Edinburgh Research Explorer

Real-Time Vocal Tract Modelling

Research Article
Real-Time Vocal Tract Modelling

A. Benkrid, A. Benallal, and K. Benkrid

To date, most speech synthesis techniques have relied upon the representation of the vocal tract by some form of filter, a typical example being linear predictive coding (LPC). 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 parts of the vocal tract.

1. INTRODUCTION

The introduction of fast digital signal processors has made possible the construction of computing platforms capable of modelling problems in real time. This presents a challenge in two fields, firstly in the design of the hardware necessary to build such a platform and secondly in the formulation of the problem to be modelled in such a way that it can take full advantage of the processing speed available. Modelling of the human vocal tract is one area, where the ability to generate real-time output is of significant advantage as it is then possible to listen to the output directly. This paper shows how transmission line modelling (TLM) may be used to model the human vocal tract in such a way that full advantage may be taken of a multi-DSP processing platform built by the authors. The formulation of the TLM algorithm makes it ideally suited to the DSP chips which were emerging in the mid-1980s. The first family of DSPs to become established as an industry standard was the TMS320 series.

Most speech synthesis techniques developed to date have relied upon the representation of the vocal tract by some form of filter, a typical example being linear predictive coding (LPC) [1]. The aim of this research work is to develop a physiologically realistic model of the vocal tract, capable of running in real-time, using transmission line modelling (TLM) [2]. 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 hardware used to run the model is based on a number of TMS320's processors working in parallel and being supported by an IBM PC compatible computer acting as host. Before the model may be run, the code for the task to be performed by each processor must be loaded from the host. When the model is running each time frame is begun by the host passing the coefficients to be used for that frame to each processor. At the end of each time-step, the output generated by one processor must be passed to the processor responsible for an adjacent section of the model. Consequently, it can be appreciated that communication between the various processors is an important consideration in the design of the hardware. There are several solutions to this problem; the use of global memory to which all processors have access or direct processor-to-processor communication using, for example, input-output ports are two alternatives. The nature of TLM [3] allows direct communication to be used to advantage, whereas global memory may be a more flexible alternative for general problems. This paper is concerned primarily with the formulation of the model such that full advantage of the multiprocessor platform can be realised rather than with the detail of that platform.

The TLM technique has previously been used by other workers to solve a wide range of field problems efficiently [4–8], 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 [9]. In order to do this successfully three conflicting demands have to be satisfied. These are the programme run-time, the accuracy of the model output, and the length of the time-step. In this particular case, however, there is the additional demand that the model must run in real-time.

The vocal tract is a multiple cavity whose first four resonances are centred at 500 Hz, 1500 Hz, 2500 Hz, and 3500 Hz. However, as the articulators move, these resonances vary. Thus, the speech spectra are not completely arbitrary but are characterised by resonant peaks called formants and most of the energy of a typical speech signal is contained within a limited bandwidth. All the formants characterising the different phonemes, with an acceptable audibility, are included within a bandwidth of about 3.5 kHz. Indeed the telephone system uses a bandwidth of 300 to 3.3 kHz to transmit speech with an acceptable clarity.

Conventionally, for acceptable performance of a TLM model, the node spacing should be about one tenth of the smallest wavelength to be modelled. With a maximum frequency of 3.4 kHz ($\lambda_{\min} = 100$ mm) the section lengths will be about 10 mm [10]. A typical vocal tract is 175 mm long so using sixteen sections produces a model of acceptable resolution and gives a time-step (T) of 32 μs. Including the nasal tract, which is typically 120 mm long, the model would increase in size and contain a total of 27 scattering nodes as illustrated in Figure 1. With a real-time model the output samples must be produced at a rate which is greater than the Nyquist limit; for a signal bandwidth of 3.4 kHz this would be about 147 μs [10]. This indicates that the model is approximately four times oversampled. A TMS32020 DSP with a clock speed of 20 MHz requires about 4 μs for a simple node calculation. The model proposed above requires 27 such calculations to be performed in a single time-step which would take about 100 μs. This is about a factor of three greater than the required time-step and has also excluded the time required for the six-port calculation at the pharynx, reflection coefficient update, generation of the excitation waveform, and other housekeeping tasks. The task, therefore, cannot be achieved in real-time by a single processor, it is thus necessary to use a multiprocessor system.

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. 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 terms of an area function, S, which describes the sectional area, perpendicular to the air
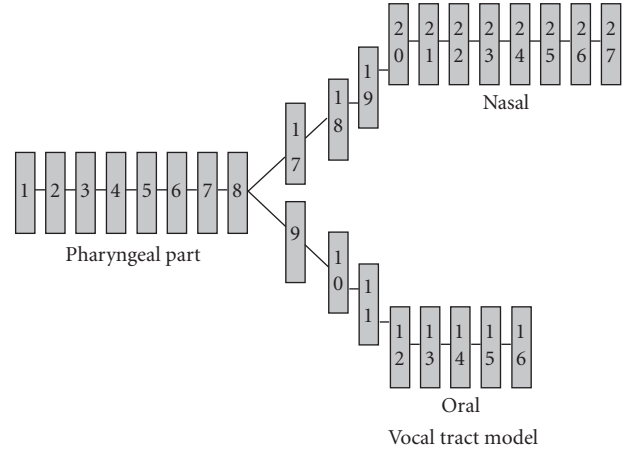

Figure 1: Vocal tract model.

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 [12–16]. 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 [13] that sound waves in the system satisfy the following pair of equations:

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

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

where

  (i) $P(x, t)$ is the pressure at position x and time t;

  (ii) $U(x, t)$ is the volume velocity at position x and time t;

  (iii) ρ is the density of the air in the tube;

  (iv) c is the velocity of the sound;

  (v) $S(x, t)$ is the area function of the tube.

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 [15]. A more modern technique uses magnetic resonance imaging [16]. It is supposed that during a certain number of time-steps, the areas will not have enough time to change. Taking this into consideration,

(1) becomes

$$-\frac{\partial P}{\partial x} = \rho \frac{1}{S} \frac{\partial U}{\partial t},$$
$$-\frac{\partial U}{\partial x} = \frac{S}{\rho c^2} \frac{\partial P}{\partial t}. \tag{2}$$

The above relations are comparable to those defining an electrical lossless transmission line [17, 18]:

$$-\frac{\partial V}{\partial x} = L \frac{\partial I}{\partial t},$$
$$-\frac{\partial I}{\partial x} = C \frac{\partial V}{\partial t}, \tag{3}$$

where

(i) $V(x, t)$ is the voltage in the line;
(ii) $I(x, t)$ is the current in the line;

L and C are, respectively, the inductance and the capacitance of the transmission line.

The electrical analogue of an acoustic tube is an electrical transmission line in which the current is analogous to the volume velocity and the voltage is analogous to the sound pressure [19]. The characteristic impedance at a point on the analogous electrical line depends upon the cross-sectional area at the corresponding point on the acoustic line. The characteristic impedance of an electrical line is given by [10]

$$Z_{\text{el}} = \sqrt{\frac{L}{C}}. \tag{4}$$

The equivalent acoustic impedance is equal to

$$Z_{\text{ac}} = \frac{\rho c}{S}. \tag{5}$$

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 $u^-(t + x/c)$ [20] as follows [10]:

$$U(x, t) = u^+\left(t - \frac{x}{c}\right) - u^-\left(t + \frac{x}{c}\right),$$
$$P(x, t) = \frac{\rho c}{S}\left[u^+\left(t - \frac{x}{c}\right) + u^-\left(t + \frac{x}{c}\right)\right]. \tag{6}$$

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 $\delta t = l/c$. Figure 2 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 [10]

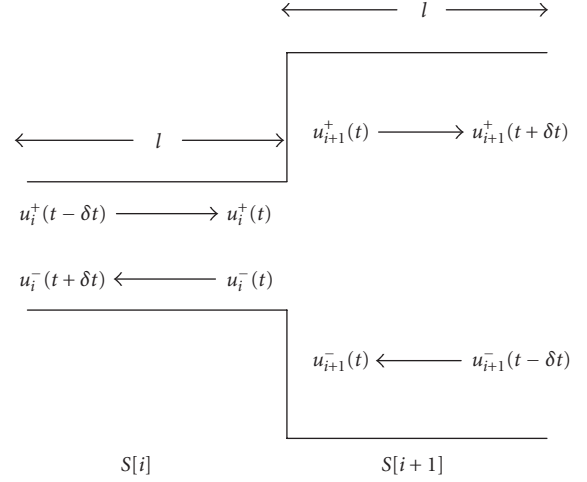$$\rho = \frac{S[i+1] - S[i]}{S[i+1] + S[i]}. \tag{7}$$

FIGURE 2: The junction between two transmission line segments of area $S[i]$ and $S[i + 1]$ showing the incident and reflected waves.

The pressure and volume velocity may be calculated at the junction between the ith and $(i + 1)$th section [19],

$$U_i(x, t) = u_i^+\left(t - \frac{x}{c}\right) - u_i^-\left(t + \frac{x}{c}\right),$$
$$P_i(x, t) = \frac{\rho c}{S[i]}\left[u_i^+\left(t - \frac{x}{c}\right) + u_i^-\left(t + \frac{x}{c}\right)\right]. \tag{8}$$

As the pressure and volume velocity must be continuous in both time and space everywhere in the system, (8) may be rewritten [19]:

$$u_{i+1}^-(t) = (1 + \rho_i)u_i^+(t - \delta i) + \rho_i u_{i+1}^+(t),$$
$$u_i^-(t + \delta i) = -\rho_i u_i^+(t - \delta i) + (1 - \rho_i)u_{i+1}^+(t). \tag{9}$$

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, relations (9) 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 and the nasal path is the dominant sound transmission channel. The three port matrix between the pharyngeal, oral, and the nasal part is calculated [10] by considering the continuity of the volume velocity and the pressure (Figure 3);

$$\begin{bmatrix} C[i] \\ D[k] \\ D[i] \end{bmatrix} = \frac{1}{W} \begin{bmatrix} G & 2S[i] & 2S[i] \\ 2S[k] & R & 2S[k] \\ 2S[i+1] & 2S[i+1] & H \end{bmatrix} \begin{bmatrix} A[i] \\ B[k] \\ B[i] \end{bmatrix}, \tag{10}$$

where, $G = S[i] - S[i+1] - S[k]$, $R = S[k] - S[i] - S[i+1]$, $H = S[i+1] - S[i] - S[k]$, $W = S[i] + S[i+1] + S[k]$, $C[i]$, $D[k]$
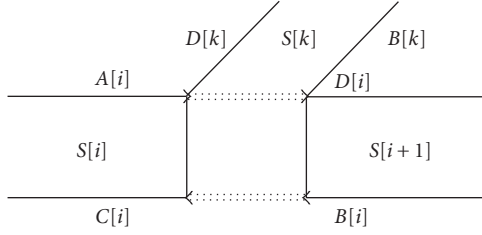
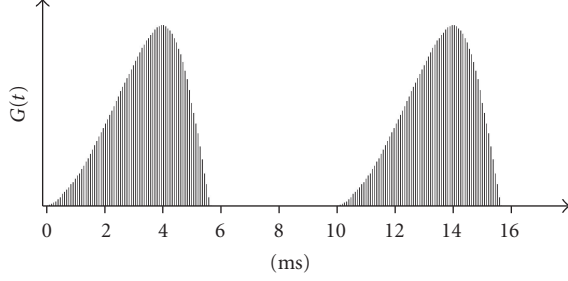FIGURE 3: Scattering node between the pharynx, oral, and nasal tracts.

FIGURE 4: The input samples of the glottal waveform.

and $D[i]$ are the backward waves and $A[i]$, $B[k]$, and $B[i]$ are the forward waves in the six-port junction according to Figure 3.

The total energy within a section of an electrical transmission line of inductance, L, and capacitance, C, is given by [10]

$$E_e = \frac{1}{2}(Li^2 + Cv^2),\qquad(11)$$

where i and v are the current and voltage within the section.

The corresponding energy within an acoustic tube of area, S, is equal to [10]

$$E_a = \frac{\rho}{S}U^2,\qquad(12)$$

where, U, is the volume velocity and ρ is the density of the air.

During speech generation, the articulators: tongue, mouth, and so forth move and the cross-sectional area of the vocal tract changes with time. In order to model these changes the cross-sectional areas of the model segments must gradually change. This change, however, can only be performed at the end of a time-step. As the change in area takes place on a much slower time-scale than the acoustic model, a small change in area is made after a number of model iterations. When the area is changed suddenly at the end of a time-step there will be no time for the mass of air within that section to change and so this is the quantity which should be preserved. This is comparable to conserving the charge on a changing electrical transmission line. However, it could also be argued that the change occurs rapidly and so there is no time for energy exchange with the travelling wave, in this case the total energy within each section will be preserved. In either case, the magnitude of the forward and backward

waves must be adjusted to account for the change in characteristic impedance. As long as the forward and reverse wave are independent of each other, the ratios $u_{i+1}^-(t)/\sqrt{S}$ and $u_i^-(t+\delta t)/\sqrt{S}$ should remain unchanged during any change of the vocal tract profile [10]. Consequently, the $u_{i+1}^-(t)$ and $u_i(t + \delta t)$ will be modified as a result of the change in S, by keeping the ratios $u_{i+1}^-(t)/\sqrt{S}$ and $u_i^-(t+\delta t)/\sqrt{S}$ constant.

The radiation model at the lips considers the free space impedance. The specific series impedance can be put in terms of equivalent parallel resistance and inductance. It has been shown [19, 21, 22] that the radiating impedance can be expressed in a normalised form,

$$Zl(\omega) = j\omega\frac{L_p R_p}{R_p + j\omega L_p},\qquad(13)$$

where $R_p = 128/9\pi^2$ is the radiation resistance, $L_p = 8r/3\pi c$ is the radiation inductance, and r is the radius of a circular orifice having the same area as the lips opening.

The radiating impedance $Zl(w)$ is modelled by a transmission line of infinite length in parallel with a short circuit stub. The characteristic impedance of this infinite line corresponds to a tube of an area given by (14). The length of the short circuit stub is equal to half that of the sections in order to allow the pulse to travel to the end of the stub and back again in one time-step [10]. Its corresponding area is given by relation (15);

$$A_{pr} = \frac{\rho c}{R_p} = \frac{9\pi^2}{128}\sqrt{A_n},\qquad(14)$$

$$A_{pl} = \frac{3\pi\sqrt{\pi}}{16}\frac{\sqrt{A_n}}{N}l.\qquad(15)$$

The determination of the glottal wave $U_g(t)$ is treated only in several papers [20, 22–24]. It is known [22] that the shape and the periodicity of the glottal vocal cord excitation are subject to large variations. The glottal waveform is characterised by closure, rising, and falling phases. In this work, the natural glottal waveform is replaced by a periodic wave, as illustrated in Figure 4. The equation representing one period is given by [24]

$$
\begin{aligned}
U_g(t) &= \frac{a}{2}\left(1 - \cos\frac{\pi m}{N1}\right) && 0 \le m < N1,\\
&= a\,\cos\left(\pi\frac{m - N1}{2N2}\right) && N1 \le m < N2,\\
&= 0 && \text{elsewhere.}
\end{aligned}\qquad(16)
$$

3. MODEL TIME-STEP

The resolution of the system corresponds to the time-step which is equal to the time taken by the wave to travel from one section to its neighbour. As part of the modelling process the vocal tract is divided into a number of sections of equal length and the velocity of sound is constant; this is necessary in order that all sections of the model have the same time interval (time-step). As the number of sections in the model is

increased, the accuracy of the model will also increase. However, when the number of sections is increased, the time-step is reduced but the time required to perform the calculation is increased. As it is necessary to perform all the calculations within a single time-step increasing the number of sections means that an increased computational load must be performed in a shorter time. Conversely, the introduction of parallel sections to model the nasal tract will increase of the number of node calculations but will not affect the time-step. The relative computational simplicity of the TLM modelling technique allows the level of parallelism to be increased as the number of sections is increased and allows the use of digital signal processors to develop a multiprocessor platform. The availability of low cost, high-power microprocessors has made possible the use of multiple parallel processors to perform complex computations. The appearance of digital signal processing chips, such as TMS32020 [25] has made possible the implementation of algorithms allowing computers to generate speech in real-time with an acceptable degree of accuracy.

4. HIERARCHICAL STRUCTURE

The development of a multiprocessor platform commenced with work on a single TMS32020, provided as a development system from Loughborough Sound Images (LSI) [26], supported by a PC host. This allowed easy development and testing of the programs. Initially a real-time model which would run on a single processor was developed. This consisted of a total of twelve sections including four for the nasal tract. The tools provided by the development system allowed the code for the model to be downloaded from the PC to the TMS32020. At the start of the modelling program the reflection coefficients are passed from the PC to the TMS320's using external (off-chip) memory as illustrated in Figure 5. The model is then allowed to run for the duration of one time frame before the reflection coefficients are updated. This time frame is typically 10 milliseconds in duration. At the start of a time frame the new reflection coefficients are moved from off-chip to on-chip memory by the TMS32020, the forward and reverse waves are adjusted and the model proceeds generating an output from the on-board digital to analogue converter (D/A) every time-step which for a twelve section model is 64 microseconds.

Once the performance of the model on a single processor was established, a second processor board was added to that from LSI. This allowed the size of the model to be increased to seventeen sections and the interprocessor communication to be thoroughly checked. In this model, five sections were used for the pharyngeal and oral tracts, and seven sections for the nasal tract, and Figure 6(a) represents the vocal tract shape characterising the nasal sound /m/. The LSI board was used to generate the excitation waveforms to model the first four sections of the pharyngeal tract and the last four of the oral tract, and also to calculate the output for the D/A converter from the output of the oral and nasal tracts. The second processor calculated the remaining parts of the model including the six-port junction. Reflection coefficients were passed from the PC to both processor boards at the start of
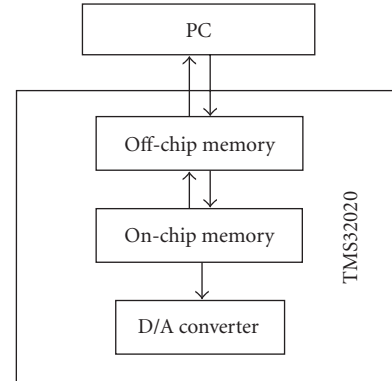

FIGURE 5: Overview of memory layout on TMS32020.

each time frame using off-chip memory in a similar manner to the single processor case.

The output from the two-processor model is shown in Figure 6(b). From Figure 6(c) formants can be observed at 270, 1090, and 2030 Hz which is in agreement with those of natural speech for this utterance [17]. Figure 7(a) shows the tract profile adopted for the fricative sound /s/ which is generated by using random noise as an excitation source. The output from this model is shown in Figure 7(b) and the spectrum in Figure 7(c).

The two-processor model was used to generate the utterance "summer" by downloading a sequence of reflection coefficient values from the PC to the two processors. The sound /s/ is generated as explained previously. Then, the sound /a/ is produced by injecting a glottal wave into the model. The nasal sound /m/ is generated by a closing of the lips and a lowering of the velum and lastly the energy built into the model is released without any additional excitation. Figure 8 shows the frequency-time intensity domain plot of the word "summer."

Further development of the platform has taken place to include six TMS32020 processors. The LSI board may now be used as a master with five slave processors thus allowing the PC to be used for general housekeeping tasks, a schematic of this arrangement is shown in Figure 9 [27]. This has not only allowed the size of the model to be incised but has also allowed other processing to be introduced. For example, the model may be run on two of the slave processors, while the LSI board is used for output generation via its D/A converter. The samples may also be converted into the frequency domain using a real-time FFT running in parallel on a further slave. This opens up the possibility of using the model to search for the vocal tract profile which matches a given utterance as defined by its frequency-time intensity data [28].

5. 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. As well as showing that the TMS320's digital signal processor can be
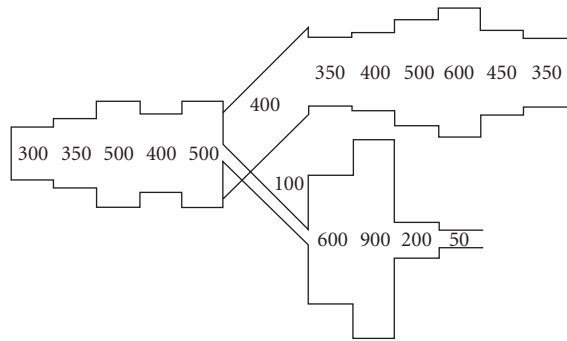
(a) Vocal tract profile corresponding to the sound /m/ [14, page 142] (areas are expressed in mm^2)
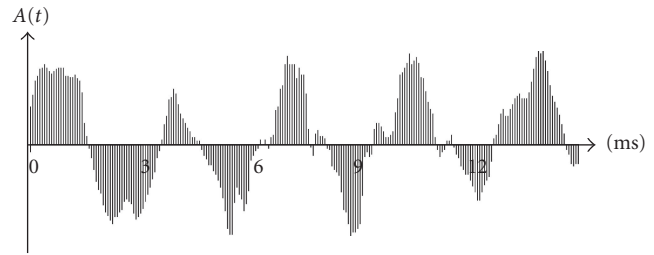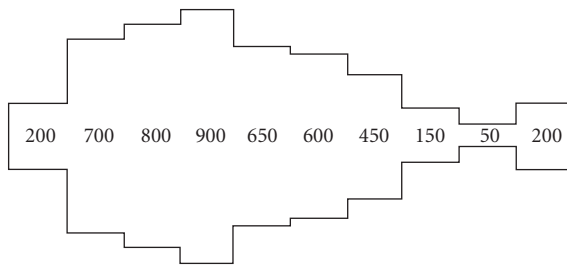
(b) The output samples from the oral tract model of Figure 6(a) terminated by the free space impedance
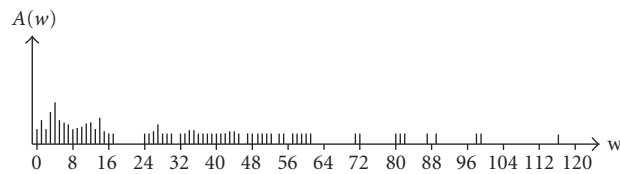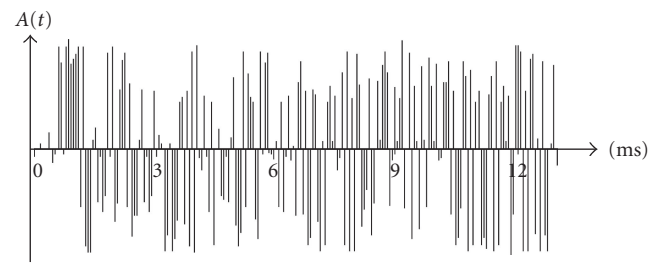

(c) 256-point FFT from the output shown above. The resolution of the FFT is equal to 78 Hz

FIGURE 6

(a) Vocal tract profile corresponding to the sound /s/ [14, page 172] (areas are expressed in mm^2)

(b) The output samples from the oral tract modeling of Figure 7(a) terminated by the free space impedance

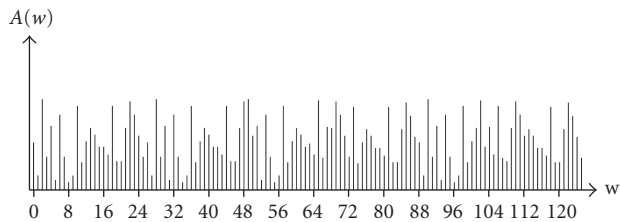(c) 256-point FFT from the output shown above. The resolution of the FFT is equal to 78 Hz

FIGURE 7

used as the basis of a multiple processor platform for running the model.

The coupling of the nasal tract has the effect of introducing zeros into the model when the oral tract is terminated by an acoustic short circuit. This is in contrast to the LPC technique which uses an all-pole filter to approximate the vocal tract. The major problem associated with LPC is in the gener-

ation of the predictor coefficients from the speech data. This is principally due to the need to test whether a particular predictor will lead to a stable filter and the minimisation of the output energy which is a nonlinear problem [5].

Using a real-time model allows the effect of a change in the tract shape on the utterance to be tested quickly and also opens up the possibility of using iterative techniques to
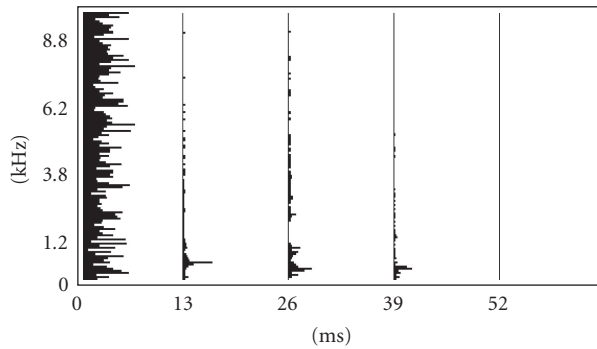
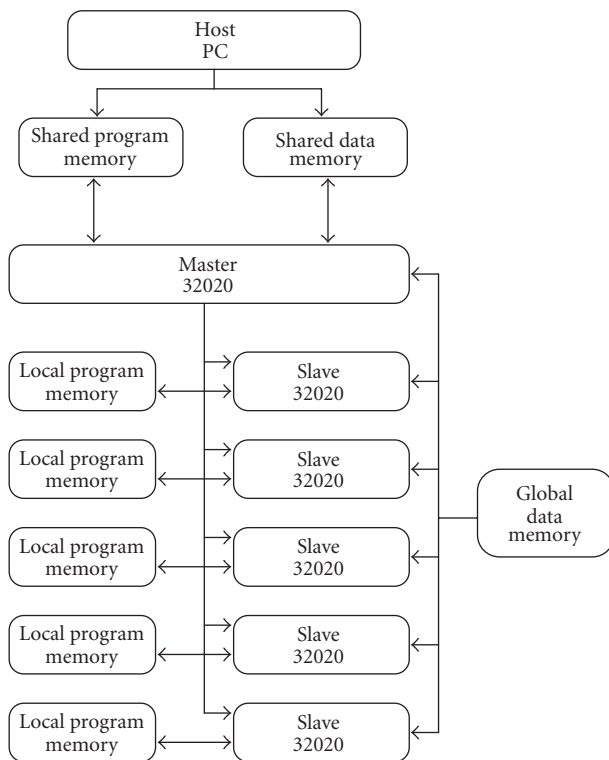FIGURE 8: Frequency-time plot of the word summer.

FIGURE 9: Overview of multiprocessor system architecture.

obtain a vocal tract area profile from an utterance. 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 appearance of DSP chips with increased word size, floating point capability, and increased support for parallelism would permit the construction of yet more powerful systems with higher performance, and the architecture of the different boards are similar to those explained elsewhere [10].

The authors believe that the use of the transmission line modelling technique to simulate the vocal tract has advantages as this technique exhibits a close relationship between the model parameters and the mechanisms of wave propagation in the vocal tract. It is also straightforward to formulate

and programme in real-time using currently available digital signal processing chips.

REFERENCES

[1] G. Bristow, *Electronic Speech Synthesis*, Granada, London, NY, USA, 1984.

[2] P. B. Johns and R. L. Beurle, "Numerical solution of 2-dimensional scattering problems using TLM," *Proceedings of IEE*, vol. 1118, no. 9, pp. 15–31, 1977.

[3] S. El-Masri, X. Pelorson, P. Saguet, and P. Badin, "Vocal tract acoustics using the transmission line matrix (TLM) method," http://www.asel.udel.edu/icslp/cdrom/vol2/522/a522.pdf.

[4] J. A. Portí and J. A. Morente, "A three-dimensional symmetrical condensed TLM node for acoustics," *Journal of Sound and Vibration*, vol. 241, no. 2, pp. 207–222, 2001.

[5] 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," *Proceedings of IEE*, vol. 122, no. 12, pp. 1349–1352, 1975.

[6] P. B. Johns and G. Butler, "The consistency and the accuracy of the TLM method for diffusion and its relation to existing methods," *International Journal for Numerical Methods in Engineering*, vol. 19, no. 10, pp. 1549–1554, 1983.

[7] P. B. Johns and Y. Rahal, "The use of diakoptics for numerical simulation of the penetration of time domain fields in aircraft," in *Proceedings of the International Conference on Networks and Electronic Office Systems*, pp. 57–63, IERE, London, UK, September 1983.

[8] P. B. Johns, "A simple and unconditionally stable numerical routine for a solution of the diffusion equation," *International Journal for Numerical Methods in Engineering*, vol. 11, no. 8, pp. 1307–1328, 1977.

[9] M. M. Sondhi, "Model for wave propagation in a lossy vocal tract," *Journal of the Acoustical Society of America*, vol. 55, no. 5, pp. 1070–1075, 1974.

[10] A. Benkrid, "Transmission line modelling of the vocal tract," Ph.D. thesis, University of Nottingham, Nottingham, UK, 1989.

[11] L. L. Beranek, *Acoustics*, McGraw-Hill, New York, NY, USA, 1975.

[12] P. M. Morse and K. U. Ingard, *Theoretical Acoustics*, McGraw-Hill, New York, NY, USA, 1968.

[13] M. R. Portnoff and R. W. Schaffer, "Mathematical considerations in digital simulation of the vocal tract," *Journal of the Acoustical Society of America*, vol. 53, no. 1, pp. 294–, 1973.

[14] J. L. Flanagan, "Voices of Men and Machines," *Journal of the Acoustical Society of America*, vol. 51, no. 5, pp. 1375–1377, 1972.

[15] G. Fant, *Acoustic Theory of Speech Production*, Mouton de Gruyter, Berlin, Germany, 1970.

[16] A. R. Greenwood, C. C. Goodyear, and P. A. Martin, "Measurements of vocal tract shapes using magnetic resonance imaging," *IEE Proceedings I: Communications, Speech & Vision*, vol. 139, no. 6, pp. 553–560, 1992.

[17] R. B. Adler, L. J. Chu, and R. M. Fano, *Electromagnetic Energy Transmission*, John Wiley & Sons, New York, NY, USA, 1983.

[18] D. T. Paris and F. K. Hurd, *Basic Electromagnetic Theory*, McGraw-Hill, New York, NY, USA, 1969.

[19] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1976.
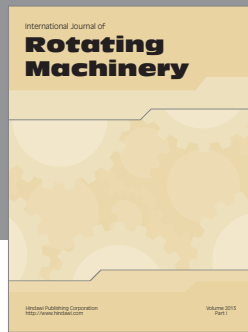
[20] H. W. Strube, "Determination of the instant of the glottal closure from speech wave," *Journal of the Acoustical Society of America*, vol. 56, no. 5, pp. 1625–1629, 1974.

[21] P. M. Morse, *Vibration and Sound*, McGraw-Hill, New York, NY, USA, 1948.

[22] I. B. Crandall, *Theory of Vibrating Systems and Sound*, D. Van Nostrand, New York, NY, USA, 1972.

[23] M. R. Matausek and V. S. Batalov, "A new approach to the determination of the glottal waveform," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 6, pp. 616–622, 1980.

[24] M. Rothenberg, "A new inverse filtering techniques for deriving the glottal airflow waveform during voicing," *Journal of the Acoustical Society of America*, vol. 53, no. 6, pp. 1632–1645, 1973.

[25] "TMS320 User's Guide," Preliminary. Texas Instruments, Dallas, Tex, USA, 1985.

[26] "TMS32020 BOARD," User Manual. Loughborough Sound, June 1985.

[27] J. M. Loasby, "Real time transmission line modelling of the vacal tract using multiple digital signal processors," Ph.D. thesis, University of Nottingham, Nottingham, UK, 1992.

[28] A. Benkrid, T. E. Cross, M. Atmani, and M. Koudil, "A multiprocessor TMS320 platform for a real-time modelling," *IEEE Transactions on Engineering Management*, vol. 46, no. 6, pp. 174–178, 1999.

AUTHOR CONTACT INFORMATION

A. Benkrid: Electronic Engineering Department, College of Technology, Dammam 31472, P.O. Box 7650, Saudi Arabia; abdelhak_benkrid@hotmail.com

A. Benallal: Electronic Engineering Department, College of Technology, Dammam 31472, P.O. Box 7650, Saudi Arabia; a_benallal@hotmail.com

K. Benkrid: School of Engineering and Electronics, University of Edinburgh, Edinburgh EH9 3JL, UK; k.benkrid@ed.ac.uk