

Applications Spatial Audio temps réel (SPARTA)

Sur

SPARTA est une collection de plug-ins audio flexibles VST destinés à la production, à la reproduction et à la visualisation audio spatiales, développée par les membres du laboratoire Acoustique de l'Université Aalto. Ces plug-ins ont déjà été utilisés en interne au sein du laboratoire d'acoustique à des fins pédagogiques et de recherche; Cependant, ils ont maintenant été publiés en tant que projet open source. Nous espérons qu'ils seront utiles aux personnes intéressées par le traitement audio spatial en temps réel.

La suite SPARTA inclut également la suite paramétrique [COMPASS](#) .

- Les liens de téléchargement peuvent être trouvés [ici](#)

Vous trouverez ci-dessous une description détaillée de chaque plug-in.

Le code source a également été publié sous licence GPLv3 et peut être trouvé [ici](#) , où les plug-ins SPARTA sont essentiellement les implémentations d'interface graphique utilisateur des exemples donnés dans Spatial_Audio_Framework, disponibles [ici](#) .

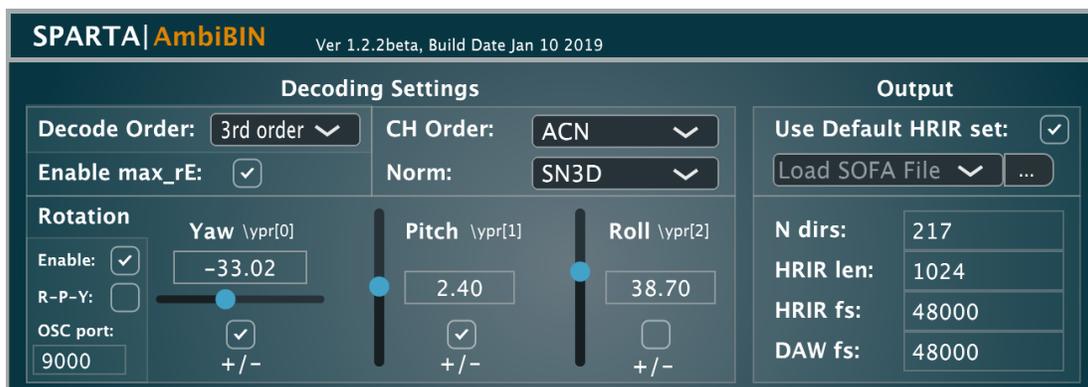
Les plug-ins SPARTA

Tous les plug-ins sont testés avec REAPER (64 bits), une DAW très abordable et flexible, qui est actuellement le seul hôte recommandé pour ces plug-ins. Actuellement, les plug-ins prennent en charge des fréquences d'échantillonnage de 44,1 / 48 kHz et des tailles de bloc multiples de 256 échantillons, sauf indication contraire.

Tous les plug-ins liés aux harmoniques sphériques sont conformes à la convention de commande du numéro de canal Ambisonic (ACN) et prennent en charge les mises à l'échelle orthonormalisée (N3D) et semi-normalisée (SN3D). remarque: AmbiX utilise ACN / SN3D. L'ordre de transformation maximal pour ces plug-ins est 7.

De plus, grâce à l'aide de Daniel Rudrich, les plug-ins correspondants prennent désormais en charge l'importation et l'exportation des instructions de haut-parleur, source et capteurs via des fichiers de configuration .json; permettant une compatibilité croisée entre SPARTA et la [suite de plug-ins IEM Ambisonics](#) . Plus d'informations sur la structure de ces fichiers sont disponibles [ici](#).

AmbiBIN



SPARTA|AmbiBIN Ver 1.2.2beta, Build Date Jan 10 2019

Decoding Settings

Decode Order: 3rd order ▾ CH Order: ACN ▾

Enable max_rE: Norm: SN3D ▾

Rotation

Enable: Yaw \ypr[0] Pitch \ypr[1] Roll \ypr[2]

R-P-Y: -33.02 2.40 38.70

OSC port: 9000 +/- +/- +/-

Output

Use Default HRIR set:

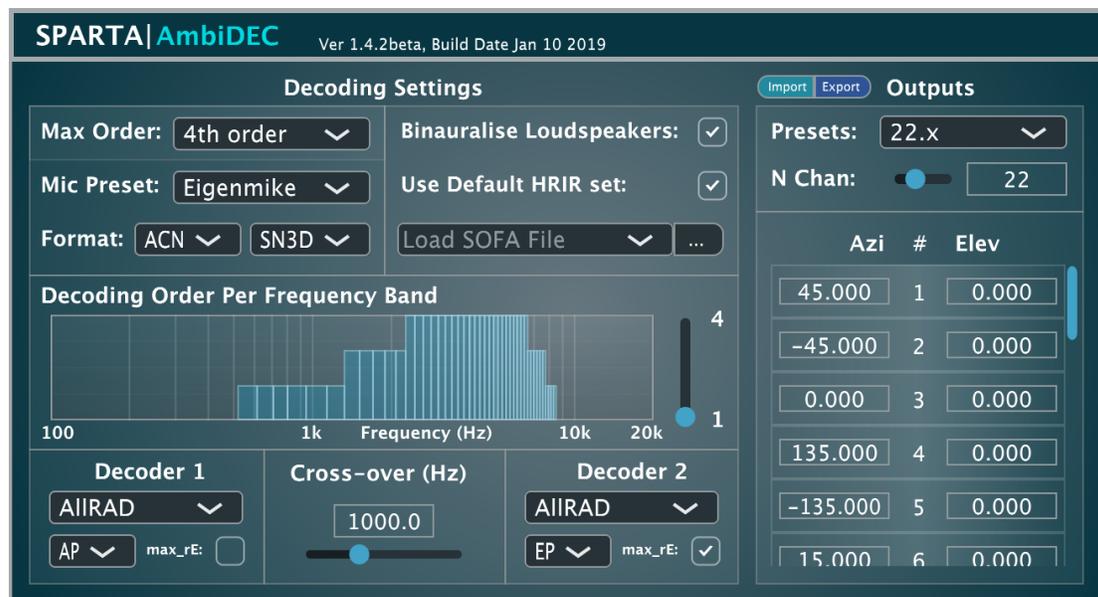
Load SOFA File ▾ ...

N dirs:	217
HRIR len:	1024
HRIR fs:	48000
DAW fs:	48000

Décodeur binaural Ambisonic pour la lecture au casque de signaux harmoniques sphériques (également appelés signaux Ambisonic), avec rotateur intégré et prise en charge du suivi de la tête via des messages OSC. Les angles de rotation sont mis à jour après la transformation temps-fréquence, ce qui permet une latence réduite par rapport à son homologue de haut-parleur 'AmbiDEC' lorsqu'il est associé à 'Rotator'. Le plug-in permet également à l'utilisateur d'importer ses propres HRIR via le standard SOFA.

Ce plug-in a été développé par Leo McCormack.

AmbiDEC



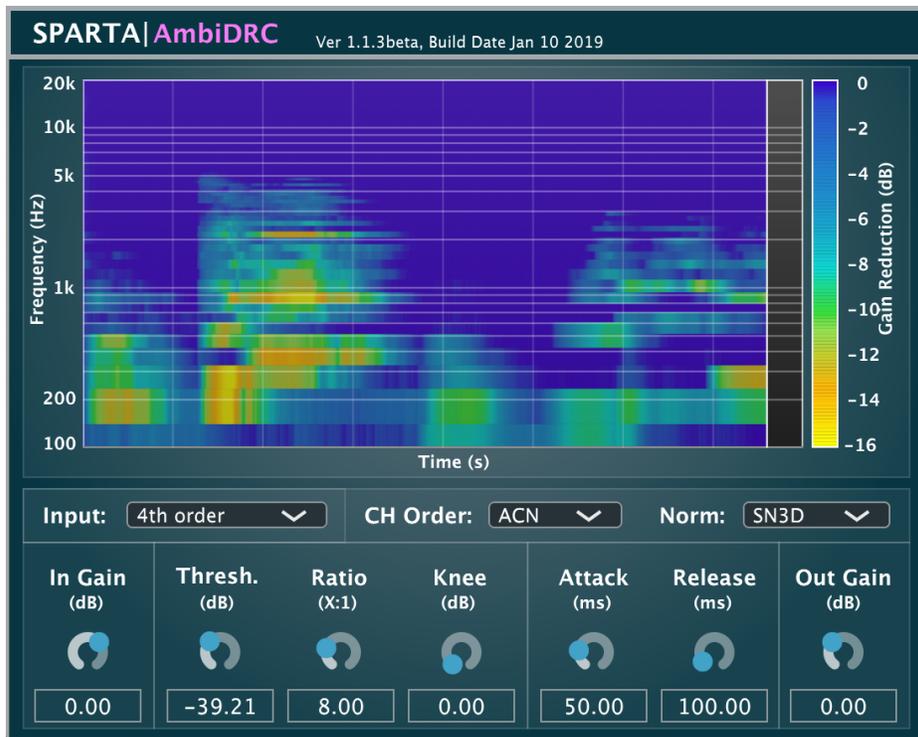
Décodeur Ambisonic dépendant de la fréquence pour les haut-parleurs. Les directions des haut-parleurs peuvent être spécifiées par l'utilisateur pour un maximum de 64 canaux. Vous pouvez également sélectionner des pré-réglages pour les configurations 2D et 3D courantes. Pour la reproduction au casque, l'audio du haut-parleur est converti par convection avec des HRTF interpolés pour chaque direction du haut-parleur (approche du haut-parleur virtuel). Le plug-in permet également d'importer des HRIR personnalisés via le standard SOFA.

Le plug-in utilise une approche de décodage double, grâce à laquelle différents réglages de décodeur peuvent être sélectionnés pour les basses et hautes fréquences; la fréquence de croisement peut être dictée par l'utilisateur. Plusieurs décodeurs ambisoniques ont été intégrés, y compris des méthodes plus motivées par la perception, telles que le décodeur ambisonique polyvalent (AllRAD) [1] et le décodeur ambisonique à conservation d'énergie (EPAD) [2]. La pondération max-RE [1] peut également être activée pour l'un ou l'autre décodeur. En outre, dans le cas de signaux d'harmoniques sphériques non idéaux en entrée (c'est-à-dire ceux qui sont dérivés de matrices de microphones physiques / simulées), l'ordre de décodage peut être spécifié pour les gammes de fréquences appropriées; où la normalisation préservant l'énergie (EP) ou préservant l'amplitude (AP) peut être sélectionnée pour maintenir la sonie entre les ordres de décodage cohérents.

Notez que lorsque les haut-parleurs sont uniformément répartis, toutes les approches de décodage mises en œuvre dans le plug-in sont équivalentes. Cela peut être démontré de manière efficace en choisissant une configuration d'enceintes de conception en T (distribution presque uniforme des points sur une sphère). Les avantages des approches de décodage par correspondance de mode (MMD), AllRAD et EPAD peuvent alors être observés pour des arrangements non uniformes (22.x par exemple).

Ce plug-in a été développé par Leo McCormack et Archontis Politis.

AmbiDRC

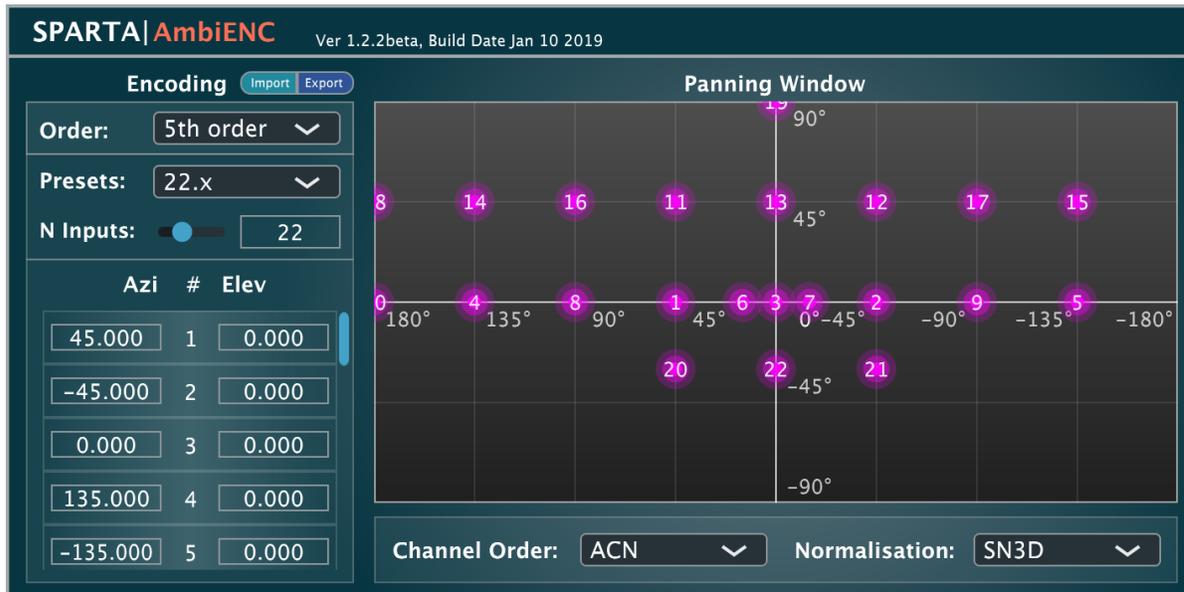


Ce plug-in est vaguement lié à [cette](#) publication.

Un compresseur de domaine dynamique (DRC) à domaine harmonique sphérique dépendant de la fréquence. Les facteurs de gain sont dérivés en analysant la composante omnidirectionnelle pour chaque bande de fréquence, qui est ensuite appliquée également aux composantes d'ordre supérieur; les propriétés spatiales des signaux d'origine restent inchangées. L'implémentation garde également une trace des facteurs de gain dépendants de la fréquence pour la composante omnidirectionnelle dans le temps, qui est ensuite tracée sur l'interface utilisateur pour obtenir un retour visuel.

Ce plug-in a été développé par Leo McCormack.

AmbiENC



Un codeur Ambisonic de base (également appelé un analyseur Ambisonic) qui prend les signaux d'entrée (jusqu'à 64 canaux) et les code en signaux harmoniques sphériques dans des directions spécifiées. Essentiellement, ces signaux harmoniques sphériques décrivent un champ sonore synthétisé dans lequel la résolution spatiale de ce codage est déterminée par l'ordre de transformation. Plusieurs préréglages ont été inclus pour plus de commodité (permettant par exemple d'encoder l'audio 22.x, etc. dans les ambisonics d'ordre 1 à 7, par exemple). La fenêtre panoramique est également entièrement pilotée par la souris et utilise une représentation équirectangulaire de la sphère pour décrire les angles d'azimut et d'élévation de chaque source.

Ce plug-in a été développé par Leo McCormack.

Array2SH



Ce plug-in est lié à [cette](#) publication.

'Array2SH' encode spatialement des signaux matriciels sphériques / cylindriques en signaux harmoniques sphériques (également appelés signaux Ambisonic ou B-Format). Le plug-in utilise des solutions analytiques, qui déterminent l'influence de la matrice sur l'estimation initiale en fonction de la fréquence et de l'ordre. Le plug-in permet à l'utilisateur de spécifier: le type de matrice (sphérique ou cylindrique), si la matrice a une enceinte ouverte ou rigide, le rayon de la matrice, le rayon des capteurs (dans le cas où ils dépassent de la matrice), les coordonnées du capteur (jusqu'à 64 canaux), la directivité du capteur (omni-dipolaire-cardioïde), la vitesse du son et l'admittance acoustique du matériau de la matrice (dans le cas de matrices rigides). Le plug-in détermine ensuite les courbes d'égalisation dépendantes de l'ordre à imposer à l'estimation initiale des signaux d'harmoniques sphériques, afin de supprimer l'influence de la matrice elle-même. Cependant, en particulier pour les ordres supérieurs, cela entraîne généralement une forte amplification des basses fréquences (y compris le bruit du capteur à ces fréquences qui l'accompagne); Par conséquent, deux approches de régularisation populaires ont été intégrées dans le plug-in, ce qui permet à l'utilisateur de faire un compromis entre l'amplification du bruit et la précision de transformation. Ces courbes d'égalisation cible et régularisée sont décrites sur l'interface utilisateur pour fournir un retour visuel. cela se traduit généralement par une forte amplification des basses fréquences (y compris le bruit du capteur à ces fréquences qui l'accompagne); Par conséquent, deux approches de régularisation populaires ont été intégrées dans le plug-in, ce qui permet à l'utilisateur de faire un compromis entre l'amplification du bruit et la précision de transformation. Ces courbes d'égalisation cible et régularisée sont décrites sur l'interface utilisateur pour fournir un retour visuel.

Le plug-in permet également à l'utilisateur d'analyser les performances de codage spatial en utilisant les mesures objectives décrites dans [8,10], à savoir: la corrélation spatiale et la différence de niveau. Ici, les matrices de codage sont appliquées à un réseau simulé, qui est décrit par des fonctions de transfert multicanal d'ondes planes pour 812 points à la surface du réseau sphérique / cylindrique. Les réponses de tableau codées résultantes devraient idéalement ressembler à des fonctions harmoniques sphériques aux points de la grille. La corrélation spatiale est ensuite obtenue en comparant les modèles de ces réponses avec les modèles des harmoniques sphériques idéales, où "1" signifie qu'elles sont parfaites et "0" complètement non corrélées; la fréquence de repliement spatial peut donc être observée pour chaque ordre, comme le point où la corrélation spatiale tend vers 0. La différence de niveau est alors la différence de niveau moyenne dans toutes les directions (différence de niveau diffuse) entre les composantes idéale et simulée. On peut observer que des limites d'amplification maximales autorisées [Max Gain (dB)] produiront des signaux plus bruyants; Cependant, cela entraînera également une plage de fréquences plus large de composants harmoniques sphériques utiles à chaque ordre. Cette analyse est principalement basée sur le code écrit pour la publication [10], qui compare les performances de différentes approches de régularisation des filtres de codage, basées à la fois sur les réponses de matrice théoriques et mesurées. Cela se traduira également par une gamme de fréquences plus large composée de composants harmoniques sphériques utiles à chaque ordre. Cette analyse est principalement basée sur le code écrit pour la publication [10], qui compare les performances de différentes approches de régularisation des filtres de codage, basées à la fois sur les réponses de matrice théoriques et mesurées. Cela se traduira également par une gamme de fréquences

plus large composée de composants harmoniques sphériques utiles à chaque ordre. Cette analyse est principalement basée sur le code écrit pour la publication [10], qui comparait les performances de différentes approches de régularisation des filtres de codage, basées à la fois sur les réponses de matrice théoriques et mesurées.

Notez que cette capacité à équilibrer l'amplification du bruit avec la précision de l'encodage spatial (pour mieux convenir à une application donnée) est très importante, par exemple: la fidélité perçue de l'audio décodé Ambisonic peut être assez médiocre si l'amplification du bruit est définie sur une valeur trop élevée. ; Par conséquent, une limite de régularisation d'amplification beaucoup plus basse est généralement utilisée dans la reproduction Ambisonics par rapport aux algorithmes de visualisation de champ sonore ou aux formateurs de faisceaux qui utilisent un post-filtrage approprié.

Pour plus de commodité, les spécifications de plusieurs modules de microphones disponibles dans le commerce ont été intégrées en tant que préréglages; notamment: Eigenmike de MH Acoustic, le module Zylia et divers modules de microphones au format A. De plus, en libérant ce plug-in, il est maintenant possible de créer / d'imprimer en 3D leur propre matrice sphérique et cylindrique, tout en disposant d'un moyen pratique d'obtenir les signaux harmoniques sphériques correspondants; par exemple, un réseau d'hydrophones à corps ouvert à quatre capsules a été présenté dans [9], qui utilisait ce plug-in Array2SH comme première étape de la visualisation et de l'auralisation d'une scène sonore sous-marine en temps réel.

Ce plug-in a été développé par Leo McCormack et Symeon Delikaris-Manias.

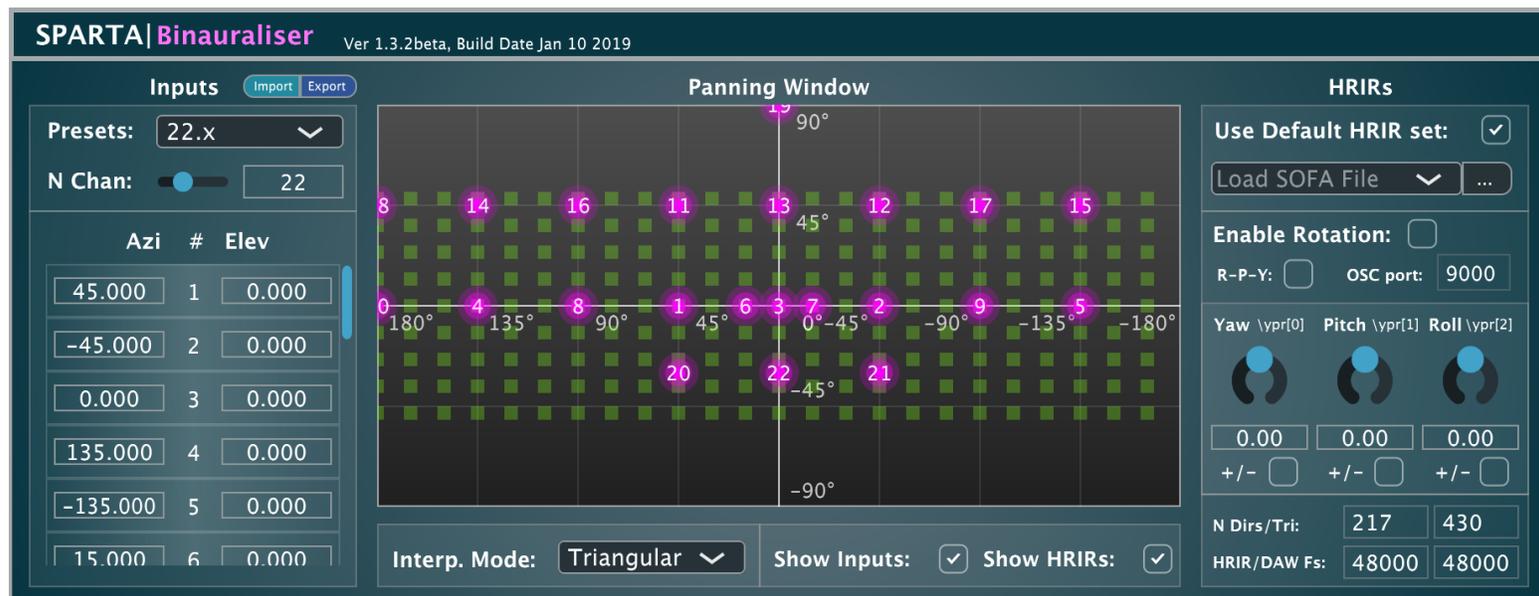
Beamformer

Un plug-in qui génère divers formateurs de faisceaux statiques et adaptatifs dans le domaine des harmoniques sphériques (jusqu'à 64 faisceaux simultanés). Le post-filtre CroPaC (Cross-Pattern Coherence) [3] peut également être appliqué pour améliorer la sélectivité spatiale des formateurs de faisceaux; particulièrement dans les environnements bruyants et / ou réverbérants.

Ce plug-in a été développé par Leo McCormack et Symeon Delikaris-Manias

(Arrive bientôt)

Binauraliser



Un plug-in qui convole l'audio d'entrée (jusqu'à 64 canaux) avec des HRTF interpolés dans le domaine temps-fréquence. Les HRTF sont interpolées en appliquant les gains de VBAP normalisés en amplitude [4] aux réponses en amplitude du HRTF et aux différences de temps inter-auriculaires (ITD) individuellement, avant d'être à nouveau combinées. Le plug-in permet également à l'utilisateur de spécifier un fichier SOFA externe pour la convolution. Des préréglages pour les formats 2D et 3D populaires sont inclus pour plus de commodité; cependant, les directions pour un maximum de 64 canaux peuvent être contrôlées indépendamment. Le suivi de la tête est également pris en charge via les messages OSC de la même manière qu'avec le plug-in Rotator.

Veuillez noter que ce plug-in ne convient que pour la convolution basée sur HRTF.

Ce plug-in a été développé par Leo McCormack et Archontis Politis.

Panner



Ce plug-in est une version légèrement modifiée du plug-in décrit dans [cette](#) publication.

«Powermap» est un plug-in qui représente l'énergie sonore relative, ou la probabilité statistique, qu'une source parvienne à la position d'écoute depuis une direction particulière, à l'aide d'un dégradé de couleurs. où le jaune indique une énergie sonore / probabilité élevée et le bleu, une énergie sonore / probabilité faible. Le plug-in intègre une variété d'approches différentes, y compris: standard basé sur le formateur de faisceau PWD (Decomposition onde onde), réponse à distorsion minimale de variance (MVDR), basé sur le pseudo spectre-Spectre à réponse de distorsion sans distorsion (MVDR), et l'algorithme de cohérence de modèle croisé (CroPaC) [3]; qui sont tous écrits pour fonctionner sur des signaux harmoniques sphériques jusqu'au 7ème ordre. Notez que l'ordre d'analyse par bande de fréquences est entièrement définissable par l'utilisateur, et des prééglages pour les matrices de microphones d'ordre supérieur ont été inclus pour la commodité (qui fournissent des valeurs de départ approximatives mais appropriées). Le plug-in utilise une grille sphérique uniformément répartie de 812 points, qui est ensuite interpolée en une carte de puissance 2D à l'aide de gains VBAP normalisés en amplitude (c'est-à-dire une interpolation triangulaire).

Notez que ce plug-in prend en charge des tailles de trame de 1024 ou 2048 uniquement. De plus, l'option 'CroPaC LCMV' est très expérimentale, vous pouvez donc voir le diable.

Ce plug-in a été développé par Leo McCormack.

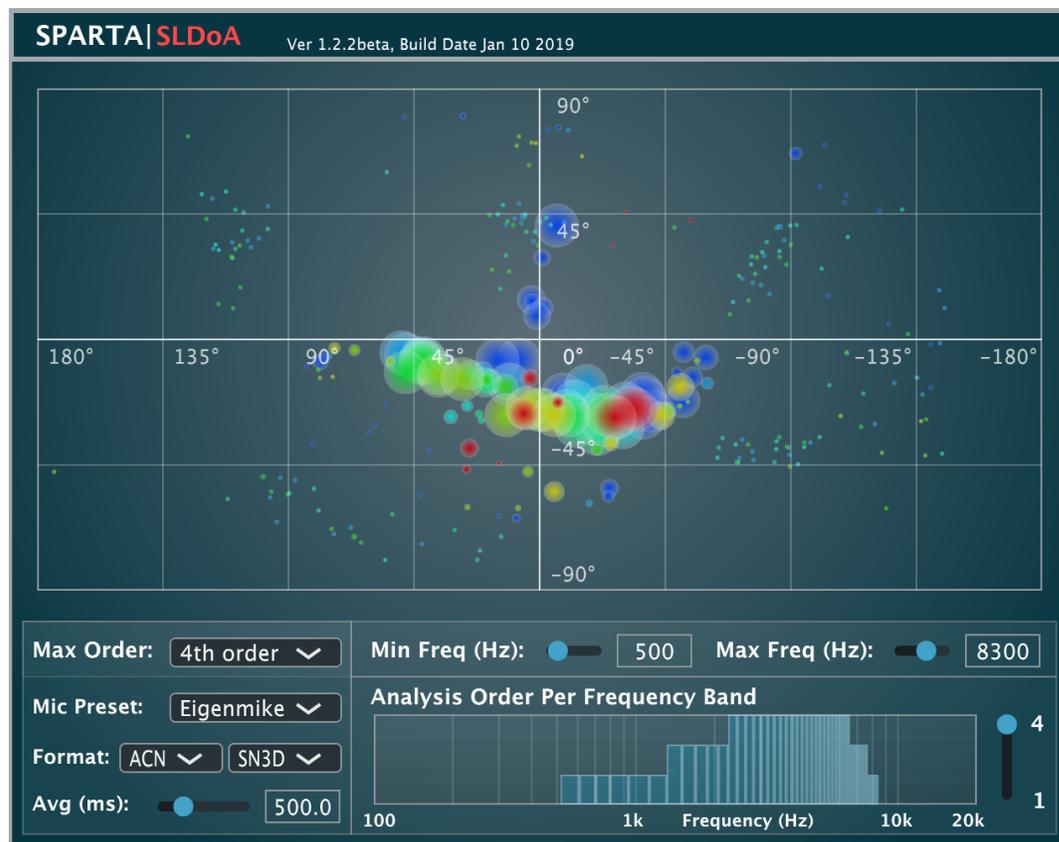
Rotateur



Ce plug-in applique une matrice de rotation harmonique sphérique [6] aux signaux harmoniques sphériques d'entrée. Les angles de rotation peuvent être contrôlés à l'aide d'un suivi de tête via des messages OSC. Configurez simplement le headtracker pour qu'il envoie un vecteur: '\ ypr [3]' au port 9000 de la CVMO (par défaut); où \ ypr [0], \ ypr [1], \ ypr [2] sont les angles yaw-pitch-roll, respectivement. Les angles peuvent également être inversés +/- afin de prendre en charge un plus grand nombre d'appareils. L'ordre de rotation (yaw-pitch-roll (par défaut) ou roll-pitch-yaw) peut également être spécifié.

Ce plug-in a été développé par Leo McCormack.

SLDoA



Ce plug-in est lié à [cette](#) publication.

Un estimateur de la direction d'arrivée (DoA) spatialement localisé. Le plug-in utilise d'abord les diagrammes de faisceau VBAP (pour les directions uniformément réparties à la surface d'une surface) afin d'obtenir des signaux zéroth et de premier ordre à polarisation spatiale, qui sont ensuite utilisés pour l'estimation vectorielle à intensité active; donc, permettant une estimation de DoA dans plusieurs secteurs limités dans l'espace pour chaque sous-bande. Les estimations des basses fréquences sont ensuite représentées par des icônes bleues, les moyennes fréquences par le vert et les hautes fréquences par le rouge. La taille de l'icône et son opacité correspondent à l'énergie du secteur, qui sont normalisés et mis à l'échelle dans l'ordre croissant pour chaque bande de fréquence. Le plug-in emploie deux fois plus de secteurs que l'ordre d'analyse, à l'exception de l'analyse de premier ordre, qui utilise l'approche traditionnelle à intensité active. L'ordre d'analyse par bande de fréquence est définissable par l'utilisateur, de même que la plage de fréquences dans laquelle analyser. Cette approche de la visualisation des champs sonores / de l'estimation DoA représente une option beaucoup plus efficace du point de vue calcul, comparée aux algorithmes intégrés au plug-in «Powermap», par exemple.

Ce plug-in a été développé par Leo McCormack et Symeon Delikaris-Manias.

UpMix

Une collection (à venir) d'algorithmes de mélange ascendant. Actuellement, le plug-in comprend uniquement l'algorithme de mixage stéréo à surround décrit dans [7], qui utilise une solution des moindres carrés pour diviser les signaux stéréo en flux directs et ambiants. Le flux direct est alors orienté vers la configuration des haut-parleurs en fonction du rapport d'amplitude entre les canaux gauche et droit, tandis que les flux ambiants sont envoyés aux haut-parleurs gauche et droit et sont également décorrélés et envoyés aux haut-parleurs arrière gauche et droit.

Ce plug-in a été développé par Leo McCormack.

(Arrive bientôt)

À propos des auteurs

- Leo McCormack: candidat au doctorat à l'Université Aalto.
- Symeon Delikaris-Manias: chercheuse postdoctorale à l'Université Aalto, spécialisée dans le traitement de réseaux de microphones compacts pour l'estimation DoA et la reproduction du champ sonore. Sa thèse de doctorat comprenait des travaux sur l'algorithme CroPaC (Cross-Pattern Coherence), qui est un post-filtre spatial optimisé pour les environnements à fort bruit / réverbération.
- Archontis Politis: chercheur post-doctorat à l'Université Aalto, spécialisé dans l'enregistrement et la reproduction spatiaux du son, l'analyse de scènes acoustiques et le traitement de réseaux de microphones.
- Ville Pulkki: Professeur à l'Université Aalto, connu pour ses comportements VBAP, SIRR, DirAC et excentrique.

Références

- [1] Zotter, F., Frank, M. (2012). Panoramique et décodage ambisoniques complets. *Journal de la société d'ingénierie audio*, 60 (10), 807-820.
- [2] Zotter, F., Pomberger, H., Noisternig, M. (2012). Décodage Ambisonic à conservation d'énergie. *Acta Acustica United avec Acustica*, 98 (1), 37-47.
- [3] Delikaris-Manias, S., Pulkki, V. (2013). Algorithme de cohérence croisée pour les applications de filtrage spatial utilisant des réseaux de microphones. *Transactions IEEE sur le traitement audio, vocal et linguistique*, 21 (11), 2356-2367.
- [4] Pulkki, V. (1997). Positionnement virtuel de la source sonore à l'aide du vecteur Panoramique avec amplitude de base. *Journal de la société d'ingénierie audio*, 45 (6), 456-466.
- [5] Laitinen, M., Vilkkamo, J., K. Jussila, A. Politis, Pulkki, V. (2014). Gain normalisé en panoramique d'amplitude en fonction de la fréquence et de la réverbération de la pièce. 55ème Conférence internationale de l'AES. Helsinki, Finlande.
- [6] Ivanic, J., Ruedenberg, K. (1998). Matrices de rotation pour les harmoniques sphériques réelles. Page Détermination directe par récurrence: Ajouts et corrections. *Journal de physique chimie A*, 102 (45), 9099-9100.
- [7] Fallor, C. (2006). Lecture de signaux stéréo par plusieurs haut-parleurs. *Journal de la société d'ingénierie audio*, 54 (11), 1051-1064.
- [8] Moreau, S., Daniel, J., Bertet, S. (2006). "Enregistrement de champs sonores 3D avec mesures objectives ambisonics d'ordre supérieur et validation du microphone sphérique", dans *Audio Engineering Society Convention 120*, Audio Engineering Society
- [9] Delikaris-Manias, S., L. McCormack, I. Huhtakallio, et V. V. Pulkki (2018). "Audio spatial sous-marin en temps réel: une étude de faisabilité" dans *Audio Engineering Society, Convention 144*, Ingénierie audio. Société.
- [10] Politis, A., Gamper, H. (2017). "Comparaison de filtres de codage harmoniques sphériques basés sur des mesures et modélisés pour réseaux de microphones sphériques. Dans l'atelier IEEE sur les applications du traitement du signal à l'audio et à l'acoustique (WASPAA).