

Vysoká škola báňská - Technická univerzita Ostrava  
Fakulta elektrotechniky a informatiky  
Katedra telekomunikační techniky



## **Adaptívne rezervácie sieťových zdrojov podľa klasifikácie typu video scény**

Ing. Lukáš Ševčík

Doktorský program: Informatika, komunikační technologie a aplikovaná matematika  
Studijní obor: Komunikační technologie

Ostrava, 2018



## **Prehlásenie**

Prehlasujem, že som túto dizertačnú prácu vypracoval samostatne. Uviedol som všetky literárne pramene a publikácie, z ktorých som čerpal.

V Ostrave, 30.3.2018

---

## **Pod'akovanie**

Týmto by som sa chcel poďakovať môjmu školiteľovi prof. Ing. Miroslavovi Vozňákovi, Ph.D. za jeho cenné rady a pripomienky počas môjho doktorského štúdia.

Moje veľké poďakovanie patrí mojej mame. Dala mi to najvzácnejšie, život. Vždy ma podporovala, stála pri mne pri každej životnej skúške a vždy verila, že to zvládnem. Bez nej by táto práca nikdy nevznikla.

## **Abstrakt**

Vyhodnotenie video kvality vyžaduje kombinovaný prístup, ktorý zahŕňa objektívne metriky, subjektívne testovanie a monitorovanie siete. Spoľahlivo vykonané subjektívne testy na hodnotenie video kvality sú mimoriadne spoľahlivé, opakovateľné. Subjektívne testy využívajú sety krátkych videosekvencií (podľa normy 10 sekúnd) pre vyhodnotenie názoru ľudí na kvalitu videa ovplyvneného rôznym spracovaním a tým aj degradovaním. Tieto testy sú zamerané na aktuálne vnímanie ľudí, na rozdiel od hodnotenia celého videa. Cieľom je nezaujatý úsudok.

Táto práca sa zaoberá hlavne subjektívnym hodnotením kvality videa pri rôznom bitovom toku. Toto hodnotenie je nápomocné pri stanovení hraníc pre zmenu bitového toku na základe časovej a priestorovej informácie videa. Výsledky subjektívnych testov som porovnal s výsledkami objektívnych testov. Na základe tohto porovnania som definoval vzájomný vzťah, koreláciu. Vytvoril som mapovaciu funkciu, ktorá na základe priestorovej a časovej informácie videa určí vhodný bitový tok. Na vyhodnotenie kvality jednotlivých videí som použil normované videosekvencie. Tieto videosekvencie boli dôsledne vybrané na základe závislosti časovej a priestorovej informácie. V práci som vykonal porovnanie kódovacích algoritmov H.264/AVC (Advanced Video Coding) a H.265/HEVC (High Efficiency Video Coding) pri FHD (Full High Definition) a UHD (Ultra High Definition) rozlíšení.

## **Kľúčové slová**

Bitový tok, subjektívne hodnotenie kvality videa, objektívne hodnotenie kvality videa, priestorová informácia, časová informácia, H.264/AVC, H.265/HEVC, neurónová sieť

## **Abstract**

Video quality assessment requires a combined approach that includes objective metrics, subjective testing, and monitoring of the network. Carefully conducted video quality subjective tests are extremely reliable and repeatable. A subjective video quality tests use a small set of short video sequences (10 sec. according to standard) to measure people's opinions of the quality of different video processing options and with degradation of this video. These tests focus on people's current opinion, as opposed, e.g., to opinions of an entire movie. The goal is to make an impartial judgment about opinion trends.

The aim of this work is the subjective evaluation of the quality at the different bitrate. This evaluation is helpful to define the threshold for setting of the bitrate and his change based on the spatial and temporal information of the video. I compared the results of subjective tests with the results of the objective tests. Based on this comparison I defined correlation. After, I created optimum mapping function for to define the threshold for setting of the variable bitrate. I used standard video sequences for the evaluation of the video quality. These video sequences were carefully select based on the location of the individual quartile of the temporal and spatial information. I compared coding algorithms H.264/AVC (Advanced Video Coding) and H265/HEVC (High Efficiency Video Coding) and also different resolutions, FHD (Full High Definition) and UHD (Ultra High Definition).

## **Keywords**

Bitrate, subjective video quality assessment, objective video quality assessment, spatial information, temporal information, H.264/AVC, H.265/HEVC, neural network

## Obsah

1	Úvod.....	15
2	Súčasný stav skúmanej problematiky .....	17
3	Metodológia .....	21
3.1	Spracovanie videa .....	21
3.1.1	Použitý typ kodeku .....	22
3.1.2	Typ snímkov používaných v MPEG .....	24
3.1.3	GOP.....	25
3.1.4	Faktory vplyvajúce na kvalitu streamu.....	27
3.2	Hodnotenie kvality videa.....	28
3.2.1	Subjektívne metódy .....	28
3.2.2	Objektívne metódy.....	30
3.3	Časová a priestorová informácia scény.....	32
3.4	Neurónová sieť.....	34
4	Ciele dizertačnej práce.....	37
5	Realizácia cieľov dizertačnej práce .....	39
5.1	Vytvorenie databázy videosekvencií.....	39
5.2	Proces kódovania .....	42
5.2.1	Bitový tok .....	42
5.2.2	Rozlíšenie.....	42
5.2.3	Parametre kódovania .....	43
5.3	Klasifikácia scény.....	44
5.3.1	Priestorová spektrálna informácia (SI) .....	44
5.3.2	Časová spektrálna informácia (TI).....	45
5.3.3	Porovnanie kvality referenčných a komprimovaných videosekvencií .....	49
5.4	Štatistické metódy vyhodnotenia.....	49
5.5	Analýza vplyvu bitového toku na kvalitu videosekvencií .....	52
6	Vytvorenie modelu klasifikátora scény pomocou neurónovej siete .....	55
6.1	Trénovanie siete .....	56
6.2	Vytvorený klasifikátor .....	58
7	Vyjadrenie korelácie výsledkov v hodnotení kvality videa medzi objektívnymi a subjektívnymi metódami .....	59
7.1	Predikcia subjektívneho hodnotenia kvality videosekvencie na základe scény a objektívneho hodnotenia videosekvencie .....	64
7.1.1	Výber vhodnej topológie.....	65

7.1.2	Vplyv aktivačnej funkcie .....	67
8	Vytvorenie optimálnej mapovacej funkcie predikcií objektívnych metrík a výsledkov subjektívnych hodnotení na určenie hraníc TI a SI .....	69
8.1	Predikcia bitového toku na základe požadovanej kvality vyjadrenej metrikou SSIM .....	69
8.2	Predikcia bitového toku na základe požadovanej kvality vyjadrenej pomocou MOS skóre .	71
9	Verifikácia a validácia modelu.....	74
9.1	Verifikácia a validácia modelu predikcie subjektívnej hodnoty na základe scény, bitového toku a objektívneho výsledku metrikou SSIM .....	74
9.2	Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou objektívnym hodnotením metrikou SSIM .....	76
9.3	Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou pomocou MOS stupnice .....	80
10	Zhrnutie výsledkov dizertácie pre prax a ďalší rozvoj vedy .....	83
11	Záver .....	87
	Referencie.....	89
	Citované príspevky autora .....	93
	Publikačná a projektová činnosť autora .....	95
	Príloha A Databáza videosekvencií.....	97
	Príloha B.1 Popis kódovania a dekódovania videosekvencií .....	107
	Príloha D.1 Verifikácia a validácia modelu predikcie subjektívnej hodnoty na základe scény, bitového toku a objektívneho výsledku metrikou SSIM pre kompresné štandardy separátne .....	113
	Príloha D.2 Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou pomocou MOS stupnice pre kompresné štandardy separátne.....	117
	Príloha E Obsah priloženého média .....	121



## Zoznam použitých skratiek

ACR	Absolute Category Rating	Absolútne kategorické hodnotenie
AES	Advanced Encryption Standard	Pokročilý šifrovací štandard
AH	Authentication Header	Autentifikačná hlavička
AVC	Advanced Video Coding	Pokročilé kódovanie videa
B	Byte	Bajt
BR	Bitrate	Bitový tok
BPNN	Back Propagation Neural Network	Neurónová sieť založená na metóde spätného šírenia chýb
CBC	Cipher Block Chaining	Reťazenie šifrovacieho bloku
CBR	Constant Bit Rate	Konštantný bitový tok
CPU	Central Processing Unit,	Procesor základnej jednotky
DCT	Discrete Cosine Transform	Diskrétna kosínusová transformácia
DCR	Degradation Category Rating	Hodnotenie degradácie
DSCS	Double Stimulus Comparison Scale	
DSCQS	Double Stimulus Continuous Quality Scale	
DSIS	Double Stimulus Impairment Scale	Dvojnásobne stimulovaná mierka zhoršenia
DVD	Digital Versatile Disc	Digitálny viacúčelový disk
EPA	Extended Pedestrian A Model	
ESP	Encapsulating Security Payload	
ETU	Extended Typical Urban model	
EVA	Extended Vehicular A model	
FPS	Frames per Second	Počet snímkov za sekundu
GOP	Group of Pictures	Súbor snímkov v MPEG
GPU	Graphics Processing Unit	Grafický procesor
HD	High Definition	Vysoké rozlíšenie
HDTV	High-definition television	Televízia s vysokým rozlíšením
HEVC	High Efficiency Video Coding	Vysoko efektívne kódovanie videa
HMAC	Hash-based Message Authentication Code	Zašifrovaný autentifikačný kód
HRC	Hypothetical Reference Circuits	Hypotetické referenčné obvody
HST	High Speed Train condition	
HVS	Human Visual System	Ľudský vizuálny systém
IP	Internet Protocol	Internetový protokol
IPsec	IP security	Bezpečnosť internetového protokolu
ITU	International Telecommunication Union	Medzinárodná telekomunikačná únia
LTE	Long Term Evolution	
MB	Megabyte	Megabajt

Mbps	Megabits per Second	Megabit za sekundu
MOS	Mean Opinion Score	Hodnotiaca stupnica MOS
MPEG	Motion Picture Experts Group	Skupina expertov pre pohyblivý obraz
MSE	Mean Square Error	Štandardná odchýlka
NN	Neural network	Neurónová sieť
OS	Operation System	Operačný systém
PAL	Phase Alternating Line	Kódovací štandard farebnej televízie
PDV	Packet Delay Variation	
PSNR	Peak Signal to Noise Ratio	Špičkový pomer signálu k šumu
PTS	Presentation TimeStamp	Prezentačná časová pečiatka
RGB	Red-Green-Blue	
RNN	Random Neural Network	Typ rekurentnej neurónovej siete
RMSE	Root Mean Square Error	
SNR	Signal to Noise Ratio	Pomer signálu k šumu
Triple DES	Triple Data Encryption Algorithm	Algoritmus šifrovania údajov
QoE	Quality of Experience	Kvalita užívateľskej prívetivosti
QoP	Quality of Experience	Kvalita vnímania
QoS	Quality of Service	Kvalita služby
SHA1	Secure Hash Algorithm 1	Bezpečnostný šifrovací algoritmus
SCSF	Spatial Contrast Sensitivity Function	
SI	Spatial Information	Priestorová informácia
SSCQE	Single Stimulus Continuous Quality Evaluation	
SQL	Structured Query Language	Štruktúrovaný vyhľadávací jazyk
SSIM	Structural Similarity Index	Index štruktúrálnej podobnosti
TCP	Transmission Control Protocol	Protokol riadenia prenosu
TI	Temporal Information	Časová informácia
UDP	User Datagram Protocol	Používateľský datagramový protokol
UHD	Ultra High Definition	Ultra vysoké rozlíšenie
VBR	Variable Bit Rate	Variabilný bitový tok
VLC	Video Lan Client	Video LAN klient
VoIP	Voice over IP	Hlas prenášaný cez IP protokol
VPN	Virtual Private Network	Virtuálna privátna sieť
VQA	Video Quality Assessment	Hodnotenie kvality videa
VQM	Video Quality Metric	Metrika video kvality

## Zoznam tabuliek

Tabuľka 1 Stupnica MOS.....	29
Tabuľka 2 Počet hodnotiteľov.....	50
Tabuľka 3 Najlepšie topológie predikcie MOS na základe SSIM, UHD.....	66
Tabuľka 4 Simulácia dát pri vybraných topológiách predikcie MOS na základe SSIM.....	66
Tabuľka 5 Vplyv aktivačnej funkcie na predikciu MOS na základe SSIM .....	68
Tabuľka 6 Najlepšie topológie predikcie BR na základe SSIM.....	70
Tabuľka 7 Simulácia dát pri vybraných topológiách predikcie BR na základe SSIM.....	70
Tabuľka 8 Vplyv akt. funkcie na predikciu BR na základe SSIM.....	70
Tabuľka 9 Najlepšie topológie predikcie BR na základe MOS .....	72
Tabuľka 10 Simulácia dát pri vybraných topológiách predikcie BR na základe MOS.....	72
Tabuľka 11 Vplyv akt. Funkcie na predikciu BR na základe SSIM .....	72
Tabuľka 12 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu MOS pomocou SSIM, H.264+H.265, UHD.....	74
Tabuľka 13 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.264+H.265, UHD .....	76
Tabuľka 14 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.264+H.265, FHD+UHD .....	78
Tabuľka 15 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264+H.265, UHD .....	80
Tabuľka 16 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264+H.265, FHD+UHD.....	81
Tabuľka 17 Databáza videosekvencií SJTU Media Lab.....	97
Tabuľka 18 Databáza videosekvencií Ultra Video Group .....	103
Tabuľka 19 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu MOS pomocou SSIM, H.264, UHD.....	113
Tabuľka 20 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.265, UHD .....	114
Tabuľka 21 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264, UHD .....	117
Tabuľka 22 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.265, UHD .....	118

## Zoznam obrázkov

Obrázok 1 Bloková schéma zdrojového kodéra a dekodéra obrazu.....	21
Obrázok 2 Ukážka kompresie.....	22
Obrázok 3 Dopredná predikcia.....	24
Obrázok 4 Skupina snímok GOP.....	25
Obrázok 5 Hierarchia MPEG video-toku .....	26
Obrázok 6 Hierarchia rámcov MPEG video-toku .....	27
Obrázok 7 Podvzorkovanie .....	28
Obrázok 8 Bloková schéma metódy SSIM.....	30
Obrázok 9 Blokový diagram VQM metriky .....	32
Obrázok 10 Ilustrácia reprezentujúca obraz v priestorovej frekvencii .....	33
Obrázok 11 Rozšírená sieť pre výpočet chybovej funkcie .....	36
Obrázok 12 Bund Nightscape      Obrázok 13 Campfire Party.....	40
Obrázok 14 Construction Field      Obrázok 15 Fountains .....	40
Obrázok 16 Marathon      Obrázok 17 Runners.....	40
Obrázok 18 Tall Buildings      Obrázok 19 Wood.....	41
Obrázok 20 Beauty      Obrázok 21 Bosphorus .....	41
Obrázok 22 Jockey      Obrázok 23 ReadySteadyGo.....	42
Obrázok 24 Priestorová informácia SI .....	44
Obrázok 25 Časová informácia TI.....	45
Obrázok 26 Proces kódovania a vyhodnotenia video kvality.....	49
Obrázok 27 Log-sigmoidná aktivačná funkcia.....	67
Obrázok 28 Tan-sigmoid aktivačná funkcia.....	67
Obrázok 29 Popis názvoslovia videosekvencií.....	97

## Zoznam grafov

Graf 1 Závislosť SI a TI pri vybraných videosekvenciách.....	48
Graf 2 SSIM vyhodnotenie Ultra Video Group sekvencií – rozlíšenie FHD.....	53
Graf 3 SSIM vyhodnotenie Ultra Video Group sekvencií – rozlíšenie UHD.....	53
Graf 4 Dosiachnuté výsledky objektívnou a subjektívnou metrikou.....	63
Graf 5 Priemerné výsledky vyhodnotenia kvality objektívnou a subjektívnou metrikou.....	64
Graf 6 Korelačný diagram, predikcia MOS pomocou SSIM, H.264+H.265, UHD.....	75
Graf 7 Krivka hustoty rozdelenia, predikcia MOS pomocou SSIM, H.264+H.265, UHD.....	75
Graf 8 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, UHD.....	77
Graf 9 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.265, UHD.....	77
Graf 10 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, FHD+UHD.....	79
Graf 11 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.265, FHD+UHD.....	79
Graf 12 Korelačný diagram, predikcia BR pomocou MOS, H.264+H.265, UHD.....	80
Graf 13 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264+H.265, UHD.....	81
Graf 14 Korelačný diagram, predikcia BR pomocou MOS, H.264+H.265, FHD+UHD.....	82
Graf 15 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264+H.26, FHD+UHD.....	82
Graf 16 Korelačný diagram, predikcia BR pomocou SSIM, H.264, UHD.....	113
Graf 17 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264, UHD.....	114
Graf 18 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, UHD.....	115
Graf 19 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.26, UHD.....	115
Graf 20 Korelačný diagram, predikcia BR pomocou MOS, H.264, UHD.....	117
Graf 21 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264, UHD.....	118
Graf 22 Korelačný diagram, predikcia BR pomocou MOS, H.265, UHD.....	119
Graf 23 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.265, UHD.....	119



# 1 Úvod

Stále rastúci záujem o služby v reálnom čase, ako je audio a video prenos cez paketové siete založené na IP (Internet Protocol), vedie k analýze týchto služieb a ich správaniu sa v takých sieťach stále intenzívnejšie. Keď porovnávame obe služby, tak väčšie nároky na podporu nástrojov QoS (Quality of Service) v sieťových prvkoch, podieľajúcich sa na prenose, kladie IP telefónia. Nájde mnoho aplikácií, ktoré majú extrémne nároky, mnohokrát vyžadujúce rezervácie sieťových zdrojov na prenosovej ceste. V lekárstve medzi ne patria napr. vzdialené vykonávané robotické operácie a v umení sú príkladom koncerty členov orchestra nachádzajúcich sa na rôznych miestach. Architektúra IP sietí nebola navrhnutá pre služby v reálnom čase a prenos údajov citlivých na oneskorenie. Tým je tam mnoho faktorov, ktoré môžu ovplyvniť výslednú kvalitu služby, obzvlášť strata paketov, oneskorenie, jitter.

V dnešnej dobe má video kvalita, ako časť multimediálnych technológií, kľúčovú úlohu. Video je majoritná časť všetkej dátovej prevádzky posielanej cez IP siete. V celkovom objeme prenosu videa jednoznačne prevláda jednosmerná služba (v porovnaní s audiom – hlasový hovor), čiže oneskorenie v sieti nie je tak dôležitým faktorom ako v prípade hlasu. Dominantnými faktormi siete, ktoré majú vplyv na kvalitu videa sú strata paketov, PDV (Packet Delay Variation) a kapacita prenosovej linky [1].

Analýza kvality videa sa zameriava na stratovosť paketov v sieti v závislosti od použitého kodeku, čo spôsobuje artefakty vo videu [2, 3]. Streamovanie videa, ako časť služieb v reálnom čase, je zabezpečovaná predovšetkým kodekmi MPEG-2 (Motion Picture Experts Group) a MPEG-4 (H.264). Video kodek H.265/HEVC (High Efficiency Video Coding) ponúka vyššiu účinnosť kompresie – lepšia kvalita s použitím rovnakého bitového toku [sev01]. Avšak potrebuje viac výpočtového výkonu a to je hlavný dôvod, prečo sa používa pre vysoké rozlíšenia ako FHD (1 920×1 080) a 4K (4 096×2 160), 8K (7 680×4 320), ... [4].

Užívatelia očakávajú zaručenú kvalitu odoberaných služieb nehľadiac na typ prenosu a použitej technológie. Na jednej strane je poskytovateľ služieb, ktorý chce dodávať služby pri minimálnych investičných nárokoch na poskytovanie, pričom implementácia nástrojov QoS, či monitorovanie QoE (Quality of Experience) sú pre neho nákladom. Na druhej strane je zákazník, ktorý pochopiteľne nie je spokojný s doručovaním obsahu v kvalite, ktorá mu neumožňuje nerušené vnímanie služby. K stanoveniu hraníc užívateľskej spokojnosti používame metriky QoE a hlavne aplikácie takých QoS nástrojov v sieti, ktoré zaručia minimálne QoE očakávané užívateľom. Využívanie internetu ako prostredia pre poskytovanie multimédií je dnes celkom bežné, nie je však úplne samozrejmé, že užívateľ dostane, v takomto prostredí, službu s požadovanou kvalitou. O to väčšie opodstatnenie dnes prináša monitorovanie QoE a hľadanie väzieb medzi QoS a QoE.

Je nevyhnutné vyhodnotiť výkon systémov pre posielanie informácií z jedného zdroja k druhému (dátová linka); či je prenos informácií efektívny a účinný. Pri systémoch určených pre doručenie ku koncovým užívateľom (IP televízna služba) hodnotíme prenosovú kvalitu z pohľadu spokojnosti užívateľa s kvalitou mediálneho obsahu pri jeho sledovaní.

Všeobecne sa predpokladá, že vysoký výkon a kvalita prenosu vedie k veľkej spokojnosti užívateľov k daným službám. Do nedávnej doby bolo slovo kvalita používané technikmi na popis „súhrnu vlastností“, čo je založené na schopnosti zvoliť stanovené alebo implicitné potreby (EN ISO 9000, 2000). Tento súbor vlastností je spätý s latinským originálom „qualitas“; anglického slova „quality“, a je to niečo, čo v dnešnej dobe môžeme nazývať ako charakteristika subjektu. Toto chápanie termínu

kvalita sa odráža aj v jeho použití v QoS, čo v sieťovej terminológii bolo dlho používané ako synonymum k množine garantovaných charakteristík sieťového spojenia. Definícia kvality bola neskôr zlepšená a definovaná ako stupeň nastavení základných charakteristík spĺňajúcich požiadavky. Z vnímania ľudského uhla pohľadu je kvalita definovaná ako výsledok posúdenia vnímanej kompozície. Táto definícia zahŕňa proces vnímania a posúdenia, počas ktorého vnímajúca osoba porovnáva vnímané udalosti s (doposiaľ neznámou) referenciou. Charakter vnímanej kompozície nie je nevyhnutne stála charakteristika objektu, v skutočnosti referencia môže mať vplyv na to, čo je aktuálne vnímané. V každom prípade ako výsledok porovnania je kvalita zvyčajne relatívna a nastane ako udalosť v určitom priestorovom, časovom a funkčnom kontexte. Pojem QoE vznikol ako protiváha k pojmu QoS, s niečím, čo adresuje ľudské vnímanie a skúsenosti pretože tieto boli považované ako viac vhodné pre navrhovanie vysoko akceptovateľných systémov a služieb. Tak by bolo možné pozorovať posun paradigmy pre poskytovateľov služieb na doručenie služieb, s nie vysokým QoS, ale s vysokým QoE, pre ich užívateľov. QoE reflektuje fakt, že vnímaná kvalita je kľúčovým faktorom pre vyhodnotenie systémov, služieb alebo aplikácií vo fáze projektovania alebo počas prevádzky. Často vykonáva merania zamerané na redukčnú perspektívu posúdenia známych služieb a identifikovanie kvality ako relevantného kritéria [5]. Preto je základ tejto práce postavený na QoE.



## 2 Súčasný stav skúmanej problematiky

V tejto kapitole je popísaná časť práci, ktoré sa venujú skúmanej problematike. Autori práce [6] poskytujú prehľad o hodnotiacich metódach audio-vizuálneho vnímania kvality. Zamerali sa na úplne referenčné modely. Tento článok poskytuje komplexný prehľad o vývoji multimediálnych metód vnímania kvality, popis formovania a vnímania kvality, vrátane rôznych kvalitatívnych faktorov, popis tvorby a vnímania kvality vrátane rôznych kvalitatívnych faktorov, všeobecný prehľad na QoS, QoE a QoP (Quality of Perception) v kontexte audiovizuálneho hodnotenie kvality. Spoľahlivý odhad kvality videa sa stáva čoraz dôležitejším, vďaka rozširovaniu online video služieb. Bolo navrhnuté množstvo metód hodnotenia objektívnej kvality videa s rôznym výkonom a zložitou. Vzhľadom na množstvo faktorov, ktoré ovplyvňujú pôžitok užívateľa z video služieb, sa zameranie výskumu postupne presúva z pojmu QoS na širšie aspekty ako je kvalita vnímania (QoP) [7] alebo kvality skúseností (QoE) [8]. ITU (International Telecommunication Union) definovala QoE ako celkovú prijateľnosť aplikácie alebo služby, ktorú subjektívne vníma konečný užívateľ [9]. Novšia definícia, ktorú navrhol Qualinet [8], definuje QoE ako stupeň potešenia alebo nepríjemnosti používateľa z aplikácie alebo služby, ktorá vyplýva z plnenia jeho očakávaní vo vzťahu k užitočnosti a/alebo vychutnaniu si aplikácie alebo služby.

V práci [10] bol popísaný účinok miery kódovania (Video Coding Rate) a straty paketov v prenosovom streame. Navrhuje sa vhodný model, ktorého koeficienty závisia od parametrov kódovania videa a zložitosti sekvencií. V [11] je popísaný účinok niekoľkých kódovaní a sieťových parametrov. Koeficienty modelu závisia hlavne od zložitosti videosekvencie z pohľadu informácií obsahu. V [12] preskúmaná závislosť VQM (Video Quality Metric) na type kodeku, rozlíšení a miere pohybu vo videosekvenciách. Sekvencie sú klasifikované v skupinách rozdelených podľa miery pohybu a hodnoty sú priradené k parametrom pre každú triedu. Nehovorí sa však o použitých kritériách klasifikácie videosekvencie.

Metódy založené na bitovom toku používajú práve bitový tok na vyhodnotenie kvality. Podľa úrovne informácií použitých na spracovanie môžu byť rozdelené na model parametrického plánovania, vrstvový model bitového toku. Parametrické plánovacie techniky používajú typ kodeku, stratu paketov a bitový tok pre hodnotenie kvality. Príkladom takéhoto modelu je vzorový model pre video-telefónne aplikácie popísaný v [13].

Vrstvové modely bitového toku dokážu robiť ľubovoľný typ analýzy bitového prúdu s výnimkou použitia pixelových údajov, a preto sú pomerne zložitejšie, ale ponúkajú lepší výkon. V práci [14] bolo popísané, že ich rámec môže byť použitý na vyhodnotenie kvality v reálnom čase, kde ako parameter QoS môže byť použitá stratosť paketov, SI (Spatial Information) a TI (Temporal Information) informácia.

V práci [15] bol popísaný model vyhodnotenia vnímanej kvality pri zohľadnení kódovania videa a kvalitatívnych parametrov. Obsahuje parametre, ktoré závisia hlavne od obsahu informácií vo videosekvenciách. Je navrhnutá metóda na ich získanie z priestorovej a časovej informácie videosekvencie. Ako objektívna metrika bola použitá VQM. Zmyslom je výpočet parametrov závisiacich na obsahu scény pre každú videosekvenciu na základe parametrov TI a SI. Najskôr sú vypočítané parametre SI a TI. Potom je každá sekvencia zakódovaná, prenášaná simulovanou sieťou a následne dekódovaná pre získanie skreslenej sekvencie. Pri hodnotení bolo použitých 21

videosekvencií v SD formáte, pričom len jedna bola prekódovaná v rozličnom formáte. Bol použitý MPEG-2.

Práca [16] navrhuje nereferenčnú metriku pre meranie artefaktov založených na analýze dekomprimovaného obrazu. Model používa SI a TI charakteristiky obsahu videa pre zlepšenie výkonnosti z hľadiska korelácie s MOS (Mean Opinion Score).

Hoci subjektívne hodnotenie kvality poskytuje spoľahlivé informácie ľudského vnímania kvality, nemožno ho použiť v reálnom čase hodnotenia kvality služby. Preto je potrebné využiť objektívne vyhodnotenie kvality, ktoré bolo navrhnuté, aby nahradilo hodnotenie užívateľov pomocou výpočtového modelu na predikciu výsledkov subjektívneho testu, teda automaticky odhadnúť hodnoty MOS, ktoré sa čo najviac približujú subjektívnemu hodnoteniu kvality [17, 18, 19].

Očakáva sa, že číselné vyjadrenie kvality získané pomocou objektívnej metódy bude lepšie korelovať s ľudskou subjektivitou. Existujú rôzne metriky na vyjadrenie korelácie medzi subjektívnym a objektívnym hodnotením. Dve najčastejšie štatistické metriky používané na vyjadrenie dosiahnutého výkonu sú RMSE (Root Mean Square Error) a Pearsonov korelačný koeficient. Vysoká hodnota korelácie (zvyčajne vyššia ako 0,8) sa považuje ako efektívna [20].

Práca [21] navrhuje nový mechanizmus, ktorý využíva údaje z verejných VQA databáz a umožňuje automaticky vytvárať všeobecné pravidlá pre mapovanie hodnôt objektívnych metód na subjektívnu hodnotu MOS. Posúdenie kvality aproximácie bolo s použitím korelácie a presnosť odhadu kvality bol okolo 95%. V modeli boli použité údaje z troch verejných databáz. Subjektívne dáta boli konvertované na rovnakú vstupnú škálu (podľa vybranej stupnice MOS, napr. 1-5). Bola použitá nelineárna regresia a inverzná interpolácia. Daný model hodnotí videosekvencie v nižšom rozlíšení s použitím kompresného štandardu H.264.

Počas môjho štúdia a písaní tejto práce som publikoval viaceré články danej problematiky. Z výberu uvádzam prácu [sev02], kde som skúmal vplyv šifrovania na kvalitu videa a poukázal na možné dopady za rôznych podmienok, v ktorých som vykonal experimenty. Cieľom práce [sev03] je preskúmať vplyv šifrovania na video prenos v IP sieti. Opisujem tu tunel IPsec pomocou ESP a AH hlavičky, ktoré poskytujú diskretnosť z hľadiska bezpečnosti, integrity a nepopierateľnosti (pomocou šifrovania HMAC-SHA1, Triple DES a AES v režime CBC). Ďalším cieľom bolo posúdiť, ako OpenVPN ovplyvňuje prenášané video. Porovnal som výsledky oboch meraní a vyjadril vplyv straty paketov na prenášané video.

V publikácii [sev04] som testoval vplyv použitia rôznej konfigurácie QoS v LTE (Long Term Evolution) sieťach na kvalitu triple play služieb. Článok [sev05] poukazuje na zmenu kvalitatívnych parametrov pri rôznych nastaveniach siete. Simuloval som pri tom LTE sieť. Na tento účel bol vybratý Matlab, v ktorom som simuloval priepustnosť, rôzne scenáre, meniac frekvenčný a časový duplex SNR a tiež boli nasimulované rôzne kanály, ako EPA (Extended Pedestrian A Model), EVA (Extended Vehicular A model), ETU (Extended Typical Urban model) a HST (High Speed Train condition). Simuláciou rôznych scenárov bol získaný požadovaný model priepustnosti a bolo určené ako ovplyvňuje kvalitu triple play služieb v LTE. V práci [sev06] som uviedol analýzu ako degraduje kvalitu prenášaného videa strata paketov, pakety mimo poradia a rôzne video rozlíšenie.

V publikáciách [sev07, sev08, sev09] sa zaoberám simuláciou sieťovej v IP sieťach, popisujem model predikcie očakávanej kvality QoE multimediálnych služieb pre rôzne politiky obsluhy prichádzajúcich paketov v smerovačoch. Na základe navrhnutého analytického modelu bola vytvorená aplikácia, pomocou ktorej je možné predpovedať kvalitatívne parametre QoS v sieťach IP, s ohľadom na politiky používané na spracovanie paketov v smerovačoch a vzhľadom na úroveň celkového využitia

siete. Príspevok tohto článku spočíva v návrhu modelu, ktorý dokáže predpovedať kvalitu služieb triple play v súlade s politikou čakania a prenosovou rýchlosťou rozhraní na smerovačoch v paketových sieťach. Publikácia [sev10] rozširuje predchádzajúce 3 publikácie Zameriava sa na sieťové javy, skúma vplyv rôznych oneskorení a strát paketov na parametre kvality služieb triple play, aby vyhodnotili výsledky pomocou objektívnych metód. Cieľom tejto práce je priniesť detailný pohľad na výkon streamingu videa cez siete založené na IP.



### 3 Metodológia

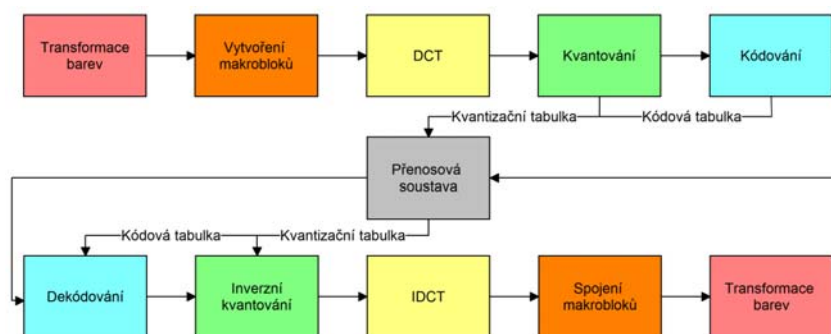
Práca je rozdelená do niekoľkých častí. Na úvod a súčasný stav riešenej problematiky nadväzuje táto kapitola, ktorá popisuje spracovanie videa, ohodnotenie kvality videa pomocou objektívnych a subjektívnych metód, charakteristiku videosekvencií. V závere kapitoly sú popísané neurónové siete, ktoré budú použité na implementáciu klasifikátora scény.

#### 3.1 Spracovanie videa

Video je séria statických obrázkov (snímok), doplnených zvukovým záznamom, ktoré sa rýchlo vymieňajú (štandardné video predstavuje 25 - 30 fps (Frames per Second)). Postup pri zápise videa do binárneho kódu je podobný ako pri zápise obrázkov. Dôsledkom toho je však hromadenie digitálnych informácií, preto pre reálne použitie digitálneho videa, je nutné toto množstvo informácií komprimovať použitím vhodnej metódy kompresie. Množstvo dát rastie pri vyššom rozlíšení. Veľké množstvo dát je hlavným problémom v prenose, spracovaní, ukladaní a zobrazovaní videa. Pomocou vhodných kompresných algoritmov môže byť odstránené veľké množstvo redundantných dát, ktoré video obsahuje [22].

#### Kompresia Videa

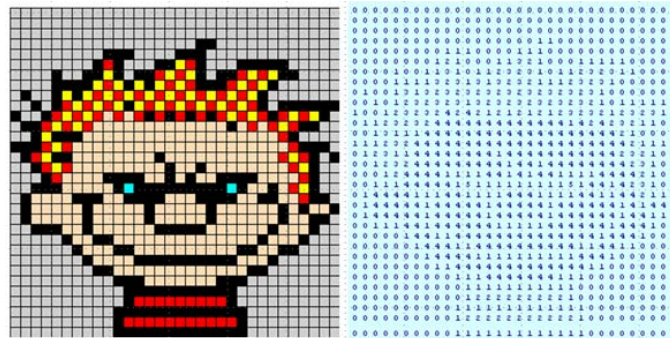
Kompresia dát je potrebná, pretože neustále rastie množstvo ukladaných a prenášaných dát. Znamená to uloženie informácie v kratšej podobe a tým zredukovanie fyzickej veľkosti bloku informácií. Video je sekvencia po sebe idúcich obrázkov. Nekomprimované video nie je vhodné na uchovávanie na disku alebo jeho šírenie, kvôli jeho veľkosti. Napríklad video pri rozlíšení 720x576 vo formáte RGB (Red-Green-Blue; 3B na obrazový bod), pri 25 fps by potrebovalo takmer 30MB na pevnom disku pre jednu sekundu ( $720 \times 576 \times 3 \times 25$  B). V tomto prípade by jedna hodina videa zaberala okolo 100 GB na pevnom disku. Spracovávať také množstvo dát nie je ekonomické a úsporné. Kompresia znižuje objem dát a videa. Je vykonávaná pomocou kompresných štandardov (kodekov). Kodek pochádza z anglických slov enCOder a DECOder (kóder a dekováder). Je to teda algoritmus, ktorý vykonáva kompresiu aj dekompresiu – zakóduje snímky do menšej podoby a následne ich dekováder v reálnom čase. Kompresný pomer je určený pomerom originálneho a komprimovaného objemu dát. Čím je väčší, tým je kvalita výsledného videa menšia (menej detailov). Oblasť kódovania som sa venoval v článku [sev11].



Obrázok 1 Bloková schéma zdrojového kodéra a dekodéra obrazu

Komprimácia môže byť stratová a bezstratová. V prvom prípade sa niektoré menej dôležité dáta vynechávajú, zachováva sa podstata informácie. Využíva nedokonalosti ľudských zmyslov, ako zrak a sluch. Obrázky nemusia obsahovať úplne všetky detaily. Zvuk je náchyľnejší na stratu informácií, ale tiež ho ľudské ucho vníma len v určitom spektre. V tomto prípade môžeme dosiahnuť veľký kompresný pomer.

Pri bezstratovej kompresii sa zachovávajú všetky pôvodné dáta. V tomto prípade sa vynechávajú redundantné informácie. Tu sa dá dosiahnuť len nízky kompresný pomer – okolo 2:1.



Obrázok 2 Ukážka kompresie

### 3.1.1 Použitý typ kodeku

Kodeky delíme na bezstratové (e.g. surové video vo formáte .RAW, HuffYUV) a stratové. Medzi stratové kodeky sa radia kodeky skupiny MPEG ktorá vyvíja štandardy, ktoré využívajú kompresný formát pre kódovanie audia a videa. Je to jeden z najpopulárnejších otvorených štandardov pre video na internete. MPEG stream je známy ako GOP (Group of Picture). Voľba GOP ovplyvňuje statické vlastnosti MPEG, ako je veľkosť rámca, veľkosť súboru a tiež dopad na bitový tok a vplyv na prehrávanie [23].

MPEG tok môže tolerovať nejaké straty, ale je náchylný na oneskorenie a jitter. Kým jeden bit môže spôsobiť poškodenie spustiteľného súboru, strata rámca vo videu nemusí byť pozorovateľná pre používateľa. Štandardný prenosový protokol TCP (Transmission Control Protocol) opäť posiela nedoručené pakety, ale vnáša oneskorenie. Z toho dôvodu sa väčšinou používa protokol UDP (User Datagram Protocol) a to najmä pre interaktívne aplikácie [13]. Popisu kodekov skupiny MPEG (MPEG-2, H.264, H.265) som sa venoval v článkoch [sev01, sev11, sev12].

Spoločnosť Google vyvinula kodeky VP8 a VP9. Práve VP9 je konkurentom H.265, je síce výkonnejší, ale nepodporuje ho každé zariadenie. Tieto kodeky sa zatiaľ nepresadili pri streamovaní videa. Bližší popis týchto štandardov je možné nájsť v publikácií [sev11].

## H.264/AVC

Je to jeden z najpoužívanejších (audio a video) kompresných metód štandardizovaných skupinou MPEG, navrhnutý pre širokú škálu aplikácií počnúc od video telefónov cez webové aplikácie až k TV vysielaniu (HDTV). Bol štandardizovaný v r.2003 Je najpopulárnejší pre vysoko dynamické videá. Môže využívať stratovú aj bezstratovú kompresiu na základe nastavení vybraných pri kódovaní, ako je počet snímkov za sekundu, kvalita a veľkosť cieľového súboru [24, 25].

Niektoré z vylepšených funkcií od predchádzajúcich kodekov:

- DCT algoritmus pracuje s pixelmi 4x4 namiesto 8x8, ale stále podporuje aj 8x8
- DCT je vrstvený pomocou Hadamardovej transformácie
- Podporuje vzorkovanie 4:2:2 a 4:4:4
- Pohybové bloky rozdielnej veľkosti
- Aritmetické kódovanie s premennou dĺžkou
- Vážená obojsmerná predikcia
- Redundantné obrázky
- Flexibilné usporiadanie makroblokov
- Priamy režim pre B-snímky
- Viac referenčných rámcov

## **H.265 (HEVC)**

Má rovnakú základnú štruktúru ako predchádzajúce štandardy, ale prináša mnoho vylepšení:

- 2-násobná efektívnosť pri kompresii so zachovaním rovnakej kvality videa – v prípade rovnakého bitového toku, bude video ešte kvalitnejšie
- Je ideálny pre UHD a vyššie formáty (8K)
- Väčšia flexibilita v režimoch predikcie a transformovaní veľkosti blokov
- Väčšia flexibilita pri delení veľkých partícií na malé
- Sofistikovanejšie predikcie a signalizácia režimov a pohybových vektorov

Je navrhnutý pre dosiahnutie viacerých cieľov ako efektívnosť kódovania, jednoduchosť pri prenose dát, odolnosť straty dát [24].

### 3.1.2 Typ snímkov používaných v MPEG

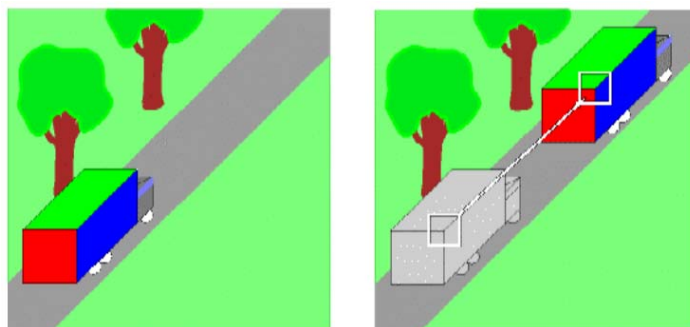
Keď je video zakódované pre prenos určený na sledovanie v TV, na internete alebo pomocou iného streamovaného zariadenia, je potrebné, aby bolo skomprimované. Ak sledujeme video s rýchlosťou 24 snímkov za sekundu, v skutočnosti vidíme sadu GOP snímkov, ktoré obsahujú referenčné (I - Intraframe), predikčné smerom dopredu (P - Predicted) a predikčné obojsmerné (Bidirectional predicted) snímky. Jednotlivé snímky sú založené na časovej a priestorovej podobnosti [23].

- Snímky I

Zahŕňajú všetky dáta pre dekódovanie a nie sú závislé na žiadnych iných rámcoch. Sú to kompletne obrázky, aj keď sú tiež komprimované, ale obsahujú každý pixel, ktorý sa v snímke nachádza (kompresný pomer je veľmi malý). Každý GOP začína touto snímkou, potom nasledujú P a B snímky – neúplné obrázky, ktoré sa odkazujú na I snímky a okolité snímky pre vyplnenie prázdneho miesta. Sú kódované priestorovo, nie je potrebné použiť okolité snímky.

- Snímky P

Pri kódovaní využíva jednosmernú (doprednú) predikciu. Aktuálna snímka P je kódovaná pomocou podobnosti s predošlou snímkou typu I alebo B. Touto kompresiou môžeme zmenšiť výsledný dátový tok zhruba o polovicu.



Obrázok 3 Dopredná predikcia

Keď je makroblok komprimovaný pohybom, komprimovaný súbor obsahuje tieto informácie:

1. Priestorový vektor medzi referenčným makroblokom a makroblokom, ktorý je kódovaný (pohybový vektor).
2. Rozdiel obsahu medzi referenčným makroblokom a makroblokom, ktorý je kódovaný (chybovosť).

- Snímky B

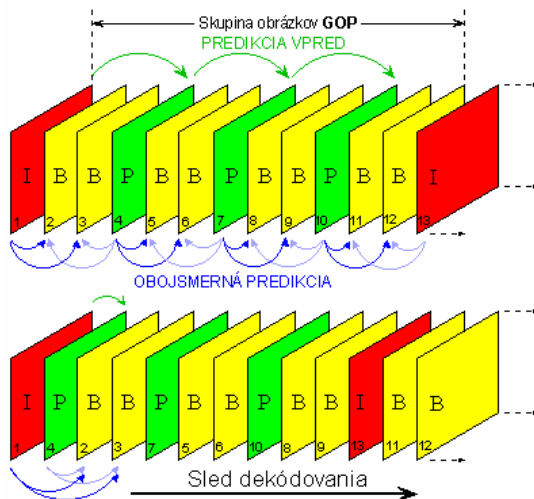
V tomto prípade je využitá obojsmerná predikcia, kde aktuálny snímok je kódovaný pomocou predošlej a predchádzajúcej snímky I alebo P. Tu je dosiahnutý najvyšší kompresný pomer, kde dokážeme znížiť dátový tok až 8-násobne. Bližšie som sa typom snímkov venoval v publikácií [sev13].



### 3.1.3 GOP

Obrázok 4 ukazuje jednoduchý GOP, kde druhý *I*-rámeček značí začiatok ďalšieho GOP-u. Pre obmedzenie kaskádového efektu chybovosti prenosu vzhľadom na závislosť rámcov, je nevyhnutné pravidelne vkladať *I*-snímky. Avšak zvýšená frekvencia *I*-snímok znamená vzdať sa väčšieho kompresného pomeru, ktorý im poskytujú *P* a *B* snímky.

Dekódovanie musí mať iné poradie ako kódovanie, čo vyplýva zo spôsobu kódovania snímok *P* a *B*. Postupnosť snímok, ktoré sa dekódujú je vyobrazená na obrázku 4 (4. snímka typu *P* musí byť dekódovaná pred 2. a 3. snímku typu *B*).



Obrázok 4 Skupina snímok GOP

Vplyv stratovosti paketov inak vplýva na jednotlivé typy snímok v dôsledku ich rôzneho významu. Strata paketov v snímke typu *I* má vplyv na celý GOP, čím je celý poškodený, zatiaľ čo strata paketu v snímke *B* poškodzuje iba snímky *B* [26].

Účinnosť GOP-u je možné vyhodnocovať vplyvom na statické MPEG videá alebo MPEG videá streamované cez stratovú sieť. Statická analýza zahŕňa kódovanie obrázkov z .raw do MPEG súborov s rozličným GOP pre porovnanie statických vlastností ako je veľkosť rámca, veľkosť súboru a kvalita. Na základe analýzy streamovania sa mení dĺžka a vzor GOP-u podľa vplyvu GOP-u na model bitového toku a snímkovej frekvencie. V súčasnosti je výber GOP-u väčšinou intuitívny proces. Niekedy sa používa predvolený GOP, ktorý je súčasťou MPEG kódera, niekedy sa mení vzor GOP-u a skúmajú sa dôsledky doručenia MPEG videa cez stratovú sieť.

Ak  $NP$  reprezentuje počet *P* rámcov v GOP a  $NBP$  reprezentuje počet *B* rámcov medzi *I* a *P* rámcami alebo dvoma *P* rámcami (rovnako ako vo väčšine MPEG videa sa predpokladá, že *B* rámce budú distribuované rovnomerne v intervale medzi *I* a *P* rámcami). S použitím týchto termínov môže byť jedinečne identifikovateľný GOP vzor pomocou vyjadrenia GOP ( $NP$ ,  $NBP$ ). Napríklad  $GOP(2, 2)$  vyjadruje GOP vzoru 'IBBPBBPBB'.

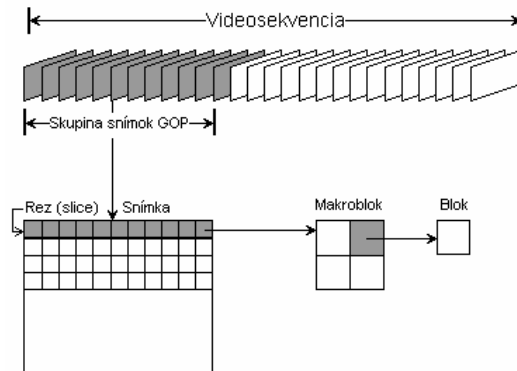
Ak  $NB$  reprezentuje počet *B* rámcov v GOP a  $NG$  reprezentuje dĺžku GOP, potom

$$NB = (1 + NP) * NBP \quad (1)$$

$$NG = 1 + NP + NB \quad (2)$$

## Videosekvencia

MPEG štandard definuje hierarchiu video-toku ako je znázornené na obrázku 5.



Obrázok 5 Hierarchia MPEG video-toku

### 1. Videosekvencia

Začína sa hlavičkou sekvencie (môže obsahovať dodatočné hlavičky sekvencie), zahŕňa jednu alebo viac skupín obrázkov a končí kódom konca sekvencie

### 2. GOP

Hlavička a séria jedného alebo viacerých snímok, ktoré majú umožniť náhodný prístup do sekvencie.

### 3. Snímok

Primárne kódovaná jednotka sekvencie. Obrázok pozostáva z 3 obdĺžnikových matic reprezentujúcich jas ( $Y$ ) a 2 chrominančné bloky  $Cb$  a  $Cr$ . Matica  $Y$  má páry počet riadkov aj stĺpcov,  $Cb$  a  $Cr$  matrice sú jedna polovica veľkosti matice  $Y$  v každom smere.

### 4. Rez

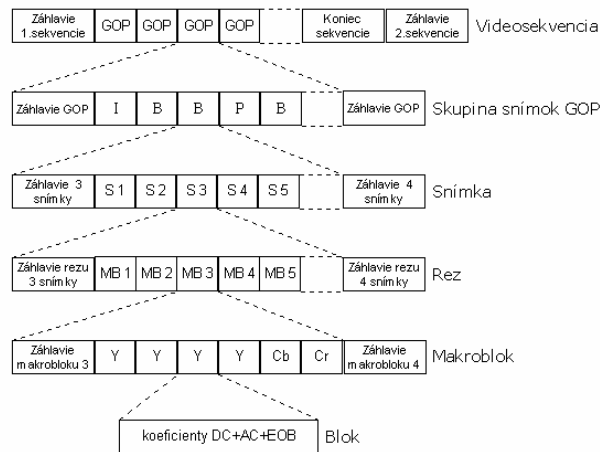
Skupina jedného alebo viacerých makroblokov. Poradie makroblokov v rámci rezu je zľava doprava a zhora nadol. Rezy sú dôležité v manipulácií s chybami. Ak stream obsahuje chybu, dekodér môže preskočiť na začiatok ďalšieho rezu.

### 5. Makroblok

Obsahuje 4 bloky  $Y$ , 1  $Cb$  a 1  $Cr$  blok. Čísla zodpovedajú usporiadaniu blokov v dátovom toku. Je základnou jednotkou kódovania vo vnútri snímku.

### 6. Blok

8x8 sada hodnôt jasovej alebo chrominančnej zložky [27].



Obrázok 6 Hierarchia rámcov MPEG video-toku

### 3.1.4 Faktory vplývajúce na kvalitu streamu

Kvalitu služby do značnej miery ovplyvňujú aktuálne parametre prístupovej siete. Prenášaný tok vyžaduje dostatočnú šírku pásma, malé oneskorenie, nízke kolísanie oneskorenia, nízku stratovosť paketov. Vplyvom zhoršenia týchto parametrov nastáva zhoršenie kvality služby u koncového používateľa. Vlastnosti ovplyvňujúce kvalitu:

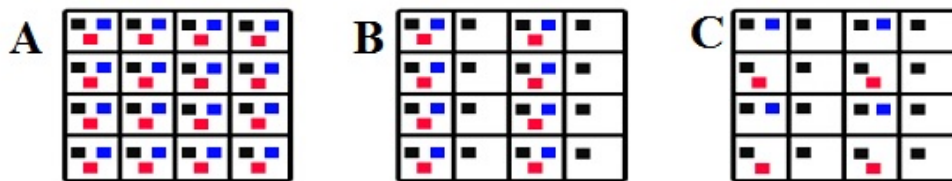
- Celkové oneskorenie (latencia) – čas, ktorý uplynie od odoslania správy od zdrojového uzla až po jej prijatie na cieľovom uzle. Zahrňuje oneskorenie na prenosovej trase a zariadeniach, ktoré sú jej súčasťou.
- Zmena poradia paketov
- Kolísanie oneskorenia (jitter) – predstavuje variabilitu v doručovaní paketov cieľovému uzlu (oneskorenie pri prenose). Spôsobuje pretečenie alebo podtečenie vyrovnávajúcej pamäti v set-top boxe.
- Stratovosť paketov – priemerný počet stratených paketov za určité obdobie vyjadrený v % vzhľadom k celkovému počtu prenesených paketov.
- Priepustnosť – udáva kapacitu prenosového kanála.
- Delay – je oneskorenie premenlivej veľkosti a vzniká naplnením paketov do fronty na odchádzajúcom interface.
- Rozlíšenie prenášaného videa (viac informácií o rozlíšení videa popísané v [28])
- Podvzorkovanie videa

## Podvzorkovanie videa

Kamery berú svetlo prichádzajúce do objektívu a konvertujú ho do 3 setov čísel, jeden pre červenú, jeden pre zelenú a jeden pre modrú. Nazýva sa to RGB. Ťažko s ním pracuje – napríklad pri potrebe znížiť rovnomerne jas na obrázku, je nevyhnutné upraviť to pre všetky 3 farby. Je možné použiť iný spôsob nazývaný  $YCbCr$ , ktorý rozdelí signál jasového kanála  $Y$  a do dvoch kanálov, ktoré ukladajú farbu bez informácií o jase – modrý a červený kanál. Niekedy je tento signál označovaný ako  $YUV$ .

Vo video priestore sa udávajú číselné pomery 4:4:4, 4:2:2 alebo 4:2:0. Vo farebných kanáloch bez akejkoľvek informácie o jase je ťažké vidieť nejaké podrobnosti. Voľným okom lepšie vnímame jas ako farbu. Táto skutočnosť platí aj pre pohyblivý obraz. Ak video používa  $YUV$  farebný priestor, najdôležitejšie dáta sú nesené v kanáli  $Y$ . Môžeme zahodiť veľa informácií o farbe a nie je možné, aby to spozoroval bežný divák [29].

Jeden používaný trik spočíva v tom, že sa v každom druhom pixely nezachovávajú informácie o farbe bez toho, aby to bolo pozorovateľné. Tento jav sa nazýva podvzorkovanie. Na obrázku 7A je znázornené podvzorkovanie 4:4:4, kde je znázornená iba malá časť rámca – matica pixelov 4x4 v obrázku. V tomto prípade má každý pixel hodnotu  $Y$ ,  $Cb$  a  $Cr$ . Ak spočítame počet výskytov  $Y$ ,  $Cb$  a  $Cr$  v riadku, pri každom dostaneme 4, čiže sa jedná o podvzorkovanie 4:4:4. Je veľmi zriedkavé vidieť záznamové zariadenie alebo kameru s výstupom 4:4:4. Vzhľadom na to, že ľudské oko skutočne nereaguje na odstránenie informácií o farbe, často sa používa podvzorkovanie 4:2:2. Ukážka tohto typu podvzorkovania je na obrázku 7B. MPEG veľmi často využíva podvzorkovanie 4:2:0. Je to postačujúce pre kvalitné vysielanie v TV. Ukážka tohto typu podvzorkovania je na obrázku 7C.



Obrázok 7 Podvzorkovanie

## 3.2 Hodnotenie kvality videa

Aká je kvalita pri sledovaní prijímanej videosekvencie? Kým subjektívne experimenty sú prohibatívne, objektívny algoritmus môže byť vypočítaný a porovnaný, vedie k výskumu nových typov analýz.

### 3.2.1 Subjektívne metódy

Subjektívne hodnotenie vychádza z výberu ľudí rôznej vekovej kategórie a pohlavia pre hodnotenie videosekvencií na základe vnímania kvality danej videosekvencie. Výhodou je popísanie ako vníma kvalitu danej videosekvencie človek, pričom sú vynechané vlastnosti, ktoré nie sú postrehnuteľné ľudským vnímaním. Nevýhodou je veľká časová náročnosť a tiež potreba vhodného vybavenia pre vykonanie tohto experimentu.

Medzinárodné odporúčania pre subjektívne hodnotenie kvality videa zahŕňajú špecifikácie ako vykonávať veľa rozličných typov subjektívnych testov. Niektoré z týchto metód sú dvojnásobne stimulové a vyžadujú hodnotenie v dvoch fázach, kde pozorujúci hodnotia kvalitu alebo zmenu v kvalite medzi dvoma videami (originálom a streamovaným videom – napríklad DSCQS - Double Stimulus Continuous Quality Scale, DSCS - Double Stimulus Comparison Scale). Iné sú jednostimulové, tu diváci hodnotia kvalitu iba jedného, streamovaného videa, napríklad SSCQE (Single Stimulus Continuous Quality Evaluation). Každá metóda má určité výhody, napríklad DSCQS je menej citlivá na obsah. SSCQE poskytuje časovo efektívnejšie hodnotenie kvality poskytovaného videa [30, 31, 32].

### **ACR (Absolute Category Rating)**

Úplne kategorické ohodnotenie - vychádza z odporúčania ITU-T P.910 [33], radí sa medzi nereferenčné metriky. Človek si pozrie videosekvenciu a následne hodnotí kvalitu videnej videosekvencie, pričom výsledky sú vyhodnotené ako priemerná hodnota na stupnici MOS (Mean Opinion Score), kde hodnotenie 5 znamená vynikajúcu kvalitu obrazu a hodnotenie 1 znamená veľmi zlú kvalitu obrazu.

*Tabuľka 1 Stupnica MOS*

<b>Ohodnotenie</b>	<b>MOS skóre</b>
Vynikajúca	5
Dobrá	4
Stredná	3
Slabá	2
Zlá	1

### **DCR (Degradation Category Rating)**

Tu sa jedná o kategorické hodnotenie degradácie. Táto metóda je tiež známa ako metóda DSIS (double stimulus impairment scale), vhodná na posúdenie málo poškodených súborov, kde sa hodnotenie posudzuje s vyššou citlivosťou. Pri tejto metrike sa posúdenie kvality vykonáva porovnaním aj referenčnej aj testovacej videosekvencie. Najskôr je prezentovaný originál, pričom hodnotiaci vie o tom, že to je originálna videosekvencia. Výsledkom je hodnota na stupnici DMOS – degradácia MOS. Škála ohodnotenia je 5-stupňová a obsahuje body:

- 5 – degradácia nie je vnímaná
- 4 – degradácia je viditeľná, ale neruší
- 3 – degradácia príliš neruší
- 2 – degradácia ruší
- 1 – degradácia príliš ruší

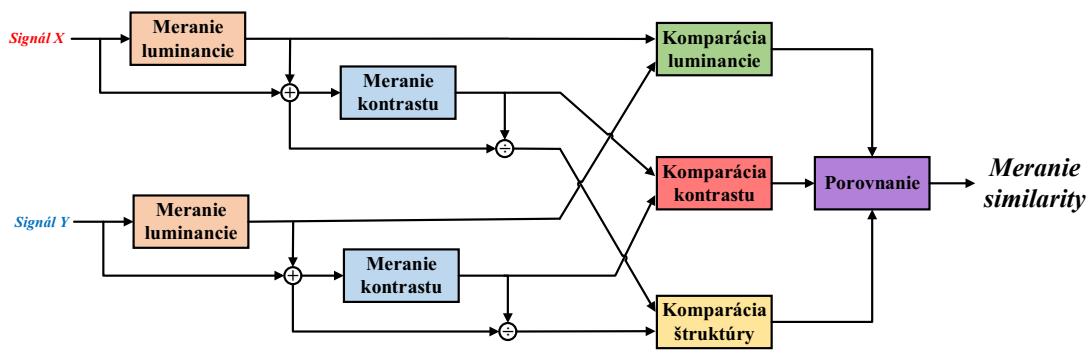
### 3.2.2 Objektívne metódy

Na objektívne hodnotenie kvality videa sa používajú objektívne metódy založené na výpočte metrík:

#### SSIM (Structural Similarity Index)

Zohľadňuje vnímanie obrazu ľudským okom. Porovnáva referenčnú videosekvenciu s prenášanou a vyhodnocuje zmenu jasů, kontrastu a iných chýb v prenesenom obraze pri porovnaní s originálnou snímku.

Používa jednoduchý funkčný model HVS (Human Visual System). Vyhodnotenie kvality videosekvencií je rozdelené do 3 vrstiev – lokálna vrstva blokového rozdelenia po obdĺžnikoch, mapa skreslenia získaná na úrovni rámca a vrstva vyhodnocujúca celú videosekvenciu. Najskôr sa získajú bloky 8\*8 pixelov z referenčnej aj prenesenej snímky. V tomto kroku sa vypočíta SSIM index pre každý blok zvlášť, pre každý komponent –  $Y$ ,  $Cb$  a  $Cr$ . V druhom kroku sa kombinujú hodnoty lokálnych vrstiev blokov ako hodnota kvality rámca. Kvalita vrstiev blokov je vypočítaná v súlade s úrovňou jasů (tmavé oblasti sú menej náchylné k degradácii kvality). V poslednej fáze sa počíta kvalita pre celú videosekvenciu. Úroveň kvality rámca je počítaná použitím pohybových vektorov rámca, pretože niektoré typy poškodenia (rozmazanie) nemajú vplyv na vnímanú kvalitu v scéne s prítomnosťou pohybu.



Obrázok 8 Bloková schéma metódy SSIM

Referenčné hodnoty SSIM sú v intervale  $[0,1]$ , kde 0 znamená nulový vzťah k originálnemu obrázku a 1 znamená, že obrázky sú totožné. Výsledná hodnota sa vypočíta:

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (3)$$

kde  $x$  je originálny signál,  $y$  je kódovaný signál, člen  $l(x, y)$  porovnáva jas signálu,  $c(x, y)$  porovnáva kontrast signálu a  $s(x, y)$  meria štruktúrnu koreláciu.

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad c(x, y) = \frac{2\zeta_x\zeta_y + c_2}{\zeta_x^2 + \zeta_y^2 + c_2} \quad s(x, y) = \frac{2\zeta_{xy} + c_3}{\zeta_x\zeta_y + c_3} \quad (4)$$

kde  $\mu_x$  a  $\mu_y$  predstavuje priemer zo vzoriek  $x$  a  $y$ ,  $\zeta_x$  a  $\zeta_y$  predstavuje rozptyl zo vzoriek  $x$  a  $y$ .

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i \quad \mu_y = \frac{1}{N} \sum_{i=1}^N y_i \quad (5)$$

$$\zeta_x = \left( \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2} \quad \zeta_y = \left( \frac{1}{N-1} \sum_{i=1}^N (y_i - \mu_y)^2 \right)^{1/2} \quad (6)$$

kde  $\zeta_{xy}$  predstavuje kovarianciu vzoriek  $x$  a  $y$ . Konštanty  $C1$ ,  $C2$  a  $C3$  sa používajú ku stabilizácií, v prípade, že priemer a rozptyl sú veľmi malé. Parametre  $\alpha > 0$ ,  $\beta > 0$  a  $\gamma > 0$  sa používajú na nastavenie relatívnej dôležitosti. Pre zjednodušenie sa v SSIM tieto parametre nastavujú zjednodušene ( $\alpha=\beta=\gamma=1$  a  $C3=C2/2$ ). Potom pôvodnú rovnicu (3) môžeme zjednodušene napísať [14, 32, 34, 35, 36]:

$$SSIM(x, y) = \left( \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left( \frac{2\zeta_x\zeta_y + C_2}{\zeta_x^2 + \zeta_y^2 + C_2} \right) \quad (7)$$

### PSNR (Peak Signal to Noise Ratio)

Je to jedna z najjednoduchších metód vyhodnocujúcich kvalitu videa. Meria pomer maximálneho výkonu signálu k šumu skreslenia signálu. Je zvyčajne logaritmicky vyjadrená v decibeloch. Používa sa na určenie vizuálnej podobnosti referenčného a preneseného snímku. Neaplikuje žiadne perceptuálne kritéria na implementáciu vnímaného šumu. Čím vyššia hodnota, tým väčšia kvalita komprimovaného obrazu. Stredná kvadratická chyba (MSE) a odstup signálu od šumu sú dve chybové metriky používané pre porovnanie preneseného obrázku a originálu. MSE predstavuje kumulatívnu kvadratickú odchýlku medzi testovanou a originálnou vzorkou, zatiaľ čo PSNR reprezentuje mieru chybovosti medzi dvoma obrazmi.

$$PSNR = 10 \log \frac{m^2}{MSE} \quad (8)$$

kde  $m$  je maximálna hodnota, ktorú môže pixel nadobudnúť (napríklad pre 8 bitové obrázky sa  $m=255$ ) a MSE, popísaná v kapitole 5.4 (21), je priemer kvadratických rozdielov medzi šedými úrovňami hodnôt pixelov v 2 snímkach alebo sekvenciách  $I$  a  $\tilde{I}$ .

Rozsah PSNR hodnoty je od 1 dB do 100 dB. Ak by hodnota bola 100 dB, tak oba obrázky sú identické.

### VQM (Video Quality Metric)

Táto metrika bola vyvinutá inštitútom pre telekomunikačné vedy (ITS). Existuje v niekoľkých rôznych implementáciách, ktoré sú optimalizované pre špecifické oblasti. Počíta viditeľnosť artefaktov v doméne DCT (Discrete Cosine Transform). Na obrázku 10 je znázornený blokový diagram tejto metriky, ktorý je rozdelený do 9 krokov.



Obrázok 9 Blokový diagram VQM metriky

Vstupom do metriky je dvojica obrazových sekvencií – referenčná sekvencia a testovacia sekvencia. Obe sekvencie sa orežú, následne sa obe sekvencie transformujú do rovnakého farebného priestoru YOZ. V ďalšom kroku sa aplikuje na farebne upravený signál DCT transformácia. Tu sa rozdelia vstupné snímky do rôznych priestorových frekvenčných komponentov. Potom sa konvertuje každý DCT koeficient na lokálny kontrast (LC) s použitím rovnice:

$$LC(i, j) = DCT(i, j) \frac{(DC/1024)^{0.65}}{DC} \quad (9)$$

kde DC je jednosmerný komponent pre každý blok. Pre 8-bitový obrázok je stredná hodnota DCT 1024. Hodnota 0,65 je najlepší parameter pre prispôbenie. Po tomto kroku sa väčšina údajov nachádza v intervale [-1, 1].

V ďalšom kroku sa aplikuje časové filtrovanie a funkcia priestorového vnímania kontrastu SCSF (Spatial Contrast Sensitivity Function). V jednom kroku sa používa matica SCSF pre statické snímky a jedna matica pre dynamické snímky. Tým sa znížia pamäťové a výpočtové nároky. DCT koeficienty vyjadrené v lokálnej kontrastnej forme sú prevedené na práve pozorovateľné rozdiely delením ich príslušných priestorových hodnôt. To implementuje priestorovú časť funkcie kontrastnej citlivosti. Z týchto dvoch sekvencií sa urobí rozdiel. Vypočíta sa vážený priemer a stanovuje sa maximálne skreslenie. Tieto časti sú odčítané ako prvé. Nasleduje maskovanie kontrastu a určenie maxima. Výstupná hodnota VQM metriky indikuje množstvo skreslenia sekvencie – pre neznižujúcu sa hodnotu je rovná nule a pre stúpajúcu hodnotu zhoršenia, výstupná hodnota stúpa [31, 37, 38, 39].

### 3.3 Časová a priestorová informácia scény

V zmysluplnej VQA (Video Quality Assessment) databáze musí byť spektrálny obsah videosekvencií charakterizovaný množstvom rozdielnych časových a priestorových informácií. Tieto informácie charakterizujú danú scénu.

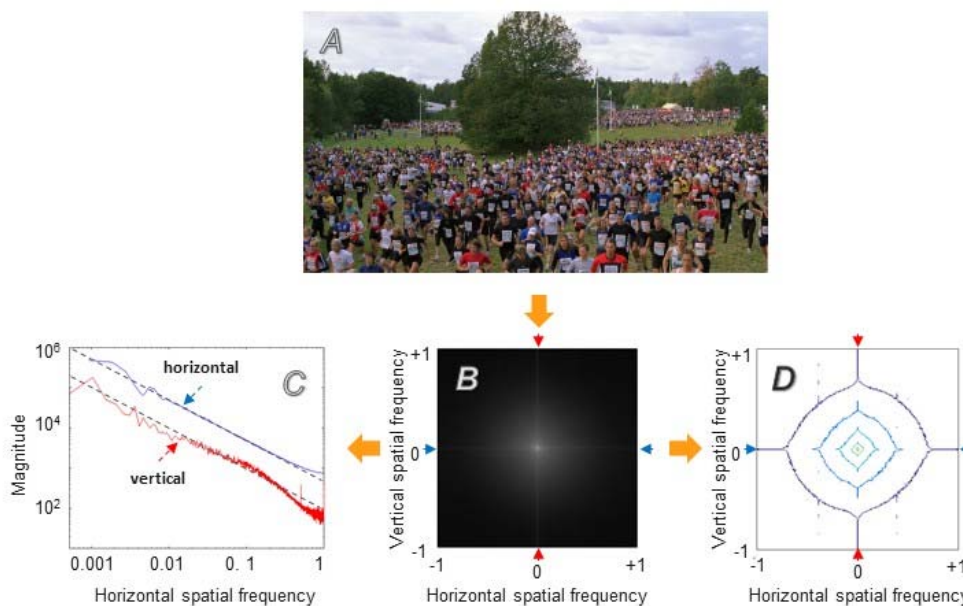
#### Štatistika prírodných scén

Obrázky prírodných scén majú zaujímavé štatistické vlastnosti – majú priestorovú frekvenciu magnitúdy spektra, ktoré majú tendenciu klesať s rastúcou priestorovou frekvenciou v pomere k inverznej priestorovej frekvencii [40]. Veľkosť spektra jednotlivých snímok sa môže výrazne líšiť, ale ako súbor – priemerné štatistické očakávania, možno povedať, že veľkosť spektra obrázkov prírodných scén spadajú pod jednu priestorovú frekvenciu. To sa vzťahuje aj na horizontálne aj na vertikálne priestorové frekvencie.



## Priestorová frekvencia

Obraz je väčšinou tvorený dvojrozmerným poľom pixelov, kde každý pixel je reprezentovaný hodnotami červenej, zelenej a modrej informácie alebo luminačným a 2 chromatickými kanálmi (ako už bolo popísané v práci). Obraz môže byť reprezentovaný tiež ako dvojrozmerné pole priestorovo-frekvenčných komponentov ako je zobrazené na obrázku 10. Obraz založený na vizuálnom pixeli a priestorovej frekvencii reprezentujúcej vizuálny obraz sú matematicky zameniteľné. Majú rovnaké informácie, ale odlišne vyjadrené.



Obrázok 10 Ilustrácia reprezentujúca obraz v priestorovej frekvencii

Priestorovo-frekvenčná prezentácia obrazu má komponent magnitúdy a komponent fázy. Komponent magnitúdy nazývaný spektrum magnitúdy poskytuje informácie koľko z celkového počtu variácií vizuálneho (pixelovo založeného) obrazu možno pripísať na určitú priestorovú frekvenciu (priestorová frekvencia je dvojrozmerná, majúca horizontálnu a vertikálnu časť). Komponent fázy nazývaný fázové centrum poskytuje informácie ako rôzne vzájomne pôsobia priestorové informácie na vytvorenie funkcií a detailov organizovaných v obraze.

Ako je ilustrované na obrázku 10/A, obraz založený na vizuálnom zobrazení môže byť reprezentovaný ako dvojrozmerné pole komplexných čísel použitím Fourierovej transformácie. Absolútna hodnota komplexných čísel je zobrazená ako dvojrozmerné spektrum magnitúdy (10/B), v ktorej jasnejšie oblasti korešpondujú k väčším hodnotám magnitúdy (Záznam o rozsahu spektra je znázornený na obrázku 10/B ako pomoc vizualizácie. Horizontálne a vertikálne frekvenčné osi sú uvedené vo vzťahu k zodpovedajúcej Nyquist frekvencii ( $\pm 1$ )). Magnitúdy hlavnej horizontálnej a vertikálnej priestorovej frekvencie sú zobrazené na obrázku 10/C. Hlavná horizontálna os priestorovej frekvencie korešponduje s nulovou vertikálnou frekvenciou (modré šípky na obrázku 10/B) a hlavná vertikálna os priestorovej frekvencie korešponduje s nulovou horizontálnou frekvenciou (červené šípky na obrázku 10/B). Prerušovaná čiara na obrázku 10/C indikuje 1/štatistické očakávanie priestorovej frekvencie pre dosiahnutie prirodzených scén [40].

Obrysové mapy záznamov magnitúdového centra je užitočné použiť pre vytvorenie dvojrozmerného zloženia priestorovej frekvencie obrázkov. Príklad týchto máp môžeme vidieť na obrázku 10/D.

### 3.4 Neurónová sieť

Neurónová sieť je jedným z výpočtových modelov používaných v umelej inteligencii a jej princípy sú inšpirované skutočnými biologickými štruktúrami. Princíp neurónovej siete spočíva v prirodzených procesoch, ktoré prebiehajú v mozgu, na úrovni elektrického signálu a jeho prenosu medzi nervovými bunkami (neurónmi) a nervovými vláknami. Model neurónovej siete zjednodušuje procesy prebiehajúce v mozgu. Neurón pozostáva zo vstupov a výstupov, kde vzťah vstupu k výstupu je daný zvolenou prenosovou funkciou. Každý vstup ma priradené váhy, ktoré znamenajú významnosť daného vstupu a simulujú tak fungovanie prepojenia jednotlivých vlákien (synapsí). Prepojením viacerých neurónov a ich súčasným využitím je možné plne využiť schopnosť neurónu čo je porovnateľné s fungovaním ľudského mozgu, čo sa neurónová sieť snaží napodobniť.

Spôsob určujúci štruktúru neurónov v sieti je spojený s použitým štruktúrovaným algoritmom učenia. Je možné definovať 3 základné druhy sieťových architektur:

- Jednvrstvové dopredné siete – najjednoduchšia forma tohto typu, obsahuje vstupnú vrstvu premietajúcu sa na výstupnú. Tento algoritmus funguje iba v jednom smere, čiže môžeme hovoriť o doprednej, resp. acyklickej sieti. Výpočet prebieha iba na výstupnej vrstve, vstupná vrstva teda ani je započítaná do celkového počtu vrstiev.
- Viacvrstvové dopredné siete – tento typ siete obsahuje aspoň jednu skrytú vrstvu, ktorých výpočtové uzly sa nazývajú skryté neuróny. Ich funkciou je zasahovať do spoja medzi externým vstupom a výstupom siete. Pri zväčšovaní veľkosti vstupnej vrstvy zvyšujú skryté neuróny výkon siete. Neuróny sú organizované formou jednotlivých vrstiev, ktoré vzájomným prepojením tvoria topológiu danej siete. Neuróny medzi susednými vrstvami sú úplne prepojené, zatiaľ čo neuróny v rámci jednej vrstvy prepojené nie sú. Obvykle vytvorená topológia pozostáva zo vstupnej vrstvy, skrytých vrstiev a výstupnej vrstvy. Skrytých vrstiev môže byť definovaných viac, neexistuje predpísaný počet. Záleží to od hierarchie a početnosti dát. Vstupná vrstva prijíma a preposiela do siete hodnoty vstupného vektoru. Na každej zo skrytých vrstiev sa aplikuje výpočet váh na základe stanovenej aktivačnej funkcie. Výstupná vrstva prijíma údaje zo skrytej vrstvy a preposiela ich na výstup. Neurónová sieť by mala mať lichobežníkový, prípadne trojuholníkový tvar klesajúci smerom k výstupu, aby sieť mohla generalizovať.
- Rekurentná sieť – na rozdiel od viacvrstvovej doprednej siete pridáva ešte rekurentné neuróny, ktoré majú prepojené ich vstupy s výstupmi neurónovo nasledujúcich vrstiev. Tieto typy sietí sú výhodné pri dátach, ktorých javy sú závislé na ich predošlej hodnote.

Niekedy je nutné zosilniť alebo utlmiť signál určitého neurónu. Pre tento účel slúžia biasové neuróny, ktoré nemajú žiadny vstup a ich výstup je '1'. Do ďalšej vrstvy prenášajú hodnotu váhy prenášaného prepojenia.

Primárnou vlastnosťou neurónových sietí je vlastnosť učiť sa a zlepšovať svoju výkonnosť na základe ponúkaných dát. Je to proces, ktorého úspešnosť závisí na zložitosti danej úlohy a tiež na veľkosti a zložitosti vytvorenej siete. Sieť sa každým krokom nastavenia váh spojení a biasov stáva dokonalejšou, čo vedie v ideálnom prípade k jej adaptácii na potrebnú úlohu. Samotné učenie siete je založené na výbere vhodného algoritmu učenia sa. Existuje viacero druhov algoritmov, ktoré sa odlišujú v spôsobe úpravy parametrov siete, čiže váh a biasov. Cieľom je natréňovať sieť, ktorá bude korektne odpovedať na súbor vstupných hodnôt s minimálnou chybovosťou, respektíve požadovanou presnosťou. Toto je typ „trénovania s učiteľom“ (supervised learning), čo znamená, že siete sú v procese trénovania poskytnuté vstupy a ich prislúchajúce výstupy[41, 42].

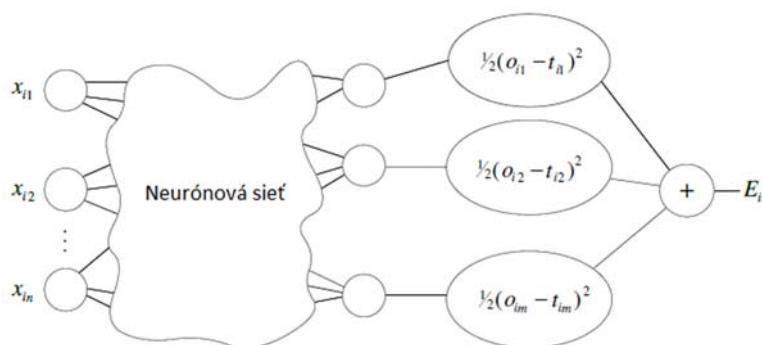
## Backpropagation algoritmus

Hľadá minimum chybovej funkcie v priestore váh s použitím metódy *gradient descent*. Funkcia sa považuje za riešenie problému učenia. Pretože táto metóda vyžaduje výpočet gradientu chybovej funkcie v každom iteračnom kroku, musíme zaručiť kontinuitu a diferenciáciu chybovej funkcie. Je zrejmé, že musíme použiť inú aktivačnú funkciu než funkciu použitú v perceptrónoch.

Backpropagation algoritmus môže byť ľahko odvodený výpočtom gradientu. Dopredná neurónová sieť je výpočtovým grafom, ktorého uzly sú počítajúce jednotky a ktorého hrany s daným smerom prenášajú číselné informácie z uzla do uzla. Každá výpočtová jednotka je schopná vyhodnotiť funkciu jej vstupu. V skutočnosti sieť predstavuje reťazec funkčných kompozícií, ktoré transformujú vstup na výstupný vektor (nazývaný vzor). Sieť je konkrétna implementácia kompozitnej funkcie od vstupu do výstupného priestoru, ktorú nazývame sieťová funkcia. Cieľ spočíva v hľadaní optimálnej kombinácie váh tak, že sieťová funkcia aproximuje danú funkciu najlepšie ako to je možné. Dopredná sieť s  $n$  vstupmi a  $m$  výstupmi môže pozostávať z ľubovoľného počtu skrytých uzlov. Dostávame tiež tréningový set  $\{(x_1, t_1), \dots, (x_p, t_p)\}$  pozostávajúci z  $p$  usporiadaných dvojíc a  $m$ -rozmerných vektorov, ktoré tvoria vstup a výstup vzoru. V každom uzle siete sú funkcie kontinuálne a diferencované. Váhy počiatočných uzlov sú reálne čísla generované náhodne. Keď je vstupná hodnota  $x_i$  z tréningovej sady prezentovaná do siete, produkuje výstup  $o_i$  odlišný ako požadovaný výstup  $t_i$ . Cieľom je predikovať  $o_i$  rovnako ako  $t_i$  pre  $i = 1, 2, \dots, p$  použitím algoritmu učenia. Presnejšie cieľom je minimalizovanie chybovej funkcie definovanú ako:

$$E = \frac{1}{2} \sum_{i=1}^p \|o_i - t_i\|^2 \quad (10)$$

Po minimalizovaní tejto funkcie pre tréningovú sadu, sú prezentované siete nové neznáme vstupné vzory s očakávaním, že sieť bude schopná predikovať požadovaný výstup pre dané hodnoty. Sieť musí rozpoznať, či je nový vstupný vektor podobný vzorkám a produkovať podobný výstup. Spätňý propagačný algoritmus sa používa na nájdenie lokálneho minima chybovej funkcie. Sieť je inicializovaná s náhodne vybranými váhami. Gradient funkcie chyby sa vypočíta a použije na úpravu počiatočných váh. Našou úlohou je rekurzívne vypočítať tento gradient.



Obrázok 11 Rozšírená sieť pre výpočet chybovej funkcie

Prvým krokom procesu minimalizácie je rozšírenie siete tak, aby automaticky vypočítala chybovú funkciu. Obr.11 ukazuje akým spôsobom je tento proces vykonávaný. Každá z  $j$  výstupných jednotiek siete je pripojená k uzlu, ktorý vyhodnocuje funkciu

$$\frac{1}{2}(o_{ij} - t_{ij})^2 \quad (11)$$

Kde  $o_{ij}$  a  $t_{ij}$  označujú  $j$ -tý komponent výstupného vektora  $o_i$  a požadovaného výstupu  $t_i$ . Výstup dodatočných uzlov  $m$  zoskupovaných v uzle, ktorý ich vyhodnotí ako súčet  $E_i$  ako výstup. Pre každý  $t_i$  musí byť vytvorené rovnaké rozšírenie. Výpočtová jednotka zhromažďuje kvadratické chyby a výstupy ako súčet chýb  $E_1, \dots, E_p$ . Výstupom tejto rozšírenej siete je chybová funkcia  $E$ . Teraz je sieť schopná vypočítať celkovú chybu pre daný tréningový set. Váhy v sieti sú jediné parametre, ktoré môžu byť modifikované pre zníženie kvadratickej chyby  $E$  na najnižšiu hodnotu ako je to možné. Pretože  $E$  je vypočítaná rozšírenou sieťou výlučne prostredníctvom funkcií uzlov, je to kontinuálna a diferencovateľná funkcia váh  $w_1, w_2, \dots, w_l$  v sieti. Môžeme teda minimalizovať  $E$  použitím iteračného procesu znižovania gradientu, pre ktorý je potrebné spočítať gradient

$$E = \left( \frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_l} \right) \quad (12)$$

Každá váha je aktualizovaná prírastkom:

$$\Delta w_i = -\gamma \frac{\partial E}{\partial w_i} \quad (13)$$

Pre  $i=1, \dots, l$ , kde  $\gamma$  predstavuje konštantu učenia, parameter proporcionality, ktorý definuje dĺžku kroku každej iterácie v smere negatívneho gradientu. S týmto rozšírením pôvodnej siete, je celý problém učenia sa zameraný na výpočet gradientu funkcie siete s ohľadom na jeho váhy. Keď máme metódu na výpočet tohto gradientu, hmotnosť sietí môžeme iteračne upraviť. Týmto spôsobom očakávame, že nájdeme minimálnu chybovú funkciu.

Keďže chceme minimalizovať chybovú funkciu  $E$ , ktorá závisí od váh siete, musíme sa zaoberať so všetkými váhami v sieti. Dopredný krok sa vypočíta obvyklým spôsobom, ale je tiež potrebné uložiť výstup každej jednotky na pravej strane. Krok spätnej propagácie vykonáme v rozšírenej verzii siete, ktorá vypočíta chybovú funkciu a potom sa zameriava na jednu z váh, napr.  $w_{ij}$  priradenú k okrajovému bodu z  $i$ -tého do  $j$ -tého uzla siete. Táto váha môže byť považovaná za vstupný kanál do podsiete vytvorenej zo všetkých ciest začínajúcich v  $w_{ij}$  a končiacich v jednej výstupnej jednotke siete [43].

## 4 Ciele dizertačnej práce

Na súčasný stav nadväzujú ciele dizertačnej práce a ďalší postup v oblasti hodnotenia kvality a videa a to výskumu efektívneho používania bitového toku. Vyrovnávajúca pamäť je veľmi dôležitá súčasť kódera a dekódera. Vyrovnáva premenlivý dátový tok v závislosti od obsahu aktuálnej scény a použitých kompresných prostriedkoch. Statický záber moderátora televíznych správ alebo politickej relácie potrebuje nižší bitový tok ako akčná vysoko dynamická scéna, reklamný spot alebo športový prenos (samozrejme vysoko dynamický šport; aj túto kategóriu je potrebné odlišovať a brať do úvahy športy s rýchlou dynamikou ako hokej, futbal; dynamický šport ako šípky a skoro statický šport ako šach). Preto je cieľom tejto práce zistiť, pri akom bitovom toku dosahujú jednotlivé scény najvyššie uspokojenie koncových užívateľov z pohľadu kvality a definovať tak hranice pre jednotlivé scény na základe priestorovej a časovej informácie obsahu. Pre jednotlivé hranice by bol priradený vhodný bitový tok, tak, aby spĺňal najvyššiu kvalitu v jednotlivom rozlíšení.

Pri nasadení VBR (Variable Bit Rate) môžeme ušetriť značnú časť priemerného bitového toku CBR (Constant Bit Rate), pričom je veľmi dôležitá priemerná a tiež maximálna hodnota bitového toku pri použití VBR. Je potrebné vybrať vhodné testovacie video scény na určenie kvality videa a ich kódovanie. Tieto video scény reprezentujú širokú škálu obsahu ako je textúra snímky, detail snímky, rýchlosť pohybu objektu v snímke, panoráma snímky, nastavenie objektívu kamery. Sú založené na rôznej dynamickosti – od statickej až po vysoko dynamické. Tieto scény budú kódované do rozlíšení, ktoré sú používané a s ktorými sa plánuje v blízkej dobe. Pri kódovaní bude použitá široká škála bitového toku, aby sa tok dal dynamicky meniť na základe dynamickosti scény. Vybraté videosekvencie budú testované objektívnymi metódami. Tieto výsledky budú porovnané so subjektívnym hodnotením testovacích videosekvencií. Subjektívne testovanie bude vykonané podľa odporúčaní ITU a hodnotiť budú rovnako ženy ako aj muži v rôznych vekových kategóriách.

Ciele realizácie dizertačnej práce sú popísané v nasledujúcich bodoch:

- Vyjadrenie korelácie výsledkov v hodnotení kvality videa medzi objektívnymi a subjektívnymi metódami

Korelácia vyhodnotí mieru závislosti medzi subjektívnym hodnotením rôznych užívateľov a objektívnym ohodnotením na základe výsledkov objektívnych metrík.

- Vytvorenie optimálnej mapovacej funkcie predikcií objektívnych metrík a výsledkov subjektívnych hodnotení na určenie hraníc TI a SI

TI a SI informácie reprezentujú video obsah. Časová a priestorová perceptuálna informácia scény sú kritickými parametrami. Tieto parametre hrajú kľúčovú úlohu v určení množstva video kompresie. Analýza klasifikácie obsahu je vykonávaná výpočtom indexu časovej a priestorovej informácie vzhľadom na luminančný komponent každej videosekvencie. Pomocou funkcie mapovania určíme hranice pre časovú a priestorovú informáciu, na základe ktorej bude navrhnuté prahovanie a škálovateľnosť bitových tokov v modeli adaptívnej rezervácie sieťových zdrojov. Návrh a realizácia nového klasifikátora scény postaveného na modeli, ktorý určuje typ scény na základe TI a SI informácie. Podľa typu scény model určí vhodný bitový tok na základe mapovacej funkcie vychádzajúcej zo subjektívnych hodnotení.

- Verifikácia a validácia modelu

Verifikácia návrhu v experimente, v ktorom bude porovnaná predikovaná kvalita videa v prenosovom prostredí využívajúcom VBR podľa navrhnutého klasifikátora s požadovanou hodnotou bitového toku.

## 5 Realizácia cieľov dizertačnej práce

Táto kapitola popisuje jednotlivé kroky vedúce k hlavným stanoveným cieľom tejto práce definovaných v kapitole 4. Čiastočnými a nevyhnutelnými cieľmi tejto práce vedúcimi k hlavným cieľom sú vytvorenie vhodnej databázy videosekvencií na základe prekódovania vybraných typov videosekvencií do vybraných kvalitatívnych parametrov (rozlíšenie, kompresný štandard, bitový tok), ich klasifikácia na základe charakteristických znakov SI a TI a ohodnotenie kvality videosekvencií.

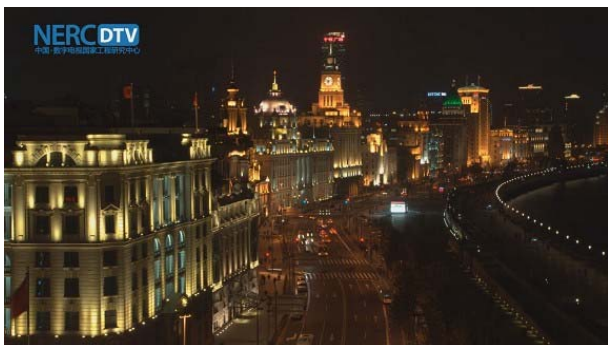
### 5.1 Vytvorenie databázy videosekvencií

Porovnanie zahŕňa rôzne typy videosekvencií a to statické aj dynamické. Vzhľadom k tomu, že obsah videa hrá kľúčovú úlohu v príslušnom výskume, sú použité videosekvencie pokrývajúce širokú škálu typu obsahu. Do úvahy sú tiež brané faktory ako textúra obrázku, detail obrázku, rýchlosť pohybu objektu, panoramatický pohyb, intenzita svetla (denné a nočné natáčanie). Výber scén je vykonaný na základe odporúčaní ITU-T [33, 52].

Pre potreby vytvorenia databázy videosekvencií boli použité vytvorené scény výskumníkmi zo *Shanghai Jiao Tong University* [44]. Tieto videosekvencie sú v nekomprimovanom \*.yuv formáte v UHD rozlíšení. Z databázy, ktorá poskytuje 15 rozdielnych videosekvencií, som vybral 8, ktoré široko pokrývajú všetky kvadranty z pohľadu TI a SI informácií. Všeobecne, tieto informácie reprezentujú video obsah. Priestorové a časové vnemové informácie o scénach sú kritickými parametrami. Tieto parametre zohrávajú rozhodujúcu úlohu určenia množstva kompresie videa a významný vplyv korešpondujúci s obsahom videa. Krátky popis použitých videosekvencií.

- Bund Nightscape - nočný záber architektúry mesta, obsahuje dynamický pohyb áut a chodcov, statické segmenty predstavujú mestské budovy. Kamera zachytáva scénu zo statickej polohy.
- Campfire Party – nočná scéna zobrazujúca plamene táborového ohňa pred skupinou ľudí, na prednej strane scény je horiaci plameň (rýchla zmena časovej a jasovej informácie). V pozadí scény je skupina ľudí, čo robí pozadie mierne statickým. Na konci sekvencie je priblížená táto skupina ľudí.
- Construction Field – pohľad na stavenisko, kde je prezentované statické pozadie budovami vo výstavbe, dynamické objekty sú prezentované vozidlami stavby a kráčajúcimi pracovníkmi. Jedná sa o scénu s miernym pohybom, zachytenú staticky.
- Fountains – sekvencia zachytávajúca mestskú fontánu pred vysokou budovou. Popredné miesto pozostáva zo striekajúcej vody, pozadie je statické, sú na ňom vyobrazené stromy a budovy. Jedná sa o statickú scénu s nízkou dynamikou pohybu.
- Marathon – bežecká súťaž v maratóne. Videosekvencia zobrazuje začiatok Medzinárodného preteku maratónu v Šanghaji. Bežci sú viacerými pohyblivými objektami s miernou dynamickosťou, na pozadí je statické zobrazenie cesty. Kamera zachytávajúca scénu je statická a zachytáva pohľad z výšky.
- Runners – Ide tiež o zábery z Medzinárodného preteku maratónu v Šanghaji. Na rozdiel od scény maratónu, je tu menej bežcov. Kamera je statická, umiestnená z prednej strany bežcov, mierne z uhla (vyššie priestorové informácie). Scéna je pomerne dynamická.

- Tall Buildings – jedná sa o záber na moderné mesto - vysoké budovy v Lujiazui Pudong New District v Šanghaji. Statické objekty sú mrakodrapy, rieky a infraštruktúry mesta. Mestská premávka predstavuje pomaly sa pohybujúce objekty. Kamera sa pohybuje pomaly zľava doprava. Scéna je charakteristická zmenou SI a TI informácie.
- Wood – lesná scenéria zachytávajúca stromy v lese, ktorým preniká slnečné svetlo. Objekty sú statické, pohyb kamery je zľava doprava a postupne zrýchľuje. Relatívne vysoká hodnota SI a TI informácií.



*Obrázok 12 Bund Nightscape*



*Obrázok 13 Campfire Party*



*Obrázok 14 Construction Field*



*Obrázok 15 Fountains*



*Obrázok 16 Marathon*



*Obrázok 17 Runners*





Obrázok 18 Tall Buildings

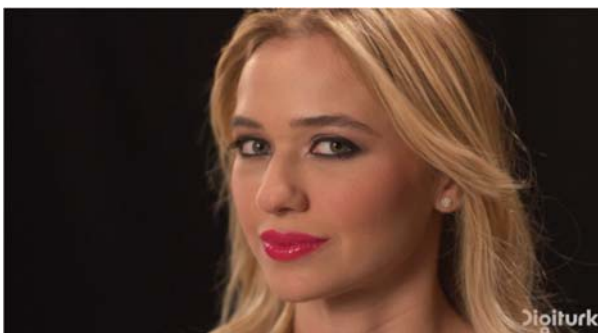


Obrázok 19 Wood

Vybrané videosekvencie boli prekódované použitím nástroja ffmpeg [47] na potrebný formát s podvzorkovaním 4:2:0, s bitovou hĺbkou 8b (bitov), rýchlosťou zobrazenia 30 fps (Frames per Second) s počtom snímok 300 (10s). Takto dlhé videosekvencie sú postačujúce pre subjektívne ohodnotenie, dlhšie sekvencie nemajú vplyv na kvalitatívne ohodnotenie videosekvencie [45].

Pre rozšírenie vytváranej databázy som vybral ďalšie 4 videosekvencie, ktoré vytvorila skupina výskumných pracovníkov z Fínskej Univerzity - *Tampere University of Technology*, skupina z laboratória *Laboratory of Pervasive Computing* (sekvencie Ultra Video Group) [46]. Tieto videosekvencie rozširujú databázu videosekvencií pre objektívne hodnotenie a predikciu bitového toku na základe objektívnej metriky SSIM.

- Beauty - detail ženskej tváre, jej vlasy pomaly fúkajú vo vetre, čierne statické pozadie
- Bosphorus – plavba loďou s obrovským mostom v pozadí, kamera sa pohybuje zľava doprava. Mierne dynamická scéna, kde sa pohybuje jeden objekt
- Jockey - bežiaci kôň s jazdcom, kamera posúvaná zľava doprava - jeden objekt s rýchlym pohybom
- ReadySteadyGo - preteky jazdcov na koňoch, niekoľko rýchlo pohybujúcich sa objektov (jazdci na koňoch), kamera sa presúva zľava doprava



Obrázok 20 Beauty



Obrázok 21 Bosphorus



Obrázok 22 Jockey



Obrázok 23 ReadySteadyGo

## 5.2 Proces kódovania

Táto podkapitola popisuje proces, akým sa vybrané typy videosekvencií kódujú do potrebného formátu, kde sa nastavuje bitový tok, rozlíšenie a kompresný štandard, následne je popísané dekódovanie, aby bolo možné porovnať prenášanú videosekvenciu s referenčnou.

### 5.2.1 Bitový tok

Pomocou bitového toku sa vyjadruje kvalita kompresie. Udáva počet bitov použitých na zakódovanie 1 sekundy sekvencie. Cieľom tejto práce je zistiť aký bitový tok je postačujúci na maximálne uspokojenie koncového užívateľa. VBR má v porovnaní s CBR lepšiu účinnosť kódovania. To udržuje konštantnú kvalitu pre celú videosekvenciu, čo je dôležité v obsahu s rýchlym pohybom (napríklad šport). Okrem toho, VBR umožňuje štatistický multiplex video streamov. Toto multiplexovanie umožňuje zachovať stálu kvalitu ponúkaného obsahu. Hlavnou nevýhodou tohto riešenia je komplikovanejšia sieťová prevádzka a dimenzovanie.

Kvalita kódovaného obsahu sa určuje podľa množstva dát, ktoré sa stratili pri komprimácii a dekomprimácii obsahu. Množstvo faktorov ovplyvňuje dáta v komprimačnom procese, ale vo všeobecnosti platí, že čím sú zložitejšie pôvodné dáta a čím je vyšší kompresný pomer, tým viac detailov je stratených pri komprimovaní.

Pri reálnom vysielaní sa používa bitový tok 10 Mbps pre rozlíšenie HD a niektoré stanice používajú bitový tok až okolo hodnoty 15 Mbps. S týmto tokom sa počíta aj pri nasadení UHD.

Cieľom meraní je zistiť efekt bitového toku pre video formát H.264/AVC - najpoužívanejší pre digitálne kódovanie televíznych staníc a H.265/HEVC, ktorý ešte nie je reálne nasadený, zatiaľ sa testuje, s jeho nasadením sa počíta v blízkej budúcnosti (pozn.: blízka budúcnosť od času, kedy bola táto práca písaná). Pri prekódovaní daných videosekvencií som použil utility x264 a x265 nástroja ffmpeg. Bitový tok som nastavil v rozmedzí od 1-20 Mbps (1, 3, 5, 8, 10, 12, 15 a 20 Mbps).

### 5.2.2 Rozlíšenie

Všetky videosekvencie som prekódoval z nekomprimovaného formátu (\*.yuv) na potrebný formát, rozlíšenie a bitový tok pomocou nástroja ffmpeg. Čoraz populárnejšie je video s vysokým rozlíšením, FHD (1920×1080), začína sa testovať a reálne aj nasadzovať rozlíšenie UHD (3840×2160),

ktoré bude trendom budúcnosti. Rastúce nároky na väčší zážitok koncového užívateľa spojeného s vysokou kvalitou obsahu (pre domáce kino alebo prémiové scény ako je šport) tlačia vývoj týchto technológií. Vybraté videosekvencie som kodoval do rozlíšenia FHD a UHD. 4K na rozdiel od UHD ma rozlíšenie 4096×2160, komerčne sa zatiaľ nevyužíva a je uprednostňované UHD z dôvodu ponúkaného obrazu 16:9. To je dôvod, prečo som použil rozlíšenie UHD a nie 4K.

### 5.2.3 Parametre kódovania

Pri kódovaní videosekvencií som použil podvzorkovanie 4:2:0, 8 bitovú hĺbku, dĺžku GOP som nastavil na hodnoty  $M=3$  a  $N=15$ . Formát *\*.ts* (MPEG-TS, transport stream) predstavuje digitálny kontajner na prenos a uloženie videosúborov, a je používaný pre prenos digitálneho vysielania. GOP som nastavil na polovicu framerate ( $M=3$ ,  $N=15$ ). Kódovanie a dekódovanie som vykonal prostredníctvom nástroja ffmpeg [47], kde vstupovali nasledujúce parametre:

- vstupné parametre:
  - `-f` typ videa
  - `-video_size` rozlíšenie sekvencie
  - `-pix_fmt` formát sekvencie
  - `-framerate` počet snímkov za sekundu
  - `-i` názov vstupnej sekvencie
- Možnosti kódovania:
  - `-vcodec libx264 (libx265)` použitý kodek
- Možnosti GOP (definovaná veľkosť GOPu v podobe počtu B snímkov za sebou) :
  - `15:min-keyint=15:bframes=3:b-adapt=1`
- Možnosti bitového toku:
  - `-bitrate`
  - `-vbv-maxrate`
  - `-vbv-buFSIZE`
- Výstupné parametre:
  - `-vf scale` prekódovaná veľkosť rozlíšenia
  - názov a formát výsledného súboru

Po vytvorení databázy, kódovanie do potrebného formátu bolo potrebné kódované videosekvencie opäť dekódovať, aby sa mohla porovnať kvalita referenčnej a degradovanej videosekvencie. Vzor príkazu potrebného na kódovanie a aj dekódovanie v nástroji ffmpeg je možné nájsť v prílohe B.1.

## 5.3 Klasifikácia scény

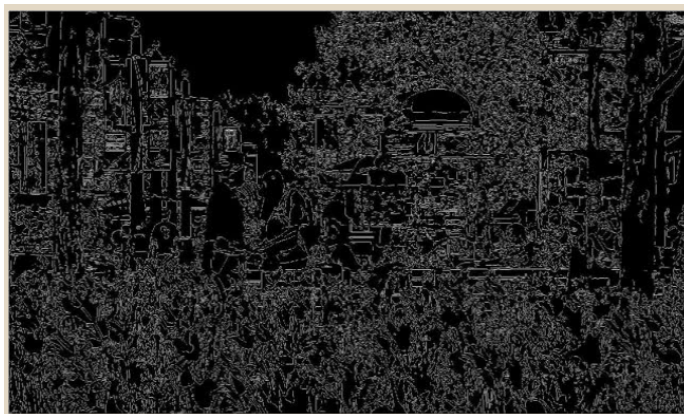
Pre potreby identifikácie scény bola vybraná priestorová (SI) a časová informácia (TI). Každá videosekvencia má tieto parametre rozdielne pre každé kódovanie, či už z pohľadu kódovania v rámci kompresného štandardu, ale aj bitového toku.

### 5.3.1 Priestorová spektrálna informácia (SI)

Vo všeobecnosti je to meranie množstva priestorových detailov snímky. Pri viac priestorovo rozsiahlych scénach je zvyčajne vyššia. Nie je to výsledok entropie, ani to nie je spojené s informáciou definovanou v teórii komunikácie. SI je založený na Sobelovom filtri (tiež nazývaný Sobelov operátor). Je to filter, ktorý používa algoritmus detekcie hrán a vytvára obrázky s dôrazom na oblasti s vysokou priestorovou frekvenciou, ktorá zodpovedá okrajom. Najskôr je každý rámeček videa filtrovaný Sobelovým filtrom podľa špecifikovaného času. Potom je vypočítaná štandardná odchýlka pixelov pre každý Sobelov filtrovaný rámeček. Spomínaná operácia sa opakuje pre každý rámeček vo video sekvencii a ako výsledok je časová rada priestorových informácií scény. V časovej sérii je vybratá najvyššia hodnota reprezentujúca SI obsah v scéne [16, 17].

$$SI = \max_{time} \{std_{space}[Sobel(f_n)]\} \quad (14)$$

kde  $\max_{time}$  je maximálna hodnota v časovej sérii,  $std_{space}$  je štandardná odchýlka nad pixelmi,  $Sobel$  reprezentuje Sobelov filter a  $f_n$  označuje počet snímok v priebehu času  $n$ .



Obrázok 24 Priestorová informácia SI

### 5.3.2 Časová spektrálna informácia (TI)

Vo všeobecnosti je to meranie množstva časových zmien vo video sekvencií. Pri scéne s veľkou dynamikou pohybu je zvyčajne väčšia. Z viac technického hľadiska je TI založená na funkcii rozdielu v pohybe. Táto funkcia je rozdiel medzi hodnotami pixelu na rovnakom mieste v priestore, ale v po sebe idúcich rámcoch videa. Funkcia rozdielu v pohybe je funkciou času  $n$  a je definovaná nasledovne:

$$M_n(i, j) = F_n(i, j) - F_{n-1}(i, j) \quad (15)$$

kde  $F_n(i, j)$  je pixel v  $i$ -tom riadku a v  $j$ -tom stĺpci  $n$ -tého rámcu v čase  $n$ .

TI sa meria ako maximálny čas štandardnej odchýlky v priestore pohybu z rozdielu medzi hodnotami pixelov pre sekvenčné rámce nad všetkými  $i$  a  $j$ . Rovnica TI je vyjadrená ako:

$$TI = \max_{time}\{std_{space}[M_n(i, j)]\} \quad (16)$$

kde  $\max_{time}$  je maximálna hodnota v časovej sérii,  $std_{space}$  je štandardná odchýlka v priestore,  $M_n(i, j)$  je rozdiel medzi hodnotami pixelov pre sekvenčné rámce a  $(i, j)$  je umiestnenie pixelu. Rovnica 16 ukazuje, že čím vyšší je odhad v príslušných rámcoch, bude aj vyššia hodnota TI [16].



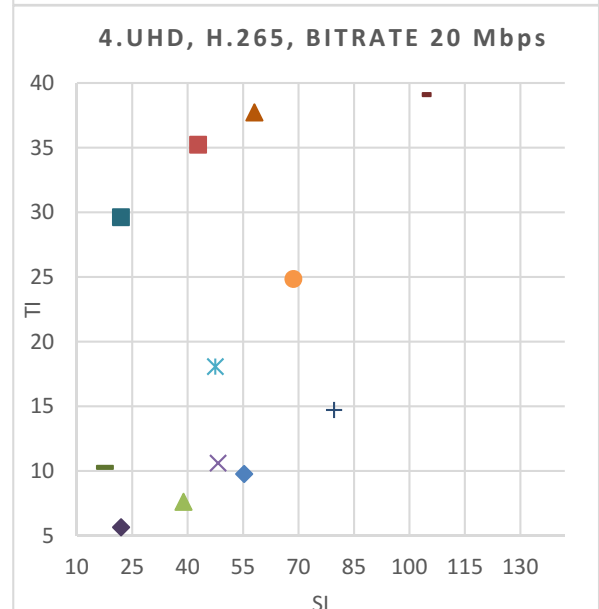
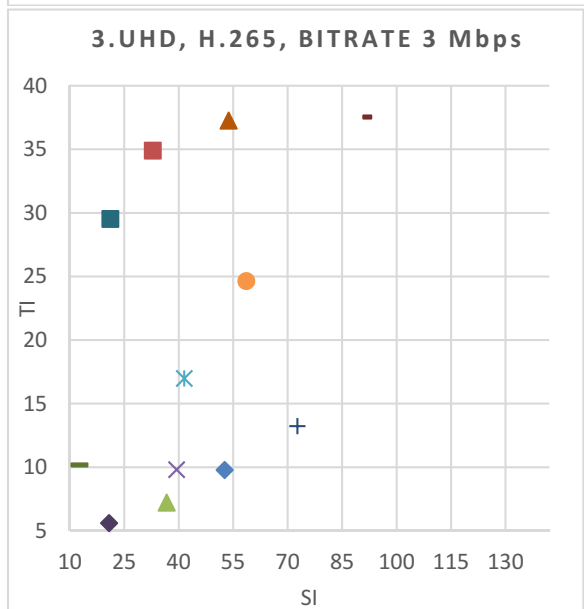
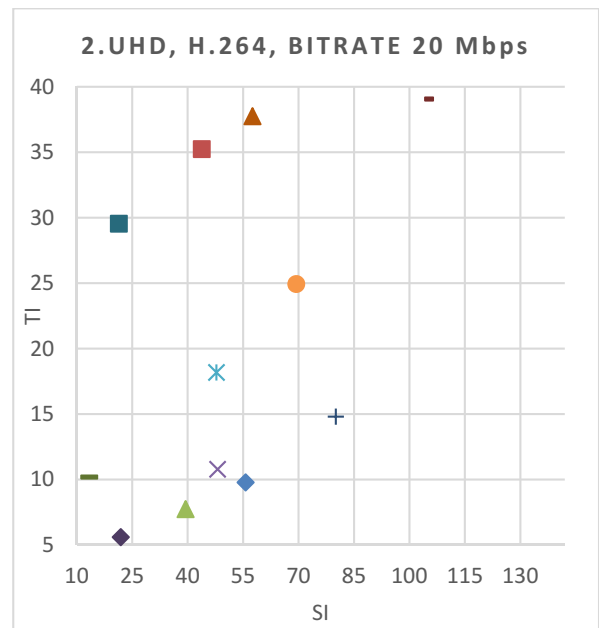
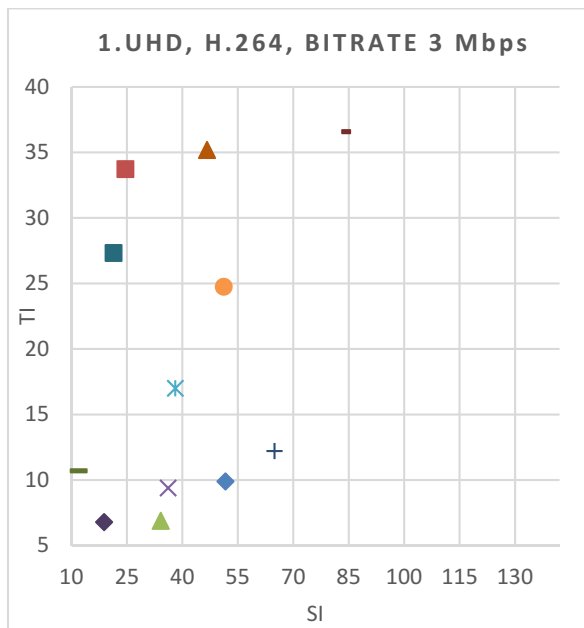
Obrázok 25 Časová informácia TI

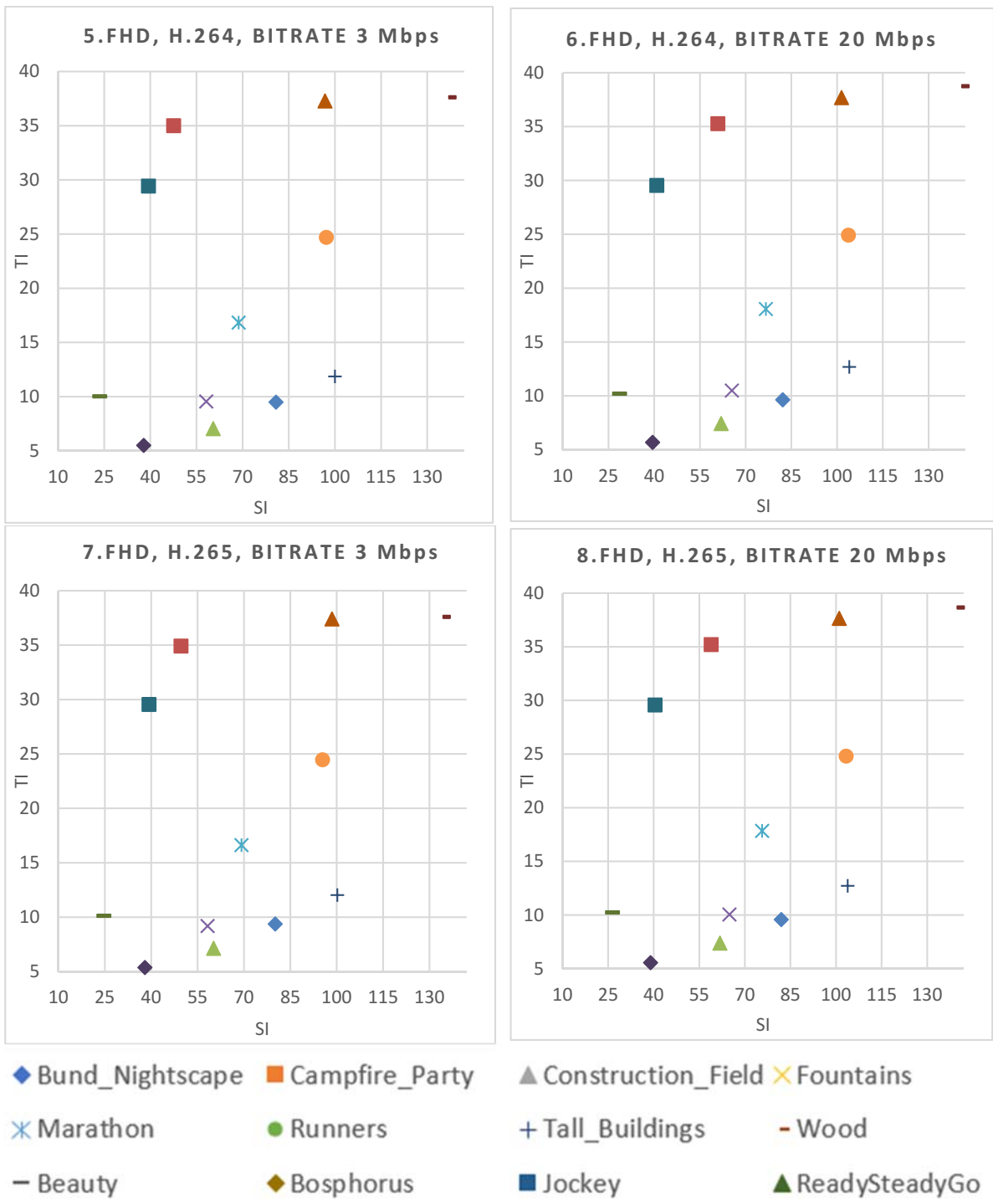
Pre testy som vyberal viaceré sekvencie s odlišnou časovou a priestorovou informáciou obsahu v nich. Pri vykonávaní subjektívnych testov som dodržiaval odporúčenia ITU [33, 52]. Pre subjektívne hodnotenie videí som použil MOS stupnicu.

Všeobecne sa priestorové a časové informácie používajú na reprezentáciu video obsahu. Tieto informácie sú kritickými a kľúčovými parametrami danej scény a určujú mieru kompresie videa. Analýzu klasifikácie obsahu som vykonal počítaním indexov SI a TI jasovej zložky pre každú videosekvenciu.

Cieľom vytvoreného klasifikátora je určiť hranice priestorovej a časovej informácie pre prepínanie bitového toku. Analýzu klasifikácie obsahu každej videosekvencie v databáze, na základe informácií SI a TI, som vykonal pomocou nástroja Mitsu tool [48]. Tieto informácie sa menia nielen v odlišnom type videosekvencie, ale aj v rovnakej scéne, pri jej prekódovaní, a sú ovplyvnené rozlíšením, kompresným štandardom a samozrejme aj bitovým tokom. V grafoch 1.1-1.8 sú zobrazené tieto

informácie pre vybrané typy prekódovaných videosekvencií, kde môžeme vidieť závislosť TI a SI pri týchto scénach. Do porovnania som zahrnul kombináciu oboch rozlíšení a oboch použitých kompresných štandardov v spojení s bitovým tokom 3 Mbps a 20 Mbps. Na týchto vzorkách môžeme potvrdiť zmenu hodnôt týchto informácií, čím sa potvrdzuje možnosť klasifikácie scény na základe týchto parametrov. Z grafov vidíme, že vytvorenú databázu tvoria rôznorodé videosekvencie a tak pokrývajú široké spektrum videosekvencií, čo umožňuje adaptáciu výsledkov v praxi. Kompletná databáza videosekvencií spolu s klasifikáciou scény sa nachádza v prílohe A.





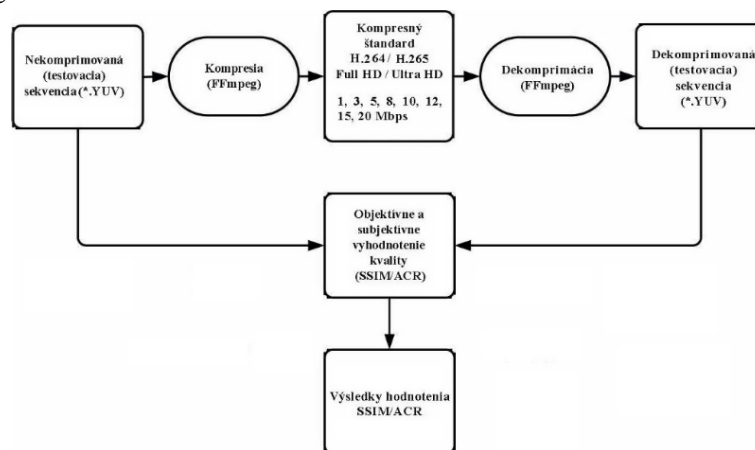
Graf 1 Závislosť SI a TI pri vybraných videosekvenciách



### 5.3.3 Porovnanie kvality referenčných a komprimovaných videosekvencií

Celý proces kódovania a následného vyhodnotenia je na obrázku 24. Na vyhodnocovanie kvality a určenie vplyvu bitového toku na výsledný obraz som použil subjektívne a tiež objektívne metódy. Zvyčajne sa preferujú objektívne metódy, ktoré okamžite môžu byť použité pre porovnanie a verifikovanie. Každé použitie týchto metód má rozličnú procedúru a používa rozličnú metriku na vyhodnotenie. Parametre ako video rozlíšenie, typ obsahu, parametre kódovania, snímková frekvencia, stratovosť a typ paketov nemôže byť dostatočne pokrytá subjektívnym hodnotím video databáz. Ako objektívna metóda je použitá SSIM, nakoľko lepšie koreluje s vnímaním ľudského oka [49, 50]. Túto metriku som vypočítal s použitím nástroja MSU Video Quality Measurement Tool [51], ktorý vie porovnať referenčnú vzorku s prekódovanou a vyhodnotiť hodnoty pre objektívne metódy. Vzor príkazu z tohto nástroja uvádzam v prílohe B.1.

Subjektívne hodnotiace metódy sú veľmi náročné, z hľadiska času a využitia ľudí. Je použitá nereferenčná metóda ACR, pretože v tomto prípade je video porovnávané iba na základe videnej videosekvencie a nie porovnaním s referenciou. V reálnom prostredí pri prijímaní signálu od poskytovateľa služieb, konečný užívateľ tiež dostáva iba prijímaný signál a neporovnáva ho s referenčným originálom.



Obrázok 26 Proces kódovania a vyhodnotenia video kvality

## 5.4 Štatistické metódy vyhodnotenia

Akákoľvek analýza údajov je rozdelená do špecifických krokov: “vyčistenie dát“, správne štatistické vyhodnotenie a interpretácia výsledkov. Je potrebné detekovať irelevantné dáta, kde skúmame, ktoré dáta nezodpovedajú intervalu spoľahlivosti s ostatnými dátami. Tieto dáta neboli zaradené do finálnej databázy ohodnotení. Najpoužívanejším spôsobom na vyhodnotenie relevantnosti dát je Pearsonov korelačný koeficient. Je založený na predpoklade, že stupnica merania je pomer a predpokladá, že dáta majú normálne rozdelenie. Ak má hodnotiacia stupnica malý rozsah, je možné Pearsonov korelačný koeficient zmeniť za Spearmanov korelačný koeficient, ktorý je založený na slabšom predpoklade, že rozsah merania je radový. Pre stupnicu s malým rozsahom je ťažko predpokladať, že rozdelenie je blízko normálnemu (špeciálne, keď väčšina odpovedí je blízko určitej hranice stupnice).

Hodnotenia sa zúčastnilo 35 respondentov, pričom po štatistickom vyhodnotení neboli brané v úvahu hodnotenia od 5 respondentov, ktoré sa vyčyľovali od ostatných hodnotení. Finálna databáza obsahuje údaje od 30 hodnotiteľov, z ktorých bolo 15 mužov vo veku 21-44 rokov a 15 žien vo veku 19-34 rokov. Priemerný vek hodnotiteľov je 25 rokov.

Tabuľka 2 Počet hodnotiteľov

	Počet mužov	počet žien
<b>FHD + UHD</b>	15	15
<b>Vek</b>	21-44	19-34
<b>Priemerný vek</b>	25	

### Výberový priemer

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (17)$$

kde  $x_i$  sú konkrétne hodnoty a  $n$  je celkový počet hodnôt

### Výberová smerodajná odchýlka

Je definovaná ako druhá odmocnina z výberového rozptylu. Výberovú smerodajnú odchýlku počítame z realizovaného výberu, teda v prípade že nemáme k dispozícii úplnú množinu možných stavov, ale len výber z nich. Vo vzorci musíme uvažovať  $n-1$  (pretože 1 meranie je už závislé s výpočtom strednej hodnoty).

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (18)$$

### Variačný koeficient

Je podiel smerodajnej odchýlky a aritmetického priemeru vyjadrený v percentách. Používa sa na charakterizáciu variability znaku a je definovaný iba pre znaky s nezápornými hodnotami.

$$v = \frac{s}{\bar{x}} * 100 \quad (19)$$

Aritmetický priemer nemôže reprezentovať heterogénny štatistický súbor. To nastane v prípade, že je variačný koeficient väčší ako 50%..

## Pearsonov korelačný koeficient

Je mierou lineárnej závislosti dvoch premenných. Čitateľ sa nazýva kovariancia a vyjadruje ako sa súčasne menia hodnoty dvoch premenných. Nula znamená, že sa menia nezávisle. Nadobúda hodnoty -1 až 1. Korelácia pod 0,1 sa často označuje ako triviálna, 0,1 – 0,3 malá, 0,3 – 0,5 stredná a nad 0,5 veľká. 0,7 – 0,9 je veľmi vysoká a 0,9 – 1 takmer dokonalá korelácia. Pearsonov korelačný koeficient (ako aj koeficient determinácie) sú ovplyvnené extrémnymi hodnotami.

$$r_{x,y} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (20)$$

kde  $\bar{x}$  predstavuje aritmetický priemer hodnôt MOS ( $x_i$ ) získaných od  $i$  užívateľov v hodnotiacom procese a  $\bar{y}$  predstavuje aritmetický priemer hodnôt MOS ( $y_i$ ) predikovaných nástrojom.

## Spearmanov korelačný koeficient

Pohybuje sa v rozmedzí  $<-1; -1>$  a vyhodnocuje pevnosť spojenia medzi premennými. Na rozdiel od Pearsonovho korelačného koeficientu, Spearmanove koeficienty sú vypočítané z radov dátových hodnôt ako zo samostatných hodnôt, čo spôsobí, že sú menej citlivé na odchýlky.

## Mean Square of Error (MSE)

Výberový reziduálny rozptyl, označovaný ako priemerná kvadratická chyba. Predstavuje nevychýlený bodový odhad rozptylu  $\sigma^2$  náhodných chýb. MSE je metrika používaná na posúdenie toho, ako dobre funguje metóda rekonštrukcie obrazu vzhľadom na pôvodný obrázok. Hovorí ako blízko je množina bodov k regresnej priamke. Spracuje vzdialenosti dát od regresnej krivky (tieto vzdialenosti sú chyby) a umocní ich, aby sa odstránili negatívne znaky.

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_i)^2 \quad (21)$$

kde  $\bar{x}$  je vektor predikcií a  $x$  je vektor nameraných hodnôt. Hodnoty sú v intervale  $<0,65025>$ ; 0 - pre identické rámce.

## Root Mean Square Error (RMSE)

Štandardná odchýlka chyby predikcií. Chyba predpovedí je vyjadrenie ako ďaleko sú jednotlivé namerané hodnoty od regresnej krivky, RMSE je teda vyjadrenie rozptýlenia nameraných dát okolo najlepšej, referenčnej, hodnoty. Inými slovami, informuje o tom, ako sú údaje sústredené okolo najlepšej línie. Táto chyba sa pre parameter  $\theta$  a metódu práce s hodnotami  $q$  vypočíta podľa vzorca:

$$RMSE(\theta, q) = \sqrt{\frac{1}{n-1} \sum_{i=1}^N (\theta - \theta_{iq})^2} \quad (22)$$

kde  $\theta_{iq}$  predstavuje odhad parametru  $\theta$  metódou  $q$  vo výbere  $i$ . V každom výbere sa teda najskôr spočíta rozdiel medzi hodnotou populačného parametra  $\theta$  a odhad parametra v tomto výbere  $\theta_{iq}$ . Po umocnení týchto rozdielov sa z nich vypočíta priemerná hodnota.

## 5.5 Analýza vplyvu bitového toku na kvalitu videosekvencií

Táto kapitola vyhodnocuje vplyv bitového toku na kvalitu videosekvencií vytvorenej databázy. Objasňuje postup subjektívneho testovania, jeho metodiku a tiež spôsob, akým boli dosiahnuté výsledky pomocou objektívnej metriky.

### Subjektívne testovanie

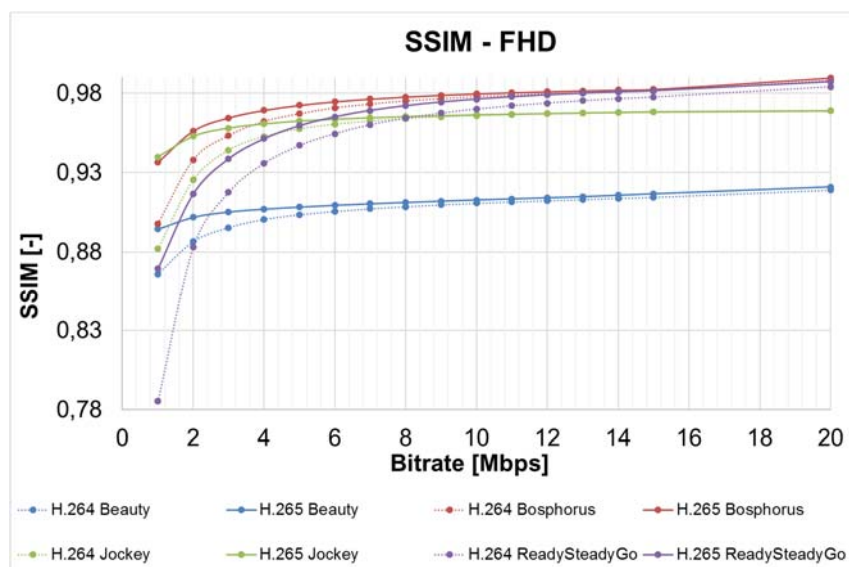
Postup a podmienky subjektívnych testov sú opísané v odporúčaní [33, 52]. Tieto odporúčania definujú požiadavky na zdrojové signály (referenčné sekvencie), hodnotiteľov kvality, počiatočné pokyny predložené pred začatím testovania, kroky testovania videosekvencie, spôsob zdokumentovania výsledkov, analýzu výsledkov a faktory, ktoré môžu mať negatívny vplyv na subjektívny test. Hodnotiacia skupina musí pozostávať z minimálne 15 hodnotiteľov. Pokyny pre hodnotiteľov musia byť jasné dané pred testovaním. Všetci hodnotitelia musia dostať rovnaké pokyny. Toto eliminuje jeden potenciálny zdroj predpojatosti subjektu. Inštrukcie by mali popisovať hodnotiaci cyklus, stupnicu kvality, ako zaznamenať hodnotenia, správanie nášho systému a prehrávania. Inštrukcie by mali obsahovať upozornenie, aby hodnotitelia nehodnotili kvalitu danej scény na základe sympatií a názoru na danú scénu, prípadne ľudí, ktorí v danej scéne vystupujú. Po inštrukciách nasleduje tréningová sekcia videosekvencií, ktorá slúži k demonštrácii úlohy, oboznámením sa s rozsahom kvality a typom zhoršenia v experimente. Videosekvencie zaradené do tréningovej sekcie sa nesmú vyskytovať počas riadneho testovania.

Subjektívne testovanie prebiehalo v prispôsobenej miestnosti, ktorá spĺňala odporúčania ITU-T P.910 a ITU-T BT.500-13, aby hodnotitelia mali vhodné svetelné podmienky, aby dodržiavali správnu vzdialenosť od monitora. Hodnotenie prebiehalo na monitoroch 24" Dell P2415Q UHD, ktoré spĺňajú odporúčania [33, 52].

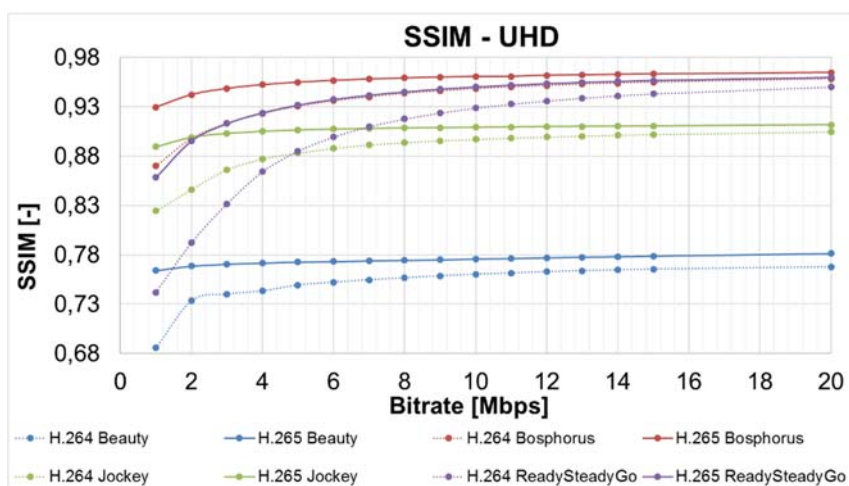
Ako subjektívna metóda bola vybraná ACR, nakoľko hodnotenie prebieha na základe videnej videosekvencie bez možnosti porovnania s referenčnou videosekvenciou, tak ako tomu je pri reálnom sledovaní televízie. ACR je subjektívna metóda založená na vyhodnotení kvality respondentmi. Kvalita je definovaná 5-stupňovou stupnicou MOS, jednu videosekvenciu je možné ohodnotiť len raz a to jednou konkrétnou hodnotou. Prínosom tejto subjektívnej metódy je rýchlosť realizácie, nezaberá toľko času v porovnaní s napr. DSIS alebo DSCQS.

### Objektívne testovanie

Na vyhodnotenie kvality jednotlivých videosekvencií som použil nástroj MSU Video Quality Measurement Tool [51], ktorý porovnáva referenčné a testovacie sekvencie na základe parametrov: názov referenčnej videosekvencie, rozlíšenie, názov porovnáwanej sekvencie, jej rozlíšenie a metriky, ktoré chceme vyhodnotiť. Rozlíšenie je potrebné uvádzať nakoľko nekomprimované video v sebe nenesie túto informáciu. Výsledky objektívneho vyhodnotenia pre Ultra Video Group sekvencie zobrazuje graf 2, pre rozlíšenie FHD, resp. graf 3 pre rozlíšenie UHD. Výsledky pre videosekvencie SJTU Media Lab sa nachádzajú v kapitole 7, v grafoch 4.1-4.16, kde sú vzájomne porovnané vyhodnotenú dáta objektívnou a subjektívnou metriku. Komplet celá databáza, ktorú som vytvoril z pôvodných videosekvencií skupiny SJTU Media Lab [44] sa nachádza v prílohe A, v tabuľke 17. Ďalšia časť vytvorenej databázy (Ultra Video Group [46]) je obsiahnutá v prílohe A v tabuľke 18.



Graf 2 SSIM vyhodnotenie Ultra Video Group sekvencií – rozlíšenie FHD



Graf 3 SSIM vyhodnotenie Ultra Video Group sekvencií – rozlíšenie UHD



## 6 Vytvorenie modelu klasifikátora scény pomocou neurónovej siete

Navrhnutý model je založený na neurónovej sieti vytvorenej v prostredí Matlab. Termín „neurónová sieť“ je odvodený od základnej funkčnej jednotky ľudského nervového systému alebo nervových buniek, ktoré sú prítomné v mozgu a iných častiach ľudského tela. Neurónová sieť sa trénuje postupným nastavením svojich váh a biasov, tento postup sa vykonáva iteračne pre získanie požadovaného výstupu. Tieto parametre sa nazývajú aj voľné parametre. Aby sa neurónová sieť naučila je potrebné v prvej fáze trénovanie, vykonávané pomocou definovaného súboru pravidiel známych aj ako algoritmus učenia. Trénovací proces všeobecne pozostáva zo 4 krokov:

1. Vytvorenie trénovacích dát
2. Vytvorenie objektu siete
3. Trénovanie siete
4. Simulácia siete s testovacími dátami

Neural network toolbox (NN Toolbox) je knižnica implementovaná v prostredí Matlab umožňujúca prácu s umelými neurónovými sieťami. Umožňuje zvolenie NN rôznych štruktúr (viacvrstvové, rekurentné, samoorganizujúce sa siete) a vykonávať algoritmy prislúchajúce k danej štruktúre. Dopredné neurónové siete (Feedforward networks) majú jednu alebo viac skrytých vrstiev sigmoidných neurónov, po ktorých nasleduje výstupná vrstva lineárnych neurónov. Viaceré vrstvy neurónov s nelineárnymi prenosovými funkciami umožňujú sieti naučiť sa nelineárne aj lineárne vzťahy medzi vstupnými a výstupnými vektormi. Lineárna výstupná vrstva umožňuje sieti vytvárať hodnoty mimo rozsahu (-1, +1). Prvým krokom pri tréovaní doprednej siete je vytvorenie sieťového objektu funkciou *newff*. Vyžaduje štyri vstupy a výstupom je sieťový objekt. Prvým vstupom je dvojrozmerná matica normovaných hodnôt pre každý prvok  $R$  vstupného vektora. Druhým vstupom je pole obsahujúce veľkosť každej vrstvy. Tretím vstupom je bunkové pole obsahujúce mená prenosových funkcií použitých v každej vrstve. Konečný vstup obsahuje názov trénovacej funkcie, ktorá sa má použiť.

Inicializácia váh – pred samotným tréovaním siete je potrebné inicializovať sieť príkazom *init*. Táto funkcia berie ako vstup sieťový objekt a vráti ho s inicializovanými váhami a biasmi. Samotná funkcia *newff* inicializuje sieť, ale niekedy je potreba reinicializovať sieť.

Model využíva algoritmus Back-propagation, ktorý môže byť v Matlab-e implementovaný viacerými druhmi. V práci som otestoval 3 z nich, líšia sa nielen zložitou a stabilitou, ale hlavne rýchlosťou konvergencie, pričom všetky tri využívajú metódu gradient descent na úpravu vnútorných parametrov siete.

- Batch Gradient Descent (*traingd*)
- Batch Gradient Descent with Momentum (*traingdm*) - tento algoritmus v sebe zahŕňa takzvané *momentum*. Momentum sa vytvára násobením predošlej hodnoty váhy konštantou v rozsahu 0–1 a súčtom tejto novej hodnoty a hodnoty zmeny váhy navrhovanej metódou spätnej propagácie. Ak je konštanta momenta rovná 0, zmena hodnoty váhy sa rovná hodnote navrhovanej metódou spätnej propagácie, naopak ak je rovná 1, zmena je rovná predošlej hodnote váhy. Tento prístup umožňuje elimináciu uviaznutí algoritmu v lokálnych plytkých minimách, ako by sa to mohlo stať bez použitia momenta.

- Gradient Descent with Variable Learning Rate (traingdx) - Výkon algoritmu s nepremennivou mierou učenia je veľmi závislý na zvolenej hodnote. Ak je miera učenia príliš veľká, algoritmus môže oscilovať a stane sa nestabilným. Ak je však príliš malá tréningovanie trvá dlho. Zlepšenie algoritmu teda dosiahneme ak sa miera učenia bude meniť dynamicky podľa jeho stavu. Princíp spočíva v porovnávaní aktuálnej a predošlej hodnoty chyby siete počas cyklov tréningovania. Ak aktuálna hodnota chyby prevyšuje predošlú o istú hodnotu, nové váhy a biasy sú zahodené a miera učenia sa zmenší (typicky násobením s  $lr\_dec=0.7$ ). Naopak ak je aktuálna chyba menšia ako predošlá, váhy sa zachovávajú a miera učenia sa zväčší (typicky násobením  $lr\_inc=1.05$ ). To znamená že tam, kde by veľká miera učenia mohla spôsobiť oscilácie, dochádza k jej zníženiu. Zvyšovanie miery učenia v oblastiach kde sa nepoškodí stabilita algoritmu zas umožňuje jeho zrýchlenie [42].

## 6.1 Tréningovanie siete

Základom tréningovania neurónových sietí je vzorová množina. Pred samotným tréningovaním musíme zadať definovať takzvanú tréningovaciu sadu vstupných a výstupných vektorov. Je to vlastne vzorka na ktorú sa bude sieť učiť správne odpovedať. Tieto premenné sa zadávajú v maticovom tvare. Túto množinu ako aj testovaciu množinu som vytvoril z databázy videosekvencií a obsahovala klasifikáciu videosekvencie v podobe SI a Ti informácie, kvalitu danej videosekvencie pri zadanom rozlíšení, použitom kodeku a bitovom toku. Pomocou neurónovej siete som následne hľadal vhodnú topológiu na základe:

- Algoritmu učenia
- Prenosovej (aktivačnej) funkcie
- Tréningovej funkcie
- Počtu skrytých vrstiev a počtu neurónov v daných vrstvách
- Času potrebného na dosiahnutie vytréningovanej siete

### Spracovanie dát

Tréningovanie neurónovej siete môže byť ešte efektívnejšie ak je vykonaná úprava vstupov a výstupov škálovaním v rovnakom intervale. V práci som otestoval škálovanie na intervaloch  $\langle -1;1 \rangle$  a  $\langle 0;1 \rangle$  (tiež som otestoval tréningovanie bez škálovania vstupov a výstupov). Tréningovanie siete som vykonával zvlášť pre skupinu kodeku H.264 a H.265, tak ako aj pre oba kodeky naraz. Takto je možné porovnať natréningovanie siete pre 2 fundamentálne skupiny vstupov. Pre presnejšie výsledky som neurónovú sieť, pre každú topológiu, trénoval 10 krát, pričom vždy boli náhodne generované váhy a biasy. Toto obmedzilo vplyv náhodne generovaných inicializačných váh, vďaka čomu sa zvýšila štatistická významnosť celého procesu. Pri tréningovaní sa používala validačná množina dát, ktorá overovala tréningovanie siete, aby nedochádzalo k jej pretréningovaniu. Výsledná databáza bola náhodne rozdelená na tréningovaciu a testovaciu sadu v pomere 75:25 v percentuálnom zobrazení. Validačná množina bola náhodne generovaná z tréningovej množiny dát v pomere 70:30 (tréningovacia sada:validačná sada). Pre štatistickú významnosť bol tento proces opakovaný opäť 10 krát, aby sa sieť natrénovala pri rozličných dátach tréningovej a validačnej množiny. Tréningovanie prebiehalo pre rôzne vytvorené topológie, ktoré obsahujú 2-4 vrstvy skrytých neurónov, kde v jednej vrstve môže byť až 160 neurónov Pri vyššom počte



vrstiev dochádzalo k pretrénovaniu siete, vyšší počet neurónov v jednej vrstve nebol potrebný nakoľko sa sieť dokázala správne natrénovať už pri nižšom počte neurónov, čo jasne dokazujú výsledky simulácie. Pomer rozdelenia dát, počet vrstiev, počet neurónov vyplýva z testov tréningu a simulácií, čo zobrazujú výsledky finálne naprogramovaného algoritmu ako aj výsledky dosiahnuté pri hľadaní vhodných funkcií, výber správneho rozloženia dát. Zvolený počet opakovaní vychádza aj z výsledkov autorov prác [53, 54, 55].

Aby sa predišlo pretrénovaniu siete, využíval som validačnú množinu, kde ak 6 krát po sebe neklesala chybovosť validačnej množiny, tréning bolo ukončené v momente, kde sieť bola natrénovaná najlepšie, výsledkom sú váhy a prahy dosiahnuté pri minimálnej chybe validačnej sady. Čiže ak chybovosť tréningových dát klesala, ale chybovosť validačnej množiny dát stúpala alebo ostávala na rovnakej úrovni, tréning sa zastavilo. Pri nastavení vyššieho počtu výskytu neklesajúcej chybovosti na validačnej sade neboli dosiahnuté lepšie výsledky. Rovnako tak bola overená aj možnosť modifikácie funkcie na výpočet chybovosti popísanej v práci. Každú simuláciu som vykonal pre obe rozlíšenia zvlášť a tiež spoločne v jednej simulácii, rovnako tak aj pre oba kompresné štandardy.

Po natrénovaní neurónovej siete je potrebné otestovať jej natréningovanie na testovacích dátach, ktoré neurónová sieť doposiaľ nemala k dispozícii. Po simulácii dát je vypočítaná korelácia na základe Pearsonovho korelačného koeficientu. Ak by chybovosť pri testovacej sade mala výrazne nižšiu hodnotu ako chybovosť pri validačnej sade, znamenalo by to zlé rozdelenie databázy videosekcencií.

## 6.2 Vytvorený klasifikátor

Klasifikátor založený na neurónovej sieti predikuje vhodný bitový tok na základe scény klasifikovanej kritickými parametrami SI a TI a kvalitatívnymi parametrami danej videosekvencie.

Model umožňuje predikciu:

- Bitového toku na základe scény a hodnotenia kvality danej scény objektívnou metrikou SSIM
- Bitového toku na základe scény a hodnotenia kvality danej scény pomocou stupnice MOS (hodnota dosiahnutá subjektívnou metrikou ACR)

Model okrem predikcie bitového toku taktiež umožňuje predikovať subjektívnu kvalitu videosekvencie na základe kvalitatívnych parametrov hodnotenej videosekvencie a kvality vyjadrenej pomocou subjektívnej metriky SSIM. V praxi je veľmi dôležité uspokojenie koncových zákazníkov a teda je nevyhnutné vyhodnocovať vnímanú kvalitu prezeranej videosekvencie užívateľmi. Samotná objektívna metrika nám nepodá informáciu o vnímanej kvalite prezeraného obsahu používateľom. Vytvorený model teda približuje vnímanie kvality videosekvencie koncovým užívateľom a predikuje:

- Subjektívnu kvalitu na základe scény (charakterizovanou SI a TI informáciami), bitového toku, kvalitatívnych parametrov a objektívneho výsledku metrikou SSIM

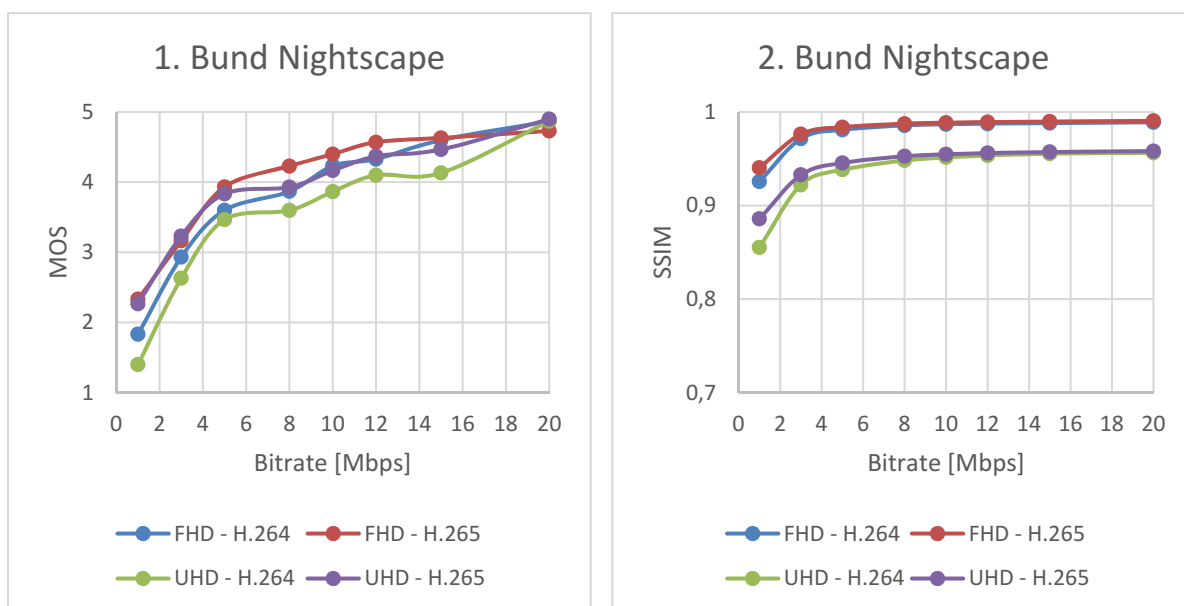
Vytvorený model okrem predikcie bitového toku a vyjadrenia kvality pomocou subjektívnej metriky na základe kvality ohodnotenej objektívnou metrikou, dokáže kvalitatívne vyhodnotiť videosekvencie a teda predikovať:

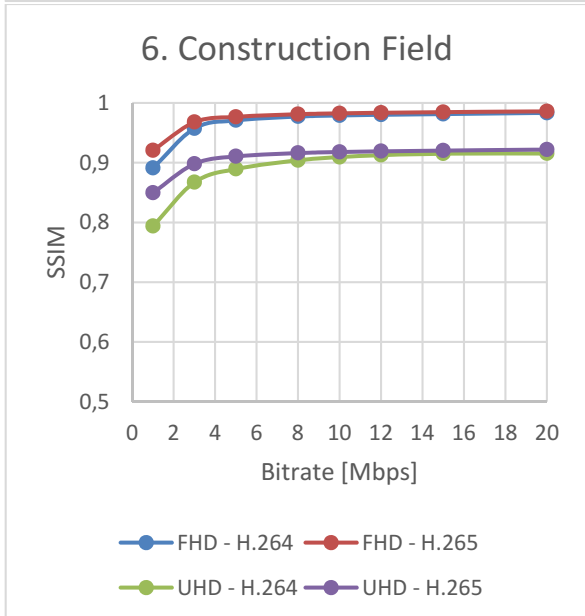
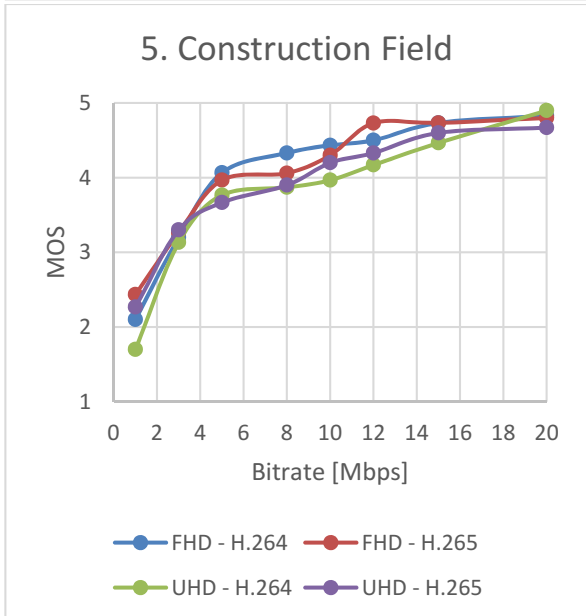
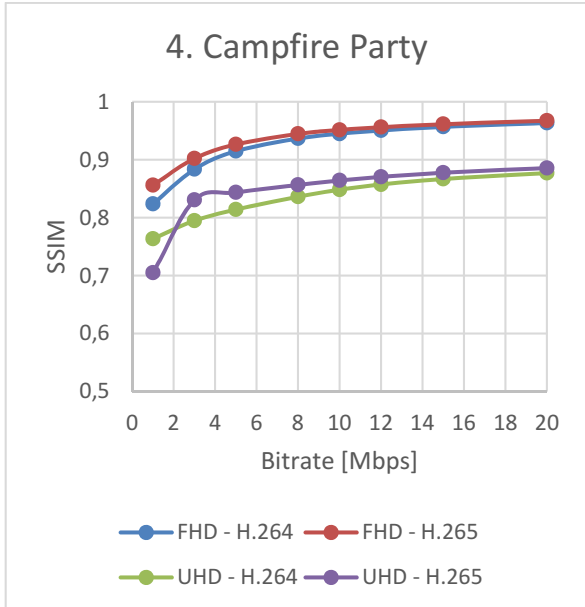
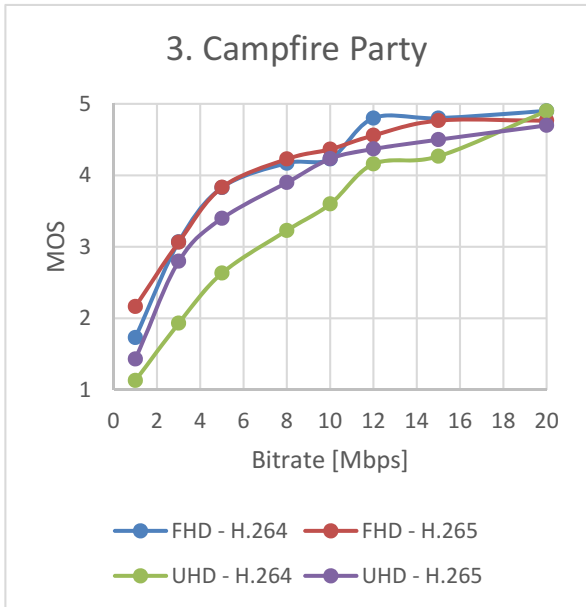
- Subjektívnu kvalitu videosekvencie na základe informácií o videosekvenciách (SI a TI) a bitového toku
- Objektívnu kvalitu na základe informácií o videosekvenciách a bitového toku

## 7 Vyjadrenie korelácie výsledkov v hodnotení kvality videa medzi objektívnymi a subjektívnymi metódami

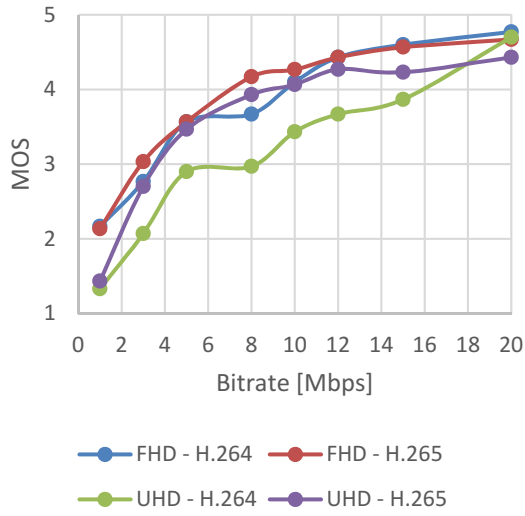
Každú videosekvenciu z vytvorenej databázy som klasifikoval pomocou SI a TI informácií a ohodnotil pomocou objektívnych testov. Vytvorená databáza obsahuje 384 scén. Tie vznikli z 12 typov videosekvencií prekódovaním do dvoch kompresných štandardov, ôsmich bitových tokov a dvoch rozlíšení. 265 videosekvencií (sekvencie SJTU Media Lab) bolo ohodnotených subjektívnou metrikou (postup som popísal v kapitole 5). V tejto kapitole sú vyhodnotené výsledky kvality videosekvencií z vytvorenej databázy, je popísaná vzájomná korelácia medzi subjektívnym a objektívnym ohodnotením kvality a je popísaný model predikcie subjektívnej kvality na základe objektívnej.

Výsledky dosiahnuté subjektívnou metrikou ACR, v podobe priemerného MOS skóre, pre jednotlivé videosekvencie, môžeme vidieť v grafoch 4.1-4.16 označené nepárnyimi číslami. Tieto výsledky môžeme porovnať s hodnotami dosiahnutými vyhodnotením video kvality s použitím objektívnej metriky SSIM, ktoré sú v grafoch 4.1-4.16 pod párnyimi číslami. Z grafov môžeme vidieť, že hodnotenia pomocou subjektívnej metriky sa veľmi približujú a podobajú hodnoteniam objektívnej metriky. Jednotlivé grafy sú k dispozícii na médiu, priloženom k tejto práci, aj s chybovými úsečkami reprezentujúcimi rozpätie chýb. Najlepšie skóre dosahuje štandard H.265 pri rozlíšení FHD, naopak najhoršie H.264 pri rozlíšení UHD.

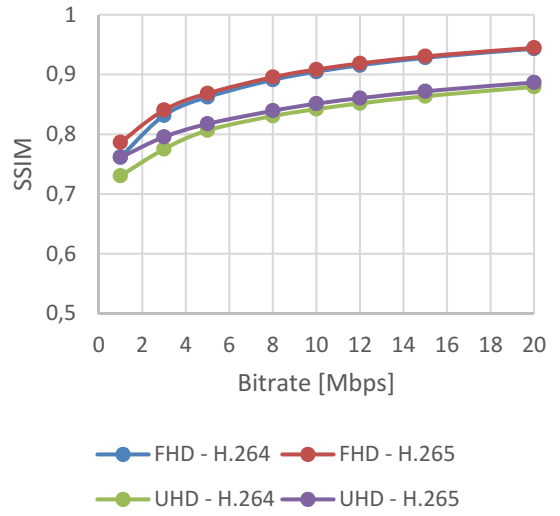




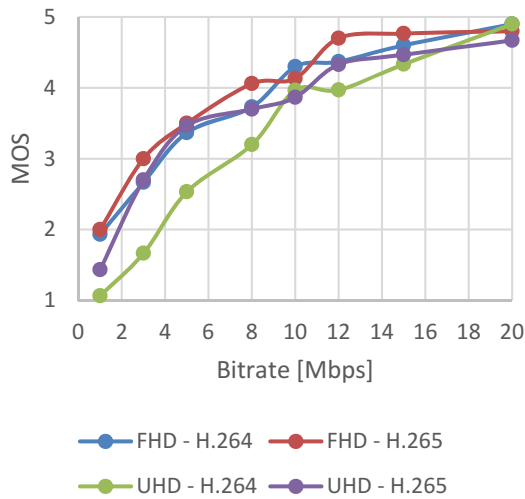
### 7. Fountains



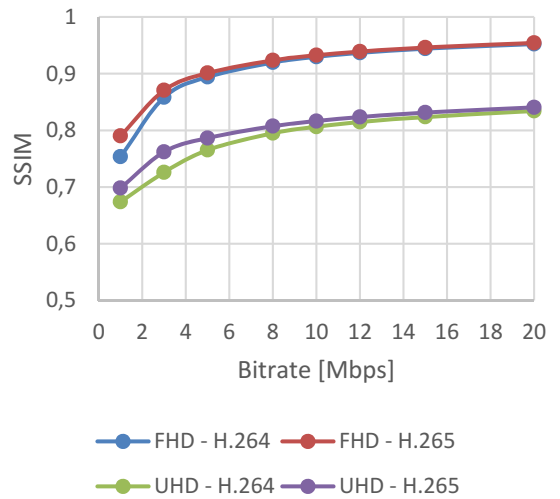
### 8. Fountains

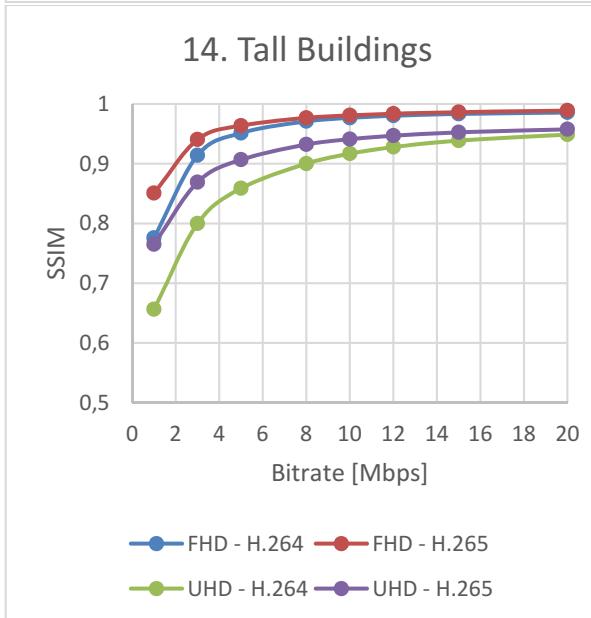
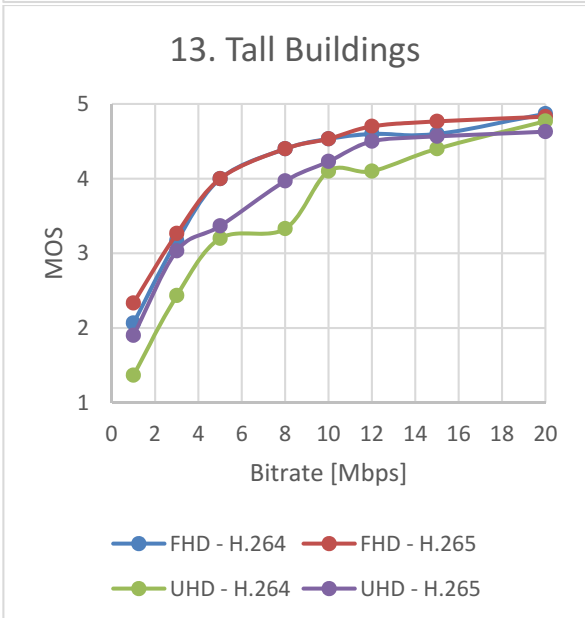
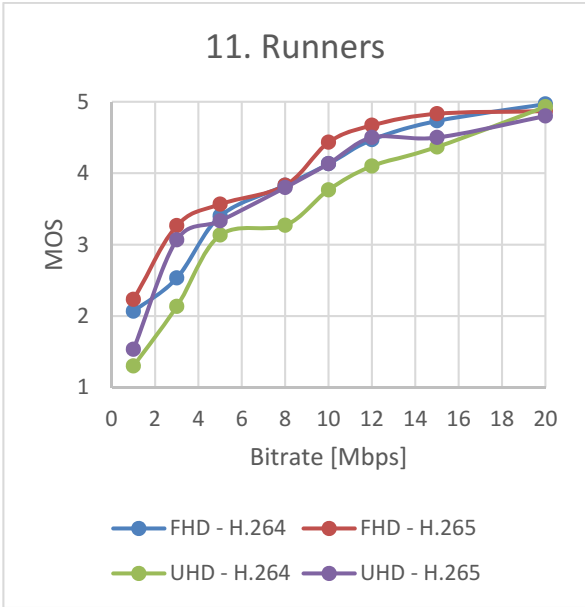


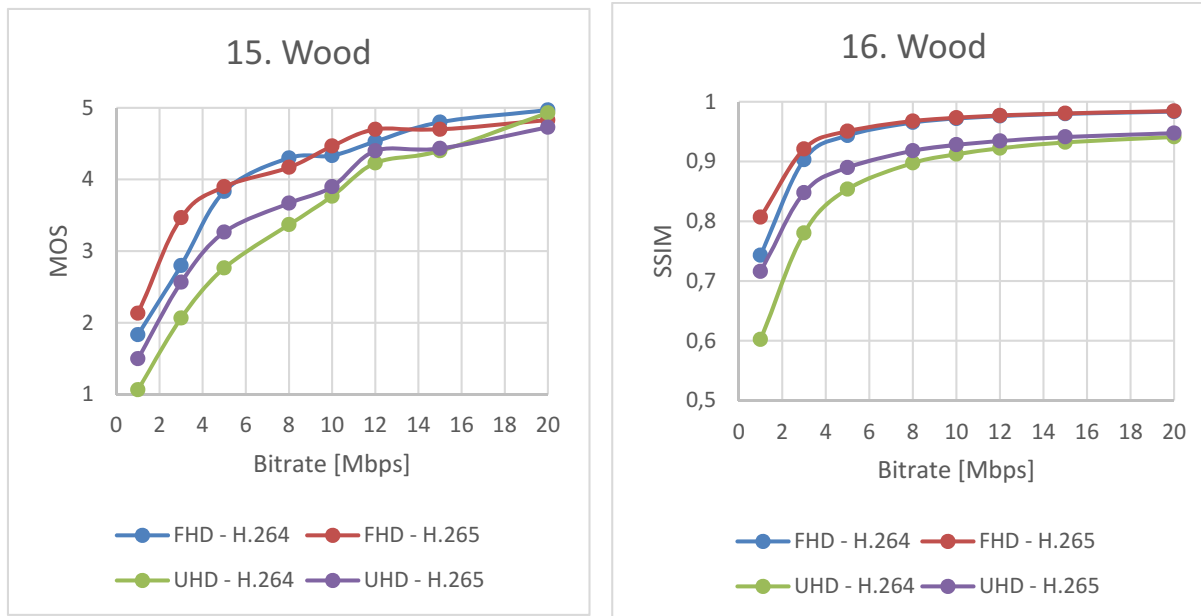
### 9. Marathon



### 10. Marathon





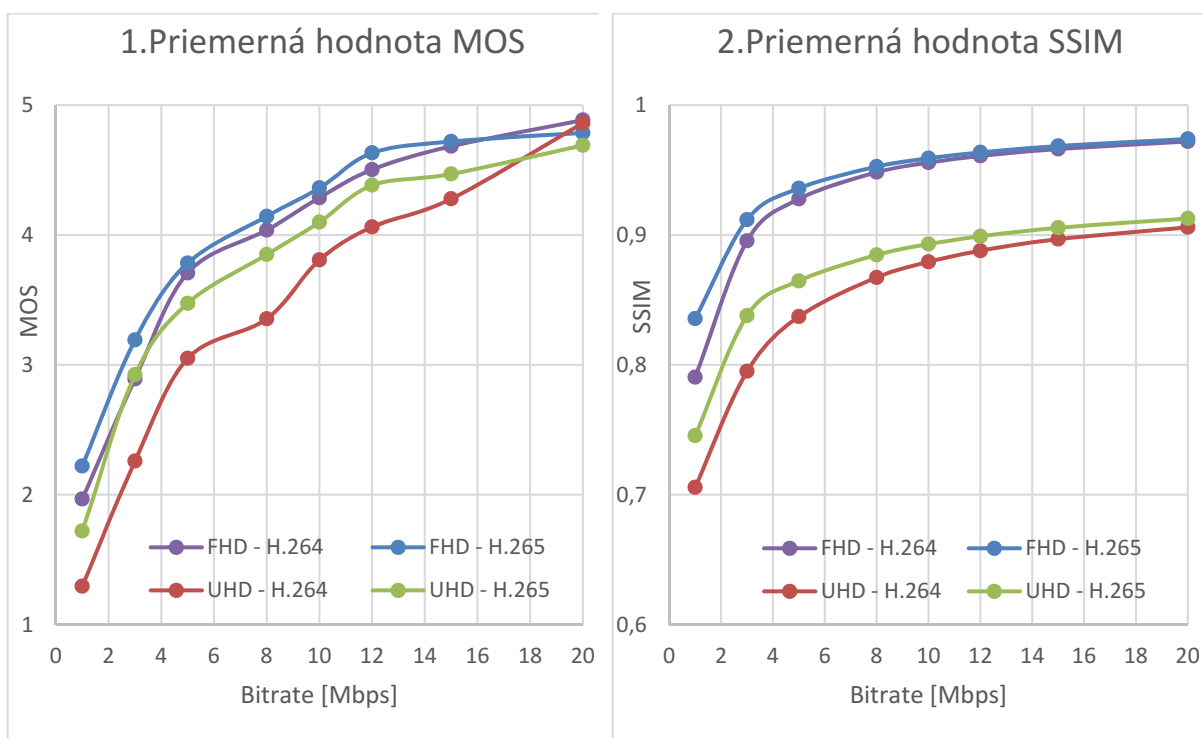


Graf 4 Dosiachnuté výsledky objektívnou a subjektívnou metrikou

Najlepšie výsledky pri FHD dosahujú scény s miernym pohybom (s malou dynamickosťou – scény s nízkou hodnotou TI) ako Construction field a Tall Buildings. Výsledná krivka sekvencie Wood, ktorá má najvyššie hodnoty SI a TI sa nachádza v okolí polovice všetkých kriviek kvality. Táto skutočnosť poukazuje na to, že ľudský mozog nevníma pohyb kamery rovnako ako pohyb objektu na statickom pozadí. Rovnako je z výsledkov zrejmé, že pri vyššom bitovom toku sú menšie rozdiely hodnotenia v škále MOS. Rovnako tak aj pri rozlíšení UHD, dosahujú najlepšie výsledky málo dynamické scény. Pri tomto rozlíšení sa jedná o scény Construction Field a Bund Nightspace.

## Priemerne výsledky vyhodnotenia kvality

V grafe 5.1 som vyhodnotil priemerné výsledky MOS skóre, graf 5.2 ukazuje priemerné výsledky SSIM skóre pre obe rozlíšenia aj kodeky. Z výsledkov je zrejmé, že subjektívne testy vykazujú veľkú podobnosť s testami objektívnymi a že kvalita všetkých analyzovaných sekvencií stúpa logaritmicky. Z grafov 5.1 a 5.2 vidíme, že používatelia hodnotili kvalitu videosekvencií veľmi podobne ako vyhodnotila metrika SSIM, čo opäť potvrdzuje dôveryhodnosť vykonaných subjektívnych testov. Tiež je z týchto grafov zrejmé, že ľudský mozog nevníma zmeny použitého kodeku pri vysokom bitovom toku, tak intenzívne ako tomu je v prípade objektívneho ohodnotenia. Tu vidíme, že sa hodnotenia užívateľov pohybujú zhruba na podobnej úrovni. Efektivita kódovania s použitím H.265 je vyššia ako pri H.264; H.265 prekonal H.264 v rozlíšení FHD ako aj UHD.



Graf 5 Priemerné výsledky vyhodnotenia kvality objektívnou a subjektívnou metrikou

## 7.1 Predikcia subjektívneho hodnotenia kvality videosekvencie na základe scény a objektívneho hodnotenia videosekvencie

V tejto podkapitole som popísal princíp fungovania modelu z pohľadu výberu správnej topológie pre ďalšie vykonanie testov, popísal simuláciu navrhnutého modelu a predstrel jeho časť, ktorá dokáže predikovať ako vníma používateľ kvalitu ponúkaného obsahu pri zadanej hodnote objektívneho ohodnotenia kvality metrikou SSIM. V kapitole sú popísané výsledky úspešnosti predikcie tejto časti modelu.



### 7.1.1 Výber vhodnej topológie

Pre dosiahnutie optimálneho výsledku je nevyhnutné nájsť správnu topológiu z pohľadu počtu skrytých vrstiev a počtu neurónov v jednotlivých vrstvách. Malý počet neurónov v jednotlivých vrstvách môže viesť k nedokonalému natrénovaniu siete, naopak, príliš veľký počet neurónov môže viesť k pretrénovaniu siete, čím by sieť bola generalizovaná na svoje vzory a neschopná predikovať nové, simulované hodnoty.

Na základe charakteru dát, ktoré sú jednoznačne určené vstupmi (5 vstupov určujúcich videosekvenciu a práve 1 výstup v podobe MOS skóre) a empirických pozorovaní, ktoré ukázali jednoduchú štruktúru skrytých vrstiev, som určil základné topológie pre ďalší výskum:

- 5-1
- 5-1-1
- 5-3-1
- 5-3-2-1
- 5-5-3-1

Simulácia dát teda prebiehala pre 2-4 vrstvy. Pre každú modelovanú topológiu bolo tréovanie neurónovej siete vykonané 10 krát, pričom zakaždým boli náhodne generované váhy a biasy neurónov. Pri každom modelovaní siete sa 10 krát náhodne vygenerujú váhy a biasy. Po natrénovaní, prebiehala simulácia opäť 10 krát, pričom zakaždým bola náhodne rozdelená databáza videosekvencií na tréovaciu a validačnú množinu, čiže pre každé z 10 opakovaní modelovania siete, bola 10 krát náhodne vygenerovaná tréovacia a validačná množina (bližšie popísané v kapitole 6.1).

Výsledky simulácie základných vrstiev pre rozlíšenie UHD sú v tabuľke 3. Tabuľka obsahuje optimálne výsledky, z pohľadu priemerného korelačného Pearsonovho koeficientu testovacej sady, informácie o priemernom čase a priemernom počte iterácií potrebnom na dosiahnutie optimálneho výsledku. Pri výbere najlepších topológií som v rámci celej práce zvolil postup:

- Výber topológie, s čo najlepšou výslednou koreláciou, s čo najmenšou chybou
- Nižší počet epoch a s tým spojený čas potrebný na dosiahnutie optimálneho výsledku
- Počet neurónov v jednotlivých vrstvách

Výsledky z tabuľky 3 naznačujú, že dvojvrstvová neurónová sieť dosahuje lepšie výsledky z pohľadu sledovaných parametrov, preto pri ďalších simuláciách som pracoval s 2 vrstvami, pričom som hľadal ešte optimálnejší výsledok. Ten je v tabuľke farebne odlišený a je vidieť, že pri vybranej topológii bol dosiahnutý daný cieľ. Tieto topológie som preto vybral ako najlepšie a je na nich vykonaný ďalší výskum.

Tabuľka 3 Najlepšie topológie predikcie MOS na základe SSIM, UHD

	Topológia	$R^2_{\text{TRAIN}} [-]$	Čas [s]	Počet epoch [-]
<b>H.264 + H.265</b>	5-1	0,97	64,962	681
	<b>51-25</b>	<b>0,996</b>	<b>12,446</b>	<b>302</b>
	5-1-1	0,974	99,480	762
	5-3-1	0,977	108,160	965
	5-3-2-1	0,974	91,926	820
	5-5-3-1	0,98	106,551	924
<b>H.264</b>	5-1	0,934	93,305	853
	<b>31-15</b>	<b>0,993</b>	<b>281,610</b>	<b>273</b>
	5-1-1	0,953	91,826	963
	5-3-1	0,934	135,354	1101
	5-3-2-1	0,957	152,567	1118
	5-5-3-1	0,953	94,658	888
<b>H.265</b>	5-1	0,971	96,904	808
	<b>39-19</b>	<b>0,992</b>	<b>30,651</b>	<b>278</b>
	5-1-1	0,969	91,443	836
	5-3-1	0,975	116,228	1037
	5-3-2-1	0,984	103,205	827
	5-5-3-1	0,989	118,936	918

Po výbere najlepších topológií z priemerných výsledkov boli nasimulované dáta pri nastavení týchto topológií. Výsledky sú uvedené v tabuľke 4, kde je uvedená chybovosť tréningovej sady, s použitím funkcie MSE, korelácia tréningovej ( $R^2_{\text{TRAIN}} [-]$ ), validačnej ( $R^2_{\text{VAL}} [-]$ ) a testovacej ( $R^2_{\text{TEST}} [-]$ ) sady vypočítaná pomocou Pearsonovho korelačného koeficientu. Tabuľka tiež ukazuje výsledky pre simuláciu oboch rozlíšení a oboch kompresných formátov.

Tabuľka 4 Simulácia dát pri vybraných topológiách predikcie MOS na základe SSIM

Rozlíšenie	Kodek	Topológia	MSE	$R^2_{\text{TRAIN}} [-]$	$R^2_{\text{VAL}} [-]$	$R^2_{\text{TEST}} [-]$
<b>UHD</b>	<b>H.264+H.265</b>	<b>51-25</b>	0,002	0,996	0,996	0,997
	<b>H.264</b>	<b>31-15</b>	0,004	0,994	0,987	0,999
	<b>H.265</b>	<b>39-19</b>	0,004	0,993	0,998	0,997
<b>FHD+UHD</b>	<b>H.264+H.265</b>	<b>47-23</b>	0,006	0,986	0,991	0,993

Z hľadiska veľkosti súborov sa výsledky ostatných topológií neuvádzajú v práci, ale sú uložené na priloženom médiu.

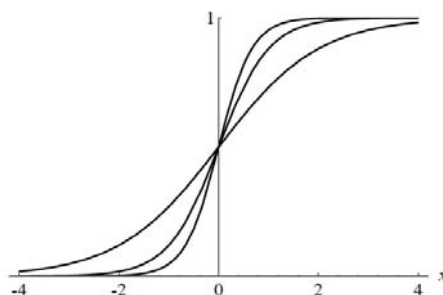
### 7.1.2 Vplyv aktivačnej funkcie

Je veľmi dôležité, aby pre potreby tréningu bola zvolená vhodná aktivačná funkcia. Pre tvorbu modelu boli použité funkcie:

- Log-sigmoidná funkcia logsig - jedna z najpoužívanejších aktivačných funkcií pre siete spätného šírenia je sigmoid, reálna funkcia sc:  $\mathbb{R} \rightarrow (0, 1)$  definovaná ako:

$$s_c = \frac{1}{1 + e^{-cx}} \quad (23)$$

Konštanta  $c$  môže byť zvolená ľubovoľne. Tvar zmeny sigmoidu sa mení podľa hodnoty  $c$ , ako je vidieť na obrázku, ktorý zobrazuje tvar sigmoidu pre  $c = 1$ ,  $c = 2$  a  $c = 3$ . Vyššie hodnoty  $c$  prinášajú tvar sigmoidu bližšie k tvaru krokovej funkcie.



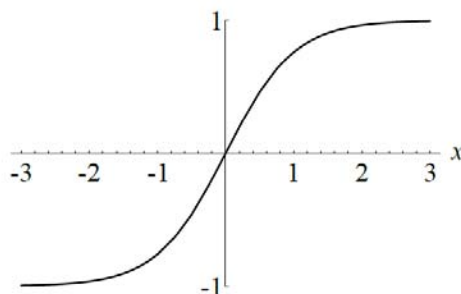
Obrázok 27 Log-sigmoidná aktivačná funkcia

Výstupný rozsah sigmoidu je medzi 0 a 1. Pri váhach  $w_1, \dots, w_n$  a prahu  $\theta$ , sigmoid vypočíta pre vstupy  $x_1, \dots, x_n$  výstup:

$$\frac{1}{1 + \exp(\sum_{i=1}^n w_i x_i - \theta)} \quad (24)$$

- Tan-sigmoid aktivačná funkcia tansig – alternatíva k sigmoidu je symetrický sigmoid  $S(x)$  definovaný ako [35]:

$$S(x) = 2s(x) - 1 = \frac{1 - e^{-x}}{1 + e^{-x}} \quad (25)$$



Obrázok 28 Tan-sigmoid aktivačná funkcia

- Lineárna aktivačná funkcia “Purelin” – funkcia použitá na poslednej vrstve viacvrstvovej siete pre odhad výstupu neurónovej siete v podobe konkrétnej hodnoty

V backpropagation sieťach je dôležitá schopnosť derivácie ľubovoľnej použitej aktivačnej funkcie. Korenšponujúce derivačné funkcie vyššie spomínaných aktivačných funkcií sú dtansig, dlogsig a dpurelin. Výsledky pri použití rozdielnych aktivačných funkcií sú veľmi malé. Lepšie výsledky v každej simulácii a v každej neurónovej sieti v tejto práci dosahuje aktivačná funkcia *tansig*. Porovnanie aktivačných funkcií pri predikcii subjektívneho ohodnotenia na základe objektívnych výsledkov je v tabuľke 5.

Tabuľka 5 Vplyv aktivačnej funkcie na predikciu MOS na základe SSIM

Rozlíšenie	Kodek	Topológia	Aktivačná funkcia	MSE	R <sup>2</sup> <sub>TRAIN</sub> [-]	R <sup>2</sup> <sub>VAL</sub> [-]	R <sup>2</sup> <sub>TEST</sub> [-]
UHD	H.264+H.265	51-25	tansig	0,002	0,996	0,996	0,997
			logsig	0,019	0,973	0,958	0,986
	H.264	31-15	tansig	0,004	0,994	0,987	0,999
			logsig	0,019	0,974	0,984	0,992
FHD+UHD	H.265	39-19	tansig	0,004	0,993	0,998	0,997
			logsig	0,020	0,952	0,950	0,976
	H.264+H.265	47-23	tansig	0,006	0,986	0,991	0,993
			logsig	0,025	0,956	0,958	0,956

Ako najlepšie boli vyhodnotené topológie s 2 skrytými vrstvami, so škálovaním vstupných a výstupných hodnôt v intervale  $\langle -1;1 \rangle$ , pri použití aktivačnej funkcie *tansig*, algoritmu Backpropagation implementovaného typom *traingdx*, nakoľko na natrénovanie potreboval oveľa menší počet iterácií, čím výrazne znížil časové prostriedky na dosiahnutie definovaného cieľa (platí pre všetky vykonané simulácie v práci).

Štatistické vyhodnotenie úspešnosti predikcie navrhnutého modelu a tejto simulácie sa nachádza v kapitole 9.1.

## **8 Vytvorenie optimálnej mapovacej funkcie predikcií objektívnych metrík a výsledkov subjektívnych hodnotení na určenie hraníc TI a SI**

Pri nasadení variabilného bitového toku je potrebné jeho správne určenie, aby sa dosiahla optimálna kvalita pre jednotlivý typ videosekvencií. K tomuto kroku bolo nutné vybrať široké spektrum videosekvencií, aby sa poskytlo pokrytie rôznorodých scén a daný model mohol byť použitý v praxi. Tieto videosekvencie som vyberal na základe kritických parametrov určujúcich ich typ, čo popisuje kapitola 5. Na základe vykonaného ohodnotenia kvality pomocou objektívnej metriky SSIM a subjektívnej metriky ACR je možné stanoviť požadovanú kvalitu, či už na základe subjektívneho vnímania koncového používateľa alebo na základe objektívnej metriky. Mapovacia funkcia dokáže vyhodnotiť optimálny bitový tok tak, aby spĺňal zadanú kvalitu vo forme MOS skóre alebo formou SSIM výsledku. Vytvorený model nám tak vyhodnotí, aký bitový tok spĺňa zadané požiadavky kvality pre zadaný typ videoscény. Teda, ak požadujeme, aby kvalita danej videoscény v MOS vyjadrení bola na prahu akceptovateľnosti (MOS=3) alebo prijateľného vnímania kvality (MOS=4) model nám určí, aký je potrebný bitový tok pre dosiahnutie tohto cieľa, na základe priemerného ohodnotenia užívateľov, pre zadaný typ scény. Pre potreby poskytovateľov týchto služieb je taktiež možné vyhodnotiť potrebný bitový tok na základe kvality zadanej formou SSIM metriky. Vytvorený model umožňuje aj opačný typ predikcie a to vyhodnotenie kvality na základe zadaného typu videosekvencie a použitého bitového toku. Model je schopný predikovať kvalitu v podobe MOS skóre alebo SSIM hodnoty.

### **8.1 Predikcia bitového toku na základe požadovanej kvality vyjadrenej metrikou SSIM**

Cieľom vytvoreného modelu je určiť postačujúci bitový tok pre danú scénu. V tejto podkapitole je popísaný model predikcie bitového toku na základe zadanej kvality vyjadrenej formou SSIM hodnoty. Parametre, ktoré vstupujú do neurónovej siete sú dané charakteristikou danej videosekvencie z pohľadu SI a TI informácie, kvalitatívnymi parametrami a požadovanou kvalitou vo forme SSIM. Výsledky dosiahnuté pre rozlíšenie UHD a pre spojenú simuláciu oboch rozlíšení sú popísané v nasledujúcich podkapitolách. Dôvodom vyhodnotenia jednotlivých typov rozlíšenia je možné porovnanie schopnosti modelu predikovať ako aj obe rozlíšenia v jednej simulácii, tak aj vyhodnotenie simulácií separátne pre obe rozlíšenia.

Keďže vývoj technológií neustále rastie rýchlym tempom a v praxi sa čoraz viac používajú lepšie technológie, je dôležité zamerať sa na predpokladané nasadenie vyššieho rozlíšenia. To je ďalší dôvod na separátnu simuláciu kompresných štandardov pri rozlíšení UHD, čo umožní pohľad pri stanovení vhodného bitového toku tohto rozlíšenia, či už by bol nasadený VBR alebo by sa predikcia tohto modelu použila pre minimálny bitový tok s použitím CBR. Preto v práci uvádzam aj výsledky pre jednotlivé kompresné štandardy pri rozlíšení UHD. Samozrejme, model umožňuje vyhodnotiť predikcie pre jednotlivé kodeky aj pri rozlíšení FHD a tiež pri spoločnej simulácii rozlíšení. Výber prezentovaných výsledkov som zamerlal na plánované použitie rozlíšenia UHD, kde som aj poukázal na rozdiely pri oddelenej simulácii kompresných štandardov. Následne vyhodnocujem výsledky pre simuláciu oboch

rozlíšení a aj kompresných štandardov spoločne. Výsledky poukazujú na vysokú úspešnosť predikcie vo všetkých spomenutých simuláciách.

*Tabuľka 6 Najlepšie topológie predikcie BR na základe SSIM*

Rozlíšenie	Kodek	Topológia	$R^2_{TEST}$ [-]	Čas [s]	Počet epoch [-]
UHD	<b>H.264+H.265</b>	43-21	0,946	191,265	338
	<b>H.264</b>	79-39	0,970	363,562	266
	<b>H.265</b>	63-31	0,939	627,246	360
FHD+UHD	<b>H.264+H.265</b>	47-23	0,965	217,427	294

Tabuľka 7 zobrazuje simuláciu najlepšie vyhodnotených topológií, z tabuľky 6, pre rozlíšenie UHD a tiež pre simuláciu oboch použitých rozlíšení. Pri UHD je možnosť porovnať simuláciu pre oba kompresné štandardy spolu alebo oddelene.

*Tabuľka 7 Simulácia dát pri vybraných topológiách predikcie BR na základe SSIM*

Rozlíšenie	Kodek	Topológia	MSE	$R^2_{TRAIN}$ [-]	$R^2_{VAL}$ [-]	$R^2_{TEST}$ [-]
UHD	<b>H.264+H.265</b>	43-21	0,041	0,949	0,944	0,963
	<b>H.264</b>	79-39	0,0001	0,999	0,999	0,999
	<b>H.265</b>	63-31	0,001	0,999	0,999	0,999
FHD+UHD	<b>H.264+H.265</b>	47-23	0,029	0,963	0,957	0,985

Tabuľka 8 popisuje vplyv aktivačnej funkcie na simuláciu predikcie BR na základe SSIM. Opäť výsledky poukazujú, že aktivačná funkcia *tansig* lepšie aproximuje hodnoty pri tréningu vzhľadom na štruktúru tréningových dát.

*Tabuľka 8 Vplyv akt. funkcie na predikciu BR na základe SSIM*

Rozlíšenie	Kodek	Topológia	Aktivačná funkcia	MSE	$R^2_{TRAIN}$ [-]	$R^2_{VAL}$ [-]	$R^2_{TEST}$ [-]
UHD	<b>H.264+H.265</b>	<b>43-21</b>	<b>tansig</b>	0,041	0,949	0,944	0,963
			<b>logsig</b>	0,257	0,576	0,504	0,759
	<b>H.264</b>	<b>79-39</b>	<b>tansig</b>	0,000	1,000	1,000	1,000
			<b>logsig</b>	0,004	0,992	0,995	0,998
	<b>H.265</b>	<b>63-31</b>	<b>tansig</b>	0,001	0,999	0,999	0,999
			<b>logsig</b>	0,009	0,987	0,995	0,996
FHD+UHD	<b>H.264+H.265</b>	<b>47-23</b>	<b>tansig</b>	0,029	0,963	0,957	0,985
			<b>logsig</b>	0,262	0,553	0,6	0,595

## Vylepšenie trérovacieho procesu

Pre dosiahnutie optimálneho výsledku je potrebný výpočet chyby pri tréovaní. V simuláciách bola použitá funkcia MSE (21), ktorú prostredie Matlab štandardne využíva pri dopredných neurónových sieťach. Jeden z problémov, ku ktorému dochádza pri tréovaní neurónovej siete je jej pretrénovanie (overfitting). Síce, chyba tréovacej množiny sa stále znižuje a efektivita natréovaných dát sa môže zvyšovať, pri simulácii nových dát, s ktorými sieť dovtedy nepracovala, chybovosť narastá. Sieť bola vyškolená na tréovacie hodnoty, ale nenaučená pre generalizáciu nových dát.

## Regulácia

Prvá metóda pre vylepšenie tréovania je regulácia – modifikácia funkcie pre výpočet chyby pri tréovaní. Štandardná funkcia MSE doplnená o strednú hodnotu váh a biasov.

$$msereg = \gamma mse + (1 - \gamma) msw \quad (26)$$

kde  $\gamma$  je pomer výkonnosti.

Použitie tejto funkcie spôsobí, že sieť bude mať menšie váhy a biasy, bude reagovať jemnejšie a zníži sa pravdepodobnosť, že sieť bude pretrénovaná. Problémom je obtiažné určenie optimálnej hodnoty parametra pomeru výkonnosti. Ak je príliš veľký, sieť môže byť pretrénovaná. Ak je naopak veľmi malý, sieť nebude adekvátne natréovaná pre tréovacie dáta.

## Skoré ukončenie tréovania

Ďalšou metódou vyhnutiu sa pretrénovaniu je skoré zastavenie tréovania. V tomto prípade sa používa validačná sada. Chybovosť tejto sady je počítaná a monitorovaná pri tréovaní siete. Chybovosť validačnej sady normálne klesá v počiatočnej fáze tréovania rovnako ako chybovosť pri tréovacej sade. Avšak ak sa sieť začne pretrénovávať, chybovosť validačnej sady zvyčajne začne stúpať. Keď sa chybovosť na validačnej sade zvýši po určitom špecifikovanom opakovaní, tréovanie je zastavené.

V práci som overil použitie oboch metód. Lepšie výsledky pri všetkých simuláciách v práci dosahoval algoritmus s použitím validačnej sady. Porovnanie výsledkov tejto simulácie sa nachádza v prílohe C (1). Štatistické vyhodnotenie úspešnosti predikcie tejto simulácie sa nachádza v kapitole 9.2.

## 8.2 Predikcia bitového toku na základe požadovanej kvality vyjadrenej pomocou MOS skóre

Pri určovaní požadovanej kvality je najdôležitejším faktorom uspokojenie koncových zákazníkov a pre poskytovateľov je prínosné poznať názor používateľa a jeho vnímanie kvality. Ak poskytovateľ vie, ako vníma jeho zákazník kvalitu, dokáže prispôbiť kvalitatívne nastavenie ponúkaného obsahu a nastaviť tak vhodné kvalitatívne parametre. Jedným z nich je aj bitový tok. Pri vytváraní tejto práce boli používateľmi ohodnotené rôzne typy scén kódované použitím rôznorodého bitového toku. Tým vznikla široká databáza ohodnotených videosekvencií na základe rôznej zmeny bitového toku. S použitím tejto ohodnotenej databázy je možné stanoviť vhodný bitový tok na základe požadovanej kvality vyjadrenej pomocou MOS stupnice a predikovať ho pomocou vytvoreného modelu.

Tabuľka 9 obsahuje vybrané najlepšie topológie pri predikcii bitového toku na základe informácií o scéne a kvalitatívnom ohodnotení scény pomocou MOS skóre. Najlepšie topológie sú vybrané na základe priemerného Pearsonovho korelačného koeficientu testovacej sady, času a počtu epoch potrebných na natrénovanie danej topológie. Výsledky sú zobrazené pre rozlíšenie UHD, kde je porovnanie spoločnej simulácie videosekvencií kódovaných v oboch kompresných formátoch ako aj separátna simulácia týchto kodekov. V tabuľke je tiež zobrazená najlepšia topológia pre spoločnú simuláciu oboch rozlíšení.

*Tabuľka 9 Najlepšie topológie predikcie BR na základe MOS*

Rozlíšenie	Kodek	Topológia	$R^2_{TEST}$ [-]	Čas [s]	Počet epoch [-]
UHD	H.264+H.265	47-23	0,988	12,163	270
	H.264	35-17	0,992	18,957	330
	H.265	71-35	0,990	20,299	347
FHD+UHD	H.264+H.265	71-35	0,990	29,113	250

Tabuľka 10 obsahuje výsledky dosiahnuté simuláciou pri vybraných topológiách z tabuľky 9. V tabuľke je vypočítaná najmenšia dosiahnutá chyba pri tréningu pomocou funkcie MSE a tiež je zobrazená korelácia z pohľadu Pearsonovho korelačného koeficientu pre tréningovú, validačnú a aj testovaciu sadu.

*Tabuľka 10 Simulácia dát pri vybraných topológiách predikcie BR na základe MOS*

Rozlíšenie	Kodek	Topológia	MSE	$R^2_{TRAIN}$ [-]	$R^2_{VAL}$ [-]	$R^2_{TEST}$ [-]
UHD	H.264 + H.265	47-23	0,013	0,982	0,991	0,992
	H.264	35-17	0,006	0,992	0,991	0,996
	H.265	71-35	0,011	0,988	0,982	0,993
FHD+UHD	H.264 + H.265	71-35	0,008	0,989	0,99	0,995

Štatistické vyhodnotenie úspešnosti predikcie navrhnutého modelu pri vybraných topológiách sa nachádza v kapitole 9.3. Vplyv aktivačnej funkcie pre vybrané topológie je v tabuľke 11 a opäť potvrdzuje lepšie výsledky pri aktivačnej funkcii tansig.

*Tabuľka 11 Vplyv akt. Funkcie na predikciu BR na základe SSIM*

Rozlíšenie	Kodek	Topológia	Aktivačná funkcia	MSE	$R^2_{TRAIN}$ [-]	$R^2_{VAL}$ [-]	$R^2_{TEST}$ [-]
UHD	H.264+H.265	47-23	tansig	0,013	0,982	0,991	0,992
			logsig	0,046	0,937	0,950	0,950
	H.264	35-17	tansig	0,006	0,992	0,990	0,996
			logsig	0,019	0,968	0,972	0,963
	H.265	71-35	tansig	0,011	0,988	0,982	0,993
			logsig	0,060	0,933	0,893	0,949
FHD+UHD	H.264+H.265	71-35	tansig	0,008	0,989	0,99	0,995
			logsig	3,011	0,683	0,612	0,762



Aj pre túto topológiu bolo vykonané porovnanie zmeny funkcie na výpočet chybovosti pri nastavení rôznych hodnôt parametra pomeru výkonnosti. Výsledky tohto porovnania nájdeme v prílohe C (2).

## 9 Verifikácia a validácia modelu

V tejto kapitole sa zaoberám overením dosiahnutých vlastností mnou navrhnutého prístupu. Táto časť obsahuje overenie jednotlivých častí predikcií, nakoľko mnou vytvorený klasifikátor, využívajúci umelú inteligenciu, dokáže predikovať subjektívne ohodnotenie kvality videosekvencií na základe objektívneho ohodnotenia, vhodný bitový tok na základe vstupných, objektívne získaných parametrov a mapovacia funkcia dokáže správne odhadnúť optimálny bitový tok na základe požadovanej kvality pre daný typ scény.

### 9.1 Verifikácia a validácia modelu predikcie subjektívnej hodnoty na základe scény, bitového toku a objektívneho výsledku metrikou SSIM

Ako najlepšia topológia pri simulovaní oboch kompresných štandardov v časti modelu popísanom v 7.1 bola vybratá topológia 51-25. Táto topológia dosahovala najlepšiu koreláciu a najlepšiu predikciu simulovaných dát. V tabuľke 3 môžeme vidieť výsledky predikcie modelu a porovnať ich s výsledkami simulácií pre oba použité kompresné štandardy separátne. Hodnoty dosiahnuté simuláciou je potrebné porovnať s referenčnými hodnotami dosiahnutými subjektívnou metrikou ACR. Verifikovanie nasimulovaných dát je možné dokázať viacerými štatistickými spôsobmi. Pri každej použitej metóde, pracujeme s 95%-ným intervalom spoľahlivosti. Overenie simulácie oboch kompresných štandardov a predikcie MOS skóre pomocou objektívneho ohodnotenia kvality:

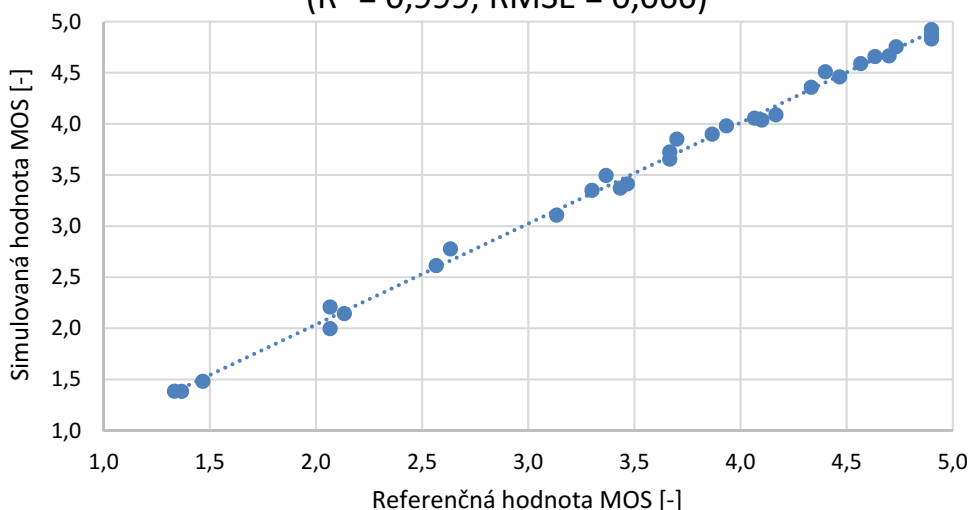
- Prvým krokom porovnania vierohodnosti dát je porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu. Ako vidieť z tabuľky 12, minimálne a maximálne hodnoty, hodnoty priemerov, mediánov a konfidenčných intervalov sa veľmi približujú, čím vidíme veľkú podobnosť medzi dátami a tým môžeme považovať simuláciu a použitý model, z pohľadu tohto štatistického vyhodnotenia, vierohodný.

Tabuľka 12 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu MOS pomocou SSIM, H.264+H.265, UHD

	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	3,481	3,501
Minimum [-]	1,333	1,382
Maximum [-]	4,900	4,920
Rozsah [-]	3,567	3,538
Medián [-]	3,700	3,850
Konfidenčný interval [-]	<3,067 ; 3,896>	<3,091 ; 3,910>
Pearsonov korelačný koeficient [-]		0,999
Spearmanov korelačný koeficient [-]		0,996

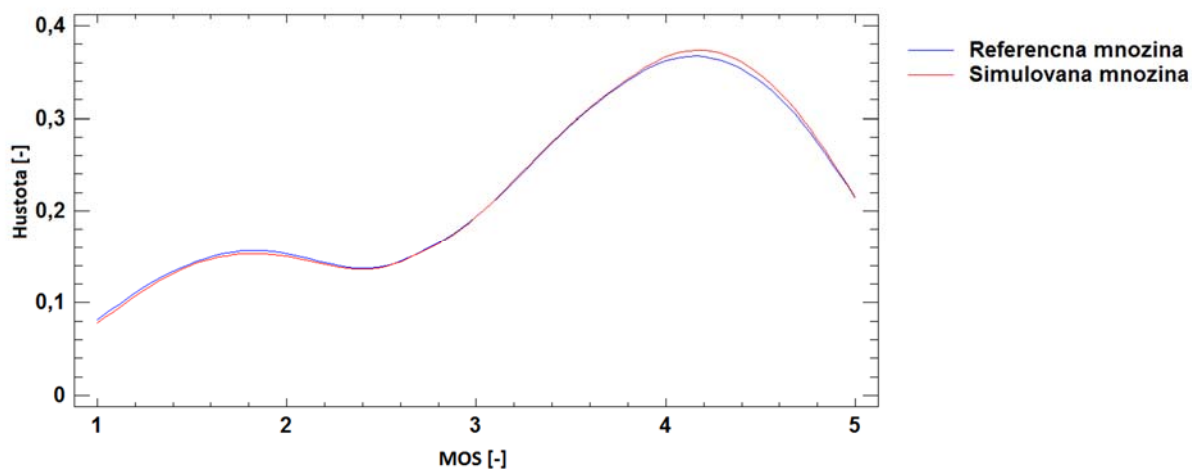
- Korelačný diagram a funkcia RMSE – na základe korelácie je možné porovnať referenčnú a simulovanú množinu. Ako vidieť na nasledujúcom grafe, korelácia vyjadruje veľmi podobné simulované dáta s referenčnými vzorkami, čo opäť dokazuje vysokú mieru úspešnosti predikcie. Korelácia je overená funkciou RMSE, ktorá v tomto prípade potvrdzuje vierohodnosť dát, keď hovorí o 1,32 %-nej chybe predikcie.

Testovacia sada - H.264 + H.265  
( $R^2 = 0,999$ ; RMSE = 0,066)



Graf 6 Korelačný diagram, predikcia MOS pomocou SSIM, H.264+H.265, UHD

- Krivka hustoty rozdelenia – neparametrický odhad funkcie hustoty pravdepodobnosti. Rozdelenie pravdepodobnosti je možné chápať ako zobrazenie, ktoré každému elementárnemu javu priradzuje určité reálne číslo, ktoré charakterizuje pravdepodobnosť tohto javu. Hustota pravdepodobnosti je funkcia, ktorej hodnotu pre ľubovoľne zvolený prvok z množiny možných vzorov (hodnôt náhodnej premennej) je možné interpretovať ako relatívnu početnosť hodnoty tohto prvku v rámci celej množiny možných vzorov. Z grafov je vidno, že simulovaná množina blízko odpovedá referenčnej, čo opäť potvrdzuje vierohodnosť modelu.



Graf 7 Krivka hustoty rozdelenia, predikcia MOS pomocou SSIM, H.264+H.265, UHD

Podobne ako pri simulácii oboch kompresných štandardov je v práci uvedené štatistické vyhodnotenie dát simulovaných separátne pre oba použité kodeky. Je tak možné porovnať správanie sa siete pri simulovaní dát spolu a simulovaní dát osobitne a tak určiť, či niektorý kompresný štandard vplýva viac na výsledky. Verifikácie modelu pre simuláciu jednotlivých kodekov a pre simuláciu oboch rozlíšení sa nachádzajú v prílohe D.1.

## 9.2 Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou objektívnym hodnotením metrikou SSIM

V tejto kapitole sa nachádza verifikovanie predikcie bitového toku na základe požadovaného kvalitatívneho vyhodnotenia videosekvencií pomocou metriky SSIM. Nachádza sa tu vyhodnotenie pre simuláciu oboch kompresných štandardov v jednom simulačnom behu pre rozlíšenie UHD a simuláciu oboch rozlíšení aj kompresných štandardov v jednom behu.

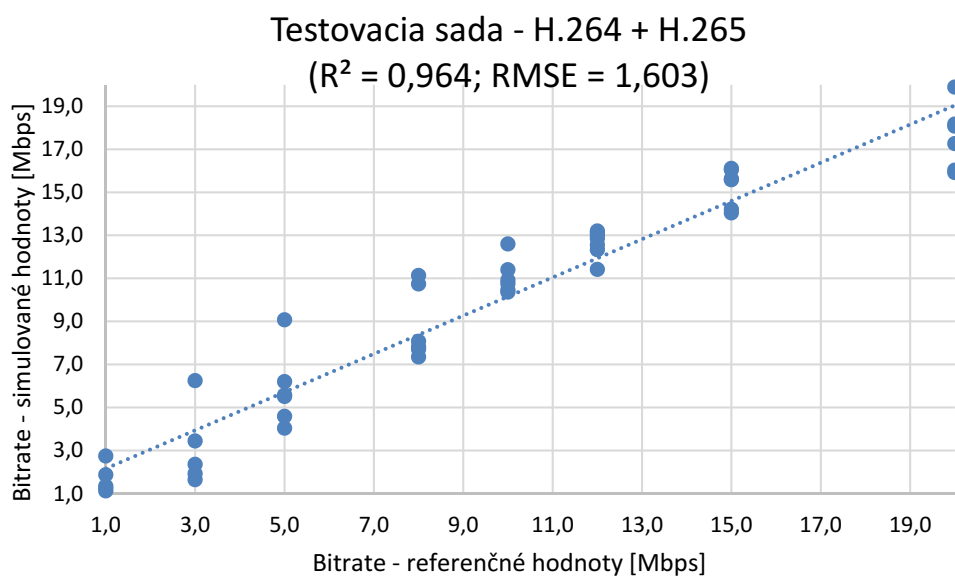
### UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu v prípade simulácie bitového toku pre požadovanú kvalitu vyjadrenú metrikou SSIM ukazuje blízkosť priemerných hodnôt, mediánov a konfidenčných intervalov

*Tabuľka 13 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.264+H.265, UHD*

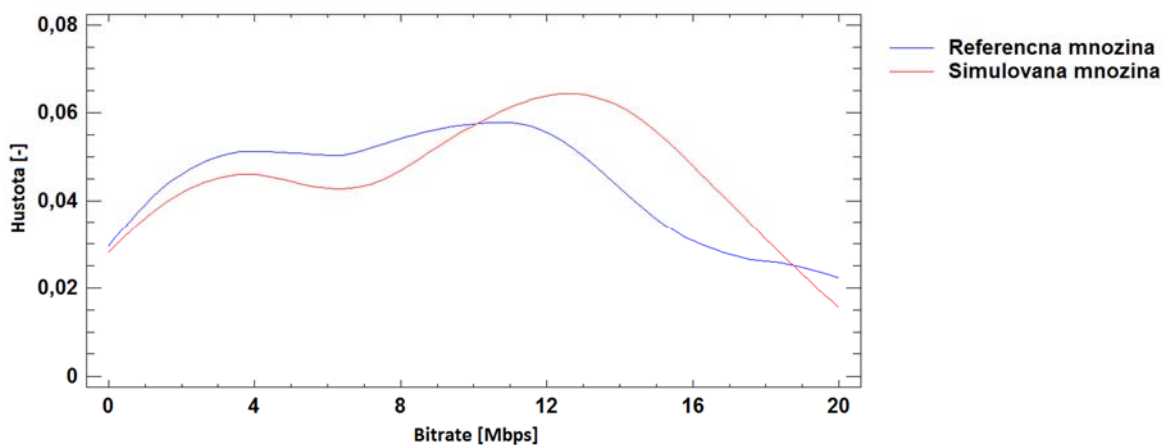
	Referenčná množina	Simulovaná množina
<b>Priemerná hodnota [-]</b>	9,438	9,664
<b>Minimum [-]</b>	1,000	0,369
<b>Maximum [-]</b>	20,000	19,884
<b>Rozsah [-]</b>	19,000	19,515
<b>Medián [-]</b>	10,000	10,744
<b>Konfidenčný interval [-]</b>	<7,708 ; 11,167>	<8,071 ; 11,257>
<b>Pearsonov korelačný koeficient [-]</b>		0,964
<b>Spearmanov korelačný koeficient [-]</b>		0,976

- Korelačný diagram a funkcia RMSE – v tejto simulácii je o niečo vyššia chyba v porovnaní s predikciou BR na základe MOS, funkcia RMSE vyjadruje 8,015 %-nú chybu predikcie.



Graf 8 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, UHD

- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu pri predikcii BR na základe požadovanej kvality zadanej formou SSIM pre UHD rozlíšenie



Graf 9 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.265, UHD

## FHD + UHD

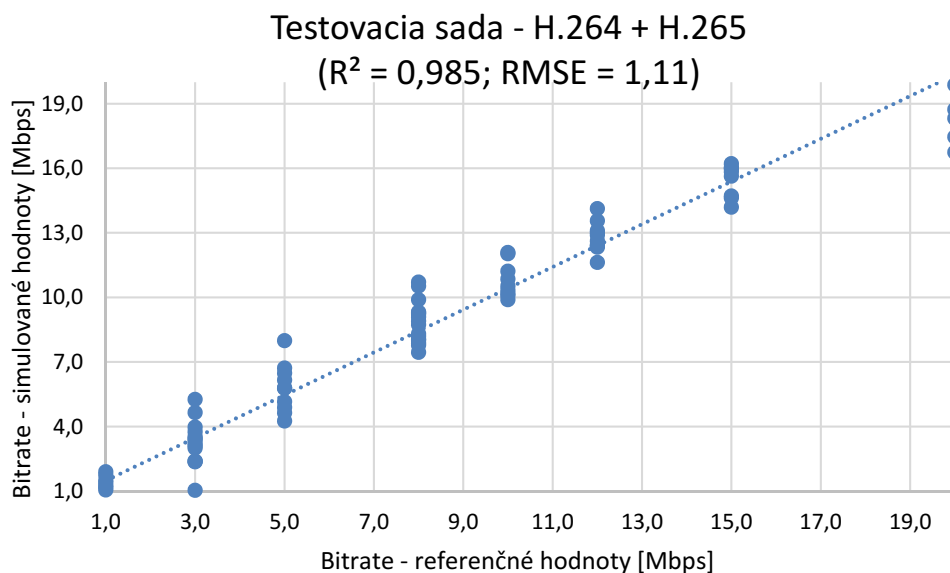
Simuláciou BR na základe vstupných parametrov SI, TI , SSIM, kompresného štandardu a použitého rozlíšenia získavame komplexný nástroj na hodnotenie kvality videosekvencií z pohľadu nastavenia kvalitatívnych parametrov pre odlišné skupiny scén, ktoré sú kódované pre rozličné kompresné štandardy a pre rozličné rozšírenia. Výsledky predikcie bitového toku nám aj napriek tomu ponúkajú vysokú presnosť predikcie a je možné teda vykonávať simulácie v jednom behu. Model sa adaptuje na základe zadaných požiadaviek a rozozná rozdiel v kvalite pre jednotlivé fundamentálne odlišné skupiny videosekvencií.

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

*Tabuľka 14 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.264+H.265, FHD+UHD*

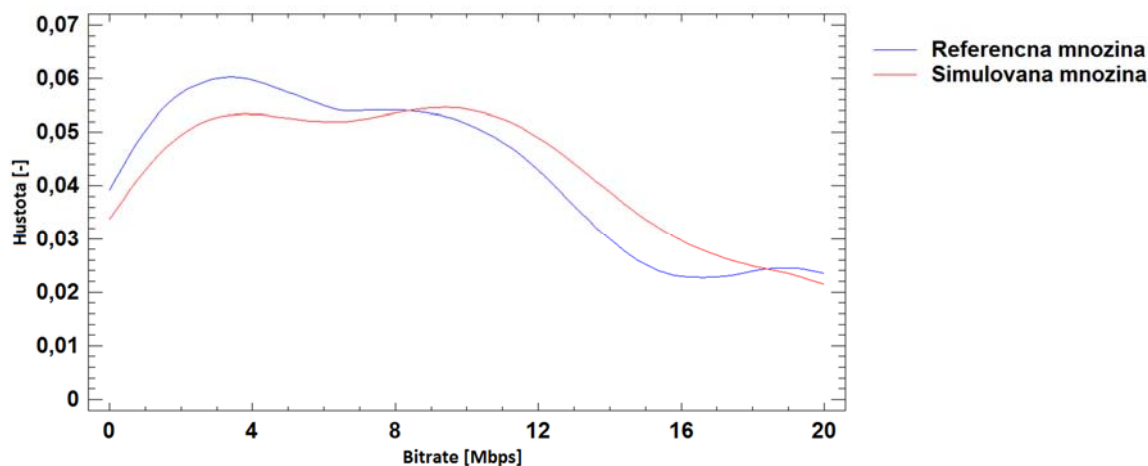
	Referenčná množina	Simulovaná množina
<b>Priemerná hodnota [-]</b>	8,583	9,14
<b>Minimum [-]</b>	1	0,73
<b>Maximum [-]</b>	20,00	22,11
<b>Rozsah [-]</b>	19,00	22,84
<b>Medián [-]</b>	8,00	8,81
<b>Konfidenčný interval [-]</b>	<5,43 ; 7,22>	<5,39 ; 7,18>
<b>Pearsonov korelačný koeficient [-]</b>		0,985
<b>Spearmanov korelačný koeficient [-]</b>		0,970

- Korelačný diagram a funkcia RMSE – ako popisuje graf 10, predikcia bitového toku na základe scény charakterizovanej pomocou SI a TI informácií, použitého rozlíšenia a kompresného štandardu dosahuje vysokej korelácie. Pomocou výpočtu funkcie RMSE je odhadovaná veľkosť chyby pri predikcii 5,55%.



Graf 10 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, FHD+UHD

- Krivka hustoty rozdelenia tejto simulácie nám vizuálne potvrdzuje výsledky korelácie a tiež vypočítanej chybovosti, keďže vidíme približovanie simulovanej množiny k referenčnej.



Graf 11 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.265, FHD+UHD

### 9.3 Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou pomocou MOS stupnice

V tejto podkapitole sú overené výsledky pre predikciu bitového toku pri vstupných informáciách, ktoré charakterizujú danú scénu a vyjadrujú ohodnotenie kvality danej videosekvencie používateľmi, vyjadrené v priemernej hodnote MOS.

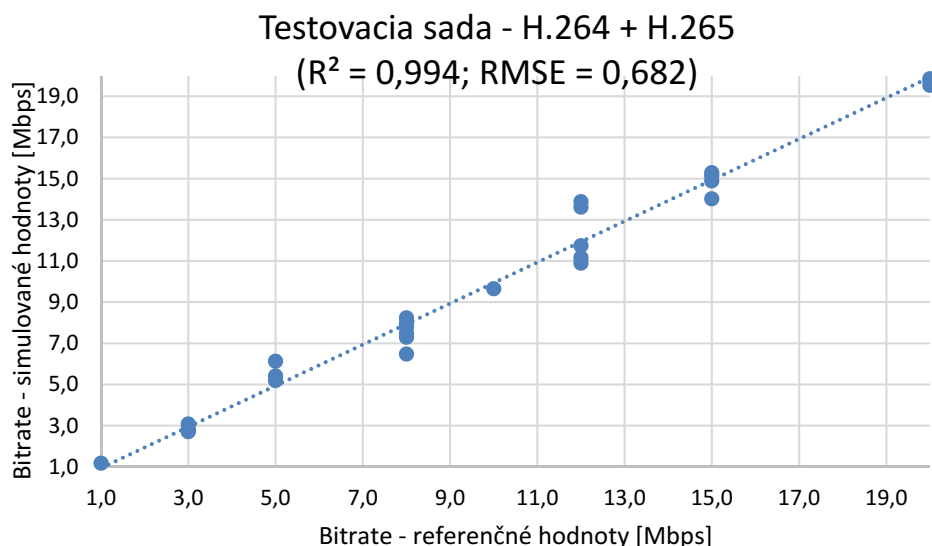
#### UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

Tabuľka 15 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264+H.265, UHD

	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	9,656	9,594
Minimum [-]	1	0,667
Maximum [-]	20	20,467
Rozsah [-]	19	19,801
Medián [-]	8	8,106
Konfidenčný interval [-]	<7,380 ; 11,932>	<7,307 ; 11,881>
Pearsonov korelačný koeficient [-]		0,994
Spearmanov korelačný koeficient [-]		0,99

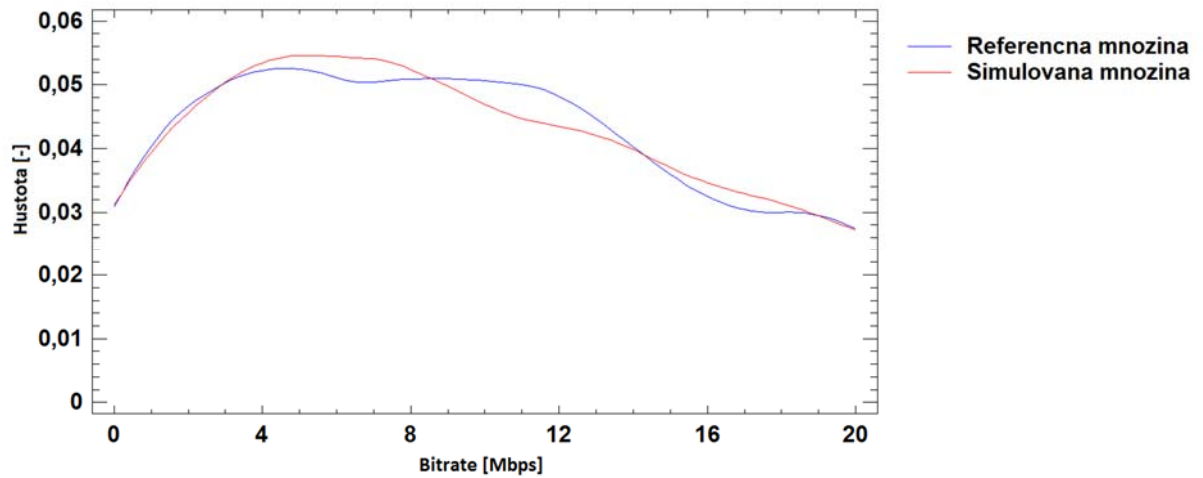
- Korelačný diagram a funkcia RMSE – v prípade výpočtu RMSE pri predikcii bitového toku na základe SI, TI a MOS skóre odpovedá 3,41 %-nej chybovosti.



Graf 12 Korelačný diagram, predikcia BR pomocou MOS, H.264+H.265, UHD



- Krivka hustoty rozdelenia



Graf 13 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264+H.265, UHD

Vyhodnotenie simulácie pre jednotlivé kompresné štandardy je uvedené v prílohe D.2

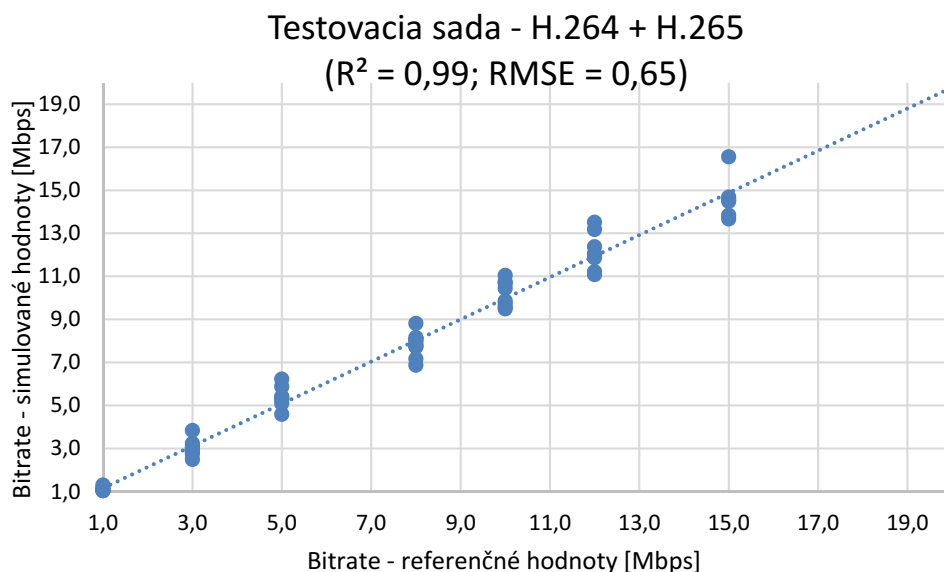
## FHD + UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu v prípade simulácie oboch rozlíšení a tiež kompresných štandardov v jednom behu

Tabuľka 16 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264+H.265, FHD+UHD

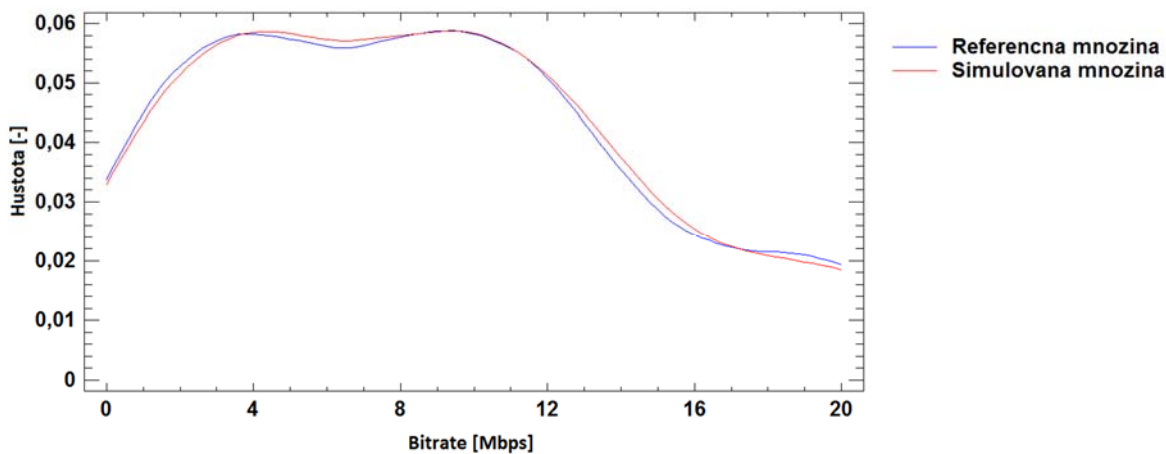
	Referenčná množina	Simulovaná množina
<b>Priemerná hodnota [-]</b>	8,73	8,74
<b>Minimum [-]</b>	1,00	0,93
<b>Maximum [-]</b>	20,00	20,76
<b>Rozsah [-]</b>	19,00	19,83
<b>Medián [-]</b>	8,00	8,05
<b>Konfidenčný interval [-]</b>	<7,28 ; 10,19>	<7,31 ; 10,18>
<b>Pearsonov korelačný koeficient [-]</b>		0,99
<b>Spearmanov korelačný koeficient [-]</b>		0,99

- Korelačný diagram a funkcia RMSE – RMSE v tejto simulácii poukazuje na 3,25%-nú chybovosť predikcie.



Graf 14 Korelačný diagram, predikcia BR pomocou MOS, H.264+H.265, FHD+UHD

- Krivka hustoty rozdelenia



Graf 15 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264+H.26, FHD+UHD

Výsledky štatistického vyhodnotenia a verifikácie modelu v každej simulácii ukazujú vysokú presnosť, čo potvrdzujú sumárne štatistiky, korelácia z pohľadu Pearsonovho a aj Spearmanovho koeficientu, RMSE a tiež krivka hustoty rozdelenia vizuálne potvrdzuje podobnosť referenčnej a simulovanej množiny pri každom type neurónovej siete. V prípade simulácie bitového toku na základe SI, TI a SSIM, pri UHD rozlíšení, vidíme väčšie vyjadrenie chybovosti, ktorá sa ale zmenší, keď simuláciu vykonávame pre obe rozlíšenia v jednom behu, kde je použitých viac dát pre potreby simulácie a následnej predikcie.

## 10 Zhrnutie výsledkov dizertácie pre prax a ďalší rozvoj vedy

Meranie vnímanej kvality audiovizuálnych signálov u koncového používateľa je dôležitým parametrom v mnohých multimediálnych sieťach a aplikáciách. Je dôležitou úlohou pri formovaní audiovizuálneho spracovania, kompresie a prenosu spolu s ich implementáciou, optimalizáciou a testovaním. Poskytovatelia služieb zavádzajú rozličné QoS riešenia na poskytovanie najlepšej kvality vnímania pre ich zákazníkov. Takto sa vytvárajú presné metriky vnímania kvality, ktoré výrazne pomáhajú zlepšovať multimediálne služby prostredníctvom káblových a bezdrôtových sietí.

Nedávny nasadenie technológie 4G v mobilných sieťach viedlo k ďalšiemu nárastu využívaniu multimediálnych služieb a aplikácií, ako napríklad IPTV, mobilných multimédií na smartfónoch, sociálnych sieťach, multimediálnych hrách, hrách založených na virtuálnej realite, videokonferenciách a vzdelávacích multimediálnych prezentáciách. Tieto multimediálne aplikácie sa stali súčasťou každodenného života a očakáva sa ešte vyšší nárast obľúbenosti týchto služieb. Poskytovatelia multimediálnych služieb formulujú rôzne techniky na zabezpečenie lepšej kvality vnímania, ktorú čoraz viac vyžadujú koncoví používatelia. Preto je pri navrhovaní a zostrojení akýchkoľvek súčasných a budúcich multimediálnych sietí a služieb veľmi dôležitý a významný názor vnímania kvality koncovým zákazníkom. Zvuk a videa sú hlavnými zložkami vo väčšine multimediálnych aplikácií. Ich kvalitu ovplyvňuje kódovanie ako aj prenos kanálmi, čo vedie k zhoršeniu kvality. Popis spracovania videa je obsiahnutý v kapitole 3.

Jedným z dosiahnutých výsledkov tohto výskumu je mapovacia funkcia, ktorá nám určuje vhodný bitový tok pre jednotlivé typy scén s cieľom dosiahnuť požadovanú kvalitu. Model nám umožňuje zvoliť, či chceme určiť vhodný bitový tok na základe zadania požadovanej kvality vyjadrenej hodnotou objektívnej metriky (SSIM) alebo pomocou subjektívneho vnímania kvality danej videosekvencie vyjadrenej hodnotou MOS. Po zadaní kvalitatívnych parametrov scény (rozlíšenie, kompresný štandard), charakteristických parametrov scény (SI a TI) a požadovanej kvality (SSIM alebo MOS) nám mapovacia funkcia určí bitový tok, pri ktorom je dosiahnutá požadovaná kvalita videa.

Pri zvolení subjektívneho vyhodnotenia kvality a nastavení prahových hodnôt  $MOS=3$  alebo  $MOS=4$  ( $3 =$  najnižšia hranica prijateľného vnímania kvality a  $4 =$  minimálna hodnota, pri ktorej je kvalita hodnotená ako dobrá) je nám poskytnutý bitový tok, ktorý zodpovedá zadanej MOS kvalite na základe priemerného vyhodnotenia danej scény hodnotiacimi užívateľmi. Na základe dynamiky v scéne je potom možné v danej scéne prepínať bitový tok tak, aby kvalita bola postačujúca a aby spĺňala požadovaný štandard koncového používateľa.

Na stanovenie vhodného bitového toku bolo potrebné poznať vnímanie užívateľov a prepojiť ho s vyhodnotením kvality na základe výpočetných metód – objektívnych metrík. Pri množstve dát, prenášaných pri vysielaní, nie je možné v reálnom čase vyhodnocovať kvalitu, napriek tomu, že je to veľmi dôležitý faktor, ktorý vplýva na spokojnosť používateľa. Hodnotenie pomocou subjektívnych metrík je časovo veľmi náročné a tak vznikli objektívne metriky, ktoré vyhodnocujú pokles kvality prenášanej snímky s referenčnou.

Nový ponúkaný model rozlišuje videosekvenciu na základe obsahových informácií TI a SI. Bolo vybratých 12 rôznych typov videosekvencií, ktoré boli, zmenou kvalitatívnych parametrov, prekódované na požadované rozlíšenie, kompresný štandard s využitím širokej škály bitového toku. Vznikla tak databáza obsahujúca 384 videosekvencií. Tieto informácie boli počítane pre každú videosekvenciu vo vytvorenej databáze nakoľko sa nelíšia len pri zmene typu scény, ale aj pri zmene

kompresného štandardu, rozlíšenia alebo bitového toku. Toto sú parametre určujúce danú videosekvenciu. Tieto hodnoty, môžu byť vypočítané bez potreby referenčnej videosekvencie. Databázu najskôr tvorilo 8 typov videosekvencií z pohľadu dynamickosti. Z nich po prekódovaní vznikla databáza obsahujúca 256 videosekvencií, ktoré boli ohodnotené subjektívnou metrikou SSIM a objektívnou metrikou ACR. Databáza bola behom práce rozšírená o ďalšie 4 typy videosekvencií (sekvencie Ultra Video Group), z ktorých pribudlo ďalších 128 vzoriek videosekvencií, ktorých kvalita bola ohodnotená pomocou objektívnej metriky SSIM, čím vznikla širšia databáza pre potreby predikcie bitového toku na základe vyhodnotenia pomocou objektívnych metrik, ktoré sú v praxi využívané častejšie (menšia časová náročnosť na vyhodnotenie, vynechaný vplyv ľudského faktora, ktorý určité poškodenie videoscény nezaregistruje).

Ďalším dosiahnutým výsledkom práce je vyjadrenie vzťahu medzi objektívnym a subjektívnym ohodnotením kvality. Navrhnutý model tak môže pomocou mapovacej funkcie určiť požadovaný bitový tok na základe zadanej hodnoty požadovanej kvality oboma spôsobmi. Model taktiež predikuje vyjadrenie subjektívneho vnímania na základe zadaných kvalitatívnych parametrov, charakteristických parametrov videoscény a vyhodnotenej kvality objektívnou metrikou SSIM. Toto umožňuje poskytovateľovi vidieť ako koncový zákazník vníma kvalitu danej videosekvencie pri výsledku ohodnotenia kvality danej videosekvencie objektívnou metrikou SSIM. Model taktiež umožňuje predikciu kvality videoscény na základe charakteristických parametrov a bitového toku. Je možné zvoliť predikciu subjektívnej ako aj objektívnej metriky a získať tak MOS skóre alebo SSIM ohodnotenie videosekvencie.

Boli otestované rôznorodé simulácie, ktoré prebiehali pre jednotlivý typ rozlíšenia a kompresného štandardu samostatne a tiež v jednom toku. V práci je tak možné porovnať výsledky pre rôzne skupiny dát (napr. každý kompresný štandard má rozdielne výsledky, čím vektory vstupov tvoria 2 odlišné množiny). Z výsledkov je možné zhodnotiť, či má dané rozlíšenie alebo kompresný štandard väčší vplyv na konečné výsledky, ale aj adaptáciu modelu pri simulovaní rôznorodých dát v jednej simulácii.

Vytvorený model využíva doprednú viacvrstvovú neurónovú sieť implementovanú v prostredí Matlab. Využíva Backpropagation algoritmus. Model porovnal veľký počet rôznych topológií, kde sa menil počet skrytých vrstiev, počet neurónov v týchto vrstvách. Boli otestované rozličné aktivačné funkcie, tréningový algoritmus, algoritmus učenia. Ako najlepšie boli vyhodnotené topológie s 2 skrytými vrstvami, pri použití aktivačnej funkcie tansig, algoritmu Backpropagation implementovaného typom traingdx, čo dokazujú prezentované výsledky v práci. Pri vyššom počte vrstiev dochádzalo k častejšiemu pretrénovaniu siete. Výsledky navrhovaného modelu dosahujú vysokú úspešnosť a predikcie sa výrazne približujú k referenčným hodnotám.

Vytvorený model umožňuje predikciu:

- Bitového toku na základe typu videosekvencie (charakterizovanou SI a TI informáciami) a hodnotenia kvality danej scény objektívnou metrikou SSIM
- Bitového toku na základe typu videosekvencie a hodnotenia kvality danej scény pomocou stupnice MOS (hodnota dosiahnutá subjektívnou metrikou ACR)
- Subjektívnej kvality na základe typu videosekvencie, bitového toku, kvalitatívnych parametrov a objektívneho výsledku metrikou SSIM
- Subjektívnej kvality videosekvencie na základe informácií o type videosekvencie a bitového toku
- Objektívnej kvality na základe informácií o type videosekvencie a bitového toku

Ďalší vývoj sa bude sústrediť na rozšírenie vytvorenej databázy o nové vzorky, čím by sa zvýšila presnosť odhadu typu videoscény. Pre komplexnosť nástroja je možné rozšíriť model o ďalšie parametre, ktoré identifikujú konkrétnu videoscénu. Ďalším možným spôsobom obohatenia prezentovaného modelu je simulácia vplyvu pôsobenia sieťovej prevádzky na prenos.



## 11 Záver

Prezentovaná dizertačná práca sa zaoberá možnosťou využívania variabilného dátového toku pre prenos videa na základe časovej a priestorovej perceptuálnej informácie scény pri zachovaní určeného QoE. Základná myšlienka vychádza zo skutočnosti, že vnímanie kvality videa užívateľom je závislé na type scény. V rámci dizertačnej práce som vytvoril klasifikátor bitového toku využívajúci umelú inteligenciu a navrhol mapovaciu funkciu, ktorá dokáže nájsť optimálny bitový tok spĺňajúci požiadavky kvality pre daný typ videoscény, čo by umožnilo stálu, koncovým používateľom vyžadovanú kvalitu prenášaného videa. Výhoda nasadenia VBR spočíva v efektívnom využití sieťových prostriedkov pri konzistentnej kvalite videa. Dôvodom výskumu v tejto oblasti je narastajúce množstvo online služieb poskytujúcich videá vo vysokej kvalite. To si vyžaduje dostatočné prenosové pásmo pre poskytnutie čo najvyššej kvality. Odhadnutie kvality prenášaných audio-vizuálnych signálov môže výrazne prispieť k multimedialným službám a komunikačným sieťam. V skutočnosti je hodnotenie kvality digitálnych signálov jedným zo základných a náročných úloh v oblasti multimedialného spracovania a jeho praktických situácií, ako je proces hodnotenia, implementácie, optimalizácie kódovania a dekódovania, testovania a monitorovania. Navyše, hodnotenie audio a video kvality predstavuje kľúčovú úlohu pri multimedialných službách, algoritmov a systémov. Poskytovanie týchto služieb s dostatočným uspokojením zákazníka vyžaduje správny odhad kvality multimedialných služieb. Je potrebné zabezpečiť čo najvyššiu možnú kvalitu vysielaného obsahu, nakoľko uspokojenie zákazníka je jeden z hlavných cieľov poskytovania triple play služieb. Riešením tejto práce je vytvorenie modelu, ktorý by vedel určiť, pri akých kvalitatívnych nastaveniach, dosiahneme tohto cieľa. Je preto nevyhnutné monitorovať danú sieť, vyhodnotiť aktuálne zobrazovaný obsah a poznať nielen dosiahnuté výsledky kvality vypočítané objektívnou metrikou, ale tiež vnímanie kvality koncového užívateľa, pri danom výsledku objektívnej metriky. Prvým cieľom tejto práce bolo práve vyhodnotiť vzájomnú koreláciu ohodnotenia kvality videosekvencií objektívnou a subjektívnou metrikou. Prezentovaný model ponúka predikciu MOS na základe parametrov videosekvencie a jej ohodnotením kvality pomocou SSIM. Vytvorený model vyhodnotí vhodný bitový tok pre požadovanú kvalitu danej videosekvencie, definovanej pomocou charakteristických a kvalitatívnych parametrov. Prenášaná scéna môže niesť príznak o dynamickosti, prípadne dynamickosť scény môže byť prepočítavaná v reálnom čase, čo by samozrejme bolo časovo a aj výpočtovo náročné a prinášalo by to určité oneskorenie. Reálny prepočet dynamickosti videa by tým nebol vhodný na online prenosy, tie by mohli byť klasifikované už spomenutým príznakom. Reálny prepočet by bol prínosný pri videách typu „video on demand“, pri hrách, pri virtuálnej realite.

Prezentovaný model dosahuje veľmi dobré výsledky, každá simulácia je štatisticky verifikovaná s vysokou úspešnosťou predikcie jednotlivých simulovaných parametrov. Podobný model, ktorý by bol založený na vyhodnocovaní bitového toku na základe charakteristických vlastností scén vo forme hraničných informácií SI a TI; ponúkal predikciu vyhodnotenia kvality videosekvencií a tiež mapovania subjektívnej kvality pomocou objektívneho vyhodnotenia, je ojedinelý a prináša tak prínos v oblasti vyhodnocovania kvality. V práci je kompletne popísaná metodika vedúca k realizácii a splneniu cieľov dizertačnej práce. Kapitola 1 nám popisuje súčasný stav, kde je prezentovaná časť vedeckých prác, ktoré sa venujú tejto problematike. Na túto kapitolu nadväzuje kapitola popisujúca metodológiu, kde je zobrazený popis spracovania videa, hodnotenie kvality videa, princíp neurónovej siete, na ktorej je

založený vytvorený model. Hlavná časť tejto práce začína stanovením jej cieľov v kapitole 4. Pri pohľade na tieto ciele je možné potvrdiť, že všetky boli splnené.

Bola vytvorená databáza videosekvencií zabezpečujúca širokú variabilitu možných scenárov vyskytujúcich sa v bežnej praxi. Obsahovala videosekvencie kódované na základe rôznych kvalitatívnych parametrov s rôznorodým bitovým tokom. Tieto videosekvencie boli ohodnotené subjektívne, značným počtom pozorovateľov, s použitím metriky ACR a objektívnou metrikou SSIM. Vyjadrenie korelácie výsledkov v hodnotení kvality videa medzi objektívnymi a subjektívnymi metódami je popísané v kapitole 7, vytvorenie klasifikátora scény je popísané v kapitole 6, mapovacia funkcia, ktorá na základe typu scény, charakterizovanej pomocou SI a TI, určí vhodný bitový tok pre zadané ohodnotenia kvality formou MOS alebo SSIM, je popísaná v kapitole 8 a kapitola 9 pozostáva z verifikácie a validácie navrhnutého modelu, kde je porovnaná určená hodnota bitového toku s požadovanou, pomocou rôznych štatistických spôsobov. Kapitola 10 popisuje zhrnutie dosiahnutých výsledkov a naznačuje smer pokračovania výskumu v tejto oblasti. Predkladaná dizertačná práca predstavuje prínos pre odbornú komunitu vo svojej oblasti, vytvorený model je možné využiť v praxi a navrhnuté postupy implementovať v reálnom nasadení. Dizertačná práca podľa môjho názoru splnila svoj cieľ a priniesla nový prístup umožňujúci adaptívne rezervácie sieťových zdrojov podľa typu scény. Myšlienky boli overené na reálnej dátovej sade a návrh klasifikátora bitového toku i mapovacej funkcie som urobil s využitím posledného stavu poznania a pomocou nástrojov, ktoré odpovedali aktuálnym trendom.



## Referencie

- [1] S. Chikkerur, V. Sundaram, M. Reisslein and L. J. Karam, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," in *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 165-182, June 2011. doi: 10.1109/TBC.2011.2104671
- [2] Y. Changhoon and A.C. Bovik, " Evaluation of temporal variation of video quality in packet loss networks," in *Journal Image Communication archive, Volume 26 Issue 1, January, 2011, Pages 24-38*
- [3] N. Feamster and N. Balakrishnan, " Packet loss recovery for streaming video," In 12th international packet video workshop, Pittsburgh, PA
- [4] I. K. Kim, S. Lee, Y. Piao and J. Chen, "Coding efficiency comparison of new video coding standards: HEVC vs VP9 vs AVS2 video," *2014 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*, Chengdu, 2014, pp. 1-6. doi: 10.1109/ICMEW.2014.6890700
- [5] S. Moller and A. Raake, " Quality of experience: Advanced Concepts, Applications and Methods," *In Springer, ISBN: 9783319026817*
- [6] J. You, U. Reiter, M. M. Hannuksela, M. Gabbouj, and A. Perkis, "Perceptual-based quality assessment for audio\_Visual services: A survey," *in Signal Process., Image Commun., vol. 25, no. 7, pp. 482\_501, 2010*
- [7] G. Ghinea and S. Y. Chen, "Measuring quality of perception in distributed multimedia: Verbalizers vs. imagers," *in Comput. Human Behav., vol. 24, no. 4, pp. 1317–1329, Jul. 2008*
- [8] P. Le Callet, S. Möller, and A. Perkis, "Qualinet white paper on definitions of quality of experience," *Eur. Netw. Qual. Experience Multimedia Syst. Services (COST Action IC 1003)*, Lausanne, Switzerland, Tech. Rep. 1.2, Mar. 2013. Dostupné na :<https://hal.archives-ouvertes.fr/hal-00977812>
- [9] Doporučenie ITU-T P.10/G.100 Amendment 2: New definitions for inclusion in Recommendation ITU-T P.10/G.100—Vocabulary for performance and quality of service," Dostupné na : <https://www.itu.int/rec/T-REC-P.10/en>
- [10] O. Verscheure, P. Frossard, and M. Hamdi, "MPEG-2 Video Services over Packet Networks: Joint Effect of Encoding Rate and Data Loss on User-Oriented QoS," *in ACM NOSSDAV'98, Cambridge, UK, July 1998*
- [11] P. Frossard and O. Verscheure, "Joint source/FEC rate selection for quality-optimal MPEG-2 video delivery," in *IEEE Transactions on Image Processing*, vol. 10, no. 12, pp. 1815-1825, Dec 2001. doi: 10.1109/83.974566
- [12] J. Joskowicz, J. C. López-Ardao, M. A. González Ortega and C. L. García, "A Mathematical Model for Evaluating the Perceptual Quality of Video," *In: Mauthe A., Zeadally S., Cerqueira E., Curado M. (eds) Future Multimedia Networking. FMN 2009. Lecture Notes in Computer Science, vol 5630. Springer, Berlin, Heidelberg, [https://doi.org/10.1007/978-3-642-02472-6\\_15](https://doi.org/10.1007/978-3-642-02472-6_15)*
- [13] Doporučenie ITU-T G.1070, "Opinion model for video-telephony applications," *Dostupné na <https://www.itu.int/rec/T-REC-G.1070/en>*
- [14] F. Yang, S. Wan, Q. Xie and H. R. Wu, "No-Reference Quality Assessment for Networked Video via Primary Analysis of Bit Stream," in *IEEE Transactions on Circuits and Systems for*

*Video Technology*, vol. 20, no. 11, pp. 1544-1554, Nov. 2010.

doi: 10.1109/TCSVT.2010.2087433

- [15] P. d. I. C. Ramos, F. G. Vidal and R. P. Leal, "Perceived Video Quality Estimation from Spatial and Temporal Information Contents and Network Performance Parameters in IPTV," *2010 Fifth International Conference on Digital Telecommunications*, Athens, TBD, Greece, 2010, pp. 128-131. doi: 10.1109/ICDT.2010.31
- [16] P. Romaniak, L. Janowski, M. Leszczuk and Z. Papir, "Perceptual quality assessment for H.264/AVC compression," *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV, 2012, pp. 597-602. doi: 10.1109/CCNC.2012.6181021
- [17] S. Bech and N. Zacharov, "Perceptual Audio Evaluation\_Theory, Method and Application," *In Book Perceptual Audio Evaluation\_Theory, Method and Application, November 2006 New York, NY, USA: Wiley, 2006. 10.1002/9780470869253.ch1*
- [18] W. Lin and C.-C. Jay Kuo, "Perceptual visual quality metrics: A survey," *in Journal of Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297\_312, 2011
- [19] U. Engelke and H.-J. Zepernick, "Perceptual-based quality metrics for image and video services: A survey," *in Proc. EuroNGI Conf. Next Gen. Int. Netw., May 2007, pp. 190\_197*
- [20] S. Möller and A. Raake, "Quality of Experience: Advanced Concepts, Applications and Methods". Berlin, Germany: Springer, 2014
- [21] A. N. Moldovan, I. Ghergulescu and C. H. Muntean, "VQAMap: A Novel Mechanism for Mapping Objective Video Quality Metrics to Subjective MOS Scale," *in IEEE Transactions on Broadcasting*, vol. 62, no. 3, pp. 610-627, Sept. 2016. doi: 10.1109/TBC.2016.2570002
- [22] L. Janowski and P. Romaniak, "QoE as a Function of Frame Rate and Resolution Changes," *In: Zeadally S., Cerqueira E., Curado M., Leszczuk M. (eds) Future Multimedia Networking. FMN 2010. Lecture Notes in Computer Science, vol 6157. Springer, Berlin, Heidelberg, doi. 978-3-642-13789-1*
- [23] The Moving Picture Experts Group. MPEG Standards. Dostupné na: <http://mpeg.chiariglione.org/standards>
- [24] M. A. Layek *et al.*, "Performance analysis of H.264, H.265, VP9 and AV1 video encoders," *2017 19th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Seoul, 2017, pp. 322-325, doi: 10.1109/APNOMS.2017.8094162
- [25] Odporučenie ITU-T H.264. H.264:Advanced video coding for generic audiovisual services. Dostupné na: <http://www.itu.int/rec/T-REC-H.264-201602-S>, 2016
- [26] Wu, H., Claypool, M. and Kinicki, "R. Guidelines for Selecting Practical MPEG Group of Pictures," *Dostupné na: http://digitalcommons.wpi.edu/computerscience-pubs/168*
- [27] Didier J. Le Gall, "MPEG Video Compression algorithm," *in Signal Processing: Image Communication, Volume 4, Issue 2, April 1992, Pages 129-140, https://doi.org/10.1016/0923-5965(92)90019-C*
- [28] Ghanbari, M., Crawford, D., Fleury, M., Khan, E., Woods, J., Lu, H. and Razavi, R., "Future Performance of Video Codecs," *Dostupné na: https://www.researchgate.net/publication/250298818\_Title\_Future\_Performance\_of\_Video\_Codecs*
- [29] E. Dumić, M. Mustra, S. Grgić and G. Gvozden, "Image quality of 4:2:2 and 4:2:0 chroma subsampling formats," *2009 International Symposium ELMAR*, Zadar, 2009, pp. 19-24.

- [30] M. Leszczuk, L. Janowski and M. Barkowsky, "Freely available large-scale video quality assessment database in Full-HD resolution with H.264 coding," *2013 IEEE Globecom Workshops (GC Wkshps)*, Atlanta, GA, 2013, pp. 1162-1167.  
doi: 10.1109/GLOCOMW.2013.6825150
- [31] J. Frnda, "Nový způsob hodnocení kvality videa," Dostupné na: <https://theses.cz/id/kk48y1/>.  
Disertační práce. Vysoká škola báňská - Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky
- [32] Z. Wang, L. Lu and A. C. Bovik, "Video quality assessment using structural distortion measurement," *Proceedings. International Conference on Image Processing*, 2002, pp. III-65-III-68 vol.3. doi: 10.1109/ICIP.2002.1038904
- [33] Doporučenie ITU-T P.910, "Subjective video quality assessment methods for multimedia applications," Dostupné na: <https://www.itu.int/rec/T-REC-P.910-200804-I/en>, 2008
- [34] Y. WANG, "Survey of Objective Video Quality Measurement," dostupné na: <ftp://ftp.cs.wpi.edu/pub/techreports/pdf/06-02.pdf>
- [35] Zhou Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," in *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, April 2004, doi: 10.1109/TIP.2003.819861
- [36] H. R. WU and K. R. RAO, "Digital Video Image Quality and Perceptual Coding" Taylor and Francis Group LLC, 2006, ISBN 0-8247-2777-0
- [37] S. Winkler and P. Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics," in *IEEE Transactions on Broadcasting*, vol. 54, no. 3, pp. 660-668, Sept. 2008. doi: 10.1109/TBC.2008.2000733
- [38] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," in *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430-444, Feb. 2006. doi: 10.1109/TIP.2005.859378
- [39] Objective metrics for Video Quality Evaluations, dostupné na: [http://compression.ru/video/quality\\_measure/info\\_en.html](http://compression.ru/video/quality_measure/info_en.html)
- [40] S. T. McCarthy, "Ultra HD: Bandwidth Planning and Verification for 4K," Dostupné na: <http://www.arris.com/solutions/ultrahd/>
- [41] J. Rozhon, "Model pro prediktivní určování kvality hovoru jako kritéria pro hodnocení výkonnosti VOIP řešení," Dostupné na: <https://theses.cz/id/0g1xmg/>. Disertační práce. Vysoká škola báňská - Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky. Vedoucí práce Miroslav Vozňák
- [42] H. Demuth and M. Beale, "Neural Network Toolbox User's Guide," dostupné na: [http://www.image.ece.ntua.gr/courses\\_static/nn/matlab/nnet.pdf](http://www.image.ece.ntua.gr/courses_static/nn/matlab/nnet.pdf)
- [43] R. Rojas, "Neural Networks, A systematic Introduction," ISBN: 9783642610684, doi:10.1007/978-3-642-61068-4, dostupné na: <https://page.mi.fu-berlin.de/rojas/neural/neuron.pdf>
- [44] SJTU Media Lab: Databáza testovacích sekvencií. Dostupné na: <http://medialab.sjtu.edu.cn/web4k/index.html>
- [45] P. Fröhlich, S. Egger, R. Schatz, M. Muhlegger, "QoE in 10 seconds: Are short video clip lengths sufficient for Quality of Experience assessment?," in *Proceedings of Quality of Multimedia Experience (QoMEX)*, 2012, pp. 242-247

- [46] Ultra Video Group: Databáza testovacích sekvencií. Dostupné na: <http://ultravideo.cs.tut.fi/#testsequences>
- [47] FFmpeg. A complete, cross-platform solution to record, convert and stream audio and video. Dostupné na: <https://www.ffmpeg.org/>
- [48] MitsuTool, Video Quality Indicators. Dostupné na: <http://vq.kt.agh.edu.pl/metrics.html>
- [49] M. Uhrina, J. Hlubik and M. Vaculik, "Correlation between objective and subjective methods used for video quality evaluation," *2012 ELEKTRO*, Rajeck Teplice, 2012, pp. 103-108. doi: 10.1109/ELEKTRO.2012.6225581
- [50] M. Uhrina, J. Bienik and M. Vaculik, "Procedure for mapping of objective video quality metrics to subjective MOS scale of VP 9 compression standard for Full HD resolution," *2016 International Conference on Information and Digital Technologies (IDT)*, Rzeszow, 2016, pp. 288-294, doi: 10.1109/DT.2016.7557188
- [51] MSU Video Quality Measurement Tool. Program for objective video quality assessment. Dostupné na: [http://www.compression.ru/video/quality\\_measure/video\\_measurement\\_tool.html](http://www.compression.ru/video/quality_measure/video_measurement_tool.html)
- [52] Doporučenie ITU-R BT.500-12, "Methodology for the subjective assessment of the quality of television pictures," Dostupné na: <https://www.itu.int/rec/R-REC-BT.500-12-200909-S/en>
- [53] J. Frnda, J. Rozhon, M. Voznak, J. C-W. Lin, T-Y. Wu, "A new tool for quality of multimedia estimation based on network behaviour," *In ELSEVIER Perspectives in Science, Volume 7, Pages 87-94, 2016*
- [54] D. Valderrama and N. Gómez, "Nonintrusive Method Based on Neural Networks for Video Quality of Experience Assessment," *in Advances in Multimedia. 2016. 1-17. 10.1155/2016/1730814. Advances in Multimedia, Volume 2016*
- [55] J. Rozhon, M. Voznak, F. Rezac and J. Slachta, "Utilizing the Neural Networks for Speech Quality Estimation Based on the Network Characteristics," *in SPRINGER : Lecture Notes in Electrical Engineering, Vol. 371, pp. 99-109, 2016. Notes in Electrical Engineering, Vol. 371, pp. 99-109, 2016*

## Citované príspevky autora

- [sev01] M. Uhrina, J. Frnda, L. Sevcik and M. Vaculik, "Impact of H.264/AVC and H.265/HEVC compression standards on the video quality for 4K resolution, " in *Advances in Electrical and Electronic Engineering*. 12. 10.15598/aeee.v12i4.1216
- [sev02] L. Sevcik, D. Uhrin, J. Frnda, M. Voznak, Homer Toral-Cruz, M. Mikulec, Sergej Jakovlev, "Encryption for confidentiality of the network and influence of this to the quality of streaming video through network", Proc. SPIE 9497, Mobile Multimedia/Image Processing, Security, and Applications 2015, 94970A (21 May 2015); doi: 10.1117/12.2177555; <https://doi.org/10.1117/12.2177555>
- [sev03] L. Sevcik, D. Uhrin, J. Frnda, M. Uhrina, Z. Chmelikova and M. Voznak, "The impact of encryption on video transmission in IP network," *2014 22nd Telecommunications Forum Telfor (TELFOR)*, Belgrade, 2014, pp. 123-126, doi: 10.1109/TELFOR.2014.7034372
- [sev04] L. Sevcik, K. Tomala, J. Frnda and M. Voznak, "QoS of triple play services in LTE networks," In: *Pan JS., Snasel V., Corchado E., Abraham A., Wang SL. (eds) Intelligent Data analysis and its Applications, Volume II. Advances in Intelligent Systems and Computing, vol 298. Springer, Cham, [https://doi.org/10.1007/978-3-319-07773-4\\_3](https://doi.org/10.1007/978-3-319-07773-4_3)*
- [sev05] L. Sevcik, K. Tomala, J. Frnda and M. Voznak, "QoS of multimedia services in 4G networks," *2015 38th International Conference on Telecommunications and Signal Processing (TSP)*, Prague, 2015, pp. 1-5, doi: 10.1109/TSP.2015.7296450
- [sev06] L. Sevcik, J. Frnda, M. Voznak, Degrading effect analysis, packet loss and out of order data on various tips and video resolution, In *Proceedings of the 2014 Networking and Electronic Commerce Research Conference*, Trieste, Italy, ISBN 978-0-9820958-0-5, pp. 130-138. M.
- [sev07] L. Sevcik, M. Voznak and J. Frnda, "QoE prediction model for multimedia services in IP network applying queuing policy," *International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS 2014)*, Monterey, CA, 2014, pp. 593-598. doi: 10.1109/SPECTS.2014.6879998
- [sev08] J. Frnda, M. Voznak, L. Sevcik, "Network Performance QoS Prediction," In *Advances in Intelligent Systems and Computing*, Vol. 297, 2014, pp. 165-174, DOI 10.1007/978-3-319-07776-5\_18
- [sev09] J. Frnda, M. Voznak, L. Sevcik, P. Fazio, Prediction Model of Triple play services for QoS Assessment in IP Based Networks, In *Journal of Networks*, Vol. 10, No. 4, pp. 232-239, April 2015, DOI 10.4304/jnw.10.4.232-239
- [sev10] J. Frnda, L. Sevcik, M. Uhrina, M. Voznak, Network Degradation Effects on Different Codec Types and Characteristics of Video Streaming, In *Journal Advances in Electrical and Electronic Engineering*, Vol. 12, Issue 4, 2014, pp. 377-383, ISSN 1336-1376
- [sev11] M. Uhrina, L. Sevcik, J. Frnda and M. Vaculik, "Impact of H.265 and VP9 compression standards on the video quality for 4K resolution," *2014 22nd Telecommunications Forum Telfor (TELFOR)*, Belgrade, 2014, pp. 905-908, doi: 10.1109/TELFOR.2014.7034551
- [sev12] J. Frnda, M. Voznak and L. Sevcik, "Impact of packet loss and delay variation on the quality of real-time video streaming," In *Telecommunication Systems* , Volume 62, Issue 2, 2016, pp. 265-275. DOI 10.1007/s11235-015-0037-2.

[sev13] M. Uhrina, J. Frnda, L. Sevcik and M. Vaculík, "The impact of GOP pattern and packet loss on the video quality of H. 264/AVC compression standard, " *In Advances in Neural Networks, Fuzzy Systems and Artificial Intelligence*, ISSN: 1790-5109, ISBN: 978-960-474-379-7, pp 150-157, dostupné na: <http://www.wseas.us/e-library/conferences/2014/Gdansk/FUNAI/FUNAI-24.pdf>

## Publikačná a projektová činnosť autora

### Účasť v projektoch

- Špecifický výskum SGS FEI VŠB TOU projekt SP2013/94 s názvom: *Výskum vplyvu okolitého prostredia na vlastnosti rádiového kanálu a vývoj nových prístupov k hodnoteniu kvality služieb (QoS) multimédií v sieťach 4G.*
- Špecifický výskum SGS FEI VŠB TOU projekt SP2014/72 s názvom: *Výskum atmosférických vplyvov na prenosy v rádiovom kanáli.*
- Špecifický výskum SGS FEI VŠB TOU projekt SP2015/82 s názvom: *Vyťažovanie informácií z komunikačných sietí, ich modelovanie a simulácia.*
- Špecifický výskum SGS FEI VŠB TUO projekt SP2016/170 s názvom: *Vyťažovanie informácií z komunikačných sietí, ich modelovanie a simulácia.*
- Špecifický výskum SGS FEI VŠB TUO projekt SP2017/174 s názvom: *Siete a ich bezpečnosť, modelovanie, simulácia, vyťažovanie znalosti a komunikačné technológie pre chytré mestá.*
- The European Union's EuropeAid/131928/C/SER/Multi Programme for EaP countries. *5th Workshop – Quality of Service and Customer protection*, ENPI / 2012 / 307-572 Kyiv, Ukraine 2014
- Projekt na podporu aktivít VŠB-TUO s Čínou za podpory Moravskosliezského kraja s názvom: *Výskum v oblasti riadenia prístupu k službám v konvergovaných sieťach NGN pomocou prediktívnych metód a efektívne vyhľadávanie v databázových štruktúrach.* (dve krátkodobé pracovné stáže na Harabin Institute of Technology Shenzhen v rokoch 2015 a 2016)

### Publikačná činnosť

(Evaluácia publikačnej činnosti behom doktorského štúdia)

- Web of Knowledge
  - 10 záznamov: 2xčasopis a 8xkonferencia
- Elsevier Scopus
  - 14 záznamov: 4xčasopis a 10xkonferencia
- IEEE-Xplore
  - 6 záznamov

### Ostatné výsledky dosiahnuté behom doktorandského štúdia a indexované na WoS, či SCOPUS

- J. Lin, L. Yang, P. Fournier-Viger, J. Frnda, L. Sevcik, M. Voznak, *An Evolutionary Algorithm to Mine High-Utility Itemsets*, In Journal Advances in Electrical and Electronic Engineering, Vol. 13, Issue 4, 2015, pp. 392-398, ISSN 1336-1376. DOI: 10.15598/aeec.v13i4.1474
- H. S. Nguyen, D. T. Do, M. Voznak, L. Sevcik, *A Straightforward Method to Evaluate the Energy Aware Two-way Relaying Networks Under Effect of Co-channel Interference*. XI International Symposium on Telecommunications – BIHTEL 2016, Sarajevo, October 24-26, 2016.
- H. S. Nguyen, D. T. Do, M. Voznak, L. Sevcik, *Two-Way Relay Networks with Energy Harvesting and Information Transfer: Throughput Performance with Distance Allocation*. XI International Symposium on Telecommunications – BIHTEL 2016, Sarajevo, October 24-26, 2016.
- M. Voznak, M. Prokes, L. Sevcik, J. Frnda, H. Toral-Cruz, S. Jakovlev, P. Fazio, M. Mehic, M. Mikulec, *Vulnerabilities in GSM technology and feasibility of selected attacks*, In Proc. SPIE 9456, Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security, Defense, and Law Enforcement XIV, 94560T, Baltimore, Maryland, USA. (May 23, 2015) doi: 10.1117/12.2177111.

### Ostatné výsledky dosiahnuté behom doktorandského štúdia

- P. Fournier-Viger, J. C-W. Lin, Q-H. Duong, T-L. Dam, L. Sevcik, D. Uhrin, M. Voznak, *PFPM: Discovering Periodic Frequent Patterns with Novel Periodicity Measures*, Proceedings of the 2nd Czech-China Scientific Conference 2016, pp. 27-38, Feb. 2017, DOI: 10.5772/66780.
- D. Uhrin, P. Partila, J. Frnda, L. Sevcik, M. Voznak, J. C-W. Lin, *Design of Emotion Recognition System*, Proceedings of the 2nd Czech-China Scientific Conference 2016, pp. 53-63, Feb. 2017, DOI: 10.5772/66782.
- J.C.-W. Lin, W. Gan, P. Fournier-Viger, L Yang, Q. Liu, J. Frnda, L. Sevcik, M. Voznak, *High Utility-Itemset Mining and Privacy-Preserving Utility Mining*, In ELSEVIER : Perspectives in Science , Vol. 7, March 2016, pp. 74-80. DOI: 10.1016/j.pisc.2015.11.013
- T.-Y. Wu, J.C.-W. Lin, C.-M. Chen, Y.-M. Tseng, J. Frnda, L. Sevcik, M. Voznak, *A brief review of revocable ID-based public key cryptosystem*, In ELSEVIER : Perspectives in Science , Vol. 7, March 2016, pp. 81-86. DOI: 10.1016/j.pisc.2015.11.014
- L. Sevcik, M. Voznak, M. Prokes, P. Fazio, H. Total-Cruz, *Study of security issues in GSM network and their practical demonstration*, In Proceedings of the 2014 Networking and Electronic Commerce Research Conference, Trieste, Italy, ISBN 978-0-9820958-0-5, pp. 297-305.



## Príloha A Databáza videosekvencií

Výsledná databáza je rozdelená do 2 tabuliek. Tabuľka 17 obsahuje názov videosekvencie, definovanie scény (SI a TI informácie), objektívne vyhodnotenie kvality v podobe SSIM, subjektívne ohodnotenie zapísané ako priemer MOS skóre, dosiahnuté hodnotením užívateľov, smerodajná odchýlka  $s$  a variačný koeficient  $v$ . Ako vidno z tabuliek, dosiahnuté dáta sú relevantné, vypočítaný priemer má reprezentatívnu funkciu (keďže variačný koeficient nepresiahol 50%). Hodnoty nad 40% sa vyskytujú u 7 videosekvencií (maximálna hodnota 41,15). Tieto vyššie hodnoty sa vyskytujú u videosekvencií, ktoré boli kódované s bitovým tokom 1 Mbps pri rozlíšení UHD. Obrázok 29 zobrazuje popis názvoslovia videosekvencií.



Obrázok 29 Popis názvoslovia videosekvencií

Tabuľka 17 Databáza videosekvencií SJTU Media Lab

Názov videosekvencie	SI [-]	TI [-]	SSIM [-]	MOS [-]	$s$ [-]	$v$ [%]
BundFHD1	76,585	9,873	0,926	1,700	0,651	38,309%
BundFHD3	80,803	9,490	0,972	2,933	0,828	28,216%
BundFHD5	81,382	9,540	0,981	3,600	0,621	17,263%
BundFHD8	81,772	9,587	0,986	3,870	0,681	17,624%
BundFHD10	81,895	9,606	0,987	4,233	0,679	16,037%
BundFHD12	81,978	9,605	0,988	4,330	0,547	12,616%
BundFHD15	82,052	9,619	0,988	4,600	0,675	14,667%
BundFHD20	82,122	9,632	0,989	4,870	0,346	7,104%
CampfireFHD1	34,843	34,701	0,824	1,430	0,568	39,650%
CampfireFHD3	47,564	34,998	0,884	3,070	0,640	20,859%
CampfireFHD5	53,268	35,117	0,915	3,830	0,986	25,709%
CampfireFHD8	57,166	35,210	0,937	4,170	0,592	14,210%
CampfireFHD10	58,592	35,226	0,945	4,230	0,728	17,195%
CampfireFHD12	59,373	35,239	0,951	4,800	0,407	8,476%
CampfireFHD15	60,079	35,253	0,957	4,800	0,407	8,476%
CampfireFHD20	60,860	35,271	0,964	4,900	0,305	6,227%
ConstructionFHD1	54,727	7,221	0,891	2,100	0,759	36,136%
ConstructionFHD3	60,385	7,042	0,958	3,200	0,961	30,041%
ConstructionFHD5	61,315	7,240	0,971	4,067	0,740	18,189%
ConstructionFHD8	61,681	7,318	0,977	4,330	0,606	13,996%
ConstructionFHD10	61,817	7,380	0,979	4,433	0,679	15,314%
ConstructionFHD12	61,896	7,400	0,980	4,500	0,509	11,301%
ConstructionFHD15	61,929	7,415	0,982	4,733	0,583	12,323%
ConstructionFHD20	61,924	7,422	0,984	4,830	0,379	7,842%

FountainsFHD1	48,192	8,710	0,762	2,167	0,747	34,460%
FountainsFHD3	58,112	9,553	0,832	2,767	0,504	18,217%
FountainsFHD5	60,783	9,899	0,862	3,567	0,679	19,035%
FountainsFHD8	62,722	10,114	0,891	3,670	0,802	21,881%
FountainsFHD10	63,719	10,256	0,905	4,100	0,607	14,815%
FountainsFHD12	64,448	10,310	0,916	4,430	0,679	15,314%
FountainsFHD15	64,998	10,388	0,928	4,600	0,675	14,667%
FountainsFHD20	65,436	10,498	0,943	4,770	0,430	9,025%
MarathonFHD1	54,468	16,003	0,754	1,933	0,450	23,264%
MarathonFHD3	68,685	16,837	0,858	2,667	0,547	20,500%
MarathonFHD5	72,291	17,334	0,894	3,367	0,718	21,339%
MarathonFHD8	74,374	17,731	0,920	3,730	0,640	17,134%
MarathonFHD10	75,164	17,864	0,930	4,300	0,651	15,146%
MarathonFHD12	75,532	17,916	0,937	4,370	0,615	14,083%
MarathonFHD15	76,100	18,006	0,944	4,600	0,724	15,739%
MarathonFHD20	76,574	18,065	0,952	4,900	0,305	6,227%
RunnersFHD1	75,173	24,915	0,649	2,067	0,583	28,224%
RunnersFHD3	97,170	24,695	0,845	2,533	0,730	28,828%
RunnersFHD5	101,097	24,642	0,903	3,400	0,855	25,147%
RunnersFHD8	102,842	24,791	0,939	3,830	0,874	22,807%
RunnersFHD10	103,288	24,795	0,952	4,133	0,681	16,487%
RunnersFHD12	103,514	24,840	0,960	4,470	0,730	16,350%
RunnersFHD15	103,611	24,884	0,968	4,733	0,450	9,502%
RunnersFHD20	103,712	24,920	0,975	4,970	0,183	3,676%
TallFHD1	86,894	10,919	0,776	2,067	0,785	37,980%
TallFHD3	99,942	11,870	0,914	3,167	0,648	20,454%
TallFHD5	102,277	12,304	0,951	4,000	0,871	21,775%
TallFHD8	103,436	12,538	0,971	4,400	0,498	11,324%
TallFHD10	103,662	12,516	0,977	4,533	0,629	13,871%
TallFHD12	103,856	12,516	0,980	4,600	0,563	12,244%
TallFHD15	103,947	12,602	0,983	4,600	0,563	12,244%
TallFHD20	103,945	12,685	0,986	4,870	0,346	7,104%
WoodFHD1	117,671	34,651	0,743	1,700	0,640	36,905%
WoodFHD3	137,088	37,615	0,903	2,933	0,551	19,674%
WoodFHD5	139,253	38,111	0,944	3,600	1,020	26,607%
WoodFHD8	141,083	38,433	0,965	3,870	0,750	17,435%
WoodFHD10	141,453	38,554	0,972	4,233	0,711	16,411%
WoodFHD12	141,364	38,639	0,976	4,330	0,571	12,603%
WoodFHD15	141,125	38,683	0,980	4,600	0,484	10,088%
WoodFHD20	141,010	38,745	0,984	4,870	0,183	3,676%
BundXFHD1	76,521	9,064	0,940	2,333	0,711	30,478%
BundXFHD3	80,076	9,374	0,976	3,167	0,699	22,072%
BundXFHD5	80,959	9,479	0,984	3,933	0,785	19,956%
BundXFHD8	81,484	9,530	0,988	4,230	0,817	19,304%

BundXFHD10	81,594	9,550	0,989	4,400	0,814	18,493%
BundXFHD12	81,691	9,555	0,989	4,570	0,626	13,709%
BundXFHD15	81,791	9,566	0,990	4,630	0,556	12,001%
BundXFHD20	81,917	9,572	0,991	4,730	0,450	9,502%
CampfireXFHD1	40,919	34,524	0,856	2,167	0,834	38,488%
CampfireXFHD3	49,599	34,911	0,902	3,067	0,828	26,990%
CampfireXFHD5	52,774	34,980	0,927	3,833	0,791	20,647%
CampfireXFHD8	55,210	35,066	0,945	4,230	0,728	17,195%
CampfireXFHD10	56,323	35,134	0,952	4,367	0,615	14,083%
CampfireXFHD12	57,300	35,157	0,956	4,560	0,679	14,867%
CampfireXFHD15	58,040	35,176	0,962	4,767	0,430	9,025%
CampfireXFHD20	58,897	35,197	0,967	4,767	0,430	9,025%
ConstructionXFHD1	56,869	6,949	0,921	2,433	0,728	29,915%
ConstructionXFHD3	60,139	7,131	0,968	3,267	0,740	22,643%
ConstructionXFHD5	60,898	7,251	0,977	3,967	0,850	21,436%
ConstructionXFHD8	61,290	7,283	0,981	4,060	0,691	17,004%
ConstructionXFHD10	61,473	7,324	0,983	4,300	0,702	16,331%
ConstructionXFHD12	61,552	7,332	0,984	4,730	0,450	9,502%
ConstructionXFHD15	61,661	7,353	0,985	4,733	0,521	11,004%
ConstructionXFHD20	61,739	7,376	0,986	4,800	0,407	8,476%
FountainsXFHD1	51,528	8,696	0,786	2,133	0,860	40,330%
FountainsXFHD3	58,184	9,185	0,841	3,033	0,615	20,273%
FountainsXFHD5	60,246	9,433	0,869	3,567	0,817	22,912%
FountainsXFHD8	62,504	9,668	0,896	4,170	0,592	14,210%
FountainsXFHD10	63,246	9,753	0,908	4,267	0,583	13,671%
FountainsXFHD12	63,802	9,829	0,919	4,430	0,568	12,819%
FountainsXFHD15	64,334	9,932	0,931	4,567	0,626	13,709%
FountainsXFHD20	64,894	10,047	0,945	4,670	0,479	10,274%
MarathonXFHD1	60,342	16,001	0,790	2,000	0,743	37,139%
MarathonXFHD3	69,131	16,617	0,871	3,000	0,695	23,160%
MarathonXFHD5	72,034	17,072	0,901	3,500	0,777	22,195%
MarathonXFHD8	73,596	17,391	0,924	4,060	0,521	12,807%
MarathonXFHD10	74,353	17,530	0,932	4,133	0,776	18,776%
MarathonXFHD12	74,708	17,621	0,939	4,700	0,466	9,917%
MarathonXFHD15	75,153	17,701	0,946	4,767	0,430	9,025%
MarathonXFHD20	75,624	17,834	0,954	4,800	0,484	10,088%
RunnersXFHD1	81,325	24,711	0,733	2,233	0,774	34,650%
RunnersXFHD3	95,386	24,456	0,874	3,267	0,640	19,582%
RunnersXFHD5	99,233	24,613	0,916	3,567	0,626	17,553%
RunnersXFHD8	101,137	24,682	0,945	3,830	0,834	21,754%
RunnersXFHD10	101,880	24,707	0,955	4,433	0,679	15,314%
RunnersXFHD12	102,426	24,743	0,961	4,670	0,479	10,274%
RunnersXFHD15	102,911	24,752	0,969	4,833	0,379	7,842%
RunnersXFHD20	103,300	24,807	0,976	4,870	0,346	7,104%

TalIXFHD1	93,816	10,398	0,851	2,333	0,844	36,179%
TalIXFHD3	100,171	12,031	0,940	3,267	0,907	27,771%
TalIXFHD5	101,611	12,277	0,964	4,000	0,788	19,696%
TalIXFHD8	102,600	12,527	0,977	4,400	0,675	15,333%
TalIXFHD10	103,028	12,570	0,981	4,533	0,681	15,032%
TalIXFHD12	103,276	12,604	0,984	4,700	0,466	9,917%
TalIXFHD15	103,552	12,665	0,986	4,767	0,430	9,025%
TalIXFHD20	103,804	12,727	0,989	4,830	0,379	7,842%
WoodXFHD1	120,465	35,311	0,807	2,333	0,730	34,233%
WoodXFHD3	134,439	37,593	0,921	3,167	0,776	22,387%
WoodXFHD5	137,214	38,126	0,951	3,933	0,712	18,256%
WoodXFHD8	138,710	38,421	0,968	4,230	0,791	18,996%
WoodXFHD10	139,148	38,515	0,973	4,400	0,681	15,256%
WoodXFHD12	139,471	38,553	0,977	4,570	0,466	9,917%
WoodXFHD15	139,819	38,609	0,981	4,630	0,466	9,917%
WoodXFHD20	139,957	38,653	0,984	4,730	0,379	7,842%
BundUHD1	42,954	10,692	0,855	1,370	0,556	40,687%
BundUHD3	51,640	9,898	0,922	2,633	0,718	27,281%
BundUHD5	53,611	9,754	0,938	3,467	0,776	22,387%
BundUHD8	54,584	9,712	0,948	3,600	0,563	15,646%
BundUHD10	54,946	9,699	0,952	3,867	0,819	21,189%
BundUHD12	55,151	9,731	0,954	4,100	0,548	13,359%
BundUHD15	55,386	9,757	0,956	4,133	0,860	20,815%
BundUHD20	55,653	9,786	0,956	4,870	0,346	7,104%
CampfireUHD1	23,763	32,480	0,764	1,133	0,434	38,309%
CampfireUHD3	24,610	33,733	0,795	1,933	0,583	30,170%
CampfireUHD5	28,482	34,470	0,814	2,633	0,615	23,353%
CampfireUHD8	35,069	34,943	0,836	3,230	0,626	19,363%
CampfireUHD10	37,809	35,056	0,848	3,600	0,894	24,845%
CampfireUHD12	39,777	35,138	0,857	4,160	0,648	15,545%
CampfireUHD15	41,701	35,166	0,867	4,267	0,583	13,671%
CampfireUHD20	43,852	35,241	0,877	4,900	0,305	6,227%
ConstructionUHD1	28,080	9,514	0,794	1,700	0,651	38,309%
ConstructionUHD3	34,146	6,894	0,867	3,133	0,730	23,307%
ConstructionUHD5	35,947	7,126	0,890	3,767	0,626	16,621%
ConstructionUHD8	37,227	7,354	0,904	3,870	0,571	14,776%
ConstructionUHD10	37,835	7,484	0,910	3,967	0,615	15,503%
ConstructionUHD12	38,274	7,570	0,913	4,170	0,461	11,067%
ConstructionUHD15	38,731	7,622	0,915	4,467	0,571	12,791%
ConstructionUHD20	39,407	7,744	0,916	4,900	0,305	6,227%
FountainsUHD1	23,342	10,117	0,730	1,300	0,535	41,153%
FountainsUHD3	36,132	9,405	0,775	2,070	0,640	30,952%
FountainsUHD5	39,827	9,697	0,807	2,900	0,759	26,167%
FountainsUHD8	43,568	10,218	0,831	2,970	0,615	20,729%

FountainsUHD10	44,660	10,388	0,842	3,433	0,504	14,680%
FountainsUHD12	45,635	10,505	0,852	3,670	0,661	18,024%
FountainsUHD15	46,780	10,636	0,864	3,867	0,900	23,264%
FountainsUHD20	48,098	10,793	0,879	4,700	0,535	11,383%
MarathonUHD1	27,961	15,361	0,674	1,067	0,254	23,785%
MarathonUHD3	38,057	17,014	0,726	1,667	0,606	36,389%
MarathonUHD5	41,477	17,084	0,765	2,533	0,937	36,991%
MarathonUHD8	44,029	17,513	0,795	3,200	0,847	26,466%
MarathonUHD10	45,080	17,734	0,806	3,967	0,964	24,310%
MarathonUHD12	45,924	17,910	0,815	3,970	0,615	15,503%
MarathonUHD15	46,697	18,047	0,824	4,333	0,758	17,495%
MarathonUHD20	47,774	18,187	0,834	4,900	0,305	6,227%
RunnersUHD1	34,347	23,453	0,570	1,230	0,504	40,865%
RunnersUHD3	51,205	24,756	0,694	2,133	0,571	26,782%
RunnersUHD5	58,792	24,567	0,770	3,133	0,819	26,148%
RunnersUHD8	63,489	24,691	0,826	3,270	0,640	19,582%
RunnersUHD10	65,545	24,778	0,848	3,767	0,626	16,621%
RunnersUHD12	66,568	24,798	0,863	4,100	0,662	16,141%
RunnersUHD15	67,850	24,875	0,879	4,367	0,615	14,083%
RunnersUHD20	69,388	24,940	0,895	4,930	0,379	7,842%
TallUHD1	42,934	10,964	0,656	1,900	0,556	40,687%
TallUHD3	64,933	12,228	0,800	3,033	0,728	29,915%
TallUHD5	72,064	13,267	0,859	3,367	0,925	28,899%
TallUHD8	75,999	14,061	0,900	3,970	0,711	21,335%
TallUHD10	77,567	14,395	0,917	4,230	0,803	19,586%
TallUHD12	78,371	14,527	0,928	4,500	0,607	14,815%
TallUHD15	79,394	14,717	0,939	4,567	0,563	12,801%
TallUHD20	80,078	14,820	0,949	4,630	0,430	9,025%
WoodUHD1	50,072	29,957	0,602	1,370	0,254	23,785%
WoodUHD3	83,258	36,585	0,780	2,633	0,740	35,791%
WoodUHD5	93,840	37,838	0,854	3,467	0,728	26,311%
WoodUHD8	99,835	38,461	0,898	3,600	0,615	18,266%
WoodUHD10	101,466	38,659	0,912	3,867	0,568	15,088%
WoodUHD12	102,392	38,779	0,922	4,100	0,568	13,425%
WoodUHD15	103,419	38,904	0,932	4,133	0,675	15,333%
WoodUHD20	104,259	39,068	0,941	4,870	0,254	5,143%
BundXUHD1	47,304	10,171	0,886	2,070	0,785	37,980%
BundXUHD3	52,631	9,788	0,933	3,233	0,817	25,274%
BundXUHD5	53,811	9,684	0,946	3,833	0,986	25,709%
BundXUHD8	54,644	9,725	0,953	3,930	0,868	22,077%
BundXUHD10	54,890	9,734	0,955	4,167	0,699	16,774%
BundXUHD12	55,057	9,761	0,956	4,370	0,615	14,083%
BundXUHD15	55,162	9,759	0,957	4,467	0,730	16,350%
BundXUHD20	55,258	9,774	0,958	4,900	0,305	6,227%

CampfireXUHD1	24,728	87,489	0,705	1,300	0,535	41,153%
CampfireXUHD3	32,864	34,905	0,831	2,800	0,887	31,667%
CampfireXUHD5	36,240	35,040	0,844	3,400	0,770	22,651%
CampfireXUHD8	39,186	35,092	0,857	3,900	0,759	19,458%
CampfireXUHD10	40,025	35,115	0,864	4,233	0,679	16,037%
CampfireXUHD12	40,805	35,175	0,871	4,370	0,669	15,313%
CampfireXUHD15	41,662	35,201	0,878	4,500	0,731	16,246%
CampfireXUHD20	42,830	35,237	0,886	4,700	0,535	11,383%
ConstructionXUHD1	34,101	6,522	0,850	2,267	0,828	36,515%
ConstructionXUHD3	36,651	7,231	0,898	3,300	0,952	28,858%
ConstructionXUHD5	37,502	7,436	0,911	3,667	0,844	23,023%
ConstructionXUHD8	38,022	7,562	0,916	3,900	0,662	16,968%
ConstructionXUHD10	38,229	7,569	0,918	4,200	0,714	17,009%
ConstructionXUHD12	38,426	7,589	0,919	4,330	0,661	15,251%
ConstructionXUHD15	38,632	7,609	0,921	4,600	0,621	13,510%
ConstructionXUHD20	38,849	7,621	0,922	4,670	0,711	15,239%
FountainsXUHD1	29,510	8,922	0,761	1,330	0,547	41,000%
FountainsXUHD3	39,397	9,816	0,796	2,700	0,837	30,987%
FountainsXUHD5	41,634	10,029	0,818	3,467	0,819	23,634%
FountainsXUHD8	43,910	10,214	0,839	3,930	0,740	18,805%
FountainsXUHD10	45,695	10,324	0,851	4,067	0,691	17,004%
FountainsXUHD12	46,398	10,424	0,860	4,270	0,521	12,207%
FountainsXUHD15	47,163	10,493	0,872	4,230	0,626	14,789%
FountainsXUHD20	48,221	10,621	0,887	4,430	0,504	11,369%
MarathonXUHD1	34,985	16,439	0,698	1,433	0,547	41,000%
MarathonXUHD3	41,513	16,989	0,762	2,700	0,651	24,121%
MarathonXUHD5	43,298	17,442	0,786	3,467	0,571	16,481%
MarathonXUHD8	44,935	17,773	0,807	3,700	0,466	12,597%
MarathonXUHD10	45,523	17,880	0,816	3,867	0,776	20,071%
MarathonXUHD12	46,139	17,912	0,824	4,330	0,479	11,065%
MarathonXUHD15	46,622	17,997	0,831	4,467	0,507	11,360%
MarathonXUHD20	47,478	18,086	0,840	4,670	0,571	12,791%
RunnersXUHD1	42,886	33,367	0,583	1,470	0,571	38,955%
RunnersXUHD3	58,610	24,639	0,767	3,067	0,785	25,595%
RunnersXUHD5	62,083	24,628	0,817	3,333	0,884	26,523%
RunnersXUHD8	64,828	24,772	0,854	3,800	0,761	20,030%
RunnersXUHD10	65,904	24,814	0,870	4,133	0,629	15,213%
RunnersXUHD12	66,680	24,837	0,880	4,500	0,572	12,719%
RunnersXUHD15	67,453	24,842	0,892	4,500	0,630	13,994%
RunnersXUHD20	68,599	24,852	0,903	4,800	0,407	8,476%
TallXUHD1	62,847	10,828	0,765	1,570	0,568	36,276%
TallXUHD3	72,685	13,244	0,869	3,033	0,999	32,948%
TallXUHD5	75,650	13,962	0,907	3,367	0,718	21,339%
TallXUHD8	77,589	14,380	0,932	3,970	0,669	16,857%

TalIXUHD10	78,206	14,462	0,941	4,230	0,817	19,304%
TalIXUHD12	78,703	14,585	0,947	4,500	0,572	12,719%
TalIXUHD15	79,181	14,658	0,953	4,567	0,568	12,445%
TalIXUHD20	79,614	14,730	0,958	4,630	0,490	10,578%
WoodXUHD1	67,209	35,030	0,716	2,070	0,479	35,960%
WoodXUHD3	90,839	37,552	0,848	3,233	0,898	34,973%
WoodXUHD5	96,357	38,378	0,890	3,833	1,015	31,066%
WoodXUHD8	100,281	38,746	0,918	3,930	0,802	21,881%
WoodXUHD10	101,421	38,865	0,928	4,167	0,759	19,458%
WoodXUHD12	102,270	38,943	0,934	4,370	0,621	14,124%
WoodXUHD15	102,943	39,042	0,941	4,467	0,626	14,122%
WoodXUHD20	103,579	39,109	0,948	4,900	0,450	9,502%

Tabuľka 18 obsahuje videosekvencie, ktoré boli doplnené počas priebehu vyhotovenia tejto práce, aby rozšírili už stávajúcu databázu o objektívne ohodnotenie. Tabuľka popisuje názov prekódovanej videosekvencie, jej charakteristické informácie SI a TI a hodnotu SSIM.

*Tabuľka 18 Databáza videosekvencií Ultra Video Group*

Názov videosekvencie	SI [-]	TI [-]	SSIM [-]
BeautyFHD1	19,781	9,802	0,866
BeautyFHD3	23,518	10,025	0,895
BeautyFHD5	23,054	10,026	0,903
BeautyFHD8	24,321	10,047	0,909
BeautyFHD10	24,934	10,053	0,911
BeautyFHD12	25,655	10,102	0,912
BeautyFHD15	26,860	10,146	0,914
BeautyFHD20	28,621	10,189	0,919
BosphorusFHD1	31,466	6,656	0,898
BosphorusFHD3	37,790	5,509	0,953
BosphorusFHD5	37,365	5,473	0,967
BosphorusFHD8	38,281	5,550	0,975
BosphorusFHD10	38,516	5,586	0,978
BosphorusFHD12	38,820	5,607	0,980
BosphorusFHD15	39,105	5,640	0,981
BosphorusFHD20	39,403	5,684	0,988
JockeyFHD1	36,221	27,658	0,882
JockeyFHD3	39,370	29,425	0,944
JockeyFHD5	39,129	29,333	0,957
JockeyFHD8	39,787	29,402	0,964
JockeyFHD10	39,999	29,472	0,966
JockeyFHD12	40,222	29,515	0,967
JockeyFHD15	40,471	29,539	0,968
JockeyFHD20	40,824	29,538	0,969

ReadySteadyGoFHD1	68,846	33,358	0,785
ReadySteadyGoFHD3	96,726	37,280	0,918
ReadySteadyGoFHD5	95,828	37,257	0,947
ReadySteadyGoFHD8	98,401	37,516	0,964
ReadySteadyGoFHD10	99,452	37,563	0,970
ReadySteadyGoFHD12	100,153	37,614	0,974
ReadySteadyGoFHD15	100,935	37,679	0,978
ReadySteadyGoFHD20	101,434	37,691	0,984
BeautyXFHD1	22,617	10,042	0,894
BeautyXFHD3	24,661	10,129	0,905
BeautyXFHD5	24,459	10,137	0,908
BeautyXFHD8	25,096	10,144	0,911
BeautyXFHD10	25,299	10,191	0,913
BeautyXFHD12	25,644	10,204	0,914
BeautyXFHD15	26,061	10,205	0,917
BeautyXFHD20	26,428	10,238	0,921
BosphorusXFHD1	34,441	5,399	0,936
BosphorusXFHD3	37,929	5,371	0,964
BosphorusXFHD5	37,710	5,338	0,972
BosphorusXFHD8	38,236	5,451	0,978
BosphorusXFHD10	38,394	5,473	0,980
BosphorusXFHD12	38,557	5,501	0,981
BosphorusXFHD15	38,672	5,526	0,982
BosphorusXFHD20	38,928	5,558	0,990
JockeyXFHD1	37,468	29,120	0,940
JockeyXFHD3	39,278	29,540	0,958
JockeyXFHD5	39,140	29,529	0,962
JockeyXFHD8	39,559	29,554	0,965
JockeyXFHD10	39,794	29,569	0,966
JockeyXFHD12	39,951	29,550	0,967
JockeyXFHD15	40,154	29,571	0,968
JockeyXFHD20	40,441	29,570	0,969
ReadySteadyGoXFHD1	85,477	36,124	0,869
ReadySteadyGoXFHD3	98,425	37,408	0,939
ReadySteadyGoXFHD5	97,730	37,328	0,960
ReadySteadyGoXFHD8	99,379	37,510	0,972
ReadySteadyGoXFHD10	99,990	37,554	0,976
ReadySteadyGoXFHD12	100,352	37,592	0,979
ReadySteadyGoXFHD15	100,650	37,615	0,982
ReadySteadyGoXFHD20	101,022	37,643	0,987
BeautyUHD1	9,512	15,035	0,686
BeautyUHD3	11,942	10,710	0,740
BeautyUHD5	11,949	10,148	0,749
BeautyUHD8	12,124	10,102	0,757



BeautyUHD10	12,206	10,127	0,760
BeautyUHD12	12,359	10,147	0,763
BeautyUHD15	12,663	10,174	0,765
BeautyUHD20	13,406	10,191	0,768
BosphorusUHD1	16,466	9,854	0,870
BosphorusUHD3	18,802	6,801	0,913
BosphorusUHD5	19,920	5,834	0,931
BosphorusUHD8	20,696	5,485	0,944
BosphorusUHD10	21,010	5,512	0,948
BosphorusUHD12	21,232	5,520	0,952
BosphorusUHD15	21,527	5,557	0,955
BosphorusUHD20	21,858	5,604	0,958
JockeyUHD1	19,862	46,958	0,825
JockeyUHD3	21,412	27,336	0,866
JockeyUHD5	21,268	28,534	0,883
JockeyUHD8	21,164	29,285	0,894
JockeyUHD10	21,213	29,387	0,897
JockeyUHD12	21,143	29,447	0,899
JockeyUHD15	21,270	29,484	0,902
JockeyUHD20	21,381	29,552	0,905
ReadySteadyGoUHD1	34,381	52,913	0,742
ReadySteadyGoUHD3	46,645	35,194	0,832
ReadySteadyGoUHD5	51,920	36,786	0,885
ReadySteadyGoUHD8	54,615	37,317	0,918
ReadySteadyGoUHD10	55,318	37,499	0,929
ReadySteadyGoUHD12	55,928	37,617	0,936
ReadySteadyGoUHD15	56,753	37,708	0,943
ReadySteadyGoUHD20	57,522	37,762	0,950
BeautyXUHD1	11,241	10,043	0,764
BeautyXUHD3	12,701	10,184	0,771
BeautyXUHD5	13,622	10,199	0,773
BeautyXUHD8	14,905	10,192	0,775
BeautyXUHD10	15,570	10,205	0,776
BeautyXUHD12	15,972	10,228	0,777
BeautyXUHD15	16,859	10,252	0,779
BeautyXUHD20	17,649	10,290	0,781
BosphorusXUHD1	18,266	6,425	0,930
BosphorusXUHD3	20,806	5,612	0,949
BosphorusXUHD5	20,900	5,572	0,955
BosphorusXUHD8	21,286	5,585	0,959
BosphorusXUHD10	21,499	5,605	0,961
BosphorusXUHD12	21,635	5,620	0,962
BosphorusXUHD15	21,766	5,634	0,963
BosphorusXUHD20	21,963	5,654	0,965

JockeyXUHD1	20,422	53,540	0,890
JockeyXUHD3	21,217	29,526	0,903
JockeyXUHD5	21,003	29,598	0,907
JockeyXUHD8	21,090	29,580	0,909
JockeyXUHD10	21,168	29,605	0,909
JockeyXUHD12	21,295	29,610	0,910
JockeyXUHD15	21,544	29,621	0,911
JockeyXUHD20	21,947	29,629	0,912
ReadySteadyGoXUHD1	41,074	56,170	0,859
ReadySteadyGoXUHD3	53,750	37,275	0,913
ReadySteadyGoXUHD5	55,363	37,557	0,932
ReadySteadyGoXUHD8	56,574	37,668	0,945
ReadySteadyGoXUHD10	57,016	37,678	0,950
ReadySteadyGoXUHD12	57,422	37,691	0,954
ReadySteadyGoXUHD15	57,710	37,711	0,957
ReadySteadyGoXUHD20	58,057	37,734	0,959

## Príloha B.1 Popis kódovania a dekódovania videosekvencií

Príkaz nástroja ffmpeg pre kódovanie vyzeral nasledovne (popis jednotlivých parametrov je obsiahnutý v práci, v kapitole 5.2.3):

```
ffmpeg -f rawvideo -video_size 3840x2160 -pix_fmt yuv420p -framerate 30 -i Bund_raw.yuv -  
vcodec libx264 -x264-params keyint=1:min-keyint=1:bframes=3:b-adapt=1:bitrate=1000:vbv-  
maxrate=1000:vbv-buFSIZE=1000 -vf scale=3840x2160 BundUHD1.ts
```

Po kódovaní boli videosekvencie opäť dekódované na bezstratový formát s použitím rovnakého nástroja ffmpeg. Príkaz dekódovania:

```
ffmpeg -i BundUHD1.ts -f rawvideo -vcodec rawvideo -pix_fmt yuv420p -video_size 3840x2160  
-framerate 30 BundUHD1_dekompres.yuv
```

Na vyhodnotenie kvality jednotlivých videosekvencií pomocou objektívnej metriky bol použitý nástroj MSU Video Quality Measurement Tool. Príklad porovnania referenčnej a testovacej sekvencie:

```
msu_metric -f "Bund_raw.yuv"  
IYUV -yw 3840 -yh 2160 -f "BundUHD1_dekompres.yuv"  
IYUV -yw 3840 -yh 2160 -metr psnr -cc YYUV -metr ssim_precise -cc YYUV -metr vqm -cc YYUV
```

kde sa nachádza názov referenčnej videosekvencie, jej rozlíšenie, nasleduje názov porovnáwanej sekvencie, jej rozlíšenie a metriky, ktoré chceme vyhodnotiť. Rozlíšenie je potrebné uvádzať nakoľko nekomprimované video v sebe nenesie túto informáciu.



## Príloha B.2 Popis objektu neurónovej siete

V prípade vytvárania doprednej (feedforward) neurónovej siete použijeme príkaz `newff()`, ktorého vstupné parametre sú:

```
Vstup
Vystup
TransFcns = {' '};
BTF = "";
BLF = "";
PF = "";
Layers = [i j];
net = newff(Vstup, Vystup, Layers, TransFcns, BTF, BLF, PF); kde
```

Vstup predstavuje vektor normovaných vstupov,  
Vystup predstavuje vektor požadovaného výstupu,  
TransFcns predstavuje aktivačnú funkciu jednotlivých vrstiev,  
BTF predstavuje tréningovú funkciu,  
BLF predstavuje učiacu sa funkciu siete,  
PF predstavuje funkciu pre určenie chyby pri tréningu,  
Layers predstavuje počet neurónov v jednotlivých vrstvách.

Do premennej „net“ sa uloží objekt neurónovej siete, ktorý obsahuje celú dátovú štruktúru neurónovej siete. Pred tréningom siete je potrebné danú sieť inicializovať, kedy dochádza k počiatočnému nastaveniu váh a prahov neurónovej siete. Inicializácia siete sa vykoná príkazom:

```
net = init(net);
```

Pre presnejšie výsledky bola neurónová sieť, pre každú topológiu, tréningovaná 10 krát, pričom vždy boli náhodne generované váhy a prahy. Pri tréningu sa používala validačná množina dát, ktorá overovala tréningovanie siete, aby nedochádzalo k jej pretréningu. Validačná množina bola náhodne generovaná z tréningovej množiny dát v pomere 70:30 v percentuálnom zobrazení. Pre štatistickú významnosť bol tento proces opakovaný opäť 10 krát, aby sa sieť natréningovala pri rozličných dátach tréningovej a validačnej množiny. Priemer týchto tréningov je braný ako výsledok tréningovania.

```
% Mapovanie vstupov
[pn,ps] = mapminmax(Vstup);
[tn,ts] = mapminmax(Vystup);

% Definovanie Vrstiev
j=pocetNeuronovPrvaVrstva;
l= pocetNeuronovDruhaVrstva;

% Tréningovanie siete
while j<maxPocetNeuronov
    Layers = [j l];
    net = newff(pn, tn, Layers, TransFcns, BTF, BLF, PF);
    net.divideParam.testRatio = x/100;
    net.divideParam.trainRatio = x/100;
    net.divideParam.valRatio = x/100;
```

```

    for k=1:n
        [net, tr] = train(net, pn, tn);
        trainInput = pn .* tr.trainMask{1};
        valInput = pn .* tr.valMask{1};
        testInput = pn .* tr.testMask{1};

        simulaciaTrainData = sim(net, trainInput);
        simulaciaValidData = sim(net, valInput);
        simulaciaTestData = sim(net, testInput);
    end

%    Uloženie váh zodpovedajúcich najlepšie natrénovanej sieti
weights = [];
Weights = [weights getwb(net)];
end

%    Načítanie váh zodpovedajúcich najlepšie natrénovanej sieti
najlepsie_vahy = xlsread('subor',topologia);

%    Simulovanie siete pri zadani vstupných parametrov
for k=1:10
    net = init(net);
    net = setwb(net, najlepsie_vahy ');
    [net, tr] = train(net, pn, tn);
end

simulacia_pozadovanej_predikcie = sim(natrenovanaSiet, TestHodnoty);

```

## Príloha C Vplyv modifikovania funkcie pre výpočet chyby pri tréningu neurónovej siete

### 1. Predikcia bitového toku na základe požadovanej kvality vyjadrenej metrikou SSIM

	Topológia	Funkcia	Ratio	MSE	$R^2_{\text{TRAIN}} [-]$	$R^2_{\text{VAL}} [-]$	$R^2_{\text{TEST}} [-]$
<b>H.264 + H.265</b>	<b>43-21</b>	MSE	-	0,041	0,949	0,944	0,963
		msereg	0,2	0,047	0,939	0,962	0,968
			0,3	0,035	0,954	0,932	0,975
			0,4	0,038	0,944	0,952	0,967
			0,5	0,046	0,943	0,959	0,965
			0,6	0,043	0,949	0,934	0,978
			0,9	0,045	0,938	0,953	0,977
<b>H.264</b>	<b>79-39</b>	MSE	-	0,0001	0,999	0,999	0,999
		msereg	0,2	0,005	0,994	0,955	0,983
			0,3	0,005	0,992	0,938	0,975
			0,4	0,007	0,992	0,916	0,984
			0,5	0,007	0,993	0,978	0,991
			0,6	0,010	0,987	0,931	0,984
			0,9	0,007	0,992	0,932	0,978
<b>H.265</b>	<b>61-31</b>	MSE	-	0,001	0,999	0,999	0,999
		msereg	0,2	0,049	0,946	0,886	0,968
			0,3	0,045	0,953	0,875	0,961
			0,4	0,030	0,959	0,875	0,954
			0,5	0,076	0,911	0,962	0,977
			0,6	0,074	0,914	0,966	0,977
			0,9	0,071	0,898	0,973	0,979

2. Predikcia bitového toku na základe požadovanej kvality vyjadrenej pomocou MOS skóre

	Topológia	Funkcia	Ratio	MSE	$R^2_{\text{TRAIN}} [-]$	$R^2_{\text{VAL}} [-]$	$R^2_{\text{TEST}} [-]$
H.264 + H.265	47-23	MSE	-	0,013	0,982	0,991	0,992
		msereg	0,2	0,020	0,972	0,981	0,971
			0,3	0,016	0,979	0,982	0,974
			0,4	0,018	0,975	0,977	0,970
			0,5	0,013	0,983	0,987	0,986
			0,6	0,019	0,975	0,981	0,969
			0,9	0,010	0,986	0,993	0,989
H.264	35-17	MSE	-	0,006	0,992	0,990	0,996
		msereg	0,2	0,009	0,978	0,984	0,976
			0,3	0,017	0,978	0,968	0,975
			0,4	0,034	0,988	0,988	0,982
			0,5	0,009	0,987	0,987	0,990
			0,6	0,014	0,982	0,985	0,981
			0,9	0,016	0,981	0,977	0,975
H.265	71-35	MSE	-	0,011	0,988	0,982	0,993
		msereg	0,2	0,009	0,989	0,985	0,989
			0,3	0,019	0,976	0,975	0,971
			0,4	0,015	0,980	0,977	0,975
			0,5	0,012	0,986	0,981	0,982
			0,6	0,015	0,981	0,978	0,983
			0,9	0,010	0,987	0,990	0,983



## Príloha D.1 Verifikácia a validácia modelu predikcie subjektívnej hodnoty na základe scény, bitového toku a objektívneho výsledku metrikou SSIM pre kompresné štandardy separátne

V tejto prílohe sa nachádza verifikácia a validácia modelu predikcie subjektívnej kvality na základe charakteristík videoscény a vyjadrením kvality pomocou objektívnej metriky SSIM pre simuláciu jednotlivých kompresných štandardov pri rozlíšení UHD a simuláciu oboch rozlíšení a kodekov v jednom simulačnom behu.

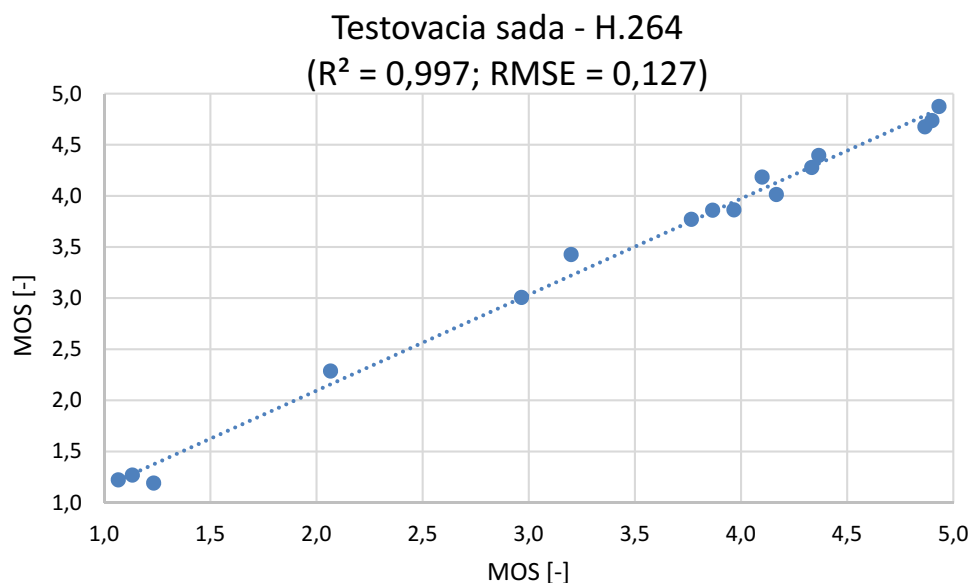
### UHD

- 1) Vyhodnotenie simulácie pre kompresný štandard H.264 a rozlíšenie UHD
  - Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

Tabuľka 19 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu MOS pomocou SSIM, H.264, UHD

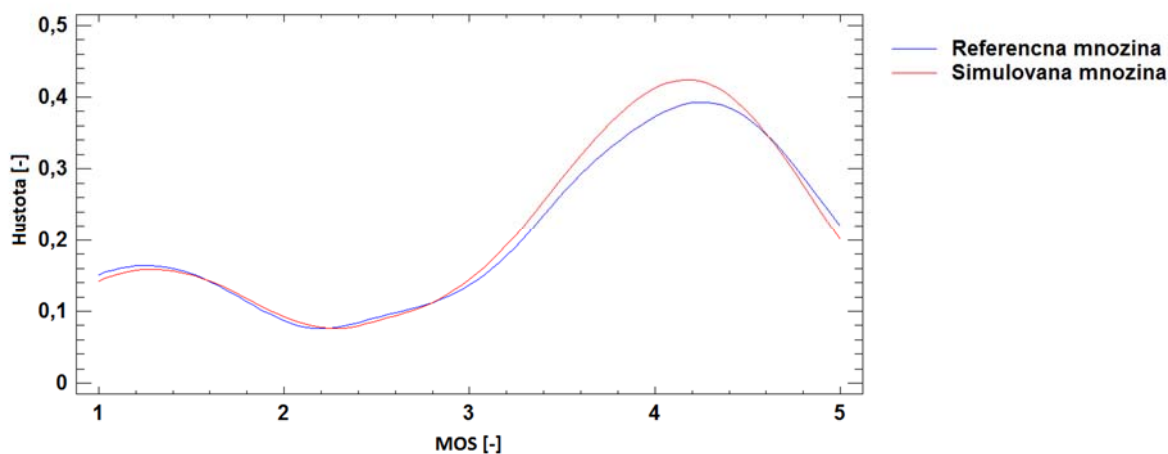
	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	3,433	3,441
Minimum [-]	1,068	1,190
Maximum [-]	4,933	4,875
Rozsah [-]	3,867	3,685
Medián [-]	3,917	3,862
Konfidenčný interval [-]	<2,711; 4,156>	<2,761 ; 4,121>
Pearsonov korelačný koeficient [-]		0,997
Spearmanov korelačný koeficient [-]		0,988

- Korelačný diagram a funkcia RMSE



Graf 16 Korelačný diagram, predikcia BR pomocou SSIM, H.264, UHD

- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu



Graf 17 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264, UHD

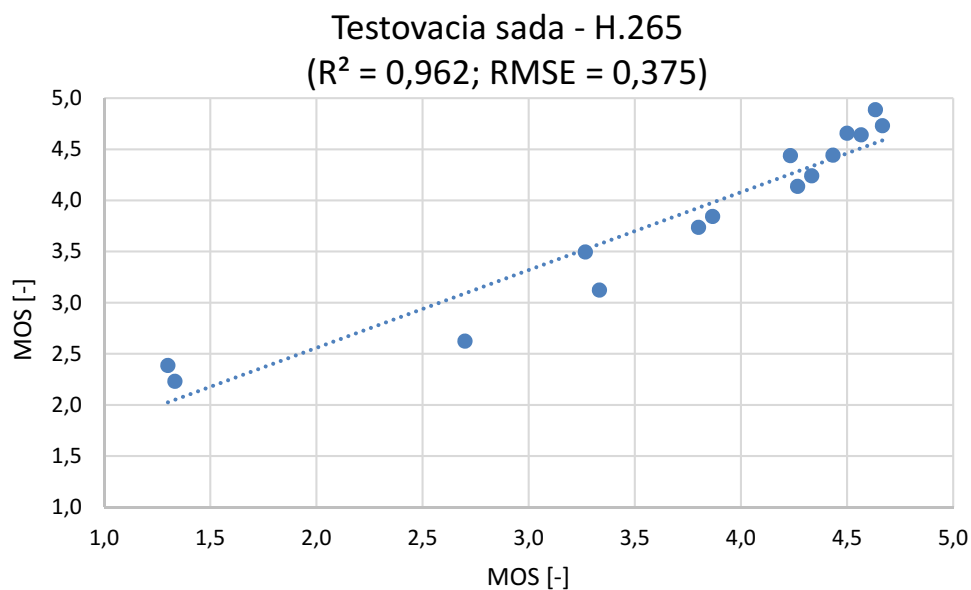
2) Vyhodnotenie simulácie pre kompresný štandard H.265 a rozlíšenie UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

Tabuľka 20 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou SSIM, H.265, UHD

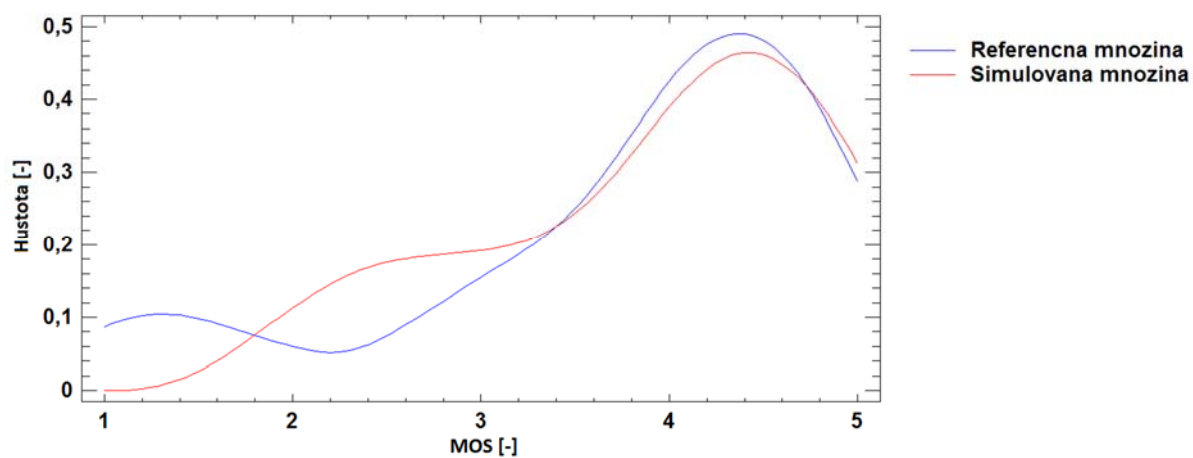
	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	3,758	3,903
Koeficient variácie [-]	0,300	0,230
Minimum [-]	1,3	2,229
Maximum [-]	4,9	4,886
Rozsah [-]	3,6	2,657
Medián [-]	4,25	4,136
Konfidenčný interval [-]	<3,161 ; 4,356>	<3,427 ; 4,378>
Pearsonov korelačný koeficient [-]		0,962
Spearmanov korelačný koeficient [-]		0,974

- Korelačný diagram a funkcia RMSE



Graf 18 Korelačný diagram, predikcia BR pomocou SSIM, H.264+H.265, UHD

- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu



Graf 19 Krivka hustoty rozdelenia, predikcia BR pomocou SSIM, H.264+H.26, UHD

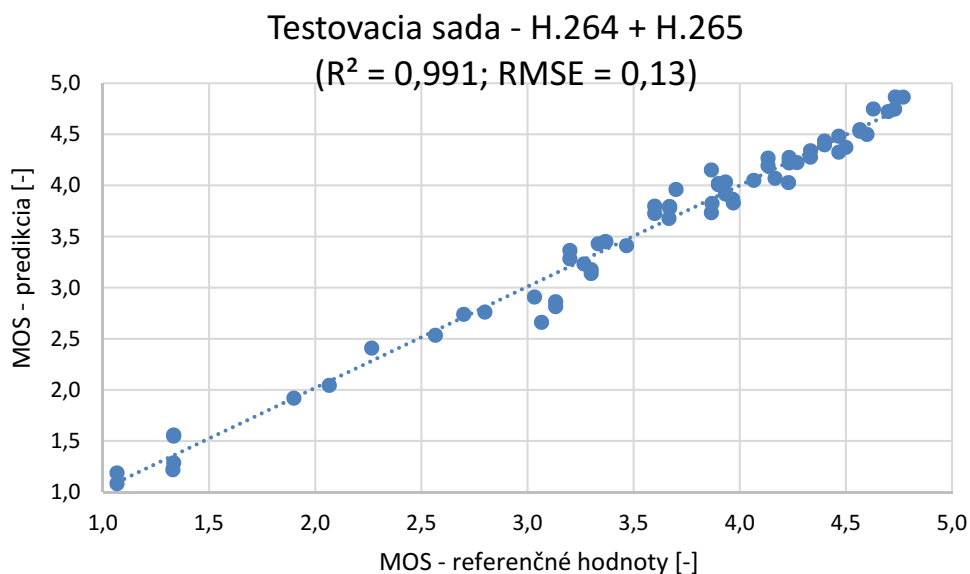
## FHD + UHD

Vyhodnotenie simulácie pre oba kompresné štandardy a rozlíšenia v jednom simulačnom behu.

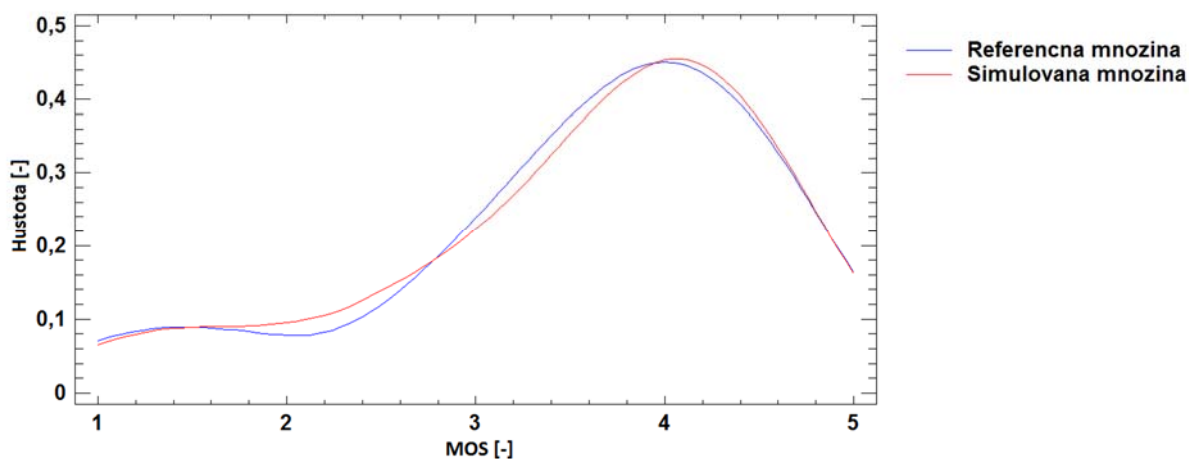
- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	3,546	3,551
Minimum [-]	1,067	1,083
Maximum [-]	4,770	4,865
Rozsah [-]	3,703	3,783
Medián [-]	3,867	3,826
Konfidenčný interval [-]	<3,297 ; 3,795>	<3,303 ; 3,8>
Pearsonov korelačný koeficient [-]	0,991	
Spearmanov korelačný koeficient [-]	0,989	

- Korelačný diagram a funkcia RMSE



- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu



## Príloha D.2 Verifikácia a validácia modelu predikcie bitového toku na základe SI a TI informácií a kvalitou videosekvencie charakterizovanou pomocou MOS stupnice pre kompresné štandardy separátne

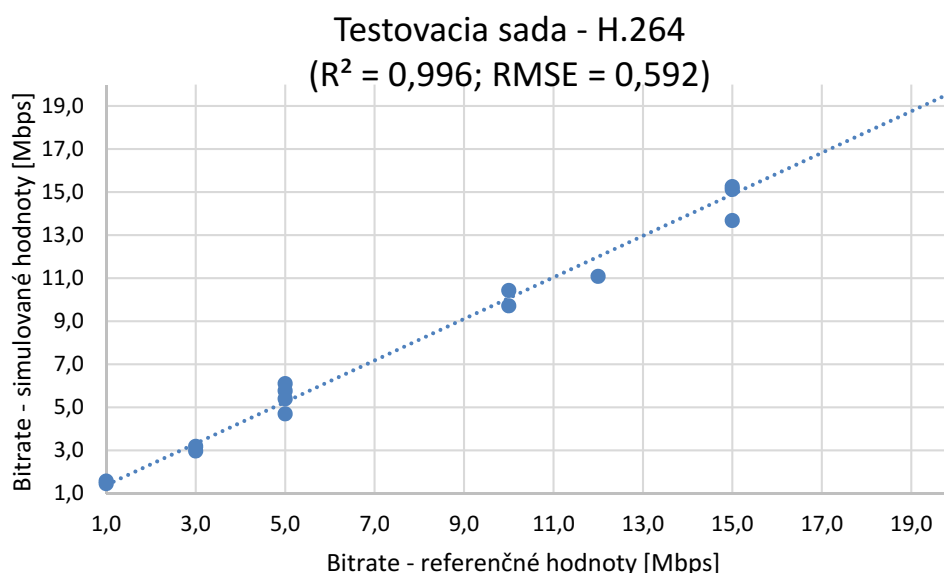
1) Vyhodnotenie simulácie pre kompresný štandard H.264 a rozlíšenie UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

Tabuľka 21 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.264, UHD

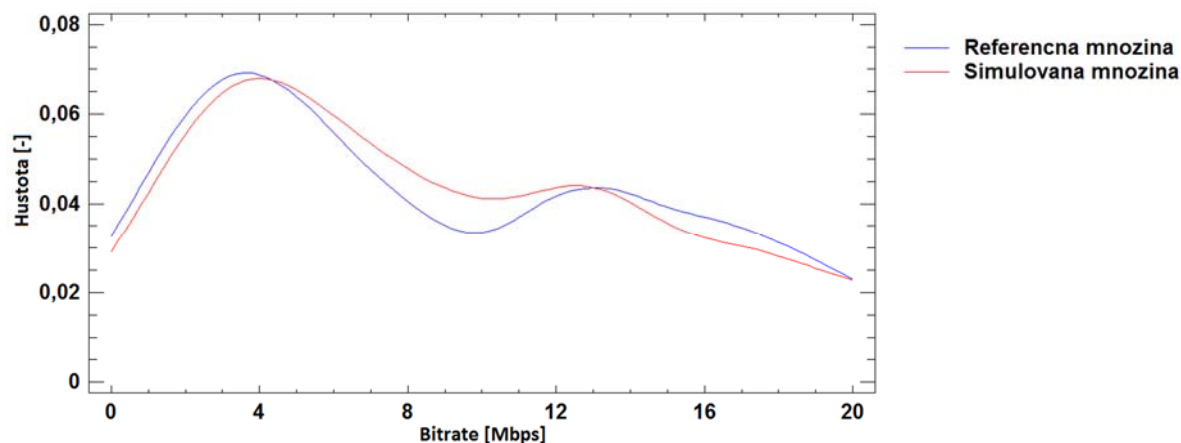
	Referenčná množina	Simulovaná množina
Priemerná hodnota [-]	9,063	9,16862
Minimum [-]	1	1,4537
Maximum [-]	20	20,18
Medián [-]	7,5	7,9033
Konfidenčný interval [-]	<5,623 ; 12,502>	<5,837 ; 12,5>
Pearsonov korelačný koeficient [-]		0,996
Spearmanov korelačný koeficient [-]		0,987

- Korelačný diagram a funkcia RMSE



Graf 20 Korelačný diagram, predikcia BR pomocou MOS, H.264, UHD

- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu



Graf 21 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.264, UHD

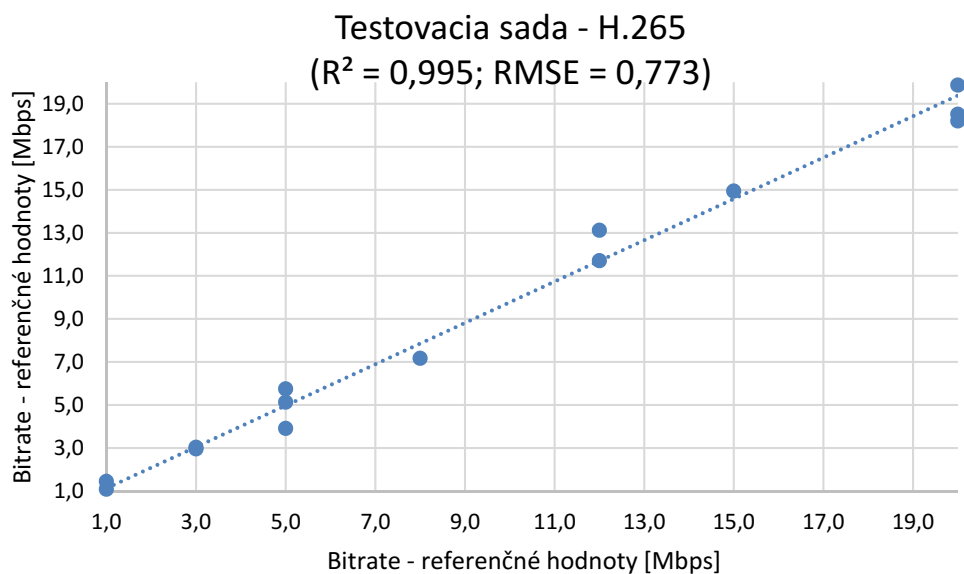
2) Vyhodnotenie simulácie pre kompresný štandard H.265 a rozlíšenie UHD

- Porovnanie súhrnných štatistík pre referenčnú a simulovanú množinu

Tabuľka 22 Súhrnné štatistiky referenčných a simulovaných hodnôt pre predikciu BR pomocou MOS, H.265, UHD

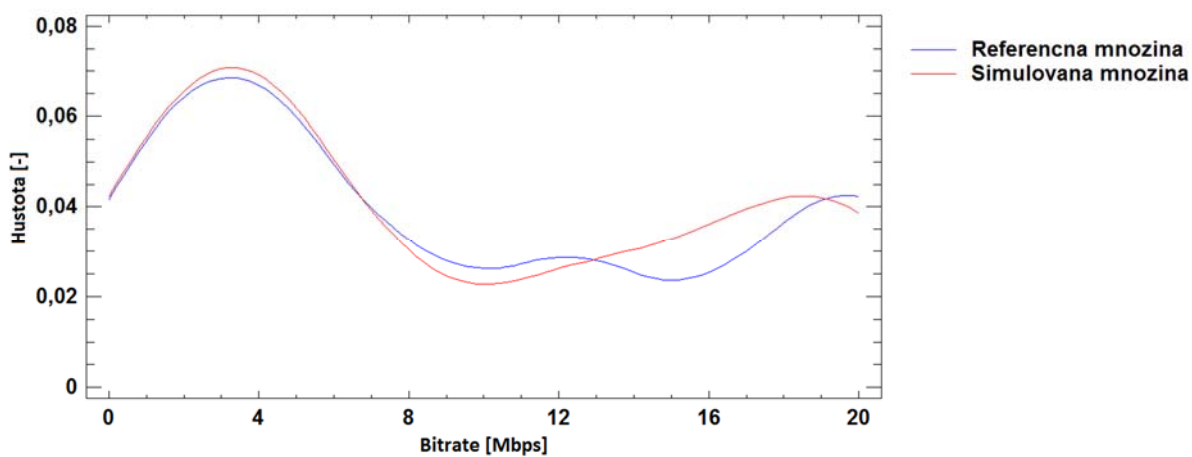
	Referenčná množina	Simulovaná množina
<b>Priemerná hodnota [-]</b>	9,438	9,238
<b>Minimum [-]</b>	1,000	0,836
<b>Maximum [-]</b>	20,000	20,180
<b>Medián [-]</b>	6,500	6,456
<b>Konfidenčný interval [-]</b>	<5,43112 ; 12,4439>	<5,37154 ; 13,1037>
<b>Pearsonov korelačný koeficient [-]</b>		0,995
<b>Spearmanov korelačný koeficient [-]</b>		0,985

- Korelačný diagram a funkcia RMSE



Graf 22 Korelačný diagram, predikcia BR pomocou MOS, H.265, UHD

- Krivka hustoty rozdelenia pre referenčnú a simulovanú množinu



Graf 23 Krivka hustoty rozdelenia, predikcia BR pomocou MOS, H.265, UHD





## Príloha E Obsah priloženého média

- Dizertačná práca – elektronická verzia dizertačnej práce
- Zložka *Subjektívne ohodnotenie* – obsahuje zložku *Videosekvencie SJTU Media Lab*, kde som uložil prepis subjektívneho ohodnotenia videosekvencií od každého hodnotiteľa pre FHD (*ohodnotenie\_ACR\_FHD*) a UHD (*ohodnotenie\_ACR\_UHD*). Ďalej obsahuje súbor *grafy.xlsx*, kde sú grafy pre jednotlivé videosekvencie, kompresné štandardy.
- *Objektívne ohodnotenie* – obsahuje zložky *Videosekvencie SJTU Media Lab* a *Videosekvencie Ultra Video Group*, ktoré obsahujú vypočítané hodnoty SSIM a prislúchajúce grafické zobrazenie pre jednotlivé videosekvencie
- *Implementácia NN v prostredí Matlab* – obsahuje príklad konfigurácie NN pre simuláciu MOS skóre, na základe vstupov SI, TI, BR a SSIM (zložka *sim\_ACR\_vstup\_SI\_TI\_BR\_SSIM*). V implementáciách iných neurónových sietí sa menili názvy vstupných súborov a jednotlivé vstupy a výstupy. Trénovanie ukladá všetky natrénované dáta a vyhodnotenie výsledkov do súboru *trenovanieVah.xlsx*. Z tohto súboru boli vyhodnotené najlepšie priemerné výsledky jednotlivých topológií. Každá topológia nesie v sebe príznak s jej číslom a rovnako sa ukladajú váhy pre každú tréňovanú topológiu do súboru *Vahy.xlsx*, kde popis hárka označuje topológiu. Tieto váhy sú potom načítané pri simulácii, ktorá predikuje požadované hodnoty. Výsledok sa ukladá do súboru *simulacia.xlsx*. Celá simulácia je popísaná v súbore *sim\_ACR\_SI\_TI\_BR\_SSIM.m*. Po natrénovaní je možné simulovať hocikáku už tréňovanú topológiu, keď zadáme počet neurónov v jednotlivých vrstvách a číslo simulácie priamo do algoritmu *konkretna\_topologia.m*, ktorý je tiež priložený.
- *Vstupné dáta pre NN* – obsahujú vstupné súbory pre jednotlivé NN. Údaje sú rozdelené pre jednotlivé rozlíšenia a kodeky, čo popisuje ich názov. Každý tento súbor obsahuje viac hárkov, ktoré slúžia ako vstupy a výstupy NN. Názov hárku popisuje dané dáta. Kodek je vyjadrený číselne: 0 (H.264) a 1 (H.265), obdobne rozlíšenie je definované číslom: 0 (FHD), 1 (UHD)
- *Najlepšie topológie dosiahnuté pri jednotlivých simuláciách* - obsahuje výsledky všetkých simulovaných topológií pre jednotlivé rozlíšenia a kompresné štandardy

Jednotlivé videosekvencie nie sú obsahom ani dizertačnej práce ani priloženého média, pretože sa jedná o príliš veľké množstvo dát. V práci je uvedená dostupná literatúra odkiaľ boli stiahnuté originály a je popísaný postup ako ich zakódovať, dekodovať a ako dospieť k jednotlivým výsledkom.