Real-Time Control of Greek Chant Vocal Synthesis

Ioannis Zannos
Ionian University
Department of Audiovisual Arts
Kerkyra, Greece
+30 26610 30406
iani@ionio.gr

Anastasia Georgaki
University of Athens
Department of Musicology
Athens, Greece
+30 210 7277583
georgaki@music.uoa.gr

Dimitrios Delviniotis
University of Athens,
Department of Informatics and Telecommunications
GR-15784 Athens, Greece
ddelvis@di.uao.gr

Georgios Kouroupetroglou
University of Athens,
Department of Informatics and Telecommunications
GR-15784 Athens, Greece
koupe@di.uoa.gr

ABSTRACT

This paper we report on an interdisciplinary project for modeling Greek chant with real-time vocal synthesis. Building on previous research, we employ a hybrid musical instrument: *Phonodeon* (Georgaki et al. 2005), consisting of a MIDI-accordeon coupled to a real-time algorithmic interaction and vocal synthesis engine. The synthesis is based on data provided by the AOIDOS program developed in the Department of the Computer science of the University of Athens, investigating Greek liturgical chant compared to *bel canto* singing. Phonodeon controls expressive vocal synthesis models based on formant synthesis and concatenated filtered samples. Its bellows serve as hardware control device that is physically analogous to the human breathing mechanism [georgaki, 1998a], while the buttons of the right hand can serve multiple functions. This paper focuses on a particular aspect of control, namely that of playing in the traditional non-tempered and flexible interval structure of Greek modes (ήχοι: *echoi*) while using the 12-semitone piano-type keyboard of the left hand. This enables the musical exploration of the relationship between the spectral structure of the vocal timbre of Greek chant and characteristic intervals occuring in the modal structure of the chant. To implement that, we developed techniques for superimposing interval patterns of the modes on the keyboard of the phonodeon. The work is the first comprehensive interactive model of antique, medieval and modern near-eastern tunings. The techniques developed can be combined with techniques for other control aspects, such as timbre and vocal expression control, phoneme or (expressive/ornamental/melodic pattern, inflection) sequence recall and combination, data record on/off, or others, which form part of the phonodeon project. This research can find applications in many computer music fields such as algorithmically controlled improvisation, microtonal music, music theory and notation of (algorithmic/computerized) real-time performance, and computer modeling of experimental or non-western musical styles.

Keywords

Vocal synthesis, expressive synthesis, microtonal tunings, real-time control interfaces, algorithmic improvisation, music notation, music theory, ethnomusicology.

1. INTRODUCTION

Voice has always been the privileged mode of Automatophones and appears like a mechanism of a machine automation. Marin Mersenne in 1678 in his Harmonia Universalis dreams of an organ which could reproduce vowels such as a, e, o, and u[1]. Von Kempelen achieved reproduction of these vowels in 1791, using a control mechanism for motorized bellows. Research on the synthesis of the singing voice has a history of over 35 years [Rodet, 2002] while that on the real-time control of synthetic voice is a relatively new field [Cook, 2000, 2005], [Yonezawa and al, 2005].

[1] *" Je m'occupe maintenant à trouver la manière de faire prononcer les syllabes aux tuyaux d'orgue. J'ai desja rencontré les voyelles a, e, o et u mais i me fait bien de la peine, et puis j'ay treuvé la syllabe vê et fê. Je ne sçay si je paourray prendre le losir de trouver les autres consonnes, à raison des différentes experiences qu'il faut faire sur ce sujet , lesquels estant de coust, je laisseray le reste à ceux qui voudront passer outre." (Marin Mersenne*)[1]

The overall objective of the phonodeon project is to create a novel hybrid musical instrument by combining the control capabilities of the accordeon with programmable real-time algorithmic generation and sound synthesis techniques. At this initial phase, we are developing a "grammar" for using the phonodeon as controller of synthesized singing by any accordeonist equipped with an accordeon-to-MIDI interface. We also want to enable composers of contemporary electroacoustic music to write for this kind of controller in order to explore and extend the capabilities of synthetic voice models and beyond and to experiment with the extrapolation and interpolation between different singing techniques of vocal styles from various parts of the world.

In previous stages of our research, we employed the Phonodeon for the control of voice-like timbres, mapping the bellows to expressive characteristics such as amplitude, virbarto and others and the left -hand buttons for changing vowel qualities [Georgaki et al. 2005].

The work described in the present paper aims to expand the musical capabilities of the phonodeon, enabling it to "sing" in the style of Greek chanting. Greek chanting (often called "Byzantine Music") refers to the liturgical music and recitation tradition that is still practiced in Greek Orthodox churches. It is a strictly vocal tradition with more than 1000 years of history. As such, it has developed its own modal system, which is related to other contemporary and older systems of the eastern Mediterannean and Near East. The theory of the system is unique because it relies heavily on *relative pitch structures*, that is, the transposition of intervallic structures to different regions relative to a transposeable reference pitch, rather than on a fixed repertoire of absolute pitches or pitch classes. Implementing this principle on a freely programmable synthesis system with hardware controller is one of our research tasks. The solutions developed may be useful in exploring microtonal tuning techniques in other genres, including experimental computer music in general.

2. A VIRTUAL GREEK CHANTER

Based on the conception of text-to speech vocal synthesizer developed in the Department of Computer science, University of Athens [Georgaki, 2004] [18] a research program named AOIDOS[2] is on evolution on the analysis and synthesis of Greek singing, aiming on, the synthesis by concatenation of Greek singing diphones and the study of performance rules in different vocal techniques of Greek music. The long-term goal is to develop a 'cognitive chanter model' with elements concerning not only timbre, but all levels of chant structure. Layers of elements beyond timbre and phonemes are: Interval structure of the modes, generative phrase and melody structure models of the modes, micro-ornamentation (ornaments added by voice fluctuation in the range of 1-3 notes, lasting about 1 second or shorter), macro-ornamentation (ornamentation and expansion of melodic skeletons). In order to achieve this, a model is being developed based on data obtained by analysis of recordings of chanters and singers in a number of related genres (rebetiko, liturgical chant, traditional folk music, etc.). As a first step data has been collected regarding voice timbre, formant structure and inflection or ornamentation details including note transitions and vibrato. An important issue arising in this context is the relationship of the intervals and tunings of the Greek modal system to the vocal timbre of Greek chant. Since Greek chant has existed as a strictly vocal genre for more than 1000 years, it is interesting to examine whether and how the intervals that it employs concur with spectral characteristics of its particular voice technique.

The implementation of a playable system based on recorded data and on the MIDI-accordeon serves as a practical test for the validity of the mentioned cognitive chanter model.

3. SYSTEM AND CONSTRAINTS

Our performance system consists of an accordion coupled to a programmable synthesis environment via MIDI. The right-hand keyboard is used in the usual manner to play pitched notes. The buttons of the left hand serve multiple purposes: Phoneme formant choice, timbral inflection, ornament and expression indications and choice of melodic pitch interval patterns. Additionally, the

4. DATA AND ANALYSIS

Singing data for the analysis were originated from a well-experienced professional Greek chant male performer, 50 years old, who was instructed to chant the five Greek singing vowels, in isolation and in an ascending scale of the eight notes C3 up to C4.

The experimental recordings were performed in an anechoic studio using a Rode K2 microphone and digitizing at 44.100 Hz with a resolution of 16 bits using model Yamaha AW16G professional audio workstation.

Formant frequencies F_i (i=1,…,5) (Hz), formant bandwidths B_i (Hz) and formant relevant to the signal gain (*third column*) amplitude A_i (dB) were derived from the above data using Linear Prediction Coding (LPC) digital signal analysis my means of PRAAT software environment [19]. Table 1 presents the results of the formant analysis at the steady-state regions of each vowel.

The same performer was asked to chant a non sense music phrase, consisting of a mix of the five Greek vowels. Fig. 1 describes this music phrase in Byzantine notation. in The conditions and the apparatus for the recording of this phrase was identical to the one described above for the case of the isolated vowels. Then, using the same methodology for the analysis of the data, formant trajectories (F_i, B_i, A_i, i=1, ..,5) were extracted for the whole duration (36,4 sec) of the musical phrase and were stored in a file for the purpose of this paper.

---

2 AOIDOS refers to the singer who composed in performance (the *aoidos*) in Homer's *Iliade.* Furthermore, it it is part of a triad of concept relating to the study of the arts in Greek antiquity, reflected in three primordial forms of the Muses represent three central aspects of the human artistic-intellectual pursuit: *Melete:* Research, Study, *Mneme :* Pattern storage and recognition mechanisms, *Aoidos*: Expression of emotional states.

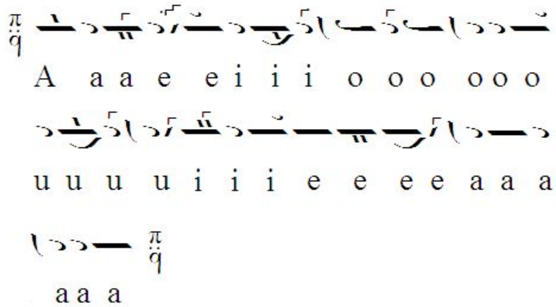

Fig. 1 Byzantine music phrase used for formant trajectories extraction.

5. SYNTHESIS MODEL

Gia ton Gianni

6. TIMBRE AND EXPRESSION (gia ton Gianni)

6.1 Synthesis Model

6.2 Phonemes

6.3 Dynamics

6.4 Vibrato

6.5 Transitions

6.6 Micro-inflections

7. MODAL SYSTEM

ASDF

7.1 Modal Structures

ADSF

7.2 Interval Modules

7.3 Module Combinations

7.4 Momentary Inflections

7.5 Melodic Patterns

8. EXPERIMENTS and SYSTEM EVALUATION

The experiments done on the control of the virtual Greek Chanter through *Phonodeon* is based upon the following :

1. Interpreting the vowel /a/ of a masculine voice in the second mode of Byzantine scale.
    a. Adding expression trough the movement of the bellows
    b. Interpolating between different types of Greek chant vowels
    c. Changing the timber of the Chant vowels to bel-canto vowels through the control of the left buttons and the special mapping of the formants
    d. Adding vibrato through the sensitive touch mechanism of the right hand.
2. Interpreting a phrase of vocalises in Byzantine singing
    a. Make gradually the interpolation form the *bel-canto singing vowels* to the *Greek Byzantine singing vowels*
    b. Ascend an occidental scale which gradually is being transformed to a Byzantine mode. Timbers the same.

3. Work on the changement of the timber quality of vowels form *bright to dark*, whisper and *wet. Explore the transition form one Byzantine vowel to other.*

The evaluation of the system is based upon the perception of our team. We van say that the controlling of Greek Chant vowels is on the point and the resulting sounds are satisfactory.

Improvements on the controlling system and on the responsitivity of the MIDI controller can give better results.

a. The expressivity of our system, trough bellows and touch sensitive control by adding more sensors in the bellows and by establishing the grammars of performance of a vowel and its transitions

b. Accessibility of Phonodeon controller. We found that this kind of voice can be played by an expert accordion player who is not specialized in the electronic world, indicating hin the way to interpret the Greek Chant vowels (by the left hand and the bellows)

c. Mapping of parameters reflect the user's experience . Fixed and non-fixed mappings will establish a new grammar for interpreting vowels, and diphones through Phonodeon.

9. DISCUSSION-FURTHER RESEARCH

Along with the opportunities will come many challenges. In order to fully exploit the musical potential of developments in technology, musicians will need to be receptive to learning about the possibilities afforded by the new technologies while, at the same time, remembering that these are just tools, and no substitute for creativity. Technological developments facilitating the transition from non-real-time to real-time [16] modes of operation in sound synthesis – and thus live synthesis for example – have also helped to broaden the range of performance possibilities. Music research has advanced in a cyber-world where the continuity between traditional modes of interpreting music and musical styles and new technological tools is note evident. The musicians need to adapt their instruments and their performance and interpretations styles in new modes which give them a progressive capability controlling new synthetic timbers in a grammar which is not far away form their skills..

Creating the conditions of a *Phonodeon* mapping parameters grammar in order to control Byzantine singing vowels enhances this new instrument to be more accessible to musicians who are not

In order to complete our new instrument Phonodeon with different grammar techniques we would like to establish the imperatives that put limits in the infinity of possibilities of use of a virtual instrument.

In order to justify our ideas, about the eventual conception of this vocal controller [Georgaki, 1999] we are going to discuss the construction this instrument according to the classical instrumental model of the acoustician E. Leipp [Leipp, 1989] which depends on the:

a) perceptive imperatives,

b) fabrication imperatives,

c) anatomo-physiological imperatives,

d) commercial imperatives and

e) the liberty domains concerning pitch, intensity, timbre, form

f) the possibility to implement new vocal techniques in a cognitive form.

According to the liberty domain concerning the *intensity and the dynamics*, we are expecting a better control than the ordinary synthesizer in order to enhance to expressive quality of singing. We have been confronted many times till now with disadvantages of expression in the most electronic instruments of the commerce. In addition, the control by sensitive touch or pedals is insufficient in the case of voice as we have already described above.

According to the *liberty of choosing the timbres* by pertinent controllers or by selecting a pre-constructed timbre-diphone from the database, we can create instrumental sounds which approach the vocal sound by extrapolation, hybridation or interpolation.

10. CONCLUSION

In this paper we have presented a new approach of connecting Midi-controllers to music Informatics and ethnomusicological research. We have presented ways of controlling Byzantine chant vowels by Phonodeon, a MIDI accordion type controller which enhances the transition between Byzantine Chant singing and Bel-canto vowels. The grammars established on fixed mappings of parameters enables a non–expert user to manipulate vocal timbers.

The results of controlling the synthesis engine by data based on Byzantine singing analysis are satisfactory.

11. ACKNOWLEDGEMENTS

The work described in this paper has been partially supported by the PYTHAGORAS project of the Operational Programme for Education and Initial Vocational Training (EPEAEK) of the Greek Ministry of Education under the 3rd European Community Support Framework for Greece.

12. REFERENCES

[1] Bowman, B., Debray, S. K., and Peterson, L. L. Reasoning about naming systems. *ACM Trans. Program. Lang. Syst., 15,* 5 (Nov. 1993), 795-825.

[2] Ding, W., and Marchionini, G. *A Study on Video Browsing Strategies.* Technical Report UMIACS-TR-97-40, University of Maryland, College Park, MD, 1997.

[3] Fröhlich, B. and Plate, J. The cubic mouse: a new device for three-dimensional iput. In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '00)* (The Hague, The Netherlands, April 1-6, 2000). ACM Press, New York, NY, 2000, 526-531.

[4] Lamport, L. LaTeX User's Guide and Document Reference Manual. Addison-Wesley, Reading, MA, 1986.

[5] Sannella, M. J. Constraint Satisfaction and Debugging for Interactive User Interfaces. Ph.D. Thesis, University of Washington, Seattle, WA, 1994.

[6] Cook, P. R., and Leider, C. Squeeze Vox: A New Controller for Vocal Synthesis Models. In Proceedings of the International Computer Music Conference, Berlin, 2000.

[7] Georgaki Anastasia (1998a) Problémes techniques et enjeux esthétiques de la voix de synthèse dans la recherche et création musicales. Thèse de doctorat, EHESS/IRCAM, Paris.

[8] Georgaki A (1999). "Towards the conception of a vocal synthesizer", Proceedings of the VIth symposium of Brazilian Computer Music, Rio de Janeiro, Brazil

[9] Georgaki A.(2004a) "Virtual voices on hands". Prominent applications on the synthesis of the singing voice, Sound and music computing Proceedings of the SMC05, IRCAM, Paris 2005.

[10] Georgaki Anastasia, Zannos Ioannis, Valsamakis Nicolas: "Phonodeon: controlling synthetic voices though MIDI-accordion" SMC 2005

[11] Leipp Emile(1989)Acoustique et Musique. Masson , Paris

[12] Sundberg J. (2001) "Sounds of singing. A matter of mode, style, accuracy, and beauty", invited paper, 4th Pan European Voice Conference,Stockholm

[13] Rodet X (2002): Synthesis and processing of the singing voice, Proc.1st IEEE Benelux Workshop on Model based Processing and Coding of Audio (MPCA-2002), Leuven, Belgium

[14] Yonezawa T., Suzuki N., Mase K., Mogure K. : ''Handysinger: Expressive Singing Voice Morphing using Personified Hand-Puppet Interface", In NIME 2005 Proceedings, Vancuver BC, Canada, 2006.

[15] Zannos, I. Intonation in Theory and Practice of Greek and Turkish Music. In: Dieter Christensen (ed.) Yearbook for Traditional Music, Vol. 22, pp. 42-59. New York: Columbia University.1990

[16] Zannos, I. Echos und Makam: Ichos und Makam. Vergleichende Untersuchungen zum Tonsystem der griechisch-orthodoxen Kirchenmusik und der türkischen Kunstmusik. Bonn: Orpheus Verlag GmbH. 1994.

[17] Zannos I.: Der Rechner in der Notationsforschung und musikalischen Analyse - Ein Beispiel. In: Helmut Schaffrath (ed.) Computer in der Musik: Über den Einsatz in Wissenschaft, Komposition und Pädagogik, pp. 45-50. Stuttgart: Metzler1991

[18] G.Xydas and G.Kouroupetroglou: "The DEMOSTHeNES Speech Composer" . In Proceedings of the 4th ISCA Tutorial and Research Workshop on Speech Synthesis, Perthshire, Scotland, August 29th - September 1st, 2001 , pp 167-172.

[19] P. Boersma - D. Weenink, Institute of Phonetic Sciences, University of Amsterdam, PRAAT, www.praat.org

[20] D. Delviniotis, (1998) "A classification of Byzantine singing voices", Proceedings of EUSIPCO-98, Rhodes 1998

		gain	F1	B1	A1	F2	B2	A2	F3	B3	A3	F4	B4	A4	F5	B5	A5
C3	/a/	59.2	494	47	59	922	109	46	2124	78	38	2700	282	28	3129	31	40
	/e/	64.8	468	44	63	1371	81	51	2169	22	63	2796	155	52	2977	43	57
	/i/	65.9	333	62	59	1577	24	58	2289	26	62	2876	66	59	3075	55	58
	/o/	65.4	552	46	66	663	277	56	2489	135	40	3077	139	57	3105	53	58
	/u/	60.9	410	59	60	2207	51	37	3009	51	47	3158	36	49	4180	74	20
D3	/a/	66.7	475	46	66	971	115	51	2387	28	60	2803	102	54	3063	35	58
	/e/	67.2	461	49	62	1363	44	56	2186	31	59	2850	73	60	2989	33	64
	/i/	66.9	333	70	59	1572	23	59	2245	81	53	2875	76	61	3019	48	62
	/o/	66.7	468	34	69	772	79	55	2397	89	41	2948	54	54	3135	35	56
	/u/	66.8	444	36	67	873	77	53	2322	92	41	3019	130	60	3044	26	63
E3	/a/	68.4	489	34	68	1066	146	48	2431	44	57	2849	64	59	3107	22	64
	/e/	66.2	484	61	60	1356	54	55	2171	42	57	2854	41	62	3049	36	60
	/i/	65.5	342	60	60	1542	33	56	2215	45	56	2894	108	56	3031	35	62
	/o/	67.4	527	50	66	2319	107	44	2926	163	52	3048	13	68	3995	164	23
	/u/	65.2	697	137	58	2210	269	42	2700	332	43	3068	6	70	3990	330	20
F3	/a/	68.9	514	45	67	1036	147	50	2423	59	56	2882	84	60	3057	38	63
	/e/	64.9	485	97	55	1496	203	44	2155	36	60	2877	88	57	3046	34	62
	/i/	67.6	360	77	59	1536	12	66	2232	48	57	2882	49	62	3082	24	65
	/o/	67.0	492	53	66	832	136	52	2348	105	43	3068	15	65	4026	172	22
	/u/	64.7	495	118	58	861	96	54	2270	74	44	3017	49	62	3055	87	60
G3	/a/	67.9	502	144	58	1020	245	47	2383	45	58	2955	84	63	3054	69	63
	/e/	67.2	505	187	53	1380	75	55	2158	28	63	2882	40	63	3108	34	62
	/i/	69.8	395	73	62	1566	21	66	2197	62	59	2871	71	62	3100	28	67
	/o/	69.4	554	220	60	823	154	57	2360	125	47	3030	201	66	3057	13	72
	/u/	65.8	455	98	61	933	128	51	2291	62	49	2933	96	53	3146	22	63
A3	/a/	66.0	506	70	63	975	115	52	2439	50	56	2750	239	48	3126	13	64
	/e/	68.6	482	79	62	1399	41	61	2203	51	59	2888	67	60	3144	30	64
	/i/	69.5	439	20	69	1557	13	68	2196	33	63	2863	77	59	3122	30	64
	/o/	70.2	495	99	65	901	61	62	2322	77	49	2995	18	69	3174	88	55
	/u/	70.0	481	58	68	973	56	61	2255	93	48	2967	31	64	3217	55	57
B3	/a/	71.1	546	116	64	1030	152	55	2341	51	60	2874	92	60	3143	14	71
	/e/	73.9	520	51	69	1597	254	53	2198	71	65	2870	44	70	3153	52	64
	/i/	71.9	442	56	64	1657	43	63	2203	70	62	2883	95	63	3131	32	69
	/o/	74.8	520	71	71	1005	49	67	2382	110	55	2955	110	62	3174	21	72
	/u/	72.9	505	87	66	1009	67	62	2260	82	53	3034	167	67	3077	21	73
C4	/a/	75.0	535	66	69	1073	67	62	2448	157	53	3021	213	72	3050	17	77
	/e/	74.8	534	30	73	1527	150	55	2274	61	66	2882	58	69	3177	49	66
	/i/	73.3	438	72	64	1610	26	68	2240	62	64	2894	56	68	3189	75	63
	/o/	75.0	538	67	70	1043	37	69	2123	132	50	2993	21	74	3173	63	64
	/u/	70.8	532	98	64	1078	126	56	2313	72	55	3000	104	60	3246	30	67

Table 1. Formant frequencies F_i (i=1,…,5) (Hz), formant bandwidths B_i (Hz) and formant relevant to the signal gain (*third column*) amplitude A_i (dB) for the signing of the five Greek vowels in eight different notes.