Virtual Analog Modeling in the Wave-Digital Domain

Giovanni De Sanctis and Augusto Sarti, *Member, IEEE*

***Abstract*—We refer to "Virtual Analog" (VA) as a wide class of digital implementations that are modeled after nonlinear analog circuits for generating or processing musical sounds. The reference analog system is therefore typically represented by a set of blocks that are connected with each other through electrical ports, and usually exhibits a nonlinear behavior. It therefore seems quite natural to consider Nonlinear Wave Digital modeling as a solid approach for the rapid prototyping of such systems. In this paper, we discuss how nonlinear wave digital modeling can be fruitfully used for this purpose, with particular reference to special blocks and connectors that allow us to overcome the implementational difficulties and potential limitations of such solutions. In particular, we address some issues that are typical of VA and physical modeling, concerning how to accommodate special blocks into WD structures, how to enable the interaction between different WD structures, and how to accommodate structural and topological changes on the fly.**

***Index Terms*—Binary connection tree, physical modeling, wave digital filters.**

I. Introduction

The term "Virtual Analog" (VA) is commonly used for describing digital implementations of analog circuitry employed for generation, enhancement or processing of musical sounds. The typical approach to the development of virtual analog algorithms consists of starting with the reference analog circuit and, from there, developing a digital implementation that will mimic the behavior of the analog reference while providing improved flexibility and, perhaps, incorporating additional features; extending its range of applicability; or enhancing its performance. This loose definition of VA extends to a wide class of musical sound synthesis methods based on physical modeling, where the physical reference system is not necessarily an electrical circuit, but can always be represented by an electrical equivalent circuit. This way VA ends up covering sensors and transducers that are often encountered in vintage circuitry, whose nonlinear behavior is of significant musical interest. Examples of VA systems are: valve or solid-state amplifiers; vintage musical instruments with special magneto-electric transduction (e.g., electromechanical pianos or "Hammond"-style electrical organs, etc.); electronic circuits for sound generation (e.g., "Moog"-style voltage-controlled oscillators, filters, amplifiers, or envelope generators); or special electromechanic or electronic circuits for vintage sound effects (e.g., spring reverberators, analog chorus circuitry, etc.).

What makes analog circuits and systems interesting to model in a digital fashion is usually their inherent nonlinear behavior, whether introduced by a valve in a final amplification stage, or by some special transducers. This nonlinear behavior, in fact, is usually responsible for characteristic sounds, or deep timbral dynamics, which are often received with great interest by musicians and listeners. It is therefore very important to plan for a flexible digital implementation of such analog systems that is aimed at preserving this nonlinear behavior as part of the modeling process. It is also equally important to adopt a design process that will enable the VA system to extend the performance and the range of applicability of the physical reference by including, for example, novel physically-plausible controls or enabling time variance in parameters that are normally static in real life. This would offer the musician a wider "timbral space" to explore.

VA systems benefit from an extensive literature on sound generation through physical modeling. The digital modeling of analog (physical) systems has been pursued by musicians and researchers for quite some time. Although this approach was initially intended for simulation purposes, it recently became palatable for applications of sound generation and processing, as massive computational power became available at low cost. The reasons of the interest in physical modeling are numerous, and are mostly related to the intuitive link that exists between control and model reaction; to the ability to generate musically interesting timbral spaces associated to the changing of just a few physically plausible parameters; and to the possibility to exploit a well-established modeling experience in musical acoustics. If, instead of focusing on simulation accuracy, we look at the physical modeling problem from the user's point of view; however, we are immediately faced with new challenges. In order to become an approach of practical interest, in fact, physical modeling needs to become a systematic and automatic methodology based on a simple modeling metaphor (instantiated by a meaningful graphical user interface). Through this methodology, the user should be spared all the algorithmic and physics-related aspects of the modeling process and should be able to construct a sound generation/processing algorithm through a limited number of decisions (structural and parametric) of immediate and intuitive impact on the final result.

Our starting point is always a lumped circuit, intended as a set of blocks that are connected to each other through ports, each characterized by a pair of dual variables: a *through* variable, such as velocity, flow, or current; and an *across* variable,

Manuscript received March 31, 2009; revised September 09, 2009. Current version published April 14, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Udo Zoelzer.

G. De Sanctis is with the Sonic Arts Research Centre, Queen's University Belfast, BT7 INN Belfast, U.K. (e-mail: gdesanctis01@qub.ac.uk).

A. Sarti is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milan, Italy (e-mail: sarti@elet.polimi.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASL.2009.2033637

such as force, pressure or voltage. Whether this circuit is natively an electrical one or whether it is the electrical equivalent representation of a mechanic or fluid-dynamic system is, for the moment, irrelevant. We are, however, interested in a solution that preserves the identity of the ports after discretization and allows us to monitor the energy flow through such ports to keep the algorithmic stability under control. A circuit described by pairs of across-through variables is called a "Kirchhoff" (K) representation, characterized by a set of "local" equations that describe the building blocks and a set of "global" equations (Kirchhoff laws, laws of dynamics) that describe how blocks are interconnected (topology). A modeling approach based on K variables consists of discretizing the whole set of (local and global) implicit equations at once. As a consequence, changing anything in the reference circuit would usually require the discretization process to restart. Furthermore, if a nonlinear element is present in the circuit, a K-based solution would usually require solving an implicit equation at every time step, with a negative impact on computational efficiency. Standard ordinary differential equation (ODE) solvers are excellent solutions for simulation problems as they are aimed at maximizing simulation accuracy over a range of problems that is as wide as possible. However, generality and accuracy are not at the top of our priorities. We are interested, in fact, in solutions that will be handled primarily by musicians, who will interact with the resulting algorithms with the intent of exploring the associated timbral space while pushing the boundaries of algorithmic controls and stability. We are therefore more interested in the inherent plausibility (self-consistency) of the result rather than in its strict adherence to the behavior of the analog reference circuit (simulation accuracy). This means that we will be willing to trade some accuracy and generality for the guarantee that whatever instability the musician will experience can be solely attributed to the reference analog circuit and not its algorithmic implementation. Another crucial aspect is interactivity, which implies real time operation and reduced delay. With this in mind, we cannot generally afford increasing the sampling frequency just to guarantee that the algorithm will not behave in an unexpected fashion when modeling nonlinear circuits. For example, a discrete implementation based on standard ODE solvers of the nonlinear subharmonic oscillator described in [7] will exhibit the expected chaotic behavior only when the sampling frequency is much higher than the one suggested by the bandwidth of the involved signals. These, and others that will be clarified later on, are the reasons why we are interested in solutions based on Digital Waves, which offer the degree of flexibility and stability preservation that we are looking for.

Fettweis introduced the theory of Wave Digital Filters [1] (WDF) a few decades ago, with the intent of working with scattering parameters in the digital domain and exploiting their properties in the hope that they would be preserved by the digital implementation. In order to build a WDF, all K pairs of variables that describe the various circuit ports are replaced by a pair of digital waves (incident and reflected), obtained as a linear transformation of the K pair. This linear transformation is designed to turn the port representation into an input/output relationship. This circuit-inspired approach to the physical modeling of linear filters exploits the scattering nature of the involved variables and appropriate adaptation conditions to produce an explicit implementation of the K laws.

The literature of physical modeling is rich with solutions developed in the Wave (W) domain which are closely related to WDFs. A well-known example of this sort is given by the Digital WaveGuides (DWG) [9]. DWG elements are particularly suitable for modeling acoustic resonating structures that are fully compatible with WDFs. Hybrid WDF/DWG [10] structures represent a good solution to the problem of sound synthesis by physical modeling as, besides referring to an acoustic instrument, they are based on a local (block-based) discretization of the physical elements that constitute the analog model. In other words, these solutions open the way to the development of a flexible synthesis technique based on the interconnection of predefined building blocks. Traditional WDF/DWG structures, however, are inherently linear; therefore, they cannot easily incorporate nonlinear elements. In the past few years, however, new solutions and methodologies have emerged, which allow us to generalize the WDF/DWG framework in order to encompass a wide class of nonlinearities with memory, without giving up flexibility and modularity in the synthesis strategy [6], [7], [10]. An embryo of these ideas was first presented [6] and then more completely formalized in [7]. More recently, the authors proposed a method that enables the automatic construction of physical models of analog circuits that are suitable for applications of musical acoustics [3], which brings the design of nonlinear WD structures to a level of practical usability while enabling the modeling of a wide variety of nonlinear physical models in a completely automatic fashion.

In this paper, we push this last approach forward by proposing WD blocks and elements that allow us to overcome some of the inherent difficulties that are encountered when dealing with VA modeling problems. We will discuss categories of NLE that are typical of VA; WD blocks that allow us to include, under certain conditions elements that normally cause problems of computability; blocks that allow us to implement new elements that are crucial for VA applications, but had not been considered before (e.g., comparators, cross controllers, etc.). Finally, we discuss the important issues of time-varying parameter control and usability of the models.

II. Wave Digital Structures and Connection Trees

In this section, we provide a brief overview on nonlinear WD structures and on the Binary Connection Tree (BCT) [3]. This overview has the purpose of making the manuscript as self-contained as possible and, at the same time, tuning the reader's perspective towards VA applications. We first discuss some basic blocks that are needed for the construction of WD structures, just enough to brush up our knowledge of WD structures. A more complete description of WDF structures and related blocks can be found in [1]. More WD blocks can be found in the literature of DWGs (see for example [9]). A more comprehensive treatment of nonlinear WD structures can be found in [4]–[7], while solutions for their automatic implementation are thoroughly described in [3], [8].

TABLE I
COMMON CIRCUIT ELEMENTS AND THEIR WAVE REPRESENTATION

	KIRCHHOFF	WAVE	W ADAPTED
R	$v = Ri$	$b = \frac{R-R_0}{R+R_0}a$	$b = 0$
REAL SOURCE	$v = v_0 - Ri$	$b = \frac{2R_0 v_0 + (R-R_0)a}{R+R_0}$	$b = v_0$
IDEAL VOLTAGE SOURCE	$v = v_0$	$b = 2v_0 - a$	—
IDEAL CURRENT SOURCE	$i = i_0$	$b = 2R_0 i_0 + a$	—
C	$i = C\frac{dv}{dt}$	$a_n - b_n = \frac{2RC}{T}\frac{b_k + a_k - b_{k-1} - a_{k-1}}{b_k + a_k + b_{k-1} + a_{k-1}}$	$b_k = a_{k-1}$
L	$v = L\frac{di}{dt}i$	$a_n + b_n = \frac{2L}{RT}\frac{b_k - a_k - b_{k-1} + a_{k-1}}{b_k + a_k + b_{k-1} + a_{k-1}}$	$b_k = -a_{k-1}$
SERIES JUNCTION	$\sum_{m=1}^{3} v_m = 0$ $i_1 = i_2 = i_3$	$\mathbf{b} = \mathbf{a} - \begin{bmatrix} \gamma_1 & \gamma_1 & \gamma_1 \\ \gamma_2 & \gamma_2 & \gamma_2 \\ \gamma_3 & \gamma_3 & \gamma_3 \end{bmatrix} \mathbf{a}$ $\gamma_m = \frac{2R_m}{R_1+R_2+R_3}$	$\gamma_3 = 1$ $\gamma_2 = 1 - \gamma_1$
PARALLEL JUNCTION	$\sum_{m=1}^{3} i_m = 0$ $v_1 = v_2 = v_3$	$\mathbf{b} = \begin{bmatrix} \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \\ \delta_1 & \delta_2 & \delta_3 \end{bmatrix} \mathbf{a} - \mathbf{a}$ $\delta_m = \frac{2G_m}{G_1+G_2+G_3}$	$\delta_3 = 1$ $\delta_2 = 1 - \delta_1$

A. *Crash Overview of Basic WD Structures*

A circuit port is defined by a K pair (v, i). A one-port element is, in fact, characterized by a relationship between such variables. This port representation can be turned into an explicit representation (an I/O relationship) through a linear transformation [1] of the form

$$
\begin{aligned}
a &= v + R_0 i \\
b &= v - R_0 i
\end{aligned}
\tag{1}
$$

where a and b are the *incident* (input) and *reflected* (output) waves, respectively, and R_0 is a nonzero free parameter called *reference resistance*; therefore, this transformation is always invertible. In general, the WD representation of a block can be obtained from its (continuous time) K representation, by applying this change of variables and discretizing the result through the bilinear transformation. The WD representation of elementary blocks such as resistors, capacitors, and inductors, can be readily obtained by following this procedure, and the result is summarized in Table I. As we can see, a careful choice of the free parameter R_0 leads to a WD representation of the block that is free from instantaneous I/O dependencies. This *adaptation* condition is crucial for constructing computable WD structures, as it can be used in order to avoid delay-free loops.

Junctions are structural elements that enable the parallel/series interconnection of multiple one-port elements in compliance with the Kirchhoff laws. It is through such blocks that the interconnection topology of the circuit is implemented. In order to obtain their WD representation, we need to start from the Kirchhoff laws that describe them and apply the above change of variables. The result will be a multi-port parallel or series junction, depending on which K laws have been used to begin with. The WD representations of such junctions in the three-port case are shown in Table I.

As the sum of all reflection coefficients γ_i is bound to be equal to 2, the port resistances can always be chosen to balance out in such a way to have $\gamma_n = 1$, for any port n, which makes port n reflection-free [1]. The condition on port resistances (admittances) in the case of series (parallel) junctions is $R_1 + \cdots + R_{n-1} + R_{n+1} + \cdots + R_N = R_n$ (for parallel junctions we have $G_1 + \cdots + G_{n-1} + G_{n+1} + \cdots + G_N = G_n$), which means that port n must match the series (parallel) of all other port resistances (admittances). A junction that satisfies this condition on port resistances is called an adaptor, which is a very important building block as it enables the interconnection between elements in such a way that the resulting structure will be computable. To be precise, in the seminal work of Fettweis [1], what we call adaptors are also referred to as "constrained adaptors." The term junction, widely used in the literature of DWGs, is a generic term that covers both unconstrained and constrained adaptors.

A generic N-port series (parallel) adaptor can always be decomposed into a chain of $N-2$ series (parallel) three-port adaptors connected in such a way to avoid delay-free loops. In fact, $N-3$ of the available $N-2$ adapted ports will be used for adaptor-to-adaptor interconnections, and only one will be left for connections with other blocks (the adapted port of the N-port adaptor).

A tree-like interconnection of parallel and series three-port adaptors is referred to as a MacroAdaptor (MA) [3], whose properties are similar to those of multi-port adaptors, as it is non-energetic (the algebraic sum of all flows of energy directed towards the adaptor is always zero) and has one adapted port available for MA-to-block connections [2]. Most WD computable circuits of interest can be seen as a number of WD blocks that are connected to each other through one such MA. A slightly wider category of circuits can be covered if we consider MAs whose outer ports connect to the elements either directly or through WD transformers or some other non-energetic two-port blocks such as gyrators [1]. In fact, even if such elements turn out to be placed between two of the adaptors that constitute the MA, they can always be "moved outward" using some simple equivalency rule (e.g., an adaptor with a transformer at one of its ports can be transparently replaced by an adaptor with the same transformer on the other two ports).

There exists a special class of non-energetic two-port junctions that further extends the range of applicability of WD structures. This is the class of Dynamical Scattering Junctions (DSJs) [6], [7].

1) Dynamical Scattering Junctions and Adaptors: Dynamical scattering junctions refer to a generalization of the concept of scattering waves, which were first introduced in [7] to accommodate a wider range of nonlinear elements into a WD structure. We start from a K pair (v, i) and introduce a linear transformation of the form

$$
\begin{aligned}
A(z) &= V(z) + R_0(z) I(z) \\
B(z) &= V(z) - R_0(z) I(z)
\end{aligned}
\tag{2}
$$

where A and B are the Z-transforms of the *incident* (input) and *reflected* (output) waves, respectively, and $R_0(z)$ is a *reference* *transfer function* (RTF), which is assumed to be causal and passive (positive real in the outer disk). A DSJ is a two-port element designed to transform the wave pair (a_1, b_1), referred to the RTF $R_1(z)$, into the wave pair (a_2, b_2), referred to $R_2(z)$. Let $A_i(z) = V_i(z) + R_i(z)I_i(z)$ and $B_i(z) = V_i(z) - R_i(z)I_i(z)$, $i = 1, 2$, be the incident and reflected waves, respectively, that flow through the junction. Using the continuity constraints $V_1(z) = V_2(z)$ and $I_1(z) + I_2(z) = 0$ we can derive the output waves as a function of the input waves

$$\begin{aligned} B_1(z) &= K(z)A_1(z) + (1 - K(z))\, A_2(z) \\ B_2(z) &= (1 + K(z))\, A_1(z) - K(z)A_2(z) \end{aligned} \tag{3}$$

where

$$K(z) = \frac{R_2(z) - R_1(z)}{R_2(z) + R_1(z)}.$$

The direct implementation of (3) is known as Kelly–Lochbaum scattering cell [9], whose reflection coefficient is replaced with the reflection filter $K(z)$. In order to avoid instantaneous reflection of the incident waves, we just need the reflection filter to exhibit no instantaneous input/output connection, i.e., $K(z) = z^{-1}\hat{K}(z)$, with $\hat{K}(z)$ causal and stable.

This definition of DSJ was generalized into that of dynamic adaptors in [7], along with an extension of the concept of adaptation (instantaneous adaptation; partial or total adaptation), and a proof of the fact that such adaptors are, in fact, non-energetic. A similar generalization was proposed for Digital WaveGuides in [13], along with a discussion of the non-energetic behavior of its junctions. We then showed in [3] that a wide range of nonlinear circuits can be modeled as a number of WD blocks that are connected to each other through one dynamic MA (interconnections of dynamic 3-port adaptors). Finally, we showed in [8] that a generic dynamic MA can be equivalently implemented as a memoryless MA surrounded by dynamic scattering cells. This is why DSJs are, in fact, very important for WD structures and deserved to be described here.

2) Mutators: A special case of DSJ is the WD mutator [6], [7], which allows us to accommodate reactive nonlinearities such as nonlinear inductors and nonlinear capacitors into a WD structure.

The $R-C$ mutator is a DSJ between a capacitive RTF and resistive one. If the analog reference port impedances are $R_1(s) = R$ and $R_2(s) = 1/(sC)$, $C > 0$, the reflection filter is an all-pass of the form

$$K(z) = \frac{p + z^{-1}}{1 + pz^{-1}}, \quad p = \frac{T - 2RC}{T + 2RC}.$$

If $R = T/2C$, then both ports turn out to exhibit no local instantaneous reflection, as $K(z) = z^{-1}$. The mutator therefore becomes a Kelly–Lochbaum scattering cell whose reflection coefficient is replaced by a delay element.

The $R - L$ mutator is a DSJ between an inductive RTF and a resistive one, therefore the situation is symmetrical with respect to that of the capacitive mutator. We start from $R_1(s) = R$ and $R_2(s) = sL$, $L > 0$, which corresponds to a scattering digital filter

$$K(z) = -\frac{p + z^{-1}}{1 + pz^{-1}}, \quad p = \frac{T - 2L/R}{T + 2L/R}. \tag{4}$$

Instantaneous reflections can be eliminated at both ports of the scattering junction by letting $R = 2L/T$, in which case we have $K(z) = -z^{-1}$, as expected from the classical WDF theory [1].

B. Topological and Structural Issues

There are two important issues that arise from using WD methods for VA modeling. One concerns the constraints in the interconnection topology and one concerns the constraints in the number of NLEs.

An interconnection between WD elements through parallel and/or serial adaptors is known to result in a computable WDF because the corresponding signal flow diagram does not contain any delay-free directed loops [1]. This is based on the assumption that the network of adaptors has a *tree-like* structure, which is true for most reference physical systems of interest that are encountered in VA and in musical acoustics [3]. To be more precise, it is worth mentioning the work of Franken and Ochs [12], according to which an arbitrary circuit structure can always be turned into a tree-like structure made of parallel and series nodes, as well as triconnected "R-type" nodes [12] (which are inherently built, for example, as a star-shaped connection embedded in a triangle-shaped connection). As such blocks are still relatively unexplored in the literature of WD implementations (particularly as far as initialization issues are concerned), in this paper we will only focus on MAs that do not include them. It is our intention, however, to explore more general types of MAs that include triconnected R-type blocks in the near future.

As already discussed above, a tree-like network of (dynamic or memoryless) adaptors is here referred to as a (dynamic) Macro-Adaptor (MA), and a port that exhibits no (instantaneous) local reflections is called an *adapted* port. As reflected waves at this port are either absent or delayed, this port can be freely connected to another port without causing computability problems. As a non-adapted port can only be connected to an adapted one, the resulting MA ends up having one adapted port only, which can be used for connecting a nonlinear element (NLE). The fact that MAs can only directly accommodate one NLE is a typical limitation of WD structures, including WDFs [4]. If more than one NLE were present, we could go for an ad-hoc implementation based on replacing the definitions of the waves into the K representation of a subportion of the circuit that includes such NLEs. Usually this results in a set of equations that needs be solved at every time step (usually with some iterative method). In conclusion, the presence of multiple NLEs reduces the effectiveness of our approach but it does not prevent us from using it. What we lose is some of the computational efficiency and the possibility to automate the implementation procedure.

One final issue concerns the fact that some WD structures are based on dynamic MAs, as envisioned in [7]. We recently showed [8], however, that a computable tree-like interconnection of dynamic adaptors is completely equivalent to a memoryless MA whose outer ports might be connected to DSJs. The proof of this equivalency is constructive and can be done in two steps:

1) replace all three-port parallel/series adaptors with memory with memoryless adaptors of the same type, whose adapted port is connected to a properly defined DSJ;

2) have all the inner DSJs (those that end up in between two adaptors) slide outward until they reach the periphery, according to specific sliding rules.

More details about this procedure can be found in [8].

C. Connection Tree

A nonlinear WD structure based on a dynamic MA that connects a number of building blocks can always be implemented using a Wave Tableau (WT), which is a reduced matrix representation in the wave domain. A method for turning this representation into an explicit form (state update equation) is described in detail in [3]. This method, however, is outperformed in computational complexity and flexibility by a method based on connection tree inspection, an overview of which is included in this subsection.

As already said above and discussed in detail in [3], a NL circuit for VA can always be turned into a WD structure that is made of a number of building blocks connected with each other by a multi-port memoryless MA, possibly through DSJs [7]. A computable N-port MA is a tree-like interconnection of $N-1$ three-port parallel and/or series adaptors [3]. This topology is therefore that of a *binary connection tree* with:

- $N-2$ nodes, which are the three-port adaptors;
- $2N-3$ branches, which represent the port-to-port connections of the circuit, namely:
  - — $N-3$ inner branches, which represent the node-to-node (adaptor-to-adaptor) connections: for reasons of computability, such connections require at least one of the two facing ports to be adapted;
  - — N outer branches, which represent the node-to-leaf connections: as $N-3$ of the $N-2$ available adapted ports have already been used for the inner branches, only one of the outer branches can be adapted;
- $N-1$ leaves, which are the linear building blocks, to be connected to the non-adapted outer branches;
- 1 root, which is the nonlinear building block, to be connected to the only adapted outer branch.

We showed in [3] that this structure can be automatically implemented through an algorithm that iteratively inspects this tree through a "forward scan" (from the leaves towards the NLE) followed by a "backward scan" (from the NLE towards the leaves). The computation always starts from the leaves of the tree, as they contain the "memory elements" with the related initial conditions. At the end of the forward scan the wave reflected from the NLE is evaluated, then the backward scan computes all the waves until it reaches the leaves again, where all memory cells are finally refreshed.

One major advantage of this iterative implementation lies in the fact that its computational cost and its memory requirements increase linearly with the number of adaptors, while the complexity of the WT method grows quadratically with the number of circuit elements. What makes this approach interesting, however, is its remarkable flexibility, which allows us to change parameters and even its topology on the fly. We will see later that the possible topological changes are:

- grafting: two separate models merge into a single one because depending on an appropriate proximity condition (e.g., a piano hammer hits a string);
- pruning: two models split into independently evolving submodels depending on an appropriate proximity condition (e.g., a piano hammer separates from a string after hitting it);
- bridging: two trees with independent roots (NLEs) are joined by outer branches (through appropriate elements).

The reference circuit normally includes dynamic elements such as capacitors or inductors. The corresponding WD representation of such elements includes memory elements that need to be initialized with values that reflect a specific initial condition of the circuit. This initialization process becomes particularly important when dealing with circuits that are the electrical equivalent of some mechanical or fluid-dynamic system. In this case, in fact, choosing the wrong initial conditions could result in the wrong placement of mechanical elements of the system. The initialization process for BCTs is described in detail in [3].

D. Needs and Requirements for VA Applications

The binary connection tree was introduced with the purpose of implementing WD structures in an automatic, efficient, and flexible fashion. Such structures accommodate all the typical blocks that were envisioned and defined in the theory of WDFs [1], as well as those that are commonly used in DWGs [9]. They also accommodate a wide range of NLEs (algebraic NLEs) through mutators [3], [8]. Some very specific elements, however, cannot be naturally accommodated in WD structures, and need to be addressed with ad-hoc solutions. This issue will be discussed in Section III. Here we will consider the problem of non-adaptable elements and that of elements that are defined in the K domain.

An important issue that we encounter in typical VA applications is that of accommodating time-variance. The adoption of the BCT approach allows us to address parametric time variance as already discussed in [3]. More interesting is the issue of topological time variance, i.e., the ability of WD structure to accommodate changes in block interconnections on the fly. Time-varying interconnections are, in fact, encountered when dealing with multiple interacting objects, whose interconnection is governed by some proximity condition [15]. This issue can be addressed by exploiting the BCT approach. How to do so will be discussed in Section IV. Another crucial aspect that is not naturally addressed in WD structures is the dependence between WD elements. This dependence, in fact, is quite common in a variety of physical models, and needs to be addressed through the definition of specific cross-control mechanisms. This issue will be thoroughly addressed in Section IV.

III. WD Building Blocks for VA

The WD blocks that we consider can be roughly classified into physical blocks, which have an analog counterpart in the reference circuit, and structural blocks, which are introduced for purposes related to the WD structure. Structural blocks are typically two-port element that are aimed at the following:

- *Modifying the behavior of the blocks that they are connected to*—this is usually aimed at simplifying the implementation of elements, by enabling parameter changes even where this feature was not included in the element's

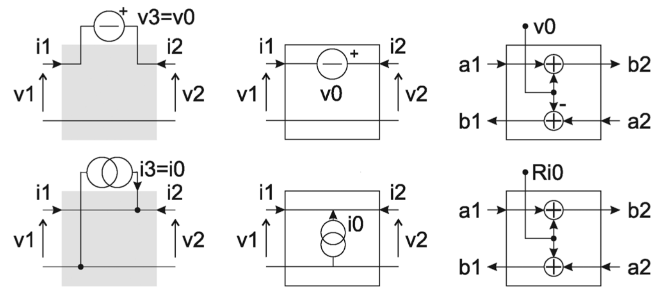


Fig. 1. Series IVS (top) and parallel ICS (bottom). From left to right: analog circuits with emphasized connections; reference circuits; resulting W structure when the ideal source is connected to the adapted port of the junction and the other two ports have the same reference resistance.

design. This is particularly useful for NLEs, as it effectively simplifies its implementation. Such modifications include rigid translations of characteristics, scaling, etc.

- *Accommodating elements that otherwise would not be possible to include*—e.g., elements whose port cannot be adapted, such as NLEs and ideal generators.
- *Connecting a block that is defined in a different domain other than the Wave Digital one*—this refers, for example, to the K-W conversion, which is aimed at incorporating in the WD design some elements (or whole sections of circuits as seen from one of their ports) that are already implemented in the K domain, without having to start over with the discretization process.

We will discuss such blocks through specific examples.

A. Ideal Sources

Ideal sources are among those WD elements that cannot be adapted, therefore it is not possible to connect them to any non-adapted port of the MA. The fact that they cannot be leaves of a connection tree generally prevents us from using them in a WD structure. However, we will show here that it is possible to define a two-port element that contains a series ideal voltage source (IVS) or a parallel ideal current source (ICS) inside its structure, which can be used instead of a three-port adaptor for accommodating the generator, as shown in Fig. 1. In the case of an IVS (ICS), these two-port elements allow us to rigidly translate along the v axis (i axis), the characteristics of an element. They can also be used when an IVS (ICS) is needed, provided that the rest of the structure can be split into two blocks that are connected in series (parallel).

The relationships for the series IVS, using the sign conventions of Fig. 1, are $v_2 = v_1 + v_0$ and $i_2 = -i_1$. Assuming that reference resistance of the first port is equal to that of the second port, $R_1 = R_2 = R$, in the WD domain the above equations become

$$\begin{cases} \frac{a_2+b_2}{2} = \frac{a_1+b_1}{2} + v_0 \\ \frac{a_2-b_2}{2R} = -\frac{a_1-b_1}{2R} \end{cases} \Rightarrow \begin{cases} b_1 = a_2 - v_0 \\ b_2 = a_1 + v_0. \end{cases} \tag{5}$$

We can follow a similar procedure for the parallel ICS, obtaining the pair of equations $b_1 = a_2 + Ri_0$ and $b_2 = a_1 + Ri_0$. The same result can also be achieved by working entirely in the WD domain. Starting from an ideal source connected to the adapted port of a three-port series adaptor and forcing the two ports to have the same reference resistance. This choice of port resistances, corresponds to the reflection coefficients $\gamma_1 = \gamma_2 = 0.5$. If we then consider the ideal source as part of the adaptor, we end up with a two-port element whose ports are both adapted, as they only depend on the incident wave of the other port and on the output of the ideal source.

The value v_0 (i_0 for the parallel ICS) or the incident wave coming from the source can be treated as a parameter, disregarding the corresponding reflected wave which, if needed, can be readily computed to be $b_3 = a_1 + a_2$. These results are consistent with the classic WDF theory, as by terminating one of the two-port of the IVS (ICS) with a short-circuit (open-circuit), we re-obtain its one-port counterpart.

B. Generic Linear Transformation

In many applications there is the need to parametrically alter or control the characteristics of an element. Aside from the two-port elements that we find in the classical circuit theory (e.g., transformers and gyrators) we can define other transformations in quite an arbitrary fashion. Examples are scalings with independent scale factors for each axis, rotations, or reflections of the characteristics. Let us consider the generic linear transformation on the $i - v$ plane

$$\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} k_v & r \\ g & k_i \end{bmatrix} \begin{bmatrix} v_2 \\ i_2 \end{bmatrix}. \tag{6}$$

Using (1), we can readily derive the corresponding W formulation

$$\begin{aligned} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} &= \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \\ t_{11} &= \frac{k_v R_2 - g R_1 R_2 - r + k_i R_1}{k_v R_2 + g R_1 R_2 - r - k_i R_1} \\ t_{12} &= -\frac{2R_1(k_v k_i - rg)}{k_v R_2 + g R_1 R_2 - r - k_i R_1} \\ t_{21} &= \frac{2R_2}{k_v R_2 + g R_1 R_2 - r - k_i R_1} \\ t_{22} &= -\frac{k_v R_2 + g R_1 R_2 + r + k_i R_1}{k_v R_2 + g R_1 R_2 - r - k_i R_1}. \end{aligned}$$

In order to guarantee that the resulting two-port structure will not affect the computability of the WD circuit, the port resistances must balance out in such a way to have $t_{22} = 0$. As an example, for the transformer with turn ratio n, whose equations are $v_1 = nv_2$ and $i_1 = -(1/n)i_2$ we have

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \frac{1}{nR_2 + \frac{R_1}{n}} \begin{bmatrix} nR_2 - \frac{R_1}{n} & 2R_1 \\ 2R_2 & \frac{R_1}{n} - nR_2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

The adaptation condition will therefore be $R_2 = (1/n^2)R_1$, which results in much simpler WD relations of the form $b_1 = na_2$ and $b_2 = (1/n)a_1$, which are valid for any nonzero value of n.

Similar is the situation of the gyrator with gyration resistance R_G, whose equations are $v_1 = R_G i_2$ and $i_1 = -(1/R_G)v_2$. The corresponding WD representation, under the adaptation condition $R_G^2 = R_1 R_2$, in this case, turns out to be given by the following pair of WD relations $b_1 = (R_1/R_G)a_2$ and $b_2 = -(R_G/R_1)a_1$. A special form of gyrator, known by the name of *dualizer*, is obtained by applying the condition

$R_G = R_1$ (which yields $R_2 = R_1^{-1}$). The dualizer has the interesting property of transforming a generic element into its *dual*. Its wave representation is simply given by the pair of WD equations of the form $b_1 = a_2$ and $b_2 = -a_1$.

C. K-W Converter

Numerous models in literature of VA and physical modeling are implemented in the Kirchhoff domain, as this is the "natural" domain for a number of distributed parameters modeling techniques. In order to use such models in a WD structure, these models need to be modified so that they can accept and return digital waves. The most obvious solution consists in reworking the whole model in the W domain, although this is not always a viable or convenient solution. For example, if we want to exploit a specific discretization technique that is particularly efficient for a certain purpose or, if the model is controlled by physically meaningful or intuitive parameters that would be lost in the conversion to the W domain, then this may not be the best way to proceed. In addition, we may not be willing to change an already existing implementation in the K domain. Hence, a second option is the interposition, between the K model and the WD structure, of a two-port element that will act as an "interpreter" between the two worlds by applying (1) at run time (see also [14]).

In the rest of this section, we will assume that the K model has the dimensions of an impedance, i.e., it returns a voltage when the input is a current. A similar result can be obtained by using an admittance model. Apparently, a direct implementation of the K-W transformation, as in Fig. 2, would introduce two delay-free loops, one on the K port and one on the W port. It can be shown (see also [16]) that if the K-model has a direct (i.e., delay-free) input-output connection, a proper choice of the W port's reference resistance R will eliminate all the delay-free loops and the element, as seen from the W port of the K-W converter, will be adapted. Assuming that the impedance filter of K-model can be decomposed into the sum of an instantaneous term and a term delayed by at least one sample, $Z(z) = (V(z)/I(z)) = R_0 + z^{-1}R_1(z)$, where R_0 is positive real and R_1 is a causal filter, the adaptation condition is achieved by imposing $R = R_0$. Referring to Fig. 2(a), this can be verified by noticing that the instantaneous term of the transfer function between a and the point labeled as w

$$\frac{W(z)}{A(z)} = \frac{2R_0^{-1}(R_0 + z^{-1}R_1)}{1 + R_0^{-1}(R_0 + z^{-1}R_1)} = 1 + \frac{z^{-1}R_1/R_0}{2 + z^{-1}R_1/R_0}$$

is equal to one and cancels out with the other term confluent to the same summation node. This implies a null instantaneous part of the transfer function between a and b, or the adaptation condition. As a simple check, when $Z = R_0$, i.e., when Z is a pure, constant resistance, we can verify that the reflected wave is null as expected from the classic WDF theory, and the K-W converter will not introduce any delay-free loops.

If $R_0 = 0$, it follows that there are no delay-free loops on the K side, but it is not possible to adapt the W port. This is in fact equivalent to an ideal voltage source [Fig. 2(c)], and it should be connected to the only adapted port of the MA, thus preventing the connection of a NLE. However, we found that under particular circumstances it is possible to remove the delay-free loop from the W port even when this is connected to a generic port of the MA. By recalling the results of Section III-A, if we connect the K-W converter to the adapted port of a three-port series junction, then the remaining two ports of this will also be adapted, provided that we choose the same reference resistance for them. The new K-W converter obtained has now two W ports characterized by the same reference resistance, which is a free parameter. Although these two ports are adapted [their reflected wave only depends instantaneously on the other port's incident wave, see Fig. 2(d)], if we connect a non-adapted element to each of them we would generate an outer delay-free loop [Fig. 2(e)]. It follows that one of the two W ports must be connected to an adapted one. However, the presence of the series junction means that in order to be able to use this particular K-W-converter, the overall structure (K element and W structure) must be decomposable into the series connection of two W blocks and the K model. We can reach a similar result also for an admittance type of K model, but in this case the W port of the K-W converter would be an ideal current source and, consequently, the junction would be a parallel one.

In conclusion, the union of the impedance (admittance) K-W converter and the series (parallel) junction, can be seen as a special series (parallel) junction between a K model and two W models, where one of the two W ports can connect a non-adapted element. In Section V-E we will provide an example of this sort.

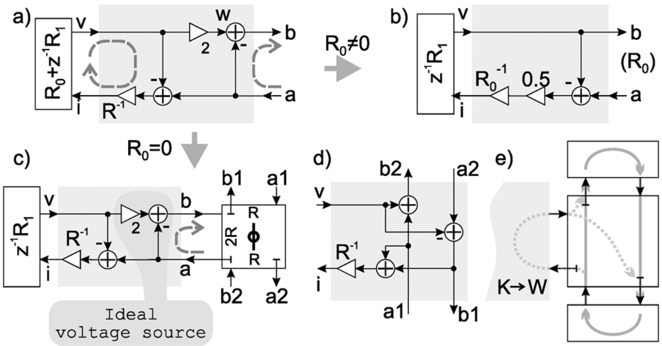


Fig. 2. K-W converter. (a) The connection of a generic impedance K model to a K-W-converter generates two delay-free loops. (b) Loops are eliminated if we choose the W port's reference resistance equal to the instantaneous term of $Z(z) = R_0 + z^{-1}R_1(z)$. (c) When $R_0 = 0$ the loop on the K port disappears, while the delay-free loop due to the ideal voltage source at the W port can be broken by connecting a suitable series junction as explained in Section III-A. (d) Resulting structure when $R_0 = 0$. (e) Schematic illustration of the fact that the junction of fig. (d) does not introduce any delay-free loops (solid lines denote instantaneous dependencies, dotted lines represent delayed dependencies), but an outer delay free loop can still be generated when both W elements connected are not adapted.

IV. Structures

As already anticipated in the introduction and in Section II-D, one critical issue that needs to be addressed when dealing with VA and, more generally, with physical models for musical acoustics, is the need of accommodating the interaction between different WD structures. In fact, the analog reference system is often inherently modeled as the interaction of several subsystems. Another reason why we need to model the interaction between connection trees is related to the fact that a WD

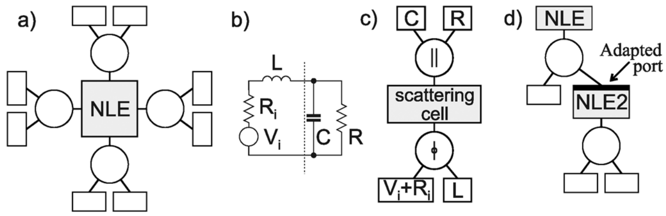


Fig. 3. Types of multi-port NLE. (a) Generic multi-port NLE. (b) Linear circuit implemented using a scattering cell that connects two trees. (d) A two-port NLE (labeled NLE2) with one of the two ports adapted, so that it can be connected as a leaf.

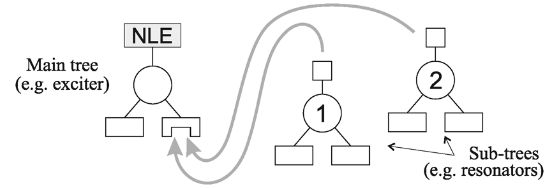


Fig. 4. Example of Root–Leaf connection, where two sub-trees can be alternatively connected to a special leaf of the main tree.

structure can only accommodate up to one nonlinear element. This limitation is inherited by WDFs and cannot be overcome directly without radically changing the philosophy of implementation of connection trees, or giving up their computability. In many situations of practical interest, what we can do instead is to overcome this limitation by *bridging* two or more trees together.

The inherent iterative nature of connection trees allows us to address the interaction between different WD structures in a quite straightforward fashion. In this section, we show that there are different ways for WD structures to interact with each other. Some are more suitable for model interaction purposes, others are more oriented to solving computability problems. Assuming that each WD structure is represented by a connection tree, one type of interaction between them is through a shared nonlinearity (root in common), which can be implemented as a multiport element. Another possibility is through a form of "structural" time variance that can be managed on the fly through properly defined operations of pruning, grafting, and bridging. Finally, interactions between WD structures can take place in the form of cross-dependencies between WD elements of different connection trees.

A. Multi-Port Nonlinearity

Sometimes the model to be implemented presents two or more NLEs and it is not possible to separate their update process (i.e., by interposing a delay cell) to guarantee its computability due to, for example, their physical proximity or contiguity. A quite straightforward solution in this cases is to implement all NLEs and their interconnections as a single multi-port element, each port of which must be connected to the root of a tree [see Fig. 3(a)]. The same solution can be applied when the reference model, or its WDF implementation, contains a nonlinear (or a non-adaptable) element which is already by definition a multi-port one. A notable example is the Kelly–Lochbaum scattering cell defined by (3), implemented as a BCT root that connects two sub-trees. Another example will be introduced in Section V-A. In the presence of a multi-port NLE, the update process of the whole BCT structure must be modified slightly, to ensure its correct update. In fact all the NLE's incident waves need to be computed first, in order to compute its reflected waves. Consequently, the update process must scan all the trees from the leaves to the NLE first, compute all the NLE's reflected waves and then scan the trees from the root to the leaves.

A special case of multi-port NLE is when one of its ports can be adapted. This port can in fact be connected as a leaf, leading to another connection topology between trees [Fig. 3(c)] although very similar in many aspects to the *cross-control*, introduced next. A very useful example of this type of element is the controlled source. In particular, the *voltage follower* which can be seen as a particular case of *Voltage Controlled Voltage Source* (VCVS), will be used often in Section V. This can be readily implemented through a very high resistance on the adapted port (although slightly more complex methods exist for accommodating an ideal generator instead), and an ideal source, controlled by the voltage across the resistor, on the other port.

B. Intra-Structure Connections

An important class of interactions require that parts of a circuit, representing a physical object or part of it, should be able to be connected to or disconnected from the main structure (Fig. 4). This is often the case in exciter-resonator type of interactions, where the exciter can interact with a number of different resonators. If the part of circuit to be (dis)connected is linear, then this operation corresponds to (dis)connecting a sub-tree, by breaking/making a parent-child connection. Such an abrupt change in the topology requires a re-initialization of the whole structure, in order to ensure consistency of voltages and currents along the new circuit(s) obtained. Unfortunately, although ensuring the physicality of the model, this process would unavoidably create discontinuities or "clicks" in the output audio. A way to prevent this situation and to avoid the re-initialization of the structure, is to carry out the (dis)connection only under a predefined proximity condition, corresponding to an instant before the actual contact between the two objects takes place [15]. The proximity condition ensures that the connection needs to be made (i.e., there is an high probability that the two objects will get in contact in the next few time samples), but there is no exchange of energy yet. A typical situation is when the sub-circuit is connected to the main one in parallel to a short-circuited port. The connection of the sub-circuit (proximity condition) will not change the behavior of the main one, until the short circuit will be gradually removed, e.g., by slowly changing the work point of a NLE (contact condition).

Finally, it may be worth noting that, depending on the specific implementation of the BCT, the disconnected sub-trees may not be able to be updated when not connected to a root element. This issue can easily be solved by providing a "dummy" root element that connects to floating sub-trees, ensuring their update.

C. Leaf–Leaf Connections

There are situations where two or more trees need to be connected without giving up the possibility to connect a NLE

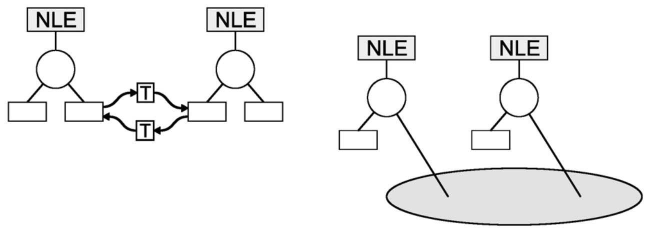


Fig. 5. Leaf–Leaf connection and a resonator with multiple excitation.

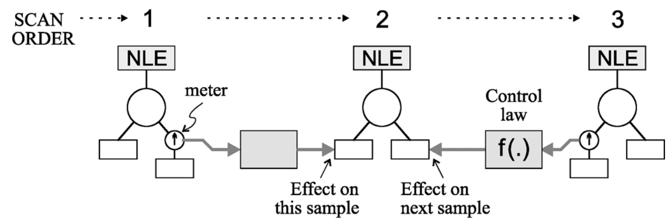


Fig. 6. Trees linked by Cross-Controls. The update order is also shown.

to each of them. By the BCT's connection rules recalled in Section II-C, nodes cannot be connected with each other through two non-adapted ports. In other words, direct connections of two (sub)trees nodes are not allowed, as this would lead to a non-computable structure. To avoid computability issues, each tree should behave in the same way as it was isolated, in particular by ensuring that all its leaves are adapted. A way to satisfy these conditions but still having connections between two or more trees, is by allowing leaves to be connected to each other, provided that the ports connected to a tree's node remain adapted. This is possible if these *leaf–leaf* connections contain at least one delay cell. Common cases where this type of connection is required are those that involve DWGs or, in fact, any distributed element, so that the delayed connection between leaves is not very restrictive for musical applications. Typical examples include a membrane connected to two or more mallets (Fig. 5).

D. Indirect Inter-Element or Inter-Structure Connections

Trees can also be connected at signal level, through a link that we refer to as *cross-control*. The idea is to use a generic physical quantity (usually a voltage) measured at a port of any element, to modify one or more parameters of another element. The controlled element does not necessarily need to belong to another tree, it can as well be connected to the same tree. In fact, it can even be the same element from which the control variable is being measured. Optionally, the control signal can be modified by a suitable nonlinear function (exponential, step, etc.) or simply a scaling factor, to ensure that the controlled element's parameters will not assume nonphysical values (e.g., a negative resistance), or to obtain more complex behaviors. An important difference with respect to the other connection topologies that we have seen so far is that the cross-control is a feed-forward type of link, since the controlling element is not affected by the controlled one.

Due to the nature of this type of interaction and to preserve computability, the effect of an element's modification will in general take place at the next sample, although it may take place during the same sample if the controlled element's (sub)tree is still to be computed (Fig. 6). This is clearly a limitation when two (sub)trees have mutually dependent elements. In particular this happens in feedback structures, where the input signal should be compensated by the output, which in turn depends on the input. This issue will be examined in more detail in Section V.

It should also be noted that in general the change of a parameter requires energy. This is the case, for example, when we want to change the capacitance of a charged capacitor (to fix ideas, ideally by changing the distance between its plates). The energy to be spent corresponds to the work needed to move the plates against the electric field, which is null only if the voltage across the capacitor's terminals is null ($\Delta E = (1/2)\Delta C V^2$). On the other hand, since there is only one variable (signal) read, the link between two elements does not carry any information about the energy involved in the process. Incidentally, changing the value of a capacitor in the W domain, corresponds to changing its port resistance. This is equivalent to varying the capacitance under a constant applied voltage ($v = (a+b)/2$, is not affected), while the current ($i = (a-b)/2R$) changes with $1/R$, resulting in a variation of the total energy of the system. For instance, in a simple L-C oscillator we would notice a change in the amplitude of the output sine wave. Unless this is the effect we want to achieve, care should be taken, e.g., by forcing the correct values for voltage and current and then running a structure initialization to fix all the other elements' variables correspondingly. Another more practical solution is to use a nonlinear resistor to control the output amplitude (see Section V-C), dissipating (or providing) the energy in excess. From what has been said above, cross controls can affect the stability of the whole structure and should be used with care.

V. Circuit Synthesis for Virtual Analog Applications

In this section, we introduce some examples of VA structures developed using the blocks and the interconnection principles set forth in the previous two sections. The examples are simple enough to illustrate the concepts seen so far, yet their implementation is, in fact, of quite a significant musical interest.

A. Simple Class-A Amplifier

Fig. 7(a) represents a classic single-transistor class-A amplifier. The circuit shown in this example is based on a *Bipolar Junction Transistor* (BJT), though the same topology is valid almost unchanged for other types of transistors and also for vacuum tubes (although in the latter case supply voltages and resistors' values would be sensibly different). In fact, here we are not interested in a specific model of a transistor or a vacuum tube, but only in the topology of the circuit and its BCT implementation. For details on how to implement a specific model see for example [18].

Starting from the schematics of Fig. 7(a), if we try to implement it as a WDF using only one-ports, three-port adaptors and a two-port small signal model for the transistor, we would end up with a non-computable loop [Fig. 7(b)]. The loop, identifiable in the same figure, is due to the fact that we need $V_{BE} = V_B - V_E$, to calculate I_C, but V_E depends on I_C. A possible solution is to implement the transistor and all the adaptors that form the

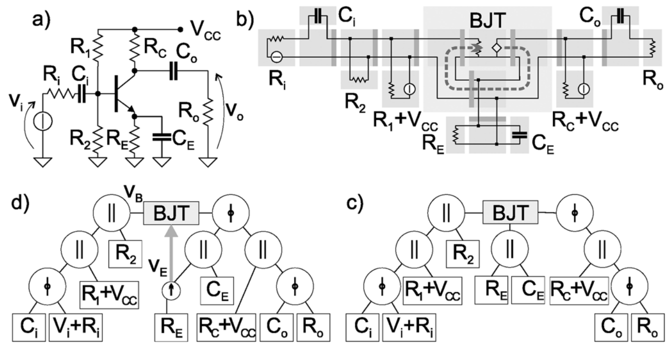


Fig. 7. Class A amplifier. (a) Electrical schematic. (b) Hybrid analog-WDF representation of the same circuit, using only one-port elements, three-port junctions, and a (over-simplified) small signal model for the transistor. (c) BCT implementation using a three-port NLE. (d) Alternative BCT implementation using a two-port NLE and a cross-control.

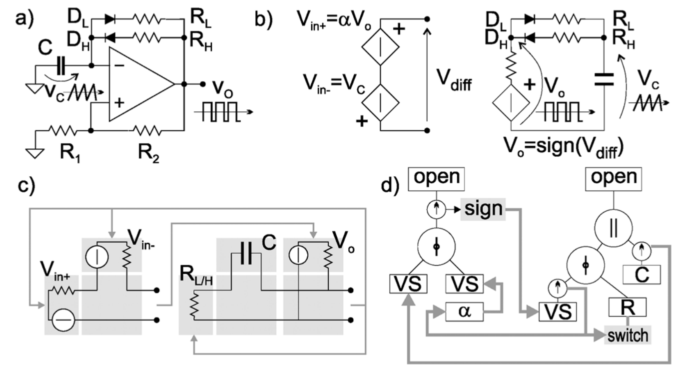


Fig. 8. Possible implementation of a square/sawtooth generator, using a comparator in configuration astable multivibrator.

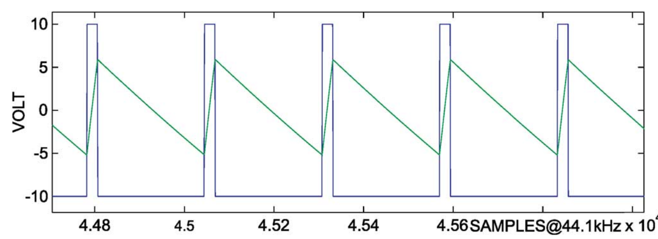


Fig. 9. Detail of a simulation for the astable multivibrator shown in Fig. 8. The values used were $V_{CC} = 10$ V; $R_o = 10^{-5}$ Ω; $R_H = 1$ kΩ; $R_L = 10$ kΩ; $C = 5$ μF; $\alpha = 0.1$. The plot represents V_o (square wave) and V_C (sawtooth wave), the latter amplified by a factor of 10 for clarity.

non-computable loop as a single three-port NLE, as highlighted in Fig. 7(b). By so doing, the non-computability would be dealt with inside the NLE, for example by using a fixed point search method. Also notice that a three-terminal element in the analog domain ended up being modeled as a three-port element in the W domain.

Another possible BCT implementation is the topology proposed in [18], which basically uses a cross-control. The voltage across R_E is fed as a parameter into the NLE, where V_{BE} is computed and used to obtain I_C (respectively marked as R_K, V_{GK} and I_P in [18]).

We have seen in Section IV-D that feedback circuits cannot be easily implemented as a BCT structure. There are a few special cases though, where a good approximation of the reference circuit can be obtained. In this example, the feedback is a very "mild" one, as in practice it only exists for very slow varying signals (below the band of interest for the signal), that is, when C_E can be considered an open circuit.

B. Square/Sawtooth Generator Using an Op-Amp Astable Multivibrator

Multivibrators are another example of feedback circuits that can be well approximated using a BCT structure. In this case the feedback is nonlinear, i.e., it is used to change the state of the system, which can assume only a finite set of values. A typical application of this class of circuits in musical acoustics is the *astable multivibrator*, used as a square or sawtooth waveform generator. Being able to correctly model the behavior of this implementation in a faithful fashion is quite important in VA applications, as its non-ideal behavior is perceived as musically interesting. Fig. 8(a) shows a classic analog implementation using an op-amp configured as a voltage comparator. Its output can only assume the two saturation values corresponding to the positive and negative power supply voltages $\pm V_{CC}$, depending on the sign of the difference between its inputs $V_{\text{diff}} = V_{\text{in}+} - V_{\text{in}-}$. In the astable multivibrator, the output is used to charge a capacitance C, through the resistance R_H (or R_L), and the voltage across C is fed back to the inverting input $V_{\text{in}-}$. The output is also used to set the "toggle threshold" voltage on the non-inverting input $V_{\text{in}+}$ to αV_{CC}, $\alpha < 1$, through the voltage divider composed by R_1 and R_2. Supposing for simplicity that initially the voltage across the capacitor is $V_C = 0$, this will increase (in magnitude) and tend asymptotically to either V_{CC} or $-V_{CC}$, until its modulus will be greater than αV_{CC} set on $V_{\text{in}+}$, which will cause a change in the sign of V_{diff} and, as a consequence, on V_0 and $V_{\text{in}+}$. This will reverse the current flowing into C and the trend of V_C, until the newly set threshold voltage will be surpassed in absolute value, and so on. It follows that V_0 changes its sign periodically, resulting in a square waveform. On the other hand, V_C is a composition of branches of negative exponential curves which, depending on the choice of α, C, R_H, and R_L, well approximate a straight line, so that V_C results in a triangular waveform (Fig. 9). The resistors R_H and R_L are selected by the current polarity of the output using the diodes D_H and D_L, in order to obtain different rise and fall timings for V_C, i.e., to obtain asymmetric waveforms.

Fig. 8(b) shows a simplified model of the circuit, more suitable for a W implementation. The input stage is composed by the series of two VCVSs, corresponding to the voltages at the two inputs referred to the 0 V reference (the "ground"). The output stage is also composed by a VCVS (in the schematic labeled as V_0), connected to the capacitor C through R_H or R_L. The model is completed by the links between the VCVSs and their controlling elements: $V_0 = V_{CC} sign(V_{\text{diff}})$, $V_{\text{in}+} = \alpha V_0$ and $V_{\text{in}-} = V_C$.

The W implementation can be easily obtained from the simplified model, by noticing the similarity between a VCVS and a "cross-controlled" voltage source. It follows that the input stage of the op-amp can be implemented by the series connection of two voltage sources while the output stage, implemented as a separate structure, is composed by another voltage source in series with a resistor and a capacitor. Notice that for the input

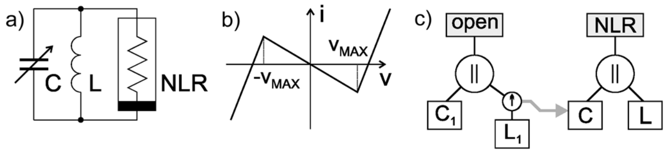


Fig. 10. Simple VCO with voltage regulation. (a) Electrical schematic. (b) Example of NLR characteristic. (c) BCT implementation, using a *low frequency oscillator* (L_1, C_1) to vary the main oscillator's capacitance.

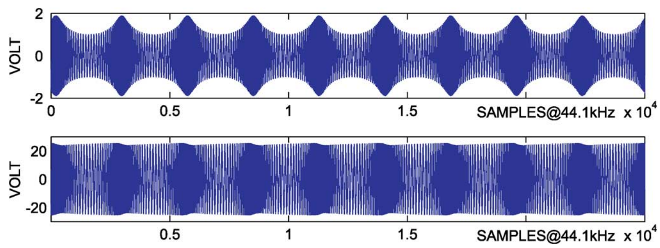


Fig. 11. Simulation of the BCT structure in Fig. 10(c). Top: without the NLR; bottom: with a NLR having slope -1 between $[-10\text{ V}, 10\text{ V}]$, and slope 1 elsewhere, while the other values used were $C = 3.6 \cdot 10^{-4}$ F; $L = 3.6 \cdot 10^{-4}$ H; $C_1 = 3.6 \cdot 10^{-2}$ F; $L_1 = 3.6 \cdot 10^{-2}$ H; $V_C(0) = 1$ V; $V_{C1}(0) = 1$ V. The Cross-control was mapped so that the entire excursion of the controlling signal $(-1\text{ V} \ldots 1\text{ V})$ modifies C from 0.1 C to 1.3 C.

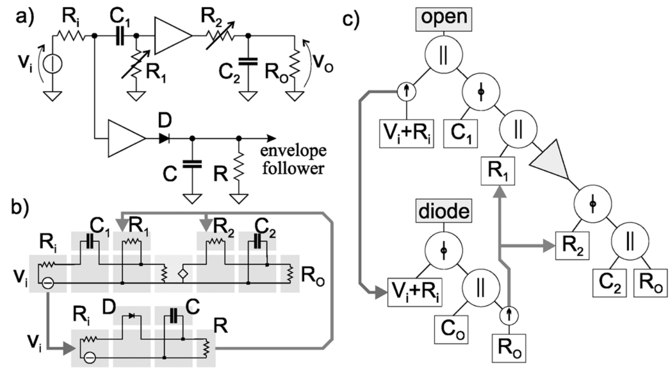


Fig. 12. Second-order filter composed by two decoupled first-order cells, controlled by an envelope follower. (a) Principle schematic. (b) The same circuit with highlighted series/parallel connections. (c) BCT implementation, where the two-port NLE that models the voltage follower has been represented with a triangular shape.

stage we can use real voltage sources, as the presence of the internal resistor would not cause any voltage drop since their output is left open and is only used to control the output voltage source. This choice greatly simplifies the implementation of the input stage without introducing any approximation. On the other hand, the output stage should be implemented as an ideal voltage source, as the internal resistance of a real source would affect the time constants CR_H and CR_L. For the sake of simplicity, in this example we used a real source whose internal resistance is much smaller than R_H and R_L. It is indeed possible to implement V_0 as an ideal voltage source as explained in Section III-A or simply by connecting it as the root of a tree. Finally, four cross-controls must be set: from V_C to $V_{\text{in}-}$ (one to one); from V_0 to $V_{\text{in}+}$ (through a coefficient $0 < \alpha < 1$); from V_0 to R (through a step function that chooses two suitable values for R, depending on the sign of V_0); and from V_{diff} to V_0 (through a sign function).

It is important to point out that, despite its lengthy description, the proposed implementation of the astable multivibrator is very simple and efficient as it only requires three adaptors and five one-port elements.

C. Voltage-Controlled Oscillator With Automatic Gain Control

A sine wave oscillator can be easily implemented by an LC circuit, taking for example the voltage as the output. To transform it into a voltage controlled oscillator (VCO) we can vary its capacitance through a cross-control, which uses a voltage (or, in fact, any electrical quantity) taken from a controlling circuit [Fig. 10(c)]. Since the oscillating frequency of an *LC* circuit doubles when the capacitance becomes a quarter of its original value, the control signal must be filtered by a suitable function if a linear voltage–frequency dependence is needed.

As mentioned in Section IV-D, a modification of the value of a reactive element will in general change the total energy of the system. This reflects on the output amplitude which will vary in an unpredictable fashion. A simple way to regulate the voltage in this type of applications is to use a nonlinear resistor (NLR) [Fig. 10(a)] having a *I-V* curve characterized by a negative slope around the origin and positive slope elsewhere, as the piecewise linear function shown in Fig. 10(b). The addition of the NLR to an *L-C* circuit forms the classical nonlinear oscillator [20] also used in practical applications, where the negative part of the NLR's characteristics provides the energy dissipated by non-ideal capacitors and inductors. Fig. 11 shows how the output changes when such NLR is connected to the circuit.

D. Voltage-Controlled Filter With Decoupling Stage

Many VA applications, from synthesis algorithms, to "graphic equalizers," to guitar effects, require a time-varying filter. In this example, we want to design a simple time-varying bandpass filter composed by the cascade of high-pass and low-pass RC stages. This configuration lets us set the high-pass and low-pass cut frequencies independently, as long as we provide adequate *decoupling* between the two stages so that the corresponding time constants due to C_1 and C_2 will be independent from each other. Other features of this configuration are also the absence of resonances and a glitchless response to sudden variations of the cut frequencies, provided that these are carried out by varying only the resistors' values, as shown in Fig. 12(a).

We know that in analog circuits an impedance decoupling can be obtained by interposing a voltage follower configuration between the two stages to be isolated. We have seen that a BCT counterpart of these decouplers can be easily obtained using a cross-control applied to an (ideal) voltage source, or very well approximated by designing a special two-port NLE composed by an adapted resistor with a very high value on the first port and an ideal voltage source on the second port, the latter directly controlled by the voltage across the first port. For this example we will choose the second solution.

To control the filter we could link both R_1 and R_2 to the output of a low-frequency oscillator, easily implemented with an *LC* circuit. Another interesting control source can be obtained from the output of an *envelope follower* fed by the same

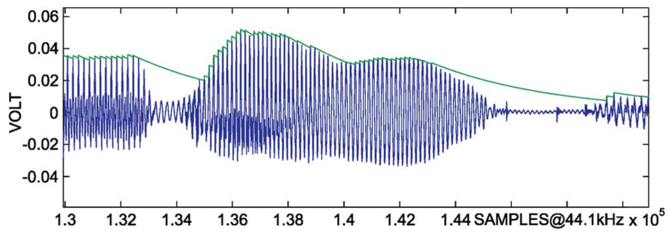


Fig. 13. Detail of input and output waveforms of the envelope follower shown in Fig. 12(a), for a musical input signal. An ideal diode was used, while $R_i = 10\ \Omega$; $C = 10\ \mu$F; $R = 10\ \mathrm{k}\Omega$.

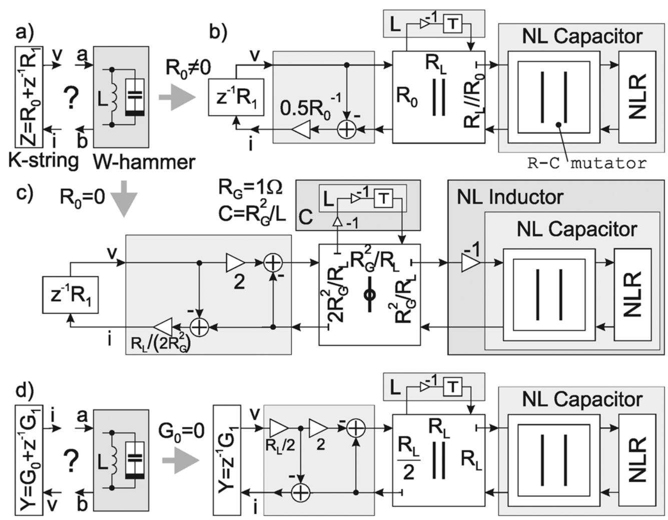


Fig. 14. W hammer connected to a K resonator through a K-W converter. (a) The two elements to be connected. (b) Resulting structure when the resonator contains a direct feed-forward block (R_0). (c) case when the resonator does not contain any direct in/out connection. (d) K-W-converter for an admittance model (only the case $G_0 = 0$ is shown).

input as the filter's, to implement the classic "auto-wah" effect. Fig. 12(a) also shows a simple example of envelope follower, based on a diode detector, of which a simulation is reported in Fig. 13. To prevent any distortion of the input signal caused by the presence of the diode, we interposed another decoupling stage before the envelope follower, this time implemented using a cross-control.

E. Hammer–Resonator Interaction Using a K Resonator

A hammer having nonlinear compliance and its contact condition to a generic resonator block can be quite easily implemented in the W domain [19]. This is composed by a NL capacitor [7] that models the felt and the contact condition, in parallel with an inductor to model the hammer mass. On the other hand, it would be useful to be able to connect a resonator (e.g., a string) modeled in the K domain, possibly through a K-W converter. In this example we suppose that the K model is an impedance one. We have seen in Section III-C that the K-W conversion can be done when the K-model contains a direct input–output term R_0, as long as we set $R = R_0$ as the reference resistance of the K-W-converter's W port. In order to do so, we need to have access to R_0, which will also need to be removed from the K model [Fig. 14(b)].

When the direct input-output connection is not available, the W port of the K-W converter cannot be adapted (it is an ideal source). It should thus be connected to the MA's adapted port, but in our example this is already taken by the NLE. From what said in Section III-C we can still obtain a computable structure if the K-W converter is connected to a series adaptor or, in other words, if the WD structure as seen from the K-W converter's W port can be split into the series connection of two blocks. Unfortunately this is not the case for the W-hammer of this example, because its only two elements are connected in parallel. However, by invoking the duality principle (see for example [20]) we can still transform the parallel of the inductor and the NL capacitor into the series of their dual counterparts, respectively a capacitor and a NL inductor. This can be done quite easily by using a dualizer, as shown in Fig. 14(c). Incidentally, this is the electrical equivalent that we would obtain using Firestone's (or force-current) electrical-mechanical equivalence [21] instead of the force-voltage one that we have used so far. Finally it is useful to note that if we had an admittance type of K model, since its corresponding K-W-converter requires a parallel junction we would be able to connect it to the hammer model right away [Fig. 14(d)], provided that we chose the NL capacitor's reference resistance equal to that of the inductor (see Section III-C).

VI. Conclusion and Perspectives

In this paper, we discussed and revisited existing methodologies for the modeling of nonlinear WD structures in view of their application to VA modeling. We emphasized the strengths but also the weaknesses of such solutions within this special category of physical modeling, and proposed novel solutions for making the WD approach more suitable for VA. In particular, we proposed novel transformation blocks and various interconnection solutions that are commonly encountered in the area of VA. We finally proposed several circuit modeling solutions for VA applications that serve as proof of concept as well as a testbed for the effectiveness of the developed solutions.

The method proposed in this manuscript extends the range of application of WD modeling for VA, and preserves the advantages that WD modeling (particularly BCT modeling) offers over K-based solutions, which are in terms of flexibility, support for time-varying parameters and topologies, computational efficiency, and responsivity. One more point of strength of WD modeling is in the availability of a solid implementation methodology based on connection trees, which allows the user to develop models without having to worry about algorithmic issues and code generation.

These techniques are meant not just to serve as a fast-prototyping circuit simulation solution, but also to address issues of flexibility and playability for VA applications in musical sound generation and processing. In fact, what makes VA interesting is the possibility to overcome the physical limits of the analog reference circuit, and open the way to a less constrained exploration of the resulting timbral space. We can achieve this through time-varying topologies, special cross-controls, and a more general definition of NLEs that enables additional parametrization. In order to exploit all such added features, it would be desirable to develop a framework that accommodates multiple connection trees with all the topological operations at hand (pruning, grafting, etc.), equipped with an intuitive and effective graphical user interface. A prototype of this interface was developed in [15] and successfully used for applications of musical acoustics [17]. Its generalization to multiple BCTs will be the next step.

REFERENCES

[1] A. Fettweis, "Wave digital filters: Theory and practice," *Proc. IEEE*, vol. 74, no. 2, pp. 270–327, Feb. 1986.

[2] A. Fettweis and K. Meerkötter, "On adaptors for wave digital filters," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-23, no. 12, pp. 516–525, Dec. 1975.

[3] A. Sarti and G. De Sanctis, "Systematic methods for the implementation of nonlinear wave-digital structures," *IEEE Trans. Circuits Syst.*, vol. 56, no. 2, pp. 460–472, Feb. 2009.

[4] K. Meerkötter and R. Scholtz, "Digital simulation of nonlinear circuits by wave digital filter principles," in *Proc. IEEE Int. Symp. Circuits Syst.*, Portland, OR, May 8–11, 1989, vol. 1, pp. 720–723.

[5] T. Felderhoff, "Simulation of nonlinear circuits with period doubling and chaotic behavior by wave digital principles," *Trans. Circuits Syst. I*, vol. 41, no. 7, pp. 485–491, Jul. 1994.

[6] A. Sarti and G. De Poli, "Generalized adaptors with memory for nonlinear wave digital structures," in *Proc. EUSIPCO, 8th Euro Signal Process. Conf.*, Trieste, Italy, Sep. 1996, vol. 3, pp. 1941–1944.

[7] A. Sarti and G. De Poli, "Toward nonlinear wave digital filters," *IEEE Trans. Signal Process.*, vol. 47, no. 6, pp. 1654–1668, Jun. 1999.

[8] A. Sarti and G. De Sanctis, "Memory extraction from dynamic scattering junctions in wave digital structures," *IEEE Signal Process. Lett.*, vol. 13, no. 12, pp. 729–732, Dec. 2006.

[9] J. O. Smith, "Principles of digital waveguide models of musical instruments," in *Applications of Digital Signal Processing to Audio and Acoustics*, M. Kahrs and K. Brandenburg, Eds. Norwell, MA: Kluwer, 1998.

[10] R. Rabenstein, S. Petrausch, A. Sarti, G. De Sanctis, C. Erkut, and M. Karjalainen, "Blocked-based physical modeling for digital sound synthesis," *IEEE Signal Process. Mag.*, vol. 24, no. 2, pp. 42–54, Mar. 2007.

[11] S. Bilbao, "Wave and scattering methods for the numerical integration of partial differential equations," Ph.D. dissertation, Stanford Univ., Stanford, CA, 2002.

[12] D. Franken and J. Ochs, "Generation of wave digital structures for networks containing multiport elements," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 3, pp. 586–596, Mar. 2005.

[13] D. Rocchesso and J. O. Smith, "Generalized digital waveguide networks," *IEEE Trans. Speech Audio Process.*, vol. 11, no. 3, pp. 242–254, May 2003.

[14] C. Erkut, "Modular interactions and hybrid models: A conceptual map for model-based sound synthesis," in *Proc. 13th Eur. Signal Process. Conf. (EUSIPCO-05)*, Antalya, Turkey, Sep. 2005.

[15] G. De Sanctis, A. Sarti, G. Scarparo, and S. Tubaro, "An integrated system for the automatic block-wise synthesis of sounds through physical modeling," in *Proc. 13th Eur. Signal Process. Conf. (EUSIPCO-05)*, Antalya, Turkey, Sep. 2005.

[16] S. Petrausch and R. Rabenstein, "Interconnection of state space structures and wave digital filters," *IEEE Trans. Circuits Syst. II*, vol. 52, no. 2, pp. 90–93, Feb. 2005.

[17] "IST Project of the Fifth-Framework Programme: ALMA, IST-2001-33059, ALgorithms for the Modeling of Acoustic Interactions," 2004. [Online]. Available: http://www-dsp.elet.polimi.it/alma/

[18] M. Karjalainen and J. Pakarinen, "Wave digital simulation of a vacuum tube amplifier," in *Proc. 31st IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP'06)*, Toulouse, France, May 14–19, 2006, vol. 5, pp. 153–156.

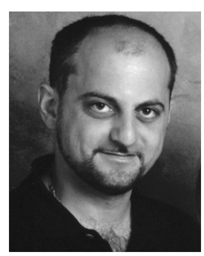
[19] F. Pedersini, A. Sarti, and S. Tubaro, "Block-wise physical model synthesis for musical acoustics," *IEE Electron. Lett.*, vol. 35, no. 17, pp. 1418–19, Aug. 1999.

[20] L. O. Chua, C. A. Desoer, and E. S. Kuh*, Linear and Nonlinear Circuits*. New York: McGraw-Hill, 1987, p. 839.

[21] F. A. Firestone, "A new analogy between mechanical and electrical systems," *J. Acoust. Soc. Amer.*, vol. 3, pp. 250–257, Jan. 1933.

Giovanni De Sanctis received the "Laurea" degree in electronic engineering from the Politecnico di Milano, Milano, Italy, in 2003. He is currently working toward the Ph.D. degree on parameters estimation for physical modeling in the Sonic Arts Research Centre, Queen's University Belfast, Belfast, U.K.

Prior to beginning his doctoral studies, he worked for three years as a Contract Researcher with the Image and Sound Processing Group, Politecnico di Milano, where he focused on sound synthesis through physical modeling and acoustic source localization for human–computer interaction. His research interests include signal processing, with an emphasis on sound–source modeling and tactile interfaces.



Augusto Sarti born in 1963. He received the "Laurea" degree (*cum laude*) and the Ph.D. degree in electrical engineering, both from the University of Padua, Italy, in 1988 and 1993, respectively, with research on nonlinear communication systems.

His graduate studies included a joint graduate program with the University of California at Berkeley, where he worked on nonlinear system theory. In 1993, he joined the Dipartimento di Elettronica e Informazione of the Politecnico di Milano, Milan, Italy, where he is currently an Associate Professor. His research interests are in the area of digital signal processing, with particular focus on sound analysis and processing, image analysis, and computer vision. He coauthored over 150 scientific publications on international journals and congresses as well as numerous patents in the image and sound processing area. He cofounded the Image and Sound Processing Group (ISPG) at the "Dipartimento di Elettronica e Informazione," "Politecnico di Milano," which has contributed to numerous national and international research Projects.