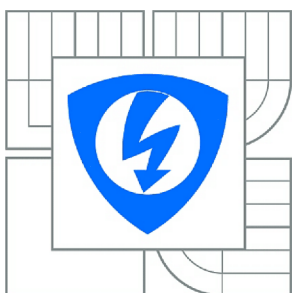




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



**FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ**

ÚSTAV RADIOELEKTRONIKY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF RADIO ELECTRONICS

IMPLEMENTACE METRIKY PRO HODNOCENÍ KVALITY VIDEOSEKVENCÍ

IMPLEMENTATION OF A METRIC FOR VIDEO QUALITY EVALUATION

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

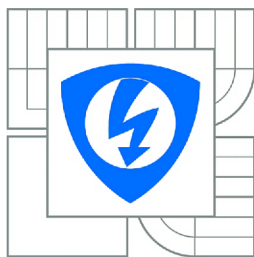
Bc. MILOŠ KACHLÍK

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. MARTIN SLANINA, Ph.D.

BRNO 2012



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav radioelektroniky

Diplomová práce

magisterský navazující studijní obor
Elektronika a sdělovací technika

Student: Bc. Miloš Kachlík

ID: 106519

Ročník: 2

Akademický rok: 2011/2012

NÁZEV TÉMATU:

Implementace metriky pro hodnocení kvality videosekvencí

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s problematikou hodnocení kvality videosekvencí s plnou referencí. Prostudujte kvalitativní metriku CPqD IES, popsanou v doporučení ITU-R BT.1683.

Metriku IES implementujte v libovolném programovacím jazyce. Ověřte výsledky Vaší implementace na základě výsledků testování VQEG [2].

Váš program doplňte vhodným uživatelským rozhraním. Pokuste se rozšířit jeho použitelnost i pro videosekvence s vysokým rozlišením a ověřte platnost výsledků s využitím subjektivních testů.

DOPORUČENÁ LITERATURA:

[1] Rec. ITU-R BT.1683. Objective Perceptual Video Quality Measurement Techniques for Standard Definition Digital Broadcast Television in the Presence of a Full Reference. Geneva: ITU, 2004.

[2] CORRIVEAU, P., WEBSTER, A. Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II [online]. VQEG, 2003 [cit. 16. prosince 2009]. Dostupné na [www](http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final_Report.pdf):
http://www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final_Report.pdf

Termín zadání: 6.2.2012

Termín odevzdání: 18.5.2012

Vedoucí práce: Ing. Martin Slanina, Ph.D.

Konzultanti diplomové práce:

prof. Dr. Ing. Zbyněk Raida

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Cílem této práce je vytvořit program schopný implementovat metriku CPqD-IES pro hodnocení kvality videosekvencí v prostředí MATLAB. Tato metrika je popsána v doporučení ITU-R BT.1683 pro metody objektivního měření kvality videosekvencí ve standardním rozlišení vysílaném digitální televizí v reálném čase s plnou referencí.

Hodnocení kvality videa je vypočítáno na základě objektivních parametrů v závislosti na segmentaci obrazu. Měřené videosekvence jsou segmentací rozčleněny do oblastí hran, ploch a textur. Objektivní parametry jsou přiřazeny ke každé z těchto oblastí. Vztah mezi každým objektivním parametrem a subjektivním hodnocením kvality je aproximován křivkou, která je výsledným odhadem úrovně kvality pro každý parametr.

KLÍČOVÁ SLOVA

CPqD-IES, video, objektivní měření, hodnocení kvality

ABSTRACT

The aim of this work is to create a program able to implement the metric CPqD-IES to evaluate the quality of video in system MATLAB. This metric is described in Recommendation ITU-R BT.1683 for objective video quality measurement techniques for standard definition digital broadcast television with presence of a full reference.

Video quality assessment is calculated objective parameters based on image segmentation. Natural scenes are segmented into edge, plane and texture regions. Objective parameters are assigned to each of these contexts. The relationship between each objective parameter and a subjective impairment level is approximated by the logistic curve, which is resulting in an estimated impairment level for each parameter.

KEYWORDS

CPqD-IES, video, objective measurement, quality evaluation

Kachlík, M. *Implementace metriky pro hodnocení kvality videosekvencí*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav radioelektroniky, 2012. 43 s. Vedoucí diplomové práce: Ing. Martin Slanina, Ph.D..

PROHLÁŠENÍ

Prohlašuji, že svoji diplomovou práci na téma Implementace metriky pro hodnocení kvality videosekvencí jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

V Brně dne 18. května 2012

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji vedoucímu diplomové práce Ing. Martinovi Slaninovi, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne 18. května 2012

.....

(podpis autora)

OBSAH

Seznam obrázků

Seznam tabulek

| | |
|---|-----------|
| Úvod..... | 1 |
| 1 Metody objektivního hodnocení kvality videosekvencí..... | 2 |
| 1.1 Doporučení ITU-R BT.1683..... | 2 |
| 2 Metrika CPqD-IES..... | 4 |
| 2.1 Popis systému IES..... | 4 |
| 2.2 Korekce posunutí a zesílení..... | 6 |
| 2.2.1 Časové posunutí snímků..... | 6 |
| 2.2.2 Prostorové posunutí snímku..... | 7 |
| 2.2.3 Zesílení..... | 8 |
| 2.3 Segmentace obrazu..... | 9 |
| 2.3.1 Plošné oblasti..... | 9 |
| 2.3.2 Oblasti hran..... | 10 |
| 2.3.3 Oblasti Textur..... | 13 |
| 2.4 Objektivní měření..... | 14 |
| 2.5 Databáze znehodnocených videosekvencí..... | 15 |
| 2.6 Odhad úrovně kvality..... | 16 |
| 2.6.1 Výpočet parametru W_i | 17 |
| 2.6.2 Výpočet parametrů F_i a G_i | 18 |
| 3 Realizace metriky CPqD-IES..... | 21 |
| 3.1 Videosekvence ve standardním rozlišení..... | 21 |
| 3.1.1 Načtení videosekvence ve standardním rozlišení..... | 21 |
| 3.2 Videosekvence ve vysokém rozlišení..... | 23 |
| 3.2.1 Načtení videosekvence ve vysokém rozlišení..... | 23 |
| 3.3 Kodekové operace..... | 23 |
| 3.3.1 Kodeky pro videosekvence ve standardním rozlišení..... | 24 |
| 3.3.2 Kodeky pro videosekvence ve vysokém rozlišení..... | 25 |

| | |
|---|-----------|
| 3.4 Vytvoření databáze znehodnocených videosekvencí..... | 25 |
| 3.4.1 Databáze SD..... | 25 |
| 3.4.2 Databáze HD..... | 29 |
| 3.5 Vyhodnocení odhadů kvality VQR..... | 31 |
| 3.5.1 Odhad VQR pro videosekvence ve standardním rozlišení..... | 31 |
| 3.5.2 Odhad VQR pro videosekvence ve vysokém rozlišení..... | 34 |
| 3.6 Uživatelský panel..... | 35 |
| 3.6.1 Popis použitých skriptů..... | 36 |
| Závěr..... | 40 |
| Literatura..... | 42 |
| Seznam symbolů, veličin a zkratk..... | 43 |

SEZNAM OBRÁZKŮ

| | |
|---|----|
| Obr. 2.1: Obecné schéma algoritmu CPqD-IES (převzato z [1])..... | 5 |
| Obr. 2.2: Blokové schéma segmentace (převzato z [1])..... | 9 |
| Obr. 2.3: Obraz jasové složky Y videosekvence O | 10 |
| Obr. 2.4: Plošné oblasti jasové složky Y videosekvence O | 10 |
| Obr. 2.5: Hranové oblasti před prahováním a dilatací..... | 12 |
| Obr. 2.6: Hranové oblasti po prahování a dilataci..... | 13 |
| Obr. 2.7: Texturní oblasti..... | 13 |
| Obr. 2.8: Obraz absolutního Sobelova rozdílu D_b | 14 |
| Obr. 3.1: Výsledky metriky pro SD videosekvence..... | 32 |
| Obr. 3.2: Vliv parametru m na odhad kvality VQR..... | 32 |
| Obr. 3.3: Odhad kvality VQR pro videosekvenci č.1..... | 33 |
| Obr. 3.4: Výsledky metriky pro HD videosekvence..... | 34 |
| Obr. 3.5: Transformované výsledky metriky pro HD videosekvence..... | 35 |
| Obr. 3.6: Vzhled uživatelského prostředí..... | 36 |
| Obr. 3.7: Vývojový diagram..... | 37 |

SEZNAM TABULEK

| | |
|---|-----------|
| Tab. 1.1: Přehled objektivních modelů hodnocení kvality videosekvencí [1]..... | 3 |
| Tab. 2.1: Přehled časových a prostorových atributů dle ITU-R BT.1683 [1]..... | 16 |
| Tab. 3.1: Použité videosekvence ve formátu 625/50i [6]..... | 26 |
| Tab. 3.2: Použité videosekvence ve formátu 1080/50i..... | 30 |
| Tab. 3.3: Porovnání DMOS a parametrů m pro videosekvenci č.1..... | 33 |

ÚVOD

Tradiční hodnocení kvality obrazu je prováděné na reprezentujícím počtu hodnotitelů, kteří subjektivně hodnotí kvalitu videa. Hodnocení může být prováděno buď s nebo bez referenční videosekvence. V prvním případě je hodnotitelům předložena referenční videosekvence a znehodnocená videosekvence. Na základě porovnání předložených videosekvencí mohou hodnotitelé testované videosekvenci udělit subjektivní bodové ohodnocení. Proto je toto hodnocení nazýváno jako subjektivní test kvality videosekvencí. Ačkoli je subjektivní test považován za nejpřesnější metodu, neboť odráží lidské vnímání, má i několik omezení. Za prvé vyžaduje určitý počet hodnotitelů, proto je časově náročné a může být poměrně drahé. Kromě výše zmíněného nemůže být prováděno v reálném čase. Z těchto důvodů je velký zájem na rozvíjení objektivních metod pro měření kvality videosekvencí. Subjektivní metody hodnocení kvality videa jsou popsány v doporučení ITU-R BT.500, případně ITU-T P.910.

Cílem této práce je vytvořit program schopný implementovat metriku CPqD-IES pro hodnocení kvality videosekvencí v prostředí Matlab. Tato metrika je popsána v doporučení ITU-R BT.1683 - pro metody objektivního měření kvality videosekvencí ve standardním rozlišení vysílaném digitální televizí v reálném čase s plnou referencí.

Práce je uspořádána následovně. První kapitola je věnována metodám objektivního hodnocení kvality videosekvencí popsaných v doporučení ITU-R BT.1683. Kapitola 2 je zaměřena na popis metriky CPqD-IES. Je zde ukázáno obecné schéma algoritmu a popsána segmentace obrazu. V kapitole 3 je popsána realizace metriky v prostředí Matlab a dosažené výsledky.

1 METODY OBJEKTIVNÍHO HODNOCENÍ KVALITY VIDEOSEKVENCÍ

Digitální video systémy postupně nahrazují všechny stávající analogové video systémy a umožňují tak vytvoření spousty nových telekomunikačních služeb (např. satelitní vysílání, digitální televize, televize s vysokým rozlišením, video telekonference, internetová komerce a video), které se postupně stávají nezbytnou součástí průmyslu a všech domácností. Objektivní metriky pro měření kvality videosekvencí jsou požadovány průmyslem pro specifikaci požadavků na výkonnost systémů, porovnání konkurenční nabídky služeb, údržbu sítě, optimalizaci využívání omezených síťových zdrojů (přenosové rychlosti). Mezi jejich velké výhody patří jednoduchý algoritmický výpočet a schopnost integrace do automatizovaných testů.

Zvýšení kvality služeb je zároveň s uplatněním nových technologií měření velkým přínosem, jak pro koncové uživatele, tak i pro poskytovatele telekomunikačních služeb a zařízení. Na přesná měření kvality videosekvencí má vliv nejen vlastní přenosová soustava a její přenosové charakteristiky (např. bitová rychlost, chybovost), ale i samotné vlastnosti přenášené videosekvence (např. prostorový detail, pohyb).

Technologiemi potřebnými k posouzení výkonu nových digitálních video systémů se zabývají specializované organizace, například *International Telecommunication Union* (ITU), zejména pracovní skupina *Video Quality Experts Group* (VQEG). Jejich cílem je jak vývoj nových technologií a standardů, tak i dohled nad již používanými.

1.1 Doporučení ITU-R BT.1683

Schopnost automaticky měřit kvalitu vysílaných videosekvencí je již dlouhou dobu považováno za cenný přínos pro telekomunikační odvětví. Telekomunikační průmysl vyžaduje, aby takové nástroje nahradily nebo doplnily nákladné a časově náročné subjektivní hodnocení kvality. Dříve bývalo objektivní měření kvality získáváno na základě výpočtu špičkového poměru signál šum (PSNR). Ačkoliv je PSNR stále užitečným ukazatelem kvality videosekvencí, bylo prokázáno, že korelace PSNR a výsledků subjektivních testů není příliš vysoká. Překonání omezení spojená s PSNR měl výzkum zaměřený na algoritmy, které mohou měřit vnímání kvality videosekvencí. Nástroje pro takové objektivní měření vnímání kvality mohou být například použita k testování kvality vysílací sítě, jako zařízení zákaznické podpory, anebo při vývoji nových technik kódování vysílání. V posledních letech bylo významné úsilí věnováno vývoji spolehlivých a přesných nástrojů, které mohou být použity k objektivnímu hodnocení vnímání kvality videosekvencí.

ITU-R BT.1683 definuje objektivní výpočetní modely, které se ukázaly být účinnější než PSNR, jako nástroje automatického měření hodnocení kvality videosekvencí. Modely byly testovány na 525-řádkových a 625-řádkových videosekvencích v souladu s doporučením ITU-R BT.601, které charakterizuje sekundární distribuci digitálně kódovaných signálů.

Výkon objektivních modelů byl posuzován pomocí dvou rozdílných způsobů hodnocení testovaných videosekvencí. V případě prvního hodnocení byla použita standardní subjektivní metoda DSCQS (The double-stimulus continuous quality-scale). Pozorovatelům jsou v náhodném pořadí prezentovány dvojice totožných videosekvencí. Prvně je promítána referenční videosekvence, která je po krátké odmlce následována testovanou videosekvencí, jež byla vytvořena z referenční videosekvence náhodnými druhy znehodnocení v různých kombinacích. Pozorovatelé mají za úkol ohodnotit úroveň kvality testované videosekvence vůči referenční videosekvenci, přičemž celý test by neměl trvat déle než půl hodiny [2]. Druhé hodnocení bylo provedeno za použití modelů pro objektivní hodnocení kvality videosekvencí. Pro každý model byla vypočítána přesnost a konzistence, s nimiž objektivní metody hodnocení předpověděly subjektivní hodnocení.

Tři na sobě nezávislé laboratoře provedly subjektivní hodnocení kvality videosekvencí metodou DSCQS. Laboratoře Communications Research Center (CRC, Kanada) a Verizon (USA) provedly zkoušky na videosekvencích o 525 řádcích s pulsnímkovým kmitočtem 60 Hz. Třetí laboratoř Fondazione Ugo Bordoni (FUB, Itálie) provedla zkoušky na videosekvencích o 625 řádcích s pulsnímkovým kmitočtem 50 Hz. Tabulka 1.1 uvádí laboratorní pracoviště, které provedly testy objektivních modelů pro hodnocení kvality videosekvencí na stejných videosekvencích jako laboratoře CRC, Verizon a FUB.

Tab. 1.1: Přehled objektivních modelů hodnocení kvality videosekvencí [1]

| Model | Název pracoviště | Označení (VQEG) | Země původu |
|--------------|---|------------------------|--------------------|
| 1 | British Telecom | D | Velká Británie |
| 2 | Yonsei University/Radio Research Laboratory/SK Telecom | E | Korea |
| 3 | Center for Telecommunications Research and Development (CPqD) | F | Brazílie |
| 4 | National Telecommunications and Information Administration/Institute for Telecommunication Science (NTIA/ITS) | H | USA |

2 METRIKA CPQD-IES

Metrika CPqD-IES, jejíž popis je uveden v doporučení ITU-R BT.1683, využívá k hodnocení kvality videosekvencí objektivní parametry získané na základě segmentace obrazu. Měřené videosekvence jsou segmentací rozčleněny na plošné oblasti, oblasti hran a oblasti textur, ke kterým se následně přiřazují objektivní parametry. Metrika je založena na předpovědi subjektivního hodnocení (podle doporučení ITU-R BT.500 a ITU-R BT.802), která je definována vztahem mezi parametry objektivních měření a výsledky subjektivních hodnocení. Objektivní měření se použijí na soubor videosekvencí zpracovaných video kodekem MPEG-2. V této metrice je navíc každý objektivní parametr vypočítán ještě pomocí dalších dvou kompresních systémů, které jsou tvořeny video kodeky MPEG-2 a MPEG-1. Vztah mezi objektivními parametry a subjektivním hodnocením kvality videosekvence je aproximován křivkou, která vede k odhadu úrovně kvality videosekvence pro každý objektivní parametr. Konečného výsledku je dosaženo prostřednictvím lineární kombinace jednotlivých odhadů úrovně kvality, kde váha každé úrovně je úměrná své statistické spolehlivosti.

2.1 Popis systému IES

Základní přehled o funkci CPqD-IES algoritmu podává obrázek 2.1. Každá testovaná videosekvence je zastoupena jednou referenční videosekvencí O a jednou znehodnocenou videosekvencí I , která je získána kodekovou operací aplikovanou na referenční videosekvenci O . Videosekvence I je podrobena korekci posunutí, čímž je vytvořena korigovaná videosekvence I' . Každý snímek f videosekvence I' odpovídá snímku f referenční videosekvence O pro $f = 1, 2, \dots, n$.

Videosekvence I a O vstupující do CPqD-IES algoritmu jsou vzorkovány ve formátu 4:2:2 $YC_B C_R$ dle doporučení ITU-R BT.601 (kódování parametrů digitální televize ve standardu 4:3 a širokouhlého obrazu s poměrem stran 16:9) [4].

Jasová složka obrazu Y , každého snímku f videosekvence O , je segmentací rozdělena do tří oblastí: plošné oblasti, oblasti hran a oblasti textur. Jedno objektivní měření je vypočítáno na základě rozdílu mezi odpovídajícími si snímky O a I' jasové složky obrazu Y . Další objektivní měření jsou vypočítány stejným způsobem pro chrominanční složky obrazu C_B a C_R . Z těchto objektivních měření je pro každý snímek f vytvořen soubor devíti parametrů objektivních měření $\{m_1, m_2, \dots, m_9\}$. Každý parametr objektivního měření m_i , $i = 1, 2, \dots, 9$ produkuje odhad úrovně kvality L_i na základě svého modelu znehodnocení, které je dáno vztahem [1]

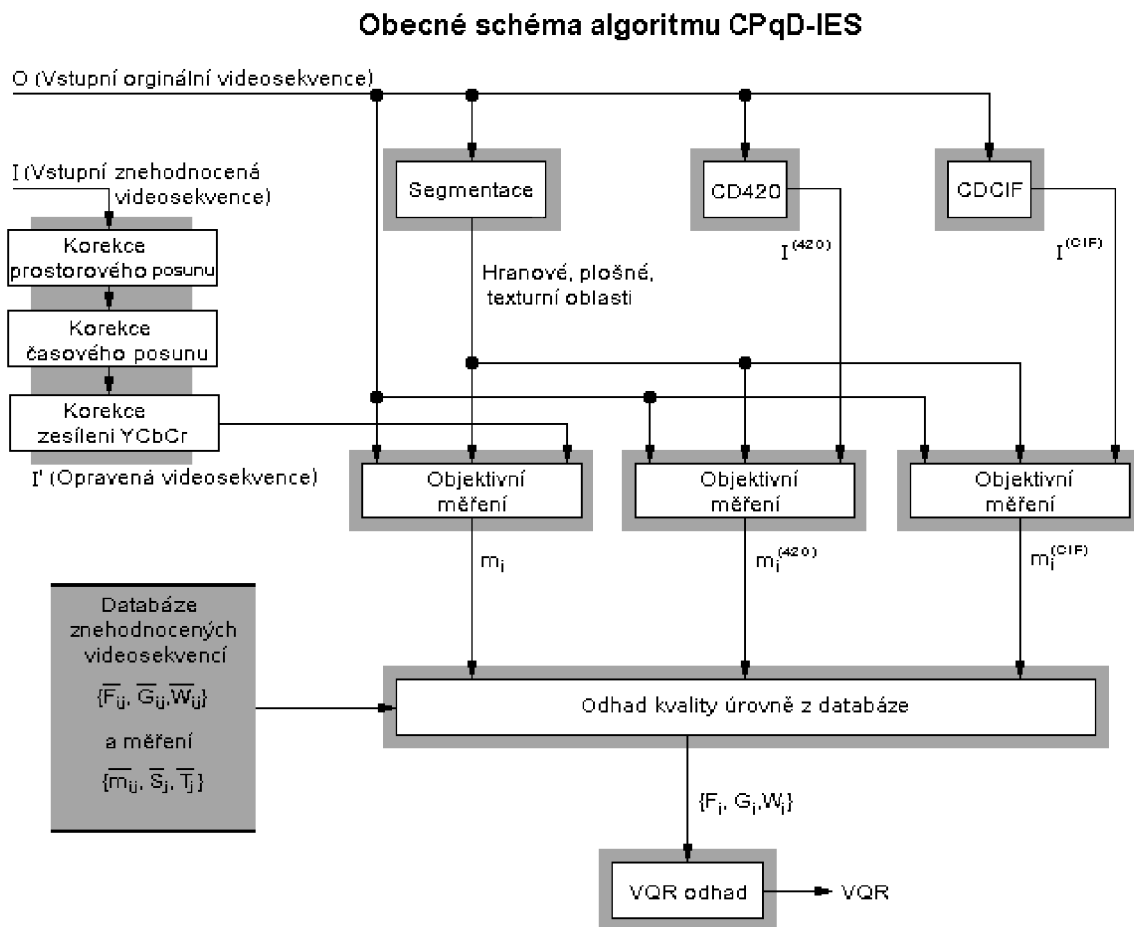
$$L_i = 100 / \left[1 + \left(\frac{F_i}{m_i} \right)^{G_i} \right], \quad (1)$$

kde F_i a G_i jsou parametry spočítané na základě dat v databázi znehodnocených videosekvencí, prostorových S a časových T atributů a objektivních měření $m_i^{(420)}$ a $m_i^{(CIF)}$ pro snímek f , které vyplývají z kodekových operací bloků CD420 a CDCIF

aplikovaných na referenční videosekvenci O .

K cílenému znehodnocení referenční videosekvence O slouží bloky CD420 (kodér/dekodér MPEG-2 4:2:0) a CDCIF (kodér/dekodér MPEG-1 CIF). Tyto kodekové operace respektují všechny procedury obsažené v MPEG2 (doporučení ITU-T H.262) a MPEG1 (ISO/IEC, 1992).

V současné realizaci algoritmu CPqD-IES pracují tyto kodekové operace v režimu s pevným kvantizačním krokem o velikosti 16. Důležitou vlastností bloků CD420 a CDCIF je, že nevnášají žádné posunutí obrazu ani rozdíly v zesílení oproti referenční videosekvenci O .



Obr. 2.1: Obecné schéma algoritmu CPqD-IES (převzato z [1])

Objektivní hodnocení kvality VQR_f snímku f se získá lineární kombinací odhadů úrovní kvality L_i , $i = 1, 2, \dots, 9$, které je následovně [1]:

$$VQR_f = \sum_{i=1}^9 W_i \cdot L_i, \quad (2)$$

kde W_i je váha odhadu úrovně kvality L_i pro každý objektivní parametr m_i , jejichž výpočty jsou popsány dále v textu.

Pořadí vypočítaných hodnot $VQR_1, VQR_2, \dots, VQR_n$ je transformováno mediánovým filtrem o velikosti 3 do jiného pořadí $VQR'_1, VQR'_2, \dots, VQR'_n$ tím, že je vyloučena střední hodnota v rámci okolí VQR_l a VQR_n . Během mediánové filtrace se algoritmus vyhýbá opakování dvou po sobě následujících středních hodnot. To znamená, že pokud je střední hodnota VQR'_{f-1} vypočítaná v rámci okolí VQR_f rovna střední hodnotě VQR'_{f-2} vypočítané v okolí VQR'_{f-1} , pak algoritmus vybere VQR'_{f-1} jako minimální hodnotu vypočtenou v rámci okolí VQR_f . Tento algoritmus lze popsat následovně.

- 1) Pro každý snímek f od 2 do $n - 1$,
- 2) Výpočet med , střední hodnoty z $VQR_{f-1}, VQR_f, VQR_{f+1}$
- 3) Jestliže je $med = VQR'_{f-2}$ potom
- 4) Výpočet VQR'_{f-1} jako minimální hodnota z $VQR_{f-1}, VQR_f, VQR_{f+1}$
- 5) Jinak
- 6) $VQR'_{f-1} \leftarrow med$.

Výsledné VQR je pak průměr hodnot z $VQR'_f[1]$.

$$VQR = \frac{1}{n-2} \cdot \sum_{f=1}^{n-2} VQR'_f, \quad (3)$$

kde n je počet snímků analyzované videosekvence.

Rovnice (1) a (2) popisují proces odhadu VQR z odhadnuté úrovně kvality dané parametry $\{F_i, G_i, W_i\}$ a parametry objektivního měření $m_i, i = 1, 2, \dots, 9$. Další kapitoly popisují detaily uvnitř zbývajících bloků na obr. 2.1.

2.2 Korekce posunutí a zesílení

Před vlastním objektivním měřením jsou provedeny korekce posunutí a zesílení mezi referenční videosekvencí O a znehodnocenou videosekvencí I . Tyto korekce slouží k eliminaci chyb, které nebyly do videosekvence vneseny během průběhu kodekových operací. Cílem je dosáhnout co možná největší shody mezi všemi obrazovými pixely videosekvence O a I , neboť je tím zpřesněn samotný odhad kvality videosekvence. Korekce se provádí nad omezeným počtem snímků obou videosekvencí.

2.2.1 Časové posunutí snímků

Časový posun dt je celé číslo v rozmezí od -2 do 2. Vstupní videosekvence s časovým posunem mimo tento rozsah nejsou brány v úvahu. Necht' I_{dt} je znehodnocená videosekvence I posunutá o f snímků. Mezi videosekvencí O a každým časově posunutým snímkem videosekvence I_{dt} je vypočítán koeficient odlišnosti. Časové posunutí s nejnižším koeficientem odlišnosti ξ_T je bráno jako časový posun a výstupní videosekvence I_{dt} je pak videosekvence I posunutá o tuto hodnotu pro další výpočty. Koeficient odlišnosti ξ_T mezi O a I_{dt} je získán následovně:

- 1) $\xi_T \leftarrow 0$
- 2) Pro každý snímek f od 1 do n
- 3) Výpočet S_b
- 4) Výpočet S'_b
- 5) Výpočet D_b
- 6) Výpočet μ , střední hodnoty pixelů v D_b
- 7) $\xi_T \leftarrow \xi_T + (\mu/n)$
- 8) Výstup ξ_T (koeficient odlišnosti mezi O a I_{dt}),

kde n je počet snímků použitých ke korekci časového posunutí mezi O a I_{dt} , S_b je velikost Sobelova gradientu z jasové složky Y f -tého snímku videosekvence O , S'_b je velikost Sobelova gradientu z jasové složky Y f -tého snímku videosekvence I_{dt} a D_b je velikost absolutního Sobelova rozdílu.

K výpočtu Sobelova gradientu je použito Sobelových operátorů, kde Sobelův operátor G_x slouží k detekci hran v horizontálním směru, a Sobelův operátor G_y k detekci hran ve vertikálním směru.

$$G_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad G_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Sobelův gradient S_b je vyjádřen následovně:

$$S_b = \sqrt{S_x^2 + S_y^2}, \quad (4)$$

kde S_x je obraz změn získaný konvolucí operátoru G_x s jasovou složkou Y f -tého snímku O a S_y je obraz změn získaný konvolucí operátoru G_y s jasovou složkou Y f -tého snímku O . Obdobně je vypočítán i Sobelův gradient S'_b , ale s využitím jasové složky Y f -tého snímku videosekvence I_{dt} .

Koeficient odlišnosti ξ_T je získán jako střední hodnota ze všech absolutních Sobelových rozdílů D_b f -tého snímku, kde absolutní Sobelův rozdíl je vypočítán jako:

$$D_b = |S_b - S'_b|. \quad (5)$$

Časový posun může nabývat hodnot v rozmezí od -2 do 2, proto může být videosekvence I_{dt} při korekci zkrácena o jeden až dva snímky k zajištění vzájemné synchronizace mezi snímky obou videosekvencí.

2.2.2 Prostorové posunutí snímku

Prostorový posun (dx, dy) nabývá hodnoty jednoho z následujících celých čísel horizontálního a vertikálního posunu $dx = -6, -5, \dots, 6$ a $dy = -6, -5, \dots, 6$. Videosekvence $I_{dx,dy}$ je dále brána jako upravená videosekvence I_{dt} se všemi snímky posunutými horizontálně a vertikálně o (dx, dy) pixelů. Mezi O a $I_{dx,dy}$ je vypočítán koeficient

odlišnosti ξ_s . Prostorové posunutí s nejnižším koeficientem odlišnosti ξ_s je použito jako prostorový posun a výstup $I_{dx,dy}$ je pak videosekvence I_{dt} posunutá o tuto hodnotu pro další výpočty. Koeficient odlišnosti ξ_s mezi videosekvencemi O a $I_{dx,dy}$ je popsán:

- 1) $\xi_s \leftarrow 0$; $c \leftarrow 0$
- 2) Pro každý snímek f od 1 do n
- 3) Pro x od x_0 do $(x_0 + w/4)$
- 4) Pro y od y_0 do $(y_0 + h/4)$
- 5) $\xi_s \leftarrow \xi_s + |Y(4x,4y) - Y'(4x + dx, 4y + dy)| +$
 $+ |Cb(4x,4y) - Cb'(4x + dx, 4y + dy)| +$
 $+ |Cr(4x,4y) - Cr'(4x + dx, 4y + dy)|$
- 6) $c \leftarrow c + 3$
- 7) $\xi_s \leftarrow \xi_s / c$
- 8) Výstup ξ_s (koeficient odlišnosti mezi O a $I_{dx,dy}$),

kde $w \times h$ jsou rozměry snímku, x_0 a y_0 jsou hodnoty pixelů, od kterých je počítána korekce prostorového posunutí mezi snímky videosekvence O a $I_{dx,dy}$. $Y(x, y)$, $Cb(x, y)$, $Cr(x, y)$ jsou hodnoty obrazových složek ve snímku f videosekvence O pro pixel (x, y) . $Y'(x + dx, y + dy)$, $Cb'(x + dx, y + dy)$, $Cr'(x + dx, y + dy)$ jsou hodnoty obrazových složek ve snímku f videosekvence $I_{dx,dy}$ pro pixel $(x + dx, y + dy)$. Proměnná c slouží k znormování koeficientu odlišnosti ξ_s .

Prostorový koeficient odlišnosti ξ_s je počítán z horizontálních a vertikálních přírůstků o velikosti 4 pixely, tudíž jako absolutní rozdíl každého 16-tého vzorku videosekvence O a videosekvence I_{dt} , která je posunutá o horizontální nebo vertikální posun v rozmezí od -6 do 6.

Jelikož jsou videosekvence ve formátu digitálního kódování 4:2:2, chrominanční složky C jsou vzorkovány s poloviční frekvencí v horizontálním směru oproti jasové složce Y . Tomu odpovídá struktura, kdy je jeden vzorek chrominanční složky C přiřazen ke dvěma vzorkům jasové složky Y . Z tohoto důvodu jsou nakonec brány v úvahu horizontální přírůstky pro složku Y o velikosti 8 pixelů oproti 4 pixelům pro složky C a výpočet pro složku Y je upraven do tvaru $|Y(8x,4y) - Y'(8x + dx, 4y + dy)|$.

2.2.3 Zesílení

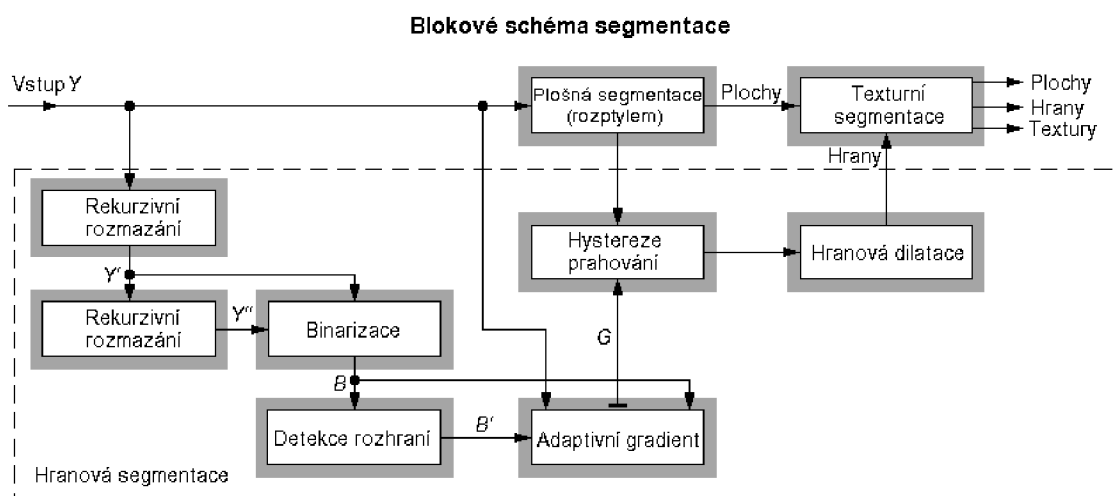
Zesílení mezi videosekvencemi O a $I_{dx,dy}$ je vypočítáno pro každou složku obrazu Y , C_B a C_R zvlášť. Algoritmus vypočítá průměrné zesílení přes všechny snímky n a podle toho opravuje každou obrazovou složku. Výstup I' je videosekvence $I_{dx,dy}$ s opraveným zesílením, která se použije pro všechny následující výpočty. Zesílení mezi obrazovou složkou snímku f ve videosekvenci $I_{dx,dy}$ s ohledem na stejnou složku snímku f ve videosekvenci O je získáno rozmazáním obou snímků použitím Gaussova filtru s maticí:

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

Při výpočtu poměru mezi rozmazanými snímky je brán v úvahu pouze jeden z 16 pixelů (horizontální a vertikální přírůstky o velikosti 4 pixely), stejně jako při výpočtu koeficientu odlišnosti ξ_s v podkapitole o prostorovém posunutí snímku (2.2.2).

2.3 Segmentace obrazu

Segmentační algoritmus nejprve klasifikuje každý pixel v obrazové složce Y daného snímku f referenční videosekvence O na plošné a neplošné oblasti. K definici oblasti hran je použit hranový detektor aplikovaný na jasovou složku Y . Jako hrany jsou poté vyhodnoceny pixely, které spadají do již dříve vymezené oblasti ploch. Oblast textur se skládá ze zbývajících pixelů obrazu Y (viz Obr. 2.2).



Obr. 2.2: Blokové schéma segmentace (převzato z [1])

Segmentace je vypočítaná z každého snímku jasové složky Y originální videosekvence O . Pro chrominanční složky C_B a C_R jsou oblasti klasifikovány pozicí pixelů složky Y po převzorkování na velikost složek C_B a C_R .

2.3.1 Plošné oblasti

Rozptyl jasu jednotlivých pixelů ve složce Y je počítán z okolí 5×5 pixelů. Rozptyl obrazu je prahován tak, že pixely s hodnotou rozptylu nižší než 25^2 jsou klasifikovány jako plošné oblasti. Tento proces vytváří malé skupiny pixelů chybně určených jako oblasti textur. K odstranění těchto malých skupin pixelů se používá mediánový filtr o velikosti 3×3 pixelů. Nakonec je binární obraz plošných oblastí podroben morfologické dilataci s využitím kruhového elementu o průměru 11 obrazových bodů [3].

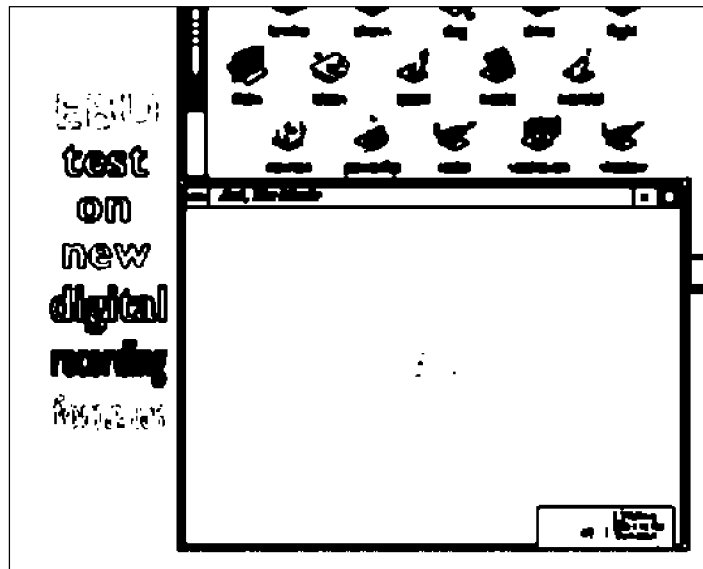
Klasifikace plošných oblastí je založena na výpočtu rozptylu jednotlivých pixelů v okolí 5×5 pixelů jasové složky Y . Jelikož okrajové pixely nemohou mít dostatečně velké okolí, jak je požadováno, počítá se jejich hodnota rozptylu z menšího okolí. Rozptyl je definován jako rozdíl kvadrátu odchylky a střední hodnoty. Z této definice vyplývá, že pokud budou uvnitř uvažovaného okolí pixely, jejichž hodnoty budou

navzájem podobné, pak se jedná o plochu s rovnoměrným jasem okolí.

Rozptyl jasového obrazu je následně prahován a pixely s hodnotou rozptylu nižší než 25^2 jsou klasifikovány jako plošné oblasti. Vzniká tím binární obraz, kde pixely patřící do plošné oblasti nabývají hodnoty jedna. Na obrázku 2.3 je vyobrazen příklad snímku jasové složky Y videosekvence O , jenž je použit k segmentaci obrazu a na obrázku 2.4 jsou bíle vyznačeny plošné oblasti pro tentýž snímek.



Obr. 2.3: Obrázek jasové složky Y videosekvence O



Obr. 2.4: Plošné oblasti jasové složky Y videosekvence O

2.3.2 Oblasti hran

Ze snímku jasové složky Y referenční videosekvence O je algoritmem, popsaným níže, vytvořen první rozmazaný obraz Y' a na tento již jednou rozmazaný obraz je poté použit stejný algoritmus za účelem vytvoření druhého rozmazaného obrazu Y'' .

Algoritmus vytváří rozmazaný obraz Y' postupným posouváním snímku jasové složky Y ve všech horizontálních a vertikálních směrech. Nejprve vznikne první posunutý obraz v horizontálním směru, který je dále brán jako zdrojový obraz pro výpočet druhého posunutého obrazu a celý proces se opakuje dokud nejsou provedeny posuny pro všechny horizontální a vertikální směry. Jelikož je každý výsledek předchozího kroku brán jako zdroj pro další krok, jedná se vlastně o jistý druh rekurzivního procesu či filtrace.

- 1) Pro y měnící se od 0 do $(h - 1)$
- 2) Pro x měnící se od 0 do $(w - 2)$
- 3) $Y(x + 1, y) \leftarrow Y(x, y) + 0.7 [Y(x + 1, y) - Y(x, y)];$
- 4) Pro y měnící se od 0 do $(h - 1)$
- 5) Pro x měnící se od $(w - 1)$ do 1
- 6) $Y(x - 1, y) \leftarrow Y(x, y) + 0.7 [Y(x - 1, y) - Y(x, y)];$
- 7) Pro x měnící se od 0 do $(w - 1)$
- 8) Pro y měnící se od 0 do $(h - 2)$
- 9) $Y(x, y + 1) \leftarrow Y(x, y) + 0.7 [Y(x, y + 1) - Y(x, y)];$
- 10) Pro x měnící se od 0 do $(w - 1)$
- 11) Pro y měnící se od $(h - 1)$ do 1
- 12) $Y(x, y + 1) \leftarrow Y(x, y) + 0.7 [Y(-1) - Y(x, y)];$
- 13) Uložení obrazu Y do obrazu Y' ,

kde $Y(x, y)$ reprezentuje jas pixelu na souřadnicích (x, y) , h je počet řádků a w je počet sloupců ve snímku jasové složky Y . Druhé provedení uvedeného algoritmu vytváří rozmazaný obraz Y'' .

Rekurzivní proces použitý při detekci hran potlačuje náhodný šum v obraze. Vyskytuje-li se v původním snímku jasové složky Y nějaká dynamicky se měnící oblast, pak se její vzájemná poloha v rozmazaných obrazech Y' a Y'' bude lišit. Porovnáním rozmazaných obrazů je vytvořen binární obraz B [1]:

$$B(x, y) = \begin{cases} 1 & \text{pokud } Y'(x, y) \geq Y''(x, y) \\ 0 & \text{jinak} \end{cases}, \quad (6)$$

K určení hranových pixelů je dále testováno okolí pixelů náležejícího do binárního obrazu B . Pokud je pixel $B(x, y)$ roven jedné a v okolí 8 sousedních pixelů je alespoň jeden pixel roven nule, pak se jedná o hranový pixel a je zapsán do druhého binárního obrazu B' [1]:

$$B'(x, y) = \begin{cases} 1 & \text{pokud } B(x, y) = 1 \text{ a } B(x', y') = 0 \\ & \text{pro kterýkoliv pixel } (x', y') \in N_8(x, y), \\ 0 & \text{jinak} \end{cases}, \quad (7)$$

kde $N_8(x, y)$ je množina 8 pixelů (x', y') , jež se nachází uvnitř oblasti 3×3 pixelů se středem v (x, y) , tj. 8 sousedních pixelů.

Obraz G je vytvořen aplikací adaptivního gradientního filtru na vybrané pixely jasové složky Y , výběr pixelů jasové složky Y závisí na binárních obrazech B a B' [1]:

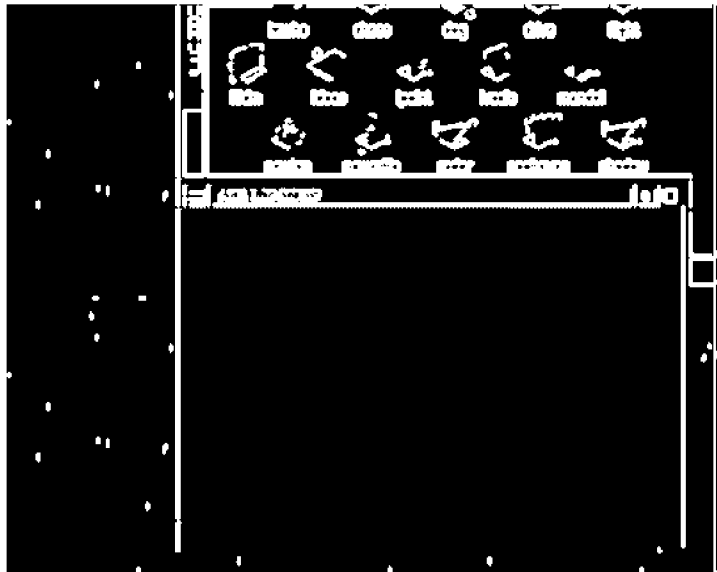
$$G(x, y) = \begin{cases} |\mu_1 - \mu_0| & \text{pokud } B'(x, y) = 1 \\ 0 & \text{jinak} \end{cases} \quad (8)$$

Algoritmus nejprve detekuje pixely v binárním obraze $B'(x, y)$ nabývající hodnoty jedna a poté vymezi okolí 8 sousedních pixelů $N_8(x, y)$, takto dané sousední pixely jsou určeny souřadnicemi (x', y') . Při vlastním výpočtu středních hodnot μ_0 a μ_1 však algoritmus používá binární obraz B , nikoliv B' . Střední hodnota μ_1 je vypočítána z pixelů jasové složky $Y(x', y')$ pro něž platí $B(x', y') = 1$ a naopak střední hodnota μ_0 je vypočítána z pixelů jasové složky $Y(x', y')$ pro něž platí $B(x', y') = 0$. Absolutním rozdílem středních hodnot μ_1 a μ_0 je vytvořen obraz G , jehož pixely nabývají hodnot od 0 do 255. Na obrázku 2.5 je vyobrazen obraz G , bíle jsou vyznačeny odpovídající hranové oblasti.



Obr. 2.5: Hranové oblasti před prahováním a dilatací

Hystereze prahování je aplikována na vyhrazené pixely v obraze G , které jsou současně klasifikovány jako plošné oblasti. Dolní prahová hranice je stanovena na hodnotu 30 a horní hranice na hodnotu 40. Algoritmus nejprve identifikuje pixely v obraze G , které splňují podmínku $G(x, y) > 40$, a poté na ně použije algoritmus pro růst oblastí. Tímto způsobem vybrané pixely jsou použity jako „semínka“ (tj. počáteční body růstu) a hranice růstu je omezena podmínkou $G(x, y) > 30$. Všechny součásti vertikálně a horizontálně propojených pixelů s méně než 6 body jsou vyloučeny z výsledku růstu. Konečný binární obraz je dilatován s využitím kruhového elementu o průměru 5 obrazových bodů, který ignoruje omezení vyhrazené plošnými oblastmi. Hranové oblasti po prahování a dilataci nabývají hodnoty jedna a v obrázku jsou vyznačeny bíle (obr. 2.6).

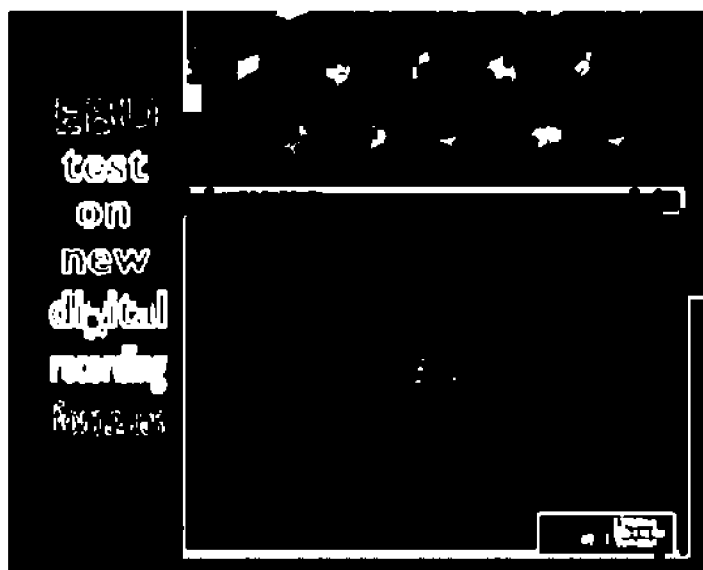


Obr. 2.6: Hranové oblasti po prahování a dilataci

Po aplikaci algoritmu prahování zmizela z obr. 2.5 spousta oblastí, jenž byla tvořena méně než 6 pixely a taktéž oblasti, které byly sice tvořeny více než 6 pixely, ale nesplňovaly podmínky prahování. Tento algoritmus tedy klasifikuje jako hranové oblasti jen velmi ostré a dynamické přechody v jasové složce Y . To může být problémem u některých druhů videosekvencí, kde je většina hranových oblastí algoritmem zamítnuta, protože hrany nejsou dostatečně výrazné.

2.3.3 Oblasti Textur

Oblasti textur se skládají z pixelů jasové složky obrazu Y , které nebyly v předchozí části segmentace klasifikovány jako pixely spadající do oblasti hran ani oblasti ploch. Výpočet texturních oblastí je realizován logickými operacemi nad oblastmi ploch a hran, tím vzniká binární obraz 2.7, kde jsou texturní oblasti zobrazeny bílou barvou.



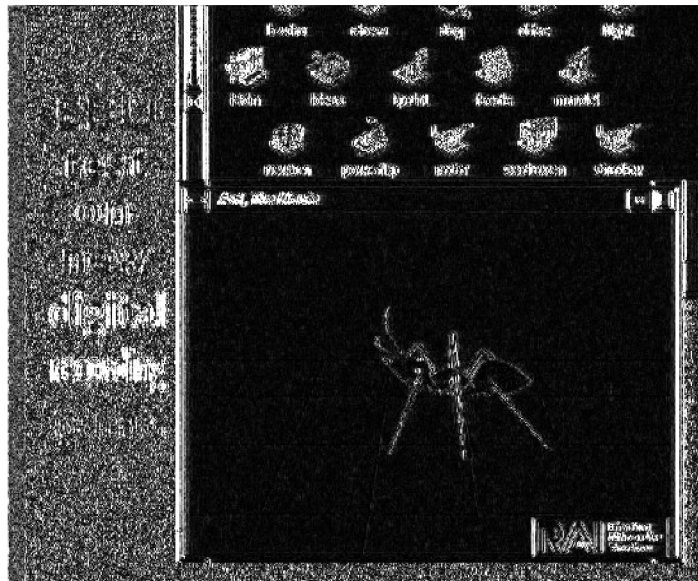
Obr. 2.7: Texturní oblasti

2.4 Objektivní měření

Základem objektivního měření je výpočet absolutního Sobelova rozdílu ve vymezeném prostoru \mathfrak{R} , který je definován rozčleněním každého snímku videosekvence během procesu segmentace. Vymezený prostor \mathfrak{R} je dán binárními obrazy příslušející ke každé oblasti segmentace (plochy, hrany a textury).

Absolutní Sobelův rozdíl je z každého snímku videosekvence O a I' získán totožně, jak je popsáno v kapitole 2.2.1 o časovém posunutí snímků. Pro vybranou obrazovou složku Y , C_B a C_R snímku f referenční videosekvence O je vypočítán obraz Sobelova gradientu S_b . Druhý obraz Sobelova gradientu S'_b je vypočítán pro vybranou složku Y , C_B a C_R daného snímku f znehodnocené videosekvence I' . Absolutní Sobelův rozdíl je reprezentován obrazem D_b (obr 2.8), který je vypočítán jako absolutní hodnota rozdílu mezi pixely obrazů S_b a S'_b . Parametr objektivního měření m je poté počítán jako střední hodnota z vymezeného prostoru \mathfrak{R} v obraze absolutního Sobelova rozdílu D_b . Do výpočtu jsou tedy zahrnuty pouze pixely příslušející jedné ze tří oblastí segmentace.

Tento postup vytváří soubor devíti parametrů objektivních měření $\{m_1, m_2, \dots, m_9\}$ pro každý snímek obrazu $f, f = 1, 2, \dots, n$, s ohledem na všechny tři segmentační oblasti a tři složky obrazu.



Obr. 2.8: Obraz absolutního Sobelova rozdílu D_b

Stejný postup se pak použije při vytváření parametrů objektivních měření $\{m_1^{(420)}, m_2^{(420)}, \dots, m_9^{(420)}\}$ a $\{m_1^{(CIF)}, m_2^{(CIF)}, \dots, m_9^{(CIF)}\}$ pro snímek f s ohledem na použité kodekové operace CD420 a CDCIF aplikované na referenční videosekvenci O (obr. 2.1). Parametrem $m^{(420)}$ jsou označovány objektivní parametry měření získané z referenční videosekvence O po průchodu blokem CD420, tzn. kódování a dekódování videosekvence kodekem MPEG-2 při převzorkování na formát 4:2:0 $Y_C_B C_R$. Parametrem $m^{(CIF)}$ jsou označovány objektivní parametry měření získané z referenční videosekvence O po průchodu blokem CDCIF, tzn. kódování a dekódování kodekem MPEG-1 při převzorkování na formát 4:2:0 a zmenšení obrazových složek na rozlišení CIF. Kódování a dekódování je provedeno pomocí programu FFmpeg [8].

Objektivní měření je prováděno pro různé velikosti obrazových složek $YC_B C_R$ (4:2:2, 4:2:0 a CIF). Vstupní videosekvence O a $I_{dx,dy}$ jsou ve formátu vzorkování 4:2:2 $YC_B C_R$, ale protože znehodnocené videosekvence $I^{(420)}$ a $I^{(CIF)}$, už mají jiný formát vzorkování, tak se musí před vlastním objektivním měřením převzorkovat obrazy absolutního Sobelova rozdílu D_b pro znehodnocené videosekvence.

Po průchodu blokem CD420 je videosekvence ve formátu vzorkování 4:2:0 $YC_B C_R$ a proto je nutno podvzorkovat obraz D_b u chrominančních složek C ve vertikálním směru na polovinu. Obraz D_b pro jasovou složku Y se po průchodu blokem CD420 nemění. Po průchodu blokem CDCIF je videosekvence také ve formátu vzorkování 4:2:0 $YC_B C_R$, ale je snížen počet obrazových bodů jak ve vertikálním, tak i horizontálním směru. Pro výpočet parametrů s označením CIF je tedy nutno podvzorkovat obraz D_b pro jasovou složku Y na polovinu v obou směrech a obrazu D_b pro chrominanční složky C na čtvrtinu ve vertikálním směru a na polovinu v horizontálním směru.

Objektivní parametry měření jsou používány společně s prostorovými S a časovými T atributy, k určení odhadu úrovně kvality znehodnocené videosekvence I' . Časový atribut T je vypočítán jako střední hodnota z absolutních rozdílů pixelů mezi oblastmi klasifikovanými segmentací snímků f a $f-1$, normalizován v intervalu $\langle 0,1 \rangle$. Prostorový atribut S je definován jako poměr $m_7^{(CIF)} / m_7^{(420)}$, který je normalizován v intervalu $\langle 0,1 \rangle$, kde $m_7^{(CIF)}$ a $m_7^{(420)}$ jsou parametry objektivních měření odpovídající střední hodnotě z vymezeného prostoru \mathfrak{R} pro oblast textur v obraze absolutního Sobelova rozdílu D_b obrazové složky Y snímku f .

2.5 Databáze znehodnocených videosekvencí

Systém metriky CPqD-IES využívá databázi znehodnocených videosekvencí pro odhad míry kvality testované videosekvence I' z referenční videosekvence O . Tato databáze, podle doporučení ITU-R BT.1683, obsahuje informace o dvanácti videosekvencích, s pulsníkovým kmitočtem 60 Hz, jenž reprezentující různé stupně pohybu (dynamické a statické videosekvence), původu (reálné a syntetické videosekvence) a klasifikace segmentace (množství texturních, plošných a hranových pixelů). Parametry databáze jsou vytvořeny zprůměrováním parametrů získaných při objektivních měřeních, proto jsou označovány vodorovným pruhem nad názvem parametru.

Databáze znehodnocených videosekvencí byla vytvořena na základě průměrných hodnot objektivních měření a subjektivních hodnocení. Průměrné hodnoty objektivních parametrů $\{\bar{m}_{1,j}, \bar{m}_{2,j}, \dots, \bar{m}_{9,j}\}$, $\{\bar{m}_{1,j}^{(420)}, \bar{m}_{2,j}^{(420)}, \dots, \bar{m}_{9,j}^{(420)}\}$ a $\{\bar{m}_{1,j}^{(CIF)}, \bar{m}_{2,j}^{(CIF)}, \dots, \bar{m}_{9,j}^{(CIF)}\}$ byly vypočítány přes snímky každé videosekvence j , $j = 1, 2, \dots, 12$. Hodnoty atributů S_j a T_j byly vypočteny jako průměr prostorových a časových atributů přes snímky každé videosekvence j . Všechny poškozené videosekvence obsažené v databázi byly rovněž předloženy k subjektivnímu hodnocení. Získané subjektivní hodnocení kvality SL_j bylo normalizováno hodnotami v rozmezí mezi 0% až 100% pro každou videosekvenci j .

Podle rovnice (1) je každé objektivní měření $m_{i,j}$, $i = 1, 2, \dots, 9$ a $j = 1, 2, \dots, 12$ spojováno do souvislosti s odhadem úrovně kvality $L_{i,j}$. Hodnoty $F_{i,j}$ a $G_{i,j}$ ve stejné rovnici byly zjištěny minimalizací očekávané střední kvadratické chyby

$E[(\overline{SL}_j - \overline{L}_{i,j})^2]$ pro každou videosekvenci j . Navíc hodnoty W_{ij} v rovnici (2) byly vypočítány s cílem minimalizovat očekávanou střední kvadratickou chybu [1]:

$$E \left[\left(\overline{SL}_j - \sum_{i=1}^9 \overline{W}_{i,j} \overline{L}_{i,j} \right)^2 \right]. \quad (9)$$

Databáze znehodnocených videosekvencí se podle doporučení ITU-R BT.1683 skládá z devíti souborů $\{\overline{F}_{i,j}, \overline{G}_{i,j}, \overline{W}_{i,j}, \overline{S}_j, \overline{T}_j\}$, o $i = 1, 2, \dots, 9$ parametrech pro každou videosekvenci $j, j = 1, 2, \dots, 12$. Tabulka 2.1 obsahuje prostorové a časové atributy \overline{S}_j a \overline{T}_j , které by měli sloužit pro výpočet parametrů $\{\overline{F}_{i,j}, \overline{G}_{i,j}, \overline{W}_{i,j}\}$.

Tab. 2.1: Přehled časových a prostorových atributů dle ITU-R BT.1683 [1]

| Scene j | T (temporal) | S_T (spatial Y) | S_{Cb} (spatial C_B) | S_{Cr} (spatial C_R) |
|-----------|----------------|----------------------|---------------------------|---------------------------|
| 1 | 27.01 | 36.79 | 25.20 | 38.01 |
| 2 | 25.33 | 26.08 | 5.93 | 67.99 |
| 3 | 45.54 | 60.97 | 10.28 | 28.75 |
| 4 | 36.40 | 30.47 | 6.46 | 63.07 |
| 5 | 32.02 | 72.50 | 11.72 | 15.78 |
| 6 | 12.63 | 84.22 | 2.85 | 12.94 |
| 7 | 28.38 | 61.53 | 11.08 | 27.39 |
| 8 | 10.19 | 46.08 | 5.45 | 48.47 |
| 9 | 0.01 | 5.89 | 5.07 | 89.03 |
| 10 | 7.26 | 4.75 | 2.00 | 93.25 |
| 11 | 7.60 | 69.16 | 9.41 | 21.43 |
| 12 | 14.27 | 69.61 | 3.89 | 26.50 |

Po důkladném prostudování doporučení ITU-R BT.1683 bylo zjištěno, že databáze poškozených videosekvencí nelze zrekonstruovat pouze ze znalosti uvedených prostorových a časových atributů v tabulce 2.1. Z tohoto důvodu bylo rozhodnuto vytvořit novou databázi znehodnocených videosekvencí z vlastních objektivních měření. Kapitola 3.2 je věnována postupu při vytváření databáze znehodnocených videosekvencí.

2.6 Odhad úrovně kvality

Odhad úrovně kvality pro daný snímek f videosekvence I' se skládá z parametrů $\{F_i, G_i, W_i\}$ obsažených v rovnicích (1) a (2) pro $i = 1, 2, \dots, 9$. Tato část textu popisuje, jak spočítat tyto parametry pomocí znehodnocených videosekvencí $I^{(420)}$ a $I^{(Clp)}$, získaných kodekovými operacemi nad referenční videosekvencí O .

2.6.1 Výpočet parametru W_i

Parametr W_i slouží k váhování odhadu úrovně kvality L_i , viz rovnice (2). Výpočet vychází z lokální vzdálenosti $D_{i,j}$, která vyjadřuje podobnost mezi jednotlivými snímky f znehodnocených videosekvencí $I^{(420)}$ a $I^{(CIF)}$, a každé videosekvence j v databázi [1]:

$$D_{i,j} = \frac{1}{2} \cdot \left(\left| L_{i,j}^{(420)} - \bar{L}_{i,j}^{(420)} \right| + \left| L_{i,j}^{(CIF)} - \bar{L}_{i,j}^{(CIF)} \right| \right), \quad (10)$$

kde index i představuje složku obrazu Y , C_B nebo C_R s ohledem na klasifikované oblasti během segmentace, tzn. $i = \{1, 2, \dots, 9\} \equiv \{(plocha, Y), (plocha, C_B), (plocha, C_R), (hrana, Y), (hrana, C_B), (hrana, C_R), (textura, Y), (textura, C_B), (textura, C_R)\}$. Index j udává vybranou videosekvenci z databáze znehodnocených videosekvencí. Proměnné $L_{i,j}$ vyjadřují odhady úrovně kvality pro znehodnocené videosekvence $I^{(420)}$ a $I^{(CIF)}$, které jsou jednou vypočítány pro objektivní parametry $\bar{m}^{(420)}$ a $\bar{m}^{(CIF)}$ pocházející z databáze a po druhé pro parametry $m^{(420)}$ a $m^{(CIF)}$ získané během vlastního objektivního měření. Při výpočtu jsou dále použity databázové parametry $\bar{F}_{i,j}$ a $\bar{G}_{i,j}$ [1]:

$$\bar{L}_{i,j}^{(420)} = 100 / \left[1 + \left(\bar{F}_{i,j} / \bar{m}_{i,j}^{(420)} \right)^{\bar{G}_{i,j}} \right], \quad (11)$$

$$\bar{L}_{i,j}^{(CIF)} = 100 / \left[1 + \left(\bar{F}_{i,j} / \bar{m}_{i,j}^{(CIF)} \right)^{\bar{G}_{i,j}} \right], \quad (12)$$

$$L_{i,j}^{(420)} = 100 / \left[1 + \left(\bar{F}_{i,j} / m_i^{(420)} \right)^{\bar{G}_{i,j}} \right], \quad (13)$$

$$L_{i,j}^{(CIF)} = 100 / \left[1 + \left(\bar{F}_{i,j} / m_i^{(CIF)} \right)^{\bar{G}_{i,j}} \right], \quad (14)$$

Algoritmus poté vybere z databáze znehodnocených videosekvencí soubor šesti nejpodobnějších videosekvencí Ω založených na vzdálenosti $D_{i,j}$ a definuje $W_{i,j}^{(1)}$ jako [1]:

$$W_{i,j}^{(1)} = \frac{a(\Omega)_{i,j} D_{i,j}^{-1}}{\sum_{k=1}^{12} a(\Omega)_{i,k} D_{i,k}^{-1}}, \quad (15)$$

$$a(\Omega)_{i,j} = \begin{cases} 1 & \text{pokud (videosekvence } k) \in \Omega \\ 0 & \text{jinak} \end{cases}. \quad (16)$$

Výběr šesti nejpodobnějších videosekvencí Ω je prováděn ze všech videosekvencí v databázi přes jednotlivé složky obrazu Y , C_B nebo C_R s ohledem na klasifikované oblasti během segmentace, tzn. pro každé $i = \{1, 2, \dots, 9\}$. Vybrané videosekvence jsou

definovány proměnnou a_k nabývající hodnoty jedna. Hodnota váhovacího parametru je poté vypočítána dle rovnice (15) pro všechny vzdálenosti $D_{i,j}$ patřící do Ω , jinak nabývá hodnoty nula a tím vylučuje ostatní videosekvence v databázi z podílu na odhadu výsledné kvality testované videosekvence. Jmenovatel je vypočítán jako suma hodnot přes všechny vzdálenosti $D_{i,j}$ patřící do výběru Ω pro danou složku obrazu a klasifikovanou oblast.

Pokud je definováno $u = \textit{textury, hrany, plochy}$ a $v = Y, C_B, C_R$, poté jsou hodnoty parametru $W_i, i = 1, 2, \dots, 9$, vypočítány jako [1]:

$$W_i = \frac{\kappa_i}{\tau} \cdot \frac{1}{E_i}, \quad (17)$$

kde E udává váhu jednotlivých složek obrazu Y, C_B nebo C_R s ohledem na klasifikovanou oblast během segmentace. Jelikož je E vyjádřeno násobením vzdáleností $D_{i,j}$ s parametrem $W_{i,j}^{(1)}$, dochází vlastně k jistému zprůměrování hodnot pro všechny složky obrazu a související klasifikované oblasti segmentace. Proměnná κ rozděluje složky obrazu zvláště na jasovou a chrominanční, což znamená, že jasové složky Y mají ve výsledku větší váhu než chrominanční složky C_B nebo C_R . Proměnná τ provádí normalizaci hodnot E pro jednotlivé oblasti segmentace $u = \textit{textury, hrany, plochy}$. Výpočet E, κ a τ je popsán následovně [1]:

$$E_i = \sum_{j=1}^{12} D_{i,j} W_{i,j}^{(1)}, \quad (18)$$

$$\kappa_i = \begin{cases} 1 & \text{pokud } v = Y_i \\ \frac{1}{2} & \text{jinak} \end{cases}, \quad (19)$$

$$\tau = \sum_u \left[\frac{1}{E_{u,Y}} + \frac{1}{2} \left(\frac{1}{E_{u,C_B}} + \frac{1}{E_{u,C_R}} \right) \right]. \quad (20)$$

2.6.2 Výpočet parametrů F_i a G_i

Odhady úrovně kvality $L(VQR)_i^{(420)}$ a $L(VQR)_i^{(CIF)}$ snímku f pro znehodnocenou referenční videosekvenci O (MPEG-1 a MPEG-2) jsou počítány podle rovnic [1]:

$$L(VQR)_i^{(420)} = \frac{1}{\gamma} \cdot \sum_{j=1}^{10} W_{i,j} L_{i,j}^{(420)}, \quad (21)$$

$$L(VQR)_i^{(CIF)} = \frac{1}{\gamma} \cdot \sum_{j=1}^{10} W_{i,j} L_{i,j}^{(CIF)}, \quad (22)$$

kde $L_{i,j}^{(420)}$ a $L_{i,j}^{(CIF)}$ jsou odhady úrovně kvality získané při výpočtu parametru W_i , rovnice (13) a (14), a parametr $W_{i,j}^{(1)}$ v rovnici (15). Koeficient γ provádí korekci odhadnuté úrovně kvality na základě prostorových a časových atributů videosekvence, přičemž je omezený v intervalu $\langle 1/2, 2 \rangle$ [1]:

$$\gamma = 1 + a - b, \quad (23)$$

kde koeficienty a a b jsou vypočítány na základě prostorových a časových atributů (S_j, T_j) hodnocené videosekvence a atributů (\bar{S}_j, \bar{T}_j) z databáze [1]:

$$a = \sum_{j=1}^{12} w_j \left[\frac{\bar{S}_j \cdot \bar{T}_j}{2} + (1 - \bar{T}_j^2) \cdot \left(1 - \frac{\bar{S}_j^2}{2} \right) \right], \quad (24)$$

$$b = \frac{S \cdot T}{2} + (1 - T^2) \cdot \left(1 - \frac{S^2}{2} \right). \quad (25)$$

kde w_j definuje váhu jednotlivých videosekvencí v databázi na základě vzdálenosti $D(ST)_j$ mezi prostorovými a časovými atributy (S_j, T_j) hodnocené videosekvence a atributy (\bar{S}_j, \bar{T}_j) z databáze [1]:

$$w_j = \frac{D(ST)_j^{-1}}{\sum_{k=1}^{12} D(ST)_k^{-1}}, \quad (26)$$

$$D(ST)_j = (S - \bar{S}_j)^2 + (T - \bar{T}_j)^2. \quad (27)$$

Výsledné parametry F_i a G_i jsou získané vyřešením soustavy o dvou rovnicích uvedených níže [1]:

$$L(VQR)_i^{(420)} = 100 / \left[1 + \left(\frac{F_i}{m_i^{(420)}} \right)^{G_i} \right], \quad (28)$$

$$L(VQR)_i^{(CIF)} = 100 / \left[1 + \left(\frac{F_i}{m_i^{(CIF)}} \right)^{G_i} \right]. \quad (29)$$

Parametry F_i a G_i jsou poté použity při odhadu úrovně kvality testované videosekvence v rovnici (1). K vyřešení soustavy rovnic je nutno zavést pomocnou proměnnou x , která definuje exponent mocniny a odmocniny, následně lze dopočítat parametry F_i a G_i jako:

$$x_i = \frac{\log\left(\frac{100}{L(VQR)_i^{(CIF)}} - 1\right)}{\log\left(\frac{100}{L(VQR)_i^{(420)}}\right)}, \quad (30)$$

$$F_i = \sqrt[x_i - 1]{\frac{m_i^{(420)}^{x_i}}{m_i^{(CIF)}}}, \quad (31)$$

$$G_i = \frac{\log\left(\frac{100}{L(VQR)_i^{(420)}} - 1\right)}{\log\left(\frac{F_i}{m_i^{(420)}}\right)}. \quad (32)$$

3 REALIZACE METRIKY CPQD-IES

Realizace metriky CPQD-IES je provedena v programu Matlab. Matlab je programové prostředí a skriptovací programovací jazyk pro vědecko-technické numerické výpočty, modelování, návrhy algoritmů, počítačové simulace, analýzu a prezentaci dat, měření a zpracování signálů.

Předchozí část textu byla věnována popisu metriky CPQD-IES pro hodnocení kvality videosekvencí ve standardním rozlišení, jak je popsáno v doporučení ITU-R BT.1683. Nyní budou v textu popsány vlastnosti videosekvencí ve standardním a vysokém rozlišení, včetně způsobu zpracování programem Matlab. Použité videosekvence jsou soubory ve formátu YUV, který je v dnešní době obecně používán k označení video souborů, které nesou digitální informace $YC_B C_R$.

V další části textu je uveden postup znehodnocení referenční videosekvence kodekovými operacemi, následně je popsáno vytvoření databáze znehodnocených videosekvencí a jaké změny je nutno provést v metrice CPQD-IES, aby byla schopna hodnotit videosekvence ve vysokém rozlišení. Poslední část je věnována popisu uživatelského rozhraní a použitých skriptů

3.1 Videosekvence ve standardním rozlišení

Vzorkování a kódování obrazových signálů $YC_B C_R$ ve standardním rozlišení (SD) popisuje doporučení ITU-R BT.601 [4]. Formát vzorkování 4:4:4 $YC_B C_R$ označuje, že jsou všechny tři obrazové složky vzorkovány o stejné frekvenci 13,5 MHz. Tento formát vzorkování se používá především k přenosu uvnitř televizního studia (mezi studiovým vybavením) a v aplikacích pro zpracování signálu ve vysoké kvalitě. Formát vzorkování 4:2:2 $YC_B C_R$ vyjadřuje, že poměr vzorkovacích kmitočtů je 13,5 MHz pro jasovou složku obrazu Y a po 6,75 MHz pro každou chrominanční složku C_B a C_R . Tento vzorkovací formát je používán jako standard pro digitální rozhraní mezi hlavním vybavením televizního studia a mezinárodní distribucí televizních programů. Například pro přenos digitálních dat mezi kamerovým vybavením a vysílacím systémem.

Systémy pracující s 525 řádky (USA, Japonsko) a opakovacím kmitočtem 60 Hz mají 486 aktivních řádků na snímek. Celkem je na jednom řádku 854 vzorků, z toho je aktivních 720 vzorků jasového signálu. Každý z chrominančních signálů má celkem na jednom řádku 429 vzorků, z toho je aktivních 360 vzorků. Systémy pracující s 625 řádky (evropská norma) a opakovacím kmitočtem 50 Hz mají 576 aktivních řádků na snímek. Celkem je na jednom řádku 864 vzorků a z toho je aktivních 720 vzorků jasového signálu. Chrominanční signály mají celkem na jednom řádku 432 vzorků a z toho je aktivních 360 vzorků.

3.1.1 Načtení videosekvence ve standardním rozlišení

Videosekvence I a O vstupující do CPQD-IES algoritmu jsou ve formátu vzorkování 4:2:2 $YC_B C_R$ a uloženy v souboru formátu YUV. Videosekvence I a O mají na řádku 720

aktivních vzorků jasového signálu a 360 aktivních vzorků pro každý z chrominančních signálů. Vzorky jasového a chrominančních signálů se v datovém souboru periodicky opakují v sekvenci $C_B Y C_R Y$ po jednotlivých vzorcích z každé obrazové složky.

Každý vzorek Y , C_B , a C_R je vyjádřen 8 bity a dekadicky může nabývat hodnot od 0 do 255. Vzorky však nejsou kvantovány v plném rozsahu, neboť krajní úrovně jsou vyhrazeny pro synchronizaci obrazového signálu. Jeden aktivní řádek videosekvence obsahuje 720 vzorků jasového signálu a po 360 vzorcích pro každý chrominanční signál. Celkem má jeden řádek videosekvence 1440 vzorků. Jeden obrazový snímek videosekvence u systému pracujícího s 525 řádky má 699840 aktivních vzorků a u systému s 625 řádky má 829440 aktivních vzorků.

Načtení videosekvence je v Matlabu realizováno skriptem, jehož výstupem je po jedné matici pro každou z obrazových složek Y , C_B , a C_R . Ke správné funkci potřebuje skript znát následující vstupní parametry: název souboru videosekvence, výšku a šířku snímku, počáteční snímek a počet snímků, který se má načíst. Výškou snímku je myšlena hodnota označující počet aktivních řádků v jednom snímku a šířkou snímku počet aktivních vzorků na řádku.

Počet načtených snímků pro odhad kvality videosekvence by měl být alespoň 12, protože 12 snímků je obvykle definována tzv. Group of pictures (GOP) a tím je zajištěno, že sekvence bude obsahovat minimálně jeden intra kódovaný snímek. Tedy snímek, který je kódován nezávisle na ostatních snímcích. Nejedná se o přikázaný počet snímků, ale pouze o doporučený počet zajišťující jistý standard při odhadu kvality videosekvence.

Počet snímků zadaných uživatelem k procesu objektivního měření kvality se musí automaticky zvětšit, protože během přípravy na výpočet a vlastním výpočtu je nutný větší počet snímků. Během objektivního měření je počítán časový atribut T , jenž vyžaduje data získané ze segmentace obrazu předcházejícího snímku, které však při prvním chodu algoritmu neexistují. Aby byly k dispozici potřebné údaje k výpočtu, je nutné provést segmentaci obrazu navíc u jednoho snímku referenční videosekvence O , to znamená zvýšit celkový počet snímků o jedna pro obě vstupní videosekvence O a I . Počet snímků je zvýšen i u testované videosekvence I , protože musí být zajištěna synchronizace snímků mezi oběma videosekvencemi. Necht' je počáteční snímek objektivního měření videosekvence označen f a počet analyzovaných snímků n , pak je kvůli výpočtu časového atributu počáteční snímek roven $f-1$ a počet snímků $n+1$.

Dále ovlivňuje pozici počátečního snímku a počet snímků i korekce časového posunutí, která pracuje v rozsahu od -2 do +2 snímků, proto je nutné mít při jejím výpočtu v paměti o čtyři snímky testované videosekvence I více. Z toho důvodu je automaticky posunuta i pozice počátečního snímku a vezme-li se v úvahu předchozí úprava hodnot, pak bude kvůli korekci časového posunutí počáteční snímek na pozici $f-1-2$ a celkový počet snímků roven $n+1+4$, tentokrát ale pouze u videosekvence I podléhající operaci korekce. Po dokončení korekce časového posunutí jsou už dále používány jen relevantní snímky, jejichž počet se u videosekvence I vrací na hodnotu $n+1$ a přebytečné snímky jsou zahozeny.

Vlastní načtení videosekvence je realizováno pomocí interních funkcí v Matlabu. Příkaz *fopen* slouží k otevření a příkaz *fread* pro binární čtení ze souboru zadané videosekvence. Čtení je prováděno po blocích o 1440 vzorcích, tzn. že se v jednom kroku naráz načte celý řádek jasové složky obrazu a po řádku pro každou z chrominančních

složek obrazu. Následně jsou ze známé sekvence $C_B Y C_R Y$ postupně vyčítány vzorky, které patří vybrané obrazové složce videosekvence.

3.2 Videosekvence ve vysokém rozlišení

Vzorkování a kódování obrazových signálů $YC_B C_R$ ve vysokém rozlišení (HD) popisuje doporučení ITU-R BT.709 [9]. Stejně jako u vzorkování obrazových signálů $YC_B C_R$ ve standardním rozlišení existují pro vzorkování obrazových signálů $YC_B C_R$ ve vysokém rozlišení formáty vzorkování 4:4:4 a 4:2:2 $YC_B C_R$. Vzorkovací kmitočty jsou dány násobky frekvence 2,25 MHz a liší se podle počtu snímků za sekundu a typu prokládání.

Systémy pracující s 50 pulsnímkami za sekundu (50i) mají pro formát 4:2:2 $YC_B C_R$ vzorkovací kmitočet pro jasovou složku obrazu 74,25 MHz a pro chrominanční složky obrazu 37,125 MHz. Systémy pracující s 50 snímky za sekundu (50p) mají vzorkovací kmitočty dvojnásobné, pro jasovou složku obrazu 148,5 MHz a pro chrominanční složky obrazu 74,25 MHz. Aktivních řádků ve snímku je 1080 a v každém řádku je celkem 2304 vzorků, z čehož je aktivních 1920 vzorků. Každý z chrominančních signálů má celkem na jednom řádku 1152 vzorků, z toho je aktivních 960 vzorků.

Systémy pracující s 60 pulsnímkami za sekundu (60i) nebo s 60 snímky za sekundu (60p) mají totožné vzorkovací kmitočty, stejný počet aktivních řádků na snímek a stejný počet aktivních vzorků na řádku jako systémy 50i a 50p. Oproti systémům s opakovacím kmitočtem 50 Hz se však liší celkovým počtem řádků a celkovým počtem vzorků na jednom řádku.

3.2.1 Načtení videosekvence ve vysokém rozlišení

Vstupní videosekvence I a O jsou opět ve formátu vzorkování 4:2:2 $YC_B C_R$ a uloženy v souboru formátu YUV. Videosekvence I a O mají 1080 aktivních řádků, na řádku 1920 aktivních vzorků jasového signálu a 960 aktivních vzorků pro každý z chrominančních signálů. Celkem má jeden řádek videosekvence 3840 aktivních vzorků a jeden celý obrazový snímek videosekvence má 4147200 aktivních vzorků.

Vzorky jasového a chrominančních signálů se periodicky opakují v sekvenci $Y C_B C_R Y$ po celých snímcích jednotlivých obrazových složek. Vlastní načtení je provedeno po blocích o 4147200 vzorcích, kdy se najednou načte celý snímek pro jasovou složku obrazu a po snímku pro každou z chrominančních složek obrazu. Následně je každý snímek obrazové složky uložen do odpovídající matice.

Načtení videosekvence v HD rozlišení je v Matlabu realizováno totožným skriptem jako pro videosekvence v SD rozlišení, pouze se musí změnit vstupní parametry funkce oznamující název videosekvence, výšku a šířku snímku. Počáteční snímek a počet snímků, který se má načíst záleží na uživateli. I zde lze doporučit, aby nastavený počet analyzovaných snímků obsahoval minimálně jeden intra kódovaný snímek z GOP.

3.3 Kodekové operace

Objektivní měření jsou prováděny kromě testované videosekvence I také na cíleně

znehodnocené referenční videosekvenci O . Znehodnocení se provádí kódováním a následným dekódováním referenční videosekvence O pomocí vybraného video kodeku. Kodekové operace jsou provedeny pomocí programu FFmpeg [8].

3.3.1 Kodeky pro videosekvence ve standardním rozlišení

Pro SD videosekvence jsou použity kodeky MPEG-1 a MPEG-2 s pevným kvantizačním krokem o velikosti 16. Během kódování a dekódování jsou vytvářeny dočasné soubory, které jsou po skončení kodekových operací a nahrání potřebných dat smazány. Kódování je prováděno nad všemi snímky referenční videosekvence pro co nejlepší stabilizaci bitového toku. Při kódování videosekvencí z formátu YUV je použit AVI formát a při dekódování videosekvence z formátu AVI je videosekvence převedena zpět do formátu YUV.

Kodek MPEG-2 převádí vstupní videosekvenci z formátu vzorkování 4:2:2 na formát vzorkování 4:2:0 $Y C_B C_R$, proto musí být skript pro načtení videosekvence pozměněn. Úprava spočívá ve změně alokování počtu vzorků pro chrominanční složky obrazu. Podobná úprava je provedena také při použití kodeku MPEG-1 CIF, kdy se mění rozlišení všech obrazových složek videosekvence. Po průchodu kodeky jsou z dekódovaných videosekvencí načteny pouze snímky potřebné k objektivnímu měření, tzn. snímky zadané uživatelem. Při dekódování zpět do souboru formátu YUV, však program FFmpeg duplikuje jeden snímek navíc jak u kodeku MPEG-1 tak kodeku MPEG-2. Tento snímek navíc je zařazen hned na začátek celé videosekvence a není brán v úvahu při načítání videosekvence.

Program FFmpeg je ovládán z příkazové řádky. Kromě názvu vstupního a výstupního souboru videosekvence jsou zadány parametry definující použitý kodek a formáty videosekvencí. Příkaz sloužící k MPEG-2 kódování vypadá následovně:

```
!ffmpeg -s 720x576 -pix_fmt uyvy422 -i test.yuv -qmin 16 -qmax 16  
-vcodec mpeg2video -an -pix_fmt yuv420p -y test420.avi
```

Znak `-s` uvozuje rozlišení kódované videosekvence, `-pix_fmt` definuje vnitřní uspořádání vzorků obrazových složek, `-i` uvozuje název vstupní videosekvence, `-qmin` a `-qmax` vymezují velikost kvantizačního kroku, `-vcodec` definuje použitý video kodek, `-an` odpovídá kódování beze zvuku a `-y` značí název výstupní videosekvence.

Kodekem MPEG-1 by měla být referenční videosekvence kódována do rozlišení CIF, které je 352×288 pixelů, a podvzorkována do formátu 4:2:0 $Y C_B C_R$. Jelikož se při objektivním měření musí shodovat velikosti obrazů absolutního Sobelova rozdílu D_b , jak referenční tak i znehodnocené videosekvence, musí se obraz D_b referenční videosekvence převzorkovat na stejné rozlišení jako má znehodnocená videosekvence po změně na rozlišení CIF. Při rozlišení referenční videosekvence 720×576 by však docházelo k vypouštění některých vzorků a poté by nemusela být zajištěna shoda pixelů na všech pozicích obrazu D_b . Z tohoto důvodu je místo rozlišení CIF použito polovičního rozlišení referenční videosekvence jak v hori-zontálním tak i vertikálním směru. To znamená, že v ukázkovém příkladě bude kodek MPEG-1 kódovat do rozlišení 360×288 pixelů.

3.3.2 Kodeky pro videosekvence ve vysokém rozlišení

Pro videosekvence v HD rozlišení není vhodné používat kodek MPEG-1 CIF, když je zřejmé, že videosekvence mohou dosahovat vysokých bitových toků. V tomto případě by byla redukce kvality až příliš vysoká, a proto byla metrika upravena. Kodek MPEG-1 byl nahrazen kodekem MPEG-4. V případě objektivního měření HD videosekvencí je referenční videosekvence znehodnocena kodeky MPEG-2 a MPEG-4 part 2 s pevným kvantizačním krokem o velikost 16. Během kódování a dekódování jsou vytvářeny dočasné soubory, které jsou po skončení kodekových operací a nahrání potřebných dat smazány. K mezi převodu během kódování a dekódování se používá soubor ve formátu AVI.

Kodeky MPEG-2 a MPEG-4 part 2 opět převádí vstupní videosekvenci z formátu vzorkování 4:2:2 na formát vzorkování 4:2:0 $YC_B C_R$. Kodek MPEG-4 part 2 provádí kódování videosekvence v nezměněném rozlišení jako má vstupní referenční videosekvence. Při kódování kodekem MPEG-4 part 2 je použito dvouprůchodové kódování, které je sice časově náročnější než použití jednorůchodového kódování, ale výhodou je lepší a konstantní kvalita obrazu pro celou videosekvenci. V prvním průchodu je zjištěna náročnost videosekvence a ve druhém průchodu pak kodek přiděluje videosekvenci bitový tok na základě informací získaných při prvním průchodu. Při dekódování kodekem MPEG-4 part 2 nedochází k žádné duplikaci snímků.

Kodekem MPEG-2 jsou videosekvence kódovány do polovičního rozlišení oproti referenční videosekvenci jak v horizontálním tak i vertikálním směru. Pokud je vstupní videosekvence v rozlišení 1920 x 1080, potom má videosekvence po průchodu kodekem rozlišení 960 x 540. Opět je zvoleno jednoduché podvzorkování z důvodu shody velikosti obrazů referenční videosekvence a videosekvence znehodnocené kodekem MPEG-2, která je nutná při objektivním měření.

3.4 Vytvoření databáze znehodnocených videosekvencí

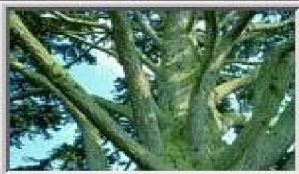






Databáze znehodnocených videosekvencí je klíčem k správnému odhadu kvality videosekvence metrikou CPQD-IES. Jednotlivé videosekvence v databázi musí mít rozličný charakter a vlastnosti, aby byla pokryta co největší šíře různých typů videosekvencí. Jelikož je databáze znehodnocených videosekvencí vytvořena zprůměrováním naměřených dat získaných během vlastního objektivního měření, bylo nutné k jejímu vytvoření použít videosekvence, které byly navzorkovány ve formátu 4:2:2 $YC_B C_R$ a měli k dispozici jak referenční videosekvenci, tak několik různých znehodnocených videosekvencí včetně průměrných subjektivních hodnocení označovaných jako Difference Mean Opinion Scores (DMOS).

3.4.1 Databáze SD

Videosekvence splňující výše popsané požadavky jsou dostupné z ftp serveru skupiny VQEG [10]. Subjektivní hodnocení videosekvencí byly získány metodou DSCQS. VQEG tyto videosekvence použilo při prvním ověřování schopností objektivních modelů odhadovat kvalitu videosekvencí a výsledky testu jsou obsaženy v dokumentu označovaném jako Final report phase 1 [6].

Použité videosekvence mají 625 aktivních řádků a obnovovací frekvenci 50 pulsů za sekundu. Vzorke jasového a chrominancních signálů se v datovém souboru periodicky opakují v sekvenci $C_B Y C_R Y$ po jednotlivých vzorcích každé obrazové složky. U těchto videosekvencí je prvních a posledních 10 snímků nevyužito, což zvětšuje celkový počet snímků videosekvence a zlepšuje stabilitu při kódování. Celá videosekvence pak má 220 snímků, to odpovídá délce videosekvence 8 sekund. Seznam a popis použitých videosekvencí je uveden v tabulce 3.1. Ukázky snímků videosekvencí zobrazených v tabulce 3.1 pocházejí z referenčních videosekvencí [10].

Tab. 3.1: Použité videosekvence ve formátu 625/50i [6]

| Číslo videosekv. | Název videosekvence | Ukázka videosekvence | Charakteristika videosekvence | Zdroj * |
|------------------|---------------------|---|---|---------|
| 1 | Tree |  | Statický obraz | EBU |
| 2 | Barcelona |  | Saturace barev, maskovací efekt | RAI |
| 3 | Harp |  | Saturace barev, zoom, detaily | CCET |
| 4 | Moving graphic |  | Syntetický původ, vertikálně se pohybující text, drobné znaky | RAI |
| 5 | Canoa Valsesia |  | Pohyb a vlnění vody v různých směrech, vysoké detaily | RAI |
| 6 | F1 Car |  | Rychlý pohyb, saturace barev, text | RAI |
| 7 | Fries |  | Dynamický obraz, barvy pleti, rychlý posun kamery | RAI |

| | | | | |
|----|------------------------|---|--|-------|
| 8 | Horizontal scrolling 2 |  | Horizontální rolování textu | RAI |
| 9 | Rugby |  | Sport, dynamický obraz a rychlý pohyb kamery | RAI |
| 10 | Mobile & calendar |  | Horizontální i vertikální pohyb, hodně detailů | CCETT |

*EBU - European Broadcasting Union (Evropa)

*RAI - Radiotelevisione Italiana (Itálie)

*CCETT - Laboratoire Evaluation et acceptabilité de la Qualité des Services (Francie)

Každá z uvedených videosekvencí je tvořena referenční videosekvencí a devíti znehodnocenými videosekvencemi, které reprezentují typické MPEG artefakty a několik dalších efektů jako analogové kódování (PAL/NTSC) či prostorové posunutí. Bitové rychlosti se pro jednotlivé znehodnocené videosekvence pohybují od 3 do 50 Mb/s.

Databáze pro objektivní hodnocení videosekvencí v SD rozlišení byla vytvořena z 10 referenčních videosekvencí (uvedených v tabulce 3.1) a jejich znehodnocených videosekvencí. Protože původně, podle doporučení ITU-R BT.1683, byla databáze utvořena z 12 znehodnocených videosekvencí a nyní jich bylo použito pouze 10, museli být upraveny vztahy pro odhad úrovně kvality v kapitole 2.6. Úpravy se týkají především výpočtů přes všechny videosekvence databáze určované indexem j , který nyní nabývá hodnoty $j = 1, 2, 3, \dots, 10$. Změnil se také počet vybraných nejpodobnějších videosekvencí Ω založených na vzdálenosti $D_{i,j}$. Místo souboru šesti nejpodobnějších videosekvencí se nyní vybírá pouze pět.

Jednotlivé dvojice referenčních a znehodnocených videosekvencí byly podrobeny objektivnímu měření, které je popsáno v kapitole 2.4. Pro každou dvojici videosekvencí byly vypočítány parametry objektivních měření $\{m_1, m_2, \dots, m_9\}$, $\{m_1^{(420)}, m_2^{(420)}, \dots, m_9^{(420)}\}$, $\{m_1^{(CIF)}, m_2^{(CIF)}, \dots, m_9^{(CIF)}\}$, včetně časových T a prostorových S atributů. Objektivní měření byly provedeny pro 24 snímků z každé videosekvence. Naměřené parametry přes těchto 24 snímků byly zprůměrovány a tím byly získány databázové parametry objektivních měření $\{\bar{m}_{1,j}, \bar{m}_{2,j}, \dots, \bar{m}_{9,j}\}$, $\{\bar{m}_{1,j}^{(420)}, \bar{m}_{2,j}^{(420)}, \dots, \bar{m}_{9,j}^{(420)}\}$, $\{\bar{m}_{1,j}^{(CIF)}, \bar{m}_{2,j}^{(CIF)}, \dots, \bar{m}_{9,j}^{(CIF)}\}$, \bar{S}_j a \bar{T}_j . Aby byla databáze kompletní, musí se ještě pro každou z videosekvencí vypočítat průměrné parametry $\bar{F}_{i,j}$ a $\bar{G}_{i,j}$. Parametry $\bar{F}_{i,j}$ a $\bar{G}_{i,j}$ jsou zjištěny na základě minimalizace očekávané střední kvadratické chyby pro každou videosekvenci j zařazenou do databáze. Upravením rovnice (9) je dosaženo vztahu:

$$E \left[\left(\overline{SL}_j - \frac{100}{1 + \left(\frac{\overline{F}_{i,j}}{\overline{m}_{i,j}} \right)^{\overline{G}_{i,j}}} \right)^2 \right], \quad (33)$$

kde \overline{SL}_j je průměrné subjektivní hodnocení kvality videosekvence a $\overline{m}_{i,j}$ je průměrné objektivní měření získané mezi referenční videosekvencí O a znehodnocenou videosekvencí I . Průměrné subjektivní hodnocení kvality videosekvence \overline{SL}_j je pro každou z videosekvencí získáno z dokumentu VQEG o testování objektivních modelů hodnocení kvality videosekvencí [6].

Minimalizace očekávané střední kvadratické chyby byla vypočítána matlabovskou funkcí *fminsearch*, která hledá minimum funkce pro neznámé $\overline{F}_{i,j}$ a $\overline{G}_{i,j}$ při zadaných počátečních podmínkách hledání. Bohužel minimalizovaná funkce má velmi ploché minimum a každá sebemenší změna počátečních podmínek hledání vede k odlišnému výsledku hledaných parametrů. Při výpočtech však bylo zjištěno, že hodnota parametru $\overline{F}_{i,j}$ se obvykle pohybuje v intervalu 0 až 1 a k určení jeho hrubé hodnoty není potřeba měnit počáteční podmínku hledání pro parametr $\overline{G}_{i,j}$. Parametr $\overline{F}_{i,j}$ je určen několika-násobným spuštěním minimalizace funkce s postupným zpřesňováním jeho počáteční podmínky hledání. Minimalizací zjištěný parametr $\overline{F}_{i,j}$ pak slouží jako nové zadání počátečních podmínek pro další kolo minimalizace. Poté je podobným způsobem přistoupeno k hledání parametru $\overline{G}_{i,j}$, jehož hodnoty už nelze jednoznačně definovat v nějakém úzkém intervalu hodnot.

Počáteční podmínky hledání jsou navíc zpřesňovány na základě zpětného výpočtu odhadu úrovně kvality L_i . K tomuto účelu jsou použity rovnice (11) a (12), kde se počítají odhady úrovně kvality $\overline{L}_i^{(420)}$ a $\overline{L}_i^{(CIF)}$ z už známých databázových parametrů $\overline{m}^{(420)}$, $\overline{m}^{(CIF)}$ a z parametrů $\overline{F}_{i,j}$ a $\overline{G}_{i,j}$ získaných během každého spuštění minimalizace. Z hodnot $\overline{L}_i^{(420)}$, $\overline{L}_i^{(CIF)}$, $\overline{m}^{(420)}$ a $\overline{m}^{(CIF)}$ jsou vypočítány parametry F_i a G_i na základě vyřešení soustavy dvou rovnic (28) a (29).

Z vypočtených parametrů F_i a G_i je kontrolováno, zda-li minimalizací nalezené parametry $\overline{F}_{i,j}$ a $\overline{G}_{i,j}$ opravdu produkují odhad kvality na jakou byly trénovány. Kontrola je prováděna výpočtem odhadu úrovně kvality L_i podle rovnice (1) s použitím stejného objektivního parametru $m_{i,j}$ jako při minimalizaci. Pokud hodnoty odhadů úrovně kvality L_i odpovídají trénované hodnotě, která je dána průměrným subjektivním hodnocením kvality videosekvence \overline{SL}_j , pak lze minimalizaci funkce ukončit jako úspěšnou a získané parametry $\overline{F}_{i,j}$ a $\overline{G}_{i,j}$ zapsat do databáze pro příslušnou videosekvenci.

Zbývá určit jaké znehodnocené videosekvence použít při vytváření databáze. Analýzou naměřených dat bylo zjištěno, že čím jsou objektivní parametry $m_{i,j}$ větší tím je produkován vyšší odhad kvality VQR. Proto jsou do databáze vybírány znehodnocené videosekvence na základě porovnání objektivních parametrů $m_{i,j}$

s průměrnými subjektivními hodnoceními kvality \overline{SL}_j . Při výběru znehodnocených videosekvencí do databáze byl kladen důraz na to, aby jejich velikosti objektivních parametrů $m_{i,j}$ a subjektivních hodnocení kvality \overline{SL}_j splňovaly výše uvedenou skutečnost.

3.4.2 Databáze HD




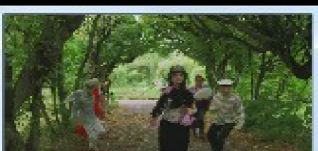



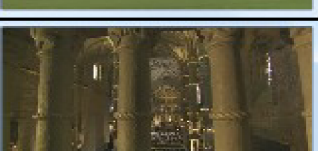

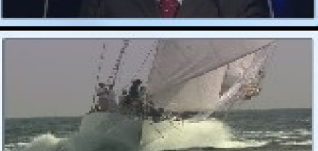
K vytvoření databáze a testování videosekvencí v HD rozlišení se musel najít soubor videosekvencí, které obsahují referenční videosekvence a k nim několik sprážených znehodnocených videosekvencí s uvedeným subjektivním hodnocením. Navíc museli být všechny videosekvence ve formátu vzorkování 4:2:2 $YC_B C_R$. Tyto podmínky splňuje soubor videosekvencí IRCCyN IVC 1080i Database [11], kde je k dispozici celkem 20 referenčních videosekvencí o různém charakteru a zároveň po sedmi znehodnocených videosekvencích ke každé z referenčních videosekvencí. Znehodnocené videosekvence byly vytvořeny video kodekem MPEG-4 part 10 (H.264) s různě nastavenou bitovou rychlostí podle typu videosekvence, aby bylo dosaženo rozdílných kvalit videosekvencí.

Videosekvence mají 1080 aktivních řádků a obnovovací frekvenci 50 pulsů za sekundu. Vzorok jasoového a chrominancních signálů se v datovém souboru periodicky opakují v sekvenci $Y C_B C_R Y$ po celých snímcích každé z obrazových složek. Videosekvence mají rozdílný počet snímků, který odpovídá délce trvání videosekvencí mezi 9 až 12 sekundami.

Subjektivní hodnocení byly získány metodou ACR (Absolute category rating) a metodou SAMVIQ (Subjective assessment methodology for video quality). U metody ACR jsou hodnotitelům předkládány znehodnocené videosekvence v náhodném pořadí bez přítomnosti referenční videosekvence. K hodnocení se používá měřítko kvality, kde se volí jedna z pěti kategorií kvality [12]. U metody SAMVIQ jsou hodnotitelům předkládány současně celé soubory videosekvencí, kdy jeden soubor obsahuje referenční videosekvenci a k ní několik znehodnocených videosekvencí. Hodnotitelé přitom mají libovolný přístup ke všem videosekvencím v souboru, mohou si mezi videosekvencemi přepínat a své hodnocení mohou měnit dokud neohodnotí všechny videosekvence z testovaného souboru. K hodnocení se používá měřítko kvality, kdy je vybrána odpovídající hodnota z rozsahu 0 až 100 [13]. Pro vytvoření databáze znehodnocených modelů byly vybrány subjektivní hodnocení získané metodou SAMVIQ, protože hodnotitelé mohou zvolit subjektivní hodnocení přesněji podle svého uvážení, a protože mají větší dynamický rozsah hodnot.

Databáze pro objektivní hodnocení videosekvencí v HD rozlišení byla vytvořena z 10 referenčních videosekvencí (uvedených v tabulce 3.2) a jejich znehodnocených videosekvencí na základě postupu uvedeného v předchozí kapitole 3.4.1. Všechny znehodnocené videosekvence prošly algoritmem metriky CPqD-IES a naměřené objektivní parametry $\{m_1, m_2, \dots, m_9\}$, $\{m_1^{(420)}, m_2^{(420)}, \dots, m_9^{(420)}\}$, $\{m_1^{(CIF)}, m_2^{(CIF)}, \dots, m_9^{(CIF)}\}$, včetně časových T a prostorových S atributů byly zprůměrovány přes 20 snímků z každé videosekvence. Následně byly zjištěny parametry $\bar{F}_{i,j}$ a $\bar{G}_{i,j}$ na základě minimalizace očekávané střední kvadratické chyby z rovnice (33) pro každou videosekvenci j zařazenou do databáze.

Tab. 3.2: Použité videosekvence ve formátu 1080/50i

| Číslo videosekv. | Název videosekvence | Ukázka videosekvence | Charakteristika videosekvence |
|------------------|---------------------|---|--|
| 1 | Above marathon |  | Dynamické změny obrazu, statické pozadí |
| 2 | Concert |  | Horší světelné podmínky, pomalý pohyb |
| 3 | Credits |  | Vertikální rolování textu |
| 4 | Dance in the woods |  | Pohyb v různých směrech, zoom |
| 5 | Duck fly |  | Pohyb a vlnění vody v různých směrech |
| 6 | Fountain man |  | Stříkající voda, rychlé dynamické změny |
| 7 | Golf |  | Téměř statická kamera, pomalý pohyb objektů |
| 8 | Movie |  | Statický obraz s pomalým nájezdem kamery |
| 9 | Standing |  | Téměř statický obraz, obličej člověka |
| 10 | Voile |  | Pohybující se voda, dynamické změny v obraze |

Výběr videosekvencí do databáze HD probíhal na základě rozdílných vlastností jednotlivých videosekvencí. Především podle analýzy zprůměrovaných objektivních parametrů $\bar{m}^{(420)}$ a $\bar{m}^{(CIF)}$ pro jednotlivé videosekvence. Rozhodujícím faktorem bylo získat do databáze co největší škálu hodnot parametrů $\bar{m}^{(420)}$ a $\bar{m}^{(CIF)}$, jenž reprezentují charakter a vlastnosti každé referenční videosekvence. Do databáze jsou vybírány znehodnocené videosekvence s nižšími subjektivními hodnoceními kvality \overline{SL}_j , protože výsledky minimalizace funkce z rovnice (33) jsou mnohem přesnější než pro vysoké hodnoty subjektivních hodnocení. Za kritickou hodnotu se dá považovat subjektivní hodnocení kvality oceněné známkou 50.

3.5 Vyhodnocení odhadů kvality VQR

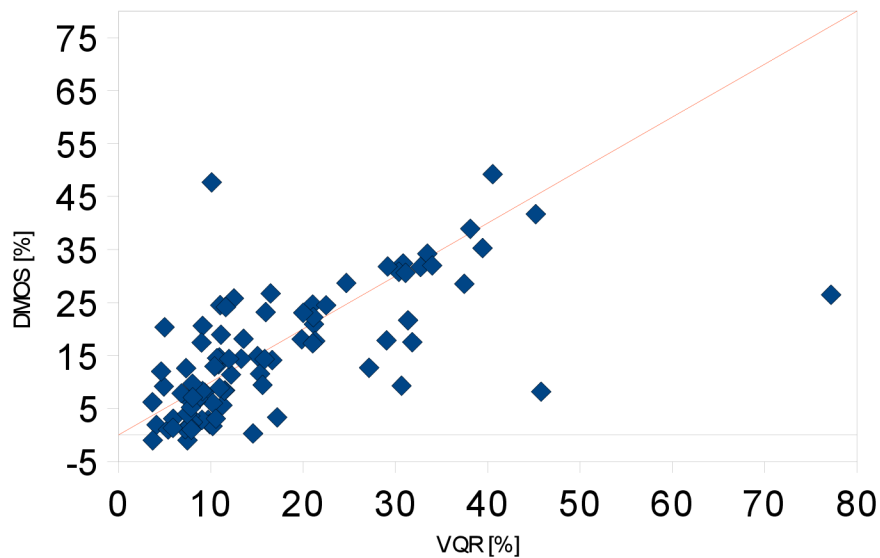
Funkčnost metriky CPqD-IES byla vyzkoušena na souboru videosekvencí, pro které byly známy hodnoty subjektivních hodnocení kvality. Odhady kvality videosekvence VQR byly porovnávány s průměrnými hodnotami subjektivních hodnocení DMOS.

Při prvních testech vytvořených databází bylo zjištěno, že u některých videosekvencí nebyly během segmentace obrazu klasifikovány žádné hranové oblasti. To vedlo k nepřesným výpočtům, které se projeví na špatném odhadu kvality VQR. Zlepšení bylo dosaženo použitím Sobelovy metody hranové detekce. Zároveň s touto změnou bylo odstraněno prahování, kdy byla většina hranových pixelů zamítnuta. Podmínka vertikálně a horizontálně propojených pixelů o nejméně 6 bodech, však zůstala zachována.

3.5.1 Odhad VQR pro videosekvence ve standardním rozlišení

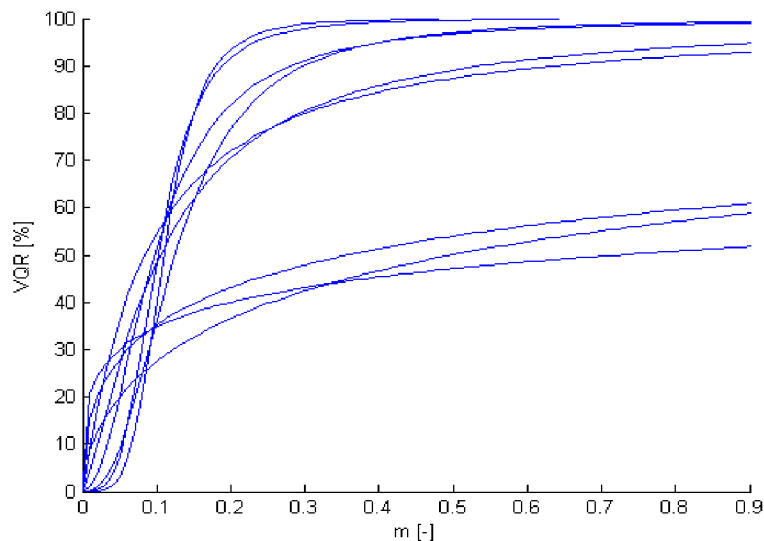
Metrika s nově vytvořenou databází znehodnocených videosekvencí pro videosekvence v SD rozlišení byla ozkoušena na stejných videosekvencích ve formátu 625/50i jaké byly použity pro tvorbu databáze. Jedná se o soubor znehodnocených videosekvencí o různých bitových tocích, které pocházející z ftp serveru skupiny VQEG [10], a jejichž základní popis je uveden v tabulce 3.1.

Vyhodnocení výsledků je zobrazeno do grafu (obr. 3.1), ve kterém jsou na horizontální ose vyneseny hodnoty odhadů kvality VQR a na vertikální ose průměrné hodnoty subjektivních hodnocení DMOS pro danou videosekvenci. Obě osy jsou cejchovány v procentech, kde 0% znamená nejhorší kvalitu videosekvence a 100% nejlepší kvalitu videosekvence. Diagonální přímkou jsou vyznačeny ideální hodnoty, kterým odpovídá situace, kdy jsou hodnoty odhadů VQR naprosto totožné jako průměrné hodnoty subjektivních hodnocení DMOS. Pokud jsou body zobrazené v grafu nad touto přímkou, znamená to, že odhad kvality VQR vypočítaný metrikou pro danou videosekvenci podhodnotil průměrné subjektivní hodnocení DMOS. V případě, že jsou body zobrazené pod touto přímkou je situace opačná a metrika svůj odhad kvality VQR nadhodnotila.



Obr. 3.1: Výsledky metriky pro SD videosekvence

Drobné odchylky od přímky ideálních hodnot se budou v grafu objevovat vždy, protože ne vždy lze subjektivní hodnocení kvality přesně postihnout. Z grafu lze také vyčíst, že několik odhadů kvality VQR je značně nepřesných a ani zdaleka se neblíží subjektivním hodnocením DMOS. To je nejlépe objasněno pokud se blíže prozkoumá vztah mezi subjektivním hodnocením znehodnocené videosekvence a jemu příslušejících parametrů objektivního měření m , které reprezentují rozdíl mezi referenční a znehodnocenou videosekvencí. Tento vzájemný vztah mezi oběma parametry je aproximován křivkou, které se stává výsledným odhadem kvality videosekvencí. Proces aproximace křivkou je definován vztahy v kapitole 2.6 a je na něm postavena celá metrika CPqD-IES. Tvary těchto křivek jsou zobrazeny na obrázku 3.2 při použití databázových parametrů, které patří první videosekvenci z databáze ($j = 1$). Každá z devíti zobrazených křivek odpovídá jedné z obrazových složek Y , C_B nebo C_R s ohledem na klasifikované oblasti během segmentace, tzn. pro každé $i = \{1, 2, \dots, 9\}$.



Obr. 3.2: Vliv parametru m na odhad kvality VQR

Z uvedených křivek vyplývá, že čím je větší velikost parametru objektivního měření m , tím je vyšší vypočítaný odhad kvality VQR. Křivky mají z počátku velkou strmou, ale potom dochází ke zvolnění růstu, v některých případech dokonce k nasycení. Pokud jsou brány v úvahu reálné hodnoty naměřené během objektivních měření, pak je stav nasycení velmi nepravděpodobný. Neboť hodnoty parametrů objektivních měření m se u většiny měřených videosekvencí pohybovaly do hodnoty 0,2, ale velmi záleží o jakou z obrazových složek Y , C_B nebo C_R se jedná. Ohled musí být brán i na klasifikované oblasti během segmentace. Podobné průběhy popisující vliv naměřených parametrů na odhad VQR jako jsou na obr. 3.2 lze sledovat u všech videosekvencí v databázi.

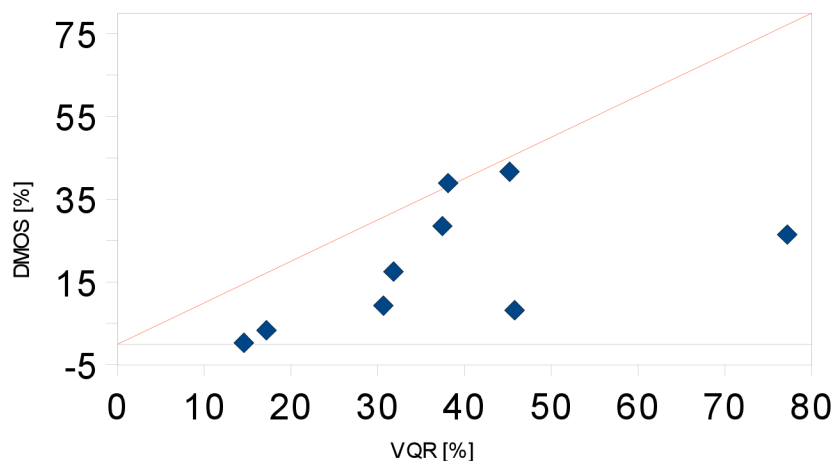
Pokud jsou tedy v grafu na obr. 3.1 nějaké odchylky od přímky ideálních hodnot je to způsobeno tím, že znehodnocená videosekvence má naměřené parametry objektivních měření m , které nabývají relativně velkých hodnot a přitom subjektivní hodnocení kvality přisuzují videosekvenci nízké hodnoty. Nebo nastává přesně opačná situace, kdy jsou hodnoty parametrů objektivních měření m malé a subjektivní hodnocení dosahuje vysokých hodnot.

Velmi dobře je to vidět při odhadu kvality VQR pro videosekvenci číslo jedna, která obsahuje statický obraz koruny stromu. Tabulka 3.3 porovnává subjektivní hodnocení DMOS a parametry $m_{i,j}$ pro znehodnocené videosekvence (ZV) od referenční videosekvence číslo jedna, kde $m_{i,j}$ jsou objektivní parametry změřené pro jasovou složku plošných oblastí, $i = 1, j = 1$.

Tab. 3.3: Porovnání DMOS a parametrů m pro videosekvenci č. 1

| ZV | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------|--------|-------|-------|--------|-------|--------|-------|--------|--------|
| DMOS | 26,477 | 3,333 | 8,176 | 38,909 | 9,301 | 41,683 | 0,307 | 28,544 | 17,544 |
| m | 0,281 | 0,039 | 0,109 | 0,138 | 0,101 | 0,158 | 0,037 | 0,183 | 0,098 |

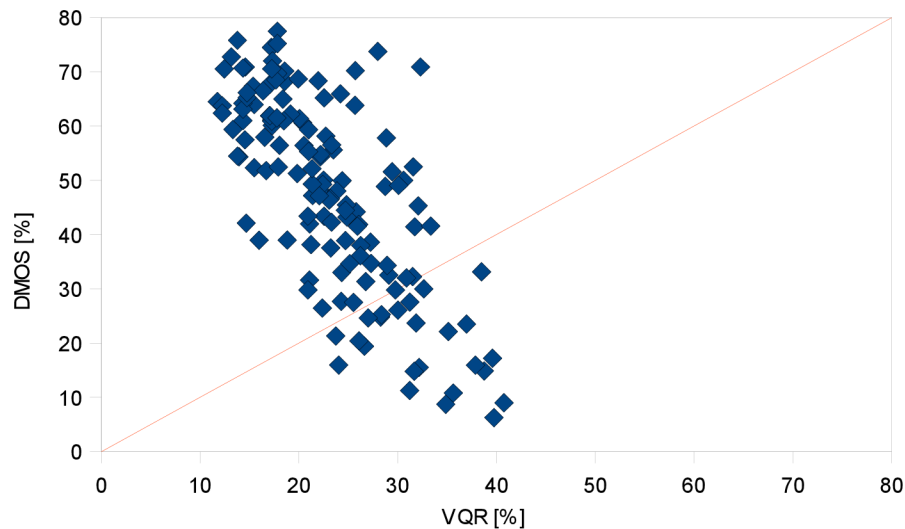
Ačkoli je velikost objektivního parametru m pro první znehodnocenou videosekvenci (ZV 1) nejvyšší, odpovídající subjektivní hodnocení nejvyšší není. Algoritmus metriky se ale řídí velikostí objektivního parametru m , a proto považuje tuto videosekvenci jako nejkvalitnější. Z tohoto důvodu metrika navrhuje poměrně nepřesné odhady kvality VQR pro nehodnocené videosekvence od referenční videosekvence číslo jedna, jak ukazuje obrázek 3.3.



Obr. 3.3: Odhad kvality VQR pro videosekvenci č. 1

3.5.2 Odhad VQR pro videosekvence ve vysokém rozlišení

Metrika s databází znehodnocených videosekvencí pro videosekvence v HD rozlišení byla vyzkoušena na všech 20 videosekvencích, jenž jsou k dispozici z IRCCyN IVC 1080i Database [11]. Kromě již zmíněných 10 videosekvencí použitých při tvorbě databáze je zde dalších 10 videosekvencí, což zvětšuje testovaný vzorek na dvojnásobek oproti SD videosekvencím.



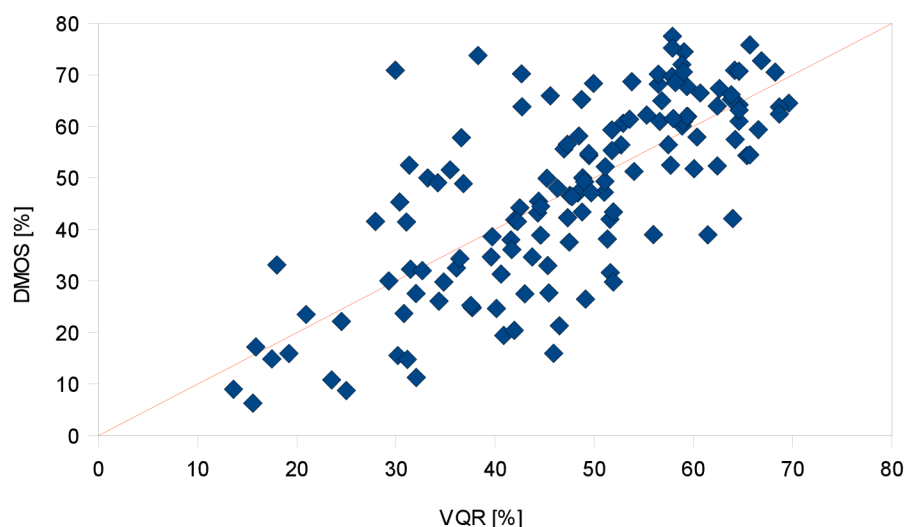
Obr. 3.4: Výsledky metriky pro HD videosekvence

Na obrázku 3.4 jsou prezentovány výsledky metriky pro HD videosekvence a na první pohled je zřejmé, že metrika naprosto selhává. Videosekvence, které měly mít odhad kvality VQR vysoký, jsou metrikou značně podhodnoceny, a videosekvence, u kterých měl být odhad kvality VQR nízký, jsou naopak nadhodnoceny.

K odhalení chyby produkující nepřesný odhad kvality VQR bylo nutno provést analýzu celého algoritmu metriky. Nejprve byla zkontrolována databáze znehodnocených videosekvencí, zda-li plní svůj účel. Vzájemný vztah mezi subjektivním hodnocením znehodnocené videosekvence a jemu příslušejících parametrů objektivního měření m byl u všech videosekvencí v databázi aproximován křivkami podobného tvaru, jako tomu bylo u SD databáze (obr. 3.2), proto se v hledání pokračovalo dál. Poté byla pozornost zaměřena na schopnost metriky vybírat z databáze nejpodobnější videosekvence, což je důležitá vlastnost při odhadu modelu kvality pro jednotlivé znehodnocené videosekvence, ale nebyla zde nalezena žádná nesrovnatelnost s normálním chodem metriky. Proto bylo přistoupeno k analýze videosekvencí vstupujících do metriky a dat získaných při objektivních měřeních. Problémem se ukázaly být parametry objektivních měření m , jejichž hodnoty měly sestupnou tendenci oproti objektivním parametrům naměřených u SD videosekvencí. Takže se zvyšující se kvalitou testovaných videosekvencí (větší bitový tok) byly naměřeny menší hodnoty parametrů objektivních měření m . Jinak řečeno, čím větší subjektivní hodnocení videosekvence od hodnotitelů dostaly, tím menší byly hodnoty parametrů objektivních měření m . To je přesně opačná situace než jaká se vyskytovala při prozkoumání SD videosekvencí. Navíc rozdíly mezi parametry objektivních měření m pro znehodnocené videosekvence s různým bitovým tokem se od sebe moc nelišily a byly si velmi

podobné. To je zajímavé zjištění, které ukazuje na to, že znehodnocené videosekvence vytvořené video kodekem MPEG-4 part 10 potřebují poměrně malý bitový tok pro dostačující kvalitu bez obrazových artefaktů, čímž odstraňují hlavní nevýhody kodeků MPEG-2 nebo MPEG-1. Tím jsou však zhoršeny možnosti úspěšně detekovat zhoršení kvality obrazu pomocí metriky CPqD-IES.

Přesto lze z parametrů objektivních měření m rozeznat jistou schopnost metriky postihnout rozdíly mezi videosekvencemi s různým bitovým tokem. Nejjednodušším řešením je transformovat výsledky odhadů kvality VQR pomocí lineární funkce. Pro body zobrazené v grafu (obr. 3.4) byla určena regresivní přímka ve směrnicovém tvaru, která poté posloužila k transformaci hodnot VQR do nové podoby. Přímka je dána ve směrnicovém tvaru rovnicí $y = -1,93x + 92,27$. Po transformaci jsou výsledky metriky pro HD videosekvence zobrazeny na obrázku 3.5.

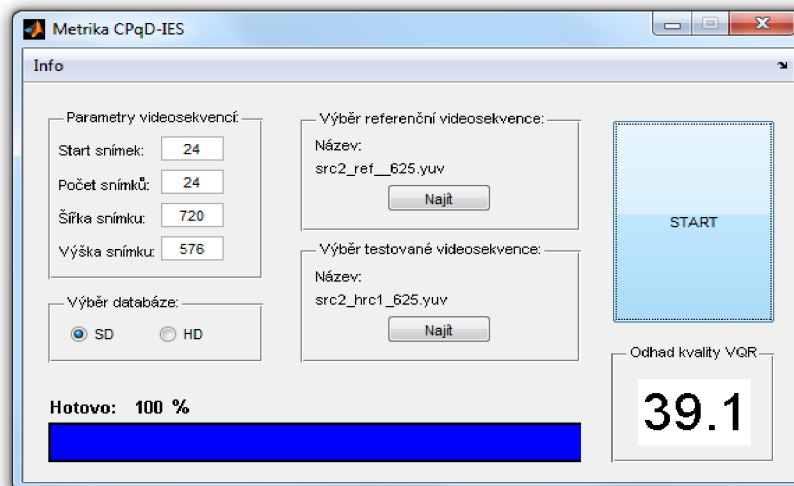


Obr. 3.5: Transformované výsledky metriky pro HD videosekvence

Regresivní přímka transformovaných výsledků má totožnou rovnici přímky jako přímka ideálních hodnot. Transformované odhady kvality VQR už daleko lépe vystihují průměrné subjektivní hodnocení DMOS, i když některé hodnoty jsou stále daleko od ideálních hodnot. Zejména pro videosekvence s nízkým subjektivním hodnocením jsou odhady kvality VQR poměrně nepřesné, ale velká shoda výsledků se dá najít u videosekvencí s vysokým subjektivním hodnocením. Obecně se dá tvrdit, že videosekvence obsahující statický obraz dosahují vyšších odhadnutých výsledků VQR než videosekvence s dynamicky se měnícím obrazem.

3.6 Uživatelský panel

Panel uživatelského prostředí slouží k jednoduchému a intuitivnímu ovládní programu, který realizuje objektivní metriku CPqD-IES. Obsahuje základní ovládací prvky k nastavení parametrů videosekvencí a přehledně zobrazuje dosažené výsledky. Na obrázku 3.6 je zobrazen uživatelský panel po dokončeném hodnocení kvality SD videosekvence.



Obr. 3.6: Vzhled uživatelského prostředí

V bloku parametry videosekvencí se nastavuje rozlišení referenční a testované videosekvence, které je dáno políčky označenými jako šířka a výška snímku. Dále se zde nastavuje počáteční snímek, od kterého se videosekvence analyzují, a celkový počet snímků použitých při objektivním měření. V dalším bloku se vybírá druh databáze znehodnocených videosekvencí, kde je na výběr z možnosti SD a HD. Volba databáze záleží na načtených videosekvencích, které se vybírají tlačítky „najít“. Po stisknutí tlačítka „najít“ se otevře nové okno, ve kterém uživatel najde cestu k hledané videosekvenci. Pokud se hodnotí SD videosekvence, musí uživatel potvrdit výběr SD databáze a naopak v případě HD videosekvencí. Proces objektivního měření a odhadu kvality se spouští tlačítkem „START“. Po jeho zmáčknutí se spustí algoritmus metriky a postup měření je zobrazován v dolní části panelu. Po dosažení stavu hotovo 100% je v okně odhadu kvality VQR zobrazeno číslo, které odpovídá naměřené kvalitě testované videosekvence v procentech. Odhad kvality VQR je zobrazován číslem o jednom desetinném místě.

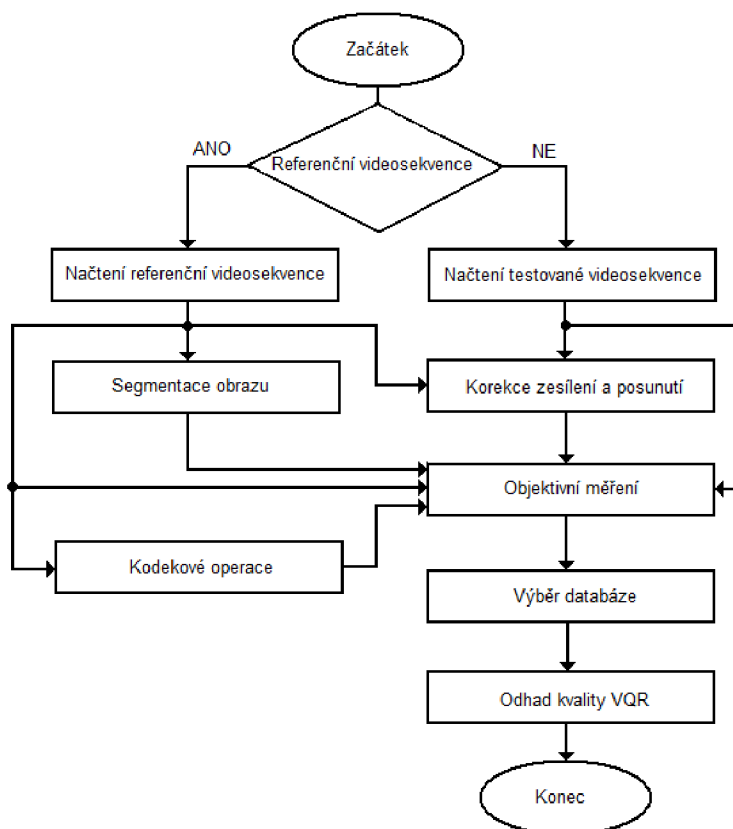
Stav objektivního měření je ukazován číslem v procentech a také ukazatelem průběhu, tzv. progress barem. Ten odpovídá zobrazovaným procentům a je vytvořen skriptem generující obraz pomocí složek RGB. Průběh je rozdělen podle počtu načtených snímků videosekvence a po dílčích částech programu je aktualizován.

3.6.1 Popis použitých skriptů

V této kapitole budou stručně popsány hlavní části programu realizující metriku objektivního měření kvality videosekvencí. Popis se bude zabývat vstupními a výstupními parametry skriptů, v jakém jsou vstupní data formátu a tvaru. Všechny použité skripty byly vytvořeny v programovém prostředí Matlab. Uživatel programu nastavuje pouze parametry vstupních videosekvencí, ostatní parametry si jednotlivé skripty předávají bez jeho účasti.

Vývojový diagram (obr 3.7) zobrazuje všechny důležité části nutné k běhu programu. Na začátku všeho je provedeno načtení videosekvencí do paměti, odkud jsou data dále brána ke zpracování. Referenční videosekvence prochází oproti testované videosekvenci ještě blokem segmentace obrazu a kodekových operací. Obě videosekvence se podílejí na výpočtu korekce zesílení a posunutí snímků, které je

následované objektivním měřením. Do něj navíc vstupují znehodnocené videosekvence po kódování a dekódování referenční videosekvence a binární obrazy získané během segmentace referenční videosekvence. Před vlastním výpočtem odhadu kvality musí být ještě vybrána databáze znehodnocených videosekvencí, podle které se budou videosekvence porovnávat.



Obr. 3.7: Vývojový diagram

Načtení videosekvencí je realizováno třemi různými skripty v závislosti na použitém formátu vzorkování videosekvence. Hlavička však vypadá pro všechny tři skripty stejně:

```
[Y, Cb, Cr] = readYUV_SD(filename, width, height, num, start)
```

Tento skript realizuje načtení videosekvence uložené v datovém formátu YUV, ve kterém nejsou uloženy žádné značky o velikosti snímků, proto musí být tyto vlastnosti zadány ručně. Videosekvenci charakterizuje proměnná *width* označující šířku snímku a proměnná *height* označující výšku snímku. Dále *num* udává počet načtených snímků a *start* počáteční snímek videosekvence, od kterého se provede načtení. Název videosekvence, včetně cesty do cílového adresáře, je řetězec znaků označovaný jako *filename*. Výstupem jsou tři matice, jedna pro každou z obrazových složek Y , C_B a C_R podle zadaných vstupních parametrů. Vzorčky obrazových složek jsou načteny ve formátu *uint8*, jeden vzorek je vyjádřen 8 bity a může nabývat hodnoty od 0 do 255.

Skript s názvem *correction* provádí korekci zesílení, prostorového a časového posunutí snímků mezi vstupními videosekvencemi.

```
[Y_Icor, Cb_Icor, Cr_Icor] = correction(Y_O, Cb_O, Cr_O, Y_I, Cb_I, Cr_I, zac, kon),
```

kde jsou jako vstupní parametry brány matice obrazových složek obou podílejících se videosekvencí. Obrazové složky Y , C_B a C_R referenční videosekvence jsou označovány jako Y_O , Cb_O , Cr_O a testované videosekvence jako Y_I , Cb_I , Cr_I . Vzorky obrazových složek jsou před výpočty převedeny z 8-bitového formátu na formát *double*, tzn. na číslo s plovoucí desetinnou čárkou. Proměnné *zac* a *kon* definují snímky, ze kterých se bude korekce provádět. V současné podobě jsou tyto parametry nastaveny na hodnotu 3 a 7, které zajišťují dostačující počet snímků k úspěšnému provedení korekce. Zároveň reflektují požadavek na výpočet časového posunutí, kdy se uvažuje posunutí snímků dt v rozsahu od -2 do 2. Výstupem korekce je korigovaná testovaná videosekvence, jejíž matice obrazových složek jsou nyní označovány jako Y_{Icor} , Cb_{Icor} , Cr_{Icor} .

Segmentace obrazu je prováděna pouze pro jasovou složku obrazu Y referenční videosekvence Y_O , proto má skript pouze jeden vstup.

```
[PY, EY, TY] = segmentation(Y_O)
```

Výstupem jsou však tři binární matice reprezentující oblasti ploch, hran a textur. Oblast ploch je uložena do matice PY , oblast hran do EY a oblast textur do TY . Jelikož se jedná o segmentaci jasové složky je v názvu písmeno Y a další písmena jsou podle anglických slov „plane“, „edge“ a „texture“.

Znehodnocení referenční videosekvence obstarávají skripty *mpeg2* a *mpeg4*. Pro videosekvence ve standardním rozlišení je použit *mpeg2*, který referenční videosekvenci zpracovává video kodekem MPEG-1 a MPEG-2. Pro videosekvence ve vysokém rozlišení je použit *mpeg4*, který referenční videosekvenci zpracovává video kodekem MPEG-2 a MPEG-4. Vstupní a výstupní proměnné jsou pro oba skripty totožné, proto bude uveden příklad hlavičky jen pro jeden z nich.

```
[Y420, Cb420, Cr420, YCIF, CbCIF, CrCIF] = mpeg2 (width, height, filename, num, start)
```

Protože se zpracovávají referenční videosekvence ze zdrojového souboru YUV, musí být zadané parametry definující vlastnosti videosekvence. Jako je šířka snímku *width*, výška snímku *height* a název referenční videosekvence *filename*. Kodekové operace se provádí nad celou referenční videosekvencí, ale jelikož je nutno dále pracovat se stejným rozsahem snímků videosekvence jako v předchozích bodech, tak je potřeba znát parametry pro načtení znehodnocených videosekvencí. Ty jsou definovány počtem snímků *num* a počátečním snímkem videosekvence *start*. Matice označované jako $Y420$, $Cb420$, $Cr420$ jsou obrazové složky videosekvence znehodnocené video kodekem MPEG-2 a matice $YCIF$, $CbCIF$, $CrCIF$ jsou obrazové složky videosekvence znehodnocené video kodekem MPEG-1.

Znehodnocení kodekovými operacemi je provedeno uvnitř programu objektivního měření *measurement* a z toho důvodu je nutno zadat jako vstupní parametry i název referenční videosekvence *filename* a počáteční snímek načtení *start*. Parametrem *HD* se vybere, jaké kodeky budou použity k znehodnocení referenční videosekvence. Pokud je *HD* rovno jedné, pak jsou použity kodeky MPEG-2 a MPEG-4. Pro *HD* rovno nule se použijí kodeky MPEG-1 a MPEG-2.

```
[m, m420, mCIF, S, T] = measurement(Y_O, Cb_O, Cr_O, PY, EY, TY, Y_Icor, Cb_Icor, Cr_Icor, filename, start, HD)
```

K objektivnímu měření je zapotřebí znát obrazové složky referenční videosekvence Y_O , Cb_O , Cr_O , dále obrazové složky korigované testované videosekvence Y_{Icor} , Cb_{Icor} , Cr_{Icor} a také binární obrazy definující jednotlivé oblasti segmentace PY , EY , TY . Vzorčky obrazových složek jsou před výpočty převedeny z 8-bitového formátu na formát *double*, tzn. na číslo s plovoucí desetinnou čárkou. Výstupem je pak soubor parametrů objektivního měření, kde m označuje objektivní parametry získané z výpočtu mezi referenční a testovanou videosekvencí, $m420$ a $mCIF$ jsou objektivní parametry mezi referenční a její znehodnocenou videosekvencí po průchodu kodekovými operacemi. Písmenem S je označován prostorový atribut a T značí časový atribut získaný během objektivního měření.

Nakonec je proveden výpočet odhadu kvality VQR testované videosekvence pomocí skriptu *VQR_estimation*.

```
[VQR_final] = VQR_estimation(m, m420, mCIF, S, T, SD, HD)
```

K výpočtu jsou potřeba parametry získané během objektivního měření m , $m420$, $mCIF$, S , T a také parametry určující, která databáze znehodnocených videosekvencí bude použita SD a HD . Databázové parametry jsou uloženy v souborech *database_SD.mat* a *database_HD.mat*. Pokud je parametr HD roven jedné, pak je ještě navíc spuštěna lineární transformace získaných hodnot odhadů VQR, která je nutná při hodnocení videosekvence ve vysokém rozlišení.

ZÁVĚR

Tato diplomová práce pojednává o objektivních technikách hodnocení kvality videa ve standardním a vysokém rozlišení digitální televize. Pozornost byla zaměřena na metriku CPqD-IES, která je založena na segmentaci obrazu do oblastí hran, ploch a textur. V práci je podrobně popsán algoritmus vyhodnocování kvality videosekvencí a vytvoření databáze znehodnocených videosekvencí. Daná problematika byla řešena v programovém prostředí MATLAB, které je uzpůsobeno k práci s různými druhy signálů a obsahuje vhodné nástroje pro zpracování videosekvencí.

Stěžejní prací bylo vytvořit databáze znehodnocených videosekvencí, neboť originální databáze nebyla k dispozici. Databáze jsou vytvořeny na základě naměřených objektivních parametrů z několika různých videosekvencí a jejich subjektivních hodnocení kvality. Při tomto procesu musí být minimalizována funkce, která nemá výrazné minimum a je proto obtížné dosáhnout požadovaných hodnot. Situace byla navíc ztížena nedostatkem dostupných videosekvencí s požadovanými parametry a údaji, což se mohlo negativně projevit na skladbě databáze znehodnocených videosekvencí.

V práci použité videosekvence ve standardním rozlišení byly kódovány video kodekem MPEG-2 a videosekvence ve vysokém rozlišení byly kódovány video kodekem MPEG-4. Vytvořené databáze byly vyzkoušeny na souboru videosekvencí, kde se ukázalo, že metrika je uzpůsobena k hodnocení videosekvencí zpracovaných video kodekem MPEG-2. Algoritmus metriky je postaven tak, aby byl schopen rozpoznat typické obrazové artefakty způsobené MPEG-2 kódováním videosekvencí. Proto metrika selhala při pokusu odhadovat kvalitu videosekvencí zpracovaných video kodekem MPEG-4, který používá vyhlazovací filtr k odstranění blokových artefaktů, čímž dosahuje lepší obrazové kvality i při nízkém bitovém toku. Naměřené objektivní parametry přesto dokázaly postihnout jisté rozdíly mezi různými stupni znehodnocení pro jednu referenční videosekvenci. Na základě této skutečnosti bylo možno upravit výsledky odhadů kvality pomocí lineární transformace.

K ohodnocení funkčnosti metriky byl vypočítán Pearsonův koeficient korelace, který udává vztah mezi subjektivními hodnoceními a odhadnutými výsledky metriky. Čím je Pearsonův koeficient korelace bližší jedné tím je silnější lineární vztah mezi hodnoceními. Pro videosekvence ve standardním rozlišení vyšel koeficient korelace roven 0,658 a pro videosekvence ve vysokém rozlišení dokonce 0,741. To znamená, že videosekvence ve vysokém rozlišení byly nakonec odhadovány lépe než ve standardním a to i přes potíže způsobené jejich zpracováním kodekem MPEG-4. Ale ani výsledky odhadů pro videosekvence ve standardním rozlišení nedopadly vyloženě špatně, neboť jejich koeficient korelace byl velmi podobný tomu, který byl pro tuto metriku a stejný typ videosekvencí uváděn ve výsledném hodnocení objektivních metrik skupinou VQEG [6]. V tomto dokumentu je koeficient korelace uváděn 0,672. Dále zde tvůrci metriky uvádějí, že metrika byla navržena a optimalizována k hodnocení videosekvencí o nízké kvalitě, zejména pro přenosové aplikace. Protože v práci byly použity videosekvence dosahující i velmi vysokých bitových toků, lze považovat výsledky implementované metriky za dobré.

Důležité je však zmínit, že odhad kvality testovaných videosekvencí se může značně lišit podle právě vybraných snímků videosekvence. Proto se může stát, že několik měření provedených na stejné videosekvenci, ale pro různé snímky, bude mít různé odhady kvality.

Dalším krokem kam by se práce mohla ubírat, je rozšíření databáze o jiné typy videosekvencí a provedení optimalizace za účelem zpřesnění výsledků odhadů kvality. Případně provést specializaci na určitý typ videosekvencí a video kodeků. Především se ale zaměřit na video kodeky používané v současnosti a analyzovat jejich typické znaky projevující se při kódování videosekvence nízkým bitovým tokem. Na základě těchto dat pak upravit stávající metriku CPqD-IES, nebo vytvořit zcela novou metriku založenou na odlišném postupu.

LITERATURA

- [1] Rec. ITU-R BT.1683. *Objective Perceptual Video Quality Measurement Techniques for Standard Definition Digital Broadcast Television in the Presence of a Full Reference*. Geneva: ITU, 2004.
- [2] Rec. ITU-R BT.500. *Methodology for the subjective assessment of the quality of television pictures*. Geneva: ITU, 2009.
- [3] GONZALEZ, R. C. and WOODS, R. E. *Digital Image Processing*. , New Jersey, Prentice-Hall, 2002. ISBN: 0-201-18075-8.
- [4] Rec. ITU-R BT.601-4. *Encoding parametrs of digital television for studios*. Geneva: ITU, 1994.
- [5] Rec. ITU-T J.144. *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*. Geneva: ITU, 2001.
- [6] ITU-T Study Group 9 Contribution 80. *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase I*. VQEG, 2000 [cit. 16. prosince 2011]. Dostupné na www.its.bldrdoc.gov/vqeg/projects/frtv_phaseI/COM-80E_final_report.doc
- [7] CORRIVEAU, P., WEBSTER, A. *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. VQEG, 2003 [cit. 16. prosince 2011]. Dostupné na www.its.bldrdoc.gov/vqeg/projects/frtv_phaseII/downloads/VQEGII_Final_Report.pdf
- [8] FFmpeg, version N-38938-ge01f478 built on Mar 19 2012 23:16:52, 2012 [cit. 30. března 2012]. Dostupné na www.ffmpeg.org
- [9] Rec. ITU-R BT.709. *Parameter values for the HDTV standards for production and international programme exchange*. Geneva: ITU, 2002.
- [10] ITU-T Study Group 9 Contribution 80. *Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase I*. VQEG, 2000. Dostupné na: ftp://vqeg.its.bldrdoc.gov/SDTV/VQEG_PhaseI/
- [11] STÉPHANE PÉCHARD, ROMUALD PÉPION, PATRICK LE CALLET. *Suitable methodology in subjective video quality assessment: a resolution dependent paradigm, International Workshop on Image Media Quality and its Applications, IMQA2008, Kyoto : Japan, 2008*. [cit. 10. května 2012]. Dostupné na: ftp://ftp.ivec.polytech.univ-nantes.fr/IRCCyN_IVC_1080i_Database/
- [12] Rec. ITU-T P.910. *Subjective video quality assessment methods for multimedia applications*. Geneva: ITU, 1999.
- [13] QUAN HUYNH-THU, MATTHEW BROTHERTON, DAVID HANDS, KJELL BRUNSTROM, MOHAMMED GHANBARI. *Examination of the SAMVIQ methodology for subjective assessment of multimedia quality*. UK s Sweden: 2007. [cit. 10. května 2012].

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

| | |
|-------|--|
| ITU | International Telecommunication Union |
| VQEG | Video Quality Experts Group |
| PSNR | Špičkový poměr signál šum |
| DSCQS | The double-stimulus continuous quality-scale |
| VQR | Video quality rating |
| CIF | Common intermediate format |
| CD420 | Kodér/dekodér MPEG-2 4:2:0 |
| CDCIF | Kodér/dekodér MPEG-1 CIF |
| SD | Standard-definition |
| HD | High-definition |