



HAL
open science

A Virtual Castrato (!?)

Philippe Depalle, Guillermo Garcia, Xavier Rodet

► **To cite this version:**

Philippe Depalle, Guillermo Garcia, Xavier Rodet. A Virtual Castrato (!?). ICMC: International Computer Music Conference, 1994, Aarhus, Denmark. pp.357-360. hal-01105432

HAL Id: hal-01105432

<https://hal.science/hal-01105432>

Submitted on 20 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Virtual Castrato (!?)

Ph. Depalle, G. García, X. Rodet

ICMC 94, Aarhus (Denemark)

Copyright © Ircam - Centre Georges-Pompidou 1994

Abstract

In this paper we present an original research work culminating in the high quality production of thirty-nine minutes of castrato voice, by the means of sound analysis, processing and synthesis. The goal of the project is the creation of a soundtrack for a film about a famous eighteenth century castrato. Two complementary voices - a coloratura-soprano and a counter-tenor - are used to cover the entire range and compensate for technical difficulties. Then their timbres are homogenised by using combined techniques such as phase-vocoding, additive analysis/synthesis and spectral envelope and pitch estimation.

1. Introduction

In this paper, we present an original work on the recreation of a castrato voice through sound analysis, processing and synthesis. It has been done to produce the soundtrack for a musical film [[Corbiau, 94](#)] about Farinelli, the famous castrato of the eighteenth century. Farinelli, issued from the castrato's music schools of Porpora and Bernachi, and as a contemporary of Haëndel, was a celebrity of the Baroque era like "pop stars" are in the present [[Dico, 91](#)]. The film, realised by Gerard Corbiau, brings back to life a repertoire which could not be sung anymore. The musical consultant of the film Marc David [[David, 94](#)] has "recovered" unedited scores from the French National Library.

Thirty-nine minutes of high quality processed singing have been created for a CD of the soundtrack, produced by Audiovis [[Audiovis, 94](#)].

In the following sections we explore some of the issues by describing the characteristics of castrati voices and the practical constraints of the film project. Then, we detail the chosen solution and the specific processing used to make such a voice.

2. Problematic

Castration has been forbidden in western countries since the last century, and the last occidental castrato died in 1922.

Castrati were generally well known for the special *timbre* of their voices : due to the surgical intervention they had undergone, their voices had not changed with puberty. Furthermore, with maturity, castrati lung capacity, chest's size, physical endurance and strength were generally greater than those of normal male [[Sauvage et al., 84](#)]. Consequently, they were able to sing very powerfully. Farinelli could sustain a note longer than one minute and he could sing long phrases of more than two hundred notes without seeming to take a breath. Their small and supple larynx along with their short vocal cords, allowed them to vocalise in a large range (up to three octaves and a half) and to sing with a great vocal flexibility (they could sing large intervals rapidly, cascading scales and trills). Furthermore, castrati were selected among the best child singers and trained very intensively.

As castrati's specific repertoire takes their expert singing technique into account, this repertoire is

extremely difficult to sing. Some pieces of music are interpreted today but they are chosen for their simplicity or are played at a lower tempo.

Another difficulty in recreating a castrato voice is the lack of recorded references. The last occidental castrato has recorded less than one hour of singing on wax cylinders between 1902 and 1904 [[Moreschi, 88](#)]. This historical recording has very little technical utility due to its extremely poor quality.

Nevertheless, we can take into account the physical characteristics of the whole vocal production system of the castrati, the global aesthetic of the historical recording and descriptions found in the literature. On the other hand, the design of the final processed voice is conditioned by the wishes of the film and music producers. The definitive choice concerning the timbre of the voice is a compromise between the two preceding constraints.

Scores are difficult to interpret and parts of them can not be sung by contemporary singers anymore, due to the large range of the castrati. Thus it is necessary to use two complementary voices to cover the entire range and compensate for technical difficulties.

3. Recording and Editing of the Voice

Several constraints had to be respected to choose the two voices. Basically, the desired voice had to sound like an ambiguous juvenile male voice. A large casting was made by the film and music production to find a counter-tenor and a coloratura-soprano with similar and good baroque singing techniques (especially in terms of vibrato and articulation). Chosen singers are Derek Lee Ragin and Eva Godlevska.

The recording was made in the concert hall "L'Arsenal" in Metz, France with the orchestra "Les Talens Lyriques", conducted by Christophe Rousset. Due to artistic constraints, sound engineer Jean Claude Gaberel recorded voice and orchestra simultaneously, despite the evident interest of a multitrack recording. One consequence is the presence of orchestra components at 20 to 30 dB under the mean average level of the singing voices. This drastic constraint imposed a certain robustness to our processing method. Technically, the recording was made on a Nagra IV-D machine with a precision of 20 bits.

The editing has been made by Jean Claude Gaberel on a Sonic solution machine. This remarkable work often reached the note-by-note editing level.

To conclude, we should notice several points inherent to the use of two singers: The perceived dynamic is different between them and one can sometimes hear in the middle of a phrase unnatural discontinuities which sound like phrase attacks.

4. Processing

Introduction

The chosen strategy (Cf. figure 1) can be divided in two steps:

First, since one of the artistic specifications was to make the finally processed voice sound close to the counter-tenor one, we modified the coloratura-soprano parts to match the counter-tenor timbre. This procedure, which we call *voice morphing*, constitutes the main and critical step of the scheme.

Secondly, we gave the voice a more juvenile aspect by using global modifications. For instance, we attenuated some high frequency bands to reduce the kind of breathiness found in Derek Lee Ragin's voice. We also made the voice sound brighter by modifying the spectral envelope.

Data Base Constitution

Due to the great predominance of vowels over consonants, we only processed the vowels. As vowel timbre is not only a function of phonemes but also of pitch and intensity, a specific processing has to be applied to each vowel note. Thus, a reference data base composed of all the combinations phoneme-pitch-intensity of the counter-tenor voice had to be set up. Phonemes are represented by spectral parameters which will be detailed further. Since song texts are written in Italian language, we only use the five following phonemes /a/, /ε/, /i/, /o/, /u/. Pitch are chosen chromatically from 185 Hz to 987 Hz and intensity has only three levels: *piano*, *mezzoforte* and *forte*.

For practical reasons, the data base does not cover the entire intensity range. To complete it, we used intensity rules [Rodet et al., 89] to compute the missing fields.

Segmentation

Once the data base of the reference voice had been built, the musical phrases to be processed had to be segmented in order to label elementary portions in terms of singer, phoneme, pitch, power, begin and end time. Precise pitch estimation was made by the new frequential method described in [Doval, 94]. A first segmentation pass was performed automatically on the fundamental frequency evolution [Cerveau, 94]. Then, a second pass was performed by hand on the signal to adjust the begin and end time of the vowels and to give the singer and phoneme labels.

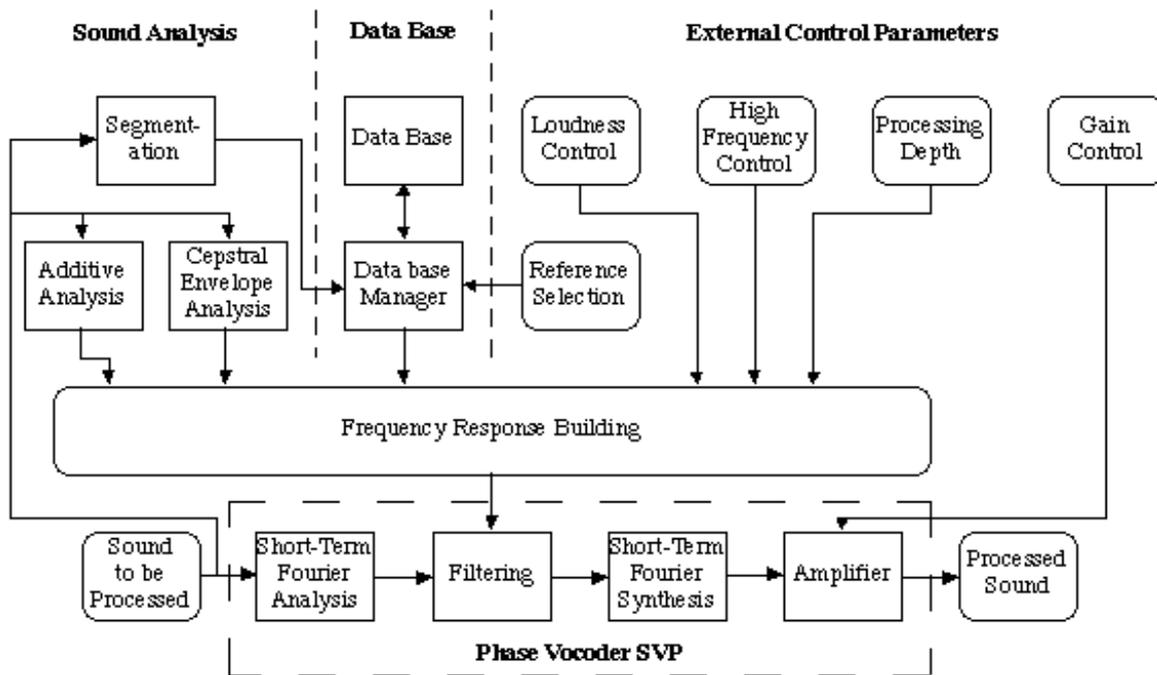


Figure 1: General Synopsis of the voice processing.

Reference Selection

The reference selection step associates each vowel segment found in the voice with a target phoneme in the data base. For example, every phoneme /a/ sung *mezzoforte* by the soprano with a pitch between 415 Hz and 520 Hz can be associated with the data base phoneme /a/ sung *mezzoforte* by the counter-tenor in the same range of pitch. In practice, this correspondence is not hard-coded but can be given in a parameter file. In this way, it makes the transformation process more flexible. For example, multiple versions of the data base elements coexist, corresponding to different recording conditions such as microphone distance, equalisation, etc. So, one of these versions can be selected by including its corresponding reference in the parameter file. Beyond standard voice morphing, this reference selection allows more dramatic transformations such as changing every phoneme /a/ into phoneme /o/.

Spectral Characteristics of the Voices

The basic idea of our voice morphing technique consists in modifying the spectral envelope of the soprano voice to match that of the counter-tenor voice. This is achieved by a frequency domain filtering through a phase-vocoder [Depalle et al., 91] (Cf. figure 1). But we use a more refined technique which is described now.

Since the scores we use are written for castrati, most of the songs are high-pitched, and it's a common fact that in this case the frequency response of the vocal tract is poorly estimated. This is particularly true in the low frequency range (below 2.5 kHz) where partials are widely spaced and formants are very narrow. The first consequence of bad estimation is that voice morphing does not reach the reference timbre; another consequence, which is specific to our context, is that transformation may emphasise some partials of the orchestra.

One possible solution to improve the estimation of the spectral envelope in the low frequency range is to use time evolution of frequency and amplitude of the partials, which scan the spectral envelope as can be seen in figure 2.

Thus, a method which estimates a spectrum envelope model by minimising its distance to the set of frequency-amplitude points, such as the discrete cepstrum [Galas, 90] could be considered.

But in practice the spectral envelope is not always stable during a vowel note. First, the coloratura-soprano often changes continuously the shape of her vocal tract when singing a cascade of notes on the same vowel (Cf. figure 3).

In addition, tremolo correlated to the vibrato effect induces a variation on the amplitude of each partial, which superimposes on the scanning of the spectrum envelope (Cf. Figure 4).

In the middle (2.5 to 5 kHz) and upper frequency range (greater than 5 kHz), the estimation of the spectral envelope remains valid, because of the wider shape of the formants. But if the spectral envelope shape is still constant for a given vowel note, its global amplitude is modulated and fluctuates according to the tremolo. Moreover, this effect is emphasised by the loudness. Finally, in the upper frequency range, the average level is perceptually more important than the precise shape of the spectrum.

A (dB)



Figure 2 : Time-trajectories of the first ten partials in the Frequency-Amplitude plane, for the phoneme /ε/ sung by the coloratura-soprano.

A (dB)

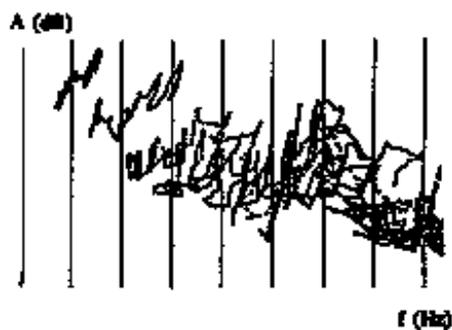


Figure 3 : Time-trajectories of the first ten partials in the Frequency-Amplitude plane, for an eight notes cascade on phoneme /a/ sung by the coloratura-soprano.

A (dB)

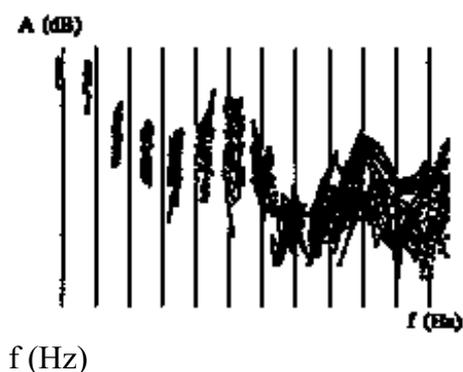


Figure 4 : Time-trajectories of the first fourteen partials in the Frequency-Amplitude plane, for the phoneme /a/ sung by the counter-tenor with a large tremolo effect.

Frequency Response Building

Taking the preceding observations into account, we decided to build a filter frequency response which only acts on the vicinity of each voice partial. In this way, partials of the orchestra, present in between partials of the voice, are not emphasised. So, we designed frequency responses of rectangular form (Cf. figure 5).

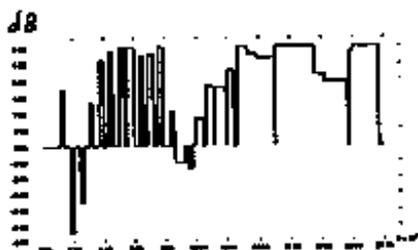


Figure 5 : Frequency response building.

Figure 5 : *Frequency response building.*

The rectangles are designed using additive synthesis parameters [Depalle et al., 93\]](#) and spectral envelopes. For each partial, a rectangle, centred around its frequency is designed. Its level is computed to impose the same relative amplitudes between harmonics on the processed sound as those of the corresponding phoneme stored in the data base. The width of each active band is computed according to the size of the temporal window currently used by the phase vocoder and the frequency deviation due to the vibrato in this window.

In the medium and high frequency range, we evaluate the height of the rectangle by dividing the spectral envelope of the desired phoneme stored in the data base by the spectral envelope of the phoneme to be processed. In addition, high frequency active bands are weighted by a coefficient which controls the breathiness of the result.

Conclusion

Our work constitutes one step towards the very difficult goal which consists in transforming one instrument into another. Furthermore thirty-nine minutes of concert quality singing voice have been produced. This work is all the more interesting because of real world production constraints implying recording conditions which were not optimal for our purposes. For instance, the presence of the orchestra imposed the development of very robust processing methods. Finally one of the most major interesting points of this work is that it now allows us to revive the whole musical repertoire of pieces originally written for castrato. Extracts of the pieces of music will be played during the conference.

Acknowledgement

The authors would like to thank Digital Equipment Corporation France for the DEC alpha 600 and a JV 300 audio board used in this work.

References

[Audiovis, 94]

Orch. "Les Talens Lyriques" conducted by Christophe Rousset. Farinelli. CD to appear in Dec. 1994.

[Cerveau, 94]

Laurent Cerveau. Segmentation des signaux musicaux. *Mémoire de DEA ATIAM*, Université de Paris VI, Jul. 1994.

[Corbiau, 94]

Gérard Corbiau. Farinelli, Primo Uomo. *film*, to appear in Dec. 1994.

[David, 94]

Marc David. Farinelli. To appear in Dec. 1994.

[Depalle et al., 91]

Philippe Depalle and Gilles Poirot. SVP: A Modular System for Analysis, Processing and Synthesis of Sound Signals. *Proc. of ICMC*, Montréal, Canada, Oct. 1991.

[Depalle et al., 93]

Philippe Depalle, Guillermo Garcia and Xavier Rodet. Analysis of Sound for Additive Synthesis: Tracking of Partials Using Hidden Markov Model. *Proc. of ICMC*, Tokyo, Japan, Sept. 1993.

[Dico, 91]

Dictionnaire de la musique. Larousse, 1992.

[Doval, 94]

Boris Doval. Détection de la fréquence fondamentale par une méthode harmonique. *Thèse de Doctorat*. Université de Paris VI, Mar. 1994.

[Galas et al., 91]

Th. Galas and Xavier Rodet. Generalized discrete cepstral analysis for deconvolution of source-

filter system with discrete spectra. *Proc. of ICMC*, San Jose, California, Oct. 1992.

[Rodet et al., 89]

Xavier Rodet and Gerald Benett. Synthesis of the Singing Voice. *Current Directions in Computer Music Research*, M. Mathews and J. Pierce Ed., MIT Press, Boston, 1989.

[Sauvage et al., 84]

Jean-Pierre Sauvage, Philippe Defaye. Les Castrats (Hypothèses phoniatriques). *Les Cahiers d'O.R.L.*, Tome XIX, N° 10, pp 925-930, Oct. 1984.

[Moreschi, 88]

Alessandro Moreschi. The Last Castrato. Complete Vatican Recordings. *OPAL, Pavillon Records Ltd.*, England, 1988.