

# AUTOMATIC DATA ACQUISITION FOR TRAINING OF A HAND DETECTOR

**Juraj Blaho**

Master Degree Programme (1), FIT BUT

E-mail: xblaho00@stud.fit.vutbr.cz

Supervised by: Michal Hradiš

E-mail: ihradis@fit.vutbr.cz

## ABSTRACT

Manual annotation of data for training a classifier is usually a very time consuming task. Automatizing this process is therefore very important. This paper presents a novel method of automatized data acquisition for training of a hand detectors. By using the proposed method it is possible to collect thousands of training samples in just a few hours. This amount of data is enough to train a usable hand detector.

## 1. ÚVOD

Rozhnanie založené na detekcii a sledovaní rúk v obraze nájdu svoje uplatnenie v situáciách, keď je potrebné zabezpečiť bezkontaktné ovládanie. Takouto situáciou je napríklad ovládanie počítača pri chirurgických operáciách [3]. Ďalšie využitie môže byť diaľkové ovládanie domácej elektroniky, rozpoznanie znakovkej reči, či výuka bojových umení [2].

Kľúčovým problémom, ktorý musia takéto rozhnanie riešiť, je detekcia ruky a prípadne rozpoznanie konkrétnej pózy. K tomu je možné použiť štatistický klasifikátor, ten však vyžaduje veľké množstvo tréningových dát. Ručná anotácia dát býva veľmi zdĺhavá a nezábavná. Preto je potrebné navrhnúť metódu, ktorá by tento proces do značnej miery automatizovala.

Tento článok predstavuje novú metódu automatického získavania dát pre tréningovanie detektorov rúk. Úspešnosť metódy je demonštrovaná natréningovaním použiteľného detektora.

## 2. NAVRHNUTÁ METÓDA ZÍSKAVANIA DÁT

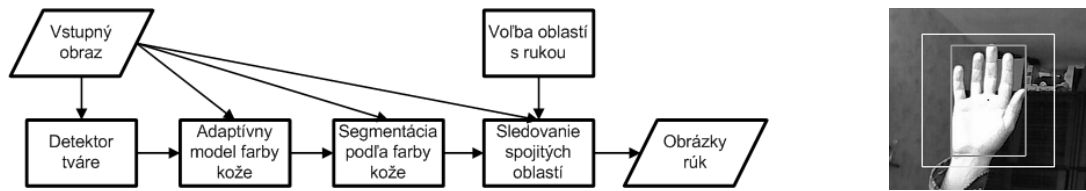
Navrhnutá metóda automatického získavania tréningových dát využíva skutočnosť, že farba rúk je podobná farbe tváre. Princíp metódy je znázornený na obrázku 1.

Vstupom metódy je obraz získaný z kamery. V tomto obraze je detekovaná tvár<sup>1</sup>. Na základe získanej pozície tváre je adaptovaný model farby kože. Následne je na základe tohto modelu segmentovaný vstupný obraz. Spojité oblasti s farbou kože sú sledované v po sebe idúcich obrazoch. Užívateľ zabezpečí výber oblastí obsahujúcich ruku. Okolo zvolených oblastí je vytvorený obalujúci obdĺžnik s potrebným pomerom strán. Tento obdĺžnik je po-

---

<sup>1</sup> Je možné použiť napríklad voľne dostupný detektor tváre z knižnice OpenCV.

tom zväčšený, aby vzniklo voľné miesto okolo ruky. Pozície výsledných obdĺžnikov potom môžu byť použité na anotáciu vstupného obrazu. Avšak z dôvodu redukcie ukladaných dát je vhodnejšie ukladať len výrezy zo vstupného obrazu.



**Obrázok 1:** Vľavo princíp metódy automatického získavania tréningových dát. Vpravo znázornený výstup metódy, kde vnútorný obdĺžnik predstavuje pôvodnú spojitú oblasť farby kože a vonkajší obdĺžnik je výsledný ukladaný výrez s potrebným pomerom strán.

## 2.1. ADAPTÁCIA MODELU FARBY KOŽE

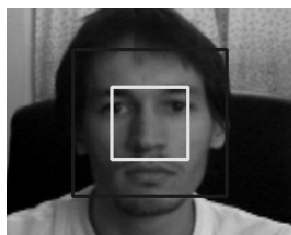
Kvôli nezávislosti na intenzite osvetlenia a zároveň rýchlosti výpočtu je pre popis farby zvolený normalizovaný RG priestor. Jeho dve zložky sú určené rovnicami:

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

Klasifikátor modeluje triedu farby kože a triedu farby pozadia. Hustoty rozložení pravdepodobností oboch tried sú modelované normalizovanými 2D histogramami. Pravdepodobnosť, že daný pixel je koža sa vypočíta podľa Bayessovho vzorca:

$$P(\text{skin}|x) = \frac{p(x|\text{skin})P(\text{skin})}{p(x|\text{skin})P(\text{skin}) + p(x|\text{background})P(\text{background})} \quad (2)$$

Model sa adaptuje počas behu systému na základe detekcie tváre vo vstupnom obraze. Princíp je znázornený na obrázku 2. Obdĺžnik predstavujúci výstup z detektora tváre je zmenšený, pričom ostáva zachovaný jeho stred. Pixely vnútri zmenšeného obdĺžnika sú použité na adaptáciu triedy farby kože. Všetky pixely mimo detekovanú tvár sú považované za pozadie. Do pozadia sú teda zahrnuté aj oblasti ako krk alebo ruky, ktoré majú farbu kože. Ale vzhľadom na malú plochu týchto oblastí v porovnaní s celým obrazom je ich vplyv na kvalitu modelu nevýznamný. Na adaptáciu sú využívané všetky dovtedy získané dáta, pričom starším dátam sa postupne znižuje ich váha.



**Obrázok 2:** Pixely vo vnútornom obdĺžniku sú považované za kožu. Vonkajší obdĺžnik predstavuje detekciu tváre. Pixely mimo vonkajší obdĺžnik sú považované za pozadie.

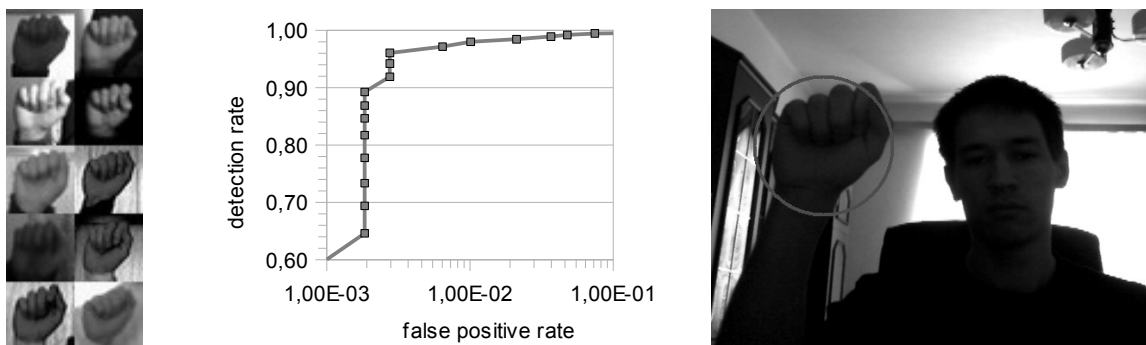
## 2.2. SLEDOVANIE OBLASTÍ

Každá spojitá oblasť farby kože má priradený príznak aktivity. Ten určuje, či je oblasť považovaná za ruku alebo nie, a teda, či budú na základe jej pozície získavané dáta pre tréningovanie. Na začiatku sú všetky oblasti neaktívne. Oblasť je aktivovaná, keď prejde cez zvolený bod v obraze. Toho užívateľ dosiahne umiestnením ruky na dané miesto. Tiež by sa dali oblasti aktivovať kliknutím myši, avšak to by znížilo mieru automatizácie metódy.

Korešpondencia oblastí v dvoch po sebe nasledujúcich obrazoch je určená na základe ich vzdialenosti. Bolo by možné využiť aj lepšie metódy korešpondencie, prípadne dokonalejšie metódy sledovania, avšak tento spôsob je dostatočný.

### 3. TESTY

Dáta získané prezentovanou metódou boli použité na natrénovanie detektora ruky zaťaťej v päst'. Bol použitý detektor založený na AdaBoost implementovaný knižnicou OpenCV. Trénovacie dáta boli získané za niekoľko málo hodín a obsahovali 5625 pozitívnych príkladov rúk. Všetky získané príklady boli zmenšené na 24x24 pixelov. Úspešnosť detektora bola testovaná na 1088 snímkoch s rozlíšením 640x480, so zložitým pozadím a s vysokým šumom kamery. Výsledky testov sú znázornené na obrázku 3. False positive rate určuje pomer medzi celkovým počtom falošných detekcií k počtu snímkov.



**Obrázok 3:** Vľavo ukážka trénovacích dát. V strede ROC krivka výsledného detektora. Vpravo ukážka detekcie ruky.

### 4. ZÁVER

V rámci tohto článku bola predstavená metóda automatického získavania trénovacích dát. Táto metóda umožňuje získavať až niekoľko trénovacích príkladov za sekundu. To v praxi znamená, že za pár hodín je možné získať až niekoľko tisíc príkladov. Bolo ukázané, že na základe takto získaných dát je možné natrénovať detektor dostatočne úspešný pre praktické použitie.

Ďalšie zlepšenie výsledkov metódy by bolo pravdepodobne možné dosiahnuť zarovnaním získaných dát. Toto však zatiaľ nebolo otestované. Princípy použité v predstavenej metóde je možné aplikovať taktiež na získavanie trénovacích dát pre iné objekty ako ruky.

### LITERATÚRA

- [1] Francke, H.; Ruiz del Solar, J.; Verschae, R.: Real-Time Hand Gesture Detection and Recognition Using Boosted Classifiers and Active Learning. In Pacific Rim Symposium on Image Video and Technology, 2007, s. 533–547.
- [2] Kohler, M.: Vision Based Hand Gesture Recognition Systems. URL <http://ls7-www.cs.uni-dortmund.de:81/research/gesture/>
- [3] Wilson, A.: Hand gesture recognition system targets medical applications. 2008. URL [http://www.vision-systems.com/display\\_article/340083/19/none/none/Techt/Hand-gesture-recognition-system-targets-medical-applications](http://www.vision-systems.com/display_article/340083/19/none/none/Techt/Hand-gesture-recognition-system-targets-medical-applications)