

Oponentský posudek na dizertační práci

Doktorand: Hasan KHADDOUR

Název práce: LOCALIZATION AND RENDERING OF SOUND SOURCES IN ACOUSTIC FIELDS

Oponent: Dr. Ing. Libor HUSNÍK

Předložená dizertační práce Ing. Hasana Khaddoura se zabývá oblastí lokalizace a polohování zdrojů zvuku. Obsahuje celkově 124 stránek, přičemž samotný text je dlouhý 92 stránek a skládá se ze sedmi kapitol. Je napsána v anglickém jazyce.

V první kapitole autor shrnuje současný stav v oblasti lokalizace a polohování zdrojů zvuku. Podává stručný přehled principů, na jakých člověk lokalizuje zdroje zvuku v horizontální rovině a dále metody, na kterých jsou založeny systémy pro lokalizaci zdrojů zvuku elektroakustickými systémy. Následuje část, kdy je podán výklad o možnostech polohování virtuálních zdrojů zvuku pomocí soustavy reproduktorů. Druhá kapitola se věnuje časově-frekvenčním transformacím a souvisejícím okénkům. V třetí kapitole jsou vytyčeny cíle dizertace. Ve čtvrté kapitole se začíná řešit jádro dizertace. Je v ní popsána simulace algoritmů a měření na systémech lokalizace zdrojů zvuku a také experimentální měření přesnosti lokalizace pomocí systémů Ambisonie a Vector Based Amplitude Panning. V páté kapitole je popsán a experimentálně otestován navržený systém pro odhad směru dopadu zvukové vlny při použití několika reproduktorů. Šestá kapitola se věnuje akustickému zoomu, konkrétně systému, který je schopen akustickému zoomu a odhadu směru zdroje zvuku. Tento systém je popsán a experimentálně ověřen. V sedmé kapitole je podáno shrnutí všech výsledků práce.

Aktuálnost

Zpracovaná tematika je aktuální, např. z hlediska použitelnosti při telekonferencích pro další vylepšení komfortu a v dalších aplikacích, kde se využívá efektů tzv. „prostorového zvuku“ při potřebě zvýraznění některé části signálu závisící na směru, např. v oblasti počítačových her a dalších simulací v asistivních technologiích. Z toho i vyplývá příslušnost tematiky dizertace oboru Teleinformatika.

Splnění cílů

Dizertant si vytyčil tyto cíle: vybrat nejvhodnější reprodukční systém pro akustický zoom, porovnat různé metody systémů pro lokalizaci zdrojů zvuku, navrhnout a ověřit novou metodu pro odhad směru zvukového zdroje, implementovat systém pro akustický zoom pomocí DirAC a dříve zmíněné metody pro odhad směru zvukového zdroje a ověřit navržený systém akustického zoomu pomocí subjektivních i objektivních testů. Dle předloženého textu je zřejmé, že tyto cíle byly beze zbytku splněny.

Přínos

Dizertant navrhl metodu odhadu příchodu směru zvukové vlny pomocí úpravy metody využívající signály B-format systému Ambisonie analýzou energie. Tuto metodu ověřil simulací i experimentálně. Dále navrhl a ověřil kombinovaný systém pro akustický zoom využívající též tuto metodu a tento systém ověřil poslechovými testy.

Publikační výstupy práce

V seznamu literatury je celkem 11 citací článků, které vyšly většinou v konferenčních sbornících, dvě z nich jsou citace časopisecké a tři internetového časopisu. Publikačních výstupů je tedy dostatečné množství a dizertant dle mého názoru prokázal schopnost publikovat výsledky své odborné práce.

Připomínky

V první kapitole jsou uvedeny principy, na jakých člověk lokalizuje zdroje zvuku. Jsou zde však bez další specifikace uvedeny pouze principy lokalizace v horizontální rovině. I když je pravda, že metody „prostorového zvuku“ se většinou orientují na horizontální rovinu, existují i systémy na zlepšení vjemu polohy zdroje ve výšce.

Práce by si zasloužila korekturu angličtiny. Příklady: state-of-the-art se nepojí s předložkou about, ... it also contains the sound reflectED (ne reflecting) to the ears... (z Kapitoly 1). 2.2: the longer the window we use (ne used) the lower (ne less) resolution we get (nebo got, pokud necháme v předešlém used). Popisovaný systém ozvučení (podstatné jméno) je AmbisonicS. Adjektivum ambisonic však možné (kap. 4.2). Kap.5: je rozdíl mezi approved a proved. Dále: Thus, we will use (ne used) it later.

Str. 5: *ILD je způsobeno útlumem během šíření [zvukové vlny] ve vzduchu.* Toto není úplně přesně vyjádřeno, na tak krátké vzdálenosti (ve volném poli zdroje, který není úplně blízko subjektu) je útlum na vzdálenosti menší než 30 cm zanedbatelný). ILD je způsobeno šířením kolem hlavy, vlnovými jevy a útlumem na překážce. Tedy ne pouze šířením vzduchem. Útlum 1kHz při 50% vlhkosti a 20C° je 0,005 dB/m. Amplituda vlny sice ubývá o 6dB se zdvojnásobením vzdálenosti od zdroje, ale to je pouze u sférické vlny (nutno specifikovat) a mnohdy nejsme tak blízko zdroje této sférické vlny, aby to bylo relevantní. Autor správně konstatuje v kapitole 1.1.2.2, že tohoto principu se dá využít pouze u blízkých zdrojů.

Vztah 1.35. překlep, ve jmenovateli v závorce se sčítá kvadrát impedance s kvadrátem akustické rychlosti.

2.1: *Time frequency transformations represent the transformed signal using...* Časově-frekvenční transformace nereprezentují signál. Jsou to prostředky na převod jedné reprezentace signálu na jinou. (Fourierova transformace nereprezentuje signál.)

Název kapitoly 5.2.3 *Simulation results in Three Dimensional Plane* a popis obrázku 5.9: zřejmě překlep plane/space.

Kap. 5.2.1 Akustická intenzita není vektor, který definuje směr energie. Energie nemůže mít směr. Pouze její tok může mít směr. A to je ta intenzita.

Názvy obrázků 6.1 a 6.2 mohou znamenat, že se jedná o jednotky pro zoom nebo lokalizaci nakreslené nebo existující ve dvourozměrné rovině (navíc rovina nemá jiný počet rozměrů).

Pro neznalého čtenáře není na první pohled snadné zjistit, kde přesně autor popisuje své výsledky a kde je přebírá. Třeba v kapitole 5, kde již čtenář předpokládá pouze výsledky autora, jsou opět zmiňovány známé pojmy (signály B formát), které mohly být začleněny do úvodních kapitol, a pak není úplně zřejmé, kde začíná autorův přínos. Je to zřejmě na konci kapitoly 5.2.1, podkapitola s názvem Estimation of the Direction of Arrival a následné simulace a experimenty.

Otázky k diskusi

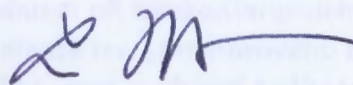
Doplňující otázka ke kapitole 1.1.1 : který psychoakustický jev pomáhá při detekci směru přímého zvuku (zdroj-místo posluchače) od odražených, přicházejících k posluchači z jiných směrů. Dal by se tento jev také nějak implementovat?

Dala by se navržená metoda jednoduše rozšířit i o parametr výšky zdroje pomocí využití signálu $z(t)$ ze sady signálů B formátu využitím stejných principů? O co by byla složitost algoritmu složitější?

Autor v samém závěru navrhuje uvážení možnosti implementace popsaného systému v reálném čase. Proč nebyla tato tematika zahrnuta do této práce?

Závěr: Předložená dizertační práce ing. Hasana Khaddoura s názvem Localization and rendering of sound sources in acoustic fields splňuje dle mého názoru zákon č. 11/1998 Sb. o vysokých školách a proto ji doporučuji k obhajobě a doporučuji dizertantovi udělit titul PhD.

V Praze, 5. května 2015



Dr. Ing. Libor Husník
ČVUT FEL