

Comparison and Evaluation of Constant-Q Transform-based Frequency Transformations

Lilia Lajmi^{1,*}, Stefan Roth¹

¹ University of applied Sciences, Wolfenbüttel, Germany
*Corresponding author, Email address: l.lajmi@ostfalia.de

Abstract

Over time, the Constant Quality Transform (CQT) has established itself as a useful tool in audio signal processing, with a variety of methods proposed to reduce computational complexity, increase its accuracy and improve its usability. The aim of this review is to compare three different CQT-based transformations in order to gain further insight on how differences in the implementation affect frequency and amplitude accuracy in terms of computation time. All methods are implemented with the same parameter set in a MATLAB framework to transform short audio frames and compute the relative deviation in amplitude and frequency in order to establish a level playing field. The results show that the limitation of the window length to the number of available samples does not only reduce the computation time, but also keeps the accuracy at a comparable level. The results are then evaluated and discussed to inspire further analysis, discussion, and verification.

Keywords: Spectral Transformation, Constant-Q Transformation

1. Introduction

The primary attribute of the CQT [1] is its ability to provide a time-frequency representation with a logarithmic frequency scale. The human auditory perception can be considered linear up to approximately 500 Hz; beyond this range, nonlinearities become more pronounced, encouraging the use of a logarithmic frequency axis, as shown in [2]. This characteristic renders the CQT to be particularly advantageous for tasks such as music and audio analysis.

However, this approach is associated with a significant computational cost. Since then, numerous contributions have been made to improve its efficiency, as evidenced by the work of researchers such as [3]. Following the initial implementations, including those by C. Schörkhuber et al. [4], subsequent contributions have focused on enhancing the applicability in real-time [5]. Additionally, an inverse transformation was missing, but methods proposed by A. Nagathil and R. Martin [6] and others [7][8] have suggested optimal reconstruction methods. The integration of these contributions culminates in the CQT being considered as a viable solution.

However, further refinement is possible. As previously mentioned, human auditory perception below 500 Hz adheres to a more linear scale. Consequently, methods have been proposed to reduce the frequency resolution in this range and thereby increase the frequency-time resolution by implementing a variable Q-factor [7]. This results in a non-constant Q transformation. A similar approach is taken in [9], where the Adaptive Quality Frequency Transform (AQFT) is proposed. This method is based on the premise of adapting the Q-factor to minimize the computational cost.

This prompts the following research question: How do different CQT-based transformations differ in terms of implementation and how does it affect the time-frequency representation? To answer this question, we set out to compare, implement, and evaluate three different methods in terms of amplitude and frequency accuracy and computation time relative to each other.

The structure of this review is as follows: Section 1 introduces the theory of the CQT and gives an overview of the important parameters. The subsequent section 2 describes the first method chosen, based on the theory proposed in [6], which resembles a CQT very close to the theory of section 1. Section 3 introduces the AQFT as a non-constant Q-factor version [9]. The following section 4 introduces a CQT based on the theory proposed in [5], which offers a different approach by performing the transformation in the frequency domain using the Fast Fourier Transform (FFT). A summary of the main differences is given in section 5, and the experimental setup and evaluation method are described in section 6. Finally, section 7 is dedicated to the discussion and evaluation of the results, followed by the conclusion in section 8.

2. The constant quality transform

The CQT is related to the Discrete Fourier Transform (DFT), which is defined as $X^{dft}(k)$ of a discrete time domain signal $x(n)$:

$$X^{dft}(k) = \sum_{n=0}^{N-1} w(n) \cdot x(n) \cdot e^{-j2\pi nk/N} \quad (1)$$

where n and k are the time and frequency parameters, respectively. $\{w(n)\}_{n \in [0, N-1]}$ is a normalized analysis window and N represents the DFT length. The constant spectral resolution Δf is defined as follows in regard to the sampling frequency f_s :

$$\Delta f = f_s/N \quad (2)$$

The DFT can be conceptualized as a filter bank with equally spaced center frequencies $f_k = k \cdot \Delta f$, and it can be formulated accordingly:

$$X^{dft}(k) = \sum_{n=0}^{N-1} w(n) \cdot x(n) \cdot e^{-j2\pi n f_k / f_s} \quad (3)$$

In order to provide a more accurate representation of human auditory perception, the center frequencies can be distributed exponentially:

$$f_k = f_0 \cdot 2^{k/b}; \quad (k = 0, \dots, K-1) \quad (4)$$

where b (bins) denotes the number of semitones per octave, corresponding to the Western musical scale of twelve semitones per octave. An octave is the interval between one musical note and another that has twice its frequency. The parameter b sets the frequency resolution based on the lower bound of frequencies f_0 and can be increased for $b \in \{24, 36, \dots\}$ to further augment the number of semitones per octave and the frequency resolution.

The parameter K , as defined in equation (5), denotes the total number of center frequencies, based on the lower bound frequency f_0 and upper bound frequency f_{max} , representing the total bandwidth.

$$K = \lceil b \cdot \log_2(f_{max}/f_0) \rceil \quad (5)$$

The CQT can be conceptualized as a filter bank with logarithmically spaced center frequencies f_k . The bandwidth of the k -th filter, denoted by Δ_k^{cq} , is defined by equation (6).

$$\Delta_k^{cq} = \Delta_{k-1}^{cq} \cdot 2^{1/b} \quad (6)$$

The factor Q is defined as the quotient between the center frequency and the bandwidth of each filter.

$$Q = f_k / \Delta_k^{cq} = f_k / (f_{k+1} - f_k) = 1 / (2^{1/b} - 1) \quad (7)$$

The quality factor Q remains constant for all filters. An increase in Q corresponds to an increase in frequency resolution, and vice versa.

In order to achieve a frequency-dependent resolution, it is necessary to select the window length in accordance with the center frequency. The desired bandwidth, which is denoted by (6) can be realized through the selection of a window length N_k .

$$N_k = f_s / \Delta_k^{cq} = Q \cdot f_s / f_k \quad (8)$$

The CQT $X^{cq}(k)$ of a discrete time-domain signal $x(n)$ can be derived from the DFT definition in equation (3) by inserting the logarithmically spaced center frequencies.

$$X^{cq}(k) = \sum_{n=0}^{N_k-1} w_k(n) \cdot x(n) \cdot e^{-j2\pi n f_k / f_s} \quad (9)$$

Accordingly, $w_k(n)$ denotes a normalized window function with the length N_k . Substituting equation (8) into equation (9), the CQT can be expressed as follows:

$$X^{cq}(k) = \sum_{n=0}^{N_k-1} w_k(n) \cdot x(n) \cdot e^{-j2\pi n Q / N_k} \quad (10)$$

In contrast to the DFT, the window length N_k is frequency-dependent and is not constant (see equation (8)). For the lower bound frequency f_0 , the maximum number of samples is required. The need for greater frame length at lower frequencies makes the CQT, as described above, not universally applicable to real-time applications. The suitability of the CQT depends on the parameters f_0 , f_{max} , b and the correlating total bandwidth and resolution.

3. Constant quality transform proposed by A. Nagathil and R. Martin

As previously stated, the publication by A. Nagathil and R. Martin [6] primarily proposes a method to reconstruct a signal from its CQT spectrum. Concurrently, a CQT implementation was presented, founded on the same theoretical underpinnings elucidated in section 1. The objective of this review is to compare different CQT implementations; therefore, it is necessary to choose one that resembles the fundamental CQT. Consequently, the term "Nagathil-CQT" will be employed henceforth to denote this particular method.

It is evident that this method does not deviate from the CQT; however, it does demonstrate discrepancies in the definition and calculation of the parameters f_0 and f_{max} . The lower bound frequency was set to $f_0 = 220$ Hz and the upper bound to $f_{max} = 7040$ Hz in the analysis, thus resembling musical notes. Accordingly, the upper bound frequency f_{max} is calculated as follows:

$$f_{max} = f_0 \cdot 2^{N_{octaves}} \quad (11)$$

$N_{octaves}$ describes the total number of octaves in the frequency range $[f_0, f_{max}]$.

On this basis, the center frequencies f_k and the associated window lengths N_k are calculated (equation (4) and (8)). Furthermore, the window function Hann was selected.

This approach establishes the desired resolution and bandwidth based on the parameters f_0 , f_{max} and b in advance, without requiring any knowledge of the available frame length of the signal. In scenarios where the provided samples fall short of the targeted resolution and bandwidth, it is posited that the frame length can be augmented by applying zero-padding.

In order to implement the Nagathil-CQT effectively, it is first necessary to realize the transform using a matrix multiplication based on equation (12). The matrix $CQT_{Nagathil}(k, n)$ can be computed beforehand based on the aforementioned parameters. The Nagathil-CQT can then be carried out by a matrix multiplication as illustrated in equation (13).

$$CQT_{Nagathil}(k, n) = w(k, n) e^{-j2\pi n Q / N_k} \quad (12)$$

$$X(k) = CQT_{Nagathil}(k, n) \cdot x(n) \quad (13)$$

4. The adaptive quality frequency transform

The AQFT proposed in [9] imposes limitations on the window length employed for the transform and maintains the logarithmic frequency resolution. The quality factor is adapted to the new length so that the resulting resolution hardly differs from the CQT frequency scaling. The time-frequency representation is still based on the CQT calculation. Consequently, the transform is well-suited for real-time applications, given its utilization of a limited number of samples, while also emulating the frequency scaling of the human auditory perception.

In the initial phase, which is analogous to the CQT, the lower bound frequency f_0 , the upper bound frequency f_{max} , and the number of frequency bins per octave b are designated. Additionally, a maximum window length, designated as N_{max} , is defined. Subsequently, the lengths of windows specified by N_k , as delineated in equation (8), are then contrasted with N_{max} . The smallest value obtained is designated as N'_k , and is subsequently employed for the transform. This step

consequently results in a reduction in the quality factor for lower frequencies. Therefore, the designation "Constant-Q-Transformation" is rendered invalid for this particular method.

In the subsequent step, the reduced quality is determined so that the center frequencies can be retained. However, this accommodates the bandwidths Δ_k^{cq} , making equation (7) not applicable in regard to b . The factor Q is adapted to the new length for each frequency. This results in the quality vector Q_k , which is then calculated for each bin k and each window length N_k , as follows:

$$Q_k = Q \cdot N'_k / N_k \quad (14)$$

It is necessary to adjust the center frequencies f_k so that the frequencies in the spectrum can be interpreted correctly. The calculation of the corrected center frequencies f'_k , is performed for the new window sizes N'_k and new factors Q_k .

$$f'_k = f_s \cdot Q_k / N'_k \quad (15)$$

The discrepancy between the frequency resolution of the AQFT and the CQT resolution remains small [9].

Equation (10) can still be used to transform the signal using the adapted parameters for each frequency. The resulting transform is designated as the AQFT. For an effective implementation, the AQFT was realized with a matrix multiplication. Based on equation (10), a matrix $AQFT(k, n)$ is defined as follows:

$$AQFT(k, n) = w(k, n) e^{-j2\pi n Q_k / N_k} \quad (16)$$

The corresponding matrix is initialized, based on the parameters f_0 , f_{max} and b . The analysis can then be carried out by means of a matrix multiplication as seen in equation (17).

$$X(k) = AQFT(k, n) \cdot x(n) \quad (17)$$

5. Constant quality transform proposed by F. Holzmüller et al.

The publication by F. Holzmüller et al. [5], proposes a real-time capable analysis algorithm based on the CQT. In addition, a reference MATLAB script is being used to implement this method. The term "Holzmüller-CQT" will be employed henceforth to denote this particular method.

The Holzmüller-CQT method involves a different approach, where the CQT equation (10) is rewritten as a convolution equation (18) [5] and computed in the frequency domain by multiplication, as shown in equation (19).

$$X(k) = x(n) * (w(k, n) \cdot \exp^{-j2\pi Q n / N_k}) \quad (18)$$

$$X(k) = \mathcal{F}^{-1}\{\mathcal{F}\{x(n)\} \cdot \mathcal{F}(w(k, n) \cdot \exp^{-j2\pi Q n / N_k})\} \quad (19)$$

The transition from the time domain to the frequency domain, and vice versa, necessitates the implementation of the FFT and its inverse, respectively.

As our primary focus is on analysis for evaluation, the proposed tuning-algorithm is not of interest. The lower bound frequency is set to musical tones, and the upper bound frequency is based on equation (11).

The calculation of the Q-factor is performed using the following equation [4]:

$$Q = (2^{1/b} - 2^{-1/b})^{-1} \quad (20)$$

When compared to equation (7), this results in a Q-factor that is half the size. This accommodates in bandwidths Δ_k^{cq} that exceed the distance between the center frequencies.

The calculation of the window functions necessitates the initial estimation of the required bandwidth Δ_k^{cq} for each frequency bin. Given that the transformation is executed in the frequency domain, it is optimal to calculate the window functions based on a substantial "lookup" window, with the length equivalent to the performed FFT. In this calculation, the closest values are selected according to the center frequencies f_k and their bandwidths Δ_k^{cq} . This approach constitutes a filter bank that is directly applied in the frequency domain.

To enhance low-frequency time resolution, it is proposed that the bandwidth be offset by a value γ (equation (21)). Making the transform a non-constant-Q transform, but reflecting the human auditory perception more accurately [4].

$$\Delta_{k,\gamma}^{cq} = f_k/Q + \gamma \quad (21)$$

For an efficient calculation, the transform is performed as a matrix multiplication, as shown in equation (23).

$$CQT_{Holzmueller}(k, f) = W(k, f) \quad (22)$$

$$X(k) = CQT_{Holzmueller}(k, f) \cdot X(f) \quad (23)$$

6. Comparison of the methods

Following the proposal of the three methods, the key differences between them are identified and categorized, as shown in Table 1.

In the initial aspect, the transformation domain is examined. The Nagathil-CQT and AQFT have been found to be congruent, as both operate within the time domain. This is in contrast to the Holzmueller-CQT, where the transformation is executed in the frequency domain.

The subsequent aspect pertains to the implementation of zero-padding, a technique not employed by the AQFT, as it adapts to the available number of samples. However, the Holzmueller-CQT and Nagathil-CQT do implement this technique, albeit for distinct reasons. The Nagathil-CQT employs zero-padding to maintain the selected resolution and ensure a constant Q factor, while the Holzmueller-CQT utilizes it to attain the requisite FFT length.

Accordingly, the AQFT does adapt the Q factor when necessary. In contrast, the Nagathil-CQT does not, and the Holzmueller-CQT proposes a method that is not applied in this scenario.

An additional element to consider is the variation in determining the parameters f_0 , f_{max} and b . However, as these parameters relate to the experimental setup (see section 6), they are equalized for all three methods.

Table 1 Key differences between Nagathil-CQT, AQFT and Holzmueller-CQT

Aspect	Nagathil-CQT	AQFT	Holzmueller-CQT
Domain	Time	Time	Frequency
zero-padding	Yes	No	Yes
Constant Q	Yes	No	If $\gamma = 0$

A more thorough investigation of these discrepancies can be conducted by comparing the calculated center frequencies and the corresponding bandwidths. As illustrated in Figure 1, the calculated center frequencies demonstrate no variation. However, minor discrepancies may be attributable to rounding errors, which introduce slight variations in the calculated values. See Table 2 for numeric values.

Table 2 Numeric values [Hz] for the center frequencies and the bandwidths for Nagathil-CQT, AQFT and Holzmueller-CQT

Nagathil-CQT		AQFT		Holzmueller-CQT	
f_k	Δ_k^{cq}	f_k	Δ_k^{cq}	f_k	Δ_k^{cq}
110.0	0.80	110.0	3.33	110.0	1.59
110.8	0.80	110.8	3.33	110.8	1.60
111.6	0.81	111.6	3.33	111.6	1.61
...
6889.1	49.92	6958.7	50.79	6889.1	99.48
6939.1	50.28	7003.2	51.12	6939.1	100.21
6989.4	50.64	7048.2	51.45	6989.4	100.93

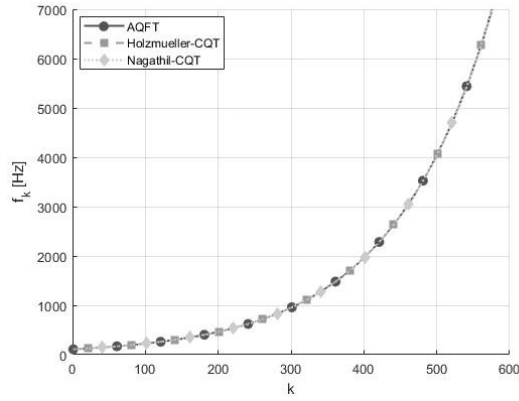


Figure 1 Center frequencies calculated by AQFT, Holzmüller-CQT and Nagathil-CQT; $f_0 = 110 \text{ Hz}$, $f_{max} = 7040 \text{ Hz}$ and $b = 96$

However, a more detailed analysis of the bandwidths reveals significant discrepancies. For the Holzmüller-CQT, the previously mentioned halving of the Q-factor is clearly visible, as the bandwidths are twice the size of those observed in the Nagathil-CQT and the AQFT for $k > 200$. In contrast, the AQFT demonstrates its restriction in its window size, resulting in a constant bandwidth in the lower frequency range ($k < 200$). See table 2 for numeric values.

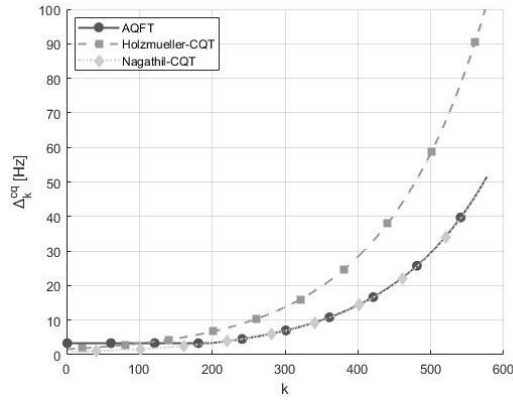


Figure 2 Bandwidths calculated by AQFT, Holzmüller-CQT and Nagathil-CQT; $f_0 = 110 \text{ Hz}$, $f_{max} = 7040 \text{ Hz}$ and $b = 96$

7. Experimental setup and evaluation method

The objective of the study was to establish a simple and consistent experimental setup to identify the correlations between implementation variations and outcomes. To this end, a total of 30 distinct static signals, with a length of $N = 4800$ samples and a sampling rate of $f_s = 16 \text{ kHz}$, were synthesized. Each signal contains up to six distinct spectral components, with known frequencies and amplitudes, amounting to a total of 119 components. The variable parameters were set as follows: $b = \{24, 48, 96\}$, $f_0 = 110 \text{ Hz}$, and $f_{max} = 7040 \text{ Hz}$, covering the required bandwidth and three different frequency resolutions. The signals are defined as short audio frames and are individually transformed using each of the proposed methods.

The computation time depends on the hardware; however, the present review exclusively compares relative differences, not absolute times. The measurement is achieved by multiplying the matrices defined in equations (12), (17) and (23) with the audio signal $x(n)$.

The dominant spectral components are extracted from the resulting spectrum by means of parabolic interpolation. The relative deviation of these components from the known amplitude and frequency values is then calculated by applying equations (24) and (25).

$$A_{rel.dev.} = \frac{|A_{known} - A_{detected}|}{A_{known}} \quad (24)$$

$$f_{rel.dev.} = \frac{|f_{known} - f_{detected}|}{f_{known}} \quad (25)$$

The undetected spectral components are counted and excluded from the calculation of the relative deviations.

Applying this procedure to all 30 audio frames, averaging the computation time, and calculating the mean and standard deviation from the relative deviation in amplitude and frequency, for only the spectral components detected by all three methods, yields the following results.

8. Results and discussion

The ensuing results are presented in this section. As illustrated in Figure 3, the computational time for all three methods and resolutions is presented. The computation time increases in proportion to higher b -values, i.e. an increase in resolution.

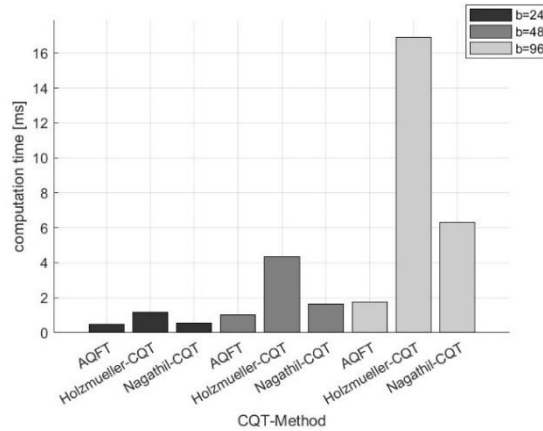


Figure 3 Computation time for AQFT, Holzmueller-CQT and Nagathil-CQT.

It is observed, that the Holzmueller-CQT method requires the most computation time for all resolutions. Conversely, the AQFT is the fastest at higher resolutions ($b = 48$ and 96). At the lowest resolution ($b = 24$) Nagathil-CQT takes approximately the same amount of time as the AQFT. The analysis suggests that the limited window size is a significant factor, showing a direct correlation with the resolution.

At the lowest resolution ($b = 24$), the window sizes for the AQFT and the Nagathil-CQT are nearly equivalent. This is due to the fact that the resolution is sufficiently low to guarantee that the window length N_k for the AQFT, barely exceeds the maximum window size N_{max} and the zero-padding for the Nagathil-CQT is applied at a minimum. Consequently, the execution time required for both methods is found to be comparable (AQFT=0.466 ms; Nagathil-CQT=0.5573 ms).

In contrast, the Holzmueller-CQT shows significant increases in computation time due to the application of zero-padding to achieve the required FFT resolution. This, in turn, results in an increased window size utilised for the transformation, which consequently leads to a prolonged computation time.

As shown in Figure 4, there are only minor discrepancies in the number of undetected spectral components. Nagathil-CQT and AQFT demonstrate a lower amount of undetected spectral components compared to Holzmueller-CQT; however, the difference appears to be statistically insignificant.

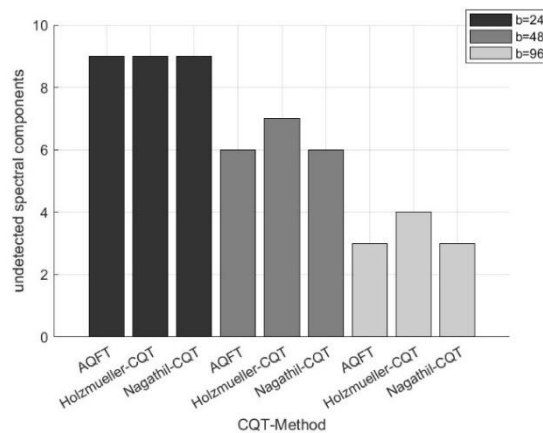


Figure 4 Amount of undetected spectral components for AQFT, Holzmueller-CQT and Nagathil-CQT.

In Figure 5, the relative deviation in amplitude for the three methods is illustrated. The Holzmueller-CQT has the highest standard deviation and mean for all resolutions. In contrast, the AQFT and Nagathil-CQT show a significantly

lower standard deviation and mean, which continue to decrease with an increasing resolution b . This may indicate a systematic error in the Holzmüller CQT, as it does not follow the trend of the AQFT and Nagathil-CQT.

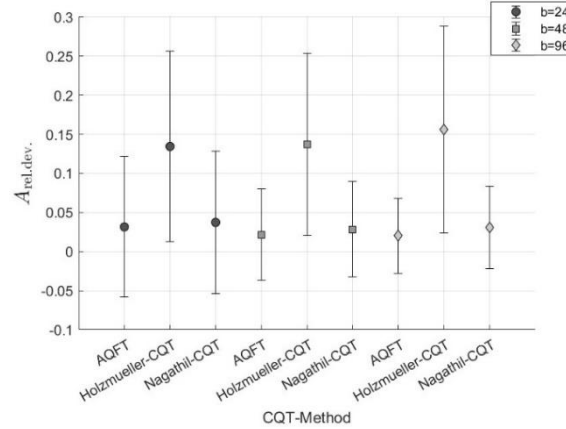


Figure 5 Relative amplitude deviation for AQFT, Holzmüller-CQT and Nagathil-CQT.

As illustrated in Figure 6, which presents the relative frequency deviation, the AQFT shows the lowest mean across all resolutions b and, for the highest resolution ($b = 96$), also exhibits the lowest standard deviation. In contrast, the Holzmüller-CQT shows a lower standard deviation for lower resolutions ($b = 24$ and 48), but also a higher mean, compared to the other methods. The Nagathil-CQT shows comparable higher standard deviation and mean for all resolutions b , however it has a slightly lower standard deviation compared to AQFT, for the lowest resolution ($b = 24$).

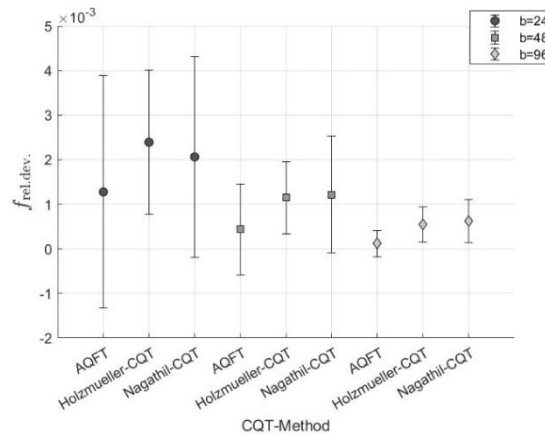


Figure 6 Relative frequency deviation for AQFT, Holzmüller-CQT and Nagathil-CQT.

Given the comparatively elevated computation time associated with the Nagathil-CQT and Holzmüller-CQT, for higher resolutions ($b = 48$ and 96), it is noteworthy that the AQFT maintains a consistently high level of accuracy. As illustrated in Figure 1, the spectral resolution of the three methods appears to be comparable. However, it is noteworthy that modifying the window length, i.e., the bandwidth (Figure 2), does not result in a substantial impact on the accuracy.

9. Conclusions

In summary, the comparison between AQFT, Holzmüller-CQT, and Nagathil-CQT demonstrated the impact of particular implementation methodologies on performance outcomes. The AQFT has been demonstrated to exhibit a distinct advantage in constraining the size of the window size. This strategy has been shown to reduce computation time while maintaining a comparable degree of accuracy. It is evident that this method possesses a notable advantage over alternative approaches.

This review examined various CQT implementations under specific conditions, and the findings can inform the development of new frameworks for ongoing development of CQT-based methods. Subsequent investigations could broaden the scope of audio signals examined to encompass longer and dynamic signals, thereby facilitating an analysis of the temporal behavior of spectral transformations. The Holzmüller-CQT could be utilized as a point of comparison, given its frequency-domain calculation approach. A more thorough exploration of computation time could be achieved by incorporating both the FFT and its inverse. The Nagathil CQT is also of interest, as it presents an inverse transformation. This could also offer the possibility of investigating signal synthesis.

References

- [1] J. C. Brown, „Calculation of a constant Q spectral transform“, *The Journal of the Acoustical Society of America*, Jg. 89, Nr. 1, S. 425–434, 1991, doi: 10.1121/1.400476.
- [2] D. B. Pisoni und R. E. Remez, Hg., *The handbook of speech perception*, 1. Aufl. Malden, Mass.: Blackwell Publ, 2008.
- [3] S. Nisar, O. U. Khan und M. Tariq, „An Efficient Adaptive Window Size Selection Method for Improving Spectrogram Visualization“ (eng), *Computational intelligence and neuroscience*, Jg. 2016, S. 6172453, 2016, doi: 10.1155/2016/6172453.
- [4] C. Schörkhuber, A. Klapuri, N. Holighaus und M. Dörfler, „A Matlab Toolbox for Efficient Perfect Reconstruction Time-Frequency Transforms with Log-Frequency Resolution“ in *AES 53rd International Conference*, London, UK, 2014.
- [5] F. Holzmüller, P. Bereuter, P. Merz, D. Rudrich und A. Sontacchi, Hg., *Computational efficient real-time capable constant-Q spectrum analyzer*, 2020. [Online]. Verfügbar unter: <http://www.aes.org/e-lib/browse.cfm?elib=20805>
- [6] A. Nagathil und R. Martin, „Optimal Signal Reconstruction from a Constant-Q Spectrum“ in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012*, Kyoto, Japan, 2012, S. 349–352.
- [7] C. Schörkhuber und A. Klapuri, „Constant-Q Transform Toolbox For Music Processing“, 2010.
- [8] N. Holighaus, M. Dorfler, G. A. Velasco und T. Grill, „A Framework for Invertible, Real-Time Constant-Q Transforms“, *IEEE Trans. Audio Speech Lang. Process.*, Jg. 21, Nr. 4, S. 775–785, 2013, doi: 10.1109/TASL.2012.2234114.
- [9] L. Lilia, K. Blum, M. Tucholke, „Introduction and evaluation of a new time-frequency transformation based on the CQT“, Wolfenbüttel, Ostfalia Hochschule für angewandte Wissenschaften, doi: 10.26271/opus-1254 (2021 June).

Authors



Lilia Lajmi

Lilia Lajmi is a professor of information technology at the University of Applied Sciences in Wolfenbüttel, Germany. She has a Ph.D. in substitution of packet losses in audio signals from the Technical University of Berlin, Germany. Her research interests include audio signal processing, audio coding, audio over IP networks, and error concealment.



Stefan Roth

Stefan Roth is a bachelor graduate in electrical engineering and information technology at the Ostfalia University of Applied Sciences in Wolfenbüttel, Germany. He is currently studying intelligent systems for his master's degree and works for the faculty of electrical engineering and information technology.