

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE FAKULTA
ELEKTROTECHNIKY A INFORMATIKY

Mgr. Adam Polakovič

Autoreferát dizertačnej práce
Adaptácia doručovania multimediálneho obsahu v prostredí 5G sietí

na získanie akademického titulu: „doktor“ („philosophiae doctor“, v skratke „PhD.“)

v doktorandskom študijnom programe: Telekomunikácie

v študijnom odbore: 18. Informatika

Forma štúdia: denná

Miesto a dátum: Bratislava, 2020

Dizertačná práca bola vypracovaná na:

Ústav multimediálnych komunikačných a
informačných technológií
Fakulta elektrotechniky a informatiky
Slovenská technická univerzita v Bratislave

Predkladateľ: **Mgr. Adam Polakovič**

Slovenská technická univerzita v Bratislave
Fakulta elektrotechniky a informatiky
Ústav multimediálnych komunikačných a
informačných technológií
Ilkovičova 3, 812 19 Bratislava

Školiteľ: **prof. Ing. Gregor Rozinaj, PhD.**

Slovenská technická univerzita v Bratislave
Fakulta elektrotechniky a informatiky
Ústav multimediálnych komunikačných a
informačných technológií
Ilkovičova 3, 812 19 Bratislava

Oponenti: **prof. Ing. Milan Dado, PhD.**

Žilinská univerzita v Žiline
Fakulta elektrotechniky a informačných technológií
Katedra multimédií a informačno-komunikačných technológií
Univerzitná 8215/1, 010 26 Žilina

doc. Ing. Katarína Žáková, PhD.

Slovenská technická univerzita v Bratislave
Fakulta elektrotechniky a informatiky
Ústav automobilovej mechatroniky
Ilkovičova 3, 812 19 Bratislava

Autoreferát bol rozoslaný dňa:

**Obhajoba dizertačnej práce sa bude konať dňa o h na Fakulte elektrotechniky
a informatiky STU, Ilkovičova 3, 812 19 Bratislava.**

.....

prof. Dr. Ing. Miloš Oravec, PhD.

dekan fakulty

Obsah

1 ÚVOD	1
2 SÚČASNÝ STAV ADAPTÁCIE DORUČOVANIA VIDEO.....	2
3 CIELE PRÁCE	4
4 METÓDY RIEŠENIA PRÁCE	5
4.1 METÓDA PRE KOMPRESIU VIDEO, KTORÁ VYUŽÍVA INFORMÁCIE O VÝZNAČNOSTI REGIÓNOV V OBRAZE.....	5
4.2 SYSTÉM PRE ADAPTÍVNE DORUČOVANIE VIDEO S VYUŽITÍM MPEG-DASH ZALOŽENÝ NA DETEKCIU POHLADU POUŽÍVATEĽA	6
4.3 METÓDA PRE HĽADANIE OPTIMÁLNYCH PARAMETROV ADAPTÁCIE PRE INDIVIDUÁLNE DORUČOVANIE VIDEO.....	8
4.4 SYSTÉM PRE ADAPTÁCIU DORUČOVANIA MULTIMEDIÁLNEHO OBSAHU, KTORÝ JE VHODNÝ PRE VŠESMEROVÉ VIDEO	10
5 ZHRNUTIE VÝSLEDKOV PRÁCE	17
5.1 NAVRHNUTÁ METÓDA KOMPRESIE VIDEO, KTORÁ VYUŽÍVA INFORMÁCIE O VÝZNAČNOSTI REGIÓNOV V OBRAZE	17
5.2 NAVRHNUTÝ SYSTÉM PRE ADAPTÍVNE DORUČOVANIE VIDEO S VYUŽITÍM MPEG-DASH ZALOŽENÝ NA DETEKCIU POHLADU POUŽÍVATEĽA	18
5.3 NAVRHNUTÁ METÓDA PRE HĽADANIE OPTIMÁLNYCH PARAMETROV ADAPTÁCIE PRE INDIVIDUÁLNE DORUČOVANIE VIDEO	19
5.4 NAVRHNUTÝ SYSTÉM PRE ADAPTÁCIU DORUČOVANIA MULTIMEDIÁLNEHO OBSAHU, KTORÝ JE VHODNÝ PRE VŠESMEROVÉ VIDEO.....	19
6 ZÁVER	22
7 PRÍNOSY DIZERTAČNEJ PRÁCE	24
8 RESUMÉ	25
8 VLASTNÁ PUBLIKAČNÁ ČINNOSŤ	26
8.1 PUBLIKÁCIE TÝKAJÚCE SA DIZERTAČNEJ PRÁCE.....	26
8.2 OSTATNÉ PUBLIKÁCIE	27
9 RIEŠENÉ VÝSKUMNÉ PROJEKTY	27
10 ZOZNAM POUŽITEJ LITERATÚRY	28

1 Úvod

Využívanie mobilných sietí sa v dnešnej dobe rozvíja veľmi rýchlym tempom, pričom vznikajú požiadavky pre video s ultra vysokým rozlíšením (UHD), ktoré vyžaduje až 16-násobnú šírku pásma oproti videu s vysokým rozlíšením (HD). Tento narastajúci trend potreby väčšej šírky pásma sa nebude spomaľovať pričom sa predpokladá, že v najbližších piatich rokoch bude 82% z objemu prenesených dát tvoriť video. Okrem neustále sa zväčšujúceho dopytu pre vyššiu kvalitu videa a všesmerového stereoskopického videa sa tiež zvýšila dostupnosť týchto multimédií pre širokú verejnosť, ktorá má narastajúce nároky na kvalitu.

Vzhľadom na toto neustále zvyšovanie nárokov vzniká potreba pre adaptáciu doručovania videa. V dnešnej dobe je multimediálny obsah pre VR okuliare, ako napríklad virtuálna realita a panoramatické video, doručovaný v plnej kvalite aj mimo pohľadu používateľa. Tento prístup je neefektívny, pretože väčšina prenesených dát nie je využitá a dokonca sa používateľovi ani len nezobrazí. Preto je potrebné urobiť určité kroky, aby doručovanie tohoto obsahu bolo adaptované na požiadavky koncového používateľa a bral sa ohľad aj na zdroje, akými sú napríklad šírka pásma, alebo výkon koncového zariadenia, pričom sa bude používateľovi prenášať len multimediálny obsah, ktorý bude skutočne využitý.

Existuje viacero stratégií pre adaptáciu doručovania multimediálneho obsahu. Vo všeobecnosti sa však dajú rozdeliť na adaptáciu na strane siete a adaptáciu na strane multimédií.

Adaptácia na strane multimédií využíva staticky upravené (komprimované) video, pre ktoré je najskôr potrebné zistiť, ktoré oblasti sú význačné, buď použitím modelov vizuálnej pozornosti, alebo pomocou sledovania pohybu očí. Z takto získaných oblastí záujmu sa tvorí mapa význačnosti, pomocou ktorej sa význačné oblasti vo videu kódujú s vyššou kvalitou ako menej význačné oblasti.

Dynamická adaptácia doručovania videa je aj na úrovni siete. Najrozšírenejšou formou adaptácie doručovania videa je MPEG-DASH, kde sa na základe rýchlosti siete dodáva najlepšia možná reprezentácia multimediálneho obsahu. V prípade škálovateľného videa je konvenčný bitový tok rozložený do vrstiev, pričom je na strane prijímača možné zvoliť a dekodovať rôzny počet vrstiev. Najnižšie podporované prevedenie videa sa nachádza v základnej vrstve, ktorá môže byť rozšírená dekódovaním zlepšujúcich vrstiev.

Pri adaptácii doručovania s cieľom dodávania minimálneho množstva dát s maximálnou využiteľnosťou je potrebné zisťovať polohu pohľadu používateľa v reálnom čase. S využitím znalosti o polohe pohľadu používateľa je následne možné posilať mu vyššiu kvalitu videa v oblastiach kde je jeho pohľad, a v oblastiach mimo jeho pohľadu posilať video s nižšou kvalitou. Toto je možné realizovať dvoma spôsobmi:

- adaptovať video na strane servera a posilať len relevantné časti,
- adaptovať získavanie videa zo servera tak, že sú sťahované len relevantné časti, pričom na strane servera je už kódované video s rôznymi možnosťami škálovateľnosti.

Z pohľadu bitového toku je výhodnejšia prvá možnosť, pretože sa doručované video môže adaptovať priamo pre používateľa. Táto možnosť vyžaduje veľký výkon na strane servera, čo by pre potenciálny veľký počet používateľov bolo nerentabilné. Ďalším problémom tejto možnosti je veľká odozva. Predmetom riešenia tejto práce bol návrh systému adaptácie doručovania videa, ktoré bude mať dobré vlastnosti z pohľadu šetrenia bitového toku, tiež bude mať krátku odozvu na zmeny pohľadu používateľa a bude škálovateľné pre doručovanie väčšiemu množstvu používateľov.

2 Súčasný stav adaptácie doručovania videa

Existuje viacero stratégií pre zníženie zaťažovania siete pri streamovaní videa, ktorými je možné zmierniť dopad na kvalitu zážitku používateľa v prípade jej preťaženia. Jedným z možných prístupov je kódovanie význačných oblastí videa. Cieľom kódovania význačných oblastí videa, je alokovanie väčšieho množstva bitov v prospech týchto význačných oblastí oproti menej význačným oblastiam. Vo všeobecnosti sú používané tieto dva prístupy:

- implicitný – predspracovanie videa pred kódovaním [64][65][66],
- explicitný – modifikovanie dát kodéra [67][68][69][70][71][72].

Prístupy adaptácie doručovania videa na strane klienta sú založené na MPEG-DASH (dynamické adaptívne streamovanie cez HTTP), ktoré bolo predstavené v štandarde ISO/IEC [73]. MPEG-DASH rozdelí konvenčný bitový tok do časových segmentov, ktoré obsahujú kódované médium a metadáta s použitím formátu MPD (popis reprezentácie médií). Video je rozdelené do viacerých reprezentácií, ktoré sú kódované s viacerými bitovými tokmi (napríklad s použitím rôzneho rozlíšenia, kvality kódovania, alebo počtu snímok za sekundu). Reprezentácie sú rozdelené do časových segmentov s danou dĺžkou trvania. Každý segment má unikátnu adresu URL (uniform resource locator), ktorá je prístupná použitím požiadavky typu „HTTP GET“ [74]. Adaptívne schémy založené na DASH umožňujú klientovi, aby si zvolil najvhodnejšie segmenty počas prehrávania videa v závislosti od aktuálnych podmienok siete pre zvýšenie pozorovanej kvality videa.

Autori článku [77] navrhli modifikácie MPEG-DASH pre vytvorenie viacerých reprezentácií rovnakého všesmerového videa. Okrem pridania viacerých reprezentácií s viacerými bitovými tokmi budú reprezentácie ponúkať regióny s vylepšenou kvalitou. Kvalita týchto regiónov sa znižuje so zvyšujúcou sa vzdialenosťou od centra regiónu a podobne ako pri MPEG-DASH sú videá rozdelené do časových úsekov – segmentov. Pri náhlom pohybe používateľa sa prehrávanie videa nezastaví, ale používateľovi je prezentované video s nižšou kvalitou. Autori článku [78] popisujú druhý dodatok MPEG-DASH štandardu číslo 1 23009-1:2014 [79], ktorý predstavuje úpravu formátu MPD na SRD (z anglického spatial relationship description - popis priestorových vzťahov). Je to prístup k streamovaniu iba priestorových častí videa s kombináciou viacerých bitových tokov podporovaných MPEG-DASH. Toto rozšírenie umožňuje DASH klientovi zvoliť iba tie časti videa, ktoré sú relevantné pre používateľa. V prvom rade sa video rozdelí na dlaždice, ktoré sú nezávisle dekódovateľné videá. Toto sa dosiahne tak, že sa video rozdelí na jednotlivé snímky, ktoré sa rozdelia na menšie rozlíšenia a následne sa pospájajú obrazy z jednotlivých regiónov do videa. Autori článku [80] navrhujú rozšírenie MPEG-DASH SRD pre využitie v prostredí 3D virtuálnej reality tak, že sa využíva prepojenie medzi pohľadom používateľa, jeho polohy v priestore a uprednostňovanie sťahovania obsahu použitím algoritmu „rozdeľuj a panuj“. Podobný prístup zvolili aj autori článku [82] využívajúci kubickú projekciu videa v prostredí virtuálnej reality. Video bolo rozdelené na 24 dlaždíc, pričom na jednu stranu kocky pripadajú 4 dlaždice kódované použitím HEVC (z anglického high efficiency video coding – vysoko efektívne kódovanie videa). V článku [83] autori popisujú, ako by bolo možné využiť softvérom definované siete SDN (z anglického Software-Defined Networking) pre streamovanie multimédií v prostredí virtuálnej reality.

Prístupy pre adaptívne streamovanie videa s nerovnomerne rozdeleným bitovým tokom využívajú vlastnosti ľudského oka, teda rozloženie svetlocitlivých buniek v oku. Prenášané video je kvalitné v centre pozornosti človeka, pričom sa jeho kvalita kontinuálne znižuje so zvyšujúcou sa vzdialenosťou od centra pozornosti. Takáto adaptácia by sa mala dosahovať v reálnom čase tak, aby každá ďalšia snímka bola prispôbená pohľadu diváka. Podobný spôsob sa využíva pre renderovanie graficky náročných hier v prostredí virtuálnej reality - foveated rendering, pričom sa ukázalo, že takýto prístup nespôsobuje signifikantný rozdiel v kvalite v porovnaní s priamym renderovaním [84][85]. Autori

článku [86] navrhli systém pre takýto typ streamovania. Systém generuje istý počet regiónov s rôznymi veľkosťami oblastí so zlepšenou kvalitou obrazu. Generované regióny sú následne zlúčené do jednej snímky. Tieto snímky sú následne kódované použitím H.264 kodéra a posielané používateľovi. Rozmazávacia maska je aplikovaná na prechody medzi zlúčenými snímkami pre hladší prechod. Autori deklarujú 5-8 násobné zníženie využitia šírky pásma a dvojnásobné zníženie využitého výkonu procesora. Ďalším prístupom je aplikácia v cloudovom hraní počítačových hier, kde sú snímky v reálnom čase kódované tak, že sú pridelené rôzne kvantizačné parametre makroblokom v závislosti od vzdialenosti pohľadu používateľa, pričom bola dosiahnutá celková odozva systému 110 ms [87]. S využitím princípu MPEG-DASH navrhli autori článku [88] metódu, kde klient pošle serveru požiadavku so sférickými súradnicami a veľkosťou regiónu záujmu. Na základe tejto požiadavky server oreže segment v regióne záujmu s vysokým rozlíšením s odozvou 700 ms. Taktiež sú klientovi prenášané segmenty s nízkym rozlíšením, ktoré sú následne spájané so segmentami s vysokým rozlíšením tak, že časť segmentu s nízkym rozlíšením je vynechaná v prospech segmentu s vysokým rozlíšením.

Jedným zo štandardných prístupov je aj využitie škálovateľného kódovania videa, ktorý bol predstavený vo viacerých formách, napríklad: H.262 (MPEG-2 Video) [89], H.263 [90], MPEG-4 Visual [91], H.264 SVC [92] a H.265 HEVC [93]. Škálovanie videa je kompresným algoritmom, ktorý umožňuje efektívny prenos videa cez siete s obmedzenou šírkou pásma. V tomto prípade je konvenčný bitový tok rozdelený na základnú vrstvu a viacero vylepšujúcich vrstiev, ktoré vylepšujú základnú vrstvu. Jedným z dôvodov pre nízku používanosť škálovateľnosti je znížená efektivita kódovania pri priestorovej škálovateľnosti a škálovateľnosti kvality, a taktiež väčšia zložitosť dekódovania v porovnaní s neškálovateľnými profilmi [94].

Cieľom všetkých popísaných metód je šetrenie prenášaného bitového toku a maximalizovanie, alebo zachovanie zážitku koncového používateľa. Žiaden z týchto systémov neponúkal dostatočne kvalitné riešenie pri zohľadnení všetkých aspektov doručovania. V prípade statickej úpravy videa chýbala individuálna adaptácia v prípade, že používateľov pohľad je aj napriek predpokladom v regiónoch, ktoré boli považované za nevýznačné. V prípade dynamickej adaptácie pozorujeme nasledovné problémy navrhovaných riešení:

- rozhranie medzi kvalitným regiónom a menej kvalitným regiónom malo výrazný pozorovateľný prechod,
- adaptácia na zmenu pohľadu používateľa bola dosiahnutá za dlhšiu dobu ako 100 ms (zvyčajne to bola približne 1 sekunda),
- nedostatočná škálovateľnosť počtu používateľov pri riešeníach s kódovaním videa v reálnom čase na strane servera,
- v prípade škálovateľného kódovania videa nebola adaptácia zameraná na región pohľadu používateľa.

3 Ciele práce

Ciele dizertačnej práce „Adaptácia doručovania multimediálneho obsahu v prostredí 5G sietí“ sú:

1. Navrhnuť novú metódu pre kompresiu videa, ktorá využíva informácie o význačnosti regiónov v obraze.
2. Navrhnuť systém pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa.
3. Navrhnuť metódu pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa (napr. priestorový uhol a stupeň zníženia bitového toku v ňom, ktorý nie je badateľný pre bežného používateľa).
4. Navrhnuť systém pre adaptáciu doručovania multimediálneho obsahu, ktorý je vhodný pre všesmerové video a maximalizuje využiteľnosť prenesených dát individuálnemu používateľovi a algoritmus načítavania videa do vyrovnávacej pamäte.

4 Metódy riešenia práce

Problém adaptácie doručovania multimediálneho obsahu je aktuálny aj v prostredí 5G sietí, kde bude možné prenášať všesmerové video s bitovým tokom do rýchlosti 1 Gbit/s. Kombinácia krátkej odozvy (1 ms) a vysokej rýchlosti v 5G sieťach s možnosťou vytvorenia sieťovej vrstvy s prispôsobiteľnými sieťovými funkciami umožní nové spôsoby adaptácie doručovania multimediálneho obsahu. Navrhované metódy riešenia dizertačnej práce sa dajú rozdeliť do dvoch častí. Prvá časť sa venuje metódam statickej úpravy videa pre adaptáciu doručovania. Druhá časť sa venuje dynamickému doručovaniu videa individuálnemu používateľovi a predpríprave videa pre takéto doručovanie.

4.1 Metóda pre kompresiu videa, ktorá využíva informácie o význačnosti regiónov v obraze

Pre statické uskladnenie videa, alebo pri nemožnosti vysielania multimediálneho obsahu adaptovaného na používateľa, je potrebné upraviť video pred jeho samotným použitím (napríklad skladovanie na disku, alebo streamovanie). Navrhli sme **metódu kompresie videa, ktorá využíva informácie o význačnosti regiónov v obraze**. Súčasťou tejto metódy je spôsob pre generovanie máp význačnosti regiónov v obraze z dát získaných pomocou sledovania polohy očí od viacerých používateľov. Nakoniec sme navrhli metódu, ktorá s využitím tejto mapy kóduje video s väčším počtom bitov vo význačných oblastiach.

Vstupom navrhovanej metódy je video sekvencia a prislúchajúce informácie o polohe pohľadu používateľov. Výstupom tejto metódy je generované video zohľadňujúce význačné oblasti. Mapa význačnosti pre danú snímku videa je počítaná zo vstupných dát o pohľade používateľa. Následne je na snímku aplikovaný filter nerovnomerného rozmazávania s rôznou úrovňou zodpovedajúcou hladine význačnosti daného pixelu. Spracované snímky sú následne presunuté do video kódéra H.264. Všetky snímky s priestorovo nerovnomerne rozdeleným bitovým tokom sú následne spojené dokopy, pričom je vyprodukované finálne video s nerovnomerne rozdeleným bitovým tokom.

4.1.1 Generovanie máp význačnosti

Pôvodné videá boli rozdelené do jednotlivých snímok. Mapa význačnosti je zložená z dvojrozmerných Gaussových funkcií, pričom pre každú snímku a každý eye-tracker bola pripočítaná dvojrozmerná Gaussova funkcia s príslušnými súradnicami. Výsledná mapa význačnosti je daná vzťahom:

$$p(x, y) = \sum_{i=1}^N e^{-\frac{1}{2\sigma_h^2} \left(\frac{(x-x_{e_i})^2}{\left(\frac{W}{H}\right)^2} + (y-y_{e_i})^2 \right)} \quad (1)$$

Kde $p(x, y)$ je hodnota pixelu na relatívnych súradniciach (x, y) idúcich od 0 po 1, pričom $(0, 0)$ znamená pixel v ľavom spodnom rohu a $(1, 1)$ znamená pixel v pravom hornom rohu snímky, i je premenná idúca od 1 po N , pričom N je počet eye-trackero, e je Eulerovo číslo, σ_h je štandardná odchýlka v relatívnych súradniciach výšky y , W je šírka videa v pixeloch, H je výška videa v pixeloch a (x_{e_i}, y_{e_i}) sú relatívne súradnice i -tej pozície eye-trackera. Mapa význačnosti je teda zložená z N prekrývajúcich sa dvojdimenzionálnych Gaussových funkcií, pričom ich hodnoty v tejto mape sú medzi 0 a N .

4.1.2 Rozmazávanie nevýznačných regiónov

Rozmazávanie nevýznačných regiónov je vykonávané dolnopriepustným filtrom riadeným mapou význačnosti. Hodnoty z vygenerovanej mapy význačnosti sú transformované na hodnoty medzi 0 a 1. Gaussovo rozmazávanie je vykonávané pre každú snímku v 20 iteráciách. Videá s rôzne rozmazanými oblasťami vytvorené týmto procesom sú následne kódované kódrom H.264 s konštantným bitovým tokom.

4.2 Systém pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa

Štandard pre dynamickú adaptáciu doručovania multimédií cez HTTP: MPEG-DASH zatiaľ nevyužíva informácie o význačnosti regiónov v obraze. Pri zhoršených podmienkach bitového toku sa znižuje celková kvalita prenášaného videa. Namiesto zníženia kvality celého videa by sa mohla znížiť kvalita iba v nevýznačných oblastiach, čím by sa tiež znížil bitový tok. Preto sme navrhli **systém pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa**. Tento systém integruje vyššie navrhovanú metódu statickej úpravy videa s dynamickou adaptáciou doručovania multimédií cez HTTP: MPEG-DASH. Súčasťou navrhnutého systému sú režimy využívajúce kódované videá s rôznymi stupňami adaptácie:

- pasívny režim s využitím vopred získanej mapy význačnosti,
- validovaný pasívny režim – v závislosti na pohľade používateľa systém prepína medzi využívaním pasívneho režimu a režimu bez využitia máp význačnosti,
- režim s individuálnymi regiónmi – používateľovi je prenášané video so zlepšenou kvalitou v oblasti jeho pohľadu.

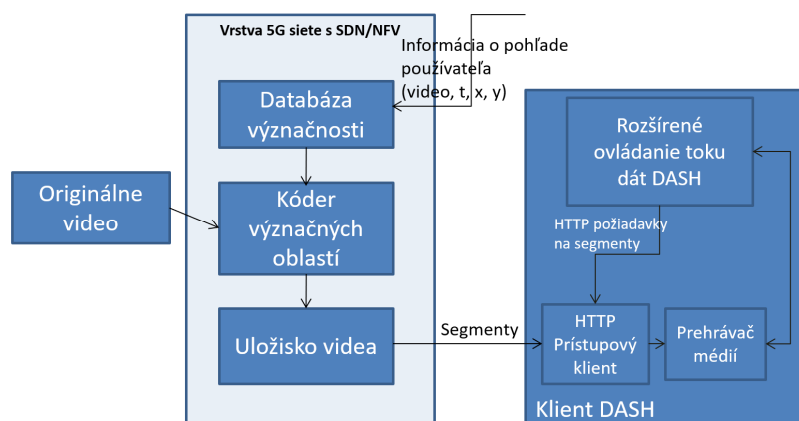
Tieto režimy adaptácie sú popísané v nasledujúcich podkapitolách.

4.2.1 Pasívny režim význačnosti

Pasívny režim význačnosti rozširuje zoznam reprezentácií MPEG-DASH o dve nové reprezentácie založené na mapách význačnosti: reprezentácia s vylepšenou kvalitou na význačných miestach a reprezentácia so zachovanou kvalitou význačných miest. V prípade reprezentácie s vylepšenou kvalitou na význačných miestach je priradený bitový tok využitý pre mierne zvýšenie kvality pre význačné miesta a mierne zníženie na nevýznačných miestach. Reprezentácia so zachovanou kvalitou význačných miest je aplikovaná v prípade zníženého bitového toku, pričom je zachovávaná kvalita najmä vo význačných oblastiach. Tieto oblasti sú určené modelom význačnosti, ktorý je reprezentovaný mapou význačnosti. Mapa význačnosti je tvorená z vyzbieraných bodov (video, čas, súradnice) od viacerých používateľov. Každý bod prispieva do mapy význačnosti Gaussovou funkciou s centrom s prislúchajúcimi súradnicami, pričom mapa je naškálovaná na hodnoty od 0 do 255. Video je kódované s použitím získanej mapy význačnosti s využitím referenčného kodéra význačných oblastí (upravený kodér x264) [72]. V pasívnom režime je možné postupne vylepšovať reprezentácie s pribúdajúcimi informáciami o pohľade používateľov na základe mapy význačnosti počas životného cyklu aplikácie.

Keď je dostupná reprezentácia so zachovanou kvalitou význačných miest, tak riadiaca jednotka streamovania DASH môže namiesto znižovania kvality a rozlíšenia v prípade zníženého bitového toku prejsť na video s nižším bitovým tokom, pričom sa zachová rovnaká subjektívna kvalita videa. Bitový tok pri zachovaní rovnakej kvality zážitku je o 23% nižší pri využití referenčného kodéra význačných oblastí [72]. Na obrázku 1 je zobrazená architektúra systému adaptácie doručovania v prostredí 5G sietí s využitím SDN/NFV (Software-Defined Networking a Network Function

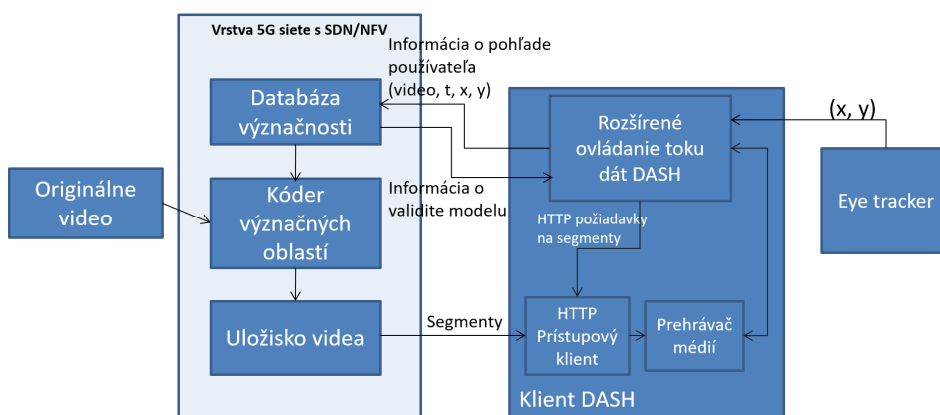
Virtualization). Súbory MPD budú obsahovať dve nové reprezentácie v adaptačnom sete videa. Klient DASH musí byť upravený pre ich využívanie na základe preferencií používateľa.



Obrázok 1 – Architektúra systému adaptívneho doručovania videa v prostredí 5G sietí s využitím DASH pri pasívnom režime s využitím vopred získanej mapy význačnosti.

4.2.2 Validovaný pasívny režim

Validovaný pasívny režim je vhodný pre používateľov s eye-trackerom. Pri tomto režime zbiera DASH klient informácie o pohľade používateľa počas prehrávania daných segmentov a ukladá ich v databáze význačnosti. Na základe informácií z databázy význačnosti je vo vytvorenej vrstve 5G siete vyhodnotené, či používateľov pohľad súhlasí s pohľadmi iných používateľov. Táto informácia je následne posielaná DASH klientovi pre finálne rozhodnutie o vhodnej reprezentácii pre načítanie. Na základe týchto informácií DASH klient prepína medzi pasívnym režimom a bežným režimom, ktorý nevyužíva informácie o význačnosti. Na obrázku 2 je zobrazená príslušná architektúra systému pre doručovanie videa v 5G sieťach. Súbory MPD budú v tomto prípade obsahovať dve nové reprezentácie v adaptačnom sete videa rovnako ako v prípade pasívneho režimu.



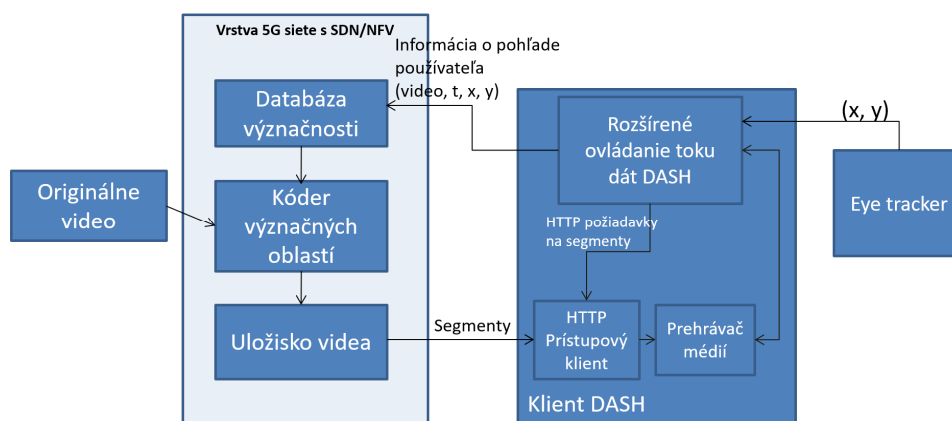
Obrázok 2 – Architektúra systému adaptívneho doručovania videa v prostredí 5G sietí s využitím DASH pri validovanom pasívnom režime.

4.2.3 Režim s individuálnymi regiónmi

Režim s individuálnymi regiónmi umožňuje vyššiu flexibilitu pri individualizácii doručovania obsahu. V tomto režime je vytvorených viacero reprezentácií s rôznymi miestami s vylepšenou

kvalitou pre všetky potenciálne pohľady používateľa. Pre kódovanie týchto blokov s vylepšenou kvalitou v centrách význačnosti je využitý referenčný kodér význačných oblastí (upravený kodér x264) [72]. Výsledné súradnice centier s vylepšenou kvalitou sú dodatočnými parametrami reprezentácií vo formáte MPD.

V tomto režime klient DASH zbiera informácie o pohľade používateľa počas prehrávania segmentu a vyhodnocuje, ktorá reprezentácia je vhodná pre používateľa. Pre fungovanie navrhnutého konceptu je potrebné upraviť implementáciu DASH klienta tak, aby bral súradnice z eye-trackera ako vstup. S touto informáciou bude DASH klient schopný posilať požiadavky na segmenty, ktoré sú najbližšie daným súradniciam v prípade, že sa pohľad používateľa nezmenil určitú dobu. Takáto reprezentácia sa bude prenášať bez ohľadu na momentálny obsah vyrovnávacej pamäte a klient bude posilať požiadavku na najskorší možný segment s ohľadom na aktuálnu šírku pásma pre hladkú tranzíciu medzi reprezentáciami. Architektúra navrhnutého systému je zobrazená na obrázku 3.



Obrázok 3 – Architektúra systému doručovania videa v prostredí 5G sietí s využitím DASH s režimom s individuálnymi regiónmi.

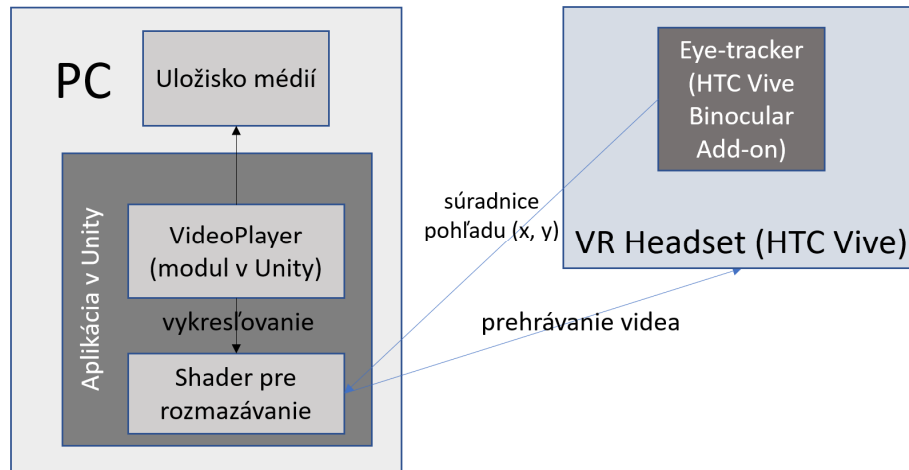
4.3 Metóda pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa

4.3.1 Princíp a architektúra systému

V súčasnom stave zatiaľ nebola venovaná dostatočná pozornosť otázkam ohľadom veľkosti a tvaru plochy so zvýšenou kvalitou a tiež ohľadom toho, aký úbytok kvality videa je možný v nevýznačných oblastiach bez vplyvu na kvalitu zážitku používateľa. Preto sme navrhli **metódu pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa**, konkrétne priestorový uhol oblasti s plnou kvalitou a gradient zníženia kvality v ňom, ktorý nie je badateľný pre bežného používateľa. Pre zistenie týchto parametrov je vytvorená simulácia dokonalého doručovania multimediálneho obsahu v prostredí virtuálnej reality, čo znamená systém s adaptáciou na zmenu pohľadu používateľa s odozvou 0 ms, a teda každá nasledujúca vykreslená snímka bude rozmazávaná podľa najnovšieho pohľadu.

Navrhnutý systém sa skladá z osobného počítača (PC) s aplikáciou implementovanou v prostredí *Unity* a jednotky pre ukladanie dát. Aplikácia využíva komponent *VideoPlayer* pre vykresľovanie všesmerového videa na povrch gule. Výsledky z vykresľovania sú následne pozmenené použitím grafického shadera, ktorý simuluje adaptáciu doručovania videa v závislosti na informáciách o polohe pohľadu používateľa z eye-trackera. Tento obraz je potom poslaný do headsetu pre virtuálnu realitu HTC Vive na zobrazenie používateľovi. Použitý VR headset obsahuje vstavaný eye-tracker

HTC Vive Binocular Add-on pre sledovanie polohy očí, ktorý posiela aplikácii informácie o polohe očí (súradnice x, y). Popísaná architektúra systému je zobrazená na obrázku 4.



Obrázok 4 – Architektúra systému pre simuláciu dokonalého doručovania multimediálneho obsahu na vyžiadanie v reálnom čase v prostredí virtuálnej reality.

4.3.2 Implementácia shadera – rozmazávanie v okolí centra pohľadu

Pre simuláciu kontinuálneho zníženia bitového toku smerom od centrálnej časti na periférne časti zobrazovaného videa je upravený výsledok z renderovania komponentu *VideoPlayer*. Cieľom je ponechať plnú kvalitu videa v centrálnej časti, pričom sa bude postupne znižovať kvalita s rastúcou vzdialenosťou od regiónu s plnou kvalitou. Význačnosť jednotlivých pixelov pre zobrazovanie všesmerového videa v okuliároch virtuálnej reality je reprezentovaná nasledujúcou funkciou:

$$\begin{aligned}
 i(x, y) &= 1 & r \in (0; R) \\
 i(x, y) &= 1 - q \cdot r(x, y) & r \in \langle R; \rho \rangle \\
 i(x, y) &= 0 & r \in (\rho; \infty)
 \end{aligned} \tag{2}$$

Kde $i(x, y)$ je význačnosť pixelu nachádzajúceho sa na súradniciach (x, y) , R je polomer, v ktorom je video dodávané v plnej kvalite od stredu pohľadu, q je kvocient degradácie kvality, hodnota ρ predstavuje polomer obrazu, ktorý je zobrazovaný v okuliároch virtuálnej reality a $r(x, y)$ je vzdialenosť pixelu od centra pohľadu používateľa a je definovaný nasledovne:

$$r(x, y) = \frac{(x - x_{e_i})^2}{\left(\frac{W}{H}\right)^2} + (y - y_{e_i})^2 \tag{3}$$

Kde uvedené hodnoty súradníc predstavujú relatívne hodnoty na škále od 0 po 1, hodnoty W a H sú šírka a výška videa v pixeloch a (x_e, y_e) sú relatívne súradnice polohy pohľadu z eye-trackera. Pre každý pixel sa nakoniec vykoná nerovnomerné rozmazávanie s aplikovaním matice 3×3 Gaussového filtra zobrazeného na obrázku 5.

$$\frac{1}{\sigma^2} \times \begin{bmatrix} e^{-\frac{2}{2\sigma^2}} & e^{-\frac{1}{2\sigma^2}} & e^{-\frac{2}{2\sigma^2}} \\ e^{-\frac{1}{2\sigma^2}} & e^{-\frac{0}{2\sigma^2}} & e^{-\frac{1}{2\sigma^2}} \\ e^{-\frac{2}{2\sigma^2}} & e^{-\frac{1}{2\sigma^2}} & e^{-\frac{2}{2\sigma^2}} \end{bmatrix}$$

Obrázok 5 – Matica s rozmermi 3×3 Gaussového filtra aplikovaná na každý pixel.

Členy v matici sme upravili tak, aby bol celkový súčet členov matice rovný presne 1. Hodnota σ v matici zodpovedá štandardnej odchýlke Gaussového filtra a je definovaná podľa význačnosti. Gaussov filter pre rozmazávanie je následne aplikovaný na každý pixel každej snímky v 20 iteráciách.

4.4 Systém pre adaptáciu doručovania multimedialného obsahu, ktorý je vhodný pre všesmerové video

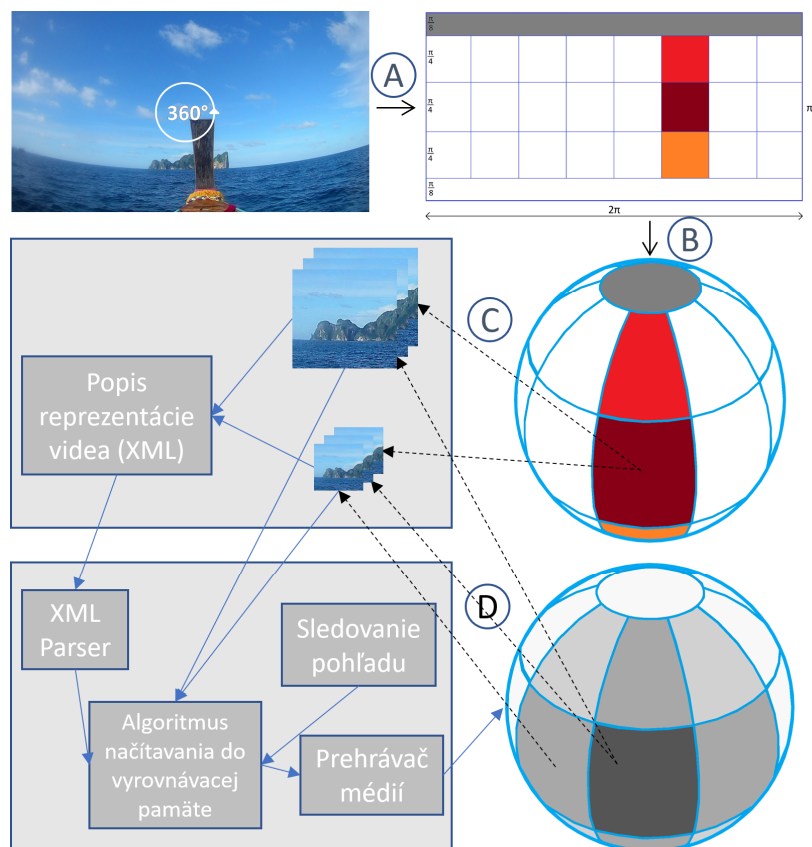
Metódy adaptáciu doručovania multimedialného obsahu popísané v kapitole o súčasnom stave problematiky majú niektoré základné nedostatky. Systémy, ktoré využívajú streamovanie videa s nerovnomerne rozdeleným bitovým tokom, ponúkajú nedostatočné možnosti adaptácie na zmenu pohľadu používateľa (v niektorých prípadoch môže trvať až 2 sekundy). V prípade škálovateľného videa sú nedostatočné možnosti pre škálovateľnosť doručovania iba význačných oblastí. Preto sme navrhli **systém pre adaptáciu doručovania multimedialného obsahu, ktorý je vhodný pre všesmerové video a maximalizuje využiteľnosť prenesených dát individuálnemu používateľovi.**

4.4.1 Princíp a architektúra systému

Hlavným cieľom navrhovaného systému je doručovanie najlepšej možnej kvality zážitku pre používateľa pri najnižšej možnej spotrebe šírky pásma. Princípom systému je doručovanie obmedzenej oblasti všesmerového videa, ktorá je v pohľade pozorovateľa, alebo s vysokou pravdepodobnosťou bude v jeho pohľade v krátkom čase. Architektúra navrhovaného systému sa skladá zo serverovej časti a klientovej časti, ktorá získava multimedialný obsah zo servera pomocou požiadaviek HTTP-GET. Funkčná schéma tohoto systému je zobrazená na obrázku 6, kde video kódované konvenčnou ekvidištančnou valcovou projekciou je rozdelené na dlaždice, ktoré sú následne rozdelené do vrstiev škálovateľného videa, ktoré je možné prenášať používateľovi v rôznom počte v závislosti od polohy pohľadu.

Serverová časť systému pozostáva z HTTP servera, ktorý uchováva všesmerové videá. Každé video uložené na serveri je rozdelené do dlaždíc s pevne stanovenou dĺžkou trvania (zvyčajne 1 až 10 sekúnd). Dlaždice sú rozdelené do vrstiev škálovateľne kódovaného videa (SVC) na základnú vrstvu s nízkou kvalitou a bitovým tokom a viacero vylepšujúcich vrstiev so zvyšujúcou sa kvalitou a bitovým tokom. Na serveri je taktiež prístupný štruktúrovaný popis reprezentácie videa vo formáte XML, ktorý obsahuje informácie o dĺžke trvania dlaždíc, plochu pokrytú dlaždicami a tiež informácie o rozdelení dlaždíc na vrstvy.

Klientská časť sa skladá zo zariadenia na prehrávanie videa (napr. VR headset, osobný počítač, smartfón) a klientskej aplikácie s algoritmom načítavania dlaždíc do vyrovnávacej pamäte. Aplikácia na strane klienta najskôr načíta súbor XML zo servera, ktorý obsahuje informácie o rozdelení videa na dlaždice. Pohľad používateľa je priebežne sledovaný príslušným zariadením, napríklad zabudovaným eye-trackerom vo VR headsete, sledovaním polohy hlavy vo VR headsete, navigovanie pomocou myši v prípade osobného počítača, alebo pomocou akcelerometra a gyroskopu v prípade smartfónu. Klient načítava základnú vrstvu a vylepšujúce vrstvy podľa ich vzdialenosti od pohľadu používateľa. So zvyšujúcou sa vzdialenosťou od pohľadu používateľa sa znižuje počet prenášaných vylepšujúcich vrstiev. Pri náhlej zmene pohľadu používateľa klient načíta dlaždicu so skorším počiatkom a synchronizuje jej prehrávanie na rovnakú snímku s práve prehrávaným obsahom. Tento princíp sa aplikuje rovnako aj v prípade vylepšujúcich vrstiev, ktoré sú pripočítavané ku základnej vrstve. Popísaný proces umožní selektívne doručovanie videa iba v regiónoch záujmu a tým umožní šetriť šírku pásma.



Obrázok 6 – Systém adaptívneho doručovania všesmerového videa: a) rozdelenie konvenčného bitového toku do dlaždíc, b) projekcia dlaždíc na povrch gule, c) rozdelenie dlaždíc do vrstiev škálovateľného videa, d) dlaždica v pohľade používateľa je zložená z viacerých vrstiev v porovnaní s dlaždicami, ktoré sú vzdialenejšie od pohľadu používateľa.

4.4.2 Vytváranie adaptovateľného video obsahu

Pre vytvorenie adaptovateľného video obsahu sme na manipuláciu s videom použili program *FFmpeg*. Vytváranie adaptovateľného obsahu pozostáva z preskupenia kľúčových snímok videa tak, aby každý segment začínal kľúčovou snímku, rozdelenia videa na časové segmenty s fixnou dĺžkou trvania pre doručovanie iba krátkych úsekov videa, rozdelenia segmentov na dlaždice podľa schémy na obrázku 7 pre doručovanie iba určitých regiónov v pohľade používateľa, vytvorenia škálovateľných vrstiev dlaždíc pre regulovanie bitového toku v jednotlivých oblastiach podľa pohľadu používateľa a vytvorenia štruktúrovaného popisu multimédií vo formáte XML, ktorý popisuje uvedené delenie videa.

Okrem video súborov je vytvorený aj súbor pre popis multimediálneho obsahu vo formáte XML. Tento súbor obsahuje informácie o formáte videa, názve videa, informácie o dĺžke trvania videa, informácie o dĺžke trvania segmentov, informácie o priestorovom usporiadaní dlaždíc a informácie o vrstvách dlaždíc. Príklad takéhoto súboru XML je zobrazený na obrázku 8. V úvode súboru je popísaná konvencia pre názvy video súborov, čo znamená, že napríklad vylepšujúca vrstva štvrtej dlaždice v šiestej sekunde videa predstavuje videom s názvom: “17_UnderwaterPark_segment6_tile4_layer2.mp4”. Video je zložené z viacerých dlaždíc, pričom každá dlaždica má definovanú veľkosť pokrývaného rozsahu (hodnoty $[x, y]$ medzi 0 a 1) a posun vzhľadom na počiatok videa v bode $[0, 0]$ (hodnoty $[x, y]$ medzi 0 a 1). Každá dlaždica je zložená z jedného alebo viacerých vrstiev, pričom prvá vrstva je základná a ďalšie sú vylepšujúce vrstvy, ktoré

sa pripočítavajú k základnej vrstve. Každá dlaždica má definovanú kvalitu (reprezentovanú hodnotou CRF) a rozlíšenie (šírka a výška v počte pixelov).



Obrázok 7 – Schémy pre tvorbu dlaždíc.

```
<XML xmlns="vda" type="mp4" duration="20000"
segmentDuration="1000" videoName="17_UnderwaterPark"
segmentName="segment%" sectorName="tile%"
layerName="layer%">
...
<Sector id="4" tiling="0.166/0.25" offset="0.333/0.666">
  <Layer id="1" crf="30" width="640" height="480"/>
  <Layer id="2" crf="18" width="640" height="480"/>
</Sector>
...
<Sector id="8" tiling="0.166/0.333" offset="0.00/0.333">
  <Layer id="1" crf="30" width="640" height="640"/>
  <Layer id="2" crf="18" width="640" height="640"/>
</Sector>
...
</XML>
```

Obrázok 8 – Príklad popisu multimediálneho obsahu vo formáte XML.

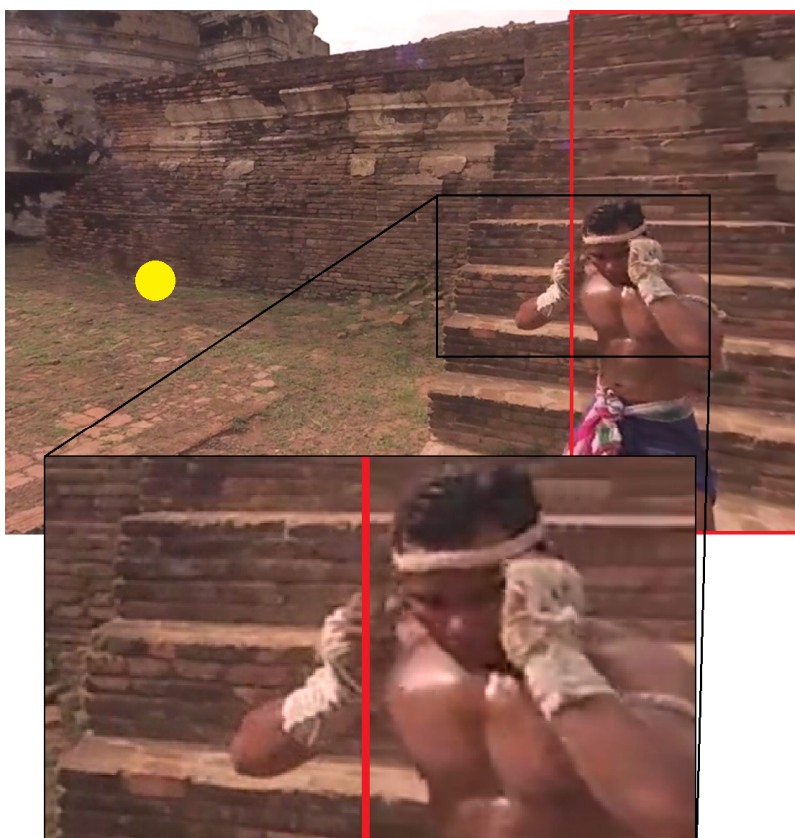
4.4.3 HTTP server poskytujúci multimediálny obsah

Vytvorené video súbory so svojimi škálovateľnými vrstvami sú dostupné na HTTP serveri spolu s XML súborom pre popis multimediálneho obsahu. Pre tento experiment sme využili Microsoft Internet Information Services (IIS) na serveri nachádzajúcom sa v rovnakej domácej sieti (LAN) ako klient. Tento server poskytuje multimediálny obsah pomocou *HTTP GET* požiadaviek. Všetky video súbory sú dostupné na serveri pre načítanie klientom spolu s príslušnými názvami popísanými v súbore pre popis mediálneho obsahu.

4.4.4 Aplikácia pre prehrávanie všesmerového videa

Aplikácia na strane klienta pre prehrávanie všesmerového videa bola implementovaná v prostredí *Unity*. Aplikácia najskôr rozoberie XML súbor s popisom multimediálneho obsahu a následne podľa schémy pre tvorbu dlaždíc vytvorí pre každú dlaždicu samostatný komponent pre renderovanie video dlaždice do samostatnej textúry. Každá textúra je priradená veľkosť rozsahu a posun v priestore vzhľadom na určený bod na povrchu gule. Textúry sú vykresľované na “skybox” (nebeská sféra – imaginárny povrch gule s nekonečným polomerom).

Všetky dlaždice videa, ktoré sú dostupné vo vyrovnávacej pamäti dlaždíc sú prehrávané, v inom prípade sa vykresľuje šedá farba. V prípade, že je dostupná základná vrstva dlaždice a vylepšujúca vrstva nie je dostupná, textúra pre vylepšujúcu vrstvu je nastavená na 50% šedú farbu, ktorá sa následne odpočítava v shaderi. Príklad prehrávania klienta a porovnanie medzi prehrávaním základnej vrstvy a základnej vrstvy s pripočítanou vylepšujúcou vrstvou sú zobrazené na obrázku 9.



Obrázok 9 – Príklad prehrávania všesmerového videa klientom. Žltá bodka predstavuje centrum pohľadu používateľa, pričom v tejto oblasti sú dlaždice zložené z oboch vrstiev – základnej a vylepšujúcej. Dlaždice v červenom obdĺžniku sú vykresľované iba zo základnej vrstvy a majú nižšiu kvalitu.

Dlaždice sa načítavajú podľa algoritmu pre načítavanie videa do vyrovnávacej pamäte, ktorý je popísaný nižšie. Pohľad vo videu je ovládateľný ťahavým pohybom myši, alebo pohybom headsetu v prípade prehrávania do VR headsetu. Prehrávanie videa rozdeľujeme na 2 typy: včasné prehrávanie videa a neskoré prehrávanie videa. *Včasné prehrávanie videa* znamená, že dlaždica daného segmentu bola už načítaná, keď sa má začať prehrávať tento segment. V tomto prípade dochádza k prehrávaniu všetkých načítaných dlaždíc s daným segmentom v rovnakom čase. *Neskoré prehrávanie videa* znamená, že určitá dlaždica s daným segmentom bola načítaná do vyrovnávacej pamäte už počas prehrávania daného segmentu. V tomto prípade sa dlaždica začne prehrávať na rovnakom čísle

snímky ako sú aktuálne prehrávané dlaždice. Toto znamená, že prehrávanie neskoro načítanej dlaždice sa synchronizuje s ostatnými dlaždicami.

4.4.5 Algoritmus pre načítavanie videa do vyrovnávacej pamäte

Algoritmus pre načítavanie videa vychádza z toho, že adaptovateľný obsah s úsekmi videa je už vygenerovaný a je dostupný na serveri. Úsekom videa budeme v nasledujúcom texte rozumieť video V s daným číslom segmentu s , dlaždice i , vrstvy l v čase prehrávania t označené ako $V_s^{i,l}(t)$. Algoritmus pre načítavanie videa do vyrovnávacej pamäte je zodpovedný za naplnenie vyrovnávacej pamäte vhodnými úsekmi videa. Pre selekciu kvality dlaždíc sme zvolili rozdelenie význačnosti do koncentrických kruhových regiónov v závislosti od vzdialenosti od pohľadu používateľa, alebo možného pohľadu v krátkej budúcnosti. Dôležitým parametrom potrebným pre adaptáciu je vzdialenosť pohľadu používateľa od stredu určitej dlaždice. Najkratšia spojnica medzi dvoma bodmi na guľovej ploche nazývaná ortodróma je tvorená kratším oblúkom hlavnej kružnice. Zo sférickej kosínusovej vety je ortodróma d je vypočítaná podľa nasledujúceho vzťahu:

$$d = r \cdot \arccos(\sin(\varphi_1)\sin(\varphi_2) + \cos(\varphi_1)\cos(\varphi_2)\cos(\lambda_2 - \lambda_1)) \quad (4)$$

Kde $[\varphi_1, \lambda_1], [\varphi_2, \lambda_2]$ sú sférické súradnice okrajových bodov ortodrómy P, Q a σ je jej stredový uhol a r je polomer gule. Dlaždica je považovaná za kandidáta pre načítanie do vyrovnávacej pamäte v prípade, že sa nachádza v regióne pohľadu používateľa a môže byť načítaná vo viacerých úrovniach kvality. Región kvality je definovaný v závislosti od vzdialenosti od pohľadu používateľa: menej ako D_{base} (my sme zvolili $1,8 \cdot r$) a dlaždici je pridelená základná kvalita zobrazeného obsahu (základná vrstva) a menej ako $D_{enhancement}$ (my sme zvolili $0,9 \cdot r$) a dlaždici je pridelená vylepšená kvalita (základná aj vylepšujúca vrstva). Algoritmus je jednoducho rozšíriteľný pre použitie s viacerými vylepšujúcimi vrstvami. Hodnoty D_{base} a $D_{enhancement}$ boli zvolené tak, aby bolo pokryté zorné pole používateľa s vylepšenou kvalitou v prípade, že nepohybuje hlavou a základnou kvalitou (pokým systém nenačíta vylepšujúcu vrstvu) v prípade, že používateľ náhle zmení polohu pohľadu. Algoritmus využíva prioritný front pre zaradenie kandidátov úsekov videa pre načítanie. Tieto úseky môžu byť načítané a neskôr prehrávané.

Zaradovanie kandidátov úsekov videa do prioritného frontu

Algoritmus pre načítavanie videa využíva vyrovnávaciu pamäť s dĺžkou BL . Aby sa úsek videa mohol načítať, musí byť číslo jeho segmentu s medzi číslom $s_{current}$ práve prehrávaného segmentu a $s_{current}$ plus dĺžka vyrovnávacej pamäte, vrátane:

$$s_{current} \leq s \leq s_{current} + BL \quad (5)$$

Okrem tejto podmienky, musí platiť, že vzdialenosť d medzi stredom úseku videa a pohľadom používateľa musí byť menší ako D_{base} pre základnú vrstvu:

$$d(V_s^{i,0}(t), \text{Pohľad}) < D_{base} \quad (6)$$

Podobný vzťah platí aj pre vylepšujúcu vrstvu:

$$d(V_s^{i,1}(t), \text{Pohľad}) < D_{enhancement} \quad (7)$$

Priorita P úseku videa v prioritnom fronte je definovaná nasledujúcim vzťahom:

$$P(V_s^{i,l}) = P_{max} - \alpha \cdot (s - s_{current}) - \beta \cdot d - \gamma \cdot l \quad (8)$$

Kde P_{max} je maximálna možná priorita (my sme zvolili $P_{max} = 1000$) a α, β, γ sú prioritizačné konštanty. Konštanta α je váha pre časovú vzdialenosť práve prehrávaného segmentu $s_{current}$ a segmentu s úseku videa vo fronte. Konštanta β je váha pre priestorovú vzdialenosť d medzi stredom úseku videa a pohľadom používateľa a γ je váha indexu kvality učeného vrstvou l . V experimentoch uvedených v kapitole 5.4 s výsledkami boli použité konštanty $\alpha = 100, \beta = 10, \gamma = 1$. Hodnoty týchto konštánt sme zvolili tak, aby boli najskôr prenášané úseky videa, ktorých segment je práve prehrávaný, potom úseky videa, ktoré sú najbližšie pohľadu používateľa a nakoniec ich vylepšujúce vrstvy.

Vyberanie kandidátov úsekov videa z prioritného frontu

Paralelne s procesom zaraďovania do frontu, algoritmus priebežne vyberá kandidátov z prioritného frontu pre načítanie (prenos a dekodovanie) a prehrávanie. Pre zachovanie krátkeho času načítavania úsekov videa sme určili maximálny počet simultánne načítavaných úsekov videa BV_{max} (v našom prípade $BV_{max} = 2$). Keďže pohľad používateľa sa mení priebežne, priorita vyberania kandidátov úsekov videa z prioritnej fronty je rozhodnutá až v čase, keď je možné začať načítavanie ďalšieho videa. Týmto spôsobom je prioritizované poradie prenosu a dekodovania videa v závislosti od dôležitosti úsekov videa pre používateľa.

Odstraňovanie neaktuálnych kandidátov úsekov videa z prioritného frontu

V procese zaraďovania je pridaných viacero kandidátov úsekov videa do prioritného frontu. Niektoré úseky videa nemuseli byť načítané kvôli ich nízkej priorite a stali sa neaktuálnymi. Úsek videa, ktorý má nižšie číslo segmentu s ako je práve prehrávaný segment $s_{current}$ je odstránený z frontu.

$$s < s_{current} \quad (9)$$

Taktiež, úsek videa pre ktorý platí $s = s_{current}$ je považovaný za neaktuálny, ak sa práve prehrávajúci segment blíži ku koncu a tento úsek by sa nestihol načítať včas pre prehrávanie. Pre takýto prípad platí vzťah:

$$T_s - 2 \cdot \mu_N > t \quad (10)$$

Kde T_s je dĺžka trvania jedného segmentu, t je aktuálny čas prehrávaného segmentu a μ_N je priemerný čas načítavania (prenosu a dekodovania) úseku videa $V_s^{i,l}(t)$ a je vypočítaný ako kľzavý priemer nasledovne:

$$\mu_{N+1} = \frac{N \cdot \mu_N + t_{prepare}}{N+1} \quad (11)$$

Kde N je počet doposiaľ načítaných dlaždíc a $t_{prepare}$ je trvanie prípravy poslednej dlaždice. Popísaný algoritmus načítavania úsekov videa do vyrovnávacej pamäte je zobrazený na obrázku 10.

$M \leftarrow$ Počet dlaždíc v schéme pre tvorbu dlaždíc
 $BV_{current} \leftarrow$ Počet aktuálne načítavaných úsekov videa

Funkcia update je volaná pri každom vykreslení snímky

```

procedure UPDATE()
  for int  $s = s_{current}$  to  $s_{current} + BL$  do
    for int  $i = 1$  to  $M$  do
      if  $d(V_s^{i,0}, Pohľad) < D_{base}$  then
         $PrioritnýFront.Zaraď(V_s^{i,0})$ 
      end if
      if  $d(V_s^{i,1}, Pohľad) < D_{enhancement}$  then
         $PrioritnýFront.Zaraď(V_s^{i,1})$ 
      end if
    end for
  end for
  for each  $V_s^{i,l}$  in  $PrioritnýFront$  do
    if  $s < s_{current}$  then
       $PrioritnýFront.Odstráň(V_s^{i,l})$ 
    else if  $s == s_{current}$  and  $T_s - 2 \cdot \mu_N > t$  then
       $PrioritnýFront.Odstráň(V_s^{i,l})$ 
    end if
  end for
  for int  $i = 1$  to  $BV_{max} - BV_{current}$  do
     $V_s^{i,l} = PrioritnýFront.Vyber()$ 
     $NačítajADekóduj(V_s^{i,l})$ 
  end for
end procedure

procedure VYBER()
   $výsledok = PrioritnýFront[0]$ 
  for each  $V_s^{i,l}$  in  $PrioritnýFront$  do
    if  $P(výsledok) < P(V_s^{i,l})$  then
       $výsledok = V_s^{i,l}$ 
    end if
  end for
   $PrioritnýFront.Odstráň(výsledok)$ 
  return  $výsledok$ 
end procedure

```

Obrázok 10 – Algoritmus načítavania do vyrovnávacej pamäte.

5 Zhrnutie výsledkov práce

5.1 Navrhnutá metóda kompresie videa, ktorá využíva informácie o význačnosti regiónov v obraze

5.1.1 Použité video a informácie o pohľade z eye-trackera

Pre overenie navrhnutej metódy sme použili informácie o pohľade z eye-trackera a sekvencie videa poskytnuté autormi článku [72]. Z poskytnutých dát sme využili 2 sekvencie videa s dĺžkami 18 sekúnd. Prvé video z filmu (Avatar, režirovaný Jamesom Cameronom, 2009) a druhé video z “LIVE Video Database” s názvom “Dance” [98]. Keďže dáta z eye-trackera obsahovali súradnice častejšie ako je frekvencia snímok videa, pre každú snímku sme dáta z eye-trackera spriemerovali tak, že pre každú snímku bude existovať súbor súradníc s polohami pohľadu, pričom jeden používateľ bude prispievať jedným párom súradníc. V tomto prípade boli použité informácie z eye-trackera od 58 používateľov. Sekvencie videa sme spracovali podľa postupu uvedeného v kapitole o metódach, pričom vo vzťahu (1) sme použili štandardné odchýlky $\sigma_h = 0,15$ a $\sigma_h = 0,30$, taktiež sme vytvorili referenčné video bez rozmazávania. Takto vytvorené videá sme kodovali pri bitovom toku 1.5 Mbps s kódrom x264, pričom spolu vzniklo 6 videí (2 rôzne originálne videá, každé s tromi rôznymi reprezentáciami). Príklad aplikácie tohoto postupu je na obrázku 11.

5.1.2 Porovnanie kvality zážitku navrhnutej metódy

Pre porovnanie kvality zážitku pri sledovaní jednotlivých videí sme využili 44 účastníkov vo vekoch medzi 21 a 61. Každý účastník videl každú reprezentáciu videa práve raz, pričom mal za úlohu vybrať najhoršie a najlepšie video z pohľadu kvality zážitku. Výsledky sú uvedené v tabuľke 1.

	Avatar			Dance		
	Najlepšie	Priemer	Najhoršie	Najlepšie	Priemer	Najhoršie
x264	4	6	34	15	14	14
$\sigma_h = 0,30$	11	25	8	18	20	5
$\sigma_h = 0,15$	29	13	2	10	9	24

Tabuľka 1 – Kvalita zážitku používateľa (najlepšia metóda je vyznačená šedou).

V prípade videa „Avatar“ bolo navrhnutou metódou vytvorené video s lepšou kvalitou zážitku pri sledovaní. Video s najlepšou perspektívnou kvalitou malo menší parameter σ_h , video s najnižšou perspektívnou kvalitou bolo video kódované bežným kódrom x264. Tieto výsledky boli testované Chí-kvadrát testom (χ^2 test) s hodnotou 74,73 a prislúchajúcou p -hodnotou $< 0,00001$, čo značí štatisticky vysoko významný vzťah na hladine významnosti $P < 0,05$. V prípade videa “Dance” z databázy “LIVE Video database” sú hodnoty χ^2 štatistiky 19,12 s p -hodnotou 0,000746.

Rozdiel vo výsledkoch týchto dvoch videí mohol byť spôsobený tým, že video “Avatar” má strih každé 2 sekundy, to znamená, že je potrebných viac kľúčových snímok, pričom video “Dance” má stálu scénu s konštantným pozadím, ktoré bolo potrebné kódovať iba raz. Tieto výsledky sú zdokumentované v článku publikovanom na medzinárodnej vedeckej konferencii ELMAR 2018 [AFC02].



Obrázok 11 – Výsledok procesu kódovania snímky s využitím navrhutej metódy: a) originálna snímka, b) originálna snímka so zobrazením polohy pohľadu, c) mapa význačnosti pre snímku vytvorenú z polôh pohľadu, d) snímka po aplikácii Gaussového filtra pre rozmazávanie, e) snímka originálneho videa kódovaného kodekom x264 s bitovým tokom 1,5 Mbps, f) snímka upraveného videa pri $\sigma_h = 0,15$ kódovaného kodekom x264 s bitovým tokom 1,5 Mbps.

5.2 Navrhnutý systém pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa

Navrhli sme systém pre adaptívne doručovanie videa s využitím MPEG-DASH. Tento systém obsahuje tri rôzne režimy adaptácie:

- pasívny režim s využitím vopred získanej mapy význačnosti,
- validovaný pasívny režim – v závislosti na pohľade používateľa sa prepína medzi využívaním pasívneho režimu a režimu bez využitia máp význačnosti,
- režim s individuálnymi regiónmi – používateľovi je prenášané video so zlepšenou kvalitou v oblasti jeho pohľadu.

Výhodou pasívneho režimu je využitie získaných údajov o pohľade používateľa použitím ďalších dvoch navrhnutých režimov, pričom používateľ nepotrebuje eye-tracker a systém vie v prípade menšej šírky pásma ponúkať rovnakú subjektívnu kvalitu. Validovaný pasívny režim umožňuje využiť získané mapy význačnosti, pričom používateľovi je prenášaný rovnaký obsah ako pri pasívnom režime. V prípade, že používateľov pohľad nesúhlasí s pohľadmi iných používateľov, bude mu doručované originálne neupravené video. Tieto dva režimy umožnia ušetriť 23% bitového toku pri zachovaní kvality zážitku používateľa. V prípade režimu s individuálnymi regiónmi je približné šetrenie bitového toku 50%, pričom odozva na zmeny pohľadu používateľa má pomerne dlhé trvanie na úrovni dĺžky vyrovnávacej pamäte. Pri použití vyrovnávacej pamäte s jedným nasledujúcim segmentom a dĺžkou segmentov 1000 ms bude odozva adaptácie minimálne 1000 ms. Porovnanie jednotlivých režimov vzhľadom na nutnosť použitia eye-trackera, odozvy adaptácie a približného šetrenia bitového toku je uvedené v tabuľke 2.

	Eye-tracker	Približné šetrenie bitového toku	Odozva adaptácie
Pasívny režim	Nie	23%	0 ms
Validovaný pasívny režim	Áno	23%	0 ms
Režim s individuálnymi regiónmi	Áno	50%	>1000 ms

Tabuľka 2 – Porovnanie navrhnutých režimov adaptívneho doručovania videa s využitím MPEG-DASH založených na detekcii pohľadu používateľa.

Vyššie popísaná navrhnutá metóda kompresie videa je využiteľná aj pri tomto systéme pre doručovanie videa s využitím MPEG-DASH využívajúcim informácie o význačnosti regiónov v obraze. Návrh uvedeného systému bol odprezentovaný na medzinárodnej konferencii IWSSIP 2018 [AFC01].

5.3 Navrhnutá metóda pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa

Metóda pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa je určená pre systémy doručujúce video individuálnemu používateľovi. Cieľom tejto metódy bolo nájsť priestorový uhol, v ktorom je potrebná plná kvalita videa a stupňa znižovania tejto kvality od centra. Výsledkom použitia tejto metódy je parameter R (polomer kruhu plnej kvality doručovaného videa) a parameter q (parameter určujúci stupeň úbytku kvality).

Vyššie popísaná metóda nebola otestovaná na dostatočnom množstve nezávislých pozorovateľov. Z tohto dôvodu nie je možné vyvodzovať závery týkajúce sa vhodnej oblasti pre dodávanie plnej kvality a gradientu úbytku kvality smerom na perifériu. Návrh uvedenej metódy a systému pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa bol odprezentovaný na konferencii REDŽÚR 2019 [AFD01].

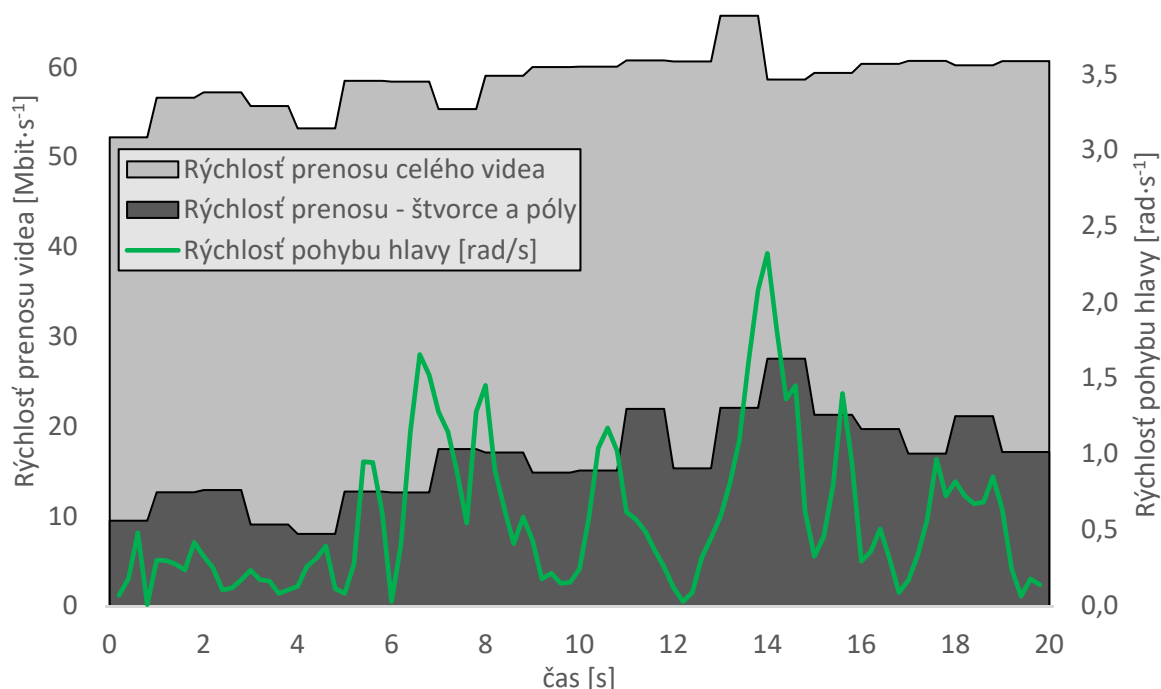
5.4 Navrhnutý systém pre adaptáciu doručovania multimedialného obsahu, ktorý je vhodný pre všesmerové video

Pre navrhnutý systém adaptácie doručovania multimedialného obsahu, ktorý je vhodný pre všesmerové video s využitím priestorových dlaždíc a škálovateľného kódovania videa bola podaná patentová prihláška [AGJ02] a prihláška úžitkového vzoru [AGJ01]. Tento systém bol

experimentálne otestovaný v reálnych podmienkach. Pre simuláciu reálnych podmienok sme použili verejne dostupný dataset pohybov hlavy a očí v prostredí virtuálnej reality [103].

5.4.1 Šetrenie šírky pásma siete

Pre navrhnutý systém adaptácie doručovania všesmerového videa sme monitorovali využitie šírky pásma pri prehrávaní videa pre všetkých používateľov. Príklad využitia šírky pásma pri schéme „štvorce a póly“ pre výrobu dlaždíc v porovnaní s prenosom celého videa je zobrazený na obrázku 12. Tento graf porovnáva aktuálnu rýchlosť prenosu videa a jej vzťah ku rýchlosti pohybu hlavy používateľa a demonštruje väčšiu rýchlosť prenosu pri veľkých zmenách pohľadu.



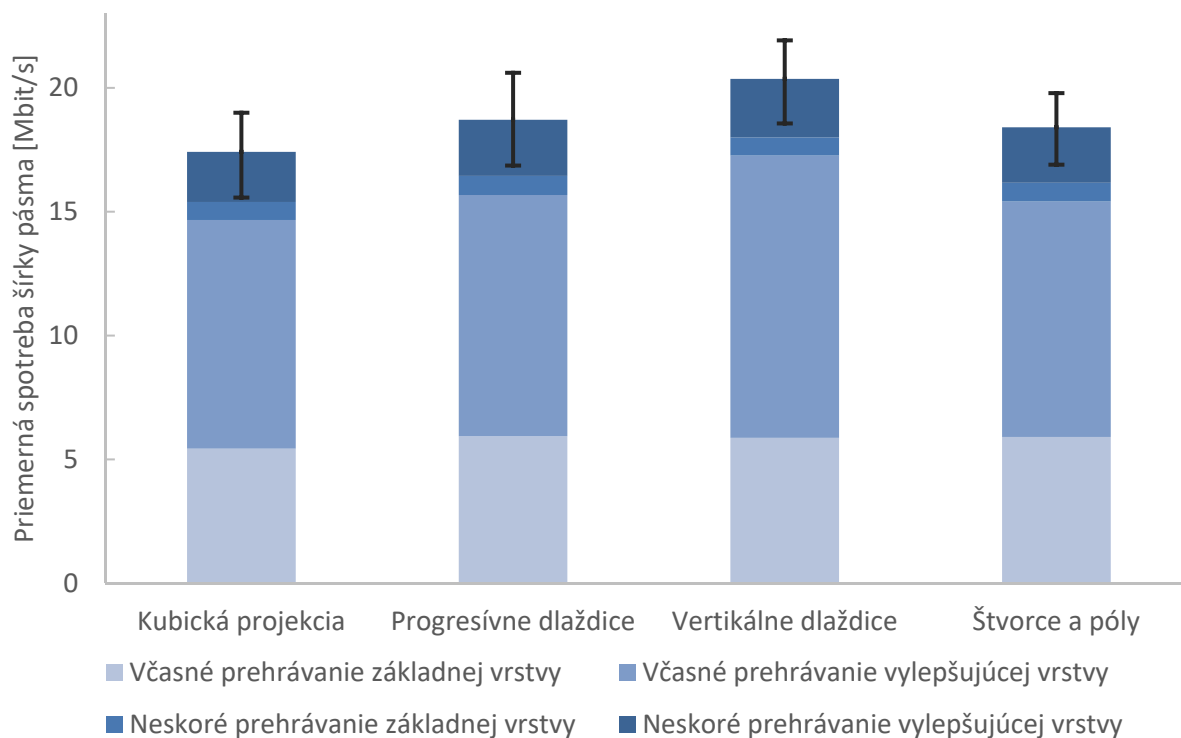
Obrázok 12 – Porovnanie rýchlosti internetového prenosu celého videa v porovnaní s prenosom častí videa so schémou pre tvorbu dlaždíc – štvorce a póly s príslušnou rýchlosťou pohybu hlavy používateľa číslo 4.

Na obrázku 13 je zobrazené porovnanie medzi priemernou spotrebou šírky pásma a štyrmi navrhnutými schémami dlaždíc. Priemerná spotreba šírky pásma je ohraničená horným a dolným kvartilom rýchlosti prenosu pre jednotlivých používateľov. Celkovo sa nám podarilo znížiť spotrebu šírky pásma na úroveň **30 až 35%** oproti spotrebovanej šírke pásma pri tradičnom prenose videa s celým priestorovým uhlom bez adaptácie.

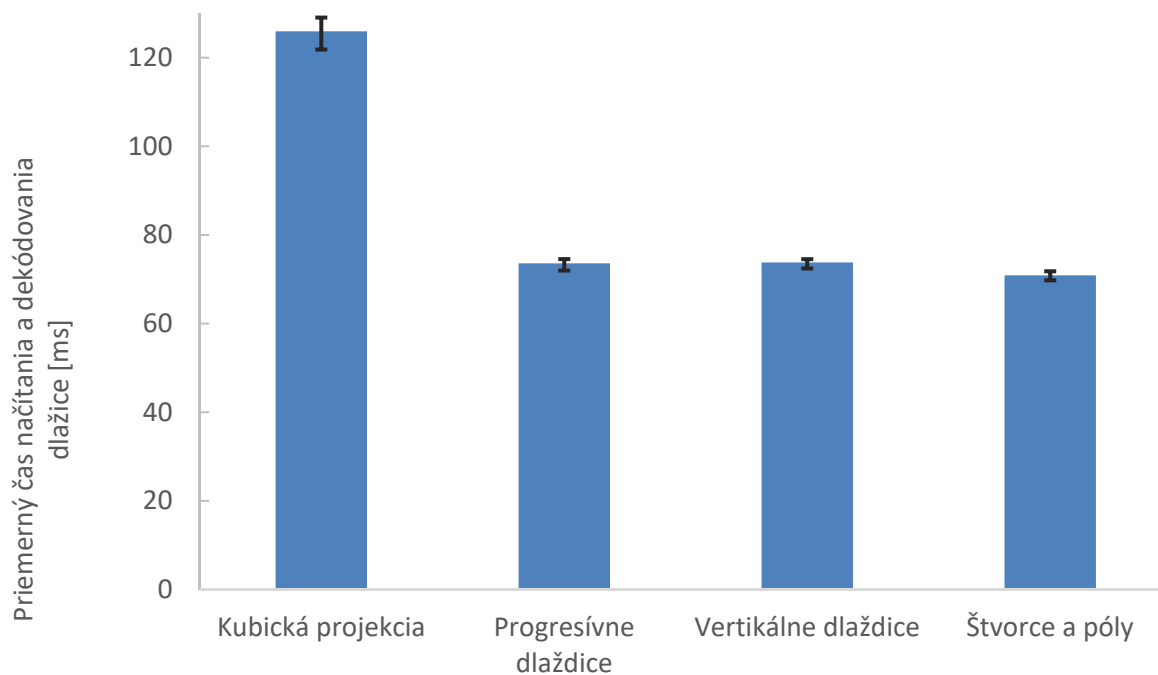
5.4.2 Odozva adaptácie

Pri doručovaní iba časti celej scény je jedným z najdôležitejších parametrov dĺžka trvania adaptácie na zmeny pohľadu používateľa. Graf na obrázku 14 zobrazuje čas, ktorý trvá prenos a dekódovanie jedného úseku videa (jedna vrstva jednej dlaždice) pre jednotlivé typy navrhnutých schém. Tento čas je ohraničený dolným a horným kvartilom času prípravy videa. Dĺžka trvania prenosu a dekódovania dlaždice pri kubickej projekcii bol najvyšší zo všetkých a predstavoval v priemere $125,94 \pm 5,66$ ms. Toto bolo spôsobené tým, že kubická projekcia rozdeľuje video iba na 6 dlaždíc, v porovnaní s progresívnymi dlaždicami (18 dlaždíc), vertikálnymi dlaždicami (18 dlaždíc) a štvorcami a pólmi (20 dlaždíc). Dlaždice ostatných troch schém pre rozdelenie na dlaždice pokrývajú menšiu plochu a teda majú menšiu veľkosť súboru a taktiež je ich možné rýchlejšie

dekódovať. Najnižší čas prípravy videa je $70,87 \pm 2,37$ ms a bol dosiahnutý pri použití schémy štvorce a póly.



Obrázok 13 – Porovnanie priemernej spotreby šírky pásma medzi štyrmi navrhnutými schémami dlaždíc.



Obrázok 14 – Porovnanie priemerného času odozvy adaptácie medzi štyrmi navrhnutými schémami dlaždíc.

6 Záver

Témou tejto dizertačnej práce bola adaptácia doručovania multimedialného obsahu. Ide o problém, ktorý je aktuálny najmä so zvyšujúcou sa kvalitou videa, tým pádom aj bitovým tokom videa a tiež zvyšujúcou sa interaktivitou videa (napríklad všesmerové video). Tento problém je aktuálny aj pre 5G siete, keďže rýchlosť prenosu videa nerastie pri nových generáciách bezdrôtových sietí takým rýchlym tempom, ako neustále zdvojnásobovanie rozlíšenia prenášaného videa. 5G siete ponúknu väčšiu šírku pásma a tiež rýchlejšiu odozvu ako siete predošlých generácií, ale napriek tomu bude potrebné adaptovať doručovanie všesmerového videa s vysokými rozlíšeniami.

Navrhli sme **metódu pre kompresiu videa, ktorá využíva informácie o význačnosti regiónov v obraze**. Táto metóda sa skladá z generovania mapy význačnosti z informácií z eye-trackera použitím sumácie dvojrozmernej Gaussovej funkcie, následným rozmazávaním nevýznačných regiónov použitím Gaussového filtra pred kódovaním a následným kódovaním pomocou kodéra x264 pri zvolenom konštantnom bitovom toku. Túto metódu sme otestovali na 44 nezávislých pozorovateľoch, ktorí hodnotili subjektívnu kvalitu sledovania videa. Metóda bola hodnotená na dvoch rôznych typoch videa, pričom jedno video malo rýchle zmeny scény (3 a menej sekúnd na jednu scénu) a druhé video malo konštantnú scénu s nemeniacim sa pozadím. Ukázali sme, že s využitím tejto metódy je subjektívna kvalita pozorovaného videa vyššia pre väčšinu pozorovateľov v porovnaní s bežným spôsobom kódovania videa s rovnakým bitovým tokom, pričom nami navrhnutá metóda je obzvlášť vhodná pre video s krátkymi scénami, resp. vysokou frekvenciou kľúčových snímkov.

Navrhli sme **systém pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa**. Tento systém využíva video pripravené metódou pre kompresiu videa, ktorá využíva informácie o význačnosti regiónov v obraze. Súčasťou navrhnutého systému sú režimy využívajúce už kódované videá s rôznymi stupňami adaptácie:

- pasívny režim s využitím vopred získanej mapy význačnosti,
- validovaný pasívny režim – v závislosti na pohľade používateľa sa prepína medzi využívaním pasívneho režimu a režimu bez využitia máp význačnosti,
- režim s individuálnymi regiónmi – používateľovi je prenášané video so zlepšenou kvalitou v oblasti jeho pohľadu.

Tieto režimy umožňujú zvýšiť kvalitu v určitých regiónoch videa, pričom je potrebná modifikácia DASH klienta a štruktúry súboru MPD. Tento prístup je vhodný, keď DASH klient potrebuje znížiť bitový tok v prípade zníženej šírky pásma, pričom sa zachováva subjektívna kvalita videa, alebo v prípade, že DASH klient chce zvýšiť kvalitu pozorovaného videa vo význačných regiónoch. Tiež je vhodné poznamenať, že tieto režimy je možné ďalej rozšíriť, napríklad pasívny režim a validovaný pasívny režim by mohli byť rozšírené o rôzne modely význačnosti (dospelí, deti, ženy, muži, ...). Validovaný pasívny režim by mohol byť rozšírený tak, aby bral preferencie používateľa pri hodnotení platnosti modelu. Validovaný pasívny režim a režim individuálnych regiónov môžu byť skombinované, aby bolo lepšie zachytené správanie individuálnych používateľov.

Prvá navrhnutá metóda a systém ponúkajú rozšírenie štandardnej metódy doručovania MPEG-DASH. Problémom týchto opísaných riešení je nedostatočná rýchlosť adaptácie na zmenu pohľadu používateľa najmä v prípade prenosu všesmerového videa. Preto sme navrhli riešenia opísané nižšie.

Navrhli sme **metódu pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa**. Súčasťou navrhutej metódy je aj navrhnutý systém pre simuláciu dokonalého doručovania multimedialného obsahu na vyžiadanie v reálnom čase v prostredí virtuálnej reality, ktorý simuluje adaptáciu na zmenu pohľadu používateľa s odozvou 0 ms. Výsledky z aplikácie tejto metódy môžu slúžiť pri adaptácii doručovania videa v reálnom čase, pričom bude používateľovi

dodávané minimum dát s maximálnou využiteľnosťou. Nami navrhnutý systém umožní nájsť optimálny priestorový uhol pre dodávanie plnej kvality a gradientu znižovania kvality videa. Tento systém ponúka dve premenné, premennú R pre zmenu polomeru plnej kvality videa od centra pohľadu používateľa a premennú q pre definovanie stupňa degradácie kvality smerom od centra. Ukázali sme stavebné bloky systému a navrhli sme metódu pre nerovnomerné rozmazávanie s centrom v určitých súradniciach. Funkcionalita tohoto systému musí byť vyhodnotená na reálnych používateľoch pre nájdenie optimálnych parametrov pre adaptáciu doručovania všesmerového videa.

Navrhli sme **systém pre adaptáciu doručovania multimediálneho obsahu, ktorý je vhodný pre všesmerové video a maximalizuje využiteľnosť prenesených dát individuálnemu používateľovi a algoritmus načítavania dlaždíc do vyrovnávacej pamäte s prioritizáciou dlaždíc pri prenose**. Tento systém je určený najmä pre doručovanie všesmerového videa a využíva časové rozdelenie videa na segmenty s následným priestorovým rozdelením videa na dlaždice, ktoré sú rozdelené do vrstiev škálovateľného videa. Taktiež sme navrhli adaptačný algoritmus, ktorý je zodpovedný za načítavanie najvhodnejších úsekov videa (správna dlaždica a jej daná vrstva) do vyrovnávacej pamäte. Tento navrhnutý systém a algoritmus sme implementovali a otestovali sme túto implementáciu na dátach s polohami hlavy pozbieraných od reálnych používateľov sledujúcich všesmerové video. Navrhli sme štyri schémy pre rozdelenie videa na dlaždice a porovnali sme efektivitu doručovania videa a tiež odozvu na zmenu pohľadu používateľa pri použití týchto schém. Navrhnutý systém bol testovaný pri pokrytí celého priestorového uhla zobrazovaného v okuliároch pre virtuálnu realitu, a v prípade, že používateľ prudko zmení polohu hlavy, bude vidieť základnú vrstvu doručovaného videa, pokiaľ sa nenačíta vylepšujúca vrstva videa. Navrhnutý systém je schopný znížiť spotrebu šírky pásma na 30% v porovnaní so spotrebou šírky pásma pri bežnom prenose všesmerového videa s odozvou adaptácie na zmenu pohľadu používateľa 70,87 ms. Cieľom tejto práce bol návrh systému pre adaptáciu doručovania videa. Navrhnutý systém nie je limitovaný na sledovanie polohy hlavy a je možné ho rozšíriť o eye-tracking, alebo o kombináciu predikčného algoritmu s kombináciou so sledovaním polohy hlavy alebo očí, čím by bolo možné znížiť spotrebu šírky pásma a tiež odozvu adaptácie, čím by sa zvýšila celková subjektívna kvalita zážitku používateľa (QoE). Pri porovnaní nami navrhnutého systému s inými systémami adaptácie doručovania videa sme ukázali, že náš systém má porovnateľné šetrenie šírky pásma s inými systémami s oveľa nižšou odozvou adaptácie na zmeny v pohľade používateľa a tiež ponúka možnosti škálovania počtu súčasne sledujúcich používateľov, keďže kódovanie videa nie je vykonávané na vyžiadanie, ale už sú pripravené kódované úseky videa.

Na základe uvedených výsledkov môžeme usúdiť, že nami navrhnutý systém ponúka konkurencieschopnú adaptáciu doručovania všesmerového videa čo sa týka parametrov šetrenia bitového toku a ponúka výnimočne krátku odozvu na zmeny pohľadu používateľa v porovnaní s inými metódami. Systém taktiež ponúka možnosť škálovateľnosti prenosu pre väčší počet klientov, keďže nekóduje video počas prenosu videa.

7 Prínosy dizertačnej práce

Hlavným cieľom tejto práce bola adaptácia doručovania videa ku koncovému používateľovi tak, aby sa maximalizovala pozorovaná kvalita multimédií pre používateľa a pritom sa znížila bitová rýchlosť prenosu. Tento cieľ je dôležitý najmä v prípade všesmerového videa, kde je väčšina prenášaného bitového toku nevyužitá, pretože nie je v pohľade používateľa.

Na základe uvedených faktov považujem za pôvodné prínosy mojej vedeckej činnosti a tejto dizertačnej práce nasledujúce body:

1. Návrh novej metódy pre kompresiu videa, ktorá využíva informácie o význačnosti regiónov v obraze.
Publikované v [AFC02].
2. Návrh systému pre adaptívne doručovanie videa s využitím MPEG-DASH založený na detekcii pohľadu používateľa.
Publikované v [AFC01].
3. Návrh metódy pre hľadanie optimálnych parametrov adaptácie pre individuálne doručovanie videa (napr. priestorový uhol a stupeň zníženia bitového toku v ňom, ktorý nie je badateľný pre bežného používateľa).
Publikované v [AFD01].
4. Návrh systému pre adaptáciu doručovania multimediálneho obsahu, ktorý je vhodný pre všesmerové video a maximalizuje využiteľnosť prenesených dát individuálnemu používateľovi a algoritmus načítavania dlaždíc do vyrovnávacej pamäte s prioritizáciou dlaždíc pri prenose.
Pre návrh systému bola podaná prihláška úžitkového vzoru [AGJ01].
Pre návrh systému bola podaná patentová prihláška [AGJ02].
Parametre navrhnutého systému boli experimentálne overené a popísané – v čase písania tejto práce je článok v review procese v [ADC01].

8 Résumé

The main goal of this thesis was adaptation of video delivery to the end user at reduced bitrate with maximized quality of experience. This goal is especially important in case of virtual reality videos, where majority of transferred data is not in the user's viewport.

According to the mentioned facts I consider the following as contributions of this dissertation thesis:

1. Proposal of a novel method for video compression, which uses saliency information.
Published in [AFC02].
2. Proposal of a system for adaptive video delivery, which makes use of MPEG-DASH and is based on detection of user gaze.
Published in [AFC01].
3. Proposal of a method for finding optimal parameters of individual multimedia delivery adaptation (e.g. spatial angle and gradient of decreasing quality, which is not observable for the user).
Published in [AFD01].
4. Proposal of a system for adaptation of multimedia delivery, which is suitable for omnidirectional video and maximizes usability of transferred data for individual user, and algorithm for buffering tiles with prioritization of tiles.
For the system proposal we have submitted a utility model application [AGJ01].
For the system proposal we have submitted a patent application [AGJ02].
Parameters of the proposed system were experimentally tested and described – at the time of writing this thesis, the manuscript is in review process in [ADC01].

8 Vlastná publikačná činnosť

8.1 Publikácie týkajúce sa dizertačnej práce

ADC Vedecké práce v zahraničných karentovaných časopisoch

ADC01 POLAKOVIČ, Adam [34%] - ROZINAJ, Gregor - MUNTEAN, Gabriel-Miro. Adaptation of Omnidirectional Video Delivery Using Spatial Tiling and Scalable Video Encoding. In *IEEE Transactions on Multimedia*, 2020, v review procese.

AFC Publikované príspevky na zahraničných vedeckých konferenciách

AFC01 POLAKOVIČ, Adam [34%] - VARGIC, Radoslav - ROZINAJ, Gregor. Adaptive multimedia content delivery in 5G Networks using DASH and saliency information. In *IWSSIP 2018: 25th International conference on systems, signals and image processing. Maribor, Slovenia. June 20-22, 2018*. Piscataway: IEEE, 2018, USB, [5] s. ISBN 978-1-5386-6978-5. V databáze: IEEE: 8439215; WOS: 000451277200008; SCOPUS: 2-s2.0-85053110520. **Kateg. B.**

Citované v:

1. BI, Ting, Fabio SILVA, Gheorghita GHINEA a Gabriel-Mira MUNTEAN. Improving Learning Experience by Employing DASH-Based Multimedia Delivery. In: *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 2018, s. 1-9.
2. CELIKCAN, Ufuk, Mehmet Bahadır ASKIN, Dilara ALBAYRAK a Tolga K. CAPIN. Deep into visual saliency for immersive VR environments rendered in real-time. *Computers & Graphics*. 2020, 88, s. 70-82.
3. ZOLFAGHARI, Behrouz, Gautam SRIVASTAVA, Swapnoneel ROY, et al. Content Delivery Networks. *ACM Computing Surveys*. 2020, 53(2), s. 1-34.

AFC02 POLAKOVIČ, Adam [25%] - VARGIC, Radoslav - ROZINAJ, Gregor - MUNTEAN, Gabriel-Miro. An approach to video compression using saliency based foveation. In *Proceedings ELMAR-2018: 60th International symposium. Zadar, Croatia. September 16-19, 2018*. 1. ed. Zagreb: University of Zagreb, 2018, S. 169-172. ISSN 1334-2630. ISBN 978-953-184-244-0. V databáze: IEEE: 8534631; WOS: 000454262700039; SCOPUS: 2-s2.0-85058711467. **Kateg. B.**

Citované v:

1. LYUDVICHENKO, Vitaliy, Mikhail EROFEEV, Alexander PLOSHKIN a Dmitriy VATOLIN. Improving Video Compression with Deep Visual-attention Models. In: *Proceedings of the 2019 International Conference on Intelligent Medicine and Image Processing - IMIP '19*, New York, USA: ACM Press, 2019, s. 88-94.

AFD Publikované príspevky na domácich vedeckých konferenciách

AFD01 POLAKOVIČ, Adam [50%] - ROZINAJ, Gregor. System for finding optimal parameters of omnidirectional video delivery adaptation. In *Redžúr 2019: 13th International workshop on multimedia information and communication technologies. Bratislava, Slovakia. May 13, 2019*. 1. vyd. Bratislava: Vydavateľstvo Spektrum STU, 2019, S. 39-42. ISBN 978-80-227-4931-2. **Kateg. C.**

AGJ Autorské osvedčenia, patenty, objavy

AGJ01 POLAKOVIČ, Adam [25%] - ROZINAJ, Gregor - VARGIC, Radoslav - MUNTEAN, Gabriel-Miro. *Systém a spôsob adaptívneho doručovania všesmerového videa: prihláška úžitkového vzoru č. 134-2019, dátum podania prihlášky: 09.09.2019, stav: v konaní*. Banská Bystrica: Úrad priemyselného vlastníctva SR, 2019. Dostupné na internete: <<https://wbr.indprop.gov.sk/WebRegistre/UzitkovyVzor/Detail/134-2019>>.

AGJ02 POLAKOVIČ, Adam [25%] - ROZINAJ, Gregor - VARGIC, Radoslav - MUNTEAN, Gabriel-Miro. *Systém a spôsob adaptívneho doručovania všesmerového videa: prihláška patentu č. 101-2019, dátum podania prihlášky: 09.09.2019, stav: v konaní*. Banská Bystrica: Úrad priemyselného vlastníctva SR, 2019. Dostupné na internete: <<https://wbr.indprop.gov.sk/WebRegistre/Patent/Detail/101-2019>>.

8.2 Ostatné publikácie

AFC Publikované príspevky na zahraničných vedeckých konferenciách

AFC03 ROZINAJ, Gregor - VANČO, Marek - VARGIC, Radoslav - MINÁRIK, Ivan - POLAKOVIČ, Adam [20%]. Augmented/virtual reality as a tool of self-directed learning. In *IWSSIP 2018: 25th International conference on systems, signals and image processing*. Maribor, Slovenia. June 20-22, 2018. Piscataway: IEEE, 2018, USB, [5] s. ISBN 978-1-5386-6978-5. V databáze: IEEE: 8439309; WOS: 000451277200018; SCOPUS: 2-s2.0-85053125363. **Kateg. B.**

AFD Publikované príspevky na domácich vedeckých konferenciách

AFD02 POLAKOVIČ, Samuel – POLAKOVIČ, Adam [33%] – ROZINAJ, Gregor. In *Redžúr 2020: 14th International workshop on multimedia information and communication technologies*. Bratislava, Slovakia. May 29, 2020. V procese vydávania. **Kateg. C.**

AFG Abstrakty príspevkov zo zahraničných vedeckých konferencií

AFG01 SERSENOVÁ, Dominika - GBELCOVÁ, Helena – POLAKOVIČ, Adam [20%] – REPISKÁ, Vanda – MACHALA, Zdenko -. Effect of plasma pen treatment and plasma activated medium (PAM) on cancer and normal cells, *International Workshop on Plasma for Cancer Treatment*, Clinical Plasma Medicine: Supplement, Vol. 9, Greifswald (Germany), Marec 20-21, p. 24 (2018).

AFG02 POLAKOVIČ, Adam [20%] – ŠTRÍŽENCOVÁ, Lucia - KOTRBANCOVÁ, Michaela – PETRŽELOVÁ, Jana – FULOVÁ, Miriam – ŠPALEKOVÁ, Margita – MARTIŠOVITŠ, Viktor – MACHALA, Zdenko -. Cold air plasma source for biomedical applications based on DC corona discharge, *6th International Conference on Plasma Medicine (ICPM-6)*, Bratislava (Slovakia), September 4-9, p. 300 (2016).

AFG03 KUČEROVÁ, Katarína – POLAKOVIČ, Adam [20%] - GBELCOVÁ, Helena – JANDA, Mário – MACHALA, Zdenko – HENSEL, Karol – JIJIE, R. – MIHAI, C. T. – TOPALA, I. -. Indirect treatment of cancer cells by air spark and corona discharge, *4th International Workshop on Plasma for Cancer Treatment IWPCT-2017*, Paris (France), March 27–28, p. 57 (2017).

BCI Skriptá a učebné texty

BCI01 PODHRADSKÝ, Pavol – MEDVECKÝ, Martin – TRÚCHLY, Peter - VARGIC, Radoslav - LONDÁK, Juraj - KADLIC, Radovan - SCHUMANN, Sebastian - SALAZAR, Jordi - SANTIAGO, Silvestre - LEVICKÝ, Dušan, - POLAKOVIČ, Adam [5%]. *Sieťové architektúry, služby a aplikácie*. Bratislava: SPEKTRUM STU, 2019. 120 s. ISBN 978-80-227-5000-4.

9 Riešené výskumné projekty

H2020 NEWTON No. 688503 - Networked Labs for Training in Sciences and Technologies for Information and Communication. Koordinátor pre STU prof. Ing. Gregor Rozinaj, PhD, H2020 ICT-20 2015 - Information and Communications Technologies, Technologies for better human learning and teaching (2016-2019).

VEGA INOMET No. 1/0800/16 - Inovatívne metódy spracovania multimediálnych signálov pre inteligentné systémy a služby. Zodpovedný riešiteľ prof. Ing. Gregor Rozinaj, PhD. (2016-2018).

PPMV 1323PS/2019, Adaptácia doručovania multimediálneho obsahu, zodpovedný riešiteľ Mgr. Adam Polakovič (2019).

10 Zoznam použitej literatúry

- [1] ANDERSON, Robert, David GALLUP, Jonathan T. BARRON, Janne KONTKANEN, Noah SNAVELY, Carlos HERNÁNDEZ, Sameer AGARWAL a Steven M. SEITZ. Jump: virtual reality video. *ACM Transactions on Graphics* 2016, 35(6), 1-13 DOI: 10.1145/2980179.2980257. ISSN 0730-0301.
- [2] ALBERT, Rachel, Anjul PATNEY, David LUEBKE a Joohwan KIM. Latency Requirements for Foveated Rendering in Virtual Reality. *ACM Transactions on Applied Perception*, 14(4), 1-13. DOI: 10.1145/3127589. ISSN 1544-3558.
- [3] HUNT, Cale. Field of view face-off: Rift vs Vive vs Gear VR vs PSVR. [online]. 25 Júl 2016. [cit. 2020-05-15]. Dostupné na internete: <https://www.vrheads.com/field-view-faceoff-rift-vs-vive-vs-gear-vr-vs-psvr>
- [4] THIBOS, L. N., F. E. CHENEY a D. J. WALSH. Retinal limits to the detection and resolution of gratings. *Journal of the Optical Society of America A*. 1987, 4(8). DOI: 10.1364/JOSAA.4.001524. ISSN 1084-7529.
- [5] BLAKE, Randolph a SEKULER, Robert. *Perception*. Boston : McGraw-Hill, 2006.
- [6] NIEHORSTER, Diederick C., Li LI a Markus LAPPE. The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research. *I-Perception*. 2017, 8(3). DOI: 10.1177/2041669517708205. ISSN 2041-6695.
- [7] JERALD, Jason J., *Scene-Motion- and Latency-Perception Thresholds for Head-Mounted Displays*. Dizertačná práca, University of North Carolina at Chapel Hill, 2009.
- [8] ZHENG, Feng, Turner WHITTED, Anselmo LASTRA, Peter LINCOLN, Andrei STATE, Andrew MAIMONE a Henry FUCHS. Minimizing latency for augmented reality displays: Frames considered harmful. In: *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 2014, s. 195-200. DOI: 10.1109/ISMAR.2014.6948427. ISBN 978-1-4799-6184-9.
- [9] KENNEDY, Robert S., Norman E. LANE, Kevin S. BERBAUM a Michael G. LILIENTHAL. Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*. 1993, 3(3), 203-220. DOI: 10.1207/s15327108ijap0303_3. ISSN 1050-8414.
- [10] BROWN, Chip. Bringing pixels front and center in VR video. *Google* [online]. 14 March 2017. [cit.2020-05-15]. Dostupné na internete: <https://www.blog.google/products/google-ar-vr/bringing-pixels-front-and-center-vr-video/>
- [11] Upload virtual reality videos, 2018, *Google Inc.*, [cit. 2019-01-15]. Dostupné na internete: https://support.google.com/youtube/answer/6316263?hl=en&ref_topic=2888648.
- [12] Rendering Omni-directional Stereo Content. *Google Inc.*, [cit. 2019-01-15]. Dostupné na internete: <https://developers.google.com/vr/jump/rendering-ods-content.pdf>.
- [13] ACHARYA U., Rajenda, Eddie Y. K. NG a Jasjit S. SURI. *Image Modeling of the Human Eye*. Artech House, 2008, 359 s. ISBN 978-1596932081.
- [14] SNELL, Richard S. a Michael A. LEMP. *Clinical Anatomy of the Eye*, 2nd Edition. New York: Wiley-Blackwell, 2013, 432 s. ISBN 978-1-118-69100-7.
- [15] LÖFGREN, Stefan, Jörgen THAUNG a Cesar LOPES. Laser pointers and eye injuries. An analysis of reported cases. Strålsäkerhetsmyndigheten, 2013. ISSN 2000-0456.
- [16] HENDERSON, John M. a Andrew HOLLINGWORTH. Eye Movements During Scene Viewing. *Eye Guidance in Reading and Scene Perception*. Elsevier, 1998, 1998, s. 269-293 DOI: 10.1016/B978-008043361-5/50013-4. ISBN 9780080433615.
- [17] DODGE, Raymond. An experimental study of visual fixation. *The Psychological Review: Monograph Supplements*. 1907, 8(4), i-95 DOI: 10.1037/h0093042. ISSN 0096-9753.
- [18] DITCHBURN, R.W. The function of small saccades. *Vision Research*. 1980, 20(3), 271-272. DOI: 10.1016/0042-6989(80)90112-1. ISSN 00426989.
- [19] YOUNG, Laurence R. a David SHEENA. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*. 1975, 7(5), 397-429. DOI: 10.3758/BF03201553. ISSN 1554-351X9.
- [20] SHARAFI, Zohreh, Timothy SHAFFER, Bonita SHARIF a Yann-Gael GUEHENEUC. Eye-Tracking Metrics in Software Engineering. In: *2015 Asia-Pacific Software Engineering Conference (APSEC)*. IEEE, 2015, s. 96-103. DOI: 10.1109/APSEC.2015.53. ISBN 978-1-4673-9644-8.
- [21] TOATES, F. M. Accommodation function of the human eye. *Physiological Reviews*. 1972, 52(4), 828-863. DOI: 10.1152/physrev.1972.52.4.828. ISSN 0031-9333.
- [22] KRUGER, Philip B., Steven MATHEWS, Karan R. AGGARWALA, Dean YAGER a Ekaterina S. KRUGER. Accommodation responds to changing contrast of long, middle and short spectral-waveband components of the retinal image. *Vision Research*. 1995, 35(17), 2415-2429. DOI: 10.1016/0042-6989(94)00316-5. ISSN 00426989.
- [23] HUNG, George K. Adaptation model of accommodation and vergence. *Ophthalmic and Physiological Optics*. 1992, 12(3), 319-326. DOI: 10.1111/j.1475-1313.1992.tb00404.x. ISSN 0275-5408.
- [24] CAMPBELL, F. W. a G. WESTHEIMER. Dynamics of accommodation responses of the human eye. *The Journal of Physiology*. 1960, 151(2), 285-295. DOI: 10.1113/jphysiol.1960.sp006438. ISSN 00223751.
- [25] VARGIC, Radoslav, Júlia KUCEROVÁ a Jaroslav POLEC. Wavelet-based image coding using saliency map. *Journal of Electronic Imaging*. 2016, 25(6). DOI: 10.1117/1.JEI.25.6.061610. ISSN 1017-9909.

- [26] SHRIVASTAVA, Abhinav, Tomasz MALISIEWICZ, Abhinav GUPTA a Alexei A. EFROS. Data-driven visual similarity for cross-domain image matching. *ACM Transactions on Graphic.* 2011, 30(6), 1-10. DOI: 10.1145/2070781.2024188. ISSN 0730-0301.
- [27] ĆULIBRK, Dubravko, Milan MIRKOVIĆ, Vladimir ZLOKOLICA, Maja POKRIĆ, Vladimir CRNOJEVIĆ a Dragan KUKOLJ. Salient Motion Features for Video Quality Assessment. *IEEE Transactions on Image Processing.* 2011, 20(4), 948-958. DOI: 10.1109/TIP.2010.2080279. ISSN 1057-7149.
- [28] GOFERMAN, Stas, Lihi ZELNIK-MANOR a Ayellet TAL. Context-Aware Saliency Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2012, 34(10), 1915-1926. DOI: 10.1109/TPAMI.2011.272. ISSN 0162-8828.
- [29] BALUJA. Shumeet a Dean POMERLAU, Using a Saliency Map for Active Spatial Selective Attention: Implementation & Initial Results, In: *Proceedings of the 7th International Conference on Neural Information Processing Systems.* 1994, s. 451-458.
- [30] ZHAO, Q. a C. KOCH. Learning a saliency map using fixated locations in natural scenes. *Journal of Vision.* 2011, 11(3), 9-9. DOI: 10.1167/11.3.9. ISSN 1534-7362.
- [31] TSOTSOS, John K. *A Computational Perspective on Visual Attention.* The MIT Press, 2011. DOI: 10.7551/mitpress/9780262015417.001.0001. ISBN 9780262295420.
- [32] ITTI, L., C. KOCH a E. NIEBUR. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 20(11), 1254-1259. DOI: 10.1109/34.730558. ISSN 01628828.
- [33] ITTI, Laurent, Bruno BOSACCHI, David B. FOGEL, Nitin DHAVALA, Frederic PIGHIN a James C. BEZDEK. Realistic avatar eye and head animation using a neurobiological model of visual attention. *Proceedings of SPIE - The International Society for Optical Engineering*, s. 64. DOI: 10.1117/12.512618.
- [34] TORRALBA, Antonio. Modeling global scene factors in attention. *Journal of the Optical Society of America A.* 2003, 20(7). DOI: 10.1364/JOSAA.20.001407. ISSN 1084-7529.
- [35] OLIVA, A., A. TORRALBA, M.S. CASTELHANO a J.M. HENDERSON. Top-down control of visual attention in object detection. In: *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429).* IEEE, 2003, 1-253-6. DOI: 10.1109/ICIP.2003.1246946. ISBN 0-7803-7750-8.
- [36] ROSENHOLTZ, Ruth. A simple saliency model predicts a number of motion popout phenomena. *Vision Research.* 1999, 39(19), 3157-3163. DOI: 10.1016/S0042-6989(99)00077-2. ISSN 00426989.
- [37] BRUCE, Neil D. B. a John K. TSOTSOS. Saliency based on information maximization, In: *NIPS'05 Proceedings of the 18th International Conference on Neural Information Processing Systems*, 2005, s. 155-162.
- [38] HOU, Xiaodi a Liqing ZHANG. Saliency Detection: A Spectral Residual Approach. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE, 2007, 2007, s. 1-8. DOI: 10.1109/CVPR.2007.383267. ISBN 1-4244-1179-3.
- [39] ACHANTA, Radhakrishna, Sheila HEMAMI, Francisco ESTRADA a Sabine SUSSTRUNK. Frequency-tuned salient region detection. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE, 2009, s. 1597-1604. DOI: 10.1109/CVPR.2009.5206596. ISBN 978-1-4244-3992-8.
- [40] KIENZLE, W., M. O. FRANZ, B. SCHOLKOPF a F. A. WICHMANN. Center-surround patterns emerge as optimal predictors for human saccade targets. *Journal of Vision.* 2009, 9(5), 7-7. DOI: 10.1167/9.5.7. ISSN 1534-7362.
- [41] PETERS, Robert J. a Laurent ITTI. Beyond bottom-up: Incorporating task-dependent influences into a computational model of spatial attention. In: *2007 IEEE Conference on Computer Vision and Pattern Recognition.* IEEE, 2007, s. 1-8. DOI: 10.1109/CVPR.2007.383337. ISBN 1-4244-1179-3.
- [42] JUDD, Tilke, Krista EHINGER, Fredo DURAND a Antonio TORRALBA. Learning to predict where humans look. In: *2009 IEEE 12th International Conference on Computer Vision.* IEEE, 2009, s. 2106-2113. DOI: 10.1109/ICCV.2009.5459462. ISBN 978-1-4244-4420-5.
- [43] AL-RAHAYFEH, Amer a Miad FAEZIPOUR. Eye Tracking and Head Movement Detection: A State-of-Art Survey. *IEEE Journal of Translational Engineering in Health and Medicine.* 2013, 1, 2100212-2100212. DOI: 10.1109/JTEHM.2013.2289879. ISSN 2168-2372.
- [44] GIPS, James, Philip DIMATTIA, Francis X. CURRAN a Peter OLIVIERI. Using EagleEyes—an electrodes based device for controlling the computer with your eyes—to help people with special needs. In: *Proceedings of the 5th International conference on Computers helping people with special needs.* 1996, s. 77-84.
- [45] RAUDONIS, Vidas, Rimvydas SIMUTIS a Gintautas NARVYDAS. Discrete eye tracking for medical applications. In: *2009 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies.* IEEE, 2009, s. 1-6. DOI: 10.1109/ISABEL.2009.5373675. ISBN 978-1-4244-4640-7.
- [46] LI, Xiaokun a William G. WEE. An efficient method for eye tracking and eye-gazed FOV estimation. In: *2009 16th IEEE International Conference on Image Processing (ICIP).* IEEE, 2009, 2009, s. 2597-2600. DOI: 10.1109/ICIP.2009.5413997. ISBN 978-1-4244-5653-6.

- [47] YANG, Xiaohui, Jiande SUN, Ju LIU, Jinyu CHU, Wei LIU a Yuling GAO. A gaze tracking scheme for eye-based intelligent control. In: *2010 8th World Congress on Intelligent Control and Automation*. IEEE, 2010, 2010, s. 50-55. DOI: 10.1109/WCICA.2010.5553932. ISBN 978-1-4244-6712-9.
- [48] CHEN, Yen-Wei a Kenji KUBO. A Robust Eye Detection and Tracking Technique Using Gabor Filters. In: *Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007)*. IEEE, 2007, 2007, s. 109-112. DOI: 10.1109/IIH-MSP.2007.58.
- [49] YOO, Dong Hyun, Jae Heon KIM, Bang Rae LEE a Myoung Jin CHUNG. Non-contact eye gaze tracking system by mapping of corneal reflections. In: *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. IEEE, 2002, s. 101-106. DOI: 10.1109/AFGR.2002.1004139. ISBN 0-7695-1602-5.
- [50] YOO, Dong Hyun a Myung Jin CHUNG. A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Computer Vision and Image Understanding*. 2005, 98(1), 25-51. DOI: 10.1016/j.cviu.2004.07.011. ISSN 10773142.
- [51] HANSEN, D.W. a QIANG JI. In the Eye of the Beholder: A Survey of Models for Eyes and Gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010, 32(3), 478-500. DOI: 10.1109/TPAMI.2009.30. ISSN 0162-8828.
- [52] KASSNER, Moritz, William PATERA a Andreas BULLING. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct Publication - UbiComp '14 Adjunct*. New York, New York, USA: ACM Press, 2014, 2014, s. 1151-1160. DOI: 10.1145/2638728.2641695. ISBN 9781450330473.
- [53] Recommendation ITU-R M.2083-0 (2015-09),: *IMT Vision – Framework and overall objectives of the future development of IMT for 2020 and beyond* [online]. ITU-R, 2015 [cit. 2019-01-15]. Dostupné z: https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf
- [54] *View on 5G Architecture (Version 2.0)* [online]. 5G PPP Architecture Working Group, 2017 [cit. 2019-01-15]. Dostupné z: https://5g-ppp.eu/wp-content/uploads/2017/07/5G-PPP-5G-Architecture-White-Paper-2-Summer-2017_For-Public-Consultation.pdf
- [55] *Minimum requirements related to technical performance for IMT-2020 radio interface(s): Draft new Report ITU-R M. [IMT-2020.TECH PERF REQ]* [online]. ITU-R, 2017 [cit. 2019-01-15]. Dostupné z: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2410-2017-PDF-E.pdf.
- [56] *3GPP TS 23.501 V15.2.0 (2018-06): System Architecture for the 5G System* [online]. 3GPP, 2018 [cit. 2019-01-15]. Dostupné z: https://www.etsi.org/deliver/etsi_ts/123500_123599/123501/15.02.00_60/ts_123501v150200p.pdf.
- [57] *3GPP TS 23.502 V15.2.0 (2018-06): Procedures for the 5G System* [online]. 3GPP, 2018 [cit. 2019-01-15]. Dostupné z: https://www.etsi.org/deliver/etsi_ts/123500_123599/123502/15.02.00_60/ts_123502v150200p.pdf.
- [58] *3GPP TS 23.503 V15.2.0 (2018-07): Policy and Charging Control Framework for the 5G System; Stage 2* [online]. 3GPP, 2018 [cit. 2019-01-15]. Dostupné z: https://www.etsi.org/deliver/etsi_ts/123500_123599/123503/15.02.00_60/ts_123503v150200p.pdf.
- [59] ZHANG, Haijun, Na LIU, Xiaoli CHU, Keping LONG, Abdol-Hamid AGHVAMI a Victor C. M. LEUNG. Network Slicing Based 5G and Future Mobile Networks: Mobility, Resource Management, and Challenges. *IEEE Communications Magazine*. 2017, 55(8), 138-145. DOI: 10.1109/MCOM.2017.1600940. ISSN 0163-6804.
- [60] *Network Slicing for 5G Networks & Services, 5G Americas White Paper* [online]. 5G Americas, 2016 [cit. 2020-05-15]. Dostupné z: https://www.5gamericas.org/wp-content/uploads/2019/07/5G_Americas_Network_Slicing_11.21_Final.pdf.
- [61] GUTZ, Stephen, Alec STORY, Cole SCHLESINGER a Nate FOSTER. Splendid isolation. In: *Proceedings of the first workshop on Hot topics in software defined networks - HotSDN '12*. New York, New York, USA: ACM Press, 2012, 2012, s. 79. DOI: 10.1145/2342441.2342458. ISBN 9781450314770.
- [62] *ETSI GS NFV-MAN 001 V1.1.1 (2014-12): Network Functions Virtualisation (NFV); Management and Orchestration* [online]. ETSI, 2014 [cit. 2019-01-15]. Dostupné z: https://www.etsi.org/deliver/etsi_gs/NFV-MAN/001_099/001/01.01.01_60/gs_NFV-MAN001v010101p.pdf.
- [63] *The Zettabyte Era: Trends and Analysis* [online]. 07.06.2017 [cit. 2019-01-15]. Dostupné z: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni-hyperconnectivity-wp.html>.
- [64] CHENLEI GUO a LIMING ZHANG. A Novel Multiresolution Spatiotemporal Saliency Detection Model and Its Applications in Image and Video Compression. *IEEE Transactions on Image Processing*. 2010, 19(1), 185-198. DOI: 10.1109/TIP.2009.2030969. ISSN 1057-7149.
- [65] ITTI, L. Automatic Foveation for Video Compression Using a Neurobiological Model of Visual Attention. *IEEE Transactions on Image Processing*. 2004, 13(10), 1304-1318. DOI: 10.1109/TIP.2004.834657. ISSN 1057-7149.
- [66] LU, Shao-Ping a Song-Hai ZHANG. Saliency-Based Fidelity Adaptation Preprocessing for Video Coding. *Journal of Computer Science and Technology*. 2011, 26(1), 195-202. DOI: 10.1007/s11390-011-9426-5. ISSN 1000-9000.
- [67] LI, Zhicheng, Shiyin QIN a Laurent ITTI. Visual attention guided bit allocation in video compression. *Image and Vision Computing*. 2011, 29(1), 1-14. DOI: 10.1016/j.imavis.2010.07.001. ISSN 02628856.

- [68] GUPTA, Rupesh, Meera THAPAR KHANNA a Santanu CHAUDHURY. Visual saliency guided video compression algorithm. *Signal Processing: Image Communication*. 2013, 28(9), 1006-1022. DOI: 10.1016/j.image.2013.07.003. ISSN 09235965.
- [69] HADIZADEH, Hadi a Ivan V. BAJIC. Saliency-preserving video compression. In: *2011 IEEE International Conference on Multimedia and Expo*. IEEE, 2011, s. 1-6. DOI: 10.1109/ICME.2011.6012194. ISBN 978-1-61284-348-3.
- [70] HADIZADEH, Hadi. *Visual saliency in video compression and transmission*, dizertačná práca, Simon Fraser University, 2013.
- [71] GITMAN, Yury, Mikhail EROFEEV, Dmitriy VATOLIN, Bolshakov ANDREY a Fedorov ALEXEY. Semiautomatic visual-attention modeling and its application to video compression. In: *2014 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2014, s. 1105-1109. DOI: 10.1109/ICIP.2014.7025220. ISBN 978-1-4799-5751-4.
- [72] LYUDVICHENKO, Vitaliy, Mikhail EROFEEV, Yury GITMAN a Dmitriy VATOLIN. A semiautomatic saliency model and its application to video compression. In: *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)*. IEEE, 2017, s. 403-410. DOI: 10.1109/ICCP.2017.8117038. ISBN 978-1-5386-3368-7.
- [73] ISO/IEC 23009-1:2014 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats [online]. ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information, 2014 [cit. 2019-01-15]. Dostupné z: <https://www.iso.org/standard/65274.html>.
- [74] SODAGAR, Iraj. The MPEG-DASH Standard for Multimedia Streaming Over the Internet. *IEEE Multimedia*. 2011, 18(4), 62-67. DOI: 10.1109/MMUL.2011.71. ISSN 1070-986X.
- [75] MPEG-DASH / Media Source demo [online]. [cit. 2019-01-15] Dostupné na: <http://dash-mse-test.appspot.com/media.html>.
- [76] KUZYAKOV, Evgeny a David PIO. Next-generation video encoding techniques for 360 video and VR. *Facebook Engineering* [online]. 21.01.2016 [cit. 2020-05-15]. Dostupné z: <https://engineering.fb.com/virtual-reality/next-generation-video-encoding-techniques-for-360-video-and-vr/>.
- [77] CORBILLON, Xavier, Gwendal SIMON, Alisa DEVLIC a Jacob CHAKARESKEI. Viewport-adaptive navigable 360-degree video delivery. In: *2017 IEEE International Conference on Communications (ICC)*. IEEE, 2017, s. 1-7. DOI: 10.1109/ICC.2017.7996611. ISBN 978-1-4673-8999-0.
- [78] NIAMUT, Omar A., Emmanuel THOMAS, Lucia D'ACUNTO, Cyril CONCOLATO, Franck DENOUAL a Seong Yong LIM. MPEG DASH SRD. In: *Proceedings of the 7th International Conference on Multimedia Systems - MMSys '16*. New York, New York, USA: ACM Press, 2016, 2016, s. 1-8. DOI: 10.1145/2910017.2910606. ISBN 9781450342971.
- [79] ISO/IEC 23009-1:2014/AMD 2:2015 Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats — Amendment 2: Spatial relationship description, generalized URL parameters and other extensions [online]. ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information, 2015. [cit. 2019-01-15]. Dostupné z: <https://www.iso.org/standard/66486.html>.
- [80] HOSSEINI, Mohammad a Viswanathan SWAMINATHAN. Adaptive 360 VR Video Streaming: Divide and Conquer. In: *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2016, s. 107-110. DOI: 10.1109/ISM.2016.0028. ISBN 978-1-5090-4571-6.
- [81] HOSSEINI, Mohammad a Viswanathan SWAMINATHAN. Adaptive 360 VR Video Streaming based on MPEG-DASH SRD. In: *2016 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2016, s. 407-408.
- [82] SKUPIN, Robert, Yago SANCHEZ, Dimitri POBORSKI, Cornelius HELLGE a Thomas SCHIERL. HEVC tile based streaming to head mounted displays. In: *2017 14th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 2017, s. 613-615. DOI: 10.1109/CCNC.2017.7983191. ISBN 978-1-5090-6196-9.
- [83] ZHAO, Shuai a Deep MEDHI. SDN-Assisted adaptive streaming framework for tile-based immersive content using MPEG-DASH. In: *2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN)*. IEEE, 2017, s. 1-6. DOI: 10.1109/NFV-SDN.2017.8169831. ISBN 978-1-5386-3285-7.
- [84] GUENTER, Brian, Mark FINCH, Steven DRUCKER, Desney TAN a John SNYDER. Foveated 3D graphics. *ACM Transactions on Graphics*. 2012, 31(6), 1-10. DOI: 10.1145/2366145.2366183. ISSN 0730-0301.
- [85] PATNEY, Anjul, Joohwan KIM, Marco SALVI, Anton KAPLANYAN, Chris WYMAN, Nir BENTY, Aaron LEFOHN a David LUEBKE. Perceptually-based foveated virtual reality. In: *ACM SIGGRAPH 2016 Emerging Technologies on - SIGGRAPH '16*. New York, USA: ACM Press, 2016, s. 1-2. DOI: 10.1145/2929464.2929472. ISBN 9781450343725.
- [86] SARAIJI, MHD. Yamen, Kouta MINAMIZAWA, Susumu TACHI, Foveated Streaming: Optimizing video streaming for Telexistence systems using eye-gaze based foveation. *Virtual Reality Society of Japan*, September 2017.

- [87] ILLAHI, Gazi, Matti SIEKKINEN a Enrico MASALA. Foveated video streaming for cloud gaming. In: *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2017, s. 1-6. DOI: 10.1109/MMSP.2017.8122235. ISBN 978-1-5090-3649-3.
- [88] ROMERO-RONDÓN, Miguel Fabian, Lucile SASSATELLI, Frédéric PRECIOSO a Ramon APARICIO-PARDO. Foveated streaming of virtual reality videos. In: *Proceedings of the 9th ACM Multimedia Systems Conference*. New York, NY, USA: ACM, 2018, 2018-06-12, s. 494-497. DOI: 10.1145/3204949.3208114. ISBN 9781450351928.
- [89] ISO/IEC 13818-2:1996 *Information technology — Generic coding of moving pictures and associated audio information: Video*. ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information. 1996. Dostupné z: <https://www.iso.org/standard/22990.html>.
- [90] *Video Coding for Low Bit Rate communication*, ITU-T Rec. H.263, ITU-T, Version 3: Nov.2000.
- [91] ISO/IEC 14496-2:2004 *Information technology — Coding of audio-visual objects — Part 2: Visual*. ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information. 2004, 706 s.
- [92] WIEGAND, T., G. J. SULLIVAN, J. REICHEL, H. SCHWARZ, a M. WIEN, Joint Draft 11 of SVC Amendment. *Joint Video Team*, Doc. JVT-X201, 2007.
- [93] ISO/IEC 23008-2:2017 *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*, ISO/IEC JTC 1/SC 29 Coding of audio, picture, multimedia and hypermedia information, 2017, 796 s.
- [94] SCHWARZ, H., D. MARPE a T. WIEGAND. Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Transactions on Circuits and Systems for Video Technology*. 2007, 17(9), 1103-1120. DOI: 10.1109/TCSVT.2007.905532. ISSN 1051-8215.
- [95] UNANUE, Iraide, Inigo URTEAGA, Ronaldo HUSEMANN, Javier DEL, Valter ROESLER, Aitor RODRIGUEZ a Pedro SANCHEZ. A Tutorial on H.264/SVC Scalable Video Coding and its Tradeoff between Quality, Coding Efficiency and Performance. DEL SER LORENTE, Javier, ed. *Recent Advances on Video Coding*. InTech, 2011. DOI: 10.5772/19227. ISBN 978-953-307-181-7.
- [96] POLEC, Jaroslav, Jarmila PAVLOVÍČOVÁ, Tatiana KARLUBÍKOVÁ. *Medzinárodné štandardy pre kompresiu obrazu II : H.261, MPEG-1, MPEG-2, H.263, MPEG-4*. Bratislava: FEI STU, 2002. 151 s. ISBN 80-227-1784-3..
- [97] VARGIC, Radoslav, Martin MEDVECKÝ, Juraj LONDÁK a Pavol PODHRADSKÝ. Advanced Interactive Multimedia Delivery in 5G Networks. *Advances in Intelligent Systems and Computing*. 2018, s. 421-430. DOI: 10.1007/978-3-319-75175-7_42. ISBN 978-3-319-75174-0.
- [98] SESHADRINATHAN, Kalpana, Rajiv SOUNDARARAJAN, Alan Conrad BOVIK a Lawrence K. CORMACK. Study of Subjective and Objective Quality Assessment of Video. *IEEE Transactions on Image Processing*. 2010, 19(6), 1427-1441. DOI: 10.1109/TIP.2010.2042111. ISSN 1057-7149.
- [99] *H.264 Video Encoding Guide* [online]. FFmpeg [cit. 2020-05-15]. Dostupné z: <https://trac.ffmpeg.org/wiki/Encode/H.264>.
- [100] KOCH, Michael. *An Introduction to FFmpeg, Timelapse and Full dome Video Production, Color Grading, Streaming, Audio Processing, Canon 5D-MK4, Panasonic LUMIX GH5S, Kodak PIXPRO SP360 4K Image Processing and Astronomy Software, Version from February 14, 2020* [online]. 2020 [cit. 2020-02-25]. Dostupné z: http://www.astro-electronic.de/FFmpeg_Book.pdf.
- [101] *Ayutthaya - Easy Tripod Paint - 360/VR Master Series* [online]. [cit. 2020-02-15]. Dostupné z: <https://www.mettle.com/360vr-master-series-free-360-downloads-page/>.
- [102] Great-circle distance [online]. [cit. 2020-02-15]. Dostupné z: https://en.wikipedia.org/wiki/Great-circle_distance.
- [103] DAVID, Erwan J., Jesús GUTIÉRREZ, Antoine COUTROT, Matthieu Perreira DA SILVA a Patrick Le CALLET. A dataset of head and eye movements for 360° videos. In: *Proceedings of the 9th ACM Multimedia Systems Conference*. New York, NY, USA: ACM, 2018, s. 432-437. DOI: 10.1145/3204949.3208139. ISBN 9781450351928.
- [104] QIAN, Feng, Lusheng JI, Bo HAN a Vijay GOPALAKRISHNAN. Optimizing 360 video delivery over cellular networks. In: *Proceedings of the 5th Workshop on All Things Cellular Operations, Applications and Challenges - ATC '16*. New York, New York, USA: ACM Press, 2016, s. 1-6. DOI: 10.1145/2980055.2980056. ISBN 9781450342490.
- [105] CAO, Haowei, Jiali LU a Nan ZONG. A Multi-User 360-Video Streaming System for Wireless Network. In: *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*. New York, NY, USA: ACM, 2019, s. 1-5. DOI: 10.1145/3359997.3365713. ISBN 9781450370028.