

Källorienterat ljud – Del 1

Alf Berntson

I musikalerna och teaterföreställningarna med förstärkt ljud är det snarare regel än undantag att ljudet kommer från fel håll. Källorienterad förstärkning gör det nu möjligt att åstadkomma en upplevelse av att ljudet kommer från artisten – inte från högtalarna. I denna artikel beskriver Alf Berntson problemet och hur hörseln lokaliserar ljudkällor. I del 2 i nästa nummer kommer dagens och framtidens tekniska lösningar att presenteras.

Problemet med fel lokalisering

Fel lokalisering av ljudkällor (exempelvis sångare) på en scen leder till:

- sämre tydlighet och textuppfattbarhet
- svårighet att skilja mellan skådespelare i en dialog
- förvirring, eftersom örats och ögats perspektiv inte överensstämmer
- att publiken får det svårare att tro på illusionen

Fel lokalisering innebär att hörselintrycket ger felaktig riktning *och/eller* felaktigt avstånd jämfört med synintrycket. Det har sagts att ”öronen är till för att peka åt ögonen”. Hörseln kan lokalisera ljud från alla håll till skillnad från synen som ju är begränsad till synfältet. Dessutom kan vi inte blunda med öronen (även om det skulle vara skönt ibland). Här har vi själva kärnan till problemet med fel lokalisering i teaterljudsammanhang. På teatern sitter vi ofta relativt långt från artisterna på scenen och dessutom känner vi inte igen rösterna. Pekar inte öronen rätt åt ögonen så börjar blicken att flacka. Vem är det som sjunger? Vi blir förvirrade och får svårt att följa vad som händer på scenen. Om hörseln däremot har möjlighet att lokalisera ljudet så kan vi koncentrera oss på händelseförloppet och textuppfattbarheten

förbättras.

Med dagens teknik går det att (med viss omsorg) åstadkomma en bra ljudkvalitet från ljudet som kommer ut ur högtalarna. Men i musikalerna och andra uppsättningar som förlitar sig på förstärkning av artister på scenen är det snarare regel än undantag att ljudet kommer från fel håll och fel avstånd. Man hör ofta sångaren i den närmsta högtalaren. Är det flera som sjunger eller talar samtidigt så är det väldigt svårt att höra vem som säger vad.

Fel lokalisering kan även förekomma i naturlig akustik utan förstärkning. Ett vanligt exempel är att vissa instrument i ett orkesterdike kan låta som de satt bakom en. Detta orsakas vanligtvis av starka reflexer från konkava ytor samtidigt som direktljudet är svagt.

Den ideala ljudförstärkningsanläggningen skall ge låg distorsion, tillräcklig ljudnivå, tillräcklig uppfattbarhet (tydlighet), naturligt ljud i rummet och korrekt lokalisering av alla ljudkällor. Eftersom artister rör sig på scenen (i alla fall för det mesta) så skall systemet kunna skapa flera samtidiga virtuella ljudkällor som hela tiden kan ändra position oberoende av varandra. Dessutom bör lokaliseringen fungera för större delen av publiken – inte bara i en ”hot spot” där kanske mixerbordet är placerat.

Hur vi lokaliserar ljud

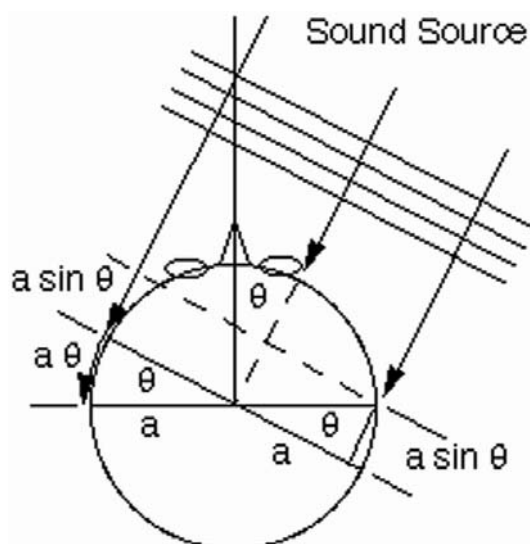
För att kunna skapa virtuella ljudkällor så måste vi tränga in i hur hörseln lokaliserar ljud. Här måste man skilja på två olika

fenomen. Dels hur hörseln lokaliserar en ljudkälla i fritt fält (helt utan reflexer, exempelvis ekofritt rum) och dels i normala rum med reflexer.

Generellt så är hörseln är bäst på att bestämma en ljudkällas riktning i horisontalplanet, sämre på riktning i höjddled och sämst på att bestämma avstånd. För låga frekvenser är riktningsupplösningen klart sämre än för höga frekvenser.

Det viktigaste för att kunna avgöra riktningen på en ljudkälla är att vi har *två* öron – vår hörsel är **binaural**. Hjärnan jämför de två öronens ljudsignaler. Riktning och i viss mån avstånd bestäms av **skillnader i nivå, fas och ankomsttid mellan öronen**. I medianplanet (höjddled och bak/fram) är det **ytteröronen** som ger skillnader i spektrum som bestämmer riktningen. Även huvudets form och överkroppen inverkar men är av mindre betydelse. **Förhållandet mellan direktljudsnivån och efterklangsnivån** är den viktigaste faktorn för att bestämma avståndet till en ljudkälla i normala rum. Att vi kan **röra på huvudet** förbättrar lokaliseringsförmågan.

Figur 1 visar ett förenklat sfäriskt huvud. Infallsvinkeln i horisontalplanet bestäms av skillnaden i ankomsttid (ITD) och nivå mellan öronen.



Figur 1. Snett infall av plan våg (ljudkälla långt bort) i horisontalplanet. Ljudet anländer först till höger öra.

För sfäriskt huvud blir

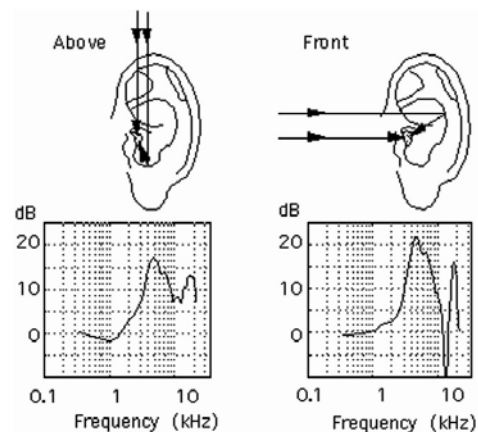
$$ITD = \frac{a}{c}(\theta - \sin \theta), \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$$

och $c = 343$ m/s. För ljud rakt från sidan blir då ITD maximalt 0,7 ms.

Lågfrekventa ljud har stor våglängd och böjs runt (diffraktion) och nivåskillnaden blir liten. Då bestäms riktningen av ITD. För högre frekvenser (ca >1500 Hz) så får vi flera våglängder mellan öronen och riktningen bestäms av nivåskillnaden.

Naturliga ljud är av komplex natur med transienter och in- och utsvängningsförlopp. När ljudet plötsligt ändrar sig så kommer denna ändring anlända till öronen med olika ankomsttid (utom i medianplanet). Det är därför ITD som är viktigast för lokalisering.

För att kunna **lokalisera i medianplanet** (upp/ner och bak/fram) så har vi **ytteröronen** till hjälp. Man kan säga att ytterörat fungerar som en akustisk antenn. Vi får resonanser i kaviteter och interferens mellan olika "ljudvägar" som ger toppar och dalar i frekvensgången. Framför allt dalarna varierar i frekvens beroende på höjdvinkel (se figur 2). Vi får naturligtvis stor skillnad mellan bak och fram.



Figur 2. Ytterörat hjälper oss att lokalisera i höjddled. Olika höjdvinklar ger interferens mellan direktljudet och reflexer från ytterörat med olika notchfrekvenser. För ljud framifrån hamnar notchen vid ca 13 kHz medan ljud uppifrån ger en notch över det hörbara området.

Eftersom höjdlökaliseringen bestäms av spektrala skillnader i högfrequensområdet så är det mycket svårt att lokalisera lågfrekventa ljud och även toner. (Rena toner i rum varierar kraftigt i nivå mellan olika positioner pga moder i rummet och är därför nästan alltid mycket svåra att lokalisera)

Avståndet bestäms huvudsakligen av **förhållandet mellan direktljudsnivån och efterklangsnivån**. Det är därför som extremt smalstrålande högtalare kan ge en känsla av att ljudet är onaturligt nära.

Lagen om första vågfronten

I normala rum har vi förutom direktljudet från en ljudkälla även en mängd reflexer från olika håll. Hur i hela friden skall hjärnan kunna få ordning på alla riktningar? Här har vi tack och lov begåvats med en psykoakustisk effekt som kallas **”lagen om första vågfronten”** eller **”Haas-effekten”**. Kortfattat innebär den att riktningen bestäms av det ljud som når oss först. Men om direktljudet har riktigt låg nivå jämfört med sent ankommande reflektat ljud fungerar inte lokaliseringen.

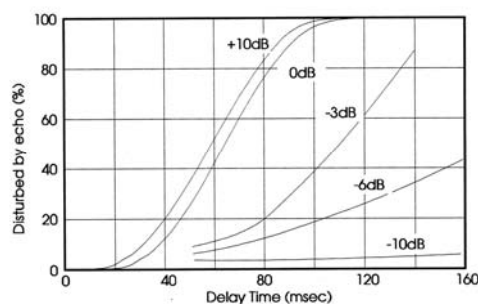
Det har utförts en mängd experiment med försökspersoner för att förstå denna effekt. I de flesta experiment har tal i normal hastighet spelats upp i två högtalare i ”stereoupställning” (ca 80° vinkel). Den ena högtalaren har simulerat direktljud. Den andra har simulerat en reflex med varierande nivå och fördröjning. Med samma nivå på direktljudet och talet så fås en summalokalisering (mellan högtalarna) för fördröjning < 1 ms. Med ökande fördröjning så uppkommer ”lagen om första vågfronten”. När den s.k. ekoträskeln passerats så upplever man direktljudet och ett eko. Slutligen för mycket lång fördröjning upplevs ekot som ett helt separat ljud. Kort fördröjning upp till ca 5 ms kan även upplevas som en färgning pga interferens mellan ljudkällorna (s.k. kamfilter), särskilt om ljudkällorna ligger nära varandra. Även om ”lagen om första vågfronten” gör att vi lokaliserar till det först anlända ljudet så kommer naturligtvis reflexer att påverka upplevd bredd på ljudkällan,

rymdkänsla och klangfärg (vi får en akustisk rumskänsla).

Observera att tröskelvärdena för fördröjningstid och nivå påverkas av direktljudets nivå, typ av signal (pulser, snabbt tal, långsam musik etc) och riktningar på direktljud och reflex.

Haas utförde sina experiment 1951 på 80 testpersoner i ett rum med 0,8 s efterklangstid. Experimentet hade inte att göra med lokalisering över huvud taget. Han ställde två frågor.

- När upplever du dig ”precis störd” av den sekundära signalen?
- När är båda signalerna lika starka?

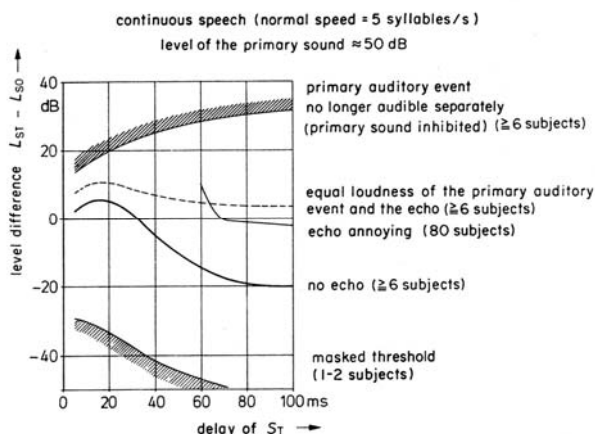


Figur 3. Andel lyssnare som upplevde ett eko som störande som funktion av fördröjningstiden mellan primär- och sekundärljudet. Parameter för de olika kurvorna är nivåskillnaden mellan primär- och sekundärljudet (Haas, 1951). Ett sekundärljud med fördröjning kortare än 30 ms bedöms inte som störande.

Figur 3 visar hur många som kändes sig ”precis störda” av den sekundära signalen (reflexen). Här visas tydligt att det är väldigt få som upplever sig störda av ”ekot” om fördröjningen är kortare än 30 ms. T.o.m. om den fördröjda signalen är 10 dB starkare.

Resultaten av flera andra experiment visas i Figur 4. Här har tal med normal hastighet och direktljudsnivån ca 50 dB använts. Den lägsta tröskeln (maskerings-tröskeln) är den nivå då reflexen inte påverkar ljudupplevelsen över huvud taget vid på- och avslag. Nästa kurva visar ekoträskeln, dvs den nivå då inget eko uppfattas. Kurvan däröver är från Haas och visar

när ekot upplevs störande. Den streckade kurvan däröver anger den nivå då direktljud och reflex upplevs lika starka. Överskrida gränsvärdet anger när direktljudet inte längre uppfattas som ett separat ljud utan reflexen bestämmer lokaliseringen (fel lokalisering).



Figur 4. Jämförelse av olika tröskelnivåer för en reflex relativt direktljudet som funktion av fördröjningstiden. Om fördröjningstiden ligger mellan ca 10-25 ms kan den fördröjda signalen vara upp till ca 4 dB högre än direktljudet utan att den uppfattas som ett separat ljud.

Sammanfattningsvis kan sägas att fördröjda ljud som kommer inom 10-25 ms kan ha upp till 4 dB högre nivå än direktljudet utan att påverka lokaliseringen. Vi kan alltså förstärka ljud utan att påverka lokaliseringen. Kutruff har visat att tidsfördröjningen kan förlängas genom att använda flera högtalare om bara varje ny högtalare håller sig inom ekoträskeln för föregående högtalare.

Cocktail-party-effekten

Alla dessa förmågor som nämnts ovan gör det möjligt för signalbehandlarna på väg in till hjärnan och i hjärnan att fokusera på de ljud som är intressanta för oss och sortera bort ointressanta. Denna förmåga kallas "Cocktail-Party-effekten". Vi har den fantastiska förmågan att skilja ut tal ur brus, skilja ut en talare ur ett sorl från många och att fokusera på en av två sam-

tidiga talare.

Denna effekt undersöktes redan på 50-talet när man försökte förstå och förbättra lyssningen för flygtrafikledare i USA. Forskningen har visat att förmågan till selektiv uppfattbarhet förbättras när:

- Rösterna kommer från olika platser i rummet
- Man har möjlighet att associera till visuella intryck som läpprörelser, kroppspråk och gester. Belysning av teaterscener har alltså betydelse för taluppfattbarheten.
- Rösterna har olika klangfärg eller dialekt.
- Texten har ett "förväntat" sammanhang och flöde. Hjärnan har vid konventionella texter förmågan att "fylla i" ord som inte uppfattas.

Cocktail-party-effekten försämras om inte upplevelsen av ljudkällans förflyttning följer talaren med korrekt hastighet och utan "språng".

Lösningar och framtida möjligheter

Med hjälp av en matricemixer med möjlighet till reglering av både nivå och fördröjning i korspunkterna kan vi få "lagen om första vågfronten" att fungera och åstadkomma en korrekt lokalisering. Kommerciella ljudprocessorer kan i dag skapa flera samtidiga virtuella ljudkällor på detta sätt. Är högtalarna korrekt placerade, dimensionerade och injusterade så kan lokaliseringen fungera för större delen av publiken – inte bara i en "hot spot" som för konventionell stereo.

Mer om detta i del 2 i nästa nummer!

Referenser

Blauert, J. (1997): Spatial Hearing. The Psychophysics of Human Sound Localization.

Cremer, L. and Müller, A. (1982): Principles and Applications of Room Acoustics. Volume 1.

Duda, R. O.: Sound Localization Research. Hemsida
<http://www.engr.sjsu.edu/~duda/Duda.Research.html>

TiMax User Manual (1998).

Arons, B (1992): A Review of the Cocktail Party Effect.



Artikelförfattaren Alf Berntson arbetar som konsult inom rumsakustik, elektroakustik och styrsystem. Han är teknologie licentiat inom psykoakustik och har de senaste 15 åren arbetat med projektering. Bland annat har han projekterat ljudanläggningar för Göteborgs-Operan, Norrköpings Konserthus och Göteborgs Stadsteater. För Göteborgs stadsteater har han även projekterat det nya inspicient- och show control systemet. Alf jobbar nu bl.a. med projektering av ljud- och inspicientanläggningar för nya operahuset i Oslo. Du kan nå Alf på alf@artifon.se eller på telefon 031-761 88 10.