



**Vilniaus
universitetas**

Ataskaitinė informatikos ir informatikos inžinerijos kryptių doktorantų ataskaitinė konferencija 2025-03-28

Kasparas Karlauskas, Informatikos inžinerija 2-as kursas

Preliminari tema:

Trimačių taškų debesų analizė ir objektų atpažinimas
(Three dimensional point cloud analysis and object detection)

Vadovas: prof. dr. Povilas Treigys

Studijų pradžios ir pabaigos metai: 2023–2027

Studijų metai: II (2024 / 2025)

Ataskaitinis laikotarpis: II studijų metų pavasario semestras

Studijų plano vykdymo suvestinė

Vilniaus
universitetas

Studijų metai	Egzaminai	
	Planas	Įvykdyta
I (2023/2024)	2	2
II (2024/2025)	2	2
III (2025/2026)		
IV (2026/2027)		
Iš viso:	4	4

Studijų plano vykdymo suvestinė

Vilniaus
universitetas

Studijų metai	Dalyvavimas konferencijose				Publikacijos					
	Tarptautinėse		Nacionalinėse		Su citavimo rodikliu			Be citavimo rodiklio		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė	Planas	Įvykdyta	Būklė
I (2023/2024)			1	1				1	0	
II (2024/2025)			1	1				1 + 1 (skola iš I metų)	2	Publikuota
III (2025/2026)	1				1					
IV (2026/2027)	1				1					
Iš viso:	2	0	2	0	2			2		

Ataskaitinio pusmečio darbų planas ir jų vykdymo suvestinė

Vilniaus
universitetas

Publikacijos 2024/2025 (II pusmetis)			
Planas	Įvykdyta	Būklė	Publikacijos tipas
SSTD '25: Proceedings of the 19th International Symposium on Spatial and Temporal Data	Karlauskas, K., Treigys, P. <i>Entropy- based Pre-filtering of High-density Forest Point Clouds for Individual Tree Detection Algorithms 2025</i>	Priimta: 2024-06-27	Be cituojamumo rodiklio
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing	Karlauskas, K., Gelšvartas, J., Treigys, P. <i>Point cloud data mining with HD map priors for making synthetic forest datasets</i> // IEEE journal of selected topics in applied earth observations and remote sensing. New York : Institute of Electrical and Electronics Engineers Inc. 2025, vol. 18, p. 19606-19617. DOI: 10.1109/JSTARS.2 025.3593827.	Publikuota: 2024-08-16	Turi cituojamumo rodiklį CA WoS duomenų bazėje

IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing

[Open Access](#)

5.3

Impact Factor

0.03176

Eigenfactor

1.19

Article Influence Score

9.3

CiteScore

Powered by Scopus®

19606

IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, VOL. 18, 2025

Point Cloud Data Mining With HD Map Priors for Making Synthetic Forest Datasets

Kasparas Karlauskas, Julius Gelšvartas, and Povilas Treigys

Abstract—Airborne laser scanning has proven to be an indispensable tool in surveying outdoor areas due to its efficient land coverage and unimpeded access to difficult-to-reach areas. The spatial information in 3-D point clouds, the data produced in airborne laser scanning surveys, can be leveraged to glean insights about the target area that would usually be unavailable in 2D-based remote sensing, such as satellite imaging. In forest point clouds, determining the locations and extents of individual trees, such as individual tree segmentation, may lend itself to forest inventory management and hazard prevention applications. However, substantial amounts of annotated and diverse data are necessary to develop classical and machine learning algorithms for individual tree segmentation. While 3-D point clouds are already notorious for how difficult it is to annotate them, high-altitude airborne laser scans make the task even more difficult due to their lower point density. This makes the structure of individual trees harder to discern, and prior information may be required to determine tree extents. This study proposes an automated approach in extracting individual trees from urban point clouds to construct synthetic datasets, combining point cloud clustering, geometrical features, and crowdsourced geospatial information of tree locations. The proposed method produces a new individual tree segmentation benchmark dataset for airborne laser scanning applications.

Index Terms—HD map, individual tree segmentation (ITS), LIDAR, synthetic data.

I. INTRODUCTION

AIRBORNE laser scanning (ALS) is a popular choice for surveys of outdoor areas, such as cities [1] or forests [2]. A particularly useful analysis enabled by the spatial information in forest point clouds is individual tree segmentation (ITS)—the task of determining each individual tree's location and spatial extent.

Various ITS methods for ALS point clouds have been developed for 2-D and 3-D domains of computer vision. 2-D raster-based approaches include growing regions on the full canopy height model (CHM) [9], [10] or analyzing the CHM's growth at several consequent height layers to determine the locations of potential treetops [11]. Algorithms that work on raw point clouds in three dimensions include rule-based region growing [12], height-adaptive density-based clustering [13], rotating the point cloud around potential treetops and analyzing height profiles to delineate the boundary between individual tree crowns [14], [15], and multistage filtering by intensity and derived geometric features to increase separation between tree instances [16]. Recent deep learning approaches tend to use 3-D voxel representations of forest point clouds [17], [18] to predict offset vectors for each point and then use classical clustering algorithms on the offset points to obtain tree instances. However, for the domain of low-density ALS point clouds, the performance of these algorithms is often not satisfactory as per the benchmarks carried out by the authors in [15], [16], [19], and [20].

The high difficulty of ITS in airborne scans, as opposed to higher density terrestrial laser scanning (TLS, [21]), could be attributed to the poor definition or complete absence of tree trunk points due to occlusion by the canopy. The characteristics of ALS forest point clouds pose significant challenges to producing data with individual tree instance labels, limiting its availability to researchers. Cao et al. [19] produced a dataset of two labeled forest plots by aligning low-density ALS and high-density labeled TLS point clouds. The two 1 ha plots containing ~500

Ataskaitinio pusmečio darbų planas ir jų vykdymo suvestinė

Vilniaus
universitetas

Dalyvavimas konferencijose 2024/2025 (II pusmetis)

Planas	Įvykdyta	Konferencijos tipas
19th International Symposium on Spatial and Temporal Data, 2025-08-26, Osaka, Japonija	Karlauskas, K., Treigys, P. <i>Entropy-based Pre-filtering of High-density Forest Point Clouds for Individual Tree Detection Algorithms</i> // 19th International Symposium on Spatial and Temporal Data, 2025-08-26, Osaka, Japonija	Tarptautinė

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

Darbo pavadinimas		Atlikimo terminai		Pastabos
		Nuo	Iki	
1.	Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):			
1.1.	Disertacijos tyrimo objekto konkretizavimas.	2023 m. IV ketvirtis	2024 m. I ketvirtis	Individualių medžių segmentavimas iš oro skanuotuose miškų trimačiuose taškų debesyse, naudojantis giliojo mokymosi metodais

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

1.2.	3D taškų debesų analizės metodų apžvalga. a) Giliojo mokymosi metodai b) Geometriniai metodai	2023 m. IV ketvirtis	2024 m. III ketvirtis	a) Apžvalga atlikta, savo eksperimentuose išmėgintas vienas segmentavimo neuroninis tinklas b) Aptikti perspektyvūs atviro kodo klasikiniai individualių medžių segmentavimo algoritmai, taip pat du straipsniai be kodo, bet detaliu algoritmo įgyvendinimo aprašymu
1.3.	Atliktos apžvalgos apibendrinimas ir pateikimas disertacijos analitinės dalies aprašyme.	2023 m. IV ketvirtis	2024 m. III ketvirtis	Algoritmų bei duomenų rinkinių apžvalga pateikta mokslinės ataskaitos skyriuje "2. Related Work"

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.	Mokslinio tyrimo vykdymas:			
2.1.	Tyrimo metodikos sudarymas:			
2.1.1.	Pagrindžiamas temos aktualumas, nurodoma darbo problema ir suformuluojamas tikslas.	2024 m. IV ketvirtis	2024 m. IV ketvirtis	Individualių medžių lokacijų nustatymas geriausiai veikia nuo žemės skenuotuose taškų debesyse, tačiau tai lėčiausias skenavimo būdas. Aktualu žinoti kokios segmentavimo algoritmų galimybių ribos didinant skenavimo aukštį (mažėja taškų tankis ir daugėja uždengtų skeneriui nepasiekiamų regionų).

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.1.2.	Atliekamas uždavinių pasirinkimas bei sprendžiamos problemos formulavimas.	2024 m. IV ketvirtis	2024 m. IV ketvirtis	Aukštos ir žemos raiškos taškų debesims: Segmentavimo algoritmų galimybių ribos idealiu atveju: koks tikslumas pasiekiamas naudojant anotuotus (tikrus arba sintetinius) duomenis? Neuroninio tinklo įsisotinimui reikalingas duomenų kiekis: palyginus su dvimatės kompiuterinės regos literatūra, trimačių taškų debesų segmentavimo mokymo aibės turi daug mažiau unikalių duomenų. Apžvelgti koks miško plotas naudojamas literatūroje, nustatyti ar neuroninis tinklas gali įsisotinti su tokio dydžio mokymosi aibe. Triukšmingų anotacijų taikymas: palyginus su neuroniniu tinklu mokytu naudojantis švariai anotuotus duomenis, kokį tikslumą galima pasiekti naudojant triukšmingas anotacijas gautas egzistuojančiais algoritmais?
--------	--	----------------------	----------------------	--

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.1.3.	Taikomos metodologijos ar tyrimo metodo pateikimas ir aprašymas	2024 m. IV ketvirtis	2024 m. IV ketvirtis	Neuroninių tinklų tikslumo palyginimo statistiniam reikšmingumui bus naudojamas Friedman testas. "Stebėjimų" aibės dydis bus užtikrinamas naudojant kelias atsitiktinių skaičių sėklas, ir kelias nepriklausomas testavimo duomenų aibes. T.y. rangai nustatomi kiekvienai sėklos ir duomenų rinkinio kombinacijai.
--------	---	----------------------	----------------------	---

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.2.	Teorinis tyrimas:			
2.2.1.	Mokslinės ir kitos informacijos analizė ir sisteminimas.	2024 m. IV ketvirtis	2025 m. III ketvirtis	Viešai prieinamų anotuo-tų (arba anotavimui palankių) duomenų rinkinių apibendrinimas;
2.2.2.	Objektų detekcijos 3D taškų debesyse metodai (geometriniai, giliojo mokymosi).	2024 m. IV ketvirtis	2025 m. III ketvirtis	Atlikta klasikinių ir giliojo mokymosi algoritmų apžvalga pateikta straipsnio „Point Cloud Data Mining with HD Map Priors for making Synthetic Forest Datasets“ įžangoje

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.2.3.	3D taškų debesyse aptiktų objektų parametrų nustatymas.↵	2024 m. IV ketvirtis↵	2025 m. III ketvirtis↵	Medžio kamieno savybių nustatymas triukšmui
				atspariais cilindru/apskritimų tinkinimo metodais:↵ 1. RANSAC (RANdom SAmple Consensus)↵ 2. RHT (Randomized Hough Transform)↵ 3. RLTS (Robust Least Truncated Squares)↵

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.3.	Empirinis tyrimas:			
2.3.1.	Eksperimentinės dalies aprašas.	2025 m. I ketvirtis	2026 m. II ketvirtis	
2.3.2.	Egzistuojančio metodo realizavimas kuriamo metodo palyginimui.	2025 m. I ketvirtis	2026 m. II ketvirtis	
2.3.3.	Metodų modifikacijos sukūrimas.	2025 m. I ketvirtis	2026 m. II ketvirtis	
2.3.4.	Sukurtų modifikacijų eksperimentinis tyrimas ir modifikacijų palyginimas su publikuotais egzistuojančiais metodais.	2025 m. I ketvirtis	2026 m. II ketvirtis	

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

2.4.	Gautų duomenų analizė, apibendrinimas, išvadų parengimas:			
2.4.1.	Teorinio tyrimo apibendrinimas.	2025 m. III ketvirtis	2026 m. II ketvirtis	
2.4.2.	Empirinio tyrimo apibendrinimas.	2025 m. III ketvirtis	2026 m. II ketvirtis	
2.4.3.	Rezultatų apibendrinimas, išvadų parengimas.	2025 m. III ketvirtis	2026 m. II ketvirtis	

Visų mokslinių tyrimų ir disertacijos rengimo etapai

Vilniaus
universitetas

3.	Atskirų daktaro disertacijos dalių (tyrimo metodikos, rezultatų, ginamų teiginių, išvadų, ir kt.) parengimas:			
3.1.	Tikslų, uždavinių, tyrimo metodikos, ginamųjų teiginių patikslinimas.	2026 m. II ketvirtis	2027 m. I ketvirtis	
3.2.	Analitinės disertacijos dalies parengimas.	2026 m. II ketvirtis	2027 m. I ketvirtis	
3.3.	Teorinės disertacijos dalies parengimas.	2026 m. II ketvirtis	2027 m. I ketvirtis	
3.4.	Eksperimentinės disertacijos dalies parengimas.	2026 m. II ketvirtis	2027 m. I ketvirtis	
3.5.	Bendrųjų išvadų formulavimas.	2026 m. II ketvirtis	2027 m. I ketvirtis	
4.	Daktaro disertacijos <u>parengimas</u> ir svarstymas padalinyje:	2027 m. II ketvirtis	2027 m. II ketvirtis	
5.	Daktaro disertacijos gynimas:	2027 m. III ketvirtis	2027 m. III ketvirtis	

Tyrimo objektas

Vilniaus
universitetas

- Trimačiai taškų debesys, užfiksuoti naudojantis lazerinio skenavimo aparatais, pritvirtintais prie dronų, leidžia atlikti sudėtingesnę analizę negu dvimačiai vaizdai. Miškų taškų debesys naudojami miškininkijoje (medžių inventorizacija) bei infrastruktūroje (elektros perdavimo laidų saugumas).
- Individualių medžių segmentacija žemos skiriamosios gebos lazerinio skenavimo 3D taškų debesyse
 - Kol kas vyrauja geometriniai algoritmai, kurių veikimas tankiose medžių zonose ne idealus, sudėtinga atitaikyti algoritmo parametrus (per smulkus/stambus segmentavimas);
 - Literatūroje pateikiami mašininio mokymosi metodai orientuoti į antžeminius lazerinio skenavimo duomenis, arba dronu užfiksuotus aukštos skiriamosios gebos duomenis (žemas skrydis);
 - Sužymėti duomenys individualių medžių segmentacijai reti, gavimo procesas sudėtingas.

Moksliniai rezultatai

Medžių kamienų taškų išskyrimas aukštos raiškos taškų debesyse

- Naujas kamienų aptikimo ir segmentavimo algoritmas, paremtas taškų pasiskirstymo entropija
 - Preliminarūs rezultatai pristatyti konferencijoje
 - Ruošiama publikacija pilnam išplėtam eksperimentui
- Įgyvendinti du algoritmai iš literatūros (greičio ir tikslumo palyginimui su naujuoju algoritmu)

Medžių kamienų taškų išskyrimas aukštos raiškos taškų debesyse

Entropija 1 x 1 x 1 m kubuose (vokseliuose)

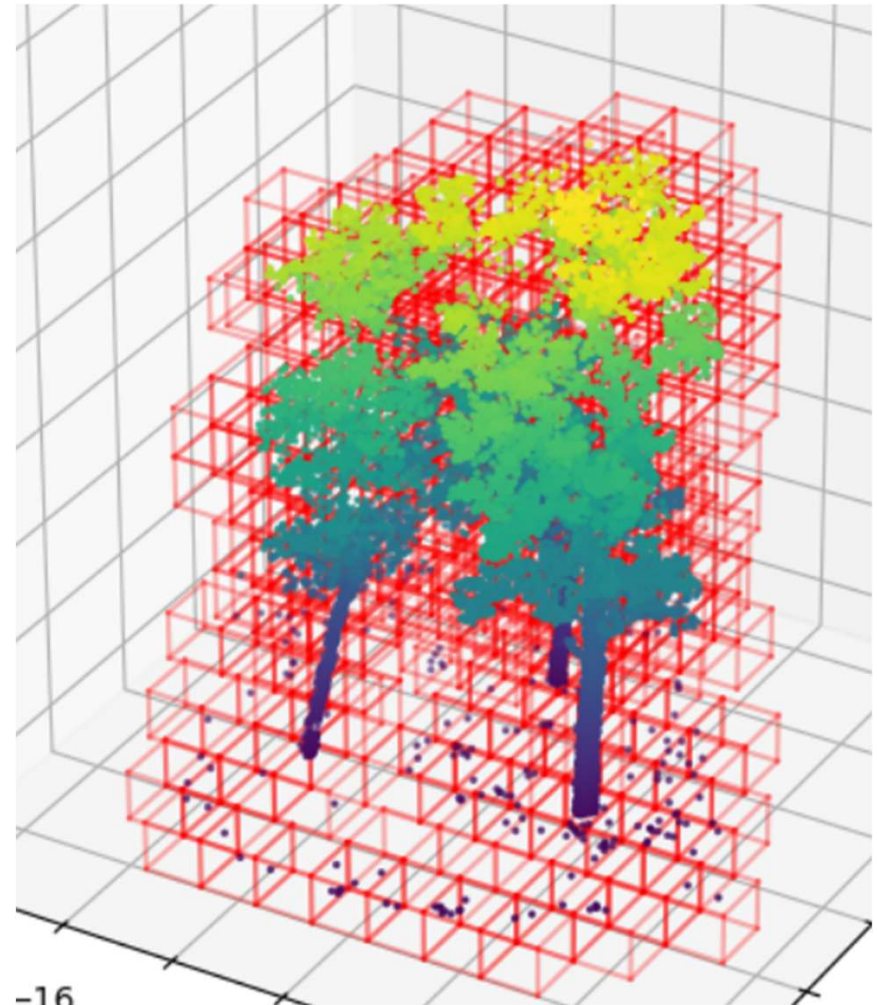
Vilniaus
universitetas

1. Suskaidyti taškų debesį į kubus
2. Kiekvieno kubo viduje suskaičiuoti entropiją:

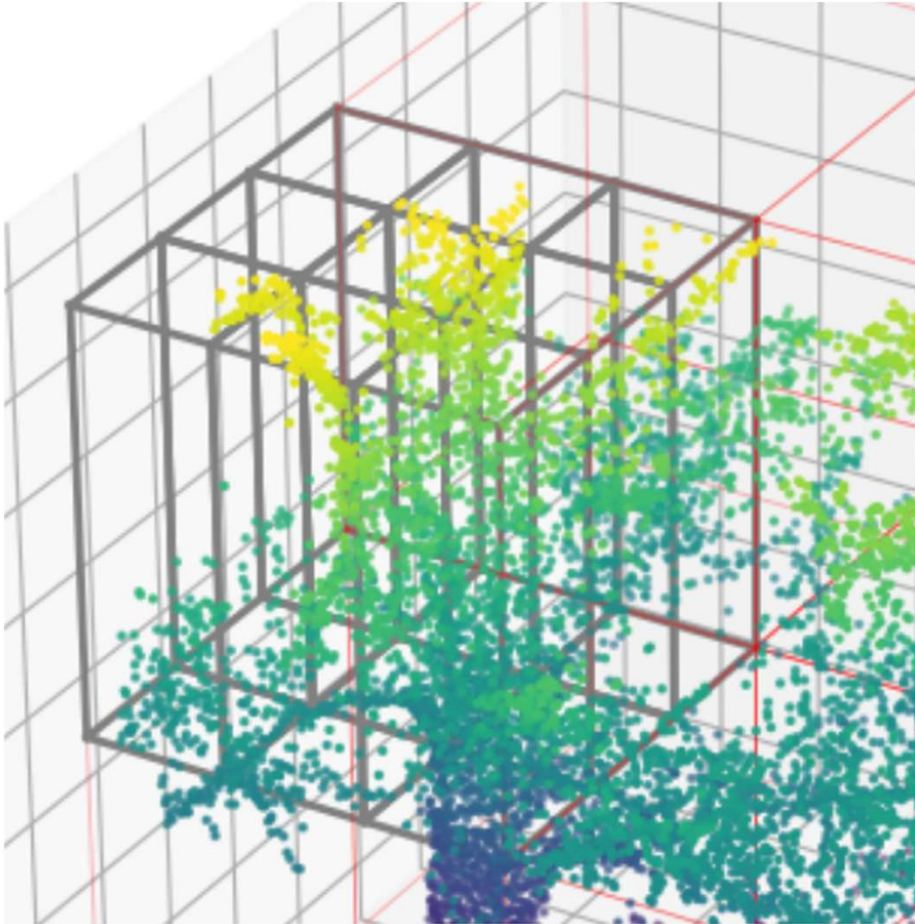
$$H(X) := - \sum_{x \in X} p(x) \log p(x)$$

$$p(S_{i'j'k'}^{ijk}) := \frac{|S_{i'j'k'}^{ijk}|}{|V_{ijk}|}$$

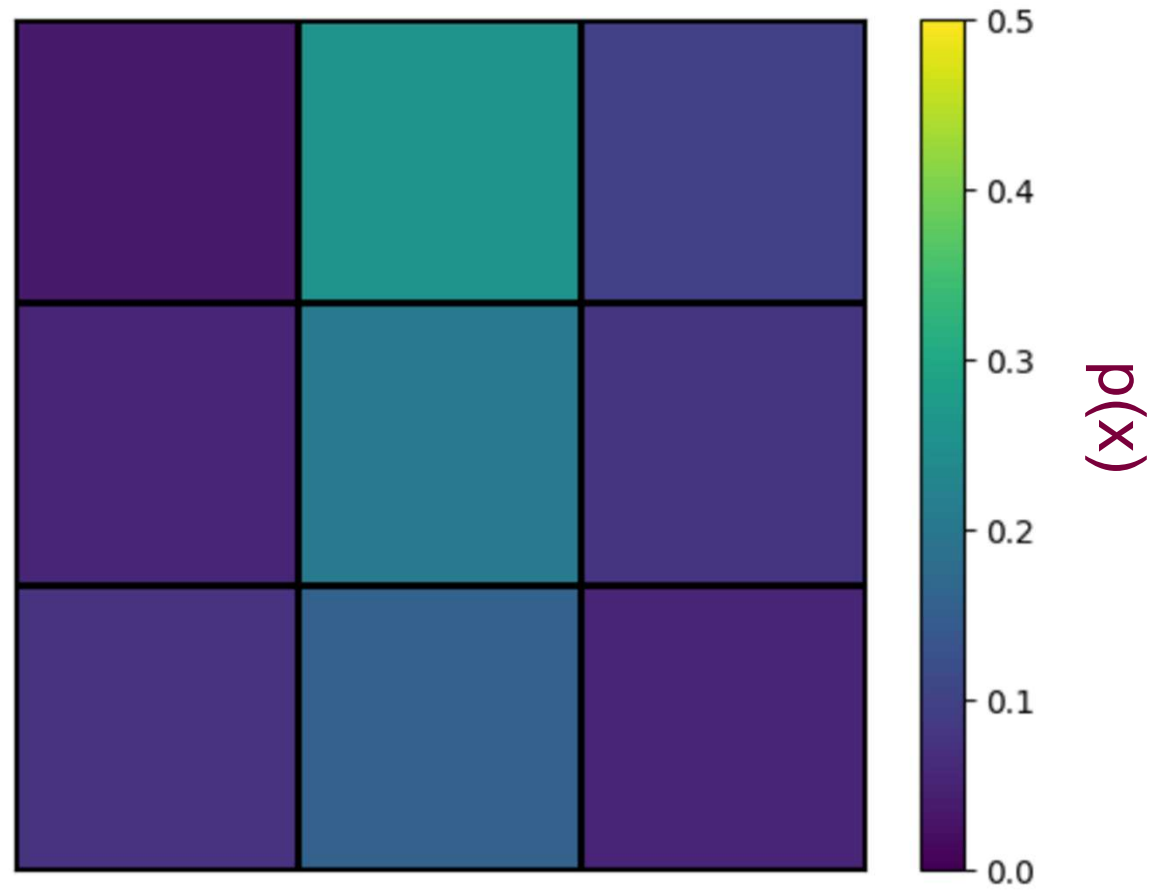
3. Kiekvienam taškui priskirti jo vokselio entropijos vertę



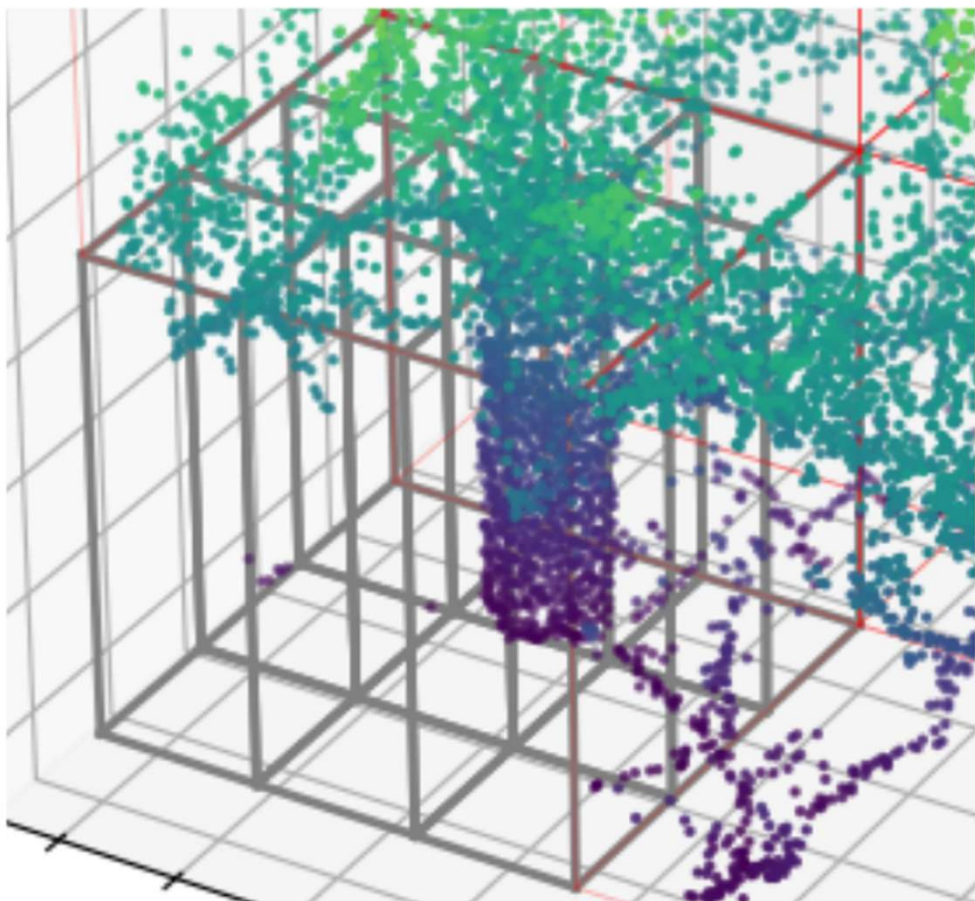
$$H = 0.91$$



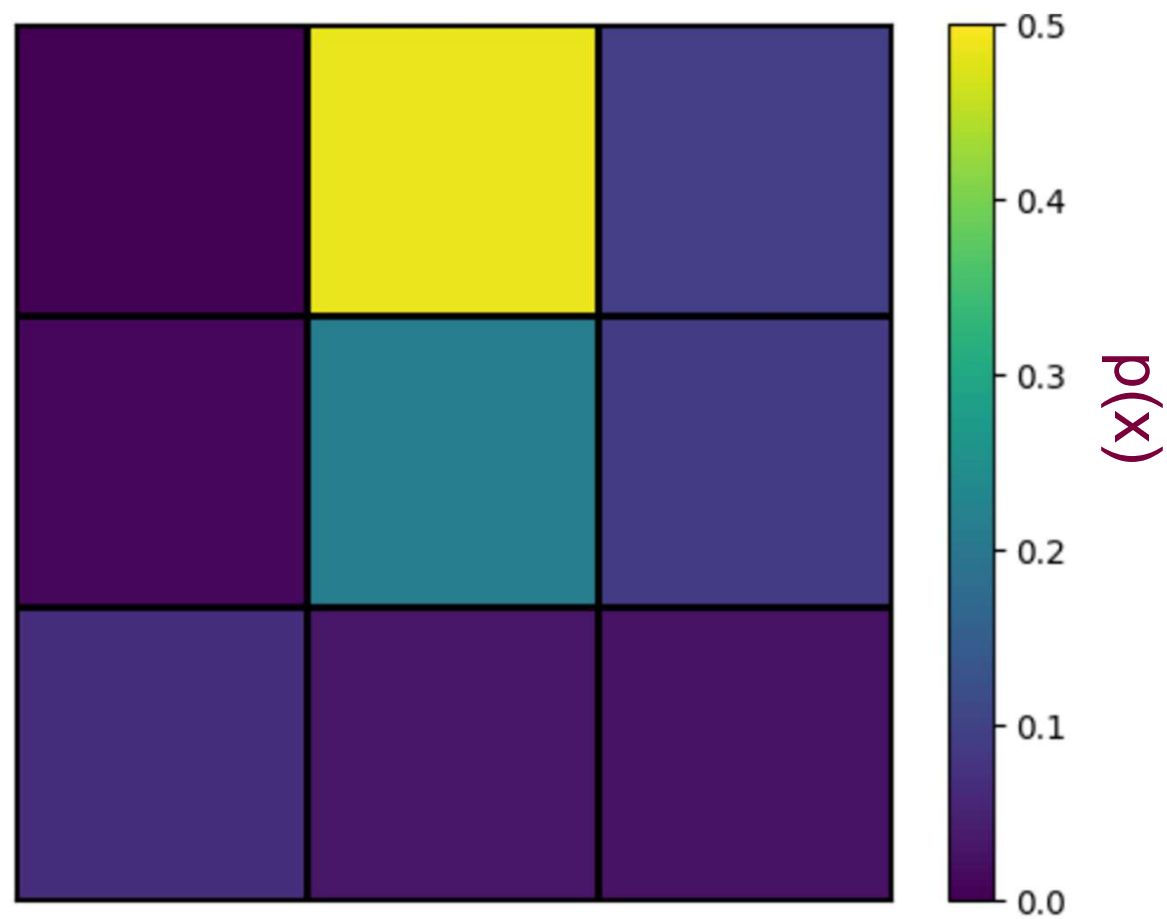
**Vilniaus
universitetas**



$$H = 0.70$$

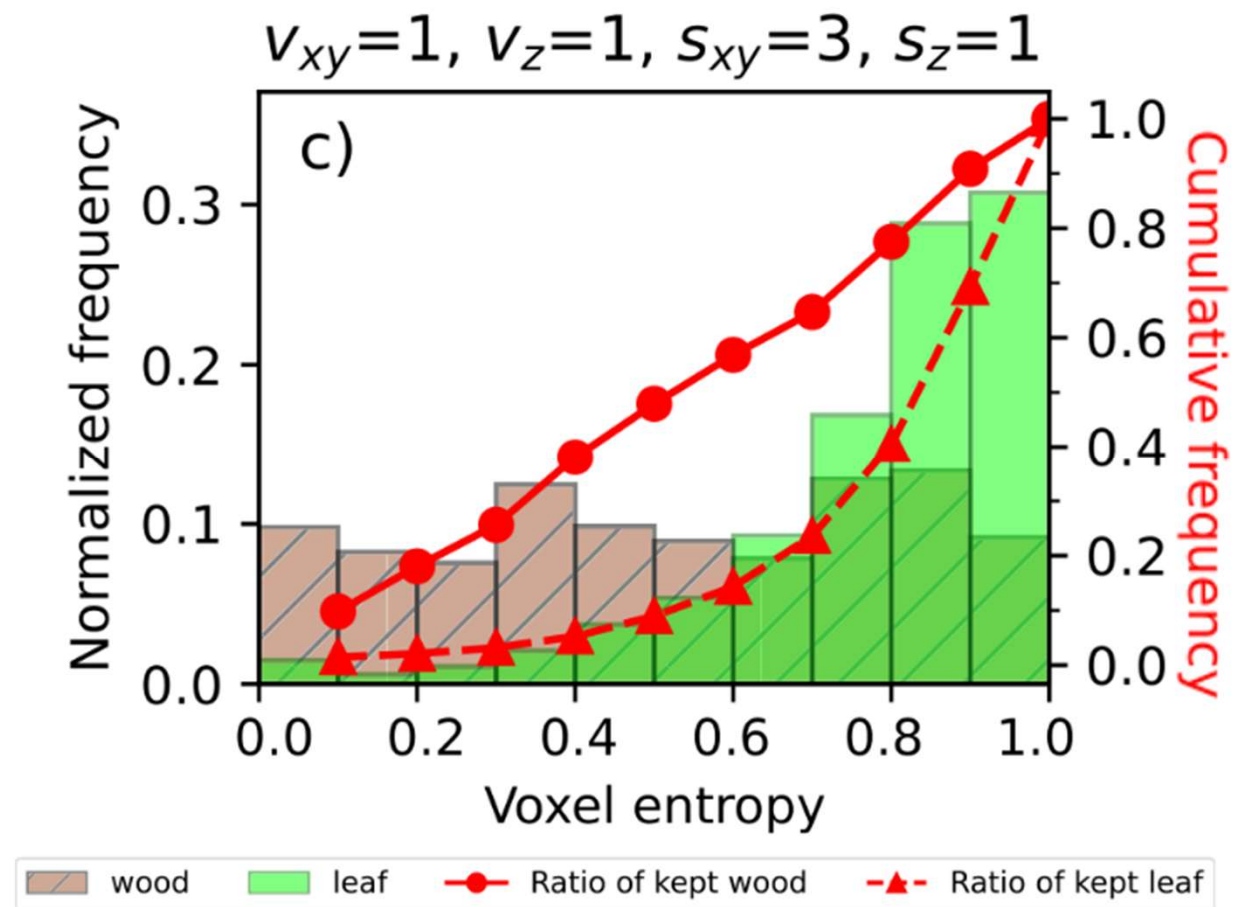


Vilniaus
universitetas



Konferencijoje pristatyti rezultatai: Pradinis taškų filtravimas esamiems algoritmams

- Pašalinus visus taškus su $H > 0.8$, virš 80 % kamieno taškų lieka, tuo tarpu iki pusės lapų taškų pašalinami
- 30 % ~ 60 % pagreitėjimas kamienų segmentavimo algoritmui, kuris pradžioj kiekvienam taškui skaičiuoja vietinį kreivumą su kd-tree užklausomis
- Be papildomų modifikacijų etaloniniam algoritmui, tikslumas sumažėja, tačiau ne smarkiai



Konferencijoje
pristatyti rezultatai:
Pradinis taškų
filtravimas esamiems
algoritmams

Table 3: Individual tree detection accuracy metrics for the study plots

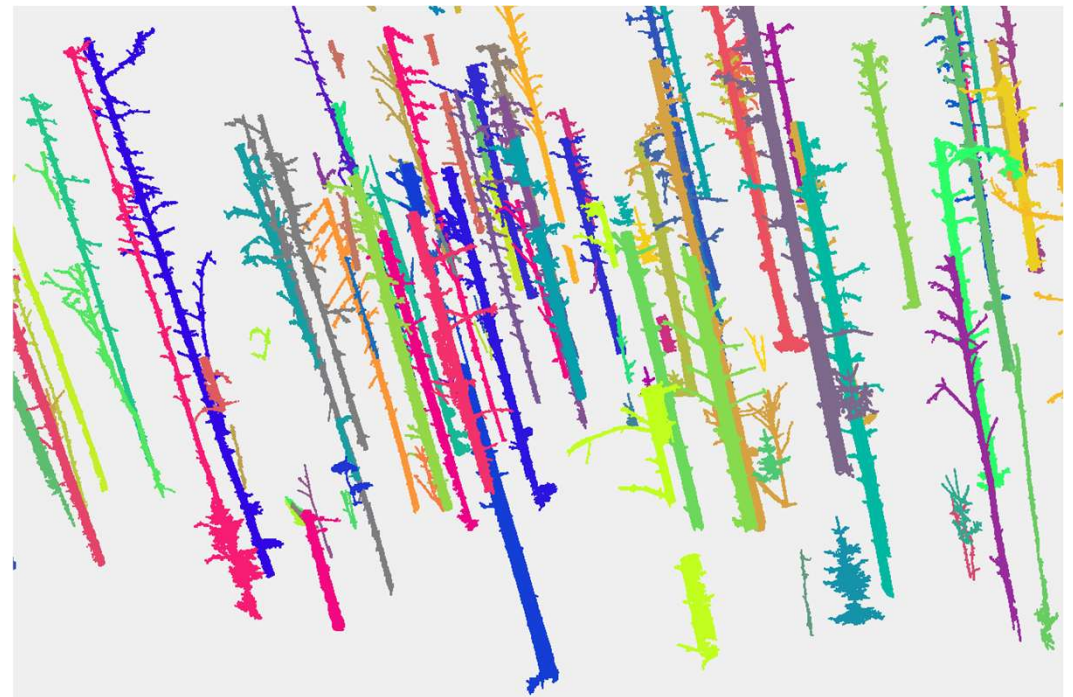
Plot #	Feature ^a	time ^b	TP	FP	FN	prec	rec	F1
1	NCR	84	46	6	5	0.88	0.90	0.89
	VE + NCR	<i>1 + 63</i>	44	9	7	0.83	0.86	0.85
2	NCR	57	67	26	17	0.72	0.80	0.76
	VE + NCR	<i>1 + 40</i>	67	23	17	0.74	0.80	0.77
3	NCR	180	87	10	61	0.90	0.59	0.71
	VE + NCR	<i>2 + 108</i>	87	9	61	0.91	0.59	0.71
4	NCR	187	46	17	32	0.73	0.59	0.65
	VE + NCR	<i>2 + 128</i>	43	9	35	0.83	0.55	0.66
5	NCR	231	64	40	67	0.62	0.49	0.54
	VE + NCR	<i>3 + 133</i>	59	30	72	0.66	0.45	0.54
6	NCR	169	78	8	158	0.91	0.33	0.48
	VE + NCR	<i>2 + 124</i>	65	7	171	0.90	0.28	0.42

a Features used for point filtering before the clustering step: NCR = Normal Change Rate, VE = Voxel Entropy.

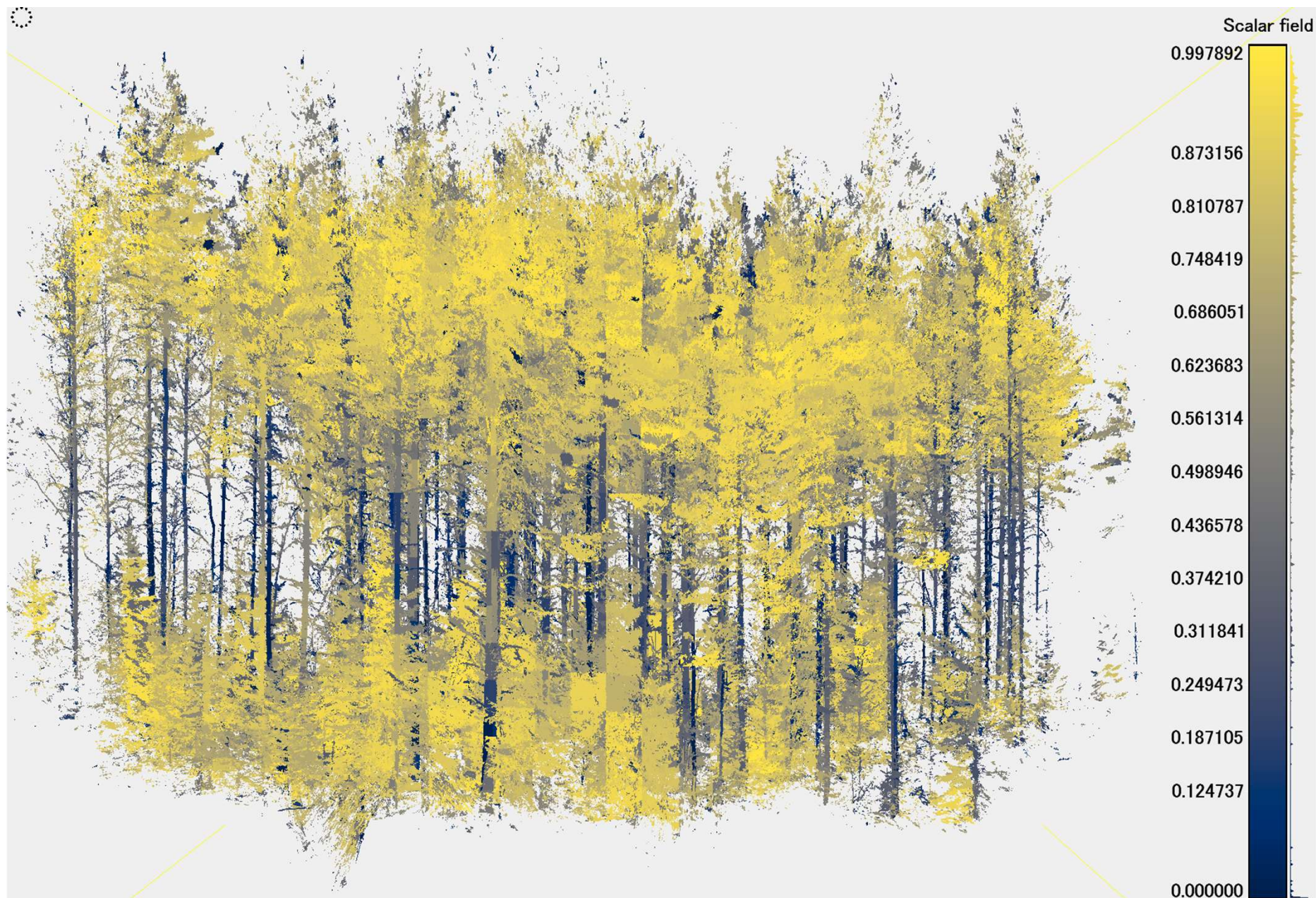
b Time taken to compute individual point features in minutes. If several features are computed, the times are given in a sum in their respective order in the Feature column. Plaintext times are for multiprocessing using 16 Intel® Xeon® Gold 6252 Processor CPU cores, while *italicized* times are computed using a single CPU core.

Metodo tęsinys: atskiras kamienų aptikimo algoritmas paremtas „vokselių entropija“

1. Pašalinti visus taškus su $H < 0.5$
2. Segmentavimas - sujungtų komponentų segmentacija (connected component segmentation)
3. Per smulkiai suskaidytų kamienų sujungimas



Suskaičiuoti
vokselių
entropiją taškų
debesyje



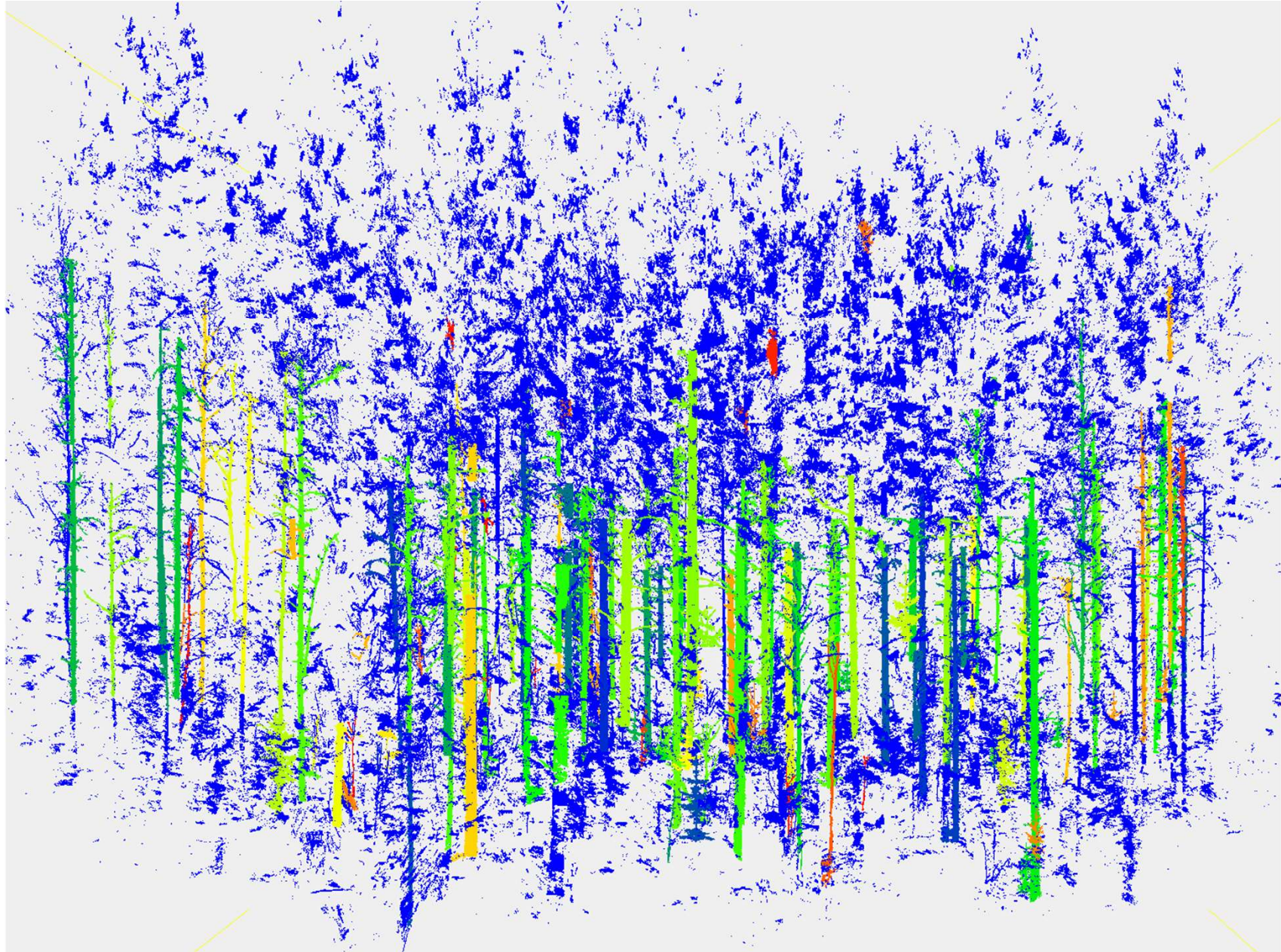
Pašalinti
visus taškus
su $H < 0.5$



Rasti gretimų
taškų grupes
(galimi
kamieniai)

Klasteriai
pašalinami
jeigu yra per
trumpi, arba
nepakankamai
vertikalūs

Gretimi xy
plokštumoje
klasteriai
sujungiami



Medžių lokacijų nustatymui,
tai gali būti patenkinamas
rezultatas.

Segmentacijai priklausomai
nuo kamienus supančios
augmenijos tankio, gali būti
aptikta tik maža viso
kamieno dalis

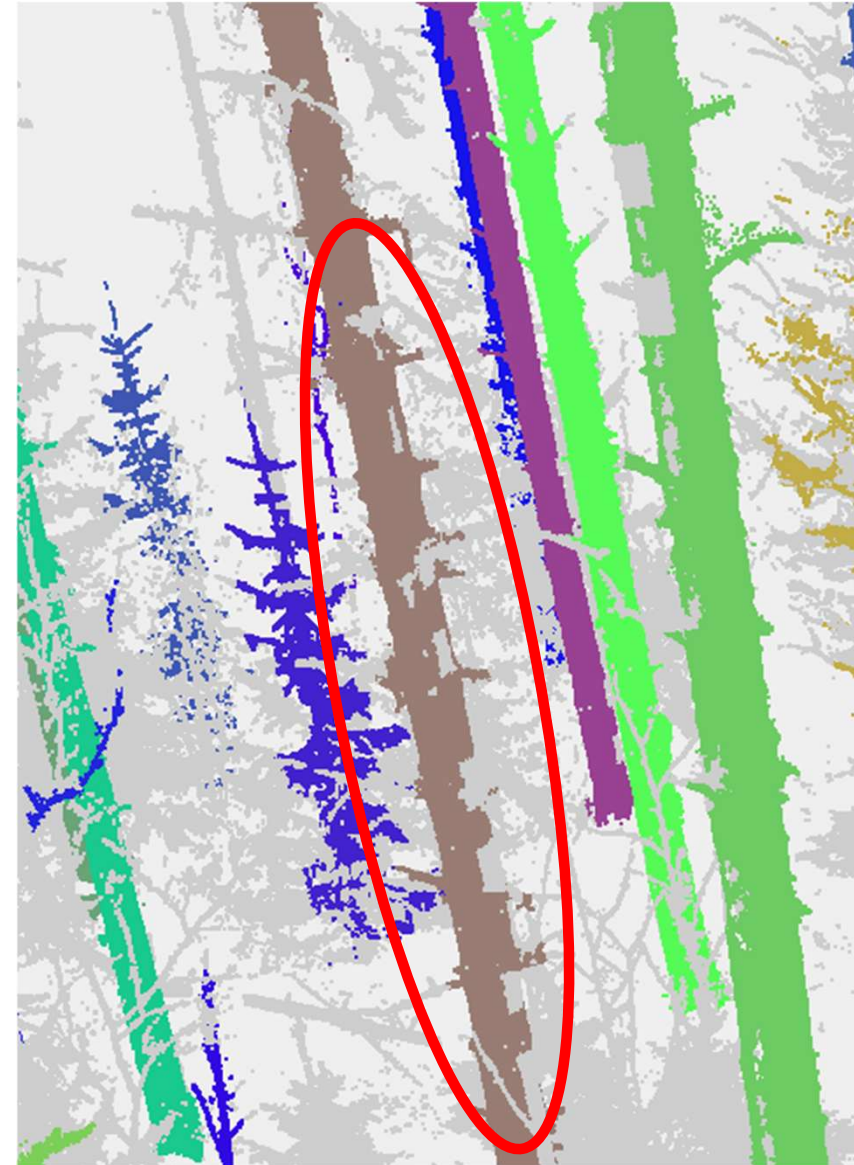


Kamieno taškų atstatymas naudojantis cilindro tinkinimą

Nustatyti medžio kamieną aproksimuojančio
cilindro parametrus iš aptiktos kamieno dalies

Populiariausi metodai reikalauja žinoti taškų
paviršiaus liestinei statmeną kryptį (surface
normal) – RANSAC metodas atsitiktinai traukiant
du taškus.

Kolkas bandau Nurunnabi ir kt. (2017) algoritimą
kuris naudoja tik taškų koordinates su triukšmui
atspariais PCA metodais. Greičio ir tikslumo
palyginimai ateinantį semestrą.

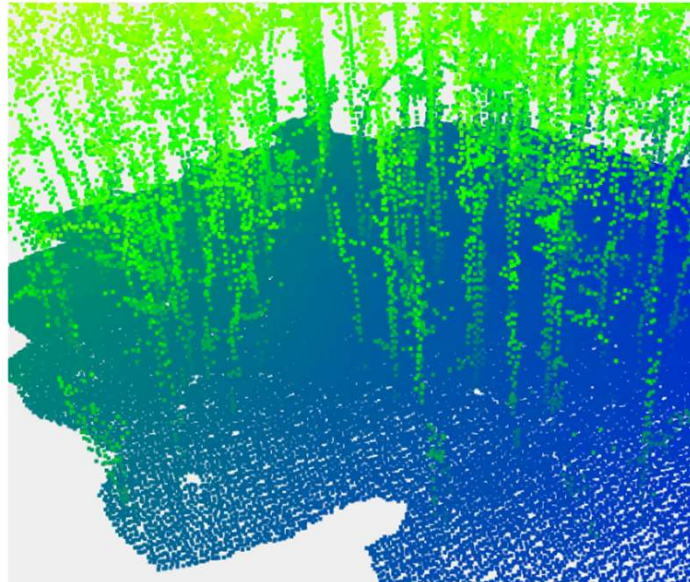


Sąsaja su pagrindine tema – žemos skyros miškų taškų debesimis

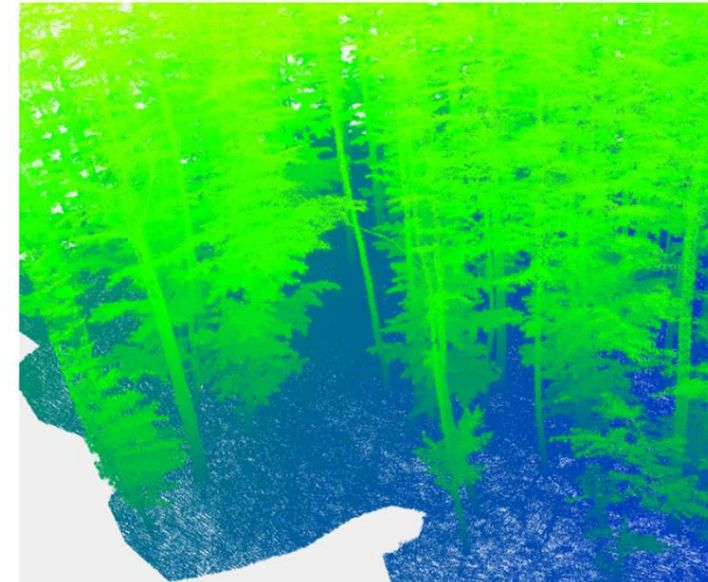
Vilniaus
universitetas

Našiose apžvalgose iš oro (*Airborne Laser Scanning, ALS*) užfiksuoti miškų taškų debesys neretai turi prastai išskirtus kamienus, arba jų visai neturi

Tiksliai anotuoti be išorinių duomenų praktiškai neįmanoma



ALS



TLS

Data from SilviLaser 2021

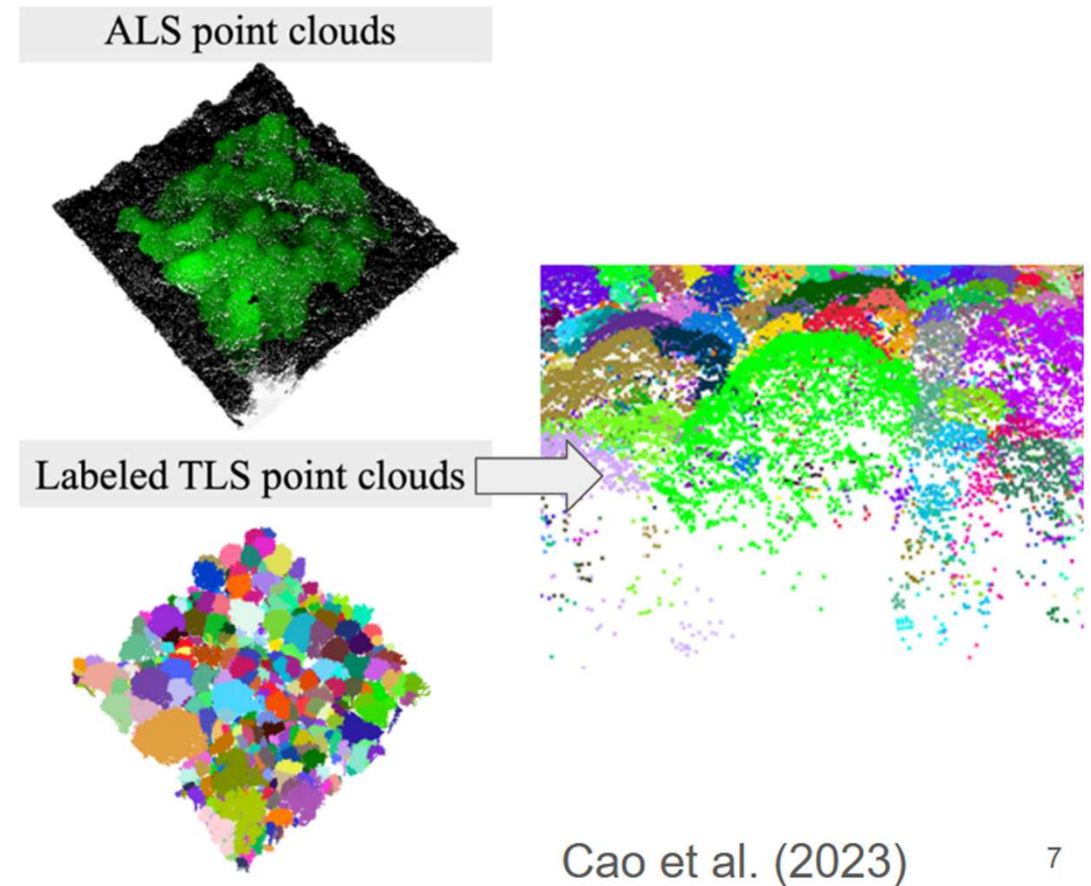
Sąsaja su pagrindine tema – žemos skyros miškų taškų debesimis

Vilniaus
universitetas

Anotacijų perkėlimas turint tos pačios vietos TLS ir ALS taškų debesis.

Medžių lokacijos – TLS kamienai, perkelti lokaciją į aukščiausią taršką aplink kamieną (individualaus medžio viršūnė)

Medžių segmentacija – susegmentuoti TLS, naudoti *k artimiausių kaimynų* klasifikavimą perkelti antoacijas į ALS



Cao et al. (2023)

Kito pusmečio darbo planas

1. Pabaigti ruošti publikaciją praeito semestro metu išvystytam segmentavimo algoritmui, pateikti WoS indeksuotam žurnalui
2. Paruošti medžių lokacijų aptikimo duomenų rinkinį mašininio mokymosi tyrimams – aptikti kamienus aukštos raiškos duomenyse, naudoti lokacijų informaciją tos pačios vietovės žemos skyros taškų debesyse
3. Atlikti mašininio mokymosi eksperimentą medžių lokacijų aptikimui
 1. Preliminarus tyrimas lyginant 2D rasterio ir 3D taškų debesų naudojimą lokacijų aptikimui, pristatyti konferencijoje
 2. Praplėstas tyrimas su didesniu duomenų rinkiniu: palyginti tikslumą naudojant sintetinius žemos skyros duomenis (i) sumažinto tankio originalūs aukštos skyros duomenys; ii) lazerinio skenavimo simulatorius su originaliais aukštos skyros duomenimis)