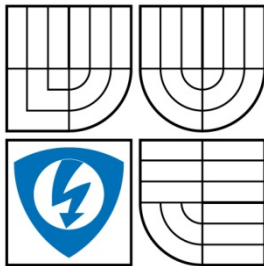


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

SNÍMÁNÍ A ZPRACOVÁNÍ AKUSTICKÝCH SIGNÁLŮ TECHNOLOGIÍ B-FORMAT

ACOUSTIC SIGNAL RECORDING AND PROCESSING USING B-FORMAT TECHNOLOGY

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

DAVID KURC

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. JIŘÍ SCHIMMEL, Ph.D.

BRNO 2009



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav telekomunikací

Bakalářská práce

bakalářský studijní obor
Teleinformatika

Student: David Kurc

ID: 72938

Ročník: 3

Akademický rok: 2008/2009

NÁZEV TÉMATU:

Snímání a zpracování akustických signálů technologií B Format

POKYNY PRO VYPRACOVÁNÍ:

Prostudujte technologii snímání zvukových signálů pomocí tří gradientních a jednoho všesměrového mikrofону označovanou SoundField, její princip a možnosti zpracování výstupních signálů této technologie označovaných B-Format. Popište algoritmy konverze B formátu do běžných vícekanálových zvukových formátů a otestujte je a implementujte v prostředí Matlab jako aplikaci s vlastním grafickým rozhraním umožňujícím také zobrazení energie signálu jednotlivých reprodukčních kanálů po konverzi v závislosti na azimutu a elevačním úhlu virtuálního zdroje zvuku. Vytvořte konstrukci mikrofonního systému Soundfield pomocí dostupných mikrofonních kapslí a pořídte testovací zvukové nahrávky.

DOPORUČENÁ LITERATURA:

- [1] Zwicker, E., Fastl, H. Psychoacoustics, Facts and Models, 2nd updated Ed. Springer, 1999. ISBN 3-540-65063-6
- [2] Rumsey, F. Spatial Audio. Focal Press, 2005. ISBN 0-240-51623-0.
- [3] Streicher, R., Everest, F., A. The New Stereo Soundbook, 3rd ed. Audio Engineering Associates, 2006. ISBN 978-0-9665162-1-0
- [4] Doňar, B., Zaplatílek, K. MATLAB pro začátečníky, 1. díl. Ben, technická literatura, Praha. ISBN 80-7300-175-6

Termín zadání: 9.2.2009

Termín odevzdání: 2.6.2009

Vedoucí práce: Ing. Jiří Schimmel, Ph.D.

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

ABSTRAKT

Tato bakalářská práce se zaměřuje na systém snímání a zpracování akustických signálů nazvaný Ambisonie. Popisuje jednotlivé fáze v řetězci zpracování akustického signálu tímto systémem. Zejména se zabývá technologií B-format, jakožto hlavním prostředkem ambisonie, jeho získáváním ze signálů snímaných speciální mikrofonní kapslí typu soundfield nebo pomocí kódovacích rovnic a jeho dekodováním do jednotlivých kanálů reproduktorů různých reproduktorových polí. Zmíněno je i použití přenosového formátu UHJ. V prostředí MATLAB byl vytvořen program B-format Dekodér, složený z několika funkcí, s vlastním grafickým uživatelským rozhraním, který je určen ke zpracování a dekodování složek B-formátu pro reproduktorové pole se zadanými parametry. Zajišťuje mimo jiné načtení, zobrazení, analýzu a přehrávání zdrojových a výsledných akustických signálů.

KLÍČOVÁ SLOVA

Ambisonie, B-format, soundfield, akustický signál, dekodér, reproduktorové pole, UHJ, azimut, elevace.

ABSTRACT

This bachelor's thesis is focused on the system of recording and processing of acoustics signals called Ambisonics. It describes single sections of this system in the chain of acoustic signal processing. At most it is aimed at the B-format technology as the main ambisonic instrument, its encoding from the signals recorded with the special microphone capsule of the soundfield type or with help of encoding equations and its decoding into single speaker channels of various speaker arrays. The UHJ transmission format is also mentioned. The application B-format Decoder was created in the MATLAB programming environment. It consists of several functions and has friendly user interface, which provides the processing and decoding of B-format components for the speaker array with entered parameters. Among others this interface ensures loading, displaying, analysis and playback of source and resulting acoustic signals.

KEYWORDS

Ambisonics, B-format, soundfield, acoustic signal, decoder, speaker array, UHJ, azimuth, elevation.

KURC, D. *Snímání a zpracování akustických signálů technologií B Format*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 64 s., 2 přílohy. Vedoucí bakalářské práce Ing. Jiří Schimmel, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou bakalářskou práci na téma „Snímání a zpracování akustických signálů technologií B-format“ jsem vypracoval samostatně pod vedením vedoucího bakalářské práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené bakalářské práce dále prohlašuji, že v souvislosti s vytvořením této bakalářské práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení §11 a následujícího autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení §152 trestního zákona č. 140/1961 Sb.

V Brně dne 26.5.2009

.....
David Kurc

PODĚKOVÁNÍ

Děkuji vedoucímu bakalářské práce panu Ing. Jiřímu Schimmelovi, Ph.D. za velmi užitečnou metodickou pomoc a cenné rady při zpracování mé práce.

V Brně dne 26.5.2009

.....
David Kurc

OBSAH

SEZNAM OBRÁZKŮ	8
ÚVOD	9
1 HISTORIE VZNIKU AMBISONIE	10
2 AMBISONIE	12
3 KÓDOVÁNÍ	14
3.1 SNÍMÁNÍ POMOCÍ MIKROFONŮ	15
3.2 KONVERZE MONOFONNÍHO ZVUKU DO B-FORMÁTU	17
4 FORMÁT UHJ	19
5 DEKÓDOVÁNÍ	21
5.1 PSYCHOAKUSTICKÁ KRITÉRIA LIDSKÉHO SLUCHU V SYSTÉMECH SURROUND SOUND	24
5.1.1 <i>Nízkofrekvenční lokalizace</i>	24
5.1.2 <i>Lokalizace na středních a vyšších kmitočtech</i>	25
5.1.3 <i>Vysokofrekvenční lokalizace</i>	25
5.1.4 <i>Lokalizace pomocí dozvuku</i>	25
5.2 TYPY REPRODUKTOROVÝCH POLÍ	26
6 PERIFONIE	28
7 VÝHODY AMBISONIE	29
8 OMEZENÍ	30
9 FORMÁTY AMBISONIE	31
10 APLIKACE B-FORMAT DEKODÉR	32
10.1 STRUKTURA APLIKACE	32
10.1.1 <i>Výpočty výstupních signálů jednotlivých reproduktorových polí</i>	38
10.2 TYPY ZOBRAZENÍ SIGNÁLŮ V APLIKACI B-FORMAT DEKODÉR	48
11 ZÁVĚR	55
SEZNAM LITERATURY	56
SEZNAM ZKRATEK A SYMBOLŮ	58
SEZNAM PŘÍLOH	60

SEZNAM OBRÁZKŮ

Obr. 1.1: Rozmístění reproduktorů při kvadrofonní reprodukci.....	10
Obr. 1.2: Rozmístění reproduktorů v systému 5.1 podle ITU-R BS 775-1.....	11
Obr. 1.3: Logo ambisonie znázorňující čtyři koincidenční mikrofony	11
Obr. 2.1: Cesta signálu základním ambisonickým řetězcem.	13
Obr. 3.1: Směrové charakteristiky mikrofonů reprezentující jednotlivé složky B-formátu	15
Obr. 3.2: Kapsle mikrofonu soundfield se čtyřmi téměř koincidenčními mikrofony.	16
Obr. 4.1: Možnosti rozšíření a kompatibility přenosu jednotlivých kanálů formátu UHJ.....	20
Obr. 5.1: Šestiúhelníkové reproduktorové pole s vyznačenými zdánlivými obrazy.....	22
Obr. 5.2: Příklad ambisonického dekodéru.....	22
Obr. 5.3: a) Hlava posluchače s pomyslnou osou, na které jsou umístěny obě uši, b) Dva mikrofony v pozicích uší nahraditelné jedním tlakovým mikrofonem.....	25
Obr. 5.4: Příklad trojrozměrného reproduktorového pole s dvanácti reproduktory.....	26
Obr. 5.5: Příklad n-stěnných reproduktorových polí	27
Obr. 6.1: Nejjednodušší trojrozměrné reproduktorové pole (tetrahedron)	28
Obr. 7.1: Stereo systém s vyznačením bodu „sweet spot“ ve vrcholu trojúhelníka.....	29
Obr. 10.1: Základní struktura aplikace B-format Dekoder	33
Obr. 10.2: Mono systém.....	39
Obr. 10.3: Stereo systém.	40
Obr. 10.4: Rectangular (obdélník).....	41
Obr. 10.5: Rectangular 3D (kvádr).....	42
Obr. 10.6: Birectangle.....	43
Obr. 10.7: Tetrahedron (čtyřstěn)	44
Obr. 10.8: Octahedron (osmistěn).....	45
Obr. 10.9: Pravidelný polygon (UHJ).....	46
Obr. 10.10: Pravidelný polygon 3D (UHJ).....	47
Obr. 10.11: Náhled časových průběhů testovacích složek B-formátu.....	48
Obr. 10.12: Náhled časových průběhů výstupních signálů Pravidelného polygonu.....	49
Obr. 10.13: Náhled časových průběhů výstup. signálů pole 5.1 Surround Sound ITU-R BS.775... 50	50
Obr. 10.14: Náhled spektrogramů signálů pole 5.1 Surround Sound ITU-R BS.775	50
Obr. 10.15: Náhled závislostí intenzity signálů reproduktorů Pravidelného polygonu 3D s osmi reproduktory (a vertikálním úhlem 60°) na azimutu zdroje	51
Obr. 10.16: Náhled závislostí intenzity signálů reproduktorů 5.1 Surround Sound ITU-R BS.775 na azimutu zdroje	52
Obr. 10.17: Náhled 2D závislostí intenzity signálů reproduktorů pole Rectangular 3D (horizontální i vertikální úhel 60°) na azimutu a elevaci zdroje zvuku.....	53
Obr. 10.18: Náhled 2D závislostí intenzity signálů 5.1 Surround Sound ITU-R BS.775 na azimutu a elevaci zdroje zvuku.....	53
Obr. 10.19: Náhled 3D závislostí intenzity signálů 5.1 Surround Sound ITU-R BS.775 na azimutu a elevaci zdroje zvuku.....	54

ÚVOD

Technologie B-format je hlavním prostředkem systému snímání a zpracování akustických signálů, který se nazývá Ambisonie. V dalším textu bude krátce přiblížen vývoj, který vedl k jejímu vzniku, budou popsány její vlastnosti, výhody a omezení. Jednotlivé fáze v řetězci zpracování signálu ambisonickou technologií jsou postupně probírány v příslušných kapitolách. Jedná se tedy především o vznik signálů B-formátu, jejich přenos a interpretaci konečnému posluchači. Součástí kapitol je také srovnání této relativně nové technologie s dnes nebo dříve populárními systémy reprodukce zvuku, jako například mono, stereo nebo 5.1 Surround Sound. Zmíním se také o dalších alternativních formátech, které mohou být ambisonií používány namísto standardního B-formátu.

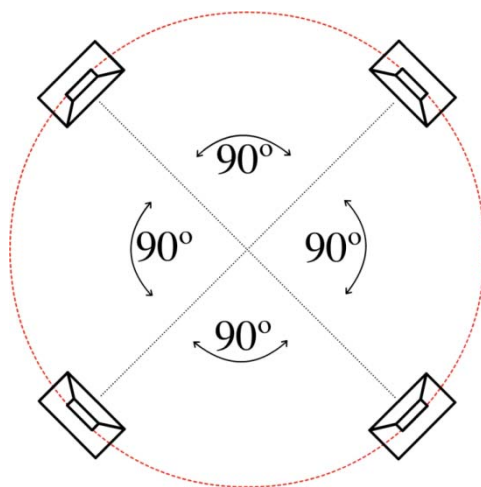
Dalším úkolem této práce bylo vytvořit uživatelskou aplikaci v prostředí MATLAB s vlastním grafickým rozhraním, která je schopna zpracovávat signály B-formátu. Jedná se vlastně o ambisonický dekodér, který dekoduje zdrojové složky B-formátu pro konkrétní soustavu reproduktorů (reproduktorové pole se zadanými parametry) určenou k reprodukci prostorového (ale i mono nebo stereo) zvuku. Aplikace měla také mít možnost zobrazit závislosti výstupních signálů (respektive jejich energie) na azimutu a elevaci zdroje zvuku. Tato zobrazení společně s dalšími (např. časovými průběhy) jsou dostupná v grafickém rozhraní. Funkce a struktura programu jsou popsány v kapitole 10, způsob jeho ovládání pak v příloze A.

K poslednímu úkolu, tedy pořízení testovacích zvukových nahrávek s mikrofonem soundfield sestaveným z dostupných mikrofonních kapslí, bohužel nedošlo, protože mikrofonní systém ještě nebyl konstrukčně přizpůsoben pro závažnější manipulaci v terénu.

1 HISTORIE VZNIKU AMBISONIE

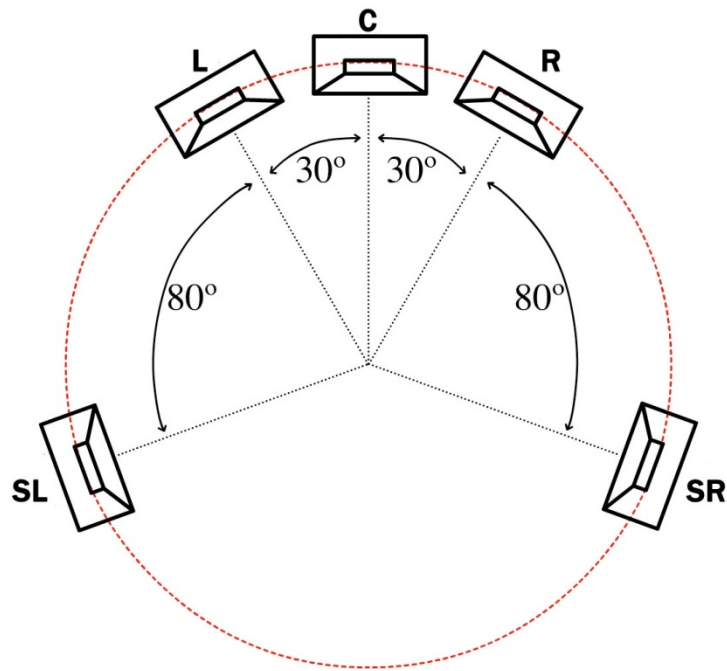
Zvukové reprodukční systémy se dnes objevují prakticky v každé domácnosti, ať už ve formě počítačů, televize, mobilních telefonů, MP3 přehrávačů, hi-fi soustav a v neposlední řadě například i systémů domácího kina s technologiemi surround sound. Všechny tyto prostředky lze shrnout do tří kategorií: jednokanálová (mono), dvoukanálová (stereo) a vícekanálová (surround sound) zařízení. Vývoj začal před více než stoletím mono systémy, pokračoval přes stereo až k nynějším, stále častěji se objevujícím, systémům multikanálovým (vícekanálovým). Primárním účelem surroundových systémů je vyprodukovat v uších a mozku posluchače iluzi vzoru zvukového pole, které obsahuje zvuky v různých směrech a vzdálenostech od posluchače. Prvním systémem využívajícím více než dvou reproduktorů byla kvadrofonie se čtyřmi reproduktory kolem posluchače [14]. Technika reprodukce byla založena na jednoduchém panorámování intenzit, podobně jako stereofonie. Tím se daly vytvářet zdánlivé zvukové obrazy mezi jednotlivými reproduktory [16]. Kromě interaurální intenzitní difference ale sluch k lokalizaci zdroje zvuku používá i další mechanismy. V případě že tyto mechanismy nejsou brány v potaz při vytváření imaginárního zvukového obrazu, výsledky bývají nevyhovující.

U kvadrofonie, kde je úhel mezi předními reproduktory více než 60° (obr. 1.1), se projevuje nestabilita zvukového obrazu spolu s efektem jakési „díry uprostřed“. Schopnost lokalizace zdrojů vzadu je také poměrně slabá, po stranách téměř žádná, celkově je ještě navíc citlivá na postavení reproduktorů a pozici posluchače [11].



Obr. 1.1: Rozmístění reproduktorů při kvadrofonní reprodukci [4].

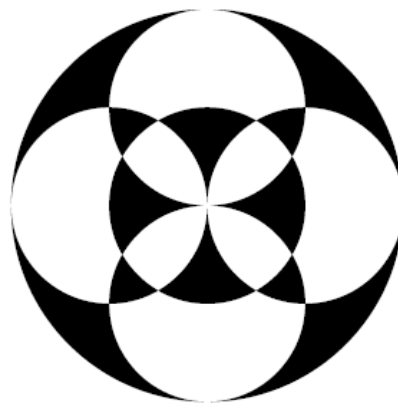
Následující systémy se snažily eliminovat tyto problémy a jsou používány dodnes (například 5.1 Surround Sound). Ke zlepšení stability předního zvukového obrazu využívají třetí, střední reproduktor. Příkladné surroundové reproduktory potom reprodukují zvuk okolního prostředí (obr. 1.2).



Obr. 1.2: Rozmístění reproduktorů v systému 5.1 podle ITU-R BS 775-1 [4].

Dalším pokusem bylo věrné přenesení komplexního snímaného zvukového pole do prostoru, kde docházelo k jeho reprodukci. To je ovšem z fyzikálního hlediska téměř nemožné z důvodu potřebného množství elektroakustických měničů. Pro reprodukci zvukového pole o průměru dva metry pro frekvence do 20 kHz bychom potřebovali 400 000 kanálů a reproduktorů [13].

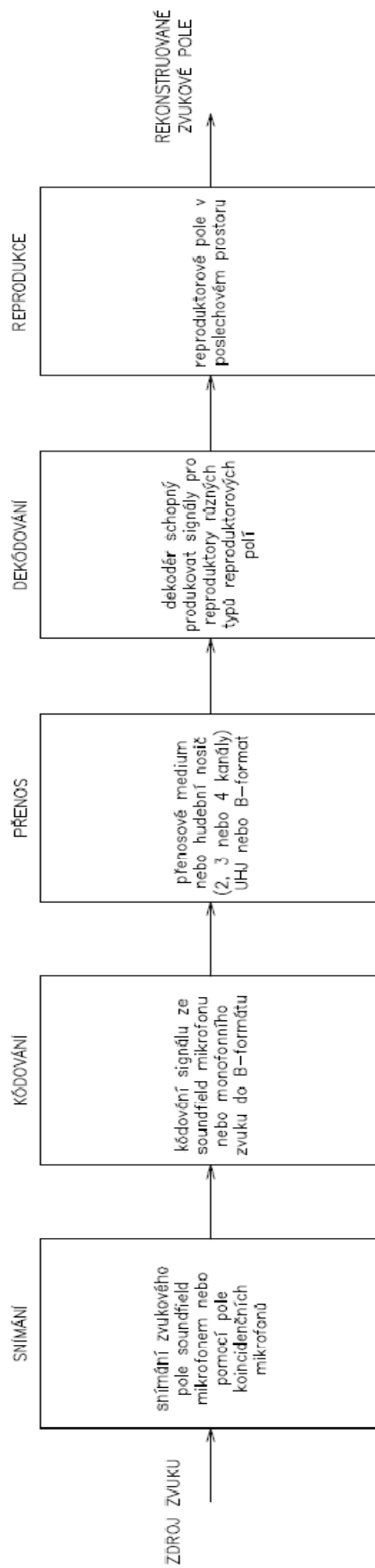
Cílem vývoje se tedy stalo nalézt takovou metodu, která by, místo věrné reprodukce původního zvukového pole, poskytovala pouze jakousi jeho podrobnou iluzi, čímž by se rapidně snížil počet potřebných měničů. Vznikly takové metody, jako například syntéza zvukových polí [21] nebo právě ambisonie (znak ambisonie je na obrázku 1.3).



Obr. 1.3: Logo ambisonie znázorňující čtyři koincidenční mikrofony [7].

2 AMBISONIE

Ambisonie je matematicky založený systém pro zpracování akustických signálů, který se snaží zachytit a reprodukovat informaci získanou z kompletního třídimenzionálního zvukového pole, včetně přesné lokalizace každého zvukového zdroje a charakteristických rysů okolního prostředí, které toto pole obsahuje. Je to vlastně proces kódování a dekódování akustického signálu. Tímto signálem může být signál získaný snímáním reálného zvukového pole speciálním způsobem nebo jakýkoliv jiný monofonní signál. Teorie a technologie související s ambisonií byly vyvinuty v 70. letech 20. století (skupinou lidí kolem Michaela A. Gerzona z matematického institutu v Oxfordu [16]) obzvláště pro nahrávání a reprodukci akustické hudby (hlavně klasické). Analogové vybavení používané v té době bylo však (a pořád zůstává) poměrně drahé, což zapříčinilo, že se ambisonie komerčně nerozšířila, tak jako například technologie Surround Sound 5.1. V dnešní době se ale spíše využívají pro zpracování digitální metody, zvláště pomocí software plug-inů. Díky tomu jsou tyto principy zpracování stále běžnější především v počítačové hudební produkci (mimo jiné například i v počítačových hrách) nebo při zvukovém snímání sportovních přenosů, což by mohlo pomoci většímu rozšíření ambisonie mezi veřejnost. Blokové schéma cesty signálu základním ambisonickým řetězcem je na obrázku 2.1.



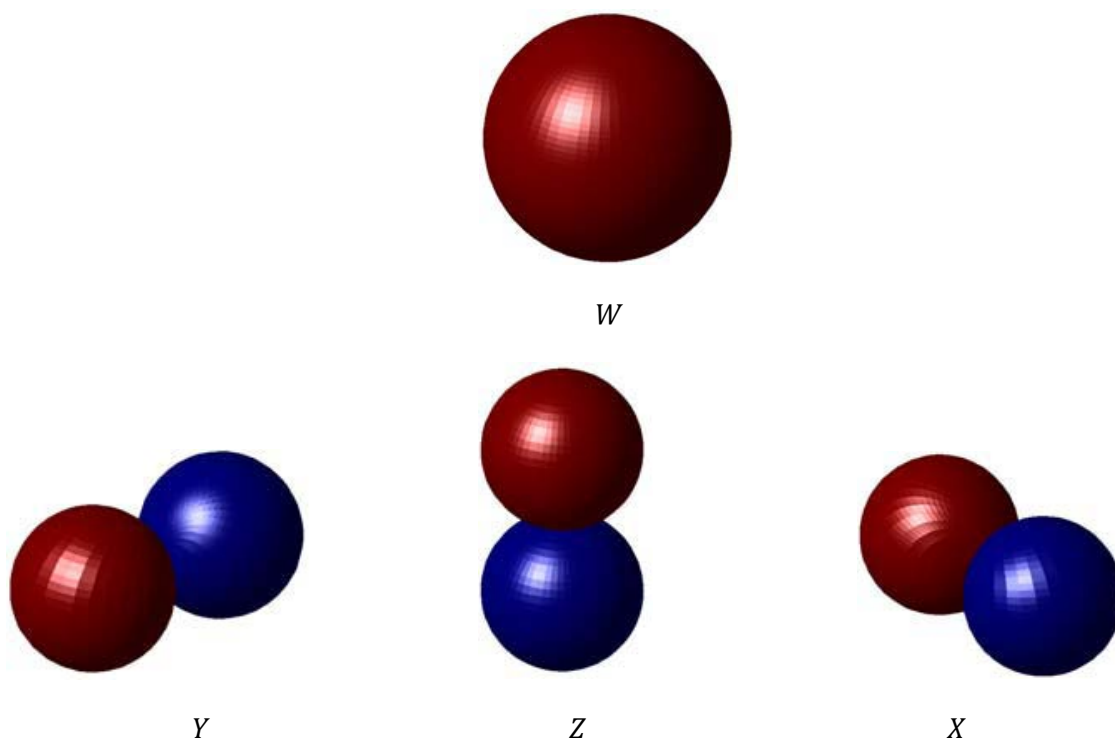
Obr. 2.1: Cesta signálu základním ambisonickým řetězcem [10].

3 KÓDOVÁNÍ

Kódování je proces, kdy každému možnému směru, pozici a vzdálenosti zdroje zvuku od posluchače ve zvukovém poli přiřadíme způsob, jakým se tento zvuk začlení do datových kanálů, které máme k dispozici. Různé umístění zdrojů zvuku pak odpovídá konkrétním zvukům zakódovaným s různou relativní fází a amplitudou v různých kanálech. Ambisonie prvního řádu používá kódování do čtyř datových kanálů (složek). V průběhu nahrávání živé akustické hudby, při zvukové syntéze nebo mixování různých zdrojů zvuku pomocí ambisonické technologie je tedy informace o lokalizaci zdrojů zvuku a okolním prostoru matematicky zpracována a zakódována do těchto samostatných kanálů (složek) obsahujících příslušná data. Na rozdíl od tradičního stera, 5.1 a dalších vícekanálových zvukových formátů, data z jednotlivých kanálů ambisonie nemohou být použity přímo jako zdrojové signály pro jednotlivé reproduktory. Tyto kanály nemají nic společného s reproduktory, pouze popisují to, co se v ambisonické terminologii nazývá sférické harmonické složky [12].

Při snímání zvukového pole se soustava mikrofonů snaží nasnímat hodnoty akustického tlaku na povrchu pomyslné koule (tzn. ze všech směrů). Pro rozlišení jednotlivých směrů pomocí snímání povrchu koule je potřeba jakási analogie k teorii Fourierovy transformace, která byla nazvána Teorie sférických harmonických složek (Theory Of Spherical Harmonics [12]). Podle ní každá poloha zdroje zvuku v prostoru může být popsána „směrovými kosiny“, které jsou definovány jako příslušné souřadnice (x, y, z) bodu na povrchu jednotkové koule, který určí jednotkový vektor vedoucí z počátku souřadného systému a směřující k tomuto zdroji zvuku. Potom velikost, s jakou bude zvuk z určitého směru (s určitými souřadnicemi) zastoupen v nahrávaném kanálu, je komplexní funkcí tohoto směru. Je zvykem tuto funkci považovat za směrovou charakteristiku mikrofonu (například velikost amplitudy v různých směrech). V ambisonii prvního řádu se setkáváme jen se sférickými harmonickými složkami nultého a prvního řádu. Nultému řádu odpovídá složka W , reprezentovaná mikrofonem s všesměrovou charakteristikou. Prvnímu řádu odpovídají složky X , Y a Z , reprezentované mikrofony s osmičkovou charakteristikou (obr. 3.1). Základní sférické harmonické složky popisují zvukové roviny, jako například umístění zvuku v rovině vpřed-vzad, vlevo-vpravo a nahoru-dolů.

Kódování (nahrávání nebo mixování) a dekodování jsou zcela odlišné procesy. Kódování není přímo spojeno s konkrétním počtem použitých reproduktorů, je to proces, který je navržen tak, aby zachytil co nejpřesněji všechny informace ze snímaného akustického pole. Pro reprodukci musí být data z těchto kanálů dekodována pro určité množství a rozestavení konkrétních reproduktorů v prostoru.



Obr. 3.1: Směrové charakteristiky mikrofonů reprezentující jednotlivé složky B-formátu. Červená barva reprezentuje kladný směr příslušné osy prostoru [21].

Základním požadavkem kódovacích metod je kompatibilita s předchozími rozšířeními zvukovými formáty. To znamená, aby bylo zakódovaná data následně možno dekodovat například ve stereo, či mono systému.

Tak může být ambisonická nahrávka dekodována do standardních stereo systémů nebo do systémů se čtyřmi, pěti, šesti, osmi, šestnácti a více reproduktory, rozmístěnými určitým způsobem v prostoru. V ambisonické terminologii se uspořádáním reproduktorů v prostoru, které jsou charakteristické nejen počtem reproduktorů, ale i jejich konkrétními umístěními uvnitř prostoru, říká reproduktorové pole (systém). Obecně platí, že čím více reproduktorů máme k dispozici, tím více přesné a detailní reprodukce můžeme dosáhnout. Více o dekodování se dozvíte v kapitole 5.

3.1 Snímání pomocí mikrofonů

Zdrojem ambisonických nahrávek může být speciálně navržené pole čtyř velmi blízkých (ideálně koincidenčních) mikrofonů. Tři z mikrofonů mají osmičkovou směrovou charakteristiku a jeden má charakteristiku všesměrovou (kulovou). Mikrofony jsou v kapsli rozmístěny tak, aby ty osmičkové byly navzájem kolmé a mohly tak snímat jednotlivé osy prostoru, mikrofon všesměrový pak snímá celkový akustický tlak v daném bodě prostoru – tzn. je nezávislý na směru příchozího zvuku. Pro tento účel se však vyrábějí vhodnější mikrofony, které mají čtyři (téměř) koincidenční mikrofony obsaženy v jedné mikrofonní kapsli (obr. 3.2).



Obr. 3.2: Kapsle mikrofonu soundfield se čtyřmi téměř koincidenčními mikrofony [20].

Jsou označovány jako soundfield mikrofony a mají za úkol snímat zvuk z jednoho konkrétního bodu akustického pole, pouze z různých směrů, daných jejich směrovými charakteristikami a umístěním (obr. 3.2). Umístění a směr jednotlivých mikrofonů jsou uvedeny v tabulce 3.1.

Tab. 3.1: Hodnoty azimutu a elevace jednotlivých mikrofonů v kapsli mikrofonu soundfield (*A*, *B*, *C*, *D* – signály A-format).

Mikrofon	Azimut	Elevace
A	45°	35,3°
B	135°	-35,3°
C	-45°	-35,3°
D	-135°	35,3°

Všechny mikrofony mají hyperkardioidní směrové charakteristiky. Snímané osy akustického pole jsou

- Osa *x* (vpředu–vzadu),
- Osa *y* (vlevo–vpravo),
- Osa *z* (nahore–dole).

Signály, které získáme z mikrofonu soundfield se nazývají A-format. Musí se dále převést do podoby určené pro další zpracování nebo přenos. Originální a stále nejvíce rozšířený formát ambisonického kódování, známý jako technologie B-format, využívá matematických rovnic za účelem převedení signálů z těchto čtyř mikrofonů do čtyř datových kanálů, odpovídajícím výše zmíněným sférickým harmonickým složkám snímaného akustického pole:

- *W* – všesměrový signál, který popisuje časovou změnu amplitudy intenzity každého zdroje zvuku,

- X – signál zakódovaný z mikrofону v ose x , představuje hloubku – popisuje jak je složka W (intenzita) rozložena podél osy x zvukového pole,
- Y – signál zakódovaný z mikrofону v ose y , představuje šířku – popisuje jak je složka W rozložena podél osy y zvukového pole,
- Z – signál zakódovaný z mikrofону v ose z , představuje výšku – popisuje jak je složka W rozložena podél osy z zvukového pole.

Poznámka:

Všimněte si, že popis os v ambisonii je nestandardní – většinou se používá pro osu vlevo–vpravo označení x , ne y .

Převod A-formátu (získaného ze soundfield mikrofону) do B-formátu se provádí pomocí rovnic:

$$W = 0.5 (A + B + C + D), \quad (3.1)$$

$$X = (A + C) - (B + D), \quad (3.2)$$

$$Y = (A + B) - (C + D), \quad (3.3)$$

$$Z = (A + D) - (B + C), \quad (3.4)$$

kde A, B, C, D jsou výstupní signály jednotlivých mikrofónů v kapsli soundfield (viz tab. 3.1).

Pro případ snímání zvukového pole ze 4 koincidenčních mikrofónů odpovídají směřům signálů z jednotlivých mikrofónů rovnice:

$$W = 1/\sqrt{2}, \quad (3.5)$$

$$X = \cos(\varphi) \cos(\varepsilon), \quad (3.6)$$

$$Y = \sin(\varphi) \cos(\varepsilon), \quad (3.7)$$

$$Z = \sin(\varepsilon), \quad (3.8)$$

kde ε je elevační úhel zdroje zvuku a φ je azimut zdroje zvuku (směrem od bodu přímo vpředu proti směru hodinových ručiček). Tyto signály jsou tedy přímo složkami B-formátu (není zde třeba konverze).

3.2 Konverze monofonního zvuku do B-formátu

Rovnice (3.5–3.8) se také dají použít pro zakódování jakéhokoliv monofonního signálu do B-formátu, tedy do určité pozice ve výsledném zvukovém poli. Jednotlivé rovnice se pouze vynásobí tímto monofonním signálem a zadá se požadovaná elevace a azimut, ve kterém se má signál objevit. Např. pro monofonní zvuk, který chceme zakódovat do azimutu 50° a elevace 0° (jde tedy pouze o horizontální pozici) budou rovnice vypadat následovně:

$$W = (1/\sqrt{2}) s = 0,7071 s, \quad (3.9)$$

$$X = \cos(50^\circ) \cos(0^\circ) s = 0,6428 s, \quad (3.10)$$

$$Y = \sin(50^\circ) \cos(0^\circ) s = 0,7660 s, \quad (3.11)$$

$$Z = \sin(0^\circ) s = 0, \quad (3.12)$$

kde s je monofonní zdroj zvuku, který chceme kódovat.

Zakódovaný signál je poté přenášen přenosovými kanály (například při přenosu na hudebním nosiči), které mohou být reprezentovány buď přímo složkami W, X, Y, Z nebo různými typy speciálního formátu UHJ.

4 FORMÁT UHJ

Byl vyvinut společností BBC kvůli lepším možnostem kompatibility hlavně se systémy mono a stereo, přičemž umožňuje i rozšíření na plnou trojrozměrnou reprodukci tam, kde jsou k tomu předpoklady a prostředky. UHJ by přitom mělo poskytovat při stereo reprodukci dokonce uspokojivější výsledky, než klasické stereo techniky. Princip spočívá v zakódování původních složek W, X, Y, Z do nových složek (datových kanálů Σ, Δ, T a Q) formátu UHJ:

$$\Sigma = 0,9397 W + 0,1856 X, \quad (4.1)$$

$$\Delta = j(-0,3420 W + 0,5099 X) + 0,6555 Y, \quad (4.2)$$

$$T = j(-0,1432 W + 0,6512 X) - 0,7071 Y, \quad (4.3)$$

$$Q = 0,9772 X. \quad (4.4)$$

Signály Σ a Δ jsou součet a rozdíl levého a pravého stereo kanálu, ale používají fázově-amplitudové kódování, aby přenesly kompletní 360° horizontální zvukové pole. Zhruba se dá říct, že poměr amplitud levého a pravého stereo kanálu reprezentuje stranovou směrovost (vlevo – vpravo) a směrovost vpředu – vzadu je určena rozdílem jejich fází [9]. Hierarchie UHJ se skládá ze čtyř formátů distribuce signálů:

- **BHJ**

Je optimalizován pro plnou sluchitelnost s mono a dvoukanálovým stereo systémem. Poskytuje přitom i dostatečně kvalitní reprodukci prostorového zvuku. Obsahuje dvě složky, Σ a Δ , zakódované z původních W, X a Y . Pro monofonní reprodukci se používá pouze signál Σ . Pro reprodukci stereofonní se pak signály určené pro levý a pravý reproduktor získají výpočtem:

$$L = 1/2 (\Sigma + \Delta), \quad (4.5)$$

$$R = 1/2 (\Sigma - \Delta). \quad (4.6)$$

- **SHJ**

Při dekódování do mono a stereo systémů zde platí stejná pravidla a výpočetní vztahy jako pro BHJ, uvedeny výše. Navíc je zde však rozšíření o kanál T , přenášený s redukovanou šířkou pásma (určen pro přenos přes média s omezenou šířkou pásma).

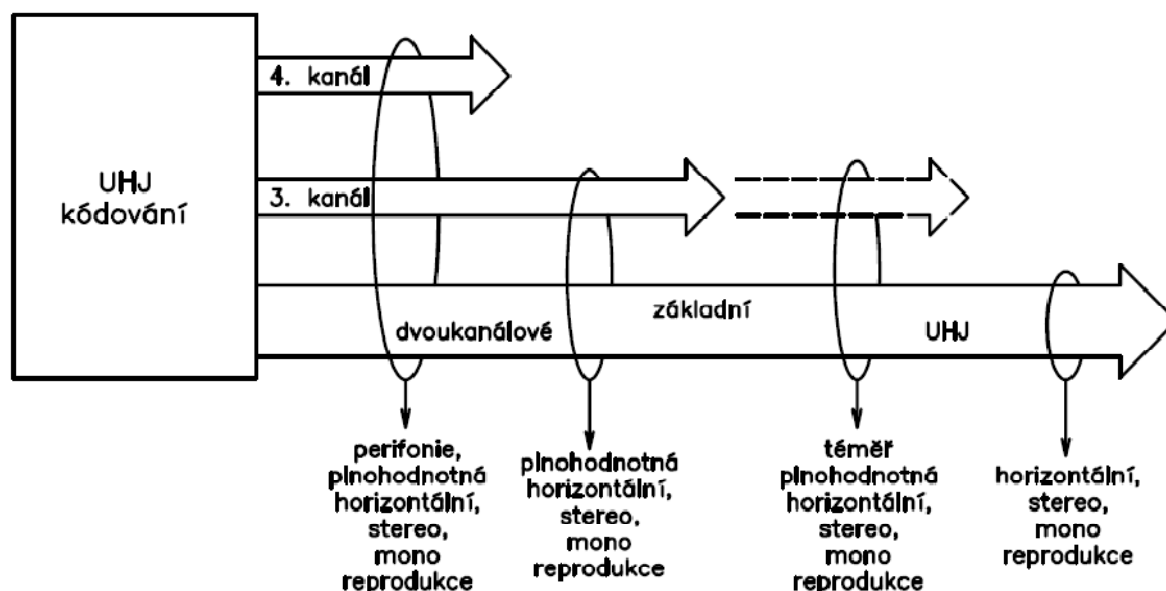
- **THJ**

Shodný se systémem SHJ, kanál T je však obsažen v plné šířce pásma.

- **PHJ**

Umožňuje oproti THJ navíc zahrnout informaci o výšce z přídatného kanálu Q , kódovaného ze složky Z .

První tři typy jsou určeny pro horizontální reprodukci, PHJ potom pro perifonii. Jak je vidět na obrázku 4.1, přijímač (respektive jeho dekodér) určený pouze pro jeden formát UHJ funguje při příjmu kteréhokoliv UHJ formátu. Vždy si totiž vybere jen kanály, které je schopen zpracovat. Nebo, v případě nejvýkonnějšího (PHJ) dekodéru, všechny kanály, které jsou při přenosu k dispozici (tzn.: pokud se přenáší jen např. THJ, poskytne pouze horizontální zvukový obraz). Dekodéry potom mají za úkol převést jednotlivé signály UHJ zpět na původní složky W , X , Y a případně i Z .



Obr. 4.1: Možnosti rozšíření a vzájemné kompatibility přenosu jednotlivých kanálů formátu UHJ [9].

Například u BHJ ovšem dochází k částečné ztrátě informace následkem toho, že jsme pomocí matice zakódovali tři signály pouze do dvou. Naštěstí se však nejedná o nic znepokojivého. Dá se to ošetřit například přidáním složky T , která může být přenášena s plnou šířkou pásma v případě THJ nebo pouze s redukovanou šířkou u SHJ. Tím pádem SHJ zajišťuje jisté zlepšení lokalizačního efektu oproti BHJ. Z THJ se potom dají získat původní složky B-formátu beze ztrát. Pokud do sady signálů tvořících THJ přidáme ještě další složku Q , dostaneme formát PHJ. Signál Q je tvořen pouze původní složkou Z , reprezentuje tedy informaci o výšce. PHJ tak nese kompletní sestavu složek ambisonie prvního řádu a umožňuje tak plnou perifonickou reprodukci sférického zvukového pole.

Na druhou stranu může být složka Q využita i jiným způsobem. Existuje způsob reprodukce, který je sice v rozporu se snahami ambisonie o to, aby v reprodukováném zvukovém poli posluchač pokud možno nemohl sluchem lokalizovat přesně rozmístění jednotlivých reproduktorů. Někteří hudební producenti však mohou požadovat naopak zdůraznění pozic reproduktorů ve zvukovém poli – tzv. efekt zdůraznění (emphasis effect [4]). Tohoto efektu lze dosáhnout právě s použitím složky Q , ovšem zakódované odlišným způsobem než u klasického PHJ, určeného pro perifonickou reprodukci.

Protože pouze varianta kódování BHJ byla rozšířena i v komerčním využití, termíny BHJ a UHJ se pomalu staly synonymy a UHJ bývá všeobecně považováno za formát BHJ.

5 DEKÓDOVÁNÍ

Dekódování je proces, kdy z jednotlivých kanálů (složek W , X , Y , Z) získáváme signály určené pro reprodukci jednotlivými konkrétními reproduktory reproduktorového pole, za účelem dosažení co nejlepší iluze zamýšleného nebo vzorového zvukového pole. Nejdříve tedy musíme určit, jaké reproduktorové pole máme k dispozici a poté vybrat vhodnou kombinaci složek, jejich fáze a amplitudy, pro každý reproduktor.

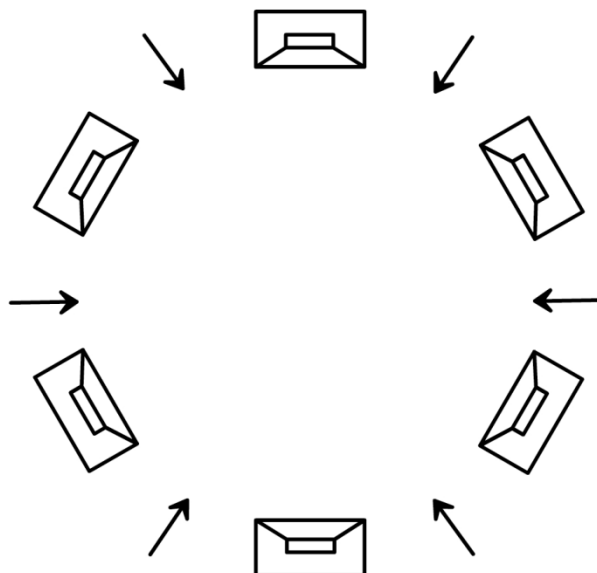
Původní ambisonické dekodéry byly vyráběny jako hardwarové jednotky, které zajišťovaly maticové techniky výpočtu správného poměru složek W , X , Y a Z se správnou fází pro každý reproduktor. Tyto se sice pořád prodávají (obzvláště v Anglii), dnes se však většina dekodování v ambisonii provádí pomocí příslušného softwarového vybavení. Aby dekodér splnil svoji funkci co nejlépe, musí znát přesnou konfiguraci reproduktorového pole (ne pouze počet reproduktorů, ale i jejich umístění uvnitř pole). Dekodéry (ať už jejich hardwarová nebo softwarová verze) proto běžně mívají možnost zadat počet reproduktorů a jejich konkrétní umístění. Upřednostňovány jsou však hlavně pole s reproduktory umístěnými symetricky a to tak, že tvoří pokud možno protilehlé dvojice (viz kapitola 5.2).

Signál určen pro konkrétní dvojici protilehlých reproduktorů se neliší pouze v intenzitě (jako např. v běžné intenzitní stereofonii), ale také ve fázi (rozdíl fází u protilehlých reproduktorů většinou činí 180°). V případě že chceme reprodukovat všechny dimenze příslušného zvukového pole, je nezbytné mít alespoň jeden reproduktor na každé z protilehlých polokoulí pomyslného sférického (kulového) reproduktorového pole. To znamená, že dekodování složky Y (šířky) vyžaduje minimálně jeden reproduktor nalevo od posluchače a druhý napravo. Podobně pro hloubku prostoru, tedy složku X , potřebujeme reproduktor vpředu a další za posluchačem. Takto vytvořená soustava ovšem nepovede k úplně kvalitní reprodukci.

Jisté zlepšení nastane už při posunutí reproduktorů jakoby do rohů čtvercové místnosti (přičemž posluchač je umístěn uprostřed) a natočením jejich osy směrem k posluchači. Tím dosáhneme toho, že úhel mezi reproduktory (vzhledem k posluchači) bude 90° a budeme mít vlastně dva reproduktory vpředu, dva vzadu a po dvou také na obou bočních stranách (obr. 1.1). Toto reproduktorové pole připomíná kvadrofonickou sestavu, ale zde jednotlivé reproduktory pracují na zcela odlišném principu. Ambisonie totiž nereprodukuje originální zvukové pole jako takové (to bychom dokázali pouze za použití obrovského množství mikrofónů a reproduktorů – viz kapitola 1), ale poskytuje souvislé informace o každém zdroji zvuku z více reproduktorů. Tím může vytvořit velmi přesvědčivou iluzi původně zaznamenaného nebo, v případě elektronicky vytvořené či panorámované hudby, zamýšleného zvukového pole.

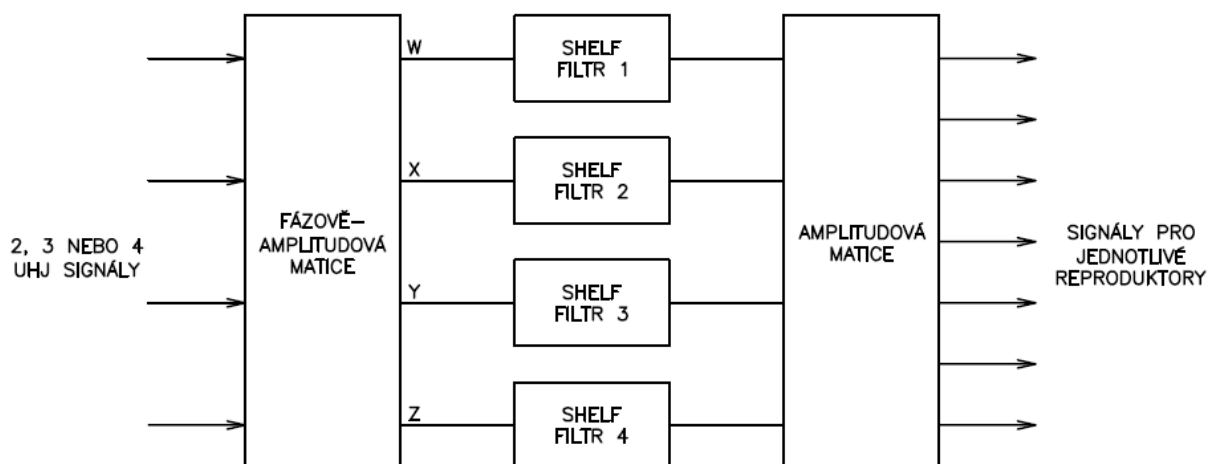
Když je například ambisonická nahrávka reprodukována přes stereo sestavu, obraz osy vlevo–vpravo je obvykle velmi dobrý. Vjem hloubky však bude pouze částečný. Reprodukce sice zprostředkuje nějakou informaci (obecně o dost lepší, než standardní stereo) o vzdálenosti zdrojů, které se nacházejí před posluchačem (máme totiž dva reproduktory v předních pozicích), protože ale chybí reproduktory zadní, nebude tato informace precizní. Zvuky ze zdrojů umístěných vzadu pak mohou mít tendenci se jakoby zrcadlit a vycházet z předních reproduktorů, pouze s redukovanou intenzitou. Informace o výšce, složka Z , nebude dekodérem vůbec vnímána. Protože stereo reproduktory jsou oba ve stejné výšce vzhledem k posluchači, neexistuje způsob, jakým by toto reproduktorové pole mohlo doplňkovou informaci o výšce využít.

Čtyři reproduktory v obdélníkovém (tedy i čtvercovém) uspořádání jsou tedy jakási minimální konfigurace reproduktorového pole, potřebná k úplné reprodukci horizontálního ambisonického obrazu. Pět reproduktorů rozestavených ve vrcholech pravidelného pětiúhelníka pak bude poskytovat o něco lepší horizontální obraz. Šestiúhelník ještě vylepší obzvláště stranový a zadní obraz – zdánlivé zvukové obrazy mezi reproduktory (obr. 5.1). Čím více reproduktorů tedy použijeme, tím detailnější iluze prostoru dosáhneme.



Obr. 5.1: Šestiúhelníkové reproduktorové pole s vyznačenými zdánlivými obrazy mezi reproduktory pomocí šipek [4].

Dekodér musí poskytovat takové signály, které vyhovují psychoakustickým kritériím prostorového slyšení lidského sluchu (viz kapitola 5.1). Je tedy nutné znát, jaké akustické podněty jsou nejlépe schopny v mozku vyprodukovat požadovanou iluzi. Tato iluze by měla být nezávislá na skladbě reproduktorového pole. Proto se musí signály určené pro jednotlivé reproduktory vždy počítat pro jejich konkrétní polohu. Tím se zajistí schopnost reprodukce požadovaného prostorového efektu na reproduktorových polích, které vyhovují širokému spektru prostorů různých velikostí a tvarů (obzvláště v domácích podmínkách). Dekodér může vypadat podobně jako na obrázku 5.2.



Obr. 5.2: Příklad ambisonického dekodéru [9].

Na vstupu fázově-amplitudová matice převádí signály z přenosových kanálů UHJ (Σ, Δ, T, Q) zpět do složek B-formátu (W', X', Y', Z') pomocí následujících rovnic.

Pro SHJ a THJ:

$$W' = 0,982 \Sigma + 0,197 j(0,828 \Delta + 0,768 t T) , \quad (5.1)$$

$$X' = 0,419 \Sigma - j(0,828 \Delta + 0,768 t T) , \quad (5.2)$$

$$Y' = 0,187 j\Sigma + (0,796 \times \Delta - 0,676 t T) , \quad (5.3)$$

$$B' = -0,694 j\Sigma + 0,140 (0,828 \Delta + 0,768 t T) . \quad (5.4)$$

Pro BHJ:

$$W' = 0,982 \Sigma + 0,164 j\Delta , \quad (5.5)$$

$$X' = 0,419 \Sigma - 0,828 j\Delta , \quad (5.6)$$

$$Y' = 0,385 j\Sigma + 0,763 \Delta , \quad (5.7)$$

$$B' = -0,694 j\Sigma + 0,116 \Delta . \quad (5.8)$$

Pro PHJ:

$$W' = 0,982 \Sigma + 0,197 j(0,828 \Delta + 0,768 T) , \quad (5.9)$$

$$X' = 0,419 \Sigma - j(0,828 \Delta + 0,768 T) , \quad (5.10)$$

$$Y' = 0,187 j\Sigma + (0,796 \Delta - 0,676 T) , \quad (5.11)$$

$$Z' = 1,023 Q , \quad (5.12)$$

kde t je koeficient množství příspěvku kanálu T ($t \in \langle 0;1 \rangle$), j je fázový posuv o 90° ($j = \sqrt{-1}$) a B' je speciální kanál pro zlepšení prostorových vlastností horizontálního zvukového pole.

Pokud jsou ovšem jako signály pro přenos použity přímo složky W, X, Y, Z , pak není tato dekódovací matice potřeba. Dalším stupněm jsou shelving (shelf) filtry, které mají za úkol upravit frekvenční vlastnosti jednotlivých složek nad a pod kmitočtem cca 500–700 Hz tak, aby vyhovovaly psychoakustickým kritériím lidského sluchu. Příslušné rovnice pro jednotlivé signály vystupující z filtrů jsou:

$$W = k_1 W' , \quad (5.13)$$

$$X = k_2 X' , \quad (5.14)$$

$$Y = k_2 Y' + k'k_3 B' , \quad (5.15)$$

$$Z = k_2 Z' , \quad (5.16)$$

kde k_1, k_2, k_3 jsou kmitočtově závislé zisky filtrů, k' je koeficient preference přední strany zvukového pole ($k' \in \langle 0; 0,7 \rangle$) a W', X', Y', Z', B' jsou složky dekodované z UHJ fázově-amplitudovou maticí a vstupující do filtrů.

Poznánka:

Pro PHJ, kde není signál B' obsažen (respektive se rovná 0), se rovnice (5.15) zjednoduší na tvar:

$$Y = k_2 Y' . \quad (5.17)$$

Jiný typ filtru (s jiným ziskem nad a pod dělicím kmitočtem) je použit pro složku W , a jiný pro složky ostatní – X, Y, Z (viz [9],[2],[8]). Konkrétní hodnoty zisku jednotlivých filtrů jsou zatím spíše otázkou experimentu a v různých literaturách se objevují různá doporučení. Amplitudová matice už potom získává signály určené pro jednotlivé reproduktory v konkrétních pozicích reproduktorového pole. Tato matice se liší podle druhu použitého reproduktorového pole. Pro pole pravidelná se používá rovnice:

$$S = 0,5 \left((\sqrt{2} W) + (\cos(\varphi) \cos(\varepsilon) X) + (\sin(\varphi) \cos(\varepsilon) Y) + (\sin(\varepsilon) Z) \right), \quad (5.18)$$

kde S je signál určený pro konkrétní reproduktor, φ je azimut tohoto reproduktoru (směrem od bodu přímo vpředu proti směru hodinových ručiček) a ε je elevace tohoto reproduktoru.

Pro nesymetrická pole se používají složitější postupy, např. heuristická metoda Tabu [21]. Ta je určená konkrétně pro systém Surround Sound 5.1 (viz kapitola 6.2). Složitější dekodéry pak mohou nabízet i možnosti relativního zdůraznění určité strany zvukového pole, nebo jeho rotaci.

5.1 Psychoakustická kritéria lidského sluchu v systémech surround sound

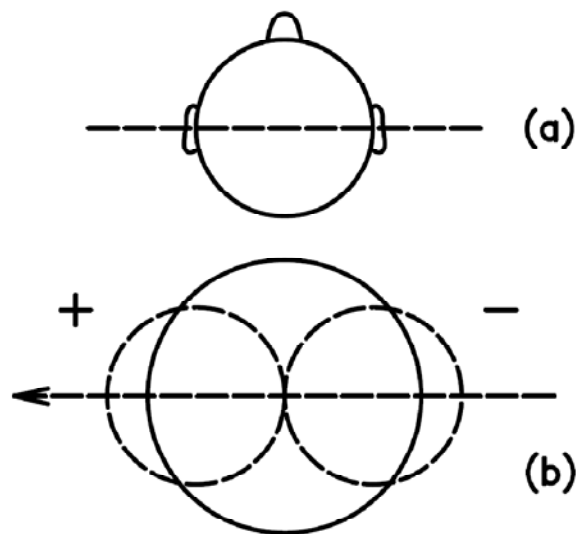
Psychoakustická kritéria prostorových zvukových systémů popisují vztahy mezi reprodukováným zvukovým polem a mezi tím, co posluchač opravdu slyší. Existuje mnoho mechanismů, na základě kterých uši lokalizují zvuky. Rozdílné mechanismy jsou využity pro nízké, střední a vysoké kmitočty. Dále je užitečná i informace získaná z odrazů zvuku od překážek. Většina současných systémů však nespĺňuje více než jedno nebo dvě tato kritéria. Pro vytvoření požadovaného zvukového obrazu ale nepotřebujeme úplně rozumět konkrétním složitým procesům zpracování zvuku v uších a mozku. Stačí nám pouze zjistit, jaký podnět je potřeba k vyvolání požadovaného vjemu. Potom se snažíme vytvořit systém, který dokáže tento podnět (a tedy i konkrétní vjem) vyvolat.

5.1.1 Nízkofrekvenční lokalizace

Vzdálenost mezi lidskýma ušima se rovná přibližně polovině vlnové délky zvuku o kmitočtu 700 Hz. Při kmitočtech nižších hlava nevytváří zvukovým vlnám překážku, takže amplituda zvuku dopadajícího do obou uší je v podstatě identická. Jediná informace, která je k dispozici, je fázový rozdíl (tzn. interaurální časová – fázová diference). Součet vln dopadajících do uší vytvoří akustický tlak, který by se naměřil v místě uprostřed posluchačovy hlavy. Tato informace se tedy dá snímat všesměrovým (tlakovým) mikrofonom (obr. 5.3).

5.1.2 Lokalizace na středních a vyšších kmitočtech

Při kmitočtech nad 700 Hz je vlnová délka zvuku dostatečně malá a interaurální časová diference už není při lokalizaci nejdůležitější faktor. Záleží spíše na směrových vlastnostech pole akustické intenzity (energie) okolo posluchače (interaurální intenzitní diference). Pokud se původní zvukový zdroj nacházel zhruba na místě některého z reproduktorů reproduktorového pole, pak se dá pole akustické energie, které vytvořil, reprodukovat pomocí poměrně malého počtu reproduktorů [13]. Pokud je konkrétní zvuk reprodukován několika reproduktory, výsledný směr tohoto zvuku se určí jako směr součtového vektoru vektorů, které směřují k jednotlivým reproduktorům a mají délku určenou akustickou energií vyzářenou tímto konkrétním reproduktorem. Poměry jednotlivých délek těchto vektorů k celkovému množství reprodukované akustické energie ve zvukovém poli by ideálně měly být shodné. Čím větší tento poměr je, tím lépe definovaný je výsledný zvukový obraz.



Obr. 5.3: a) Hlava posluchače s pomyslnou osou, na které jsou umístěny obě uši, b) Dva mikrofony v pozicích uší (+, -), které se dají nahradit jedním tlakovým mikrofonom [13].

5.1.3 Vysokofrekvenční lokalizace

U kmitočtů nad cca 5 kHz se projevuje nový mechanismus, založený na zabarvení zvuku vlivem odrazů od ušních boltců. Tento způsob lokalizace dosud není přesně prozkoumán, ale je založen na tom, že zvuk z různých směrů přichází do ucha s různým zabarvením (díky boltci). Takže pokud toto zabarvení dokážeme reprodukovat, dosáhneme požadovaných směrových účinků. V prostorových systémech (ani v ambisonii) se ale tato metoda ještě prakticky nevyužívá.

5.1.4 Lokalizace pomocí dozvuku

V místnostech s jakousi optimální dobou dozvuku (spíše kratší – cca 1 s) se zvuky dají lokalizovat přesněji než ve volném poli (bezodrazový prostor). Mechanismus opět není dokonale prozkoumán, je však známo, že správně snímaný a reprodukováný dozvuk lokalizaci pomůže, nedokonalý umělý dozvuk ale zvukový obraz degraduje. Bylo zjištěno, že nejlepších výsledků se dosáhne, pokud je intenzita dozvuku rozložena rovnoměrně (tzn.: není příliš koncentrována v žádném směru). Aby systém fungoval správně, musí být tedy schopen zaznamenat všechny

detaily dozvuku a reprodukovat je rovnoměrně kolem posluchače. Těto požadované rovnoměrnosti však zatím systémy (včetně ambisonie) dostatečně nevyhovují a výsledný efekt je tedy spíše negativní. Proto se tento mechanismus také prakticky nevyužívá [13].

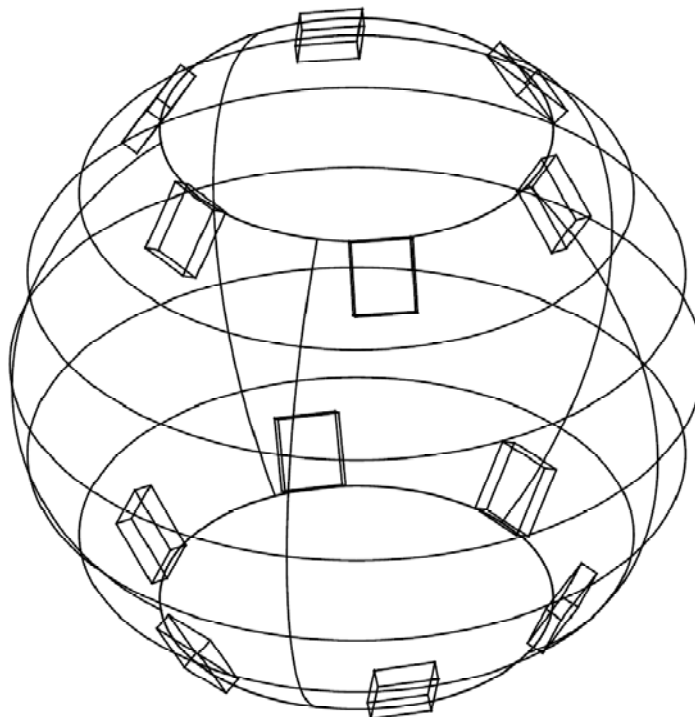
5.2 Typy reproduktorových polí

Ambisonické dekodéry mohou být navrženy pro celou řadu různých reproduktorových polí, které se dělí na:

- pravidelné nebo nepravidelné
- horizontální (2D) nebo perifonické (3D)

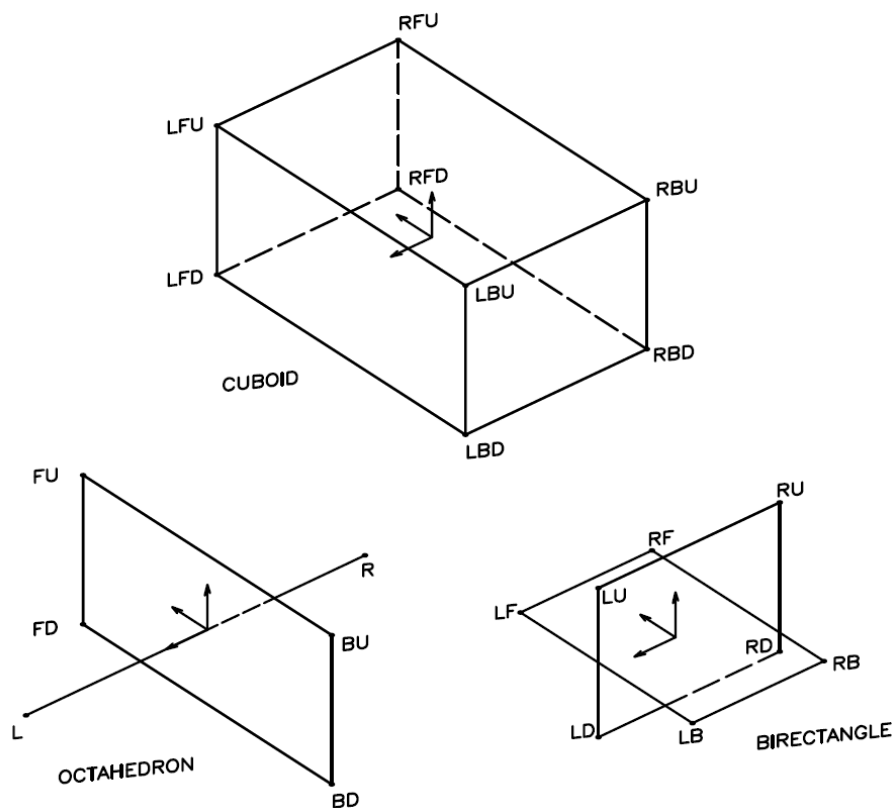
Pravidelná pole jsou v horizontálním případě pravidelné n -úhelníky (např. obr. 5.1), ve všesměrovém (3D) např. dva n -úhelníky ve dvou různých výškách (tzn. na hranách pomyslného válce uvnitř kulového zvukového pole – obr. 5.4) nebo reproduktory rozmístěné do pomyslných vrcholů n -stěny (obr. 5.5). Reproduktory by přitom ideálně měly být v protilehlých párech. Osy reproduktorů směřují ke středu reprodukováného zvukového pole.

Nepravidelná reproduktorová pole jsou častá hlavně v domácnostech kvůli často velké rozmanitosti a nepravidelnosti obytných prostor. Používají se pouze k horizontální reprodukci. Získání signálů určených pro jednotlivé reproduktory tak, aby výsledné zvukové pole vyhovovalo požadavkům, je zde poměrně složité. Ke zjištění koeficientů amplitudové matice pro výpočet poměrů jednotlivých složek, určených pro tyto signály, se používají různé metody. Asi nejlepších výsledků (konkrétně pro systém 5.1) se dosahuje použitím heuristických metod v takzvaném Tabu algoritmu [21].



Obr. 5.4: Příklad trojrozměrného reproduktorového pole s dvanácti reproduktory.

Obecně platí, že čím vyšší je počet kanálů, tím vyšší musí být minimální počet reproduktorů, aby nedocházelo k přílišnému zdůraznění pozic jednotlivých reproduktorů. Pokud je počet kanálů shodný s počtem reproduktorů, zdánlivé zvukové obrazy mezi reproduktory jsou vtahovány před nejbližší reproduktor.



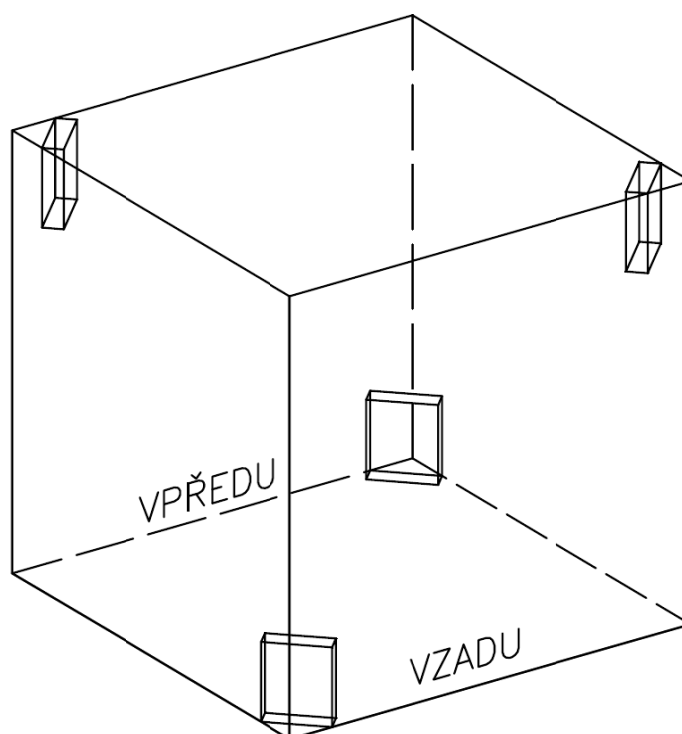
Obr. 5.5: Příklad n-stěnů reproduktorových polí [9].

Popisy u vrcholů n-stěnů a reproduktorů polí značí pozice reproduktorů: B – zadní (back), F – přední (front), L – levý (left), R – pravý (right), D – spodní (down), U – horní (up), S – surround.

6 PERIFONIE

Jak už bylo zmíněno, dekodování složky Z vyžaduje, aby alespoň některé reproduktory byly umístěny nad a pod úrovní pomyslné osy posluchačova slyšení (tzn. například na zemi a nad jeho hlavou). Zatímco horizontální ambisonie poskytuje dvojdimenzionální zvukové pole, ve kterém zvuky mohou být lokalizovány v jakémkoliv bodě v horizontální rovině uvnitř pole, přidání výškové komponenty zvětší reprodukční prostor na úplné třídimenzionální (kulové - sférické) zvukové pole. Tento způsob reprodukce (nebo snímání) se nazývá perifonie (pochází z řečtiny).

Perifonie je možné dosáhnout v nejjednodušším případě rozestavením reproduktorů do vrcholů čtyřstěnu (tetrahedronu), což znamená použití čtyř reproduktorů (obr. 6.1). Pro dostatečné zastoupení výškové složky je ale vhodné použít spíše osm a více reproduktorů rozdělených do dvou okruhů, jeden pod a druhý nad úrovní hlavy (např. obr. 5.4).



Obr. 6.1: Nejjednodušší trojrozměrné reproduktorové pole (tetrahedron) [12].

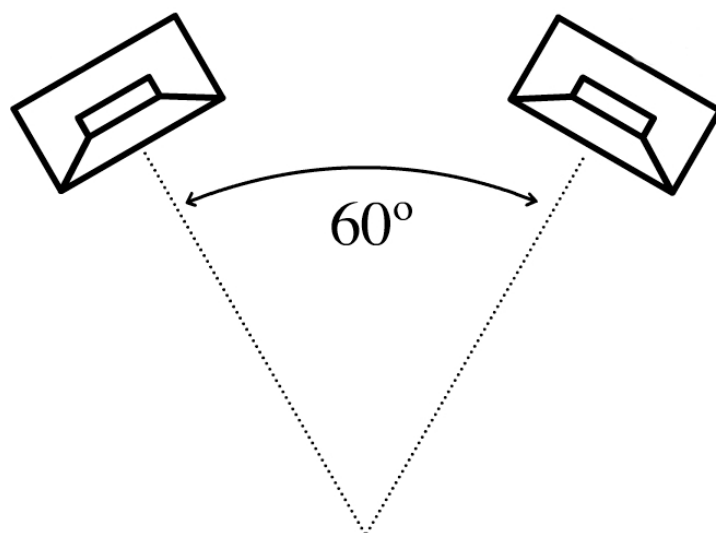
Pro záznam například sólové hry na piáno nebo kterékoliv jiné hudby, ve které se všechny zvukové zdroje zdají být přibližně ve stejné výšce, se může zdát dekodování výškové složky Z zbytečné, zvláště z hlediska nárůstu ceny vybavení a dalších komplikací. Nicméně, posluchači, kteří se setkali s perifonickým poslechem, často pozorovali, že hudba se zdá nejen „plnější“ a více „živá“ (částečně díky různým odrazům a dalším vlastnostem prostoru snímaného pole, reprodukováným zvýšenými a sníženými reproduktory), ale že jsou také schopni slyšet více detailů (pozorování podobná jako při srovnání stereo a monofonní reprodukce) [12]. Navíc, v produkci hudby s pomocí počítače, znamená možnost vytvoření zvukového 3D obrazu a schopnost umístit konkrétní zvuk do jakéhokoliv směru vznik úplně nových, kreativních výrazových prostředků.

7 VÝHODY AMBISONIE

Kromě toho, že ambisonie nabízí vylepšené možnosti iluze zvukového pole ve srovnání se standardními stereo nebo multikanálovými reprodukcími systémy, jsou zde i další významné výhody.

Ambisonie není limitována konkrétním počtem přenosových kanálů. Čím větší počet jich bude, tím vyšší rozlišení směrových informací získáme. Navíc platí vzájemná kompatibilita, tzn. přenosové médium se třemi kanály (nebo jakékoliv jiné tříkanálové vysílání) je možno dekódovat systémem určeným pro 2 nebo 4 kanály a podobně.

Rozměry prostoru určeného pro posluchače aby slyšel zvukové pole přesně (nazývaného „sweet spot“) jsou mnohem větší, než například u stera, kde se jedná pouze o relativně malý bod v pomyslném vrcholu rovnostranného trojúhelníka, v jehož zbývajících vrcholech jsou reproduktory směřující osou k posluchači (obr. 7.1). Výsledný obraz zůstává přesný a stabilní v poměrně široké oblasti, která často obklopuje většinu prostoru místnosti nebo koncertního sálu. Tento prostor můžeme ještě zvětšit použitím ambisonie vyššího řádu. Posluchači sedící v libovolném místě uvnitř pak přijímají v podstatě stejnou zvukovou informaci.



Obr. 7.1: Stereo systém s vyznačením bodu „sweet spot“ ve vrcholu trojúhelníka [4].

Hlavně pro domácí využití je vhodné, že posluchač nemusí být umístěn pouze čelem k předním reproduktorům. Může být natočen kterýmkoliv směrem.

Další výhodou je, že počet reproduktorů, jejich umístění a směrová orientace nejsou natolik kritické faktory, jak tomu bývá u ostatních systémů. V rámci co nejkvalitnější reprodukce sice v tomto směru existují určitá pravidla a doporučení (zmněna výše), ale je k dispozici mnoho různých možností, jak reproduktory umístit prakticky kamkoliv okolo s téměř zanedbatelnou degradací reprodukováného zvukového pole. Některá reproduktorová pole jsou ale samozřejmě efektivnější, některá méně.

Ambisonie také redukuje zdůraznění umístění reproduktorů ve zvukovém poli – aby zdánlivé zvukové obrazy nebyly „vtahovány“ směrem před jednotlivé reproduktory. Zdánlivé zvukové obrazy mezi reproduktory jsou tedy stabilní (na rozdíl např. od kvadrofonie [14]).

8 OMEZENÍ

Protože ambisonické dekódované pole může být dost složité a vyžaduje přesnou reprodukci, je důležité, aby k reprodukci byly použity kvalitní reproduktory. Ty by pokud možno měly být stejného typu a konstrukce (nejlépe zcela identické modely) se shodnými směrovými a kmitočtovými charakteristikami. Musí být fázově srovnané a ideálně pokrývají celé spektrum slyšitelných kmitočtů. Reprodukce může znít značně špatně v případě použití levných či rozdílných reproduktorů.

Jak bylo řečeno, nahrávky realizované s pomocí ambisonických technik poskytují jejich vlastní prostorovou informaci a informaci o okolním prostředí. Cílem dekóderu a reprodukce je přesné přenesení těchto informací do cílového prostoru ve formě vzniklého zvukového pole. Cílový prostor pak do výsledného vjemu přidá pouze malé procento jeho vlastností. Sál tak může například poskytovat ideální podmínky pro symfonický orchestr, ale když jsou tyto jeho vlastnosti překryty lokalizačními vjemy a vlastnostmi okolního prostoru prezentovanými reprodukcí ambisonického materiálu, může se objevit jakési „rozmazání“ zvukového obrazu a celková kvalita obklopujícího reprodukčního prostředí se změní a vjem se jakoby rozptýlí.

Další nevýhodou jsou sály, ve kterých se hlediště postupně od zadu dopředu svažuje. Tím vzniknou problémy s umístěním reproduktorů, protože uši posluchačů v různých místech jsou v různých výškách, což je pro ambisonii horší než pro jiné systémy. Dokonce i pro pouze horizontální reprodukci. Zadní reproduktory, blízké úrovni uší posluchačů v zadních řadách, budou sice dobře slyšitelné i pro posluchače v předních řadách, avšak přední reproduktory budou ponořeny pod úroveň nohou posluchačů v zadních řadách. Pro perifonii pak bude situace ještě složitější (není zde možnost rozestavět např. 8 reproduktorů do rohů pomyslné krychle nebo kvádru). Tyto potíže ovšem nejsou neřešitelné, zvláště také proto, že ambisonie, jak bylo zmíněno, toleruje určité chyby v rozmístění reproduktorového pole.

Všechna tato omezení by se také měly vzít v potaz při vytváření lokalizace v průběhu mixování, aby se předešlo tomu, že nahrávka, která výborně zní ve studiu, ve velkém sále pokulhává.

9 FORMÁTY AMBISONIE

Signály, které získáme z mikrofonu soundfield, jsou označovány A-format. Tyto se následně kódují do složek B-formátu. Ačkoliv je ambisonie prvního řádu (B-format) nejrozšířenější, existuje i řada jiných formátů. Formáty vyšších řádů poskytují mimo *W, X, Y, Z* navíc přídatné složky pro ještě konkrétnější lokalizaci a věrohodnost zvukového pole. Například ambisonie druhého řádu obsahuje devět kanálů (složek) zakódovaných zvukových dat. Pro reprodukci ambisonie na nesymetrickém surroundovém systému 5.1 se B-formát kóduje do tzv. G-formátu, který svými vlastnostmi lépe kompenzuje problémy tohoto systému s asymetrickým rozložením reproduktorů. Zejména kvůli kompatibilitě s přenosovými systémy s různým počtem kanálů (2, 3 nebo 4) byl vyvinut formát kódování UHJ, určen pro přenos ambisonických dat těmito systémy. Občas je také popisován jako C-format (zprostředkovaný posluchači hudebním nosičem nebo jiným médiem - „consumer format“). Dekódováním formátu UHJ znovu získané původní složky B-formátu jsou pak nazývány E-format a po průchodu shelf filtry F-format. Signály určené jednotlivým reproduktorům, které jsou produktem dekodéru, jsou často označovány jako D-format.

10 APLIKACE B-FORMAT DEKODÉR

V prostředí MATLAB jsem vytvořil aplikaci s vlastním grafickým uživatelským rozhraním, která je určena ke zpracování signálů B-formátu. Umožňuje mimo jiné zobrazit zdrojové složky B-formátu, konkrétně jejich časové průběhy a spektrogramy, výpočet signálů určených pro konkrétní reproduktory v různých typech reproduktorových polí (symetrických a nesymetrických) a následně zobrazení průběhů jejich intenzit v závislosti na čase, spektrogramů, závislosti intenzit na azimutu a elevačním úhlu zdroje zvuku ve trojdimenzionální i dvojdimenzionální podobě a samotné závislosti intenzit pouze na azimutu zdroje. Kromě zmíněného dekódování nabízí program B-format Dekodér i zakódování zdrojových složek B-formátu do třech typů formátu UHJ, určeného pro případný přenos přes média s různou šířkou pásma a zajišťující kompatibilitu s různými typy reprodukcí systémů (viz kapitola 4). Tyto formáty jsou BHJ, THJ a PHJ. Mezi další funkce aplikace patří zobrazení podporovaných typů reproduktorových polí a nastavení mezních kmitočtů shelving filtrů, použitých při dekódování.

10.1 Struktura aplikace

Struktura aplikace je naznačena na obrázku 10.1. Skládá se z 22 m-souborů. Hlavní m-soubor, který obsahuje funkci se samotným grafickým uživatelským rozhraním, má název

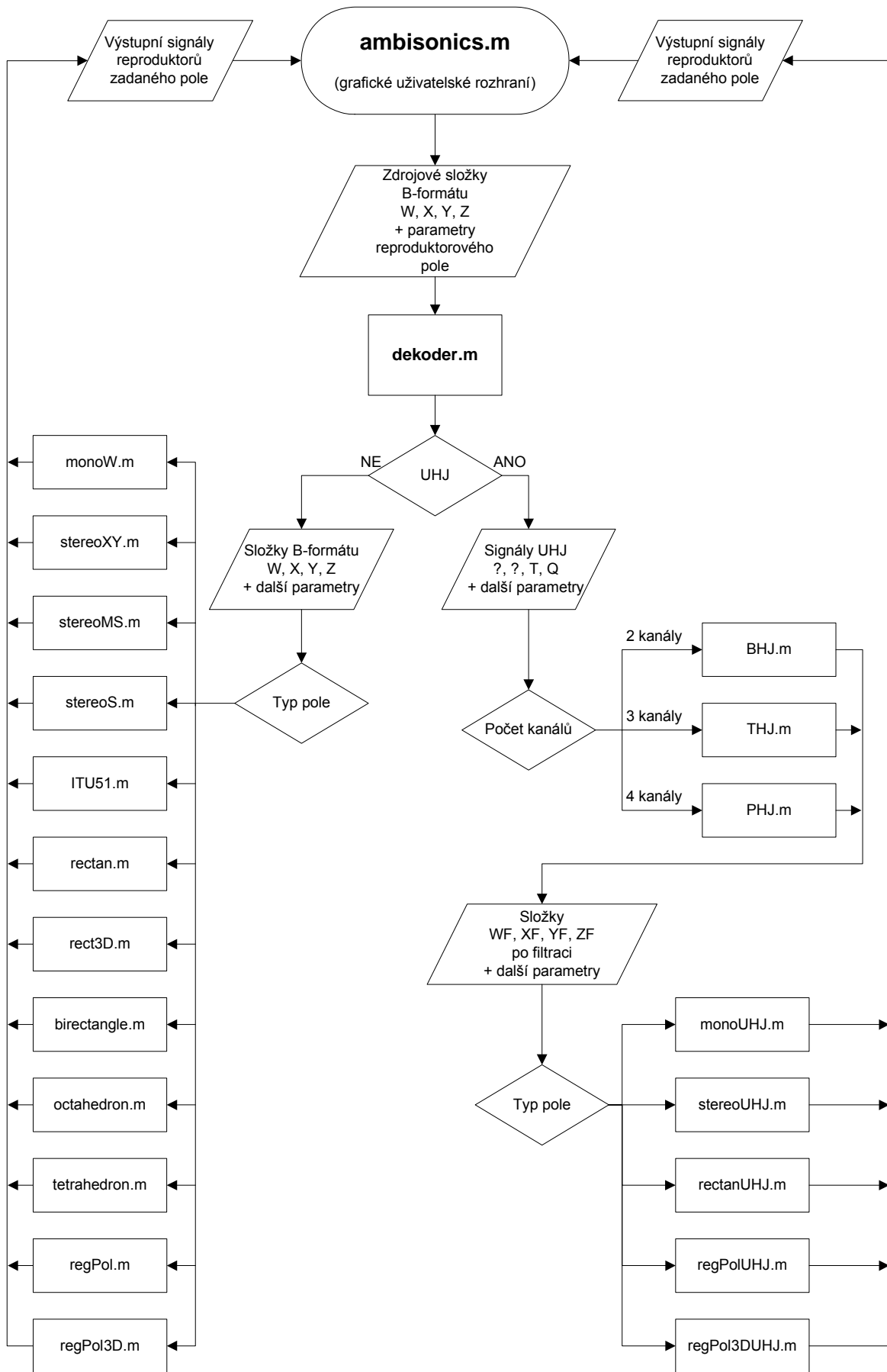
- `ambisonics.m`.

Jde tedy o funkci, s jejíž pomocí uživatel celou aplikaci ovládá, slouží k načtení a zobrazení zdrojových signálů a zobrazení, analýze, přehrávání a uložení výsledných signálů.

Další důležitou funkci, kterou volá funkce *ambisonics*, obsahuje soubor jménem

- `dekoder.m`.

Ten je určen ke zpracování vstupních parametrů zadaných v grafickém rozhraní (tzn. uživatelských požadavků na výstupní signály pro konkrétní reproduktorové pole). Zde se také uskutečňuje případná konverze složek W , X , Y , Z do koeficientů formátů UHJ podle rovnic (4.1), (4.2), (4.3) a (4.4). Podle parametrů zadaných uživatelem se funkce nejdříve rozhoduje, zda provést konverzi do UHJ. V případě, že tato varianta konverze byla zvolena, je ještě nutné určit počet kanálů UHJ, tedy formát BHJ pro 2 kanály, THJ pro tři nebo PHJ pro čtyři kanály. Podle této hodnoty se volají funkce v souborech `BHJ.m`, `THJ.m` nebo `PHJ.m`. Ty se postarají o výpočet (rekonstrukci) složek B-formátu z příslušného počtu kanálů UHJ. Dalším krokem je pak určení reproduktorového pole, pro které se mají zdrojové akustické signály dekódovat. Nakonec dochází k volání funkce pro požadované pole s příslušnými koeficienty (tedy konkrétním rozmístěním jednotlivých reproduktorů v poli).



Obr. 10.1: Základní struktura aplikace B-formát Dekoder.

Jak bylo řečeno výše, funkce *BHJ*, *THJ* a *PHJ* se starají o zpětnou konverzi z příslušného počtu kanálů UHJ zpět na složky B-formátu. Jim odpovídající soubory jsou:

- *BHJ.m*,
- *THJ.m*,
- *PHJ.m*.

Takto získané signály B-formátu ovšem nejsou například v případě funkce *BHJ* identické s těmi původními (zdrojovými), jak je zdůvodněno v kapitole 4. Dekódování ve funkci *BHJ* je provedeno pomocí rovnic (5.5), (5.6), (5.7) a (5.8). *THJ* obsahuje dekodovací rovnice (5.1), (5.2), (5.3), (5.4) a *PHJ* rovnice (5.9), (5.10), (5.11) a (5.12). Součástí všech těchto funkcí je navíc filtrace rekonstruovaných složek B-formátu shelving filtry (podle rovnic (5.13), (5.14), (5.15), (5.16) a případně (5.17)) s koeficienty zesílení *G* získanými z [9] pro *BHJ* a *THJ* a [8] pro *PHJ*. Následkem filtrace je zlepšení lokalizačních vlastností signálů na základě psychoakustických kritérií lidského sluchu. Zdrojový kód z prostředí MATLAB určený k filtraci rekonstruovaných složek B-formátu z dvoukanalového UHJ (tedy *BHJ*) vypadá následovně:

```
% shelving filtry

k = 0.6;
m = tan(pi*fc/48000);

% low-frequency cut k1

G1 = 20*log10(0.646);
k1 = 10^(-G1/20);
jmenovatel1 = 1 + sqrt(2*k1)*m + k1*m^2;
b01 = (1 + sqrt(2)*m + m^2) / jmenovatel1;
b11 = (2*(m^2 - 1)) / jmenovatel1;
b21 = (1 - sqrt(2)*m + m^2) / jmenovatel1;
a11 = (2*(k1*m^2 - 1)) / jmenovatel1;
a21 = (1 - sqrt(2*k1)*m + k1*m^2) / jmenovatel1;

b1 = [b01 b11 b21];
a1 = [1 a11 a21];

% low-frequency boost k2

G2 = 20*log10(1.263);
k2 = 10^(G2/20);
jmenovatel2 = 1 + sqrt(2)*m + m^2;
b02 = (1 + sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
b12 = (2*(k2*m^2 - 1)) / jmenovatel2;
b22 = (1 - sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
a12 = (2*(m^2 - 1)) / jmenovatel2;
a22 = (1 - sqrt(2)*m + m^2) / jmenovatel2;

b2 = [b02 b12 b22];
a2 = [1 a12 a22];

k2YE = filter(b2,a2, YE);

% low-frequency cut k3

G3 = 20*log10(0.775);
```

```

k3 = 10^(-G3/20);
jmenovatel3 = 1 + sqrt(2*k3)*m + k3*m^2;
b03 = (1 + sqrt(2)*m + m^2) / jmenovatel3;
b13 = (2*(m^2 - 1)) / jmenovatel3;
b23 = (1 - sqrt(2)*m + m^2) / jmenovatel3;
a13 = (2*(k3*m^2 - 1)) / jmenovatel3;
a23 = (1 - sqrt(2*k3)*m + k3*m^2) / jmenovatel3;

b3 = [b03 b13 b23];
a3 = [1 a13 a23];

BF = filter(b3,a3, BE);
kk3BF = k*BF;

```

% vysledek

```

WF = filter(b1,a1, WE);
XF = filter(b2,a2, XE);
YF = k2YE + kk3BF;

```

Pro filtraci tří kanálů THJ jsou použity oproti BHJ pouze dva filtry realizovány kódem:

% shelf filtry

```

m = tan(pi*fc/48000);

```

% low-frequency cut k1

```

G1 = 20*log10(0.661);
k1 = 10^(-G1/20);
jmenovatel1 = 1 + sqrt(2*k1)*m + k1*m^2;
b01 = (1 + sqrt(2)*m + m^2) / jmenovatel1;
b11 = (2*(m^2 - 1)) / jmenovatel1;
b21 = (1 - sqrt(2)*m + m^2) / jmenovatel1;
a11 = (2*(k1*m^2 - 1)) / jmenovatel1;
a21 = (1 - sqrt(2*k1)*m + k1*m^2) / jmenovatel1;

b1 = [b01 b11 b21];
a1 = [1 a11 a21];

```

% low-frequency boost k2

```

G2 = 20*log10(1.293);
k2 = 10^(G2/20);
jmenovatel2 = 1 + sqrt(2)*m + m^2;
b02 = (1 + sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
b12 = (2*(k2*m^2 - 1)) / jmenovatel2;
b22 = (1 - sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
a12 = (2*(m^2 - 1)) / jmenovatel2;
a22 = (1 - sqrt(2)*m + m^2) / jmenovatel2;

b2 = [b02 b12 b22];
a2 = [1 a12 a22];

```

% vysledek

```

WF = filter(b1,a1, WE);
XF = filter(b2,a2, XE);
YF = filter(b2,a2, YE);

```

A v případě filtrace čtyřkanálové perifonie PHJ potřebujeme dva filtry s kódem:

```
% shelf filtry

    m = tan(pi*fc/48000);

% high-frequency boost k1

    G1 = sqrt(2);
    k1 = 10^(G1/20);
    jmenovatel1 = 1 + sqrt(2)*m + m^2;
    b01 = (k1 + sqrt(2*k1)*m + m^2) / jmenovatel1;
    b11 = (2*(m^2 - k1)) / jmenovatel1;
    b21 = (k1 - sqrt(2*k1)*m + m^2) / jmenovatel1;
    a11 = (2*(m^2 - 1)) / jmenovatel1;
    a21 = (1 - sqrt(2)*m + m^2) / jmenovatel1;

    b1 = [b01 b11 b21];
    a1 = [1 a11 a21];

% low-frequency boost k2

    G2 = sqrt(3);
    k2 = 10^(G2/20);
    jmenovatel2 = 1 + sqrt(2)*m + m^2;
    b021 = (1 + sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
    b121 = (2*(k2*m^2 - 1)) / jmenovatel2;
    b221 = (1 - sqrt(2*k2)*m + k2*m^2) / jmenovatel2;
    a121 = (2*(m^2 - 1)) / jmenovatel2;
    a221 = (1 - sqrt(2)*m + m^2) / jmenovatel2;

    b21 = [b021 b121 b221];
    a21 = [1 a121 a221];

% vysledek

    WF = filter(b1,a1, WE);

    XF1 = filter(b21,a21, XE);
    YF1 = filter(b21,a21, YE);
    ZF1 = filter(b21,a21, ZE);

    XF = filter(b1,a1, XF1);
    YF = filter(b1,a1, YF1);
    ZF = filter(b1,a1, ZF1);
```

kde f_c je mezní kmitočet filtrů (lze nastavit v grafickém prostředí), G je zesílení propustného/potlačení nepropustného pásma (boost/cut), k je koeficient preference přední strany zvukového pole, k_1 , k_2 , k_3 jsou parametry filtrů (liší se ve znaménku u G pro filtr typu boost a cut), WE , XE , YE , ZE , BE jsou složky rekonstruované z UHJ (příslušné rovnice (23–26)), WF , XF , YF , ZF jsou výsledné rekonstruované složky B-formátu po filtraci a a , b jsou koeficienty přenosové funkce filtrů.

Poznámka:

Všechny shelving filtry použité v aplikaci B-format Dekodér mají vzorkovací kmitočet f_{vz} nastaven na 48 kHz, což je vzorkovací kmitočet se kterým byly pořízeny všechny testovací nahrávky akustických signálů B-formátu, používané v tomto programu.

Ostatní m-soubory jsou funkce, určené pro vytváření koncových signálů pro reproduktory v jednotlivých reproduktorových polích (přičemž název souboru odpovídá příslušnému reproduktorovému poli). Jsou to tyto:

- monoUHJ.m,
- stereoUHJ.m,
- rectanUHJ.m,
- regPolUHJ.m,
- regPol3DUHJ.m,
- monoW.m,
- stereoXY.m,
- stereoMS.m,
- stereoS.m,
- ITU51.m,
- rectan.m,
- rect3D.m,
- birectangle.m,
- octahedron.m,
- tetrahedron.m,
- regPol.m,
- regPol3D.m.

Všechny obsahují výpočty pozic konkrétních reproduktorů v soustavě (jejich azimutů a elevací vzhledem k pozici posluchače) s hodnotami parametrů zadanými uživatelem v grafickém rozhraní. Tyto m-soubory můžeme rozdělit podle toho, zda se jedná o pole pravidelná nebo nepravidelná (viz kapitola 5.2). Další dělení je možné na základě zkratky UHJ, kterou mohou obsahovat ve svém názvu. Pokud tomu tak je, jedná se o funkci určenou k dekódování složek B-formátu rekonstruovaných z kanálů UHJ. Jestliže se navíc jedná i o pravidelné pole, znamená to, že tyto funkce mají jako vstupní parametry výstupní signály WF , XF , YF a ZF funkcí BHJ , THJ nebo PHJ (podle počtu kanálů, které zadá uživatel v grafickém prostředí). Jak je popsáno výše, tyto vstupní složky jsou již odfiltrovány shelving filtry v příslušných funkcích, proto v těchto funkcích už k další filtraci nedochází. Opačně je tomu u m-souborů bez zkratky UHJ, které přitom současně reprezentují funkce pro výpočet signálů pro pravidelné soustavy. V tomto případě funkce obsahují vlastní shelving filtry s koeficienty zisku G , získanými z literatury [2], [8] a [15]. Přitom soubory tetrahedron.m, octahedron.m, birectangle.m, rect3D.m a regPol3D.m pro pravidelná perifonická pole (tzn. pro pole Tetrahedron, Octahedron, Birectangle, Rectangle 3D a Pravidelný polygon 3D) obsahují ve svých funkcích stejné shelving filtry se shodnými koeficienty jako funkce PHJ . Tyto filtry s jejich zdrojovým kódem byly uvedeny výše. Soubory rectan.m a regPol.m obsahují shelving filtry s těmito koeficienty a zdrojovým kódem:

```

% shelf filtry

    m = tan(pi*fc/48000);

% high-frequency boost k1

    G1 = 1.76;
    k1 = 10^(G1/20);
    jmenovatel1 = 1 + sqrt(2)*m + m^2;
    b01 = (k1 + sqrt(2*k1)*m + m^2) / jmenovatel1;
    b11 = (2*(m^2 - k1)) / jmenovatel1;
    b21 = (k1 - sqrt(2*k1)*m + m^2) / jmenovatel1;
    a11 = (2*(m^2 - 1)) / jmenovatel1;
    a21 = (1 - sqrt(2)*m + m^2) / jmenovatel1;

    b1 = [b01 b11 b21];
    a1 = [1 a11 a21];

% high-frequency cut k2

    G2 = 1.25;
    k2 = 10^(-G2/20);
    jmenovatel2 = k2 + sqrt(2*k2)*m + m^2;
    b02 = (1 + sqrt(2)*m + m^2) / jmenovatel2;
    b12 = (2*(m^2 - 1)) / jmenovatel2;
    b22 = (1 - sqrt(2)*m + m^2) / jmenovatel2;
    a12 = (2*(m^2 - k2)) / jmenovatel2;
    a22 = (k2 - sqrt(2*k2)*m + m^2) / jmenovatel2;

    b2 = [b02 b12 b22];
    a2 = [1 a12 a22];

% vysledek

    WF = filter(b1,a1, W);
    XF = filter(b2,a2, X);
    YF = filter(b2,a2, Y);

```

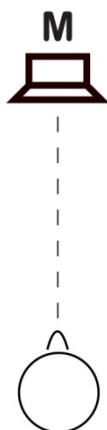
Funkce pro výpočet nepravidelných polí, tedy soubory monoUHJ.m, stereoUHJ.m, monoW.m, stereoXY.m, stereoMS.m, stereoS.m a ITU51.m, shelving filtry neobsahují.

10.1.1 Výpočty výstupních signálů jednotlivých reproduktorových polí

V dalším textu budou uvedeny konkrétní zdrojové kódy z MATLABu pro výpočet výstupních signálů určených pro reproduktory v jednotlivých reproduktorových soustavách, které jsou obsaženy v odpovídajících funkcích, respektive jejich m-souborech. U zdrojových kódů jsou také vyobrazeny samotná reproduktorová pole kvůli lepší představě rozmístění jednotlivých reproduktorů vzhledem k posluchači a pochopení zkratkovitého popisu proměnných nesoucích příslušné signály. Jako první budou uvedeny funkce pro nesymetrická pole.

Mono UHJ

```
M = SUMA;
```



Obr. 10.2: Mono systém.

Stereo UHJ

$$L = (SUMA + DELTA)/2;$$

$$R = (SUMA - DELTA)/2;$$

Proměnné SUMA a DELTA reprezentují první a druhou složku UHJ, tedy Σ a Δ (jejich výpočet viz vztahy (4.5) a (4.6) z kapitoly 4).

Mono

$$M = \text{sqrt}(2) * W;$$

Jediný reproduktor tohoto mono systému tak bude obsahovat pouze původní akustický signál snímáný tlakovým všesměrovým mikrofonom. Rozmístění stejné jako na obrázku 10.2.

Stereo XY

$$L = (X + Y)/\text{sqrt}(2);$$

$$R = (X - Y)/\text{sqrt}(2);$$

Stereo MS

$$M = (W * \text{sqrt}(2)) + X;$$

$$L = M + Y;$$

$$R = M - Y;$$

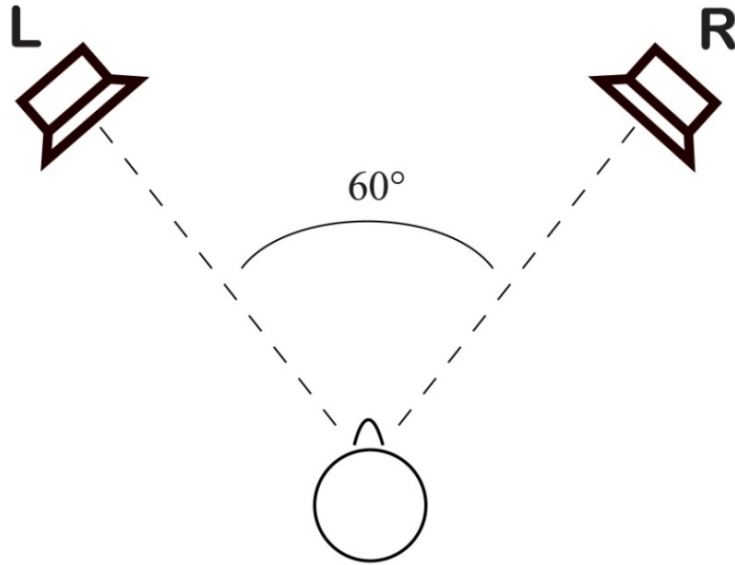
Systémy dekódování stereo XY a stereo MS jsou založeny na příslušných stereofonních snímacích mikrofonních technikách (viz například [1]).

Super Stereo

$$L = (0.0928 + i*0.255)*X + (0.4699 - i*0.171)*W + (0.3277)*Y;$$

$$R = (0.0928 - i*0.255)*X + (0.4699 + i*0.171)*W - (0.3277)*Y;$$

Systém Super stereo má podobné vlastnosti jako Stereo UHJ, jeho dekódovací rovnice jsou tedy založeny na UHJ (viz [16]). Umístění reproduktorů je shodné u všech Stereo systémů a je zobrazeno na obrázku 10.3.



Obr. 10.3: Stereo systém.

5.1 Surround Sound ITU-R BS.775

```
% numerical optimized Bruce Wiggins' coefficients (Tabu algorithm)
% tested and recommended by Richard Lee
```

```
L = (0.3928*W) + (0.4868*X) + (0.3450*Y);
R = (0.3928*W) + (0.4868*X) + (-0.3450*Y);
SL = (0.6570*W) + (-0.2890*X) + (0.4507*Y);
SR = (0.6570*W) + (-0.2890*X) + (-0.4507*Y);
```

```
% Classic ITU-R 5.1 BS.775 (angles +-30 and +-110)
% Tabu optimized
```

```
% L = (0.5018*W) + (0.6218*X) + (0.4406*Y);
% R = (0.5018*W) + (0.6218*X) + (-0.4406*Y);
% SL = (0.8392*W) + (-0.3692*X) + (0.5757*Y);
% SR = (0.8392*W) + (-0.3692*X) + (-0.5757*Y);
```

```
% Real World 5.1 Systems (angles +-30 and +-110) - almost square
% Tabu optimized
```

```
% L = (0.5018*W) + (0.6218*X) + (0.4406*Y);
% R = (0.5018*W) + (0.6218*X) + (-0.4406*Y);
% SL = (0.4465*W) + (-0.1964*X) + (0.3063*Y);
% SR = (0.4465*W) + (-0.1964*X) + (-0.3063*Y);
```

Všechny koeficienty pro výpočet signálů pro 5.1 Surround Sound ITU-R BS.775 byly zjištěny pomocí již zmiňovaného Tabu algoritmu [21]. Všimněte si, že centrální reproduktor se při ambisonii vůbec nepoužívá. Je to proto, že mívá v drtivé většině případů odlišnou konstrukci a vlastnosti oproti ostatním reproduktorům, což je nežádoucí (z důvodů uvedených v kapitole 8). V našem případě máme k dispozici tři sady koeficientů, které můžeme použít po „odkomentování“ příslušné čtveřice řádků. První čtveřice, která je „odkomentovaná“, obsahuje koeficienty otestované v literatuře [2]. Další čtveřice s nadpisem „Classic“ obsahuje koeficienty určené pro reproduktory v azimutech ± 30 a ± 110 stupňů a ležících na pomyslné kružnici

s posluchačem v jejím středu (stejně jako předchozí, ale nejsou otestovány – viz obr. 1.2). Poslední čtveřice označena „Real World 5.1 Systems“ je určena pro reproduktory ve stejných azimutech jako předchozí, tentokrát však neleží na pomyslné kružnici, ale blíže k posluchači, a to tak, že vytvářejí téměř čtvercové uspořádání. Tato varianta je obvyklá obzvláště v domácích podmínkách, kde například přední reproduktory jsou umístěny po stranách televizní obrazovky a ty surroundové jsou z důvodu nedostatku místa mnohem blíže k posluchači, nejsou tedy na pomyslné kružnici v určité vzdálenosti od posluchače. Dále pokračuje výčet zdrojových kódů pro výpočty pravidelných polí, které jsou založeny na vztahu (5.18).

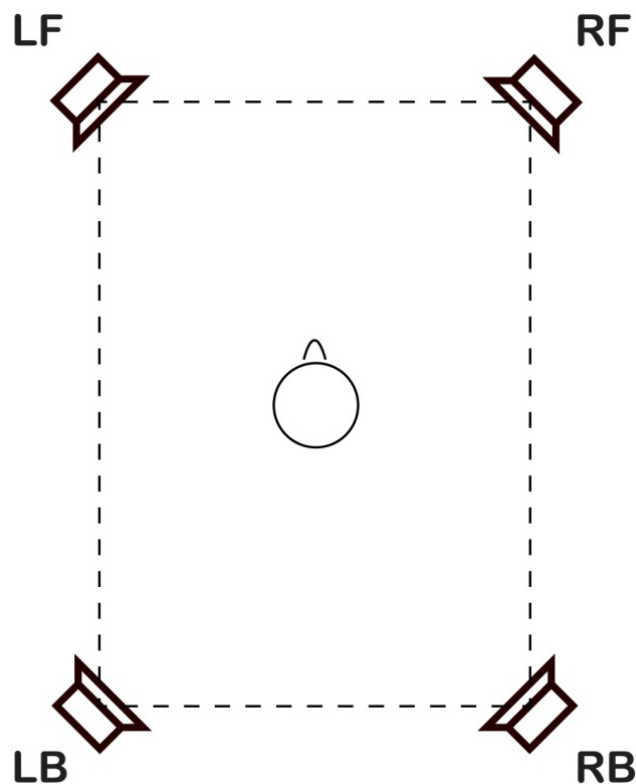
Rectangular a Rectangular UHJ

```
% loudspeaker at fi/2

uhelLF = (fi/2) * (pi/180);
uhelRF = -(fi/2)*(pi/180);
uhelLB = pi - ((fi/2)*(pi/180));
uhelRB = pi + ((fi/2)*(pi/180));

% loudspeaker feeds

LF = 0.5*(WF + ((1/sqrt(2))*cos(uhelLF)) * XF) +
      ((1/sqrt(2))*sin(uhelLF)) * YF);
RF = 0.5*(WF + ((1/sqrt(2))*cos(uhelRF)) * XF) +
      ((1/sqrt(2))*sin(uhelRF)) * YF);
LB = 0.5*(WF + ((1/sqrt(2))*cos(uhelLB)) * XF) +
      ((1/sqrt(2))*sin(uhelLB)) * YF);
RB = 0.5*(WF + ((1/sqrt(2))*cos(uhelRB)) * XF) +
      ((1/sqrt(2))*sin(uhelRB)) * YF);
```



Obr. 10.4: Rectangular (obdélník).

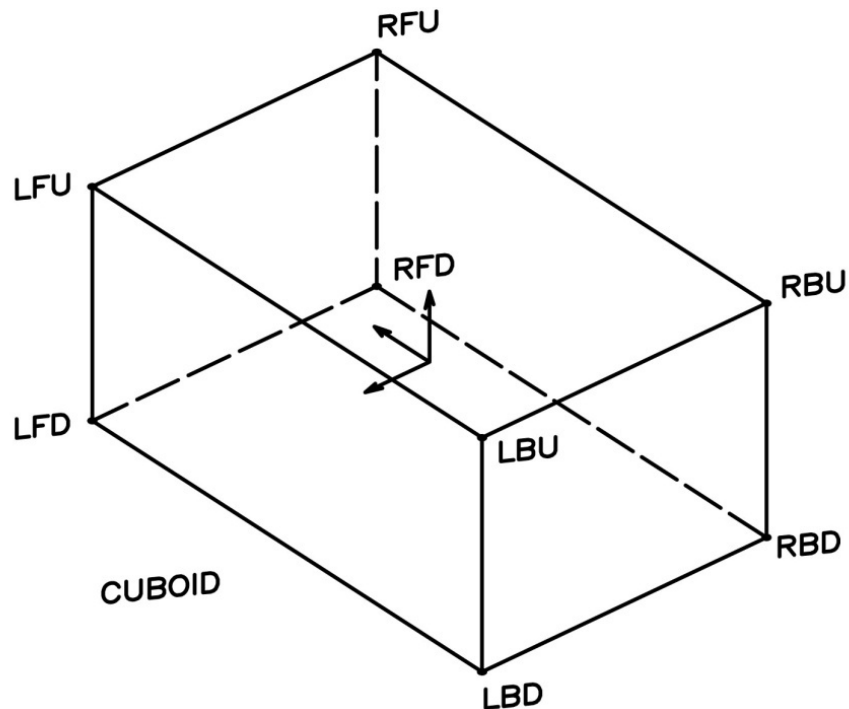
Rectangular 3D

`% loudspeaker angles`

```
azimLF = (fi/2) * (pi/180);  
azimRF = -(fi/2)*(pi/180);  
azimLB = pi - ((fi/2)*(pi/180));  
azimRB = pi + ((fi/2)*(pi/180));  
  
elevFD = -(eps/2)*(pi/180);  
elevFU = (eps/2) * (pi/180);  
elevBD = -(eps/2)*(pi/180);  
elevBU = (eps/2) * (pi/180);
```

`% loudspeaker feeds`

```
LFD = 0.5*((sqrt(2)*WF) + (cos(azimLF)*cos(elevFD)*XF) +  
          (sin(azimLF)*cos(elevFD)*YF) + (sin(elevFD)*ZF));  
RFD = 0.5*((sqrt(2)*WF) + (cos(azimRF)*cos(elevFD)*XF) +  
          (sin(azimRF)*cos(elevFD)*YF) + (sin(elevFD)*ZF));  
LBD = 0.5*((sqrt(2)*WF) + (cos(azimLB)*cos(elevBD)*XF) +  
          (sin(azimLB)*cos(elevBD)*YF) + (sin(elevBD)*ZF));  
RBD = 0.5*((sqrt(2)*WF) + (cos(azimRB)*cos(elevBD)*XF) +  
          (sin(azimRB)*cos(elevBD)*YF) + (sin(elevBD)*ZF));  
LFU = 0.5*((sqrt(2)*WF) + (cos(azimLF)*cos(elevFU)*XF) +  
          (sin(azimLF)*cos(elevFU)*YF) + (sin(elevFU)*ZF));  
RFU = 0.5*((sqrt(2)*WF) + (cos(azimRF)*cos(elevFU)*XF) +  
          (sin(azimRF)*cos(elevFU)*YF) + (sin(elevFU)*ZF));  
LBU = 0.5*((sqrt(2)*WF) + (cos(azimLB)*cos(elevBU)*XF) +  
          (sin(azimLB)*cos(elevBU)*YF) + (sin(elevBU)*ZF));  
RBU = 0.5*((sqrt(2)*WF) + (cos(azimRB)*cos(elevBU)*XF) +  
          (sin(azimRB)*cos(elevBU)*YF) + (sin(elevBU)*ZF));
```



Obr. 10.5: Rectangular 3D [kvádr, převzato z [9]].

Birectangle

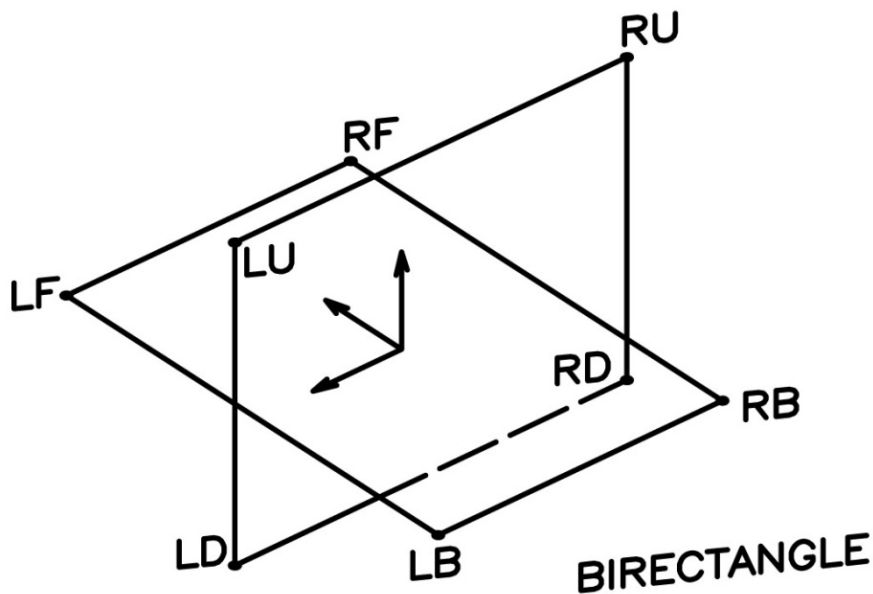
```
% loudspeaker angles
```

```
azimLF = (fi/2) * (pi/180);  
azimRF = -(fi/2)*(pi/180);  
azimLB = pi - ((fi/2)*(pi/180));  
azimRB = pi + ((fi/2)*(pi/180));
```

```
elevLU = (fi/2) * (pi/180);  
elevLD = -(fi/2)*(pi/180);  
elevRD = -(fi/2)*(pi/180);  
elevRU = (fi/2) * (pi/180);
```

```
% loudspeaker feeds
```

```
LF = 0.5*((sqrt(2)*WF) + (cos(azimLF)*cos(0)*XF) +  
         (sin(azimLF)*cos(0)*YF) + (sin(0)*ZF));  
RF = 0.5*((sqrt(2)*WF) + (cos(azimRF)*cos(0)*XF) +  
         (sin(azimRF)*cos(0)*YF) + (sin(0)*ZF));  
LB = 0.5*((sqrt(2)*WF) + (cos(azimLB)*cos(0)*XF) +  
         (sin(azimLB)*cos(0)*YF) + (sin(0)*ZF));  
RB = 0.5*((sqrt(2)*WF) + (cos(azimRB)*cos(0)*XF) +  
         (sin(azimRB)*cos(0)*YF) + (sin(0)*ZF));  
LU = 0.5*((sqrt(2)*WF) + (cos(pi/2)*cos(elevLU)*XF) +  
         (sin(pi/2)*cos(elevLU)*YF) + (sin(elevLU)*ZF));  
RU = 0.5*((sqrt(2)*WF) + (cos(-(pi/2))*cos(elevRU)*XF) +  
         (sin(-(pi/2))*cos(elevRU)*YF) + (sin(elevRU)*ZF));  
LD = 0.5*((sqrt(2)*WF) + (cos(pi/2)*cos(elevLD)*XF) +  
         (sin(pi/2)*cos(elevLD)*YF) + (sin(elevLD)*ZF));  
RD = 0.5*((sqrt(2)*WF) + (cos(-(pi/2))*cos(elevRD)*XF) +  
         (sin(-(pi/2))*cos(elevRD)*YF) + (sin(elevRD)*ZF));
```

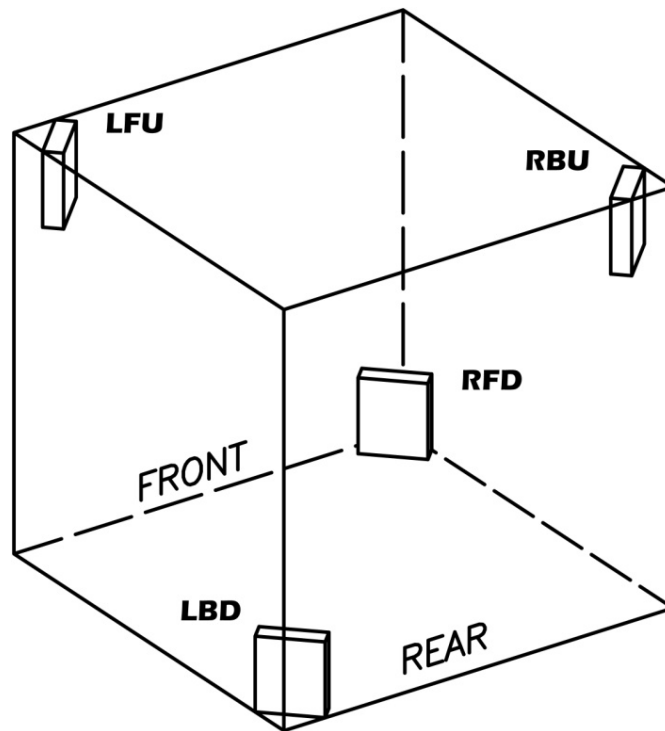


Obr. 10.6: Birectangle (převzato z [9]).

Tetrahedron

```
% loudspeaker feeds
```

```
LFU = 0.5*((sqrt(2)*WF) + (cos(pi/4)*cos(pi/4)*XF) +  
          (sin(pi/4)*cos(pi/4)*YF) + (sin(pi/4)*ZF));  
RFD = 0.5*((sqrt(2)*WF) + (cos(-pi/4)*cos(-pi/4)*XF) +  
          (sin(-pi/4)*cos(-pi/4)*YF) + (sin(-pi/4)*ZF));  
LBD = 0.5*((sqrt(2)*WF) + (cos(pi-(pi/4))*cos(-pi/4)*XF) +  
          (sin(pi-(pi/4))*cos(-pi/4)*YF) + (sin(-pi/4)*ZF));  
RBU = 0.5*((sqrt(2)*WF) + (cos(pi+(pi/4))*cos(pi/4)*XF) +  
          (sin(pi+(pi/4))*cos(pi/4)*YF) + (sin(pi/4)*ZF));
```



Obr. 10.7: Tetrahedron (čtyřstěn, převzato z [9]).

Octahedron

```
% loudspeaker angles
```

```
elevFU = (eps/2) * (pi/180);  
elevFD = -(eps/2)*(pi/180);  
elevBU = (eps/2) * (pi/180);  
elevBD = -(eps/2)*(pi/180);
```

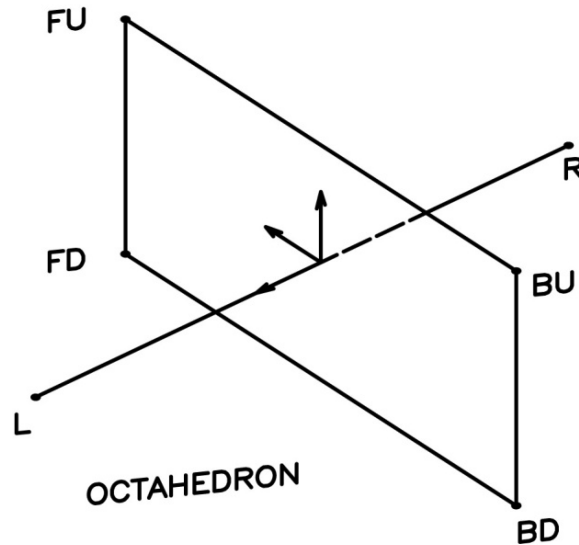
```
% loudspeaker feeds
```

```
FU = 0.5*((sqrt(2)*WF) + (cos(0)*cos(elevFU)*XF) +  
          (sin(0)*cos(elevFU)*YF) + (sin(elevFU)*ZF));  
FD = 0.5*((sqrt(2)*WF) + (cos(0)*cos(elevFD)*XF) +  
          (sin(0)*cos(elevFD)*YF) + (sin(elevFD)*ZF));
```

```

BU = 0.5*((sqrt(2)*WF) + (cos(pi)*cos(elevBU)*XF) +
          (sin(pi)*cos(elevBU)*YF) + (sin(elevBU)*ZF));
BD = 0.5*((sqrt(2)*WF) + (cos(pi)*cos(elevBD)*XF) +
          (sin(pi)*cos(elevBD)*YF) + (sin(elevBD)*ZF));
L  = 0.5*((sqrt(2)*WF) + (cos(pi/2)*cos(0)*XF) +
          (sin(pi/2)*cos(0)*YF) + (sin(0)*ZF));
R  = 0.5*((sqrt(2)*WF) + (cos(-pi/2)*cos(0)*XF) +
          (sin(-pi/2)*cos(0)*YF) + (sin(0)*ZF));

```



Obr. 10.8: Octahedron (osmistěn, převzato z [9]).

Proměnná f_i má význam horizontálního úhlu mezi předními reproduktory pole vzhledem k posluchači (tedy například 60° pro stereo soustavy). Používá se kurčení „referenčního“ azimutu předního levého reproduktoru, který má azimut $f_i/2$, a tím také všech ostatních (výpočet příslušných proměnných azimXX). Proměnná eps má stejný význam, pouze ve vertikálním směru (určuje elevace reproduktorů v proměnných $elevXX$). Obě veličiny mohou z praktického hlediska nabývat hodnot od 0° do 180° . Větší hodnoty by vedly k prohazování pozic reproduktorů (tedy nemají smysl). Hodnoty f_i i eps zadává uživatel v grafickém rozhraní pro výpočet konkrétního pole.

Pravidelný polygon a Pravidelný polygon UHJ

```

n = nargout;
a = (2*pi)/n ;

% loudspeaker at 0 rad
%   for i=0:(n-1)
%       uhell(i+1) = i*a;
%   end

% loudspeaker at 0 + a/2 rad

b = a/2;
for i=0:(n-1)
    uhel(i+1,1) = b + (i*a);
end

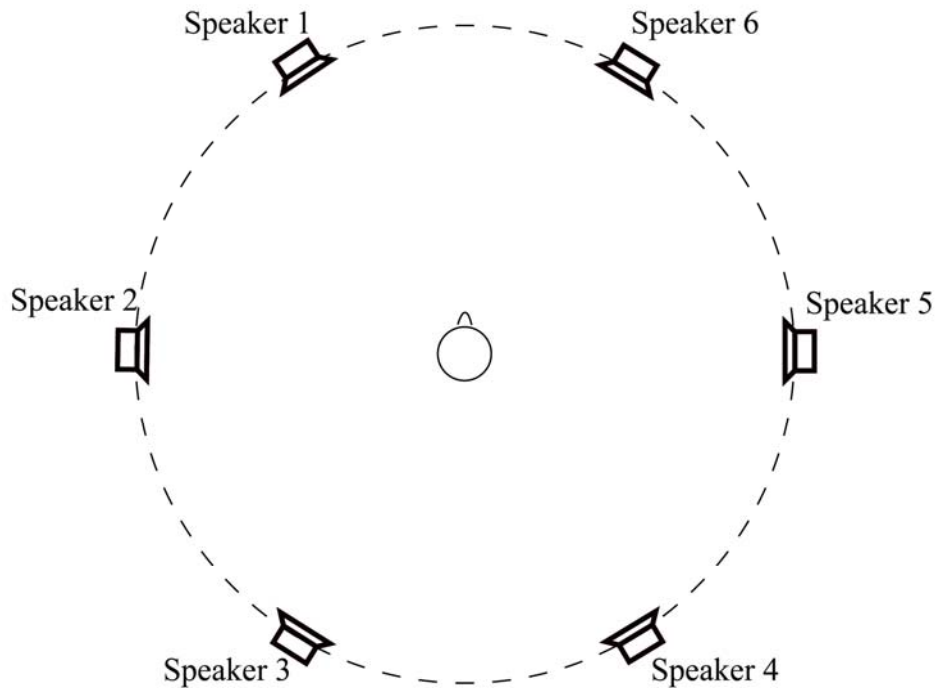
```

```

% loudspeaker feeds

for k=1:n
    varargout{k}= 0.5*((sqrt(2)*WF) + (cos(uhel(k,1))*XF) +
                    (sin(uhel(k,1))*YF));
end

```



Obr. 10.9: Pravidelný polygon (UHJ).

Pravidelný polygon 3D a Pravidelný polygon 3D UHJ

```

n = nargout;
a = (2*pi)/(n/2);

% loudspeaker at 0 rad
% for i=0:(n-1)
%     azim1(i+1) = i*a;
% end

% loudspeaker at 0 + a/2 rad

b = a/2;
for i=0:(n/2 - 1)
    azim(i+1,1) = b + (i*a);
end

% loudspeaker feeds
%     down loudspeakers

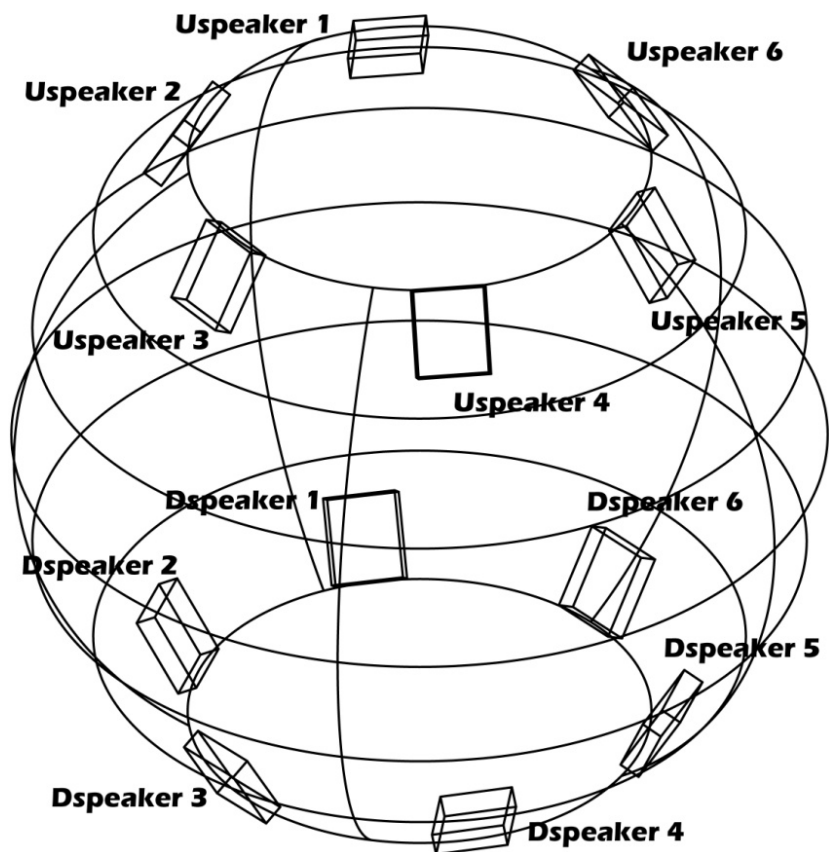
```

```

for k=1:(n/2)
    varargout{k}= 0.5*((sqrt(2)*WF) +
        (cos(azim(k,1))*cos((-eps/2)*(pi/180))*XF) +
        (sin(azim(k,1))*cos((-eps/2)*(pi/180))*YF) +
        (sin((-eps/2)*(pi/180))*ZF));
end

%     up loudspeakers
for k=((n/2)+1):n
    varargout{k}= 0.5*((sqrt(2)*WF) +
        (cos(azim(k-(n/2),1))*cos((eps/2)*(pi/180))*XF) +
        (sin(azim(k-(n/2),1))*cos((eps/2)*(pi/180))*YF) +
        (sin((eps/2)*(pi/180))*ZF));
end

```



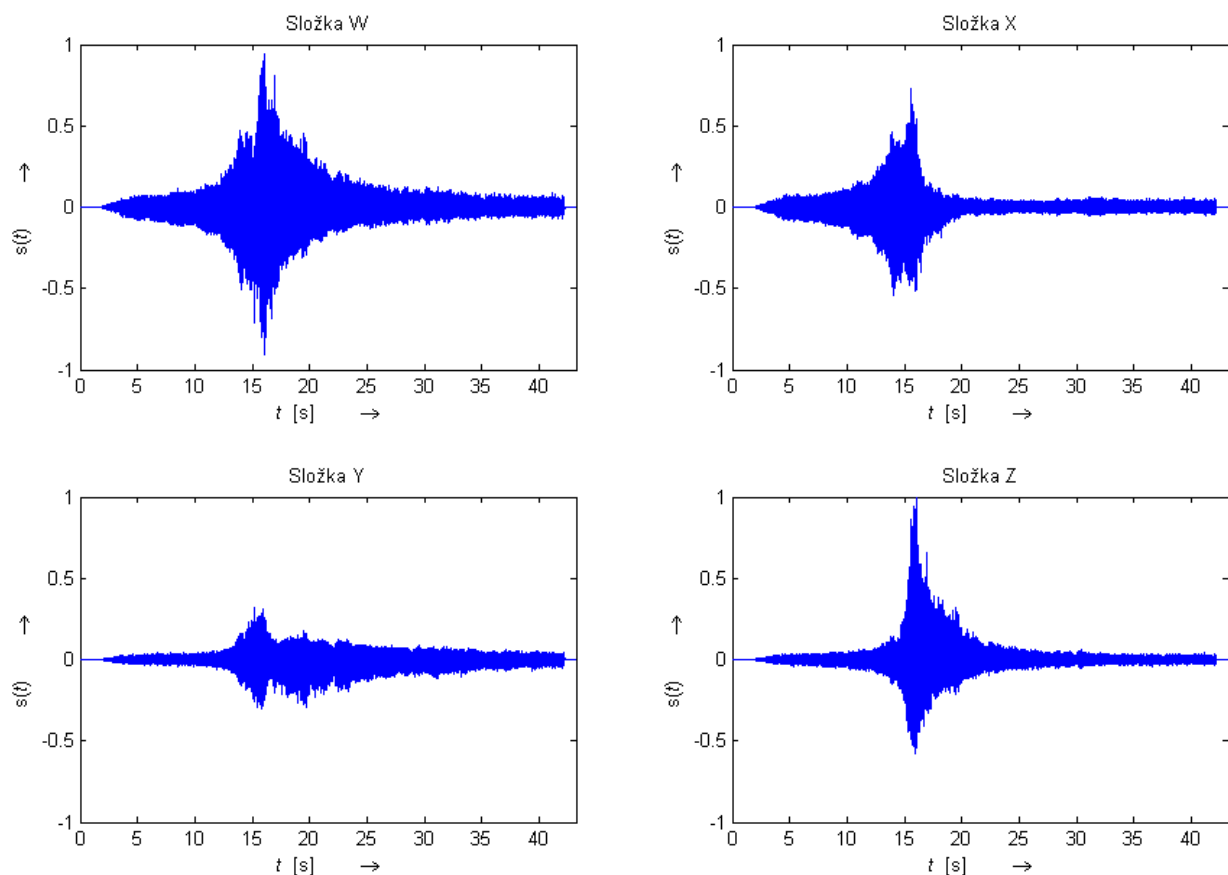
Obr. 10.10: Pravidelný polygon 3D (UH).

V předchozích rovnicích proměnná n reprezentuje počet reproduktorů pole, a je horizontální úhel mezi dvěma předními reproduktory (závisí na počtu reproduktorů), b je azimut prvního reproduktoru (předního levého), $azim$ a $uhel$ jsou matice azimutů všech reproduktorů a eps má stejný význam jako výše. V proměnných $varargout\{k\}$ (kde $k = 1, 2, \dots, n$) jsou uloženy

výstupní signály, přičemž v Pravidelném polygonu odpovídá číslo k proměnné `varargout` přímo číslo reproduktoru (s popisem `Speaker(k)`). U Pravidelného polygonu 3D se počet proměnných `varargout` rozdělí na dvě poloviny. První polovina odpovídá příslušným reproduktorům se zápornou elevací (dolní reproduktory `Dspeaker(k)`) a druhá polovina pak odpovídá reproduktorům s kladnou elevací (`Uspeaker(k)`), jak je naznačeno v obrázku 10.10.

10.2 Typy zobrazení signálů v aplikaci B-format Dekodér

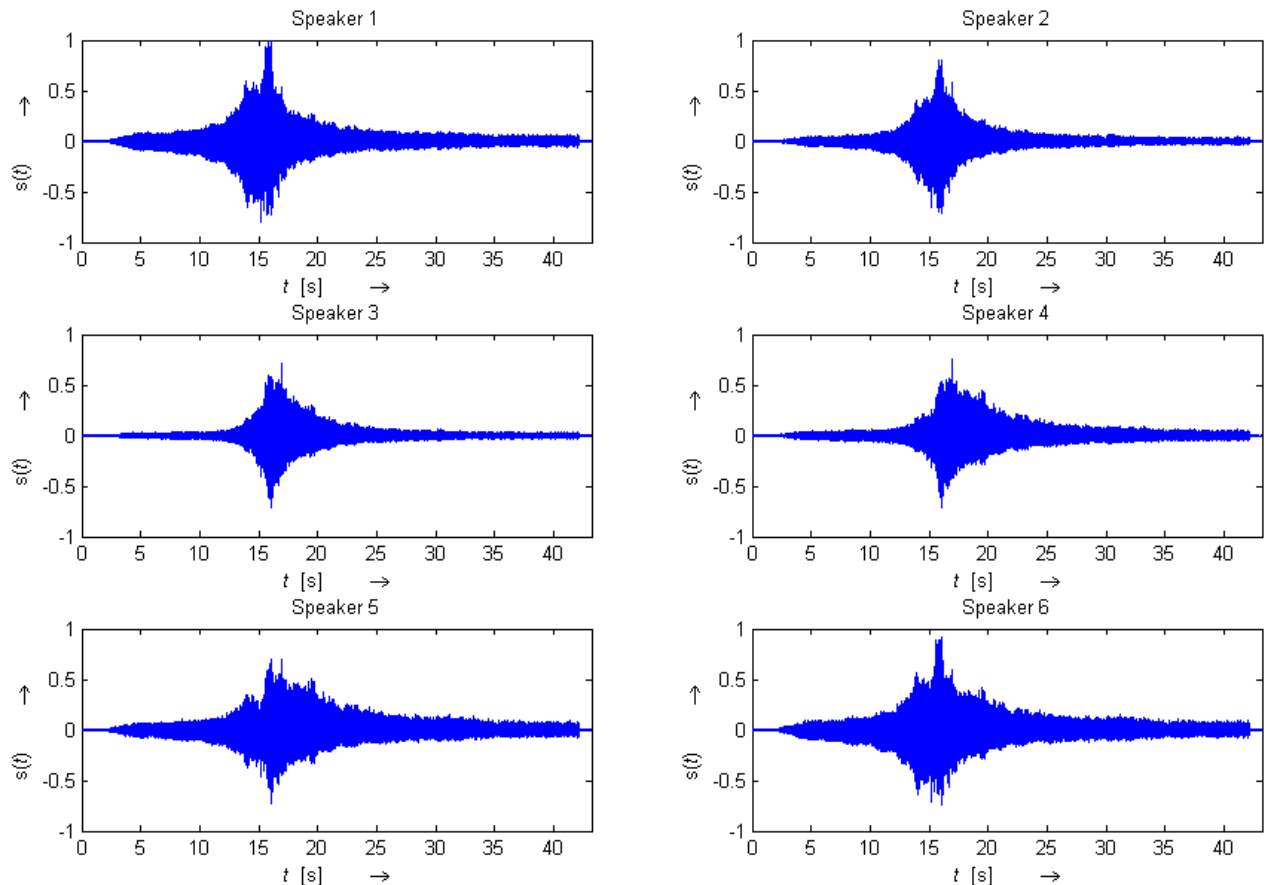
Pro zobrazení vstupních a výstupních signálů máme v tomto programu k dispozici několik druhů jejich závislostí na různých parametrech. V případě signálů vstupních (tedy zdrojových složek B-formátu) můžeme zobrazit závislosti jejich intenzity na čase nebo jejich spektrogramy, což je časový sled krátkodobých spekter signálu. Velikost intenzity jednotlivých signálů je vztahena k velikosti vzorku s největší absolutní hodnotou intenzity ze všech čtyř složek. Příklad časových průběhů testovacích zdrojových složek B-formátu ze souboru `bformatAERO.mat` je uveden na obrázku 10.11. Jednotlivé náhledy grafů jednoho pole budou v textu zobrazeny pohromadě, aby bylo lépe vidět jejich vzájemné souvislosti.



Obr. 10.11: Náhled časových průběhů testovacích složek B-formátu.

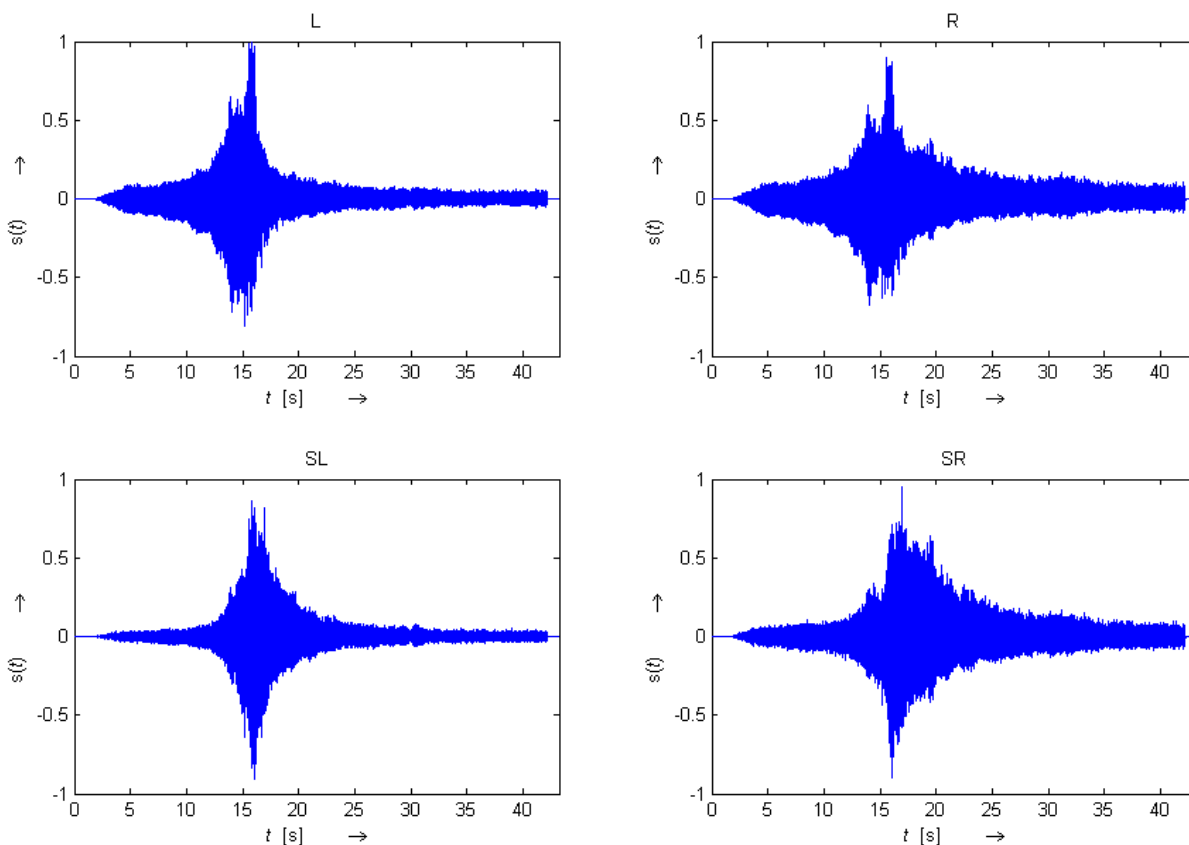
Jde o akustický signál zvuku motoru letadla nasnímaný mikrofonom soundfield, získaný ze zdroje [20]. Z jednotlivých průběhů můžeme například odhadnout, že letadlo přilétalo k mikrofону v ose x (zepředu nebo zezadu), v prvních cca 12ti sekundách je totiž signál složky X nejsilnější, v 15ti sekundách pravděpodobně přelétlo nad mikrofonom (nejsilnější signál složky Z) atd.

Výstupní signály pro reproduktory v konkrétních reproduktorových polích můžeme zobrazit opět v podobě již zmíněných časových průběhů a spektrogramů. Kromě toho však navíc máme k dispozici obecné závislosti intenzity signálu jednotlivých reproduktorů na azimutu nebo zároveň na azimutu i elevaci zdroje zvuku, zobrazené dvoj i trojdimenzionálně. Na obrázku 10.12 jsou časové průběhy výstupních signálů pro Pravidelný polygon tvořený šesti reproduktory. Hodnoty intenzity jsou vždy vztaženy k velikosti vzorku s největší absolutní hodnotou intenzity ze všech výstupních signálů.

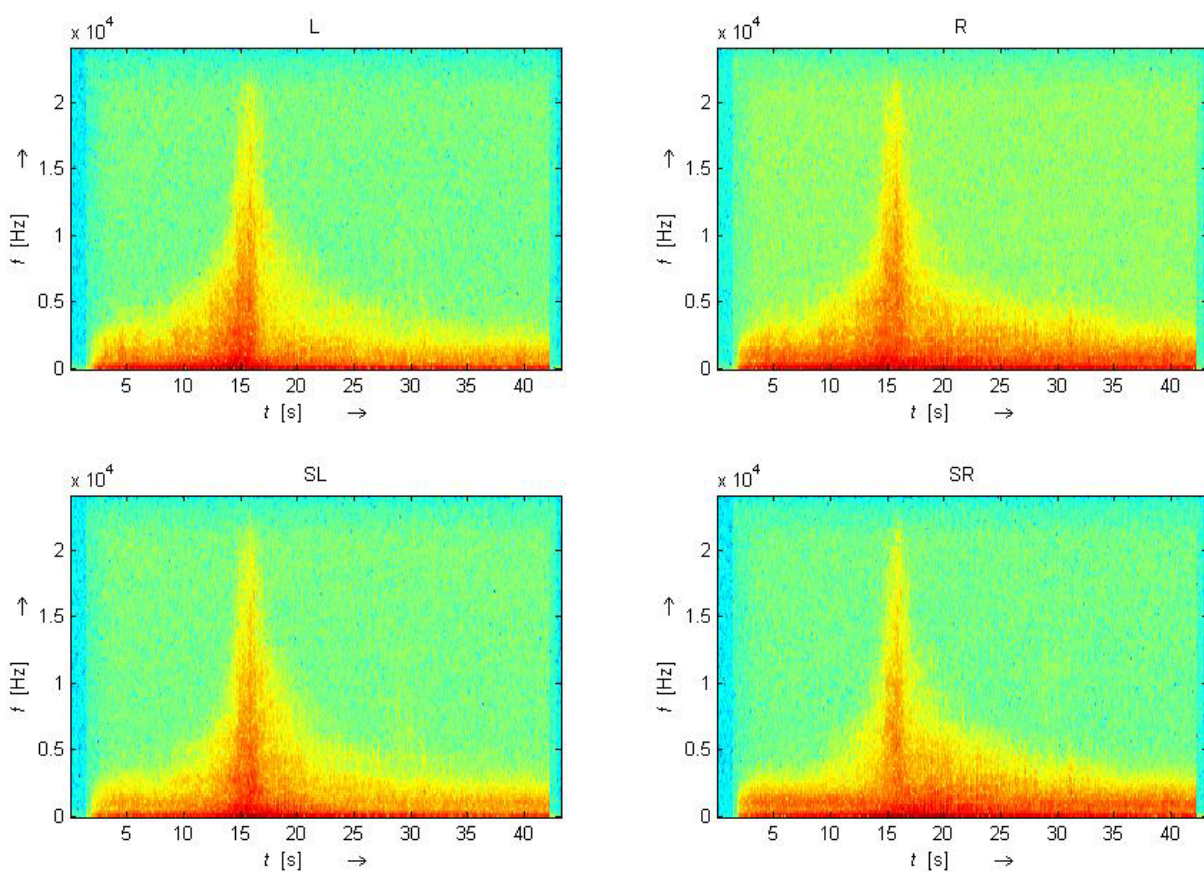


Obr. 10.12: Náhled časových průběhů výstupních signálů Pravidelného polygonu (šestiúhelníka).

Opět můžeme pozorovat, že největších hodnot dosahuje intenzita signálů reproduktorů 1, 6 a od cca 15té sekundy i reproduktoru 5 (to znamená přibližnou představu o poloze zdroje zvuku). Na dalším obrázku 10.13 můžeme pozorovat časové průběhy výstupních signálů reproduktorového pole podle specifikace 5.1 Surround Sound ITU-R BS.775. Obrázek 10.14 pak znázorňuje jejich spektrogramy. Z časových průběhů je opět vidět stejná přibližná pozice zdroje zvuku v průběhu jeho trvání. Ve spektrogramech sledujeme, že většina energie výstupních signálů vypočtených ze složek B-formátu v souboru `bformatAERO.mat` je obsažena v kmitočtech do cca 5 000 Hz. Pouze v čase 15 sekund, kdy byl zdroj zvuku pravděpodobně nejbližší mikrofonu, se na chvíli zvýší i energie na kmitočtech do 20 kHz.

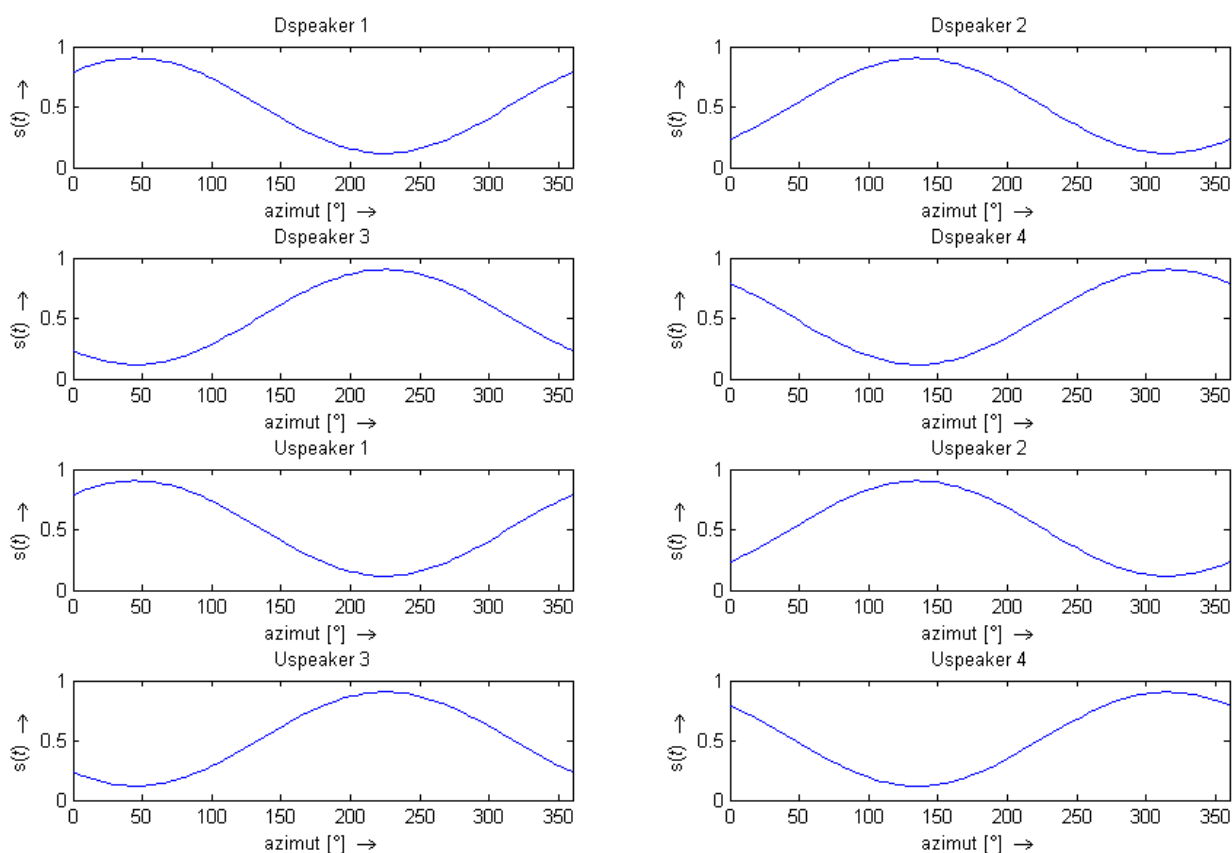


Obr. 10.13: Náhled časových průběhů výstup. signálů pole 5.1 Surround Sound ITU-R BS.775.



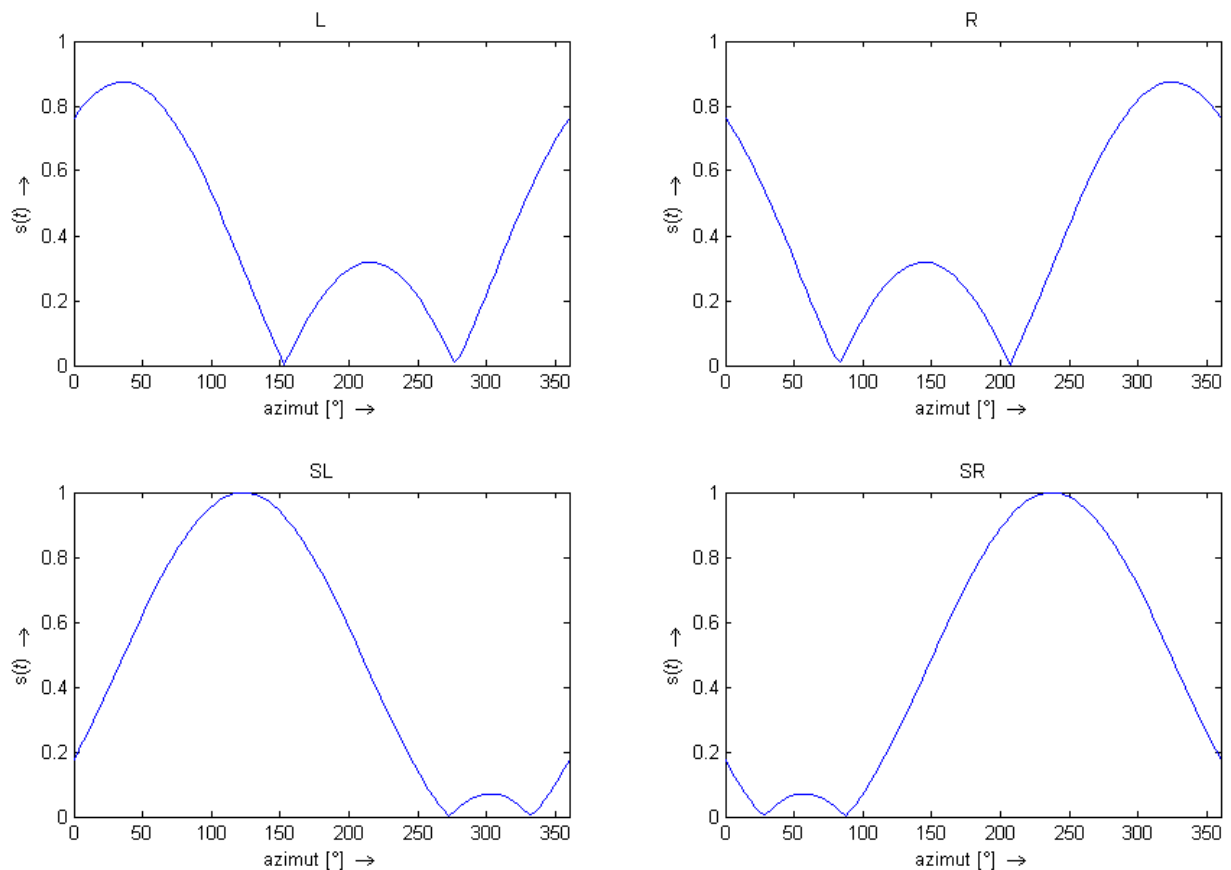
Obr. 10.14: Náhled spektrogramů signálů pole 5.1 Surround Sound ITU-R BS.775.

Další zobrazení, závislost intenzity výstupního signálu na azimutu zdroje zvuku, umožňuje sledovat, jak se změna azimutu zdroje projeví na intenzitě signálu reproduktoru v určitém umístění v prostoru pole. Obrázek 10.15 obsahuje náhledy těchto závislostí pro Pravidelný polygon 3D.



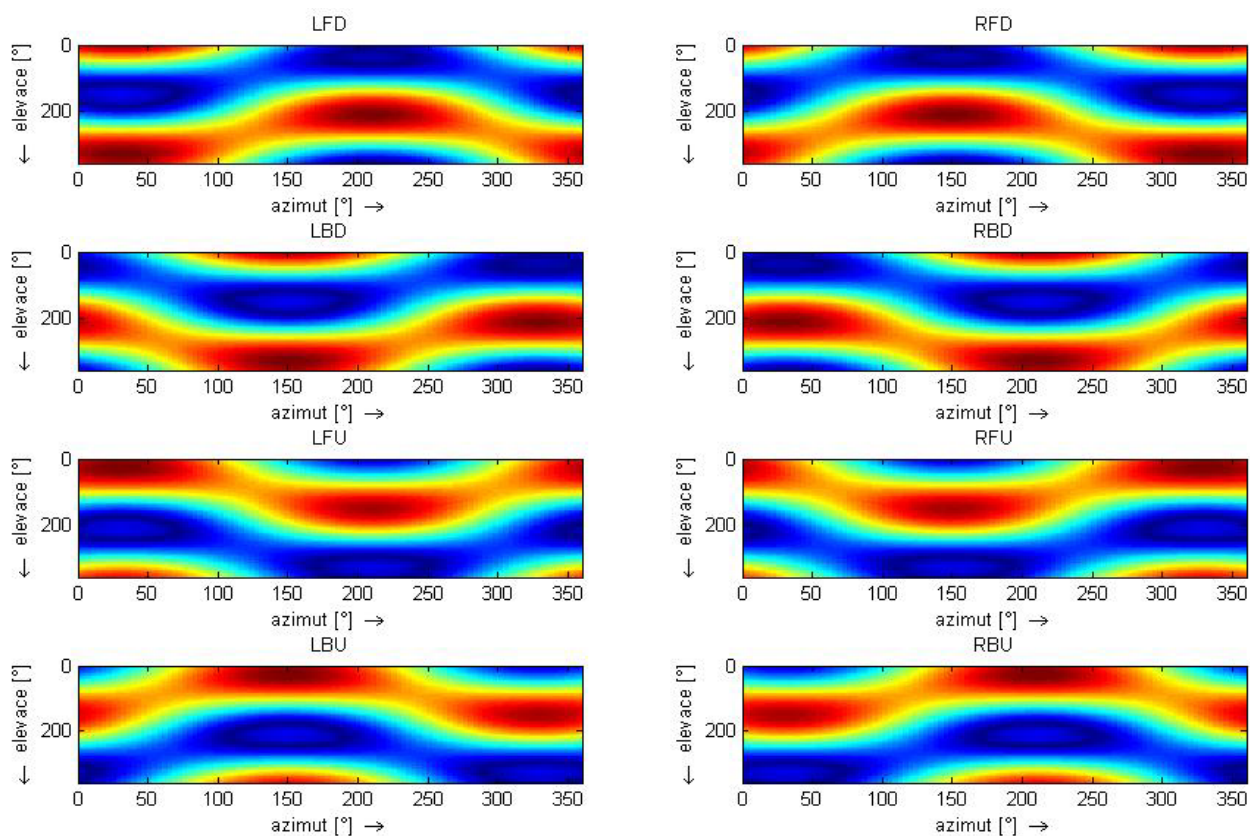
Obr. 10.15: Náhled závislostí intenzity signálů reproduktorů Pravidelného polygonu 3D s osmi reproduktory (a vertikálním úhlem 60°) na azimutu zdroje.

Z těchto náhledů vidíme, že intenzita výstupního signálu každého reproduktoru má nejvyšší hodnotu vždy, když se azimut zdroje rovná azimutu umístění reproduktoru v poli. To znamená například pro první dolní reproduktor (Dspeaker1) azimut 45° (dohromady je totiž 8 reproduktorů, 4 dolní a 4 horní, takže dva pravidelné čtyřúhelníky – čtverce). Dále si můžeme všimnout, že intenzita nikdy nedosáhne své maximální hodnoty (tedy zde hodnoty 1). To je způsobeno tím, že mezi reproduktory horního a dolního čtverce byl nastaven vertikální úhel 60° (reproduktory jsou v elevacích $\pm 30^\circ$ vzhledem k posluchači). Zdroj přitom mění svůj azimut, ale jeho elevace je stále nulová – proto se zdroj v tomto případě nikdy nedostane přesně do umístění některého z reproduktorů (ty mají totiž elevaci nenulovou), což by znamenalo maximální úroveň intenzity. Samozřejmě čím menší bude vertikální úhel (elevace reproduktorů), tím budou intenzity dosahovat vyšších hodnot. To ale v tomto případě není našim cílem, znamenalo by to zhoršení prostorové iluze třetího rozměru. Stejný typ náhledu pro pole 5.1 Surround Sound ITU-R BS.775 je na následujícím obrázku.

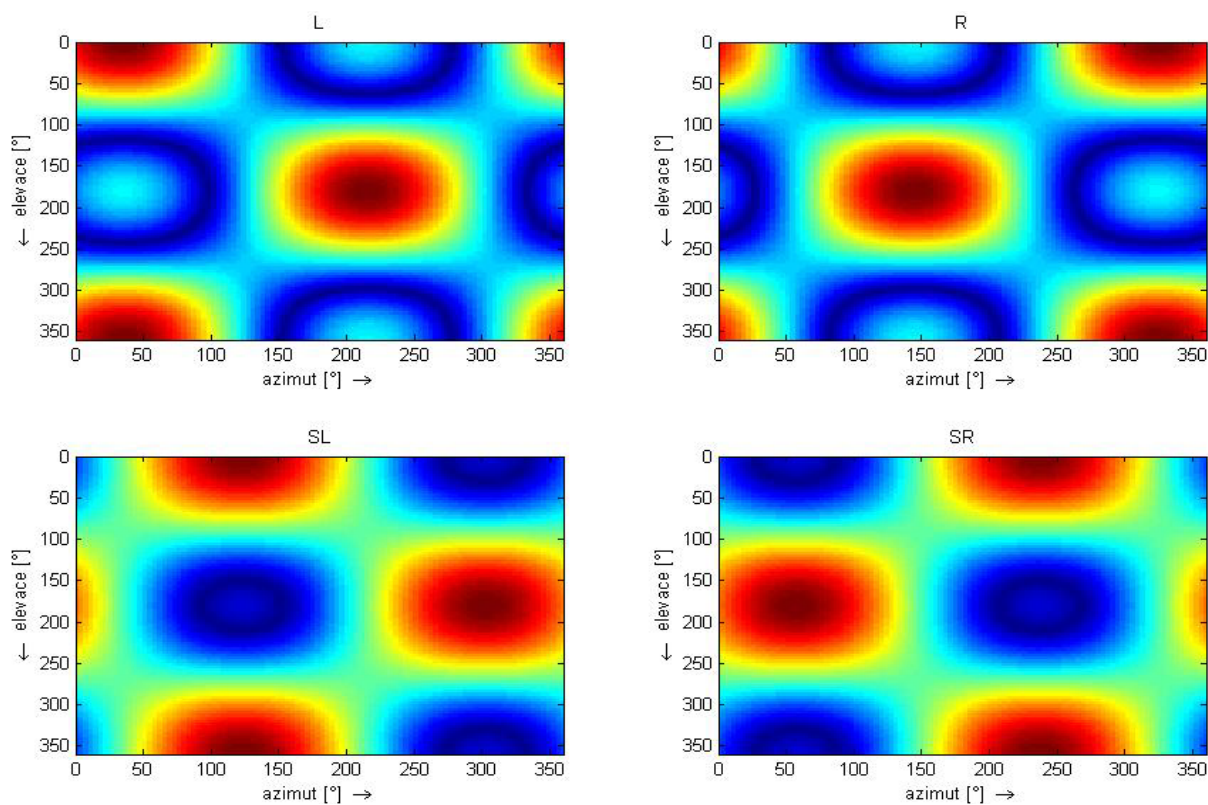


Obr. 10.16: Náhled závislostí intenzity signálů reproduktorů 5.1 Surround Sound ITU-R BS.775 na azimutu zdroje.

Zde křivky závislostí nejsou tolik plynulé, protože se jedná o nesymetrické pole. Platí ale stejné závěry, jako jsme pozorovali u předchozího pole. Reproductory v azimutech lišících se pouze znaménkem mají charakteristiky zrcadlově převrácené (stejně jako v předchozím případě). Tentokrát však intenzita dosahuje svého maxima, reproduktory mají nulovou elevaci (jedná se pouze o horizontální reprodukci). Závislost intenzity výstupního signálu na azimutu i elevaci zdroje zvuku, umožňuje sledovat, jak se změna azimutu a elevace zdroje projeví na intenzitě signálu reproduktoru v určitém umístění v prostoru pole. Tuto závislost je možné zobrazit 2D i 3D. Dvojdímní verze náhledu pro pole Rectangular 3D (s horizontálním i vertikálním úhlem 60°) je zobrazena na obrázku 10.17. Velikost intenzity je v obrázku reprezentována barvou, přičemž tmavě modrá znamená nejnižší hodnotu a tmavě červená hodnotu nejvyšší. Maxima sledujeme opět v okolí místa, kde se pozice zdroje zvuku blíží pozici konkrétního reproduktoru. Například pro levý přední spodní reproduktor (LFD – azimut 30° , elevace -30°) jsou tedy maxima při azimutu 30° a elevaci 330° nebo azimutu 210° a elevaci 210° (jinými slovy stejná pozice popsána jinými souřadnicemi). Opět můžeme pozorovat zrcadlově převrácené charakteristiky dvojic reproduktorů s azimuty nebo elevacemi, lišícími se jen ve znaménku. Obrázek 10.18 ukazuje náhledy horizontální reprodukce 5.1 Surround Sound ITU-R BS.775.

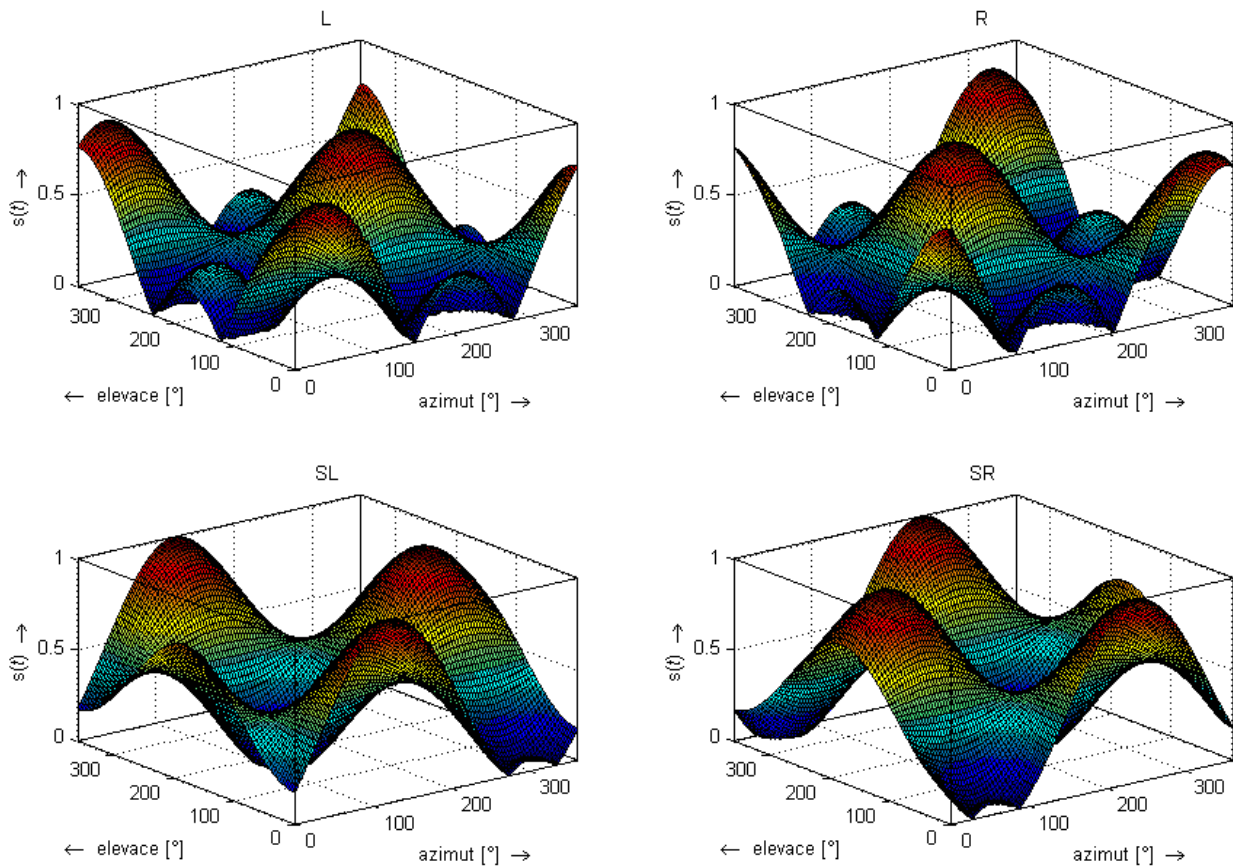


Obr. 10.17: Náhled 2D závislosti intenzity signálů reproduktorů pole Rectangular 3D (horizontální i vertikální úhel 60°) na azimutu a elevaci zdroje zvuku.



Obr. 10.18: Náhled 2D závislosti intenzity signálů 5.1 Surround Sound ITU-R BS.775 na azimutu a elevaci zdroje zvuku.

V tomto případě je intenzita největší v příslušných azimutech (podle umístění reproduktoru) a současně v elevacích 0° nebo 180° . Pro doplnění jsou na obrázku 10.19 uvedeny ještě 3D verze předchozích závislostí.



Obr. 10.19: Náhled 3D závislostí intenzity signálů 5.1 Surround Sound ITU-R BS.775 na azimutu a elevaci zdroje zvuku.

11 ZÁVĚR

V této práci byly popsány všechny kroky při zpracování akustického signálu pomocí vícekanálového zvukového systému Ambisonie od snímání pomocí mikrofону typu soundfield, až po reprodukci výsledných akustických signálů konkrétních reproduktorových polí. Dále byla popsána technologie označovaná B-format, která je jeho hlavní součástí. Byly uvedeny způsoby konverze akustických signálů do složek B-formátu i způsoby jeho dekodování do ostatních formátů, které ambisonie využívá, nebo přímo do cílových signálů, určených k reprodukci.

Pro možnost zpracování zdrojových složek B-formátu a produkci výsledných signálů pro různé typy reproduktorových polí byla vytvořena aplikace B-format Dekodér. Pro tuto aplikaci bylo z důvodu snadné ovladatelnosti vytvořeno vlastní grafické uživatelské rozhraní, kterým lze ovládat všechny její funkce. Kromě výše zmíněného kódování a dekodování s nastavením volitelných parametrů převodu lze také zdrojové i výsledné signály analyzovat pomocí zobrazení několika druhů závislostí jejich energie na různých parametrech. Podrobněji jsou jednotlivé typy analýzy spolu s případnou reprezentací jejich výsledků probírány v příslušné kapitole. Konečným produktem aplikace potom mohou být výstupní signály jednotlivých reproduktorů v požadovaném reproduktorovém systému, uložené ve formátu **.mat* nebo formátu **.wav*, připravené k reprodukci tímto polem. V praxi je pak možné s jejich pomocí experimentálně ověřit platnost závěrů zjištěných při analýze konkrétních reprodukcí soustav.

Po kompletním sestavení mikrofónu soundfield bude v budoucnu tato aplikace zpracovávat i akustické signály nasnímané touto soustavou mikrofónů, zatím jsou ale k dispozici pouze testovací nahrávky společnosti SoundField Ltd. získané z internetu, které najdete spolu s programem na přiloženém CD.

SEZNAM LITERATURY

- [1] BEHENG, Dieter. Stereo recording techniques. *Deutsche Welle Radio Training Centre* 2002. Dostupný z WWW:
<http://www9.dw-world.de/rtc/infotheque/stereo/stereo_recording.pdf>.
- [2] BENJAMIN, Eric M; LEE, Richard; HELLER Aaron J. Localization in Horizontal-Only Ambisonic Systems. *Corrected version of the convention paper presented at the 121st Convention of the Audio Engineering Society, San Francisco , October 2006*. Dostupný z WWW:
<<http://www.ai.sri.com/pubs/files/1374.pdf>>.
- [3] CAPRA, Andrea; FONTANA, Simone; ADRIAENSEN, Fons; FARINA, Angelo; GRENIER, Yves. Listening tests of the localization performance of Stereodipole and Ambisonic systems. *Convention paper presented at the 123rd Convention of the Audio Engineering Society, New York, October 2007*. Dostupný z WWW:
< <http://www.kokkinizita.net/papers/AES-123-conv.pdf> >.
- [4] DIGENIS, Aristotel. *The Implementation Of Ambisonics For Restoring Quadraphonic Recordings*. [s.l.], 2002. 88 s. Middlesex University/SAE. Bakalářská práce. Dostupný z WWW:
<<http://www.digenis.co.uk/documents.php>>.
- [5] DOŇAR, B., ZAPLATÍLEK, K. *MATLAB pro začátečníky*, 1. díl. Ben, technická literatura, Praha. ISBN 80-7300-175-6
- [6] DOŇAR, B., ZAPLATÍLEK, K. *MATLAB tvorba uživatelských aplikací*, 2. díl. Ben, technická literatura, Praha. ISBN 80-7300-133-0
- [7] ELEN, Richard. *Ambisonic.net* [online]. 1998-2007 [cit. 2008-11-16]. Dostupný z WWW:
<<http://www.ambisonic.net/>>.
- [8] FARINA, Angelo; UGOLOTTI, Emanuele. Software implementation of B-Format encoding and decoding. *Convention paper presented at the 104th Convention of the Audio Engineering Society*, April 1998. Dostupný z WWW:
< http://www.acoustics.net/objects/pdf/article_farina01.pdf>.
- [9] GERZON, Michael A. Ambisonics in multichannel broadcasting and video. In RUMSEY, Francis. *SPATIAL SOUND TECHNIQUES PART 2 : Multichannel Audio Technologies*. [s.l.] : [s.n.], 2006. s. 101-113.
- [10] GERZON, Michael A. Ambisonics, Part two: *Studio techniques*. *Studio sound*, August 1975. Dostupný z WWW:
<<http://www.michaelgerzonphotos.org.uk/articles/Ambisonics%202.pdf>>.
- [11] GERZON, Michael A. Criteria for evaluating surround-sound systems. In RUMSEY, Francis. *SPATIAL SOUND TECHNIQUES PART 2 : Multichannel Audio Technologies*. [s.l.] : [s.n.], 2006. s. 77-85.

- [12] GERZON, Michael A. Periphony: With-height sound reproduction. In RUMSEY, Francis. *SPATIAL SOUND TECHNIQUES PART 2 : Multichannel Audio Technologies*. [s.l.] : [s.n.], 2006. s. 92-100.
- [13] GERZON, Michael A. Surround-sound psychoacoustics. In RUMSEY, Francis. *SPATIAL SOUND TECHNIQUES PART 2 : Multichannel Audio Technologies*. [s.l.] : [s.n.], 2006. s. 86-91. Dostupný z WWW:
<http://www.audiosignal.co.uk/Resources/Surround_sound_psychoacoustics_A4.pdf>.
- [14] GERZON, Michael A. What's wrong with quadrasonics. *Studio sound*, May 1974. Dostupný z WWW:
<http://www.audiosignal.co.uk/Resources/What_is_wrong_with_quadrasonics_A4.pdf>.
- [15] HUMPHREY, ROBERT A. . *Automatic Loudspeaker Location Detection for use in Ambisonic Systems*. [s.l.], 2006. 200 s. Royal Institute of Technology School of Computer Science and Communication. Dizertační práce. Dostupný z WWW:
<http://www.nada.kth.se/utbildning/grukth/exjobb/rapportlister/2006/rapporter06/humphrey_robert_06114.pdf>.
- [16] MALHAM, D.G.. *Spatial hearing mechanisms and sound reproduction*. 1998. Dostupný z WWW:
<<http://music.nuim.ie/musictec/docs/documents/SPATIALHEARINGMECHANISMSandSOUNDREPRODUCTION.pdf>>.
- [17] PENHA, Rui. Distance Encoding in Ambisonics Using Three Angular Coordinates. *Paper presented at the 5th Sound and Music Computing Conference, Berlin, July 2008*. Dostupný z WWW:
<http://smc.afim-asso.org/smc08/images/proceedings/session2_number3_paper32.pdf>.
- [18] SCHINDLER, Allan. *Spatial ambience and localization programs*. Eastman computer music center (ECMC) users' guide. August 2005. Dostupný z WWW:
<<http://ecmc.rochester.edu/onlinedocs/USERSGUIDE/section7.pdf>>.
- [19] SMETANA, C. a kol. *Praktická elektroakustika*. Státní nakladatelství technické literatury, Praha 1981.
- [20] SoundField Ltd. *SoundField Technology and B-Format* [online]. Dokument dostupný na <<http://www.soundfield.com/soundfield/technology.php>>.
- [21] WIGGINS, Bruce. *An investigation into the real-time manipulation and control of three-dimensional sound fields*. [s.l.], 2004. 370 s. UNIVERSITY OF DERBY. Dizertační práce. Dostupný z WWW:
<<http://sparg.derby.ac.uk/SPARG/PDFs/BWPhDThesis.pdf>>.

SEZNAM ZKRATEK A SYMBOLŮ

2D	dvojdímenzionální (např. pouze horizontální rovina)
3D	trojdímenzionální (kompletní prostor)
*.mat	formát souboru pro ukládání dat v binární podobě v prostředí MATLAB
*.wav	formát souboru pro ukládání akustických signálů v Microsoft Windows
a	koeficienty jmenovatele přenosové funkce filtru
A, B, C, D	signály A-formátu snímané mikrofony A, B, C, D koincidenčního pole mikrofonů nebo kapsle soundfield
A-format	soubor signálů nasnímaných pomocí mikrofону soundfield nebo pole koincidenčních mikrofonů
b	koeficienty čitatele přenosové funkce filtru
B'	kanál pro zlepšení prostorových vlastností horizontálního zvukového pole
B-format	soubor signálů, zakódovaných ze složek A-formátu, určených pro přenos a dekodování do koncového reproduktorového pole
BHJ	dvoukanálová verze přenosového formátu UHJ
C-format	soubor signálů, zakódovaných ze složek B-formátu, určených pro přenos nebo distribuci (častěji nazývaný UHJ)
D-format	soubor signálů, dekodovaných ze složek B-formátu, určených pro reprodukci jednotlivými reproduktory koncového reproduktorového pole
ε	elevace zdroje zvuku (úhel ve stupních ve vertikální rovině)
E-format	soubor signálů, dekodovaných ze složek formátu UHJ, určených pro zpracování shelving filtry
eps	proměnná aplikace B-format Dekodér reprezentující vertikální úhel mezi reproduktory v reproduktorovém poli
φ	azimut zdroje zvuku (úhel ve stupních v horizontální rovině)
f_c	mezní (dělicí) kmitočet filtru
f_{vz}	vzorkovací kmitočet
f_i	proměnná aplikace B-format Dekodér reprezentující horizontální úhel mezi reproduktory v reproduktorovém poli
F-format	soubor signálů získaných ze složek E-formátu průchodem přes shelving filtry

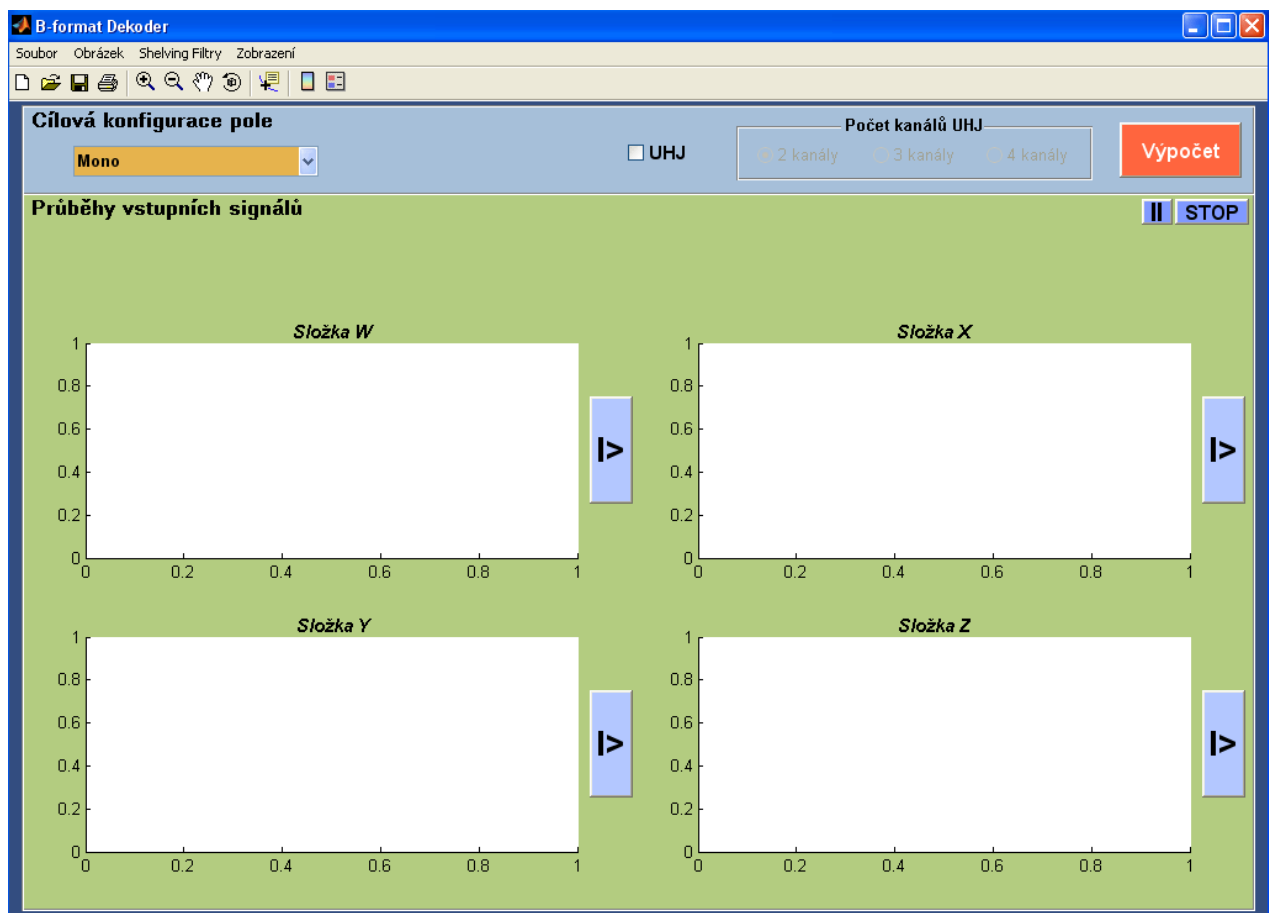
G	zisk (zesílení/zeslabení) filtru
G-formát	soubor signálů, dekodovaných ze složek B-formátu, určených pro reprodukci jednotlivými reproduktory pole 5.1 Surround Sound
ITU-R BS 775-1	specifikace standardu pro reproduktorové pole 5.1 Surround Sound
j	komplexní jednotka – fázový posuv o 90°
k_1	kmitočtově závislý zisk filtru pro filtraci signálu WE nebo W
k_2	kmitočtově závislý zisk filtru pro filtraci signálů XE, YE, ZE nebo X, Y, Z
k_3	kmitočtově závislý zisk filtru pro filtraci signálu B'
k'	koeficient preference přední strany zvukového pole
m-soubor	soubor obsahující zdrojový kód příslušné funkce spustitelný v prostředí MATLAB
n	proměnná aplikace B-format Dekodér reprezentující počet reproduktorů v reproduktorovém poli
PHJ	čtyřkanálová verze přenosového formátu UHJ
Σ, Δ, T a Q	signály přenosového formátu UHJ, zakódované z B-formátu
s	monofonní signál zdroje zvuku
S	signál D-formátu, dekodovaný z B-formátu, určený pro konkrétní reproduktor v reproduktorovém poli
SHJ	tříkanálová verze přenosového formátu UHJ s omezenou šířkou pásma pro třetí kanál
t	koeficient množství příspěvku kanálu T při dekodování UHJ
THJ	tříkanálová verze přenosového formátu UHJ
UHJ	soubor signálů, zakódovaných ze složek B-formátu, určených pro přenos nebo distribuci (také nazývaný C-format)
W, X, Y, Z	složky B-formátu reprezentující všesměrovou složku a tři směry prostoru
W', X', Y', Z'	složky E-formátu dekodované z UHJ fázově-amplitudovou maticí (také WE, XE, YE, ZE)
WE, XE, YE, ZE	složky E-formátu dekodované z UHJ fázově-amplitudovou maticí (také W', X', Y', Z')
WF, XF, YF, ZF	složky F-formátu získané ze složek W', X', Y', Z' (nebo WE, XE, YE, ZE) průchodem přes shelving filtry

SEZNAM PŘÍLOH

A	OVLÁDÁNÍ GRAFICKÉHO UŽIVATELSKÉHO ROZHRANÍ APLIKACE B-FORMAT DEKODÉR	61
B	OBSAH PŘILOŽENÉHO CD	64

A Ovládání grafického uživatelského rozhraní aplikace B-format Dekodér

Pro spuštění programu musíme všechny soubory nahrát do složky v počítači, kterou používá MATLAB. Aplikaci spustíme pomocí její hlavní funkce v souboru `ambisonics.m` a to například tak, že do příkazového řádku MATLABu napíšeme příkaz `ambisonics` (v případě použití *StandAlone* verze aplikace stačí pouze spustit soubor `BfDekoder.exe`). To vede k otevření grafického rozhraní, které má podobu jako na obrázku A.1.

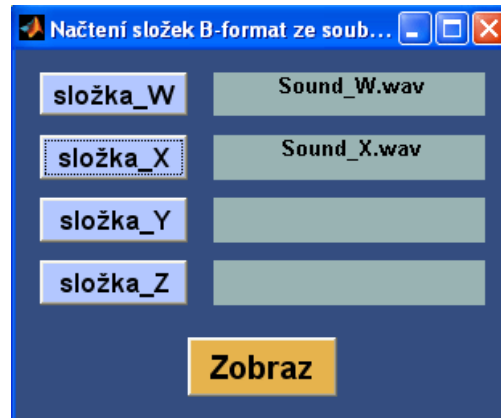


Obr. A.1: Grafické uživatelské rozhraní aplikace B-format Dekodér.

Okno aplikace je rozděleno na lištu s jednotlivými menu, lištu funkcí pro rychlou manipulaci s grafy a dva panely, nazvané „Cílová konfigurace pole“ a „Průběhy vstupních signálů“. Jednotlivé panely a ovládací prvky jsou od sebe z důvodu snadnější orientace barevně odlišeny. Hlavním ovládacím prvkem programu jsou menu v horní liště.

Pomocí prvního menu *Soubor* lze načíst zdrojové soubory (složky *W*, *X*, *Y*, *Z* B-formátu) buď ze souboru typu **.mat* (položka *Načíst zdrojový soubor MAT*), kde jsou jednotlivé složky reprezentovány proměnnými *W*, *X*, *Y*, *Z* spolu se vzorkovacím kmitočtem f_{vz} , nebo z jednotlivých souborů typu **.wav* (položka *Načíst zdrojové soubory WAV*). V druhém případě dojde k otevření okna (viz obr. A.2), které obsahuje dialog pro načtení jednotlivých složek B-formátu s pomocí funkce `wavread` (současně se načte i f_{vz}). Po potvrzení dialogu načtení souboru **.mat* nebo kliknutí na tlačítko *Zobraz* v dialogu pro soubory **.wav* se zobrazí časové průběhy všech složek B-formátu v panelu *Průběhy vstupních signálů*. Další dvě položky menu, což jsou *Uložit cílový*

soubor *MAT* a *Uložit cílové soubory WAV* slouží k uložení jednotlivých výstupních signálů (určených pro jednotlivé reproduktory pole). V prvním případě do jednoho souboru typu **.mat*, kde jsou uloženy ve formě sloupcových vektorů reprezentovaných příslušnými proměnnými (viz zdrojové kódy a obrázky u konkrétních polí, které byly uvedeny v kapitole 10.2). V druhém případě do více souborů typu **.wav*, pojmenovaných stejně jako vektory v předchozím případě. Jejich počet tedy odpovídá počtu reproduktorů pole. Poslední položkou je položka *Konec*, s jejíž pomocí se program ukončí a dojde k uzavření všech jeho otevřených oken.



Obr. A.2: Dialog pro načtení složek B-formátu ze souborů **.wav*.

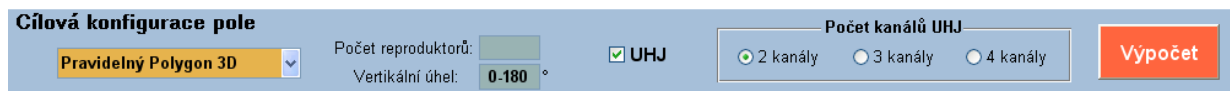
Následující menu *Obrázek* umožňuje zobrazit jednotlivá reproduktorová pole se zkratkovitým popisem. Zobrazení konkrétního obrázku se přitom vyvolá i po ukončení výpočtu výstupních signálů tohoto pole.

V menu *Shelving Filtry* lze nastavit mezní kmitočet shelving filtrů, použitých při výpočtech. K tomuto slouží okno, k jehož zobrazení dojde po kliknutí na položku *Nastavení parametrů*. Stačí pouze zapsat do kolonky nový mezní kmitočet a potvrdit ho kliknutím na tlačítko *OK*. Jako defaultní nastavení má mezní kmitočet (f_c) hodnotu 500 Hz.

Poslední menu *Zobrazení* umožňuje přepínat mezi zobrazením vstupních a vypočtených výstupních signálů kliknutím na položku *Přepnout vstupní/výstupní signály*. Tím dojde k přepnutí panelu *Průběhy vstupních signálů* na panel *Průběhy výstupních signálů* (liší se i barvou), kde jsou po výpočtu zobrazeny průběhy signálů pro reproduktory zvolené soustavy. K tomuto přepnutí dochází také automaticky po výpočtu. Následující položka *Zobrazit v novém okně* je určena k zobrazení aktuálních grafů závislostí do nového okna, které může sloužit zejména pro tisk. Ostatní položky slouží k přepínání mezi náhledy různých typů závislostí odpovídajících názvu jednotlivých položek. Přitom pro zobrazení vstupních složek B-formátu jsou určeny pouze první dva typy analýzy, tedy časová závislost intenzity a spektrogram. Pro výstupní signály je možno zobrazit všechny typy.

Důležitým ovládacím prvkem je také panel *Cílová konfigurace pole*, který realizuje samotný výpočet výsledných signálů podle předem nastavených parametrů a má podobu například jako na obrázku A.3. Před zahájením výpočtu je samozřejmě nejprve nutné načíst zdrojové složky B-formátu pomocí příslušných položek v menu *Soubor*. Parametry výpočtu jsou v panelu umístěny podle důležitosti zleva doprava. První a nejdůležitější parametr je typ (konfigurace) cílového pole, který uživatel vybírá z rozbalovacího menu. Počet položek menu je závislý na volbě konverze do UHJ. Při volbě této konverze lze vybírat jen z pěti typů

reproduktorových polí, které v rámci této aplikace konverzi do UHJ podporují. V závislosti na zvoleném typu soustavy se může jako další parametr vpravo od rozbalovacího menu zobrazit položka pro zadávání horizontálního úhlu (ϕ_i), vertikálního úhlu (ϵ_s) nebo počtu reproduktorů pole (n). Přesný popis těchto položek (proměnných) byl uveden v kapitole 10.1.1 u příslušných polí. Význam parametru počet reproduktorů je zřejmý z jeho názvu. Parametr horizontální úhel je horizontální úhel ve stupních mezi dvěma předními reproduktory vybraného pole vzhledem k pozici posluchače (může nabývat hodnoty od 0° do 180°). Vertikální úhel je totéž, ale pro reproduktory ve vertikálním směru (má tedy smysl pouze v perifonických soustavách). Dalším parametrem v řadě je zaškrťovací políčko s názvem *UHJ*. Slouží k zavedení konverze zdrojových složek B-formátu do formátu UHJ (tedy jakýsi mezistupeň mezi zdrojovými a cílovými signály). Pokud je políčko zaškrtnuté, aktivuje se ještě panel pro výběr počtu kanálů UHJ, který byl dosud nedostupný. V něm si zvolíme počet kanálů UHJ, do kterých se mají zdrojové signály zakódovat – tedy typ BHJ, THJ nebo PHJ (viz kapitola 4 a 10.1). Poslední ovládací prvek v panelu *Cílová konfigurace pole* je tlačítko *Výpočet*. Jeho stiskem dojde k výpočtu výstupních signálů pro reproduktory v poli se zvolenými vlastnostmi (parametry), přepnutí na panel *Průběhy výstupních signálů*, kde se zobrazí jednotlivé průběhy (typ závislostí závisí na zaškrtnuté volbě v menu *Zobrazení*) a k načtení obrázku popisujícího toto pole.



Obr. A.3: Panel pro výpočet výstupních signálů v konkrétním reproduktorovém poli.

Hlavní funkcí panelů *Průběhy vstupních signálů* a *Průběhy výstupních signálů* je zobrazovat vstupní a výstupní signály a jejich vlastnosti v podobě zobrazení různých závislostí jejich parametrů. Panely dále obsahují kromě grafů příslušných typů signálů pouze ovládací prvky určené k jejich reprodukci. Těmito prvky jsou tlačítko *play* s popisem „I“, které se nachází vedle každého grafu časového průběhu, případně spektrogramu, tlačítko *STOP* a tlačítko *pause* s popisem „II“. Tlačítka mají standardní funkci, která je zřejmá z jejich názvů – slouží k přehrávání příslušných akustických signálů (vstupních i výstupních) a k ovládní jeho průběhu. Jako indikátor stavu slouží nápis „Přehrávání...“, který se objeví vlevo od tlačítka *pause* při aktivování reprodukce tlačítkem *play*, nebo nápis „Pauza“ při pozastavení tlačítkem *pause*.

V případě aktivace některé funkce programu dojde k zobrazení nového okna indikátoru aktivity s nápisem „Probíhá výpočet“, které opět zmizí po dokončení výpočtu či funkce. Některé výpočty mohou být z důvodu značné velikosti vzorků zdrojových signálů poměrně zdlouhavé. Při nesprávném zadání parametrů výpočtu nebo aktivaci některé funkce programu, která vyžaduje určité předchozí akce, dojde k vyvolání chybového okna s textem, který chybu zdůvodňuje, případně navrhuje její řešení.

B Obsah přiloženého CD

Adresáře:

- B-formatDekoder – zdrojové soubory aplikace B-format Dekodér spustitelné v prostředí MATLAB (lze je zároveň prohlížet v jakémkoli textovém editoru)
- B-formatDekoder_StandAlone – zdrojové soubory aplikace B-format Dekodér spustitelné samostatně (vytvořené v prostředí MATLAB)
- Zdrojové soubory B-format – zdrojové testovací soubory složek B-formátu (W, X, Y, Z) ve formátu *.mat nebo *.wav získané z internetu [20]

Soubory:

- BakalarskaPrace.pdf* – vypracovaná bakalářská práce
- B-formatDekoder_Manual.pdf* – manuál k programu B-format Dekodér