Proceedings of the Institute of Acoustics

CYBERSPACE VISUALISATION OF VOCAL DEVELOPMENT DATA

David Rossiter and David M Howard

Signal Processing : Voice and Hearing Research Group
Department of Electronics, University Of York, Heslington, York, YO1 5DD

ABSTRACT

Developments in the field of data visualization for scientific research are continuing at an ever increasing rate. Within this context, there is growing interest in the concept of cyberspace as a medium within which data may be viewed and manipulated. Visualization techniques, popular in other scientific disciplines, appear not to have been previously applied to vocal data analysis. This paper considers cyberspace representations of data determined from longitudinal recordings of students undergoing regular singing training. Approaches taken towards encapsulating this data within a cyberspace domain for visualisation are described, together with an assessment of associated benefits and drawbacks. With regard to the cyberspace paradigm, consideration is given to the possible design of future real-time visual feedback systems for use with singing students.

1. INTRODUCTION

As part of ongoing work towards the development of real–time visual displays for voice development, regular electrolaryngographic recordings [1] have been made of singers undergoing training. Through analysis of specific qualities (listed in figure 1) it is hoped to discern trends which can be quantised and utilised in the implementation of real–time visual feedback systems for voice development. Statistical analysis of the data [2] provides valuable information towards this goal, but only within the confines implicit in the use of non–immediate pre–determined paths of data analysis. A more refined, responsive and flexible visualisation should help to verify these trends, and also provide clues towards subtler patterns implicit within the data.

The concept of cyberspace was first introduced in an acclaimed science fiction novel [3] in which the author describes use and abuse of the medium and the sub–cultures that arise. Cyberspace was originally conceived as a data manipulation and viewing medium accessible by millions of people throughout the world – a "mass consensual hallicination", a space that does not exist in a physical sense but which is brought into being through artificial stimulation to the senses of the human body. Gibson wrote of "...unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data", a "...transparent 3D chessboard

VISUALISATION OF VOCAL DEVELOPMENT

Label	Description
CQ	Closed quotient
F0	Fundamental frequency
Amp	Amplitude
Time	Time passed
Gradient	Gradient of laryngeal slope

Figure 1 Attributes of electrolaryngographic recordings monitored for developmental trends

stretching to infinity". Several definitions of the medium itself exist, and formalised definitions of related protocols continue to emerge (for examples see [4] and [5]). A definition is quoted in part here:

> "*Cyberspace is a completely spatialized visualisation of all information in global information processing systems, along pathways provided by present and future communications networks, enabling full coprescence... allowing input and output from and to the full human sensorium... and total integration and intercommunication with a full range of intelligent products and environments in real space*" [6].

Explorations in scientific visualisation techniques have developed at a phenomenal rate over the last few years and hardware and software products able to satisfy the exacting demands of visualisation have now become readily available. Scientific visualization involves investigations into the examination and transformation of data with the aim of extracting a maximum of information from the data set. One concise definition is :

> "*The computer presentation, for human display and interaction, of large volumes of scientific data to gain a better understanding of the data*" [7].

The overlap between this field of study and cyberspace should be apparent. A fusion of the two can host an interactive and iterative process of discovery in which the user interacts with the source data to promote a clear interpretation of what the data reveals. This paper does not concern itself in the main with multi-participatory aspects of cyberspace which will demand processing power several orders of magnitude greater than current technology can provide. Instead, it concentrates on a single user subset of the model with regard to a specific area of data visualisation, that of the analysis of laryngeal recordings of singers in training – an application to which users of current technology *are* able to address themselves.

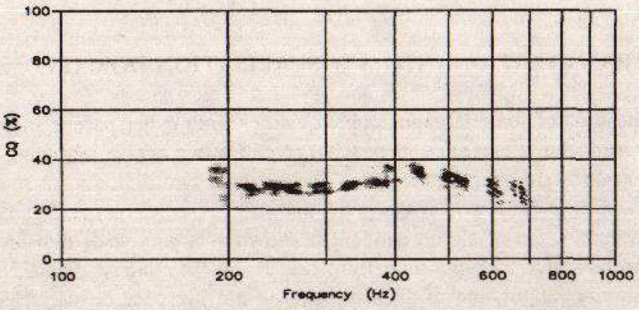
2. TRANSPOSITION INTO A VISUALISATION DOMAIN

An example of transposition of some laryngographic qualities into cyberspace is shown in figure 2. Data determined from a closed quotient (CQ) plot of a two octave ascending and descending G major scale is shown presented as a conventional two dimensional scattergram, where frequency is mapped to the x axis of the graph and CQ to the y axis (see figure 2a). The degree of statistical accumulation of closed quotient is shown as a grey scale measure, with greater levels of CQ represented by proportionally darker levels. In contrast, figure 2b shows one possible transposition configuration of the same data to a cyberspace domain. Here, a decision has been made to retain the same x–y mapping, but with CQ density projected into the third dimension, with the level of protrusion directly proportional to the degree of CQ density. The data has now become a tangible object, which the user is able to interactively move about and around (as indicated by the x, y and z rotation direction arrows) and further alter for maximum portent of meaning as required.

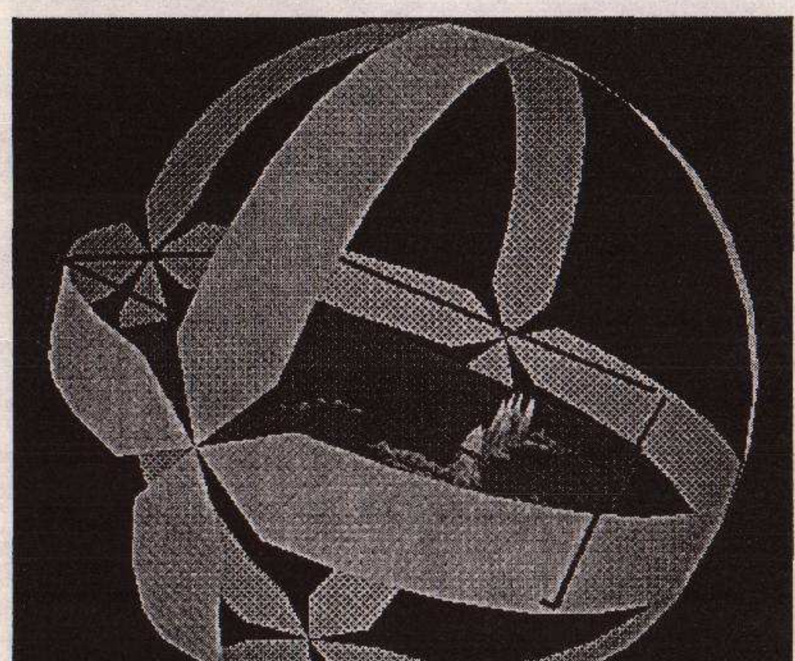
Figure 2a presents an amount of data in an accessible and informative format. For a certain degree of data examination, it provides a maximum of portent for a minimum of input. However, this low–interaction, single plane representation has limitations in comparison to an interactive cyberspace visualisation of the same data. For example, in figure 2a CQ is mapped to a grey scale, with weighting proportional to the density of sample data. However, this may result in low information resolution as the eye/brain couple is poor in precisely discriminating between neighbouring levels of similar greyness. In the simple cyberspace model extension of CQ into the third dimension enables precise interpretation of different CQ levels, as the human visualisation mechanism (heightened by the ability to zoom in and alter viewport, perspective and angle of elevation both specifically *and* implicitly, by heuristic, real–time 'flying' through space) is able to discriminate between a selection of linearly projected vector quantities with greater ease and precision than it can for continuous grey levels.

Colour has been utilised in figure 2b to further emphasise the gradients of CQ data. Through the interactive arrangement of colour schemes, the user can specifically accent or diminish the visual attraction of certain data ranges. Figure 3 shows one of the simplest mappings, with data mapped to an approximately linear grey scale. Research into using aural signatures as a further indicator of information exist (for examples, see [8] and [9]) although this is an emerging field. Both extensions of data representation need to be carefully employed to increase information density, and help accent important data. The creation of an aurally and visually *loud* fusion of cyberspace attributes is likely to be easy, but the creation of a cyberspace construct able to support *new* and *useful* insights is another task entirely.

VISUALISATION OF VOCAL DEVELOPMENT



(a) Closed quotient scattergram of male singer for G major two octave scale



(b) Transposition of the same data into a cyberspace domain

Figure 2 Closed quotient visualisation

VISUALISATION OF VOCAL DEVELOPMENT

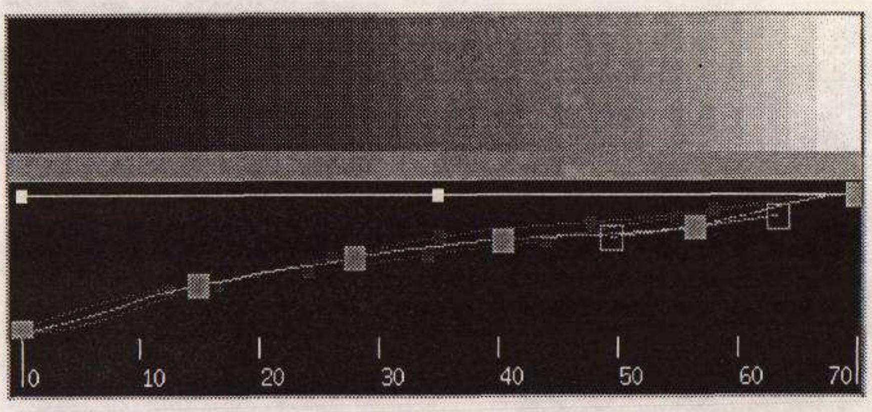
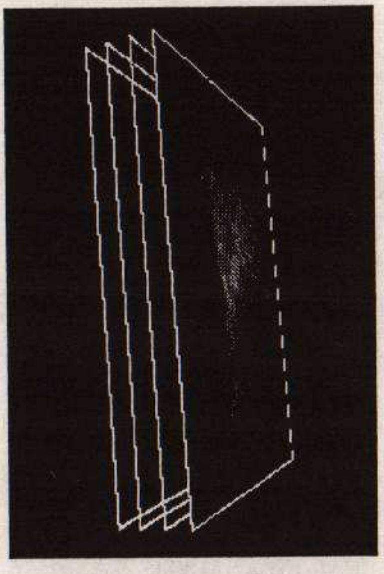


Figure 3 Mapping of data to a chromatic scheme
(in this example, red, green and blue attributes combine to produce a linear grey scale)

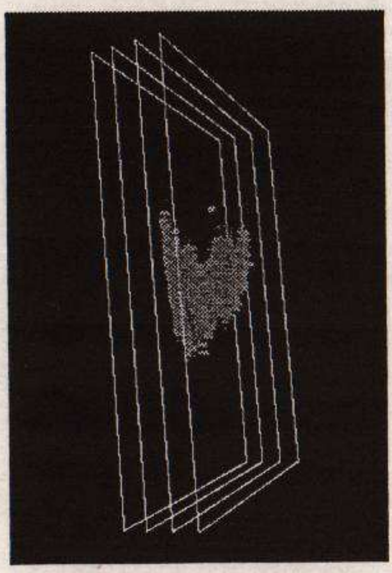
3. VISUALISATION OF LONGITUDINAL CQ ANALYSIS

This approach is only one of a myriad of possible techniques for use in considering longitudinal studies of CQ data.

A CQ scattergram is created for each of the longitudinal recordings, with CQ density represented by colour contouring. This data is placed in a lattice construct, and the sequence of these lattices ordered by time of the original recording (see figure 4a). Manipulative techniques can then be applied to the lattice sequence in order to help determine trends in the data. For instance, the cyberspace ethos of dimensionality, the creation of a tangible 3D object as representative of data, can be interpreted directly. Common cartographic techniques can be applied to the data, such as the determination of isosurfaces through the lattice sequence. An *isosurface* is a three dimensional equivalent to a contour, indicating a region of constant value in space. A low threshold isosurface would, for example, tend to highlight the 'base' of 3D protrusion of CQ resulting in a display of a larger diameter isosurface thus giving an indication of low accuracy in pitching. High thresholds will highlight precision areas of high intensity activity, with the highest thresholds revealing the maximum CQ density, apparent as a trend throughout the longitudinal recordings. This helps in providing clues towards discernable trends in the data. An example of an isosurface, with median threshold, is shown in figure 4b.



(a) CQ data represented in a lattice sequence



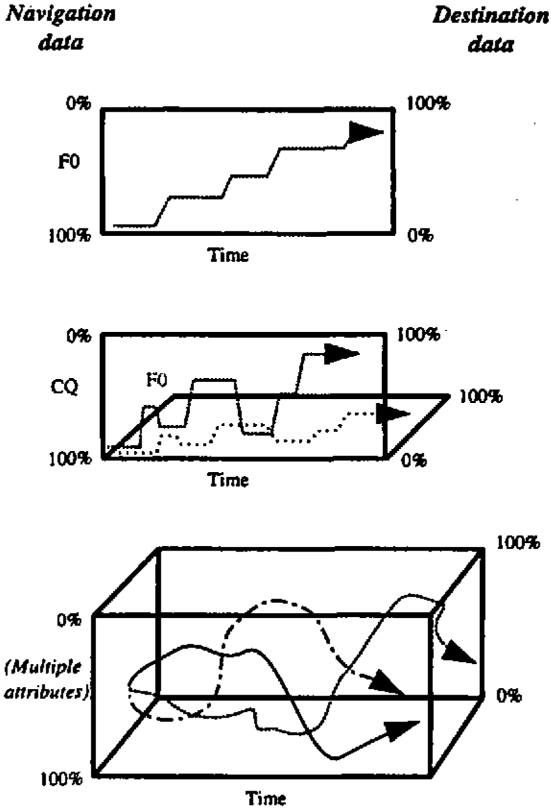
(b) Median threshold isosurface through CQ lattice sequence

Figure 4 Visualisation of longitudinal CQ analysis

4. THE FUTURE - SINGING LESSONS IN CYBERSPACE?

Previous work into real–time visual displays for singing development at York University has resulted in a sequence of several vocal quality monitoring systems. Each successive system has employed a further embodiment of dimensionality than the previous one (schematically outlined in figure 4). For example, the SINGAD software [10] supported a simple real–time plot of frequency against time, projected to the screen as a single flat plane (see figure 4a). More recently, work on a real–time laryngographically derived closed quotient display [11] has resulted in a system able to monitor and display both CQ and fundamental frequency against time, displayed in two seperate but parallel windows, each with a common time axis (see figure 4b). The logical extension of this progression of real–time feedback systems would be a cyberspace environment in which the desired subset of qualities (those listed in figure 1 and others both currently known and yet to be determined) are mapped to visualisation parameters. Thus a representative model is created which accurately portrays the development of quantified aspects of vocal output. This can be manipulated in real–time as the singer progresses.

VISUALISATION OF VOCAL DEVELOPMENT



(a) SINGAD fundamental frequency monitoring

One monitored and displayed vocal quality (the 'navigation data') leads in stages to the required goal (the 'destination data'), in this case pitching (and possibly vibrato) development.

(b) Real time closed quotient and fundamental frequency monitoring

Extends model (a) by a further plane in which to project CQ. Both F0 and CQ planes have a common x–axis, time.

(c) Cyberspace visualisation system

Model shows multiple destinations through the visualisation of many vocal qualities, at different distances and degrees of apparency. Navigation and destination data always available and have a maximum potential complimention to each other.

Figure 5 Schematic representation of real–time feedback systems for voice development

True cyberspace is a social medium permitting interaction between different users. Consequently, a singing lesson of the future might involve both pupil and tutor immersed in cyberspace, with both able to see representations of each other as well as adaptable visualisations of monitored vocal qualities of whoever is currently singing. As the scene is individually configurable, it would be possible, for example, for the tutor to select visualisation of certain qualities such as CQ and a spectral display, while the pupil only sees a fundamental frequency trace from which to aid the development of pitching accuracy.

VISUALISATION OF VOCAL DEVELOPMENT

5. CONCLUSIONS

This paper has briefly considered some of the visualisation issues predominant in the use of cyberspace domains with regard to the analysis of quantifiable trends readily monitorable in developing singers. With recent developments such as head mounted stereoscopic displays and tactile feedback data manipulation devices human–computer interfaces are moving towards a fully integrated synaesthetic interface. The power of scientific analysis through the medium of cyberspace looks set to continue to increase at the quasi–exponential rate it has already experienced. However, the processing techniques required to isolate and quantify the required parameters from the abundancy of data is likely to require at least as much development work in the future.

6. ACKNOWLEDGEMENTS

This work was supported by SERC award GR/H14878.

7. REFERENCES

[1] A J FOURCIN. 'Laryngographic assessment of vocal fold vibration', In: B WYKE, (Ed.) 'Ventilatory and phonatory control systems', Oxford: Oxford University Press (1974).
[2] D M HOWARD, & D ROSSITER. 'Results from a pilot longitudinal study of electrolaryngographically derived closed quotient for adult male singers in training' *these proceedings* (1992).
[3] W GIBSON. 'Neuromancer', Great Britain: Victor Gollancz, (1984).
[4] M SHOSWELL. 'Overview of Cyberterm, a Cyberspace Protocol Implementation', *unpublished paper*, (1992).
[5] M BENEDIKT. 'Cyberspace: Some Proposals', In: M BENEDIKT, (Ed.) 'Cyberspace: First Steps', MIT Press, pp119–224 (1991).
[6] M NOVAK. 'Liquid Architectures in Cyberspace', In: M BENEDIKT, (Ed.) 'Cyberspace: First Steps', MIT Press, pp225–254 (1991).
[7] J LATTA. 'A New Periscope', *Multimedia Review*, 2 (4)/3 (1) pp41–47 (1991/1992).
[8] G KENDALL.'Visualization by Ear: Auditory Imagery for Scientific Visualization and Virtual Reality', *Computer Music Journal*, 15 (4) pp70–73 (1991).
[9] G KRAMER, & S ELLISON. 'Audification: The Use of Sound to Display Multivariate Data', *Proceedings of the International Computer Music Conference*, pp214–221 (1991).
[10] D M HOWARD, & G F WELCH. 'Microcomputer–based Singing Ability Assessment and Development', *Applied Acoustics*, 27 pp89–102 (1989).
[11] D M HOWARD, & P GARNER. 'DSP–56000 based real–time electrolaryngographically derived closed quotient' *these proceedings* (1992).