MDCT Domain Enhancements for Audio Processing

Abstract

Modified discrete cosine transform (MDCT) derived from DCT IV has emerged as the most suitable choice for transform domain audio coding applications due to its time domain alias cancellation property and de-correlation capability. In the present research work, we focus on MDCT domain analysis of audio signals for compression and other applications. We have derived algorithms for linear filtering in DCT IV and DST IV domains for symmetric and non-symmetric filter impulse responses. These results are also extended to MDCT and MDST domains which have the special property of time domain alias cancellation. We also derive filtering algorithms for the DCT II and DCT III domains. Comparison with other methods in the literature shows that, the new algorithm developed is computationally MAC efficient. These results are useful for MDCT domain audio processing such as reverb synthesis, without having to reconstruct the time domain signal and then perform the necessary filtering operations.

In audio coding, the psychoacoustic model plays a crucial role and is used to estimate the masking thresholds for adaptive bit-allocation. Transparent quality audio coding is possible if the quantization noise is kept below the masking threshold for each frame. In the existing methods, the masking threshold is calculated using the DFT of the signal frame separately for MDCT domain adaptive quantization. We have extended the spectral integration based psychoacoustic model proposed for sinusoidal modeling of audio signals to the MDCT domain. This has been possible because of the detailed analysis of the relation between DFT and MDCT; we interpret the MDCT coefficients as co-sinusoids and then apply the sinusoidal masking model. The validity of the masking threshold so derived is verified through listening tests as well as objective measures.

Parametric coding techniques are used for low bit rate encoding of multi-channel audio such as 5.1 format surround audio. In these techniques, the surround channels are synthesized at the receiver using the analysis parameters of the parametric model. We develop algorithms for MDCT domain analysis and synthesis of reverberation. Integrating these ideas, a parametric audio coder is developed in the MDCT domain. For the parameter estimation, we use a novel analysis by synthesis scheme in the MDCT domain which results in better modeling of the spatial audio. The resulting parametric stereo coder is able to synthesize acceptable quality stereo audio from the mono audio channel and a side information of approximately 11 kbps. Further, an experimental audio coder is developed in the MDCT domain incorporating the new psychoacoustic model and the parametric model.

Student : K. Suresh
Guided by : Prof T. V. Sreenivas