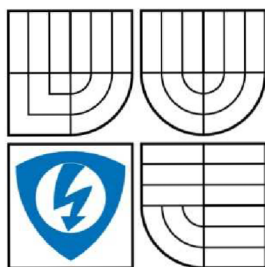


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLGIÍ
ÚSTAV TELEKOMUNIKACÍ
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

METODY A APLIKACE DETEKCE MRKÁNÍ OČÍ S VYUŽITÍM ČÍSLICOVÉHO ZPRACOVÁNÍ OBRAZU

METHODS AND APPLICATIONS OF EYE BLINK DETECTION WITH DIGITAL IMAGE
PROCESSING

DOKTORSKÁ PRÁCE
DOCTORAL THESIS

AUTOR PRÁCE
AUTHOR

Ing. JAN VLACH

VEDOUCÍ PRÁCE
SUPERVISOR

Mgr. PAVEL RAJMIC, Ph.D.

BRNO 2008

Děkuji vedoucímu disertační práce Mgr. Pavlovi Rajmicovi, Ph.D. za cenné připomínky a za vytvoření podmínek pro řešení disertační práce. Dále děkuji rodině za podporu při studiu na vysoké škole.

P r o h l á š e n í

Prohlašuji, že jsem předloženou disertační práci zpracoval sám s konzultační pomocí vedoucího projektu a uvedených konzultantů. Použité literární prameny jsou uvedeny v literárních odkazech.

V Brně dne :

Podpis:

ABSTRAKT

Disertační práce pojednává o oblasti detekce mrkání jako součásti komplexního oboru detekce a rozpoznávání obličejů. Práce je zaměřena směrem k číslicovému zpracování obrazu. Součástí je analýza problematiky a popis dostupných obrazových databází vhodných k testovacím účelům. Ve dvou hlavních kapitolách jsou návrhy metod pro detekci mrkání s využitím číslicového zpracování obrazu jak bez IR technologie, tak s IR technologií.

ABSTRACT

The thesis deals with eye blink detection, which is part of complex topic of face detection and recognition. The work intents on digital image processing. There is analyse of the topic and description of image databases for testing. Two main chapters describe design of eye blink detection with digital image processing with IR technology and without IR technology.

Seznam použitých zkratek, symbolů a značení

\mathbb{R}	množina všech reálných čísel
\mathbb{Z}	množina celých čísel
\mathbb{C}	množina komplexních čísel
\mathbb{E}^2	2D euklidovský prostor
HSV	barevný tón, sytost, jas (H ue, S aturation, V alue)
HSL	barevný tón, sytost, jas (H ue, S aturation, L ightness) = HSV
PCA	P rincipal C omponent A nalysis
RGB	červená, zelená, modrá (R ed, G reen, B lue)
YCbCr	chromatický model
ROI	oblast zájmu (R egion O f I nterest)
NIR	blízké infračervené záření
SWIR	krátké vlnové délky
MWIR	střední vlnové délky
LWIR	dlouhé vlnové délky
FIR	dlouhé infračervené záření
I	jasová složka
ADF	metoda akumulace rozdílových snímků
r_b	frekvence mrkání

Vysvětlení některých pojmů

- *pixel* – jeden bitmapový obrazový bod
- *oční krajina* – oblast očí, popř. očníce (oko, oční víčka, řasy, oční klenba, někdy i obočí)
- *detekce živosti* (liveness detection) – podpůrný podsystém biometrických systémů, který odhalí snahu o podvádění při snímání biometrického rysu
- *kompenzační technologie* (assistive technology) – technologie, která umožňuje ovládat určité zařízení nebo vykonávat určitou činnost při pohybových, smyslových či mentálních omezeních
- *metoda rozdílových snímků* – nejprimitivnější elementární metoda při zpracování videosekvencí (obvykle při detekci pohybu)
- *frekvence mrkání* – průměrná hodnota, poměr počtu mrknutí za časový interval
- *binární funkce* – jednorozměrný vektor, výsledek po prahování, obvykle nabývá hodnot 0 nebo 1
- *binární mapa* – dvourozměrná matice, výsledek po prahování, obvykle nabývá hodnot 0 nebo 1
- *barevný tón, odstín* (hue) – barva odražená nebo procházející objektem, měří se jako poloha na standardním barevném kole (0° až 360°); obecně se odstín označuje názvem barvy
- *syťost barvy* (saturation) – syťost barvy, příměs jiné barvy; někdy též chroma, síla nebo čistota barvy; představuje množství šedi v poměru k odstínu, měří se v procentech od 0% (šedá) do 100% (plně syťá barva); na barevném kole vzrůstá syťost od středu k okrajům (např. červená s 50% syťostí bude růžová)
- *hodnota jasu* (value) – množství bílého světla; relativní světlost nebo tmavost barvy; jas vyjadřuje kolik světla barva odráží, dalo by se také říct přidávání černé do základní barvy
- *kolorometrie* – fyzikální teorie barev, měření barev
- *eigenfaces* – je sada vlastních vektorů, které se využívají v statistické metodě PCA pro detekci a rozpoznávání obličejů
- *PCA* (Principal Component Analysis) – je technika redukce dimenze příznakového prostoru
- *neuronové sítě* – je jedním z výpočetních modelů používaných v umělé inteligenci
- *klasifikátor* – procedura (popř. celý systém), která má za úkol vyhodnotit data (obvykle část obrazu) a přiřadit mu určité vlastnosti či příznak (je obličej, není obličej)
- *Beziérová křivka* – je jednou z mnoha parametrických křivek; umožňuje interaktivní vytváření a modifikaci jejího tvaru; pomocí Bézierovy křivky je také možno reprezentovat i interpolační křivky
- *ohraničující box* (*bounding box*) – ohraničuje oblast zájmu (ROI)
- *kvantizační zkreslení* – charakteristické zkreslení číslicového signálu vznikající v signále vyjádřením signálové funkce konečným počtem diskrétních hladin
- *multispektrální obraz* – barevný obraz

Obsah

<i>Seznam použitých zkratk, symbolů a značení</i>	<i>iv</i>
<i>Vysvětlení některých pojmů</i>	<i>v</i>
<i>Obsah</i>	<i>vi</i>
<i>Seznam obrázků</i>	<i>ix</i>
<i>Úvod</i>	<i>1</i>
1 Přehled současné problematiky	2
1.1 Metody detekce bez vazby na určitou aplikaci	3
1.2 Aplikace detekce mrkání.....	4
1.2.1 Havarijní řídicí systémy.....	4
1.2.2 Detekce živosti v biometrických systémech.....	5
1.2.3 Interakce člověk-počítač.....	7
2 Cíle disertační práce	10
3 Analýza problematiky detekce a rozpoznávání mrkání	11
3.1 Mrkání očí	11
3.2 Rozdělení problematiky podle přístupu	13
3.2.1 Invazivní monitorování.....	13
3.2.2 Digitální zpracování obrazu.....	14
3.2.3 IR technologie.....	14
3.3 Analýza modelu postupu detekce mrkání	14
3.3.1 Detekce obličeje.....	15
3.3.2 Detekce očí	17
3.4 Cíle detekce a s ním spojené problémy.....	17
3.5 Systém pro rozpoznávání mrkání očí.....	18
3.5.1 Kvalita obrazových dat	18
3.5.2 Měřené události mrkání očí	19
3.6 Lidské oko a záření	21
3.6.1 Oko a zrak.....	21
3.6.2 Elektromagnetické záření	23
3.6.3 Infračervené záření	23
3.6.4 Nebezpečné účinky infračerveného záření	24

3.6.5	Reakce oka na infračervené záření	24
4	Číslicové zpracování obrazu.....	26
4.1	Digitální reprezentace obrazu	27
4.1.1	Vzorkování	27
4.1.2	Kvantizace	27
4.2	Základy zpracování obrazu (předzpracování obrazu).....	28
4.2.1	Obrazová funkce.....	28
4.2.2	Konvoluce.....	29
4.2.3	Filtrace	29
4.2.4	Matematická morfologie.....	31
4.2.5	Detekce hran	32
4.2.6	Obrazové derivace	33
4.3	Metody detekce obličeje a očí.....	34
4.3.1	Detekce v šedotónových (monochromatických) obrazech.....	34
4.3.2	Klasifikace v barevných obrazech.....	37
4.4	Metody detekce mrknutí	42
4.4.1	Detekce mrknutí pomocí rozpoznání duhovky.....	42
4.4.2	Detekce mrknutí pomocí diferenční matice.....	43
4.4.3	Detekce očního víčka.....	44
5	Obrazové databáze.....	46
5.1	Dostupné obecné obrazové databáze	46
5.1.1	Popis dostupných databází.....	47
5.2	Dostupné obrazové databáze pro účely detekce mrkání	52
5.2.1	Popis databází pro účely detekce mrkání.....	52
6	Detekce mrkání bez využití IR technologie.....	55
6.1	Hardwarová realizace.....	55
6.2	Softwarová realizace.....	55
6.2.1	Návrh softwarové realizace	55
6.2.2	Funkce Detekce	57
6.2.3	Funkce Lokalizace očí	60
6.2.4	Podprogram Detekci změny oblasti očí.....	65
6.2.5	Podprogram Vyhodnocení mrknutí z průběhu shody	66

6.2.6	Podprogram na odstranění malých oblastí.....	67
6.2.7	Časová náročnost programu	67
6.3	Závěr	68
7	Detekce mrkání s využitím IR technologie.....	70
7.1	Hardwarová realizace.....	70
7.1.1	Náležitosti hardwarové realizace	70
7.1.2	Možnosti hardwarové realizace	72
7.1.3	Konečná hardwarová realizace	73
7.2	Softwarová realizace	74
7.2.1	Návrh softwarové realizace	74
7.2.2	Konečná softwarová realizace	77
7.2.3	Větev analýzy	83
7.3	Závěr	84
8	Závěr	85
	<i>Literatura.....</i>	<i>86</i>
	Seznam vlastních prací	95
	CURRICULUM VITÆ.....	97

Seznam obrázků

obr. 1.1: Fáze detekce mrknutí z pohledu problematiky detekce a rozpoznávání obličejů (první přiblížení).....	2
obr. 1.2: Vysvětlení nejednoznačnosti zařazení konkrétní metody do konkrétní fáze zpracování.	4
obr. 3.1: Popis důležitých částí oka.....	12
obr. 3.2: Modelový průběh rychlého a pomalého mrknutí.	13
obr. 3.3: Univerzální model postupu při detekci mrkání, využívající zpracování obrazu (druhé přiblížení)	15
obr. 3.4: Schematické rozdělení metod detekce obličeje podle formátu vstupních dat	16
obr. 3.5: Příklady kvality obrazových dat.	18
obr. 3.6: Příklady měřených událostí mrkání očí.	19
obr. 3.7: Postup získání frekvence mrkání.	20
obr. 3.8: Příklad analýzy pečlivě zaznamenaného průběhu mrkání.	21
obr. 3.9: Řez okem s popisem jednotlivých částí.....	22
obr. 3.10: Zjednodušené spektrum elektromagnetického záření.....	23
obr. 3.11: Závislost relativní škodlivosti infračerveného záření na vlnové délce [83].	25
obr. 4.1: Ukázka čtvercové a hexagonální mřížky.	27
obr. 4.2: Filtrace standardním průměrováním.	30
obr. 4.3: Příklady mediánového filtru.	31
obr. 4.4: Příklady strukturních elementů.....	31
obr. 4.5: Dilatace v praxi, strukturní element ve tvaru disku o poloměru 2.	32
obr. 4.6: Eroze v praxi, strukturní element ve tvaru disku o poloměru 2.....	32
obr. 4.7: Detekce hran – Sobelův operátor.....	34
obr. 4.8: Převod z modelu RGB do modelů YCbCr a HSV. Horní řada vždy představuje jednotlivé složky modelu v monochromatickém rozlišení; dolní řada představuje postupně každou složku maximálně potlačenou a pro lepší představu je výsledek převeden zpět do modelu RGB.	38
obr. 4.9: Barevný model HSV.....	39
obr. 4.10: Podprostor barvy kůže v modelu RGB (a), podprostor barvy kůže v modelu YCbCr (b).....	40
obr. 4.11: : Vzorová šablona barvy kůže.....	40
obr. 4.12: Otevřené oko a částečně přivřené s rozmezím.	43
obr. 5.1: Příklad z AT&T The Database of Faces.....	47
obr. 5.2: Příklad z BANCA Database	47
obr. 5.3: Příklad z Cohn-Kanade AU-Coded Facial Expression Database.....	48
obr. 5.4: Příklad z databáze Collection of Facial Images.....	48
obr. 5.5: Příklad z Georgia Tech face database.....	49
obr. 5.6: Příklad z MMI Face Database	49

obr. 5.7: Příklad z NRC-IIT Facial Video Database	50
obr. 5.8: Příklad z The BioID Face Database.....	50
obr. 5.9: Příklad z The University of Oulu Physics-Based Face Database	51
obr. 5.10: Příklad z The VidTIMIT Audio-Video Dataset.....	51
obr. 5.11: Příklad z The Yale Tace Database.....	51
obr. 5.12: Příklad z The Yale Face Database B	52
obr. 5.13: Příklad z ZJU Eyeblink Database	53
obr. 5.14: Příklad z ZJU Eyeblink Database.....	53
obr. 5.15: Příklad z SKY Database	54
obr. 6.1: Akumulační diferenční rámec se zvoleným malým segmentem (a) a velkým segmentem (b).	56
obr. 6.2: Plocha prezentace výsledků programu.	57
obr. 6.3: Obraz diferenční matice se špatným nastavením (a) a dobrým nastavením (b) prahu.	58
obr. 6.4: Obraz diferenční matice před filtrací (a) a po odstranění malých oblastí (b).	58
obr. 6.5: Obraz diferenční matice po odstranění oblastí mimo oblast zájmu.....	59
obr. 6.6: Jednotlivé nenormované diferenční snímky v pořadí 2 až 9.	59
obr. 6.7: Akumulační diferenční matice, která vznikla sečtením normovaných diferenčních snímku.	60
obr. 6.8: Prahovaná akumulční matice s prahem 0,4.....	60
obr. 6.9: Filtrace bwSum před dilatací (a), po dilataci (b).	61
obr. 6.10: Obraz rozdělených oblastí.	61
obr. 6.11: Zobrazení mezí pro určení očí.	62
obr. 6.12: První fáze určení středu obličeje, průměrným sloupcem.....	63
obr. 6.13: Rozdělený obraz levá převrácená strana (a), pravá strana (b).....	63
obr. 6.14: První fáze určení středu obličeje, průměrným sloupcem.....	64
obr. 6.15: Výsledek vynásobení dilatované masky a akumulčního rámce.....	64
obr. 6.16: Oblast očí a jejich binární podoba s prahovanou jasovou hodnotou 0.4. ..	65
obr. 6.17: Počátek překrytí šablony s aktuální oblastí.	66
obr. 6.18: Výstup podprogramu na detekci změny a jeho následné zpracování.	66
obr. 7.1: Webová kamera	70
obr. 7.2: CCD snímač použité webové kamery.....	71
obr. 7.3: CCD snímač použité webové kamery.....	72
obr. 7.4: Možnosti hardwarové realizace	73
obr. 7.5: Konečná hardwarové realizace	74
obr. 7.6: Názorná ukázka filtrování pomocí rozdílu intenzit objektu a jeho okolí	75
obr. 7.7: Diagram algoritmu hledajícího očí ve snímku.....	77
obr. 7.8: Bloky algoritmu v Simulinku	79
obr. 7.9: Ukázka 4-směrného záplavového vyplňování	81
obr. 7.10: Názorný příklad, vlevo detekované očí, vpravo analytický pohled.....	84

Úvod

V posledních letech dochází ke zlevňování a zlepšování kvality počítačové a snímací techniky. Počítače musí být stále častěji schopny interakce s okolním prostředím a to zejména s člověkem. Tyto důvody mají za následek zvětšující se zájem o výzkum na poli počítačového rozpoznávání objektů. Snad nejvíce energie bylo věnováno rozpoznávání lidského obličeje. Zejména proto, že nachází uplatnění v mnoha oborech (biometrie, dohledové systémy, uživatelské rozhraní, a další).

Detekce obličeje tvoří první krok v systémech pro rozpoznávání obličeje (popř. obličejových rysů) a jeho náležitou analýzu. Metody detekce a případného dalšího zpracování lze považovat za komplexní problematiku, neboť využívají znalostí a postupů z mnoha vědních oblastí jako je zpracování obrazu, pravděpodobnost a statistika, maticový počet, umělá inteligence, fuzzy logika nebo spektrální analýza.

Tato disertační práce pojednává o oblasti detekce a rozpoznávání mrkání, pro kterou je nalezení obličeje popř. očí v obraze klíčovým krokem. Studie, zabývající se detekcí a rozpoznáváním mrkání, sice nejsou tak mohutně rozšířeny, přesto lze nalézt několik motivačních přístupů, proč se výzkumem v této oblasti zabývat.

Text otevírá kapitola 1, ve které je uveden současný stav řešené problematiky, tj. přehled používaných metod pro detekci očí a mrkání a dále popis hlavních oblastí, ve kterých se detekce mrkání využívá jako důležitá část postupu.

Na zhodnocení současného stavu problematiky v oblasti detekce mrkání přirozeně navazuje kapitola 2, obsahující vymezení konkrétních cílů této dizertační práce.

Kapitola 3 analyzuje problematiku detekce mrkání z různých úhlů, včetně vysvětlení mrkání a popisu modelu pro detekci mrkání. Dále navazuje kapitola 4 s rozborem základů číslicového zpracování obrazu a metod pro detekci obličeje a očí.

V kapitole 5 jsou popsány nejdostupnější obrazové databáze pro zpracování statických obrazů i videosekvencí.

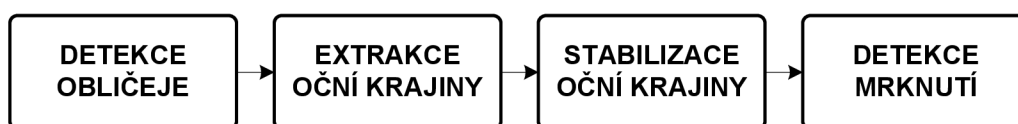
Návrh metody pro detekci mrkání s využitím číslicového zpracování obrazu bez IR technologií a dále s IR technologií je popsán v kapitole 6 a 7.

Na závěr je proveden souhrn zhodnocení přínosů dizertační práce. Rovněž jsou navrženy směry, kterými by se mohl ubírat další výzkum v této oblasti.

1 Přehled současné problematiky

Obsahem této kapitoly je uvedení a zhodnocení současného stavu problematiky, tj. dnes v praxi více či méně používaných postupů a jejich technik v oblasti detekce a rozpoznávání mrkání.

Jak již bylo řečeno, detekce mrkání spadá do komplexní oblasti detekce a rozpoznávání obličejů. Z tohoto pohledu lze proces zpracování rozdělit až do čtyř fází, kde každá má v závislosti na použitých algoritmech jinou váhu popř. složitost (obr. 1.1). Některé nemusí být součástí zpracování vůbec nebo je jich více sruženo do jedné.



obr. 1.1: Fáze detekce mrknutí z pohledu problematiky detekce a rozpoznávání obličejů (první přiblížení¹)

Komplexnost oblasti rozpoznávání obličejů spočívá v tom, že každá fáze zpracování je sama předmětem vědeckého vývoje a lze ji tak považovat za vlastní vědní disciplínu.

Techniky detekce obličejů (popř. lidí) v obraze je možné rozdělit na metody, které obraz rozdělí na oblasti, kde se můžou nacházet lidé a na metody, které používají detekci založenou na rysech člověka a nepotřebují předzpracování obrazu v takové míře. První typ metod se označuje jako metody s oddělením pozadí a druhý typ se označuje jako metody založené na přímé detekci [20], [36].

Lokalizace očí, nebo přesně řečeno extrakce oční krajiny, se užívá v postupech pro normalizaci obličejů, čímž je míněno co nejpřesnější vymezení oblasti obličejů v obraze, dále pro uživatelské rozhraní člověk-počítač, bezpečnostní systémy užívající lidskou oční duhovku k identifikaci apod.

Detekce mrkání pak umožňuje měření *frekvence mrkání* a spád očních víček. Tyto vjemy jsou často využívány např. k vyhodnocení bdělosti řidičů, kteří ovládají různé stroje – nemusí se jednat jen o řidiče automobilů, [86], [98].

¹ Jedná o povrchní náhled do problematiky.

1.1 Metody detekce bez vazby na určitou aplikaci

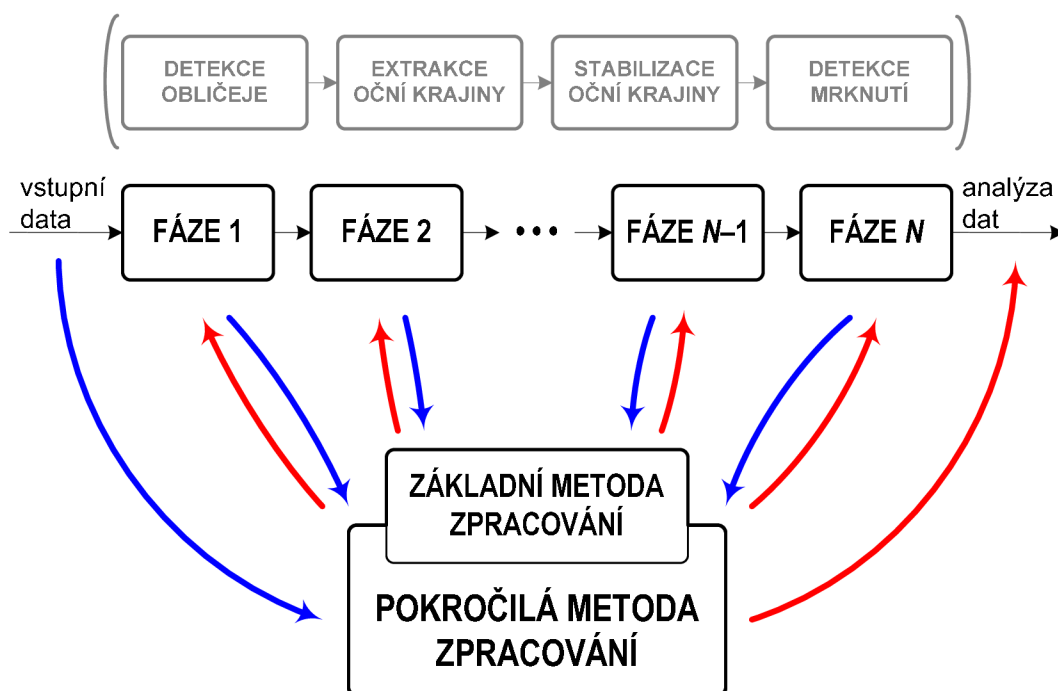
K tématu detekce mrkání bylo publikováno několik metod. V [16] využívají pro detekci mrkání *metodu rozdílových snímků*, oko (popř. oči) je zde bráno jako jeden celek. Složitější způsob [54] popisuje sledování očních příznaků (pohyb očních víček), kde při určitých změnách tvaru očí, dojde k vyhodnocení, že došlo k akci mrknutí. V práci [19] používají k detekci mrkání optický tok. Tento systém však pracuje správně za omezených podmínek, a to takových, aby se testovaná osoba dívala stále přesně přímo do objektivu snímacího zařízení. Překvapivě i po splněních takto přísných předpokladů, je ve zprávě uvedeno, že úspěšnost správně určeného mrkání je 65%. Tato hodnota je příliš nízká pro využití v praktických úlohách. Systém detekce mrkání navržený dle [23] dosahuje se svým přístupem úspěšnosti 96,5% a dokonce téměř v reálném čase. V jejich systému je metoda rozdílových snímků nejprve využita k získání segmentů, ve kterých došlo k pohybu. Dále je vybrán předem natrénovaný vzor „otevřených očí“ pro výpočet míry *Mahalanobisovy vzdálenosti*, která je použita pro získání nejpravděpodobnějších kandidátů očí. Ty jsou v dalších snímcích sledovány a pomocí míry korelace mezi aktuálním segmentem očí a odpovídajícím vzorem „zavřených očí“ je vyhodnoceno, zda došlo k mrknutí nebo ne. Nicméně tato metoda vyžaduje seřízení sledovacího zařízení v čase před zpracováním. Pokud sledující osoba změní svoji polohu, musí se změnit nastavení parametrů systému včetně natrénování nového vzoru. K selhání systému může také docházet, pokud se změní etnická národnost testujících osob, což může znamenat různorodost ve velikosti zkoumaných segmentů očí a vzdáleností mezi očima. Vylepšením metody [19] o sledovací algoritmus KTL (Kanade –Thomasi-Lucas) [29], [97], vznikl postup, který detekuje mrkání s přesností 97% [17]. Moriyamaova metoda² detekce mrkání [38] je založena na změně průměrného jasu v oblasti oční krajiny, je citlivá na světelné podmínky a šum. V další práci [53] byla zkoušena aktivní IR kamera k detekci mrkání pro předpověď únavy řidiče.

Mnoho prací, zabývajících se detekcí mrkání při zpracování obrazu, vyžaduje vysoce kontrolované podmínky a vstupní data v nejvyšší kvalitě, viz například systém pro automatické rozpoznávání lidských obličejových rysů [48].

V celé problematice zpracování obličejů je naráženo na problém rozdělení jednotlivých základních a pokročilých metod do jasně stavených fází zpracování. Téměř žádnou metodu nelze jasně považovat za metodu detekce obličeje a podobně. Navíc i ve fázích detekce a vyhodnocování (*klasifikace*) nejsou pevně stanoveny hranice. Pro klasifikaci se obvykle využívají statistické metody, ty však můžou být jinde považovány za metody detekce. Posledním paradoxem je skutečnost, že některé metody detekce mrkání vyžívají vícefázovou proceduru detekci obličeje a sledování očí

² Jediná dostupná a publikovaná práce, týkající se rozpoznávání mrkání a analýzy průběhu akce mrknutí

k získání informace o mrkání a na druhé straně jiné metody využívají právě detekci mrkání k nalezení obličeje. Grafické znázornění této problematiky je na obr. 1.2.



obr. 1.2: Vysvětlení nejednoznačnosti zařazení konkrétní metody do konkrétní fáze zpracování.

1.2 Aplikace detekce mrkání

S detekcí mrkání je spojeno několik aplikací, ve kterých je sledování mrkání očí hlavním pilířem. Nejpůsobivější technické aplikace, se týkají ochranných systémů, sloužících k zabránění automobilové havárie z důvodu selhání řídicových smyslových schopností. V oboru bezpečnostních systémů je detekce očí popř. mrkání vhodná jako kontrolní příznak v biometrických systémech. Z lékařsko-technického hlediska je velice zajímavým nápadem využít mrkání k interakci postižený člověk a počítač. Pro všechny tyto zmíněné výzkumy je stěžejní smysl aplikace, tudíž zde není snaha o vylepšení již hotových metod, ale spíše o jejich přizpůsobení daným aplikacím.

1.2.1 Havarijní řídicí systémy

Statistiky ukazují, že více než 60% nehod motorových vozidel, je způsobeno řídicovou otupelostí³ nebo ospalostí [116]. Dnešní bezpečnostní systémy jako napří-

³ Výrazem je míněna řídicí psychická/fyzická vyčerpanost, projevující se únavou.

klad airbagy, bezpečnostní pásy, kšíry a další pomáhají předejít ztrátě života když se stane dopravní nehoda, nedokážou však předejít nehodě samotné. Práce [55] popisuje kompletní návrh havarijního řídicího systému, který je schopen upozornit řidiče, že jeho ospalost přerostla přes únosnou hranici. V tomto systému je řidičova bdělost průběžně odhadována s využitím technik na sledování očí [21]. Pomocí Sobelovy masky jsou určeny okrajové body očních víček, které jsou extrapolované ve dvě parabolické funkce pro každé oko zvlášť; jedno pro vrchní a druhé pro spodní oční víčko. Hlavní nedostatkem existujících systémů, které využívají měření šířky oka, je nemožnost odhadu střední části oka. Takové systémy signalizují pouze skutečnost, že oko je otevřené nebo zavřené [38], nerozpoznají však, zdali je oko zavřené z důvodu spontánního mrkání nebo kvůli ospalosti. Navrhnutý systém překonává tento nedostatek aktivním monitorováním řidičových očí přes určité časové rozpětí. Iterační algoritmus odhaduje prahovou hodnotu šířky řidičova oka podle ustavičné odezvy ze senzorů. Když naměřené hodnoty dosahují nižší hodnoty než stanovený práh po určité době, systém vyšle potvrzující informaci o řidičově únavě. Havarijní řídicí systém spustí brzdný systém motorového vozidla. Současně je pomocí namontovaného elektromagnetického ventilu odpojen přívod paliva z palivové nádrže. Zároveň se deaktivuje plynový pedál, i kdyby se ho řidič snažil sešlápnout.

1.2.2 Detekce živosti v biometrických systémech

Biometrie je technologie, která rozpoznává totožnost člověka na základě skutečných fyziologických vlastností nebo charakteristických rysů chování, například tváře, otisku prstů, duhovky, hlase apod. [4]. Nicméně možnost napodobení takových vlastností je stále velkou hrozbou pro spolehlivost některých biometrických systémů [40]. Pro zvýšení spolehlivosti se zavádí *detekce živosti*, která zajistí, že biometrický systém získává a zpracovává data (charakteristické rysy) pouze od živého člověka, čímž se zabrání možnosti využít napodobeninu biometrického rysu k oklamání systému. Tyto detektory se stále více rozšiřují v oblastech rozpoznávání otisků prstů nebo duhovky [40], [41], [30], [44].

Ve vědeckých oborech, zabývajících se rozpoznáváním obličejů, bylo představeno mnoho metod, ale snaha o zabezpečení proti napodobení je stále velmi malá [46]. Nejběžnější způsob padělání je použití fotografie obličeje oprávněného uživatele k podvedení rozpoznávacího systému – v dnešní době není pro širokou veřejnost problém stáhnout z internetu alespoň jednu fotografii téměř kohokoliv, popř. si tajně dotyčnou osobu vyfotografovat. Útok fotografií je jedním z nejlevnějších a nejnadanějších způsobů ošálení biometrického systému. Podvodník může před kamerou při prověřování identity s fotkou platného uživatele točit, pohybovat a ohýbat jako živá osoba. Pro detekční systém je velice těžké určit, zdali vstupní obličej patří živé osobě nebo jen fotografii.

Mnoho současných systémů pro rozpoznávání obličejů, je založeno na zpracování monochromatických obrazů a vybaveno obyčejnou kamerou. Vývoj algoritmů

pro odhalování podvodníků musí tyto parametry respektovat. Pro člověka je velice jednoduché rozlišit živou tvář od fotografie, protože si může snadno všimnout fyziologických vodítek živosti, například změny výrazu obličeje, pohyb úst, natočení hlavy, těkání očí apod. Nicméně pro počítač je velmi obtížné si těchto záchytných bodů všimnout, dokonce i při laboratorních podmínkách. Ze statického pohledu je základní rozdíl v tom, že živá tvář je plně trojrozměrná, zatímco fotografie je dvojrozměrné rovinné struktury. Tato skutečnost je využita v metodě měření struktury pochybu pro získání informace o hloubce⁴ [56] tváře k rozlišení živé tváře od fotografie. Problémy nastávají, pokud je hlava uživatele v klidu a bez pohybu. Zpracování je navíc velmi citlivé na šum a světelné podmínky. Metoda se tím lehce může stát nespolehlivá.

Další významnou vlastností živé tváře ve srovnání s fotografií je neztuhlý výraz a možnost změny zjevu jako jsou pohyb ústy a změny výrazu (gesta). K přesnému a spolehlivému odhalení těchto rysů je však potřeba vysoce kvalitní vstupní materiál nebo spolupráce uživatele při detekci. V [43] je popsána aplikace optického toku vstupních dat k získání informace o pohybu obličeje pro detekci živosti, metoda je však zranitelná v případě napodobení pohybu s fotografií jako je změna hloubky nebo ohýbání. Některé výzkumy využívají multi-modální metody obličej-hlas k odhalení podvodu [59], [25]. Během zpracování dat jsou sledovány rty dotyčného uživatele. Tento druh metod vyžaduje hlasový nahrávač a uživatelskou spolupráci. Interaktivní způsob funguje tak, že systém vyzívá uživatele k vykonání zřetelného pohybu hlavou [61].

Se zcela jiným přístupem přichází metody, které se nesnaží ověřit živou tvář, ale naopak zdiskreditovat fotografii. Pomocí Fourierova spektra lze odlišit živé tváře od padělaných obrazů za předpokladu, že fotografie obsahuje více vyšších frekvencí než nasnímaný obraz živé tváře systémem [65]. S přídatným zařízením lze pomocí obličejového termogramu [66] získat mapu žil obličeje při blízkém infračerveném zobrazení.

tab. 1.1: Srovnání přístupů pro odhalování podvodu při rozpoznávání obličejů [26]

Přístup	Kvalita dat	Přídavné zařízení	Spolupráce uživatele
Výraz obličeje	vysoká	ne	střední
Informace o hloubce	vysoká	ne	nízká
Pohyb úst	střední	ne	střední
Pohyb hlavy	vysoká	ne	střední
Mrkání očí	nízká	ne	nízká
Degradace	vysoká	ne	nízká
Víceúčelový	-	ano	střední/vysoká
Obličejový termogram	-	ano	nízká

⁴ Hloubkou je míněn třetí rozměr 3D prostoru (délka, šířka, hloubka)

Obličejová mapa žil	-	ano	střední
Interaktivní odpověď	-	ano	vysoká

Přehled přístupů pro odhalování podvodu při rozpoznávání obličejů je shrnut v tab. 1.1. Z té je patrné, že přístup využívající mrkání očí nevyžaduje vysokou kvalitu vstupních dat, přídavné zařízení a náročnou spolupráci uživatele.

Vysvětlení pojmu mrkání očí a dalších souvislostí je součástí kapitoly 3. Jak je vysvětleno v kapitole 3.1, průměrná délka mrknutí trvá asi 250 ms. Současné obyčejné kamery můžou snadno zachytávat tvář ve formě videosekvencí s frekvencí ne méně než 15 snímků za sekundu. To je asi 70 ms na jeden snímek a to znamená, že na jedno mrknutí může připadnout kolem 4 snímků. Pro takovou kameru je potom snadné zachytit dva nebo více snímků mrknutí, když se uživatel dívá obličejem přímo do kamery. Mrknutí očí lze tedy považovat za přijatelný přístup při odhalování falešných ověřování. Výhodami metod detekce živosti založených na mrkání očí jsou:

- Snímání biometrických rysů může být doprovázeno ne příliš dotěrnou kontrolou živosti, obvykle bez uživatelské spolupráce.
- Žádné další přídavné zařízení není nutné.
- Mrkání je význačným rozlišovacím znakem živé tváře od fotografie, což je výhodné, pokud máme obyčejné snímací zařízení.

1.2.3 Interakce člověk-počítač

Poslední léta dochází také k rozšiřování tradičních rozhraní člověk-počítač, jako jsou klávesnice a myš s inteligentním rozhraním, které umožňují uživatelům komunikace s počítačem mnohem pohodlněji a efektivněji. Cílem je vyvinout systém počítačového vidění, který umožní počítačům vnímat uživatelské přirozené komunikační vjemy, jako jsou gesta, výrazy tváře, směr pohledu. Takové systémy jsou obzvlášť významné pro lidi, kteří nemohou používat klávesnici nebo myš kvůli svému postižení.

Klasické rozhraní člověk-počítač vyžadují od uživatelů dobrou manuální zručnost a přesné motorické ovládání. Pro lidi s určitým postižením to znamená nepřekonatelné problémy. Proto vznikla snaha o návrh alternativních komunikačních prostředků pro lidi, jejichž pohybové schopnosti jsou značně omezeny počínaje pacienty s dětskou mozkovou obrnou, mozkovým traumatem a konče roztroušenou nebo amyotrofickou laterální sklerózou [92]. *Kompenzační technologie*, umožňující přístup k informacím a rozšířenou komunikaci, je pro postižené občany praktické i posilující zároveň.

Mnoho těchto lidí s mozkovým postižením je tak motoricky limitováno, že to jediné, čeho jsou schopni, je mrkat. K tomu, aby měli možnost používat různá zařízení a provádět osobní komunikaci „tváří v tvář“, využívají právě mrkání [23]. Nejjedno-

duším příkladem je elementární komunikace, kde jedno mrknutí znamená „ano“ a dvě „ne“. Existují dva typy komunikačních systémů:

- tradiční systém – pomocí mřížky s písmeny se za účasti asistence komunikačního partnera vytváří slova popř. celé věty;
- moderní systém – přímo vykonává uživatelské příkazy, na základě způsobu, kterým mrká.

Protože mrkat se dá různými způsoby, navrhnout obecné techniky detekce mrkání je velmi obtížné. Komunikační systém je proto obvykle navržen tak, aby se přizpůsobil každému uživateli zvláště. Obvykle je mrkání rozděleno do dvou skupin:

- typ 1 – pacient dobře ovládá obě oční víčka a může mrkat oběma očima i každým okem zvláště, někdy doprovázeno mírným zvednutím obočí, žádný jiný další pohyb
- typ 2 – pacient má kontrolu pouze nad jedním víčkem, zavírání je doprovázeno lícním pohybem s/nebo pohybem obočí.

Detekce mrkání pro řízení okolí jedince je komplikovaný úkol a mnoho technologických postupů se ho snažilo vyřešit, ne však s výrazným výsledkem. Jedna z nejčastěji prezentovaných metod pro lidi s poškozením mozkového kmene je metoda, založená na údajích z Elektromyografie⁵ (EMG) [93], které se získávají využitím tří malých elektrod přilepených na kůži páskou s mikropóry kolem svalu obicularis oculi muscle [91]. Tento sval obalí oko a je využit k zavírání očí. Ačkoliv z principu by systém založený na EMG měl být schopen efektivně detekovat mrkání, realita je jiná. Důvodem je elektrický šum v pacientově okolí a různé faktory související s elektrodami a kvalitou EMG signálu. EMG je rovněž zodpovědná za zvýšení poměru signál/šum, z důvodu změn vodivosti kůže a mírnému pochybu elektrod. Kůže pacientů s poškozením mozkového kmene má tendenci být mastná, což způsobuje změny vodivosti kůže. Pravděpodobnost fungování EMG jednotky po posunutí elektrod je velmi malá.

Nedávný pokrok v této problematice přinesl nový koncept, využívající webovou kameru a první testy ukazují slibné výsledky pro lidi, kteří nemají úspěch s řízením a komunikací pomocí jiných moderních řešení [69]. Metody využívající webové kamery mají oproti metodám s EMG výhodu v menším počtu přednastavení. Dále webová kamera má pevné umístění a je vzdálena od pacienta, kterého nemusí obtěžovat a není s ním nikterak fyzicky spojena. Kamera je navíc levné a dostupné zařízení. Další výhodou je, že když se pacient pohne, například z důvodu křečů, kamera vyrovná snímací cíl.

K tomu, abychom rozpoznaly příkazy, založené na mrkání uživatele, byl vytvořen následující koncept. Prvním úkolem při rozpoznávání příkazů založených na mrká-

⁵ Elektromyografie studuje funkci svalů tím, že vyšetřuje elektrické biosignály, které ze svalů vycházejí.

ní uživatele po zachycení scény kamerou je detekce vizuálních změn způsobených pohybem obličeje. K tomu, aby byla lokalizována oblast změny v obraze, kde by se mělo provést následné zpracování, je důležité před tím detekovat obličej. Druhým úkolem je analyzovat detekovanou změnu tak, aby se zjistilo, zdali obsahuje informaci o pohybu očních víček. Pokud obsahuje, potom z uchované historie změn pohybu očních víček je zjištěno, jestli zavírání nebo otvírání očí bylo spontánní nebo úmyslné. Pokud bylo úmyslné, dojde k rozpoznání příkazu. Co je však nejdůležitější, tyto úkoly musí být provedeny tak rychle, aby detekce mrkání probíhala v reálném čase. Konkrétně veškeré zpracování videosekvence by mělo trvat nanejvýš 100 ms. Jinak, jestliže uživatel vědomě zpomalí pohyby mrkání, nemusí být tyto informace vůbec zachyceny.

Obvyklé metody detekce pohybu ve videosekvencích jsou založeny na detekci změny úrovně jasu (velikost intenzity), způsobené pohybem. Nejotevřenější způsob detekce takové změny je binárně prahovaný rozdíl po sobě jdoucích snímků⁶ [85], [57], [58]. Tato technika nicméně při detekci selhává, pokud dojde k pohybu hlavy, protože se objeví mnoho prvků kolem hranice obličeje, jako jsou nos, ústa a další části. Vzhledem k tomu, že hlava je zřídka kdy v klidové poloze v [69] jsou prezentovány techniky detekce, nazývané jako „změny druhého řádu“, kde je rozpoznán pohyb hlavy od pohybu očních víček. Využívají se tři po sobě jdoucí snímky. Podmínkou je zachytávání více snímků, než je obvyklé (asi 50 snímků za sekundu). Hlava je považována za rovinný objekt. Výsledky experimentů ukazují, že mrkání je dobře zjistitelné, pokud intraokulární vzdálenost v obraze je asi 20 pixelů. Proto, při pozorování pacientovy tváře ze vzdálenosti asi 30 cm (obličej-obrazovka), což je nejčastější případ, stačí snímky o velikosti 160×120 pixelů. Z toho důvodu neplynou žádné výhody, pokud by se zpracovával obraz s vyšším rozlišením.

⁶ Upravená základní metoda ve zpracování videosekvencí – rozdíl snímků.

2 Cíle disertační práce

Na základě rozboru stavu současně problematiky uvedené v kapitole 1 byly formulovány cíle této disertační práce:

- Analyzovat problematiku detekce a rozpoznávání mrkání, jaké jevy lze měřit, popsat; jakým způsobem navrhnout systém pro detekci mrkání, v jakých aplikacích je využít.
- Prostudovat dostupné obrazové databáze a nalézt vhodné kandidáty na testování systémů pro detekci mrkání; v případě neúspěchu, vyvinout vlastní databázi.
- Vyvinout algoritmus pro detekci mrkání s využitím číslicového zpracování obrazu a zhodnotit jeho přínos.
- Zhodnotit význam zájmu o problematiku detekce a rozpoznávání mrkání.

3 Analýza problematiky detekce a rozpoznávání mrkání

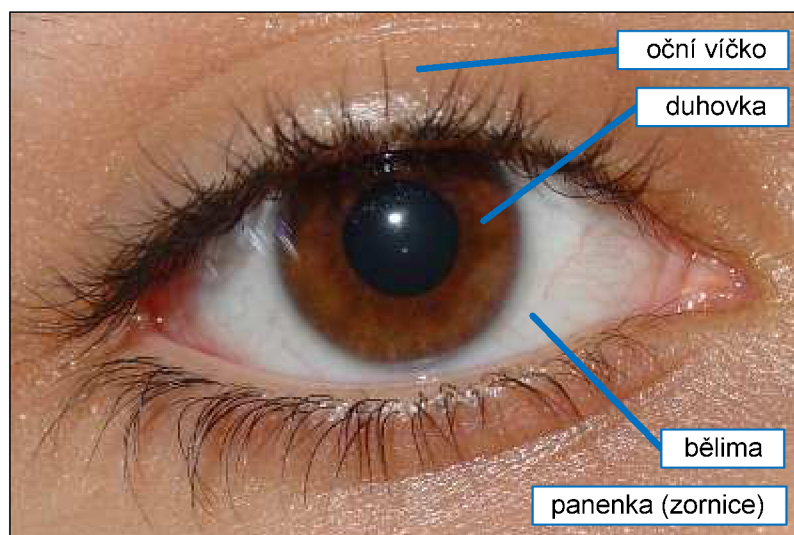
V této kapitole je vysvětleno, co je to mrkání očí, jakými cestami lze postupovat v problematice detekce a rozpoznávání mrkání a jakým způsobem navrhnout systém pro automatickou detekci.

3.1 Mrkání očí

Mrknutí lze definovat jako rychlé zavírání a otevírání očních víček. Je to důležitá funkce očí, která napomáhá k šíření slzní tekutiny po povrchu oční rohovky a spojivek a tím odstraňuje různé nežádoucí částičky, které způsobují podráždění očí. Jedno mrknutí obvykle trvá 200 až 400 ms [11], zornice je plně zakryta očním víčkem 100 až 150 ms (obr. 3.2) [82]. Pokud nepočítáme pláč, je několik důvodů, proč lidé mrkají. Hlavní důvod je čištění a zvlhčení oční tkáně [45]. Jelikož se jedná obvykle o nevědomý děj, proces mrknutí není myslí často vůbec zaznamenáno. Člověk průměrně mrkne (*frekvence mrkání*) každých dvě až deset sekund. Ve speciálních případech může být mrkání ovlivněno faktory jako vyčerpání, oční zranění, vliv léků nebo nemoc. V neposlední řadě je třeba brát v úvahu fakt, že mrkání lze ovlivnit vůlí, což může znemožnit správné výsledky sledování a vyhodnocení mrkání. Taková osoba se na své mrkání soustředí a poruší nenucenost procesu mrkání. Mrkání lze tedy rozdělit do třech skupin [8]:

- spontánní mrkání – nejčastější, nenucené, vůlí téměř nezaznamenáno,
- vědomé mrkání – ovládané vůlí,
- reflexní mrkání – přivozeno nějakým impulsem (světlo, zvuk, strach, nemoc apod.).

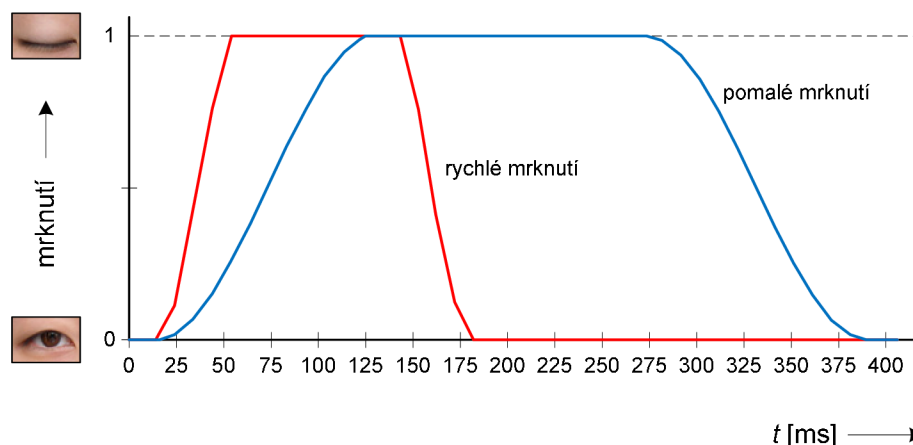
Se všemi typy mrkání se lze v praktických úlohách setkat: bezpečnostní systémy (vědomé), havarijní systémy (spontánní a reflexní) a lékařské systémy (reflexní a vědomé).



obr. 3.1: Popis důležitých částí oka.

Při porovnání mrkání u mladých a starších lidí, lze vypořádat určité rozdílnosti. Malé děti nemrkají se stejnou frekvencí jako dospělí. Ve skutečnosti mrknou jednou až dvakrát za minutu. Vysvětlení pro tento úkaz není zatím zcela znám, předpokládá se však, že děti nepotřebují tak velké množství zvlhčující látky jako dospělí, neboť jejich oční víčka jsou menší v porovnání s dospělými. Nemluvnata navíc neprodukuje žádné slzy první měsíc života. Těm se také dopravuje významně více spánku než dospělým a jak bylo výše zmíněno, unavené oči mrkají častěji. Nicméně postupným vývojem se frekvence mrkání zvyšuje a v době dospívání se téměř vyrovná dospělým [70].

Co se týče rozdílů mezi muži a ženami, nebyly zaznamenány žádné podstatné rozdíly [71]. V laboratorních podmínkách lidé mrknou asi desetkrát za minutu. Pokud se však člověk soustředí na nějaký objekt, například při čtení, frekvence mrkání klesne na 3 až 4 mrknutí za minutu. Což je hlavní příčinou toho, proč oči vysychají a stávají se unavenými při čtení.



obr. 3.2: Modelový průběh rychlého a pomalého mrknutí.

Mrkání očí může být také měřítkem při diagnostice lékařských onemocnění [12]. Přílišné mrkání může například pomoci k odhalení počátků Tourettova syndromu, mrtvice, nebo poruchy nervového systému [92]. Snížení frekvence mrkání je zase spojováno s Parkinsonovou chorobou. Pacienti s touto chorobou mají výrazně strnulý výraz v očích, čehož si lze velice jednoduše všimnout.

3.2 Rozdělení problematiky podle přístupu

Podle toho, jaká zařízení a příslušenství máme k dispozici, lze přistupovat k detekci mrkání třemi způsoby:

- 1) Invazivní monitorování
- 2) Digitální zpracování obrazu
- 3) IR technologie

Každý přístup má své výhody i nevýhody v závislosti na dané aplikaci. Přístupy 2 a 3 se také nazývají souhrnně metody detekce mrkání založené na vidění [69].

3.2.1 Invazivní monitorování

Tento přístup nalézá uplatnění hlavně v lékařství, kde je pohyb očních víček sledován snímacím zařízením. Ze své podstaty lze tento přístup považovat za nejpřesnější, ne však nej pohodlnější či nejvhodnější. Část snímacího zařízení je v určitých případech třeba voperovat do očních orgánů (oko, oční víčko). Příkladem může být metoda, která využívá magnetické pole a tenký drátek voperovaný do očního víčka [85]. Jestliže pacient mrkne, dojde ke změně vlastností magnetického pole a tyto data jsou dále zpracována.

3.2.2 Digitální zpracování obrazu

Přestože procesy detekce mrkání popř. sledování očí vyžadují i časový rozměr, tedy přesněji sekvenci snímků (lze se setkat i s výrazem pohyblivé snímky nebo jednoduše video⁷), většina metod pro zpracování pohyblivého obrazu vychází nebo je shodná se základními a pokročilými technikami digitálního zpracování obrazu. Hlavní myšlenkou tohoto přístupu je hledání kompromisu mezi počtem a komplexností použitých metod a rychlostí výpočtu daného algoritmu. Jinak řečeno (automatický) systém pro detekci mrkání musí být tak robustní, jak je pro danou aplikaci z hlediska času únosné. Tento předpoklad je samozřejmě platný pro většinu systémů pro automatické zpracování signálů obecně. Výhodou je však skutečnost, že zdrojem pro zpracování obrazu může být jakýkoliv dostupný obrazový materiál, popř. materiál nový, pořízený obyčejnou kamerou.

3.2.3 IR technologie

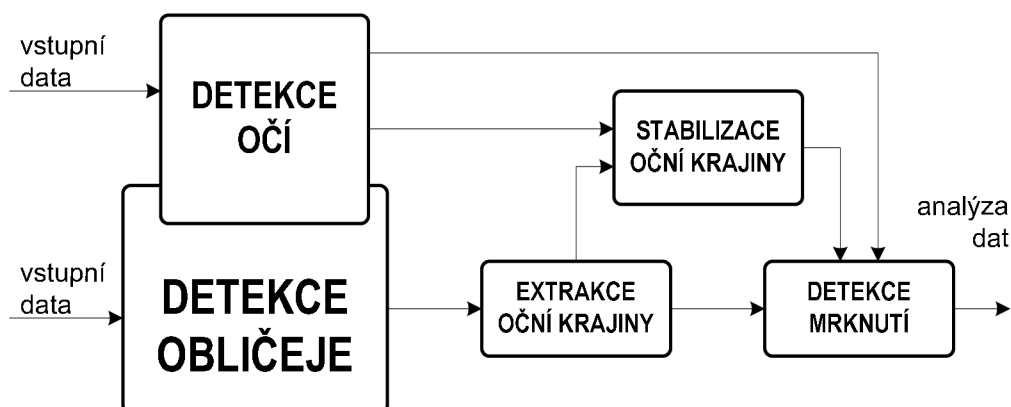
Posledním přístupem lze uvažovat kategorii metod, využívající IR technologii. Na rozdíl od téměř drastického invazivního monitorování, metody používající přisvětlení infračerveným zářením, využívají odrazových vlastností oční duhovky k odhadu mrknutí [73], [77], [80], [81]. Dopadající infračervené světlo se od oční duhovky odrazí tak mohutně, že snímající kamera zaznamená v oblasti oka výrazný bílý kruh. Zavedením jednoduchého prahování, můžeme tyto kruhy oddělit a sledovat jejich posuv. K mrknutí pak dochází, když je duhovka zatemněna očním víčkem a IR detektor nezaznamená odraz. Tímto získáváme mocný nástroj pro návrh systému detekující mrkání. Takový systém vyžaduje speciální vybavení jako LED diody emitující IR záření a kameru snímající i IR spektrum.

Tuto kategorii metod lze určitým způsobem považovat za shodnou s druhým přístupem. Proces, kdy díky IR technologii získáme velice kvalitní materiál, můžeme považovat za preemfázi následujícího zpracování. Tím je míněno opět číslicové zpracování obrazu. Jako hlavní výhodu lze považovat skutečnost, že následné metody by měly být triviální na výpočet.

3.3 Analýza modelu postupu detekce mrkání

Vzhledem k velké složitosti problematiky je velice obtížné popsat jednoduchým způsobem postup zpracování při detekci mrkání. Různé přístupy a jejich metody, které mohou fungovat každá zcela jinak, komplikují návrh tohoto modelu. Přesto byl vytvořen univerzální model postupu detekce mrkání, který zohledňuje 2. i 3. přístup. Jelikož invazivní monitorování je přímá metoda detekce, která navíc nevyužívá zpracování obrazu, nebyla při návrhu brána v úvahu.

⁷ Z oblasti teorie televizní techniky vyplývá, že pojmem video je míněn pořízený záznam amatérskou/polo/profesionální kamerou (též i webkamerou), který splňuje jisté televizní normy [10].



obr. 3.3: Univerzální model postupu při detekci mrkání, využívající zpracování obrazu (druhé přiblížení)

3.3.1 Detekce obličeje

Již podle obr. 3.3 lze soudit, že procedura detekce obličeje bude nejsložitější částí (v případě, že bude v procesu zpracování zahrnuta). Lze ji definovat jako proces, jehož cílem je určit přítomnost nebo absenci obličejů v předloženém obraze. Pokud jsou v obraze obličeje, detekční systém má určit polohu a ohraničující oblast každého z nich. K detekci obličeje se váže několik blízce souvisejících pojmů, jejich význam je také vhodné vymezit. Lokalizace obličeje představuje zjednodušený problém detekce, kdy se předpokládá, že obraz obsahuje pouze jedinou tvář. Sledování obličeje (stabilizace oční krajiny) se zabývá konstantním zjišťováním polohy obličeje většinou v reálném čase. Rozpoznávání obličeje si klade za cíl poznat v obraze nějakého konkrétního jedince.

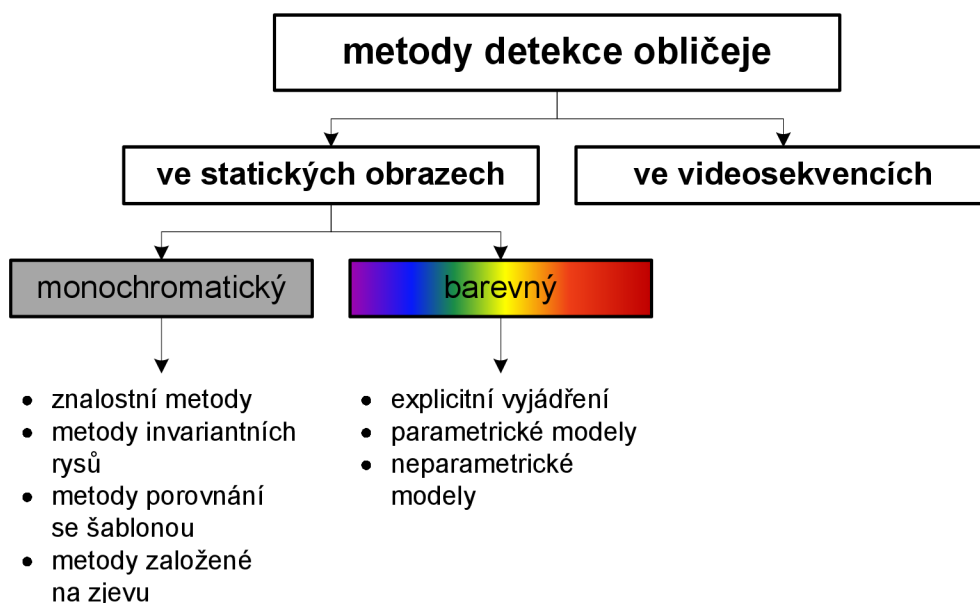
Na detekci obličeje lze nahlížet jako na problém klasifikace oblastí obrazu do dvou tříd: „tvář“ a „ostatní“. Oproti ostatním klasifikátorům se příslušníci třídy vyznačují velkou rozmanitostí, což klade na detektor vysoké nároky. Detekce obličeje také přináší zajímavý úkol pro s ní související metody učení a rozpoznávání vzorů, protože předložený obraz představuje obrovské množství vstupních dat a kombinačních možností [13]. Detekční systémy musí být proto schopny efektivně redukovat dimenzi prostoru vlastností vstupních dat.

Vzhledem k nepřehlednému množství metod používaných k detekci obličeje se jeví užitečné rozdělit tyto metody do několika tříd podle jejich základního přístupu.

Nejobecnějším kritériem dělení metod detekce je charakter vstupních dat (obr. 3.4). Detektory pracují v zásadě se třemi druhy

- jasový obraz ve stupních šedi (monochromatický obraz),
- barevný obraz,
- videosekvence.

Hlavním polem pro detekci obličeje je obraz ve stupních šedi. Detektory klasifikují obličej na základě tvaru, informace o barvě (barevný obraz) nebo pohybu (video) je ve většině případů využita pouze k redukci prohledávaného prostoru [31]. Přesto je detekce barvy kůže často využívanou součástí detektorů a intenzivně zkoumanou oblastí.



obr. 3.4: Schematické rozdělení metod detekce obličeje podle formátu vstupních dat

Metody znalostního základu

Tyto metody jsou založené na základě definovaných pravidel, které popisují „typickou tvář“. Pravidla většinou vyjadřují vztahy mezi částmi obličeje.

Metody invariantních rysů

Tyto metody jsou zaměřeny na nalezení základních strukturních prvků obličeje, které se neztrácejí při různém úhlu pohledu, různých světelných podmínkách nebo pozici obličeje.

Metody porovnání se šablonou

Šablona představuje celý obličej případně pouze některou jeho část. Tato šablona se pomocí korelace porovná se vstupním obrazem. Metody se používají, jak pro lokalizaci, tak i detekci obličeje a očí.

Metody založená na zjevu

Modely obličeje jsou získány učením z trénovací množiny, která obsahuje různé vzory tváří. Detekce probíhá srovnáváním částí obrazu s tímto modelem.

3.3.2 Detekce očí

Jelikož metody detekce obličeje jsou velkým inspiračním zdrojem pro další zpracování, procedura detekce očí se s detekcí obličeje částečně prolíná (obr. 3.3). V závislosti na aplikaci lze detekovat střed oka, oční víčka či celou plochu oční krajiny.

Nejrychlejší a nejjednodušší metody detekce očí využívají segmentaci pomocí barvy kůže, [47], [52], [68], [79], [90]. Protože jsou oči většinu času otevřeny, procedura segmentace pomocí barvy kůže objeví v oblasti kandidáta na obličej dvě díry. Statistický model barvy kůže (obr. 4.10) se vytvoří pomocí trénovací sady obrazů s obličejí. Barevné složky těchto obrazů musí být normalizovány, tím se odstraní potíže s pestrostí jasu. Barva lidské kůže různých etnických skupin se téměř neliší barevným tónem (odstínem), ale je velmi odlišná, co se týče úrovní jasu. Pokud se jasová složka potlačí, je možné vytvořit specifický model dle barvy kůže, nezávislý na etniku. K odstranění jasové složky je třeba normalizovat barvy určitou hodnotou. Více o barevných modelech a detekce dle barvy kůže v kapitole 4.3.2

3.4 Cíle detekce a s ním spojené problémy

Cílem celé detekce je schopnost v libovolném snímku rozhodnout, zda se v něm oko resp. jiná výrazná část obličeje, s jejíž pomocí dosáhneme správné detekce, nachází. Pokud tomu tak je, dochází k navrácení lokací a rozměrů všech nalezených objektů. S tímto cílem se ovšem pojí nemalé problémy, které by se daly vypsát do následujících bodů [22]:

- **Pozice** – Pozice hlavy vůči kameře může způsobit částečné vytracení hledané části obličeje ze záběru a její velice obtížnou detekci.
- **Přítomnost jednotlivých strukturních prvků obličeje** – Obličejové znaky jako například vousy, knírek, brýle nebo make-up mohou, ale nemusí být v obraze přítomny. To dává velký prostor rozmanitosti těchto prvků zahrnující tvar, barvu, velikost apod.
- **Výraz obličeje** – Detekce je vždy velice ztížena, pokud má snímaný člověk nestandardní výraz v obličejí, například pokud se mračí.
- **Jiné objekty v okolí obličeje** – Například v obrázku s více lidmi může být obličej obklopen dalšími obličejí.
- **Orientace obrazu, vzdálenost snímaného cíle** – Obraz obličeje a tedy i očí se mění s natočením kamery, resp. natočením obličeje, se vzdáleností snímaného člověka od kamery.
- **Podmínky při snímání** – Všechny podmínky, jako je například osvětlení nebo vlastnosti a kvalita kamery, se projeví také ve snímku a v obtížnosti následného zpracování.

- **Typ vstupního obrazu** – Na šedotónovém obraze se provádí jiné operace detekce než na obraze barevném.

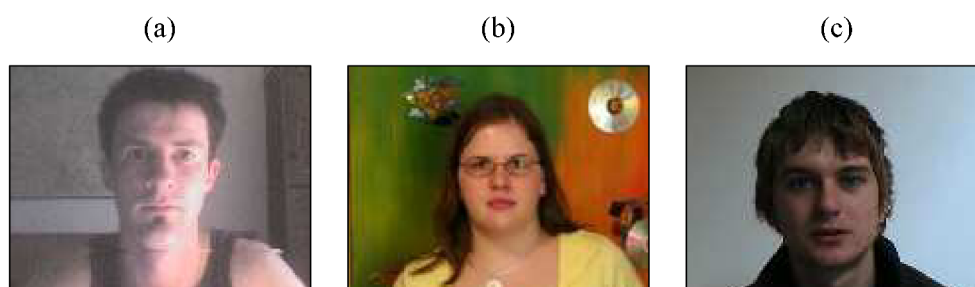
3.5 Systém pro rozpoznávání mrkání očí

Při návrhu takového systému je třeba uvážit několik věcí. Zaprvé je důležité brát v úvahu kvalitu obrazových dat (video data). S tím souvisí i velikost rozlišení obrazu. Podle kvality těchto dat, lze zaznamenat akci mrknutí (obr. 3.6a) nebo průběh mrkání (funkci mrkání), viz obr. 3.6b.

3.5.1 Kvalita obrazových dat

Kvalitu obrazových dat lze rozdělit podle zdroje snímání do tří úrovní kvality:

- nízká – obyčejné snímače, webkamery, levné mobilní telefony (obr. 3.5a)
- střední – kvalitní webkamery, poloprofesionální kamery, drahé mobilní telefony a digitální fotoaparáty (obr. 3.5b)
- vysoká – profesionální kamery (obr. 3.5c)



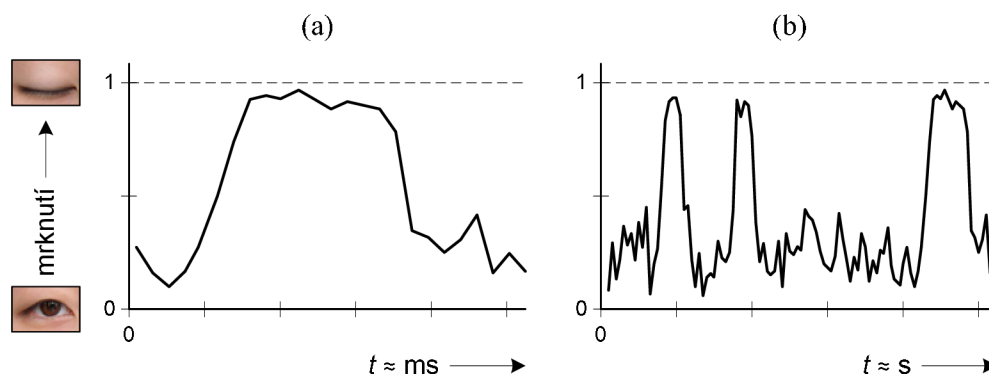
obr. 3.5: Příklady kvality obrazových dat.

Rozdělení podle velikosti rozlišení obrazu je komplikovanější. Není žádná norma definující dostatečnou velikost. U různých metod je experimentálně zjišťováno dostatečné rozlišení pro optimální zpracování. Jediným vodítkem jsou normy velikostí snímacích a zobrazovacích zařízení (webkamery, kamery, digitální fotoaparáty, televize, monitory, LCD obrazovky apod.) [9].

Poměr velikosti obličeje ke zbytku obrazu je rovněž těžko uchopitelný problém. Opět záleží na robustnosti algoritmů a měřené události mrkání očí. Podle porovnání různých metod lze usoudit, že v obdélníkovém obraze (obvykle 4:5 nebo 3:4) by měl obličej zaujímat alespoň 1/5 popř. 1/4 obrazu.

Systémy založené na procesu mrknutí také zahrnují biometrii segmentu očí, analýzu nebo syntézu očí (modely) [27], [75]. Kvalita souboru dat by měla být ve vyšším rozlišení nebo aspoň tak kvalitní, aby bylo možné vytvořit model segmentu očí (oblast kolem oční krajiny).

3.5.2 Měřené události mrkání očí



obr. 3.6: Příklady měřených událostí mrkání očí.

Nelze vždy získat nebo pracovat s kvalitním obrazovým materiálem. V určitém případě to však není úplně nutné. Pro odhad frekvence mrkání není potřeba požadavek na vysokou kvalitu zpracovávaného materiálu dodržet. Na druhé straně to znamená, že navržený automatický systém detekce mrkání musí být velmi robustně sestaven. Procedura odhadu frekvence mrkání tak může být velice výpočetně náročná.

Frekvence mrkání

Jestliže mrknutím je definováno rychlé otevírání a zavírání očních víček, pak frekvence mrkání reprezentuje poměr počtu mrknutí za určitý časový interval. Pro odhad frekvence mrkání z měření slouží rovnice

$$r_b = \frac{b}{t} \text{ [mrknutí/s] = [Hz] ,} \quad (3.1)$$

b je počet mrknutí, t doba měření a r_b je frekvence mrkání. Pro účely digitálního zpracování obrazu, při práci s videosekvencemi je více užitečný přepis předchozí rovnice na tvar

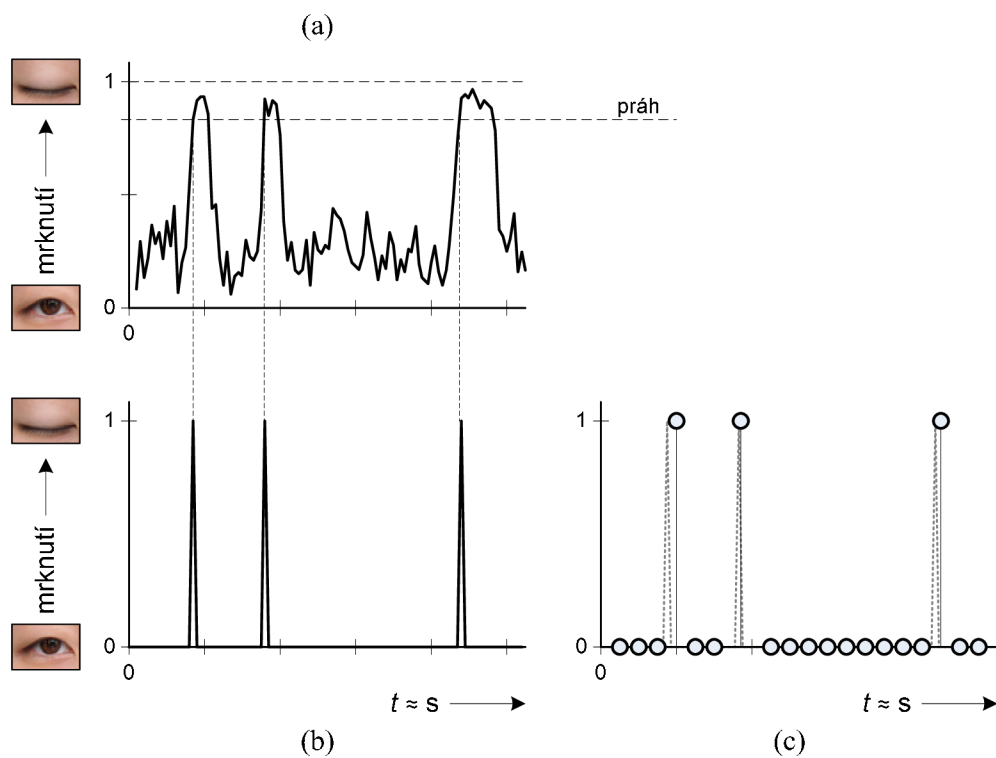
$$r_b = 60 \cdot b \cdot \frac{f_r}{f} \text{ [mrknutí/min] ,} \quad (3.2)$$

kde f_r je snímková frekvence a f počet snímků.

Průměrná frekvence mrkání pro celou populaci byla odhadnuta na 26,13 mrknutí za minutu. Nicméně tato hodnota závisí na okolnostech, jako jsou prostředí a stav jedinců, jak již bylo uvedeno výše. Například lidé při konverzaci mají průměrně 20,9 mrknutí za minutu a při čtení dokonce jen 16,4 mrknutí za minutu [76].

Frekvenci mrkání lze měřením získat z průběhu mrkání (obr. 3.7a), které nemuselo být zcela přesně zaznamenáno. Pomocí nastaveného prahu, který lze získat pouze experimentálním způsobem vzhledem k dané aplikaci a algoritmu, odhadneme okamžiky mrknutí (obr. 3.7b). Výsledek lze uložit to tzv. *binární funkce*, což je

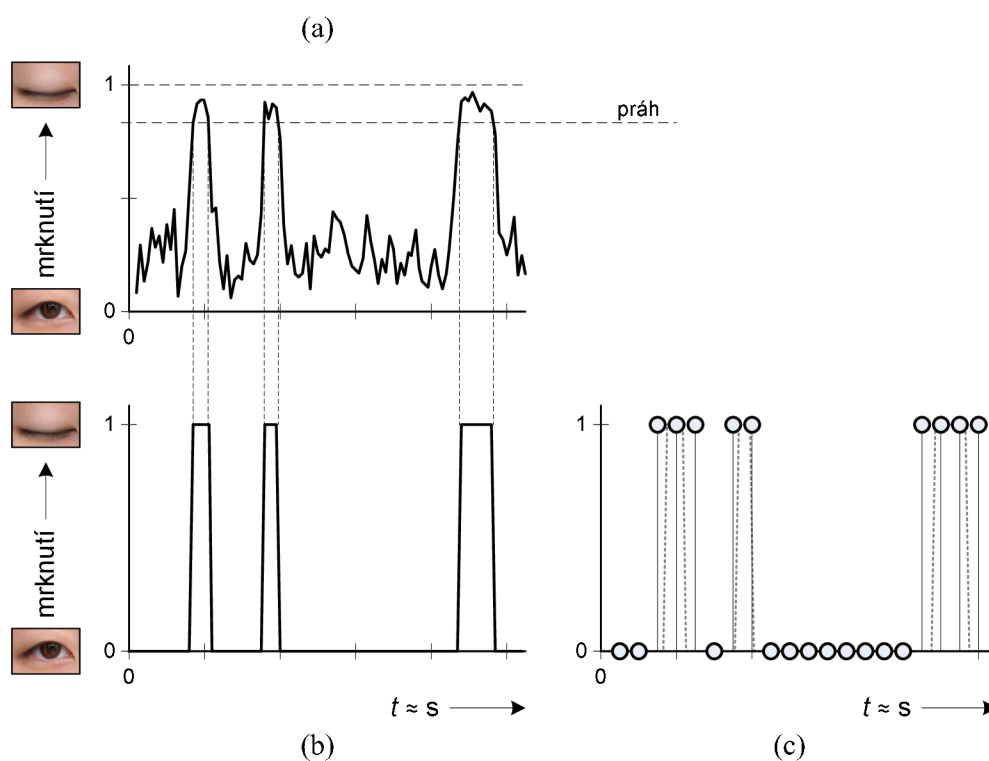
funkce, která nabývá hodnoty 1 pro pozitivní jev mrknutí a hodnoty 0 pro negativní jev mrknutí (obr. 3.7c). Tuto funkci lze v případě nutnosti dále vyhodnocovat nebo jednoduše zjistit frekvenci mrkání pomocí (3.2).



obr. 3.7: Postup získání frekvence mrkání.

Průběh mrkání

Nejdůležitější a z principu nejzákladnější událost, kterou lze naměřit je průběh mrkání (obr. 3.6b). Čím kvalitnější zdroj obrazových dat a robustnost navrženého systému, tím lépe lze průběh mrkání dále analyzovat. Postup vypadá podobně jako v případě měření frekvence mrkání, z výsledku průběhu mrkání lze však vyvodit mnohem více úsudků. Pečlivě zaznamenaný průběh mrkání (obr. 3.8a) lze samotný využít k dalšímu vyhodnocení nebo získat jako v předešlém případě binární funkci (obr. 3.6c), pomocí prahování a normování (obr. 3.6b). V této funkci lze sledovat celé doby trvání akcí mrknutí, ne jen samotnou informaci, že došlo k mrknutí.

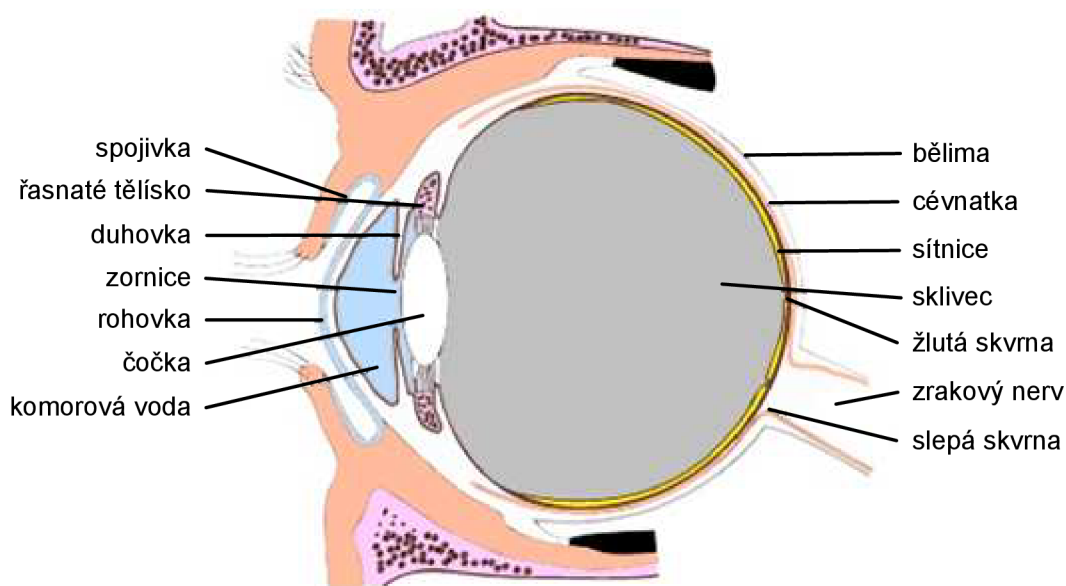


obr. 3.8: Příklad analýzy pečlivě zaznamenaného průběhu mrkání.

3.6 Lidské oko a záření

3.6.1 Oko a zrak

Zrak je pro člověka nejdůležitějším smyslem. Asi 80% všech informací z okolí získáváme prostřednictvím zraku. Orgánem zraku je oko, reagující na viditelné světlo, část spektra elektromagnetického záření - to se v oku transformuje v nervové signály. Příjmu a zpracování vizuálních informací se účastní v každém oku více než 100 milionů receptorových buněk, tyčinek a čípků v sítnici. Tyto buňky jsou s mozkem spojeny 1 600 000 nervovými vlákny. Již z těchto čísel lze usuzovat, že proces zrakového vnímání je mimořádně náročný [89].



obr. 3.9: Řez okem s popisem jednotlivých částí

Oko je uloženo v očníci a je složeno z šesti základních částí.

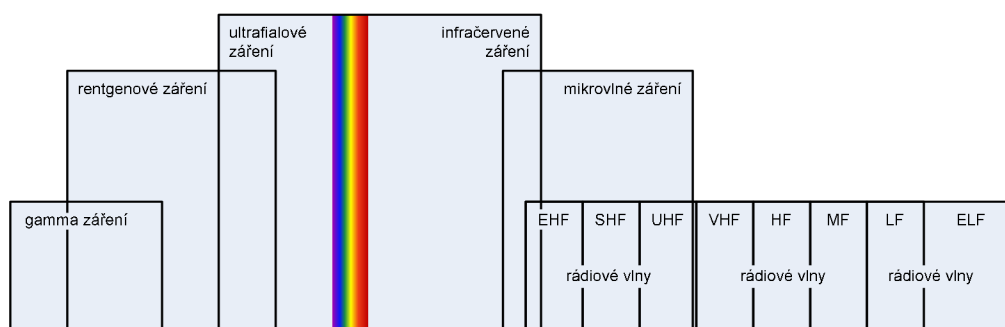
- Bělima je tvořena vazivovou blánou, udržuje tvar oka a pozorujeme ji jako bílý obal oka. V přední části přechází v průhlednou rohovku.
- Cévnatka tvoří vnitřní vrstvu oční koule. Jejím úkolem je zásobovat zevní vrstvy oka krví a zabránit rozptylu světla uvnitř oka za pomoci pigmentových buněk. Vpředu přechází cévnatka v řasnaté tělísko.
- Duhovka má tvar kruhového terčíku s kruhovým otvorem uprostřed - zornicí. Je tvořena hladkým svalstvem. Na povrchu duhovky se nacházejí pigmentové buňky, které dávají očím barvu.
- Čočka je zavěšena na vazivových vláknech vycházejících z řasnatého tělíska. Je tvořena rosolovitou dokonale průhlednou hmotou.
- Sklivec, průhledná rosolovitá hmota, tvoří vnitřní výplň většiny prostoru oční koule
- Sítnice, nejvnitřnější vrstva oční koule, tvoří vlastní světločivný systém oka. Pokrývá zadní 2/3 oka, kromě místa, kterým z oční koule vychází zrakový nerv. Toto místo se nazývá slepá skvrna.

Světlolomný systém oka

Světelné paprsky přicházející do oka procházejí nejprve rohovkou do přední komory vyplněné komorovou vodou a dále čočkou a sklivcem. Pomocí rohovky a čočky jsou přitom světelné paprsky soustředěny na sítnici. Obraz, jenž se na sítnici promítá, je zmenšený a obrácený obraz pozorovaného předmětu [89].

3.6.2 Elektromagnetické záření

Světelný paprsek je postupná vlna tvořená elektrickým a magnetickým polem. S tímto zjištěním přišel v roce 1860 James Clerk Maxwell a spojil tak dvě odvětví vědy - optiku a elektromagnetismus. Jeho teorii poté rozšířil o další druhy elektromagnetického záření Heinrich Hertz. Jejich zásluhou dnes známe pojem Maxwellovy duhy, tedy širokého spektra elektromagnetického záření [94].



obr. 3.10: Zjednodušené spektrum elektromagnetického záření

Infračervené záření, jehož vlastností se využívá v této práci, je právě součástí rozsáhlého spektra. Různé části tohoto spektra se liší různými hodnotami určité proměnné veličiny. Tato veličina se nazývá vlnová délka. S její změnou ve viditelném spektru mění světlo svou barvu od červené k fialové. Pokud budeme zkoumat spektrum systematicky od dlouhých vlnových délek ke kratším, začneme dlouhými vlnami. Následují vlny rádiové (rozhlasové). Běžné vysílače používají vlnovou délku okolo 500 metrů, k rádiovým vlnám ovšem patří také krátké vlny, radarové, poté milimetrové atd. Skutečné hranice mezi jednotlivými rozsahy neexistují a nejsou tak určena žádná ostrá rozhraní. Daleko pod milimetrovými vlnami najdeme infračervené vlny, poté viditelné světlo a dále oblast, které říkáme ultrafialová. Pod ní se nacházejí rentgenové paprsky, které se dále dělí na měkké, normální a velmi tvrdé. A na konci se nachází gama záření. Samozřejmě by bylo možné pokračovat dál ke stále menším hodnotám vlnové délky. [94]

3.6.3 Infračervené záření

Infračervené záření je elektromagnetické záření s vlnovou délkou delší než viditelné světlo a zároveň kratší než mikrovlny. Jeho vlnová délka je mezi 750 nm a 1 mm. Využití IR (infračerveného) záření a jeho snímání je velice mnoho jak v civilní tak i vojenské oblasti. V civilní je to například dálkové měření teploty, bezdrátová komunikace nebo předpovídání počasí. Ve vojenské oblasti je to pak zaměřování cíle, systémy nočního vidění a sledovací zařízení [83], [95].

Objekty (např. LED dioda, člověk) vyzářují infračervené záření ve velkém spektru vlnových délek, ale jen některé oblasti tohoto spektra jsou pro nás v určité chvíli

zajímavé. Proto rozdělila Mezinárodní komise pro osvětlování (CIE) IR záření do následujících tří pásem [83]:

- **IR-A:** 700-1400 nm

Blízké infračervené záření neboli **NIR** (Near-infrared) pásmo je absorbováno vodou a jeho hlavní využití najdeme v oblasti telekomunikací při použití optických kabelů. Právě toto pásmo se používá z důvodu nízkých ztrát na přenosovém médiu z SiO₂. K dalším příkladům patří systémy nočního vidění.

- **IR-B:** 1400-3000 nm

Krátké vlnové délky neboli **SWIR** (Short-wavelength) pásmo je také využíváno v telekomunikacích a nejvíce v rozsahu vlnových délek od 1530 nm do 1560 nm.

- **IR-C:** 3000 nm - 1 mm

Střední vlnové délky neboli **MWIR** (Mid-wavelength) pásmo (někdy označováno zkratkou **IIR**) se rozkládá na vlnových délkách od 3000 nm do 8000 nm. Využívá se hlavně ve vojenské technice například k tepelnému navádění raket a označování cílů.

Dlouhé vlnové délky neboli **LWIR** (Long-wavelength) pásmo je v intervalu 8000-15000 nm. Jeho vlastností využívají například termokamery, které dokážou vytvořit obraz bez jakýchkoliv externích přisvětlování.

Dlouhé infračervené záření neboli **FIR** (Far IR) zaujímá vlnové délky od 15000 nm do 1 mm a využívá se v armádní technice.

3.6.4 Nebezpečné účinky infračerveného záření

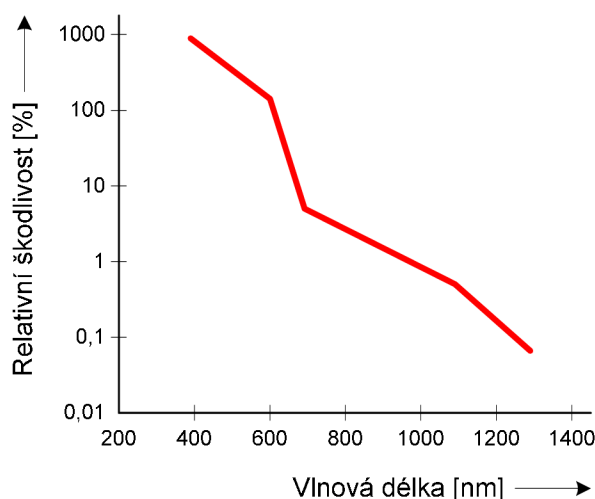
Lidé jsou již od dávné historie vystavováni vysokým dávkám IR záření. Ať už na silném denním světle nebo při zpracovávání oceli a skla, jednoduše všude tam, kde jsou v blízkosti velmi horkých objektů. Infračervené záření je nebezpečné kvůli teplotě, které sálá ve směru jeho působení. Tohoto jevu se například využívá v IR přímotopech. Jeho silné nebo dlouhodobé působení však může na kůži vytvořit popáleniny a puchýře. Mnohem hůř ovšem působí na oči [14].

3.6.5 Reakce oka na infračervené záření

Jasně denní světlo apod., obsahující pro oči škodlivé infračervené záření, způsobí, že začneme očima mrkat, stáhne se duhovka, tak aby propouštěla méně světla, a donutí nás odklonit hlavu. Efekt, kdy se stáhne duhovka, můžeme dobře pozorovat při focení s bleskem, kdy oko nedokáže tak rychle reagovat a ve vyfocené snímku se objeví silný odraz od zornice. Tento podmíněný způsob obrany proti IR záření je ale rázem pryč v případě, že odstraníme ze spektra právě viditelnou složku. Oko se přestane bránit a je v tu chvíli vystaveno záření, které může při rychlé a silné expozici přímo popálit sítnici a způsobit zákal. Při delším ale menším výkonu expozice

dochází k zahřívání komorové vody, které taktéž způsobuje velká poškození zrakového orgánu [83], [96].

Následující obr. 3.11 znázorňuje relativní škodlivost závislou na vlnové délce. Z tohoto grafu můžeme jasně vidět, že nejškodlivější vlnové délky jsou právě ty nejbližší viditelnému spektru. Těm se ovšem, jak již bylo popsáno výše, dokáže člověk bránit. Proto opravdu nebezpečné vlnové délky začínají až na 750 nm a výše, jejich škodlivost však logaritmicky klesá. IR LED diody použité v této práci jsou však bezpečné. Výkon vyzařování jedné LED diody je dle technické dokumentace 30 mW na vlnové délce 940 nm. Udávaná hodnota pro škodlivé působení IR záření začíná na 300 mW. Tato čísla jsou ale minimální v porovnání s IR lampami používanými ve zdravotnictví, kde jejich výkony dosahují i 500 W, těmi se ale přirozeně nesmí ozářovat oči a jejich expozici je člověk vystaven na velice malou chvíli [83].



obr. 3.11: Závislost relativní škodlivosti infračerveného záření na vlnové délce [83].

Shrnutí parametrů podílejících se na škodlivosti IR záření:

- Výkon zářiče, čím vyšší, tím škodlivější.
- Vlnová délka, čím vyšší, tím bezpečnější.
- Úhel vyzařování, čím vyšší tím, bezpečnější (při stálém výkonu)
- Vzdálenost pozorovatele od zářiče, čím vyšší, tím bezpečnější.
- Délka vystavení záření, čím vyšší, tím škodlivější.

4 Číslicové zpracování obrazu

Úkolem zpracování a případného rozpoznávání obrazu je obrazová informace o reálném světě, která do počítače vstupuje nejčastěji televizní či jinou kamerou. Cílem je porozumění obsahu obrazu. Postup zpracování a rozpoznávání obrazu reálného světa lze obvykle rozdělit do několika základních kroků:

- Snímání, digitalizace a uložení obrazu v počítači.
- Předzpracování obrazu.
- Segmentace a popis objektů.
- Porozumění obsahu obrazu, případně jen klasifikace objektů.

Prvním krokem zpracování obrazu je snímání, digitalizace a uložení obrazu v číselné formě do počítače. Při snímání se převádějí vstupní optické veličiny na elektrický signál spojený v čase i úrovni. Vstupní informací může být jas (z kamery, scanneru), nebo několik spektrálních složek (červená, zelená, modrá) při barevném snímání.

Digitalizací se převádí vstupní spojitý signál do diskrétního tvaru. Vstupní analogový signál je popsán funkcí $f(x, y)$ dvou proměnných – souřadnic v obraze. Funkční hodnota odpovídá např. jasu. Vstupní signál je vzorkován a kvantován. Výsledkem je matice čísel popisujících obraz. Jednomu prvku matice se říká obrazový element – pixel.

Druhým základním krokem je předzpracování obrazu. Cílem předzpracování je potlačit šum a zkreslení vzniklé při digitalizaci a přenosu obrazu. Jindy se předzpracování snaží zvýraznit určité rysy obrazu podstatné pro další zpracování. Příkladem může být hledání hran.

Třetím a asi nejnáročnějším krokem postupu zpracování je segmentace, která umožňuje v obraze nalézt objekty. Za objekty lze považovat ty části obrazu, které nás z hlediska dalšího zpracování zajímají. Popis nalezených objektů v obraze je ovlivněn tím, na co se bude užívat. Lze je popsat buď kvantitativně pomocí souboru číselných charakteristik, nebo kvalitativně pomocí relací mezi objekty. Za krajně jednoduchý popis lze považovat velikost plochy objektu.

V některých případech není nutné provádět všechny uvedené kroky. Například při klasifikaci neuronovou sítí se segmentace a popis objektů explicitně neprovádí, neboť je součástí chování a struktury sítě. O to větší důraz může být kladen na kvalitní předzpracování obrazu.

Velkou komplikací při zpracování obrazu je totiž vztah mezi jasem, který měří kamera, a tvarem povrchu 3D objektů v obraze. Jas bodu totiž závisí na mnoha vlivech (odrazivost povrchu, poloha a vlastnosti zdrojů světla).

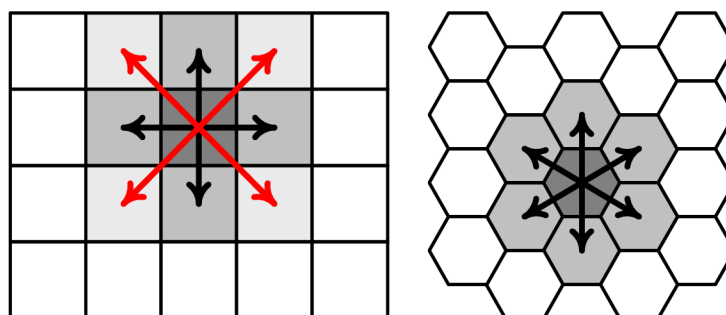
4.1 Digitální reprezentace obrazu

Abychom mohli obraz zpracovávat na počítači, je nejprve nutné jej zdigitalizovat. Digitalizace se skládá ze dvou částí:

- vzorkování obrazu do matice $M \times N$ prvků
- kvantování navzorkovaného obrazu do jasových úrovní

4.1.1 Vzorkování

Při vzorkování plošně spojitě obrazové funkce $f(x, y)$, se musí vyřešit dva problémy. Nejdříve musíme určit plošnou vzorkovací frekvenci, která reprezentuje vzdálenost mezi nejbližšími vzorkovacími body v obraze. Aby nedocházelo ke zkreslení obrazu, musí tato frekvence být nejméně dvakrát větší než nejvyšší frekvence ve vzorkovaném signálu (což je také nejmenší detail zastoupený v obraze). Tuto problematiku řeší *Shannonova věta* o vzorkování, známá z teorie řízení či teorie signálů. Druhým problémem je plošné uspořádání bodů při vzorkování (ve vzorkovací mřížce). Existuje několik typů mřížek, čtvercová a hexagonální (obr. 4.1). Čtvercová mřížka je častější díky snadnější technické realizaci a reprezentaci v počítači, ale na rozdíl od hexagonální mřížky nemá mezi okolními obrazovými body stejné vzdálenosti.



obr. 4.1: Ukázka čtvercové a hexagonální mřížky.

Jeden obrazový bod v uspořádané vzorkovací mřížce se nazývá obrazový element (anglicky *picture element*, *pixel*). Jednomu pixelu odpovídá hodnota funkce $f(x, y) = b$, kde x, y představují prostorové souřadnice a b jeho barvu, popřípadě odstín.

4.1.2 Kvantizace

Druhou část digitalizace představuje kvantizace obrazu, což je přiřazení spojitěmu signálu diskrétní hodnoty. Signál se nejprve rozdělí do několika (kvantizačních) úrovní, které svoji šířkou pokrývají celou amplitudu vzorku. Výzkumem se zjistilo, že pro lidské oko je dostatečný počet úrovní 230 stupňů šedi. Proto se ve výpočetní technice používá 256 (2^8 bitů) úrovní. Pro barevné obrazy potom pro každou složku

2^8 bitů. Při kvantizaci vzniká tzv. *kvantizační zkreslení*, kdy je vstupním spojitým vzorkům v určitém intervalu přiřazena pouze jedna hodnota odpovídající jednomu kvantizačnímu stupni. Z tohoto důvodu je nutno na základě daného počtu kvantizačních úrovní stanovit rozhodovací a rekonstrukční úrovně, tak aby kvantizační chyba byla minimální [2]. Takto digitalizovaný obraz dále používáme v počítačovém zpracování.

4.2 Základy zpracování obrazu (předzpracování obrazu)

Nejprve pár slov k literatuře. Publikací zabývajících se zpracováním obrazu je na českém trhu několik. Na rozdíl od světových titulů [5], které se obvykle snaží postihnout kompletní problematiku číslicového zpracování signálů, jsou české prameny většinou směřovány k určité problematice jako například restaurace signálů [1] nebo počítačová grafika a vidění [2]. V základech jsou však jednotné a jejich rozdílnost spočívá zejména v rozdělení a uspořádání metod spojených se zpracováním obrazů.

Výklad problematiky zpracování obrazu se opírá o některé základní pojmy, které je třeba vysvětlit. Prvním pojmem je *obraz*. Obrazem rozumíme optický obraz v obvyklém smyslu. Příkladem může být obraz vidění televizní kamerou. Takový obraz má plošnou (dvourozměrnou) povahu.

4.2.1 Obrazová funkce

Při formálním vymezení je použit matematický model obrazu. Tím je spojitá funkce dvou proměnných, tzv. *obrazová funkce* $f(x, y)$. Hodnotou obrazové funkce je nejčastěji *jas* (intenzita). Jas je obrazová veličina, která souhrnně vyjadřuje vlastnosti obrazového signálu způsobem, jenž odpovídá vnímání člověkem.

Obrazové funkce dělíme na spojitě a diskrétní. *Spojitá* funkce má definiční obor i obor hodnot spojitý. Pokud tvoří definiční obor množina diskrétních bodů a je-li oborem hodnot diskrétní množina, potom funkci nazýváme *diskrétní*.

Obraz může být v jednodušším případě *monochromatický*. Je reprezentován jedinou obrazovou funkcí. Ve složitějším případě se jedná o obraz barevný, *multispektrální*. Každé dvojici souřadnic odpovídá vektor hodnot neboli jasů pro jednotlivé barevné složky obrazu. Barevné složky obrazu mohou být reprezentovány trojicí jasů červeného, zeleného a modrého světla (*RGB*). Protože lidské oko vnímá intenzitu jednotlivých barevných složek světla různě, vyjadřuje se celkový jas pomocí empirického vztahu

$$I = 0,3R + 0,59G + 0,11B. \quad (4.1)$$

Spojitou obrazovou funkci je nutné reprezentovat v počítači, získat její diskrétní ekvivalent. Diskretizace dosahujeme digitalizací obrazu. Digitalizace spočívá ve

vzorkování obrazu v matici bodů a ve kvantování spojité úrovně jasové funkce. Čím jemnější je vzorkování a kvantování, tím přesnější je aproximace původního spojitého obrazu.

4.2.2 Konvoluce

Důležitou operací při lineárním přístupu ke zpracování obrazu je *konvoluce*. Konvoluce dvourozměrných spojitých funkcí f a h je definována integrálem

$$f(x, y) * h(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x - a, y - b)h(a, b) da db. \quad (4.2)$$

kde funkce $h(x, y)$ se nazývá *konvolučním jádrem*.

Při práci s digitálním obrazem se používá tzv. *diskrétní konvoluce*, která je diskrétní dvojrozměrnou podobou předchozího vztahu,

$$I(x, y) * h(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k I(x - i, y - j)h(i, j). \quad (4.3)$$

V tomto případě je $I(x, y)$ diskrétní obraz a $h(x, y)$ jádro konvoluce.

4.2.3 Filtrace

Část operací předzpracování obrazu se nazývá *filtrace*. Jedná se o metody využívající pro výpočet jasu bodu ve výstupním obraze jen lokálního okolí odpovídajícího bodu ve vstupním obraze. Na filtraci lze také pohlédnout jako na diskrétní konvoluci. Přičemž konvoluční jádro definuje použité lokální okolí.

V mnoha praktických případech se využívá pravoúhlého okolí. Aby bylo použité okolí symetrické vůči jeho středovému elementu (x, y) , volí se nejčastěji rozměry okolí z množiny lichých přirozených čísel.

Konvoluční lineární operace (*filtry*) jsou použitelné pro vyhlazování, detekci hran nebo čar (gradientní operace). *Vyhlazování obrazu* vede k potlačení vyšších frekvencí obrazové funkce. Žádaným výsledkem vyhlazování je potlačení náhodného šumu. Současně naneštěstí dochází k potlačení ostatních náhlých změn jasové funkce, neboli k rozmazání hran. Příkladem vyhlazování je obyčejné *průměrování*. Každému bodu přiřazuje nový jas, který je průměrem původních jasů ve zvoleném okolí. V některých případech se zvyšuje váha středového bodu masky (konvolučního jádra) nebo jeho sousedů, aby se lépe aproximovaly vlastnosti šumu s gaussovským rozložením. Příkladem často používaných průměrovacích filtrů jsou

$$h_{16} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, h_{10} = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Výsledek vyhlazování filtrem h_{16} je znázorněn na obr. 4.2. Ve filtrovaném obraze (vpravo) je patrné částečné odstranění náhodného impulsního šumu, ale i výrazné rozmazání hran.



obr. 4.2: Filtrace standardním průměrováním.

Gradientní operace a s nimi těsně související ostření obrazu naopak vedou ke zdůraznění vyšších frekvencí. Současně jsou zvýrazněny ty obrazové elementy, ve kterých se jasová funkce náhle mění a má zde velký modul gradientu. Žádaným výsledkem je zvýraznění hran v obraze. Bohužel jsou zvýrazněny i šumové body. Z uvedeného srovnání vyhlazovacích a ostřících operací je patrná jejich protichůdnost. Řešení tohoto rozporu umožňují některé algoritmy, které oba postupy kombinují.

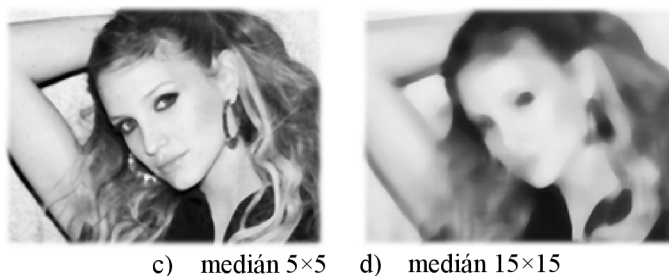
Medián

Společným principem všech nelineárních filtračních metod je snaha nalézt v analyzovaném okolí jen tu jeho část, do které reprezentativní bod patří. Jen pixely v této oblasti se využijí k hledání jasové hodnoty. Nejjednodušší a nejpoužívanější operací ve zpracování obrazu je *medián* [3].

Výpočet mediánu je pro diskrétní obrazovou funkci jednoduchý. Stačí vzestupně uspořádat hodnoty jasu v lokálním okolí a medián určit jako prvek, který je uprostřed této posloupnosti. Aby se snadno určil prostřední prvek, používají se posloupnosti s lichým počtem prvků. Výpočet ještě urychlí skutečnost, že k nalezení mediánu stačí částečné uspořádání posloupnosti. Metoda redukuje stupeň rozmazání hran a dobře potlačuje impulsní šum. Její nevýhodou je ovšem to, že porušuje tenké čáry a ostré rohy v obraze. Tento nedostatek se dá vyřešit použitím jiného tvaru okolí [3].



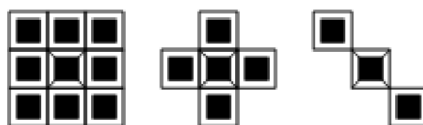
a) originální obraz b) medián 3×3



obr. 4.3: Příklady mediánového filtru.

4.2.4 Matematická morfologie

Matematická morfologie je soubor funkcí, pracujících s bodovými množinami, v našem případě binárními obrazy. Definičním oborem těchto obrazů je číslo z množiny \mathbb{E}^2 (Eukleidovský prostor) a může nabývat hodnot $\{0, 1\}$. Využívá se tzv. *morfologické transformace*, která je dána relací mezi obrazem a typicky menším *strukturním elementem* B . Ten je vždy vztažen k lokálnímu počátku (označen křížkem). Dvěma základními operacemi jsou *dilatace* a *eroze*. [3]

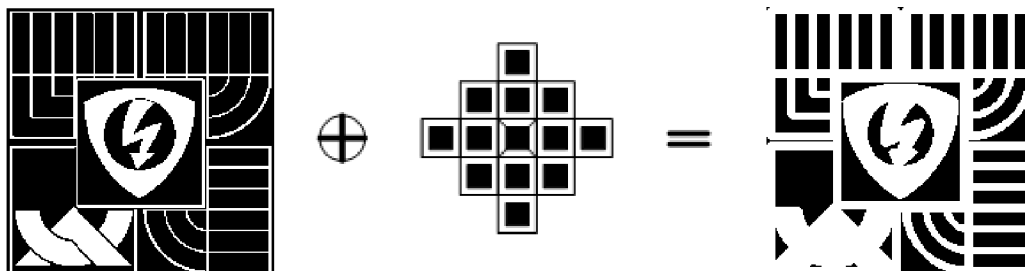


obr. 4.4: Příklady strukturních elementů.

Dilatace

Sčítá dvě bodové množiny.

$$X \oplus B = \{p \in \mathbb{E}^2: p = x + b, x \in X \wedge b \in B\}. \quad (4.4)$$



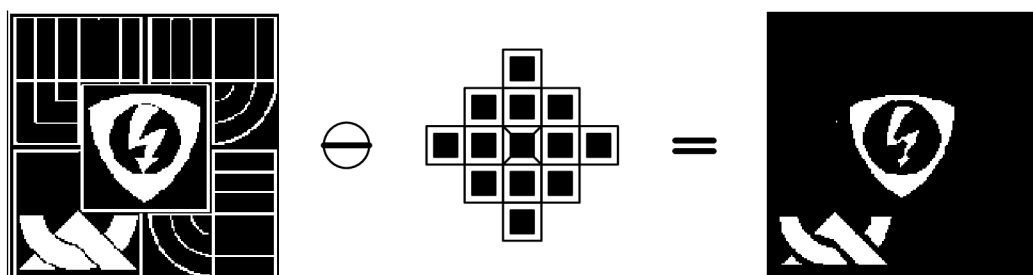
obr. 4.5: Dilatace v praxi, strukturní element ve tvaru disku o poloměru 2.

Eroze

Skládá dvě množiny pomocí *Minkovského rozdílu*. Jde o *duální morfologickou transformaci* k dilataci.

$$X \ominus B = \{p \in \mathbb{E}^2: p + b \in X, \forall b \in B\}. \quad (4.5)$$

Pro každý bod obrazu p se ověřuje, zda pro všechna možná $p + b$ leží výsledek v X . Pokud ano, je výsledek 1, jinak 0. Z praktické ukázky je vidět, že objekty menší



obr. 4.6: Eroze v praxi, strukturní element ve tvaru disku o poloměru 2.

než *strukturní element* vymizely. *Eroze* je proto použitelná zejména na zjednodušení struktury obrazu.

Otevření

Eroze následovaná dilatací

$$X \circ B = (X \ominus B) \oplus B. \quad (4.6)$$

Uzavření

Dilatace následovaná erozí

$$X \bullet B = (X \oplus B) \ominus B. \quad (4.7)$$

4.2.5 Detekce hran

Hrana v obraze je vlastnost obrazového elementu a jeho okolí. *Hrana* je vektorová veličina a je určena velikostí a směrem. Tyto veličiny vycházejí z gradientu (operátor *nabla*) obrazové funkce. Hrana indikuje body obrazu, ve kterých dochází ke změnám obrazové funkce. Gradient je dán

$$\nabla f(x, y) = \left(\frac{\partial f(x, y)}{\partial x}, \frac{\partial f(x, y)}{\partial y} \right). \quad (4.8)$$

Z matematického hlediska udává gradient směr, v němž funkce f v bodě (x, y) nejrychleji roste. Jeho velikost se pak počítá jako

$$|\nabla f(x, y)| = \sqrt{\left(\frac{\partial f(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f(x, y)}{\partial y} \right)^2}. \quad (4.9)$$

Příklad detekce hran v obraze demonstruje obr. 4.7. Výpočet potřebných obrazových derivací a použitý operátor bude dále diskutován.

4.2.6 Obrazové derivace

Způsob reprezentace obrazu diskrétní obrazovou funkcí znemožňuje přesný výpočet obrazových derivací. Tyto derivace lze proto pouze aproximovat, a to např. pomocí *diferencí* neboli *rozdílů*. Nejjednodušší jsou tzv. *dvoubodové* diference, jejich matematické vyjádření je

$$I_x(i, j) = I(i, j) - I(i - k, j), \quad (4.10)$$

$$I_y(i, j) = I(i, j) - I(i, j - k), \quad (4.11)$$

kde k je malé celé číslo, většinou 1. Výpočet derivací použitím těchto rovnic je velmi nepřesný, lepších výsledků lze dosáhnout použitím tzv. *centrálních* diferencí,

$$I_x(i, j) = \sum_k w(k) (I(i - k, j) - I(i + k, j)), \quad (4.12)$$

$$I_y(i, j) = \sum_k w(k) (I(i, j - k) - I(i, j + k)), \quad (4.13)$$

kde $k \in 1, \dots, n$ a w je případná váhová funkce.

Operátory pro detekci a ohodnocení hran v diskrétních obrazech vycházejí právě z diferencí. Příkladem dobré aproximace gradientu je *Sobelův* operátor, který je směrově závislý. Sobelův operátor lze vytvořit pro různě velké rozměry masek. Příkladem konvoluční masky 3×3 pro dva z osmi směrů jsou

$$h_1 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, h_2 = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -2 & -1 & 1 \end{bmatrix}.$$

Chování operátoru h_1 ilustruje obr. 4.7. Čím výraznější je bod výsledného obrazu, tím větší je odezva hranového operátoru.



obr. 4.7: Detekce hran – Sobelův operátor.

4.3 Metody detekce obličeje a očí

4.3.1 Detekce v šedotónových (monochromatických) obrazech

Metody znalostního základu

Tyto algoritmy využívají k detekci explicitně definovaná pravidla založená na znalosti struktury lidského obličeje. Jedná se většinou o jednoduché souvislosti mezi různými částmi obličeje (například vzájemná poloha očí, nosu a úst).

Práce [28] prezentuje metodu založenou na integrální projekci. Označením intenzity jasu $I(x, y)$ v obraze o rozměrech $m \times n$ na pozici (x, y) je možno definovat horizontální projekci $HI(x)$ a vertikální projekci $VI(y)$ jako

$$HI(x) = \sum_{y=1}^n I(x, y), \quad (4.14)$$

$$VI(y) = \sum_{x=1}^m I(x, y), \quad (4.15)$$

Dvě lokální minima v oblasti velkých změn křivky horizontální projekce určují strany hlavy. Z vertikální projekce jsou získány polohy nosu a úst a očí. Těmito pravidly se určí možná poloha obličeje ve scéně, která je ještě ověřena dalšími pravidly pro vzájemnou polohu částí obličeje. Tento algoritmus dokáže velmi přesně lokalizovat jednu tvář ve frontálním pohledu v prostředí bez komplexního pozadí (100% obličejů, 85% obličejů včetně správné lokalizace jejich částí).

Extrahovat z obecných znalostí pravidla použitelná v algoritmech tak, aby definovala pokud možno pouze strukturu obličeje a zároveň aby brala v úvahu jeho různorodost, je největším problémem znalostních metod. Explicitně definovaná pravidla se navíc těžko rozšiřují na případy, kdy je obličej natočen nebo částečně zakryt. Znalostní metody proto nacházejí uplatnění hlavně při lokalizaci obličeje ve scénách s homogenním pozadím a pro detekci v přesném slova smyslu se nehodí.

Metody invariantních rysů

Společným jmenovatelem těchto metod je předpoklad, že existují nějaké znaky, které mohou obecně popsat obličej nezávisle na jeho podobě nebo parametrech scény. Jednotlivé metody se ovšem rozcházejí v tom, jak invariantní rysy definovat a jakým způsobem je v obraze hledat. Některé metody vidí tyto příznaky v geometrickém uspořádání jednotlivých částí tváře, jako jsou oči, obočí nebo ústa. Extrakcí těchto částí hranovými detektory a vytvořením statistických modelů jejich vztahů je pak ověřována přítomnost obličeje v obraze. K rozlišení obličejů od ostatních objektů může být použita i jejich specifická textura nebo rozložení jasu.

Algoritmus [39] vychází z pozorování, že ačkoliv je jas jednotlivých oblastí obličeje ovlivněn různým osvětlením, relativní hodnoty jasu mezi těmito oblastmi zůstávají na různém osvětlení nezávislé. Tvář je rozdělena na 16 regionů a mezi jednotlivými páry je definováno, který z nich má být světlejší a který tmavší. Oblast v prohledávaném obraze je prohlášena za tvář, pokud vyhovuje těmto párovým pravidlům.

Metody porovnání se šablonou

V tomto přístupu je vzor obličeje manuálně předdefinován šablonou nebo parametrizován funkcí. Kritériem pro lokalizaci obličeje je pak hodnota korelace prohledávaného obrazu a šablony. Většinou jsou vytvářeny hierarchické šablony – pro obrys hlavy a části obličeje. Analýzou vzájemné pozice šablon (respektive maximálních hodnot korelace) v obraze je pak lokalizována tvář.

Výhodou těchto metod je snadná implementace, na druhou stranu mají problémy detekovat obličeje, které se natočením nebo velikostí liší od šablony. Tento nedostatek se snaží překlenout deformovatelné šablony, nebo šablony pracující s více stupni rozlišení.

Ve starších přístupech [63] popisují šablony tvar obličeje a jeho částí pomocí obrysů. Prohledávaný obraz je zpracován filtrem, který extrahuje hrany, hledání obličeje pak probíhá hierarchicky ve dvou krocích. Nejprve je hledán většinou oválný obrys obličeje, poté se v nalezené oblasti hledají části podobné šablonám jednotlivých prvků obličeje.

Metoda založená na deformovatelných šablonách je prezentována v [74]. Příznaky obličeje jsou popsány parametrizovanou šablonou. Je definována energetická funkce, která přiřazuje hrany, jasová maxima a minima prohledávaného obrazu parametrům šablony. Největší podobnosti je dosaženo minimalizací energetické funkce této elastické šablony.

Metody založené na zjevu

Oproti srovnávání šablon, kde jsou šablony tvořeny na základě heuristiky - expertních znalostí, využívají metody založené na zjevu k tvorbě šablon učící algoritmy. Tyto metody se opírají zejména o strojové učení a statistickou analýzu. Naučené charakteristiky mají většinou formu modelů pravděpodobnostního rozložení,

kteřé jsou pak užity ke klasifikaci. S tím je spojena snaha redukovat vysokou dimenzi modelů za účelem snížení rychlosti výpočtu. Tato oblast se oproti předchozím třem v poslední době nejvíce rozvíjí, proto bude popsána podrobněji.

Mnohé metody nahlížejí na detekci obličejů z pravděpodobnostního hlediska. Na obraz nebo na jeho parametrizovaný protějšek je nahlíženo jako na náhodnou veličinu x . Veličina x je pro členy třídy „obličej“ a „ostatní“ charakterizována rozložením hustoty podmíněných pravděpodobností $p(x|\text{obličej})$ a $p(x|\text{ostatní})$. Ke klasifikaci jsou pak použity klasifikátory jako Bayesův nebo ML (Maximum Likelihood) detektor [67]. Přímá implementace tohoto přístupu bohužel není možná, neboť neexistuje přirozená parametrizace obou tříd a dimenze veličiny x je obrovská.

Třídy se proto parametrizují aproximacemi, které jsou získány experimentálně učením. Jednou z často používaných metod jak redukovat dimenzi proměnné veličiny je vytvoření vlastních vektorů obličejů (takzvané *eigenfaces*) za použití PCA (Principal Component Analysis) [67]. Obraz o rozměrech m a n je zde chápán jako vektor ve vektorovém prostoru všech obrazů o dimenzi $m \times n$. Vzhledem k tomu, že obličejů mají některé společné znaky, sdílí v tomto prostoru podprostor o menší dimenzi. Obličejů v tomto podprostoru lze pak vyjádřit lineární kombinací vektorů báze (ty představují jakési báze obličejů). K získání vektorů báze slouží právě PCA. Bázi tvoří vlastní vektory kovarianční matice spočítané ze sady trénovacích vzorů obličejů. Detekce pak probíhá na základě prahování vzdálenosti od oblasti obličejů v takto vzniklém podprostoru.

Složitější metody budují v tomto podprostoru oblasti tříd „obličej“ a „ostatní“ na základě trénovacích množin obou tříd. Obě oblasti jsou pak modelovány funkcemi rozložení hustoty pravděpodobnosti. Předložený obraz je poté klasifikován do té třídy, v jejíž funkci rozdělení hustoty má větší hodnotu.

Mnoho problémů klasifikace vzorů (rozpoznávání písma, objektů) je s úspěchem řešeno *neuronovými sítěmi*. Jejich výhodou je především schopnost popsat třídu s velkou rozmanitostí zástupců. Neuronové sítě musí být ovšem pro správný výkon pečlivě vyladěny a natrénovány, což je pro velký počet parametrů (počet vrstev, počet neuronů v jednotlivých vrstvách, parametry učení) nesnadný úkol.

Většina neuronových sítí určených pro detekci pracuje s okny o pevně nastavené velikosti (do 20×20 pixelů), sítě bývají hierarchické s několika skrytými vrstvami, většinou obsahují podsítě, které řeší nějakou specifickou část problému. Aby byl systém schopen detekovat i větší obličejů, bývá vstupní obraz iterativně podvzorkován. Například v [62] je ze vstupního obrazu vyjmuto okno o velikosti 20×20 pixelů, ve kterém je korigován jas. Jako vstup do sítě slouží jednak samotné okno 20×20 pixelů, jednak vektor průměrného jasu v řádkách a jednak matice 2×2 , která obsahuje průměrné hodnoty jasu ve čtvrtinách obrazu. Každý vstup klasifikuje jedna samostatná podvrstva, jejíž výstup vstupuje do vyšší vrstvy. Ta pak klasifikuje celkový obraz.

Ke klasifikaci založené na učení se také používají architektury jako SVM (Support Vector Machine), SNoW (Sparse Network of Winnows), nebo skryté Markovovy modely (Hidden Markov Models). Detektory založené na učení lépe klasifikují obličej v různých podmínkách, systémy jsou adaptabilnější. Daní za tyto výhody jsou zvýšené nároky na výpočetní kapacitu. Výkon systémů značně závisí na trénovací množině a parametrech učení.

4.3.2 Klasifikace v barevných obrazech

Detekce na základě barvy kůže⁸ je často používanou součástí systémů pro detekci obličejů. Její největší předností je rychlost zpracování při zachování přesnosti detekce. Navíc je invariantní vůči natočení nebo poloze obličejů ve scéně. Klasifikace barvy kůže se používá k předzpracování obrazu, vede k zúžení prohledávané oblasti a tím ke zrychlení výpočtu a větší přesnosti detekčních algoritmů. Samotná detekce obličejů se pak provádí některou z výše uvedených metod.

Základní princip všech metod vychází z pozorování, že barva lidské kůže sdílí určitý kompaktní podprostor v barevném prostoru [50]. Ohraničením tohoto prostoru lze modelovat barvu lidské kůže a pak klasifikovat barvu jednotlivých pixelů obrazu. Klasifikaci na základě barvy kůže lze rozdělit do tří hlavních kroků. Za prvé jde o výběr vhodného barevného modelu, za druhé je nutno zvolit model barvy kůže a jeho přesnost a za třetí definovat kritéria pro klasifikaci.

Barevné modely

Výzkum v oblasti *kolorimetrie*, zobrazování a zpracování obrazu přinesl mnoho barevných modelů s rozdílnými vlastnostmi. Tvar a kompaktnost podprostoru barev kůže se v jednotlivých modelech liší. Výběr vhodného barevného modelu je tedy zásadní pro mnoho metod klasifikace. Následující část shrnuje nejčastěji používané barevné modely.

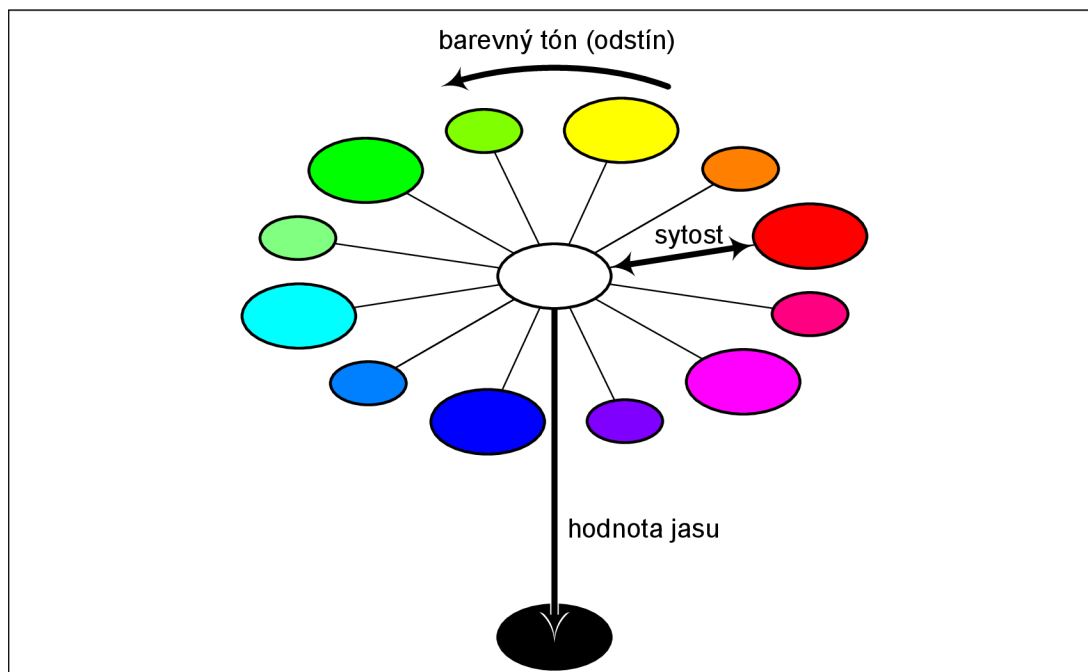
⁸ Lze se setkat i s pojmem plet'ové barvy.



obr. 4.8: Převod z modelu RGB do modelů YCbCr a HSV. Horní řada vždy představuje jednotlivé složky modelu v monochromatickém rozlišení; dolní řada představuje postupně každou složku maximálně potlačenou a pro lepší představu je výsledek převeden zpět do modelu RGB.

Za základní barevný model používaný ve zpracování obrazu lze označit barevný model RGB. Barva je definována složením tří základních barev (červená, modrá a zelená) o různé intenzitě. Vzhledem k tomu, že tento barevný model explicitně neodděluje jasovou a barevnou složku, není příliš využíván.

Dalším často používaným modelem je HSV (HSL), obr. 4.9. Tento model přináší intuitivnější popis barvy pomocí barevného tónu (hue), sytosti (saturation) a jasu (value, lightness). Explicitní oddělení jasové složky a další vlastnosti podporují použitelnost tohoto modelu [50].

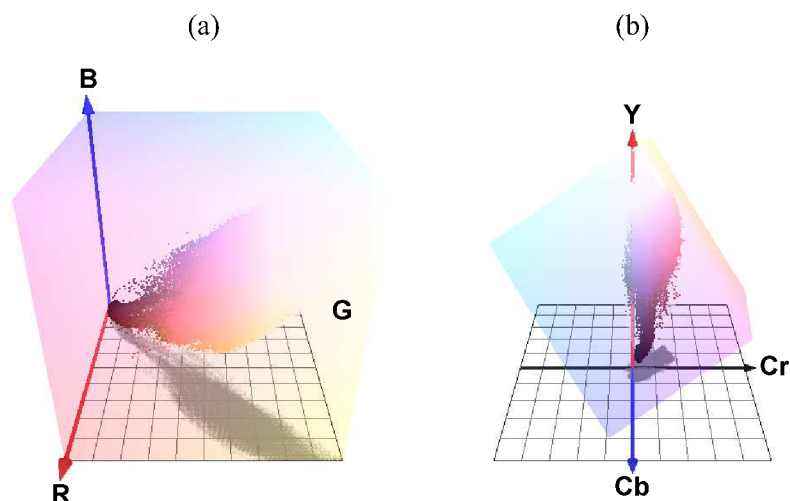


obr. 4.9: Barevný model HSV.

Televizní standard NTSC i obrazový kompresní algoritmus JPEG používají barevný model YCbCr. Jde o transformovaný RGB model, kde Y reprezentuje jasovou složku a Cb a Cr složky chrominanční. Vyznačuje se jednoduchým převodním vztahem mezi oběma modely:

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,500 \\ 0,500 & -0,419 & -0,081 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (4.16)$$

Jednoduchost transformace a fakt, že z tohoto modelu vychází často používané formáty, zajistily velkou oblibu tohoto modelu mezi systémy pro klasifikaci (např. [60]). Z podobných modelů lze ještě zmínit YUV nebo YIQ použitý v [35].



obr. 4.10: Podprostor barvy kůže v modelu RGB (a), podprostor barvy kůže v modelu YCbCr (b).

Explicitní vyjádření oblasti s barvou kůže

Jednou z metod jak implementovat klasifikátor je explicitní definice pravidel ohraničujících podprostor barvy kůže v nějakém barevném modelu. Barva pixelu je pak označena jako barva kůže, pokud tato pravidla splňuje. Jednoduché rozhodovací pravidlo je prezentováno například v [87], kde je použit barevný model YIQ: pokud je barevná složka I v intervalu $\langle 20, 100 \rangle$, je pixel klasifikován jako kůže.



obr. 4.11: : Vzorová šablona barvy kůže.

Pravidla jsou vytvořena na základě rozložení vzorových barev kůže v barevném modelu. Toto rozložení je získáno experimentálně srovnáním různých testovacích vzorů. Největší výhodou je lehká implementace a tím i velmi výrazná rychlost klasifikace. Kvalita výsledku je velmi závislá na vybraném modelu a přesnosti pravidel, která je nutno určit experimentálně. Klasifikátor je tak neflexibilní a nedokáže reagovat na změněné podmínky osvětlení. Přesto je pro jeho rychlost často používán.

Neparametrické modely rozložení barvy kůže

Klíčová myšlenka tohoto přístupu je zjistit rozložení barvy kůže z trénovací množiny bez explicitního vyjádření barevného podprostoru pomocí pravidel. Namísto toho je většinou konstruována takzvaná *pravděpodobnostní mapa* – jednotlivým bodům diskretizovaného barevného prostoru je přiřazena pravděpodobnost, s jakou má daný bod barvu kůže.

Příkladem tohoto přístupu je například metoda zvaná „Normalized Lookup Table“ (princip využitý v [49]). Barevný prostor (většinou pouze chrominanční rovina) je kvantizován do diskretních oblastí. Během předkládání trénovacích vzorů je ke každé oblasti přiřazena četnost, s jakou se barva dané oblasti vyskytovala v předložených vzorech. Výsledkem trénování je tedy histogram obsahující četnosti výskytu jednotlivých oblastí v trénovací množině. Z histogramu se poté normováním získá pravděpodobnostní rozložení.

Při klasifikaci je pak barvě každého pixelu přiřazena pravděpodobnost příslušné oblasti. Pokud tato pravděpodobnost překročí určenou hranici, je pixel klasifikován jako kůže. Mezi další neparametrické metody patří také Bayesův klasifikátor nebo přístup založený na Kohonenových neuronových sítích.

Síla neparametrického modelu tkví v rychlosti učení a nezávislosti na tvaru podprostoru barev kůže. Tento model bohužel vyžaduje značné množství paměti na uložení pravděpodobnostní mapy (velikost závisí na hustotě kvantizace barevného modelu) a není schopen interpolace chybějících hodnot.

Parametrické modely

Parametrické modely se snaží popsat tvar rozložení barvy kůže v rámci barevného modelu vhodným matematickým modelem. Parametry tohoto modelu jsou pak získány z trénovací množiny. To, že podprostor barvy kůže má ve většině používaných modelů eliptický tvar a přibližně normální rozložení, vedlo některé vědce [32] k aplikaci eliptické Gaussovy funkce hustoty pravděpodobnosti:

$$p(c|k) = \frac{1}{2\pi|\Sigma_s|^{1/2}} e^{-\frac{1}{2}(c-\mu_s)^T \Sigma_s^{-1}(c-\mu_s)}, \quad (4.17)$$

kde c je vektor barvy a $p(c|k)$ značí pravděpodobnost, s jakou je barva c barvou kůže. Vektor středních hodnot μ_s a kovarianční matice Σ_s jsou parametry funkce, které ovlivňují její tvar. Tyto parametry jsou získány z trénovací množiny. Více sofistikovaná metoda používá součet několika takovýchto funkcí, aby bylo možno modelovat komplexnější tvary rozložení. Větší přesnost je avšak kompenzována větší složitostí a většími nároky na rychlost výpočtu.

Velmi úspěšný model je představen v [42]. Podprostor barev kůže je sice eliptický, hustota pravděpodobnosti je v něm však rozdělena asymetricky. To má za následek, že jednoduchá symetrická Gaussova funkce neaproximuje dostatečně vlastnosti

podprostoru barev kůže. [42] nabízí řešení v podobě nového modelu rozložení pravděpodobnosti, takzvaný „Eliptic boundary model“, kde je tento aspekt brán v potaz.

Kladnou vlastností parametrických modelů je oproti neparametrickým jejich schopnost interpolovat chybějící data trénovací množiny. Hodnotu pravděpodobnosti je možno určit pro jakoukoliv barvu. Parametrizace umožňuje dramaticky redukovat potřebnou paměť pro uložení modelu, protože postačuje uložit pouze několik parametrů funkce. Nevýhodou jsou naopak výpočetní nároky (zvláště u modelu používajícího několik Gaussových funkcí) a značná závislost na použitém barevném modelu.

Shrnutí klasifikace v barevných obrazech

Klasifikace na základě barvy kůže se ukázala jako velmi účinný prostředek k redukci prohledávaných oblastí při detekci obličeje. Explicitně vyjádřené modely těží z jednoduchosti a rychlosti, neparametrické z nezávislosti na barevném modelu a tvaru podprostoru barev kůže, parametrické modely dovedou interpolovat chybějící data a jsou nenáročné na paměť. Výkonnost explicitních i parametrických modelů ovlivňuje volba barevného modelu, pokud je ovšem v modelu rozložení barev kůže nekompaktní nebo dochází k překrývání barev kůže a ostatních barev, může ovlivňovat i modely neparametrické.

Většina modelů pracuje pouze s chrominanční rovinou barevného prostoru. Argumentace správnosti tohoto zjednodušení vychází především z pozorování, že se barva kůže různých jedinců liší především v hodnotě jasu. Zanedbání jasové složky redukuje dimenzi prostoru a tak jednoznačně zvyšuje rychlost výpočtu. Model nezávislý na jasové složce je obecnější a platnější i pro jiné jasové podmínky osvětlení než u trénovacích vzorů.

4.4 Metody detekce mrknutí

Detekovat mrknutí oka, lze využitím několika v základu rozdílných algoritmů, kde každý má své výhody a nevýhody. S úspěšností algoritmu rostou i požadavky na kvalitu obrazu. Například u přesného zaměření duhovky je potřeba vysoké rozlišení, které poskytuje úplný popis pohybu očních víček.

4.4.1 Detekce mrknutí pomocí rozpoznání duhovky

Detekce mrknutí pomocí rozpoznání duhovky využitím výpočtu kruhové intenzity

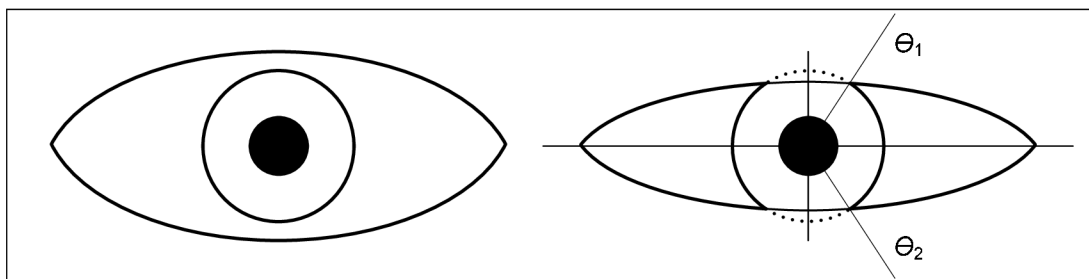
V první řadě se pro algoritmus [88] musí určit střed oka pomocí detekce panenky, která je kruhová a tmavá, nebo pomocí detekce duhovky, která může být podobně tmavá jako panenka. Na výřez oka se aplikuje jeden z vyhlazovacích filtrů, buď mediánový, nebo rozostřovací, a to v takové míře, aby nedošlo k velikému posunu hrany duhovky. Dále se pomocí horizontální osy, která prochází vertikálním středem nejtmaší oblastí panenky a změny jasu na ní určí začátek a konec duhovky.

V dalším kroku algoritmus určí přesnou pozici duhovky z již získaného středu oka a prvního obrázku, jenž není filtrován a je součástí předchozích algoritmů na detekci očí. Algoritmus je založen na vlastnostech oka. Duhovka je tmavší než okolní bělmo, je kulatá a nachází se v oblasti výřezu oka.

Kulatosti duhovky využívá vztah, který určuje funkci intenzity kruhové oblasti kolem bodů x_0 a y_0 s poloměrem r , v rozmezí $\theta = 0$ až 2π :

$$f_{\theta}(x_0, y_0, r) = \int_{\theta \in \theta} I(x_0 + r \cos \theta, y_0 + r \sin \theta) d\theta, \quad (4.18)$$

Rozmezí θ až 2π se většinou nenalézá v plném rozsahu, protože duhovka může být částečně až plně překryta spodním či horním víčkem, takže výpočet probíhá od horizontální osy oka nahoru i dolů, vlevo i vpravo (obr. 4.12) a testuje se jaká intenzita je za hranicí poloměru a jaká před hranicí, aby se určilo, zda se ještě jedná o duhovku nebo o jedno z víček.



obr. 4.12: Otevřené oko a částečně přivřené s rozmezím.

Určení pozice duhovky pomocí detekce hran

Tato metoda [115] slouží k detekci a vyjmutí duhovky, která je pak roztažena po obvodu panenky. Tato data jsou dále zpracována a použita k identifikaci majitele. Na rozdíl od předchozího algoritmu se neuvažuje se soustřednými a kruhovými vlastnostmi panenky a duhovky. Prvním krokem této metody je hledání středu využitím aproximace obvodu duhovky podle horizontální osy.

Obraz je také předzpracován vyhlazovacími filtry a poté je navýšen jeho kontrast k výraznějšímu přechodu přes hrany. Detekce hran je provedena filtrováním Haarovou vlnkou po horizontální přímce, která zdůrazní vzrůst či pokles světlosti a tím detekuje hranu. Následně po znalosti filtrovaného průběhu ze středu panenky, jsou v průběhu vyhodnoceny zlomy a tím určeny poloměr duhovky i panenky.

4.4.2 Detekce mrknutí pomocí diferenční matice

Pokud pro detekci mrkání je již připraven výřez oka, zpracovávat se můžou přímo tyto data pomocí algoritmu [67]. Algoritmus má za úkol určit pozici obou očí a to

i za okolností kdy je hlava nakloněna do strany. Výpočet algoritmu používá diferenčních map pro detekci.

Prvním krokem je inicializace matice na nulové hodnoty diferenční mapy σ^2 o velikosti $M \times N$, velikosti výřezu. Dalším krokem je inicializace průměrovaného obrázku μ o velikosti $M \times N$ s rozsahem 24 bitů na pixel, kterému jsou přiřazeny hodnoty prvního obrazu.

Dalším krokem je upravení diferenční matice a průměrové matice podle druhého obrazu podle vzorců (4.19) a (4.20). Vzorec (4.19) má za úkol zpracovat hodnoty, které jsou obsaženy v posloupnosti matic vyjadřujících videosekvenci

$$\mu_{j+1}(x, y) = \frac{j\mu_j(x, y) + I_{j+1}(x, y)}{j + 1}, \quad (4.19)$$

$$\sigma^2_{j+1}(x, y) = \left(1 - \frac{1}{j}\right)\sigma^2_j(x, y) + (j + 1)\left(\mu_{j+1}(x, y) - \mu_j(x, y)\right), \quad (4.20)$$

kde μ označuje matici průměrovacího obrazu, σ je matice varianční mapy a I je vstupní obraz. V dalším kroku se diferenční mapa nechá projít prahováním s prahem 255⁹, aby v matici nebyly přesycené hodnoty. Poté se diferenční mapa podrobí druhému prahování dle velikostí oblastí, kde došlo k zanechání informací ve varianční mapě, tak aby malé a velké oblasti byly vynechány. Počet pixelů, jenž prošel přes prahovací filtr, určuje v poměru s celkovým počtem pixelů v okně výřezu, že došlo k mrknutí oka. Diferenční mapa je nyní matice s binárními pixely (binární mapa), které jsou buď šum anebo jsou spojeny do větších celků. Na tyto pixely jednotlivých oblastí lze aplikovat algoritmus, který pixel po pixelu nesoucí stejnou informaci jako jeho soused v diferenční mapě, začne procházet oblast ve všech osmi směrech tak, aby odlišil samostatné pixely nebo menší oblasti od oblastí určité velikosti a polohy. Tím, že algoritmus projde celou oblast, vznikne kontura této oblasti. V obraze by po zpracování měly zůstat dvě tyto výrazné kontury v horní třetině z tohoto výřezu z detekce na základě pohybu. Tyto dvě kontury by měly být přibližně horizontálně elipsovité se špičkami na stranách. Tyto špičky lze využít ke zjištění, jak jsou oči nakloněny a kde je jejich střed. Po určení bočních mezí oka se určí dolní a horní maximum oka.

4.4.3 Detekce očního víčka

Metoda [51] pracuje na základě detekce hran, která je uskutečněna horizontálním Sobelovým jádrem. Filtrace Sobelovým operátorem zvýrazní hrany dané oběma víčky a jejich řasami, ale také hranu způsobenou přehybem měkké tkáně mezi obo-

⁹ 255 je hraniční hodnota při 8-bitovém vyjádření, které se obvykle pro jednu obrazovou vrstvu používá.

čím a víčkem. Ostatní hrany vrásek nejsou tak masivní, takže je nutné jen správně odlišit hrany přehybu od víčka. Identifikace hran probíhá určením, v jaké změně kontrastu se každá hrana nachází. Vrásky a přehyb jsou obklopeny kůží, ale víčko je z jedné strany obklopeno kůží přechodem přes řasy nakonec k oku. Druhou možností jak odlišit vzniklé hrany je přes barevný model, tak že z barevného modelu odstraníme zelenou a modrou složku, tak zůstane červená, která je hodně v oblasti kůže zastoupena. Z použití barevného kanálu plyne podmínka, že touto metodou lze zpracovávat jen barevný obraz.

Z již získaných souřadnic hran spodního i horního víčka vzniknou *Beziérov*y *křivky*, z kterých můžeme určit do jaké míry je oko otevřeno.

5 Obrazové databáze

Pro vývoj a testování algoritmů určených k rozpoznávání obličejových rysů je zapotřebí dostatečně velká, pokud možno co nejkompaktnější, databáze obličejů. Aby bylo možné simulovat použití algoritmu v různých podmínkách, měla by obrazová databáze obsahovat různé výrazy obličejů, různé podmínky osvětlení a různé druhy pozadí. Do dnešní doby byl vytvořen velký počet obličejových databází, v některých případech jsou obrazové databáze přizpůsobeny pro testování jednoho určitého algoritmu (detekce natočení hlavy, rozpoznání řeči, atd.). Mnohé z již vytvořených obrazových databází jsou pro nekomerční využití volně dostupné na internetu.

5.1 Dostupné obecné obrazové databáze

V následující části je uveden seznam vybraných na internetu dostupných obličejových databází. U každé databáze je krátký přehled vlastností a ukázka z databáze.

tab. 5.1: Srovnání parametrů veřejně dostupných videodatabází

název databáze	počet osob [–]	délka [snímků/os]	formát [–]
AT&T The Database of Faces	40	10	PGM
BANCA Database	52	500	PNG
Cohn-Kanade AU-Coded Facial Expression Database	100	-	-
Collection of Facial Images	395	20	JPEG
Georgia Tech face database	50	15	JPEG
MMI Face Database	19	182	-
NRC-IIT Facial Video Database	11	300	AVI
The BioID Face Database	23	66	PGM
The University of Oulu Physics-Based Face Database	125	16	BMP
The VidTIMIT Audio-Video Dataset	40	1500	JPEG+WAV
The Yale Face Database	15	11	GIF
The Yale Face Database B	10	576	PGM

5.1.1 Popis dostupných databází

AT&T The Database of Faces

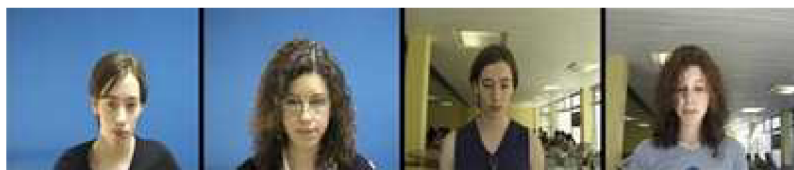
Databáze [107] obsahuje vzorky obličejů zachycených v období duben 1992 až duben 1994 v laboratorních podmínkách. Pro každou ze 40 osob obsahuje deset vzorků. Pro některé osoby byly vzorky zachyceny v různém čase, s různým osvětlením, výrazem v obličeji (otevřené/zavřené oči, úsměv) a obličejovými doplňky (s/bez brýlí). Všechny vzorky byly zaznamenány proti tmavému jednotnému pozadí s čelním pohledem. Velikost obrazu každého vzorku je 92×112 pixelů s 256 odstíny šedi na pixel.



obr. 5.1: Příklad z AT&T The Database of Faces

BANCA Database

Databáze [106] obsahuje vzorky 52 osob (26 mužů a 26 žen) zachycených při různých nastavení scény. Celkově zahrnuje 208 vzorků. Pro vytvoření databáze byly použity dvě videokamery a dva mikrofony o různých kvalitách. Zvuk z obou mikrofonů byl nahráván zároveň do dvou zvukových stop na pásku DV. V průběhu natáčení byly osoby dotazovány na čtyři otázky, na které odpověděly jednou pravdivě a jednou nepravdivě. Průměrná délka jednoho vzorku je 20 sekund. Velikost celé databáze je přibližně 4,7GB (jeden disk DVD). Všechny vzorky jsou uloženy jako sekvence barevných obrázků ve formátu PNG s rozlišením 720×576 pixelů. Pro vědecké účely je databáze dostupná za cenu 100 \$ pro komerční použití za 200 \$.



obr. 5.2: Příklad z BANCA Database

Cohn-Kanade AU-Coded Facial Expression Database

Databáze [100] obsahuje obrazové vzorky sta osob, pro každou osobu 5 sekvencí. Věk zaznamenaných osob se pohybuje v rozmezí 18 až 30 let, 65% osob v databázi tvoří ženy. Pro záznam byly použity dvě videokamery připojené k videorekordérům S-VHS. Jedna z kamer byla namířena přímo na pozorovanou osobu, druhá byla pootočena pod úhlem 30° . Každá osoba se pokouší vyvolat 23 různých obličejových výrazů, pokaždé vychází z neutrálního výrazu obličeje. Zá-

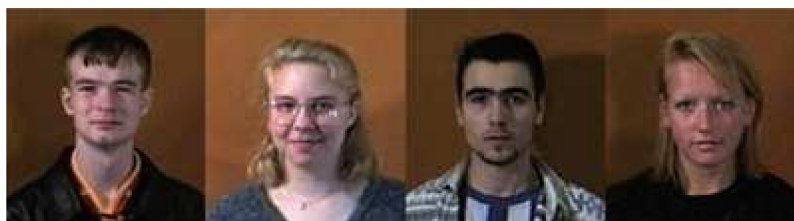
znam byl z pásek S-VHS digitalizován do klipů s rozlišením 640×480 pixelů v odstínech šedi. Databázi je po podepsání smlouvy možné využít k výzkumným účelům. V současné době (2008) jsou dostupné pouze vzorky z čelního pohledu.



obr. 5.3: Příklad z Cohn-Kanade AU-Coded Facial Expression Database

Collection of Facial Images

Obsah [101] databáze se skládá z obličejových vzorků 395 osob, pro každou osobu je zachyceno 20 obrazových vzorků, celkový počet vzorků je 7900. Věk osob v databázi se pohybuje v rozmezí od 18 do 20 let, zastoupeni jsou jak muži, tak i ženy. Vzorky jsou pořízeny před komplexním i jednotným pozadím. K vytvoření databáze byla použita videokamera se záznamem S-VHS. Jednotlivé vzorky jsou uloženy ve formátu JPEG s rozlišením 180×200 pixelů. Obsah celé databáze je pro výzkumné účely volně dostupný.



obr. 5.4: Příklad z databáze Collection of Facial Images

Georgia Tech face database

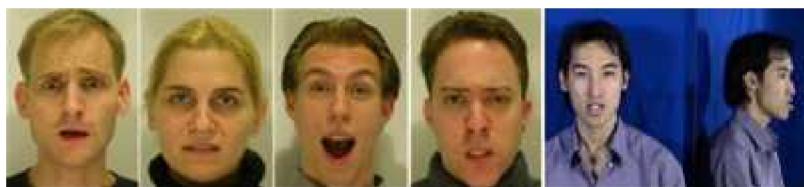
Obsahuje [99] vzorky padesáti osob zabíraných ve dvou až třech termínech v Centru zpracování signálů a obrazu Georgijského institutu technologie. Všechny osoby v databázi jsou reprezentovány patnácti barevnými obrázky ve formátu JPEG s rozlišením 640×480 pixelů a komplexním pozadím. Průměrná velikost obličeje v této databázi je 150×150 pixelů. Obrázky zobrazují čelní nebo boční pohled s různým výrazem v obličeji, světelnými podmínkami a měřítkem.



obr. 5.5: Příklad z Georgia Tech face database

MMI Face Database

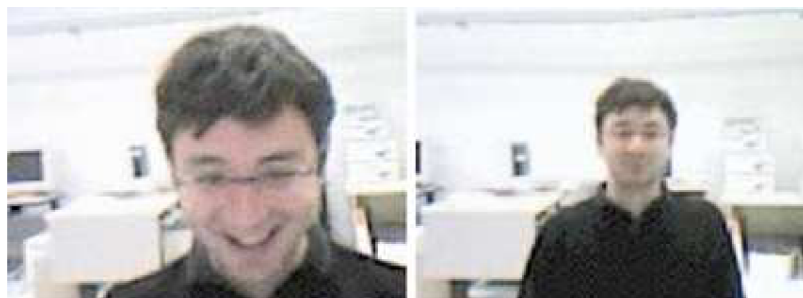
Obsahuje 740 statických obrázků a 848 videosekvencí 20 osob ve věku od 16 do 62 let. Všechny osoby jsou zachyceny z čelního pohledu, u některých je doplněn i boční pohled, jsou obsaženy vzorky s jednotným i komplexním pozadím. Videodatabáze [105] obsahuje mnoho výrazů obličejů jednotlivých osob. K záznamu videodatabáze byla použita videokamera formátu PAL. Rozlišení videosekvencí je 720×576 pixelů a jejich délka se pohybuje od 2 do 10 sekund. Videodatabáze je neustále rozšiřována.



obr. 5.6: Příklad z MMI Face Database

NRC-IIT Facial Video Database

Videodatabáze [104] obsahuje vzorky 11 osob, každá osoba je zastoupena dvojicí krátkých video klipů. Každý klip zobrazuje obličej uživatele počítače před monitorem. Zachycuje široký rozsah výrazů a orientace obličejů nahraný webovou kamerou nainstalovanou na počítačový monitor. Celkově obsahuje 22 vzorků. Průměrná délka jednoho vzorku je 15 sekund při snímkové frekvenci 20 snímků za sekundu. Rozlišení videosekvencí je 160×120 pixelů. Velikost celé videodatabáze je 28MB.



obr. 5.7: Příklad z NRC-IIT Facial Video Database

The BioID Face Database

Obsahuje obrazové vzorky 23 osob, pro každou osobu je obsaženo průměrně 66 vzorků. Celkově se v databázi [103] nachází 1521 obrazových vzorků, které jsou uloženy v odstínech šedi ve formátu PGM s rozlišením 384×286 pixelů. Vzorky v databázi obsahují komplexní i jednotné pozadí. Celková velikost databáze je 166MB a je volně dostupná.



obr. 5.8: Příklad z The BioID Face Database

University of Oulu Physics-Based Face Database

Databáze [108] obsahuje obličejové vzorky 125 osob, pro každou osobu jsou nasimulovány 4 různé světelné podmínky (umělé světlo, denní světlo, žárovka, zářivka) a 4 nastavení kamery. Pro každou osobu je tedy vytvořeno 16 vzorků zachycující čelní pohled na pozorovanou osobu. Vzorky databáze byly zachyceny pomocí tří čipové digitální videokamery. Rozlišení jednotlivých obrazových vzorků je 328×569 pixelů a jsou uloženy ve formátu BMP. Vzorky jsou zachyceny proti jednotnému šedému pozadí. Databáze je dostupná pro výzkumné účely, celý obsah je uložen na dvou CD. Databázi je možné objednat za cenu 35 €.



obr. 5.9: Příklad z The University of Oulu Physics-Based Face Database

The VidTIMIT Audio-Video Dataset

Databáze [109] se skládá z video a audio nahrávek 43 osob, přednášejících krátké fráze. Video sekvence jsou uloženy jako sekvence obrázků ve formátu JPEG s rozlišením 512×384 pixelů. Audio sekvence jsou ve formátu WAV. Databáze byla nahrávána ve třech termínech. Věty byly vybrány z testovací části TIMIT corpus. Každá osoba přednáší deset frází. Prvních šest frází bylo zachyceno v prvním termínu, další dvě ve druhém a poslední dvě fráze ve třetím termínu. První dvě fráze jsou pro všechny osoby stejné.



obr. 5.10: Příklad z The VidTIMIT Audio-Video Dataset

The Yale Face Database

Obsahuje 165 vzorků 15 osob ve formátu GIF v odstínech šedi. Rozlišení jednotlivých vzorků je 320×243 pixelů. V databázi [110] je 11 vzorků na osobu, jeden na každý výraz obličeje nebo nastavení kompozice: čelní osvětlení, boční osvětlení s brýlemi, bez brýlí, veselý, smutný, ospalý, překvapený a mrkající. Databáze je pro nekomerční použití veřejně dostupná na internetu.



obr. 5.11: Příklad z The Yale Face Database

The Yale Face Database B

Databáze [111] obsahuje 5760 vzorků deseti osob zachycených v 576 konfiguracích scény (devět výrazů a 64 druhů osvětlení). Pro každou osobu byla v jednotné pozici zaznamenána mapa okolního osvětlení. Z toho důvodu je celkový počet ob-

rázků ve výsledku 5850. Vzorky mají rozlišení 640×480 pixelů a jsou uloženy v odstínech šedi. Celková velikost komprimované databáze je přibližně 1GB.



obr. 5.12: Příklad z The Yale Face Database B

tab. 5.2: Vlastnosti veřejně dostupných videodatabází

zkratka databáze	počet osob [-]	délka [snímků/os.]	rozlišení [pixelů]	velikost	barevné vzorky	pozadí
AT&T	40	10	92×112	7,1 MB	NE	jednotné
BANCA	52	500	720×576	4,7 GB	ANO	jednotné/komplexní
Cohn-Kanade	100	-	640×480	-	NE	jednotné
Collection	395	20	180×200	48,5 MB	ANO	jednotné/komplexní
Georgia	50	15	640×480	128 MB	ANO	komplexní
MMI	19	83	720×576	-	ANO	jednotné/komplexní
NRC-IIT	11	± 15 sec	160×120	28 MB	ANO	komplexní
BioID	23	66	384×286	166,6 MB	NE	jednotné/komplexní
Oulu	125	16	328×567	2×CD	ANO	jednotné
VidTIMIT	40	± 60 sec	512×384	3 GB	ANO	jednotné
Yale	15	11	320×243	6,1 MB	NE	jednotné
Yale B	10	576	640×480	1 GB	NE	komplexní

5.2 Dostupné obrazové databáze pro účely detekce mrkání

Jelikož je detekce mrkání nepatrnou součástí detekce a rozpoznávání obličejů, nelze předpokládat přílišný zájem vývojářů o zahrnutí mrkání při návrhu obrazové databáze.

5.2.1 Popis databází pro účely detekce mrkání

ZJU Eyeblink Database

Databáze videosekvencí [112] obsahuje 80 klipů 20 jedinců v AVI formátu (nekomprimovaný RGB formát), natočených kamerou Logitech Pro 5000. Pro každou

osobu jsou k dispozici 4 videosekvence: čelní pohled bez brýlí, čelní pohled s brýlemi s tenkými obroučky, čelní pohled s brýlemi s černými skličky a poslední sekvence, kdy se osoba dívá směrem vzhůru a nemá brýle. Všechny videosekvence obsahují spontánní mrkání snímaných osob. Rozlišení obrazu je 320×240 pixelů při frekvenci 30 snímků za sekundu. Každý vzorek trvá asi 5 sekund. Počet mrknutí se liší vzorek od vzorku od jedné do počtu šesti mrknutí. Celkem databáze obsahuje 255 případů mrknutí. Data byla natáčena v místnosti bez dodatečného osvětlení.



obr. 5.13: Příklad z ZJU Eyeblink Database

Canadian Forces - Combat Camera Video Archive

V tomto případě se nejedná o databázi, ale o archív videosekvencí, jehož účel není vědecký. Tento archív [113] patří Ministerstvu obrany Kanady a jedná se o záběry bojové kamery kanadských vojáků z vojenských misí. Některé záznamy obsahují až dvouminutové monology vojáků přímo dívajících se do kamery (např. různé vzkazy rodinám). Osoby na těchto záběrech hovoří přirozeně a tedy i spontánně mrkají. Proto se tento archív hodí jako zajímavý zdroj pro vědecké účely této práce, například pro odhad frekvence mrkání.

Nasbíráno bylo 35 různých vzorků videosekvencí s délkami 15 sekund až 2 minuty. Rozlišení obrazu bylo z důvodu příliš vysoké kvality sníženo na 320×240 pixelů. Videosekvence samozřejmě obsahují i zvukovou stopu.



obr. 5.14: Příklad z ZJU Eyeblink Database

Databáze SKY

Protože žádná z obecně dostupných databází nespĺňuje požadavky pro měření všech událostí detekce mrkání (např. frekvence mrkání), bylo přistoupeno k vývoji

vlastní videodatabáze pro měření obličejových rysů se zaměřením pro výzkum v oblasti detekce mrkání.

K zachycení obrazových vzorků byla použita poloprofesionální digitální videokamera CANON XM1. Videosekvence se na DV pásku zaznamenávalo s rozlišením 720×576 pixelů a snímkovou frekvencí 25 snímků za sekundu. Natáčení videodatabáze probíhalo ve třech termínech a byly pořízeny vzorky celkem 30 osob. Při prvním natáčení bylo zvoleno komplexní pozadí, další natáčení probíhalo před jednotným bílým pozadím. V průběhu natáčení byla videokamera nastavena v automatickém režimu. U 21 vzorků, které byly natočeny před bílým pozadím, bylo provedeno vyklíčování obličeje z pozadí a jeho umístění před uměle vytvořené pozadí komplexní¹⁰. Tímto došlo k navýšení počtu vzorků ve videodatabázi na 51. Délka všech vzorků videodatabáze včetně vyklíčovaných vzorků je 6 hodin 28 minut, průměrná délka jednoho vzorku je 7 minut 19 sekund. V průběhu natáčení osoby odpovídaly na 29 jednoduchých otázek. Surová data o velikosti 44 GB byla komprimována na 917MB při rozlišení 384×288 pixelů formátem WMV (Windows Media Video).



obr. 5.15: Příklad z SKY Database

¹⁰ Umělé přimíchání komplexního pozadí bylo z důvodu využití databáze např. pro testování algoritmů detekce obličeje. Lidským zrakem lze sice rozpoznat mírně zkrácené obrysy osob, nicméně pro navržené algoritmy to může znamenat nepředstavitelné problémy.

6 Detekce mrkání bez využití IR technologie

V této kapitole je popsán návrh systému pro detekci mrkání využívající číslicové zpracování obrazu bez využití IR technologií. Návrh je založen na základu dvou algoritmů. Prvním z nich je algoritmus dle [78], kde je použita metoda ADF (accumulated difference frame). Druhým je algoritmus [23], který má pomáhat lidem s omezeným pohybem ovládat počítač, který následně využívá rozdílnost v jednotlivých snímcích.

6.1 Hardwarová realizace

Jelikož tento přístup není založen na žádné přídavné technologii a nevyžaduje další zařízení kromě počítače pro zpracování dat, obecně není potřeba hardwarové požadavky řešit. Jedinými požadavky tedy jsou počítač s dostatečnou výpočetní silou a zdrojová data (nízké, střední nebo vysoké kvality).

6.2 Softwarová realizace

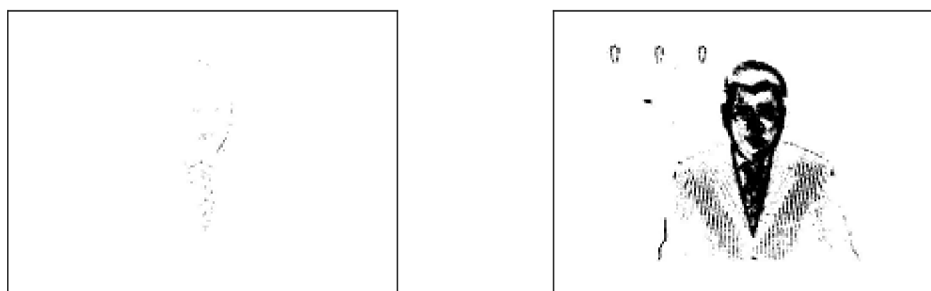
Návrh algoritmu (program) je rozdělen do několika částí, tak aby se na jednotlivých částech dalo samostatně pracovat s výstupy předchozích podprogramů (funkcí). Rozdělení programu do několika podprogramů bylo nutné i z časových důvodů, protože prostředí MATLAB nemůže načíst větší část videosekvence, jelikož se ukládá do nekomprimovaných matic, což je výrazně paměťově náročné. Videosekvence se zpracovává postupně po částech o velikosti 10-20 snímků. Načtení videosekvence do MATLABu a následné celkové zpracování dat je rovněž velice náročné na čas. V následujících podkapitolách jsou popsány jednotlivé kroky těchto podprogramů a jejich vlastnosti.

6.2.1 Návrh softwarové realizace

Funkcí hlavního algoritmu je, aby načítal vstupní videosekvence, inicializoval proměnné, které se budou v hlavním programu zobrazovat. Načítání vstupu probíhá postupně (podle délky videa dělené velikostí segmentu). Takže pokud je délku segmentu zvolena 10 snímků, ze vstupu se vyjme 10 snímků, které se uloží do matic, které jsou zpracovávány dalšími volanými funkcemi.

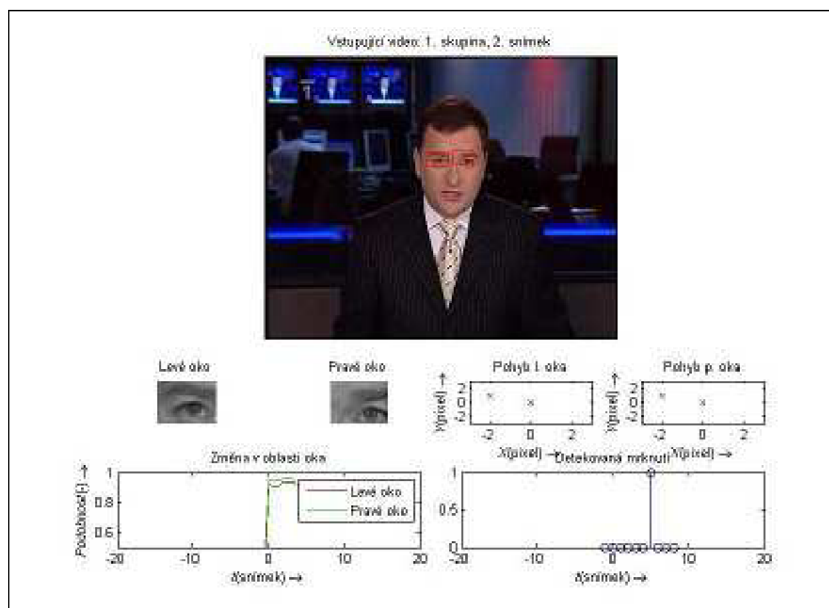
Délka segmentu je prvním parametrem ovlivňujícím funkčnost celého algoritmu, jelikož se v dalším stupni počítá akumulární diferenční matice, délka segmentu zásadně ovlivní to, jak bude vypadat diferenční akumulární rámec. Délka segmentu ovlivňuje akumulární rámec v závislosti na rozmezí pohybu sledovaného člověka. Například, pokud je zvolen malý segment a pohyb člověka nebude dostatečně výrazný, akumulární rámec bude mít velice malé hodnoty, takže další algoritmus na detekci oblasti očí nebude schopen označit oblast, kde se nacházejí oči (obr. 6.1a). Naopak pokud je zvolen segment ke zpracování příliš velký a subjekt se bude během této doby výrazněji pohybovat, oblasti změny budou velké s velkými hodnotami v jednotlivých pixelech (obr. 6.1b).

Algoritmus je náchylný zvláště pokud se jedná o pohyb v horizontální ose, jelikož jako testovací zdroj byly použity televizní zprávy uloženy v archivu v horší kvalitě (jak je uvedeno v kapitole 3.5.1, jedná se o nízkou kvalitu obrazových dat), kde je tmavé pozadí, může stát, že jednotlivé oblasti se překryjí a stane se z nich jedna.



obr. 6.1: Akumulární diferenční rámec se zvoleným malým segmentem (a) a velkým segmentem (b).

Další úlohou algoritmu je prezentovat hodnoty vrácené funkcemi, které měly za úkol detekovat mrknutí. Prezentační plochu (obr. 6.2) je rozdělena na čtyři sloupce a čtyři řádky. V horní polovině je zpracovávané video s vyznačenou oblastí detekovaných očí z akumulárního rámce. Dále jsou vlevo pod videosekvencí zobrazeny vyjmuté oční oblasti ve stupních šedi, vpravo je vyobrazen detekovaný posun levého a pravého oka. Vlevo dole je průběh velikosti podobnosti oblasti očí. A vpravo dole jsou vykreslena detekovaná mrknutí.



obr. 6.2: Plocha prezentace výsledků programu.

Data jsou předzpracována po segmentech, potom jsou zobrazena po jednotlivých snímcích a průběhy výsledných hodnot jsou posouvány, tak že na nule je vždy aktuální hodnota. Postup algoritmu je podle návrhu obr. 3.3.

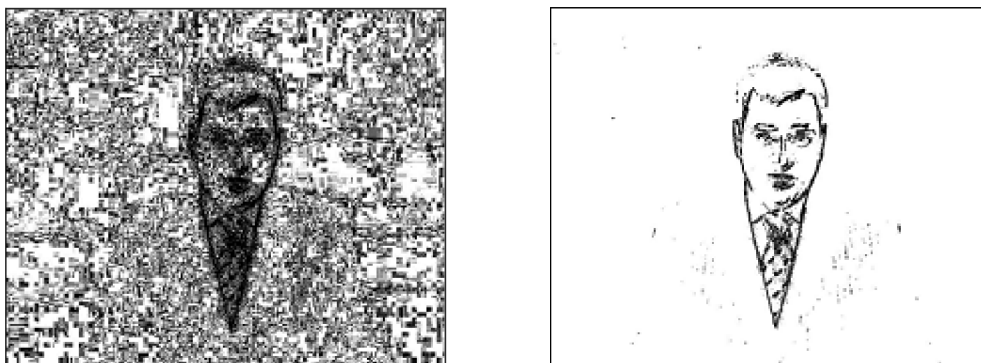
6.2.2 Funkce Detekce

Další část procesu shrnuje všechny detekční algoritmy celého programu. Na začátku procesu se převede barevné video na čtyřdimenzionální matici normovanou na zlomek jednotlivých barevných složek, tzn. každou hodnotu dělenou maximální možnou hodnotou jasu (255). Po inicializaci dalších proměnných převádíme matici na třídimenzionální se stupni šedi.

Největší část programu se zpracovává uvnitř smyčky, která prochází video od druhého snímku do snímku určeného velikostí vyjmutého segmentu. Prvním příkazem uvnitř smyčky je odčítání aktuálního snímku od předchozího, proto smyčka začíná druhým snímkem. Matice vzniklá odečtením dvou snímků se prahuje s hodnotou, která byla přímo zvolena jako $4/255$.

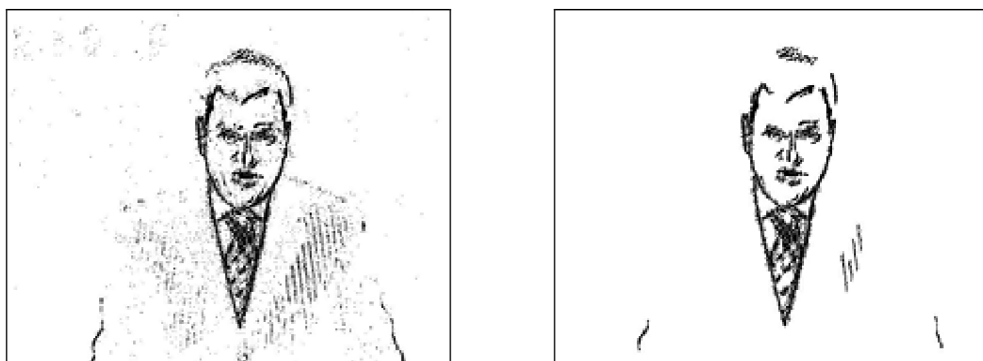
Hodnota prahu pro převod diferenční matice na binární obraz je dalším parametrem programu a pro jeho úplnou automatizaci by se tato hodnota měla v programu počítat pomocí funkcí v závislosti na míře pohybu ve videosekvenci, anebo za použití statistických metod aplikovaných na rozdíl dvou odečtených rámců. Nastavení správného prahu může velice urychlit práci a zvýšit procentuální šanci na správnou detekci. Nastavení vyšších hodnot prahu ($2/255$ a více) je nutné také kvůli odstranění šumu, který vznikl kvantizací a ve větší míře kompresí videa (obr. 6.3). Hodnot prahu byla zvolena na základě vzorků čtyř videosekvencí, tak aby byla co možná nejmenší kvůli velké ztrátě informací, ale dostatečně velkou na odstranění šumu nebo

jiných nepatrných projevů změny v jednotlivých pixelech snímků. Rozdíl mezi správným a špatným nastavením je patrný z obr. 6.3.



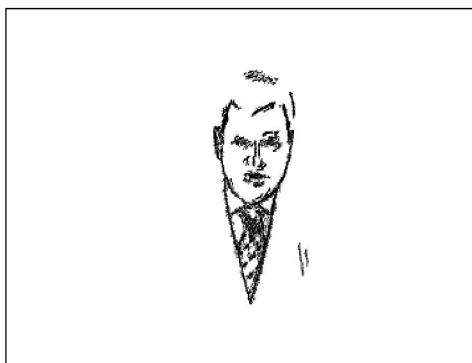
obr. 6.3: Obraz diferenční matice se špatným nastavením (a) a dobrým nastavením (b) prahu.

Po binarizaci rozdílové matice je dalším krokem „odstranění malých oblastí z obrazu“ – odfiltrování oblastí s menší plochou než 80 pixelů, což výrazně diferenční matici pročistí (obr. 6.4). Odstranění malých oblastí probíhá na základě sousedství kladných pixelů.



obr. 6.4: Obraz diferenční matice před filtrací (a) a po odstranění malých oblastí (b).

Potom jsou binární data diferenční matice analyzována jako oblasti dotýkající se se pixelů. Pro každou oblast kladných hodnot jsou nalezeny sloupce, v kterých se nachází kladná hodnota, a podle nich jsou určeny průměrné hodnoty sloupců pro tuto oblast. Tyto průměrné hodnoty sloupců jsou porovnávány s číselnou hodnotou levé a pravé hranice obličeje, a ty binární oblasti, které leží mimo hranici, jsou odstraněny (rozdíly mezi obr. 6.4b a obr. 6.5). Meze obličeje jsou určeny pomocí horizontální hustoty.

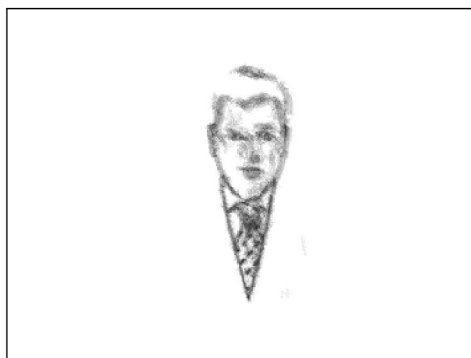


obr. 6.5: Obraz diferenční matice po odstranění oblastí mimo oblast zájmu.

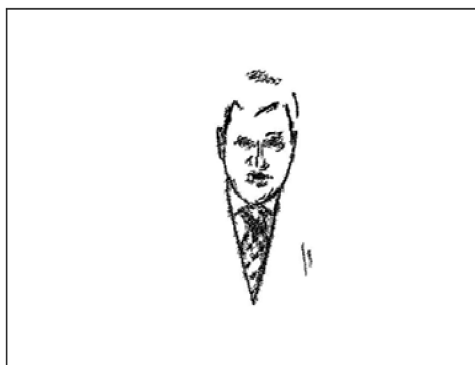
Takto upravená binární matice se po normování počtem snímků přičítá k akumulární diferenční matici. Přičítání k akumulární diferenční matici probíhá tolikrát, kolik je do funkce na vstupu vloženo snímků. Pro testované videosekvence bylo zvoleno 10 a 20 snímků, podle míry pohybu. Po ukončení smyčky se zavolají podprogramy na hledání očí a detekce změny, jejichž vstupem je akumulární rámec pohybu. Vznik akumulárního rámce z jednotlivých diferenčních matic (obr. 6.6) je zobrazen na obr. 6.7.



obr. 6.6: Jednotlivé nenormované diferenční snímky v pořadí 2 až 9.



obr. 6.7: Akumulační diferenční matice, která vznikla sečtením normovaných diferenčních snímků.



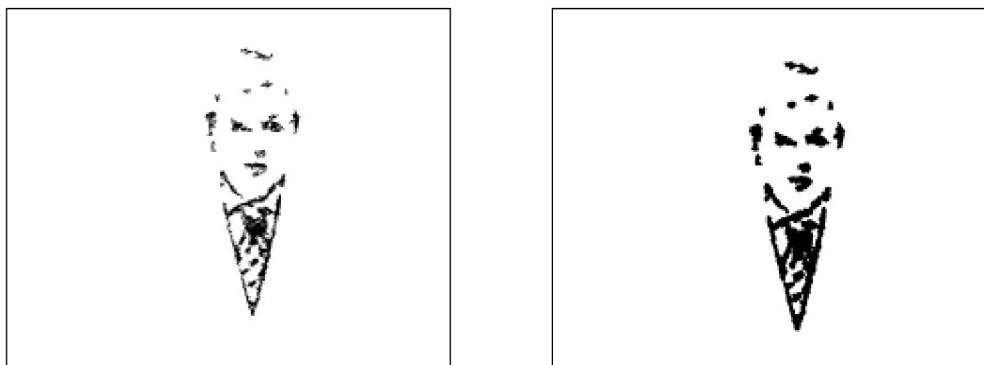
obr. 6.8: Prahovaná akumulční matice s prahem 0,4.

6.2.3 Funkce Lokalizace očí

Funkce na hledání očí je nejdůležitější částí celého programu na detekci mrknutí. Jeho úkolem je detekovat pozici očí ze vstupních dat, což je akumulční rámec a číslo obsahující velikost segmentu videosekvence, které se zpracovává.

Prvním krokem programu je výpočet parametru pro práh, od kterého se bude počítat hodnota pixelu jako logická 1. Tedy podobně jako u předešlé kapitoly, zde nedochází k binarizaci jednotlivých diferenčních snímků, ale k binarizaci celé akumulční matice (Tato matice jsem v programu nazvaná jako *bwSum*. Dalším rozdílem je, že tato data neznázorňují jen velikost změny v pixelu, ale také v kolika snímcích se pixel nachází.

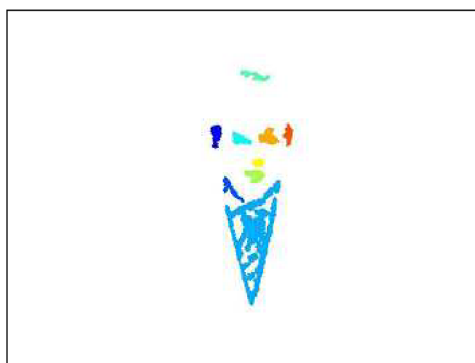
Jako dalšímu zpracování se *bwSum* podrobí filtraci, ve které jsou odstraněny oblasti menší než 20 pixelů. Kvůli vyplnění a rozšíření oblastí byla použita dilatace kruhem o průměru 2 pixely, která kruhové okolí logické jedničky vyplní jedničkami viz. obr. 6.9.



obr. 6.9: Filtrace *bwSum* před dilatací (a), po dilataci (b).

Odstraněním oblastí, které se pomocí dilatace nespojily s ostatními z obrázy odstraníme – oblasti menší než 200 pixelů. Odstraněním těchto oblastí je také urychlen chod programu, jelikož potom není nutné procházet tak velké množství dat při jejich analýze. Po filtraci a odstranění malých oblastí z *bwSum* jsou získána, která jsou už vhodná k rozpoznávání (obr. 6.10).

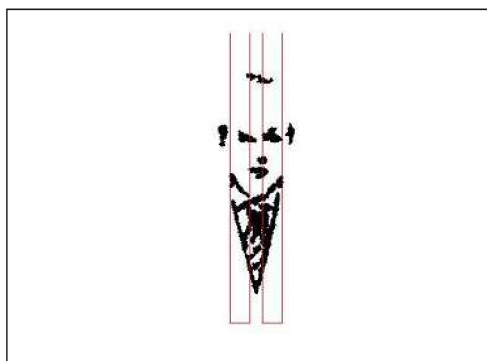
Analýza oblastí probíhá z toho důvodu, abych bylo zjištěno, kde se nachází levá a pravá strana obličeje. Levá část je zajištěna pomocí nejnižšího sloupce a pravá pomocí nejvyššího. Rozbor oblastí je už jednou použit ve funkci Detekce k filtrování pixelových oblastí a je použit ještě několikrát. Vstupní data tvoří matice, která obsahuje hodnoty od jedné až do počtu jednotlivých pixelových oblastí podobně jako na obr. 6.10, kde rozdílné hodnoty jsou převedeny na odlišné barevné odstíny. Analýza oblastí probíhá ve smyčce, která určuje, s jakou oblastí se právě pracuje. Dále jsou vyhledány všechny sloupce a řádky, přes které se oblast rozprostírá. Hodnoty řádků a sloupců jsou zprůměrována a uložena.



obr. 6.10: Obraz rozdělených oblastí.

Dalším krokem k určení pozice očí je vzdálenostní poměrování vlastností jednotlivých oblastí, tedy zda se některá z oblastí nachází v rozmezí v určitých horizontálních mezích. Tyto proporce obličeje jsou odhadnuty na základě několika vzorků akumulčních map a vyšly hodnoty 2/12, 5/12 pro levé oko a symetricky pro pravé oko 7/12, 10/12 (obr. 6.11). Tyto hodnoty jsou pomocí zjištěných hodnot pravé a

levé strany obličeje porovnávají s průměrnými sloupci oblastí. Jelikož dochází jen k porovnávání v horizontální ose, může se stát, že budou do mezí spadat i jiné oblasti z horní části hlavy nebo takové, které vznikly pohybem šatů, jak je vidět na obrázku část límce. Následně je oblast testována, zda její poměr výšky a šířky je v rozmezí jedna 1:1 až 1:2.



obr. 6.11: Zobrazení mezí pro určení očí.

Špatná detekce zvláště vadí v případech, když je vybrána oblast jen v jednom rozmezí levém či pravém a potom nedojde ke správnému určení osy obličeje. Hlavní idea této části programu je, že osa obličeje se nachází přesně mezi očima. A pomocí osy jsme schopni dále symetricky zpracovávat obraz.

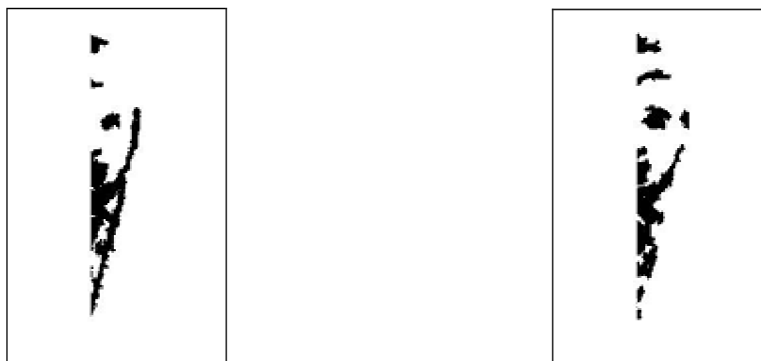
V prvním návrhu algoritmu přímou detekci oblastí očí nebyla použita, protože osu obličeje byla určována pomocí celkového průměrného sloupce, což se při testování na více videosekvencích ukázalo jako nevhodné řešení (obr. 6.12). V konečné fázi programu se osa obličeje pořád určuje pomocí průměrného sloupce, ale pokud bude tento test na meze oblastí pozitivní, tak se hodnota osy upraví na střed mezi oblastmi detekované jako oči. Další podmínkou pro úpravu hodnoty osy je, že levá oblast se výrazně neodlišuje od pravé oblasti svou průměrnou pozicí ve vertikální ose.



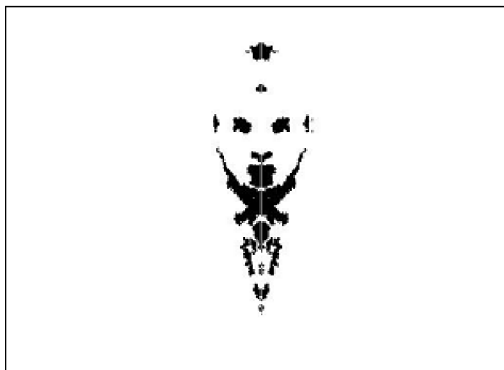
obr. 6.12: První fáze určení středu obličeje, průměrným sloupcem.

Pokud je určen střed obličeje, tak pro další zpracování je vypočítána vzdálenost stran obličeje zleva i zprava od středu a otestováno, zda obě hodnoty nepřesahují velikost matice. Následně je testováno, zda hodnoty osy obličeje a vzdálenost od ní vlevo a vpravo se vejdu do rozměrů obrazu, a když ano, obraz je dále zpracován podle jeho souměrnosti. Když nedojde ke správnému výpočtu hodnot rozměrů obličeje je další zpracování matice přeskočeno.

Další zpracování probíhá rozdělením obrazu na dvě části, tak jak byl obličej rozdělen jeho osou (obr. 6.13). Po rozdělení obrazu na dvě části se jeho levá polovina převrátí a vynásobí s pravou polovinou, čímž vznikne středově vyrovnaný obraz (obr. 6.14).

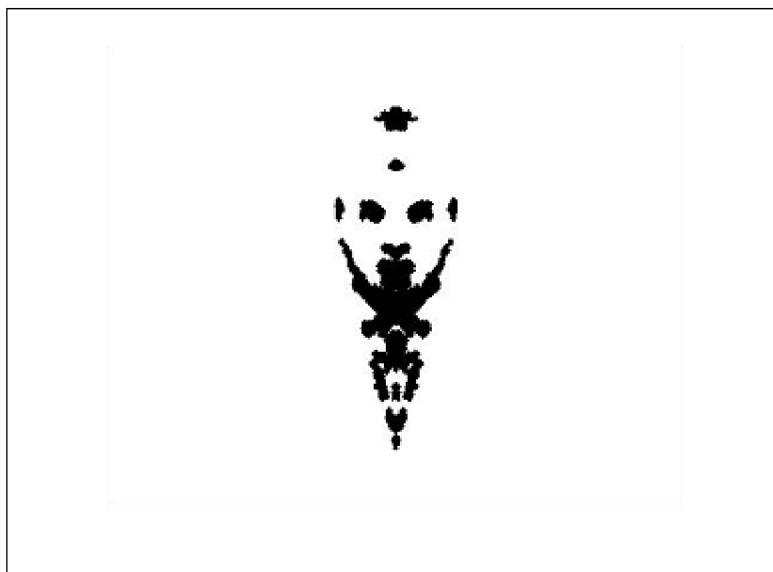


obr. 6.13: Rozdělený obraz levá převrácená strana (a), pravá strana (b).



obr. 6.14: První fáze určení středu obličeje, průměrným sloupcem.

Součin obou polovin obrazu se spojí v jeden binární obraz, který slouží jako binární maska původní akumulární matice. To znamená, že po úpravě středově souměrného obrazu tak, aby z něj vymizela mezera označující osu užitím dilatace, se tím vynásobí původní akumulární rámec (obr. 3.18). Použitím středové souměrnosti při zpracování obrazu je docíleno odstranění nežádoucích výčnělků oblastí a také rozdělení nesouvisejících oblastí, což je nezbytné k přesnému vytyčení očí.



obr. 6.15: Výsledek vynásobení dilatované masky a akumulárního rámce.

Po získání binárního obrazu daného středově souměrnou maskou násobené akumulárním rámcem, tento obraz je podroben další oblastní analýze, která vypočítá jejich průměrné sloupce. Potom se každá oblast podrobí testování, zda jejich průměrný sloupec spadá do horizontální oblasti určené jako rozmezí $2/12$ až $5/12$ z celé šíře obličeje. Testuje se jen levé oko z toho důvodu, že obraz je symetrický, takže oblast pravého oka lze jen dopočítat.

V cyklu na testování oblastí se ukládají souřadnice obdélníku, který obkresluje hranice oblasti a při splnění podmínky pro oblast se z cyklu vyskočí. Získanou oblas-

tí posléze násobíme původní akumulční rámec a s pomocí výsledných dat vzniklých násobením je průnikem s původním rámcem určena oblast, která zachycuje celkovou pozici oka v průběhu části zpracovávaného videa. Okolí oka je definováno taktéž jako obdélník určený z oblasti původního akumulčního rámce nejnižším a nejvyšším sloupcem a řádkem.

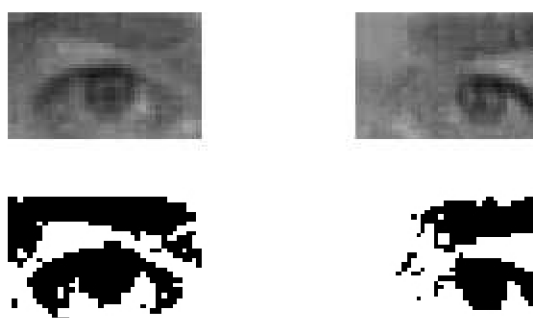
Na závěr algoritmu se testuje, zda jsou výstupní souřadnice obdélníků správné. Pokud nejsou správné, jsou všechny vynulovány pro pozdější snazší určení toho, jaké byly použity souřadnice z minulého hledání očí.

6.2.4 Podprogram Detekci změny oblasti očí

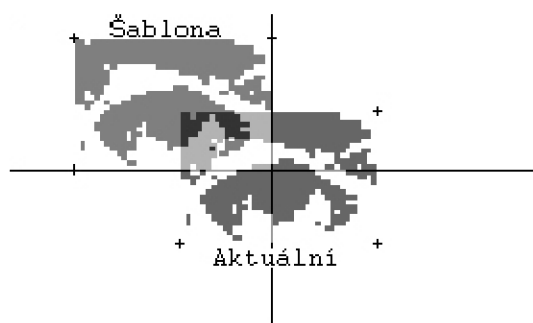
Úkolem tohoto podprogramu je sledovat k jaké změně došlo v okolí očí. Používá se k tomu detekce změny pixelů v oblasti pohybujícího se oka z aktuálního a předchozího snímku.

Prvním návrhem, jak detekovat mrknutí bylo určit kruh opisující duhovku a sledovat jeho úplnost. Ale při jeho implementaci na vzorek videosekvence bylo zjištěno, že pro tuto detekci je zapotřebí vysoce kvalitní videomateriál s rozlišením více než 30 pixelů na šíři duhovky, což je nejméně trojnásobné rozlišení. Proto jako metodu detekce bylo zvoleno porovnávání šablony s maticí reprezentující oblast oka.

Vstupní data jsou tvořena částí zpracovávané videosekvence ve stupních šedi, souřadnicemi čtverců obklopující oči a prahem, kterým je převeden šedý obraz na binární (obr. 6.16). V cyklu na detekci změny oproti šabloně dochází k porovnávání matice binarizovaného oka z aktuálního snímku s předchozím, tak že je pohybováno šablonou (obr. 6.17) po aktuální matici, přitom obě násobíme a zaznamenáme nejvyšší hodnotu shody. Při nalezení maximální shody zaznamenáme vektor pohybu, který nastal posuvem šablony vzhledem k matici. Fáze pohybující se šablony probíhá jednou pro každé oko.

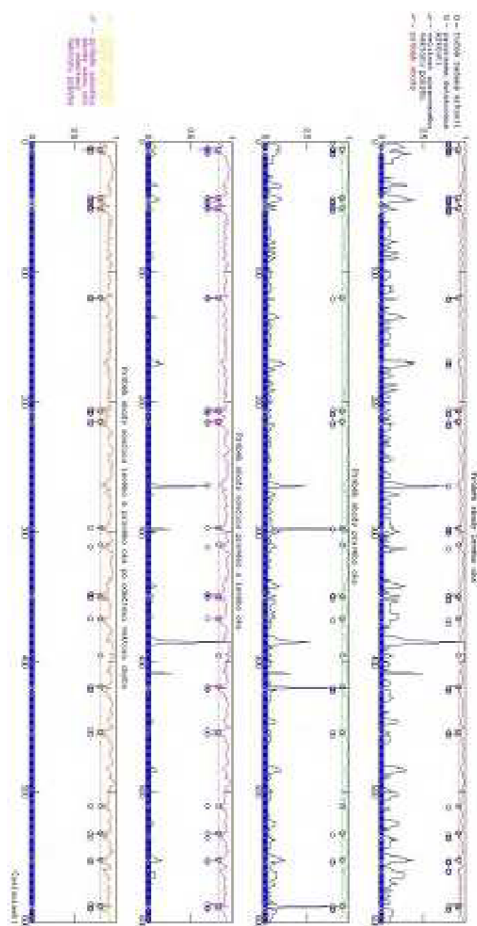


obr. 6.16: Oblast očí a jejich binární podoba s prahovanou jasovou hodnotou 0.4.



obr. 6.17: Počátek překrytí šablony s aktuální oblastí.

Výstup tedy tvoří hodnoty nejvyšší shody společně s vektorem pohybu aktuální matice oka vůči šabloně. Shoda je vyčíslená jako počet shodných pixelů a před ukončením algoritmu je normována celkovým počtem pixelů v matici.



obr. 6.18: Výstup podprogramu na detekci změny a jeho následné zpracování.

6.2.5 Podprogram Vyhodnocení mrknutí z průběhu shody

Poslední částí programu na detekci mrkání je funkce určená k vyhodnocení mrknutí z průběhu shody šablony s výřezem oka (obr. 6.18). Na obrázku se nachází vý-

stup podprogramu pro určení změny v oblasti levého a pravého oka. Druhým průběhem je velikost pohybu levého a potom pravého oka. Po vynásobení průběhů změn levého a pravého oka se odečtou hodnoty velikosti vektoru pohybu také vzniklých vynásobením normovaných hodnot pro levé a pravé oko. Tento nový průběh se testuje na pokles pod 80%-ní hranicí, a když tento stav nastane, je detekován jako mrknutí.

Po vyhodnocení, že došlo k mrknutí, se následující hodnoty pro další dva snímky automaticky ukládají jako „nemrkání“. Tím se ošetří chybná detekce vzniklá změnou ze zavřeného oka na otevřené.

Vyhodnocování mrknutí z průběhu pomocí testování poklesu pod určitou úroveň není vůbec robustní. Program je přizpůsoben k detekci rychlého a častého mrkání, které se v testovaném vzorku nachází. Pro detekci delších mrknutí by bylo třeba použít jednu šablonu k testování na více snímků. A místo detekce poklesu pod určitou úroveň by bylo lepší testovat tvar průběhu.

6.2.6 Podprogram na odstranění malých oblastí

Tento skript slouží k odstranění oblastí ze vstupní matice, které jsou menší než druhá vstupní hodnota funkce. Oblast je tvořena pixely, které se dotýkají aspoň v jednom z osmy směrů. Pro co nejrychlejší zpracování je skript plně vektorizován, tudíž nelze nakreslit jednoduchý vývojový diagram.

6.2.7 Časová náročnost programu

Účelem programu nebyla co nejrychlejší práce, to už bylo částečně dáno použitím prostředí MATLAB, u kterého načtení deseti snímků z videa trvá přibližně jednu sekundu na počítači s průměrnou konfigurací.

Časová náročnost podprogramu detekce tvoří polovinu celkové. Doba zpracování je zaprvé způsobena transformací dat z AVI formátu do čtyřrozměrné matice, která je posléze převedena na třírozměrnou ve stupních šedi. Nejvíce časově náročnou částí detekce je cyklus, kde probíhá zpracování snímků a ukládání jejich zpracovaných diferencí do akumulárního rámce. Cyklus tvoří polovinu časové spotřeby podprogramu detekce.

Čas spotřebovaný podprogramem určeným k hledání oka činí 2,2 sekundy. Takto velká hodnota je způsobena opakovaným použitím algoritmu na analýzu oblastí, u kterého je rychlost závislá na počtu a velikosti zpracovávaných oblastí, takže doba strávená hledáním očí je značně proměnlivá.

Podprogram na porovnávání výřezů oka pracuje ve smyčce, v které se prochází jednotlivými snímky, v tomto se nachází ještě další dva cykly pro posouvání šablony po výřezu. Uvnitř těchto cyklů je hlavní výkonná část na testování podobnosti. To znamená, že kubická náročnost algoritmu i s tak malou porcí dat zpracovávanými

informacemi, jakými jsou matice o rozměru 40 pixelů znatelně, prodlužuje dobu nutnou ke zpracování.

Vyšších jak kvadratických časových náročností je v programu několik a to se spojením s vektorizací a zpracováním velkého množství dat je důvodem k pomalé rychlosti programu. Ostatní podprogramy jako ten na vyhodnocení mrkání z průběhu změny a skript na odstraňování malých oblastí jsou z časového hlediska vcelku zanedbatelné, protože první zpracovává jen krátké úseky jednorozměrných hodnot. U druhého se čas potřebný na to, aby program odstranil malé oblasti, pohybuje od 100 milisekund do 200 milisekund.

6.3 Závěr

Dobrá identifikace mrknutí je závislá na dobrém nastavení parametrů, ale ani po dobrém nastavení parametrů stále není stoprocentní pravděpodobnost správné detekce. Prvním místem důležitým pro správnou detekci je program na hledání očí. Je to také nejslabší článek celého programu.

První možností jak může dojít k nesprávné detekci je, že oblast oka nebude vůbec v akumulacním rámci zastoupena, potom dojde k vybrání jiné oblasti v horizontálním rozmezí. Dalším možným stavem, který může nastat, je, že oblast oka bude spojena s další oblastí, což posune její těžiště mimo horizontální rozmezí, v kterém by měla být umístěna oblast oka. Pokud dojde k nesprávnému určení očí a označí se jiná oblast, která splní všechny podmínky, program předá souřadnice této oblasti dál a podprogram na detekci změny potom hledá mrknutí tam, kde se oko vůbec nenachází. Když nastane druhá možnost, že není identifikována žádná oblast, podprogram detekce potom použije souřadnice z minulého úspěšného hledání oka. Obě tyto možnosti mohou nastat po sobě, to potom znamená, že program po neúspěšném hledání oblasti použije minulé souřadnice, které jsou vadné.

Podprogram určený k detekci změny zase byl nastaven na poměrně rychlé mrknutí, takže pokud došlo k pomalému mrknutí, nedošlo k jeho detekci. Nastavení pevné hranice asi není nejlepším řešením, proto algoritmus není dostatečně robustní.

Vyhodnocování mrknutí nebylo zcela funkční při testování na videosekvenci s moderátorkou s tmavými očními stíny, protože při prahování výřezu okolí oka je jako kladná logická hodnota vyhodnoceno velké množství pixelů. Kvůli velkému počtu pixelů v šabloně i výřezu nedochází při mrknutí k výraznějšímu procentuálnímu skoku změny. A protože se podobnost šablony s výřezem pohybují nad 90%, tak nejsou vyhodnocena mrknutí, protože nedojde k poklesu průběhu pod definovanou hranici.

Při zkoušení přizpůsobeného algoritmu na první videosekvenci byl algoritmus úspěšný při hledání na 60 segmentech ze 130, jinak byly použity souřadnice minulých detekcí. Při nastavení parametrů a testování na druhé videosekvenci byl algoritmus ještě méně úspěšný, kvůli očním stínům byly oblasti pohybu očí velké a často

se spojovaly s ostatními. V třetí videosekvenci se moderátor málo pohyboval a pohyb očí tak nebyl dostatečně zřetelný a špatně se detekovaly.

Výsledky programu na vyhodnocení mrkání byly při správném určení okolí oka s 80% procentní úspěšností.

7 Detekce mrkání s využitím IR technologie

V této kapitole je popsán návrh systému pro detekci mrkání využívající IR technologii a základní metody zpracování obrazu.

7.1 Hardwarová realizace

Základ je podobný jako v případě realizace detekce mrkání bez využití IR technologie, jinak celá hardwarová realizace týkající se IR technologie je relativně jednoduchá a účinná. Hlavními částmi jsou webová kamera, 16 LED diod zářících infračerveným světlem a zdroj napájení v podobě jedné AA baterie.

7.1.1 Náležitosti hardwarové realizace

Kamera

Výchozím typem kamery pro celý tento přístup byla webová kamera. Tento typ kamery, nehledě na výrobce, je velmi dobře použitelný, k testovacím účelům. Ke kladům patří cena kamery, kompaktní rozměry, jednoduchost její výroby a možnost propojení s USB porty počítače.

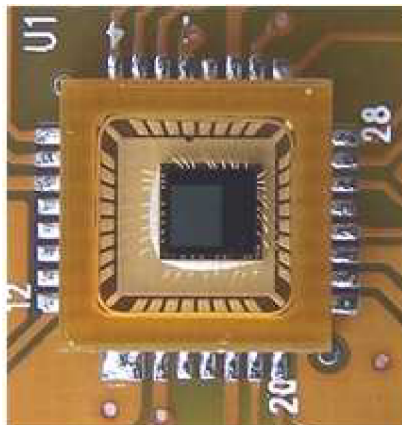


obr. 7.1: Webová kamera

Použitý CCD snímač

Na obr. 7.1 je vidět CCD snímač použité kamery. Tento snímač je založen na typu FT (Field Transfer) [64] a skládá se ze dvou shodných částí, snímací a zatemněné paměťové části. Po akumulaci se náboje přesunou ze snímací do paměťové části. Jednotlivé světločivné buňky se přitom využívají k přenosu náboje jako posuvné

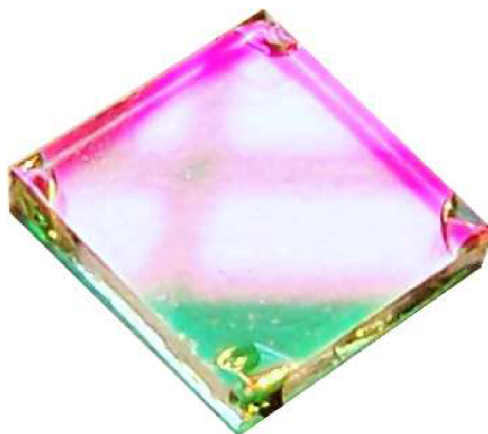
registry. Nevýhodou tohoto typu snímače je mazání obrazu ve svislém směru při snímání velkých jasných ploch, má ovšem maximální využití plochy snímače. [10]



obr. 7.2: CCD snímač použité webové kamery

Úskalí použití webové kamery

Obecně většina těchto kamer snímá také obraz v infra oblasti. Aby u kamery infra osvětlení fungovalo, je nutné otestovat, zda je čočka opatřena RGB filtrem. U černobílých kamer se tento filtr obvykle nenachází, proto pokud je použita černobílá kamera, bude infra osvětlení s velkou pravděpodobností fungovat. Barevné kamery ovšem tímto filtrem vybaveny být musí, protože je citlivost snímacího čipu nerovnoměrná. Funkčnost kamery pro snímání IR záření lze ověřit jednoduchým způsobem. Stačí namířit do čočky kamery dálkové ovládání a vyslat například volbu programu. V tu chvíli je vidět na monitoru blikající světlo infračerveného záření, které lze pouhým okem stěží zahlédnout. I přesto ale nemusí být problém ještě vyřešen. Kamera totiž nad CCD snímačem skrývá výše zmíněný RGB filtr, většinou nekvalitní plastový, který může neúplně potlačovat IR záření. To se projevuje například tím, že přímé IR záření do kamery viditelné je, ovšem jeho odrazy (od očí, obličeje, okolí) již viditelné nejsou. Proto je nutno tento filtr nahradit jiným filtrem (skleněným), nebo jako v případě tohoto návrhu nenahrazovat žádným, což ovšem vede k obrazu s deformovanými barvami (a tedy nutnosti použití desaturovaného obrazu). [114]



obr. 7.3: CCD snímač použité webové kamery

7.1.2 Možnosti hardwarové realizace

Následující možnosti hardwarové realizace byly čerpány z publikací [24] a [34], jejichž autoři (Flickner, Moromoto) a další jsou uznávanými vědci v tomto oboru.

Základní přístup

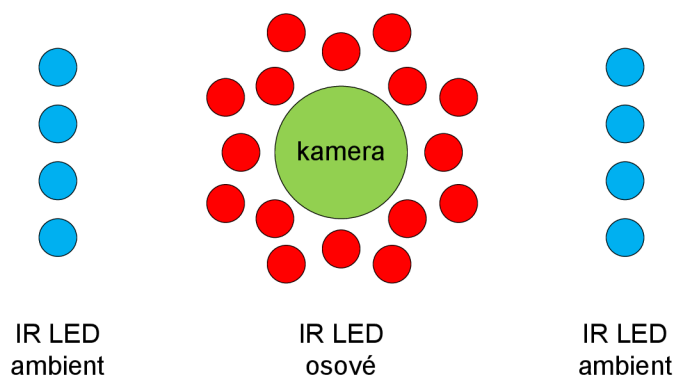
- kamera se stále svítícími IR LED diodami nejbližší ose snímání
- nejjednodušší realizace
- závislost na světelných podmínkách (detekce je slabší při horším osvětlení, vychází z fyziologie oka, rozšíření zorničky)
- vhodné doplnit dalšími náročnějšími metodami

Pokročilý přístup

- kamera s IR LED diodami nejbližší ose snímání, diody blikají synchronizovaně s expozicí kamery
- nutnost synchronizace blikání IR LED s kamerou
- znatelné zvýšení účinnosti
- není nutné přidávat další náročné metody

Zdokonalený pokročilý přístup

- kamera s vnitřním a vnějším kruhem IR LED diod, kruhy se střídají synchronně s expozicí
- vnější kruh (vzdálenější od osy snímání) přidá do snímané oblasti tzv. ambientní (okolní) osvětlení, které zajistí ještě účinnější zpracování než v pokročilém přístupu.



obr. 7.4: Možnosti hardwarové realizace

Možnosti synchronizace

Problém: kamera mění automaticky frekvenci snímání s měnícími se podmínkami, z toho plyne nemožnost použití pevně dané frekvence blikání IR LED diod.

Řešení:

- nastavení konstantní frekvence snímání za jakýchkoliv světelných podmínek v ovladačích kamery
- synchronizace přímo z kamery, neúčinnější (okamžitá odezva), obtížné nalézt kontakty, na kterých se provádí expozice CCD snímačem
- synchronizace pomocí rozboru přijímaných dat, pravděpodobně nerealizovatelné, Matlab komunikuje s kamerou asynchronně

7.1.3 Konečná hardwarová realizace

Ke konečné hardwarové realizaci byla vybrána možnost „základního přístupu“, tedy přisvětlování infračerveně stále svítícími diodami rozmístěnými ve dvojitěm kruhu, co nejbliže ose snímání. Bylo použito 16 diod paralelně spojených napájených jednou klasickou AA baterií. Použité diody vyzařují na vlnové délce 940 nm v úhlu 20° . Jejich počet je sice dostatečný, ovšem v tomto případě platí pravidlo „čím více, tím lépe“.



obr. 7.5: Konečná hardwarové realizace

7.2 Softwarová realizace

Tato podkapitola se stává ze dvou celků. Prvním z nich je návrh, jak by měl algoritmus fungovat a co se od něj očekává. Také bude dále popsána jedna z rozšiřujících funkcí, kterou lze použít ke zpřesnění celé detekce očí a detekci obličeje. Druhým celkem je samotná realizace v prostředí MATLAB Simulink. Tento celek popisuje více do hloubky jednotlivé stavební bloky a funkce algoritmu.

7.2.1 Návrh softwarové realizace

Softwarová realizace spočívá v rozboru již IR LED diodami nasvíceného cíle. Je započata spuštěním snímání z kamery. Na obr. 7.7 je algoritmus rozboru, který se opakuje pro každý snímek přijatý z kamery.

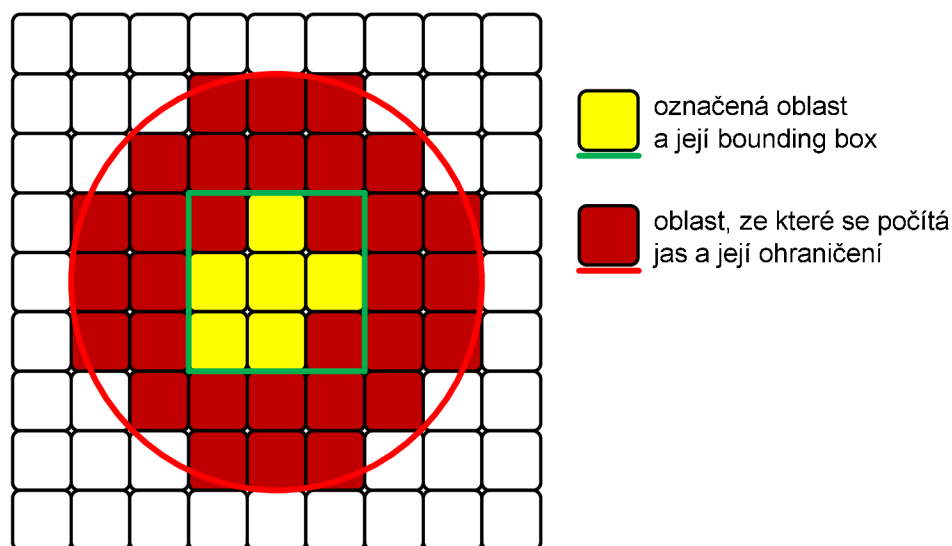
Vstupní snímek je ovšem nutné upravit do šedotónového obrazu (jasový obraz). Je tedy nutné snímek desaturovat. Toho je docíleno přepočítáním každého pixelu dle rovnice (4.1) a získáme tak ze třísloužkového (*RGB*) snímku jednosložkový (*I*).

V dalším kroku dojde k odprahování nízkých intenzit, které se vynulují (*False*) a nebudou už tak brány v úvahu při dalších výpočtech. Zbylé pixely se v dalším kroku nastaví na jednotnou nenulovou hodnotu (*True*) a tak získána binární mapa.

Na této mapě jsou označeny jednotlivé oblasti svým vlastním identifikátorem, a na nich dále provedena měření. Počet těchto oblastí je omezen na 128. Úkolem je změřit a přiřadit každé oblasti s identifikátorem položku s celkovou plochou (*area*) a tzv. bounding box, tedy 4 hodnoty určující levý, pravý, horní a spodní okraj oblasti v celém snímku. Tyto hodnoty jsou pro jednoduchost v diagramu v dalších krocích sjednoceny do proudu dat LAB.

Díky tomu, že je známa plocha jednotlivých oblastí, lze odfiltrovat další nevhodné oblasti. Protože je také znám tvar ohraničujícího boxu, lze další oblasti odfiltrovat na základě nevyhovujícího tvaru. U hledaného oka se očekává, že bude mít tvar velice blízký čtverci.

V další fázi stejně jako v té minulé se bude pracovat pouze s proudem hodnot LAB. Následuje výpočet středů oblastí společně s jejich poloměry. Každá oblast má dva poloměry, jeden na horizontální souřadnici x , druhý na vertikální souřadnici y . Vybrán je ten nižší, čímž je zajištěna vyšší pravděpodobnost, že se kruh nachází celou plochou v objektu. O předem danou hodnotu tento se tento poloměr zvýší a jsou vybrány všechny pixely okolí oblasti (v původním obraze), jejichž střed se nachází v kružnici. Na těchto pixelech je spočten jejich průměrný jas a ten je odečten od hodnoty jasu středového pixelu oblasti (v původním obraze), u kterého se počítá s tím, že v případě oka bude mít opravdu nejvyšší jas. Díky nedostatečnému rozdílu jsou odfiltrovány další oblasti.



obr. 7.6: Názorná ukázka filtrování pomocí rozdílu intenzit objektu a jeho okolí

Posledním krokem je vyznačení zbylé oblasti do původního obrazu a ten vykreslen na obrazovku.

Návrh rozšíření

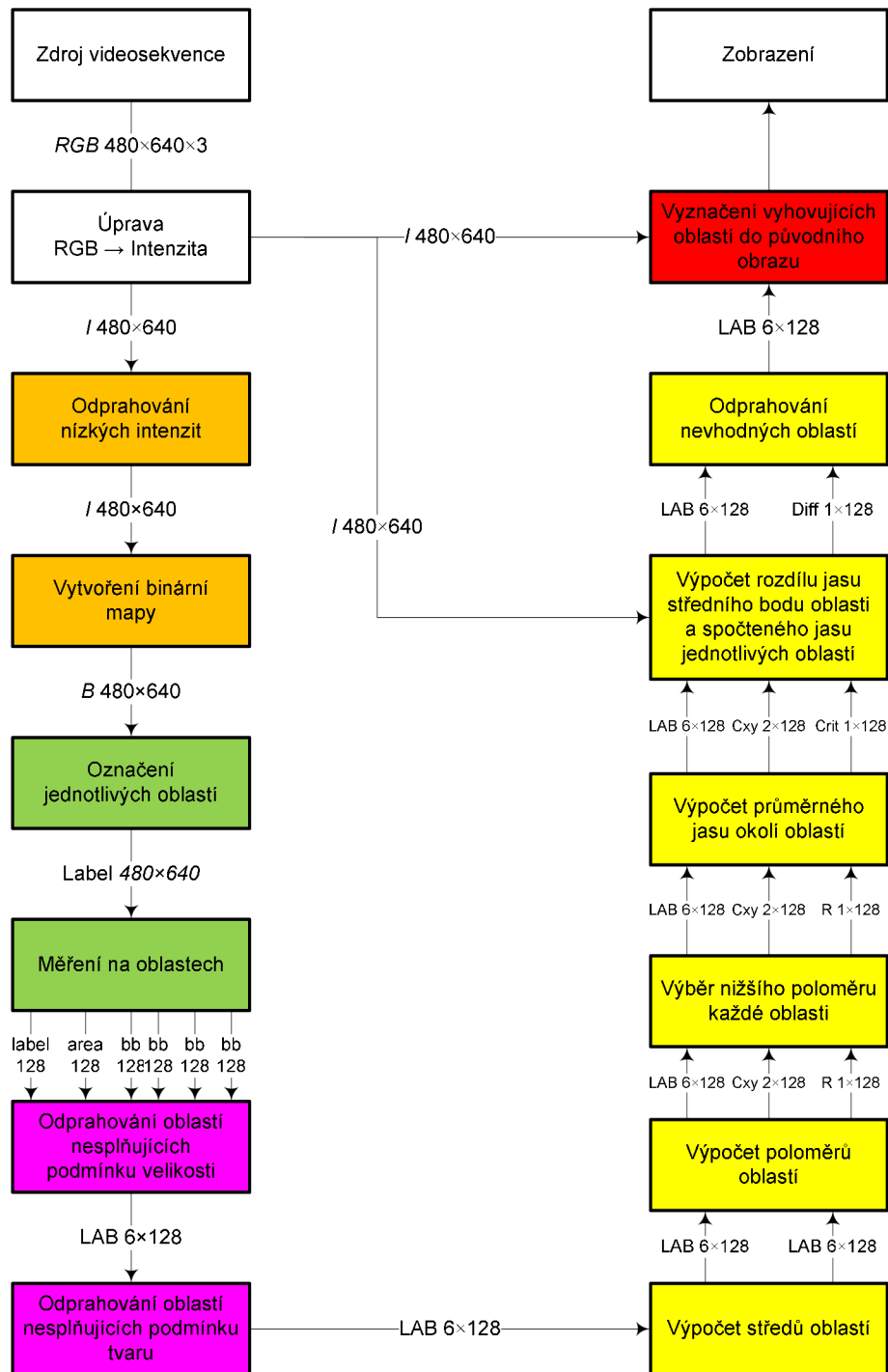
Prvním rozšířením, které se přímo nabízí, je rozšířit detekci o pravděpodobnosti matematický model ustavení očí a celého obličeje. Obecně totiž při detekci obličeje platí pravidlo, že je nutné nalézt oči (ústa, nos a jiné výrazné objekty), než může být objekt vyhodnocen jako obličej.

Ve vyfiltrovaném obraze zůstalo několik oblastí, o kterých se předpokládá, že se jedná o oči. Nicméně prakticky je dokázáno, že ne vždy je detekce ideální.

Cílem toho rozšíření bude zjistit, která dvojice očí patří k sobě. Na začátku se musí nejprve určit pevné hodnoty těchto vlastností a nastavit pravděpodobnost jejich odchylky:

- nejvyšší možný úhel mezi dvěma očima (omezí možnosti správně detekovaného naklání hlavy vlevo/vpravo),
- nejnižší vzdálenost mezi dvěma očima (sníží nejvyšší možnou vzdálenost snímaného cíle),
- nejvyšší vzdálenost mezi dvěma očima (zvýší nejnižší možnou vzdálenost snímaného cíle),
- podobnost velikostí oblastí vyhodnocených jako oko (obě zornice jednoho člověka jsou si geometricky velice podobné).

Ve vstupním vyfiltrovaném obraze jsou vypočteny tyto čtyři hodnoty pro každou možnou dvojici, dále porovnány s nastavenými hodnotami a vyhodnoceny, s jakou pravděpodobností se jedná o dvojici očí. Všechny dvojice, které nebudou kolidovat a jejich pravděpodobnosti budou vyšší než daný práh, budou vyhodnoceny jako dvojice očí. V případě kolize se samozřejmě vybere dvojice s větší pravděpodobností. Tato rozšiřující metoda není dokonalá a její funkčnost bude v porovnání s korelačními metodami (porovnání s databází snímků očí) značně nepřesná. Jejím velkým kladem je ovšem velice nízká výpočetní náročnost a možnost použití dalších filtrů.



obr. 7.7: Diagram algoritmu hledajícího oči ve snímku

7.2.2 Konečná softwarová realizace

Konečná softwarová aplikace téměř odpovídá návrhu výše. Pro její realizaci bylo zvoleno laboratorní vývojové prostředí Matlab, přesněji jeho součást Simulink. Barvy jednotlivých bloků obr. 7.7 jsou pro jednoduchost sladěny s realizací v Simulinku.

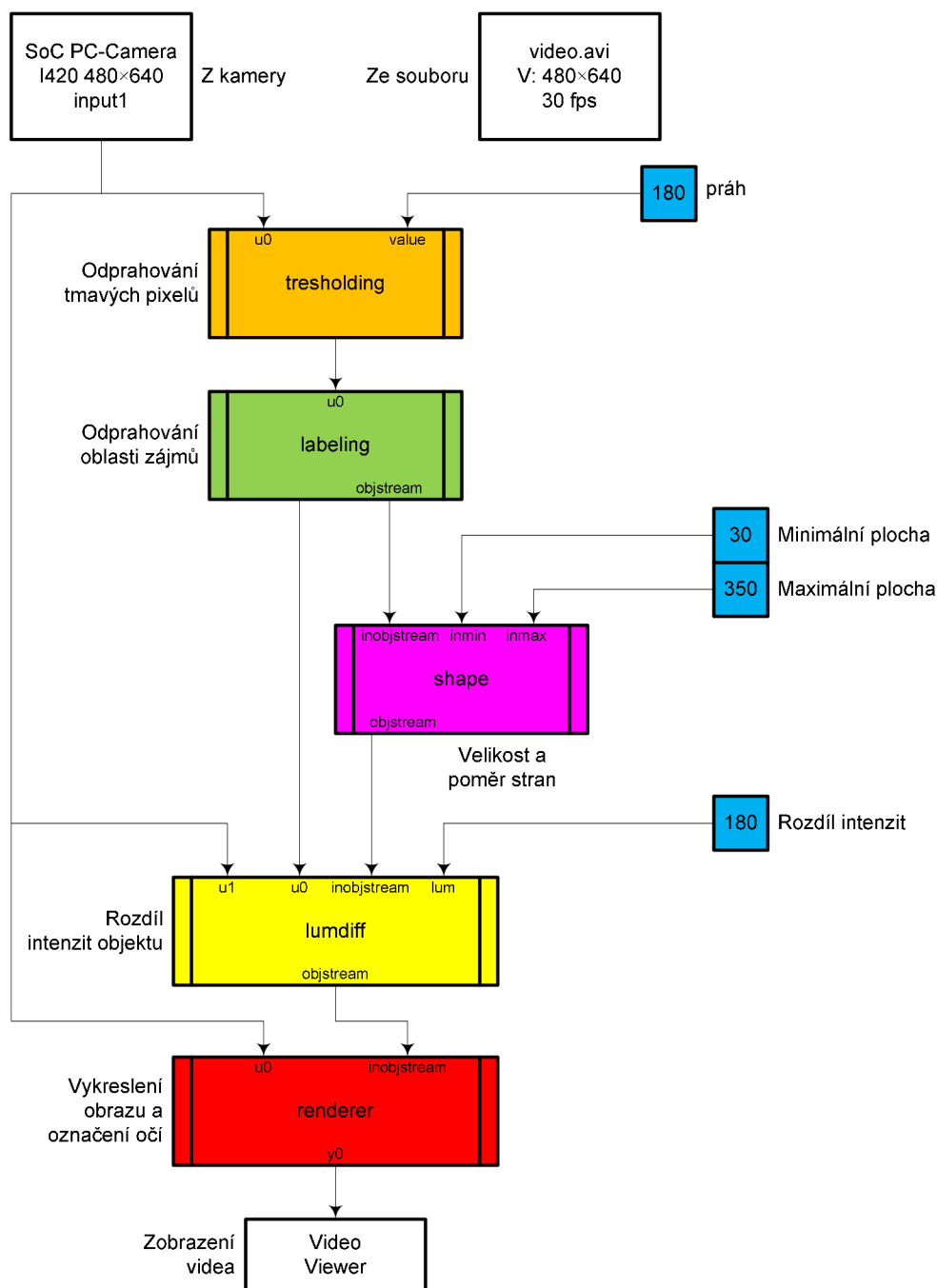
Vstupní bloky

Zdrojem videosekvencí, které dále zpracováváme, je právě tento blok, zpracovaný buď z kamery, nebo pro účely testování ze souboru. Jeho výstupem je 2D signál I , který je obrazem intenzity. Pro snímání bylo vybráno rozlišení 640×480 pixelů. S takto vysokým rozlišením je detekce přesnější, oči jsou lépe viditelnější, na druhou stranu je ovšem práce výpočetně mnohem náročnější. Proto byl zvolen pro další důležité bloky jazyk C, který v porovnání s funkcemi MATLABu i bloky Simulinku zpracovával celý algoritmus o poznání rychleji.

Konstanty

Dalšími bloky ze Simulinku jsou bloky konstant.

- Práh – tento blok nastavuje práh, který určí, že hodnoty intenzit nižších než práh budou vynulovány.
- Minimální a maximální plocha - nastavují, jak velké mohou být oblasti uvažovaného oka. V tomto případě odfiltruje oblasti, které nebudou mít poloměr v rozsahu cca od 3 do 11 pixelů.
- Rozdíl intenzit – tato konstanta bude popsána dále v textu u bloku Rozdíl intenzit objektů (*lumdiff*).



obr. 7.8: Bloky algoritmu v Simulinku

Bloky a funkce hlavního algoritmu

Odprahování tmavých pixelů

Tento blok je prvním uživatelsky definovaným blokem v jazyce C tohoto algoritmu. Stará se o odprahování pixelů s intenzitou nižší než je nastavený práh (vstupní hodnota value). V této fázi je obraz zbaven největšího počtu oblastí, a to tím, že jsou hodnoty pixelů vynulovány. Pokud ovšem hodnota pixelu splňuje zmíněné podmínky, nastaví se na hodnotu 255. Výstupem je odprahovaný signál y_0 , nabývající v každém pixelu hodnoty 0 nebo 255.

Označení oblastí zájmu (*ROI*)

V tomto bloku probíhá jedna z nejdůležitějších částí celého algoritmu. Touto částí je proces vyhledání a odlišení jednotlivých oblastí, který je založen na algoritmu záplavového vyplňování (funkce *FloodFill*) a použití FIFO fronty. Každá nalezená oblast je následně označena unikátní hodnotou. Výstupem je původní obraz y_0 a struktura *objstream*, která v sobě nese hodnotu oblasti, její plochu v pixelech a ohraničující box (bounding box).

Pro každý pixel se v tomto bloku mimo další operace vykoná následující proces. Tento úsek má za úkol zasít první výchozí semínko každé jednodílné oblasti s hodnotou 255. Těchto výchozích semínek může zasít maximálně 128, tedy maximální počet oblastí na výstupu je 128. Poté co je nalezena výchozí pozice pro semínko, je uložena do proměnných struktury *Seed*. Po vykonání záplavového vyplňování se počet objektů v poli *objects* zvýší o 1 a identifikační hodnota *label* se přičtením 1 připraví pro další oblast.

Fronta FIFO

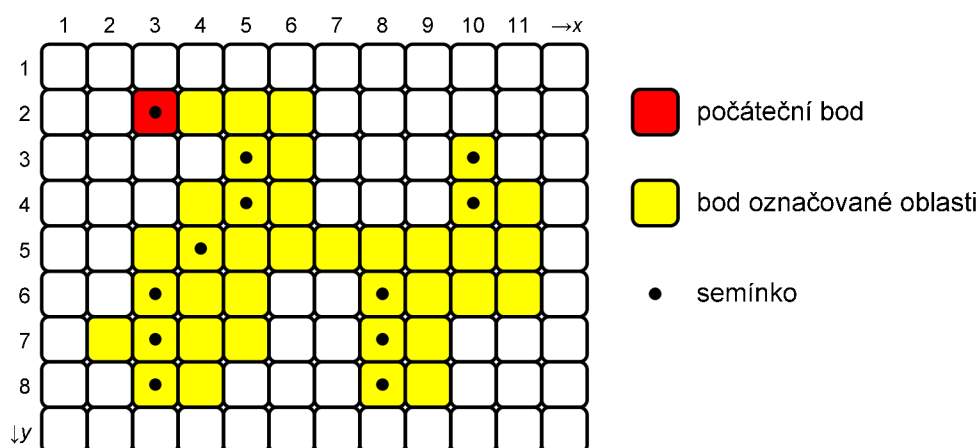
Jedná se o soubor funkcí, které zajišťují, že se budou data (semínka) zařazovat do fronty, to znamená, budou se zpracovávat formou FIFO, tedy první dovnitř, první ven. Opakem fronty je zásobník (LIFO) [13].

První funkce se stará o vymazání celé fronty. Druhá funkce vrací počet zpracovávaných semínek jednoduchým odečtením pozice následujícího semínka od semínka startovního. Třetí funkce se stará o přidání semínka do fronty a čtvrtá jej z fronty vrací k dalšímu zpracování.

Funkce FloodFill

Flood fill neboli česky záplavové vyplňování je algoritmus, který rozlišuje spojitou plochu jednotné barvy v obraze pro většinu uživatelů lépe představitelný jako „plechovka“ nebo „kouzelná hůlka“ ve většině bitmapových grafických nástrojů.

Pro celý algoritmus jsou důležité tři parametry: počáteční semínko (uzel), cílová barva a barva, kterou se bude vyplňovat. Je mnoho způsobů, kterými může být celý proces vyplňování proveden, ale všechny využívají řazení semínek formou fronty nebo zásobníku. Existují ovšem dva způsoby zaplavování, které dávají různý výsledek. Prvním je 4-směrné vyplňování, které je použito v této metodě a jehož průběh je popsán níže. Druhým způsobem je 8-směrné vyplňování, které na rozdíl od prvního kontroluje všech 8 směrů okolo semínka.



obr. 7.9: Ukázka 4-směrného záplavového vyplňování.

Podrobný popis názorné ukázky fungování (obr. 7.9) záplavového vyplňování

Obsahuje popis prvního a čtvrtého prohledávaného úseku. Pro upřesnění se pro podobné zpracování algoritmu, které je používáno v této práci vžil název „Scanline fill“.

1. Algoritmus dorazí po řádcích k pixelu s hodnotou 255 a zaseje semínko. Proměnné *seed_up* resp. *Seed_down* určují, zda může algoritmus zasít semínko v případě, že nalezne o pixel výše resp. níže sečí hodnotu 255. Tyto proměnné jsou automaticky po spuštění nastaveny na 1 (True).
2. Zkontroluje, zda není počáteční bod na začátku řádku. Pokud není, zkontroluje pixely vlevo na hodnotu 255 (vykonává se i pro počáteční pixel). Pokud zde žádné nejsou, provede kontrolu vpravo.
3. Při pohybu vpravo narazí na pixel (2,5). Algoritmus stejně jako u každého pixelu zkontroluje horní a spodní pixel na naší sečí hodnotu.
4. Je vidět, že pixel (3,5) má hodnotu 255 a jelikož je hodnota proměnné *seed_down* 1 (True), zaseje se v tomto místě semínko, které se zařadí do fronty (v tomto případě na první místo).
5. Dokončí se pohyb vpravo. Přestože pod pixelem (2,6) je při kontrole hodnoty hodnota 255, semínko už se nezaseje, protože v proměnné *seed_down* je již hodnota 0 (False).
6. Pokud se změní řádek, nastaví se proměnné *seed_up* i *seed_down* na 1 (True). Pokud algoritmus narazí o pixel výše resp. níže na překážku (tzn. pixel s hodnotou 0), uloží hodnotu 1 do *seed_up*, resp. *seed_down*.
7. Dalším pixelem ve frontě je tedy (3,5). Průběh v tomto i následujícím úseku je podobný jako výše.
8. Abychom si objasnili výhodu použití fronty, pokročíme k semínku (5, 4). Stejně jako výše zkontrolujeme, zda nejsme na začátku řádku, a zkontrolu-

jeme pixely vlevo na hodnotu 255. Tím se dostaneme na pixel na levém okraji úseku a provádíme test s proměnnou *seed_down*.

9. Podmínka, že spodní pixel (6,3) má hodnotu 255 a *seed_down* je roven 1 (True) je splněna, v bodě (6,3) se tedy zaseje semínko a bod se zařadí do fronty.
10. Proměnné *seed_down* se nastaví 0 (False) a algoritmus se pohybuje po úseku vpravo až do bodu (5,6), kde narazí na spodní pixel (6,6) s hodnotou 0, který tvoří překážku. Proměnná *seed_down* se tak nastaví zpět na hodnotu 1 a pokračuje dále vpravo.
11. V bodě (5, 8) algoritmus zjistí splnění podmínky pro zasetí semínka na pixelu (6,8). Toto semínko se zařadí do fronty na druhé místo. Proměnná *seed_down* se nastaví na 0 (False).
12. V bodě (5, 10) je splněna podmínka s proměnnou *seed_up* a je zaseto semínko na pixel (4,10) a zařazeno do fronty na třetí místo.
13. Algoritmus poté dokončí prohledávání úseku (5, 3 – 11) a přejde k semínku, které je ve frontě na řadě.

Funkce EncodeStream

Funkce *EncodeStream* má za úkol zakódovat používaná data do proudu, jednak z důvodu lepší přehlednosti, než byl původní, ale také z důvodu jednodušší práce s tímto proudem. Parametry této funkce jsou odkaz na výstupní proud *pStream*, pole nalezených oblastí ve snímku a jejich počet. Celá funkce vrací počet zakódovaných objektů.

Funkce DecodeStream

Funkce *DecodeStream* má za úkol zakódovaná data z proudu zase rozložit pro následné zpracování. Jedná se o opačný proces než u funkce *EncodeStream*.

Test velikosti a poměru stran

V tomto bloku dojde na základě vstupních konstantních parametrů *inmin* a *inmax* k odfiltrování objektů s plochou vyšší resp. nižší než je stanovené maximum resp. minimum a dále k odfiltrování objektů, jejichž poměr stran je nižší než 3:4 resp. vyšší než 4:3. Stejně jako se na vstupu bloku musela vstupní struktura *inobjstream* rozkódovat, tak se nyní na konci bloku znovu zakóduje (narovná) do struktury *objstream*.

Tento úsek zdrojového kódu se stará právě o vyfiltrování oblastí s nevyhovující plochou a tvarem. Provede se pro každý objekt. Pokud objekt splňuje podmínku plochy, zjistí si z pole *inobjects* hodnoty jeho krajních bodů na obou souřadnicích a vypočte rozměry bounding boxu. Z těchto rozměrů spočítá poměr mezi stranami a

přejde k podmínce. Pokud je podmínka splněna uloží se zpracovávaný objekt na místo v připraveném poli struktury *Object_objects*.

Rozdíl intenzit objektů

Funkce tohoto bloku je jednoduše pochopitelná z obr. 7.6. Jas uvažovaného oka odvodíme z jeho středového bodu a porovnáme s průměrnou hodnotou jasu v okolí oblasti, obě tyto operace provádíme na originálním obraze. Pokud je rozdíl mezi těmito intenzitami vyšší než nastavená konstanta *lum* - rozdíl intenzit, vyhodnotí se tato oblast jako oko. Všechna takto nalezená místa se opět zakódují do struktury *objstream*.

V tomto úseku kódu se provádí výpočet parametrické rovnice kružnice a do proměnné *crit* se ukládá hodnota 1 (True), pokud se pixel nachází svým středem kružnici. Pokud tomu tak je a pixel zároveň není označen identifikátorem zpracovávané oblasti, zapíše se do proměnné *luminance* hodnota jasu pixelu a zvýší se čítač *count* jejich počtu.

Pokud alespoň jeden pixel vyhovoval podmínkám výše, provede se výpočet průměrného jasu.

Tato hodnota se odečte od jasu středového pixelu objektu a porovná se s hodnotou konstanty, na kterou odkazuje *lum*. Pokud je podmínka splněna dojde k zapsání zpracovávaných hodnot na určené místo pole *eyeObjects* a čítač *numEyeObjects* se zvýšením o 1 připraví pro další objekt.

Vykreslování a označení očí

Tento blok se již stará pouze o označení očí v originálním obraze a jeho posílání na výstup. Jednoduché vykreslení hran bounding boxu invertovanou barvou.

Výstupní obraz

Díky tomuto bloku vidíme finální výsledek celého algoritmu.

7.2.3 Větev analýzy

Větev analýzy se skládá ze tří bloků:

- Blok rozboru rozkódovaných streamů v jednotlivých fázích a vyplnění bounding boxů jednotlivých fází jednotnou barvou pomocí funkce *DrawRect*. Na výstupu je tedy signál s tímto analytickým pohledem a počet detekovaných objektů v jednotlivých fázích.
- Přidání textu s informacemi o počtu nalezených objektů v jednotlivých fázích.
- Zobrazení analytického výstupu.

Tato část algoritmu byla k realizaci přidána hlavně z důvodu kontroly fungování jednotlivých bloků. Na obr. 7.10 je příklad analytického pohledu, kde jsou v horním

rohu zobrazeny informace o počtu detekovaných objektů v jednotlivých fázích. Bílá barva v analytickém pohledu označuje nalezené oči. Je patrné, že právě v tomto snímku je detekce nepřesná, protože je detekován objekt, který okem očividně není. Kvůli těmto chybám v detekci je nutné vyladit prahovací konstanty algoritmu a k tomuto účelu je zde větev analýzy.



obr. 7.10: Názorný příklad, vlevo detekované oči, vpravo analytický pohled.

7.3 Závěr

Metoda popisovaná v této části je především zajímavá tím, že i jednoduchý a výpočetně minimálně náročný algoritmus může být dostatečně spolehlivý. Velkým nedostatkem metody detekce barvy kůže byla mimo vysokou výpočetní náročnost její silná závislost na ideálním prostředí a osvětlení, a i minimální odchylky, ať již na straně špatného osvětlení nebo nekvalitní kamery, trasování znemožňovaly. Cílem proto bylo tento problém alespoň v malé míře eliminovat. Použitý algoritmus sice nefunguje ideálně ve všech světelných podmínkách, jeho rozsah je však dostatečný. Největší předností je, že při dostatečné vzdálenosti (asi 1 m) je pravděpodobnost správné detekce 98%.

Díky těmto výsledkům, se může popsaný postup využít v biometrických v systémech pro detekci živosti. Momentálně se používá detekce s využitím číslicového zpracování signálů bez IR technologie. Současně s např. rozpoznáváním oční duhovky je výhodné použít tento způsob, který je levný, rychlý a přesný.

8 Závěr

V disertační práci byla analyzována problematika detekce mrkání očí s využitím číslicového zpracování obrazu a to v přístupech bez i s IR technologií. Pro oba přístupy byla navržena metoda detekce mrkání.

Po uvedení do problematiky a stanovení cílů disertační práce byly v kapitolách 3 a 4 uvedeny základy číslicového zpracování obrazu společně s metodami a postupy, užití pro komplexní oblast detekce a rozpoznávání obličejů. V kapitole 5 jsou rozebrány dostupné obrazové databáze statických obrazů i videosekvencí. Rovněž je zde popis vlastní vytvořené videodatabáze.

V kapitole 6 je podrobně vysvětlen postup detekce mrkání s využitím číslicového zpracování bez IR technologie. Podle předpokladů bylo zjištěno, že hledání kompromisu mezi výpočetní rychlostí a přesností záleží na dané aplikaci. V tomto případě byl postup navíc poměrně pomalý, což je daň za skutečnost, že tento přístup je po hardwarové stránce nenáročný. Výzkum v tomto přístupu je dále vyvíjen, stejně jako v celé problematice detekce a rozpoznávání obličejů.

V kapitole 7 je podrobně vysvětlen postup detekce opět s využitím číslicového zpracování signálů tentokrát s IR technologií. Tento přístup využívá přídavná zařízení, která však nejsou nijak drahá nebo nedostupná, jedná se o přípravek s IR diodami a obyčejnou kameru. Hardwarová realizace je sice náročnější než v předešlém přístupu, výsledky jsou však mnohem příznivější.

Výsledky z posledního přístupu jsou tak příznivé, že lze metodu využít v biometrických systémech jako podsystém detekce živosti. Například při identifikaci pomocí oční duhovky lze sledovat oči osoby a v případě mrknutí tak potvrdit jeho skutečnou přítomnost. Ve vzdálenosti 1 m od snímacího zařízení je dosaženo téměř 100% úspěšnosti detekce mrknutí, což je nad rámec požadavků, neboť osoba během snímání duhovky mrkne více než jednou.

Další práce spočívá ve vývoji algoritmů pro detekci mrkání vždy přizpůsobené pro určitou aplikaci, tj. dnes hlavně ve třech oborech: bezpečnostní havarijní systémy, biometrické systémy a v lékařské technice. V celém komplexním oboru detekce a rozpoznávání obličejů se neobjevují zcela nové techniky, pouze se vylepšují ty obecně známé.

Literatura

- [1] JAN, J. *Číslicová filtrace, analýza a restaurace signálů*. Brno: VUTIUM Publishing, 2002. 427 s. ISBN: 80-214-1558-4.
- [2] HLAVÁČ, V., ŠONKA, M. *Počítačové vidění*, Praha: Grada, 1993. 272 s. ISBN: 80-85424-67-3.
- [3] HLAVÁČ, V., SEDLÁČEK, M. *Zpracování signálů a obrazů*. Praha: České Vysoké Učení Technické, 2000. 221 s. ISBN: 80-01-02114-9.
- [4] ROSS, A., NANDAKUMAR, K., JAIN, A. K. *Handbook of Multibiometrics*. New York: Springer Verlag, 2006. ISBN: 978-0-387-22296-7.
- [5] GONZALEZ, R. C., WOODS, R. E. *Digital Image Processing*. Prentice Hall, 2007. 779 s. ISBN: 978-0131687288.
- [6] GONZALEZ, R. C., WOODS, R. E., EDDINS, S. L. *Digital Image Processing Using MATLAB*. Prentice Hall, 2003. 609 s. ISBN: 978-0130085191.
- [7] HALLIDAY D., RESNICK R., WALKER J. *Fyzika: Elektromagnetické vlny, Optika, Relativita*. Brno: VUTIUM Publishing, 2003. 141 s. ISBN: 80-274-1868-0.
- [8] Japan Society of Vision Science ed. *Shikaku Jyouthou Syori Handobukku (Handbook of Visual Information Processing)*. Tokyo: Asakura Publishing, 2001.
- [9] VÍT, V., KUBA, P. *Televizní technika – studiové zpracování televizního signálu*, 1. vyd. Praha: BEN – technická literatura, 2002. 208 s. ISBN 80-86056-88-0.
- [10] VÍT, V. *Televizní technika – přenosové barevné soustavy*, 1. vyd. Praha: BEN – technická literatura, 2002. 720 s. ISBN: 80-86056-04-X.
- [11] KAUFMAN, P. L., ALM, A. *Adler's Physiology of the Eye: Clinical Application*, 10. vyd. Mosby, 2003. 888 s. ISBN: 978-0-323-01136-5.
- [12] GALLOWAY, N. R. et al. *Common Eye Diseases and Their Management*, 3. vyd. Springer-Verlag, 2006. 224 s. ISBN: 1-85233-985-3.
- [13] GIBSON, W. *Pattern Recognition*. G. P. Putnam's Sons, 2003. 368 s. ISBN: 0-399-14986-4.
- [14] KRUSE, R. L. *Data Structures & Program Design*. Englewood Cliffs: Prentice-Hall, 1987. 150 s. ISBN: 0-13-195884-4.
- [15] SIVIA D. S., *Data Analysis: A Bayesian Tutorial*. Oxford: Clarendon (Oxford University Press), 1996. ISBN: 0-19-851762-9.
- [16] YANO, K et al. Detection of Eye Blinking from Video Camera with Dynamic ROI Fixation. In *Proceedings of the 1999 IEEE International Confere-*

- rence on Systems, Man and Cybernetic (SMC 1999)*. Tokyo, vol. 6, s. 335-339. ISBN: 0-7803-5731-0.
- [17] BHASKAR, T. N., KEAT, F. T., RANGANATH, S. Blink Detection and Eye Tracking for Eye Localization. In *Proceedings of the Conference on Convergent Technologies for Asia-Pacific Region (TENCON 2003)*. Taj Residenc, 2003, vol. 2, s. 821-824. ISBN: 0-7803-8162-9.
- [18] CHAI D., BOUZERDOUM, A. A Bayesian Approach to Skin Color Classification in YCbCr Color Space. In *Proceedings of the TENCON 2000*. Kuala Lumpur, 2000, vol. 2, s. 421-424. ISBN: 0-7803-6355-8.
- [19] BLACK, M. J., FLEET, D. J., YACOOB, Y. A Framework for Modeling Appearance Change in Image Sequences. In *Proceedings of the 1998 IEEE International Conference on Computer Vision (ICCV'98)*. Bombai, 1998, s. 660-667.
- [20] VIOLA, P., JONES, M. J., SNOW, D. Detecting Pedestrians using Patterns of Motion and Appearance. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV'03)*. Nice, 2003, vol. 2, s. 734-741. ISBN: 0-7695-1950-4.
- [21] GONI, S., et al. Robust Algorithm for Pupil-Glint Vector Detection in a Video-oculography Eyetracking System. In *Proceedings of the 17th international Conference on Pattern Recognition, (ICPR'04)*. Washington DC, 2004, vol. 04, s. 941-944. ISBN: 0-7695-2128-2.
- [22] YANG, M.-H. Detecting Faces in Images: A Survey. *IEEE Transactions on pattern analysis and machine intelligence (PAMI)*. January 2002, vol. 24, no. 1, s 34-58. ISSN 0162-8828.
- [23] GRAUMAN, K. et al. Communication via Eye Blinks - Detection and Duration Analysis in Real Time. In *Proceedings of the 2001 IEEE Computer Conference on Computer Vision and Pattern Recognition (CVPR'01)*. Kauai Marriott, 2001, vol. 1, s. 1010-1017. ISBN: 0-7695-1272-0.
- [24] HARO, A., FLICKER, M., ESSA, I, Detecting and Tracking Eyes by Using Their Physiological Properties, Dynamics, and Appearance. In *Proceedings of the 2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)*. Hilton Head Island, 2000, vol. 1, s. 163-168. ISBN: 0-7695-0662-3.
- [25] CHETTY, G., WAGNER, M. Multi-level Liveness Verification for Face-Voice Biometric Authentication. In *Proceedings of the Biometrics Symposium 2006*. Baltimore MD, September 2006, s. 1-6. ISBN: 978-1-4244-0487-2.
- [26] PAN, G. et al. Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcam. In *Proceedings of the IEEE 11th International Confe-*

- rence on Computer Vision (ICCV'07)*. Rio de Janeiro, October 2007, s. 1-8. ISBN: 978-1-4244-1631-8.
- [27] HEISHMAN, R., DURIC, Z., WELCHSLER H. Using Eye Region Biometrics to Reveal Affective and Cognitive States. In *Proceedings of the Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04)*. Washington DC, 2004, vol. 5, s. 69-69. ISBN 0-7695-2158-4.
- [28] KONTROPOULOS, C., PITAS, I. Rule-based Face Detection in Frontal Views. In *Proceeding of the 1997 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '97)*. Washington DC, 1997, vol. 4, s. 2537-2540, ISBN: 0-8186-7919-0.
- [29] LUCAS, B. D., KANADE, T. An Iterative Image Registration Technique with an Application to Stereo Vision. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'81)*. Vancouver, 1981, vol. 2, s. 674-679.
- [30] PARTHASARADHI, S. et al. Time-Series Detection of Perspiration as a Liveness Test in Fingerprint Devices. *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, 2005, vol. 35, no. 3, s. 335-343. ISSN 1094-6977.
- [31] QUINTILLIANO, P., SANTA-ROSA, A. Face Recognition Based on Eigeneyes. *Pattern Recognition and Image Analysis*. 2003, vol. 13, no. 2, s. 335-338. ISSN 1054-6618.
- [32] YANG, M.-H., AHUJA, N. Detecting Human Faces in Color Images. In *Proceedings of the International Conference on Image Processing (ICIP '98)*. Chicago, 1998, vol. 1, s. 127-130.
- [33] KARSON, C. Spontaneous Eye-Blink Rates and Dopaminergic Systems. *Brain*. 1983, vol.106, s.643-653. ISSN 0006-8950.
- [34] MORIMOTO, C. et al. Real-time detection of eyes and faces. In *Proceedings of the Workshop on Perceptual User Interfaces (PUI'98)*. 1998, s. 117-120.
- [35] GARCIA, C., ZIKOS, G., TZIRITAS, G. Face Detection in Color Images using Wavelet Packet Analysis. In *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*. Florence, 1999, vol. 1, s. 703-708. ISBN: 0-7695-0253-9.
- [36] PICCARDI, M. Background subtraction techniques: a review. In *Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetic*. The Hague, 2004, vol. 4, s. 3099-3104. ISBN: 0-7803-8566-7.
- [37] ŽÁRA, J. et al. *Moderní počítačová grafika*, 2. vyd. Praha: Computer Press, 1998. 628 s. ISBN: 80-251-0454-0.
- [38] MORIYAMA, T. et al. Automatic Recognition of Eye Blinking in Spontaneously Occurring Behavior. In *Proceedings of the International Conferen-*

ce on Pattern Recognition (ICPR'2002). Quebec, 1992, vol. IV, s. 78-81. ISBN: 0-7695-1695-X.

- [39] SINHA, P. Object Recognition via Image Invariants: A Case Study. *Investigative Ophthalmology & Visual Science*. 1994, vol. 35, no. 4, s. 1735-1740. ISSN 0146-0404.
- [40] SCHUCKERS, S. Spoofing and Anti-Spoofing Measures. *Information Security Technical Report*. 1992, vol. 7, no. 4, s. 56-62. ISSN 1363-4127.
- [41] BIGUN, J., FRONTHALER, H., KOLLREIDER, K. Assuring Liveness in Biometric Identity Authentication by Real-Time Face Tracking, In *Proceedings of the IEEE Conference on Computational Intelligence for Homeland Security and Personal Safety (CIHSPS'04)*. Venice, July 2004, s. 104-111. ISBN: 0-7803-8381-8.
- [42] LEE, Y. J., YOO, S. I. An Elliptical Boundary Model for Skin Color Detection. In *Proceedings of the 2002 International Conference on Imaging Science, Systems, and Technology (CISST'2002)*. Las Vegas, 2002.
- [43] KOLLREIDER, K., FRONTHALER, H., BIGUN, J. Evaluating Liveness by Face Images and the Structure Tensor, In *Proceedings of the 4th IEEE Workshop on Automatic Identification Advanced Technologies*. Washington DC, October 2005, s.75-80. ISBN: 0-7695-2475-3.
- [44] ANTONELLI, A. et al. Fake finger detection by skin distortion analysis. *IEEE Transactions on Information Forensics and Security*, 2006, vol. 1, no. 3, s. 360-373. ISSN 1556-6013.
- [45] NAKAMORI, K., et al. Blinking is controlled primarily by ocular surface conditions. *American Journal of Ophthalmology*. 1997, vol. 124, iss. 1., s. 24-30. ISSN 0002-9394.
- [46] ZHAO, W. et al. Face Recognition: A Literature Survey. *ACM Computing Surveys*. 2003, vol. 35, no. 4, s.399-458. ISSN 0360-0300.
- [47] FERIS, R.-S., et al. Detection and tracking of Facial Features in Video Sequences. *Lecture Notes in Artificial Intelligence*. 2000, vol. 1793, s. 197-206. ISSN 0302-9743.
- [48] TIAN, Y., KANADE, K., COHN, J. F. Recognizing Action Units for Facial Expression Analysis. *IEEE Transactions on pattern analysis and machine intelligence (PAMI)*. 2001, vol. 23, no. 2, s. 97-115. ISSN 0162-8828.
- [49] SIGAL, L., SCLAROFF, S., ATHITSOS, V. Estimation and Prediction of Evolving Color Distributions for Skin Segmentation under Varying Illumination. In *Proceedings of the 2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)*. Hilton Head Island, 2000, vol. 2, s. 152-159. ISBN: 0-7695-0662-3.

- [50] VEZHNEVETS, V., SAZOMOV, V., ANDREEVA, A. A Survey on Pixel Based Skin Color Detection Techniques. In *Proceedings of the GraphiCon 2003*. Moscow, 2003, s. 85-92.
- [51] VEZHNEVETS, V., DEGTIAREVA, A. Robust and Accurate Eye Contour Extraction. In *Proceedings of the GraphiCon 2003*. Moscow, 2003, s. 81-84.
- [52] SEDLÁČEK, M. Evaluation of RGB and HSV models in Human Faces. In *Proceedings of the 8th Central European Seminar on Computer Graphics (CESGG 2004)*, Budmerice, 2004, s. 125-131.
- [53] JI, Q., ZHU, Z., LAN, P. Real Time Nonintrusive Monitoring and Prediction of Driver Fatigue, *IEEE Transactions on Vehicular Technology*. 2004, vol. 53, no. 4, s. 1052-1068. ISSN 0018-9545.
- [54] AL-QAYEDI, A. M., CLARK, A. F. Constant-Rate Eye Tracking and Animation for Model-Based-Coded Video. In *Proceedings of the 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing (CASSP'00)*. 2000, vol. 4, s. 2353-2356.
- [55] MURALI, S., SHANKAR, R. Performance Analysis of a Vehicle Crash Control System using Image. In *Proceedings of the 4th IEEE International Symposium on Electronic Design, Test & Applications (DELTA 2008)*. Hong Kong, 2008, s. 61-66. ISBN: 0-7695-3110-5.
- [56] CHOUDHURY, T. et al. Multimodal Person Recognition Using Unconstrained Audio and Video. In *Proceedings of the 4th International Conference on Audio- and Video-based Biometric Person Authentication (AVB-PA'99)*. Washington DC, 1999, s.176-181.
- [57] CROWLEY, J. L., BERARD, F. Multi-modal Tracking of Faces for Video Communications. In *Proceedings of the 1997 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'97)*. Puerto Rico, 1997, s. 640-645. ISBN: 978-08-1867-8226.
- [58] BALA, L.-P., TALMI, K., LIU, J. Automatic Detection and Tracking of Faces and Facial Features in Video Sequences. In *Proceedings of the Picture Coding Symposium (PCS'97)*. Berlin, 1997.
- [59] FRISCHHOLZ, R. W., DIECKMANN, U. BioID: A Multimodal Biometric Identification System. *IEEE Computer*. 2000, vol. 33, no. 2, s.64-68. ISSN 0018-9162.
- [60] WANG, H., CHANG, S. A Highly Efficient System for Automatic Face Region Detection in MPEG Video. *IEEE Transactions on Circuits and Systems for Video Technology*. 1997, vol. 7, no. 4, s. 615-628. ISSN 1558-2205.
- [61] FRISCHHOLZ, R. W., WERNER, A. Avoiding Replay-Attacks in a Face Recognition System using Head-Pose Estimation. In *Proceedings of the*

- IEEE International Workshop on Analysis and Modeling of Faces and Gestures (AMFG'03)*. 2003, s. 234- 235. ISBN: 0-7695-2010-3.
- [62] ROWLEY, H., BALUJA, S., KANADE, T. Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998, vol. 20, no. 1, s. 23- 38. ISSN 0162-8828.
- [63] CRAW, I., ELLIS, H., LISHMAN J. R. Automatic extraction of face features. *Pattern Recognition Letters*. 1987, vol. 5, iss. 2, s. 183-187. ISSN 0167-8655.
- [64] SHIRATORI, T. et al. Video Completion by Motion Field Transfer. In *Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition*. New York, 2006, vol. 1, s. 411-418. ISBN: 0-7695-2597-0.
- [65] LI, J. et al. Live Face Detection Based on the Analysis of Fourier Spectra. *Biometric Technology for Human Identification, Proceedings of the SPIE (The International Society for Optical Engineering)*. 2004, vol. 5404, s. 296-303. ISSN 0277-786X.
- [66] SOCOLINSKY, D. A., SELINGER, A., NEUHEISEL, J. D. Face Recognition with Visible and Thermal Infrared Imagery. *Computer Vision and Image Understanding (CVIU)*. 2003, vol. 91, no. 1-2, s. 72-114. ISSN 1077-3142.
- [67] TURK, M., PENTLAND, A. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*. 1991, vol. 3, no. 1, s. 71-86. ISSN 0148-9267.
- [68] SAAD, A.-S., et al. Eye Detection in a Face Image using Linear and Nonlinear Filters. *Pattern Recognition*. 2001. vol. 34, iss. 7, s. 1367-1391. ISSN 0031-3203.
- [69] GORODNICHY, D. O. Towards Automatic Retrieval of Blink-Based Lexicon for Persons Suffered from Brain-Stem Injury using Video Cameras. In *Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPR'04)*. Washington DC, 2004, vol. 5, s. 68-75. ISBN: 0-7695-2158-4.
- [70] JUAN, S. *The Odd Body: Mysteries of Our Weird and Wonderful Bodies Explained*. Andrews McMeel Publishing, 2004. 304 s. ISBN: 0-7407-6159-5.
- [71] DOUGHTY, M. J., NAASE, T. Further Analysis of the Human Spontaneous Eye Blink Rate by a Cluster Analysis-Based Approach to Categorize Individuals With 'Normal' Versus 'Frequent' Eye Blink Activity. *Eye & Contact Lens: Science & Clinical Practice*. 2006, vol. 32, iss. 6, s. 294-299. ISSN 1542-2321.
- [72] YEDNEKACHEW A., CHEN B., ADLER A. Impact of Pose and Glasses on Face Detection Using the Red Eye Effect. In *Proceedings of the Canadi-*

- an Conference on Electrical and Computer Engineering (IEEE CCECE 2003)*. May 2003, vol. 3, s. 1211-1214. ISBN: 0-7803-7781-8.
- [73] RYAN, S. B., et al. A long-range, wide field-of-view infrared eye blink detector. *Journal of Neuroscience Methods*. 2005, vol. 152, no. 1, s. 74-82. ISSN 0165-0270.
- [74] YUILLE, A., HALLINAN, P., COHEN, D. Feature Extraction from Faces using Deformable Templates. *International Journal of Computer Vision*. 1992, vol. 8, no. 2, 99-111. ISSN 0920-5691.
- [75] XIONG, L. et al. Eye Synthesis Using the Eye Curve Model. In *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007)*. Washington DC, 2007, vol. 2, s. 531-534. ISBN: 0-7695-3015-X.
- [76] AL-ABDULMUNEM, M.; BRIGGS, S. T. Spontaneous blink rate of a normal population sample. *International Contact Lens Clinic*. 1999, vol. 26, iss. 2, s. 29-32. ISSN 0892-8967.
- [77] KAWATO, S., TETSUTANI, N., Detection and Tracking of Eyes for Gaze-camera Control. *Image and Vision Computing*. 2004, vol. 22, iss. 12, s. 1031-1038. ISSN 0262-8856.
- [78] MORRIS, T., BLENKHORN, P., ZAIDI, F. Blink Detection for Real-time Eye Tracking. *Journal of Network and Computer Applications*. 2002, vol. 25, iss. 2, s. 129-143.
- [79] HSU, R. L., ABDEL-MOTTALEB, M. Face detection in color images. In *Proceedings of the 2001 International Conference on Image Processing (ICIP '01)*. Thessaloniki, 2001, vol. 1, s. 1046-1049.
- [80] OYEKOYA, O. K., STENTIFORD F. W. M. Eye Tracking as a New Interface for Image Retrieval. *British Telecommunications Technology Journal*. 2004, vol. 22, iss. 3, s. 161-169. ISSN 1358-3948.
- [81] ZHU, Z., JI, O. Robust Real-Time Eye Detection and Tracking under Variable Lighting Conditions and Various Face Orientations. *Computer Vision and Image Understanding*. 2005, vol. 98, iss. 1, s. 124-154. ISSN 1077-3142.
- [82] BRISTOW, D., FRITH, C., REES, G. Two distinct neural effects of blinking on human visual processing. *NeuroImage*. 2005, vol. 27, iss. 1, s. 136-145. ISSN 1053-8119.
- [83] SHUOCKER, D. *Handbook of the Eurolaser Academy*. Kluwer Academic Publishers, 1998. 448 s. ISBN: 978-0-412-82590-3.
- [84] LI, S. Z., JAIN, A. K. *Handbook of Face Recognition*. New York: Springer Science+Business Media Inc., 2005. 394 s. ISBN: 0-387-40595-X.

- [85] GRAUMAN, K., et al. Communication via eye blinks and eyebrow raises: video-based human-computer-interfaces. *Universal Access in the Information Society*. October 2003, vol. 2, no. 4, s. 359-375. ISSN 1615-5289.
- [86] UENO, H., KANEDA, M., TSUKINO, M. Development of Drowsiness Detection System. In *Proceeding of the Conference on Vehicle Navigation & Information Systems (VNIS'94)*. Yokohama, 1994, s. 15-20. ISBN: 0-7803-2105-7.
- [87] HE, X., LIU, Z. M., ZHOU, J. L. Real-time Human Face Detection in Color Image. In *Proceeding of the 2nd International Conference on Machine Learning and Cybernetics*. Xi'an, 2003, vol. 5, s. 2915-2920.
- [88] DENG, X., CHANG, C., BRANDLE, E. A New Method for Eye Extraction from Facial Image. In *Proceedings of the 2nd IEEE International Workshop on Electronic Design, Test & Applications (DELTA 2004)*. Perth, 2004, s. 29-34. ISBN: 0-7695-2081-2.
- [89] NOVOTNÝ, I., HRUŠKA M. *Biologie člověka pro gymnasia*, 3. vyd. Praha: Nakladatelství učebnic Fortuna, 2002. 240 s. ISBN: 80-7168-819-3.
- [90] GUO, C. F., PONG, C. Y. Multi-cues Eye Detection on Gray Intensity Image. *Pattern Recognition*. 2001, vol. 34, iss. 5, s. 1033-1046. ISSN 0031-3203.
- [91] ČIHÁK, R., *Anatomie 1*, 2. vyd. Praha: Grada, 2001. 516 s. ISBN: 80-7169-970-5.
- [92] ČIHÁK, R., DRUGA, R. *Anatomie 3*, 2. vyd. Praha: Grada, 2004. 692 s. ISBN: 80-247-1132-X.
- [93] KELLER, O. *Obecná elektromyografie*. Praha: Triton, 1999. 176 s. ISBN: 80-7254-047-5.
- [94] HALLIDAY D., RESNICK R., WALKER J. *Fyzika: Elektromagnetické vlny, optika, relativita*. Brno: VUTIUM, 2003. 141 s. ISBN: 80-274-1868-0.
- [95] SMITH, W. J. *Modern Optical Engineering*. 4th edition. McGraw-Hill Professional, 2007. 764 s. ISBN: 978-0-07-147687-4.
- [96] DEVEREUX, H., SMALLEY, M. Are Infra Red Illuminators Eye Safe? In *Proceedings of the 29th Annual International Carnahan Conference on Security Technology*. October 1995, s. 480-481. ISBN: 0-7803-2627-X.
- [97] TOMASI, C., KANADE, T. Detection and Tracking of Point Features. *Technical Report CMU-CS-91-132*. Carnegie Mellon University, 1991.
- [98] NAKANO, T. et al. Blink Measurement by Image Processing and Application to Warning of Driver's Drowsiness in Automobiles. In *Proceedings of the IEEE International Conference on Intelligent Vehicles*. Stuttgart, 1998, s. 285-290.

- [99] Ara Nefian Face Recognition Page. 1999 [cit. 2008-04-15]. Dostupný z WWW: <http://www.anefian.com/face_reco.htm>. Georgia Tech face database.
- [100] CMU Image Data Base: face [online]. 1997 [cit. 2008-04-10]. Dostupný z WWW: <http://vasc.ri.cmu.edu/idb/html/face/facial_expression/index.html>. Cohn-Kanade AU-Coded Facial Expression Database.
- [101] Face Recognition Data [online]. 1997 [cit. 2008-04-12]. Dostupný z WWW: <<http://cswww.essex.ac.uk/mv/allfaces/index.html>>. Collection of Facial Images.
- [102] Face Recognition Homepage [online]. 2005 [cit. 2008-03-22]. Dostupný z WWW: <<http://www.face-rec.org/>>.
- [103] HumanScan : BioID : Downloads : BioID Face Database [online]. c2004 [cit. 2008-04-22]. Dostupný z WWW: <<http://www.bioid.com/downloads/facedb/index.php>>. The BioID Face Database.
- [104] IIT-NRC: facial Video Database [online]. 2004 [cit. 2008-04-20]. Dostupný z WWW: <<http://www.perceptual-vision.com/db/video/faces/cvglab/>>.
- [105] MMI Face Database [online]. 1998 [2004] [cit. 2008-04-16]. Dostupný z WWW: <<http://www.mmifacedb.com/>>.
- [106] The BANCA Database [online]. 2004 [cit. 2008-04-01]. Dostupný z WWW: <<http://www.ee.surrey.ac.uk/CVSSP/banca/>>.
- [107] The Database of Faces [online]. c2002 [cit. 2008-04-03]. Dostupný z WWW: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>. MMI Face Database.
- [108] The University of Oulu Physics-Based Face Database [online]. c1994 [cit. 2008-04-20]. Dostupný z WWW: <<http://www.ee.oulu.fi/research/imag/color/pbfd.html>>.
- [109] The VidTIMIT Audio-Video Database [online]. c2001 [cit. 2008-04-20]. Dostupný z WWW: <<http://www.itee.uq.edu.au/~conrad/vidtimit/>>.
- [110] Yale Face Database [online]. 1997 [cit. 2008-05-19]. Dostupný z WWW: <<http://cvc.yale.edu/projects/yalefaces/yalefaces.html>>.
- [111] Yale Face Database B [online]. 2001 [cit. 2008-04-19]. Dostupný z WWW: <<http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html>>.
- [112] ZJU Eyeblink Database [online]. 2007 [cit. 2008-07-20]. Dostupný z WWW: <http://www.stat.ucla.edu/~gpan/data/db_blink.html>.
- [113] Canadian Forces - Combat Camera [online]. 2008 [cit. 2007-03-16]. Dostupný z WWW: <<http://www.combatcamera.forces.gc.ca/common/combatcamera/news/>>.
- [114] JENŠÍK L. IR světlo pro noční vidění bezpečnostní kamery [online]. Elektroamater.cz, [cit. 2. 06. 2008] Dostupné z URL:

<<http://www.elektroamater.cz/ir-svetlo-pro-nocni-videni-bezpecnostni-kamery/>>.

- [115] LIPINSKI, B. Iris Recognition: Detecting the Iris [online]. 2004 [cit. 2008-29-04]. Dostupné z WWW: <<http://cnx.org/content/m12489/latest/>>.
- [116] Driving Related Facts and Figures [online]. 2001 [cit. 2008-07-20]. Dostupný z WWW: <<http://www.driveandsurvive.co.uk/cont5.htm>>.

Seznam vlastních prací

- [117] VLACH, J.; RAJMIC, P.; PŘINOSIL, J.; VYORAL, J.; MÍČA, I. Optimized Discrete Wavelet Transform to Real-time Digital Signal Processing, Springer Verlag. *Mobile and Wireless Communication Networks*. 2007, roč. 2007, č. 9, s. 514-520. ISSN: 1571-5736.
- [118] PŘINOSIL, J.; VLACH, J. Face Detection in Image with Complex Background, Springer Verlag. *Mobile and Wireless Communication Networks*. 2007, roč. 2007, č. 9, s. 533-544. ISSN: 1571-5736.
- [119] RAJMIC, P.; VLACH, J. Forward Segmented Wavelet Transform. In *Proceedings of the 7th Nordic Signal Processing Symposium*. Reykjavik, 2006. s. 162-165.
- [120] VLACH, J.; RAJMIC, P. Using the Wavelet Transform in Real-time Digital Signal Processing. In *ICSES '06 International Conference on Signals and Electronic Systems (Proceedings volume 2 of 2)*. Lodz, 2006. s. 721-724. ISBN: 83-921172-5-5.
- [121] RAJMIC, P.; VLACH, J. Real-time Audio Processing via Segmented Wavelet Transform. In *Proceedings of the 10th International Conference on Digital Audio Effects (DAFx-07)*. Bordeaux: University of Bordeaux 1, 2007. s. 55-58. ISBN: 978-88-901479-1-3.
- [122] MÍČA, I.; PŘINOSIL, J.; VLACH, J. Framework for Support of Digital Signal Processing Education. In *Proceedings of the Telecommunications and Signal Processing(TSP)*. Brno, 2007, s. 97-100. ISBN: 978-80-214-3445-5.
- [123] PŘINOSIL, J.; VLACH, J. Human face localization in color images. In *Proceedings of the 12th IFIP International Conference*. Prague, 2007, s. 533-544. ISBN: 978-0-387-74158-1.
- [124] VLACH, J.; RAJMIC, P.; PŘINOSIL, J.; VYORAL, J.; MÍČA, I. New algorithm of discrete wavelet transform optimized to real-time digital signal processing. In *Proceedings of the 12th IFIP International Conference*. Prague, 2007, s. 514-520. ISBN: 978-0-387-74158-1.
- [125] VLACH, J.; PŘINOSIL, J. Lokalizace obličejů v obraze s komplexním pozadím. *Elektrorevue - Internetový časopis* (<http://www.elektrorevue.cz>), 2007, roč. 2007, č. 04, s. 1-12. ISSN: 1213-1539.

- [126] RAJMIC, P., VLACH, J. Method of Segmented Wavelet Transform for Real-time Signal Processing. In *Proceedings of Audio Technologies and Processing (ATP 2005)*. Prahe, 2005, s. 36-40.
- [127] VLACH, J. Modení hudební efekty využívající waveletovou transformaci. In *Sborník konference: Elektrotechnika a informatika 2005, část druhá - Elektronika*. Plzeň, 2005. s. 149-152. ISBN: 80-7043-374-4.
- [128] RAJMIC, P.; VLACH, J. Forward and Inverse Segmented Wavelet 2006 Transform. In *Research in Telecommunication Technology 2006 Proceedings - Part I*. Nové město na Moravě, 2006, s. 134-139. ISBN: 80-214-3243-8.

CURRICULUM VITÆ

Jméno: Ing. Jan Vlach

Narozen: 7.6.1982 v Prostějově

Kontakt: jan.vlach@phd.feec.vutbr.cz

Vzdělání

- 1993 – 2000 Gymnázium Jiřího Wolker, Prostějov
- 2000 – 2005 Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, magisterské studium oboru Elektronika a sdělovací technika
- 2005 – 2008 Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, magisterské studium oboru Teleinformatika

Účast na řešení projektů

V rámci doktorského studia jsem se podílel na řešení těchto projektů:

- MŠMT 1442/2007/F1 Inovace výuky předmětu Základy počítačové sazby a grafiky.
- MŠMT 972/2007/G1 Využití moderních algoritmů waveletové transformace pro kompresi multimediálních signálů a jejich začlenění do výuky.
- GA 102/07/1303 Nelineární metody zvýraznění řeči.
- GA102/04/1097 Zvýrazňování řečového signálu zamaskovaného v šumu.
- AV 1ET301710509 Sofistikované potlačovače hluku a poruch při přenosu řečových signálů pro pevné a mobilní sítě nové generace.
- MŠMT MSM0021630513 Elektronické komunikační systémy a technologie nových generací (ELKOM)

Pedagogické aktivity

Po celou dobu doktorského studia jsem se podílel na výuce předmětů Komunikační sítě a techniky, Komunikační a sdělovací technika, Číslicové zpracování signálů. Vedl jsem 12 bakalářských prací a 5 diplomových prací.

Další aktivity

V roce 2006 jsem zúčastnil dvoutýdenního workshopu na téma Verbální a neverbální komunikace ve Vietri sul Mare (Itálie). V roce 2007 jsem absolvoval tříměsíční odbornou stáž na IIASS Vietri sul Mare (Itálie). Svou práci publikuji na zahraničních i tuzemských konferencích.