An interdisciplinary approach to audio effect classification

Vincent Verfaille⁹, Catherine Guastavino⁹
Caroline Traube⁹,⁹

⁹SPCL / CIRMMT, McGill University
⁹GSLIS / CIRMMT, McGill University
⁹LIAM / OICM, Université de Montréal

Montréal, Qc, Canada

♩ Sept. 18, 2006
Motivation

- **audio effects**: tools used by composers, performers, sound engineers to modify sounds
  
  $\implies$ "effect" = technique (cause) vs. "effect" on perception

  [Verfaille et al., IEEE-TASLP, 2006]

- generally classified on the basis of **underlying techniques**
  
  $\iff$ musicians rely on **perceptual** attributes

  $\implies$ gap between techniques & perception

  $\implies$ poor communication between researchers and artists
An interdisciplinary approach

– Goal: to link various types of classifications based on:
  – underlying techniques
  – type of control
  – perceptual attributes

– intersection between:
  – digital signal processing
  – acoustics
  – auditory perception and cognition
  – psycholinguistics
Existing discipline-specific classifications

Based on:

1. underlying techniques
   
   [Moore, 1990; Orfanidis, 1996; DePoli et al., 1996; Roads, 1996; Zoelzer, 2002]
   
   1.1 analog technologies
   1.2 implementation techniques
   1.3 domain of application / processing type
   1.4 operations applied to a model

2. type of control

3. perceptual attributes
1. Classifications based on underlying techniques

1.1 Analog technologies

- mechanics/acoustics
  e.g. musical instruments, effects due to room acoustics

- electromechanics
  e.g. vinyls: pitch-shifting by changing rotation speed

- electromagnetics
  e.g. magnetic tapes: flanging

- electronics
  e.g. filters, vocoder, ring modulators
1. Classifications based on underlying techniques

1.2 Implementation techniques, from [Zoelzer, 2002]

- filters
- delays
- modulators and demodulators
- nonlinear processing
- spatial effects
- time-segment processing, e.g. SOLA, PSOLA
- time-frequency processing, e.g. phase vocoder
- source-filter processing, e.g. LPC
- spectral processing, e.g. $\sum \sin + \text{noise}$
- time and frequency warping
1. Classifications based on underlying techniques

1.3 Domain of application and processing type

- **time domain:**
  - block processing  \(\text{(e.g. OLA, SOLA, PSOLA)}\)
  - sample processing  \(\text{(e.g. delay line, nonlinear processing)}\)

- **frequency domain** (block processing):
  - frequency domain synthesis (IFFT)  \(\text{(e.g. phase vocoder)}\)
  - time domain synthesis  \(\text{(oscillator bank)}\)

- **time and frequency domain**  \(\text{(e.g. phase vocoder + LPC)}\)

\[\Rightarrow\] choice depends on the artifacts
1. Classifications based on underlying techniques

1.4 Operations applied to a model

e.g. source-filter model based audio effects: [Verfaille & Depalle, DAFX-04]

- **basic operations:**
  - scale, shift, warp, multiply, interpolate

- **applied to** the filter, the source or both **components**
1. Classifications based on underlying techniques

Pros:
- see technical similarities of various effects
- better understand / implement multi-effects

Cons:
- audio effects may appear in more than one class
- steep learning curve for non-DSP experts
- non-intuitive for musicians
2. Classification based on the control type

from [Verfaille, 2003; Verfaille et al., JNMR 2006]

- constant

- variable, provided by:
  - wave generators:
    - periodic or low frequency oscillator (LFO)
  - other generators:
    - gestural control: realtime user-defined
    - automation: offline user-defined
    - adaptive: sound-defined
2. Classification based on the type of control

Pros:

– complements previous classifications
– appeals to developers, performers and composers
– defines a general framework to design new audio effects, e.g. adaptive audio effects [Verfaille et al., IEEE-TASLP, 2006]

Cons:

– useful mainly in a HCI & real-time context
– no link to implementation techniques / perception
Discipline-specific classifications

Quizz: what do you hear?

Sound examples from [Verfaille, 2003]

🎶 bell from Varèse’s *Poème Électronique*
freq.-dependent tremoli controlled by $C(f) = \sum_{\nu=0}^{f} S(t, \nu)$

⇒ tremolo? flanging? both?
implementation technique + control type

🎶 Sylvain Boeuf’s *Like Someone In Love*
adaptive time-scaling + synchronization points (both)
control type + sound feature ⇒ **performed** differently

We need to take perception into account
3. Classification based on perceptual attributes

Modified perceptual attribute(s) [Amatriain et al., JNMR, 2003]

- **pitch**: *e.g.* melody, intonation, harmony
- **loudness**: *e.g.* dynamics, *tremolo*
- **time**: *e.g.* duration, rhythm
- **space**: *e.g.* localization, room effect
- **timbre**: *e.g.* formants, brightness, texture
3. Classification based on perceptual attributes

Examples of effects modifying timbre: [Verfaille et al., JNMR, 2006]

<table>
<thead>
<tr>
<th>DAFx name</th>
<th>Perceptual Attr.</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Main</td>
<td>Other</td>
</tr>
<tr>
<td>chorus</td>
<td>T</td>
<td></td>
</tr>
<tr>
<td>equalizer</td>
<td>T</td>
<td>L</td>
</tr>
<tr>
<td>filter</td>
<td>T</td>
<td>L</td>
</tr>
<tr>
<td>flanger</td>
<td>T</td>
<td>P</td>
</tr>
<tr>
<td>spectrum shift</td>
<td>T</td>
<td>P</td>
</tr>
<tr>
<td>adaptive ring modulation</td>
<td>T</td>
<td>P</td>
</tr>
<tr>
<td>comb filter</td>
<td>T</td>
<td>L,P</td>
</tr>
<tr>
<td>resonant filter</td>
<td>T</td>
<td>L,P</td>
</tr>
<tr>
<td>wah-wah</td>
<td>T</td>
<td>L,P</td>
</tr>
</tbody>
</table>
3. Classification based on perceptual attributes

Pros:
- complements to previous classifications
- appeals to all listeners
- represents artifacts (e.g. time-scaling)

Cons:
- one effect can modify several attributes (control-dependent)
- difficult to find a graphical representation
Interdisciplinary audio effect classification

⇒ links discipline-specific classifications:
- semantic descriptors
- perceptual attributes
- control type
- operation / processing applied
- processing domains
- digital implementation techniques
Interdisciplinary audio effect classification

Chorus implementations:
- white noise controlling delay line(s) length modulation
- mixing pitch-shifted & time-scaled versions
Interdisciplinary audio effect classification

Adaptive time-scaling implementation:
- sound-defined control 🎵
- timbre → duration

Semantic Descriptors
Perceptual Attribute
Control Type
Applied Processing
Processing Domain
Digital Implementation Technique

Warm Sound
Several Performers
Duration
Adaptive
A-time-scaling Effect
Transposition
Time-Scaling
Resampling
Time-Frequency
Phase Vocoder
SOLA
Delay Line
Interdisciplinary audio effect classification

Pros:
- combines different standpoints
- links layers of discipline-specific features
- compact representation of audio effects

Cons:
- using a shoehorn to fit an elephant in a glass
- collaborative efforts
Conclusions

– review existing classifications

– introduce transverse classification:
  – from signal processing to semantics
  – best meet the need of a wider variety of users

– implications for
  – teaching and knowledge sharing
  – design of more intuitive user interfaces

– future directions:
  – correlate verbal descriptors and lower-level attributes
  – develop navigation tools (Wiki, trees)
  – retrieve information
An interdisciplinary approach to audio effect classification

Vincent Verfaille‡, Catherine Guastavino§
Caroline Traube¶,♯

‡SPCL / CIRMMT, McGill University
§GSLIS / CIRMMT, McGill University
¶LIAM / OICM, Université de Montréal

Montréal, Qc, Canada
dni Sept. 18, 2006