IDENTIFYING TRANSIENT PATTERNS IN IDIOMATIC GREEK AND SLAVONIC MUSIC

*Dionysios Politis** *Alexander Dolia*[†] *Panagiotis Linardis***

{*Multimedia Lab, **VLSI Lab},
Computer Science Dept., Aristotle University of Thessaloniki, GR-540 06, GREECE.
e-mail: {dpolitis, linardis}@csd.auth.gr
[†] Dept.504, State Aerospace University,
17 Chkalova St., Kharkov, 61070, Ukraine.
e-mail: lukin@mmds.kharkov.ua

Abstract—**This paper describes a method for determining the content affinity of musical patterns in vocal reproduction in Greek and Slavonic by using wavelet decomposition techniques. The goal of the analysis is to determine the exact musical description of the patterns in order to reproduce them by musical synthesizers acting as synthetic singers. The exemplar pattern of this paper is that of "petasti", which was selected because it is a transient phenomenon particular in traditional Greek songs and chants, in both instrumental and vocal performances. A comparison of the pattern is attempted with Slavonic chants by determining time and frequency localizations of the pattern.**

Index Terms—**Wavelet decomposition, musical patterns, "petasti", voice synthesis.**

I. INTRODUCTION

This paper copes with the identification of musical patterns that literally exist in Greek and Slavonic musical traditions. The origin of these patterns can be found in manuscripts that are at least 10 centuries old. In Fig. 1.(a) a fragment of a manuscript is shown which implicitly denotes these patterns [1]. In Fig. 1.(b) explicit marks of "petasti" are recorded in Byzantine music notation [2].

The creation of a musical notation system proved to be a painful and long procedure in the West and especially in the East, where as the manuscripts that have survived indicate (Fig. 1), we have had different notations evolving one from the other in a shorthand like form. This musical system was not confined only to ecclesiastical music; it was a generalized musical system originating directly from ancient Greek Music and was used as the usual music surface by all the people living in the vast areas of the Byzantine empire, from Southern Italy and the Balkans up to Ukraine and Russia and down to Middle East and Egypt.

Finally the Byzantine notation prevailed in the East as the dominant musical notation and was reformed in 1814 and 1881 from committees of the Ecumenical Patriarchate of Constaninople to an analytical notation system whose symbols came out of the numerous symbols of the earlier shorthand-like notations [3].

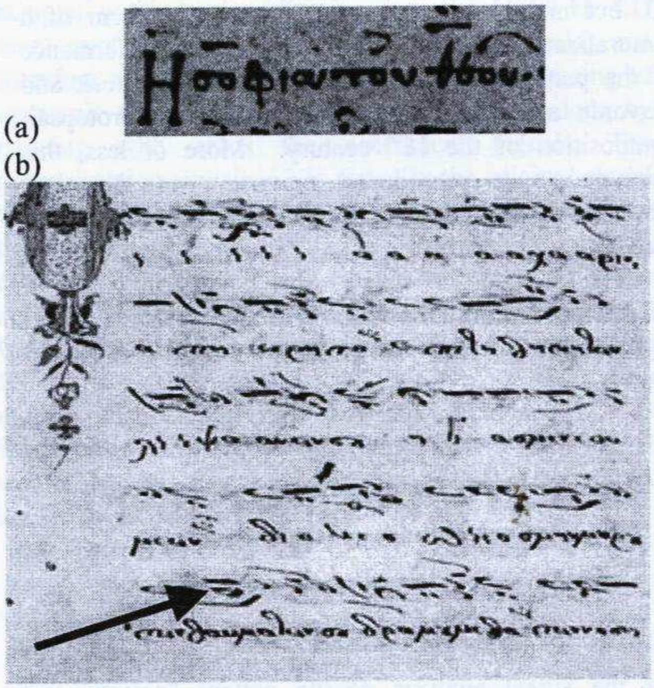

Fig. 1. Byzantine Music Manuscripts: (a) Detail from an early manuscript, the so-called 'Chartres fragment', with musical notation, beginning of the sticheron "Η σοφια του Θεου, Mode plagal D", Copenhagen. (b) "Petasti" pointed out in a Doxastarion composed by Jacov Protopsalt, Mount Athos, 1805 AD.

Major differences between the polyphonic Western music tradition and the Byzantine system are:

- the intervalistic system (system of musical scales and their microtonal distribution) that contains predominantly musical intervals that are smaller than the well tempered ones. Successors of the ancient Greek Dorian, Lydian, Phrygian, Mixolydian Modes and their plagal ones are encountered [3][4]. This characteristic introduces a subjective criterion of orientalization in the psycho-acoustics of the hearings, and
- the use of larger, formalized transitory patterns as main elements of the musical structure that are

inherent in vocal performance and render a qualitative nuance.

A musical pattern that has a quantitative and qualitative as well character is that of "petasti". This pattern has been detected in instrumental performance of traditional Greek songs [5] and in Byzantine music vocal performance [6]. According to Byzantine Music theory [2], "petasti" is a transient phenomenon that assigns a phonation quality to the raising or lowering of pitch. In this paper, analysis of this pattern is performed in order to decipher its qualitative nature. By the term qualitative we mean the characteristics that enable the specialized listener to recognize the "petasti" pattern not simply as a pitch fluctuation, a fluctuation of fundamental frequency F0, but mainly as a parametric transitory form of a gutturalization. Analysis is focused on the performance of the pattern in Mode A (Dorian) [4] in Greek and Slavonic according to the classical Jacov Protopsalt composition of the 18th century. More or less, this performance is reproduced nowadays in the slow rendition of the vesper hymn "Lord, I have cried unto thee".

The exact description of this pattern in terms of microtonal distributions will enable the correct musical reproduction by computer programs acting as synthetic singers.

II. EXTRACTION OF AUDIO FEATURES USING THE DISCRETE TIME WAVELET TRANSFORM

For the signal analysis of the pattern "petasti" the Discrete Time Wavelet transform (DTW) will be used in conjunction with the classical Short Time Fourier Transform (STFT). One major advantage afforded by wavelets is the ability to perform local analysis. This is useful near the discontinuity areas of the signal, namely the consonants for vocal signals. A second advantage is that with each stage DWT the signal is seperated into *approximations* and *details* while downsampling is performed. The DWT can be seen as an equivalent of a tree-structured multirate filter bank (Fig. 2) [7].

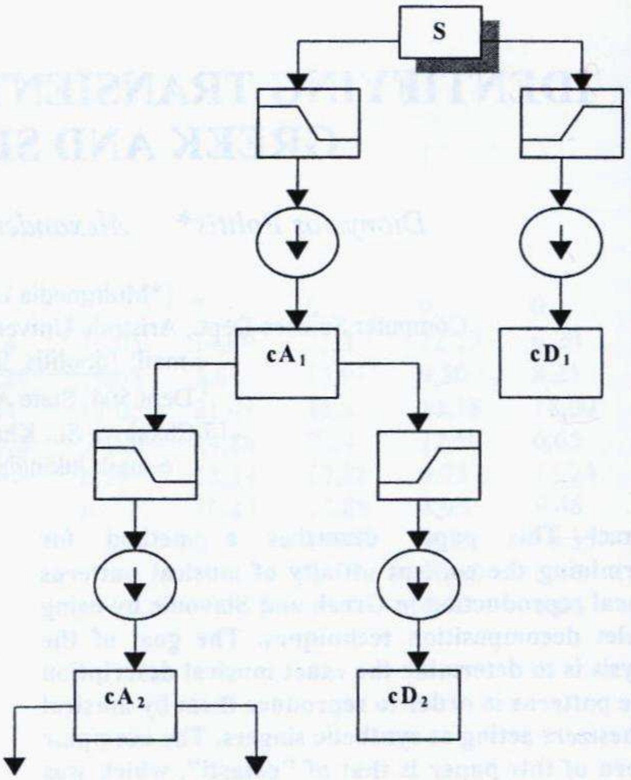

Fig. 2. The 2-level binary tree structured filter bank for signal S.

The mother wavelet used for decomposition was db4, the fourth of the Daubechies family wavelets, and the wavelet decomposition tree is presented in Fig. 2, albeit for convenience only two levels are drawn. D(z) is the outcome of a high pass filter and downsampling, thus a detail component, and A(z) is the outcome of a low pass filter and downsampling, thus an approximation component. This procedure produces at each stage DWT coefficients which are denoted in Fig. 2 as **cA**$_i$ for the approximation components and as **cD**$_i$ for the detail components for the **i**-th level analysis of signal S.

In order to analyze the transient nature of "petasti", in both Greek and Slavonic reproductions of classical musical phrases, prototypal performances of these melodies have been recorded.

The waveform of the prototypal performance in Greek is shown in Fig. 3(a) and its spectrogram is shown in Fig. 3(b). In Fig. 3(b) the segment around the first phoneme /i/, the one that is performed with "petasti", fluctuates around note E3 in a manner close to wavelet db4 (Fig. 4). This is the reason for choosing the Daubechies family of wavelets.

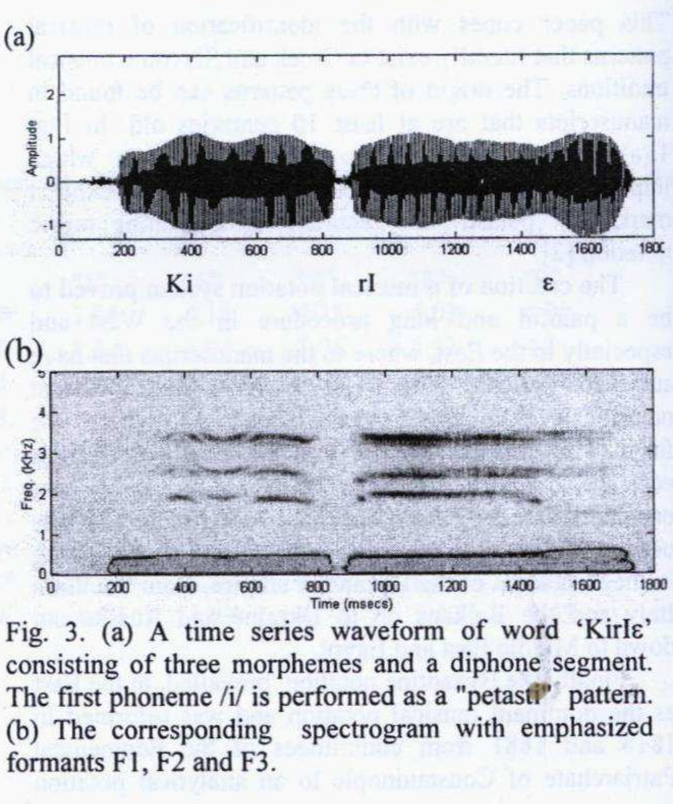

Fig. 3. (a) A time series waveform of word 'KirIε' consisting of three morphemes and a diphone segment. The first phoneme /i/ is performed as a "petasti" pattern. (b) The corresponding spectrogram with emphasized formants F1, F2 and F3.

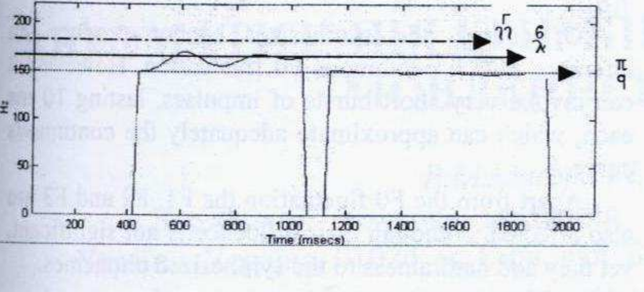

Fig. 4. F0 fluctuations around note Boυ, the equivalent of note E3. Note Boυ corresponds to 158 Hz, while the pitch levels of notes Πα and Γα, shown with dotted lines, correspond to 144 Hz and 170 Hz respectively. The interval Boυ-Γα is slightly bigger than a semitone.

The pattern we examine consorts the first morpheme of the utterance in Fig. 3(a). By using a cepstrum based approach, fundamental frequency F0 is estimated, and the "petasti" fluctuation around note Boυ is estimated. Boυ is the equivalent of note E3 of the well tempered scale and it is assigned the frequency of 158 Hz [2]. The subject N. Kougias who has performed the musical pattern has achieved a remarkable tuning with the resonant frequencies.

As seen from the formant curves of Fig. 3(b), it is clear that the musical pattern affects not only F0 but also the first three formants, F1, F2 and F3.

The waveform of an equivalent prototypal performance in Slavonic is presented in Fig. 5(a). The word uttered is 'Gospodi' with the first /o/ performed with "petasti". For this utterance, the cepstrum based technique for the determination of F0 yields poor results.

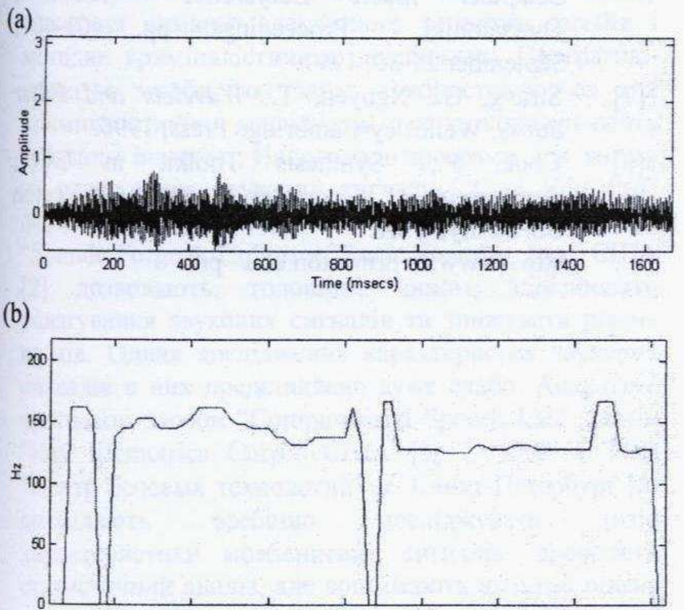

Fig. 5. (a) Time series of morpheme /Go/ from the slow rendition of the word 'Gospodi' having the first /o/ performed with "petasti". (b) F0 cepstrum based estimation of approximation cA3 of the original signal.

In this case, the signal was decomposed by using the fourth wavelet of the Daubechies family and was analyzed using the third-level approximation of it. Prior to estimating F0, upsampling takes place. Then F0 is readily estimated. The final F0 contour is drawn in Fig. 5(b).

Apart from F0 fluctuations, the first three formants of the utterance also have important information. Following the same decomposition scheme, we obtain the detail coefficients of the analyzed signal in order to estimate the first three formants of the utterance.

We focus our analysis on signals cA3 and cD3. cA3 consists of the approximation coefficients downsampled and decomposed from the original signal by the Discrete Wavelet Transform three times. Both signals are thus sampled at 5512.5 Hz (deriving from the 44100 Hz CD level quality recordings of the original signals) and therefore a resampling procedure takes place aiming to upsample them to at least 11025 Hz. This is done by padding with zeros the coefficient signals. The approximation signal cA3 is virtually the original signal that has become less noisy after the successive decompositions, and the high frequency information concerning formants F1, F2 and F3 is filtered out of the signal in the form of the detail coefficients signal cD3.
In Fig. 6 we see an estimation of the formant fluctuation for the utterance with "petasti" of /o/ (Fig. 6(a)). It is evident that in this case the performed analysis is inadequate. Consequently, we dilate the signal estimating its approximation and detail coefficients according to the binary-tree structure of Fig. 2.

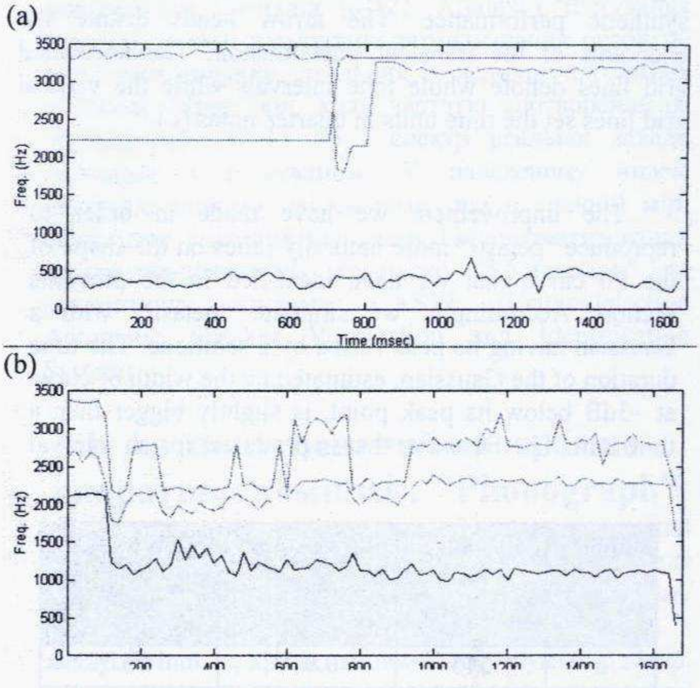

Fig. 6. (a) Formants F1, F2 and F3 estimated from the prototypal signal in Slavonic (b) Formants F1, F2 and F3 estimated from cD3, the detailed third level high frequency component of the signal.

III. CONCLUSIONS: ADVANCES IN SYNTHETIC VOICE REPRODUCTION

The synthetic musical performer we have used to reproduce melodic themes with “petasti” was a formant synthesizer that had physically modeled the glottis and shape files of a Greek singer. This program was compiled using P. Cook’s Synthesis Toolkit in C++ [8].

This voice synthesizer, in order to produce a “petasti”-like phonation quality, produced discrete sounds like the ones shown in Fig. 7(a). If we consider that phoneme /o/ has the typical formant distribution described in Table I, then the conventional synthetic reproduction of /o/ with “petasti” would merely yield three intermediate notes, with the second note raised by a semitone. This means that a quarter-note influenced by “petasti” is analyzed as the sequence [sixteenth-note, eight-note sharp, sixteenth-note.]

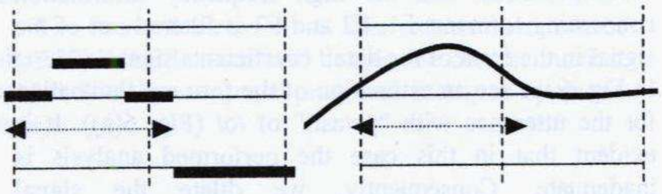

Fig. 7. (a) Quantum leap versus (b) continuous F0 synthetic performance. The arrow heads denote the endpoints of the transient phenomenon. The horizontal grid lines denote whole tone intervals while the vertical grid lines set the time units in quarter-notes (♪).

The improvement we have made in order to reproduce “petasti” more naturally relies on the shape of the F0 curve that we have identified in the previous section. Accordingly, we simulate “petasti” with a Gaussian having its peak raised by a semitone. The time duration of the Gaussian, estimated by the width of curve at –3dB below its peak point, is slightly bigger than a time unit. The F0 curve that is produced for an interval of a major second (i.e. a whole tone) is shown in Fig. 7(b).

Phoneme	formants (Hz)	Freq. Width	Relative amplitudes (dB)
/o/	515	0.977	0
	1805	0.810	-10
	2526	0.875	-10

Table I. Estimated mean values for phoneme /o/. The frequency width of the formants is not denoted in Hz but as it is inserted into the synthesizer, i.e. exp(2W/Fp).

Of course, the synthesizer cannot produce an utterance with a continuous F0 fluctuation. However, it can invoke very short bursts of impulses, lasting 10 ms each, which can approximate adequately the continuous curve.

Apart from the F0 fluctuation the F1, F2 and F3 are also affected. Although their influence is not significant, yet they add naturalness to the synthesized phonemes.

REFERENCES

[1] “Monumenta Musicae Byzantinae”, Institute for Greek and Latin, University of Copenhagen. http://www.igl.ku.dk/MMB.

[2] Manuscript from Iveron Monastery of Mount Athos, cod. 437, 1805 AD, reproducing the Doxastarion of Jacov protopsalt (†1800).

[3] Panagiotopoulos, D., *Theory and Praxis of Byzantine Ecclesiastical Music*, 3rd ed, Athens, 1982 (In Greek).

[4] Jacobs, A., *The Penguin Dictionary of Music*, Penguin, USA, 1991.

[5] Pikrakis, A., Theodoridis, S., Kamarotos, D., “Recognition of Isolated Musical Patterns in the context of Greek Traditional Music”, *Third IEEE International Conference on Electronics, Circuits and Systems ICECS ’96*, Rhodes, pp. 1223-1226, October 13-16, 1996.

[6] Politis, D., Tsoukalas, A., Linardis, P., Bakalakos, A., “VIDI-A Voice Instrument Digital Interface for Byzantine Music”, *International Computer Music Conference ICMC97*, Thessaloniki, Proceedings, pp. 403-407, September 25-30, 1997.

[7] Strang, G., Nguyen, T., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[8] Cook, P., “Synthesis Toolkit in C++”, Department of Computer Science, Princeton University, USA. http://www.cs.princeton.edu/~prc