

UNIVERSITATEA BABEȘ-BOLYAI



REZUMATUL TEZEI DE DOCTORAT

**Tehnici și aplicații avansate de procesare a imaginilor și
viziune computerizată**

Doctorand:

Alexandru I. MARINESCU

Conducător științific:

Prof. Univ. Dr. Anca M. ANDREICA

Facultatea de Matematică și Informatică
Departamentul de Informatică

2023

Rezumat

Teza de doctorat de mai jos își propune să ofere o perspectivă asupra cercetărilor anterioare și actuale ale autorului în domeniul în continuă expansiune al vederii computerizate și al procesării imaginilor. Ne vom concentra mai întâi asupra subiectului sensibil al segmentării imaginilor, folosind automate celulare special concepute, o abordare care va fi utilizată ulterior într-un modul de reconstrucție facială 3D. Deși automatele celulare au fost înlocuite în prezent de abordările de învățare profundă (deep-learning), demonstrăm numeric că simplitatea lor aparentă este compensată de viteza lor de calcul și exploatăm ușurința de paralelizare pe unitatea de procesare grafică. Automatul celular hibrid Grow Cut propus de noi este un instrument generic, nefiind limitat la segmentarea imaginilor medicale. Cu o ușoară adaptare, suntem capabili să îl aplicăm cu succes la segmentarea părului unui subiect uman și să extragem în continuare conturul părului și al frunții, ceea ce reprezintă o completare importantă la scopul principal al acestei teze: un modul de reconstrucție facială 3D.

Dorim să obținem o reconstrucție rapidă și precisă a feței unui subiect uman, păstrând cât mai mult posibil trăsăturile faciale, ceea ce reprezintă încă o problemă deschisă în domeniul vederii computerizate și, în special, în analiza criminalistică și în sistemele de recomandări de modă. Încercarea noastră de a crea un modul de reconstrucție geometrică a feței provine din observația că abordările actuale de învățare profundă pot capta doar caracteristicile statistice medii ale feței umane, fără a lua în considerare caracteristicile excepționale. Există o corelație directă între acuratețea reconstrucției și rezultatul măsurătorilor faciale. În plus, trecerea de la imagini bidimensionale la un model tridimensional ne permite să aliniem modelul la o postură perfect neutră, o sarcină imposibilă în cazul imaginilor bidimensionale pure. Modulul de reconstrucție este validat pe două seturi de date foarte cunoscute, oferind rezultate promițătoare. Se va face o comparație cu tehnicile de ultimă generație care pretind că pot realiza reconstrucția dintr-o singură imagine, prin intermediul rețelelor neuronale convoluționale. După cum vom demonstra numeric, aceste tipuri de modele sunt limitate la reconstrucția fețelor medii, în timp ce modulul nostru utilizează preponderent transformări geometrice.

În timpul fazei experimentale a cercetării noastre, ne-am confruntat cu necesitatea de a avea o cameră video bine calibrată înainte de a achiziționa imaginile care urmează să fie trimise la modulul de reconstrucție. O astfel de calibrare trebuie să ne permită, cel puțin, să anulăm distorsiunile imaginii, inerente oricărui obiectiv de cameră. Această etapă trebuie efectuată o singură dată pentru fiecare model de cameră, prin realizarea de instantanee ale unui model de calibrare în diferite poziții în fața camerei. Spre deosebire de abordările obișnuite, în care o aplicație de calibrare îi spune explicit utilizatorului unde să plaseze modelul pentru a minimiza eroarea de reproiecție a calibrării, noi pornim de la un set mai mare de imagini și utilizăm un algoritm genetic care găsește setul optim de imagini de calibrare, având în vedere anumite criterii globale. Algoritmul genetic este supus la două adaptări pentru a-l face mai robust în ceea ce privește căderea în atractori de minim local, iar algoritmul este evaluat atât pentru calibrarea mono cât și stereo, în funcție de patru parametri de performanță.

În cele din urmă, este analizat un prototip de sistem de inferență a formei feței, care va constitui subiectul unor cercetări în curs. În faza sa actuală, acest sistem de inferență ia ca date

de intrare distanțe personalizate ale reperelor faciale și, prin utilizarea unui clasificator naiv Bayes, poate deduce forma feței unui subiect uman cu o precizie raportată de 85%.

Subiect: *Viziune computerizată*

Cuvinte cheie: *automate celulare, segmentarea imaginilor, reconstrucția facială, structura din mișcare, calibrarea camerei, inferența formei feței*

Motivația cercetării

Îl familiarizăm pe cititor cu conceptele de segmentare a imaginilor prin intermediul automatelor celulare, reconstrucția tridimensională a feței umane prin intermediul unui sistem simplu și preponderent geometric și facem o punte între cele două subiecte aparent disparate, prin dezvoltarea unei tehnici îmbunătățite de segmentare a liniei părului și a frunții, care poate fi utilizată pentru a spori reconstrucția facială.

Motivul pentru care am adus în discuție un subiect considerat bine înțeles și studiat, cum este segmentarea imaginilor prin intermediul automatelor celulare, este că am considerat că acest subiect nu a fost încă exploatat pe deplin. Am remarcat, de asemenea, progresele recente ale hardware-ului de calcul GPU de uz general, care se potrivesc perfect acestui scenariu, deoarece automatele celulare au o caracteristică inerentă care le face candidați perfecți pentru paralelizare. Prelucrarea imaginilor și viziunea computerizată, indiferent dacă sunt realizate prin intermediul automatelor celulare, al transformărilor geometrice sau al abordărilor de învățare profundă, au aplicații potențial nelimitate. Din punct de vedere filozofic, trebuie să ne întoarcem la însăși rațiunea care a stat la baza apariției calculatoarelor, mai întâi ca simplu instrument de calcul, vizând, în cele din urmă, crearea unei mașini pe deplin conștiente, ca urmaș al speciei umane. Vederea este unul dintre cele cinci simțuri umane și o componentă cheie în multe aplicații bazate pe calcul. Vorbind în terminologie algoritmică, acesta primește ca date de intrare o imagine sau o secvență/un flux de imagini și emite o decizie, care este apoi transmisă serviciilor executoare ulterioare, care efectuează o acțiune.

După cum am identificat mai târziu, automatul celular modificat pe care l-am dezvoltat nu se limitează la analiza imaginilor medicale, ci poate fi reprofilat pentru a se potrivi cu sistemul nostru de reconstrucție facială 3D în sarcini precum segmentarea feței/părului, unde pixelii imaginii nu sunt legați prin diferența de intensitate a semnalului (de exemplu, imaginile de rezonanță magnetică), ci prin culoarea părului/pigmentarea pielii. Un mare interes a fost stârnit în domeniul reconstrucției faciale, din păcate, în timpul recente pandemii care a afectat societatea. Din nou, am pornit de la o sarcină specifică (de exemplu, efectuarea de măsurători faciale ale subiecților umani, minimizând în același timp interacțiunile la distanță mică între oameni) și am generalizat la domenii de interes mai largi, cum ar fi analiza criminalistică și sistemele de recomandare bazate pe forma feței. Există, de asemenea, o nișă pentru crearea de avatare personale în curs de dezvoltare în cadrul trendului meta-versului.

Conținutul cercetării

Vom aborda noțiuni introductive cu privire la cele două subiecte principale de cercetare, trecând în revistă definiția automatului celular ca și construcție matematică, principalele sale caracteristici, cum ar fi simplitatea și ușurința de paralelizare, făcând câteva observații cheie cu privire la stabilitatea și convergența sa pe măsură ce se desfășoară iterațiile. Investigația noastră inițială în acest domeniu a fost realizată în cadrul unui grant de cercetare în care obiectivul principal a fost segmentarea precisă și rapidă a scanărilor medicale obținute prin rezonanță magnetică a inimii umane. Cerințele au fost foarte stricte și au solicitat două caracteristici foarte specifice ale algoritmului de segmentare, și anume nesupervizată și autonomă, trăsături care vor fi detaliate în conținutul acestei teze, ceea ce înseamnă, în esență, că ar trebui să existe o intervenție redusă sau chiar inexistentă a utilizatorului final în timpul configurării și execuției algoritmului, dar care se dovedește a fi utilă pentru noi în sensul că aceeași tehnică poate fi utilizată într-un scenariu cu totul aparte - segmentarea părului unui subiect uman în imagini 2D. Datorită naturii agnostice față de datele de intrare a automatului celular de segmentare, această tranziție s-a făcut fără probleme.

Discutăm în detaliu conceptul de segmentare a imaginilor prin intermediul automatelor celulare, menționând câteva lucrări de referință în acest domeniu, peste care aducem câteva optimizări special concepute, care se dovedesc a fi benefice prin experimentare numerică. Algoritmul de bază care servește drept fundament pentru tehnica propusă de noi este cunoscut în literatura de specialitate sub numele de Grow Cut. Rezultatele numerice pentru tehnica noastră hibridă, care va fi denumită în continuare "Grow Cut nesupervizat hibrid", sunt foarte promițătoare, după cum reiese din creșterea valorii coeficientului Dice, o măsură bine cunoscută a preciziei de segmentare a imaginilor.

A doua jumătate a tezei se concentrează pe sarcina complexă de reconstrucție facială umană din achiziții faciale 2D sub formă de imagini simple, cu o aplicație concretă pentru inferența formei feței umane. La momentul scrierii acestei lucrări, literatura de specialitate privind inferența formei feței este redusă, deoarece determinarea formei feței umane este o chestiune foarte subiectivă. Am dezvoltat sistemul nostru de reconstrucție facială pornind de la ipoteza că, în procesul de realizare a unei fotografii, informațiile de adâncime sunt eliminate, rămânând doar măsurătorile 2D ale reperelor faciale care sunt insuficiente.

Astfel, am început să construim un modul de reconstrucție facială 3D, concentrându-ne să îl facem robust, bazându-ne pe transformări geometrice cât mai mult posibil. Așa cum era de așteptat, deoarece informațiile de adâncime se pierd în timpul achiziționării unei imagini, acestea trebuie recuperate din altă parte. Abordările recente pretind că pot realiza acest lucru prin intermediul tehnicilor de învățare profundă, dar, așa cum am concluzionat în urma replicării lor, aceste modele învață o medie statistică a feței umane și, prin urmare, sunt foarte predispuși la anomalii (adică indivizi cu particularități faciale). Acest lucru ne-a determinat să construim sistemul nostru de reconstrucție. Deoarece nu vom obține reconstrucția noastră prin intermediul rețelelor neuronale convoluționale sau al altor metode similare, informațiile de adâncime lipsă trebuie recuperate din mai multe ipostaze - mai multe achiziții 2D ale aceluiași subiect sub diferite orientări față de cameră.

Rezultatul va consta într-un set de repere faciale 3D, cunoscut sub numele de nor de puncte de reper facial. În continuare, ne confruntăm cu sarcina de a crește densitatea de vertecși a modelului facial. Am reușit să traversăm această etapă prin integrarea modelelor deformabile în sistemul nostru. După cum sugerează și numele lor, acestea sunt modele de fețe umane oarecum generice (adică, care împărtășesc trăsături atât ale bărbaților, cât și ale femeilor, determinate ca o medie statistică umană), care au o densitate mare de vertecși și, pe baza unei cartografieri a reperelor, sunt deformate cu ajutorul interpolării funcției de bază radială pentru a le aduce la forma norului de puncte de reper 3D. Texturarea se va realiza folosind petece de imagine “cusute” din achizițiile inițiale.

O secțiune întreagă este dedicată subiectului calibrării camerei, fie că este vorba de mono sau stereo. O cameră bine calibrată presupune că rectificarea imaginii de intrare va furniza cu precizie sistemului de înregistrare a norului de puncte 3D reperele dorite. În loc să optimizăm parametrii intrinseci și extrinseci ai camerei, așa cum procedează majoritatea abordărilor, ne bazăm pe determinarea submulțimii optime de imagini de calibrare care produce calibrarea cu cea mai mică eroare medie de reproiecție. Realizăm acest lucru prin utilizarea unui algoritm genetic, în care cromozomii sunt o submulțime dintr-o mulțime mai mare de imagini de calibrare, iar indicatorul de optimalitate este eroarea medie de reproiectare, în pixeli. Astfel, abordarea noastră se situează la granița dintre explorarea prin forță brută și căutarea aleatorie a întregului spațiu de soluții. Am validat această tehnică, atât pentru calibrarea camerelor mono cât și stereo, și am ajuns la concluzia că aduce o îmbunătățire semnificativă, atât în ceea ce privește timpul de execuție, cât și în ceea ce privește calitatea, față de metodele interactive de calibrare a camerelor existente, care se bazează pe indicarea utilizatorului unde să plaseze modelul de tablă de control de calibrare pentru a minimiza eroarea medie de reproiecție în calibrarea finală.

În cele din urmă, facem o punte între cele două direcții de cercetare distincte, adăugând la norul de puncte 3D reperele de pe frunte și păr, care lipseau modelului nostru, prin segmentarea părului din imagini folosind același automat celular descris inițial. După ce am depășit cu succes toate aceste etape, putem încerca acum să deducem forma feței unei persoane prin intermediul măsurătorilor faciale 3D, utilizând, de asemenea, informațiile de adâncime, și să deschidem calea pentru cercetări viitoare în domenii precum sistemele de recomandare a modei și analiza criminalistică.

Structura tezei

Teza este organizată în capitole, după cum urmează, fiecare capitol fiind împărțit în secțiuni care se referă la ceea ce am considerat a fi entități logice distincte:

- **Capitolul 2** familiarizează cititorul cu stadiul actual al tehnologiei atât în ceea ce privește automatele celulare pentru segmentarea imaginilor, cât și în ceea ce privește reconstrucția facială 3D la momentul scrierii acestui articol; aici vom discuta o scurtă istorie a automatelor celulare ca și concept, încercând să răspundem de ce o entitate atât de simplă poate produce rezultate atât de imprevizibile, urmată de cele mai comune tipuri de vecinătăți și de aplicarea lor la segmentarea imaginilor; continuăm cu un scurt interludiu despre segmentarea părului uman în imagini și începem să descriem sistemul de reconstrucție facială 3D, unde vom analiza subiecte precum stadiul actual al învățării profunde pentru extragerea facială dintr-o singură imagine, seturile de date existente pentru modele deformabile și abordările de inferență a formei feței.
- **Capitolul 3** se concentrează asupra implementării și validării numerice a tehnicii hibride nesupravegheate Grow Cut pentru segmentarea imaginilor, detaliind separat pragul adaptiv și pragul prioritar, și studiind efectele aplicării lor separate și simultane pentru a obține automatul nostru numit "Grow Cut nesupervizat hibrid".
- **Capitolul 4** se ocupă de funcționarea internă a sistemului de reconstrucție facială și a algoritmului de inferență a formei feței propus de noi, acoperind subiecte precum detectarea reperelor faciale, înregistrarea norului de puncte, ajustarea norului de puncte, modele deformabile cuplate cu interpolarea funcției de bază radială și, în cele din urmă, inferența formei feței construită pe baza feței reconstruite 3D.
- **Capitolul 5** concluzionează această teză și indică unele direcții de cercetare viitoare promițătoare.

Cuprinsul tezei

1	Introducere	1
1.1	Motivația cercetării	1
1.2	Conținutul cercetării	2
1.3	Structura tezei	4
1.4	Contribuții originale	5
1.5	Publicații	6
2	Situația actuală în domeniu	9
2.1	Definiția și proprietățile automatelor celulare	9
2.2	Automate celulare bidimensionale pentru segmentarea imaginilor	18
2.3	Tehnici de reconstrucție facială tridimensională	21
2.4	Calibrarea camerelor mono și stereo	28
2.5	Modelarea unei probleme NP-complete	30
2.6	Inferența formei feței în literatura de specialitate	33
3	Un automat celular nesupervizat hibrid bazat pe Grow Cut	35
3.1	Segmentarea imaginilor	35
3.2	Modificări ale regulilor	38
3.3	Hiper-parametrii algoritmului	43
3.4	Măsuri de performanță	44
3.5	Experimente numerice	46
3.6	Segmentarea nesupervizată cu clusterizare k-means	53
3.7	Segmentarea nesupervizată cu ajutorul învățării profunde	57
4	Un sistem geometric de reconstrucție facială 3D	61
4.1	Descrierea sistemului	61
4.2	Extragerea liniei părului	68
4.3	Rezultatele reconstrucției 3D	69
4.4	Calitatea reconstrucției 3D	76
4.5	Calibrarea optimă a camerei	81
4.6	Calitatea calibrării camerei	87
4.7	Discuție privind inferența formei feței	93
5	Concluzii și direcții viitoare de cercetare	101
5.1	Concluzii	101
5.2	Direcții viitoare de cercetare	105
	Referințe	107

Contribuții originale

Autorul ar dori să rezume aici, extrasă din fiecare publicație corespunzătoare relevantă, contribuția originală sub forma unor îmbunătățiri ale algoritmilor existenți, demonstrații experimentale sau tehnici noi. Elementele sunt enumerate în ordinea relevanței lor respective pentru traseul de cercetare al autorului, primul dintre ele (**reconstrucția facială umană 3D**) fiind obiectivul principal, urmat de cercetarea fundamentală și de potențialele scenarii de aplicare:

- 1 – un sistem robust, preponderent geometric, pentru reconstrucția facială 3D din imagini cu mai multe ipostaze, ca o condiție prealabilă pentru crearea de avatare și inferența formei feței prin analiza reperelor faciale.
- 2 - o analiză aprofundată a avantajelor unei astfel de sistem geometric față de tehnicile de reconstrucție facială prin învățare profundă, care sunt în prezent cele mai favorizate.
- 3 - o analiză amănunțită a importanței unei calibrări corecte a camerei și un algoritm evolutiv nou care asigură o calibrare optimă a camerei, având în vedere un set mare de imagini de calibrare inițiale.
- 4 - un nou set de reguli pentru automatul celular Grow Cut, și anume pragurile adaptive și prioritare, care, atunci când sunt aplicate la sarcina de segmentare a imaginilor medicale prin rezonanță magnetică, dau rezultate semnificativ mai precise.
- 5 - un mecanism automat de însămânțare pentru implementarea hibridă Grow Cut, care asigură plasarea optimă a semințelor inițiale de regiune în imaginea de intrare; acest lucru, la rândul său, este în conformitate cu paradigma nesupravegheată și complet autonomă.
- 6 - o versiune nouă, pregătită pentru paralelizare, a algoritmului hibrid Grow Cut, care rulează la rate de cadre interactive pe hardware OpenCL standard.
- 7 - o extensie și o analiză a performanței algoritmului hibrid Grow Cut pe imagini medicale 3D și generalizarea ulterioară la N dimensiuni.
- 8 - un sistem naiv Bayes de inferență a formei feței construit pe baza caracteristicilor faciale de referință, ca precursor pentru un sistem de recomandare a modei.

Contribuțiile originale ale autorului sunt împărțite între subiecte de cercetare fundamentală și aplicativă. Punctul 1 discută subiecte de cercetare fundamentală, concentrându-se pe un nou model geometric pentru reconstrucția facială, indiferent de scopul final, fie că este vorba de inferența formei feței, crearea de avatare virtuale etc. Punctul 3 reprezintă încercarea noastră de a răspunde la întrebarea dacă tehnicile actuale de calibrare a camerei sunt suficient de precise și stabile și propune un algoritm genetic pentru a găsi subsamblul optim de imagini de calibrare care, la rândul său, produce calibrarea cu cea mai mică eroare. Din nou, acest subiect este centrat pe cercetarea fundamentală, deoarece analizează complexitatea procesului de calibrare a camerei, sensibilitatea sa în raport cu imaginile de intrare și soluția propusă de algoritmul nostru genetic care abordează aceste dezavantaje. Punctele 5 și 6 menționate anterior

constituie modificări ale unui algoritm existent de segmentare a imaginilor care îi îmbunătățesc precizia (punctul 5) și performanța de calcul (punctul 6).

Elementul 2 este un subiect de cercetare aplicată care se concentrează pe o comparație între un sistem de reconstrucție a feței, în principal geometric, și abordări de învățare profundă care îndeplinesc aceeași sarcină. Punctele 4 și 7 constituie studiul aplicat al unui automat celular hibrid de segmentare a imaginilor, concentrându-se pe ajustarea acestuia și pe observarea rezultatelor aplicării sale pe imagini medicale. În cele din urmă, punctul 8 este un studiu de caz al implementării unui sistem de bază de inferență a formei feței, aplicat pe un cadru de detectare a reperelor care ia în considerare disponerea reperelor pentru a clasifica fața subiectului într-una dintre cele 7 categorii de forme de față general recunoscute.

Concluzii

Într-o manieră ascendentă, am definit conceptele de automate celulare, segmentarea imaginilor, algoritmul Grow Cut și cele două variante ale sale, supravegheată [1] și nesupravegheată [2], urmată de reconstrucția facială 3D și tehnicile de ultimă generație pentru realizarea acesteia, cu o discuție aprofundată și exemple ilustrative cu privire la motivele pentru care aceasta este greșită, făcând câteva remarci cheie cu privire la importanța calibrării camerei înainte de achiziția imaginii și, în cele din urmă, discutând structura unui modul de bază de inferență a formei feței construit pe baza tehnicilor anterioare. Toate aceste concepte trebuie considerate în cadrul subiectelor mult mai largi ale prelucrării imaginilor și vederii computerizate, așa cum sugerează și titlul tezei.

Am prezentat o versiune îmbunătățită a algoritmului clasic UGC [2] care este performant, dar nu se limitează la segmentarea imaginilor medicale. Principala noastră contribuție este o pragare adaptivă în care pragul este redus odată cu numărul de iterații și o pragare prioritara, în care pixelilor neetichetați din imaginea originală li se acordă o prioritate mai mare decât celor deja etichetați [3]. Am obținut o creștere semnificativă a calității rezultatului, așa cum reiese din metrica coeficientului Dice, printr-o modificare dinamică a parametrului pragului local, care determină granularitatea segmentării. Ideea principală care stă la baza acestor modificări provine din inspecția vizuală a evoluției automatului celular Grow Cut simplu, în care am dorit să minimizăm numărul de iterații până când automatul ajunge la o stare adecvată, lucru realizat prin acordarea de prioritate celulelor/pixelilor neetichetați față de cei etichetați și, în al doilea rând, să încurajăm fuziunea segmentelor de imagine vecine, întârziată prin scăderea progresivă a pragului local după trecerea unui anumit număr de iterații. Acest lucru oferă segmentelor/regiunilor de imagine suficient timp pentru a se forma, a crește și a se unifica.

Am oferit o analiză aprofundată a algoritmului UGC și a variantelor sale, urmată de o discuție privind beneficiile fiecărei combinații distincte de tehnici și cu scopul de a depăși calitativ dezavantajele algoritmului de clusterizare k-means, utilizat de regulă. Este expusă

rațiunea din spatele comparației cu clusterizarea k-means. Judecând după rezultate, abordarea noastră hibridă UGC depășește atât metoda de clusterizare k-means, cât și algoritmul simplu nesupravegheat Grow Cut [4].

În al doilea rând, am propus un sistem de reconstrucție a feței 3D din mai multe imagini capturate, folosind o abordare de tip Active Appearance Model (AAM) orientată spre clasificarea formei feței [5]. Propunerea noastră rezultă din observația că majoritatea lucrărilor publicate în prezent, bazate pe CNN-uri și pe abordări de învățare profundă pornind de la o singură imagine sau de la mai multe imagini, nu recuperează corect forma conturului geometriei 3D rezultate [15]. Motivul din spatele acestei situații, după cum am explicat, este faptul că tehnicile bazate pe învățare profundă "învață" o reprezentare statistică a feței umane și, prin urmare, sunt foarte predispuse la valori aberante (adică persoane cu trăsături faciale neobișnuite). Interesul nostru este să acoperim cât mai mult posibil spectrul tuturor trăsăturilor faciale umane.

În experimentele noastre, sunt utilizate două seturi de date remarcabile, și anume modelul facial Basel 2009 [11] și setul de date de imagini faciale 3D BU-3DFE [12]. Reamintim pe scurt aici caracteristicile acestora. Modelul facial Basel 2009 este un model generic parametrizat dobândit din mai multe scanări faciale 3D, care este redus la 200 de parametri prin intermediul analizei componentelor principale (PCA). Acest lucru înseamnă că, prin variația oricărei combinații a acestor parametri, putem genera procedural un set teoretic infinit de modele viabile de fețe umane. Am exploatat acest aspect în timpul configurației noastre experimentale [6]. Celălalt set de date, și anume setul de date BU-3DFE, este format din 100 de subiecți 3D (56 de femei și, respectiv, 44 de bărbați, de diferite etnii și vârste) achiziționați printr-o metodă de scanare laser 3D. Metoda de achiziție introduce zgomot semnificativ, iar numărul variabil de vertecși pentru modelul 3D final limitează gama de experimente pe care le putem efectua pe acest set de date.

Clasificarea formei feței este de obicei modelată și implementată pe baza rapoartelor distanțelor dintre punctele periferice de pe față sau învățată din seturi de imagini adnotate folosind proporțiile feței [7]. Teoretic, clasificarea fețelor ar putea fi realizată pe o imagine 2D într-o poziție absolut neutră. Cu toate acestea, astfel de cerințe sunt aproape imposibil de implementat pe sistemele care funcționează "în natură". În consecință, este vitală abordarea unei metode de reconstrucție 3D a feței care să păstreze forma conturului, indiferent de condițiile de captare a imaginii.

Am implementat reconstrucția 3D pentru a obține conservarea formei conturului pe modelul 3D reconstruit pentru a putea calcula corect proporțiile feței sau pentru a aplica o metodă de învățare automată pe fața aliniată reconstruită. Utilizăm un model generic de cap 3D care poate fi ușor de schimbat cu orice alt model pentru a realiza atât calibrarea camerei, cât și optimizarea norului de puncte 3D obținut. Acest model generic de cap 3D este modelul de față Basel 2009, cu toate componentele din analiza componentelor principale setate la zero. Conform documentației modelului, acest lucru ar trebui să producă o față umană medie din punct de vedere statistic, neutră din punct de vedere al genului, ceea ce este exact ce necesită abordarea noastră. Acesta servește drept bază pentru creșterea densității de vertecși a reconstrucției faciale 3D prin interpolare cu funcție de bază radială și pentru etapa finală de texturare a feței reconstruite [5, 6].

Optimizarea și rafinarea norului de puncte 3D se realizează cu ajutorul optimizării prin metoda celor mai mici pătrate medii, prin intermediul algoritmului Levenberg-Marquardt. În forma sa cea mai simplă, problema de optimizare se reduce la rezolvarea unui sistem neliniar de ecuații, în care necunoscutele sunt coordonatele 3D ale reperelor și matricele intrinsecă și extrinsecă ale camerei, iar funcția obiectiv de minimizat este eroarea medie de reproiecție (adică RMS).

De asemenea, adăugăm la tehnicile existente aspectul de detectare și reconstrucție a frunții, care este crucial pentru o reconstrucție facială completă. O determinare precisă a punctelor frunții este un aspect vital în calcularea metricilor faciale necesare pentru clasificarea formei feței. Metoda noastră a fost testată în mediul aplicativ și este capabilă să reconstruiască cu succes modele faciale 3D ale unor persoane aleatorii. Principala problemă asociată cu reconstrucția frunții este că majoritatea detectoarelor de repere faciale ignoră complet acest subiect, concentrându-se, așa cum sugerează și numele lor, strict pe reperele faciale. Cu toate acestea, putem reutiliza automatul celular UGC [3, 4] pentru sarcina de segmentare a părului, acordând o atenție deosebită plasării semințelor inițiale ale UGC concentrate în regiunea de deasupra feței, unde se preconizează că se află părul subiectului. Acest lucru ne permite, la rândul său, să determinăm linia părului și să selectăm un set de puncte care se află pe aceasta, evitând astfel necesitatea de a re-antrena un detector de repere personalizat care detectează și punctele de pe frunte.

Am conceput în același timp o abordare bazată pe algoritmi genetici pentru selectarea imaginilor de calibrare stereo optime dintr-un set mai mare de achiziții. Am demonstrat că o metodă de optimizare bazată pe algoritm genetic pentru selectarea subansamblului optim de imagini de calibrare ar putea duce la o îmbunătățire a rezultatelor calibrării (mono sau stereo) cu ordine de mărime între 2 și 10 [10]. Elementul central al acestei tehnici este un algoritm genetic special adaptat, cu mecanisme suplimentare pentru a face față căderii în atractori de minim local. În acest scop, aducem două contribuții sub forma unui eveniment de cataclism, declanșat ori de câte ori fondul de cromozomi devine învechit sau suprapopulat cu indivizi similari, care distruge un anumit procent din cei mai apți indivizi (la fel ca evenimentele cataclismice naturale), și a grădinii Edenului, care servește ca depozit izolat pentru cele mai bune exemplare exact înainte de declanșarea unui cataclism.

Spre deosebire de abordările de ultimă generație cu care am comparat [13, 14], sistemul nostru este unul unificat și complet automat, nefiind nevoie de nicio intervenție din partea utilizatorului final, în afară de realizarea mai multor instantanee de imagine ale unui model de calibrare, în poziții rezonabil distincte. Timpul de execuție al abordării noastre este foarte scurt în comparație cu o explorare completă prin forță brută a spațiului de soluții și produce parametri de calibrare optimizați nu numai în funcție de RMS sau de eroarea medie a distanței liniei epipolare, ci și de cerințe suplimentare de modelare din lumea reală, cum ar fi eroarea medie a distanței dintre coordonatele 3D reale ale colțului din stânga sus al modelului și toate celelalte colțuri rămase ale modelului de tablă de șah.

Am abordat procesul de calibrare încercând să arătăm că un proces obișnuit de capturare a imaginilor (mono/stereo) ar putea introduce imagini care ar putea fi în detrimentul procesului de calibrare și am propus un algoritm de selecție fără a explora întregul spațiu de soluții - care ajută la îmbunătățirea calității rezultatului calibrării. După cum s-a afirmat, o astfel de abordare se situează între explorarea pură prin forță brută și căutarea aleatorie naivă a spațiului de soluții

și presupune că, prin schimbul de material genetic (adică imagini de calibrare) între cromozomi buni (adică seturi de imagini de calibrare), am putea obține indivizi și mai buni. Operatorii de mutație și de încrucișare sunt mutația clasică de tip "bit flip" și încrucișarea cu un singur punct de tăietură, iar procesul de selecție este unul elitist. Există o lipsă în tehnicile de ultimă generație pe această temă, în timp ce majoritatea abordărilor, în loc să se concentreze pe setul optim de imagini de calibrare dintr-un set mai mare, se concentrează pe evoluția directă a matricelor intrinseci și extrinseci ale camerei.

În cele din urmă, modulele de estimare a formei feței disponibile în prezent fac mai multe presupuneri: în primul rând, acestea presupun că persoana reprezentată în imagine are o poziție aproape frontală. În al doilea rând, deoarece se bazează pe imagini, ceea ce implică proiecții 2D ale feței umane, este destul de dificil să se extragă informații despre măsurătorile legate de adâncime, cum ar fi lungimea bărbiei.

În ceea ce privește viitoarele direcții de cercetare, metoda noastră hibridă UGC ar putea fi ușor extinsă la segmentarea imaginilor 3D, necesitând o adaptare minimă a parametrilor datorită numărului crescut de interacțiuni între voxeli vecini. Abordarea noastră are potențialul de a fi utilizată ca element constitutiv al unui sistem de diagnosticare asistată de calculator (CAD). Având în vedere că am ales calea nesupravegheată și complet autonomă, prioritatea noastră este de a implementa un mijloc de evaluare nesupravegheată a calității pentru segmentarea noastră. Un alt pas următor important va fi găsirea parametrilor optimi (cum ar fi pragul global inițial) care pot fi partajați între achiziții. Ne propunem să realizăm acest lucru prin evoluția parametrilor cu ajutorul unui algoritm genetic. Am înregistrat deja un succes semnificativ în această direcție de cercetare cu algoritmul genetic modificat pe care l-am aplicat la sarcina de calibrare optimă a camerei, un alt subiect abordat în conținutul prezentei teze [10].

Sistemul nostru de reconstrucție facială 3D atinge o acuratețe foarte bună, luând în considerare distanța și metricile de deviație normală la evaluarea pe CPU, deoarece aceasta este cea mai importantă condiție prealabilă. Ne propunem să furnizăm un șablon de bază pentru aplicații care să sprijine clasificarea formei feței și crearea de avatare 3D personalizate. Pentru inferența formei feței, am utilizat un clasificator naiv Bayes pe baza unor distanțe de repere faciale selectate în mod specific și a anumitor raporturi. Setul de date pe care am antrenat acest sistem de inferență este alcătuit din 604 mostre de date de antrenament și 115 mostre de date de testare, pentru care s-a obținut o precizie de 85% [7].

Odată ce modelul feței este extras cu precizie, putem măsura toate distanțele și unghiurile necesare direct pe modelul 3D și, prin urmare, putem dezvolta un algoritm clasic bazat pe reguli. Deși extragerea unui model 3D al feței pentru a estima fața subiectului poate duce la dezvoltarea unui algoritm simplu de determinare a formei feței bazat pe reguli, problema este că regulile utilizate în determinarea formei feței sunt foarte subiective. Prin urmare, avem în vedere dezvoltarea unui model CNN bazat pe grafuri pentru a analiza relațiile dintre toate reperele faciale relevante și pentru a recunoaște automat forma feței [8, 9].

Publicații

Mai jos, cititorul poate găsi o listă care detaliază materializarea cercetărilor noastre actuale sub formă de lucrări susținute la conferințe sau articole în reviste. Alături de acestea, este trecut punctajul publicației, așa cum a fost determinat din schema de notare valabilă la momentul înscrierii autorului ca doctorand:

- **A. Marinescu**, Z. Bálint, L. Dioșan, A. Andreica, Dynamic autonomous image segmentation based on Grow Cut, ESANN 2018, 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, i6doc, ISBN 978-287-587-047-6, Bruges, April 25-27, 2018, pp. 67-72. *Conferință rang B, 4 autori - 2p.*
- **A. Marinescu**, Z. Bálint, L. Dioșan, A. Andreica, Unsupervised and Fully Autonomous 3D Medical Image Segmentation based on Grow Cut, SYNASC 2018, 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, NCA Workshop, IEEE Computer Society, ISBN 978-1-7281-0624-3, Timișoara, September 20-23, 2018, pp. 401-408. *Conferință rang C, 4 autori, workshop publicat în volumul conferinței - 0.5p.*
- **A. Marinescu**, T. Ileni, A. Dărăbant, A Versatile 3D Face Reconstruction from Multiple Images for Face Shape Classification, SOFTCOM 2019, 2019 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Split, September 19-21, 2019, pp. 1-6. *Conferință rang B, 3 autori - 4p.*
- **A. Marinescu**, T. Ileni, A. Dărăbant, A Fast and Robust, Forehead-Augmented 3D Face Reconstruction from Multiple Images using Geometrical Methods, SOFTCOM 2020, 2020 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Hvar, September 17-19, 2020, pp 1-6. *Conferință rang D, 3 autori - 1p.*
- **A. Marinescu**, Automatic Face Shape Classification via Facial Landmark Measurements, Studia Universitatis Babeș-Bolyai Informatica, Volume LXVI, Number 2 (December 2021), pp. 69-78. *Jurnal categoria D, 1 autor - 1p.*
- D. Borza, A. Dărăbant, T. Ileni, **A. Marinescu**, Effective Online Knowledge Distillation via Attention-Based Model Ensembling, Mathematics 2022, 10(22):4285, DOI: 10.3390/math10224285. *Jurnal categoria A, 4 autori - 4p.*
- D. Borza, T. Ileni, **A. Marinescu**, A. Dărăbant, Teacher or supervisor? Effective online knowledge distillation via guided collaborative learning, Elsevier ScienceDirect, Computer Vision and Image Understanding, Volume 228, 2023, 103632, ISSN 1077-3142, DOI: 10.1016/j.cviu.2023.103632. *Jurnal categoria B, 4 autori - 2p.*

Referințe

- [1] Vezhnevets, V. and Konouchine, V. (2005). *GrowCut - interactive multi-label N-D image segmentation by cellular automata*. In *Russian Academy of Sciences*, pages 1–7. Russian Academy of Sciences.
- [2] Ghosh, P., Antani, S. K., Long, L. R., and Thoma, G. R. (2011). *Unsupervised GrowCut: Cellular automata-based medical image segmentation*. In *Healthcare informatics, imaging and systems biology (HISB), 2011 first IEEE international conference on*, pages 40–47.
- [3] A. Marinescu, Z. Bálint, L. Dioșan, A. Andreica, *Dynamic autonomous image segmentation based on Grow Cut*, *ESANN 2018, 26th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, i6doc, ISBN 978-287-587-047-6, Bruges, April 25-27, 2018*, pp. 67-72.
- [4] A. Marinescu, Z. Bálint, L. Dioșan, A. Andreica, *Unsupervised and Fully Autonomous 3D Medical Image Segmentation based on Grow Cut*, *SYNASC 2018, 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, NCA Workshop, IEEE Computer Society, ISBN 978-1-7281-0624-3, Timișoara, September 20-23, 2018*, pp. 401-408.
- [5] A. Marinescu, T. Ileni, A. Dărăbant, *A Versatile 3D Face Reconstruction from Multiple Images for Face Shape Classification*, *SOFTCOM 2019, 2019 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Split, September 19-21, 2019*, pp. 1-6.
- [6] A. Marinescu, T. Ileni, A. Dărăbant, *A Fast and Robust, Forehead-Augmented 3D Face Reconstruction from Multiple Images using Geometrical Methods*, *SOFTCOM 2020, 2020 International Conference on Software, Telecommunications and Computer Networks, IEEE Xplore, ISSN 1847-358X, Hvar, September 17-19, 2020*, pp 1-6.
- [7] A. Marinescu, *Automatic Face Shape Classification via Facial Landmark Measurements*, *Studia Universitatis Babeș-Bolyai Informatica, Volume LXVI, Number 2 (December 2021)*, pp. 69-78.
- [8] D. Borza, A. Dărăbant, T. Ileni, A. Marinescu, *Effective Online Knowledge Distillation via Attention-Based Model Ensembling*, *Mathematics 2022, 10(22):4285, DOI: 10.3390/math10224285*.
- [9] D. Borza, T. Ileni, A. Marinescu, A. Dărăbant, *Teacher or supervisor? Effective online knowledge distillation via guided collaborative learning*, *Elsevier ScienceDirect, Computer Vision and Image Understanding, Volume 228, 2023, 103632, ISSN 1077-3142, DOI: 10.1016/j.cviu.2023.103632*.
- [10] A. Marinescu, A. Dărăbant, T. Ileni, *Optimal Stereo Camera Calibration via Genetic Algorithms*, *IJCAI 2021 AI4AD Workshop, Artificial Intelligence for Autonomous Driving, Montreal, Canada, August 20, 2021, <https://www.ai4ad.net/accepted-papers>*.
- [11] Paysan, P., Knothe, R., Amberg, B., Romdhani, S., and Vetter, T. (2009). *A 3D face model for pose and illumination invariant face recognition*. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, pages 296–301.
- [12] Yin, L., Wei, X., Sun, Y., Wang, J., and Rosato, M. J. (2006). *A 3D facial expression database for facial behavior research*. *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pages 211–216.
- [13] Peng, S. and Sturm, P. (2019). *Calibration wizard: A guidance system for camera calibration based on modelling geometric and corner uncertainty*. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1497–1505.
- [14] Abeles, P. (2016). *BoofCV v0.25*. <http://boofcv.org/>. Accessed: 2021-04-13.
- [15] Feng, Y., Wu, F., Shao, X., Wang, Y., and Zhou, X. (2018). *Joint 3D face reconstruction and dense alignment with position map regression network*. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 534–551.