

PROJET CAANEVA, POSSIBILITÉS ET LIMITATIONS DE L'USAGE DES INTERFACES CERVEAU-MACHINE DANS UN CONTEXTE DE CRÉATION D'ENVIRONNEMENT VISUEL ET SONORE

Armand Bazin

Université Gustave Eiffel
bazin.armand@gmail.com

Ulysse Del Ghingaro

Université Gustave Eiffel
ulyssedgo@gmail.com

RÉSUMÉ

Les techniques d'analyse électroencéphalographiques (EEG) permettent aujourd'hui de mesurer l'activité électrique neuronale avec une certaine précision à partir de dispositifs peu chers et polyvalents. De nombreux artistes se sont appropriés ces technologies de mesure corticale en créant des œuvres souvent basées sur l'imaginaire auquel elles sont associées. Cet article s'intéresse aux possibilités réelles et aux limitations de tels dispositifs dans un contexte de création visuelle et sonore. Nous y présentons le projet *CAANEVA* (Création Autonome par Analyse Neuronale d'Environnement Visuel et Auditif) à travers lequel nous avons essayé de mettre en œuvre des boucles de rétroaction entre les perceptions cognitives d'un utilisateur et une création audiovisuelle à l'aide d'une interface cerveau-machine. La mesure de l'activité cérébrale est de façon générale assez complexe à analyser, toutefois il est possible de discriminer des états de concentration qui peuvent être associés à différents paramètres compositionnels. Au sein du projet, l'environnement sonore est directement synthétisé d'après les ondes corticales captées par différentes électrodes de contact, tandis que la synthèse visuelle est, pour sa part, pensée comme étant une représentation matricielle des composantes fréquentielles de ces mêmes ondes. Le système, pensé à partir des interactions entre l'environnement audiovisuel et sa perception, s'auto-organise et laisse émerger un écosystème sonore et visuel.

1. INTRODUCTION

Depuis maintenant plus d'un siècle, les mesures électroencéphalographiques ont gagné en qualité et en précision à mesure des avancées en électronique, en traitement des signaux numériques et en algorithmique statistique [4]. La question se pose alors : peut-on retranscrire, à partir d'une mesure physique et objective, comme l'EEG, les volontés, les images ou encore les sons auxquels un individu peut penser ?

Après la sortie de l'exclusivité d'usage des EEG dans le domaine médical, des artistes comme Alvin Lucier [8], Pierre Henry [6] ou David Rosenboom [10] se sont emparés des technologies d'analyse corticale à des fins créatives [9]. Cependant, c'est bien à partir des années 2010

que les projets artistiques se sont mis à pleinement se développer, grâce notamment à des entreprises privées telles qu'*OpenBCI*, *Emotiv* ou *Neurosky* proposant des outils de mesures accessibles à un plus grand public et laissant entrevoir des possibilités de contrôle cognitif en temps réel ou quasi réel de systèmes interactifs sonores et visuels. On pourra notamment penser au projet *EEG KISS* de Lancel et Maat (2014-2016) ou encore à *Staalhemel* de Christoph De Boeck (2009-2011) [9].

La plupart de ces nouveaux dispositifs intègrent les EEG à des *Interfaces Cerveau-Machine* (ICM)¹. Celles-ci permettent d'interpréter plus facilement les données de tension électrique captées par des électrodes. Le projet *CAANEVA* (Création Autonome par Analyse Neuronale pour Environnement Visuel et Auditif) s'inscrit dans la continuité de ces installations d'arts numériques.

Cette étude, encore en cours, a commencé en octobre 2019 et a été réalisée à deux. Elle est en lien avec les thématiques de nos recherches personnelles de master, traitant de l'approche écosystémique de la composition musicale et de l'utilisation de l'apprentissage automatique et profond à des fins créatives.

CAANEVA se présente comme un écosystème perceptif visuel et sonore à travers lequel un individu porteur de l'ICM sera à l'origine de la synthèse de son environnement audiovisuel. L'idée est de créer, grâce aux technologies électroencéphalographiques, une boucle de rétroaction entre la génération de l'environnement audiovisuel et sa perception.

CAANEVA est une proposition d'utilisation des EEG comme interfaces pour la composition². Au sein de cet environnement perceptif autotrophique le porteur du casque compose en temps quasi réel et de manière autonome l'espace sensitif dans lequel il se trouve immergé. Il est ainsi créateur et spectateur de son milieu (figure 1).

L'idée de cet article est de proposer une méthodologie et un protocole de test des ICM modernes grand public à travers la mise en place d'un dispositif audiovisuel basé sur le neurofeedback.

Finalement, à travers cette étude nous chercherons

¹Les interfaces cerveau-machine sont parfois appelées *Interfaces Cerveau-Ordinateur* (ICO) ou *Brain-Computer Interface* (BCI) en anglais.

²Ici le terme « composition » est à prendre dans son sens large : sonore et visuelle.

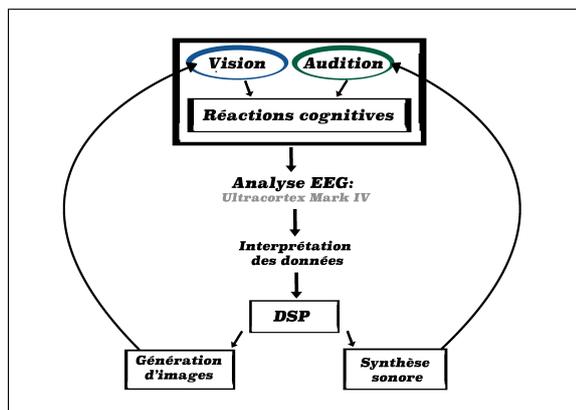


Figure 1. Schéma fonctionnel de l'écosystème CA-NEEVA.

à mettre en avant les possibilités réelles et les limitations de l'utilisation des interfaces cerveau-machine dans un contexte artistique. Afin de mieux appréhender les concepts et analyses exposés dans la suite de cet article il est cependant nécessaire de présenter certains pré-requis théoriques de l'analyse EEG.

2. MATÉRIELS ET MÉTHODES DE MESURES CORTICALES

2.1. Electrodes et Captations ERSF

L'analyse électroencéphalographique permet de mesurer les variations de potentiels électriques synaptiques de zones corticales [4] plus ou moins étendues selon les méthodes d'analyse et la résolution spatiale des électrodes utilisées. Nous nous concentrerons exclusivement sur les EEG dans la suite de cet article.

Les EEG permettent ainsi de mesurer une tension électrique en fonction du temps. Cette fonction complexe, appelée *time serie* captée par chacune des électrodes, est alors analysable comme étant une somme de fonctions sinusoïdales simples. Aussi, à la manière d'une analyse spectrale sonore, il est possible d'appliquer une analyse de Fourier sur le signal capté.

L'analyse spectrale permet l'identification de certains types d'ondes. Ces types d'ondes sont définis par leurs fréquences d'activation qui correspondent aux rythmes d'activation synaptique (aussi appelés rythmes cérébraux) et dont l'amplitude est corrélée à certains états mentaux. Les ondes alpha correspondant aux rythmes cérébraux de 8 à 12 Hz et qui sont souvent corrélées à un état de relaxation et d'apaisement. Les ondes bêta sont elles représentées de 12 à 30 Hz et sont associées à la pensée active. Couplées entres elles, les ondes alpha et beta permettent de distinguer de façon relativement fiable différents états de concentration d'un sujet. Les ondes gamma sont quant à elles des marqueurs d'une activité mentale intense; elles se situent dans la bande passante de 24 à 40 Hz. Ce sont principalement ces trois types d'ondes, alpha, bêta et gamma qui nous intéresseront ici.

Certains comportements et attitudes sont discriminables d'après une analyse EEG. Par exemple, il est re-

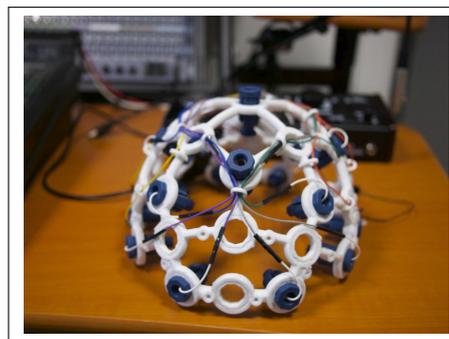


Figure 2. Casque ultracortex « Mark IV ».

connu que les fluctuations du niveau de vigilance, modifient de façon importante l'aspect du tracé de l'électroencéphalogramme ainsi que son contenu fréquentiel. L'association des types d'ondes à des fonctions cognitives distinctes reste toutefois encore mal connue et discutée aujourd'hui [4].

Par ailleurs, il est important de remarquer que les amplitudes maximales des différents types d'ondes dépendent des individus, du contexte et de nombreux facteurs desquels il peut être difficile d'isoler un motif. Aussi, pour optimiser la précision des mesures il est souvent nécessaire de passer par une phase de calibrage de l'EEG. Une fois le calibrage fait, l'analyse corticale s'effectue par la discrimination de perturbations liées à des événements perceptifs ou moteurs (ERSF³). Les ERSF sont des fluctuations visibles du spectre en réponse à un stimulus. Elles sont toutefois à différencier des artefacts qui sont des perturbations du spectre causées par des contractions musculaires et pas uniquement par l'activité corticale. Par exemple, changer d'expression faciale entraîne une activité électrique des muscles du visage qui pourra être détectée par les électrodes créant ainsi des pics d'intensité lors de l'enregistrement des données. Ces pics ne reflètent pas l'activité cérébrale et peuvent ainsi être considérés comme des artefacts⁴. Ces artefacts ne sont pas pour autant dénués d'intérêt dans un cadre artistique puisqu'ils sont identifiables facilement et contrôlables sur commande par le porteur du casque.

2.2. Spécificité matérielle et disposition des électrodes

2.2.1. Hardware

L'expérimentation s'est faite grâce au casque *ultracortex* « Mark IV », produit d'Open BCI⁵ (figure 2).

Le casque permet une disposition de seize électrodes en respectant toutes les conventions du système 10-20⁶ (figure 3).

Les électrodes dans un système 10-20 sont labellisées en fonction de leur zone d'analyse [7]. Ainsi, on trouve les

³ERSF est l'abréviation d'*Event Related Spectral Perturbation*

⁴C'est en ces termes qu'on référencera les captations dont l'origine n'est pas cognitive dans la suite de cet article

⁵openbci.com, accédé le 21 octobre 2020.

⁶10-20 réfère ici aux rapport de distance entre deux électrodes adjacentes qui représente 10 ou 20 pourcents de la longueur et de la largeur totale du crâne

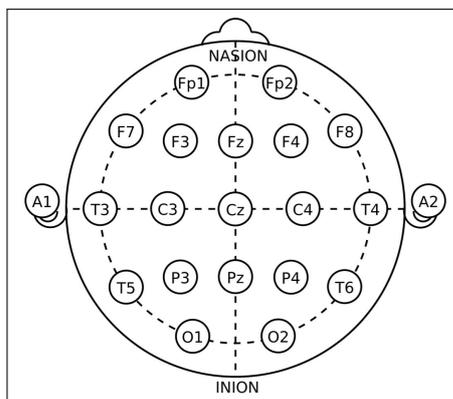


Figure 3. Schéma de disposition des électrodes suivant le système 10-20.

labels « Fp » pour les électrodes analysant le lobe préfrontal, « F » pour celles placées sur le lobe frontal, « T » pour les lobes temporaux, « O » pour le cortex occipital, « P » pour le cortex pariétal et « C » pour les zones centrales. Fz/Cz/Pz sont quant à eux les labels des points de références nulles, les « terres » du système électroencéphalographique.

Cette labellisation permet d'avoir une séparation systématique du cerveau en zones fonctionnelles. Chacune de ces zones est responsable de fonctions cognitives très différentes. Aussi, en sachant précisément quelle zone est mesurée par quelle électrode il devient possible d'interpréter les variations d'amplitudes des signaux neuronaux. Les limitations des différentes zones corticales ont été déduites des mesures d'ERSP et d'observations anatomiques. Ces zones ainsi que leur fonctions associées sont un bon point de départ d'analyse. Aussi, dans CAANEVA nous nous intéressons principalement aux variations d'amplitudes des électrodes placées sur les lobes temporaux dans lesquels se situe le cortex auditif primaire responsable de l'audition (T3, T4); aux lobes occipitaux responsables des fonctions associées à la vision (O1, O2) et finalement aux lobes préfrontaux responsables des fonctions dites supérieures (Fp1, Fp2), qui rassemblent notamment les fonctions exécutives et les fonctions associées au raisonnement.

2.2.2. Software

Le logiciel *open source OpenBCI-GUI*, qui fonctionne sous *processing* (langage assimilé à Java), nous permet de récupérer les données des *times series* associées à chaque électrode. Ces dernières représentent les informations de tension électrique (en μV) mesurées par les électrodes au cours du temps. Le taux d'échantillonnage des *times series* est de 125 Hz. *OpenBCI-GUI* sert de logiciel d'acquisition et de pré-traitement des données.

Parmi les premiers traitements applicables directement sur *OpenBCI-GUI* pour débruiter le signal on peut utiliser le *notch filter*. Cet outil permet de réduire le bruit lié à l'impédance électrique en appliquant un filtre à 50 ou 60 Hz sur le signal. Il est à noter qu'il est possible d'effectuer directement une transformée de Fourier rapide (FFT en anglais) dans *OpenBCI-GUI*. Cette FFT a de base une

résolution fréquentielle de 0.97 Hz et est constituée de 125 échantillons. Autrement dit, cette FFT permet d'analyser les différentes amplitudes des composantes fréquentielles situées sur une bande passante de 0 à environ 121 Hz. Ces données peuvent ensuite être aussi communiquées au logiciel *Max/MSP* via *Open Sound Control* (OSC), en utilisant le widget *Networking* d'*OpenBCI-GUI*. *Max/MSP* fait figure de logiciel de composition principal, il est utilisé à la fois dans le traitement des données et dans la création de l'environnement audiovisuel. Son utilisation sera décrite plus précisément dans la suite de l'article.

3. PROTOCOLES EXPÉRIMENTAUX

3.1. Tests des dispositifs

Afin d'analyser les réponses du casque, deux hypothèses ont été émises.

3.1.1. Identification des artefacts

Hypothèse 1 : les changements d'expression faciale sont isolables et identifiables grâce aux artefacts qu'ils produisent sur l'électroencéphalogramme. Ces artefacts sont suffisamment caractéristiques pour être reconnus.

Protocole 1 : les expressions faciales testées ont été choisies selon leur simplicité d'exécution et leur reproductibilité : il s'agit de clignements des yeux, de la contraction de la mâchoire et de la fermeture des paupières pendant un temps long. Il a été demandé au porteur du casque de faire ces mouvements plusieurs fois de suite afin de pouvoir observer les artefacts (ERSP) qu'ils produisaient. Les données EEG ont été enregistrées et une analyse a été faite des *time series* pour identifier des paramètres propres à chaque artefact selon son origine. Un test témoin a été fait où le porteur du casque garde un visage fixe tout au long de l'enregistrement (à l'exception du clignement des yeux).

3.1.2. Détection des réponses aux stimuli

Le but étant de capter l'influence que l'écoute [1] et la vision peuvent avoir sur l'activité cérébrale, il est crucial de savoir s'il est possible de détecter cette influence via le casque EEG.

Hypothèse 2 : une activité intense au niveau du cortex préfrontal, auditif ou visuel peut être détectée selon les données captées par les électrodes placées sur les zones associées à ces fonctions. C'est à dire lors d'une activité monopolisant une de ces fonctions cognitives, le casque EEG est capable de détecter un changement dans les valeurs des ondes cérébrales, et ce changement est suffisamment caractéristique pour être identifié comme une conséquence de la concentration du sujet sur cette fonction.

Il est difficile de définir un état de concentration, qui peut paraître bien subjectif. Pour palier ce défaut, nous avons décidé de définir la concentration comme la réalisation d'une action qui monopolise grandement l'attention (rendant tout autre action impossible ou très complexe).

Au contraire un état de relaxation n'est défini par aucune action spécifique.

Pour l'ensemble des protocoles, les données enregistrées et analysées proviennent des électrodes situées sur les zones Fp1 et Fp2 (cognition), T3 et T4 (audition), O1 et O2 (vision). Les tests durent tous 35 secondes et ont été répétés trois fois chacun dans un souci de normalisation et d'exhaustivité des résultats. Des échantillons de 35 secondes sont suffisamment longs pour être représentatifs d'une activité mentale spécifique. Cette durée facilite les mesures et permet d'augmenter la taille de la base de données et par là même sa représentativité.

Protocole de calibration : comme vu précédemment un état d'attention particulier peut être identifié via les valeurs des ondes alpha et beta. Un protocole de calibration a été effectué sur Armand Bazin, utilisateur final du dispositif. Il lui a été demandé de réaliser des opérations de calcul mental, et pour le test témoin de se relaxer.

Protocole 2 : le premier groupe de tests a été basé sur une stimulation visuelle. Le porteur du casque devait suivre des yeux un visuel mouvant (point noir sur fond blanc). Ce même test a aussi été réalisé avec un autre visuel plus complexe dont il devait interpréter les formes (nuages de points mouvants). Pour les deux tests témoins le sujet était soumis aux mêmes visuels mais ne devait pas se focaliser dessus : ne pas suivre les points, ne pas chercher à interpréter les formes, etc.

Le second groupe de tests a été basé sur une stimulation auditive. Le sujet devait écouter de façon active (anticiper, analyser les timbres, discriminer les plans sonores) une musique. Pour le témoin le sujet était soumis à la même musique mais il lui était demandé d'essayer de ne pas écouter activement la pièce. La musique choisie est celle composée à partir du patch *Max/MSP* du projet afin d'avoir des timbres qui correspondent aux sons de la composition finale.

Deux tests complémentaires ont été réalisés afin d'essayer d'isoler les réponses à des fonctions précises. Pour isoler la réponse à un stimulus visuel, le sujet a mis des bouchons d'oreilles avant d'être exposé au visuel complexe (nuages de points mouvants). Et pour isoler la réponse à un stimulus auditif, il a été demandé au sujet de fermer les yeux avant d'écouter activement la musique.

Les enregistrements des données FFT ont été faits avec le logiciel *OpenBCI-GUI*. Pour l'étude des artefacts seule l'observation des "time series" a été nécessaire. Pour la mesure de l'activité cérébrale un patch *Max/MSP* a été créé pour un premier traitement des données. Ce patch est construit comme suit. Les FFT d'*OpenBCI-GUI* sont reçues via OSC puis les données de chaque électrode sont séparées (objet : route). Les plages de fréquences correspondantes aux ondes alpha et beta sont isolées dans deux listes séparées pour chaque électrode. Puis un sous-patch fait la moyenne de ces fréquences et l'enregistre (objet : coll) pour un pas de temps donné. Le pas de temps choisi pour l'ensemble des tests est de 110 ms. Une première étape d'analyse a consisté à faire la moyenne des valeurs des ondes alpha et beta sur les 35 secondes pour tous les

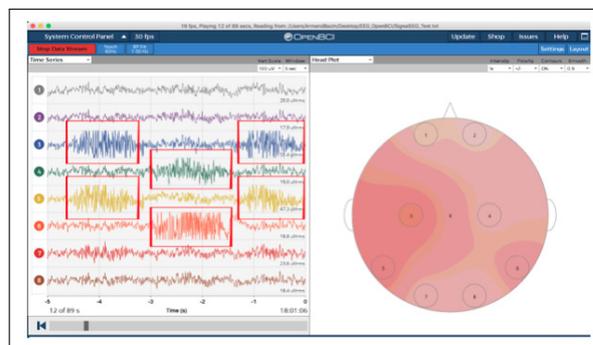


Figure 4. Mise en évidence d'un artefact lié à la contraction des mâchoires.

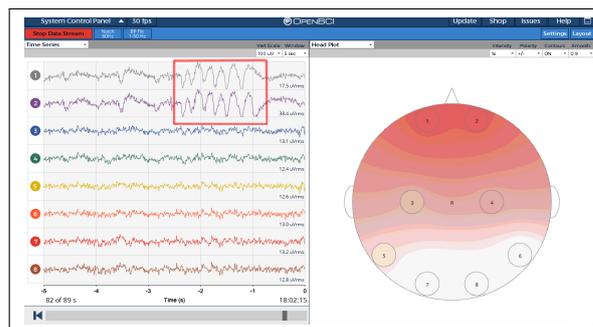


Figure 5. Mise en évidence d'un artefact lié au clignement des yeux.

tests afin d'avoir une valeur associée au stimulus et à l'activité demandée au sujet.

3.2. Résultats

Résultats du protocole 1 : les artefacts de serrage de mâchoire sont visibles sur les électrodes situées en T3, T4, T5, T6, C3 et C4 respectivement les électrodes numérotées 3, 4, 5, 6, 7 et 8 (figure 4).

Ces artefacts sont assez durs à décrire de manière précise puisqu'ils dépendent de l'intensité du serrage de mâchoire et de sa durée. Cependant l'augmentation du voltage captée par les électrodes est telle (supérieure à $30 \mu\text{V}$) qu'elle ne peut pas être confondue avec une activité cérébrale aussi intense soit elle. Ces artefacts sont donc facilement identifiables. Pour les clignements des yeux les réponses aux artefacts du casque sont précises et très semblables sur les *time series* (figure 5).

Les artefacts de clignements des yeux sont présents sur les électrodes placées en Fp1 et Fp2 soit les électrodes 1 et 2. Ils durent environ 25 ms et se repèrent grâce à une augmentation d'environ $80 \mu\text{V}$ sur les *time series*. De plus ils possèdent le même profil sur les deux électrodes qui les captent.

Résultats du protocole 2 : pour l'instant seuls les résultats du protocole de calibration ont pu être analysés (table 1).

En se basant sur les résultats du protocole de calibration, on constate une tendance chez les ondes alpha à être supérieures en phase de relaxation comparativement à la phase de concentration et cela pour les deux électrodes. Cepen-

abrite le granulateur, la 3 le système d'enregistrement des échantillons et enfin la section 4 le système de lecture des échantillons. Pour que les sons immergent complètement la salle, une diffusion quadraphonique est envisageable pour le dispositif. La répartition des sons dans la panoramique quadraphonique se fera à l'aide du granulateur. Le but étant non pas de pouvoir faire des mouvements de panoramique mais de créer différentes répartitions des masses sonores entre les enceintes variables en fonction du temps.

La combinaison de ces trois sources sonores intriquées permet l'obtention de sons riches à partir du son direct qui est assez simple et possède peu de variations mais qui sert d'ancrage fort pour l'auditeur et de point de départ toujours identique pour toutes les compositions. La granulation permet de synthétiser des sons au timbre qu'on pourrait qualifier d'aquatique, métallique jusqu'à de discrets crépitements. Enfin c'est la concordance des nombreux paramètres qui permet une grande diversité d'environnement sonore, aussi bien dans l'aspect harmonique que rythmique ou timbral. Au final, le dispositif permet par lui-même de créer des compositions dignes d'intérêt ou de servir à produire de façon originale des matériaux de base pour des compositions électroacoustiques⁷.

4.1.2. Environnement visuel

Le visuel du dispositif CAANEVA est synthétisé à partir d'une représentation matricielle des composantes fréquentielles des ondes captées au niveau du cortex préfrontal. Les *times series* étant des signaux périodiques, l'analyse de Fourier nous permet de les décomposer en une multitude de fonctions sinusoïdales, les composantes fréquentielles des signaux. Chacune de ces composantes fréquentielles représentent ainsi une fonction sinusoïdale simple dont on fait correspondre l'amplitude maximum à une valeur RGB de 255 (pixel blanc) et l'amplitude minimale à une valeur de RGB de 0 (pixel noir) [3]. Les amplitudes intermédiaires correspondent alors à un dégradé de gris. Ces fonctions sinusoïdes sont ensuite étendues à une troisième dimension par projection. L'image résultante d'une composante fréquentielle donnée sera le plan x,z (figure 9).

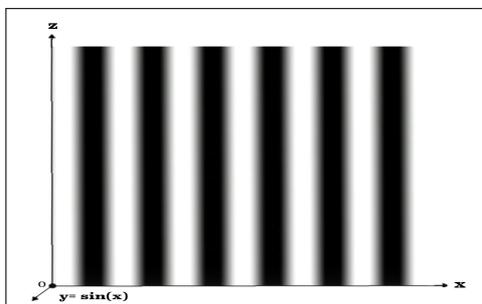


Figure 9. Passage d'une représentation 2D d'une sinusoïde à une représentation 3D (plan x,z)

⁷Un exemple d'une composition faite à partir de montage d'expérimentations sonores obtenues avec CAANEVA, youtu.be/AUy1BLFvNrU, accédé le 21 octobre 2020.

Chacune des fréquences est de cette manière représentée par une série de lignes plus ou moins proches. L'aspect « flou » des bandes noires et blanches provient du gradient de gris issu des valeurs intermédiaires entre amplitude maximum et minimum⁸. Chacune de ces représentations de fréquence est, dans un second temps et après avoir subi une rotation, disposée sur un canevas. La rotation de chacune des représentations fréquentielles se fait en relation avec l'activité du cortex occipital (responsable de la vision)⁹. Finalement, l'espace visuel émerge [14] de cette intrication entre chacune des représentations des composantes fréquentielles du signal capté par Fp1 et Fp2. Dans nos expérimentations nous travaillons également sur l'implémentation de différents opérateurs permettant la génération de visuels plus riches et variés permettant une meilleure stimulation corticale des différentes régions analysées par le casque¹⁰.

4.1.3. Modulation de l'environnement

L'environnement de synthèse issu de la FFT des signaux de Fp1 et Fp2 est modifié d'après les *times series* reçues d'O1 et O2 (partie visuelle) et de T3 et T4 (partie sonore). Ces *times series* sont analysées par un algorithme de type K-mean (K-moyenne en français) permettant de partitionner les données reçues par les électrodes Fp1 et Fp2.

L'algorithme de type K-mean à apprentissage non supervisé [2] est entraîné à partir d'enregistrements d'activités corticales issus de phases de test. Son but est de partitionner les données selon leur niveau de proximité sur une échelle de temps et d'amplitude afin de fournir des moyennes associées à certaines activités neuronales distinctes. Ces moyennes permettent ensuite d'estimer la présence de certains comportements neuronaux chez un même utilisateur lors de prochaines utilisations. Les activités ainsi reconnues déclenchent des modifications telles que décrites dans les parties en amont.

4.2. EEG et boucles de rétroactions

L'utilisation des électroencéphalogrammes en temps réel permet de créer une boucle de rétroaction entre le porteur du casque et l'œuvre. Nous pensons que c'est sur ce point précis que les EEG trouvent une grande partie de leur intérêt dans un cadre artistique. A partir de ce constat nous avons pensé le dispositif en nous inspirant des travaux d'Agostino Di Scipio dont les oeuvres sont souvent basées sur des boucles de rétroactions approfondies créant

⁸Exemple d'un visuel dans sa première version, sans rotation, fait à partir d'une captation en temps réel (opérateur d'addition entre les fréquences) : youtu.be/d84kU5tdRtg, accédé le 21 octobre 2020.

⁹Exemple d'un visuel, avec rotations, à partir d'une captation en temps réel (opérateur d'addition entre les fréquences) : youtu.be/mTAYrdqYa7U, accédé le 21 octobre 2020.

¹⁰Un exemple d'un visuel dans une version plus avancée, avec rotations, (changement des opérateur entre les fréquences) : www.youtube.com/watch?v=JpHBiL9d5MU, accédé le 21 octobre 2020 – Avertissement : les visuels générés dans ce test comportent des stimulations lumineuses intermittente qui sont susceptibles de provoquer des crises d'épilepsie chez les personnes y étant sensibles.

des écosystèmes sonores dont les interactions complexes font émerger une auto-organisation [11]. Ici, nous avons voulu pousser l'idée encore plus loin en intégrant la perception de l'individu au centre d'un écosystème que nous avons qualifié de « perceptif sonore et visuel ». Tout le dispositif s'articule en différentes boucles d'interactions qui s'influencent et d'où émerge une forme compositionnelle. Le but du dispositif n'est donc pas tant de créer une composition en soit mais bien les conditions nécessaires à son émergence. En effet, l'émergence est définie en science cognitive comme l'apparition, au sein d'un système complexe, de propriétés nouvelles non prédictibles et irréductibles à la somme des composantes du système et/ou de ses états antérieurs [5, 12].

Aussi, à travers ce dispositif nous avons cherché à mettre en place les bases d'un système qui, à travers la perception du porteur du casque, intègre son environnement. On peut alors définir ce système comme un écosystème, qui ne sera pleinement fonctionnel que lorsque l'utilisateur y prendra place.

De plus l'espace sonore et l'espace visuel dans lesquels se trouvera immergé l'utilisateur sont synthétisés à partir des mêmes « times series » issus de la mesure des électrodes placées sur son cortex préfrontal. De cette manière, la cognition et le niveau d'attention du porteur du casque font office d'une source unique de synthèse pour l'ensemble de l'écosystème. Cet espace audiovisuel s'auto-influence, de par l'action réciproque du son sur l'image et de l'image sur le son (on pourra penser ici à l'effet McGurck comme exemple de ces interactions). Que le spectateur considère le son et l'image comme causale ou non, des liens émergent de l'interaction entre ces deux espaces.

Finalement, c'est de l'interprétation de ces relations par le spectateur que sera issue la composition.

Malgré une certaine forme d'auto-organisation de la pièce, le porteur du casque possède la capacité d'influencer la composition de façon plus ou moins consciente par le déclenchement d'artefacts ou encore par son niveau d'attention.

5. CONCLUSION

Avec CAANEVA nous avons cherché à utiliser les ICO comme des interfaces de composition. Cependant, ce projet connaît quelques limitations en partie dues à la nature des EEG. La principale source d'imprécision de ce travail provient de la résolution spatiale de la captation par électrodes. En effet, cette résolution étant très faible il est compliqué de cibler précisément une zone du cerveau et d'isoler individuellement des phénomènes neurocognitifs. Les données récupérées sont ensuite enregistrées sans amplification ce qui peut entraîner un manque de précision notamment au niveau des micro-variations d'amplitude du signal. Par ailleurs, et de façon général, le port d'un casque EEG peut influencer de façon significative l'état mental du porteur, en le mettant dans une situation peu propice à la relaxation par exemple. L'objet de mesure influence ainsi la mesure, ce qui peut nuire à la composition. De plus, le

dispositif CAANEVA est constitué d'une longue chaîne de traitement entraînant des biais d'interprétation en ajoutant du bruit au signal analysé.

Pour ce qui est des résultats expérimentaux, isoler l'attention auditive de l'attention visuelle s'est avéré compliqué avec notre première approche d'analyse de données. Une autre limite réside dans la nécessité du calibrage en amont de la composition pour améliorer la précision de la captation et des contrôles cognitifs sur un individu précis. Le calibrage peut être un exercice long qui peut potentiellement résulter à une mauvaise généralisation du modèle sur d'autres individus.

Une étude complémentaire est donc nécessaire. Dans un premier temps pour augmenter le nombre de tests et de protocoles afin de tirer des tendances à confirmer ou infirmer. De plus, augmenter le nombre d'échantillons présents dans la base de données pourrait permettre de pallier le mauvais rapport signal-bruit des enregistrements. Une autre solution réside dans l'accès à un matériel EEG de meilleure qualité, avec une meilleure résolution et un meilleur rapport signal-bruit. Au niveau de l'analyse du signal pour contrôler les transformations du son et du visuel, l'utilisation d'un système d'apprentissage automatique semble être judicieux [13]. Cependant, l'implémentation de l'algorithme de type K-mean permettant faire des moyennes par regroupement n'est pas complètement optimale. Au niveau du patch *Max/MSP*, la connexion avec ce système d'analyse reste à implémenter pour qu'il puisse contrôler la synthèse granulaire ainsi que la modification du visuel.

Pour autant les résultats esthétiques sont encourageants. Le dispositif crée une grande variété de sons avec des timbres intéressants, qui peuvent servir de matériaux de base pour des compositions. Les compositions en temps réel sont prometteuses bien que l'imprécision des données récoltées limite en partie la pertinence des interactions entre le porteur du casque et l'environnement sonore et visuel. Il reste tout de même un travail à effectuer sur l'enchaînements des différentes phases de l'environnement aussi bien visuel que sonore pour le rendre plus cohérent.

6. REMERCIEMENTS

Nous tenons à remercier Miguel Almiron, Kévin Dahan, Martin Laliberté pour leur aide durant la rédaction de cet article et Januibe Tejera pour son aide lors de la confection du patch *Max/MSP*.

7. RÉFÉRENCES

- [1] Alluri, V., Toivianen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., Brattico, E. « Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm », *Neuroimage* 56/4 (2012), p. 3677-3689.
- [2] Bishop, C. M. *Pattern recognition and machine learning*. Springer, New York, 2006.

- [3] Bradley, J., Brislawn, C. M., Hopper, T. « The FBI Wavelet/Scalar Quantization Standard for Gray-scale Fingerprint Image Compression », *Visual Information Processing II*, Huck, F.Ö., Juday, R. D. (dir.), SPIE, Bellingham, 1993.
- [4] Grandchamp R. « Electroencéphalographie et Interfaces Cerveau-Machine : nouvelles méthodes pour étudier les états mentaux », thèse de doctorat, sous la dir. de A. Delorme, Université Paul Sabatier, Toulouse, 2012.
- [5] Hautbois, X., Laliberté, M., Stransky, L., Stransky, V., et coll. (dir.). *Emergence en Musique, Dialogue des sciences*. ACROSS Delatour, Paris, 2019.
- [6] Henry, P. *Mise en musique du Corticalart de Roger Lafosse* (1971).
- [7] Jasper, H. « Report of the committee on methods of clinical examination in electroencephalography », *Electroencephalography and Clinical Neurophysiology* 10/2 (1958), p. 370-375.
- [8] Lucier, A. *Music for Solo Performer* (1965).
- [9] Miranda, E. et Brouse, A. « Interfacing the brain directly with musical systems: on developing systems for making music with brain signals », *Leonardo* 38/4 (2005), p. 331-336.
- [10] Rosenboom, D. *Portable Gold and Philosophers Stones* (1974).
- [11] Meric, R. et Solomos, M. « Audible Ecosystems and emergent sound structures in Di Scipio's music. Music philosophy enhances musical analysis », *Proceedings of the Conference for Interdisciplinary Musicology*, Thessaloniki, Greece, 2008.
- [12] Tiberghien, G. *Dictionnaire des sciences cognitives*. Armand Collin, Paris, 2002.
- [13] Werbos, P. J. « Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences », thèse de doctorat, sous la dir. de K. Deutsch, Harvard, Cambridge, USA, 1974.
- [14] Whitelaw, M. *Metacreation: art and artificial life*, MIT Press, Cambridge, 2004.

Texte édité par Corentin Guichaoua.