

Introducció de l'elevació als sistemes de reproducció de so espacial basats en Wave-Field Synthesis mitjançant l'ús de l'HRTF.

Autor: Cabanes Sempere, Maria

Director: López Monfort, José Javier

Abstract

The study of elevation is an interesting subject of research in many signal processing applications, ranging from location and reproduction of sound sources to object location in electromagnetic exploration systems. Applied to the Head Related Transfer Function HRTF, it can be used to get 3D spatial reproduction systems based on Wave-field Synthesis (WFS). WFS system is a spatial sound reproduction system that can synthesize an acoustic field by means of horizontal loudspeaker arrays. However, even when several horizontal arrays are located in parallel in the vertical dimension, the desired reproduction effect cannot be obtained in elevation. In this work, we propose to design HRTF filters in order to fill the gap in the reproduction in elevation. Some HRTF filters have been designed by different methods, having been evaluated by subjective tests that show the high complexity of recognizing or locating sound sources in elevation in the space.

Resum

L'estudi de l'elevació és un tema interessant a la investigació en moltes aplicacions de tractament de senyals, que van des de la localització i reproducció de fonts sonores a la localització d'objectes en sistemes d'exploració electromagnètica. Aquest estudi s'aplica a la Funció de Transferència del Cap per a poder obtenir sistemes de reproducció espacial 3D basats en *Wave-field Synthesis* (WFS). Els sistemes WFS poden sintetitzar un camp acústic mitjançant agrupacions d'altaveus en horitzontal. No obstant, al col·locar diverses agrupacions en paral·lel, l'efecte aconseguit en el pla vertical és molt distant del desitjat. En aquest treball es proposa el disseny de filtres HRTF per omplir el buit de la reproducció en elevació. S'han dissenyat filtres per diferents mètodes i han sigut avaluats mitjançant tests subjectius que demostren com de complex és el reconeixement o localització de fonts sonores a l'espai en elevació.

Autor: Cabanes Sempere, Maria, email: macasem@dcom.upv.es

Director: López Monfort, José Javier, email: jjlopez@dcom.upv.es

Data d'enviament: 07-09-08

A Pau.

Índex

1	Introducció	1
2	Funció de transferència del cap, HRTF	4
2.1	HRTF mesurades front a modelades	4
2.2	Models d'HRTF	5
3	Wave Field Synthesis	7
3.1	Integrals de Kirchhoff-Helmholtz i Rayleigh	8
3.2	Derivació de la funció del senyal director	12
3.3	Potencial i restriccions de WFS	16
3.4	Propietats especials de WFS	17
3.5	Implementació pràctica de WFS	19
4	Disseny de Filtres HRTF en elevació	20
4.1	Anàlisi de dades	20
4.2	Síntesi de dades	21
4.2.1	Síntesi I: consideració dels angles d'elevació junts	23
4.2.2	Síntesi II: consideració dels angles d'elevació per separat	24
4.3	Exaltació de filtres	25
4.4	Filtrat RSS	27
5	Aplicacions de l'HRTF a tècniques WFS	28
5.1	Avaluació subjectiva	28
5.2	Resultats	29
6	Resum i conclusions	31
A	Optimització de funcions	32
A.1	Teoria general de minimització de funcions	32
A.2	Mètode de descens del símplex	33
	Bibliografia	37

Capítol 1

Introducció

En el camp de la reproducció de so 3D, es vol que l'oient experimente una immersió vertadera en un entorn acústic, a través de les tècniques de reproducció. Per aquesta raó resulta imprescindible conèixer com funciona el sistema d'audició humana, per a què d'aquest mode es puguem comprendre els sistemes de reproducció 3D i es puga investigar i desenvolupar en aquest camp.

Molts científics s'han encarregat d'investigar el funcionament de l'oïda humana, per a poder modelar-la de forma que puga tractar-se com un sistema més, mitjançant les tècniques de l'acústica i del processat de senyal [5]. Aquests estudis han possibilitat el desenvolupament dels sistemes de so 3D.

Com a conseqüència de les investigacions fetes s'han desenvolupat sistemes electrònics de reproducció que intenten reconstruir la sensació espacial del so. Fins a dia d'avui, s'ha aconseguit una reproducció molt realista al pla horitzontal fent ús de dos factors que influeixen en la localització al pla horitzontal: la Diferència de Temps Interaural (ITD) i la Diferència d'Intensitat Interaural (IID). No obstant al pla vertical és més complicat obtenir l'efecte desitjat, a causa que no es poden fer servir els mitjans binaurals que s'utilitzen per a la detecció en azimuth. En aquest cas, els mitjans són monoaurals.

Els mitjans en elevació es basen en el fet que el pavelló auricular (orella) actúa com una antena acústica, és a dir, poseeix certa directivitat. Les cavitats ressonants de l'orella amplifiquen certes freqüències, la geometria dels plecs de l'orella, com es veu a la Figura 1.1, produeix efectes d'interferència que atenuen altes freqüències.

A més aquesta resposta en freqüència és dependent de la direcció amb la que arriben els sons al cap. En la Figura 1.2 es mostren les respostes en freqüència mesurades per a dos angles d'arribada a l'oient en elevació diferents. En la Figura 1.1 es pot observar que hi ha dos camins des de la font sonora fins el canal auditiu: un camí directe i un camí més llarg seguit d'una reflexió a l'orella. A freqüències baixes moderades l'orella recull energia sonora addicional i els senyals dels dos camins arriben en fase. En canvi, a altes freqüències, el senyal retardat està desfasat amb el senyal directe, i es produeix una interferència destructiva. La major interferència ocorre quan la diferència en la longitud de camins d , és la meitat de la longitud d'ona, és a dir, quan $f = c/2d$. En l'exemple mostrat, es produeix un nul al voltant de 10 kHz. Per a valors típics de d el nul de freqüències es troba normalment entre 6 kHz i 16 kHz.

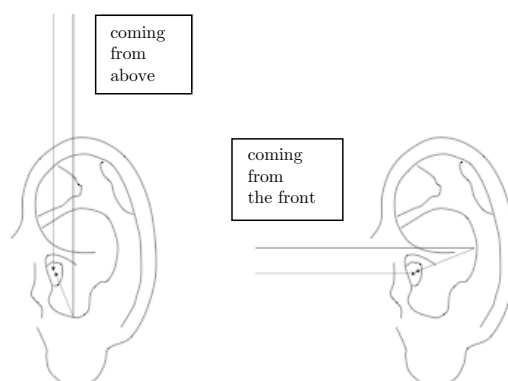


Figura 1.1: Reflexions en el pavelló auricular per a dos angles d'elevació diferents. En la Figura de l'esquerra la incidència del so es produeix des del zenit, mentre que en la dreta es produeix de davant de l'oient. Les fletxes indiquen els camins directes als plecs de l'orella.

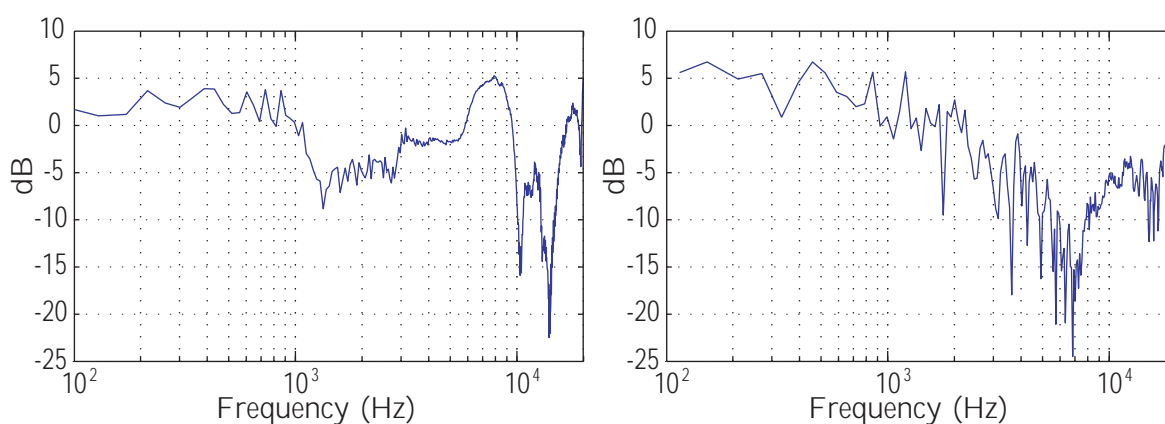


Figura 1.2: Resposta en freqüència del cap, per a dos angles d'elevació diferents en la incidència del so. Les dues gràfiques es corresponen respectivament amb els casos mostrats a la Figura 1.1. Es pot apreciar la gran atenuació que es produeix a 8 kHz en la incidència frontal.

Donat que l'orella és millor reflector per als sons que arriben des del front que des dels que arriben des de dalt, el nul és molt més pronunciat per a les fonts situades en front que per a les fonts que es situen dalt. A més, les diferències de camins varien amb l'angle d'elevació i en conseqüència el nul es desplaça en freqüència amb l'elevació. A pesar d'açò, existeixen encara debats respecte a quines característiques són perceptualment més importants, encara que s'admet que l'orella proporciona els mecanismes de percepció de l'elevació.

Tenint en compte tot l'exposat abans, en aquest treball s'intentarà dissenyar filtres de la millor forma, per a que puguin ser aplicats als sistemes existents de *Wave Field Synthesis* (WFS) i s'obtinguen els efectes d'elevació desitjats a la reproducció. Aquest treball està estructurat de la següent forma: en el Capítol 2 es veu un repàs sobre la teoria que fa possible tractar l'oïda com un sistema

de processat de senyal, seguidament en el Capítol 3 és veuen els principis dels sistemes de síntesi d'ones. Al Capítol 4 es descriuen els filtres dissenyats per a dur a terme aquest treball. Finalment al Capítol 5, es realitzen diferents experiments on s'apliquen els filtres dissenyats als sistemes WFS i les conclusions obtingudes d'aquests experiments es mostren al Capítol 6.

Capítol 2

Funció de transferència del cap, HRTF

El sistema auditiu pot modelar-se com qualsevol sistema lineal i invariant per a conèixer el senyal que arriba a l'oïda. La resposta a l'impuls d'aquest sistema és coneguda com a HRIR (Head Related Impuls Response), i la seua transformada de Fourier $H(f)$, és coneguda com HRTF (Head Related Transfer Function).

L'HRTF comprèn tots els aspectes físics de la localització del so per a una posició concreta del cap (diferència interaural, ombra del cap, difraccions als plecs de l'orella, canal auditiu,...). Una vegada es coneixen les HRTF per a l'oïda esquerra i la dreta, i la posició d'una font virtual en l'espai, es poden sintetitzar els senyals corresponents a l'oïda per a un so emès per eixa font.

L'HRTF és una equació complexa de quatre variables: tres coordenades espacials i la freqüència, com es veu a l'equació (2.1). En coordenades esfèriques, com es veu a l'equació (2.2), per a distàncies superiors a un metre, es suposa que la font es troba en camp llunya i es suposa que l'HRTF cau inversament amb la distància. La majoria de les mesures de l'HRTF es fan en camp llunyà, i per tant es redueix l'HRTF a una funció d'azimut, elevació i freqüència, tal com es veu a l'equació (2.3).

$$HRTF = H(x, y, z, f) \quad (2.1)$$

$$HRTF = H(\phi, \theta, d, f) \quad (2.2)$$

$$HRTF = H(\phi, \theta, f) \quad (2.3)$$

2.1 HRTF mesurades front a modelades

Cada persona té una HRTF diferent, perquè físicament les mesures dels seus òrgans auditius són diferents. Per tant, per a una bona simulació de fonts virtuals, es necessitaria conèixer les HRTF individualitzades de la persona sobre la que es vol produir la impressió sonora.

No obstant, per a sistemes comercials o d'ús generalitzat, açò no és possible o bé resulta molt

dificultós. Per tal motiu es recorre a utilitzar models més o menys senzills de l'HRTF que funcionen amb la major part d'oients.

A causa que les HRTF són molt complexes, molts sistemes de so espacial depenen de dades mesurades experimentalment, com són les mesures del maniquí acústic KEMAR realitzades en el M.I.T. per Bill Gardner [12]. La principal raó d'utilitzar mesures de les HRTF és per a capturar els efectes d'elevació i azimuth. No obstant els aspectes de localització de l'elevació són particularment més sensibles a cada persona, ja que depenen de la morfologia del pavelló auricular. Per resoldre aquest problema han aparegut quatre aproximacions diferents:

- Utilitzar una HRTF estàndar de compromís. Açò produeix resultats pobres en elevació, però és útil per a sistemes de baix cost.
- Utilitzar una HRTF d'entre un conjunt de HRTF estàndar. Açò requereix mesurar les HRTF d'una petita quantitat de persones, que siguin representatives, cadascuna d'aquestes, d'un grup de població diferent. A més s'ha de proposar un mètode senzill per a seleccionar l'HRTF d'entre totes per a un oient donat. A dia d'avui, no existeix un conjunt d'HRTF estàndar.
- Ús de l'HRTF individualitzada. Açò requereix la mesura de l'HRTF de cada oient, amb les molèsties que açò pot ocasionar i el temps necessari per a mesurar-les. No obstant açò produeix excelents resultats.
- Ús d'un model d'HRTF que continga paràmetres que es puguin adaptar a cada oient individualment.

En aquest cas s'utilitzarà la primera solució proposta. S'utilitzarà una HRTF estàndar de compromís, obtinguda com a mitja de una quantitat donada de subjectes. A partir d'aquesta HRTF es cercaran els filtres necessaris per a sintetitzar-la.

2.2 Models d'HRTF

L'objectiu dels models d'HRTF és aconseguir un sistema que represente al conjunt de funcions de transferència relatives al cap d'una forma simplificada i el més ajustada possible. Existeixen diversitat de models que ho aconsegueixen amb més o menys èxit, però es poden classificar en tres grans grups [8, 10]:

Funció Racional o model de pols i zeros. El problema del modelat s'aborda com un problema d'identificació de sistemes, amb diversos mètodes clàssics per a obtenir la solució. Les funcions que apareixen, solen ser molt complicades i s'han de tabular per a diferents elevacions i azimuth, minvant la seua flexibilitat.

Expansió en sèries. L'expansió en sèries de Fourier o en sèries de Krhunen-Loeve (també conegut com ànlisi de components principals o PCA) permeten una representació de l'HRTF com una suma ponderada de funcions simples i bàsiques. Encara que és útil per examinar dades, la complexitat del processat en temps real de cada model limita la viabilitat del mateix.

Models estructurats. És un intent de crear una funció de transferència que es corresponga fidelment al mecanisme físic de localització (ombra del cap, reflexions als múscles...) [14].

De entre els models de l'últim grup es poden citar els següents:

Model ITD. Sols utilitza el mecanisme de diferència de temps interaural. El càlcul de la diferència de temps d'arribada es realitza a partir de l'equació 2.4, la qual és funció de l'angle d'arribada.

$$ITD = \frac{\tau}{c}(\theta + \sin \theta) \quad (2.4)$$

Model IID. S'utilitza el mecanisme de diferència d'intensitat interaural. El càlcul de l'efecte d'ombra del cap es realitza a través d'una equació obtinguda tenint en compte els efectes de difracció del cap i proposta per Rayleigh en [20].

Model de cap esfèric. Simplement es col·loca els models IID i ITD i s'obté una bona aproximació del model de cap esfèric. Encara que no hi ha sensació d'externalització o elevació, s'elimina el problema de la confusió davant/darrere.

Models de pavelló auricular. Amb objecte de simular l'elevació, s'intenten modelar els efectes de l'orella com un sumatori d'un o més ecos (reflexos) en els plecs de l'orella mitjançant filtres FIR en funció de la direcció. El problema consisteix en determinar la variació dels guanys i retardaments per a cadascun dels elements del filtre FIR. Aquest model és el que es veurà en aquest treball.

Models combinats. Mitjançant la combinació del model del cap i el model del pavelló (i sumant el model de difracció del tors, el model de la ressonància del canal auditiu, el model d'una habitació...) podem obtindre successivament millors aproximacions per al model de l'HRTF.

En els Capítols 4 i 5 es veuran els resultats dels experiments que s'han dut a terme en aquest treball amb l'objecte d'avaluar l'HRTF en elevació aplicada a sistemes WFS.

Capítol 3

Wave Field Synthesis

Wave Field Synthesis (WFS) és una tècnica de so espacial que genera un camp sonor vertader a partir d'una agrupació d'altaveus. La base de la síntesi de camps sonors està en col·locar fonts sonores virtuals en diferents posicions d'un escenari sonor, bé darrere de l'àrea que tanca l'agrupació d'altaveus, o dins d'aquesta, a la zona d'escolta. Aquestes fonts les anomenarem més tard fonts focals. Amb la reproducció del so amb WFS, els camps sonors poden generar-se de forma correcta espacial i temporalment. A conseqüència d'açò, l'oient pot experimentar el sentiment que l'origen de la font està a la posició de la font virtual.

Wave Field Synthesis pot ser utilitzar en aplicacions com són: realçament de so directe en teatres, cines, *home theaters*, teatres de realitat virtual, simuladors o sistemes de teleconferència.

El desenvolupament de la reproducció de so espacial va començar prompte. Els resultats obtinguts amb aquests sistemes depenien del material enregistrat i la forma de reproduir-lo.

En 1953, Snow va publicar una visió general de tècniques estereofòniques[23] i va proposar la cortina acústica com a tècnica ideal de reproducció estereofònica (veure Figura 3.1). El seu objectiu era transportar l'escenari enregistrat a una zona de reproducció amb l'ús d'agrupacions de micròfons i altaveus. Aquests científics es van adonar ràpidament, que degut als requeriments tècnics no era factible posar les seues idees en pràctica. Com a compromís van aplicar tres canals d'estereofonia i van acceptar que l'objectiu del que partien per recrear camps sonors reals ja no podria ser realitzat.

Als anys 80 el concepte de WFS va ser introduït per la Universitat Tècnica de Delft. L'origen d'aquesta història va ser en 1987 i 1988 quan Berkhout va publicar els llibres [1] i [2]. Ací va suggerir el concepte "*acoustical holography*", encara no conegut com WFS, com a última eina per a sistemes de control acústic en teatres. Berkhout va introduir les bases físiques del WFS al aplicar algorismes coneguts d'ones sísmiques al camp de l'acústica. El treball bàsic de WFS va ser continuat en [3] i [4].

WFS proveeix una notable millora en estereofonia en termes de reproducció espacial d'escenes sonores[24]. No obstant, abans de jutjar l'ús d'aquesta tècnica, han d'avaluar-se els avantatges i inconvenients respecte el sistema de percepció. Les següents seccions introduiran els conceptes de *Wave Field Synthesis*.

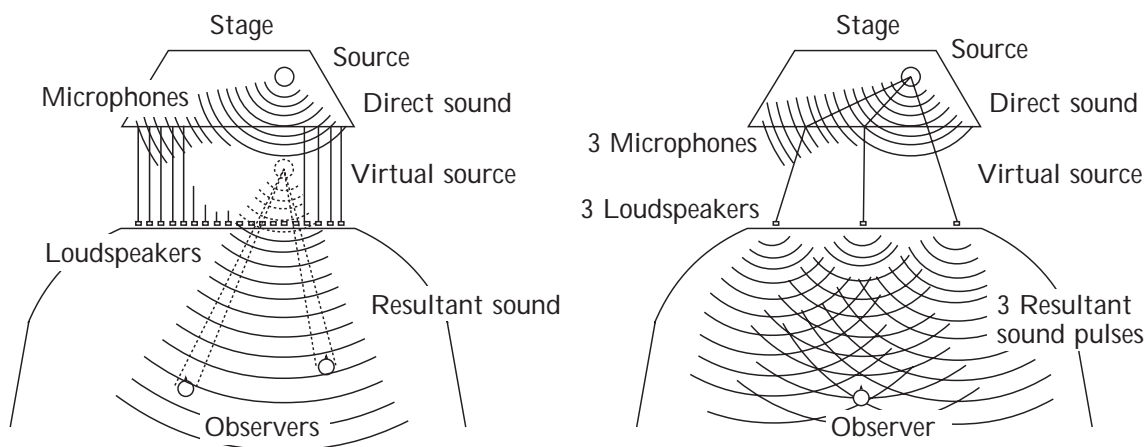


Figura 3.1: Sistema estereofònic dissenyat (esquerra) i implementat (dreta) de Snow, Steinberg (reproduït a partir de l'article de Snow [23]).

3.1 Integrals de Kirchhoff-Helmholtz i Rayleigh

La teoria de WFS està relacionada amb el principi de Huygens, formulat en 1678. Aquest principi diu que cada element d'un front d'ones que es propaga a través d'un medi particular pot ser vist com el centre d'una ona esfèrica individual. Conseqüentment el front d'ones generat per la font sonora primària pot ser vist com una sèrie de fonts secundàries elementals.

No obstant, no és gens pràctic posicionar les fonts acústiques en els fronts d'ones per a la síntesi. Si es situen els altaveus en una corba fixa arbitrària, i a cada altaveu li s'assigna un retardament i un pes adequats, el front d'ones acústic pot ser sintetitzat per una agrupació d'altaveus. La Figura 3.2a) mostra aquest principi.

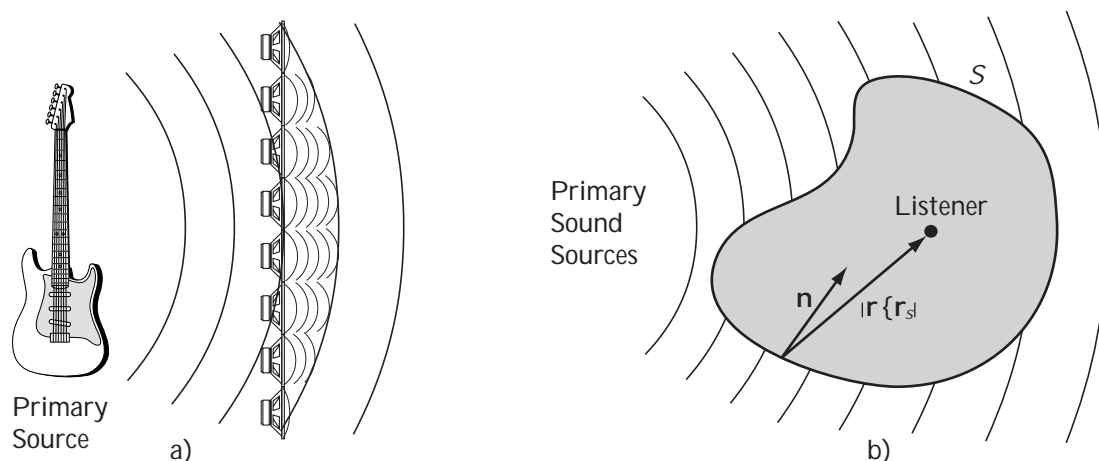


Figura 3.2: a) Principi bàsic de *wave field synthesis*, b) D'acord amb la integral de Kirchhoff-Helmholtz, el camp d'ones dins del volum lliure de fonts V pot ser definit pels paràmetres del camp mesurats al voltant del límit de la superfície tancada S [4].

Matemàticament, la pressió dins o fora de la superfície pot ser determinada com:

$$P(\mathbf{r}) = \iint_S \left[G \frac{\partial P(\mathbf{r})}{\partial n} - P(\mathbf{r}) \frac{\partial G}{\partial n} \right] dS, \quad (3.1)$$

on G és la funció de Green en espai lliure, donada per

$$G(\mathbf{r}|\mathbf{r}_s) = \frac{1}{4\pi} \frac{e^{jkR}}{R}, \quad (3.2)$$

i $R = |\mathbf{r} - \mathbf{r}_s|$. L'equació (3.1) planteja, que si es considera el problema interior, el camp acústic en V , generat per les condicions fora de la superfície S , pot ser processat únicament si es reemplacen aquestes condicions per una distribució de superfícies de monopols simples sobre $P(\mathbf{r})G(\mathbf{r}|\mathbf{r}_s)$ i si es sumen les contribucions sobre S . Per aquest motiu un camp d'ones acústiques arbitrari pot ser recreat dins d'un volum V , lliure de fonts, per fonts sonores secundàries distribuïdes en una superfície tancada en la seua frontera S (veure la Figura 3.2b). Aquest és expressat per l'anomenada integral de Kirchhoff-Helmholtz:

$$P(\mathbf{r}, \omega) = \oint_S \frac{1}{4\pi} \left[P(\mathbf{r}_s, \omega) \frac{\partial}{\partial n} \left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_s|}}{|\mathbf{r}-\mathbf{r}_s|} \right) - \frac{\partial P(\mathbf{r}_s, \omega)}{\partial n} \left(\frac{e^{-jk|\mathbf{r}-\mathbf{r}_s|}}{|\mathbf{r}-\mathbf{r}_s|} \right) \right] dS, \quad (3.3)$$

on $P(\mathbf{r}, \omega)$ és la transformada de Fourier de la pressió sonora $p(\mathbf{r}, t)$, k és el nombre d'ona, \mathbf{r} és el vector coordinat d'un punt d'observació i \mathbf{r}_s és el vector coordinat de les funcions integrades a la superfície S . La primera part de l'expressió representa una distribució de dipols amb intensitat de font donada per la pressió sonora, mesurada sobre la superfície S . El segon terme representa una distribució de monopols, la intensitat dels quals ve donada per una component normal de velocitat particular del camp sonor que és proporcional a $\partial P/\partial n$ [10].

En la representació gràfica de la integral de Kirchhoff-Helmholtz (Figure 3.2b) les ones sonores de les fonts sonores primàries externes són recreades dins del volum V . Al mateix temps el contorn de la superfície pot ser vist com un pla infinit que separa una part de l'espai amb fonts sonores de l'àrea d'escolta lliure de fonts i forma el volum V . En aquest cas (3.3) pot ser simplificada i el camp d'ones en l'àrea lliure de fonts pot ser representat per les integrals 3D Rayleigh I i II[1]. La integral Rayleigh I està formulada en termes d'una distribució planar de fonts monopols secundàries i la integral de Rayleigh II en termes de fonts dipol secundàries.

En la pràctica, la integral de Kirchhoff-Helmholtz demostra que en qualsevol punt d'escolta dins del volum V lliure de fonts la pressió sonora $P(\mathbf{r}, \omega)$ pot ser calculada si es coneix la pressió sonora i el gradient de la superfície que envolta el volum. Açò pot utilitzar-se per sintetitzar un camp d'ones dins de la superfície S si es selecciona la distribució de pressió apropiada $P(\mathbf{r}_s, \omega)$. Aquest fet és utilitzat per a la reproducció basada en WFS. Si la superfície S degenera en un pla $z = z_1$, que separa l'àrea d'escolta de l'àrea de fonts primàries, (Figura 3.3a), aleshores l'equació (3.3) es pot escriure segons la integral de Rayleigh II[4]:

$$P(\mathbf{r}, \omega) = |z - z_1| \int_{S_1} P(\mathbf{r}_s, \omega) \frac{1 + jk|\mathbf{r} - \mathbf{r}_s|}{2\pi|\mathbf{r} - \mathbf{r}_s|^3} e^{-j|\mathbf{r}-\mathbf{r}_s|} dS_1. \quad (3.4)$$

En la Figura 3.3b, es mostra una geometria orientada a un auditori on les superfícies S_2 , S_3 i S_4 són completament absorbents. El camp d'ones de l'àrea d'escolta pot ser generat per una

distribució de fonts secundàries en z_1 , cada font secundària representa un dipol, la font sonora del qual ve donada per la pressió sonora primària en la seua localització (Figure 3.3c).

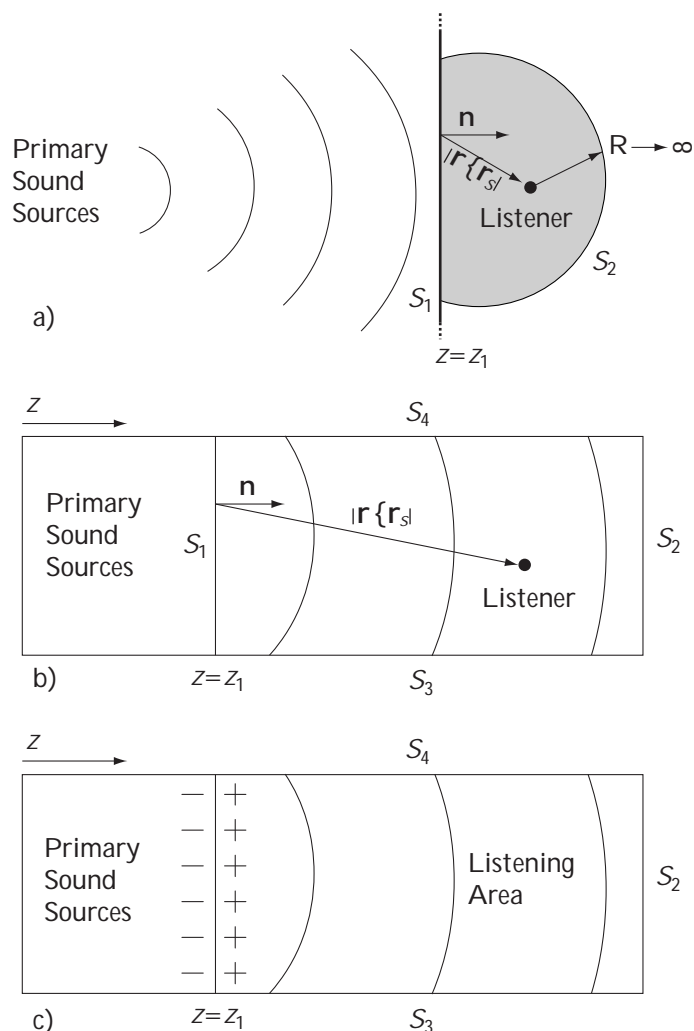


Figura 3.3: a) Simplificació de la integral de Kirchhoff-Helmholtz per a mig camp, b) Geometria orientada a un auditori, c) Distribució secundària de fonts dipol.

A causa d'açò, és possible sintetitzar físicament els fronts d'ones a qualsevol punt d'escolta si es rradia la pressió sonora, enregistrada per micròfons en $z = z_1$, amb altaveus que tinguen característiques de dipol, com s'indica en la Figura 3.4a.

L'agrupació de transductors que són utilitzats per enregistrar i sintetitzar els fronts d'ones poden també ser situats en plans amb diferents coordenades z_0 i z_1 , respectivament, com es veu a la Figura 3.4b. Després, a través de l'equació (3.4), els senyals dels micròfons podrien ser transferits als altaveus a través d'un processador i es simularia la propagació del front d'ona des de z_0 a z_1 numèricament. Aquest procés es coneix com *extrapolació*.

Ara, les posicions dels altaveus \mathbf{r}_n en $z = z_1$ actuen com a posicions d'escolta virtuals i a

causa d'açò el senyal dirigit a cada altaveu en $z = z_1$, per exemple, $P(\mathbf{r}_n, \omega)$, és calculat amb el processament dels senyals de pressió $P(\mathbf{r}_l, \omega)$ enregistrades per tots els micròfons en $z = z_0$ d'acord amb la integral de Rayleigh II.

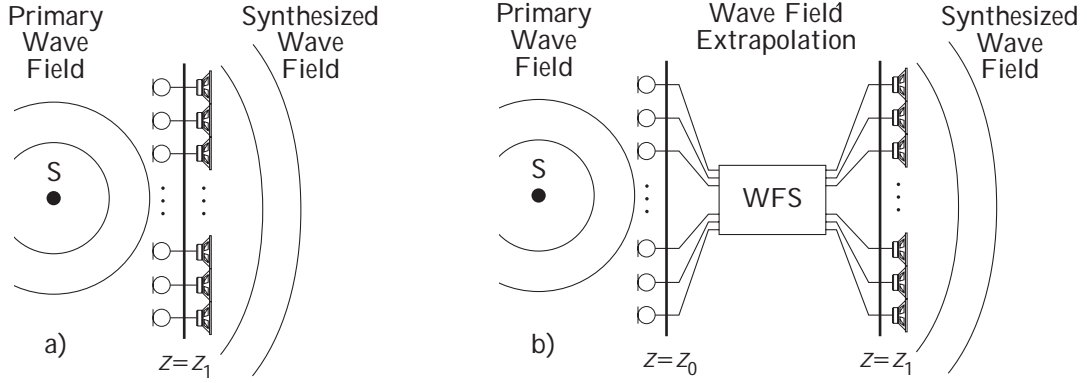


Figura 3.4: a) Representació pràctica de *wave field synthesis* d'acord amb (3.3), b) Generalització del diagrama en a): extrapolació del camp d'ones anterior a l'emissió del camp.

Combinació de la primera integral de Rayleigh

La primera integral de Rayleigh (*Rayleigh I*) indica que el camp d'ones en la meitat de l'espai d'escolta pot ser reconstruït a partir del camp sonor original si es mesura sols la velocitat particular del pla d'interès i s'utilitzen aquestes mesures com a senyals font per a una distribució de monopols en el pla de reproducció[7]:

$$P(\mathbf{r}, \omega) = \frac{jk\rho c}{2\pi} \iint_S u_n(\mathbf{r}_s, \omega) \frac{e^{-j|\mathbf{r}-\mathbf{r}_s|}}{|\mathbf{r}-\mathbf{r}_s|} dx dy, \quad (3.5)$$

o en la seua forma discreta:

$$P(\mathbf{r}, \omega) = \frac{jk\rho c}{2\pi} \sum_n u_n(\mathbf{r}_n, \omega) \frac{e^{-j|\mathbf{r}-\mathbf{r}_n|}}{|\mathbf{r}-\mathbf{r}_n|} \Delta x \Delta y, \quad (3.6)$$

on l'índex n indica els punts de mostreig en el pla S . Les fonts secundàries poden ser construïdes amb petits altaveus monopols on la velocitat del volum és $U_n(\omega) = u_n(\mathbf{r}_n, \omega) \Delta x \Delta y$.

S'ha de tindre en compte que per a altaveus dinàmics, sobre la seua freqüència de ressonància, la velocitat del volum està relacionada amb el voltatge d'entrada $E_n(\omega)$ en el moviment de la seua bobina per $U_n(\omega) = K_m E_n(\omega) / j\omega$, on K_m és una constant que depèn de les propietats electro-mecàniques dels sistemes d'altaveus. Finalment es pot deduir:

$$E_n(\omega) = \frac{j\omega}{K_m} \Delta x \Delta y u_n(\mathbf{r}_n, \omega), \quad (3.7)$$

Aquest resultat mostra que les velocitats particulars en el pla de registre i reconstrucció han de ser diferenciades (filtrades amb +6 dB/oct) abans de ser aplicades als altaveus.

Combinació de la segona integral de Rayleigh

La segon integral de Rayleigh (*Rayleigh II*) dóna una relació similar entre la pressió sonora en el pla mesurat i la distribució de dipols en el pla de reproducció:

$$P(\mathbf{r}, \omega) = \frac{jk}{2\pi} \iint_S P(\mathbf{r}_s, \omega) \frac{1 + jk|\mathbf{r} - \mathbf{r}_s|}{jk|\mathbf{r} - \mathbf{r}_s|} \cos \phi \frac{e^{-jk|\mathbf{r} - \mathbf{r}_s|}}{|\mathbf{r} - \mathbf{r}_s|} dx dy, \quad (3.8)$$

on $\cos \phi$ ve definit com $|z - z_s|/|\mathbf{r} - \mathbf{r}_s|$, sent $|z - z_s|$ la distància entre el pla de reproducció i el punt d'observació. La forma discreta ve donada per:

$$P(\mathbf{r}, \omega) = \frac{jk}{2\pi} \sum_n P(\mathbf{r}_n, \omega) \frac{1 + jk|\mathbf{r} - \mathbf{r}_n|}{jk|\mathbf{r} - \mathbf{r}_n|} \cos \phi_n \frac{e^{-jk|\mathbf{r} - \mathbf{r}_n|}}{|\mathbf{r} - \mathbf{r}_n|} \Delta x \Delta y. \quad (3.9)$$

En aquest cas, les fonts secundàries poden ser construïdes amb petits altaveus sense caixa la força del volum dels quals és:

$$F_n(\omega) = P(\mathbf{r}_n, \omega) \Delta x \Delta y. \quad (3.10)$$

La relació entre la força del volum i el voltatge d'entrada sobre la freqüència de ressonància mecànica ve donada per:

$$F_n(\omega) = K_d E_n(\omega), \quad (3.11)$$

on K_d és una constant electro-mecànica dels sistemes d'altaveus dipol. A partir de les últimes dues equacions,

$$E_n(\omega) = \frac{1}{K_d} \Delta x \Delta y P(\mathbf{r}_n, \omega). \quad (3.12)$$

Cal observar què, respecte dels resultats de Rayleigh I, en aquest cas no es necessita freqüència de pes.

3.2 Derivació de la funció del senyal director

Per a la derivació del senyal dirigit a cada altaveu, es considera la geometria mostrada a la Figura 3.5. El camp de pressió de la font virtual en \mathbf{r}_m en el pla $z = z_0$ podria ser reconstruït en el pla d'escolta horitzontal d'un subjecte situat en \mathbf{r} en el pla $z = z$, si s'utilitza una agrupació lineal d'altaveus paral·lela a l'eix x en el pla $z = z_1$. La línia que connecta la font amb l'oient forma un angle θ amb l'eix z . S'ha de tindre en compte que la font, l'agrupació i l'oient estan localitzats en el pla $y = 0$, és a dir, a la mateixa altura. En la pràctica, l'agrupació d'altaveus es montaran sobre els nivells d'escenari i audiència. En [4], es veu que l'altura de l'agrupació y_1 pot ser despreciada quan és molt menor que les distàncies en horitzontal entre la font i l'agrupació, i entre l'oient i l'agrupació, què és el cas més usual.

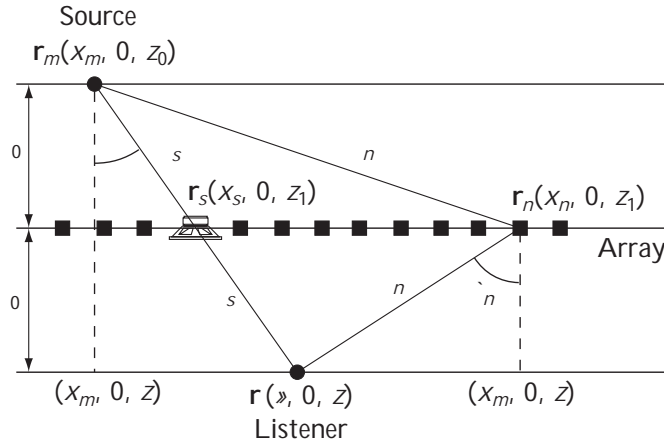


Figura 3.5: Configuració per a *wave field synthesis*. L'agrupació d'altaveus a $z = z_1$ sintetitza el camp d'ones de la font en \mathbf{r}_m en el pla receptor per a $z > z_1$.

D'acord amb el teorema de Rayleigh, els senyals dirigits als altaveus es representen com una versió amb diferents pesos del camp de pressió de la font situada en una determinada posició de l'agrupació.

$$Q(\mathbf{r}_n, \omega) = A_n P(\mathbf{r}_n, \omega) = A_n S(\omega) \frac{e^{-jk|\mathbf{r}_n - \mathbf{r}_m|}}{|\mathbf{r}_n - \mathbf{r}_m|}, \quad (3.13)$$

on A_n és una funció de pesos, que depèn de la posició lateral de l'altaveu n -èsim,

$$A_n = A(x_n, \omega). \quad (3.14)$$

Per sintetitzar un front d'ones esfèric, ha de ser considerada la seua formulació matemàtica:

$$P(\mathbf{r}, \omega) = S(\omega) \frac{e^{-jk|\mathbf{r} - \mathbf{r}_m|}}{|\mathbf{r} - \mathbf{r}_m|}, \quad (3.15)$$

on $S(\omega)$ és l'espectre de la font en qüestió. D'acord amb l'equació de Rayleigh, el front d'ones esfèric es pot sintetitzar com

$$P(\mathbf{r}, \omega) = \sum_{n=1}^N \left[Q(\mathbf{r}_n, \omega) G(\phi_n, \omega) \frac{e^{-jk|\mathbf{r} - \mathbf{r}_n|}}{|\mathbf{r} - \mathbf{r}_n|} \right] \Delta x, \quad (3.16)$$

on N és la quantitat d'altaveus en l'agrupació, $Q(\mathbf{r}_n, \omega)$ és el senyal conduït a l'altaveu n -èsim, ϕ_n és l'angle entre l'eix principal de l'altaveu n -èsim i la seua connexió amb la posició de l'oient, i Δx és l'interval espacial entre l'agrupació d'elements[9]. S'ha d'observar que els únics elements desconeguts en l'operació de síntesi són els senyals conduïts als altaveus.

L'última expressió és utilitzada amb la següent notació abreujada

$$r = |\mathbf{r} - \mathbf{r}_m| = \sqrt{(\xi - x_m)^2 + (z - z_0)^2}, \quad (3.17)$$

$$\rho_n = |\mathbf{r}_n - \mathbf{r}_m| = \sqrt{(x_n - x_m)^2 + (z_1 - z_0)^2}, \quad (3.18)$$

$$\sigma_n = |\mathbf{r} - \mathbf{r}_n| = \sqrt{(\xi - x_n)^2 + (z - z_1)^2}, \quad (3.19)$$

$$(3.20)$$

que condueix a

$$\frac{e^{-jkr}}{r} = \sum_{n=1}^N \left[A_n G(\phi_n, \omega) \frac{e^{-jk(\rho_n + \sigma_n)}}{\rho_n \sigma_n} \right] \Delta x. \quad (3.21)$$

Aquesta equació integral discretitzada pot ser aproximada si s'utilitza la seua representació de fase estacionària[6]:

$$\frac{e^{-jkr}}{r} \simeq \sqrt{\frac{j2\pi}{\alpha''(x_s)}} f(x_s) e^{j\alpha(x_s)}. \quad (3.22)$$

on $\alpha(x_n) = -k(\rho_n + \sigma_n)$, $f(x_n) = A_n G(\phi_n, \omega) / \rho_n \sigma_n$ i la prima significa diferenciació amb x_n . Per altra banda, x_s és la coordenada lateral del punt de fase estacionària, és a dir, el punt on $\alpha'(x_n)=0$, per al qual segueix com

$$x_s = x_m + \frac{\rho_0}{\rho_0 + \sigma_0} (\xi - x_m), \quad (3.23)$$

on $\rho_0 = |z_1 - z_0|$ i $\sigma_0 = |z - z_1|$. La comparació de l'última equació amb la geometria de la Figura 3.5 mostra que x_s és la coordenada lateral del punt d'intersecció del vector $\mathbf{r} - \mathbf{r}_m$ amb el pla $z = z_1$, per exemple, amb l'agrupació d'altaveus. Per tant, $r = \rho_s + \sigma_s$, $-kr = \alpha(x_s)$ i $\theta = \phi_s$.

L'aproximació de fase estacionària físicament significa que el front d'ones està sintetitzat per tots els altaveus de l'agrupació junts, però hi haurà una contribució dominant de l'altaveu posicionat en el punt de fase estacionària.

Després d'una substancial manipulació matemàtica, $\alpha''(x_s)$ ve donada per[9]:

$$\alpha''(x_s) = -k \left(\frac{\rho_0 + \sigma_0}{\rho_0 \sigma_0} \right) \left[1 + \frac{(\xi - x_m)^2}{(\rho_0 + \sigma_0)^2} \right]^{-3/2} \quad (3.24)$$

i és possible escriure $f(x_s) = A_s G(\theta, \omega) / \rho_s \sigma_s$.

Si es substitueix açò en l'aproximació de fase estacionària de la integral discretitzada, dóna una expressió per a la funció de pesos A_s :

$$A_s = A \left(x_m + \frac{\rho_0}{\rho_0 + \sigma_0} \right) = \frac{1}{G(\theta, \omega)} \sqrt{\frac{jk}{2\pi}} \rho_0 \sqrt{\frac{\sigma_0}{\rho_0 + \sigma_0}} \frac{1}{\sqrt{\rho_s}}. \quad (3.25)$$

S'ha d'observar que el punt de l'agrupació (x_s, z_1) actua com a punt de fase estacionària de la coordenada lateral de l'oient específica x . Una variació del valor x també variarà el valor de x_s . A causa d'açò cada posició d'altaveu serà de fase estacionària segons la posició de l'oient en el pla $z = z$. Per aquest motiu, l'equació (3.25) pot ser generalitzada de la següent forma:

$$\begin{aligned}
A_n = A(x_n, \omega) &= \frac{1}{G(\theta_n, \omega)} \sqrt{\frac{jk}{2\pi}} \rho_0 \sqrt{\frac{\sigma_0}{\rho_0 + \sigma_0}} \frac{1}{\sqrt{\rho_n}} \\
&= \frac{1}{G(\theta_n, \omega)} \sqrt{\frac{jk}{2\pi}} |z_1 - z_0| \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{1}{\sqrt{|\mathbf{r}_n - \mathbf{r}_m|}},
\end{aligned} \tag{3.26}$$

on θ_n és l'angle entre la línia que connecta la font amb l'altaveu n -èsim i l'eix z . Si es substitueix aquest resultat en (3.13), el senyal conduït $Q(\mathbf{r}_n, \omega)$ a l'altaveu n -èsim pot ser:

$$Q(\mathbf{r}_n, \omega) = S(\omega) \frac{1}{G(\theta_n, \omega)} \sqrt{\frac{jk}{2\pi}} |z_1 - z_0| \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{e^{-jk|\mathbf{r}_n - \mathbf{r}_m|}}{\sqrt{|\mathbf{r}_n - \mathbf{r}_m|^3}}, \tag{3.27}$$

La proporció entre ρ_0 i ρ_n dóna $\cos \theta_n = |z_1 - z_0|/|\mathbf{r}_n - \mathbf{r}_m|$, i, per aquest motiu, el senyal conduït pot ser reformulat com

$$Q(\mathbf{r}_n, \omega) = S(\omega) \frac{\cos \theta_n}{G(\theta_n, \omega)} \sqrt{\frac{jk}{2\pi}} \sqrt{\frac{|z - z_1|}{|z - z_0|}} \frac{e^{-jk|\mathbf{r}_n - \mathbf{r}_m|}}{\sqrt{|\mathbf{r}_n - \mathbf{r}_m|}}, \tag{3.28}$$

Aquesta funció conductora conté un factor de divergència cilíndrica espacial $|\mathbf{r}_n - \mathbf{r}_m|^{-1/2}$ i d'aquesta manera, el senyal conduït a l'altaveu n -èsim pot ser interpretat com una versió amb pesos del camp de pressió sonora en \mathbf{r}_n causada per una font lineal virtual en \mathbf{r}_m . El fet que sols una línia horitzontal en el pla de reconstrucció $z = z_1$, és utilitzat en el procés de síntesi del front d'ones, es compensa per la extensió espacial de la font virtual d'un punt en línia vertical.

L'equació dels senyals directors (3.28) pot ser separada en dues parts, com es mostra:

$$Q(\mathbf{r}_n, \omega) = C(z, z_0, z_1) P_{line}(\mathbf{r}_n, \omega), \tag{3.29}$$

amb

$$C(z, z_0, z_1) = \sqrt{\frac{|z - z_1|}{|z - z_0|}}, \tag{3.30}$$

i

$$P_{line}(\mathbf{r}_n, \omega) = S(\omega) \sqrt{\frac{jk}{2\pi}} \frac{\cos \theta_n}{G(\theta_n, \omega)} \frac{e^{-jk|\mathbf{r}_n - \mathbf{r}_m|}}{\sqrt{|\mathbf{r}_n - \mathbf{r}_m|}}. \tag{3.31}$$

La font virtual té un patró de directivitat que està inversament relacionat amb les característiques de directivitat de l'agrupació d'altaveus,

$$G_{line}(\phi, \omega) = \sqrt{\frac{jk}{2\pi}} \frac{\cos \phi}{G(\phi, \omega)}. \tag{3.32}$$

Açò vol dir que en principi qualsevol tipus d'altaveu pot ser utilitzat per a formar una agrupació per a *wavefield synthesis*. Si s'adapta la funció de senyals directors d'acord amb (3.28) les característiques de directivitat de l'altaveu són virtualment transferides a les fonts virtuals.

3.3 Potencial i restriccions de WFS

El principi característic de WFS és que l'escena acústica roman constant per a l'àrea total d'escolta, és a dir, el muntatge absolut de l'escena acústica és independent de la posició de l'oient. La perspectiva relativa acústica és percebuda pels canvis de l'oient amb moviments d'aquest. Aquesta variació engloba un canvi realístic del nivell de pressió sonora quan la distància a la font virtual varia.

Les aptituds teòriques de WFS per a crear un camp sonor quasi-realístic o per recrear un camp sonor existent són prou grans.

Des d'un punt de vista creatiu, WFS ofereix una millora de flexibilitat: direcció i localització de fonts estables poden ser reproduïdes. El disseny de l'escena acústica està menys limitat a les restriccions de les tècniques de reproducció en comparació a l'estèreo. La simulació d'una escena acústica real és més factible.

Per altra banda, si es té en compte que l'objectiu de WFS és la creació d'una còpia vertadera d'un camp sonor natural, aquest objectiu tan extrem, en la pràctica, sols pot ser aconseguit amb certes restriccions. Algunes d'aquestes restriccions són les següents:

1. *Aliasing espacial*

La distància entre transductors Δx defineix la freqüència de mostreig per al camp d'ones a nivell d'enregistrament. Per aquest motiu el camp reconstruït serà físicament correcte a la freqüència de Nyquist:

$$f_{nyq} < \frac{c}{2\Delta x} \leftrightarrow \Delta x < \frac{\lambda_{\min}}{2}, \quad (3.33)$$

on λ_{\min} és la longitud d'ona mínima d'interès. Sobre la freqüència f_{nyq} , l'efecte d'*aliasing* espacial pot ocórrer, per exemple, després d'una reconstrucció correcta d'un front d'ones, on pot emetre erròniament un front d'ones addicional[4].

En la pràctica la recreació del camp d'ones sense *aliasing* espacial és possible per a freqüències menors a la "Freqüència d'*aliasing* espacial (f_{al})". Aquesta freqüència de tall ve donada per

$$f_{al} = \frac{c}{2\Delta x |\sin \alpha_{\max}|} \quad (3.34)$$

on α_{\max} indica l'angle màxim entre un pla d'ones incident i la superfície de l'agrupació de micròfons.

2. *Efectes de Truncament*

En teoria, la síntesi de camp d'ones sorgeix del sumatori d'una quantitat de senyals provinents d'altaveus. No obstant, en la pràctica l'agrupació d'altaveus utilitzada tindrà sempre dimensions finites. L'agrupació finita pot ser vista com una finestra, a través de la qual la font primària virtual pot ser o no vista per l'oient. Conseqüentment, existeix un àrea que és il·luminada per la font virtual, amb la seua corresponent àrea d'ombra[24]. Si s'aplica aquesta analogia, s'originen ones difractades a causa dels extrems de l'agrupació d'altaveus.

Aquestes contribucions d'error apareixen com a pre-ecos i post-ecos de les fonts focals, i segons el seu nivell i el temps de retardament en el punt de localització del receptor, poden provocar una pujada de la coloració de la font.

3. Elevació

Han sigut moltes les proves per a aconseguir l'efecte d'elevació a WFS, però fins ara no s'ha vist cap resultat satisfactori. No obstant en aquest treball, com es veurà als Capítols 4 i 5, es proposa una nova tècnica per a aconseguir l'efecte desitjat.

3.4 Propietats especials de WFS

Localització de fonts virtuals

A través de WFS l'enginyer de so té una eina poderosa per a dissenyar una escena sonora. Una de les propietats més importants, respecte les tècniques convencionals, és la seua capacitat per a proveir una localització realística de fonts virtuals. Problemes i limitacions típiques de la imatge estereofònica desapareixen en una escena sonora WFS.

En contrast amb els sistemes estereofònics WFS està capacitat per a:

- produir fonts virtuals que estan localitzades a la mateixa posició a tota l'àrea d'escolta. En la Figura 3.6, les fletxes indiquen les direccions dels esdeveniments de l'auditori quan fonts virtuals puntuals són reproduïdes.
- produir ones planes que estan localitzades a la mateixa direcció a través de l'àrea d'escolta sancera.
- millorar la localització de fonts virtuals i el sentiment de presència i entorn a través d'una reproducció realística de la distribució d'amplituds de la font virtual. En altres paraules, quan l'oient s'aproxima a la localització de la font virtual l'amplitud s'incrementa com si fos el cas real. D'acord amb açò, l'amplitud de l'ona plana, la qual pot ser vista com una font a distància infinita, varia molt poc en diferents posicions de l'oient.

Aquestes propietats fan possible la síntesi d'escenes sonores complexes que poden ser una experiència per a l'oient mentre es maneja dins de l'àrea d'escolta. La Figura 3.6 mostra com varia la imatge sonora segons la posició de l'oient.

A més a més, com s'ha vist, la resolució de la localització comparada amb l'estereofonia fa que l'oient distingisca més fàcilment entre les diferents fonts virtuals, que fan l'escena sonora significativament més transparent.

Fonts sonores virtuals front a l'agrupació d'altaveus (Fonts Puntuals)

La Figura 3.7a) i b) mostren els fronts d'ones d'una font puntual darrere i davant de l'agrupació respectivament, segons una simulació. Els fronts d'ona còncaus de la Figura 3.7a) aconseguixen la síntesi del senyal d'una font virtual darrere de l'agrupació. No obstant, WFS és també capaç

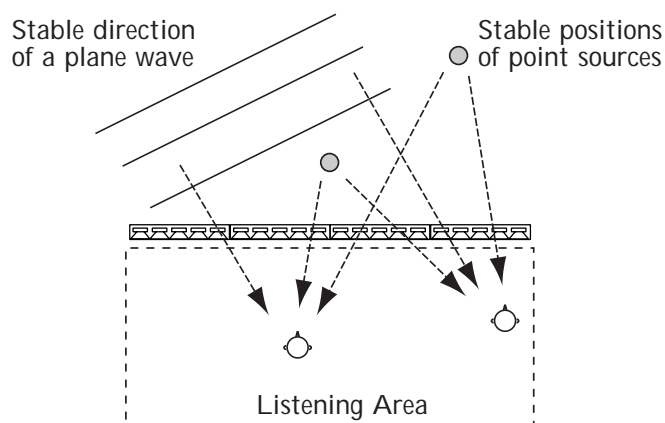


Figura 3.6: WFS és capaç de reproduir posicions estables de les fonts puntuals i les direccions estables d'ones planes.

de sintetitzar una font virtual davant de l'agrupació. A conseqüència d'açò l'agrupació WFS emet fronts d'ona convexos el focus dels quals pot estar localitzat a la posició de la "font focal" (Figura 3.7b)). Naturalment, la localització pot no ser correcta per a posicions d'escolta entre el punt focal i l'agrupació perquè l'emissió sonora de la font virtual ocorre en aquests punts de manera contrària.

Per a l'aplicació pràctica hi ha un enorme progrès en què les fonts sonores virtuals en el camp entre l'oient i els altaveus pot ser recreat. Els enginyers de so poden disposar completament de noves eines per al disseny de so espacial.

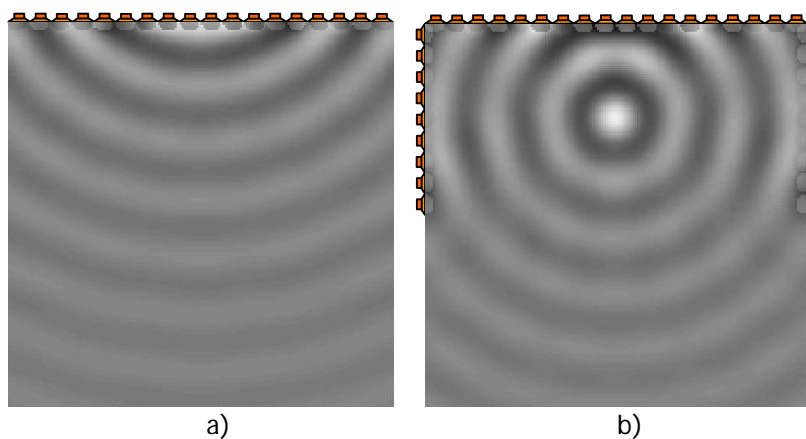


Figura 3.7: Fronts d'ones de les fonts virtuals, a) darrere de l'agrupació, b) en front de l'agrupació.

S'ha de veure que aquestes propietats són sols vàlides per a una interpretació de la localització respecte la direcció de la font virtual. En el que respecta a la distància de percepció real, les propietats particulars de WFS que s'han descrit, no són suficients. Aquest és un punt sota investigació.

Després d'aquesta revisió de teoria acústica, es presenten dos enfocaments de processament: basat en models i basat en dades. Ambdós enfocaments segueixen unes especificacions apropiades

de la versió discreta de la integral de Rayleigh I.

3.5 Implementació pràctica de WFS

WFS s'ajusta al concepte de MPEG-4 [15] que pot ser molt atractiu per a produccions audiovisuals, on es dona als enginyers de so un nou nivell de llibertat en l'àudio espacial. Amb l'ús d'aquesta tècnica composta, es pot crear el camp sonor mitjançant mètodes basats en un model o en bases de dades, que són també comuns en els sistemes de reproducció d'àudio espacial relacionats amb el cap.

Wave Field Synthesis basat en models

En l'enfocament de WFS basat en models, les dades del camp d'ones es divideixen en tres parts diferents: el so directe en l'entorn lliure de reverberacions, les primeres reflexions (< 100 ms) i la reverberació o entorn real [15]. Els altaveus individuals de l'agrupació seran alimentats per senyals que contenen aquesta informació.

Si les dues primeres components poden ser separades per a cada altaveu, les ones planes són generades per grups d'altaveus.

Wave Field Synthesis basat en dades

En l'enfocament basat en dades la creació del camp acústic es basa també en senyals separats que representen fonts sonores discretes. No obstant els senyals conductors a cada altaveu individual són obtinguts amb la convolució de senyals secs i respostes a l'impuls que representen l'entorn acústic. Aquestes respostes a l'impuls són pre-mesurades i emmagatzemades en bases de dades.

En aquest treball, com es veu al Capítol 4, s'aportarà l'efecte vertical als sistemes WFS. Per a dur-ho a terme es basarem tant en dades, respostes a l'impuls pre-mesurades, com en models amb l'aplicació de reverberacions i reflexions per fer més real el so.

Capítol 4

Disseny de Filtres HRTF en elevació

En aquest capítol s'utilitzen diverses tècniques per al disseny d'un filtre que represente l'HRTF en elevació generalitzada per a una quantitat elevada de subjectes. Per a desenvolupar aquesta síntesi caldrà avaluar l'HRIR (*Head Related Impuls Response*) de diferents subjectes.

4.1 Anàlisi de dades

Es recorre a la base de dades de IRCAM [16] per a descarregar la informació sobre l'HRIR en 3D de 48 subjectes, que està inclosa en *The Listen project* desenvolupat per AKG i Ircam. Les mesures de l'HRIR han sigut fetes en una cambra anecoica de dimensions $(8.1 \times 6.2 \times 6.45m^3)$ que absorbeix ones sonores sobre 75 Hz. L'altaveu es mou a través d'una estructura metàl·lica en forma d'U. L'elevació s'aconsegueix amb un parell de motors pas a pas que són controlats a través de l'ordinador. El subjecte mesurat se'u sobre una cadira rotatòria fixada sobre una taula giratòria, model B&K 9640, i connectada a un controlador model B&K 5960. El cap del subjecte es recolza per a ajudar a que no es meneje. A més es col·loca un receptor dalt del cap per a validar la posició d'aquest.

Per mesurar l'HRIR s'han utilitzat un parell de micròfons, la part frontal dels quals s'ha col·locat a l'entrada del canal auditiu, de forma que s'aconsegueix l'oclusió del canal auditiu per previndre la ressonància. La càpsula s'inserta en un motlle personalitzat de silicona, per a aconseguir sempre la mateixa posició del micròfon. S'inclourà també a la base de dades, l'HRIR del maniquí mesurat per KEMAR [13].

A partir de les mesures obtingudes a la base de dades de Ircam, s'aconsegueix l'HRTF (*Head Related Transfer Function*) de cada subjecte per a una freqüència de mostreig de 44100 Hz, angle 0° en azimuth i 10 angles d'elevació diferents. A la Figura 4.1 es pot veure l'HRTF obtinguda per a un subjecte a l'atzar de la base de dades de Ircam.

Una vegada obtingudes les HRTF dels 48 subjectes es fa una mitja de totes les respostes, per a cada freqüència f i cada angle d'elevació φ , segons l'expressió 4.1 i s'obté l'HRTF d'interès que més tard es vol sintetitzar.

$$HRTF_{total}(f, \varphi) = \frac{\sum_{i=1}^N HRTF_i(f, \varphi)}{N} \quad (4.1)$$

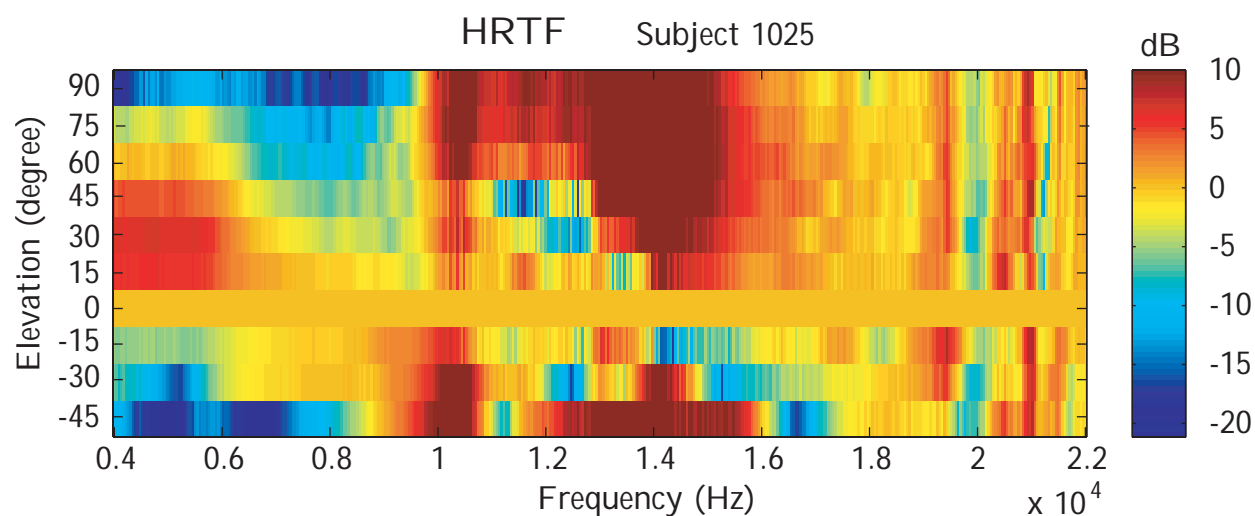


Figura 4.1: HRTF per a un subjecte amb angle azimuth 0° i diferents angles en elevació

A la Figura 4.2 es pot veure l'HRTF mitjana obtinguda.

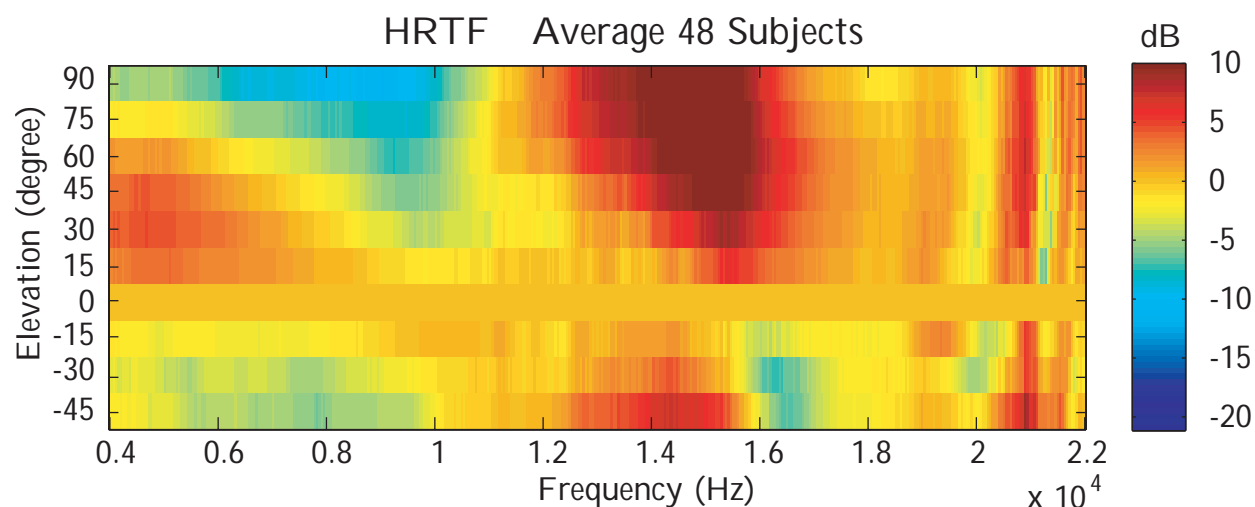


Figura 4.2: Mitja de l'HRTF per a 48 subjectes amb angle azimuth 0° i diferents angles en elevació

4.2 Síntesi de dades

Es realitzarà la síntesi de l'HRTF en elevació a partir de la combinació de múltiples filtres *peak*, que segueixen diferents funcions de primer i segon ordre, com poden ser rectes o paràboles.

Els filtres *peak* junt als filtres *notch*, els filtres de *Shelving* i els filtres passa tot, són filtres de segon ordre molt utilitzats per a aplicacions d'àudio que procedeixen de models analògics. Aquests

filtres es defineixen per una sèrie de paràmetres, com són el guany G , la freqüència central f_c i el factor de qualitat Q . Per aquest motiu són coneguts també com filtres paramètrics. Els filtres *peak* són utilitzats als equalitzadors paramètrics i realcen una banda molt estreta. A la Figura 4.3 es pot veure la representació d'un filtre *peak*.

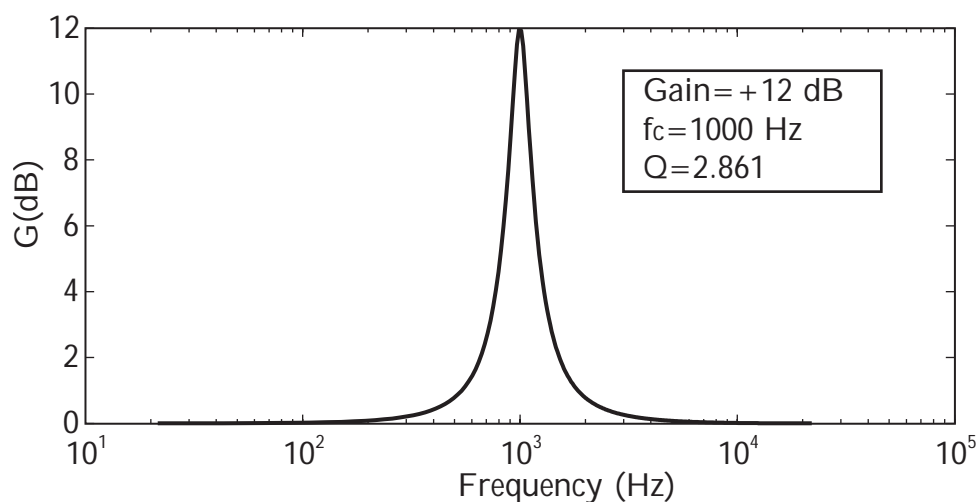


Figura 4.3: Filtre *peak* per a uns paràmetres determinats

El seu model analògic ve determinat per l'equació 4.2:

$$H(s) = \frac{s^2 + \frac{A}{Q}s + 1}{s^2 + \frac{1}{AQ}s + 1} \quad (4.2)$$

on A és l'arrel quadrada del guany.

El seu model digital ve definit per l'equació 4.3:

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{a_0 + a_1 z^{-1} + a_2 z^{-2}} \quad (4.3)$$

on els coeficients $a_0, a_1, a_2, b_0, b_1, b_2$, es defineixen a partir dels paràmetres G, f_c i Q de la següent forma:

$$a_0 = 1 + \frac{\alpha}{A} \quad (4.4)$$

$$a_1 = \frac{-2 \cos \omega}{a_0} \quad (4.5)$$

$$a_2 = \frac{1 - \frac{\alpha}{A}}{a_0} \quad (4.6)$$

$$b_0 = \frac{1 + A\alpha}{a_0} \quad (4.7)$$

$$b_1 = \frac{-2 \cos \omega}{a_0} \quad (4.8)$$

$$b_2 = \frac{1 - A\alpha}{a_0} \quad (4.9)$$

on ω , A i α estan definides segons:

$$\omega = \frac{2\pi f_c}{f_s} \quad (4.10)$$

$$A = \sqrt{G} \quad (4.11)$$

$$\alpha = \frac{\sin \omega}{2Q} \quad (4.12)$$

on f_s es refereix a la freqüència de mostreig utilitzada.

4.2.1 Síntesi I: consideració dels angles d'elevació junts

Per a la síntesi de l'HRTF en elevació on es consideren tots els angles al mateix temps, a la Figura 4.2, obtinguda com a mitja de les HRTF dels subjectes, es cerquen diferents funcions característiques. Després s'aplica un filtre *peak* per a cadascuna d'aquestes funcions, i mitjançant la correlació dels coeficients del filtre, s'obté un filtre final que representa l'HRTF de la resposta de l'oïda en elevació.

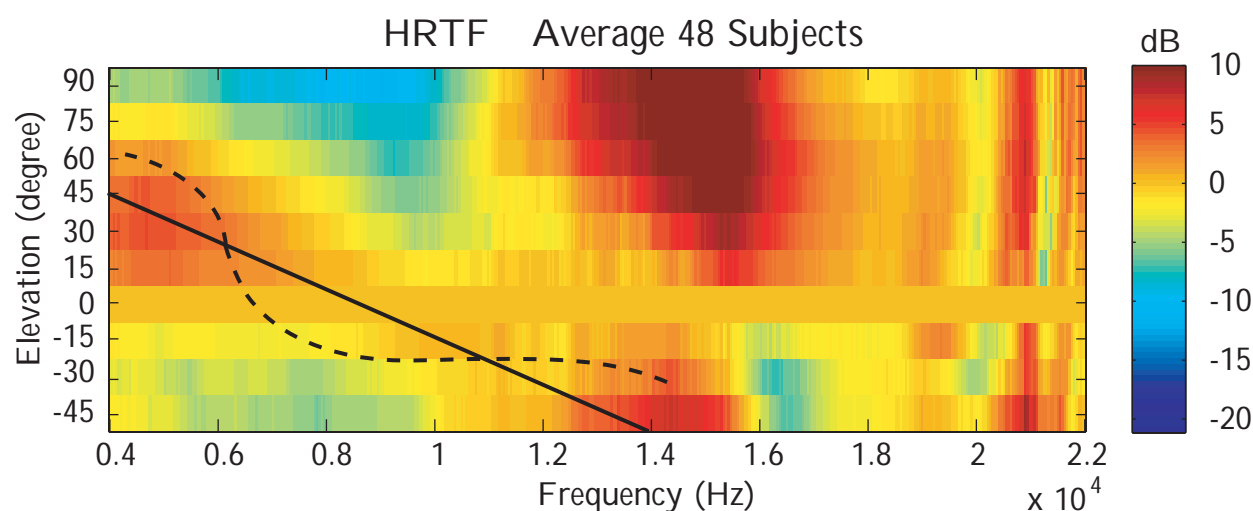


Figura 4.4: Extracció de funcions a la representació de l'HRTF

Per exemple, a la Figura 4.4 es pot observar com hi ha una recta de pendent negativa (veure línia contínua), l'equació de la qual es pot obtenir a partir de dos punts significatius, com poden

ser $P_1 = (0.4 \times 10^4, 45^\circ)$ i $P_2 = (1.4 \times 10^4, -45^\circ)$. També es pot observar que als extrems de la recta l'amplitud és màxima i s'anula al centre, de forma que segueix l'equació d'un cosinus (veure línia discontinua). Es calcula la freqüència central f_c i el guany G amb la consideració d'aquestes dues suposicions i s'aplica la funció filtre *peak* amb la f_c i la G obtingudes i un factor de qualitat segons s'ajuste més al resultat desitjat. En aquest cas $Q = 1$, f_c segueix la funció de la recta i G segueix la funció del cosinus.

De la mateixa forma es procedeix amb la resta de funcions visibles, es fa la convolució dels coeficients del filtre i s'aconsegueix, com es veu a la Figura 4.5 el filtre final representatiu de l'HRTF.

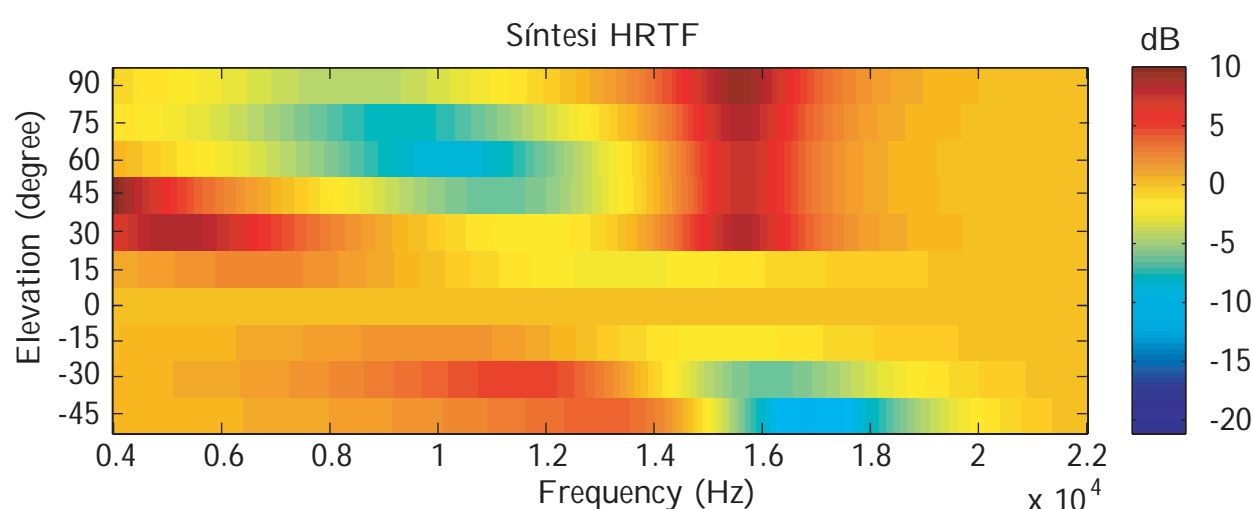


Figura 4.5: Síntesi de l'HRTF obtinguda amb la consideració de tots els angles d'elevació al mateix temps

4.2.2 Síntesi II: consideració dels angles d'elevació per separat

En aquesta segona opció per fer la síntesi, s'avalua cada angle en elevació per separat. S'apliquen tècniques d'optimització (veure Annex A) per a aplicar de la millor forma els filtres *peak* a cada funció en freqüència segons l'angle en elevació.

A la Figura 4.6 es representa en color blau la funció real en freqüència per a $+15^\circ$ en elevació. La funció en color verd és l'obtinguda mitjançant filtres *peak* i tècniques d'optimització.

De la mateixa forma que per a l'angle de $+15^\circ$ es procedeix a obtenir les funcions de la resta d'angles en elevació, fins completar tot el filtre HRTF de l'oïda al pla vertical. Finalment, es pot veure a la Figura 4.7 la resposta total de l'HRTF en freqüència obtinguda.

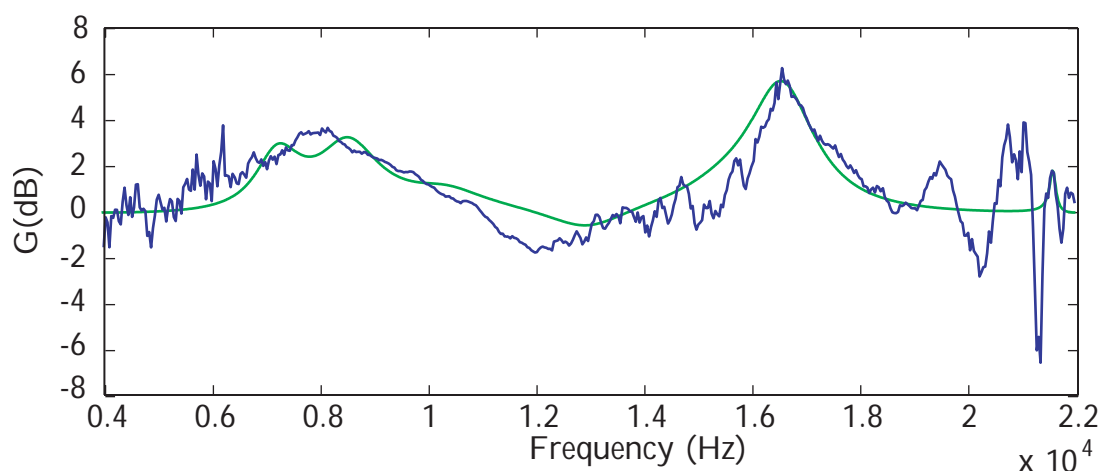


Figura 4.6: Síntesi de l'HRTF per a un sol angle d'elevació (15°) obtinguda amb tècniques d'optimització

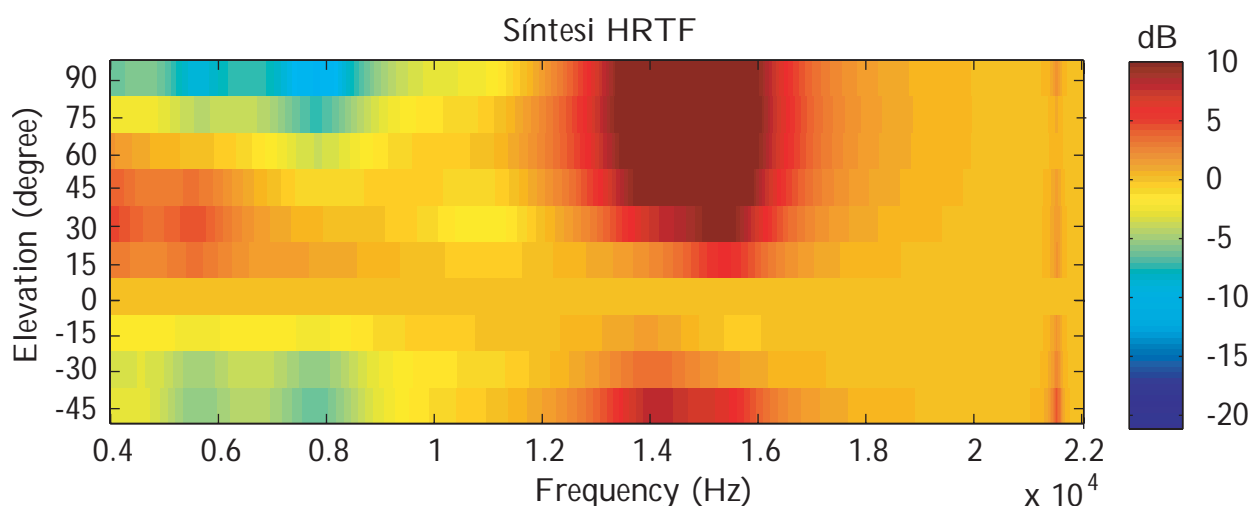


Figura 4.7: Síntesi de l'HRTF obtinguda amb els angles d'elevació per separat

4.3 Exaltació de filtres

Com més tard es veurà als resultats del Capítol 5, els efectes obtinguts amb els filtres dissenyats són un poc pobres. A causa d'açò, es decideix aplicar una funció d'exaltació tant per als filtres dissenyats pel mètode de síntesi I com per als filtres dissenyats pel mètode de síntesi II. La funció d'exaltació és una transformació del guany (g) dels filtres *peak* que conformen el filtre complet, basada en filtrat *Warped* segons la següent expressió:

$$T(g) = \text{sign}(g) \frac{|g| + \lambda|g|}{2|g| + \lambda - 1} \quad (4.13)$$

on λ marca el pendent del filtre. A menor λ , major pendent té la transformació, és a dir, com es mostra a la Figura 4.8 a), els nivells positius pròxims a zero tendeixen a assolir un nivell alt molt ràpidament i, els nivells negatius pròxims a zero, tendeixen a assolir un nivell molt més negatiu.

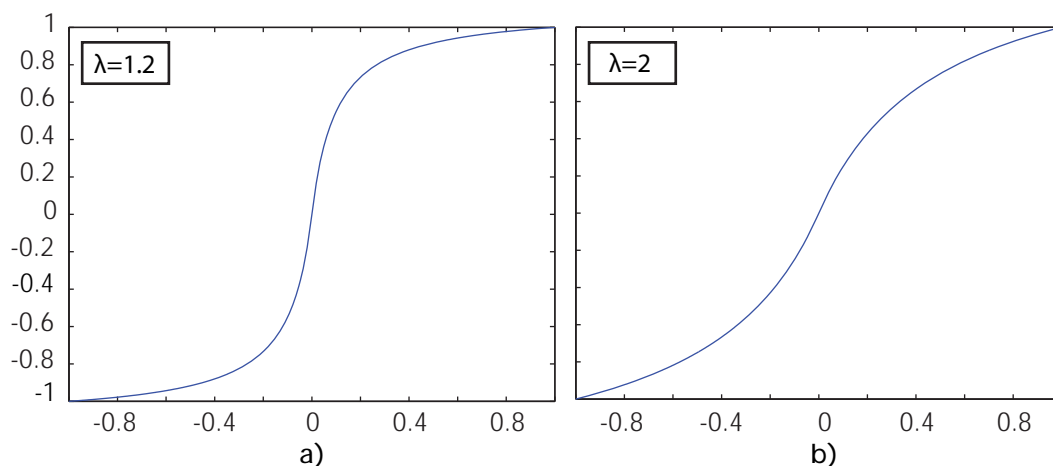


Figura 4.8: Transformació del guany aplicada als filtres per a un major ressalt de l'elevació

A la Figura 4.9 es veu el filtre obtingut al aplicar la funció d'exaltació al mètode de síntesi I. Es pot distingir com les zones positives augmenten el seu nivell (van més cap al roig) i les zones negatives el minven (es fan més blaves).

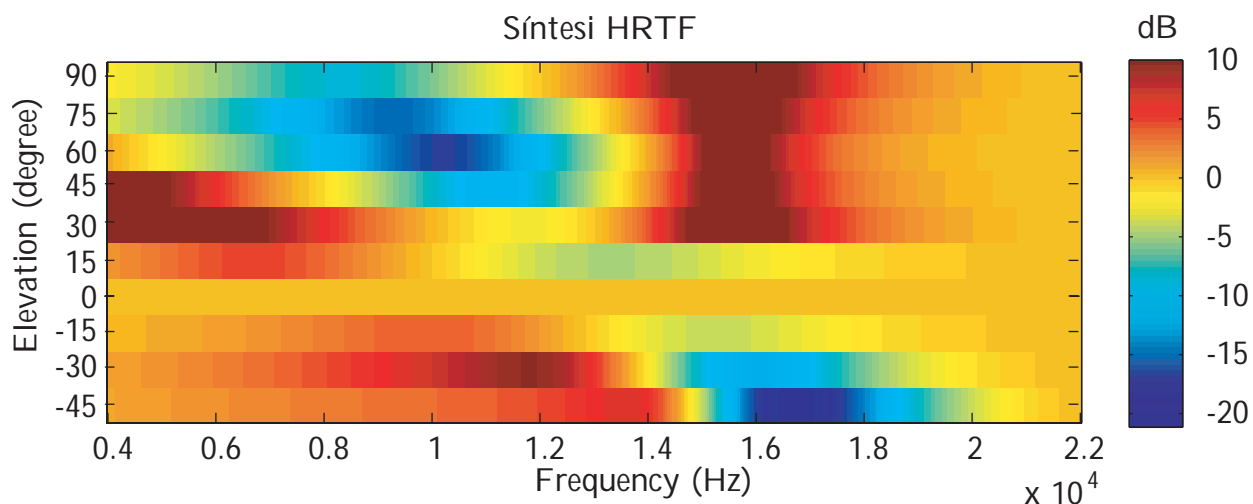


Figura 4.9: HRTF amb exageració per al mètode de Síntesi I

De la mateixa forma, a la Figura 4.10 es pot veure el filtre obtingut per al cas del mètode de síntesi II.

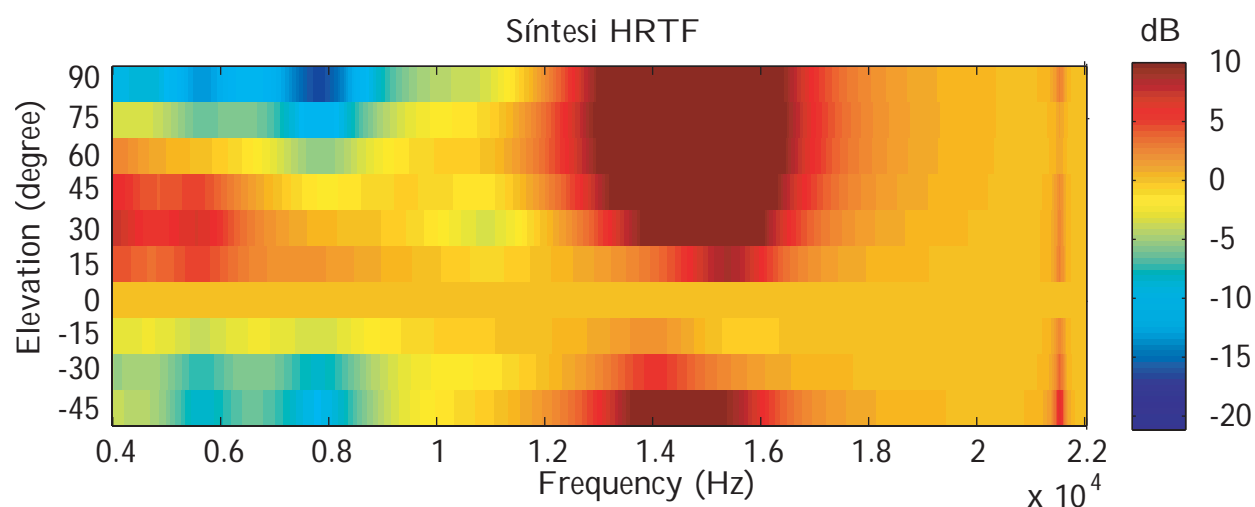


Figura 4.10: HRTF amb exageració per al mètode de Síntesi II

S'ha d'eleger una λ que no sature molt el senyal, perquè distorsionaria el resultat i no s'obtidria l'efecte desitjat.

4.4 Filtrat RSS

Per a poder comparar els filtres abans dissenyats, s'ha optat per crear un nou filtre a partir de les respostes a l'impuls mesurades, mitjançant tècniques MLS [17], del procesador de so espacial RSS-10 *Roland Sound Space*.

El procesador RSS-10 utilitza la tecnologia de procesat de so tridimensional de Roland [22]. Amb l'ús d'un sistema convencional de 2 canals, el so pot ser situat o mogut per qualsevol punt del camp tridimensional. Al Capítol 5 es veurà l'avaluació sobre aquest filtre feta per diversos subjectes.

Capítol 5

Aplicacions de l'HRTF a tècniques WFS

Són nombrosos els intents per dur a terme l'efecte de l'elevació als sistemes basats en WFS. Entre aquests es pensava que la combinació de diverses agrupacions d'altaveus en paral·lel podia donar un bon resultat. No obstant al posar la idea en pràctica no es va aconseguir l'efecte desitjat. Molts articles nomenen que l'única forma d'aconseguir so en elevació és a partir d'holofonia en sistemes binaurals i estereofònics.

En aquest capítol es posa en pràctica la nova solució plantejada al Capítol 4 per a poder aplicar variacions en elevació als sistemes WFS. Aquesta solució es basa en aplicar els filtres HRTF dissenyats als senyals que s'envien a una única agrupació d'altaveus horitzontal. La qualitat percebuda com a resultat és avaluada.

5.1 Avaluació subjectiva

A continuació es proposen diferents criteris per a avaluar el funcionament dels diferents filtres dissenyats al Capítol 4. Per a l'avaluació es fa ús de la sala d'àudio del GTAC (*Grupo de Tratamiento de Audio y Comunicaciones*) de la Universitat Politècnica de València.

L'agrupació d'altaveus utilitzada als tests subjectius es pot veure a la Figura 5.1

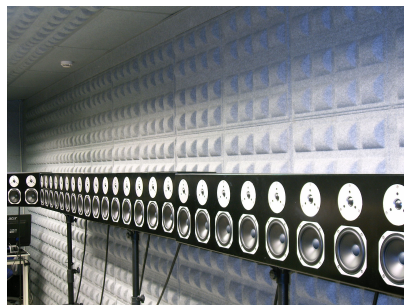


Figura 5.1: Agrupació WFS en el grup d'investigació GTAC.

L'avaluació subjectiva s'ha fet a través de diversos tests aplicats a un conjunt de 12 persones

entre les quals hi ha gent que treballa al camp de l'acústica i hi ha gent totalment aliena a aquest camp d'investigació. Els subjectes es col·loquen front l'agrupació WFS i a continuació es presenten diferents escenes de forma successiva.

- La primera avaluació consisteix en reproduir tres sons curts amb els tres filtres dissenyats. Cada so es reproduïx de 0° a 90° a intervals de 15° . Els subjectes han d'avaluar si els filtres han aconseguit el seu propòsit i si s'aprecia pas a pas com canvia l'angle en elevació a la reproducció. Els tres sons curts que es reproduïxen són:
 - Primera prova: *ding.wav* so agut utilitzat per expressar error en alguns sistemes operatius.
 - Segona prova: *noise.wav* soroll blanc el qual conté tot l'espectre de freqüències.
 - Tercera prova: *bat.wav* colp de caixa d'una bateria.
- La segona avaluació consisteix en reproduir dos sons iguals sintetitzats amb els filtres dissenyats, amb dos angles diferents en elevació. Els subjectes han de reconèixer si el pas d'un so a l'altre és ascendent (primer so amb menor angle d'elevació que el segon) o descendent (primer so amb major angle d'elevació que el segon so).
- La tercera avaluació consisteix en sintetitzar un so de llarga duració de forma dinàmica amb diferents angles en elevació. El subjecte ha de dibuixar la sensació que percep, és a dir, la trajectòria del so en funció del temps.

La puntuació permitida per a la primera avaluació és: excel·lent (5), bé (4), regular (3), pobre (2), roïn (1). A més a més es demana als subjectes avaluats que facen qualsevol comentari positiu o negatiu sobre els filtres, de manera que es justifique la seua puntuació.

5.2 Resultats

A continuació es mostren en diferents taules els resultats obtinguts de les tres proves d'avaluació dels filtres dissenyats en aquest treball. En la Taula 5.1 es mostren les mitges de les puntuacions donades per cada subjecte a cada filtre segons el senyal reproduït. Es pot apreciar que cap disseny

	Primera prova	Segona prova	Tercera prova
Filtre per Síntesi I	2,42	1,83	2,58
Filtre per Síntesi II	2,50	2,42	2,83
Filtre RSS	2,33	3,08	3,00

Taula 5.1: Avaluació subjectiva dels filtres dissenyats. Rang de 1 (resultat roïn) a 5 (resultat excel·lent).

funciona a la perfecció. El filtre dissenyat pel mètode de Síntesi I (segons els comentaris fets pels subjectes) funciona bé entre 0° i 60° , i per als últims dos angles d'elevació avaluats no es percep

l'efecte. El filtre dissenyat pel mètode de Síntesi II al començament no dona sensació d'elevant-se, però a partir dels 30 graus es percep l'efecte. No obstant, per a algun senyal reproduït, es distorsiona molt el senyal conforme augmenta l'angle d'elevació, ja que s'atenuen molt les freqüències baixes i l'efecte produït a l'oïda no és el d'un so comú. Finalment els comentaris fets sobre el filtre RSS són, en general, que s'aprecia perfectament l'elevació al començ, però als últims àngles d'elevació el resultat s'estanca i sembla que és constant i no varia amb l'elevació.

Seguidament, a la Taula 5.2 es veu el resultat de la segona avaluació. En la primera columna es descriuen les 4 situacions reproduïdes, les quals han de ser reconegudes pels subjectes. El senyal utilitzat ha sigut un senyal de recepció de cridada telefònica antiga (*ringin.wav*). Dos són intervals ascendents i dos descendents. Els dos primers són intervals sintetitzats amb el mètode de Síntesi II, i els dos segons amb el mètode de Síntesi I. En general els resultats han sigut satisfactoris excepte

	Quantitat d'encerts	Quantitat d'errades
↑1r to en 0°, 2n to en 60°	9	3
↓1r to en 45°, 2n to en 15°	10	2
↓1r to en 60°, 2n to en 30°	2	10
↑1r to en 15°, 2n to en 45°	8	4

Taula 5.2: Avaluació subjectiva de dos tons

per al tercer interval, al qual un dels dos tons es sentia saturat després d'haver-se filtrat.

Finalment, en la Taula 5.3 es veuen els resultats de la tercera avaluació. En la primera columna es veuen fletxes que indiquen com ha de ser més o menys el dibuix fet pel subjecte.

	Quantitat d'encerts	Quantitat d'errades
↗ De 0° a 60°	8	4
↘ De 0° a 60° i de 60° a 0°	9	3
↘ De 60° a 0°	7	5

Taula 5.3: Avaluació subjectiva de so dinàmic.

Dels tres tests d'avaluació fets, l'últim ha sigut el que millors resultats ha aportat.

Capítol 6

Resum i conclusions

En aquest treball s'ha fet un anàlisi de la capacitat dels oïents de percebre sons que arriben per diferents anles d'elevació. En un primer moment s'ha fet una revisió bibliogràfica (Capítol 1) per veure quins estudis sobre àudio 3D s'havien fet fins a l'actualitat, i decidir de quina forma millorar-los per a aconseguir l'efecte de l'elevació.

A continuació s'ha analitzat el model de l'HRTF (Capítol 2) que ha de ser aplicat segons les nostres previsions per al disseny i realització dels filtres per al procesament de senyals en elevació. Seguidament s'ha fet un estudi de les tècniques de síntes de camp d'ones (Capítol 3) que seran combinades amb els filtres dissenyats per a aconseguir l'efecte d'elevació en un camp més extens que les tècniques binaurals i estereofòniques.

Al disseny dels filtres HRTF en elevació (Capítol 4) s'ha fet un anàlisi de dades per a crear una base de dades, lo suficientment extensa, per a crear un filtre que s'ajuste a una gran quantitat de subjectes. Després s'ha fet una descripció dels filtres paramètrics utilitzats per al disseny dels filtres HRTF. A més a més s'han proposat dos camins diferents per a la creació d'aquestos filtres, un basat en la recerca de funcions elementals dins de la funció mitja d'HRTF de diversos subjectes, i un segon basat en tècniques d'optimització.

Finalment, s'ha analitzat la capacitat dels filtres amb el procesament de senyals reals i el conseqüent test subjectiu (Capítol 5), que consta de 3 proves diferents per a analitzar l'efecte que produeixen els filtres dissenyats en un conjunt de subjectes, al ser aplicats a diferents tipus de senyals sonores que més tard es reproduïxen en un sistema de Wave-Field Synthesis.

Els filtres dissenyats i utilitzats en aquest treball no donen resultats d'alta qualitat, però si satisfactoris, ja que fins ara sols s'havien obtingut efectes d'elevació en tècniques binaurals i estereofòniques. El treball desenvolupat dona esperances positives, i com que els resultats no són definitius obri pas a la reproducció de l'elevació en sistemes WFS en un futur pròxim.

Apèndix A

Optimització de funcions

En aquest annex es presenta la teoria bàsica per a la minimització de funcions multidimensionals, que serveix com a suport per al disseny dels filtres d'HRTF en elevació. A més a més s'explica el funcionament del mètode de minimització sense restriccions utilitzat al disseny d'un dels filtres d'aquest treball, com és el *Mètode de descens del símplex*.

A.1 Teoria general de minimització de funcions

L'optimització de funcions és una eina matemàtica amb gran àmbit d'aplicació. El seu objectiu es trobar la millor solució per a un determinat problema matemàtic, expressat normalment mitjançant una equació. Aquesta millor solució sol ser el punt en el que la funció pren el valor mínim o màxim. Per aquest motiu s'utilitza en el disseny de filtres HRTF en elevació, per a cercar els màxims i mínims del filtre que representa l'oïda de l'esser humà i així, d'aquesta manera, poder fer la síntesi del filtre ajustant de la millor forma els filtres *peak*. Com cercar el màxim d'una funció és el mateix que cercar el mínim de la funció inversa, es parla generalment de *minimització de funcions*.

El problema consisteix en cercar el mínim d'una funció objectiu, a la que li se passa com a paràmetre d'entrada un vector de variables a ajustar. A més a més, és important que aquests mínims es cerquen de la forma més ràpida i amb el menor cost computacional.

Hi ha dos tipus de mínims: mínim global, si no existeix cap punt de la resta de la funció que siga menor que aquest; mínim local, si per a una regió determinada és el punt mínim de la funció. Les funcions amb un sol mínim s'anomenen *funcions unimodals* i aquelles en les que apareixen mínims locals i global s'anomenen *funcions multimodals*. Mitjançant derivades de primer i segon ordre, s'avaluen les pendents i la curvatura de la funció en cada punt.

Per a que un punt minimitze la funció objectiu, ha de complir dues propietats o *Condicions necessàries*:

- La derivada direccional de f al llarg de *qualsevol* direcció s , avaluada en eixe punt x^* , serà 0 (és a dir, el punt x^* és un extrem de la funció f).

- La funció *mai serà còncava* en eixe punt, la curvatura de la funció en dit punt al llarg de qualsevol direcció és *no negativa*.

En general, tots els mètodes de minimització de funcions tenen una mateixa estructura: Parteixen d'un punt inicial, van prenent aproximacions successives de la solució segons la informació del valor de la funció i d'algunes de les derivades i s'obté una successió de punts que convergeixen al mínim de la funció. Aquest procés iteratiu finalitza quan es compleix algún *criteri de terminació*, com poden ser, parar si el gradient de la funció en el punt actual és suficientment pròxim a zero o finalitzar si, entre dues iteracions successives, la distància entre dos punts x és menor que un llindar imposat amb anterioritat x_{tol} .

Es pot consultar més informació sobre teoria de minimització de funcions i optimització a [21]

A.2 Mètode de descens del símplex

El *Mètode de Descens del Símplex* és un mètode de búsqueda directa molt robust, que sols avalua la funció objectiu per a cercar el mínim d'aquesta. Aquest mètode és adequat per a funcions amb una alta no linealitat o amb discontinuïtats, com són les funcions que formen el filtre HRTF en elevació.

Un símplex és una figura geomètrica de n dimensions i $n + 1$ vèrtex, on n és la quantitat de variables de les que depèn la funció a minimitzar. És important resaltar que un símplex no té per que ser regular. Així mateix, no pot haver diversos vèrtex coincidents en un mateix punt, és a dir, no pot ser una figura degenerada.

Per començar amb l'algorisme és necessari formar un símplex inicial. La forma més senzilla d'obtindre els punts és la següent: es parteix d'un punt inicial P_0 , que serà un vector de n variables, cadascun dels n vèrtex restants es calcularà desplaçant-se una determinada quantitat en cadascuna de les n direccions de l'espai segons la següent expressió:

$$P_i = P_0 + \lambda \hat{e}_i, \quad i = 1, 2, \dots, n \quad (\text{A.1})$$

on cada \hat{e}_i és un vector unitari de l'espai n -dimensional, i λ és una constant, que podria ser diferent per a cada direcció. Gràficament açò es pot veure en la Figura A.1.

El següent pas és ordenar els vèrtex del símplex. Per a açò, es calcula el valor de la funció en cadascun d'aquests. En funció dels valors aconseguits, s'anomenarà vèrtex 0 al vèrtex en què la funció pren el menor valor, i així fins al vèrtex n , en el qual la funció pren el major valor.

A partir d'aquesta estructura, comencen les iteracions del mètode. La particularitat està en la forma en la que s'avança cap al mínim de la funció: en cada iteració, el pitjor vèrtex (aquell de major valor en la funció) es substitueix per un altre en què el valor siga menor. Açò és el que s'anomena *moviment del símplex*. D'aquesta forma, poc a poc el símplex va lliscant fins al mínim de la funció.

Com no s'avança en la direcció del millor vèrtex, sinó que s'elimina el pitjor, el mètode no convergeix amb massa rapidesa. No obstant, es pretén amb aquesta estratègia guanyar robustesa. Gràcies a aquest tipus de moviments s'evita que el mètode es quede atrapat en un mínim local, cosa molt important.

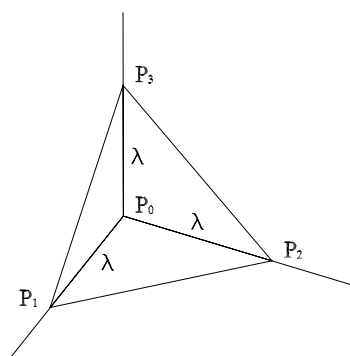


Figura A.1: Formació del símplex inicial.

En cada iteració s'ha de substituir el vèrtex \mathbf{P}_n per un altre punt, al que s'arriba mitjançant una sèrie de moviments. Els distints moviments possibles són: reflexió, expansió, contracció externa, contracció interna i contracció multidimensional.

A continuació es veu la forma en la que s'opta per un moviment o per un altre per a cada iteració.

En primer lloc, s'estableix la direcció en la que es buscarà un punt millor. Com interessa allunyar-se el màxim possible del pitjor punt (\mathbf{P}_n), la direcció escollida és la que uneix dit punt amb el punt mig dels restants vèrtex del símplex. Aquest punt mig té la següent expressió:

$$\mathbf{P}_m = \frac{\sum_{i=0}^{n-1} \mathbf{P}_i}{n} \quad (\text{A.2})$$

Per tant, la direcció en la que es meneja ve descrita per $\vec{\mathbf{d}} = \overrightarrow{\mathbf{P}_n \mathbf{P}_m}$.

Una vegada definida la direcció de búsqueda, es comença amb un moviment de *reflexió*. En aquest cas, el punt reflexat correspon a:

$$\mathbf{P}_r = \mathbf{P}_m + \rho \vec{\mathbf{d}} \quad (\text{A.3})$$

on ρ és un paràmetre que es pot ajustar, encara que sol prendres $\rho = 1$. El moviment de reflexió es pot observar gràficament en la Figura A.2(a).

Una vegada tenim el punt reflexat, es calcula el valor de la funció en dit punt. Si $f(\mathbf{P}_r) < f(\mathbf{P}_{n-1})$, és a dir, si el nou punt és millor que el pitjor dels $n - 1$ restants, s'accepta el punt reflexat. Si, a més a més, aquest nou punt és millor que la resta, és a dir, $f(\mathbf{P}_r) < f(\mathbf{P}_0)$, sembla lògic pensar que el símplex s'ha mogut en una direcció adequada. Per tant, el que es fa a continuació és provar amb un moviment de *expansió*, com es veu a la Figura A.2(b), per a veure si amb un avanç un poc major s'obté un punt encara millor. El punt expandit té la següent expressió:

$$\mathbf{P}_e = \mathbf{P}_m + \Psi \rho \vec{\mathbf{d}} \quad (\text{A.4})$$

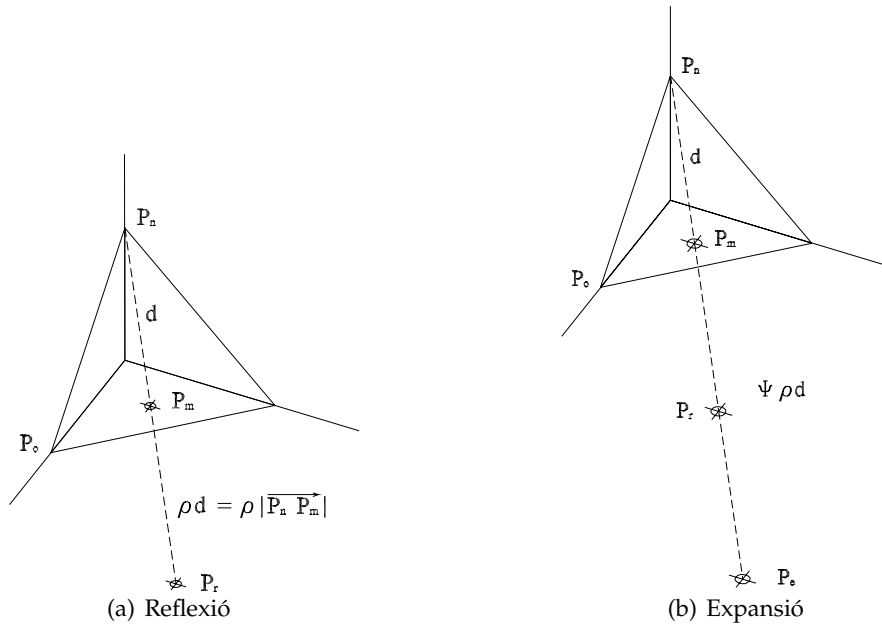


Figura A.2: Moviments de reflexió i d'expansió

on Ψ és un altre paràmetre que es pot ajustar en funció de les necessitats particulars de cada funció. El seu valor serà sempre major a 1 (sol ser 2).

Si pel contrari, la primera comprovació no es compleix, és a dir, el punt reflexat no és millor que el pitjor dels $n - 1$ restants però sí és millor que \mathbf{P}_n ($f(\mathbf{P}_r) > f(\mathbf{P}_{n-1})$ però $f(\mathbf{P}_r) < f(\mathbf{P}_n)$), es suposa que s'ha avançat massa en la direcció escollida. En aquest cas el que es fa, és retrocedir un poc i comprovar el que succeeix amb un moviment de *contracció externa*, el qual apareix a la Figura A.3(a). L'expressió d'aquest nou punt és la següent:

$$\mathbf{P}_{ce} = \mathbf{P}_m + \phi \rho \vec{\mathbf{d}} \quad (\text{A.5})$$

on ϕ és un altre paràmetre característic del mètode, el valor del qual serà sempre menor que 1 (normalment 0.5).

Un altra possibilitat és que el punt reflexat \mathbf{P}_r no siga tampoc millor que el pitjor de tots els punts, \mathbf{P}_n . En aquest cas, no té sentit avançar allunyant-se de \mathbf{P}_n . Es provarà amb un apropament, mitjançant un moviment anomenat *contracció interna*, representat en la Figura A.3(b). D'aquesta forma, el punt de contracció interna té la següent expressió:

$$\mathbf{P}_{ci} = \mathbf{P}_m - \phi \rho \vec{\mathbf{d}} \quad (\text{A.6})$$

Finalment, si el moviment de contracció interna tampoc obté un punt millor que el \mathbf{P}_{n-1} , es realitza un moviment de *contracció multidimensional*. Amb aquest moviment, plasmat a la Figura A.4, els punts del símplex s'apropen al millor d'ells, \mathbf{P}_0 . D'aquesta forma s'incrementen les probabilitats que els valors de la funció siguin millors en la següent iteració.

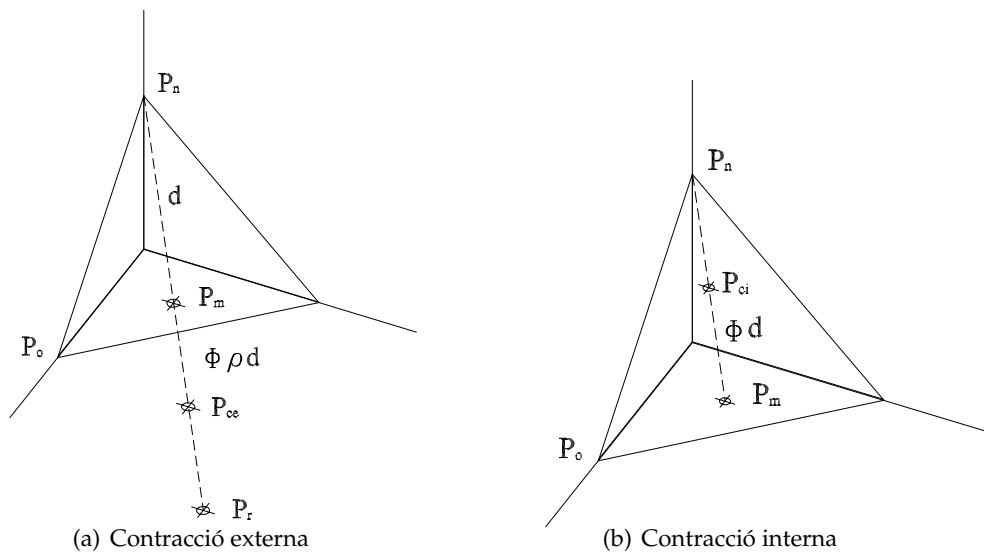


Figura A.3: Moviments de contracció externa i de contracció interna

En el moviment de contracció multidimensional, tots els vèrtex excepte P_0 es modificaran de la següent forma:

$$P_i = P_0 + \sigma \overrightarrow{P_0 P_i} \tag{A.7}$$

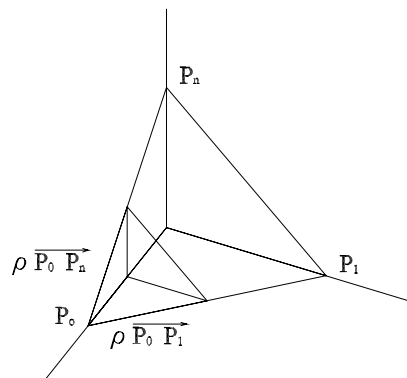


Figura A.4: Moviment de contracció multidimensional

Es fan comprovacions en cada iteració, per determinar quin dels moviments es realitza. El mètode finalitza quan el símplex té unes dimensions inferiors a un llindar especificat, o també si es supera la quantitat d'iteracions desitjades, o d'avaluacions de la funció objectiu.

La qualitat d'aquest mètode és la robustesa. Gràcies a l'estratègia que es du a terme per a determinar el moviment a realitzar, es minimitza la probabilitat que el símplex quede atrapat en un mínim local, encara que per a que siga possible, siga necessari sacrificar la velocitat de la convergència.

Per a una informació més extensa es poden consultar [11], [18] i [19].

Bibliografía

- [1] BERKHOUT, A. J. *Applied Seismic Wave Theory*. 1987.
- [2] BERKHOUT, A. J. A holographic approach to acoustic control. *Journal of the Audio Engineering Society*, 36 (1988), 977–995.
- [3] BERKHOUT, A. J., DE VRIES, D., AND VOGEL, P. Wave front synthesis: a new direction in electro-acoustics. In *In 93th Conv. Audio Eng. Soc.* (San Francisco, USA, Mar. 1992), no. 3379.
- [4] BERKHOUT, A. J., DE VRIES, D., AND VOGEL, P. Acoustic control by wave field synthesis. *Journal of the Acoustic Society of America* 93 (1993), 2765–2778.
- [5] BLAUERT, J. *Spatial hearing*. MIT Press, Cambridge, MA (1997).
- [6] BLEISTEIN, N. *Mathematical Methods for Wave Phenomena*. 1984.
- [7] BOONE, M., AND VERHEIJEN, E. N. G. Multichannel sound reproduction based on wave field synthesis. In *In 95th Conv. Audio Eng. Soc.* (New York, USA, Oct. 1993), no. 3719.
- [8] CARLILE, SIMON, E. *Virtual auditori space: Generation and applications*. R.G. Landes Company, Austin TX (1996).
- [9] DE VRIES, D. Sound reinforcement by wave field synthesis: Adaptation of the synthesis operator to the loudspeaker directivity characteristics. *J. Audio Eng. Soc.* 44(12) (Dec. 1996), 1120–1131.
- [10] DUDA, R. Modelling head related transfer functions. In *Proceedings Twentyseventh Annual Asilomar Conference on Signals, Systems and Computers*, Asilomar, CA (1993).
- [11] ESTEBAN, H. Análisis y síntesis de circuitos de microondas. *Curso de Máster, programa de comunicaciones, Universidad Politécnica de Valencia* (2007-2008).
- [12] GARDNER, B., AND MARTIN, K. Hrtf measurements of a kemar dummy-head microphone. *Technical Report 280, MIT Media Lab Perceptual Computing* (May 1994).
- [13] GARDNER, B., AND MARTIN, K. Hrtf measurements of a kemar dummy-head microphone. In *MIT Media Lab Perceptual Computing - Technical Report #280* (May 1994).

- [14] GENUIT, K. A model for the description of outer-ear transmission characteristics. *PhD thesis, Rheinisch-Westphalian Technical University, Aachen, Germany* (1984).
- [15] HORBACH, U., AND BOONE, M. Future transmission and rendering formats for multichannel sound. In *In AES 16th Int. Conf. on Spatial Sound Reproduction* (Rovaniemi, Finland, Apr. 1999).
- [16] IRCAM, R. A. T. In <http://recherche.ircam.fr/equipes/salles/listen/index.html> (Paris, France, 2003).
- [17] LOPEZ MONFORT, J. J. Reproducción de sonido 3-d mediante técnicas de control local. *Tesis Doctoral. Universidad Politécnica de Valencia* (1999).
- [18] NELDER, J., AND MEAD, R. A simplex method for function minimization. *Computer J.* 7 (January 1965), 308–313.
- [19] PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T., AND FLANNERY, B. P. Numerical recipes in c: The art of scientific computing. *Cambridge University Press* (1992).
- [20] RAYLEIGH, L. On the acoustic shadow of a sphere. *Phil. Transactions of the Royal Society*, 203A:87-99 (1904).
- [21] ROGER, F. Practical methods of optimization. *John Wiley and Sons, second edition* (1987).
- [22] RSS-10, R. <http://www.rolandus.com/products/productdetails.aspx?objectId=411> (1995).
- [23] SNOW, W. Basic principles of stereophonic sound. *Transactions on Audio* (Mar. 1955), 44–55.
- [24] WITTEK, H. *Perception of spatially synthesized sound fields*. 2003.