



UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



RNDr. Petra Heribanová

Autoreferát dizertačnej práce

METÓDY HODNOTENIA KVALITY VIDEA

na získanie akademického titulu philosophiae doctor
v odbore doktorandského štúdia:9.2.7 Geometria a topológia

Bratislava 2014

Dizertačná práca bola vypracovaná v dennej forme doktorandského štúdia na Katedre algebry, geometrie a didaktiky matematiky, Fakulty matematiky, fyziky a informatiky Univerzity Komenského v Bratislave.

Predkladateľ: **RNDr. Petra Heribanová**
Katedra algebry, geometrie a didaktiky matematiky
FMFI UK v Bratislave
Mlynská dolina, 842 48 Bratislava
petra.heribanova@gmail.com

Školiteľ: **prof. Ing. Polec Jaroslav, PhD.**
Fakulta elektrotechniky a informatiky
Ilkovičova 3
812 19 Bratislava
jaroslav.polec@stuba.sk

Oponenti:
.....
.....
.....
.....
.....

Obhajoba dizertačnej práce sa koná o h
pred komisiou pre obhajobu dizertačnej práce v odbore doktorandského štúdia vymenovanou
predsedom odborovej komisie

9.2.7 Geometria a topológia

na Fakulte matematiky, fyziky a informatiky Univerzity Komenského, Mlynská dolina, 842
48 Bratislava

Predseda odborovej komisie:

prof. RNDr. Július Korbaš, CSc.
Katedra algebry, geometrie a didaktiky matematiky
FMFI UK, Mlynská dolina, 842 48 Bratislava

OBSAH

1. ÚVOD	2
2. ZHODNOTENIE SÚČASNÉHO STAVU PROBLEMATIKY	3
3. CIELE DIZERAČNEJ PRÁCE	4
4. ZVOLENÉ METÓDY RIEŠENIA	5
5. EXPERIMENT A DOSIAHNUTÉ VÝSLEDKY	7
6. KONKRÉTNE ZÁVERY PRE ĎALŠÍ ROZVOJ VEDY	12
7. REFERENCIE	13
8. PUBLIKÁCIE AUTORKY A PROJEKTY	15
9. RESUMÉ	18

1 ÚVOD

V súčasnosti sa pri určovaní kvality preneseného videa pristupuje k videu bez ohľadu na jeho obsah, predpokladá sa, že obsahuje fotorealistické obrazy a kritériom kvality je vo všeobecnosti to, ako sa video páči pozorovateľovi vizuálne.

Existujú však videá, ktoré svojim obsahom nezodpovedajú takejto filozofii. Reprezentatívnym príkladom sú rôzne realizácie metód augmentatívnej a alternatívnej komunikácie (AAK) na diaľku [Ang'97, Bat'02, Tar'08]. Alternatívna a augmentatívna komunikácia má za cieľ umožniť osobám so závažnými poruchami komunikačného procesu aktívne dorozumievanie a zapojenie sa do života spoločnosti. Pri takejto komunikácii ide viac o zrozumiteľnosť než o ľúbivosť videa.

Hlavný rozdiel medzi pojmami kvalita a zrozumiteľnosť je v tom, že pod pojmom kvalita rozumieme, „ako“ divák vníma video signál (ako vyzerá spracovaný signál). Pod pojmom zrozumiteľnosť rozumieme „čo“ divák vidí (či to, čo vidí, má aj informačný zmysel). Z výsledkov subjektívnych hodnotení vyplýva, že neprirodzený obraz, tj. so sýtymi farbami, sa páči viac. A tiež, že pri video ukážke s kvalitným zvukom môžeme poľaviť do určitej miery z kvality obrazu, keďže zvuk znižuje schopnosť používateľa rozlíšiť znehodnotenia.

Kritériá hodnotenia kvality sú úzko prepojené s kódovaním a kompresiou, pretože sú prostriedkom pre určenie hranice, nakoľko je možné znižovať veľkosť dátového toku videa. Tieto kritériá však nezohľadňujú potrebu zrozumiteľnosti posunkov, znakov alebo gest. Neexistuje žiadne odporúčanie ITU (International Telecommunication Union) pre hodnotenie kvality a zrozumiteľnosti videa obsahujúceho alternatívne a augmentatívne spôsoby komunikácie.

V práci sa zameriavame na hodnotenie kvality videa obsahujúceho prstové znaky Slovenského posunkového jazyka (SPJ), keďže nepočujúcich kvalitný zvuk neovplyvní a v prvom rade je pre nich dôležité, či je video ukážka (video konferencia) zrozumiteľná, a nie či je pekná, kvalitná.

2 ZHODNOTENIE SÚČASNÉHO STAVU PROBLEMATIKY

Video môže prejsť mnohými fázami spracovania, kým sa dostane k pozorovateľovi, pričom každý stupeň spracovania môže zaviesť rušenia, ktoré by mohli znížiť kvalitu konečného videa. Kvalita videa je teda závislá od metódy zachytávania a ukladania obrazu. Meranie kvality obrazu je dôležité pri video kompresii, keďže tam vždy ide o kompromis medzi mierou kompresie a kvality, a teda rôzne formáty majú rôzny pomer kvality a objemu.

Podľa [MSU'08] MSU Video Grup z Lomonosovovej Moskovskej štátnej univerzity je *kvalita videa* charakteristika videa (prenášaného / spracovaného systémom) definovaná ako formálna alebo neformálna miera vnímanej degradácie obrazu (typicky, oproti pôvodnému videu). Objektívne metódy na hodnotenie kvality zhŕňajú pod pojem metriky.

Pojem metrika spojený s hodnotením kvality videa sa odlišuje od geometrickej definície pojmu metrika a využíva sa v prenesenom význame, teda METRIKA ako MIERA, METÓDA. *Metrika* (hodnotenie kvality obrazu a videa) je akýkoľvek algoritmus použitý na modelovanie obrazu, po ktorom nasleduje cenová funkcia a výsledkom je číselná hodnota, teda na základe nej je možné modelovať poškodenie v obraze.

V zásade sa používajú na hodnotenie kvality videa dve skupiny metód:

- Subjektívne metódy merajú reakcie pozorovateľov sledujúcich testovaný systém. Tieto metódy sú veľmi náročné na čas a na realizáciu. Medzinárodné odporúčania ITUR BT.500-12 (2009) a ITUT P.910 (2008) špecifikujú, ako vykonávať rôzne typy subjektívnych testov na hodnotenie kvality videa [ITUR'09, ITUT'08].
- Objektívne metódy sú automatické metódy bez účasti pozorovateľov. Zohľadňujú vnútornú štruktúru obrazovej informácie a rôzne vlastnosti ľudského vizuálneho systému, prípadne sa realizujú pomocou niektorej vzdialenostnej metriky. [Avc'02, ITUR'09, Wink'05, Kawa'08, Ries'08].

Snahou je nájsť pre príslušnú aplikáciu metódu s najväčšou koreláciou s výsledkami subjektívnych metód. Relevantné korelácie pre hodnotenie úspešnosti objektívneho modelu sú Personov korelačný koeficient pre presnosť a Spermanov pre monotónnosť.

Každá videosekvencia je kombináciou „záberov“ získaných z určitého pohľadu, každý záber sa skladá z niekoľkých segmentov, z ktorých každý zodpovedá sade podobných snímokov.

Extrakcia KF všeobecne zahŕňa výber jednej snímky z každého záberu segmentu, tzv. „cluster“, ktorý reprezentuje video [Omi'11].

Keďže posunkový jazyk využíva na dosiahnutie komunikačného cieľa statické a dynamické gestá/znaky. Preto môžeme predpokladať, že pri videách s prstovou abecedou Slovenského posunkového jazyka je možné extrahovať sadu kľúčových snímok, ktoré predstavujú znaky prstovej abecedy a teda zachytávajú obsah videa.

3 CIELE DIZERAČNEJ PRÁCE

Na základe zhodnotenia súčasného stavu som si stanovila nasledovné ciele:

1. Určiť kritériá na potrebnú kvalitu video signálov degradovaných kódovaním s rôznou bitovou rýchlosťou tak, aby bola prstová abeceda pre nepočujúcich zrozumiteľná, t.j. určiť podmienky pre zrozumiteľnosť, kde je potrebné presne pochopiť/ zachytiť jednotlivé znaky, pričom sa nedajú odvodiť z kontextu. Vytvoriť chýbajúcu metodiku pre takéto hodnotenie.
2. Na základe metódy úplnej referencie nájsť metriku, ktorá by dobre korelovala so zrozumiteľnosťou.
3. Určiť podmienky pre automatické objektívne hodnotenie kvality a zrozumiteľnosti videa obsahujúceho prstovú abecedu.
4. Využiť skúsenosti z publikovaných metód na hľadanie kľúčových snímok a významných oblastí vo videu a určiť oblasti záujmu pre video s prstovou abecedou pomocou segmentácie a podľa významnosti ich váhovať.
5. Návrh metódy merania kvality (zrozumiteľnosti) videa obsahujúceho prstové znaky SPJ na základe nájdenia kľúčových snímok, ktorá bude zohľadňovať subjektívne hodnotenie zrozumiteľnosti a význam regiónov záujmu na základe segmentácie a sledovania dominantnej ruky vo videu.
6. Implementácia navrhutej objektívnej metódy merania, experimentálne overenie dosiahnutých výsledkov na statických obrazoch a videách a porovnanie výsledkov navrhutej metódy s výsledkami iných objektívnych a subjektívnych metód.

4 ZVOLENÉ METÓDY RIEŠENIA

Prstová abeceda

Prstové znaky sú vizuálne symboly, ktoré neobsahujú informáciu o slove, ani o emóciách, ale symbolizujú fonémy (hlásky), alebo grafémy (písmená), číslice. Prstová abeceda je najznámejším systémom prstových znakov. Ide o systém konfigurácií prstov rúk a pohybov, ktoré reprezentujú abecedné znaky, pričom počet znakov súvisí s počtom hlások daného jazyka. Svetovo rozšírené a používané sú dva systémy prstovej abecedy: *jednoručná prstová abeceda (daktyl)* a *dvojručná prstová abeceda*, pričom je veľmi časté miešanie znakov z oboch abecied a využívanie len základných znakov [Hef'09, Tar'02, Tar'08].

Zrozumiteľnosť (akustická)

Pojem zrozumiteľnosť sa používa pri hodnotení reči v akustike alebo pri syntetickej reči. Zrozumiteľnosť je jedným z hlavných faktorov na určenie kvality prenosových kanálov a definuje sa ako percentuálny pomer porozumeného k vyslovenému, respektíve percento správne rozpoznávaných slov (foném, viet). So zvyšujúcou sa zrozumiteľnosťou reči sa znižuje námaha pri posúvaní. *Prahová hodnota zrozumiteľnosti* je v akustike definovaná ako bod, pri ktorom poslucháč síce ešte počuje, ale už nerozumie. *Modifikáciou kritérií* hodnotenia zrozumiteľnosti reči, používaných v akustike (hlásková, vetná zrozumiteľnosť) môžeme skúmať zrozumiteľnosť pri videách [Mak'95].

Rozoznatel'nosť

Metriky používané na vyhodnotenie úspešnosti detekcie strihov [Bro'06] a tiež extrakcie kľúčových snímok sú *Citlivosť / Rozoznanie (Recall) R* (hovorí o počte stratených odpovedí), *Presnosť (Precision) P* (hodnota klesá so zvyšujúcimi sa počtom falošných odpovedí) a *F1 Miera* (je kombinácia P a R, globálny pohľad na správnosť skúmaných odpovedí)[Olson'08]. Modifikáciou kritérií rozoznatel'nosti sme skúmali úspešnosť výberu kľúčových snímok (KF).

Návrh metódy 1: Pomocou masky so štandardnými metrikami

Metóda 1: *Pomocou masky so štandardnými metrikami* je založená na predpoklade, že časť obrazu predstavujúca pozadie je irelevantná a zrozumiteľnosť je počítaná iba v ROI. Na

výpočet kvality na ROI boli vytvorené dva typy masky: statická elipsovité a dynamická maska.

Návrh metódy 2 : Pomocou kľúčových snímok

Metóda 2: *Pomocou kľúčových snímok* je založená na nájdení kľúčových snímok (KF) a ich optimálnom využití pri hodnotení zrozumiteľnosti videa. Séria nájdených KF nám reprezentuje video. Využitím bežných objektívnych metód na hodnotenie kvality videa len na nájdené kľúčové snímky môžeme rýchlo a efektívne určiť kvalitu pre celé video, kdeže:

1. každá KF reprezentuje sériu snímok so „stabilným“ znakom
2. kvalita KF by mala reprezentovať kvalitu všetkých snímok zo série pre daný znak

V prvom kroku je potrebné rýchlo, efektívne a s maximálnou úspešnosťou nájsť kľúčové snímky reprezentujúce jednotlivé znaky vo videu. V druhom kroku určiť kvalitu použitím objektívnych metód na hodnotenie kvality videa. Navrhnuté, špecifikované a implementované boli dve metódy na extrakciu KF s následným hodnotením zrozumiteľnosti len na detekovaných KF:

- ***So sledovaním a prahovaním so štandardnými metrikami a priemerovaním adaptívnym prahom*** je založená na stanovení polohy rúk (obdĺžnikového ROI), segmentácii a využití štandardných metrik na hodnotenie kvality videa. Výpočet metriky sa vykonáva iba v oblastiach záujmu a len v časti "viditeľného okna". Pre jednoduchosť sme zvolili negatívny alfa obraz, teda čierna predstavuje body, v ktorých je metrika počítaná. Princíp posuvného okna pre nájdenie lokálneho extrému a adaptívne prah založený na lokálnom priemerovaní sa používa pre výber kľúčových snímok. Účinnosť metódy je porovnaná s aplikovaním metriky na všetky snímky [Pol'13].
- ***So sledovaním a prahovaním s metrikou DFT a Dugadovym prahom*** je založená na sledovaní ruky pomocou Centertrackeru [Kon'11], porovnaní dvoch po sebe idúcich segmentovaných regiónov záujmu obsahujúcich ruku so znakom a nájdení a výberom KF prahovaním podľa modifikovaného Dugadovho modelu [Dug'98], teda nájdením vybraných snímok z množiny lokálnych extrémov na vhodne zvolenom počte sledovaných snímok. Porovnanie sa robí pomocou spektrálnej metriky DFT a jej modifikáciou pomocou DCT s predpokladom, že v ustálenom stave dochádza medzi snímkami jedného znaku k minimálnej zmene fázy [Her'14].

5 EXPERIMENT A DOSIAHNUTÉ VÝSLEDKY

Videá použité pri testovaní boli špeciálne vytvorené na tento účel. Každá video ukážka obsahuje sedem rôznych logatémov so 41 znakmi slovenskej jednotučnej abecedy. Dĺžka videa je približne minútu. Základné video bolo vytvorené vo formáte " 4 - CIF " (704x576) s 25 snímkami za sekundu.

Pomocou programu vVirtualDub sme vytvorili video ukážky v ďalších formátoch a to " CIF " (352x288) a " QCIF " (176x144). Následne boli videá kódované kodekom H.264 v rôznych bitových rýchlostiach (QP = 30 , 40 , 50 , ktorá zodpovedá dátovému toku od 390 kbit / s na 4,5 kbit / s v uvedenom poradí) .

Pracujeme na základe objektívnej metódy úplnej referencie (FR metódy). Ako originálne videá používame prvotné videá vo formáte 4 CIF (704x576). Bilineárnou interpoláciou sme formáty CIF a QCIF previedli na spoločný formát 4CIF.

Subjektívne hodnotenie

Subjektívne testy na hodnotenie zrozumiteľnosti známkovaním sú pre nás bezpredmetné a výsledky z nich nemôžeme brať ako správne hodnotenie zrozumiteľnosti.

Výsledky objektívneho hodnotenia znakovej zrozumiteľnosti boli realizované subjektívnym hodnotením metódy ACR na skupine nepočujúcich respondentov. Zameraním sa na kritérium hláskovej zrozumiteľnosti, pomocou umelých jednoslabičných slov bez významu, tzv. logatémov, vylúčime náchylnosť človeka opravovať nesprávne prijaté slabiky, či slová podľa obsahového významu.

Z výsledkov vyplynulo, že klesajúcou zrozumiteľnosťou sa zvýšil počet výmen spoluhlások reprezentovaných podobným znakom, predovšetkým 'a' a 's', 'o' a 'f', a tiež dochádzalo k častejšiemu vynechávaniu alebo pridávaniu hlások, pretože respondentovi nebol jasný prechod medzi znakmi. Svetelné podmienky, nastavenie kamery a farba pozadia majú veľký vplyv na celková zrozumiteľnosť, rovnako ako na viditeľnosti a schopnosť interpretovať znaky.

Objektívne hodnotenie

Na základe výsledkov objektívneho hodnotenia zrozumiteľnosti hľadáme metódu, ktorá by najlepšie korelovala so zrozumiteľnosťou, a teda by mohla predstavovať metódu na automatické hodnotenie zrozumiteľnosti videa s posunkovou rečou a prstovou abecedou.

Na objektívne vyhodnotenie kvality pomocou vybraných metrik bol použitý akceptovaný softvér MSU Video Quality Measurement tool od MSU Graphics & Media Lab (Video Group) [MSU'08]. Podľa odporúčania ITU-T pre hodnotenie úspešnosti objektívnej metriky vzhľadom na subjektívne hodnotenie Pearsonovým korelačným koeficientom zisťujeme hodnotu korelácie danej metriky a zrozumiteľnosti.

Podľa návrhu metódy 1 : *Pomocou masky so štandardnými metrikami* sme hodnotili kvalitu zrozumiteľnosti bežne používanými metrikami pre celé video s rovnakou váhou pixelov, so statickou elipsovitou maskou a s dynamickou maskou. Dynamická maska bola vytvorená aplikovaním prahu s hodnotou okolo 130 a následne statickou elipsovitou maskou z predchádzajúceho testovania, čím sa vytvoril ROI pre každú snímku videa. V experimente boli použité následne nastavenia: ROI = 1, pozadie BG = 0 [Her'12a, Her'12b]. Výsledky obsahuje Tab. 5.2.

metódy metriky	Bez trackingu					Tracking
	DFT $\lambda=0$		DFT $\lambda=1$	DCT		DFT $\lambda=0$
WINDOW	31	37	25	31	37	25
CS	40	40	37	40	39	39
DS	0	0	3	0	0	2
MS	1	1	1	1	2	0
CP	8	7	8	8	7	7
DP	0	0	0	0	0	1
MP	0	0	0	0	0	0
FP	0	0	0	0	0	0
Recall (R)	97,96	97,92	97,96	97,96	95,83	100
Precision (P)	100	100	96,970	100	100	98
F1	98,969	98,947	97,462	98,969	97,872	98,990

metódy metriky	Tracking		Celé video	S maskou		
	DCT		SSIM	M-VQM	M-SSIM	M-MSE
WINDOW	25	31	21	21	21	19
CS	39	38	35	37	37	35
DS	2	2	6	3	4	6
MS	0	1	0	1	0	0
CP	6	6	7	6	5	7
DP	2	2	1	2	3	1
MP	0	0	0	0	0	0
FP	0	0	0	0	0	0
Recall (R)	100	97,96	100	97,959	100	100
Precision (P)	98	98	94,231	96,970	96,078	94,231
F1	98,990	97,959	97,030	97,462	98,000	97,030

Tab. 5.1 Výsledky úspešnosti extrakcie KF

Podľa návrhu metódy 2: *Pomocou kľúčových snímok* sme hodnotili úspešnosť extrakcie KF navrhnutých metód 1. *So sledovaním a prahovaním so štandardnými metrikami a priemerovaním adaptívnym prahom* a 2. *So sledovaním a prahovaním s metrikou DFT a Dugadovym prahom* pomocou modifikovaných kritérií rozoznatel'nosti P, R a F1 miery. Výsledky obsahuje Tab. 5.1.

Pre nájdenie lokálneho extrému a hodnoty prahu sme použili okno W v rozmedzí veľkostí 17 až 43 a konštantu $k = 1$. Pre určenie kľúčových snímok sme zvolili priestorovo podvzorkované video vo formáte CIF kvôli potrebe rýchlosti, ktorá následne umožní ich využitie v reálnom čase v kódéri a overenie, či táto veľkosť je naozaj použiteľná pre dostatočnú zrozumiteľnosť video komunikácie tak, ako je všeobecne odporúčaná ITU.

Metriky DFT a DCT sme implementovali v prostredí MATLAB [Mat'08]. Výsledky ukazujú, že extrakcia KF pomocou DCT a DFT pre $\lambda = 0$ s trackovaním pomocou Centertrackeru [Kon'11], prahovaním podľa modifikovaného Dugadovho modelu [Dug'98] dosahuje najlepšie výsledky. Pričom pri testovaní sa zistilo, že výsledky úspešnosti extrakcie KF pomocou DCT a DFT pre $\lambda = 0$ sú veľmi podobné až zhodné, keďže pri extrakcii je minimálny počet rozdielne detegovaných snímok. Overením predpokladu, že rozdiel fáz medzi spektrami susedných snímok je takmer nulový, môžeme z oboru komplexných čísiel, v ktorých je počítaná DFT prejsť do oboru reálnych čísiel a použiť DCT ako metriku na extrakciu KF (nejedná sa o DCT metriku uvádzanú v prehľade objektívnych metód). DCT je ideálnou náhradou za DFT pre $\lambda = 0$, a to z nasledujúcich dôvodov:

1. jednoduchší výpočet bázy v obore reálnych čísiel
2. veľmi podobné až zhodné výsledky, a teda minimálny počet rozdielne detegovaných snímok
3. DCT je implementovaná v kódéri, a preto už v procese kódovania môže posielat' informácie o snímkach a extrahovat' KF, teda kódér by si mohol sám kontrolovať a korigovať žiadanú kvalitu pri prenose aj kódovaní .

Navrhnutá metóda hodnotenie kvality videa pomocou hľadania KF a DCT je dostatočne úspešná a môže byť použitá v priebehu procesu hodnotenia kvality videa s prstovou abecedou. Môžeme teda hodnotiť len kvalitu jednotlivých snímok videa, ktoré predstavujú jednotlivé známky (kľúčové snímky) namiesto toho, aby sa hodnotila celá video sekvencia, prípadne metódu extrakcie KF použiť ako predspracovanie pre automatické rozpoznanie znakov vo videách s posunkovým jazykom.

				BEZ MASKY				
Logatom								
Format	QP	Transmission rate [kbit/s]	recognizability [%]	MSAD	VQM	SSIM	PSNR	MSE
4cif	30	199,707	94,84	1,651	1,115	0,956	38,414	12,036
4cif	40	74,662	73,40	2,671	1,642	0,905	33,764	29,146
4cif	50	27,172	59,23	4,701	2,589	0,848	29,743	70,780
Correlation				-0,953	-0,961	0,989	0,997	-0,938
cif	30	113,205	91,46	2,431	1,380	0,922	34,958	21,889
cif	40	30,680	73,77	3,528	2,062	0,871	31,497	46,869
cif	50	11,108	39,39	5,809	3,205	0,835	28,151	101,021
Correlation				-1,000	-0,999	0,961	0,981	-1,000

				Maska staticka – elipsa				
Logatom								
Format	QP	Transmission rate [kbit/s]	recognizability [%]	MSAD	VQM	SSIM	PSNR	MSE
4cif	30	199,707	94,840	0,618	0,595	0,987	43,729	5,223
4cif	40	74,662	73,400	0,915	0,862	0,977	39,504	9,013
4cif	50	27,172	59,230	1,741	1,503	0,962	34,984	21,682
Correlation				-0,928	-0,939	0,978	0,991	-0,913
cif	30	113,205	91,460	0,783	0,685	0,981	40,683	5,983
cif	40	30,680	73,770	1,153	1,085	0,969	37,382	12,312
cif	50	11,108	39,390	2,117	1,965	0,958	33,250	31,651
Correlation				-0,998	-1,000	0,981	0,993	-0,995

				Dynamická maska				
Logatom								
Format	QP	Transmission rate [kbit/s]	recognizability [%]	MSAD	VQM	SSIM	PSNR	MSE
4cif	30	199,707	94,840	0,301	0,435	0,994	47,061	2,686
4cif	40	74,662	73,400	0,462	0,663	0,988	42,648	5,113
4cif	50	27,172	59,230	0,766	1,216	0,981	38,578	11,645
Correlation				-0,957	-0,939	0,988	0,996	-0,930
cif	30	113,205	91,460	0,400	0,520	0,990	43,906	3,606
cif	40	30,680	73,770	0,592	0,859	0,984	38,676	7,467
cif	50	11,108	39,390	0,971	1,648	0,980	36,817	17,461
Correlation				-1,000	-0,999	0,961	0,900	-0,998

				DCT KF				
Logatom								
Format	QP	Transmission rate [kbit/s]	recognizability [%]	MSAD	VQM	SSIM	PSNR	MSE
4cif	30	199,707	94,840	2,104	1,084	0,955	38,179	10,030
4cif	40	74,662	73,400	3,373	1,648	0,887	33,285	30,735
4cif	50	27,172	59,230	5,854	2,743	0,794	29,052	81,079
Correlation				-0,955	-0,955	0,979	0,997	-0,938
cif	30	113,205	91,460	3,057	1,532	0,904	34,071	26,994
cif	40	30,680	73,770	4,524	2,333	0,831	30,592	57,770
cif	50	11,108	39,390	7,216	3,403	0,769	27,456	117,128
Correlation				-1,000	-0,995	0,974	0,977	-1,000

Tab. 5.2 Porovnanie výsledkov objektívneho hodnotenia kvality videa

Pôvodné vedecké prínosy :

1. Návrh a realizácia metódy hodnotenia kvality videa s prstovou abecedou prostredníctvom kľúčových snímok.

Metóda je založená na sledovaní pohybu ruky metódou založenou na korelácii, výbere kľúčových snímok a následného hodnotenia kvality len na tejto vybranej množine.

2. Návrh a implementácia novej metódy extrakcie kľúčových snímok z videa s prstovou abecedou, ktorá rozhoduje v priestore diskkrétnej kosínusovej transformácie II.
3. Pôvodné spektrálne metriky sú založené na Fourierovej transformácii, v práci realizovaná DCT ukazuje, že počítanie spektra v obore reálnych čísiel je postačujúce, a je to veľmi komfortnou náhradou za Fourierovu bázu ortogonálnych funkcií v priestore komplexných čísiel. DCT je v súčasnosti nevyhnutnou súčasťou video kóderov. Je preto jasné, že kóder videa s takýmto typom obsahu, je schopný priamo označiť snímky s relevantnou informáciou. To je dôležité aj pre automatické prekladače z prstovej abecedy do písaného textu.

Prínosy v metodologickej oblasti :

1. Pôvodná metodika bola založená na kontrole všetkých snímok. V práci je ukázané, že meranie len na významných snímkach je dostačujúce pre predpoveď zrozumiteľnosti videa s prstovou abecedou.
2. Modifikácia mier "Recall", "Precision" a "F1 rate" pre použitie vo vyhodnotení rozoznatel'nosti komunikácie prstovou abecedou.

6 KONKRÉTNE ZÁVERY PRE ĎALŠÍ ROZVOJ VEDY

V práci je navrhnutá metóda hodnotenia kvality videa, ktorá bola špeciálne vyvinutá pre videá s prstovou abecedou. Ukázali sme, že pre hodnotenie kvality je postačujúce hodnotiť len kľúčové snímky videa. Tie sú v tejto práci získané modifikáciou Dugadovej metódy extrakcie ostrých strihov vo videu. Získané snímky úspešne určujú snímky s jednotlivými písmenovými znakmi slov prstovej abecedy a medzier medzi slovami. Po zhodnotení rozpoznateľnosti znakov si dovoľujeme predpokladať, že takto získané snímky môžu byť vstupom pre automatické rozpoznávanie znakov prstovej abecedy. Úspešnosť takéhoto postupu však závisí od kvality pôvodného videa, z ktorého sú znaky extrahované ako kľúčové snímky.

Po zohľadnení vlastností rôznych iných typov znakov a gest vytváraných jednoručne aj dvojručne by mohla byť metóda použitá ako predspracovanie aj iných typov rozpoznávačov rôznych alternatívnych spôsobov komunikácie prostredníctvom statických znakov alebo posunkov. Po úprave v spolupráci s databázou znakov alebo gest by zrejme bolo možné modifikovať navrhnutú metódu aj pre dynamické znaky a gestá.

Navrhnutá metóda používa diskretnú kosínusovú transformáciu. Jej úpravou pre blokové video kodéry by mohla byť integrálnou súčasťou video kodérov, kedy by kódér priamo mohol predpovedať okrem kvality videa aj jeho rozpoznateľnosť v prípade, že ide o video s alternatívnou komunikáciou. Sledovanie znakov môže byť využité aj pre určenie regiónov záujmu napríklad v spravodajských reláciách s podporou pre sluchovo postihnutých divákov.

7 REFERENCIE

- [Ang'97] Angelo, Dianne H. (1997). "AAC in the family and home". In Glennen, Sharon; DeCoste, Denise C. Handbook Of Augmentative And Alternative Communication. San Diego: Singular Publishing Group. ISBN 1-56593-684-1.
- [Avc'02] Avcibaş, I. et al.: Statistical evaluation of image quality measures. In: Journal of Electronic Imaging, vol. 11, 2002, no.2, pp. 206-223.
- [Bat'02] Batshaw, M. L.; Shapiro, B. (2002). Batshaw, M. L.. ed. Children with disabilities (5th ed.). Baltimore: Paul H. Brookes Publishing Company. ISBN 0864331371.
- [Bro'06] BROWNE, Paul, Alan F SMEATON, Noel MURPHY, Noel O'CONNOR, Sean MARLOW a Catherine BERRUT. Evaluating and combining digital video shot boundary detection algorithms. In: *Proceedings of the 8th ACM SIGMM International Workshop on Multimedia Information Retrieval MIR '06*. New York, N.Y: ACM Press, 2006, s. 231-238. ISBN 1-59593-495-2.
- [Dug'98] DUGAD, Rakesh, Krishna RATAKONDA a Narendra AHUJA. Robust video shot change detection. In: *Multimedia Signal Processing, 1998 IEEE Second Workshop on*. 1998, s. 376-381.
- [Hef'09] Hefty, M. : Prstová abeceda. Organizácia Myslím – rozvoj myslenia nielen pre sluchovo postihnutých , 2009. [Online] [Dátum: 6.4. 2011] [http:// www.zzz.sk](http://www.zzz.sk)
- [Her'12a] Heribanová, Petra - Polec, Jaroslav - Tarcsiová, Darina: An Evaluation of Finger Alphabet Intelligibility Using Quality Assessment of Video with Masked Content. In: GraphiCon 2012 : 22nd International Conference on Computer Graphics and Vision. Moscow, Russia, October 1-5, 2012. - Moscow : MAKS Press, 2012. - ISBN 978-5-317-03808-3. - S. 256-261
- [Her'12b] Heribanová, Petra - Polec, Jaroslav: Real-Time Finger Alphabet Logatom Recognizability Measurement by Distance Geometry. In: ELITECH'12 [elektronický zdroj] : 14th Conference of Doctoral Students. Bratislava, Slovak Republic, 22 May 2012. - Bratislava : Nakladateľstvo STU, 2012. - ISBN 978-80-227-3705-0. - CD-ROM, [4] s.
- [Her'14] Heribanová, Petra – Jagelková, Jana - Polec, Jaroslav: Key frame extraction for finger-spelling alphabets recognition. In: ELITECH'14 – článok podaný

- [ITUT'08]** ITU-T P.910: 2008, Subjective video quality assessment methods for multimedia applications.
- [ITUR'09]** ITU-R BT.500-12: 2009, Methodology for the subjective assessment of the quality of television pictures.
- [Kawa'08]** Kawayoke, Y., Horita, Y: NR objective continuous video quality assessment model based on frame quality measure. In: 15th IEEE International Conference on Image Processing, 2008, pp. 385-388.
- [Kon'11]** Konen Stephan: Centertracker, 2011 [online 27.4.14]
- [Mak'95]** MAKÁŇ, F., Elektroacoustika, Bratislava: Vydavateľstvo STU, 1995.
- [Mat'08]** Mathworks. 2008. Matlab 7.7.0 R2008b. <http://www.mathworks.com>
- [MSU'08]** MSU Video Quality Measurement Tool. MSU Graphics & Media Lab (Video Group) Moskva, 2008. [Online][Dátum: 5.2.2011]
- [Olson'08]** Olson, David L.; and Delen, Dursun (2008); Advanced Data Mining Techniques, Springer, 1st edition (February 1, 2008), page 138, ISBN 3-540-76916-1
- [Omi'11]** M. Omidyeganeh, S. Ghaemmaghami, and S. Shirmohammadi, "Video Keyframe Analysis Using a Segment-Based Statistical Metric in a Visually Sensitive Parametric Space," IEEE Transactions on Image Processing, vol. 20, issue 10, pp. 2730:2737, 2011.
- [Pol'13]** Polec, Jaroslav - Heribanová, Petra - Hirner, Tomáš: Key Frames Extraction for Sign Language Video Analysis and Recognition. In: World Academy of Science, Engineering and Technology [elektronický zdroj]. - ISSN 2010-376X. - Iss. 78 (2013), s. 479-483
- [Ries'08]** Ries, M. et al.: Video quality estimation for mobile H.264/AVC video streaming. In: Journal of Communications, vol.3, 2008, no.1, pp. 41-50.
- [Tar'02]** Tarcisová Darina: Prstová abeceda. Nitra: Effeta, 2002. ISBN 80- 9698584 -4 -0
- [Tar'08]** Tarcisová, D.: Pedagogika sluchovo postihnutých. MABAG spol. s r. o., Bratislava, 2008. ISBN 978-80-89113-52-1
- [Wink'05]** Winkler, S.: Digital video quality vision model and metrics. 1. vyd. Chichester : John Wiley & Sons Ltd., 2005. ISBN 0-470-02404-6

8 PUBLIKÁCIE AUTORKY A PROJEKTY

Publikácie autorky vygenerované z Centrálného registra publikácií

Spolu registrované v SCOPUS bez WoS 3

WoS a SCOPUS 1

ADE Vedecké práce v zahraničných nekarentovaných časopisoch (5)

ADE1 HERIBANOVÁ, PETRA [40%] - POLEC, JAROSLAV [40%] - ONDRUŠOVÁ, SANDRA [10%] - Host'ovecký, Marek [10%]: *Intelligibility of Cued Speech in Video*.

In: World Academy of Science, Engineering and Technology [elektronický zdroj]. - ISSN 2010-376X. - Iss. 55 (2011), s. 492-496 (v indexe Scopus) [Projekt č.: 1/0602/11]

ADE2 HERIBANOVÁ, PETRA [40%] - POLEC, JAROSLAV [10%] - POCTAVEK, JÁN [10%] - Mordelová, Angela [40%]: *Intelligibility Threshold for Cued Speech in H.264 Video Conference*. In: International Journal of Electronics and Telecommunications. - ISSN 0867-6747. - Vol. 57, Iss. 3 (2011), s. 383-387 (v indexe Scopus) [Projekt č.: 1/0602/11. - 119-005TVU-4/2010]

ADE3 HERIBANOVÁ, PETRA [40%] - KUČEROVÁ, JÚLIA [40%] - POLEC, JAROSLAV [20%]: *Key Frames Extraction for Sign Language Using Visual Attention Approach*. In: World Academy of Science, Engineering and Technology [elektronický zdroj]. - ISSN 2010-376X. - Iss. 82 (2013), s. 561-567 [Projekt č.: 1/0602/11]

ADE4 HERIBANOVÁ, PETRA [50%] - POLEC, JAROSLAV [40%] - MARTINOVIČ, MICHAL [10%]: *Video Quality Control Using a ROI and Two-Component Weighted Metrics*. In: World Academy of Science, Engineering and Technology [elektronický zdroj]. - ISSN 2010-376X. - Iss. 79 (2013), s. 708-711 [Projekt č.: 1/0602/11]

ADE5 POLEC, JAROSLAV [50%] - HERIBANOVÁ, PETRA [40%] - HIRNER, TOMÁŠ [10%]: *Key Frames Extraction for Sign Language Video Analysis and Recognition*. In: World Academy of Science, Engineering and Technology [elektronický zdroj]. - ISSN 2010-376X. - Iss. 78 (2013), s. 479-483 [Projekt č.: 1/0602/11]

AFC Publikované príspevky na zahraničných vedeckých konferenciách (3)

AFC1 HERIBANOVÁ, PETRA [60%] - POLEC, JAROSLAV [20%] - TARCSIOVÁ, DARINA [20%]: *An Evaluation of Finger Alphabet Intelligibility Using Quality Assessment of Video with Masked Content*. In: GraphiCon 2012 : 22nd International Conference on Computer Graphics and Vision. Moscow, Russia, October 1-5, 2012. - Moscow : MAKS Press, 2012. - ISBN 978-5-317-03808-3. - S. 256-261 [Projekt č.: 1/0602/11]

AFC2 HERIBANOVÁ, PETRA [40%] - POLEC, JAROSLAV [10%] - POCTAVEK, JÁN [10%] - MORDELOVÁ, ANGELA [40%]: *Intelligibility Threshold for Cued Speech in H.264 Video*. - Vyšlo aj v: International Journal of Electronics and Telecommunications. - Vol. 57, No. 3 (2011), p. 383-387. (v indexe Scopus) In: SPS 2011 Proceedings : Signal Processing Symposium. Jachranka Village, Poland, June 8-10, 2011. - Warsaw : Institute of Electronic Systems, 2011. - CD-Rom [Projekt č.: 1/0602/11. - 119-005TVU-4/2010]

AFC3 HERIBANOVÁ, PETRA [40%] - POLEC, JAROSLAV [40%] - KRULIKOVSKÁ, LENKA [20%]: *Logatom Intelligibility of Single-Handed Finger Alphabet*. In: Proceedings ELMAR-2011 : 53rd International Symposium ELMAR-2011, 14-16 September 2011, Zadar, Croatia. - Zadar : Croatian Society Electronics in Marine, 2011. - ISBN 978-953-7044-12-1. - S. 71-74 (v indexe Web of Science and Scopus) [Projekt č.: 1/0602/11. - 119-005TVU-4/2010]

AFD Publikované príspevky na domácich vedeckých konferenciách (4)

AFD1 HERIBANOVÁ, PETRA [40%] - POLEC, JAROSLAV [10%] - POCTAVEK, JÁN [10%] - MORDELOVÁ, ANGELA [40%]: *Intelligibility of Cued Speech in Videoconference*. In: Inovačný proces v e-learningu, Bratislava, 10. marec 2011. - Bratislava : Ekonóm, 2011. - ISBN 978-80-225-3112-2. - CD-Rom [Projekt č.: 119-005TVU-4/2010]

AFD2 HERIBANOVÁ, PETRA [50%] - POLEC, JAROSLAV [50%]: *Intelligibility of Finger Alphabet in Videoconference Based on Logatom Recognizability*. In: ELITECH'11 : 13th Conference of Doctoral Students Faculty of Electrical Engineering and Information Technology. Bratislava, Slovak Republic, 17 May, 2011. - Bratislava : Nakladateľstvo STU, 2011. - ISBN 978-80-227-3500-1. - S. 1-4 [Projekt č.: 119-005TVU-4/2010]

AFD3 HERIBANOVÁ, PETRA [40%] - Polec, Jaroslav [10%] - Mordelová, Angela [40%] - Poctavek, Ján [10%]: *Intelligibility of Single-Handed and Double-Handed Finger Alphabets*.

In: Redžúr 2011 : proceedings; 5th International Workshop on Multimedia and Signal Processing. May 12, 2011, Bratislava, Slovak Republic. - Bratislava : Nakladateľstvo STU, 2011. - ISBN 978-80-227-3506-3. - S. 37-40 [Projekt č.: 119-005TVU-4/2010]

AFD4 HERIBANOVÁ, PETRA [50%] - POLEC, JAROSLAV [50%]: *Real-Time Finger Alphabet Logatom Recognizability Measurement by Distance Geometry*. In: ELITECH'12 [elektronický zdroj] : 14th Conference of Doctoral Students. Bratislava, Slovak Republic, 22 May 2012. - Bratislava : Nakladateľstvo STU, 2012. - ISBN 978-80-227-3705-0. - CD-ROM, [4] s. [Projekt č.: 1/0602/11]

AFK Postery zo zahraničných konferencií (1)

AFK1 HERIBANOVÁ, PETRA [50%] - Polec, Jaroslav [50%]: *Logatom Recognizability of Finger Alphabet in Video*. In: ICVSS 2011. Registration, Recognition and Reconstruction in Image and Video : International Computer Vision Summer School 2011. Sicily, 11-16 July, 2011. - Catania : University of Catania, 2011. - S. 36 [Projekt č.: 1/0602/11]

Výskumné projekty

- The Optimization of Video Coding Efficiency for Transmission and Record. VEGA 1/0602/11. Duration: 2011-2013 (J. Polec)
- The Development of Digital Literacy in Primary and Secondary School. KEGA 119-005TVU-4/2010. Duration: 2010-2011 (J. Polec)
- Evaluačné kritéria hodnotenia videa. GRANT UK/106/2012. Doba: 2011/2012

Mobilitné projekty

- International Computer Vision Summer School ICVSS 2011. Štipendium Ľ. Čuchranovej. Slovensko- český ženský fond. Doba: 2010/2011
- International Computer Vision Summer School ICVSS 2011. Štipendium Hlavička č. 68/2011. Nadácia SPP. Doba: 2010/2011
- Moskovská štátna univerzita (Lomonosov). Štipendium Študenti idú do sveta č. 11Sds078. Nadácia Tatra Banka. Doba: 2011/2012

9 RESUMÉ

Title: VIDEO QUALITY ASSESSMENT METHODS

Keywords: cued speech, sign alphabet, intelligibility, video quality, DCT, key frame

This work deals with the evaluation of quality for videos containing alternative communication. Particular form of communication deaf sign language and finger alphabet , which is important approach to quality assessment based on clarity of communication is not " beauty" image than photorealistic images with content.

The work provides an extensive overview of the current stage and often less frequently used methods of subjective and objective assessment of image quality and video. Also disclosed are methods for extracting key frames and thresholding.

The second part deals with the design, specification and implementation methods of assessing video quality by extracting key frames, taking into account the region of interest, choosing appropriate metrics to measure the quality and use of tracking and thresholding. Also shows an objective method of testing logatómvej intelligibility (as used in telefonometrii the articulation of speech sounds) using a subjective score according to ACR method using " sign logatómov " for the deaf. Shows the results obtained, their assessment and a description of the advantages and disadvantages of the chosen method of measuring the quality of video sequences containing the finger alphabet. The thesis also shows possible directions for future research at the end.

The paper proposed video quality assessment method that has been specially developed for videos with the finger alphabet. We have shown that the quality assessment is sufficient to examine a selected video frames. These are obtained in this work by modifying Dugadovej extraction methods sharp cuts in the video. Acquired images successfully identify places with different alphabetic characters finger alphabet words and spaces between words . After evaluating discoverable characters take the liberty to assume that the resulting images can be input for automatic character recognition finger alphabet. The success of this procedure depends on the quality of the original video from which the characters are extracted as key frame.

The proposed method uses the Discrete Cosine Transform (DCT). By adjusting the block The units could be an integral part of a video encoder , the encoder which can directly predict in addition to its video quality and the visibility in the case of the video communication with the alternative.