RAPPORT

Kvantifiera prestanda för akustiska mikrofonuppsättningar

Mer än bara antalet mikrofoner

Patrick W.A. Wijnings, MSc
Doktorand inom elektroteknik

Introduktion

Under de senaste årtiondena har det skett imponerande tekniska framsteg – från datorprestanda [1] till antalet pixlar i en digitalkamera [2]. Det här är en följd av Moores lag [3], som är både observationen av och prognosen för att antalet transistorer i en integrerad krets ökar exponentiellt med tid. En liknande trend kan observeras för akustiska mikrofonuppsättningar: i och med framstegen inom integrerade kretsar har digitala mikrofoner baserade på MEMS-teknik (Micro-Electro-Mechanical Systems) [4] utvecklats som ett ekonomiskt, småskaligt alternativ till analoga kondensatormikrofoner för användning i mikrofonuppsättningar.

Det här har möjliggjort för dagens mikrofonuppsättningar att bestå av 32–128 mikrofoner, eller ännu fler: världens största mikrofonuppsättning, enligt Guinness rekordbok [5], består av imponerande 4 096 mikrofoner. Det är dock viktigt att inte bli bländad av dessa siffror: ungefär som Megahertz-myten för processorer eller megapixelmyten för fotografering kan det skapas en mängdmyt för akustiska mikrofonuppsättningar. Till skillnad från vad man kanske frestas att tro berättar inte enbart antalet mikrofoner hela historien om akustiska prestanda. I synnerhet är placeringen av mikrofonerna en lika viktig faktor i och med fysikens lagar, vilket i slutändan begränsar de mätprestanda som går att uppnå.

Den här rapporten ger en översikt över effekterna av mikrofonernas antal och placering på mätprestanda och tillhandahåller de grundläggande verktygen för att kvantifiera denna.

Proportionell fördelning minskar brus

Varje mikrofon ger oundvikligen upphov till en del brus vid mätningar av ljudtryck: på grund av produktionstoleranser varierar känsligheten något från mikrofon till mikrofon [6], och mikrofonens elektronik ger också upphov till eget brus. När ett (tyst) ljud överskuggas av det här bruset går det inte att identifiera den tillhörande ljudkällan. Eftersom ljudet klingar av när det rör sig bort från källan kan det också begränsa detekteringsområdet för den akustiska mikrofonen. Som tur är finns det ett välkänt fenomen inom signalbehandling som innebär att snittet av proportionellt fördelade mätningar från flera mikrofoner tenderar att minska bruset:

$$\text{Brusminskning (i dB)} = 20 \log_{10} \left(\sqrt{\text{Antal mikrofoner}}\right)$$

Den här formeln innebär att varje gång antalet mikrofoner fördubblas minskas bruset med 3 dB, vilket det mänskliga örat knappt uppfattar under normala förhållanden [7]. Även en mindre akustisk mikrofonuppsättning överträffar därför en enda mikrofon avsevärt (t.ex. minskar en uppsättning med 64 mikrofoner bruset med 18 dB), men effekten avtar om antalet mikrofoner ökas ytterligare (t.ex. är skillnaden mellan uppsättningar med 128 och 64 mikrofoner bara 3 dB). Förr eller senare börjar dessutom brus som är gemensamt för alla mikrofoner (t.ex. från nätaggregat) ta över, eftersom det inte kan minskas genom proportionell fördelning. Slutligen kräver fler mikrofoner mer databearbetning, vilket antingen försämrar batteriets livslängd och bärbarhet eller kräver kompromisser vad gäller bildhastighet eller upplösning.

Sammanfattningsvis kan fler mikrofoner minska bruset, men vid en viss punkt minskar fördelarna och väger inte upp nackdelarna. Med dagens teknik anser vi att det optimala antalet ligger runt 64 mikrofoner.

Strålformning och vågutbredning

Att bara fördela mikrofonens mätningar proportionellt möjliggör inte visualiseringen av ljud som projiceras över en kamerabild. Istället måste en algoritm för strålformning [8] användas. Strålformning kombinerar signalerna från alla mikrofoner i en uppsättning så att data från källor i särskilda vinklar får konstruktiv interferens medan andra får destruktiv interferens.

Den underliggande fysiken som möjliggör strålformning fastställs av vågekvationen.[9] Ljud sprids i synnerhet genom luften med en fast hastighet:

$$\text{Ljudhastighet} = 343 \text{ m/s}.$$

Det innebär att ljudet inte bara har en frekvens utan även en våglängd:

$$\text{Våglängd (i m)} = \frac{\text{Ljudhastighet (i m/s)}}{\text{Frekvens (i Hz)}}$$

Om man skulle frysa tiden är våglängden ljudvågens fysiska längd i den riktning som den sprids.

Strålformning utnyttjar det faktum att de olika mikrofonerna i en akustisk uppsättning mäter olika punkter i den här vågen. Därför är rätt förhållande mellan avstånden mellan mikrofoner i uppsättningen och våglängden avgörande för att få bra akustiska prestanda. Vid låga frekvenser är våglängden stor (t.ex. 3,4 m vid 100 Hz) och uppsättningen har långa avstånd mellan mikrofonerna. Vid höga frekvenser och ultraljudsfrekvenser är våglängden liten (t.ex. 17 mm vid 20 kHz) och uppsättningen har korta avstånd mellan mikrofonerna.

Uppsättningens diameter avgör upplösningen i låga frekvenser

Det längsta avståndet mellan mikrofonerna motsvarar två mikrofoner på motsatta sidor av den akustiska uppsättningen. Det innebär att uppsättningens diameter (fig. 1) är relaterad till upplösningen i låga frekvenser: när uppsättningen är för liten smälter separata akustiska källor samman med varandra i den projicerade ljudbilden. En analogi finns inom astronomi, där ett teleskop med en större diameter (dvs. bländare) ger upplösning för mindre detaljer.

Rayleigh-kriteriet kvantifierar den här effekten genom att approximera den minsta vinkeln mellan två akustiska källor under vilken de smälter samman med varandra:

$$\text{Vinkel (i grader)} = 69{,}88 \times \frac{\text{Våglängd (i m)}}{\text{Mikrofonuppsättningens diameter (i m)}}$$

En ökning av uppsättningens diameter minskar förstås bärbarheten avsevärt. Som tur är avger många intressanta ljudkällor, som luftläckage och elektriska urladdningar, främst ljud vid höga frekvenser. För de här tillämpningarna påverkar en kompakt akustisk mikrofonuppsättning inte prestanda avsevärt.

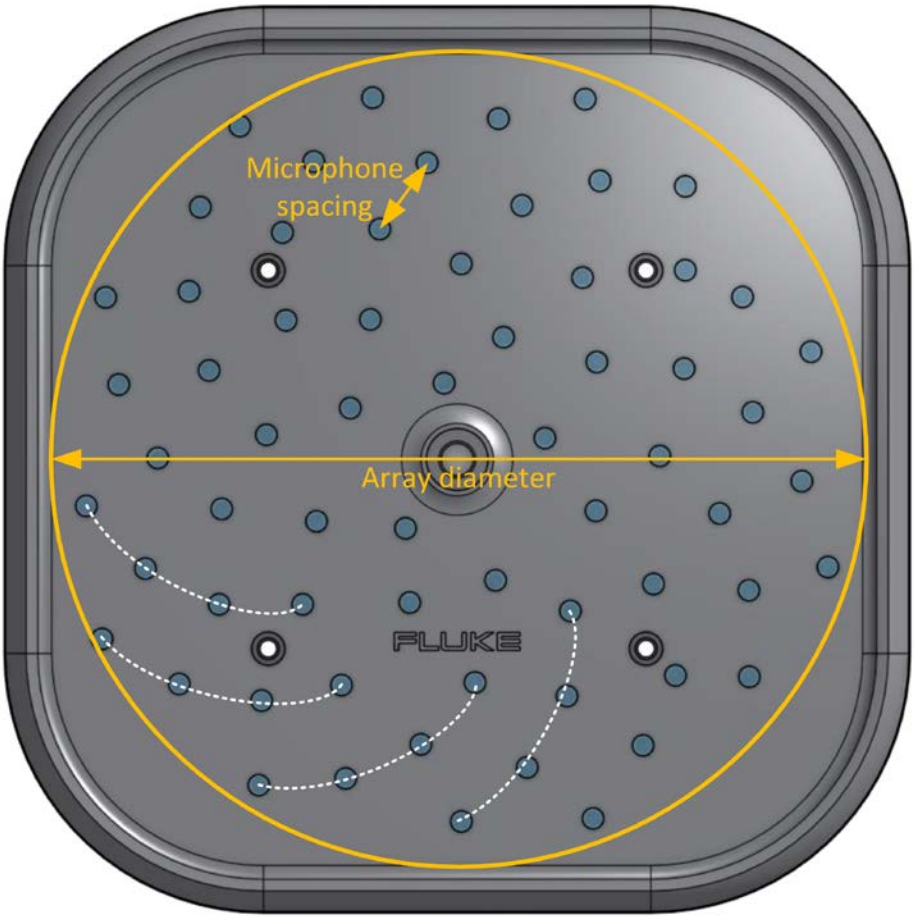

Fig. 1: Uppsättningens diameter och mikrofonernas avstånd är två viktiga parametrar för en akustisk uppsättning. Stora uppsättningar har bättre upplösning i låga frekvenser och uppsättningar med tätare placerade mikrofoner förhindrar vikningsartefakter i höga frekvenser. Prestanda i höga frekvenser kan förbättras ytterligare genom att mikrofonerna placeras i en spiral (anges med de vita streckade linjerna).

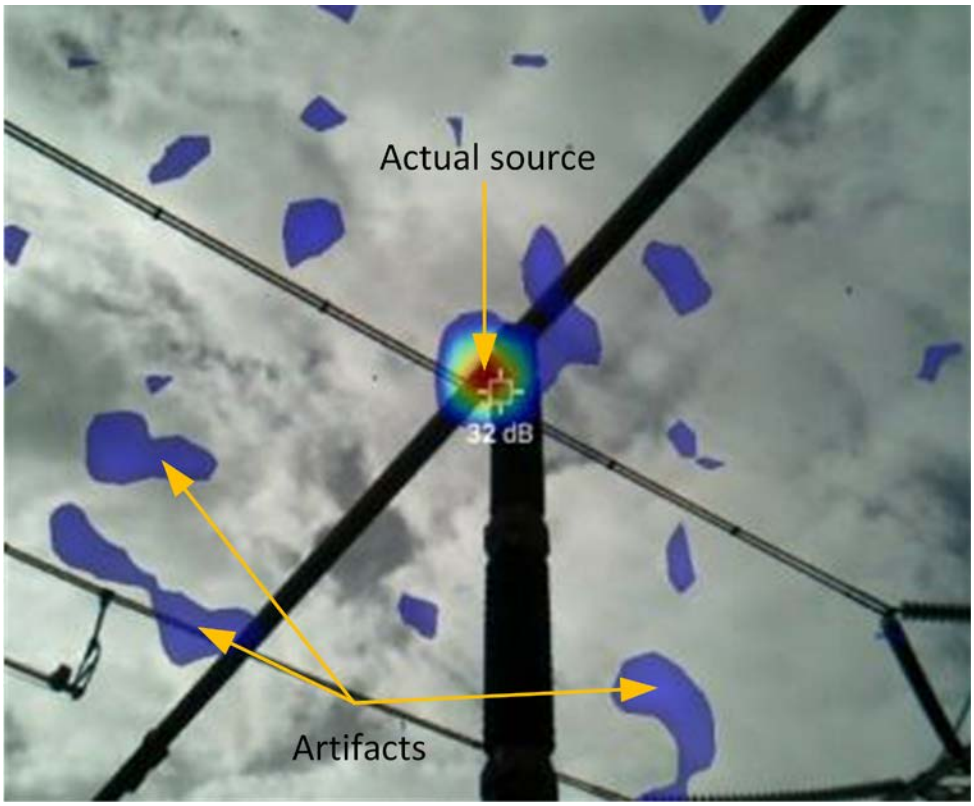

Fig. 2: Exempel på artefakter i den projicerade ljudbilden.

Mikrofonernas avstånd avgör mängden artefakter i höga frekvenser

Det minsta avståndet mellan mikrofonerna motsvarar två närliggande mikrofoner (fig. 1). Därför är mikrofonavståndet relaterat till prestanda i höga frekvenser: när avståndet är för stort kan ljudvågorna inte tolkas unikt, vilket resulterar i vikning (även kallat starka sidolober). Detta yttrar sig som artefakter eller ”spökkällor” (som egentligen inte finns där) i den projicerade ljudbilden (fig. 2). En analogi finns inom fotografering, där artefakter kan förekomma om pixlarna i bilden är för stora. Om mikrofonerna placeras i ett vanligt rutnät kvantifieras den frekvens över vilken vikning kan inträffa med hjälp av Nyquist-kriteriet:

$$\text{Vikningsfrekvens (i Hz)} = 0{,}5 \times \frac{\text{Ljudhastighet (i m/s)}}{\text{Mikrofonavstånd (i m)}}$$

För att minska mikrofonavståndet måste antingen uppsättningens diameter minskas (vilket påverkar upplösningen i låga frekvenser) eller antalet mikrofoner ökas (vilket påverkar batteriets livslängd). Båda alternativen har oönskade effekter.

Som tur är finns det en smartare lösning: Nyquist-kriteriet kan överskridas genom att regelbundenheten i mikrofonrutnätet bryts. Vid signalbehandling kallas detta fenomen för gles sampling (även kallat komprimerad avkänning) eftersom ett oregelbundet rutnät kan konstrueras med sampling från punkter (och där andra förbises) ur ett mycket finmaskigare regelbundet rutnät. Vid vissa frekvenser kan givetvis även gles sampling fungera dåligt, men vid korrekt implementering kan denna frekvens användas mycket långt in i ultraljudsintervallet.

Det finns många sätt att välja ut ett lämpligt oregelbundet rutnät. [10] En intressant lösning hämtar inspiration från naturen: Fermats spiral, som beskriver fördelningen av frön i ett solroshuvud. Med den här ”solrosspiralen” fördelas mikrofonerna effektivt och nästan jämnt över uppsättningens yta, men på ett sätt som gör att avståndet mellan mikrofonparen varierar något (fig. 1). Det gör att en solrosuppsättning minskar vikningsartefakterna avsevärt jämfört med en vanlig uppsättning med samma antal mikrofoner, eller omvänt, ger samma akustiska prestanda med färre mikrofoner och därmed längre batterilivslängd.

Slutsatser

Vi har gett en översikt över aspekterna kring hur mikrofonernas antal och placering påverkar systemets prestanda. Vi har också tillhandahållit de grundläggande verktygen för att kvantifiera detta. En kompakt solrosuppsättning med cirka 64 mikrofoner som Fluke ii900 ger en utmärkt balans mellan akustiska prestanda – särskilt för källor som orsakas av luftläckage eller elektriska urladdningar – och användbarhetsaspekter som batteritid och bärbarhet.

Referenser

[1] Karl Rupp, ”48 Years of Microprocessor Trend Data”, https://github.com/karlrupp/microprocessor-trend-data [Konsulterades: 7 augusti 2020], augusti 2020.

[2] Nathan Myhrvold, ”Moore’s Law Corollary: Pixel Power”, The New York Times, juni 2006.

[3] Gordon E. Moore, ”Cramming more components onto integrated circuits”, Electronics, vol. 38, nr 8, april 1965.

[4] Siti Aisyah Zawawi, Azrul Azlan Hamzah, Burhanuddin Yeop Majlis och Faisal Mohd-Yasin, ”A Review of MEMS Capacitive Microphones,” Micromachines, vol. 11, nr 5, s. 484, maj 2020, nummer: 5 Utgivare: Multidisciplinary Digital Publishing Institute.

[5] Guinness World Records, ”Largest microphone array”, https://www.guinnessworldrecords.com/world-records/largest-microphone-array/ [Konsulterades: 18 september 2020].

[6] Patrick W.A. Wijnings, Sander Stuijk, Rick Scholte och Henk Corporaal, ”Characterization of MEMS Microphone Sensitivity and Phase Distributions with Applications in Array Processing”, 2020, under granskning.

[7] ABD Engineering and Design, ”How Much is a Decibel? Sound Perception Vs. Reality”, https://www.abdengineering.com/blog/perception-vs-reality/ [Konsulterades: 8 november 2020].

[8] B. D. van Veen och K. M. Buckley, ”Beamforming: a versatile approach to spatial filtering”, IEEE ASSP Magazine, vol. 5, nr 2, s. 4–24, april 1988.

[9] Hans Bodén, Ulf Carlsson, Ragnar Glav, Hans-Peter Wallin och Mats Åbom, Sound and Vibration, Marcus Wallenberg-laboratoriet, KTH, Stockholm, Sverige, 2007.

[10] Zebb Prime och Con Doolan, ”A comparison of popular beamforming arrays”, in Proceedings of Acoustics 2013 Victor Harbor, Victor Harbor, South Australia, 2013, Australian Acoustical Society.

Fluke. *Keeping your world up and running.®*

Fluke Sverige AB
c/o Gilbarco Veeder-Root
Johannesfredsvägen 11 A
16869 Bromma
Tel: 08 5663 7400
E-mail: cs.se@fluke.com
www.fluke.se

©2021 Fluke Corporation. Med ensamrätt. Data kan komma att ändras utan föregående meddelande. 05/2021 210512-sv

Ändringar får inte göras i det här dokumentet utan skriftligt medgivande från Fluke Corporation.