

SYNTÉZA AUDIO SIGNÁLŮ

R. Čmejla

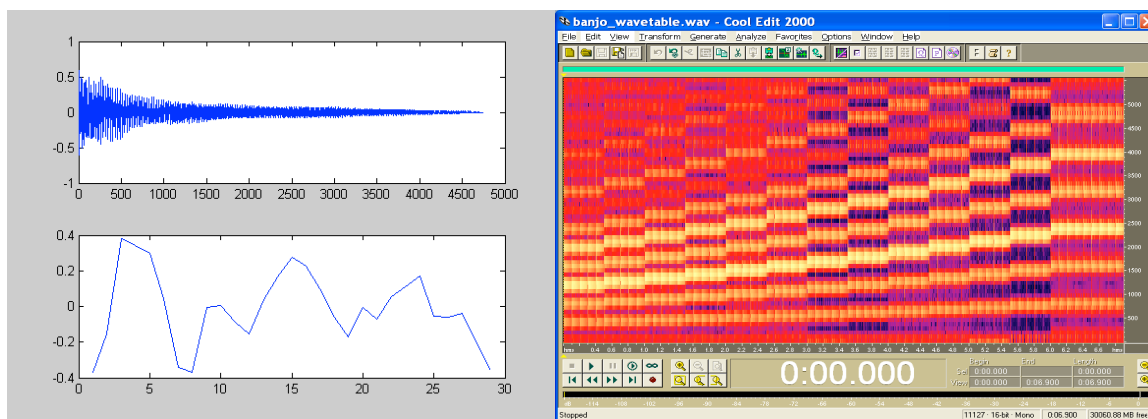
Fakulta elektrotechnická, ČVUT v Praze

Abstrakt

Příspěvek pojednává o technikách číslicové audio syntézy vyučovaných v předmětu Syntéza multimediálních signálů na Elektrotechnické fakultě ČVUT v Praze. Teoretické koncepty z přednášek jsou ve cvičeních doplněny praktickým programováním úloh v MATLABu.

1 Tabulková syntéza

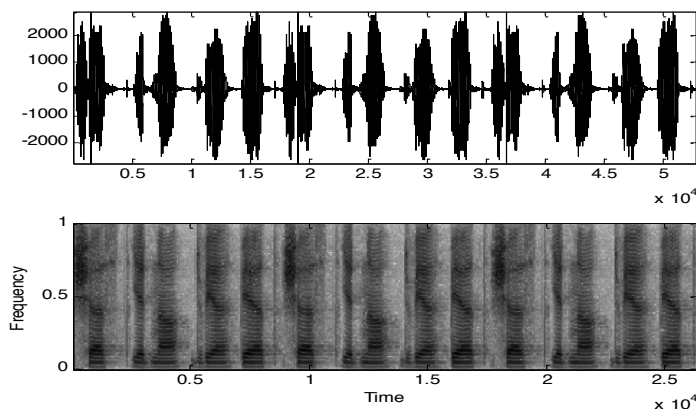
Tabulková syntéza, která je jednou z nejpoužívanějších technik, používá vzorkovaný záznam skutečného nástroje. Studenti v prostředí MATLAB pomocí lineární interpolace transformují harmonické průběhy na libovolnou frekvenci a prodlužují na libovolnou délku, dále generují uvedenou technikou zvuk banja a dalších hudebních nástrojů, viz obr. 1.



Obrázek 1: Tabulková syntéza: reálný signál banja – vlevo nahoře, vyjmutí jedné periody reálného signálu – vlevo dole, periodogram syntetického signálu – vpravo

2 Konkatenční syntéza

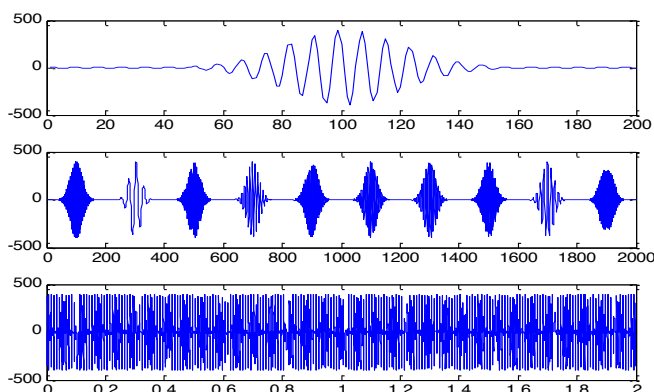
Konkatenční syntéza je základem většiny současných syntezátorů řeči. Signál se vytváří řetězením (konkatenací) akustických elementů uložených v databázi řečových jednotek. Princip syntézy je demonstrován na jednoduchém spojování slov v prostředí MATLAB, viz obr. 2.



Obrázek 2: Časový průběh a spektrogram řetězeného signálu

3 Granulační syntéza

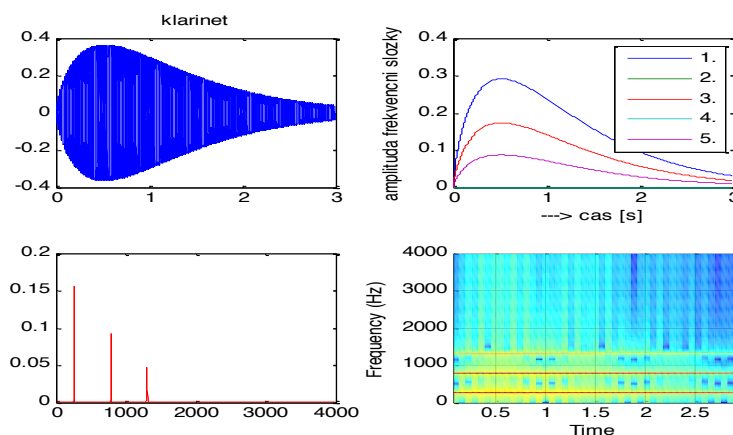
Granulační syntéza pracuje s velmi malými zvukovými elementy (v délce od 1 do 50 ms), které při rychlém přehrávání vnímáme jako spojitý zvuk. Studenti vytváří řadu umělých zvuků typu "sci-fi" změnou parametrů u připravených algoritmů v časové i ve spektrální oblasti, viz obr. 3.



Obrázek 3: Syntéza signálu ze zvukových elementů

4 Součtové syntézy

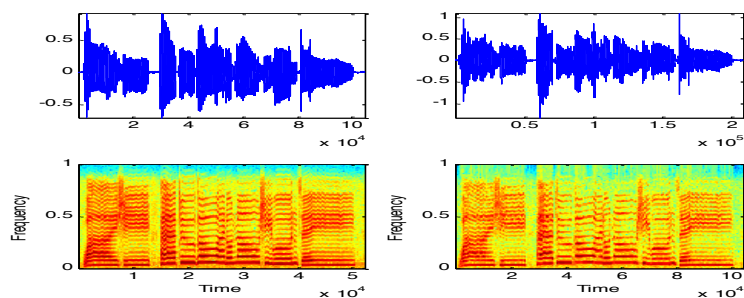
Součtové syntézy představují zobecnění konceptu Fourierových řad, při kterém se sčítají časově proměnné parciály s různou fází, amplitudou a frekvencí. Ve cvičení studenti provádí syntézu základních periodických průběhů v prostředí MATLAB a sledují vliv fáze na generovaný audio signál. Součtová syntéza banja, zvonku, zvonů, klarinetu a drknutí struny je realizována na základě resyntézy s využitím spektrogramů reálných signálů, viz obr. 4.



Obrázek 4: Součtová syntéza signálu klarinetu:
časový průběh syntetického signálu – *vlevo nahoře*
časový průběh obálek harmonických – *vpravo nahoře*
periodogram syntetického signálu – *vlevo dole*
spektrogram syntetického signálu – *vpravo dole*

5 Fázový vokodér

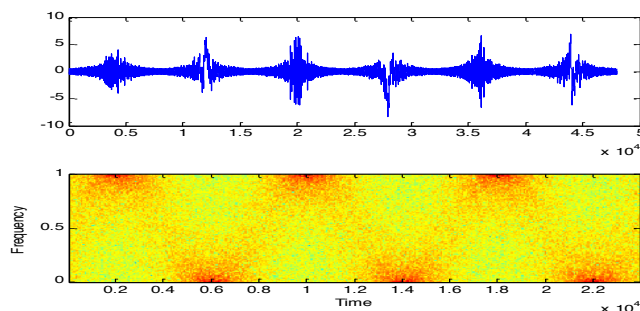
Fázový vokodér je založen na přímé a na inverzní krátkodobé Fourierově transformaci. Ve cvičení studenti nejprve realizují resyntézu audio signálů pomocí přímé a zpětné Fourierovy transformace, libovolně mění délku signálu (při zachování frekvenčního obsahu spektra), realizují frekvenční posunutí (při zachování délky signálu) a realizují audio efekty „robotizace“ a „whisperizace“, viz obr. 5.



Obrázek 5: Použití fázového vokodéru pro prodloužení délky signálu při zachování původního frekvenčního obsahu: původní signál – *vlevo*, 2x prodloužený signál – *vpravo*

6 Rozdílová syntéza

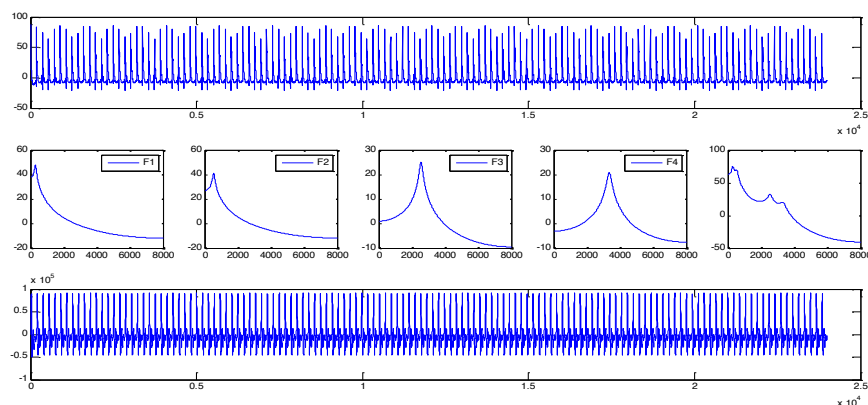
Rozdílová syntéza je jednou z nejstarších stále používaných technik a je založena na průchodu spektrálně bohatého budícího signálu filtrem. Při praktickém ověření v MATLABu studenti realizují syntézy perkusních nástrojů, tryskového letadla, parních strojů, výstřelu, hodin, větru, vln a frikativních souhlásek se šumovým buzením. S použitím pulzního buzení lze realizovat syntetické výstupy odpovídající žesťům, klarinetu (obdélník) a smyčcovým nástrojům (pila), viz obr. 6.



Obrázek 6: Syntetický signál vzniklý průchodem bílého šumu IIR filtrem 1.řádu s proměnným koeficientem

7 Formantová syntéza

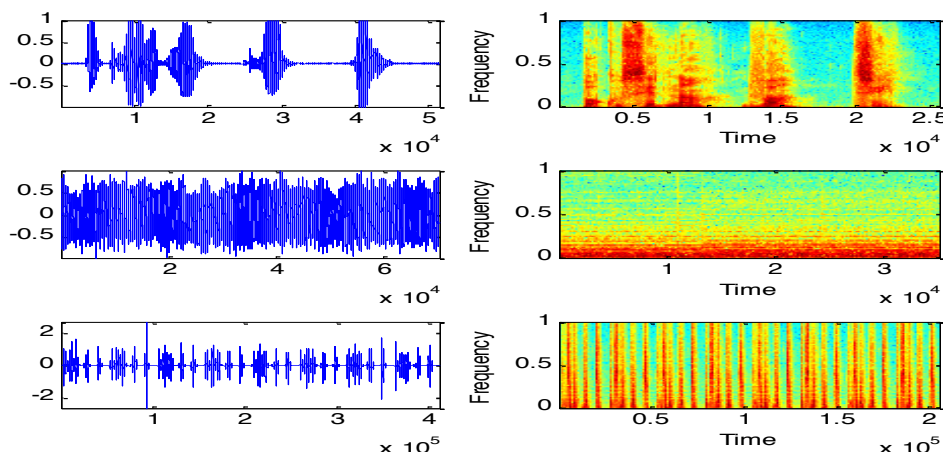
Formantová syntéza je založena na akustické teorii vytváření řeči a na zjednodušeném modelování hlasového traktu pomocí formantů. Ve cvičení studenti nejprve pomocí formantové syntézy generují samohlásky (viz obr. 7) a poté pomocí principů formantové syntézy modelují hudební nástroje, kdy budící signál je tvarován rezonátory.



Obrázek 7: Kaskádní syntéza samohlásek: budící signál – *nahoře*, frekvenční charakteristiky formantových filtrů – *uprostřed*, syntetická samohláska – *dole*

8 LPC syntéza

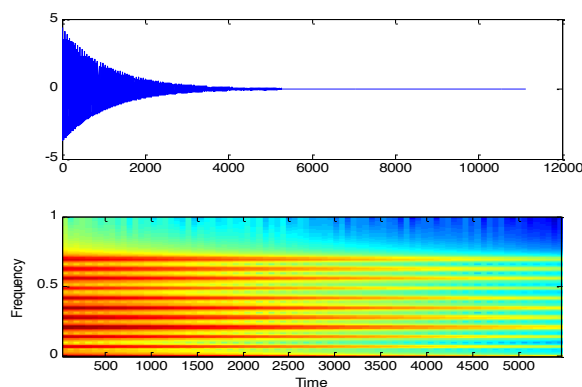
LPC syntéza představuje typ rozdílové syntézy u které výstupní audio signál je modelován pomocí lineární predikce a časově proměnné filtry jsou nastavovány podle spektrální obálky reálných signálů. Ve cvičení studenti realizují LPC vokodér a sledují jeho výstup při různém typu buzení, realizují časové a frekvenční změny audio signálů a experimentují se vzájemnou syntézou různých typů audio signálů, kdy aplikují spektrální obálku jednoho signálu na druhý, čímž vytváří např. zpívající hudební nástroje, viz obr. 8.



Obrázek 8: Časové průběhy a spektrogramy pro:
řeč – *nahoře*, hudbu – *uprostřed*,
vzájemnou syntézu řeči a hudby – *dole*

9 Tvarovací syntéza

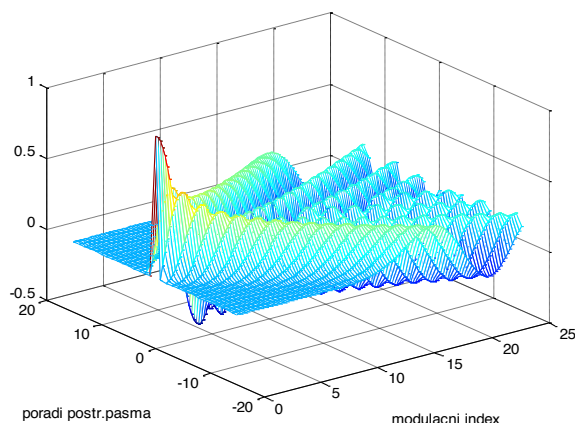
Tvarovací syntéza se řadí mezi metody “nelineárního zkreslení”. Představuje modifikaci audio signálu nelineární přenosovou funkcí. Studenti si ve cvičení ověřují vliv různých nelinearit na tvar spektra a pomocí Čebyševových polynomů generují signál s přesně specifikovaným frekvenčním obsahem, který odpovídá vybranému hudebnímu nástroji, např. banju, viz obr. 9.



Obrázek 9: Syntetický signál banja vygenerovaný pomocí Čebyševových polynomů

10 Frekvenční modulace

Frekvenční modulace získala v audio syntéze velikého rozšíření, neboť na rozdíl od ostatních syntéz nepotřebuje ke generování signálu velké množství parametrů. Charakteristická frekvenční pásma se vytváří v závislosti na poměrech nosné, modulační frekvence, hloubky modulace a tvaru signálu. Ve cvičení probíhá praktická realizace modulačních efektů tremola, vibráta a dále různých FM syntetických signálů (zvony, klepání na různé materiály, žestě, klarinet, alarm), viz obr. 10.



Obrázek 10: Besselovy funkce zobrazující vztah mezi amplitudovým spektrem signálu a modulačním indexem

11 Audio efekty

Audio efekty, které studenti realizují ve cvičení, jsou založeny na časovém zpoždění realizovaném prostřednictvím konvoluce, hřebenových filtrů a fázovacích článků, kdy malé zpoždění přináší oživení a rozjasnění zvuku a zpoždění delší než 50 ms je vnímáno jako echo, viz tab. 1.

Tabulka 1: AUDIO EFEKTY ZALOŽENÉ NA ČASOVÉM ZPOŽDĚNÍ (DOZVUKY A ECHA)

<i>Zvukový efekt</i>	<i>zpoždění</i>	<i>koeficient</i>
pod mostem	0,400	0,30
v chrámu	0,250	0,30
elektronický umělý dozvuk	0,200	0,90
klasické echo	0,150	0,50
v podzemní chodbě	0,120	0,70
v koncertní síni	0,100	0,40
elektronický efekt	0,085	0,90
ve sprše	0,030	0,60
v malé místnosti	0,010	0,50
mikrofonní zpětná vazba	0,001	0,97

12 Fyzikální modelování

Fyzikální modelování představuje audio syntézu, která pomocí matematického modelu - soustavy rovnic a algoritmů - simuluje zdroj zvuku. Ve cvičení studenti realizují základní Karplusův-Strongův algoritmus kytary, vycházejícího z fyzikálního modelu struny, upraveného pro zadávání vstupních fyzikálních parametrů jako jsou síla a místo drnknutí a také tuhost strun a tlumení na kobylce, viz tab. 2.

Tabulka 2: PŘÍKLAD VSTUPNÍCH PARAMETRŮ PŘI FYZIKÁLNÍM MODELOVÁNÍ KYTARY

<i>délka tónu</i>	<i>1500 ms</i>
<i>Frekvence</i>	<i>110 Hz</i>
<i>místo drnknutí</i>	<i>38 %</i>
<i>poč.amplituda</i>	<i>100 %</i>
<i>útlum kobylky</i>	<i>5 Hz</i>
<i>tuhost strun</i>	<i>5 %</i>

13 Závěr

Absolventi výše popsaného kurzu se seznamují s technikami syntézy multimediálních signálů, jejichž principy se používají v moderních číslicových systémech, systémech virtuální reality, počítačových animacích, hrách a ve filmu. Předmět je zakončen zkouškou a také semestrální prací, ve které studenti realizují syntetický audio signál pomocí vybraných technik výhradně v prostředí MATLAB. Výstupní audio signály ve formátu wav lze nalézt na stránce:

<http://sami.fel.cvut.cz/sms/semestralky>.

Poděkování

Tato práce byla podporována z grantů GAČR 102/12/2230 a SGS12/185/OHK4/3T/13.

Roman Čmejla
cmejla@fel.cvut.cz