the original signal $\boldsymbol{x}$(t) from the given observation sequence $\{y(\tau) : 0 \le \tau \le t\}$. The solution can be obtained by a way of the Kalman-Bucy filter in which the random fluctuation term $dB(t)\boldsymbol{x}(t)$ is treated as a state-dependent noise[6]. It yields the following differential equation:

$$\frac{d\hat{\boldsymbol{x}}(t)}{dt} = \Omega(t)\hat{\boldsymbol{x}}(t) + \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}[y(t) - \mathbf{1}^T \hat{\boldsymbol{x}}(t)]. \quad (3)$$

$P(t)$ means a covariance error matrix:

$$P(t) = \mathrm{E}[(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t))(\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t))^*] \quad (4)$$

which satisfies the following Riccati equation:

$$\begin{aligned}\frac{dP(t)}{dt} &= \Omega(t)P(t) + P(t)\Omega^*(t) \\ &\quad - \frac{1}{\sigma_v^2(t)} P(t)\mathbf{1}\mathbf{1}^T P(t) + \Sigma(t), \quad (5)\end{aligned}$$

where

$$\begin{aligned}\Sigma(t) = \mathrm{diag}\{&\{\sigma_{a1}^2(t) + \sigma_{\omega}^2(t)\}\bar{x}_1^2(t), \\ &\cdots, \{\sigma_{aN}^2(t) + N^2\sigma_{\omega}^2(t)\}\bar{x}_N^2(t)\},\end{aligned}$$

$\bar{x}_n^2(t) = \mathrm{E}[|x_n(t)|^2]$ and superscripts $*$ means the complex conjugate transpose.

3. STEADY STATE SOLUTION

If every parameter's fluctuation is sufficiently small, the interaction between different harmonic components can be neglected and then Eq.(3) separates into independent scalar equations:

$$\begin{aligned}\frac{d\hat{x}_n(t)}{dt} &= [-\frac{1}{2}\sigma_n^2(t) + jn\omega_0(t)]\hat{x}_n(t) \\ &\quad + \frac{p_n(t)}{\sigma_v^2(t)}[y(t) - \hat{x}_n(t)] \quad (6)\end{aligned}$$

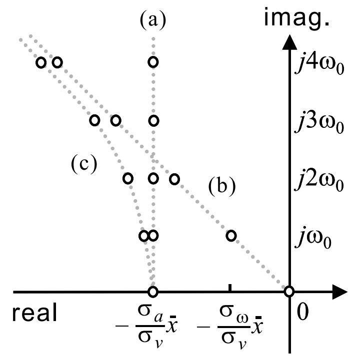

Figure 1: Pole allocation of the transfer function in the z-plane. (a)Constant-BW type. (b)Constant-Q type. (c)Composite type.

on each harmonic component, where $p_n(t)$ satisfies the scalar Riccati equation:

$$\frac{dp_n(t)}{dt} = \sigma_n^2(t)\bar{x}_n^2(t) - \sigma_n^2(t)p_n(t) - \frac{1}{\sigma_v^2(t)}p_n^2(t), \quad (7)$$

where

$$\sigma_n^2(t) = \sigma_{an}^2(t) + n^2\sigma_{\omega}^2(t).$$

When any parameter can be regarded as nearly constant values, i.e., $\omega_0(t) \to \omega_0, \bar{x}_n^2(t) \to \bar{x}_n^2, \sigma_v^2(t) \to \sigma_v^2$, $\sigma_{an}^2(t) \to \sigma_{an}^2, \sigma_{\omega}^2(t) \to \sigma_{\omega}^2$, Eq.(7) asymptotically converges to

$$p_n(t) \to \sqrt{\bar{x}_n^2\sigma_n^2\sigma_v^2 + \frac{1}{4}\sigma_n^4\sigma_v^4} - \frac{1}{2}\sigma_n^2\sigma_v^2. \quad (8)$$

In this steady state condition, Eq.(6) is simplified as

$$\begin{aligned}\frac{d\hat{x}_n(t)}{dt} &= \left[-\sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4} + jn\overline{\omega}_0\right]\hat{x}_n(t) \\ &\quad + \left[\sqrt{\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 + \frac{1}{4}\sigma_n^4} - \frac{1}{2}\sigma_n^2\right]y(t). \quad (9)\end{aligned}$$

For $\frac{\bar{x}_n^2}{\sigma_v^2}\sigma_n^2 \gg \frac{1}{4}\sigma_n^4$, the transfer function to estimate the harmonic component $\hat{x}_n(t)$ from the observation $y(t)$ can be approximately described as

$$H_n(s) = \frac{\hat{X}_n(s)}{Y(s)} \approx \frac{\frac{\bar{x}_n}{\sigma_v}\sigma_n}{s + \frac{\bar{x}_n}{\sigma_v}\sigma_n - jn\omega_0} \quad (10)$$

through the Laplace transform to Eq.(9), It is shown that the transfer function involves a band-pass characteristic of which center frequency is just located at each harmonic frequency $jn\omega_0$, i.e., a comb-like structure.

(a) For only amplitude fluctuation

When only the amplitude fluctuation exists uniformly in all harmonic components ($\sigma_{a1}^2 = \cdots = \sigma_{aN}^2 = \sigma_a^2$) but the pitch fluctuation is omitted ($\sigma_\omega^2 = 0$), every pole is allocated at equal spaces along the straight line parallel to the imaginary axis on the left side of s-plane (see Figure 1(a)). It means the transfer function consists of pass-bands with a constant bandwidth just located at each harmonic frequency. It is, therefore, called "constant-BW comb filter." The bandwidth is dilated in proportion to the spacing between each pole and imaginary axis, i.e., κ_a.

(b) For only pitch fluctuation

When only the pitch fluctuation exists but the amplitude fluctuation is omitted ($\sigma_{a1}^2 = \cdots = \sigma_{aN}^2 = 0$), every pole is allocated at equal spaces along the skew line across the origin (see Figure 1(b)). It means the bandwidth of pass-band just located at each harmonic frequency is homogeneously dilated in proportion to the harmonic frequency. It is so called "constant-Q comb filter." The more the gradient of pole allocation sharpens, i.e., the smaller κ_ω, the more the Q-factor enlarges.

(c) For both fluctuations

When both amplitude and pitch fluctuations exist, the transfer function shows the following tendency. Since the effect of amplitude fluctuation σ_a becomes dominant in small order n, the pole allocation is more close to the constant-BW type in lower frequency. Oppositely, the effect of pitch fluctuation σ_ω becomes dominant in large order n and thus the pole allocation is more close to the constant-Q type in higher frequency (see Figure 1(c)). It means, therefore, the filter characteristic gradually changes from the constant-BW type to the constant-Q type as the harmonic frequency varies from low to high.

In either case, the ideal periodic signal input never produce any distortion to the filter output because the filter gain on each harmonic frequency always holds $H_n(jn\omega_0) = 1$.

4. ALL-POLE DIGITAL COMB FILTERS

From Eq.(10), the digital filter formulation can be directly derived by the s-z transform using the impulse invariant method, thus

$$H_n(z) = \frac{\kappa_n T_s}{1 - e^{(-\kappa_n + jn\omega_0)T_s} z^{-1}}, \qquad (11)$$

where T_s means a sampling interval and $\kappa_n = \frac{\overline{x}_n}{\sigma_v}\sigma_n$. The overall filter architecture for harmonics extraction is given by $H(z) = \sum_{n=1}^{N} H_n(z)$. In the z-plane, each pole is allocated at equal angles corresponding to the harmonic frequency and pulled into more inside of the unit circle in higher harmonic frequency. This structure is indeed suitable for real-time processing.

Another description of the digital filter can be derived as follows. Using $\hat{x}_n(t - T_s)$ as the initial value, the analytical solution of Eq.(6) at the time t can be represented as

$$\hat{x}_n(t) = \phi(t, t - T_s)\hat{x}_n(t - T_s) + \int_{t-T_s}^{t} \phi(t, \tau)\frac{p_n(\tau)}{\sigma_v^2(\tau)} y(\tau) d\tau, \qquad (12)$$

where

$$\phi(t, \tau) = e^{\int_\tau^t (jn\omega_0(s) - \frac{p_n(s)}{\sigma_v^2(s)})ds}. \qquad (13)$$

Since any parameters can be regarded as varying smoothly for the assumption of quasi-periodicity and holding constant values for $t - T_s \le \tau \le t$, Eq.(12) can be rewritten as

$$\hat{x}_n(t) = e^{\{-\kappa_n(t) + jn\omega_0(t)\}T_s}\hat{x}_n(t - T_s) + \kappa_n(t) T_s y(t), \qquad (14)$$

where $\kappa_n(t) = p_n(t)/\sigma_v^2(t)$. This representation means the adaptive filter operation such that pole parameters can be adjusted to the time-varying phenomena. In the steady state condition, especially, the above coincides with Eq.(11).

5. EVALUATIONS

In the following simulations, we assume that the periodic signal is synthesized by cutting out one cycle part from a waveform of real voiced sound 'a' and connecting it periodically, and the time-varying harmonics is made artificially by imposing some amplitude and/or pitch fluctuations to the periodic signal. Our goal here is to extract the original time-varying harmonics from the contaminated version with white Gaussian noise by using the filters proposed above.

The first trial is performed on the harmonics including only the amplitude fluctuation, Kalman gains are determined as adjusted to the amplitude fluctuation. The extracted waveform seems to be considerably close to the original signal in the case of the constant-BW comb filter, but in the constant-Q type some noise remains in higher frequency. This is because the constant-Q type causes the noise component in higher frequency to be more absorbed for the wide bandwidth as seen in the spectrum distribution (see Figure 2). This result maintains an advantage of the constant-BW comb filter to the amplitude fluctuation.

The next one is performed on the harmonics including only the pitch fluctuation. Kalman gains are

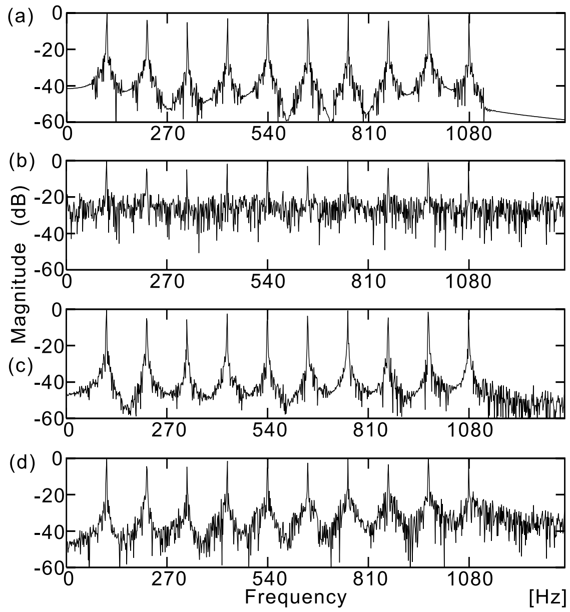

Figure 2: Spectral deformation caused by amplitude fluctuation. (a)Original signal. (b)Noise-corrupted signal. (c)Extracted result by the constant-BW comb filter. (d)Extracted result by the constant-Q comb filter.

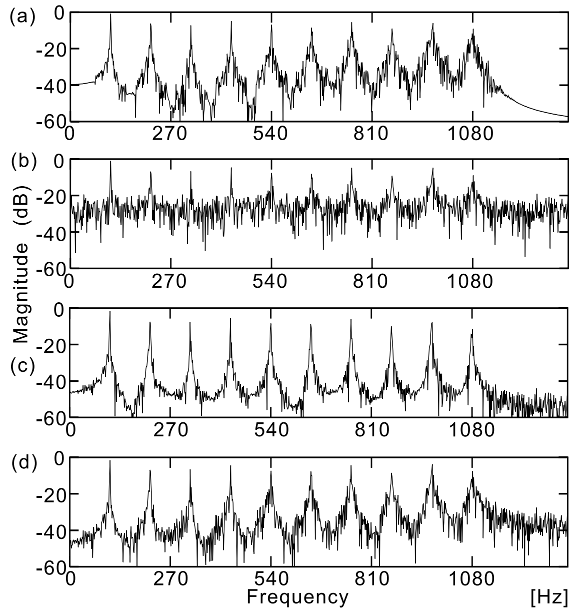

Figure 3: Spectral deformation caused by pitch fluctuation. (a)Original signal. (b)Noise-corrupted signal. (c)Extracted result by the constant-BW comb filter. (d)Extracted result by the constant-Q comb filter.

same with the above. The waveform collapse occurs in the constant-BW type, but in the constant-Q type the original signal is restored accurately as compared with the constant-BW type. Since the pitch fluctuation dilates each bandwidth of harmonic component in proportion to the harmonic frequency, the constant-Q structure is suitable for holding the spectral shape of the time-varying harmonics without any deformation (see Figure 3). The constant-BW structure, however, forces the spectral deformation significantly in higher frequency. It means the constant-Q comb filter has a robustness to the pitch fluctuation.

6. CONCLUSION

We have presented optimal filters for quasi-periodic signal extraction with a robustness to the amplitude and pitch fluctuations. To consider the time-varying phenomena, the Itô stochastic differential equation is employed as the signal model. The filter equation giving the optimum estimate is derived by the Kalman-Bucy filter method. It has been shown that the transfer function leads to the constant-BW comb structure, the constant-Q comb structure and these mixture, corresponding to the amplitude fluctuation, the pitch fluctuation and both of them, respectively. Finally, usefulness of the constant-BW and constant-Q comb filters are verified through some simulations.

7. REFERENCES

[1] J.S.Lim, A.V.Oppenheim and L.D.Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," IEEE Trans. ASSP, vol.ASSP-26, no.4, pp.354–365, 1978.

[2] A.Nehorai and B.Porat, "Adaptive comb filtering for harmonic signal enhancement," IEEE Trans. ASSP, vol.ASSP-34, no.5, pp.1124–1138, 1986.

[3] R.R.Bitmead, A.C.Tsoi and P.J.Parker, "A Kalman filtering approach to short-time Fourier analysis," IEEE Trans. ASSP, vol.ASSP-34, no.6, pp.1493–1501, 1986.

[4] S.H.Park et al., "Short-time Fourier analysis via optimal harmonic FIR filters," IEEE Trans. Signal Processing, vol.45, no.6, pp.1535–1542, 1997.

[5] K.Nishi and S.Ando, "An optimal comb filter for time-varying harmonics extraction," IEICE Trans. Fundamentals, vol.E81-A, no.8, pp.1622–1627, 1998.

[6] P.J.McLane, "Optimal linear filtering for linear systems with state-dependent noise," Int. J. Control, vol.10, no.1, pp.41–51, 1969.