Pavel Ivanov

3 April 1998

An Idea of a Computer Program to Play Visual Music

A computer program that would allow performing musical text in a visual form should be based on a theory uncovering the intrinsic mechanisms common for both hearing and visual perception. One such theory relates the perception of musical pitch to the perception of an angle on the plane. This theory could be implemented as follows.

Visual tones

The internal representation of an elementary musical tone is described by the Gauss distribution on the pitch axis (the logarithm of sound frequency)

$$G_{\sigma,h}(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-h)^2}{2\sigma^2} \right),$$

where h is the centre of the tone and σ is the dispersion. For a direction on the plane, the same distribution could be chosen, with x being the angle counted from a definite reference direction (which actually does not matter since it is only the differences $x - h$ that enter the formula); however, one cannot visually distinguish the directions differing by the angle 2π radians (360°), and the distribution $G_{\sigma,h}(x)$ could be replaced by its circular counterpart

$$\Gamma_{\sigma,h}(x) = \sum_{k=-\infty}^{\infty} G_{\sigma,h}(x + 2\pi k) = \sum_{k=-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-h+2\pi k)^2}{2\sigma^2} \right),$$

with $x - h$ within [–π, +π].

For programming purposes, one could either use the exact values of $\Gamma_{\sigma,h}(x)$ calculated numerically, or introduce a simpler distribution possessing the same qualitative behaviour:

1. the function is symmetric around h;
2. the only maximum of the function occurs in $x = h$;
3. the only minimum of the function occurs in $x = h+\pi$;
4. the distribution becomes wider with increasing σ, so that one obtains a δ-function for σ = 0 (all the directions except $x = h$ have zero weight) and the uniform distribution for σ = ∞ (all the directions have the same weight).

If all the directions on the plane are represented with the same circular distribution Γ, one could consider a linear combination of the directions $h + 2\pi\cdot\log k$ with $k = 1,2,3,\ldots$ as an analogue of the overtone series for the direction h (binary logarithms will be used throughout this paper). However, one must ensure that the principal tone is much stronger than the rest of the partials.

Visual scales

Just like in music, any sensible composition must be based on certain *scales*, that is, the collections of tones forming a structural and functional integrity. One can introduce the same notation for directional scales as for musical pitch, taking, for definiteness, the direction from the left to the right for the analogue of the note *do*, and the downward direction for the note *la*. The well-tempered scale is then defined unambiguously:

0°	*c*	120°	*e*	240°	*gis*, *as*
30°	*cis*, *des*	150°	*f*	270°	*a*
60°	*d*	180°	*fis*, *ges*	300°	*ais*, *b*
90°	*dis*, *es*	210°	*g*	330°	*h*

One can define the same sub-structures in the 12-tone scale, such as diatonic modes, various kinds of pentatonic, and chords.

However, this is not the only possibility. Thus one can use natural diatonic rather than embedded in the 12-tone scale. The corresponding directions are:

0° 51.7° 102.2° 153.4° 210.2° 263.7° 310.4°

This structure can be rotated, depending on the variety of natural diatonic to represent. In the same way, there is natural pentatonic too:

0° 71.4° 144° 212.1° 288°

One could also consider another universal scale, dividing the visual “octave” (the angle of 360°) into 19 equal parts. This scale may include 12-tone hyper-modes, as well as the usual 7-tone and 5-tone modes, and triads.

Actually, every scale is a collection of *zones*, so that all the variations of angle (pitch) within the zone will be perceived as the variants of the same tone. One can calculate the zones for every scale, and the positions and widths of the zones will be different for each degree of the same scale.

Single-voice instruments

In this paper, a *voice* is a part of the composition played with a single instrument. An *instrument* is modelled with a point on the screen that can emit coloured waves of a standard shape in different directions. An instrument is characterised with its spectral parameters, dynamic parameters and phase parameters.

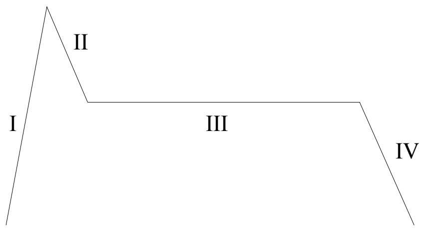
Spectral parameters

1. *Texture*. This is what visually distinguishes one voice from another. An analogue of the instrumental timbre in music. In the simplest case, colour is constant for the instrument; more complex instruments may have colours dependent on the direction of wave emission (in analogy with the registers of a musical instrument) or its brightness (volume). For instance, one dimension of colour (e.g. hue) might be used for visual timbre, some other dimension (e.g. saturation) indicating the visual octave (a kind of register specification). For an elementary tone, the intensity if the texture is to smoothly decrease from the pitch direction to the periphery.

2. *Fuzziness*. The value of the dispersion σ is the parameters of the instrument. Small values of σ give instruments with definite pitch (voice instruments), while large σ give noise-like performance (e.g. for percussion).
3. *Overtones*. Every instrument may have its own overtone spectrum (the collection of amplitudes of the partials represented by the circular Gauss distributions with the same dispersion σ). This spectrum may depend on volume.
4. *Random noise*. There may be many ways of admixing noise to the colour wave emitted by the instrument (bow noise, pump noise, reverberation etc.). Normally, the intensity of noise will depend on the volume.

Dynamic parameters

1. *Volume*. The intensity of the texture (colour) will correspond to sound volume. The brighter is the wave, the "louder" it displays. An instrument may be characterised by its volume range.
2. *Profile*. The dynamic profile of every note could be parameterised in the standard way:



I **attack**

II **relaxation**

III **plateau**

IV **cut-off**

The changes in the amplitude are to be modelled by the corresponding changes of texture intensity. Stages I, II and IV are characterised by their slope and relative or absolute duration; the amplitude and duration of the plateau are the parameters of the note. Voice instruments will have rather short attack, compared to the duration of the plateau; percussion instruments can well be modelled by the profile with zero amplitude of the plateau or its zero duration. A piano will have the peak-to-plateau ratio much higher than a violin. One could also introduce slope in the plateau stage too.

3. *Manner of performance*. The peak-to-plateau ratio may also depend on the manner of performance (the strength of stroke). The duration of the cut-off may be different for legato and staccato.

Phase parameters

1. *Spatial distribution*. The positions of the nodes (instruments) on the screen are analogous to the positions of the performers on the stage (like the instruments in an orchestra). The only difference is that two-dimensional spatial distributions become possible. However, two- and three-dimensional music can easily be performed using electronic delay circuits; one can also consider three-dimensional distributions of light sources modelled by the perspective on the screen.
2. *Spatial correlation*. The performance will always correlate the activity of different nodes on the screen. The simplest case of such a correlation is the many-voice composition, where the parties of the instruments are carefully synchronised; in particular, one can set up various node patterns on the screen acting in unison. However, there may be other ways of correlating the nodes—still, in general, the other nodes will serve as triggers for single nodes or node combinations. A particular way of phase correlation is passing the tune from one instrument to another; on the plane, one could arrange this so that a node gets activated when the wave from another node achieves it—still, there is no need to correlate only adjacent nodes.

Like in music, phase effects become less important when one observes the picture from a distance.

Special effects

The parameters of the instrument may be modified during performance in a number of standard ways, which could model various instrumental effects (pedals, mute etc.).

The volume (intensity) and pitch (direction) of the note may vary during performance. This will give such effects as *crescendo/diminuendo*, *glissando*, amplitude and pitch *vibrato* etc.

The changes in the phase parameters will be perceived as *movement*. Generally, the instruments may move in any way, which is to be defined by the composer or performers. However, there may be some standard movements defined for the instrument. Thus, one could implement chord performance through cycling around a small plane figure.

The instrument may emit waves in any directions—or, alternatively, it may be tuned to a fixed scale (for instance, the 12-direction scale). However, even for an arbitrary-direction (arbitrary-pitch) instrument, it would be convenient to specify directions in terms of some scale, with possible deviations from the exact angles. For the fixed-scale instruments, there may be the zones specified for each possible direction, and the standard ways of adjusting intonation within the zone would then be introduced as performance effects.

Ensembles

The same node on the screen may be associated with two or more instruments performing independently. In particular, the parameters of the instruments may be the same, so that the node may emit several waves simultaneously, which may be an analogue of musical chords. So, a many-voice instrument will be represented by the superposition of single-voice instruments.

The way of superposing different instruments (or the waves emitted simultaneously with the same instrument) have to be specially decided upon. The straightforward solution is to sum the intensities, in analogy with incoherent sound summation. However, there may be phase-dependent mechanisms of superposition, similar to the usual wave interference. In the visual case, this would mean that the colours get summed with the weights depending on the distance from the corresponding instruments; this dependence could be of the sine type, with negative weights possible, and hence destructive interference.

Performance

Now, any musical score can be translated into visual dynamics. The transcription of traditional music will require certain decisions concerning the selection of the instruments and there placement on the screen. However, this is not very different from the re-casting a musical piece for another set of instruments.

One could also combine light and music in the same performance. It should be observed that the light parties do not merely duplicate those of the musical instruments, being reasonably synchronised with them.

For the pitch-like mechanisms of angle perception to work, the screen must be not too close to the observer, so that the whole picture could always be grasped with one glance. However, the observer must be close enough to appreciate any phase effects in the performance.

The solutions with the observer within the light composition are not described with the theory lying in the basis of the application described. One will have to study the laws of such “embracing” art separately. However, there are reasons to suppose that locally, in every single fixation of the sight, the laws of directional scaling would be still applicable, and their computer implementations would work.

The computer applications for playing visual music may be of two major types. First, one could make an application to perform a specially prepared script, similar to those used by various scoring programmes for computer music. The script could be entered either from a file, or interactively, as plain text or through a graphic interface. In the applications of the second type would allow live performance, using either the usual MIDI keyboards or the computer keyboard and the mouse. The usual MIDI channel switching could be used to control different instruments (nodes on the screen). The parameters of the instruments must be specified in the special patch files beforehand. Of course, live-performance applications must be capable of storing the sequences played as scripts, for future replay or editing.

It would be reasonable to make he application as compatible with the existing MIDI coding scheme as possible. In this case, the same MIDI file could be performed both in music and in light, or in any their combination, the re-casting being effectuated quite simply through using a different patch file.

---

http://unism.pjwb.net/arc
http://unism.pjwb.org/arc
http://unism.narod.ru/arc