

ŠIAULIŲ UNIVERSITETAS
TECHNOLOGIJOS, FIZINIŲ IR BIOMEDICINOS MOKSLŲ FAKULTETAS
INFORMATIKOS IR MATEMATIKOS KATEDRA

Ignas Dovydas Strakšas

Informatikos specialybės II kurso nuolatinųjų studijų studentas

Žmogaus kūno rekonstrukcijos kompiuterinis modeliavimas

Computer modeling of human body reconstruction

MAGISTRO DARBAS

Darbo vadovė:

Doc. dr. Sigita Turskienė

Recenzentas:

Prof. Genadijus Kulvietis

Šiauliai

2017m.

„Tvirtinu, jog darbe pateikta medžiaga nėra plagijuota ir paruošta, naudojant literatūros sąrašė pateiktus informacinius šaltinius bei savo tyrimų duomenis“

Darbo autoriaus _____
(vardas, pavardė, parašas)

DARBO TIKSLAI IR UŽDAVINIAI

Tikslas

Sukurti žmogaus kūno rekonstrukcijos kompiuteriu modelį.

Uždaviniai

- Atlikti literatūros analizę žmogaus kūno modeliavimo metodams pasirinkti.
- Išanalizuoti naudojamą kompiuterinės regos įrankį.
- Sumodeliuoti žmogaus kūno rekonstrukcijos modelį.
- Realizuoti sumodeliuotą sistemą.

Darbo vadovė _____
(vardas, pavardė, parašas)

TURINYS

TERMINŲ ŽODYNAS	6
IVADAS.....	7
1. ANALITINĖ DALIS.....	8
1.1. Žmogaus veiksmų sekimo vaizdo medžiagoje metodų klasifikavimai.....	8
1.1.1. Optine tēkme paremti metodai	8
1.1.2. Siluetais paremti metodai	8
1.1.3. Bruoų sekimu paremti metodai	9
1.1.4. Trimačiai ir dvimačiai metodai	9
1.2. OpenCV bibliotekos galimybių analizė.....	10
1.2.1. Automatinis atminties tvarkymas.....	10
1.2.2. Sotinimo aritmetika	10
1.2.3. Fiksuoti tipai ir ribotas šablonų naudojimas	11
1.2.4. Klaidų tvarkymas	11
1.2.5. Pakartotinas kviestinumas	11
1.2.6. Pagrindiniai komponentai	11
1.2.7. Paveikslėlių apdorojimo funkcijos	12
1.2.8. Duomenų išgavimo funkcijos	12
1.2.9. Mašinų mokymosi funkcijos.....	12
1.2.10. Grafinės vartotojo sąsajos kūrimo įrankiai	13
2. PROJEKTINĖ DALIS	14
2.1. Darbinė sritis	14
2.2. Segmentatoriaus apmokymas ir taikymas.....	15
2.3. Aktoriaus šešėlių problemos sprendimas	16
2.4. Segmentatoriaus tikslumo įvertinimo formulė	16
2.5. Rekonstruktorius modelis	17
2.5.1. Projekcijos matricos radimas	18
2.5.2. Trimatė rekonstrukcija.....	19
2.6. Duomenų rinkiniai validavimui.....	21
2.6.1. Reikalavimai renkantis duomenų rinkinius.....	21
2.6.2. Parinkti duomenų rinkiniai.....	21
3. REALIZACIJOS DALIS	23
3.1. Segmentatorius realizacija	23
3.2. Tolesnis segmentatorius tobulinimas.....	23

3.3. Rekonstruktorius	24
IŠVADOS	25
PUBLIKACIJOS	26
LITERATŪRA	27
ANOTACIJA	29
ANNOTATION.....	29
PRIEDAI	30

TERMINŲ ŽODYNAS

KPS - kadrai per sekundę.

Kaukė - dvejetainė dvimatė matrica paveikslėlio daliai pažymėti.

Rekonstravimas, trimatė rekonstrukcija - procesas, kurio metu, panaudojant kelias trimačio objekto formos projekcijas, aproksimuojama originali objekto forma.

Pirmas planas - analizei aktuali vaizdo medžiagos dalis. Priešakyje veikiantis žmogus.

Antras planas - analizei neaktuali vaizdo medžiagos dalis. Scena, pirmo plano užnugaris.

Antro plano modelis - scenos, kurios pirmame plane neveikia aktorius, reprezentacija.

Segregavimas, segmentavimas - procesas, kurio metu kadre esantys taškai išskiriami kaip priklausantys pirmajam ar antrajam planui. Pirmojo plano kaukės sudarymas.

Stereokalibravimas - procesas, kurio metu randamas santykis tarp individualios kameros vaizdo medžiagos kadro ir jame rodomos trimatės erdvės. Projekcijos matricos sudarymas.

Projekcijos matrica - rekonstravime naudojama matrica, kurioje saugomos visos kameros vidinės ir išorinės savybės.

Pikselis - dvimatės matricos elementas; paveikslėlio taškas; taškas, priklausantis plokštumai.

Vokselis - trimatės matricos elementas; taškas, priklausantis tūriui.

ĮVADAS

Kompiuterinė rega ir automatinis atpažinimas yra daug dėmesio sulaukusi modernių tyrimų sritis. Geras šios problemos sprendimas turi dar tik pradedamą tyrinėti pritaikymo potencialą daugybei taikomųjų programų, pavyzdžiui, didelių vaizdo medžiagos archyvų automatinį struktūravimą ir paiešką, lauko ir vidaus objektų stebėjimą ir nuotolinį jutimą, nefizinį žmogaus ir kompiuterio sąveikavimą, gestų atpažinimą ir pažangų vaizdo medžiagos redagavimą.

Šio darbo tikslas - sukurti žmogaus kūno rekonstrukcijos kompiuteriu modelį, pateikiantį žmogaus kūno reprezentaciją, naudotiną realiu laiku ar artimu realiam laikui veikiančioje sistemoje, skirtoje žmogaus veiksmams atpažinti.

1. ANALITINĖ DALIS

1.1. Žmogaus veiksmų sekimo vaizdo medžiagoje metodų klasifikavimai

Metodai, naudojami žmogaus veiksmams vaizdo medžiagoje sekti, gali būti skiriami, remiantis klasifikatoriaus naudojamais deskriptoriais, t.y., kokiais iš vaizdo medžiagos gautais duomenimis sistema remiasi, atskirdama žmogų nuo jo fono. Taip pat šie metodai gali būti išskirti pagal deskriptorių matavimų skaičių į dvimačius ir trimačius metodus.

1.1.1. Optinė tēkme paremti metodai

Optinės tēkmės užduotį galima formuluoti kaip to pačio taškų rinkinio kadre aptikimą kitame kadre. Radus šio rinkinio pozicijos pokytį tarp kadro ir žinant šio pokyčio laiką, galima nustatyti jų greitį. Analizuojant visų kadro optinės tēkmės matmenis, galima sėkmingai klasifikuoti konkrečiuose duomenų rinkiniuose esančius veiksmus. [1]

Taip pat, derinant šį vaizdo medžiagos analizės metodą su bruožinių taškų aptikimu vaizdo medžiagoje (pvz., spalvos analize odos paieškai, kontūrų ar veido aptikimo metodais), galima vertinti šių objektų judėjimą. [1,2]

Tačiau, projektuojant tiriamą kūną dvimačiame kadre, prarandamas atstumas nuo jo iki kameros. Dėl tos pačios priežasties objektai, esantys arčiau kameros, atrodo judantys greičiau nei tokiu pat greičiu judantys objektai, esantys toliau nuo kameros. Todėl optinė tēkme paremti metodai priklauso nuo vaizdo medžiagos įrašymo sąlygų. [2]

1.1.2. Siluetais paremti metodai

Stacionarios kameros atveju žmogaus siluetui iš vaizdo medžiagos išgauti paprastai naudojamas antro plano modelis, kuris sudaromas iš antro plano nuotraukos ar kadro sekos. Šis modelis taip pat gali būti sudaromas ir vaizdo medžiagos srauto analizavimo metu. Tokiu atveju modelis sudaromas iš prieš vertinamą kadro ežusių kadro, tačiau, šiuo atveju, daroma prielaida, kad žmogus juda. Silueto išskyrimas vyksta vertinant, ar einamojo kadro kiekvienas taškas pakankamai gerai apibrėžtas antro plano modelio. [1, 3]

Žmogaus veiksmas gali būti apibrėžiamas kaip laiko atžvilgiu išdėstyta jo silueto seka. Ši seka gali būti atvaizduojama kaip figūra trimatėje erdvėje, kurios du matmenys yra kadro erdvė, o trečiasis - laikas. [1, 3]

1.1.3. Bruožų sekimu paremti metodai

Žmogaus veiksmai gali būti analizuojami, sekant vaizdo medžiagos bruožinius taškus. Analizuojant šių taškų erdvinį išsidėstymą vienas kito atžvilgiu, galima atskirti paprastus veiksmus. Šių taškų pozicijos ir judėjimo greičiai taip pat gali būti panaudojami žmogaus kūno modelio pozos sudarymui. Veiksmai gali būti klasifikuojami remiantis šio modelio dalių pozicijomis, sudaromais kampais ar judėjimu. [1]

Bruožų sekimą apsunkina didelė sekamų bruožų ir žmogaus kūno judėjimo variacija. [4]

1.1.4. Trimačiai ir dvimačiai metodai

Žmogus veikia realioje trimatėje erdvėje, tačiau tradicinės kameros įrašo tik dvimatę šios erdvės projekciją, todėl kompiuterine rege paremta dvimatė analizė neatsižvelgia į visus erdvinio veiksmo parametrus ir tampa priklausoma nuo stebėtojo pozicijos. [2,4,5]

Šiam trūkumui spręsti duomenys išreiškiami trimatėje erdvėje, pasitelkiant daugiau nei vieną kamerą, tačiau trimatė forma išreikštų duomenų analizė priklauso nuo objekto orientacijos žinojimo. Taip pat, ši problema sprendžiama, analizuojant kelių perspektyvų vaizdą. [2,5]

Trimačius metodus apimanti charakteristika yra jų rėmimasis statiniais bruožais, pvz.: forma ar poza. Tiksliausi dvimačiais duomenimis besiremiantys metodai paremti judesio informacija arba jos ir statinių duomenų kombinacija. [2]

1.2. OpenCV bibliotekos galimybių analizė

OpenCV yra atviro kodo kompiuterinės regos biblioteka, parašyta C ir C++ kalbomis, veikianti Linux, Windows ir Mac OS X operacinėse sistemose ir saugoma BSD licencijos. OpenCV Python, Ruby, Matlab sąsajos aktyviai plėtojamos. OpenCV naudojamas IBM, Microsoft, Intel, SONY, Siemens ir Google bei tyrimų centrų Stanford, MIT, CMU, Cambridge, INRIA kompanijų. Šios bibliotekos kodas optimizuotas, palaiko lygiagrečius skaičiavimus per NVIDIA CUDA platformą, todėl išnaudoja daugiabranduolinius procesorius ir yra tinkamas realiu laiku veikiančioms sistemoms kurti. OpenCV kodas moduliškas, t. y., turi kelias bendras ir statiškas bibliotekas.

1.2.1. Automatinis atminties tvarkymas

Funkcijų ir metodų naudojamos duomenų struktūros turi destruktorius, kurie atlaisvina atminties buferius, tik kai to prireikia. Tai reiškia, kad šie destruktoriai ne visada atlaisvina atminties buferius. Jie atsižvelgia į galimą duomenų dalijimąsi. Destruktorius skaičiuoklyje vienetu sumažina nuorodą į struktūros atminties buferį skaičių, kuris atlaisvinamas tik kai skaičiuoklis pasiekia nulį, t.y., kai jokie kiti objektai neberodo į šią struktūrą.

Kai duomenų struktūra kopijuojama, jos duomenys nėra priskiriami kitam atminties buferiui. Nuorodų skaičius padidinamas, kadangi į struktūrą kreipiasi kitas objektas. Norint kopijuoti atminties buferį, atliekama struktūros klonavimo operacija.

OpenCV automatiškai išskiria vietą didžiajai daliai išvesties funkcijų parametru. Šių duomenų dydis ir tipas priklauso nuo išvesties funkcijai paduodamų duomenų dydžio ir tipo bei funkcijoje atliekamų procesų. Kilus dviprasmybėms, šios funkcijos priima papildomų parametru išvesties bruožams nusakyti.

1.2.2. Sotinio aritmetika

OpenCV kompiuterinės regos biblioteka dažnai apdorojami kompaktiškai (8-iais ar 16-a bitų vienam kanalui) koduotų paveikslėlių taškai, kurie turi ribotą reikšmių spektrą. Taip pat, tam tikros operacijos su paveikslėliais, pavyzdžiui spalvų erdvės keitimas, šviesumo ar kontrasto derinimas, aštrinimas ar sudėtinė interpoliacija, gali gauti reikšmes, netelpančias į šį spektrą, ir taip sukelti klaidų. Šiai problemai spręsti visoje bibliotekoje naudojama taip vadinama "sotinio aritmetika" - artimiausių leistinų reikšmių parinkimas.

1.2.3. Fiksuoti tipai ir ribotas šablonų naudojimas

C++ leidžia naudoti šablonus galingoms, efektyvioms ir saugioms duomenų struktūroms kurti ir naudoti. Tačiau platus šablonų naudojimas itin didina kompiliacijos laiką, kodo dydį ir padaro išskirtinai šablonus naudojančią interfeisą ir jo realizaciją sunkiai atskiriamus. Dėl lengvesnio Python, Java ir Matlab (kalbų, kurios neturi arba turi ribotus šablonus) sąsajų kūrimo ir minėtų priešasčių, OpenCV remiasi polimorfizmu ir vykdymo metu atliekamais kvietimais, o ne šablonais.

Masyvai su sudėtingais elementais negali būti konstruojami ar apdorojami naudojant OpenCV. Kiekviena funkcija ar metodas gali naudotis tik tam tikru visų masyvų tipų poaibiu.

1.2.4. Klaidų tvarkymas

OpenCV pranešimams apie kritines klaidas naudoja išimtis. Algoritmas, gavęs tinkamus įvesties duomenis, tačiau, dėl tam tikrų priešasčių, negalintis sėkmingai užsibaigti, grąžina klaidos kodą. Šios išimtys gali būti `cv::Exception` klasės objektai ar jos dariniai. Savo ruožtu `cv::Exception` yra `std::exception` klasės darinys. Todėl jos gali būti grakščiau apdorotos, naudojant kitų standartinių C++ bibliotekų komponentus. Dėl automatinio atminties tvarkymo, visi klaidoje dalyvavę tarpiniai buferiai yra atlaisvinami.

1.2.5. Pakartotinas kviestinumas

Didžioji dalis OpenCV struktūrų yra pilnai pakartotinai kviestinos. T.y., ta pati funkcija, tas pats klasės objekto konstantinis metodas, arba tas pats nekonstantinis kitos klasės objekto metodas gali būti kviečiami iš skirtingų lygiagrečiai veikiančių gijų.

1.2.6. Pagrindiniai komponentai

OpenCV realizuoti dažnai sutinkami baziniai duomenų tipai, jų formatavimo įrankiai ir operatoriai su jais, pvz.: eilutė, taškas, trimatis taškas, skaliaras, dydis (t.y. ilgio ir pločio pora), kvadratas, pasuktas kvadratas, matrica, maža matrica, reta matrica, kompleksinis skaičius bei iteratoriai kai kuriems iš šių tipų. Baziniai duomenų tipai įvardinami pagal šabloną: `CV_bitų-skaičius{U|S|F}C(<kanalų skaičius>)` - čia, U - beženklis sveikasis skaičius, S - sveikasis skaičius, arba F - slankiojančio kablelio skaičius.

OpenCV turi šabloninę klasę `DataType` duomenų tipams kurti, virtualią klasę `Algorithm` su dukterinėmis klasėmis (pvz.: `TermCriteria` - iteratyvių algoritmų užbaigimo kriterijų klasė), kuri skirta kompleksiniams algoritmams realizuoti, bei "gudriąją" rodyklę `Ptr`, kuri turi kelis pranašumus prieš panašią funkciją atliekančią C++11 `std::shared_ptr` klasę (pvz.: Naudojant `Ptr` rodyklę į objektą, vietoje pačio objekto, automatiškai gaunami visi reikalingi konstruktoriai ir priskyrimo operatoriai $O(1)$ sudėtingumu).

OpenCV turi operacijoms su masyvais išskirtą komponentą, kuriame galima rasti aritmetines, logines ir transformacines operacijas, ekstremumų bei kitų bruožų paieškos įrankius. Taip pat OpenCV pagrindiniame komponente realizuotos klasterizavimo, optimizavimo bei duomenų saugojimo XML arba YAML formatais funkcijos.

1.2.7. Paveikslėlių apdorojimo funkcijos

OpenCV turi funkcijas, leidžiančias įvairiai apdoroti dvimačius paveikslėlius, išreikštus matricomis: paveikslėlių filtravimą, geometrines ar spalvų transformacijas, primityvų piešimo ir teksto spausdinimo funkcijas.

1.2.8. Duomenų išgavimo funkcijos

OpenCV turi histogramų sudarymo, normalizacijos, atgalinės projekcijos ir lyginimo funkcionalumą. Taip pat pateikiamas įrankių rinkinys paveikslėlių struktūrinei analizei atlikti - darbui su kontūrais ir daugiakampiais, jų formos momentais. Taip pat realizuoti vieno kadro segmentavimo metodai bei bruožų išskyrimo metodai kontūrų kraštams ir kampams rasti.

Realizuoti metodai objektui sekti vaizdo medžiagoje remiantis antro plano modelio sudarymu, optinės tėkmės ar iš anksto žinomo šablono radimu.

1.2.9. Mašinių mokymosi funkcijos

Mašinių mokymosi algoritmai yra apmokomi pavyzdiniais duomenimis. Patogiam jų pateikimui ir administravimui OpenCV turi struktūras šiems duomenims tvarkyti.

Kiekvienas apmokomasis pavyzdys yra reikšmių rinkinys - vektorius. OpenCV "ml" modelio apmokymo pavyzdžių duomenų struktūra sukonstruota su prielaida, kad visi šie vektoriai turi tą patį skaičių komponentų arba bruožų. Kiekvienas bruožas gali būti rūšiuojamas arba kategorizuojamas.

OpenCV turi plačiausiai naudojamą dirbtinio neuronų tinklo realizaciją - daugiasluoksnių perceptronų tinklą. Šis tinklas susideda iš įvesties sluoksnio, vieno ar kelių tarpinių, paslėptų sluoksnių ir išvesties sluoksnio. Kiekvienas sluoksnis susideda iš vieno ar kelių kryptingai su praeitu ir sekančiu sluoksniu susietų neuronų.

Taip pat realizuota klasė sprendimų medžiui konstruoti. Medis konstruojamas rekursyviai. Visi apmokymo duomenys yra naudojami padalinti šakninę šaką. Sudarant kiekvieną šaką, randama optimali, labiausiai diskriminuojanti klasifikacija. Visi duomenys padalinami gautais klasifikatoriais, tuomet rekursyviai dalinami gauti poaibiai.

Taip pat realizuotas normalus Bayes'o klasifikatorius - modelis su prielaida, kad bruožų vektorius yra normaliai pasiskirstęs, todėl visa duomenų išdėstymo funkcija yra Gauso mišinys, turintis vieną komponentą klasei, ir K-artimiausių kaimynų klasifikatorius, kuris įvertina visus apmokymo duomenis ir nuspėja naujo pavyzdžio rezultatą. Šis rezultatas analizuojamas tam tikrą skaičių artimiausių to pavyzdžio kaimynų, naudojant balsavimą ir sumuojant svorines reikšmes.

1.2.10. Grafinės vartotojo sąsajos kūrimo įrankiai

OpenCV pasižymi funkcionalumu kurti langams dvimatei vaizdinei bei trimatei scenos informacijai atvaizduoti. Šiuose languose automatiškai atvaizduojama pagal duomenų struktūrų tipą interpretuota spalvinių kanalų vaizdinė informacija.

OpenCV taip pat turi funkcionalumą pamatiniams grafinės vartotojo sąsajos elementams konstruoti. Yra funkcijos nusakančios mygtuką, slankiklį, teksto lauką.

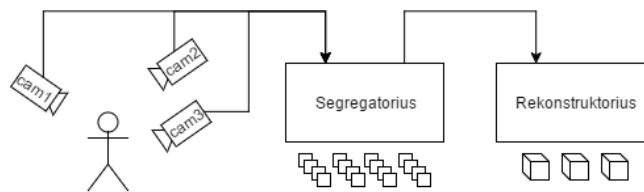
2. PROJEKTIINĖ DALIS

2.1. Darbinė sritis

Kameros sinchronizuotai siunčia vaizdinę informaciją apie sceną. Segmentatorius, remdamasis kiekvienos kameros antro plano modeliu, kuris sudarytas iš scenos, kurioje iš anksto žinoma, kad neveikia joks aktorius, kadru sekos, išskiria pirmąjį ir antrąjį planą. Kitaip tariant, iš kamerų gautų aktoriaus kadru grupių sekų ir iš anksto turimų tų pačių atitinkamų kamerų scenos kadru grupių sekų skirtumų, kiekvienam kadrui sudaroma pirmo plano kaukė. Kiekviena ši kaukė - kameros matomo aktoriaus siluetas.

Kiekvienos kameros projekcijos matrica žinoma iš anksto. Todėl rekonstruktorius, gavęs pirmo plano kaukes, gali jas naudoti projektuojant aktoriaus siluetus atgal į trimatį tūrį. Iteruodamas per kiekvienos kameros kiekvieną aktyvų kaukės tašką, rekonstruktorius, naudodamas to taško koordinatas ir projekcijos matricą, gali rasti visas įmanomas to taško pozicijas trimačiame tūryje tam tikrame rėžyje.

Apskaičiavus visas galimas kiekvieno kameros matomo silueto taško pozicijas trimačiame tūryje, turime to silueto į trimatį tūrį metamą erdvinį kūgį. Suradus kiekvienos kameros silueto erdvinių kūgių susikirtimus, gauname originalaus trimačio kūno aproksimaciją.



Pav. 1 Darbinės srities iliustracija

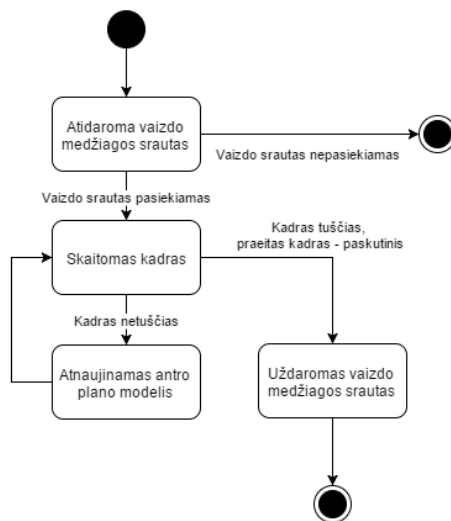
2.2. Segmentatoriaus apmokymas ir taikymas

Dvimate kelių vaizdų informacija paremti metodai dėl mažesnio kompiuterinių resursų reikalavimo ir sąlyginai aukšto patikimumo labiau tinka realiu laiku veikiančioms sistemoms. Didžioji dalis šių metodų remiasi silueto pokyčiais judesiui nusakyti. [2,6]

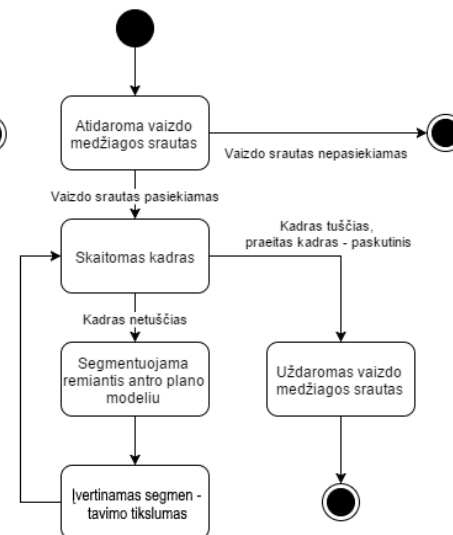
Darant prielaidą, kad kameros pozicijos fiksuotos, taikomas fono segregavimas remiantis antro plano modelio sudarymu. Realiam laiku veikiančioms sistemoms fono segregavimui rekomenduojamas Gauso modelis, kuris gali prisitaikyti prie palaipsniui besikeičiančios scenos. Svarbu pastebėti, kad šis modelis neatsparus dinamiškam fonui pvz., fonui kuriame siūbuoja medžio šakos ir, nors ir yra būdų pritaikyti šį modelį minėtoms aplinkoms, dėl papildomų resursų kaštų ir, tuo pačiu, žemesnio stabilumo, šis sprendimas netinkamas. [3]

Antro plano modeliavimas susideda iš dvejų pagrindinių etapų. Antrojo plano modelio apsimokymo ir atnaujinimo. Pirmajame žingsnyje naudojant kadro sekas, kuriuose rodoma scena, kurioje žinoma iš anksto, kad neveikia joks aktorius, sudaromas pradinis antrojo plano modelis. Kadangi atsižvelgiama į kadro seką, o ne į individualų kadrą, modelis tampa atsparus triukšmui. Antrasis etapas vykdomas naudojant antrojo plano modelį pirmajam planui išskirti. Po kiekvieno kadro įvertinimo einamasis kadras įtraukiamas į antrojo plano modelį. Taip modelis tampa atsparus palaipsniui vykstantiems scenos antro plano pokyčiams (pvz., apšvietos svyravimams) bei statiškiems fono pakeitimams (pvz., objekto pastatymui scenoje), kadangi šį objektą nusakančių taškų aibė pamažu bus įtraukta į antrojo plano modelį. Kitaip tariant – neaktyvus pirmojo plano aktorius pamažu pereis į antrąjį planą.

Posistemė skiriama į apsimokymo (Pav. 2) ir segregavimo (Pav. 3) dalis.



Pav. 2 Apsimokymo posistemė



Pav. 3 Segregavimo posistemė

2.3. Aktoriaus šešėlių problemos sprendimas

Šešėlių atskyrimas nuo pirmo plano yra dažna antro plano segmentavimo problema. Asmeniui judant erdvėje, jo šešėlis taip pat juda. Dėl šio judėjimo šešėlis klaidingai įvardinamas kaip judančiojo aktoriaus dalis. Aptinkant pirmojo plano kaukę panaudotiną trimatėje rekonstrukcijoje, svarbu išskirti pirmajame plane veikiančio aktoriaus šešėlį nuo pačio aktoriaus.

Taško vaizdo medžiaga saugoma RGB formatu atvaizduojama kaip trijų baitų rinkinys, kurių kiekviena reikšmė gali įgyti teigiamą sveikąjį skaičių. Kiekvienas baitais nusako vienos iš pirminių spalvų (raudonos, žalios ir mėlynos) intensyvumą bite. Kai ant taško metamas šešėlis, jis pasidaro tamsesnis nei anksčiau. RGB erdvėje tai reiškia, kad visų šių spalvų kanalų intensyvumas sumažės proporcingai vienas kito atžvilgiu. Tačiau, palyginus matomo kameros kadro taško informaciją su antro plano modeliu, bus matomas spalvinis skirtumas. Todėl šešėlis bus klaidingai įvertintas kaip priklausantis pirmajam planui.

Šešėliui priklausantys taškai gali būti išskirti iš pirmojo plano, naudojant normalizuotą kryžminę koreliaciją. Kad lyginamas matomo kadro taškas būtų įvertintas kaip antro plano modelio taškas su šešėliu, lyginamų taškų chromatinė informacija turi būti tapati, o kanalų šviesumo intensyvumas pakliūti į iš anksto parinktą standartinį nuokrypį.

2.4. Segmentatoriaus tikslumo įvertinimo formulė

Segmentatoriaus *tikslumas* vertinamas pagal kadro ploto wh ir segmentatoriaus gautos dvejetainės kaukės matricos (M) ir duomenų rinkinio pateikto dvejetainės kaukės matricos (V) skirtumo (R) aktyvių taškų sumos Δ skirtumo santykį su kadro plotu wh .

$$R_{w,h} = M_{w,h} \oplus V_{w,h}$$

$$\Delta = \sum_{i=0}^w \sum_{j=0}^h r_{i,j}$$

$$r_{i,j} \in R_{w,h}$$

$$r_{i,j} = 1$$

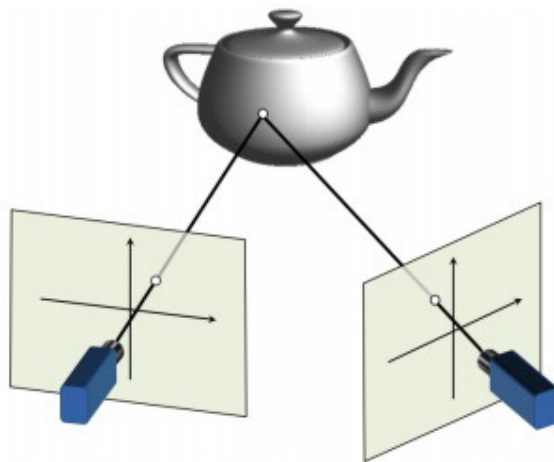
$$\text{tikslumas} = \frac{wh - \Delta}{wh}$$

2.5. Rekonstruktoriaus modelis

Stereo rekonstravimas - trimačio medelio atkūrimas iš kelių dvimačių to modelio projekcijų. Iš anksto žinant fiksuotas stebėjimo kamerų pozicijas bei orientaciją, įmanoma jų filmuojamo aktoriaus projekcijas panaudoti atkuriant trimatį modelį.

Šio modelio atkūrimui aktoriaus silueto, t.y. kameros perspektyva matomo aktoriaus formos kontūro, nepakanka, būtina aktoriaus vaizdo medžiaga. Todėl siluetas naudojamas kaip kaukė pradinei vaizdo medžiagai. Taip išgaunama tik aktorių vaizduojanti, rekonstrukcijai aktuali informacija.

Stereo rekonstrukcija atliekama aptinkant silueto ir jo kaip kaukės dengiamos vaizdo medžiagos bruožinių taškų ir jų kameros perspektyva metamų "spindulių" susikirtimų trianguliacija (Pav. 4)[10].



Pav. 4 Projektijų susikirtimo vizualizacija

2.5.1. Projektijos matricos radimas

Stereokalibravimas reikalingas siekiant sukurti matomo objekto trimatį modelį iš kelių lygiagrečiai veikiančių kamerų teikiamos vaizdo medžiagos. Stereokalibravimas atliekamas randant santykį tarp individualios kameros vaizdo medžiagos kadro ir jame rodomos trimatės erdvės. Šio proceso rezultatas yra projektijos matrica, kurioje saugoma informacija apie kameros poziciją, pokrypį. Turint projektijos matricą, yra žinoma, kur dvimačiame kadre bus matomas kiekvienas trimatės stebimos scenos taškas. [12,11]

Projektijos matrica susideda iš vidinių savybių matricos (angl. *intrinsic*) ir išorinių savybių matricų (angl. *extrinsic*). Vidinių savybių matrica yra nuo kameros savybių priklausančių parametrų matrica, sudaryta iš dydžio iškreipimo koeficientų ilgio ir pločio atžvilgiu, kreivumo koeficiento ir centrinio taško.

$$I = \begin{bmatrix} a_x & s & u_x \\ 0 & a_y & u_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$a_x = f \frac{r_x}{w_x} \quad a_y = f \frac{r_y}{w_y} \quad s = \tan \theta$$

Čia I vidinių savybių matrica, a - pritraukimo koeficientas x ir y atžvilgiu, f - atstumas tarp kameros lęšio ir jutiklio, r - nuotraukos raiška taškais x ir y atžvilgiu, w_y nuotraukos aukštis ir w_x plotis. s iškreipimo koeficientas, o θ kampas tarp x ir y ašių.

Išorinių savybių matricos - tai nuo kameros savybių nepriklausančių parametrų matricos: pokrypio matrica R ir poslinkio matrica T .

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{bmatrix} \cdot \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \cdot \begin{bmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \end{bmatrix}$$

Čia $\theta_x, \theta_y, \theta_z$ - kameros kampai aplink, o t_x, t_y, t_z - poslinkiai pagal x, y ir z ašis.

Pati projektijos matrica yra visų šių matricų sandauga:

$$P = I \cdot R \cdot T$$

2.5.2. Trimatė rekonstrukcija

Trimatė rekonstrukcija yra stereokalibravimo gautų projekcijos matricų m panaudojimas objekto trimačiam modeliui išgauti iš kelių kamerų perspektyvų. Santykis tarp kameros matomo kadro, jos projekcijos matricos ir trimačio objekto išreiškiamas funkcijų sistema:

$$\begin{cases} x_i = \frac{m_{1,1}X_i + m_{1,2}Y_i + m_{1,3}Z_i + m_{1,4}}{m_{3,1}X_i + m_{3,2}Y_i + m_{3,3}Z_i + m_{3,4}} \\ y_i = \frac{m_{2,1}X_i + m_{2,2}Y_i + m_{2,3}Z_i + m_{2,4}}{m_{3,1}X_i + m_{3,2}Y_i + m_{3,3}Z_i + m_{3,4}} \end{cases}$$

Čia X_i, Y_i, Z_i yra trimačio vokselio koordinatės, o x_i, y_i jį atitinkančio divmačio projekcijos taško. Kadangi, turint dvimates projekcijas, ieškomas šių projekcijų apirbėžtas trimatis tūris, iš turimos funkcijų sistemos išreiškiamos vokselio koordinatės. Kadangi dvejų funkcijų nepakanka trimis parametrams nusakyti, trečioji koordinatė imama kaip sutartinė. Taip gaunamos funkcijos:

$$\begin{cases} X_i = \frac{a_1 Z_i + b_1}{c_1} \\ Y_i = \frac{a_2 Z_i + b_2}{c_2} \end{cases}$$

, kai

$$a_1 = (y_i m_{3,3} + m_{2,2})(x_i m_{3,2} + m_{1,2}) - (x_i m_{3,3} + m_{1,3})(y_i m_{3,2} + m_{2,2})$$

$$b_1 = (y_i m_{3,4} + m_{2,4})(x_i m_{3,2} + m_{1,2}) - (x_i m_{3,4} + m_{1,4})(y_i m_{3,2} + m_{2,2})$$

$$c_1 = (x_i m_{3,1} + m_{1,1})(y_i m_{3,2} + m_{2,2}) - (y_i m_{3,1} + m_{2,1})(x_i m_{3,2} + m_{1,2})$$

$$a_2 = (y_i m_{3,3} + m_{2,3})(x_i m_{3,1} + m_{1,1}) - (x_i m_{3,3} + m_{1,3})(y_i m_{3,1} + m_{2,1})$$

$$b_2 = (y_i m_{3,4} + m_{2,4})(x_i m_{3,1} + m_{1,1}) - (x_i m_{3,4} + m_{1,4})(y_i m_{3,1} + m_{2,1})$$

$$c_2 = (x_i m_{3,2} + m_{1,2})(y_i m_{3,1} + m_{2,1}) - (y_i m_{3,2} + m_{2,2})(x_i m_{3,1} + m_{1,1})$$

Kiekvienai Z_i reikšmei galime apskaičiuoti X_i ir Y_i reikšmes, o visi įmanomi sprendimai šioms reikmėms sudaro liniją trimatėje erdvėje. Ši linija yra matomo kadre taško atgalinės projekcijos spindulys į trimatį tūrį.

Prieš panaudojant kamerų matomus vaizdus trimatėje rekonstrukcijoje, iš jų turi būti išgauti tik modeliuojamą objektą nusakantys taškai - pirmasis planas. Šie taškai išgaunami segmentatoriumi.

Panaudojant iš segmentatoriaus gautą esminių taškų kaukę ir projekcijos matricą, galima gauti kiekvieno kameros matomo objekto tašo projekcijos spindulį, kuriame yra visos galimos to taško pozicijos trimatėje erdvėje.

Gautų spindulių visuma sudaro individualios kameros matomo silueto erdvinį kūgį. Trimatė rekonstrukcija įvykdoma, randant skirtingų kamerų matomų siluetų erdvinių kūgių sankirtas. Kiekvienas erdviniui kūgiui priklausantis taškas yra galimas kameros projekcijoje matomo objekto taškas, todėl visų šių kūgių sankirtai priklausantis taškas privalo pakliūti į kamerų projekcijų apibrėžtą objektą.

Vienas trimatis tūris nusako teritoriją, kurioje gali būti matomas objektas remiantis viena kamera. Trimatis tūris sukuriamas iš kelių kūgių sankirtų, todėl gaunama objekto aproksimacija yra didesnė už patį stebimą objektą. Didinant naudojamų kamerų skaičių pasiekama geresnė, tikslesnė aproksimacija. Tačiau papildomų kamerų kūgiams įvertinti reikalinga daugiau kompiuterio išteklių. Didelio kamerų skaičiaus panaudojimas realiu laiku veikiančioje trimatėje rekonstrukcijoje yra nepraktiškas. [1, 2, 11, 13]

2.6. Duomenų rinkiniai validavimui

2.6.1. Reikalavimai renkantis duomenų rinkinius

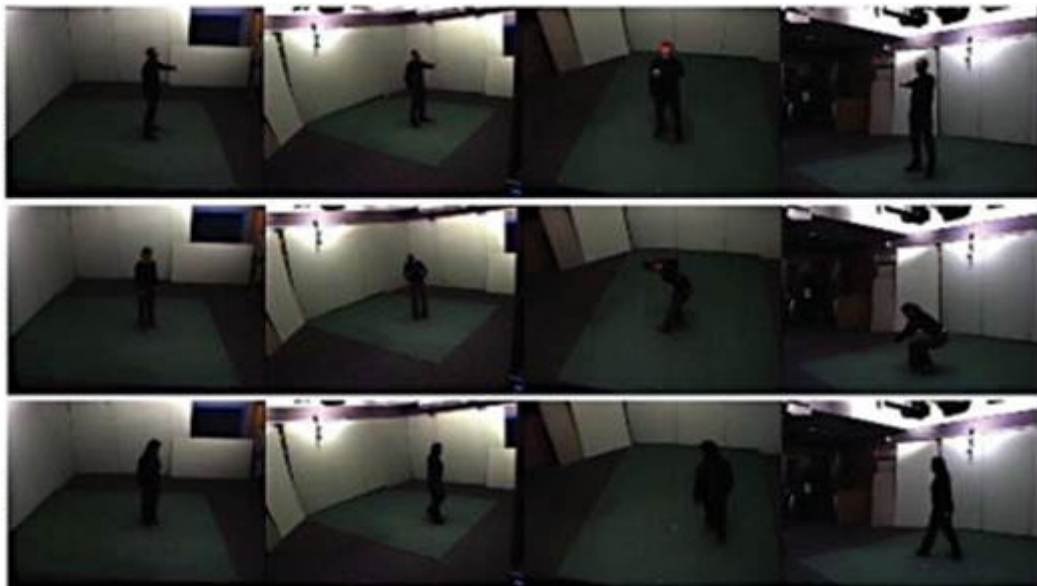
Reikalavimai duomenų rinkiniui keliami, remiantis galutiniais sistemos tikslais. Reikalavimai duomenų rinkiniui:

- duomenys pateikiami kaip vaizdo medžiaga,
- vaizdo medžiagoje rodoma žmogaus veikla,
- patiekiami projekcijos matrica ar jos komponentės,
- vaizdo medžiaga pateikiama su verifikavimo duomenimis.

Duomenų rinkinys renkamas iš Teksaso Universiteto teikiamo žmogaus veiklos vaizdo medžiagos duomenų rinkinių indekso.

2.6.2. Parinkti duomenų rinkiniai

IXMAS(žr. Pav. 4) yra plačiai naudojamas daugiaperspektyvis vaizdo medžiagos duomenų rinkinys, kuriame 11 aktorių 3 kartus atlieka 13-os kasdinių veiksmų seką iš anksto nenusakyta tvarka. Pateikiamas originalus 5-ių kamerų vaizdas PNG formato paveikslėlių seka ir kiekvienam vaizdui atitinkančios priešakyje (pirmame plane) esančių aktorių siluetų kaukės – juodai baltų PBM formato paveikslėlių seka. Taip pat pateikiama tuščio kiekvienos perspektyvos antro plano paveikslėlių seka. Rinkinys filmuotas 23kps, 390x291 taškų raiška.



Pav 5. IXMAS duomenų rinkinio pavyzdys.

MuHAVi yra didelis daugiaperspektyvis žmogaus veiklą vaizduojančios vaizdo medžiagos duomenų rinkinys, kuriame 14 aktorių vaidina 17 veiksmų klasių. Taip pat pateikiama tuščio kiekvienos perspektyvos antro plano paveikslėlių seka, skirta antro plano modelio sudarymui. Medžiaga pateikiama kaip JPG kadru seka. Taip pat pateikiamos kiekvienos perspektyvos priešakyje esančių aktorių siluetų kaukės. Rinkinys filmuotas 25kps, 720x576 taškų raiška, išskyrus aštuntąją kamerą, kurios raiška yra 704x576.

Abu šie duomenų rinkiniai pateikia galutines kiekvienos naudojamos kameros projekcijos matricas, neišskirdami kiekvienų šios matricos komponentų.

	KPS	Raiška	Aktoriai	Veiksmai	Siluetai	Formatas
IXMAS	23	390x291	11	13	Yra	PNG / PBM
MuHAVi	25	720x576	14	17	Yra	JPG

Lent. 1 Duomenų rinkiniai

3. REALIZACIJOS DALIS

3.1. Segmentatorius realizacija

Segmentatorius realizuotas, naudojant OpenCV bibliotekoje realizuotą pagerintą Gauso Mišinio Modelį (GMM) bibliotekoje pavadintą MOG2.

Šiuo metodu, naudojant k Gauso skirstinių, modeliuojamas kiekviena antro plano taško spalvinė informacija laiko atžvilgiu. Gauso mišinių svoriai nusako scenoje matomų spalvų išsilaikymo proporcijas laiko atžvilgiu. Daroma prielaida, kad antrajam planui priklausančios nusakančios spalvos yra tos, kurios išsilaiko scenoje ilgiau ir su mažiau pakitimų.

Šio modelio algoritmas veikia sparčiau nei nemodifikuotas Gauso Mišinio Modelis ir turi galimybę automatiškai nustatyti Mahalonobio ribą, kuri nusako, ar kaukės taškai pakankamai gerai apibrėžti antro plano modelio, t.y., jam priklauso. Taip pat šis modelis leidžia pasirinkti, ar reikalinga atsižvelgti į scenoje aptiktus šešėlius, t.y., įvertinamame kadre matomus antrojo plano taškus, kurie sutampa su antrojo plano modeliu pagal spalvą, bet ne pagal šviesumą. Kadangi numatomas rekonstruktorius neatsižvelgia į šešėlio informaciją ir veikia remiantis tik sekamą objektą nusakančių taškų rinkiniu, šis MOG2 segmentatoriaus funkcionalumas įjungiamas, o gautos kaukės reikšmės normalizuojamos į 0 ir 255 nustatant šešėlio reikšmę į 0, kadangi MOG2 pagal nutylėjimą gautas rezultatas – vieno baido kanalo matrica, kur pikseliai pažymėti kaip 0 nusako antrąjį planą, 127 – šešėlį, o 255 – pirmąjį planą.

Realizuota posistemė verifikuota naudojant IXMAS rinkinio „julien1“ paketą. IXMAS duomenų rinkinio dvejetainės kaukės matricos pateiktos didesne raiška nei originalus kamerų vaizdas, todėl validuojant formatuotos į 390x291 taškų raišką.

Posistemė pasiekia ~0.993 tikslumą, tai yra ~99.3% ir apdoroja daugiau kadrų per sekundę nei pateiktame duomenų rinkinyje.

3.2. Tolesnis segmentatorius tobulinimas

Atsparumas triukšmui gali būti pagerintas naudojant metodus, skirtus apsisaugoti nuo atsitiktinio klaidingo smulkių taškų telkinių klasifikavimo. Naudojant eroziją ir suliejimą galima pašalinti smulkius neatitikimus pirmo plano kaukėje.

Erozija naudojama pašalinti kaukės silueto kontūrus. T.y. visi pirmam planui priskirti taškai, kurie turi bent vieną kaimyną, kuris priskirtas antrajam planui, yra priskiriami antrajam planui. Naudojant šią funkciją galima pašalinti pirmo plano kaukėje matomą „sniegą“ – pavienius klaidingai teigiamai įvardintus taškus, kurie paprastai sukeliama ribą viršijusių šešėlių. Priklausomai nuo erozijos stiprumo (kiek kartų ši operacija atliekama) taip pat pašalinami ir teisingai įvardinti smulkūs telkiniai, todėl šis metodas netinkamas jei kameros matomų sekamų objektų siluetai smulkūs.

Suliejimas yra erozijos priešingybė – silueto kontūrai praplečiami. T.y., visi antrajam planui priskirti taškai, kurie turi bent vieną kaimyną, kuris priskirtas pirmam planui yra priskiriami

pirmam planui. Suliejimas naudojamas pašalinti klaidingai neigiamiems netikslumams silueto viduje. Suliejant šios ertmės pamažu užpildomos.

Šie metodai leidžia pašalinti smulkius netikslumus siluetuose ir pasiekti stabilesnę trimatę rekonstrukciją, kadangi nepaisant formos aštrumo ir smulkmenų praradimo, silueto nuoseklumas yra svarbesnis nei tiksli kraštinė.

3.3. Rekonstruktorius

Rekonstruktorius realizuotas, naudojant OpenCV funkciją `findNonZero` gražinančią kiekvienos kaukės pirmo plano aktyvių taškų koordinates (x_i, y_i) ir juos pateikiančią į projekcijos matricą išreikštą formulių sistemą trimačiam (X_i, Y_i, Z_i) taškui gauti. Kadangi šios trys koordinatės išreikštos dvejomis lygtimis, trečioji koordinatė Z_i yra parenkama kaip papildomas sutartinis diskretus parametras.

Gautos trimatės koordinatės pagal reikšmę saugomos elementų rinkiniuose - taip užtikrinama, kad erdviniam kūgiui priklausantis taškas nebūtų apdorojamas dukart. Radas visuose rinkiniuose esančius taškus, t.y., šių rinkinių susikirtimus, randame sąrašą taškų, priklausančių aktorius kūnui.

Dėl aukšto skaičiavimų apkrovimo kompiuterio resursams, naivus šio trimačio kūno atnaujinimas yra nepraktiškas.

IŠVADOS

1. Išanalizavus žmogaus veiksmų aptikimui naudojamus metodus nustatyta:
 - 1.1. Našiausi žmogaus veiksmų analizės metodai pasitelkia pilną trimatę informaciją ir neapsiriboja keliomis dvimatėmis projekcijomis.
 - 1.2. Našiausi žmogaus veiksmų analizės metodai remiasi trimate rekonstrukcija išgautu žmogaus kūnu.
2. Išanalizavus darbinio įrankio galimybes nustatyta:
 - 2.1. OpenCV biblioteka pritaikyta darbui su vaizdo informacija, orientuota į objektinį programavimą ir paruošta vartotojo klasių kūrimui ir plėtimui.
 - 2.2. OpenCV turi plačiai naudojamų kompiuterinės regos problemų sprendimo įrankių bei dažnai sutinkamų metodų tenkinančių numatomų panaudos atvejų reikalavimus.
 - 2.3. OpenCV yra tinkamas įrankis žmogaus kūno rekonstrukcijos kompiuteriniam modeliui kurti.
3. Po realizuotos programinės įrangos testavimo paaiškėjo, kad:
 - 3.1. Realizuotas segmentatorius pakankamai našus, kad būtų tinkamas naudoti realiu laiku veikiančioje sistemoje.
 - 3.2. Rekonstruktorius nėra pakankamai našus naudoti realiu laiku veikiančioje sistemoje, jei apdorojama kiekviena kadrų grupė.

PUBLIKACIJOS

- Šiaurės Lietuvos kolegijos VII-oje tarptautinėje mokslinėje praktinėje konferencijoje "Verslas, teisė ir informacinės technologijos šiuolaikinėje visuomenėje 2016" skaitytas pranešimas "Žmogaus elgesio stebėjimo modeliavimo problematika".
2016m. vasario 24d., P3-2937.
- Šiaulių Universiteto respublikinėje mokslinėje-praktinėje konferencijoje "Informacinių technologijų iššūkiai kūrybos ekonomikoje" skaitytas pranešimas "Stereokalibravimo ir trimatės rekonstrukcijos realizacijos problematika".
2017m. kovo 17d.

LITERATŪRA

1. An Efficient Approach for Video Action Classification based on 3D Zernike Moments.
I. Lassoued, E. Zagrouba, Y. Chahir.
[Preeiga internetu] link.springer.com/chapter/10.1007%2F978-3-642-22309-9_24
2. Human Action Recognition using Multiple Views: A Comparative Perspective on Recent Developments, Michael B. Holte, Thomas B. Moeslund, Cuong Tran, Mohan M. Trivedi.
[Preeiga internetu] cvrr.ucsd.edu/publications/2011/Holte_HGBU2011.pdf
3. Real Time Robust Human Detection and Tracking System. Jianpeng Zhou, Jack Hoang
[Preeiga internetu] ieeexplore.ieee.org/abstract/document/1565316/
4. Challenges of Human Behavior Understanding.
Albert Ali Salah, Theo Gevers, Nicu Sebe, Alessandro Vinciarelli.
[Preeiga internetu] disi.unitn.it/~sebe/publications/HBU10.pdf
5. Multi-view Human Action Recognition System Employing 2DPCA
Mohamed A. Naiel, Moataz M. Abdelwahab, Motaz El-Saban.
[Preeiga internetu] microsoft.com/en-us/research/wp-content/uploads/2016/02/2WACVPaperID20276_CameraReady.pdf
6. A new pose-based representation for recognizing actions from multiple cameras
Selen Pehlivan, Pinar Duygulu.
[Preeiga internetu] yoksis.bilkent.edu.tr/pdf/files/10.1016-j.cviu.2010.11.004.pdf
7. Monitoring Human Activity - A project of the Artificial Intelligence, Robotics and Vision Laboratory.
University of Minnesota, Department of Computer Science and Engineering.
[Preeiga internetu] mha.cs.umn.edu
8. OpenCV techninis gidas.
[Preeiga internetu] docs.opencv.org/
9. TU žmogaus veiklos vaizdo medžiagos duomenų rinkiniai.
[Preeiga internetu] cs.utexas.edu/~chaoyeh/web_action_data/dataset_list.html

10. 3D Reconstruction from Multiple Images.
Theo Moons, Luc Van Gool, Maarten Vergauwen.
[Preeiga internetu]www.ee.oulu.fi/research/imag/courses/Sturm/moons09.pdf

11. Human motion capturing: Visual hull reconstruction.
Niels Van Malderen.
[Preeiga internetu]www.eavise.be/mastertheses/VanMalderen.pdf

12. Multiple View Geometry in Computer Vision.
R. I. Hartley, A. Zisserman.
[Preeiga internetu]
[cvrs.whu.edu.cn/downloads/ebooks/Multiple%20View%20Geometry%20in%20Computer%20Vision%20\(Second%20Edition\).pdf](http://cvrs.whu.edu.cn/downloads/ebooks/Multiple%20View%20Geometry%20in%20Computer%20Vision%20(Second%20Edition).pdf)

13. Visual Hull. David Schneider [Preeiga internetu]
homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0809/schneider.pdf

ANOTACIJA

Autorius: Ignas Dovydas Strakšas

Tema: Žmogaus kūno rekonstrukcijos kompiuterinis modeliavimas

Šiaulių universitetas 2017

Šio darbo tikslas - sukurti žmogaus kūno rekonstrukcijos kompiuteriu modelį, pateikiantį žmogaus kūno reprezentaciją, naudotiną realiu laiku ar artimu realiam laikui veikiančioje sistemoje, skirtoje žmogaus veiksmams atpažinti.

ANNOTATION

Author: Ignas Dovydas Strakšas

Subject: Computer modeling of human body reconstruction

Šiauliai university 2017

The purpose of this work is to create a computer model for human body reconstruction, which produces a human body representation useful in human action recognition software that operates in real or nearly real time.

PRIEDAI

- Kompaktinis DVD diskas:
 - Sistemos programos projektas
 - Referatas "OpenCV galimybių analizė"
 - Segmentatoriaus veikimo demonstracija