33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Compression of surface myoelectric signals using MP3 encoding

Adrian D. C. Chan, *Senior Member IEEE*

Department of Systems and Computer Engineering, Carleton University,
Ottawa, CANADA, e-mail: adcchan@sce.carleton.ca

Abstract— **The potential of MP3 compression of surface myoelectric signals is explored in this paper. MP3 compression is a perceptual-based encoder scheme, used traditionally to compress audio signals. The ubiquity of MP3 compression (e.g., portable consumer electronics and internet applications) makes it an attractive option for remote monitoring and telemedicine applications. The effects of muscle site and contraction type are examined at different MP3 encoding bitrates. Results demonstrate that MP3 compression is sensitive to the myoelectric signal bandwidth, with larger signal distortion associated with myoelectric signals that have higher bandwidths. Compared to other myoelectric signal compression techniques reported previously (embedded zero-tree wavelet compression and adaptive differential pulse code modulation), MP3 compression demonstrates superior performance (i.e., lower percent residual differences for the same compression ratios).**

Keywords – data compression, electromyography, MP3, MPEG, myoelectric signals, remote monitoring, telemedicine

I. INTRODUCTION

Myoelectric signals (MES) are electrical signals associated with the contraction of a muscle, which are used in various research areas and clinical applications, including gait analysis, diagnosis of neuromuscular disorders, ergonomics, and rehabilitation. Data compression of MES is desirable to reduce storage and/or transmission bandwidth requirements. When performing long term MES data collection [1], multi-channel recordings [2][3], and in telemedicine applications [4][5], where the amount of data is large and/or the data transmission bandwidth is limited, the need is even greater.

MES is the spatiotemporal superposition of action potentials from multiple motor units, which results in a signal with strong stochastic characteristics. Signals that resemble random processes are often challenging to compress. Previous works in MES compression have demonstrated success when adopting techniques used in speech and audio processing. Norris and Lovely [6] applied adaptive differential pulse code modulation (ADPCM) to compress MES from static, isometric contractions. This work was later extended to include dynamic contractions [7]. ADPCM was able to reduce the MES from 12 bits per sample to 4 bits per sample, resulting in a compression ratio (CR) of 66.7%. Guerrero and Mailhes [8] compared the performance of three linear predictive coding methods (differential pulse code modulation, multi-pulse coder, and code excited linear predictive coder) and two techniques based on transform coding (discrete cosine transform and discrete wavelet transform). They found that transform coding techniques outperformed the linear predictive coding methods. Norris *et al.* [9] applied embedded zero-tree wavelet (EZW) encoding, with CR ranging from 60-95%. The ADPCM method was found to have a lower amount of signal distortion than EZW, for an equivalent CR; however, the EZW has the advantage of being able to easily adjust the CR to the signal fidelity required by a specific application. Carotti *et al.* [10] employed a technique based on the algebraic code excited linear prediction (ACELP), which is used for coding speech signals. They reported a CR of 87.3%, which corresponded to an ACELP bit rate of 12.2 kbps. A higher CR is possible by lowering the ACELP bit rate. Naturally, a higher CR would come at the expense of higher signal distortion but this was not examined in their work.

In this paper, MPEG-1 Audio Layer 3 (MP3) encoding is used to perform MES compression. MP3 is a lossy audio encoding scheme that leverages human psychoacoustic models to discard or reduce audio signal components. The success of other audio and speech compression techniques for MES originally inspired this investigation of MP3 compression. In addition, in other MES research, mel-frequency cepstral coefficient signal features, which are also based on psychoacoustic models, have been applied successfully in myoelectric speech recognition [11][12][13]. This suggested that there was merit in exploring MP3 as a technique for MES compression. As well, the ubiquity of MP3 in consumer audio applications would enable simple design and implementation of solutions for MES. There already exist real-time MP3 encoders and decoders in portable electronics, such as mobile phones and pocket portable computers, which would facilitate the implementation of a data logger. Applications that stream MP3 over the internet would enable remote monitoring applications.

II. METHODS

A. MES Compression

MP3 is a perceptual audio coding scheme that compresses signals in a manner that efficiently reduces the

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

amount of data (i.e., compresses the data as much as possible), while attempting to maintain the perceived audio (i.e., the reconstructed signals sounds exactly the same, or very similar, to the original signal for human listeners) [14][15]. MP3 compression consists of four main stages: 1) analysis filterbank, 2) perceptual model, 3) quantization and coding, and 4) bitstream encoding.

The analysis filterbank for MP3 belongs to a class of hybrid filterbanks and is comprised of a polyphase filterbank followed by a modified discrete cosine transform. The filterbank decomposes the input signal into subsampled spectral components, which are used with the perceptual model to compress the signal, leveraging the phenomena of auditory masking. Auditory masking is a result of the auditory systems being unable to separate components of a complex sound. The perceptual model provides estimates of masking thresholds, which are used by the quantization and coding block to remove signal components that are perceptually insignificant. An attempt is made to keep the noise below the masking threshold. A nonlinear power-law quantizer is also employed, along with Huffman coding. The result is then assembled into the MP3 bitstream. For more details on MP3 encoding and decoding, the reader is referred to [14] and [15].

In this work, MP3 encoding and decoding was performed using LAME 3.92 (The LAME Project, http://lame.sourceforge.net/), which converts audio files between the Windows WAVE (.wav) format and the MP3 format (.mp3). The default parameters for LAME were employed.

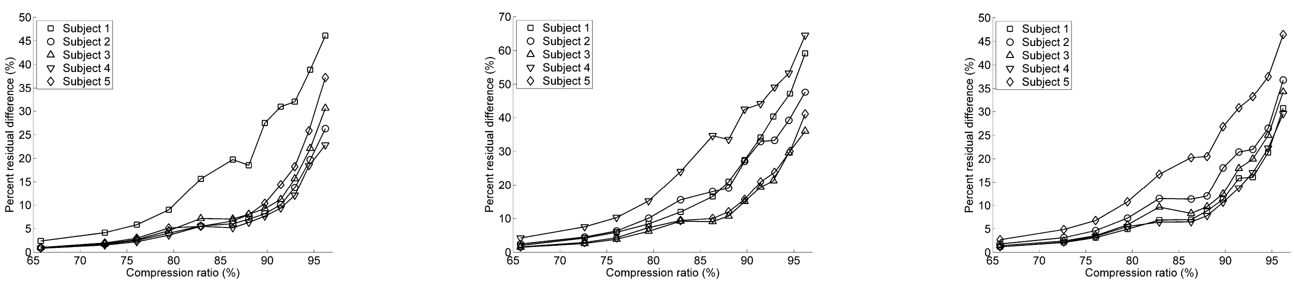
To perform MP3 encoding, digitized MES were first converted into the Windows WAVE format. The resultant file serves as the input to the LAME MP3 encoder. The bandwidth of MES (around 500 Hz) is much smaller than audio signals. In order to adapt the MP3 encoder to MES, the sampling frequency of the Windows WAVE file was modified to be 16 times the actual MES sampling rate (note that this does not involve any sample rate conversion, rather the sampling rate parameter of the file was simply set at 16 times higher than the actual sampling rate). Substituting a higher sampling frequency value, results in the MP3 encoder interpreting the MES as having frequencies that are 16 times higher than they actually are. This effect can be appreciated when listening to the WAVE files. The MES WAVE files sound like they have higher frequency components compared to WAVE files where the MES were stored using their actual sampling rate. Therefore, from the perspective of the MP3 encoder, incoming MES samples would appear to have a bandwidth that is comparable to speech (around 8 kHz = 16 $\times$ 500 Hz).

MP3 decoding used the reverse process of encoding. LAME was used to convert from the MP3 format to the Windows WAVE format. The MES was then extracted from the Windows WAVE file.

B. Dataset

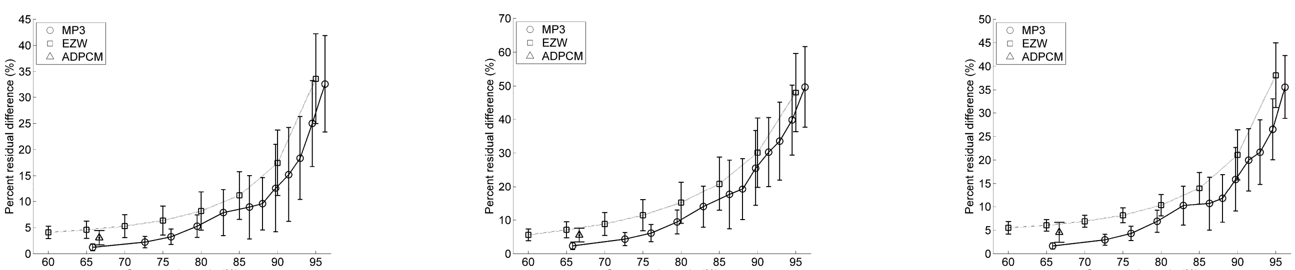
MP3 compression was evaluated on MES data from different contractions types (static and dynamic), as well as different muscle groups (biceps and triceps). Data used in this work were the same data used in [9]. Data were collected using AgAgCl electrodes in a standard bipolar electrode configuration. The electrodes were spaced 2.4 cm apart. The gain of the amplifier was adjusted to ensure the maximum use of the dynamic range of the analog-to-digital converter (A/D), without overranging the A/D or the amplifier. The MES were filtered with cutoff frequencies of 0.1 Hz and 1000 Hz. MES were sampled at 2000 Hz and stored in 16-bit integer format.

Data were collected from five subjects that underwent two contraction tasks: 1) a static contraction task, and 2) a



(a) (b) (c)

Figure 1. PRD from MP3 compression as a function of CR: a) biceps static contractions, b) triceps static contractions, and c) biceps dynamic contractions.



(a) (b) (c)

Figure 2. Performance comparison of three compression techniques (MP3, EZW, and ADPCM) for MES from: a) biceps static contractions, b) triceps static contractions, and c) biceps dynamic contractions. Mean PRD are plotted as a function of CR with standard deviation bars.

dynamic contraction task. In the static contraction task, 20 s of MES data were simultaneously acquired from the biceps and triceps of the right arm while the subject sustained a moderate isometric and isotonic co-contraction of both muscles. In the dynamic contraction task, 20 s MES data were acquired only from the biceps of the right arm. Subjects performed a cyclic task, moving their elbow between 40° and 90° flexion at a rate of 2 s/cycle (time regulated by an electronic metronome), holding a 2.27 kg dumbbell in their hand. Further details on the data collection methodology can be found in [9].

C. Analysis

The degree of compression was quantified using the CR:

$$CR = \frac{U_S - C_S}{U_S} \times 100\% \quad (1)$$

where U_S is the number of bytes in the original MES data and C_S is the number of bytes in the compressed data. A higher CR value corresponds to a higher amount of compression.

The bitrate for encoding was varied from 16 to 160 kbps; these bitrates produced a CR range from 65.8% to 96.2%, which was similar to the CR range used in [9]. Using the results reported in [9], MP3 compression was also compared against the EZW and ADPCM techniques.

Lossless compression techniques are capable of perfect reconstruction of the original signal. Lossy compression methods, such as MP3 encoding, are capable of obtaining higher CRs but at the expense of not being able to perform perfect reconstruction. In this work, the percent residual difference (PRD) is used to quantify the error in reconstruction:

$$PRD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - y_i)^2}{\sum_{i=1}^{N} x_i^2}} \times 100\% \quad (2)$$

where x_i are the original data samples, y_i are the data samples from the reconstructed signal, and N is the signal length. A lower PRD corresponds to a lower reconstruction error.

III. Results

MP3 compression results are shown in Fig. 1 for all subjects and contraction tasks. As expected, as the CR increases, there is a tradeoff of increasing PRD. Fig. 1a is a plot the PRD as a function of CR for MES from the biceps during static contractions and Fig. 1b is the plot for the triceps. The results show that the PRD for the biceps are lower than the PRD for the triceps. Fig. 1c is a plot of the PRD as a function of CR, for all subjects, for MES from the biceps during dynamic contractions. With the exception of subject 1, the PRD for MES from the biceps during the static contractions were lower than the dynamic contractions.

Fig. 2 are plots of the PRD as a function of CR, averaged across subjects; one standard deviation bars are also plotted along with the mean values. These figures also compare the performance of MP3 compression to the EZW and ADPCM techniques [9]. For all contraction types, MP3 compression outperforms the other two compression techniques. The mean PRD curve of the MP3 technique is 2 to 10% lower than the EZW, with larger differences occurring for higher CR. As the CR increases, the standard deviation of the PRD for both MP3 and EZW increases. This is not surprising, as the magnitude of the variation in PRD would be anticipated to increase as the magnitude of the PRD increases. At high CR, the standard deviations are large relative to the mean differences between compression methods; however, when examining the results on a per subject basis, the MP3 compression consistently outperforms EZW for all subjects and all CR. ADPCM has a CR of 66.7% and its PRD falls near the middle between the MP3 and EZW techniques.

IV. Discussion

The difference in PRD for the MP3 compression, between the biceps and triceps, is similar to the result found in the EZW compression [9]. The EZW also found that the MES from the biceps during static contractions had a lower PRD than MES from the biceps during dynamic contractions. Norris *et al.* [9] hypothesized that these differences in PRD were due to the bandwidth of the MES. The EZW performs coefficient reduction using a decreasing threshold to achieve a particular CR. Since high frequency components are generally smaller than low frequency components in MES, EZW compression results in the removal or distortion of high frequency content. The PSD of these triceps data were shown to have a larger bandwidth than these biceps data [9]. Likewise, the MES bandwidths from the dynamic contractions were larger than those from the static contractions [9].

A limitation of MP3 compression is the loss of high frequency content [14][16]. As with the EZW technique, the differences in PRD for MP3 compression between the different muscle groups (biceps versus triceps) and contraction types (static versus dynamic) can be also explained by the differences in bandwidth; the higher the bandwidth, the higher the PRD. Closer examination on a per subject basis provides further supporting evidence. Fig. 3 is a plot of the power spectral density (PSD) of the biceps MES from static contractions, for all subjects. Subject 1 has a noticeably larger bandwidth than the remaining subjects and in Fig. 1a, we see that there is a corresponding larger PRD for this subject. Plotting the PSD for the other contraction tasks revealed a similar pattern. For the triceps MES from static contractions, subject 4 had the largest high frequency components and correspondingly, the highest PRD (Fig. 1b). For the biceps MES from dynamic contractions, the high frequencies in the PSDs for subject 2 and 5 were very similar and distinctly higher than other subjects. Correspondingly, these two subjects had the highest two PRDs (Fig. 1c). The PRD for subject 5 was higher than subject 2, which could be explained by the larger low frequency components for subject 2 compared to subject 5.

V. Conclusions

Previous works have examined the use of popular audio compression techniques for compressing MES. Application of these audio compression techniques (e.g. ADPCM, EZW, and ACELP) was sensible because audio signals and MES

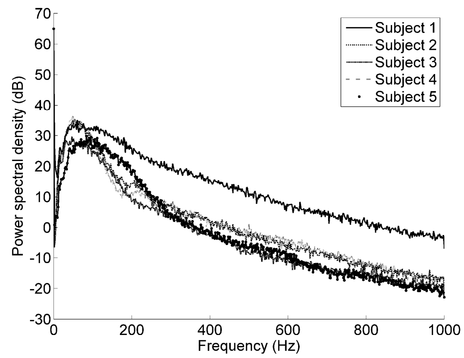


Figure 3. Power spectral density for biceps during static contractions.

both exhibit similar stochastic characteristics. To our knowledge, this work is the first to explore a perceptual audio encoder for MES compression. Perceptual encoders leverage the knowledge from psychoacoustics to achieve high levels of compression, while maintaining differences that are inaudible; however, audio perception of MES is of course irrelevant. Despite being designed using psychoacoustic models, MP3 has been demonstrated to be an effective technique for compressing MES. Results show that MP3 compression, compared to EZW and ADPCM, has a higher performance (quantified by PRD). Part of the increased performance can attributed to the Huffman encoding employed in MP3 compression, which is absent in the EZW and ADPCM implementations.

While PRD provides a quantitative measure of performance, it is difficult to discern the implications from these results. Future work will provide an examination of MP3 compression of MES within specific applications. For example, we will examine the effect of MP3 compression on the mean and median frequencies, which are conventional measures used for muscle fatigue estimation. Preliminary results on MES data used for research on myoelectrically controlled prostheses [17], demonstrated that a CR of up to 82.5% can be obtained with negligible impact; uncompressed MES had a classification accuracy of 93.87% across 30 subjects and the compressed MES were within 0.38% of this value. As the CR can be adjusted to improve signal fidelity, MP3 compression can be configured to the needs of any application.

The popularity of MP3, including portable consumer electronics and internet applications, makes this technique an attractive option for MES compression. Off-the-shelf solutions can be easily adapted to MES. In this work, to adapt MES to MP3 encoding, the input sampling rate was set at 16 times the actual MES sampling rate; equivalently, this would correspond to running the encoder or decoder at 1/16 the normal rate. MP3 encoding of multiple MES channels can be also performed, using a data buffer and a multiplexer, by interleaving segments of MES data into a data stream for a single MP3 encoder. With the current implementation one could implement a 16 channel system with mono MP3 encoding or a 32 channel system with a stereo MP3 encoder.

Performance of MP3 compression appears to be sensitive to high frequency content, similar to EZW. Newer audio compression formats, such as Windows Media Audio Pro (Microsoft, Redmond WA), Vorbis (Xiph.org Foundation, http://www.xiph.org), Real Audio (RealNetworks, Seattle WA) may have better high frequency performance and can be explored as alternative methods.

Acknowledgments

The author would like to thank Dr. Kevin Englehart at the University of New Brunswick, for access to data and results published in [9]. The author would also like to acknowledge the funding support of the Natural Sciences and Engineering Council of Canada, the Canadian Foundation for Innovation, and the Ontario Innovation Trust.

References

[1] P. J. Mork, R. H. Westgaard RH, “Long-term electromyographic activity in upper trapezius and low back muscles of women with moderate physical activity”, J Appl Physiol, vol. 99, pp. 570-578, 2005.

[2] G. Drost, D. Stegeman, B. van Engelen, M. Zwarts, “Clinical applications of high-density surface EMG: A systematic review”, J Electromyography and Kinesiology, vol. 16, no. 6, pp. 586-602, 2006.

[3] D. Staudenmann, A. Idsart Kingma Daffertshofer, D. F. Stegeman, J. H. van Dieen, “Improving EMG-based muscle force estimation by using a high-density EMG grid and principal component analysis”, IEEE Trans Biomed Eng, vol. 53, no. 4, pp. 712-719, 2006.

[4] S. C. Simmons, D. P. Moore, J. Wells, “Evaluation of technical approaches to tele-electromyography via internet protocol networks”, Telemedicine and e-Health, vol. 13, no. 1, pp. 33-40, 2007.

[5] H. van Dijk, H. J. Hermens, “Distance training for the restoration of motor function”, J Telemed Telecare, vol. 10, pp. 63-71, 2004.

[6] J. F. Norris, D. F. Lovely, “Real-time compression of myoelectric data utilizing adaptive differential pulse code modulation”, Med Biol Eng Comput, vol. 33, pp. 629-635, 1995.

[7] A. D. C. Chan, D. F. Lovely, B. Hudgins, “Errors associated with the use of adaptive differential pulse code modulation in the compression of isometric and dynamic myo-electric signals”, Med Biol Eng Comput, vol. 36, pp. 215-219, 1998.

[8] A. P. Guerrero, C. Mailhes, “On the choice of an electromyogram data compression method”, Proc IEEE Int Conf Eng Med Biol Soc, vol. 4, pp. 1558-1561, 1997.

[9] J. A. Norris, K. Englehart, D. Lovely, “Myoelectric signal compression using zero-trees of wavelet coefficients”, Med Eng Phys, vol. 25, pp. 739-746, 2003.

[10] E. Carotti, J. C. De Martin, R. Merletti, D. Farina, “Compression of surface EMG signals with algebraic code excited linear prediction”, Med Eng Phys, vol. 29, pp. 253-258, 2007.

[11] A. D. C. Chan, K. B. Englehart, B. Hudgins, D. F. Lovely, "Multi-expert automatic speech recognition using acoustic and myoelectric signals," IEEE Trans Biomed Eng, vol. 53, no. 4, pp. 676-685, 2006.

[12] K.-S. Lee, “EMG-based speech recognition using Hidden Markov Models with global control variables,” IEEE Trans Biomed Eng, vol. 55, no. 3, pp. 930-940, 2008.

[13] H. Manabe, Z. Zhang, “Multi-stream HMM for EMG-based speech recognition,” Proc IEEE Intl Conf Eng Med Biol Soc, pp. 4389-4392, 2004.

[14] K. Brandenburg, “MP3 and AAC explained”, AES 17th Intl Conf High Quality Audio Coding, 1999.

[15] D. Pan, “A tutorial on MPEG/audio compression”, IEEE Multimedia, vol. 2, no. 2, pp. 60-74, 1995.

[16] S.-H. Oh, W.-J. Yoon, Y.-H. Cho, K.-S. Park, K.-M. Kim, “A new spectral enhancement algorithm in MP3 audio”, IEEE Trans Consumer Elec, vol. 52, no. 1, pp. 196- 199, 2006.

[17] A. D. C. Chan, G. C. Green, “Myoelectric control development toolbox”, 30th Conf Cnd Med Biol Eng Soc, M0100, 2007.