INVERSION OF A PHYSICAL MODEL OF A TRUMPET.

T. Hélie*, C. Vergez*, J. Lévine†, X. Rodet*

*IRCAM, Analysis/Synthesis Team, 1, Place Igor-Stravinsky, 75004 Paris, France

†Centre Automatique et Systèmes, 35, rue Saint Honoré 77305 Fontainebleau cedex, France

helie,vergez,rodet@ircam.fr,levine@cas.enscm.fr

Abstract

In this paper, we deal with the inversion of a physical model of a trumpet, i.e. how should the player control the model in order to obtain a given sound? After having shown that the inversion is a *ill-posed* problem, we add a physically based constraint which leads to a physically pertinent solution.

1 Introduction

Accurate parameter estimation methods have been developed for sound synthesis by signal models (e.g. additive synthesis) and have been very successful by allowing high quality and very flexible processing. But, in spite of some previous studies ([1], [2], [3]), there is a lack of similar estimation methods for most of physical models, particularly those describing instruments with self-sustained sound. In this paper, we propose a method to estimate the time-varying inputs of physical models of self-sustained musical instruments.

We first describe the physical model of trumpet-like instruments which is to be inverted. Then, we show that the model is *non invertible*: an infinity of inputs can produce the same sound. However, if it is assumed that the musician gestures are slower than the sound signal, a single input can be selected. We present a formal description of the method and then apply it to our model. Finally, we present some simulation results and conclude by looking at the efficiency of the method.

2 The physical model

We have built a model of trumpet-like instruments (cf. figure 1) with the aim of retaining most of the essential characteristics of the natural instrument sound production. The physical model for the lips is displayed in figure 2.

Stroboscopic visualisation of an artificial mouth confirmed that one lip often oscillates with an amplitude larger than the amplitude of the other lip (cf. [4]). Therefore, the oscillation of the upper lip only is simulated. Our lip model includes a *single*

Figure 1: General parts of the system

parallelepipedic mass m attached to a spring k and a damper r . The bore of the trumpet is modelled by its *acoustic reflection function* $h(t)$ derived from measurement on real trumpets in an anechoic room [5]. An air jet is assumed between the lips and the air velocity in the mouth is neglected. Then, the nonlinear coupling between the lips and the instrument is modelled by the *Bernoulli equation* which connects air flow $u(t)$ between the lips, lip's position $x(t)$, mouth pressure $p_m(t)$ and pressure $p(t)$ at the mouthpiece entry (cf. [6]).

Under the hypothesis of a plane wave propagation, we can write $p(t)$ as the sum of an *outgoing* wave $p_o(t)$ and an *incoming* wave $p_i(t)$. Let $\tilde{x}(t)$ be the ratio between the lip position

Figure 2: Model for the lips of the trumpet player: a single mass m with position x , is attached to a spring k and a damper r

$x(t)$ and the mouthpiece diameter at the lips, and $\xi = \text{sgn}(p_m - p_o - p_i)$. It is easily shown that $\xi = \text{sgn}(p_o - p_i) = \text{sgn}(p_m - 2p_i)$. Then we obtain the following physical model:

for opened lips ($\tilde{x}(t) \geq 0$),

$$\begin{cases} \ddot{\tilde{x}} + v\dot{\tilde{x}} + \omega^2\tilde{x} = Ap_m + B(p_o + p_i) \\ p_o = p_i - \xi \frac{C}{2} \left[C\tilde{x} - \sqrt{(C\tilde{x})^2 + 4|p_m - 2p_i|} \right] \\ p_i = h * p_o \end{cases} \quad (1)$$

and, for closed lips ($\tilde{x}(t) < 0$),

$$\begin{cases} \ddot{\tilde{x}} + 5v\dot{\tilde{x}} + 4\omega^2\tilde{x} = A(p_m - p_o - p_i) \\ p_o = p_i \\ p_i = h * p_o \end{cases} \quad (2)$$

where (cf. [7])

$v = \frac{r}{m}$ is the massic viscosity

$\omega = \sqrt{\frac{k}{m}}$ is proportional to a lips' contraction factor

A , B and C are constants which depend on the lip geometry, on the mass of the modelled lip, and on physical constants (*air density* and *sound velocity*)

h is the trumpet reflexion function which has been measured

In systems (1) and (2), the first equation describes the lip mechanics, the second equation describes the coupling between the lip and the instrument and the last equation describes the acoustic response of the instrument. At the output of the bell, the sound is given by $p_b = g * p_o$ where g is the acoustic transfer function of the instrument.

3 Problem statement

Ideally, the problem would consist in the recovery of the inputs (p_m , ω , v) of the trumpet physical model presented above by observing its outputs (p_b). The aim is that these inputs be adjusted automatically in the model in order to produce a desired sound, as if it was done by a real trumpet player.

However, we assume that p_m is measured and that p_o is the output of the model: this allows to lessen the calculation complexity without any loss of generality for the forthcoming method.

3.1 Non invertibility of the model

We now show that the inversion problem is *ill-posed*.

Proposition Assume that A , B , C , h and g are known and that neither internal nor output noise are considered. When $\tilde{x}(t) > 0$, the set of pairs (ω, v) corresponding to a given (p_o, p_m) is described by :

$$\omega^2 + \beta v + \gamma = 0 \quad (3)$$

$$\text{where } \begin{cases} \beta = \frac{1}{2} \frac{\partial \ln |q|}{\partial t} \\ \gamma = \frac{1}{2} \frac{\partial^2 \ln |q|}{\partial t^2} + \frac{1}{4} \left(\frac{\partial \ln |q|}{\partial t} \right)^2 \\ - \xi C \frac{\sqrt{|p_1|}}{p_1^2} ((A+B)p_m - Bp_1) \\ p_1 = p_m - p_o - p_i \\ p_2 = p_o - p_i \\ q = \frac{p_2^2}{p_1} \end{cases}$$

We therefore need to introduce an additional criterion to select a unique pair (ω, v) among the infinity of pairs satisfying (3).

Sketch of proof From system (1), we obtain:

$$\tilde{x}(t) = \frac{|p_o(t) - p_i(t)|}{C \sqrt{|p_m(t) - p_o(t) - p_i(t)|}} \quad (4)$$

$\tilde{x}(t)$ given by (4) is positive.

In fact, if $\tilde{x}(t) \leq 0$, we can write that $p_o(t) = p_i(t)$ but no further indications on ω and v can be obtained.

On the contrary, if $\tilde{x}(t) > 0$, we can express $\tilde{x}(t)$, $\dot{\tilde{x}}(t)$ and $\ddot{\tilde{x}}(t)$ as functions of p_o , p_m , ω and v . Substituting these functions for $\tilde{x}(t)$, $\dot{\tilde{x}}(t)$ and $\ddot{\tilde{x}}(t)$ in the first equation of (1), leads to (3).

4 Input reconstruction method

4.1 An additional constraint

Even though this is difficult to check, it is rather natural to assume that the internal dynamics of the model ($p_o(t)$, $p_i(t)$, $\tilde{x}(t)$) evolves as fast as the sound $s(t)$, and much faster than the musician's "gestures" i.e. than ω and v .

Thus, for a short-enough time window $\mathbb{I} = [t_b, t_e]$, the inputs ω and v are assumed to be constant.

4.2 Principle of the method

Let's consider a dynamical system described by

$$\mathcal{F}(\underline{S}, \tau, \underline{I}_{\mathbb{I}} \mid 0 \leq \tau \leq t) = \underline{0} \quad \forall t \in \mathbb{I} \quad (5)$$

where $\mathcal{F}$ is a *functional* (in particular $\mathcal{F}$ can include derivation or convolution operators), $\underline{S}$ is the vector of the *known* state variables (deduced from the observation), and $\underline{I}_{\mathbb{I}}$ is the vector of the *inputs*, constant on $\mathbb{I}$.

Let us introduce the partial mapping $\forall t \in \mathbb{I}$

$$f_{t, \underline{S}} : \underline{I}_{\mathbb{I}} \mapsto \mathcal{F}(\underline{S}, \tau, \underline{I}_{\mathbb{I}} \mid 0 \leq \tau \leq t) \quad (6)$$

Let $\mathbb{M}_t \triangleq f_{t, \underline{S}}^{-1}(\underline{0})$. It is obvious that for each $t \in \mathbb{I}$, the components of $\underline{I}_{\mathbb{I}}$ satisfy $\underline{I}_{\mathbb{I}} \in \mathbb{M}_t$. Since $\mathbb{M}_t$ evolves with t whereas $\underline{I}_{\mathbb{I}}$ is constant, $\underline{I}_{\mathbb{I}}$ is necessarily an *intersection point* of all the $\mathbb{M}_{t \in \mathbb{I}}$. Therefore $\underline{I}_{\mathbb{I}}$ belongs to $\cap_{t \in \mathbb{I}} \mathbb{M}_t$. Since this intersection

may sometimes be empty (which might result from model inaccuracies), the method we propose consists in minimizing the sum of distances between $\underline{I}_t$ and each $\mathbb{M}_t$.

Let us consider the *criterion* :

$$C_I : \underline{I} \mapsto \frac{1}{W} \int_I \|f_{t,\underline{I}}\|^2 dt \quad (7)$$

where $W = \int_I dt$ is a normalization factor.

We look for $\hat{\underline{I}}$ which minimizes C_I . Slowly varying inputs can also be considered in the same framework by considering sequences $\hat{\underline{I}}_j$ with constant inputs on each sub-interval $\mathbb{I}_j$.

4.3 Trumpet inputs reconstruction

The inputs are represented by $\underline{I} = \begin{bmatrix} \omega \\ v \end{bmatrix}$. We suppose that p_o is the observed signal, and $p_i = h * p_o$. So, from equation (4), we can reconstruct $\tilde{x}(t)$:

$$\tilde{x}(t) = \frac{|p_o(t) - (h * p_o)(t)|}{C \sqrt{|p_m(t) - p_o(t) - (h * p_o)(t)|}}$$

This is done when $\tilde{x}(t) > 0$ only. Therefore, instead of considering C_I , we use C_{I_+} where $\mathbb{I}_+ = \mathbb{I} \cap \mathbb{T}_+$ and $\mathbb{T}_+ = \{t \in \mathbb{R} / \tilde{x}(t) > 0\}$. Practically, it appears that the largest part of a period of the signal corresponds to positive $\tilde{x}(t)$, so that using $\mathbb{I}_+$ instead of $\mathbb{I}$ is not an inconvenience.

From now on, let us suppose $t \in \mathbb{T}_+$. Then, from equation (3), we obtain:

$$f_{t,\underline{I}} = \omega^2 + \beta(t)v + \gamma(t) \quad (8)$$

Thus, for the trumpet, $\mathbb{M}_t$ is a parabola in the (ω, v) plane, symmetrical about the axis $\omega = 0$, for each $t \in \mathbb{I}_+$. As described in (4.2), for constant inputs, $\hat{\underline{I}}$ is the intersection point between the $\mathbb{M}_t$ for all $t \in \mathbb{I}_+$ (as shown in figure 3).

Figure 3: $\hat{\underline{I}}$ is the intersection point between all the parabola $\mathbb{M}_t$.

Instead of looking for the intersection of the set of parabola, we can consider that (3) represents

a straight line with slope coefficient $-v$ and origin ordinate $-\omega^2$, passing through the measured points $(\beta(t), \gamma(t))$, $\forall t \in \mathbb{I}_+$.

In practice, the points are never perfectly aligned and a linear regression is required. Moreover, this allows to take into account the case of slowly varying inputs (ω, v) .

Since these computations must be implemented in discrete-time, we introduce the discrete set $\mathbb{I}_+^d = \mathbb{I}_+ \cap \mathbb{T}_s \mathbb{Z}$ where $\mathbb{T}_s$ is the sampling period. Consequently, the criterion (7) becomes

$$C_{I_+^d} : \underline{I} \mapsto \frac{1}{N} \sum_{t \in \mathbb{I}_+^d} \|f_{t,\underline{I}}\|^2 dt$$

where $N \triangleq \text{Card } \mathbb{I}_+^d$.

5 Results

Our method has been tested in two configurations where the inputs ω and v evolve slowly compared to the measured p_o (which agrees with our assumption). For the first test, we consider a signal p_o synthesized by the trumpet physical model for rigorously constant inputs. For the second test, we consider a constant input v and a slowly varying input ω .

In both cases, the model is parametrized by the same A, B, C, h which are typical of a real instrument in usual conditions. The sampling frequency is 10^5 Hz and the duration of each interval $\mathbb{I}$ is 10 ms.

Aberrant measurements are eliminated by a simple statistical criterion (cf. [7]). Since this aspect is not fundamental in our approach, we have omitted it in this paper.

5.1 Constant inputs

Figure 4 shows the estimated ω and v (represented by crosses) for the signal synthesized with $\omega = 2784.2$ I.S.U. and $v = 660$ I.S.U. (represented by a plain line). The bias and the variance of the estimates of ω are -0.0671 and 0.0118 respectively. Similarly, bias and variance of v are 1.2888 and 0.1929 respectively. In relative error, the worst result on bias is $2 \cdot 10^{-3}$ for v .

5.2 Slowly varying inputs

Figure 5 shows similar results for the signal synthesized with $v = 660$ I.S.U. and ω varying linearly from 2784.2 I.S.U. to 3627.9 I.S.U. in 0.5 s. From these estimations, the bias and the variance of the estimates of ω are -0.4923 and 15.9879 respectively. Similarly the bias and the variance of the estimates of v are 7.5459 and 8.6492 . In relative error, the worst result on bias is $1.14 \cdot 10^{-2}$ for v .

Figure 4: Analysis result for constant inputs (ω, ν)

Figure 5: Result of the analysis for slowly varying inputs (ω, ν)

The previous result seems to indicate that the inversion method is reliable even for slowly varying inputs. Moreover, we note that the evolution of the waveform of p_o seems not to influence the quality of the estimation.

6 Conclusion

We have presented a simplified model of the trumpet and of the lips of the player. This model gives satisfactory perceptive results but it is *non invertible*. However, we have developed a method which allows us to obtain a unique solution by using an additional constraint. We are currently working on an extension of the method to allow the recovering of the blowing pressure p_m as well. An extension of the inversion algorithm to noisy measurements and faster inputs is being developed.

References

- [1] J. Schroeter and M. M. Sondhi. *Advances in Speech Signal Processing*, chapter Speech coding based on physiological models of speech production, page p231:268. Marcel Dekker Inc, NewYork, 1992.
- [2] A. T. Cemgil and C. Erkut. Calibration of physical models using artificial neural networks with application to plucked string instruments. In *ISMA '97 Proceedings*, Edinburgh, 1997.
- [3] P. Guillemain, R. Kronland-Martinet, and S. Ystad. Physical modelling based on the analysis of real sounds. In *ISMA '97 Proceedings*, Edinburgh, 1997.
- [4] C. Vergez and X. Rodet. Experiments with an artificial mouth for trumpet. In *ISMA98*, Leavenworth, Washington State, USA, june 1998.
- [5] C. Vergez and X. Rodet. Comparison of real trumpet playing, latex model of lips and computer model. In *Proceedings ICMC'97*, Thessalonike, September 1997. ICMC.
- [6] A. Hirschberg, X. Pelorson, and J. Gilbert. Aeroacoustics of Musical Instruments. *Meccanica*, 31:131-141, 1996.
- [7] T. Hélie. Étude de méthodes d'estimation des paramètres d'un modèle physique d'instrument de musique-application à la trompette. Master's thesis, DEA Automatique et Traitement du Signal, Université de Paris XI, Orsay, septembre 1998.