Proceedings of the Institute of Acoustics

USE OF PALATE SHAPE DATA IN AN ENHANCED ELECTROPALATOGRAPHY SYSTEM

W S C Chiu and C H Shadle

Department of Electronics and Computer Science, The University, Highfield, Southampton, SO9 5NH, UK

1. INTRODUCTION

In working towards improved models of speech production, it is necessary to distinguish between errors in the shape-to-sound conversion algorithm and errors in the initial estimates of vocal tract shape; we wish to minimize the latter so we can improve the former. Further, it now appears that for sounds such as fricatives the conventional area function representation may not be sufficient; details of the shape of at least some parts of the vocal tract can have a big influence on the sound produced [1]. For both these reasons, more accurate vocal tract shape data are needed. There is a variety of methods that provide information on the vocal tract shape. The technique of Electropalatography (EPG) is cheap, relatively simple, non-invasive and highly informative [2]. Using EPG on its own, it is possible to deduce information about the shape, movement and position of tongue-palate contact during continuous speech. However, the two-dimensional representation of the tongue-palate contact patterns does not provide direct measurement of the vocal tract area. In this paper, we describe the development of an enhanced electropalatography (eEPG) system, which retains most of the advantages of electropalatography while overcoming some of the disadvantages by representing the three-dimensional (3D) shape of the palate. By combining EPG data obtained during speech with 3D coordinates of the EPG palate obtained separately, we can derive a 3D time-varying tongue contact pattern. This is one step closer to our goal of obtaining the actual vocal tract shape during speech by a safe and inexpensive method.

The principles of EPG are conceptually very simple. An artificial plate moulded to fit the upper palate is worn by the subject. For the Reading system, as used in this study [3], the artificial plate has 62 electrodes mounted on the surface for detection of lingual contact. A computer is used to display the lingual contact patterns in real-time and to store the patterns for display, or to provide a permanent record of the patterns as a printout. Patterns of lingual contact are shown on small palate diagrams, as shown in Fig. 1. EPG has been increasingly used in speech research to provide information about acoustically significant aspects of the contact pattern such as location of a constricted region. In the past few years, some feature coding methods [4] have been introduced to try to reduce the amount of data while retaining the important aspects.

In this paper we discuss the development of the enhanced EPG system. We compare its output to that of conventional EPG for two subjects; since every palate shape is unique, the use of a 3D display allows any differences between individual palates to be seen where this is not possible on a 2D representation. The contact patterns can be related more easily to articulatory features such as the alveolar ridge since the ridge is visible on the 3D display. Finally, we show that feature coding of the patterns can be more sophisticated since it can be both related to articulatory features and expressed in absolute distances.

```
.00.000.        .00.00..
000.0000        000.0000
00....00        000.0000
0......0        00...000
0......0        00...000
0......0        00....00
00....00        00....00
00....00        00....00
```

Figure 1: EPG contact pattern of subjects WC (right) and CHS (left) producing the sustained fricative /s/. On each palate diagram the top row is the most anterior and the bottom is most posterior. Actual lingual contacts are indicated by zeros.

2. DEVELOPMENT OF ENHANCED EPG

2.1 Acquisition and Display of 3D Palate Shape

One disadvantage with EPG is that the data obtained are shown in a two-dimensional representation where the depth information is lost. One of the aims of the development of the eEPG system is to address this problem by using a 3D representation of the palate shape. In order to do this, we must bring in EPG data; we must obtain 3D coordinates of the palate shapes; we must use these coordinates to reconstruct the palate shape which we can draw on the computer screen as a graphical object. Lighting is then provided to the palate shape to give it colour and the illusion of three-dimension. Finally, we must show EPG contact patterns on the palate shape. Efforts to develop an eEPG system in the Vision, Speech and Signal Processing (VSSP) Group at the University of Southampton began in 1989. For the early eEPG system, the 3D coordinates and contact locations were measured manually from the artificial palate of one subject CHS. The graphical computation was done on a μVax computer. For the current eEPG system, the 3D coordinates are acquired automatically using the colour-encoded structured light (CESL) system [5, 6]. The computation has been changed to Silicon Graphics high-performance graphics UNIX workstations. The workstations have special hardware that is optimised for high speed graphics operations, and a set of very powerful routines called the Graphics Library that provide support for real-time interactive graphics.

The eEPG display system accepts a 3D set of coordinates describing the shape of the palate. These coordinates are then processed to represent the whole surface. In order to do this, it is necessary to derive a set of small planar elements that approximate to that surface; this is the process of triangulation. The simplest solution is to divide the surface into abutting triangular tiles by selecting groups of three points to make up these tiles. Automatic triangulation for a generic set of 3D coordinate data is very involved and computationally heavy. The solution implemented has been optimised for this application only. Once the triangulated surface is rendered, the palate shape can be displayed as a shaded opaque object. For the eEPG system, the palate shape is lit from both sides with different colour sources making it easier to distinguish the top and bottom sides of the palate shape. The palate shape can be rotated around the $x - y - z$ axes in real-time by using a mouse.

The CESL technique involves projecting a known pattern of light, in this case a grid of parallel stripes, onto an object and measuring the distorted pattern using a video system. If the geometry of the camera, the light source and the position of the pattern in the measured picture is known, the exact coordinates of each recognised point in the pattern can be calculated. It is then possible to build up a 3D representation of the subject. A colour-coded light pattern is used because the identity of each stripe must be known uniquely so that one can measure the extent of its displacement from the reference plane.

Data from the CESL system is more highly structured than shape data acquired manually: the x and y coordinate data lie on a regular grid and the corresponding depth information (z coordinate) varies according to the way in which the colour stripes are distorted by the object. The palate shape reconstructed using the CESL data has higher resolution than the early system. The palate shape now consists of approximately 10000 tiles, as opposed to approximately 600 tiles using the earlier method.

2.2 Representing the EPG contact pattern on the 3D palate

The next stage in the development of the eEPG system was to superimpose dynamic EPG contact patterns on the 3D palate shape. EPG frame data provides information on which electrodes are contacted during continuous speech. The aim here is to develop a representation of EPG contact patterns on the 3D palate shape. The first problem is locating the position of contact electrodes on the palate shape. These were hand-measured for the early system. In the current system, the contact positions are generated semi-automatically as post-processing using the CESL acquisition system. The contact positions are not acquired with the palate shape data because contact electrodes are made from thin reflective silver discs and this presents a problem for the optical stage of the CESL system. This will be rectified in future and contact positions will be acquired automatically using the CESL system along with the palate shape data.

Since one of the aims of the eEPG system is to arrive at an estimate of the vocal tract shape, we would like to use the EPG contact patterns on the 3D palate shape to estimate the actual area contacted by the tongue. However, it is non-trivial to outline a contacted region in a pattern consistent with tongue anatomy and ballistics. We must first solve the problem of simply showing the contact pattern while retaining a sense of depth and keeping computation cost down.

Several methods have been used to represent EPG patterns. These methods include showing only the location of electrodes when contacted, and showing some defined region around the contacted electrode. The first method involved placing clear or black circles at contact locations on the 3D image of the palate. It was found, however, that the circles did not looks as though they were attached to the surface, and if large parts of the palate were contacted (so that most of the circles were black) the illusion of depth was destroyed. For the second method, abutting quadrilaterals were defined so that each quadrilateral surrounded an electrode and the coordinates of the vertices are the geometric means of the coordinates of the four electrodes around each vertex. If an electrode was contacted, its quadrilateral was coloured. Although this method gave a rough approximation to the tongue outline (albeit with a jagged edge), the illusion of depth was again lost in the coloured region. Another way to represent EPG patterns is to assign a number of tiles to represent each contact electrode. This set of adjacent tiles is called a contact patch. The patch is then displayed in a different colour to that of the palate shape if its electrode is contacted by the tongue. The neighbouring tiles method has the advantage of being compatible with the existing geometry system; there is no need to create other shapes to represent the patches. This reduces computation cost and is suitable for animated display. The disadvantages with this method are that the patches have sharp corners and the smooth outline of estimated tongue contact area is not achieved.

In the current system with decreased tile size the patch defined by all tiles touching a contact is likewise decreased in size. Colouring all contact patches will thus give an even poorer approximation to a contiguous tongue-contact region. Since the size of a tile in the current system is roughly the same as that of a contact electrode, the current method is to assign only one tile to each contact electrode and to display this tile with a different colour when the electrode is contacted. An example from the current system is shown in Fig. 2.

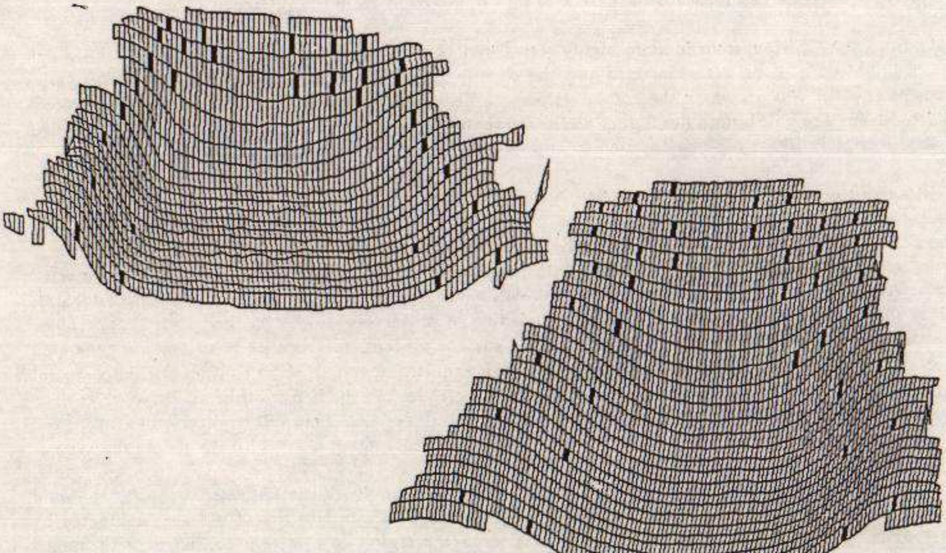

Figure 2: Contact pattern shown in Fig. 1 superimposed on palate shapes from current system. WC is on the right and CHS on the left, producing sustained fricative /s/.

3. COMPARISON OF EPG AND eEPG

Figure 2 shows contact patterns superimposed on two palate shapes from the current system. The palate shapes are shown as wire-frame meshes with darkened tiles representing contacted electrodes. The palate shapes are slightly tilted forward to provide a better sense of depth; note that the shapes can be rotated to any orientation in real-time on the computer screen. The first thing to note is the overall difference in the shapes. WC's palate appears to be wider than CHS's; this difference is confirmed by reference to the dental impressions. WC's palate shape is wider and deeper toward the posterior end, and the slope from the sides towards the middle of the palate is more gradual. The overall size of CHS's palate is smaller, the posterior middle part of the palate is flatter and the slope towards the middle of the palate is steeper than that for WC.

Note also some of the artifacts in the palate shapes; there are a few oddly shaped tiles along the perimeter of WC's palate, and there are some floating and extended tiles around the perimeter of CHS's palate. The oddly shaped tiles are caused when noisy data are considered as being part of the palate shape. The floating tiles are caused by the CESL digitisation system recognising the wire and metal clips on the

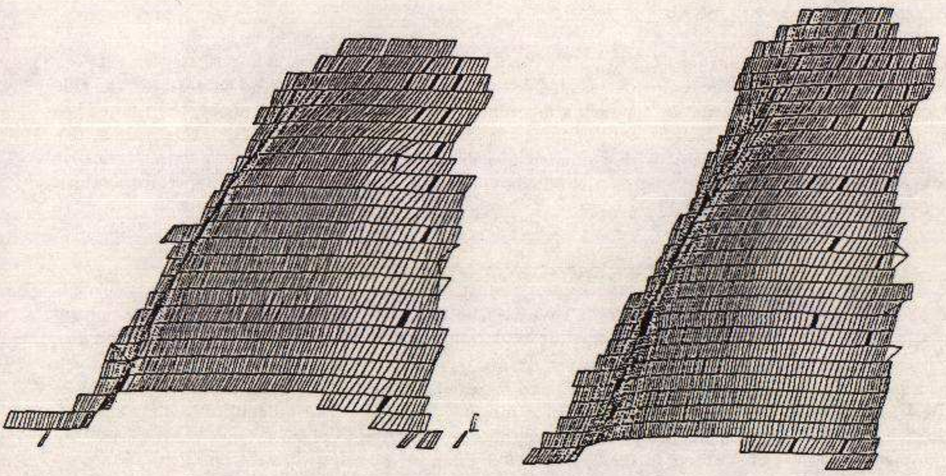

Figure 3: Contact patterns superimposed on palate shapes from current system. WC is on the right producing sustained fricative /s/ and CHS on the left, producing fricative /s/ in a VCV sequence.

artificial palate as being part of the palate shape.

Looking at the 2D EPG contact patterns in Fig. 1, we can deduce that WC has a longer constriction channel than CHS for the sustained /s/ and there is more tongue contact towards the sides of the palate for WC. This is confirmed when looking at the eEPG representation of the contact patterns on 3D palate shapes in Fig. 1. The constriction channel for WC appears to be longer, and along the sides of the palate a greater area is contacted by the tongue. Imagine the tongue forming these contact patterns. For WC, because of the shallow angle of the palate sides, the rounded shape and the depth of the palate, the surface of the tongue could be flat to produce this contact pattern. For CHS, however, since the inner posterior pairs of the contact electrodes are along the 'edges' of a relatively flat central palate, the tongue must clearly be grooved in order not to contact the central portion of the palate. We therefore see that patterns that appear similar with conventional EPG may actually appear different enough with enhanced EPG to suggest a different articulatory strategy for differently-shaped palates.

Figure 3 shows the same palate shapes but in a different orientation, where the palate shapes are upright and rotated slightly to show a side view. Note that in this figure the contact pattern for CHS is different: it corresponds to /s/ in the VCV sequence /asa/. The slope of the post-alveolar region can be partially seen in Fig. 3; for CHS the slope is steeper than that of WC's, and for WC the slope is more gradual and extends further along the palate resulting in a deeper central portion. It can be seen clearly from this angle that the constriction channel for WC is longer than that of CHS. It can also be seen that there are no contacts on the slope for CHS.

By looking at the palate shapes for CHS in both Figs. 2 and 3, we can observe a few differences between the sustained /s/ and /s/ in the sequence /asa/. For the /s/ in /asa/, there are no contacts on the side slopes of the palate. One explanation for such a difference is that for the VCV sequence, the tongue has to travel quickly from the low back vowel /a/ to articulate the alveolar fricative /s/. This may be the cause of the tongue being in contact with a smaller area in the posterior region of the palate. Note that in this case there are no differences in constriction channel width, length and location between sustained /s/ and /s/ in /asa/.

4. FEATURE CODING OF EPG DATA

In the past few years feature coding methods have been introduced to parameterize and extract useful information from EPG data. In a study aimed at exploring methods of parameterizing parallel palatographic and acoustic descriptions of the fricatives /s/ and /ʃ/ [7], five parameters were defined, of which two were position of maximum constriction (CP) and constriction width (CW). In a more recent study of fricatives [8, 4], EPG contact patterns were coded to represent changes in the amount of contact between tongue and palate for the front and back of the mouth. The front/back parameter was characterised by determining the number of contacts in the front three rows and the back five rows. The EPG data used in the study are part of a large database recorded by two subjects, of whom CHS is one, as part of a joint EEC project. Because of our access to the corpus we were able to compare features derived by Scully from EPG and our own features developed using eEPG.

The top graph in Fig. 4 shows a plot of the front/back parameters for the nonsense word /pisi/ by subject CHS. The number of contacts in the front three rows and the back five rows are plotted on the same graph. The beginning and end of the noise segment is marked in the plot. It can be seen that the back contact decreases during the frication noise segment, which may indicate a lowering of the tongue dorsum. It can also be seen that the front contact increases during the frication. This suggests that the constricted region may have shifted from the back portion of the palate to the more anterior part.

Two parameters, location of constriction (LC) and degree of constriction (DC), have been defined in the system described here and initial study of these parameters have taken place. The *location of constriction* is defined as the row that has the greatest number of electrodes contacted by the tongue. The *degree of constriction* parameter is an extension to the location of constriction parameter. For the most constricted row, as indicated by the location of constriction parameter, the degree of constriction is the number of electrodes not contacted by the tongue.

The algorithm for coding the LC and DC parameters is still under development. It accepts as input the EPG binary data file containing contact patterns and frame numbers from which analysis should begin and stop. The algorithm decodes and reconstructs the EPG contact patterns. It considers each contact frame: the row number with the maximum number of 'on' contacts is taken to be the most constricted row. The degree of constriction is the number of contacts that are not 'on' for that row. The algorithm also codes the Scully front/back parameter automatically. Initial testing showed that for some patterns the algorithm failed when there was a tie. A tie can happen when maximum number of contacts occur in more than one row. Modifying the algorithm so that it averages row numbers when a tie occurs provides the wrong result when tied rows are non-adjacent and are at opposite ends of the palate. The algorithm was again modified so that the averaging action only takes place when the tied rows are adjacent to

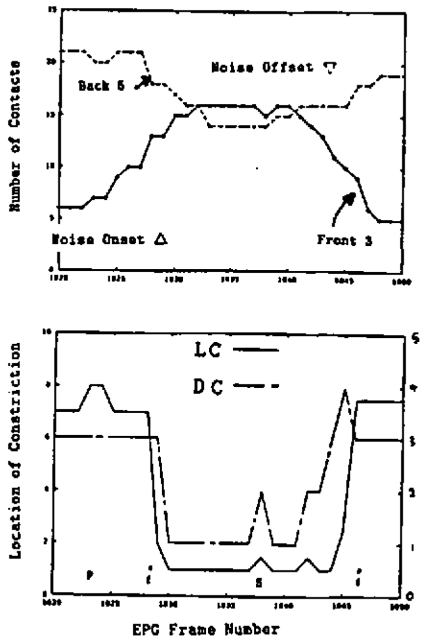

Figure 4: Top plot shows EPG contact traces for front and back contacts of /pisi/. Bottom graph show plot of parameters Location of Constriction and Degree of Constriction.

each other. Furthermore it was decided that the most anterior row, row 1, should have more weighting placed on it since it has the least amount of space between adjacent contacts. The algorithm does this by treating row 1 as having eight contacts although it only has six physical contacts, and if any contacts are 'on' in row 1 the extra two contacts are immediately considered as 'on' also.

The bottom graph in Fig. 4 shows a plot of the parameters LC and DC versus EPG frame numbers for the same item /pisi/. It can be seen from the plot that the location of constriction moves from the back of the mouth (rows 7 and 8) to the front (rows 2 and 1) during the frication noise segment. For the most constricted rows, the degree of constriction increases from around 3 'off' contacts to just 1'off' contact, which is a very narrow opening, during the frication noise segment. This suggests that the most constricted region not only shifts anteriorly, it also gets narrower during frication.

The feature coding algorithm is still under development and at present there may be cases where the algorithm will fail to pick out the most constricted region correctly. It will be necessary to build more rules into the algorithm before it can be made to code automatically and provide correct information. Also, although eEPG was used to check the results of each algorithm, none of the 3D information is at present used in the coding algorithm. Sudden changes in the slope of the palate are clearly significant articulatorily and acoustically, and should be incorporated. The parameters could also be coded in terms of absolute distance instead of contact units. Parameters such as LC and DC, when generated in millimetres, will provide information on the dimensions of the constriction which is useful for fricative production models. Indeed, some features, such as abruptness of constriction, are difficult to quantify without using absolute distances, as discussed by Hoole *et al* [6].

5. SUMMARY

In this paper we have described the development of the eEPG system. We have discussed why superimposing EPG contact patterns on a 3D representation of the palate shape is useful. Steps for enhancing EPG were outlined; they include the need to acquire 3D coordinates of the palate shape, to represent the palate shape on the computer screen and to represent superimposed EPG contact patterns. We have shown examples contrasting conventional EPG and eEPG and have shown that more information can be obtained using eEPG. Finally we have discussed feature coding of EPG contact patterns. We have shown how information can be extracted from contact patterns using acoustically relevant parameters. By studying a combination of these parameters along with using the eEPG system it will be possible to build up a picture of the tongue-palate contact area.

Acknowledgements: this work was supported in part by a Croucher Foundation Fellowship and by an SERC grant number GR/G 21155.

6. REFERENCES

[1] Shadle, C.H. (1991) "The effect of geometry on source mechanisms of fricative consonants", *Journal of Phonetics*, 19, pp. 409-424.

[2] Hardcastle, W. (1972) "The use of Electropalatography in Phonetic Research", *Phonetica*, 25, pp. 197-215.

[3] Hardcastle, W., Jones, W., Knight, C., Trudgeon, A. and Calder G. (1989) "New development in Electropalatography: a state-of-the-art report", *Clinical Linguistics & Phonetics*, 3(1), pp. 1-38.

[4] Scully, C. Grabe-Georges, E. and Badin, P. (1991) "Movement paths: Different phonetic contexts and different speaking styles", to appear in: *Proc. of Symposium on Current Phoentics Research Paradigms: Implications for Speech Motor Control, Stockholm.*

[5] Monks, T., Carter, J. and Shadle, C. (1992) "Colour-Encoded structured light for digitisation of real-time 3D data", Proc. of 4th Int'l. Conf. on Image Processing and Its Applications, Institution of Electrical Engineers, pp 327–330.

[6] Monks, T., Shadle, C. and Carter, J. (1989) "Measuring the shape of the mouth using structured light", IEE Colloquium on Biomedical Applications of Digital Signal Processing, Digest No. 1989/144, London.

[7] Hoole, P., Ziegler, W., Hartmann, E. and Hardcastle, W. (1989) "Parallel electropalatography and acoustic measures of fricatives", *Clinical Linguistics & Phonetics*, 3(1), pp. 59-69.

[8] Scully, C., Georges, E. and Castelli, E. (1991) "Fricative consonants and their articulatory trajectories", *Proc. of XIIth Int'l Congress of Phonetic Sciences, Aix-en-Provence*, 3, pp. 58–61.