

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

ROZPOZNÁVÁNÍ GEST LIDSKÉ RUKY

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

JAKUB ADÁMEK

BRNO 2011



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ
FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

ROZPOZNÁVÁNÍ GEST LIDSKÉ RUKY

HAND GESTURE RECOGNITION

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE
AUTHOR

JAKUB ADÁMEK

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. MICHAL ŠPANĚL Ph.D.

BRNO 2011

Abstrakt

Tato práce je zaměřena na rozpoznávání gest lidské ruky. Její stěžejní částí je segmentace obrazu ve video sekvenci. Pro segmentaci obrazu je použita metoda rozpoznání obličeje v obrazu s následnou detekcí lidské kůže a metoda odečítání pozadí. Pro odstranění šumu po segmentaci obrazu jsou použity metody matematické morfologie. Práce se zabývá rozpoznáváním čistě dynamických gest. V rámci práce byla navržena aplikace pro rozpoznávání dynamických gest lidské ruky a sada modelů dynamických gest. Rozpoznávání gest je inspirováno skrytými Markovovými modely. Závěrečná kapitola je pak věnována míře přesnosti rozpoznávání gest.

Abstract

This thesis is focused on human hand gesture recognition. The main part of the work deals with image segmentation of videosequences for further gesture recognition. For the image segmentation, techniques such as face detection followed by skin detection in combination with background subtraction method are used. In order to eliminate noise, methods of mathematic morphology are applied. The work focuses only on dynamic hand gesture recognition. The proposed gesture recognition system is inspirited by hidden Markov model method. The last chapter of the thesis discusses the accuracy of the gesture recognition.

Klíčová slova

Počítačové vidění, rozpoznávání gest lidské ruky, segmentace obrazu, detekce lidské kůže, odečítání pozadí, matematická morfologie, skrytý Markovův model.

Keywords

Computer vision, hand gesture recognition, image segmentation, skin detection, background subtraction, mathematic morphology, hidden Markov model

Citace

Adámek Jakub: Rozpoznávání gest lidské ruky, bakalářská práce, Brno, FIT VUT v Brně, 2011

Rozpoznávání gest lidské ruky

Prohlášení

Prohlašuji, že jsem tuto bakalářskou práci vypracoval samostatně pod vedením Ing. Michala Španěla Ph.D. Uvedl jsem všechny literární prameny a publikace, ze kterých jsem čerpal.

.....
Jakub Adámek
18. května 2011

Poděkování

Chtěl bych poděkovat Ing. Michalu Španělovi Ph.D. za trpělivost a odborné konzultace i za ztížených podmínek. Dále bych chtěl poděkovat všem co mi pomohli s vytvořením testovacích sad a nebáli se stoupnout si před kameru.

© Jakub Adámek, 2011

Tato práce vznikla jako školní dílo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práce je chráněna autorským zákonem a její užití bez udělení oprávnění autorem je nezákonné, s výjimkou zákonem definovaných případů..

Obsah

1	Úvod.....	2
2	Ovládání pomocí gest lidské ruky.....	3
3	Zpracování a segmentace obrazu	5
3.1	Detekce lidské kůže.....	5
3.2	Detekce pohybu v obraze	7
3.3	Detekce obličeje	8
3.4	Histogram.....	9
3.5	Morfologické operátory.....	10
3.6	Skryté Markovovy Modely	13
4	Existující aplikace pro rozpoznávání gest lidské ruky	14
4.1	Nástroj HandVu.....	14
4.2	Aplikace Sonic Gesture	14
4.3	Další komerční řešení.....	15
5	Návrh aplikace pro rozpoznávání gest lidské ruky	17
5.1	Sada Gest.....	18
5.2	Metody pro zpracování obrazu.....	19
5.3	Rozpoznání gest a zobrazení výsledků.....	20
6	Implementace.....	23
6.1	Použití aplikace	24
7	Výsledné hodnocení.....	25
7.1	Testování a výsledky segmentace obrazu.....	25
7.2	Testování a výsledky rozpoznávání gest	29
8	Závěr	31

1 Úvod

Tématem této bakalářské práce je „Rozpoznávání gest lidské ruky.“ Jak již z názvu vyplývá, jedná se o práci, která úzce souvisí s počítačovým viděním reálného světa. Lidská gesta jsou nedílnou součástí života každého z nás. Většina lidí sice tato gesta vyhodnocuje pouze podprahově, tudíž si některých nemusí ani všimnout, jindy však právě neverbální komunikace může dokonce zastúnit komunikaci verbální [1]. Ale co kdybychom byli nuceni používat pouze neverbální komunikaci, tak jako to dělají neslyšící? Museli bychom vyjadřovat všechny naše myšlenky, pocity a nálady pouze gesty. Gesty našeho těla a rukou. Český znakový jazyk je přirozený a plnohodnotný komunikační systém tvořený specifickými vizuálně-pohybovými prostředky s velkým množstvím znaků a gest [2].

Cílem práce je vytvořit aplikaci, která bude schopná rozpoznávat gesta lidské ruky. Aplikace bude využívat jako vstup dat videokameru připojenou k počítači nebo video soubor uložený na pevném disku počítače. Rozpoznávání gest lidských rukou z video souborů či z kamer a následné použití těchto gest pro ovládání jiných aplikací či strojů není zatím masově rozšířeno, ale usilovně se pracuje na vývoji aplikací, které by tento stav změnily.

Cílem aplikace je vyhledání lidské postavy v dané video sekvenci, nalezení její dlaně a vyhodnocení základních navržených gest sledováním pohybu dlaní. Navržená gesta by se mohla uplatnit jako ovládací prvek pro multimediální přehrávač nebo snímkovou prezentaci.

Tato technická správa se zaměřuje na problematiku jednotlivých metod pro zpracování obrazu a rozpoznání gest lidské ruky. Problematika a rozdělení gest jsou uvedeny ve druhé kapitole. Nejdůležitější částí zpracování obrazu je segmentace obrazových dat, která je rozebrána ve třetí kapitole. Ve čtvrté kapitole jsou uvedeny již existující nástroje pro rozpoznávání gest. Na základě experimentů a poznání již existujících nástrojů je v páté kapitole uveden kompletní návrh aplikace, která je součástí práce. Tento návrh je postaven na detekci lidské kůže a skrytých Markovových modelů. Následující kapitola se zabývá implementací celé aplikace. V závěrečné kapitole jsou shrnuty výsledky rozpoznávání a experimentů, které byli pomocí aplikace provedeny.

2 Ovládání pomocí gest lidské ruky

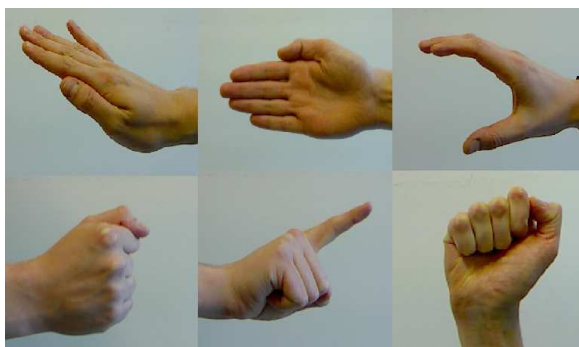
S vývojem počítačové techniky se také vyvíjí a dostávají na trh nové prostředky pro její ovládání. Na počátku byly těmito prostředky jednoduché klávesnice, které se postupně vyvinuly až do nám známého klasického rozložení, i tak se však výrobci předhánjí v pohodlnosti používání od bezdrátových zařízení až po ergonomicky tvarované ovladače, které lépe vyhovují lidské fyziologii a dopřávají tak uživateli větší pohodlí. Je toto ovšem opravdu nejpohodlnější možné ovládání, byť se sebelepším ovladačem?

Nejsnadnějšími a nejpřístupnějšími úkony jsou většinou ty, ke kterým nic nepotřebujeme. A proto by se i ovládání aplikací a elektroniky mohlo ubírat tímto směrem – směrem k minimalizaci množství používaných technických prostředků. Psaní by mohlo být nahrazeno diktováním a pohybové ovládání by nahradila právě gesta lidské ruky. Touto kombinací bychom mohli nahradit aktuální ovládání počítače, ovládáním bez užití jakýchkoliv ovládacích prvků.

Ovládání pomocí gest lidské ruky se objevuje již několik let a v poslední době se dostává na trh ve velmi úspěšných formách. Ovšem aby se takto uživatelsky jednoduché ovládání mohlo dostat do každodenní praxe, je třeba při jeho realizaci vzít v potaz několik kritérií.

Základem takového ovládání je určení sady gest, na která bude daný systém reagovat. Gesta je možno rozdělit do tří základních kategorií – na gesta statická, dynamická a kombinovaná.

Jak již z názvu vyplývá, jedná se o gesta, které nemění svůj význam a dají se zachytit v jednom časovém okamžiku (viz Obrázek 2.1). U těchto gest se nejčastěji zaměřujeme na rozložení prstů na ruce, natočení dlaně, popř. na postavení dlaně vůči tělu. Nevýhodou této kategorie je malý počet znázorněných a posléze rozpoznávaných gest, což je způsobeno nízkou flexibilitou lidské dlaně. Po rozpoznání takového gesta nedostaneme žádnou informaci o pohybu, rychlosti, případně zrychlení, nemůžeme jím tedy nahradit ovládání pomocí myši.



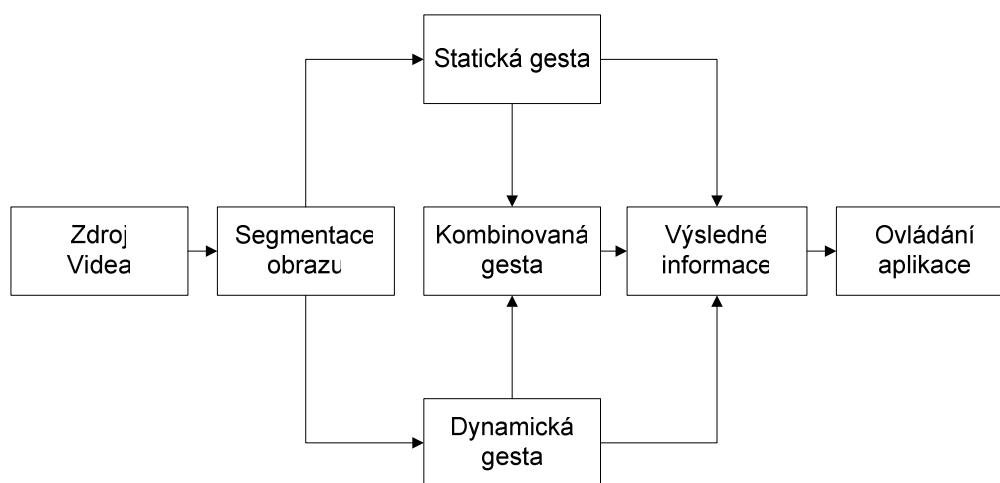
Obrázek 2.1 Sada statických gest lidské ruky

Základním principem dynamických gest je pohyb. U těchto gest se můžeme zaměřit na polohu dlaně v prostoru či pohyb dlaně vůči tělu. Danými gesty můžeme ovládat výpočetní techniku podobně jako myši, to znamená, že jimi lze určovat směr a rychlost pohybu kurzoru. Nevýhodou této sady je, že nám na rozdíl od statických gest nedokáže určit zlomové okamžiky, jako je například kliknutí myši, posun o snímek v prezentaci a tak dále. Tuto nevýhodu je však možno odstranit navržením speciální sady dynamických gest, která tyto akce zastoupí.

Sloučením těchto dvou kategorií dostaneme gesta kombinovaná. Jedná se tedy o gesta složitější, u kterých se určuje postavení prstů na ruce a jejich pohyb, natočení dlaně, postavení dlaně v prostoru či vůči tělu. Tato gesta jsou základem veškeré neverbální komunikace. Národní znakové řeči jsou

založeny na velkém množství právě takových gest, jejichž kombinací je velké množství a vyrovnají se i složitosti mluveného projevu.

Abychom vůbec mohli jednotlivá gesta rozpoznávat a následně je použít k ovládní výpočetní techniky nebo jiných zařízení, musíme nejprve správně zpracovat zdrojový obraz, kde nalezneme hledanou část ruky. Následně musíme vyhodnotit postavení prstů a natočení dlaně, pokud se jedná o gesta statická, nebo postupně vyhodnocovat rychlost a polohu ruky u gest dynamických. Pokud bychom chtěli rozpoznávat kombinovaná gesta, museli bychom použít oba tyto postupy současně. Takto získané informace bychom dále porovnávali s předem určenými matematickými modely. Na základě rozpoznání jednotlivých modelů ve video sekvenci bychom mohli ovládat dané aplikace nebo celé systémy pomocí gest lidské ruky. (Obrázek 2.2).



Obrázek 2.2 Ovládání aplikace pomocí gest lidské ruky

Pro zpracování zdrojového obrazu, se používají metody segmentace obrazových dat. Metody pro segmentaci obrazu detekují v obraze objekty podobných vlastností. V obraze můžeme detekovat například pohyb, který je nejčastěji detekovaný pomocí metody odečítání pozadí [18]. Dále je možné detekovat objekty s podobností obrazových informací, jako je například detekce lidské kůže [11]. Nebo je možné detekovat objekty s podobnými rysy jak barevnými tak tvarovými například detekce obličeje [24] a [26]. Pro rozpoznávání gest jak statických tak dynamických se nejčastěji používá metoda skrytých Markovových modelů, i když se s touto metodou častěji setkáváme u rozpoznávání mluvené řeči [22]. Pokud bychom se zaměřili na systémy, které nevyužívají pouze klasické zpracování dat obrazu, ale i jiné metody vylepšující výsledné rozpoznání gest jistě stojí za zmínku systém Kinect od společnosti Microsoft [10].

3 Zpracování a segmentace obrazu

Segmentace obrazu je jedním z nejdůležitějších kroků vedoucích k analýze obsahu zpracovávaných obrazových dat. Snahou segmentace obrazu je rozčlenit obraz do částí, které mají úzkou souvislost s předměty či oblastmi reálného světa zachyceného na obraze. Výsledkem segmentace má být soubor vzájemně se nepřekrývajících oblastí. Pokud tyto oblasti jednoznačně korespondují s objekty vstupního obrazu, jedná se o segmentaci kompletní. Vytvořené segmenty však nemusí přímo souhlasit s objekty obrazu; pak jde o segmentaci částečnou [4].

Obrazovou segmentací se zabývá mnoho aplikací; jak komerčních, tak volně šiřitelných projektů. Tyto aplikace se vesměs snaží získat z obrazových dat informace o objektech, které mají stejné či podobné vlastnosti. Může se jednat o různé vlastnosti nebo prvky, které jsou v obraze vyhledány. Nejčastěji se jedná o detekci pohybujících se objektů v obraze, vyhledávání částí lidského těla, detekci základních geometrických útvarů či vyhledávání předem známých a specifických objektů nebo značek. Pomocí segmentace se hledané objekty buďto zvýrazní, nebo se naopak odstraní objekty neodpovídající daným kritériím. Následně můžeme zjišťovat další informace o hledaných objektech, jako polohu jejich středu, směr či vektor pohybu, velikost, umístění v daném obraze.

V našem případě se zajímáme o segmentaci obrazu, která přímo souvisí s rozpoznáváním gest lidské ruky. Většina aplikací pro detekci objektů v obraze využívá OpenCV, což je volně šiřitelná knihovna pro práci s obrazem. Tuto knihovnu využívá i většina známých aplikací dostupných na trhu či na internetu.

3.1 Detekce lidské kůže

Při rozpoznávání gest lidské ruky je základním stavebním kamenem nalezení a vyhodnocení objektů, které odpovídají vlastnostem lidské kůže. K detekci lidské kůže můžeme přistupovat různými způsoby, jako jsou například barevné modely lidské kůže, datové sady lidské kůže či vytvoření histogramu předem známé oblasti, u které si můžeme být jisti, že se jedná o lidskou kůži. V tomto případě se může jednat o kalibraci aplikace nebo její nastavení na základě známých poznatků z reality. Nejpoužívanější metoda detekce lidské kůže je založena na základě matematických vzorců pro její barvu a odstín. Složitější verze jsou buď kombinací různých metod s jejich postupným filtrováním, nebo se jedná o metody rozsáhlejší, využívající různých známých aspektů.

Detekce lidské kůže je vhodná metoda pro segmentaci obrazových dat, ve kterých je jasně vidět odkrytá část lidského těla. Základním prvkem pro zobrazování obrazových dat na počítači je adaptivní míchání tří základních barev. Tento model se nazývá RGB a obsahuje informace o mohutnosti červené, zelené a modré. Pro detekci barvy lidské kůže není tento model zcela vyhovující, jelikož RGB nese informaci nejen o barvě, ale také o její zářivosti. Informace o zářivosti lidské kůže jsou v tomto případě nežádoucí z důvodu jejích měnících se vlastností na základě osvětlení, opálení a také rasy člověka [11]. Tuto složku můžeme odstranit pomocí převodu trojrozměrného modelu RGB do chromatického barevného prostoru. Převod do chromatických (tzv. čistých) barev, neobsahujících informace o zářivosti barvy lidské kůže, je definován převodem do dvou normalizovaných složek (1) a (2).

$$r = \frac{R}{(R + G + B)} \quad (1)$$

$$g = \frac{G}{(R + G + B)} \quad (2)$$

R, G, B jsou zde informace o barevném prostoru RGB, kde R zastupuje barvu červenou, G zelenou a B modrou. Výsledkem jsou normalizované složky r a g dvourozměrného euklidovského prostoru v chromatických barvách. Normalizovaná modrá barva m je zbytečná, jelikož platí (3).

$$r + g + b = 1 \quad (3)$$

I když se nám zdá barva kůže mnohých lidí velmi odlišná, největší rozdíly jsou v jasu, ne v barvě. Zmíněnou problematikou se již zabývalo mnoho projektů. V [12] je popsána podobnost barvy kůže v chromatickém barevném prostoru lidí různých ras. V tomto článku je zanesena podobnost barvy do grafu (viz Obrázek 3.1). Toto rozložení je možné charakterizovat vícerozměrným Gaussovým rozložením pravděpodobnosti.

$$N(m, \Sigma) \quad (4)$$

kde m je střední hodnota, která se určí podle (5) a \bar{r} , \bar{g} se určí podle (6) a (7). Σ je kovariační matice, kterou určíme pomocí rovnice (8) [13]. Pomocí těchto hodnot jsme schopni určit rozložení pravděpodobnosti.

$$m = (\bar{r}, \bar{g}) \quad (5)$$

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i \quad (6)$$

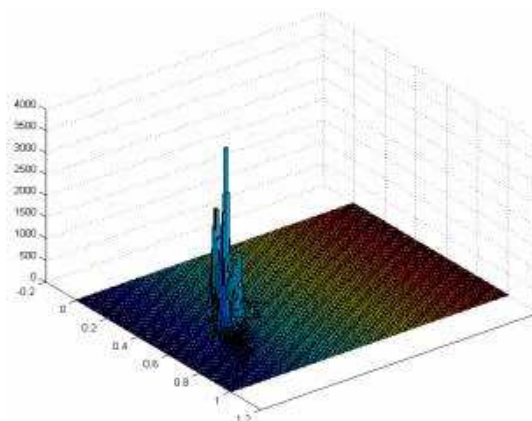
$$\bar{g} = \frac{1}{N} \sum_{i=1}^N g_i \quad (7)$$

$$\Sigma = \begin{pmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{pmatrix} = \frac{1}{N} \sum_{i=1}^N (x_i - m)(x_i - m)^T \quad (8)$$

$P(x)$ nám určuje pravděpodobnost jevu, že pixel x (9) má právě barvu kůže. $P(x)$ určíme pomocí vzorce rovnice (10).

$$x = (r, g) \quad (9)$$

$$P(x) = \exp\left(-0,5(x_i - m)^T \Sigma^{-1}(x_i - m)\right) \quad (10)$$

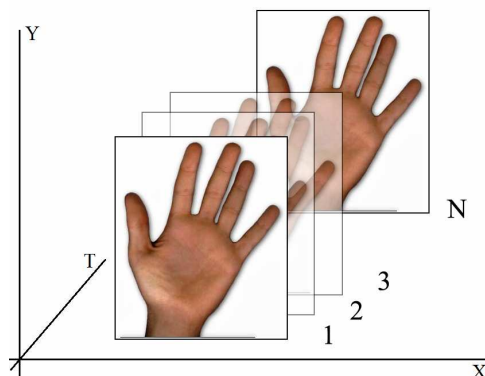


Obrázek 3.1 Barevné spektrum lidské pokožky [12]

3.2 Detekce pohybu v obraze

Detekce pohybu v obraze je další důležitá část potřebná ke správné segmentaci obrazu. Detekce pohybu nám segmentuje obraz na objekty, které se v dané video sekvenci pohybují, jinak řečeno, že daný objekt mění svou polohu v obraze. Pokud v obraze vyhledáváme gesta lidských rukou, která nejsou statická, tato metoda se osvědčuje jako velmi užitečná. Nejčastěji se ovšem s detekcí pohybu setkáme v bezpečnostních aplikacích, které kontrolují pohyby lidí ve vymezeném prostoru.

Aby bylo možné detekovat pohyb v obraze, je nutné použít dynamický obraz, což je množina po sobě jdoucích statických snímků se známým časovým intervalem rozestupu (Obrázek 3.2). Pro analýzu pohybu jsou také kromě obrazových dat velmi důležité i další informace, jako například pohyb kamery – je-li statická nebo pohybující se. Podle těchto informací můžeme zvolit vhodnou metodu pro analýzu pohybu v obraze. Metod, které se dají použít pro analýzu pohybu, je celá řada. Mají různá využití a různé náročnosti na početní výkon [18].



Obrázek 3.2 Dynamický obraz

Základní metodou pro detekci obrazu je rozdíllová metoda. Tato metoda nám sice neurčí směr, rychlost ani zrychlení objektu, dokáže však správně segmentovat obraz a zvýraznit tak pohybující se objekty v reálném času. Zpracování obrazu touto metodou není náročné na početní výkon.

Rozdíllová metoda pracuje na principu odečtení dvou po sobě následujících obrázků f_1 a f_2 (Obrázek 3.3) a vznikne tak rozdíllový obraz d , který můžeme určit pomocí (11).

$$d = \begin{cases} 0 & \text{pro } |f_1(x, y) - f_2(x, y)| < e \\ 1 & \text{pro } |f_1(x, y) - f_2(x, y)| \geq e \end{cases} \quad (11)$$

kde e je kladné číslo které určuje práh rozdílu, a x, y jsou souřadnice analyzovaného bodu. Hodnota 0 odpovídá místům v obrazech f_1 a f_2 , ve kterých nedošlo k výrazným změnám v úrovních jasu.

Na podobném principu funguje metoda kumulativního rozdíllového obrazu. Tato metoda je velmi podobná s tím rozdílem, že první obrázek v řadě se označí jako referenční a následující obrázky se od něj odečítají. Výsledný kumulativní obraz je pak dán součtem všech rozdíllových snímků (12). Tato metoda se liší od té klasické i faktem, že zde lze získat informace o směru pohybu a rychlosti [19].

$$d_a(x, y) = \sum_{i=1}^N a_i d_i \quad (12)$$



Obrázek 3.3 Ukázka rozdíllové metody [28]

3.3 Detekce obličeje

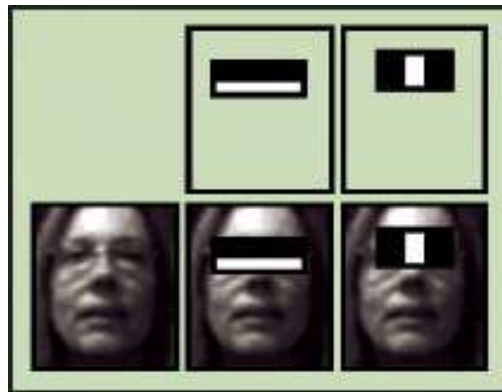
Metod pro rozpoznávání a detekci obličeje se v poslední době objevuje stále více, a tato oblast výzkumu se stále rozrůstá, neboť nachází stále větší uplatnění v různých oborech jako kriminalistika, identifikace, snímací techniky, biometrické zabezpečení a další. Pro detekci obličeje existuje celá řada postupů a způsobů, z nichž každý je použitelný v jiné oblasti či jiných podmínkách, podle vhodnosti. Metody pro detekci obličeje pracují na mnoha principech od detekce barvy lidské kůže, hledání specifických bodů obličeje po metody strojového učení a porovnávání databází. O různých typech metod pro detekci obličeje dále v [24].

Metoda použitá v knihovně OpenCV se nazývá AdaBoost. Jedná se o nejznámější variantu metody boosting, což je zlepšení klasifikační přesnosti algoritmů strojového učení [26]. Tato metoda se používá pro detekci libovolných objektů v obraze. Jde o sekvenční trénování modelů. Vezme jednoduchý (Obrázek 3.4) a slabý klasifikátor a tyto klasifikátory spojí, aby celek dosáhl požadované přesnosti [25]. Tuto kombinaci můžeme znázornit pomocí rovnice (13).

$$h(x_i) \in \{-1, 1\}$$

$$H(x) = \text{sign} \sum_{i=1}^N \alpha_i h(x_i) \quad (13)$$

Kde x_i jsou jednotlivá data, $h(x_i)$ jsou jednoduché klasifikátory a α je váha, kterou potřebujeme nastavit trénováním.



Obrázek 3.4 Ukázka metody AdaBoost z [26]

3.4 Histogram

Histogram je graf četností, který zobrazuje rozložení barevných kanálů červené, zelené, modré (RGB), stupnice šedi a dalších veličin, jako jsou například sytost, světlost a odstín barvy v obraze. Histogram se používá k analýze obrazových dat před jejich úpravou. Může poskytnout důležité informace o daných obrazových datech.

Histogram můžeme také chápat jako vektor (14) absolutních četností výskytu každé barvy v obraze [16].

$$[H(1), H(2), \dots, H(c)] \quad (14)$$

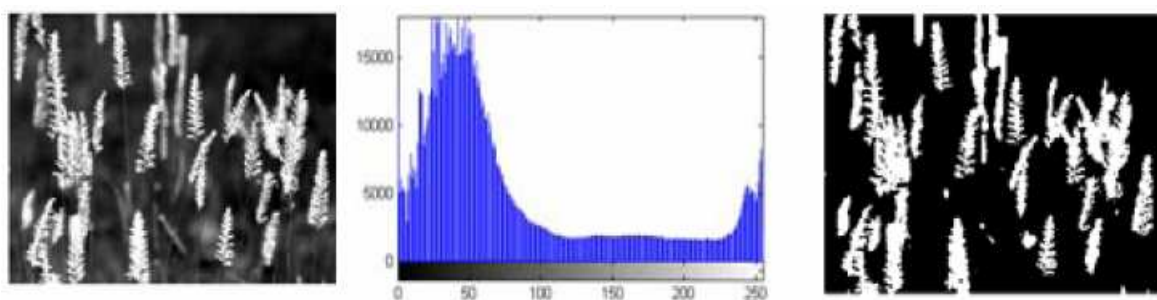
$$c \leq m.n \quad (15)$$

$$\sum_{i=1}^c H(i) = m.n \quad (16)$$

Kde c je počet barev a m, n značí velikost neboli rozlišení daného obrázku. Jak správně rozdělit obrazová data a vypočítat hodnoty histogramu s danou pravděpodobností řeší [17]. Vytvořený histogram k daným obrazovým datům (Obrázek 3.5) je velmi důležitý pro další segmentaci obrazu. Z histogramu můžeme určit potřebný práh (17), který nám obraz dokáže segmentovat na námi hledané objekty.

$$y = \begin{cases} L & \text{pro } x < T \\ H & \text{pro } x \geq T \end{cases} \quad (17)$$

Kde x je vstupní hodnota intenzity, y je výstupní hodnota intenzity a T je hodnota hledaného prahu.



Obrázek 3.5 Vytvoření histogramu a následné zpětná projekce [16]

3.5 Morfologické operátory

Při zpracování obrazových dat pomocí výpočetní techniky se velmi často setkáváme s šumem. Tento jev může nastat z několika důvodů, jako například kvůli nekonstantnosti pozadí, splývání barev či jasů při segmentaci obrazu, a v mnoha jiných případech. Abychom tomuto většinou nežádoucímu jevu zabránili, můžeme použít metody matematické morfologie. Pojmeme matematická morfologie rozumíme nástroj pro extrakci hledaných částí rozpoznávaného obrazu, který je založen na nelineárních matematických operacích. Tyto nástroje jsou využívány k předzpracování obrazu a při finálních úpravách po segmentaci obrazu. Mezi nejzákladnější operace matematické morfologie patří dilatace a eroze, které se také nejčastěji používají při zpracování obrazu, ale můžeme se také setkat s jinými operacemi, jako jsou například otevření, uzavření, transformace tref či miň a jiné [20].

Dilatace

Morfologická operace dilatace skládá body dvou množin pomocí vektorového součtu. Tento, tzv. Mikowského množinový součet, je znázorněn rovnicí (18). Vstupem této operace je obraz X , který má být transformován, a strukturní element B [21].

$$X \oplus B = \bigcup_{b \in B} X_b \quad (18)$$

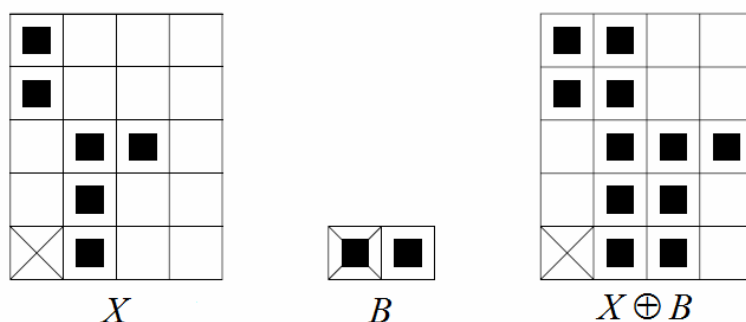
$$X \oplus B = \{ p \in E^2 : p = x + b, x \in X \text{ a } b \in B \}$$

Typicky je tato metoda použita s binárními obrazy, ale můžeme se setkat i s šedotónovou variantou. Pomocí dilatace můžeme objekty zvětšovat a zaplňovat malé díry či úzké štěrby vzniklé v obraze po segmentaci (Obrázek 3.6). Zjednodušeně můžeme tuto operaci chápat jako porovnání vstupního obrazu se strukturálním elementem, který je obrazem posunován.

$$X = \{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4)\} \quad (19)$$

$$B = \{(0,0), (1,0)\} \quad (20)$$

$$X \oplus B = \{(1,0), (1,1), (1,2), (2,2), (0,3), (0,4), \\ (2,0), (2,1), (2,2), (3,2), (1,3), (1,4)\} \quad (21)$$



Obrázek 3.6 Příklad matematické dilatace

Eroze

Morfologická operace eroze skládá dvě množiny využitím rozdílu dvou vektorů neboli Minkowského rozdílu. Jde o duální morfologickou transformaci k dilaci nikoli inverzní [21]. Pro výpočet eroze můžeme použít následující rovnici (22).

$$X \ominus B = \bigcap_{b \in B} X_{-b} \quad (22)$$

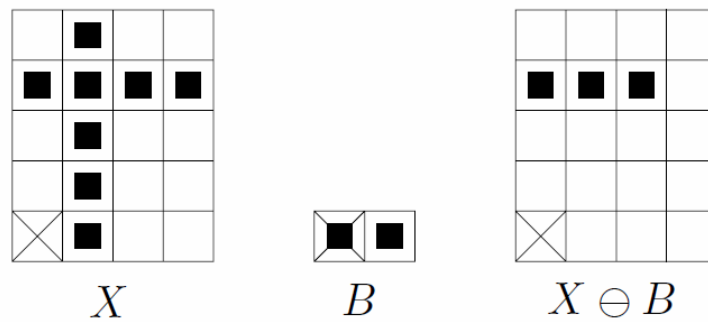
$$X \ominus B = \{p \in E^2 : p = x + b \in X \text{ pro každé } b \in B\}$$

Pomocí eroze můžeme docílit ořezání hranic objektu, čímž se objekt zmenší. Malé objekty nebo šum menší než strukturní element touto operací zanikají. Eroze se často používá na rozložení složitějších objektů na jednodušší pro následné zpracování; viz příklad (Obrázek 3.7).

$$X = \{(1,0), (1,1), (1,2), (0,3), (1,3), (2,3), (3,3), (1,4)\} \quad (23)$$

$$B = \{(0,0), (1,0)\} \quad (24)$$

$$X \ominus B = \{(0,3), (1,3), (2,3)\} \quad (25)$$



Obrázek 3.7 Příklad matematické eroze

Otevření a uzavření

Tyto morfologické operace jsou kombinacemi základních morfologických operací dilatace a eroze [21]. Operace otevření je složením eroze a následující dilatace (21) a operace uzavření je dilatace následovaná erozí (22).

$$X \circ B = (X \ominus B) \oplus B \quad (26)$$

$$X \bullet B = (X \oplus B) \ominus B \quad (27)$$

3.6 Skryté Markovovy Modely

S pojmem skryté Markovovy modely se setkáváme nejčastěji ve spojitosti s rozpoznáváním mluveného jazyka. Avšak v poslední době, s rozvojem počítačového vidění, se můžeme s tímto pojmem setkat jako s nástrojem pro rozpoznávání jednotlivých obrazových dat. Ovšem nejsou to jediná odvětví, ve kterých se s touto metodou můžeme setkat. Nachází uplatnění například v medicíně jako analyzátor EKG a EEG signálů.

Markovovské modely jsou zvláštním případem stochastických konečných automatů, dovolující vyjádřit statické závislosti dané pořadím pozorování jako například v časových řadách. [30]. Značením posloupností podle rovnice (28), můžeme vyjádřit pozorování (29) a skryté stavy (30).

$$(x_a, x_{a+1}, \dots, x_b) = x_a^b \quad (28)$$

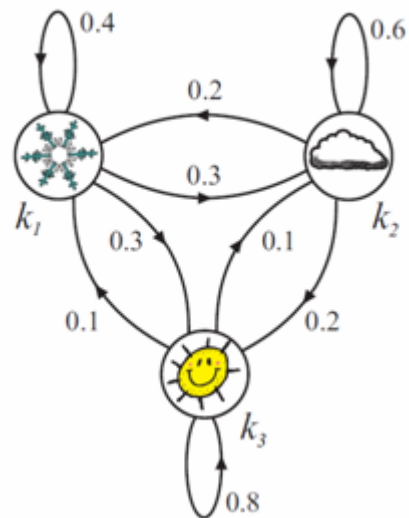
$$x = x_1^n \quad (29)$$

$$k = k_0^n \quad (30)$$

Příklad skrytého Markovova modelu si můžeme ukázat na jednoduchém modelu počasí. [30]. Máme tři stavy k , odpovídající možnému aktuálnímu počasí. Mezi jednotlivými stavy jsou hodnoty pravděpodobnosti přechodu z jednoho stavu do druhého. Tyto informace můžeme zapsat do přechodové matice stavů A (Obrázek 3.8).

k_1 – dešťové nebo sněhové srážky
 k_2 – zataženo
 k_3 – slunečno

$$A = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0.2 & 0.6 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$$



Obrázek 3.8 Příklad Markovových modelů [30]

Stejným způsobem jako je znázorněno v příkladu, můžeme klasifikovat rozpoznávání gesta, které je natrénováno. Mezi jednotlivými stavy gesta jsou hodnoty pravděpodobností přechodů z jednoho stavu do druhého. Při rozpoznávání je cílem určit s jakou pravděpodobností pozorování odpovídá natrénovaným modelům. Další využití a podrobné informace o zpracování skrytých Markovových modelů nalezneme v [30] nebo [22]. Případně na webových stránkách věnující se této problematice [23].

4 Existující aplikace pro rozpoznávání gest lidské ruky

Aplikací, které se zabývají rozpoznáváním gest lidské ruky, bychom nejspíš našli bezpočet. Proto si uvedeme pouze ty, které by mohly být budoucím přínosem pro tento obor. Aplikace se různě zaměřují na různé sady gest, které se mají rozpoznávat. Jedná se jak o gesta statická, tak o gesta dynamická. V podstatě záleží na tom, jakou aplikaci nebo systém chceme pomocí gest ovládat. Gesta statická nejsou příliš vhodná na ovládání aplikací, u kterých vyžaduje ovládání znalost rychlosti či zrychlení, takové aplikace rozpoznávají gesta dynamická. S tímto rozpoznáváním se můžeme setkat u herních systémů jako je Nintendo Wii nebo Xbox 360. Oproti tomu gesta statická jsou rozpoznávána pro ovládání aplikací jako HandVu nebo Sonic Gesture.

4.1 Nástroj HandVu

Tento nástroj se zabývá rozpoznáním gest lidské ruky a následným vytvořením uživatelského rozhraní pro jednotlivé aplikace. Vstupním zařízením této aplikace je velké množství podporovaných kamer. Aplikace využívá knihovnu OpenCV 2.1.0, což také odpovídá velkému výběru podporovaných kamer.

Nástroj HandVu se může pyšnit širokou škálou použitelnosti, jako například ovládání multimediálních aplikací a jiných programů, u kterých by tento nástroj mohl být přijatelným vylepšením. V neposlední řadě je možno použít knihovny HandVu ve vlastní aplikaci. Aplikace je postavena na principu klient-server, kde klient detekuje a trackuje pohyb ruky a server je databází rozpoznávaných gest. Rozpoznávaná gesta jsou především statická. Sada gest je uložena na serverové části aplikace, jedná se o data určující rozložení prstů na dlani a natočení dlaně. Po následné detekci statického gesta může uživatel začít trackovat pohyb dlaně. Tato kombinace je velmi flexibilní a můžeme s ní ovládat velké množství aplikací. Podle autorů je aplikace dostatečně přesná a proto není nutná kalibrace při spuštění [5].

Aplikace detekuje ruce podle modelu kůže, a to na jednotném pozadí. Navíc je pro inicializaci programu nutné, aby se ruka objevila ve video sekvenci v přesně dané poloze, která se nesmí vychýlit o více než 15 stupňů. Tuhle část bych považoval za kalibraci programu. HandVu je též velmi citlivé na osvětlení či změnu pozadí. Nástroj HandVu rozpoznává šest gest uvedených v dokumentaci programu [6].

4.2 Aplikace Sonic Gesture

Aplikace je založena na knihovně OpenCV a uživatelském rozhraní QT. Rozpoznává gesta lidské ruky a na základě přiřazeného příkazu k jednotlivým gestům generuje zvuk. Tato aplikace umožňuje vytváření hudebních skladeb bez kontaktu s počítačem. Aplikace je schopna učít se nová gesta, ale už bohatě implicitní sada, která je značně obsáhlá, by měla postačit k vytvoření jednoduché hudební skladby. Aplikaci je možné propojit i se softwarovým mixážním pultem a tím vytvořit různé hudební styly bez kontaktu s počítačem [7].

Vstupem pro aplikaci může být video soubor nebo kamera. Aplikace vyhledává v obrazu gesta rukou, která jsou natrénována a uložena v datovém souboru. Natrénovaná gesta jsou čistě statická. Sonic Gesture se vůbec nezabývá pohybem jednotlivého gesta. Vyhodnocování je jednoduché a velmi rychlé. Aplikace v každém obrázku videa rozpoznává jednotlivá gesta z datového souboru. Aplikace nevyužívá žádný barevný model lidské kůže, pouze pomocí OpenCV vyhledá obličej a z obličeje je zhotoven histogram. Ten je následně použit pro vytvoření modelu kůže. Tím aplikace získá přesný vzor lidské kůže. Program následně vrací tři detekované části, a to obličej a dvě ruce, ze kterých jsou následně rozpoznávána natrénovaná gesta. Sonic Gesture je velmi přesný a vyhledává gesta i na poměrně velkou vzdálenost (Obrázek 4.1), za různých světelných a jiných nepříznivých podmínek. Tato metoda segmentace obrazu mě zaujala a částečně inspirovala.



Obrázek 4.1. Ukázka Sonic Gesture [7]

4.3 Další komerční řešení

V dnešní době se na trhu objevuje čím dál více řešení ovládání aplikací, spotřební elektroniky a jiných zařízení pomocí gest lidské ruky. Tato představa, před pár lety ještě futuristická, se dnes stává každodenním standardem. Tyto aplikace mohou být využívány v širokém spektru zaměření, např. v lékařství, kde je možné používat tyto aplikace pro rehabilitaci [8], dále pro ovládání spotřební elektroniky v domácnosti (televizorů, mobilních telefonů a samozřejmě počítačů), i pro ovládání celých operačních systémů. Ovšem největší podíl na vývoji tohoto uživatelského rozhraní má odvětví počítačových her.

Průlomem v této sféře byla konzole Nintendo Wii (Obrázek 4.2) v listopadu roku 2006, která se na rozdíl od největších konkurentů zaměřila na ovládání prostřednictvím pohybů těla a nikoli jen na ovládání pomocí mačkání tlačítek na ovladači. Její ovladače jsou schopné snímat vlastní polohu, natočení i akceleraci pohybu. Detekce probíhá ve dvou úrovních: akcelerometr uvnitř ovladače detekuje polohu ovladače a jeho změnu. Míření ovladače směrem k obrazovce pak zajišťuje panel s infračervenými diodami, který se instaluje například pod televizi, a snímač v přední části dálkového ovladače. Panel obsahuje celkem deset diod, vždy po pěti na každém konci. Vnitřní diody jsou

orientovány mírně do středu, vnější naopak ven. Snímač pak změří vzdálenost obrazu na své ploše, současně zná vzdálenost diod na panelu. Díky tomu může procesor vypočítat vzdálenost ovladače od plochy a především změnu jeho polohy. Nicméně prudké změny polohy, například při hře box, již měří akcelerometr.

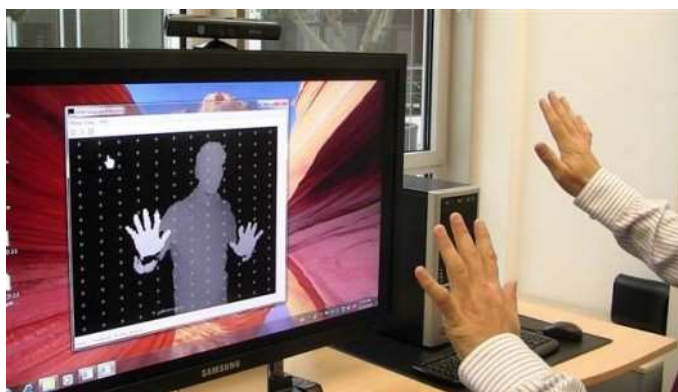


Obrázek 4.2 Herní konzole Nintendo Wii s nutností použití ovladače [29]

Následně se na trhu objevily další ovladače, například připojitelný nunchuk nebo volant, do kterého se základní ovladač jednoduše vloží. I když konzole Wii stále nutí hráče držet v ruce ovladač, stala se velmi populární a propagovanou. Revoluce by se ale nekonala bez skvěle propracovaných her, které v plné míře ovládání pohybu využívají [9].

Skutečné ovládání pouze a jen pohybem těla přináší až Microsoft v podobě systému Kinect for Xbox 360 (Obrázek 4.3). O detekci pohybu se stará malá krabička připojitelná ke všem stávajícím herním konzolím Xbox 360. Uvnitř této 23 centimetrů široké tyčinky se nachází infračervený projektor a černobílá snímací CMOS kamera. Právě práce v infračerveném spektru umožní činnost v různých světelných podmínkách. Systém si při snímání vytvoří kostru hráče s celkem 48 body a jeho pohyb snímá s frekvencí 30 snímků za vteřinu [10]. Toto zařízení si vydobýlo u uživatelů velkou popularitu a díky tomu se stalo nejrychleji prodávaným elektronickým zařízením vůbec.

Kinect se stal velmi populární díky přesnosti segmentace obrazu, rozpoznávání gest a také obličejů. Nezaostává ani v možnosti vytvoření osobních profilů s biometrickým zabezpečením. Ovšem většina her vyvíjených pro Kinect a konzoly Xbox 360 je založena na detekci čistě dynamických gest.

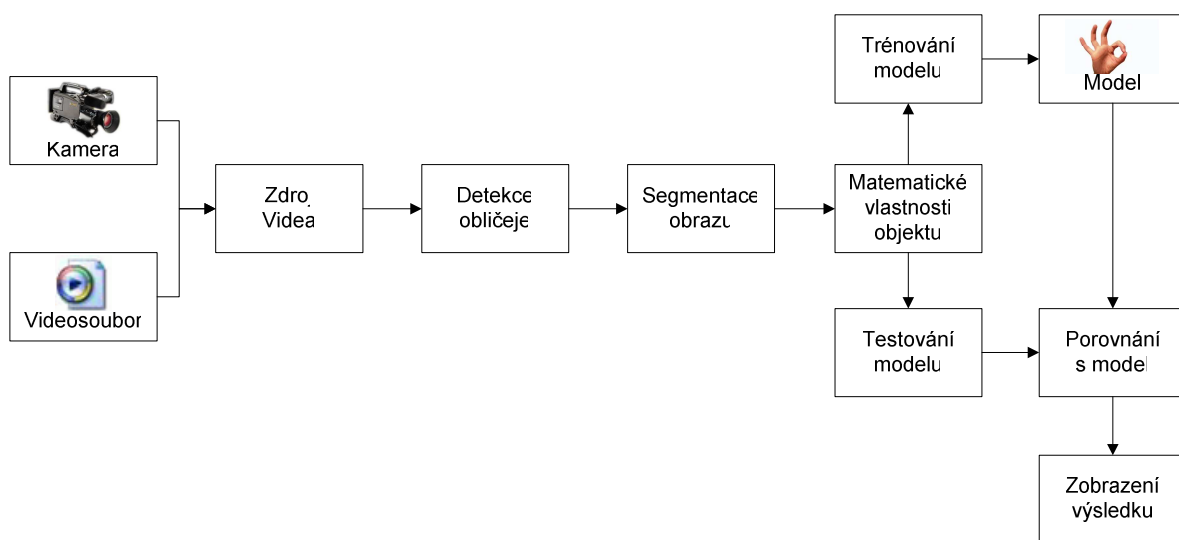


Obrázek 4.3 Použití systému Kinect pro Xbox 360 [27]

5 Návrh aplikace pro rozpoznávání gest lidské ruky

Aby bylo možné vytvořit aplikaci pro rozpoznávání gest lidské ruky, je nutné uvědomit si několik základních principů jak postupovat. Nejdůležitější je uvědomit si cíl práce a gesta, které chceme rozpoznávat. Po nastudování různých materiálů a vyzkoušení několika aplikací pro rozpoznávání gest lidské ruky, jsem se rozhodl, že budu rozpoznávat gesta čistě dynamická. Nebudu sledovat pouze jejich pohyb a směr, ale budu se snažit natrénovat sadu složitějších dynamických gest, které budu následně rozpoznávat v obraze.

Asi jako první krok je nutné určit zdroj dat. Za zdroj dat je považována video sekvence ze souboru nebo z připojené kamery k počítači. Na zdrojovém videu je spuštěna detekce lidského obličeje pro detekování postavy a spuštění následné segmentace obrazu, která je prováděna pro každý snímek zvlášť. Po správné segmentaci obrazu, je nutné určit matematické vlastnosti hledaného objektu, v našem případě dlaní. Po získání těchto informací můžeme model natrénovat nebo pokud natrénováno již máme můžeme gesta porovnávat s matematickým modelem. Rozpoznávání gest je inspirováno teorií skrytých Markovových modelů. Pokud bude gesto rozpoznáno do zdrojového videa bude přidán text o detekci a informacích rozpoznaného gesta (Obrázek 5.1).

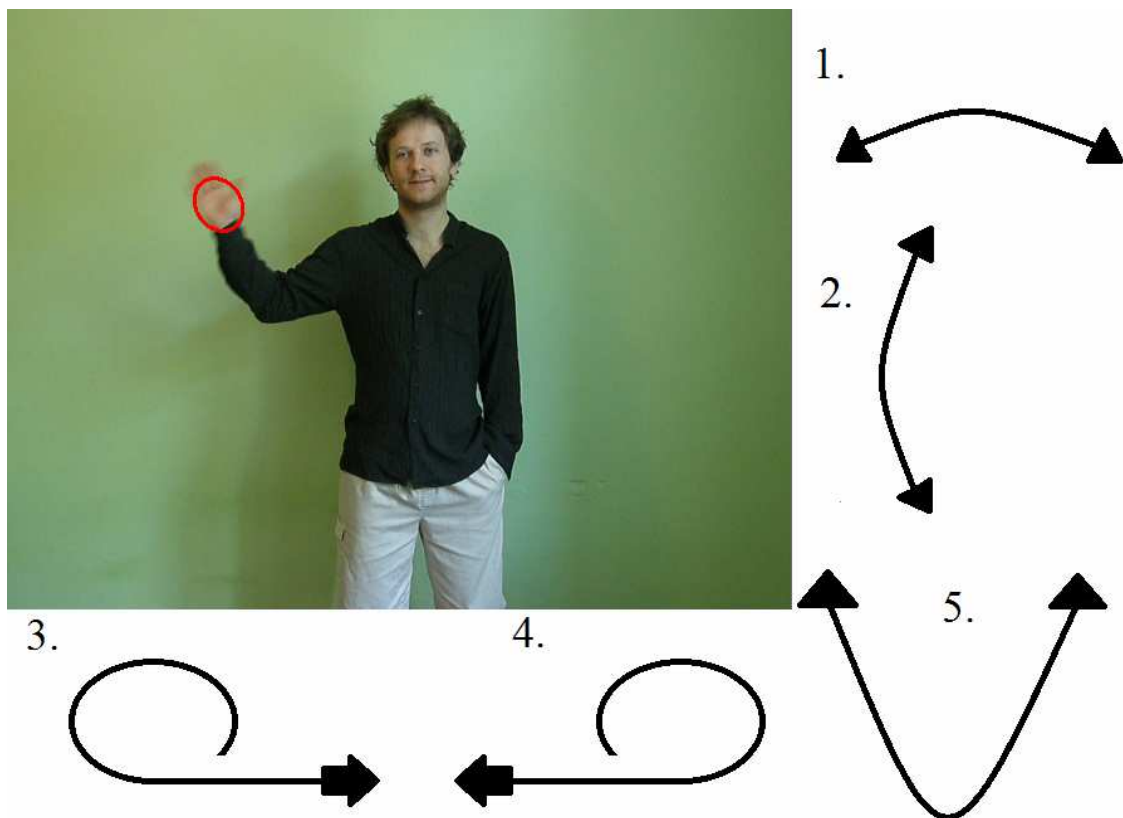


Obrázek 5.1 Návrh běhu aplikace

5.1 Sada Gest

Aby celá aplikace správně fungovala, bylo nutné vybrat a navrhnout vhodnou sadu gest, které by se dali rozpoznávat. Jelikož byly vybrány gesta dynamická, u kterých se nebude hodnotit jen rychlost a směr pohybu, ale i následnost těchto pohybů, je nutné tyto gesta správně zanést do modelu. Jelikož gesta budou rozpoznávána pomocí algoritmu inspirovaným skrytými Markovovými modely, je nutné vytvořit pro každé gesto jeden model, se kterým se bude nadále pracovat. Jelikož se jedná o dynamická gesta, je nutné aby model obsahoval informace o pohybu případně rychlosti. Zdrojové video je rozděleno na několik menších částí. Přejdem mezi jednotlivými částmi se dostáváme do jednotlivých stavů modelu (Obrázek 5.2). Tyto stavy následně uložíme tak, jak jdou za sebou a získáme model jednotlivého gesta, který uložíme do datového souboru.

Pro aplikaci je navrženo pět gest, které budou namodelovány a následně rozpoznávány ve video sekvencích. Jednotlivá gesta jsou uvedena na Obrázku 5.3, kde šipky označují pohyb dlaně. Jednotlivé modely jsou pojmenovány následovně: 1. Horizont, 2. Vertical, 3. Left, 4. Right, 5. Konvex.

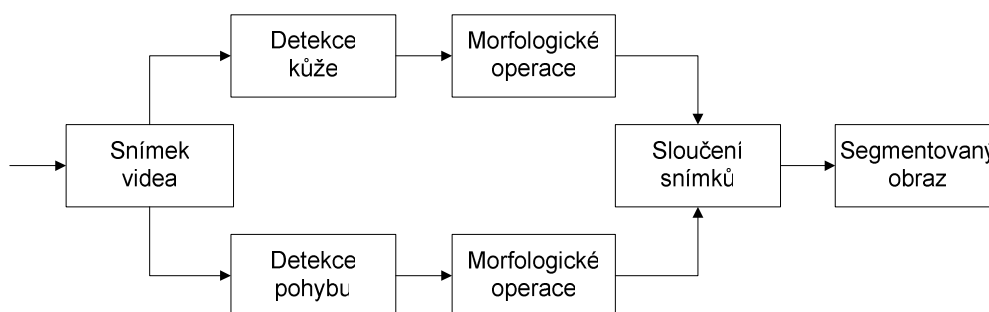


Obrázek 5.3 Ukázka modelovaných gest

5.2 Metody pro zpracování obrazu

Abychom mohli rozpoznávat výše uvedená gesta je nejdůležitějším krok správně segmentovat obraz. V Našem případě se jedná o detekci dlaní, je tedy nutné si uvědomit jaké vlastnosti hledaný objekt má. Jistě můžeme určit, že dlaně mají barvu kůže. A jelikož se jedná o detekci dynamických gest, můžeme taky určit, že hledaný objekt bude v pohybu. Pro úplnou segmentaci obrazu, ve kterém chceme najít pohybující se dlaň můžeme použít tyto dvě metody současně. Abychom docílili správné segmentace je nutné aby každá z těchto metod zpracovávala původní snímek obrazu, nikoliv už upravený jednou z těchto metod. V logickém postavení bychom mohli říct, že tyto metody použijeme paralelně.

Po nasazení každé z těchto metod použijeme na úpravu výsledného obrazu morfologické operace, které nám zajistí celistvost hledaných komponent. Po té oba výsledné obrazy sloučíme a tím získáme segmentovaný obraz s hledanými objekty (Obrázek 5.4).



Obrázek 5.4 Průběh segmentace obrazu

Pro správnou a přesnou segmentaci obrazu je nutné, aby jednotlivé metody, jak již samotné detekce nebo kombinační funkce pracovali co nejlépe. Pro detekci lidské kůže by měla být použita metoda, která bude méně citlivá na rozdílné vlastnosti prostředí jako je osvětlení, barva pozadí a jiné. V tomto případě bude srovnána klasická metoda detekce pomocí barevného modelu lidské kůže a detekce postavená na principu získání informací o barvě pleti pomocí nalezení obličeje v obraze.

Matematický model barvy lidské kůže obsahující informace o odstínu a sytosti barev, který je aplikován na celý obraz vyhodnotí segmenty obrazu, které by měli odpovídat barvě lidské kůže. Po provedení této operace můžeme obraz dále upravit pomocí morfologických operací, které nám zajistí spojitost hledaných komponent.

Pomocí detekce obličeje dostaneme informace o hranicích obličeje, které budou následovně zúženy kvůli vlasům a blízkému okolí, které se do obdélníku ohraničujícího obličej dostaly. Tyto data jsou následně označeny jako bod zájmu a je z nich vygenerován nový obraz, který obsahuje pouze obličej. Pro tento nový obraz necháme vypočítat histogram pro odstín a sytost barev. Tím byla odstraněna složka jasu, která je u lidí různé barvy pleti, nebo při změnách osvětlení odlišná. Data tohoto histogramu, nesou informaci o barvě kůže člověka, který je právě na snímku. Použitím zpětné projekce histogramu na původní obraz, dostaneme segmentovaný obraz, kde je hledaný segment barva lidské kůže. Po použití morfologických operací je obraz připraven k dalšímu zpracování. Pro detekci pohybu budou použity funkce obsažené v knihovně OpenCV.

Aby byla výsledná segmentace obrazových dat pokud možno co nejlepší budou vyzkoušeny různé metody pro sloučení dílčích obrazů segmentace. Tyto metody budou založeny na principu logického součtu a součinu.

V dalším kroku je nutné odstranit případný šum, který mohl vzniknout nepříznivými podmínkami. Tento šum odstraníme pomocí morfologických operátorů. Ve výsledném obrazu nám stále může zůstat hodně objektů, které by mohly odpovídat objektům zájmu a proto můžeme odstranit ty, které jsou nám známe jako například obličeje. Po těchto výsledných úpravách bychom měli dostat obraz ve kterém budou pouze dlaně.

5.3 Rozpoznání gest a zobrazení výsledků

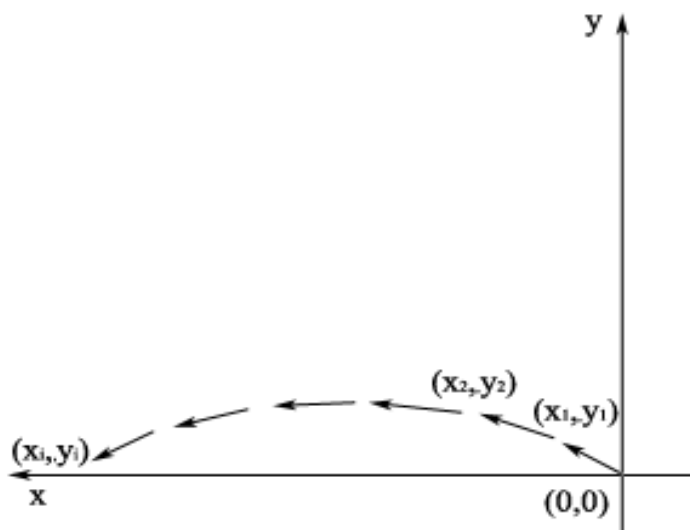
Pro rozpoznávání gest lidské ruky bych chtěl využít dva odlišné algoritmy. První algoritmus je pouze inspirovaný zjednodušenými skrytými Markovovými modely. Princip a návrh algoritmu je popsán v následujících řádcích. Druhá metoda pro srovnání je implementace skrytých Markovových modelů v knihovně OpenCV.

Abychom mohli začít rozpoznávat případně trénovat jednotlivá gesta, je nutné nejprve určit vlastnosti objektů našeho zájmu. Jelikož budeme rozpoznávat gesta dynamická musíme z hledaných objektů získat informace o jejich pohybu případně postavení vůči tělu. Po správné segmentaci obrazových dat nám může ještě ve výsledném obrazu zůstat několik malých objektů, které můžeme odstranit. Tyto objekty odstraníme nastavením minimální hranicí velikosti plochy, kterou musí každý objekt mít. Pokud je menší než námi uvedená hranice, bude objekt odstraněn. Stejnou metodou můžeme získat informace o největších objektech. Tyto objekty při správné segmentaci můžeme označit za dlaně. Takto nalezené objekty označíme elipsou (Obrázek 5.5), u které můžeme sledovat pohyb těžiště elipsy.



Obrázek 5.5 Ukázka nalezeného objektu označeného červenou elipsou

Označením objektu zájmu elipsou získáme informace o umístění objektu v obraze jelikož je nám známo těžiště elipsy. Tímto způsobem můžeme v čase určit pohyb dlaně. Při vytvoření datového souboru s modelem je nutné mít k dispozici zdroj videa, který bude obsahovat pouze trénované gesto. Jinak řečeno je nutné aby video začínalo obrazem obsahující výchozí stav gesta a končilo posledním stavem námi trénovaného gesta. Pomocí tohoto videa se nám podaří správně natrénovat model prováděného gesta. První snímek bude označen jako výchozí a jeho souřadnice budou označeny jako počátek os kartézského souřadnicového systému. Následující pohyb dlaně je popsán posunem od počátku os. Takto vytvořený systém nám dá přesné informace o pohybu dlaně. Aby byl pohyb směrodatný a nebyl příliš proměnlivý tyto informace se budou zapisovat do datového souboru s model s určitým časovým rozestupem. Tímto postupem dostane tedy soubor dat, který bude odpovídat posunu těžiště elipsy od středu v jednotném časovém intervalu (Obrázek 5.6). K jednotlivým informacím o posunu budou přidány data obsahující aktuální stav modelu, ve kterém se můžeme nacházet a také informaci o velikosti nalezeného obličeje. Tato informace je důležitá jelikož díky ní získáme přehled o tom v jaké vzdálenosti je postava od kamery v poměru s natrénovaným modelem. Tuto informaci budeme dále používat při rozpoznávání gest.

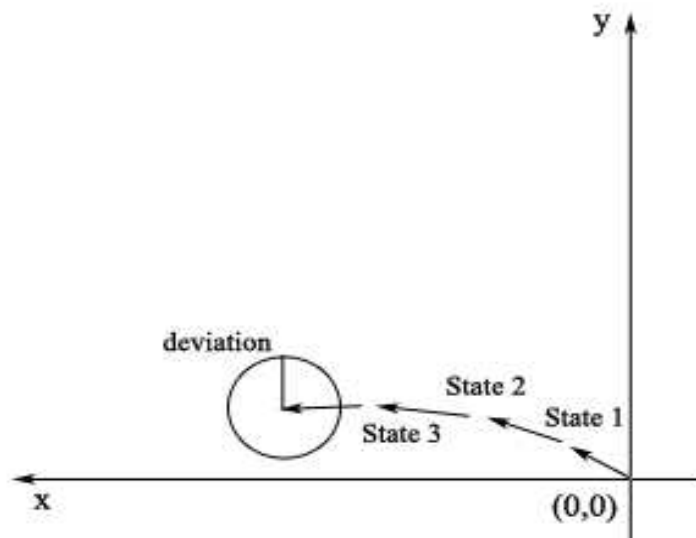


Obrázek 5.6 Ukázka pohybu gesta

Správné rozpoznávání gest je cílem celé práce. Použitý algoritmus, který je inspirován skrytými Markovovými modely, jakožto konečným automatem s daným počtem stavů a přechodů mezi nimi. Při rozpoznávání jednotlivých gest, je načtena určitá časová sekvence, jejíž první stav je označen za výchozí stav. Jelikož testovací a trénovací videa jsou omezena pouze na jednotlivá gesta, můžeme s jistotou říci, že začátek videa je výchozí stav. V dané sekvenci několika po sobě jdoucích snímků se snažíme najít stav odpovídající dalšímu stavu modelu. Pokud takový stav najdeme, můžeme se dále posouvat ve video sekvenci skokově, ze stavu na stav. Jednotlivé stavy jsou v datovém souboru určeny jako přesná pozice těžiště nalezené elipsy, abychom vůbec mohli porovnávat jednotlivá gesta, musíme hledat porovnávaná gesta s jistou odchylkou, která bude úzce souviset s poměrem velikosti nalezených obličejů (Obrázek 5.7).

Tato odchylka nám určuje kružnice, ve které by se porovnávané těžiště mělo nacházet aby splňovalo kritérium hledaného stavu. S velikostí odchylky souvisí přesnost rozpoznávání gest. Pokud

bude odchylka příliš velká, gesta můžou být chybně rozpoznána, ale na druhou stranu pokud bude odchylka příliš malá a tím pádem porovnávání přísnější může se stát, že jednotlivé gesto nebude detekováno.



Obrázek 5.7 Ukázka odchylky pro hledaný stav

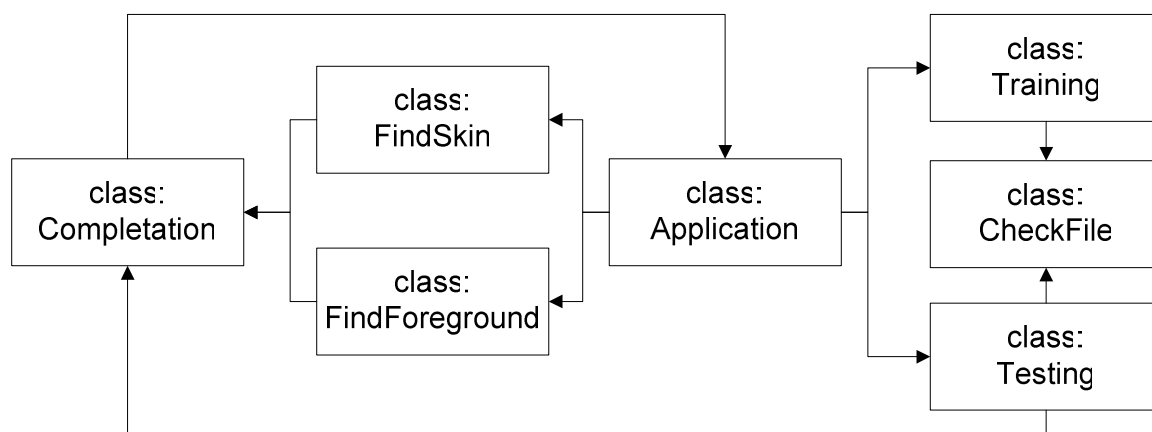
Pokud bude testované gesto rozpoznáno, bude uživatel informován o názvu modelu, který danému gestu odpovídá. Tento název se vypíše jak na standardní výstup, tak přidáním textu do samotného zdrojového videa. Aby text byl ve zdrojovém videu dostatečně čitelný bude umístěn v pravém horním rohu videa a vypsán sytě zelenou barvou (Obrázek 5.8). Text bude připsán do několika framů jdoucích po sobě, aby byl čitelný po dobu alespoň dvou sekund.



Obrázek 5.8 Ukázka informace o detekování gesta

6 Implementace

Aplikace pro rozpoznávání gest lidské ruky je konzolovou aplikací, implementovanou pomocí programovacího jazyka C++. Aplikace využívá nejnovějších knihoven OpenCV 2.2 a jejich využití v objektově orientovaném programování. Aplikace byla vytvářena v prostředí Microsoft Visual Studio 2008 v 9.0 pod operačním systémem Microsoft Windows 7. Objektově orientovaný návrh je rozdělen do sedmi tříd, přičemž každá z těchto tříd se stará určitou část návrhu aplikace. Návrhy a rozdělení tříd je uvedeno na Obrázku 6.1.



Obrázek 6.1 Propojení tříd

- 1) `Application` je řídicí třídou celé aplikace, v níž se nachází hlavní cyklus, načítající jednotlivé framy ze zdroje videa. Tyto framy jsou následně posílány do dalších tříd pro zpracování a segmentaci obrazu. Třída `Application`, se stará o kontrolu správně zadaných argumentů při spuštění aplikace, o volání konstruktorů tříd pro segmentaci obrazu, trénování popřípadě testování gest. V neposlední řadě tato třída obsahuje metodu s nápovědou pro použití programu.
- 2) `FindSkin` je první třída vyvolaná pro segmentaci obrazu. V této třídě dochází k detekci lidské kůže. Nejprve se detekuje obličej, ze kterého se přečtou informace o barvě kůže, a následně se zpětnou projekcí vyznačí místa, která odpovídají barvě lidské kůže. Obsahuje metody pro úpravy obrazu pomocí morfologických operací. Tato třída vrací obraz po segmentaci obsahující objekty s barvou lidské kůže.
- 3) `FindForeground` je třídou, která paralelně s `FindSkin` zpracovává zdrojový obraz. Tato třída se zabývá detekcí pohybu v obraze, využitím funkcí implementovaných v OpenCV. Třída vrací segmentovaný obraz po morfologických úpravách.

- 4) `Completion` se stará o spojení výše uvedených obrazů, ve kterých je následně vyznačena dlaň a získány informace o poloze dlaně. Třída se také zabývá zobrazením výsledných informací o detekci hledaného gesta.
- 5) `Training` je třída, která se stará o správný zápis do datového souboru. A vytvoření jedinečných modelů pro jednotlivá gesta. Třída umožňuje trénování gest uživatelem.
- 6) `CheckFile` obsahuje metody pro práci se soubory. Jelikož je povoleno trénování modelů uživateli, bylo nutné vytvořit třídu pro správu datového souboru. Tato třída kontroluje kontinuitu souboru. Opravuje nedokončený zápis.
- 7) `Testing`, je třída obsahující algoritmus, který se zabývá porovnáváním modelu z datového souboru a aktuálního videa a vyhodnocuje, zda-li je gesto rozpoznáno nebo nikoliv. V třídě `Testing` měly být implementovány dvě metody pro rozpoznávání gest. Zjednodušená metoda inspirovaná skrytými Markovovými modely je naimplementována. Avšak původní skryté Markovské modely obsažené v OpenCV se nepodařilo zprovoznit a propojit s objektově orientovaným návrhem.

6.1 Použití aplikace

Aplikace pro rozpoznávání gest lidské ruky je navržena jako konzolová aplikace, která očekává při spuštění správně zadané argumenty. Tyto argumenty určují některé podstatné vlastnosti chování aplikace. Aplikace obsahuje nápovědu, která je vyvolána buď špatně zadanými argumenty nebo při zadání argumentu `-h`. Nápověda obsahuje základní informace o aplikaci a jejím použití.

Aplikace očekává od jednoho argumentu, který vyvolá nápovědu, dva nebo tři další argumenty. Pokud první argument není zadán, pro zobrazení nápovědy aplikace očekává, že bude zadán zdroj videa. Zdroj videa můžeme brát tedy z video souboru, tím pádem bude jako první argument uvedena cesta k video souboru. V případě, že budeme chtít načítat video z připojené kamery, bude první argument vypadat následovně: `-cam`.

Třetí argument nám zajišťuje načtení souboru s haar klasifikátorem, který je nutný pro rozpoznávání obličeje a následnou detekci lidské kůže. Klasifikátor se nachází v kořenovém adresáři aplikace, tudíž stačí uvést jeho název.

Posledním, nepovinným argumentem je příznak pro trénování modelu. Pokud chceme natrénovat model pomocí videa, které máme uvedené jako první argument, je nutné uvést `-training` jako třetí argument. Pro názornost jsou níže uvedeny ukázky.

```
GestureRecognition -h
GestureRecognition -cam haar.xml
GestureRecognition D:\data\1.avi haar.xml
GestureRecognition D:\data\1.avi haar.xml -training
```

7 Výsledné hodnocení

Celkové výsledné hodnocení je zaměřeno na dvě rozdílné části celé aplikace. První část hodnocení se věnuje výsledkům segmentace obrazu, experimentům s použitými metodami a ukázkám výsledků segmentace v různých prostředí se ztíženými podmínkami. Druhá část se zabývá vyhodnocením rozpoznávání gest podle natrénovaných modelů a přesností rozpoznávání.

Pro celou práci a výsledné hodnocení byly vytvořeny tři testovací sady video sekvencí. První sada byla pořízena za účelem ohodnocení implementace detekce obličeje. Obsahovala frontální záběry na horní část lidského těla a několik záběrů prostředí. Druhá a třetí sada byly pořízeny pro kontrolu segmentace obrazu. Sady jsou vytvořeny v různých prostředích s proměnlivým světlem. Po testování segmentace byla lepší sada rozšířena a využita pro detekci gest lidské ruky. Informace o jednotlivých sadách jsou uvedeny v Tabulce 7.1.

	Celkový počet sekvencí	Počet osob v sekvencích
sada 1	32	26
sada 2	61	5
sada 3	130	6
sada 2 + sada 3	181	11

Tabulka 7.1 Informace o použitých sadách

7.1 Testování a výsledky segmentace obrazu

Aplikace je navržena tak, aby segmentace obrazu pro detekci lidské kůže a pohybu byla spuštěna v okamžiku, kdy dojde k detekci postavy ve video sekvenci. Postava je detekována pomocí frontální detekce obličeje. Tato detekce byla vyzkoušena pomocí první testovací sady. Cílem tohoto experimentu bylo zjištění přesnosti a spolehlivosti detekce obličeje v obraze. V následující tabulce jsou uvedeny výsledky rozpoznání obličeje. Sada obsahovala 26 sekvencí s obličejem, pozitivně byly rozpoznány všechny kromě jednoho, což bylo v tomto ojedinělém případě způsobeno sčesáním dlouhých vlasů přes značnou část obličeje. Dále sada obsahovala šest sekvencí prostředí bez osob. V žádné z těchto sekvencí nebyl detekován obličej. Z těchto poznatků můžeme určit, že detektor obličeje je plně vyhovující a jeho úspěšnost se blíží sto procentům.

Pozitivní	Falešně pozitivní	Negativní	Falešně negativní
25	0	6	1

Tabulka 7.2 Výsledky detekce obličeje

Jak již bylo řečeno, segmentace obrazu je jedním ze základních stavebních kamenů celé aplikace pro rozpoznávání gest lidské ruky. Samozřejmostí dobře pracující segmentace obrazu je dobré zpracování jednotlivých dílčích metod detekce – jak lidské kůže, tak pohybu. Možností, jak detekovat lidskou kůži, je mnoho. Zřejmě nejvhodnější bude však použití metody, která bude nejméně citlivá na

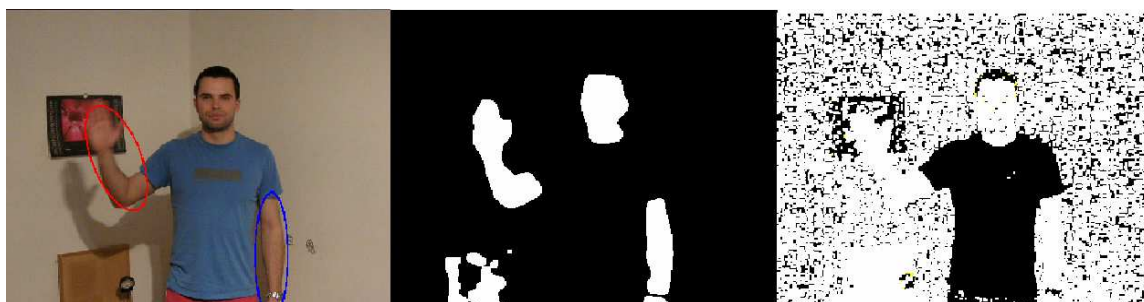
změnu osvětlení, barvu pleti a jiné nepříznivé podmínky. Porovnáním dvou odlišných metod, které detekují barvu kůže, můžeme vybrat přesnější a spolehlivější metodu. V levé části obrázku se nachází zdrojový obraz za příznivých podmínek. Metoda, která je zobrazena ve středové části, je založena na získání informací o barvě kůže z detekovaného obličeje. V poslední části vidíme detekci pomocí matematického modelu barvy kůže. Z obrázku je zřejmé, že obě metody za těchto podmínek pracují správně.

(Obrázek 7.1). U metody využívající matematický model barvy kůže můžeme vidět i členitost prstů.



Obrázek 7.1 Srovnání detekce kůže v příznivém prostředí

Při detekci za zhoršených podmínek se projevila jedna z metod jako nevyhovující. Rozdíl mezi jednotlivými obrázky je velmi patrný a rozeznatelný pouhým okem (Obrázek 7.2).



Obrázek 7.2 Srovnání detekce kůže v nepříznivém prostředí

Levá část obrázku znázorňuje zdrojové video s již označenými body zájmu. Zdrojové video bylo pořízeno v místnosti s umělým žlutým osvětlením a lehce krémovou barvou stěn. Tyto podmínky by se daly považovat za zhoršené pro segmentaci dat v obraze. Na středové části je ukázka segmentace obrazu pomocí detekce obličeje a následné zpětné projekce histogramu ze získaných dat. Na této části jsou bílou barvou označeny plochy, které odpovídaly v daný okamžik získané barvě obličeje s jistou malou odchylkou způsobenou prahováním. Jasně zde rozeznáme obličej, horní končetiny a též další blob, který se v daný okamžik nacházel na videu a pod umělým osvětlením měl podobnou barvu jako lidská kůže. Jak je vidno na původním obrázku, jedná se o korkovou nástěnku. Ve třetí části je vidět obraz po segmentaci s použitím barevného modelu kůže. Bohužel, rozeznání končetin nebo obličeje je složité i pro lidské oko. Bílá barva by opět měla znázorňovat plochu pokrytou barvou lidské kůže. Tato část obrazu pro nepřehlednost neprošla úpravou pomocí morfologických operací, jelikož po jejich provedení již nebyla znatelná ani silueta postavy. Tato metoda se v dané situaci projevila jako zcela nepoužitelná z důvodů splnutí pozadí s hledanými objekty.

I když se zprvu mohla zdát metoda založená na principu matematického modelu barvy lidské kůže jako lépe vyhovující z důvodu přesnější detekce, při nepříznivých podmínkách se ukázala být

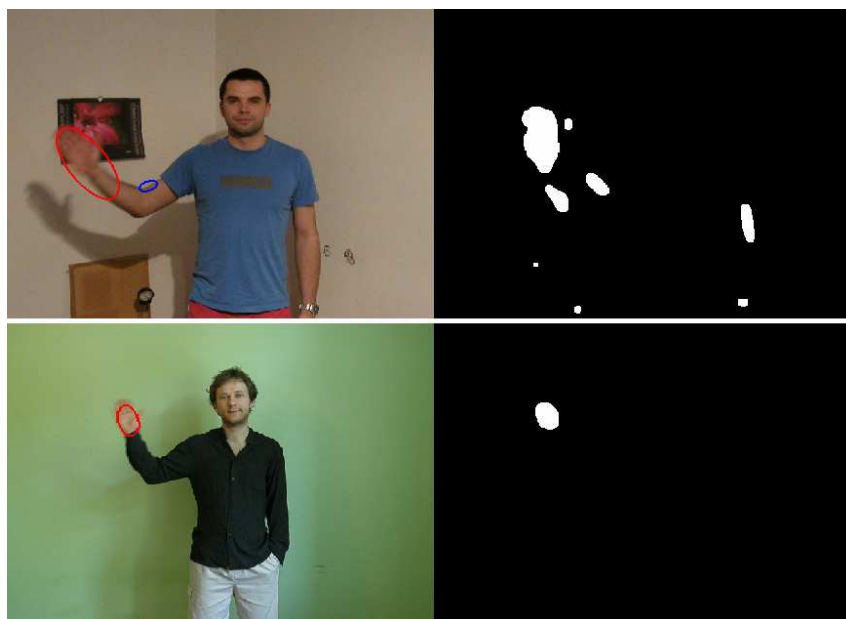
zcela nevyhovující. Jelikož aplikace rozpoznává gesta dynamická, je výhodnější použít metodu získávající informace o barvě kůže z detekovaného obličeje.

Další dílčí částí pro segmentaci obrazu je detekce pohybu. Pro tuto část jsem zvolil metodu odečítání pozadí, která je popsána v kapitole 3.2 Detekce pohybu v obraze. Základním stavebním kamenem této metody je odečítání po sobě jdoucích snímků. Proto je důležité, aby zdroj videa byl pořízen statickou kamerou. Jedná se do určité míry o omezení, ovšem při ovládání aplikace gesty by s největší pravděpodobností kamera ležela či stála na stativu. Na Obrázku 7.3 můžeme vidět příklad segmentace obrazu pomocí detekce pohybu. Levá část obrázku je zdrojové video, které je opět pořízené za zhoršených podmínek, v tomto případě nestatickou kamerou. Ve střední části je znázorněna detekce pohybu v okamžiku, kdy kameraman posunul kameru jedním směrem. Tím pádem došlo k detekování pohybu celé postavy i výrazných bodů na pozadí. V pravé části je pohyb kamery omezen, ale i tak je zde jasně vidět šum vzniklý třesem kamery nebo pohybem uživatele. Ovšem největší bílý blob představuje pohyb ruky, v tomto případě pohyb ruky i vzniklý stín. Detekce pohybu je tedy dostačující za podmínek statické kamery.



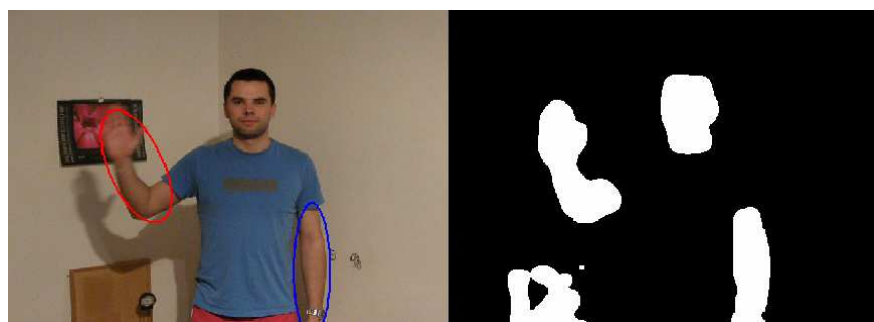
Obrázek 7.3 Ukázka detekce pohybu

Po provedení detekce lidské kůže a detekce pohybu jsou následně provedeny morfologické operace, které zajistí sjednocení hledaných objektů. Nyní můžeme přistoupit ke sloučení těchto dvou dílčích celků. Spojením těchto celků bychom měli dostat výsledný segmentovaný obraz. Následně by měl obraz obsahovat pouze pohybující se objekty v barvě lidské kůže. Experimentálně byly zjištěny dva způsoby, jak spojit tyto dvě části segmentace obrazu. Dle výsledků porovnání obou metod je každá použitelná za jiných podmínek. Prvním způsobem pro spojení jednotlivých obrazů je sloučení všeho, co mají společné. Jinak řečeno, mezi dvěma binárními obrazy provedeme logický součin. Díky této operaci dostaneme opravdu pouze objekty, které se pohybují a mají barvu kůže. Tato metoda není příliš vhodná za takových podmínek, kdy jedna z metod segmentace obrazu zaostává: v takovém případě dostaneme nepříliš odpovídající informace. Na Obrázku 7.4 můžeme vidět dva příklady použití této metody. Příklad v horní polovině obrázku nemá při záznamu statickou kameru, čímž dochází k nežádoucímu šumu při detekci pohybu. Po logickém součinu s detekovanou barvou kůže dostáváme nepříliš odpovídající obraz pro hledané objekty. Ukázka v dolní polovině obrázku má již statickou kameru a obě detekce jsou přesné, čemuž odpovídá také přesně detekovaná pohybující se dlaň.



Obrázek 7.4 Sloučení dílčích obrazů pomocí logického součinu

Další metodou po sloučení jednotlivých dílčích obrazů je sečtení těchto částí. Ve výsledném obrázku sice dostaneme informace jak o objektech, které mají barvu kůže, tak o objektech, které se pohybují. Ve výsledku získáme daleko více blobů, které je třeba rozpoznávat, ovšem tato metoda se osvědčila při zhoršených podmínkách pro detekci pohybu. Na Obrázku 7.5 můžeme vidět příklad pro stejné podmínky jako v horní polovině Obrázku 7.4.



Obrázek 7.5 Sloučení dílčích obrazů pomocí sečtení

Při použití druhé metody sloučení je jasné, že nemůžeme použít všechny části po segmentaci videa. Proto si musíme zvolit další kritéria tak, abychom dostali pokud možno jasnou odpověď o tom, který z blobů opravdu představuje pohybující se dlaň lidské ruky. Po odstranění obličeje ze segmentovaného videa můžeme určit, že největší pohybující se blob je právě hledaná dlaň.

Pro testování jednotlivých metod segmentace byly použity sady video sekvencí. Obě sady byly pořízeny v interiéru pod umělým osvětlením. Pro sadu číslo dvě bylo požádáno pět různých osob o provedení jednotlivých gest. Na pozadí světle krémové barvy této sady je vidět obraz a položená korková nástěnka, která byla využita jako metoda pro falešně pozitivní detekci lidské kůže. Místnost je osvětlena žlutým světlem snižujícím kvalitu pořízených dat. Sada číslo tři je také natočena v interiéru s proměnlivým osvětlením. V této sadě jsou použity tři druhy osvětlení, a to přirozené osvětlení, osvětlení odrazem světla bílého reflektoru, a umělé žluté osvětlení. V následující tabulce

(Tabulka 7.3) je uvedeno procentuální vyjádření vyrovnání se jednotlivých metod s uvedenými stěžujícími podmínkami. Jinak řečeno, v tabulce jsou uvedena procenta ukazující, v kolika případech byla detekce lidské kůže správná.

	Sada č. 2		Sada č. 3		
	korek	pozadí	žárovka	reflektor	slunce
Počet sekvencí s daným prvkem	61	61	30	64	36
Model barvy kůže	16,39%	19,67%	96,66%	93,75%	91,66%
Histogram barvy obličeje	14,75%	91,80%	93,33%	89,06%	91,66%

Tabulka 7.3 Účinnost detekce barvy lidské kůže

Z těchto dat vyplývá, že obě metody jsou vyrovnané a liší se pouze v jednom bodě, kdy barva pozadí byla velmi podobná barvě lidské kůže. Vzhledem k tomuto rozdílu jsem se rozhodl dále použít detekci založenou na získání informací o barvě lidské kůže z detekovaného obličeje, která má v porovnání s druhou metodou uspokojivé výsledky.

7.2 Testování a výsledky rozpoznávání gest

Rozpoznávání gest lidské ruky je provedeno pomocí zjednodušeného algoritmu inspirovaného skrytými Markovovými modely. Implementace algoritmů skrytých Markovových modelů obsažených v knihovnách OpenCV se nepodařila sloučit s objektově orientovaným návrhem aplikace. Algoritmus byl podle dostupných informací v pořádku, avšak při vytváření modelu selhal a nemohl být následně použit. Z tohoto důvodu budou diskutovány pouze výsledky získané pomocí výše zmíněné metody.

Metoda rozpoznává gesta lidské ruky ve video sekvenci. Pro rozpoznávání jednotlivých gest byla použita sada číslo 3, která obsahuje sekvence s různě se opakujícími gesty. Na pořízení této sady se podílelo šest různých osob, které byly seznámeny s jednotlivými gesty. Každá osoba udělala minimálně jednu každé z natrénovaných gest; jednotlivá gesta prošla rozpoznávacím algoritmem.

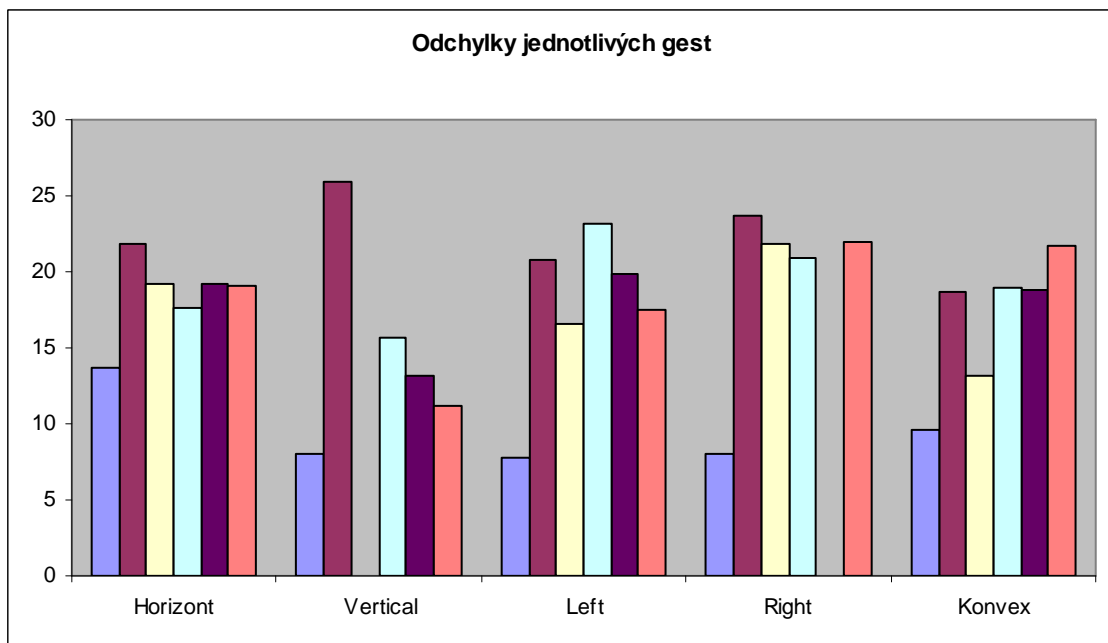
Při rozpoznávání je nastavená odchylka, která nám určuje prostor, o který se může daný stav posunout. S nejvhodněji zvolenou odchylkou tato metoda dosáhla slabších, ale uspokojivých výsledků. Pokud bychom chtěli vyčíslit úspěšnost tohoto rozpoznávání, je možno k tomu přistupovat několika způsoby. Nejzákladnějším způsobem je porovnání počtu jednotlivých rozpoznávaných gest s celkovým počtem těchto gest. Toto porovnání je uvedeno v Tabulce 7.4.

	Horizont	Vertical	Left	Right	Konvex
Počet gest v sadě	30	26	29	22	23
Úspěšnost rozpoznání	43,33%	34,61%	34,48%	31,82%	60,87%

Tabulka 7.4 Úspěšnost rozpoznání

Podobnost některých gest nebo nevhodně nastavená odchylka může vést k falešně pozitivnímu rozpoznání. Pokud je ovšem odchylka nastavena na malé hodnoty, výskyt jevu záměny gest je téměř nulový.

Algoritmus byl rozšířen o počítání průměrné odchylky od modelu. Tato odchylka je vzdálenost, ve které se pozorování nachází. Průměrná odchylka je počítána z rozpoznaných gest. V Grafu 7.1 jsou znázorněny průměrné odchylky jednotlivých gest pro každého uživatele. Můžeme tím sledovat přesnost (jak je který uživatel přesný a schopný dané gesto správně ukázat). U gesta Vertical si můžeme všimnout, že jeden z uživatelů nemá definovanou odchylku. To je způsobeno špatně pořízenou sadou gest, kdy dané gesto nebylo u tohoto uživatele správně natočeno. U gesta Right také nemá jeden z uživatelů spočítanou odchylku, v tomto případě je to z důvodu neschopnosti rozpoznat dané gesto.



Graf 7.1 Průměrné odchylky od modelu

8 Závěr

Tématem bakalářské práce bylo rozpoznávání gest lidské ruky. Gesta rukou tvoří značnou část komunikace mezi lidmi. Stejná gesta se dají použít při dorozumívání člověka s elektronikou. Důležitou částí této práce bylo nastudování metody pro zpracování obrazu a rozpoznávání gest ve video sekvenci. Cílem práce bylo vytvořit jednoduchou aplikaci zpracovávající vstupní obraz a následně v něm rozpoznávat dynamická gesta.

V rámci práce byla vytvořena aplikace na rozpoznávání gest lidské ruky. Aplikaci je ve svém principu plně funkční s možností různých vylepšení. Součástí aplikace je návrh systému pro segmentaci obrazových dat. Pro použití správných metod segmentace obrazu byly provedeny experimenty a porovnání různých druhů přístupu. Metody použité ve výsledné aplikaci jsou diskutovány v závěrečném hodnocení práce. Ve výsledku se domnívám, že navržená segmentace obrazu pracuje správně i za nepříznivých podmínek a je plně dostačující.

Vytvořený algoritmus pro rozpoznávání gest byl inspirován skrytými Markovovými modely. Bohužel se nepodařilo propojit aplikaci s Markovovými modely implementovanými v OpenCV, tím pádem nemohla být porovnána úspěšnost těchto algoritmů. Rozpoznávání gest lidské ruky pomocí zjednodušeného algoritmu je méně přesné, při dobrých podmínkách je však dostačující.

Přínos této práce spočívá v porovnání metod segmentace obrazu a použití více metod současně, dále v návrhu čistě dynamických gest, které by mohly být použity pro ovládání multimediálních přehrávačů nebo prezentací, a jejich rozpoznávání ve video sekvenci.

V budoucnu bych se rád znovu vrátil k tématu rozpoznávání gest lidské ruky a snažil se tak tuto práci rozšířit a vylepšit. Možnost vylepšení práce vidím v nasazení plné podpory skrytých Markovových modelů a jejím porovnání s navrženým algoritmem. Dalším zajímavým vylepšením by mohlo být použití metod pro zrychlení zpracování obrazu, jako jsou například TBB knihovny, či využití grafických akceleratorů. Jako rozšíření by bylo vhodné zvětšení trénování sady gest, popřípadě přidání jiných typů gest k rozpoznávání.

Literatura

- [1] MLČOCH, Z.: *Neverbální komunikace - co prozradí gesta, mimika a postoj* [online]. Olomouc: 2003-2011, [cit. 2011-04-24]. Dostupné na URL: <http://www.zbynekmlcoch.cz/info/psychologie_aneb_ty_a_ja/neverbalni_komunikace_co_prozradi_gesta_mimika_a_postoj.html>.
- [2] Kulturní centrum RUCÉ: *Struktura a fungování znakového jazyka* [online]. Praha: 2004-2011, [cit. 2011-04-24]. Dostupné na URL: <<http://ruce.cz/znakovy-jazyk/odborne>>.
- [3] FÍŠER, M.: *Revoluce v hraní her, ovladače nahradí lidská gesta* [online]. Seznam.cz, Praha: 2003 - 2011, [cit. 2011-04-24]. Dostupné na URL: <<http://www.novinky.cz/internet-a-pc/hry-a-herni-systemy/170455-revoluce-v-hrani-her-ovladace-nahradi-lidska-gesta.html>>.
- [4] E-learning TUL: *Segmentace obrazu* [online]. Technická univerzita Liberec, Liberec: 2011, [cit. 2011-04-27]. Dostupné na URL: <http://e-learning.tul.cz/cgi-bin/elearning/elearning.fcgi?ID_tema=67&ID_obsah=1202&stranka=publ_tema&akce=polozka_vstup>.
- [5] KÖLSH, M.: *Overview, The HandVu* [online]. [cit 2011-04-27]. Dostupné na URL: <<http://www.movesinstitute.org/~kolsch/HandVu/HandVu.html#overview>>.
- [6] KÖLSH, M.: *Gestore Posture, The HandVu* [online]. [cit 2011-04-27]. Dostupné na URL: <http://www.movesinstitute.org/~kolsch/HandVu/doc/HandVu_Postures.pdf>.
- [7] MOLENAAR, G.: *About, Sonic Gesture* [online]. Amstrodam: 2010, [cit 2011-04-27]. Dostupné na URL: <<http://code.google.com/p/sonic-gesture/>>.
- [8] GestureTek Health: *Rehabilitation, GestureTek* [online]. USA: 2011, [cit 2011-04-27]. Dostupné na URL:<<http://www.gesturetekhealth.com/products-rehab.php>>.
- [9] KURUC, J.: *Ovládání dotykem a gesty: od mobilů a her míří k počítači* [online]. Computer, Praha: 2011, [cit 2011-04-27]. Dostupné na URL: <<http://www.zive.cz/clanky/ovladani-dotykem-a-gesty-od-mobilu-a-her-miri-k-pocitaci/kapacitni-displeje-ovladani-pohybem-tela/sc-3-a-152778-ch-71900/default.aspx>>.
- [10] MicrosoftWave: *No Controller Necessary* [online]. USA: 2011, [cit 2011-04-27]. Dostupné na URL: <<http://www.microsoft.com/uk/wave/hardware-kinect.aspx>>.
- [11] CHANG H., ROBES U.: *Face detection* [online] Stanford University, California, USA: 2000, [cit 2011-04-27]. Dostupné na URL: <<http://www-cs-students.stanford.edu/~robles/ee368/main.html>>.

- [12] SLOVÁČEK, M.: *Evaluation of RGB and HSV models in Human Faces Detection* [online]. Faculty of Informatics and Information Technologies, Slovak University of Technology, Bratislava, Slovensko: 2004, [cit 2011-04-27]. Dostupné na URL: <<http://www.cg.tuwien.ac.at/hostings/cescg/CESCG-2004/web/Sedlacek-Marian/>>.
- [13] LEE J.Y., YOO S.I.: *An Elliptical Boundary Model for Skin Color Detection*, The 2002 International Conference on Imaging Science, Systems, and Technology , Las Vegas, USA: 2002, [cit 2011-04-27]. Dostupné na URL: <<http://ailab.snu.ac.kr/publication/down/CISST02-169CT.pdf> >.
- [14] HINNER J.: *Detekce a rozpoznávání obličejů osob a jejich identifikační význam*, [online] Kriminalistika Praha: 2008, [cit 2011-04-27]. Dostupné na URL: <http://aplikace.mvcr.cz/archiv2008/casopisy/kriminalistika/2003/03_01/hinner.html>.
- [15] ZAACHI: *Základy detekce obličeje v obraze* [online], 2008, [cit 2011-04-29]. Dostupné na URL: <<http://www.zaachi.com/cs/items/zaklady-detekce-obliceje-v-obraze.html>>.
- [16] MUDROVÁ, M.: *Histogram a jeho zpracování* [online], Vysoká škola chemicko-technická v Praze, Praha: 2004, [cit 2011-04-27]. Dostupné na URL: <<http://uprt.vscht.cz/mudrova/zob/prednasky/03-HISTOGRAM/histogramcz-tisk.pdf>>.
- [17] KOUTKOVÁ H. PEŠL J.: *Histogram* [online] Fakulta stavební, Vysoké učení technické v Brně, Brno: 2001, [cit 2011-05-01]. Dostupné na URL: <<http://math.fce.vutbr.cz/vyuka/matematika/statistika/histogram.html>>.
- [18] VILA, J. G.: *Foreground Segmentation and Tracking based on Foreground and Background Modeling Techniques* [online]. ETSETB TSC: 2009, [cit 2011-05-02]. Dostupné na URL <<http://upcommons.upc.edu/pfc/bitstream/2099.1/7006/1/Proyecto%20MERIT%20Master%20Jaume%20Gallego%20Final%20Version%2013Febrero09%202.pdf> >.
- [19] YOUNG, D.: *Motion* [online]. University of Sussex, UK: 1994, [cit 2011-05-02]. Dostupné na URL: <http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/YOUNG/vision6.html#heading3 >.
- [20] MUDROVÁ, M.: *Matematická morfologie a segmentace obrazu* [online]. Vysoká škola chemicko-technologická Praha: 2004, [cit 2011-05-02]. Dostupné na URL: <<http://uprt.vscht.cz/mudrova/zob/prednasky/10-MORFOLOGIE/morfologie-tisk.pdf>>.
- [21] HLAVÁČ, V.: *Matematická Morfologie* [online]. Fakulta elektrotechnická, ČVUT, Praha: 2006, [cit 2011-05-02]. Dostupné na URL: <<http://cmp.felk.cvut.cz/~hlavac/TeachPresCz/11DigZprObr/71-03BinMatMorfolCesky.pdf>>.
- [22] LI, Jia: *Hidden Markov Model* [online]. Department of Statistics, The Pennsylvania State University. USA: 2007, [cit 2011-05-02]. Dostupné na URL: <<http://www.stat.psu.edu/~jiali/course/stat597e/notes2/hmm.pdf>>.

- [23] WARAKAGODA, N.: *Hidden Markov Models* [online]. 1996, [cit 2011-05-02]. Dostupné na URL: < <http://jedlik.phy.bme.hu/~gerjanos/HMM/node2.html>>.
- [24] GRIGIC, M.: *Face Recognition Homepage – Algorithms* [online]. University of Zagreb. Záhřeb, Chorvatsko: 2008, [cit 2011-05-02]. Dostupné na URL: < <http://www.face-rec.org/algorithms/#Video>>.
- [25] JORGENSEN, A.: *AdaBoost and Histograms for Fast Face Detection* [online]. School of Computer Science and Communication. Stockholm. Sweden: 2006, [cit 2011-05-02]. Dostupné na URL: <http://www.nada.kth.se/utbildning/grukth/exjobb/rapportlistor/2006/rapporter06/jorgensen_anders_06110.pdf>.
- [26] T & L Publications: *How Face Detection Works* [online]. SERVO Magazine: 2007, [cit 2011-05-02]. Dostupné na URL: <http://www.cognotics.com/opencv/servo_2007_series/part_2/sidebar.html>.
- [27] DARELL, R.: *Kinect Hack Makes Future Navigation Just Like Minority Report!* [online]. 2010, [cit 2011-05-03] Dostupné na URL: < <http://www.bitrebels.com/technology/kinect-hack-makes-future-navigation-just-like-minority-report/> >.
- [28] BRADSKI, G.; KAEHLER, A.: *Learning OpenCV: Computer Vision with the OpenCV Library*. O'Reilly Media, Sebastopol, USA: 2008, ISBN 0596516134.
- [29] HILL, J.: *Nintendo Wii: Revolutionary* [online], The Age Company, Austrálie: 2006, [cit 2011-05-03]. Dostupné na URL: <http://blogs.theage.com.au/screenplay/archives/battle_stations/002911.html> .
- [30] HLAVÁČ. V.: *Rozpoznávání s Markovskými modely* [online]. ČVUT, Praha: 2008, [cit 2011-05-03]. Dostupné z URL: < <http://cmp.felk.cvut.cz/~hlavac/TeachPresCz/31Rozp/61MarkovianPR.pdf> >.
- [31] Kolektiv autorů: *Pravidla českého pravopisu*. Academia, Praha, 1993. ISBN 80-200-0475-0.

Seznam příloh

Příloha 1. DVD Rozpoznávání gest lidské ruky

Obsah DVD:

Data – sady gest

Aplikace – spustitelná aplikace pro rozpoznávání dynamických gest lidské ruky

Zdrojové soubory – zdrojové soubory s komentářem celé aplikace

Technická zpráva – kompletní technická zpráva ve formátu PDF

Plakát – presentační plakát bakalářské práce