

TITRE DU SUJET :

Communication {homme/véhicule} intuitive et multimodale : Nouveaux paradigmes de synthèse audio-tactile

TUTEUR INDUSTRIEL RESPONSABLE DE LA THESE :

Nom, Prénom, email. : DENJEAN Sébastien, sebastien.denjean@stellantis.com

Direction, Département, Service : DRIA/SD2T/SCHF

DIRECTEUR DE THESE :

Nom, Prénom, email : YSTAD Solvi, ARAMAKI Mitsuko

Laboratoire, équipe adresse :

PRISM « Perception, Représentations, Image, Sound, Music » (UMR 7061 CNRS-AMU)

Date de démarrage souhaité : 09/2022

DEVELOPPEMENT DU PROJET DE RECHERCHE

Contexte

Nos véhicules deviennent de plus en plus interactifs et communicants. Les retours d'information vers le conducteur ne cessent d'augmenter, et représentent différents niveaux de criticité, pouvant aller de la simple notification aux alertes graves. Cette augmentation du nombre de retours d'information peut aller jusqu'à la surcharge d'information pour l'utilisateur et nécessite de créer des retours pertinents et intuitifs pour les conducteurs, qui informent au juste niveau de criticité. Dans le même temps, les modalités de retour ne cessent de se développer, qu'elles soient sonores, vibratoires ou lumineuses. L'utilisation de ces différentes solutions permet d'exploiter la multimodalité pour informer au mieux l'utilisateur, d'une part en donnant du sens dans les messages véhiculés par l'association et la scénarisation des informations dans les différentes modalités et d'autre part en ajustant les capacités d'attirer son attention et la gêne en fonction de la criticité du message à transmettre.

De plus, avec l'émergence de l'automatisation des véhicules automobiles, les conducteurs sont amenés à devenir superviseurs et non plus acteurs lors des phases de conduite autonome, les rendant moins vigilants et moins conscients de la situation qui les entoure. La multimodalité se présente aussi comme une solution idéale pour attirer/ramener l'attention du conducteur sur une information ciblée en jouant sur les spécificités de chaque modalité pour attirer au mieux l'attention et porter du sens : un message visuel portera naturellement du sens, quand un son pourra par exemple attirer l'attention quelle que soit la focalisation d'attention du conducteur ou une vibration pourra permettre d'accéder dans sa sphère personnelle... Dans ce contexte, la multimodalité peut permettre de répondre à ces problématiques en proposant des signaux plus ou moins alertants et porteurs de sens. Nous chercherons aussi à développer des solutions de remédiation multimodales permettant de maintenir le superviseur dans les meilleures conditions d'attention lorsque c'est nécessaire, par exemple dans le cadre d'une reprise en main programmé d'un véhicule autonome.

Objectifs

L'objectif du projet de thèse est de développer un véritable système de **communication non-verbale et multimodale** entre le véhicule et le conducteur. Nous privilégierons les développements dans les

modalités auditives et vibratoires afin de travailler les leviers qui leur permettent de donner du sens, mais ces solutions seront mises en scène et évaluées dans un contexte résolument multimodal en les associant à des solutions visuelles de lighting et retours IHM.

Cette communication se veut la plus naturelle possible, intuitive et efficace, aussi bien dans le cas d'une conduite manuelle que pour remettre le superviseur en condition de reprendre le véhicule dans le cadre d'une conduite autonome. Cette communication sera développée de sorte à pouvoir moduler le niveau de criticité du retour d'information allant de la 'simple' notification à l'alerte d'urgence.

Les solutions sonores seront restituées par le système audio du véhicule en exploitant les solutions innovantes à l'étude au sein de la direction scientifique (spatialisation, haut-parleurs appui-tête). Les solutions tactiles seront focalisées sur des systèmes de restitution de vibration dans le siège. Nous nous appuyerons pour cela sur les démonstrateurs développés au sein de la direction scientifique embarquant jusqu'à 8 actuateurs vibrants répartis dans l'assise et le dossier et contrôlables en niveau et fréquences indépendamment. Ces solutions seront intégrées dans un démonstrateur et associées aux retours d'informations visuels (messages combiné, écran ou VTH, lighting habitacle) et évaluées dans un contexte situé représentatif des situations réelles de conduite.

D'un point de vue fondamental, nous nous proposons en effet d'aborder ici la perception multimodale par une approche écologique et cognitivement située. Il s'agit ainsi d'aborder le problème dans sa complexité dans le cadre d'un paradigme de type analyse par synthèse. Dans la lignée de cette approche écologique, les chercheurs de PRISM ont acquis une expertise dans la génération de stimuli calibrés basés sur la définition d'invariants perceptifs liés aux évocations. De même, la prise en compte du contexte et de l'état (cognitif ou émotionnel) de l'auditeur est fondamentale pour que la communication soit 'établie', i.e., les informations sont correctement perçues et comprises par l'auditeur malgré des perturbations extérieures.

Parmi les points centraux à un véritable dialogue {homme/véhicule}, la notion de proximité (zone intime) joue un rôle prépondérant. Cette notion a rarement été utilisée dans le domaine du sonore, pourtant on peut raisonnablement faire l'hypothèse qu'une voix chuchotée à l'oreille émergera plus facilement (au sens cognitif du terme) d'un contexte bruité qu'une voix parlée à même niveau de sonie. Dans le même ordre d'idée, l'interaction tactile est par construction liée à un phénomène de proximité, donc susceptible d'induire une réactivité accrue. La recherche proposée vise à la modélisation et à la construction d'outils de « synthèse multimodale » basés sur la génération de stimuli sonores et tactiles calibrés et contrôlés à haut niveau afin d'optimiser le dialogue {homme/véhicule}. Les nouveaux paradigmes mis en œuvre constitueront une véritable rupture méthodologique en assurant à chaque modalité son identité propre plutôt qu'un simple effet d'amplification. Nous ne nous attacherons en effet pas seulement à dupliquer l'information de façon synchrone en visuel, audio et vibratoire, mais chercherons à exploiter les capacités propres à chaque modalité dans le cadre d'une scénarisation où les différentes modalités peuvent se répondre et se compléter afin de véhiculer le bon niveau d'alerte.

Cette communication multimodale sera mise en œuvre et évaluée sur une série de cas d'usage, allant de la notification aux alertes graves. La criticité de ces dernières sera modulée et mise en scène avec deux situations spécifiques au cas du véhicule autonome : l'alerte d'urgence pour une reprise en main immédiate (quelques secondes), et la remédiation pour ramener le conducteur dans les meilleures conditions d'attention dans le cas d'une reprise en main programmée (quelques minutes). Ces solutions seront mises en place sur des moyens représentatifs en situations, et leur efficacité sera évaluée d'un point de vue perceptif et UX.

Organisation des travaux de recherche

Afin de répondre aux besoins industriels en termes de solutions sécuritaires et de remédiation, nous nous intéresserons aux cas d'usage suivants :

- La notification et les sons fonctionnels, retours se voulant purement informatifs et ne nécessitant pas d'action de l'utilisateur,
- L'alerte (notion de temps court) nécessitant une réponse rapide de l'utilisateur dans une situation potentiellement stressante et chargée d'un point de vue cognitif,
- La régulation continue (notion de temps long) pour laquelle l'attention de l'utilisateur est surveillée afin d'être maintenue dans un état optimal.

Dans le cas de l'alerte, nous chercherons à exploiter les capacités de la multimodalité (audio-vibrotactile) afin de proposer des alertes capables d'attirer l'attention d'un superviseur distrait et de l'informer au mieux sur la situation ou les actions à réaliser. Associée au cas de la notification, cette situation d'alerte nous permettra de moduler la criticité des retours proposée sur tous les degrés d'alertes rencontrés en situation.

Dans le cas de la remédiation, nous nous appuierons sur des solutions existantes de monitoring de l'état du conducteur (rythme cardiaque, respiratoire, conductance électrodermale...) afin de proposer des solutions multimodales permettant de modifier l'état d'attention du conducteur afin de le remettre dans des conditions optimales dans les situations qui le nécessitent. Ces mesures physiologiques permettront de valider l'apport de ces solutions.

Nous proposerons des solutions multimodales adaptées à ces situations sur la base d'études fondamentales permettant de mettre en évidence les attributs spécifiques à chacune des modalités convoquées. Les solutions multimodales proposées seront ensuite implémentées en environnement représentatif avant d'être évaluées au moyen de tests perceptifs réalisés en situation immersive (simulateur de conduite ou véhicule réel). L'efficacité des solutions proposées sera évaluée en prenant en compte les différences interindividuelles des utilisateurs (morphologie, seuils de perception, ...) et des potentiels masqueurs ambiants (bruit et vibrations moteur/route), en environnement représentatif.

Bibliographie

- Aramaki M., Besson M., Kronland-Martinet R., Ystad S. (2011). Controlling the Perceived Material in an Impact Sound Synthesizer. *IEEE Trans. Audio, Speech and Language Processing*, 19(2):301–314
- Arnal L. (2015) Human Screams Occupy a Privileged Niche in the Communication Soundscape. *Current Biology*, 25:2051-2056.
- Audry E., Garcia J.(2019). Congruent Audio-Visual Alarms for Supervision Tasks. In Proc. 25th International Conference on Auditory Display, Jun 2019, Newcastle, United Kingdom.
- Baldwin C.L. (2007) Acoustic and Semantic Warning Parameters Impact Vehicle Crash Rates. In Proceedings of 13th International Conference on Auditory Display, Montreal, Canada, June 26-29, 2007, 143-145.

- Cancar L., Diaz A., Barrientos A., Travieso D., Jacobs D. (2013). Tactile-Sight: A Sensory Substitution Device based on Distance-Related Vibrotactile Flow. *International Journal of Advanced Robotic Systems*, 14, 21.
- Danna J., Fontaine M., Paz-Villagrán V., Gondre C., Thoret E., Aramaki M., Kronland-Martinet R., Ystad S., Velay JL.(2015). The Effect of Real-Time Auditory Feedback on Learning New Characters. *Human Mov. Sc.*, 43:216-228.
- Dehais F., Causse M., Vachon F., Régis N., Menant E., Tremblay S. (2014). Failure to Detect Critical Auditory Alerts in the Cockpit: Evidence for Inattentive Deafness. *Human Factors*, 56(4):631-44.
- Dehais F., Roy RN., Scannella S. (2019) Inattentive Deafness to Auditory Alarms: Inter-Individual Differences, Electrophysiological Signature and Single Trial Classification. *Behavioural Brain Research*, 360:51-59.
- Durantin G., Dehais F., Gonthier N., Terzibas C., Callan D.E. (2017). Neural Signature of Inattentive Deafness. *Human Brain Mapping*, 38(11):5440–5455.
- Ernst M., Banks M. (2002). Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. *Nature*, 415:429-33.
- Fagerlönn J.(2011) Making Auditory Warning Signals Informative: Examining the Acceptance of Auditory Icons as Warnings Signals in Trucks. *Proc.6th Int Driving Symposium on Human Factors in Driver Assessment*, 2011:96-101.

INTITULE ET RESUME DU SUJET EN ANGLAIS :

Intuitive and multimodal {human/vehicle} communication: New paradigms of audio-tactile synthesis

The aim of the thesis project is to develop a real system of non-verbal and multimodal communication between the vehicle and the driver. Since the vision is dedicated to the driving task, we will favor communication centered on hearing and tactile modalities. This communication from the vehicle to the driver is intended to be as natural as possible, intuitive and effective.

Among the central points to a real dialog {man/vehicle}, the notion of proximity (intimate area) plays a leading role. This concept has rarely been used in the field of sound, yet it can reasonably be assumed that a whispered voice in the ear will emerge more easily (in the cognitive sense of the term) from a noisy context than a voice spoken at the same sound level. In the same vein, tactile interaction is constructed as a function of proximity, which can lead to increased responsiveness. The proposed research aims at modeling and constructing "multimodal synthesis" tools based on the generation of high-level calibrated and controlled so-called and tactile stimuli to optimize the {man/vehicle} dialog. The new paradigms implemented will be a true methodological break by ensuring that each modality has its own identity rather than merely an amplification effect.

We will focus on the following cases of use: alerting and remediation. We will propose multimodal solutions adapted to these situations on the basis of fundamental studies to highlight the attributes specific to each of the modalities involved. The proposed multimodal solutions will be evaluated by means of perceptual tests carried out in an immersive situation (driving simulator or real vehicle).