IIR QMF-BANK DESIGN FOR SPEECH AND AUDIO SUBBAND CODING

Heinrich W. Löllmann, Matthias Hildenbrand, Bernd Geiser, and Peter Vary

Institute of Communication Systems and Data Processing (IND)
RWTH Aachen University, 52056 Aachen, Germany
loellmann@ind.rwth-aachen.de

ABSTRACT

A new speech and audio codec has been submitted recently to ITU-T by a consortium of Huawei and ETRI as candidate proposal for the super-wideband and stereo extensions of ITU-T Rec. G.729.1 and G.718. This hierarchical codec with bit rates from 8-64 kbit/s relies on a subband splitting by means of a quadrature-mirror filter-bank (QMF-bank). For this, an allpass-based QMF-bank is used whose design and implementation is presented in this contribution. This IIR filter-bank allows to achieve a significantly lower signal delay in comparison to the traditional FIR QMF-bank solution without a compromise for the speech and audio quality.

***Index Terms*— QMF-banks, IIR filter-banks, allpass filters, speech and audio coding, ITU-T**

1. INTRODUCTION

Analysis-synthesis filter-banks are commonly employed for subband coding. For such purposes, *quadrature-mirror filter-banks* (QMF-banks) are of special interest. Such filter-banks have been considered already in early proposals for subband coding, e.g., [1]. A QMF-bank has been used later for the first standardized 7 kHz wideband audio codec, ITU-T Rec. G.722 [2]. Tree-structured QMF-banks including the discrete wavelet (packet) transform can achieve an octave-band or a critical-band frequency resolution, which is commonly exploited for perceptual audio coding, cf., [3]. More recently, QMF-banks are used for hierarchical speech and audio coding as, for example, in ITU-T Rec. G.729.1 [4, 5].

At present, FIR QMF-banks are the most common choice and only a comparatively low number of publications deals with IIR filter-banks for subband coding (see, e.g., [3] for a survey). Hence, the application of IIR filter-banks to subband coding is a subject that is considered to be "largely unexplored" [3].

FIR filter-banks are often preferred as their design and implementation is apparently much easier in comparison to IIR filter-banks, cf., [6, 7]. However, IIR subband filters offer distinctive advantages over their FIR counterparts. They can achieve a comparable frequency selectivity with a much lower filter degree, that is, a lower signal delay and a lower computational complexity. Moreover, their non-linear phase characteristic is mostly tolerable for speech and audio processing due to the insensitivity of the human auditory system towards moderate phase distortions.

These benefits have motivated us to employ an IIR QMF-bank design in a recent proposal for a new speech and audio codec [8]. This codec has been developed by a consortium of Huawei (China) and ETRI (South Korea) in collaboration with the IND (RWTH Aachen University) and submitted to ITU-T as candidate proposal

for the upcoming super-wideband and stereo extensions of ITU-T Rec. G.729.1 and Rec. G.718. This hierarchical speech and audio codec employs the allpass-based IIR QMF-bank design of [9] instead of a more common FIR QMF-bank as used, for example, in ITU-T Rec. G.729.1 [5] and G.722 [2]. This approach has helped to achieve a high signal quality with a low signal delay.

In this contribution, the design and implementation of this IIR QMF-bank are presented in more detail. In Sec. 2, a brief overview of the considered coding system is given. The structure and design of the devised IIR QMF-bank are treated in Sec. 3 and Sec. 4. The implementation by fixed-point arithmetic is discussed in Sec. 5, and the paper concludes with Sec. 6.

2. SYSTEM OVERVIEW

The proposal of [8] is a *hierarchical* speech and audio codec with bit rates from 8-64 kbit/s and a bit-stream of 17 layers. A block diagram of this codec for mono coding is depicted in Fig. 1, adopted from [8] where a more detailed description can be found. For a sampling rate of 32 kHz, the input signal is split by an allpass-based IIR analysis QMF-bank into two critically subsampled signals with bandwidths of 0-8 kHz and 8-16 kHz. The subband signal $s_{\text{wb}}(n)$ is encoded by the *core codec* which is almost identical to ITU-T Rec. G.729.1 [5]. The subband signal $s'_{\text{swb}}(n)$ is encoded by a new *super-wideband* (SWB) encoder [8].

At the receiver side, only the core decoder is active for bit rates of 8-32 kbit/s as specified in [5]. For bit rates of 36 kbit/s and more, the SWB decoder provides the additional signal $\tilde{s}_{\text{swb}}(n)$ with a bandwidth of 8-14 kHz (due to the lowpass filtering at the encoder). A group delay compensation is performed to account for the signal delay of the (FIR or IIR) QMF-bank used in the core codec. The time-aligned output signal of the SWB decoder and the core decoder are merged by an IIR synthesis QMF-bank.

Due to the use of an IIR QMF-bank, the algorithmic signal delay of the G.729.1 core codec of 48.9375 ms is only increased by 2.21875 ms. The proposed codec provides an option to replace the *inner* FIR QMF-bank of the G.729.1 core codec by our IIR QMF-bank and to adapt the group delay compensation for the signal $s'_{\text{swb}}(n)$ accordingly (cf., Fig. 1). This enables a further signal delay reduction of about 1.5 ms. In the following, we will mainly treat this part of the proposal, which allows a comparison with the FIR filter-bank of a *standardized* codec for which a fixed- and floating-point reference implementation (in C) exist. Besides, the *outer* IIR QMF-bank is designed in a similar manner.

3. ALLPASS-BASED QMF-BANK

The allpass-based IIR QMF-bank design of [9] is considered. A *design I* is proposed which can make phase, amplitude and aliasing
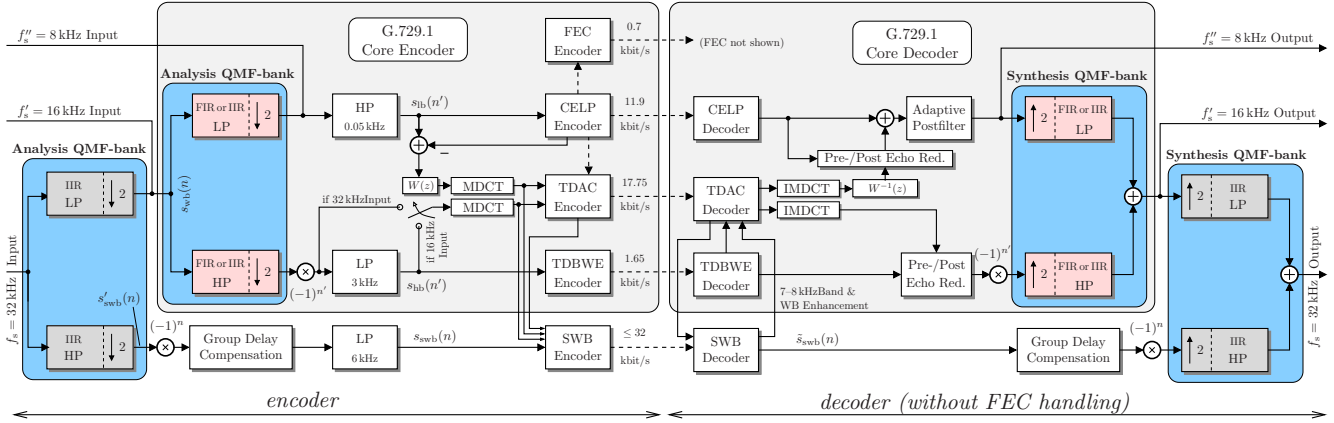
Figure 1: *High level block diagram of the codec proposed in [8]. Solid lines mark time-domain signals and dashed lines parameters. For the sake of clarity, the QMF-banks are depicted in a direct implementation and not their actual polyphase network implementation.*

distortions arbitrarily small, while a *design II* causes no aliasing and amplitude distortions and minimizes phase distortions. This design II is used here for the subband splitting in Fig. 1, despite its higher delay in comparison to the design I, as non-cancelled aliasing components can cause audible distortions.

Fig. 2 shows the polyphase network implementation of the considered two-channel IIR QMF-bank. The polyphase components of the *analysis filter-bank* are allpass filters of degrees K_0 and K_1 whose transfer functions in the z-domain read

$$A_i(z) = \prod_{\kappa=1}^{K_i} \frac{1 - \alpha_i(\kappa) \cdot z}{z - \alpha_i(\kappa)} ; \quad i \in \{0, 1\} ; \quad K_i \in \mathbb{N} \quad \text{(1a)}$$

$$|\alpha_i(\kappa)| < 1 \quad \wedge \quad \max\left\{\left|\alpha_i(\kappa)\right|\right\} < |z| \, \forall \, i, \kappa .$$

Here, allpass filters with real poles $\alpha_i(\kappa)$ are considered, but the QMF-bank design of [9] can also be applied to complex poles. The analysis lowpass and highpass filter are given by (cf., Fig. 2)

$$H_0(z) = A_0\left(z^2\right) + z^{-1} A_1\left(z^2\right) \quad \text{(2)}$$

$$H_1(z) = A_0\left(z^2\right) - z^{-1} A_1\left(z^2\right) . \quad \text{(3)}$$

These subband filters, each of degree $2\left(K_0 + K_1\right) + 1$, have the property $|H_0(e^{j(\pi-\Omega)})| = |H_1(e^{j\Omega})|$ which reasons the naming as *quadrature-mirror filters* (introduced in [1]).

The polyphase components of the *synthesis filter-bank* are given by [9]

$$B_0(z) = \frac{1}{2} \cdot P_{\text{AP}}^{(0)}(z) \cdot T_{\text{AP}}^{(1)}(z) \quad \text{(4)}$$

$$B_1(z) = \frac{1}{2} \cdot P_{\text{AP}}^{(1)}(z) \cdot T_{\text{AP}}^{(0)}(z) \quad \text{(5)}$$
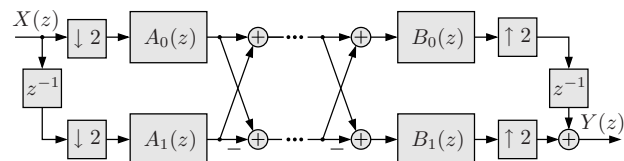

Figure 2: *Polyphase network implementation of the allpass-based two-channel QMF-bank (without spectral processing).*

with the allpass *phase equalizer*

$$P_{\text{AP}}^{(i)}(z) = \prod_{\kappa=1}^{K_i} \prod_{l=0}^{J_i(\kappa)-1} \frac{1 + \left(\alpha_i(\kappa) \cdot z\right)^{2^l}}{z^{2^l} + \left(\alpha_i(\kappa)\right)^{2^l}} ; \quad J_i(\kappa) \in \mathbb{N}_0, \quad \text{(6)}$$

and the allpass transfer function given by

$$T_{\text{AP}}^{(i)}(z) = A_i(z) \cdot P_{\text{AP}}^{(i)}(z) \quad \text{(7a)}$$

$$= \prod_{\kappa=1}^{K_i} \frac{1 - \left(\alpha_i(\kappa) \cdot z\right)^{N_i(\kappa)}}{z^{N_i(\kappa)} - \left(\alpha_i(\kappa)\right)^{N_i(\kappa)}} ; \quad N_i(\kappa) = 2^{J_i(\kappa)} \quad \text{(7b)}$$

for the two subbands $i \in \{0, 1\}$. If $J_i(\kappa) = 0 \, \forall \, \kappa \in \{1, \ldots, K_i\}$, Eq. (6) gives the empty product

$$P_{\text{AP}}^{(i)}(z) = 1 . \quad \text{(8)}$$

If no spectral processing takes place, the overall transfer function of the QMF-bank is equal to

$$T_0(z) = \frac{Y(z)}{X(z)} = z^{-1} \cdot T_{\text{AP}}^{(0)}\left(z^2\right) \cdot T_{\text{AP}}^{(1)}\left(z^2\right) . \quad \text{(9)}$$

This transfer function represents an allpass filter which causes only phase distortions for the reconstructed signal. It is obvious from Eq. (9) and Eq. (7) that these distortions depend on the absolute values for the allpass poles $|\alpha_i(\kappa)|$ and the filter degrees $N_i(\kappa)$. The determination of these coefficients is treated in the following.

4. FILTER DESIGN

The QMF-bank design has to ensure (almost) perfect signal reconstruction as well as a good separation between the subband signals (no 'spectral leakage') in order to avoid audible aliasing distortions despite (non-linear) coding operations. Thereby, the achievable frequency selectivity is implicitly limited by the need for a system with a low latency.

This trade-off is addressed in ITU-T Rec. G.729.1 by using the well-known least-squares (LS) design of Johnston [10].[1] This

---

[1]The design method for the QMF-bank is not specified in [5], but the filter coefficients of the C source code are given by the 64D design in [6].

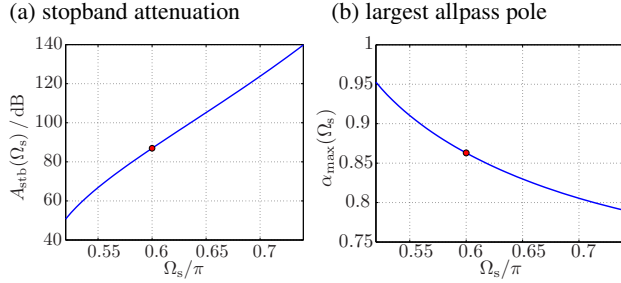(a) stopband attenuation (b) largest allpass pole

Figure 3: *Stopband attenuation A_{stb} and largest allpass pole $\alpha_{max} = max\{|\alpha_i(\kappa)|\}$ obtained by the equiripple IIR lowpass filter design of [11] for $K_0 = 3$, $K_1 = 2$, and different stopband frequencies Ω_s. The dots mark the chosen design at $\Omega_s = 0.6\,\pi$.*

FIR QMF-bank design achieves nearly perfect reconstruction with complete aliasing cancellation, and the subband filters can be efficiently implemented due to their linear phase property [5, 7].

The analysis filters of the IIR QMF-bank of Sec. 3 can be designed by a LS minimization of the stopband energy, cf., [7, 9]. However, this approach turned out to be less suitable for the design of subband filters with a high stopband attenuation as needed here. Instead, a superior performance is achieved by the *equiripple IIR lowpass filter* design of [11].

Fig. 3 shows the evaluation of this algorithm for different stopband frequencies Ω_s. Fig. 3-a illustrates the well-known trade-off between the conflicting goals for a stopband frequency close to $\pi/2$ and a high stopband attenuation. In addition, the absolute values for the allpass poles are also important here. As explained in Sec. 3, a value close to one requires synthesis polyphase filters of high degrees to achieve a sufficient phase equalization. Besides, poles near the unit circle are also problematic w.r.t. a fixed-point implementation. Fig. 3-b reveals that a higher stopband frequency Ω_s yields a lower magnitude for the largest allpass pole. The design marked by the dots has been found most suitable for our application, which yields the following allpass coefficients for Eq. (1):

$$\alpha_0(1) = -0.05423717361391; \; \alpha_1(1) = -0.62112612175302$$
$$\alpha_0(2) = -0.39882741266158; \; \alpha_1(2) = -0.19971974999496$$
$$\alpha_0(3) = -0.86293152951751 \,. \tag{10}$$

For the corresponding synthesis filter-bank, the values

$$J_i(\kappa) = \begin{cases} 4 & \text{for } i = 0 \;\wedge\; \kappa = 3 \\ 0 & \text{else} \end{cases} \tag{11}$$

are taken for Eq. (6), i.e., only the largest pole $\alpha_0(3)$ is considered.

Fig. 4 shows that the equiripple IIR subband filters achieve a higher stopband attenuation as well as a lower transition bandwidth in comparison to the FIR subband filters of the Johnston design as used originally for the G.729.1 coder.

The overall frequency responses of FIR and IIR QMF-bank are examined in Fig. 5. The transfer function of the IIR QMF-bank is an allpass filter according to Eq. (9). This transfer function causes no magnitude distortions, in comparison to the transfer function of the FIR QMF-bank (see Fig. 5-a), but it results a non-linear phase response. However, informal listening tests with different audio samples have shown that such phase distortions are inaudible.

Fig. 5-b shows that the group delay of the IIR QMF-bank is significantly lower than for the FIR QMF-bank but frequency
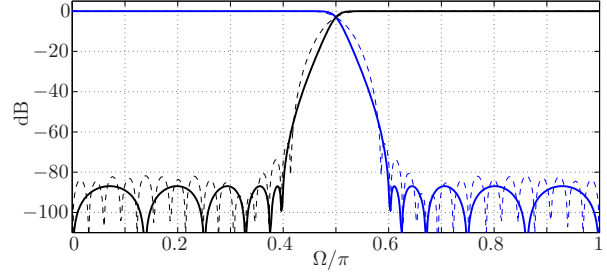

Figure 4: *Magnitude responses of the analysis subband filters: The dashed lines correspond to the FIR design of ITU-T Rec. G.729.1 and the solid lines mark the equiripple IIR subband filters.*
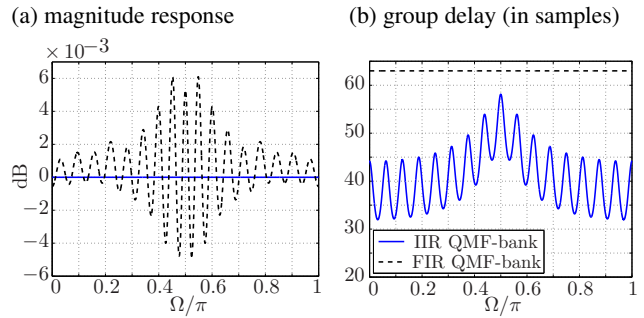
(a) magnitude response (b) group delay (in samples)

Figure 5: *Analysis of the overall transfer functions of the FIR QMF-bank of Rec. G.729.1 (dashed lines) and the proposed allpass-based IIR QMF-bank (solid lines).*

dependent. Therefore, the *signal delay* of the filter-banks is determined by the cross-correlation between input and output signal for a unit sample sequence as input signal. This resulted an algorithmic signal delay of 3.9375 ms for the FIR filter-bank and a signal delay of only 2.4375 ms for the IIR filter-bank ($f_s = 16$ kHz). The count of the *weighted million operations per second* (WMOPS) on basis of the C floating-point code submitted to ITU-T lead to a value of 1.358 WMOPS for the FIR QMF-bank and a value of only 0.977 WMOPS for the IIR QMF-bank.

The outer IIR QMF-bank is designed in the same manner as the inner IIR filter-bank taking into account the different sampling frequency of 32 kHz (cf., Fig. 1). The prototype lowpass filter is designed with parameters $K_0 = K_1 = 3$ and $\Omega_s = 0.57\,\pi$. Again, only the largest allpass pole is considered by the phase equalization by taking a value of $J_1(3) = 5$ and zero otherwise, cf., Eq. (11). The obtained outer filter-bank alone has a delay of only 2.2188 ms and a complexity of 2.290 WMOPS. The group delay compensation for the inner IIR QMF-bank is derived from Eq. (9): At the encoder, a filter with transfer function $T_{AP}^{(1)}(z^2)$ is used and a filter with transfer function $z^{-1} T_{AP}^{(0)}(z^2)$ is employed at the decoder, which leads to an additional complexity of only 0.388 WMOPS. For comparison, the G.729.1 coder has a fixed-point complexity of 36 WMOPS [4] and the SWB extension adds a complexity of 14.6 WMOPS (floating-point), see [8].

5. FIXED-POINT IMPLEMENTATION

Effects due to fixed-point arithmetic are an important issue (and frequent concern) regarding IIR filter-banks. The proposed IIR filter-bank consists solely of (different) allpass filters. Hence, the
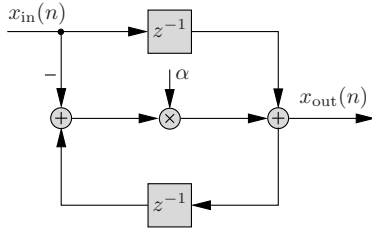
Figure 6: *Used implementation of an allpass filter of first order.*

Table 1: *SQNR for different audio signals (EBU-SQAM database).*

audio sequence	FIR QMF-bank	IIR QMF-bank
sinus sweep	84.52 dB	83.10 dB
guitar (Sarasate)	68.09 dB	62.01 dB
female speech	73.90 dB	68.25 dB

choice of the *filter form* to implement these filters has a significant influence on the performance of the system. Regarding a fixed-point realization, it is beneficial to implement the allpass filters in a cascade form according to Eq. (1). The single allpass filters in turn can also be implemented in different forms, which have been analyzed theoretically by a linear quantization noise model. A comparison of (five) different implementations for an allpass filter of first order has revealed that the realization of Fig. 6 is of special interest here. Its *noise gain*, which describes the amplification of the quantization noise by the filter [7], is given by

$$\mathcal{G}_{\text{noise}} = \frac{\sigma_{\text{out}}^2}{\sigma_q^2} = \frac{1}{1 - \alpha^2} \quad (12)$$

where σ_q^2 represents the variance of the input (quantization) noise and σ_{out}^2 marks the variance of the output noise. Due to Eq. (10), the noise gain for the used allpass filters of first order lies only between $1 < \mathcal{G}_{\text{noise}} < 4$. Besides, the input signal of the implementation of Fig. 6 needs not be scaled to reduce overflows by means of the $\mathcal{L}_2$ scaling rule, see [7].

The discussed (inner) IIR QMF-bank has been implemented in C using the 16 bit fixed-point operations of ITU-T [12].[2] Limit cycles are avoided by using accumulators with double precision, i.e., 32 instead of 16 bit word length. The quantization noise is measured by the *signal-to-quantization-noise ratio* (SQNR) which is determined by the difference between the output signals of the floating-point and fixed-point implementation, $y(n)$ and $\bar{y}(n)$:

$$\text{SQNR} = \frac{\sum_n y^2(n)}{\sum_n \left(\bar{y}(n) - y(n) \right)^2} . \quad (13)$$

Table 1 shows that the FIR QMF-bank achieves a slightly higher SQNR in comparison to the IIR QMF-bank. However, informal listening tests with different speech and audio signals have revealed no audible differences between the input and output signals obtained by the fixed-point implementations of the FIR and IIR QMF-bank. Besides, the signal distortions due to the non-perfect reconstruction of the QMF-bank and its fixed-point implementation are much lower than the distortions due to coding operations, especially at lower bit rates.

---

[2]The submission of a fixed-point implementation of the codec was not part of the candidate proposal for ITU-T [8].

6. CONCLUSIONS

The design and implementation of an allpass-based IIR QMF-bank and its application for subband coding is presented. The proposed design can achieve a signal reconstruction without aliasing and amplitude distortions, while the remaining phase distortions can be made arbitrary small (i.e., inaudible). It is shown how this IIR QMF-bank can be designed to achieve a comparable or even better frequency selectivity than a competitive FIR QMF-bank. The fixed-point implementation reveals that the devised IIR QMF-bank reaches a similar signal-to-quantization-noise ratio (SQNR) as the contrasted FIR QMF-bank of ITU-T Rec. G.729.1. However, the IIR QMF-bank achieves a significantly lower algorithmic signal delay than its FIR counterpart. This has been exploited by a recent proposal for a new speech and audio codec [8]. However, the ability to trade signal delay and phase distortions in a simple manner makes the presented filter-bank design also interesting for other speech and audio processing applications.

7. REFERENCES

[1] D. Esteban and C. Galand, "Application of Quadrature Mirror Filters to Split Band Voice Coding Schemes," in *Proc. of Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Hartford (Connecticut), USA, May 1977, vol. 2, pp. 191–195.

[2] ITU-T Rec. G.722, "7 kHz Audio Coding within 64 kbit/s," Blue Book, vol. Fascicle III.4, pp. 269-341, 1988.

[3] A. Spanias, T. Painter, and V. Atti, *Audio Signal Processing and Coding*, Wiley, Hoboken, New Jersey, 2007.

[4] S. Ragot, B. Kövesi, R. Trilling, D. Virette, N. Duc, D. Massaloux, S. Proust, B. Geiser, M. Gartner, S. Schandl, H. Taddei, Y. Gao, E. Shlomot, H. Ehara, K. Yoshida, T. Vaillancourt, R. Salami, M. S. Lee, and D. Y. Kim, "ITU-T G.729.1: An 8-32 Kbit/s Scalable Coder Interoperable with G.729 for Wideband Telephony and Voice Over IP," in *Proc. of Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu (Hawaii), USA, Apr. 2007, vol. 4, pp. 529–532.

[5] ITU-T Rec. G.729.1, "G.729 based Embedded Variable Bit-Rate Coder: An 8-32 kbit/s Scalable Wideband Coder Bitstream Interoperable with G.729," Mar. 2006.

[6] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, Upper Saddle River, New Jersey, 1983.

[7] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Upper Saddle River, New Jersey, 1993.

[8] B. Geiser, H. Krüger, H. W. Löllmann, P. Vary, D. Zhang, H. Wan, H. T. Li, and L. B. Zhang, "Candidate Proposal for ITU-T Super-Wideband Speech and Audio Coding," in *Proc. of Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, Apr. 2009, pp. 4121–4124.

[9] H. W. Löllmann and P. Vary, "Design of IIR QMF Banks with Near-Perfect Reconstruction and Low Complexity," in *Proc. of Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas (Nevada), USA, Apr. 2008, pp. 3521–3524.

[10] J. D. Johnston, "A Filter Family Designed for Use in Quadrature Mirror Filter Banks," in *Proc. of Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Denver (Colorado), USA, Apr. 1980, vol. 5, pp. 291–294.

[11] X. Zhang and T. Yoshikawa, "Design of Orthonormal IIR Wavelet Filter Banks using Allpass Filters," *Signal Processing*, vol. 78, no. 1, pp. 91–100, Oct. 1999.

[12] ITU-T Rec. G.191, "Software Tools for Speech and Audio Coding Standardization," 2005.