TRANSMISSION LINE MODELLING OF THE HUMAN VOCAL TRACT

A Benkrid* M Koudil ** M Atmani* and TE Cross***
*Electronic Engineering Dpt, College of technology at Dammam, PO box 7650, Saudi Arabia.
E-mail : benkridah@hotmail.com
** Institut National de formation en Informatique INI Oued smar BP 68M Algiers, Algeria.
*** Electrical and Electronic Engineering Dpt, University of Nottingham, England.

Abstract: This paper describes the development of a physiologically realistic model of the vocal tract using the well established technique of transmission line modelling (TLM). This technique is based on the principle of wave scattering at transmission line segment boundaries and may be used in one, two or three dimensions. This work uses this technique to model the vocal tract using a one dimensional transmission line. A six port scattering node is applied in the region separating the pharyngeal, oral and the nasal part of the vocal tract.

1 Introduction

The aim of this research work is to develop a physiologically realistic model of the vocal tract (figure 1.), capable of running in real time, using transmission line modelling (TLM)[1]. The relative computational simplicity of this modelling technique opens up the possibility of using the latest generation of digital signal processors to build a multiprocessor platform.

The TLM technique has previously been used by other workers to solve a wide range of field problems efficiently[2-5], with a rapid rate of convergence and unconditional stability. TLM simulates the analogue process of the system under study by splitting it into a number of sections which are separated from each other by scattering nodes. It can be applied in one, two or three dimensions depending on the nature of the problem and the accuracy required from the model. In order to guarantee the performance of the vocal tract model, it is essential to calculate the number of sections which are required to model the system with an acceptable degree of accuracy. It is necessary to establish the precision at which the TLM model has to be used to minimise the error arising only from the ability of the transmission line model to represent the system and not from the numerical solution of the model[6].

2 Modelling of the system

The length of the vocal tract is comparable to the wavelength of the sound at audio frequencies. It is therefore possible to represent the tract as a one-dimensional acoustic tube of variable cross-sectional area (figure 2.). This makes it convenient to apply TLM to the vocal tract.

The mathematical treatment of the sound production process involves the following successive operations. The vocal tract must first be viewed in terms of an area function, S, which describes the sectional area, perpendicular to the air stream, from the glottis to the radiating surface at the lips. Secondly, this function has to be approximated by a sufficient number of successive parts, each of a constant cross-sectional area. Lastly, the model is excited by an appropriate input and an output obtained which is characteristic of the shape of vocal tract.

A set of partial differential equations can be obtained that describe the motion of air in such a system[6-10]. So long as the greatest cross-sectional dimension of the system is appreciably less than the wavelength and the internal wave reflection is well known, it has been shown[7] that sound waves in the system satisfy the following pair of equations:

$$\frac{\partial P}{\partial x} = r\,\frac{\partial (U/S)}{\partial t}$$

$$\frac{\partial U}{\partial x} = \frac{1}{r\,c^2}\frac{\partial (P.S)}{\partial t} + \frac{\partial S}{\partial t}$$

Where:

P:- P(x,t) is the pressure at position x and time t.
U:- U(x,t) is the volume velocity at position x and time t.
ρ:- Density of the air in the tube.
c:- Velocity of the sound.
S:- S(x,t) is the area function.

A complete solution of the differential equations requires that pressure and volume velocity must be worked out for x and t between the glottis and the lips, taking into consideration the boundary conditions. The sectional areas which are changing, not only along the vocal tract but also in time, must be known. Measurements of S(x,t) are difficult to obtain even for continuant sounds when the vocal tract shape can be reasonably assumed not to change with time. Historically the use of X-ray motion pictures has provided data of this form[11]. A more modern technique uses Magnetic Resonance Imaging[12]. It is supposed that during a certain number of time-steps, the areas will not have enough time to change. Taking this into consideration, the previous set of equations becomes:

$$-\frac{\partial P}{\partial x} = r\frac{1}{S}\frac{\partial U}{\partial t}$$

$$-\frac{\partial U}{\partial x} = \frac{S}{r\,c^2}\frac{\partial P}{\partial t}$$

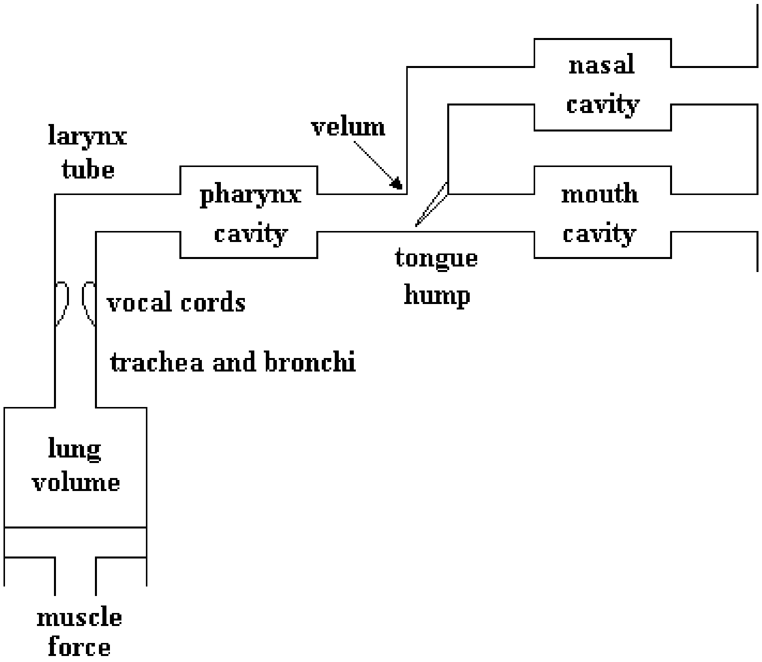


Figure 1.

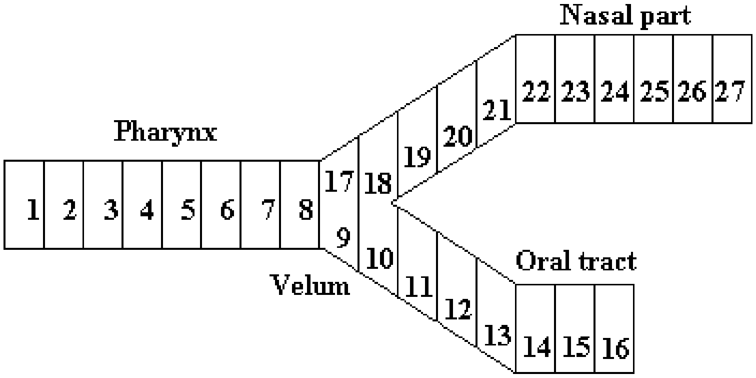


Figure 2.

The volume velocity, U(x,t), and pressure, P(x,t), in an acoustic tube may be represented by forward and reverse waves, $u^+(t-x/c)$ and $u^-(t+x/c)$[13] as follows:

$$U(x,t) = u^+(t-x/c) - u^-(t+x/c)$$

$$P(x,t) = \frac{rc}{S}\left[u^+(t-x/c) + u^-(t+x/c)\right]$$

An acoustic tube may be divided into a number of cascaded transmission line segments. Using TLM all segments are of equal length and within a segment the area remains constant. The time step of the model, δt, is therefore given in terms of the segment length, l, and wave velocity, c, as: $dt = l/c$. Figure 3. illustrates two adjacent segments of different characteristic area S[i] and

S[i+1] carrying volume velocity waves in the forward direction $u^+_i(t)$, $u^+_{i+1}(t)$ and reverse direction, $u^-_i(t)$, $u^-_{i+1}(t)$. The reflection coefficient, ρ, for the volume velocity waves (analogous to a current wave) is given by:

$$r = \frac{S[i+1] - S[i]}{S[i+1] + S[i]}$$

As the pressure and volume velocity must be continuous in both time and space everywhere in the system, the previous set of equations may be re-written:

$$u^+_{i+1}(t+dt) = (1+r)u^+_i(t) + r\,u^-_{i+1}(t)$$
$$u^-_i(t+di) = -r\,u^+_i(t) + (1-r)u^-_{i+1}(t)$$

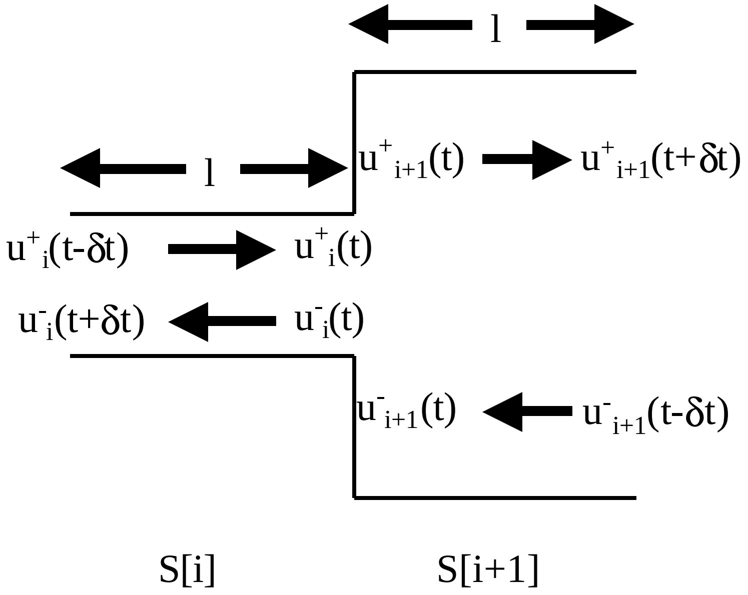
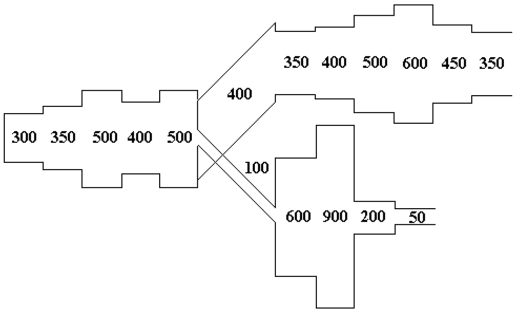


Figure 3.

As long as the interest is only in values of pressure and volume velocity at the input and output of the sections constituting the vocal tract system, the previous set of equations describe totally each section of the system. This is not restrictive since the aim is to relate the output of the last section and the input of the first section. The model is extended to consider a representation of the nasal cavity. This is necessary for generating sounds such as /m/, /n/ and /ng/ when the area of part of the oral tract is reduced to zero at and the nasal path is the dominant sound transmission channel.

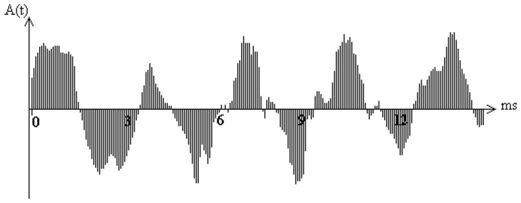
3 Results

In order to test the performance of the modelling platform, simulations were run so the results of which could be compared with previous non real-time results [14]. This was in agreement with results obtained from non real-time simulations[14]. The performance of the system was established by simulating several vowel sounds[14]. The nasal sound /m/ is generated by a closing of the lips., a lowering of the velum to inject energy into the nasal tract and a glottal wave excitation (figure 4..). Figure 5 represents a frequency-time-intensity domain plot of the word summer. This is produced by the successive generation of the sounds /s/, /a/, /m/ followed by the opening of the lips to release the stored energy. The sounds /s/ and /m/ are as described above. The sound /a/ is produced by injecting a glottal wave into an appropriately shaped mode[15].



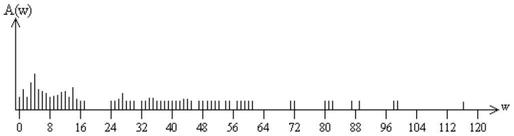
Vocal tract profile corresponding to the sound /m/. The areas are in mm^2

Figure 4a.



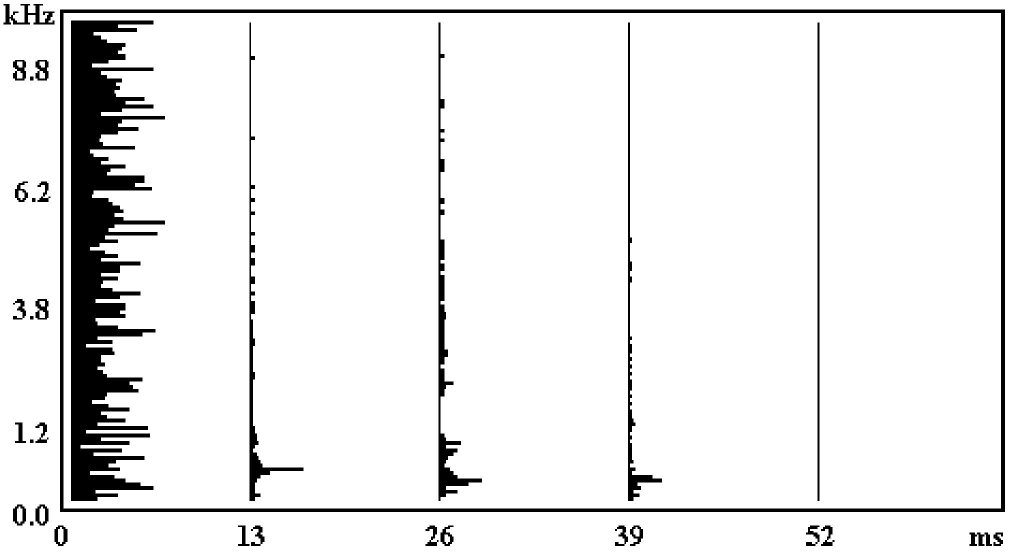
The output samples from the oral tract transmission line modelling, of figure 4a. terminated by the free space impedance run in real-time

Figure 4b.



256 pts FFT from the I/O model above. The resolution of the FFT is equal to 78 Hz

Figure 4c.



Frequency-time-intensity domain plot of the word summer

Figure 5.

4 Conclusion

This work has demonstrated the feasibility of using transmission line modelling for the development of a physiologically realistic model of the vocal tract running in real-time. A further increase in the number of processors will allow the model to be run faster than real-time making the use of iteration a feasible proposition. The authors believe that the use of TLM technique to simulate the vocal tract advantages as this technique exhibits a close relationship between the model parameter and the mechanisms of the wave propagation in the vocal tract. It is also straightforward to formulate and programme in real-time using currently available digital signal processing chip.

REFERENCES:

[1]: PB Johns and RL Beurle, “Numerical Solution of 2-dimensional scattering problems using TLM”, Proc IEE Vo1118, No 9, pp15-31, July 1977

[2]: S Akhtarzad and P.B Johns, “The Solution of Maxwell’s equations in three space dimensions and time by the TLM method of numerical analysis” Proc. IEE 122, pp1349-1352 (1975).

[3]: P.B Johns and G Butler, “The consistency and the accuracy of the TLM method for diffusion and its relation to existing methods”, Int J num. Mod., 19 pp1549-1554 (1983).

[4]: P.B Johns and Y Rahal, “ The use of diakoptics for numerical simulation of the penetration of time domain fields in aircraft” London England, IERE, pp57-63 (1983).

[5]: P.B Johns, “A simple and unconditionally stable numerical routine for a solution of the diffusion equation” Int. J. Numerical Methods Eng., No 11, pp1307-1328, (1977).

[6]: M.M Sondhi "Model for wave propagation in a lossy vocal tract" J. Acoust Soc. Amer, vol. 55, No 5, pp 1070-1075 (May 1974).

[7]: L.L Beranek "Acoustics",Mc Graw Hill Book co,N.Y 1975.

[8]: P.M Morse and K.U Ingard "Theoretical Acoustics" Mc Graw Hill .Book co,N.Y,1968.

[9]: M.R Portnoff and R.W Schaffer "Mathematical Considerations in Digital Simulation of the Vocal Tract" J. Acoust Soc. Amer Vol. 53,No 1 (abstract) p 294 January 1973.

[10]: J.L Flanagan "Voices of Men and Machines" J acoust Soc Amer 51,pp 1375-1377 (1972).

[11]: G Fant "Acoustic Theory of Speech Production" Mouton. 1970.

[12]: A.R. Greenwood, C.C. Goodyear and P.A. Martin “Measurements of Vocal Tract Shapes Using Magnetic-Resonance-Imaging, IEE Proc.1 Comms Speech and Vision Vol 139, No6 pp553-560, 1992 .

[13]: M.V Mathews, J.E Miller and E.E David "Pitch synchronous analysis of voiced sounds" J. Acoust. Soc Amer.,Vol 33, pp 179-186, 1961.

[14]:A.Benkrid, «Real-time TLM Vocal Tract Modelling », PhD Thesis, University of Nottingham, 1989.

[15]: A Benkrid, TE Cross, M Atmani and M Koudil " A multiprocessor TMS320 platform for real-time modeling". IEEE, TEM, No 6, pp 174-178, 1999.