



(12) CERERE DE BREVET DE INVENȚIE

(21) Nr. cerere: a 2022 00292

(22) Data de depozit: 30/05/2022

(41) Data publicării cererii:
30/01/2023 BOPI nr. 1/2023

(71) Solicitant:
• UNIVERSITATEA TEHNICĂ DIN
CLUJ-NAPOCA, STR.MEMORANDUMULUI
NR.28, CLUJ-NAPOCA, CJ, RO;
• ROBOTICS.AI S.R.L.,
STR.BIS.ORTODOXE, NR.11,
CLUJ-NAPOCA, CJ, RO

(72) Inventatori:
• MOLNAR SZILARD, STR.BERZEI, NR.10,
SFÂNTU-GHEORGHE, CV, RO;
• KELENYI BENJAMIN, STR.BUCOVINA,
NR.6, CLUJ-NAPOCA, CJ, RO;
• TAMAS LEVENTE,
STR.BISERICII ORTODOXE, NR.18,
CLUJ-NAPOCA, CJ, RO

(54) **METODĂ PENTRU ESTIMAREA NORMALELOR
PENTRU CAMERE CU INFORMAȚIE DE DISTANȚĂ
EMIȚĂTOARE DE IMPULSURI, FOLOSIND REȚELE
NEURONALE CONVOLUȚIONALE**

(57) Rezumat:

Invenția se referă la o metodă de calculare automată a normalelor de pe suprafețe situate pe scanări 3D, create de camere cu senzori de tip ToF (Time of flight). Metoda conform invenției cuprinde o componentă care folosește rețele neuronale convoluționale pentru calcularea normalelor dintr-un nor de puncte cu date spectrale, norul fiind creat de o cameră ToF și stocat ca o imagine de adâncime, completată cu informații spectrale și constă în:

- preluarea imaginii de adâncime și de spectru color cu ajutorul camerei și crearea unui set de date necesare antrenării rețelei neuronale;

- sincronizarea imaginilor de adâncime cu cele de spectru color pentru antrenarea unui model de tip FPN (Feature Pyramid Network);
- realizarea modelului neuronal de tip FPN pentru date multimodale;
- exploatarea modelului neuronal pentru estimarea normalelor, codificat în spațiul RGB;
- obținerea unui nor de puncte prin combinarea normalelor în spațiul 3D împreună cu informația de distanță obținută de la imaginile de intrare.

Revendicări: 2
Figuri: 3



OFICIUL DE STAT PENTRU INVENȚII ȘI MĂRCI	
Cerere de brevet de invenție	
Nr.	2022 00292
Data depozit	30-05-2022

Metodă pentru estimarea normalelor pentru camere cu informație de distanță și de spectru color, folosind rețele neuronale convoluționale

Inventatori: Szilard Molnar¹, Benjamin Kelenyi¹, Levente Tamas²
Universitate Tehnică din Cluj-Napoca, România¹
Robotics.AI SRL²

Invenția se referă la o metodă de calculare automată a normalelor de pe suprafețele situate pe scanări 3D, create de către camere cu senzori de tipul "Time-of-Flight" (ToF). Metoda cuprinde o componentă care folosește rețele neuronale convoluționale (Convolutional Neural Networks - CNN) pentru calcularea normalelor dintr-un nor de puncte cu date spectrale. Norul de puncte este creat de o cameră ToF și stocat ca o imagine de adâncime și informația spectrală, coloră. CNN-ul de tip Feature Pyramid Networks (FPN) se bazează pe stocarea informației pe 3 canale: informația de adâncime pe 1 canal, și informația spectrală pe două canale. Informațiile obținute pe cele 3 canale au fost antrenate folosind seturi mari de date, atât reale cât și sintetice. În faza evaluării, CNN-ul este capabil să estimeze normalele pentru un nor de puncte, astfel încât aceasta să fie mai rapidă și mai robustă decât metodele existente.

Pentru calcularea normalelor există multe metode, din care o subcategorie aparține metodelor cu învățare automată conform referințelor [5],[6],[7],[8] iar cealaltă categorie fiind cea a metodelor tradiționale.

Metodele tradiționale:

Abordarea clasică de calculare a normalelor este bazată pe analiza componentei principale (Principal Component Analysis - PCA) [4], care afirmă că se poate considera un suport pentru fiecare punct într-un nor de puncte, iar analizând covarianța punctelor din interiorul acestui suport, se obține vectorul propriu care corespunde celei mai mici valori proprii. Acest vector este considerat ca fiind normala punctului respectiv. Câteva metode au analizat efectul mărimii bazei de suport asupra calității normalei estimate [9,10,11], dar asta înseamnă că și complexitatea programului crește [12].

Alte metode tradiționale folosesc celule de tip Voronoi [13], [14],[15]. Aceste metode apar ca fiind instabile, dar dacă sunt bazate pe PCA, atunci se obțin rezultate îmbunătățite [16]. Alții au folosit abordări multiscale, și au obținut metode robuste, în ceea ce privește zgomotul [17,18], aici putem să menționăm și transformările Hough [19].

Presupunând că suporturile sunt planare, chiar și pe margini, se obțin rezultate mai robuste [20,21].

Metode bazate pe învățare automată:

Prin folosirea învățării automate se pot obține îmbunătățiri considerabile [22],[5],[23],[24],[6],[25],[7],[3]. Metodele cu învățare automată sunt populare chiar dacă sunt lente și au nevoie de seturi mari de date.

În primul rând, se poate folosi proiecția seturilor de puncte în 2D, cu care sunt obținute imagini de adâncime, astfel încât se pot folosi și operații de convoluție tradiționale, făcând estimarea normalelor ușoară și rapidă.

În al doilea rând, câteva metode se concentrează pe seturi de puncte discrete nestructurate [26], începând de aici, s-au obținut metode pentru estimarea normalelor [22,25], introducând abordări orientate înspre seturile multiscalare. Cu ajutorul acestor metode și prin adaugarea vectorilor de tip Fischer [3] combinați cu anumiți experți (metode de decizie) în faza de estimare a normalelor au evoluat noile metode.

Abordări multimodale:

Abordările multimodale folosesc date heterogene, de exemplu intrări radiometrice sau spectrale [27].

Brevete similare:

Majoritatea metodelor sunt bazate pe metoda Random sample consensus (RANSAC), diferența fiind fie stocarea datelor și abordările propuse.

[32] estimează cromaticitățile pentru fiecare pixel. Metoda acesta are nevoie de cel puțin trei surse de lumină și două camere, din care cel puțin o cameră trebuie să fie o cameră stereo. În metoda noastră avem nevoie numai de o cameră de tip ToF.

[30,31] folosesc imagini rotite pentru estimarea poziției, care este robustă cu privire la zgomot, dar calculele sunt foarte complexe. Metodele [30,31] caută marginile în fiecare plan, și produsul lor este considerat ca un vector normal. Căutarea marginilor are un timp de procesare foarte mare, iar metoda noastră este mult mai rapidă.

În [29] este prezentată o aproximare pentru normale, similară cu modul de funcționare al algoritmului RANSAC. Modelul potrivește un plan pentru punctele din vecinătatea unui punct. Dacă planul este mic, atunci calitatea aproximării nu va mai fi atât de robustă, numai cu un plan mai mare. Pentru a obține rezultate mai bune, această metodă netezește norul de puncte, în schimb, pentru a efectua această operațiune sunt necesare calcule robuste. În contrast, metoda noastră moștenește un comportament adaptiv de la mai multe straturi ale FPN-ului.

O altă metodă este [33], care lucrează cu imagine RGB și imagini de adâncime, similar cu metoda noastră. Această metodă calculează o funcție de cost ca să aproximeze planurile, dar iarăși, planurile sunt folosite cu RANSAC ca să obțină normalele.

Mai trebuie menționat și faptul că în această metodă este integrată și asumarea de spațiu Manhattan, adică planurile sunt aliniată cu coordonatele lumii, simplificând calculele.

[28] procesează punctele 3D, deci datele au nevoie de preprocesare, în contrast metoda noastră lucrează cu datele pure obținute de la camera ToF. Din aceste cadre, metoda creează niște plasturi, care sunt utilizați în calcularea normalelor. Metoda calculează cu o mai mare precizie normalele, însă acestea sunt des interpolate, ceea ce duce la pierderi mari de informație și rezoluție.

[34] perfecționează normalele suprafețelor, și nu le generează. Metoda constă în faptul că normalele obiectelor fizice reprezintă câmpuri de normale conservative. De aceea, o hartă de înălțime bazată pe componenta conservatoare a suprafeței inițiale se poate să îmbunătățească direcțiile normalelor. Metoda noastră calculează normalele din prima procesare, nu are nevoie de normale preexistente.

[35] folosește imagini secvențiale cu condiții de lumini diferite pentru a reproduce informațiile spațiale de la imaginile RGB. Imaginile sunt împărțite în diferite zone, dintre care se pot folosi informații diferite, mai ales de la diferite imagini. Este de menționat faptul că o cameră cu laser și o tehnică de triangulație este mai inferioară decât metoda aceasta, deoarece este mult mai complexă, și se poate să piardă informații referitoare la suprafețele transparente sau non-reflectorizante. Problema asta este prezentă și la camerele RGB-D, deci și metoda noastră suferă de acest fenomen. În orice caz, dacă verificăm numai datele comune în cele două metode, ajungem la concluzia că metoda noastră funcționează mai bine, și nu are nevoie de imagini atât de multe precum metoda [35]. Faptul că metoda noastră folosește o cameră care înglobează toate informațiile necesare, metoda noastră este mult mai potrivită pentru procese de timp real sau procese dinamice.

[36] tratează punctele din spațiul 3D drept sarcini electrostatice. Dacă se consideră un punct drept centrul unei sfere și se calculează potențiale electrostatice în interiorul sferei, atunci un vector care pornește din centrul sferei până la punctul unde potențialul electrostatic este minim, vectorul respectiv poate fi considerat ca normală pentru centrul sferei. Metoda este bazată pe modelul RANSAC.

[37] folosește mai multe puncte de vedere pentru a capta o imagine a unei scene sau a unui obiect. Dacă se aliniază scena de la mai multe imagini, din diferențele minore, de exemplu luminescență, se pot calcula normalele. Un dezavantaj este când în anumite zone pozele nu se aliniază corect, iar atunci se setează un prag, care decide care dintre imagini este folosită,

în anumite cazuri se poate ca numai o poză să fie folosită pentru zona problematică.

În [38] este calculată variația medie a punctelor pe axele x , y și z din care rezultă matricea de covarianță. Prin descompunerea ei în valori proprii se obțin normalele. Acest fapt se datorează deoarece normalele tind să se alinieze pe direcția cu variația cea mai mare. Această metodă se afirmă prin faptul că este eficientă, nu are nevoie de calcule complexe, și este flexibilă, din cauză că folosește o mărime variabilă a vecinătății de puncte.

Metoda [28] este cea mai apropiată metodă de invenția curentă. În această lucrare, se propune o rețea de fuziune ierarhică cu reponderare adaptivă a caracteristicilor pentru estimarea normalelor dintr-o singură imagine de tip RGB-D. Mai exact, caracteristicile din imaginea color și din imaginea de adâncime sunt integrate succesiv, pentru a asigura netezirea globală a suprafeței, păstrând în același timp detaliile vizuale importante. Între timp, caracteristicile de adâncime sunt reponderate cu o hartă de încredere estimată din imaginea de adâncime înainte de a fi integrate în ramura de culoare pentru a evita artefactele cauzate de imaginea de adâncime. Diferența majoră și benefică al metodei propuse în invenția curentă, este folosirea unei rețea de tip FPN, care asigură construirea de hărți de caracteristici semantice de nivel înalt la toate scările, astfel oferind o robustețe superioară cu un timp redus de calcul.

De mai sus se vede că numărul metodelor care se ocupă cu calcularea normalelor este mare. Acest fapt nu este o coincidență, din cauză că calcularea normalelor este un aspect important în domeniul viziunii artificiale. Pe cunoștința normalelor se poate baza segmentarea imaginilor, sau aplicarea efectelor de lumină în modelarea 3D. Astfel, putem trage concluzia, cu cât mai rapidă este estimarea normalelor, cu atât se poate accelera și celelalte procese, și cu cât mai exactă este o predicție despre o normală, cu atât va îmbunătăți efectele aplicate. Ca și consecință, putem afirma ca este foarte important să avem o metodă care nu numai că este precisă, ci este agilă și robustă către zgomote.

Problema pe care o rezolvă invenția este că metodele prezentate fie procesează imaginile în puțin timp, fie sunt mai precise, dar să existe amândouă simultan, nu exista. Pentru acesta se folosește arhitectura FPN [1] și o codificare a normalelor în spațiul RGB. Adicional, este folosit și o funcție de pierdere specifică, care transformă informația de la 2D de adâncime și spectru color în 3D. Invenția se bazează pe o abordare multimodală, deoarece au fost introduse mai multe tipuri de date. Pe lângă imaginea de adâncime a fost introdusă imaginea

de infraroșu. Aceasta imagine produsă ca urmare a detectării radiațiilor electromagnetice emise sau reflectate de o anumită suprafață țintă în poziția infraroșu a spectrului electromagnetic. Abordarea menționată a servit la îmbunătățirea algoritmului pentru estimarea normalelor.

Urmează descrierea mai elaborată a invenției.

Metoda se bazează pe abilitatea de a captura caracteristici multi-scalare de la camerele ToF folosind Feature Pyramid Networks (FPN) (1) pentru estimarea normalelor conform Figurii 1. Din cauza tipului specific al camerelor ToF, măsurătorile spațiului 3D sunt codificate sub forma imaginilor de adâncime (2), care se pot proiecta în 3D ca un nor de puncte organizat. Chiar dacă între straturile piramidei și bazele de suport a normalelor, din norul de puncte, nu este o legătură strictă definită, modelul imită estimarea normalelor folosind metode multi-scalare în spațiul 3D. După o mai bună analizare a structurii modelului, se poate observa că intrarea modelului este constituit din 3 canale. Aceste canale pot fi date diferite, ceea ce certifică faptul că algoritmul nostru este o metodă multimodală. În cazul metodei propuse, pe cele 3 canale avem după cum urmează: pe 1 canal este introdusă imaginea de adâncime iar pe celelalte 2 canale avem aceeași imagine de infraroșu. Această abordare este benefică din mai multe considerente. Un beneficiu important ar fi robustețea la zgomot, deoarece imaginile de infraroșu nu prea au zgomot față de imaginile de adâncime, îmbinarea celor două tipuri de date crește performanța algoritmului. Această concluzie a fost luată după ce au fost realizate mai multe scenarii de teste. Schimbând datele de intrare pe cele 3 canale, am ajuns la consecința că asocierea imaginii de adâncime cu 2 canale de infraroșu este cea mai benefică combinație de tipuri de date din toate punctele de vedere.

Codificarea informațiilor referitoare la normale înseamnă că în spațiul 3D un vector are trei coordonate: x , y și z . Vectorii sunt considerați ca fiind vectori normalizați, adică valorile coordonatelor sunt în intervalul $[-1,1] \in \mathbb{R}$, iar lungimea vectorului este de o unitate. După stabilirea limitelor în spațiul 3D, putem să stabilim limitele în spațiul RGB. La o imagine color, de obicei sunt 3 canale (R-roșu (red), G-verde (green), B-albastru (blue)), fiecare canal fiind pe 8 biți, cu valori posibile în intervalul $[0,255] \in \mathbb{N}$ (3). Folosindu-ne de informațiile menționate, se poate face o mapare între coordonatele x , y , z și canalele R, G, B cu o pierdere minimală, care este sub un grad.

Detaliile modelului FPN prezentate în Figura 2 sunt prezentate după cum urmează:

Dimensiunea imaginii de intrare poate fi aleatoare (depinde de camera ToF folosită), pentru că la ieșire rezultatul va avea aceeași dimensiune. Ca bază, putem să folosim orice model, însă noi am folosit ResNet [2] care este un model pre-antrenat pe setul de date ImageNet.

Construirea piramidelor (1) înseamnă că trebuie să avem o cale de jos în sus (4) și o altă cale de sus în jos (5) unite prin mai multe căi laterale (6).

Calea de jos în sus (4):

Această cale constă în calculul feedforward al rețelei convoluționale, combinând hărți caracteristice de la diferite straturi.

Pentru faza de upsampling am folosit pixel shuffle cu interpolare biliniară, ca să evităm mărimile inconsistente de la hărțile caracteristice. Ieșirea unui nivel este intrarea pentru următorul nivel. Din cauză că am folosit ResNet la bază, am avut acces la caracteristica de activare pentru ieșirea blocului rezidual de la ultima etapă, notat cu C_i de la stratul i .

Calea de sus în jos (5):

Calea aceasta imită caracteristici cu rezoluții înalte, fiindcă modelul face un upsampling de la hărțile caracteristice rare, dar mai descriptive. Straturile sunt notate cu P_i . Conexiunile laterale impun aceste caracteristici incluzând informația de la calea de jos în sus.

Operatorul între ultimele două nivele conține două module convoluționale consecutive de 3×3 pentru procesarea ultimei hărți caracteristice, urmat de o funcție de activare de tip Sigmoid.

Conexiunile laterale (6):

Conexiunile de la straturile P_i sunt asigurate printr-o convoluție de 1×1 și o sumă de element cu element. Dimensiunile straturilor verticale sunt moștenite și de către straturile laterale.

Funcția de calculare a erorii:

$$L_{normal} = \frac{1}{N} \sum_1^N \left(1 - \frac{\|N_i \times N_{GT}\|}{\|N_i\| \cdot \|N_{GT}\|} \right) \quad (1)$$

În ecuația (1) se descrie funcția de calculare a erorii pentru antrenarea modelului. Funcția calculează diferențele (7) între unghiurile dintre vectorii normalelor din norul de puncte de referință (8) și din norul de puncte estimat (9). În ecuație, N este numărul punctelor din norul de puncte, N_i reprezintă vectorii normalelor estimați, iar N_r reprezintă vectorii normalelor de referință.

Ideea inovatoare iese în evidență prin faptul că metoda noastră codifică normalele de la spațiul 3D în spațiul RGB (3), care este un fapt unic. Metoda noastră ca multe dintre metodele prezentate folosește CNN, dar diferit de ele, folosește

arhitectura FPN (1) pentru asta. Am mai adăugat și o funcție de pierdere nouă, care înseamnă că în timpul antrenării rețeaua calculează normalele într-o imagine de RGB, care este proiectată înapoi în spațiul 3D, iar erorile sunt calculate în spațiul proiectat.

Metoda conform invenției presupune realizarea următoarelor etape:

1. Preluarea imaginii de adâncime și de spectru color cu ajutorul camerei și crearea setului de date necesare antrenării rețelei neuronale;
2. Sincronizarea imaginilor de adâncime cu cele de spectru color pentru antrenarea modelului de tip FPN;
3. Realizarea modelului neuronal de tip FPN pentru date multimodale;
4. Exploatarea modelului neuronal pentru estimarea normalelor codificat în spațiul RGB;
5. Obținerea unui nor de puncte prin combinarea normalelor în spațiul 3D împreună cu informația de distanță obținută de la imaginile de intrare.

Prin invenție, se obține o metodă cu care putem să estimăm normalele pentru o anumită suprafață extrasă de la o imagine făcută de către o cameră de tip ToF multispectrală. Metoda se poate folosi ca o fază de preprocesare pentru mai multe teme în domeniul viziune artificială, de exemplu clasificarea imaginilor, segmentarea obiectelor sau calcularea reflexiei luminii în modelare 3D.

Referințe

- [1] Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 936–944.
- [2] He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, 2016.
- [3] . Ben-Shabat, M. Lindenbaum, and A. Fischer, "Nesti-net: Normal estimation for unstructured 3d point clouds using convolutional neural networks," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 06 2019
- [4] . Hoppe, T. DeRose, T. Duchamp, J. McDonald, and W. Stuetzle, "Surface reconstruction from unorganized points," in Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques, ser. SIGGRAPH '92. New York, NY, USA: Association for Computing Machinery, 1992, p. 71–78. [Online]. Available: <https://doi.org/10.1145/133994.134011>
- [5] J. E. Lenssen, C. Osendorfer, and J. Masci, "Deep iterative surface normal estimation," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun 2020.
- [6] H. Zhou, H. Chen, Y. Feng, Q. Wang, J. Qin, H. Xie, F. L. Wang, M. Wei, and J. Wang, "Geometry and learning co-supported normal estimation for unstructured point cloud," 2020 IEEE/CVF Conference on

Computer Vision and Pattern Recognition (CVPR), pp. 13 235-13 244, 2020.

[7] D. Lu, X. Lu, Y. Sun, and J. Wang, "Deep feature-preserving normal estimation for point cloud filtering," *Comput. Aided Des.*, vol. 125, p.102860, 2020

[8] Z. Wang and V. A. Prisacariu, "Neighbourhood-insensitive point cloud normal estimation network," in *BMVC*, 2020.

[9] Cazals and M. Pouget, "Estimating differential quantities using polynomial fitting of osculating jets," *Computer Aided Geometric Design*, vol. 22, no. 2, pp. 121-146, 2005.

[10] R. Bormann, J. Hampp, M. Hägele, and M. Vincze, "Fast and accurate normal estimation by efficient 3d edge detection," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2015, pp. 3930-3937

[11] G. Guennebaud and M. Gross, "Algebraic point set surfaces," *ACM Trans. Graph.*, vol. 26, no. 3, p. 23-es, July 2007. [Online]. Available: <https://doi.org/10.1145/1276377.1276406>

[12] Jordan and P. Mordohai, "A quantitative evaluation of surface normal estimation in point clouds," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 4220-4226.

[13] Q. Mérigot, M. Ovsjanikov, and L. J. Guibas, "Voronoi-based curvature and feature estimation from point clouds," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 6, pp. 743-756, 2011.

[14] T. K. Dey, G. Li, and J. Sun, "Normal estimation for point clouds: a comparison study for a voronoi based method," in *Proceedings Eurographics/IEEE VGTC Symposium Point-Based Graphics*, 2005, pp. 39-46.

[15] T. K. Dey and S. Goswami, "Provable surface reconstruction from noisy samples," *Computational Geometry*, vol. 35, no. 1, pp. 124-141, 2006, special Issue on the 20th ACM Symposium on Computational Geometry

[16] Alliez, D. Cohen-Steiner, Y. Tong, and M. Desbrun, "Voronoi-based variational reconstruction of unoriented point sets," in *Proceedings of the Fifth Eurographics Symposium on Geometry Processing*, ser. SGP '07. Goslar, DEU: Eurographics Association, 2007, p. 39-48.

[17] Mura, G. Wyss, and R. Pajarola, "Robust normal estimation in unstructured 3d point clouds by selective normal space exploration," *The Visual Computer*, vol. 34, pp. 961-971, 2018.

[18] J. Zhang, J. Cao, X. Liu, H. Chen, B. Li, and L. Liu, "Multi-normal estimation via pair consistency voting," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, pp. 1693-1706, 2019.

[19] Boulch and R. Marlet, "Fast and robust normal estimation for point clouds with sharp features," *Computer Graphics Forum*, vol. 31, pp. 1765-1774, 08 2012.

[20] H. Avron, A. Sharf, C. Greif, and D. Cohen-Or, "Sparse reconstruction of sharp point set surfaces," *ACM Trans. Graph.*, vol. 29, no. 5, Nov. 2010.

[21] H. Chen, J. Huang, O. Remil, H. Xie, J. Qin, Y. Guo, M. Wei, and J. Wang, "Structure-guided shape-preserving mesh texture smoothing via joint low-rank matrix recovery," *Computer-Aided Design*, vol. 115, pp. 122-134, 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0010448519302088>

[22] Hyeon, W. Lee, J. H. Kim, and N. Doh, "Normnet: Point-wise normal estimation network for three-dimensional point cloud data,"

International Journal of Advanced Robotic Systems , vol. 16, no. 4, p. 1729881419857532, 2019.

[23] A. Boulch and R. Marlet, "Deep learning for robust normal estimation in unstructured point clouds," Computer Graphics Forum , vol. 35, 2016.

[24] X. Wang, D. F. Fouhey, and A. Gupta, "Designing deep networks for surface normal estimation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) , pp. 539-547, 2015.

[25] P. Guerrero, Y. Kleiman, M. Ovsjanikov, and N. Mitra, "Pcpnet learning local shape properties from raw point clouds," Computer Graphics Forum , vol. 37, 2018.

[26] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," 2017.

[27] M. Gao, J. Jiang, G. Zou, V. John, and Z. Liu, "RGB-D-Based Object Recognition Using Multimodal convolutional Neural Networks: A Survey," IEEE Access, vol. 7, pp. 43 110-43 136, 2019.

[28] Zeng, Jin and Tong, Yanfeng and Huang, Yunmu and Yan, Qiong and Sun, Wenxiu and Chen, Jing and Wang, Yongtian, "Deep Surface Normal Estimation With Hierarchical RGB-D Fusion, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[28] US8665267B2

[29] US9767598B2

[30] US20120330447A1

[31] US20170227942A1

[32] US10097813B2

[33] US9117281B2

[34] US9875554B2

[35] US9626767B2

[36] US009043186B2

[37] US20180025496A1

[38] US005283837A

REVEDICĂRI

1. Metoda de estimare a normalelor ce folosesc rețele neuronale convoluționale, cu o arhitectură de tip Feature Pyramid Network (FPN), prin care se extrag caracteristicile 3D dintr-o imagine de adâncime 2D pentru puncte 3D **caracterizată prin aceea că** informația de intrare este o imagine de adâncime și o imagine de tip spectru color.

2. Metoda conform revendicării 1 **caracterizată prin aceea că** intrarea modelului este o intrare multimodală pe 3 canale, adică arhitectura permite să avem la intrare mai multe tipuri de date fuzionate, care ajută la îmbunătățirea rezultatelor de estimare ale normalelor 3D.

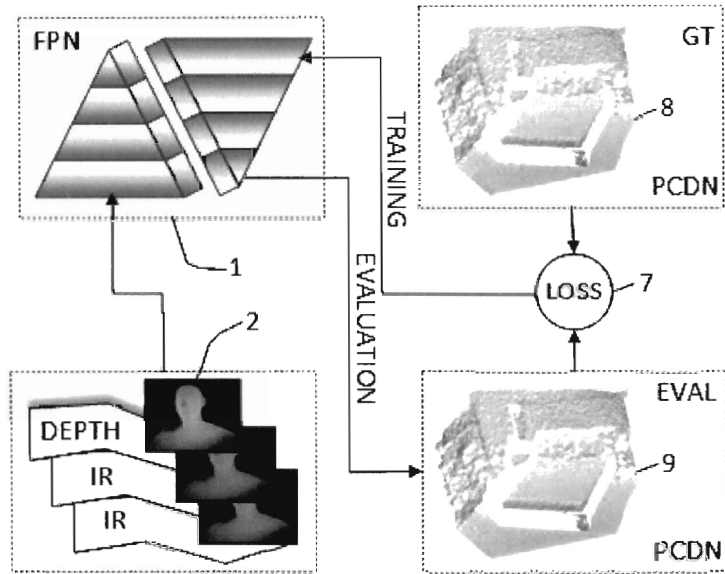


Fig .1

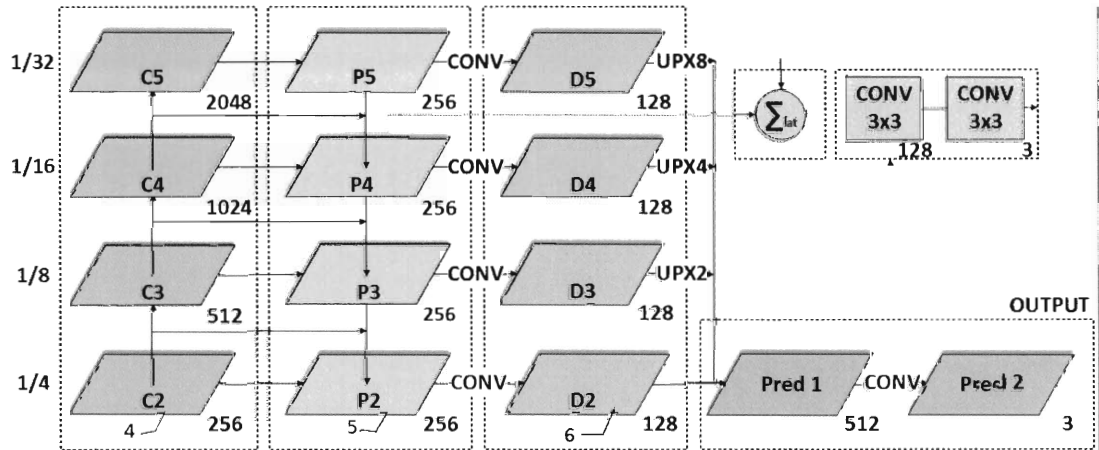
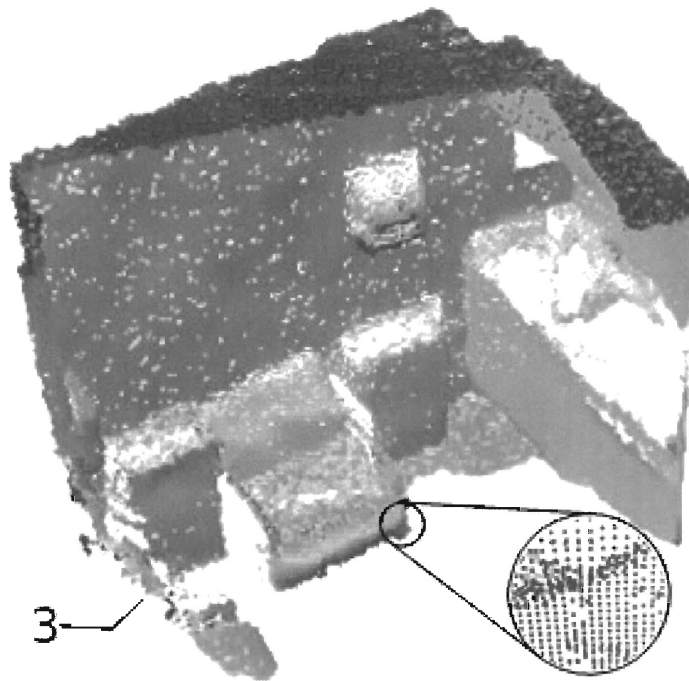


Fig.2.



3

Fig. 3