mathematics

MDPI

Article

Color and Timbre Gestures: An Approach with Bicategories and Bigroupoids

Maria Mannone [1,2,*], Giovanni Santini [3], Esther Adedoyin [4] and Carmine E. Cella [5]

1 Department of Engineering, University of Palermo, 90128 Palermo, Italy
2 European Centre for Living Technology (ECLT), Dipartimento di Scienze Ambientali, Informatica e Statistica (DAIS), Ca' Foscari University of Venice, Via Torino, 107, 30172 Venice, Italy
3 Music Department, Faculty of Arts, Ho Sin Hong Campus, Kowloon Tong, Hong Kong; info@giovannisantini.com
4 First Irish Business School, Egbeda-Idimu, Lagos 100276, Nigeria; adedoyinesther42@gmail.com
5 Center for New Music and Audio Technologies (CNMAT), University of California, 1750 Arch Street, Berkeley, CA 94709, USA; carmine.cella@berkeley.edu
* Correspondence: maria.mannone@unive.it or mariacaterina.mannone@unipa.it

Abstract: White light can be decomposed into different colors, and a complex sound wave can be decomposed into its partials. While the physics behind transverse and longitudinal waves is quite different and several theories have been developed to investigate the complexity of colors and timbres, we can try to model their structural similarities through the language of categories. Then, we consider color mixing and color transition in painting, comparing them with timbre superposition and timbre morphing in orchestration and computer music in light of bicategories and bigroupoids. Colors and timbres can be a probe to investigate some relevant aspects of visual and auditory perception jointly with their connections. Thus, the use of categories proposed here aims to investigate color/timbre perception, influencing the computer science developments in this area.

Keywords: cross-modal correspondences; equivalence classes; category theory; bicategories; complexity

MSC: 18D05; 20L05; 58H99; 76-XX; 78-XX

check for updates

Citation: Mannone, M.; Santini, G.; Adedoyin, E.; Cella, C.E. Color and Timbre Gestures: An Approach with Bicategories and Bigroupoids. *Mathematics* **2022**, *10*, 663. https://doi.org/10.3390/math10040663

Academic Editors: Luca Andrea Ludovico and Paolo Mercorelli

Received: 17 January 2022
Accepted: 16 February 2022
Published: 20 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The complexity of colors, the complexity of sound timbres, and the mystery of their possible interaction have been fascinating physicists, mathematicians, artists, and poets for centuries [1–7]. Visible light is a small portion of the electromagnetic spectrum; each color corresponds to a different frequency. Light is a transverse wave, while sound waves are longitudinal waves. A complex sound is constituted by the superposition of multiple frequencies, the harmonics or partials. White light can be decomposed into its colored components through a prism, as proved by Newton. A complex sound can be decomposed into its components through the Helmholtz resonators or computationally through the Fourier analysis [8]. Interestingly, a compact device with the function of an acoustic prism has been built [9]. Despite evident physical differences, colors and timbres have something in common: the concept of superposition of several components, the existence of decomposition tools, but also the possibility to create shadings and mixing of colors/timbres. In addition, we can think of some 'perceptive' analogies between timbres and colors: there are bands of colors which stimulate the attention and the feeling of alert more than others (e.g., red/violet more than light gray or blue), as well as some timbre bands with respect to other ones (e.g., trumpets, trombones and percussion with a *ff* dynamic, rather than a flute playing with a *pp* dynamic or a diapason). These analogies are not related with cultural associations between colors/sound and semantic meaning [10].

In this article, we try to sketch these analogies using the language of category theory, because of its power of abstraction. Category theory is an abstract branch of mathematics [11–13], that nowadays counts both theoretical and applied developments, including the arts [14–19].

1.1. Mix and Comparisons

We can use the language of category theory to describe model color mixing, timbre mixing, and comparisons between them. In particular, color mixing and transitions used in painting seem to have their correspondent operations in timbre superposition and timbre morphing in orchestration and computer music. The operation of mixing two colors or timbres can be described as a weighted algebraic sum of their coordinates; the morphing through a suitable morphism from a point to another one. Intuitively, to map colors and paths between colors (morphisms) from timbres and paths between timbres (morphisms), we may need the structure of a functor. We also map classes of colors to equivalence classes of timbres, and operations on colors to operations on timbres. We provide mathematical details in Section 2.

The perceptive correspondence between classes of (visual) colors and classes of timbres has been the object of a recent experiment, summarized here. The idea is based on the *chromo-gestural similarity*, an extension of the conjecture of gestural similarity. We will define the bicategory of colors, whose objects are colors and morphisms are mappings between them (color gestures), and the bicategory of timbres, whose points are orchestral timbres, and morphisms are the morphing between them, by using topological spaces and the structure of bigroupoids (Section 2). These ideas can lead to new interfaces for color-based computer-assisted orchestration, and for the definition of musical souvenirs.

Categories abstractly help investigate some fields, including applications in the domain of computer science. Because there are more and more connections with computer science which are related with music and the visual arts as probes to investigate perception, an application of the language of categories to the arts can thus be fruitful for computer science developments in the humanities [20–22]. We can envisage techniques to translate (or, better, define a mapping between) a visual form and a(n equivalence class of) musical structure(s) [23]. However, none of the colors were mentioned before, and concerning the musical side, timbre had also been left apart. The origin of the present research started from two questions. The first question was: "What about color?" The second question was the following. At the end of a MCM 2019's talk on the Synesthesizer, a tool to map colors into timbres [24], the speaker was asked if it would have been possible to apply a perceptive-relevant association criteria to make the sonification more "effective". Following the conference, a cognitive experiment was run [25]. It showed that, given an orchestral timbre, there is no one-to-one correspondence with a color, and neither a random answer; but instead, there is a correspondence with an *equivalence class of colors*, sharing some perceptive similarity with the proposed timbre. For instance, "red" requires more attention than "gray" or "light blue", and thus, low–middle timbre brass and percussions were more likely to be associated with red or dark yellow. However, high-pitched winds may appear as being closer to light-blue or white. This appeared as being independent from cultures: in fact, there were participants from Europe, America, Africa, and Asia.

Already, the psychologist Palmer suggested to compare sounds and colors according to a similarity of perception [5]. Rosenblum and coauthors thought of a supramodal brain [26], with different senses mapping information and contributing to the elaboration of a central "mental image" of the external reality. When the vision fails, the hearing takes its role for this mind mapping, and vice versa. Moreover, several studies in the field of crossmodal correspondences investigate widely shared perceived analogies between different sensory stimuli [27]. Timbre occupies a relevant position in recent studies [28,29].

The problem of sound–color mapping has been addressed in light of synesthesia, of subjective associations, but also of experiments addressing cross-modal—and thus more widely shared—associations. Artistic approaches include Kandinsky's contribution ([1], p. 70; [30], p. 6). Kandinsky was inspired by Skrjabin, who proposed a table of equivalent tones in color and

music applied in *Prometheus: A Poem of Fire* (1908). The painter aimed to create in painting what Skrjabin had done for music, exploiting visual colors to express feelings.

Kandinsky developed the *color theory* [3] which influenced works such as his experimental theater piece, *A Yellow Sound*. For instance, according to Kandinsky, yellow reminds one of a feeling of disturbance, and blue of spiritual aspirations. Re-reading these ideas in light of our research, we may be tempted to associate 'yellow' with higher tension, and 'blue' with lower tension. To support this conjecture, a recent experiment on crossmodal perception [31] shows that the lyrical section of a Rachmaninoff Prelude is more likely to be associated with the blue.

The term 'color' is often used in music to denote some *emerging properties*, as a resulting effect including articulations and harmonic choices. This is, for instance, exemplified by Erik Christensen, talking of *timbral fluctuation and darkening harmonic color* while referring to a Schönberg's orchestral piece, *Summer Morning by a Lake* [32], p. 87.

In this article, for the sake of simplicity, we will only consider the *visual color* and the *timbre* related to instrument choice and orchestration.

The philosopher Michel Henry, who studied Kandinsky's paintings and Briesen's musical drawings in detail, considered music and visual art as an expression of some external forces, thus in this sense being *abstract* [33].

1.2. Some Empirical Evidence

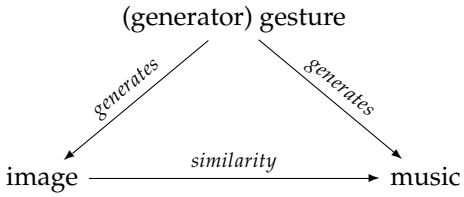
A recent experiment proposed students to associate colors with specific Rachmaninoff's passages; their answers showed similarities [31]. Other experiments combined lines and colors, asking students to improvise music according to elements from paintings by Kandinsky and Mondrian, and then comparing the results [34].

Here, we will not use the term "color" in the musical meaning. Instead, we consider color in the visual domain and timbre in the sound domain. We focus on what biologists would call *emerging properties*, and a visual artist might consider as essential lines. In the realm of color, this could be the overall effect, and for timbre, the whole effect of the instrumental combinations.

Our approach regards *classes* of timbres of colors, and shifts/mappings between them. We can think of color shifts as a "movement" in the space of colors, to be compared with a "movement" in the space of timbres. Thus, we may want to extend the definition of musical "gestures" to this new framework.

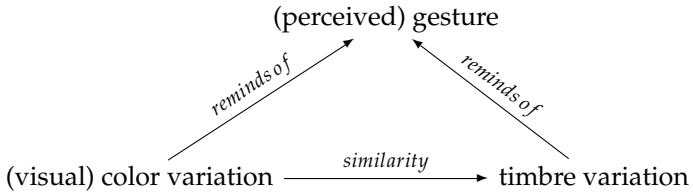
The first mathematical definition of gesture, published in the Reference [14], described a gesture as the *embodiment* of a digraph (the *skeleton*) in a topological *space*. To simplify, there are discrete points in the space of musical parameters connected by continuous curves, which represent the configuration and/or the movement of the body. This theory is based on the mathematical language of category theory [11,17,19,35].

With this in mind, we can compare gestures of different musical performers, envisaging similarities of articulation, dynamics, and so on. For example, a violinist and a pianist, to play their instruments, make different movements; but if they have to both play a *crescendo*, they will perform similar variations of movement: the pianist will increase hand pressure, and the violinist will increase bow pressure on the string. The idea of *gestural similarity* [17] can be extended to the comparisons between music and the visual arts. A sketched drawing and a short musical sequence can be considered as "similar" if they appear as being generated by the same gesture (Diagram (1)). For example, a louder sequence and a thicker pencil-made line on paper are produced by similar higher-pressure movements. As another example, the same detached movement can produce a *staccato* sequence of notes, or a collection on points on a piece of paper.



(1)

The gestural similarity criterion between music and visuals considers lines, speed, articulation, directions, but not visual color or sound timbre. An extension of gestural similarity should involve these parameters. We can consider a color variation to be *similar* to a timbre variation if they produce a similar perceptive effect; see Diagram (2), the scheme of what we can call the *chromo-gestural similarity*.



(2)

2. Categorical Depictions of Color and Timbre Gestures

In this Section, we define the bicategory of colors and the bicategory of gestures, with color gestures and timbre gestures as morphisms (paths) within them.

Let us consider the three-dimensional real Euclidean space of RGB colors, hereinafter called Colors space, and the three-dimensional Grey's space for timbres [36,37], the Timbres space. (If we add luminosity for colors, and intensity for timbres, we should consider R^4). More precisely, we can consider a variety contained in a subspace of R^4. For example, because RGB coordinates only contain positive numbers, we should focus on the octant with only positive values for color coordinates. Examples of interest deal with visible light (and thus visible colors) and audible sounds, and thus the choice of a subspace is a mandatory step.

R^3, R^4 and their subspaces are topological spaces. Let us consider a point in the R^3 Colors space as a term of coordinates; we can do the same for Timbres. The objects are the points, the 0-cells; the morphisms are the paths (gestures), the 1-cells; the bands are the 2-paths (hypergestures), that is, the 2-cells. The composition of paths is not associative: associativity is verified for equivalence classes of homotopies [38]. Thus, the bicategory COLOR and the bicategory TIMBRE can be considered as bigroupoids [39,40], that is, bicategories whose morphisms are weakly invertible (up to iso). A bigroupoid is a "bi-category [...] such that the 2-cells are strictly invertible and the 1-cells are invertible up to coherent isomorphism" [39], p. 313. According to [39], p. 312:

> In dimension 0 points p, q of X will be identified if they can be joined by a path, i.e., a continuous map $f : I \to X$ from the unit interval $I = [0, 1]$ of real numbers such that $f(0) = p$ and $f(1) = q$. This gives rise to the set of path-components, $\Pi_0(X)$, of X. In dimension 1 the points of X will be retained, but paths f, f' between fixed points p, q will be identified if there is a homotopy of rel end points between them. This gives rise to the fundamental groupoid, $\Pi_1(X)$, of X. The class of a path will be called a 1-track. Hence, the most natural approach to 2-dimensional homotopical algebra of a space X is to retain points and paths between them and identify homotopies' rel end points under a suitable homotopy relation. This gives rise to the notion of a 2-track. In this way we obtain a two-dimensional structure with points in dimension 0 (0-cells), paths in dimension 1 (1-cells), and 2-tracks in dimension 2 (2-cells). [...] Horizontal pasting is neither strictly associative, nor do we have strict identities. However, horizontal pasting is still reasonably well-behaved in the sense that associativity does hold and strict inverses do exist up to coherent isomorphisms. Thus, we obtain a bicategory,

$\Pi_2(X)$, in the sense of Bénabou. The bicategory $\Pi_2(X)$ has the additional feature that the 2-cells are strictly invertible with respect to vertical pasting and the 1-cells are invertible up to coherent isomorphism, that is, $\Pi_2(X)$ is a bigroupoid which will be called the *homotopy bigroupoid* of the topological space X.

The structure of *groupoid*, a category where all arrows are invertible, had interestingly been proposed already in 1927 (see [41] cited in the Reference [12]), page 17, well before the birth of category theory in 1945 [13]. The notion of *bigroupoid* has been independently proposed by Stevenson [42] and Hardie and coauthors [39] in 2000 and 2001, respectively; the sketch of the idea was, however, already present in a manuscript by Grothendieck of 1983 [43]. The idea of a bigroupoid has been extended to n-dimensions, with the n-groupoids, by Metere and coauthors [44,45].

We can define "color bands" as equivalence classes of colors, the 2-paths described above (Figure 1). We need the concept of homotopy categories, which generalize the notion of equivalence class [19].

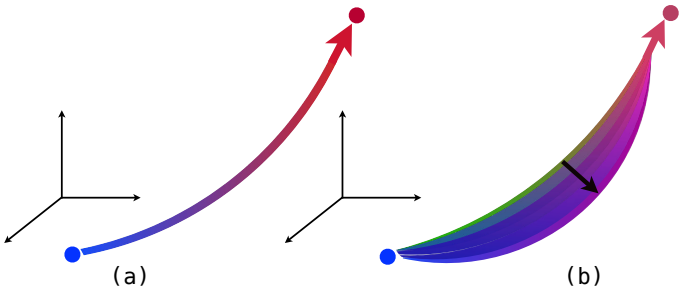


Figure 1. An example of the color gesture as a 1-path (**a**) and color band gesture as a 2-path (**b**).

Formally, to prove that COLOR and TIMBRE are bigroupoids, we should first prove that they are bicategories [39,42]. The references for these concepts are Definitions 1.1 and 1.2 from the Ref. [39] and Definitions 8.1 and 8.2 from the Reference [42]. Following Definitions 1.1 from [39], we can easily check that COLOR (TIMBRE) is a bicategory:

- There is a set of objects of COLOR, the points, that is, the 0-paths, or 0-cells (and similarly for TIMBRE);
- For each pair of objects in COLOR (TIMBRE), there is a 1-path between them, that is, an arrow or 1-cell;
- A morphism between two 1-paths exists and it is a 2-cell, here called a *color band* (timbre band);
- The composition of two 2-paths $\beta : path_1 \Rightarrow path_2$ and $\beta' : path_2 \Rightarrow path_3$ is additive: $\beta + \beta' : path_1 \Rightarrow path_3$. In fact, we can add a color band to another adjacent one, creating a larger color band (similarly for timbre bands);
- The identity element exists and it is a 1-cell: it corresponds to the lazy path for colors (timbres);
- For each triple of color points (color1, color2, color3), there is a composition functor $path(color_1, color_2) \times path(color_2, color_3) \rightarrow path(color_1, color_3)$ (same for timbres);
- The identity 2-cell (the identity 2-arrow) can be considered as a mapping from a 1-cell to itself (as the identity 1-cell maps a 0-cell to the same 0-cell);
- For each quadruple $(color_1, color_2, color_3, color_4)$, there are natural isomorphisms, the associativity isomorphisms: $\alpha : path_1 \cdot (path_2 \cdot path_3) \Rightarrow (path_1 \cdot path_2) \cdot path_3$, where $path_1 : color_1 \rightarrow color_2$, $path_2 : color_2 \rightarrow color_3$, $path_3 : color_3 \rightarrow color_4$ (same for timbre paths);
- For each pair $(color_1, color_2)$ of objects of COLOR, there are two natural isomorphisms, the left and right identities: $\lambda : [identity\ arrow\ on\ color_2] : path_1 \Rightarrow path_1$ and $\rho : f \cdot [identity\ arrow\ on\ color_1] \rightarrow path_1$ (same for timbres);
- Isomorphisms α, λ, ρ satisfy pentagonal and triangular identities, similarly to the conditions required for monoidal categories.

Following Definition 1.2 from [39], we can therefore prove that COLOR (TIMBRE) is a bigroupoid, that is, a bicategory such that:

- for each pair $(color_1, color_2)$ of objects of COLOR, the bicategory COLOR$(color_1, color_2)$ is a groupoid, that is, any 2-cell is invertible (same for TIMBRE);
- for each pair $(color_1, color_2)$ of COLOR, there is a (covariant) functor

$$F^{-1} : COLOR(color_1, color_2) \to COLOR(color_2, color_1)$$

(and similarly for TIMBRE);

- for each pair $(color_1, color_2)$ of COLOR (same for TIMBRE) there are two natural isomorphisms, the cancellation isomorphisms: $\iota : path_1^{-1} \cdot path_1 \Rightarrow identity_{color_1}$ and $\iota' : path_1 \cdot path_1^{-1} \Rightarrow identity_{color_2}$, where $path_1 : color_1 \to color_2$ is a 1-cell, that is, a 1-path, with the composition of $path$, $path^{-1}$ verifying the pentagonal relationship of Diagram (3), where p stands for $path_1$, i is the identity arrow from and to $color_1$, i' is the identity arrow from and to $color_2$, 0_p is the identity 2-cell $0_p : i \to i$.

$$\begin{array}{ccc} p \cdot (p^{-1} \cdot p) & \xRightarrow{\alpha} & (p \cdot p^{-1}) \cdot p \\ \Downarrow{\scriptstyle 0_p \cdot \iota} & & \Downarrow{\scriptstyle \iota' \cdot 0_p} \\ p \cdot i & & i' \cdot p \\ & {\scriptstyle \rho} \searrow \quad \swarrow {\scriptstyle \lambda} & \\ & p & \end{array} \tag{3}$$

In the bicategory of colors, each object is a set of coordinates that uniquely indicate a point in the RGB space (let us think of visible light, considering thus additive synthesis), and each morphism is an arrow that gradually blends a color into another one; that is, a bridge between a point with (x,y,z) coordinates and a point with (x′,y′,z′) coordinates. We have the same idea for timbres. In the case of timbre, we choose as the Euclidean space the space of Grey's study. Each point is a specific timbre, each morphism is an arrow between two points in the topological space of timbres, gradually morphing a timbre into another one. Morphisms in both spaces are (weakly) invertible because we can produce (up to iso) the inverse transformations, going either way: for example, from blue to red and from red to blue, or from a clarinet sound to a trombone sound, and from trombone to clarinet. These effects for timbres are easily produced through electronics. Regarding colors, we have two choices: either considering the painting space of colors, with subtractive synthesis (the sum of all colors is black), or the colors of visible light, with additive synthesis (the sum of all colors is white). The two spaces are dual between them. In summary, the 0-paths are points in the chosen Euclidean spaces; the 1-morphisms are paths, and the 2-morphisms are homotopy between two paths with common endpoints; as mentioned above, we describe these 2-morphisms as bands. For example, we can just move from blue to red (1-morphism), or we can create a color band as a "strip" that, going from blue to red, expands in the middle reaching shadows of violet (Figure 1). The same applies to Timbres: from clarinet to trumpet, we can enrich harmonics of the sound adding a 'color' of trumpet in our path toward trombone sound. Homotopy classes are associative up to homotopy equivalence. From a point of view of the mathematical theory of gestures, 1-paths are gestures, and 2-paths are hypergestures [14].

For each pair of objects p, q of a bigroupoid $\mathbf{S}$, Hardie et al. [39] define a covariant functor $F^{-1} : \mathbf{S}(p, q) \to \mathbf{S}(q, p)$. For the bicategory COLOR, that would correspond to a mapping from a path, let's say, red→blue to blue→red, and similarly for TIMBRES.

Thus, COLOR is the (bi)category and the bigroupoid whose objects (points) are the RGB coordinates, whose morphisms (arrows) are the paths between colors, and whose 2-morphisms are paths of paths (color bands). 1-morphisms are the 1-paths, that we call *gestures*, and the 2-morphisms are the 2-paths, that we associate to *hypergestures* [14]. Re-

garding the associativity of gestures and hypergestures (gestures of gestures), see Theorem 2 from the Reference [22].

A color path is the equivalent of "color morphing", a *shade* from a tone to another one. In the visual arts, *mixing* also has a relevant role. If we consider printing, we should talk of color layers' superposition; dealing with visible light, we have superpositions of waves with different wavelengths.

Regarding color models, we chose the simplest one. It is, under ideal conditions, that of the convex cone in the vector space, an extended RGB space [7]. According to Schrödinger: "the space of perceived colors is a regular convex cone embedded in a real vector space of dimension less or equal to 3; [...] we can endow [the space] with the operations of *sum* and *multiplication by a positive real scalar*" [7]. The mixture of two lights (superposition of light beams) corresponds to the sum of vectors, and the scalar multiplication corresponds to intensity modulation by the scalar. For the RGB model, the computer programs use an average to represent color mixtures. The intensity has a practical limitation in such devices (a maximum 255), so an average intensity is used (allowing computer-reproducibility). We remind that Pollard and Jansson proposed the *tristimulus method*, in analogy with the theory of three primary colors [46]. We might wonder if primary colors could be categorically described as (categorical) *limits*, and if there is a timbral equivalent of primary colors. However, given the structure of (bi)groupoids, all points are equivalent, and thus it is not possible to define limits or colimits.

In fact, if we had used the groupoid structure, coproducts would have existed, but they would have been ontologically equivalent to any other object. In the case of bigroupoids, we cannot even define limits and colimits in general. In this article, we are just interested in rending the artistic concept of "summing up" things, as the sound wave superposition in orchestral playing, or the superposition of different layers in printing (with a little abuse we can consider color mixing in painting as a superposition, but it's a bit more complex), then the categorical sum may not be the apt tool. Instead, we consider the weighted algebraic sum of components (Figure 2). The superposition of colors 1 and 2 leads to another color, color 3, that is, another point in the space of colors. See, for an example, https://meyerweb.com/eric/tools/color-blend/ (accessed on 16 January 2022). Selecting two colors and midpoint = 1, the coordinates of the resulting color are obtained as the halved component-wise sum of the two input colors. The sum is halved in order to obtain, let's say, red if we are summing up red with itself. With more midpoints, we obtain blended colors with different weights (that is, with weights $\neq \frac{1}{2}$). Thus, the general formula is (coordinates of color 1 * weight 1 + coordinates of color 2 * weight2). The same applies to the space of timbres. The process is similar to the structure of a monoid: there are a set and some operations on it, such that the result of an operation having as input two elements of the set gives as output another element of the set. We can think of a bicategory with only one object. Therefore, to include in our analysis color "sums" and timbre "sums", we should talk of a *monoidal structure on a bi-groupoid* [47,48]. The definition of a monoidal groupoid is given in the Reference [49]. According to [49] (p. 1), "monoidal groupoids $G = (G, \otimes, I, a, l, r)$, [are] categories G in which all arrows are invertible, enriched with a monoidal structure by a tensor product $\otimes : G \times G \to G$, a unit object I, and coherent associativity and unit constraints $a : (X \otimes Y) \otimes Z \cong X \otimes (Y \otimes Z)$, $l : I \otimes X \cong X$, and $r : X \otimes I \cong X''$.

To prove the monoidal structure of bigroupoid, let us consider Definition C.1 in the Reference [47], p. 272:

- Our structure is a bigroupoid, and thus it is a special case of bicategory;
- We can define a tensor functor as $\otimes : COLOR \times COLOR \to COLOR$, which in our case can be the following: adding two colors in COLOR gives as output another color in COLOR obtained as the weighted algebraic sum of the first two colors;
- The tensor product of a color (with itself) is the color itself: mixing white with white gives white;
- The tensor product of two objects $color_1$, $color_2$ is $color_1 \otimes color_2$;
- We have the associator $a : (color_1 \otimes color_2) \otimes color_3 \to color_1 \otimes (color_2 \otimes color_3)$;

- The pentagonator of [47] (p. 272) is verified;
- Similarly, the associahedron [47] (p. 273) is verified as well, because there is no dependency on the organization of brackets;
- There is a monoidal unit I, for example, in this case mixing white with a transparent color, as a transparent acrylic;
- There are two unitor elements $l : I \otimes color_1 \to color_1$, $r : color_1 \otimes I \to color_1$;
- There are two unitor invertible modifications λ, μ, ρ verifying triangular correspondences as shown in the Reference [47] (p. 275);
- There are four equations of modifications as shown in the Reference [47] (p. 276).

The same applies to TIMBRE. The transparent color is substituted by a superposition with a silent instrument, or muted and playing too softly to be heard.

The concept of chromo-gestural similarity allows one to pick up functors: for example, the idea of "enrichment" and "intensification" can help us map an enrichment path in COLOR (e.g., from grey to blue) with an enrichment path in TIMBRE (e.g., from a diapason, a pure sound, to clarinet, with added odd harmonics). If we consider the painting colors, we can keep adding up colors until we reach the black; if we consider the light colors, until we reach the white. From a first experimental study [25], a cluster with lower notes on a piano keyboard (with a high level of dissonance and harmonic superposition) is almost always associated with black (or dark brown, dark blue). This idea supports our intuition of chromo-gestural similarity between painting colors and timbres.

The inverse operation of superposition, the decomposition of a mixed color into its components, may remind us of the prism, with white light decomposition into rainbow colors. A similar effect can be achieved in the domain of sound with the Helmholtz resonators, as it will be mentioned later.

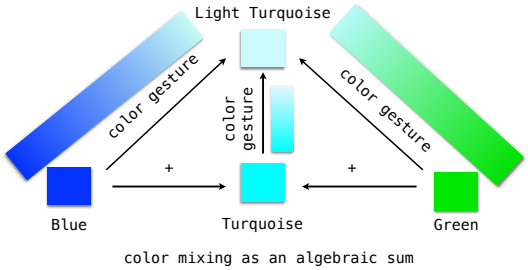


Figure 2. Color mixing as an algebraic sum, color transition as a morphism.

We can envisage a space whose axes are harmonic structure, frequency level, and intensity level, to be associated with the described bigroupoid COLOR. This could be the starting idea for the definition of a Euclidean *space of timbres*, considering the Grey's model. We can measure timbre as a function of these parameters [10]. Timbre is a complex set of properties, fundamental in music orchestration and computer music [50]. It has also been proposed to measure timbre in terms of masked absolute threshold (minimal intensity to discern a frequency among others) versus masked differential threshold (m inimal intensity to discern a change of that frequency), to measure the sensitivity of the ear to different frequencies [10]. Or, we can consider a space where directions are given by color tones of orchestral instruments, and this is at the base of orchestration [50].

We can define "timbre bands" and equivalence classes of timbres. Also in this case, we can define homotopy categories [19].

We can thus compare color gestures with timbre gestures, that is, instances of sound morphing.

The role of mixing is relevant for timbres as well, and it is fundamental in the domain of orchestration. Regarding visual/mathematical representations of spaces of timbres, we can mention the pioneering works by Xenakis [51]. The idea of timbre spaces was also used in [52].

Timbre is a difficult concept to formalize since it assumes different meanings when it is used in psycho-acoustics, in music, in audio processing, and in other disciplines [53].

Historically, timbre has been thought of as the perceptual quality of sounds that allows listeners to tell the difference between different musical instruments and, ultimately, recognize the instrument.

In the 19th century, however, Fourier analysis was used to study musical instrument sounds and to describe both the acoustics of sound production and the physiological underpinnings of sound perception [54]. This interpretation mainly applies to the steady-state portion of sounds and does not work well with the attack and decay portions, or with other temporal variations. More recently, studies showed the importance of temporal variations in the recognition of musical instruments [55,56]. These variations, together with the spectral description of sound provided by Fourier analysis, are often grouped into the expression *sound color* [57]. This expression, however, does not always successfully coexist with quantitative descriptions of sound, that are important in specific creative tasks. An attempt to provide such types of descriptions was done via the concept of *timbre spaces*.

2.1. Timbre Spaces

Describing timbre from a quantitative point of view is intrinsically complex. A first attempt, in this sense, was done by Grey by measuring the dissimilarity between pairs of musical instrument sounds [36,37]. Using this approach, it has been possible to generate a spatial configuration, called a *timbre space* (see Figure 3), where similar timbres are close together and dissimilar timbres are farther apart. Many more other strategies to build timbre spaces are in use today, involving multidimensional feature representations and scaling.

Timbre spaces are often interpreted as descriptors of *sensory qualities* of timbre. In this interpretation, two sounds can be qualitatively dissimilar independently from any sources. However, another interpretation of timbre spaces involves the absolute *categorization* of a sound into musical instruments. This interpretation is essential for the tracking over time of the identity of a sound source [58].

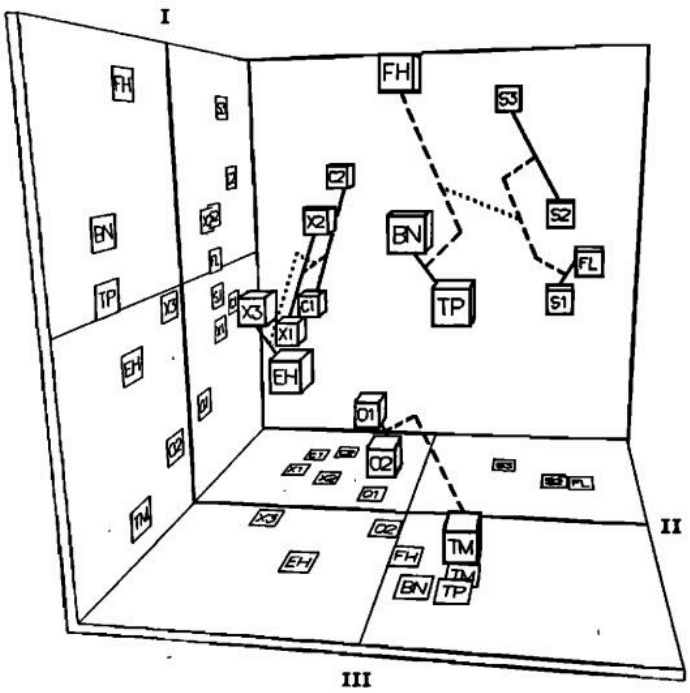


Figure 3. Grey's timbre space. Each point represents a musical instrument sound, such that similar timbres are close together and dissimilar timbres are farther apart. Reproduced from [36], with the permission of the Acoustical Society of America.

The complexity of timbre perception plays a major role in the difficulty to formalize musical orchestration, as we shall see.

3. Mapping of Color Classes onto Timbre Classes

In this Section, we provide a definition and a couple of examples of mapping Color→ Timbres and vice versa. In this framework, the COLOR-TIMBRE functor maps 1-paths in COLOR to 1-paths in TIMBRE, and the TIMBRE-COLOR maps 1-paths in TIMBRE to 1-paths in COLOR. The same goes for 2-paths, that is, bands, that is, hypergestures. The definition of functors between bigroupoids is not trivial. This is discussed in detail in the Reference [39]. If we think of ∞-groupoids consisting of their simplicial complexes (also with associativity defined up to homotopy), then the comparison between timbres and colors is realized through an ∞-functor between their ∞-groupoids. Properties of ∞-groupoids are described in detail in the Reference [40].

Let us give the definition of a 2-functor between groupoids provided in the Ref. [39] (p. 316): Given two bigroupoids $\mathbf{S}$, $\bar{\mathbf{S}}$, a 2-functor $F : \mathbf{S} \to \bar{\mathbf{S}}$ is given by the following data:

- A map $F : Ob(\mathbf{S}) \to Ob(\bar{\mathbf{S}})$ for the objects;
- For each pair (p, q) of objects of $\mathbf{S}$, a functor $F : \mathbf{S}(p, q) \to \bar{\mathbf{S}}(Fp, Fq)$ for the morphisms.

In our research, we can thus define the 2-functor $F_{C,T} : COLOR \to TIMBRE$ and the 2-functor $F_{T,C} : TIMBRE \to COLOR$. $F_{C,T}$ maps' colors (as 0-paths, that is, points in the RGB space) to timbres (as points in the timbre space), and color gestures (as 1-paths in the color space) to timbre gestures (1-paths in the timbre space), and vice versa, respectively. See Figure 4 for an intuitive representation of this idea.

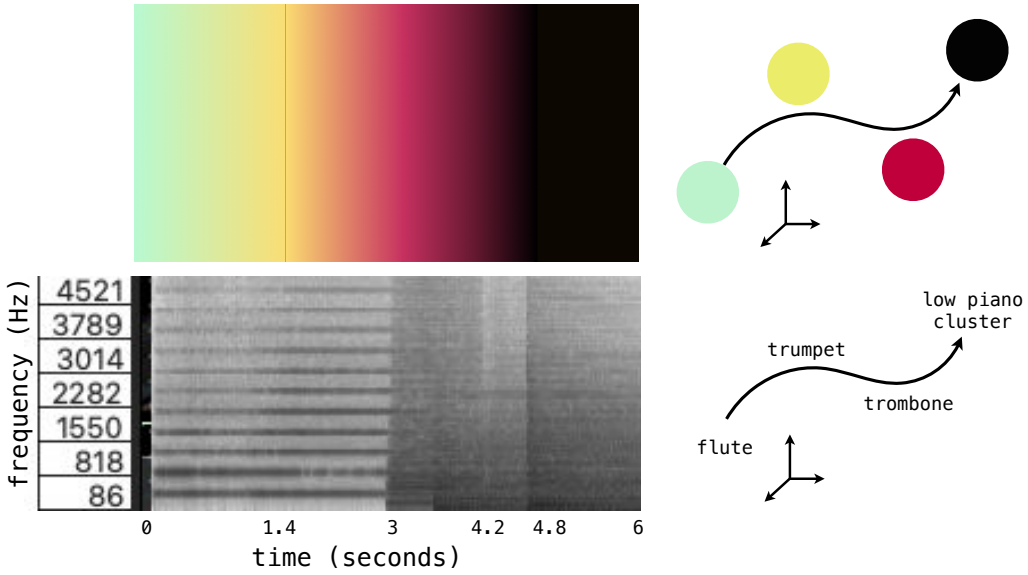


Figure 4. Top, **left**: a sequence and partial superposition of colors; **bottom**, **left**: a similar sequence and partial superposition of timbres; bottom, right: the corresponding timbre gesture; **up**, **right**: the corresponding color gesture.

Let us make a little note regarding similarity. The idea of similarity comes in while comparing the idea of superposition in the space of Colors to a superposition in the space of Timbres: for example, comparing a diagram of color sum with a sum of timbres of Clarinet and Horns. In our framework, we start from the (qualitative) definition of "gestural similarity" between images and sound [22]: a visual sketch in the domain of visuals and a short musical sequence can be judged as "similar" if they appear as being produced by the same creative gesture. Regarding colors and timbres: we can associate a path in the space of colors with a path in the space of timbres if they appear as instances of the

same (perceived) process, for example, "intensification", "rarefaction", "progress toward arousal", "suspension", and so on.

In Figure 4, we have a sequence of colors, as different points touched in a gesture, from light blue/green to black, and the spectrogram of a sequence of timbres, with a progressive enrichment of harmonics, in the upper register (from flute to trumpet) and in the lower register (with trombone and low piano clusters). This example can also be coded by using subtractive synthesis, simultaneously generating a sound enriched toward the "dark" low cluster on a piano and a color superposition reaching black. The perceptive correspondence between a "darkening" sound and a visual progression toward the black (or dark brown, or dark blue) has been observed in the aforementioned experiment [25]. The dual example, using additive synthesis instead of subtractive synthesis, is less intuitive in terms of similarity, but still feasible and coherent. It would consist in the progressive partials addition reaching white noise (timbres), and a color superposition reaching white light (colors). Such a structured application has recently been coded in the Reference [20].

We argue that TIMBRE/COLOR mapping can be validated by some underlying universal property, that is, a reference to a *common perception*, see Figure 5. The idea of *universal property* is fundamental in category theory; here, it is used more metaphorically.

Diagram (2) is based on the idea of perceived similarities between color bands and timbre bands. Variations can thus be seen as paths within these bands, or between a band and another one. To test perceived similarities, an experiment was run [25]. Let us now describe this experiment in more detail. It uses the bicategory of timbres on one side, and the bicategory of colors on the other one. The perception of a synesthetic person can be depicted as an arrow picking up some points from the timbre bicategory, and mapping them onto (always the same points given the same outputs, as in a function) points of the color bicategory. For a non-synesthetic person, the target is not the same: there is *a range* of targets. The experiment focuses on the identification of these clusters. Other studies focused on pitch-range and timbre association with color shades, and aimed to evaluate the difference between shared and individual associations [21], but also for non-synesthetes, there are identifiable trends and color-gradients. The experiment design is constituted by two parts. We notice that Figure 1 from the Reference [21] contains a triangular diagram, with unimodal auditory, unimodal vision, and multi-modal regions (spatially organized), to investigate synesthesia, which could also be adapted to our case.

In the first part, some general information about participants were gathered: age, artistic education (if any), synesthetic experiences (if any), and visual or auditory impairment (if any). In the second section, participants were asked to match sounds to colors. They were assigned eight different orchestral realizations of a C-major chord (simulated in a DAW with a professional sample library). Participants had to match sounds with a freely chosen color (using a color picker) or with a color gradient among four gradients provided. Results were plotted in a 3D graph where the RGB (Red, Green, Blue) values of selected colors were used as coordinates. The results showed that color associations tended to occupy some portions of space (i.e., they were not randomly distributed) and tended to condense in the neighborhood of some specific colors (i.e., RGB values-coordinates). A quite high correlation (0.5) was also found between synesthetic experiences and some associations. A non-negligible correlation (0.32) was found between age and some color associations.

Figure 5 shows the mapping from a point and a loop in the space of timbres, to a point and a (non-loop) arrow in the space of colors. These paths produces similar perception, which justifies the mapping itself. The timbre with a loop indicate a single sound sample, as the ones used in the experiment [25]; the box with the color space is one of the experimental outcomes—answers belong to a specific region of space.

Of course, the different registers of instruments, which can produce different perceptions, have been taken into account.

A possible next experiment could involve the idea of natural transformations and several repetitions of the test for each participant. If we ask a participant to repeat his or her associations of timbres and colors, and we have several participants, we can compare their

different answers by using natural transformations. Let be G_1, G_2 the functors representing answers of participants 1 and 2, respectively. We can define a natural transformation $\nu_{1,2} : G_1 \to G_2$ comparing the two functors. According to the chromo-gestural similarity, for non-synesthetic participants we expect a significant overlap between the color bands, and a small amount of functors' difference, quantified by $\nu_{1,2}$. In the case of synesthetic participants, however, the overlap might be missing: given the same timbre, each synesthete is expected to choose the same color, and the color chosen by the first person could be different from the color chosen by the second person. We would also expect that both colors lie within the same color band which can be find out comparing answers of multiple participants, as it happened in the first experiment (Figure 5).

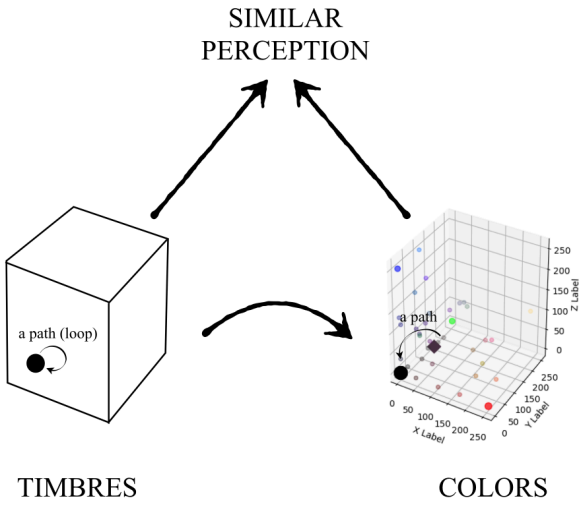
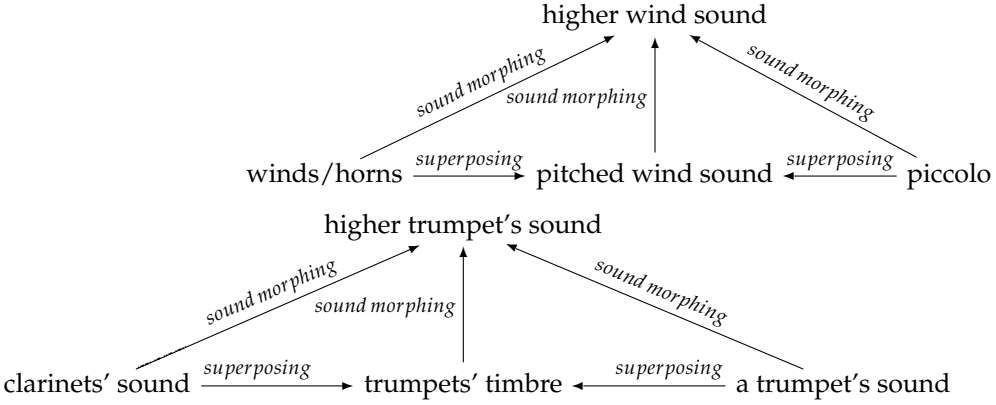


Figure 5. The functor which maps points in the space of colors to points in the space of timbres, and color gestures to timbre gestures, implies the existence of a similar perception which justifies the association. The experiment described in Section 3 aimed to show the association between color points (with loop arrows) and equivalence classes of timbre points (with arrows between them). The color cluster is one of clusters actually found in the experiment.

4. Morphing versus Mixing

We can create a chromatic (visual, non musical) progression from a color to another one. However, we can also mix colors. Similarly, we can create a progression from a timbre to another one, as a timbre morphing; but we can also superpose timbre to obtain new ones. The equivalent of color gesture in the space of timbres is the timbre morphing (e.g., gradually transitioning from an orchestral timbre to another one). The equivalent of color mixing in the space of timbres is timbre mixing (e.g., adding different orchestral instruments together in orchestration). Diagrams (4) and (5) show two examples.



Orchestration is a substantial part of musical composition, and automatic orchestration exploits machine learning techniques. The developments of these techniques may profit from a multisensory approach, such as the chromo-gestural similarity; see more details in the Discussion section.

The equivalent of the prism's action is the use of multiple Helmholtz resonators. The main difference is the following: with a single resonator, we separate a sound frequency from the rest; with a set of resonators, we can separate all frequencies—and that has been done through the recent acoustic prism [9]. A prism acts as a separator of color components, that is, it acts as the dual of the algebraic sum for color superposition. In fact, the prism separates white light into colors. We can wonder if a "local prism" may isolate just some color components. More details on morphing and hybridization are provided in the following section.

4.1. On Orchestration, Morphing and Hybridization

Generally speaking, orchestration can be seen as the blending of instrument timbres together [59]. Originally being the simple assignment of instruments to pre-composed parts of the score, orchestration has become an integral part of the compositional process [60,61]. Around the beginning of the 20th century, composers started imagining new timbres made of extended instrumental tec hniques and unusual instrument combinations and this motivated the need for a more systematic approach to orchestration, able to take into account the evolving timbre of musical instruments (*sound color*) and the logical principles of music composition.

An example of music composition that involves such advanced orchestration principles is the piece *Metastaseis* (1953–1954) for 61 instruments, by the Greek composer and architect Iannis Xenakis. The piece is influenced by the Einsteinian view of time and by Xenakis' memory of the sounds of warfare. New orchestration and compositional practices are created by the composer by using specific mathematical ideas derived from architecture. The left side of Figure 6 shows a sketch of string glissandi in measures 309–314 of the piece. The right side of the same figure shows the *Philips Pavilion* designed by Le Corbusier. It is evident how the musical *gestures* are derived by the hyperbolic paraboloids of the building.

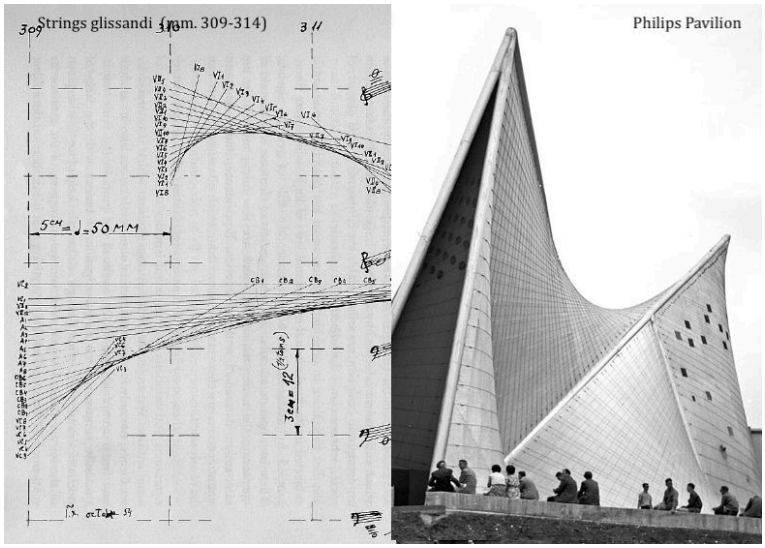


Figure 6. Left side: a sketch of strings glissandi (mm. 309–314) from Xenakis' *Metastaseis*. **Right** side: a photo of the *Philips Pavilion* by Le Corbusier.

The string glissandi in Xenakis' piece are used to transition from a sound to another in a continuous manner. In this respect, it is interesting to note how the new orchestrational practices defined by composers also deal with spectromorphological concepts such as *morphing* and *hybdridization*. Sound morphing aims at creating a new sound that is a

perceptual intermediate between two (or more) sounds. Usually, morphing is based on interpolation and assumes a gradual transition between the involved sounds. Sound hybridization, on the other hand, refers to a process that combines two (or more) distinct objects in a discrete fashion by taking elements of each one [62]. Being able to formalize perceptual processes in the context of orchestration would be an important step toward a generalized theory of music composition. Examples of timbre spaces are proposed by [63,64].

4.2. Assisted Orchestration and Sound Colors

The main interface used in *Orchidea* is similar to a score [65]. The user is asked to specify the orchestra to be used in the process, together with the target sound and other symbolic constraints. This is, usually, a convenient way for composers to model their creative thinking. However, this strategy enforces an orchestration approach that focuses on the categorical interpretation of timbre described in Section 2.1, through the selection of instruments. An alternative approach could also be possible, that is closer to the *sensorial* interpretation of timbre. Instead of specifying musical instruments, the composer could be asked to select orchestral *registers* and leave the selection of the instruments to the search algorithm. In this way, the focus would shift on all the similar timbres that each orchestral register contains and would underline the relationships between timbres. This approach is currently under development and will be included in a future release of *Orchidea*.

5. The Souvenir Theorem

We can simplify a visual form through a superposition of simple form with suitable coefficients, telling where, how big, and how many the simple forms are. These simple forms and their coefficients can be mapped to simple musical sequences and musical coefficients (telling when, how loud, how often a sequence is played), constituting a rough musical piece, which can be later refined [23]. There is a minimum number of simple forms that allows the original object recognition. Of course, to each visual form, there is an equivalence class of short musical sequence which can work well. The idea of form simplification often inspired *souvenir* production. According to the *Souvenir Theorem*, a souvenir is more likely to be sold if: (1) it shows a minimum number of simple forms; (2) it has enough material quality; (3) it is cheap enough to fit the traveler's budget. Applying the mapping from visual to musical forms, one can create *musical souvenirs*. Their support can be musical cards, and their budget can be small enough. The mapping visual form/musical form is based on gestural similarity. The mapping from colors to timbres is based on chromo-gestural similarity.

On Computer-Assisted Orchestration

The manipulation of highly complex timbral mixtures has become more and more difficult to predict while composing. In such a context, having a tool able to simulate the result of such mixtures during the compositional phase appeared to be a necessity. This need, fostered by the advancements in machine learning techniques, converged into a field of research that today is known as *computer-assisted orchestration* (CAO).

Broadly speaking, CAO is the process of helping the composer orchestrate by means of computer programs and machine learning. Among the different possible types of assisted orchestration, there is one called *target-based*. In the target-based approach, a sound is used as reference to derive the orchestration parameters. By means of sophisticated search algorithms, a combination of orchestral sound samples is chosen from a large database of instrument notes to perceptually match the target sound. Each sample from the database comes with symbolic information such as the instrument name, the pitch, the dynamics and the playing style. As such, the combination can be then symbolically represented in a score [66].

CAO is a long-standing problem for composers and has its roots in *spectral music*, a specific music composition practice where musical parameters are derived by the analysis of sound spectra [67]. It has been of paramount importance at the *Institut de Recherche et Coordination Acoustique/Musique* (IRCAM) in Paris, http://www.ircam.fr (accessed on 16

January 2022), where a number of software packages, papers and PhD thesis have been produced on the subject in the last twenty years. The most recent and most complete system for target-based computer-assisted is *Orchidea*, http://www.orch-idea.org (accessed on 16 January 2022), developed jointly between IRCAM, the Haute École de Musique de Genève and the University of California, Berkeley. *Orchidea* implements a number of advanced features such as joint spectro-temporal orchestration, that allows to orchestrate dynamically evolving timbres [65].

A piece composed with the support of *Orchidea* is *Stades d'ombre, stades de lumière* (2018) for 14 instruments, by [one of the authors], performed for the first time at *Milano musica* in 2018 by the *Ensemble orchestral de Lyon* under the direction of Daniel Kawka, http://www.milanomusica.org/it/sezione-media/news/560-lunedi-29-ottobre-doppio-concerto-al-teatro-elfo-puccini.html (accessed on 16 January 2022). The work is made of different sections, each one based on a dynamically evolving sound used as a target, including field recordings of different natural environments, such as bells, docking boats, water drops and musical instruments.

6. An Audio Example

As an additional example regarding timbre gestures, we show morphing from a note played with the theremin to the same note played with the flute. Figure 7 shows the sonograms of the initial and final sounds, recorded by one of the authors, and a sequence of an intermediate timbre morphing between them, made with the spectral model synthesis (SMS Tools) https://github.com/MTG/sms-tools (accessed on 16 January 2022). Figure 8 shows a superposition of the two audio sequences and Figure 9 presents two different timbre gestures to make the sound morphing, using different parameters of the SMS tools. The presented sonograms are made with SMS tools and with an online software: https://convert.ing-now.com/audio-spectrogram-creator/(accessed on 16 January 2022).

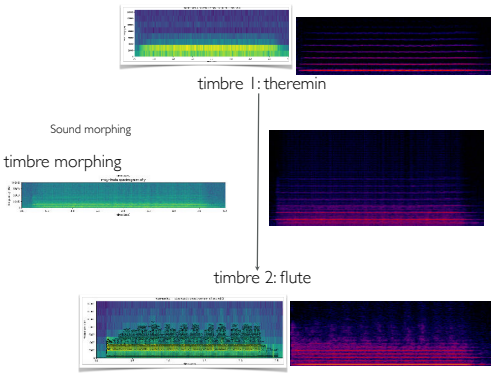


Figure 7. Timbre morphing of an A played with theremin to an A played with flute, with sonograms.

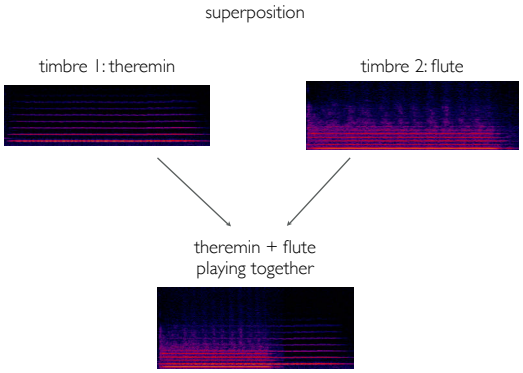


Figure 8. Simple juxtaposition of the A played with theremin and with flute.

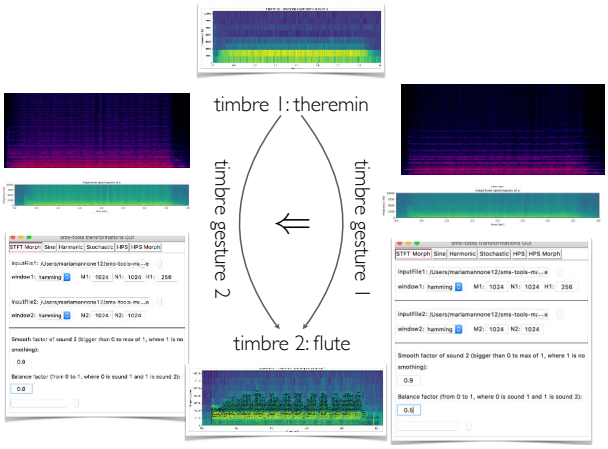


Figure 9. A change of parameters of sound morphing gives different "timbre gestures", that is, transitions from a timbre to another one.

7. Cultural Influences

Before presenting discussions and conclusions, let us mention some cultural influences on color perception. In the aforementioned experiment, participants were from different continents; they were given a slider to select colors, because the use of words could have been misleading. In fact, because we wanted to analyze shared aspects, is important to avoid bias linked to different cultures. It is interesting to note that culture can affect color perception. In particular, different languages and cultural groups measure out the color spectrum differently. In 1969, Berlin [68] published the book *Basic Color Terms: their Universality and Evolution*. Berlin and Kay concluded: "The perception of color mainly occurs inside our heads, so its existence in the mind is influenced by personal feelings, tastes or opinions." In later studies, cross-cultural studies comparing the speakers of different languages challenged the Berlin and Kay's theory and revealed that language may affect color perception. A rich culture widens color perception. Unlike some cultures, in Nigeria, color choices or perceptions are often influenced by Religion. Other influencing factors are 2-languages influences in bilingual people. As an example of non-Western languages and cultures, the following table shows the visual associations to color and their corresponding words in languages used in Nigeria.

English	Blue	English	Yellow
Yoruba	Olomi Aro (blue)	Yoruba	Elezuru (yam special)
Igbo	Oji (something dark)	Igbo	Onashara (light white)
Tiv	Kwar Kwaodo (like sky)	Tiv	Oyha (like banana fruit)
Owan	Iblue (sky)	Owan	(paint of banana)
Urhobo	Oda dibo (paint of banana)	Ijaw	Pinapina (just like white)
English	Green	English	Red
Yoruba	Alawo ewe (color of leaves)	Yoruba	Pupa
Igbo	Akwukwo Ndu	Igbo	Uhie (color of blood)
Tiv	Ngu-er-ka Ikya uwer nahan	Tiv	Nyian
Owan	Ebesugbo (leaf)	Heusa	Ja
Ijaw	Deibide (like a particular cloth)	Urhobo	Oda Obara (blood-like)
Hausa	Igreen	Ijaw	Kwekwe

8. Discussion and Conclusions

In this article, we compared mixing and transformations for colors and timbres, as paths in the spaces of colors and of timbres, respectively, and we formalized them in the theoretical framework of category theory. Timbre mixing and variations are fundamental in the domain of orchestration, while color mixing and shadings in the domain of painting. Theoretical formalization can lead to further technological developments.

While analyzing similarities in color and timbre transformations, we proposed the idea of *chromo-gestural similarity*. Possible applications of the proposed chromo-gestural similarity involve further cognitive studies, hospital patients' self-reports based on color variations and sounds, visually-aided interfaces for automatic orchestration, and new approaches to disability studies and multisensory interfaces.

Such interfaces may be applied to the Synesthesizer [24], a tool that uses machine learning to generate sounds from color RGB values. The three-dimensional input (RGB coordinates) produces, as an output, 35 parameters that control five different physical models of different instruments and their cross-synthesis. The training of the machine learning model is performed on few values and the remaining combinations are generated through a linear regression. The resulting sounds are not predictable, although they tend to share similarities with the training set. The training step could profit from the database acquired for the experiment [25], and a new interface can be conceived after chromo-gestural similarity. The precise timbre choices can exploit tools of computer-assisted orchestration, creating a new instrument made of different modules: a probe to scan a complex visual image (as the one used in the Synesthesizer) as input, visual to musical-sequence mapping, color evaluations, color-to-timbre mapping, timbral refinement, and final sound result as the output.

The existence of automatic techniques to orchestrate and to produce sounds and specific timbre combinations starting from a visual source, may in fact open new frontiers to rethinking music composition and music analysis.

In our vision, the interdisciplinary dialogue between different fields of human creativity, in the arts and in the sciences, might profit from these new perspectives. Additionally, the exchanges between different musical cultures can be enhanced via new tools for multisensory expression empowered by the unifying perspective of mathematics.

Author Contributions: Conceptualization, M.M.; methodology, M.M., G.S., E.A., and C.E.C.; software, M.M.; validation, M.M.; formal analysis, M.M.; investigation, M.M.; resources, M.M., G.S., E.A., C.E.C.; data curation, M.M.; writing—original draft preparation, M.M., G.S., E.A., C.E.C.; writing—review and editing, M.M.; visualization, M.M.; supervision, M.M. and C.E.C.; project administration, M.M. and C.E.C.; funding acquisition, C.E.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We are grateful to Giuseppe Metere and Juan Sebastián Arias-Valero for their suggestions on mathematical sections.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Davis, J.; Leshko, J.; Fabing, S.J. *Smith College Museum of Art: European and American Painting and Sculpture 1760–1960*; Hudson Hills Press: Boston, MA, USA, 2000.
2. von Goethe, W. *Theory of Colours (Zur Farbenlehre)*; J.G. Cotta'schen Buchhandlung: Stuttgart, Germany, 1810.
3. Kandinsky, W. *Complete Writings on Art*; Lindsay, K.C., Vergo, P., Eds.; Da Capo Press: New York, NY, USA, 1994.
4. Newton, I. Opticks, or, A Treatise of the Reflections, Refractions, Inflections and Colours of Light. 1704. Available online: https://www.loc.gov/resource/rbctos.2017gen39060/?st=gallery (accessed on 16 January 2022).
5. Palmer, S.E.; Schloss, K.B.; Xu, Z.; Prado-Leon, L. Music-color associations are mediated by emotion. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 8836–8841. [CrossRef] [PubMed]
6. Pinna, B. What colour is it? Modal and amodal completion of colour in art, vision science and biology. *Int. J. Arts Technol.* **2010**, *3*, 2–3. [CrossRef]
7. Provenzi, E. Geometry of color perception. Part 1, Structures and metrics of a homogeneous color space. *J. Math. Neurosci.* **2020**, *10*, 1–19. [CrossRef]

8. Sethares, W. *Tuning, Timbre, Spectrum*; Springer: Berlin/Heidelberg, Germany, 2005.
9. Esfahlani, H.; Karkar, S.; Lissek, H.; Mosig, J.R. Acoustic dispersive prism. *Nat. Sci. Rep.* **2016**, *6*, 18911. Available online: https://www.nature.com/articles/srep18911 (accessed on 16 January 2022). [CrossRef] [PubMed]
10. Lewis, D.; Larsen, M.J. The Measurement of Timbre. *J. Acoust. Soc. Am.* **1937**, *8*, 207. Available online: https://asa.scitation.org/doi/10.1121/1.1901995 (accessed on 16 January 2022). [CrossRef]
11. Mac Lane, S. *Categories for the Working Mathematician*; Springer: New York, NY, USA, 1971.
12. Grandis, M. *Higher Dimensional Categories*; World Scientific: Singapore, 2020.
13. Eilenberg, S.; Mac Lane, S. General Theory of Natural Equivalences. *Trans. Am. Math. Soc.* **1945**, *58*, 231–294. Available online: https://www.ams.org/journals/tran/1945-058-00/S0002-9947-1945-0013131-6/S0002-9947-1945-0013131-6.pdf (accessed on 16 January 2022). [CrossRef]
14. Mazzola, G.; Andreatta, M. Diagrams, gestures and formulae in music. *J. Math. Music* **2010**, *1*, 23–46. [CrossRef]
15. Jedrzejewski, F. *Hétérotopies Musicales: Modèles Mathématiques de la Musique*; Hermann: Paris, France, 2019.
16. Popoff, A.; Andreatta, M.; Ehresmann, A. A Categorical Generalization of Klumpenhouwer Networks. In Proceedings of the Conference Mathematics and Computation in Music (MCM 2015), London, UK, 22–25 June 2015; Collins, T., Volk, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2015.
17. Mannone, M. Introduction to Gestural Similarity in Music: An Application of Category Theory to the Orchestra. *J. Math. Music* **2018**, *12*, 63–87. [CrossRef]
18. Mannone, M.; Turchet, L. Shall We (Math and) Dance? In Proceedings of the International Conference on Mathematics and Computation in Music (MCM 2019), Madrid, Spain, 18–21 June 2019; Montiel, M., Gómez-Martín, F., Agustín-Aquino, O.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2019.
19. Arias-Valero, J.S.; Lluis-Puebla, E. Simplicial Sets and Gestures: Mathematical Music Theory, iNfinity-Categories, Homotopy, and Homology. Compositionality (2020 to Appear). Available online: https://jusearva.files.wordpress.com/2020/08/simpgest.pdf (accessed on 16 January 2022).
20. Mannone, M.; Arias-Valero, J.S. Some mathematical and Computational Relations between Timbre and Color. *Math. Comput. Music. Conf.* 2022, *under review*.
21. Ward, J.; Huckstep, B.; Tsakanikos, E. Sound-Colour Synaesthesia: To What Extent Does it Use Cross-Modal Mechanisms Common to Us All? *Cortex* **2006**, *42*, 264–280. [CrossRef]
22. Mannone, M. Knots, Music and DNA. *J. Creat. Music Syst.* **2018**, *2*, 32–54. Available online: https://www.jcms.org.uk/article/id/523/ (accessed on 16 January 2022). [CrossRef]
23. Mannone, M.; Favali, F.; Di Donato, B.; Turchet, L. Quantum GestART: Identifying and applying correlations between mathematics, art, and perceptual organization. *J. Math. Music* **2020**, *15*, 62–94. [CrossRef]
24. Santini, G. Synesthesizer: Physical Modelling and Machine Learning for a Color-Based Synthesizer in Virtual Reality. In Proceedings of the International Conference on Mathematics and Computation in Music MCM 2019, LNAI 11502, Madrid, Spain, 18–21 June 2019; Montiel, M., Gómez-Martín, F., Agustín-Aquino, O.A., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 229–235.
25. Mannone, M.; (Department of Engineering, University of Palermo, Palermo, Italy); Santini, G.; (Ca' Foscari University of Venice, Via Torino, 107, Venezia Mestre, Italy). Perceived Similarities between Classes of Colors and Classes of Timbres. Available from the Authors upon Request. 2020, *unpublished*.
26. Rosenblum, L.D.; Dias, J.W.; Dorsi, J. The Supramodal Brain: Implications for Auditory Perception. *J. Cogn. Psychol.* **2016**, *29*, 65–87. [CrossRef]
27. Parise, C.V. Crossmodal Correspondences: Standing Issues and Experimental Guidelines. *Multisensory Res.* **2016**, *29*, 7–28. [CrossRef] [PubMed]
28. Itoh, K.; Sakata, H.; Kwee, I.L.; Nakada, T. Musical pitch classes have rainbow hues in pitch class-color synesthesia. *Nat. Sci. Rep.* **2017**, *7*, 17781. [CrossRef]
29. Thoret, T.; Caramiaux, B.; Depalle, P.; McAdams, S. Learning metrics on spectrotemporal modulations reveals the perception of musical instrument timbre. *Nat. Hum. Behav.* **2020**, *5*, 369–377. [CrossRef]
30. Favali, F. Shapes and Notes: Transforming Images into Musical Structures. Ph.D. Dissertation, University of Birmingham, Birmingham, UK, 2019. Available online: https://etheses.bham.ac.uk/id/eprint/9571/ (accessed on 16 January 2022).
31. Crnjanski, N.; Tomas, D. Musical Perception and Visualization. Music and Spatiality. Paper read at Music and Spatiality. In Proceedings of the 13th Biennale International Conference on Music Theory and Analysis, Belgrade, Serbia, 4–6 October 2019.
32. Christensen, E. *The Musical Timespace: A Theory of Musical Listening*; Aalborg University Press: Aalborg, Denmark, 1996.
33. Seyler, F. *"Michel Henry" The Stanford Encyclopedia of Philosophy (Fall 2019 Edition)*; Zalta, E.N., Ed.; Stanford University: Palo Alto, CA, USA, 2019; Available online: https://plato.stanford.edu/archives/fall2019/entries/michel-henry/ (accessed on 16 January 2022).
34. Fortuna, S. Images to play: An Improvisation Model between Visual Arts and Music. In *Konteksty Kultury I Edukacji Muzycznej*; Parkita, E., Parkita, A., Sztejnbis-Zdyb, J., Eds.; Uniwersytetu Jana Kochanowskiego: Kielce, Poland, 2020; pp. 109–118.
35. Arias-Valero, J.S.; Lluis-Puebla, E. Some Remarks on Hypergestural Homology of Spaces and Its Relation to Classical Homology. *J. Math. Music* **2020**. [CrossRef]
36. Grey, J. Multidimensional Perceptual Scaling of Musical Timbres. *J. Acoust. Soc. Am.* **1977**, *61*, 1270–1277. [CrossRef]

37. McAdams, S.; Winsberg, S.; Donnadieu, S.; De Soete, G.; Krimphoff, J. Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities, and Latent Subject Classes. *Psychol. Res.* **1995**, *58*, 177–192. [CrossRef]
38. Hatcher, A. *Algebraic Topology*; Cambridge University Press: Cambridge, UK, 2018.
39. Hardie, K.A.; Kamps, K.H.; Kieboom, R.W. A Homotopy Bigroupoid of a Topological Space. *Appl. Categ. Struct.* **2001**, *9*, 311–327. [CrossRef]
40. Porter, T. Spaces as Infinity-Groupoids. In *New Spaces in Mathematics Formal and Conceptual Reflections*; Cambridge University Press: Cambridge, UK, 2021; pp. 258–321. Available online: http://nlab-pages.s3.us-east-2.amazonaws.com/nlab/files/Spaces+as+infinity-groupoids.pdf (accessed on 16 January 2022).
41. Brandt, H. Über eine Verallgemeinerun des Gruppenbegriffes. *Math. Ann.* **1927**, *96*, 360–366. [CrossRef]
42. Stevenson, D. The Geometry of Bundle Gerbes. Ph.D. Thesis, University of Adelaide, Adelaide, Australia, 2000. Available online: https://arxiv.org/pdf/math/0004117.pdf (accessed on 16 January 2022).
43. Grothendieck, A. *À la Poursuite des Champs*; 1983. Typescript, English Translation in. Available online: https://thescrivener.github.io/PursuingStacks/ps-online.pdf (accessed on 16 January 2022).
44. Kasangian, S.; Metere, G.; Vitale, E.M. The ziqqurath of exact sequences of n-groupoids. *Cahiers de Topologie et Géométrie Différentielle Catégoriques* **2011**, *52*, 2–44. Available online: http://www.numdam.org/item?id=CTGDC_2011__52_1_2_0 (accessed on 16 January 2022).
45. Metere, G. The Ziqqurath of Exact Sequences of n-Groupoids. Ph.D. Thesis, University of Milano, Milano, Italy, 2008. Available online: https://arxiv.org/pdf/0802.0800.pdf (accessed on 16 January 2022).
46. Pollard, H.F.; Jansson, E.V. A Tristimulus Method for the Specification of Musical Timbre. *Acustica* **1982**, *51*, 162–171
47. Schommer-Pries, C.J. The Classification of Two-Dimensional Extended Topological Field Theories. Ph.D. Dissertation, University of California, Berkeley, CA, USA, 2014.
48. Hesse, J. Group Actions on Bicategories and Topological Quantum Field Theories. Ph.D. Dissertation, Staats-und Universitätsbibliothek Hamburg Carl von Ossietzky, Hamburg, Germany, 2017. Available online: https://d-nb.info/1137625104/34 (accessed on 16 January 2022).
49. Calvo, M.; Cegarra, A.M.; Heredia, B.A. Structure and Classification of Monoidal Groupoids. *Semigroup Forum* **2013**, *87*, 35–79. [CrossRef]
50. McAdams, S. Timbre as a structuring force in music. *Proc. Meet. Acoust.* **2013**, *19*, 035050.
51. Xenakis, I. *Formalized Music*; Pendragon Press: New York, NY, USA, 2001.
52. Wessel, D. Timbre Space as a Musical Control Structure. *Comput. Music. J.* **1974**, *3*, 45–52. Available online: https://www.cnmat.berkeley.edu/sites/default/files/attachments/Timbre-Space.pdf (accessed on 16 January 2022). [CrossRef]
53. McAdams, S.; Giordano, B.L. The Perception of Musical Timbre. In *The Oxford Handbook of Music Psychology*; Hallam, S., Cross, I., Thaut, M., Eds.; Oxford University Press: New York, NY, USA, 2009; pp. 72–80.
54. Siedenburg, K.; Saitis, C.; McAdams, S. The present, past, and future of timbre research. In *Timbre: Acoustics, Perception, and Cognition*; Siedenburg, K., Saitis, C., McAdams, S., Popper, A.N., Fay, R.R., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 1–19.
55. Risset, J.C. Computer study of trumpet tones. *J. Acoust. Soc. Am.* **1965**, *38*, 912. [CrossRef]
56. Mathews, M.V.; Miller, J.E.; Pierce, J.R.; Tenney, J. Computer study of violin tones. *J. Acoust. Soc. Am.* **1965**, *38*, 912–913. [CrossRef]
57. Slawson, W. *Sound Color*; University of California Press: Berkeley, MA, USA, 1985.
58. Siedenburg, S.; McAdams, S. Four distinctions for the auditory "wastebasket" of timbre. *Front. Psychol.* **2017**, *8*, 1747 [CrossRef] [PubMed]
59. Nouno, G.; Cont, A.; Carpentier, G.; Harvey, J. Making an Orchestra Speak. In Proceedings of the Sound and Music Computing, Porto, Portugal, 23–25 July 2009.
60. Kendall, R.A.; Carterette, E.C. Identification and blend of timbres as a basis for orchestration. *Contemp. Music. Rev.* **1993**, *9*, 51–67. [CrossRef]
61. Rose, F.; Hetrik, J.E. Enhancing Orchestration Technique via Spectrally Based Linear Algebra Methods. *Comput. Music. J.* **2009**, *33*, 32–41. [CrossRef]
62. Caetano, M.; Rodet, X. Musical instrument sound morphing guided by perceptually motivated features. *IEEE Trans. Audio Speech Lang. Process.* **2013**, *21*, 1666–1675. [CrossRef]
63. Grey, J.; Gordon, J.W. Perceptual effects of spectral modifications on musical timbres. *J. Acoust. Soc. Am.* **1978**, *65*, 1493–1500. [CrossRef]
64. Grey, J.; Moorer, J.A. Perceptual evaluation of synthesized musical instrument tones. *J. Acoust. Soc. Am.* **1977**, *62*, 454–462. [CrossRef]
65. Cella, C.-E. Orchidea: A comprehensive framework for target- based computer-assisted dynamic orchestration. *J. New Music. Res. (JNMR), under review*.
66. Maresz, Y. On computer-assisted orchestration. *Contemp. Music. Rev.* **2013**, *32*, 99–109. [CrossRef]
67. Anderson, J. A provisional history of spectral music. *Contemp. Music. Rev.* **2000**, *9*, 7–22 [CrossRef]
68. Berlin, B.; Kay, P. *Basic Color Terms: Their Universality and Evolution*; California University Press: Berkeley, CA, USA, 1999.