

Le son en trois dimensions

Son en relief, les différentes tentatives

Né à la fin des années 1800, le principe du son en relief basé sur deux canaux s'est standardisé avec le disque microsillon stéréo à la fin des années 50. Le home cinéma et le son à haute définition SACD et DVD-Audio ont multiplié le nombre de canaux, mais sans combler pour autant une lacune, le son en relief incluant la possibilité d'une vraie localisation en largeur, en profondeur comme en hauteur.

Par Jean Hiraga



Démonstration de son en relief 5.0 et 7.0 canaux à composante verticale, proposée par Philippe Pélissier en janvier 2005 sur les bases d'un procédé expérimenté initialement en 1979. Les enceintes disposées sur plan vertical forment un trapèze avec trois enceintes inférieures et deux supérieures.

Photos : Jean Hiraga

Comme pour l'image en relief, le son en vrai relief est un rêve qu'un très grand nombre de chercheurs ont cherché à concrétiser tout en butant chaque fois sur de multiples problèmes étalés entre la recherche du financement d'un projet jusqu'à l'éventualité d'une commercialisation liée, par la force des choses, à des paramètres de standardisation, de technologie, de coût de revient, de fiabilité ou de facilité d'utilisation. Lors du Salon International de l'Electricité qui se déroula à Paris en 1881, Clément Ader, l'un des pères de l'aviation et du moteur à explosion, fit sensation avec la transmission directe, depuis la scène de l'Opéra Garnier, d'un son stéréophonique sur deux canaux. Les visiteurs furent stupéfaits d'entendre se former autour d'eux des sources sonores virtuelles à partir d'écouteurs conçus eux aussi par Clément Ader. Il était difficile de faire plus court et plus simple, car les microphones étaient reliés directement aux écouteurs, étant donné que les premiers tubes triodes et les premiers montages amplificateurs ne furent inventés qu'un quart de siècle plus tard.

En 1931, un génie anglais, Alan Dover Blumlein, fut le premier à déposer une demande de brevet relative à un système de gravure et de reproduction phonographique basé sur le principe stéréophonique deux canaux selon la méthode dite 45/45. En 1933, ce brevet fut accepté sous la référence 394 325. Citons au passage qu'il fallut attendre la fin des années 50 pour que ce procédé se standardise sous la forme du disque microsillon stéréophonique. Une idée concurrente, celle de l'américain Cook qui proposa en 1954 le principe des deux cellules et du disque muni de deux plages distinctes, possédait trop d'inconvénients pour pouvoir être commercialisée. A partir de 1933, aux États-Unis, William B. Snow, des laboratoires de la Bell Labs,

fut le premier à aborder une étude sérieuse relative à l'enregistrement et à la reproduction du son en relief. Elle commença par l'étude de la fonction de transfert de la tête, la fameuse fonction HRTF (Head Related Function Transfer) dont on fait tant état aujourd'hui, puis par celle du pouvoir de localisation de nos oreilles dans le plan de la largeur et de la profondeur. L'expérience décisive fut la mise en place d'une scène carrée de 9 m de côté sur laquelle prenaient place trois rangées de 3 sources sonores écartées d'environ 1,9 m. L'enregistrement de ces neuf sources sonores s'effectuait sur les bases de deux ou trois microphones (gauche, centre et droit). Le son était restitué par deux ou trois haut-parleurs (gauche, centre et droit également), dans une grande salle. Un panel d'auditeurs était chargé d'évaluer en aveugle la position virtuelle de chaque source sonore selon le mode d'enregistrement et de reproduction. La grande surprise fut de constater que l'utilisation d'une paire de microphones G + D pour l'enregistrement suivie de la reproduction à partir de deux enceintes apportait une perte sensible de la localisation dans le sens de la profondeur, en particulier sur les sources centrales pour lesquelles il fut constaté un effet de masque des sources centrales situées au second et au troisième rang par celle située au centre, au premier rang. La localisation la plus précise dans le plan de la largeur et de la profondeur était obtenue en adoptant le principe des trois microphones et des trois haut-parleurs. Quelques années plus tard, Paul W. Klipsch fut parmi les premiers à revendiquer l'importance de ces recherches, de la stéréophonie à trois canaux en proposant un système simplifié, un peu moins précis, sur lequel, lors de la reproduction, le canal central était obtenu à partir du

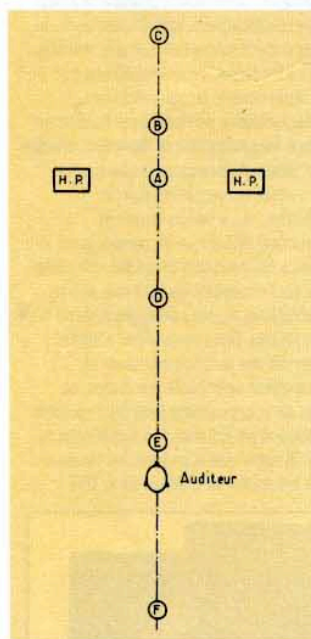


Figure 1 : en stéréophonie 2 canaux standard, il ne faut pas confondre effet de profondeur et vraie localisation en profondeur, cette dernière devant permettre de localiser non seulement un son aux points A, B ou C, mais également au point E (juste devant l'auditeur) ou F (derrière lui).

mixage des canaux gauche et droit. Cette expérience a été détaillée dans l'article intitulé "Le son en relief, 2 oreilles = 2 haut-parleurs? Pas si simple!" de notre revue n° 260 de février 2002. Après l'arrivée du disque microsillon stéréophonique à la fin des années 50, l'ingénieur Robert Fine assisté de Wilma Cozart surent tirer le meilleur parti de ces recherches en adoptant le principe des trois microphones gauche-centre-droit pour leurs prises de son leurs sur la célèbre série Mercury Living Presence, laquelle étonne toujours par le réalisme tridimensionnel de ses enregistrements.

L'avènement du home cinéma et du son multicanaux a permis de bénéficier sur les voies frontales des avantages offerts par ces recherches. Rappelons cependant que le son issu des trois voies frontales est souvent le résultat d'un savant mixage destiné à créer seulement des sources virtuelles en rapport avec l'image. Bien plus intéressante a été la vulgarisation du son haute définition aux formats DVD-Audio et surtout SACD, ce dernier étant un standard auquel se rallient aujourd'hui plus de 280 éditeurs pour un total proche de 2900 titres, ce qui correspond au succès obtenu par le disque compact trois années après sa lancée. En Super Audio CD, plusieurs labels comme Pentatone en Hollande ou Fine NF au Japon respectent à la lettre le principe de prise de son basé sur les 5 canaux, ce qui permet d'hériter, en plus des qualités de définition et de dynamique permises par ce standard, de la possibilité de localisation frontale en profondeur basée sur les recherches de William B. Snow.

Du son multicanal à la recherche du vrai relief sonore tridimensionnel

A la fin des années 70 on proposa, pour contrer la baisse des ventes du disque microsillon stéréophonique, différents procédés de son tétraphonique sur 4 canaux. Les démonstrations selon le procédé SQ proposées par Sansui s'effectuaient parfois en plaçant les quatre enceintes à une certaine hauteur, en dirigeant celles-ci vers le bas, donnant ainsi à des enregistrements de passages d'avions l'impression d'entendre ces avions passer au-dessus de la tête des auditeurs. En dehors du fait que la trop grande multiplicité des standards en son tétraphonique finit par conduire à un "flop" magistral, des chercheurs pointèrent le doigt sur le fait que le principe de la stéréophonie à deux canaux et de la tétraphonie ne procurait finalement qu'une illusion du relief sonore, car n'étant en réalité qu'un son à une ou à deux dimensions.

Stéréophonie, son en une dimension ?

Si stéréophonie signifie "son en relief", le principe à deux canaux reste, malgré les apparences, un son en une dimension. Même si, sur le plan frontal, les aménagements sur les bases de trois canaux à l'enregistrement avec remixage de la voie centrale sur les voies gauche et droite pour la reproduction permettent effectivement une certaine localisation en profondeur sur le plan frontal, la stéréophonie sur deux canaux reste un son en une dimension. C'est la largeur, avec la distorsion de géométrie que l'on connaît. Pour mettre en évidence cette distorsion de géométrie, une méthode consiste à enregistrer à l'aide d'une paire microphonique un son se déplaçant de gauche à droite sur une scène, puis à le reproduire sur une paire d'enceintes. La localisation de la source sonore réelle est excellente, en particulier dans le cône d'attention de l'auditeur, surtout lorsque le son est reproduit dans un environnement acoustiquement mat. La même source, écoutée sur une paire d'enceintes, montre que l'impression de source virtuelle vraiment "punctiforme", assimilable à un point dans l'espace, n'est obtenue que lorsque cette source provient de l'enceinte gauche ou droite (en prenant l'exemple d'enceintes de qualité). Si la source virtuelle centrale reste assez précise en localisation, on s'aperçoit que le déplacement de la source de gauche à droite n'est pas du tout conforme à la réalité, le son ayant tendance à "sauter" de la gauche vers le centre puis du centre vers la droite en passant par des zones de flou. Ce phénomène de distorsion géométrique dans le plan hori-

zontal explique pourquoi des ingénieurs du son comme André Charlin allaient jusqu'à déplacer la position des instruments sur la scène jusqu'à obtention d'un positionnement plus réaliste d'un orchestre entre les enceintes. La figure 1 montre qu'à partir du point d'écoute, seule la source virtuelle de référence A, située entre les enceintes, est localisable, une véritable localisation en profondeur signifiant que, si l'on prétend pouvoir localiser réellement des sources positionnées derrière les enceintes, aux points B ou C, la transcription d'une seconde dimension, la profondeur, signifie que l'on doit être en mesure de localiser des sources situées aux points D ou même E (à quelques centimètres devant l'auditeur) voire même au point F, c'est-à-dire derrière l'auditeur. Ce genre de test montre bien que le son "stéréophonique" sur 2 canaux reste un son en une dimension et qu'il ne faut pas confondre effet de profondeur et vraie localisation en profondeur.

Tétraphonie, stéréophonie à 5 ou 7 canaux, du son en deux dimensions ?

L'expérience représentée sur la figure 2 montre qu'en tétraphonie ou en son "stéréophonique" multicanaux, les mêmes limitations se constatent. La source virtuelle peut, à volonté, provenir d'une des quatre enceintes ou de positions intermédiaires situées dans le plan horizontal, en se déplaçant sur les quatre lignes qui encadrent les enceintes. Une vraie localisation dans deux dimensions, la largeur et la profondeur, signifie, comme le montre cette figure, qu'une source virtuelle mobile de caractère vraiment punctiforme censée suivre le trajet A est impossible à transcrire entre quatre, cinq ou sept enceintes disposées dans le même plan horizontal.

Octophonie, du son en trois dimensions en mode "restreint" ?

La figure 3 montre qu'en octophonie, un principe basé sur une prise de son à huit microphones disposés chacun aux huit sommets d'un cube et reproduit à partir de huit enceintes placées aux huit angles d'une pièce, l'auditeur situé au centre peut localiser le son émis par chacune des huit enceintes et également d'autres sources virtuelles intermédiaires localisées au niveau des surfaces inscrites dans chacune des six surfaces formées par les huit enceintes. En revanche, comme pour l'expérience précédente, l'auditeur sera incapable de localiser une source virtuelle aux contours précis se déplaçant selon le trajet A : au lieu d'entendre la source se rapprocher de lui, l'auditeur entendra un son plus fort émis par quatre enceintes latérales, puis par les huit, avant que ce son ne bascule sur les quatre enceintes opposées. Cette "octophonie", que certains pourraient appeler le "son en relief du futur" reste seulement une approche du son en trois dimensions, mais sous une forme restreinte.

Phénomènes subjectifs et illusion d'effet vertical en stéréophonie deux canaux ou multicanaux

Au Japon, au début des années 80, la N.H.K. proposa, à titre expérimental, des programmes télévisés dont les images en couleurs pouvaient être obtenues à partir d'écrans noir et blanc. Sur des dessins animés réalisés à cet effet, il avait été mis à profit l'effet stroboscopique, la persistance rétinienne et des effets combinés de fausses couleurs découverts par Ewald Hering, formant ainsi des images en couleurs "imaginaires", peu nombreuses et de teinte pâle. La N.H.K. continua ce genre d'expérience en adoptant cette fois les illusions d'effet stéréophonique vertical sur les bases de 2 canaux, une expérience rendue possible grâce à l'arrivée du son stéréophonique FM multiplexé. Bien qu'excellente, l'idée n'était pas nouvelle. Sur les premiers enregistrements stéréophoniques, on s'était aperçu que, pour mieux centrer un soliste entre les enceintes tel qu'un violon ou une chanteuse, il était fréquent de les enregistrer en monaural, puis de les remixer sur les canaux gauche et droit en y ajoutant un peu de réverbération pour mieux les fondre dans la masse orchestrale. C'est suite à cette expérience que l'on s'aperçut qu'il se produisait sur ces sources centrales et dans certaines plages de fréquences, un effet d'élévation de la source virtuelle. Il n'est pas rare de noter ces effets sur deux

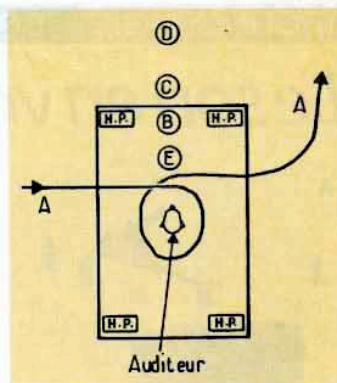


Figure 2 :

Avec le principe de la stéréophonie à 4 canaux ou tétraphonie, on pourrait penser pouvoir accéder à une localisation réelle en profondeur, ce que l'on pourrait croire en écoutant des sons provenant des points B ou E. En fait, il ne s'agit que d'un effet de profondeur lié à l'esthétique, à la forme de sons plus ou moins emprunts de réverbération, vu que la transcription réelle d'un son dans deux dimensions devrait permettre à la source mobile A de se déplacer hors du cadre formé par les quatre enceintes tout en formant une source parfaitement punctiforme, c'est-à-dire assimilable à un point dans l'espace.

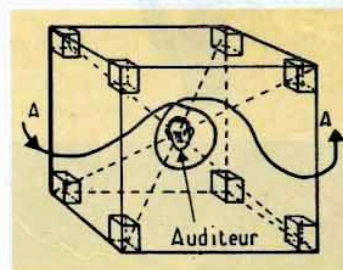


Figure 3 :

Avec le principe de l'octophonie à 8 haut-parleurs, on peut localiser le son provenant de chacune des haut-parleurs de même que celui provenant de directions intermédiaires. Des essais montrent toutefois que, même dans cette configuration "3D" pouvant être considérée comme idéale, la localisation se limite aux 6 surfaces formées par les 8 haut-parleurs. Comme en tétraphonie, on s'aperçoit que la localisation en direction et en distance est impossible, ce qui explique pourquoi on ne peut pas transcrire une source en mouvement comme le ferait la source A sur la figure : au lieu de se déplacer librement dans l'espace en donnant l'impression de traverser la pièce, le son se localise sur une face encadrée de quatre haut-parleurs, on ressent ensuite progressivement le son apparaître au niveau des quatre haut-parleurs du côté opposé, mais on n'obtient pas l'effet recherché, une source ponctuelle aux contours précis s'approchant de l'auditeur, se déplaçant autour de lui avant de l'éloigner dans une autre direction.

Le son en vrai relief : les différentes tentatives



Figure 4 :
En A, un système de plateforme d'enregistrement 5 canaux pour le SACD, le Atmos 5.1 de la firme allemande SPL (www.soundperformancelab.com).

En B, un autre système de prise de son en 5.1 ou en 7.1, le H2-Pro commercialisé par la firme américaine Holophone (www.holophone.com).

En C, enregistreur Yamaha DM 2000 utilisé par Philippe Pélissier pour ses démonstrations de stéréophonie à composante verticale 5.0 ou 7.0. Photo réalisée à Radio France, au Studio 107, en janvier 2005.

canaux, une des raisons étant un effet d'élévation des sources frontales ressenti par certains sujets. J. Blauert en 1969), l'autre concernant un autre phénomène dont il sera question à la fin de cet article. Découvert par hasard par une équipe de la RTF (Radio Télévision Française) et baptisé "Effet J" (commande de dosage d'écho sur la source centrale d'une table de mixage), cet effet fut mis à contribution au début des années 60 pour un spectacle Son et Lumières qui se déroula dans le parc du Château de Versailles : les spectateurs furent très étonnés d'entendre, à un moment donné, une voix, celle de Dieu, qui semblait provenir du ciel.

Beaucoup de chercheurs dont J. Blauert aux Etats-Unis, M. Tsujimoto et M. Nakabayashi (N.H.K., Japon) étudièrent ce sujet de près. Dans ses suites d'émissions expérimentales, la N.H.K. proposa des séquences filmées ou des dessins animés au cours desquels la simple paire d'enceintes situées de part et d'autre de l'écran du téléviseur suffisait pour donner l'impression d'un train ou d'un vaisseau spatial passant au-dessus de la tête du spectateur. A partir de 1992, la firme américaine SRS Labs, mit à profit des idées du même genre sur ses processeurs équipés à leurs débuts de deux commandes variables, l'une pour l'effet d'élargissement de l'effet stéréophonique dans le plan horizontal, l'autre pour le plan vertical, le tout poussé à fond permettant d'obtenir un effet global similaire à une sorte de portion de géode sonore. Toutes ces expériences ont servi de base pour élaborer par la suite différents processeurs sonores dont ceux, très actuels, conçus pour créer à partir d'une seule enceinte placée devant l'auditeur, un effet de couverture sonore proche de celui obtenu en 5.1, un inconvénient étant de contraindre l'auditeur à rester bien au centre, face à l'enceinte. Entre temps, pour les besoins de la prise de son multicanaux, les systèmes de plateformes microphoniques multicanaux se sont multipliés, deux exemples étant le système allemand Atmos 5.1 de SPL et le système américain Holophone (www.holophone.com). On trouvera les photos concernées sur la figure 4.

Une autre approche du son en relief : le système Surround de Philippe Pélissier

A Radio France, dans ce lieu polyvalent et prolifique en inventions, plusieurs chercheurs ont essayé d'apporter à la stéréophonie d'autres dimensions. Il est tout d'abord essentiel de retenir que, dans une salle de concert, l'auditeur est plongé dans un environnement de sons directs et diffus dans lequel le cône d'attention couvert par le champ de vision privilégie un plan sonore vertical situé devant lui. Il perçoit, au niveau du sol ou d'une suite d'estrades de hauteur appropriée, les différents pupitres qui composent l'orchestre ainsi que leur positionnement respectif dans le plan horizontal ou vertical. C'est le cas des percussions, à la fois surélevées et situées très en arrière-plan par rapport aux pupitres de cordes. Ce pouvoir de localisation dans le plan vertical de notre système auditif est encore plus évident lors de l'écoute d'un chœur situé sur des gradins très surélevés et positionnés derrière l'orchestre. Suite à cette constatation et à l'absence d'effet vertical sur la stéréophonie deux canaux classique, Philippe Pélissier, musicien et metteur en ondes à Radio France a cherché à remplacer cette ligne sonore horizontale en un plan vertical, tout en y ajoutant les bénéfices d'une possibilité de localisation (et non pas d'effet) en profondeur. En 1979, à une époque en plein effervescence au niveau des différents standards proposés en tétraphonie, il expérimenta un jour une prise de son de type "plan sonore vertical" pour essayer de s'affranchir des effets de masque engendrés par les premiers plans de l'orchestre, nécessitant d'ordinaire la mise en place de microphones dits "d'appoint". Rappelons que sur les grandes formations orchestrales, il n'est pas du tout évident de se passer de ces microphones d'appoint dès l'instant où il est néces-

saire de préciser plusieurs instruments qui tendent à trop se fondre dans la masse orchestrale. Il est important de mentionner ici que, lors de la prise de son sur un plan vertical, Philippe Pélissier a respecté le principe de la base stéréophonique à trois microphones dont il a été question plus haut, deux autres microphones un peu moins écartés prenant place plus haut. Le tout se complète, pour certains enregistrements, de deux canaux arrière prévus en option. Parmi les différentes prises de son d'époque et récentes proposées, celle de "La Vierge" de Massenet, réalisée à l'Eglise de la Trinité à Paris a révélé un réalisme tridimensionnel stupéfiant, bien supérieur à celui obtenu à partir d'un système conventionnel à 5 ou à 7 canaux : localisation des plans sonores des premiers rangs de l'orchestre, des derniers rangs surélevés, d'une voix très haut perchée chantée depuis un niveau supérieur de l'église, absence d'effet de masque des plans entre eux. Différents essais ont montré que la suppression de la voie centrale rendait les plans sonores moins précis et nécessitaient de se placer parfaitement au centre. L'ajout des voies arrière n'a apporté qu'un certain complément de réalisme tridimensionnel, l'essentiel étant apporté par les 5 canaux disposés verticalement. On trouvera, sur la figure d'entrée de cet article, une photo de cette démonstration en 5.0 présentée à Radio France en mi-janvier 2005, à l'occasion du Satis.

Ajoutons, à ce propos de cette expérience de stéréophonie à composante verticale, qu'un autre passionné de prise de son et de son en relief, Jacques Chaponnay, a proposé lui aussi un système proche, mais sur les bases de 4 microphones et de 4 haut-parleurs, lequel a été déposé le 17 juillet 1990 à l'INPI sous l'appellation "Holophonie".

A propos de la localisation dans le plan vertical des enceintes multi-voies

Ce sujet a fait jusqu'à présent beaucoup parler de lui à travers différents articles étalés sur plusieurs décennies. Le plus récent, rédigé par Sam Ferguson et Densil Cabrera de l'Université de Sydney, en Australie, vient d'être publié dans la revue de l'AES (Volume 53, number 3, mars 2005). Les essais effectués en chambre anéchoïque ont consisté à placer de manière invisible, derrière un rideau acoustiquement transparent, des enceintes 2 voies alignées verticalement, le test consistant à demander à un panel d'auditeurs placés devant ou de côté la hauteur virtuelle de la source. Devant le rideau occultant, une grande règle positionnée verticalement servait à mieux préciser l'endroit d'où semblait provenir le son. Cette étude a mis en œuvre une procédure très stricte et différentes sortes de stimuli sonores sollicitant simultanément ou séparément les registres de grave-médium ou d'aigu sous forme de sons continus ou bien encore d'impulsions de bruits. L'analyse des résultats et des erreurs importantes de localisation a conduit en majeure partie aux mêmes conclusions que celles d'autres chercheurs ayant étudié ce domaine comme C.C. Pratt (Caractère spatial des sons graves et aigus, 1930), S.K. Roffler et R.A. Butler (Localisation des sons dans le plan vertical, JASA, 1968) ou M. Morimoto et H. Aokata (localisation des sources sonores dans l'hémisphère supérieur, JASA, 1984). Il a surtout été noté un effet de dissociation marqué des sources virtuelles graves et aiguës sur des systèmes 2 voies sur lesquels la fréquence de coupure haute était supérieure à 500 Hz.

En guise de conclusion

La course au son en vrai relief est un sujet passionnant qui a conduit à l'étude, puis à la commercialisation de plusieurs procédés, depuis ceux que nous connaissons actuellement en passant par des systèmes expérimentaux à 12, voire à 24 canaux distincts. Le procédé Pélissier reste, compte tenu des moyens mis en œuvre une approche particulièrement intéressante.

Jean Hiraga