

RAČUNALNIŠKI VID NEKDAJ IN DANES

Franc Solina

Laboratorij za računalniški vid
Fakulteta za računalništvo in informatiko, Univerza v Ljubljani
E-pošta: franc.solina@fri.uni-lj.si
URL: <http://lrv.fri.uni-lj.si/~franc/>

POVZETEK: Podan bo pregled razvoja računalniškega vida od 70-tih let prejšnjega stoletja do danes. Računalniški vid je precej heterogena znanstvena disciplina, za katero ne obstaja niti standardna definicija, niti standardna formulacija, kako naj se rešuje probleme na tem področju. Obstaja cela vrsta specifičnih metod za reševanje ozko definiranih nalog, ki se jih težko generalizira na širši krog aplikacij. Mnogo metod in aplikacij je še na stopnji bazičnih raziskav, toda mnogo metod se na drugi strani že uporablja v komercialnih produktih.

Razloženo bo, kako je računalniški vid povezan z drugimi disciplinami kot so umetna inteligenca, fizika, biologija oziroma psihologija, procesiranjem slik in računalniško grafiko.

Računalniški vid je bil zaradi visokih zahtev po procesorski in spominski kapaciteti nekdanj rezerviran le za področja, ki so zmogla visoke vložke (vojaške, medicinske, industrijske aplikacije). Danes pa je možno zaradi višjih in cenejših procesorskih kapacitet ter cenениh vizualnih senzorjev metode računalniškega vida implementirati že na namiznih računalnikih. Zato je spekter problemov, ki se ga danes lotevamo z metodami računalniškega vida veliko širši in bogatejši. Zaradi večje zmogljivosti računalnikov se razvijajo metode za interpretacijo video sekvenc in metod, ki delajo v realnem času.

Tipične naloge računalniškega vida kot so razpoznavanje in sledenje objektov ter interpretacija scen se danes ne uporabljajo le na tradicionalnih področjih uporabe, na primer v industriji in medicini, temveč segajo danes tudi na tako različna področja kot sta arheologija na eni in zagotavljanje varnosti in nadzor na drugi strani.

1. UVOD

Računalniški vid (angl. computer vision) se je začel razvijati v 70-tih letih 20. stoletja, ko je procesorska moč računalnikov postala dovolj močna, da so lahko začeli procesirati večje zbirke podatkov, kar slike nedvomno so. Osnovni motiv za nastanek računalniškega vida je povezan z idejo umetne inteligence, ki si je zadala nalogo, s pomočjo računalnikov reševati kompleksne probleme, ki jih je sposoben reševati človeški razum. Končni cilj računalniškega vida je torej s pomočjo vizualnih informacij reševati naloge, ki jih s pomočjo vida lahko rešuje človek [8]. Vid je namreč naše najpomembnejše čutilo, ki nam omogoča inteligentno interakcijo z našim okoljem brez fizičnega dotika. Naše vizualno dožemanje okolice poteka navidez tako enostavno in preprosto, saj sam proces, kako vidimo, ni odprt naši samorefleksiji. Zato tudi pojem računalniškega vida nekemu najlažje na hitro pojasnimo tako, da ga postavimo kot nasprotje računalniške grafike. Pri računalniškem vidu poteka procesiranje ravno v obratni smeri kot pri računalniški grafiki (slika 1).



Slika 1: Smer procesiranja pri računalniškem vidu je ravno obraten kot pri računalniški grafiki, saj želimo iz slike pridobiti informacije v simbolični obliki o temu, kaj je na sliki. Računalniška grafika pa s pomočjo začetnih informacij (modelov, simbolov) ustvari sliko.

Danes smo še zelo daleč od tega cilja, saj sistemi računalniškega vida še zdaleč niso tako prilagodljivi in univerzalni kot biološki sistemi. Uspešni, predvsem komercialni sistemi računalniškega vida so namenjeni ozkim in dobro definiranim problemom. Za reševanje takih konkretnih uporabniških problemov se uporabljajo raznovrstne in za posamezne vrste nalog zelo specifične metode [3, 9, 11]. Zato posledično težko govorimo o neki standardni metodologiji, kako reševati naloge računalniškega vida in zato na področju računalniškega vida tudi ne obstajajo univerzalni produkti, ki bi se jih lahko uporabilo za reševanje zelo širokega spektra vizualnih nalog.

2. ZAKAJ SO NALOGE RAČUNALNIŠKEGA VIDA TAKO ZAHTEVNE?

Pri vidnem zaznavanju se običajno neka 3D scena preko optičnega sistema kot sta oko ali kamera projicira na 2D sliko. Pri tej projekciji s 3D na 2D se zgubi toliko informacij, da gledano s povsem matematičnega vidika ni možno poiskati inverzne transformacije. Torej iz 2D slike ni možno zopet pridobiti nazaj enoznačne 3D informacije, ki bi bila koristna za reševanje številnih nalog vizualnega sistema. Da lahko iz 2D informacij kljub temu sklepamo o 3D naravi našega okolja so potrebne dodatne informacije. Te dodatne informacije so lahko v obliki raznih metodologij ali pa tudi geometrijskih modelov. Pri razumevanju človeškega vida, ki še nikakor ni do konca razvozan, ni jasno, v kolikšni meri in za kako zahtevne naloge je potrebno posredovanje zavesti, do kakšne mere pa je procesiranje slike, ki se projicira na mrežnico v očesu odvisno le od vhodnih podatkov.

Ko zajamemo digitalno sliko, so informacije zajete v milijonih slikovnih pikah, ki zahtevajo tudi na milijone spominskih bitov. Na koncu procesiranja pa lahko relevantne rezultate zapišemo običajno z nekaj sto biti. Gre torej za ogromno redukcijo podatkov, ki iz digitalne slike izlušči zahtevano informacijo. Večina slikovnih pik je namreč običajno povsem redundantnih. Pri tej redukciji pa igra osrednjo vlogo segmentacija slike, ki želi slikovne pike združiti v segmente, ki ustrezajo segmentaciji odslikanega realnega sveta. Regije, ki jih želimo identificirati na sliki, naj bi ustrezale predmetom oziroma njihovim delom v realnem 3D svetu.

3. POVEZNE DISCIPLINE

Računalniški vid zaradi svojega osnovnega poslanstva sodi v umetno inteligenco (angl. artificial intelligence). Številne metode, razvite na področju umetne inteligence, se uporabljajo v računalniškem vidu, kot so strojno učenje in nevronske mreže. Po aplikacijski plati pa je računalniški vid bližje robotiki, saj so industrijski roboti in avtonomna vozila danes običajno opremljena s kamerami. V tem kontekstu se računalniški vid pogosto imenuje strojni vid (angl. machine vision) ali robotski vid (angl. robot vision).

Kot možni model delovanja se računalniški vid še vedno zgleduje tudi po bioloških sistemih. Zato je računalniški vid povezan tudi z nevrofiziologijo, ki preučuje biološke vizualne sisteme na organskem nivoju in psihologijo zaznavanja (angl. psychophysics), ki preučuje vizualne sisteme na sistemskem nivoju. Preučevanje oči, nevronov in možganskih struktur posvečenih procesiranju vizualnih signalov je v zadnjih sto letih pripeljalo do dobrega, vendar še ne popolnega razumevanja, kako sistemi vidnega zaznavanja delujejo pri človeku in različnih živalih.

Ker je kamera oziroma oko senzor za svetlobo, lahko zajem in formiranje slike kot rezultata interakcije 3D scene, iluminacije in objektivna kamere preučujemo s povsem fizikalnega vidika. Na ta način se računalniški vid navezuje na fiziko. Številne probleme

računalniškega vida pa lahko preučujemo tudi s povsem matematičnega vidika, saj mnoge rešitve temeljijo na statistiki, optimizacijskih metodah in geometrijskih modelih.

Uporaba istih ali podobnih geometrijskih modelov računalniški vid povezuje tudi z računalniško grafiko. Kreiranje virtualnih modelov realnih scen pa tudi dejanskih oseb je veliko lažje z metodami računalniškega vida.

Z računalniškim vidom je povezano tudi področje digitalnega procesiranja signalov. Mnoge obstoječe vrste procesiranja eno-dimenzionalnih signalov lahko razširimo tudi na slikovne informacije. Digitalno procesiranje slik je pogosto potrebno v začetnih stopnjah procesiranja pri računalniškem vidu, da bi izboljšali kvaliteto slike, ali da bi izločili nizkonivojske značilke, kot so na primer robovi. Rezultat digitalnega procesiranja slik je še vedno slika, računalniški vid pa s pomočjo modelov in drugih predpostavk želi priti do informacij, kaj je na sliki.

Računalniški vid se zopet tesneje povezuje z računalniško grafiko, saj je tako veliko lažje hitro in avtomatično graditi modele realnih scen. Računalniški vid se zato uporablja v filmski in televizijski produkciji pri kreaciji posebnih efektov za združevanje slikovnih informacij iz različnih posnetkov [6].

4. APLIKACIJSKA PODROČJA

Računalniški vid se je začel najprej uporabljati na tistih področjih, ki so prenesla visoko ceno implementacije. Potrebna procesorska moč, spominske kapacitete ter senzorji so bili še v nedavni preteklosti zelo dragi. Med tovrstne aplikacije štejejo v prvi vrsti vojaški sistemi za interpretacijo satelitskih slik, za detekcijo sovražnih vojakov in vozil, za vodenje raket (npr. angl. cruise missiles) in za interpretacijo dogajanja na bojišču (angl. battlefield awareness) na osnovi katerih se lahko sprejemajo strateške odločitve.

Med novejšje aplikacije štejejo avtonomna vozila, med katere štejemo tako podmornice, vozila za ceste in brezpotja (hodeči roboti in roboti s kolesi, avtomobili, tovornjaki, vozila za raziskovanje Marsa), do zračnih vozil (angl. UAV -- unmanned aerial vehicle). Avtonomija je lahko popolna, kar pomeni da ni voznika ali pilota, do tega da sistem računalniškega vida nudi pomoč vozniku v določenih situacijah. Med tipične naloge avtonomnih vozil je, da najdejo ovire pred seboj in se jim umaknejo (angl. obstacle avoidance) ali da na osnovi slik zgradijo model svojega okolja [2]. Pri takih vozilih so še posebej koristni senzorji, ki vidijo na vse strani (angl. omnidirectional, panoramic cameras).

V industriji je vloga računalniškega vida že zelo uveljavljena. Naloge, ki so jih v preteklosti opravljali ljudje, predvsem za kontrolo kvalitete sestavnih delov ali končnih izdelkov, so prevzeli računalniki, saj so za enostavne in ponavljajoče se naloge veliko bolj natančni, hitri in neutrudljivi. Vizualna kontrola se uporablja v računalniški, avtomobilski in farmacevtski industriji, pa tudi v prehrambeni industriji, kjer opazovane komponente kot npr. sadje nima uniformne oblike in barve. Bolj kot je naloga natančno

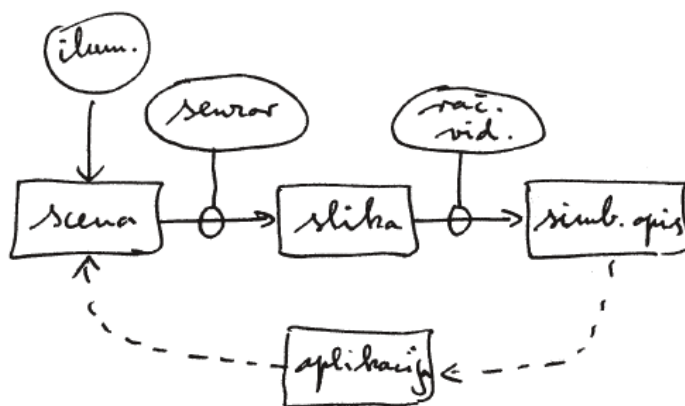
definirana (štetje objektov na tekočem traku, branje serijske številke, ugotavljanje površinskih napak), lažje je sistem programirati.

Medicinske aplikacije računalniškega vida so povezane z vedno večjo vlogo raznovrstnih slik, ki zdravnikom pomagajo pri diagnosticiranju bolezni in v zadnjem času tudi pri načrtovanju in izvajanju operacij. Cilj teh aplikacij je torej iz slik izločiti take informacije, da zdravniki lažje postavijo medicinsko diagnozo posameznemu pacientu. Zato je potrebno velikokrat primerjati sliko nekega pacienta s prototipično anatomsko sliko zdravega človeka [1]. Primeri slik so mikroskopske slike, rentgenske slike, angiografske slike, ultrazvočne slike in tomografske slike. Informacije, ki jih zdravniki želijo so detekcija tumorjev, maligne spremembe tkiv in organov, meritve dimenzij organov, meritve pretoka krvi ipd. Take informacije podpirajo tudi medicinske raziskave o strukturi možganov ali kvaliteti zdravstvene terapije.

Zaradi političnih in družbenih sprememb postajajo vse pomembnejši varnostni sistemi, ki lahko s pomočjo kamer in na osnovi obraza in drugih biometričnih značilnosti avtomatsko preverjajo identiteto ljudi ali opazijo nenavadno obnašanje posameznikov in skupin ljudi.

Nenazadnje postaja nova informacijska tehnologija tudi teren za sodobno umetnost, to je umetnost novih medijev. V vizualni umetnosti video podobe že nekaj desetletij igrajo osrednjo vlogo. Zato ni presenetljivo, da v raznih instalacijah umetniki uporabljajo tudi sisteme računalniškega vida za analizo teh slik, da bi kot nevsiljiv uporabniški vmesnik obiskovalci razstav lahko le s svojo prisotnostjo in gibanjem vplivali na dogajanje [7, 10].

5. KOMPONENTE SISTEMOV RAČUNALNIŠKEGA VIDA



Slika 2: Struktura tipičnega sistema računalniškega vida.

Osnovne komponente sistema računalniškega vida so prikazana na sliki 2. Videz 3D scene je odvisen od iluminacije in položaja ter usmeritve senzorja. Sliko, ki jo pridobimo s senzorjem, običajno je to kamera, procesira programska oprema sistema za računalniški vid. Rezultat procesiranja je nek simboličen opis scene. Kakšen je ta opis, je v največji meri odvisno od namena oziroma vrste aplikacije. Če gre na primer za pobiranje predmetov na tekočem traku s pomočjo robotskega oprijemala, potreben simboličen opis scene sestavlja položaj x - y robotskega prijemala ter orientacija in razmik prstov oprijemala. V sistemu za identifikacijo ljudi na osnovi njihovega obraza pa je npr. Želeni simbolični opis ime človeka na sliki. Ko gre za pobiranje predmetov z robotskega traku s pomočjo robota je sistem računalniškega vida tesno integriran z aplikacijo v zaključeno zanko in mora zato delovati dovolj hitro. V drugem primeru pa aplikacija niti nujno ne predstavlja nek računalniški sistem ampak človeško aktivnost.

5.1 Strojne komponente

Med strojne komponente sistema za računalniški vid štejemo osvetlitev, aktuatorje za manipulacijo scene, senzorje, naprave za digitalizacijo slik (angl. frame grabber), ter procesorje, na katerih je nameščena programska oprema.

Ker je videz neke scene v veliki meri odvisen od osvetlitve, lahko težavnost interpretacije scene v veliki meri zmanjšamo, če poskrbimo za konstantno in čim bolj enakomerno osvetlitev brez senc. To je običajno možno zagotoviti le v zaprtih prostorih, kjer deluje na srečo večina industrijskih aplikacij. Sistemi, ki delujejo zunaj pri naravni osvetlitvi, imajo veliko težjo nalogo. Izgled nekega cestnega odseka, po katerem mora peljati npr. avtonomno vozilo, je zelo različen ob jutranji ali večerni svetlobi.

Tako kot ljudje vrtimo po rokah manjše predmete ali hodimo okoli večjih predmetov, ko jih želimo spoznati, tako aktivni sistemi računalniškega vida (angl. active vision systems) lahko s spremembo smeri pogleda ali manipulacijo scene svojo nalogo olajšajo.

Veliko večino senzorjev v sistemih računalniškega vida predstavljajo črno-bele kamere CCD. Za določene vrste aplikacij se uporablja tudi barvne kamere, vendar pri teh igra vrsta osvetlitve še veliko večjo vlogo. Črno-bele kamere dajejo intenzitetno sliko (angl. intensity image), običajno z 256 nivoji sivine, kar zahteva 8 bitni zapis posamezne slikovne pike (angl. pixel). Velikost slike je odvisna od števila slikovnih pik v x in y dimenziji slike. Barvne kamere pa dajejo podoben zapis za rdeč, zelen in moder kanal.

Laserski senzorji, senzorji s strukturirano svetlobo in stereo kamere dajejo globinske slike (angl. range image), ki jim pravimo tudi $2\frac{1}{2}D$ slike. Globinske slike sestavlja množica 3D točk, ki ležijo na površini slikanih predmetov. Sodobne medicinske naprave pa omogočajo zajem 3D slik notranjih organov in struktur človeškega telesa (tomografija).

Za digitalizacijo slik so bile v preteklosti običajno potrebni posebni vmesniki (angl. frame grabber), danes pa lahko kamere priključimo direktno na enega od vhodov

računalnika. Procesorska moč sodobnih namiznih računalnikov je že tako visoka, da se na njih lahko izvajajo takšne aplikacije računalniškega vida, ki bi v preteklosti zahtevali posebne paralelne računalniške arhitekture. Če so mikroprocesorji na katerih tečejo aplikacije računalniškega vida vgrajeni že kar v kamere, govorimo o pametnih kamerah (angl. smart camera).

Pri nekaterih sistemih je potrebno na začetku kalibrirati kamere, saj so pri nadaljnjem procesiranju potrebne informacije o položaju in notranjih parametrih kamere (npr. goriščna razdalja).

5.2 Programske komponente

Kaj se dogaja z digitalno sliko, ko je enkrat že v računalniku? Najprej je na vrsti nizkonivojsko procesiranje ali predprocesiranje slike. Cilj te faze je zmanjšati ali povsem izločiti šum na sliki in zmanjšati celotno količino podatkov. Te naloge opravljajo razne vrste digitalnih filtrov. Podatke pa lahko zreduciramo z raznimi vrstami transformacij in pragovnimi funkcijami. Primeri so segmentacija na osnovi pragovnih vrednosti (npr. binarizacija slike), računanje lastnih funkcij, računanje optičnega toka ali disparitete v stereo parih slik.

V naslednji fazi izločamo značilke, da bi še bolj zmanjšali količino podatkov. Izločene značilke naj bi bile čim bolj stabilne in čim manj odvisne od osvetlitve, pozicije kamere, orientacije predmetov na sceni, šuma in drugih motenj. Značilke bi lahko razdelili na 2D značilke (robovi, oglišča, konture, regije) in 3D značilke (geoni, superkvadrki [4]). V zvezi z metodami, kako izračunati določeno vrsto značilk iz določene vrste slik, je potekala v preteklosti večina raziskav na področju računalniškega vida.

Zadnja faza procesiranja ali visokonivojsko procesiranje daje končni rezultat, to je simbolično informacijo o vsebini slike. Če gre za razpoznavanje objekta, mora sistem ugotoviti korespondenco med izračunanimi značilkami in značilkami znanih objektov iz podatkovne zbirke in na tej osnovi narediti hipotezo, kateri predmet je na sliki. Hipotezo sistemi običajno verificirajo tako, da identificirani objekt preslikajo nazaj na sliko in še enkrat izračunajo ujemanje (angl. registration) med projekcijo in vhodno sliko.

6. TIPIČNE NALOGE

Najbolj pogoste naloge računalniškega vida so razpoznavanje objektov, sledenje objektom, interpretacija scen ter rekonstrukcija lastnega gibanja kamere iz zaporedja slik (angl. egomotion). Med naloge, ki so sicer nekje na robu zanimanja računalniškega vida, vendar so med najuspešnejšimi aplikacijami, so razpoznavanje črk (angl. optical character recognition) ter tudi branje črtnih kot (angl. bar code).

Med tipične naloge razpoznavanje objektov sodi ugotavljanje prisotnosti tistih objektov na sliki, za katere obstajajo njihovi modeli v podatkovni bazi, ter njihova 3D pozicija. Naloga je še težja, če se medsebojno razmerje členov objekta lahko med seboj spreminja, ko na primer okončine pri človeškem telesu [5]. Preiskovanje slikovnih

podatkovnih zbirk, da bi našli določeno vsebino (angl. content-based image retrieval) tudi sodi med naloge razpoznavanja, ki zaradi vse večjega deleža slikovnih informacij postaja izredno aktualna. Kot konkretno in aktualno nalogo omenimo prepoznavanje človeških obrazov. Za rešitev te naloge moramo najprej ugotoviti, kje sploh so obrazi na sliki. Šele nato lahko detektirane obraze primerjamo z obrazi v podatkovni zbirki, da bi ugotovili njihovo identiteto. Prvi del naloge detekcija obrazov na sliki sodi med tako imenovane naloge kategorizacije, ki je za primer obrazov relativno enostaven, ker so si človeški obrazi po strukturi dokaj podobni. Težji problem pa je na primer poiskati vse vrste avtomobilov na slikah, saj so si avtomobili po obliki in barvi lahko veliko bolj različni kot obrazi.

Med sledenje znanih objektov na sliki sodi na primer sledenje osebi, ki se sprehaja po nakupovalnem centru ali sledenje določenemu avtomobilu, ko vozi po cesti.

Interpretacija scene pomeni rekonstrukcija 3D modela okolja iz ene ali več slik. Primer je gradnja modela okolja iz slik, ki jih zajame kamera na mobilnem robotu.

7. METODOLOGIJE

Rekonstrukcija 3D oblike je bila med prvimi problemi, s katerimi se je ukvarjal računalniški vid, saj je tedaj prevladovalo mnenje, da je 3D rekonstrukcija oblike nujna stopnja za nadaljnje naloge kot je razpoznavanje objektov [8]. Prve metode računalniškega vida v 70-tih letih 20. stoletja so poskušale interpretirati poliedrske scene s pomočjo enumeracije robov in oglišč. Že ti prvi poskusi so pokazali, da sicer povsem pravilno zasnovane metode odpovedo, če so na primer detektirani robovi zaradi šuma v slikah prekinjeni. Zato so raziskovalci poglobljeno preučevali, iz katerih vrst informacij na slikah (tekstura, senčenje, konture, stereo 3D točke, globinske slike) je možno robustno rekonstruirati 3D obliko (angl. shape from X). To nalogo je možno olajšati tudi tako, da se sistem aktivno prilagaja nalogi (angl. active vision). Kot biološki sistemi, se tudi kamera na robotski platformi lahko prilagaja nalogi, tako da spremeni pozicijo in zorni kot, ter prilagodi parametre kamere (zaslonko in goriščno razdaljo).

Za razliko od metod na osnovi geometrijskih značilnk, ki objekt prepoznajo na podlagi lastnosti posameznih značilnk ter njihovih medsebojnih razdalj (npr. iz rekonstruiranega 3D modela), metode na osnovi videza (angl. appearance based methods) uporabljajo za prepoznavo sliko kot celoto. Metode na osnovi videza torej delujejo nad celotno sliko, torej na nivoju slikovnih elementov. Ker velja, da so slikovni elementi slike med seboj zelo redundantni s pomočjo metod kot so metoda glavnih komponent (PCA) lahko redundanco zmanjšamo. Redundanca podatkov je toliko večja, če delujemo na nekem ožjem področju normaliziranih slik.

Na sliko lahko gledamo kot na posamezno točko v visoko-dimenzijskem prostoru. Cilj metod na osnovi videza je predstaviti slike iz visoko dimenzijskega prostora v nižje dimenzijski podprostor, tako da pri tem ohranimo kar največ vhodne informacije slik. Tako predstavljeni podatki so med seboj neodvisni in s tem odpravimo tudi redundanco vhodnih podatkov. Preslikovalno funkcijo, ki preslika slike v podprostor, izračunamo s

pomočjo učne množice dovolj raznolikih slik. Ker so tako predstavljeni objekti sedaj opisani z manj podatki je tudi sama primerjava računsko manj zahtevna. Prednost metod na osnovi videza je tudi ta, da so tako predstavljeni podatki manj občutljivi na šum vhodnih slik, ki se lahko pojavi zaradi naprave same ali pa zaradi delnega prekrivanja objektov. Med najbolj znanimi metodami na osnovi videza spadajo še linearna diskriminantna metoda (LDA) in metoda neodvisnih komponent (ICA). Metoda LDA poskuša najti takšen podprostor, ki med seboj najbolje loči posamezne razrede objektov. Metoda ICA, pa poskuša poleg statistične odvisnosti drugega reda, podatke ločevati tudi na podlagi višje nivojskih statističnih odvisnosti.

8. ZAKLJUČEK

Računalniški vid je aktivno raziskovalno in aplikacijsko področje. Za uspeh na tem področju je potrebno veliko izkušenj, saj rešitve, ki so v teoriji ali na papirju odlične, zaradi šuma na slikah in drugih mejnih vplivov kaj hitro odpadejo.

Biološki sistemi vida še vedno dajejo navdih za reševanje problemov, sama implementacija teh rešitev pa običajno zelo odstopa od bioloških rešitev, katerih delovanje velikokrat samo po sebi še ni niti do konca razumljeno.

Uspešni sistemi računalniškega vida skušajo problem, ki ga rešujejo čimbolj natančno omejiti oziroma ga definirati. S pravilno izbrano osvetlitvijo, s primernim kontroliranjem konteksta (samo en objekt na sceni, v primerni orientaciji itd.) je možno probleme poenostaviti oziroma zagotoviti veliko boljše rezultate.

LITERATURA

1. R. Bajcsy, S. Kovačič (1989), Multiresolution elastic matching, *Computer Vision, Graphics and Image Processing*, vol. 46, str. 1-21.
2. O. Faugeras, Q.-T. Luong, T. Papadopoulos (2001), The Geometry of Multiple Images : The Laws That Govern the Formation of Multiple Images of a Scene and Some of Their Applications, *MIT Press*.
3. D. A. Forsyth, J. Ponce (2002), Computer Vision: A Modern Approach, *Prentice Hall*.
4. A. Jaklič, A. Leonardis, F. Solina (2000), Segmentation and Recovery of Superquadrics, *Computational imaging and vision*, vol. 20, Kluwer.
5. J. Krivic, F. Solina. Part-level object recognition using superquadrics (2004), *Computer vision and image understanding*, vol. 95(1), str. 105-126.
6. A. Leonardis, F. Solina, R. Bajcsy (Eds.) (2000), Confluence of Computer Vision and Computer Graphics, *NATO Science Series 3. High Technology*, vol. 84, Kluwer.

7. T. Y. Levin, U. Frohne, P. Weibel (Eds.) (2002). CTRL [SPACE], Rhetorics of Surveillance from Bentham to Big Brother, *MIT Press*.
8. D. Marr (1982), Vision. A Computational Investigation into the Human Representation and Processing of Visual Information, W.H. Freeman, 1982.
9. L. G. Shapiro, G. C. Stockman (2001), Computer Vision, *Prentice Hall*.
10. F. Solina (2004), 15 seconds of fame, *Leonardo*, vol. 37(2), str. 105-110, 125.
11. E. Trucco, A. Verri (1998), Introductory Techniques for 3-D Computer Vision, *Prentice Hall*.