複雜系統與聲響美學：以微粒合成法產生多體即時互動聲響
Complex Systems and Sound Aesthetics: Granular Synthesis and Many-Body Induced Sound

郝從恩 Chung-En Hao [1]、劉辰岫 Ivan Chen-Hsiu Liu [2]
國立陽明交通大學應用藝術研究所 研究生 [1]、國立陽明交通大學應用藝術研究所 助理教授 [2]

摘 要

近年來互動科技的興起為科學與藝術的跨領域融合開創了更多的可能性，本計畫以互動裝置與數位聲響科技作為媒材，透過藝術手法探討自然界複雜系統（Complex System）現象以及複雜系統美學的湧現，如多體振盪及突崩效應，並建構出一套適用於此類型研究及創作的互動聲響科技。在聲響技術上，我們利用聲音微粒合成法（Granular Synthesis），把聲音樣本在時間軸上切分成數個小段，而每段聲音皆以點狀微粒來表示，並依照每個段聲音的頻率、音量及其他特徵，各別分佈在二維描述符座標（Descriptor Space）。我們將複雜現象中，多體的位移投射在此座標上，而任何被動態位移所碰觸到的微粒，皆會即時地發出該聲音段落，使動態位移產生即時且連續性的聲音反饋。也讓原先完整被預錄的聲音，經由拆解過後產生出新的樣貌。參與者在操作過程中可感受到視覺與聽覺共同發生的連鎖效應，甚至進一步探討都市文明噪音與自然間互動的可能性。

關鍵詞: 複雜系統、聲音藝術、人機互動、科技藝術、微粒合成法

Abstract

The development of interactive design and technology in recent years has opened up new possibilities for the cross-disciplinary fusion of science and art. This project investigate the phenomena in complex systems and the emergence of aesthetics from such systems. In particular, we aim to develop an interactive sound installation based to simulate the earthquake model widely used in physics. We utilize the granular synthesis technique to divide sampled sound temporally into small “grains”, which are analyzed according to tis frequency, volume or other features. These grains are distributed onto a 2D descriptor space based on which two descriptors are chosen. The few-body movement of the earthquake model are mapped onto the descriptor space to trigger the sound grains instantaneously. The result is an overall continuous and realistic sound experience that opens up new possibility for story-telling.

Keywords: complex system, sound art, human computer interaction, tech-art, granular synthesis

一、 研究動機與目標

近年來互動科技的興起為科學與藝術的跨領域融合開創了更多的可能性，以數位科技應用操作人類感官，實際體驗抽象飄渺的科學現象，使之與人重新產生連結感，而不再只是書本中過於形式化的公式。其中體感科技與虛擬實境等穿戴裝置已是娛樂產業重點發展目標，期盼新這些興科技可以在藝術上開闢新的道路。

日常中所聽見的聲音皆由不同程度的震動所產生，而其中只有震動頻率介於 20～20,000 赫茲範圍的震動會傳入耳裡，形成聽覺感官，低於或高於這頻率區間的震動缺少了被「聽見」的功用，然而這些震動與波動接與聲音一樣，皆為「運動」。

能量的傳遞在生活中無處不在，從微觀的分子結構裡，到巨觀的地震現象；從不會發聲的，到會發聲的，皆為萬物在空間中運動的體現。人們習慣了區別聲音與運動，卻往往不會把兩者銜接為一整體的概念去理解。大氣中傳遞振動的空氣無法被肉眼所及，我們也就把從遠端傳來的聲音視為理所當然，領會不到空氣所扮演介質的角色。

本計畫嘗試製作一套可互動式模型，將不直觀的聲音傳導給具象化，同時也在互動中反饋出聲響。我們參考了地震模型相關文章（Brown et al, 1991; Olami et al, 1992），目標設計一套彈簧塊體力學模型（Spring-Block Model）（圖 1），以多個塊狀物（在本篇以「盒塊」來簡稱）作為與人進行互動的硬件，再將多個盒塊相互與彈簧伸縮結構連接，在平面上形成網狀結構，以模擬物質與物質之間受力與能量的傳導行為，連帶出多體震盪效應。我們以仿地震現象作為首要實驗目標，將多個盒子相互與彈簧連接為了讓體驗者在觸摸及搖晃或移動盒塊的過程中，同時觸發周圍盒子擺動。配置多個 IMU(Inertial Motion Unit)感測器置入盒塊裡，來使搖晃與移動產生力的變化，產生數據，將與聲音工程上微粒合成法（Granular Synthesis）搭配（Schwarz, 2011），形成多點聲音觸發系統，再結合成一套可以手動操作來模擬巨觀自然現象的美學實驗。此外也嘗試摸索出都市文明噪音與人體感官間產生新的可能性，也呼應自然界裡複雜系統對社會間微妙且深遠的影響。此篇論文為本計畫近期研究成果。

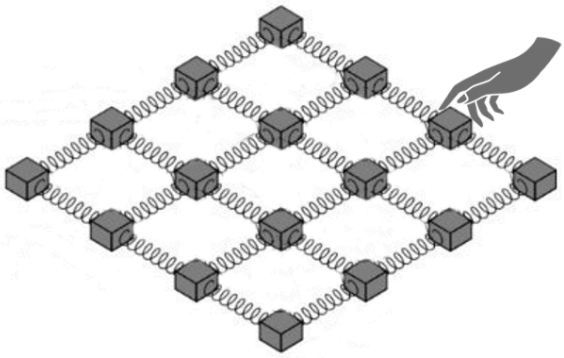

圖 1. 將多個盒塊以彈性結構相互連接

二、 微粒合成法與 CataRT 介紹

3-1 微粒合成法

一段完整的音擋經過拆分成數個長度約小於 1 秒鐘的小段子，每小段可被稱為「微粒」（如圖 2）。多個微粒可以相互疊加，並可以不同的速度、頻率以及其他參數播放（Barry, 1988）。在一般的電子音樂中，微粒合成法往往針對某些段子取樣循環播放，通過聲音處理技術，改變微粒的波形、時間、密度等，可生成許多不同的聲音，使原本聲音被賦予更刺激、動感的效果。若是擷取一段間隔低於 0.1 秒的聲音，以近乎零間隔的方式循環播放時，就會產生具有特定音色的聲音，聽者無法察覺背後是經由微粒合成的聲音段子。

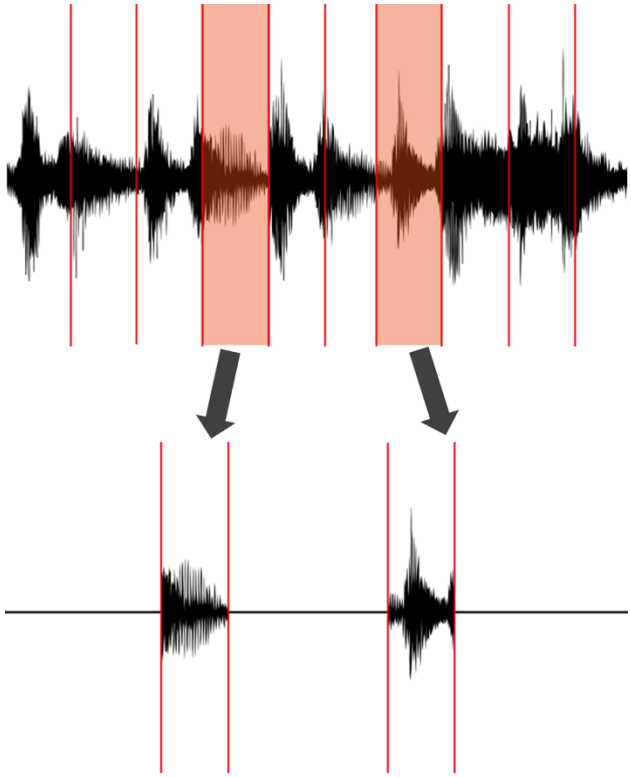

圖 2. 微粒合成法基本概念－將音檔拆分成數等份後再重組

3-2 CataRT 介紹

CataRT 是一套由法國媒體聲音研究中心開發的圖形化介面函式庫，其主要可套用在 Max/MSP 上運行（Schwarz et al, 2006）。CataRT 使用 Mubu 函式庫開發成的應用工具，主要將聲音檔案經過微粒合成法之處理後，系統將各別分析出每段微粒的特徵，諸如頻率、嘹亮程度、粗糙程度等等，針對每個特徵量給予相對應的數值。再自行挑選兩個特徵量來代表平面座標系統上的橫軸與縱軸，將所有聲音微粒置入座標系統中，也可稱為二維描述符座標（descriptor space）（流程如圖 3 所示）。程式自動偵測滑鼠所經過的所有座標上所對應的特徵值，並且會瞬間播放出與之相鄰的聲音微粒。若滑鼠在座標上連續的移動，系統也會即時回饋出所有與滑鼠特徵值相鄰的微粒，形成有別於以往的聲響互動體驗。

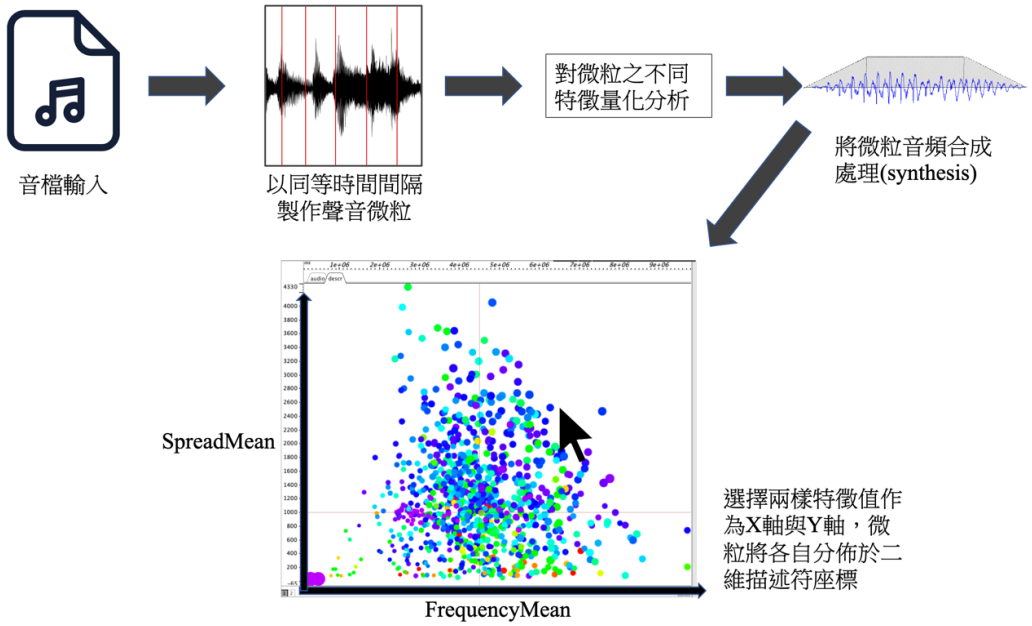

圖 3. CataRT 程式基本流程，最後所有聲音微粒子會分佈於二維描述符座標上，使用者直接以滑鼠觸發聲音微粒。

CataRT 的設計概念發端自對於聲音檔案本身的解構再重組，更真實的模擬出「運動」、「聲音」之間的因果關聯，針對不同程度的位移量與位置，反饋出有連續性之音色。而在傳統互動科技應用上的編程設計往往較生硬直接。例如當影像偵測到人體移動時去呼叫某個完整音效檔案播放，影響參容易因此種不自然的突兀感，影響體驗品質。

四、研究方法

目標將盒塊感測器所偵測到的動態過程轉換成數據，再去觸發 CataRT 上的聲音微粒來產生音效，因此要將系統原本的滑鼠座標數據給代換掉，而以感測器中所得到的位移數據來取代。

4-1 幀差法捕捉影像變化

在使用加 IMU 實測動態數據之前，必須編寫新的程式程式去擴充原本 CataRT 內容，把程式設置成可以同時輸出多點座標，並且同時觸發多個聲音微粒；另外一點是要找到取代滑鼠位移座標的輸入端。為了測試以上兩項的可行性，我們先在 Max/MSP 內建立影像幀差（Frame Difference），將每當下 Webcam 影像與上一幀的影像做比對，把不同顏色的像素以白色輸出，之後再使用 OpenCV，標示出其對應座標，並傳至 CataRT 的二維描述符座標中，圖 4 為此程式流程示意圖。

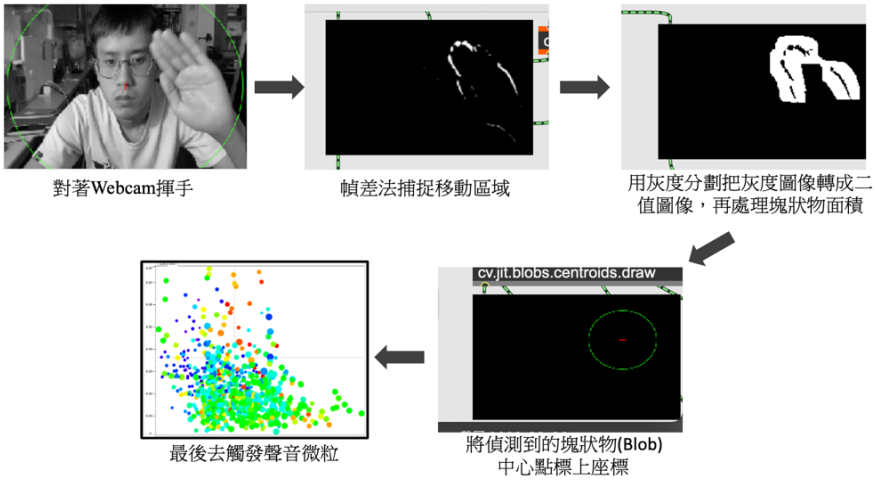

圖 4. 利用幀差法將移動區域匯至 CataRT 的座標中

4-2 多點力學動態偵測

我們選用 ESP32 開發板，並以 MPU6050 六軸感測器作為偵測動態的 IMU，其中我們只以 X, Y 軸的加速度值進行輸出處理。我們使用多對一的方式讓多片 MPU6050 數據藉由 ESP32 以無線方式，集中傳輸數據到一台以序列埠連接電腦的 ESP32 板裡（示意如圖 6.）。目前已測試同時將兩塊 MPU6050 的數據傳輸至電腦的成效。

圖 6. 用 ESP-NOW 無線傳輸示意圖

地殼運動中在地應力長期緩慢的作用下，兩塊不同板塊交界帶造成地殼的岩層發生彎曲變形，當地應力超過岩石本身能承受的強度時便會使岩層產生瞬間的斷裂錯動，其巨大的能量突然釋放，形成構造地震。為了在互動過程中模擬出板塊構造中瞬間的錯動的感覺，我們設想當在平面上瞬間移動感測器時，座標點位移可同步與感測器移往相同方向，為此我們需求出某點瞬間移動的 X－Y 單位向量 $\hat{x}$ 與 $\hat{y}$（公式 1 與公式 2），與移動角度 θ（公式 3），其中 N_x 與 N_y 分別為當下感測器測得 X 方向與 Y 方向加速度值，而 B_x 與 B_y 分別為 X 方向與 Y 方向測得之前 20 個加速度值。最後乘上一個自訂位移參數 P，就能使該點延著某方向移動 P 的距離（程式流程如圖 7）。

$$\hat{x} = \frac{\mathrm{N}_x - \mathrm{B}_x}{\sqrt{\left(\mathrm{N}_x - \mathrm{B}_x{}^2\right) + \left(\mathrm{N}_y - \mathrm{B}_y{}^2\right)}} \tag{1}$$

$$\hat{y} = \frac{\mathrm{N}_y - \mathrm{B}_y}{\sqrt{\left(\mathrm{N}_x - \mathrm{B}_x{}^2\right) + \left(\mathrm{N}_y - \mathrm{B}_y{}^2\right)}} \tag{2}$$

$$\theta = \tan^{-1}\left(\frac{\mathrm{N}_y - \mathrm{B}_y}{\mathrm{N}_x - \mathrm{B}_x}\right) \tag{3}$$

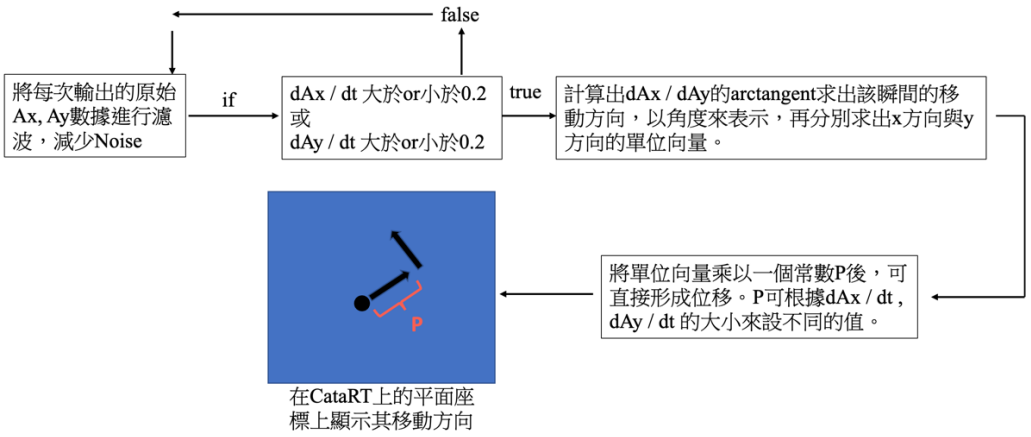

圖 7. 程式設計概念流程

五、結果與討論

5-1 實驗結果

我們實測了將得出的數據直接在 CataRT 上進行互動，將感測器在桌面上以繞圓圈的方式，移動（如圖 9），並觀察它在 CataRT 上的二維描述符座標所走出的軌跡，再將軌跡中每次停留點（圖 10）後製在一張圖上（圖 11）。另外，也將兩片 MPU6050 移動軌跡後製在圖 12 上，同時也將聲音錄下，轉換成波形圖（圖 13），以此證明我們成功運用多個動態感測器（IMU），以多點來即時觸發聲音微粒之系統。

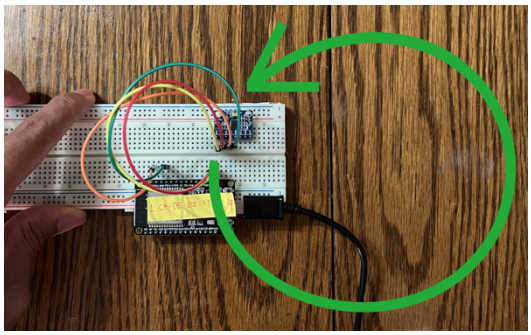

圖 9. 將 MPU6050 於桌面上以多個瞬間位移環繞一圈，觀察數據結果

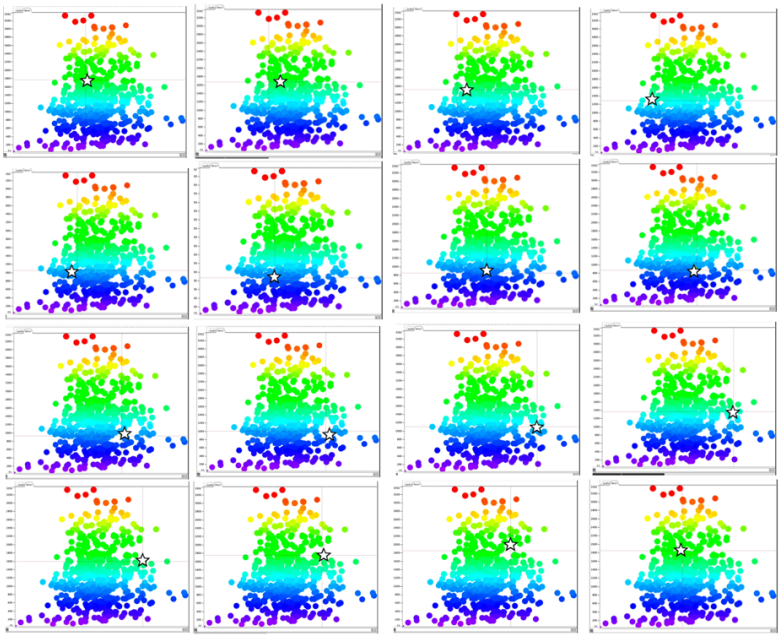

圖 10. 將感測器繞圈時所停留 16 個點都截圖下來。

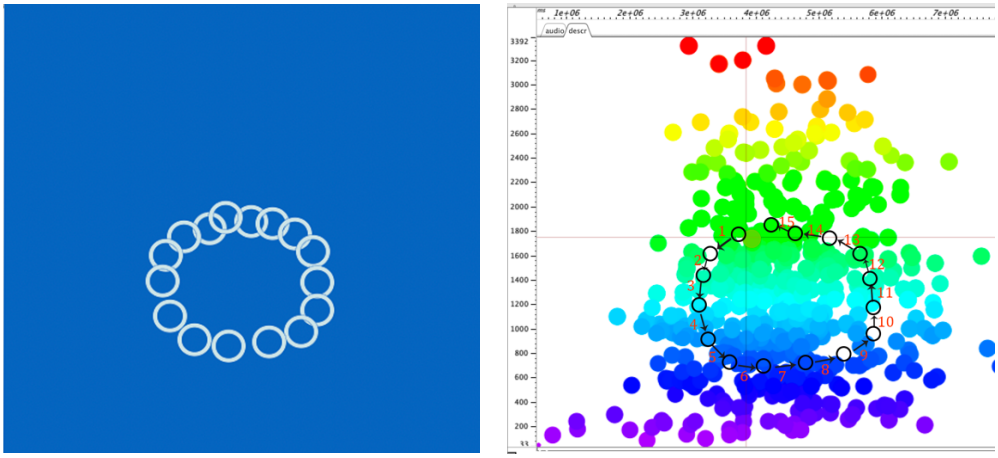

圖 11. 將感測器繞圈時所停留 16 個點都截圖下來，後製在此圖上。

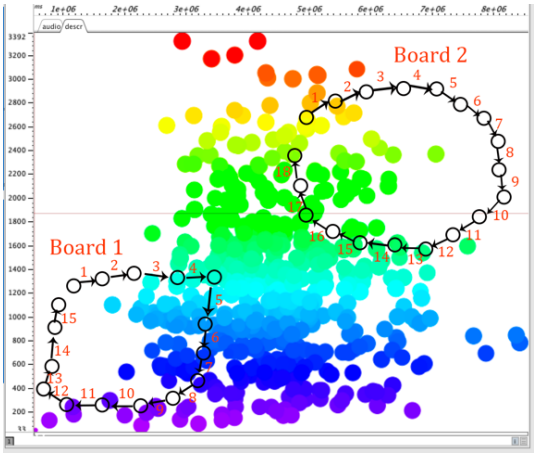

圖 12. 最後使用兩片 MPU6050 與 ESP32，將數據用無線方式傳輸至 CataRT 介面上。

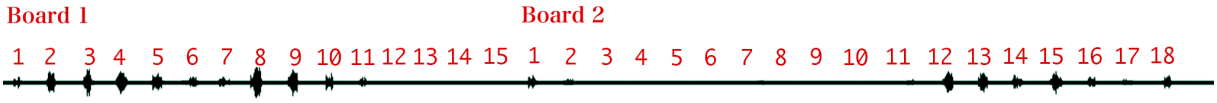

圖 13. 將 Board 2 與 Board 2 觸發聲響錄下來，由錄音圖可發現無產生波形的段落（例如 Board 2 的 3 號到 10 號的位移），是由於座標移動中未觸及到微粒的因素。

5-2. 未來改善方向

1. 位移數據圖形化

若要讓 MPU6050 感測器在位移過程可直接輸出位移量，在基礎理論上，直接將加速度的值積分兩次可得到位移值，但由於感測器上仍有諸多造成測量誤差的因素難以被排除，包含噪音(Noise)與零偏(Bias)。由於位移結果是以原始數據經過二次積分求得的，因此當原始數據產生些微誤差時，積分過後的誤差會呈指數性成長，使數值嚴重飄移(float)，而誤差值也因此難易被歸零校正。後續優化方法可選用九軸感測器，透過磁力計的輔助偵測，或是用感測器整合技術(Sensor Fusion) 減少可被排除的誤差。

2. 聲音微粒在座標上分佈

目前所實測的聲音檔案經由 CataRT 轉換成聲音微粒，顯示在二維描述符座標上時，皆有在空間上分佈不均勻的情況，導致當滑鼠或是感測器位移時，也會使聲音微粒不均勻地被觸發，在實驗過程中也得把座標定在微粒集中處，以免發不出聲音。若可使用特殊演算法把座標符中各微粒平均分散在空間當中，便可排解這樣的問題。

參考文獻

1. Brown, S. R., Scholz, C. H., & Rundle, J. B. (1991). A simplified spring-block model of earthquakes. *Geophysical Research Letters*, *18*(2), 215–218

2. Olami, Z., Feder, H. J. S., & Christensen, K. (1992). Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Physical Review Letters*, *68*(8), 1244–1247
3. Schwarz, D. (2011). State of the Art in sound texture synthesis. *Proc. of the 14th Int. Conference on Digital Audio Effects (DAFx-11),* 1-11
4. Schwarz, D., Beller, G., Verbrugghe, B., & Britton, S. (2006). Real-Time Corpus-Based Concatenative Synthesis with CataRT. *9th International Conference on Digital Audio Effects (DAFx)*, 279-282
5. Truax, B. (1988). Real-Time Granular Synthesis with a Digital Signal Processor. *Computer Music Journal*, *12*(2), 14