



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

**Fakulta Elektrotechnická
K13137 Katedra Radioelektroniky**

Metody konverze 2D videa na stereoskopické 3D

Methods of 2D-to-3D video conversion

Diplomová práce

Studijní program: **Komunikace, Multimédia a Elektronika**

Studijní obor: **Multimediální technika**

Vedoucí práce: **Ing. Karel Fliegel, Ph.D.**

Bc. David Röhrich

Praha 2014

Abstrakt

Cílem této práce je seznámit čtenáře s problematikou konverze videa z 2D do 3D stereoskopického. Popsána jsou teoretická monokulární obrazová vodítka, která mohou pomoci při konverzi pro stanovení hloubkové mapy. Dále jsou zmíněny některé konkrétní metody, jež využívají rozdílné přístupy a také hardware a software, které umožňují konvertovat video do 3D. Programy vhodné ke konverzi v domácích podmínkách jsem společně se svou navrženou metodou porovnal v subjektivních testech, které jsem následně doplnil i objektivním měřením.

Klíčová slova: 3D video, Stereoskopie, Hloubková mapa, Obrazová vodítka, Subjektivní hodnocení.

Abstract

In this work I am going to present some facts about conversion of the video stream to stereoscopic 3D. I have described a few theoretical monoscopic cues which are usually used when we want to estimate a depth map. The next section is about a few methods to convert video sequences including a description of commercial software and hardware designed for a same purpose. Two commercial softwares were tested in subjective quality assessment including my proposed method. The subjective quality assessment was followed by an objective one too.

Index terms: 3D video, Stereoscropy, Depth map, Pictorial cues, Subjective quality assessment.

Poděkování

Rád bych poděkoval vedoucímu práce Ing. Karlu Fliegelovi, Ph.D za jeho čas a trpělivost, kterou mi věnoval při řešení zadané problematiky. Dále děkuji kolegovi Bc. Dominiku Vozákovi za cenné připomínky, Mgr. Filipu Duškovi za pomoc při korektuře textu a v neposlední řadě všem respondentům, kteří se zúčastnili subjektivního testu.

Prohlášení

„Prohlašuji, že jsem předloženou práci vypracoval samostatně a že jsem uvedl veškeré použité informační zdroje v souladu s Metodickým pokynem o dodržování etických principů při přípravě vysokoškolských závěrečných prací.“

V Praze dne

David Röhrich

.....

.....

Zadání

Obsah

PODĚKOVÁNÍ	3
PROHLÁŠENÍ	4
ZADÁNÍ	5
OBSAH	6
SEZNAM ZKRATEK	8
SEZNAM OBRÁZKŮ	9
ÚVOD	10
1 STEREOSKOPIE	11
1.1 TEORETICKÝ PRINCIP	11
1.2 BINOKULÁRNÍ PARALAXA	13
2 KONVERZE 2D DO 3D	15
2.1 ROZDĚLENÍ METOD	15
2.2 PRINCIPY PRO URČENÍ HLOUBKY SCÉNY	17
2.2.1 Hloubka z ostrosti objektů	18
2.2.2 Hloubka z geometrie snímku	19
2.2.3 Hloubka z barev a jejich intenzity	22
2.2.4 Hloubka z pohybu ve scéně	23
2.2.5 Generování stereoskopického obrazu	25
2.2.6 Přehled vodítek a jejich zdrojových publikací	28
3 POPIS KONKRÉTNÍCH IMPLEMENTACÍ	29
3.1 TEORETICKÉ METODY	29
3.1.1 Prostorově - časová konverze	29
3.1.2 Metoda pro konverzi v reálném čase	31
3.1.3 Neparametrické vzorkování	32
3.1.3.1 Podrobnější teoretický popis	33
3.1.3.2 Testování v praxi	35
3.1.3.3 Příklad I	36
3.1.3.4 Příklad II	36
3.1.3.5 Příklad III	36
3.1.3.6 Příklad IV	37
3.1.3.7 Subjektivní pozorování výsledků	38
3.1.3.8 Zhodnocení metody	39

3.2 NAVRŽENÁ METODA.....	40
3.3 KOMERČNÍ ŘEŠENÍ	42
3.3.1 <i>Profesionální 3D procesory</i>	42
3.3.1.1 Blackmagic Teranex 3D processor.....	42
3.3.1.2 JVC IF-2D3D1 Image Processor	43
3.3.1.3 SONY MPE200 Multi Image Processor	43
3.3.2 <i>Software - automatické metody</i>	45
3.3.2.1 TriDef 3D	45
3.3.2.2 PowerDVD 13.....	46
3.3.3 <i>Software - poloautomatické metody</i>	47
3.3.3.1 DVDFab 3D Converter	47
3.3.3.2 Další komerční software	48
4 PRAKTICKÉ MĚŘENÍ	49
4.1 SUBJEKTIVNÍ HODNOCENÍ	49
4.1.1 <i>Metodika subjektivních testů</i>	49
4.1.2 <i>Postup měření</i>	51
4.1.2.1 Sada testovacích sekvencí.....	51
4.1.2.2 Průběh testu	52
4.1.3 <i>Výsledky</i>	54
4.1.4 <i>Statistické vyhodnocení</i>	57
4.2 OBJEKTIVNÍ MĚŘENÍ	59
4.2.1 <i>VQM metrika</i>	59
4.2.2 <i>Postup měření</i>	61
4.2.3. <i>Výsledky</i>	61
5 ZÁVĚR.....	63
6 SEZNAM ZDROJŮ	65
7 PŘÍLOHY	69
7.1 DOTAZNÍK PRO SUBJEKTIVNÍ HODNOCENÍ.....	69
7.2 OBSAH DVD.....	70
7.3 UKÁZKY TESTOVACÍCH SEKVENCÍ	71
7.3.1 <i>Automobil</i>	71
7.3.2 <i>Basketbal</i>	71
7.3.3 <i>Dvůr</i>	71
7.3.4 <i>Pracovna</i>	72
7.3.5 <i>Tělocvična</i>	72
7.3.6 <i>Deštník</i>	72

Seznam zkratek

ANOVA	Analysis of Variance
ANSI	American National Standards Institute
BIQI	Blind Image Quality index
DIBR	Depth Image Based Rendering
ITU-R	International Telecommunication Union - Radiocommunication
MSE	Mean Square Error
NTIA	National Telecommunications and Information Administration
RGBD	Red Green Blue Depth
SBS	Side-by-Side
SSD	Solid-State Drive
TB	Top-Bottom
VQM	Video Quality Metric

Seznam obrázků

Obr. 1 Vznik trojrozměrného vjemu.	12
Obr. 2 Schéma binokulární paralax.	13
Obr. 3 Diagram monokulárních vodítek.	17
Obr. 4 Příklad scény se zaostřenými objekty v popředí.	18
Obr. 5 Ostření na blízké/vzdálené objekty.	19
Obr. 6 Příklad využití lineární perspektivy ve scéně.	20
Obr. 7 Scéna obsahující vodítka výška v obraze, interposition a relativní velikost.	20
Obr. 8 Ukázka falešných textur na omítce.	21
Obr. 9 Příklad atmosférického rozptylu.	22
Obr. 10 Rozdělení scény na do tří bloků se stejnou hloubkou.	23
Obr. 11 Příklad snímku a jeho hloubková mapa.	25
Obr. 12 Schéma pro výpočet horizontálního posunu pixelů.	26
Obr. 13 Blokové schéma principu DIBR - nakresleno podle.	27
Obr. 14 Princip posunutí obrazu s chybějící informací.	29
Obr. 15 Princip prostorového posunutí.	30
Obr. 16 Blokové schéma metody Neparametrického vzorkování.	33
Obr. 17 Výsledek příkladu I - 1. frame.	36
Obr. 18 Výsledek příkladu II - 1. a 7. frame.	36
Obr. 19 Výsledek příkladu III - 1, 15. a 30. frame.	37
Obr. 20 Výsledek příkladu IV - 1, 15. a 30. frame.	37
Obr. 21 Princip získání snímku A a L.	41
Obr. 22 Ilustrace změn regionů.	41
Obr. 23 Blackmagic Teranex 3D.	42
Obr. 24 JVC IF-2D3D1 Image Processor.	43
Obr. 25 SONY MPE200 Multi Image Processor.	43
Obr. 26 TriDef 3D nastavení výstupního zobrazení.	45
Obr. 27 Schéma uspořádání pracoviště pro subjektivní testy.	53
Obr. 28 Graf výsledků pro Hloubku scény.	54
Obr. 29 Graf výsledků pro Artefakty.	55
Obr. 30 Graf výsledků pro Celkový dojem.	55
Obr. 31 Graf celkových výsledků.	57
Obr. 32 Graf metriky VQM pro jednotlivé metody a sekvence.	62
Obr. 30 Sekvence 1 - Automobil.	71
Obr. 31 Sekvence 2 - Basketbal.	71
Obr. 32 Sekvence 3 - Dvůr.	71
Obr. 33 Sekvence 4 - Pracovna.	72
Obr. 34 Sekvence 5 – Tělocvična.	72
Obr. 35 Sekvence 6 – Deštník.	72

Úvod

V současné době se na trhu objevuje stále větší množství 3D televizorů nebo podobných zařízení, lákající zákazníky ke koupi. Zároveň však není dostatek materiálu, se kterým by se daly využít všechny vlastnosti těchto 3D přístrojů. Zdaleka ne všechny filmy jsou pořizovány ve 3D a televizní vysílání v tomto formátu v České republice zatím není dostupné pro širokou veřejnost. Existuje ovšem možnost konverze z 2D do 3D, čemuž bych se chtěl v této práci věnovat.

Protože převod starých filmových nebo jiných materiálů je pro studia velmi finančně a časově náročný proces, vznikají různé další nástroje, které jsou schopné konvertovat obsah do 3D levněji, ovšem s výrazně nižší kvalitou. Rád bych se proto podíval na možnosti, které má běžný uživatel, v případě, že chce sledovat materiál, který ve 3D pořízen nebyl.

Cíle pro tuto práci jsou rozděleny do těchto dílčích úkolů: **popis a rozdělení metod, kterými je možné konvertovat video do 3D, dále vysvětlení principů, na jejichž základě je možné rozpoznat prostorové uspořádání ve sledované scéně a nakonec samotné otestování konkrétních algoritmů.**

Hodnocení algoritmů **bude prováděno pomocí subjektivních testů i objektivní metriky**, takže budu moci vyhodnotit kvalitu jednotlivých metod.

Věřím, že tato práce pomůže čtenářům získat představu o komplexnosti této problematiky a lepší orientaci v aktuálních možnostech konverze v domácích podmínkách pro běžné uživatele.

1 Stereoskopie

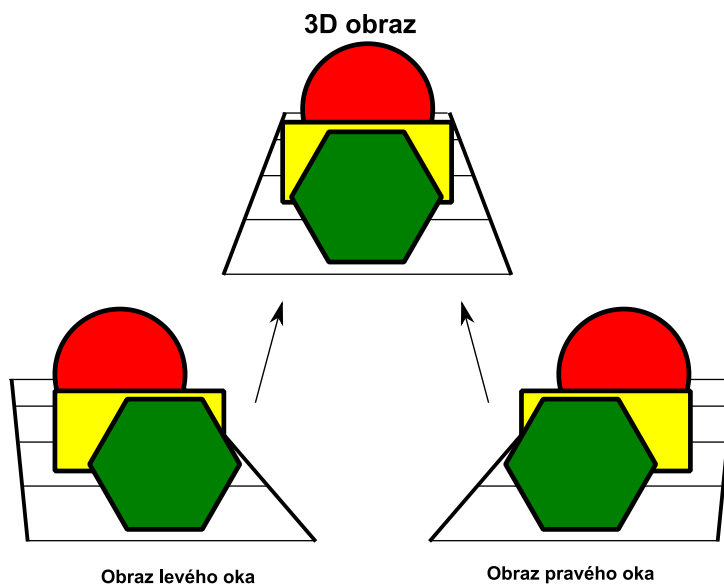
V první kapitole teoretické části se nejprve stručně podíváme na to, jak funguje stereoskopické zobrazení, kterému se dnes říká 3D. Tato část vychází z publikací [1], [2] a také článku [3]. Rovněž posloužila i moje bakalářská práce [4].

Díky technologickému pokroku jsme se dostali do doby, kdy nás začíná zajímat nejen kvalita filmu a dalších vizuálních produktů, ale také jejich forma. Trh se nasytil množstvím 3D televizorů, ale nyní je třeba zajistit dostatečné množství obsahu, který je schopen využít potenciál těchto zařízení.

Samotné pořízení tohoto trojrozměrného obsahu není levné, a proto by bylo ideální využít již stávající materiály, abychom je mohli také sledovat ve 3D a vychutnali si je tak ještě více. Existuje celá řada principů, díky kterým můžeme obraz konvertovat z 2D do 3D, a jejich problematiku bych v této práci rád popsal.

1.1 Teoretický princip

V první řadě je nutné si ujasnit, jak funguje samotný 3D vjem. To, co můžeme sledovat v TV či v kině pomocí speciálních brýlí, není trojrozměrná realita v pravém slova smyslu, ale pouze vhodně zobrazené stereoskopické snímky (dále v textu ovšem budeme tento typ zobrazování označovat jako 3D). Díky páru očí vidíme svět ze dvou odlišných úhlů, tedy některé objekty jsou v pozorované scéně pro každé oko lehce posunuté a lidský mozek z těchto rozdílů, kterým se říká *disparita*, vygeneruje trojrozměrný vjem scény. Na *obr. 1* je schéma, které ilustruje vznik prostorového vjemu. Proto se 3D materiál natáčí pomocí dvou kamer, aby byly získány dva navzájem posunuté obrazy.



Obr. 1 Vznik trojrozměrného vjemu - nakresleno podle [4].

Abychom následně mohli získaný obsah sledovat, potřebujeme obraz z levé kamery dostat do levého oka a obraz z pravé kamery do pravého oka. K tomu nejčastěji využijeme systémy s brýlemi, nicméně existují i TV, kde brýle nejsou třeba¹. Nejčastější typy projekce pro 3DTV, které využívají brýle jsou: **aktivní technologie** a **pasivní technologie**. Pasivní technologie zobrazuje levý i pravý kanál na obrazovce současně. Pasivní brýle obsahují dva polarizační filtry, kde obraz jednoho oka je promítán s levotočivou polarizací světelných vln a obraz druhého oka s pravotočivou polarizací. Naproti tomu aktivní technologie střídavě zatemňuje na brýlích levé nebo pravé oko a na obrazovce se velmi rychle střídá levý kanál s pravým. Brýle jsou synchronizovány s 3DTV, a proto je vidět každým okem příslušný obraz ve správnou chvíli. Dále existují i systémy, které pro svou funkci brýle nepotřebují, ty se nazývají **autostereoskopické**. Využívají přesně definovaných zón, kde sedí divák, a pomocí speciální masky, která je umístěna před obrazovkou, lámou světelné paprsky současně zobrazených obrazových kanálů tak, že pozorovatel vnímá výsledek trojrozměrně. Jako poslední zmíníme **anaglyf**, který funguje podobně jako pasivní technologie, jen místo polarizačních filtrů a rozdílné polarizace vln využívá filtry barevné a oba kanály jsou barevně tónované. Důvod, proč je stále populární i anaglyf je ten, že si díky němu můžeme prohlížet 3D obsah i na papíře.

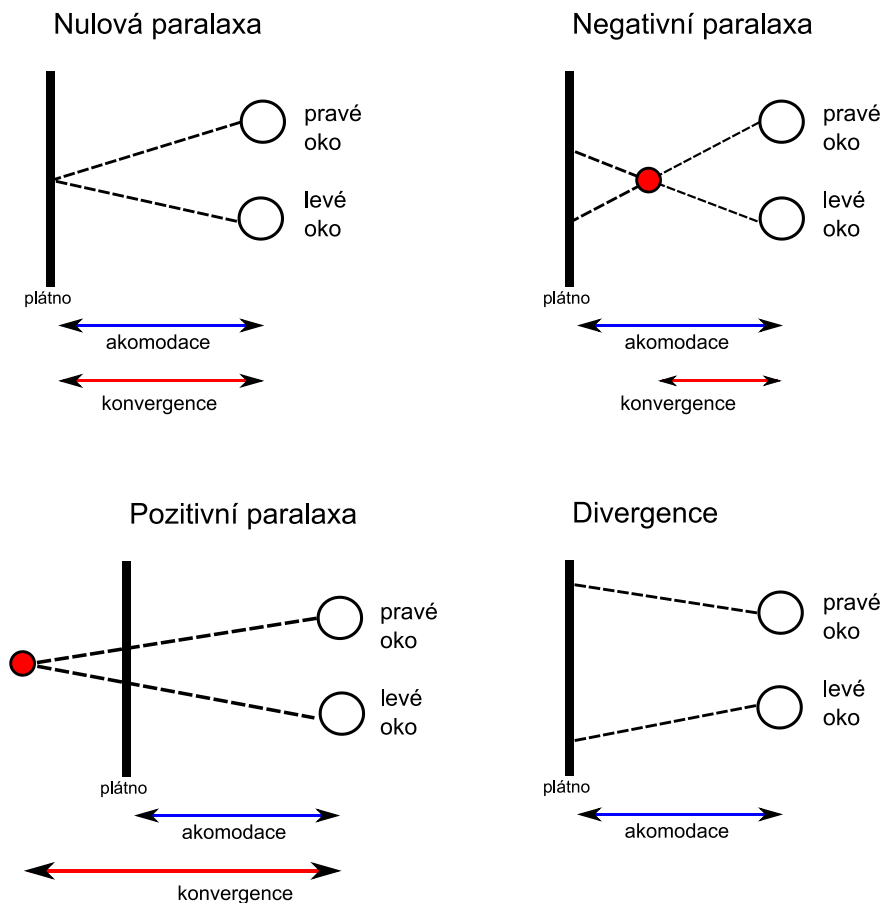
Jak již bylo zmíněno výše, protože jsou nyní 3DTV dostupné široké veřejnosti a rychlost s jakou je vyráběn 3D obsah zdaleka nesplňuje poptávku, je nutné převádět staré filmy či živé

¹ Podrobnější informace o této problematice jsou v mé bakalářské práci [4], kde jsem srovnával a testoval různé metody zobrazování, které používají 3DTV

vysílání z 2D do 3D. To znamená, že ke standardním pořízeným snímkům musíme vytvořit druhý obrazový kanál, abychom vytvořili disparitu, prostřednictvím které se nám dostane požadovaného 3D vjemu.

1.2 Binokulární paralaxa

Vnímání věcí před plátnem nebo za plátnem ovlivňuje binokulární paralaxa, což je úhel, který mezi sebou svírají oční osy, při sledování objektů. S tím je úzce spjat pojem konvergence očí neboli sbíhavost očních os. *Obr. 2* schematicky naznačuje čtyři typy paralaxy - **negativní**, **nulová**, **pozitivní** a čtvrtý případ, kdy dochází k **divergenci očních os**.



Obr. 2 Schéma binokulární paralax - překresleno z [5].

Z obrázku na první pohled vidíme, že akomodace čočky, tj. zaostření na objekt, je stále stejné, protože ostříme na plátno a mění se pouze konvergence očních os. V případě, že pozorujeme obraz před plátnem, jedná se o negativní paralaxu, a naopak pokud pozorujeme obraz za plátnem, nazýváme to pozitivní paralaxou. Nulová paralaxa zobrazuje obraz přímo na plátně, kdy si můžeme povšimnout, že akomodace a konvergence jsou ekvivalentní. Většina nežádoucích jevů spojených se sledováním stereoskopického videa, jako je například nevolnost, bolest hlavy, únava očí apod., plynou z toho, že akomodace a konvergence nejsou v rovnováze [7]. V reálném světě se vždy při ostření zraku na různé objekty mění i sbíhavost očních os, což ve stereoskopii není možné, protože oči jsou zaostřeny stále na plátno nebo TV. Poslední případ, který označujeme jako divergence, je situace, kdy se oční osy rozbíhají. Ten nastává většinou jen ve chvíli, kdy se snažíme zaostřit na velmi blízký předmět, případně je scéna špatně nasnímána. Následkem této situace může být opět nevolnost apod.

2 Konverze 2D do 3D

Kapitola podává přehled základního třídění metod konverze 2D materiálu do 3D stereoskopického. Vysvětluje teoretické principy, pomocí kterých můžeme získat informaci o uspořádání scény díky obrazovým vodičkům. Dále popisuje pojem hloubková mapa a jak ji využít pro převod do 3D. Tato kapitola převážně čerpá z článků [3] a [6]. Tuto část jsem zpracovával již ve svém semestrálním projektu [8] a zde jsem ji rozšířil.

2.1 Rozdělení metod

Metody pro konverzi obsahu z 2D na 3D můžeme rozdělit do tří základních skupin podle toho, zda je nutné, aby uživatel zasahoval do samotné konverze (automatický převod, poloautomatický převod a kompletně manuální).

Automatický převod je ideální v případě, že chceme konvertovat větší množství snímků, a zároveň nevyžaduje, aby uživatel znal přesné informace o hloubce scény, jak je tomu u manuálních technik. Zde vše ponecháme na použitém algoritmu a záleží pouze na principu, kterým může být konverze řešena (více v kapitole 2.2.5). Většinou je automatický převod realizován v reálném čase. Tento typ metod je použit v 3DTV nebo Blu-Ray přehrávačích či externích boxech, které mají možnost automatické konverze obsahu do 3D. Kvalita tohoto převodu nicméně zatím není příliš vysoká. Pro tento účel můžeme rovněž využít některé dostupné komerční programy, jako je Power DVD² nebo TriDef³, které obsahují přehrávač, jenž umí přehrávané video rovnou zobrazovat jako stereoskopické.

Poloautomatické metody se od předešlých příliš neliší. Většinou umožňují podrobnější definování vstupních podmínek a jsou tedy složitější. Opět zde existují komerční programy, mezi které patří například DVDFab Converter⁴, MakeMe3D⁵, 3D Video converter⁶ a 4VideoSoft⁷. Zde stačí vybrat video, zvolit jeho výstupní video kodek a případně nastavit posuvníkem velikost hloubky scény. Měnit se dá výstupní rozlišení, vložení titulků nebo zda bude video v SBS či TB formátu.

² http://www.cyberlink.com/downloads/trials/index_en_EU.html

³ <http://www.tridef.com/download/TriDef-3D3latest.html>

⁴ <http://www.dvdfab.com/2d-to-3d-converter.htm>

⁵ <http://www.engelmann.com/en-us/makeme3d.php?affChecked=1#download>

⁶ <http://convert-to-3d.com>

⁷ <http://www.4videosoft.com/3d-converter>

Poslední skupinou jsou **manuální techniky** převodu a ty vyžadují kompletní řízení uživatelem. Sem můžeme zařadit například program SilhouetteFX⁸. Tento nástroj umožňuje převod video sekvence, kde v každém framu nejprve určíme globální podobu hloubkové mapy (více o hloubkové mapě je v kapitole 2.2.5), podle několika předem nastavených šablon. Poté můžeme postupně vybírat libovolné objekty scény a zaznamenávat je do hloubkové mapy. Zásadní nevýhodou je vysoká cena těchto nástrojů a tento přístup konverze má navíc další nevýhodu v podobě časově náročného procesu, protože celou sekvenci musíme konvertovat snímek po snímku. Navíc záleží pouze na uživateli, jak celou scénu uspořádá, a z toho vyplývá, že některé vzdálenosti mezi objekty mohou být značně zkreslené. Pokud se ovšem podaří vytvořit kvalitní hloubkovou mapu, tak jsou výsledky výrazně lepší, než při použití automatického algoritmu. Proto filmová studia využívají manuální techniky, a celý proces převodu do 3D některých starších filmových děl je tak nákladný⁹.

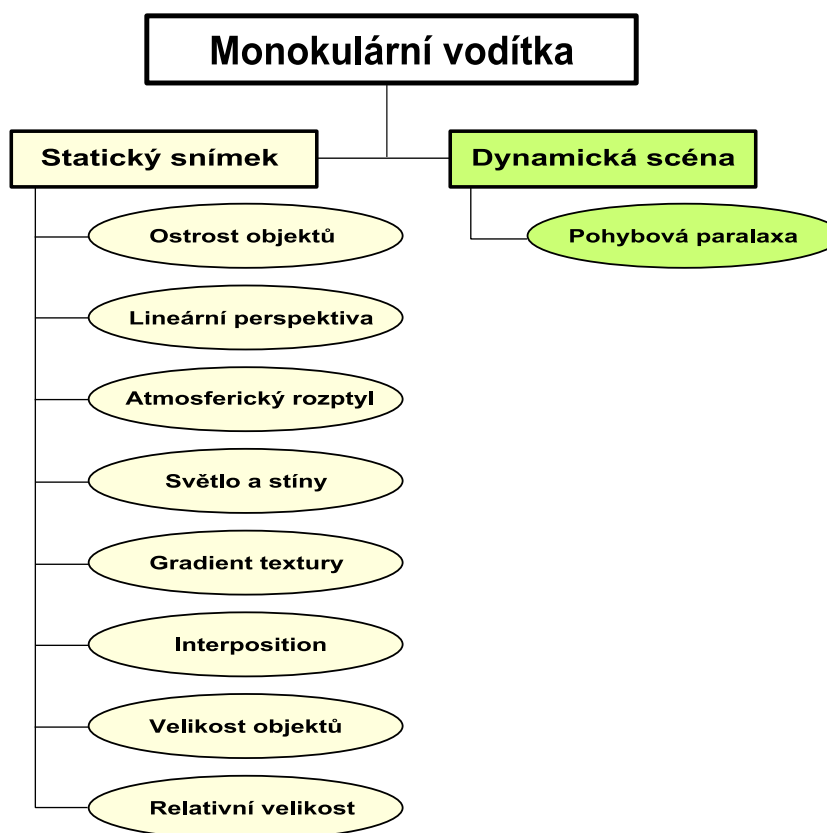
⁸ Trialová verze je k dispozici na stránkách <http://www.silhouettefx.com/>. Pomocí nástroje jsem zkoušel konvertovat některé sekvence, ale jeho použití je složité a časově náročné.

⁹ Článek zmiňující nákladnost konverze starších filmů <http://www.highdefdigest.com/blog/convertng-movies-to-3d/>

2.2 Principy pro určení hloubky scény

Tato podkapitola podává seznam a popis principů, které se nejčastěji užívají k vytvoření stereoskopického snímku. Vycházíme z vlastností lidského zraku, který je schopný získat představu o hloubce scény i monokulárně (například ve chvíli, kdy vidíme fotografii, dokážeme určit, co je v pozadí a co popředí, přestože scéna byla zaznamenána jen jedním objektivem). Anglická terminologie využívá výrazu *monoscopic cue* nebo také *pictorial cue*, které budeme dále nazývat jako *monokulární vodítka* nebo *obrazová vodítka*. Kapitola vychází z textu [3] a [6]. Na konci této části se podíváme na vytvoření stereoskopického snímku pomocí *DIBR*.

Stanovení hloubkové mapy je závislé na výchozích podmínkách. Buď máme k dispozici **statický snímek**, a tedy používáme statická vodítka, a nebo se jedná o delší sekvenci, kde detekujeme **pohyb** a využíváme pohybových vodítek. Diagram na *obr. 3* podává přehled o nejčastěji používaných monokulárních vodítkách. Metody, které se v praxi používají, jsou obvykle kombinací několika níže uvedených vodítek.



Obr. 3 Diagram monokulárních vodítek - nakresleno podle[3].

2.2.1 Hloubka z ostrosti objektů

Lidské oko jako čočka je schopné zaostřit na určité objekty, ať už blízké nebo vzdálené. Stejně tak kamera při snímání scény určité místo zaostří a zbytek zůstává rozmazaný. Na *obr. 4* uvádím příklad scény, kde můžeme vidět, že část scény je ostrá (tyto předměty leží v popředí) a zbytek je rozmazaný. Existuje zde několik praktických možností využití. Princip se v literatuře nazývá *Depth from Blur* [6] nebo také *Depth from Focus/Defocus* [3].



Obr. 4 Příklad scény se zaostřenými objekty v popředí.

K prvnímu přístupu zmíněnému v [3] potřebujeme scénu nasnímat s různým optickým nastavením, abychom mohli získat rozdíly v rozmazání scény. Tedy při zaostření na blízko budou ostré jen ty části scény, které jsou skutečně blízko, zatímco vzdálené budou rozmazané. Při ostření na dálku zůstanou rozmazané bližší části scény a vzdálené budou ostré. Z těchto rozdílů následně získáme informaci o hloubce scén, protože ostré části obrazu z blízkého zaostření prohlásíme za bližší a ostré obrazy ze vzdáleného zaostření za vzdálené. *Obr. 5* ukazuje příklad využití rozdílného zaostření. První fotka je ostřena na blízko, a proto je socha ostrá s rozmazaným okolím a naopak druhá fotka je ostřena na vzdálenější objekty, čímž se rozmazala socha. Podrobnější informace o tomto přístupu jsou v publikaci [9]. Z praktického hlediska tato metoda není příliš použitelná, protože nejsme většinou schopni snímat scénu více optickými zařízeními současně.



Obr. 5 Ostření na blízké/vzdálené objekty.

Pokud tedy máme k dispozici pouze jediný snímek, hloubka obrazu se určuje výpočtem velikosti rozmazání pro každý pixel. Článek [3] uvádí několik možností tohoto výpočtu. Například *dekonvoluce* ve frekvenční oblasti a následné použití *inverzního filtru* [10], [11], dále *detekce hran* na místech s různou úrovní rozmazání či *vlnková analýza* [12], [13] různých oblastí snímku, kde oblasti s vysokým zaostřením získají mnoho nenulových vlnkových koeficientů, kdežto oblasti rozmazané jich budou mít málo.

Musíme si ovšem uvědomit, že rozmazání obrazu může být způsobeno chybou ostření kamery, pohybem a podobně. Tedy ne všechno rozmazání je způsobeno rozdílnou hloubkou.

2.2.2 Hloubka z geometrie snímku

Zde vycházíme z geometrických rozměrů ve scéně. Anglicky se tento princip nazývá *Depth from geometric cues*. Mezi nejznámější vodítka patří **lineární perspektiva**, **výška v obraze**, **gradient textury**, **pozice mezi objekty**, *interposition* a **relativní velikost**. Jak uvádí článek [3], jsou některé přístupy lepší než druhé.

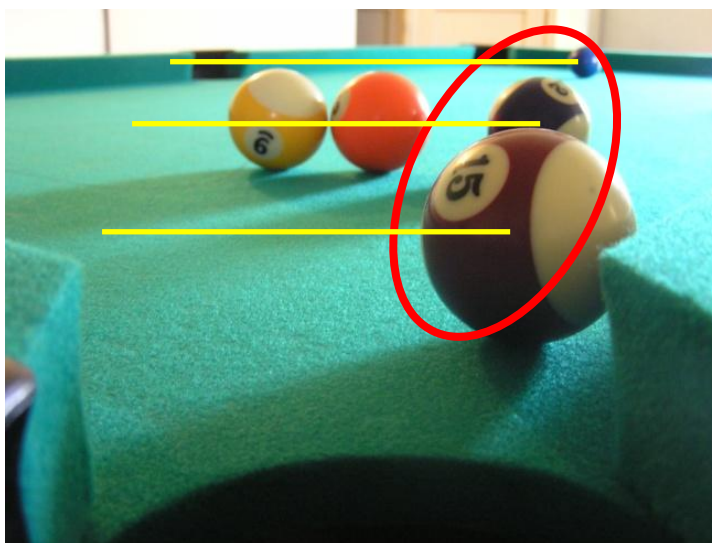
Nejlépe se pracuje s **lineární perspektivou** [14], [15], [16]. V obraze se snažíme najít rovnoběžné linie a k nim jejich společný bod, kde by se linie potkaly. (Velmi častý je tento jev například u vlakových kolejí, zábradlí na mostě nebo okraje cest). Hledaný bod se v literatuře nazývá *vanishing point*, ten je nejvzdálenější a poté pomocí linií k němu vedoucích - *vanishing lines* přiřazujeme hloubkou dalším objektům na stejné úrovni ve vzdálenosti od tohoto bodu.

Obr. 6 ukazuje příklad scény, kde můžou být použity *vanishing lines*. Na scéně je zachycen most a jeho zábradlí sbíhající se v dálce, kde u budovy můžeme určit *vanishing point*.



Obr. 6 Příklad využití lineární perspektivy ve scéně.

Zjišťování hloubky scény se dá určit i pomocí **výšky v obraze** [17]. Předpokládáme, že objekty bližší se na snímku nacházejí dole a objekty vzdálené jsou na snímku výše. To je použitelné hlavně u venkovních scén a krajín. Obraz tedy rozdělíme na pásy, kde jsou objekty se společnou vzdáleností od pozorovatele. Obr. 7 představuje příklad scény, kde jsou nejbližší objekty na snímku dole.



Obr. 7 Scéna obsahující vodítka výška v obraze, interposition a relativní velikost.

Lidské oko sice umí rozpoznat z textury povrchu informaci o objektu a jeho tvarech, ale pro použití ve video konverzi se to příliš nepoužívá [3]. Naopak metoda v článku [18] s texturami do jisté míry pracuje. Na *obr. 8* je zachycena budova, která má místo omítky malbu falešných reliéfů a balkónů. Pozorovatel toto snadno odhalí, ale pro automatické algoritmy to znamenalo chybné výsledky.



Obr. 8 Ukázka falešných textur na omítce.

Další geometrické obrazové vodítko se v literatuře uvádí jako *interposition*. Můžeme ho pozorovat v případě, že vidíme několik objektů za sebou. První v řadě je viditelný celý, zatímco následující je částečně zakrytý prvním. Tím víme, které objekty jsou blíže a které dále, nicméně informaci o vzdálenosti mezi nimi nemáme. *Obr. 7* s kulečnickovými koulemi ilustruje *interposition* v praxi.

Jako poslední zmíníme, že je možné určit hloubku scény pomocí **relativní velikosti objektů** v obraze. Opět se můžeme podívat na *obr. 7*, kde jsou stejné kulečnickové koule různé velké. Toto vodítko nám říká, že velikost objektu se zmenšuje se vzdáleností od pozorovatele. V tomto případě nám tato znalost pomůže určit hloubku scény, ale většinou máme na snímku spoustu nesouvisejících objektů a jejich přesnou reálnou velikost či alespoň poměrnou k ostatním předmětům neznáme.

2.2.3 Hloubka z barev a jejich intenzity

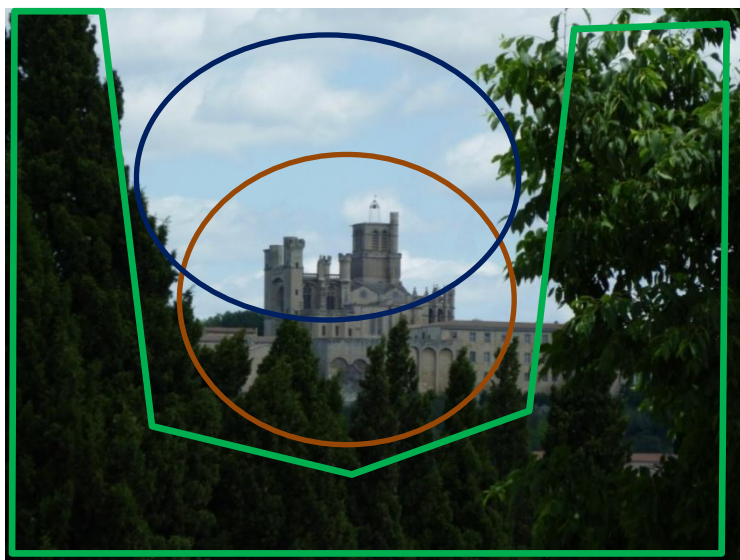
Hloubkovou informaci můžeme získat také ze změny intenzity světla nebo ze změn barev. Principy zmíněné v [3] jsou **atmosférický rozptyl paprsků**, **rozdělení světla a stínu**, vnímání *figure-ground* a **kontrast**.

Atmosférický rozptyl [14], [19] vychází z přírodního chování a rozptylu atmosférických paprsků, kde vzdálenější objekty mají namodralý odstín a horší kontrast, kdežto bližší objekty kontrast lepší. Praktická ukázka tohoto principu je na *obr. 9*. Zde vidíme poloostrov, který je krásně barevný a v modrém oparu v dálce vidíme most i zámek.



Obr. 9 Příklad atmosférického rozptylu.

Článek zmiňuje i další princip, který z těchto postřehů o barvách vychází. Algoritmus se z velkého počtu venkovních snímků naučí určovat šest různých částí krajiny a přiřadit jim hloubku. Pak je například nebe považováno za nejvzdálenější část krajiny. Jelikož se jedná o venkovní snímky nebo záběry krajiny, není nutné přesně a detailně určovat každý kousek scény zvláště, což je typické pro interiérové scény či záběry detailu. Na *obr. 10* je fotka, kterou podle tohoto principu může také rozdělit na určité regiony, jež mají jinou vzdálenost od pozorovatele. Konkrétně zde je les nejbliže, poté následuje hrad (ten je zároveň také v modrém oparu) a nejvzdálenější je nebe.



Obr. 10 Rozdělení scény na do tří bloků se stejnou hloubkou.

Jako další informaci můžeme využít poznatky o **teplotě barev** [20] u venkovních scén, kdy teplejší barvy bereme jako bližší a barvy studené jako vzdálenější. **Ze světla a stínů** [21] dovedeme určit prostorové uspořádání objektů, ale jedná se o poměrně složitý proces, protože nejprve musíme přiřadit konkrétní stíny k objektům, poté najít místo, odkud vychází světelný zdroj, a nakonec z toho vypočítat hloubku scény.

Rovněž můžeme hloubku scény určovat z principu, který se v literatuře nazývá vnímání *figure-ground* [22], [23]. Pokud máme dvě sousedící oblasti v obraze, tak hrany, které je od sebe oddělují, vymezují jeden objekt, kterému se říká *figure* a který se považuje za blízký. Druhé je okolí, které je v pozadí a nazývá se *ground*. Poté pro ohraničené objekty *figures* určujeme umístění ve scéně.

Důležitá zmínka v článku [3] k těmto vodítkům upozorňuje na fakt, že principy a postupy vycházející z barev většinou nedokážou přesně stanovit reálnou hloubku scény, ale přesto jsou použitelné pro konverzi do 3D, protože lidský zrak je dokáže správně vnímat a vyhodnotit.

2.2.4 Hloubka z pohybu ve scéně

Když máme k dispozici sérii snímků či celou video sekvenci, můžeme využít další obrazové vodítko, které nám umožňuje identifikovat hloubku scény z pohybu. Nazýváme ho pohybovou paralaxou (*motion parallax*). To funguje tak, že pokud se pohybuje pozorovatel, tak vidí, že se blízké předměty pohybují rychleji než ty vzdálené, a tím získává informaci o hloubce scény.

Pro zaznamenaný snímek to tedy funguje obdobně, protože pokud se ve scéně pohybuje kamera, tak se bližší objekty pohybují rychleji, než ty vzdálenější. Nejlepších výsledků při zkoumání pohybové paralaxy získáme ve chvíli, kdy se kamera volně pohybuje napříč scénou. Také se může stát, že v obraze je objekt, který se pohybuje nezávisle na kameře. Toho se dá využít k získání informace o jeho prostorovém umístění, ale zde je riziko možných chyb. Je však důležité, aby ve scéně byl přítomný pohyb kamery, jinak pohybovou paralaxu nezískáme. Kamera otáčející se kolem optické osy pohybovou paralaxu nemá [3].

Detekce pohybu se provádí různými způsoby. Můžeme sledovat pohybové vektory napříč celou scénou, pro bloky pixelů, konkrétní objekty a nebo pro každý pixel zvlášť. Stanovení pohybu není snadné, ale v zásadě se dělí na přímé a nepřímé metody, na které dále odkazuje publikace [3].

Při přímých metodách se využívá matematických principů. kdy jsou snímky rozloženy do bloků a následně jsou podrobeny dalším algoritmům (v publikaci jsou zmíněny: *block-matching algorithm*, *pixel recursive algorithm*, *MAP/MRF založený na Bayesově odhadu a odhad optického toku*). U nepřímých metod se sledují určité konkrétní prvky napříč snímky a z nich se odhaduje pohyb.

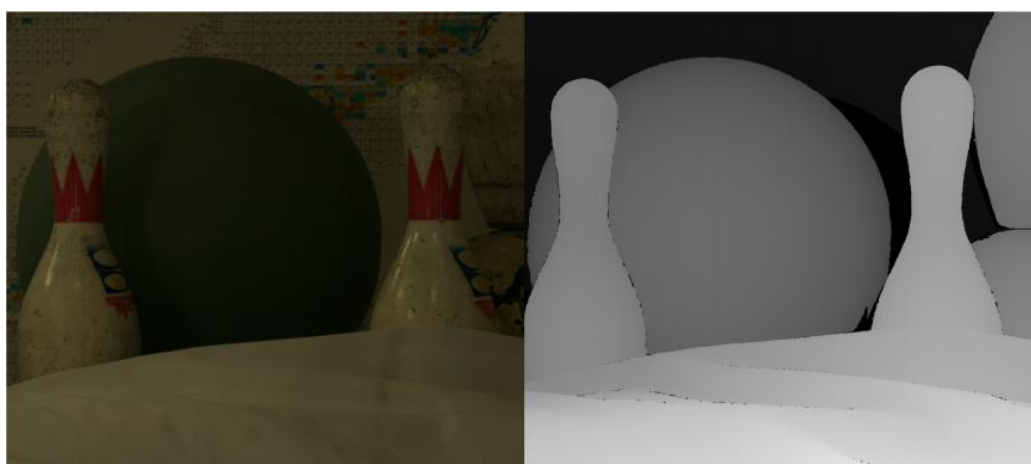
Získání hloubky scény pomocí pohybu je jedním ze základních principů používaných ve 3DTV, kde se většinou doplňuje ještě některým z výše uvedených statických vodítek. 3DTV používá konverzi v reálném čase, a tak je kladen důraz na rychlost přepočítávání, proto nejsou statická vodítka využita všechna. Je však velice složité zjistit konkrétní informace, protože výrobci zobrazovačů tyto detaily volně nezveřejňují.

V článku [18], který se zabývá implementací metody pro konverzi v reálném čase (této metodě se věnuje kapitola 3.1.2), se autoři soustředí hlavně na to, aby pozorovatel vnímal scénu pohodlně a subjektivně příjemně, než aby byla scéna kompletně fyzikálně a geometricky přesně znázorněna.

2.2.5 Generování stereoskopického obrazu

V předchozí části jsme se stručně podívali na principy, kterými můžeme rozhodnout o hloubce předmětů ve scéně, a to následně využijeme ke generování druhého pozměněného kanálu, abychom získali stereoskopické video.

DIBR je velmi častý postup, kterým můžeme vygenerovat stereoskopický obraz, a je používán napříč různými metodami. Stačí nám k tomu pouze 2D snímek a k němu odpovídající hloubková mapa. Ta většinou vypadá jako zachycený obraz s tím rozdílem, že je v šedo-tónových barvách a z původní scény zobrazuje jen obrysy bez textur. Úroveň bílé nebo černé udává zároveň vzdálenost od pozorovatele. Během pročitání odborné literatury jsem narazil na případy, kdy se hloubkové mapy napříč články lišily. Většinou byla bílá barva použita pro nejbližší objekty a černá pro vzdálené (to je mnohem častější přístup a objevuje se např. v [25], [27]). Na druhou stranu opačný trend používala metoda v publikaci [41], kde černá barva představovala nejbližší objekty. V zásadě v tom není žádný rozdíl a obě mapy nám poslouží stejně dobře, je jen nutné udržet stejný trend v celé konverzi a také se dopředu ujistit, zda je bílá barva použita pro nejbližší objekty, nebo pro nejbližší. V případě video sekvence je nutné, aby hloubková mapa byla také video sekvence se stejným počtem framů. Velmi často je pro několik po sobě jdoucích snímků videa použita totožná hloubková mapa, ale o tom bude řeč později. *Obr. 11* ukazuje, jak vypadá původní snímek a k němu následná hloubková mapa.



Obr. 11 Příklad snímku a jeho hloubková mapa [24].

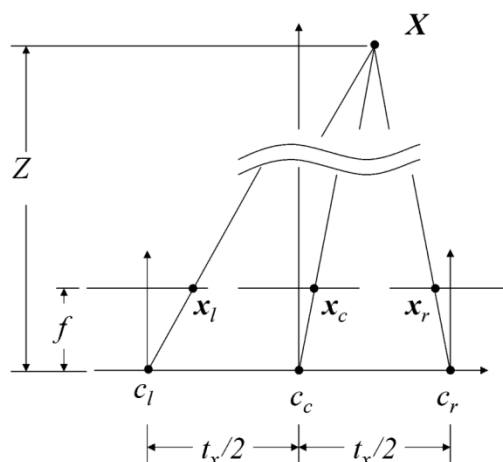
Po získání hloubkové mapy můžeme využít např. autostereoskopickou 3DTV, které stačí pro vytvoření 3D obrazu pouze zdrojová video sekvence a k ní příslušná hloubková mapa. Pokud

k dispozici nemáme zařízení, jež umí přehrávat video společně s hloubkovou mapou, musíme pomocí matematického výpočtu vygenerovat druhý snímek.

Obr. 12 představuje pohled shora na scénu, kterou chceme přepočítat pro dva snímky. Bod C_c je místo, odkud byl zaznamenán původní originální snímek. C_r a C_l představují místa, odkud by byl snímán obraz pravou a levou kamerou. Vzdálenost mezi nimi je t . Při znalosti ohniskové vzdálenosti f a hloubky scény Z můžeme vypočítat posun pixelů X_l a X_r podle vzorců (1) a (2).

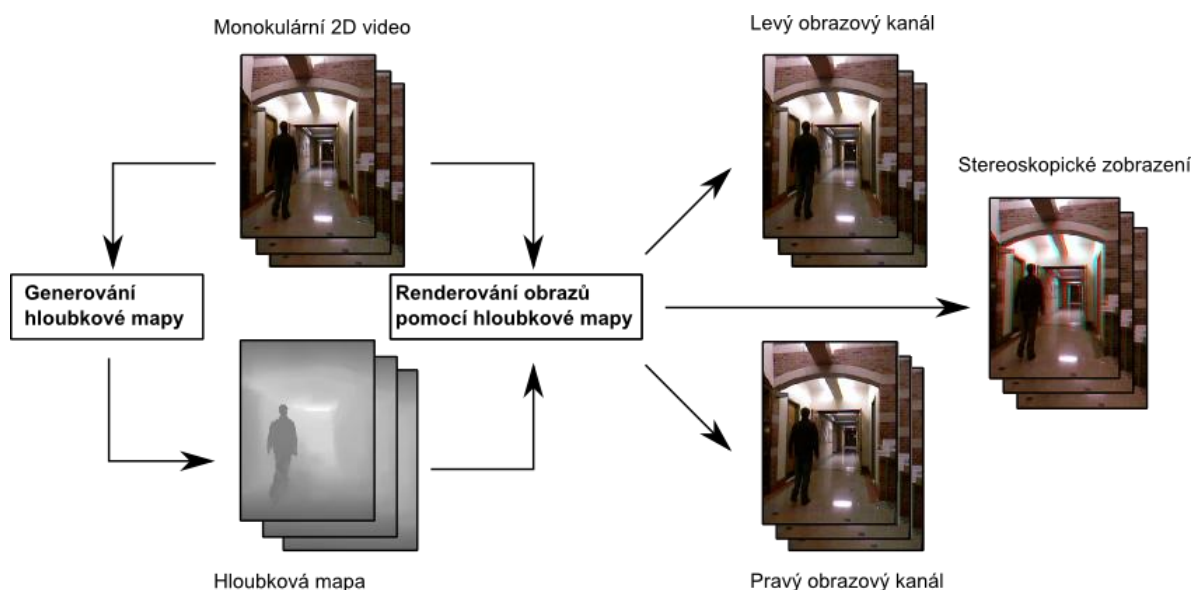
$$X_l = X_c + \frac{t f}{2 Z}, \quad (1)$$

$$X_r = X_c - \frac{t f}{2 Z}, \quad (2)$$



Obr. 12 Schéma pro výpočet horizontálního posunu pixelů [3].

Celá rekonstrukce stereoskopického obrazu pomocí DIBR tedy probíhá tak, že z monokulárního video kanálu vybíráme postupně každý frame a k němu přiřazujeme hloubkovou mapu, ze kterých vygenerujeme výsledný pravý a levý kanál. Poté už záleží jen na výběru zobrazovací metody, kterou budeme daný stereoskopický materiál prezentovat. *Obr. 13* tento postup v blokovém schématu přibližuje s tím, že jako stereoskopické zobrazení je zde použit anaglyf.



Obr. 13 Blokové schéma principu DIBR - nakresleno podle [25].

V této chvíli je nutné zdůraznit, že při generování 3D videa z 2D vstupu a hloubkové mapy (označované jako 2D+depth), se některé objekty ve druhém kanálu musí posunout, což vede ke vzniku prázdných míst. Ty je tedy nutné dále některými interpolačními metodami doplnit, aby snímek neobsahoval prázdná místa. Naštěstí se opět můžeme opřít o vlastnosti lidského vnímání, protože mozek je schopen těmto zkreslením nepřikládat velkou váhu.

Nejtěžším krokem a zároveň i výpočetně nejnáročnějším je fáze generování hloubkové mapy. U některých konkrétních implementací může být jedna mapa použita i pro několik po sobě jdoucích framů, to vše je velmi individuální.

Článek [26] shrnuje, že v případě konverze pomocí hloubkové mapy, její získání provádíme třemi možnými přístupy.

- 1) Využíváme monokulární vodítka, která byla popsána výše, a tak simulujeme funkci lidského zraku pro vnímání hloubky scény z 2D obrazu.
- 2) Máme k dispozici velkou databázi snímků společně s jejich hloubkovou mapou. Potom neznámou mapu hledáme tak, že vybereme co nejpodobnější snímky z databáze a na jejich základě stanovíme hledanou hloubkovou mapu scény. Tento přístup má výhodu v tom, že nepotřebuje využívat konkrétních obrazových vodítek. Ovšem na druhou stranu jsme velmi závislí na kvalitě a komplexnosti databáze známých scén s jejich hloubkovými mapami (více o tomto přístupu bude v kapitole 3.1.3).

- 3) Nakonec můžeme získat hloubkovou mapu scény tak, že ji nasnímáme velkým počtem kamer z několika různých úhlů. Díky tomu jsme schopni získat kompletní trojrozměrný záznam scény. Zde ovšem článek správně připomíná, že je to velmi výpočetně náročné a musíme znát přesné informace o parametrech kamer. Navíc je kompletní představa 3D uspořádání scény pro získání hloubkové mapy zbytečný luxus. O tomto principu je zmínka v [3].

2.2.6 Přehled vodítek a jejich zdrojových publikací

Přiložená tabulka kombinuje detailní odkazy z publikací [3] a [6], které byly hlavními prameny při hledání informací o monokulárních vodítkách. *Tabulka 1* podává přehled zdrojů, ve kterých je možné najít podrobnější informace k jednotlivým vodítkům, včetně matematických popisů. Pro pohybovou paralaxu nejsou uvedeny zdroje, protože není blíže specifikováno její konkrétní použití. V posledním řádku jsou uvedeny vybrané konkrétní implementace, včetně jejich zdrojových publikací.

Tab. 1 - Seznam publikací pro obrazová vodítka.

Hloubka z ostrosti objektů	
<i>Záznam scény více optickými systémy</i>	[9]
<i>Inverzní filtrování</i>	[10], [11]
<i>Vlnkové koeficienty</i>	[12], [13]
Hloubka z geometrie snímku	
<i>Lineární perspektiva</i>	[14], [15], [16]
<i>Výška obrazu</i>	[17]
Hloubka z barev a jejich intenzity	
<i>Atmosférický rozptyl</i>	[14], [19]
<i>Stíny</i>	[21]
<i>Hrany objektů</i>	[22], [23]
<i>Teplota barev</i>	[20]
Konkrétní implementace	[18], [25], [26], [27], [33], [34], [35], [36], [37]

3 Popis konkrétních implementací

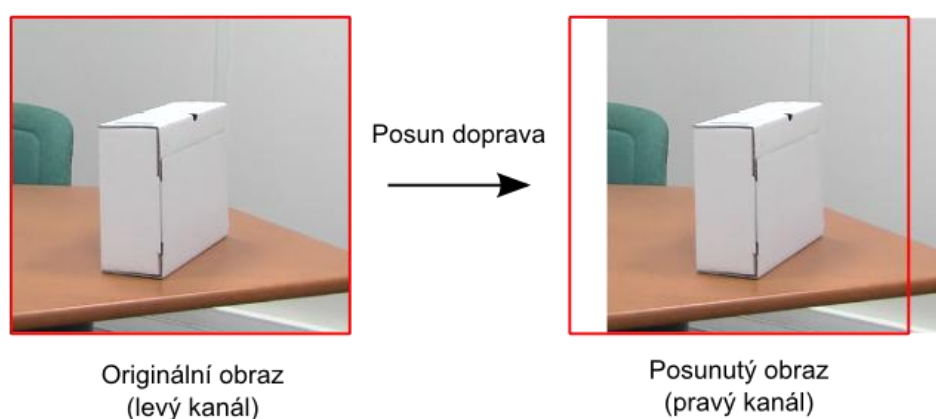
V této části se stručně podíváme na několik konkrétních případů, skrze které lze převést obsah do 3D. Nejprve dvě metody krátce popíši, dále přiblížím podrobněji postup, který získává mapu pomocí databáze, a nakonec přijdou na řadu hardwarová a softwarová řešení.

3.1 Teoretické metody

3.1.1 Prostorově - časová konverze

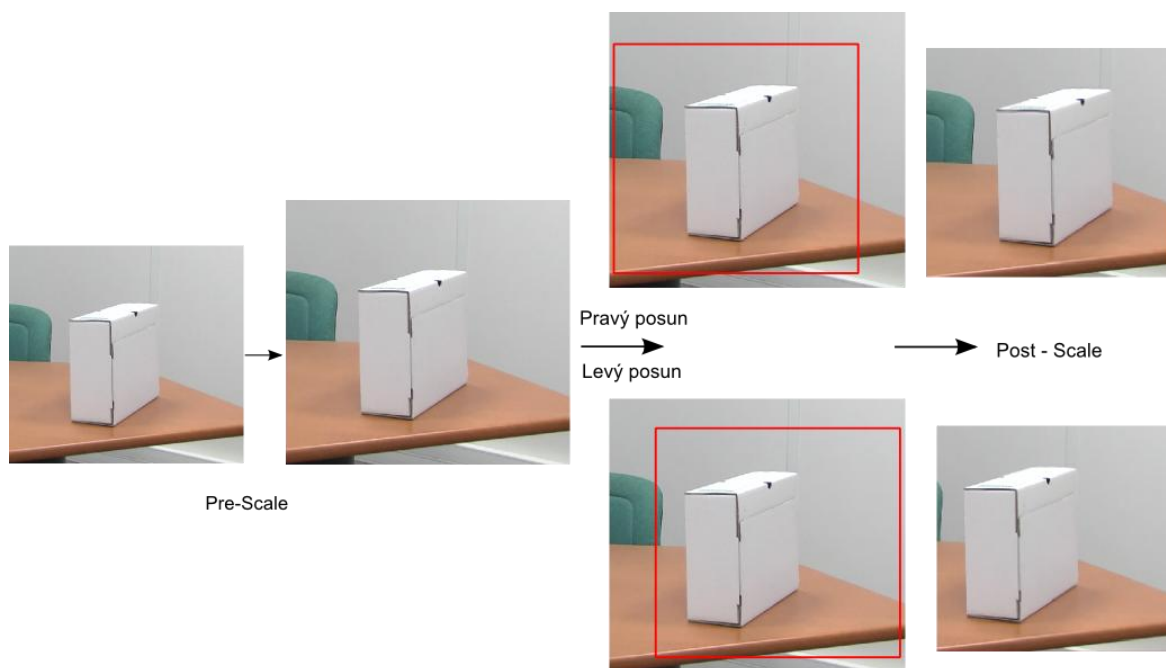
V článku [27] je popsán komplexnější způsob, jakým lze generovat 3D. Celá metoda je rozdělena na dvě části v závislosti, zda pracujeme se statickým snímkem, nebo využijeme detekci pohybu.

Pro statické snímky, případně statické části obrazu se konverze řeší pomocí prostorového posunu lokálních částí snímku. Původní snímek reprezentuje levý kanál a pravý se vytvoří posunutím části obrazu doleva a přebytečná část na pravé straně se ořízne. Samozřejmě tím vznikne na levé straně volné místo, kde obrazová informace kvůli posunutí chybí (*obr. 14*).



Obr. 14 Princip posunutí obrazu s chybějící informací.

Proto je nutné na začátku obraz lehce upravit a roztáhnout (*pre-scale*), abychom z něj poté jako levý kanál vyřízli pouze část více vpravo a pro pravý kanál část více vlevo. Tím nám nebude chybět obrazová informace na okrajích. *Obr.15* ilustruje princip postupu.



Obr. 15 Princip prostorového posunutí

V případě, že máme v sekvenci pohyb, můžeme využít konverze pomocí hloubkové mapy. Algoritmus předpokládá, že pohyblivé objekty jsou v popředí obrazu, kdežto statická část leží v pozadí. Nejprve se detekuje pohyb ve snímku. Dále se naleznou pohyblivé objekty a následně se zvýrazní hrany pomocí konvoluce s Laplaceovým jádrem ve čtyř okolí. Poté se detekované objekty vyplní a zvolí se pro ně vhodná hloubka v závislosti na okolí. Posledním krokem je *post-processing*, který vyhlazuje vytvořenou mapu. Pomocí mapy a **DIBR** pak algoritmus získá druhý obrazový kanál.

3.1.2 Metoda pro konverzi v reálném čase

Jako druhá bude přiblížena metoda popsána v článku [18]. Od předchozí i od následující se liší hlavně tím, že je schopná konvertovat video v reálném čase. Tato metoda používá spojení několika monokulárních vodítek ze statických snímků. Pro rychlou a nenáročnou konverzi videa v reálném čase se autoři zaměřili na to, aby výsledek byl pro diváka na pohled příjemný, spíše než aby hledali přesnou geometrii scény. Postup samotný se skládá ze dvou kroků, nejprve se stanoví globální hloubka scény a poté se ještě doplňují a upravují lokálně další místa pomocí odlišných obrazových vodítek.

Globální stanovení hloubkové mapy je na principu uvedeném v kapitole 2.2.2 jako **výška v obraze**. Tedy lidé vnímají jako bližší objekty ty, které jsou v zorném poli dole spíše než objekty nahoře. Pro další zpřesnění se zkoumají barevné a světelné gradienty a bližšími objekty jsou prohlášeny ty, které mají teplé barvy, zatímco studené barvy představují vzdálenější (vodítko popsané v kapitole 2.2.3 jako **teplota barev**).

Z těchto poznatků je následně sestavena hloubková mapa a dále probíhá výpočet stereoskopických snímků podle **DIBR** schématu popsáném v kapitole 2.2.5.

Aby bylo možné konverzi provést v reálném čase, autoři svůj algoritmus optimalizovali, aby CPU pracovní stanice měl za úkol pouze získání hloubkové mapy, zatímco vygenerování pravého a levého snímku provádí grafická karta. V článku je následně také porovnán tento postup pro získání mapy s dvěma dalšími algoritmy, kde jeden stanovuje mapu na základě pohybových vektorů a druhý na základě detekce hran. Pro všechny tři byly provedeny subjektivní testy pro různé typy sekvencí a tento algoritmus se jevil jako nejkompaktnější. Navíc byl i nejrychlejší a konverze jednoho snímku včetně nalezení mapy byla okolo 900ms při použití pouze jednoho vlákna procesoru. Při výše zmíněném rozdělení úkolů mezi GPU a více jádrové CPU by byl výsledný čas výrazně rychlejší a bylo tedy možné konvertovat 2D na 3D v reálném čase.

3.1.3 Neparametrické vzorkování

V kapitole 2.2 bylo popsáno mnoho různých přístupů a vodítek, kterými můžeme získat hloubkovou informaci ze snímků pořízených jednou kamerou, abychom je mohli následně konvertovat na stereoskopické. Metoda popsána v [25] a také v doprovodném textu [28] ovšem využívá řešení, které jsem zmínil na konci kapitoly 2.2.5, kdy je hloubka scény určena díky předem pořízeným testovacím sekvencím. Postup je možné aplikovat jak na jednotlivé snímky, tak i video. Není třeba žádného doplňujícího zásahu uživatele při běhu algoritmu, krom nastavení několika počátečních parametrů, tj. metoda je poloautomatická. Samotní autoři potvrzují, že metodu vyzkoušeli pro velké množství sekvencí a následně i část filmu *Charade* s velice dobrými výsledky.

Klíčový prvek celého algoritmu je databáze video sekvencí, které autoři nazývají jako RGBD snímky, protože obsahují RGB složky + hloubku (*depth*). K pořízení porovnávací databáze byla použita dvě zařízení Kinect, která umí pomocí senzoru měřit hloubku scény. Díky tomu obsahuje databáze videí jak hloubku určenou ze senzoru, tak i hloubku určenou ze dvou obrazových kanálů.

Postup pro stanovení výsledné mapy má v zásadě tři základní kroky:

1) V databázi pořízených snímků se snažíme najít ty, které jsou v RGB soustavě co nejpodobnější tomu, u kterého chceme získat výslednou mapu.

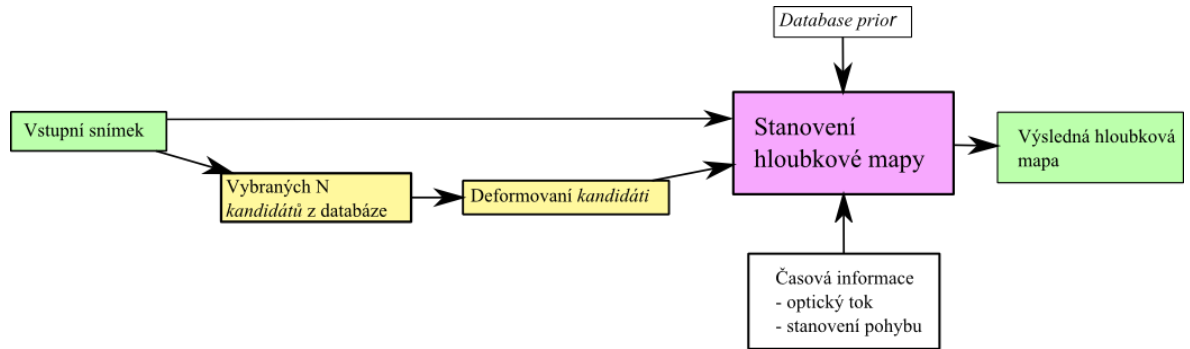
2) Pomocí *SIFT Flow* algoritmu upravujeme databázové snímky a jejich mapy tak, aby se co nejvíce blížili konvertovanému snímku. Metoda vybírá vždy 7 nejpodobnějších snímků k hledanému s tím, že mezi sedmi vybranými nesmí být více snímků ze stejné video sekvence, a s těmi dále pracuje pro získání výsledné mapy.

3) Poslední krok má za úkol pouze vyhledit a upravit hloubkovou mapu pro předem zadaný snímek.

Další velmi zajímavou částí této metody je úprava pro lepší využití pro video, protože při stanovení hloubky pro každý snímek zvlášť by mohlo docházet ke skokovým změnám hloubky. Ve výsledku je vlastně přidán mezikrok, ve kterém je stanovena obecná hloubka přes všechny framy s tím, že jsou detekovány i pohyblivé objekty.

3.1.3.1 Podrobnější teoretický popis

V předchozích odstavcích jsme se krátce podívali na princip, jakým je vytvářena hloubková mapa při použití algoritmu **Neparametrické vzorkování**. Tyto tři části nyní podrobněji probereme. Na *obr. 16* je blokové schéma konverze při použití všech doplňujících úprav.



Obr. 16 Blokové schéma metody Neparametrického vzorkování - nakresleno podle [25].

Algoritmus tedy probíhá tak, že nejprve hledá ke vstupnímu snímku co nejpodobnější snímky v RGB prostoru z databáze. Následně tvaruje a deformuje vybrané snímky tak, aby byly zarovnané podle vstupního obrázku (stejně kroky se provádí i s hloubkovými mapami snímků). Nakonec přichází konečné úpravy pro vyhlazení výsledné mapy.

Nyní se podíváme na algoritmus, který je schopen upravit a vytvarovat vybrané snímky tak, aby korespondovaly se vstupním obrázkem. (Funkci výběru vhodných snímků z databáze se věnovat nebudu, ale podrobnější informace obsahuje druhý článek s dodatky [28].)

Když je tedy nalezeno dostatečné množství K nejpodobnějších obrazů (typický autor používá $K = 7$), jsou tyto obrazy prohlášeny za *kandidáty*. Žádný z těchto 7 *kandidátů* nesmí být ze stejné video sekvence. Předpokládáme tedy, že když se *kandidáti* podobají vstupu, tak i hloubkové mapy budou odpovídat. Chceme, aby přesně korespondovaly pixely *kandidátů* s pixely vstupního obrazu, a proto je používán **SIFT flow** proces. Každý *kandidát* se zdeformuje, a získáme tak deformační funkce, ψ_i , $i \in \{1, \dots, K\}$. Tím budeme mít i deformované hloubkové mapy každého *kandidáta*.

Ted' přichází na řadu určení výsledné mapy ze získaných hloubkových map deformovaných *kandidátů*. K tomu poslouží vzorec (3), \mathbf{L} je vstupní obraz a \mathbf{D} je hledaná hloubková mapa, kde Z je normalizační konstanta pravděpodobnosti, α a β jsou váhové parametry.

$$-\log(P(\mathbf{D}|\mathbf{L})) = E(\mathbf{D}) = \sum_{i \in \text{pixely}} (E_t(\mathbf{D}_i) + \alpha E_s(\mathbf{D}_i) + \beta E_p(\mathbf{D}_i) + \log(Z)), \quad (3)$$

Pro jediný vstupní obraz se zanedbávají úpravy pro pohyb a zůstávají nám tu tak tři podmínky: Část E_t měří, jak moc je podobná výsledná mapa jednotlivým kandidátům (podrobnější matematický výpočet pro všechny tři složky je ve [26]). E_s je prostorové vyhlazení v místech, kde jsou malé intenzitní gradienty, a E_p je porovnávání s průměrem všech map v databázi (v *obr. 16* je to nazváno jako *Database prior*).

Při použití tohoto postupu na video sekvenci by vznikaly různé časové nesrovnalosti a skokové změny v mapách mezi jednotlivými framy. Proto autoři doplňují výše zmíněný vzorec a tím získávají nový (4). Doplněný postup je schopný konvertovat i sekvence, kde jiné metody selhávaly.

$$E_{\text{video}}(\mathbf{D}) = E(\mathbf{D}) + \sum_{i \in \text{pixely}} \nu E_c(\mathbf{D}_i) + \eta E_m(\mathbf{D}_i), \quad (4)$$

E_d zůstává z předchozího příkladu a pouze ho doplníme o E_c , což popisuje časovou návaznost, zatímco E_m využívá pohybových vodítek k vylepšení hloubkové informace pohybujících se objektů. Koeficienty ν a η opět pouze váhují jednotlivé složky (autoři použili $\nu = 100$ a $\eta = 5$).

V článku jsou opět podrobněji matematicky popsány jednotlivé části vzorce. Dále se budu věnovat praktickému vyzkoušení konverze, zda dosahuje avizovaných kvalitních výsledků, či nikoliv.

3.1.3.2 Testování v praxi

Pro testování metody **Neparametrické vzorkování** použiji předem připravené pokusné m-fily¹⁰ od jejího autora, které demonstrují použití algoritmu pro stanovení hloubky obrazu a následně hloubkové mapy.

K dispozici mám stolní počítač s procesorem Intel Core i5-2500K@3,3 Ghz, s operační pamětí 8GB RAM. Celý MATLAB 2012a je nainstalován na SSD disku, kde je zároveň pro zvýšení rychlosti i celá databáze videí (její velikost je 19GB).

V době testování tohoto algoritmu jsem byl v pravidelném kontaktu s autorem a ten mi poskytnul další kód pro MATLAB, díky kterému lze ze získané mapy a původního snímku vypočítat levý a pravý obrazový kanál a také jejich zobrazení pomocí anaglyfu.

Autor připravil čtyři příklady k prokázání účinnosti algoritmu. Každý z příkladů se lišil typem sekvence a počtem snímků, z čehož se odvíjela i doba konverze. Navíc první tři příklady využívaly pouze omezenou databázi, která byla nutná ke správnému chodu. *Příklad IV* pak pracoval s kompletní sadou videí, kterou jsem stáhnul z výše uvedeného odkazu. Při prvním spuštění vždy program načítal dlouho databázi (cca 2 minuty databáze pro *Příklady I – III* a cca 30 minut pro *Příklad IV*). I přesto program nepracoval se všemi dostupnými snímky (základní databáze se dala rozšířit v nastavení m-filu). I při vypnutí MATLABu ovšem zůstala databáze nahrána, a tak jsem tento čas nezapočítal do celkového trvání konverze jednotlivých příkladů.

¹⁰ Na webových stránkách <http://research.microsoft.com/en-us/downloads/29d28301-1079-4435-9810-74709376bce1/> je k dispozici kód v MATLABu společně s celou databází videí nutnou ke správnému běhu programu.

3.1.3.3 Příklad I

Zde je cílem ukázat konverzi na sekvenci, kde je statická kamera a pohybující se objekt. Ukázka se skládá ze tří framů, rozměr jednoho snímku je 430×579 pixelů a celková doba převodu trvala skoro 2,5 minuty. *Obr. 17* ukazuje původní snímek, jeho hloubkovou mapu a poté výstup v anaglyfu.



Obr. 17 Výsledek příkladu I - 1. frame.

3.1.3.4 Příklad II

Ve druhém příkladu už sekvence obsahuje framů 7, velikost snímku je 640×480 pixelů a ve scéně se kamera pohybuje zleva doprava. Čas nutný pro výpočet sedmi map vzrostl na 6 minut a 45 sekund. Na *obr. 18* je ukázka výsledků.



Obr. 18 Výsledek příkladu II - 1. a 7. frame.

3.1.3.5 Příklad III

V tomto příkladu je konvertována sekvence, která čítá 40 snímků. Bohužel při konverzi všech 40 snímků došlo k přetečení celé paměti RAM, a proto jsem vybral pouze prvním 30 snímků.

Velikost v pixelech je opět 640×480 . Sekvence obsahuje pohyblivé objekty a kamera při tom scénu přibližuje. Čas výpočtu byl tentokrát celých 33 minut! *Obr. 19* zobrazuje 1., 15., a 30. frame.



Obr. 19 Výsledek příkladu III - 1, 15. a 30. frame.

3.1.3.6 Příklad IV

Při posledním příkladu je použita velká část databáze. Sekvence obsahuje 60 snímků, ale z výše uvedených důvodů jsem musel vybrat pouze prvních 30. Velikost jednoho snímku je 215×290 pixelů. Statická kamera sleduje pohybující se postavu. Doba konverze byla 30 minut. Na *obr. 20* je opět 1., 15., a 30. frame.



Obr. 20 Výsledek příkladu IV - 1, 15. a 30. frame.

3.1.3.7 Subjektivní pozorování výsledků

Předchozí případy byly testovány pouze v domácích podmínkách a zobrazeny pomocí anaglyfu. Protože barevné filtry vedou k degradaci barev, bylo nutné ohodnotit výsledky experimentu pomocí jiné metody. Na fakultě jsem tedy provedl testy na aktivním plasmovém zobrazovači značky Panasonic.

K dispozici jsem měl tři video sekvence, které jsem vytvořil stejně jako předešlé. Jedna ze sekvencí byla stejná jako v *Příkladu IV* na *obr. 20*. Další dvě byly velmi podobné. Jejich délka byla 60 framů. Přípravné práce na těchto celkových šesti sekundách byly velmi dlouhé. Jak jsem zmiňoval dříve, získání mapy pro každý snímek trvá zhruba minutu, takže čistý čas přípravy mapy byl 180 minut. K tomu je nutné připočítat čas, kdy se z mapy a původního snímku počítal levý a pravý kanál. Dále bylo třeba vytvořit sekvenci v rozlišení 1920×1080 pixelů pro správné zobrazení na 3DTV, protože ta při SBS zobrazování měnila stranový poměr snímku.

Na zobrazovači jsem poté sledoval, jak se jevíly sekvence z hlediska kvality vnímané hloubky scény, zda byly v obraze přítomny artefakty či snímky nezpůsobují únavu očí.

Ve všech snímcích byl patrný 3D vjem. Jeho kvalita předčila očekávání, nicméně obrazová disparita byla chvílemi nepříjemná, a to vedlo k bolestem očí. Pro zlepšení by bylo nutné lépe nastavit primární získávání mapy. V obraze vznikalo několik artefaktů, z nichž nejvýraznější bylo nerovnoměrné šedé orámování obrazu. Navíc nebylo po celou dobu konstantní ani symetrické a působilo velmi rušivě. Poslední typ artefaktu vznikal okolo pohybující se postavy, kdy se v jednu chvíli začala část obrazu okolo postavy ohýbat. Při letmém pohledu to však nebylo tak markantní jako zmiňovaný rám. Díky nedostatku paměti byla sekvence rozdělena na dvě části, a tak algoritmus nemohl využít vyhlazování přes všechny framy. Při podrobném zkoumání byla v polovině sekvence pozorována lehká nespojitost, ale nešlo o nic zásadního.

3.1.3.8 Zhodnocení metody

Na předchozích příkladech jsem vyzkoušel, jak funguje tato nová a netypická metoda v praxi.

Její hlavní výhodou je kvalita výsledků. I přes malý vzorek provedených testů mohu potvrdit, že by měla být použitelná v širokém spektru typů snímků, jak uvádí autor metody. Bohužel hlavní nevýhodou je čas nutný k výpočtu hloubkové mapy a také hardwarová náročnost. V praxi trvá stanovení hloubkové mapy pro jeden snímek zhruba minutu, a to je obrovské číslo, zároveň není možné počítat najednou přes více než 30 snímků, protože RAM 8GB více nezvládne. Při konverzi delšího videa by tedy bylo nutné počítat mapu po částech, ale v tom případě není možné zaručit, že bude výsledek plynule navazující. Časová náročnost v podstatě vylučuje její větší rozšíření při konverzi filmových materiálů

Možné řešení by bylo rozdělit sekvenci po 30 snímcích a tyto skupiny konvertovat každou zvlášť. Ale vzhledem k době trvání to také není optimální přístup. Navíc pak dochází k nespojitostem scény.

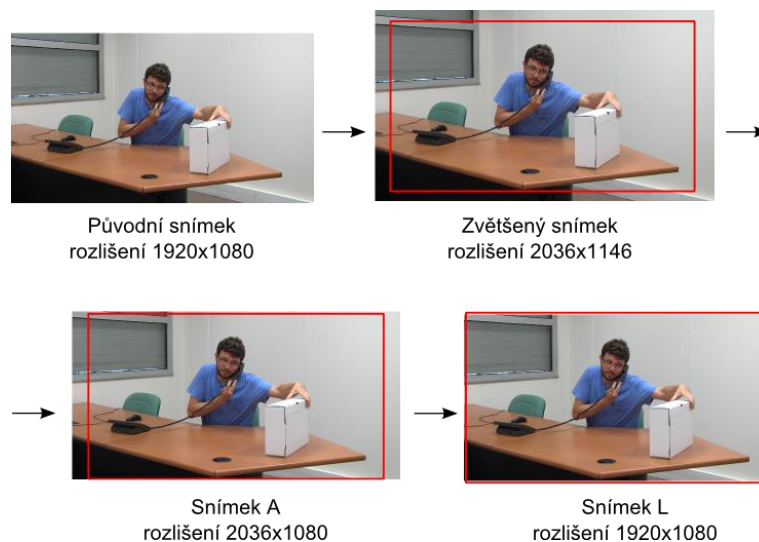
Poslední velkou nevýhodou je fakt, že databáze není uzpůsobená pro konverzi venkovních sekvencí.

3.2 Navržená metoda

Protože nebylo z technických důvodů možné použít Blu-Ray přehrávač, do subjektivního srovnání (viz kapitola 4.1.2.1), rozhodl jsem se navrhnout jednoduchý algoritmus, který by princip konverze tohoto zařízení suploval. Můj předpoklad byl rovněž takový, že pokud přehrávač převádí video do 3D v reálném čase, nepočítá hloubkovou mapu pro scénu, ale spíše simuluje hloubku obrazu pomocí určitého posouvání a natáčení scény. Oporu jsem našel v článku [27] a také v Datasheetu k procesoru Sony [32]. V prvně jmenované publikaci se popisuje princip pro vytvoření druhého kanálu pomocí posunu lokálních částí obrazu a ve specifikacích od firmy Sony je zase uvedeno, že lidský zrak lze oklamat určitým nakloněním či pootočením obrazu, a tím vznikne 3D vjem. Dále jsem sledoval materiál, který přehrávač konvertoval, a při přepínání mezi 2D a 3D režimy jsem mohl sledovat rozdíly mezi kanály. Levý a pravý kanál byl ve 2D zobrazení logicky totožný a při přepnutí do 3D zůstal rovněž levý kanál stejný, ale pravý byl pozměněný. Právě tyto experimentálně zjištěné změny byly základem pro můj algoritmus, který do jisté míry kopíruje viditelné kroky Blu-Ray přehrávače.

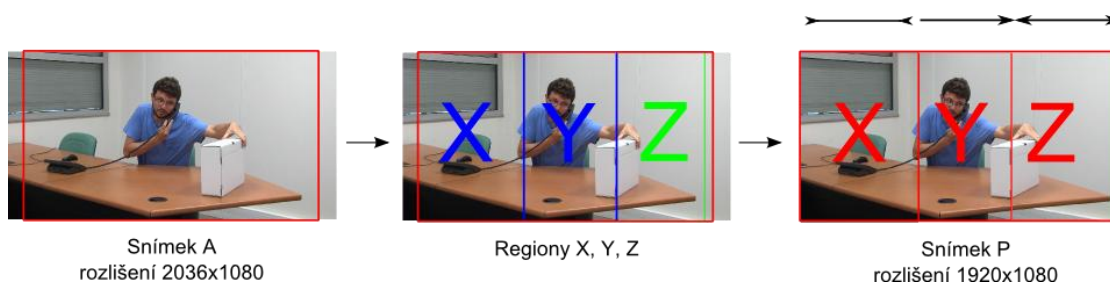
V následujícím odstavci popíši princip tohoto postupu, který je prováděn na každém jednotlivém snímku vybrané sekvence. Metoda začíná tím, že pomocí interpolace zvětší obraz (podobně jako *pre-scale* u metody v kapitole 3.1.1). Já jsem zvolil zvětšení o 6% a tato velikost má do jisté míry vliv na pozorovanou hloubku scény¹¹. Tedy konkrétně bylo nové rozlišení framu 2036×1146 pixelů (červený rám označuje standardní velikost FullHD – 1920×1080). Protože princip funguje na bázi horizontálního posunu, můžeme ihned oříznout symetricky ze snímku horní a dolní pruh tak, že vznikne obraz s 2036 sloupci a 1080 řádky. Výsledek nazveme obraz A. Následuje výběr části o rozměrech standardního FullHD rozlišení, tj. 1920×1080 pixelů tím, že odřízneme symetricky pruhy na levém i pravém konci, a tento výběr prohlásíme levým snímkem a budeme ho nazývat obraz L. *Obr. 21* vysvětluje získání obrazu A a L.

¹¹ Při konvertování některými programy popsány níže byl k dispozici náhled výsledku společně s posuvníkem pro velikost hloubky scény. Při zvyšování hloubky se výsledný obraz více roztáhl, čehož jsem využil také.



Obr. 21 Princip získání snímku A a L.

Pro získání pravého snímku je nutné provést dále několik dalších úprav. Obraz A, který má rozlišení 2036×1080 pixelů rozdělím na tři sektory a tím dostanu části X (802×1080 pixelů), Y (432×1080 pixelů) a Z (802×1080 pixelů). S každou touto částí provedu jinou úpravu. Část X musím zmenšit na rozměry 773×1080 pixelů, část Y zůstane beze změny a z části Z vyberu jen 686 sloupců a tedy dostanu oblast o rozměrech 686×1080 pixelů. Tuto novou část navíc musím roztáhnout na rozměr 715×1080 pixelů, aby po složení vedle sebe všechny regiony daly dohromady šířku 1920. Na obr. 22 je tento proces schematicky popsán, kde šipky u snímku P představují, co se s jednotlivými částmi stalo¹².



Obr. 22 Ilustrace změn regionů.

Toto pokroucení jednotlivých částí obrazu je nutné, protože kdybych jen posunul celý snímek, tak bychom ho na 3DTV viděli pouze jako vystupující nebo naopak zapuštěný v obrazovce podle směru posunu. Kód pro Matlab je přiložen na DVD.

¹² Oblast X se zúžila, ale protože obsahovala i část, která v levém kanálu na levé straně chybí, tak po zúžení stejně posunula celý blok Y o kus vpravo, který jinak zůstal nezměněn. Blok Z se naopak trochu oříznul oproti levému kanálu, a musel se tedy rozšířit, aby snímek zachoval FullHD rozlišení.

3.3 Komerční řešení

Abychom konvertovali video z 2D do 3D, můžeme použít různá hardwarová zařízení, která jsou schopná převádět obsah v reálném čase. Jedná se buď o 3D procesory, které patří do skupiny profesionálních zařízení, nebo boxy pro domácí použití, případně samotné 3DTV nebo Blu-Ray přehrávače. V následující části se podíváme, jaké procesory se dají pořídit a co nám mohou nabídnout. Veřejně bohužel není dostupná podrobná dokumentace, a tak můžeme čerpat pouze z oficiálních informací od výrobce, které poskytují na svých webových stránkách.

3.3.1 Profesionální 3D procesory

3.3.1.1 Blackmagic Teranex 3D processor

Firma Blackmagic zabývající se převážně profesionálními řešeními nabízí procesor Teranex 3D. Toto zařízení, které je na *obr. 23*, se v České republice dá pořídit u tří dodavatelů¹³ a jeho cena v současné době je skoro 99 000,- korun včetně DPH.



Obr. 23 Blackmagic Teranex 3D [29].

Zařízení je vyspělejší verzí procesoru Teranex 2D a přináší větší možnosti právě v práci s 3D obrazem. Pro 3D obraz procesor nabízí široké nastavení pro konverzi mezi jednotlivými formáty, je vybaven rozhraními 3D/HD/3G-SDI a HDMI a také rozhraním Thunderbolt pro další digitální zpracování. Při práci s jednotlivými 3D kanály procesor umí měnit jejich rozlišení a také poměr stran, synchronizaci různého *frame rate* mezi snímky, případně úplně změnit počet snímků za vteřinu. V kanálech také dovede odstraňovat pulsnímky. A konečně v režimu 3D simulace dovede konvertovat 2D signál na 3D stereoskopický, u kterého lze upravovat i geometrii záběrů.

O tento procesor jsem žádal oba distributory k zapůjčení, abych mohl otestovat jeho vlastnosti, ale bohužel ani jedna firma neměla momentálně verzi 3D k volnému zkoušení.

¹³ www.disk.cz, www.syntex.cz a www.audioworks.cz

3.3.1.2 JVC IF-2D3D1 Image Processor



Obr. 24 JVC IF-2D3D1 Image Processor [30].

Dalším zařízením na trhu je procesor od firmy JVC IF - 2D3D1, který rovněž provádí konverzi obsahu z 2D do 3D. Stejně jako předchozí stroj i tento nabízí ve svém obchodě firma Disk. Zařízení ovšem také nebylo možné zapůjčit a jeho cena se pohybuje dokonce okolo 787 000,- korun i s daní.

Konverze v reálném čase je samozřejmostí a nadto nabízí přístroj několik výstupních 3D formátů (*Line-by-line*, *Side-By-Side*, *Top-Bottom*, *Checkerboard*), dále nástroje pro upravení paralaxy a 3D intenzity a možnost vytvořit společný výstup pro levý a pravý kanál včetně řádné synchronizace. Uživatel má k dispozici rovněž vektorskop a *waveform* monitor.

3.3.1.3 SONY MPE200 Multi Image Processor



Obr. 25 SONY MPE200 Multi Image Processor [31].

Poslední zařízení z této kategorie nabízí firma, která je na trhu známá převážně díky domácí elektronice, konkrétně Sony. Procesor MPE200 jsem v žádném českém katalogu nenašel,

nicméně na oficiálních zahraničních stránkách výrobce je udána cena 38 000\$ + 8000\$¹⁴, protože k procesoru je nutné dokoupit firmware, který umožňuje konverzi do 3D.

Přístroj tedy může konvertovat 2D do 3D pomocí vlastního algoritmu, dále měnit a upravovat jasové a chrominanční složky snímků, zvládá práci se čtyřmi signály najednou a samozřejmě je zde velká nabídka výstupních formátů. Výhodou oproti předchozím procesorům má tento stroj v tom, že nabízí GUI rozhraní, pomocí kterého se procesor ovládá přes PC.

Firma Sony byla navíc jediná, která reagovala na mou žádost a pro studijními účely mi poslala krátký dokument [32], kde je něco málo řečeno o principu, na kterém zařízení konvertuje obsah do 3D. Principem pro vytvoření 3D výstupu je určení nejprve tzv. lokální hloubky, tedy hloubky, která se týká jen některých výrazných a dominantních objektů. Jako obrazové vodítko se používá zaostření objektů a výrobce připomíná, že pokud uživatel zostruje hrany těchto objektů (změnou jasové složky), tak tím zároveň mění i hloubku ve scéně. Vznikat mohou i různé nedostatky, kdy je na objektu část neostrá, a tím se tedy vyhodnotí jako hlubší místo než okolí, a naopak velmi ostré prvky budou vystupovat z obrazu. Po určení lokální hloubky přichází na řadu globální hloubka scény. Ta je založena na optickém klamu, kdy lidský zrak vnímá scénu jako prostorovou, pokud jsou levý a pravý kanál nakloněny. Pravděpodobně je princip k získání globální hloubky častou implementací v ostatních zařízeních či programech.

¹⁴ <http://pro.sony.com/bbsc/ssr/product-MPE200/> a <http://pro.sony.com/bbsc/ssr/product-MPESFX01%2F01/>

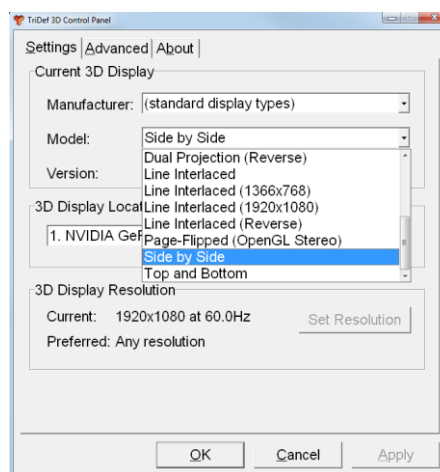
3.3.2 Software - automatické metody

3.3.2.1 TriDef 3D

TriDef 3D je komerční software, který umožňuje automaticky konvertovat obsah z 2D do 3D. Pro firmy prodávající TV techniku nabízí TriDef firmware, který je implementován do 3DTV, pomocí něhož TV konvertuje libovolný zdroj do 3D. Mezi podporované značky patří LG a Samsung a firmware lze nahrát i do několika Blu-Ray přehrávačů. V případě, že chceme využít konverzi do 3D v domácích podmínkách, je možné stáhnout do PC TriDef 3D software, jenž je schopný převést velké množství různého obsahu. Konkrétně nabízí konverzi fotografií, video souborů, nebo celých DVD a také podporuje hraní her ve 3D. Jeho cena je v době psaní této práce 40 dolarů. V případě, že chceme software pouze vyzkoušet, je k dispozici časově omezená Trial verze se všemi dostupnými funkcemi.

Díky tomuto softwaru tedy můžeme okamžitě přehrávat filmy ve 3D, protože konverze probíhá v reálném čase. Z toho plynou vyšší nároky na výkon domácího PC než při přehrávání běžného videa, ale na druhou stranu je nutné říci, že v dnešní době velké procento PC v klidu splňuje minimální požadavky doporučené výrobcem¹⁵.

Z hlediska nastavení je uživatel poměrně omezen. Můžeme volit pouze výstup grafické karty, rozlišení displeje a režim zobrazení. Na *obr. 26* je ukázka jednoduchého rozhraní pro nastavení možností zobrazení 3D.



Obr. 26 TriDef 3D nastavení výstupního zobrazení.

¹⁵ Doporučené požadavky jsou na <http://www.tridef.com/product-support/faqs>.

V podstatě jedinou možností je vybrat výrobce displeje, kde už je několik profilů pro zobrazení přednastaveno. V případě, že chceme zvolit jiné zobrazení, můžeme ponechat výrobce neznámého a vybrat si konkrétní režim 3D zobrazení. Zde máme k dispozici klasické profily jako je Side-by-Side, Top-Bottom a mnohé jiné.

Pro subjektivní testování je tato metoda použitelná, ale musíme zaznamenat výstup z přehrávače, abychom ho měli uložený jako video sekvenci, kterou později s ostatními subjekty přehrajeme na PC.

3.3.2.2 PowerDVD 13

Tento komerční software je velmi podobný předchozímu, protože opět umožňuje v reálném čase přehrávat videa a DVD ve 3D. I zde stačí zvolit obsah a integrovaný přehrávač ihned začne přehrávat obsah ve 3D. U PowerDVD 13 dokonce můžeme v nastavení volit velikost hloubky scény. Z hlediska nastavení zobrazovaného režimu je to u tohoto programu složitější.

Běžná praxe je taková, že pokud pustíme video v režimu SBS na panelu, který neumí 3D zobrazovat, uvidíme jen dva obrazové kanály vedle sebe s poloviční šířkou a teprve po zapnutí režimu 3D na TV se obrazové kanály proloží přes sebe. U PowerDVD přehrávače to tak nefunguje. Při zvolení režimů se automaticky rovnou přepne obraz do stavu, jako kdybychom měli 3DTV a s brýlemi sledovali obsah. Pouze anaglyf funguje správně v případě, že nemáme k dispozici 3DTV. Zobrazení v režimu SBS v 2D provedení není možné, a tím tato metoda není použitelná pro naše subjektivní hodnocení, protože nelze zobrazit a následně uložit video výstup v režimu SBS, kde by byly dva kanály vedle sebe. Cena tohoto softwaru se pohybuje okolo 70 dolarů, ale lze opět použít časově omezenou trial licenci.

3.3.3 Software - poloautomatické metody

Do poloautomatických komerčních metod zařadíme ty programy, které umožňují převod videa do 3D a zároveň je možné nastavit více parametrů týkajících se konverze. Tyto metody už tedy nejsou schopné pracovat v reálném čase, ale po zvolení a nastavení procesu následně nějaký čas potřebují k tomu, aby vyexportovaly výsledky, se kterými pak lze pracovat nezávisle na těchto programech.

3.3.3.1 DVDFab 3D Converter

Stejně jako PowerDVD je DVDFab primárně určen jako software pro vypalování a kopírování DVD disků, ovšem jeho další funkcí je i konverze z 2D do 3D. Díky tomu, že konverze neprobíhá v reálném čase, si můžeme zvolit a nastavit velkou škálu parametrů. Nejprve vybereme soubor, který chceme převést do 3D, a následně zvolíme jeho výstupní video formát (na výběr je *mp4*, *mkv*, *avi* a další), je zde zároveň na výběr několik typů audio kódování. V dalším kroku máme na výběr úpravy, kde můžeme měnit výstupní rozlišení, bitový tok a také snímkovou frekvenci. V záložce 3D nastavení lze opět vybrat z mnoha 3D zobrazovaných formátů, dále velikost hloubky scény a dokonce můžeme zvolit, zda do videa chceme vložit pevné titulky. Nakonec můžeme vybrat další soubor do fronty úkolů, a když jsme hotovi, lze celou frontu spustit a všechny soubory se jeden po druhém podle našeho nastavení převedou.

Tento software lze používat ve třicetidenní časově omezené trialové verzi, kde ovšem máme k dispozici všechny funkce plné verze. Cena plné verze Converter balíčku se pohybuje okolo 50 dolarů.

Všechny tyto zmíněné programy jsem vybral do subjektivního testování (bude popsáno níže), abych vyzkoušel, jak kvalitní jsou metody, které si můžeme dovolit využívat v domácích podmínkách. Protože jsou tyto metody prodávány, výrobci si nemohou dovolit poskytnout široké veřejnosti podrobnou dokumentaci, a tak o principu jejich konverze není k dispozici mnoho detailů. Ovšem z hlediska kvality a rychlosti, kterou jsou programy schopné generovat 3D video, můžeme usoudit, že se nevytváří kompletní hloubková mapa scény, ale maximálně se určí dominantní prvky a poté se pouze použije posunu a případně informace o pohybu, jak je popsáno v metodě Prostorově-časová konverze v kapitole 3.1.1.

3.3.3.2 Další komerční software

Pro poloautomatickou konverzi je na internetu k dispozici celá řada dalších nástrojů, nicméně proti zařazení do subjektivních testů bylo několik faktorů. V první řadě šlo o to, že jejich volně dostupné verze neumožňovaly plnohodnotnou konverzi. Mezi tyto programy patří 4Videosoft Video Converter, jehož výstup poté vůbec nevypadal jako 3D video, nebo 3D Video Converter, kde výstup vůbec nešel přehrát.

Poslední software, který stojí za zmínku je YUV Soft 3D¹⁶, což je nástroj, který nabízí automatickou i poloautomatickou konverzi. Využívá výše popsaná teoretická obrazová vodítka, jako je **pohyb v obraze, zaostření, blízkost objektů** a dalších. Bohužel autoři nabízí i samotnou demoverzi pouze pro firmy, takže nebylo možné software vyzkoušet. Vznosl jsem dotaz, zda je možné produkt poskytnout pro studijní účely, ale jejich oficiální emailový kontakt nefungoval, stejně jako webový formulář.

¹⁶ <http://www.yuvsoft.com/>

4 Praktické měření

4.1 Subjektivní hodnocení

4.1.1 Metodika subjektivních testů

Pro porovnání kvality jednotlivých metod z hlediska pozorovatele se nejčastěji používá metoda subjektivního hodnocení kvality (*Subjective Quality Assessment*). ITU-R vydalo pro účely televizní techniky soubor doporučení označovaný jako ITU-R BT.500-13 [38], který obsahuje několik typických přístupů k metodice testování v závislosti na tom, co a jak chceme porovnávat. Samozřejmostí jsou i dodatky, kde je kompletní popis statistického vyhodnocení včetně vzorců pro výpočet intervalů spolehlivosti a podobných údajů pro zvýšení vypovídající hodnoty subjektivních testů.

Tato specifikace bohužel nepopisuje přímo subjektivní hodnocení 3D materiálů, a proto musím sestavit metodiku na základě dříve prováděných testů, které jsou popsány v publikacích u metod převodu 2D/3D.

V článku [39] bylo použito objektivní i subjektivní hodnocení a také objektivní hodnocení několika sekvencí konvertovaných různými metodami. Použito bylo 10 video sekvencí, kde vždy klíčové framy byly posunuty o 20 – 30 snímků. Tyto ukázky obsahovaly různé obrazové vlastnosti, jako jsou ostré hrany, velké posuny objektů, oblasti bez textur apod.

Pro objektivní měření se vypočítala MSE (mean square error - střední kvadratická chyba), kde se porovnávaly jednotlivé metody s referencí a také objektivní metrika BIQI (*blind image quality index* - index slepého snímku) [40], který hodnotí hloubkové mapy na stupnici 0 – 100.

Subjektivně metody porovnávalo 16 pozorovatelů na 3DTV s aktivními přepínanými brýlemi. Referenční sekvence byly promíchány s ostatními zkonvertovanými sekvencemi. Ve výsledné tabulce je pouze zmíněno, kolikrát byla která metoda zvolena v prvním výběru a kolikrát v druhém, z čehož vyplývá, že sekvence nejspíš pouze porovnávali v párech (článek bohužel přesnější informace neuvádí).

Článek [41] popisuje další možnou metodiku subjektivního testování poloautomatických metod pro převod 2D na 3D. K dispozici má 5 video sekvencí, které jsou ve třech skupinách.

V jedné skupině jsou originální sekvence použité jako reference. Ve druhé a třetí skupině jsou stejné sekvence vytvořené pomocí rozdílných algoritmů. Jako vstupní 2D video je pro každou metodu použitý vždy levý kanál z referenční skupiny. V článku je rovněž popsáno, že pro vytvoření druhé a třetí skupiny snímků byly využity další skupiny dobrovolníků a hodnotil se čas, za který byli schopni konvertovat sekvenci, ale pro naše účely není tato část nezbytná.

Patnáct pozorovatelů mělo za úkol porovnat prostorový vjem u původních referenčních sekvencí spolu s nově vytvořenými videi. Subjektivní testování bylo prováděno na aktivní stereoskopické 3DTV v temné místnosti a hodnotící seděli 2 metry od TV panelu.

Porovnávání reference s novými metodami proběhlo celkově třikrát (dvakrát pro nové metody a jednou při automatické konverzi pomocí 3DTV). Pozorovatelé hodnotili konvertované video na stupnici 1 – 5, kde 5 znamená, že sekvence byly nerozeznatelné, a 1, že rozdíly byly veliké.

V další publikaci [37] je popsán opět subjektivní test kvality představeného algoritmu. Důležité je zmínit, že pozorovatelé napřed viděli sadu tréninkových sekvencí, aby měli možnost vyzkoušet si kvalitní stereoskopický vjem. Dále potom hodnotili prostorový vjem u testovaných videí, kterých bylo šest ve třech skupinách, podle tří rozdílných postupů vytváření videí. Sekvence byly hodnoceny na stupnici 1 – 10, kde 10 byla nejlepší známka pro silný prostorový zážitek a 1 pro nulový. Podle výsledných tabulek vychází, že pozorovatelů bylo 6.

Testování z publikace [18] bylo prováděno na základě metodiky z doporučení ITU-R BT.500, konkrétně lehce upraveného postupu *single-stimulus presentation*. Porovnávány byly tři algoritmy pro konverzi 2D obrazu v reálném čase. K dispozici bylo 6 video sekvencí, které byly promítány na 3D displeji s aktivními brýlemi. Pozorovatelů, kteří se zúčastnili testování, bylo 20. Videi jim byla promítána v náhodném pořadí a měli za úkol každé ohodnotit na stupnici 1 – 5 (5 byla nejlepší známka, 1 nejhorší), podle toho, jak jim připadala kvalitní. Výsledná čísla byla vztažena do stupnice 1 – 100 s pěti segmenty. Zajímavý fakt, který přinesly tyto testy bylo zjištění, že při obrácené hloubkové mapě je lidský mozek stále schopný vnímat správnou stereoskopickou scénu.

Při hodnocení nové metody konverze v titulu [36] bylo opět využito subjektivní hodnocení podle směrnice ITU-R. Testování probíhalo na 22" monitoru s NVIDIA GeForce 3D vision s aktivními brýlemi. 18 pozorovatelů nejprve vidělo 7 sekvencí v různém rozlišení a nakonec ohodnotili 2 sekvence ve třech různých podobách (originální stereoskopický snímek a dvě

metody, které uměle vytvořily 3D obraz z levého kanálu originálu). Opět byla použita stupnice 1 – 10, kde 10 bylo nejvyšší hodnocení.

Krátká zmínka o subjektivním hodnocení se nachází také v článku [42]. Zde 10 pozorovatelů hodnotilo jak 3D efekt, tak také kvalitu obrazu, což znamená, že neobsahuje šum či zkreslení. Oba parametry pozorovatel známkoval na stupnici 0 – 10, kde 10 znamená nejlepší a 0 nejhorší. Videá byla rozdělena do tří skupin: 2D, 3D referenční a 3D konvertované z 2D.

4.1.2 Postup měření

Ve svém testování subjektivní kvality metod pro konverzi 2D do 3D jsem vycházel z předpokladů a postřehů, které jsem získal při psaní předchozí podkapitoly.

4.1.2.1 Sada testovacích sekvencí

Jako testovací sekvence jsem zvolil vzorek z volně dostupná databáze¹⁷ z francouzské univerzity v Nantes, která obsahuje 10 sekvencí v rozlišení 1920 × 1080 pixelů a levý i pravý kanál jsou ve dvou souborech *avi*. Délka pro 9 sekvencí je 16 vteřin a 10. má pouze 11 vteřin. Tato databáze původně sloužila k testování různé kvality komprimování videa. Z toho bohužel vyplývá, že má každý obrazový kanál velikost 1,5 GB, a tedy v případě přehrávání obou kanálů zároveň je datový tok vyšší, než je maximální rychlost, kterou může číst běžný plotnový pevný disk. Protože jen SSD disk dokázal zajistit plynulé přehrávání, musel jsem videa komprimovat do *mp4* s kodekem *h.264*, abych datový tok snížil. Videá bylo možné rozdělit podle obsahu, zda jsou venkovní (3), nebo vnitřní (7) a také jestli se v sekvenci pohybuje kamera (1), či nikoliv (9). Z těchto 10 sekvencí jsem pro své účely vybral 6, kde byly 2 venkovní a 4 z interiéru. Scéna s pohybem kamery (č. 2 - *Basketbal*) byla mezi vybranými sekvencemi z interiéru. V kapitole 7.3 jsou screenshoty ze všech použitých sekvencí.

Z 6 referenčních snímků jsem vzal vždy levý kanál a použil ho pro vytvoření nového kanálu pomocí tří rozdílných metod. Konkrétně jsem zvolil program DVDFab 3D Converter, který je přiblížen v kapitole 3.3.3.1, dále svou vlastní metodu, která je popsána v kapitole 3.2, a nakonec také program TriDef 3D z kapitoly 3.3.2.1. Protože TriDef 3D konvertuje v reálném čase a ve vlastním přehrávači, musel jsem zaznamenat výstup do souboru, aby byla i tato metoda vhodná pro přehrávání později za stejných podmínek jako ostatní videa. K tomu jsem použil

¹⁷ ftp://ftp.ivc.polytech.univ-nantes.fr/NAMA3DS1_COSPAD1/Avi_videos/HRC_00_Reference/

záznamovou kartu Blackmagic Intensity Pro, díky které byl nahrán kompletní výstup grafické karty. Zde opět nastal problém s velkým datovým tokem, a bylo tedy nutné výsledky také převést do *mp4* s kodekem *h.264*. Tyto tři metody jsem do testu zařadil, abych zjistil, zda běžně dostupné komerční metody mohou divákům zajistit srovnatelný zážitek jako při sledování originálního stereoskopického videa.

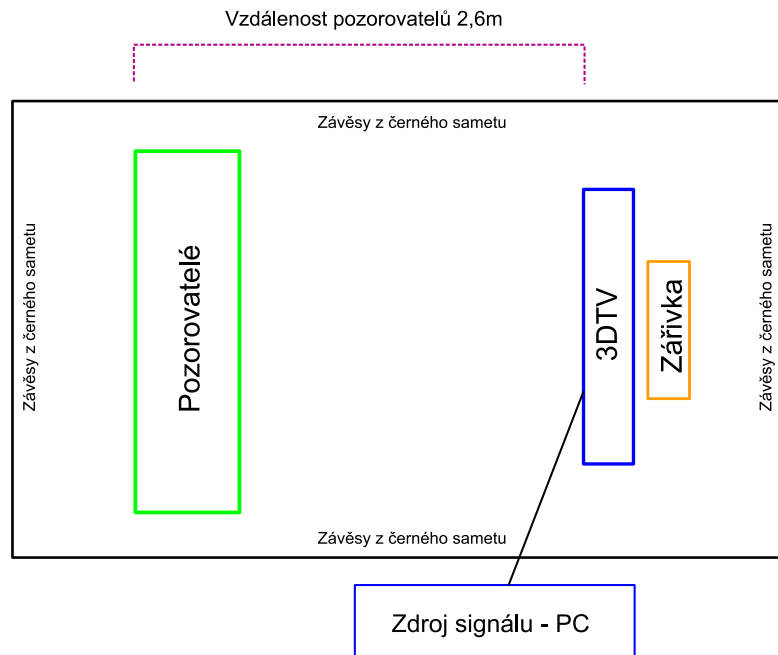
V původním plánu bylo zařadit také konverzi Blu-Ray přehrávače, který byl na pracovišti k dispozici a spadá do kategorie domácího promítání. Abych měl sekvence zkonvertované a připravené pro další přehrávání, musel jsem výstup přehrávače uložit stejně jako TriDef 3D. Naneštěstí karta Blackmagic Intensity Pro nepodporovala výstupní kombinaci rozlišení a snímkové frekvence, a proto nemohl být výstup uložen. Přestože jsem zkoušel různé nastavení rozlišení výstupu a obnovovací frekvenci, nezobrazil se kýžený výstup správně.

Rozlišení sekvencí bylo 1920×1080 pixelů v SBS režimu, čili každý kanál byl s poloviční šířkou a 3DTV ho sama interpolovala do správného FullHD rozlišení. Na přiloženém DVD je kompletní sada sekvencí, které jsem v testu použil.

4.1.2.2 Průběh testu

Když jsem tedy měl k dispozici celkem 24 sekvencí (6 originálních a 18 uměle vytvořených), mohl jsem spustit test. V publikacích zabývajících se subjektivním testováním jsem často četl, že hodnotitelé byli za svůj přínos do projektu finančně ohodnoceni, což jsem si nemohl dovolit. Hodnocení se nakonec zúčastnilo 20 dobrovolníků (17 mužů a 3 ženy), převážně studentů, kteří ve svém volném čase souhlasili s hodnocením videí. Věk pozorovatelů byl mezi 20 – 27 lety.

Pro sledování jsem vybral pasivní 3DTV značky LG typ 47LW650S. Uspořádání bylo velmi podobné tomu, které jsem použil pro testování 3D televzí v [4]. Pozorovatelé seděli ve vzdálenosti 2,6 metru od obrazovky za stolem, na kterém byl pro ně připraven dotazník. Okolí bylo zatemněno a za TV byla umístěna regulovatelná zářivka. Zdrojem signálu pro TV byl stolní počítač umístěný mimo zatemněnou oblast. Schéma na *obr. 27* ilustruje toto uspořádání.



Obr. 27 Schéma uspořádání pracoviště pro subjektivní testy.

Metodika mého testování byla postavena na základě dříve popsanych testů v kapitole 4.1.2. Vzal jsem tedy původních 6 sekvencí i 18 uměle zkonvertovaných do 3D a náhodně je promíchal (všechny sekvence obsahují náhledy v příloze 8.3).

Všichni diváci pak hodnotili každou sekvenci zvlášť na stupnici 1 – 5 pro tři zvolené parametry. Tyto parametry byly:

Vnímaná hloubka - pozorovatel měl ohodnotit, jak moc scéna vystupuje nebo zalézá do obrazovky a jak intenzivní je 3D vjem.

(5 - velmi intenzivní, 4 - intenzivní, 3 - znatelný, 2 - sotva znatelný, 1 - nepostřehnutelný)

Artefakty v obraze - pozorovatel posuzoval, zda obraz vykazuje určité nežádoucí jevy (například *ghosting*) či některé části scény nepříjemně působí na jeho oči.

(5 - perfektní, 4 - dobrý, 3 - ucházející, 2 - slabý, 1 - špatný)

Celkový dojem - nakonec pozorovatel hodnotil celkový dojem ze zhlédnuté sekvence.

(5 - perfektní, 4 - dobrý, 3 - ucházející, 2 - slabý, 1 - špatný)

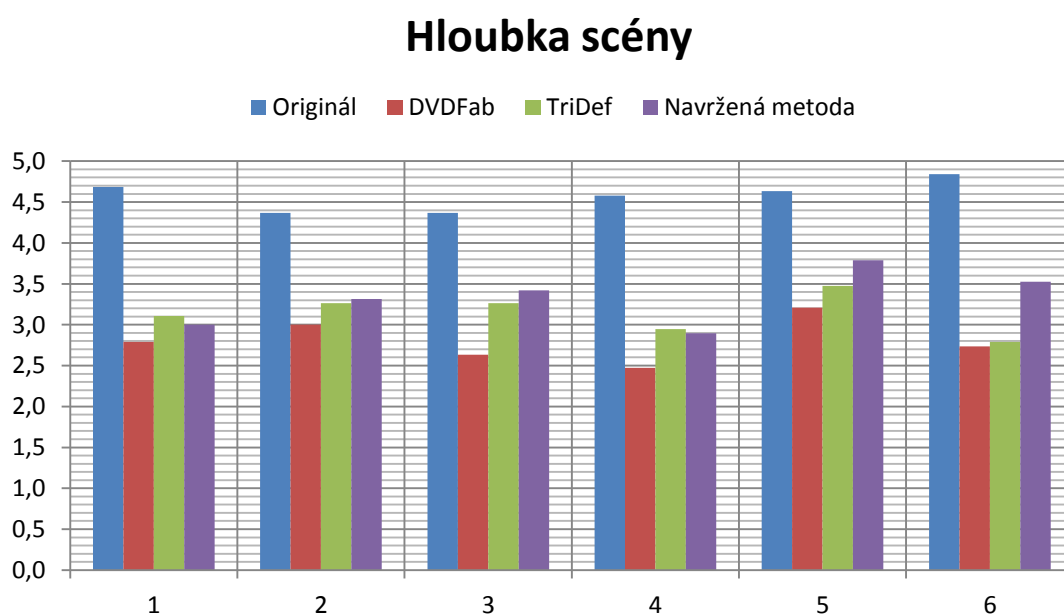
Na začátku bylo také zařazeno několik sekvencí jako zkušební, aby pozorovatelé získali představu o tom, co je čeká. Mezi každou sekvencí byla zařazena obrazovka s neutrální šedou

barvou, jenž trvala 20 sekund a při které diváci mohli ohodnotit předchozí testovanou scénu. Doba testu tedy byla okolo 20 minut, včetně vysvětlení a popsání, na co se má pozorovatel soustředit.

Na závěr této části bych rád zmínil, že nejobtížnějším úkolem bylo logistické plánování průběhu měření. Jak jsem psal výše, testu se účastnili dobrovolníci a musel jsem tedy plánovat hodnocení podle jejich časových možností. To však bylo ztíženo tím, že učebna, kde jsem testy prováděl, nebyla vždy volná. Proto často hodnotili studenti sami, maximálně ve dvou, i když testovací soustava umožňovala až 4 hodnotitele současně.

4.1.3 Výsledky

V testech, ze kterých vycházela má měření, se vždy jako výsledek porovnávala střední hodnota pro jednotlivé sekvence. Z toho jsem vycházel i já. Pozorovatelů bylo 20, nicméně jeden měl výsledky natolik odlišné, že nemohly být zahrnuty do konečných výsledků a dále jsem pracoval pouze s výsledky od 19 hodnotitelů, a ty jsou přiloženy na DVD.

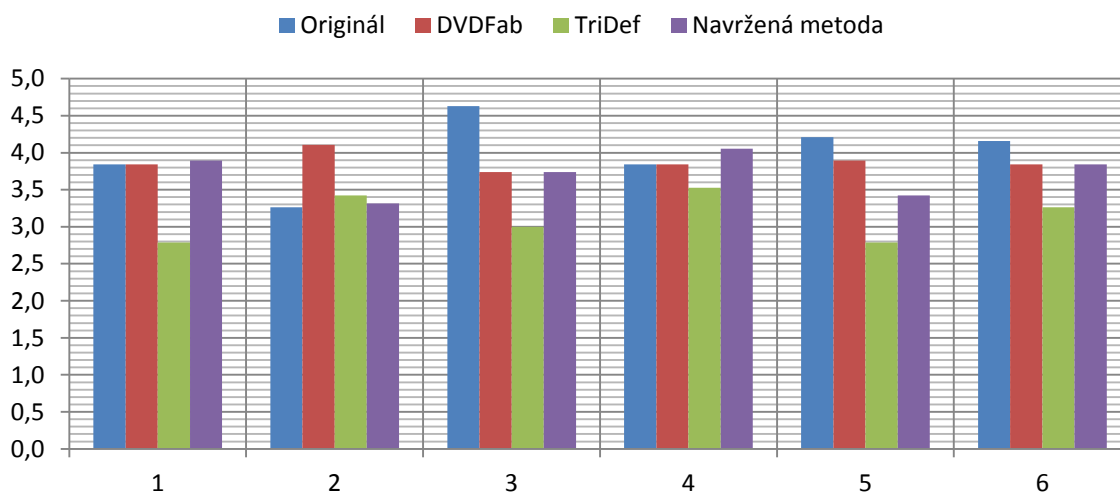


Obr. 28 Graf výsledků pro Hloubku scény.

Nejprve jsem určil střední hodnoty pro všechny sekvence v rámci každého parametru. Obr. 28 je graf výsledků pro **Hloubku scény**, kde vidíme skóre každé sekvence v závislosti na metodě. Originál dosahuje podle předpokladů nejlepšího skóre a na druhém místě je

ve většině případů navržený metoda. Naopak DVDFab v tomto parametru příliš neuspěl a skončil nejhůře u všech sekvencí.

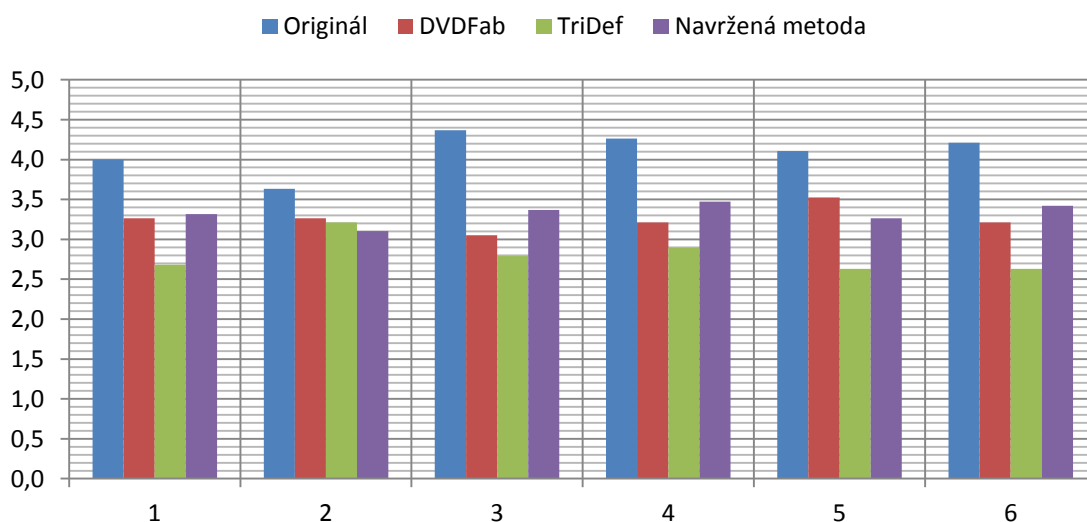
Artefakty



Obr. 29 Graf výsledků pro Artefakty.

Graf pro **Artefakty** je na obr. 29. Zde opět celkově vede originál, ovšem u sekvence *Basketbal* je DVDFab nejlepší. Tato ukázka byla jediná, která byla pořízena pohyblivou kamerou, a proto má pohybovou paralaxu. Z toho můžeme usuzovat, že DVDFab obsahuje určité vyhlazování, které paralaxu využívá (oproti např. navržené metodě).

Celkový dojem



Obr. 30 Graf výsledků pro Celkový dojem.

Z grafu **Celkového dojmu**, který je na *obr. 30*, opět pozorujeme, že originální videa dosahují nejlepších výsledků a navržená metoda je těsně druhá. Znovu si můžeme všimnout, že *Basketbal* sekvence má jiný charakter výsledků, než ostatní.

Dále se budu věnovat celkovému ohodnocení metod, které jsem stanovil jako střední hodnoty z výsledků jednotlivých sekvencí. Pro tyto hodnoty jsem následně určil intervaly spolehlivosti. Nakonec jsem pro každou metodu stanovil *hlavní skóre*, což byla střední hodnotou přes všechny tři parametry.

Podle vzorce (5) jsou spočítány střední hodnoty \bar{u}_{jkm} přes všech 19 pozorovatelů. *Tabulka 2* obsahuje výsledky pro jednotlivé metody včetně intervalů spolehlivosti.

$$\bar{u}_{jkm} = \frac{1}{N} \sum_{i=1}^N u_{ijkm} \quad (5)$$

u_{ijkm} je hodnota, kterou vyplnil pozorovatel i pro daný parametr j u sekvence k , pro metodu m . Výsledky včetně intervalů spolehlivosti jsou ve tvaru $\bar{u}_{jkm} \pm \delta_{jkm}$. δ_{jkm} spočítáme pomocí vzorce (6).

$$\delta_{jkm} = 1,96 \frac{S_{jkm}}{\sqrt{N}} \quad (6)$$

kde S_{jkm} je směrodatná odchylka – vzorec (7).

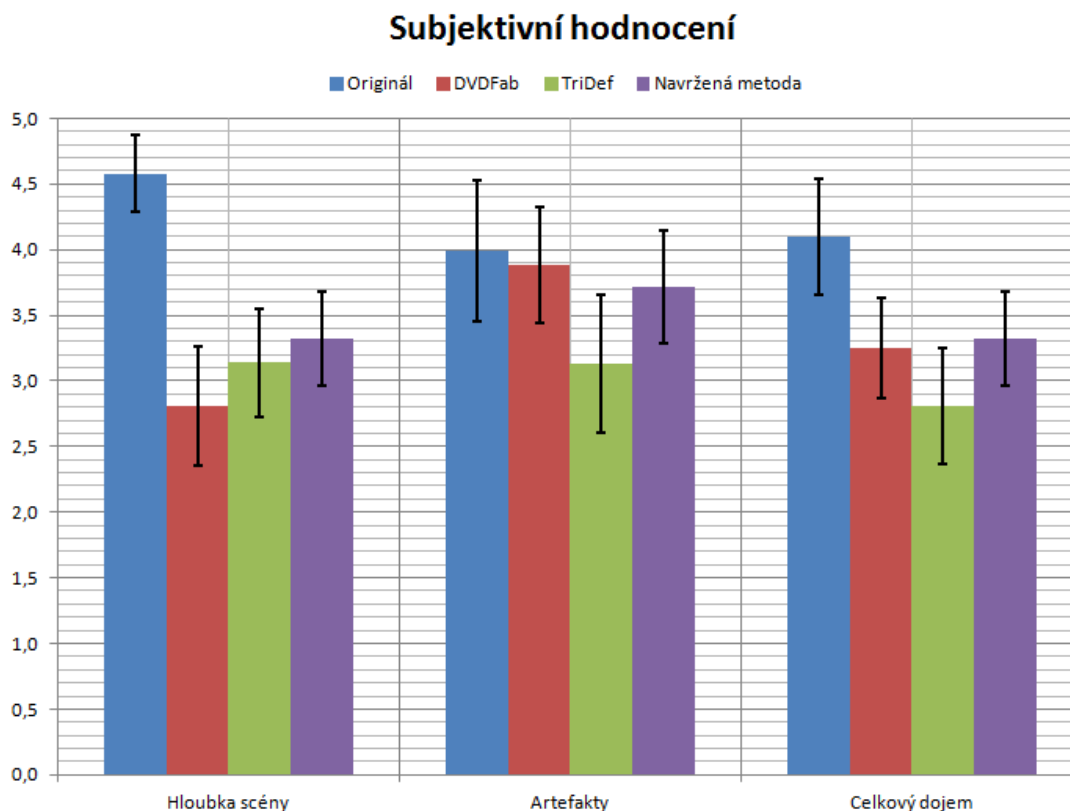
$$S_{jkm} = \sqrt{\sum_{i=1}^N \frac{(\bar{u}_{jkm} - u_{ijkm})^2}{(N-1)}} \quad (7)$$

Tab. 2 Výsledky subjektivních testů.

	Hloubka scény	Artefakty	Celkový dojem	Hlavní skóre
Originál	4,58 ± 0,29	3,99 ± 0,54	4,10 ± 0,44	4,22
DVDFab	2,81 ± 0,45	3,88 ± 0,44	3,25 ± 0,38	3,31
TriDef	3,14 ± 0,41	3,13 ± 0,52	2,81 ± 0,44	3,03
Navržená metoda	3,32 ± 0,36	3,71 ± 0,43	3,32 ± 0,36	3,45

Z tabulky na první pohled vidíme, že ukazuje očekávané výsledky, které jsem již zmiňoval u každého parametru zvlášť. Tedy originální snímky jsou z hlediska **hloubky** i **celkového dojmu** jednoznačně nejkvalitnější, zaostávají však u artefaktů, což je pravděpodobně způsobeno tím, že ne každý je zvyklý sledovat takto intenzivní stereoskopické materiály. Navíc má na tento

parametr vliv i samotná 3DTV. Ve **vnímané hloubce, celkovém dojmu**, je na druhém místě překvapivě má navržená metoda, která ovšem u sekvence č. 2 nedosahovala dobrých výsledků. Poté se umístil DVDFab Converter, který sice nevykazoval tak výrazný **3D vjem**, ale pro diváky byl příjemnější než poslední TriDef 3D. Obr. 31 graficky znázorňuje výsledky a černou úsečkou jsou vyznačeny 95% intervaly spolehlivosti.



Obr. 31 Graf celkových výsledků.

4.1.4 Statistické vyhodnocení

Abych mohl rozhodnout, zda jsou výsledky statisticky vypovídající a na jejich základě zvolit, která metoda je nejlepší, musím použít statistické vyhodnocení. Pro tento účel se velmi dobře hodí **analýza rozptylu jednoduchého třídění** (one-way ANOVA) [43]. Tato analýza je schopná určit, zda na soubor náhodných veličin, které se mění podle určitých znaků, mají statisticky významný vliv ony znaky, či nikoliv [44].

Proto budu testovat nulové hypotézy, že *střední hodnoty 6 sekvencí se pro každý měřený parametr rovnají*, což znamená, že mezi metodami použitými v subjektivním testu nejsou statisticky významné rozdíly. Měření provedu pro hodnoty každého parametru zvlášť **hloubka**

scény, artefakty a celkový dojem. Test je možný provést pomocí Excelu nebo jiného kalkulátoru. Já jsem použil volně přístupný webový kalkulátor¹⁸. Více informací [45].

Pro každý parametr, který diváci hodnotily, vezmu 6 výsledků (od každé sekvence výsledná střední hodnota) a rozdělím je do tří skupin podle metody, která byla použita pro konverzi. *Tabulka 3* je příklad hodnot pro stanovení testu u parametru **Hloubka scény**.

Tab. 3 Vstupní hodnoty do analýzy rozptylu.

číslo sekvence	DVDFab	TriDef	Navržená metoda
1	2,79	3,11	3,00
2	3,00	3,26	3,32
3	2,63	3,26	3,42
4	2,47	2,95	2,89
5	3,21	3,47	3,79
6	2,74	2,79	3,53

Pro parametr **Hloubka scény** vyšla hodnota $F = 5,138$ a při 17 stupních volnosti parametr $p = 0,02$, který následně porovnám s $\alpha = 0,05$. Pokud je $p < \alpha$, můžu odmítnout nulovou hypotézu, v opačném případě ji musím přijmout a prohlásit, že tento parametr pochází ze souboru stejných středních hodnot, a proto nemá vliv na výsledek. Zde tedy odmítnu nulovou hypotézu.

Pro parametr **Artefakty** vyšla hodnota $F = 13,92$ a při 17 stupních volnosti parametr $p = 0,00$. Z toho vyplývá, že mohu odmítnout nulovou hypotézu.

Pro parametr **Celkové hodnocení** vyšla hodnota $F = 15,79$ a při 17 stupních volnosti parametr $p = 0,00$. Z toho vyplývá, že mohu odmítnout nulovou hypotézu.

Podrobnější výsledky z protokolu z kalkulátoru ANOVA testu jsou přiloženy na DVD. Z výsledků tedy vyplývá, že všechny tři parametry jsou ovlivněny testovanými metodami a nulová hypotéza pro žádný neplatí.

¹⁸ <http://www.physics.csbsju.edu/stats/anova.html>

4.2 Objektivní měření

4.2.1 VQM metrika

Pro doplnění výsledků ze subjektivního testování jsem zařadil do hodnocení vybraných metod i objektivní metricku. Konkrétně se jedná o VQM (celým názvem *General Video Quality Model*), který byl vyvinut v ústavu Národních telekomunikací a informační administrace (NTIA) a později standardizován Institutem pro americké národní standardy (ANSI). Tuto problematiku popisují publikace [46], [47].

Toto objektivní měření se běžně používá pro hodnocení kvality 3D stereoskopického videa [48], [49], a proto ho mohu zařadit do svého hodnocení metod pro konverzi do 3D i já. V obou zdrojových publikacích se navíc píše, že výsledky, které jsou získány měřením pomocí této metody, úzce korespondují se subjektivním vnímáním člověka, a tím je tedy ideální jako doplněk k mým subjektivním testům. Metrika je vhodná pro hodnocení kvality video systému s různým rozlišením, počtem snímků za vteřinu, datovým tokem a technikou kódování [40].

VQM hodnotí zkreslení v signálové cestě mezi pořízeným videem a výsledkem zkresleným kvůli různému komprimování či jiným změnám, ke kterým dochází před zobrazením. To je obdobné i při konverzi z 2D do 3D, protože nový pravý kanál získáváme z původního levého, a porovnáváme tedy rozdíly mezi pravými kanály.

Výsledkem měření pomocí VQM je lineární kombinace sedmi parametrů, kterou popisuje rovnice (7).

$$\begin{aligned} VQM = & \\ & -0.2097si_loss + 0.5969hv_loss + 0.2483hv_gain + 0.0192chroma_spread - \\ & 2.3416si_gain + 0.0431ct_ati_gain + 0.0076chroma_extreme \end{aligned} \quad (7)$$

Parametr **si_loss** popisuje míru rozmazání, což je detekce ztráty prostorové informace v jednotlivých framech. **Hv_loss** detekuje změnu orientace hran a jejich otočení z vodorovných a svislých do úhlopříčných.

Oproti tomu **hv_gain** zjišťuje otáčení hran z diagonálního směru do horizontálního a vertikálního.

Si_gain sleduje zlepšení kvality díky zostření hran.

Údaj o rozprostření barev ve snímcích udává parametr **chroma_spread**.

Chroma_extreme popisuje velké rozdíly v barvách mezi porovnávanými snímky.

Poslední parametr je **ct_ati_gain**, který měří prostorové chyby v závislosti na pohybu, stejně jako chyby v časovém posunu snímků mezi videi.

Hodnota parametru **si_loss** je vždy menší nebo rovná 0, takže jen tato část rovnice vždy zlepšuje výslednou hodnotu VQM, naproti tomu ostatní parametry jsou vždy větší nebo rovny 0, takže jen **si_gain** může snížit VQM výsledek [41].

Získání každého parametru je trochu odlišné, ale základní princip je stejný. Nejprve se snímky musí převést z RGB barevného prostoru do $YCrCb$. Podle parametrů, které zrovna počítáme, algoritmus volí, zda bude pracovat s jasovou složkou nebo chrominančními. Pouze parametry **chroma_spread** a **chroma_extreme** využívají chrominanční komponenty C_r a C_b , ostatních pět potřebuje jen jasovou složkou. Dále se provede vstupní filtrace na jasovém (chrominančním) kanálu a ten se rozdělí do prostorově časových regionů S-T, což jsou bloky pixelů (prostorová část) jdoucí přes několik snímků (časová část). Následně se spočítá střední hodnota nebo směrodatná odchylka každého regionu a ty se poté porovnávají mezi referenčním a novým videem. Na základě předem definovaných prahů pro každý parametr a dalšími úpravami získáme finální skóre dílčích složek nutných ke stanovení celkové metriky VQM.

Na webových stránkách NTIA¹⁹ je zdarma k dispozici implementace v MATLABu, která umí VQM hodnotu pro hodnocené video stanovit. Existují rovněž i další programy, umožňující objektivní měření video sekvencí (například MSU Video Quality Measurement Tool), ale jeho plná verze je zpoplatněna.

Čím blíže je výsledek hodnotě 0, tím méně jsou mezi referencí a nově získaným videem znatelné rozdíly. Při větších rozdílech a tedy i menší kvalitou konverze dostáváme číslo vyšší.

¹⁹ <http://www.its.bldrdoc.gov/resources/video-quality-research/request-software.aspx>

4.2.2 Postup měření

Protože metrika VQM je k subjektivnímu hodnocení spíše okrajové doplnění, rozhodl jsem se ji spočítat pomocí programu MSU Video Quality Measurement Tool²⁰. V předešlém odstavci jsem psal, že plná verze je zpoplatněna, nicméně pro moje podmínky stačila demoverze. Její limitací bylo, že nešlo počítat metriku pro FullHD sekvence.

Vzhledem k tomu, že měření VQM hodnoty porovnává rozdíly mezi jednotlivým snímky referenčního a zkresleného videa, musel jsem porovnávat celé stereoskopické video v režimu SBS. Původně jsem měl v plánu porovnávat pouze pravé kanály, protože ty byly vytvořeny z levého kanálu reference, a tedy by měl být stejný, nicméně i v levém kanálu docházelo k jistým změnám (např. *pre-scale* popsány v kapitole 3.1.1).

Protože musely být oba kanály zobrazeny najednou ve FullHD rozlišení, jeden kanál měl rozměr 960×1080 pixelů. Abych je mohl použít v programu pro výpočet VQM, zmenšil jsem celé video na $1/4$, tedy do velikosti 960×540 pixelů.

Vzhledem k tomu, že tato metrika je pouhé doplnění a zmenšení proběhlo pro všechny metody včetně reference, rozhodl jsem se zanedbat možné drobné odchylky, které vznikly změnou velikosti sekvencí.

4.2.3. Výsledky

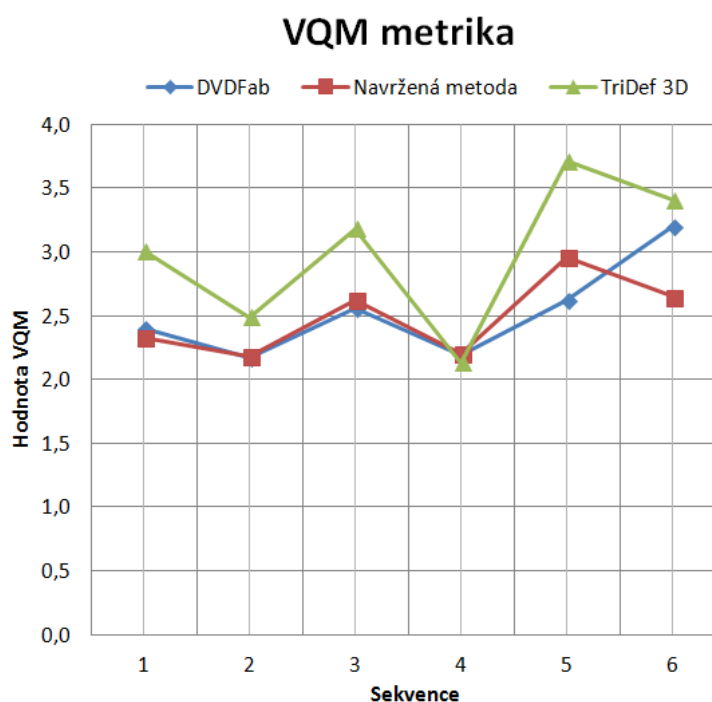
V *tabulce 4* jsou vypočítané hodnoty pro každou sekvenci zvlášť. Zároveň je z nich poté spočítán celkový průměr VQM metriky pro celou metodu. Čím nižší je výsledné číslo, tím bližší je hodnocená metoda referenci. Vidíme, že navržená metoda má srovnatelné výsledky jako DVDFab converter.

²⁰ http://compression.ru/video/quality_measure/vqmt_download_en.html#start

Tab. 4 Vypočítané hodnoty VQM metricky.

	DVDFab	Navržená metoda	TriDef 30
1	2,40	2,33	3,01
2	2,17	2,18	2,49
3	2,56	2,62	3,18
4	2,19	2,20	2,13
5	2,62	2,96	3,71
6	3,20	2,65	3,41
Průměr	2,52	2,49	2,99

V následujícím grafu (obr. 32) jsou zároveň vidět místa, kde je výsledek pro DVDFab a navrženou metodu téměř totožný. Pouze u 4. sekvence výše zmíněným metodám TriDef dokázal konkurovat, ostatních případech zaostával. To tedy potvrzuje výsledky, které jsme získali ze subjektivních testů.



Obr. 32 Graf metricky VQM pro jednotlivé metody a sekvence.

5 Závěr

V této práci jsem se pokusil vysvětlit základní principy, pomocí kterých jsme schopni rozeznat prostorovou informaci pozorované scény v případě, že nemáme stereoskopický záznam nebo binokulární pohled na scénu. Vodítka, kterým se v literatuře říká *monoscopic cues*, jsem rozdělil do dvou skupin v závislosti na pohybu kamery a doplnil je vhodnými ilustračními obrázky. Pomocí těchto vodítek můžeme získat hloubkovou mapu scény a z ní vygenerovat druhý obrazový kanál. S generováním pomocí **DIBR** jsou spojeny i další přístupy, kterými lze mapu získat.

Dále jsem rozdělil a popsal některé konkrétní metody pro převod do 3D, které lze v současnosti použít. Jejich množství je obrovské a stále přibývají další. Rozhodl jsem se popsat i některá softwarová a hardwarová řešení, která se zabývají konverzí do 3D.

V praktické části jsem nastudoval a sestavil metodiku pro subjektivní hodnocení vybraných metod konverze. Pro samotný test jsem si vybral dvě komerční řešení, protože jsem chtěl porovnat, zda je tento způsob získávání 3D obsahu konkurenceschopný originálnímu záznamu. Tyto dvě metody jsem doplnil vlastní navrženou metodou, která vycházela z několika teoretických předpokladů a také z pozorování principu, na kterém funguje konverze 3D Blu-Ray přehrávače. Testy jsem poté rozšířil i o objektivní.

Pomocí subjektivních testů a objektivní metriky VQM jsem provedl hodnocení kvality tří metod, které jsou použitelné v domácích podmínkách pro konverzi 2D videa do 3D. Pomocí dvou komerčních řešení a mnou navržené implementace na základě metody z kapitoly 3.1.1 jsem zkonvertoval šest referenčních videí. Pro tři hodnocené parametry jsem provedl i statistické vyhodnocení ANOVA, které určilo, že jsou mezi jednotlivými metodami statisticky významné rozdíly a tedy lze srovnávat metody na základě měřených parametrů - **hloubka scény, celkový dojem, artefakty**.

V subjektivním hodnocení diváci snadno poznali referenční snímky mezi ostatními a dali jim vyšší hodnocení. Zbylé tři metody vyšly z hodnocení velmi podobně, i když navržená metoda byla nakonec o něco lepší hlavně ve vnímané hloubce scény a celkovém dojmu. Naproti tomu program TriDef nebyl hodnocen dobře a nakonec skončil nejhůře, což potvrdily i objektivní testy. Pravděpodobně to je způsobeno tím, že je obsah konvertován v reálném čase. Obraz nemohl být vyhlazeno na konci konverze, jak tomu je u poloautomatických metod. Proto

diváci této metodě přiřadili nejslabší hodnocení z hlediska artefaktů v obraze a to se podepsalo i na celkovém hodnocení scén.

Z těchto výsledků vyplývá, že konverze dostupná pro běžné uživatele (zařízení, popřípadě software se pohybuje cenově okolo několika desítek dolarů) rozhodně nemá možnost dokonale převést původní obraz do 3D. Zároveň ovšem kýžený vnímaný 3D efekt zajistily, přestože ve slabší kvalitě než originál. Někteří dobrovolníci mi po testu sdělili, že jako dobré 3D berou hlavně obraz, který vystupuje dopředu. Všechny tři testované metody měly pozitivní paralaxu, což mohlo vést k tomu, že obdržely nižší hodnocení než reference.

Dále musím zmínit fakt, že všechny sekvence, které byly použity v testu krom jedné, byly zachyceny statickou kamerou, z čehož plyne, že scény neměly pohybovou paralaxu. Navržená metoda využívala principu, který byl zaměřen spíše pro scény se statickou kamerou, a je velmi pravděpodobné, že by se výsledky změnily, což potvrzuje sekvence č. 2, kde navržená metoda skončila poslední.

Metoda **Neparametrické vzorkování** nemohla být zahrnuta do subjektivních testů, protože její databáze, pomocí které generuje hloubkovou mapu, není zatím autory připravena pro venkovní snímky. Také je velmi náročná na výkon pracovní stanice, který jsem nemohl zajistit takový, abych byl schopen plně využít potenciál metody. Troufám si tvrdit podle drobného vzorku příkladů, na kterých jsem metodu vyzkoušel, že by se její výsledky výrazněji přiblížily referenci, než tomu bylo u testovaných metod.

Je škoda, že mi distributoři nezapůjčili k vyzkoušení některý z profesionálních 3D procesorů, protože by bylo velmi zajímavé porovnat, jak by si vedly proti běžným softwarovým řešením.

Možnosti rozšíření či další práce do budoucna vidím ve srovnání těchto metod s použitím jiné sady testovacích snímků. Vhodné by byly scény, kde je pohyblivá kamera, a lze tedy využít pohybové paralaxy. Dále by bylo dobré porovnat mé dosažené výsledky např. s profesionálními zařízeními či se zpracováním některé obtížnější metody.

6 Seznam zdrojů

- [1] Ozaktas, H.M., Onural, L.: Three-Dimensional Television. Capture, Transmission, Display, Springer, 2008.
- [2] Defaux, F., Pesquet-Popescu, B., Cagnazzo, M.: Emerging technologies for 3D video: creation, coding, transmission and rendering. Wiley, 2013.
- [3] Zhang, L., Vazquez, C., Knorr, C.: 3D-TV Content Creation: Automatic 2D-to-3D Video Conversion. IEEE Transactions on Broadcasting, vol. 57, no. 2, pp. 372-383, 2011.
- [4] Röhrich, D.: Technologie pro zobrazování stereoskopického obrazu. Praha, 2012, Bakalářská práce, ČVUT v Praze, Fakulta elektrotechnická, Katedra radioelektroniky.
- [5] [online] - Stereoscopic Parallax | 3D Forums, <http://www.3d-forums.com/threads/stereoscopic-parallax.4/>
- [6] Tam, W. J., Zhang, L.: 3D-TV content generation: 2D-3D conversion. IEEE International Conference on Multimedia and Expo, pp. 1869-1872, July 2006.
- [7] Mikšíček, D.: Causes of Visual Fatigue and Its Improvements in Stereoscopy. University of West Bohemia in Pilsen, Department of Computer Science and Engineering, Technical Report No. DCSE/TR-2006-04, 2006.
- [8] Röhrich, D.: Metody konverze 2D videa na stereoskopické 3D. Praha, 2013, ČVUT v Praze, Fakulta elektrotechnická, Katedra radioelektroniky.
- [9] Favaro, P.: Shape from focus/defocus. Washington University, Department of Electrical Engineering Electronic Signals and Systems Research Lab, Jun. 25, 2002.
- [10] Pentland, A. P.: A new sense for depth of field. IEEE Trans. Pattern Anal. Mach. Intell., vol. 9, pp. 523–531, 1987.
- [11] Ens, J., Lawrence, P.: An investigation of methods for determining depth from focus. PAMI, Vol.15(2), pp. 97-108, 1993.
- [12] Valencia, S. A., Rodriguez-Dagnino, R. M.: Synthesizing stereo 3D views from focus cues in monoscopic 2D images. in Proc. SPIE, 2003, vol. 5006, pp. 377–388.
- [13] Guo, G., Zhang, N., Huo, L., Gao, W.: 2D to 3D conversion based on edge defocus and segmentation. in IEEE Int. Conf. Acoust., Speech Signal Process., Mar. 31–April, 4 2008, pp. 2181–2184.
- [14] Battiato, S., Curti, S., La Cascia, M., Tortora, M., Scordato, W.: Depth map generation by image classification. in Proc. SPIE, Apr. 2004, vol. 5302, pp. 95–104.

- [15] Huang, X., Wang, L., Huang, J., Li, D., Zhang, L.: A depth extraction method based on motion and geometry for 2D to 3D conversion. in 3rd Int. Symp. Intell. Inf. Technol. Appl., 2009, pp. 294–298.
- [16] Battiato, S., Curti, S., La Cascia, Capra, M.: 3D stereoscopic image pairs by depth-map generation. Second International Symposium on 3D Data Processing, Visualization and Transmission, pp. 124-131, 2004.
- [17] Jung, Y.-J., Baik, A., Park, D.: A novel 2D-to-3D conversion technique based on relative height-depth-cue. in SPIE Conf. Stereoscopic Displays Appl. XX, San José, CA, 2009, vol. 7237, 72371U.
- [18] Tsai, S.-F., Cheng, C.-C., Li, C.-T., Chen, L.-G.: A real-time 1080p 2D-to-3D video conversion system. IEEE Trans. Consum. Electron., vol. 57, no. 2, pp. 915–922, May 2011.
- [19] Cozman, F., Krotkov, E.: Depth from scattering. in IEEE Conf. Comput. Vis. Pattern Recog. (CVPR), 1997, pp. 801–806.
- [20] Yamada, K., Suehiro, K., Nakamura, H.: Pseudo 3D image generation with simple depth models. in Int. Conf. Consum. Electron., Dig. Techn. Papers, Jan. 8–12, 2005, pp. 277–278.
- [21] Zhang, R., Tsai, P.-S., Cryer, J. E., Shah, M.: Shape-from-shading: A survey. IEEE Trans. Pattern Anal. Mach. Intell., vol. 21, no. 8, pp. 690–706, 1999.
- [22] Rubin, E.: Figure and ground. in Visual Perception Philadelphia, S. Yantis, Ed. London, U.K.: Psychology, 2001, pp. 225–229.
- [23] Tam, W. J., Yee, A. S., Ferreira, J., Tariq, S., Speranza, F.: Stereoscopic image rendering based on depth maps created from blur and edge information. in Proc. Stereoscopic Disp. Appl., Jan. 2005, vol. 5664, pp. 104–115.
- [24] [online] vision.middlebury.edu/stereo/data.
- [25] Karsch, K., Liu, C., Kang, S.B., "Depth Extraction from Video Using Non-parametric Sampling," European Conference on Computer Vision, 2012.
- [26] Rzeszutek, R., Androustos, O.: Efficient Automatic Depth Estimation for Video. Digital Signal Processing (DSP), 2013 18th International Conference on July 2013.
- [27] Liu, L., Ge, Ch., Zheng, N., Li, Q., Yao, H.: Spatio-temporal Adaptive 2D to 3D Video Conversion for 3DTV. in Proc. IEEE Int. Conf. 2012, pp. 465-466.
- [28] Karsch, K., Liu, C., Kang, S.B.: Depth Extraction from Video Using Non-parametric Sampling. Supplemental material, European Conference on Computer Vision, 2012.
- [29] [online] http://library.creativecow.net/articles/petty_grant/NAB_Blackmagic-Design-Interview/assets/teranexfrontandrear_3Dmodel.jpg

- [30] [online] <http://shop.disk.cz/pictureprovider.aspx?z=680&path=//ABRA/dokumenty/karty/obrazky/VJVV494/IF-2D3D1%5B1%5D.jpg>
- [31] [online] [http://pro.sony.com/bbsc/imageController?path=Asset%20Hierarchy\\$Professional\\$SEL-yf-generic-153714\\$SEL-yf-generic-153770SEL-asset-203556.jpg&id=StepID\\$SEL-asset-203556&dimension=600x407](http://pro.sony.com/bbsc/imageController?path=Asset%20Hierarchy$Professional$SEL-yf-generic-153714$SEL-yf-generic-153770SEL-asset-203556.jpg&id=StepID$SEL-asset-203556&dimension=600x407)
- [32] Sony Corporation MPES-2D3D12D/3D Converter Software V1.0 White Paper, Confidential information.
- [33] Zhang, Z., Wang, Y., Jiang, T., Gao, W.: Visual Pertinent 2D-to-3D Video Conversion By Multi-cue Fusion. ICIP, IEEE, 2011.
- [34] Zhao, Z., Chen, M., Yang, L., Fan, Z., Ma, L.: 2D to 3D Video Conversion based on Interframe Pixel Matching. Harbin University of Commerce, 2010 IEEE.
- [35] Wu, C., Er, G., Xie, X., Li, T., Cao, X., Dai, Q.: A novel method for semi-automatic 2-D to 3-D video conversion. in Proc. 3DTV Conf. True Vis. Capture, Transmission Display 3D Video, May 2008, pp. 65–68.
- [36] Pourazad, M., Nasiopoulos, P., Bashashati, A.: Random forests-based 2D-to-3D video conversion. in Proc. 17th IEEE Int. Conf., Dec. 2010, pp. 150–153.
- [37] Feng, Y., Ren, J., Jiang, J.: Object-based 2D-to-3D video conversion for effective stereoscopic content generation in 3D-TV applications. IEEE Trans. Broadcasting, Vol. 57, no. 2, pp. 500–509, Jun. 2011.
- [38] ITU-R.: Methodology for the subjective assessment of the quality of television pictures. Tech. Rep. BT.500-11, ITUR, 2002.
- [39] Li, Z., Cao, X., Dai, Q.: A novel method for 2D-to-3D video conversion using bi-directional motion estimation. in IEEE Conf. ASSP., Mar. 2012, pp. 1429–1432.
- [40] Moorthy, A., Bovik, A.: A two-step framework for constructing blind image quality indices. IEEE Signal Processing Letters, vol. 17, no. 5, pp. 513–516, 2010.
- [41] Zhang, Z., Zhou, Ch., Wang, Y., Gao, W.: Interactive Stereoscopic Video Conversion. IEEE Transaction on circuits and systems for video technology, vol. 6, No. 1, January 2013
- [42] Jiang, X., Lambers, M.: Synthesis of stereoscopic 3D videos by limited resources of range images. Int. Conf. on Pat. Rec., (2), pages 1220-1224, Aug. 2006.
- [43] [online] <http://cit.vfu.cz/statpotr/POTR/Teorie/Predn3/ANOVA.htm>
- [44] [online] http://cs.wikipedia.org/wiki/Anal%C3%BDza_rozptylu
- [45] [online] http://en.wikipedia.org/wiki/One-way_analysis_of_variance

- [46] Vranješ, M., Rimac-Drlje, S., Žagar, D.: Objective Video Quality Metrics. University of Osijek, Faculty of Electrical Engineering.
- [47] Pinson, M. H., Wolf, S.: A New Standardized Method for Objectively Measuring Video Quality. *IEEE Trans. on Broadcasting*, Vol. 50, No. 3, 2004, pp. 312-322.
- [48] Yasakethu, S. L. P., Hewage, C., Fernando, W., Kondozi, A.: Quality analysis for 3-D video using 2-D video quality models. *IEEE Trans. Consumer Electron.*, vol. 54, no. 4, pp. 1969–1976, Nov. 2008.
- [49] Hewage, C., Worrall, S. T., Dogan, S., Villette, S., Kondozi A.: Quality evaluation of color plus depth map-based stereoscopic video. *IEEE J. Selected Topics in Signal Process.*, vol. 3, no. 2, pp. 304-318, Apr. 2009.

7 Přílohy

7.1 Dotazník pro subjektivní hodnocení

- a) Jsem: muž , žena Věk:
- b) Jaké máte zkušenosti s 3D obrazem? Žádné , Kino IMAX , 3DTV , jiné

1) Vnímaná hloubka scény (jak moc vnímáte různou hloubku objektů ve scéně)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2) Artefakty v obraze (nepříjemné a rušivé viditelné prvky ve scéně)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3) Celkový dojem (kvalita a celkový dojem ze scény)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

7.2 Obsah DVD

Elektronická podoba práce PDF	- rohrich_dp.pdf
Elektronická podoba práce docx	- rohrich_dp.docx
Dotazník pro hodnotitele	- dotaznik.pdf
Tabulka výsledků subjektivních testů	- results.xlsx
Testovací sekvence	- /videos
Výsledky ANOVA hodnocení	- /ANOVA
M-file navržené metody	- method.m

7.3 Ukázky testovacích sekvencí

7.3.1 Automobil



Obr. 30 Sekvence 1 - Automobil.

7.3.2 Basketbal



Obr. 31 Sekvence 2 - Basketbal.

7.3.3 Dvůr



Obr. 32 Sekvence 3 - Dvůr.

7.3.4 Pracovna



Obr. 33 Sekvence 4 - Pracovna.

7.3.5 Tělocvična



Obr. 34 Sekvence 5 – Tělocvična.

7.3.6 Deštník



Obr. 35 Sekvence 6 – Deštník.

