FORMS: GENERATIVE VISUAL MUSIC THROUGH GRAPHIC NOTATION AND SPECTROGRAM SYNTHESIS

Santiago Vilanova Ángeles
Independent Researcher
santi@playmodes.com

ABSTRACT

This paper presents an in-depth analysis of FORMS, an extensive generative visual music research project that blends graphic notation with spectrogram synthesis. Rooted in the principles of generative music, FORMS encompasses a broad spectrum of digital art projects, unified by a unique software framework. The research bifurcates into two pivotal domains: spectrogram synthesis via digital image sonification and graphic notation tailored for musical instruments. This comprehensive study sheds light on the transformative capabilities of visual-to-auditory conversions in contemporary generative music, providing novel perspectives on animated algorithmic notation systems and some of their possible artistic applications.

1. INTRODUCTION

Embarking on a journey through generative [1] graphic notation and spectrogram synthesis, FORMS fuses digital art with contemporary music practices. Incepted amidst the 2020 pandemic lockdown, this initiative has since birthed a number of eclectic projects, ranging from online streaming bots to audiovisual installations or string quartet performances. Central to these projects is a bespoke software system, powering the expressive and aesthetic potential of algorithmic Visual Music [2] notation systems.

This article narrows its focus on the intricacies of the FORMS software systems and their diverse artistic manifestations. Broadly, the tools and their resultant projects can be categorized into:

- Spectrogram synthesis and image sonification. [3]
- Graphic notation tailored for instruments.

A unifying theme across these categories is generativity, with each FORMS project incorporating elements of randomness and probability to achieve unpredictable musical compositions.

This article also presents several case studies, comprising visual and musical compositions by the author, to illustrate the discussed concepts and techniques.

The research showcased in this article spans four years and encompasses the development of mathematical techniques, software algorithms and tools, audiovisual installations and instruments or contemporary music concerts. Given the breadth and depth of this work, each aspect could warrant its own detailed article. In fact, the author is currently working on a PhD thesis to comprehensively address all facets of this research. However, this article aims to provide a concise overview of the project and its recent artistic achievements.

Many of the software developments that are described in the article have been granted open access through online repositories, which will be provided punctually.

Along the text, you will also find a number of footnotes which point to website links showcasing video documentation of the described elements or artworks. We encourage the consultation of these links for a proper understanding of the research presented here, given its animated nature.

2. SPECTROGRAM SYNTHESIS AND IMAGE SONIFICATION

While traditionally a spectrogram serves as a representation of pre-existing sound, within the realm of FORMS it is the genesis of music. This paradigm shift involves first crafting the visual representation of sound, followed by the application of image sonification and sound synthesis techniques.

In order to effectively create spectrograms that 'sound good' once sonified, we first defined a number of rules for the image generation and its subsequent sonification, which establish the basic axioms to take into account:

- The X-axis denotes time, *sonogram units* positioned along it to express musical rhythms. The resulting images can either scroll horizontally towards a static play header or be traversed by the header itself.
- The Y-axis represents pitch, where each pixel row aligns with a microtone within the overall pitch range, or *tessitura*. A higher position of a pixel row indicates a higher corresponding pitch. Given that our measurement is based on pitch rather than frequency, the generated images conform to the principles of Mel-spectrograms [4].
- Each pixel row is associated with a sinusoidal sound wave. The frequency of this wave is determined by the tessitura and the total number of microtones,

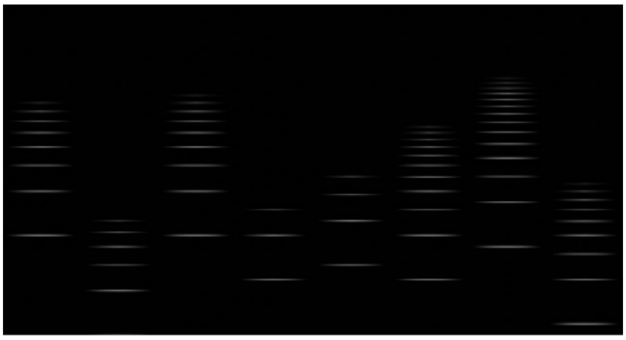

Figure 1. An example of partial clustering, to design bass sounds.

which is equivalent to the height of the sonified image in pixels. For instance, if an image stands at 120 pixels in height and covers a tessitura of one octave (12 semitones), this results in 10 microtones per semitone, calculated as 120/12.

- Luminance determines volume, with darker shades indicating lower volumes and brighter shades denoting higher volumes.

Considering these guidelines, we developed a comprehensive collection of procedural sonogram units, as we will see in section 2.2.2. Drawing upon Fourier Transform principles, we can synthesize intricate harmonic content by layering partials and adjusting their amplitudes through assigned luminance variations (Fig. 1). This approach enables the precise visual representation and synthesis of any sound by detailing its partials on the Mel-spectrogram.

This image sonification technique can be perceived as a variant of additive or spectral synthesis [5], where an array of sinusoidal waves, equivalent to the image's pixel height, plays concurrently guided by a play header that traverses the image. The luminance data from the analyzed pixel column serves as the amplitude bins for the sinusoidal wave bank, re-scaling the 0–255 luminance range to a sound linear amplitude range between 0 and 1. In turn, the frequency bins are ascertained by each pixel row's index after performing a pitch-to-frequency conversion. Please see section 2.2.2 for an in-depth review of FORMS image sonification methods.

In regards to time sequencing, and using the grid depicted in Fig. 2 as a foundational layout, we place the aforementioned sonogram units on the canvas. By incorporating inherent randomness into generative music, the system determines the presence of specific sonogram units (e.g., a kick drum, a bass note, a bell tone) at particular positions on the time axis, thanks to a series of generative sequencing algorithms that have been also developed and integrated into the software system. This same generative approach has been implemented for the pitch height of sound objects, giving the software system a certain range of freedom after the selection of a musical scale. Please see section 2.1 for details on the implemented generative music techniques.

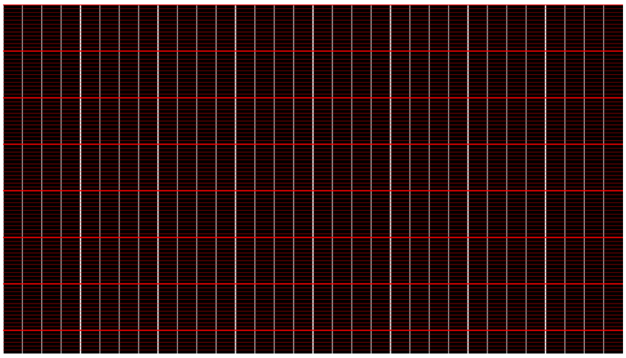

Figure 2. The sonification grid. Horizontal red lines denote steps on the chromatic scales, the thicker ones being the note C of successive octaves. White vertical lines denote quarter notes, the thicker ones being the start of a bar.

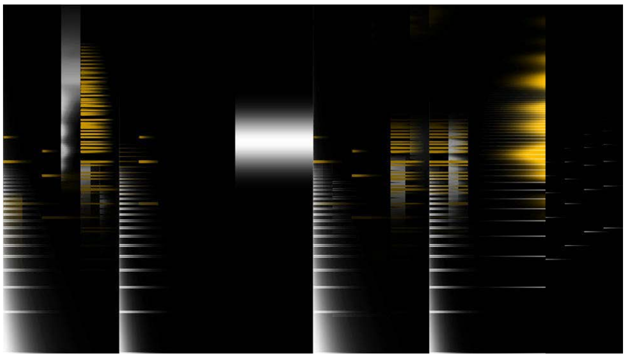

Figure 3. Generatively synthesized spectrogram showcasing varied percussive elements. Observe the multi-partial bass aligned with the kickdrum and the distinct melodic arpeggio in the final bar.

All in all, this approach crafts a realm of potential resulting images [1], as seen in Fig. 3.

2.1 Generative Music Algorithms

At the heart of the spectrogram synthesis system lies a number of principles related to musical generativity. This is understood as the capacity for automatic music generation through rule-based processes, controlled randomness and other algorithm-based automated procedures. Our objective has consistently been to ensure that every software system within FORMS incorporates algorithms that enable it to operate as a musical automaton. By interpreting user-defined rules, the system can produce endless variations within the confines of these guidelines. To achieve this, we've explored and integrated various generative music techniques, laying the foundation for this automated music generator. Below, we delve into the most pivotal of these techniques:

2.1.1 Euclidean Rhythms

Euclidean rhythms [6] are a unique approach to rhythm generation based on the Euclidean algorithm, which deter-

[1] the following video playlist displays a collection of FORMS spectrogram synthesis results: https://shorturl.at/JOQZ2

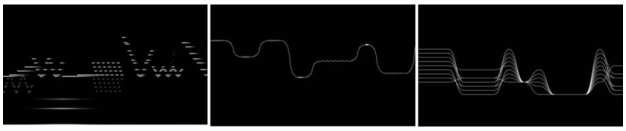

Figure 4. Examples of harmonic sequencing, featuring arpeggiation, root progressions and chords.

mines the greatest common divisor of two numbers. By distributing a set number of beats as evenly as possible over a rhythmic cycle, intriguing and musically relevant patterns emerge. These rhythms have been identified in various traditional music cultures worldwide, making them a versatile tool for generating diverse rhythmic structures in contemporary compositions.

2.1.2 Probabilistic Sequencing

Probabilistic sequencing introduces an element of chance into the sequencing process. Instead of following a fixed pattern, each step in the sequence has a probability associated with it, determining its likelihood of being played. This approach allows for dynamic and ever-changing musical patterns, ensuring that compositions remain fresh and unpredictable, yet within defined boundaries set by the composer.

2.1.3 Generative Harmony

Various generative harmony techniques have been incorporated. Users have the flexibility to select a musical scale for each synthesized spectrogram from a diverse list, which spans from ancient Greek modes to contemporary scales and world-music modes. Based on this selection, random chord progressions are generated by stacking notes at random intervals. While there's room for further refinement, such as the incorporation of functional harmony, serial and tone-row tonal systems – just to name a few –, the modal harmonies produced by this approach give a first taste of pitch organization within the FORMS system. Additionally, a user-defined weighted chance algorithm determines the likelihood of the composition shifting its root note to any given degree or undergoing a specific semitone transposition. This means that each scale degree and every chromatic step have individually assignable probabilities. The culmination of these features produces spectrograms with dynamic and scale-quantized harmonic progressions (Fig. 4).

2.1.4 Mute/Solo Chance

Considering that compositions frequently consist of multiple layers (such as percussive elements, melodies, chords, arpeggios, bass or textures), an automated mute/solo feature has been integrated. This functionality either emphasizes specific layers (solo) or suppresses them (mute), adding rhythmic depth to multi-layered compositions. Importantly, the mute/solo algorithms are cognizant of time positioning, allowing the probability of a mute/solo event to be influenced by the composition's temporal context. For instance, a solo might be more likely during the final bar of an eight-bar sequence.

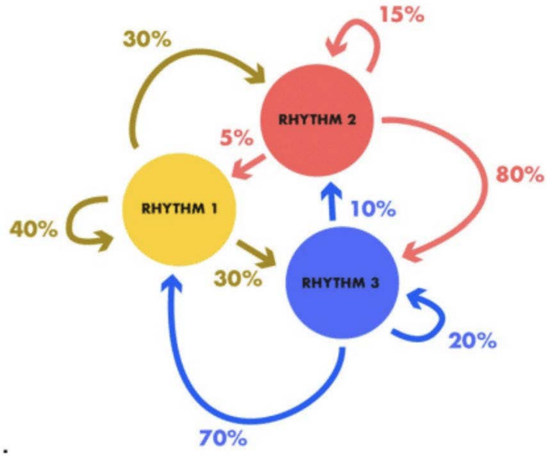

Figure 5. An example of 1st order Markov Chains application for rhythmic transitions.

2.1.5 Algorithmic Arpeggios

Algorithmic arpeggios utilize mathematical formulas to craft arpeggiated sequences, deconstructing chords into a series of notes played consecutively. The system boasts a parametric approach to arpeggiation, encompassing variables like pattern length, direction, polyphony, tempo, and interval randomness. By assigning random values to these parameters, the system can generate a limitless array of dynamic and evolving arpeggios (Fig. 4).

2.1.6 First Order Markov Chains

Markov Chains serve as a potent instrument for intelligent sequencing. Primarily used for navigating between presets (where a preset refers to a specific set of parameters defining a spectrogram), Markov Chains facilitate the determination of transition probabilities from one state to another in a structured manner (Fig. 5). Within a musical framework, this mechanism proves invaluable for guiding chord transitions, tonal shifts, and stylistic evolution.

2.1.7 Noise Functions

Noise functions, commonly utilized in computer graphics, have found applications in musical domains as well. By harnessing algorithms such as Perlin noise [7], we generate a sequence of values that produce organic and evolving control signals. These signals can be channeled to modulate various parameters, such as the volume of individual instrumental layers, leading to a natural interplay of voices. They can also influence the rate of LFOs that adjust volume or the number of sound partials, or even the tempo of an arpeggio. Consequently, compositions acquire a sense of organic fluidity and unpredictability, enhancing their overall appeal.

2.2 Software Implementation

The applications underpinning spectrogram synthesis and sonification are developed across multiple coding platforms, interconnected through the OSC network protocol.

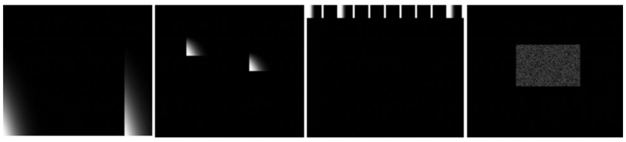

Figure 6. Examples of percussive sonogram units. Kick-drum, Snaredrum, Hihat, Noise.

Although the software architecture remains in a prototypical phase, with aspirations towards a unified application handling all processes, its current iteration has demonstrated robust functionality. This makes it suitable for practical applications, such as audiovisual installations and live concerts. The entire software infrastructure, as illustrated in Fig. 8 and Fig. 7, is optimized for real-time execution, establishing it as a tool for live generative visual music performances.

2.2.1 Image Synthesis with Processing

Image synthesis is executed using Processing [8], an open-source Java-based framework tailored for procedural graphic design. We developed multiple functions that define diverse bitmap graphics generation algorithms. As we explained in section 2.1, all these functions integrate a degree of controlled randomness, ensuring the outcomes are never identical, aiming for a generative music system.

At the center of the spectrogram synthesis methodology lies the concept of the *sonogram unit*. Various functions have been developed to define these visual entities. Each function can generate a segment of a spectrographic image, which can then be strategically positioned on the canvas, aligned to a time or pitch grid, or overlaid with other sonogram units through additive or alpha visual blending techniques. A plethora of these sonogram unit functions have been crafted, encompassing elements such as kick drums, snare drums, hi-hats, bass, bells, and pads. Typically, these functions can be customized with parameters like x and y coordinates, duration, or specific attributes tailored to each object, such as the number of partials for elements that adhere to the harmonic series. Leveraging the generative sequencing algorithms outlined in section 2.1, these sonogram units are rhythmically and tonally arranged on the canvas, leading to a boundless array of musical expressions.

Typically, the final output from this Processing application is a FullHD image, 1920x1080 pixels, saved as a .png file. The time required for this software system to synthesize a new image upon request varies between 1/2 to 5 seconds, contingent on the graphic algorithms' complexity and the available CPU.

2.2.2 Image Sonification with Max/MSP

Image sonification is managed by a custom Max/MSP [9] patch. This patch retrieves the path of the newly generated image from Processing, loads the image into memory, and performs a luminance analysis on a specific column of pixels, referred to as the *header*. The header can either move from the image's left to right at a set BPM or remain fixed while the image scrolls from right to left. In both scenarios, the movement of the header or the image is driven by an audio rate phasor, quantized to pixel column positions.

The results of the header's luminance analysis are used to adjust the amplitudes of a predetermined frequency bank of sinusoidal oscillators. The frequency values of these oscillators are based on the user-defined tessitura, while the bank's size (and the number of microtones within the tessitura) corresponds to the image's pixel height. As the header progresses across the image, the amplitudes of the oscillator bank evolve, producing a sonic representation that mirrors the visual content of the image. Following a simple audio mastering process that includes limiting and subtle reverb adjustments, the resulting sound is channeled to the computer's audio output. From there, it can be further manipulated on external mixers or directly relayed to speaker systems.

2.2.3 Control Architecture in Max/MSP

Also acting as a conductor, the Max/MSP patch requests Processing for new images, transmitting an OSC data bundle with details about the graphic methods to be employed, the image's bar count, the musical scale to be utilized, among other characteristics. The user-friendly GUI design enabled by Max/MSP has facilitated the creation of an interface where users can precisely define the details of the spectrograms to be synthesized. A preset system simplifies the storage of this features. Each preset will consequently produce an infinite collection of images (owing to the integrated randomness in the Processing graphic algorithms), which, once sonified, will give rise to varied musical streams, from raging rhythms to pleasant harmonic ambient or complex melodic sequences.

2.2.4 Image Scroller Coded in C++

The final software component in this suite is the *Scroller*. Developed in C++, this straightforward application manages the scrolling of the .png format spectrogram images generated by Processing, aligning with the speed and phase set by the audio rate phasors in Max/MSP. Configurable for full-screen display, the Scroller is particularly advantageous for public concerts or installations due to its efficient GPU-based operation.

2.2.5 Open Source Software Repository

All the developments described previously on this section have been made public on a Github repository [2] .

2.3 Case Study: FORMS – Screen Ensemble

Forms – Screen Ensemble [3] (Fig. 9) made its debut at the 2020 edition of the Ars Electronica Garden [10] in Barcelona, courtesy of a grant from New Art Foundation [11], Institut Ramón Llull [12], and Hangar.org [13]. It is conceptualized as a generative visual music jukebox triptych. Governed by chance and probability, this automaton fabricates endless, unique spectrograms which are instantaneously converted into sound via sonification algorithms,

[2] https://github.com/PlaymodesStudio/FORMS
[3] https://www.playmodes.com/home/forms-screen-ensemble

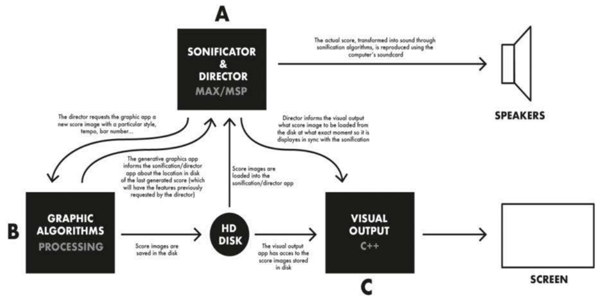

Figure 7. A diagram showing the data flow between the different software applications that conform the spectrogram synthesis system.

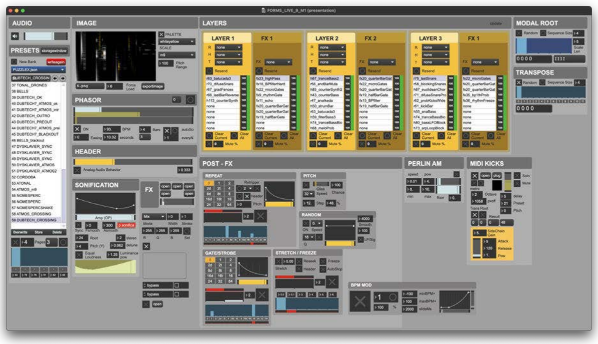

Figure 8. The FORMS Max/MSP patch, showing the GUI with the parameters that control the spectrogram synthesis.

enabling listeners to audibly perceive what they visually observe. Each screen within this interconnected ensemble assumes a specific instrumental role, be it Rhythm, Harmony, or Texture. Inside each of these 3 screen boxes, there is a FullHD monitor, a soundbar speaker, a computer and an Ethernet switch which interconnects the 3 computers into a single network (Fig. 10). Orchestrated by this trio of automatons, a visual music symphony continually evolves, giving rise to singular sonic landscapes that will never be replicated: from tonal ambient music to intense rhythms, surreal electronic transitions, or dance-floor beats.

Extensive mathematical research has been undertaken to code timbrically and musically coherent graphics that generate captivating rhythms, harmonies, or textures. Additionally, a suite of sequence control routines governs the compositions' evolution over time, leveraging weighted probabilities, Markov chains, and chance algorithms. Its generative nature yields endless, unique, non-repetitive, and rich music, liberated from the confines of conventional written musical language.

2.4 Case Study: Sonògraf

The *Sonògraf* [4] is conceptualized as an electronic audiovisual instrument. Designed as a musical learning tool for primary schools, it facilitates the transformation of

[4] https://www.playmodes.com/home/sonograf

Figure 9. A photo of the Screen Ensemble as exhibited in Arts Santa Mònica during 2020's Ars Electronica Garden.

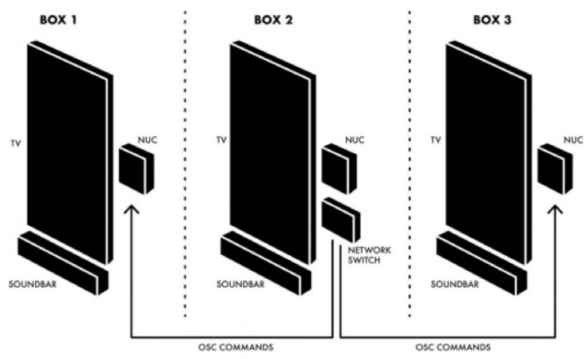

Figure 10. A diagram showing the different devices that compose the Screen Ensemble piece, and its network connections.

handmade drawings or object collages into music, converting gestural strokes and geometric figures into electronic sounds. A collection of buttons and potentiometers enables live manipulation of the drawing's sonification characteristics, allowing users to accelerate, decelerate, or pause the resultant music and also choose its scales and tonalities. The Sonògraf is equipped with speakers and a compact video projector, enabling both intimate practice and staging in an audiovisual concert for a diverse audience. The conception, engineering, and design of the Sonògraf have been meticulously crafted to ensure intuitive play for all, offering a novel paradigm for understanding musical notation and deviating from the academic nature of traditional notation based on staves.

The Sonògraf was awarded the "People's Choice Award" and a jury honorable mention at the 2024 edition of the Guthman Musical Instrument Competition [5] held by Georgia Tech University, in Atlanta. A performance with the instrument, and in collaboration with clarinet player Lauren McCall, can be seen in a recording published by Georgia Tech. [6]

[5] https://guthman.gatech.edu

[6] https://www.youtube.com/watch?v=j4gHsl5CQRs

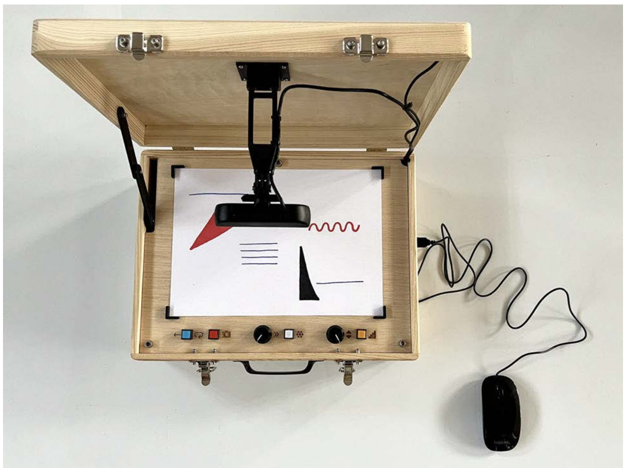

Figure 11. The “Sonògraf”, an educational tool for primary schools oriented towards music learning through hand-drawn, automatically sonified, graphic scores.

Finally, a Sonògraf DIY guideline has also been made public – using a Creative Commons license – on a Github repository. [7]

3. GENERATIVE GRAPHIC NOTATION FOR INSTRUMENTAL INTERPRETERS

Diverging from the spectrogram synthesis techniques, the FORMS initiative also delves into generative graphic notation. The primary distinction between these methodologies is that while spectrogram synthesis is predominantly employed to craft digital electronic music, the generative graphic notation techniques facilitate the creation of simple and succinct notation for human instrumentalists.

The generative notation approach within the FORMS series aims for an integration with the results of the spectrogram synthesis. The aim is to amalgamate both methodologies into multi-layered musical compositions, blending live instrumental music with sonified synthetic spectrograms. In anticipation of this integration, the rules governing the generation of graphic notation closely mirror those of spectrogram synthesis: Time is represented on the X-axis of the image, while Pitch is encoded on the Y-axis. However, a fundamental deviation from the spectrogram synthesis technique is that the volume or dynamics of the composition are not represented using luminance but through shape weights12. Additionally, to ensure clarity for the interpreters, when tonal precision is mandated in the composition, an annotated text preceding each musical object indicates the exact pitch to be played. A color code is also employed to distinguish different instruments, allowing each instrumentalist to identify their part without necessitating individual scores.

3.1 Software Implementation

The creation of the graphic notation scores is executed exclusively using Processing (Fig. 13). As the image sonification is performed by human interpreters, a software algorithm for image-to-sound synthesis is redundant. The approach for this set of graphic algorithms diverges from the previous spectrogram synthesis technique in that the generated graphics are vector-based, as opposed to bitmaps. Concurrently, each musical piece to be composed is treated as a unique project. While randomness and probability remain integral, each musical passage is meticulously programmed and its final visual output is curated from the multitude of musical/visual outcomes produced by the integrated randomness factors. The final selected .svg files are subsequently edited using a vector graphics editing software (e.g., Adobe Illustrator or Inkscape), ensuring macro-structural coherence for the entire composition. These files are then animated using an animation or post-production suite (e.g., Adobe After Effects).

Currently, efforts are underway to develop a new software system for macro-structured real-time generation of graphic notation, tailored for adept human interpreters. This new set of software tools would free us from the need of manual curation of results and its subsequent manual edition and animation. It is anticipated that in the upcoming months, a series of concerts will be made possible with these novel tools.

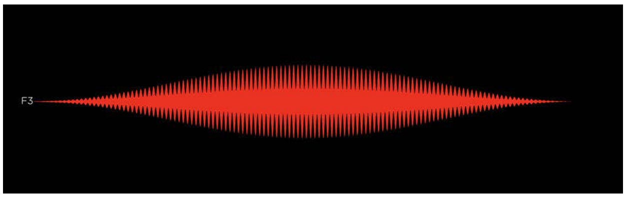

Figure 12. Dynamic volume contours are defined through the weight of the shapes. Note the accurate dynamic progression from ppp>ff>ppp and the fast tremolo contour. For tonal precision, the English notation is added at the beginning of the shape.

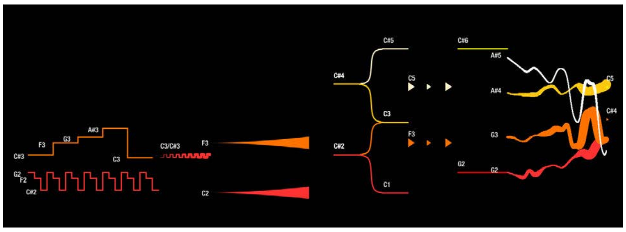

Figure 13. A screen capture of the Processing generative scoring software for string quartet.

3.2 Case Study: FORMS – String Quartet

A collaboration with Quartet Brossa [14], was showcased at Barcelona’s CosmoCaixa Science Museum [15] in April 2021, as part of the art+science NEO cycle [16], curated by Irma Vilà. The project received support from the Department of Culture of the Generalitat de Catalunya [17].

FORMS – String Quartet [8] (Fig. 16) is a live multimedia concert designed for a string quartet, electronic music, and panoramic visuals, rooted in the concept and techniques of

[7] https://github.com/PlaymodesStudio/Sonograf

[8] https://www.playmodes.com/home/forms-string-quartet

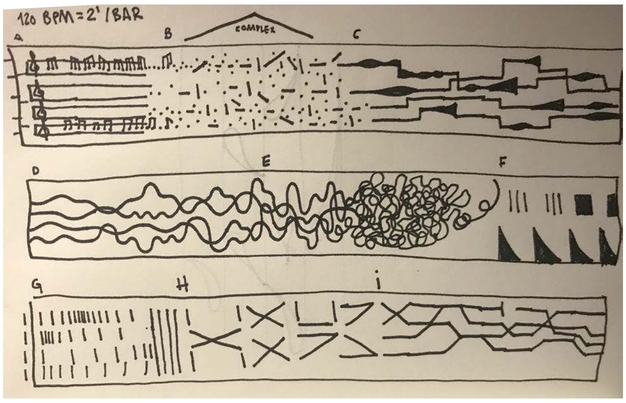

Figure 14. A preparatory hand-drawn sketch for the FORMS – String Quartet concert.

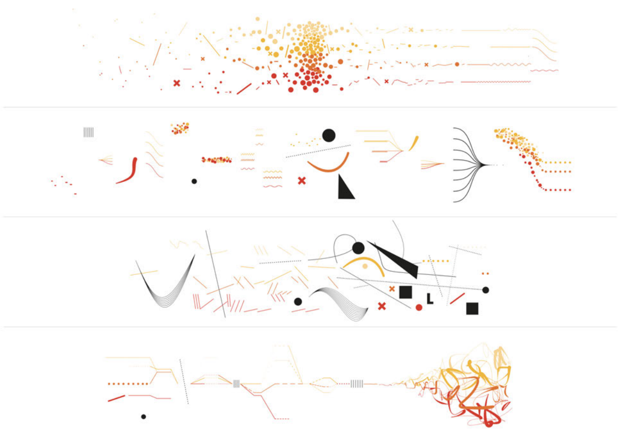

Figure 15. The final score of the FORMS – String Quartet's first movement. Each of the 4 warm colors is the score for the quartet: cream, first violin; yellow, second violin; orange, viola; red, cello. Black is electronic music, to be sonified using the methods discussed at 2.2.

image sonification. Musicians interpret a series of graphic scores, which concurrently construct the visual scenery, offering the audience an immersive experience where they can anticipate the 'music to come'. These scores, in conjunction with the generative spectrograms that form the electronic accompaniment of the piece, have been crafted through the use of the software tools seen in section 2.2. The composition process for this concert diverged from traditional methods. Following the initial commission, a decision was made to assign distinct colors to each instrument's score part so that all four melodic lines of the string quartet could coexist on a single drawing. This idea was first visualized through hand-drawn preparatory sketches, as shown in Fig. 14. These sketches were subsequently modeled algorithmically in Processing, as depicted in Fig. 13, with added randomness to generate varied outcomes for each section. The scores, produced by Processing in .svg format, were curated and served as the foundation for the macro-structural composition, developed using standard vector graphic editing software. The final vector score (Fig. 15) was animated using video editing software, ans resulting in a panoramic format video. Two versions

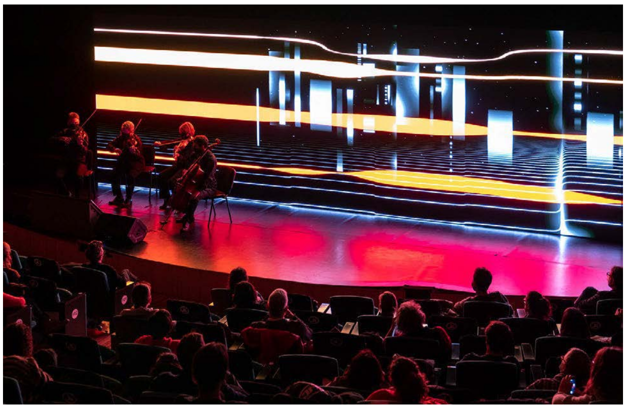

Figure 16. A photography of the 2021 *FORMS – String Quartet* premiere concert at the "NEO" art and science cycle in Cosmocaixa, Barcelona's science museum.

of this animated score were produced: a high-resolution panoramic version for concert visuals, and a 16:9 full HD version for musicians' reference. It's noteworthy that the musicians' version employed a distinct color scheme for clarity. While the scenic video displayed instruments in four shades of red, the musicians' version utilized a varied palette: red for the cello, blue for the viola, green for the second violin, and yellow for the first violin.

Alongside the string quartet score, a layer of visual electronic music was synthesized. This synthesis was based on the methods of spectrogram synthesis detailed in section 2 and was synchronized with the string quartet components.

Instead of traditional paper scores on a lectern, musicians refer to a 45" TV screen during rehearsals and concerts. This screen displays the animated score in synchronization with the panoramic scenic version.

The concert, which has a duration of 25 minutes, explores the expressive power of generative graphic notation, and deploys an extensive research on instrumental techniques – and its representation – from the classical legato or pizzicato, to extended techniques such as col·legnos, glissandi or Bartok pizzicati.

Since its premiere in 2021, this concert has been represented dozens of times in auditoriums and contemporary music festivals, and it has been widely acclaimed by audiences and specialized press.

4. "MIDIFICATION" OF GENERATIVE GRAPHICS

Within the FORMS framework, another avenue of exploration is the conversion of images to MIDI. The scoring rules mirror those of the spectrogram synthesis and graphic notation techniques (X-axis represents time, Y-axis denotes pitch, with luminance serving as the metric for volume dynamics. Review sections 1 and 3). Given the event-based nature of the older MIDI protocol [18], where MIDI velocity remains static during the playback of a note, luminance gradients cannot be employed to continuously modify volume expressions. Thus, dynamics are fixed at the onset of each note. Although newer iterations of the MIDI

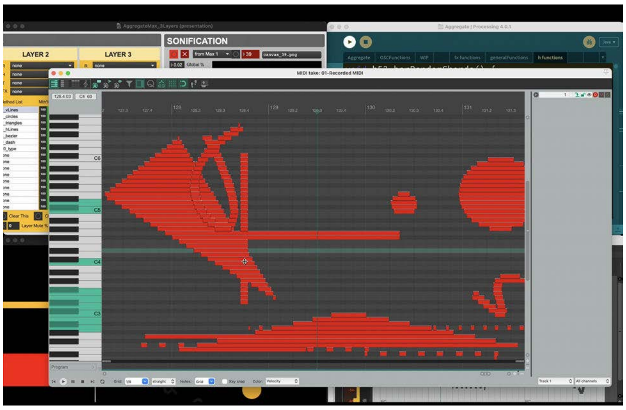

Figure 17. A screen capture of the resulting MIDI data from the sonification of generative graphics.

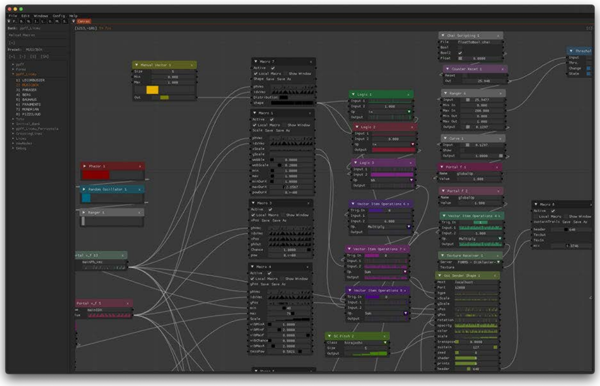

Figure 18. *Oceanode,* the nodal development environment software developed by Playmodes, which controls the graphic score generation of the *ppff* installation.

protocol (MIDI 2.0 and MPE) permit such continuous dynamic control along the duration of a note, the majority of existing MIDI instruments have yet to adopt these features.

4.1 Software Implementation

The MIDI-centric generative visual scoring tools are the latest additions to the FORMS suite. This set comprises two applications. Firstly, a nodal programming environment developed in Open Frameworks/C++ [19], responsible for defining shape types (e.g., circles, triangles, rectangles) and their XY positions, sizes, luminance, etc., for a maximum of 4096 shapes (Fig. 18). This nodal environment, called Oceanode, can be found as an Open-source tool at Playmodes' GitHub repository [20]. Secondly, a custom Processing/Java application processes the OSC data bundles from the C++ nodal environment, rendering this data as a visual output in real-time. Processing also oversees the analysis of luminance crossing the read header, converting this data into MIDI events.

This novel, streamlined approach has proven to be highly efficient and adaptable, setting a new benchmark for the future of the FORMS graphic generation software.

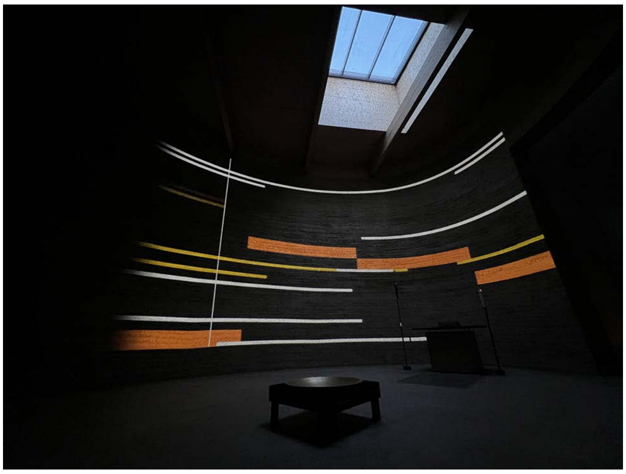

Figure 19. An image of the *Clash/Blend* installation at the Reconciliation Chapel in Berlin, 2022. The scores scroll towards the organ position and are sonified once the graphics cross the vertical header.

4.2 Case Study: Clash/Blend

Clash/Blend[9] (Fig. 19) is a visual music installation that commandeers the pipe organ at the *Kapelle der Versöhnung* in Berlin. Featured in the 2022 edition of the *Aggregate – New Works for Automated Pipe Organs* [21] festival program, this installation aligns with the Color Organ tradition. It deploys a video-projected visual canvas on the chapel walls. A cascade of geometries scrolls towards the Chapel's pipe organ, undergoing a transformation into MIDI (see Fig. 17) as these drawings intersect with the organ's position. Each color on the score is routed to a different MIDI channel, activating different sound registers per color. Owing to its generative nature, these graphic scores, and the sounds they produce, are unique and non-repetitive. This installation was developed using Processing (for real-time image generation) and Max/MSP (for MIDI translation of images), as taught in section 4.

in 2023 *Clash/Blend* was awarded a Golden LAUS [22] design prize, the most prestigious recognition in Spanish design.

4.3 Case Study: FORMS – ppff

Premiered on July 2023 at Barcelona's *Liceu* Opera [23], *ppff*[10] – an acronym for the musical notation *pianissimo-fortissimo* – is a generative visual music installation. The centerpiece of this installation is a *Steinway and Sons Spirio* player piano, a MIDI-automated grand piano capable of striking its hammers in response to real-time music data. This piano derives its musical input from an unconventional source: a continuous cascade of generative graphics (Fig. 20), showcased on a vertical LED screen. As these graphics reach the piano's position, they are transmuted into MIDI notes, facilitated by Oceanode, the aforementioned set of custom software developed in C++ (Open Frameworks), and Java (Processing) (See section 4).

[9] https://www.playmodes.com/home/clash-blend
[10] https://www.playmodes.com/home/ppff

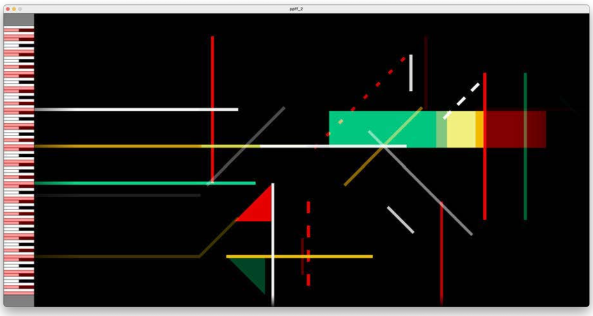

Figure 20. an example of the resulting graphic scores of ppff. Note the keyboard diagram at the left, which shows the positioning of the image-to-midi conversion.

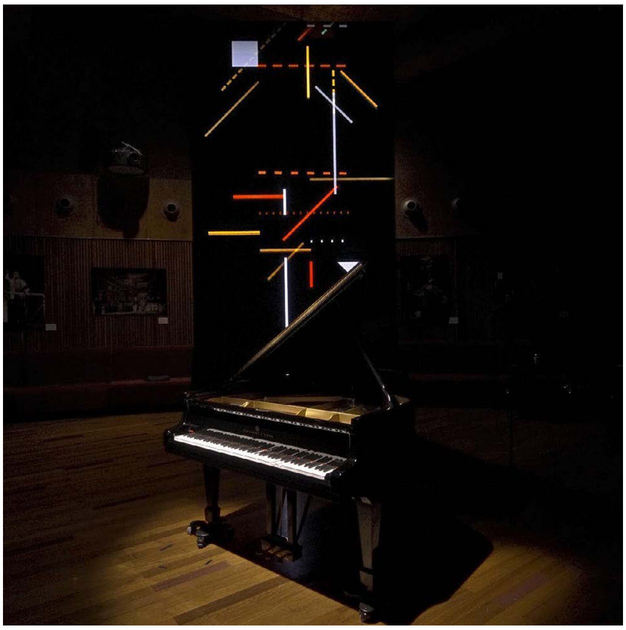

Figure 21. An image of the *ppff* setup in Barcelona's Liceu Opera. A vertical LED screen displays the downwards scrolling scores. When the graphics arrive at the piano position they are converted into MIDI to control a Steinway&Sons *Spirio* grand piano.

Musically, this solo piano piece comprises seven generative behaviors or miniatures. Each miniature delves into a distinct expressive, conceptual, or harmonic idea, ranging from atonal and indeterminate music to modal or rhythmic compositions which refer to and get inspiration from composers like Morton Feldman, György Ligeti or Arvo Pärt. All miniatures are heavily influenced by randomness, chance, and probabilities, ensuring that these musical pieces will always remain unique.

ppff remains a dynamic work-in-progress. Having been presented only once, it is set to undergo further evolution in the upcoming years. This progression will improve the image-to-MIDI methodologies and allow further research into the expansive potential of visual musical expressions inherent to such systems.

5. CONCLUSION

The FORMS initiative represents an advancement in the integration of contemporary generative music and digital art, particularly through its exploration of procedural graphic notation and Mel-spectrogram synthesis. Originating during the unique challenges of the 2020 pandemic lockdown, FORMS has demonstrated the adaptability and potential of software tools and algorithms in the realm of visual-to-auditory conversions.

This paper has provided a comprehensive examination of the various components and applications within the FORMS framework. From the foundational principles of image sonification to the diverse projects it has given rise, it is evident that FORMS seeks to address both the technical and artistic challenges inherent in this interdisciplinary field.

As the FORMS project progresses, it is anticipated to contribute further to the ongoing dialogue between visual art and music. Its continued development and refinement will likely offer valuable insights and methodologies for artists, researchers, and technologists aiming for a more cohesive and nuanced understanding of visual and auditory artistic expressions.

6. REFERENCES

[1] M. A. Boden and E. A. Edmonds, "What is generative art?" *Digital Creativity*, vol. 20, pp. 21–46, 2009. [Online]. Available: http://dx.doi.org/10.1080/14626260902867915

[2] K. Brougher *et al.*, *Visual Music: Synesthesia in Art and Music Since 1900*. London: Thames and Hudson, 2005.

[3] C. Roads and J. Strawn, *Graphic Sound Synthesis*. MIT press, 2003, vol. 32, no. 6.

[4] S. S. Stevens, J. Volkmann, and E. B. Newman, "A scale for the measurement of the psychological magnitude pitch," *Journal of the Acoustical Society of America*, vol. 8, no. 3, pp. 185–190, 1937.

[5] C. Roads, *The Computer Music Tutorial*. MIT Press, 1996, see sections on additive synthesis.

[6] G. T. Toussaint, "The euclidean algorithm generates traditional musical rhythms," in *Proceedings of BRIDGES: Mathematical Connections in Art, Music, and Science*, 2005. [Online]. Available: http://cgm.cs.mcgill.ca/~godfried/publications/banff.pdf

[7] K. Perlin, "An image synthesizer," in *ACM SIGGRAPH Computer Graphics*, vol. 19, no. 3. ACM, 1985, pp. 287–296. [Online]. Available: https://dl.acm.org/doi/10.1145/325165.325247

[8] C. Reas and B. Fry. Processing webpage. [Online]. Available: https://processing.org

[9] Max-msp. [Online]. Available: https://cycling74.com/products/max

[10] Ars electronica festival. [Online]. Available: https://ars.electronica.art

[11] New art foundation. [Online]. Available: https://www.newartfoundation.art

[12] Institut Ramon Llull – Llengua i cultura catalanes. [Online]. Available: https://www.llull.cat/catala/home/index.cfm

[13] Hangar – Centre de producció i recerca d'arts visuals. [Online]. Available: https://hangar.org/ca

[14] Brossa Quartet de Corda. [Online]. Available: http://www.brossaquartet.com

[15] CosmoCaixa – Museu de la Ciència de Barcelona. [Online]. Available: https://cosmocaixa.org

[16] NEO – CosmoCaixa. [Online]. Available: https://cosmocaixa.org/ca/neo

[17] Oficina de Suport a la Iniciativa Cultural – Generalitat de Catalunya. [Online]. Available: https://cultura.gencat.cat/ca/tramits/oficina-de-suport-a-la-iniciativa-cultural/index.html

[18] Specifications – MIDI Association. [Online]. Available: https://midi.org/specifications

[19] openFrameworks. [Online]. Available: https://openframeworks.cc

[20] PlaymodesStudio, "ofxOceanode," https://github.com/PlaymodesStudio/ofxOceanode, 2023.

[21] Aggregate – Gamut Inc. [Online]. Available: https://gamutinc.org/aggregate

[22] Premiados – Laus Awards - ADG-FAD. [Online]. Available: https://www.adg-fad.org/es/laus/premiados

[23] Gran Teatre del Liceu – Barcelona. [Online]. Available: https://www.liceubarcelona.ca