

System pro detekování a identifikaci osob na základě digitálního obrazu obličeje

Abstract

The diploma thesis deals with the issues of face recognition in image, especially detection, classification (gender, age and people) and verification of faces. For detection were tested two methods: detection based on Haar features using the boost algorithm AdaBoost and detection using Histograms of Oriented Gradients (HOG). In classification task was used method for reduction dimension of data - PCA (Principal Component Analysis) and subsequently tested these methods of classification: Euclidean distance, Mahalanobis distance, GMM (Gaussian Mixture Models) and linear SVM (Support Vector Machine). For verification task has been tested solution which uses LARK (Locally Adaptive Regression Kernel) descriptors in combination with PCA and linear SVM. Next part of the thesis was to create a modern Web interface that implements the most successful tested methods for detection, classification and verification. This system enables users to recognize their own images with faces of people.

Úvod

Rozpoznání osob v obraze je v současné době často řešený problém v oblasti počítačového vidění, jelikož je stále více kladena poptávka právě na identifikaci osob, pohlaví, věku či emocí. Diplomová práce se zabývá problematikou rozpoznání obličejů v obraze, a to zejména algoritmy pro detekci a identifikaci obličejů.

Cíle práce

- Seznámit se s problematikou rozpoznání obrazu a knihovny.
- Navržení algoritmů pro detekci a identifikaci osob z obrazu obličeje.
- Otestování algoritmů na volně dostupných databázích obrazů obličejů.
- Vytvoření webového systému, ve kterém bude uživatel moci rozpoznávat vlastní obrazy s obličejí osob.

Metodika rozpoznání obličejů

Pojem rozpoznávání obličejů v obraze pod sebou skrývá několik různých samostatných úloh:

- Detekce** – nalezení obdélníkové oblasti v obraze obsahující obličej.
- Klasifikace** – správné zařazení detekovaného obličeje do předem určených skupin (tříd). V případě této práce bylo klasifikováno pohlaví, věkový interval a konkrétní osoby.
- Verifikace** – ověření, zdali dva detekované obličejí náleží stejné osobě či nikoliv.

Pro nalezení nejvhodnějších řešení těchto úloh byly testovány různé algoritmy. V testech byly postupně použity tři různé obrazové databáze obličejů: Face Detection Data Set and Benchmark (FDDB), Unfiltered faces for gender and age classification (UFGA) a Labeled Faces in the Wild (LFW). Pro zjištění úspěšnosti bylo využito jedné z následujících metrik:

$$přesnost = \frac{TP}{TP + FN} \quad (1)$$

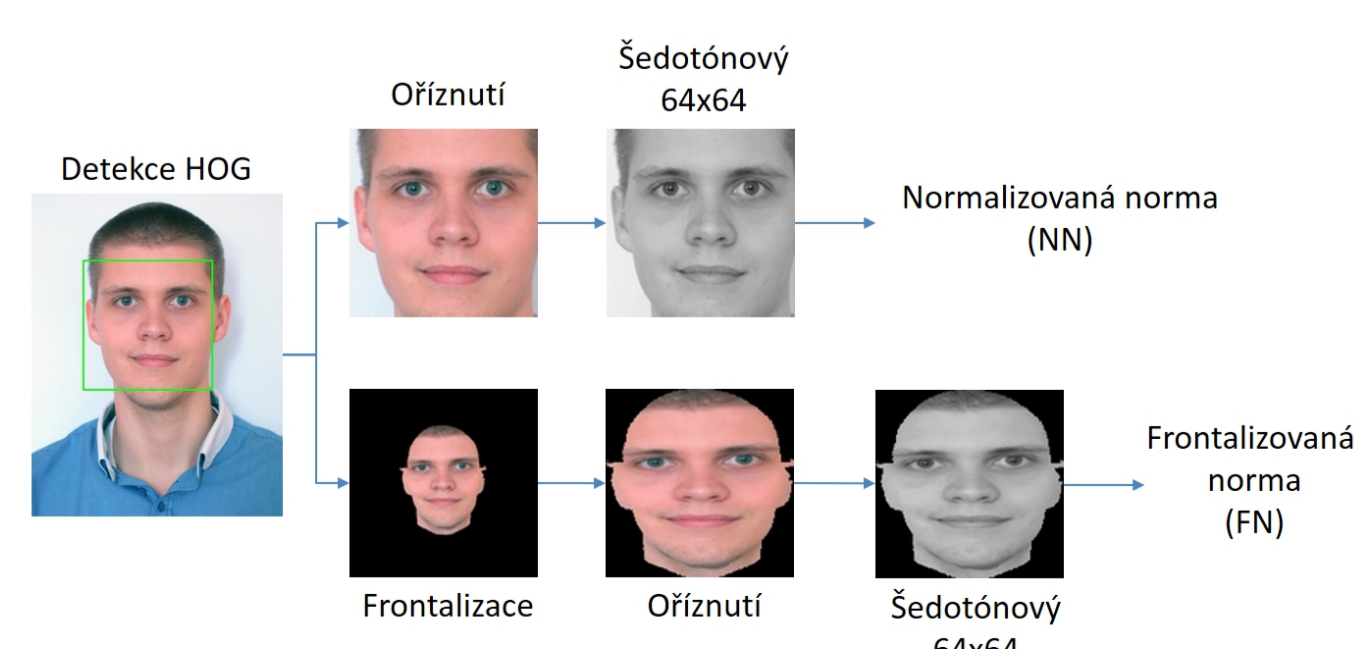
kde TP je počet pravdivě pozitivních vzorků a FN je počet falešně negativních vzorků. Při klasifikaci do více tříd je přesnost vypočtena pro každou klasifikovanou třídu zvlášť a následně je vypočten aritmetický průměr přesnosti pro všechny třídy. Druhou metrikou je:

$$úspěšnost = \frac{1}{1+n} \sum_{i=1}^{n-1} 1(\hat{y}_i = y_i) \quad (2)$$

kde y je vektor správných hodnot, \hat{y} je vektor předpovězených hodnot a n je celkový počet vzorků.

Testování a diskuze

Pro detekci byla otestována metoda na základě Haar příznaků s využitím boostovací metody AdaBoost, která dosáhla přesnosti **93,70 %**. Druhou otestovanou metodou byla detekce využívající Histogramů orientovaných gradientů, která dosáhla o něco vyšší přesnosti **96,30 %**. V úloze klasifikace obličejů byla nejdříve vytvořena normalizovaná a frontalizovaná norma obrazových dat (viz obr. 1) a následně byla využita metoda PCA (Principal Component Analysis) pro redukci dimenze dat.



Obr. 1: Předzpracování obrazu do dvou obrazových norem

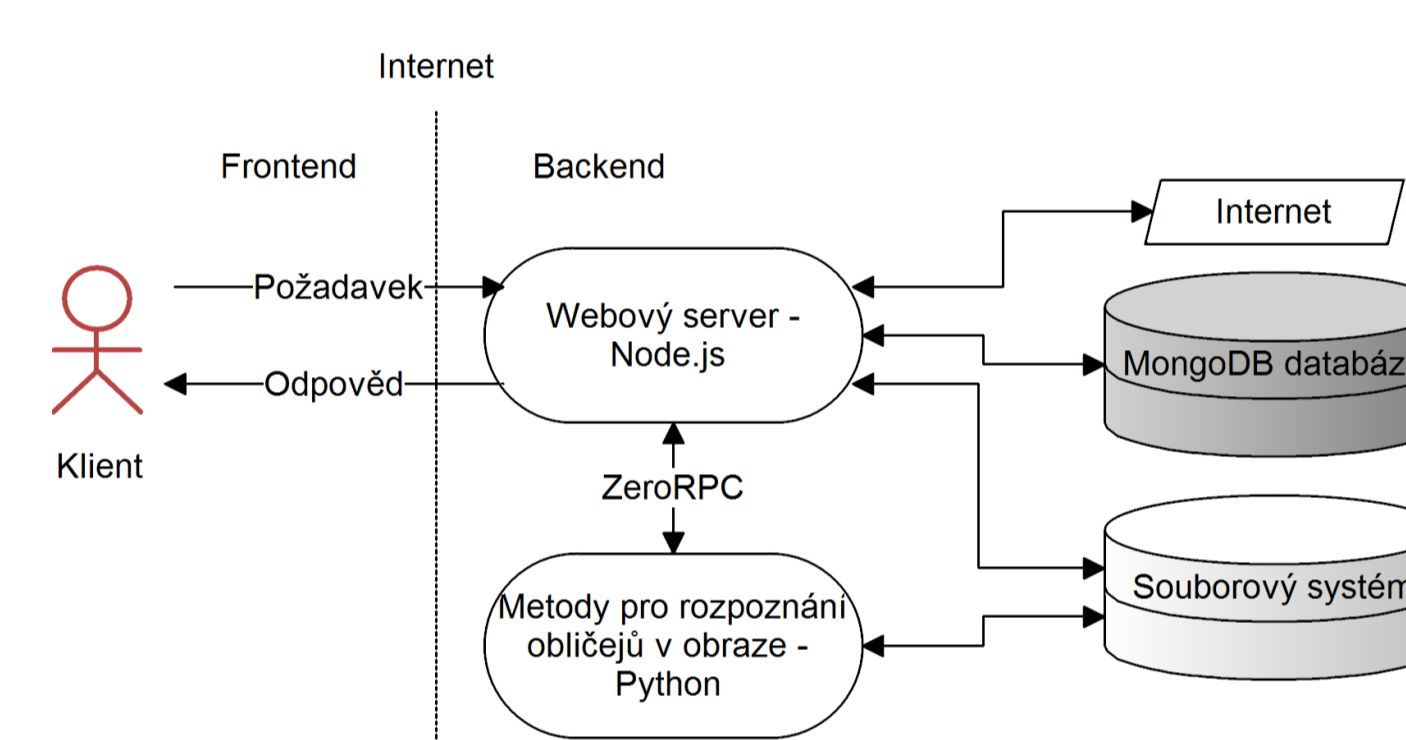
Dále byly provedeny testy úspěšnosti klasifikace pohlaví, 8 věkových intervalů (0-3, 3-7, 7-14, ..) a 84 konkrétních rozpoznávaných osob. Při testování bylo zjištěno, že při využití frontalizované normy je dosaženo ve všech experimentech vyšší úspěšnosti, proto jsou v následující tabulce zahrnuty pouze výsledky této normy.

Tabulka 1: Dosažené úspěšnosti v experimentech klasifikace

	Klasifikace pohlaví	Klasifikace věku (8 věk. intervalů)	Klasifikace osob (84 osob)
FN+PCA+Euklidovská vzdálenost	62,4 %	23,6 %	-
FN+PCA+Mahalanobisova vzdálenost	61 %	26,1 %	-
FN+PCA+GMM	70,8 %	32,3 %	-
FN+PCA+lineární SVM	79,8 %	34,5 %	86,62 %

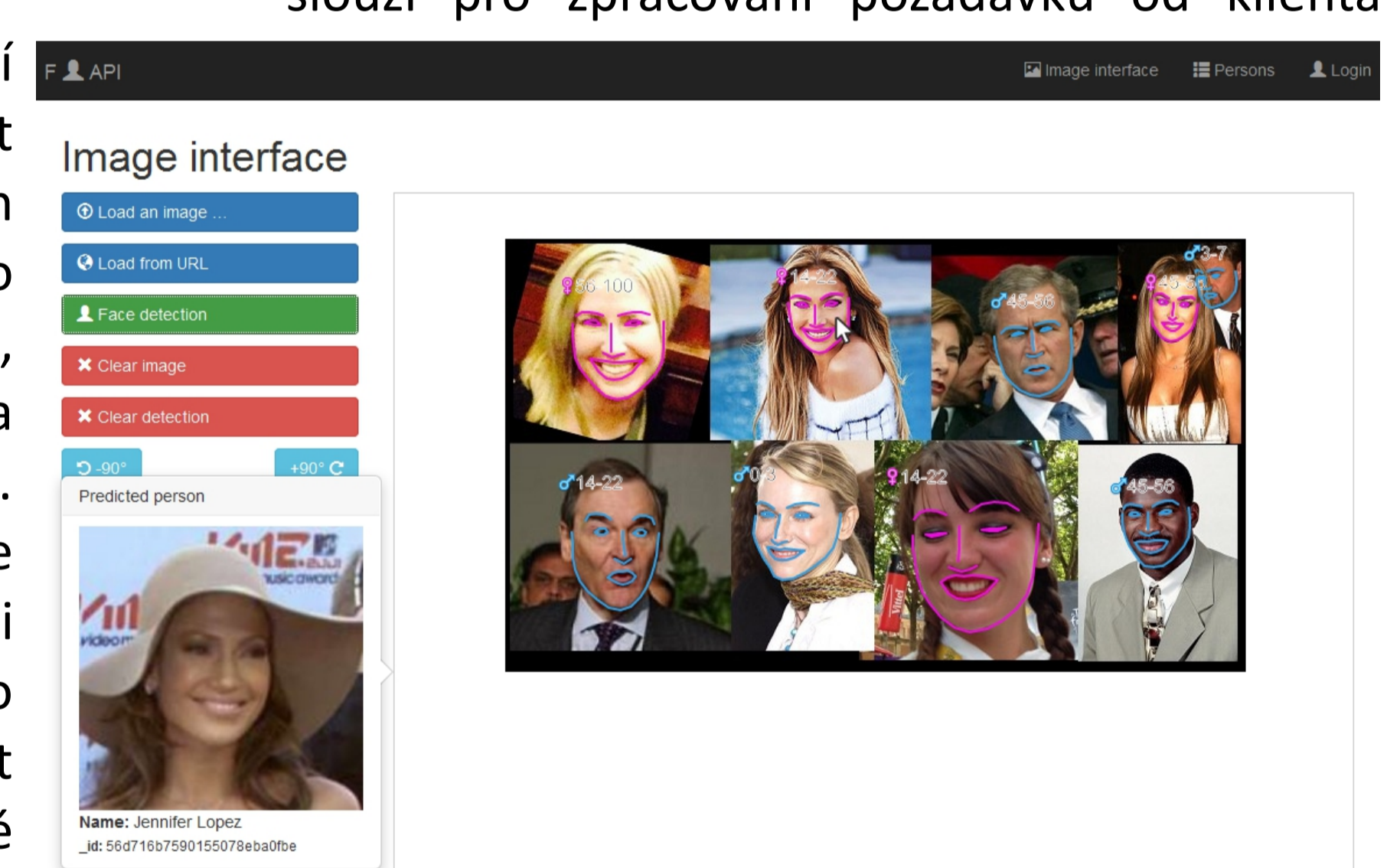
Poslední otestovanou metodou byla verifikace, ve které bylo při využití frontalizované normy, příznaků LARK, PCA a lineárního SVM dosaženo úspěšnosti **77,74 %**.

Webový systém



Obr. 2: Architektura webového systému

a také klientovi poskytuje responzivní HTML5 webovou aplikaci. Třetí část byla naprogramována v jazyce Python a stará se o pokročilé metody pro rozpoznání obličejů v obraze (detekce, klasifikace a verifikace). Architektura systému je znázorněna na obr. 2. Webový systém uživatelům umožňuje rozpoznávat vlastní obrazy s obličejí osob. Po přihlášení do webového systému má uživatel možnost spravovat systémem rozpoznávané osoby a obrazy jejich obličejů. Na obr. 3 je zobrazena ukázka rozpoznání obrazu s obličejí osob webovým systémem.



Obr. 3: Ukázka rozpoznání obrazu s obličejí osob webovým systémem

Závěr

V rámci diplomové práce byly úspěšně navrženy algoritmy pro detekci a identifikaci osob z obrazu obličeje. Tyto algoritmy byly vhodně otestovány na nalezených databázích obrazů obličejů osob. Dále byl navržen a vytvořen webový systém umožňující uživatelům rozpoznávat vlastní obrazy s obličejí osob.

- ŠONKA, Milan, Václav HLAVÁČ a Roger BOYLE. Image processing, analysis, and machine vision. Fourth edition. Stamford, CT: Cengage Learning, 2015. ISBN 978-1-133-59360-7.
- VIOLA, Paul a Michael J. JONES. Robust Real-Time Face Detection. International Journal of Computer Vision [online]. 2004, 2004(57 (2)), 137–154.
- SEO, Hae Jong a Peyman MILANFAR. Face Verification Using the LARK Representation. IEEE Transactions on Information Forensics and Security [online]. 2011, 6(4), 1275–1286
- HASSNER, Tal, Shai HAREL, Eran PAZ a Roee ENBAR. Effective face frontalization in unconstrained images. [online]. IEEE, 2015, ISBN 978-1-4673-6964-0.

martin.bumba@tul.cz | 776 460 408

Tato práce byla podpořena z projektu Studentské grantové soutěže (SGS) na Technické univerzitě v Liberci v roce 2016.