

MOBILE AUGMENTED REALITY IN COMMERCIALS

Michael Angelov

Master Degree Programme (2), FIT BUT

E-mail: xangel00@stud.fit.vutbr.cz

Supervised by: Vítězslav Beran

E-mail: beranv@fit.vutbr.cz

Abstract: This paper presents a possible application of augmented reality in domain of commercials. It presents designed architecture of a mobile application that is able to detect and track specific objects (e.g. printed commercials, logos) in mobile's phone camera in real time and provide some extra information concerning the detected object towards the user.

Keywords: mobile application, augmented reality, object recognition, object tracking

1 ÚVOD

Posledných niekoľko rokov môžeme byť svedkami rýchleho vývoju v oblasti mobilných telefónov a s nimi súvisiacich aplikácií, vďaka čomu sa stali dnes už neodmysliteľnou súčasťou každodenného života. Zároveň nemenej rýchlo postupuje aj akademický výskum v oblasti počítačového videnia, pričom dochádza ku každodennému nasadzovaniu nových algoritmov do bežnej praxe. Ďalším zaujímavým trendom zavádzanie rozšírenej reality, ktorá je po neúspechu virtuálnej reality novým a ešte nie celkom preskúmaným fenoménom v oblasti informačných technológií.

Tento príspevok prezentuje konkrétnu možnosť využitia spomenutých trendov v praxi - je ňou mobilná aplikácia schopná rozpoznávať objekty (napr. fotografia v časopise, billboard na ulici) v kamere mobilného telefónu a zobrazovať o nich rozširujúce informácie na displeji v štýle rozšírenej reality. Primárne nasadenie aplikácie sa predpokladá v oblasti reklamy, prípadne však bude môcť tento nový informačný kanál nájsť uplatnenie aj v iných oblastiach.

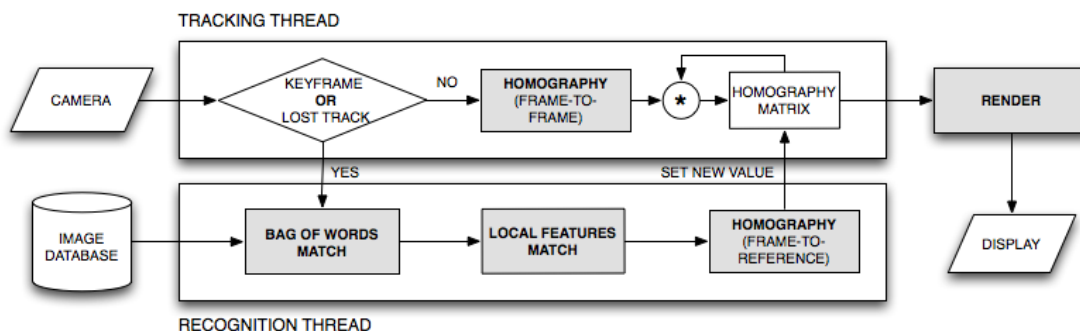
2 ARCHITEKTÚRA

Skutočnosť, že aplikácia bola navrhnutá na fungovanie v reálnom čase a na mobilnom zariadení so sebou prináša zvýšené nároky na efektivitu návrhu, primárne z dôvodu obmedzených výpočetných prostriedkov, ktorými tieto zariadenia v súčasnosti disponujú. Z tohto dôvodu bolo nutné venovať zvýšenú pozornosť výberu vhodných metód, ktoré boli následne pospájané do funkčných blokov tvoriacich základ celého mechanizmu aplikácie. Ten možno vo svojej podstate rozdeliť do dvoch nezávislých fáz:

- off-line spracovanie – vytvorenie obrazovej databázy z referenčných snímok (training)
- on-line spracovanie – detekcia snímaného objektu v referenčnej databáze a jeho sledovanie (matching, tracking)

Výstupom procesu tréningu je obrazová databáza, ktorá sa následne predá na vstup procesu rozpoznávania. Po úspešnom rozpoznaní objektu pokračuje proces sledovania objektu v snímku.

Z dôvodu zvýšenia efektívnosti spracovania v on-line fáze sú oba procesy (sledovanie a rozpoznávanie) realizované ako samostatné vlákna (Obrázok 1). Kým proces sledovania prebieha periodicky a je



Obrázek 1: On-line fáza sledovania a detekcie.

synchronizovaný s obrazovým vstupom z kamery, výpočetne náročnejší proces rozpoznávania musí byť aktivovaný podstatne zriedkavejšie. Dôležité je taktiež, že rozpoznávanie prebieha na pozadí a viditeľne neovplyvňuje plynulosť aplikácie. Za účelom detekcie a deskripcie lokálnych charakteristík sa používa robustná a prijateľne rýchla metóda SURF (Speeded-Up Robust Features, [1]). Sledovanie obrazu v kamere (tracking) je realizované s použitím rýchleho rohového detektoru FAST (Features from Accelerated Segment Test, [3]) v kombinácii s binárnym deskriptorom BRIEF (Binary Robust Independent Elementary Features, [2]). Na hľadanie homografie medzi sledovaným a referenčným snímkom je použitý bežne zaužívaný postup s využitím metódy RANSAC (RANDOM SAMPLE CONSENSUS). Navrhnutý postup je realizovaný s využitím knižnice OpenCV 2.2.

2.1 OBRAZOVÁ DATABÁZA

Úlohou obrazovej databázy je poskytnúť predspracované dáta referenčných obrazov na vstup procesu rozpoznávania. Pre každý z referenčných obrazov sa tu ukladajú lokálne charakteristiky popísané príslušnými deskriptormi, avšak veľký počet (rádovo stovky až tisíce na jeden obraz) a vysoká dimenzia deskriptorov (64 až 128-dimenzionálny vektor typu float pre SURF), prináša so sebou výkonnostné problémy pri vyhľadávaní podobností v obrazovej databáze.

Z tohto dôvodu je vyhľadávanie v obrazovej databáze, založené na metodike vizuálneho slovníka a metódy Bag of Words. Takto možno každý referenčný obraz popísať jedným histogramom výskytu vizuálnych slov v obrázku, čo umožňuje rýchle vyhľadanie podobných obrázkov z databázy. S vybranými obrázkami sa následne vykoná klasické hľadanie zhody a úrovni lokálnych charakteristík.

2.2 SLEDOVANIE POHYBU KAMERY A OBJEKTU

Kým sledovanie pohybu kamery je nevyhnutnou súčasťou v procese rozpoznávania, sledovanie objektu v kamere je kľúčové z hľadiska následného zobrazenia rozšírenej reality na displeji mobilného zariadenia.

Sledovanie pohybu kamery sa principiálne opiera o hľadanie homografie medzi snímkami (frame-to-frame), pričom získaná transformačná matica (matica homografie) určuje druh a smer pohybu kamery medzi jednotlivými snímkami. Primárne sa takto zistí prítomnosť pohybu v kamere, čo má význam pri identifikácii tzv. kľúčových snímkov (keyframes), teda tých snímkov, ktoré sa ďalej spracovávajú pre účely rozpoznávania.

Po detekovaní objektu v kľúčových snímkoch sa následné sledovanie objektu realizuje pomocou výpočtu homografie medzi dvojicami snímkov (frame-by-frame) pomocou už spomínaných metód FAST, BRIEF a RANSAC, pričom lokálne charakteristiky sa detekujú v okolí detekovaného objektu, čím sa odfiltruje rušivý pohyb z okolia objektu.

Bežne sa môže stať, že sledovanie objektu v kamere z nejakého dôvodu zlyhá alebo sa stane nepresným (napr. náhly pohyb kamery, vyčnievanie časti objektu z kamery). V takomto prípade sa nová matica homografie vypočíta z aktuálnej snímky a referenčného obrázku (frame-to-reference) a proces sledovania začne odznova.

2.3 ROZPOZNÁVANIE A VYHL'ADÁVANIE V OBRAZOVEJ DATABÁZE

Kľúčovou fázou aplikácie je proces rozpoznania a vyhľadania objektu zo snímky z kamery v obrazovej databáze. Po výbere kľúčovej snímky sa z nej vypočíta histogram Bag of Words, podľa ktorého sa v obrazovej databáze rýchlo vyhľadá niekoľko najlepších kandidátov, v rámci ktorých sa bude vyhľadávať zhoda lokálnych charakteristík. Po nájdení dostatočnej zhody je ďalším krokom určenie homografie medzi kľúčovou snímkou a referenčným obrázkom, pričom výstupom metódy je projekčná matica popisujúca geometrickú transformáciu medzi snímkou z kamery a referenčným obrázkom z databázy.

3 ZÁVER

Navrhnutá aplikácia je v súčasnosti realizovaná ako desktopová aplikácia na platforme Unix, pričom v najbližšej dobe bude prenesená do mobilného zariadenia iPhone 4G. Funkcionalita aplikácie bola testovaná s kamerou s rozlíšením 640x480, pričom rýchlosť spracovania dosahuje 10 až 15 fps na platforme Mac (Core2Duo 2GHz, 2GB RAM). Z testov, ktoré prebehli na platforme iPhone vyplýva, že sledovanie objektov s kamerou s rozlíšením 320x480 dosahuje rovnaké framerate ako na platforme Mac. Aplikácia bola navrhnutá tak, aby vďaka podpore vlákien v iPhone mohla fungovať v reálnom čase aj na tomto zariadení.



Obrázek 2: Detekcia obálky časopisu .tyždeň. Vstupná snímka z kamery (vľavo) a detekovaná obálka časopisu prekrytá rozšírenou realitou, v tomto prípade referenčným obrázkom pre ukážku (vpravo).

Podakovanie: táto práca vznikla za podpory grantu Centrum excelence IT4Innovations, CZ 1.05/1.1.00/02.0070.

REFERENCE

- [1] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. In *In ECCV*, pages 404–417, 2006.
- [2] Michael Calonder, Vincent Lepetit, Christoph Strecha, and Pascal Fua. Brief: Binary robust independent elementary features. In *ECCV (4)*, pages 778–792, 2010.
- [3] Edward Rosten and Tom Drummond. Fusing points and lines for high performance tracking. In *IEEE International Conference on Computer Vision*, volume 2, pages 1508–1511, October 2005.