

Apport des réseaux de neurones dans le maillage spatial des paysages sonores urbains

Laurent Brocolini

Catherine Lavandier

LMRTE

Université de Cergy-Pontoise

5, mail Gay Lussac

Neuville sur Oise

95031 Cergy Pontoise CEDEX

E-mail : laurent.brocolini@u-cergy.fr

E-mail : catherine.lavandier@u-cergy.fr

Mathias Quoy

ETIS

Université de Cergy-Pontoise

6, avenue du Ponceau

95014 Cergy Pontoise CEDEX

E-mail : mathias.quoy@u-cergy.fr

Carlos Ribeiro

Bruitparif

9, impasse Milord

75018 Paris

E-mail : carlos.ribeiro@bruitparif.fr

Résumé

Établir un lien entre caractéristiques perceptives et paramètres acoustiques nécessite de réaliser des mesures sur le terrain dans différents environnements sonores.

Ces mesures peuvent tenir compte d'indicateurs globaux et d'indices spécifiques aux différentes sources sonores. Généralement, les lieux d'études comme les marchés, les parcs ou les rues sont choisis selon l'utilisation faite par des citoyens. Mais qu'en est-il des différences le long d'une rue ? L'entrée d'un parc est-elle similaire à la rue juxtaposée ? À quelle distance d'une rue l'environnement sonore d'un parc peut-il être considéré comme typique ? S'il est facile visuellement de faire la différence entre deux situations proches physiquement (un boulevard le long d'un parc par exemple), qu'en est-il d'un point de vue acoustique ? Les zones urbaines typiques (parc, rues, boulevards, ...) correspondent-elles à des environnements sonores homogènes ?

Afin de caractériser différents espaces urbains et de déterminer la distance optimale qui les sépare, une série d'enregistrements a été effectuée. Sur une distance globale d'environ 600 mètres, 70 enregistrements sonores de 15 minutes ont été réalisés tous les 7 à 8 mètres. Les sources sonores ont été identifiées par un expert. Un ensemble de 18 indicateurs a été extrait des enregistrements.

Afin de réaliser une classification des différentes situations sonores, et définir un maillage spatial pertinent, nous avons utilisé une méthode basée sur les réseaux de neurones, ou plus précisément des cartes de Kohonen. Quatre classes séparent les paysages sonores urbains : le parc, le boulevard, la rue piétonne et une zone de transition (entre le parc et le boulevard, ou aux alentours du carrefour).

Abstract

Linking perceptive features and acoustic measurements in natural environments requires measurements obtained during field studies. These measurements can take into account global indicators and indices devoted to different sources, and generally, locations such as markets, parks or streets are selected with regards to their use by city dwellers. But what about the measurement differences along a street? Is the entrance of a park similar to the street beside it? How far from the street can a typical park sound environment be heard? Therefore, in order to determine the optimized distance between urban locations, a series of recordings has been carried out. To establish a very fine space resolution along the streets, the parks and the transition areas between them, 15 minutes sequences were recorded every 7 meters. 70 locations in total were measured along a distance of 600 m. Sources on the recordings were identified by an expert listener. A set of 18 indicators was extracted from the recordings.

A classification of these locations, based on these indicators, was carried out to set a relevant space resolution. The clustering of the data was performed through a neural network using a Kohonen map. Four classes discriminate the urban soundscapes: park, thorough fare, pedestrian streets and "transition" (between park and street, or at crossroads).

Un paysage sonore peut être défini comme un environnement sonore qui est perçu dans un contexte particulier. Parmi les critères qui caractérisent le contexte, le lieu et sa morphologie sont particulièrement importants [1]. L'influence du contexte visuel sur la perception sonore a déjà été étudiée [2]. Par exemple, la présence d'oiseaux ou de véhicules n'est pas évaluée de la même façon dans un parc ou dans une rue [3]. Lorsque deux types de lieux différents sont à proximité, ils sont facilement discriminés d'un point de vue visuel, mais cela est-il aussi évident d'un point de vue acoustique ?

L'objectif de ce travail est de répondre à cette question par la détermination d'environnements sonores homogènes qui correspondent aux zones urbaines typiques (rues, parcs, rues piétonnes, etc.).

Ainsi, une campagne de mesure a été effectuée à Paris. Dix-huit variables acoustiques ont été extraites de soixante-dix enregistrements de quinze minutes chacun. Les enregistrements ont été pris en des positions distinctes espacées de sept à huit mètres. Les données ont été analysées au moyen d'une carte de Kohonen.

Protocole expérimental

Lieux

Les quartiers ont été choisis selon les contraintes suivantes : réunir sur une zone limitée différents types de lieux typiques (boulevard, rue piétonne, marché, parc, ...). Un premier site a été sélectionné dans le XII^e arrondissement de Paris allant du parc de Bercy, en passant par le boulevard Joseph Kessel (2 x 2 voies avec deux voies dédiées pour les bus) et se finissant dans la rue de Bercy (1 voie). Le deuxième site a été choisi dans le V^e arrondissement, commençant le long de la rue de l'épée du bois, rue à une voie et tournant dans la rue Mouffetard, rue piétonne avec une forte concentration de commerces. Le parcours se termine sur une place où un marché alimentaire a lieu tous les matins.

Campagne de mesures

Les enregistrements sonores de 15 minutes ont été effectués tous les 7 à 8 mètres, sur une distance totale d'envi-

ron 600 mètres (figures 1 et 2) pour un total de 70 points de mesure, répartis de la façon suivante : 46 mesures sur le premier site et 24 sur le second.

Comme nous souhaitons étudier l'effet du maillage spatial, nous voulions nous assurer que la différence entre deux mesures était due à la distance dans l'espace et non pas dans le temps (l'environnement sonore du matin est différent de celui de la soirée). Une précédente étude effectuée à Aix-en-Provence (13) a montré que cinq périodes peuvent caractériser l'évolution de l'environnement sonore au cours d'une journée dans un même lieu : l'éveil de la ville (6h-9h), une longue période dans la journée durant laquelle les mesures sont assez similaires (9h-17h), la fin de journée (17-20h), le soir (20h-23h), et enfin la nuit (23h-6h). Dans cette étude [5], il a été démontré qu'au cours de chaque période, des enregistrements de 15 minutes étaient suffisants pour être représentatifs de l'environnement sonore du lieu. Par conséquent, dans notre travail, tous les enregistrements ont été effectués de juillet à octobre (sauf août), les mardis, mercredis et jeudis entre 9h et 11h.

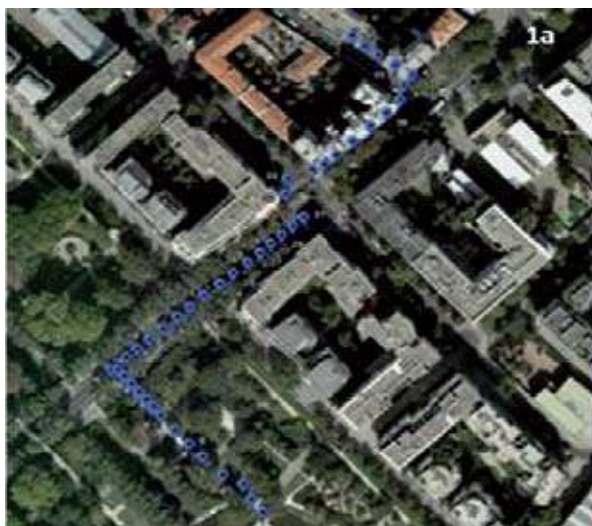


Fig. 1 : Description du site 1 : la figure 1a présente les 46 points de mesure ; la figure 1b présente le boulevard Joseph Kessel



Fig. 2: Description du site 2 : la figure 2a présente les 24 points de mesure; la figure 2b présente la rue piétonne Mouffetard

Indicateurs	Sources	Mesures	01	02	...	34	35	...
(1)	Indicateurs globaux	LAeq (dB(A))	54,0	56,8		69,5	68,9	
(2)		Ecart-type (dB (A))	2,1	3,3		4,6	5,2	
(3)		LA90 (dB (A))	51,2	51,2		60,8	59,0	
(4)		LA10 (dB(A))	56,3	58,0		72,4	72,4	
(5)	Véhicules légers (VL)	%T	0,0	0,0		40,5	47,7	
(6)		Émergence (dB(A))	-15,0	-15,0		1,4	3,7	
(7)	Poids lourds (PL)	%T	0,0	0,0		11,1	17,8	
(8)		Émergence (dB(A))	-15,0	-15,0		6,4	5,2	
(9)	Cyclomoteurs	%T	0,0	0,0		4,1	1,9	
(10)		Émergence (dB(A))	-15,0	-15,0		4,8	11,4	
(11)	Klaxons / Sirènes	%T	11,6	1,2		0,3	0,1	
(12)		Émergence (dB(A))	4,6	-2,5		5,2	12,5	
(13)	Voix	%T	1,6	11,7		22,6	6,4	
(14)		Émergence (dB(A))	3,4	-2,3		4,7	4,4	
(15)	Oiseaux	%T	100,0	100,0		50,0	10,0	
(16)		Em. (unités perceptives)	5	5		1	1	
(17)	Activités	%T	8,8	17,2		0,0	0,0	
(18)		Émergence (dB(A))	6,0	9,6		-15,0	-15,0	

Tabl. 1 : Données extraites de la base de données des 18 indicateurs

Indicateurs

À l'échelle de la ville et pour des heures de flux continu du trafic routier, le niveau équivalent L_{Aeq} peut suffire à discriminer différents lieux, mais pour des paysages sonores urbains changeants, il est nécessaire d'identifier les sources [3]. C'est pourquoi, outre les indicateurs énergétiques globaux comme le L_{Aeq} , l'écart-type de $L_{A,100ms}$ sur 15 minutes, le L_{A90} , ou le L_{A10} , nous avons caractérisé toutes les sources suivantes : les véhicules légers, les poids lourds, les cyclomoteurs, les klaxons et sirènes, les voix, les oiseaux et les activités diverses (livraisons de marchandises, activités commerciales, activités de travail, etc.). De plus, des indicateurs spécifiques à chaque type de source ont été calculés à partir de l'évolution temporelle $L_{A,100ms}$:

- le pourcentage de temps de présence (T %) : Pourcentage du temps durant lequel un seul type de source est présent dans les 15 minutes d'enregistrement ;

- et l'émergence : Différence entre le L_{A10} particulier pour un type de source et le niveau global L_{Aeq} .

Chaque source a été identifiée en réécoute par un expert. S'il est relativement aisé de coder le passage d'un véhicule, il n'en est pas de même pour les chants d'oiseaux. Les chants d'oiseaux ne sont pas assez forts pour masquer les autres sources, mais ils sont clairement remarqués par les auditeurs.

Aussi, nous avons décidé de noter ces variables de façon globale sur la durée de l'enregistrement. Le pourcentage de temps de présence a été globalement estimé par un auditeur «expert», variant de 0 à 100 % et l'émergence a été distribuée suivant 6 unités perceptives de 0 à 5.

En outre, lorsqu'une source n'est pas entendue, le pourcentage de temps de présence a été évidemment fixé à 0 et l'émergence a été déterminée en appliquant une valeur

de -15 dB (A), car nous avons supposé qu'une source qui est de 15 dB (A) inférieure au niveau équivalent n'est pas entendue.

Ainsi nous avons une base de 18 indicateurs pour chacun des 70 lieux d'enregistrements, comme présentés dans le tableau 1 montrant un échantillon des 18 indicateurs extraits des enregistrements.

Réseaux de neurones/Cartes de Kohonen

Afin de discrétiser différentes zones tout le long du parcours à partir de la base de données de 70 (lieux) x 18 (indicateurs), nous avons utilisé des cartes auto-organisatrices de Kohonen de 30 neurones. La carte est organisée avec un nombre réduit de neurones [6] par rapport au nombre d'objets présentés en entrée.

Il s'agit d'une classification non supervisée. Cela regroupe les lieux similaires dans un même neurone ou dans un groupe de neurones qui sont proches les uns des autres. Plus clairement, on considère 70 objets décrits par p variables (18 dans notre cas). Chaque neurone de la carte est caractérisé par un vecteur poids de dimension p qui relie ce neurone à tous les indicateurs. Par exemple, la première variable du poids de chaque vecteur correspond au niveau équivalent L_{Aeq} .

De même, la quatrième variable est le niveau L_{A10} . L'algorithme peut être décrit comme ceci : de façon aléatoire un objet est présenté au réseau. Le neurone de la carte qui est le plus proche de l'objet présenté, en termes de distance euclidienne, est désigné comme le neurone gagnant. Les poids de ce neurone sont alors modifiés de telle sorte que la distance entre lui et l'objet présenté diminue. Les poids des neurones voisins sont également modifiés dans une moindre mesure.

Puis les objets suivants sont présentés un à un et la même opération est répétée jusqu'à ce que les poids ne varient plus de façon significative.

Par opposition à une classification des «K-means» par exemple, cette classification conserve la même topologie que l'espace des données (d'où le nom de «carte»). Dans notre cas, cette méthode de classification est intéressante car le passage continu d'un point de mesure à un autre peut être représenté par l'activation de neurones voisins.

Résultats

Poids des neurones

Le poids final des neurones peut expliquer pourquoi ils ont répondu à un lieu ou à un autre. Sur la figure 3, chaque carré de couleur représente le poids entre un indicateur et chaque neurone de la carte. Une couleur jaune caractérise les valeurs fortes de l'indicateur alors qu'une couleur noire caractérise des valeurs faibles. Les neurones sont numérotés de 1 à partir du coin gauche en bas à 30 dans le coin droit en haut (voir figure 6).

Sur la figure 3c, nous pouvons voir que le pourcentage de temps de présence des véhicules légers caractérise les neurones 1, 2, 3, 6 et 7 avec une valeur forte et les neurones 13, 19, 25, ..., 30 avec une valeur faible. Cependant cette seconde série de neurones est divisée en deux groupes si on regarde le pourcentage de temps de présence des oiseaux (figure 3h). Les neurones 13, 19, 25 et 26, à gauche de la carte, ont un temps de présence des oiseaux faible alors que les neurones 23, 24, 29 et 30 ont un temps de présence des oiseaux élevé. Enfin, les trois neurones situés en bas à droite de

la carte (neurones 5, 6 et 12) sont caractérisés par un temps de présence ou une émergence de nombreuses sources différentes (véhicules légers/figure 3c, poids lourds/figure 3d, cyclomoteurs/figure 3e, klaxons et sirènes/figure 3f, voix/figure 3g, activités/figure 3i, ...)

Carte de Kohonen et classification de Ward

Les résultats de la classification de Kohonen sont reportés sur la figure 6. Sur cette figure, nous indiquons pour chaque lieu de mesure le neurone qui a codé ce lieu. Par exemple, les lieux 1 et 2 sont codés par le neurone 24. En outre, pour regrouper les 30 neurones en un nombre de classe plus restreint, nous avons eu recours à une classification hiérarchique ascendante de Ward (figure 4). L'avantage d'opérer cette classification sur les neurones plutôt que sur les objets initiaux (dans notre cas les lieux de mesures) est que, comme montré précédemment, nous pouvons comprendre avec les poids des neurones quelles variables sont responsables du regroupement en différentes classes [7].

La classification de Ward propose 3 ou 4 classes (figure 5), mais nous avons décidé d'en garder 4 (après le nœud 56) afin de minimiser la perte d'information due à l'agrégation des groupes. La représentation simplifiée du réseau de Kohonen (figure 6) basée sur les 18 indicateurs présente donc les 4 zones typiques établies.

La première zone (vert) correspond au parc. Les enregistrements dans le parc (lieux 1 à 12) sont regroupés dans cette zone. Le point de mesure 47 (neurone 18) est également regroupé dans cette zone bien qu'il se situe physiquement dans la rue de l'épée de bois. Cela peut être expliqué par une forte présence d'oiseaux au moment de l'enregistrement (figure 3h).

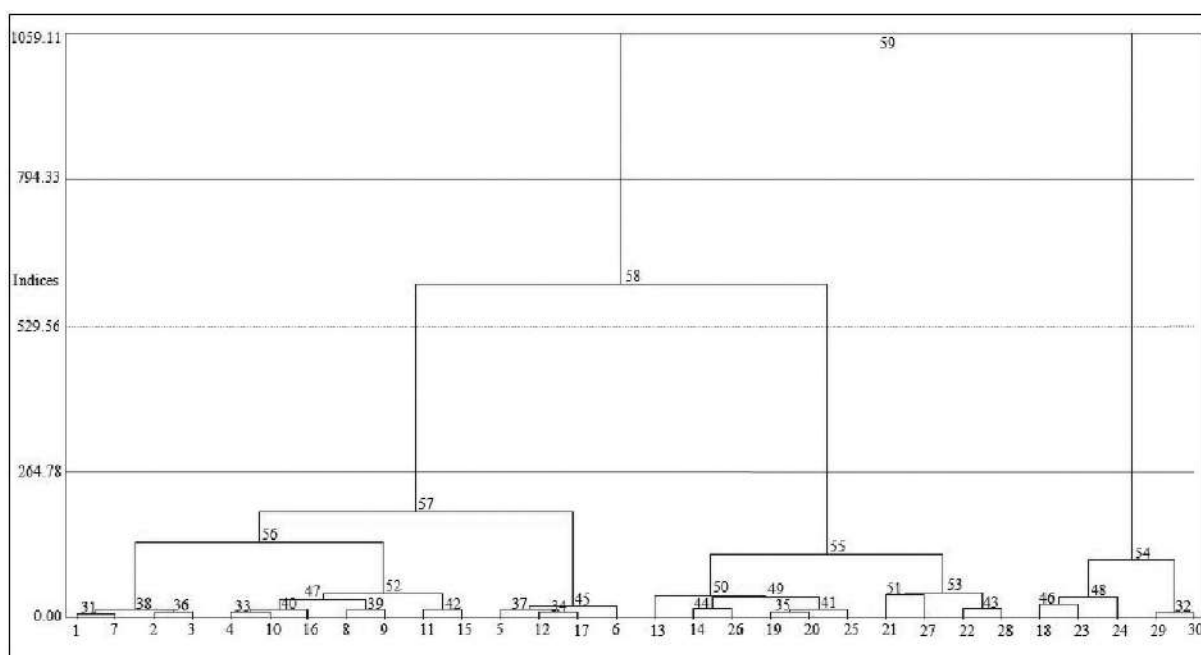


Fig. 4 : Dendrogramme de la classification de Ward sur les neurones

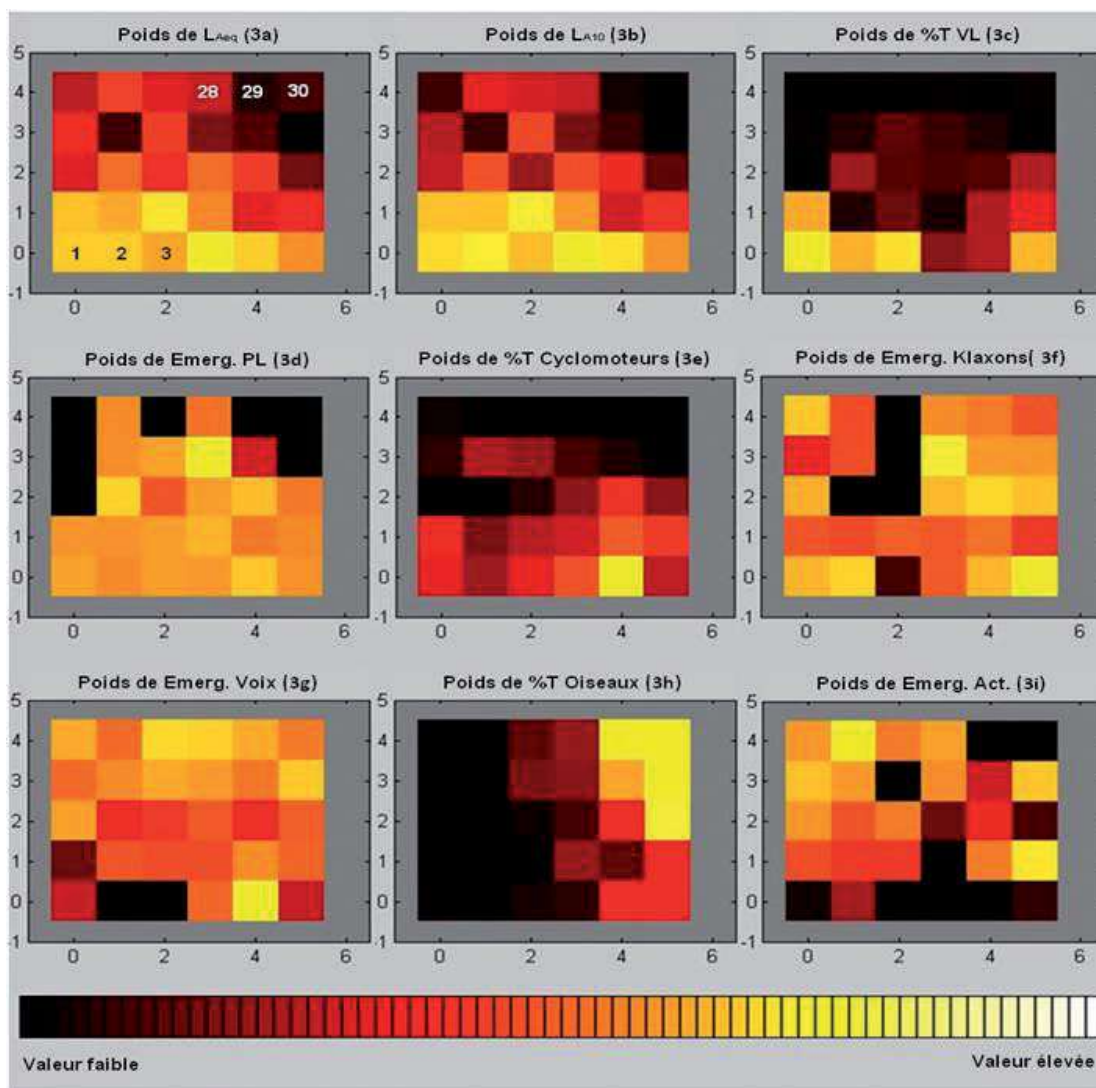


Fig. 3 : Poids des neurones pour différentes variables, avec l'échelle de couleur

La seconde zone (rouge) concerne le boulevard Joseph Kessel, depuis le lieu 19 et la rue de Bercy. Elle regroupe les enregistrements dans lesquels se trouvent des niveaux sonores élevés (figures 3a et 3b) avec un temps de présence des véhicules légers important (figure 3c). Il est intéressant à noter qu'une petite rue comme la rue de l'épée de bois (lieux 47 à 50) a été répartie dans 3 classes différentes. Il semble donc que la caractérisation sonore d'une petite rue n'est pas simple. Le nombre de mesures dans cette rue n'est peut-être pas suffisant pour que celle-ci soit codée par un neurone de la carte de Kohonen.

La troisième zone (bleu) réunit les lieux de la rue piétonne et commerçante. Elle est caractérisée par forte présence des voix et des activités diverses (figure 3g et 3i).

Enfin, la quatrième zone (orange) demande une attention particulière. Les lieux 13 à 18 (neurones 6 et 12) sont regroupés dans cette zone et caractérisent la transition entre le parc et le boulevard. Les lieux 30, 33, 34 et 36 (neurones 5 et 6) sont également situés dans cette zone. Ils correspondent à un carrefour avec des feux de signalisation où le flux de véhicules n'est pas

continu. Le point de mesure 50 appartient aussi à cette quatrième zone. Il correspond à l'angle entre la rue circulaire de l'épée de bois et la rue piétonne Mouffetard. Nous avons déjà mentionné que ces neurones sont caractérisés par un grand nombre de sources différentes. Il peut être noté que les lieux 23 et 24 qui correspondent à un feu de signalisation sur le boulevard sont associés à la troisième zone (rouge) et non pas avec cette dernière zone, dite de «transition», bien que le flux de véhicules n'y soit pas continu. Il semble donc que même si le flux de véhicule s'arrête au feu tricolore le long de la route, cela n'est pas suffisant d'un point de vue acoustique pour appartenir à une zone de transition.

Une analyse spatiale montre que la transition acoustique entre le parc et le boulevard s'étend sur une cinquantaine de mètres. En ce qui concerne le carrefour, cette transition semble également s'étaler sur une cinquantaine de mètres. Par conséquent, ce n'est qu'à 25 mètres de part et d'autre du carrefour ou des limites du parc que le paysage sonore semble se stabiliser dans une catégorie plus typée. Ce résultat est important car il ne correspond pas avec la distance de transition visuelle, qui elle est bien plus courte.

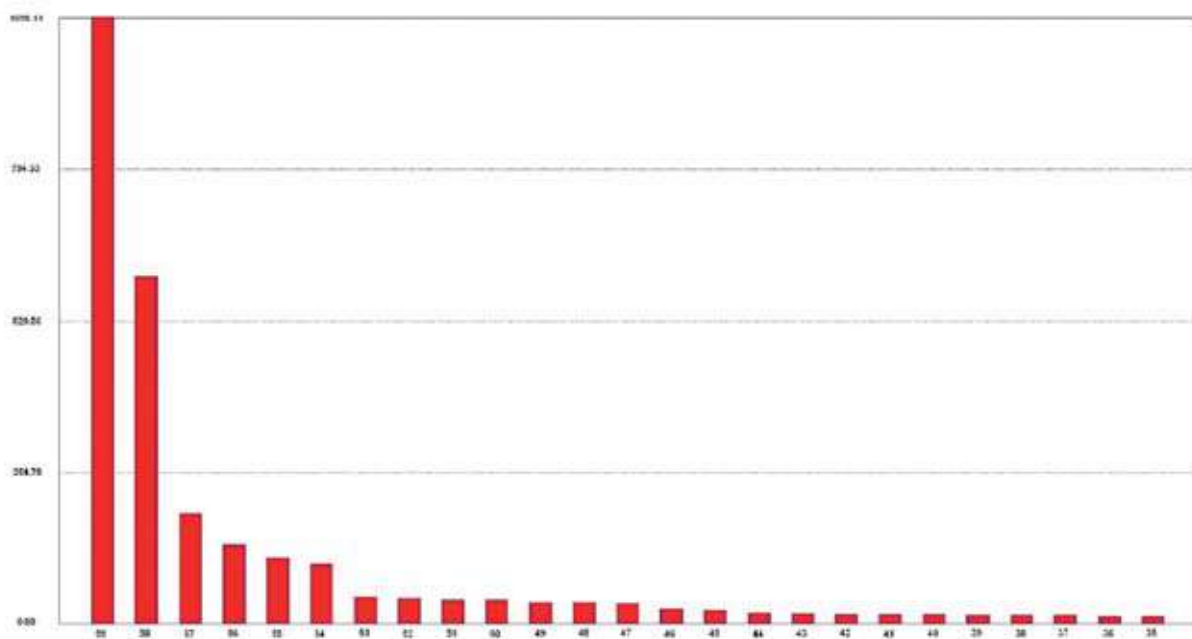


Fig. 5 : Diagramme des indices des noeuds



Fig. 6 : Neurones de la carte de Kohonen avec les lieux correspondants de Ward sur les neurones

zones caractérisées par la présence de nombreuses sources, comment s'orientent-ils ? Écoutent-ils certaines sources ? Toutes les sources ? Cette zone est-elle identifiée en tant que zone de transition, ou bien est-elle une « non-zone » qui sert de passage d'une zone typée à une autre ? Cette étude perceptive est actuellement en cours de réalisation afin de proposer des indicateurs de qualité sonore en milieu urbain, adaptés aux différents contextes temporels et spatiaux de la ville.

Remerciements

Cette recherche fait partie d'un programme financé par l'ADEME (Agence de l'environnement et de la maîtrise de l'énergie) dans le cadre du PREDIT. Nous remercions ici les différentes autorités de Paris qui nous ont autorisé à réaliser les différentes mesures dans les lieux publics.

Références bibliographiques

- [1] Brown L., Kang J., Gjestland T., Towards some standardization in assessing soundscape preference, submitted to Euronoise 2009, Ottawa, 2009.
- [2] Viollon S., Lavandier C., Drake C., « Influence of visual setting on sound ratings in an urban sound environment ». *Applied Acoustics*, Vol. 63 (5), pp. 493-511, mars 2002.
- [3] Lavandier C., Defreville B., « The contribution of sound source characteristics in the assessment of urban soundscapes », *Acta Acustica united with Acustica*, Vol. 92 (6), pp. 912-921, novembre-décembre 2006.
- [4] Kohonen T., *Self-organization and associative memory*, Springer.
- [5] Lavandier C., Barbot B., Influence of the temporal scale on the relevance of acoustic parameters selected to characterize urban sound environments, *Proceedings of Euronoise 2003*, Naples, 2003.
- [6] Nakache J-P, Confais J., *Approche pragmatique de la classification*, Technip, pp. 165-170, 2005.
- [7] Cottrell M., Ibbou S., Letremy P., SOM-based algorithms for qualitative variables, *Neural Networks, New Developments in Self-Organizing Systems*, Volume 17(8-9), pp. 1149-1167, October-November 2004.

Conclusion

Comme attendu, trois zones typiques ont été extraites des enregistrements sonores. Les zones sont caractérisées par des environnements sonores homogènes ou des sources typiques sont présentes : des oiseaux dans le parc, des véhicules sur le boulevard et dans la rue circulée, des voix et des activités dans la rue piétonne. La quatrième zone regroupe des environnements sonores de transition où différents types de sources peuvent être entendus. Compte tenu de ces résultats, il serait maintenant intéressant d'étudier ces zones de transition d'un point de vue perceptif. Que se passe-t-il lors du passage d'une zone à une autre ? Les gens portent-ils leur attention sur la morphologie visuelle et/ou sonore du lieu. Dans ces