

Vícekanálová databáze impulzních odezev měřená v hustě pokrytém prostoru

Jaroslav Čmejla <jaroslav.cmejla@tul.cz>, Tomáš Kounovský <tomas.kounovsky@tul.cz>

Vedoucí projektu SGS - Karel Paleček

V tomto příspěvku představujeme databázi vícekanálových nahrávek vytvořených v akustické laboratoři s nastavitelným dozvukem. Nahrávky nesou akustickou prostorovou informaci o pozici zdroje zvuku (reproduktoru). Nahrávky byly pozice v hustě pokrytém prostoru o velikosti $46 \times 36 \times 32$ cm, celkem 4104 pozic. Nahrávky budou posléze využity k detailní analýze beamformerů a jiných metod využívajících vícerozměrných signálů, případně trénování a testování akustických matematických modelů.

Klíčová slova: Impulzní odezva místnosti, Akustická přenosová funkce, Mikrofonní pole, Hustá síť pozic

Úvod

V oboru zlepšování řečových nahrávek je jednou z nejčastěji zkoumaných situací případ, kdy se řečník nachází na jednom statickém místě v zarušeném prostoru. V reálných situacích se však samozřejmě stává, že se řečník pohybuje. Metody pro zlepšení signálu řečníka, které využívají prostorové akustické informace, jsou obvykle zaměřeny pouze v jednom směru, popř. na jeden bod v prostoru. Je žádoucí, aby nepatrné změny pozice zaměřeného řečníka nezpůsobily přílišné zhoršení výsledků navrhovaných metod (tzv. prostorová robustnost). Je tedy vhodné znát schopnosti navrhovaných metod nejenom pro jeden bod, ale i pro celý prostor okolo něj.

K evaluaci prostorové robustnosti algoritmů je třeba znalost šíření zvuku v daném akustickém prostředí, která je vzhledem k mnoha ovlivňujícím faktorům netriviální. Akustická „cesta“ mezi dvěma pozicemi je ovlivněna nejen velikostí místnosti, jejím rozložením, množstvím překážek, ale také materiály, ze kterých jsou tvořeny. Akustickou přenosovou funkcí (angl. Acoustic Transfer Function - ATF) nazýváme popis akustické „cesty“, propagaci zvuku, mezi dvěma pozicemi. Označení ATF se používá pro popis propagace ve frekvenční oblasti, v případě vyjádření popisu v časové oblasti, jej nazýváme impulzní odezvou místnosti (angl. Room Impulse Response - RIR). Při znalosti dané ATF je možné simulovat libovolný zvuk tak, jak by zněl v určené místnosti a mezi určenou pozicí zdroje (reproduktoru) a senzoru (mikrofonu).

Pro výpočet RIR/ATF je zapotřebí získat nahrávku pro danou místnost, danou pozici reproduktoru a mikrofonního pole. To vede k nekonečnému množství různých kombinací, v praxi se tedy místo reálných nahrávek používají generátory umělých impulzních odezev (např. [1]). Ty však představují pouze zjednodušený model, který nemusí vystihovat všechny vlastnosti reálných ATF. Databáze nahrané v reálných podmínkách pak obvykle obsahují kombinaci pouze několika pozic a druhů místností a nejsou příliš vhodné pro evaluaci prostorové robustnosti algoritmů. Z tohoto důvodu jsme se rozhodli připravit databázi, jež zahrnuje pozice zdroje v hustě pokrytém prostoru tvaru kvádru, nahrávaném více mikrofonními poli a v místnosti s více různými dozvukovými (reverberačními) časy.

Metodika

Podobné existující databáze, příklady

- „Multichannel audio database in various acoustic environments“ [3]
- „Database of omnidirectional and b-format room impulse responses“ [2]

První databáze obsahuje nahrávky dvou zdrojů vzdálených 1 a 2 m před lineárními mikrofonními poli na 13 pozicích umístěných v úhlech od -90° do 90° tj. databáze umožňuje umísťovat zdroje zvuku na kružnici před mikrofony. Druhá databáze obsahuje přes 700 RIR nahraných v přednáškových halách s uniformně rozloženými pozicemi zdroje po celém prostoru. Uvedené databáze se nehodí k evaluaci prostorové robustnosti algoritmů, neboť vzdálenosti mezi jednotlivými pozicemi jsou příliš velké (> 0.5 m).

Model místnosti a vybavení

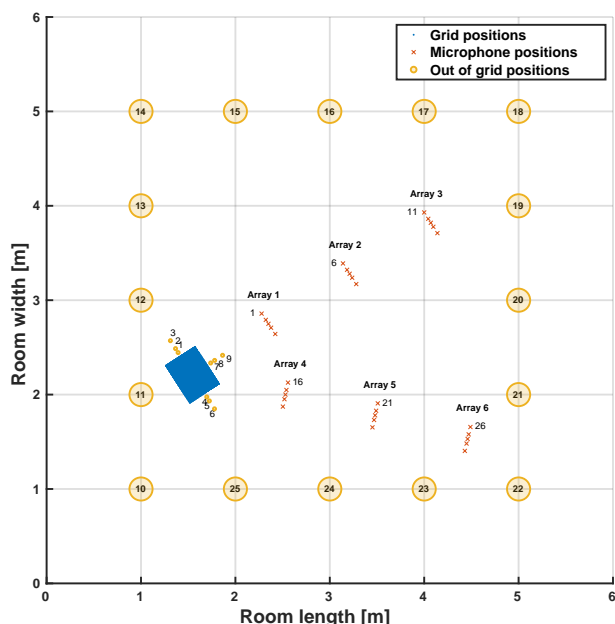
Nahrávání probíhalo v akustické místnosti ($6 \times 6 \times 2.4$ m; 100, 200 a 300 ms reverberační čas) na Univerzitě Bar Ilan v Izraeli. Vybavení v nahrávací místnosti bylo umístěno podle modelu zobrazeném na obrázku č. 1 a 2.

Polohování reproduktoru v rámci husté sítě (obrz. 2) bylo zajištěno upravenou konstrukcí 2D plotteru a ručně ovládaným zdvihacím stolem. Pozice byly nahrávány po jednotlivých řezech v z ose (rozestup 4 cm). Rozestup mezi pozicemi v osách x a y byl 2 cm. Ostatní pozice byly obsluhováno ručně.

Mikrofonní pole (o 5-ti mikrofonech) byla umístěna před sítí pozic ve vzdálenostech 1, 2 a 3 m pod úhly 0° a -45° (celkem 6 polí).

Nahrávání

Nahrávání probíhalo sekvenčně, tj. každá pozice zvlášť. Nahrávány byly dva typy budících signálů: bílý šum (všechny frekvence), logaritmický chirp (postupné zvyšování buzené frekvence v čase po logaritmické křivce od 200-16000Hz). Vzorkovací frekvence byla 48 kHz (formát WAV).



Obrázek 1: Rozložení nahrávací místnosti. Modře je znázorněna oblast, kde se nacházel přesně pozicovaný zdroj zvuku (hlavní výstup, 4104 pozic). Žlutě jsou znázorněny oblasti, kde se nacházel ručně pozicovaný zdroj zvuku (vedlejší výstup, 25 pozic). Červené křížky znázorňují umístění mikrofónů.

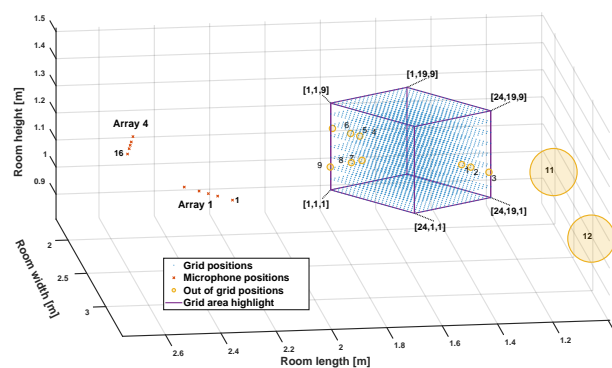
Výsledky a diskuze

Výsledkem této práce je zveřejněná strukturovaná databáze zvuků (celkem cca 750GB). Integrita databáze byla experimentálně ověřena.

Dále byl představen software umožňující zjednodušený výběr parametrů (pozice, metoda, ...) pro výpočet ATF.

Závěr

Článek představující databázi právě prochází recenzním řízením. Databáze je již využívána na ústavu ITE FM (skupina ASAP).



Obrázek 2: Rozložení nahrávací místnosti detailní pohled na síť pozic. Barevné značení je shodné se značením obrázku č. 1.

Poděkování

Poděkování patří panu Prof. Sharonu Gannotovi z Univerzity Bar Ilan (Ramat Gan, Izrael), který nám umožnil využít akustickou laboratoř. Práce vznikala v rámci výjezdu do Izraele podpořeném projektem Erasmus+ a Fondem mobilit Technické Univerzity v Liberci.

Tato práce byla podpořena z projektu Studentské grantové soutěže (SGS) na Technické univerzitě v Liberci v roce 2019.

Reference

- [1] HABETS, A. *Room impulse response generator*. 2006.: no. 2.4, vol. 2. Tech. Rep. Technische Universiteit Eindhoven.
- [2] STEWART, R. a M. SANDLER. Database of omnidirectional and b-format room impulse responses. *IEEE International Conference on Acoustics*. 2010, 165-168.
- [3] HADAD, E., F. HEESE, P. VARY a S. GANNOT. Multichannel audio database in various acoustic environments. *International Workshop on Acoustic Signal Enhancement*. 2014, 14., 313–317.