Стручни критички осврт

Иван Т. Ћирић, *Inteligentno upravljanje mobilnim robotima na osnovu neuro-fazi-genetskog prepoznavanja objekata i praćenja ljudi u robotskoj viziji*. Докторска дисертација. Ниш: Машински факултет Универзитета у Нишу, 2015; 156 стр. основног текста; литература стр. 157–168; биографија стр. 169; на крају приложен скениран извештај комисије (10 страна без пагинације). Ментор проф. др Жарко Ћојбашић; комисија: Д. Антић — председник, Ж. Ћојбашић — ментор, В. Николић, Д. Ристић-Дурант и С. Станковски. Одбрана 15.12.2015. Национални репозиторијум дисертација у Србији (НаРДуС).

I Предмет, грађа и мерило

Дисертација узима за предмет једну целину: развој система роботске визије и највишег, на визији заснованог нивоа управљања мобилном платформом, и то тако што се детекција, препознавање и просторно лоцирање објеката и људи, па и праћење људи, решавају спрегом више техника вештачке интелигенције — неуронских мрежа, fazi логике, генетских алгоритама, ANFIS-а, метода потпорних вектора и наивног Бајеса — над четири различита сензора визије (монокуларна, стерео и термовизијска камера и 3D сензор са структурираном светлошћу) на двема експерименталним платформама, DaNI у Нишу и FRIEND у Бремену. Рад сам своје полазиште сажима захтевом да систем визије омогући издвајање и класификацију особина објеката „bez potrebe za velikim bazama podataka, koje zahteva većina postojećih robotskih sistema“ (стр. 2) — дакле свесним опредељењем за приступе без великих јавних скупова, и с изричитим сужавањем на „klasifikacije objekata i ljudi u laboratorijskom okruženju“ (стр. 3). То опредељење одређује како се читају све замерке које следе.

Пре свега осталог ваља изрећи мерило, изричито и без изузетка. Рукопис је пред комисијом био крајем августа 2015 (комисија за оцену дисертације именована је 25. августа 2015, одлуком Наставно-научног већа бр. 612-508-6/2015, а извештај комисије датиран је 18. септембра 2015). Рад се стога овде мери искључиво према литератури, стандардима, прописима и стању знања који су постојали и били доступни пре тог тренутка. Ништа што је настало касније није мерило: ни доцнија издања међународних норми, ни методолошки радови објављени после рока, ни сензори и приступи који су уследили. То правило важи и за списак литературе овог осврта — свака његова јединица објављена је пре августа 2015, изузев саме дисертације, извештаја комисије и података о раду у репозиторијуму (који се временом не мењају, уз назнаку).

Наводи из дисертације преузети су дословно, у наводницима, у писму рада (латиница), с бројем стране према штампаној пагинацији (штампана страна = страна PDF-а – 20).

Још једно ваља рећи на почетку, јер одређује како се чита методолошки слој рада. Дисертација је, по нацрту, инжењерски рад који доказ поставља демонстрацијом: развојем и приказом радних система на платформи, с измереним тачностима и дијаграмима у контролисаним условима. Рад нема формално нумерисане хипотезе; носеће тврдње постављене су у уводним одељцима 1.2–1.3 и проглашене остварeним у резимеу и закључку. Демонстрација функционалног система на експерименталној платформи јесте легитиман исход инжењерског истраживања и без класичног статистичког апарата хипотеза (F. Amigoni, M. Reggiani, V. Schiaffonati, *Autonomous Robots*, 27/4, 2009; T. Bartz-Beielstein и др. (ур.), *Experimental Methods for the Analysis of Optimization Algorithms*, Springer, 2010). Замерке које следе стога се не постављају као захтев за формалном хипотезом коју ова врста рада не мора имати, него као питање **поређења и мера**: да ли оно што рад тврди има покриће у ономе што рад показује.

II Шта рад доноси

Осврт је дужан да почне оним што дисертација јесте, јер је то и највећи део онога што она вреди.

Прво и најважније, рад доноси једну носећу замисао која је истовремено његов методолошки печат: увођење повратне спреге из теорије управљања у сегментацију слике. Алгоритам препознавања објеката „zasniva se na robusnoj segmentaciji slike sa povratnom spregom“ (стр. 41), у којој дводимензионална ентропија сегментоване слике служи као мера квалитета која затвара петљу — поступак који природно повезује машинско инжењерство и управљање с рачунарском визијом и који је дат као оквир за читав низ каснијих решења.

Друго, рад је широк по обухвату сензора и платформи, а притом поштен у разграничавању сопственог и туђег доприноса. Платформа FRIEND и њен систем визије изричито се приписују бременској групи — „razvijena je na Institutu za automatiku Univerziteta u Bremenu“ (стр. 98) — а прелазак са сегментације по боји на сегментацију мапе диспаритета изведен је из стварног ограничења, не произвољно: „segmentacija po boji ne može razmatrati kao dovoljno robustan algoritam, jer namena ove mobilne robotske platforme nije da uvek posmatra čoveka spreda“ (стр. 99). Где надограђује туђу идеју, рад то каже: „Ovde je razrađena i modifikovana ideja predstavljena u radu grupe autora iz Bremena [47], da se vrši segmentacija mape dispariteta“ (стр. 87). Подела рада Ниш/Бремен транспарентно је наведена и у захвалници.

Треће, тамо где су резултати дати, дати су методолошки чисто. За класификаторе на платформи FRIEND и за детекцију људи стерео камером приказане су пуне поделе на обучавајући, валидациони и тест скуп с матрицама конфузије и ROC кривама, и то бројчано доследно: „Obučavanje, validacija i testiranje vršeno je sa 3474 objekata“ (стр. 104), при чему се та подела 70/15/15 и укупне тачности могу независно преконтролисати из приказаних матрица. Оптимизација прага сегментације у термовизији постављена је с мерљивом функцијом циља и стварном референцом — на педесет ручно сегментираних слика за сваки од четири снимка.

Четврто, рад показује методолошку самокритичност. Аутор отворено именује уско грло целог система: „klasifikacija nije kritična, već je segmentacija ta koja u pojedinim frejmovima ne daje dobre rezultate, pa se nikakvim klasifikatorom ne mogu nadomestiti greške koje u sistem uvodi algoritam segmentacije“ (стр. 93); признаје осетљивост тачности на поделу података; и уздржава се пред литературом, упозоравајући да се туђи (медицински) резултати о надмоћи ANFIS-а „ovo poređenje klasifikatora ne može se generalizovati, jer detekcija i prepoznavanje objekata na slici zavisi od vrste slike“ (стр. 75). Тај трезвен однос појачава веродостојност целог рада и, што је важно, унапред одговара на део замерки које би се иначе могле поставити оштрије.

Пето, рад доноси и промишљене замисли у делу праћења: свођење праћења човека на предвиђање временске серије положаја нелинеарном ауторегресивном мрежом, уз коришћење независног мерног еталона — „Stvarna pozicija čoveka je merena digitalnim laserskim senzorom distance Bosch PLR 50 i služila je za evaluaciju rezultata“ (стр. 137) — што је коректна метролошка основа поређења.

Најзад, потребно је поштено рећи и оно што надилази појединачне вредности: техничке тврдње рада у највећем делу одговарају стручним и произвођачким изворима на које се рад позива. То важи за спецификације носеће опреме (стерео камера Point Grey Bumblebee 2, термовизијска камера FLIR E50, контролер sbRIO-9632 платформе DaNI, 3D сензор Asus Xtion PRO LIVE, Kinect), за носеће једначине — стерео триангулација и реконструкција из диспаритета, претварање у сиву слику по коефицијентима BT.601, хибридно учење ANFIS-а — и за приказе туђих радова (нпр. преглед метода прага са четрдесет техника у шест категорија, приступ учењу хвата новог објекта, детекција позе из скелета), који верно преносе методе и налазе изворника. Ова поузданост фактографске основе није споредна: она је оно што уопште омогућава да се о раду расправља овако детаљно.

III Циљ и метод

Циљ рада, изведен из демонстрација, углавном је остварен у оном свом делу који се тиче изводљивости. Тврдња да је „moguće je razviti upravljački sistem inteligentnog robotskog sistema sa senzorom vizije koji ima sposobnosti detekcije, prepoznavanja i

prostornog lociranja objekata i ljudi“ (стр. 154) поткрепљена је радним примерима на четири сензора и двема платформама, с приказаним тачностима у контролисаним условима. То је стваран и вредан исход.

Средишња слабост методолошког слоја тиче се, међутим, једне једине речи у истој реченици закључка. Рад тврди да је спрегом више техника вештачке интелигенције могуће обезбедити препознавање и праћење „koje je superiorno u odnosu na postojeća rešenja“ (стр. 154). Та квалификација нигде у раду није поткрепљена поређењем сопствених метода с референтним методама на истим или јавним подацима. Пријављене тачности (реда 89–98%) измерене су на сопственим лабораторијским скуповима, а једино место где се реч о надмоћи изворно јавља (стр. 75) преузето је из туђих радова на магнетно-резонантним сликама, уз ауторову сопствену ограду да се не може уопштити. Изостаје и сваки тест значајности разлике међу методама, за који постоје устаљени поступци (J. Demšar, *Journal of Machine Learning Research*, 7, 2006). Мерило за ову замерку постоји и именовано је. Методолошки оквир поређења класификатора тражи да се у поређење укључи метода најсроднија предложеној и да се оцена изведе на независном тесту (S. Salzberg, *Data Mining and Knowledge Discovery*, 1/3, 1997, где стоји и да „Whenever tuning takes place, every adjustment should really be considered a separate experiment“). Протокол вредновања детекције VOC (Everingham и др., 2010) изричито упозорава да проста тачност није примерена мера кад расподела класа није равномерна: „the prior distribution over classes is significantly nonuniform so a simple accuracy measure (percentage of correctly classified examples) is not appropriate“ (M. Everingham и др., *International Journal of Computer Vision*, 88/2, 2010), и уводи просечну прецизност (AP) уз јасан критеријум поготка. Тежину замерке ваља, ипак, поштено сузити: не постоји један обавезан протокол — „In general, there is no single ‘correct’ evaluation protocol“ (P. Dollár, C. Wojek, B. Schiele, P. Perona, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34/4, 2012) — па је инжењерски избор мере легитиман ако је транспарентан; рад је, уз то, изричито усмерен на приступ без великих база и на сценарио неповољан за фронталне детекторе. Оно што остаје јесте да тврдњу о супериорности треба свести на оно што је стварно показано: изводљивост и високу тачност на сопственим подацима, а не доказану надмоћ над постојећим решењима.

Из истог круга је и уопштавање с контролисаних услова. Закључак говори о „autonomno i pouzdano“ раду, док су сва истраживања изричито сужена на лабораторију (стр. 3), а термовизијски систем „razvijan isključivo za upotrebu u zatvorenom prostoru“ (стр. 156). Овде рад на другом месту сам зна оно што у закључку пропушта — граница важења наведена је више пута — па је реч о томе да завршна реченица тај квалификатор треба изричито да понесе, а не о неоправданом закључку.

Трећа замерка тиче се кохерентности целине. Рад је по нацрту скуп засебних огледа на четири сензорска система и двема платформама, сваки са сопственим скупом, метриком и условима. Заједнички оквир — највиши, на визији заснован ниво управљања — постоји као замисао, али недостаје јединствено квантитативно мерило које би повезало доприносе међу собом и омогућило оцену да ли целина надмашује збир делова. То не поништава вредност појединачних решења, али оставља отвореним питање интегрисаности система као таквог.

Четврта се тиче поновљивости стохастичких поступака. За генетски алгоритам наведени су величина популације, елитизам, тип укрштања и мутације и интервал, али недостају број генерација, услов заустављања, вероватноће оператора и добијене нумеричке вредности прагова, а поступак није поновљен ни праћен анализом осетљивости; исто важи за генетско подешавање PID појачања, дато као оквир с упутом на пратећи рад. Мерило је овде минимални стандард поновљивости: „Reproducibility has the potential to serve as a minimum standard for judging scientific claims when full independent replication of a study is not possible“ (R. D. Peng, *Science*, 334/6060, 2011), при чему тај стандард тражи да подаци и програмски код буду доступни другима; за вредновање и подешавање стохастичких оптимизационих алгоритама исто траже и монографије-мерила (Bartz-Beielstein и др., 2010). У корист рада иде то што Peng понављивост изричито поставља тек као *минимум* и признаје да пуна независна репликација често није изводљива, а рад именује коришћени софтвер (LabVIEW 2011, MATLAB, OpenCV, OpenNI) и бројчано описује скупове — чиме тај минимум делом испуњава.

Пето, две тврдње о методи траже прецизирање. Предиктор реализован у поглављу 7 јесте нелинеарна ауторегресивна мрежа (NAR, без егзогеног улаза, како показује топологија на слици 7.4), док наслов замисли рада, Abstract и кључне речи користе NARX — а то није исти модел. Уз то, тврдња да NAR мрежа надмашује Калманов филтар изведена је из дијаграма на којима се обе процене готово поклапају с мереном позицијом (слика 7.6), без нумеричке метрике грешке; аутор и сам наводи да би разлика била израженија на геометријски неправилној путањи — дакле да у приказаном огледу није очигледна. Предност би требало показати на нелинеарним трајекторијама уз меру грешке; до тада је реч о демонстрацији применљивости, не о доказаној супериорности над устаљеним методом. (Ваља приметити и да је у стварно верификованом огледу праћења уграђен Проширени Калманов филтар, док NAR, како аутор изричито каже на стр. 156, тек треба да се имплементира на 3D платформи.)

Најзад, два наведена доприноса у раду су замисли, не изведени огледи. Оптималан избор улазног вектора карактеристика генетским алгоритмом (одељак 5.3) изложен је као поставка вишекритеријумске оптимизације — „problem se svodi na višekriterijumsku optimizaciju, a cilj je naći vektor karakteristika sa najmanjim brojem

karakteristika koji će dati najveću tačnost“ (стр. 110) — али без покренутог алгоритма, добијеног вектора и постигнуте тачности; исто важи за интелигентну адаптивну сегментацију из одељка 4.4, дату с функцијом циља али без експеримента. Оба су у попису доприноса означена речју „Predlog“ (стр. 155), што је коректно; али их у оцени научног доприноса треба држати одвојено од експериментално потврђених резултата, а не наводити раме уз раме с њима.

IV Основе обраде слике и препознавања објеката

Уводна поглавља о дигиталној аквизицији и обради слике те о детекцији и препознавању објеката носе прегледну, коректну грађу. Опис формирања мапе диспаритета преклапањем блокова јасан је и практично употребљив: „nxn blok piksela sa leve slike, gde je piksel od interesa u sredini, upoređuje sa svim mogućim blokovima iste dimenzije desne slike“ (стр. 35), уз тачно наведена ограничења епиполарне геометрије и анализу утицаја величине блока. Реконструкција тродимензионалне тачке из диспаритета и триангулација постављене су исправно, а преглед метода сегментације прагом тачно је приписан прегледном раду Sezgin–Sankur.

Замерке у овом делу су лаке и претежно чињеничне. Спецификације сензора углавном су пренете из једне систематизујуће дисертације (Filko, 2013), а не из примарних произвођачких извора; за прегледно поглавље то је прихватљив избор, а вредности су у суштини тачне, али би упућивање на примарне спецификације било поузданије. Кључне спецификације Kinect-a (640×480, 30 Hz, 11 бита односно 2048 нивоа дубине) тачне су; видно поље (58°/45°) и радно подручје рад наводи с позивом на произвођача — „Prema specifikacijama proizvođača preporučeno radno područje je od 0.7m do 5m“ (стр. 22) — али без навођења самог документа, па би упућивање на конкретну спецификацију, као и за сродне сензоре PrimeSense/Xtion за које рад каже да имају „identično rešenje“, било поузданије.

V Интелигентни системи за сегментацију и класификацију објеката

Поглавље 4 језгро је онога што рад експериментално показује, и његови су резултати документовани тамо где су најзахтевнији: неуронска мрежа на платформи FRIEND даје тачност изнад 97% с матрицом конфузије (слика 4.18), а на мапи диспаритета с четири класе изнад 98% с уредном поделом 560/120/120 која се поклапа с укупним бројевима у матрицама (слика 4.28). Уз то, поглавље носи и најбоље методолошке увиде рада — већ поменуто именовање сегментације као уског грла (стр. 93) и намерно тестирање наивног Бајеса на сировим, непречишћеним подацима да би се показало да „nikakav drukčiji klasifikator ne bi mogao da da značajnije unapređenje“

(стр. 85). Ниже тачности наивног Бајеса (65–75%) стога нису недоследност него намеран налаз, и не смеју се читати као скривена нестабилност; уз саму табелу вредело би то и изричито нагласити, јер те бројке нису рангирање класификатора — потичу с различито припремљених улаза.

Три замерке овде стоје, све умерене. Прва: резултат почетног ANFIS класификатора на платформи DaNI дат је само описно, „tačnost oko 95%“ (стр. 63), без табеле или матрице конфузије — за разлику од каснијих, захтевнијих класификатора. Реч је о првом, двокласном огледу мале носивости, поткрепљеном и провером у реалном раду робота, али би и тај резултат ради проверљивости требало приказати квантитативно. Друга: код наивног Бајеса требало је навести и бројчани однос обучавајућег и тест скупа у десет подела. Трећа: опис структуре ANFIS класификатора понавља се готово дословно у два одељка (стр. 63 и 76), што је сувишно и могло је бити скраћено упућивањем — иако има оправдање у томе што су оба огледа намерно вођена истом структуром ради упоредивости. Најзад, одељак 4.4 (интелигентна адаптивна сегментација) јесте, како је већ речено, замисао с функцијом циља, коју рад и сам затвара као правац даљих истраживања; вредност му је у поставци проблема, реализација остаје наредни корак.

VI Детекција и праћење људи: стерео и термовизија

У делу о детекцији људи стерео камером (поглавље 5) метод је изведен из стварног ограничења платформе и уредно извештен: скуп од 3474 објекта, подела 70/15/15, матрице конфузије и ROC криве по класама, укупна тачност изнад 89%. Ту, међутим, стоје две умерене замерке о поузданости баш тог броја. Прво, рад не документује да ли су обучавајући и тест скуп раздвојени по видео снимцима и особама; при случајној подели узорака из једне видео секвенце суседни, јако корелисани кадрови могу доспети и у обуку и у тест, што пријављену тачност чини оптимистичном проценом — независност теста ту није ни оповргнута ни потврђена (о стратификованој унакрсној валидацији као поузданијем поступку оцене в. R. Kohavi, *Proceedings of the 14th IJCAI*, 1995). Друго, класе су небалансиране (човек чини око три четвртине узорака), па би тривијалан класификатор „увек човек“ већ достигао око 76%, а тачност се извештава без основне вредности за поређење. Мерило је овде изричито: „As a class distribution changes these measures will change as well, even if the fundamental classifier performance does not“ (T. Fawcett, *Pattern Recognition Letters*, 27/8, 2006), а VOC протокол (Everingham и др., 2010) из истог разлога тражи прецизност/одзив уместо голе тачности при неравномерној расподели класа; о избору мере према задатку и о мерама учинка при небалансу постоји и посебна систематика (M. Sokolova, G. Lapalme, *Information Processing & Management*, 45/4, 2009; N. Japkowicz, M. Shah, *Evaluating Learning Algorithms*, Cambridge University Press, 2011). У корист рада иде то што су приказане и матрице конфузије и ROC

криве, из којих се одзив мањинске класе (72,4%) види као знатно изнад тривијалног — дакле потпунија слика ипак постоји; недостаје јој само основна вредност и осврт на небаланс. Поређења с референтним детекторима људи на истим или јавним подацима нема, што је донекле оправдано декларисаним циљем (детекција на мапи диспаритета без великих база), али вреди приметити да преглед стања у 5.1 не помиње ни стандардне јавне скупове за детекцију пешака (INRIA — N. Dalal, B. Triggs, 2005; Caltech — P. Dollár и др., 2009), што би контекст употпунило; захтев да обухвати и дубоке детекторе био би, међутим, анахрон за рукопис из средине 2015 и овде се не поставља.

Термовизијски део (поглавље 6) доноси заокружен ток обраде — fazi сегментација која опонаша људско расуђивање о температури, генетска оптимизација параметара сегментације с референтним ручно сегментираним сликама, и SVM класификација — на сопственим снимцима, с прецизно наведеним спецификацијама камере („Termovizijska osetljivost kamere je manja od 0.05°C a tačnost čitanja je ±2°C“, стр. 118).

Најозбиљнија замерка целог рада налази се, међутим, управо овде. Одељак 6.5, насловљен „Evaluacija rezultata“, не садржи ниједан бројчани показатељ за термовизијски SVM класификатор — ни тачност, ни матрицу конфузије, ни прецизност, одзив или F-меру, нити поделу скупа, избор језгра и хиперпараметара; евалуација се своди на општу расправу и једну илустрацију успешног праћења. За класификатор чији су улази и обим података наведени (512 објеката, четири обележја, из двеста кадрова), то је стварна празнина у извештавању, а не ситница, јер управо о том делу почива закључак поглавља о поузданом препознавању људи термовизијом. Минимум који се разумно очекује именован је и невелик: „Scaling before applying SVM is very important“ и препорука да се спроведе „a ‘grid-search’ on C and γ using cross-validation“ (C.-W. Hsu, C.-C. Chang, C.-J. Lin, *A Practical Guide to Support Vector Classification*, издање од 15.04.2010) — дакле бар скалирање, језгро и пар (C, γ) уз поделу скупа. Пошто тога у поглављу нема, закључак о „поузданом“ препознавању термовизијом у овом делу није квантитативно поткрепљен и мора се ослонити на бројчане резултате сродних класификатора из поглавља 4–5 и из пратећих радова, или свести на доказ применљивости термовизије у мобилној роботици — при чему је за такво поређење постојао наменски термовизијски јавни скуп (OTCBVS Benchmark Dataset Collection, скуп 01: OSU Thermal Pedestrian Database, Ohio State University, 2005).

Из истог круга су и две мање замерке које не дирају тезу. Fazi сегментација ради над каналима боја једне фиксне палете лажне боје, а не над сировим температурним вредностима, па fazi правила и прагови зависе од избора палете и температурног опсега — при њиховој промени морали би се изнова подешавати; преносивост стога није демонстрирана, иако је у оквиру огледа палета намерно фиксирана, а генетска

оптимизација прага делом компензује варијације. Ознака верзије комплета уз плочу sbRIO-9632 наведена је као „1.0“, док према документацији произвођача (National Instruments, LabVIEW Robotics Module — одреднице о иницијализацији комплета Starter Kit 1.0 за плочу sbRIO-9631 и Starter Kit 2.0 за плочу sbRIO-9632) тој плочи (400 MHz) одговара комплет 2.0 — тривијална омашка у ознаци верзије, без утицаја на резултате; сама плоча описана је тачно.

VII Праћење, позе и гестови с 3D сензором

Поглавље 7 добро је постављено као инжењерска целина: праћење човека сведено је на предвиђање временске серије положаја рекурентном мрежом, детекција 3D сензором повезана је с естимацијом и предвиђањем и с двама PID контролерима за одржавање растојања и поравнање, а стварни положај мерен је независним ласерским еталоном. Аутор поштено наводи границе спроведеног — да NAR систем још није интегрисан на 3D платформи (стр. 156) — чиме унапред одговара на део напетости између NAR-а и Калмановог филтра.

Замерке су, осим већ наведених терминолошке (NAR/NARX) и оне о недоказаној предности над Калмановим филтром, претежно о обиму верификације. Интегрисани систем праћења проверен је у једном лабораторијском сценарију, на једном растојању (1,5 m), без броја понављања, варирања особа, растојања и брзина и без метрике грешке праћења; из тога се изводи општа тврдња о аутономном и поузданом праћењу, коју би требало читати као доказ концепта. Естимација и предикција NAR мрежом демонстриране су без наведеног обима скупа и поделе у самој тези (шира верификација дата је у пратећим радовима). Прегледни приказ Кинектовог алгоритма позе из скелета садржински је тачан, али је ослоњен на спецификациону референцу; за такав приказ туђег комерцијалног решења требало би навести и изворни научни рад (Shotton и др., 2011) уз једначину карактеристике дубине. Најзад, у делу о гестовима навод библиотеке Coding4Fun.Kinect.Wpf — чији назив упућује на Kinect for Windows SDK — стоји уз стварно коришћени ток OpenNI/Xtion, што су различита окружења и што би ради прецизности имплементационог описа требало раздвојити.

VIII Слике, табеле, једначине, бројеви и језик

Овај слој рада најмање је брижљив, мада ниједна од омашки у њему не дира садржину — све су, по правилу, локално инвертибилне или се самим текстом разрешавају.

Једначине. У дефиницији глобалног прага прозни опис ставља објекат испод прага, док једначина 3.7 бинарну јединицу додељује пикселима изнад прага, супротно конвенцији „Beli pikseli odgovaraju segmentiranom objektu“ (стр. 37) — реч је

о замени смера доделе коју читалац лако инвертује, а стварно коришћена сегментација по нијанси тиме није погођена. У једначини 3.3 (сума апсолутних разлика) границе суме нису усклађене с индексирањем сабирка, а за непарно n није употребљено (n−1)/2; намера је, међутим, из текста јасна, а минимизација у једначини 3.4 исправна. Под једначином 3.32 наведено је осам инваријанти названих „Hu-овим“, док их је M. K. Hu (1962) дефинисао седам (осми потиче од Flusser–Suk, 1993), уз погрешне знаке у појединим сабирцима — али у класификацији су и иначе коришћене прве три, односно седам инваријанти, па осми не улази у резултате; исти инваријант тачно је записан на другом месту (једн. 3.32 наспрам једн. 5.1, где недостаје коефицијент 3 уз η12). У приказу ANFIS структуре горњи индекси излаза слојева 4 и 5 не прате број слоја, а унакрсни позив на једначине (3.32) и (3.33) на стр. 80 треба да гласи (4.12) и (4.13); сам извод делта-правила је коректан. Запис једначине 5.2 нотационо је непрецизан (даје скалар уместо селекције обележја; намеравано је $\mathrm{diag}(\theta)\cdot x^{T}$), а вектор θ привидно има десет чланова услед поновљеног индекса — из димензије јединичне матрице и броја обележја једнозначно следи да их је девет. У једначини 6.6 величина „4320“ словна је омашка за 43.200, што сам текст на стр. 118 демантује. Ниједна од ових омашки не мења ниједан резултат, али све заједно траже пажљиву математичку редактуру, у духу доследног писања ознака и симбола (ISO 80000-2:2009).

Слике и табеле. Две различите слике носе исти број (Слика 4.9 на стр. 68 и 69; Слика 4.27 на стр. 90 и 94), недостаје ознака слике 5.1, а нумерација пододељака поглавља 7 је неуредна (изостаје 7.3.2, а одељак о управљању означен је „7.4.2“ и стоји ван редоследа). Табеле су нумерисане глобално, а слике по поглављима — интерно исправна, али недоследна шема. У опису слике 2.15 панел означен као сегментација по боји заправо приказује сиву верзију претходног панела — омашка у опису преузете илустрације прегледног поглавља.

Бројеви. Величине скупова наведене за различите класификаторе (око 860 за ANFIS, 940 триплета, „preko 800“ у уводу) нису у тексту изричито усаглашене, што отежава праћење на којим тачно подацима почива који резултат; разлика проистиче из различитог одбацивања лоше сегментираних узорака, а свака појединачна подела интерно је доследна. Помен „90 iteracija“ обуке не поклапа се са сликом 5.8, која документује 43 епохе (најбоља валидација на 37) уз рано заустављање; тачност изнад 89% од тог броја не зависи. На дијаграму слике 7.6 вредности на оси дубине (до око 23 m) несагласне су с лабораторијским опсезима и дометом стерео камере, што указује на проблем размере осе, а апсцисе дијаграма праћења нису доследно означене.

Језик. Кроз рад постоји низ куцачких омашки (нпр. „disetraciji“, „struktuiranom“/„strktiranom“, „block maching“, „addaboost“, „porepoznavanje“, „ra čuna“), недоследан запис назива (једном „FREND“ уместо „FRIEND“; „Bumblebee“/„Bumbebee“), а на

насловној страни у енглеском наслову стоји „Gentic“ уместо „Genetic“; закључна реченица о термовизији на стр. 128 недовршена је (изостала је реч „superiorno“). Реч је о лекторским пропустима који не мењају смисао, уз стручне англицизме (диспаритет, кадар/фрејм, кластеровање) који су делом устаљени у домаћој техничкој литератури, па њихова употреба сама по себи није мана ако је доследна (в. *Правопис српскога језика*, Матица српска, 2010, који у латиничком тексту допушта изворно писање страних имена, али тражи доследност). Једна језичка омашка ипак мења значење и вреди је издвојити: рад наводи да глувонеме особе користе „glasovnog jezika“ (стр. 130), док је реч о знаковном језику, чему одговарају и наведени извори — то треба исправити.

IX Научни апарат

Научни апарат рада у основи је уредан, што ваља рећи пре замерки. Свако позивање у тексту разрешава се на постојећу јединицу, нема цитата ван опсега [1]–[120], а од сто двадесет јединица искоришћено је сто деветнаест; нумерисање по редоследу првог навођења доследно је спроведено — управо онако како допушта међународна норма за библиографске референце: „numerals in the text… refer to information resources in the order in which they are first cited“ (ISO 690:2010, т. A.3.1). Нема дуплих јединица ни посредног цитирања; литература је савремена (око три петине из последњих пет година пре предаје) и покрива темељна дела уже области. Табела 1 прегледно повезује раније објављене радове аутора с деловима рада по сензору, алгоритму и класификатору.

Замерке су формалне, али бројне. Једна јединица (Tang и др., 2000) наведена је у попису, а није цитирана у тексту, где тематски припада набрајању о детекцији глава стерео визијом. Два интернет извора наведена су без датума приступа, а један од њих (Wikipedia) и без наслова, аутора и године — што одступа од норме, која за онлајн изворе подложне промени тражи „the network location of the particular version that was used and the date on which the document was accessed“ (ISO 690:2010, т. 3.4). Кључне хардверске и софтверске компоненте (Asus Xtion, NI Robotics Starter Kit, LabVIEW, OpenCV, OpenNI, MATLAB, FLIR E50) именоване су у тексту, често с верзијом, али без засебних библиографских јединица; седам јединица наведено је без године, а сопствена техничка решења и без издавача и страна. Транслитерација имена није уједначена (Gräser/Graeser/Graser; Ristić-Durrant/Durant; Nikolić/Nikolic), као ни стил записа јединица (мешају се тачка-зарез, велика слова презимена и APA-облик с иницијалима) — а норма изричито тражи супротно: „A uniform style, format and punctuation scheme shall be used for all references in a document, regardless of the particular style guide being used“ (ISO 690:2010, т. 3.5). У једној јединици распон страна записан је с погрешним знаком.

Садржински осетљивија је навика да се темељне методе на месту помена упуте на радове аутора или на несродне јединице уместо на изворна дела. Три критеријума оптималне детекције ивица порекло имају у раду J. Canny (1986), а не у цитираној јединици о сегментацији геометријских примитива из даљинских података; стандардне методе на набрајању класификатора (SVM, AdaBoost) упућене су на радове аутора уместо на изворна дела (Cortes–Vapnik, 1995; Freund–Schapire, 1997), при чему је каскада исправно приписана Виоли и Џонсу; а прегледни приказ Кинектовог алгоритма позе требало би да носи изворни научни рад (Shotton и др., 2011) поред спецификационе референце. У истом реду вреди приметити и да семинални радови носећих метода рада — изворни рад о ANFIS-у (Jang, 1993), о fazi скуповима (Zadeh, 1965) и о генетским алгоритмима (Goldberg, 1989) — нису у попису, иако су цитиране њихове примене; то је напомена за размишљање, а не пропуст који дира садржину, будући да су методе примењене исправно и потврђене на местима формалне разраде. Сами називи метода свуда су тачни, па је реч о усмерењу цитата, не о садржинској грешци.

Најзад, ваља изрећи и оно што иде у прилог раду: висока самоцитираност није мана него испуњење прописаног услова — допуна универзитетског правилника из 2013 (Гласник Универзитета у Нишу, бр. 3/13) захтева објављен или прихваћен рад пре одбране, а рад ту поновну употребу изричито исказује (Табела 1, одељак 1.4); категорије радова аутора уредно су сврстане по Правилнику о вредновању научноистраживачких резултата („Службени гласник РС“, бр. 38/2008).

X Однос рада и реферата комисије

Извештај комисије тачно наводи основну грађу рада — осам поглавља, осам табела, сто двадесет референци — и оцењује да су циљеви постављени у пријави остварени. Тај део оцене, у делу који се тиче изводљивости и обима, одговара стању у тексту: рад доноси низ радних решења по сензорима и платформама, с документованим тачностима тамо где су резултати дати.

Две ствари, међутим, траже опрез. Прво, оцена да су циљеви остварени преузима из резимеа и квалификацију о супериорности над постојећим решењима, а та квалификација, како је показано, нема покриће у поређењу с референтним методама; извештај је не проверава посебно. Друго, извештај (и сажетак делова рада) узима да је спроведена упоредна анализа различитих интелигентних алгоритама, док рад сам две стране пред крај признаје да би тек будуће истраживање подразумевало „primenu svih spomenutih klasifikatora na sve probleme, njihovu uporednu analizu performansi“ (стр. 156) и да је „realno upoređenje performansi klasifikatora moguće tek nakon većeg broja slučajnog odabira trening, validacionog i test skupa“ (стр. 156) — унутрашњу напетост коју извештај не примећује. У истом светлу, два доприноса која извештај наводи као остварена — оптималан избор улазног

вектора и адаптивна сегментација — у раду су замисли, што и сам рад ознаком „Predlog“ признаје. Најзад, званични метаподаци нису усклађени међу собом (извештај наводи 170 страна и 98 слика, страница с подацима 169 страна и 96 слика), што је формална ствар коју би требало ускладити. Оцену комисије о самосталности и креативности аутора осврт не проверава — реч је о вредносном суду који се из самог рада не може ни потврдити ни оповргнути, а подела рада Ниш/Бремен транспарентно је наведена.

XI Шта остаје

Кад се све сабере, добија се оцена коју вреди изрећи у једном комаду, без ублажавања и без претеривања.

Дисертација Ивана Ћирића солидан је и самостално изведен инжењерски рад који домаћој роботици и аутоматском управљању доноси широку целину: систем роботске визије и на њему заснован ниво управљања мобилном платформом, изведен спрегом више техника вештачке интелигенције над четири сензора визије и на двема експерименталним платформама. Њене су трајне вредности методолошки печат — повратна спрега у сегментацији слике — поштено разграничавање сопственог и туђег доприноса, методолошки чисто извештавање тамо где су резултати дати (матрице конфузије, ROC криве, уредне поделе скупова, независни мерни еталон), и методолошка самокритичност која уско грло система именује тачно. Фактографска основа рада у највећем делу држи.

Слабости су, по оцени овог осврта, четвороструке, и ниједна не руши изводљивост као носећи исход. Прва и најтежа јесте раскорак између онога што рад тврди и онога што показује: квалификација да је праћење „superiorno u odnosu na postojeća rešenja“ (стр. 154) нема покриће ни у једном поређењу с референтним методама на истим или јавним подацима, а закључак о „autonomno i pouzdano“ раду уопштава с изричито лабораторијских услова. Друга је празнина у извештавању баш на осетљивим местима — термовизијски SVM класификатор без иједног бројчаног показатеља у одељку насловљеном „Evaluacija rezultata“, почетни ANFIS без табеле, детекција људи без документоване независности теста и без основне вредности при небалансу класа, генетски поступци без потпуних параметара и понављања. Трећа је кохерентност целине: недостаје јединствено мерило које би повезало засебне огледе и показало да интегрисани систем надмашује збир делова. Четврта је уредност — нумерација пододељака и слика, недоследан библиографски стил и транслитерација, усмерење цитата темељних метода на сопствене или несродне радове, и низ лекторских омашки, међу којима једна (знаковни наспрам гласовног језика) мења значење.

Изнад свега тога стоји једна правилност која иде у прилог раду: тамо где рад мери и разграничава, чини то брижљивије него тамо где закључује и уопштава. Управо зато највећи део онога што треба поправити већ је негде у раду тачно урађено или тачно речено — само на погрешном месту, недоследно спроведено, или изражено јаче него што бројеви носе.

Пет измена биле би, по оцени овог осврта, најплодније у евентуалном монографском издању или наставку истраживања. Прво, свести тврдњу о супериорности на оно што је показано и, где год је изводљиво, спровести поређење с најсроднијом референтном методом на истим или јавним подацима (за термовизију постоји наменски јавни скуп — OTCBVS, скуп 01: OSU Thermal Pedestrian Database, 2005), уз меру примерену небалансу класа. Друго, употпунити евалуацију термовизијског класификатора бројчаним показатељима (тачност, матрица конфузије, прецизност/одзив, подела скупа, језгро и хиперпараметри), а почетни ANFIS резултат приказати квантитативно. Треће, документовати независност теста од обуке код детекције људи (подела по снимцима/особама, а не по узорцима) и допунити параметре и понављања стохастичких поступака. Четврто, увести јединствено мерило које би повезало појединачне огледе и оправдало реч „интегрисани“, а у попису доприноса доследно одвојити замисли (оптималан избор улаза, адаптивна сегментација) од потврђених резултата; ускладити и ознаку модела (NAR наспрам NARX) кроз наслов, кључне речи и тело рада. Пето, спровести уредничку и лекторску редактуру — нумерација, једначине, транслитерација, уједначавање библиографског стила и датуми приступа код интернет извора, уз изворне референце темељних метода.

Уз ту врсту редакције — која у највећем делу не тражи ново истраживање него доследно спровођење и мерење онога што је рад већ започео — ова дисертација остаје вредан и употребљив допринос, чија би методолошка замисао и појединачна решења имали шта да кажу и стручном читаоцу изван уже теме.

Литература

Предмет осврта

Ћирић, Иван Т., *Inteligentno upravljanje mobilnim robotima na osnovu neuro-fazi-genetskog prepoznavanja objekata i praćenja ljudi u robotskoj viziji* [докторска дисертација]. Машински факултет Универзитета у Нишу, 2015. (Национални репозиторијум дисертација — НаРДуС, https://nardus.mpn.gov.rs/handle/123456789/5301.)

Извештај комисије о оцени докторске дисертације Ивана Т. Ћирића, Машински факултет Универзитета у Нишу, 18. септембар 2015.

Методологија поређења, вредновања и експеримента

Salzberg, Steven L., „On Comparing Classifiers: Pitfalls to Avoid and a Recommended Approach“, *Data Mining and Knowledge Discovery*, 1/3, 1997, 317–327. DOI 10.1023/A:1009752403260.

Kohavi, Ron, „A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection“, *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI)*, Montreal, 1995, 1137–1145.

Demšar, Janez, „Statistical Comparisons of Classifiers over Multiple Data Sets“, *Journal of Machine Learning Research*, 7, 2006, 1–30.

Amigoni, Francesco; Reggiani, Monica; Schiaffonati, Viola, „An insightful comparison between experiments in mobile robotics and in science“, *Autonomous Robots*, 27/4, 2009, 313–325. DOI 10.1007/s10514-009-9137-8.

Bartz-Beielstein, Thomas; Chiarandini, Marco; Paquete, Luís; Preuss, Mike (yp.), *Experimental Methods for the Analysis of Optimization Algorithms*, Springer, Berlin/Heidelberg, 2010. DOI 10.1007/978-3-642-02538-9.

Japkowicz, Nathalie; Shah, Mohak, *Evaluating Learning Algorithms: A Classification Perspective*, Cambridge University Press, Cambridge, 2011.

Мере учинка, ROC и протоколи детекције

Fawcett, Tom, „An introduction to ROC analysis“, *Pattern Recognition Letters*, 27/8, 2006, 861–874. DOI 10.1016/j.patrec.2005.10.010.

Sokolova, Marina; Lapalme, Guy, „A systematic analysis of performance measures for classification tasks“, *Information Processing & Management*, 45/4, 2009, 427–437. DOI 10.1016/j.ipm.2009.03.002.

Everingham, Mark; Van Gool, Luc; Williams, Christopher K. I.; Winn, John; Zisserman, Andrew, „The PASCAL Visual Object Classes (VOC) Challenge“, *International Journal of Computer Vision*, 88/2, 2010, 303–338. DOI 10.1007/s11263-009-0275-4.

Dollár, Piotr; Wojek, Christian; Schiele, Bernt; Perona, Pietro, „Pedestrian Detection: An Evaluation of the State of the Art“, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34/4, 2012, 743–761. DOI 10.1109/TPAMI.2011.155.

Dollár, Piotr; Wojek, Christian; Schiele, Bernt; Perona, Pietro, „Pedestrian Detection: A Benchmark“, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009, 304–311. DOI 10.1109/CVPR.2009.5206631.

Dalal, Navneet; Triggs, Bill, „Histograms of Oriented Gradients for Human Detection“, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, 886–893. DOI 10.1109/CVPR.2005.177.

OTCBVS Benchmark Dataset Collection, скуп 01: *OSU Thermal Pedestrian Database*, Ohio State University (Computer Vision Laboratory), 2005, http://vcipl-okstate.org/pbvs/bench/.

Hsu, Chih-Wei; Chang, Chih-Chung; Lin, Chih-Jen, *A Practical Guide to Support Vector Classification*, Department of Computer Science, National Taiwan University, издање од 15.04.2010.

Поновљивост

Peng, Roger D., „Reproducible Research in Computational Science“, *Science*, 334/6060, 2011, 1226–1227. DOI 10.1126/science.1213847.

Научни апарат, ознаке и језик

ISO 690:2010, *Information and documentation — Guidelines for bibliographic references and citations to information resources*, Third edition, ISO, Geneva, 2010-06-15.

ISO 80000-2:2009, *Quantities and units — Part 2: Mathematical signs and symbols to be used in the natural sciences and technology*, ISO, Geneva, 2009-12-01.

Правопис српскога језика, измењено и допуњено издање (приредили М. Пешикан, Ј. Јерковић, М. Пижурица), Матица српска, Нови Сад, 2010.

Формални оквир доктората (важећи до августа 2015)

Правилник о поступку и начину вредновања, и квантитативном исказивању научноистраживачких резултата истраживача, „Службени гласник РС“, бр. 38/2008 (14.04.2008).

Правилник о допунама Правилника о поступку давања сагласности на одлуку о усвајању теме докторске дисертације и на одлуку о усвајању извештаја о урађеној докторској дисертацији, „Гласник Универзитета у Нишу“, бр. 3/13, 26.04.2013.

Изворне методе (семинални радови)

Canny, John, „A Computational Approach to Edge Detection“, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8/6, 1986, 679–698. DOI 10.1109/TPAMI.1986.4767851.

Cortes, Corinna; Vapnik, Vladimir, „Support-Vector Networks“, *Machine Learning*, 20/3, 1995, 273–297. DOI 10.1007/BF00994018.

Freund, Yoav; Schapire, Robert E., „A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting“, *Journal of Computer and System Sciences*, 55/1, 1997, 119–139. DOI 10.1006/jcss.1997.1504.

Jang, Jyh-Shing Roger, „ANFIS: Adaptive-Network-Based Fuzzy Inference System“, *IEEE Transactions on Systems, Man, and Cybernetics*, 23/3, 1993, 665–685. DOI 10.1109/21.256541.

Zadeh, Lotfi A., „Fuzzy Sets“, *Information and Control*, 8/3, 1965, 338–353. DOI 10.1016/S0019-9958(65)90241-X.

Goldberg, David E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, Reading (MA), 1989.

Hu, Ming-Kuei, „Visual Pattern Recognition by Moment Invariants“, *IRE Transactions on Information Theory*, 8/2, 1962, 179–187. DOI 10.1109/TIT.1962.1057692.

Flusser, Jan; Suk, Tomáš, „Pattern Recognition by Affine Moment Invariants“, *Pattern Recognition*, 26/1, 1993, 167–174. DOI 10.1016/0031-3203(93)90098-H.

Shotton, Jamie; Fitzgibbon, Andrew; Cook, Mat; Sharp, Toby; Finocchio, Mark; Moore, Richard; Kipman, Alex; Blake, Andrew, „Real-Time Human Pose Recognition in Parts from Single Depth Images“, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, 1297–1304. DOI 10.1109/CVPR.2011.5995316.

Техничка документација опреме

National Instruments, *LabVIEW Robotics Module Help* (одреднице о иницијализацији комплета Starter Kit 1.0 за плочу sbRIO-9631 и Starter Kit 2.0 за плочу sbRIO-9632) и *NI LabVIEW Robotics Starter Kit — Data Sheet (sbRIO-9632)*, издања важећа до 2015.