

# PREMIÈRE MÉTHODE POUR LE PROLONGEMENT DU GESTE MUSICAL

Charles Bascou  
gmem-CNCM-marseille  
prénom.nom@gmem.org

## 1. PROBLÉMATIQUE

Dans le domaine de la lutherie numérique, et plus largement de l'informatique musicale, les questions de contrôle et d'écriture sont au coeur de nombreuses recherches. En effet, de nouveaux outils de production, transformation ou spatialisation sonores ne peuvent pas se penser sans leurs stratégies de contrôle et d'écriture propres. Depuis l'invention du Theremin en 1919, de nombreuses interfaces d'expressions ont vu le jour. Les interfaces tactiles, depuis quelques années, se sont montrées très efficaces et versatiles pour le contrôle du son électronique, notamment par le fait qu'elles permettent d'avoir, avec un seul geste, un contrôle direct sur plusieurs paramètres en même temps. Par exemple, en ce qui concerne le contrôle de système de spatialisation sonore, on peut considérer une surface tactile représentant la salle de diffusion vue de dessus où les points de contact sont autant de sources sonores virtuelles. Le mouvement du doigt sert donc à déplacer la source sonore dans la salle.

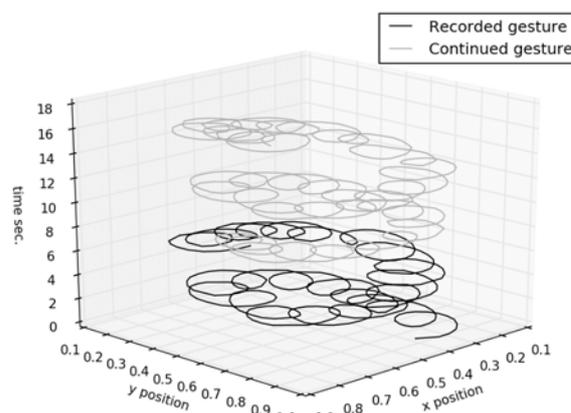
Le contrôle direct est bien entendu intéressant mais l'ajout de fonctions permettant de contrôler les mouvements automatiquement devient vite essentiel. Pour reprendre l'exemple précédent, on souhaiterait faire tourner une source sonore virtuelle à une certaine vitesse autour des auditeurs.

Des systèmes d'enregistrement de mouvement et de relecture en boucle ont déjà été expérimentés et montrent un réel intérêt. Mais nous souhaiterions pousser le modèle plus loin. L'objectif est que la machine puisse produire des mouvements (ou automatisations) à *la manière de*, que la machine puisse prolonger le geste dans le même style.

Dans un scénario élémentaire, j'initie un geste circulaire avec ses caractéristiques spatiales et temporelles propres sur la surface tactile et lorsque je lève le doigt la machine poursuit ce mouvement en reprenant ces mêmes caractéristiques. La notion de cycle ou de périodicité est donc très importante. Cependant nous souhaitons ne pas nous limiter à cet aspect du geste. Ainsi, une pluie de tapotements aléatoires est également un geste qu'il faudrait reproduire à partir de ses propriétés stochastiques.

Nous nous plaçons ainsi dans un cadre large, qui aborde le geste humain à la fois dans ses aspects périodiques et réguliers et dans ses aspects aléatoires et erratiques.

Nous tentons de montrer de plus que le prolongement du geste musical peut être vue comme une nouvelle forme



**Figure 1.** Exemple d'un geste 2d (noir) et son prolongement (gris). L'espace est sur les axes  $x$  et  $y$ , le temps sur  $z$ .

d'écriture, écriture que ne se situe ni dans le champ de la réalisation d'un processus (partition), ni dans la description de ce processus (système génératif paramétrique) mais dans reproduction d'un comportement par l'exemple.

## 2. CADRE DU DEMONSTRATEUR

### 2.1. Délimitations du problème

Même si nous souhaitons pouvoir prolonger n'importe quel type de geste multidimensionnel, nous nous sommes pour l'instant limité aux gestes mono-touche, mono-traités en 2d sur surface tactile.

L'aspect temps-réel est primordial dans cette problématique. Le système doit pouvoir prendre le relais du geste du performeur de manière complètement transparente. Par ailleurs, la finalité temps réel influence notamment le choix de méthodes signal relativement simples et réputées pour leur faible demande en ressource processeur. Étant pour l'instant dans une phase d'évaluation des méthodes de prolongement, nous n'avons pas développé encore cette fonctionnalité dans le démonstrateur proposé. Le système traite le geste une fois exécuté et nous renvoie sa continuation concaténée au geste effectué.

## 2.2. Méthode KNN

Le prolongement du geste proposé dans ce démonstrateur est basé sur une méthode d'estimation de type  $K$  plus proches voisins ( $K$  Nearest Neighbors *eg.* KNN). La phase d'apprentissage est la phase au cours de laquelle le musicien performe le geste. L'algorithme récupère les positions  $\mathbf{p}(t)$  captées et calcule alors certains descripteurs instantanés (vitesse et accélération) à chaque instant. L'algorithme construit un ensemble de points  $\mathbf{v}^{\text{inst}}(t)$  dans  $\mathbb{R}^6$  tel :

$$\mathbf{v}^{\text{inst}}(t) \triangleq \begin{bmatrix} \mathbf{p}(t) \\ \mathbf{p}(t) - \mathbf{p}(t-1) \\ \mathbf{p}(t) - 2\mathbf{p}(t-1) + \mathbf{p}(t-2) \end{bmatrix} \in \mathbb{R}^6,$$

A été ajouté la possibilité de donner une mémoire à chaque point. Il s'agit en fait d'étendre le vecteur de position, vitesse et accélération à l'instant  $t$  avec les positions vitesses et accélérations aux instants  $t - \tau$ . La taille de la mémoire - *eg.* le nombre de points passés et leur portée, est un paramètre que l'on étudie actuellement et qui pour l'instant est réglé manuellement.

S'en suit la phase de prédiction, où l'on souhaite estimer le point  $\mathbf{p}(t+1)$  sachant  $\mathbf{p}(t)$ . On cherche les  $K$  plus proches voisins de  $\mathbf{p}(t)$  et on calcule le mouvement moyens de tous ces points pour ainsi déduire  $\mathbf{p}(t+1)$ . Et l'algorithme réitère.

## 2.3. Moteur sonore

Le moteur sonore utilisé est basé sur une resynthèse granulaire de sons pré-enregistrés. La stratégie de correspondance (*mapping*) est la suivante. Chaque touche (doigt) sur la surface tactile 2d déclenche et contrôle une tête de lecture indépendante. Pour chacune, le *mapping* suivant est défini :

$$\begin{aligned} x &\rightarrow \text{panoramique binaural} \\ y &\rightarrow \text{ratio } \textit{dry/wet} \text{ de réverbération} \\ \|v\| &\rightarrow \text{amplitude du son} \\ \alpha = \textit{ang}(v) &\rightarrow \text{position de lecture dans le son} \end{aligned}$$

Une spécificité intéressante de ce *mapping* est le principe de parcours du son : un mouvement circulaire lira le son en boucle, une rature oscillera entre deux positions dans le fichier son par exemple. Il a été choisi pour permettre *a priori* une bonne cohérence entre les formes spatiales et sonores.

## 3. REMERCIEMENTS

Ce travail a été mené en collaboration avec Valentin Emiya (LIF-Marseille) et Mathieu Lauriere (LJLL-Paris). Il a été partiellement supporté par le Gdr ISIS, projet progest, et par l'Agence Nationale de la Recherche (ANR), avec le projet MAD ANR-14-CE27-0002, Inpainting de

données audio manquantes. Nous aimerions remercier François Denis (LIF) pour ses riches apports scientifiques.