Proposition de thèse :

Détection d'événements sonores par un système interface/traitement neuromorphique à faible consommation

A new Neuromorphic Interface for Detection and Classification of Acoustic Scenes and Events

I) CONTEXTE DE LA THESE

Dans un certain nombre d'applications liées à la capture de l'information, les résolutions et les traitements associés sont tellement importants qu'il parait très difficile de les réaliser sans recourir à une consommation énergétique en conséquence. Cela limite donc le développement de systèmes autonomes en énergie ainsi que leur déploiement en réseau. Cependant en étudiant l'activité d'un grand nombre de capteurs, il apparait que les besoins en traitement n'apparaissent que lorsqu'une situation particulière est identifiée : arythmie pour un capteur ECG, fonctionnement anormal pour la surveillance d'un équipement industriel, ou encore lors de la survenue d'un évènement sonore (le projet se positionnera sur cette dernière application).
Il apparait alors plus judicieux de développer deux systèmes complémentaires :

- un système dit "on-demand" qui va effectuer les traitements nécessaires lorsque la situation "particulière" est identifiée. Ce système aura une conception plus classique. N'étant activé que périodiquement, le critère de consommation est relâché au profit de contraintes de performances plus sévères.
- un système dit "always-on" qui va détecter, analyser et classifier la situation dans lequel se trouve le capteur. C'est lui qui va déclencher le fonctionnement du système "on-demand". Sa faible consommation est impérative.

C'est sur ce dernier système que le projet compte se focaliser. D'autant plus que pour certaines applications dont le rôle est simplement d'alerter ou de prévenir de la présence d'un événement, le système « always-on » peut à lui seul réaliser la fonction.

II) POSITIONNEMENT

Le projet, bien qu'adapté à plusieurs types de signaux, souhaite se positionner sur la **captation et le traitement de signaux audios** attestant de la présence d'un évènement**.** Le traitement se résumant dans un premier à temps des opérations de classification. Une communauté récente (2013) s'est d'ailleurs spécialisée dans ce domaine (Detection and Classification of Acoustic Scenes and Events [1]. Ceci permet de pouvoir se comparer aux solutions déjà développées en termes d'approche architecturale.

Certains systèmes de détection d'évènements sonores peuvent être très simples comme par exemple le franchissement d'un seuil attestant de la présence d'un signal. Néanmoins la détection de situations plus complexes (détection d'un message audio) nécessite immédiatement une certaine sophistication. Dans la mesure où ce message va avoir tendance à apparaitre souvent, un apprentissage des situations à détecter peut permettre de réduire drastiquement la complexité du traitement. Les systèmes à base de **réseaux de neurones** s'avèrent donc particulièrement intéressants pour effectuer les traitements dans ce type de système. Mais dans la plupart de ces

systèmes, **l'interface analogique** est la plupart du temps négligée ou sinon conçue de façon trop simpliste. Classiquement, le signal passe directement dans un convertisseur analogique numérique et est ensuite traité par un réseau de neurones (apprentissage profond par exemple) ou traitement de type software [2]. Cette solution est évidemment fonctionnelle mais la plupart du temps la consommation électrique du CAN domine excessivement et le traitement est lourd. Il y a donc un intérêt certain à **concevoir conjointement le système de traitement neuromorphique et son interface** afin d'alléger la classification, de réduire drastiquement la consommation en optimisant la nature de l'information à traiter. **L'originalité de la proposition se situe dans l'articulation de ces deux éléments**.

Aujourd'hui des techniques de conversion et de traitement dans le domaine temporel permettent d'obtenir des interfaces très numérisées [3] qui peuvent ainsi fonctionner à des très basses tensions d'alimentation. La basse tension peut ainsi être utilisée aussi bien sur la partie traitement que sur la partie interface notamment en favorisation la conception utilisant des transistors MOS fonctionnant sous le seuil [4]. Egalement tout type d'encodage et de prétraitement peut alors être envisagé. Il y a donc un intérêt stratégique à co-concevoir le traitement du signal et l'interface du système : en répartissant de façon optimale le traitement entre analogique et numérique, en rapprochant les méthodes de conception (numérique et analogique) et enfin en utilisant **un format d'encodage commun aux deux traitements**.

III) OBJECTIFS DE LA THESE

Le projet vise ainsi à développer, dans l'architecture représentée sur la figure 1, le système « always-on ». Le travail consiste donc à réaliser une co-conception interface/réseau de neurones spécifique. Elle s'appuie sur la génération par l'interface de signaux dans un encodage Pulse Density Modulated ou Pulse Width Modulated [5]. Ce type de signal est parfaitement adapté au traitement par un réseau de neurones à impulsions [6]. Un réseau de neurones à cliques utilisé comme une mémoire associative peut également convenir [7]. Rappelons que le traitement de l'information consiste ici des opérations de classification. Un **prototype sera réalisé sous forme de circuit intégré**. Par ailleurs le système devra avoir une contrainte en énergie forte. Une consommation très faible de l'ordre de la dizaine de µW est particulièrement visée. Dans ces conditions la récupération d'énergie radiofréquence [8] peut être envisagée pour alimenter le système. Le domaine applicatif associé est l'analyse de signaux audios dans l'optique de détecter un évènement aussi complexe soit-il.

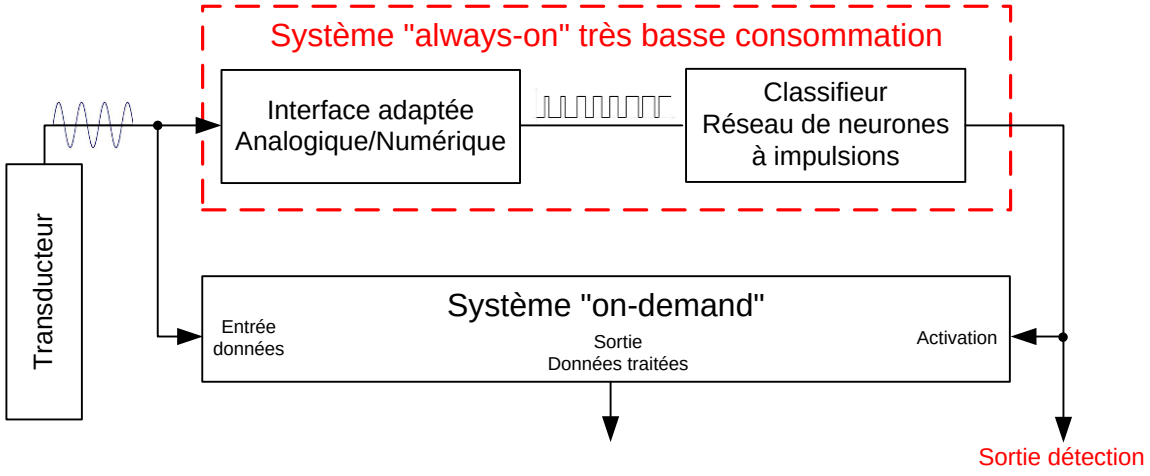

Figure 1 : Interface et traitement de signaux audios provenant d'un transducteur

IV) ORGANISATION

L'encadrement sera assuré par l'IMT Atlantique (site de Brest) et par le CEA LETI (Grenoble). La plus grande partie se déroulera sur le site de Grenoble.
Début de la thèse prévue en septembre/octobre 2018

Contact :
Fabrice SEGUIN,
IMT Atlantique, département électronique,
02 29 00 10 83
Fabrice.seguin@imt-atlantique.fr

Bibliographie

[1] http://www.cs.tut.fi/sgn/arg/dcase2017/dcase

[2] Bae, Soo Hyun and Choi, Inkyu and Kim, Nam Soo"Acoustic Scene Classification Using Parallel Combination of LSTM and CNN". Proceedings of the Detection and Classification of Acoustic Scenes and Events 2016 Workshop (DCASE2016) September 2016

[3] Prabha, P., Seong Joong Kim, K. Reddy, S. Rao, N. Griesert, A. Rao, G. Winter, and P.K. Hanumolu. "A Highly Digital VCO-Based ADC Architecture for Current Sensing Applications." IEEE Journal of Solid-State Circuits 50, no. 8 (August 2015); pp. 1785–95.

[4] B. Larras, C. Lahuec, F. Seguin and M. Arzel, "Design of analog subthreshold Encoded Neural Network circuit in sub-100nm CMOS," *2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, 2015, pp. 1-7.

[5] S. Patil, A. Ratiu, D. Morche and Y. Tsividis, "A 3–10 fJ/conv-step Error-Shaping Alias-Free Continuous-Time ADC," in IEEE Journal of Solid-State Circuits, vol. 51, no. 4, pp. 908-918, April 2016.

[6] Q. Yu, R. Yan, H. Tang, K. C. Tan and H. Li, "A Spiking Neural Network System for Robust Sequence Recognition," in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 3, pp. 621-635, March 2016.

[7] Larras Benoît, Lahuec Cyril, Seguin Fabrice, Arzel Matthieu "Ultra-Low-Energy Mixed-Signal IC Implementing Encoded Neural Networks". IEEE transactions on circuits and systems I - regular papers, 2016, vol. 63, n° 11, pp. 1974-1985

[8] V. Kuhn, C. Lahuec, F. Seguin and C. Person, "A Multi-Band Stacked RF Energy Harvester With RF-to-DC Efficiency Up to 84%," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 63, no. 5, pp. 1768-1778, May 2015.

[9] https://www.usinenouvelle.com/article/le-marche-des-capteurs-pour-automobile-devrait-atteindre-23milliards-de-dollars-en-2022.N580843