

Od računalniškega vida k umetnemu spoznavnemu vidu *

Danijel Skočaj in Aleš Leonardis

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

Tržaška 25, 1001 Ljubljana, Slovenija

danijel.skocaj@fri.uni-lj.si

From computer vision towards artificial cognitive vision

In this paper we briefly describe an emerging new scientific field of cognitive vision. We briefly present some main characteristics of the mature and recognised field of computer vision and show some motivation for its natural development into cognitive vision. We underline some characteristics of cognitive vision systems, which make them different from the classical machine and computer vision systems. We also present some typical applications of cognitive vision systems and cognitive systems in general and indicate several possibilities for their employment.

1 Uvod

Računalniški vid se je v preteklih desetletjih že dodobra uveljavil kot pomembna znanstvena disciplina [1]. Velik napredek je bil narejen na raznih področjih računalniškega vida, veliko problemov je bilo zadovoljivo rešenih. V industrijskih obratih, medicini in v vsakdanjem življenju je nameščenih veliko sistemov računalniškega vida, ki v predvidljivih pogojih uspešno delujejo. Lahko rečemo, da so mnoga področja računalniškega vida prišla v zrelo fazo, ko so trenutne zmogljivosti sistemov poznane, razvoj posameznih komponent dokaj predvidljiv, pričakovanja pa realna. Mnoga področja so že zelo dodelana in uporaba rešitev v praksi je zanesljiva ter razširjena.

Zgornja opažanja veljajo predvsem za specifična področja računalniškega vida s točno določenim problemom ter omejenim področjem delovanja. Ob predpostavki, da je okolje, v katerem takšen sistem računalniškega vida deluje, nadzorljivo in je nabor možnih vhodnih signalov nekako omejen in nadzorovan, lahko takšni sistemi zelo zanesljivo delujejo. Večje težave pa so pri razvoju bolj splošnih sistemov, ki naj bi delovali v ne najbolj predvidljivih pogojih in bi se znali odzvati na nenadne nenavadne spremembe

v okolici. V takšnih primerih preizkušene nizkonivojske rešitve računalniškega vida¹ ne zadostujejo več. Potrebna je nadgradnja klasičnih rešitev in razvoj višjenivojskih sistemov računalniškega vida². V ta namen se pogosto uporabljajo tudi metode umetne inteligence, ki pa ponavadi na vходу zahtevajo čiste simbolne opise. Teh pa računalniški vid ne more v popolnosti zagotavljati, zato mora za učinkovito kombinacijo teh metod priti do ustreznih prilagoditev.

Tako se je v zadnjih letih začela razvijati znanstvena disciplina, kateri se je prijelo ime *umetni spoznavni vid* (umetni kognitivni vid, ang. cognitive vision)³. Lahko bi rekli, da gre pri tem za povezovanje računalniškega vida z umetno inteligenco in kognitivnimi znanostmi. Cilj ni samo videti, ampak tudi razumeti kar vidimo (oz. vidi umetni agent). Želja je torej približati sposobnosti umetnih agentov človeškim sposobnostim vizualnega zaznavanja in spoznavanja.

Seveda pa te želje in cilji niso novi in tudi raziskave na tem področju se niso začele s pojavom termina umetni spoznavni vid. Klasični računalniški vid in umetni spoznavni vid se v marsičem prekrivata in veliko raziskav na področju visokonivojskega računalniškega vida je bilo opravljenih tudi v preteklosti. Razlika je ta, da se je v zadnjih letih povečala potreba po intenzivnejših raziskavah na tem področju, zato se je tudi izoblikovala nova disciplina, ki je bolj interdisciplinarno naravnana in ki skuša s sodelovanjem računalniškega vida, umetne inteligence, kognitivnih znanosti, pa tudi drugih sorodnih področij narediti konkreten korak naprej pri razvoju sistemov spoznavnega vida in umetnih spoznavnih sistemov nasploš.

Cilj tega članka je osvetliti ta korak in željo po kvalitnem preskoku. Prikazali bomo dosednji razvoj, ki pelje v to smer, tako na primeru raziskovalnih izizov kot tudi končnih aplikacij, in poskušali predvideti nadaljnji razvoj na tem področju.

¹Procesiranje slik, izločanje lokalnih značilnic, izračun osnovnih značilnosti (barva, oblika, tekstura,...), ipd.

²Razumevanje ter interpretacija slik, naprednejše predstavitve, semantični opisi, ipd.

³Iz SSKJ: *spoznaven, kognitiven*: nanašajoč se na spoznavanje, spoznanje; *spoznavati*: na osnovi zaznav, podatkov in umske dejavnosti prihajati do poznavanja česa in vedenja [2].

* To delo so deloma podprli MVZT (Raziskovalni program *Računalniški vid-1539-506*), EU projekta *CoSy* in *MOBVIS* ter projekt *CONEX*.

2 Od računalniškega vida ...

V tem razdelku bomo naredili zelo hiter oris različnih področij računalniškega vida. Naš namen seveda ni narediti podroben pregled razvoja računalniškega vida niti ne natančno razčleniti posamezna področja te discipline. Radi bi samo izpostavili nekaj tipičnih problemov in predstavili nekatere uspešne aplikacije ter pokazali v čem se novonastajajoča disciplina umetni spoznavni vid razlikuje od klasičnega pojmovanja računalniškega vida.

Z računalniškim vidom so se resneje začeli ukvarjati v 60-tih letih prejšnjega stoletja, ko je nivo razvoja računalnikov in druge strojne opreme to omogočil. Ker je bila ta oprema še vedno dokaj enostavna in počasna, so prve raziskave v glavnem potekale na poenostavljenih prizorih s predmeti preprostih oblik. Raziskovalci so sistematično začeli raziskovati principe nastanka slike in fizikalnih lastnosti procesa vizualnega zaznavanja. Naredili so se prvi koraki pri analizi slik (segmentacija, izločevanje enostavnih značilnic - robovi, črte). Z večjo dostopnostjo barvnih kamer se je začelo proučevanje barv in odbojnih lastnosti predmetov. V 80-tih letih je bil velik poudarek na rekonstrukciji prizorov, stereo vidu, pridobivanju oblike predmetov iz različnih modalnosti, analizi gibanja. Na področju razpoznavanja predmetov so se do takrat ukvarjali predvsem z enostavnimi predmeti, ki se jih je dalo modelirati z geometrijskimi modeli. V 90-tih so se bolj uveljavili modeli, ki temeljijo na videzu predmetov (ang. appearance-based) ter paradigma glediščno-osnovanega vida (ang. view-based).

Vzporedno s temi raziskovalnimi izzivi je potekal razvoj na konkretnih aplikativnih domenah - v biomedicinskih sistemih, biometričnih sistemih (razpoznavanje obrazov, prstnih odtisov), optičnem razpoznavanju znakov, enostavnemu video nadzoru. Zelo se je razširila uporaba sistemov t.i. strojnega vida, ki bi ga lahko označili kot inženirsko disciplino, ki uporablja funkcije nizkonivojskega računalniškega vida za učinkovito, hitro in zanesljivo delovanje konkretnih sistemov na omejenih domenah v nadzorovanih pogojih.

Ugotovimo lahko, da so največji uspehi računalniškega vida doseženi ravno na področjih, ki se ne zelo oddaljijo od nivoja zaznavanja - zajemanja slike. Podobno je tudi z razumevanjem zaznavno/spoznavnih procesov pri človeku. Lahko bi rekli, da so biološki oz. fizikalni aspekti zajemanja vizualne informacije in obdelava vizualnih signalov tako pri človeku kot pri kameri dovolj dobro razumljeni (leča/objektiv, retina/slikovna ravnina, čepnice in paličnice/CCD elementi, zgodnje stopnje vidne skorje/filtri), ravno tako problemi računalniškega vida ter njihove rešitve, ki izhajajo neposredno iz tega (optično merjenje, 3-D rekon-

strukcija...). Precej manj pa so razumljeni spoznavni procesi, ki potekajo v človeških možganih na osnovi informacij, ki jih dobijo iz vizualnega zaznavnega sistema. Podobno velja tudi za sisteme umetnega spoznavnega vida. In ravno z boljšim razumevanjem in razvojem bolj zanesljivih metod umetnega spoznavnega vida lahko pričakujemo kvalitativen preskok na tem področju.

3 ... k umetnemu spoznavnemu vidu

Trend povečevanja povpraševanja po razvoju na področju visokonivojskega računalniškega vida se kaže že od preloma stoletja naprej. Tako se je npr. na področju razpoznavanja predmetov središče raziskav premaknilo v smeri kategorizacije predmetov, pri čemer naloga sistema za razpoznavanje ni le razpoznati že prej videne predmete, ampak tudi njim sorodne, ki spadajo v isto kategorijo. Ta naloga seveda zahteva precej višji nivo abstrakcije in gradnja modela je precej bolj zahtevna. Poudarja se tudi vloga učenja, torej samodejno pridobivanje znanja o predmetih oz. prizorih, ki jih senzorski sistem zajema. Velik poudarek je na integraciji različnih modalnosti, na povečanju robustnosti in splošnosti sistemov računalniškega vida. Ta razvoj je pripeljal do oblikovanja nove discipline - umetnega spoznavnega vida.

Seveda je meja med klasičnim računalniškim vidom in umetnim spoznavnim vidom zelo zabrisana. Tudi veliko raziskav, ki so bile narejene vse od prvih začetkov računalniškega vida (pa tudi umetne inteligence) sodi v domeno umetnega spoznavnega vida. Kljub temu pa bomo poskušali ti dve področji približno razmejiti in pokazati razlike najprej z izpostavljanjem različnih poudarkov na konkretnih raziskovalnih izzivih, nato pa tudi na primeru aplikacij. Pri tem se bomo naslonili na belo knjigo umetnega spoznavnega vida *A Research Roadmap of Cognitive Vision* [3], ki so jo oblikovali raziskovalci s tega področja in poskuša zajeti bistvo te discipline in načrtati smernice za njen razvoj.

3.1 Lastnosti sistemov umetnega spoznavnega vida

Kot je rečeno v [3], je bil izraz umetni spoznavni vid uveden, da bi zajel poskuse za doseg bolj robustnih, prožnih in prilagodljivih sistemov računalniškega vida z njihovo nadgradnjo s spoznavnimi zmožnostmi: sposobnostjo učiti se, prilagajati se, primerjati alternativne rešitve, in celo razvijati nove strategije za analizo in interpretacijo.

Ključna lastnost sistema umetnega spoznavnega vida je sposobnost robustnega in zanesljivega delovanja tudi v nepredvidljivih okoliščinah. Tak sistem

mora biti sposoben predvideti dogodke in temu ustrezno prilagajati njegovo delovanje. V idealnem primeru bi moral biti vizualen spoznavni sistem sposoben razpoznavanja in prilagajanja na nove variacije v trenutni vidni okolici, posploševanja v nove kontekste in aplikativne domene, interpretacije obnašanja in namenov s ciljem predikcije prihodnjih konfiguracij vidne okolice in komuniciranja razumevanja okolice z ostalimi sistemi vključno s človekom. Sistem umetnega spoznavnega vida naj bi težil k postavljenemu cilju, se prilagajal na nepredvidene spremembe v vidni okolici in bi lahko predvidel pojavljanje novih objektov in dogodkov. Te sposobnosti naj bi pridobil preko učenja semantičnega znanja (t.j., kontekstualiziranega razumevanja oblike, funkcije in obnašanja), preko ohranjanja znanja o okolici, o sebi samem in o njegovih odnosih z okolico, ter preko razmišljanja o predmetih in dogodkih v okolici.

V beli knjigi *A Research Roadmap of Cognitive Vision* [3], je nanizanih kar nekaj raziskovalnih temeljev na katerih naj bi temeljili sistemi umetnega spoznavnega vida: vizualno zaznavanje, arhitektura sistema, predstavitev, spomin, učenje, razpoznavanje, razmišljanje in sklepanje, načrtovanje, komunikacija ter povratna zanka zaznavanje-odziv. Pri vsem tem igra vizualno učenje zelo pomembno vlogo. Obseg vizualnih podatkov, ki so umetnemu agentu na voljo (in jih mora obdelovati), je namreč odločno prevelik, da bi se znanje o teh podatkih zajelo v ustrezne modele zgolj s programiranjem in ročnim specificiranjem določenih lastnosti. Umetni vizualni spoznavni sistem si mora veliko večino znanja čimbolj samodejno pridobiti z učenjem. Na začetku mora vsebovati določeno stopnjo predefiniranih (prirojenih) funkcionalnosti, mora pa biti nato sposoben razvijati nove koncepte in prilagajati se spremembam. Pomembno je torej, da je učenje kontinuiran, nenehen, brezkončen proces, ki neprestano izboljšuje sposobnosti spoznavnega sistema.

3.2 Aplikacije umetnega spoznavnega vida

Da bi malo bolj konkretizirali teoretično predstavitev umetnega spoznavnega vida bomo v nadaljevanju prikazali nekaj potencialnih aplikacij ter pokazali kako se le-te razlikujejo od klasičnih aplikacij računalniškega vida.

Športne igre. Vzemimo za primer spremljanje nogometne tekme. Že sedaj obstajajo sistemi, ki znajo mozaičiti slike z več kamer in jih združevati v enotno sliko, ki pokriva celotno igrišče. Taki sistemi znajo tudi uspešno slediti igralcem in žogi, računati razne statistike, itn. To so klasični problemi računalniškega vida. Pred precej drugačen problem pa se postavimo, če od sistema zahtevamo tudi interpretacijo igre. Če želimo, da sistem razume kaj se dogaja, da ve katera

ekipa napada, kdaj je priložnost za gol, kdaj je bil dosežen zadetek in kdaj storjen prekršek. Če želimo torej sistem, ki bi lahko igral vlogo avtomatskega televizijskega komentatorja. Takšni problemi so precej bolj zahtevni in spadajo v domeno umetnega spoznavnega vida.

Promet. Podobno velja za spremljanje prometa s statično kamero. Klasične aplikacije računalniškega vida znajo samodejno slediti avtomobilom na sliki, šteti promet in delati razne druge statistike. Aplikacije obogatene s prvinami umetnega spoznavnega vida pa poleg tega tudi poskušajo razumeti dogajanje na cesti in predvideti prihodnje dogodke. Tako lahko s spremljanjem vozil, ki vozijo v bližini nekega avtomobila, uspešno predvidijo njegovo prihodnje ravnanje ter na ta način tudi zaznajo nenavadno obnašanje na cesti.

Video nadzor. Trg s sistemi za video nadzor se iz leta v leto povečuje in samodejni sistemi za video nadzor pridobivajo na pomembnosti. Že zelo veliko takšnih sistemov je nameščenih po raznih javnih in zasebnih stavbah in zunanjih površinah. Ti sistemi pa so pogostokrat zelo enostavni – sposobni so le zaznati gibanje, mogoče tudi slediti objektom oz. osebam v vidnem polju. Te sisteme bi se dalo zelo lepo nadgraditi s prvinami spoznavnega vida – lahko bi bili sposobni samodejnega razpoznavanja sumljivega obnašanja, učinkovitega razpoznavanja konkretnih oseb na slikah, lahko bi bili bolj splošni, prilagodljivi, s samodejnim učenjem bi se lahko hitro in enostavno samo-konfigurirali in prilagodili na nove situacije.

Preverjanje kakovosti v industriji. V industrijskih obratih je že nameščenih zelo veliko sistemov strojnega vida, ki učinkovito opravljajo predvsem kvantitativne meritve za potrebe optičnega nadzora kakovosti in sorodnih nalog. Z uvedbo spoznavnega vida bi lahko opravljali tudi kvalitativne meritve, estetske preglede (tekstil, keramične ploščice, itn.). Ravno tako bi lahko bili sistemi bolj prijazni, bolj prilagojeni željam posameznih uporabnikov, na enostaven način (z učenjem ter povratno zanko) bi se jih lahko dalo učinkovito rekonfigurirati ipd.

Iskanje po slikovnih podatkovnih zbirkah. Večina obstoječih sistemov za razpoznavanje objektov je sposobnih razpoznati samo tiste objekte, ki so bili predstavljeni v fazi učenja. Tudi sistemi za iskanje po slikovnih podatkovnih zbirkah so precej omejeni in so sposobni iskanja le tistih slik, ki imajo podobne nekatere enostavne značilnosti (recimo barvo, mogoče teksturo, ipd.). Ne obstajajo pa še metode, ki bi znale učinkovito kategorizirati predmete, ki bi znale recimo poiskati vse slike na katerih so prikazani psi (ki so lahko zelo različni). Z razvojem takšnih metod bo omogočeno dejansko razumevanje in semantično označevanje slik in iskanje po slikovnih ter video podatkovnih zbirkah.

Film, TV in zabava. Na področju zabave se pojavljajo prenekateri možnosti uporabe umetnega spoznavnega vida – od sistemov za samodejno označevanje, indeksiranje in iskanje po video podatkovnih zbirkah, preko sistemov za detekcijo določenih dogodkov ali predmetov, do nudenja dodatnih informacij gledalcu (samodejno pridobljenih v realnem času med neposrednim prenosom).

Raziskave vesolja in vojaški sistemi. Umetni spoznavni sistemi bi bili zelo uporabni v aeronavtiki in seveda v raziskavah vesolja, kjer so avtonomni mobilni roboti, ki se znajo inteligentno gibati po popolnoma neznanem in nepredvidljivem okolju in samostojno ter hitro odreagirati na nepredvidljive dogodke, nujno potrebni. Seveda pa so avtonomni in decentralizirani inteligentni sistemi, sposobni samostojnega odločanje v neznanem, nepredvidljivem in spreminjajočem se okolju, izrednega pomena tudi za vojsko.

3.3 Umetni spoznavni sistemi

Za večino zgoraj navedenih aplikacij velja, da je umetni spoznavni vid le en del večje celote – spoznavnega sistema [4]. Za učinkovito in zanesljivo delovanje avtonomnega umetnega agenta je namreč potrebno zagotoviti učinkovito povezavo med različnimi komponentami spoznavnega sistema. Za povečanje robustnosti ima tako tak sistem lahko več senzorjev s katerimi zaznava okolico (vizualni, haptični, infrardeči, akustični, globinski senzorji, itn.). Pri utelešenih (ang. embodied) umetnih agentih je zelo pomemben robotski del (aktuatorji, robotske roke, mobilne platforme). Veliko vlogo pri načrtovanju in sklepanju igrajo tudi metode umetne inteligence. Nenazadnje, za učinkovito komunikacijo (verbalno in neverbalno) s človekom (pa tudi med agenti samimi) je potreben tudi prijazen in zmožljiv način komuniciranja, ki je čim bližje človeškemu naravnemu jeziku. Tu igrajo pomembno vlogo lingvistične metode.

Združevanje vseh teh komponent v delujoč sistem je vse prej kot lahko delo. Pri tem si lahko pomagamo z nekaterimi dognanji ved, ki proučujejo do sedaj edini zares delujoči spoznavni sistem – človeka (kognitivne znanosti, psihologija, psihofiziologija, nevroznost, nevrofiziologija, itn.). Za razvoj umetnih spoznavnih sistemov je torej potreben zelo multidisciplinarni pristop in sodelovanje raziskovalcev z različnih področij. Ker pa je delovanje kateregakoli agenta v realnem okolju zelo odvisno od uspešnega procesiranja informacij, ki jih agent zaznava, in je med vsemi vrstami informacij daleč najbolj pomembna ravno vizualna (70% informacij iz okolja dobi človek v vizualni obliki), igra umetni spoznavni vid ključno vlogo pri razvoju umetnih spoznavnih sistemov.

4 Zaključek

V tem članku smo poskušali opisati novo znanstveno disciplino, ki se poraja v zadnjih letih – umetni spoznavni vid. Na kratko smo okarakterizirali računalniški vid kot že uveljavljeno disciplino in pokazali razloge za njeno nadgradnjo. Izpostavili smo nekatere značilnosti sistemov umetnega spoznavnega vida po katerih se razlikujejo od klasičnih sistemov strojnega in računalniškega vida, ter te razlike demonstrirali tudi na izbranih primerih aplikacij. Pokazali smo, da igra umetni spoznavni vid osrednjo vlogo v umetnih spoznavnih sistemih.

Naj ta članek zaključimo z opisom zelo ambicioznega a še precej oddaljenega cilja raziskovalcev s področja umetnega spoznavnega vida in spoznavnih sistemov nasploh – zgraditi pravega kognitivnega asistenta. Tak asistent naj bi bil sposoben samostojno raziskati okolico in zgraditi ustrezen zemljevid po katerem bi se v prihodnje znal orientirati in navigirati po prostoru. Ravno tako bi se bil sposoben naučiti ter nato razpoznavati in kategorizirati predmete ter razumeti njihove namene in funkcije. Znal bi tudi razpoznavati in razumeti akcije v njegovi okolici, ter se te akcije naučiti in ponoviti. Tak asistent bi tudi znal verbalno in neverbalno komunicirati z ljudmi in drugimi roboti. Zaznaval bi nove situacije in ustrezno odreagirati nanje. Neprestano bi se učil in posodabljal svoje znanje ter sposobnosti in temu ustrezno prilagajal svoje vedenje.

Čeprav ljudje vsa ta opravila (bolj ali manj) z lahkoto opravljamo, je z znanstvenega in inženirskega vidika zgraditi tak umeten sistem zelo zahtevna naloga. Raziskave na področju umetnih spoznavnih sistemov in še posebej umetnega spoznavnega vida nas peljejo v to smer. In čeprav je pot do končnega cilja še dolga, pa v vmesnem času nastajajo in se razvijajo mnoge komponente, ki bodo našle pot do takojšnje uporabe v različnih praktičnih aplikacijah.

Literatura

- [1] D. Forsyth and J. Ponce. *Computer Vision A Modern Approach*. Prentice Hall, 2003.
- [2] *Slovar slovenskega knjižnega jezika*. DZS, Ljubljana, 1995.
- [3] D. Vernon et al. A research roadmap of cognitive vision. Technical report, ECVision, http://www.ecvision.org/research_planning/Research_Roadmap.htm, February 2005.
- [4] Cognitive Systems for Cognitive Assistants - CoSy, <http://cognitivesystems.org/>