

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE

Brno, 2016

Stanislava Serečunová



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

AUTOMATICKÁ REGULACE HLASITOSTI PODLE VZDÁLENOSTI POSLUCHAČE

MUSIC VOLUME ADJUSTMENT BASED ON DISTANCE DETECTION

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Stanislava Serečunová

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Oto Janoušek, Ph.D.

BRNO 2016

Bakalářská práce

bakalářský studijní obor **Biomedicínská technika a bioinformatika**

Ústav biomedicínského inženýrství

Studentka: Stanislava Serečunová

ID: 165021

Ročník: 3

Akademický rok: 2015/16

NÁZEV TÉMATU:

Automatická regulace hlasitosti podle vzdálenosti posluchače

POKYNY PRO VYPRACOVÁNÍ:

1) Nastudujte problematiku detekce obličeje ve videosekvenci. 2) Navrhněte způsob odhadu vzdálenosti uživatele od kamery, založený na vyhodnocení parametrů detekovaného obličeje. Preferujte řešení pracující v reálném čase. 3) Navrhněte způsob regulace hlasitosti hudby podle vzdálenosti uživatele tak, aby posluchač vnímal hudbu pořád stejně hlasitě. 4) V programovém prostředí Matlab vytvořte aplikaci pro přehrávání hudby, jejíž hlasitost se bude v reálném čase adaptovat podle vzdálenosti posluchače od kamery. 5) Ověřte funkčnost aplikace na skupině alespoň deseti dobrovolníků. 6) Proveďte diskuzi získaných výsledků a na jejich základě vytvořte doporučení pro zajištění optimálních podmínek snímání videosekvence.

DOPORUČENÁ LITERATURA:

[1] VIOLA, Paul a Michael J. JONES. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer society conference on computer vision and pattern recognition. 2001, (1): 511-518.

[2] LI, S. a Anil K. JAIN. Handbook of face recognition. 2nd ed. New York: Springer, 2011, 699 s. ISBN 978-08-7299-314.

Termín zadání: 8.2.2016

Termín odevzdání: 27.5.2016

Vedoucí práce: Ing. Oto Janoušek, Ph.D.

Konzultant bakalářské práce:

prof. Ing. Ivo Provazník, Ph.D., předseda oborové rady

UPOZORNĚNÍ:

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

ABSTRAKT

Táto práca sa zaoberá dvoma metódami detekcie tváre vo videosekvencii. Na základe testovania je vybraná metóda KLT. Práca je ďalej zameraná na návrh prepočtu vzdialenosti užívateľa od kamery v reálnom čase pomocou vyhodnotenia parametrov detegovanej tváre. Podľa vzdialenosti užívateľa od kamery je navrhnutá regulácia hlasitosti tak, aby užívateľ pri jej zmene vnímal hudbu stále rovnako.

KLÚČOVÉ SLOVÁ

detekcia tváre v reálnom čase, Viola-Jones, CAMShift, KLT, vzdialenosť, hlasitosť

ABSTRACT

This work is about two methods of face detection in video sequencing. According to testing the KLT method is chosen. The work is focused on user distance calculation from camera in real time by means of evaluating parameters of detected face. According the distance of user from camera sound regulation is set, so the user can perceive and feel the music on the same level.

KEYWORDS

face detection in real time, Viola-Jones, CAMShift, KLT, distance, volume

SEREČUNOVÁ, S. *Automatická regulace hlasitosti podle vzdálenosti posluchače*: bakalárska práca, Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrstva, 2016. 56 s. Vedúci bakalárskej práce bol Ing. Oto Janoušek, Ph.D.

PREHLÁSENIE

Prehlasujem, že som svoju bakalársku prácu na tému „Automatická regulace hlasitosti podle vzdálenosti posluchače“ vypracovala samostatne pod vedením vedúceho bakalárskej práce, využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej bakalárskej práce ďalej prehlasujem, že v súvislosti s vytvorením tejto bakalárskej práce som neporušila autorské práva tretích osôb, najmä som nezasiahla nedovoleným spôsobom do cudzích autorských práv osobnostných alebo majetkových a som si plne vedomá následkov porušenia ustanovenia § 11 a nasledujúceho autorského zákona č. 121/2000 Sb., o autorskom práve, o právach súvisiacich s autorským právom a o zmene niektorých zákonov (autorský zákon), v znení neskorších predpisov, vrátane možných trestnoprávnych dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákonníka č. 40/2009 Sb.

V Brne dňa 27. 5. 2016

.....

(podpis autora)

POĎAKOVANIE

Ďakujem vedúcemu bakalárskej práce pánovi Ing. Otovi Janouškovi, Ph.D. za odborné vedenie, podnetné pripomienky a motivujúci prístup. Zároveň ďakujem svojej rodine a priateľom za ich podporu a pomoc pri písaní práce.

V Brne dňa 27. 5. 2016

.....

(podpis autora)

OBSAH

Zoznam obrázkov	VIII
Zoznam tabuliek	X
1 Úvod	1
2 Metódy detekcie a sledovania tváre	2
2.1 Detektor Viola-Jones	2
2.1.1 Integrálny obraz	3
2.1.2 Haarove vlnky a ich črty	3
2.1.3 Adaptive Boosting	4
2.1.4 Kaskádové zapojenie klasifikátorov	7
2.2 Metóda sledovania tváre pomocou farby pleti.....	7
2.3 Metóda sledovania tváre pomocou funkčných bodov	9
3 Testovanie a selekcia metód	11
3.1 Vplyv osvetlenia scény	11
3.2 Zmena otočenia tváre	12
4 Detekcia vzdialenosti objektu	15
5 Fyzikálna akustika	18
5.1 Zmena intenzity hlasitosti so vzdialenosťou.....	18
6 Implementácia programu v prostredí MATLAB	20
6.1 Kalibrácia kamery	21
6.1.1 Kalibračná metóda pomocou známej vlastnosti scény	22
6.1.2 Kalibračná metóda pomocou šachovnice	22
6.2 Porovnanie kalibračných metód.....	26
6.2.1 Priebeh kalibrácie kamery	27
6.3 Nastavenie zvuku a jeho regulácia.....	30

6.4	Detekcia tváre	31
7	Funkčnosť programu	33
7.1	Optimálne podmienky snímania videosekvencie	34
8	Záver	36
	Literatúra	37
	Zoznam skratiek	39
	Zoznam príloh	40

ZOZNAM OBRÁZKOV

Obr. 2.1: Kumulatívne súčty [3]	3
Obr. 2.2: Haarove črty na tvári	4
Obr. 2.3: Haarove črty [5].....	4
Obr. 2.4: Kaskádové zapojenie klasifikátorov [1]	7
Obr. 2.5: HSV model [7]	8
Obr. 2.6: Blokový diagram algoritmu CAMShift [7]	9
Obr. 2.7: Princíp metódy KLT	10
Obr. 3.1: (a) presvetlenie zľava, (b) frontálne presvetlenie, (c) presvetlenie sprava,.....	11
Obr. 3.2: (a) presvetlenie zľava, (b) frontálne presvetlenie, (c) presvetlenie sprava,.....	12
Obr. 3.3: Detekcia tváre pri rôznych natočeniach hlavy,.....	12
Obr. 3.4: Detekcia tváre splývajúcej s pozadím pri rôznych natočeniach hlavy,	13
Obr. 3.5: Mierne otáčanie hlavy počas detekcie	13
Obr. 3.6: Otáčanie hlavy o 90° vľavo a vpravo	14
Obr. 3.7: Detekcia tváre pri rôznych natočeniach hlavy,.....	14
Obr. 4.1: Exponenciálna závislosť obsahu detegovaného štvorca na vzdialenosti.....	16
Obr. 4.2: Lineárna závislosť obsahu detegovaného štvorca na vzdialenosti	17
Obr. 5.1: Vlnoplochy izotropného bodového zdroja [11].....	19
Obr. 6.1: Vývojový diagram navrhnutého programu	20
Obr. 6.2: Navrhnutý program	21
Obr. 6.3: Kalibrácia kamery	22
Obr. 6.4: Parametre šachovnice [16]	23
Obr. 6.5: Priebeh kalibrácie pomocou šachovnice	24
Obr. 6.6: Priemerná chyba reprojekcie v snímkach.....	25
Obr. 6.7: Rohové body šachovnice porovnávané medzi snímkami.....	25
Obr. 6.8: Umiestnenie kalibračnej šachovnice	27
Obr. 6.9: Graf závislosti obsahu detegovaného štvorca na vzdialenosti po kalibrácii	

kamery.....	28
Obr. 6.10: Kalibrácia kamery	29
Obr. 6.11: Presvetlená scéna počas kalibrácie kamery	29
Obr. 6.12: Výber zvukového záznamu	30
Obr. 6.13: Detekcia tváre	31
Obr. 6.14: Detegované tváre odlišnej farby	31
Obr. 6.15: Vpravo detegovaná tvár, vľavo nedetegovaná tvár	32
Obr. 6.16: Vpravo detegovaná tvár užívateľa, vľavo detegovaná tvár na podtlačí trička	32
Obr. 7.1: Detegovaná osoba v rôznych vzdialenostiach	33
Obr. 8.1: Detegované tváre	41

ZOZNAM TABULIEK

Tab. 6.1: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 1. osoby	26
Tab. 6.2: Priemer jednotlivých vypočítaných vzdialeností u 10 osôb kalibračnou metódou pomocou šachovnice	27
Tab. 7.1: Namerané hodnoty regulácie hlasitosti v rôznych vzdialenostiach	34
Tab. B.1: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 2. osoby	42
Tab. B.2: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 3. osoby	42
Tab. B.3: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 4. osoby	43
Tab. B.4: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 5. osoby	43
Tab. B.5: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 6. osoby	43
Tab. B.6: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 7. osoby	44
Tab. B.7: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 8. osoby	44
Tab. B.8: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 9. osoby	44
Tab. B.9: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 10. osoby	45

1 ÚVOD

Počítačové videnie je technická disciplína, ktorej snahou je priblížiť sa ľudskému videniu. Zrak je zdrojom prevládajúcej väčšiny informácií o okolitom svete. Veľkú časť pri vyhodnotení vizuálnej informácie tvorí inteligencia človeka, ktorá reprezentuje nadobudnuté poznatky a skúsenosti o okolitom svete. [1] Medzi technologické aplikácie počítačového videnia patrí medicínske modelovanie (počítačová tomografia), medicínska diagnostika (aplikácie v oblasti rádiológie, dermatológie, forenznej medicíny), interaktívne systémy s ľuďmi, porovnanie a vyhodnotenie vizuálnych a iných vlastností snímaných objektov a mnohé ďalšie aplikácie v iných odvetviach priemyslu. [2]

Snahou počítačového videnia je porozumieť trojrozmernej scéne, napríklad osoba nachádzajúca sa v miestnosti. Ak sa pozorovateľ zameria len na tvár osoby, dokáže ju v okamihu identifikovať. V scéne sa v miestnosti nachádzajú okrem osoby aj predmety, ktoré ju obklopujú, napríklad skrine, tapeta na stene, lampa a iné. Aby systém dokázal rozpoznať osobu v scéne, musí o nej vedieť bližšie informácie, t.j. akú má farbu, jej veľkosť voči ostatným objektom v miestnosti a ďalšie. Na základe týchto parametrov dokáže systém rozpoznať tvár a určiť jej súradnice v trojrozmernom priestore.

Úlohou tejto bakalárskej práce je oboznámiť sa s metódami detekcie tváre vo videosekvencii. Práca je rozdelená na osem kapitol vrátane úvodu a záveru práce. Detektor Viola-Jones, ktorý sa používa na detekciu polohy tváre vo videosekvencii, je podrobne opísaný v druhej kapitole. Tento detektor tvorí základ sledovania metód detekcie tváre, v tomto prípade metóda detekcie a sledovania tváre pomocou farby kože (Continuously Adaptive Mean Shift algoritmus) a metóda detekcie a sledovania tváre pomocou bodov (Kanade-Lucas-Tomasi detektor).

V tretej kapitole je vykonané testovanie jednotlivých metód a následná selekcia metódy, ktorá bude najvhodnejšia pre ďalšie kroky tejto práce. Práca je v štvrtej a piatej kapitole zameraná na návrh prepočtu vzdialenosti detegovanej osoby od kamery a následnej synchronizácie hlasitosti hudby so vzdialenosťou poslucháča. Implementácia programu v programovacom prostredí MATLAB a zároveň jeho obsluha v grafickom užívateľskom rozhraní GUI je opísaná v šiestej kapitole. Siedma kapitola je venovaná overeniu funkčnosti navrhnutého programu pre reguláciu hlasitosti.

2 METÓDY DETEKCIE A SLEDOVANIA TVÁRE

K sledovaniu tváre sa v súčasnosti využívajú rôzne metódy, z ktorých každá pozostáva z odlišných parametrov jej rozpoznania.

- Metóda detekcie a sledovania tváre pomocou farby kože (Continuously Adaptive Mean Shift algoritmus)
- Metóda detekcie a sledovania tváre pomocou bodov (Kanade-Lucas-Tomasi detektor)

Všetky tieto metódy používajú na detekciu polohy tváre vo videosnímke detekčný algoritmus Viola-Jones.

2.1 Detektor Viola-Jones

Objektový detektor Viola-Jones, pôvodne nazývaný „Rapid Object Detection using a Boosted Cascade of Simple Features“, bol prvýkrát zverejnený P. Violou a M. J. Jonesom v roku 2001. [3] Výhodou tohto detektora je jeho rýchlosť, nezávislosť korektnosti detekcie na osvetlení a veľkosti sledovaného objektu. Nevýhodou je škálová a rotačná nezávislosť, čo je výpočtovo náročné. Popríklad prekrytie objektov alebo zmena tvaru nerigidných objektov veľmi znižujú pravdepodobnosť správnej detekcie. [4] Tento objektový detektor pracuje len so šedo-tónovými obrazmi.

Viola-Jones metóda detekcie je založená na globálnom opise. Používa metódu klzavého okna, ktoré vyberie určitý región obrazu a ten sa vyhodnocuje porovnávaním výskytu príznakov, t.j. svetlých a tmavých obdĺžnikových oblastí. [4]

Detektor sa skladá z troch základných častí:

- vytvorenie integrálneho obrazu
- určenie Haarových črt (Haar features) pomocou Haarových vlniek (Haar-like features)
- klasifikácia príznakov pomocou algoritmu AdaBoost (Adaptive Boosting) [3]

2.1.1 Integrálny obraz

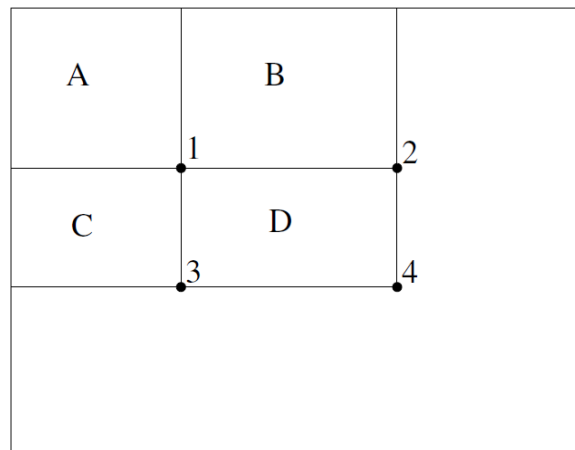
Každý bod integrálneho obrazu odpovedá súčtu hodnôt všetkých predchádzajúcich bodov:

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (2.1)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y), \quad (2.2)$$

kde $s(x, y)$ je kumulovaný súčet hodnôt v riadku obrazu, $i(x, y)$ sú hodnoty vstupného obrazu a $ii(x, y)$ sú hodnoty integrálneho obrazu. [3]

Pre ľubovoľne veľkú oblasť stačí poznať štyri referenčné polia. Súčet pixelov v obdĺžniku D (viď Obr. 2.1) je počítaný týmito štyrmi referenčnými poľami. Hodnota v bode A je vypočítaná ako súčet pixelov v obdĺžniku A, hodnota v bode 2 ako súčet obdĺžnikov A + B, hodnota v bode 3 ako súčet obdĺžnikov A + C a hodnota v bode 4 je vypočítaná ako súčet obdĺžnikov A + B + C + D. Suma obdĺžnika D je vypočítaná ako $4 + 1 \cdot (2+3)$. [3]

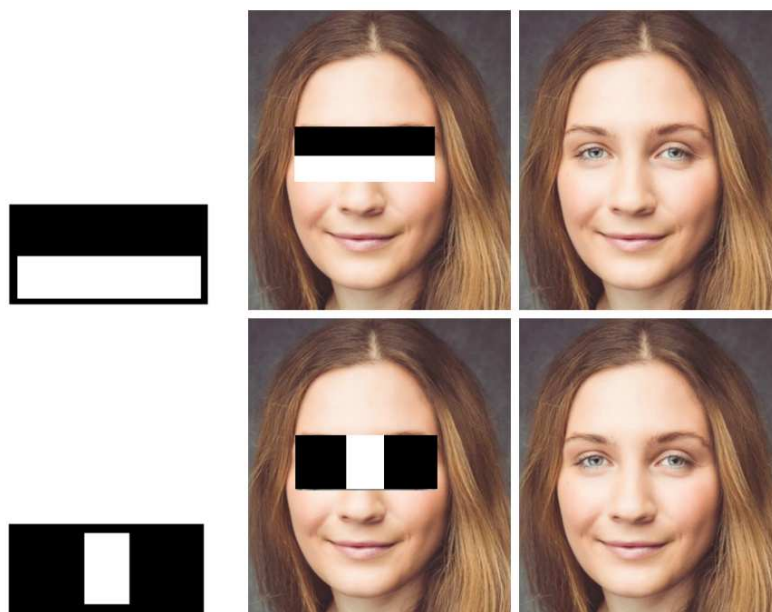


Obr. 2.1: Kumulatívne súčty [3]

2.1.2 Haarove vlnky a ich črty

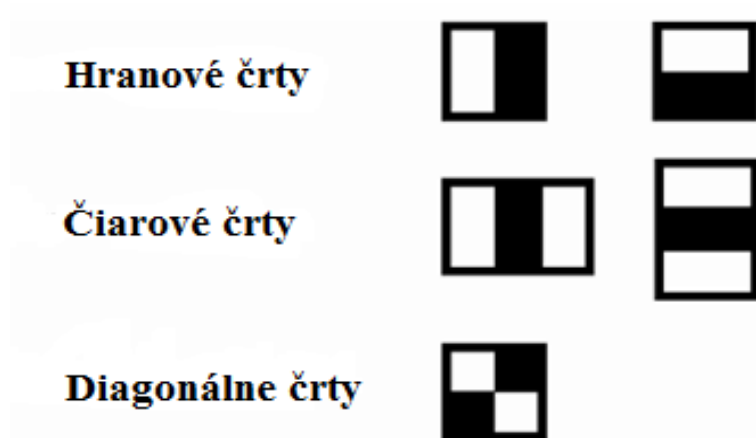
Haarove vlnky využívajú pre rýchly výpočet integrálový obraz, o ktorom už bolo zmienené v kapitole 2.1.1. Výhodou týchto statických vlniek je ich použitie pri detekcii ľudskej tváre, pretože pravidelné tvary deteguje vynikajúco. Avšak nevýhodou je detekcia rozličných tvarov ľudského tela.

Hodnota Haarovej črty sa počíta ako rozdiel sumy pixelov obrazu, ktorý odpovedá svetlej časti a sumy pixelov, ktorým odpovedá tmavá časť (viď Obr. 2.2).



Obr. 2.2: Haarove črty na tvári

Tieto vlnky môžu byť tvorené rôznymi počtami svetlých a tmavých obdĺžnikových oblastí, ako je možné vidieť na Obr. 2.3. Vlnky môžu byť tvorené dvoma (hranová črta), tromi (čiarová črta), štyrmi (diagonálna črta) obdĺžnikovými oblasťami. [5]



Obr. 2.3: Haarove črty [5]

2.1.3 Adaptive Boosting

Adaptive Boosting, skrátene nazývaný ako AdaBoost, je algoritmus, ktorý vychádza z metódy strojového učenia nazývaného *boosting*. [5] Vstupom algoritmu je trénovaná množina a výstupom je klasifikátor dvoch tried:

$$H(x) = \text{sign} \left(\sum_{t=1}^T a_t h_t(x) \right), \quad (2.3)$$

kde výsledný klasifikátor je založený na lineárnej kombinácii rozhodnutí niekoľkých slabších klasifikátorov $h_t(x)$. Tieto slabšie klasifikátory sú vyberané z množiny klasifikátorov H , pričom ich lineárnou kombináciou vzniká nelineárny silný klasifikátor $H(x)$. T odpovedá počtu iterácií a a_t lineárnej kombinácii kroku učenia.

Kaskádovým zret'azením slabších príznakov dosahuje výsledný silný klasifikátor lepšie výsledky.

Popis procesu učenia klasifikačného algoritmu Adaboost je realizovaný v niekoľkých krokoch: [3] [6]

Vstup:

$$S = ((x_1, y_1), \dots, (x_m, y_m)), \text{ počet iterácií } T, \quad (2.4)$$

kde S je vstup, x_1 až x_m je meranie a y_1 až y_m je skutočná príslušnosť merania.

Inicializácia váh:

$$D_i(i) = \frac{1}{m}, \quad (2.5)$$

kde D_i sú váhy a m je veľkosť trénovanej množiny.

Cyklus pre $t = 1, \dots, T$:

a. Výber klasifikátorov na základe váženej trénovanej chyby

$$\varepsilon_j = \sum_{i=1}^m D_t I[y_t \neq h_j(x_i)], \quad (2.6)$$

$$h_j = \arg \min \varepsilon_j$$

$$h_j \in H$$

kde ε_j je vážená trénovaná chyba, m je veľkosť trénovanej množiny, y_t je skutočná príslušnosť merania, h_j je slabý klasifikátor, I je výrok, ktorý vracia 1 pokiaľ je výrok pravdivý a 0 ak je nepravdivý. [6]

b. Pokiaľ:

$$\varepsilon_t = 0 \text{ alebo } \varepsilon_t \geq \frac{1}{2}, \text{ potom ukončiť cyklus} \quad (2.7)$$

Nastavenie:

$$a_t = \frac{1}{2} \log \left(\frac{1 - \varepsilon_t}{\varepsilon_t} \right), \quad (2.8)$$

kde a_t je lineárna kombinácia kroku učenia a ε_t je vážená trébovaná chyba.

c. Úprava váh:

$$D_{t+1}(i) = \frac{D_t e^{a_t y_i h_t(x_i)}}{Z_t}, \quad (2.9)$$

$$Z_t = \sum_{i=1}^m D_t(i) e^{a_t y_i h_t(x_i)}, \quad (2.10)$$

kde a_t je lineárna kombinácia kroku učenia, Z_t charakterizuje úpravu váh, ktorá spôsobuje zväčšenie váhy D_t chybné klasifikovaného merania x_i a zmenšenie správne klasifikovaného merania y_i . Preto bude v ďalšom kroku vyhľadávaný slabší klasifikátor h_t , ktorý bude musieť lepšie klasifikovať doteraz chybné klasifikované merania x_i .

Výstup:

$$H(x) = \text{sign} \left(\sum_{t=1}^T a_t h_t x \right), \quad (2.11)$$

kde výstupom algoritmu je trébovaná množina silného klasifikátora $H(x)$, ktorá sa skladá z dvojíc (x_i, y_i) , kde x_i znamená meranie a y_i skutočná príslušnosť merania x_i patriacej k jednej z dvoch tried $\{-1, 1\}$, T je počet iterácií a a_t lineárna kombinácia kroku učenia. [6] AdaBoost na rozdiel od *boostig-u* váhuje trébované množiny pomocou váh D_t . Tieto váhy sú nastavené na začiatku rovnomerne. Učenie algoritmu Adaboost prebieha v slučke, kde je v každej jednej slučke potrebné vykonať nasledujúce kroky:

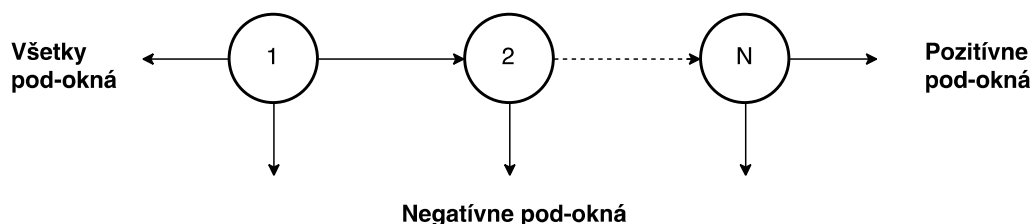
- vyhľadanie slabého klasifikátora s najmenšou chybou pri váhovaní D_t trébovaných dát
- overenie chyby klasifikátora, ktorá nepresiahla hodnotu 0,5
- súčet koeficienta slabého klasifikátora v lineárnej kombinácii $H(x)$
- aktualizovanie váh trébovanej množiny D_t

Algoritmus vyhľadáva slabý klasifikátor podľa zvolených parametrov. Ak chyba klasifikátora prekročí hodnotu 0,5, tak vstupná požiadavka je nespĺnená a tým nie je garantovaná konvergencia algoritmu. Dôsledkom aktualizácie váh je, že váha chybné

klasifikovaného merania sa zväčší a váha správne klasifikovaného merania sa zmenší. Následne bude v ďalšom kroku vyhľadávaný slabý klasifikátor, ktorý bude lepšie klasifikovať doteraz chybne uskutočnené meranie. Z toho vyplýva, že Adaboost redukuje trénovanú chybu pomocou exponenciálnej závislosti na rastúcom počte klasifikátorov. [6]

2.1.4 Kaskádové zapojenie klasifikátorov

Algoritmus prechádza celý snímok pomocou jednotlivých pod-okien. Takto sa celý proces stáva výpočtovo náročný. Napríklad pre snímok, ktorý má rozmery 320 x 240 pixelov je nevyhnutné spracovať vyše 500 000 pod-okien. Pre zefektívnenie a zníženie priemernej doby, počas ktorej detektor prehľadáva každé pod-okno bolo zavedené kaskádové zapojenie klasifikátorov. Toto zapojenie je riešené tak, že sa vytvoria klasifikátory, ktoré dokážu vybrať všetky pozitívne črty, t.j. pod-okná obsahujúce tvár a zároveň negatívne črty, ktoré tvár neobsahujú. Detektor počíta s faktom, že negatívnych pod-okien sa v snímku nachádza podstatne viac a na základe toho sa v prvom stupni kaskády snaží zamietnuť čo najväčší počet negatívnych pod-okien. Až v ďalších stupňoch kaskády spracováva pozitívne pod-okná (viď Obr. 2.4).



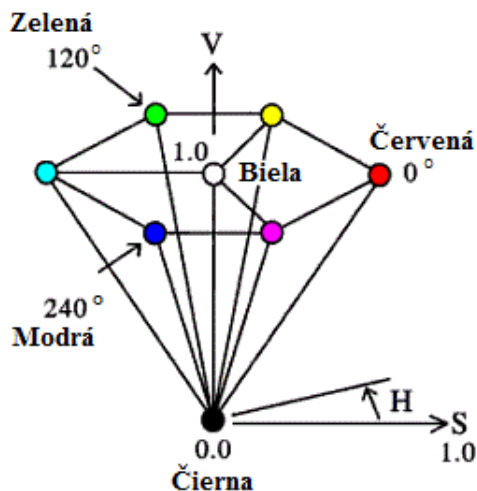
Obr. 2.4: Kaskádové zapojenie klasifikátorov [1]

2.2 Metóda sledovania tváre pomocou farby pleti

Metóda bola vyvinutá pre efektívne sledovanie tváre na základe percepčného užívateľského rozhrania. V origináli je nazývaná ako Continuously Adaptive Mean Shift algoritmus, ktorý sa označuje tiež ako CAMShift. Hlavnú časť predstavuje neparametrická technika, ktorá sa snaží nájsť v narastajúcej hustote gradientov vrchol pravdepodobnosti rozdelenia hustoty. Tento spôsob sa uvádza v literatúre ako *the mean-shift algorithm*. [7]

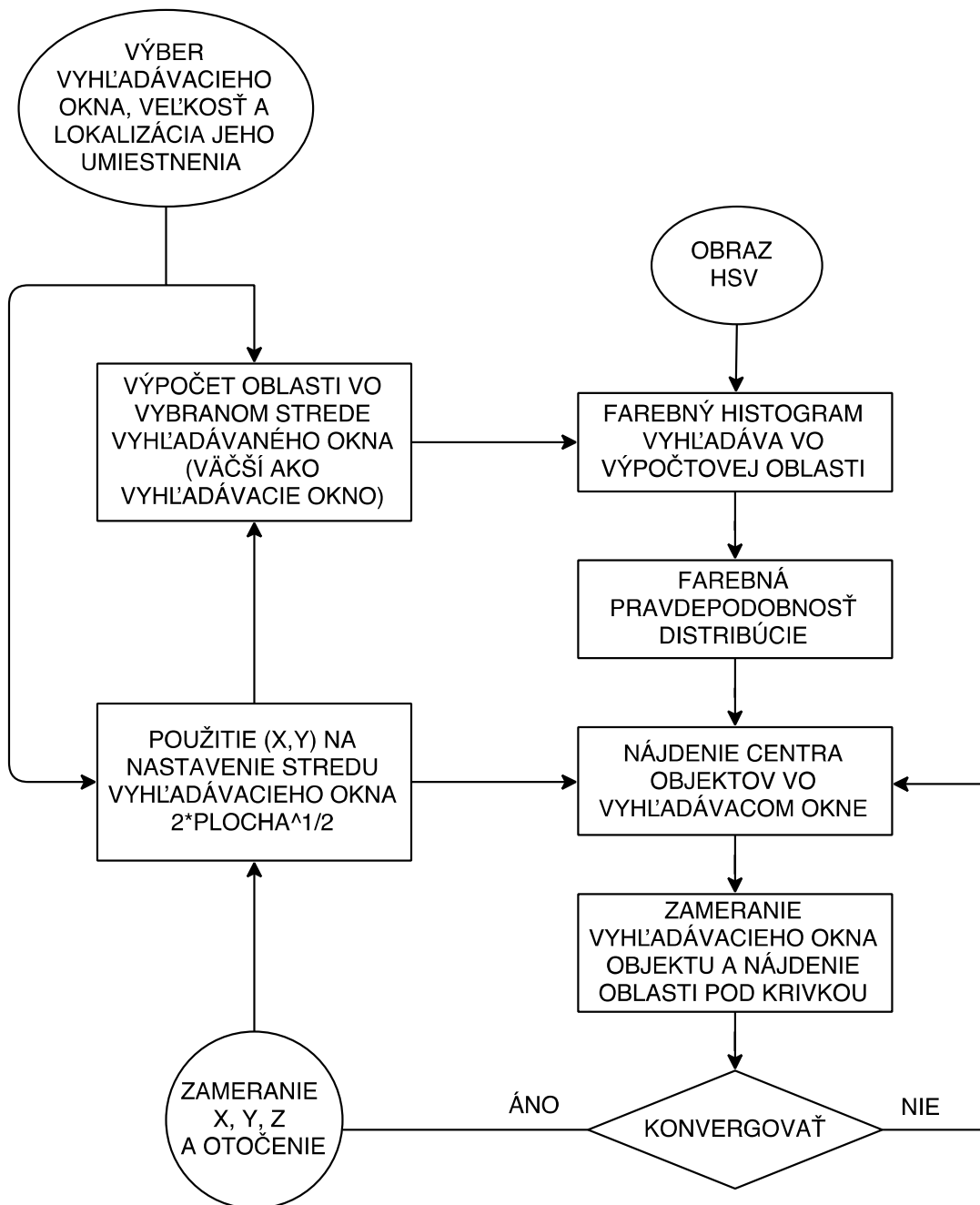
Tento algoritmus používa pre sledovanie objektu jednorozmerný histogram získaný z objektového modelu. Histogram sa skladá z prvého kanála modelu HSV vo farebnom priestore, t.j. kanál *hue*, ktorý v tomto modeli reprezentuje farebný odtieň

(viď Obr. 2.5). Tento odtieň udáva prevládajúcu farbu v rozmedzí uhlov 0° - 360° .



Obr. 2.5: HSV model [7]

Pre sledovanie tváre algoritmus CAMShift sleduje roviny X, Y a priestor, ktorý predstavuje farebné rozloženie tónu pleti na tvári. Priestor je úmerný rovine Z, t.j. vzdialenosti od kamery (viď Obr. 2.6). Schopnosť algoritmu sledovať otáčanie hlavy je vnímaný ako ďalší stupeň výhody tohto algoritmu.



Obr. 2.6: Blokový diagram algoritmu CAMShift [7]

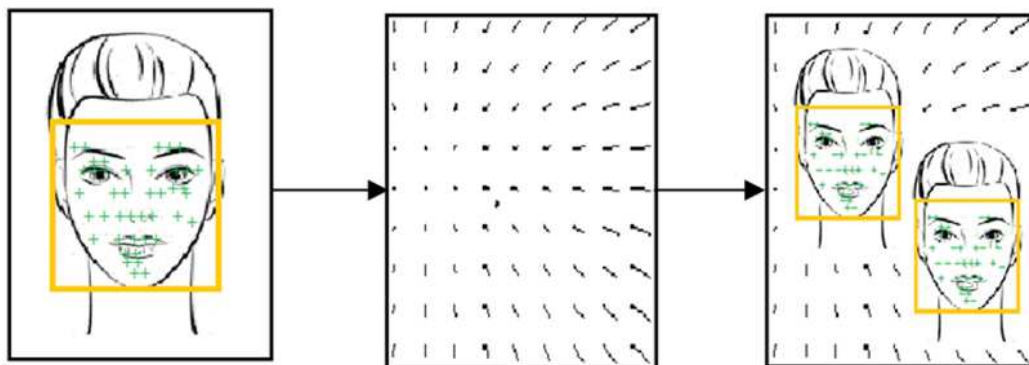
2.3 Metóda sledovania tváre pomocou funkčných bodov

Základnou myšlienkou tejto metódy je sledovanie snímok, ktoré sú snímané v okamžikoch za sebou a úzko spolu súvisia. Táto metóda nazývaná tiež ako KLT detektor (*Kanade-Lucas-Tomasi*) sleduje rad hraničných bodov vo videosnímke na základe výpočtu optického toku, ktorý získava informácie o rozmeroch pohybu

objektov za časový interval dt (viď Obr. 2.7). V dôsledku čoho môžeme vyjadriť vzájomný vzťah:

$$I(x, y, t) = I(x + dx, y + dy, t + dt), \quad (2.12)$$

kde $I(x,y,t)$ je obraz v snímke. Ak by sa v obraze posunul pixel o vzdialenosť dx a dy v ďalšej následnej snímke v čase dt , tak by sme mohli tvrdiť, že pixely v oboch obrazoch v časoch t a $t + dt$ sú identické. [8]



Obr. 2.7: Princíp metódy KLT

Sledovanie jedného pixelu nemôže byť použité, pretože je nepravdepodobné, že má výrazný jas vo vzťahu s pixelom predchádzajúceho snímku. Práve preto detektor využíva na sledovanie celé okno pixelov. To vedie k použitiu bohatého modelu s veľkým počtom parametrov. Tento model môže byť definovaný ako:

$$J(x) = I(x - d) + n(x), \quad (2.13)$$

kde $n(x)$ je šum, $J(x)$ je rovné $I(x,y,t)$ a $I(x - d)$ je rovné $I(x + dx, y + dy, t + dt)$. [8]

Vektor posunutia d je potom zvolený tak, aby sa minimalizoval chybný zvyšok určený nasledujúcim integrálom nad daným oknom W :

$$\varepsilon = \int_W [I(x - d) - J(x)]^2 w \, dx, \quad (2.14)$$

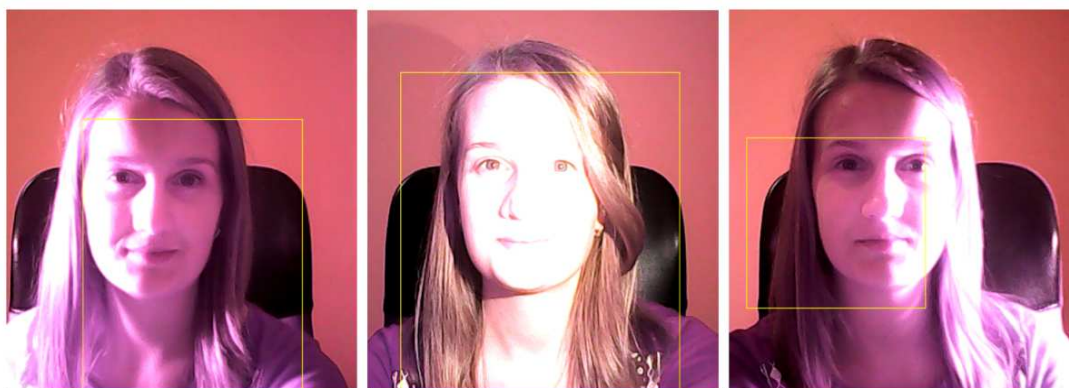
kde w je váhová funkcia, ktorá by v najjednoduchšom prípade mohla byť nastavená na hodnotu 1. [8]

3 TESTOVANIE A SELEKCIA METÓD

Metódy KLT a CAMShift boli testované na rôznych škálach detegovaných osôb. Experimentovalo sa ako bude ovplyvnená detekcia pri snímaní rôzneho pohlavia, človeka s okuliarmi, človeka s bradou, človeka s rozpustenými vlasmi, človeka s učesanými vlasmi vo vrkoči, zmeny otočenia tváre a zmeny osvetlenia detegovanej osoby. Zo všetkých experimentov sa výrazne líšili zmeny pozície tváre a osvetlenie scény pri detekcii snímanej osoby, ktoré budú zanalyzované v nasledujúcich častiach kapitoly.

3.1 Vplyv osvetlenia scény

Metóda CAMShift je citlivejšia na zmeny v jase ako metóda KLT. Pokiaľ snímanej osobe svieti svetlo do tváre, dochádza k presvetleniu scény (vid' Obr. 3.1). Pri presvetlení scény dochádza k saturácii CCD snímačov vedúcich k vypáleniu určitých pixelov, ktoré následne nie sú detegované ako tón pleti. Presvetlenie scény má za následok chybné detegovanie snímanej oblasti, tzn. zväčšenie alebo zmenšenie, či úplná strata snímanej osoby. Aj napriek presvetleniu scény metóda KLT zachováva významné hraničné body a tvár snímanej osoby deteguje spoľahlivo (vid' Obr. 3.2).



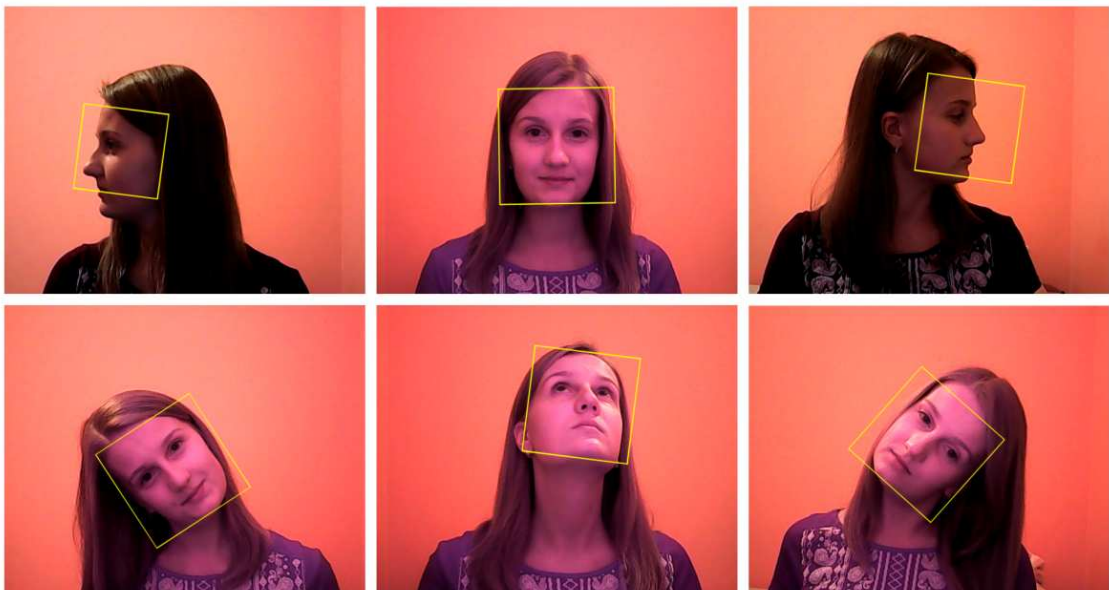
Obr. 3.1: (a) presvetlenie zľava, (b) frontálne presvetlenie, (c) presvetlenie sprava, metóda CAMShift



Obr. 3.2: (a) presvetlenie zľava, (b) frontálne presvetlenie, (c) presvetlenie sprava,
metóda KLT

3.2 Zmena otočenia tváre

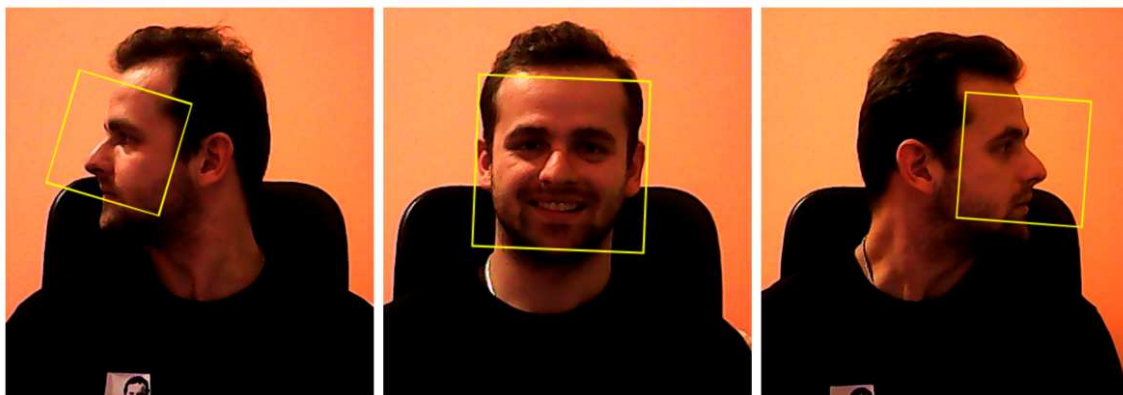
Pri testovaní bol kladený dôraz na rôznorodosť detekcie otočenia tváre. Metódu KLT je možné z pohľadu detekcie žltého rovinného útvaru, ktorý sa na videosnímke v oblasti tváre javí ako štvorec, označiť za spoľahlivú (viď Obr. 3.3).



Obr. 3.3: Detekcia tváre pri rôznych natočeniach hlavy,
metóda KLT

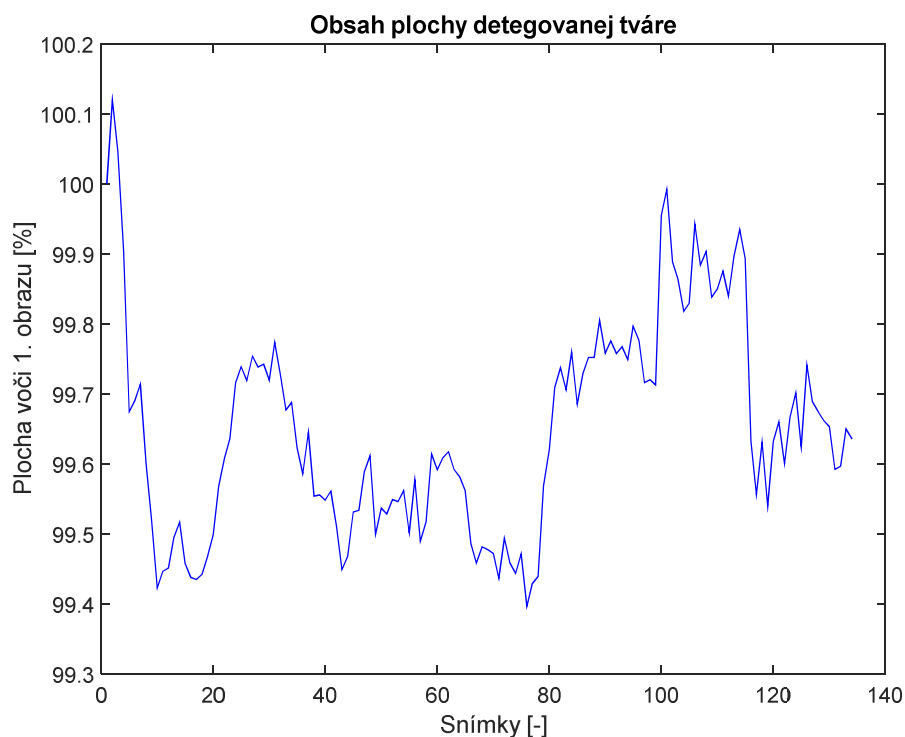
Pri snímaní osoby mužského pohlavia KLT metóda deteguje spoľahlivo zmeny pozícií tváre, ale dochádza k zmenšeniu detegovanej oblasti (viď Obr. 3.4). To môže byť spôsobené z dôvodu splývania brady s kreslom, resp. tým, že detegovaný žltý rovinný útvar sa javí ako štvorec, t.j. má rovnako dlhé strany. Je možné, že algoritmus

upravuje výšku detegovaného štvorca podľa jeho šírky, ktorá môže byť odvodená napríklad z rozstúpenia očí.

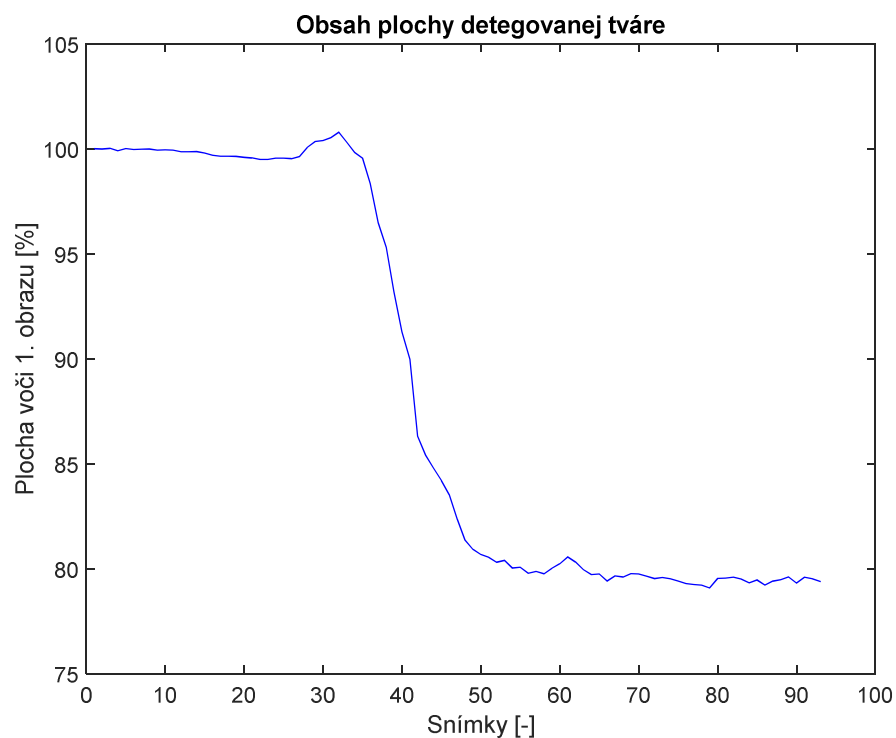


Obr. 3.4: Detekcia tváre splývajúcej s pozadím pri rôznych natočeniach hlavy, metóda KLT

Ďalej boli pozorované zmeny obsahu štvorca pri miernom otáčaní hlavy „vyjadrujúcej nesúhlas“ (vid' Obr. 3.5), kde štvorec svoje rozmery menil maximálne o cca 0,6 % a tento fakt by nemusel ovplyvniť ďalšie kroky detekcie, ktorej budú venované kapitoly 4 a 5. Ak bola hlava otáčaná o 90° na pravú stranu alebo na ľavú stranu obsah štvorca sa výrazne menil približne o 20% (vid' Obr. 3.6).

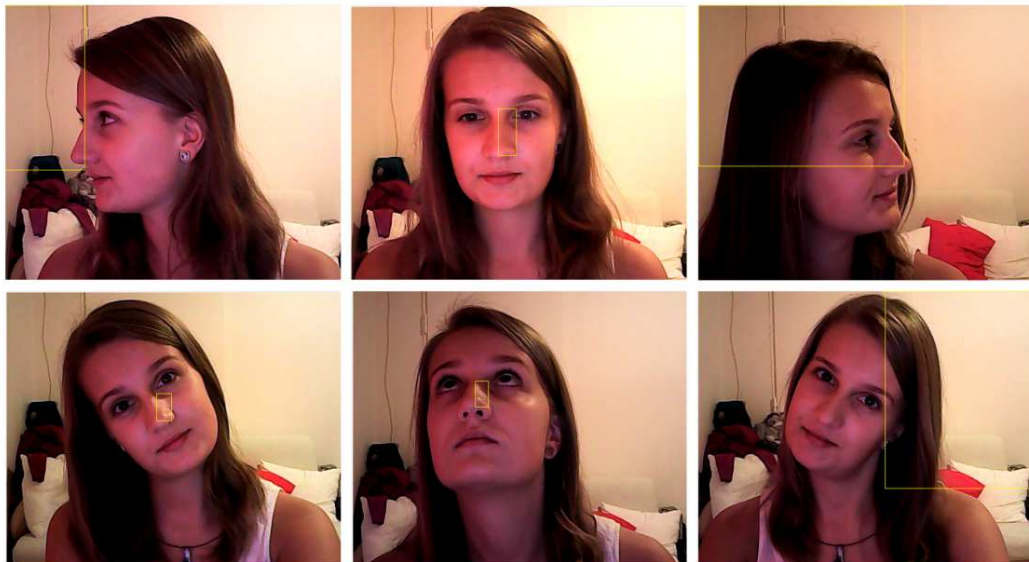


Obr. 3.5: Mierne otáčanie hlavy počas detekcie



Obr. 3.6: Otáčanie hlavy o 90° vľavo a vpravo

Metóda CAMShift je z pohľadu detekcie mnohoúhelníka, ktorý sa počas detekcie v čase neustále mení, nevyhovujúca. V tomto prípade bolo bezpredmetné zisťovať zmenu obsahu štvoruholníka, pretože už z Obr. 3.7 je vidieť jeho nestálosť.



Obr. 3.7: Detekcia tváre pri rôznych natočeniach hlavy,
metóda CAMShift

4 DETEKCIA VZDIALENOSTI OBJEKTU

K detekcii vzdialenosti bola využitá metóda sledovania KLT. Z funkčných bodov, ktoré algoritmus vyhľadá vo videosnímke, vytvorí spojením rohových bodov $x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4$ štvorec. Veľkosť strany štvorca bola vypočítaná pomocou Pytagorovej vety, ktorá je definovaná ako:

$$c = \sqrt{a^2 + b^2}, \quad (4.1)$$

kde strana a bola získaná rozdielom bodov x_2 a x_1 , strana b rozdielom bodov y_2 a y_1 .

Detekcia vzdialenosti bola následne implementovaná pomocou výpočtu obsahu štvorca S detegovaného vo videosnímke:

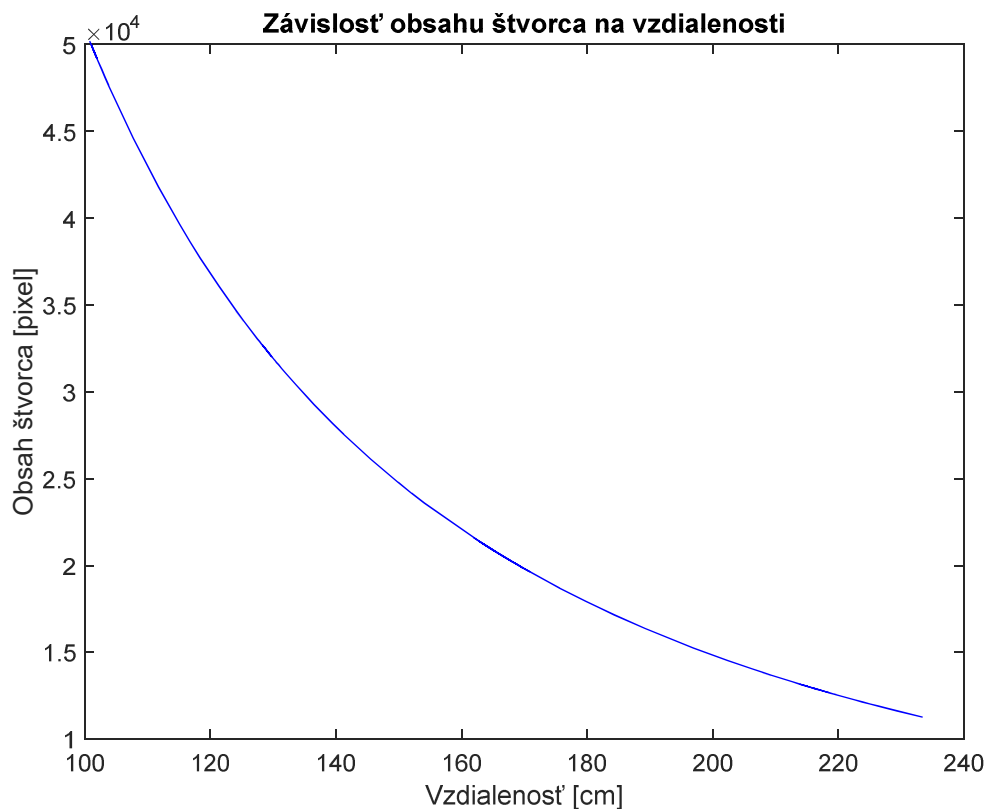
$$S = c^2, \quad (4.2)$$

kde c je strana detegovaného štvorca.

Pre následnú detekciu vzdialenosti bolo nutné zistiť, na základe akej funkcie sa mení vypočítaný obsah detegovaného štvorca v závislosti na vzdialenosti objektu od kamery. Táto funkcia bola zistená tak, že v miestnosti boli umiestnené detekčné značky, ktoré určovali presnú vzdialenosť sledovaného objektu od kamery. Objekt sa vzdľoval o určitú vzdialenosť, pričom bola pri každej vzdialenosti určená veľkosť obsahu štvorca v danom momente. Na základe takto nameraných údajov bolo možné určiť funkciu závislosti. Na Obr. 4.1 je vidieť, že obsah štvorca so vzdialenosťou od zdroja exponenciálne stúpa. Funkciu je možné definovať ako:

$$y = a^x, \quad (4.3)$$

kde rovnica každej reálnej hodnote x , ktorá odpovedá obsahu štvorca S , priraduje práve jednu hodnotu y odpovedajúcu vzdialenosti, a je reálna konštanta, $a > 1$, $a \neq 1$.

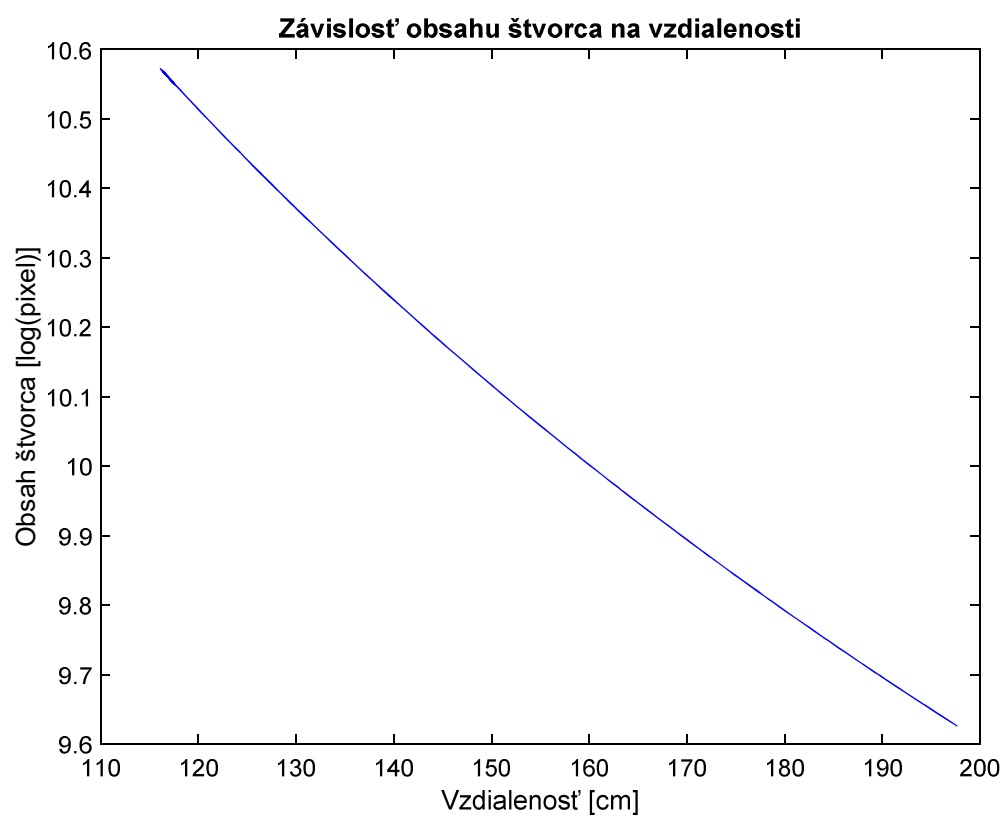


Obr. 4.1: Exponenciálna závislosť obsahu detegovaného štvorca na vzdialenosti

Vzhľadom k realizácii programu v reálnom čase a vlastnostiam programovacieho prostredia MATLAB, v ktorom je práca realizovaná, je vhodnejšie pre zjednodušenie výpočtu obsah detegovaného štvorca zlogaritmovat'. Na základe tohto kroku sa prenosová funkcia zmení na lineárnu funkciu:

$$y = ax + b, \quad (4.4)$$

kde rovnica každej reálnej hodnote x , ktorá odpovedá obsahu štvorca S , priraduje práve jednu hodnotu y odpovedajúcu vzdialenosti, a a b sú reálne konštanty (vid' Obr. 4.2).



Obr. 4.2: Lineárna závislosť obsahu detegovaného štvorca na vzdialenosti

5 FYZIKÁLNA AKUSTIKA

Akustika patrí do odboru fyziky, ktorý sa zaoberá štúdiom zvuku. Fyzikálna akustika sa zaoberá fyzikálnymi zákonitosťami šírenia zvuku ako intenzita zvuku, frekvencia rýchlosť a iné. Zvuk je možné definovať ako mechanické vlnenie, ktoré je charakterizované vibráciami pružného prostredia vo frekvenčnom rozsahu ľudského ucha od 16 Hz do 20 kHz. Toto frekvenčné rozmedzie je schopné vyvolať sluchový vnem. Vibrácie pružného prostredia vo frekvenciách pod počuteľnú hranicu 16 Hz sú označované ako infrazvuk a vibrácie nad počuteľnú hranicu ako ultrazvuk. [8][10]

5.1 Zmena intenzity hlasitosti so vzdialenosťou

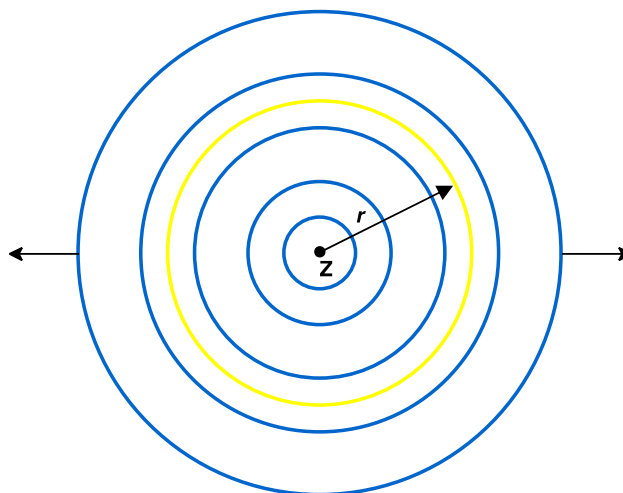
Intenzita zvuku je charakterizovaná priemernou energiou vlnenia, ktorá prejde za jednotku času jednotkovou plochou kolmou ku smeru šírenia. Tento vzťah je definovaný ako:

$$I = \frac{P}{S}, \quad (5.1)$$

kde I je intenzita zvuku [$\text{W} \cdot \text{m}^{-2}$] a P je výkon zvukovej vlny [W] dopadajúcej na plochu S [m^2]. [11]

Vzhľadom ku reálnemu zvuku je zmena intenzity hlasitosti závislej na vzdialenosti zložitým predmetom určenia. V skutočnom prostredí sa zvuk ku poslucháčovi šíri konečnou rýchlosťou od zdroja. Zvuková vlna sa šíri k poslucháčovi rýchlosťou c 340 m/s a s rovnakým oneskorením po zaniknutí zdroja vlna prichádza k zadnej časti hlavy poslucháča. Vo voľnom priestore sa zvuk šíri od zdroja priamou vlnou, ktorá postupuje po priamke od miesta, kde zvuk vzniká ku miestu jeho príjmu, t.j. sluchový orgán poslucháča. Šírenie zvuku vo voľnom priestore je možné popísať vlnoplochami. Vlnoplochy sú definované ako spojnice všetkých miest, ktoré majú rovnaké parametre v daný okamih. Napríklad zvuk z reproduktorov sa šíri len do určitého smeru, ale v uzavretom priestore umožňuje vzniku odrazených zvukových vln, pri ktorých dochádza k ozvene. V uzavretom priestore je preto nutné k priamej vlne šírenia zohľadniť aj odrazové vlny, ktoré tento priestor ohraničujú. Vlny, ktoré sa odrážajú od stien priestoru sa nešíria priamo, ale na základe ich odrazu. Odrazom sa ich cesta predĺži a dochádza k oneskoreniu. Odrazená vlna sa sčítava s priamou vlnou a tým sa mení charakter prijímaného zvuku. V priestore sa tak vytvára zvukové pole. Pre zjednodušenie situácie je možné zanedbať vplyv odrazených zvukových vln a predpokladať, že zvuk sa šíri od zdroja izotropne. Ak je zdroj zvuku bodový,

vlnoplochy majú tvar gule so stredom Z a polomerom r (viď Obr. 5.1), kde od bodového zdroja Z sa zvukové vlny šíria do všetkých smerov. [10]



Obr. 5.1: Vlnoplochy izotropného bodového zdroja [11]

Prognóza je, že celková mechanická energia vln sa pri šírení od zdroja zachováva a musí prechádzať povrchom gule. Z toho je možné vyvodiť rovnicu pre intenzitu zvuku I , ktorá sa v každom bode povrchu gule rovná:

$$I = \frac{P}{4\pi r^2}, \quad (5.2)$$

kde $4\pi r^2$ určuje veľkosť povrchu gule [m^2] a P je výkon vln [W], ktoré prechádzajú jej povrchom. Z tohto vzťahu je možné definovať, že intenzita zvuku izotropného bodového zdroja klesá so štvorcom vzdialenosti polomeru r [m] od zdroja. [11]

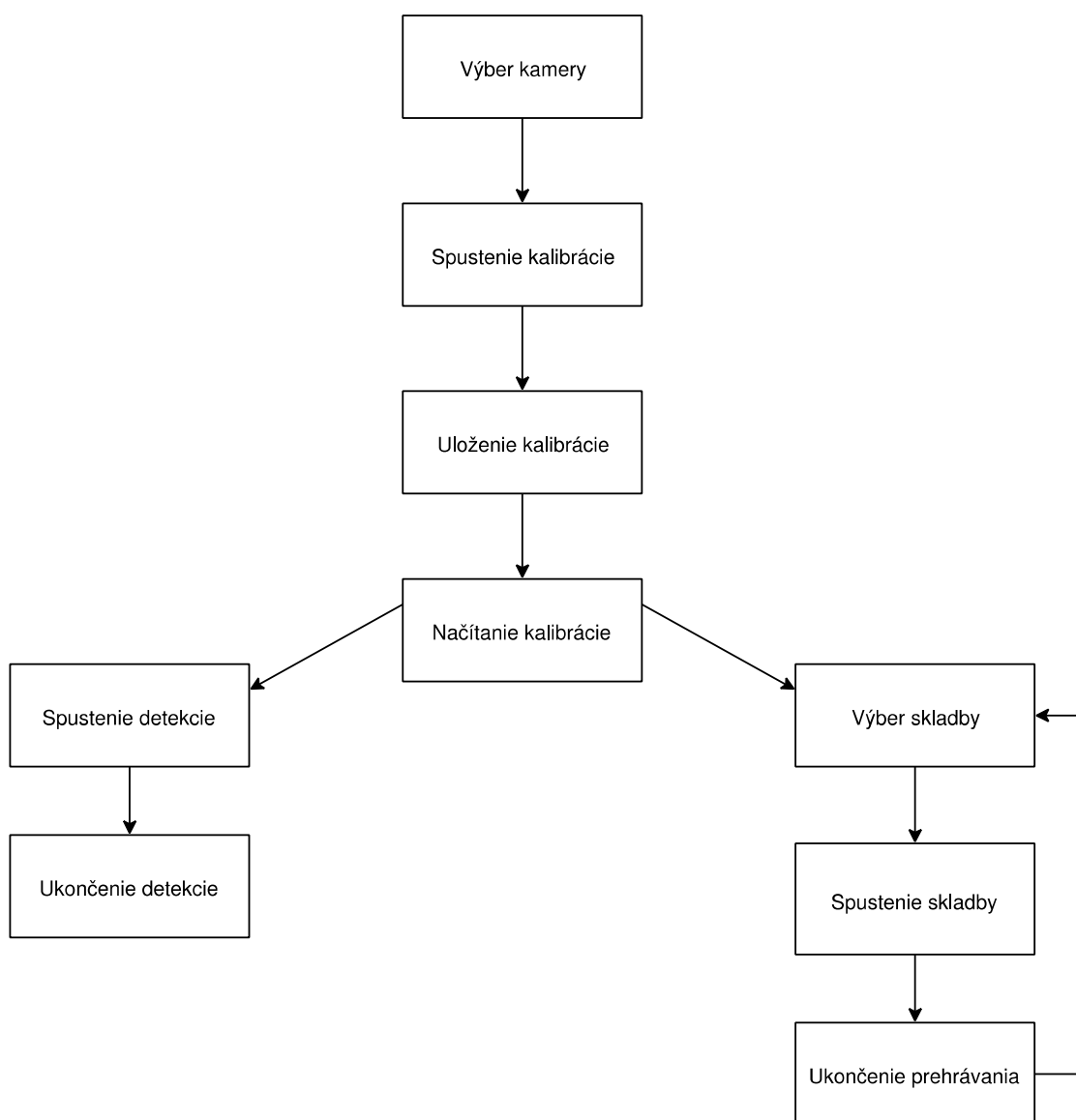
Sluch nie je rovnako citlivý pre tóny rôznych výšok. Z tohto dôvodu je sila zvuku vnímaná poslucháčom subjektívne, a tým môže byť rôzna pri dvoch zvukoch s rovnakou intenzitou, pričom subjektívna sila zvuku rastie neúmerne s jeho fyzikálnou intenzitou. V oblasti najväčšej citlivosti ucha je intenzita zvuku 10^{12} . Práve kvôli zachádzaniu s takou veľkou oblasťou hodnôt je vhodné vyjadrovať hladinu intenzity zvuku pomocou funkcie logaritmu. Ak vynásobíme intenzitu zvuku $\log 10^{12}$, hodnota sa zvýši len o 12. Na základe týchto faktov hladinu intenzity zvuku definujeme pomocou Weber-Fechnerovho zákona ako:

$$L = 10 \log \frac{I}{I_0}, \quad (5.3)$$

kde I je intenzita zvuku [$\text{W} \cdot \text{m}^{-2}$], I_0 je intenzita prahu počutia [$\text{W} \cdot \text{m}^{-2}$]. Jednotkou hlasitosti zvuku je bel, no v praxi sa častejšie využíva pre nižšie hodnoty jednotka decibel [dB]. [12]

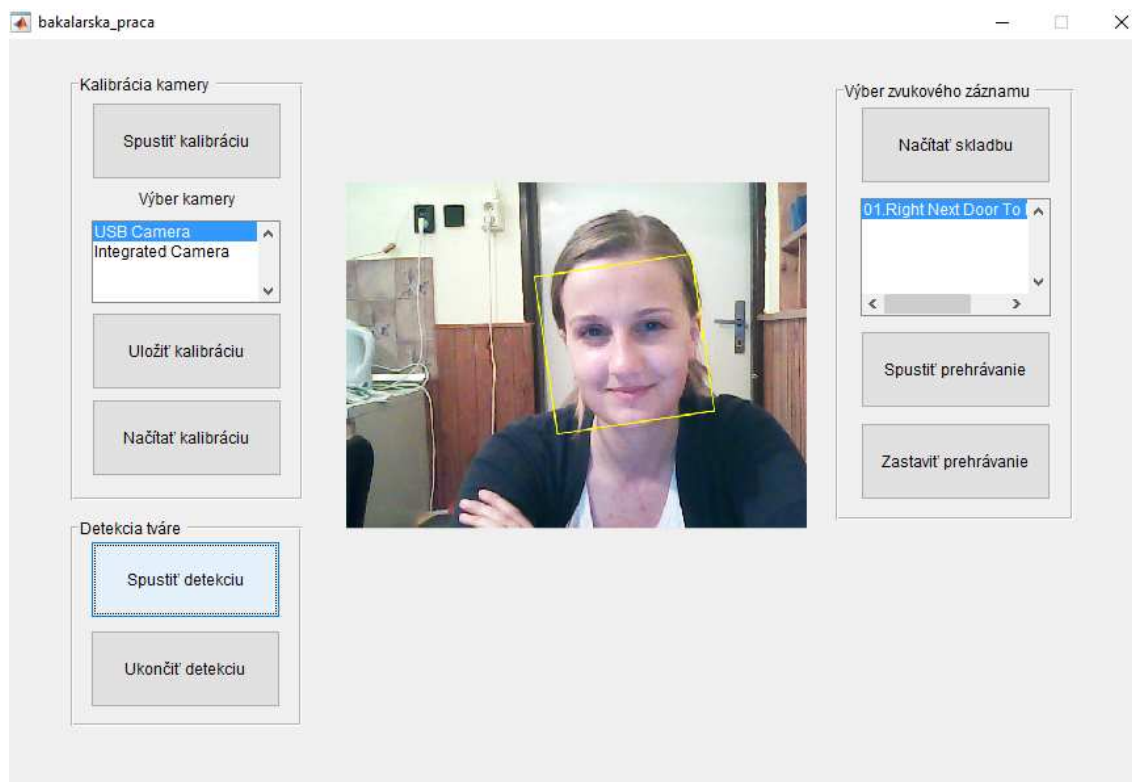
6 IMPLEMENTÁCIA PROGRAMU V PROSTREDÍ MATLAB

Program je navrhnutý v programovacom prostredí MATLAB a pre zjednodušenie obsluhy užívateľa je ovládaný grafickým užívateľským rozhraním GUI (Graphical User Interface). GUI obsahuje ovládacie prvky, ako sú panely, tlačidlá, posuvníky, ktoré intuitívne vedú k správne použitiu daného programu a eliminujú tým potrebu užívateľa učiť sa jazyk alebo spúšťanie jednotlivých príkazov. Na Obr. 6.1 je možné vidieť následnú postupnosť krokov, ktoré sú nutné pre správne fungovanie navrhnutého programu.



Obr. 6.1: Vývojový diagram navrhnutého programu

Na Obr. 6.2 je zobrazené grafické prostredie navrhnutého programu pre reguláciu hlasitosti podľa vzdialenosti poslucháča. Jednotlivé funkcie tlačidiel sú podrobnejšie popísané v nasledujúcich kapitolách 6.1 až 6.3.



Obr. 6.2: Navrhnutý program

6.1 Kalibrácia kamery

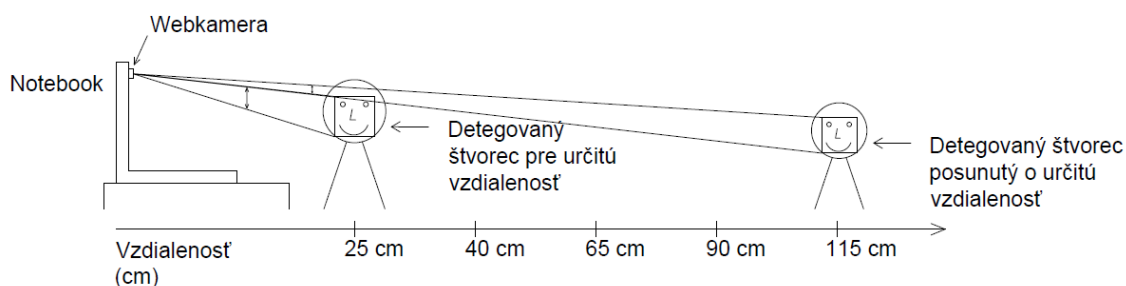
Kamery delíme na dve základné skupiny: metrické a nemetrické. Užívateľ si pred začatím kalibrácie volí jednu z ponúkaných kamier. Výber kamery závisí na type aplikácie, pre ktorú má byť použitá. Metrické kamery majú stabilnú opticko-mechanickú konštrukciu. Vnútorne parametre, ako je ohnisková vzdialenosť a stred zobrazenia, sú kalibrované v laboratóriu a predpokladá sa, že sú stále dlhú dobu. Prevažná väčšina bežne používaných kamier patrí do skupiny nemetrických, ktorých obrazy nemajú súradnicový systém. [13] Tieto kamery je nutné pred použitím kalibrovať. Kalibrácia kamery je potrebná pre zistenie jej vnútorných a vonkajších parametrov, na základe ktorých prebieha meranie vzdialenosti slúžiacej k realizácii

d'alších krokov programu. Kalibráciu kamery je možné uskutočniť pomocou dvoch metód:

- Kalibračná metóda pomocou známej vlastnosti scény
- Kalibračná metóda pomocou šachovnice

6.1.1 Kalibračná metóda pomocou známej vlastnosti scény

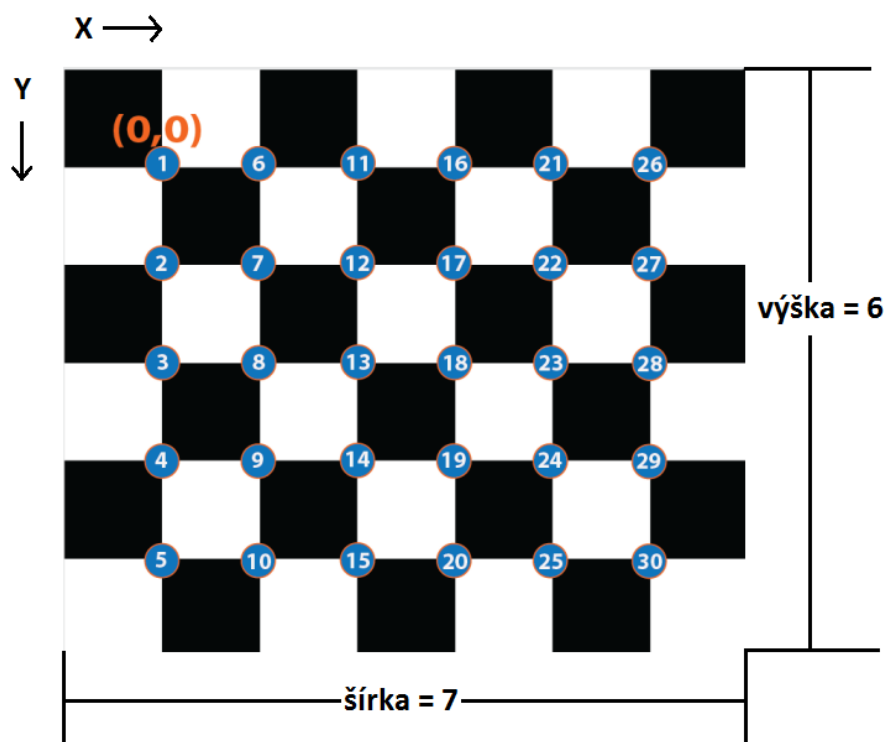
Princípom tejto kalibrácie je umiestnenie detekčných značiek, ktoré určujú presnú vzdialenosť užívateľa od kamery. Počas kalibrácie sa sníma v určených vzdialenostiach obsah detegovaného štvorca na tvári užívateľa (viď Obr. 6.3). Na základe poznanej vzdialenosti a zmeny obsahu štvorca je možné vypočítať pomocou metódy najmenších štvorcov funkciu priamky. Táto funkcia slúži v priebehu regulácie hlasitosti na prepočet vzdialenosti v reálnom čase vo videosekvencii.



Obr. 6.3: Kalibrácia kamery

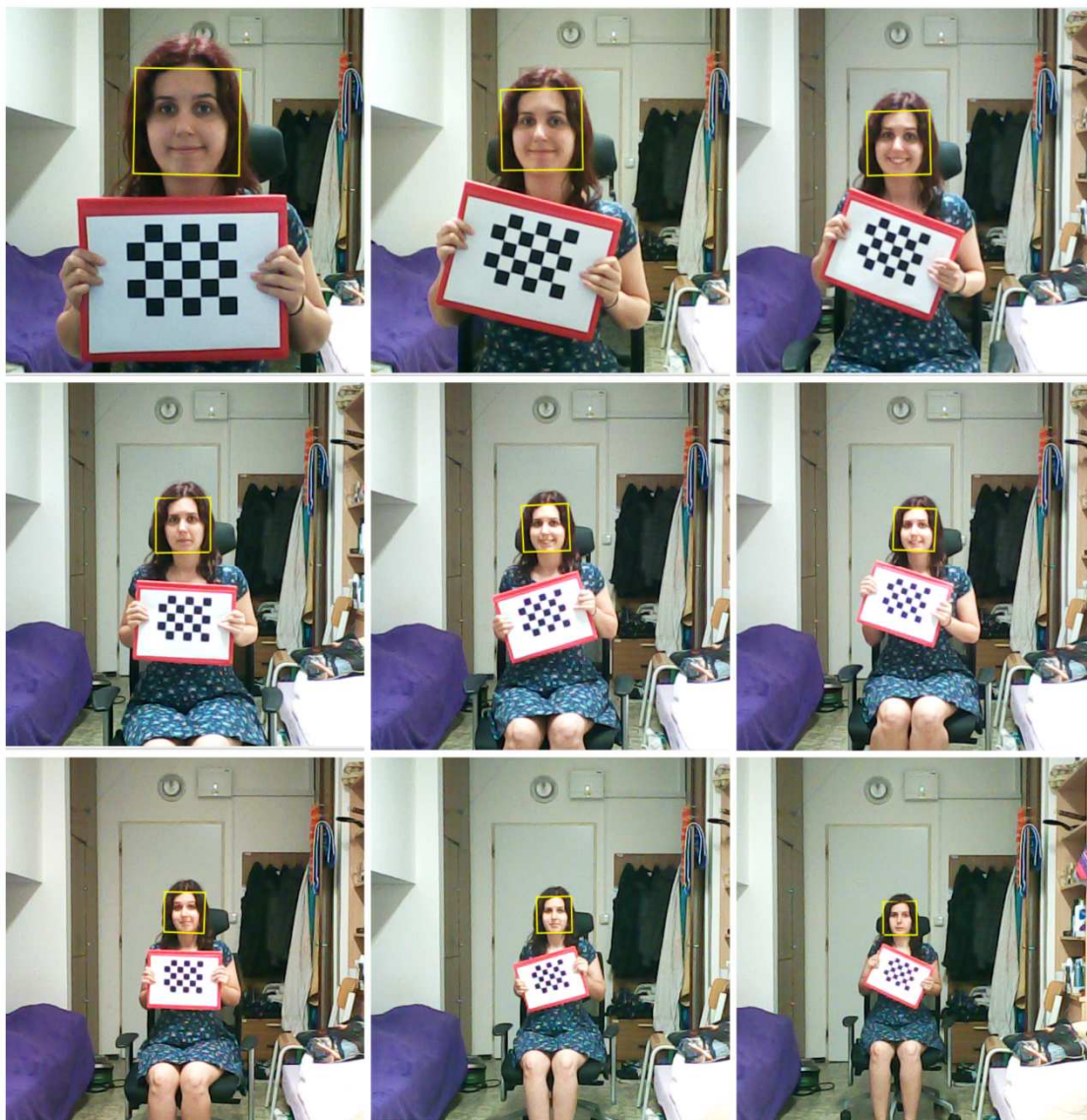
6.1.2 Kalibračná metóda pomocou šachovnice

Táto kalibrácia je realizovaná pomocou analýzy asymetrickej šachovnice na ploche, ktorá je snímaná z rôznych pohľadov. Jedna strana šachovnice obsahuje párny počet štvorcov, t.j. rovnaký počet bielych aj čiernych štvorcov a druhá strana obsahuje nepárny počet štvorcov (viď Obr. 6.4). Na šachovnici je nutné poznať skutočnú veľkosť jedného štvorca. Pre zachovanie vzťahu medzi kamerou a priestorom je vhodnejšie pohybovať vzorom v priestore pred kamerou ako samotnou kamerou. [14] Algoritmus počíta maticu pixelov nasnímaných kamerou pomocou vonkajších a vnútorných parametrov, kde vonkajšie parametre predstavujú ohniskové vzdialenosti zariadení a vnútorné parametre predstavujú vektory popisujúce rotáciu a transláciu 3D rozmerov z kamery do súradníc 2D obrázku. [15]



Obr. 6.4: Parametre šachovnice [16]

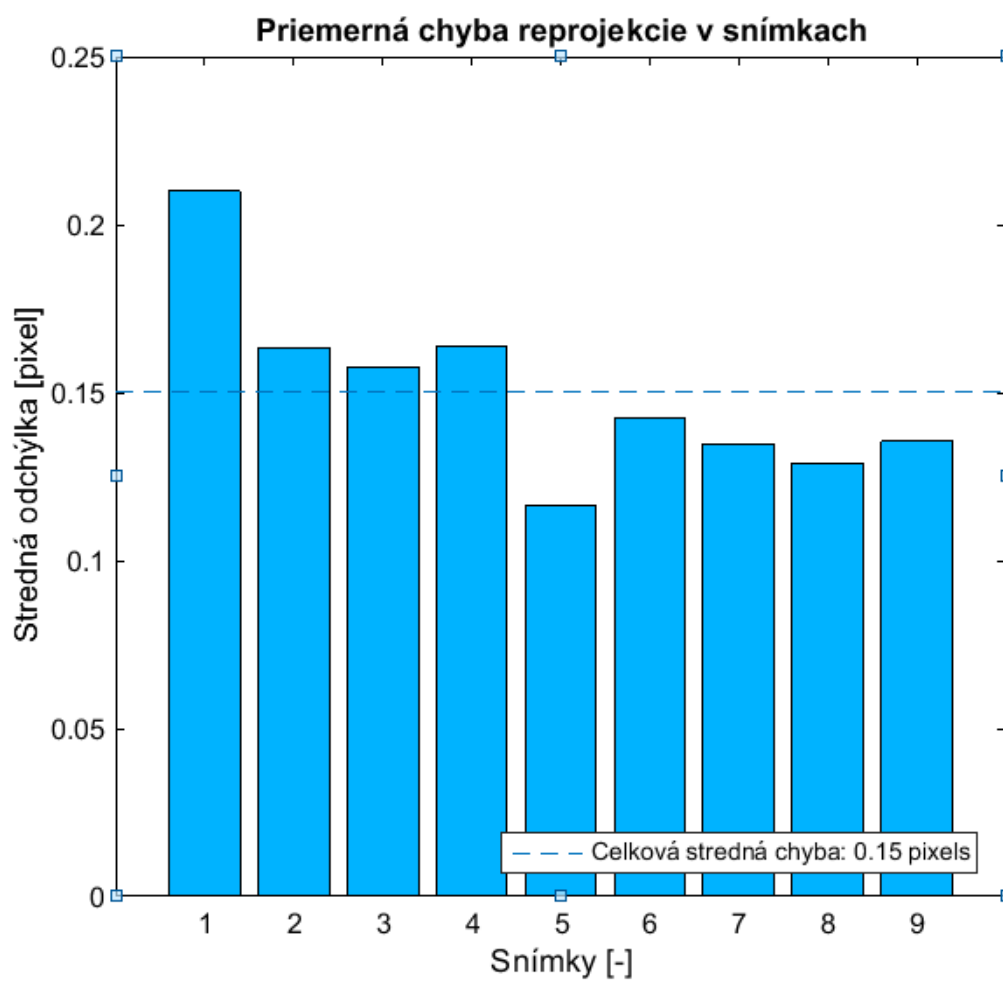
Snímky kalibrácie musia obsahovať detegovanú tvár a šachovnicu v jednej úrovni. Snímky by nemali byť zosnímané v jednej polohe, resp. v jednej a tej istej vzdialenosti. Počas kalibrácie užívateľ mierne rotuje šachovnicu do strán. Pohyb šachovnice je možné prirovnať k otáčaniu volanta pri šoférovaní. Užívateľ sa počas kalibrácie posúva o ním volenú vzdialenosť (viď Obr. 6.5).



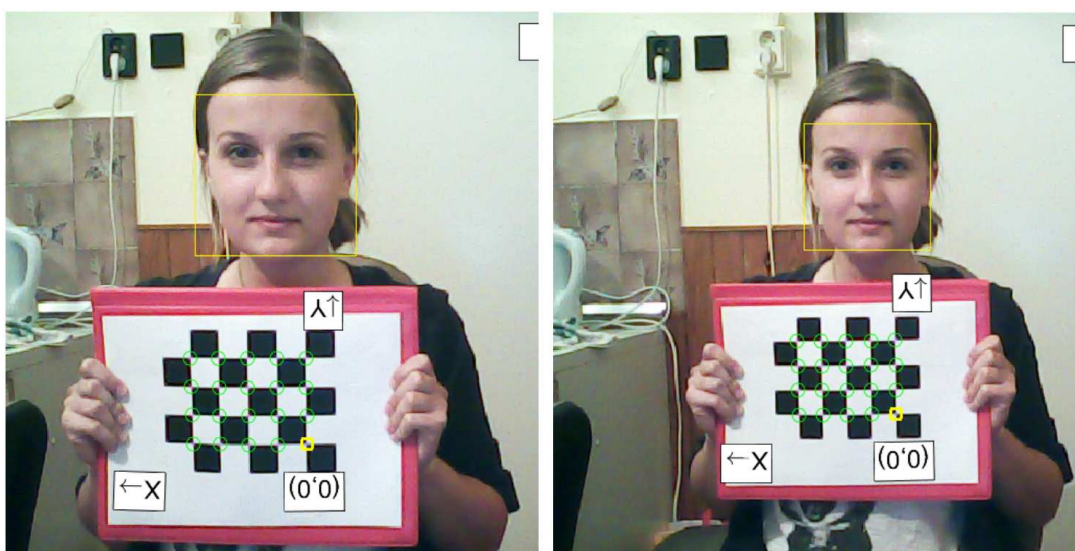
Obr. 6.5: Priebeh kalibrácie pomocou šachovnice

Touto metódou je možné kameru nakalibrovat' už dvoma snímkami, ale pre lepšie výsledky je vhodnejšie urobiť viac snímok. V nami vykonanej kalibrácii je z videosekvencie snímaných deväť snímok, z ktorých sú dosiahnuté už veľmi presné výsledky (viď Obr. 6.6), kde stĺpcový graf indikuje presnosť kalibrácie. Každý stĺpec ukazuje priemernú chybu reprojekcie pre zodpovedajúci kalibračný snímok vo videosekvencii. Chyba reprojekcie je priemerná odchýlka každého rohového bodu (ktoré sú znázornené na Obr. 6.6) od epipoláry¹ odpovedajúcich bodov v druhej snímke (viď Obr. 6.7), t.j. priemerná vzdialenosť zameraných bodov od ich skutočnej pozície.

¹ Priamka v priestore, ktorá sa zobrazuje do jedného bodu v 1. snímke zobrazená v 2. snímke.



Obr. 6.6: Priemerná chyba reprojekcie v snímkach



Obr. 6.7: Rohové body šachovnice porovnávané medzi snímkami

6.2 Porovnanie kalibračných metód

Porovnanie kalibračných metód bolo vykonané pre získanie presnosti kalibračnej metódy používajúcej šachovnicu voči kalibračnej metóde pomocou poznanej vlastnosti scény, ktorou bola vzdialenosť. Pre kalibráciu kamery bola použitá integrovaná kamera s rozlíšením 640 x 480 pixelov. Kalibrácia kamery bola v tomto prípade vykonaná na siedmich snímkach. Je to z dôvodu vplyvu osvetlenia v miestnosti, kde sa približne v troch metroch nachádzal nad užívateľom svetelný zdroj ovplyvňujúci meranie. V Tab. 6.1 sú zobrazené hodnoty vypočítanej a známej vzdialenosti pre jedného užívateľa spolu s vypočítanou odchýlkou vyjadrenou v centimetroch a percentách. Namerané hodnoty kalibrácie ďalších deviatich osôb sú v zobrazené v prílohe B. Počas kalibrácie kamery vznikala chyba zapríčinená ľudským faktorom. Meraná osoba sa v reálnom čase nikdy neposunula presne na poznanú vzdialenosť. Napriek tomu maximálna priemerná odchýlka celého procesu kalibrácie vzdialenosti voči poznanej vzdialenosti u 10 osôb bola len 1,01% (viď Tab. 6.2). Z tohto dôvodu bude efektívnejšie a pohodlnejšie pre užívateľa používať kalibračnú metódu pomocou šachovnice, kde sa počas kalibrácie môže vzdďaľovať o ľubovoľnú vzdialenosť. To kalibrácia pomocou poznanej vlastnosti scény neumožňuje. Užívateľ by si musel pred počiatkom kalibrácie buď vyznačiť detekčné značky alebo na povrázok spraviť uzlíky, ktoré by znázorňovali danú vzdialenosť. Povrázok by potom musel byť pripevnený v úrovni snímajúcej kamery a užívateľ by sa v priebehu kalibrácie posúval.

Tab. 6.1: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 1. osoby

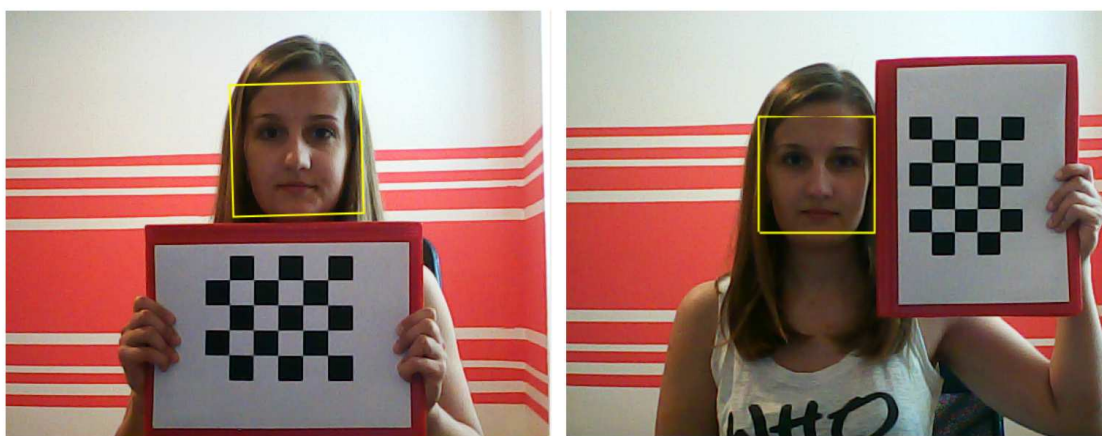
Meraná osoba č. 1			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,92	100	1,92	1,01
125,72	125	0,72	1,02
150,85	150	0,85	1,02
176,01	175	1,01	1,01
199,43	200	0,57	1,01
224,91	225	0,09	1,01
248,20	250	1,80	1,01

Tab. 6.2: Priemer jednotlivých vypočítaných vzdialeností u 10 osôb kalibračnou metódou pomocou šachovnice

Priemerné hodnoty 10 kalibrovaných osôb			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,22	100	1,15	1,01
126,23	125	1,48	1,01
151,75	150	2,03	1,01
175,75	175	1,80	1,01
200,61	200	1,91	1,01
224,42	225	1,01	1,01
250,66	250	2,26	1,01

6.2.1 Priebeh kalibrácie kamery

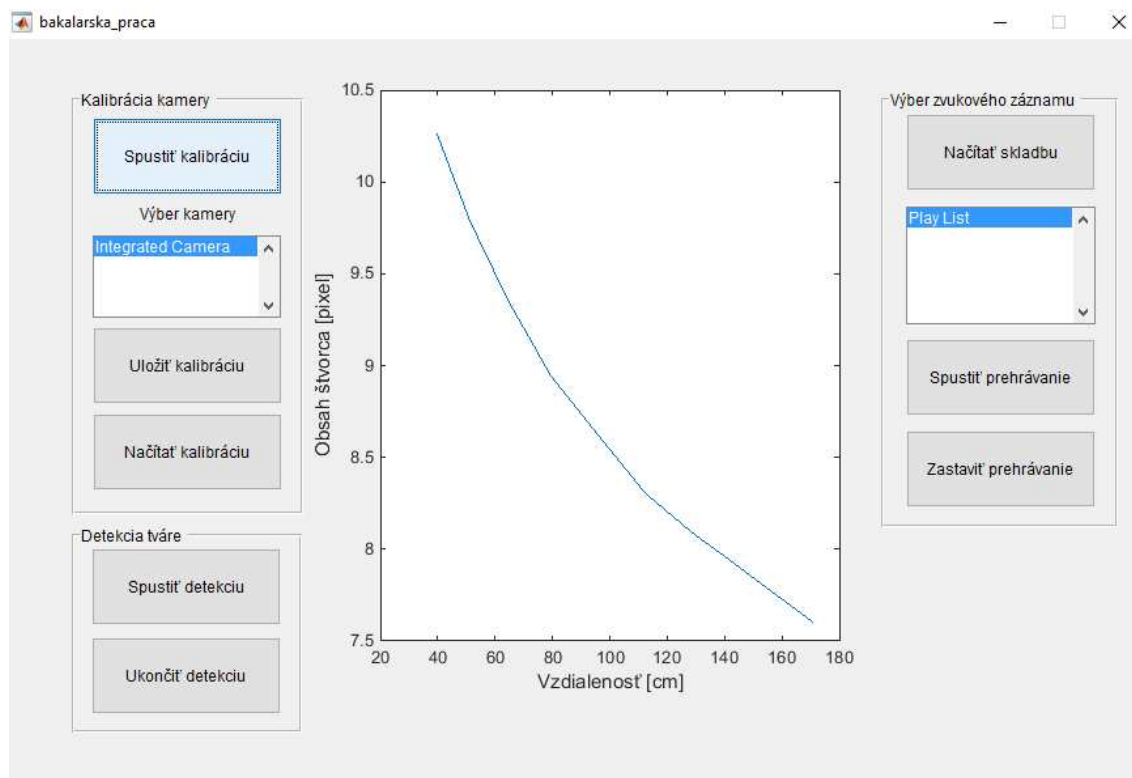
Pred kalibráciou kamery si užívateľ vyberá v paneli *Kalibrácia kamery* z kamier ponúkaných počítačom. Po vybratí kamery užívateľ vykoná kalibráciu kamery tlačidlom *Spustiť kalibráciu* (viď Obr. 6.10). Ako je vidieť na Obr. 6.8 počas kalibrácie drží detegovaná osoba v rukách šachovnicu umiestnenú pri tvári alebo pod tvárou tak, aby bola v rovnakej frontálnej rovine ako sa nachádza jeho tvár. Detegovaná tvár a šachovnica sa musia nachádzať vo videosnímke, ktorú vidí užívateľ na paneli programu.



Obr. 6.8: Umiestnenie kalibračnej šachovnice

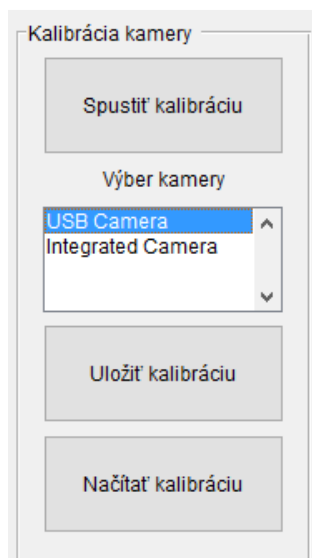
Počas kalibrácie sa užívateľ pomalým pohybom posúva smerom dozadu od kamery. Šachovnica je pri kalibrácii používaná na prepočet vzdialenosti v závislosti od obsahu detegovaného štvorca. Po ukončení kalibrácie sa zobrazí v programovom paneli graf nameraných hodnôt závislosti obsahu detegovaného štvorca na vzdialenosti

od kamery, ktorý má klesajúci charakter (viď Obr. 6.9). Kalibráciu je nutné zopakovať v prípade, ak by graf nadobúdal iné charakteristiky, alebo ak sa na paneli nezobrazí graf kalibrácie. V takom prípade bude užívateľ počuť zvukový signál, ktorý oznamuje, že počas kalibrácie nebola zosnímaná buď šachovnica alebo detegovaná tvár vo videosnímke.



Obr. 6.9: Graf závislosti obsahu detegovaného štvorca na vzdialenosti po kalibrácii kamery

Po úspešne prevedenej kalibrácii kamery užívateľ uloží kalibráciu použitím tlačidla *Uložiť kalibráciu*. Kalibrácia sa ukladá pod názov kalibrovannej kamery do súboru .mat. Tlačidlom *Načítať kalibráciu* sa kalibrácia kamery automaticky načíta do detekcie tváre. Pri opätovnom použití programu nie je potrebné, aby užívateľ vykonával znova kalibráciu kamery. Stačí ju načítať tlačidlom *Načítať kalibráciu*. Po kliknutí na toto tlačidlo sa zobrazí okno s výberom uloženej kalibrácie vo formáte .mat.



Obr. 6.10: Kalibrácia kamery

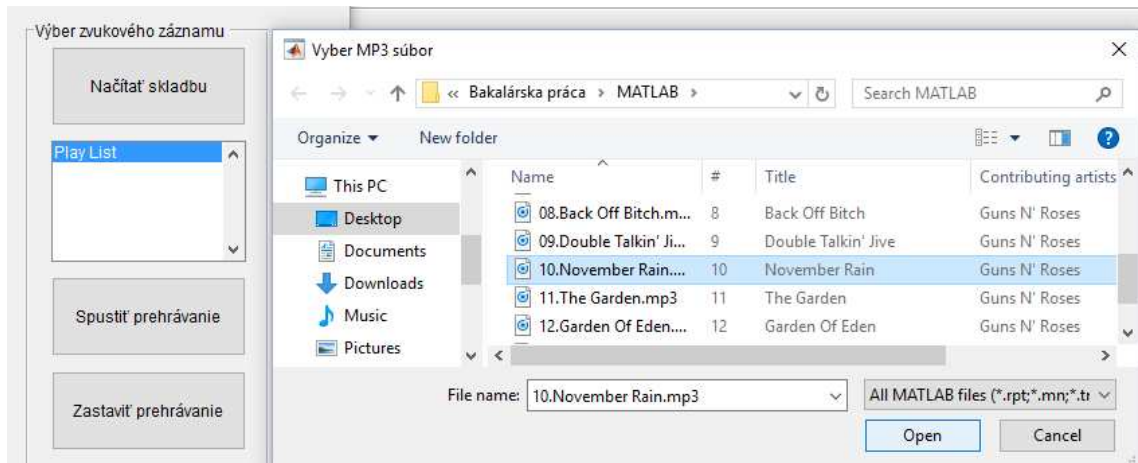
Počas kalibrácie kamery je veľmi dôležitý vplyv osvetlenia tváre, ktoré je zmienené v kapitole 3 a následne vhodné osvetlenie miestnosti. Kamera by sa nemala nachádzať pred oknom alebo vedľa okna, kde v danom momente svietia priame svetelné lúče. V takom prípade je síce osoba detegovaná, ale počas kalibrácie je tvár a šachovnica natoľko presvetlená, že dochádza k veľmi nepresnej kalibrácii (viď Obr. 6.11). Vtedy je lepšie zatiahnuť žalúzie nachádzajúce sa na okne a zapnúť umelé osvetlenie. Taktiež pri umelom osvetlení by sa počas kalibrácie užívateľ nemal nachádzať priamo pod svetlom. V tomto prípade môže vzniknúť chyba sledovania detegovanej tváre.



Obr. 6.11: Presvetlená scéna počas kalibrácie kamery

6.3 Nastavenie zvuku a jeho regulácia

Program otvorí ľubovoľný zvukový záznam vo formáte .mp3, ktorý je uložený v rovnakej zložke ako zdrojový kód programu (viď Obr. 6.12). Po kliknutí na tlačidlo *Open* sa skladba načíta v okne *Play List*. Vybraný zvukový záznam je možné prehrať kliknutím tlačidla *Spustiť prehrávanie*. Pre možnosť prerušenia prehrávania zvukovej stopy je na paneli tlačidlo *Zastaviť prehrávanie*.



Obr. 6.12: Výber zvukového záznamu

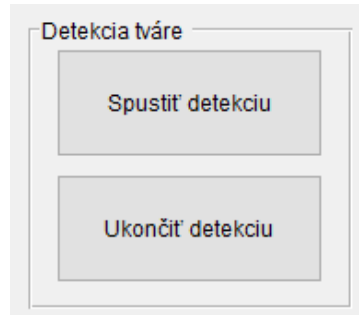
Hlasitosť zvuku je regulovaná pomocou funkcie *SoundVolume*. Pri spustení zvukového záznamu funkcia získava aktuálnu hlasitosť reproduktora. Hodnoty nastavenia hlasitosti sú numerické v rozsahu od 0.0 do 1.0. [17] Meraním pomocou hlukomeru (VOLCRAFT Sound Level Meter 322 Datalog) bola zistená hodnota decibelov v jednotlivých numerických hodnotách nastaveniach. Úbytok decibelov so vzdialenosťou bol zistený pomocou vzorca:

$$L_p = L_w + 10 \cdot \log \frac{Q}{4\pi \cdot r^2}, \quad (6.1)$$

kde L_p je hladina akustického tlaku v decibeloch [dB], r je vzdialenosť od bodového zdroja v centimetroch [cm], L_w je hladina akustického výkonu v decibeloch [dB], Q je činiteľ smerovosti, ktorý sa zavádza z dôvodu nerovnomerného šírenia zvuku do všetkých smerov od zdroja zvuku. [18]

Táto funkcia reguluje zvýšenie alebo zníženie hlasitosti zvuku z reproduktorov v rozsahu 15 dB, čo odpovedá regulácii hlasitosti do 5 metrov.

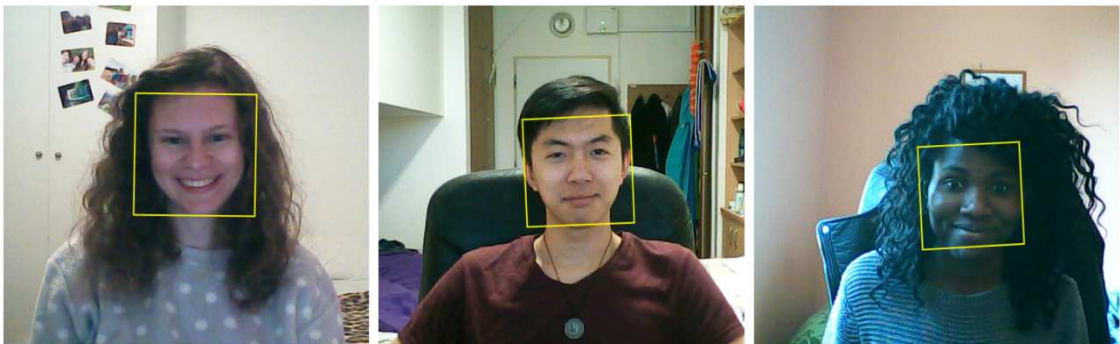
6.4 Detekcia tváre



Obr. 6.13: Detekcia tváre

Posledným krokom funkčnosti programu je spustenie detekcie pomocou tlačidla *Spustiť detekciu* (viď Obr. 6.13). Po vykonaní tohto kroku sa detegovaná osoba zobrazí pre kontrolu v užívateľskom rozhraní programu. Pre ukončenie programu slúži tlačidlo *Ukončiť detekciu*.

V jednej z predchádzajúcich kapitol bolo zmienené, že pre sledovanie tváre bola vybraná metóda KLT. Táto metóda je schopná detegovať aj rôzne rasy ako je vidieť na Obr. 6.14.



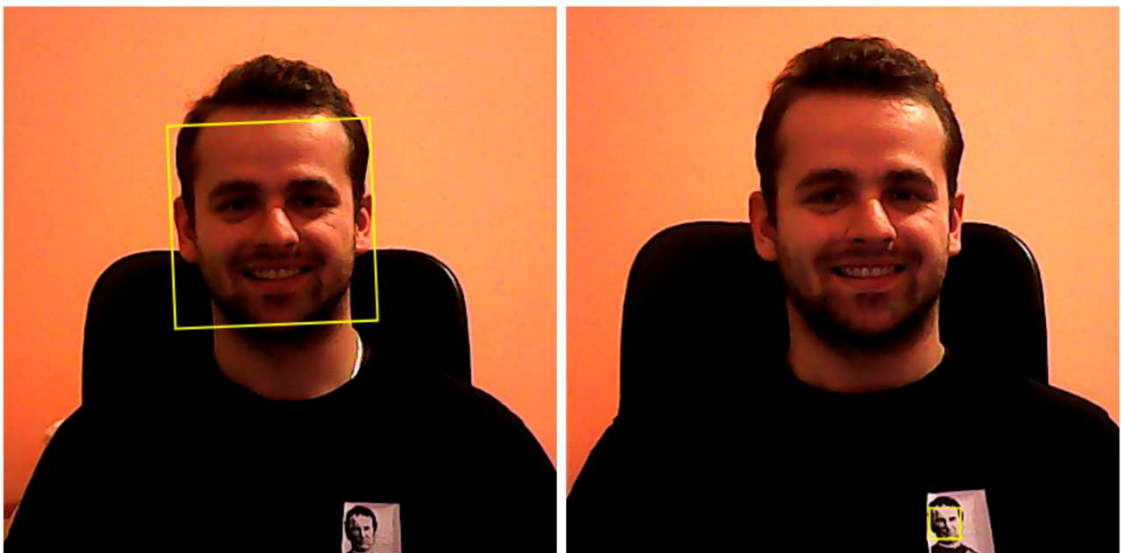
Obr. 6.14: Detegované tváre odlišnej farby

Podmienkou pre správne detegovanie tváre je, aby mala detegovaná osoba vo videosekvencii viditeľné oči a nos (viď Obr. 6.15). V prípade zahalenia nosu metóda nie je schopná tvár detegovať z dôvodu absencie jedného z Haarových príznakov, ktoré sú vysvetlené v kapitole 2.1.2. Tejto chybe je možné predísť ideálne nezahal'ovaním si tváre počas detekcie.



Obr. 6.15: Vpravo detegovaná tvár, vľavo nedetegovaná tvár

Častá chyba detekcie tváre nastávala, keď mal užívateľ oblečené tričko s podtlačou ľudskej tváre. Namiesto tváre detegovanej osoby metóda KLT rozpoznala tvár na tričku a detegovala ju (viď Obr. 6.16). V takom prípade stačí, ak užívateľ mávne jednou rukou pred kamerou, druhou rukou si zakryje podtlač a program ho opätovne nadeteguje správne.



Obr. 6.16: Vpravo detegovaná tvár užívateľa, vľavo detegovaná tvár na podtlačí trička

7 FUNKČNOSŤ PROGRAMU

Navrhnutý program bol testovaný na 13 osobách, ktorých detegované tváre sú uvedené v prílohe A. Na Obr. 7.1 je zobrazený jeden užívateľ, pri ktorom bola zmena hlasitosti zvuku zaznamenávaná pomocou hlukomeru značky VOLCRAFT Sound Level Meter 322 Datalog. Tento typ hlukomeru meria hladinu hlasitosti s chybou 1,5 dB.



Obr. 7.1: Detegovaná osoba v rôznych vzdialenostiach

Regulácia hlasitosti bola testovaná na konštantnom tóne v šiestich vzdialenostiach, ktoré sú zobrazené v Tab. 7.1. Hlasitosť zvuku sa má automaticky regulovať od počiatočnej hlasitosti. Ako počiatočná hlasitosť je v tomto prípade braná hladina zvuku zaznamenaná hlukomerom vo vzdialenosti 50 cm od kamery. Táto vzdialenosť odpovedá pozícii poslucháča pri začiatku detekcie a bola pri nej nameraná hladina hlasitosti 46,6 dB. Počiatočná hlasitosť je v tomto prípade vo vzdialenosti 50 cm od prvej detekcie, pri ktorej je hlukomerom nameraná hladina hlasitosti na 46,6 dB. Program pri vzdďaľovaní užívateľa od kamery zvyšuje hlasitosť reproduktorov, tak aby sa hodnota hlasitosti rovnala počiatočnej hladine hlasitosti, t.j. v tomto prípade 46,6 dB. Program je schopný regulovať hlasitosť s odchýlkou 0,43-1,97 % od počiatočnej hlasitosti zvuku.

Tab. 7.1: Namerané hodnoty regulácie hlasitosti v rôznych vzdialenostiach

Snímky	Vzdialenosť [cm]	Hlasitosť [dB]	Presnosť [%]
1	50	46,60	100,00
2	75	47,20	98,73
3	100	46,80	99,57
4	125	46,30	101,08
5	150	45,70	101,97
6	175	47,40	98,31

7.1 Optimálne podmienky snímania videosekvencie

Na základe analýzy vplyvu experimentálnych podmienok na robustnosť detekcie tváre a korektnosť odhadu vzdialenosti boli vytvorené nasledujúce odporúčanie pre optimálnu funkciu programu.

Priebeh kalibrácie kamery:

- Počas kalibrácie musí byť vo videosekvencii zobrazená tvár a šachovnica. Šachovnica sa musí nachádzať v rovnakej frontálnej úrovni ako detegovaná tvár. Kalibrovaná osoba sa nesmie dotýkať poľa šachovnice počas kalibrácie.
- Kalibračnú šachovnicu je vhodné podložiť, resp. pripevniť na tvrdšiu dosku, aby sa počas kalibrácie neprehýbala.
- Počas kalibrácie by kamera nemala byť orientovaná tak aby smerovala k oknu. Vplyvom priameho svetla dochádza k presvetleniu scény, ktoré má za následok nepresnosť kalibrácie. V takom prípade je vhodnejšie zatiahnuť žalúzie na okne a prípadne zapnúť umelé osvetlenie.
- Počas kalibrácie by sa kalibrovaná osoba nemala nachádzať priamo pod umelým svetlom. V tomto prípade môže vzniknúť chyba sledovania detegovanej tváre.
- V priebehu kalibrácie sa snímaná osoba pohybuje pomalým pohybom smerom od kamery a mierne rotuje kalibračnú šachovnicu do strán. Pohyb šachovnice je možné prirovnať k otáčaniu volanta pri šoférovaní.

Výber zvukového záznamu:

- Zvukový záznam musí byť vo formáte .mp3. Aby bolo možné prehrať zvukový záznam, musí sa nachádzať v zložke počítača, v ktorej sa nachádza zdrojový kód programu.

Detekcia tváre:

- Detegovaná tvár užívateľa musí byť zobrazená vo videosekvencii.
- Pri detekcii tváre je najvhodnejšie, ak užívateľ nemá zakrytú tvár žiadnym predmetom.
- V prípade zakrytia krku šatkou nesmie šatka presahovať úroveň nosa. V tomto prípade by tvár nebola detegovaná a regulácia hlasitosti by nefungovala. Je nutné, aby boli počas detekcie viditeľné aspoň oči a nos užívateľa.
- Pokiaľ má užívateľ odev, na ktorom sa nachádza podtlač ľudskej tváre, je nutné túto podtlač prekryť rukou na začiatku detekcie, aby program správne detegoval tvár užívateľa.

8 ZÁVER

Témou bakalárskej práce bola automatická regulácia hlasitosti podľa vzdialenosti poslucháča. Úlohou práce bolo oboznámenie sa s metódami detekcie tváre, ktoré sú realizované na rozdielnych parametroch. Na základe testovania jednotlivých metód v programovacom prostredí MATLAB bola vybraná najvhodnejšia metóda detekcie tváre v reálnom čase založená na detekcii a sledovaní bodov, t.j. Kanade-Lucas-Tomasi detektor.

Odhad vzdialenosti užívateľa od kamery pomocou metódy sledovania KLT bol implementovaný výpočtom obsahu detegovaného štvorca vo videosnímke, ktorý sa zväčšuje, resp. zmenšuje na základe prenosovej funkcie, ktorá bola testovaním zistená.

Pri návrhu spôsobu regulácie hlasitosti hudby podľa vzdialenosti poslucháča tak, aby poslucháč vnímal hudbu ustavične rovnako, bol pre zjednodušenie situácie zanedbaný vplyv odrazených vln v miestnosti. Ďalej bolo predpokladané, že zdroj zvuku je bodový a šíri sa izotropne. Na základe Weber-Fechnerovho zákona bolo určené, že hlasitosť, resp. intenzita zvuku bude exponenciálne klesať so štvorcom vzdialenosti r od zdroja zvuku.

Práca je realizovaná v programovacom prostredí MATLAB, v ktorom je následne grafickým užívateľským rozhraním GUI vytvorený panel pre jednoduché ovládanie programu. Pred zapnutím prvej detekcie tváre je nutné kalibrovať kameru na tvár užívateľa. Je to z dôvodu, že používaná kamera je nemetrická a jej obraz nemá súradnicový systém. Pred samotnou detekciou je užívateľom vybraný zvukový záznam. Posledným krokom k plnej funkčnosti programu je spustenie detekcie tváre. Samotný program slúži pre prehrávanie hudby, v ktorom sa hlasitosť adaptuje v reálnom čase voči vzdialenosti poslucháča od kamery.

Funkčnosť programu bola testovaná na 13 dobrovoľníkoch. Úspešnosť detekcie a sledovania tváre bola 100% u všetkých testovaných dobrovoľníkov. Pri správnej kalibrácii kamery dosahuje program až 99,87% presnosti adaptácie zvukovej hladiny. Práca zároveň rozšírila problematiku detekcie tváre v reálnom čase o ďalšie možné použitie.

LITERATÚRA

- [1] HLAVÁČ, Václav; ŠONKA, Milan. Počítačové vidění. Praha: grada, 1992. 272 s. ISBN 80-85424-67-3.
- [2] FIBAMO: Počítačové videnie. Fibamo: Elektronická bezpečnosť, Počítačové videnie [online]. 2013-02-20 [cit. 2016-01-03]. Dostupné z: <http://www.fibamo.eu/sk/pocitacove-videnie>
- [3] VIOLA, Paul; JONES, Michael J. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer society conference on computer vision and pattern recognition. 2001, (1): 511-518.
- [4] ŠIKUDOVÁ, Elena; et al. Počítačové videnie. Computer Vision, in Slovak. 397 p. Praha: Wikina, 2013.
- [5] PŘINOSIL, Jiří. Analýza emocionálních stavů na základě obrazových předloh. 2009. Ph.D Thesis. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií.
- [6] ŠOCHMAN, Jan; MATAS, Jiri. with AdaBoost. Cvičení z RPZ, 2005.
- [7] BRADSKI, Gary R. Computer vision face tracking for use in a perceptual user interface. 1998..
- [8] TOMASI, Carlo; KANADE, Takeo. Detection and tracking of point features. Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
- [9] BENEŠ, Jiří; JIRÁK, Daniel; VÍTEK, František. Základy lékařské fyziky. Charles University in Prague, Karolinum Press, 2015.
- [10] BERNAT, Petr. Akustika, vznik a šíření zvuku, frekvenční analýza a syntéza, sluchový vjem zvukového signálu.[online]. 2008.[cit. 2015-13-12]. Dostupné z WWW:< http://www.anatomie-varhan.cz/texty/varhany/anatomie/pistaly_akustika.htm.
- [11] HALLIDAY, David; RESNICK, Robert; WALKER, Jearl. Fyzika: část 1: Vlny II.: Vysokoškolská učebnice obecné fyziky. Vutium, 2000.
- [12] REICHL, Jaroslav; VŠETIČKA, Martin. Encyklopedie fyziky [online]. 2006 [cit. 2015-14-12]. Mechanické kmitání a vlnění. Dostupné z WWW:< <http://fyzika.jreichl.com/index.php>.
- [13] LUHMANN, Thomas, et al. Close Range Photogrammetry: Principles, techniques and applications. Dunbeath, Caithness KW6 6EG, Scotland, UK: Whittles Publishing, 2011. Fundamental methods, p. 510. ISBN 978-184995-057-2
- [14] ZHANG, Zhengyou. A flexible new technique for camera calibration. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000, 22.11: 1330-1334.

- [15] CAMERA CALIBRATION. Mathworks [online]. [cit. 2016-05-16]. Dostupné z: <http://www.mathworks.com/help/vision/ug/camera-calibration.html>
- [16] DETECT CHECKERBOARD POINTS: Detect checkerboard pattern in image. *Mathworks* [online]. [cit. 2016-05-10]. Dostupné z: <http://www.mathworks.com/help/vision/ref/detectcheckerboardpoints.html>
- [17] ALTMAN, Yair. SoundVolume: SoundVolume set/get the system speaker sound volume. In: Mathworks [online]. 2014 [cit. 2016-05-10]. Dostupné z: <http://www.mathworks.com/matlabcentral/fileexchange/25584-soundvolume-set-or-get-the-system-speaker-sound-volume>
- [18] VAŠINA, Martin. Environmentální fyzika: Pohlcování zvuku, šíření zvuku ve volném a uzavřeném prostoru [online]. Zlín [cit. 2016-05-16]. Dostupné z: http://ufmi.ft.utb.cz/texty/env_fyzika/EF_04.pdf

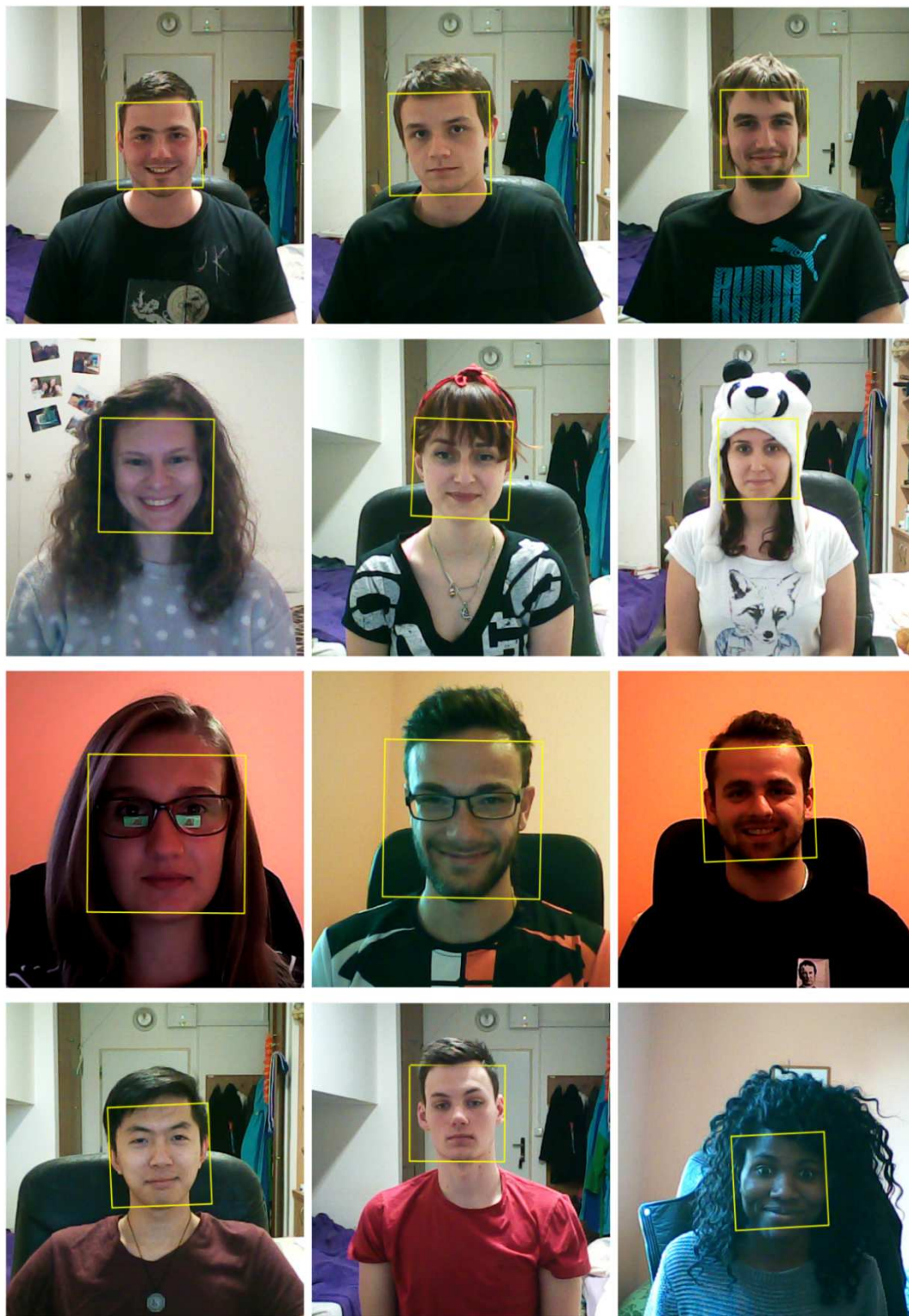
ZOZNAM SKRATIEK

HSV	Farebný model HSV (farebný odtieň, sýtosť farby, jas farby)
AdaBoost	Adaptive Boosting
KLT	Kanade-Lucas-Tomasi detektor
CAMShift	Continuously Adaptive Mean Shift algoritmus

ZOZNAM PRÍLOH

A	Detegované tváre	41
B	Kalibrácia osôb	42
C	Obsah priloženého DVD	46

A DETEGOVANÉ TVÁRE



Obr. 8.1: Detegované tváře

B KALIBRÁCIA OSÔB

Tab. B.1: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 2. osoby

Meraná osoba č. 2			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
100,81	100	0,81	1,02
127,92	125	2,92	1,01
153,62	150	3,62	1,01
177,40	175	2,40	1,01
201,32	200	1,32	1,00
222,54	225	2,46	1,00
251,58	250	1,58	1,01

Tab. B.2: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 3. osoby

Meraná osoba č. 3			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,94	100	1,94	1,02
126,21	125	1,21	1,01
152,38	150	2,38	1,02
177,79	175	2,79	1,02
202,17	200	2,17	1,01
224,83	225	0,17	1,00
253,96	250	3,96	1,02

Tab. B.3 Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 4. osoby

Meraná osoba č. 4			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,87	100	1,87	1,02
127,26	125	2,26	1,02
149,62	150	0,38	1,00
175,70	175	0,70	1,00
203,93	200	3,93	1,02
224,35	225	0,65	1,00
254,72	250	4,72	1,02

Tab. B.4: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 5. osoby

Meraná osoba č. 5			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
100,16	100	0,16	1,00
123,17	125	1,83	1,02
152,27	150	2,27	1,02
173,01	175	1,99	1,01
196,69	200	3,31	1,02
225,86	225	0,86	1,00
247,69	250	2,31	1,01

Tab. B.5: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 6. osoby

Meraná osoba č. 6			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
100,47	100	0,47	1,01
125,15	125	0,15	1,00
147,69	150	2,31	1,02
178,29	175	3,29	1,02
198,20	200	1,80	1,01
221,85	225	3,15	1,01
250,98	250	0,98	1,00

Tab. B.6: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 7. osoby

Meraná osoba č. 7			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,45	100	1,45	1,02
126,85	125	1,85	1,02
153,08	150	3,08	1,02
177,49	175	2,49	1,01
201,58	200	1,58	1,01
223,39	225	1,61	1,01
251,32	250	1,32	1,01

Tab. B.7: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 8. osoby

Meraná osoba č. 8			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
101,72	100	1,72	1,02
127,66	125	2,66	1,02
149,25	150	0,75	1,01
175,24	175	0,24	1,00
202,39	200	2,39	1,01
224,54	225	0,46	1,00
252,18	250	2,18	1,01

Tab. B.8: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 9. osoby

Meraná osoba č. 9			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
100,56	100	0,56	1,01
125,61	125	0,61	1,01
153,74	150	3,74	1,03
173,04	175	1,96	1,01
198,84	200	1,16	1,01
224,68	225	0,32	1,00
246,34	250	3,66	1,02

Tab. B.9: Porovnanie známej vzdialenosti s vypočítanou vzdialenosťou kalibračnou metódou pomocou šachovnice, meranie 10. osoby

Meraná osoba č. 10			
Vypočítaná vzdialenosť pomocou šachovnice [cm]	Známa vzdialenosť [cm]	Odchýlka [cm]	Odchýlka [%]
100,56	100	0,56	1,01
125,63	125	0,63	1,01
150,94	150	0,94	1,01
176,09	175	1,09	1,01
199,15	200	0,85	1,00
224,69	225	0,31	1,00
249,92	250	0,08	1,00

C OBSAH PRILOŽENÉHO DVD

Priložené DVD obsahuje nasledujúce položky:

- Bakalárska práca
- Zdrojový kód navrhnutého programu
- Zdrojové kódy detekcie tváre
- Namerané dáta kalibrácie kamery 10 dobrovoľníkov
- Kalibračná šachovnica