

Etat de l'art de la séparation aveugle de sources audio

Introduction

Dans ce chapitre, nous présentons un état de l'art des principales méthodes de séparation aveugle de sources. Nous nous concentrerons ensuite sur les méthodes qui ont été appliquées dans le cadre de l'audition des robots. Mais tout d'abord, nous commençons par le modèle de mélange et les principales notations que nous adopterons le long de ce rapport.

2.1 Formulation du problème

2.1.1 Modèle des signaux

Supposons que l'on dispose de N sources audio $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$ et d'un réseau de M microphones. Les sorties du réseau de capteurs sont notés $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$, t étant l'indice de temps. Nous supposons que nous sommes dans un cas de séparation sur-déterminé $M \geq N$. Comme nous travaillons dans un environnement réel, les mélanges à la sortie des microphones dans le domaine temporel sont modélisés comme la somme des convolutions entre les signaux sources et les réponses impulsionnelles des différents chemins de propagation entre les sources et les capteurs (*cf.* figure 2.1).

On note $\mathbf{h} = [\mathbf{h}(0), \dots, \mathbf{h}(L-1)]$ les réponses impulsionnelles tronquées à la longueur L , où $\mathbf{h}(l) = [h_{ij}]_{1 \leq i \leq M, 1 \leq j \leq N}$, est une matrice de dimension $M \times N$ contenant les $l^{\text{ième}}$ coefficients des réponses impulsionnelles des différents

chemins acoustiques entre les N sources et M capteurs. Les mélanges à la sortie des capteurs s'écrivent dans le domaine temporel comme suit :

$$\mathbf{x}(t) = \sum_{l=0}^{L-1} \mathbf{h}(l) \mathbf{s}(t-l) + \mathbf{n}(t) \quad (2.1)$$

où $\mathbf{n}(t)$ est un vecteur de bruit stationnaire. Nous considérons un bruit diffus, spatialement décorréolé, dont l'énergie est supposée être négligeable par rapport à celle des sources. Si le bruit est ponctuel, il sera considéré comme une source sonore. Ce scénario correspond à notre processus expérimental et à l'application finale.

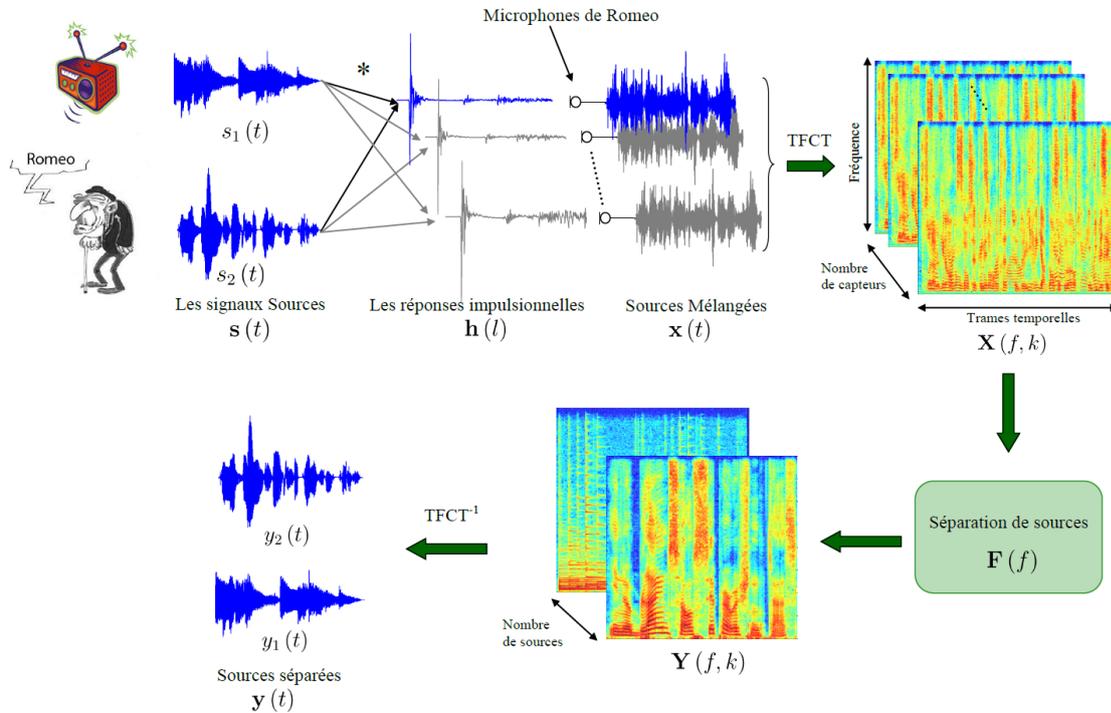


FIGURE 2.1 – Principe de la séparation de sources

Dans le domaine temporel, le problème de séparation de sources se résume à estimer des filtres de séparation $\mathbf{f} = [\mathbf{f}(0), \dots, \mathbf{f}(L' - 1)]$ de longueur L' , avec $\mathbf{f}(l) = [f_{ji}]_{1 \leq i \leq N, 1 \leq j \leq M}$ une matrice de dimension $N \times M$ contenant les $l^{\text{ième}}$ coefficients des filtres de séparation des N sources à partir des M capteurs.

Dans le domaine fréquentiel, quand la longueur de la fenêtre d'analyse N_f de la Transformée de Fourier à Court Terme (TFCT) est au moins deux fois plus grande que la longueur L des filtres de mélanges (les réponses impulsionnelles), le mélange convolutif est approché à un mélange instantané à chaque fréquence f . Le problème de séparation devient donc plus simple que celui posé dans le domaine temporel et

nous estimons des matrices de séparation au lieu de filtres de séparation dont la longueur devient importante si nous travaillons avec des signaux réels.

Dans le domaine temps-fréquence, les signaux du mélange à l'indice fréquentiel f et à la trame temporelle k peuvent être approchés par :

$$\mathbf{X}(f, k) \simeq \mathbf{H}(f) \mathbf{S}(f, k) \quad (2.2)$$

$\mathbf{X}(f, k) = [X_1(f, k), \dots, X_M(f, k)]^H$ est la TFCT de $\{\mathbf{x}(t)\}_{1 \leq t \leq T}$ à la fréquence $f \in \left[1, \frac{N_f}{2} + 1\right]$ et la fenêtre d'analyse $k \in [1, N_T]$, $\mathbf{S}(f, k) = [S_1(f, k), \dots, S_N(f, k)]^H$ est la TFCT de $\{\mathbf{s}(t)\}_{1 \leq t \leq T}$. \mathbf{H} est la transformée de Fourier des filtres de mélanges $\{\mathbf{h}(l)\}_{0 \leq l \leq L-1}$.

Notre objectif est d'utiliser un critère de séparation approprié et de trouver, à chaque fréquence, une matrice de séparation $\mathbf{F}(f)$ qui conduira à l'estimation des sources originales dans le domaine temps-fréquence :

$$\mathbf{Y}(f, k) = \mathbf{F}(f) \mathbf{X}(f, k) \quad (2.3)$$

où $\mathbf{Y}(f, k) = [Y_1(f, k), \dots, Y_N(f, k)]^H$. La Transformée de Fourier à Court Terme Inverse (TFCTI) des signaux sources estimés $\mathbf{Y} = \{\mathbf{Y}(f, k)\}_{1 \leq f \leq N_f, 1 \leq k \leq N_k}$ permet de retrouver les sources $\mathbf{y}(t) = [y_1(t), \dots, y_N(t)]^T$ dans le domaine temporel :

$$\mathbf{y}(t) = \sum_{l=0}^{L'-1} \mathbf{f}(l) \mathbf{x}(t-l) \quad (2.4)$$

2.1.2 Les problèmes relatifs à la séparation de sources dans le domaine fréquentiel

La séparation de sources dans le domaine fréquentiel présente un avantage par rapport à la séparation de sources dans le domaine temporel qui est l'estimation de matrices de séparation au lieu de filtres, ce qui engendre des algorithmes de séparation plus simples et donc plus rapides. Cependant, travailler dans le domaine fréquentiel présente deux problèmes bien connus : le problème d'échelle et le problème de permutation. Ces problèmes se traduisent en écrivant l'équation (2.3) des sources estimées dans le domaine temps-fréquence sous la forme :

$$\mathbf{Y}(f, k) = \mathbf{F}(f) \mathbf{X}(f, k) = \mathbf{P}(f) \mathbf{D}(f) \mathbf{S}(f, k) \quad (2.5)$$

où $\mathbf{P}(f)$ une matrice de permutation et $\mathbf{D}(f)$ est une matrice diagonale d'échelle.

Problème d'échelle A chaque fréquence f , la multiplication des sources estimées par des constantes $\mathbf{D}(f)$ n'affecte le résultat de l'algorithme de séparation que par un filtrage des signaux estimés une fois convertis dans le domaine temporel.

Problème de permutation L'ordre des sources estimées peut être différent à des fréquences adjacentes. Ceci est modélisé par la multiplication des sources originales dans le domaine fréquentiel $\mathbf{S}(f, k)$ par une matrice de permutation $\mathbf{P}(f)$. Si le problème de permutation n'est pas résolu, la reconstruction dans le domaine temporel d'une source donnée contiendra des contributions des autres sources.

2.2 Séparation aveugle de sources audio

Le livre de Pierre Comon and Christian Jutten [24] présente un bon résumé de l'histoire de la séparation de source dont voici les principaux points. Le problème de la séparation de sources a été formulé pour la première fois en 1982 par Bernard Ans, Jeanny Héroult et Christian Jutten dans le cadre de la modélisation neuronale pour le décodage de mouvements dans les vertèbres. A Grenoble en septembre 1987, les acteurs sont J.F. Cardoso, J. Héroult, P. Comon et C. Jutten. Lors d'un workshop, J.F. Cardoso visite J. Héroult et C. Jutten qui lui expliquent le principe de la séparation de sources et lui montrent une démonstration en temps-réel accompli par un appareil purement analogique que C. Jutten a construit en 1985. L'appareil était capable de séparer en temps-réel deux sources audio dans un mélange contrôlé par un potentiomètre. Immédiatement mais indépendamment, J.-F. Cardoso et P. Comon deviennent très enthousiastes à propos de la séparation aveugle de sources audio. Ensuite en 1990, le problème de séparation de sources a été traité par plusieurs chercheurs de plusieurs domaines : traitement de signal, statistique, réseau de neurones, etc, ... En 1991, le premier papier international consacré à ce problème a été publié dans le journal *Signal Processing* et en 1999, la première conférence internationale complètement dédiée à la séparation aveugle de sources a été organisée à Aussois dans les Alpes Français et a attiré 130 chercheurs du monde entier.

Initialement, le problème de séparation de sources a été posé pour les mélanges linéaires instantanés, puis il a été généralisé aux mélanges convolutifs au début des années 90. Puisque nous travaillons avec des signaux enregistrés dans des milieux réels réverbérants, nous nous intéressons aux méthodes de séparation de mélanges convolutifs. Les algorithmes de séparation de mélanges convolutifs peuvent être classés [66] selon :

- le domaine de séparation : domaine temporel, domaine fréquentiel ou domaine temps-fréquence,
- le critère sur lequel se base la séparation : l'indépendance des sources, la non-corrélation des sources, la parcimonie dans le domaine temps-fréquence ou les principes d'analyse de scènes sonores et la psychoacoustique.

Dans cette section, nous allons présenter un aperçu de l'état de l'art de la séparation de sources dans les mélanges convolutifs en les classant selon leurs critères de séparation. Si le lecteur veut plus de détails sur ces méthodes et est curieux d'explorer d'autres méthodes pas nécessairement dans notre sujet, il peut consulter les livres [24, 50].

2.2.1 Algorithmes basés sur l'indépendance des sources

Dans la première classe d'algorithmes, la séparation de sources est basée sur l'hypothèse d'indépendance statistique des sources. Nous nous intéresserons dans cette section aux statistiques de 4^{ième} ordre et aux méthodes basées sur la théorie de l'information. Nous parlerons aussi de l'Analyse en Composantes Indépendantes (ACI) qui correspond à un cadre général pour résoudre les problèmes de séparation aveugle de sources basés sur l'indépendance statistique des sources à estimer. Elle a été introduite en 1987 par Jutten et formalisée pour les mélanges linéaires par Comon en 1991 [22, 23].

Statistique du 4ème ordre L'indépendance statistique peut se traduire en annulant tous les moments croisés entre les modèles des sources :

$$E \left[y_n(t)^\alpha y_m(t-\tau)^\beta \right] = 0, \quad n \neq m, \quad \{\alpha, \beta\} \subset \mathbb{N}, \quad \forall \tau \quad (2.6)$$

où $E(\cdot)$ est l'espérance mathématique.

Cependant, pour atteindre la séparation, il n'est pas nécessaire de minimiser tous les moments croisés, plusieurs algorithmes se basent sur la minimisation des statistiques de quatrième ordre. La minimisation des cumulants de quatrième ordre a été étudiée par plusieurs chercheurs [19, 25, 41, 75]. Ye *et al.* ont utilisé les cumulants du second et troisième ordre pour la séparation de sources [96]. L'algorithme le plus connu basé sur cette méthode est JADE de Cardoso et Souloumiac [21] pour les signaux complexes appliqué dans le domaine temps-fréquence pour séparer les mélanges convolutifs. Une mesure des statistiques de quatrième ordre est le kurtosis. Cette mesure de la non-Gaussianité des signaux a été utilisée dans plusieurs algo-

rithmes [77, 82]. D'autres algorithmes utilisent les statistiques d'ordre supérieur avec des fonctions non linéaires, le développement de Taylor de ces fonctions capture les moments d'ordre supérieur et c'est suffisant pour séparer les mélanges convolutifs. [43].

Du côté de la théorie de l'information Dans la théorie de l'information, l'indépendance entre les sources est exprimée en fonction de leurs densités de probabilité. Augmenter l'indépendance entre les sources \mathbf{y} équivaut à minimiser leur information mutuelle ce qui revient à maximiser l'entropie. Le maximum d'entropie est obtenu lorsque la somme des entropies des variables y_n est égale à l'entropie jointe de \mathbf{y} . Un algorithme bien connu qui se base sur cette idée est Infomax de Bell et Sejnowski [42] qui a été amélioré en vitesse de convergence par la méthode du gradient naturel d'Amari [13]. L'algorithme Infomax peut aussi être dérivé en utilisant le maximum de vraisemblance [9] ou la divergence de Kullback-Leibler entre les distributions empiriques et le modèle d'indépendance [20]. L'algorithme de Bell et Sejnowski a été étendu au cas convolutif par Torkkola [84].

2.2.2 Algorithmes basés sur la non-corrélation des sources

Dans cette deuxième classe d'algorithmes, la séparation de sources est basée sur l'hypothèse de la non corrélation des sources plutôt que la condition plus forte de leur indépendance. La décorrélation des sources s'obtient en utilisant les statistiques de second ordre, elle n'implique pas l'indépendance et est insuffisante pour la séparation (le système d'équation est dans ce cas sous-déterminé [49]). Cependant, quand les sources sont non-stationnaires, les statistiques de second ordre sont différentes à chaque trame temporelle. Ceci implique que plus d'équations sont disponibles pour résoudre le problème et que le système d'équations pour estimer la matrice de séparation peut être résolu. Les statistiques variables au cours du temps donnent des informations additionnelles pour la séparation : c'est la décorrélation basée sur la non-stationnarité des sources.

Plusieurs algorithmes se basent sur les statistiques de second ordre avec une condition de non-stationnarité des sources. L'idée de la séparation basée sur la décorrélation des signaux non-stationnaires a été proposée par Weinstein *et al.* en 1993 [92]. Ils ont proposé de minimiser les puissances des sources estimées pendant des périodes de stationnarité du signal différentes. Wu et Principe [93] étendent le principe de la diagonalisation multiple à la séparation des mélanges convolutifs dans le domaine temps-fréquence en se basant sur la méthode proposée par Souloumiac

en 1995 pour les mélanges à bande étroite et qui exploite la structure des valeurs propres de deux matrices de covariance [76]. L'utilisation des statistiques de second ordre sur les mélanges non-stationnaires convolutifs s'est poursuivie avec Pham et Cardoso et Matsuka *et al.* qui se basent sur la non stationnarité des sources appliquée aux mélanges instantanés [51, 68]. Cette méthode a été appliquée dans le domaine temporel [45] et fréquentiel [44]. Une version adaptative a été proposée par Aichner *et al.* en 2003 [10]. Un algorithme de séparation bien connu qui se base sur la minimisation de la corrélation croisée des sources à des instants multiples dans le domaine fréquentiel est celui de Parra et Spence [64]. Les statistiques de second ordre sont capturées par le spectre de puissance à des instants multiples et les matrices de mélange et de séparation sont estimées avec une méthode des moindres carrés. Une approche adaptative a été proposée la même année [65]. En 2004, Wang *et al.* proposent une version de ce même algorithme en incorporant une fonction de pénalité dans la fonction de coût [91]. Ceci élimine les contraintes dans une optique de programmation non-linéaire et améliore les résultats.

Si les sources sont non-blanches, nous avons une corrélation décalée pour de multiples délais temporels. Les statistiques de second ordre sont différentes dans chaque décalage temporel et donc plus d'équations sont disponibles et le système d'équation peut être résolu. C'est la décorrélation en se basant sur la non-blanchité des sources ou sur le décalage-temporel [53]. Aichner *et al.* ont combiné l'approche basée sur la non-stationnarité des sources et l'approche basée sur leur blanchiment pour améliorer les résultats de la décorrélation [11, 18].

D'autres méthodes utilisant la non corrélation des signaux sources à estimer se basent sur d'autres principes d'estimation que les statistiques de second ordre tels que la phase minimale du système de mélange et la stationnarité cyclique des signaux sources [66].

2.2.3 Algorithmes basés sur la parcimonie dans le domaine temps-fréquence

Ce type d'algorithme se base souvent sur le non recouvrement des sources dans le domaine temps-fréquence. Les méthodes les plus utilisées sont le masquage temps-fréquence et le clustering.

Nous pouvons voir ce type d'algorithme du point de vue de leur critère : une fonction de coût qui agit sur la parcimonie des mélanges pour les séparer (norme l_p , l_1 , etc). [26] présente un aperçu des méthodes de séparation de sources en se basant

sur la parcimonie.

2.2.4 Algorithmes basés sur l'analyse de scènes sonores et la psychoacoustique

Certains algorithmes de séparation de sources se basent sur les idées issues de l'étude du système auditif et les principes de l'audition humaine : c'est l'analyse de scènes auditives computationnelle que nous avons introduite dans le chapitre précédent. Le terme analyse de scènes auditives introduit par Bregman en 1990 [17] fait référence à la capacité d'un humain à former des représentations perceptuelles des sources présentes dans un mélange acoustique comme l'effet "cocktail party". Selon Bregman, l'analyse de scènes auditives est un processus à deux étapes. Dans la première étape, le mélange acoustique est décomposé en éléments qui décrivent un événement acoustique signifiant. Dans la deuxième étape, les éléments qui peuvent provenir de la même source sont regroupés pour former une structure perceptuelle.

Parmi les outils perceptuels qui ont été utilisés dans la séparation de sources, nous citons [37, 90] qui ont utilisé le masquage perceptuel pour résoudre le problème de permutation, [83] qui a utilisé le pitch dans un algorithme de séparation semi-aveugle pour séparer les sources pendant les parties voisées, [67] qui a utilisé les fonctions de transfert de tête pour poser des contraintes géométriques à la séparation et [71] qui a utilisé la différence interaurale d'intensité (IID : Interaural Intensity Difference) et la différence interaurale de phase (IPD : Interaural Phase Difference) pour différencier les sources.

2.3 Séparation de sources pour l'audition des robots

Dans cette section, nous présentons l'état de l'art des algorithmes de séparation de sources qui ont été publiés spécifiquement dans le cadre de l'audition des robots.

2.3.1 Les premiers essais

Une première discussion sur l'importance d'un module de traitement de signal audio dans un robot autonome a été lancée en 1994 par Brooks et Stein [8]. Les auteurs proposent une réflexion sur un système physique intégrant la vision, une entrée/sortie audio et des manipulations adroites, non seulement dans un but d'ingénierie pour construire un prototype général d'un robot autonome mais aussi dans un but scientifique pour comprendre la cognition humaine.

En 1997, Huang *et al.* proposent un système de localisation et de séparation de sources en utilisant 3 capteurs installés verticalement en haut d'un robot [39]. La localisation est effectuée grâce à la différence de temps d'arrivée des sources par rapport aux deux capteurs qui reçoivent la plus grande énergie des signaux. Ce robot fait de l'*audition active* : il modifie son comportement par rapport aux conditions d'écoute, par exemple il se dirige vers la source si celle-ci est loin pour mieux la localiser. Quand à la séparation de sources, elle se fait grâce à la disparité de phase. Les auteurs ne sont pas rentrés dans les détails concernant ces techniques. La séparation a été faite sur 2 signaux de parole séparés en azimuth de 38° dans un milieu anéchoïque et une salle réverbérante. Les auteurs affirment avoir obtenu un bon résultat de séparation dans les deux milieux en se basant sur des mesures subjectives (l'allure des signaux séparés et l'écoute).

2.3.2 Utilisation des différences interaurales d'intensité et de phase

Si le robot dispose de deux microphones placés au niveau des oreilles humaines, on peut penser à utiliser la différence interaurale d'intensité (IID : Interaural Intensity Difference) et la différence interaurale de phase (IPD : Interaural Phase Difference) pour résoudre le problème de localisation et de séparation des sources. C'est ce qu'ont fait en 2003, Nakadai *et al.* [57]. Ces auteurs utilisent la localisation par IID et IPD afin d'extraire les sources par un filtre passe direction actif qui sépare le signal émis d'une direction bien déterminée. Pour estimer les IID et IPD tout en n'ayant pas recours au calcul des fonctions de transferts de tête comme l'ont fait Matsusaka *et al.* [52], les auteurs proposent deux méthodes. Tout d'abord, ils ont appliqué à l'audio la géométrie épipolaire connue dans le traitement des signaux vidéos [56]. Cette méthode leur a permis d'estimer l'IPD en continue pour toutes les directions d'arrivées des sources et l'IID pour seulement le centre, la droite et la gauche du robot. Ensuite, les auteurs ont utilisé la théorie de la diffraction : le champ à proximité des capteurs de droite et de gauche est calculé en tenant compte du champ incident et du champ diffracté. Les IID et IPD sont ainsi calculés en tenant compte de l'effet qu'a la tête sur ce champ sonore, c'est à dire en tenant compte de la diffusion. L'évaluation de cette technique a été effectuée sur le robot SIG, avec un son harmonique composé de 30 fréquences de 100 Hz à 3000 Hz dans une chambre dont le temps de réverbération est de $300ms$. Les auteurs affirment que les résultats de localisation et de séparation obtenus en utilisant la théorie de la diffraction sont

comparables à ceux obtenus en utilisant les fonctions de transfert de tête.

2.3.3 Séparation de sources à deux étapes

Valin *et al.* [85] proposent en 2004 un système de séparation de sources à deux étapes pour l'audition des robots. La première étape consiste en une séparation de sources géométrique (GSS : Geometric Source Separation) basée sur l'algorithme de Parra et Alvino [63]. Cet algorithme consiste en la décorrélation des signaux à la sortie des capteurs en imposant une contrainte de gain unitaire dans la direction des sources visées et des zéros dans la direction des sources interférentes. La deuxième étape consiste en un filtrage multi-canal utilisant l'estimation du bruit ambiant et les sources interférentes pour améliorer le signal produit lors de la séparation géométrique [54]. Cet algorithme a été appliqué à Pioneer 2, un robot avec 8 capteurs pour séparer 3 sources. Cette méthode a été utilisée en 2007 comme un prétraitement de la reconnaissance de la parole dans le robot SIG 2 [86].

En 2005, Saruwtari *et al.* proposent aussi un algorithme de séparation à deux étapes [73]. Les signaux binauraux observés aux "oreilles" d'un humanoïde sont traités d'abord avec un modèle à une seule entrée et multiple sorties SIMO (Single-Input and Multiple-Output) basé sur ICA [79] et ensuite avec un masquage binaire [31]. Les expériences ont été faites avec la tête et le torse d'un mannequin dans une salle dont le temps de réverbération est de 200 ms et avec deux sources. Les auteurs montrent que l'utilisation du modèle SIMO basé sur ICA avec masquage binaire donne de meilleurs résultats de séparation que l'utilisation d'une ICA conventionnelle suivie du même masquage.

2.3.4 Localisation et séparation

Tamai *et al.* se sont intéressés à la séparation de sources pour l'audition des robots en utilisant deux configurations de réseau de capteurs différentes. En 2004, ils proposent un réseau de 32 microphones fixés sur un cercle de 50 cm de diamètre et qui peut être monté sur un robot mobile. Dans un premier temps, ils proposent de localiser les sources par une formation de voies type *delay-and-sum* [80]. En 2005, une partie de l'équipe précédente change la configuration du réseau de capteurs qui devient un réseau de 3 anneaux avec le même nombre de microphones [81]. La localisation est effectuée par la formation de voies *delay-and-sum* et la séparation est faite à la suite par un algorithme de sélection de bande de fréquence [74].

En 2007, Rudzyn *et al.* décrivent un système d'audition des robots nommé RRAS

[70]. Avec un réseau de 4 capteurs, ils procèdent à une localisation 3D et se focalisent sur la source d'intérêt. Cette focalisation se fait suite à une caractérisation de la voix leur permettant de distinguer si la source sonore détectée est de type parole ou pas. La même année, [38] utilise ICA pour la séparation de source pour l'audition des robots avec un module de reconnaissance du locuteur.

2.3.5 Le système d'audition complet HARK

Yamamoto *et al.* proposent une technique de séparation de sources basée sur des contraintes géométriques comme un prétraitement au module de reconnaissance de la parole de leur système d'audition des robots [86, 94, 95]. Ce système a été implémenté dans les humanoïdes SIG2 et ASIMO avec un réseau de 8 microphones, comme un module d'un système d'audition des robots plus complet nommé HARK [59]. La référence [62] propose une analyse de ce problème d'un point de vue analyse computationnelle de scènes auditives en utilisant l'algorithme MUSIC pour la localisation des sources et l'algorithme proposé par Parra et Alvivo [63] pour la séparation.

De 2008 à 2010, Nakajima *et al.* dans [58, 60, 61] présentent un algorithme de séparation de sources géométrique en réglant automatiquement le pas d'avancement et le poids de l'algorithme d'optimisation. Cet algorithme a été implémenté dans ASIMO (implémenté notamment dans HARK). Dans [78], Takahashi *et al.* proposent une modification dans le module de séparation de sources géométrique utilisé dans HARK et implémenté dans le robot HRP-2. Ils utilisent les fonctions de transfert de tête précalculées du robot dans le GSS pour estimer les matrices de séparation.

Conclusion

Nous avons présenté un aperçu de l'état de l'art relatif à la séparation de sources audio dans un mélange convolutif. Dans un premier temps, nous sommes intéressés aux algorithmes de séparation en les classant selon leurs critères de séparation.

Ensuite, nous nous sommes focalisés sur les algorithmes de séparation de sources qui ont été utilisés dans le cadre qui nous intéresse dans cette thèse, c'est à dire l'audition de robots. Dans les deux chapitres suivants, nous nous intéressons à deux types d'algorithmes de séparation de sources : des algorithmes basés sur l'information spatiale des sources et les algorithmes basés sur l'information structurelle des sources (la parcimonie des signaux sources dans le domaine temps-fréquence).
