

V minulém čísle jsme otiskli první část článku o renderování prostorového zvuku. Jejím obsahem byl popis základních principů šíření zvuku v akustickém prostoru a prostorového slyšení.

Dnes se budeme věnovat metodám modelování zvukového pole. Nejdříve uvedeme několik algoritmů a pak pohovoříme o existujících technologiích a uvedeme některé aplikace.

Nejen grafika je 3D (2. část)

Metody modelování

Počátky modelování zvukových polí lze rozpoznat v době, kdy si vědci uvědomili, že nutnou podmínkou pro realistickou reprodukci zvuku je zaměstnat obě uši rozdílnými signály.

Podle počtu kanálů záznamu zvuku rozlišujeme záznamy na monofonní, stereofonní atd. Vzhledem k obsahu předchozích odstavců je evidentní, že pouhý monofonní signál není s to nést informaci o umístění zdroje zvuku. Ani prostý stereofonní signál, jehož zdroj může být umístěn více vlevo nebo vpravo, není dostatečně realistický. A přesto nám stačí pouze dvě uši, abychom vnímali prostor kolem sebe. Proto nutně musí také stačit dva kanály zvukového signálu, ve kterém však musí být přítomny veškeré výše popsané stopy. Takový signál se nazývá *binaurální*.

Existují však ještě metody, které použití sluchátek nepředpokládají. V první řadě se jedná o nejrůznější extendery stereofonního signálu, např. o *Spatializer* a *Qsound*, které marketing chybně nazývá "3D Sound" nebo "3D Stereo". Tyto metody spočívají v modifikaci již existujícího stereofonního signálu pomocí různých filtrů. Nejde však o přesné umístění zdrojů do prostoru, jen o jakési "zprostornění" zvukového vjemu, které se však úspěšně rozchází s fyzikální realitou.

Dále existují technologie založené na zvýšení počtu kanálů a reproduktorů, takže zvuk přichází z více zdrojů současně. To je tzv. *surround*, na němž jsou založeny technologie *Prologic* a *AC-3*. Jsou vhodné pro záznam a reprodukci zvukového signálu tam, kde je více posluchačů na různých místech – tj. např. v kinech. Popis těchto metod se však vymyká zaměření tohoto článku, proto se jim věnovat nebudeme.

Zaměříme se na metody generování binaurálního signálu. Většina takových metod se skládá ze dvou částí: z výpočtu impulzní odezvy akustického prostoru pro zadanou polohu posluchače a zdroje zvuku a z její následné aplikace na zvuk emitovaný zdrojem. Většinou předpokládáme bodový zdroj zvuku a homogenní prostředí, takže zdroj zvuku je vlastně zdroj "monofonního" signálu.

Věnujme se nyní výpočtu impulzní odezvy prostoru. Řada metod je založena na principech geometrické akustiky pracující se *zvukovými paprsky*, které se chovají velmi podobně jako paprsky

světelné: Mají zdroj, zachovávají směr vyslání až do jejich odrazu, jsou postupně utlumovány průchodem prostředím a odrazy a nesou určitou energii (kompletní monofonní signál). Dopadne-li takovýto odražený paprsek do ucha, slyšíme ozvěnu. Celková délka takového paprsku určí zpoždění příchodu signálu do ucha vzhledem k době vyslání.

Pro geometricky jednoduché scény je možné použít *metodu zrcadlových obrazů*. Princip této metody, běžně užívané např. i v teorii elektromagnetického pole, je znázorněn na **obr. 1**. Pro každou stěnu, která se vyskytuje ve scéně, zkonstruujeme zrcadlový obraz, kde osa zrcadlení prochází stěnou. V místě, kde paprsek z virtuálního zdroje prochází stěnou, spočítáme jeho utlumení, jako kdyby tam byl odraz. Nevýhoda této metody spočívá jednak v požadavku jednoduchosti modelované scény, jednak, a to hlavně, v exponenciální závislosti složitosti na požadovaném maximálním řádu odrazů v ozvěnách.

Další možností je pak *metoda trasování paprsku*. Je založena na pokusném vysílání paprsků ze zdroje – vypočítáváme jejich odrazy a útlumy a zjišťujeme, zda nezasáhnou model posluchačova ucha (**obr. 2**). Generátor paprsků však musí zajistit, aby tvořily stejnoměrné radiální pole. Buď jsou směry voleny náhodně (*Monte Carlo raytracing*), nebo např. postupně procházejí všemi vrcholy geodetické koule.

Obě tyto metody však neumožňují počítat s fenoménem ohybu zvuku. To částečně řeší *metoda trasování (obecného) kužele* (viz [2]). Ta předpokládá, že zdroj zvuku je množina kuželů vyplňující celý prostorový úhel. Metoda spočívá ve vyšetřování průniku podstavy kužele s překážkami ve scéně. Celý princip zobrazuje **obr. 3**. Vyskytne-li se posluchač uvnitř kužele, je započítána ozvěna. Ohyb zvuku je možné simulovat přidáním určité části k té části kužele, která pokračuje dále prostorem (na obrázku znázorněna modře).

Přesnější metody, i když náročnější na výpočetní výkon, jsou nejrůznější numerické aproximace, např. *metoda konečných prvků*.

Jak jsme řekli, druhou částí modelování prostorového zvuku je aplikace impulzní odezvy na signál vystupující ze zdroje. Toho docílíme *konvolucí* signálu s impulzní odezvou. Konvoluci nám přiblíží **obr. 4**, na kterém vidíme základní princip diskrétní konvoluce. Jedná se o proces velmi náročný na výpočetní výkon počítače vzhledem k tomu, že pro každý vzorek vstupního signálu je nutné provést velké množství aritmetických operací.

Na vstup konvoluce přicházejí vzorky vstupního signálu a vzorky impulzní odezvy. Pro každý vzorek impulzní odezvy se provede zesílení (zeslabení) vstupního signálu vynásobením všech jeho vzorků prvním vzorkem impulzní odezvy. Vzniklý signál je uložen do mezipaměti. Pak se stejná operace provede pro druhý vzorek impulzní odezvy, ale výsledný signál je uložen do mezipaměti posunutý o jedno paměťové místo. Podobně se pokračuje až do konce impulzní odezvy. Nakonec jsou sečteny všechny stejnohlé vzorky všech signálů v mezipaměti a jejich součet je signál vystupující z konvoluce.

Celým procesem je nutné projít dvakrát vzhledem k tomu, že nejdříve je aplikována impulzní odezva levého a pak pravého ucha.

Obsah předchozího textu shrnuje blokové schéma typického binaurálního reverberátoru na **obr. 5**. Model scény přibližně uprostřed obrázku přijímá informaci o poloze zdroje zvuku a posluchače a na jejím základě vypočítá impulzní odezvu pro levé a pravé ucho. Impulzní odezva spolu s monofonním signálem ze zdroje vstupují do příslušných konvolucí, ze kterých pak vystupují dva kanály stereofonního binaurálního signálu.

Výše popsané metody umožňují velmi kvalitní simulaci zvukového pole, nicméně proti nim hovoří jejich velká výpočetní náročnost, která prakticky vylučuje jejich použití v systémech požadujících výpočet prostorového zvuku v reálném čase (virtuální realita atd.). Proto bylo vyvinuto mnoho metod založených na podstatném zjednodušení metody výpočtu impulzní odezvy a konvoluce.

Je možné např. rozdělit scénu na určité objemové jednotky, pro každou vypočítat impulzní odezvu a výsledky uložit do databáze scény (tzv. celulární aproximace). Výpočet impulzní odezvy v daném bodě v reálném čase je pak redukován na nalezení nejbližšího bodu v databázi, pro kterou již impulzní odezvu známe, a na její načtení z databáze (**obr. 6**).

Další možností je najít pouze několik odrazů nižších řádů a ostatní aproximovat statistickým modelem. Je totiž známo, že délka dozvuku je úměrná objemu místnosti a obálka impulzní odezvy s časem exponenciálně klesá. Snížení náročnosti konvoluce je dosaženo např. zkrácením filtru IR (na obr. 4 by bylo méně řádků) nebo zmenšením vzorkovací frekvence vstupního signálu (méně sloupců).

Aplikace

Binaurální vnímání přináší celou řadu dalších schopností sluchu než jen schopnost lokalizovat zvukový zdroj v prostoru. V roce 1953 popsal E. C. Chery efekt koktejlového večírku (Cocktail-party effect): "Slyší-li člověk zvuk z několika nezávislých zdrojů najednou, je schopen je od sebe rozlišit." Posloucháme-li např. v recepci řeč jednoho člověka, rozumíme mu, i když kromě něj mluví v místnosti ještě osm dalších lidí. Když však ze stejného místa pořizujeme monofonní nahrávku, přijdeme o možnost řečníky od sebe rozlišit.

Renderování zvukových polí umožňuje řešit velmi mnoho zajímavých a důležitých úloh. Mnoho jich spadá do oblasti architektury, např. návrh koncertních sálů a poslucháren. Dnes již existuje řada síní postavených podle výsledků, které tato technologie poskytla. Jako příklad uveďme systém *Ramsete*, který navrhl prof. Andreo Farina z univerzity v italské Parmě. Tento systém umožňuje interaktivně simulovat šíření zvuku v uzavřených prostorech, jako jsou divadelní sály, tovární haly apod.

Tentýž autor navrhl metodu simulace šíření městského hluku. S využitím jeho metody je např. možné efektivně navrhovat nejrůznější akustické bariéry kolem silnic.

Hudební studia využívají poznatků prostorové akustiky k vytváření nejrůznějších umělých dozvuků již mnoho let. Zvukový snímek je pořízen v hudebním studiu, které má pokud možno nulový dozvuk, a je poté tzv. *nahalen*. Posлуhač pak má pocit, že se nahrávalo např. v kostele. Teprve však výpočetní technika přinesla větší flexibilitu a vyšší kvalitu takových systémů.

Přirozené jsou aplikace prostorového zvuku ve virtuální realitě a v počítačových hrách. V nich je schopnost zvukového systému poskytovat výsledky v reálném čase nutnou podmínkou použitelnosti.

Nicméně i pro podstatně závažnější obory lidské činnosti znamená aplikace prostorového zvuku přínos. Jedná se o nejrůznější profese, jejichž hlavní náplní je rychlé řešení množiny očekávatelných problémů s málo predikovatelným výskytem. Máme zde na mysli piloty letadel, dispečery letového provozu, zaměstnance elektráren, lékaře atd.

Vyskytne-li se problém, je zvykem nejrůznějších kontrolních panelů zapnout bzučák a rozblikat příslušnou kontrolku. Odpovědná osoba vypne zvukový alarm a hledá v záplavě svítících kontrolků tu, která indikuje závadu. Kdyby však byl zmíněný zvukový signál vyslán pomocí systému generujícího prostorový zvuk, nebyl by problém zaměřit se přímo na oblast kontrolního panelu, kde se vyskytl

problém. Pilotovi stíhacího letounu by zvukové upozornění na špatný stav motoru přišlo z míst, kde se skutečně motor vyskytuje. Navíc by mohl přijmout několik takových upozornění najednou vzhledem k tomu, že by každé bylo lokalizováno jinde.

Technologie

Současný vývoj již přiměl technologie opustit stadiu počátečních experimentů a objevují se první standardy. Firma Aureal Semiconductor specifikovala standard *A3D Interactive*, popisující soubor vlastností a funkcí, které by měl zvukový systém splňovat a podporovat (např. maximální počet zvukových zdrojů, složitost scény, Dopplerův efekt apod.), a zároveň vyvinula technologii, která tento standard implementuje.

Jedná se o zvukový engine určený pro osobní počítače, který renderuje třírozměrný zvuk v reálném čase v závislosti na zadané geometrii scény (seznam stěn a jejich fyzikálních vlastností) a polohách posluchače a zdrojů zvuku.

Existují dva hlavní typy implementací – softwarovými knihovnamí a hardwarově v podobě čipu, který je integrován ve zvukové kartě. Hardwarová implementace přináší velké odlehčení procesoru, který tak nemusí provádět výpočet impulzní odezvy a konvoluci. Softwarové knihovny mají tu výhodu, že pro jejich funkci plně postačuje obyčejná zvuková karta, ale většinou jsou nějakým způsobem omezeny (např. neprovádí výpočet odrazů).

V současnosti již existuje na trhu několik typů zvukových karet, které standard A3D podporují. Jedná se např. o výrobky firem Diamond Multimedia (řada Monster Sound), Turtle Beach (Mondego A3D), TerraTec Promedia (XLerate).

Z her jmenujme např. *Descent: FreeSpace* (Interplay), *Half-life* (Sierra), *Blood 2* (patche od firmy Monolith), *Duke Nukem Forever* (GT Interactive), *Jedi Knight* (LucasArts) nebo *Quake II* (Activision).

Tímto odstavcem uzavíráme základní přehled termínů a technologií týkajících se prostorového zvuku. Článek jsme považovali za vhodné vydat vzhledem k velmi rychlému vývoji této oblasti. Doplnuje jej soubor ukázek binaurálních zvukových signálů uložených na Chip CD.

Závěrem by autor rád poděkoval docentu Pavlu Slavíkovi z katedry počítačů Fakulty elektrotechnické ČVUT za mnoho cenných rad.

Adam J. Sporka (sporkaa@cs.felk.cvut.cz)

Reference

- [1] F. Kolmer, J. Kyncl – Prostorová akustika.
- [2] A. Farina, P. Galaverna, G. Truffelli – “Ramsete” un nuovo software per la previsione del campo sonoro in teatri, ambienti industriali ed ambiente esterno.
- [3] T. Funkhouser, I. Carlbom, G. Elko, G. Pingali, M. Sondhi, J. West – A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments, SIGGRAPH '98.
- [4] E. C. Cherry – Some experiments on the recognition of speech with one or two ears (1953).

[5] J. Žára, B. Beneš, P. Felkel – Moderní počítačová grafika.