Audio Encoding for Heart and Breath Sounds Acquired with Digital Stethoscope

Ignacio Palacios Santos[1], Doroteo Torre Toledano[1] and Andrés Martínez Fernández[2]

[1] ATVS, Escuela Politécnica Superior, Universidad Autónoma de Madrid, Madrid, SPAIN
`ignacio.palacios@estudiante.uam.es, doroteo.torre@uam.es`
[2] Universidad Rey Juan Carlos, Fundación Enlace Hispano Americano de Salud, Fuenlabrada, SPAIN
`andres.martinez@urjc.es`

Abstract. This paper presents a real-time audio encoding system for cardiorespiratory sounds acquired with a low cost Bluetooth digital stethoscope developed by EHAS (Enlace Hispano Americano de Salud) Foundation and Fundatel Foundation. The system is currently working with a G.722 speech coder for compression and subsequent transmission of cardiorespiratory signals. However, these signals have different frequency characteristics from speech. Therefore, we sought a better adapted alternative to encode these signals optimally, with an encoder not subject to the payment of any license, with low computational cost, good quality and low bandwidth. We have evaluated the proposed solution both using objective measures such as root mean squared error and using subjective opinions from four expert physicians.

Keywords. Cardiorespiratory Sounds, Stethoscopy, ADPCM Encoding, Development Cooperation, VoIP.

1   Introduction

EHAS Foundation is a non-profit organization which aims to promote the appropriate use of new Information and Communication Technology (ICT) to improve health processes in isolated rural areas of developing countries. In 2011, a pilot project was conducted in the Peruvian area of River Napo, with the objective of providing cardiorespiratory tele-diagnosis for the local population. The diagnosis is made between rural health posts (without a physician) and health centres of reference (with medical presence). With the audio encoding system, the physician is able to guide the patient via videoconference by means of a digital stethoscope at each end, transmitting the cardiorespiratory audio over VoIP. This solution will allow many isolated people to have a quick diagnosis, sparing them the travelling of hundreds of miles to get to the nearest hospital, going through poorly served areas. EHAS Foundation, together with Fundatel Foundation, developed a low cost bluetooth digital stethoscope and a peer-to-peer transmission system for the cardiorespiratory signals. This paper sets forth a cardiorespiratory real-time audio encoding alternative for digitalized signals that is

able to send these signals over VoIP with good quality, in real-time and using much less bandwidth than previous solutions available at EHAS Foundation. The development of a digital tele-stethoscope is not a new idea in itself [1, 2]. This work does not focus in the development of the digital tele-stethoscope, but in the optimization of the audio coding for the cardiorespiratory signals involved in digital tele-stethoscopy.

2　　Characterization of Cardiorespiratory Signals

Cardiorespiratory sounds are obtained through auscultation which is the term used for the listening of internal sounds of the body, usually by means of a stethoscope. Auscultation is performed for the purpose of examining the circulatory system and respiratory system (heart and breath sounds). There are four auscultation areas: mitral, aortic, tricuspid and pulmonary. The sounds detected in these four areas will have different frequency characteristics than voice signals.
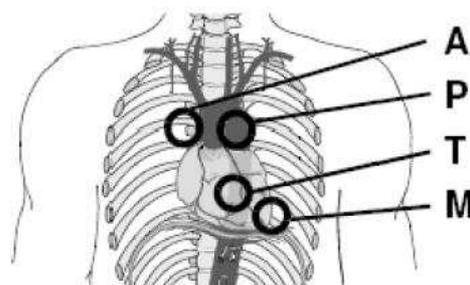

Fig. 1. Auscultation areas [3]

Generally, the components of heartbeat and lung sounds that are useful for diagnostic purposes are in the range of 20-1000 Hz. The first heart sounds (S1 and S2), are produced by the closing of the atrioventricular valves and semilunar valves respectively. These sounds fall in the range of 20-115 Hz. Disorders such as heart murmurs occur in the range of 140-600 Hz [4]. Thus, a suitable listening range for heart sounds would be approximately 20-600 Hz. For respiratory sounds, the strongest part of the signal is typically under 100 Hz, although the signal can have useful components up to 1.2 kHz [5].
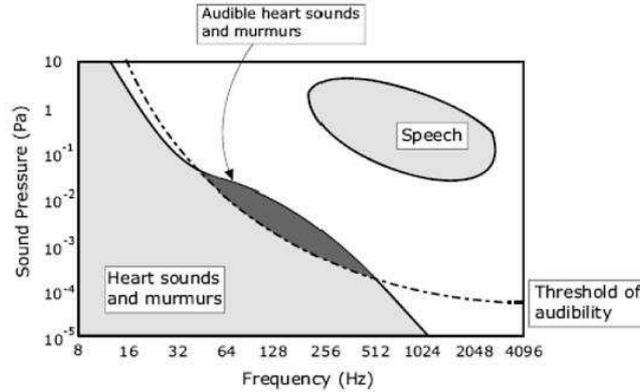
Fig. 2. Cardiorespiratory sound frequency distribution [3]

Hence, by the Nyquist theorem [6]:

$$f_s > 2f_M \qquad (1)$$

we could derive the minimum sampling frequency (f_s) to which we can downsample cardiorespiratory signals without losing relevant information for the diagnosis. In our case the maximum frequency present in the signal (f_M) is 1200 Hz. Hence by Nyquist theorem the cardiorespiratory signas can be completely recovered from samples at a sampling frequency (f_s) higher that 2400 Hz.

3   EHAS-Fundatel System Architecture

As mentioned in [3], the analogical cardiorespiratory sound captured by the chestpiece of the digital Bluetooth stethoscope EHAS-Fundatel is picked up by a microphone, which converts the mechanical motion of the sound pressure waves into a voltage variation. After the analog signal is captured, the following process is applied:

1. This voltage is transmitted, after an analogical amplification stage, to a TLV320AIC33 hardware codec. The codec hardware, at its first stage, performs the conversion of the analog signal emitted by the microphone to a digital signal that can be processed by other electronic devices. This signal is downsampled to 8 kHz and 16 bits per sample.

2. The resulting signal is sent to a OEMSPA310 Bluetooth chip that sends the audio to a PC using the Serial Port Profile (SPP) Bluetooth protocol. This audio is sent encoded in a format that adapts the sound to the features of Bluetooth transmission. During this communication, flow control information is also sent, allowing the PC to ensure it is receiving the audio without any alteration caused by a poor

management of the speed of transmission/reception, as well as the necessary information to ensure the audio can be rebuilt.

3. There is a Stethoscope's software [3] that performs a pre-encoding protocol before sending it to a VoIP application (Ekiga Softphone [7]) when the digital sound reaches the PC. The aim is to support any sound server, running at any sample rate and adjust the sending rate of data at 44100 samples per second and 32 bits per sample (which is the most common case).

4. Once the signal reaches the softphone, the latter adapts the digital audio input to the needs of the audio codec used (G.722 in our example, which requires an entry with 16000 samples per second and 14 bits per sample, providing an output of 64 Kbps).

5. After sending the final digitalized audio to the network, the audio is received by the remote computer, starting the reverse process (decoding).

4   Encoders Test

In order to find an alternative to G.722 codec several simulations of EHAS-Fundatel system architecture were performed with different encoders. We used 33 audio cardiorespiratory samples, 15 of which were obtained with the digital stethoscope's bluetooth EHAS-Fundatel (3 mitral-type, 3 aortic-type, 3 tricuspid-type and 6 pulmonary-type), and 18 downloaded from [8] (10 cardiac-type, 8 pulmonary-type). We created several Linux scripts where the audio samples were encoded and decoded using the audio manipulation software SoX [9] (Sound eXchange), obtaining different figures of merit as results that allow objective comparison of the encoders. These figures of merit are Compression Factor (CF), Bit Rate (R), Root Mean Square Error (RMSE), Percent RMS Difference (PRD) and Real Time Factor (RTF):

$$CF = \frac{source\_data\_size}{encoded\_data\_size} \qquad (2)$$

$$RMSE = \sqrt{\frac{\sum_{n=1}^{N}\left(x[n] - \tilde{x}[n]\right)^2}{N}} \qquad (3)$$

$$PRD = \sqrt{\frac{\sum_{n=1}^{N}\left(x[n] - \tilde{x}[n]\right)^2}{\sum_{n=1}^{N}x^2[n]}} \times 100\% \qquad (4)$$

$$RTF = \frac{computation\_time}{audio\_duration} \qquad (5)$$

Where $x[n]$ is the original signal and $\tilde{x}[n]$ is the signal obtained after coding and decoding the audio, and N is the number of samples in the audio.

The EHAS-Fundatel audio samples are digitalized at 44.1 kHz and 32 bits per sample, while the heart sounds [8] are digitalized at 16 kHz and 16 bits per sample; finally respiratory signals are digitalized at 11.025 kHz and 16 bits per sample. Initially we tested with G.722, G.726, IMA-ADPCM encoders and downsampled at 1, 2 and 4 kHz.

Table 1. Comparison encoders assuming a 16 bit per sample at 16 kHz input.

	CF	R (Kbps)	RMSE	PRD (%)	RTF
G.722	4	64	0.039405267	22.7856303	0.2122
G.726-2	16	16	0.004458867	3.22418553	0.03503
G.726-3	10.66	24	0.002614943	1.93012648	0.03636
G.726-4	8	32	0.002327377	1.71252903	0.03420
G.726-5	6.4	40	0.003868983	2.84869895	0.03633
4 kHz	4	64	0.000347993	0.29957613	0.00448
2 kHz	8	32	0.002195643	1.98224353	0.00446
1 kHz	16	16	0.010389677	5.048618	0.00441
IMA-ADPCM	4	64	0.000396343	1.93784919	0.00613

The G.726 codec worked well, but it requires as input audio previously encoded with G.711 μ-law or A-law. Furthermore, the decoded audio seemed to have worse subjective quality. The G.711 codec is a simple PCM with minimum processing load on the CPU of the PC, and its use is not subject to the payment of any license. However, the quality of the received audio is very dependent on the level of pressure applied to the chestpiece. This codec uses an 8-bit not uniform quantification, so that low amplitude signals suffer less distortion than those with large amplitude. As the cardiorespiratory signal presents large amplitude on low frequency components, this requires auscultation under low pressure levels, preventing signal distortion caused by the encoder working in its less linear region. In addition, working under low chestpiece pressure levels means, in practice, using a smaller number of bits per sample encoded, so the resolution is severely affected.

IMA-ADCPM codec and downsamplings had better results than the G.722 regarding the RMSE, PRD and RTF figures of merit, so we decided to design an integrated

codec consisting of an IMA-ADPCM codec previously downsamped at 4 kHz, thus respecting the bandwidth of cardiorespiratory signals.

Table 2. Comparison between the current codec (G.722) and alternative (Proposed Stethoscope-Codec)

	CF	R (Kbps)	RMSE	PRD (%)	RTF
G.722	4	64	0.039405267	22.7856303	0.02122
Proposed Stethoscope-Codec	16	16	0.002175277	1.93784919	0.00588

The new Proposed Stethoscope-Codec provides better objective results than the G.722 codec when comparing all figures of merit. In particular, provides better objective quality with four times less bit rate and less computational complexity.

5   Medical Evaluation

After performing the objective assessment of the encoders, a subjective evaluation of these was carried out, since it is intended that the encoding of the cardiorespiratory audio does not distort too much the quality of the audio signal and that the decoded audio is useful for remote diagnosis. The 33 original samples and the respective samples decoded with the G.722 codec and Stethoscope-Codec were evaluated by a group of four physicians of the Hospital Clínico San Carlos in Madrid. The doctors were asked to grade from 1 to 10 the quality of the audio heard, both the original and decoded without knowing each sample's encoding. These physicians were three pulmonologist and one internal medicine doctor.

Only one doctor graded, in general, the new Proposed Stethoscope-Codec with a higher grade, stating that there was less noise in the audio heard and that, in such a way, the cardiac and respiratory audio where she needed to focus were highlighted. The rest of the physicians gave, in general, a slightly better grade to the G.722 codec with respect to the new Proposed Stethoscope-Codec. Other doctor explained that in the case of samples related to pathologies where very high-frequency noises were detected, a distortion was observed with such noises taking a deeper register. Nevertheless, being familiar with the pathology, it was possible to diagnose it.

Table 3. Average grades for 33 samples (1 to 15 EHAS-Fundatel samples, 16 to 25 cardiac samples [8], 26 to 33 pulmonary samples [8])

	1	2	3	4	5	6	7	8	9	10	11
Original	7.5	8.75	8.5	7.75	7	7	6.75	8.25	8	7.5	6.75
G.722	7	7.5	7.5	7.25	6.25	6.25	5.5	7.75	7	7	6.5
Stethoscope-Codec	6.5	7.5	7.25	7	6	6	5.75	7	7	6.75	6.5

	12	13	14	15	16	17	18	19	20	21	22
Original	6.25	7.5	**7.25**	8.25	6.5	7.5	8	6	5.75	5.5	6.75
G.722	6.25	7.25	6.75	7.75	6.25	7	7.75	5.75	5	4.75	6.25
Stethoscope-Codec	6	6.75	6.5	7.25	6	6.75	7.25	5.5	5	5	5.5
	23	24	25	26	27	28	29	30	31	32	33
Original	6.75	6.5	5.5	5.25	6.75	7	6.25	6	6.75	5.75	6.25
G.722	6.25	6.25	6	5	5.5	6.25	5.75	5.75	6	5.75	6.25
Stethoscope-Codec	6.5	6.25	6	5.25	5.25	6	5.75	5.75	5.75	5.5	6

Table 4. Average grades per physician

	Original	**G.722**	**Stethoscope-Codec**
Dr. 1	7.33333333	7.42424242	7.90909091
Dr. 2	7	5.75757576	5.42424242
Dr. 3	6.72727273	6.24242424	5.87878788
Dr. 4	6.54545455	6.21212121	5.60606061
Average	**6.90151515**	**6.40909091**	**6.20454545**

6   Conclusion and Future Work

Despite the new Proposed Stethoscope-Codec provides better objective results than G.722 when comparing figures of merit such as RMSE and PRD, subjective medical assessment of the former is slightly lower. The new Proposed Stethoscope-Codec reduces the bit rate in a 4:1 proportion, going from 64 Kbps to 16 Kbps, and decreasing the RTF in a greater than 3:1 proportion, going from 0.02122 to 0.00588. The Proposed Stethoscope-Codec could be a good alternative to G.722 in cases where communication networks are precarious and with small bandwidth. Such is the case of most rural areas in developing countries. Our codec would therefore allow a higher number of auscultations at the same time, reducing the latency and requiring lower-processing-capacity computers thanks to the RTF reduction. The medical assessment presented should be taken with care, since the number of physicians consulted is still small. A more in-depth clinical study with patients suffering of a certain kind of pathology would be needed in order to validate the medical use of the Proposed Stethoscope-Codec.

References

1. Hedayioglu FL, Mattos SS, Moser L, de Lima ME. 2007. Development of a tele-stethoscope and its application in pediatric cardiology. in Indian J Exp Biol. 2007 Jan;45(1):86-92.

2. Sankaran, P. RMK Eng. Coll., Chennai, India Chandrasekaran, K.;Baig, A.H.; Moll, C.L. 2010. Development of a tele-stethoscope: Indian perspective, 2010, in Proc. IEEE/ICME International Conference on Complex Medical Engineering (CME), 2010.

3. Ignacio Foche Pérez: Desarrollo de un teleestetoscopio digital Bluetooth para zonas rurales aisladas de países en desarrollo, Proyecto Fin de Master en Redes de Telecomunicación para Países en Desarrollo, Escuela Técnica Superior de Ingeniería de Recomunicaciones, Universidad Rey Juan Carlos.

4. M. Abella and J. Formolo: Comparsion of the Acoustic Properties of Six Popular Stethoscopes, Department of Internal Medicine, St. John Hospital, Detroit, Michigan 48236.

5. I. Mazic, S. Sovilj and R. Magjarevic: Analysis od Respiratory Sounds in Asthmatic Infants, Measurement Science Review, Vol. 3, Section 2, 2003.

6. Alan V. Oppenhheim, Alans S. Willsky, S. Hamid Nawab: Señales y Sistemas, pp. 527-528, Segunda edición, Pearson Educación.

7. http://www.ekiga.org/, Ekiga Softphone.

8. http://www.med.umich.edu/lrc/psb/heartsounds/index.htm, University of Michigan Heart Sound and Murmur Library.

9. http://sox.sourceforge.net/, Sound eXchange.