

Rapport Recherche bibliographique

Convertir une image de profondeur en une représentation sonore

Jieli XIA
Christine CHEN
Chanborina TIENG

12/10/2023

MU5EEB12 : Designing Mechatronics Systems
Sorbonne université, Master IPS
MSR

Comment convertir l'image en sons ?

Ce projet cherche à aider les aveugles à se déplacer dans l'espace et éviter les obstacles en leur envoyant des informations sonores. Nos principaux axes de recherche vont être de savoir comment traiter une image de profondeur récupérer à l'aide d'une caméra RGB-D et de traduire l'information pour qu'une personne puisse la comprendre rien qu'avec son audition.

Plusieurs questions se posent :

- Comment traiter l'image ? Comment récupérer les données importantes de l'image, les cibler ?
- Comment émettre un signal sonore et qu'elle soit compréhensible par l'utilisateur ?
- Comment l'optimiser pour que son utilisation soit accessible pour le traitement par l'ordinateur et l'utilisateur ?

D'autres études scientifiques se sont portées sur ce sujet et nous allons voir quelles réalisations ont été apporter de ce domaine.

Trois articles nous ont particulièrement intéresser étant très proche de notre sujet d'études et pertinentes. Nous allons donc apporter une analyse sur chacun de ses textes.

Sommaire :

I. Article 1	p4
II. Article 2	p6
III. Article 3	p8
IV. Conclusion	p11
V. Bibliographies	p12

I – Article 1 : Navigating from a Depth Image Converted into Sound [1].

Appl Bionics Biomech. 2015; Published online 2015 Feb 5.

Chloé Stoll, Richard Palluel-Germain, Vincent Fristot, Denis Pellerin, David Alleysson and Christian Graff.

1. Résumé :

L'article décrit une étude visant à évaluer l'efficacité d'un dispositif portable appelé MeloSee, conçu pour aider les personnes aveugles à naviguer dans des environnements inconnus. Ce dispositif utilise un capteur de profondeur de type Kinect pour convertir une image de profondeur en une mélodie en temps réel. Ce capteur est monté sur la tête de l'utilisateur pour capturer des informations en profondeur de l'environnement. La distance par rapport au capteur est traduite en intensité sonore, modulée latéralement en stéréo, et la hauteur est représentée par la hauteur tonale. Vingt et un adultes aveuglés ont été testés sur quatre chemins différents lors de deux sessions séparées d'une semaine.

Les résultats ont montré que les participants ont appris à utiliser le système pour naviguer dans des chemins nouveaux et familiers. La performance, mesurée en termes de temps de déplacement et d'erreurs, s'est améliorée d'une semaine à l'autre. De plus, les participants ont réussi à effectuer une tâche secondaire tout en utilisant le dispositif, montrant ainsi la robustesse de leur apprentissage. Les auteurs concluent que l'utilisation de capteurs de type Kinect pour mettre en œuvre des dispositifs de substitution sensorielle est prometteuse, bien que limitée à une utilisation en intérieur et inefficace à très courte distance.

2. Méthode et résultats :

Les participants ont été aveuglés et équipés du dispositif MeloSee. Le capteur Kinect a converti une image de profondeur en une mélodie, et les participants ont été testés sur leur capacité à naviguer le long de trajets inconnus. Des tâches de distraction ont également été introduites pour évaluer l'effet sur la navigation. L'article mentionne des stimuli tactiles, sous forme de bourdonnements sinusoïdaux à 100 Hz d'une durée de 200 ms chacun, sont émis et doivent être détectés par le participant. Il doit répondre uniquement aux bourdonnements doubles en tapant sur sa cuisse avec sa main libre tout en naviguant.

Les résultats ont montré une amélioration de la performance des participants au fil du temps, indiquant un apprentissage à court terme et à long terme. La tâche de distraction n'a eu qu'un impact mineur sur la performance de navigation.

3. Critique :

L'article présente une méthodologie claire et des résultats significatifs. Cependant, la taille de l'échantillon pourrait être discutée, et il pourrait être intéressant d'explorer davantage les limitations du dispositif, telles que son incapacité à fonctionner en extérieur car le capteur infrarouge est perturbé par les signaux plus forts du soleil. Malgré tout, les résultats suggèrent que MeloSee peut être une aide prometteuse pour la navigation des personnes aveugles. Des études supplémentaires pourraient explorer davantage son efficacité par rapport à d'autres dispositifs d'aide à la navigation.

II – Article 2 : Sensory Substitution: The Spatial Updating of Auditory Scenes “Mimics” the Spatial Updating of Visual Scenes [2].

*Front Behav Neurosci. 2016; Published online 2016 Apr 21.
Achille Pasqualotto and Tayfun Esenkaya.*

1. Résumé :

Ce texte explore l'utilisation de la substitution sensorielle visuel-auditif pour transmettre des informations visuelles par l'audition. Il mentionne spécifiquement l'utilisation du logiciel "The vOICe" qui convertit des images visuelles capturées par une caméra vidéo en paysages sonores équivalents. Développé par le Dr Peter Meijer, ce système repose sur une caméra pour capturer des images en temps réel, puis traduit ces informations visuelles en sons. Les pixels de luminosité élevée produiront un son plus fort que les pixels de luminosité faible, les pixels à gauche seront joués avant ceux à droite, et les pixels en haut auront une hauteur de son plus élevée que ceux en bas. Ainsi, les images visuelles sont sonifiées en utilisant trois paramètres : le volume, le temps et la hauteur de son.

L'étude a été réalisée avec des participants voyants les yeux bandés, utilisant la substitution sensorielle pour apprendre la position spatiale de formes simples disposées sur le sol. Les résultats suggèrent que même si l'apprentissage était initialement effectué de manière égocentrique, la représentation spatiale de la scène a été mise à jour de manière allocentrique.

On rappelle que dans un contexte egocentrique, la représentation spatiale est centrée sur l'observateur. Les relations spatiales sont déterminées par rapport à la position de l'observateur lui-même. Tandis que dans un contexte allocentrique, la représentation spatiale est centrée sur les relations entre les objets eux-mêmes, indépendamment de la position de l'observateur. Cela implique que l'information auditive a été convertie en une représentation spatiale équivalente à celles obtenues par d'autres modalités sensorielles comme la vision ou la somesthésie. De plus, cela pourrait avoir des implications pratiques pour améliorer les méthodes de formation avec les dispositifs de substitution sensorielle.

L'étude a également souligné la nécessité de former les participants à reconnaître les images sonifiées et a souligné le potentiel de la substitution sensorielle pour la représentation spatiale équivalente, indépendamment de la modalité sensorielle utilisée. Enfin, les résultats ont des implications pratiques pour améliorer les méthodes de formation avec les dispositifs de substitution sensorielle, en particulier en fournissant des informations allocentriques pour faciliter la mise à jour spatiale.

2. Méthode et résultats :

Dix-huit participants voyants ont été recrutés, formés à reconnaître des images via The vOICe, et ont ensuite utilisé le dispositif pour explorer un environnement avec des images disposées au sol. Une carte représentant la disposition spatiale des images a été utilisée pour déclencher la mise à jour spatiale. Les erreurs de pointage et les temps de réaction ont été enregistrés pendant une tâche de prise de perspective.

Les résultats ont montré une mise à jour spatiale réussie des scènes auditives apprises de manière égocentrique vers une représentation allocentrique, similaire aux résultats obtenus avec d'autres modalités sensorielles. Les participants ont présenté des erreurs de pointage et des temps de réaction plus faibles lorsqu'ils imaginaient des perspectives alignées avec les axes allocentriques de la scène.

3. Critique :

Les forces incluent la démonstration d'une mise à jour spatiale réussie avec la substitution sensorielle, suggérant que ces dispositifs peuvent être utilisés plus efficacement avec l'ajout d'informations allocentriques. Cependant, les limites comprennent l'utilisation de participants voyants plutôt que de personnes malvoyantes, et la possibilité que la représentation mentale ait pu jouer un rôle. Les résultats suggèrent que la substitution sensorielle peut bénéficier de l'ajout d'informations allocentriques pour améliorer la représentation spatiale. Cela pourrait avoir des implications pratiques pour l'utilisation de tels dispositifs par des personnes malvoyantes, en réduisant la nécessité d'une formation prolongée et en améliorant l'efficacité de la représentation spatiale.

III – Article 3 : Spatial Competence and Brain Plasticity in Congenital Blindness via Sensory Substitution Devices [3].

*Front Neurosci. 2020; Published online 2020 Jul 30.
Daniel-Robert Chebat, Fabien C. Schneider and Maurice Ptito.*

1. Résumé :

L'article explore les mécanismes de plasticité cérébrale chez les individus atteints de cécité congénitale, mettant en évidence la capacité du cerveau à compenser le manque d'informations visuelles en réinterprétant les informations tactiles et auditives. Il souligne que bien que la cécité congénitale puisse retarder l'acquisition des compétences spatiales, la plasticité cérébrale déclenchée par l'entraînement permet éventuellement d'atteindre une compétence spatiale. Les dispositifs de substitution sensorielle (SSD) jouent un rôle clé en traduisant les indices visuels en informations tactiles ou auditives, déclenchant ainsi la plasticité cérébrale.

Il explore également les différentes SSD, telles que le Tongue Display Unit (TDU), l'EyeCane, l'EyeMusic et vOICe, qui transforment les informations visuelles en stimuli tactiles ou auditifs. Malgré leurs différences, ces dispositifs nécessitent tous une intégration active de l'information visuelle par l'utilisateur, impliquant une forme d'apprentissage synesthésique.

L'impact de la privation sensorielle sur la plasticité cérébrale, le recrutement de régions cérébrales pour des tâches spécifiques via les SSD, et les changements anatomiques et fonctionnels associés sont explorés. En dépit de la réduction volumétrique de certaines régions cérébrales, les personnes aveugles congénitales conservent des capacités de navigation et de représentation spatiale. Les résultats suggèrent que la cécité congénitale n'entraîne pas un déficit cognitif spatial, mais plutôt un déficit perceptif, et que les capacités spatiales peuvent être améliorées avec l'entraînement.

L'article explore ensuite comment le cerveau traite l'information spatiale. Il met en lumière les régions cérébrales cruciales pour la navigation, la mémoire spatiale et la reconnaissance des scènes tel que l'hippocampe, le complexe parahippocampique et le cortex pariétal.

Pour finir, différentes stratégies de perception spatiale et l'utilisation de SSDs sont explorées. Elle examine comment les individus aveugles utilisent des informations tactiles ou auditives pour détecter et éviter les obstacles, discutant de l'efficacité de ces stratégies dans des contextes réels. De plus, elle aborde les avancées technologiques des SSDs, soulignant leur impact potentiel sur la compétence spatiale des personnes aveugles. En conclusion, la section évoque les perspectives futures des SSDs, suggérant des applications potentielles dans des environnements naturels et soulignant la nécessité d'améliorations technologiques pour une utilisation plus répandue.

2. Méthode et résultats :

L'article examine l'impact de la privation visuelle sur la compétence spatiale, en particulier chez les individus aveugles de naissance. Les chercheurs explorent différentes théories sur l'acquisition de la compétence spatiale chez les aveugles, notamment les modèles cumulatif, persistant et convergent. La recherche utilise des modèles tels que le modèle convergent de Schinazi et al. (2016) pour argumenter en faveur de la possibilité d'acquérir une compétence spatiale malgré la privation visuelle. Des études sur l'utilisation de dispositifs de substitution sensorielle (SSD), tels que le vOICe, sont également examinées.

Concernant les résultats plusieurs points ressortent. Tout d'abord, il soutient l'idée que la privation visuelle ne conduit pas nécessairement à une incapacité totale d'acquérir une compétence spatiale. Le modèle convergent suggère que, bien que cela puisse prendre plus de temps, les individus aveugles peuvent atteindre un niveau de compétence spatiale similaire à celui des individus voyants avec suffisamment d'expérience.

3. Critique :

Une revue approfondie des théories existantes sur l'acquisition de la compétence spatiale chez les aveugles est présentée et ainsi que des modèles tels que le modèle convergent, cumulatif et persistant. Cependant, certaines limites peuvent inclure des biais potentiels dans la sélection des études examinées et des conclusions tirées. De plus, la généralisation des résultats peut être limitée par la variabilité individuelle parmi les personnes aveugles. Les résultats précédemment cités suggèrent que, malgré la privation visuelle, les individus aveugles peuvent

développer des compétences spatiales significatives en utilisant des mécanismes de compensation, tels que l'écholocalisation. Les implications de ces résultats pourraient informer le développement de dispositifs de substitution sensorielle et souligner l'importance d'explorer davantage les mécanismes de la plasticité cérébrale chez les personnes aveugles.

IV – Conclusion :

En conclusion, les trois sources examinées apportent des éclairages complémentaires sur le sujet de la conversion d'une image de profondeur en son dans le contexte de la navigation pour les personnes aveugles. L'étude sur le MeloSee démontre que l'utilisation d'un capteur de profondeur Kinect monté sur la tête des utilisateurs, offre une méthode prometteuse. Les résultats positifs, montrant une amélioration de la performance au fil du temps, soulignent la faisabilité et la robustesse de l'apprentissage de l'utilisation de ce dispositif.

La deuxième source explore l'utilisation de la substitution sensorielle visuel-auditif avec le logiciel "The vOICe". Elle met en lumière la manière dont cette technologie permet de sonifier des informations visuelles, offrant aux utilisateurs la possibilité de créer des représentations spatiales par d'autres modalités sensorielles. L'étude avec des participants les yeux bandés, utilisant les SSDs suggère une mise à jour spatiale efficace, notamment de manière allocentrique.

La troisième source explore la plasticité cérébrale chez les personnes atteintes de cécité congénitale, mettant en évidence la capacité du cerveau à compenser la vision en réinterprétant les informations tactiles et auditives. Elle souligne également l'importance des dispositifs de substitution sensorielle, tels que le Tongue Display Unit (TDU), dans la traduction des indices visuels en stimuli tactiles, déclenchant ainsi la plasticité cérébrale.

En combinant ces perspectives, il apparaît clairement que la conversion d'une image de profondeur en son, utilisée dans des dispositifs tels que MeloSee, ainsi que la substitution sensorielle visuel-auditif, jouent un rôle crucial dans le développement de compétences spatiales pour les personnes aveugles. Ces approches innovantes exploitent la plasticité cérébrale pour compenser la privation visuelle et ouvrent des perspectives prometteuses pour l'amélioration de la mobilité et de la perception spatiale chez les individus aveugles. Les avancées technologiques dans le domaine des SSDs semblent être des points clés pour cette amélioration, ouvrant la voie à des applications plus étendues dans des environnements naturels. Cependant, des défis subsistent, notamment en ce qui concerne l'adaptabilité et l'efficacité à très courte distance. Des améliorations continuent dans la technologie des SSDs, ainsi que des efforts pour comprendre et améliorer l'apprentissage des utilisateurs, seront cruciaux pour maximiser l'impact de ces dispositifs dans le futur.

V – Bibliographies :

- [1] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4745448/>
- [2] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4838627/>
- [3] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7406645/>