

# 3D paviršiaus formavimo 2D vaizdų pagrindu algoritmų tyrimas

**Arnas Nakrošis, Andrius Kriščiūnas**

Kauno technologijos universitetas,  
K. Donelaičio g. 73, Kaunas  
*arnas.nakrosis@ktu.lt, andrius.krisciunas@ktu.lt*

---

**Santrauka.** Tobulėjant technologijoms pramonės ir viešajame sektoriuose informacijos atvaizdavimui vis dažniau yra naudojami trimačiai vaizdai, leidžiantys dirbuotojams lengviau suprasti gautus rezultatus ar surinktą informaciją. Nors šiuo metu jau egzistuoja algoritmai leidžiantys standartinėmis priemonėmis surinktų 2D vaizdų pagrindu formuoti trimačius vaizdus, tačiau tokių algoritmų pateikiamų rezultatų kokybę dar vis smarkiai atsilieka nuo rinkos poreikio. Šio tyrimo metu buvo atliekama esamų algoritmų analizė siekiant pasiūlyti sprendimą, kuris leistų pasiekti geresnius rezultatus lyginant su literatūroje nagrinėjama atvejais. Rezultate autorių pasiūlyta algoritmu kombinacija leidžia suformuoti 3D paviršiu remiantis santykinai nedideliu 2D vaizdų kiekiu.

**Raktiniai žodžiai:** Taškinės nuotraukos, 3D paviršius, 3D taškų debesis, Poissono algoritmas, persidengiančios nuotraukos.

---

## 1 Įvadas

Tobulėjant technologijoms yra siekiama atrasti algoritmus, leidžiančius skaitmeninti objektus naudojant taškinius vaizdus. Objektų skaitmeninimas yra naudingas siekiant perteikti didesnę kiekį informacijos naudotojams viešajame ar pramonės sektoriuose. 2021 metais atlikti rinkos tyrimai pastebi, kad per ateinančius metus yra numatomas 20,9% metinis augimo tempas net rinką paveikus globaliems veiksniams [1], [2]. Atsižvelgiant į šiuos rinkos tyrimus yra matomas technologijos, kuri vis dar sparčiai tobulėja, poreikis. Siekiant metodologijas pritaikyti sparčiai besiplečiančiai rinkai ir spartėjančiam gamybos bei paslaugų tiekimo greičiui, tyrėjai nagrinėja skirtingas galimybes technologijų tobulinimui. 3D paviršiaus formavimo 2D vaizdų pagrindu generavimo algoritmus galima skirstyti į dvi pagrindines grupes: matematinėmis savybėmis ir mašininio mokymo algoritmais paremtus algoritmus.

Darbuose [3], [4] matematinėmis savybėmis paremti algoritmai naudoja 2.5D nuotraukas (naudojamas ~40 nuotraukų to paties objekto rinkinys),

kuriose kartu su vaizdine informacija yra pateikiamos gylio reikšmės bei parametras *delta*. Šiuo atveju algoritmas, pagal pateiktą vaizdų rinkinį erdvių vektorių pagalba apskaičiuoja atstumo reikšmes siekiant tolimesniuose žingsniuose sumažinti atminties sunaudojimą, bei skaičiavimams reikalingų resursų kiekį. Delta parametru valdomas erdvinių vektorių sudarymas, kuris nulemia arčiau esančių detalių svarbą ir atstumo užtikrintumą, lyginant su geometriškai tolimesnėms detalėms. Taip rezultate algoritmas erdviniams taškams priskiria ir užtikrintumo reikšmę, kuri tostant mažėja, bei kuomet elementas nėra matomas iš nagrinėjamos perspektyvos yra artima nuliui. Kitos studijos autoriai [5] pristatė algoritmą, kuris siekia sujungti privalumus iš retų 3D tinklus formuojančių algoritmų [6] ir tankius 3D tinklus formuojančių algoritmų [7]. Tai buvo pasiekta naudojant 2D nuotraukas (daugiau nei 26) ir 3D taškus. Pirmiausia, šis algoritmas, iš gautų nuotraukų išskiria atraminius taškus, kurie yra retai išdėstyti nuotraukų erdvėje, siekiant kad tie taškai turėtų kiek įmanoma didesnę reikšmę žiūrint į objektą iš skirtingų perspektyvų (nuotraukų). Taškų palyginimui skirtingose nuotraukose yra naudojamas ZNCC (angl. Zero – Mean Normalized Cross Correlation) metodas bei kryžminė patikra [8]. Taip yra sudaromas svarbių taškų sąrašas, kuris yra papildomas interpoliaciniais taškais. Šiuo procesu yra siekiama sutankinti formuojamą tinklą, gerinant paviršiaus tikslumą [9].

Atskira algoritmų grupė yra mašininio mokymusi paremti algoritmai. Darbe [10] autoriai pasiūlė algoritmą, kuris naudojant neuroninius tinklus sudarytų 3D paviršių iš retų magnetinio rezonanso medicininių vaizdų (toliau MRI) rinkinių. Šiam tikslui pasiekti yra naudojamos trys, skirtingų perspektyvų pjūvinės dalinai ortogonalios nuotraukos, kai MRI yra apdorojami gilioju konvoliuciniu neuroniniu tinklu (toliau CNN) [11] ir visų trijų MRI rezultatai yra bendrai apdorojami dviem pilnai sujungtais neuroninio tinklo sluoksniais. Gauti neuroninio tinklo rezultatai yra apdorojami pagrindinių komponentų analize (toliau PCA) [12], ir suformuojamas paviršius [13]. CNN yra apmokomas siekiant prognozuoti vidurki, poslinkį ir kovariacijos matricą. Apmokymo tikslo funkciją sudaro dvi pagrindinės dalys: paviršiaus prognozės patikimumas, siekiant maksimizuoti sąlyginę tikimybę, bei atstumo tarp prognozuoto ir tikrojo paviršiaus, matuojamo Kullback-Leibler pasiskirstymu [14]. Atlikus MRI apdorojimą naudojant CNN algoritmas apskaičiuoja tikimybes paviršiui, naudojant PCA įverčius. CNN yra sudarytas iš trijų šakų, skirtų kiekvieno MRI apdorojimui atskirai. Kiekvieną šaką sudaro

9 konvoliuciniai sluoksniai ir trys sutelkimo pagal maksimumą (angl. max pooling) sluoksniai. CNN šakų rezultatai yra sujungiami naudojant du neuroninio tinklo sluoksnius. Šis algoritmas yra pranašesnis už matematinius algoritmus tuo, kad formavimui naudoja nedidelį kiekį nuotraukų (priklausomai nuo objekto sudėtingumo reikia nuo 3 iki 9 nuotraukų) [10]. Tačiau norint apmokyti šį neuroninį tinklą yra reikalingas didelis kiekis apmokymui skirtų nuotraukų ir gautas apmokytas neuroninis tinklas nėra universalus (sunkus pritaikymas skirtingo tipo objektų formavimui) [10].

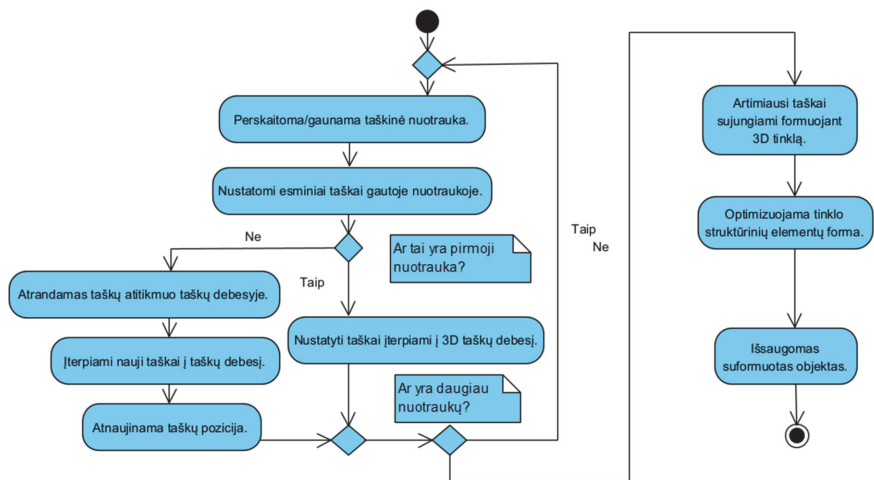
Apibendrinant 3D paviršių analizę galima teigti, kad dauguma algoritmų naudoja didelį kiekį įvesties nuotraukų, papildomų įvesčių arba yra pritaikyti tik vienai objektų grupei, kas smarkiai apriboja minimų algoritmų taikymą praktikoje. Tai lemia esamų sprendimų tobulinimo poreikį ir paklausą, kas tiesiogiai seka iš einamuoju momentu šioje srityje atliekamų tyrimų gausos. Šiame darbe autoriai nagrinėja matematinėmis savybėmis paremtą algoritmų kombinacijos sprendimą, kuri leidžia įvairių sričių objektams sudaryti 3D paviršius naudojant nedidelį kiekį taškinių nuotraukų (priklausomai nuo formuojamo objekto formos sudėtingumo), svyruoja nuo 8 iki 16 vaizdų.

## **2 Glotnaus 3D paviršiaus formavimo algoritmas**

Darbe sudaryto algoritmo skirto glotnaus 3D paviršiaus formavimui nedidelio kiekio įvesties vaizdų pagrindu, schema pateikta 1 pav. Šis algoritmas gali būti pritaikytas skirtingų formų objektams ir susideda iš trijų pagrindinių dalių: 3D taškų debesies formavimas, 3D debesies taškų optimizavimas, debesies sujungimas į 3D paviršių. 3D taškų debesies formavimas yra paremtas struktūros iš judėsio (angl. structure from motion) [15] algoritmu, kuris pasikartojančius kritinius taškus atrenka iš persidengiančių nuotraukų. Suformuotas taškų debesis yra optimizuojamas pašalinant taškus, esančius šalia kitų taškų ir esančius vienoje tiesėje su daugiau nei dviem taškais. Atlikus debesies optimizavimą, yra suformuojamas trimatis paviršius sujungiant taškus į trikampus naudojant Poissono [16] algoritmą.

## **3 Eksperimentiniai tyrimai**

Sudaryto algoritmo analizei buvo pasirinktas formuoti sudėtingos ir paprastos struktūrų objektai (2 pav.).





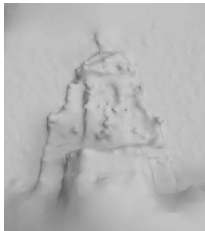
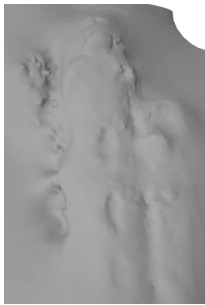


1 pav. Duomenų rinkinio pavyzdys



2 pav. Duomenų rinkinių pavyzdžiai

Pasirinkto objekto duomenų rinkinius sudarė 11 vaizdų aibės, kurių pagrindu buvo suformuotas 3D paviršius. 1 lentelėje pateikti 3D paviršiaus sudarymo rezultatai (laikas, ir nustatytų esminių taškų skaičius) taikant skirtingus algoritmus kartu pateikiant ir sudaryto paviršiaus vaizdo vizualizaciją.

**1 lentelė.** Algoritmų palyginimo lentelė, kai objektams sudaryti naudojama 11 nuotraukų/

Pavadinimas	Aurių siūlomas algoritmas		Žygiuojančių kubų algoritmas		TV-L	
Objektas 1						
	Formavimo laikas (s)	Esminių taškų skaičius	Formavimo laikas (s)	Esminių taškų skaičius	Formavimo laikas (s)	Esminių taškų skaičius
	62,32	8381	74,97	9345	69,02	7675
Objektas 2						
	Formavimo laikas (s)	Esminių taškų skaičius	Formavimo laikas (s)	Esminių taškų skaičius	Formavimo laikas (s)	Esminių taškų skaičius
	187,71	4738	192,56	5526	196,74	4257

Iš lentelėje pateiktų rezultatų yra matoma, kad autorių siūlomas algoritmų derinys vertinant vizualiai labiausiai atitinka objektus pateiktus 2 pav., bei suformuoja paviršių greičiausiai (5,66 s. vienos pirmo objekto nuotraukos apdorojimui) iš lygintų algoritmų. Antras objektas formuojamas beveik tris kartus ilgiau (17,06 s vienos nuotraukos apdorojimui), tačiau tokie rezultatai gaunami dėl to, kad nuotraukų dimensijos yra ~3 kartus didesnės.

## 4 Išvados

Darbe išanalizavus matematinėmis savybėmis bei mašiniu mokymu paremtus 3D paviršiaus formavimo algoritmus, buvo pasirinkta matematinėmis savybėmis paremtų algoritmų klasė, ir apjungus struktūros iš judėsio ir Poissono algoritmus, pasiūlytas sprendimas kuris yra pritaikomas kiek įmanoma platesniame skirtingų objektų intervale naudojant santykinai nedidelį taškinių nuotraukų įvesties kiekį. Sudarytas algoritmas suformuoja glotnų paviršių naudojant nuo 6 iki 16 nuotraukų priklausomai nuo formuojamo objekto formos sudėtingumo. Darbe pateiktiems duomenų rinkinių pavyzdžiui algoritmas suformuoja trimatį paviršių vidutiniškai vienos nuotraukos apdorojimui užtrukdamas nuo 5,66 s iki 17,06 s priklausomai nuo naudojamų nuotraukų dimensijų.

## Literatūra

- [1] "3D Mapping and Modeling Market Trends, Growth, Industry Analysis and Forecast 2022-2027." <https://www.imarcgroup.com/3d-mapping-modeling-market> (accessed Apr. 15, 2022).
- [2] "3d Mapping and 3d Modelling Market | 2022 - 27 | Industry Share, Size, Growth - Mordor Intelligence." <https://www.mordorintelligence.com/industry-reports/3d-mapping-and-3d-modelling> (accessed Apr. 15, 2022).
- [3] S. I. Young, D. B. Lindell, B. Girod, D. Taubman, and G. Wetzstein, "Non-Line-of-Sight Surface Reconstruction Using the Directional Light-Cone Transform," 2020, pp. 1407–1416. Accessed: Apr. 15, 2022. [Online]. Available: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/html/Young\\_Non-Line-of-Sight\\_Surface\\_Reconstruction\\_Using\\_the\\_Directional\\_Light-Cone\\_Transform\\_CVPR\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPR_2020/html/Young_Non-Line-of-Sight_Surface_Reconstruction_Using_the_Directional_Light-Cone_Transform_CVPR_2020_paper.html)
- [4] Curless B. and M. Levoy, "A Volumetric Method for Building Complex Models from Range Images." <https://graphics.stanford.edu/papers/volrange/> (accessed Jan. 13, 2021).
- [5] G. Zampokas, K. Tsiolis, G. Peleka, I. Mariolis, S. Malasiotis, and D. Tzovaras, "Real-Time 3D Reconstruction in Minimally Invasive Surgery with Quasi-Dense Matching," in *2018 IEEE International Conference on Imaging Systems and Techniques (IST)*, Oct. 2018, pp. 1–6. doi: 10.1109/IST.2018.8577106.
- [6] M. Kulawiak, "A Cost-Effective Method for Reconstructing City-Building 3D Models from Sparse Lidar Point Clouds," *Remote Sens.*, vol. 14, no. 5, Art. no. 5, Jan. 2022, doi: 10.3390/rs14051278.
- [7] W. Zeng, W. Ouyang, P. Luo, W. Liu, and X. Wang, "3D Human Mesh Regression With Dense Correspondence," 2020, pp. 7054–7063. Accessed: Apr. 15, 2022. [Online]. Available: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/html/Zeng\\_3D\\_Human\\_Mesh\\_Regression\\_With\\_Dense\\_Correspondence\\_CVPR\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPR_2020/html/Zeng_3D_Human_Mesh_Regression_With_Dense_Correspondence_CVPR_2020_paper.html)
- [8] J.-C. Yoo and T. H. Han, "Fast Normalized Cross-Correlation," *Circuits Syst. Signal Process.*, vol. 28, no. 6, p. 819, Aug. 2009, doi: 10.1007/s00034-009-9130-7.

- [9] P. Henderson, V. Tsiminaki, and C. H. Lampert, "Leveraging 2D Data to Learn Textured 3D Mesh Generation," 2020, pp. 7498–7507. Accessed: Apr. 15, 2022. [Online]. Available: [https://openaccess.thecvf.com/content\\_CVPR\\_2020/html/Henderson\\_Leveraging\\_2D\\_Data\\_to\\_Learn\\_Textured\\_3D\\_Mesh\\_Generation\\_CVPR\\_2020\\_paper.html](https://openaccess.thecvf.com/content_CVPR_2020/html/Henderson_Leveraging_2D_Data_to_Learn_Textured_3D_Mesh_Generation_CVPR_2020_paper.html)
- [10] K. Tóthová *et al.*, "Probabilistic 3D Surface Reconstruction from Sparse MRI Information," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham, 2020, pp. 813–823. doi: 10.1007/978-3-030-59710-8\_79.
- [11] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-Based Channel Estimation for mmWave Massive MIMO Systems," *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 5, pp. 989–1000, Sep. 2019, doi: 10.1109/JSTSP.2019.2925975.
- [12] B. M. S. Hasan and A. M. Abdulazeez, "A Review of Principal Component Analysis Algorithm for Dimensionality Reduction," *J. Soft Comput. Data Min.*, vol. 2, no. 1, Art. no. 1, Apr. 2021, Accessed: Apr. 15, 2022. [Online]. Available: <https://publisher.uthm.edu.my/ojs/index.php/jscdm/article/view/8032>
- [13] K. Tóthová *et al.*, "Uncertainty Quantification in CNN-Based Surface Prediction Using Shape Priors," in *Shape in Medical Imaging*, Cham, 2018, pp. 300–310. doi: 10.1007/978-3-030-04747-4\_28.
- [14] Y. Huang, Y. Zhang, and J. A. Chambers, "A Novel Kullback–Leibler Divergence Minimization-Based Adaptive Student's t-Filter," *IEEE Trans. Signal Process.*, vol. 67, no. 20, pp. 5417–5432, Oct. 2019, doi: 10.1109/TSP.2019.2939079.
- [15] M. Cali and R. Ambu, "Advanced 3D Photogrammetric Surface Reconstruction of Extensive Objects by UAV Camera Image Acquisition," *Sensors*, vol. 18, no. 9, Art. no. 9, Sep. 2018, doi: 10.3390/s18092815.
- [16] H. Fang and F. Lafarge, "Connect-and-Slice: An Hybrid Approach for Reconstructing 3D Objects," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, Jun. 2020, pp. 13487–13495. doi: 10.1109/CVPR42600.2020.01350.