

(This article was written originally in 1997 as part of a larger text to accompany a Ph.D. thesis at the [Music Department of the University of California, San Diego](#). It was published in "Monochord. De musica acta, studia et commentarii", Vol. XIX, ss. 43-50. Adam Marszelek Publications, Torun 1998. It has been reedited for this web site in the fall of 2001 .)

BASIC THEORY OF INTERMEDIA

COMPOSING WITH SOUNDS AND IMAGES

by
JAROSLAW KAPUSCINSKI

Premise

My artistic premise is to compose structurally integrated intermedia works, in which sounds and images are given equal importance and are developed either simultaneously or in constant awareness of each other. I am particularly interested in working with narratives that emerge BETWEEN aural and visual layers.

Sounds and images

One of the most natural ways to compose structurally integrated intermedia works is by using sounds and images that are expressively charged with at least a minimum amount of change and a clear sense of direction. Musical motives or visual gestures may serve better this purpose than static elements because their motion can serve to effectively bridge the gap between media. It is more efficient to link sounds and images through correspondence in occurring change rather than perceived equivalence of quality. If both elements are in motion they may have a "common fate" and consequently undergo integration. The clarity of direction or perceived simplicity of change help establish strong intermedia links which are the foundation for intricate constructs on higher formal levels.

The premise that images and sounds have equal importance may suggest that even as separate elements their expressive value must be comparable. This is not a necessary condition. A very expressive element in one dimension may be treated in such a way that a simpler element in the other dimension is given the same weight. The goal is not to use equally charged elements but to emancipate them to a point where they can be balanced according to a specific compositional need.

Intermodal Gestures

Sounds and images that are expressively charged and have clear direction can be linked to create an intermodal gesture. Such an object is not just a sum of aural and visual elements that constituted it but rather a perceptual intersection of the two. Through integration it gains a new dimension.

This resembles the interaction of two-dimensional images that in stereoscopic vision recreate three-dimensional space. The phenomenon is based on the combination of observed similarity and difference. The two-dimensional images present views from two points of view separated by a distance corresponding to the distance between human eyes. Each of the viewer's eyes is then shown the appropriate image and through comparison of similarities and differences the brain recreates the third dimension, *i. e.*, depth.

Analogous phenomena can be observed in an audiovisual interaction. The original idea for an intermodal gesture (or the first image in our analogy) that identifies the expressive space to be embodied may come from one of the senses alone. It may have a musical or a visual form. Then, an element from the other sense (or the second image) which contains both similar and different characteristics is linked (superimposed). The added dimension that appears in the process is difficult to define but its presence can be quite evident and compelling.

Linking

Linking is a crucial agent in intermodal composition. Its function is analogous to superimposition in stereoscopic vision. It causes the sounds and images to interact and create the added dimension. In the process two elements are unified within this new dimension.

The basic rule for linking sounds and images is temporal coincidence. The two participating elements do not have to start exactly together or be of equal duration but they need to happen in temporal proximity with some overlapping part which permits them to enter into a relationship of reference.

Two exceptions from the temporal coincidence rule need to be mentioned; in a relationship of cause and effect an action in one modality may precede reaction in the other modality, and in an imitation relationship one modality can be recognized to repeat what the other modality has shown before. In all other situations a link without some simultaneity is difficult to perceive.

Links can be formed, reinforced and given unique quality by similarities between what is expressed in both modalities. These useful correspondences fall in the following categories:

Internal correspondences:

1. Temporal correspondences (*e.g.*, parallels between rhythmic and/or metric accents, parallels in tempo changes)
2. Textural (vertical) correspondences (*e.g.*, matching media forces, homophonic/polyphonic characteristics, similarities in density)
3. Structural (horizontal) correspondences (*e.g.*, parallels of formal symmetry/asymmetry, formal resolution, formal complexity)
4. Qualitative correspondences (*e.g.*, parallels of change in sharpness, brightness, size, motion, shape)

External correspondences:

1. Correspondence through a common physical object (*e.g.*, an animal and its sound)
2. Correspondence through a common cultural archetype (*e.g.*, church and organs)
3. Correspondence through a common emotional/psychological state (*e.g.*, anger, sadness, pleasure)
4. Complicity in depicting the plot line (*e.g.*, a cardboard box falls with a sound of breaking glass revealing the unseen content of the box)

It maybe important to note that synesthesia has marginal importance for intermedia linking as it differs from person to person. Linking can sucessfully happen through correspondences created between sounds and images that are not conventionally or in any other way associated.

Even an unlikely collision of sound and image can cause both of them to be evaluated with equal attention. It may even combat the usual dominance of sight over hearing. Scientific research suggests that in average circumstances vision amounts to about 80% of the information that is channeled to our brain in an audiovisual message.

Linking relationships can be established on every formal level. A unifying correspondence can be created between:

1. Single note and single object/motion (*e.g.*: appearance or disappearance, simple interaction, accent in motion etc.)
2. Motive and gesture
3. Phrase and sequence of gestures
4. All other levels up to the highest formal level of the work as a whole.

It maybe advantageous to reinforce the sense of intermodal unity by establishing links on several formal levels simultaneously. These can be changed in time. Some fragments may be linked on multiple levels and some may be relatively free of intermodal references. This way intermodal linking becomes a major contributor to formal development.

Perceptual strata

Foreground and background objects can be found both in musical and visual dimensions of intermodal works. The global hierarchy of strata is not however a simple addition of both. The important factor in the perception of strata is intermodal linking. Intermodal gestures, *i.e.*, objects that are linked dominate over those which are not. Consequently the general structure of perception consists of the following strata shown in order of predominance:

1. Linked foreground objects
2. Unlinked foreground objects
3. Linked background objects
4. Unlinked background objects.

Adding intermodal linking to such traditional foreground-creating tools as change or motion allows to build new strata within existing ones. It also permits to bring to the foreground elements or qualities which otherwise would have required distortion or muting of others.

Counterpoint

The expression of motion of an aural or visual object depends on the motion of other objects in the field of perception. Counterpoint can be defined as the relationship of independent motion of a minimum of two objects. If the rules of linking are observed and sounds and images are referring to each other, because they belong to different modalities they can always be perceived independently, therefore they naturally create counterpoint.

Most of the theoretical knowledge about counterpoint comes from polyphonic music. According to its theory there are three basic categories of contrapuntal motion of melodic lines *a* and *b* :

a is in motion while *b* is static (side motion)

a and *b* are in motion moving in parallel directions (parallel motion)

a and *b* are in motion moving in opposite directions (contrary motion)

In intermodal composition these categories apply to cases where objects in both senses change within one of the common attributes (*e.g.* brightness, size, tempo, duration etc.). When a linked change in both senses happens in different attributes it is a counterpoint of change rather than motion. The above categories, based on direction of motion, cannot apply and new ones must be created.

Interdependent Development

To create works with structurally integrated sounds and images their development must be worked out in the constant presence of both layers. From the lowest to the highest formal levels the expressive value of one is dependent on the other. Unless the composition is done by a single artist both the visual artist and composer would need to work together throughout the whole development process.

Conclusions

The unique property of structurally integrated intermedia composition is the creative focus on the content emerging between the senses. As explained earlier, this intermodal dimension is accomplished by linking, which means that while the freedom and individuality are the base for each dimension, it is the relationships between dimensions that constitute the medium.

The intermodal experience can be highly immersive. The multisensory content can make it extremely rich and intense. The intermodal gestures act as very strong stimuli. They can have a very direct, almost physical impact.

Throughout history, sounds and images were combined in various art forms. The studies that informed this article dealt with film, dance, theater, opera, video, installation art and interactive multimedia. Each of these forms offer a different approach, which additionally varies from artist to artist or epoch to epoch. Structurally integrated intermedia composition is not a genre in itself. It is a particular way of combining sounds and images that can be found in many genres. With the invention of computer technology and the possibility of putting sounds and images on a single canvas, it can be refined and developed with new flexibility. Structurally integrated intermedia composition as an aesthetic paradigm maybe one of the idiomatic artistic approaches to digital technology.