Multi-Band Spectral Subtraction Based Zoom-Noise Suppression for Digital Cameras

Kwang Myung Jeon*, *Student Member, IEEE*, Nam In Park*, *Student Member, IEEE*,
Hong Kook Kim*, *Senior Member, IEEE*, Myung Kyu Choi**, and Kwang Il Hwang**
*Gwangju Institute of Science and Technology, Gwangju 500-712, Korea
**Digital Imaging Business, Samsung Electronics, Gyeonggi-do 443-742, Korea

***Abstract*--This paper proposes a new noise suppression method to reduce zoom noise generated when audio signals are recorded with a digital camera. The proposed method is based on multi-band spectral subtraction that can suppress spectral components of noise related to reference zoom-noise in the modified discrete cosine transform domain. In particular, in the proposed method, each frame is classified as either a noise frame or a non-noise frame, and depending on this classification, the reference zoom-noise is updated and the degree of suppression is controlled. It is shown from performance evaluation that noise due to a zooming operation of digital cameras is successfully suppressed while maintaining audio quality.**

I. INTRODUCTION

Today's digital cameras are increasingly used to record video and audio, diminishing the use of camcorders. One drawback to audio recorded by digital cameras is in that a significant level of mechanical noise is introduced by the camera's zoom operations. One intuitive solution is to limit the speed of the zoom motor [1]. However, such a solution decreases the zooming speed of digital cameras, making it difficult to capture fast moving objects. Therefore, additional effort should be needed to overcome the trade-off between the zoom speed and the noise level.

As an alternative to reducing the zoom-noise level without decreasing the zoom speed, a mechanical noise suppressor for digital cameras was proposed by adapting reference noise [1]. Zoom-noise reduction in this approach was carried out by assuming that a priori information on the intervals of zoom motor operation was precisely known and that only zoom noise existed during those intervals to update the reference noise. However, it is difficult to measure the exact timing of the zoom operation owing to unexpected time delay and/or jittering between the zoom-motor movement and its activation time. Moreover, audio signals and zoom noise are commonly mixed during the zoom-noise intervals. These factors cause performance degradation of zoom-noise reduction.

In order to address the aforementioned issues, we propose a zoom-noise suppression method by incorporating a zoom-noise detection algorithm. By doing this, information pertaining to zoom-noise operation is unnecessary. The proposed method is based on multi-band spectral subtraction (MBSS), which suppresses the spectral components of noise related to reference zoom-noise in the modified discrete cosine transform (MDCT) domain [2]. Moreover, for a given audio frame, the zoom-noise detection algorithm first estimates the sub-band signal-to-noise ratios (SNRs). Then, it controls the degree of suppression in the MBSS and determines whether the audio frame is a zoom-noise frame or not, according to the distribution of the sub-band SNRs over frequency. In other words, the reference zoom-noise is updated only if this audio frame is declared as a zoom-noise frame.

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2012-010636) and by Samsung Electronics.

*Corresponding Author: Hong Kook Kim (hongkook@gist.ac.kr)

II. PROPOSED ZOOM-NOISE SUPPRESSION METHOD

Fig. 1 shows a flowchart of the proposed zoom-noise suppression method that operates in the MDCT domain because of higher performance of energy compaction and spectral resolution than in the Fourier transform domain [3]. First, the proposed method segments audio signals into a frame whose number of samples is 1024, corresponding to 32 ms at a sampling rate of 32 kHz. Next, it applies an MDCT to audio signals deteriorated by zoom noise and divides the MDCT coefficients into 49 sub-bands whose bandwidths are identical to those in MPEG advanced audio coding (AAC) [4]. After that, for a given l-th frame, the proposed method estimates SNR of the sub-bands, i.e., $SNR(l,k)$, $k = 0,\cdots,48$, by comparing the sub-band power of the audio signal and that of the reference zoom-noise. Note here that the reference zoom-noise is the zoom-noise signal recorded by a digital camera in a quiet environment. The estimated SNR for each sub-band is then used for zoom-noise detection. According to the result of zoom-noise detection, the reference zoom-noise is updated and the

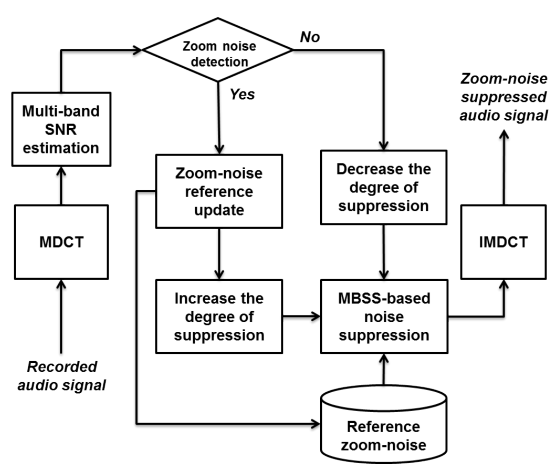

Fig. 1. Flowchart of the proposed zoom-noise suppression method.

degree of noise suppression is controlled for MBSS-based noise suppression. Finally, an inverse MDCT (IMDCT) is applied to obtain a zoom-noise suppressed version of the recorded audio signal.

As described above, the performance of the proposed method is highly dependent on that of the zoom-noise detection algorithm. The detection algorithm first counts the number of sub-bands whose SNR is below a predefined threshold, SNR_{thres}. That is,

$$N(l) = \frac{1}{49}\sum_{k=0}^{48} I(SNR(l,k), SNR_{thres}) \quad (1)$$

where $I(x,y)=1$ if $x \le y$, otherwise $I(x,y)=0$. The l-th frame is declared as a zoom-noise frame if $N(l) \ge N_{thres}$. In this paper, the parameters are set as $SNR_{thres}=0$ and $N_{thres}=0.7$ from the exhaustive preliminary experiments.

If the current frame is a zoom-noise frame, the recorded audio signal is averaged with the reference zoom-noise. Subsequently, this averaged reference zoom-noise is used for the MBSS-based noise reduction. The degree of noise suppression in MBSS is controlled depending on whether the current frame is a zoom-noise frame or not. In other words, the suppression factor is increased for a zoom-noise frame, but deceased otherwise.

Finally, the MBSS-based noise suppression is performed using sub-band SNRs, a suppression factor, and updated reference zoom-noise. By taking an IMDCT, we obtain a zoom-noise suppressed version of the recorded audio signal, as shown in Fig. 1.

III. PERFORMANCE EVALUATION

In order to evaluate the performance of the proposed method, the method was implemented using a commercially available compact digital camera with a zoom function. The camera was equipped with two electret condenser microphones for audio recording. The initial reference zoom-noise was obtained by averaging the zoom noise recorded using five different cameras of the same model. Test audio signals were recorded in an office environment while performing zoom operations. Even though the method was applied to audio signals, the proposed method had short latency enough not to cause video and audio synchronization problem. In other words, the latency was 36.94 ms in total, which summed up an algorithmic delay of 32 ms by the MDCT/IMDCT operation and the processing delay of 4.94 ms measured in the digital camera.

Fig. 2 shows a comparison of the spectrograms of zoom noise recorded in a quiet environment, audio signals recorded without any zoom operation, audio signals recorded during a zoom operation and zoom-noise suppressed audio signals by the conventional method [1] and by the proposed method, respectively. Compared to the spectral components shown in Fig. 2(c), the spectral components of zoom noise shown in Fig. 2(e) were clearly suppressed while the other spectral components were preserved. Moreover, the performance of the proposed method was more preferred than the conventional method, in terms of spectral similarity with audio signal without zoom-noise, as shown in Fig. 2(b).

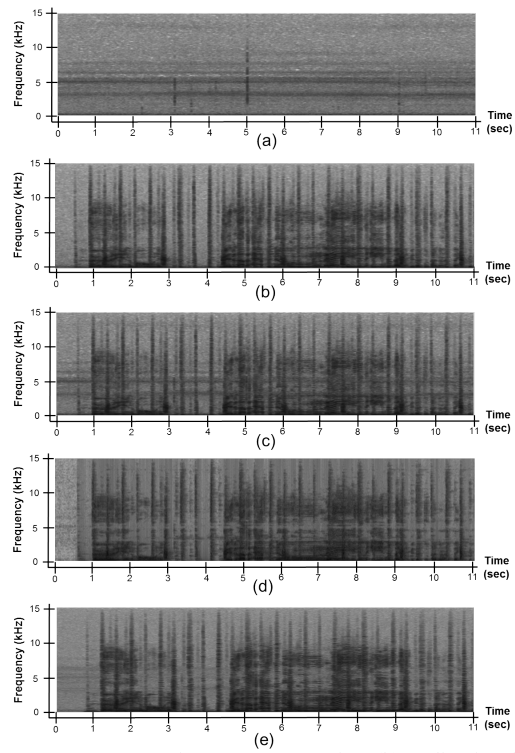

Fig. 2. Spectrogram comparison: (a) zoom noise, (b) audio signal without zoom-noise, (c) audio signal recorded during a zoom operation, (d) zoom-noise suppressed audio signal by the conventional method, and (e) zoom-noise suppressed audio signal by the proposed method.

IV. CONCLUSION

In this paper, a zoom-noise suppression method was proposed to reduce the mechanical noise generated by the zoom operation of digital cameras. The proposed method was performed by detecting zoom noise frames using sub-band SNRs, followed by updating the reference zoom-noise and controlling the degree of suppression. After applying the proposed method to audio signals recorded on a commercially available digital camera, it was shown that the proposed method could successfully reduce the zoom noise, resulting in better audio quality.

REFERENCES

[1] A. Sugiyama, T. Maeda, and K. Park, "A mechanical-noise suppressor based on *a priori* noise information for digital still cameras and camcorders," in *Proc. of International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, pp. 415-416, Jan. 2011.

[2] K. M. Jeon, N. I. Park, H. K. Kim, M. K. Choi, L. C. Hwang, and S. R. Kim, "MDCT-domain noise reduction with block switching for the application to MPEG audio coding," in *Proc. of International Conference on Advanced Signal Processing*, Seoul, Korea, p. 98, Mar. 2012.

[3] I. Y. Soon, S. N. Koh, and C. K. Yeo, "Noisy speech enhancement using discrete cosine transform," *Speech Communication*, vol. 24, no. 3, pp. 249-257, June 1998.

[4] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, and Y. Oikawa, "ISO/IEC MPEG-2 advanced audio coding," *Journal of the Audio Engineering Society*, vol. 45, no. 10, pp. 789-814, Oct. 1997.