A Study of Spectral Properties of the Voice Source

Peter Murphy

Department of Electronic and Computer Engineering,
University of Limerick, Limerick, Ireland.
Peter.Murphy@ul.ie

Abstract

An investigation of the relationships between the glottal flow waveform, and its derivative and their corresponding spectra are presented. The study presents a number of complimentary investigations; i) an interpretation of how the Fourier series coefficients are derived from the glottal waveform and its derivative, ii) Fourier series and Fourier transform estimates of synthesized glottal waveforms and derivatives, iii) an analytical approximation result and iv) investigation of the relationship between zeros in the glottal spectrum and open quotient (OQ) and asymmetry.

1. Introduction

An investigation of the spectral properties of the glottal flow and the voice source (derivative of the glottal flow) are useful for a number of reasons. The need for accurate voice source characterization is a recognized goal in voice analysis research. However, practical limitations prohibit the widescale use of a glottal source/vocal tract filter implementation. As yet no automatic inverse filtering technique is readily available for a range of voice types and problems exist with recording the phase accurately with a standard microphone in a general setting.

Frequency domain processing can be used to:

i)   replace time-domain processing and model frequency domain behaviour directly

ii)   overcome the problem of extracting the timing events of the glottal flow through implementing a frequency domain representation and parameterisation of the glottal flow waveform models (analytical spectral formulations for existing time domain glottal model are presented in [1] and [2]).

iii)   derive time domain model parameters through fitting spectra to best-fitting time domain model spectra (hence avoiding phase problems in the recording [3]).

iv)   supplement time-domain glottal processing e.g. compliment bandlimited measurements of glottal flow [4] or fine tune aspects of time domain processing e.g. estimating glottal parameters such as the return phase time from spectral details [5]

v)   find (or check validity of) correlates with time domain parameters (e.g. [6],[7],[8], H1-H2 has been used as an indicator of OQ)

vi)   check if symmetry and/or OQ can be predicted from the pattern of zeros in the glottal source spectrum [9]

vii)   make preliminary relations to the perception of voice quality [10]

The present paper addresses a number of the above issues. Based on a conceptual development involving taking the Fourier series for a single period of the glottal flow waveform with varying OQ and symmetry quotient, the corresponding spectral characteristics are hypothesized. Subsequent to this, the Fourier series coefficients are calculated for the synthesized glottal waveforms in order to test the hypotheses. Definite spectral differences are found for each parameter variation; based on these findings differential quantitative spectral measurements are suggested. Further supportive evidence is obtained through use of a Fourier transform analysis and analytical expression for approximate glottal waveforms.

2. Method

2.1. Spectral characterization

The possible spectral characterizations of each voice source measure are first postulated based on a graphical development that refers directly to the analytical expression for Fourier series expansion [11]. This is a useful exercise because it provides a graphical description of how the Fourier series gathers its spectral estimates and hence motivates an intuitive feeling for the time/frequency relationships under investigation. The spectra for the synthesis data are then produced i.e. Fourier series are evaluated in order to test the hypotheses.

In applying the Fourier series one period of a symmetrical glottal pulse with a total time record length *T* is considered. Asymmetry is introduced in later figures and the voice source is considered in later developments.

3. Results

It is reported that as the open quotient (OQ) for the glottal volume velocity (gvv) varies from 30% to 70% the H1-H2 ratio increases by 10 dB [12].

In the present graphical interpretation of the symmetrical pulse (to be presented at the conference) the following is surmised; As OQ varies between 0.0 and 0.25 both H1 and H2 (H2 is high) increase. As OQ varies between 0.25→0.5, H1 increases and H2 decreases. When OQ is 0.5, H1 is high and H2 is zero. As OQ varies between 0.75 and 1.0 both H1 and H2 decrease.

Next with OQ set at 50% the skewness of the waveform is examined. The overall implication is that as the differentiated glottal volume velocity, gvv varies from 0.25 to 0.50, H1-H2 increases and as OQ varies from 0.50 to 0.75, H1-H2 decreases.

For the differentiated glottal volume velocity, dgvv as OQ varies from 0.25 to 0.50, H1-H2 increases and as OQ varies from 0.50 to 0.75, H1-H2 decreases. However the waveshape is different for dgvv and it is F_g (=2/T_p, where T_p is time from glottal opening to peak flow, the glottal frequency) that is important. This highlights the primacy of the interval from zero to peak flow in determining F_g.

4. Discussion/Conclusion

The conceptual interpretation of the Fourier series for the glottal and voice source pulses with varying OQs and symmetry quotients reveal spectral characteristics as reported in the literature (in particular H1-H2 variation and spectral zero placement) and highlights the nature of the time-frequency relationship. The Fourier series coefficients and Fourier transforms are calculated for the synthesized glottal and voice source waveforms. The results support the graphical development. The analytical expressions supply further support. These issues have relevance to natural voice synthesis and possibly to vocal performance. Finally, a possible mechanism for octave jumping is suggested from the analysis.

5. Acknowledgements

6. References

[1] Doval, B. and d'Alessandro, C., 1997, Spectral correlates of glottal waveform models: An analytic study, ICASSP, 1285-1298.

[2] Walker, J. and Murphy, P.J., 2003, An analytical spectral formulation of glottal flow, *Proc. Irish Signals and Systems Conference*, Limerick, 364-367.

[3] Swerts, M. and Veldhuis, R., 2001, The effect of speech melody on voice quality, Speech Communication, 33(4):297-303.

[4] Holmberg,E.B., Hillman,R.E., Perkell, J.S., and Goldman,S.L., 1995, Comparisons among aerodynamic, electroglottographic, and acoustic spectrum measures of female voice, *Journal of Speech and Hearing Research* 38, 1212-1223.

[5] Fant, G. and Q. Lin, 1988, Frequency domain interpretation and derivation of glottal flow parameters, STL-QPSR, 2-3, 1-21, 1988.

[6] Henrich, N., d'Alessandro, C. and B. Doval, Spectral correlates of voice open quotient and glottal flow asymmetry: theory, limits and experimental data, *EUROSPEECH*, 2001.

[7] Hanson, H.M., 1997, Glottal characteristics of female speakers; Acoustic correlates, *J. Acoust. Soc. Amer.*, vol. 101, pp. 466-481, 1997.

[8] Sundberg, J., 1999, Effects of subglottal pressure variation on professional baritone singers' voice sources, *J. Acoust. Soc. Amer.,* 105:965-71

[9] Miller, RL., 1959, Nature of the vocal cord wave, *J. Acoust. Soc. Amer.*, 31(6):667-677.

[10] M. Epstein, B. Gabelman, N. Antonanzas-Barroso, B. Gerratt and J. Kreiman, 1999, Source model adequacy for pathological voice synthesis", International Congress of Phonetics Science.

[11] Murphy, PJ., 2000 Spectral characterisation of jitter, shimmer and additive noise in synthetically generated voice signals, *J. Acoust. Soc. Am.,* 107, 978-988.

[12] Klatt, DH. and Klatt, LG., 1990 Analysis, synthesis and perception of voice quality variations among female and male talkers, *J. Acoust. Soc. Amer.,* 87(2):820-857.