



**ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE**

---

**Fakulta elektrotechnická  
Katedra radioelektroniky**

**Kompresie 3D videa**

**Compression of 3D video**

Diplomová práce

Studijní program: Komunikace, multimédia a elektronika  
Studijní obor: Multimediální technika

Vedoucí práce: prof. Ing. Miloš Klíma, CSc.

**Bc. Bohuslav Gruber**

---

**Praha 2016**

České vysoké učení technické v Praze  
Fakulta elektrotechnická

katedra radioelektroniky

## ZADÁNÍ DIPLOMOVÉ PRÁCE

Student: **Bohuslav Gruber**

Studijní program: Komunikace, multimédia a elektronika  
Obor: Multimediální technika

Název tématu: **Kompresa 3D videosekvencí**

Pokyny pro vypracování:

Implementujte videokodek HEVC H.265 a ověřte jeho funkčnost. Vybrané 3D videosekvence komprimujte pro zvolenou sestavu vstupních parametrů a dosaženého datového toku. Výsledné videosekvence vyhodnoťte vybranou objektivní kvalitativní metrikou. Získané výsledky porovnejte se zjednodušenými subjektivními testy.

Seznam odborné literatury:

Fernando A., Worrall S.T., 3D TV: Processing and transmission of video 3D videosequences, John Wiley, 2013

Vedoucí: prof. Ing. Miloš Klíma, CSc.

Platnost zadání: do konce zimního semestru 2016/2017



V Praze dne 30. 9. 2015

# 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 .....

.....

# Abstrakt

Hlavní příčina zvýšených nároků spojených s přenosem a skladováním 3D videí je jejich vysoký bitový tok. Proto jsou stále vyvíjeny nové metody kódování umožňující redukci dat. Tato diplomová práce se z uvedených důvodů zabývá kompresí 3D videa pomocí nejnovějšího kodéru H.265/HEVC. V teoretické části práce jsou popsány informace o vnímání prostoru, popis standardu H.265/HEVC pro kódování videa a principy kódování 3D videí. V experimentální části práce je použito nesymetrické kódování stereoskopického videa pomocí metody smíšeného rozlišení pravého snímku stereoskopického videa. Účinnost metody byla ověřena pomocí objektivní metriky SSIM s přidaným binokulárním modelem. Výsledky byly porovnány se zjednodušeným subjektivním testem kvality stereoskopického videa. Na základě porovnání získaných dat z objektivního a subjektivního testu byly vybrány parametry, při nichž použitá metoda nesymetrického kódování dokáže efektivně redukovat bitový tok a zachovat přijatelnou vizuální kvalitu stereoskopického videa.

**Klíčová slova:** komprese stereoskopického videa, H.265/HEVC, stereoskopická objektivní metrika, stereoskopická subjektivní metrika

# Abstract

The main reason of increasing requirements related to a transmission and storing of 3D videos is their high bitrate. Therefore, new techniques of video encoding, which are enable to reduce data, are being continuously developed. Based on these arguments, this master thesis is focus on a compression using the latest 3D video encoder H.265/HEVC. The theoretical part describes the information of space perception, describes the H.265/HEVC standard using for video encoding and describes the principles of 3D videos encoding. In the experimental part was used asymmetric stereoscopic video encoding by the method of miscellaneous original shot resolution of stereoscopic video. An efficiency was verified by an objective SSIM metric including binocular model. The final results were compared with a simplified subjective quality test of stereoscopic video. Based on the results gained from objective and subjective tests, the suitable parametr were chosen. Within the parametr the assymmetric encoding method can effectively reduce final bitrate and still can maintain acceptable visual quality of stereoscopic video.

**Key words:** compression of stereoscopic video, H.265/HEVC, stereoscopic objective quality assessment, stereoscopic subjective quality assessment

## Obsah

1. Úvod .....	7
2. Vnímání prostoru .....	8
2.1 Binokulární disparita .....	8
2.2 Konvergence.....	8
2.3 Akomodace .....	8
2.4 Binokulární paralaxa.....	9
2.5 Další monokulární vodítka .....	9
2.6 Accomodation convergence rivalry.....	10
3. H.265/HEVC.....	12
3.1 Rozdělení obrazu.....	13
3.1.1 Stromová struktura .....	13
3.1.2 Řezy a dlaždice .....	14
3.2 Vnitrosnímková predikce .....	15
3.3 Mezišímková predikce .....	15
3.4 Transformace .....	16
3.5 Kvantování.....	16
3.6 Rekonstrukční filtry .....	16
3.7 Entropické kódování.....	16
4. Komprese vícepohledového a stereoskopického videa.....	18
4.1 MPEG-2 profil multiview.....	18
4.2 Multiview video coding.....	19
4.2.1 Profily a úrovně .....	19
4.3 MVC kódování s využitím hloubkové mapy .....	20
4.3.1 3D Advanced Video Coding.....	20
4.3.2 Multiview-High Efficiency Video Coding .....	21
4.3.3 3D High Efficiency Video Coding .....	21
5. Nesymetrické kódování.....	23
6. Kompatibilní formát pro distribuci 3D videa.....	24
7. Objektivní metody hodnocení videa .....	25
7.1 2D objektivní metriky.....	25
7.1.1 Peak Signal to Noise Ratio.....	25
7.1.2 Structural Similarity Index.....	25
7.2 Stereoskopické objektivní metriky.....	26
7.2.1 Objektivní metody hodnocení stereoskopického videa.....	26
8. Subjektivní metody hodnocení videa.....	26

8.1	Subjektivní metoda hodnocení stereoskopického videa .....	27
8.1.1	Popis databáze .....	27
8.1.2	Test zraku .....	28
9.	Experimentální část.....	29
9.1	Funkčnost kodéru HEVC.....	29
9.2	Objektivní test kvality stereoskopického videa.....	29
9.3	Subjektivní test kvality stereoskopického videa .....	29
9.3.1	Testovací sekvence.....	30
9.3.2	Popis testování .....	30
9.3.3	Účastníci subjektivního testu .....	31
10.	Výsledky a diskuze.....	33
10.1	Výsledky testu funkčnosti kodéru HEVC .....	33
10.2	Výsledky subjektivního testu kvality stereoskopického videa .....	34
10.2.1	Vliv dominance oka na výsledky subjektivního testu.....	39
10.3	Výsledky objektivního testu stereoskopického videa .....	41
10.4	Porovnání výsledků subjektivního a objektivního testu .....	43
10.5	Úspora bitového toku .....	49
11.	Závěr.....	53
12.	Použitá literatura.....	54
13.	Příloha – obsah CD .....	57

## 1. Úvod

V posledních letech nastal veliký zájem o 3D videa, který byl vyvolán příchodem filmu Avatar do kin. Prostorový vjem, jenž se divákovi dostává, umocňuje zážitek z filmů. Dnes již navíc existuje možnost vybrat si z celé řady televizí, které přehrávání 3D videa podporují. Avšak nejde jen o zdroj zábavy. 3D zobrazování nachází své uplatnění a stále častější využití také ve strojírenství či lékařství. 3D videa můžeme rozdělit na stereoskopická videa, které využívají pouze dvou pohledů, a na vícepohledová videa, která jsou složena z více než dvou pohledů.

Jelikož se stereoskopické video skládá ze dvou snímků, které se promítají na každé oko zvlášť, je nutné pracovat s dvojnásobným objemem dat. To přináší problémy jak s přenosem 3D videa, tak i s jeho ukládáním. Tato skutečnost vyžaduje efektivní kompresní metody. Protože však 3D videa vznikají snímáním stejné scény pomocí více kamer, obrazy z jednotlivých kamer jsou si značně podobné a obsahují vysokou dávku redundance. Z důvodu snížení objemu dat mohou být redundantní data z videa odebrána bez znatelného snížení kvality videa.

Za tímto účelem jsou vyvíjeny nové verze standardů dosahujících větší účinnosti při symetrickém kódování videa. Nejnovější kodér nese označení H.265/HEVC, jež by měl oproti předchozímu kodéru H.264/AVC zmenšit bitový tok na polovinu a zachovat vizuální kvalitu videa. Hlavním vylepšením kodéru H.265/HEVC je hierarchické dělení bloků. Díky tomu dokáže zpracovávat bloky až o velikosti 64x64 pixelů. Transformační bloky mohou mít velikost od 4x4 do 32x32 pixelů. Dále má kodér H.265/HEVC více možností vnitrosnímkových režimů predikce a filtrů než H.264/AVC.

Kódovací standard H.265/HEVC je možné použít při kódování jednotlivých snímků stereoskopického videa zvlášť. Kódovat všechny snímky 3D videa najednou umožňují profily standardu označené jako MV-HEVC a 3D HEVC. MV-HEVC metoda zajišťuje kompatibilitu s jednopohledovým kódováním H.265, neboť používá pro kódování videa všechny funkce z tohoto standardu. 3D HEVC je vznikající standard s rozšiřující funkce standardu H.265 o funkce pro zvýšení výkonnosti kódování.

Kromě metod symetrického kódování se dá využít princip nesymetrického kódování stereoskopického videa, které vychází z teorie binokulárního vidění. Využití nedokonalosti lidského zraku při nesymetrickém kódování umožňuje větší kompresy stereopáru.

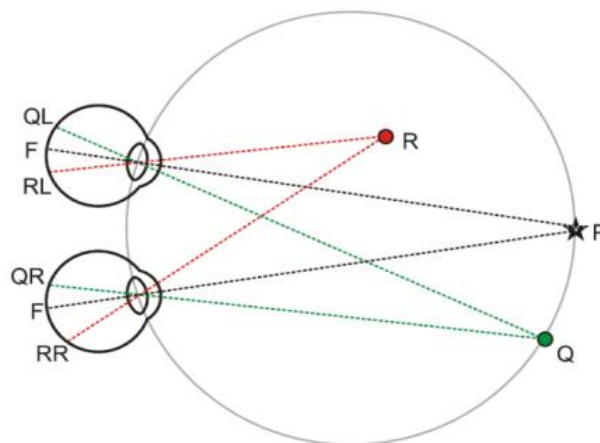
S kompresí videa souvisí hodnocení jeho kvality. Kvalita videa se může testovat za pomoci objektivních nebo subjektivních metod. Objektivní metody jsou založené na matematických modelech. Přesnost těchto modelů je však omezena. Subjektivní metody hodnocení videa jsou založeny na hodnocení kvality skupinou pozorovatelů. Oproti objektivním metodám jsou subjektivní metody časově a finančně náročnější.

## 2. Vnímání prostoru

V této kapitole budou vysvětleny základní pojmy a metody, které jsou spojené s vnímáním prostoru v reálném prostředí. Vnímání prostoru je možné hlavně díky faktu, že vidíme dvěma očima. Ty jsou u dospělého člověka od sebe vzdálené 56-75 mm, v průměru 64 mm, a tím pádem vidí každé naše oko podobné, ale horizontálně posunuté obrazy. V mozku se tyto obrazy spojí a vznikne prostorový obraz s informací o hloubce. Ale i člověk vidící jen na jedno oko dokáže hloubku vnímat. Podle toho, jestli je hloubka prostoru určena za pomoci obou očí nebo každého oka nezávisle na sobě, se metody vnímání dělí na binokulární vodítka a monokulární vodítka. Větší důraz na vnímání hloubky však mají binokulární vodítka, která vychází z fyziologie zraku.

### 2.1 Binokulární disparita

Binokulární disparita je hlavní metodou určování prostoru pomocí binokulárních vodítek. Princip je zobrazen na **obrázku 1**. Bod P leží v prostoru a je na něj zaměřen zrak. Takto jsou vytvořeny obrazy bodu P na sítnici obou očí F. Dále leží v prostoru bod Q. I obrazy QL a QR vznikají na sítnici obou očí. Vzdálenosti mezi obrazy na sítnici se nazývá disparita. Disparita obrazů F-QL a F-QR je stejná. Pokud mají body stejnou disparitu, znamená to, že jsou ve stejné hloubce, mluvíme tedy o nulové disparitě. V případě bodu R, který leží v jiné hloubce prostoru než bod P a Q, je disparita rozdílná a bod R je tudíž vnímán blíže[1].



**Obrázek 1: Binokulární disparita [1]**

### 2.2 Konvergence

Další schopnost oka, která se řadí do binokulárních vodítek, patří konvergence. Konvergence je rotace obou očí směrem dovnitř, tak aby obraz dopadl na žlutou skvrnu oka. Žlutá skvrna je oblast sítnice s největší hustou čípků a tím i místem nejostřejšího vidění. Při konvergenci jsou vždy zapojeny obě oči. Pohyb očí je zajištěn kontrakcí okohybných svalů. Opakem konvergence je divergence. Divergence může nastat při pohledu na hodně vzdálený cíl, kdy jsou oční osy až rovnoběžné. Kontrakce a dilatace okohybných svalů je velice namáhavá a dochází při ní k celkové únavě očí [2].

### 2.3 Akomodace

Akomodace je úzce spojená s konvergencí, avšak řadí se do skupiny monokulárních vodítek, protože probíhá nezávisle na obou očích. Akomodace je schopnost zaostřit na předmět nacházející se v prostoru. Zaostřit je možné díky fyziologickým schopnostem čočky měnit tvar a

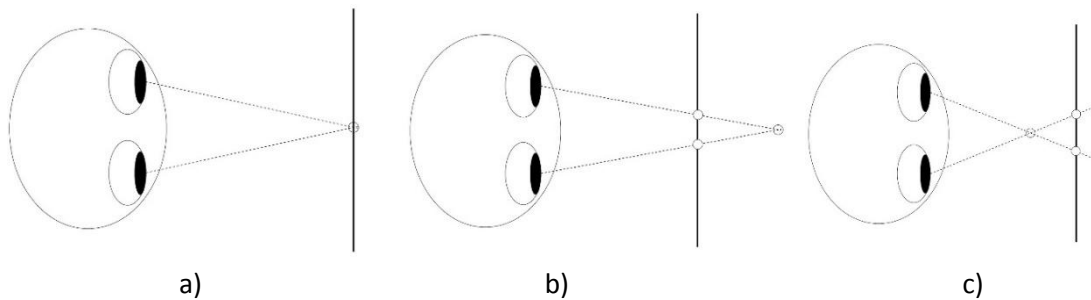


tím svou mohutnost. Při ostření na dálku se oční svaly napnou a čočka se zploští. Naopak při ostření na blízko jsou oční svaly povoleno a čočka se rozšíří [2].

## 2.4 Binokulární paralaxa

Dalším důležitým souvisejícím pojmem je binokulární paralaxa. Paralaxa a disparita jsou související jevy. Disparita je rozdíl vzdálenosti obrazů na sítnici, zatímco binokulární paralaxa je úhel, který svírají osy očí při pohledu na objekt. Paralaxu může rozdělit na tři základní typy: nulovou, pozitivní a negativní[3].

**Nulová paralaxa** - osy očí se protnou na rovině pozorování. Sledované body se promítnou přesně do této roviny pozorování, viz **obrázek 2 a)**.



**Obrázek 2: typy paralaxy a) nulová, b) pozitivní, c)negativní**

**Pozitivní paralaxa** - osy očí se protnou za rovinou pozorování. Při této paralaxe se trojrozměrný obraz odehrává v hloubce displeje, viz **Obrázek 2 b)**. Mezníkem pozitivní paralaxy je stav, kdy jsou osy levého a pravého oka rovnoběžné.

**Negativní paralaxa** - osy očí se protnou ještě před rovinou pozorování. V tomto případě trojrozměrný obraz vystupuje před pozorovací rovinou, viz **Obrázek 2 c)**. Negativní paralaxa bývá ve 3D filmech velice efektivní, ale je pro oči velice namáhá.

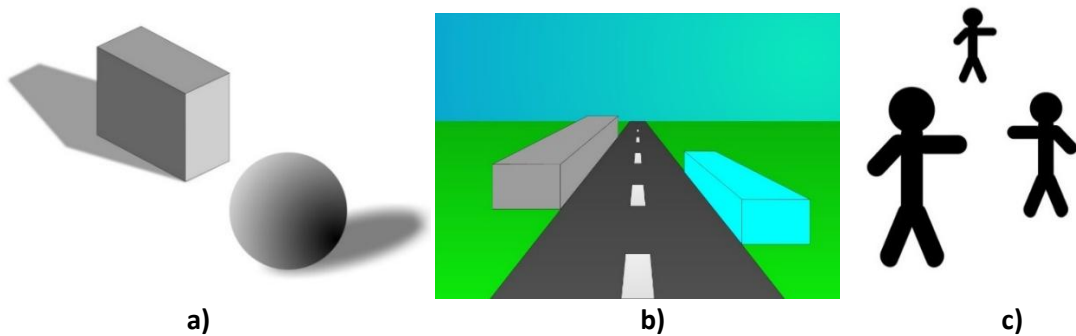
## 2.5 Další monokulární vodítka

Jak již bylo řečeno výše, monokulární vodítka slouží k rozpoznání hloubky i za pomoci jen jednoho oka. Patří do nich akomodace (viz **kapitola 2.3 Akomodace**). Dále do monokulárních vodítek patří například: světlo a stín, perspektiva, relativní velikost (**obrázek 3**). Další monokulární vodítka můžete nalézt zde [3,4].

**Světlo a stín** – podle tvaru stínu je možné rozeznat tvar objektu a jeho orientaci. Jas objektu poskytuje informaci o jeho vzdálenosti od pozorovatele. Jasnější objekty se zdají být k pozorovateli blíže než ty tmavší.

**Perspektiva** – jedná se o vztah mezi popředím a pozadím. Bližší objekty se jeví větší, než objekty vzdálenější, a to způsobuje sbíhání linií. Perspektiva také způsobuje zkracování linií, tzn. při poskládání objektů za sebe do jedné řady, mají větší rozestupy bližší objekty než objekty, které jsou dál. Počítačem generované obrazy se silnou perspektivu podávají větší hloubkový efekt a tak je snadněji dosaženo trojrozměrného vjemu.

**Relativní velikost** – pozorovatel vnímá objekty jako větší, pokud jsou k němu blíže a jako menší, když jsou od něj vzdálené. Tento jev napomáhá při rozhodnutí o vzdálenosti známých objektů. Vzdálenou osobu proto pozorovatel zobrazí jako hodně malou.

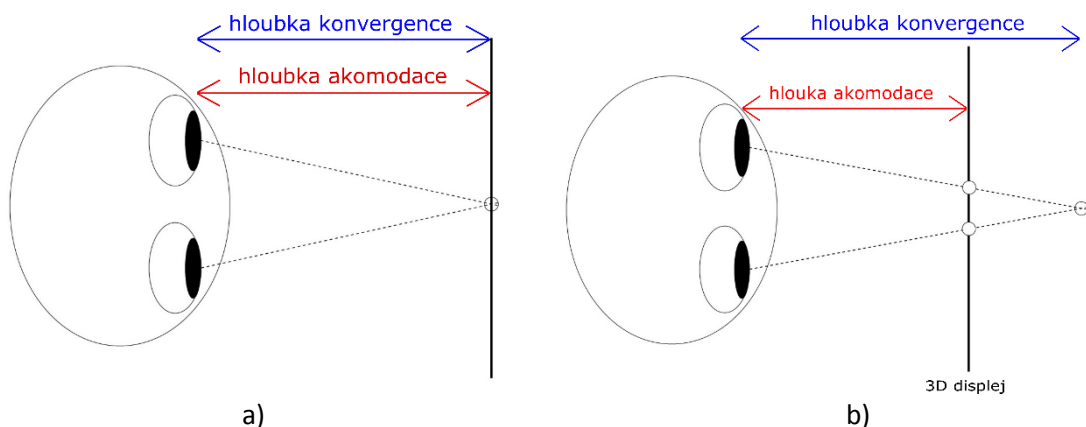


**Obrázek 3: monokulární vodítka a) světlo a stín b) perspektiva c) relativní velikost**

## 2.6 Accomodation convergence rivalry

Při vysokých hodnotách paralaxy dochází u diváků k vizuálnímu nepohodlí. Toto nepohodlí ovšem není zcela způsobeno velikostí paralaxy, ale nerovnoměrným využitím akomodace a konvergence očí. Vzhledem k tomu, že tyto vlastnosti očí se mění pomocí reflexivně spojených oko-hybných svalů, změna jednoho systému automaticky vyvolá změnu v druhém systému. Při sledování přírodní scény hloubka akomodace odpovídá hloubce konvergence, viz **obrázek 4 a)** [5,6], avšak při sledování stereoskopického videa dochází k rozdílu mezi hloubkou konvergence a akomodace, viz **obrázek 4 b)**.

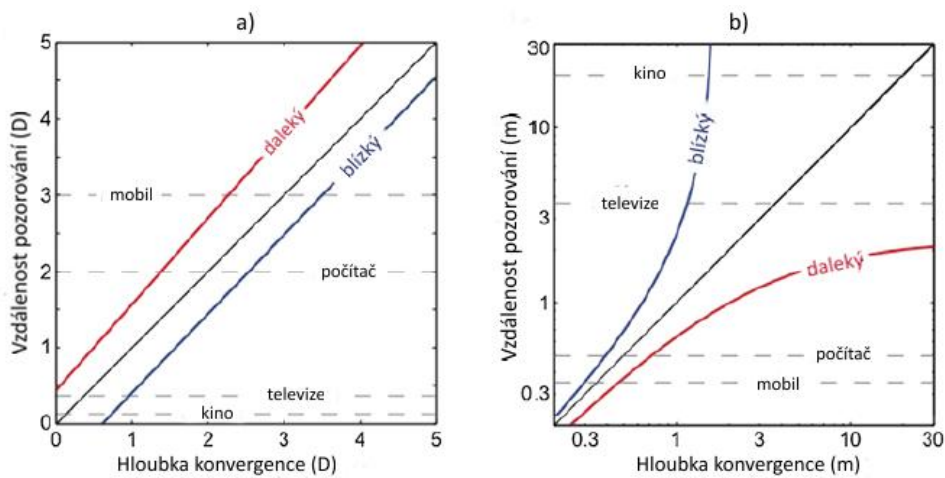
Většina 3D zobrazovačů promítá obrazy určené pro jednotlivé oči v rovině a až poté je pomocí časové či prostorové filtrace adresuje do požadovaného oka. Při sledování 3D displeje musí pozorovatel zaostřit na rovinu pozorování, tedy na displej. Vlivem paralaxy, a tedy uměle vyvolané hloubky, se oči snaží zaměřit na objekty v prostoru. Výsledkem je rozmazanost sledovaných objektů v různých hloubkách. Míra rozostření je přímo úměrná rozdílu mezi skutečnou hloubkou pozorování (vzdálenost pozorovatele a displeje) a zdánlivou hloubkou obrazu [5,6].



**Obrázek 4: Hloubka konvergence a akomodace při sledování a) přírodní scény b) 3D displeje**

V práci [7] byl proveden výzkum vlivu vzdálenosti sledování stereoskopického videa na hloubku konvergence a na základě výsledků byla určena zóna komfortu (comfort zone). Zóna komfortu představuje limit, ve kterém vzdálenost a hloubka konvergence nevyvolává vizuální nepohodlí. Zóna komfortu je graficky znázorněna na **obrázku 5**. Ukazuje se, že pozitivní paralaxa má pouze malou schopnost vyvolat nepohodlí, zatímco negativní paralaxa má velmi rušivý účinek, pokud

není pod kontrolou. Při častých výskytech hodnot paralaxy mimo komfortní zónu, může dojít k únavě očí, bolestem hlavy či nevolnosti[7,8].



**Obrázek 5: Grafické znázornění zóny komfortu**

### 3. H.265/HEVC

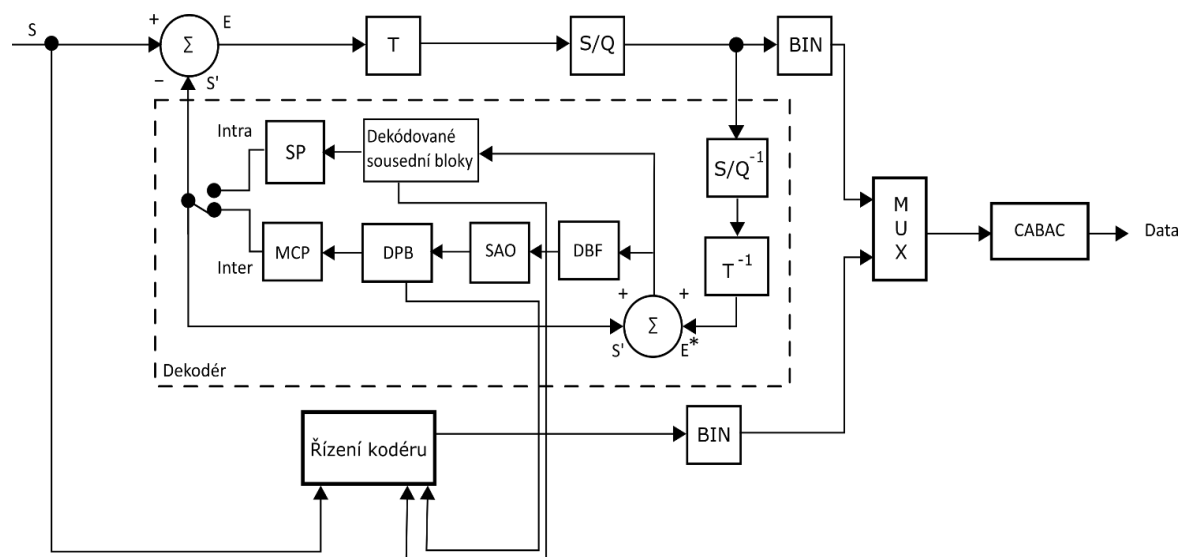
Nejnovější standard pro kódování videa, jehož první verze vyšla v roce 2013, nese název **H.265/HEVC (High Efficiency Video Coding)**. Byl vydán podobně jako jeho předchůdce standard **H.264/AVC (Advanced Video Coding)** vytvořenou skupinou **JCT-VC (Joint Collaborative Team on Video Coding)**, která vznikla ze skupin **MPEG (Moving Picture Experts Group)**, spadající do normalizační organizace ISO, a **VCEG (Video Coding Experts Group)**, která patří do normalizační společnosti ITU-T. H.265/HEVC je samostatná norma, která je ovšem schválena oběma normalizačními společnostmi a proto se jí říká H.265 podle normy ITU-T H.265 nebo HEVC podle normy ISO/IEC 23008 Part 2: High efficiency video coding [9,11].

Nový standard si klade za cíl snížit datový tok na polovinu oproti H.264/AVC při stejné kvalitě videa. Standard H.265/HEVC má široký záběr použití přes mobilní aplikace, videokonference, internetové služby, média pro ukládání až po domácí vysílání či digitální kina a to v **UHD (Ultra High Definition)** rozlišení 8k (8192x4320 pixelů). Dále standard vylepšuje paralelní zpracování. To vše však za cenu zvýšení výpočetní náročnosti [12,11].

Stejně jako jeho předchůdci používá standard H.265/HEVC hybridní kodér (vnitrosnímková nebo mezisnímková predikce a 2D transformace). Blokové schéma hybridního kodéru H.265/HEVC je zobrazeno na **obrázku 6**.

U tohoto nového standardu došlo k následujícím vylepšením:

- vysoká flexibilita rozdělování bloků,
- vysoká flexibilita v režimech predikce,
- vylepšení predikce a pohybových vektorů,
- vylepšení antiblokového filtru a přidání filtru pro potlačení přeslechů.



BIN - binarizace

CABAC - (Context Adaptive Binary Arithmetic Coding) kontextově adaptivní binární aritmetický kodér

DBF - (Deblocking Filter) protiblokový filtr

DPB - (Decoded Picture Buffer) vyrovnávací paměť již dekodovaných snímků

MCP - (Motion Compensated Prediction) pohybově kompenzovaná predikce

MUX - multiplexer

SAO - (Sample Adaptive Offset) filtr - s adaptivním offsetem

SP - (Spatial Prediction) vnitrosnímková (prostorová) predikce

E, E\* - chyba predikce, rekonstruovaná chyba predikce

S, S', S\* - vstupní snímek, predikce rekonstruovaný snímek

T - dvoudimenzionální transformace s proměnou velikosti bloku

S/Q - škálování a kvantizace

**Obrázek 6: Blokové schéma hybridního kodéru H.265/HEVC**

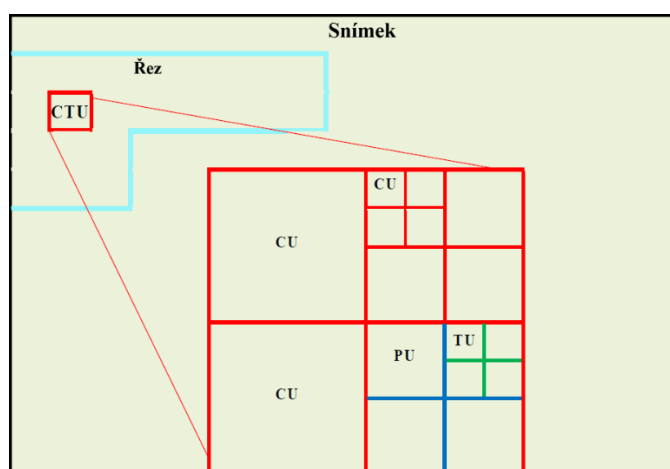
### 3.1 Rozdělení obrazu

Za účelem zpracování obrazu je ve standardu HEVC možné rozdělit obraz dvěma způsoby a to pomocí bloků vytvořených stromovou strukturou anebo pomocí řezů a dlaždic. Způsob dělení je popsán v následujících podkapitolách.

#### 3.1.1 Stromová struktura

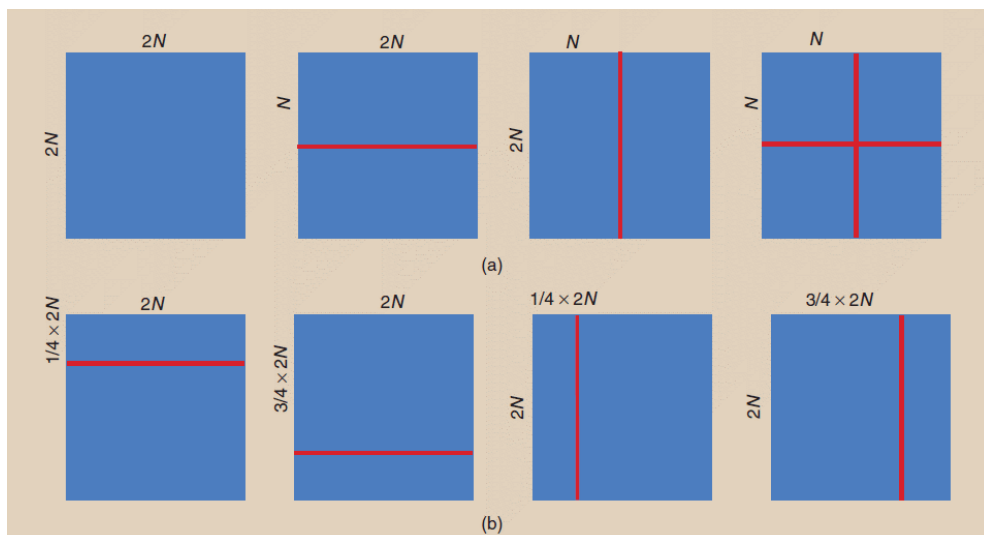
Největším rozdílem u HEVC oproti dřívějším standardům je, že se obraz nedělí na makrobloky, ale každý snímek z videa je rozdělen do stromových jednotek nazývaných **CTU (CodingTreeUnits)**, které nabízí větší flexibilitu zpracování. Velikost CTU jednotky je vybraná kodérem a může být větší než makroblok ze standardu MPEG 4 – AVC. CTU jednotka je složena z bloků **CTB (Coding Tree Block)**, které obsahují jasovou a chrominanční složku. Jasová složka má velikost  $L \times L$ , kde  $L$  může být 16, 32 nebo 64 pixelů. Barevné složky při standardním vzorkování 4:2:0 mají vždy poloviční počet pixelů než jasová složka.

Blok CTU je dále stromově dělen na menší bloky, tzv. kódovací jednotky **CU (Coding Unit)**. Velikost kódovací jednotky může být rovna velikosti celé jednotky kódovacího stromu (CTU), tedy maximálně  $64 \times 64$  pixelů, nebo se může dále rozdělit do kvadrantů CTU. Proto se tomuto dělení říká kvadrantový strom. Kódovací jednotka se dále pomocí kvadrantového stromu může dělit na další kódovací jednotky až do minimální velikosti CU, která je  $8 \times 8$  pixelů jasu. Všem CU jednotkám je přiřazena mezisnímková (inter) nebo vnitrosnímková (intra) predikce. Kódovací jednotky jsou složeny z jasové a dvou chrominančních složek, které se nazývají kódové bloky **CBs (CodingBlocks)**. Dále se kódovací jednotky dělí na predikční jednotky **PU (Prediction Unit)** a na transformační jednotky **TU (Transform Unit)**. Obě tyto jednotky se opět skládají z predikčních a transformačních bloků tedy PB a TB.



Obrázek 7: Rozdělení obrazu [10]

Jedna kódovací jednotka CU se může shodná s jednou predikční jednotkou PU nebo se může dělit na více PU, dokud PU nemají nejmenší možný rozměr  $4 \times 4$  pixelů jasu. Predikční jednotky mohou být symetrické či nesymetrické. Symetrické predikční jednotky jsou využívány jak pro mezisnímkovou tak i pro vnitrosnímkovou predikci. Avšak nesymetrické PU mohou být použity jen pro mezisnímkovou predikci. Pokud má CU velikost  $2N \times 2N$ , symetrické PU má rozměry  $2N \times N$  nebo  $N \times 2N$ . V případě nesymetrického rozdělení CU o velikost  $2N \times 2N$ , má PU jednu část o velikosti  $(N/2 \times 2N)$ ,  $(3N/2 \times 2N)$ ,  $(2N \times N/2)$ ,  $(2N \times 3N/2)$  a druhou část o velikosti  $(3N/2 \times 2N)$ ,  $(N/2 \times 2N)$ ,  $(2N \times 3N/2)$  a  $(2N \times N/2)$ , viz **obrázek 8**.



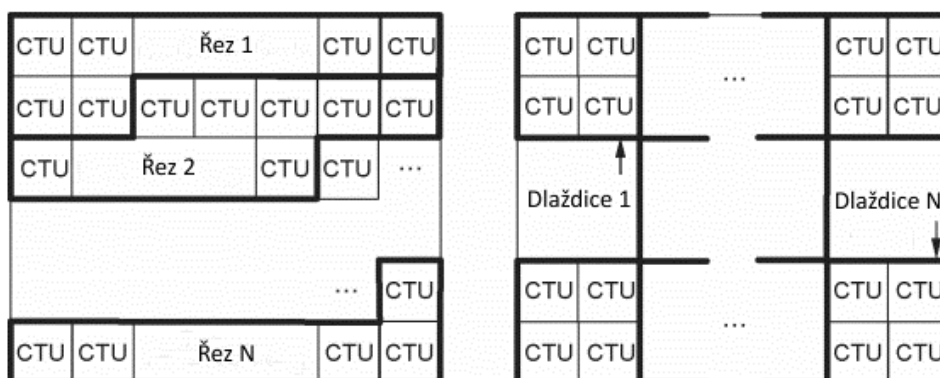
**Obrázek 8: Predikční jednotky a) symetrický a b) asymetrický [9]**

Transformační jednotka TU je základní jednotkou pro proces transformace a kvantování. Velikost jednotky TU je odvozena od jednotky CU, kdy TU může být stejně velká jako CU nebo se může dále stromově větvit. Jednotky TU jsou čtvercové o velikosti jednotek od 4x4 po 32x32 jasových pixelů[9,10].

### 3.1.2 Řezy a dlaždice

Další z možností u standardu HEVC je rozdělení obrazu na řezy (slices) či dlaždice (tiles). Při použití metody řezů je snímek možné rozdělit na jeden či více řezů, při čemž setoto rozdělení děje na úrovni jednotek CTU v rasterovém snímání. Dělení snímku na řezy má výhodu při synchronizaci dat například při streamování videa. Řezy bylo možné využít již v předchozím standardu H.264/AVC, avšak kvůli zvýšení výpočetní náročnosti a nižší kódovací účinnosti nebyla v praxi tato metoda příliš podporována.

Při druhé metodě se obraz dělí na obdélníkové oblasti, tzn. dlaždice (**obrázek 9**). Dlaždice mohou být lépe prostorově celistvé než řezy. Navíc je každá dlaždice soběstačná, tudíž je možné jejich dekódování nezávisle na sobě. Použití dlaždic zvyšuje rychlost kódování a dekódování dat pomocí paralelního zpracování [11,14].

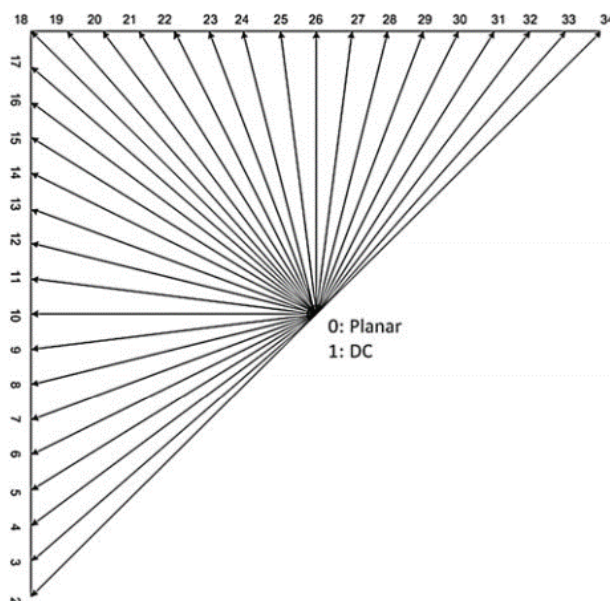


**Obrázek 9: Příklad rozdělení snímku do řezů a dlaždic [11]**

### 3.2 Vnitrosnímková predikce

Vnitrosnímková predikce, stejně jako u předcházejících standardů (MPEG-2, H.264/AVC), využívá podobnosti již dekodovaných bloků ze stejného snímku. U standardu HEVC je navýšena flexibilita, neboť standard podporuje až 36 možných režimů predikce pro jasovou složku. Možné režimy jsou znázorněny na **obrázku 10**. Všechny 36 režimů lze využít jen u PU o velikosti 8x8 až 32x32. U menších bloků 4x4 je možné použít jen režimy 0-16 a 34. Dále u bloků 64x64 lze využít pouze režimy 0-2 a 34.

I pro chromatické složky je u HEVC možnost více režimů, a to DM (Direct Mode), LM (Linear Mode), vertikální režim-26, horizontální-10, DC-1 a rovinný režim-0. Režimy DM a LM se používají, pokud struktura jasové složky koreluje s chromatickými složkami [9,11].



**Obrázek 10: Vnitrosnímkové módy pro jasovou složku[9]**

### 3.3 Mezisnímková predikce

Mezisnímková predikce, jejíž princip je využíván od standardu MPEG-1, používá snímky I (intra), P (predicted) a B (bidirectional). Referenční snímek I, neboli intra frame, je komprimován pouze metodami vnitrosnímkové predikce. U snímku P dochází k dopředné predikci, tzn. snímek je zakódován rozdílem snímku I nebo jiného snímku typu P. Pro ještě větší snížení redundance se používá obousměrná predikce, kdy aktuální snímek je kódován jako průměr ze snímku minulého a následujícího. Takový snímek je označován B a je kódován pomocí snímku I a P. Kvůli kódování z následujícího snímku musí být část snímků uložena v paměti. Všechny tyto typy snímků vytváří skupinu snímku – **Group of Pictures (GOP)**.

U standardu HEVC může být použito více referenčních snímků a to až 16. Dále se u standardu HEVC používají dva referenční seznamy (L0 a L1) a tři režimy mezisnímkové predikce. Režim inter využívá pohybového vektoru a referenčního snímku, kdežto režim merge vytváří seznam již dříve zakódovaných predikčních jednotek, tzn. kandidátů. Kandidáti jsou časově nebo prostorově spojeni s aktuální predikční jednotkou. Zakódování se provede za pomoci indexu kandidáta a pohybová data pro aktuální jednotku se jen zkopírují. Třetí režim skip se používá pro jednotky,

keré se v obraze moc nehýbou, takže se jen odkáže na informace z předcházejícího snímku.

Kompenzace pohybu má přesnost  $\frac{1}{4}$  pixelu pro jasovou složku. Pro výpočet se používá interpolační filtr osmého nebo sedmého řádu. Barevná složka má přesnost  $\frac{1}{8}$  pixelu a používá filtr čtvrtého řádu [9,13].

### 3.4 Transformace

Po provedení predikce zůstává rozdílový signál, který se zakóduje pomocí transformačních jednotek TU. Transformační blok může mít velikost 4x4, 8x8, 16x16 a 32x32. Pro bloky o velikosti 32x32 až 8x8 je použita celočíselná diskrétní kosinová transformace (**Discrete Cosine Transform, DCT**), avšak při vnitrosnímkové predikci pro jasové složky může být u standardu HEVC použita také matice o velikosti 4x4. U této matice je poté použita celočíselná diskrétní sinusová transformace (**Discrete Sine Transform, DST**). DST lépe odpovídá statistickým vlastnostem a dokáže zmenšit bitový tok o 1 %proti DCT [9].

### 3.5 Kvantování

Po přetransformování do bloků je provedeno kvantování, jež má za úkol odstranit transformační koeficienty, které obsahují větší množství detailů. Při kvantování dochází ke trvalé ztrátě části informací z videa.

### 3.6 Rekonstrukční filtry

Mezi rekonstrukční filtry řadíme protiblokový filtr (**Deblocking Filter, DBF**) a filtr s adaptivním offsetem (**Sample Adaptive Offset, SAO**). Filtry slouží k odstranění zkreslení hlavních kroků při kódování, kterými jsou rozdělení snímku do bloků, predikce, transformace a kvantování. Protiblokový filtr, který se používal už v předcházejícím standardu, má za úkol potlačit blokovou strukturu. Hranice bloku pro tento filtr představují jednotky CU, PU nebo TU. U HEVC se nově nefiltrují bloky o velikosti 4x4, tím se sníží výpočetní náročnost a vizuálně se obraz nezhorší. Volba síly filtrace je závislá na gradientu vzorků obrazu na hranici bloku.

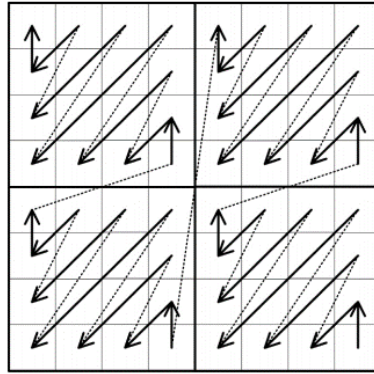
Filtr s adaptivním offsetem je nově u standardu HEVC zařazen po protiblokovém filtru. Jeho hlavní cíl je lepší rekonstrukce amplitudy původního signálu. SAO filtr klasifikuje obrazové body do různých kategorií na základě intenzity nebo vlastností hran. Podle kategorie zařazení se k obrazovým bodům přidává takový posun, aby amplituda co nejvíce odpovídala originálu. Tento filtr je určen pro potlačení artefaktu, který způsobuje pruhování.

Díky těmto filtrům u standardu HEVC dostáváme lepší obraz a tím kvalitnější informace pro odhadování pohybu [9,10].

### 3.7 Entropické kódování

Po transformaci a kvantování se bloky převádí na sériový tok dat. Za tímto účelem se u standardu HEVC pro mezisnímkovou predikci již nepoužívá čtení bloků zig-zag, ale diagonální vyčítání, které je zobrazeno na **obrázku 11**. Diagonální vyčítání se provádí na úrovni bloků TB o velikosti 4x4. Pokud je blok větší, dojde k jeho rozdělení na bloky o velikosti 4x4, které se mohou zpracovávat paralelně. U vnitrosnímkové predikce může být použito horizontální nebo vertikální vyčítání a to také pro bloky TU o velikosti 4x4 [15].





**Obrázek 11: Ukázka diagonálního vyčítání a rozdělení do bloků 4x4 [15]**

U standardu HEVC se používá metoda pro entropické kódování podle algoritmu **CABAC (Context Adaptive Binary Arithmetic Coding)**. Algoritmus CABAC se skládá ze tří kroků: binarizace, kontextové modelování a aritmetické modelování. V prvním kroku jsou všechny hodnoty převedeny na binární kód. Po převedení dat se vybere kontextový model z nabídky už dostupných modelů v závislosti na nedávno kódovaných datových symbolech. Kontextový model ukládá pravděpodobnost každého datového symbolu. Poté aritmetický kódér zakóduje každé datové slovo podle zvoleného kontextového modelu. A nakonec je kontextový model aktualizován na základě kódované hodnoty, a tím se stále přizpůsobují aktuální statistiky kontextového modelu.

Algoritmus CABAC byl použit i u předcházejícího standardu H.264/AVC. U H.265/HEVC došlo k vylepšení binarizace a kontextového modelování, aritmetické kódování zůstalo stejné. CABAC má dva provozní režimy, při čemž první funguje jako klasický CABAC a druhý režim přiřadí stejný kontextový model více kódovým slovům. Tento režim přináší zrychlení, snazší paralelizaci a tím také větší propustnost [9,10].

## 4. Kompresce vícehledového a stereoskopického videa

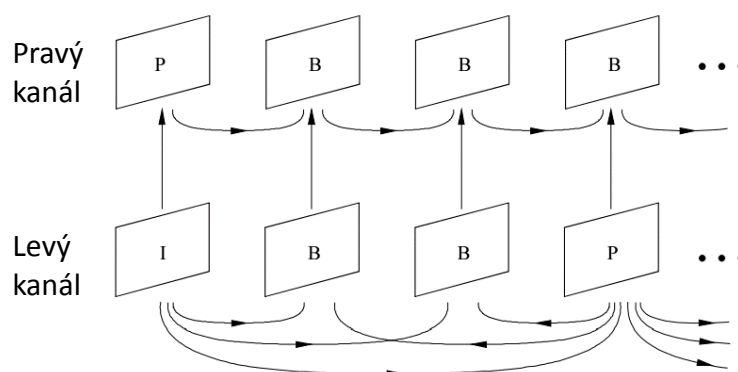
V posledních letech došlo k nárůstu tzv. 3D videí, které využívají princip stereoskopického vnímání, jímž se vytváří dojem prostorového vjemu. Stereoskopické video seskládá ze dvou obrazů, jednoho pro každé oko. Tato videa se z tohoto důvodu natáčejí pomocí dvou kamer snímajících stejnou scénu vzdálených od sebe průměrně 64 mm, což odpovídá průměrné vzdálenosti očí. Stereoskopické video je speciální případ vícehledového videa (Multiview Video). Vícehledová videa vznikají snímáním stejné scény pomocí více kamer, které jsou od sebe vzdálené v určitých intervalech. Obrazy z jednotlivých kamer jsou velice podobné a kvůli tomu obsahují vysokou dávku redundantních informací.

Skupina MPEG, přední mezinárodní standardizační skupina pro multimédia, se také zabývá kódováním vícehledového videa včetně stereoskopického videa. První standard pro kódování stereoskopického videa MPEG-2 se objevuje jako jeden z možných profilů s názvem obsahující multiview. Další standard pro kódování vícehledového videa vychází ze standardu MPEG-4 AVC a nese název **Multiview** video coding (MVC). V rozšíření MVC se začal používat formát obrazových dat v podobě textury a hloubkové mapy (MVC+D a 3D-AVC). Kódováním vícehledového videa se zabývá i nejnovější standard pro kódování videa H.265/HEVC a dané profily se nazývají MV-HEVC a 3D HEVC. V následujících podkapitolách jsou blíže popsány způsoby kódování stereoskopického videa.

### 4.1 MPEG-2 profil multiview

MPEG-2 profil multiview je rozšiřujícím profilem a do standardu MPEG-2 byl přidán v roce 1996. Kódování stereoskopického videa probíhá velmi podobně jako u videa monoskopického. Obraz levého kanálu je standardně kódován hlavním profilem, takže ke snížení bitového toku se využívá vektorů pohybu pro stejné makrobloky. U pravého kanálu se také používají vektory pohybu, a to jak mezi snímky v pravém kanále, tak i ze snímků z levého kanálu. Princip je zobrazen na **obrázku 12**. Pro přenos obou kanálů do jednoho toku se využívá časové škálovatelnosti, při níž se navýší počet snímků za sekundu. Levý kanál je vysílán v základní vrstvě a pravý kanál je vysílán ve vrstvě vyšší. Díky tomuto způsobu kódování lze sledovat na přijímači bez podpory multiview profilu alespoň levý kanál, tedy jedнопohledové video [23].

Multiview profil nebyl příliš používán, protože nebyly vyvinuty zobrazovací technologie a větší důraz se kladl na přechod ze standardního rozlišení na rozlišení HD. Přesto na tento profil navazuje Multiview video coding (MVC).

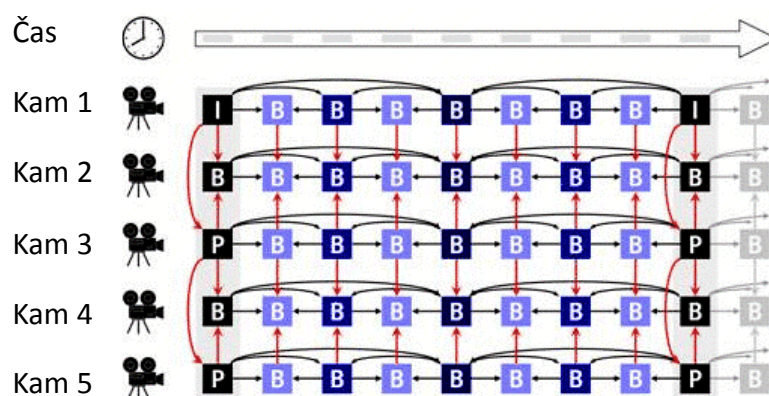


Obrázek 12: Ilustrace přenosu vektorů pohybu [23]

## 4.2 Multiview video coding

Multiview video coding (MVC) je rozšířením standardu MPEG-4 AVC. Toto rozšíření přináší možnost vylepšeného kódování vícepohledového a stereoskopického videa a také kódování **Free Viewpoint Video (FVV)**. FVV je typ videa, u kterého si uživatel může libovolně zvolit místo, z něž bude scénu pozorovat. Tuto volbu vícepohledová videa neumožňují. MVC kódování využívá korelace mezi jednotlivými snímky. Ve formátu MVC jsou kódované 3D filmy na blu-ray.

Jak už bylo uvedeno výše, MVC vychází z MPEG-4 AVC, a proto MVC shodně používá flexibilní funkce a způsob použití predikce. Za účelem snížení dat skloubilo MVC časovou a inter-view predikci. Spojení predikcí je možné, neboť k dříve dekodovaným snímkům jsou k dispozici referenční snímky, které mohou být použity jak pro časovou predikci snímků z konkrétní kamery, tak i pro inter-view predikce mezi jednotlivými kamerami. Princip časové a inter-view predikce je naznačen na **obrázku 13**. Konstrukce MVC umožňuje pouze inter-view predikci mezi aktuálními snímky. Dále je nutné nejprve kódovat video z první kamery, až poté se kóduje video z třetí kamery a následně dochází ke kódování video z kamery druhé. Mohou být použity i jiná různá schémata pro organizaci I, P a B snímků, a to jak v časové, tak i v prostorové doméně. Volba vhodného schématu je omezena pouze velikostí paměti a zpožděním. Aby bylo možné jednotlivé snímky zrekonstruovat, jsou zavedeny 3 typy snímků: IDR, kotevní a predikované. IDR jsou snímky, které využívají vnitrosnímkové kódování bez časových i prostorových referencí. Na rozdíl od IDR, kotevní snímky mají povoleno používat prostorové reference. Všechny ostatní snímky jsou označovány jako predikované.



Obrázek 13: Schéma časové a inter-view predikce pro MVC [25]

### 4.2.1 Profily a úrovně

Stejně jako video standardy používá pojmy profily a úrovně i rozšíření MVC. Každý profil určuje skupinu kódovacích nástrojů. MVC obsahuje dva profily: **Multiview High Profile (MHP)** a **Stereo High Profile (SHP)**.

Oba tyto profily vychází z High Profile (HP) standardu MPEG-4 AVC a liší se v těchto bodech:

- Multiview High Profile podporuje více pohledů, ale nepodporuje prokládané kódování,
- Stereo High Profile je omezen pouze na dva pohledy a podporuje prokládané kódování.

Pro základní video, z kterého jsou použity další predikce, se v profilu SHP používá High Profile. Kdežto pro základní video profilu MHP je použit profil Constrained Baseline. MHP používá i další nástroje z HP kromě nástrojů pro prokládání (Field Coding a MBAFF), protože profil prokládání videa nepodporuje.

Úrovně definují určitá omezení, z kterých vyplývá výkonnost dekodéru. Omezení se vztahují na velikost paměti, počet makrobloků za sekundu, maximální velikost snímku, celkovou přenosovou rychlost atd.

### 4.3 MVC kódování s využitím hloubkové mapy

Další možná metoda kódování využívá 2D barevný snímek (texturu) a hloubkové mapy. Hloubková mapa je snímek, který obsahuje 256 odstínů (8 bitů) šedé, kde bílá barva představuje nejbližší bod před obrazovou rovinou a černá nejdálčenější bod od obrazové roviny. Textura a hloubková mapa jsou znázorněny na **obrázku 14** [26].



**Obrázek 14: Textura a hloubková mapa [26]**

Výhoda tohoto formátu spočívá v možnosti přenosu malého počtu textur s hloubkovou mapou a ve vykreslení více pohledů pomocí **algoritmu DIBR (Depth Image Based Rendering)**. Tento algoritmus je převážně využíván pro zobrazování na auto-stereoskopických displejích, protože klasické MVC by zabíralo větší šířku pásma úměrnou počtu pohledů zobrazených na displeji. Standardy s hloubkovou mapou jsou kompatibilní jak pro stereoskopické displeje, které potřebují jen dva pohledy, tak i pro klasické monoskopické displeje.

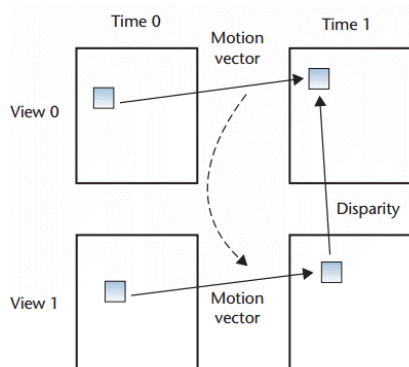
Standardy využívající hloubkovou mapu jsou rozděleny do dvou kategorií. První kategorie je kompatibilní se standardem MPEG-4 AVC a druhá kategorie se standardem HEVC. Tyto kategorie jsou dále děleny na dva paralelní vývoje. Standard **3D Advanced Video Coding (3D-AVC)** je kompatibilní se standardem AVC, oproti tomu jsou standardy **Multiview High Efficiency Video Coding (MV-HEVC)** a **3D High Efficiency Video Coding (3D HEVC)** kompatibilní se standardem HEVC. V následujících odstavcích jsou tyto standardy blíže popsány.

#### 4.3.1 3D Advanced Video Coding

Také 3D Advanced Video Coding (3D-AVC) je kompatibilní s MVC. Tento standard se odlišuje od MVC+D tím, že využívá redundanci mezi texturou a hloubkovou mapou. Z toho důvodu se textura a hloubková mapa kódují na sobě závisle. Pomocí několika nástrojů tak 3D-AVC poskytuje lepší kompresy než jaké je dosaženo u MVC+D. Těmito nástroji jsou Neighboring Block-Based Disparity Vector Derivation, Inter-View Motion Prediction, Motion Prediction from Texture to Depth, View Synthesis Prediction a kompenzace osvětlení [27].

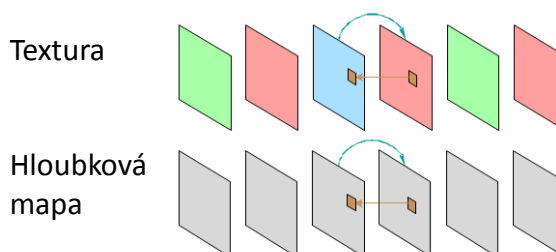
**Neighboring Block-Based Disparity Vector Derivation (NBDV)** při kódování textury využívá podobné sousední bloky, které již jsou zakódované. Místo v bitovém toku ušetří tím, že aktuální blok nekóduje zvlášť, ale pouze se odkazuje na pohybové vektory sousedních bloků.

**Inter-View Motion Prediction (IVMP)** při kódování využívá vysokou míru korelace pohybové informace mezi jednotlivými pohledy. Tato metoda odvozuje informace o pohybu z jednoho pohledu a tyto informace vkládá do pohledů dalších. Na **obrázku 15** je zobrazen princip motion prediction mezi dvěma pohledy.



**Obrázek 15: Motion prediction mezi dvěma snímky [28]**

**Motion Prediction from Texture to Depth (MPTD)** využívá faktu, že textura a k ní přidružená hloubková mapa zobrazují podobné siluety, které se mohou ve snímku pohybovat podobně. Proto je zde povolena pohybová predikce z textury do přidružené hloubkové mapy. Tato predikce je zobrazena na **obrázku 16**.



**Obrázek 16: Motion Prediction from Texture to Depth [27]**

**View Synthesis Prediction (VSP)** je technika, která poskytuje další informace o možnosti predikce, a to vytvořením syntetických pohledů pomocí algoritmu DIBR využívajícího hloubkové mapy.

**Kompensace osvětlení** je technikou používanou pro potlačení jasových nesrovnalostí proto, aby nedocházelo k selhání predikce. K selhání predikce mezi snímky dochází, když nejsou kamery stejně barevně kalibrovány nebo když jsou na snímku různé světelné efekty.

#### 4.3.2 Multiview-High Efficiency Video Coding

V roce 2014 uvedla společnost JCV-VC druhé vydání normy HEVC i s rozšířením pro profil Multiview High Efficiency Video Coding (MV-HEVC). Jak již z názvu vyplývá, standard MV-HEVC vychází ze standardu HEVC. MV-HEVC tudíž využívá všechny funkce pro kódování videa ze standardu HEVC. MV-HEVC používá stejný konstrukční princip jako v případě rozšíření MVC, který umožňuje provádět inter-view predikci z aktuálních snímků. Tato metoda zajišťuje kompatibilitu s jedнопohledovým kódováním HEVC [23].

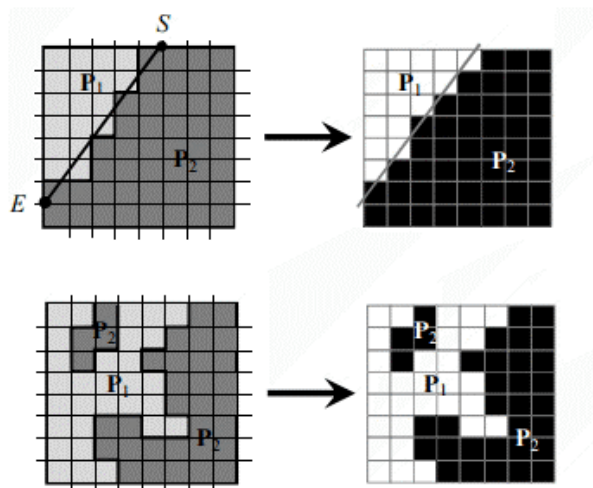
#### 4.3.3 3D High Efficiency Video Coding

3D High Efficiency Video Coding (3D HEVC) je vznikající standard pro kódování 3D videa. Jedná se o možné rozšíření standardu HEVC. 3D HEVC ze standardu HEVC využívá všechny funkce pro kódování videa – textur a přidává další funkce pro zvýšení výkonnosti kódování. Stejně jako u standardu 3D-AVC je kódování textury a hloubkové mapy na sobě závislé. Standardy závislost kódování využívají ke snížení objemu dat. Dále 3DHEVC využívá stejné nástroje jako 3D-AVC

(IVMP, VSP a kompenzace osvětlení), navíc přidává nové kódovací techniky hlavně pro kódování hloubky. Těmito techniky jsou Depth Motion Prediction, Partition-Based Depth Intra Coding, Adjustment of QP of texture on depth data, Post Processing in Loop Filter [23].

**DepthMotionPrediction** používá predikci pohybu pro hloubkové mapy, podobně jako se predikce používá pro textury.

**Partition-BasedDepth Intra Coding** potlačuje artefakty, které mohou vzniknout při vnitrosnímkovém kódování hloubkové mapy. Když je při kódování zjištěna významná hrana, za použití této techniky je blok rozdělen na dva oddíly, které jsou pak reprezentovány s konstantní hodnotou, viz **obrázek 17**.



**Obrázek 17: Partition-BasedDepth Intra Coding [23]**

**Adjustment of QP of texture on depth data** je algoritmus zvyšující kvalitu obrazu v popředí a snižující kvalitu obrazu v pozadí. To je možné díky použití přizpůsobivého kvantizačního parametru QP.

**Post Processing in Loop Filter** je technika používající dva filtry pro syntetické pohledy za účelem snížení nechtěných artefaktů. Filtr Depth-Gradient-based Loop back Filter (DGLF) se používá pro hloubkové mapy. Je to gradientní filtr, který se aplikuje na oblasti hran. Druhý filtr Availability Deblocking Loop back Filter (ADLF) zmenšuje artefakty po blokovém kódování. Tento filtr vytváří hladký přechod mezi syntetizovanou a kódovanou oblastí.

## 5. Nesymetrické kódování

Podstatnou nevýhodou stereoskopického videa je s ním spojený veliký objem dat. Pro jeho přehrávání je potřeba větší šířka pásma, ale také větší úložné prostory pro jeho skladování. Z těchto důvodů dochází u stereoskopického video ke značné komprimaci, která však nesmí ovlivnit vnímanou kvalitu videa. U stereoskopického videa mohou být k odstranění redundantní informace použity stejné metody jako u klasických 2D videí. Jedná se o snížení časové nebo prostorové vzorkovací frekvence, různé druhy predikce či o snížení bitové hloubky. U stereoskopického videa můžeme navíc využít nesymetrické kódování, při němž se kvalita pravého a levého snímku liší. Nesymetrické kódování vychází z teorie binokulárního vidění. Z teorie vyplývá, že u dvou obrazů, jejichž ostrost je různá, se kvalita vnímání blíží ostřejšímu zobrazení. V následujících odstavcích jsou uvedeny metody nesymetrického kódování [19].

Při aplikaci metody **smíšeného rozlišení** dojde u jednoho z pohledů k podvzorkování, a tím ke snížení prostorové informace. Tento pohled je při přehrávání interpolován do původního rozlišení. Vlivem interpolace dochází ke odstranění detailů ve snímku.

Při metodě **smíšeného kvantování** dochází u obou pohledů ke kvantizaci s rozdílným kvantizačním krokem. Následkem mohou být viditelné artefakty u více kvantovaného snímku.

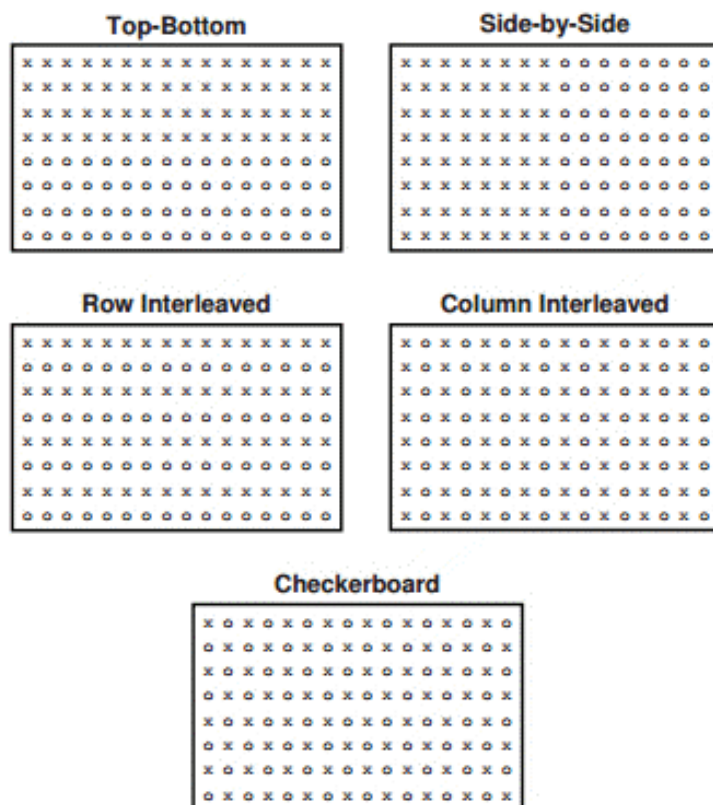
Metoda **smíšeného rozlišení chrominančních složek** vychází z faktu, že lidské vidění je citlivější na jasovou složku v obraze než na barevné složky. Proto jsou v rámci metody v jednom z pohledů podvzorkovány chrominanční složky.

## 6. Kompatibilní formát pro distribuci 3D videa

Potřeba vytvořit kompatibilní formát pro distribuci 3D videa vyplývá z nutnosti distribuovat stereoskopické video pomocí současné přenosové infrastruktury, která je založena na standardu MPEG-2 (DVB-T, DVB-C) a nemá tudíž potřebnou kapacitu. Pro potřeby stávajících přijímačů byly vytvořeny kompatibilní formáty.

Kompatibilní formát je založen na multiplexování dvou pohledů (levého a pravého) do jednoho snímku nebo sekvence snímků. V případě multiplexování kanálů do jednoho snímku musí dojít k podvzorkování snímků. Existuje řada možností jak podvzorkované snímky proložit do sebe. Například se mohou snímky podvzorkovat a proložit ve vertikálním směru. Tím vznikne snímek, ve kterém se střídají řádky levého a pravého kanálu. Nebo se snímky mohou podvzorkovat a proložit ve směru horizontálním, tak vznikne snímek se střídajícími se sloupci. Také je možné decimované snímky vložit vedle sebe a dostat formát Side-by-Side nebo vložit pod sebe a tím dostat formát Top-Bottom. Další možností je jednotlivé snímky rozložit do šachovnice (checkerboard). Možné typy kompatibilních formátů jsou ukázány na **obrázku 18**.

Tyto formáty usnadňují distribuci stereoskopického videa prostřednictvím stávající infrastruktury. Nevýhodou může být ztráta obrazové informace, která může vést až ke ztrátě vnímání hloubky obrazu.



**Obrázek 18:** Kompatibilní formáty kde X představuje vzorek jednoho kanálu a O představuje vzorek druhého kanálu [22]



## 7. Objektivní metody hodnocení videa

Objektivní hodnocení videa je založené na matematických modelech jako je střední kvadratická chyba, špičkový poměr ku šumu nebo také na modelech lidského vnímání vidění. Výhodou objektivních metod může být rychlost dosažení výsledku i fakt, že matematické modely nevyžadují zapojení lidí, kteří by podávali subjektivní hodnocení. Díky tomu jsou objektivní metody i levnější. Nevýhodou objektivních metod je, že získaný výsledek se může pouze přiblížit k hodnotám subjektivních testů. Objektivní metody můžeme rozdělit na metody s referencí a metody bez reference. Metody s referencí hodnotí kvalitu videa porovnáním s referencí (nezkresleným videm). Metody bez reference hodnotí kvalitu na základě získaných informací z videa a jsou vhodné zejména pro určování kvality online aplikací [3]. V dalších kapitolách budou popsány jen metody s referencí.

### 7.1 2D objektivní metriky

2D objektivní metriky slouží k posouzení kvality 2D videa. Matematických modelů používaných v rámci 2D objektivních metrik je velké množství. Mezi nepoužívanější patří **Peak Signal to Noise Ratio** a **Structural Similarity Index**. Tyto metriky jsou v následujících podkapitolách popsány.

#### 7.1.1 Peak Signal to Noise Ratio

Metrika Peak Signal to Noise Ratio (PSNR), neboli špičkový poměr signálu k šumu, vyjadřuje maximální energii signálu vůči energii šumu. K výpočtu PSNR je potřeba nejprve vypočítat střední kvadratickou chybu **Mean Squared Error** (MSE), viz vztah 1 [16].

$$MSE = \frac{1}{M} \cdot \frac{1}{N} \sum_{i=1}^M \sum_{j=1}^N (I_{ref} - I_{por})^2 \quad (1)$$

Ve vztahu 1 jsou  $M$  a  $N$  hodnoty rozlišení obrazu,  $I_{ref}$  je referenční snímek,  $I_{por}$  je snímek, který porovnáváme s referenčním. Pokud má MSE hodnotu 0, znamená to, že referenční a hodnocený snímek jsou stejné. Hodnoty vyšší než 0 znamenají zhoršení kvality hodnoceného snímku.

PSNR se vypočítá dle vzorce (2).

$$PSNR = 20 \log_{20} \frac{MAX_I}{\sqrt{MSE}} \quad (2)$$

$MAX_I$  je maximální hodnota pixelu. Jestliže je pixel prezentován 8 bity na vzorek, jeho maximální hodnota je 255. Metrika PSNR vychází v jednotkách dB. Čím vyšší je hodnota PSNR, tím jsou si obrazy podobnější.

#### 7.1.2 Structural Similarity Index

Metrika Structural Similarity Index (SSIM) bere v potaz charakteristické reakce lidského oka. SSIM je založen na tom, že lidské vnímání (Human Visual System, HVS) je citlivé na strukturální informace v obrazu. Výsledek SSIM je získán z porovnání jasu, kontrastu a struktury snímku. Metoda SSIM používá náročné metody a proto zde bude uveden jen výsledný vztah a to jako vzorec 3 [17].

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

SSIM nabývá hodnot -1 až 1, přičemž hodnota 1 znamená, že se porovnávané obrazy shodují.

## 7.2 Stereoskopické objektivní metriky

Stereoskopické objektivní metriky slouží k posouzení kvality stereoskopického videa. Oproti 2D videu má stereoskopické video prostorový vjem. Stereoskopické metriky vycházejí ze základních 2D metrik jako je PSNR a SSIM a aby v kvalitě videa obsáhly i artefakty typické pro stereoskopické video, používají další algoritmus pro vypočítání disparity, hloubkové mapy nebo například crosstalku. Takto může být například kombinace metriky SSIM, které se použijí na každý jednotlivý pohled, na hloubkovou mapu vypočítanou z referenčních snímků i na hloubkovou mapu vypočítanou ze snímků testovaných.

### 7.2.1 Objektivní metody hodnocení stereoskopického videa

Pro objektivní analýzu nesymetrického stereoskopického videa je vhodné použít metriku vycházející z 2D objektivní metriky SSIM a to na každý pohled stereoskopického videa zvlášť, navíc s přidáním binokulárního modelu [18].

**Binokulární model** spočívá ve vytvoření lokální mapy energie snímku. Lokální mapa energie se dá vytvořit z výpočtu lokálního rozptylu obrazu. V programu Matlab je možné použít funkci *stdfilt.m*, která vytvoří normovaný lokální rozptyl obrazu  $E_{i,r,l}$  a  $E_{i,r,r}$  pro levý a pravý referenční snímky, dále  $E_{i,d,l}$  a  $E_{i,d,r}$  pro porovnávané levé a pravé referenční snímky. Výpočet lokálních poměrů map energie pro oba pohledy lze provést podle vzorce 4.

$$R_{i,l} = \frac{E_{i,d,l}}{E_{i,r,l}} \quad R_{i,r} = \frac{E_{i,d,r}}{E_{i,r,r}} \quad (4)$$

Poměry map energie poskytují užitečné informace, které mohou být kombinovány s jednopohledovými metrikami. Regiony s vysokou energií budou s největší pravděpodobností obsahovat větší množství informací. Z tohoto důvodu je nutné těmto regionům věnovat zvýšenou pozornost. Za tímto účelem se provádí váhování podle vzorce 5.

$$g_{i,l} = \frac{\sum E_{i,d,l} R_{i,l}}{E_{i,d,l}} \quad g_{i,r} = \frac{\sum E_{i,d,r} R_{i,r}}{E_{i,d,r}} \quad (5)$$

Kde  $g_{i,l}$  a  $g_{i,r}$  představují úroveň dominance pravého i levého snímku. S  $g_{i,l}$  a  $g_{i,r}$ . Další váhování se provádí podle vzorce 6.

$$w_{i,l} = \frac{g_{i,l}^2}{g_{i,l}^2 + g_{i,r}^2} \quad w_{i,r} = \frac{g_{i,r}^2}{g_{i,l}^2 + g_{i,r}^2} \quad (6)$$

Následným váhováním  $w_{i,l}$ ,  $w_{i,r}$  zvolenou 2D metrikou podle vzorce 7 získáme výslednou metriku  $Q_i^{3D}$  pro každý snímek zvlášť.

$$Q_i^{3D} = w_{i,l} Q_{i,l}^{2D} + w_{i,r} Q_{i,r}^{2D} \quad (7)$$

Jednu hodnotu výsledné metriky pro celé video je možné vypočítat jako průměr hodnot  $Q_i^{3D}$  [18].

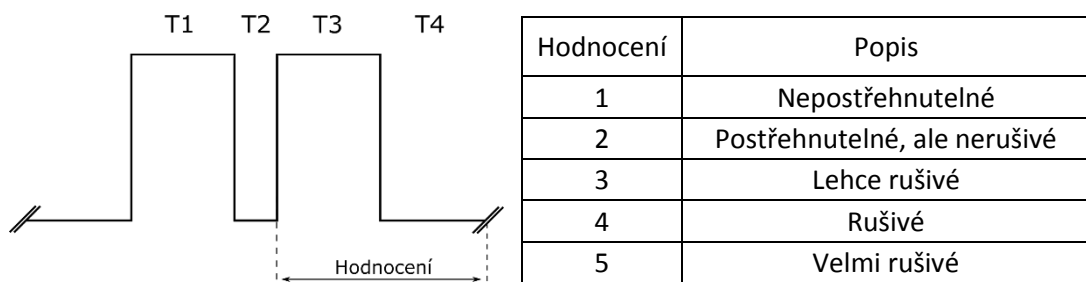
## 8. Subjektivní metody hodnocení videa

Subjektivní metody hodnocení videa jsou založeny na hodnocení kvality skupinou pozorovatelů. To je spojeno se značnou časovou a finanční náročností těchto metod. Podle normy ITU-R BT. 2021-1 je definováno několik metod umožňujících subjektivní hodnocení [29]. V následujících

podkapitole bude popsána metoda pro subjektivní hodnocení stereoskopického videa vycházející z této normy.

### 8.1 Subjektivní metoda hodnocení stereoskopického videa

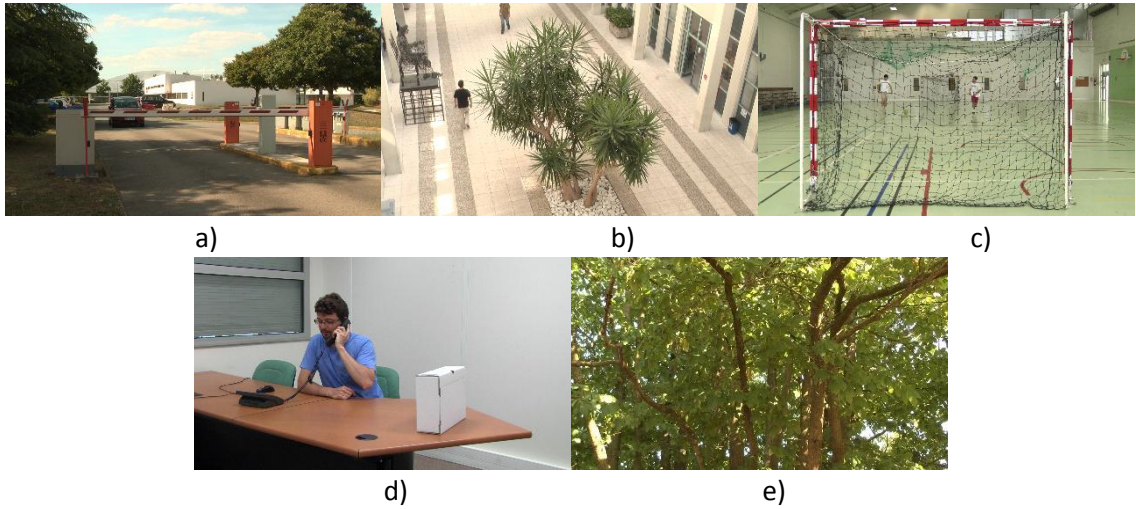
Popsaná subjektivní metoda pro hodnocení stereoskopického videa vychází z metody **Double Stimulus Impairment Scale** (DSIS) podle doporučení ITU-R BT. 2021-1 [1]. Při této metodě se porovnávají dvě videosekvence. První videosekvence je vždy referenční, tedy ničím nezkraslená, a druhá videosekvence obsahuje určité zkraslení. Celá videosekvence začíná referenční videosekvencí (T1). Po jejím skončení následuje tři sekundová pauza (T2) v podobě střední úrovně šedé. Za pauzou se objevuje zkraslená videosekvence (T3) a na ní navazuje sekvence (T4) se střední úrovní šedé, při které je pozorovatel vyzván k hodnocení videosekvence T3. Sekvence T4 obvykle trvá 5 až 11 sekund a poté následuje nová dvojice videosekvencí T1 a T3. Vyhodnocení videosekvence probíhá na stupnici 1-5 (viz tabulka v **obrázku 19**). Oproti doporučení [29] je tato stupnice převrácená. Pozorovatel hodnotí soubor videosekvencí, který obsahuje všechny požadované kombinace zkraslení i neporušené oddíly v pseudonáhodném pořadí. Celé pozorování by nemělo přesahovat 30 minut kvůli možné únavě pozorovatelů.



**Obrázek 19: Ukázka průběhu DSIS testu a tabulka hodnocení**

#### 8.1.1 Popis databáze

Pro subjektivní testy jsou voleny referenční videa z volně dostupných databází, například z databáze **Nama3ds1-cospad1** [30]. Databáze obsahuje 10 stereoskopických snímků v rozlišení 1920x1080 pixelů pro každý pohled s obnovovací frekvencí 25 snímků za sekundu. Stereoskopické videosekvence byly pořízeny kamerou Panasonic AG-3DA1E.



**Obrázek 20: Ukázka stého levého snímku z videosekvencí a) Barrier b) Hall c) Soccer d) Phone call e) Tree branches**

### 8.1.2 Test zraku

Před samotným subjektivním testem musí proběhnout test zraku pozorovatelů. Účelem testu je odhalit vady zraku, které by mohly mít negativní vliv na vnímání stereoskopického videa, což by vedlo k zanesení chyb do výsledků hodnocení. Test zraku se skládá ze zkoušek na oční dominanci, zrakovou ostrost a vnímání hloubky stereoskopického obrazu.

## 9. Experimentální část

### 9.1 Funkčnost kodéru HEVC

Ke kodování, dekódování a stříhání videí použitých při subjektivním testu kvality stereoskopického videa byl zvolen volně dostupný **software ffmpeg**. Tento software je vyvíjen pod operačním systémem Linux, ale je zkompileován pod většinu operačních systémů včetně operačního systému Windows. Software ffmpeg dokáže pracovat s videem i audiem. Tento software obsahuje knihovny, které dokážou kódovat a dekódovat většinu audio a video formátů, dále převádět audio a video do různých formátů a také obsahují různé filtry pro práci s videem. V balíčku ffmpeg najdeme i analyzátor multimediálních streamů nebo jednoduchý přehrávač. Software ffmpeg se ovládá pomocí příkazů z příkazové řádky.

Pro kódování videa do formátu H.264/AVC slouží knihovna libx264 vyvíjená společností x264. Pro kódování videa do formátu H.265/HEVC, jež je nástupcem formátu H.264/AVC, slouží knihovna libx265, která je vyvíjena společností x265. Pro určení funkčnosti kodéru bylo provedeno porovnání formátu H.264/AVC a H.265/HEVC.

Pro porovnání obou formátů bylo použito video Barrier0 z databáze **Nama3ds1-cospad1**. Toto video má rozlišení 1920x1080 pixelů s obnovovací frekvencí 25 snímků za sekundu. Video je dlouhé 16 sekund a tudíž se skládá ze 400 snímků. Pro dané video bylo použito kódování s podvzorkováním chrominančních složek na 4:2:0 u obou formátů se zvoleným bitovým tokem. Oba kodéry byli nastavené v defaultním nastavení. Pro porovnání byla použita objektivní metrika Peak Signal to Noise Ratio (PSNR).

### 9.2 Objektivní test kvality stereoskopického videa

Při objektivním testování byla referenční videa porovnána s upravenými videosekvencemi. Testované snímky se dělí na symetricky a nesymetricky kódované videosekvence. Symetricky kódované videa byla komprimována kvantizačním parametrem Qp 20, 30 a 40. U nesymetricky kódovaných videí bylo sníženo rozlišení pravého snímku na hodnoty  $\frac{1}{2}$ ,  $\frac{1}{3}$  a  $\frac{1}{4}$  a také byla komprimována kvantizačním parametrem Qp 20, 30 a 40. Použitá objektivní metoda je popsána v **kapitole 8.3 Objektivní metody hodnocení stereoskopického videa**. K výpočtu objektivní metody byl použit matlabovský skript 3DSSIM, který vypočítává metriku pro každý snímek. Výsledná hodnota metriky je aritmetický průměr ze všech snímků.

### 9.3 Subjektivní test kvality stereoskopického videa

Subjektivní test hodnocení stereoskopického videa byl proveden podle postupu popsaného v **kapitole 9.1 Subjektivní metoda hodnocení stereoskopického videa**. Referenční snímky byly vybrány z databáze zmíněné v této kapitole a to z databáze **Nama3ds1-cospad1**.

### 9.3.1 Testovací sekvence

Do subjektivního testu bylo vybráno 5 videí, které se liší v zobrazených barvách, textuře a hloubce. Videá se dále lišila tím, zda se video odehrává v interiéru či exteriéru. Popis testovaných videí je uveden v **tabulce 1**.

**Tabulka 1: Popis použitých videí**

Název	Barrier gate	Hall	Phone call	Soccer	Tree branches
Počet scén	1	1	1	2	1
Prostředí	Exteriér	Interiér	Interiér	Interiér	Exteriér
Popis	Auto a závora	Osoby se potkají v hale	Telefonování v kanceláři	1. 2 hráči hrají 2. vstřelení gólu	Listy stromů ve větru
Hloubka obrazu	Střední	Velká	Krátká	Velká	Střední

Všechna stereoskopická videa byla získána z databáze ve formátu AVI. Pomocí aplikace ffmpeg byli přeformátováni na formát YUV 4:2:0, který podporuje hlavní profil kodeku H.265/HEVC.

V testu bylo pro zpracování videosekvencí použito symetrické i nesymetrické kódování. Symetrické kódování spočívalo v prostorové kompresi do formátu Side-by-Side (SBS) a v komprimaci konstantním kvantizačním parametrem QP (10,20,30,40) pomocí kompresního standartu H.265/HEVC. Komprese byla provedena softwarem ffmpeg a pomocí knihovny libx265. U nesymetrického kódování bylo zmenšeno rozlišení (na  $\frac{1}{2}$ ,  $\frac{1}{3}$  a  $\frac{1}{4}$ ) jednoho pohledu ve vertikálním i horizontálním směru a posléze byl obraz interpolován do původního rozlišení 1920x1080. Až poté byly oba pohledy prostorově stlačeny do formátu SBS a komprimovány konstantním kvantizačním parametrem QP za pomoci kodeku H.265/HEVC, stejně jako u symetrického kódování. Jako reference posloužili symetricky kódovaná videa s konstantním kvantizačním parametrem QP 10.

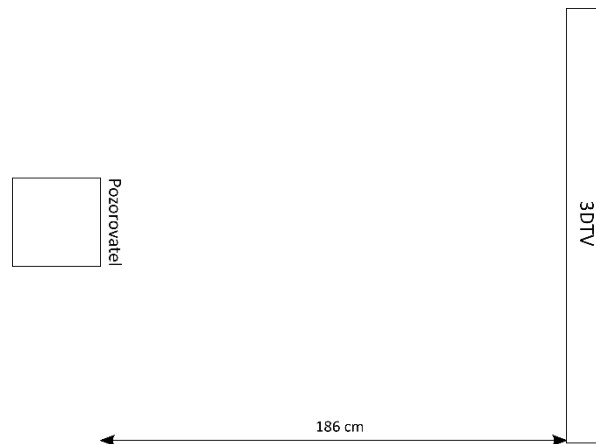
Test obsahoval celkově 82 videosekvencí, z toho dvě videosekvence byly zkušební. Z dvaceti symetricky kódovaných videosekvencí bylo 5 videosekvencí stejných jako referenční videosekvence. Zbylé videosekvence byly asymetricky kódované.

### 9.3.2 Popis testování

Před samotným subjektivním testem každý pozorovatel prošel testem zraku. Test zraku se skládal ze zkoušek uvedených v **kapitole 9.3 Test zraku**.

Testování dominance oka proběhlo následovně: pozorovatel si dal ruce před sebe tak, že si mezi spojením palců a ukazováčků obou ruku vytvořil trojúhelník, skrz který díval oběma očima na předmět vzdálený 6 metrů. Poté se pozorovatel na stejný předmět podíval jen pravým okem, respektive jen levým okem. Dominance byla určena v závislosti na míře vnímaného posunu pozorovaného předmětu. Dominantní oko je to, u něž pozorovatel nezaznamená posun předmětu.

Samotný test proběhl na 3D televizi Sony KDL-50W805C, která využívá aktivní technologii k zobrazení stereoskopického videa s použitím brýlí Sony TDB-BT500A. Televize má úhlopříčku 50" (127 cm). Výška displeje televizoru je 62 cm, pozorovatel seděl ve vzdálenosti 183 cm od displeje. Vzdálenost pozorovatele testu a zobrazujícího displeje tedy byla trojnásobek výšky displeje (**Obrázek 21**). Na začátek testu byly připraveny dvě zkušební sekvence pro seznámení pozorovatele s principem testu. Ve zkušební sekvenci byl zařazen snímek s nejhorší kvalitou a také s nejlepší kvalitou. Pozorovatelé byli na tento fakt upozorněni. Testovací videosekvenci zhlédl pozorovatel vždy jen jednou. Subjektivní testy probíhaly v temné komoře.



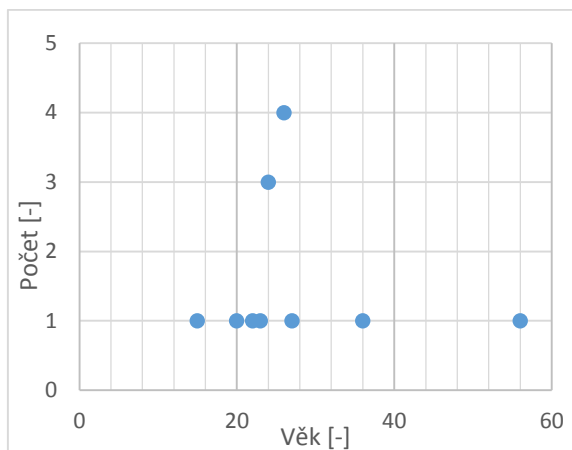
**Obrázek 21: Vzdálenost pozorovatele a televize**

Cíl testu se skládal ze tří dílčích kroků:

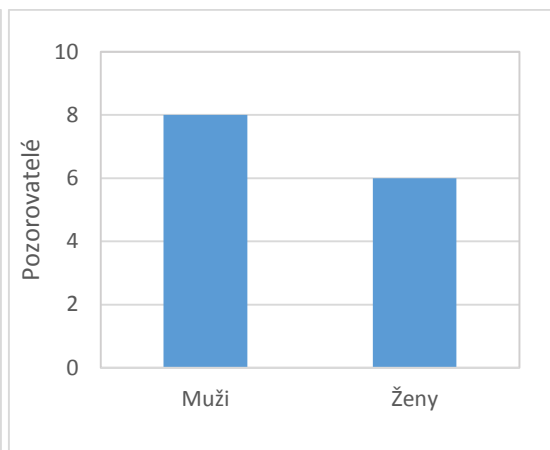
- 1) Zjistit, při jakých datových tocích zaniká hloubka, tedy 3D efekt.
- 2) Prozkoumat, zda dochází ke zkreslení snímků.
- 3) Určit celkový vizuální komfort pozorovatele.

### 9.3.3 Účastníci subjektivního testu

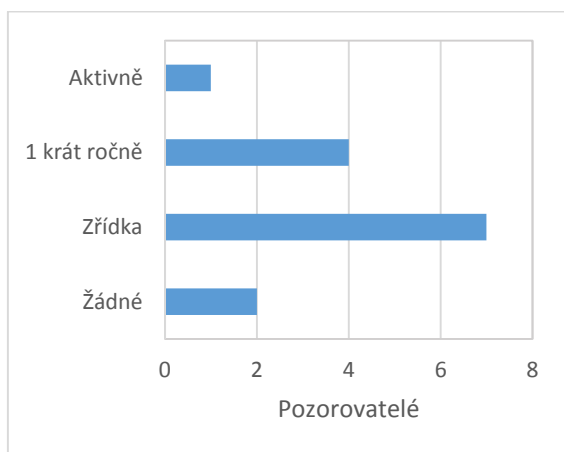
Subjektivního testu se zúčastnilo celkem 14 osob, 8 mužů a 6 žen. Nejmladšímu účastníku testu bylo 15 let, nejstaršímu účastníku testu bylo 56. Průměr stáří účastníku byl 26,7 let. Test na zjištění dominantního oka ukázal, že polovina (7 pozorovatelů) má dominantní oko levé a druhá polovina má dominantní oko pravé. Dva účastníci neměli žádné předchozí zkušenosti s 3D videem, 7 účastníků pozoruje 3D videa občas, 4 účastníci pouze jednou za rok a jeden sleduje 3D videa aktivně. Statistiky jsou zobrazeny na **grafech 1-4**.



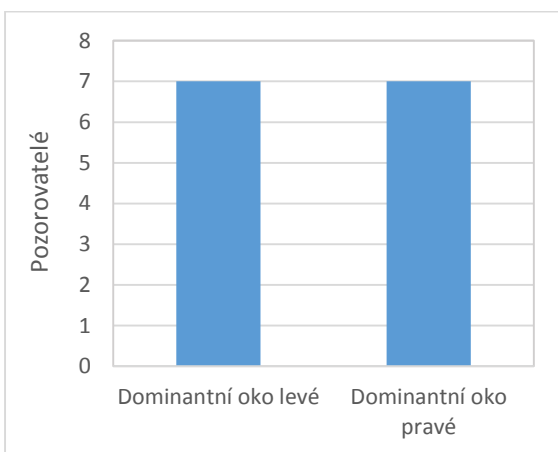
**Graf 1: Rozložení věku pozorovatelů**



**Graf 2: Rozložení pohlaví pozorovatelů**



**Graf 3: Rozložení zkušeností pozorovatelů s 3D videem**



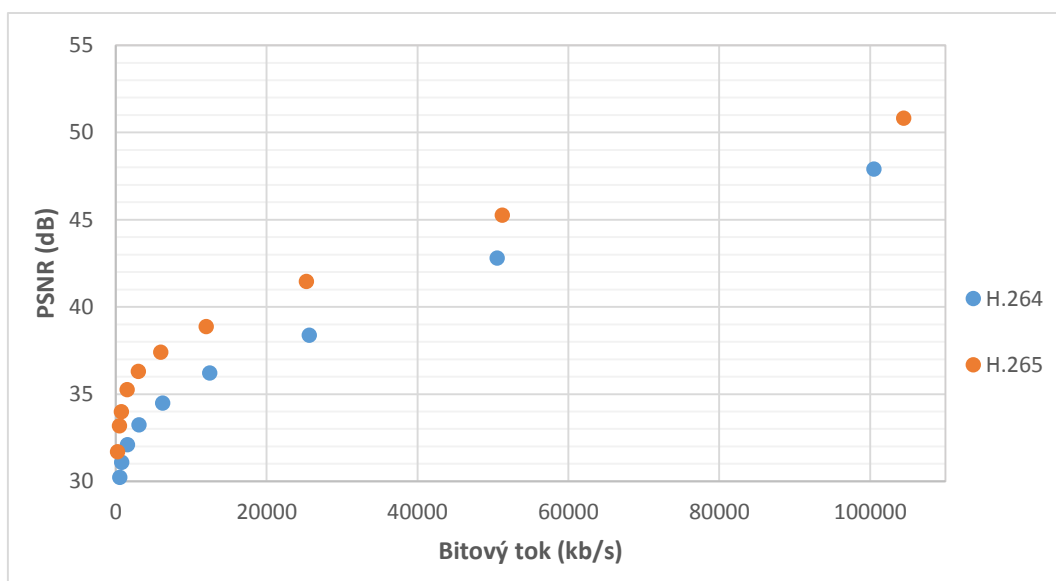
**Graf 4: Rozložení dominance oka u pozorovatelů**



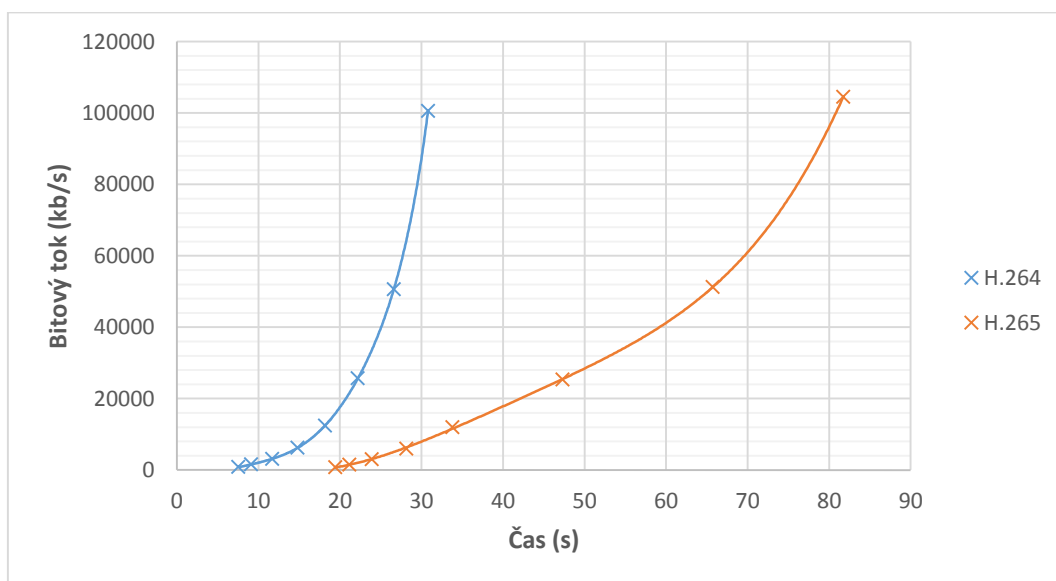
## 10. Výsledky a diskuze

### 10.1 Výsledky testu funkčnosti kodéru HEVC

Příprava videa pro porovnání formátů H.264/AVC a H.265/HEVC je popsána v **kapitole 10.1 Funkčnost kodéru HEVC**. Porovnání je graficky znázorněno na **grafu 5** a **grafu 6**. Z grafů je patrné, že kompresní účinnost H.265 dosahuje až 50% oproti H.264. Při bitovém toku 51265 kb/s pro H.265 byla hodnota metriky PSNR 45,2 dB a čas kódování byl 65,7 sekundy. Pro bitový tok 50595 kb/s pro H.264 vyšlo PSNR 42,7 dB s časem pro kódování 26,6 sekundy, což je o 39,1 sekundy kratší čas. Díky tomuto porovnání je vidět vysoká hardwarová náročnost kodeku H.265.



Graf 5: Porovnání metriky PSNR formátu H.264 a H.265 v závislosti na bitovém toku



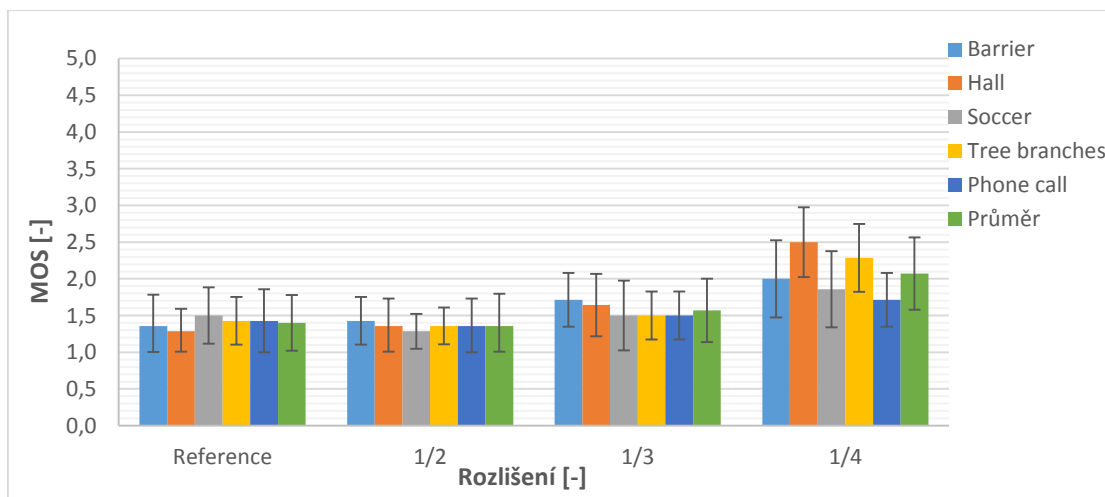
Graf 6: Porovnání časové náročnosti formátů H.264 a H.265 v závislosti na bitovém toku

## 10.2 Výsledky subjektivního testu kvality stereoskopického videa

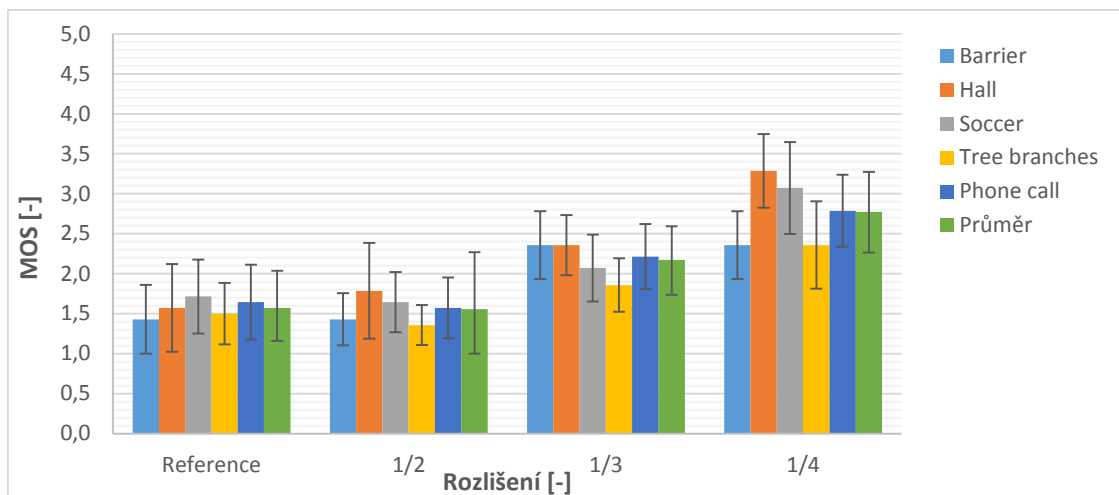
Výsledky subjektivních testů jsou vyneseny do **grafů 7-18**. Grafy jsou rozděleny do kategorií podle použitého parametru Qp a podle hodnotícího parametru tedy hloubky, ostrosti a celkového vizuálního komfortu. Ve všech grafech je zobrazen vliv nesymetrického kódování tedy parametru změny rozlišení ( $\frac{1}{2}$ ,  $\frac{1}{3}$  a  $\frac{1}{4}$ ). V grafech jsou zobrazeny hodnoty průměru, který odpovídá hodnotě průměru všech pěti testovaných videosekvencí s příslušnou kompresí a rozlišením.

Z výsledků subjektivního hodnocení kvality stereoskopického videa vyplývá, že vnímání hloubky u symetricky komprimovaných videí s parametrem Qp 20, 30 se významně nemění a zůstává nepozorované až pozorované, ale v nerušící zóně. Až při parametru Qp 40 dochází ke skokovému zhoršení vnímání hloubky. Dále lze pozorovat, že vnímání ostrosti má výrazný vliv na vizuální komfort, neboť rušivá ostrost vede k vizuálnímu nekomfortu. Vnímání ostrosti zůstává stejné pro QP 20, 30. Při hodnotě Qp 40 je míra ostrosti hodnocena jako rušivá.

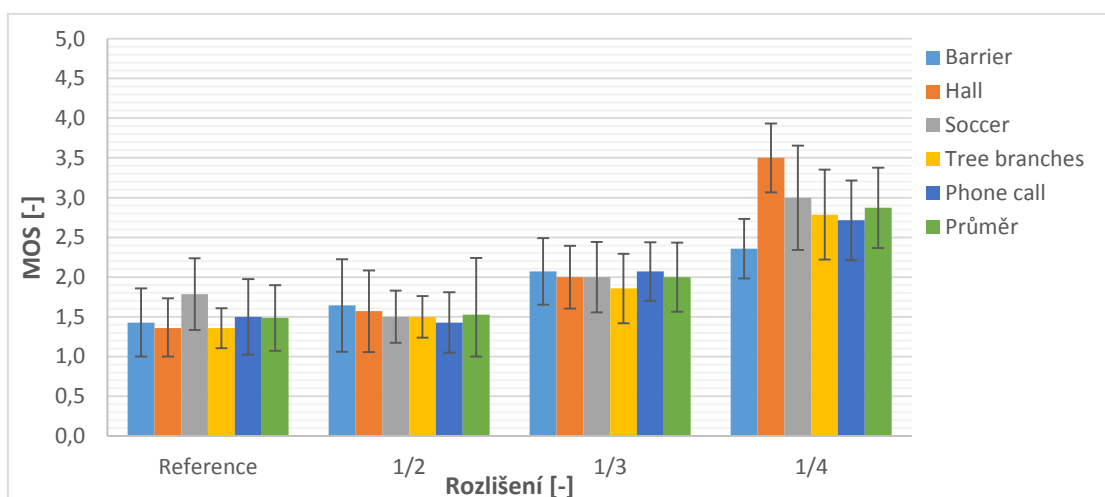
Při nesymetrickém kódování s  $\frac{1}{2}$  rozlišením pro pravý snímek referenčního videa a pro videa s kompresním s parametrem Qp 20 a 30 je vnímání hloubky totožné s referencí. Při  $\frac{1}{3}$  rozlišení pravého snímku u referenčního i při parametrech Qp 20 a 30 je vnímání hloubky jen o trochu horší. Pro obě rozlišení při kompresy Qp 40 je vnímaná hloubka v pásmu lehkého rušení. Vnímání hloubky prostoru pro pravý pohled s rozlišením  $\frac{1}{4}$  je od komprese Qp 30 hodnocena jako lehce rušivá. I u nesymetrického kódování je vidět souvislost mezi vnímanou ostrostí a vizuálním komfortem. Míra negativního vnímání ostrosti, a tudíž i snížení vizuálního komfortu, se lineárně zvyšuje v rozmezí Qp 20 a 30. Stejný jev lze pozorovat i u zmenšování rozlišení pravého snímku stereoskopického videa. Avšak i při plném rozlišení videa komprimovaného parametrem Qp 40 se ostrost i vizuální komfort jeví rušivě.



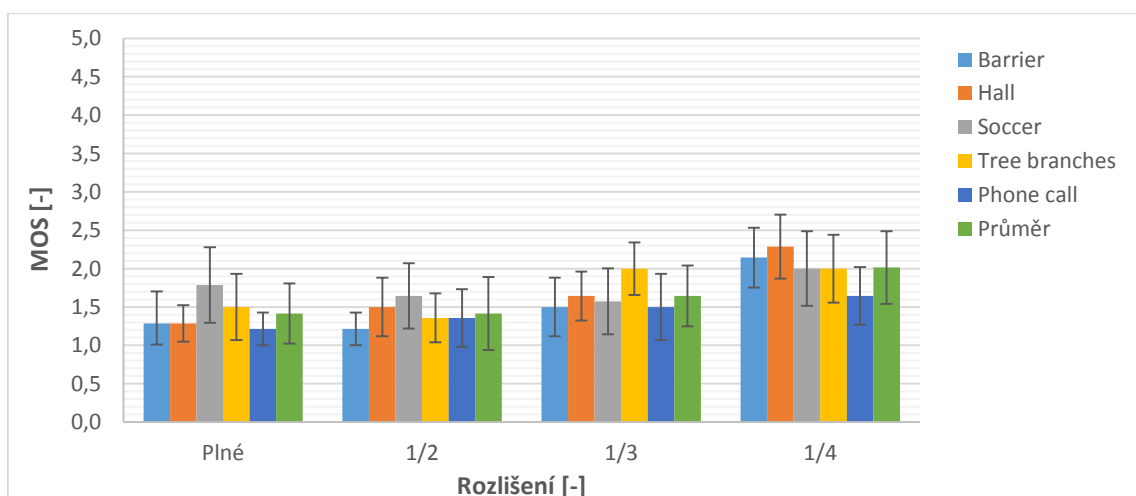
**Graf 7:** Subjektivní hodnocení celkového vjemu hloubky pro referenční snímek v závislosti na rozlišení pravého pohledu



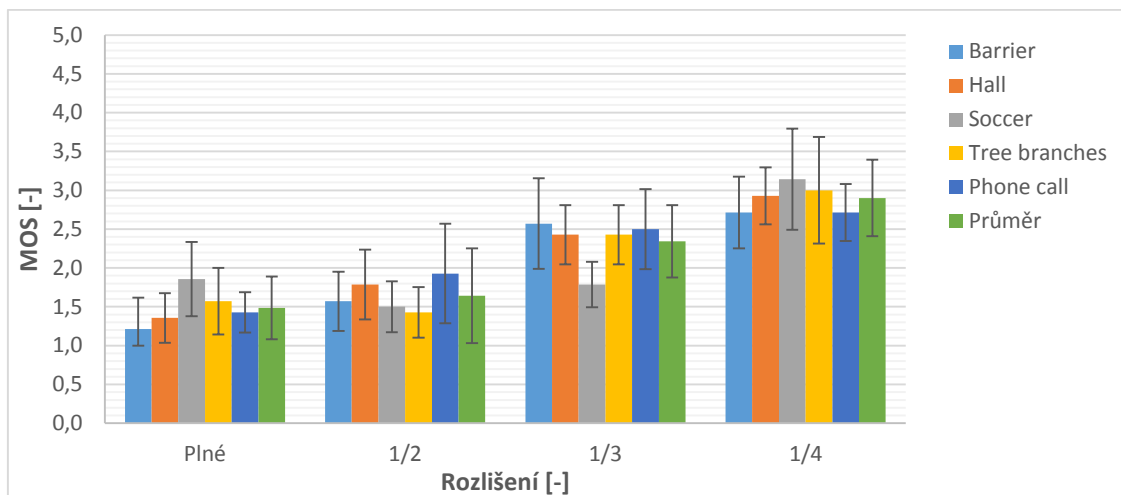
**Graf 8:** Subjektivní hodnocení celkové ostrosti pro referenční snímek v závislosti na rozlišení pravého pohledu



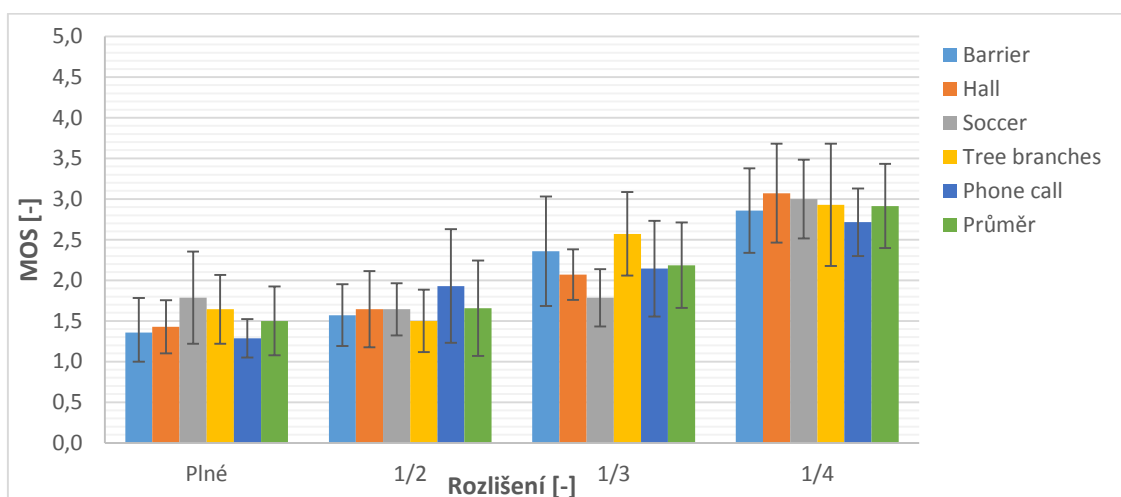
**Graf 9:** Subjektivní hodnocení celkového vizuálního komfortu pro referenční snímek v závislosti na rozlišení pravého pohledu



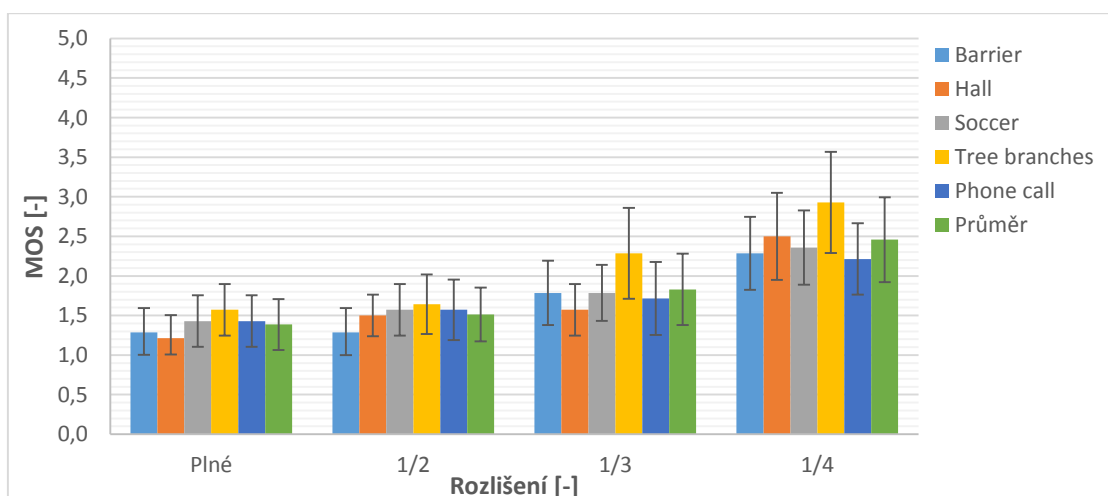
**Graf 10:** Subjektivní hodnocení celkového vjemu hloubky pro kvantizační parametr Qp 20 v závislosti na rozlišení pravého pohledu



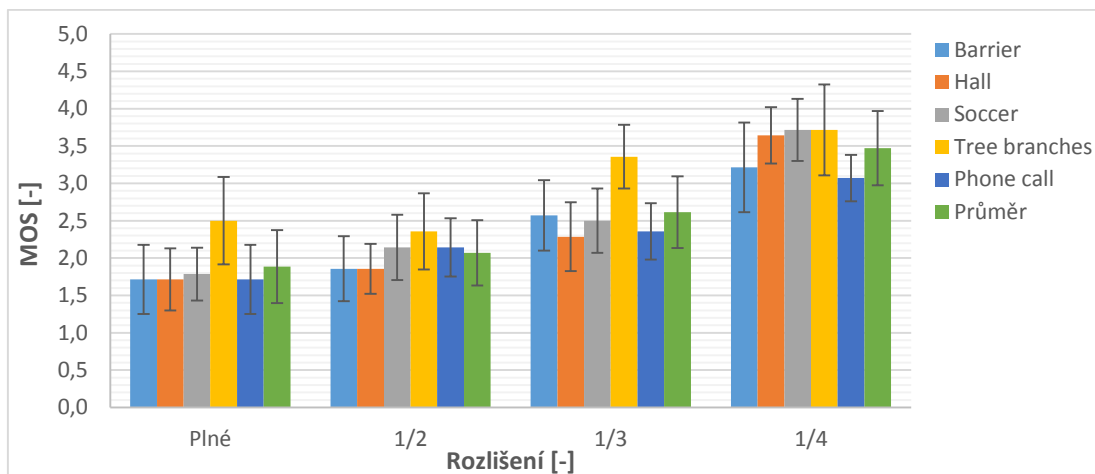
**Graf 11:** Subjektivní hodnocení celkové ostrosti pro konstantní kvantizační parametr  $Q_p$  20 v závislosti na rozlišení pravého pohledu



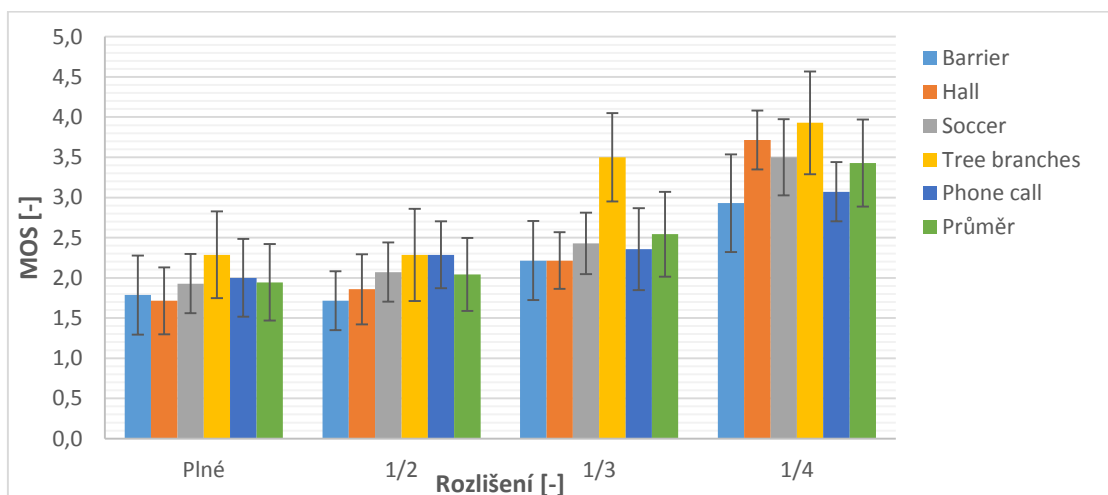
**Graf 12:** Subjektivní hodnocení celkového vizuálního komfortu pro kvantizační parametr  $Q_p$  20 v závislosti na rozlišení pravého pohledu



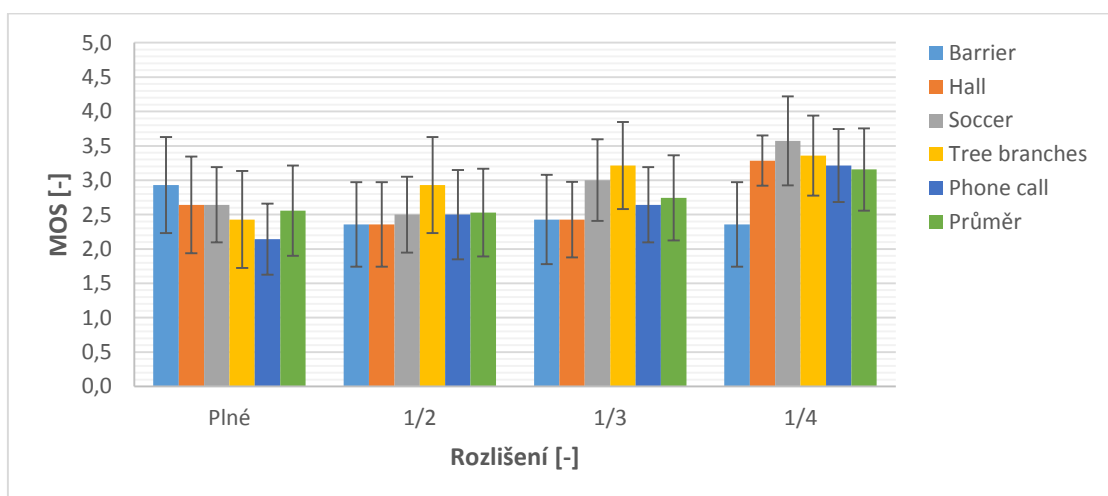
**Graf 13:** Subjektivní hodnocení celkového vjemu hloubky pro kvantizační parametr  $Q_p$  30 v závislosti na rozlišení pravého pohledu



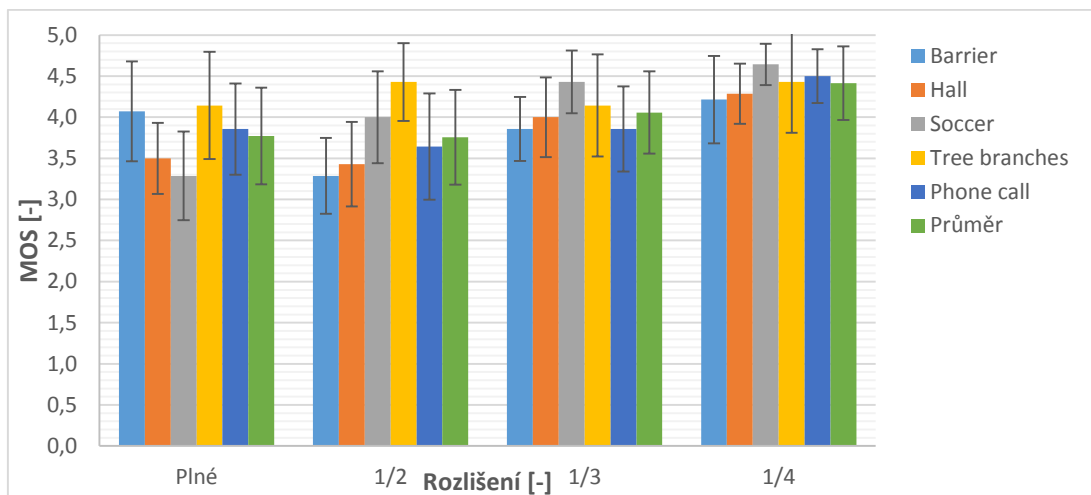
**Graf 14:** Subjektivní hodnocení celkové ostrosti pro kvantizační parametr Qp 30 v závislosti na rozlišení pravého pohledu



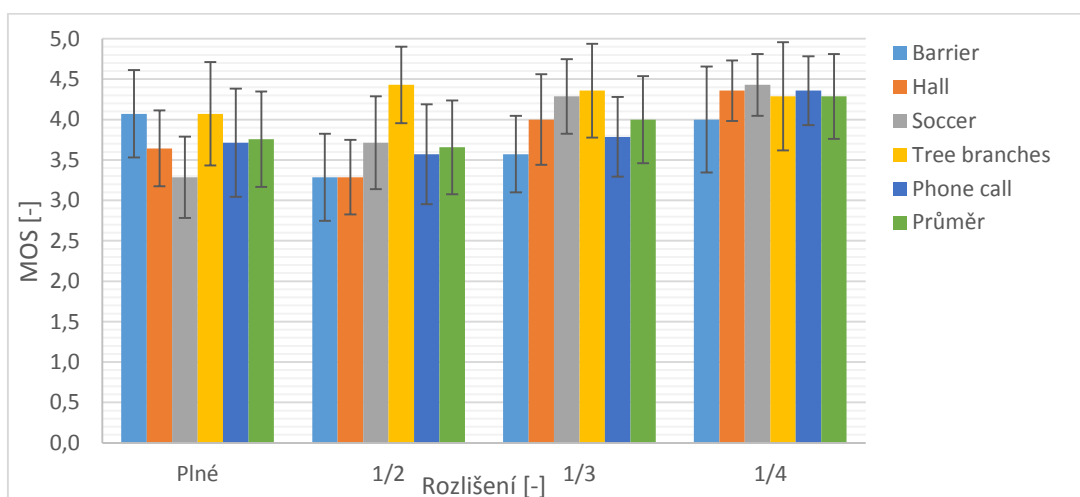
**Graf 15:** Subjektivní hodnocení celkového vizuálního komfortu pro kvantizační parametr Qp 30 v závislosti na rozlišení pravého pohledu



**Graf 16:** Subjektivní hodnocení celkového vjemu hloubky pro kvantizační parametr Qp 40 v závislosti na rozlišení pravého pohledu



**Graf 17:** Subjektivní hodnocení celkové ostrosti pro kvantizační parametr Qp 40 v závislosti na rozlišení pravého pohledu

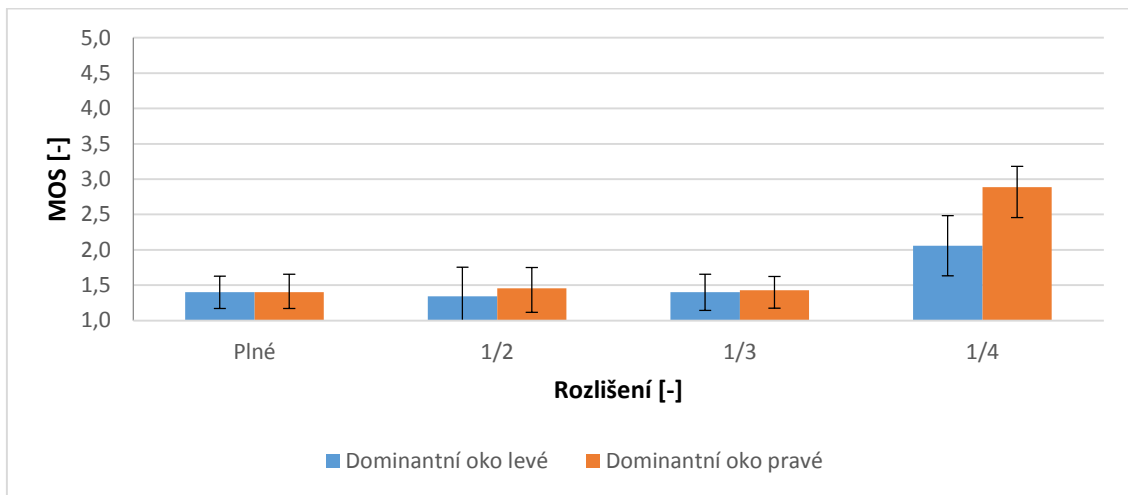


**Graf 18:** Subjektivní hodnocení celkového vizuálního komfortu pro kvantizační parametr Qp 40 v závislosti na rozlišení pravého pohledu

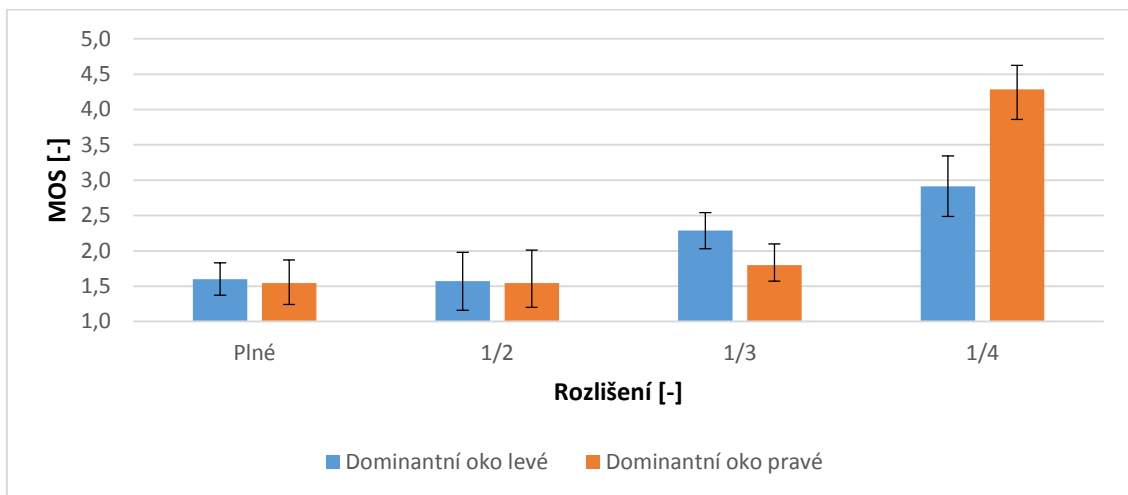
### 10.2.1 Vliv dominance oka na výsledky subjektivního testu

Výsledky subjektivních testů byly rozděleny podle použitého rozlišení v pravém snímku (plné,  $\frac{1}{2}$ ,  $\frac{1}{3}$  a  $\frac{1}{4}$ ) a tyto výsledky byly dále rozděleny podle dominantního oka pozorovatele. Poté byl vypočítán průměr ze všech testovaných videosekvencí.

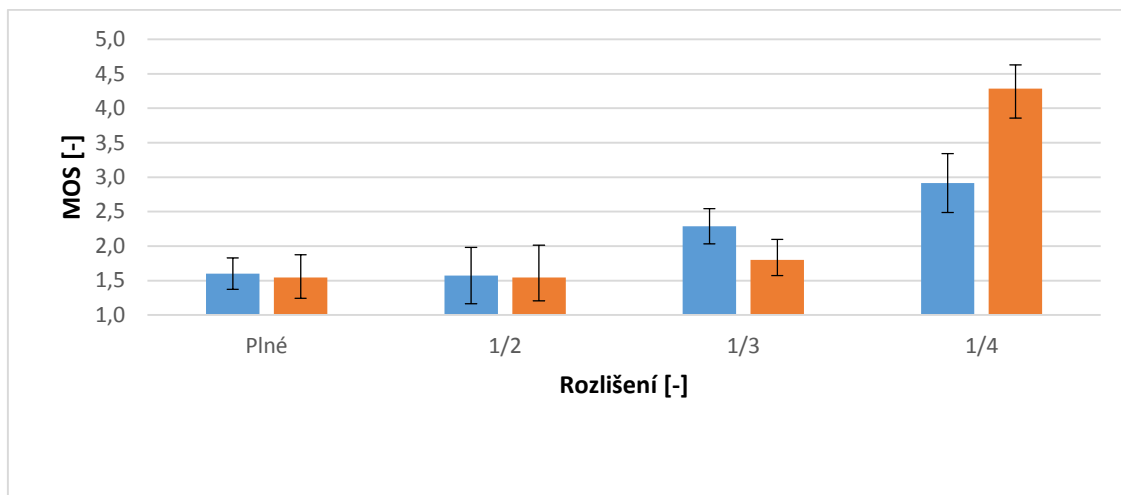
Z grafů 19-21 vyplývá, že při zmenšení rozlišení pravého snímku na  $\frac{1}{4}$  hodnotili pozorovatelé s dominantním pravým okem hloubku, ostrost i vizuální komfort jako více rušivé než pozorovatelé s dominantním levým okem. Proto může být subjektivní hodnocení pro rozlišení  $\frac{1}{4}$  zkreslené.



**Graf 19:** Vliv dominantního oka na vnímání hloubky



**Graf 20:** Vliv dominantního oka na vnímání ostrosti

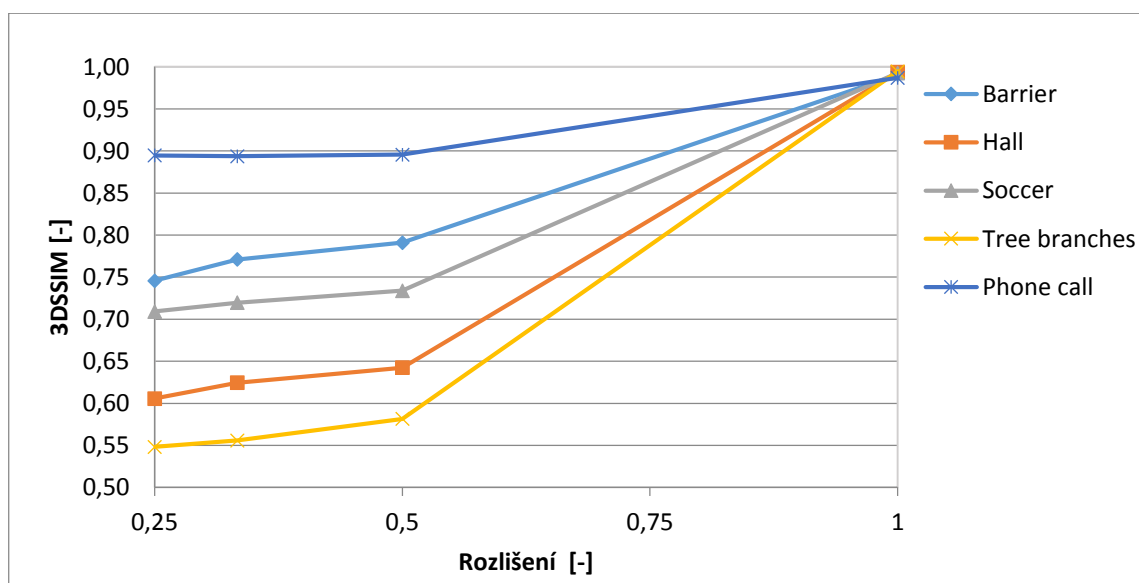


**Graf 21:** Vliv dominantního oka na vizuální komfort

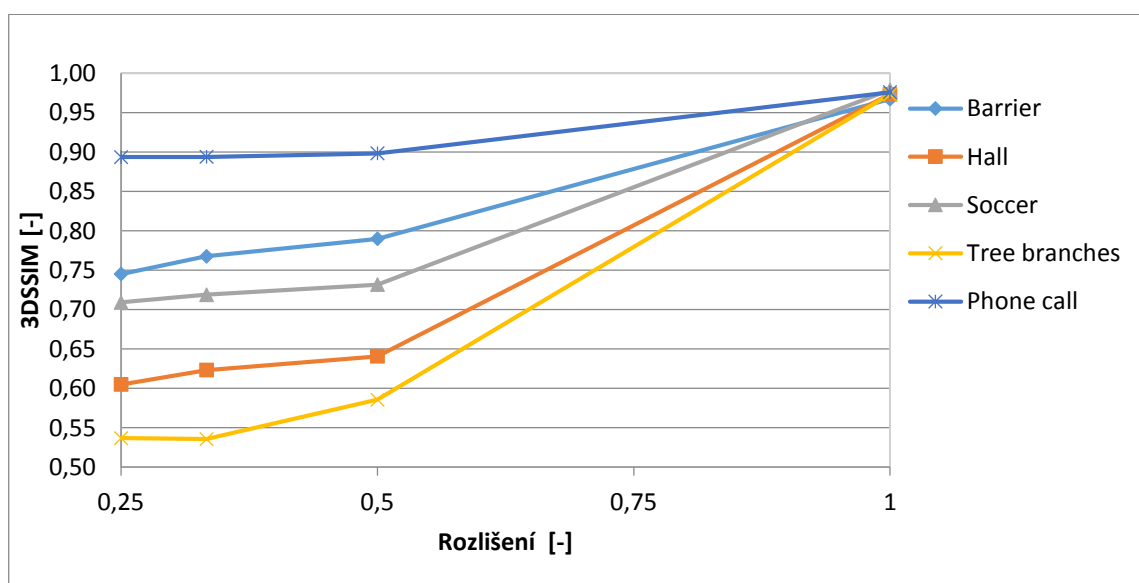


### 10.3 Výsledky objektivního testu stereoskopického videa

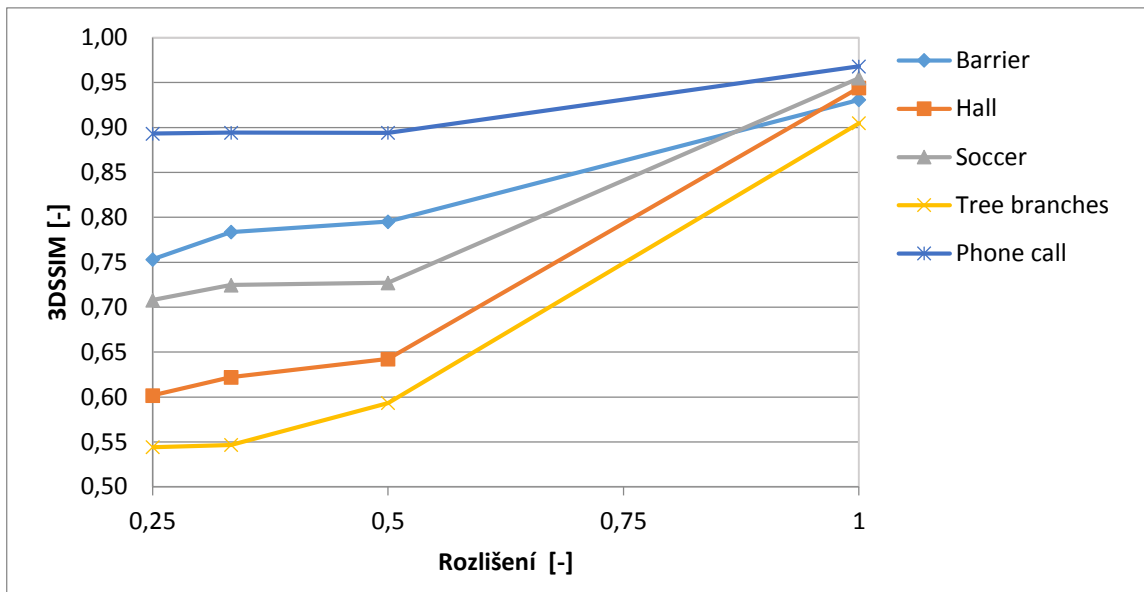
Výsledky objektivního testu (**graf 22-25**) potvrdily předpoklad, že se zmenšováním rozlišení pravého snímku stereopáru bude kvalita stereoskopického videa klesat oproti referenčnímu videu. Snížování kvality bylo předpokládáno i pro zvětšující se kvantizační parametr  $Q_p$ . Snížení hodnot 3DSSIM podle parametru  $Q_p$  je nejvíce patrné u symetricky kódovaných videí. U nesymetricky kódovaných videí je snížení kvality velice malé. Výsledné hodnoty 3DSISIM jsou různé v závislosti na obsahu scény a pohybu v použitých videích. Nejhorších výsledných hodnot dosahuje video Tree branches, které obsahuje velké množství detailů i pohybu. Nejlepších výsledných hodnot dosahuje video s názvem Phone call, protože video obsahuje pomalý pohyb a pouze velice malou proměnu scény. U všech kompresí tohoto videa pozorujeme snížení kvality při zmenšení rozlišení pravého snímku na polovinu. Při dalším snižování rozlišení se už hodnoty kvality nemění. Příčinou by mohla být právě malá změna scény.



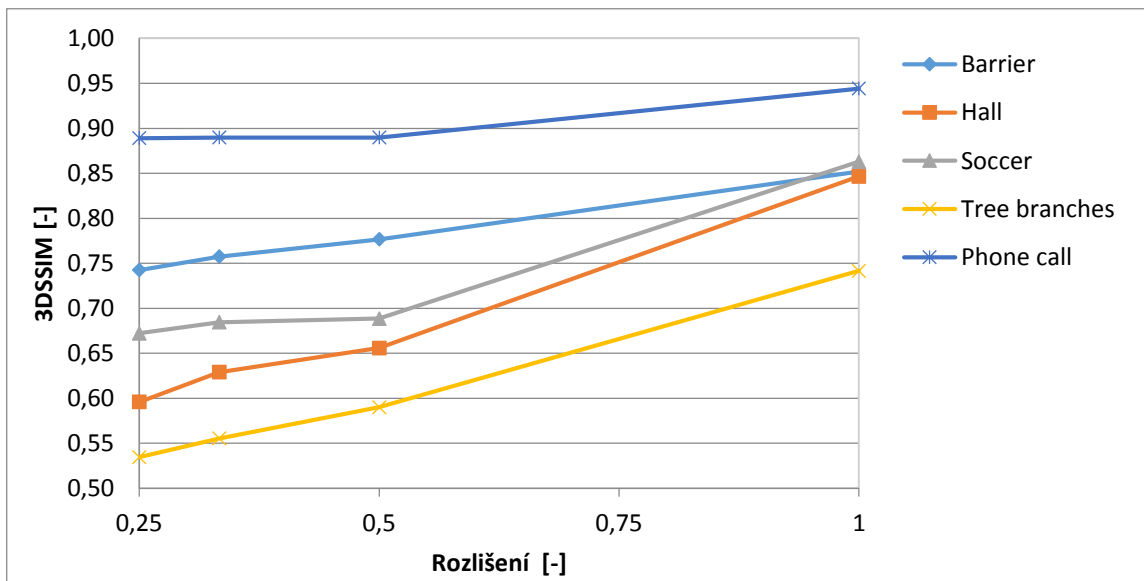
**Graf 22:** Výpočet metriky 3DSSIM referenčních videí v závislosti na rozlišení pravého snímku



**Graf 23:** Výpočet metriky 3DSSIM v závislosti na rozlišení pravého snímku s vlivem komprese  $Q_p$  20



**Graf 24:** Výpočet metriky 3DSSIM v závislosti na rozlišení pravého snímku s vlivem komprese Qp 30

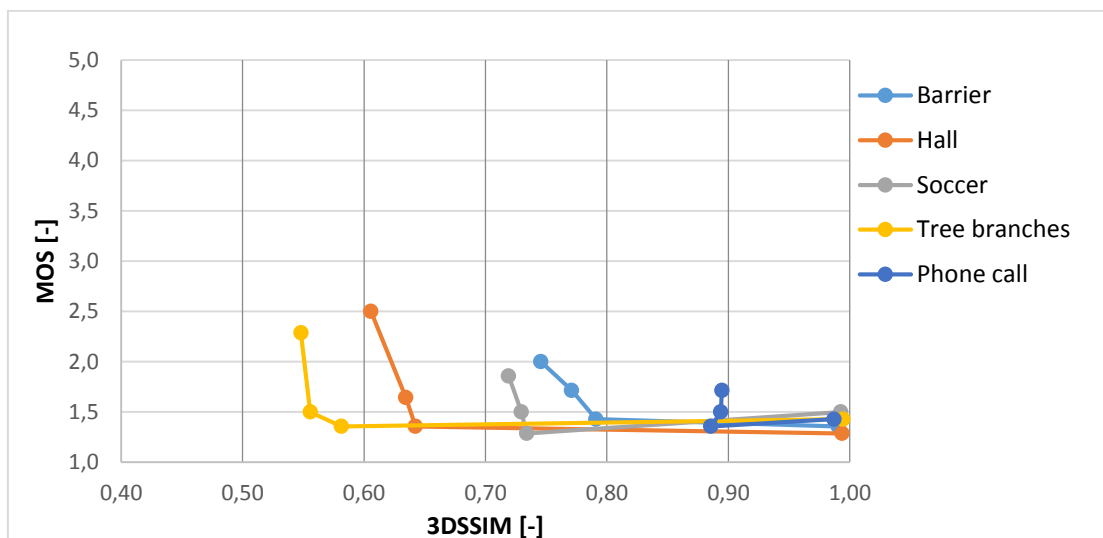


**Graf 25:** Výpočet metriky 3DSSIM v závislosti na rozlišení pravého snímku s vlivem komprese Qp 40

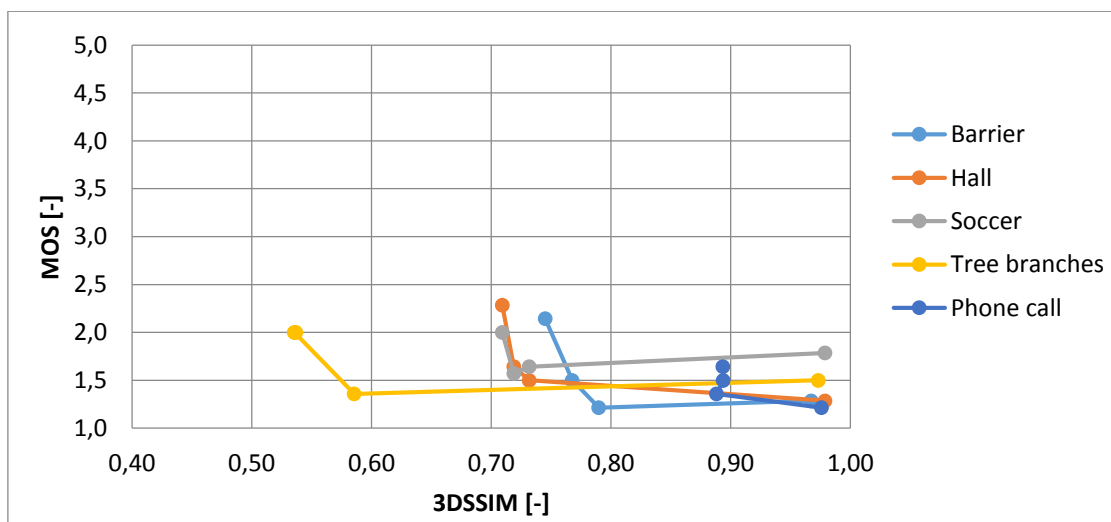
#### 10.4 Porovnání výsledků subjektivního a objektivního testu

Aby bylo možné porovnávat data ze subjektivního a objektivního testu, je důležité získat informace o jejich vzájemném vztahu. Nejpřehlednější metodou je vynesení výsledků subjektivního (MOS) i objektivního testu (3DSSIM) do grafů.

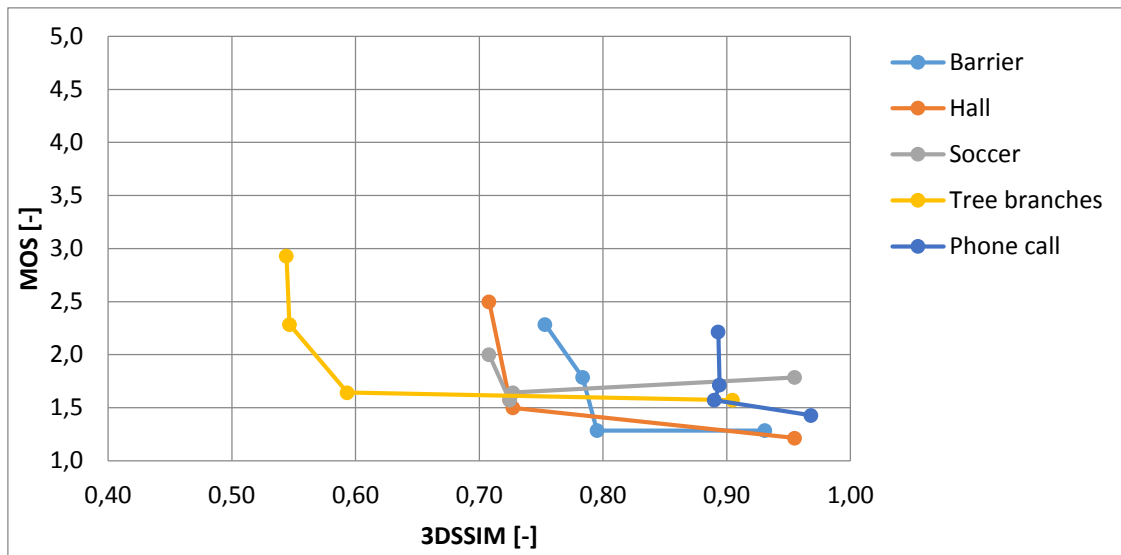
Z uvedených **grafů 26-37** vychází, že objektivní metrika 3DSSIM udává skokové zhoršení kvality videa pro zmenšení rozlišení pravého snímku na polovinu. Oproti tomu v subjektivních testech vychází parametr MOS do čísla 2, což znamená, že není vnímáno rušení v obraze. Při dalším zmenšování rozlišení pravého snímku dochází k mírným změnám metriky 3DSSIM. Pro videa komprimovaná parametrem Qp 40 metrika 3DSSIM nekoreluje se subjektivními testy.



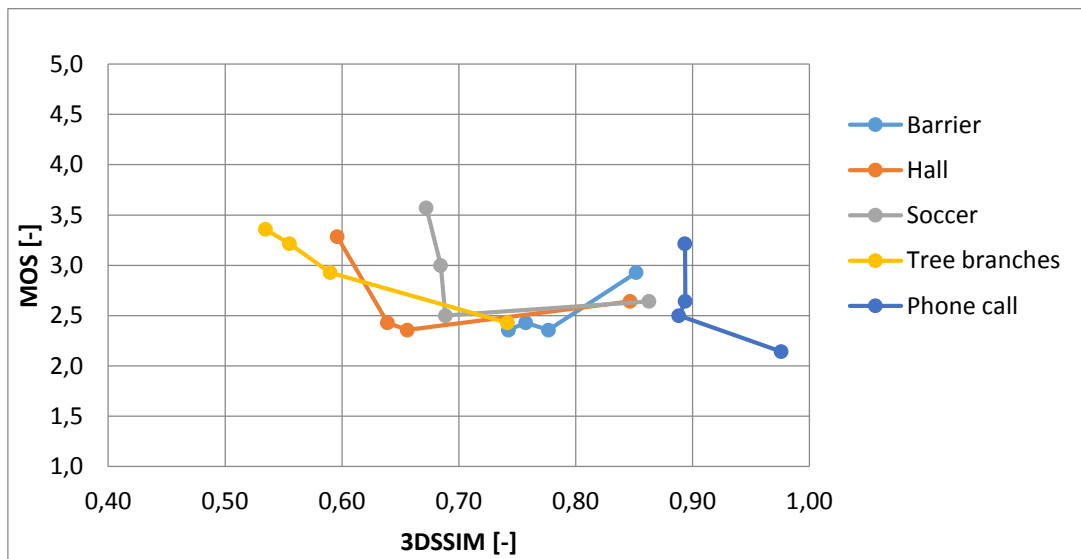
**Graf 26:** Celková hloubka - porovnání objektivních a subjektivních testů pro upravená referenční videa



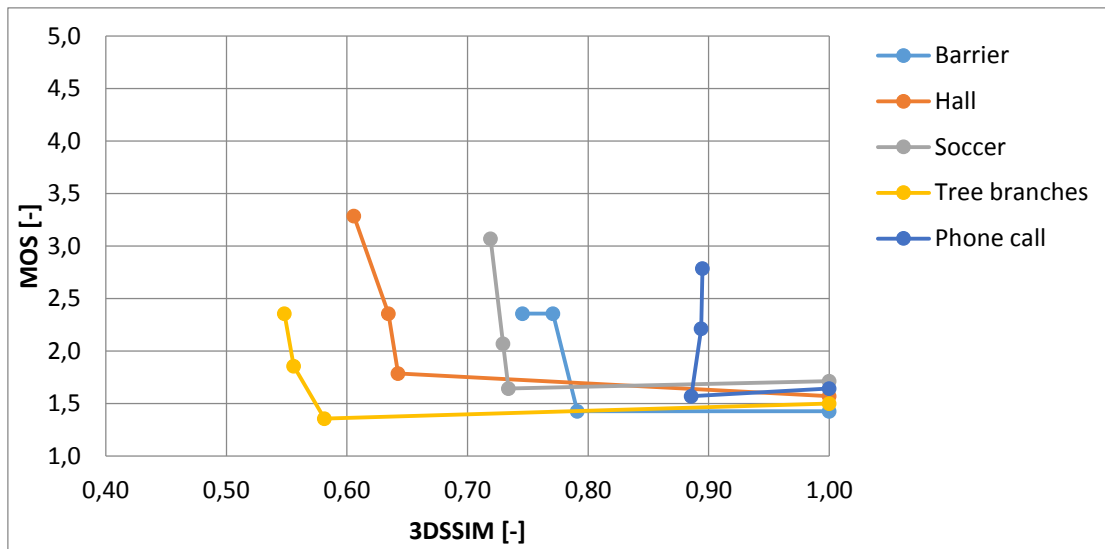
**Graf 27:** Celková hloubka - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 20



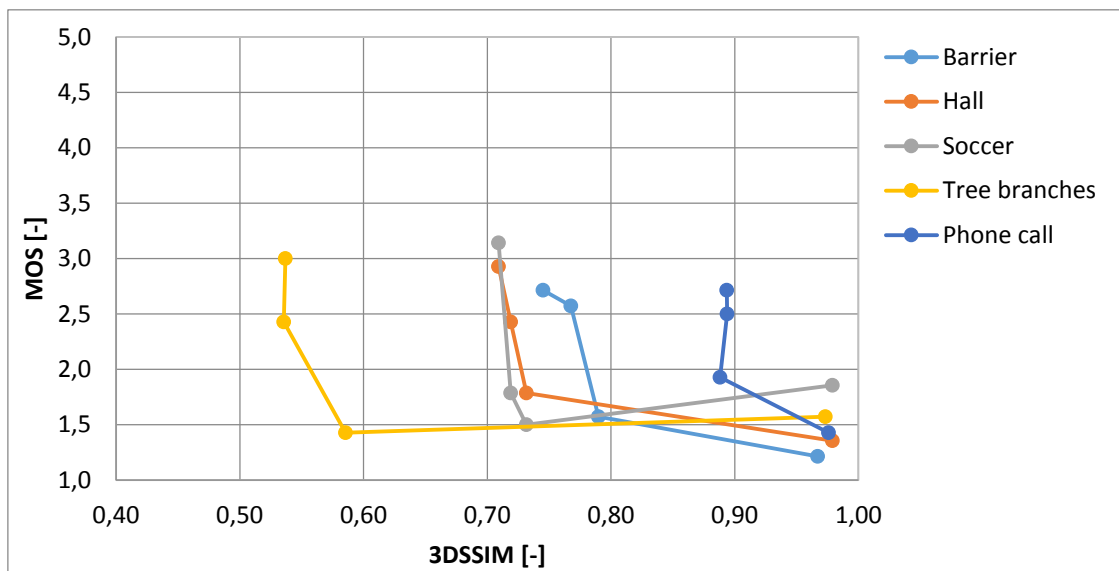
**Graf 28:** Celková hloubka - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 30



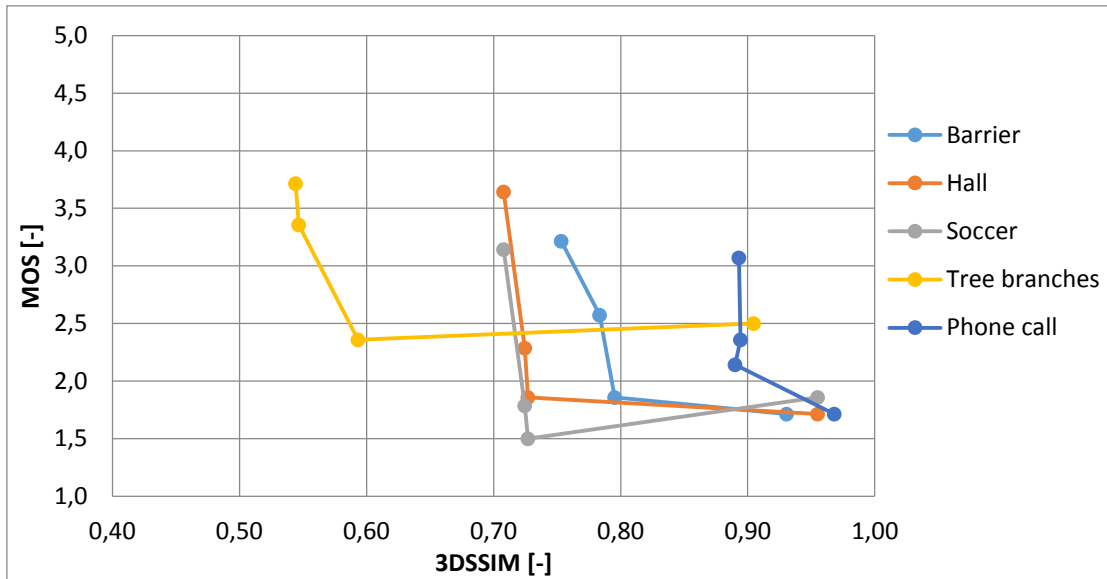
**Graf 29:** Celková hloubka - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 40



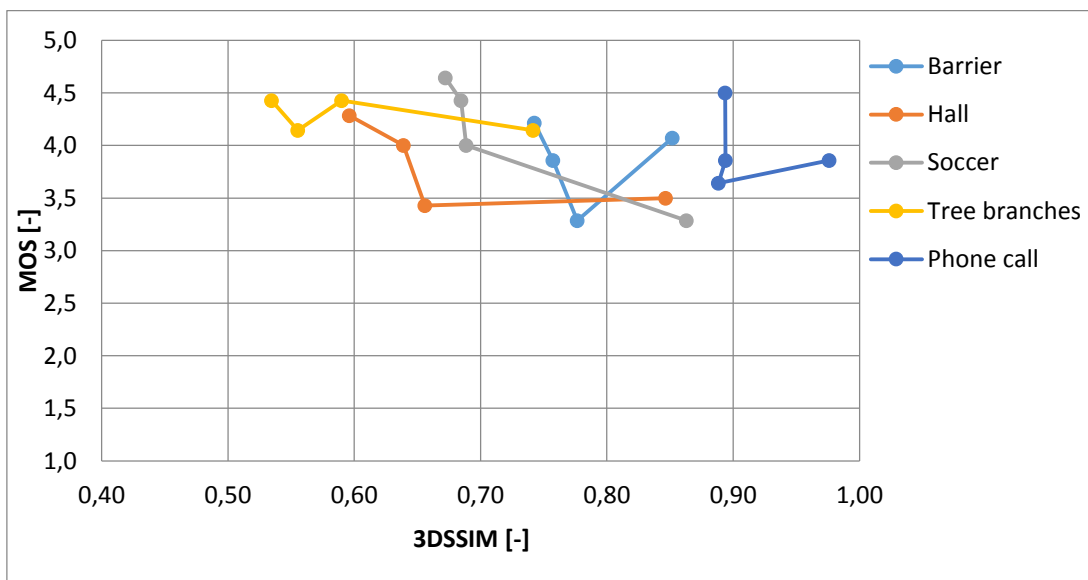
**Graf 30:** Celková ostrost - porovnání objektivních a subjektivních testů pro upravená referenční videa



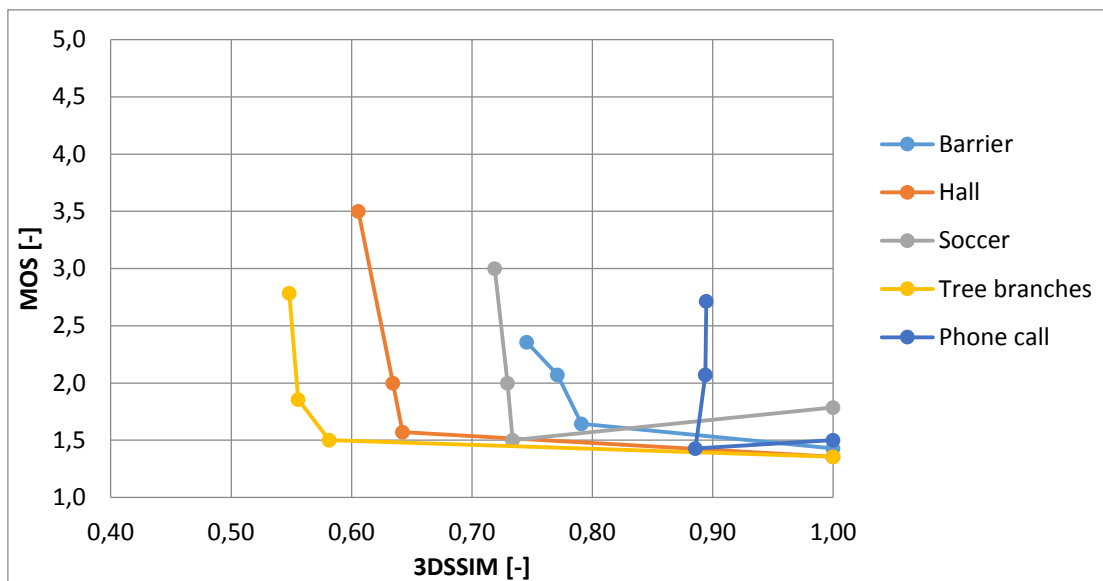
**Graf 31:** Celková ostrost - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 20



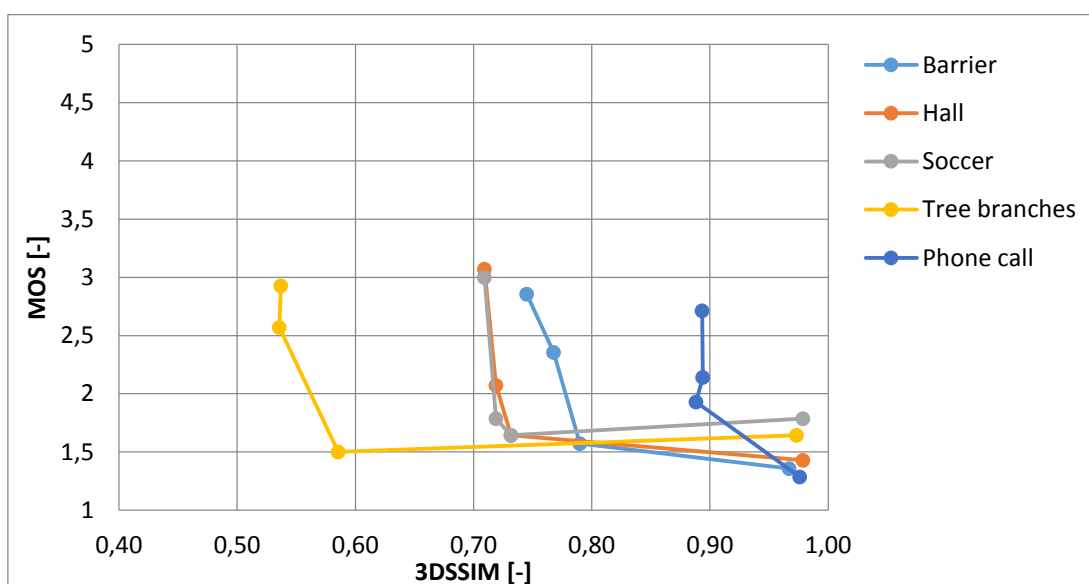
**Graf 32:** Celková ostrost - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 30



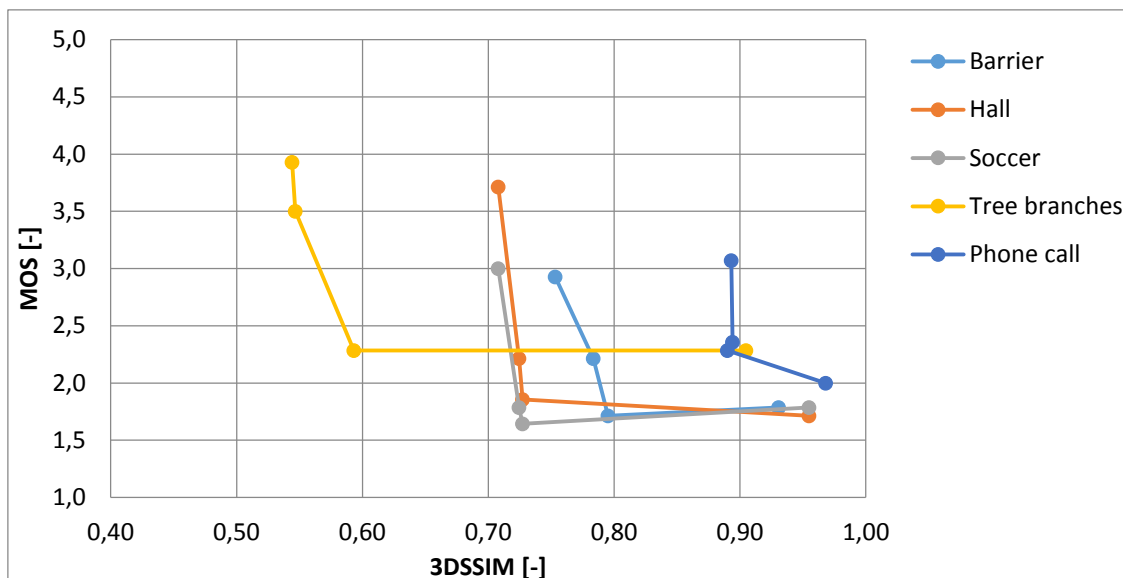
**Graf 33:** Celková ostrost - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 40



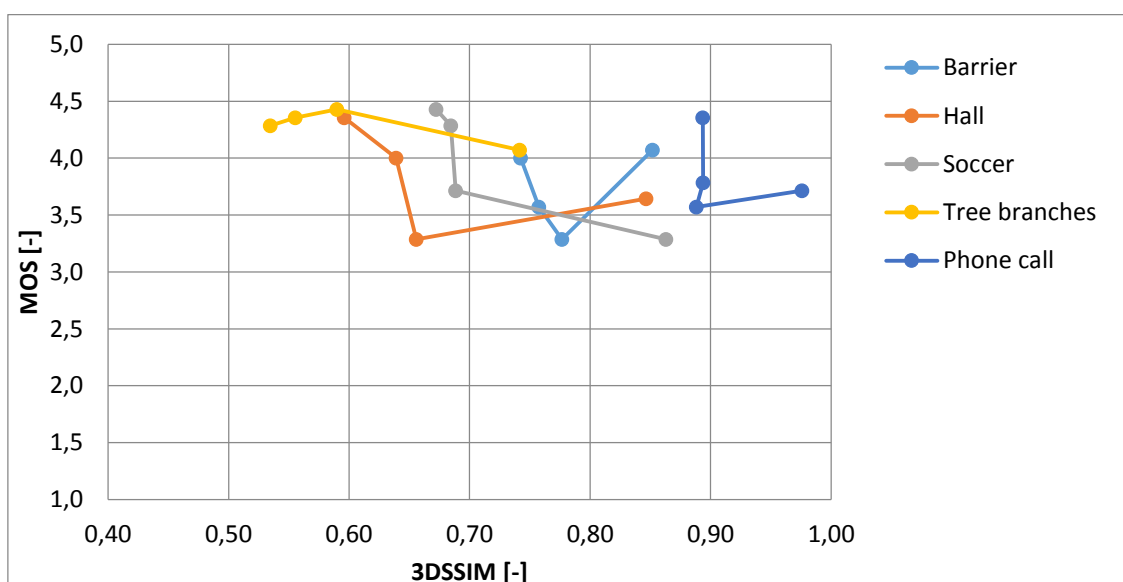
**Graf 34:** Celkový vizuální komfort - porovnání objektivních a subjektivních testů pro upravená referenční videa



**Graf 35:** Celková vizuální komfort - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 20



**Graf 36:** Celkový vizuální komfort - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 30

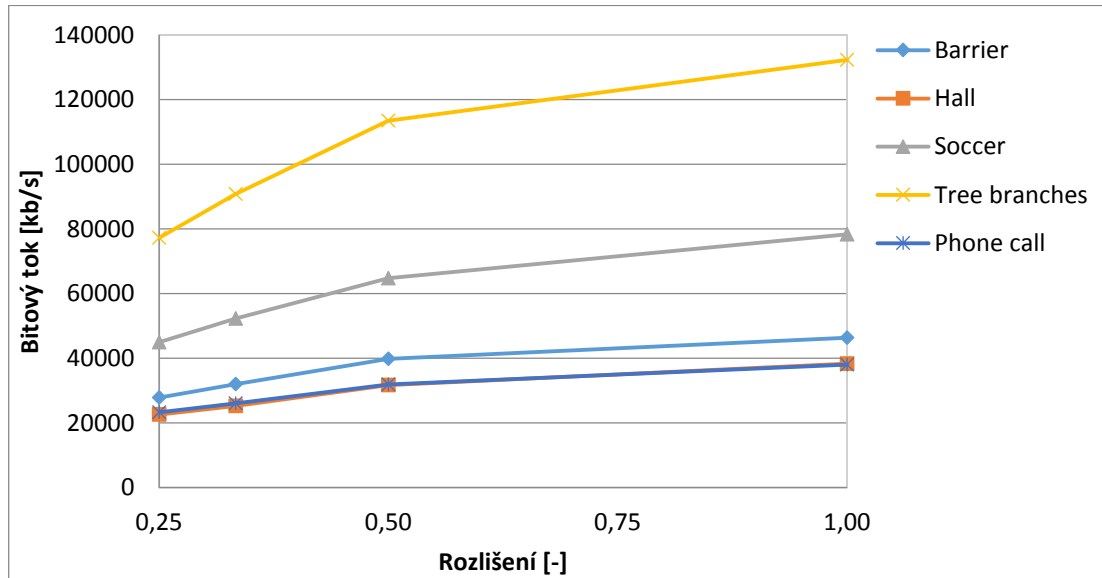


**Graf 37:** Celkový vizuální komfort - porovnání objektivních a subjektivních testů pro videa s kompresí Qp 40

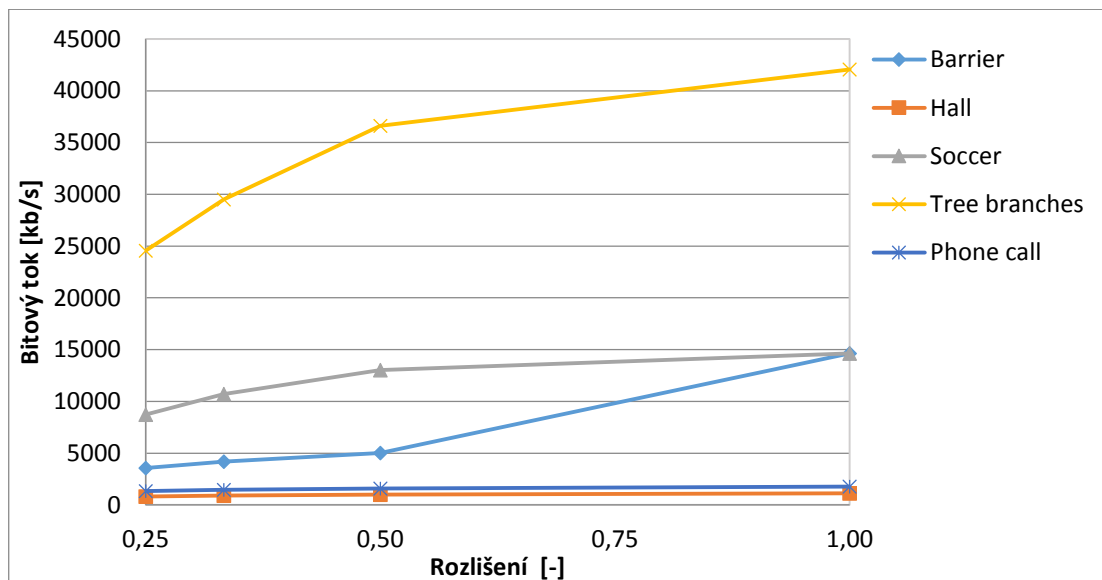


## 10.5 Úspora bitového toku

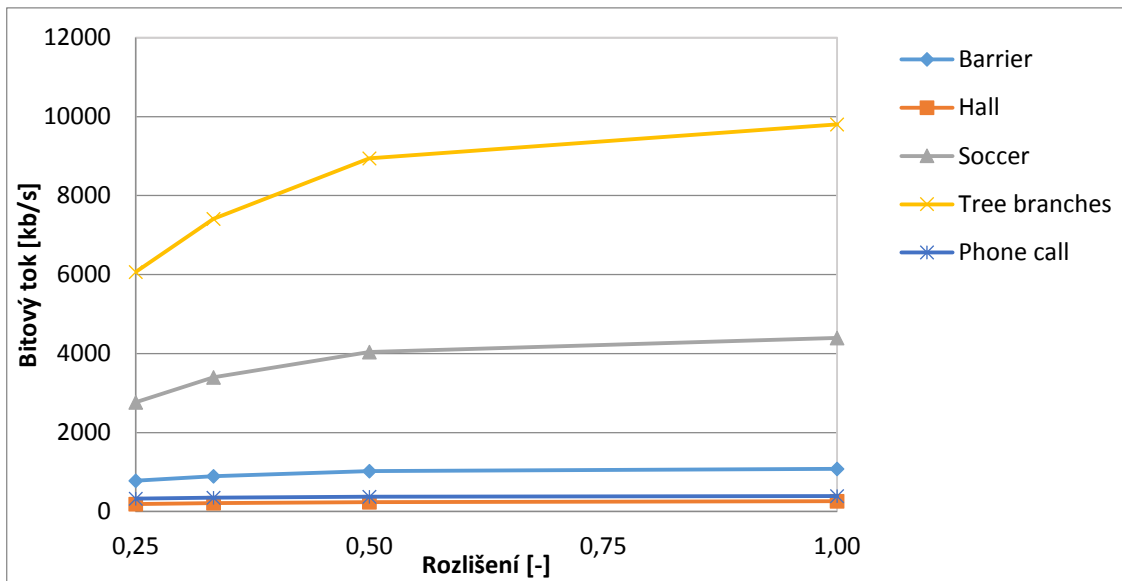
Kompresce videa se provádí za účelem snížení bitového toku, zároveň je však nutné zachovat kvalitu videa na co nejvyšší úrovni. Proto bylo provedeno porovnání velikosti bitového toku upravených videí s výslednými hodnotami subjektivního testu (**graf 38-41**). Následně byla určena varianta s nejnižším bitovým tokem, při němž obraz nebyl vnímán rušivě, což odpovídá hodnotě MOS 2.



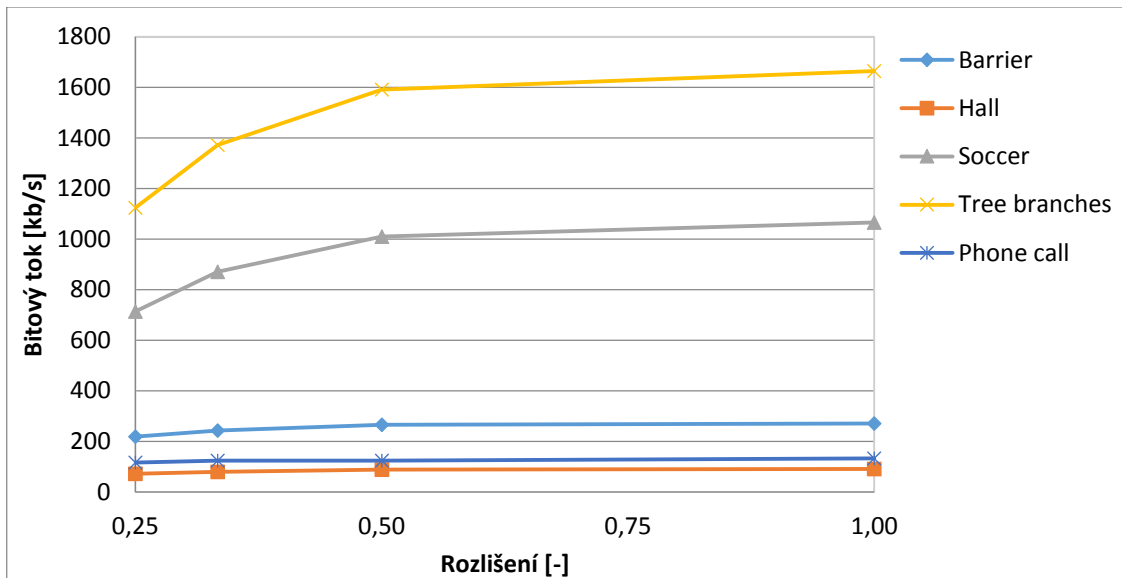
**Graf 38:** Závislost bitového toku na změnu rozlišení pro referenční videa



**Graf 39:** Závislost bitového toku na změnu rozlišení pro kompresy Qp 20



**Graf 40:** Závislost bitového toku na změnu rozlišení pro kompresy Qp 30



**Graf 41:** Závislost bitového toku na změnu rozlišení pro kompresy Qp 40

**Tabulka 2:** Výsledné bitové toky pro referenční videa upravená videa s kompresí Qp 20

Název videa	Komprese	Referenční video		Qp 20	
	Rozlišení pravého snímku	Bitový tok [kb/s]	Úspora [%]	Bitový tok [kb/s]	Úspora [%]
Barrier	Plné	46369	-	14647	68,4
	1/2	39761	14,3	5013	89,2
	1/3	31918	31,2	4169	91,0
	1/4	27805	40,0	3577	92,3
Hall	Plné	38284	-	1118	97,1
	1/2	31714	17,2	994	97,4
	1/3	25280	34,0	891	97,7
	1/4	22523	41,2	796	97,9
Soccer	Plné	78289	-	14647	81,3
	1/2	64772	17,3	13012	83,4
	1/3	52265	33,2	10717	86,3
	1/4	44941	42,6	8737	88,8
Tree branches	Plné	132285	-	42065	68,2
	1/2	113483	14,2	36610	72,3
	1/3	90825	31,3	29502	77,7
	1/4	77213	41,6	24556	81,4
Phone call	Plné	38008	-	1777	95,3
	1/2	31864	16,2	1599	95,0
	1/3	26070	31,4	1461	94,4
	1/4	23206	38,9	1335	94,2

**Tabulka 3:** Výsledné bitové toky pro upravená videa s kompresí Qp 30 a Qp 40

Název videa	Komprese	Qp 30		Qp 40	
	Rozlišení pravého snímku	Bitový tok [kb/s]	Úspora [%]	Bitový tok [kb/s]	Úspora [%]
Barrier	Plné	1082	97,7	270	99,4
	1/2	1027	97,8	265	99,4
	1/3	896	98,1	243	99,5
	1/4	779	98,3	219	99,5
Hall	Plné	265	99,3	91	99,8
	1/2	240	99,4	88	99,8
	1/3	216	99,4	79	99,8
	1/4	188	99,5	72	99,8
Soccer	Plné	4398	94,4	1065	98,6
	1/2	4041	94,8	1010	98,7
	1/3	3400	95,7	870	98,9
	1/4	2769	96,5	714	99,1
Tree branches	Plné	9802	92,6	1664	98,7
	1/2	8941	93,2	1591	98,8
	1/3	7413	94,4	1372	99,0
	1/4	6062	95,4	1124	99,2
Phone call	Plné	389	99,0	132	99,7
	1/2	374	99,0	124	99,7
	1/3	348	99,1	124	99,7
	1/4	324	99,1	116	99,7

V **tabulce 2 a 3** jsou zaznamenány bitové toky upravených videí a jejich datová úspora vůči referenčnímu videu. Z hodnot vyplývá, že pouhé snížení rozlišení na  $\frac{1}{2}$  původního snímku jednoho z pohledů dokáže zmenšit bitový tok přibližně o 15 %, přibližně o 32 % pro zmenšené rozlišení zmenšené na  $\frac{1}{3}$  původního snímku a přibližně o 40 % pro zmenšené rozlišení na  $\frac{1}{4}$  původního snímku.

Větších datových úspor je dosaženo při kompresy snímků kvantizačním parametrem Qp. Úspora bitového toku při kompresy závisí na obsahu videa. Například při kompresy videa Phone call kvantizačním koeficientem Qp 20 dochází ke snížení bitového toku o 95,3 %. Tato velká redukce dat je dána malou mírou pohybu ve videu a celistvým pozadím. Při kompresy kvantizačním koeficientem Qp 20 videa s názvem Tree branches dochází ke zmenšení bitového toku o 68,2 % z důvodu velkého množství pohybu zachyceného na videu.

Již při kompresy parametrem Qp 30 a 40 symetricky kódovaných videí je snížení bitového toku přes 90%. Přidávat proto nesymetrické kódování nemá podstatný význam pro dosažení datové úspory.

## 11. Závěr

Diplomová práce měla za cíl implementaci kodéru H.265/HEVC a ověření jeho funkčnosti. Vybrané 3D videosekvence byly kódovány podle zvolených parametrů a výsledné videosekvence byly následně vyhodnoceny objektivní kvalitativní metrikou. Získané výsledky byly porovnány s údaji, které poskytl zjednodušený subjektivní test.

Implementace kodéru H.265/HEVC byla provedena za pomoci softwaru ffmpeg a použití knihovny libx265. Kodér H.265/HEVC byl porovnán s jeho předchůdcem H.264/AVC. Porovnáním byla potvrzena skutečnost, že kodér H.265/HEVC dokáže snížit bitový tok o 50 % oproti kodéru H.264/AVC a to při zachování stejné vizuální kvality. Porovnání obou kodérů zahrnovalo i měření času kódování. S rostoucím bitovým tokem se čas kódování kodéru H.265/HEVC zvětšuje. Vyšší účinnost kodéru H.265/HEVC dosažená díky inovativním prvkům, je však spojena s většími nároky na hardware.

Dále byl software ffmpeg použit pro testovací sadu videosekvencí použitých pro subjektivní test hodnocení kvality stereoskopického videa. Z výsledků subjektivního testu vyplynulo, že při odstranění velkého množství informace vlivem použití nesymetrického kódování jen pro jeden snímek stereopáru může dojít k zhoršení vnímané kvality v případě, že se nesymetricky kódovaný snímek shoduje s dominantním okem pozorovatele. V provedeném subjektivním testu k tomu došlo při snížení rozlišení pravého snímku na  $\frac{1}{4}$  původního rozlišení.

Z výsledků subjektivního testu pro nesymetrické kódování se jako nejlepší kombinace parametrů jeví zmenšení rozlišení snímku na polovinu a použití komprese s kvantizačním parametrem Qp 20. Při tomto kódování nedochází k rušivým změnám v obraze, zároveň je bitový tok dostatečně snížen (v průměru o 87,4 %). Při použití komprese s kvantizačním parametrem Qp 30 a 40 na celý stereopár dojde ke snížení bitového toku o více jak 90 %, avšak míra ostrosti a vizuálního komfortu je v tomto případě již na značně rušivé úrovni.

Z výsledků objektivního testu stereoskopického videa vyplývá, že lze metodu použít pro zhodnocení kvality nesymetricky kódovaného stereoskopického videa při kompresi s kvantizačním parametrem Qp 20 a 30. Za těchto podmínek hodnoty objektivního testu korelují s výsledky subjektivního testu. Při použité kompresi s kvantizačním parametrem Qp 40 však již hodnoty objektivního a subjektivního testu nekorelují.

## 12. Použitá literatura

- [1] HANUS, Rostislav. *Binokulární vidění*. Brno, 2010. Bakalářská práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií.
- [2] RANDULOVÁ, Jana. *Akomodace a konvergence*. Brno, 2008. Bakalářská práce. Masarykova univerzita v Brně. Lékařská fakulta.
- [3] LIPTON, Lenny. *StereoGraphics Developers' Handbook* [online]. s. 7-13 [cit. 2012-01-12]. Dostupné z: <http://iglasses.weirdoz.org/docs/stgrHandbook.pdf>
- [4] TEITTINEN, Marko. *Depth Cues in the Human Visual System* [online]. In: . [cit. 2016-05-13]. Dostupné z: [http://www.hitl.washington.edu/projects/knowledge\\_base/virtual-worlds/EVE/III.A.1.c.DepthCues.html](http://www.hitl.washington.edu/projects/knowledge_base/virtual-worlds/EVE/III.A.1.c.DepthCues.html)
- [5] TAM, W J, F SPERANZA, S YANO, K SHIMONO a H ONO. Stereoscopic 3D-TV: Visual Comfort. *IEEE Transactions on Broadcasting* [online]. 2011, **57**(2), 335-346 [cit. 2016-05-14]. DOI: 10.1109/TBC.2011.2125070. ISSN 0018-9316. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5743036>
- [6] BOEV, Atanas, Danilo HOLLOSI a Atanas GOTCHEV. *Classification of stereoscopic artefacts*[online]. [cit. 2016-05-14]. Dostupné z: [http://sp.cs.tut.fi/mobile3dtv/results/tech/D5.1\\_Mobile3DTV\\_v1.0.pdf](http://sp.cs.tut.fi/mobile3dtv/results/tech/D5.1_Mobile3DTV_v1.0.pdf)
- [7] SHIBATA, T., J. KIM, D. M. HOFFMAN a M. S. BANKS. The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of Vision* [online]. 2011, **11**(8), 11-11 [cit. 2016-05-14]. DOI: 10.1167/11.8.11. ISSN 1534-7362. Dostupné z: <http://jov.arvojournals.org/Article.aspx?doi=10.1167/11.8.11>
- [8] LÓPEZ, Juan, Juan RODRIGO, David JIMÉNEZ a José MENÉNDEZ. Stereoscopic 3D video quality assessment based on depth maps and video motion. *EURASIP Journal on Image and Video Processing* [online]. 2013, **2013**(1), 62- [cit. 2016-05-14]. DOI: 10.1186/1687-5281-2013-62. ISSN 1687-5281. Dostupné z: <http://jivp.urasipjournals.com/content/2013/1/62>
- [9] POURAZAD, Mahsa, Colin DOUTRE, Maryam AZIMI a Panos NASIOPOULOS. HEVC: The New Gold Standard for Video Compression. *IEEE Consumer Electronics Magazine* [online]. 2012, **1**(3), 36-46 [cit. 2016-05-17]. DOI: 10.1109/MCE.2012.2192754. ISSN 2162-2248. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6222536>
- [10] SALEH, Mohammed A., Habibah HASHIM, Nooritawati Md. TAHIR a Ezril HISHAM. Review for High Efficiency Video Coding (HEVC). *2014 IEEE Conference on Systems, Process and Control (ICSPC 2014)* [online]. IEEE, 2014, , 141-146 [cit. 2016-05-17]. DOI: 10.1109/SPC.2014.7086246. ISBN 978-1-4799-6106-1. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7086246>
- [11] SULLIVAN, Gary J., Jens-Rainer OHM, Woo-Jin HAN a Thomas WIEGAND. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2012, **22**(12), 1649-1668 [cit. 2016-05-17]. DOI: 10.1109/TCSVT.2012.2221191. ISSN 10518215. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6316136>
- [12] Rec. ITU-T H.265. (04/2015), Geneva: ITU, 2015.

- [13] OHM, J. a G. J. SULLIVAN. High efficiency video coding: the next frontier in video compression [Standards in a Nutshell]. *IEEE Signal Processing Magazine* [online]. 2013, **30**(1), 152-158 [cit. 2016-05-17]. DOI: 10.1109/MSP.2012.2219672. ISSN 1053-5888. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6375943>
- [14] MISRA, Kiran, Andrew SEGALL, Michael HOROWITZ, Shilin XU, Arild FULDSETH a Minhua ZHOU. An Overview of Tiles in HEVC. *IEEE Journal of Selected Topics in Signal Processing*[online]. 2013, **7**(6), 969-977 [cit. 2016-05-18]. DOI: 10.1109/JSTSP.2013.2271451. ISSN 1932-4553. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6547985>
- [15] SOLE, Joel, Rajan JOSHI, Nguyen NGUYEN, Tianying JI, Marta KARCZEWICZ, Gordon CLARE, Félix HENRY a Alberto DUENAS. Transform Coefficient Coding in HEVC. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2012, **22**(12), 1765-1777 [cit. 2016-05-18]. DOI: 10.1109/TCSVT.2012.2223055. ISSN 1051-8215. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6324418>
- [16] LUKES, T., K. FLIEGEL a M. KLIMA. Objective image quality assessment of multiframe super-resolution methods. *2013 23rd International Conference Radioelektronika (RADIOELEKTRONIKA)* [online]. IEEE, 2013, , 267-272 [cit. 2016-05-24]. DOI: 10.1109/RadioElek.2013.6530929. ISBN 978-1-4673-5519-3. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6530929>
- [17] WANG, Z., A.C. BOVIK, H.R. SHEIKH a E.P. SIMONCELLI. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* [online]. 2004, **13**(4), 600-612 [cit. 2016-05-24]. DOI: 10.1109/TIP.2003.819861. ISSN 1057-7149. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1284395>
- [18] WANG, Jiheng, Shiqi WANG a Zhou WANG. Quality prediction of asymmetrically compressed stereoscopic videos. *2015 IEEE International Conference on Image Processing (ICIP)* [online]. IEEE, 2015, , 3427-3431 [cit. 2016-05-25]. DOI: 10.1109/ICIP.2015.7351440. ISBN 978-1-4799-8339-1. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=7351440>
- [19] AFLAKI, Payman, Miska M. HANNUKSELA a Moncef GABBOUJ. Subjective quality assessment of asymmetric stereoscopic 3D video. *Signal, Image and Video Processing* [online]. 2015, **9**(2), 331-345 [cit. 2016-05-25]. DOI: 10.1007/s11760-013-0439-0. ISSN 1863-1703. Dostupné z: <http://link.springer.com/10.1007/s11760-013-0439-0>
- [20] JIANG, Lianlian, Jiangqian HE, Nan ZHANG a Tiejun HUANG. An overview of 3D video representation and coding. *3D Research* [online]. 2010, **1**(1), 43-47 [cit. 2016-05-26]. DOI: 10.1007/3DRes.01(2010)6. ISSN 2092-6731. Dostupné z: [http://link.springer.com/10.1007/3DRes.01\(2010\)6](http://link.springer.com/10.1007/3DRes.01(2010)6)
- [21] TOMAN, Jan. Kompresa stereoskopického videa. Praha, 2015. Bakalářská práce. ČVUT, fakulta elektronická.

- [22] VETRO, A, T WIEGAND a G J SULLIVAN. Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard. *Proceedings of the IEEE* [online]. 2011, **99**(4), 626-642 [cit. 2016-05-27]. DOI: 10.1109/JPROC.2010.2098830. ISSN 0018-9219. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5705534>
- [23] MULLER, Karsten, Heiko SCHWARZ, Detlev MARPE, Christian BARTNIK, Sebastian BOSSE, Heribert BRUST, Tobias HINZ, Haricharan LAKSHMAN, Philipp MERKLE, et al. 3D High-Efficiency Video Coding for Multi-View Video and Depth Data. *IEEE Transactions on Image Processing* [online]. 2013, **22**(9): 3366-3378 [cit. 2015-06-24]. DOI: 10.1109/tip.2013.2264820
- [24] STEREOSCOPIC TELEVISION MPEG-2 MULTI-VIEW PROFILE. 1998. Dostupné také z: [https://www.itu.int/dms\\_pub/itu-r/opb/rep/R-REP-BT.2017-1998-PDF-E.pdf](https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-BT.2017-1998-PDF-E.pdf)
- [25] JIANG, Lianlian, Jiangqian HE, Nan ZHANG a Tiejun HUANG. An overview of 3D video representation and coding. *3D Research* [online]. 2010, **1**(1): 43-47 [cit. 2015-06-24]. DOI: 10.1007/3dres.01(2010)6.
- [26] MERKLE, Philipp, Aljoscha SMOLIC, Karsten MULLER a Thomas WIEGAND. Multi-View Video Plus Depth Representation and Coding. In: *2007 IEEE International Conference on Image Processing* [online]. 2007 [cit. 2015-06-19]. DOI: 10.1109/icip.2007.4378926.
- [27] Miska M. HANNUKSELA et al. 3D-AVC Draft Text 8. *Joint Collaborative Team on 3D Video Coding Extensions (JCT-3V) document JCT3V-F1002*, 6<sup>th</sup> Meeting, Geneva, CH, 25 Oct. 1 Nov. 2013
- [28] Next-Generation 3D Formats with Depth Map Support. *IEEE MultiMedia* [online]. 2014, **21**(2): 90-94 [cit. 2015-06-19]. DOI: 10.1109/mmul.2014.31
- [29] ITU-R BT.2021-1. Subjective methods for assessment of stereoscopic 3DTV systems. Geneva: International Telecommunication Union, 2015.
- [30] M. Urvoy, M. Barkowsky, R. Cousseau, Y. Koudota, V. Ricordel, P. Le Callet, J. Gutierrez, N. Garcia: Nama3ds1-cospad1: subjective video quality assessment database on coding conditions introducing freely available high quality 3d stereoscopic sequences
- [31] MITTAL, Anish, Anush K. MOORTHY, Joydeep GHOSH a Alan C. BOVIK. Algorithmic assessment of 3D quality of experience for images and videos. *2011 Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)* [online]. IEEE, 2011, , 338-343 [cit. 2016-05-23]. DOI: 10.1109/DSP-SPE.2011.5739236. ISBN 978-1-61284-226-4. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5739236>
- [32] BALOTA, Gabriel, Mario SALDANHA, Gustavo SANCHEZ, Bruno ZATT, Marcelo PORTO a Luciano AGOSTINI. Overview and quality analysis in 3D-HEVC emergent video coding standard. In: *2014 IEEE 5th Latin American Symposium on Circuits and Systems* [online]. 2014 [cit. 2015-06-25]. DOI: 10.1109/lascas.2014.6820260.
- [33] FLIEGEL, Karel. *Současný vývoj obrazových kompresních metod* [online]. [cit. 2016-05-27]. Dostupné z: [http://www.konferadio.cz/files/prezentace/Fliegel\\_Soucasny-vyvoj-obrazovych-kompresnich-metod.pdf](http://www.konferadio.cz/files/prezentace/Fliegel_Soucasny-vyvoj-obrazovych-kompresnich-metod.pdf)



## 13. Příloha – obsah CD

Bohuslav\_Gruber\_Komprese\_3D\_videa.pdf

SSIM\_3D.m

Subjektivní test.xlsx