Expressive Breath Modeling

Josep M Comajuncosas

MASTER THESIS UPF / 2009
Master in Sound and Music Computing

Master thesis supervisor:
Jordi Janer
Department of Information and Communication Technologies

Universitat Pompeu Fabra, Barcelona

Expressive Breath Modeling

Josep M Comajuncosas Nebot

MASTER THESIS UPF / 2010

DIRECTOR DE LA TESI

Dr. Jordi Janer

UNIVERSITAT POMPEU FABRA

Acknowledgments

I would thank my supervisors Drs. Jordi Janer and Jordi Bonada for their valuable insight and patient guidance throughout the development of this Master Thesis.
My acknowledgments to my colleagues as well for their infinite patience with my endless recordings, spirometries and user evaluations.
Finally, I must give credit to Dan Ellis, whose LPC analysis and synthesis routines, as well as the peak tracking and sinewave resynthesis routines were used for this implementation[1].

[1] Sinewave Speech Analysis/Synthesis in Matlab. Dan Ellis. Available online at http://labrosa.ee.columbia.edu/matlab/sws/

Abstract

This Master Thesis is devoted to the research in normal and pathologic breath noises in order to develop a sound generator which is able to recreate a variety of respiratory sounds and patterns. The provided breathing model consists of a breathing pattern generator, which computes the target airflow, a normal sound generator, which computes the noisy part of the breathing sounds, and a wheeze generator, which computes the continuous adventitious sounds characteristic from some respiratory pathologies.

The synthesis engine uses a single recorded respiratory cycle to model a whole breathing pattern by inferring a mapping from airflow to RMS and resynthesizing the proper frames from the reference sample. Time-interpolated, noise-excited LPC analysis/resynthesis is employed for the noisy part and time-domain based sinusoidal resynthesis is used to model asthmatic wheezes.

A Matlab prototype was implemented and some user tests evaluated the realism and expressiveness of the generated sounds.

Resum

Aquesta Tesi de Màster està dedicada a la recerca dels sons respiratoris normals i patològics per tal de desenvolupar un generador de so que sigui capaç de recrear diversos sons i patrons respiratoris. El model respiratori que es proporciona consisteix en un generador de patrons de respiració, que calcula el flux d'aire objectiu, un generador de so normal, que calcula la part sorollosa dels sons de la respiració, i un generador de xiulets, que computa els sons accidentals continus característics d'algunes patologies respiratòries.

El motor de síntesi empra l’enregistrament d’un sol cicle respiratori per modelar patrons de respiració sencers tot inferint un mapejat del flux d'aire a l’RMS i resintetitzant els *frames* adequats de la mostra de referència. S’empren tècniques d’anàlisi/resíntesi per LPC amb interpolació temporal i excitació per soroll blanc per la part sorollosa i resíntesi sinusoïdal en el domini temporal per modelar els xiulets asmàtics.

S’ha implementat un prototipus en Matlab i s’han realitzat diversos tests d’usuari per avaluar el grau de realisme i expressivitat dels sons generats.

Preface

Today speech synthesizers are able to produce a realistic spoken voice using techniques such as concatenative synthesis. As a further step towards emotional speech synthesis, this research addresses the generation of breathing sounds as a particular case of non-speech human sounds.
The main goal of this project is thus to develop a human breath sound synthesizer.

As expressive synthesis is the main goal of this research, apart from successfully modeling the sounds of human breath it will be investigated whether and what physiological and emotional states could be more successfully imparted with the use of such human breath model.

The interest and social relevance of this field of research is related to the techniques that would make synthetic speech more informative and to the techniques to make man-machine communication more useful.
Breath modeling is a topic that yet remains to be addressed, and its contribution to the aforementioned goals cannot be neglected.
Two different contexts that would greatly benefit from such a parametrizable model of human breathing are voice synthesis for speech and singing models and immersive environments for Virtual Reality and Videogames.

The rest of this Thesis is organized as follows: Chapter 1 summarizes the State of the Art in normal and pathologic breathing sounds, Chapter 2 summarizes the State of the Art in the study of breathing patterns and their relation to physiologic, pathologic and affective states, Chapter 3 presents a new modeling strategy to synthesize realistic breathing sounds and patterns, Chapter 4 presents the Matlab prototype developed as a proof of concept of the aforementioned sound model and Chapter 5 evaluates the model. Finally, Chapter 6 presents my conclusions and avenues for future work.

Index

1. Breathing sounds

Our knowledge of breathing sounds, even if rooted to ancient ages, owes its definitive consolidation as a scientific and therapeutic discipline to the empirical discovery of the relationship between pulmonary diseases and auscultated lung sounds by Laënnec in 1819 [5,11]. He himself invented the stethoscope three years later making it possible to describe the basic categories of lung sounds in a more systematic and objective way than before [6].

With the developments in pulmonary physiology and by extensive clinical measurements over the next century, lung sounds could be better interpreted from a physiological and anatomical perspective. [11]

Nowadays, stethoscope auscultation combined with DSP techniques helps with the evaluations and diagnosis of patients suffering from lung diseases. Thus, computer-aided analysis shows a growing interest by specialists, namely in the development of techniques to detect, remove or isolate adventitious lung sounds such as wheezes or crackles which have a valuable diagnostic usefulness at the early stages of disease [5]

Respiratory system modeling is currently a strong area of interest for its clinical implications, either by deriving formal models of pathological related adventitious sounds or, more recently, by deriving electrical models of the lung for disease detection.

The clinical terminology used to in the literature describe lung sounds was not standardized till the late 70s [6], where the American Thoracic Society established the current classification of lung sounds depicted in table(1), though some of those qualitative distinctions where not shown to be clear indicators of distinct disease categories and considerable disagreement was found among different studies [6]

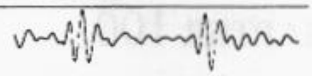
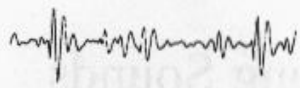
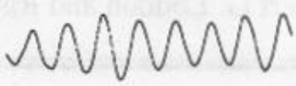
Acoustic Characteristics	*American Thoracic Society Nomenclature*	*Common Synonyms*	*Laënnec's Original Term*	*Time Expanded Waveforms*
Discontinuous, interrupted exposive sounds (loud, low in pitch)	Coarse crackle	Coarse rale	Rale muquex ou gargouillement	
Discontinuous, interrupted explosive sounds (less loud than above and of shorter duration; higher in pitch than coarse crackles or rales)	Fine crackle	Fine rale, crepitation	Rale humide ou crepitation	
Continuous sounds (longer than 250 msec, high-pitched; dominant frequency of 400 Hz or more, a hissing sound)	Wheeze	Sibilant rhonchus	Rale sibilant sec ou sifflement	
Continuous sounds (longer than 250 msec, low-pitched; dominant frequency about 200 Hz or less, a snoring sound)	Rhonchus	Sonorous rhonchus	Rale sec sonore ou ronflement	

Table 1: Classification of common lung sounds [11, itself cited from Murphy RLH, Holford SK. Lung sounds. Basics RD 1980;8:1 - 6.]

Normal lung sounds

It has been proved that the so-called normal breath sounds are produced at the parts of the respiratory tract where there is an aperture change, specifically by the pressure drop caused by the transition from a narrower to a wider radius. This noise turbulence may appear at inspiratory and expiratory flow rates over 0.2L/sec [11]. According to [16], turbulence is originated beyond a critical air velocity of about 10m/s.

This phenomena explains why the expiration phase, where only the glottis passage is narrower than the preceeding sections, is considerably softer than the inspiratory phase, when all the inner airways contribute to the overall breathing noise [16]. Vortex shedding at the vocal folds is thus the main source of noise in voiced sounds [17].
Of similar nature are other noisy components in the human speech such as the aspirated consonants, fricatives, sibilants and noise bursts in plosives [16, 17], which are beyond the scope of this Thesis. They are generated in a similar way, forcing a suitably high particle velocity by constricting a number of supraglottal airways of the vocal tract.
Modern recording and analysis techniques comparing the noise arising from the thorax, chest, trachea and at the mouth, as done by Gabvriely, Druzgalski and coworkers confirm that all the recordings show similar frequency content, with expiratory noises being mainly tracheal and inspiratory noises being closer to the lung surface, apparently originated between the alveoli and the upper airways. [6]
Further researches confirmed that the sound intensity from a breathing recording is a suitable indicator of the patient ventilation, as it is strongly correlated to airflow rates. [6] The respiratory sound itself is a product of both the source characteristics and the radiation through both the airways and the tissues [11]. For example, lung consolidation does improve the transmission of the bronquial breath sounds, because of the different filtering characteristics of the tissue [6].

Adventitious lung sounds

Adventitious lung sounds, either continuous or discontinuous, may appear in the presence of certain cardiorespiratory diseases.
Before going into detail on them, it should be noted that, even if they are a key cue to identify a pathologic breathing from auscultation, the normal noisy components may be also significantly altered under such conditions. According to [20,21], higher frequency components may be observed in the power spectra of auscultated breathing sounds from asthmatic subjects, because of the change in the chest wall transmission characteristics. As shown in fig(1), such pathological breath shows a shift in the frequency components for both the inspiratory and the expiratory phase, as well as changes in the relative spectral densities of both phases.

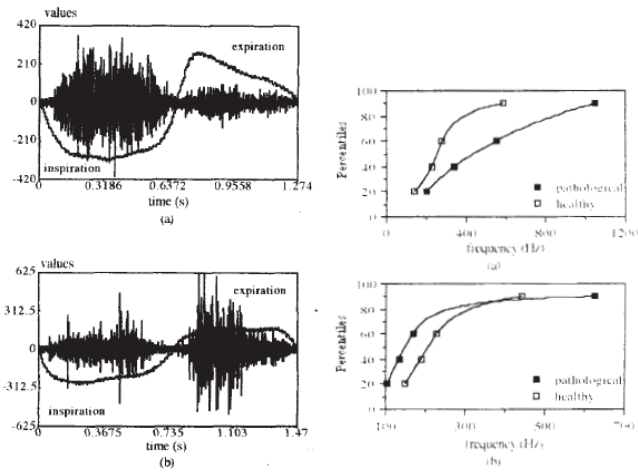


Figure 1: Respiratory sound waveforms with corresponding flow signals, (a) healthy. (b) pathological, and percentiles of PSD for inspiratory and expiratory sounds, from [21]

Wheezes

Wheezes, which are continuous high frequency sounds, are a common symptom for patients having chronic obstructive pulmonary diseases such as asthma or chronic bronchitis, and can be audible even at some distance. Some other diseases may cause a wheezing breath, which can be even produced by healthy subjects by forced expiratory maneuvers [6,18]. Of similar characteristics, but with much lower frequencies and generated in the upper airways, are the sounds known as ronchi.

Their duration is usually superior to 250ms, and several of them can be produced simultaneously, in which case they are known as polyphonic wheezes. They are mostly inharmonic as each single wheeze is caused by a different airway. If a single wheeze is detected, it is usually called a monophonic wheeze, even if it can show some harmonic content [6]. Without medication (bronchodilators), wheezes do have a highest frequency average of 480+-128Hz and may occupy up to 58+-20% of the respiratory cycle [6].

Wheezes are possibly caused by airway flutter mechanisms, that is, by air passing at a sufficient speed through an abnormally narrowed central airway, much as the metal reed from a toy trumpet [6, 9, 11]. Its pitch is mostly related to the flow velocity through the airway and the resonance frequencies of the airway itself, ranging from 400Hz in the upper airway up to 1.7 kHz in bronquial airways, through in the latter case they are of very low amplitude due to tissue absorption [18]. Several formal models related to airflow dynamics have been proposed but are still oversimplifications. The precise frequency of the wheeze could help to determine the airway involved in its production, but much research remains to be done on their physiological origins [6, 9].

Wheeze models

A number of algorithms have been developed to analyze and detect wheezes. Such algorithms are required to be of high sensitivity and independent from respiratory sound power to be of clinical usefulness [9]. The segment of interest to characterize the wheezes is the ending of the expiratory phase in a forced exhalation maneuver, for example in [9] authors concentrate on the exhalation flow interval from 1.2 to 0.2l/s and use mainly traquial recordings, where there is less high frequency absorption by the tissues [9,10].

Peak detection and tracking techniques, employing long windows (51.2ms with 50% overlap) and fine tuned with a number of heuristic rules, are used in the algorithms proposed by Shabtai-Musih, Homs-Corbera and the LAWDA method [9,18]. As a validation method, the authors compare the performance of the proposed algorithms in asthmatic patients and in normal subjects performing forced maneuvers. The methods prove to be more sensitive than auscultation. [9] The detected wheezes can be observed in fig(2)

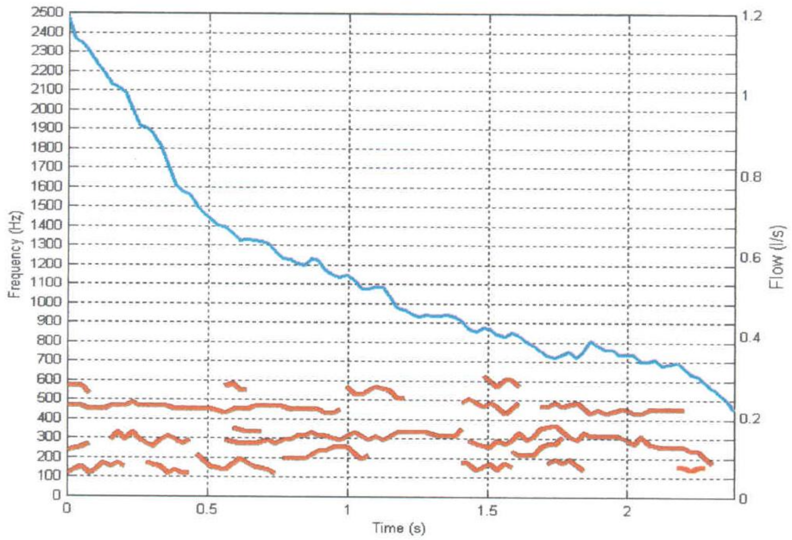


Figure 2: Wheeze indentification in a normal subject (top) and a patient with asthma (bottom) [18]

The mean values for the detected wheezes are summarized in tables(2-3), where significant differences may be observed between asthmatics and control patients, the first showing much more wheezes for a longer period but of lower frequency. Similar results may be observed in a later paper from the same authors, as shown in figure(3).

Variables	Controls Subjects	Asthma Group
Wheezes, No.	2.9 ± 2.0†‡	8.4 ± 6.4†‡
Change in Absolute No. of wheezes	0.5 ± 0.6†‡	2.8 ± 2.2†
Time Without wheezes, %	44.7 ± 40.4	22.1 ± 22.7
Time occupied by monophonic wheezes, %	35.7 ± 29.4	55.9 ± 17.7
Time occupied by polyphonic wheezes, %	15.6 ± 18.3†‡	21.9 ± 15.3†
Mean frequency prebronchodilator, Hz	750.7 ± 175.7†‡	560.9 ± 140.8†‡

Table 2: Number, duration and frequency of tracheal sounds during forced expiratory maneuvers for healthy and asthmatic subjects [18]

	ASTHMATICS	CONTROL
Number of wheezes	17.43 ± 12.64	5.04 ± 4.91
MFWHPP (Hz)	522.32 ±176.76	778.09 ± 312.03
MFWHMP (Hz)	519.00 ± 166.42	746.91 ± 264.91
AMF (Hz)	593.51 ± 177.36	781.42 ± 249.81
Wheezing (%)	88.48 ± 18.17	63.64 ± 37.84

Table 3: Mean and Standard Deviation of number of wheezes, mean frequency of the wheeze with the highest peak power, with the highest mean power, average mean frequency and percentile of maneuver occupied by wheezes [9]

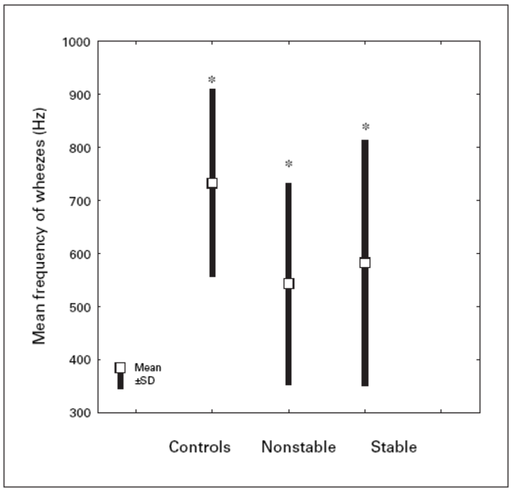


Figure 3: Mean frequency distribution in stable and nonstable asthmatics and healthy subjects during forced expiratory maneuvers. [10]

In [19], a comparison of the possible wheeze-producing mechanisms and their predicted frequencies concludes than the flutter theory may explain well these phenomena, though it is yet far from providing a complete, predictive model of the wheeze frequencies and distribution and the underlying obstructed channels. As an approximation, the critical flutter frequency for a given flutter velocity and airway length is $F = a \cdot \frac{Vf}{Lf}$, were a depends on the channel and fluid properties. For typical values of Lf=5.5cm and Vf=120m/s, we obtain a predicted frequency of 150Hz, well within the range of characteristic wheezes.

Crackles

Of clinical interest, even if much less clearly heard from the mouth at a distance, are the discontinuous sounds known as crackles. Their duration is inferior to 20ms and they have a distinctive explosive sound. They have been observed through auscultation of patients with interstitial lung diseases (fibrosing alveolitis, asbestosis…) or with presence of secretions caused by pulmonary edema or obstructive lung disease, for example. The two possible mechanisms of production of crackles are the sudden opening of closed airways and the bubbles produced by internal secretions [6].

Such discontinuous sounds are usually analyzed and modeled in time domain. The usual parameters to describe them are their quality, their timing relative to the respiratory cycle, their density per breath cycle and their location on the thorax [6], most of them automatically detectable. Monitoring devices might be developed for IC patients [11]. Fig(4) shows a typical sequence of crackles.

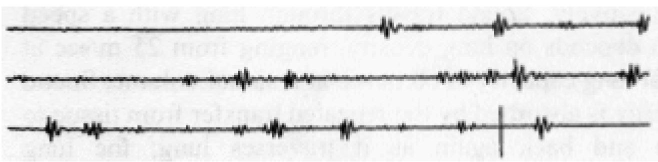


Figure 4: Sequence of coarse crackles recorded from a patient with an acute exacerbation of chronic bronchitis [11]

Measuring parameters

For a single breath cycle, the relevant parameters are:

- Vt, tidal volume. Resting inspiratory capacity.
- Fr, respiratory frequency. Number of inspiratory+respiratory cycles per second
- Ti, inspiratory time. Time from the lowest to the highest point in the volume graph
- Tt, total breath time. The period of a breathing cycle
- Ti/Ttot, duty cycle. The fraction of inspiratory time relative to the total breath cycle.
- Vt/Ti. Mean inspiratory flow.
- V·min. Minute ventilation. $V\min = Vt * f = \frac{Vt}{Ti} * \frac{Ti}{Ttot}$ [7]

Several empyrical relationships can be inferred between these parameters. For example, minute ventilation V.E is roughly proportional to the tidal volume Vt and inspiratory time Ti is slightly inversely proportional to Vt [1]
From these parameters one can carry the so called *pulmonary function tests*, designed to provide medically useful information on the respiratory system performance and to help in the diagnosis of common respiratory pathologies. The most usual measured parameters are vital capacity (VC), forced vital capacity (FVC) and forced expiratory volume (FEV).

Measuring devices

The standard methods to measure the minute ventilation V.e do require a sort of face mask or mouthpiece. Spirometry does provide reliable absolute airflow and volumetric information.
Non-invasive methods such as respiratory inductance plethysmography do indirectly measure airflow through changes of the rib cage and abdomen [7]. Multiple linear regression of spirometric reference measurements are required for proper calibration [1] to obtain reliable measurements.

Spirometry is known to alter the breath measurements as the use of a mouthpiece increases V.e, Vt, Ti and Tt, with no noticeable changes on the duty cycle [1, 7]. Thus it induces a much deeper and slower breathing, especially during quiet breathing, but equally significant for different tasks and task intensities, except at the highest workloads. Also, the body posture when taking the measurements must be taken into account, with Vt being larger in the sitting position than in the supine position [7].

2. Breathing patterns

Table(4) shows the mean values for various breathing pattern components for young, healthy subjects, exemplified in the pattern displayed in fig(5). Notice how inspiratory time is always shorter than expiratory time ($\frac{Ti}{Ttot} < 0.5$). Even if similar for the mean values, the study showed occasional presence of sights, short end-expiratory pauses, fluctuations of tidal volume and occasional apneas in old adults, as seen in fig(6).

	Young
f (breaths/min)	16.7 ± 2.7
V_T (ml)	383 ± 85
$\dot{V}_{min}$ (L/min)	6.02 ± 1.32
T_I (sec)	1.60 ± .30
T_I/T_{TOT}	.424 ± .032
V_T/T_I (ml/sec)	249 ± 54
RC/V_T(%)	42 ± 3
MCA/V_T	1.04 ± .04

Table 4: Pulmonary breathing pattern components in young, normal subjects [7]

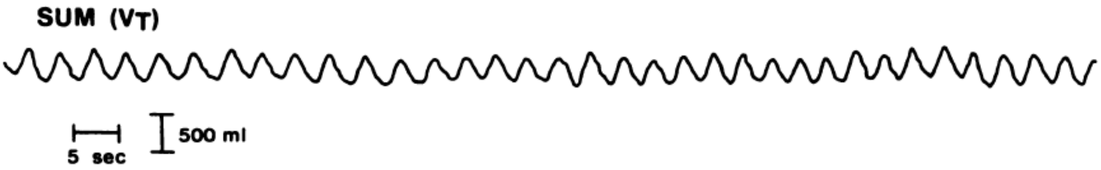


Figure 5: Representative breathing pattern in a young normal adult [7]

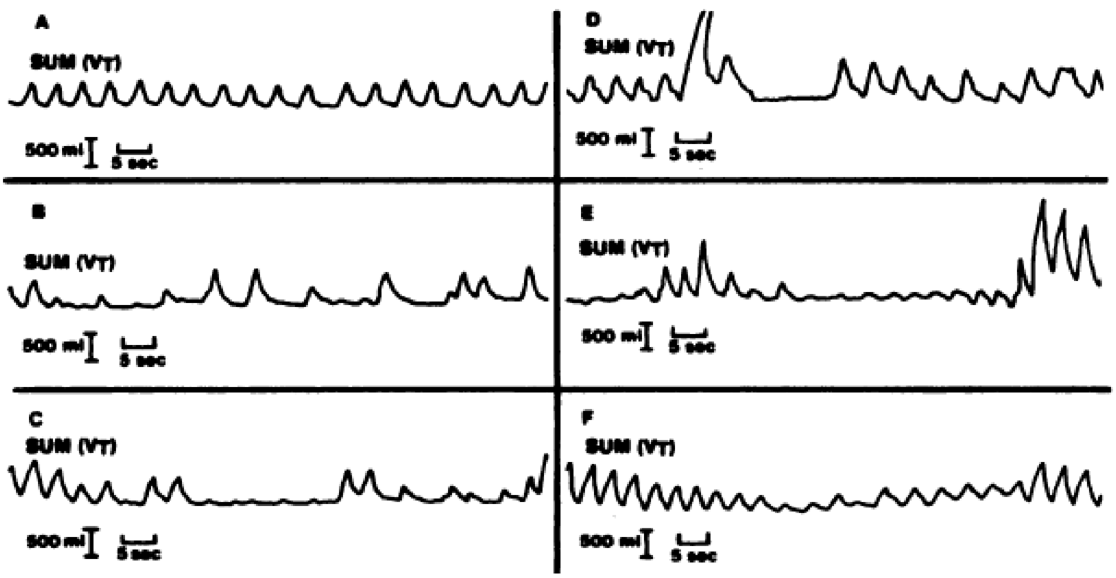


Figure 6: Types of breathing patterns in old normal adults [7]

For healthy subjects, breathing patterns as well as levels of ventilation are strongly dependent on metabolic demands, among a number of cardiorespiratory adjustments. They are also task-dependent, as shown in figure(7), where greater variation of tidal volume and residual capacity (the running offset in the graph) is observed for lifting compared to bicycling [1].

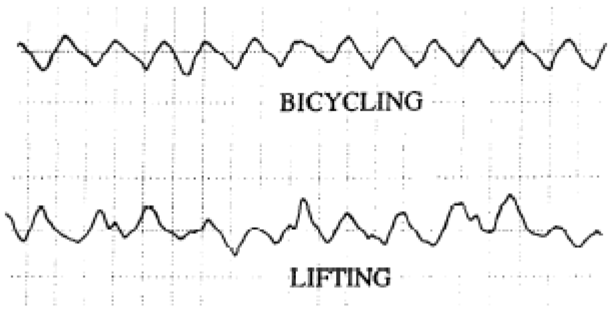


Figure 7: Calculated spirogram during biking and lifting maneuvers.

Healthy subjects

Even for healthy subjects, there is a significant variability related to age and sex. Old people show reduced ventilator efficiency, especially for females, who have smaller lung volumes and exhibit higher respiratory frequency and reduced flow rates (see table(5)). An extensive experiment, carried on a large number of subjects as described in [2] offers a linear prediction model for the main pattern and timing parameters taking into account the level of ventilator stress, age and sex, see table(6).

	Males			Females		
	20–39 yrs	40–59 yrs	60–80 yrs	20–39 yrs	40–59 yrs	60–80 yrs
V_T/IC						
20 L·min^{-1}	0.35±0.10	0.35±0.06	0.38±0.10	0.27±0.09	0.42±0.08	0.45±0.10
40 L·min^{-1}	0.56±0.10	0.54±0.09	0.61±0.11	0.57±0.11	0.57±0.10	0.65±0.10
60 L·min^{-1}	0.68±0.11	0.69±0.09	0.72±0.11	0.66±0.11	0.70±0.10	0.73±0.11
80 L·min^{-1}	0.71±0.11	0.76±0.10	0.75±0.10	0.67±0.09		
f_R breaths·min^{-1}						
20 L·min^{-1}	19±5	19±4	21±4	24±4#	23±5#	25±5#
40 L·min^{-1}	22±5	24±5	26±5	28±4#	31±5#	33±4#
60 L·min^{-1}	28±4	28±4	31±4	37±5#	37±4#	40±3#
80 L·min^{-1}	34±5	34±5	38±5	44±5#		
f_R/V_T breaths·min^{-1}·L^{-1}						
20 L·min^{-1}	18.9±6.1	18.5±6.6	21.6±6.7*	26.6±6.9#	27.3±7.3#	30.3±6.1*,#
40 L·min^{-1}	12.6±5.1	14.3±6.6	16.3±6.2*	20.3±5.8#	22.8±7.3#	26.7±5.8*,#
60 L·min^{-1}	13.1±5.7	13.4±5.8	16.6±6.1*	21.9±6.0#	24.2±6.6#	28.1±4.8*,#
80 L·min^{-1}	14.3±4.3	14.3±4.3	18.3±8.6*	24.9±5.9#		

Table 5: Some pattern and timing of breathing at different levels of ventilation according to sex and age, from [2].Notice the higher respiratory rate (fR) and lower tidal volume (indirectly from proportionally higher fR/Vt) for females.

	Age yrs	Weight kg	Height cm	Sex#	Constant	R^2	95% CI
V_T/IC							
20 L·min^{-1}	0.00107±0.001	-0.00312±0.001			0.55±0.06	0.175	0.17
40 L·min^{-1}	0.00147±0.001	-0.00382±0.001			0.78±0.07	0.177	0.21
60 L·min^{-1}	0.00124±0.001	-0.00297±0.001			0.83±0.08	0.115	0.24
f_R breaths·min^{-1}							
20 L·min^{-1}			-0.177±0.073	-2.364±1.236	52±10	0.183	8
40 L·min^{-1}			-0.233±0.069	-4.074±1.166	67±9	0.358	8
60 L·min^{-1}			-0.252±0.082	-6.071±1.332	77±12	0.420	8
f_R/V_T breaths·min^{-1}·L^{-1}							
20 L·min^{-1}			-0.362±0.140	-4.359±2.376	84.9±22.1	0.189	19.7
40 L·min^{-1}			-0.376±0.093	-4.782±1.560	82.4±14.6	0.358	12.9
60 L·min^{-1}			-0.326±0.094	-6.452±1.530	75.6±14.9	0.421	12.4
T_I s							
20 L·min^{-1}				0.33±0.08	1.19±0.06	0.117	0.42
40 L·min^{-1}				0.28±0.04	0.96±0.03	0.251	0.26
60 L·min^{-1}				0.25±0.04	0.77±0.02	0.283	0.20
T_E s							
20 L·min^{-1}				0.32±0.08	1.46±0.06	0.119	0.45
40 L·min^{-1}				0.32±0.05	1.12±0.04	0.241	0.28
60 L·min^{-1}				0.28±0.04	0.89±0.03	0.291	0.21

Data are presented as mean±SEM. CI: two-sided confidence limits; V_T: tidal volume; IC: inspiratory capacity; f_R: respiratory frequency; T_I: inspiratory time; T_E: expiratory time. #: males 1, females 0.

Table 6: Linear Prediction equations for patter and timing of breathing at different levels of ventilator stress according to age and sex. From [2].

Diseased subjects

For subjects with respiratory diseases the breathing patterns can be significantly altered [8]. As seen in table(7), diagnostic discrimination may be provided by analyzing the distinct breathing patterns for several pathologies, even for smokers. For example, patients with COPD do show an increased and more fluctuant respiratory rate, compared to mean values from table(4).

			Asthmatics		COPD				
	Normals	Smokers	Asymptomatic	Symptomatic	Nonhypercapnic	Hypercapnic	Restrictive Lung Disease	Pulmonary Hypertension	Chronic Anxiety
f (breaths/min)	16.6± 2.8	18.3± 3.0†	16.6± 3.4	16.0± 4.1	20.4± 4.1§	23.3± 3.3§	27.9± 7.9§	25.1± 6.4§	18.3± 2.8
V_T(ml)	383± 91	484± 157§	386± 133	679± 275§	447± 139†	476± 158‡	395± 70	431± 106	403± 133
V_{min} (L/min)	6.01±1.39	8.18±2.54§	6.07±1.39	9.45±2.50§	8.59±2.92§	10.14±2.96§	10.28±1.90§	10.14±2.90§	6.65±1.84
T_I (sec)	1.62± .31	1.44± .27†	1.62± .46	1.55± .35	1.12± .25§	1.10± .20§	.96± .24§	.96± .15§	1.44± .24
T_I/T_{TOT}	.421±.033	.397±.033§	.416±.023	.371±.043§	.346±.044§	.354±.037§	.409±.020	.383±.025	.397±.041†
V_T/T_I (ml/sec)	250± 58	345± 95§	244± 54	450± 135§	483± 117	479± 132§	434± 74§	461± 126§	310± 135†
RC/V_T (%)	42± 17	48± 13	55± 15†	51± 15	50± 20	52± 29	38± 14	49± 15	43± 13
MCA/V_T	1.05± .04	1.05± .02	1.02± .02†	1.15± .18§	1.14± .13§	1.11± .12†	1.04± .04	1.07± .06	1.04± .04

Table 7: Breathing pattern components, diseased subjects [8]

Recordings of breathing patterns for symptomatic asthmatics show an increased Vt and a rapid onset of volume, as seen in fig(8). A more significative alteration of the breathing patterns can be seen in fig(9), recorded from a patient with chronic anxiety. Rynthmic irregularities, episodic dyspnea, frequent sighing and alternation of shallow breathing with long apneic pauses can be observed.

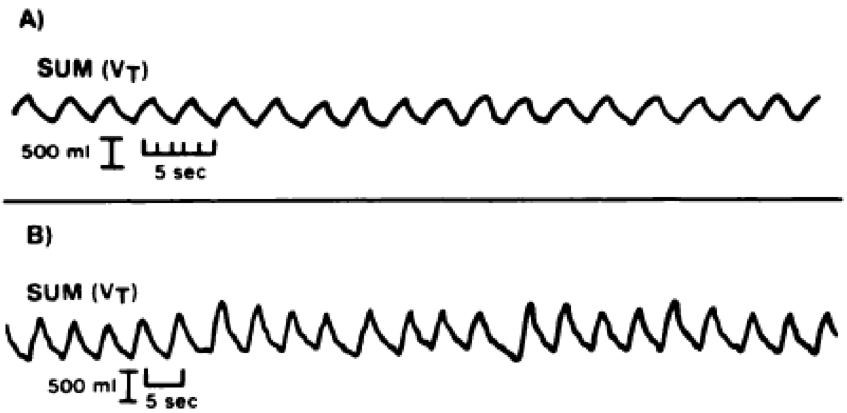


Figure 8: Breathing patterns in two sympthomatic asthmatic patients [8]

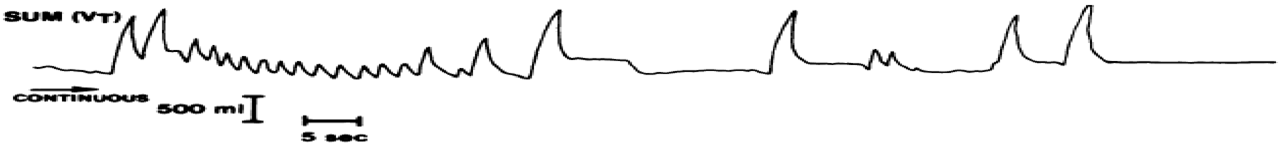


Figure 9: Breathing pattern in a patient with chronic anxiety [8]

Affective responses

Respiratory responses (both breathing cycle and breathing patterns) are also affected by emotional states. According to a number of researches [3][4], high-arousal induced states by visualizing film scenes do shorten the expiratory time and enlarge the inspiratory duty-cycle, the mean expiratory flow and the minute ventilation. As another example of emotionally altered respiratory functions, fig(10) shows a clear increase in the respiratory frequency for anticipatory anxiety states.

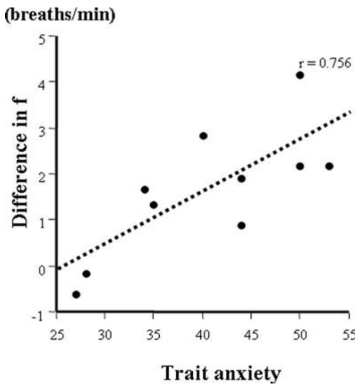


Figure 10: Relationship between respiratory frequencies during anticipatory anxiety [12]

According to [3], no significant changes were caused by the emotional valence (pleasant/unpleasantness), but in a more recent research [4] clear effects were observed for emotions such as disgust and amusement, thought it is challenging to isolate distinct induced emotions in a conducted experiment and there is a large individual variability [12]. For the amusement the breath was showed to have faster but more superficial inspirations, whereas in the case of disgust showed an increase in breath-holding (inspiratory) pauses. An increase in the variability of the main breathing pattern parameters is clearly observed in tidal volume measurements (see table(8)) for amusing stated, though they are most possibly caused by sudden bursts of laughter. The recorded pattern for disgust clearly exemplifies the aforementioned inspiratory pauses, which the authors relate to an unconscious attempt to avoid inhalation of noxious gases, or suppress nausea (see fig(11)):.

Respiratory measures	Pretask base-line	Neutral	Vomiting	Torture	Funny
T_i (s)	1.30 (0.21)	1.32 (0.27)	1.24 (0.13)	1.32 (0.27)	1.19 (0.16)
T_e (s)	2.18 (0.41)	2.14 (0.45)	2.10 (0.29)	2.13 (0.39)	1.95 (0.32)
P_i (s)	0.01 (0.01)	0.01 (0.02)	0.13 (0.22)	0.02 (0.02)	0.01 (0.04)
P_e (s)	0.28 (0.28)	0.38 (0.31)	0.35 (0.28)	0.28 (0.22)	0.38 (0.22)
T_{tot} (s)	3.77 (0.64)	3.85 (0.76)	3.82 (0.52)	3.74 (0.63)	3.53 (0.50)
T_i/T_{tot} (%)	35.2 (3.2)	35.1 (2.8)	33.7 (3.6)	35.6 (2.8)	34.6 (3.1)
R_c/V_t (%)	49.1 (14.1)	46.4 (14.2)	48.5 (13.9)	49.3 (13.3)	48.4 (15.1)
V_t (ml)	669 (131)	659 (171)	615(108)	637(158)	562(105)
V_{min} (l/min)	11.62 (2.35)	10.52 (2.10)	9.95 (1.82)	10.37 (1.56)	9.89 (2.02)
V_t/T_i (ml/s)	527 (99.7)	506 (97)	503 (101)	488 (77)	482 (102)

Table 8: Means and standard deviations of respiratory parameters by selected conditions [4]

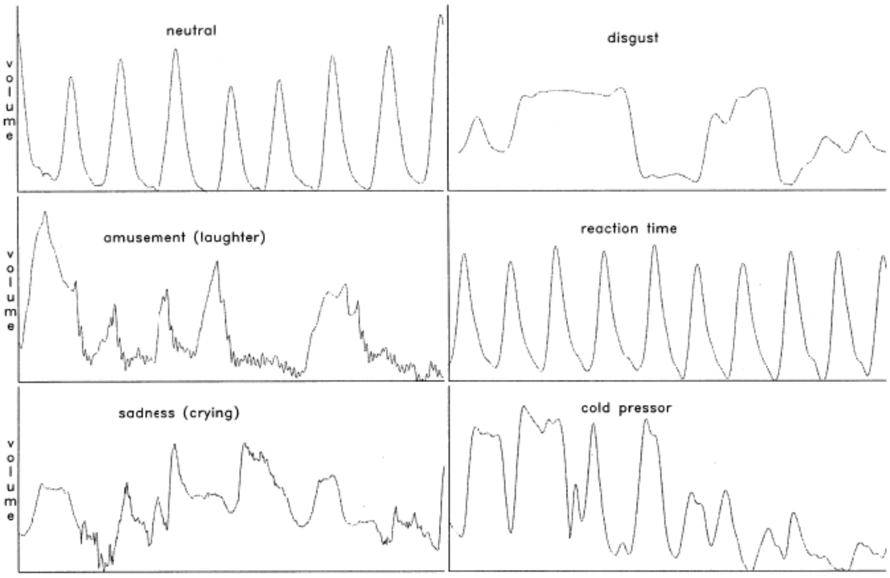


Figure 11: Task related breathing patterns in a normal young adult [12]

3. Modeling strategy

The noisy components of the speech, including breath sounds, are usually modeled as filtered noise, either in the spectral domain or in the time domain.
In spectral modeling synthesis [13,14], subtracting the deterministic part from the magnitude spectra of the original signal yields a noisy residual which is considered to be of stochastic nature, and thus can be modeled by piecewise approximating or curve-fitting the magnitude envelope of white noise. LPC is considered in [14] to be less flexible and more prone to numerical instabilities but simpler and more compact than the spectral enveloping technique. The noise component itself is generated by randomizing the phase spectrum. Other authors do employ LPC to model the residual [15].
In complete articulatory models such as the one described in [16] the breath noise is also incorporated by adding Gaussian white noise to the system, indirectly low pass filtered to about 6 kHz by the integration methods employed. In [17], a number of noise generation theories are reviewed and a custom model which depends both on the flow parameters and the geometry changes of the glottis is incorporated into a noise module as part of a complete articulatory model.

This Thesis will focus on LPC-based resynthesis techniques to model the normal respiratory sounds and time-domain sinusoidal modeling of the pathologically-related adventitious continuous sounds. Linear Predictive Coding is widely used to encode the speech for bandwidth reduction in telephone communications. In such a context, the emitting device encodes the time-varying, all-pole filter coefficients which model the spectral envelope of the voice as well as the speaker pitch and a voiced/unvoiced descriptor, while the receiver resynthesizes the speech by exciting an inverse filter with either a pulse waveform or a glottal impulse at the proper frequency for the voiced passages or white noise for the unvoiced ones. It has also been used by some Computer Music composers, such as Paul Lansky and Charles Dodge, as a kind of subtractive synthesizer, even for cross synthesis, by imparting the spectral envelope from an analyzed sound to another sound which replaces the usual residual or noise, much as with a vocoder.

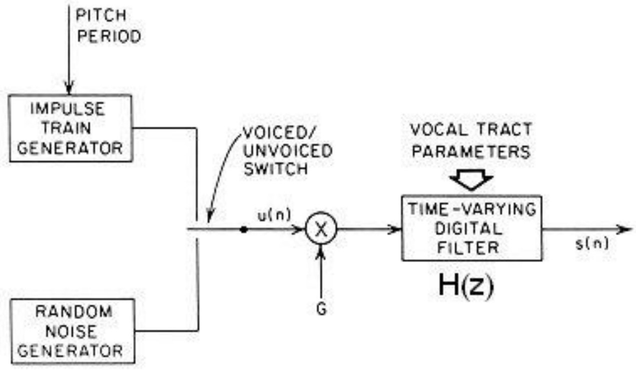


Figure 12: usual scheme of a LPC resynthesis engine. The inverse prediction filter is fed either by a train impulse or by noise to regenerate the analized voice.

The most relevant difference between this model and the usual voice synthesis approach is the fact that in a breathing model the noisy part is the predominant part of the sound output, while sinusoidal components are incorporated to this basis to model some additional pathological cues. In a voice synthesis algorithm the sinusoidal part, related to vocal fold vibrations, is considered to be the main sonic component, while the noisy components (air breath noise, fricatives) are modeled as filtered residuals.

Two basic components need to be modeled: the breathing sound and the breath dynamics. The breath dynamics will take care of the mechanisms which drive the breath synthesizer, and will be responsible of the generation of breathing patterns.
The components should be designed to take advantage of the wide availability of spirometric measurements and stethoscope recordings in the medical literature, as it is the most accurate and detailed data available.
Spirometric measurements are usually provided as airflow or tidal volume graphs, which are related by the expression $Vol = \int Flow \cdot dt$
thus, designing a driving component which sends flow or volume data to the sound synthesizer would allow us to reuse available spirometric measures to directly model different breathing patterns, and would behave as a sort of diafragm model that controls our breathing engine.

Another possibility could be do implement a dynamic model of the respiratory mechanics [23,24], which do require intensive parameterization and was discarded because of its complexity. Its main advantage would be a more transparent integration to animated 3D models as we would deal with more physically meaningful parameters.

Modeling normal breath sounds

Most available breath sounds come from auscultation recordings.
The basic noise components of breath sounds can be well characterized as filtered noise with a decreasing exponential slope, at least for healthy subjects [19]. However, stethoscope recordings, except in the case of lung consolidation of in traqueal recordings, suffer from the poor sound conductivity of tissues, with a clear lowpass filtering effect [11].
The main components of auscultated lung sounds are within the 40Hz-1.6 kHz range. Below 40Hz heart sounds and muscle friction noise mask the breath and beyond 2 kHz there is hardly components for the lung sounds, as they are mostly absorbed through the surrounding tissues [21]. For traqueal sounds and at high flow rates, frequencies can extend beyond 2 kHz and, in the case of breathing sounds recorded with a microphone at the mouth, they will incorporate the vocal tract resonances as well.

Experimental setup

Simultaneous spirometric and audio measures were taken to six adult subjects breathing at rest, to evaluate the relationship between the measured flowrate and the breath noise (fig(13)).

Figure 13: Experimental setup. Edirol with microphone in front of a Vernier educational USB spirometer (direct acquisition rates of up to 200Hz)

The airflow data provided by the spirometer shows, as expected, a strong correlation with the energy content from the recorded sound file, though unfortunately, rms is not robust to noise and recording conditions. Fig(14) displays the sensor data together with the framed RMS of the soundfile.

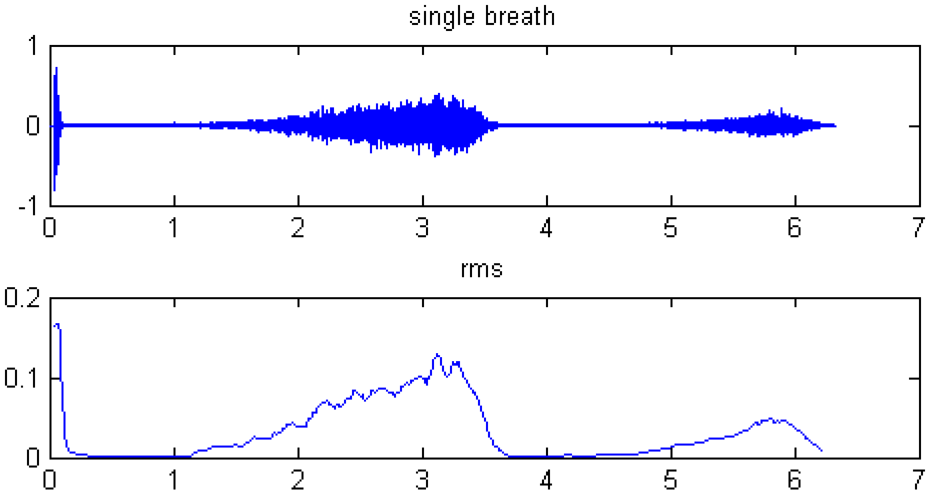


Figure 14: waveform and RMS for a whole respiratory cycle. The initial onset was later used for synchronization purposes.

The next fig(15) shows the dependence between the measured flow and the RMS of the recording, for a single breathing cycle (left) and for multiple breathing cycles (right).

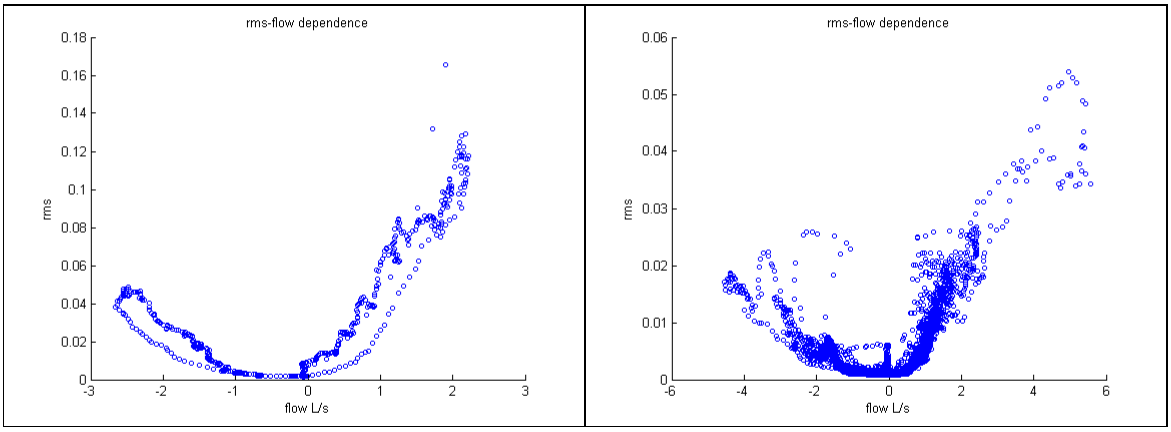


Figure 15: Airflow rate vs. RMS for a single (left) and a number of respiratory cycles (right)

This dependence between a waveform descriptor and the measured flow rate suggests a practical way to map spirometric measures to audio: once the desired flow rate is established, we just map it to RMS and find an audio frame with the desired energy. A recording of a single breathing cycle, ideally a slow, deep breath and with a somewhat linear airflow slope to keep good resolution for all possible RMS values would then be enough to recreate arbitrary breathing patterns.

Of course this method is oversimplified: It has been shown that for the same airflow, different lung volumes may cause different sound intensity levels [11]. Moreover, even if properly calibrated, each subject and vocal tract configuration will exhibit a particular mapping.

Residual

Only resynthesis from the residual is guaranteed to give nearly perfect resynthesis, the differences being caused by the artifacts introduced by the overlap and add process. Employing white noise assumes that the LPC perfectly "whitens" the residual, or conversely, that it completely models the spectral shape of the breathing noise. As this is not the case, resynthesising with white noise gives a less uniform, less detailed resynthesis even for breathing sounds. Notice for example in fig(16) how, apart from the global amplitude differences, the shape of the noise excited resynthesis doesn't follow accurately the amplitude envelope from the original waveform.

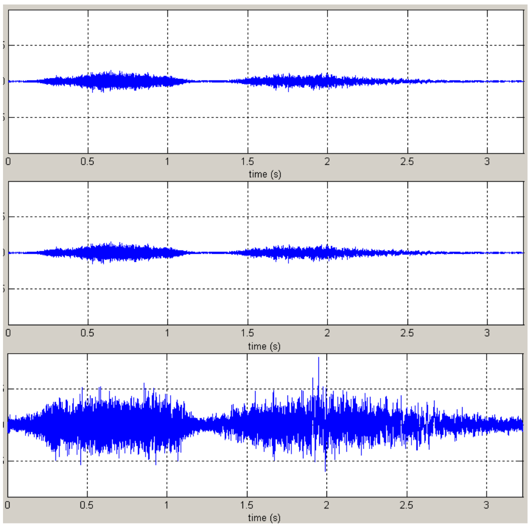


Figure 16: Top: original waveform. Middle: residual based resynthesis. Down: noise based resynthesis. Generated with the Matlab SSUM[2] package.

Recreating the breathing part of the voice with noise is a requirement when LPC is used to save bandwidth in mobile communications, where only the LPC coefficients are sent to the receiver, but we are no longer constricted to this requirement in our context.
Unfortunately, using the residual is challenging when time stretching and free frame indexing is applied as in our synthesis approach: if two or more successive frames are the same, pitch periodicities will appear, and for frame shifts smaller than the overlap index, noticeable shifting in the pitch are generated. Several strategies were tested to minimize those artifacts, namely replacing the residual with white noise when the breath was kept steady and modifying the breathing pattern by avoiding totally stable passages and none of them provided the desired quality.

[2] SSUM: Signals and Systems Using MATLAB. Bob L. Sturm and Dr. Jerry Gibson Graduate Media Arts & Technology Program (MAT) University of California, Santa Barbara. Available online at http://www.mat.ucsb.edu/~b.sturm/SSUM/docs/SSUM.htm

For that reason Gaussian white noise will be used to resynthesize the breathing, despite its lack of accuracy.

Synchrony

To precisely map the recording and the spirometer measures a sort of synchronization must be set up. As a first approximation it may suffice just asking the subjects to spell a short [k] before starting the breathing measures, as the plosive will be nearly simultaneously captured both by the pressure sensor as an positive onset, and by the microphone as a low frequency impulse. Some further manual alignment will be required after the measurements.

Spirometer response

As an unavoidable side-effect of recording through the spirometer, the breathing noise will incorporate the comb-like resonances of the cylindrical device, plus the resonances due to the bacterial filter and the sensor cavities.
To compensate by the device response we can measure it and apply an inverse filtering to the recording. To determine its frequency response a 60Hz-4 kHz sine sweep was recorded with and without the mouthpiece+spirometer attached in front of the microphone, at the precise position of the recording microphone used later on. The resulting inverse filter is displayed in fig(17)

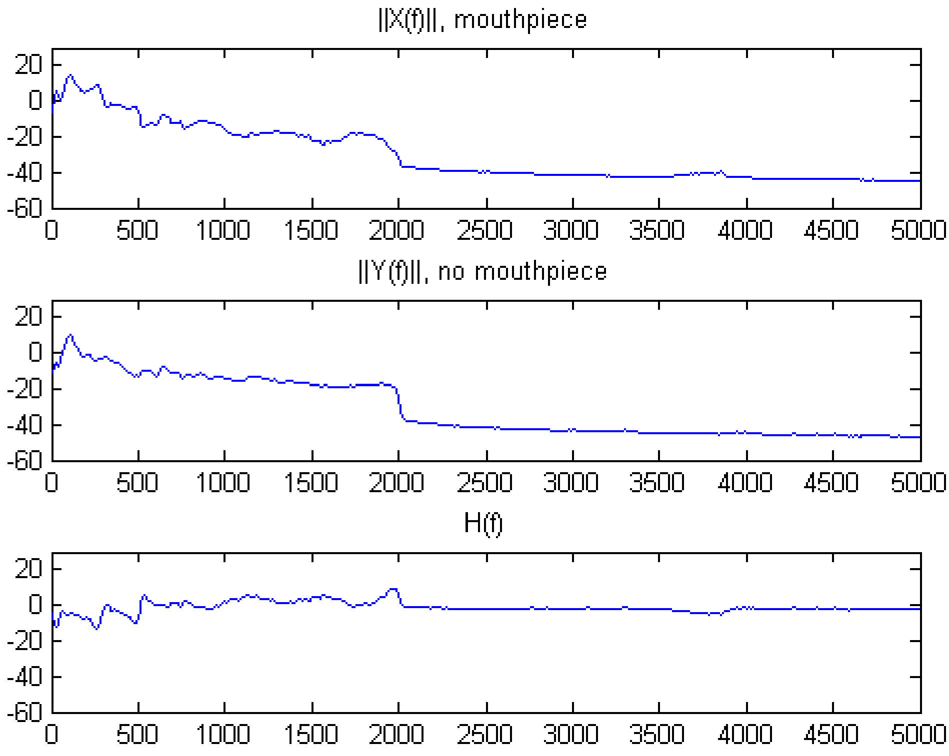


Figure 17: inverse filtering to compensate by the spirometer response. Top: spectrum of breath noise at the spirometer ending, Middle: spectrum of breath noise without spirometer, Bottom: transfer function to apply to spirometric recordings.

The mean magnitude of the spectral difference between both recordings is displayed in fig(). This transfer function will attenuate strong bore resonances at around 259 and 475 Hz. Deconvolving the recorded breathings with the measured frequency response of the mouthpiece + sensor system do noticeably improve the naturalness of the recordings, though the results are far from being optimal.

Flow to RMS mapping

Once the reference breath is ready, we may perform the mapping to an arbitrary measured flow pattern. All the process is displayed in fig(18-19):

- From the tidal volume pattern to be modeled, we obtain the flow rate by integration
- The target flow rate is mapped to a target RMS for each audio frame
- By searching to the closest RMS amplitude, a suitable LPC frame from the reference inspiratory/expiratory cycle is assigned for the target frame, performing inter-frame interpolation for fractional frame locations, by interpolating linear spectrum frequencies which guarantee the stability of the LPC filter coefficients.

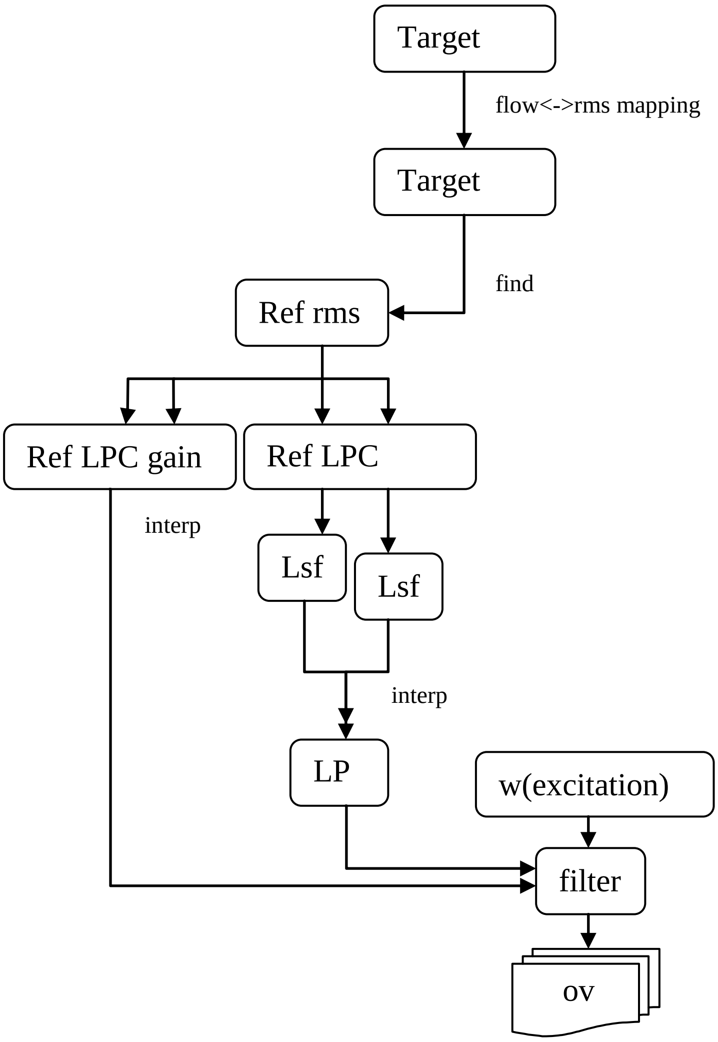


Figure 18: scheme of the synthesis model, *without* the wheeze generator.

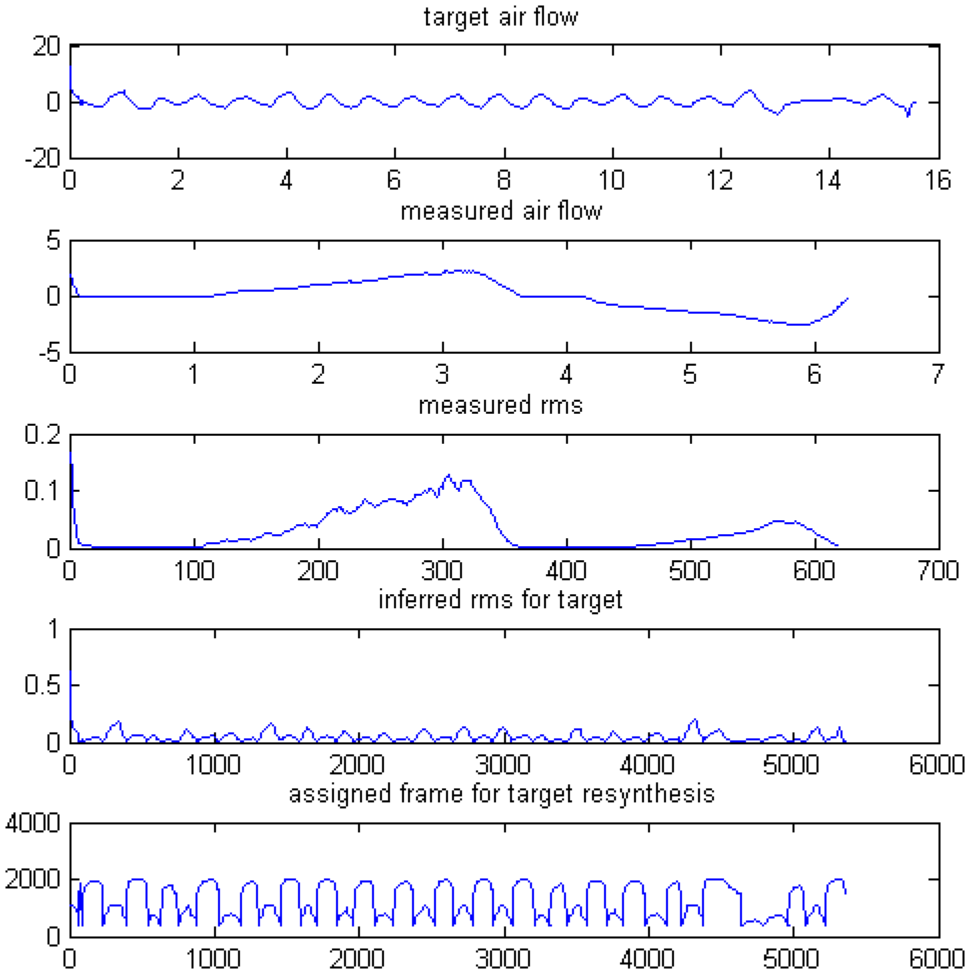


Figure 19: flow to RMS mapping process. From top to down: desired flow rate for the target pattern, reference flow rate in the recorded respiratory cycle, computed RMS for the recorded respiratory cycle, inferred RMS for the target pattern and frames chosen from the recorded respiratory cycle to emulate the desired flow rate.

Direct flow rate estimation

This audio + spirometer capture method has some disadvantages, though. The recorded breath has an unavoidably tubular characteristic, an unnatural presence of low frequencies and some high frequency ringing caused by the inner discontinuities of the discrete parts of the device which don't fit exactly. The inverse filtering technique described above cannot completely remove those artifacts.
Another disadvantage is the fact that the users cannot breath in different mouth apertures, which could improve the naturalness and add some additional clues to convey physiological or emotional states.

In order to avoid these limitations, it should be possible to directly estimate flow rates just from audio recordings without spirometric measures, given some known, reference values. Even if the accuracy will be much lower, under careful control of the subject training and the recording conditions the mapping could be acceptable. This method has already been employed by some researchers, but employing pneumotachometers and a visual feedback to the users [20].

Five users wearing noise clips were asked to keep a steady peak flow at a number of target levels in increments of 0.5L/s, both for inspiratory and expiratory phases. For each level, and after some training, the spirometer is removed and the users were asked to keep the same flow rate without the mouthpiece. Thus it was a failed experiment as without feedback it was not possible for any user to keep the desired flow rate with acceptable accuracy.

The lack of absolute measurements is unfortunate and makes a precise mapping an unfeasible task, but knowing from the previous spirometric measures that there is a rather linear relationship between flow rate and RMS, a software that infers a linear mapping between some usual minimum and maximum flow values and allowing the user to modify the default mapping manually could give acceptable results. Assuming that no breathing sounds are produced for flow rates smaller than 0.2 or 0.1L/s and that 2L/s gives rather high ventilation a simple linear mapping could infer the flow from the recorded RMS

$$flow = 0.2 + (2 - 0.2)\frac{rms - rms\_\min}{rms\_\max - rms\_\min}$$

Using a single recorded respiratory cycle from a subject, as before starting as softer as possible and gradually increasing the flow rate to obtain a rather linear RMS increment spanning a flow rate range as large as possible, should be enough to recreate any breathing and pattern.

Modeling pathological breath sounds

This Master Thesis will concentrate on asthmatic breath modeling, and specifically on time-domain wheeze modeling.

Wheeze extraction and analysis

To model asthmatic wheezes a number of recordings, either stethoscope recordings from the literature and from some medical training software and audio CDs [25, 26, 27], and a custom recording on an asthmatic subject where employed. The latter was indented to provide some insight into the relationship between wheezes and flow rate and served as an aural reference for the model.
From the asthmatic stethoscope recordings, several empirical rules where inferred as a guide to model the wheezes:

- Tracheal recordings are better suited for peak extraction, as the low-pass characteristic of the lung and surrounding tissues muffles some high-frequency wheezes if auscultation is done on the chest or back.
- Wheezes are more prevalent (longer and clearer) in expiratory passages, but also present in inspiratory passages, where they are more complex and irregular.
- Wheezes in expiratory passages do start a while after the expiratory onset, as soon as the airflow decreases enough to cause airway flutter.
- Wheezes may be mostly monophonic (ie. a single peak) or moderately polyphonic. Much denser polyphonic textures are only produced with very

collapsible airways (patients with chronic obstructive lung disease or emphysema).

- The pitch frequencies and trajectories of the wheezes show large variations, but for a given subject, the wheezes from consecutive respiratory cycles are quite uniform.
- The more severe stages of asthma, the more random the wheeze frequencies are.
- Even if the duty cycle is not significantly altered even for moderate asthmatics, the expiration is perceived to be longer because the wheeze is prolonged beyond the breath sound. They can even overlap with the next respiratory phase in case of more severe occlusion. This known phenomena is called *expirium*.[26]

Simultaneous flow rate measurements and sound recordings were made to a symptomatic asthmatic subject, without taking bronchodilators for at least a whole day, after having some exercise and lying on a table during the recording. To obtain distinct wheezes slightly forced respiratory maneuvers proved to be necessary, causing higher frequency wheezes to arise. Also, the spirometer clearly interfered with the measurements, as the greater ventilation caused by breathing through it made it more difficult for the subject to control the optimal flow rates to produce good results.
Once wheezes were produced, they could be clearly heard both in the inspiratory and expiratory phases, as may be observed in fig(20) right. The wheeze behavior seems to be in good agreement to monophonic wheezes available in the medical literature, as the one shown in fig(20) left

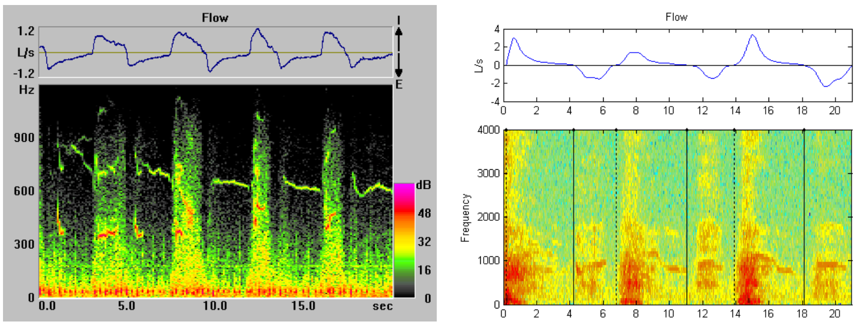


Figure 20: Respirosonograms from a 17 year old boy with acute asthma (R.A.L.E) recorded by an stethoscope over the right anterior upper chest, and an asthmatic 36 year old man after some exercise (own), recorded by microphone at the mouth. Note that the flow is inverted in the first figure.

A peak extraction and tracking algorithm on the frequency domain may be used to isolate the wheezes, carry on some statistics and resynthesize them alone by a sinewave oscillator bank. The main rules for the peak tracker to work properly would be limiting the trajectories to the usual range of wheezes (100Hz-1.2kHz) and requiring a sufficient duration for them (250ms or more). Incidentally, most wheezes could be somehow removed from the audio file itself to provide only the normal breath sound as a residual, as shown in fig(21).

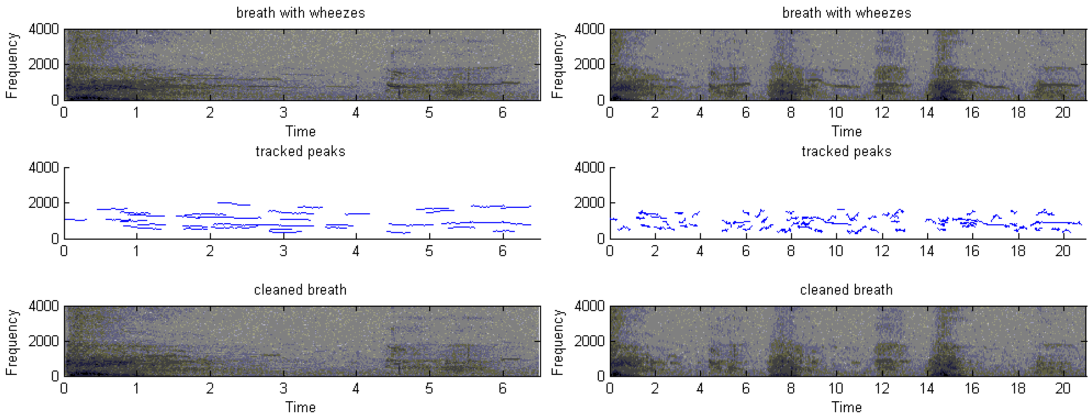


Figure 21: Wheeze tracking and removal for two asthmatic breath recordings

Time domain wheeze modeling

Deriving a parametric model for all possible wheezes is beyond the scope of this work. Instead, a perceptual model able to imitate the wheezes produced by different levels of severity will be implemented.

Wheezes will not be really modeled but resynthesized after some preprocessing. As wheezes where not found to be simply flow-dependent, they are by now the only part of the synthesis engine which not depends on flow, making its generation currently unsuitable for real-time applications.

To synthesize wheezes, as their number is kept rather low, a time domain approach is efficient enough. Once we have a matrix of pitch and amplitude trajectories for all the detected wheezes, the engine will select the wheezes from a corresponding respiratory phase, the user will have the ability to select just some (to simulate a quasi monophonic wheeze sound, or restricting wheezes only in the expiratory phase for a less severe pathologic sound), eventually modifying them (lowering their frequencies to emulate a less forced breathing, or increasing them to make the sound more tense), and they will be time-stretched to fit within the duration of the desired pattern. The wheeze data must include both audio and spirometry for reliable onset and phase detection, and a number of respiratory cycles should be included to add some randomness to the resynthesis.

For a more refined method one should take into account that asthmatic breathing patterns differ from healthy subjects, like sharper inspiratory onsets in case of severe obstruction, yet no attempt was made to implement it.

Pattern modeling

Once the main respiratory parameters for a given physiological or emotional state are known, either from the literature or from custom airflow measurements, a breathing pattern may be reconstructed from them by statistical modeling.

Mean and standard deviation values for Vt, the inspiratory time Ti and the inspiratory phase to respiratory cycle ratio Ti/Ttot may suffice for a single Gaussian model of the pattern. At each new cycle, new values for those parameters are generated according to their distribution, the actual airflow is obtained by differentiating the tidal volume and a suitable airflow shape is obtained.

As an exemplification, pattern generation and interpolation from a high level to a low level of ventilatory stress will be implemented. To know the real evolution of the main pattern parameters over time, which was not found in the literature, a 5 minute spirometry with a coarse resolution of 0.01sec was done to a subject after intensive exercise. Measurements of the evolution of the respiratory frequency Fr, the peak tidal volume Vt and the inspiratory/expiratory ratio Ti/Te are displayed in fig(22). Compared to the data provided by the medical literature, both the tidal volume and the inspiratory to expiratory ratio seemed much more stable in my experiment, and it didn't help to determine any relationship between the evolution of all the respiratory parameters.

Because of this lack of evidence of any relationship between the respiratory parameters evolution beyond the data provide by the literature (for the prototype data from tables(5,6) was used), simple linear interpolation between some target sets of parameters will be used in the pattern modeler.

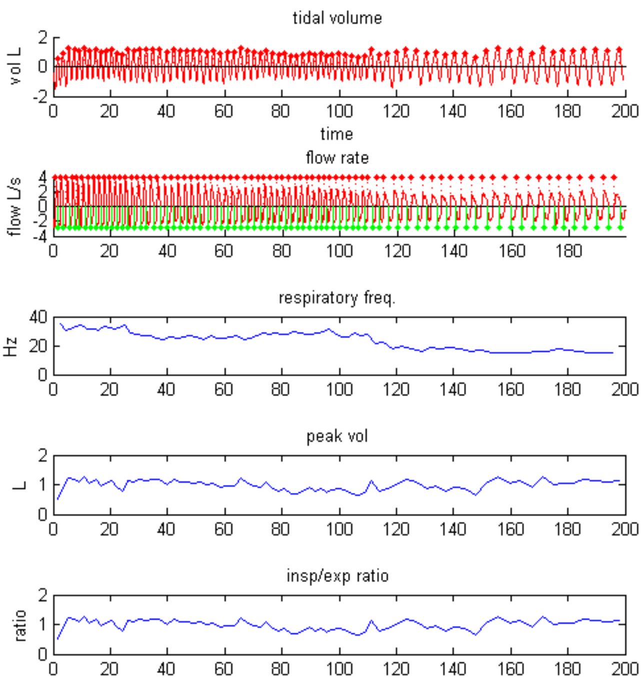


Figure 22: evolution of the main respiratory parameters for a long breathing pattern measured on a subject recovering after intensive exercise.

The tidal volume and airflow evolution within a respiratory cycle is usually analyzed with global statistical descriptors, but for our modeling purposes a better knowledge of its detailed temporal evolution might seem necessary. Unfortunately, the variability is too large to allow us to define something like a *prototypical respiratory cycle*. Spirometric measures were taken to six healthy adults, at rest. The volume, flow and time-normalized average Vt and Flow of the inspiratory and expiratory phases, is displayed in fig(23).

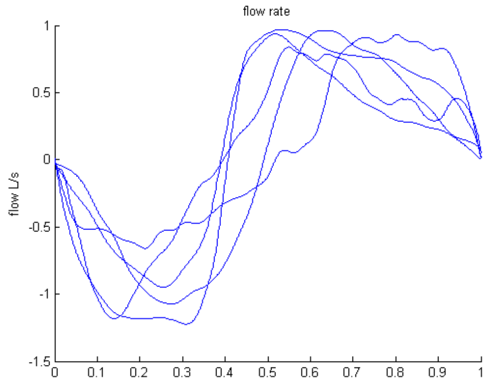


Figure 23: averaged respiratory cycles for several adult, healthy subjects at rest.

As the overall flow shapes do show such large variability even for a restricted physiological state, it seems reasonable to obviate this level of detail and synthesize simple breathing shapes that just take into account the main global descriptors. Thus a simple piecewise exponential shape properly scaled to the desired respiratory frequency, tidal volume and inspiratory/expiratory ratio will be used. Fig(24) displays two volumetric and flow graphs modeling a tired and a relaxed state according to the literature.

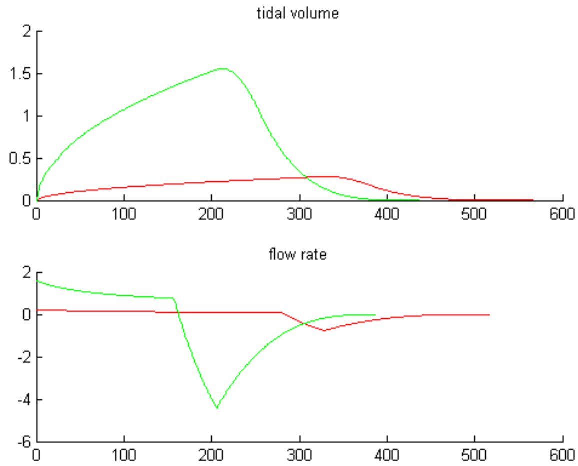


Figure 24: two synthesized breathing cycles. Top: tidal volume, Down: flowrate for tired (green) and relaxed (red) states, time normalized

The final implementation is thus able generate respiratory patterns for relaxed and exhausted subjects, as well as smoothly interpolating between both states, as shown in the synthetic pattern displayed in fig(25)

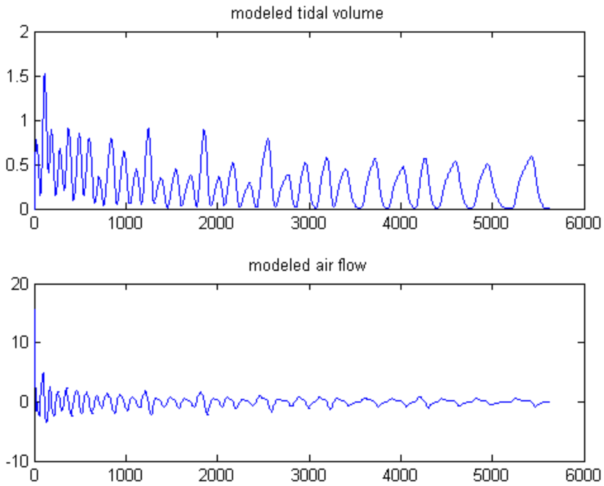


Figure 25: a synthesized breathing pattern emulating a subject exhausted and then relaxed. Top: tidal volume, Down: flowrate.

4. Software implementation

The breathing synthesizer has been implemented as a Matlab prototype. The visual interface has been designed for debugging purposes and to show the capabilities of the synthesis engine. Exhaustive options for selecting the analysis parameters are not provided in the GUI, to make it simpler to operate (see fig(26)). All the generated data is evaluated in the Workspace for better debugging.

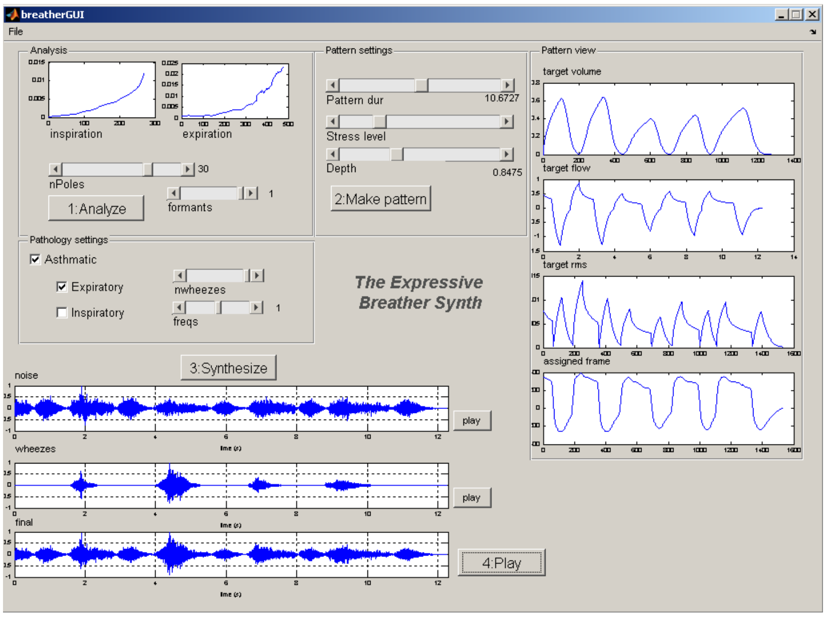


Figure 26: Graphical user interface for the Matlab breathing synthesizer prototype

The steps required to generate a breathing are

- Open the prototype by double-clicking on the breatherGUI.m file
- Open one of the provided inspiration and expiration samples from the *File* menu. You may want to assign a different vocal aperture for the inspiratory and expiratory phase for example.
- Choose a suitable number of poles for the LPC analysis (minimum 12, but more may improve the result) and press the *Analyze* button. The formant scaler will take effect in the synthesis phase but it is place there because it operates on the LPC coefficients.
- Choose a duration for the pattern, in seconds, a stress level (from relaxed, left, to exhausted, right) and a depth (how intense the breathing is to be) and press the *Make pattern button*. The tidal volume, flow rate, inferred rms and selected frames will be displayed on the Pattern view panel.
- Should you desire to incorporate wheezes to the breath sound, click on the *Asthmatic* checkbox in the *Pathology settings* panel. The first time it will take a time to load the sample and spirometry files. You may also want to configure some basic parameters like checking whether the wheezes will appear in the inspiratory or expiratory phases or both, their number and scale their frequencies.

- Press the *Synthesize* to compute the audio. First, the noisy part is computed, and if the *Asthmatic* button has been checked, the wheezes are incorporated. The normal (noise), wheezing and final synthesized waveforms will appear on the graphs below.
- Press the *Play* buttons to hear the result. Either the noise excited LPC resynthesis for the normal sounds, the sinewave reconstructed wheezes and the final result can be played-

5. Evaluation

To evaluate the quality of the Expressive Breathing synthesizer two factors were taken into account. First, one must evaluate the sound quality (specifically, how similar or easy to recognize was a breathing sound), and second, the expressiveness of that sound, its ability to convey a given physiological or even emotional state to the listener.

The evaluation was done to 7 adult users with no known auditory impairment, listening to a number of samples with headphones at a soft, natural level. The samples were chosen in random order for each user.

The first evaluation required a number of users to judge how realistic were a number of synthesized breathing sounds, taking into account that the quality of the synthesis is far from optimal, being the main drawbacks the inherent limitations of the LPC resynthesis, the artifacts introduced by the time stretching and the low sampling rate (16k) used in the prototype.

The realism was not intended to be quantified in reference to real recordings, but compared to user's experience and expectations. That is, we are evaluating perceived realism but not comparative similarity.

Six short patterns of no more than two or three respiratory phases were provided, four for normal breathing (slow, slow but deep, fast, fast and deep) and two for asthmatic breathing (a slow breath with the whole wheeze data incorporated, and a fast and deep breath with only some wheezes of lower frequency, thus more masked by the normal breath noise). Likert scales where provided to evaluate the quality and realism of the sound samples.

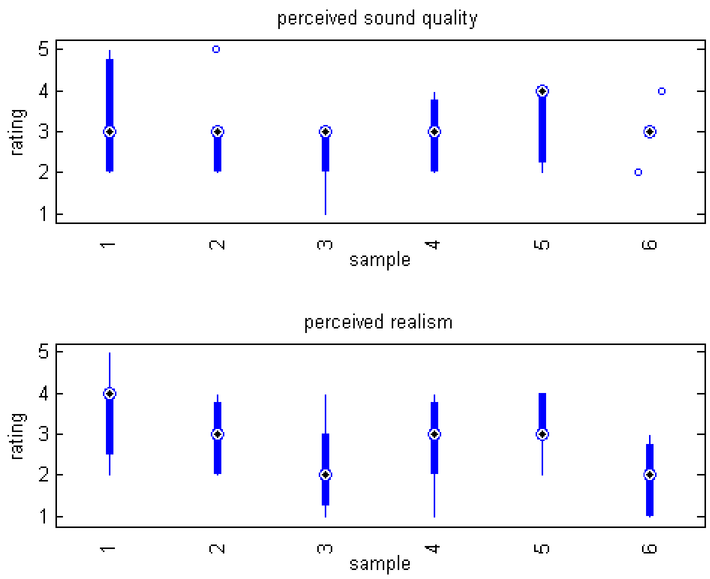


Figure 27: Boxplots for quality and realism user tests

There is good agreement in the good quality of the asthmatic samples (sounds 5&6). The reason might be that the addition of pure sinusoids adds clarity to the sound. For the realism, it is surprising that the sample 1 is generally considered as quite realist, despite being one with a large amount of noise and low detail caused by the softer flow rate mapping, one must conclude that the pattern play a more important role to impart

naturalness to the sample. This fact is confirmed by samples 3&6, those with the depth parameter at max, and thus generating abrupt inspiratory and expiratory onsets: they are perceived as less realistic.
Finally, the soft asthmatic sample is logically perceived as the most convincing sound, which according to users was because of the expressiveness and detail added by the incorporated wheezes.

The second evaluation was more difficult to determine. The question here is to determine whether the breathing sounds, so much rooted to basic physiologic and affective states and so deeply familiar to users, could express some basic emotional states. Because the synthesizer operates basically in a two dimensional timbral space, that is, it is able to synthesize breathing patterns for different levels of stress and pathology (as the amount of wheezes may be correlated to the severity of a respiratory disease), we hypothesize that the stress and depth level could be perceived as an activity cue and the pathology amount could be related to a negative/positive valence cue (but possibly also the stress level), following somehow the two-dimensional emotion space suggested by Juslin [22].
Four patterns where provided to the users: soft, slow, normal breathing, deep, fast, normal breathing, soft, slow, asthmatic breathing and deep, fast, asthmatic breathing. They were asked to evaluate how active and how anguished the patterns sounded to them.

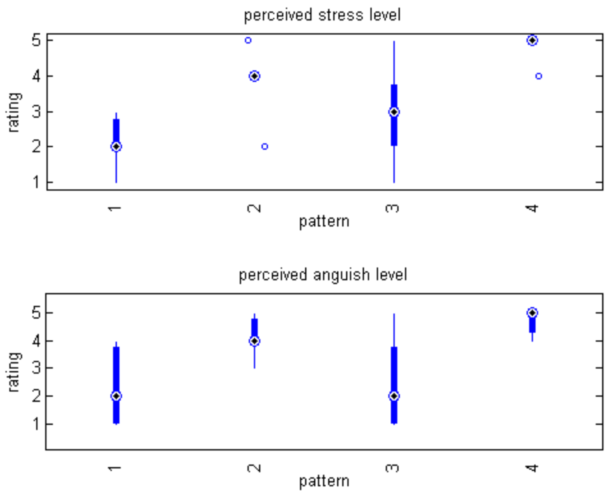


Figure 27: Boxplots for stress and anguish user tests

The four patterns are evaluated with more consistency than the previous short samples. Greatest agreement, as expected, is in the perceived stress level of patterns 2&4 (the faster ones), possibly the provided respiratory parameters for the rest state are a bit fast, according to users, and sound a bit too noisy for a quiet, relaxed breathing, which is one of the drawbacks of the RMS based mapping.
The anguish level was consistently related to the stress level, that is, both variables are strongly dependent. The conclusion is that wheezing sounds don't contribute as much to the negative valence of the generated patterns as the patterns themselves (in fact, for some users they surprisingly were not perceived as any indicator of discomfort), and thus we are restricted to a one dimensional emotional state: relaxed and comfortable versus exhausted and anguished.

6. Conclusions and Future Work

The developed prototype is expected to serve as a useful framework to explore a variety of respiratory sounds and patterns, as the implemented patterns are just a proof of concept and extending the engine to support more patterns shouldn't be difficult.

The sound quality is mostly limited by the inherent limitations of the noise excited, time-stretched LPC resynthesis approach. Maybe further abstracting the spectral envelopes and generating normal breathing sounds completely from scratch would provide a cleaner sound, but at the expense of flexibility and easiness to incorporate more reference respiratory cycles easily.

The user evaluations do show that the most presumably successful and effective way to convey expressive information to a synthetic breathing is the breathing pattern modeled. For that reason, a promising line of research could be the study of the expressiveness of pathological or emotional breathing patterns. Greater care should be needed to be able to model subtler patterns and gestures (sights, shivers…), which would require a precise RMS mapping, with careful recordings and the use of non-invasive flow measurements which I unfortunately didn't have access to.

Finally, porting the prototype to a realtime synthesis environment would let the user automate or control manually, with different levels of abstraction, the main respiratory parameters, even driving them in synchrony with 3D characters. Only the wheeze model should be redefined to work in realtime, the remaining modules are ready to be ported.

REFERENCES

[1] *Breathing patterns during varied activities*, Paek Dumyung. Journal of Applied Physiology, Vol 73, Issue 3 Pages 887-893

[2] *The pattern and timing of breathing during incremental exercise: a normative study*, J.A. Neder, S. Dal Corso, C. Malaguti, S. Reis, M.B. De Fuccio, H. Schmidt, J.P. Fuld, L.E. Nery European Respiratory Journal 2003, Vol 21, Issue 3, Pages 530

[3] *Respiratory responses associated with affective processing of film stimuli* Gomez, P.; Zimmermann, P.; Guttormsen-Schar, S.; Danuser, B.; Biological psychology 2005, Volume 68, Issue 3, March 2005, Pages 223-235

[4] *The effects of emotional behaviour on components of the respiratory cycle* Boiten, F.; Biological psychology Volume 49, Issues 1-2, September 1998, Pages 29-51

[5] *Respiratory sound analysis [Introduction for the Special Issue]* Moussavi, Zahra; IEEE Engineering in Medicine and Biology Magazine 2007

[6] *State of the Art Lung Sounds* Robert Loudon and Raymond l. h. Murphy, jr. http://www.stethographics.com/publications/p2.pdf Retrieved Date 6/29/2010

[7] *Breathing Patterns 1. Normal Subjects* MartinJ. Tobin, M.D.; Tejvir S. Chadha, M.D.; GilbertJenouri, BA;StephenJ. Birch, B.S.; Hacik B. Gazeroglu, B.S.; and Marvin A. Sackner, M.D., F.C.C.P. Downloaded from chestjournal.chestpubs.org on May 4, 2010

[8] *Breathing patterns. 2. Diseased subjects.* M J Tobin, T S Chadha, G Jenouri, S J Birch, H B Gazeroglu and M A Sackner Chest 1983;84;286-294 Downloaded from chestjournal.chestpubs.org by guest on May 4, 2010

[9] Algorithm *for Time-Frequency Detection and Analysis of Wheezes* Antoni Homs-Corbera, Raimon Jané , José Antonio Fiz, José Morera Proceedings of the 22"d Annual EMBS International Conference, July 23-28,2000, Chicago IL.

[10] *Analysis of Forced Wheezes in Asthma Patients* J.A. Fiz a R. Jané b J. Izquierdo a A. Homs b M.A. García a R. Gomez a E. Monso a J. Moreraa. Respiration 2006;73:55–60 Received: March 10, 2004 Accepted after revision: February 9, 2005 Published online: August 17, 2005

[11] *Lung Sounds* Robert G. Loudon and Raymond L. H. Murphy CHAPTER 100 THE LUNG: Scientific Foundations Second Edition edited by R. G. Crystal, J. B. West, et al. Lippincott - Raven Publishers, Philadelphia 1997

[12] *Breathing rhythms and emotions* Ikuo Homma and Yuri Masaoka Exp Physiol 93.9 pp 1011–1021

[13] *Spectral Modeling Synthesis. A Sound Analysis/Synthesis Based on a Deterministic plus Stochastic Decomposition* Serra, X. & Smith, J. Computer Music Journal, 1990, 14, 12-24

[14] *A System for Sound Analysis/Transformation/Synthesis based on a Deterministic plus Stochastic Decomposition* Serra, Xavier. Phd thesis (304) Stanford University, 1989

[15] *Efficient Speech Synthesis System using the Deterministic plus Stochastic Model* Daniel Erro, Asunción Moreno Speech Prosody 2006 Dresden, Germany May 2-5, 2006 ISCA Archive http://www.isca-speech.org/archive

[16] Functional *Phonology.Formalizing the interactions between articulatory and perceptual drives*. P.Boersma. PhD Thesis. University of Amsterdam, 1998

[17] *Physical Modeling of the Singing Voice. Malte Kob*. PhD Thesis. Aachen University of Technology, Germany, 2002

[18] *Detection of wheezing during maximal forced exhalation in patients with obstructed airways* Jose A Fiz; Raimon Jane; Antonio Homs; Jose Izquierdo; et al Chest; Jul 2002; 122, 1; ProQuest Health and Medical Complete pg. 186

[19] *Measurement and theory of wheezing breath sounds* Noam Gavriely, Yoram Palti, Gideon Alroy, James B.Grotberg 1984, American Physiological Society

[20] *Significant differences in flow standardised breath sound spectra in patients with chronic obstructive pulmonary disease, stable asthma, and healthy lungs* L Pekka Malmberg, Leena Pesu, Anssi R A Sovijarvi Thorax 1995;50:1285-1291 Downloaded from thorax.bmj.com on February 1, 2010

[21] *Spectral Analysis of Respiratory Sounds in healthy and pathological subjects*. F.C.Güler, T.Engin, B.Sankur, Y.P.Kahya. Proceedings of the 1992 International Biomedical Engineering Days, 1992

[22] *Music and Emotion* P.N. Juslin & J.A. Sloboda. Oxford: Oxford University Press, 2001

[23] *Impact of respiratory mechanics model parameters on gas exchange efficiency* S. Martin, T. Similowski, C. Straus and B. Maury ESAIM: Proceedings, June 2008, Vol. 23, p. 30-47

[24] *Breathe Easy: Model and control of simulated respiration for animation* Victor B. Zordan, Bhrigu Celly, Bill Chiu, and Paul C. Di Lorenzo.University of California Riverside Eurographics/ACM SIGGRAPH Symposium on Computer Animation, 2004

[25] Steven Lehrer, Understanding Lung Sounds Audio CD. Elsevier Health Sciences. 2002. ISBN: 1416068384

[26] R.A.L.E. Lung Sounds 3.12 Professional Edition. Software available at The R.A.L.E. Repository http://www.rale.ca/
[27] Auscultation Skills, Second Edition. Breath & Heart Sounds Audio CD. Lippincott Williams & Wilkins; Second edition, 2002. ISBN: 1582551588