OPTOPALATOGRAPH: REAL-TIME FEEDBACK OF TONGUE MOVEMENT IN 3D

A. A. Wrench, A. D. McIntosh, C. Watson, W.J. Hardcastle

Department of Speech and Language Sciences, Queen Margaret College,

Clerwood Terrace, Edinburgh, EH12 8TS

ABSTRACT

In this paper the latest prototype Optopalatograph (OPG) is described and its operation is demonstrated graphically and in comparison to theoretical predictions. The system is divided into three parts - the optopalate itself; a separate self contained unit composed of 16 switched infra-red light sources, associated control logic and 16 receivers; and a computer with A/D card running software to analyse and interpret graphically the sensor outputs. The current prototype measures distances of up to 20mm between all of the 16 pre-selected points on the hard palate and the surface of the tongue at a frame rate of 100Hz. We conclude that the new prototype provides a practical measurement system with a subjectively informative real-time display but further development is required in order to obtain objective accuracy.

1. INTRODUCTION

If we believe that "to teach a speech sound target one should 'know' how the sound is made" [3] then the Optopalatograph (OPG) can be considered as an extension of the work of Chuang and Wang [5] and Fletcher *et al* [4] as a further step towards that goal. One of the primary functions of the instrument is to make visible lingual activity that is concealed within the oral cavity. The OPG, based on sensors embedded in an artificial palate, directly measures and relays information about the size and shape of the air cavities defined by the tongue palate configuration in the important region of the vocal tract that lies between the front and rear of the hard palate.

In this paper we have chosen to display the 3-dimensional oral cavity measures in the form of a selection of simultaneous 2-dimensional displays. These displays provide an applied focus for evaluation of the 3-D potential before proceeding to the more complex 3-D cavity volume or tongue surface modelling processes. These displays have the benefit of simplifying the 3-D information for the viewer. We are able, in this way, to represent the configuration of the tongue with respect to the hard palate more clearly and focus on regions of interest.

2. THE PROTOTYPE

The optopalatograph (OPG) is a device which provides direct measurement of the distance between the hard palate and the tongue. Figure 1 shows a block diagram of the overall design of the latest prototype. In contrast to its predecessor [2], there are individual high intensity infra-red LED sources for each sensor and all sources are switched in sequence. This avoids ambiguity in distance sensing due to unknown proportions of light reflected from neighbouring sensor sources.

The system operates at a data capture frame rate of 100Hz although graphical feedback is somewhat slower. For visual feedback frame rates of 20Hz are adequate.

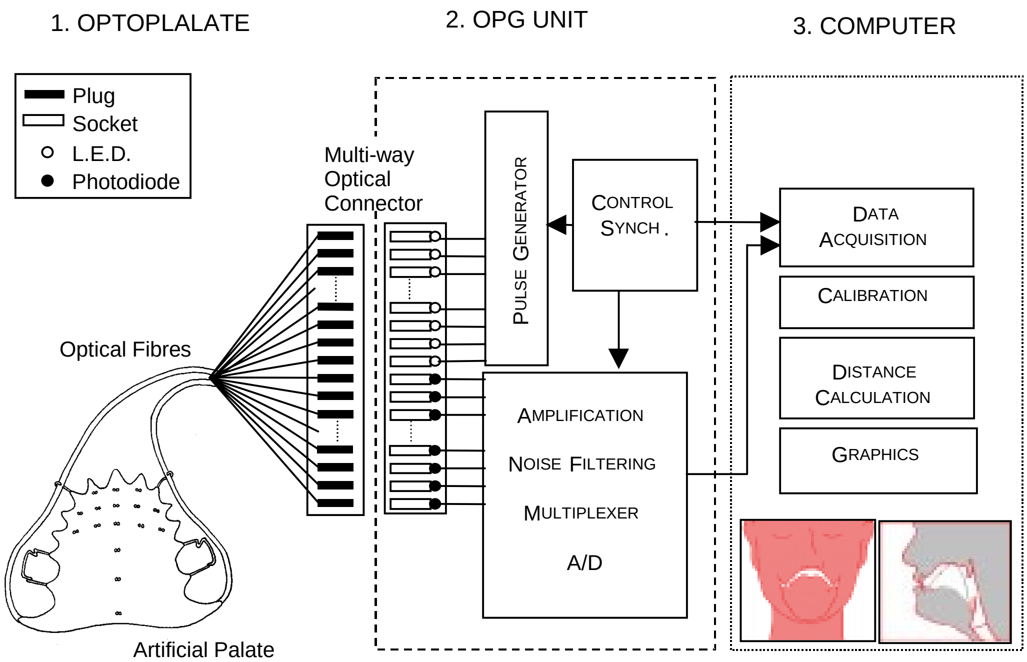

Figure 1: Block diagram describing the general configuration of the optopalatograph

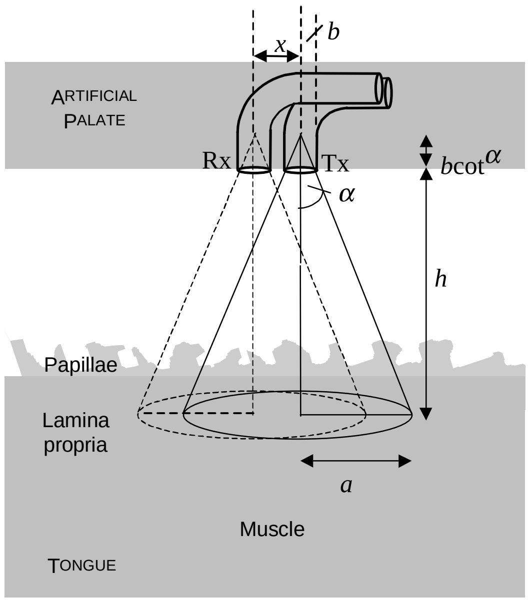

Figure 2: Model of distance sensing using optical fibres showing a transmitting (Tx) and receiving (Rx) fibre of radius *b* and separation *x* with beam spread of α producing a spotlight of radius a at a distance *h* from the ends of the fibres.

The palate is connected to the unit containing the sensor electronics via optical fibres rather than wires. The advantage of this method is the low complexity and cost of the palate construction. It has the disadvantage that there are discontinuities in the optical channel causing greater light loss than an arrangement where the optical fibres are bonded with refractive-index matched adhesive to the emitter and photodiode. However, if the latter approach is taken, not only does the cost increase but also the presence of electrical power on the client side of the device raises safety issues. Each sensor consists of a transmitting and receiving fibre of 0.5mm diameter. It is likely, in future, that smaller diameter fibres may be used, leading to the possibility of more sensors [2].

The OPG unit handles the light pulse output and received signal amplification and conditioning circuitry. The computer is used to interpret the received signal as a distance measure and provide the graphical output.

2.1. Modelling the Distance Sensing Function

We have calculated the received power in terms of the radius (*b*) of the transmitting and receiving fibres, the separation (*x*) of the transmitting and receiving fibres, the beam spread (α) from the transmitting fibre, the surface reflectivity (ρ) and the distance (*h*) from the reflecting surface (Figure 2).

In our model we assume that the light rays exiting the fibre do so as if they originate from a source point a distance $b\cot\alpha$ beyond the end of the fibre, which results in a beam, spread angle of α. I.e. radius (a) of the beam at height *h* is $b+h\tan\alpha$.

Integrating over the illuminated surface gives the gives equation 1

$$P_1 = P_0 k \frac{h}{\left(h^2+x^2\right)^{\frac{3}{2}}}\left(1-\phi\right) \qquad \text{Eqn. (1)}$$

where k is a constant which depends on the fibre diameter, the reflectivity of the surface and the angle of beam spread.

$$k = \frac{\rho b^2}{4}\frac{\tan^2\alpha}{1-\cos\alpha} \qquad \text{Eqn. (2)}$$

and where

$$\phi = \frac{3}{4}\left(\tan^2\alpha + \frac{a^2}{2}\frac{\left(2h^2-3x^2\right)}{\left(h^2+x^2\right)^2}\right) \qquad \text{Eqn. (3)}$$

and

$$a = b + h\tan\alpha \qquad \text{Eqn. (4)}$$

The term ϕ can, to a first approximation, be considered a constant with a value of

$$\phi \approx \frac{3}{2}\tan^2\alpha \qquad \text{for } x < h \text{ and } h > b \qquad \text{Eqn. (5)}$$

From a 0.5mm optical fibre we estimate the angle of beam spread to be approximately 25°. Using this value for α and using the constant approximation for ϕ

$$k(1-\phi) = 0.024 \qquad \text{Eqn. (6)}$$

Eqn. (1) assumes that all the light from the emitter can be reflected into the receiver but the receiver has a limited receiving angle and so for small *h* only a fraction of the reflected light is received. If we assume that the receiving angle of acceptance is the same as the emitter spread then P in Eqn. 1 must be modified according to the overlapping fraction given by:

$$P = nP_1 \qquad \text{Eqn. (7)}$$

Where

n = overlap area / illuminated area.

$$= \left(2a^2\cos^{-1}\left(\frac{x}{2a}\right) - x\sqrt{\left(a^2-\left(\frac{x}{2}\right)^2\right)}\right) \Big/ \pi a^2 \qquad \text{Eqn. (8)}$$

Testing 0.5mm and 0.25mm fibres we found that the latter resulted in about a 4-fold decrease in the sensing range. With the current Tx/Rx components this provided a best case range of about 5mm which is short of the target performance. The dynamic range for distance measuring with a pair of 1m length, 0.5mm diameter fibres was determined with the optimum emitter setting and using the photoreceivers provided. Figure 3 shows a typical response curve with a theoretical model curve superimposed and two further model curves showing the effect of changing the separation value (x).

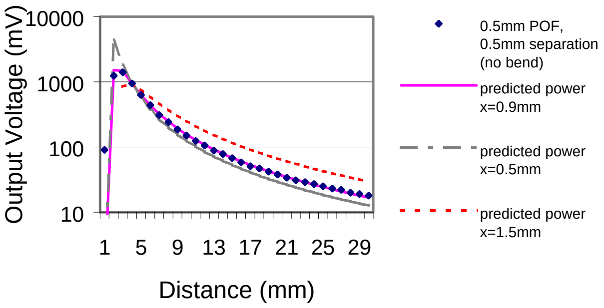

Figure 3: Measured vs. predicted distance sensing function for 3 values of separation (x) with α = 0.43 rad (25°); b = 0.25mm

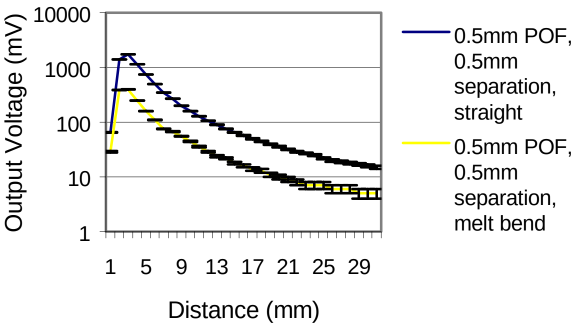

Figure 4: Reflected light intensity with and without a 90° bend. The error bars show a +/- 1mV noise level present on the raw output signal. This typical melt-bend pair shows an effective range of about 20mm but each individual pair currently varies somewhat in the range they achieve between 15 and 25mm.

The fibres also required a tight 90 degree bend to be made in the fibres at the point on the palate where they exit to form the sensor. As can be seen from the graph in Figure 4, there is appreciable loss in performance due to light leakage at the bend. Electrical and optical noise in the current prototype is approximately +/- 1mV and this noise is indicated by the error bars in Figure 4. If the desired accuracy of the system is +/-5% then the range of the current system is approximately 20mm (i.e. we accept +/- 1mm at 20mm) If we impose a tolerance of +/-0.5mm for all distance measures then the effective range is only 12mm.

We are aiming for a system which has a range of 25mm +/-10% and there is wide scope for improvements which lead us to believe that we will meet and exceed this target performance.

3. VISUAL FEEDBACK

3.1. The Dynamic Display

Previous displays using palate-based optical distance sensing built by Chuang and Wang [5] and Fletcher *et al* [4] were restricted to midsagittal views. Both systems used linear markers to indicate tongue height. By contrast, the OPG prototype offers the opportunity for a variety of 2-D and 3-D displays as sensors can be placed at any point on the hard palate.In order to demonstrate the 3-D capability of the device without engaging in the complex task of volume or surface modelling, the prototype displays two orthogonal cross-section views of the oral cavity. Figure 6 shows three pairs of snapshots from a dynamic 2D display of the midsaggital section using 4 sensors and a coronal post-alveolar section using 5 sensors. The location of the sensors on the Optopalate is shown in Figure 5.

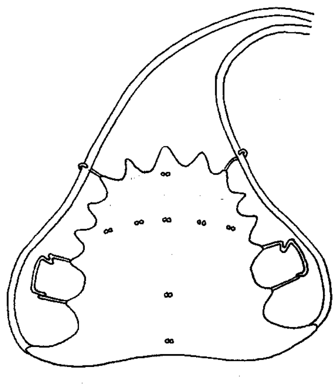

Figure 5: Diagram of the Optopalate used for the demonstration displays showing the active sensor positions.

A small circle indicate the distance of the reflecting surface from each sensor within hand drawn sections.

Bezier curve fitting is used in this demonstration display to enhance the subjective feedback. Curve fitting is applied to 4 midsaggital sensor distances and a further 3 fixed points to provide an interpretation of the whole tongue shape. In the coronal section display, curve fitting is applied to the 5 measurement points and 2 fixed end points. Bezier curve fitting was used for programming convenience but is not ideally suited to this task. The coronal section for /k/ shows a discontinuity at the central sensor because this sensor is used as an anchor point for two independently calculated curves. The Bezier algorithm uses two anchors at each end of the curve with two interval measures used to influence the curve shape. This results in curves that do not necessarily intersect the measurement points. The effect of this can be seen in Figure 6. For example, the curve for the midsaggital display of /t/ does not touch the palate. Other curve fitting algorithms are now being developed which will fit closer to the measurement points and avoid any discontinuities.

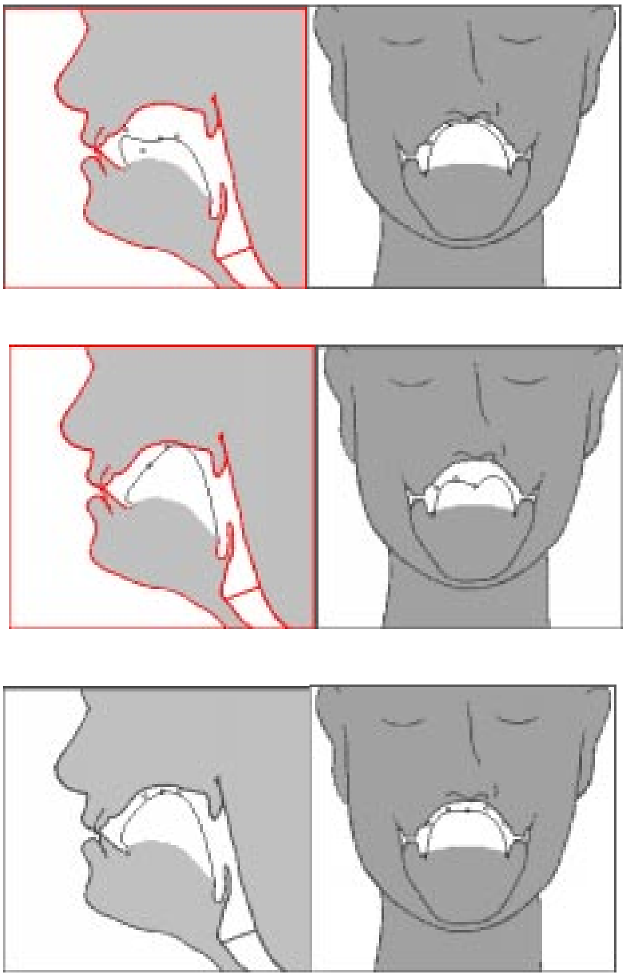

Figure 6. Captured stills from dynamic 2-D displays of tongue configured for the phoneme /t/ and /k/ and /i/.

3.2. Calibrating the System

The system can be calibrated in two stages. Firstly an *ex vivo* test is performed to establish the fibre separation, *x,* for each sensor pair. Secondly, an *in vivo* measurement is made, prior to each recording, to establish the gain factor due to the reflectivity of the individual's tongue.

For this feedback application we calibrate the sensors by getting the subject to rest the tongue on the palate and zeroing all the sensors. We have modelled the received power curve with four variables: fibre radius, fibre separation, beam spread angle and surface reflectivity coefficient. For each sensor embedded in the palate, the fibre radius is known. The separation and beam spread can be established by measuring the sensor output at a number of distances. The values are initially estimated along with an appropriate gain factor and then empirically adjusted for best visual fit. Having established the varaibles associated with the palate construction, tongue reflectivity of the individual remains as a single free variable. So with a single measured value we can predict the distance corresponding to any output intensity reading. We are currently trying to establish the reliability of this calibration technique and the variability of the reflectivity of the tongue.

4. CONCLUSIONS

In previous papers [1][2] we showed that the principle of distance sensing by reflected light intensity was sound but we had not established that the idea could be turned into a practical instrument that was suitable for use in the speech science laboratory or clinic.

With this prototype we have demonstrated the ability to show 3-D articulatory data in the form of two simultaneous orthogonal 2-D displays using only 9 out of 16 available sensors. It is clear that the potential exists to use all 16 points to reconstruct 3-D volumes.

Even in its current unrefined form, the display is able to indicate the position of the tongue dynamically during speech in a way which is very intuitive. In this form, the OPG has the unique ability to be used to display key aspects of vowel production *and* consonant production.

Some aspects of the display methodology require improvement. For example, palate shape could be measured for each individual and included as part of the display. We intend to carry out further development to improve all aspects of the design (mechanical, electronic, software) in order to enhance the accuracy, comfort and convenience of the system.

5. ACKNOWLEDGEMENTS

Development of the Optopalatograph has been supported by The Royal Society, Paul Instrument Fund.

6. REFERENCES

1. Wrench, A.A., McIntosh, A.D., and Hardcastle, W.J.,"Optopalatography: A new apparatus for speech production analysis", *Proc. Int. Conf. Spoken Lang. Processing*, Vol. 3, 1589-1592, 1996
2. Wrench, A.A., McIntosh, A.D., and Hardcastle, W.J., "Optopalatography: Development of a Device for Measuring Tongue Movement in 3D", *Proc. Int. Conf. on Speech Communication and Technology (Eurospeech97)*, Vol. 2, 1055-1058, 1997.
3. Klein, R.M., Lederer, S.H., and Cortese, E.E., "Children's knowledge of auditory/articulatory correspondences: Phonologic and metaphonologic." *Journal of Speech and Hearing Research*, 34, 559-564, 1991.
4. Fletcher, S.G., Dagenais, P.A., and Critz-Crosby, P., "Teaching vowels to profoundly hearing-impaired speakers using glossometry", *Journal of Speech and Hearing Research*, 34, 943-956, 1991.
5. Chuang, C-K., and Wang, W.S-Y., "Use of optical distance sensing to track tongue motion", *Journal of Speech and Hearing Research*, 21, 482-496, 1978.