

Metódy segmentácie snímkov a ich využitie pri analýze chôdze človeka

ATTILA KOTRBA

LUKÁŠ KROČKA

*Slovenská technická univerzita
Fakulta informatiky a informačných technológií
Ilkovičova 3, 842 16 Bratislava
attila.kotrba@centrum.sk
krofi@orangemail.sk*

Abstrakt. Pri analýze pohybu ľudskej nohy, tj. analýze chôdze je najväčším problémom dostatočne presne a kvalitne získať polohy značiek na nohe. V tejto práci sa zameriavame na existujúce metódy segmentácie pri extrahovaní značiek umiestnených na ľudskom tele a na experimenty, s ktorými sme pracovali počas projektu. Práca uvažuje o niektorých existujúcich metódach segmentácie a ukazuje ich výhody a nevýhody, možnosti ich použitia. Naše prístupy k segmentácii sú podrobne popísané, uvádzame aj experimenty, ktoré sme vykonali, a snažíme sa zhodnotiť úspešnosť jednotlivých využitých metód. Práca nakoniec zhodnotí použité prístupy a navrhne možné pokračovanie projektu a potrebné zdroje.

Úvod

Cieľom tejto práce je publikovať naše výsledky experimentovania so segmentačnými algoritmami pri analýze chôdze človeka. Chceme poukázať na problémy, s ktorými sme sa počas experimentov stretli a poskytnúť prehľad tých metód, s ktorými sme úspešne dokázali segmentovať snímky videozáznamov a animovať pomalý pohyb ľudskej nohy. Na identifikáciu umiestnenia nohy v scéne a identifikáciu umiestnenia jednotlivých kĺbov na nohe sme použili kontrastné značky. Tieto značky zodpovedali pozíciám jednotlivých kĺbov nohy. Cieľom segmentačných algoritmov potom bolo identifikovať tieto značky a zistiť súradnice ich stredov. Pri experimentovaní sme zatiaľ používali iba jednu kameru, teda sme zisťovali len 2D súradnice. Analýzou existujúcich riešení v oblasti animácie chôdze sme dospeli k zisteniam, že v praxi sa na tento účel používajú aktívne značky a špeciálne hardvérové vybavenie (vysokorýchlostné kamery a iné špeciálne zariadenia), ktoré umožňujú vykonávať animáciu v reálnom čase a aj pri rýchlych pohyboch. Keďže v praxi sa používajú

Metódy segmentácie snímkov, máj 2006, s. 1-7.

aktívne značky, tak nie je problém s identifikáciou konkrétneho kľbu. Aktívne značky sa totiž vyznačujú tým, že v danom čase všetky videokamery zachytia len jednu značku. Ostatné značky nie sú v tomto čase viditeľné. Postupným sekvenčným spínaním týchto aktívnych značiek sa docieli, že vieme jednoznačne identifikovať 3D súradnice každej značky. Keďže v rámci nášho projektu sme nemali k dispozícii takéto hardvérové vybavenie, snažili sme sa použiť iné princípy. Museli sme navrhnuť vhodný spôsob, ako dosiahnuť uspokojivé výsledky pri segmentácii snímok a pri použití bežných kamier (s frekvenciou 25 snímok za sekundu).

Naše prístupy k segmentácii

Na začiatku projektu sme boli vystavení ťažkej úlohe, a to upraveniu existujúcich spôsobov a metód, ktoré v podstate vôbec nefungovali. Našou úlohou bolo sprevádzkovanie celého cyklu animácie pohybu ľudskej nohy, tj. nasnímanie pohybu nohy, analýza a segmentácia jednotlivých snímok, a nakoniec animácia pohybu prostredníctvom modelu kostry človeka. Po náročnom štúdiu, keďže sme nemali z danej oblasti prakticky žiadne skúsenosti, sme sa snažili využiť značky jednej farby. Uvažovali sme nad využitím svietiacich značiek, ktoré by generovali svetlo, či už z viditeľného spektra, alebo neviditeľného. Takéto prístupy ale vždy zlyhávali buď na nedostatku prostriedkov, alebo na náročnej technickej realizácii potrebných elementov.

Rozhodli sme sa preto, že využijeme značky vyrobené z reflexných materiálov. Reflexné materiály sú také, ktorých pomer odrazeného svetla je omnoho väčší oproti pohltému svetlu a po osvetlení takýchto materiálov sú tieto omnoho svetlejšie, ako ostatné. V snímanej scéne sa vhodne osvetlené reflexné značky javia ako zdroj svetla a majú väčšiu intenzitu jasú, ako okolie. Našou úlohou bolo tiež skonštruovanie takýchto značiek. Nemohli sme použiť ploché značky, ale špeciálne vyrobené vypuklé značky polguľatého tvaru, aby sme zabezpečili ich viditeľnosť viacerými kamerami súčasne. Ďalším hlavným dôvodom použitia vypuklých značiek bolo, aby kamera z bočného pohľadu zachytila aj značky, z ktorých jedna bola umiestnená na päte a druhá na palci nohy. Podmienkou pre dosiahnutie maximálneho kontrastu medzi vypuklými značkami a pozadím v scéne je, aby zdroje svetla boli dostatočne intenzívne a aby správne osvietili značky. Osvietenie značiek sme zabezpečili tak, aby bolo odrazené svetlo dostatočne silné na rozpoznanie kamerami a segmentačným algoritmom. Zdroje svetla museli byť umiestnené rovnobežne s kamerami a čo najbližšie k nim, aby odrazené lúče svetla od značiek dopadali kolmo do objektívov kamier. Keďže sme použili značky vypuklých tvarov, od každej z nich sa istá časť lúčov odrazila kolmo do objektívu. Vďaka tomu boli značky na videozázname omnoho intenzívnejšie, ako pozadie scény. Po úspešnom osvetlení scény a nasímaní príslušných videozáznamov sme sa museli zamerať na segmentačný algoritmus.

Po teoretických úvahách a preštudovaní prác predchádzajúcich tímov pracujúcich na tomto projekte sme sa rozhodli, že využijeme algoritmus, ktorý bude sledovať hodnotu intenzity jednotlivých pixelov, a na základe tejto hodnoty sa rozhodne, či daný pixel môže byť značkou, alebo nie. Táto metóda sa nazýva prahovanie. Tiež sme

skúšali použiť algoritmus odčítavania snímkov od pozadia scény, avšak po vyhodnotení experimentov s týmto algoritmom sme zistili, že táto metóda nie je použiteľná. Odčítaním snímkov sa výrazne znížila intenzita značiek, zmenila sa aj ich farba a výsledný obraz po odčítaní snímkov bol veľmi zašumený (obsahoval veľké množstvo bodov rôznych farieb, ktoré boli náhodne rozmiestnené po celej scéne). V nasledujúcich podkapitolách preto opisujeme len tie algoritmy, s ktorými sme dosiahli úspešné výsledky. V princípe ide o rôzne modifikácie metód prahovania.

Algoritmus segmentácie pomocou rekurzívnej záplavovej metódy hľadania značiek

Tento algoritmus je založený na nasledovnom princípe: ak je hodnota jas pixelu väčšia ako preddefinovaná hodnota, označí tento pixel ako súčasť značky, ak hodnota jas nie je väčšia, tak tento pixel vynechá. V prvom kroku algoritmus prechádza všetky pixely aktuálneho snímku. Ak zistí, že intenzita jas niektorého prechádzaného pixelu je väčšia ako preddefinovaná hodnota, označí tento pixel ako súčasť značky a hľadá ďalej v okolí tohto pixelu (tj. kontroluje 8 pixelov susediacich s daným pixelom). Ak v okolí nájde pixely, ktoré vyhovujú, označí ich ako súčasť tej istej značky a takýmto spôsobom rekurzívne prehľadáva aj okolie týchto vyhovujúcich pixelov. Postupne takto prejde všetky značky, a výsledkom algoritmu je dvojrozmerné pole, kde sú umiestnené polohy jednotlivých pixelov, ktoré patria k jednotlivým značkám. Potom algoritmus pokračuje a získa stredy jednotlivých značiek. Tieto stredy vypočíta jednoduchým matematickým prepočtom, keď pre každú značku zoberie pri x-ovej aj y-ovej súradnici najväčšie a najmenšie hodnoty a spraví ich priemer. Výsledkom celého procesu je pole stredov značiek.

Tento algoritmus vykazoval veľmi premenlivé výsledky. Jeho základné problémy boli nasledujúce:

A, Hodnota jas značiek sa menila v závislosti od rýchlosti ich pohybu. Keďže značky sú umiestnené na rôznych miestach (klboch) na nohe, tak sa aj pohybujú rôznymi rýchlosťami a dosahujú rôzne hodnoty jas. V prípade použitia silnejších svetiel bol jas celej scény omnoho vyšší, ako keď sa záznam natáčal v tmavej miestnosti s diódami, ako zdrojmi svetla. Preto fixná hodnota minimálnej intenzity jas pre značku nebola vyhovujúca.

B, Na snímkoch sa ukazovalo, že veľkým problémom je aj rušenie. Rôzne kovové predmety, ktoré sa dostali do scény a pri osvetlení odrážali svetlo, boli pre algoritmus problémom, pretože ich označil za značky, aj keď značkami neboli.

C, Niektoré značky na snímkoch sa niekedy ukazovali byť akoby rozdvojené, tj. niektoré časti značiek odrážali svetlo viac, a niektoré menej. Toto bolo spôsobené nedokonalou výrobou značiek. Tento problém sme odstránili výmenou danej značky, pri ktorej tento jav nastával.

D, Algoritmus rozoznal aj jeden pixel ako značku, pričom bolo jasné, že jeden samotný pixel značkou nemôže byť. Toto bolo spôsobené náhodným šumom v snímkoch.

E, Algoritmus bol pomalý, prakticky nepoužiteľný na dlhších sekvenciách

F, Pri rýchlych pohyboch sa značky rozmazávali. Toto spôsobovali kamery, ktoré nestíhali rýchlejšie zábery. Tento problém sa najviac prejavil pri tmavších záberoch, tj. pri použití diód ako zdrojov svetla.

Tieto problémy sme riešili postupne úpravami algoritmu. Problémy rušenia, a rozdvojenia značiek sme vyriešili úpravou algoritmu tak, že v prípade, ak sa za značku označí malý počet bodov, značka sa zruší a nebude sa považovať za značku. Týmto sa vyriešili náhodné javy, problémy C,D a čiastočne aj B.

Rozmazávanie značiek (F) sme nevedeli vyriešiť, keďže toto bolo spôsobené kamerami, ktoré rýchle pohyby rozmazávali. Pokúsili sme sa o získanie rýchlych kamier a ich využitie, ale kvôli nedostatku finančných prostriedkov sa nám to nepodarilo.

Rôzne intenzity značiek v scéne (A) nebolo možné v tomto algoritme nejakým vhodným spôsobom úplne eliminovať. Zaviedli sme možnosť konfigurácie hraničnej hodnoty jasu, avšak aj napriek tomu sa občas stalo, že niektorá rýchlejšie sa pohybujúca značka nebola algoritmom rozpoznávaná. Jej hodnota jasu v snímku totiž v dôsledku rýchlejšieho pohybu poklesla pod prahovú hodnotu.

Softvér bol po týchto úpravách lepší, ale rekurzívny algoritmus bol stále veľmi pomalý, a v prípade veľkého šumu, tento nebolo možné softvérovo odstrániť. Preto sme navrhli nový algoritmus, ktorého princípom je, že na začiatku sa segmentuje len jeden snímok (použitím vyššie popísaného algoritmu). Tento snímok sa potom zobrazí používateľovi. Používateľ tu uvidí, ako sa rozoznali značky na snímku. V prípade, ak algoritmus nájde značku nesprávne, používateľ má možnosť ju presunúť na správne miesto, alebo vymazať. Ak algoritmus nenájde značku, používateľ môže vytvoriť novú značku a umiestniť ju tam, kde sa skutočne nachádza. Ak sa v prvom snímku rozpoznajú aj ďalšie značky, ktoré nie sú považované za značky (napríklad rôzne lesklé predmety v scéne), používateľ ich môže manuálne vymazať. Po týchto úpravách používateľ zavrie okno, a softvér pokračuje v segmentácii novým, upraveným spôsobom. Pritom sa už použije algoritmus, ktorý je popísaný v nasledujúcej podkapitole.

Algoritmus segmentácie s využitím lokálnych selekcií a výberu lekárom

Po vyznačení značiek na prvom snímku lekárom sa tento prvý snímok uzavrie, a za referenčné pozície bodov sa berú práve tie pozície, ktoré určil lekár. Algoritmus v nasledujúcom snímku nerobí kompletnú segmentáciu celého snímku, ale oblasti segmentácie sú presne určené. Oblasti sú určené tak, že z predchádzajúceho, už segmentovaného snímku (snímok, ktorý upravil manuálne lekár do presnej podoby alebo snímok, ktorý už bol segmentovaný týmto algoritmom), sa získajú súradnice stredov značiek. Okolie každého stredu značky potom tvorí segmentačnú oblasť, v ktorej sa realizuje segmentačný algoritmus. Algoritmus pracuje na princípe, že pohyb nohy je spojený a rýchlosť tohto pohybu neprekročí určitú hodnotu. Pokusne bola určená hodnota vzdialenosti v x-ovej aj y-ovej osi od súradnice stredu značky (čiže sme pokusne určili veľkosť oblasti, v ktorej strede ležala daná značka). V tejto oblasti sa teda nachádzala daná značka na predchádzajúcom snímku. Keďže rýchlosť pohybu nohy je vždy istým spôsobom obmedzená, tak vhodnou voľbou veľkosti tejto oblasti

vieme zabezpečiť, že sa v nej značka bude nachádzať aj v nasledujúcom snímku. Na predchádzajúcom snímku teda máme nájdené oblasti (štvorce), pričom v strede každej oblasti je umiestnená jedna značka. V týchto oblastiach sa potom vykonáva segmentácia. Ak sa noha so značkami pohybuje, tak v nasledujúcom snímku už značky nie sú umiestnené v stredoch definovaných oblastí, ale tieto značky sú posunuté v smere pohybu nohy. Algoritmus v týchto oblastiach hľadá, kde sa nachádzajú stredy značiek. Na základe týchto nových (posunutých) stredov značiek definuje algoritmus nové oblasti (oblasti sa teda pohybujú spolu so značkami a stred každej oblasti sa vždy nachádza v strede príslušnej značky). Segmentačný algoritmus teda používa metódu lokálneho prahovania (pre každú oblasť sa zvlášť určuje prahová hodnota). Segmentácia je vykonávaná rekurzívnym algoritmom, ale len v danej oblasti. Takto sa rýchlosť segmentácie výrazne zvýši. Keď uvažíme, že celý snímok má veľkosť $640 * 480$ pixelov, teda 307200, a veľkosť okna je stanovená pokusne na 50 pixelov, počet značiek, ktoré sa rozoznávajú je 5, potom celkový počet pixelov, ktoré sa musia prečítať rekurzívnym algoritmom možno jednoducho vypočítať ako $50 * 50 * 5$. Teda spolu 12500 pixelov, čo predstavuje asi 4 % celkového počtu pixelov na snímku. Výhodou tohoto algoritmu oproti predchádzajúcemu je:

A, Výrazné zrýchlenie. Pokusy, ktoré sme vykonávali, ukazovali zlepšenie času segmentácie až na 20% pôvodnej hodnoty. Keďže počet prehľadávaných pixelov je 4% z celkového počtu, zdá sa číslo 20% ako veľké. To však nie je, keď sa nad tým trochu zamyslíme. V oblastiach, kde sa nachádzajú značky, je väčší počet bodov s vysokým jasom, a teda rekurzívny algoritmus tu pracuje ťažšie a pomalšie.

B, Vylúčenie náhodného šumu a objektov, ktoré nepatria do značiek na scéne. Ak sa na scéne nachádza kovový výrazne lesklý objekt, ktorý nie je značkou, tento algoritmus ho nemusí brať do úvahy. Algoritmus totiž hľadá len v konkrétnych oblastiach, ktoré si nadefinuje používateľ (lekár) definovaním značiek na prvom snímku. Nemusí tak prehľadávať celé snímky.

Experimenty, ktoré sme vykonali, nám ukázali, že cesta k výsledku je možná pomocou takýchto, sofistikovaných algoritmov. Výsledky experimentov týchto algoritmov boli ale kolísavé. Výrazné ovplyvnenie výsledkov spôsobovalo osvetlenie scény, smer osvetlenia, jeho druh a tiež rýchlosť pohybu značiek na nohe. Taktiež výsledky závisia od prirodzeného svetla v pokusnej miestnosti a od pozadia. Keďže sme nemali potrebné prístroje, nemohli sme zmerať a odporučiť najlepšie charakteristiky pre tieto vlastnosti. Pri experimentoch sme používali osvetlenie vysokosvietivými diódami, ktoré ale vonkoncom nepovažujeme za vhodné, pretože ich intenzita svetla je slabá a až príliš bodová. Na druhej strane svetlo z klasických žiaroviek 60W, bolo príliš rozptýlené. Intenzita ale v tomto prípade postačovala. S týmto druhom osvetlenia sme dosiahli najlepšie výsledky. Nemali sme možnosť odskúšať bodové svetlá, s miernym rozptylom, asi $30-45^\circ$, ktoré predpokladáme, že by mohli byť najvhodnejšie pre tieto účely.

Záver

Na začiatku projektu sme boli trochu zmätený, pretože s danou problematikou nemal nikto z členov nášho tímu žiadne skúsenosti. Hlavnou činnosťou, bolo štúdium materiálov a prác predchádzajúcich tímov a tiež analýza súčasných prístupov, ktoré sa bežne používajú v praxi. Ako sme zistili, táto problematika rozpoznávania pohybu a animácie nie je jednoduchá. Postupne sme sa snažili zlepšovať výstupy a hlavným cieľom bolo rozchodenie celého procesu animácie. Keďže predchádzajúce tímy nedosiahli kompletný proces animácie, čakali nás výzvy v podobe navrhnutia vhodných značiek, úprav scény a navrhnutia a dlhého experimentovania so segmentačnými algoritmami, skúšania rôznych alternatív a posudzovania možností ich využitia. Úpravy animácie a vytvorenie teoretického základu pre epipolárnu geometriu bolo taktiež prioritou. Počas dvoch semestrov sme postupne zisťovali možnosti, stretli sme sa s nedostatkom niektorých zariadení, ako aj výsledkov profesionálnych firiem, ktoré si svoje postupy chránia. Nakoniec sme s danými možnosťami vytvorili vhodný algoritmus, ktorý na snímkoch fungoval a tak isto fungoval aj proces animácie. Celkovo sme vyriešili väčšinu problémov, s ktorými si nevedeli rady predchádzajúce tímy. Našli sme obmedzenia, ktoré spočívajú hlavne po hardvérovej stránke. Použitie dostupných kamier je totiž obmedzené na veľmi pomalý pohyb, pri bežnej chôdzi sú zábery rozmazané a nepoužiteľné. Segmentácia je tiež dlhotrvajúci proces, ktorý vyžaduje výkonné počítače. V našich podmienkach sme dokázali segmentovať len krátke zábery s maximálnou dĺžkou niekoľkých stoviek obrázkov, čo je len zopár sekúnd pri bežných kamerách, a pri rýchlych ani toľko. Napriek tomu si myslíme, že sa nám úspešne podarilo dospieť k novým poznatkom pri analýze chôdze človeka a touto prácou sme umožnili ďalej pokračovať na tomto projekte. V rámci segmentačných algoritmov je ešte možné vylepšovať mnoho drobných detailov, ktoré v rámci experimentovania ešte neboli dotiahnuté úplne dokonca. Napríklad je možné vylepšiť algoritmus segmentácie s využitím lokálnych selekcií o dynamicky sa meniacu veľkosť oblasti pre každú značku a pod. Tiež je možné použiť segmentačné algoritmy na samotnú 3D animáciu (segmentovať záznamy z viacerých kamier), pričom veľkú perspektívu kladieme na použitie epipolárnej geometrie pre tento účel. Pokračovať v tomto projekte má zmysel hlavne vtedy, ak budú k dispozícii aj rýchlejšie videokamery, aby bolo možné animovať reálne pohyby na bežiacom páse.

Referencie

- [1] Animácia a vizuálna analýza chôdze človeka. Tímový projekt FIIT STU. Mareták, J., Matuška, M., Petreje, J., Sás, I. Bratislava 2004.
- [2] Animácia a vizuálna analýza chôdze človeka. Tímový projekt FIIT STU. Lučenič, L. a kol. Bratislava 2005.
- [3] <http://www.sportsci.com/adi2001/adi/services/support/manuals/gait/markersets.asp>
- [4] <http://www.innsport.com/>

- [5] <http://www.mie-uk.com/kin/>
- [6] <http://www.laboratorium.dist.unige.it/~piero/Teaching/Gait/BONTRAGER%20Instrumented%20Gait%20Analysis%20Systems.htm>
- [7] Borghese N. A., Cerveri P. - Calibrating a video camera pair with a rigid. 1999
http://homes.dsi.unimi.it/~borghese/References/Papers/00_SelfCal_PattRecogn.pdf
- [8] Dařílková K. – Modelovanie 3D objektov z fotografii. 2004
<http://fractal.dam.fmph.uniba.sk/~darilkova/RigoroznaPracaFINE.pdf>
- [9] Zhang Z. - Determining the epipolar geometry and its uncertainty: A review. 1996
<http://graphics.cs.msu.su/library/publications/determ.pdf>
- [10] <http://neuron.tuke.sk/~pluchta/Pocitacove%20Videnie/Prednasky/minimovka.doc>
- [11] <http://fractal.dam.fmph.uniba.sk/~darilkova/RigoroznaPracaFINE.pdf>