VISUAL REPRESENTATIONS FOR DIGITAL AUDIO EFFECTS AND THEIR CONTROL

ARFIB D.

C.N.R.S Laboratoire de Mécanique et d'Acoustique, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France
arfib@lma.cnrs-mrs.fr

ABSTRACT

This article gathers some reflections on how the graphical representations of sound and the graphical interfaces can help making and controlling digital audio effects.

Time-frequency visual representations can help understand digital audio effects but are also a tool in themselves for these sound transformations. The basic laws for analysis-transformations-resynthesis will be recalled.

Graphical interfaces help for the control of effects. The choice of a relation between user control parameters and the effective values for the effect show the importance of the mapping and the graphical design.

Graphical and sonic editor programs have both in common the use of plug-ins, which use the same kind of interface and deal with the same kind of approach.

1 some definitions : digital audio effects, their control, their implementation

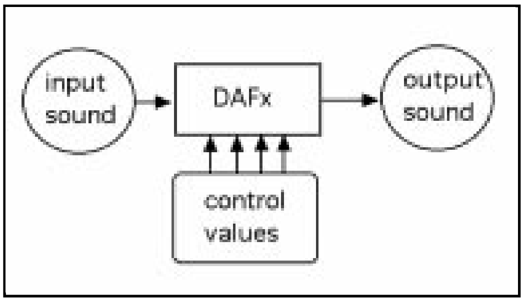


fig 1 : a digital audio effect and its control

a digital effect can be viewed as the transformation of one (or more) sound input via a process controlled by parameters values that can be fixed or variable. The output is one (or more) sounds. The definition of the value or the variation of the control parameters can be either at the algorithmic level, or at a higher level. In this case there is a mapping between the two. But in both cases the visual representation of these values or variation is a good tool for the musical use of digital audio effects.

The implementation of effects is usually in two parts : the process itself, and its control. So the implementation is often based on a graphical interface, eventually with different skins.

2 visual representations of sounds

The basis of these representations is to obtain images ($z=f(x,y)$) or curves $z=f(x)$ which are intermediates step between an input and an output. In this paragraph we deal with the process (and not the control) and see if visual representation can help designing effects.

2a an audio effect as a picture modification

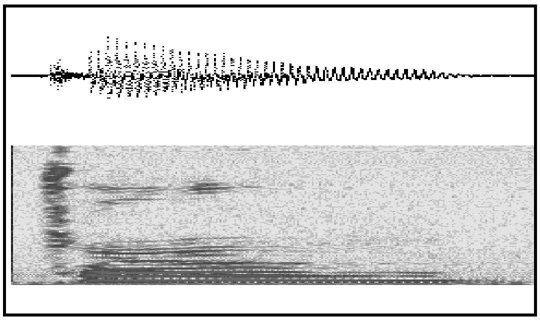


fig 2 : time and time-frequency representation

The sliding FFT is a well-known technique that can render such picture, but other exists like the wavelet transform. But the principle is the same.

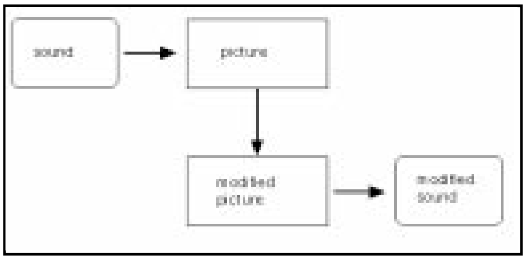


fig 3 : DAFx using a picture transformation

The operation on these representations provide a tool for the resynthesis of transformed sounds. Some programs (like Metasynth, but also experimental programs such as Phonogramme) use the modification of representation to produce sound transformations. The underlying assumption is that "what you see is the real transformation". It is not always so.

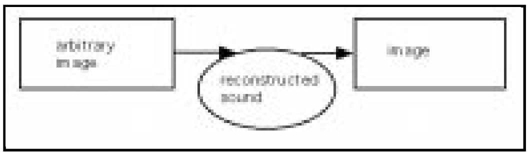


fig4 : invalidity of an arbitrary image

When an arbitrary image is given, and the synthesis of the sound forced by a specific method, usually an inverse Fourier transform, the time-frequency representation of this sound is not equal to the arbitrary image, and can even be very different.

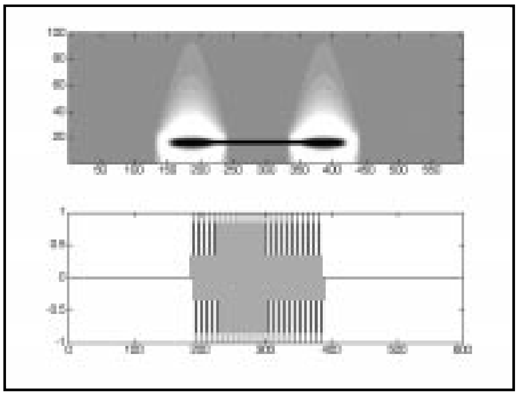


fig5 : the reconstruction of a spectral line

Modified images are always arbitrary. This in short explains why for example many time-stretching programs give undesired artefacts. One way to understand the is to consider only one point of the arbitrary image.

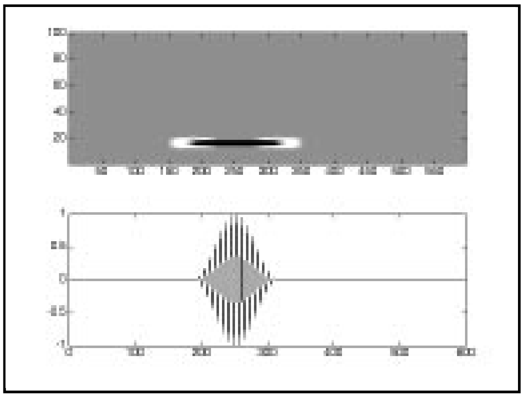


fig6 : the reconstruction of a spectral point

The sound associated to a point is a small "gaboret" (or "wavelet" in case the wavelet transform is used), the time frequency representation of it being a two dimensional zone around the co-ordinates of the initial point. This blurring zone is called the reproducing kernel associated to the transform. The reconstructed image is then the sum of these small kernels.

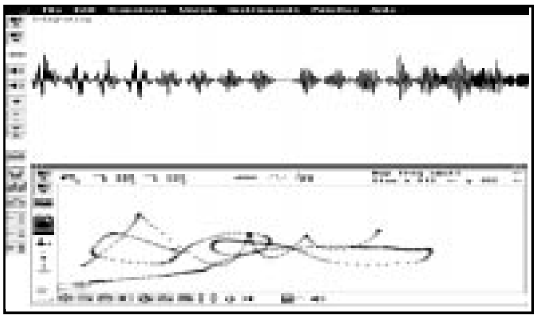


fig7 : metasynth synthesis from an image

From there we can see that one must be very careful when designing an effect based upon time-frequency representations, or even use these artefacts as the goal of the effect. This does not prevent to use digital audio effects based on the transformation, or even the synthesis of curious sounds.

2c sinusoidal noise and transients models

There are other ways to represent the sound as an intermediate step between the analysis of an input sound and the synthesis of a processed sound. One of them is the representation of the "sinusoidal" part of a sound. But to find out where the supposed lines are, one must use signal processing tricks to guess the spectral lines.

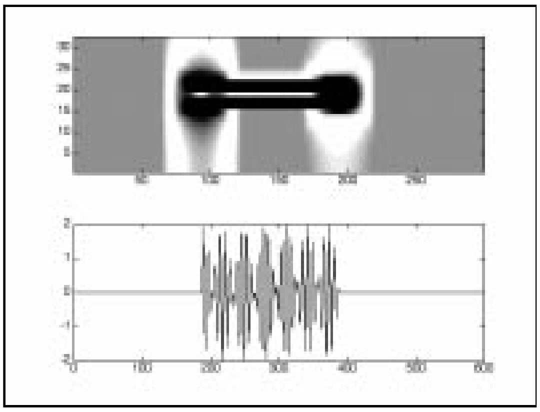


fig8 : time-frequency representation of 2 lines.

Noise is another problem, and here the question is how to distinguish a sinusoidal component from the noise?

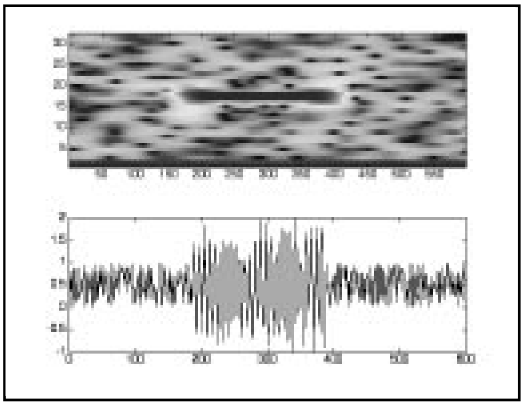


fig9 : *T/F representation of a noisy sine wave.*

Nevertheless, digital audio effects can realistically rely on these sinusoidal + transients + noise representations. The distinction between a transient part, a sinusoidal part and a noisy part is becoming standard for the definition of a sound, like the SDIF format that can be used by different programs .

2d curve extraction for adaptive effects

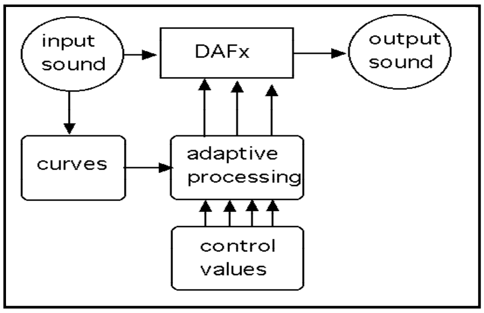


fig10 : curves extraction adds adaptive controls

Digital audio effects can use other representations, explicitly or implicitly to produce effects that use the sensitive content of a sound. In order not be misleading with what is called "content-based" information, one can call them "intelligent", adaptive" or "sensible" effects.

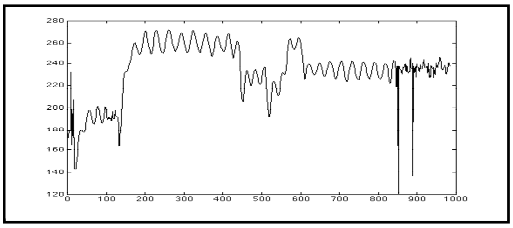


fig11 : pitch extraction for vibrato modification

This is the case of the effects using pitch detection. It is obvious in that case that this information is a very good track to make an alteration of a sound. Even "intelligent harmonising" can be done which totally rely on the accuracy of such an estimation.

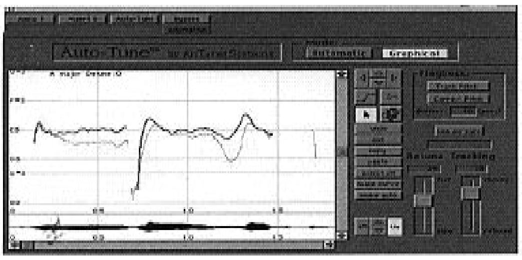


fig12 : Autotune effect

But other curves are really a must for other effects. For example the evolution of the centroïd of a sound can be very helpful for a modification of the expressivity of a sound. Even a voice/unvoiced curve can provide a very good indication for an adaptive treatment of the voice.
The images or curves are very descriptive of acoustics characteristics of sounds. At this step we are very near to the physical description of sounds, which doe not always correspond to its control. But from them we can derive perception clues that can be controlled in an easier manner.

3 graphical interfaces for effects control

Here we deal with the control parameters. The action devices are mouse, keyboard (numerical or Midi), Midi controllers, gesture sensors and so on. The feedback is a visual representation of this control and/or its consequence, and the sound by itself.

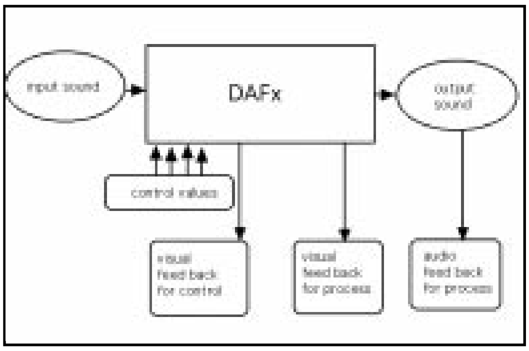


fig13 : different kind of feedback for control

3a action and visual feedback of the control.

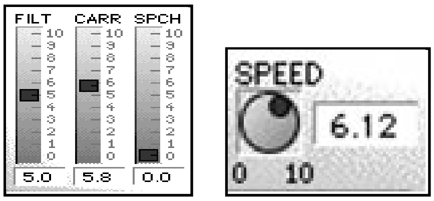


fig14 : linear and rotating sliders

A very simple case of feedback is the one of a mouse action on a slider. This slider is a 1D value, and it can be represented horizontally, vertically, on in a rotating position. The user must grab some part of the slider, but sometimes he/she can point on a non specific part and move the mouse (even in a rotating potentiometer). The control can also come from a sliding or rotating Midi slider

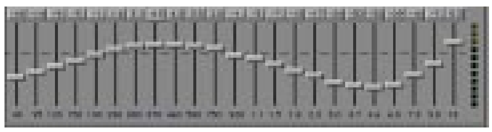


fig15 : A series of potentiometers of course is a good representation of the equalisation

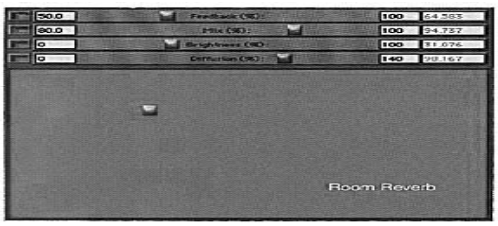


fig16 : 2D control represented on a rectangle

2D control is usually a rectangle on which a position can be specified via a mouse or a stylus on a tablet.

3b control and feedback of its effect.

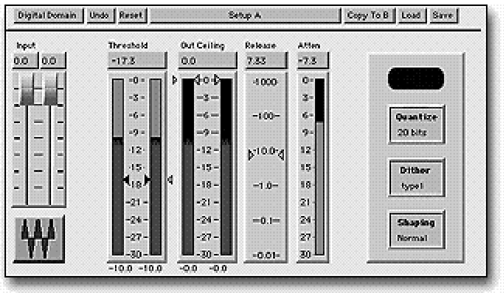


fig17 : feedback with vu meters

It is important that the feedback not only replicates the action but also takes in account the effect. As a basic example a potentiometer volume is better if there is also a vu meter that indicates the actual sound. One can find some good examples with GRM tools

fig 18 : visual feed back of the envelope of the sound with the (movable) limits for freezing

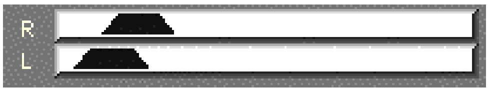


fig19 : Real time representation of the shuffling areas helps understand this effect (GRM).

Some designers take the aspect of a visual feedback of the transformation process in account.

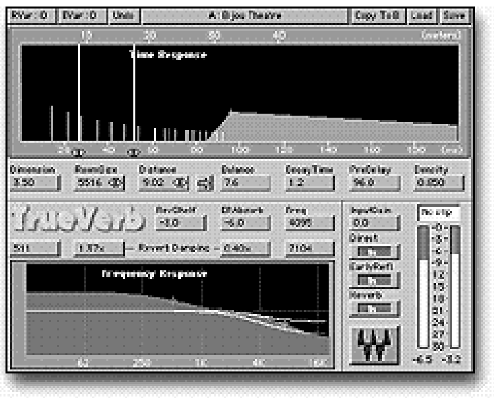
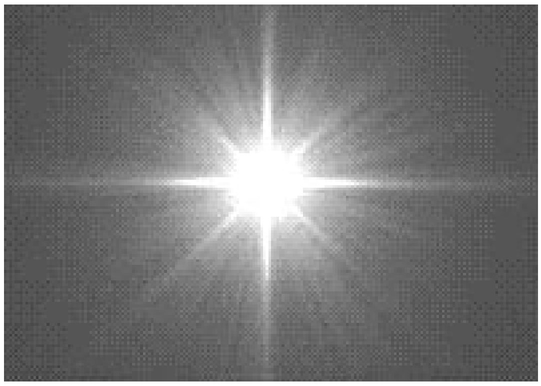


fig 20 : visual feedback of a reverb (Waves)

4 plug-ins for sound and/or vision

A final parallel between digital audio effects and vision effects is easy to make : plug-ins are add-on components for video and image processing as well as for audio processing. The interfaces that are used use the same operative systems, sot the same design, and one can find out many resemblance in the control parameters.



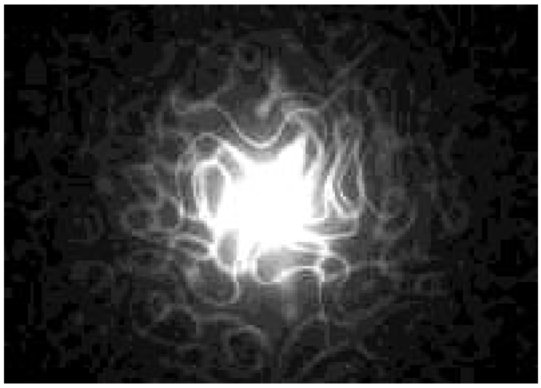


fig21 :visual effects on a source of light (Flare)

Moreover some image effects can give some ideas to audio plug-ins. As an example one can find image representation of reverberation, and at the inverse, one could think about the audio equivalence of a visual effect. Though this is not too conventional, many good ideas can come out from that reflection.

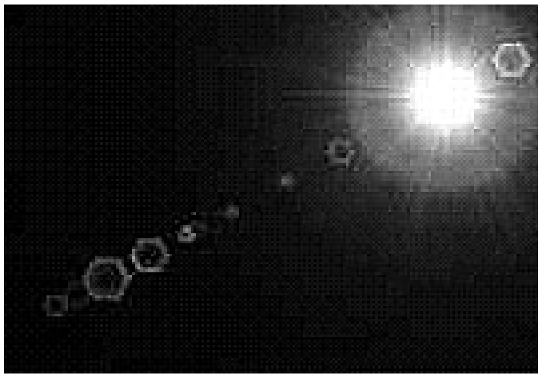


fig22 a last picture of a visual effect

conclusion

One hopes that the reader has now an intuitive catch on the possible relations between sounds and their representations and the problems and solutions that such representations bring out for digital audio effects. Some of the figures here presented have been taken from the web. Thanks to their authors.
