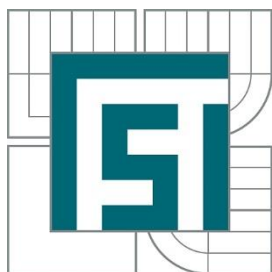


VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA STROJNÍHO INŽENÝRSTVÍ
ÚSTAV AUTOMATIZACE A INFORMATIKY
FACULTY OF MECHANICAL ENGINEERING
INSTITUTE OF AUTOMATION AND COMPUTER SCIENCE

DETEKCE MIMIKY OBLIČEJE

DETECTION OF FACIAL EXPRESSION

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. ROMAN KOZÁK

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. JAN ROUPEC, Ph.D.

BRNO 2015

Vysoké učení technické v Brně, Fakulta strojního inženýrství

Ústav automatizace a informatiky

Akademický rok: 2014/2015

ZADÁNÍ DIPLOMOVÉ PRÁCE

student(ka): Bc. Roman Kozák

který/která studuje v **magisterském navazujícím studijním programu**

obor: **Aplikovaná informatika a řízení (3902T001)**

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

Detekce mimiky obličeje

v anglickém jazyce:

Detection of facial expression

Stručná charakteristika problematiky úkolu:

Práce se bude zabývat digitálním zpracováním obrazu. V obrazu bude detekován obličej (obličeje) a bude analyzován jeho výraz (mimika). K řešení bude použita knihovna OpenCV.

Cíle bakalářské práce:

1. Popište problematiku zpracování obrazu.
2. Uveďte metody detekce obličeje.
3. Popište knihovnu OpenCV.
4. Vytvořte ukázkovou aplikaci.

Seznam odborné literatury:

1. Bradski G., Kaehler A.: Learning OpenCV. O'Reilly, Cambridge, 2008.
2. Baggio D. L., Emami S., Escrivá D. M.: Mastering OpenCV with Practical Computer Vision Projects. PACKT Publishing, Birmingham, 2012.

Vedoucí bakalářské práce: Ing. Jan Roupec, Ph.D.

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2014/2015.

V Brně, dne 25. 12. 2014

L.S.

Ing. Jan Roupec, Ph.D.
Ředitel ústavu

doc. Ing. Jaroslav Katolický, Ph.D.
Děkan fakulty

ABSTRAKT

Tato diplomová práce se zabývá návrhem a tvorbou aplikace pro rozpoznávání výrazu obličeje. V úvodu práce je popsána problematika zpracování obrazu, především algoritmy, které jsou použity v dalších částech. Dále je věnována pozornost metodám detekce obličeje, zejména detektoru Viola-Jones. Poté následuje stručný popis základních obličejových výrazů, které budou aplikací rozpoznávány. Poslední dvě části se věnují samotnému návrhu a realizaci aplikace a zhodnocení výsledků.

ABSTRACT

This master's thesis deals with the design and implementation of computer application for facial expression recognition. The first section describes the image processing, especially algorithms used in other parts of this thesis. The next section describes methods used for face detection, particularly the Viola-Jones detector. This is followed by a brief description of the basic facial expressions recognized by the application. The last two sections are dedicated to the design and implementation of the facial expression recognition algorithms and to the evaluation of the results.

KLÍČOVÁ SLOVA

zpracování obrazu, strojové učení, počítačové vidění, detekce obličeje, mimika obličeje, C++, OpenCV

KEYWORDS

image processing, machine learning, computer vision, face detection, facial expressions, C++, OpenCV

PROHLÁŠENÍ O ORIGINALITĚ

Prohlašuji, že jsem diplomovou práci vypracoval samostatně dle pokynů vedoucího práce s použitím uvedené odborné literatury.

Brno 29. května 2015

.....
Roman Kozák

BIBLIOGRAFICKÁ CITACE

KOZÁK, R. *Detekce mimiky obličeje*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2015. 51 s. Vedoucí diplomové práce Ing. Jan Roupec, Ph.D.

PODĚKOVÁNÍ

Na tomto místě bych rád poděkoval vedoucímu práce Ing. Janu Roupčovi, Ph.D. za jeho čas, podněty a věcné připomínky. Dále bych chtěl poděkovat svojí rodině a přátelům za podporu během studia.

OBSAH

1	Úvod.....	13
2	Zpracování obrazu a strojové učení	15
2.1	Barevné modely.....	15
2.1.1	Model RGB.....	15
2.1.2	Model YCbCr	15
2.2	Řetězec zpracování obrazu.....	16
2.3	Ekvalizace histogramu	17
2.4	Konvoluce	18
2.5	Strojové učení.....	19
3	Metody detekce obličeje.....	21
3.1	Detekce obličeje podle barvy kůže.....	21
3.2	Detektor Viola-Jones.....	22
3.2.1	Haarovy vlnky	22
3.2.2	Integrální obraz.....	23
3.2.3	Adaptivní boosting	24
3.2.4	Kaskáda klasifikátorů	24
4	Výrazy obličeje	27
4.1	Hněv	27
4.2	Opovržení	28
4.3	Znechucení	28
4.4	Strach.....	29
4.5	Radost.....	29
4.6	Smutek.....	29
4.7	Překvapení.....	30
5	Návrh aplikace.....	31
5.1	Trénování klasifikátorů	31
5.1.1	Trénovací množina	31
5.1.2	Lokalizace obličeje	32
5.1.3	Extrakce příznaků	32
5.1.4	Trénování klasifikátorů.....	34
5.2	Rozpoznávání výrazů obličeje.....	34
5.2.1	Klasifikace	34
6	Implementace	35
6.1	Knihovna OpenCV	35
6.2	Struktura dat	35
6.3	Sdílené funkce	36
6.4	Trénování klasifikátorů	36

6.4.1	Oříznutí obličejů.....	37
6.4.2	Extrakce příznaků.....	37
6.4.3	Příprava trénovacího setu.....	37
6.4.4	Trénování klasifikátorů	38
6.4.5	Test.....	38
6.5	Rozpoznávání výrazů obličeje	38
6.6	Výsledky.....	39
6.6.1	Statické snímky	39
6.6.2	Webkamera	40
6.6.3	Vliv šumu	43
6.6.4	Závislost na osvětlení	43
7	Závěr	45
	Seznam použité literatury	47
	Seznam příloh	49

1 ÚVOD

Mimika je důležitou součástí neverbální komunikace. Základní emoce a jejich projevy prostřednictvím obličejových výrazů jsou univerzální napříč kulturami po celém světě. Jelikož s výpočetní technikou všeho druhu trávíme stále více času, začíná být důležité, aby výrazům našich obličejů rozuměli nejen lidé, ale i počítače. Lokalizace obličeje a jeho analýza jsou proto v posledních letech intenzivně zkoumané oblasti počítačového vidění.

Cílem této práce je navrhnout a vytvořit aplikaci pro rozpoznávání výrazu obličeje. Jako zdroj obrazu bude sloužit webkamera. Pro úlohy zpracování obrazu a strojového učení bude použita knihovna OpenCV.

V úvodních částech práce bude popsána problematika zpracování obrazu a metody detekce obličeje, poté bude následovat rozbor základních obličejových výrazů. Praktická část bude věnována návrhu a tvorbě samotné aplikace a zhodnocení výsledků.

2 ZPRACOVÁNÍ OBRAZU A STROJOVÉ UČENÍ

Zpracování obrazu je ve své podstatě zpracování vícerozměrného signálu. Obraz (pro zjednodušení monochromatický) lze popsat spojitou funkcí dvou nebo tří proměnných, která se nazývá obrazová funkce. Pro statický obraz postačí obrazová funkce $f(x, y)$, kde proměnné x a y určují polohu bodu v rovině, pro obraz měnící se v čase t se využije funkce $f(x, y, t)$. Hodnoty obrazové funkce v jednotlivých bodech představují velikost fyzikální veličiny, kterou je obraz tvořen. Například u obrazu zachyceného černobílým fotoaparátem nebo kamerou se jedná o jas, u obrazu z termokamery se jedná o teplotu. [1]

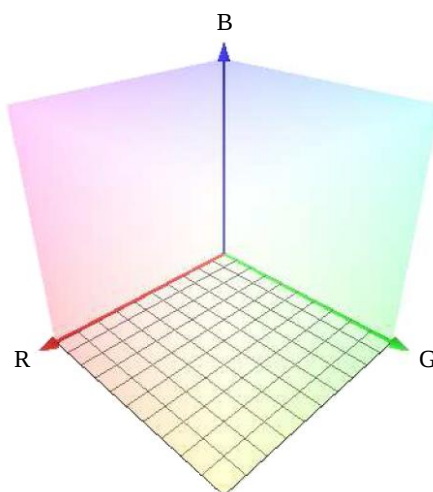
Digitální obraz, s kterým pracuje počítač, je reprezentován maticí obrazových elementů (pixelů, z anglického *picture element*), které jsou dále nedělitelné.

2.1 Barevné modely

Barvy mohou být v obraze reprezentovány různými modely. Každý model je vhodný pro jiný typ zpracování. Nejvýznamnějšími modely z hlediska zpracování obrazu jsou modely RGB a YCbCr.

2.1.1 Model RGB

Model RGB je základním a nejpoužívanějším modelem reprezentace barev. Výsledná barva je definována intenzitami tří základních složek (červenou, zelenou a modrou).

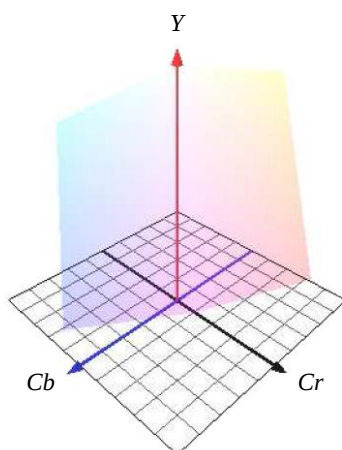


Obr. 1 Barevný prostor modelu RGB. [2]

2.1.2 Model YCbCr

Model YCbCr lze vytvořit transformací modelu RGB a jeho velkou výhodou je oddělení jasové informace, kterou reprezentuje složka Y , od informace o barvě, kterou

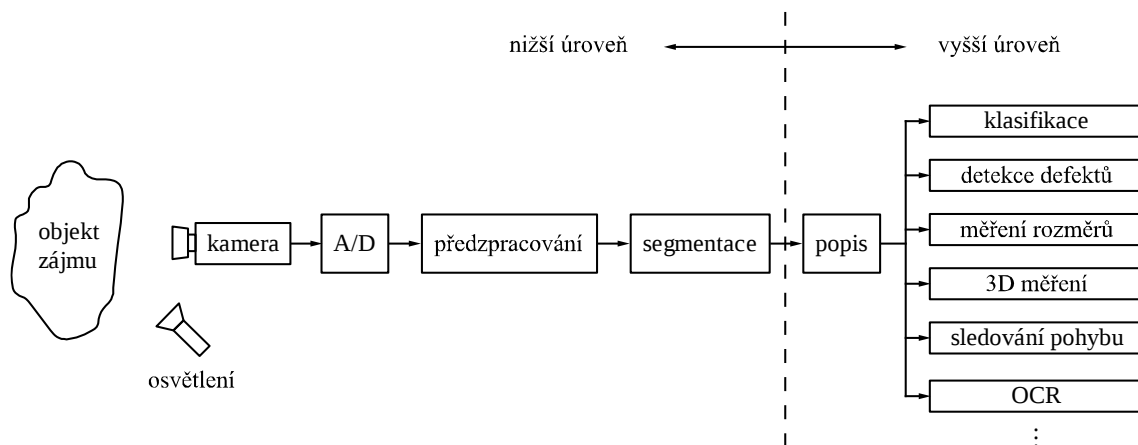
reprezentují složky Cb a Cr . Model je využíván při ztrátové kompresi fotek a videa a při detekci obličeje podle barvy kůže.



Obr. 2 Barevný prostor modelu YCbCr. [2]

2.2 Řetězec zpracování obrazu

Zpracování obrazu je součástí počítačového vidění, které se technickými prostředky snaží napodobit schopnosti lidského vidění. Zpravidla má několik částí, které po sobě následují. Ty lze rozdělit do dvou úrovní – nižší úroveň a vyšší úroveň.

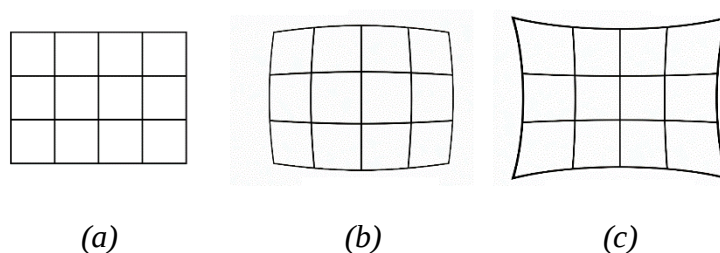


Obr. 3 Řetězec zpracování obrazu. [3]

Získání obrazu

Počátek řetězce tvoří snímací zařízení (fotoaparát nebo kamera), jež snímá scénu s objektem zájmu. Důležitým parametrem je osvětlení scény, které musí být vhodně přizpůsobeno typu úlohy. Na vstupu snímacího zařízení se nachází objektiv, který obraz

snímaného objektu soustředí na obrazový senzor. Objektiv může do obrazu vnášet různé vady, které bývají kompenzovány při předzpracování obrazu.



Obr. 4 Zkreslení objektivu: (a) reálný obraz, (b) soudkovitost, (c) poduškovitost.

Předzpracování obrazu

Obraz je zdigitalizován a následuje jeho předzpracování. Cílem předzpracování je zlepšit kvalitu obrazu například odstraněním šumu, vyrovnaním zkreslení objektivu, ekvalizací histogramu či doostřením.

Segmentace

Poslední fází nižší úrovně zpracování obrazu je segmentace, jejímž úkolem je rozdělit obraz na požadované oblasti nebo oddělit objekty v obraze od pozadí.

Popis

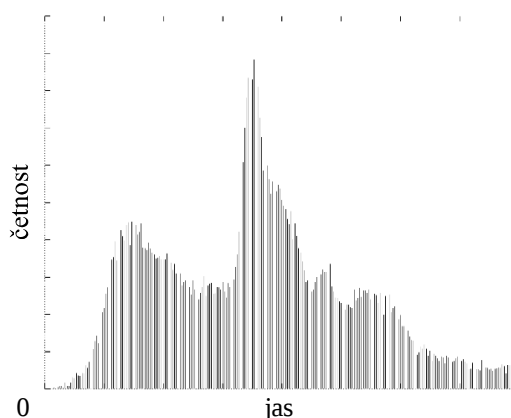
Cílem další fáze je popsat vlastnosti objektů, které jsou v poslední fázi procesu vyhodnoceny.

Vyhodnocení

Mechanismem vyhodnocování může být například klasifikace objektů, měření jejich rozměrů, detekce defektů nebo sledování jejich pohybu.

2.3 Ekvalizace histogramu

Histogram černobílého obrazu poskytuje informace o četnosti výskytu jednotlivých jasových úrovní v obraze. Často se zobrazuje jako sloupcový graf, jehož vodorovná osa znázorňuje všechny jasové složky od nejtmaší po nejsvětlejší a svislá osa reprezentuje jejich četnosti. [1]



Obr. 5 Histogram. [1]

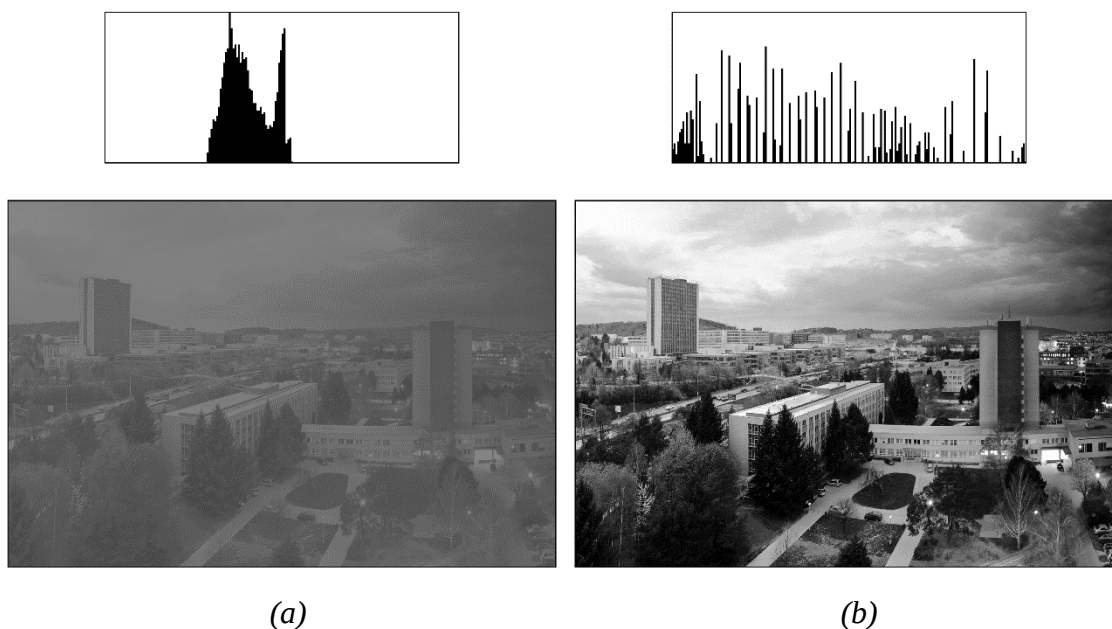
Ekvalizace histogramu je bodovou transformací, při které dojde ke změně rozložení jednotlivých jasových úrovní tak, aby zabíraly co největší rozsah a měly co nejpodobnější četnosti.

Transformace má za následek zvýšení kontrastu obrazu a používá se zejména pro normalizaci obrazu pro účely dalšího zpracování. Výpočet transformace popisuje následující rovnice [4]:

$$I' = \frac{I}{X \cdot Y} \sum_{i=I_0}^{i=I} H(i) \quad (1)$$

kde I' je nová hodnota jasu obrazového bodu,
 I je původní hodnota jasu obrazového bodu,
 I_0 je nejnižší hodnota jasu v původním obraze,
 $H(i)$ je četnost i -té jasové úrovně,
a X, Y jsou rozměry obrazu.

Efekt ekvalizace histogramu je znázorněn na následujícím obrázku:

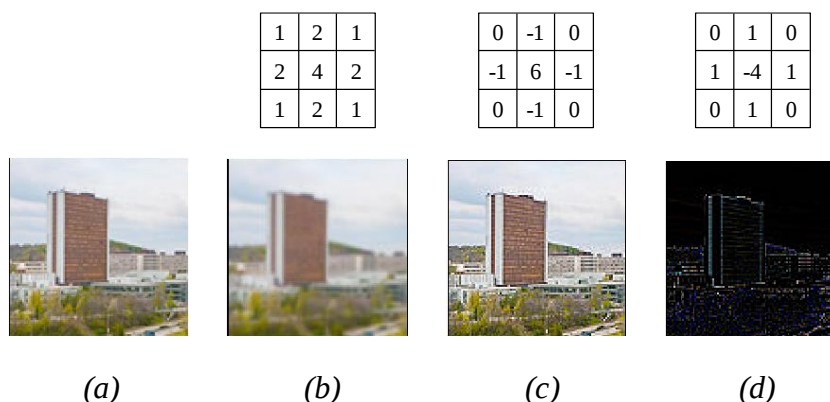


Obr. 6 Ekvalizace histogramu: (a) před ekvalizací, (b) po ekvalizaci.

2.4 Konvoluce

Konvoluce je operace, při které je počítána nová hodnota jasu obrazového bodu pomocí hodnot jasů bodů okolních. Váhy bodů jsou dány konvolučním jádrem, což je matice o lichém počtu řádků a sloupců, v jejímž středu leží váha aktuálního bodu. Výsledná hodnota středního bodu je tedy váženým průměrem hodnot všech bodů překrytých konvolučním jádrem.

Konvoluce se používá například pro rozostření nebo doostření obrazu, pro detekci hran a obecně pro aplikaci jakéhokoli filtru.



Obr. 7 Ukázka konvoluce: (a) původní obrázek, (b) rozostření, (c) doostření, (d) detekce hran.

2.5 Strojové učení

Strojové učení je oblast umělé inteligence zabývající se metodami, které umožňují počítačovým programům učit se. Využívají při tom ve velké míře hlavně statistiku.

Metody lze podle způsobu získávání znalostí rozdělit do několika skupin [5]:

- *Učení s učitelem*
Data, ze kterých se algoritmus učí, jsou předem ručně označena. Získané znalosti jsou poté využity při klasifikaci. Do této skupiny patří například umělé neuronové sítě nebo *Support Vector Machines*.
- *Učení bez učitele*
Na rozdíl od učení s učitelem jsou učicímu algoritmu poskytnuta neoznačená data a algoritmus nedostává žádnou zpětnou vazbu. Sám tedy musí být schopen v datech najít nějaké vzory.
- *Kombinace učení s učitelem a učení bez učitele*
Z množství dat, která jsou algoritmu poskytnuta, je označena jen malá část.

3 METODY DETEKCE OBLIČEJE

Lokalizace obličeje je prvním krokem při analýze jeho výrazu, proto je velice důležité, aby byl obličej lokalizován kvalitně a zároveň rychle. Důležitými faktory, které kvalitu a rychlost ovlivňují, jsou poloha a natočení obličeje a jeho osvětlení.

Pro detekci obličeje v barevném nebo černobílém obraze existuje řada metod, které je možno rozdělit do čtyř hlavních skupin [6]:

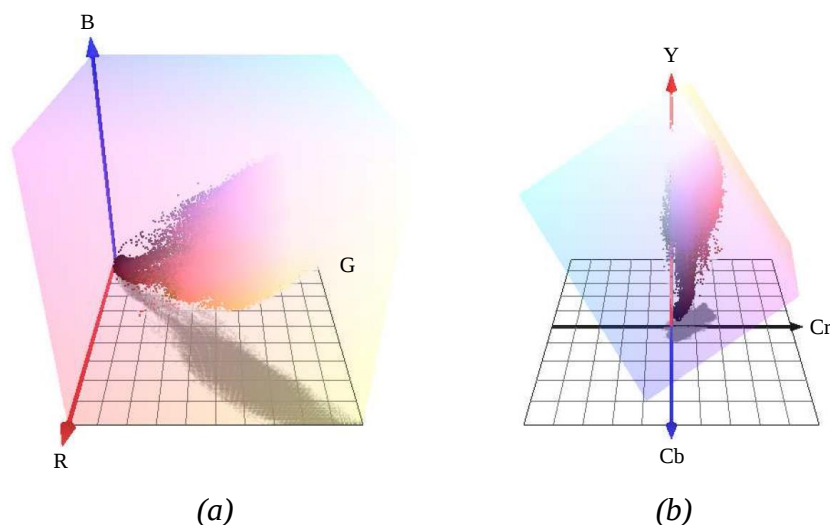
- *Znalostní metody*
Tyto metody hledají obličej podle pravidel, která pomocí vztahů mezi jednotlivými částmi obličeje definují typickou tvář. Mezi tato pravidla patří například fakt, že obličej má ve většině případů dvě symetricky umístěné oči, mezi nimi se nachází nos a pod ním jsou ústa. Nevýhodou těchto metod je složitost definice takových pravidel, pomocí kterých bude detekován jakýkoli obličej a zároveň nedojde k detekci objektů, které obličejem nejsou.
- *Metody invariantních rysů*
Tyto metody vycházejí z poznatku, že člověk je schopen detekovat tvář nezávisle na její poloze, úhlu pohledu a v jakémkoli osvětlení. Aby toto bylo možné, musí v obličejích existovat obecné rysy, které i při různém natočení tváře a rozdílných světelných podmínkách zůstávají neměnné. Těmito rysy jsou například oči, obočí, nos, ústa nebo barva kůže.
- *Metody porovnávání se vzory*
Obličej je lokalizován na základě podobnosti s ručně vytvořenou šablonou, která většinou obsahuje kontury tváře, očí, nosu a úst. V obraze je poté hledána oblast, která by mohla šabloně odpovídat. Tato oblast je se šablonou porovnávána a pro jednotlivé části předpokládané tváře jsou spočítány korelace s odpovídajícími částmi v šabloně. Podle vypočtených hodnot je rozhodnuto, zda se opravdu jedná o obličej či ne.
- *Metody založené na zjevu*
Na rozdíl od metod porovnávání se vzory pracují tyto metody se šablonami získanými učením z trénovací množiny. Detekce obličeje pak probíhá srovnáváním částí obrazu s těmito naučenými šablonami. Mezi metody založené na zjevu patří zejména detektor Viola-Jones.

3.1 Detekce obličeje podle barvy kůže

Tato metoda je založena na poznatku, že kůže si v různých světelných podmínkách zachovává svoje vlastnosti a v barevném prostoru zabírá určitý podprostor [2]. Prvním krokem detekce obličeje podle barvy kůže je volba vhodného barevného prostoru, ve kterém bude detekce probíhat.

Model RGB je pro tuto úlohu nevhodný, protože při změně osvětlení scény u něj dojde ke změně všech složek barvy, jelikož informaci o jasu nese každá z nich. Naopak velice vhodným se ukázal být model YCbCr, v němž je informace o barvě díky oddělené jasové složce velkou měrou nezávislá na různé intenzitě osvětlení nebo na přepalech či stínech. [2]

Druhým krokem je určení podprostoru, který v barevném modelu bude reprezentovat barvu kůže. Tuto informaci lze získat z množiny vzorků, která musí obsahovat co nejvíce typů kůže, aby byla detekce dostatečně robustní.



Obr. 8 Podprostor barvy kůže: (a) v modelu RGB, (b) v modelu YCbCr. [2]

Jednotlivé body obrazu jsou poté rozděleny do dvou skupin podle toho, zda svou barvou náleží do podprostoru barvy kůže, nebo ne. Oblasti v obraze, které barvě kůže vyhovují, jsou podrobeny další analýze a sérii operací, které mají za cíl tyto oblasti dále prověřit. Na konci tohoto procesu zůstane několik oblastí, které by mohly obsahovat obličej. V nich jsou pomocí červené barvy hledána ústa a pomocí bílé barvy oči. Pokud byly ústa i oči nalezeny, je oblast prohlášena za obličej. [2]

3.2 Detektor Viola-Jones

V roce 2001 Paul Viola a Michael Jones představili detektor objektů, který jako první dokázal lokalizovat obličej v reálném čase bez použití informací o rozdílech mezi jednotlivými snímky ve videu nebo o barvě jednotlivých pixelů. Zároveň se přesností detekce minimálně vyrovnal svým předchůdcům, které byly výpočetně mnohem náročnější. Vysoké rychlosti bylo dosaženo především díky využití Haarových vlnek a tří mechanismů, které tvoří kostru algoritmu: integrálního obrazu, adaptivního boostingu a skládání klasifikátorů do kaskády. [7]

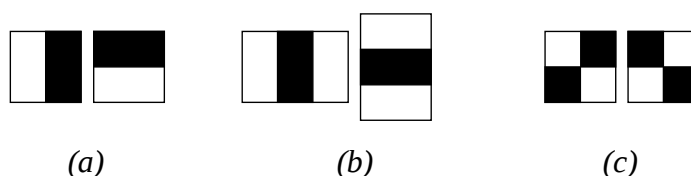
Algoritmus dokáže detekovat různé objekty, nejlépe se ale hodí pro detekci obličeje.

3.2.1 Haarovy vlnky

Detektor pro detekci obličeje používá jednoduché obdélníkové příznaky, které jsou v různých velikostech a poměrech stran přikládány k vstupnímu obrazu. Příznaky jsou podobné Haarovým vlnkám a nazývají se *Haar-like příznaky* (z anglického *Haar-like features*).

Příznaky, znázorněné na Obr. 9, se skládají z dvou, tří nebo čtyř obdélníků. Hodnota příznaku se vypočítá tak, že se od součtu hodnot pixelů v oblastech, které

odpovídají černým obdélníkům, odečte součet hodnot pixelů v oblastech odpovídajících bílým obdélníkům. [7]



Obr. 9 Haar-like příznaky: (a) hranové, (b) čárové, (c) diagonální.

3.2.2 Integrální obraz

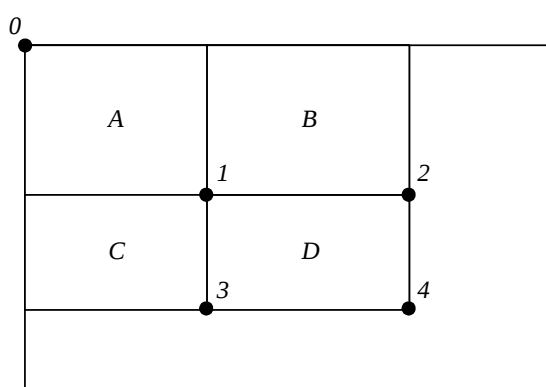
Pro zrychlení výpočtu hodnot jednotlivých oblastí se vstupní obraz převádí na tzv. *integrální obraz*, kde hodnota každého bodu odpovídá součtu hodnot všech bodů nad bodem a nalevo od bodu. Toho je dosaženo pomocí následujících dvou rovnic [7]:

$$s(x, y) = s(x, y - 1) + i(x, y) \quad (2)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \quad (3)$$

kde $ii(x, y)$ je integrální obraz,
 $i(x, y)$ je původní obraz,
 $s(x, y)$ je kumulativní součet hodnot v řádku,
 $s(x, -1) = 0$
a $ii(-1, y) = 0$.

Pro výpočet hodnoty jakékoli obdélníkové plochy z integrálního obrazu pak stačí pouze tři operace sčítání nebo odčítání a výpočet je proveden v konstantním čase.



Obr. 10 Příklad oblastí v integrálním obraze. [7]

Příklad výpočtu je naznačen na Obr. 10. Hodnota v bodě 1 odpovídá hodnotě oblasti A, hodnota v bodě 2 představuje součet hodnot oblastí A a B, hodnota v bodě 3 je součet oblastí A a C a hodnota v bodě 4 zahrnuje všechny čtyři oblasti. Hodnota oblasti D se tedy spočítá jako $4 + 1 - (2 + 3)$.

3.2.3 Adaptivní boosting

Detektor pracuje s obrazem daných rozměrů, standardně 24×24 pixelů, do kterého lze umístit přes 160 000 příznaků [7]. Protože vyhodnocování každého z nich by bylo časově velice náročné, bylo nutné nalézt klasifikační algoritmus, který z obrovského množství příznaků dokáže vybrat ty nejdůležitější, jež jsou nezbytně nutné k detekci obličeje.

Takovým algoritmem se stal AdaBoost (z anglického *Adaptive Boosting*), publikovaný v roce 1996 Yoavem Freundem a Robertem E. Schapirem. [8]

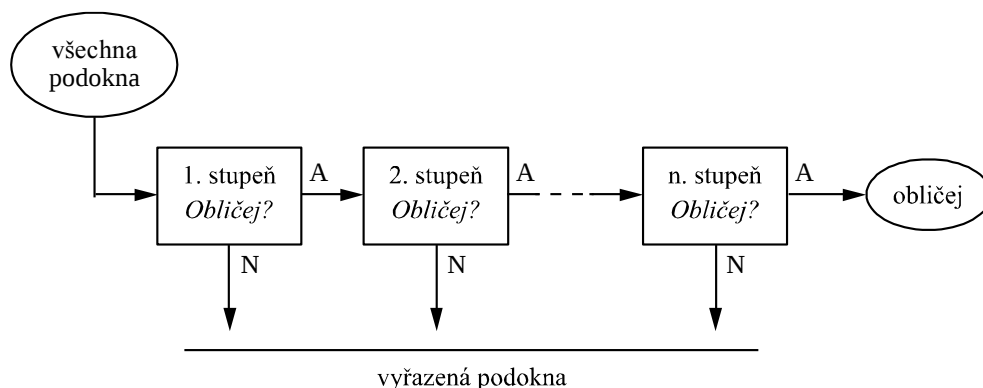
AdaBoost je schopen pomocí lineární kombinace velkého množství vážených slabých klasifikátorů vytvořit jeden silný klasifikátor. Slabé klasifikátory mají přesnost jen o málo větší, než je přesnost odhadu, což je v případě dvoustavového klasifikátoru 50 %. Na začátku jsou jednotlivým klasifikátorům přiděleny stejné váhy, ty jsou poté v dalších krocích upravovány tak, aby nejpřesnější klasifikátory měly největší váhu. [9]

Učení probíhá na vzorcích výše zmíněných rozměrů – pozitivních (obličeje) a negativních (jiné části těla, pozadí).

Detekce obličeje v obraze probíhá po částech, které jsou vybírány pomocí posuvného podokna, jehož základní velikost je stejná jako velikost vzorků použitých pro trénování. Po otestování celého obrazu se podokno zvětší a s ním se zvětší i všechny slabé klasifikátory. Zvětšení je dáno předem určeným koeficientem (např. 1,25 [7]). Následuje další průchod podokna obrazem. Postup se opakuje do té doby, než podokno obsáhne celý obraz.

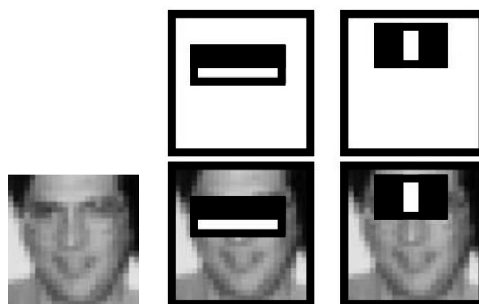
3.2.4 Kaskáda klasifikátorů

Protože podoken, která procházejí obraz, je také velké množství, je potřeba jejich vyhodnocování zefektivnit. Toho je dosaženo tak, že se klasifikátory zapojí do kaskády (Obr. 11). Do ní na začátku vstupují všechna podokna, do dalších stupňů ale postupují jen ta, o kterých klasifikátory v předchozích stupních rozhodly, že by se v nich mohl nacházet obličej. [7]



Obr. 11 Zapojení klasifikátorů do kaskády.

V prvním stupni kaskády je použit jednoduchý dvou-příznakový klasifikátor, který dokáže detekovat téměř 100 % obličejů a zároveň vyřadit zhruba 50 % negativních podoken.



Obr. 12 První dva příznaky vybrané AdaBoostem. První příznak měří rozdíl jasu mezi oblastí očí a horní částí tváře. Při tom vychází z poznatku, že oči jsou obvykle tmavší než tváře. Druhý příznak porovnává jas očí s jasnem nosu. [7]

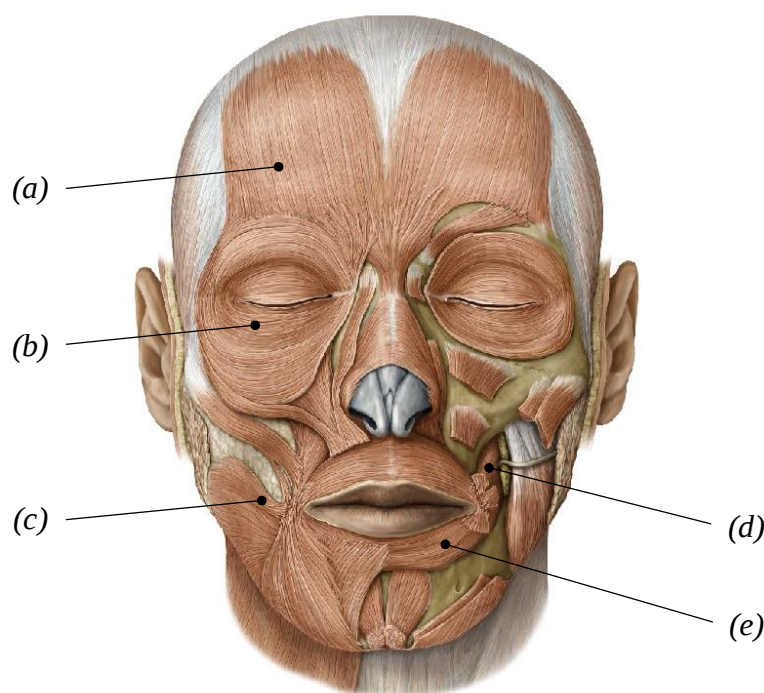
Za prvním klasifikátorem následuje klasifikátor s deseti příznaky, který opět detekuje téměř 100 % obličejů a vyřadí až 80 % negativních podoken. Klasifikátory v prvních stupních kaskády tedy mají za úkol vyřadit co největší množství negativních podoken. V dalších stupních pak následují složitější klasifikátory, které dále zpřesňují detekci obličeje. [7]

Na konci kaskády pak v ideálním případě zůstanou pouze ta podokna, která skutečně obsahují obličej.

4 VÝRAZY OBLIČEJE

Mimika obličeje je velice důležitou součástí neverbální komunikace. Používáme ji vědomě i nevědomě a odráží psychický stav člověka a jeho emoce. Výrazy tváře jsou z velké části univerzální a tudíž rozeznatelné napříč různými kulturami. Mimikou kromě člověka disponují i jiní živočichové, nejvíce je však vyvinutá u primátů. U člověka je pak díky plochému obličeji a nepřítomnosti srsti nejvýraznější. [10]

Mimiku zprostředkovává několik mimických svalů. Mezi ty hlavní patří sval čelní, který zdvihá obočí, kruhový sval oční, který ovládá oční víčka a okolí očí, sval smíchový, který táhnutím koutků do stran umožňuje úsměv, sval tvářový, který tlačí tvář k dásním a kruhový sval ústní, který svírá a špulí rty. [11]



Obr. 13 *Mimické svaly: (a) sval čelní, (b) kruhový sval oční, (c) sval smíchový (d) sval tvářový, (e) kruhový sval ústní.* [12]

Souvislost výrazů obličeje a mimických svalů, které je vytváří, popisuje systém FACS (*Facial Action Coding System*). Ten definuje tzv. akční jednotky, které se vážou na jednotlivé pohyby svalů. Všechny výrazy obličeje tak mohou být jednoznačně popsány kombinací akčních jednotek. [13]

Podle Paula Ekmana existuje sedm základních výrazů obličeje, které vyjadřují následující emoce a kterými se tato práce zabývá: hněv, opovržení, znechucení, strach, radost, smutek a překvapení. [14][15][16]

4.1 Hněv

Hněv (anglicky *anger*) je obecně reakcí na překážku, která jedinci brání v dosažení jeho cílů. Může se projevovat s různou intenzitou, hlavním znakem jakéhokoli hněvu jsou však stažená obočí a plně otevřené oči. Výrazné jsou i rty, které většinou

bývají pevně semknuté k sobě, nebo naopak otevřené do tvaru obdélníku, pokud se jedná o intenzivnější hněv.



Obr. 14 Hněv. [17]

4.2 Opovržení

Opovržení (anglicky *contempt*) vyjadřuje povýšenost či nadřazenost jedince nad ostatními nebo nesouhlas s činy druhých. Projevuje se na jedné straně lehce zdviženým koutkem úst. Zdvižen může být i horní ret, který odhalí zuby. Mírný úsměv výrazu dodá na síle. Opovržení je mnohdy doprovázeno pozvednutím hlavy, čímž je ještě více podpořena ona povýšenost.



Obr. 15 Opovržení. [17]

4.3 Znechucení

Výraz znechucení (anglicky *disgust*) je reakcí na cokoli nepříjemného, ať už se jedná o pach, chuť či o jiného člověka nebo jen myšlenku. Projevuje se zvednutím horního rtu a vráskami na horní části nosu mezi očima.



Obr. 16 Znechucení. [17]

4.4 Strach

Strach (anglicky *fear*) je všeobecně reakcí na nebezpečí a tělo se při něm připravuje na útěk. Na výrazu obličeje se strach projevuje zvednutím a zakřivením obočí, plně otevřenými očima a do šířky roztaženými ústy, která mohou být zavřená i otevřená.



Obr. 17 Strach. [17]

4.5 Radost

Radost (anglicky *joy*) je nejpozitivnější emocí vůbec. Jejím nejviditelnějším znakem je úsměv. Důležitá je ale i horní část obličeje, podle které se pozná pravý úsměv. V ní dochází ke kontrakci svalů kolem očí, které stáhnou obočí a zvednou tváře.



Obr. 18 Radost. [17]

4.6 Smutek

Smutek (anglicky *sadness*) je opakem radosti a většinou doprovází ztrátu někoho nebo něčeho významného. Smutek se projevuje zvednutím vnitřních konců obočí, zúženými ústy s propadlými koutky. Horní víčka očí mohou být přivřená.



Obr. 19 Smutek. [17]

4.7 Překvapení

Překvapení (anglicky *surprise*) nastává při prožití nějaké nečekané události a vyznačuje se zdviženým obočím, široce otevřenýma očima a mnohdy i otevřenými ústy.

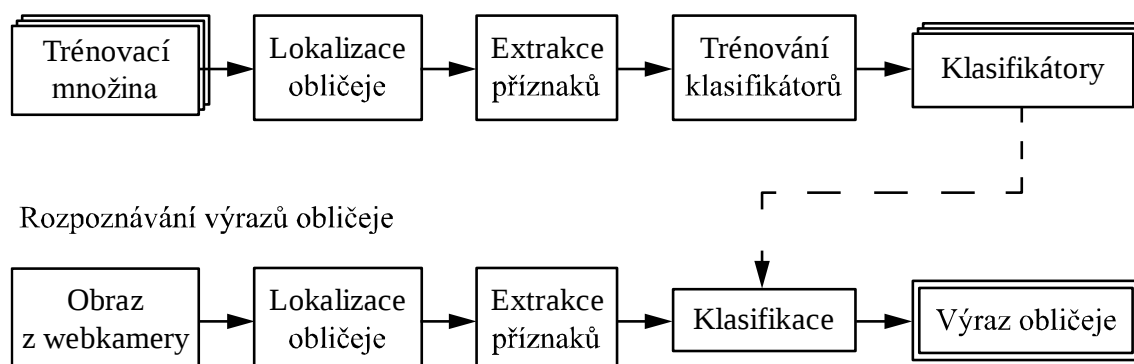


Obr. 20 Překvapení. [17]

5 NÁVRH APLIKACE

Aplikace pro rozpoznávání výrazů obličeje se skládá ze dvou částí. V první části je potřeba aplikaci naučit, jak jednotlivé výrazy vypadají, v druhé části jsou tyto vědomosti použity pro odhadování výrazů v obraze z webkamery.

Trénování klasifikátorů



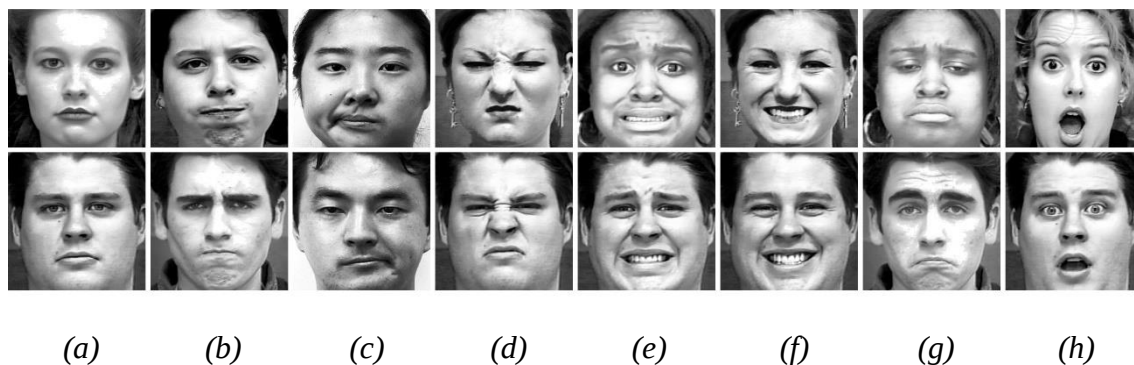
Obr. 21 Blokové schéma aplikace.

5.1 Trénování klasifikátorů

Trénování je proces, při kterém je klasifikátoru předložena množina pozitivních a negativních vzorků a klasifikátor se snaží najít optimální způsob, kterým se tyto vzorky dají rozlišit. Vzorky pocházejí z tzv. trénovací množiny a před tím, než postoupí ke klasifikátorům, musí být vhodným způsobem zpracovány.

5.1.1 Trénovací množina

Jako trénovací množina byla vybrána databáze obličejových výrazů Cohn-Kanade, konkrétně její rozšířená verze CK+ [17]. Ta obsahuje 593 sekvencí od 123 účastníků projektu. Každá sekvence začíná neutrálním výrazem a končí jedním ze sedmi základních výrazů popsanych v kapitole 4.



Obr. 22 Příklady obličejových výrazů z databáze CK+ [17]: (a) neutrální výraz, (b) zlost, (c) opovržení, (d) znechucení, (e) strach, (f) radost, (g) smutek, (h) překvapení.

Pro učení výrazů bylo použito 321 sekvencí, u kterých je v databázi označen zachycený výraz. Ostatní sekvence použity nebyly, protože by mohlo dojít ke špatnému určení výrazu, což by vedlo ke zkreslení výsledků. Složení trénovací množiny je zobrazeno v Tab. 1.

Tab. 1 Složení trénovací množiny.

Výraz	Počet vzorků
Neutrální	321
Hněv	44
Opovržení	17
Znechucení	58
Strach	24
Radost	68
Smutek	28
Překvapení	82
<i>Celkem</i>	<i>642</i>

Databáze je určena pouze k výzkumným účelům a není povoleno její další šíření.

5.1.2 Lokalizace obličeje

Obličej je lokalizován detektorem Viola-Jones, vyříznut z obrazu, zmenšen na velikost 48×48 pixelů a je u něj provedena ekvalizace histogramu. Takto zpracovaný obličej je připraven k extrakci příznaků.

5.1.3 Extrakce příznaků

Pro extrakci příznaků je využit Gaborův filtr, lineární filtr používaný pro detekci hran. Ten je pro tuto úlohu velice vhodný, protože je svou podstatou podobný systému zpracování zraku v lidském mozku [18]. Filtr je založen na Gaborových vlnkách, které jsou tvořeny komplexní sinusoidou v obálce z Gaussovy funkce a jsou definovány následující rovnicí [19]:

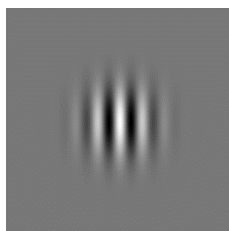
$$G(x, y, \lambda, \theta, \psi, \sigma, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi \frac{x'}{\lambda} + \psi\right)\right) \quad (4)$$

kde $x' = x \cos \theta + y \sin \theta$,
 $y' = -x \sin \theta + y \cos \theta$,
 x, y jsou souřadnice bodu v rovině,
 λ je vlnová délka sinusoidy v pixelech
 θ je natočení v radiánech,
 ψ je fázový posun v radiánech,
 σ je rozptyl Gaussovy funkce
a γ je eliptičnost.

Parametry σ a λ jsou na sobě závislé, jejich poměr lze vyjádřit následujícím vztahem [19]:

$$\frac{\sigma}{\lambda} = \frac{1}{\pi} \cdot \sqrt{\frac{\ln 2}{2}} \cdot \frac{2^b + 1}{2^b - 1} \quad (5)$$

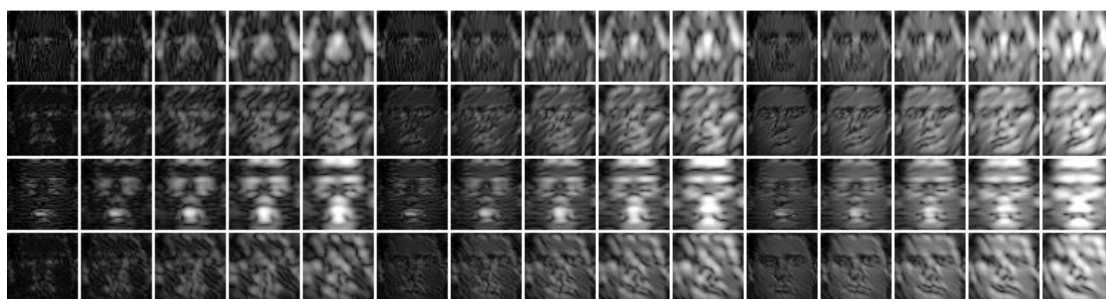
kde b je šířka pásma filtru v oktávách.



Obr. 23 Gaborova vlnka.

Za účelem extrakce příznaků je vygenerován soubor Gaborových vlnek s proměnlivou šířkou pásma b , vlnovou délkou λ a natočením θ . Během testování podával dobré výsledky soubor tří hodnot šířky pásma b v intervalu od 1,1 do 1,5, pěti hodnot vlnové délky λ v intervalu od 3,1 do 5,5 a čtyř hodnot natočení θ v intervalu od 0 do $3\pi/4$. Pro fázový posun ψ byla zvolena hodnota $\pi/2$ a pro eliptičnost γ hodnota 1.

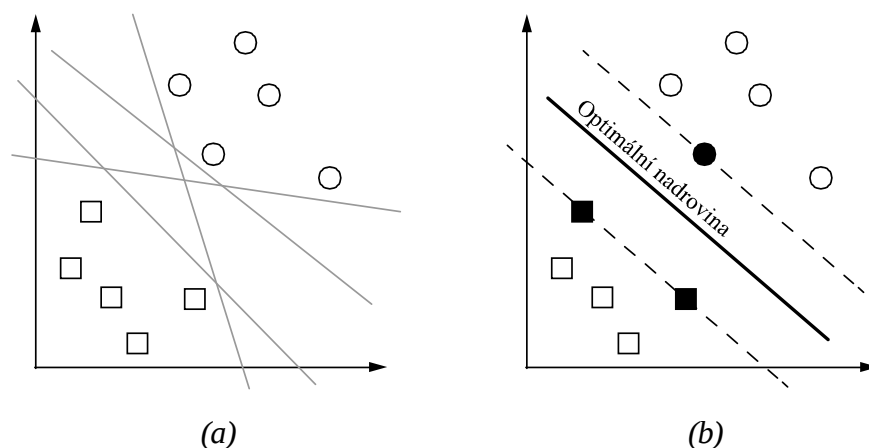
Tento soubor je poté konvolucí aplikován na obličej. Výsledný soubor příznaků je zobrazen na následujícím obrázku:



Obr. 24 Soubor příznaků po aplikování Gaborova filtru s proměnlivou šířkou pásma, vlnovou délkou a natočením.

5.1.4 Trénování klasifikátorů

V této práci byly pro roli klasifikátoru vybrány lineární *Support Vector Machines* (SVM). SVM je binární klasifikátor, který v prostoru trénovacích dat hledá nadrovinu, jež je optimálně rozděluje tak, že na opačných stranách od nadroviny leží vzorky opačných tříd, které jsou od ní co nejvíce vzdáleny. Vzorky, které jsou nadrovině nejbližší, se nazývají *podpůrné vektory*. [20][21]



Obr. 25 Hledání optimální nadroviny: (a) možné nadroviny, (b) optimální nadrovina. [20]

Jelikož binární klasifikátor rozlišuje pouze dvě třídy, pozitivní a negativní, musela být zvolena strategie, která umožní klasifikovat vzorky do K tříd.

Nejjednodušší, ale velice efektivní, strategií je tzv. strategie *One-versus-All* (jeden versus všechny), která spočívá ve vytvoření K binárních klasifikátorů, z nichž každý k -tý je trénován pozitivními vzorky patřícími do k -té třídy a negativními vzorky patřícími do ostatních tříd. [22]

Na konci tohoto procesu existuje pro každý výraz obličeje jeden natrénovaný klasifikátor.

5.2 Rozpoznávání výrazů obličeje

Druhá část aplikace se zabývá samotným rozpoznáváním výrazu obličeje, přičemž první dvě fáze jsou stejné jako při trénování klasifikátorů. Ve vstupním obrazu jsou opět lokalizovány obličeje a z nich jsou stejným způsobem extrahovány příznaky. Ty jsou předány natrénovaným SVM klasifikátorům.

5.2.1 Klasifikace

Příznaky jsou postupně zkoumány všemi klasifikátory. Třída, která je příznakům příslušným klasifikátorem přiřazena s nejvyšší mírou jistoty, je prohlášena za správnou. Pokud příznakům není během klasifikace přiřazena žádná třída, vyhodnotí se výraz obličeje jako neznámý.

U SVM klasifikátorů lze za míru jistoty považovat vzdálenost příznaku od nadroviny. Čím větší tato vzdálenost je, tím si je klasifikátor „jistější“, že příznaky do příslušné třídy patří.

6 IMPLEMENTACE

Aplikace Facepression (složení slov *Face* a *Expression*) je vytvořena v jazyce C++ ve vývojovém prostředí Microsoft Visual Studio Express 2013.

Řešení je rozděleno do tří projektů. První dva projekty implementují trénování klasifikátorů a rozpoznávání výrazů obličeje, třetí projekt tvoří statickou knihovnu, která obsahuje funkce sdílené mezi projekty.

6.1 Knihovna OpenCV

Pro úlohy zpracování obrazu a strojového učení byla využita knihovna OpenCV verze 2.4.9 [23]. OpenCV je open source knihovna pro počítačové vidění se zaměřením na zpracování obrazu v reálném čase. Je napsána v jazyce C/C++, disponuje ale i rozhraními pro Python, Javu nebo Matlab. V současné době je ve vývoji verze 3.0, která v dubnu 2015 pokročila do fáze RC.

Knihovna je rozdělena na několik modulů, z nichž byly využity následující:

- *core*
Jádro knihovny, definuje základní datové struktury, zejména vícerozměrné pole *Mat*, a základní funkce používané ostatními moduly.
- *imgproc*
Modul pro zpracování obrazu, obsahuje funkce pro filtrování, převody mezi barevnými prostory, práci s histogramem, geometrické transformace apod.
- *highgui*
Modul obsahující základní prvky uživatelského rozhraní a nástroje pro načítání, zobrazování a ukládání obrázků a videa.
- *objdetect*
Modul s funkcemi pro detekci objektů, implementuje například detektor Viola-Jones. S knihovnou jsou dodávány natrénované kaskády klasifikátorů pro detekci obličeje, očí, úst a dalších objektů.
- *ml*
Modul, který implementuje několik metod strojového učení včetně SVM klasifikátoru.

Knihovna byla nainstalována podle návodu v [24].

6.2 Struktura dat

Veškerá data, s kterými aplikace pracuje, jsou umístěny ve složce *data*, jež se nachází v hlavní složce řešení, tedy na stejné úrovni jako složky projektů. Složka obsahuje podsložky, které slouží jako úložiště výstupů jednotlivých fází trénování klasifikátorů, snímků obrazovky apod.:

- *cascades*
Kaskády klasifikátorů pro detekci obličeje (jsou součástí knihovny OpenCV).
- *screenshots*
Snímky obrazovky pořízené při rozpoznávání výrazů.

- *original*
Snímky trénovací množiny rozdělené do podsložek podle jednotlivých výrazů (*neutral, anger, contempt, disgust, fear, joy, sadness, surprise*). Každá z těchto podsložek musí obsahovat soubor *list.txt* se seznamem souborů (jeden soubor na řádek).
- *faces*
Oříznuté obličeje, opět rozdělené do podsložek.
- *features*
Příznaky extrahované z obličejů, opět rozdělené do podsložek.
- *training*
Soubory ve formátu CSV, které obsahují seznam souborů s příznaky a určení, zda se jedná o pozitivní či negativní vzorek. Pro každý výraz včetně neutrálního je vytvořen jeden soubor.
- *classifiers*
Natrénované klasifikátory.
- *test*
Snímky testovací množiny rozdělené do podsložek, které musí obsahovat soubor *list.txt*.

6.3 Sdílené funkce

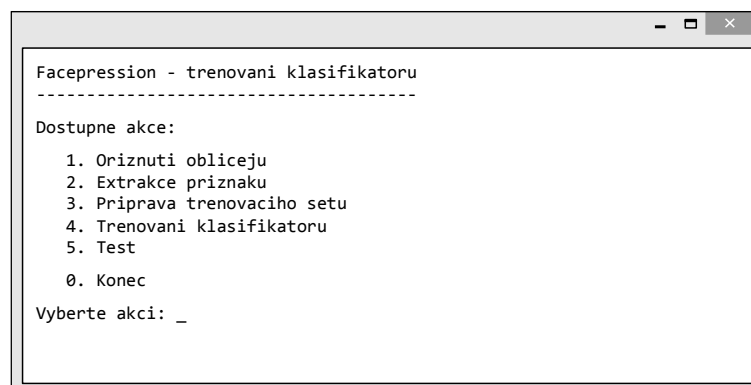
Sdílené funkce se nachází v projektu s názvem *functions*. Ten čtyři moduly:

- *Settings*
Definuje globální proměnné, ve kterých jsou nastaveny názvy obličejových výrazů, parametry Gaborových vlnek, parametry SVM klasifikátorů a názvy datových složek.
- *FaceProcessing*
Obsahuje funkci `findFace` pro lokalizaci a oříznutí obličeje pomocí detektoru Viola-Jones.
- *GaborFilter*
Obsahuje funkci `getGaborKernelReIm`, která pro zadané parametry vytvoří z reálné a imaginární části Gaborovy vlnky konvoluční jádra, funkci `fillGaborBank`, která konvolučními jádry s různými parametry naplní soubor filtrů, a konečně funkci `gaborFilter`, která tento soubor filtrů aplikuje na obličej a vytvoří tak soubor příznaků.
- *Classifier*
Obsahuje funkci `createClassifiers`, která pro každý obličejový výraz vytvoří SVM klasifikátor, a funkci `predictExpression`, která pomocí těchto klasifikátorů z předložených příznaků odhadne výraz obličeje.

6.4 Trénování klasifikátorů

Trénování klasifikátorů je implementováno v projektu *FacepressionTraining*. Aby během testování různých parametrů nebylo nutné provádět operace, u kterých se nic

nezměnilo, je proces v aplikaci rozdělen na několik fází. Výběr akce obsluhuje funkce `getAction`, jednotlivé akce se nachází v modulu *Actions*.



Obr. 26 Volba akce po spuštění trénovací aplikace.

6.4.1 Oříznutí obličejů

V první fázi se postupně ve všech obrázcích databáze obličejových výrazů lokalizují a oříznou obličej.

Nejprve je načtena kaskáda klasifikátorů (*haarcascade_frontalface_default.xml*), poté jsou procházeny všechny snímky databáze ve složce *original*, které jsou spolu s načtenou kaskádou klasifikátorů předány funkci `findFace`. Ta snímek převede do stupňů šedi, ekvalizuje histogram, pomocí metody `detectMultiScale` z knihovny OpenCV na snímku lokalizuje největší obličej, který je poté oříznut, zmenšen a je u něj opět provedena ekvalizace histogramu. Výsledný obrázek o velikosti 48×48 pixelů je nakonec uložen na své místo do složky *faces*.

6.4.2 Extrakce příznaků

Na začátku této fáze je vytvořen vektor párů matic, který je pomocí funkce `fillGaborBank` naplněn reálnými a imaginárními konvolučními jádry pro aplikaci Gaborova filtru. Poté jsou procházeny všechny obličejové ve složce *faces* a spolu s vektorem jader jsou předány funkci `gaborFilter`. Ta na obličej postupně aplikuje všechny verze filtru a vrátí soubor příznaků. Soubor je pak ve formátu YAML uložen na své místo do složky *features*.

6.4.3 Příprava trénovacího setu

V této fázi je pro každý obličejový výraz připraven seznam vzorků, z kterého poté čerpá klasifikátor. Seznamy jsou uloženy ve složce *training* v souborech ve formátu CSV. Každý řádek obsahuje název souboru příznaků, určení, zda se jedná o pozitivní (1) či negativní (0) vzorek, a název výrazu obličejové, který vzorek představuje.

321	S999_003_00000001.yml	0	neutral
322	S010_004_00000019.yml	1	anger

Obr. 27 Příklad dvou řádků souboru *anger.csv*.

Díky strategii *One-versus-All* mají seznamy totožný obsah, liší se pouze označením pozitivních a negativních vzorků, tedy prostředním sloupcem.

6.4.4 Trénování klasifikátorů

Poslední fází trénovacího procesu je samotné trénování klasifikátoru. Na počátku této fáze je nejdříve načten první soubor příznaků, z něhož jsou zjištěny rozměry. Ty jsou použity pro vytvoření matice, do které se v dalším kroku postupně načtou všechny soubory příznaků.

Po načtení všech vzorků se pro každý obličejový výraz načtou jejich označení a vytvoří se SVM klasifikátor. Ten je v knihovně OpenCV implementován v objektu `CvSVM`. Klasifikátoru se předá matice vzorků a matice označení a je spuštěno jeho trénování. Jakmile je trénování dokončeno, uloží se klasifikátor ve formátu XML do složky *classifiers*.

6.4.5 Test

Aplikace umožňuje provést test klasifikace na poskytnuté testovací množině, která se nachází ve složce *test*. Výsledky jsou uloženy do souboru *results.csv*. Každý řádek obsahuje název testovaného snímku, jeho skutečný výraz, klasifikovaný výraz a vzdálenost vzorku od dělící nadrovině.

57	S053_003_00000046.png	anger	anger	-1.05233
58	S110_006_00000027.png	anger	unknown	0

Obr. 28 Příklad dvou řádků souboru *results.csv*.

6.5 Rozpoznávání výrazů obličeje

Rozpoznávání výrazů obličeje je implementováno v projektu *Facepression*. Bezprostředně po spuštění aplikace je pomocí objektu `VideoCapture` knihovny OpenCV inicializována webkamera. Následuje načtení kaskády klasifikátorů pro detekci obličeje. Ta musí být stejná jako ta, která byla použita při trénování klasifikátorů. Dále je vytvořen a funkcí `fillGaborBank` naplněn vektor reálných a imaginárních konvolučních jader Gaborova filtru a vektor SVM klasifikátorů, který je naplněn funkcí `createClassifiers`.

Následuje nekonečný cyklus `while`, v jehož každé iteraci je načten a analyzován snímek z webkamery. Nejprve se v něm funkcí `findFace` lokalizuje obličej. Pokud je obličej ve snímku nalezen, ořízne se a zmenší. Proces pokračuje extrakcí příznaků funkcí `gaborFilter`.

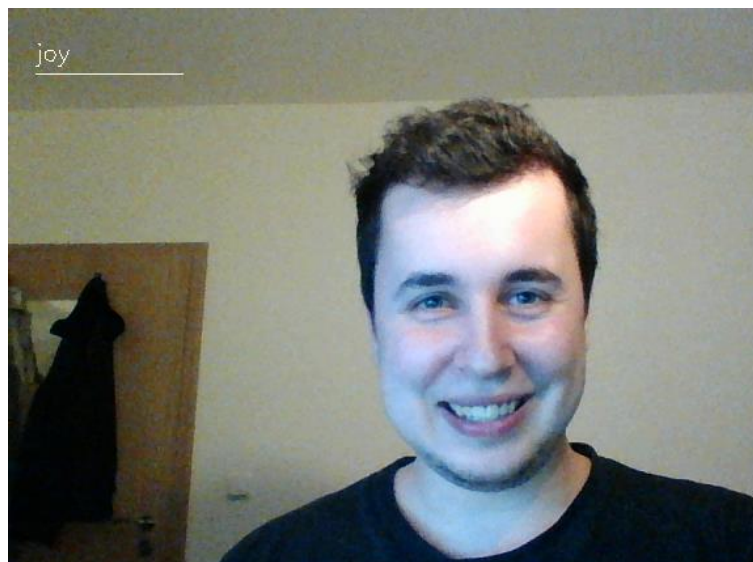
Příznaky jsou předány funkci `predictExpression`, která je postupně předloží všem klasifikátorům a pro každý z nich si do vektoru výsledků uloží vzdálenost příznaků od nadrovině.

Pokud jsou příznaky klasifikovány jako pozitivní, je vzdálenost od nadrovině záporná. Čím je tato vzdálenost větší (více záporná), tím si je klasifikátor „jistější“. Jako výsledek je zvolen výraz, pro který klasifikátor určil nejvíce zápornou vzdálenost.

Pokud žádný klasifikátor nedokázal příznaky zařadit do pozitivní třídy (vzdálenosti byly větší nebo rovny nule), je vzdálenost nastavena na nulu a výraz obličeje je považován za neznámý (anglicky *unknown*).

Funkce vrací výsledek ve tvaru uspořádané dvojice [*název_výrazu_obličeje*, *vzdálenost_od_nadroviny*].

Název výrazu je vložen do levého horního okraje snímku a je k němu přidána grafická reprezentace jistoty klasifikace v podobě vodorovné linky, jejíž délka je určena vzdáleností od nadroviny. Snímek je poté zobrazen v okně aplikace.



Obr. 29 Příklad správně určeného výrazu obličeje. Vlevo nahoře se nachází jeho název a grafická reprezentace míry jistoty.

Při stisknutí klávesy S je pořízen snímek aktuálního obrazu, který je uložen do složky *screenshots*. Aplikace je ukončena stisknutím klávesy Escape.

6.6 Výsledky

Rozpoznávání výrazů obličeje v obraze z webkamery pracuje v reálném čase. Analýza jednoho snímku zabere průměrně 120 milisekund (konfigurace s procesorem Intel Core i7-4710HQ), čímž je dosaženo zhruba osmi snímků za sekundu. Tato rychlosti je umožněna zejména díky tomu, že jsou obličeje před aplikací Gaborova filtru nejdříve zmenšeny.

Klasifikace byla testována především prostřednictvím webkamery, využita byla ale i možnost testování statických snímků (kapitola 6.4.5).

6.6.1 Statické snímky

Během vývoje aplikace byla k dispozici pouze obličejová databáze Cohn-Kanade [17], z které byly do testovací množiny použity sekvence, jež nebyly označené druhem výrazu. Většina výrazů v těchto sekvencích však není zcela jednoznačná, tudíž

byly vybrány jen ty nejvýraznější, které byly manuálně označeny. Výsledky jsou tedy spíše jen orientační.

Tab. 2 Úspěšnost klasifikace na testovací množině.

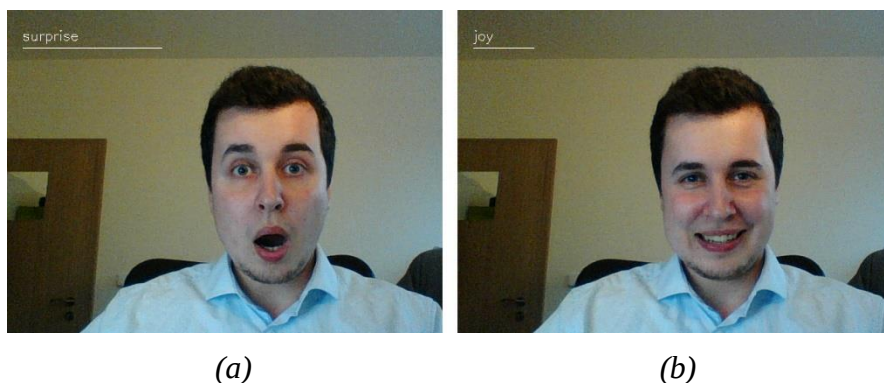
Výraz	Počet vzorků	Správně klasifikovaných	Úspěšnost
Neutrální	39	39	100 %
Hněv	6	3	50 %
Opovržení	1	0	0 %
Znechucení	7	6	86 %
Strach	15	5	33 %
Radost	11	10	91 %
Smutek	6	4	67 %
Překvapení	9	7	78 %

Neutrální výraz je obecně rozpoznáván velmi dobře, nicméně stoprocentní úspěšnost je způsobena z velké části tím, že testované neutrální obličeje patří téměř ve všech případech subjektům, jejichž obličeje byly použity pro trénování klasifikátoru neutrálního výrazu. Úspěšnost ostatních výrazů zhruba odpovídá jejich úspěšnosti při klasifikaci webkamerou. Pro výraz opovržení byl nalezen pouze jeden snímek, který byl určen jako neznámý. Úspěšnost rozpoznání tohoto výrazu obecně není velká.

6.6.2 Webkamera

Při testování rozpoznávání výrazu webkamerou bylo bez potíží dosaženo všech obličejových výrazů.

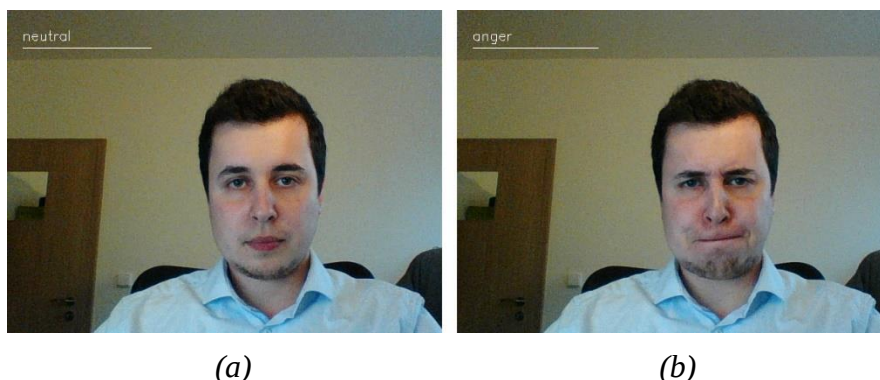
Výrazy překvapení a radosti patří mezi nejúspěšnější a k jejich rozpoznání dochází okamžitě. Důvodem je skutečnost, že jsou velice výrazné a nezaměnitelné.



Obr. 30 Správně klasifikované výrazy: (a) překvapení, (b) radost.

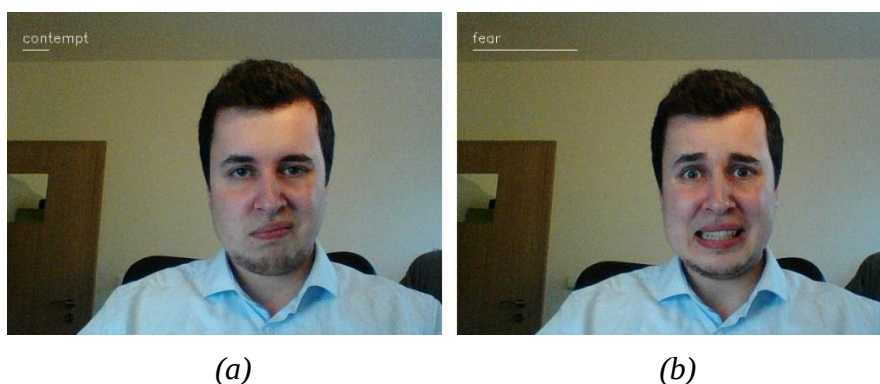
Mezi velmi dobře rozpoznávané patří i neutrální výraz a výraz hněvu. Mezi neutrálním výrazem a výrazem hněvu je úzká hranice a mírná forma hněvu občas může

být zaměněna právě za výraz neutrální. Intenzivnější hněv, při kterém dojde k „nafouknutí“ tváří, je ale téměř vždy klasifikován správně.



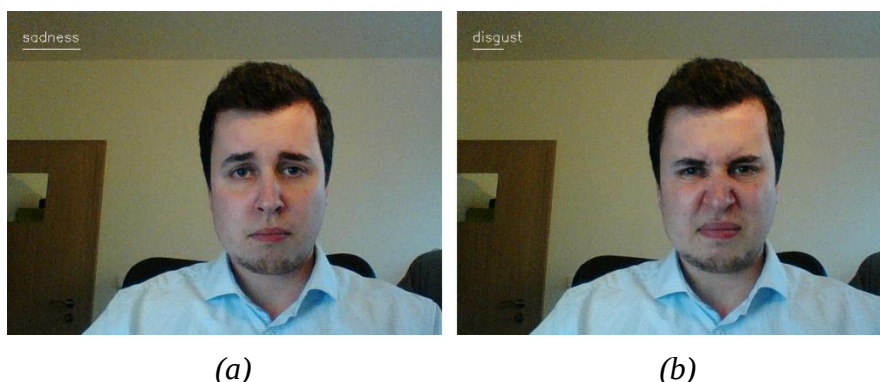
Obr. 31 Správně klasifikované výrazy: (a) neutrální, (b) hněv.

Mezi další dobře rozpoznávané výrazy patří opovržení a strach. Opovržení je hodně podobné neutrálnímu výrazu, zvednutý koutek úst ale ve většině případů postačuje k správné klasifikaci. Strach bývá rozpoznán bez větších problémů.



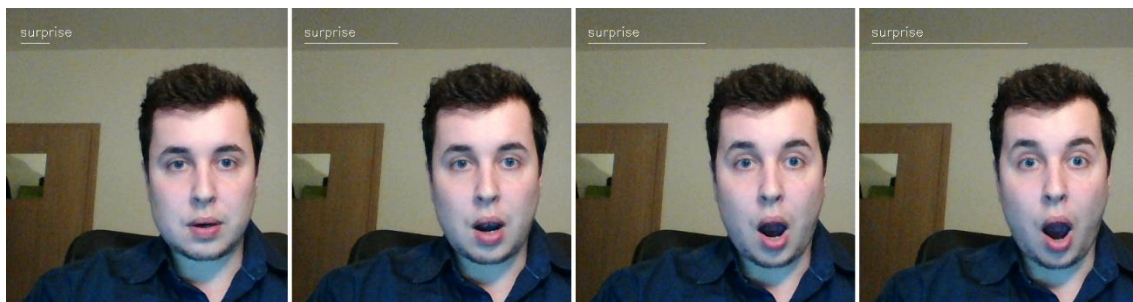
Obr. 32 Správně klasifikované výrazy: (a) opovržení, (b) strach.

Nejhůře rozpoznávané výrazy jsou smutek a znechucení. U smutku je to způsobeno tím, že je velice obtížné ho správně napodobit [15] a je proto často zaměňován za neutrální výraz. Znechucení je někdy považováno za hněv, v množství případů je ale interpretováno jako neznámý výraz.



Obr. 33 Správně klasifikované výrazy: (a) smutek, (b) znechucení.

Některé výrazy, jako například překvapení, jsou detekovány už v jejich počátku a s přibývajícím intenzitou se pouze zvyšuje jistota klasifikace:



Obr. 34 Fáze překvapení.

U osob ženského pohlaví podává rozpoznávání stejné výsledky jako u osob pohlaví mužského:



Obr. 35 Další příklady správně klasifikovaných výrazů: (a) neutrální, (b) hněv, (c) opovržení, (d) radost, (e) překvapení, (f) smutek.

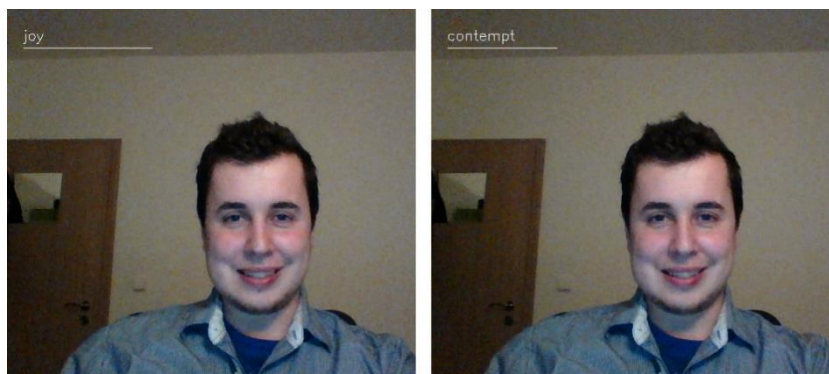
Výraznější mimika je úspěšně identifikována i u osob se slunečními brýlemi:



Obr. 36 Klasifikace výrazů u osob se slunečními brýlemi: (a) hněv, (b) radost, (c) strach, (d) překvapení.

6.6.3 Vliv šumu

Pokud obraz obsahuje nezanedbatelné množství šumu, stává se klasifikace značně nestabilní. Stejný výraz pak může být v každém snímku klasifikován jinak:



(a)

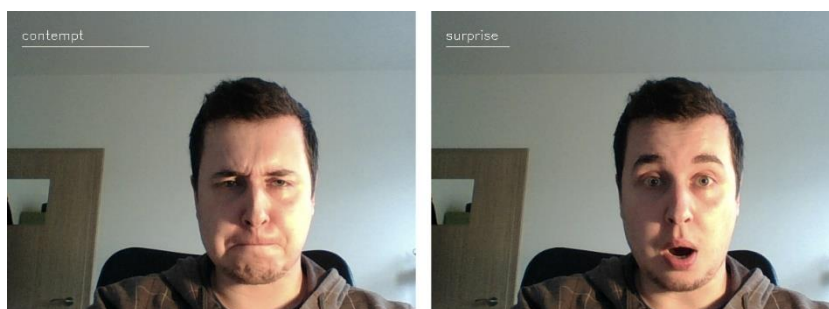
(b)

Obr. 37 Vliv šumu: (a) správně klasifikovaná radost, (b) špatně klasifikované opovržení.

6.6.4 Závislost na osvětlení

Při rozpoznávání výrazu obličeje hodně záleží na jeho osvětlení. Nejlepších výsledků je dosaženo, když je zdroj světla umístěn před obličejem nebo když je obličej rovnoměrně osvětlen ze všech stran. Ve všech případech je vhodnější měkké světlo.

Boční světlo je nevhodné, protože v obličejí vytváří stíny, které mají u většiny výrazů matoucí účinek. Nejodolnější vůči osvětlení je výraz překvapení.



(a)

(b)

Obr. 38 Boční osvětlení: (a) hněv klasifikovaný jako opovržení, (b) překvapení.

7 ZÁVĚR

Úkolem této práce bylo navrhnout a s pomocí knihovny OpenCV vytvořit aplikaci pro rozpoznání výrazu obličeje.

Prvním krokem je lokalizace tváře. Pro tuto úlohu aplikace využívá detektor Viola-Jones. Z oříznutých obličejů jsou pomocí Gaborových vlnek extrahovány příznaky, které jsou předloženy skupině SVM klasifikátorů. Pro každý obličejový výraz existuje jeden klasifikátor. Jednotlivé klasifikátory rozhodnou, zda příznaky přísluší výrazu, pro který byly natrénovány. Výraz, který byl klasifikován s nejvyšší mírou jistoty, je prohlášen za správný a jeho název a grafická reprezentace míry jistoty jsou zobrazeny uživateli.

Aplikace dokáže rozpoznat neutrální výraz a všech sedm základních výrazů, kterými jsou hněv, opovržení, znechucení, strach, radost, smutek a překvapení.

Nejúspěšněji klasifikovaný je výraz překvapení, jehož detekce není znemožněna ani silným bočním světlem. Velmi vysoké úspěšnosti dosahuje i klasifikace výrazů radosti, hněvu a neutrální výraz. Horší úspěšnost byla pozorována u výrazů smutku a znechucení, kdy docházelo buď k záměně za jiné výrazy, nebo výrazy nebyly rozpoznány vůbec.

Úspěšnost klasifikace hodně záleží na osvětlení obličeje. Nejvhodnější je osvětlení zepředu. Při bočním světle jsou výrazy často interpretovány jako opovržení nebo nejsou rozeznány vůbec. Jistý vliv na detekci má také šum, který se projevuje nestabilitou klasifikace neměnných výrazů.

Rozpoznávání výrazů obličeje v obraze z webkamery dosahuje rychlosti osmi snímků za sekundu, což je považováno za dostatečné.

Aplikace může být využita například při testování programů nebo webových stránek, kdy kromě pohybu očí budou zaznamenány i uživatelské emoce. To se týká i testování her nebo zkušebních promítání filmů. Informace o emocích mohou být využity také v komunikaci dvou osob prostřednictvím chatu, kde mohou nahradit emotikony. Další možností využití aplikace by mohla být analýza výrazů studentů při reakcích na výklad nové látky. Upravená verze aplikace by mohla být využita v automobilech pro odhalení únavy řidiče.

SEZNAM POUŽITÉ LITERATURY

- [1] HLAVÁČ, Václav a Miloš SEDLÁČEK. *Zpracování signálů a obrazů*. 1. vyd. Praha: Vydavatelství ČVUT, 2000, 220 s. ISBN 80-010-2114-9.
- [2] VLACH, Jan a Jiří PŘINOSIL. Lokalizace obličejů v obraze s komplexním pozadím. *Elektrorevue* [online]. 2007, **2007**(12) [cit. 2015-04-28]. ISSN 1213-1539. Dostupné z: <http://www.elektrorevue.cz/cz/clanky/zpracovani-signalu/45/lokalizace-obliceje-v-obraze-s-komplexnim/>
- [3] KALOVÁ, Ilona. *Předzpracování obrazu* [online]. 2013 [cit. 2015-04-03]. Dostupné z: http://midas.uamt.feec.vutbr.cz/POV/lectures-pdf/02_Predzpracovani_obrazu.pdf
- [4] PAVELEK, Milan, Eva JANOTKOVÁ a Josef ŠTĚTINA. *Vizualizační a optické měřicí metody* [online]. Druhé vydání. Brno, 2007 [cit. 2015-04-28]. Dostupné z: <http://ottp.fme.vutbr.cz/~pavelek/optika/>
- [5] JOSHI, Ajay, Anoop CHERIAN a Ravishankar SHIVALINGAM. *Machine Learning in Computer Vision: A Tutorial* [online]. 2009 [cit. 2015-05-20]. Dostupné z: <http://www-users.cs.umn.edu/~cherian/ppt/MachineLearningTut.pdf>
- [6] YANG, Ming-Hsuan, David J. KRIEGMAN a Narendra AHUJA. Detecting faces in images: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2012, **24**(1): 34-58 [cit. 2015-04-28]. DOI: 10.1109/34.982883. ISSN 01628828. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=982883>
- [7] VIOLA, Paul a Michael JONES. Robust Real-Time Face Detection. *International Journal of Computer Vision* [online]. 2004, **57**(2): 137-154 [cit. 2015-03-18]. DOI: 10.1023/B:VISI.0000013087.49260.fb. ISSN 0920-5691. Dostupné také z: <http://link.springer.com/10.1023/B:VISI.0000013087.49260.fb>
- [8] FREUND, Yoav a Robert E. SCHAPIRE. A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences* [online]. 1997, **55**(1): 119-139 [cit. 2015-04-18]. DOI: 10.1006/jcss.1997.1504. ISSN 00220000. Dostupné také z: <http://linkinghub.elsevier.com/retrieve/pii/S002200009791504X>
- [9] ŠOCHMAN, Jan a Jiří MATAS. AdaBoost and Face Detection. *Research Reports of CMP, Czech Technical University in Prague* [online]. 2003, **2003**(6) [cit. 2015-04-18]. ISSN 1213-2365. Dostupné z: <ftp://cmp.felk.cvut.cz/pub/cmp/articles/matas/Sochman-TR-2003-06.ps.gz>
- [10] Facial Expression Introduction. HAGER, Joseph C. *DataFace* [online]. 2003 [cit. 2015-04-23]. Dostupné z: <http://face-and-emotion.com/dataface/expression/expression.jsp>
- [11] Svaly hlavy – muscoli capitis. STOHWASSER, Petr. *Nabla* [online]. 2013 [cit. 2015-04-23]. Dostupné z: <http://www.nabla.cz/obsah/biologie/kapitoly/biologie-cloveka/svaly-hlavy-musculi-capitis.php>
- [12] SCHÜNKE, Michael, Erik SCHULTE, Udo SCHUMACHER a Lawrence M. ROSS. *Thieme atlas of anatomy*. Second edition. New York: Thieme Medical Publishers, Inc., 2014, 609 pages. ISBN 978-1-60406-922-8.

- [13] FACS. *Paul Ekman Group, LLC* [online]. 2015 [cit. 2015-04-23]. Dostupné z: <https://www.paulekman.com/product-category/facs/>
- [14] EKMAN, Paul. Basic emotions. DALGLEISH, Tim (ed.) a Mick POWER (ed.). *Handbook of cognition and emotion*. New York: Wiley, 1999. ISBN 0471978361.
- [15] Universal Expressions. *Facs Coding Group* [online]. VisualEmotion LLC, 2010 [cit. 2015-04-23]. Dostupné z: <http://www.facscodinggroup.com/universal-expressions>
- [16] Recognize these emotions. BEAUMONT, Leland R. *Emotional Competency* [online]. 2009 [cit. 2015-04-23]. Dostupné z: <http://www.emotionalcompetency.com/recognizing.htm>
- [17] LUCEY, Patrick, Jeffrey F. COHN, Takeo KANADE, Jason SARAGIH, Zara AMBADAR a Iain MATTHEWS. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops* [online]. IEEE, 2010, s. 94-101 [cit. 2015-04-20]. DOI: 10.1109/CVPRW.2010.5543262. ISBN 978-1-4244-7029-7. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5543262>
- [18] MARČELJA, S. Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America* [online]. 1980, **70**(11): 1297-1300 [cit. 2015-05-02]. DOI: 10.1364/JOSA.70.001297. ISSN 00303941. Dostupné také z: <http://www.opticsinfobase.org/abstract.cfm?URI=josa-70-11-1297>
- [19] GRIGORESCU, S. E., N. PETKOV a P. KRUIZINGA. Comparison of texture features based on Gabor filters. *IEEE Transactions on Image Processing* [online]. 2002, **11**(10): 1160-1167 [cit. 2015-05-02]. DOI: 10.1109/TIP.2002.804262. ISSN 10577149. Dostupné také z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1042386>
- [20] Introduction to Support Vector Machines. *OpenCV documentation* [online]. 2014 [cit. 2015-05-03]. Dostupné z: http://docs.opencv.org/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html
- [21] BURGESS, Christopher J. C. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery* [online]. 1998, **2**(2): 121-167 [cit. 2015-05-03]. DOI: 10.1023/A:1009715923555. ISSN 1384-5810. Dostupné z: <http://link.springer.com/10.1023/A:1009715923555>
- [22] ALY, Mohamed. *Survey on Multiclass Classification Methods* [online]. 2005 [cit. 2015-05-10]. Dostupné z: <http://www.vision.caltech.edu/malaa/publications/aly05multiclass.pdf>
- [23] *OpenCV* [online]. 2015 [cit. 2015-05-16]. Dostupné z: <http://opencv.org/>
- [24] How to build applications with OpenCV inside the Microsoft Visual Studio. *OpenCV documentation* [online]. 2014 [cit. 2015-04-15]. Dostupné z: http://docs.opencv.org/doc/tutorials/introduction/windows_visual_studio_Opencv/windows_visual_studio_Opencv.html#the-local-method

SEZNAM PŘÍLOH

A	Obsah přiloženého CD	51
---	----------------------------	----

A OBSAH PŘILOŽENÉHO CD

OpenCV/

- └ opencv-2.4.9.exe – knihovna OpenCV

Facepression/

- └ data/ – složka pro kaskády, klasifikátory a trénovací data
 - └ cascades/ – kaskády klasifikátorů pro detekci obličeje
 - └ classifiers/ – natrénované klasifikátory
 - └ faces/ – oříznuté obličeje
 - └ features/ – extrahované příznaky
 - └ original/ – originální obrázky trénovací množiny
 - └ screenshots/ – pořízené snímky obrazovky
 - └ test/ – obrázky testovací množiny
 - └ training/ – seznamy vzorků s jejich třídami
- └ Facepression/ – aplikace pro rozpoznávání výrazů obličeje
- └ FacepressionTraining/ – aplikace pro trénování klasifikátorů
- └ functions/ – statická knihovna s funkcemi sdílenými oběma aplikacemi
- └ Facepression.sln – řešení
- └ OpenCV_Debug.props – nastavení OpenCV knihovny (debug)
- └ OpenCV_Release.props – nastavení OpenCV knihovny (release)

RomanKozakDP.pdf – text práce