Compression of Medical Diagnostic Signals inspired by High Quality Audio Encoding Schemes

Abhejit Rajagopal, Jerry D. Gibson

Department of Electrical and Computer Engineering, University of California, Santa Barbara. USA

{*abhejit,gibson*}*@ece.ucsb.edu*

Abstract

We investigate the applicability of frequency-based compression algorithms to vital medical diagnostic signals. Filter banks and tree structures that are popular in the compression of high quality audio signals (MP3, AAC) have considerable similarity to the multi-scale analysis (wavelets) used in modern medical diagnostic software. In some sense, these wavelet representations are sub-optimal both in terms of compression and analysis because the filter bandwidths are fixed even though the signals being inspected are highly non-stationary. In this paper, we investigate the structure and implementation of the MP3 (MPEG-1 Layer-3) encoding standard and consider the construction of an audio-inspired medical signal compression scheme. Application of such an algorithm to electrocardiogram (ECG) and electroencephalogram (EEG) signals is studied, and the diagnostic reconstruction accuracy is evaluated using open-source diagnostic software.

Keywords: compression, filter banks, medical diagnosis

1. Introduction

Analysis of medical diagnostic signals is a hot topic of research both from a physiological and a purely mathematical perspective. Due to the complexity of biological systems, exact models for source signal generation are extremely difficult to develop. Instead, clinicians focus their attention on waveform features that indicate compromised vitals (i.e. biomarkers). When biomarkers are physical pieces of evidence, such as the existence of a protein or antibody, methods to physically extract or detect them are fundamentally possible (however difficult they may be to develop). Examples of these include allergen testing via visible skin reactions [1], modern disease detectors using qPCR [2, 3].

However, the problem is significantly more challenging when the evidence is implicitly embedded in a signal. Often times, researchers use data science and signal processing concepts to extract salient [4] (or recently, non-salient [5]) biomarkers, but this result is not always the most optimal. The source of the trouble remains that an appropriate model of the source-generator is not known, so relevant waveform features cannot be extracted efficiently. Instead, many groups have tried to parametrize known inputs in order to develop a receiver-side model of the biological information, as to at least discriminate between symbols [6].

Even still, adequate models of continuous signals is still a big challenge in many measurement schemes, particularly when the signal does not exhibit quasi-periodicity (e.g. beats) such as in electroencephalogram (EEG) signals. Recently, a popular trend is to apply wavelet transforms to signal streams as to help identify correlations between channels and detect salient features [7]. However, there is considerable ambiguity in the structure of such transforms, the relative frequency content being searched, and optimality of such diagnosis [8]. As such, methods of discovering adequate biomarkers plagues the field of medical science because, fundamentally, appropriate pass-band locations are still not exactly known. The wavelet-based analysis provides a good insight into the complexity of the problem (non-stationarity of signals, frequency-time resolution, multi-frequency multi-scale analysis) [9], but still requires significant tuning to be considered clinically robust. Examples of the tuning process include threshold and window-size optimization, which may vary significantly between sensors, patients, and environments.

In this paper, we consider the construction of a dynamic filter bank to both develop an architecture

for learning relevant frequency bands and compress (parametrize) medical diagnostic signals. Our methodology is inspired from the encoding of high quality audio signals, where the entropy is minimized with an error bound by applying perceptual coding to the output of a filter bank via a quantization mask. Our initial goal is to develop an variable-frequency encoder (i.e. signals are encoded on a frame-by-frame basis) which allows for diagnostic (as opposed to perceptual) quantization criteria. Although a full encoder-decoder pair is required for practical storage and reconstruction of such signals, in this paper we consider only the fundamental building block: the filter bank and diagnostic constraints.

Outline:
Midterm – filter banks theory, single-channel pseudo-QMF
Final – mutli-channel ECG/EEG encoding and testing

2. Analysis Subband Filter Bank in Audio Encodings

One of the most popular tools for time-frequency mapping is a filter bank, which divides the signal spectrum into multiple frequency subbands. Both the MPEG-1 audio codec (e.g. MP3) and Advanced Audio Codec (AAC) use a time-frequency analysis block to extract parameters from the time-domain input signal that are suitable for quantization and encoding based on perceptual criteria. [10]. In this section we review their implementations in order to understand how the exploitation of perceptual criteria might be traded for designations of diagnostically (ir)relevant information. The idea here is that biological signals may be coded more optimally if we knew what information was relevant (e.g. subband energy, subband correlation, entropy).

2.1. MP3

In the MPEG-1 audio encoder, the input audio stream is passed through a filter bank that divides the audio input into 32 uniformly spaced frequency subbands. Typically, these are realized through band-pass filters. Although the filters can be practically implemented with any bandpass filter, the MP3 codec uses a cosine-modulated lowpass prototype to select each frequency band (Figure 1). In essence, one prototype baseband lowpass filter is created and the signal is subsequently modulated from the M center frequencies for the filtering step. This effectively corresponds to a series of

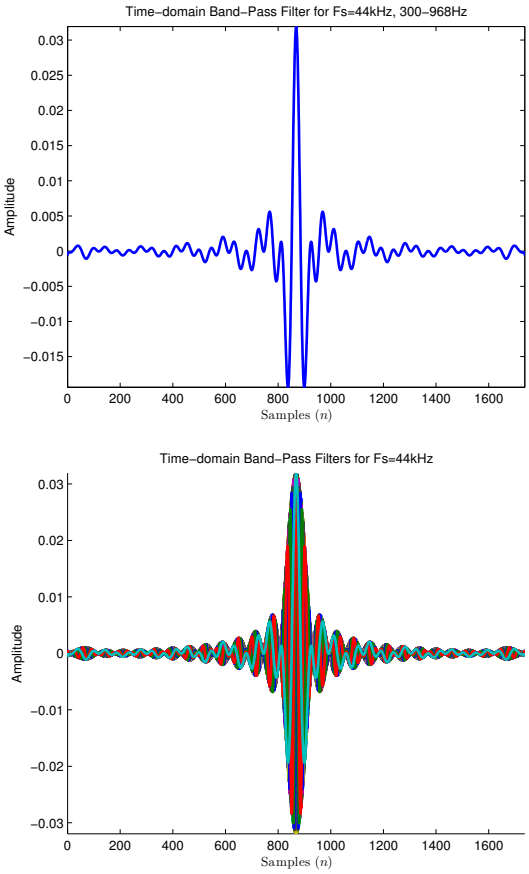


Figure 1: A modulated-cosine lowpass prototype implemented in MATLAB is shown, along with its orthogonal variants (modulo modulated). A) A bandpass filter from 300Hz - 968Hz, B) Overlay of all effective bandpass filters.

bandpass filters (Figure 2). The prototype impulse response should contain 512 samples, with sidelobe suppression better than 96-dB in the stop-band and the output ripple is less than 0.07 dB [11].

One challenge in designing these filters is the overlap region between consecutive bands. Because all the filters must be implemented as FIR systems (discrete convolution), the frequency response within the ideal stopband has a slope (Figure 3). This causes redundant information to be stored in consecutive subbands, which in turn produces aliasing during decimation and amplitude noise during reconstruction at the encoder. The incorrect frequency-amplitude information is corrected in the proceeding analysis block (Modified DCT) which divides each subband into several frequencies and interfaces with the perceptual audio coder to decide which frequencies to keep within each band [12]. Although this seems like a potential source of errors or

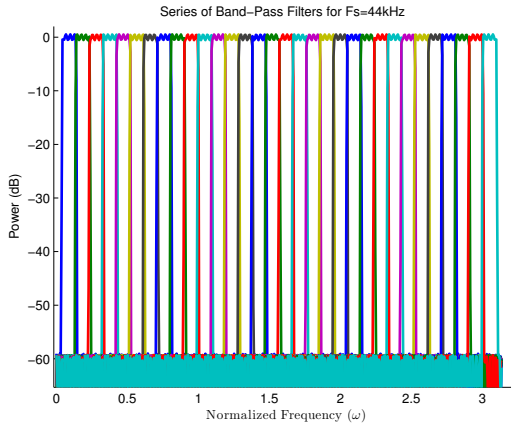


Figure 2: A series of band-pass filters was constructed in MATLAB. This implementation has a pass-band response is near 0db (no gain) while the stop-band response is roughly -60db.

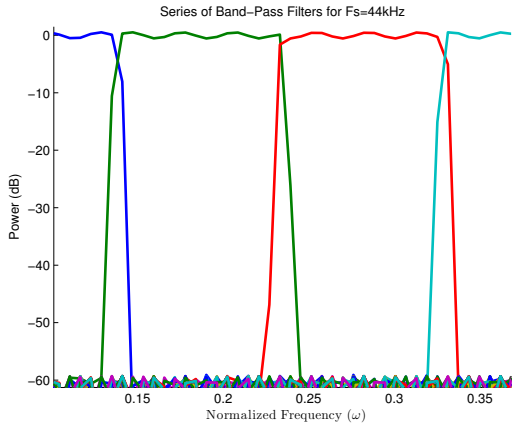


Figure 3: There is a considerable overlap region when constructing bandpass filters directly, even using the modulated-cosine prototype.

redundancy, it has been shown that a 50% overlap ratio is ideal and produces near-perfect reconstruction using specially developed algorithms, such as the Overlap-Add algorithm (OLA) [13, 14].

Following the modulation, each subband signal can be maximally decimated by the filter bank modulus factor (M) because the information is now stored as a baseband signal. There are technical issues here which increase the effect of aliasing such as the small, but non-zero, stop-band frequency components, but these are typically corrected for in the decoder. Ideally, the decoder simply performs the inverse operation of the encoder, effectively synthesizing or reconstructing the waveform by upsampling, modulating, and adding waveforms. The overall operation of the filter bank

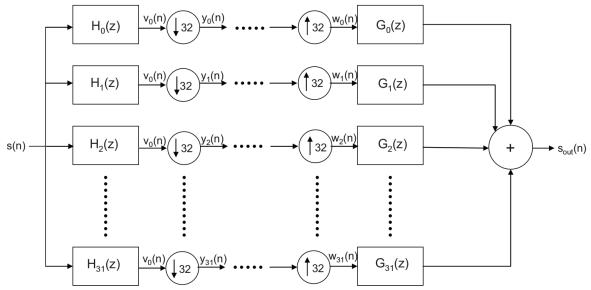


Figure 4: A uniform 32-band critically sampled analysis-synthesis filter bank. [10]

encoder-decoder can be understood as filtering, modulation, decimation, upsampling, modulation, and addition (Figure 4).

By effectively combining the steps, this process can be implemented efficiently using fast algorithms using a pseudo- quadrature mirror filter(QMF). The implementation is consider a pseudo-QMF because the signal there is near-perfect reconstruction (due to aliasing, quantization, mirroring synthesis filters, etc)[15, 16]. The filter bank outputs can be derived as (see [10] for derivation),

$$s_i = \sum_{k=0}^{63} \sum_{j=0}^{7} M_{ik}(c(k+64j)x(k+64j)) \tag{1}$$

where i is the subband index and ranges from 0 to 31, c contains the analysis window coefficients, x is the audio input sample vector, and $M \in \mathcal{R}^{32x64}$ holds the analysis coefficients. M_{ik} is computed as,

$$M_{ik} = cos\left[\frac{(2i+1)(k-16)\pi}{64}\right] \tag{2}$$

Noting that the index i indicates the subband number, Equation 1 can be reformulated as a filter convolution (Equation 3), which agrees with our initial formulation of the filter bank (Figure 4).

$$s_i = \sum_{m=0}^{511} x(n-m)h_i^c(m) \tag{3}$$

In this view, each $h_i^c(m)$ is has a modulated impulse response. That is, for $i = 0:31$ and $n = 0:511$,

$$h_i^c = cos\left[\frac{(2i+1)(n-16)\pi}{64}\right] \tag{4}$$

It is clear from this expression (Equation 4) that the re-

sponse of each subband is a modulated version of the prototype response, with a cosine term to shift the low-pass response to the appropriate frequency band. The center frequencies of these filters are at odd multiples of $\pi/(64 * Fs)$, where Fs is the audio sampling rate and each has a bandwidth $\pi/(32 * Fs)$. Following this step, each subband output is further divided by 18 spectral lines in the MDCT block to provide finer frequency resolution. One obvious question here is why the MDCT isn't used initially to provide high resolution, such as in the implementation of the AAC codec.

2.2. The Psychoacoustic Model

Fundamentally, the efficiency of subband and perceptual coding depends on the match between the properties of the designed filter bank and the characteristics of the input signal [17]. As such, a feature of the MP3 encoder is to code distant frames (typically greater than three) using a separate set of filter coefficients. In addition to the frame based approach, MP3 employs the MPEG-1 psychoacoustic model to disregard frequencies which would be masked (inaudible) to the human ear during playback.

This module produces a time-adaptive spectral pattern that emulates the sensitivity of the human ear. The MPEG audio coder compresses audio data by removing its perceptually irrelevant components. The masking characteristics of a given signal depend both on the location of frequency components in time and their intensity. Essentially, the encoder estimates a masking threshold for each subband, below which particular frequencies cannot be perceived by the average human listener. Calculation of the masking threshold involves the critical band analysis of the signal.

Critical band analysis is a model developed by experimentation and theory of human hearing. The human cochlea, which responsible for transducing sound to the basilar membrane and in turn to the sterovilli, can be viewed as a log-spaced filter bank (Figure 5). The bandwidth of each critical band (i.e. each filter), denoted as $BW_c(f)$, has been derived empirically for the average listener [18, 14] and can be approximated using the following expression.

$$BW_c(f) = 25 + 75\left[1 + 1.4\left(\frac{f}{1000}\right)^2\right]^{0.69} \tag{5}$$

In addition to measuring the energy content in each one of these frequency subbands, the MPEG-1 psychoacoustic model calculates the proportion of tonal and noise-like components, known as the tonality index of

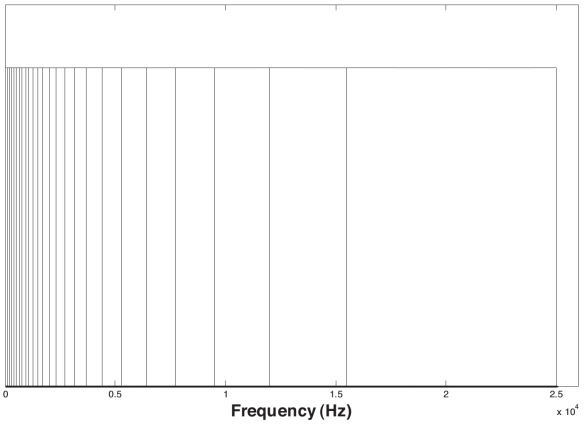


Figure 5: The human cochlea can be viewed as a log-spaced filter bank. [10, 14]

the frame. In general, tonal components are more predictable than broadband signals (which are more noise-like) and there exhibit different masking characteristics. The unpredictability measure of each spectral component is computed using the complex spectra of the time-aligned and windowed audio data. To reduce abrupt changes between frames, a Hanning window is applied prior to the FFT computation. A measure of the unpredictability is given by,

$$cw(f) = \frac{\sqrt{\left[\left(R_j(f)cos(\phi_j(f) - \hat{R}_j(f)cos(\hat{\phi}_j(f)\right)^2 + \left(R_j(f)sin(\phi_j(f) - \hat{R}_j(f)sin(\hat{\phi}_j(f)\right)^2\right]}}{(R_j(f) + |\hat{R}_j(f)|} \tag{6}$$

where

$$\hat{R}_j(f) = 2R_{j-1}(f) - R_{j-2}(f),$$
$$\hat{\phi}_j(f) = 2\phi_{j-1}(f) - \phi_{j-2}(f)$$

In short, the MPEG-1 Layer-3 algorithm achieves compression by allowing quantization noise in subbands where the human ear is the least sensitive. This quantization is directly dependent on the bit allocation algorithm; the two modules interact to minimize audible (perceived) noise. The design criteria in this module that is of interest to the application of medical diagnostic signals is the computation of the mask. Interestingly, the psychoacoustic model II does not make a dichotomous distinction between tonal and non-tonal components. Instead, it transforms the spectral data into a so-called partition domain and estimates the tonality index in each partition [10]. This measure actually determines

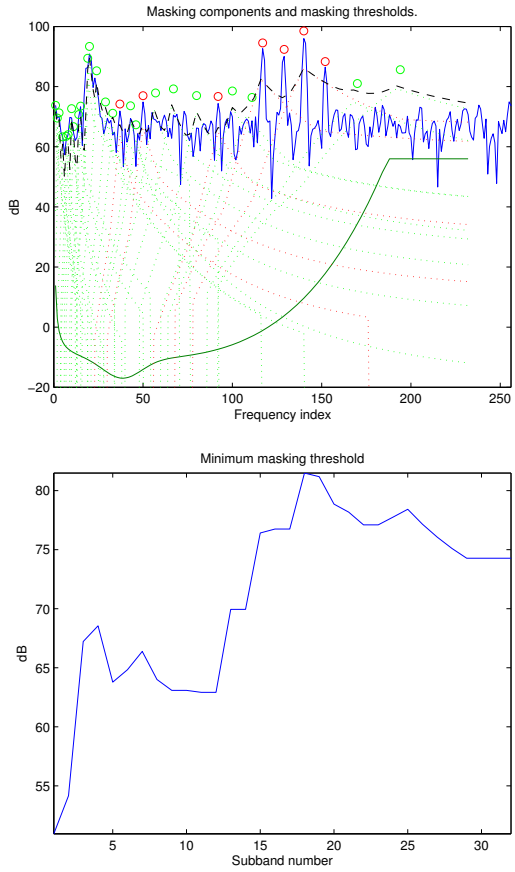


Figure 6: A signal is analyzed using [19]. A) The tonal (pink) and non-tonal (green) maxima and masking thresholds calculated using a spreading function. B) The resulting 32 subband minimum-masking threshold for the current frame.

the amount of masking and critical band ranges.

In this work, the psychoacoustic model II implementation by Petitcolas was used [19] due to its convenience and interoperability with the previously developed filter bank in MATLAB. Running this code, we are able to generate the appropriate subband thresholding masks which would be used by the quantizer (Figure 6).

In conclusion, the time-to-frequency mapping provided by the filter bank not only allows for a reduction in signal redundancy, but also enables an estimate of the perceptual noise at the encoder (i.e. so that a tighter error bound may be achieved by modifying the quantization mask).

2.3. Midterm Conclusions

Thus far, we have studied in depth the implementation of filter banks and perceptual auditory models in the high quality audio compression MP3 codec. Although much of this information has been previously reviewed and analyzed [10], we have specifically identified the components of the codec which are applicable to the coding of medical diagnostic signals.

In particular, these codecs employ a filter bank which effectively divides the frequency spectrum into several subbands. In the case of MP3, this corresponds to 32 uniform subbands which are further divided by 18 spectral lines in the MDCT block (not studied here). The psychoacoustic model produces a mask which allows for several frequencies to be subdued in amplitude, or left out entirely, from the coding stage. The masking threshold is determined using measures of the tonality and predictability of the input signal, suggesting that the quantizer minimize some measure of perceptual entropy.

In light of this, we suggest that defining an analogous tonal index, frequency-spreading, and masking relationship will enable the compression of non-stationary medical signals such as EEG. As opposed to minimizing the perceptual audible entropy, one strategy may be to maximize interband correlations. That is, we should first divide the signal into several fine frequency subbands (such as in the AAC implementation) and measure their energy content. Using this spectral energy histogram, combinations of harmonic and non-harmonic bins could be considered as a possible filter.

This is an improvement over the generic wavelet structure provided by Daubechies wavelets [7]. Although Daubechies wavelets are effective in identifying harmonics and grouping together several frequency bands to produce an effective multi-scale analysis, such a transform is fundamentally non-data specific. As such, there is no guarantee that particular outputs of a Daubechy filter bank produce any useful information. The interband combination strategy suggest here can be made to follow a smooth function, but the fundamental concept is that both harmonic and non-harmonic subbands should be considered. Its worth nothing that considering all possible combinations results in a set of non-orthogonal filters, although the basis used to generate them span the entire frequency band.

A popular feature of many EEG detection algorithms [7] is to detect maxima of the resulting subband coded signals. We note here the similarity to the operation of finding spectral peaks in the MPEG-1 psychoacoustic model. Instead of using these to determine acoustic masking thresholds, we can optimize the correlation of peaks across many channels. In the detection of epileptic seizures, for example, high frequency oscillations (HFOs) are correlated across many channels. By clus-

tering phase-locked peaks, a high sensitivity, patient-specific biomarker can be computed [8]. Typically, this is done via optimizing the threshold and phase-margins at various steps.

Instead, we are suggesting that the filters which are producing these peaks be optimized to produce the highest correlation ratio. In cases where a solution is not able to be found (i.e. non epileptic patients), high energy subbands can simply be returned without further computation. Such a scheme both provides compression to medical diagnostic signals as well as aides the analysis process.

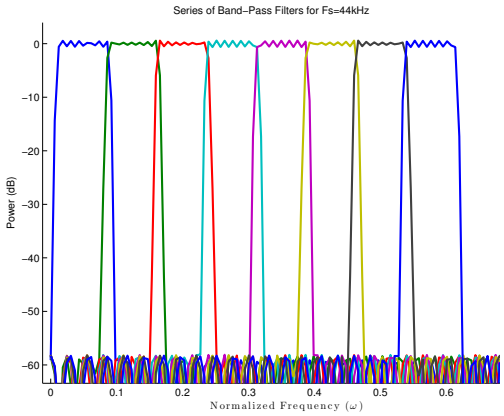


Figure 7: A uniform filter bank from 0-25Hz with some overlap. The overlap could be reduced with more computation time (i.e. a sharper filter).

3. Compression of Medical Diagnostic Signals

In light of our previous discussion, we implemented and measured the performance of a spectral compression scheme on real EEG data. In particular, we consider patients suffering from sporadic seizures, and measure the performance of a machine learning tool at identifying seizure and non-seizure instances in time. Both the machine learning tool and compression algorithms were developed in MATLAB for this project. Inspiration for the machine learning algorithm was provided by [20].

3.1. Data and preprocessing

Data for 25 patients was acquired from the CHB-MIT Scalp EEG database. Each patient's scalp EEG was monitored continuously at 256Hz across 23 channels, in 1-hour increments, for roughly 46 hours. Each patient was known to have seizures and, as such, includes both non-seizure and seizure data. In a preprocessing step, we collect all 46 hours, but split the data into two groups for testing and training, separately.

Each 23-channel signal stream was further divided into 2-second windows upon which a filter bank was provided. At the suggestion of [20] a 8-bin uniform filter bank ranging from 0 to 25 Hz was applied (Figure 7). The 2-second frame width effectively provides a 0.5Hz spectral resolution, so our filter bank is a fairly good representation of the frequency content in the signal (3.125 Hz per bin, ideally).

3.2. Machine Learning Diagnostic

Based on previous work [20], the spectral content of each frame was used as a feature representing each 2-second window. Using annotations provided by the dataset, a kernalized support-vector machine (SVM) was trained to discriminate seizure instances from non-seizure instances. Specifically, we concatenated three 2-second spectral frames (as opposed to computing the PSD of one 6-second frame), and used these as feature vectors representing each 2-second window. This captures both the transitioning features over 6-seconds and has been shown to be more robust than considering only individual pulses. The classifier system was implemented using LibSVM 3.20.

3.3. EEG Compression Schemes

In short, passing each 23-channel EEG signal through the 8-filter filterbank produces an 8x23 matrix representing the PSD in each channel. For a requested compression level, the lowest z coefficients are set to 0. Although not implemented or studied here, this reduces the number of bits required to be coded at the encoder (i.e. all zeros requires only 1 bit to store). We do not compare storage and entropy encoding schemes here, rather we compare the performance of the medical diagnostic tool at different compression levels, measured by the accuracy of seizure detections and number of spectral coefficeints retained. We considered two compression schemes. The first scheme, termed *Naive Coefficient Knockout* performs a simple reduction in the number of stored spectral coefficients. The second scheme, termed *Chaos Correlation Knockout* uses a diagnostic insight generate a quantization mask for the multi-channel spectrum. It should be noted here, the spectral coefficients are not directly used for signal reconstruction– the reconstruction estimate is produced by projection to the kernel space inside the SVM block.

3.3.1. Naive Coefficient Knockout

In short, the lowest z spectral coefficients of the multi-channel spectrum were eliminated by setting them to zero. The number z is set by the requested compression percentage. In this work, a compression of 1% corresponds roughly to 1 or 2 coefficents being knocked out, out of 184. We do not test at high distortion rates, because we do not want to sacrifice the robustness of our diagnostic. The idea here is to understand how much compression could be achieved while still maintaining a high accuracy.

Algorithm 1 Naive Coefficient Knockout

function NCK(*spectrum*,*cmpr*)
  let *spectrum* be the 8 x 23 multi-channel spectrum associated with a given 2-second EEG window.
  let *cmpr* be the % compression requested.
  $z \leftarrow floor(numel(tmp_spectrum) \times cmpr)$
  $minX, minY \leftarrow argsort(spectrum(:))$
  $minX \leftarrow minX(1:z)$
  $minY \leftarrow minY(1:z)$
  $finalspectrum \leftarrow spectrum$
  $finalspectrum(minInds_x, minIndx_y) \leftarrow 0$
end function

3.3.2. Chaos Correlation Knockout

We developed a new compression scheme, termed *Chaos Correlation Knockout*, which provides a quantization mask to the multi-channel spectrum based on the filter-coefficient cross-correlation [8]. In essence, the filter coefficient value is treated as a random vector and its covariance matrix is computed. Because the filter-coefficient random-vector contains samples from all 23 channels, the covariance matrix provides a robust measure of the correlation between channels.

In particular, we use the diagonal elements of the covariance matrix (auto-covariance of each random variable) to measure how correlated particular frequencies are. We then generate a mask which weights the frequencies in proportion to how correlated they are (cross-correlation) with the maximally correlated frequency band (max autocorrelation). To see how this can be implemented efficiently, refer to Algorithm 2.

After applying this quantization mask, Naive Coefficient Knockout is employed to remove the lowest z coefficients corresponding to the requested compression level. Such a spectral feature is postulated to accentuate highly correlated frequencies amongst both channels (auto-correlation entries on the digonal) and frequency bins (cross-correlation entries off the diagonal).

Algorithm 2 Chaos Correlation Knockout

function CCK(*spectrum*,*cmpr*)
  let *spectrum* be the 8 x 23 multi-channel spectrum associated with a given 2-second EEG window.
  let *cmpr* be the % compression requested.
  $covspectrum \leftarrow cov(spectrum^H)$
  $maxInd \leftarrow argmax(diag(covspectrum)$
  $mask \leftarrow covspectrum(:, maxInd)$
  $mask \leftarrow mask/max(mask)$
  $mask \leftarrow mask \otimes 1_{8x23}$
  $Qspectrum \leftarrow mask. * spectrum$
  $finalspectrum \leftarrow NCK(Qspectrum, cmpr)$
end function

4. Testing

Finally, we tested our system on varying levels of compression. In essence, we compressed each spectrum, constructed 6-second epochs, and trained an SVM to discriminate using identicle training parameters. The accuracy of seizure detection was compared to evaluate the effect of compression. As previously mentioned, the spectral coefficients are not used to directly reconstruct the a time-domain EEG waveform. Rather, the compressed spectrum is used train a SVM classifier. The ability to still discriminate effectively between seizure-like and non-seizure like samples provides a strong indication of how well the compression scheme can work in a clinical environment where dignostic accuracy is extremely important. Due to the time constraints of this project, only 5 patients were able to be tested thoroughly (~24 hours training data, ~12 hours testing data). The results, depicted in Figure 8 are averaged over these five patients. As is depicted, the CCK scheme actually increases the prediction accuracy at low compression levels, although this claim must still be rigorously tested and verified to make a definitive conclusion.

5. Conclusion

In summary, we applied the compression scheme inspired by high quality audio coding to medical diagnostic signals. We employed a basic compression scheme (naive coefficient knockout) and developed our own scheme (chaos correlation knockout) using inspiration from the recently developed chaos theory of EEG signals. Although initial results look promising, additional testing is required to confirm the validity of the correlation measure.

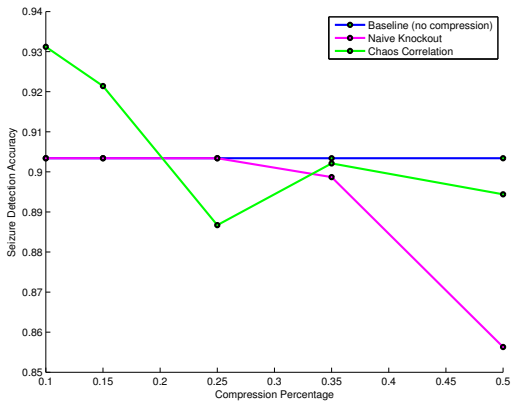


Figure 8: The baseline seizure prediction accuracy is plotted along with the accuracy of the NCK and CCK compression schemes.

References

[1] S. J. Arbes, P. J. Gergen, L. Elliott, and D. C. Zeldin, "Prevalences of positive skin test responses to 10 common allergens in the us population: results from the third national health and nutrition examination survey," Journal of Allergy and Clinical Immunology, vol. 116, no. 2, pp. 377–383, 2005.

[2] A. Rajagopal, C.-F. Chang, O. Plettenburg, S. Petry, A. Scherer, and C. L. Tschirhart, "Nanopillar field-effect and junction transistors with functionalized gate and base electrodes," July 12 2013. US Patent App. 13/941,295.

[3] A. Rajagopal, A. Scherer, A. Homyk, and E. Kartalov, "Supercolor coding methods for large-scale multiplexing of biochemical assays," Analytical chemistry, vol. 85, no. 16, pp. 7629–7636, 2013.

[4] A. Rajagopal, A. C. Nguyen, and D. M. Briggs, "Neuropass: A secure neural password based on eeg," UCLA Department of Electrical Engineering, Reports, 2013.

[5] T. Akiyama, H. Otsubo, A. Ochi, E. Z. Galicia, S. K. Weiss, E. J. Donner, J. T. Rutka, and O. C. Snead, "Topographic movie of ictal high-frequency oscillations on the brain surface using subdural eeg in neocortical epilepsy," Epilepsia, vol. 47, no. 11, pp. 1953–1957, 2006.

[6] X. Hu, P. Xu, F. Scalzo, P. Vespa, and M. Bergsneider, "Morphological clustering and analysis of continuous intracranial pressure," Biomedical Engineering, IEEE Transactions on, vol. 56, no. 3, pp. 696–705, 2009.

[7] H. Adeli, Z. Zhou, and N. Dadmehr, "Analysis of eeg records in an epileptic patient using wavelet transform," Journal of neuroscience methods, vol. 123, no. 1, pp. 69–87, 2003.

[8] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy," Biomedical Engineering, IEEE Transactions on, vol. 54, no. 2, pp. 205–211, 2007.

[9] M. Akin, "Comparison of wavelet transform and fft methods in the analysis of eeg signals," Journal of medical systems, vol. 26, no. 3, pp. 241–247, 2002.

[10] J. J. Thiagarajan and A. Spanias, "Analysis of the mpeg-1 layer iii (mp3) algorithm using matlab," Synthesis Lectures on Algorithms and Software in Engineering, vol. 3, no. 3, pp. 1–129, 2011.

[11] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, and M. Dietz, "Iso/iec mpeg-2 advanced audio coding," Journal of the Audio engineering society, vol. 45, no. 10, pp. 789–814, 1997.

[12] R. E. Crochiere and L. R. Rabiner, "Multirate digital signal processing, 1983," Pretice Hall, pp. 127–192.

[13] R. E. Crochiere, "A weighted overlap-add method of short-time fourier analysis/synthesis," Acoustics, Speech and Signal Processing, IEEE Transactions on, vol. 28, no. 1, pp. 99–102, 1980.

[14] L. R. Rabiner and R. W. Schafer, "Theory and application of digital speech processing," Preliminary Edition, 2009.

[15] P. Vaidyanathan, "Quadrature mirror filter banks, m-band extensions and perfect-reconstruction techniques," ASSP Magazine, IEEE, vol. 4, no. 3, pp. 4–20, 1987.

[16] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: A tutorial," Proceedings of the IEEE, vol. 78, no. 1, pp. 56–93, 1990.

[17] J. D. Johnston, S. R. Quackenbush, G. A. Davidson, K. Brandenburg, and J. Herre, "Mpeg audio coding," in Wavelet, subband and block transforms in communications and multimedia, pp. 207–253, Springer, 1999.

[18] J. D. Johnston, "Estimation of perceptual entropy using noise masking criteria," in Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on, pp. 2524–2527, IEEE, 1988.

[19] F. Petitcolas, "Mpeg for matlab, v. 1.2.." http://www.petitcolas.net/fabien/software/mpeg, 2003.

[20] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 975–982, 2010.