



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

**Fakulta elektrotechnická
Katedra radioelektroniky**

**Hodnocení kvality obrazu mluvčího
znakového jazyka
Quality evaluation of image of sign
language speaker**

Bakalářská práce

Studijní program: Komunikace, multimédia a elektronika

Studijní obor: Multimediální technika

Vedoucí práce: Ing. Martin Bernas, CSc.

Student: Daria Bashkatova

2016

České vysoké učení technické v Praze
Fakulta elektrotechnická

katedra radioelektroniky

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Student: **Daria Bashkatova**

Studijní program: Komunikace, multimédia a elektronika

Obor: Multimediální technika

Název tématu: **Hodnocení kvality obrazu mluvího znakového jazyka**

Pokyny pro vypracování:

Prostudujte specifické charakteristiky obrazu mluvího znakového jazyka a objektivní metody hodnocení kvality obrazu. Na jejich základě navrhnete metodu vhodnou pro objektivní hodnocení kvality obrazu mluvího znakového jazyka, metodu naprogramujete v prostředí Matlab a ověřte porovnáním se subjektivním hodnocením.

Seznam odborné literatury:

[1]Wu,H.R.;Rao,K.R.: Digital Video Image Quality and Perceptual Coding. Taylor&Francis, 2006, ISBN-13: 978-0-824-7-2777-2

[2]Rousek M.: Hodnocení kvality obrazu mluvího znakového jazyka. Bakalářská práce, FEL ČVUT v Praze, 2015.

[3]Kuchařová,L. Jazyk neslyšících. Výzkumná zpráva. Ústav českého jazyka a teorie komunikace, FFUK, Praha 2005.

Vedoucí: Ing. Martin Bernas, CSc.

Platnost zadání: do konce letního semestru 2016/2017

L.S.

doc. Mgr. Petr Páta, Ph.D.
vedoucí katedry

prof. Ing. Pavel Ripka, CSc.
děkan

V Praze dne 19. 2. 2016

Anotace

Tato bakalářská práce se ve svém teoretickém základu věnuje rozebírání modelu lidského zrakového systému HVS, který je brán v úvahu navrhovanou metodou hodnocení kvality obrazu znakového mluvčího. Po zkoumání celého přehledu různých objektivních metod a metrik byla zvolena a podrobně zpracována metrika MSSIM. Daná metrika se používá při návrhu a realizaci objektivní metody hodnocení kvality obrazu mluvčího znakového jazyka. Nakonec teoretická část se uzavře řešením dělení sledovaného obrazu s mluvčím znakového jazyka na oblastí zájmu neslyšících.

V praktické části práce se navrhne, následovně se realizuje a ověří objektivní metoda hodnocení kvality videosekvence mluvčího znakového jazyka. K naprogramování metody je použito prostředí MATLAB. Metoda k získání výsledných hodnot používá metriku MSSIM, toleruje lidský zrakový systém a bere v úvahu dělení obrazu na oblasti zájmu. Navržená metoda je otestována pomocí série videosekvencí, které obsahují mluvčího znakového jazyka. Videosekvence mají různou bitovou rychlost a nastavené parametry. Na výstupu výsledné hodnoty těchto testů jsou porovnány s původními subjektivními testy, které taky slouží ke kalibraci metody.

Klíčová slova: SSIM, MSSIM, objektivní hodnocení kvality obrazu, MATLAB, ROI, znakový jazyk, videosekvence, obrazová komprese

Abstract

Bachelor's thesis is in its theoretical foundation dedicated to dismantling the model of the human visual system HVS, which is taken into account in the proposed method to evaluate image quality of sign language speaker. After searching through the entire overview of different methods and objective metrics, MSSIM metric was chosen and elaborated. This metric has been used to design and implement an objective evaluation method for image quality of a sign language speaker. Finally, the theoretical part was closed with the solution of dividing a picture into areas of interest of sign language speaker.

In the practical part of the work a method for evaluating the quality of a movie with a sign language speaker was suggested, subsequently verified and realised. To program this method was used MATLAB. The method to obtain final values was using a MSSIM metric, that tolerates human visual system and takes into account the division of the image and the area of interest. The proposed method is tested on series of movies that include sign language speaker. Movies have different bit rate and settable parameters. The output of resulting values from these tests is compared with the original subjective tests, which also serves to calibrate the method.

Keywords: SSIM, MSSIM, objective quality assessment, MATLAB, ROI, sign language, movie, image compression

Poděkování

Ráda bych poděkovala vedoucímu práce, panu Ing. Martinu Bernasovi, CSc., za velmi přátelský přístup, důležité rady při zpracování mé bakalářské práce a možnost použití školní techniky a potřebný testovací materiál.

Prohlášení

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

V Praze dne

.....
Podpis studenta

Obsah

Úvod	8
1 Lidský zrakový systém	10
1.1 Funkce oka.....	10
1.2 Psycho-fyzikální vlastnosti HVS.....	11
1.2.1 Foveální a periferní vidění.....	11
1.2.2 Přizpůsobení světlu.....	12
1.2.3 Kontrast.....	13
1.2.4 Maskování.....	13
1.3 Oční pohyby.....	14
1.3.1 Základy očních pohybů.....	14
1.3.2 Fixace.....	14
1.3.3 Sakády.....	14
1.3.4 Rozlišovací mez oka.....	15
2 Subjektivní metody hodnocení kvality obrazu	16
2.1 Subjektivní testy.....	16
2.2 Metody DS (Double Stimulus).....	16
2.2.1 DSIS.....	16
2.2.2 DSCQS.....	17
2.3 Metody SS (Single Stimulus).....	17
2.3.1 SSM.....	17
2.3.2 SSCQE.....	17
2.4 Výhody a nevýhody subjektivního testování.....	17
3 Objektivní metody hodnocení kvality obrazu	18
3.1 MSE a PSNR.....	18
3.2 No reference metody (NR).....	19
3.3 Reduced reference metody (RR).....	19
3.4 Full reference metody (FR).....	20
3.4.1 Lubinův model HVS.....	20
3.4.2 VDP (Visible Differences Predictor).....	21
3.4.3 Watsonův DCT model.....	22
3.4.4 Sarnoffův JND model.....	23
3.4.5 SSIM (Structural Similarity Index).....	24
4 Charakteristika obrazu mluvího znakového jazyka	28
4.1 Znakový jazyk.....	28
4.2 Oblasti zájmu neslyšících diváků.....	29
4.3 Určení oblastí zájmu neslyšících diváků.....	30
4.4 Praktická implementace oblastí zájmu neslyšících diváků.....	31
4.4.1 HbbTV a asistivní technologie.....	32

5. Testované videosekvence	33
5.1 Detekce ROI.....	33
5.2 Hpar.....	33
5.3 Kódování videosekvencí.....	34
6 Praktická realizace objektivní metody	36
6.1 Předpoklady pro návrh objektivní metody.....	36
6.2 Popis navržené metody.....	36
6.2.1 Načtení videosekvencí.....	36
6.2.2 Načtení masek.....	37
6.2.3 Výpočet algoritmu metriky.....	38
6.2.4 Zajištění aditivnosti metody.....	39
6.2.5 Konečné hodnota objektivní kvality obrazu.....	39
7 Vyhodnocení objektivní kvality videosekvencí	40
7.1 Výsledky objektivní kvality obrazu u originální SSIM.....	40
7.2 Výsledky objektivní kvality obrazu u SSIM MATLABu...	48
7.3 Ohodnocení funkčnosti metody.....	51
Závěr	52
Seznam použitých obrázků	55
Seznam tabulek	56
Seznam použitých symbolů a zkratek	57
Seznam použité literatury	58
Obsah přiloženého DVD	62

Úvod

V současné době probíhá velký rozvoj multimediální techniky. Parametry multimediálních zařízení se každým rokem o stupeň zlepšují a nabízejí stále více funkcí a možností. S tím souvisí rozvoj v mnoha odvětvích, ve kterých lze tato zařízení využít díky jejich možnostem. Ve vyspělých zemích lze již nějakou dobu sledovat snahu vytvářet společensky rovné šance a příležitosti. Člověk se zdravotním handicapem, jakým je např. postižení sluchu nebo zraku, se za určitých

okolností často dokáže plnohodnotně zapojit do pracovního života. Důležité ale je také zajistit handicapovaným lidem zlepšení v odvětví společenském. Ve většině vědních oborů je věnována pozornost oblastem, které by přispěly ke zvýšení kvality života zdravotně postižených lidí. Pokud přeskočíme pomůcky na medicínské bázi, označují se dané postupy a principy termínem *asistivní technologie*. Tento termín zahrnuje veškeré hardwarové a softwarové prostředky, které nějakým způsobem pomáhají zdravotně postižené populaci usnadňovat jejich život. V rámci výzkumu týkající se neslyšící komunity takovou službu představuje přirozené tlumočení do znakového jazyka, které lze stále častěji

spatřit ve vizuálních sdělovacích prostředcích, především v televizi.

Tato bakalářská práce je součástí rozsáhlého projektu, jenž si klade za úkol vytvořit automatizovaný animovaný model tlumočnicka do znakového jazyka, který by mohl být posléze použit tam, kde dochází ke kontaktu se sluchově postiženými a není k dispozici tlumočnick (pošta, úřady, banky, nádraží aj.). Za primární se považuje nasazení modelu tlumočnicka znakového jazyka do televizního

vysílání. Tato práce vychází z již zpracovaných částí tohoto projektu. Cílem

této práce je nalezení vhodné objektivní metody pro hodnocení obrazu mluvího znakového jazyka. Tato metoda by mohla poté nahradit subjektivní hodnocení kvality obrazu, které je v současné době jediným standardem pro určení kvality obrazu.

Zkoumání objektivní kvality obrazu je v současné době důležitým tématem díky prudkému rozvoji multimédií, internetu a mobilní komunikace. Důsledkem tohoto rozvoje je zavádění nových kompresních metod obrazu splňujících kritéria efektivního přenosu a dostatečné kvality obrazu. Díky tomu je tedy snaha nalézt vhodné efektivní, levné a časově nenáročné kritérium, podle kterého by bylo možné kvalitu obrazu hodnotit.

Teoretická část práce pojednává v 1. kapitole o vlastnostech lidského zrakového systému, v 2. kapitole jsou probírány subjektivní metody hodnocení kvality obrazu, 3. kapitola je venována objektivním metodám. Kapitoly 4 a 5 probírají oblast zájmu neslyšících diváků a modely a kodování videosekvencí pro přenos mluvího znakového jazyka.

Praktická část se skládá ze dvou kapitol 6 a 7 s návrhem objektivní metody hodnocení obrazu v MATLABu a výsledky a jejich porovnání s výsledky subjektivních testů.

Kapitola 1

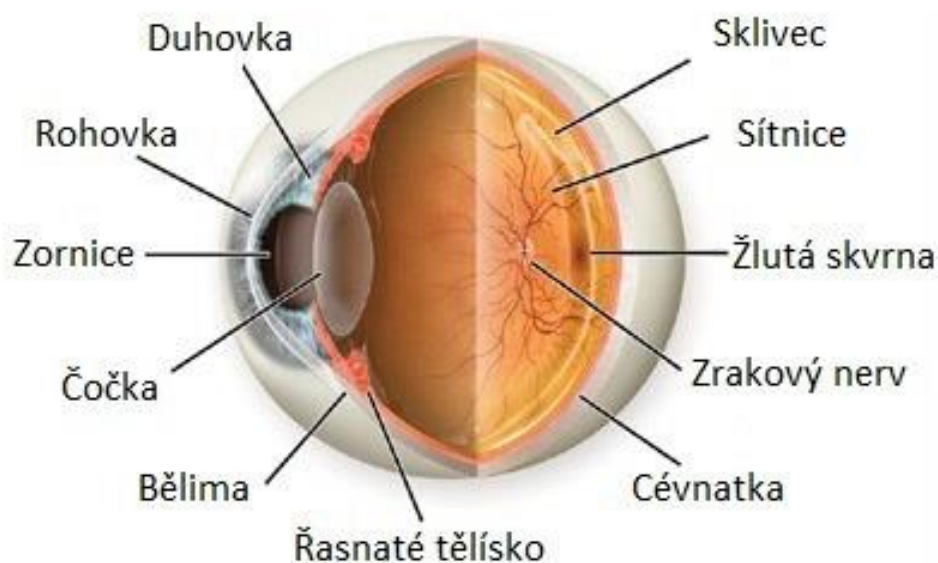
Lidský zrakový systém

Základem pro budování správné objektivní metody hodnocení a měření kvality obrazu je vyšetření lidského vnímání a zrakového systému, známého jako HVS (Human Visual System).

Lidský zrakový systém dává našemu tělu schopnost vidět fyzické prostředí kolem nás. Tento systém vyžaduje komunikaci mezi hlavním smyslovým orgánem (okem) a jádrem centrálního nervového systému (mozku) pro interpretaci vnějších podnětů (světelných vln) jako obrazů. Lidé jsou velmi vizuální tvorové ve srovnání s mnoha jinými zvířaty, která více spoléhají na čich nebo sluch a nad naši evoluční historii jsme vyvinuli neuvěřitelně složitý systém zrak.

1.1 Funkce oka

V HVS oko dostává fyzikální stimuly v podobě světla a posílá tyto podněty jako elektrické signály do mozku, kde se signály interpretují jako obrázky.



Obr. 1.1: Schématický řez lidským okem [2]

Na Obr. 1.1 je znázorněn příčný řez optikou lidského oka. Světlo dopadající na oko je zaostřováno pomocí čočky dále prochází sklivcem, což je průhledná výplň oka, a na konci své cesty dopadne na sítnici. Sítnice je velice důležitá tenká vrstva tkáně, která obsahuje fotoreceptory – tyčinky a čípky. Fotoreceptory obsahují fotonpigmenty rodopsiny a jodopsiny a kaskádní biochemický systém drah, umožňující vznik nervového impulsu, který dále převrací světlo na elektrické impulsy. Ty jsou přenášeny pomocí optického nervu do mozku.

Nerovnoměrně rozložené tyčinky a čípky na sítnici mají obrovský vliv na vnímaný obraz. Právě tyto buňky snímají dopadající světlo. Citlivé na jas tyčinky mají za úkol se starat o skotopické vidění, tj. vidění za špatných světelných podmínek. Čípky jsou barevně citlivé, zajišťující vidění za normálních světelných podmínek (fotopické vidění). Dělí se na 3 druhy podle vlnových délek, na které jsou citlivé: L-čípky (*long*, dlouhé vlnové délky), M-čípky (*medium*, střední) a S-čípky (*short*, krátké). Čípky dělí obraz projektovaný na sítnici na 3 vizuální části, které jsou ekvivalencí 3 barevných složek světla (*RGB*, červená, zelená a modrá). Člověk má kolem 6 milionů čípků (což je mnohem méně, než je tyčinek), mají nejvyšší hustotu rozložení ve foveě, která je na vizuální ose vidění. Vzdálením od fovey se zmenší koncentrace čípků a vzroste hustota rozložení tyčinek. Napodobně šíření neuronů přenášejících impulsy z oka do mozku přes optický nerv je největší ve foveě a pak jejich počet klesá.

Lidský zrakový systém není schopen prozkoumat celkový vizuální podnět rovnoměrně [3]. Fotopické vidění právě hraje největší roli při budování modelů tohoto systému. Stručnější popis daného procesu je líp vykládán v [1] a [2].

1.2 Psycho-fyzikální vlastnosti HVS

Lidské vidění je složitý systém složený z několika jevů ovlivňujících výsledný obrazek vjemu. Ty jsou skvělou napovědou při budování hodnotícího modelu. Základní jevy jsou probírány dále.

1.2.1 Foveální a periferní vidění

Z uvedených výše vlastností rozložení a vyskytu fotoreceptorů a neuronů je jasné, že rozlišení obrazu se mění se změnou vzdálenosti pozorovatele od pozorovaného bodu (s rostoucí vzdáleností klesá). To je důsledkem toho, že čípky a gangliové buňky nejsou rovnoměrně rozloženy na sítnici. Jejich hustota je největší v jamce. Tím pádem při oddalování od jamky se hustota fotoreceptorů snižuje. Jedná se o periferní vidění. Naopak přesně zaměřené na pozorovaný bod vidění se nazývá foveální.

I když v současné době objektivní metody pracují jen na principu foveálního vidění, ale jsou tady i metody, které si dokažou poradit i s periferním viděním za podmínky přijetí informace o směru, ve kterém se pozorovatel dívá. Z toho vyplývá možnost rozlišení, kam resp. na kterou část obrazu se díváme.

1.2.2 Přizpůsobení světlu

Přizpůsobení (adaptace) světlu je velice rychlý proces, který potrvá jen několik sekund, popř. desítek vteřin. Z velkého spektra hodnot intenzity světla lidské oko je citlivé na docela velké množství. To přizpůsobení funguje na principu regulací světla dopadajícího na oko přes zornicí. Na tom se podělí i adaptační mechanismy v buňkách

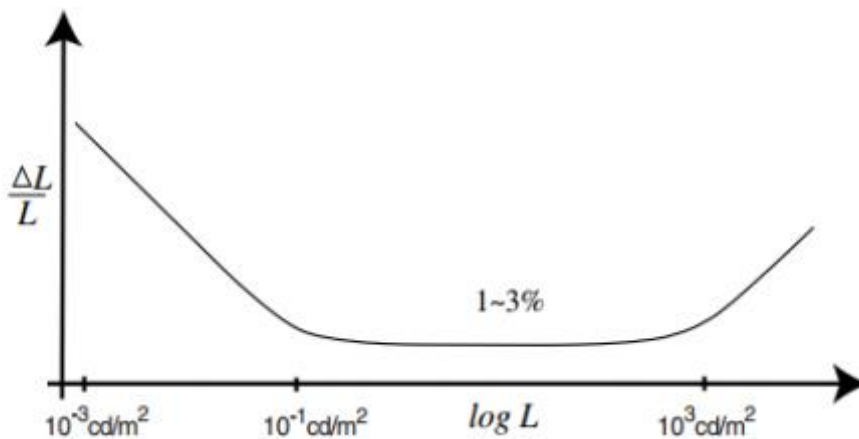
sítnice. Z toho principu dostáváme závislost Weber-Fechnerovo zákona ve tvaru [4]:

$$\Delta S = k \cdot \Delta L/L \quad (1.0)$$

kde ΔS vyjadřuje změnu intenzity subjektivního vjemu, k je Weberova konstanta, ΔL - nejméně viditelný rozdíl jasu na pozadí obrazu a L je fyzikální intenzita světla na receptor oka.

Weber-Fechnerův zákon popisuje přímou úměrnost mezi změnou požitku a intenzity podnětu a nepřímou úměrnost změny požitku a intenzity požitku.

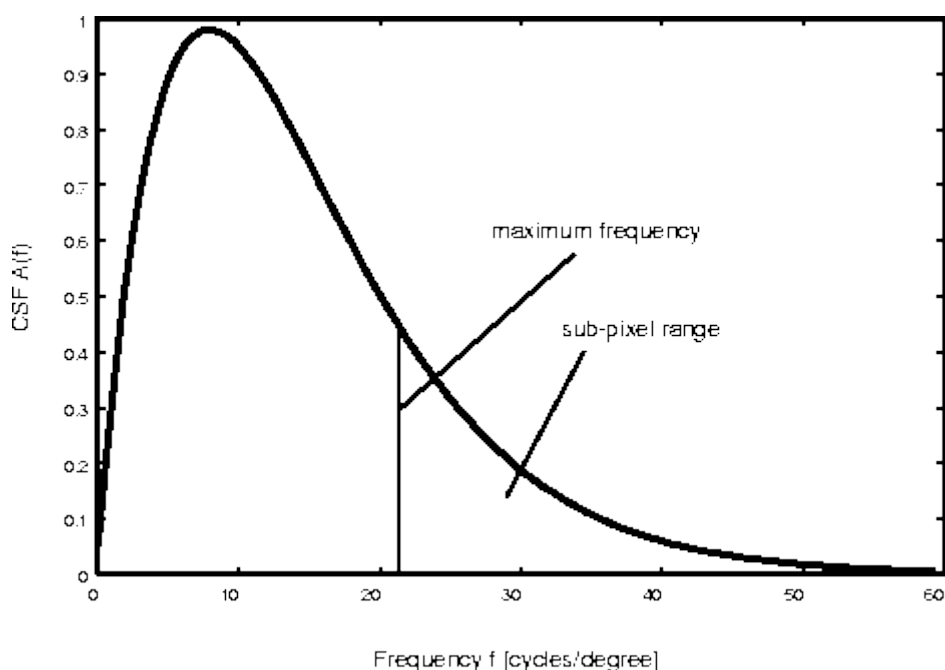
Jinými slovy subjektivně pozorovaná malá odchylka je přímo úměrná malé odchylce intenzity světelného paprsku a nepřímo - intenzitě paprsku [34]. K pochopení tohoto vztahu stáčí provést pár psychovizuálních testů, kde obraz je sestaven z pozadí se známým jasem L a nějaké vadné skvrny s jasem $L + \Delta L$. K všimání dané skvrny nejdřív se změní právě rozpoznatelný rozdíl JND podle závislosti $\Delta L/L$ (rychla změna parametru jasu z velmi tmavého do velmi světlého). Rozdíl jasů 2 % je postačitelny k vytvoření obrazového vjemu, ale osvětlení pozadí musí být z rozsahu 0,1 až 1000 cd/m². Je jasné, že lokální jas hraje významnou roli pro vnímání kvality obrazu [8].



Obr.1.2 Experiment pro Weber-Fechnerův zákon

1.2.3 Kontrast

Weber-Fechnerův zákon slouží jenom počátečním bodem při studiu přizpůsobení oka světlu, protože nebere v úvahu kompletní jasové vlastnosti. Přizpůsobení oka světlu prokazuje jeho citlivost na kontrast, ale to nezávisí jen na parametrech pozadí obrazu, a taky na ostatních jeho částech. Z toho vyplývá definice kontrastů v různých frekvenčních pásmech. V daném bodě se používá funkce kontrastní citlivosti (CFS, Contrast Sensitivity Function) z obrazku na obr. 1.3. Funkce kontrastní citlivosti nám říká, jak jsme citliví na různé kmitočty vizuálních podnětů. Pokud je frekvence vizuálního podnětu příliš vysoká, pak pozorovatel není schopen rozpoznat podnět. Převrácená hodnota kontrastního prahu se nazývá kontrastní citlivost HVS. [] Weber-Fechnerův zákon nám říká, že oko je nejvíc citlivé na středních frekvencích. Z Obr. 1.3 je vidět, že na vyšších frekvencích funkce citlivosti CSF rychle klesá, na nižších hodnotách frekvence naopak klesá pomalu.



Obr. 1.3: Grafické znázornění CSF [5]

V důsledku uvedených vlastností některé modely HVS považují funkci kontrastní citlivosti za filtr, jiné naopak používají CSF jako zatěž v subpásmech. CSF také bere v úvahu vzdálenost od jamky, ale pro foveální vidění (v blízkosti jamky) se CSF modeluje jako prostorově invariantní funkce pásma propustnosti. Proto pro metodu objektivního hodnocení kvality obrazu jsou zásadnější změny ve vzdálenosti pozorování.

1.2.4 Maskování

Masking je jednou z důležitých vlastností u sluchu a vidění, která představuje, že vjem původního signálu je potlačován jiným maskovacím signálem

(maskovačem). Na řadě s tím existuje spousta různých efektů pro maskování, jako například: maskování kontrastu, maskování aktivity, maskování barvy, časové maskování atd. Každá maskovač snižuje viditelnost původního signálu ve prospěch skutečnosti, kdy přítomna nebuda.

Má-li maska a signal stejný obsah a frekvenci, tak to usnadní detekci. Přesně z model těchto metod vznikly metoda hodnocení kvality obrazu. Nejvyužívanější je ta metoda při korekci a odstránění artefaktů.

Maskování kontrastu se využívá třeba v případě, kdy signál i maskovač mají skoro stejné vlastnosti nebo vzhled a lze tím se potlačit nežádoucí rušení originálu. Maskování barvy funguje na stejném principu, akorat poškození se nachází jenom v jednom kanálu. Pro lidský zrakový systém se využívá zejména zkrvení v oblasti pleťové barvy.

1.3 Oční pohyby

Problematika dané bakalářské práci vyžaduje nejen soustředění na jevech, které se dějou po dopadu na sítnici a zpracovávají přijatý signal, ale i to, kam je upoutána pozornost divaka.

1.3.1 Základy očních pohybů

Z důvodu potřeby získání celkové informace z pole, oči jsou v neustálém pohybu. Jsou několik druhů očních pohybů, známých jako přesun pohledu, fixace, stabilizace a adaptace. Za každý z pohybů odpovídají svaly řízené impulsy centrální nervové soustavy. Je jich 6 a stárají se o určitý pohyb jako např. nahoru, dolů, doleva, doprava a rotace v obou směrech optické osy. Lidská zraková soustava získá informaci, která by stáčila k docílení pozornosti v zorném úhlu. Navíc pohyby hlavy zvětší množství té informace a HVS bude mít kompletní představu o scéně. Střední výchylka očního pohybu je 15 %. Pokud zvětšíme výchylku do stran, zorné pole bude omezeno nosem.

1.3.2 Fixace

Z důvodu získání informace o pozorovaném objektu fixace je nejdůležitější druh očního pohybu. Zbylé druhy pohybů přesouvají oči pro fixaci na jiném objektu. Dojde k zastavení pohybu očí, následnému zaostření na object. Pohyb očí se úplně nezastaví, protože i tady jsou miniaturní pohyby, způsobené impulsy, které člověk nemůže ovládnout.

1.3.3 Sakády

Sakadický pohyb slouží k vyhledávání místa pro další fixaci, při němž dojde k rychlému (10 až 100 ms) přesunu oka bez adaptace zrakových drah, ke které dochází při dlouhém soustředění na konkrétní objekt. Jsou reflexivní (nemůžeme je ovlivnit) a vědomé.

1.3.4 Rozlišovací mezí oka

Schopnost oka vidět zvláště dva nacházející se blízko k sobě svítící body se označuje jako rozlišovací mezí oka. Svítící bod se zobrazí na sítnici normálního oka jako rozptylový kroužek. Dva body se dají vzájemně rozeznat jen tehdy, když je na sítnici oka mezi jejich rozptylovými kroužky alespoň jeden čípek nezasážený světlem [38].

Kapitola 2

Subjektivní metody hodnocení kvality obrazu

Pro návrh a následovné testování nové metody je nezbytné mít k dispozici další metodu, která by hrála roli reference. Z důvodu neexistence jediné univerzální objektivní metody hodnocení obrazu je potřeba získat hodnoty subjektivních testů a použít je jako původní výsledné hodnoty k porovnání kvality obrazu.

Existuje několik populárních subjektivních metod hodnocení kvality obrazu, které budou rozebrány níže.

2.1 Subjektivní testy

Během subjektivních testů skupina pozorovatelů sleduje a hodnotí kvalitu promítaného obrazu, což je v popisu doporučení pro určení subjektivní kvality obrazu [6]. Pozorovatelé k usnadnění hodnoticího procesu dostanou speciální stupnici, na základě které budou porovnovat a hodnotit obraz.

Subjektivní metody se dělí na dvě skupiny: metody DS (provádí se hodnocení dvojicí scén) a metody SS (hodnocení samostatného obrazu). Seznam dalších subjektivních metod je uveden v [7].

2.2 Metody DS (Double Stimulus)

2.2.1 DSIS

Pozorovatelé opakovaně uvidí dva obrazy pro různou scénu, přičemž pár obrazů obsahuje referenční a testovaný poškozený obraz. K získání výsledných hodnot se používá stupnice od 5, které odpovídá nepostřehnutelnému zkreslení do 1, což je velmi nepříjemné zkreslení. Daná metoda se často používá v případech změny takových parametrů jako jsou bitová rychlost nebo kompresní formát. V průběhu se odhalí, kdy nastane zlom ve vnímání kvality. Metodu DSIS je možné použít k hodnocení širokého rozsahu zkreslení.

2.2.1 DSCQS

Měření u metody DSCQS se na rozdíl od DSIS metody provádí na kontinuální stupnici hodnot od 0 do 100. Pozorovatelé hodnotí obrazy A a B a předem nevědí, který je referenční a který je zkreslený. Během testu pozorovatelé

hledají rozdíly mezi dvěma promítanými obrazy. Z toho důvodu metodu DSCQS se dá použít k hodnocení obrazů, které jsou málo poškozené.

2.3 Metody SS (Single Stimulus)

2.3.1 SSM

Pozorovatelům seriově ukazují různé nezávislé scény. Pak na základě dvou přístupů - SS (neopakuje se žádná scéna) a SSMR (scény jsou opakovány vícekrát) se používají další metodiky. To jsou metodika nepřímá (adjectival) - zahrnuje 5-ti stupňovou stupnici, možný krok hodnocení je 0,5 stupně a metoda nekategorická (non-categorical), což je plynulá stupnice bez čísel nebo s velkými obsahy jako např. 0-100.

2.3.2 SSCQE

Tentokrát pozorovatelé mají za úkol hodnotit jenom testovací scény. Provádí se i průběžné hodnocení kvality obrazu a pozorovatelé mají na to posuvné ovládací prvky se spojitou stupnicí (0 odpovídá nejhorší kvalitě, 100 naopak nejlepší kvalitě). Hodnocení obrazu trvá poměrně dlouhou dobu - 10 až 20 min, záznamové zařízení zapisuje výsledky každé poloviny sekundy. Metoda SSCQE se skvěle hodí k testování proměnné v čase kvality digitálních systémů pro kódování videa (MPEG-2, MPEG-4 atd) [36].

2.4 Výhody a nevýhody subjektivního testování

Výhodou subjektivního testování kvality obrazu je skutečnost, že výsledky dané metody přesně odpovídají a nachází se v souladu s lidským zrakovým systémem, protože tyto testy nejen tolerují HVS, ale jsou přímo na něj zaměřené. Z hodnot výsledků subjektivních testů je možné vypočítat skalární hodnotu parametru MOS (Mean Opinion Score).

Nevýhodou dané metody je časová a finanční náročnost, nutnost dodržení předepsaných pozorovacích podmínek a přísný výběr pozorovatelů. Navíc výsledky těchto testů většinou platí jenom pro dané pozorovací podmínky a daný typ scény[36].

Kapitola 3

Objektivní metody hodnocení kvality obrazu

Již od začátku vyvoje televizního vysílání se pokladaly otázky přizpůsobení obrazu neslyšícím divákům a výroby hodnoticích metod a modelů pro zjištění kvality obrazu. Tento problém zvedl svoji pozici v seznamu nových cílů v důsledku rychlého rozvoje telekomunikační stranky v oblasti digitální techniky. Je potřeba poskytovat vyšší kvalitu obrazu s rozšířením možnosti a vlastnosti přijímací a vysílací techniky. Tím pádem se rozvíjí i metody hodnocení obrazu.

Nabídku objektivních metod hodnocení obrazu lze rozřadit podle typu a způsobu zpracování dat a rozsahem vstupních parametrů do tří modelových skupin: metrika bez reference NR (no reference), s částečnou referencí RR (reduced reference) a s plnou referencí FR (full reference).

3.1 MSE a PSNR

Je potřeba na začátku zmínit, že existují metody (metriky), které neberou v úvahu a nerespektují lidský zrakový systém. Nejvýznamnější z těchto metric jsou tzv. pixelově orientovaná plnoreferenční metrika MSE (Mean Squared Error) a metrika PSNR (Peak Noise to Ratio). Jsou stejně nejvíc používané pro hodnocení kvality kvůli chybám vyskytujícím se při zpracování lidského zrakového systému.

V základu pixelově orientovaných metrik leží tzv. střední kvadratická chyba MSE, která je definována jako střední hodnota druhých mocnin rozdílů dvou šedotónových hodnot pixelů obrazu [36]. V případě zpracování třírozměrného obrazového signálu definujeme chybu MSE vztahem [9]:

$$MSE = \frac{1}{TXY} \sum_t \sum_x \sum_y [I(t, x, y) - \hat{I}(t, x, y)]^2, \quad (3.0)$$

kde I a \hat{I} odpovídají hodnotám jasů dvou obrazů, které mají rozměry x , y a t . Jestli I považujeme za původní nezkrácený signál, tím pádem \hat{I} je jeho zkrácená verze, potom lze kvadratickou chybu MSE brát ve funkci měřítka kvality obrazu. Kvůli rozdílu v obsažené informaci u obou signál vznikne chyba, hodnotě které odpovídá výraz $I - \hat{I}$.

Takový přístup hodnocení se dá použít i u barevných obrázků, v tomto případě se výše uvedený vztah aplikuje pro výpočet každé barevné složky obrázku a

následovně bude upraven přidáním sumy, která sčítá chyby u jednotlivých složek[36].

Druhá rozebíraná metrika PSNR (špičkový poměr signálu k šumu) je odvozená z metriky MSE. Hodnota PSNR se většinou uvádí v decibelech a je definována následujícím vzorcem:

$$PSNR_{dB} = 10 \cdot \log_{10} \left(\frac{m^2}{MSE} \right), \quad (3.1)$$

kde m reprezentuje maximální hodnotu, kterou může nabývat pixel (např. 255 pro 8-bitový obraz). Kvalitní standardní obraz má hodnotu PSNR kolem 50 dB.

Metriky MSE a PSNR se hodí k hodnocení kvality obrazu díky své dobré rychlosti výpočtů hodnot a přehlednosti. Ale pro případ hodnocení obrazu mluvčího ZJ nejsou tyto metriky použitelné, protože netolerují a neberou v úvahu lidský zrakový systém. Při porovnání těchto metrik a subjektivních testů narážíme na korelační hodnoty v rozsahu $r = 0,4 - 0,7$, což je důsledkem zpracování všech numerických chyb se stejnou vahou.

Příkladem takových chyb je šum ve členité oblasti obrazu (např. obraz se skalami) nebo v pozadí snímků. Divák si takové chyby nevšimne, ale metrika PSNR ji detekuje a tím ovlivní výslednou hodnotu kvality obrazu, čímž ji zhorší [9], [10].

3.2 No reference metody (NR)

Z názvu dané skupiny metod je jasné, že pro určení kvality obrazu nepotřebují mít žádné záznamy o původním snímku nebo obrazu, což je mnohým lepší nabídkou, protože mizí nutnost mít najednou původní a testovaný obraz. Ale vzniká problém v hledání způsobu rozlišování oblasti obrazu s a bez rušení, z čeho vyplývá další problém – nedá se navrhnout univerzální bezreferenční metrika. Je dobré, že metody bez reference se dají použít při vyhledávání rušení, ale je nutné předem vědět, co je nejvýznamnějším bodem hodnocení – šum, artefakty atd.

Stejně metriky NR jsou docela populární právě z toho důvodu, že nevyžadují informaci o referenčním obrazu a nejvíce se hodí v případě hodnocení zkráceného obrazu, přičemž referenční obraz chybí. Je to také jediná možnost, jak zhodnotit obrazy, kde již originál není k dispozici [12], [13], [14] a [15].

3.3 Reduced Reference metody (RR)

Metoda s částečnou referencí nese v sobě vlastnosti metod bez reference a s plnou referencí. Metoda RR má jenom částečnou informaci o původním obrázku, kterou přenáší spolu se zkráceným obrazem. Tato informace ve výsledném výpočtu slouží k hodnocení kvality obrazu.

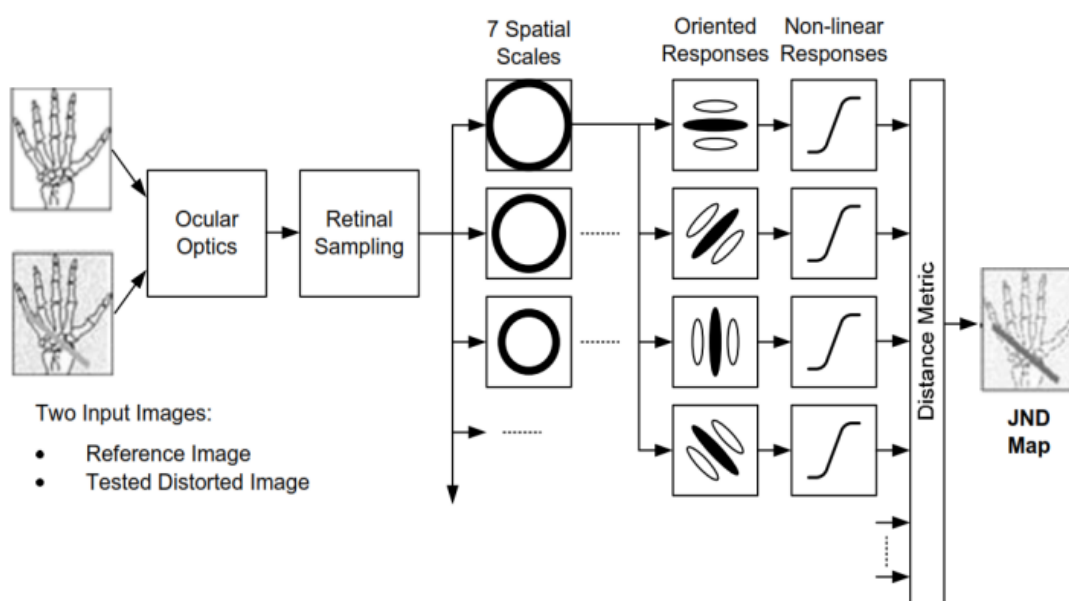
Metoda hodnocení obrazu s částečnou referencí se používá pro získání informací o hranách, o míře pohybu atd. z referenčního obrazu [36]. Všechna získaná informace se pak použije k hodnocení testovaného obrazu [13],[14].

3.4 Metody měření s plnou referencí (FR)

Metody pro měření kvality obrazu s plnou referencí se vyznačují tím, že jsou nejpoužívanější a již v současné době dobře rozvinuté. Pro svoji správnou funkci vyžaduje přítomnost obou referenčního a testovaného obrazu. Metody s plnou referencí je možné porovnat se subjektivními metodami DS, což na druhou stranu hodnotí a porovnává testované FR metody. Z důvodu velkého množství nabízených FR metod, které řeší spoustu různých problémů je potřeba zmínit o těch nejdůležitějších, svým způsobem se podělily na vývoji metody objektivního hodnocení obrazu.

3.4.1 Model lidského zrakového systému od Lubina

Za prvé zmíníme o multikanálovém modelu Lubina, který se dá nejlíp pochopit z vlastního blokového schéma na obr. 3.1[16], [17]:



Obr. 3.1: Příslušné blokové schéma pro model Lubina[18]

Jak je vidět z obrázku, celý systém se spouští pomocí optických a vzorkovacích procesů. Je potřeba nastavit vstupní procesy takovým způsobem, aby splňovaly určité podmínky jako jsou třeba pozorovací vzdálenost či vlastnosti monitoru. Pak se vstupní signály odebírají, což připomíná princip snímání obrazu fotoreceptory na sítnici lidského oka.

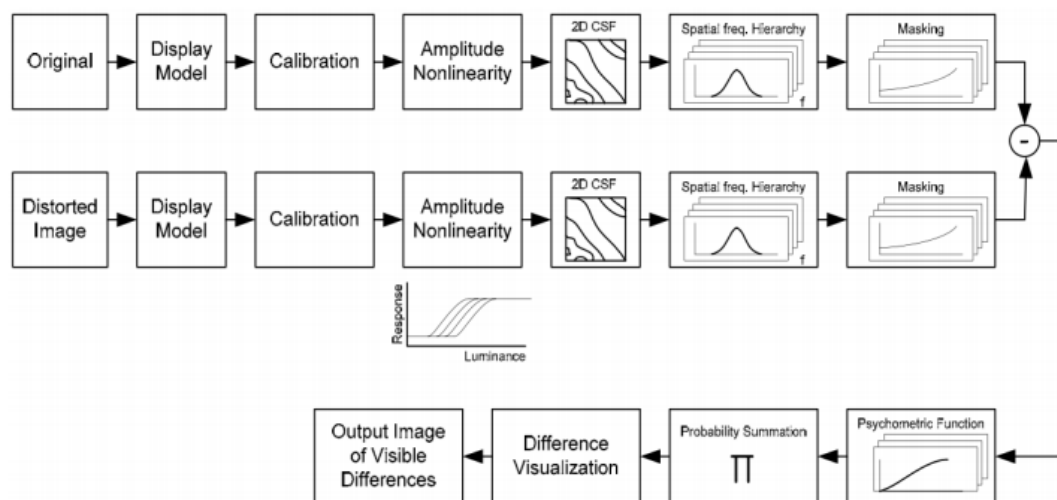
Dále přichází okamžik, kde se použije Laplaceova pyramida s cílem dekomponovat obraz na sedm jiných s jiným rozlišením. Hodnota každého nového rozlišení je pokaždé o polovinu nižší na rozdíl od vyššího rozlišení na stupnici. “Jasový signál je dále konvertován na lokální kontrast pomocí techniky podobné Peliho LBC algoritmu”[36]. Pak se objeví osm směrových filtrů, zajišťujících selektivitu ve čtyřech orientacích. Tím pádem ve výsledku budeme mít 28 kanálů modelu, v každém ze kterých se provádí maskování sigmoidní nelinearitou. Sigmoidní nelinearita funguje na principu simulaci vyhledávání dipper efektu (pokles hodnot). Na výstupu se všechny chybové hodnoty zapíšu do distorzní mapy, v každém kanálu se používá na to Minkowského simulace s exponentem 2.4. Dostáváme výslednou JND mapu chybových hodnot. Zbývá jenom získat jedno číslo, které by reprezentovalo výslednou kvalitu obrazu. Toto číslo se pak porovnává s výsledkami subjektivních testů.

Na konci experimentu [18] obdržené výsledky se vyznačují dobrou korelací se subjektivními testy ($r = 0,94$). Pro porovnání metrika MSE má nižší hodnotu korelace ($r = 0,81$).

Experimenty provedené později ukázali, že Lubinův model víc toleruje poškození přítomnost chyb v důsledku práce s moderními obrazy. Lubinův model např. se používá k zjištění kvality rentgenových snímků [18].

3.4.2 VDP (*Visible Differences Predictor*)

VDP je Dalyho multikanálový model hodnotící kvality obrazu[5], [20], [36]. Hned se budeme věnovat blokovému schemata, co je na obr. 3.2.



Obr. 3.2: VDP multikanálový model [5]

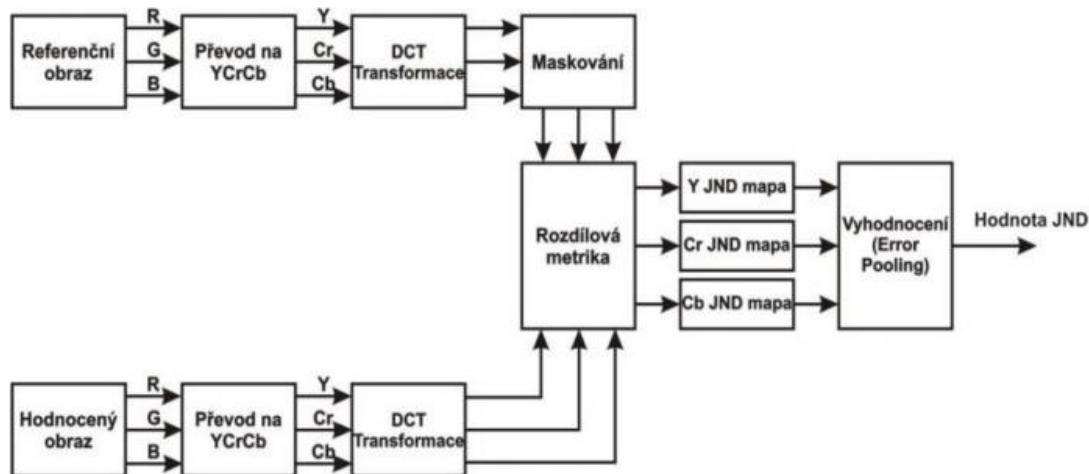
Jako u Lubinovo modelu HVS, v případě VDP modelu je taky nutno připravit system nastavením určitých parametrů a podmínek. Nasleduje řada procesu aplikovana na referenční a testovaný obraz, jako jsou jasová nelinearita, selektivní kanálová dekompozice, modelování maskovacího efektu, CSF filtr.

Kanálová dekompozice přináší na výstupu 31 nezávislý kanál (5 prostorových úrovní krát 6 úrovní orientace + základní pásmo) [36]. Maskovací proces používá prahový zdvih, který pro každý kanál najde z CSF práh viditelnosti na základě velikosti hodnoty kontrastu v příslušném kanálu. Jako maskovače pro maskovací proces jsou použity přesně referenční i testované obrazy. Realizace je správná a má smysl jenom když chyba, vyskytující se v obrazu je maskována jak referenčním, tak i testovaným snímkem (ale záleží na typu zkruslení). Dále se pro každý kanál vytvoří mapa pravděpodobnosti detekce chyby, ale předtím se spočítají veškeré chyby a se porovnávají s prahem viditelnosti. Mapy pravděpodobnosti se vytvoří díky psychometrické funkci. Vzniká celková mapa JND hodnot, každá hodnota které odpovídá nějaké pravděpodobnosti, se kterou pozorovatel vníma rozdíl mezi původním a změněným obrazem.

U přirozených obrazu je velmi těžko zkontrolovat pravdivost JND mapy. Autor dané metody doporučoval detekci maskovacích prahů provádět s použitím špičkových hodnot z VDP mapy [21]. “Díky tomuto přístupu Daly dokázal, že model je schopen správně predikovat data vycházející z širokého rozsahu psychofyzikálních experimentů [8]” [36].

3.4.3 Watsonův DCT model hodnocení obrazu

Další plnreferenční model hodnocení obrazu funguje na principu DCT transformaci a jeho blokové schéma je představeno na obr. 3.3.



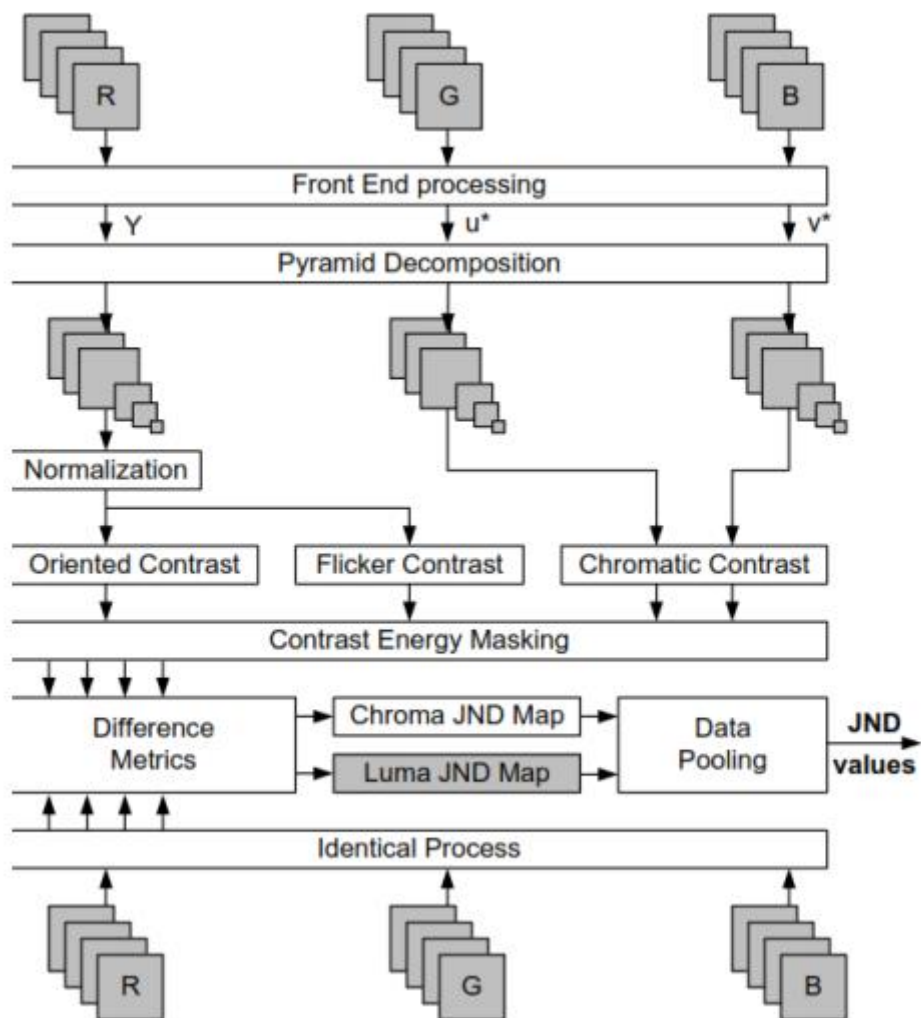
Obr. 3.3: DCT modelu ve frekvenční oblasti [23]

Na začátku se každá složka barevného obrazu $YCrCb$ rozloží do bloků o velikosti 8×8 pixelů, v každém z nichž se provádí DCT transformace, která produkuje frekvenční koeficienty z referenčního a testovaného obrazu. Prahová úroveň, kterou se dá odvodit i z dalších parametrů jako jsou kontrastní citlivost a jasové maskování se určuje z koeficientů, získaných z referenčního obrazu. Na každý frekvenční koeficient obrazu se aplikují jasové a kontrastní maskování. Tím pádem každý koeficient z každého bloku získává vlastní nastavený prah vnímání.

Zbývá jenom spočítat procento poškození každého jednotlivého bloku, ta hodnota je pak váhována celkovým prahem citlivosti. Výstupem je, jak i v předchozích případech JND mapa, která je vypočtená ve 2 krocích. Nejdřív se počítá suma chyb frekvenčních koeficientů pro každý bod a kanál. Potom matice chyb se vyhodnotí přes frekvence. Tak se dosáhne hodnoty celkové chyby vnímání pro tři kanály (Y , C_r a C_b), po sečtení kterých vznikne celková chyba JND.

3.4.4 Sarnoffův JND model

Tentokrát se seznámíme s nejdůležitějším multikanálovým modelem hodnocení kvality barevných obrazů, což je Sarnoffův JND model vidění [24], [25]. Podíváme se na blokové schéma modelu.



Obr. 3.4: Blokové schéma Sarnoffova multikanálového modelu [24]

Daný model hodnocení obrazu byl navržen a zpracován výzkumnou skupinou kolem Lubina v Sarnoffových laboratořích (odtud pochází název) v Princetonu (USA) společně s firmou Tektronix. JND model je pokračováním původních

Lubinových modelů (viz obr. 3.1) a se hodně liší od předchozích modelu dle způsobu zpracování obrazu.

Nejdřív metoda rozdělí každý obraz na 3 větve, které reprezentují jasová a dvě chrominanční složky. Prvním úkolem je transformovat složky Y' , C_r' , C_b' na signály R' , G' , B' . Bodová nelinearita se postará o odstránění nízkých hodnot jasu v každé ploše zobrazovače. "Hodnoty z barevného prostoru R , G , B jsou poté lineárně transformovány přes barevný prostor XYZ do konečného barevného prostoru CIE L^* , u^* , v^* (Wyszecki a Stiles, 1982). Větev L^* prochází jasovým zpracováním a větve u^* , v^* chrominančním zpracováním"[36].

Každá jasová hodnota musí projít kompresní nelinearitou a následovně se rozdělí na prostorové frekvence v důsledku 4úrovňové Gaussovské pyramidové dekompozici. Pak každá úroveň se vyfiltruje, spočítá se kontrast a provede se maskování. To je co se týká jasového zpracování.

Chrominanční zpracování má na rozdíl od jasového zpracování paralelní charakter v různých směrech. Odlišnost chrominančních složek u^* , v^* prostoru CIE LUV se používá k definici detekčních prahů pro chrominanční model. Jako u jasového modelu jsou chrominanční kontrasty, definované s použitím diferencí u^* , v^* podléhají aplikaci maskování. Chrominanční zpracování překonalo jasové zpracování z toho důvodu, že v chrominančním případě se provádí až 7úrovňová pyramidová dekompozice. Ve výsledku se pozoruje, že chrominanční kanály jsou citlivější na nižších prostorových frekvencích ve srovnání s jasovými kanály. Pak pro vytváření barevných rozdílů se používá prostorová filtrace Laplaceovým jádrem. Každé hodnotě barevného rozdílu se přiděluje určitá váha, která je pak předána do části maskování kontrastu. Hodnota váhy má stejnou funkci jako v jasovém zpracování. Na výstupu obdržíme JND mapu, která je sestavená z jasové a chrominanční části. Navíc JND model vytváří jediné číslo Q_{norm} , udávající výslednou kvalitu obrazu.

3.4.5 SSIM (Structural Similarity Index)

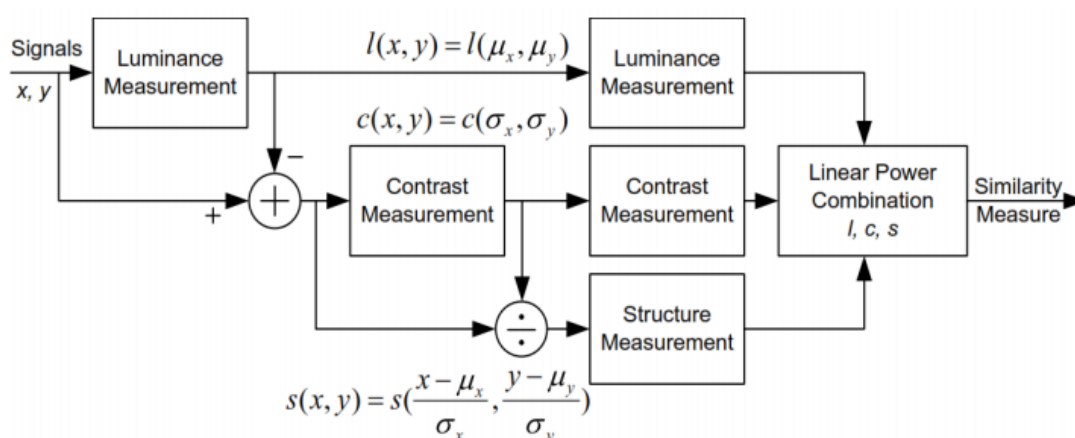
Obecně řečeno, že SSIM se používá pro měření podobnosti mezi dvěma obrazy. Index SSIM spadá do FM (Full Reference) třídy metod hodnocení kvality obrazu. Samotné měření nebo predikce kvality obrazu je založeno na počátečním nekomprimovaném (nezkresleném) obrazu, který se považuje za referenci. Metoda strukturální podobnosti je navržena s cílem zlepšení tradičních metod, např. jako psnr (PSNR) a znamenají čtverečkováné chyba (MSE), které se ukázaly být v rozporu s lidskou zrakového vnímání.

Metoda SSIM, která vznikla pod vedením Alana C. Bovika a Zhou Wang [26], [27] na rozdíl od již zmíněných výše metrik nepoužívá modelování pomocí bloků k výpočtu kvality obrazu. V tomto případě se vystačíme se statistickými výpočty k získání informací o struktuře a podobnosti mezi referenčním a testovaným obrazem. Samotné měření nebo predikce kvality obrazu je založeno na počátečním nekomprimovaném (nezkresleném) obrazu, který se považuje za

referenci. Metoda strukturální podobnosti je navržena s cílem zlepšení tradičních metod jako PSNR (SSIM částečně bere v úvahu model HVS[36]).

Stavíme konkrétní příklad měření kvality SSIM z hlediska vytváření obrazu. Potřebujeme zobecnit algoritmus a tím poskytnout rozsáhlejší výsledky ověřování. Jas povrchu pozorovaného objektu je produktem osvětlení a odrazu, ale samotné struktury objektů scény jsou nezávislé na osvětlení. Tudiž k prozkoumání strukturální informace obrazu potřebujeme oddělit vliv osvětlení. Definujeme strukturální informaci obrazu jako parametry představující strukturu objektů scény nezávisle na průměrném jasu nebo kontrastu. Vzhledem k tomu, že hodnota jasu a kontrastu se může lišit přes celou scénu, k definici používáme ponětí místního jasu a kontrastu.

Blokové schéma modelu SSIM je na obr. 3.5.



Obr 3.5: Blokové schéma SSIM modelu [26]

Předpokládáme, že \mathbf{x} a \mathbf{y} jsou dva nezáporné obrazové signály, které se navzájem překrývají (např. prostorové skvrny získané z každého snímku). Pokud považujeme jeden ze signálů za signál z dokonalou kvalitou, pak měřitko podobnosti může sloužit jako kvantitativní měření kvality druhého signálu. Systém dělí úkol měření podobnosti na 3 porovnání: jas, kontrast a struktura.

Za prvé, porovnává se jas obou signálů. Za předpokladu diskretnosti signálů se stanoví střední intenzita:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i. \quad (3.2)$$

Porovnovací funkce jasu $l(\mathbf{x}, \mathbf{y})$ je pak funkcí μ_x a μ_y a je definována jako:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (3.3)$$

kde konstanta C_1 je zavedena aby nedošlo k nestabilitě, když $\mu_x^2 + \mu_y^2$ se blíží k nule. Konkrétně zvolíme:

$$C_1 = (K_1 L)^2, \quad (3.4)$$

kde L je dynamický rozsah hodnot pixelů (255 pro osmibitový černobílý obraz), a $K_1 \ll 1$ je malý kontrast [27].

Používáme směrodatnou odchylku (druhá odmocnina rozptylu) k odhadu kontrastu signálu. Nezaujatý odhad v diskrétní formě je dan vztahem:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}}. \quad (3.5)$$

Srovnání kontrastu dvou obrazů poté provede funkce $c(x, y)$ na základě porovnání hodnot jejich směrodatných odchylek σ_x a σ_y následujícím způsobem:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (3.6)$$

kde $C_2 = (K_2 L)^2$ a $K_2 \ll 1$.

Důležitou vlastností dané funkce je, že se stejnou hodnotou změny kontrastu $\Delta\sigma = \sigma_y - \sigma_x$, tato míra je méně citlivá pro případ vysokého základního kontrastu σ_x ve srovnání s nízkým základním kontrastem. To je v souladu s rysem kontrastního maskování v HVS.

Srovnání struktury se provádí po odečtení jasu a rozptylové normalizaci. Konkrétně spojíme dvě vektorové jednotky $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$, každá z nich leží v nadrovině definované v (3), se strukturou dvou obrazů. Korelace (vnitřní produkt) mezi nimi je jednoduchá a účinná míra pro kvantifikaci strukturální podobnosti. Musíme si všimnout, že korelace mezi $(\mathbf{x} - \mu_x)/\sigma_x$ a $(\mathbf{y} - \mu_y)/\sigma_y$ je ekvivalentní korelačnímu koeficientu mezi \mathbf{x} a \mathbf{y} . Tedy, definujeme funkci strukturálního porovnání následujícím způsobem:

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (3.7)$$

Konstanta $C_3 = (K_3 L)^2$ slouží, podobně jako konstanty v předchozích případech, k zajištění stability výpočtu. Stejně jako pro míru jasu tak i kontrastu obrazu, jsme zavedli malou konstantu v čitateli a jmenovateli. V diskrétní formě σ_{xy} odhadujeme jako:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y). \quad (3.8)$$

Geometricky, korelační koeficient odpovídá kosinu úhlu mezi vektory $\mathbf{x} - \mu_x$ a $\mathbf{y} - \mu_y$. Poznamenejme si, že funkce $s(\mathbf{x}, \mathbf{y})$ může být určena i v případě záporných i .

Nakonec spojíme tři porovnání a pojmenujeme výslednou míru podobnosti jako index SSIM mezi signály \mathbf{x} a \mathbf{y}

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma, \quad (3.9)$$

kde $\alpha > 0$, $\beta > 0$ a $\gamma > 0$ jsou parametry určené k upravení relativního významu tří složek. Je snadno ověřit, že tato definice splňuje tři podmínky uvedené výše. S cílem zjednodušení výrazu, nastavíme $\alpha = \beta = \gamma = 1$ a $C_3 = C_2/2$ v této práci. To má za následek specifickou formu indexu SSIM:

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3.10)$$

Metoda SSIM je symetrická a z toho důvodu nezáleží na tom, který z načítaných obrazů ze dvou je referenční a který je testovaný. Je omezená v rozsahu hodnot $-1 < SSIM(x, y) \leq 1$, kdy 1 odpovídá skutečnosti, že oba porovnávané obrazy jsou stejné, a naopak čím více se blížíme k -1, tím jsou obrazy víc rozdílné.

Místní statistiky μ_x , σ_x a σ_{xy} jsou počítány v rámci lokálního 8x8 čtvercového okna, které se pohybuje pixel-po-pixelu přes celý obraz. Na každém kroku lokální statistika a SSIM index se vypočítávají v místním okně. Jedním z problémů dané metody je, že mapa výsledného indexu SSIM často vykazuje nežádoucí "blokovací" artefakty. V tomto článku používáme 11x11 kruhově symetrickou Gaussovou funkci vážení $\mathbf{w} = \{w_i | i = 1, 2, \dots, N\}$ se standardní odchylkou 1,5 vzorků, normalizovanou na jednotky součtu ($\sum_{i=1}^N w_i = 1$). Odhady místních statistik μ_x , σ_x a σ_{xy} jsou modifikovány jako:

$$\mu_x = \sum_{i=1}^N w_i x_i, \quad (3.11)$$

$$\sigma_x = \left(\sum_{i=1}^N w_i (x_i - \mu_x)^2 \right)^{\frac{1}{2}}, \quad (3.12)$$

$$\sigma_{xy} = \sum_{i=1}^N w_i (x_i - \mu_x)(y_i - \mu_y). \quad (3.13)$$

Pomocí takového okenního přístupu mapy kvality místně vykazují izotropní vlastnosti. V této práci při měření SSIM se používá následující nastavení parametrů: $K_1 = 0.01$; $K_2 = 0.02$; $K_3 = 0.03$. Tyto hodnoty jsou poněkud libovolné, ale zjistíme, že v našich experimentech, výkonnost algoritmu indexu SSIM je poměrně necitlivý na kolísání těchto hodnot.

V praxi pozorovatel obvykle vyžaduje jednotné celkové měření kvality celkového obrazu. Používáme střední (průměrný) SSIM index (MSSIM) k ohodnocení celkové kvality obrazu.

$$MSSIM(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^M SSIM(\mathbf{x}_j, \mathbf{y}_j) \quad (3.14)$$

kde \mathbf{X} a \mathbf{Y} jsou referenční a zkreslený obraz, v uvedeném pořadí; \mathbf{x}_j a \mathbf{y}_j jsou obsah obrazu v j -tém místním okně; M je počet lokálních oken obrazu. V závislosti na aplikaci, je také možné vypočítat vážený průměr různých vzorků v mapě indexu SSIM. Například systémy pro zpracování obrazu přes oblast zájmu dávají různé váhy pro různé segmentové oblasti obrázu.

Kapitola 4

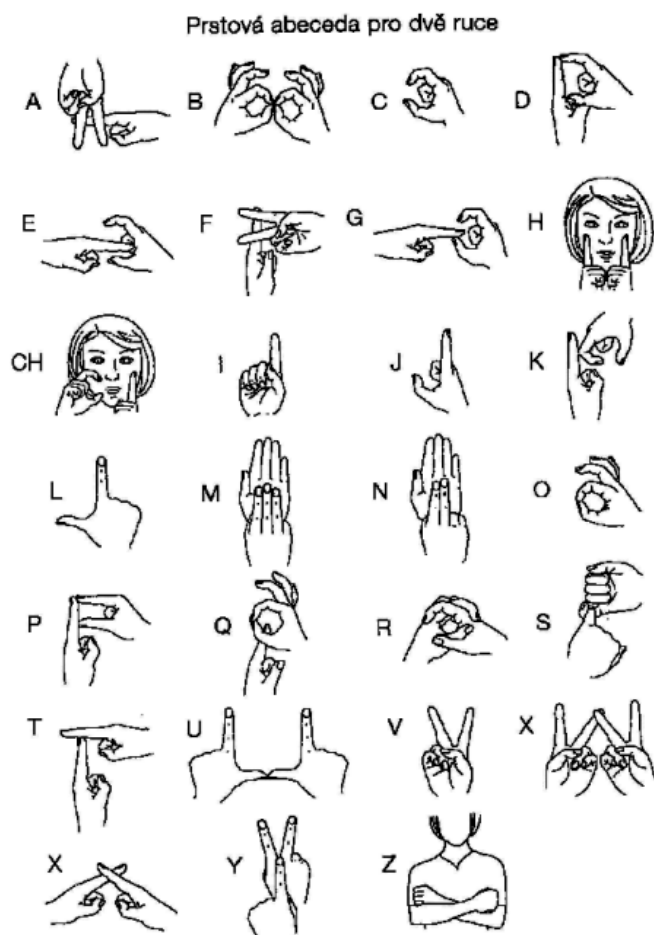
Charakteristika obrazu mluvího znakového jazyka

Základem správného návrhu metody objektivního hodnocení obrazu znakového mluvího je zjistit, jakým částem promítaného obrazu neslyšící divák věnuje největší pozornost.

4.1 Znakový jazyk

Znakový jazyk se považuje za plnohodnotný komunikační systém pro neslyšící, který je vytvořen tvary a pohyby rukou, mimikou obličeje, hlavou a také neodmítá použití horní části trupu. Tento způsob komunikace používají nejen neslyšící lidi a lidi s poruchou řeči, ale také tlumočníci nebo rodinné příslušníci neslyšícího člověka.

Znakové jazyky se liší podle zemí, jako i mluvené jazyky, což je příčinou vzniku českého znakového jazyka a spousty dalších ZJ. Abeceda ČZJ (české znakové jazyka) je představena na obr. 4.1.



Obr. 4.1: Abeceda českého znakového jazyka

Průzkum Českého statistického úřadu z roku 2007 ukázal, že v České republice bydlí skoro 26 tisíc těžce nebo velmi těžce sluchově postižených lidí [28] [36]. Z toho 8 % neslyšících lidí nebylo spokojeno s nabídkou kompenzačních pomůcek pro život. Za největší omezení zúčastnění neslyšící občani považují získávání informace a komunikaci, což je důsledkem rozdílu v komunikaci slyšícího a neslyšícího člověka.

Znakový jazyk je vnímán převážně pomocí zrakového systému a proto je to vizuálně-motorický jazyk, který se hodně liší od běžného mluveného jazyka. Neexistuje psaná forma znakového jazyka, protože není možné zachytit všechny tvary, pohyby a orientaci a z toho důvodu je snadněji natočit mluvčího ZJ na kameru a uložit videosekvenci nebo série obrazů.

4.2 Oblasti zájmu neslyšících pozorovatelů

Hlavním vychozím bodem je skutečnost, že neslyšící pozorovatel vnímá obraz jinak, než běžný slyšící divák, který používá při pozorování scény sluch i zrak. Zajímavé je, že tyto smysly řídí různá mozková centra a z toho důvodu jsou vnímány paralelně. Když je neslyšící divák závislý na zraku a nemůže si dovolit přestat sledovat scénu, aniž by přišel o informaci, běžný divák se spolehá na

sluch. Takže při odvrácení zraku stále přijímá informaci, což je důkazem odlišnosti ve vnímání stejné scény obou diváků.

Neslyšící pozorovatel nemůže vnímat znakovou řeč seriově, protože všechny komponenty se uskutečňují najednou – pohyby rukou, poloha hlavy, obličej. Bodem nejvyšší pozornosti neslyšícího diváku je konkrétní část obrazu, kterou vnímá foveálním viděním. Ostatní části obrazu se vnímají periferním viděním. Z toho faktu vyplývá nutnost rozdělení obrazu podle důležitosti při pozorování scény na oblasti zájmu tzv. ROI, což je doba, kterou divák věnuje určité části obrazu.

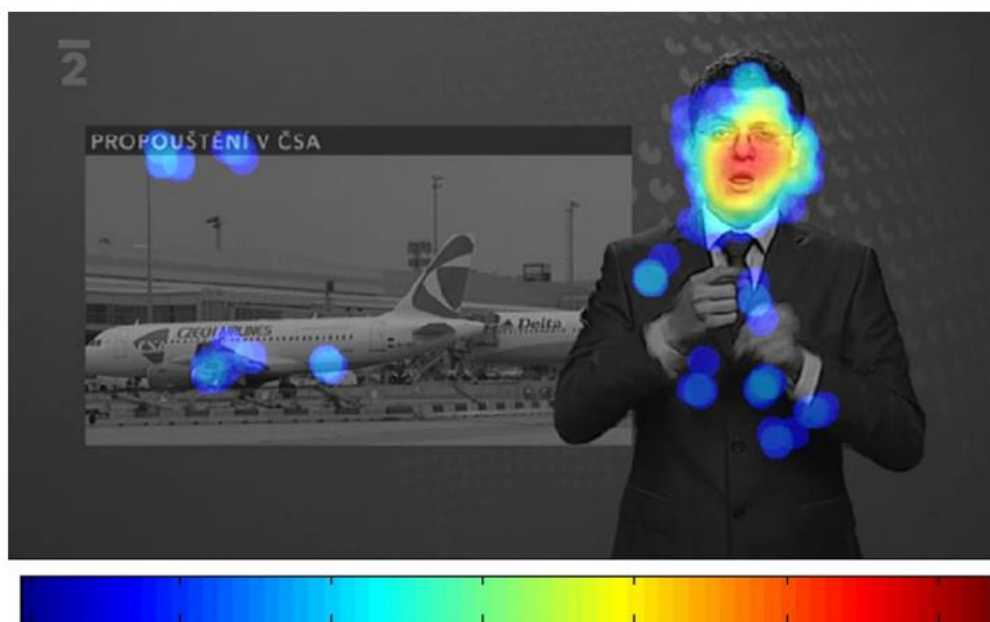
Zkoumání oblasti ROI v obrazu směřuje nejen k určení bodů nejvyšší pozornosti při sledování televizních pořadů nebo prohlížení internet, ale k vyvoji speciálních kompresních standardů. Praktické zjišťování oblasti ROI slouží k vytváření speciálních videosekvencí pro testování objektivní metody hodnocení obrazu znakového mluvčího.

4.3 Určení oblastí zájmu ROI neslyšících diváků

V práci [28] je zaveden úkol nalézt oblasti zájmu neslyšících pozorovatelů ve vizuálních pořadech, ve kterých je přítomen tlumočnický znakový jazyk. Jak už bylo zjištěno, neslyšící divák se soustředí svoji pozornost takovým způsobem, aby zachytil co nejvíce informací a porozuměl jí a to je jediný faktor, který je brán v úvahu. Největší pozornost neslyšící divák věnoval mluvčímu znakovému jazyku a konkrétně jeho obličej.

Pro uskutečnění experimentu byly použity předpřipravené videosekvence, reprezentující základní typy kompozice scén s tlumočnickem do znakového jazyka. Pomocí speciálního systému ViewPoint EyeTracker (měří pozici oka vůči hlavě) se pozorovaly pohyby očí diváků. Na výstupu byly získány data ve formě mapy bodů, charakterizujících směr a část obrázku, na kterou byly zaměřeny oči pozorovatelů.

Přehled výsledků nejlépe vytváří teplotní mapa pozornosti, “která přímo na dané videosekvenci vytvoří barevnou paletu překrývající podnět a zároveň definuje četnost výskytu pohledu na danou část scény (viz obr. 4.2)”[36].



Obr. 4.2: Teplotní mapa videosekvence po celou dobu klipu [28]

Oblast obrazu, které pozorovatel věnoval nejmíň pozornosti je vyznačena modrou barvou na obrázku. Naopak, tam, kam se pozorovatel díval nejdéle je nakreslen tmavě červený flek pro označení bodu nejvyšší pozornosti. Šedivá oblast vůbec ohodnocena pozornosti neslyšícího diváka [28] [36].

Ve výsledku teplotní mapy každé z použitých videosekvencí ukazují, že neslyšící divák se soustředí nejméně přesně na oblast kolem úst tlumočnicka. Přičemž nezáleží nacementové náladě promítané scény nebo velikosti tlumočnicka. Krátkodobě divák přemístí pozornost na další část obrazku, která ho nově zaujala, ale stejně se zase obrátí na mluvčího ZJ [28].

4.4 Praktická uplatnění oblastí zájmu neslyšících diváků

Objektivní hodnocení obrazu znakového mluvčího je počátečním bodem při plnění úkolu správného zvolení komprese, což zajistí dobrou vizuální přehlednost a srozumitelnost mluvčího znakového jazyka.

4.4.1 HbbTV a asistivní technologie

HbbTV (hybridní televize, Hybrid Broadcast Broadband TV) dokáže nakombinovat klasické televizní vysílání spolu s širokopásmovým internetem. Důvodem objevení takového typu vysílání je potřeba sjednocení různých výrobců přijímací techniky a poskytovatelů vysílání, aby zákazník se zbavil nutnosti řešit, co mu ve výsledku bude fungovat a jakou variantu musí odmítnout hned. Použivatel potřebuje televizní přijímač nebo tzv. set-top-box s podporou HbbTV technologie, důležité je připojení k internetu.

S rychlým rozšířením a zvýšením popularity dané technologie, zvětšila se i nabídka služeb a možností, nechybí tady ani asistivní technologie. Cílem je

rozšířit funkcí této technologie takovým způsobem, aby neslyšící divák mohl libovolně zobrazovat znakového mluvčího spolu s promítaným pořádem, což by bylo možné uskutečnit při oddělovaném přenosu videa se znakovým mluvčím a datového toku pořádu. Navíc by se dalo přizpůsobit obraz mluvčího ZJ konkrétnímu divákovi podle velikosti a polohy na obrazovce. Možnost volitelného zobrazení mluvčího ZJ již není dostupná a video se znakovým mluvčím se přenáší současně s televizním pořádem.

Kapitola 5

Testované videosekvence

Z předchozí kapitoly byla odnesena důležitá myšlenka, že neslyšící divák při pozorování obrazu nejvyšší procento své pozornosti věnuje té části obrazu, kde se nachází mluvěcí znakového jazyka a bodem největší pozornosti je hlava znakového mluvěcího, kterou pozorovatel vnímá foveálním viděním. Dalším důležitým bodem jsou ruce vnímané periferně.

Na základě daných vlastnosti byly připraveny videosekvence, na kterých je přítomen jenom mluvěcí znakového jazyka před modrým pozadím [30]. Videá pro testování dané metody jsou ve 2 rozlišeních: 544x544 a 640x640 pixelů. Daná rozlišení byla zvolená neslyšícími diváky při subjektivních testech. V základě kodování videosekvencí leží princip oblasti zájmu pozorovatele (ROI).

5.1 Objevování ROI oblasti

Obličej, pleťová barva (ruce) a pozadí vytváří tři nejdůležitější oblasti zájmu diváka a k jejich detekci a odlišování je zvolen Viola-Jonesův detektor. Daný detektor je schopen najít oblast obrazu s barvou pleti nebo oblast, která vůbec neodpovídá této barvě. Navíc detektor v programu má implementaci, při které dokáže najít barvu pleti obličeje a ostatních části těla [30][36].

Ve výsledku algoritmus dává masku, což je tzv. mapa čísel, každé číslo které odpovídá makrobloku o velikosti 16x16 pixelů. Právě podle hodnoty makrobloku se rozhoduje, do jaké skupiny ROI oblasti makroblok spadá.

5.2 Parametr h_{par}

Vyhledávání oblasti ROI se provádí na etapě kódování testované videosekvence. Jak již bylo zmíněno, oblasti ROI jsou rozděleny na 3 části. Důležitým parametrem v dalším kroku je kvantizační parametr QP . Má za úkol se stárat o správnou kompresi každé ROI oblasti ve snímku a k tomu slouží následující rovnice [30][36]:

$$QP_{obličej} = QP \cdot \frac{1}{h_{par}}, \quad (5.1)$$

$$QP_{ruce} = QP \cdot \frac{\frac{1}{hpar} + 1}{2}, \quad (5.2)$$

$$QP_{pozadí} = \frac{N_b \cdot QP - (a \cdot QP_{obličej} + b \cdot QP_{ruce})}{c}. \quad (5.3)$$

Vzácným parametrem pro tyto rovnice je $hpar$, který definuje rozdíl v kvalitě každé oblasti ROI. N_b odpovídá celkovému počtu makrobloků masky a a , b , c jsou hodnoty počtů makrobloků v každé ROI části. [36]

Důležitou poznámkou z rovnic je, že hodnota $hpar$ je přímo úměrná kvalitě obličejů a částečně i kvalitě rukou a nepřímo úměrná kvalitě pozadí [36].

5.3 Zakódování videosekvencí

Referenční (původní) videosekvence se znakovým mluvčím jsou kódovány do formátu UYVY se vzorkovacími hodnotami 4:2:2 a navíc jsou zabaleny v AVI kontejneru. Formát UYVY je skoro stejný jako formát YUY2. Rozdíl spočívá v uspořádání vzorků. U použitého formátu UYVY jasové a chrominanční vzorky “se periodicky opakují v sekvenci U Y V Y po jednotlivých vzorcích v každé obrazové složce”. Vzorky Y, V a U jsou vyjádřeny osmi bity. V dekadické podobě vzorky nabývají hodnot od 0 do 255 [36].

V případě testovaných videosekvencí byl použit otevřený zdrojový kód x264 ve formátu YUV420 se vzorkováním 4:2:0. Zakódovaná testovaná videa jsou následovně zabalená v MP4 kontejneru. “Jasové a chrominanční vzorky jsou v tomto případě přenášeny ve formě, kdy na 8 vzorků jasového signálu připadají 2 vzorky U a 2 vzorky V chrominančních signalů” [36]. Stejně jako u referenčních videí vzorky Y, V a U jsou osmibitové. Přesně z důvodu budoucího uplatnění daného systému hodnocení ve standardu HbbTV byl použit kód x264 k dosažení kompatibility [36].

Původní videa se znakovým mluvčím obsahují jenom jednu větu ze skupiny minimálních párů ve znakovém jazyce, což zároveň umožňuje nejen ohodnotit kvalitu obrazu znakového mluvčího, ale i srozumitelnost řeči [36].

Výsledky subjektivních testů jsou představené v poslední kapitole této práce.

Kapitola 6

Realizace objektivní metody hodnocení obrazu

Ted´se dostáváme k nejdůležitější části bakalářské práce a to je praktický návrh a realizace pomocí programu v prostředí MATLAB metody objektivního hodnocení obrazu mluvího znakového jazyku.

6.1 Předpoklady pro návrh objektivní metody

Vdůsledku neexistence standardní metriky hodnocení obrazu je vhodné nejdřív zvolit a rozhodnout, jaké vlastnosti jsou nejzásadnější pro dojem z obrazu mluvího ZJ. Na základě znalostí o funkcích lidského zrakového systému HVS se dá určit řadu vchozích bodů při stanovení teoretické bazi pro realizaci modelu hodnocení obrazu. Člověk nevnímá pozrovaný obraz lineárně, vjem je ovlivněn dalšími okolnostmi. V případě neslyšícího pozorovatele je změna v bodě pozornosti – část obrazu, která nejvíc upoutává pozornost divaka se přemístí na mluvího ZJ. Referenci při objektivním hodnocení obrazu slouží hodnoty výsledků subjektivních testů. Samozřejmě důležitou roli hraje i kodování obrazu mluvího znakového jazyka.

6.2 Popis navržené metody

K praktické realizaci navržené metody byl použit program naprogramovaný v prostředí MATLAB (verze 2014a), což je hodně populární interaktivní programové prostředí s vlastním programovacím jazykem, které se dá použít pro celou řadu funkcí jako např. technické vypočty, simulace, modelování, analýza dat a signálů, jejich zpracování atd.

6.2.1 Načtení videosekvencí

Za prvé se načítají referenční a testované videosekvence, kvalitu kterých je potřeba prozkoumat. (viz kap. 5) Pro ten účel je zavedena funkce `videoreader`, která každou videosekvenci převede do 4D matice, jejíž první dva rozměry odpovídají výšce a šířce snímku, třetí dimenze pak odpovídá barevné informaci, čtvrtá dimenze - počet snímků videosekvenci. Výškou snímku ve skutečnosti je počet aktivních řádků v daném snímku a šířkou snímku vyznačuje počet aktivních vzorků v jednom řádku daného snímku. Barevná informace je uložena u obou videosekvencí ve formátu RGB24 v buňkách `uint8` (8-bitový integer),

kteře odpovídat velikosti kvantování jednotlivých barevných vzorků ve videosekvencích [37].

6.2.2 Načtení masek

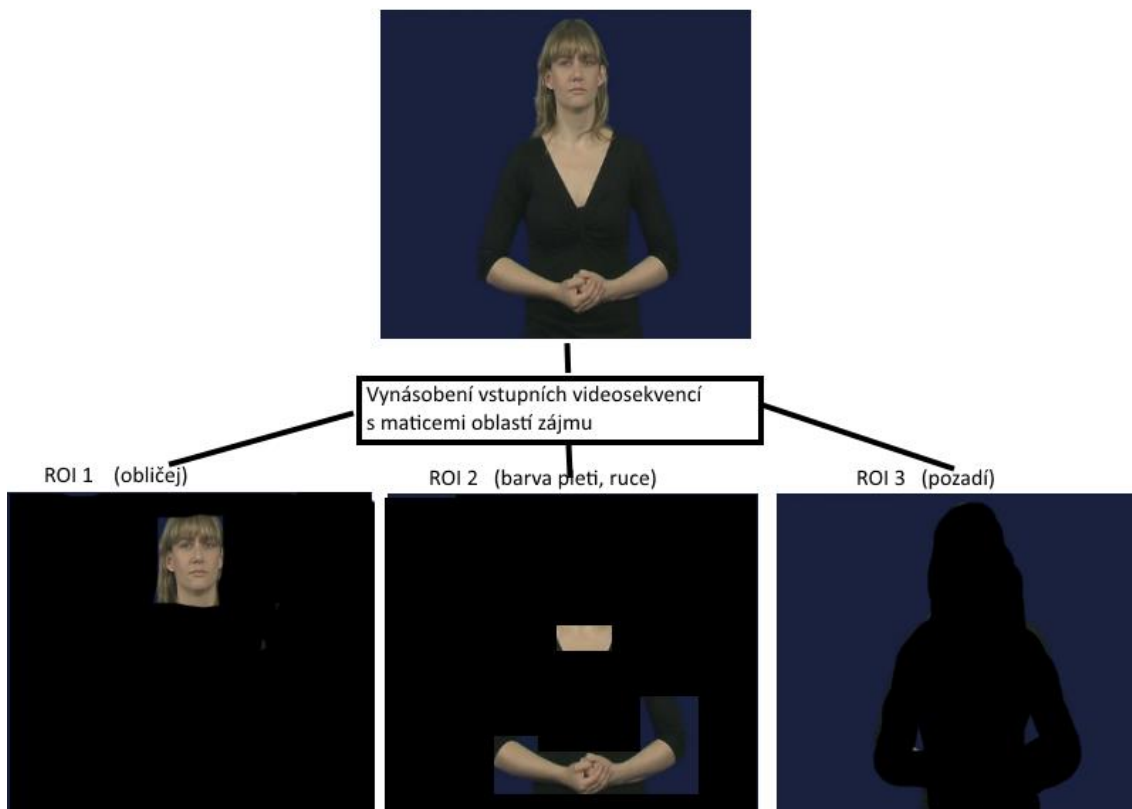
Jak již bylo zmíněno, navržená metoda hodnocení kvality obrazu se obrací k takovým vlastnostem lidského zrakového systému HVS, jako jsou foveální a periferní vidění. Daná metoda navíc dělí snímky testovaných videosekvencí na tři části z toho důvodu, že neslyšící diváci se soustředí na určitém bodě obrazu.

Tyto tři části (ROI) jsou obličej, ruce (barva pleti) a pozadí. ROI prokazují ve výsledku důležitost takového rozdělení při hodnocení kvality obrazu.

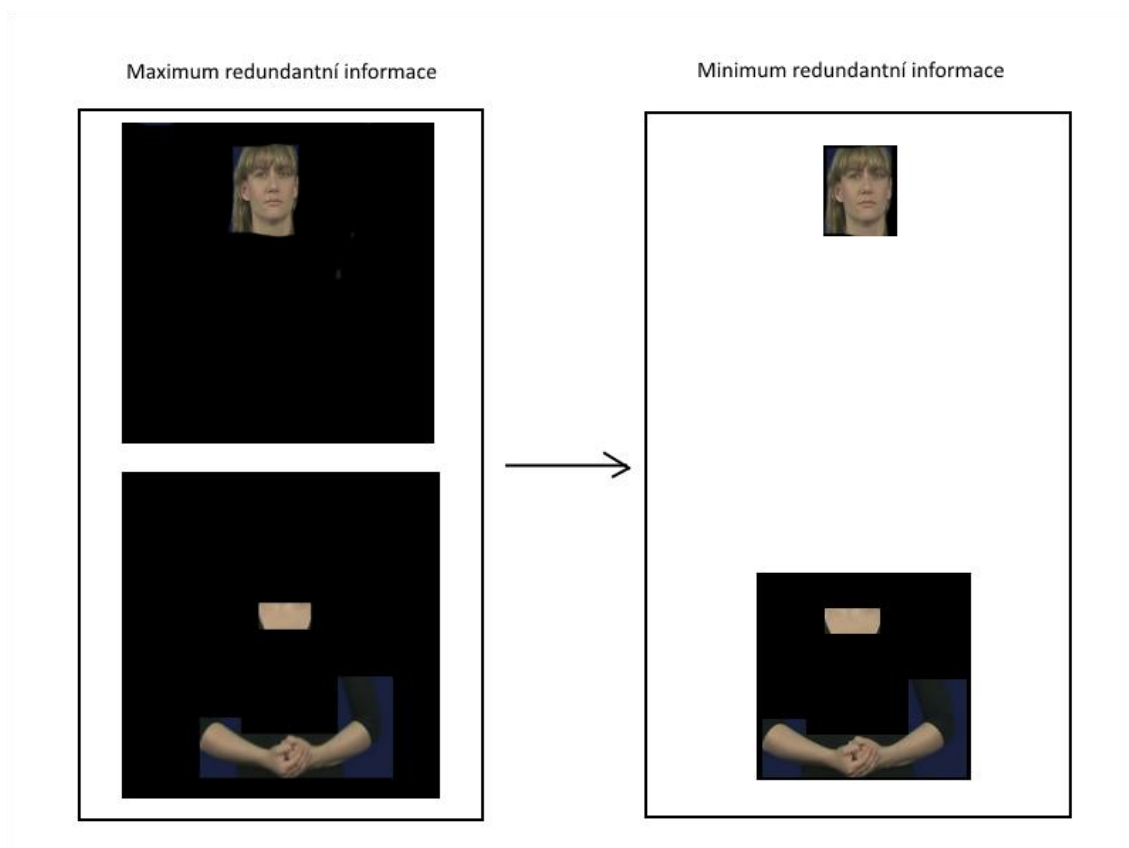
K detekci ROI v každém snímku byla použita data z výstupních textových souborů, která načítá skript `kompletmaska.m`. Dané textové soubory obsahují řadu čísel (1156 pro rozlišení 544x544 a 1600 pro rozlišení 640x640), každé číslo řady popisuje oblast zájmu pro makroblok 16x16 pixelů. Po roznásobení počtu čísel s velikostí makrobloků dostáváme hodnotu reprezentující všechny pixely v jednom snímku (295936 pixelů pro rozlišení 544x544 a 408600 pixelů pro rozlišení 640x640). Textové soubory se načítají postupně pro správné přiřazení k příslušným snímekům podle pořadového čísla. Dále probíhají matematické operace, v důsledku kterých data každého textového souboru vytvářejí 4D matice o stejných rozměrech všech dimenzí jako mají načtené videosekvence [37].

Tvořba ROI se uzavírá vzájemným vynásobením matic z textových dat během čtyř `for` cyklů (každý pro jednu dimenzi matice) spolu s maticemi referenčních a testovaných videosekvencí. Výstupem z každé videosekvence vznikne soustava tří 4D matic, která charakterizuje každou oblast zájmu.

Takovým postupem snímek se rozdělí na tři další snímky, které mají stejné parametry velikostí jako původní snímek. Při sečtení nové snímky ROI nesou stejné množství informace jako původní snímek. Dále se používá cyklus `for`, během kterého se vyhledávají maximální a minimální body v prvních dvou dimenzích u snímku ROI1 a ROI2 k vynechání redundantní informace, v důsledku čehož se zmenší rozlišení na tu nejmenší hodnotu množiny pixelů, při které se zachová informace snímku. Pro ROI3 proces zbavení nadbytečné informace neprobíhá, protože v daném případě nejde použít cyklus `for`. Grafické znázornění procesu rozdělení snímku na oblasti zájmů a odstranění redundantní informace představují obr. 6.1 a obr. 6.2.



Obr. 6.1: Schéma vytvoření oblastí zájmu (ROI)



Obr. 6.2: Odstranění redundantní informace z vytvořených oblastí zájmu

Veškeré nadbytečné informace se podařilo zbavit jenom v první oblasti zájmu ROI1, což je obličej znakového mluvčího. V dalších dvou ROI2 a ROI3 redundantní informace zůstává, ale i ten problem se dá vyřešit právě aditivností algoritmu, která bude probírána v jedné z dalších podkapitol.

Tím pádem je v tuto chvíli 6 čtyřdimenzionálních matic, každá z nich je pro jednu oblast zájmu diváka – 3 pro referenční a 3 pro testovanou videosekvenci, přičemž matice obsahují minimum nadbytečné informace, kterého se dalo dosáhnout v procesu popsaném výš. Taková vychozí informace je postačující pro zahájení samotného algoritmu hodnocení u dané metody.

6.2.3 Hodnotící algoritmus

Každá z již zmíněných metod hodnocení obrazu a jeho kvality má své specifické vlastnosti a je zaměřená na obraz s určitými parametry. Tak i v případě obrazu pro neslyšícího diváka musí být použita speciální metrika, která bude brát v úvahu požadavky člověka používajícího znakový jazyk. V jeho případě periferní vidění nehraje tak důležitou roli jako vidění foveální, protože pozorovatel se potřebuje soustředit jenom na určitou oblast zájmu. Tím pádem byla zvolena metoda SSIM, která nejen vyhovuje HVS a souhlasí s hodnotami experiment [31], ale také ve výsledku představuje jednu určitou hodnotu výsledné kvality obrazu.

Metoda strukturální podobnosti prošla později mnoha modifikacemi v důsledku zájmu jiných autorů. V této práci byla použita původní implementace metriky SSIM [31] a implementace MATLAB od verze 2014a [32]. Původní implementace má jako základ pro výpočty [31] funkci `ssimO.m`, která na vstupu má dva snímky, které je potřeba porovnat a na výstupu má hodnotu MSSIM metody a celou mapu indexů SSIM metody.

Na začátku se nastavuje velikost procházejícího okna w , které slouží jako jádro pro zjištění parametrů testovaného snímku na hodnotu 11×11 pixelů pomocí předdefinované funkce MATLABu:

```
w = fspecial('gaussian', 11, 1.5). [36]
```

Rovnice (3.11), (3.12) a (3.13) jsou použity k výpočtu parametrů $\mu_x, \mu_y, \sigma_x, \sigma_y$ a σ_{xy} snímku pomocí konvoluční funkce, která je realizována funkcí `filter2`. Například pro výpočet hodnoty parametru μ_x se používá následující příkaz:

```
ux = filter2(w, img_x, 'valid'),
```

kde `img_x` je snímek, který je potřeba otestovat. Stejným způsobem se zjistí hodnoty ostatních parametrů. Rovnice (3.10) a (3.14) slouží k výpočtu mapy indexů a průměru tech indexů pro sestavení metriky MSSIM.

Implementace v MATLABu [32] má analogický postup jako originální SSIM a základem je funkce `ssimM.m`. Na rozdíl od originální metody SSIM indexu tady se konvoluční jádro aplikuje současně na tři snímky, které obsahují pokaždé jednu barevnou složku – červenou, zelenou a modrou (R , G , a B). Tm pádem u originální implementaci SSIM se musí nejdřív každý snímek převést na šedotónovou stupnici. Používá se vzorek [19] v následujícím tvaru:

$$Y = 0.2126 R + 0.7152 G + 0.0722 B \quad (6.1)$$

Implementace v MATLABu konvoluci tří barevných složek vypočítává pomocí funkcí `imfilter`. Hodnoty v program se zjišťují pomocí příkazu:

```
Ux = imfilter(img_x, w, 'conv', 'replicate');
```

Výpočet hodnoty MSSIM a mapy indexů SSIM je stejný jako u originální implementaci SSIM indexu. Je potřeba zajistit výpočet přes všechny testované snímky použitím cyklu `parfor`, použití kterého vypadá takhle:

```
parfor i=1:frame
[ssimval(i)] = ssim(ref{i}(:, :, :), test{i}(:, :, :));
end
```

Při podrobném rozebírání daného cyklu zjistíme, že `frame` je počet snímků ve videosekvenci, `ssimval` je výsledná hodnota MSSIM, `ref` je původní snímek a `test` je testovaný snímek. Rozdíl mezi cyklem `for` a cyklem `parfor` spočívá v použití paralelního výpočtu algoritmu. Daná schopnost značně urychlí výpočet výsledných hodnot. Díky tomuto postupu jsou na výstupu tři matice s hodnotami všech MSSIM indexů jednotlivých snímků ve videosekvenci (každá matice pro jednu ROI). Cyklus `parfor` se používá třikrát z toho důvodu, že jsou šest snímků (3 ROI pro každou videosekvenci), které chceme porovnat, výstupy se poznamenávají do jednodimenzionální matice. Řadky tyto matice představují hodnoty MSSIM pro i -tý snímek ve videosekvenci.

6.2.4 Aditivnost metody SSIM

Nové snímky, popisující oblasti zájmu neslyšícího diváka obsahují jednu ROI z původního snímku, ale nesou v sobě také i redundantní informaci (černé pozadí). Změní se i velikost snímků (viz obr. 6.2). K garanci aditivnosti metody hodnocení obrazu nejde sečíst samotné hodnoty metriky MSSIM určitých ROI. To je z toho důvodu, že výsledek by neodpovídal metrice hodnocení referenčního snímku, který ještě nebyl rozdělen podle oblastí zájmu. Zavádí se parametr – váha, který slouží pro kontrolu aditivnosti metody. Váhování v program se docílí při použití cyklu `for`:

```
for i=1:framenorm
MSSIMA(i) = (1-ssimval1(i))*percent1(i);
MSSIMB(i) = (1-ssimval2(i))*percent2(i);
```

```
MSSIMC(i) = (1-ssimval3(i));
end
```

kde $ssimval1(i)$, $ssimval2(i)$ a $ssimval3(i)$ jsou hodnoty MSSIM pro ROI z původního i -tého snímku videosekvence. $percent1(i)$ a $percent2(i)$ jsou procentuálním zastoupením velikosti ROI v porovnání s původním i -tým snímkem (ROI 3 má vždy stejnou velikost jako původní i -tý snímek). $framenorm$ odpovídá počtu snímků ve videosekvenci. $MSSIMA(i)$, $MSSIMB(i)$ a $MSSIMC(i)$ jsou normované hodnoty MSSIM metriky pro každý i -tý snímek určité části oblasti zájmu.

Hodnoty MSSIM se nachází v rozsahu $0 < MSSIM(x, y) \leq 1$. Čím větší je hodnota MSSIM, tím menší je rozdíl u porovnávaných obrazů, což znamená, že obraz má lepší kvalitu. Osa hodnot MSSIM metriky se musí invertovat, aby objektivní a subjektivní stupnice se nesmíchaly, což vysvětluje to, že ve výpočtu hodnoty $ssimval$ se odečítají od jedné.

K potvrzení aditivnosti metody SSIM stačí, aby hodnota následujícího součtu hodnot $MSSIMA(i)$, $MSSIMB(i)$ a $MSSIMC(i)$ odpovídala hodnotě MSSIM původního i -tého snímku, který nebyl rozdělen na snímky ROI:

$$MSSIMZ(i) = 1 - ((MSSIMA(i) + MSSIMB(i) + MSSIMC(i))). \quad (6.2)$$

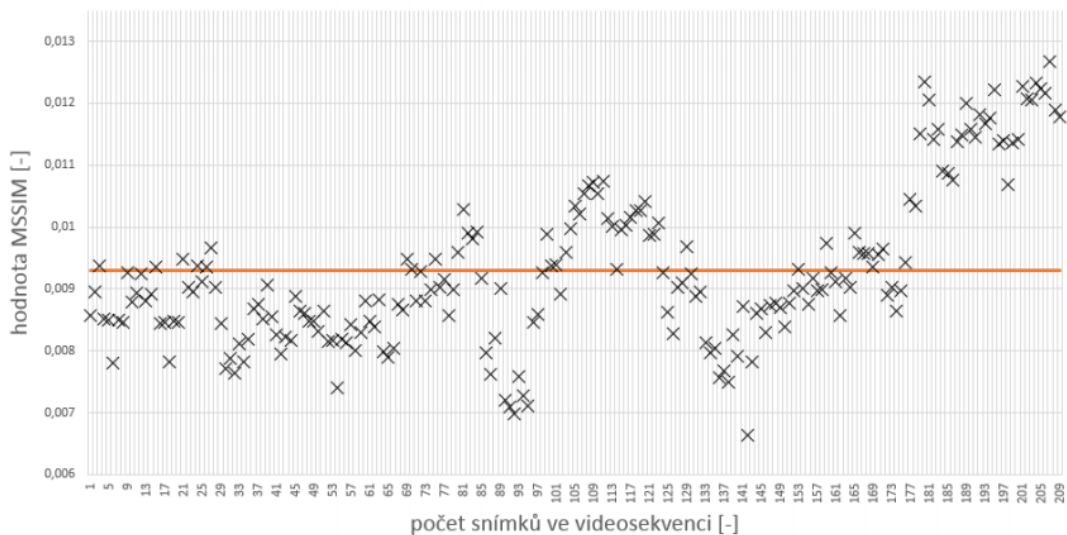
Jak již bylo zmíněno, aditivnost metody zajistí zbavení se veškeré nadbytečné informací, kterou obsahovaly snímky ROI. Hlavním důvodem potřeby odstranit redundantní informaci z ROI 1 a ROI 2 je urychlení výpočtu metody SSIM v prostředí MATLAB. V případě neúplného odstranění nadbytečné informace ze snímku by stačilo jenom sečíst hodnoty MSSIM bez procentuálního zastoupení a tím by se získala hodnota $MSSIMZ(i)$. Všechny části by zabíraly celkovou velikost původního obrazu.

V daném bodě výstup obsahuje tři normalizované matice všech MSSIM indexů jednotlivých snímků ve videosekvenci (každá matice pro jednu ROI). Poslední částí programu se zabývá výpočtem konečné hodnoty pro objektivní hodnocení kvality obrazu.[36]

6.2.5 Výsledek objektivního hodnocení kvality obrazu

Pro přehlednost, jednoduchost a soulad s výsledkami subjektivních testů je nutné přizpůsobit výstup objektivní metody takovým způsobem, aby algoritmus dával ve výsledku jenom jedno určité číslo, reprezentující hodnotu kvality obrazu.

První úkol při výpočtu hodnoty objektivní kvality obrazu je určení samotné normované hodnoty MSSIM pro každou část oblastí zájmu v celé videosekvenci. K tomu slouží jednoduchý součet průměru všech normovaných MSSIM hodnot ve všech snímcích v určité části ROI. Tím se získávají jenom tři normované hodnoty MSSIM, které budou použité dále. Výsledky daného postupu jsou graficky zobrazeny na obr. 6.3.



Obr. 6.3: Průměrné normované hodnoty MSSIM ze všech snímků jedné části oblasti zájmu [36]

Na obrázku body naznačují normované hodnoty MSSIM každého i -tého snímku určité části oblasti zájmu ve videosekvenci a barevná čárka ukazuje průměrnou hodnotu MSSIM v dané části oblasti zájmu.

Druhý úkol je váhování výsledných hodnot metriky MSSIM podle důležitosti oblastí zájmu, účelem kterého je výpočet výsledných hodnot FQR takovým způsobem, aby e výsledku hodnota hodnocení kvality co nejmíc odpovídala subjektivním testům. Váhování se realizuje v programu následujícím příkazem:

$$FQR = k * ((A * MSSIM1) ^2 + B * MSSIM2 + C * MSSIM3) ;$$

$MSSIM1$, $MSSIM2$ a $MSSIM3$ odpovídají celkovým normovaným hodnotám metriky MSSIM pro každou část ROI, FQR je výsledná hodnota objektivní kvality obrazu, k je konstanta a A , B , C jsou parametry váhy pro každou část oblasti zájmu.

Váhovací příkaz je rozdělen na dvě části. V první části každá ROI je vynásobena parametrem, který označuje důležitost dané oblasti zájmu v celkovém hodnocení kvality obrazu. Váhovací parametry A , B a C se určují vyřešením soustavy lineárních rovnic:

$$S = (A * MSSIM1) ^2 + B * MSSIM2 + C * MSSIM3, \quad (6.3)$$

Hodnota S v této soustavě odpovídá výsledku subjektivního hodnocení daného obrazu. $MSSIM1$, $MSSIM2$, $MSSIM3$ jsou normované hodnoty metriky MSSIM pro příslušnou ROI a A , B , C jsou vyžadované váhovací parametry. Počet rovnic soustavy odpovídá součtu testovaných videosekvencí pro porovnání s referenční videosekvencí. Dále součet se vynásobí počtem původních videosekvencí, které je třeba vyhodnotit. Pro lepší pochopení se uvádí příklad:

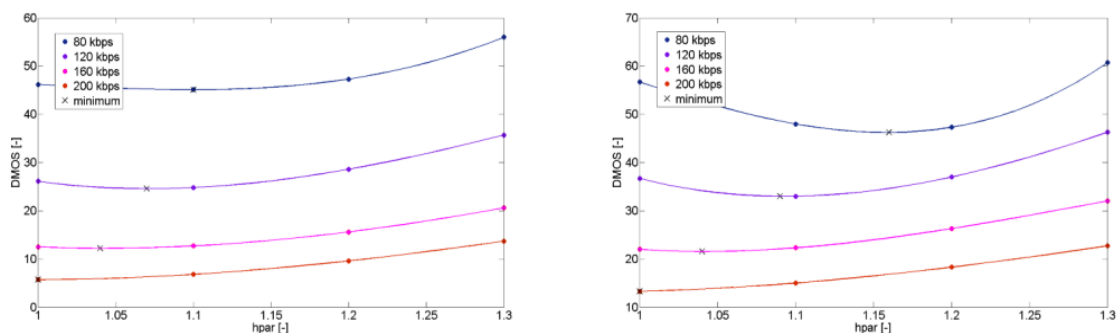
testované videosekvence mají odlišnou bitovou rychlost (80, 120, 160 a 200 kbps), hodnota parametru $hpar$ se mění (1.0, 1.1, 1.2, 1.3) a je potřeba otestovat 4 referenční videosekvence. Výsledná soustava lineárních rovnic v tomto případě obsahuje 64 členů. Daná soustava lineárních rovnic se vyřeší metodou nejmenších čtverců. Popis této metody je v [33]. V programu daný výpočet se provádí pomocí skriptu `ctverce.m`. Na vstupu jsou vektory b , které odpovídají hodnotám, získaným během subjektivních testů a také matice A reprezentující normované hodnoty metriky MSSIM. Počet řádků (a také počet prvků ve vektoru b) je celkovým počtem vyřešených lineárních rovnic. Počet sloupců dané matice je ekvivalentní počtu hledaných parametrů. Počet vyřešených rovnic se rovná 108 pro náš případ. K výpočtu parametrů A , B , C je zaveden jednoduchý příkaz [36] :

```
par=A\b;
```

Proměnná `par` nese v sobě prvky, které určují hodnoty parametrů A , B a C .

Nakonec zbyvá vydělit všechny parametry konstantou k , která odpovídá nejmenší kladné hodnotě nalezených parametrů a ve výsledku dostáváme váhové hodnoty parametrů odpovídající jejich důležitosti (parametr s nejmenší vahou má tak hodnotu 1). Pro dodržení platnosti rovnice 6.3 je nutné nalezené parametry vynásobit konstantou k . [36]

Váhování pokračuje druhou částí, která má na práci povahu výsledků subjektivních testů. Obr. 6.4 ukazuje zpracování výsledků subjektivních testů pro testovaná rozlišení 544x544 pixelů a 640x640 pixelů [30][36].



Obr. 6.4: Závislost hodnot subjektivních testů DMOS na parametru $hpar$ pro videosekvence o velikosti 544x544 pixelů (vlevo) a 640x640 pixelů (vpravo) [30][36]

Chování křivek grafu ukazuje, že hodnota parametru $hpar$ roste spolu s klesající bitovou rychlostí. Tyto výsledky pomáhají pochopit, že při snižující se hodnotě kbps roste důležitost kvality obličeje ve srovnání s kvalitou pozadí. Hodnoty daných grafů jsou v navržené metodě hodnocení kvality představeny jako kvadrát první normované hodnoty MSSIM (ROI 1 - obličej). Díky tomuto způsobu se na nižších bitových rychlostech zvýrazní přítomnost chyb objektivního hodnocení kvality obrazu v hlavní oblasti zájmu pozorovatele, což lépe odpovídá popsaným

výsledkům subjektivních testů.[36]

Ve výsledku algoritmus má na výstupu jedno určité číslo, odpovídající výsledné objektivní kvalitě (FQR) testovaného obrazu (videosekvence). Normované hodnoty metriky MSSIM pro každou oblast zájmu diváka, nalezené parametry a jejich porovnání s výsledkami subjektivních testů se rozebírá v další kapitole.

Kapitola 7

Vyhodnocení objektivní kvality videosekvencí

Navrženou metodu hodnocení obrazu je potřeba vyzkoušet a zkontrolovat její funkčnost. K tomu se používá soubor videosekvencí, jejichž hodnocení (FQR) se pak porovnává s hodnocením subjektivních testů (DMOS). Subjektivní testy se hodnotily pomocí metody DSCQS [30].

Pro nastavení parametrů potřebných k hodnocení obrazu se používají celkem 8 referenčních videosekvencí (4 v rozlišení 544x544 pixelů a 4 v rozlišení 640x640 pixelů) [36]. “U tří referenčních videosekvencí byly testované videosekvence k dispozici ve všech možnostech nastavení parametrů (bitová rychlost: 80, 120, 160 a 200 kbps, *hpar*: 1.0, 1.1, 1.2 a 1.3), ve zbylých chybělo nastavení *hpar* 1.0”[36].

Výsledné hodnoty hodnocení metrikou MSSIM testovaných videosekvencí se nacházejí v tabulkách, stejně jako výsledné hodnoty subjektivních testů. Výsledné porovnání subjektivních testů a zjištěných objektivních hodnot po váhování je zobrazeno do grafu. V grafu jsou vždy na horizontální ose vyneseny hodnoty FQR a na vertikální ose průměrné hodnoty DMOS pro danou videosekvenci. Obě osy jsou cejchovány v procentech, kde 100% znamená nejhorší kvalitu videosekvence a 0% znamená nejlepší kvalitu videosekvence. Diagonální modrou přímkou jsou vyznačeny ideální hodnoty, kterým odpovídá situace, kdy jsou hodnoty FQR naprosto totožné jako průměrné hodnoty subjektivních hodnocení DMOS.

7.1 Výsledky objektivního hodnocení kvality obrazu u originální SSIM metody

Následující tabulky obsahují hodnoty výstupů normovaných hodnocení MSSIM pro jednotlivé části ROI testovaných videosekvencí. *br* je bitová rychlost v kbps, *hpar* značí hodnotu parametru *hpar*, která byla nastavena u testované videosekvence a MSSIM1, MSSIM2 a MSSIM3 odpovídají hodnotám výstupů normovaných hodnocení MSSIM pro jednotlivé části ROI.

Tabulka výstupů subjektivních testů obsahuje hodnoty DMOS.

Tab. 7.1: Normované hodnoty MSSIM metriky pro videosekvenci marie-544-1a

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,003803	0,003240	0,002845	0,002485	0,002696	0,002277	0,002046	0,001829
MSSIM2	0,013061	0,012860	0,013151	0,013059	0,009515	0,009335	0,009415	0,009528
MSSIM3	0,010076	0,011318	0,012768	0,013684	0,008344	0,009131	0,010240	0,011096
br [kbps]	160				200			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,002109	0,001851	0,001631	0,001520	0,001788	0,001574	0,001427	0,001312
MSSIM2	0,007691	0,007598	0,007516	0,007783	0,006631	0,006501	0,006545	0,006612
MSSIM3	0,007453	0,008029	0,008798	0,009829	0,006841	0,007297	0,007850	0,008640

Tab. 7.2: Normované hodnoty MSSIM metriky pro videosekvenci marie-544-2a

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,003746	0,003245	0,002870	0,002497	0,002696	0,002281	0,002013	0,001782
MSSIM2	0,012065	0,012106	0,012443	0,012600	0,009020	0,009335	0,008927	0,009011
MSSIM3	0,010356	0,011807	0,013279	0,014894	0,008681	0,009541	0,010675	0,011632
br [kbps]	160				200			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,002071	0,001777	0,001584	0,001441	0,001709	0,001500	0,001361	0,001224
MSSIM2	0,007146	0,007026	0,007158	0,007230	0,006113	0,006053	0,006127	0,006078
MSSIM3	0,007585	0,008141	0,009008	0,009875	0,006899	0,007462	0,008084	0,008754

Tab. 7.3: Normované hodnoty MSSIM pro videosekvenci marie-544-1b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,003146389	0,002553953	0,002488789	0,002242175	0,001987378	0,001751624
MSSIM2	0,012500029	0,011602798	0,013092846	0,009058694	0,009061385	0,009069769
MSSIM3	0,011207506	0,012122513	0,014156376	0,009094390	0,009985960	0,011120916
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,001751384	0,001540056	0,001401484	0,001486512	0,001316144	0,001209906
MSSIM2	0,007153633	0,007154530	0,007200395	0,006224878	0,006160136	0,006215061
MSSIM3	0,007882271	0,008521802	0,009298135	0,007178602	0,007692777	0,008363256

Tab. 7.4: Normované hodnoty MSSIM pro videosekvenci marie-544-2b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,003280362	0,002942937	0,002458800	0,002165363	0,001898244	0,001663497
MSSIM2	0,013566453	0,014138904	0,013768110	0,009237568	0,009391766	0,009273095
MSSIM3	0,012453058	0,014827595	0,015969432	0,009557763	0,010752321	0,011749751
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,001690626	0,001490682	0,001356640	0,001398449	0,001290248	0,001191439
MSSIM2	0,007434215	0,007372198	0,007471571	0,006263836	0,006341452	0,006476839
MSSIM3	0,008210424	0,008890439	0,009985657	0,007462482	0,008124790	0,008845819

Tab. 7.5: Normované hodnoty MSSIM pro videosekvenci marie-640-1a

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,004783	0,004020	0,003476	0,003065	0,003385	0,010052	0,002494	0,002240
MSSIM2	0,013910	0,013698	0,013854	0,014030	0,010290	0,009335	0,010087	0,010194
MSSIM3	0,010356	0,013320	0,015308	0,016725	0,009839	0,011030	0,012115	0,013250
br [kbps]	160				200			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,002701	0,002304	0,002052	0,001858	0,002259	0,001950	0,001759	0,001616
MSSIM2	0,008368	0,008285	0,008276	0,008365	0,007213	0,007125	0,007115	0,007250
MSSIM3	0,008775	0,009565	0,010550	0,011782	0,008094	0,008733	0,009331	0,010195

Tab. 7.6: Normované hodnoty MSSIM pro videosekvenci marie-640-2a

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,004337688	0,003656024	0,003144130	0,002919990	0,002526999	0,002237381
MSSIM2	0,013761618	0,013385482	0,013468037	0,009349435	0,009394961	0,009342406
MSSIM3	0,014250340	0,016114627	0,017974325	0,011082641	0,012253334	0,013421515
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,002253883	0,002001968	0,001837829	0,001927540	0,001707883	0,002018912
MSSIM2	0,007447901	0,007569130	0,007750617	0,006508907	0,006438280	0,009216723
MSSIM3	0,009619391	0,010520409	0,011871436	0,008722788	0,009434935	0,013834642

Tab. 7.7: Normované hodnoty MSSIM pro videosekvenci marie-640-1b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,003799350	0,003327171	0,002891411	0,002709728	0,002356776	0,002090150
MSSIM2	0,013488873	0,013551395	0,013485574	0,009790310	0,009759796	0,009653946
MSSIM3	0,013144566	0,014748248	0,016662680	0,010673579	0,011780081	0,012929478
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,002139982	0,001922912	0,001723710	0,001820514	0,001651588	0,001506751
MSSIM2	0,007859273	0,007925913	0,007964172	0,006823337	0,006861608	0,006908389
MSSIM3	0,009240755	0,009939840	0,011164326	0,008361505	0,009077269	0,009828883

Tab. 7.8: Normované hodnoty MSSIM pro videosekvenci marie-640-2b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,003878420	0,003108275	0,002894601	0,002629966	0,002224677	0,002119927
MSSIM2	0,013200220	0,012190544	0,012838046	0,009111782	0,008815280	0,009373172
MSSIM3	0,014137325	0,015484203	0,017695644	0,010805721	0,012118095	0,013964398
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,002190849	0,001834989	0,001682435	0,001832975	0,001637886	0,001501864
MSSIM2	0,007669871	0,007287103	0,007364762	0,006489787	0,006443709	0,006443710
MSSIM3	0,009606779	0,010529069	0,011740931	0,008668078	0,009292212	0,012724092

Tab. 7.9: Hodnoty DMOS pro videosekvence o rozlišení 544x544 pixelů

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
DMOS	46,1	44,8	47,3	55,8	26,2	24,5	28,6	35,9
br [kbps]	160				200			
DMOS	12,3	13,1	15,7	21,3	5,7	6,6	9,8	13,5

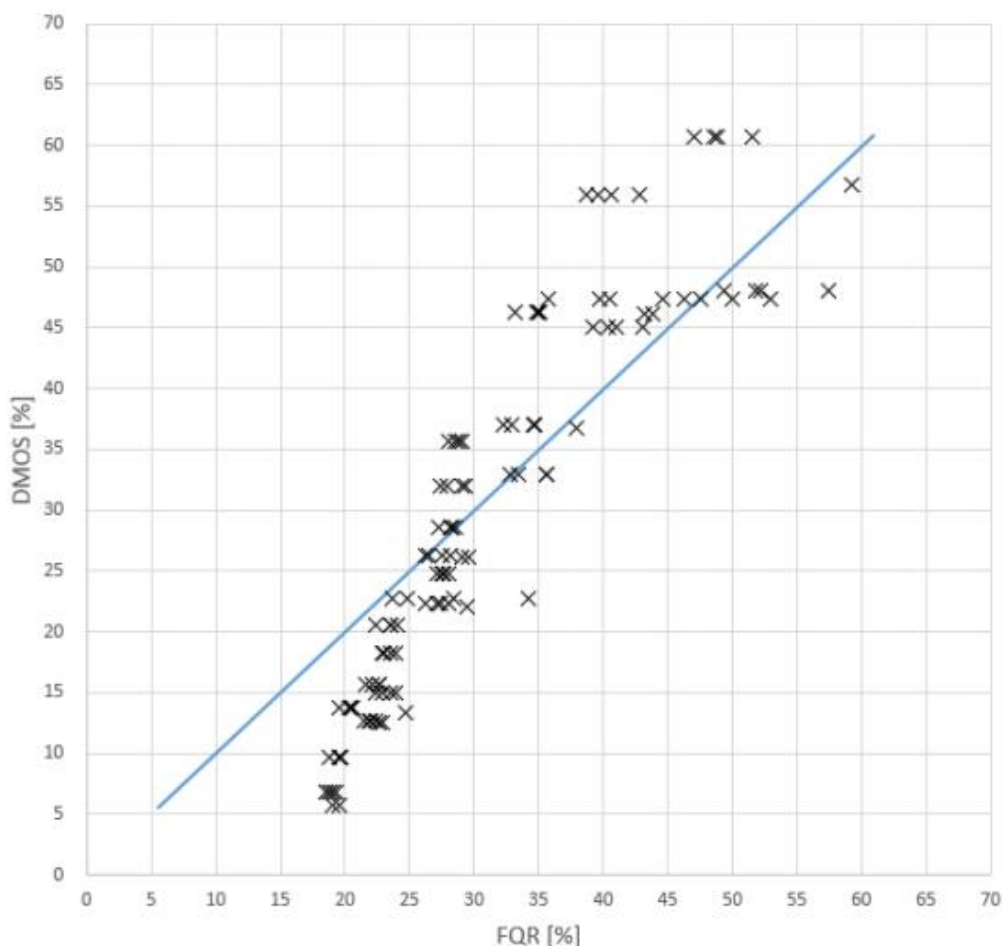
Tab. 7.10: Hodnoty DMOS pro videosekvence o rozlišení 640x640 pixelů

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
DMOS	57,3	49,1	47,0	62,0	36,7	33,0	37,2	46,3
br [kbps]	160				200			
DMOS	23,5	22,2	26,9	32,3	13,3	15,2	19,9	22,7

Parametry A , B , C a k se následovně vypočítají metodou nejmenších čtvců z hodnot metrik MSSIM a DMOS. Výsledné hodnoty těchto parametrů jsou zapsané do tabulky 7.11.

Tab. 7.11: Výsledky parametrů A , B , C a konstanty k

Parametry	A	B	C	k
Hodnoty	53,81065033	1	3,663807150	477,0119661



Obr. 7.1: Výsledky DMOS subjektivních testů a FQR objektivního hodnocení kvality obrazu originální SSIM metodou

Graf z obr. 7.1 vizuálně ukazuje vztah výsledků objektivních a subjektivních testů při hodnocení kvality obrazu metodou SSIM vůči ideálnímu stavu (výsledky subjektivních testů odpovídají výsledcům objektivních testů). Ideální

stav je zobrazen jako modrá čára. Kvůli nepřesnostem subjektivních testů jsou mezi hodnotami těch testů a hodnotami objektivního hodnocení malé odchylky. Jak je vidět z grafu v oblasti 20-25% objektivní metoda jakoby podhodnocuje výsledky a to je z toho důvodu, že při takové bitové rychlosti obrazu metoda stále nachází chyby i když neslyšící pozorovatel už si těch chyb nevšimne.

Tabulky 7.11., obsahující vypočtené parametry oblastí zájmu potvrzují, že nejvyšší pozornost diváka získává podle předpokladů část ROI 1, což je oblast obličeje. Je 50krát důležitější než oblast rukou, což potvrzuje i další předpoklady z předchozích kapitol (viz kap. 4). Podle algoritmu parametr pro druhou oblast zájmu (ruce, barva pleti) je 3,6krát méně důležitý, než parametr u oblasti zájmu 3 (pozadí). Během subjektivních testů se naopak zjistilo, že pozadí hraje důležitou roli a má váhu v hodnocení kvality, kterou nesmíme zanedbat. Parametr váhy pozadí je vyšší než váha oblasti rukou z toho důvodu, že na vyšších hodnotách parametru *hpar* klesá kvalita pozadí, jde to pozorovat i z tabulek 7.9 a 7.10.

Ale i přesto váha pozadí je mnohem menší než váha oblasti obličeje znakového mluvčího. Stále je parametr zdůrazňující pozadí několikrát menší než parametr váhování oblasti obličeje. Grafu na obr. 7.1 nam ukazuje, že výstupní hodnoty objektivního hodnocení kvality obrazu netvoří přímou lineární závislost s výsledky subjektivních testů. “Důvodem může být nedostatek zhodnocených videosekvencí s parametrem *hpar* 1.0, které nebyly k dispozici”[36]. Tento problém se snadno vyřeší použitím metody nejmenších čtverců, pak výstupní hodnoty parametrů *hpar* 1.0 jsou převáženy ve srovnání s výsledky hodnot parametrů *hpar* 1.1, 1.2 a 1.3.

7.2 Výsledky objektivního hodnocení kvality obrazu u SSIM MATLABu

I v případě metody SSIM verze MATLABu jsou představeny tabulky obsahující hodnoty výstupů normovaných hodnocení metriky MSSIM pro každou oblast ROI u testovaných videosekvencí. Také v těch tabulkách *br* je bitová rychlost [kbps], *hpar* odpovídá hodnotě parametru *hpar* a MSSIM1, MSSIM2 a MSSIM3 jsou výstupní hodnoty normovaných hodnocení metriky MSSIM pro jednotlivé části oblasti zájmu. “U výstupů subjektivních testů jsou v tabulce místo hodnot MSSIM hodnoty DMOS”[36].

Tab. 7.12: Normované hodnoty metriky MSSIM pro videosekvenci marie-544-1a

br [kbps]	80				120			
	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,001358	0,001131	0,001050	0,000899	0,000950	0,000699	0,000709	0,000633
MSSIM2	0,006755	0,006750	0,006973	0,006952	0,005002	0,004948	0,005060	0,005118
MSSIM3	0,011210	0,013126	0,016031	0,018273	0,008946	0,010066	0,011853	0,013033
br [kbps]	160				200			
	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1	0,000737	0,000640	0,000567	0,000524	0,000615	0,000538	0,000486	0,000448
MSSIM2	0,004088	0,004079	0,004051	0,004202	0,003579	0,003531	0,003551	0,003597

MSSIM3	0,007714	0,008473	0,009681	0,011375	0,006967	0,007550	0,008365	0,009376
--------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------	-----------------

Tab. 7.13: Normované hodnoty MSSIM pro videosekvenci marie-544-2a

br [kbps]	80				120				
	hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1		0,001272	0,001126	0,000994	0,000880	0,000914	0,000784	0,000696	0,000621
MSSIM2		0,006111	0,006239	0,006390	0,006526	0,004609	0,004554	0,004585	0,004685
MSSIM3		0,011441	0,013424	0,015647	0,019150	0,008995	0,010357	0,011585	0,013508
br [kbps]	160				200				
	hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1		0,000707	0,000610	0,000550	0,000495	0,000585	0,000509	0,000465	0,000414
MSSIM2		0,003723	0,003659	0,003744	0,003788	0,003206	0,003189	0,003212	0,003190
MSSIM3		0,007613	0,008461	0,009696	0,010927	0,006809	0,007512	0,008389	0,009423

Tab. 7.14: Normované hodnoty MSSIM pro videosekvenci marie-640-1a

br [kbps]	80				120				
	hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1		0,001666	0,001391	0,001202	0,001068	0,001138	0,000977	0,000853	0,000771
MSSIM2		0,006553	0,006520	0,006589	0,006861	0,004894	0,004810	0,004854	0,004967
MSSIM3		0,014083	0,016396	0,018491	0,023838	0,010363	0,012225	0,014011	0,016106
br [kbps]	160				200				
	hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
MSSIM1		0,000908	0,000779	0,000690	0,000624	0,000748	0,000652	0,000583	0,000536
MSSIM2		0,003954	0,003945	0,003969	0,004039	0,003429	0,003407	0,003408	0,003499
MSSIM3		0,008839	0,010006	0,011566	0,013046	0,007975	0,008857	0,009977	0,011321

Tab. 7.15: Normované hodnoty MSSIM pro videosekvenci marie-544-1b

br [kbps]	80			120			
	hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1		0,001096341	0,000884362	0,000864946	0,000772911	0,000684288	0,000600122
MSSIM2		0,006524153	0,006154987	0,006866008	0,004770903	0,004834109	0,004815633
MSSIM3		0,012836748	0,014003983	0,017046382	0,009892330	0,011240556	0,012533040
br [kbps]	160			200			
	hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1		0,000591037	0,000526631	0,000473119	0,000499460	0,000445276	0,000408075
MSSIM2		0,003811676	0,003838405	0,003863695	0,003317619	0,003293734	0,003340149
MSSIM3		0,008267641	0,009203213	0,010283293	0,007382576	0,008045704	0,008920438

Tab. 7.16: Normované hodnoty MSSIM pro videosekvenci marie-544-2b

br [kbps]	80			120			
	hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1		0,001116218	0,001035492	0,000862077	0,000744251	0,000648698	0,000572619
MSSIM2		0,006738290	0,007216470	0,007070576	0,004730601	0,004824350	0,004734176
MSSIM3		0,014608240	0,018799595	0,020912881	0,010377541	0,011947610	0,013609698
br [kbps]	160			200			
	hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1		0,000572554	0,000512100	0,000464308	0,000476033	0,000435438	0,000401853
MSSIM2		0,003849636	0,003805619	0,003880113	0,003267044	0,003331306	0,003390723
MSSIM3		0,008533220	0,009584463	0,010860705	0,007475308	0,008484573	0,009721371

Tab. 7.17: Normované hodnoty MSSIM pro videosekvenci marie-640-1b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,001324583	0,001147994	0,001028465	0,000945717	0,000811199	0,000715417
MSSIM2	0,006432523	0,006478426	0,006451778	0,004677886	0,004670592	0,004641143
MSSIM3	0,015322299	0,017680852	0,021540158	0,011584505	0,013267379	0,015159110
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,000726932	0,000649994	0,000586402	0,000612242	0,000558118	0,000506476
MSSIM2	0,003712993	0,003739569	0,003792159	0,003232553	0,003241426	0,003291193
MSSIM3	0,009460569	0,010760246	0,012259675	0,008322733	0,009331277	0,010577656

Tab. 7.18: Normované hodnoty MSSIM pro videosekvenci marie-640-2a

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,001535230	0,001300387	0,001120475	0,001016848	0,000870732	0,000775048
MSSIM2	0,006344205	0,006324023	0,006395645	0,004380130	0,004412691	0,004434520
MSSIM3	0,016576000	0,020386342	0,023199160	0,011851012	0,013739230	0,015580541
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,000772895	0,000696854	0,000632618	0,000664131	0,000584540	0,000501288
MSSIM2	0,003502455	0,003534544	0,003662122	0,003044025	0,003031004	0,003032630
MSSIM3	0,009900458	0,011395913	0,013199878	0,008619596	0,009771447	0,019624656

Tab. 7.19: Normované hodnoty MSSIM pro videosekvenci marie-640-2b

br [kbps]	80			120		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,001298140	0,001028320	0,000971561	0,000865522	0,000731408	0,000706230
MSSIM2	0,006156811	0,005804857	0,006212201	0,004258557	0,004131925	0,004488365
MSSIM3	0,016708479	0,019744628	0,023823436	0,011775674	0,013782340	0,016464518
br [kbps]	160			200		
hpar	1.1	1.2	1.3	1.1	1.2	1.3
MSSIM1	0,000716532	0,000595692	0,000554202	0,000592829	0,000525591	0,000482532
MSSIM2	0,003564091	0,003410166	0,003485625	0,003042901	0,003015974	0,003016020
MSSIM3	0,010077154	0,011109558	0,013157550	0,008681656	0,009692950	0,011216444

Tab. 7.20: Hodnoty DMOS pro videosekvence o rozlišení 544x544 pixelů

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
DMOS	46,2	45,1	47,3	56,0	26,2	24,5	28,6	35,9
br [kbps]	160				200			
DMOS	12,5	12,7	15,7	20,6	5,7	6,8	9,8	13,5

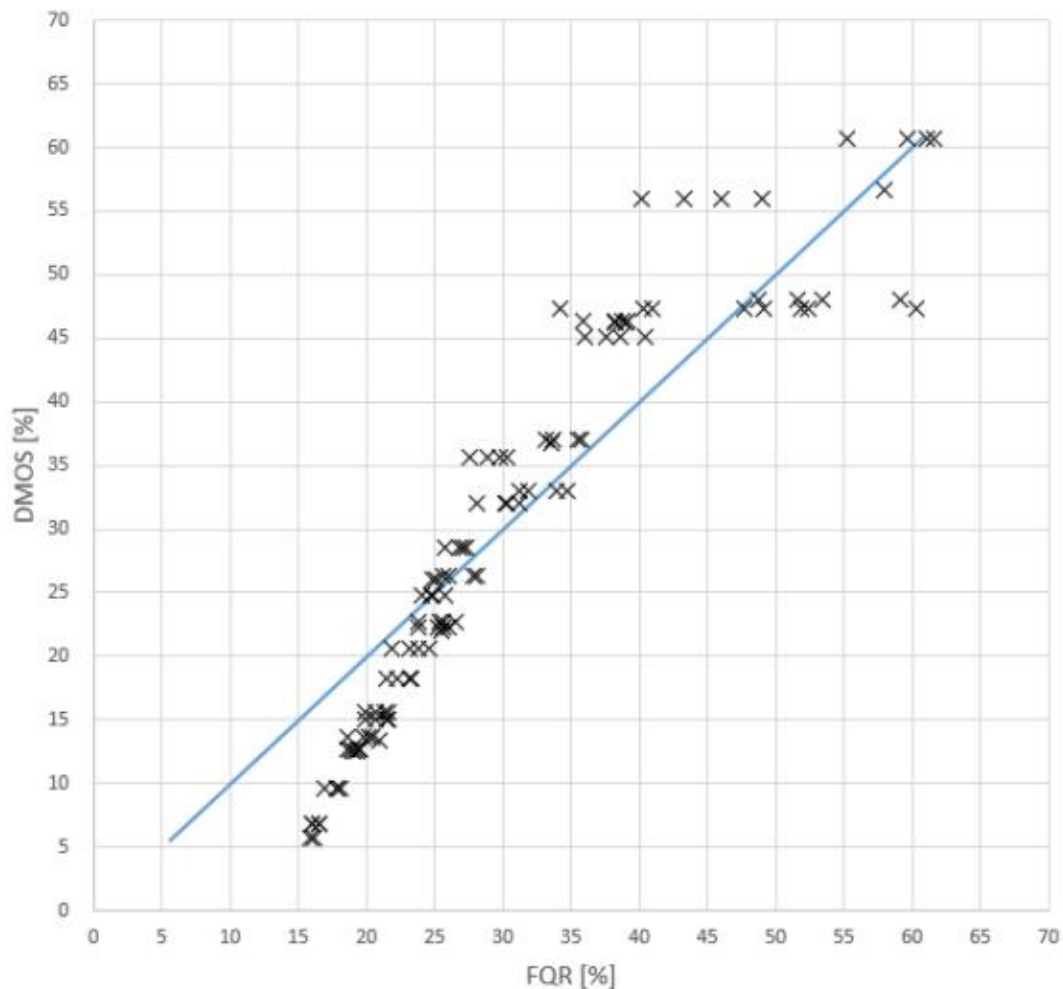
Tab. 7.21: Hodnoty DMOS pro videosekvence o rozlišení 640x640 pixelů

br [kbps]	80				120			
hpar	1.0	1.1	1.2	1.3	1.0	1.1	1.2	1.3
DMOS	56,7	48,0	47,3	60,7	36,7	33,0	37,0	46,3
br [kbps]	160				200			
DMOS	22,0	22,3	26,3	32,0	13,3	15,0	18,3	22,7

Dále se vyplní tabulka hodnot parametrů A , B , C a konstanty k , vypočítávají se pomocí tabulek hodnot MSSIM.

Tab. 7.22: Výsledky parametrů A , B , C a konstanty k

Parametry	A	B	C	k
Hodnoty	71,00174906	-0,594768075	1	2396,773847



Obr. 7.2: Porovnání výsledků subjektivních testů DMOS a FQR objektivního hodnocení kvality obrazu SSIM MATLABu[36]

Na obr. 7.2. se ukazuje graf sestavený z hodnot výsledků objektivních a subjektivních testů hodnocení kvality obrazu znakového mluvčího při použití metody SSIM od MATLABu. Daný graf souží k porovnání získaných výstupů ve stejném měřítku. Jak i v předchozím případě, kde se používala jiná verze metriky, modrá čára v grafu odpovídá ideálnímu stavu, kdy subjektivní testy jsou ekvivalentní testům objektivním. „Výsledky objektivních testů jsou i zde v drobných odchylkách vůči testům subjektivním, protože ne vždy lze subjektivní hodnocení kvality obrazu přesně postihnout“[36].

Výsledky v tomto případě mají stejný charakter chování jako na obr. 7.1. Znovu největší váhu ve výsledku má oblast zájmu ROI 1 (obličej). Druhá nejvyšší hodnota váhy odpovídá ROI 3 - pozadí. Jak je vidět z tabulky, hodnota parametru B (oblast zájmu ROI 2) je záporná, což se jedinečně dá vysvětlit snahou metody

nejmenších čtvrců najít nejlepší možné řešení problému. Také to znamená, že se používá operace odečtení v algoritmu a ne sčítání všech výsledků s různou váhou. Na takovém záporném výsledku se podělí i skutečnost, že u testů zase chybí videosekvence s parametrem $hpar$ 1.0

7.3 Funkčnost metody

Je potřeba nejen ověřit funkčnost hodnotící metody, ale i ohodnotit ji pomocí Pearsonova koeficientu korelace. Daný koeficient ukazuje vztah mezi výsledkami subjektivních testů a objektivním hodnocením kvality a může nabývat hodnot od 0 do 1. Hodnota 1 odpovídá přímé závislosti obou hodnocení a hodnota 0 naopak – nulová závislost mezi oběma hodnoceními kvality obrazu. Koeficient spočítáme ze vztahu [35]:

$$\rho(A, B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right) \quad (5.3)$$

kde A představuje výsledky subjektivních testů, B značí výsledky objektivních testů, μ_A a σ_A jsou hodnoty standardní odchylky A , μ_B a σ_B jsou hodnoty standardní odchylky B , N značí počet výsledků a ρ udává hodnotu výsledného Pearsonova koeficientu korelace. Pearsonův koeficient korelace je spočten v MATLABu ve skriptu `ctverce.m`. Následující tabulka uvádí výsledné hodnoty u obou způsobů objektivního měření využívající originální metodu SSIM a metodu SSIM MATLABu.

Tab. 7.23: Porovnání Pearsonova koeficientu korelace u objektivních měření

Metoda	Original SSIM	SSIM MATLABu
Hodnoty	0,913756057	0,935041832

Z tabulky 7.23 vyplývá, že Pearsonův koeficient korelace je v obou verzích metody vysoký (přesahuje hodnotu $\rho = 0,91$). Dokonce vyšel lépe pro metodu využívající metodu SSIM MATLABu (i přes zvláštní přístup výpočtu parametrů zdůrazňujících kvalitu jednotlivých částí ROI) než metodu využívající originální SSIM. V tuto chvíli se už musí rozhodnout samotný uživatel, zda chce využít metodu využívající metodu SSIM MATLABu (pomalejší výpočet, $\rho = 0,935$) nebo metodu využívající originální metodu SSIM (rychlejší výpočet, $\rho = 0,914$).

Závěr

Tato bakalářská práce pojednává o návrhu objektivní metody hodnocení kvality obrazu znakového mluvčího. Základem při návrhu metody sloužila skutečnost odlišné doby, během které neslyšící divák věnuje svoji pozornost určité části obrazu, což dovolilo rozdělit daný obraz na tři oblasti zájmu pozorovatele. Pro potvrzení těchto předpokladů byly použité dvě verzé již vytvořené metody hodnocení SSIM.

V první části práce se popisuje struktura a funkce lidského zrakového systému HVS, který bere v úvahu navržená metoda hodnocení obrazu. Pak jsou ukázány objektivní metody hodnocení kvality obrazu, nejvíce pozornosti obdržela metoda SSIM, nechybí i kompletní popis výpočtu příslušné MSSIM metriky.

V další části se získaly výsledky subjektivních testů neslyšících diváků a následovně byly použity pro výpočet výstupních hodnot objektivní metody hodnocení kvality obrazu.

Všechny výpočty jsou provedeny v programovacím prostředí MATLAB. Dané prostředí velice vyhovuje pro řešení naše problematiky z toho důvodu, že obsahuje veškeré nástroje k zpracování videí.

Nejdůležitější předpoklad a základní myšlenka celé práce souhlasí se samotným cílem návrhu objektivní metody hodnocení kvality obrazu, rozděleného na určité části s příslušnou váhou. Použitím metriky MSSIM je docíleno správná funkčnost metody a soulad s vlastnostmi lidského zrakového systému. Metrika MSSIM je představena ve dvou verzích - jako originální SSIM metoda (přímo od autora), a SSIM metoda v MATLABu, což je součást samotného programu MATLAB. Obraz, rozdělený na tři oblasti zájmu ROI neslyšícího diváka a následovně váhovaný po částech podle důležitosti určité oblasti je začleněn do algoritmu metody. Důležitým bodem je správný výpočet a komprimace oblasti ROI pro zajištění funkčnosti metody.

Pro spuštění a ověření metody hodnocení obrazu byly použity soubory videosekvencí s mluvčím znakového jazyka před modrým pozadím. K zakódování každé videosekvenci byl použit zdrojový kód x264, který umí různě nastavit hodnotu rychlosti toku informací a také parameter *hpar*, určující kvalitu jednotlivých ROI částí vůči sobě. Výsledky testování metody hodnocení navíc byly kalibrovány na základě výsledků získaných ze subjektivních testů pomocí metody nejmenších čtverců.

Byl použit Pearsonův koeficient korelace pro další ověření funkčnosti navržené objektivní hodnotící metody pro srovnání výsledků subjektivních testů a objektivní metody. Nakonec oba způsoby výpočtu metody měly daný koeficient nad hodnotou 0,91, u verzi SSIM MATLAB dokonce nad 0,935. Výsledné koeficienty prokazují vysokou hodnotu korelace se subjektivními testy, což

znamená i dobrou funkčnost navržené metody pro hodnocení obrazu mluvího ZJ.

Táto práce by mohla posloužit začátkem pro návrh dalších hodnotících metod a metrik a zároveň i zlepšení již navržené metody SSIM a její aplikace na animovaného znakového mluvího a úkol oddělit videosekvenci se znakovým mluvího od příslušného televizního programu.

Seznam použitých obrázků

- Obr. 1.1: Schématický řez lidským okem [2] 10
- Obr. 1.2: Weber-Fechnerův zákon pro dvoubarevný experiment a jeho grafické vyjádření [2] 12
- Obr. 1.3: Grafické znázornění CSF dle Dalyho [5] 13
- Obr. 3.1: Blokové schéma Lubinova modelu HVS [18] 20
- Obr. 3.2: Blokové schéma Dalyho VDP multikanálového modelu [5] 21
- Obr. 3.3: Blokové schéma DCT modelu ve frekvenční oblasti [23] 22
- Obr. 3.4: Blokové schéma Sarnoffova multikanálového modelu [24] 23
- Obr. 3.5: Blokové schéma SSIM modelu [26] 25
- Obr. 4.1: Základní znakovací prostor [28] 30
- Obr. 4.2: Teplotní mapa videosekvence pro všechny subjekty a celou dobu klipu [28] 31
- Obr. 6.1: Schéma postupu vytvoření oblastí zájmu v programu 37
- Obr. 6.2: Odstranění redundantní informace z vytvořených oblastí zájmu 38
- Obr. 6.3: Zjištění průměrné normované hodnoty MSSIM ze všech snímků jedné ROI části videosekvence 42
- Obr. 6.4: Závislost hodnot subjektivních testů DMOS na parametru *hpar* pro videosekvence o velikosti 544x544 pixelů a 640x640 pixelů [30] 43
- Obr. 6.5: Vzhled uživatelského prostředí 45
- Obr. 7.1: Srovnání výsledků DMOS a FQR hodnocení kvality obrazu originální SSIM 49
- Obr. 7.2: Srovnání výsledků DMOS a FQR hodnocení kvality obrazu SSIM MATLABu 53

Seznam použitých tabulek

- Tab. 7.1: Normované hodnoty MSSIM pro videosekvenci marie-544-1a 47
- Tab. 7.2: Normované hodnoty MSSIM pro videosekvenci marie-544-2a 47
- Tab. 7.3: Normované hodnoty MSSIM pro videosekvenci marie-640-1a 47
- Tab. 7.4: Normované hodnoty MSSIM pro videosekvenci marie-544-1b 47
- Tab. 7.5: Normované hodnoty MSSIM pro videosekvenci marie-544-2b 48
- Tab. 7.6: Normované hodnoty MSSIM pro videosekvenci marie-640-1b 48
- Tab. 7.7: Normované hodnoty MSSIM pro videosekvenci marie-640-2a 48
- Tab. 7.8: Normované hodnoty MSSIM pro videosekvenci marie-640-2b 48
- Tab. 7.9: Hodnoty DMOS pro videosekvence o rozlišení 544x544 pixelů 49
- Tab. 7.10: Hodnoty DMOS pro videosekvence o rozlišení 640x640 pixelů 49
- Tab. 7.11: Výsledky parametrů A , B , C a konstanty k 49
- Tab. 7.12: Normované hodnoty MSSIM pro videosekvenci marie-544-1a 51
- Tab. 7.13: Normované hodnoty MSSIM pro videosekvenci marie-544-2a 51
- Tab. 7.14: Normované hodnoty MSSIM pro videosekvenci marie-640-1a 51
- Tab. 7.15: Normované hodnoty MSSIM pro videosekvenci marie-544-1b 51
- Tab. 7.16: Normované hodnoty MSSIM pro videosekvenci marie-544-2b 52
- Tab. 7.17: Normované hodnoty MSSIM pro videosekvenci marie-640-1b 52
- Tab. 7.18: Normované hodnoty MSSIM pro videosekvenci marie-640-2a 52
- Tab. 7.19: Normované hodnoty MSSIM pro videosekvenci marie-640-2b 52
- Tab. 7.20: Hodnoty DMOS pro videosekvence o rozlišení 544x544 pixelů 53
- Tab. 7.21: Hodnoty DMOS pro videosekvence o rozlišení 640x640 pixelů 53
- Tab. 7.22: Výsledky parametrů A , B , C a konstanty k 53
- Tab. 7.23: Porovnání Pearsonova koeficientu korelace u objektivních měření 54

Seznam použitých symbolů a zkratek

1D, 4D – jednodimenzionální, čtyřdimenzionální
AVI - Audio Video Interleave
B - Blue (modrá)
BSL - Britský znakový jazyk
C_r - chrominanční složka barevného prostoru Y C_r C_b
C_b - chrominanční složka barevného prostoru Y C_r C_b
CIE L*u*v, CIE Luv - CIE barevný prostor stejných barevných rozdílů
CQE - Continuous Quality Evaluation
CSF - Contrast Sensitivity Function (Funkce kontrastní citlivosti)
ČSÚ - Český statistický úřad
ČZJ - Český znakový jazyk
DCT - Discrete Cosine Transform (diskrétní kosínová transformace)
DMOS – Differential Mean Opinion Score
DS - Double Stimulus
DSCQS - Double Stimulus Continuous Quality Scale
DSIS - Double Stimulus Impairment Scale
FQR – Final Quality Rating
FR - Full Reference
G - Green
H.264, x264 - High 264 (formát multimedialního kodeku)
Hbb TV – Hybrid broadcast broadband TV
HVS - Human Visual System (lidský vizuální systém)
JND - Just Noticeable Difference(s) (právě pozorovatelný rozdíl(y))
kbps – kilobits per second (kilobit za sekundu)
L - jasový kanál barevného prostoru CIE L*u*v
LBC - Local Band-limited Contrast (pásmově omezený kontrast)
MATLAB – MATrix LABoratory
MOS - Mean Opinion Score
MP4 – Motion Picture 4
MPEG - Motion Picture Experts Group
MSE - Mean Squared Error (střední kvadratická odchylka)
MSSIM - Mean Structural Similarity Index (celkový index strukturální podobnosti)
NR - No Reference
PSNR - Peak Signal to Noise Ratio (odstup signál šum)
R - Red
RGB - barevný prostor základních barev R, G, B
ROI - Region of Interest (oblast zájmu)
RR - Reduced Reference
SS - Single Stimulus
SSCQE - Single Stimulus Continuous Quality Evaluation
SSIM - Structural Similarity Index (index strukturální podobnosti)
SSM - Single Stimulus Method
SSMR - Single Stimulus with Multiple Repetitions
u - červeno-zelený kanál barevného prostoru CIE L*u*v
v - žluto-modrý kanál barevného prostoru CIE L*u*v
VDP - Visible Differences Predictor (predictor viditelných rozdílů)
XYZ - barevný prostor nereálných základních barev X, Y, Z
Y - Luminance (jas)
ZJ - znakový jazyk

Seznam použité literatury

[1] WANG, Zhou, BOVIK, Alan C.: *Modern image quality assessment*. [s.l.] : Morgan & Claypool publishers, 2006, s. 18-26

[2] NADENAU, Marcus.: *Integration of Human Vision Models into High Quality Image Compression*, These no. 2269, Ecole polytechnique federale de Lausanne, 2000, str. 9-29.

[3] WANG, Zhou, SHEIKH, Hamid R., BOVIK, Alan C.: *Objective video quality assessment. The Handbook of Video Databases: Design and Applications*, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, Texas, USA, 2003, kapitola 41, s. 1041-1078.

[4] HECHT, Selig.: *The visual discrimination of intensity and the Weber-Fecher law*, Department of Physical Chemistry in the Laboratories of Physiology, Harvard medical School, Boston, 1924.

[5] DALY, S.: *The visible difference predictor: An algorithm for the assessment of image fidelity*, Digital Images and Human Vision, ed. A.B. Watson, MIT press, Cambridge, MA, USA, 1993, s.179-206.

[6] ITU-R Recommendation BT.500-11.: *Methodology for the Subjective Assesment of the Quality of the Television Pictures*, 2002, s. 48.

[7] ITU-T Recommendation P.910.: *Subjective video quality assessment methods for multimedia applications*, 1999.

[8] DUŠEK, Jaroslav.: *Objektivní hodnocení subjektivní kvality obrazu na základě modelu*, Disertační práce, Ústav radioelektroniky, ČVUT FEL, Praha, 2008.

[9] WINKLER, Stefan.: *Digital video quality, Vision Models and Metrics*, London, Wiley, 2005, ISBN 0-47002404-6, s. 54.

[10] ZACH, Ondřej.: *Měření kvality obrazu při příjmu analogového a digitálního televizního vysílání*, Bakalářská práce, Ústav radioelektroniky, Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně, Brno, 2011, s. 57.

[11] CIARAMELLO, Francis M., HEMAMI, Sheila S.: *'Can you see me now?' An objective metric for predicting intelligibility of compressed american sign language video*, Visual Communication Laboratory, School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 2007.

- [12] WU, H.R., YUEN, M.: *A generalized block-edge impairment metric for video coding*. In: *IEEE Signal processing Letters*, 4(11):317-320, Nov. 1977.
- [13] WINKLER, Stefan.: *Perceptual Video Quality Metrics - A review*. In: *Digital video image quality and perceptual coding*, Boca Raton: Taylor, 2006, 640 s ISBN 0 82472777-0.
- [14] SLANINA, Martin.: *Methods and tools for image and video quality assesment*, Disertační práce, Ústav radioelektroniky, Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně, Brno, 2008.
- [15] ZACH, Ondřej.: *Nástroje pro měření kvality videosekvencí bez reference*, Diplomová práce, Ústav radioelektroniky, Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně, Brno, 2013.
- [16] LUBIN, Jeffrey.: *The use of psychological data and models in the analysis of display system performance*, In *Digital images and Human Vision*, ed. By A.B. Watson, MIT press, Cambridge, MA, USA, 1993, s. 163-178.
- [17] LUBIN, Jeffrey.: *A visual discrimination model for imaging system design and evaluation*, In *Vision models for target detection and recognition*, ed. By E.Peli, World Scientific, River Edge, NJ, USA, 1995, s. 245-283.
- [18] JACKSON, W. B., BEEBEE, P., JARED, D. A., BIEGELSEN, D.K., TRIMER, J. O., LUBIN, Jeffrey, GILLE, J.L.: *X-ray Image Systém Design Using a Human Visual Model*, Proc. SPIE, 2708 Medical Imaging, vol. 2706, 1995, s. 29-40.
- [19] ITU-R Recommendation BT.709.: *Parameter values for the HDTV standards for production and international programme exchange*, 2015.
- [20] DALY, S.: *Method and apparatus for determining visually perceptible differences between images*, USA, 1995.
- [21] DALY, S.: *A visual model for optimising the design of image processing algorithm*, In *ICIP 94*, Austin, Texas, USA, 1994.
- [22] WATSON, A. B.: *Perceptual optimization of DCT color quantization matrices*, in Proc. IEEE Int. Conf. Image Processing, Austin, TX, 1994.
- [23] ŠEVČÍK, Martin.: *Modelování vlastností modelu HVS v MATLABu*, Diplomová _____práce, Ústav radioelektroniky, Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně, Brno, 2008, s. 20.

- [24] TEKTRONIX.: *A Guide to Picture Quality Measurement for Modern Television Systems*, Tektronix, 4/97 FL5372 25W-11419-0, 1997, s. 20.
- [25] LUBIN, Jeffrey.: *Sarnoff JND Vision Model*, Sarnoff Corporation,, Princeton, NJ, USA, 1998.
- [26] WANG, Zhou, BOVIK, Alan C., SHEIKH, Hamid R., SIMONCELLI, E. P.: *The SSIM Index for Image Quality Assessment*, New York Univ. & Univ. Of Texas, Austin, 2003.
- [27] WANG, Zhou, BOVIK, Alan C., SHEIKH, Hamid R., SIMONCELLI, E. P.: *Image Quality Assessment: From Error Measurement to Structural Similarity*, IEEE Trans. On Image Proc., Vol. 13, No. 4, 2004, s. 600-612.
- [28] ZATLOUKAL, Petr.: *Oblasti zájmu ve videosekvencích znakového jazyka*, Diplomová práce, Ústav radioelektroniky, ČVUT FEL, Praha, 2011.
- [29] UYVY Video Picture Encoding [online], dostupné z: <http://www.digitalpreservation.gov/formats/fdd/fdd000365.shtml>
- [30] ZATLOUKAL, Petr, BERNAS, Martin.: *Optimized H.264 compression of sign language video*, (při zpracování práce nebylo zatím publikováno)
- [31] The SSIM Index for Image Quality Assessment [online], dostupné z: <http://www.cns.nyu.edu/lcv/ssim/>
- [32] SSIM MATLAB [online], dostupné z: <http://www.mathworks.com/help/images/ref/ssim.html>
- [33] MONSPORTOVÁ, Jana.: *Aproximace funkcí – metoda nejmenších čtverců*, Bakalářská práce, Ústav matematiky a statistiky, Přírodovědecká fakulta, Masarykova univerzita, Brno, 2013.
- [34] Weber-Fechnerův zákon [online], dostupné z: http://www.wikiskripta.eu/index.php/Weber-Fechner%C5%AFv_z%C3%A1kon
- [35] Pearsonův korelační koeficient [online], dostupné z: <http://www.mathworks.com/help/matlab/ref/corrcoef.html>
- [36] ROUSEK, Marek: *Objektivní hodnocení kvality obrazu mluvího znakového jazyka*, Bakalářská práce, Katedra radioelektroniky, Fakulta elektrotechnická, ČVUT, Praha, 2016

[37] KLEJMOVÁ, Eva.: *Měření kvality pro HEVS*, Diplomová práce, Ústav radioelektroniky, Fakulta elektrotechniky a komunikačních technologií, Vysoké učení technické v Brně, Brno, 2014
https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=83222

[38] WOLFOVÁ, Lucie.: *Optické vlastnosti lidského oka*, Semestrální projekt z předmětu Obrazové inženýrství, 2002
http://www.fch.vut.cz/~zmeskal/obring/presentace_2003/02_opticke_vlastnosti_oka.pdf

Seznam obsahu přiloženého DVD

Program	skripty, funkce a příslušné soubory použité při výpočtu výsledků navržené metody
Testovaná videa	soubor všech použitých referenčních a testovaných videí
Text	text práce ve formátu .pdf