



**Vilnius
universitetas**

**Vilnius
universitetas**

**Algoritmų, leidžiančių
nustatyti medžiagos
cheminę sudėtį, iš hiper-
spektrinių duomenų tyrimas**

Vytautas Paura

Darbo vadovas: Dr. Virginijus Marcinkevičius

Darbo konsultantas: Dr. Valdas Rapševičius

Doktorantūros pradžios ir pabaigos metai: 2020–2024

Studijų planas

Studijų metai	Egzaminai		Dalyvavimas konferencijose		Publikacijos		
	Planas	Įvykdyta	Planas	Įvykdyta	Planas	Įvykdyta	Būklė
I (2020/2021)	2	2	1	1	1	1 (konferencijoje)	
II (2021/2022)	2	1	1	1	1	1 (konferencijoje)	
III (2022/2023)			1	0	1	0	
IV (2023/2024)			1	0	1	0	

Pusmečio ataskaita

Egzaminai		
Planas	Įvykdyta	Būklė
Fundamentalieji informatikos ir informatikos inžinerijos metodai	2022-01-28	Egzaminas <u>išlaikytas</u> .

Dalyvavimas konferencijose		
Planas	Įvykdyta	Konferencijos tipas
DAMSS, 2021-12-01, Druskininkai	V. Paura, V Marcinkevičius. “ <u>The Analysis of Impact of Noise on Hyperspectral Unmixing Algorithms</u> ”. 2021-12-02	Konferencija <u>nacionalinė</u>

Disertacijos rengimo etapai

Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):

1. Disertacijos tyrimo objekto detalizavimas.
2. Medžiagų cheminės sudėties nustatymo algoritmų apžvalga ir analizė.
3. Algoritmų naudojančių hiper-spektrinius duomenis medžiagų cheminei sudėčiai nustatyti apžvalga ir analizė.
4. Metodų analizės apibendrinimas ir mokslinių problemų susietų su tyrimo objektu identifikavimas, ir tyrimo tikslo ir uždavinių suformavimas.

Tyrimo objektas

Tyrimo objektas:

- Hiper-spektriniai vaizdai ir medžiagų cheminės sudėties iš hiper-spektrinių vaizdų nustatymo algoritmai.

Sprendžiamos problemos:

- Grynujų pikselių skaičius nustatytas iš pateikto hiper-spektrinio vaizdo ar duomenų rinkinio.
- Randant grynujų medžiagų spektrines žymes (angl. signatures).
- Suskaičiuojant atrastų grynujų medžiagų sumaišymo proporcijas visame hiper-spektriniame vaizde.

2021/2022 m. m. darbo planas

- Mokslinių tyrimų disertacijos tema apžvalga ir analizė (Lietuvoje ir užsienyje):
 - Medžiagų cheminės sudėties iš hiper-spektrinių duomenų algoritmų standartizuoto eksperimento metodikos kūrimas ir įgyvendinimas.
 - Medžiagų cheminės sudėties nustatymo algoritmų eksperimentinis tyrimas.
 - Algoritmų naudojančių hiper-spektrinius duomenis medžiagų cheminei sudėčiai nustatyti apžvalga ir analizė.
 - Problemų, kylančių iš tikslo, suformulavimas būsimiems eksperimentiniams ir analitiniams tyrimams.
- Išlaikyti egzaminus:
 - “Fundamentalieji informatikos ir informatikos inžinerijos metodai”
 - “Netiesiniai statistikos modeliai masinių duomenų analizėje”
- Publikacijų rengimas
 - Mokslinių tyrimų ir etaloninio eksperimento disertacijos tema apžvalga (publikavimas žurnale).

Cheminės sudėties nustatymas

Sudėties nustatymo algoritmų tipai:

1. Neneigiamų matricų faktorizavimas.
2. Autoenkoderių tinklai.
3. Retų duomenų regresija.

Metrikos:

1. Vidutinės kvadratinės paklaidos šaknis (Root mean squared error - RMSE)
2. S rekonstravimo paklaida (Signal reconstruction error SRE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2}$$

$$SRE = 10 \log_{10} \left(\frac{E[\|x\|_2^2]}{E[\|x - \hat{x}\|_2^2]} \right)$$

Cheminės sudėties nustatymo eksperimentas

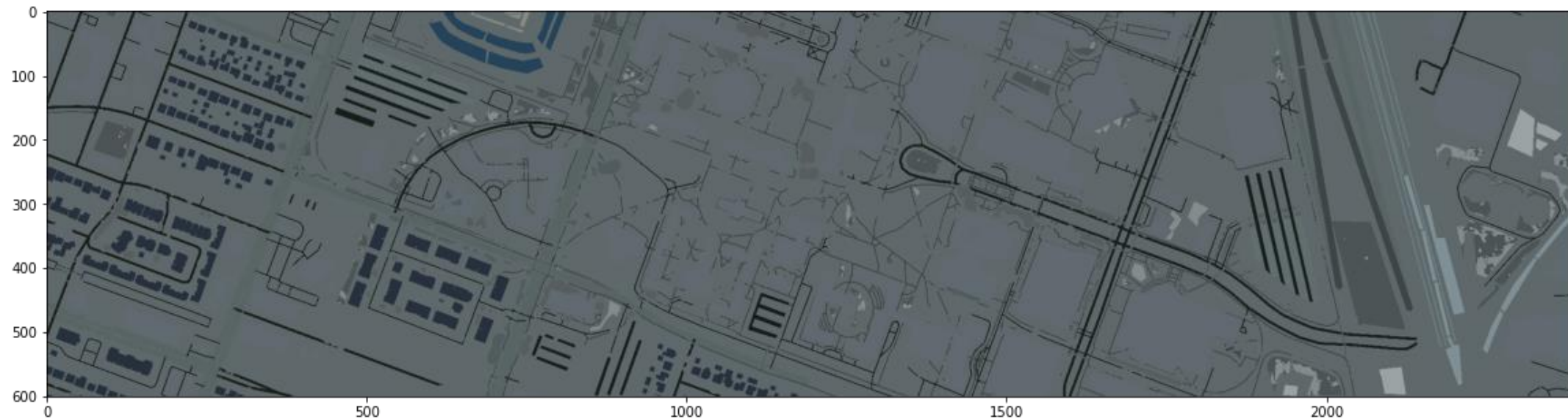
Duomenų rinkiniai:

- Dirbtiniai duomenų kubai sudaryti iš USGS spektrinės bibliotekos duomenų (Raymond F. Kokaly et al., 2017)
- “2018 IEEE GRSS data fusion” konkurso klasifikuotas miesto vaizdo paviršiaus elementų fragmentas.

Etaloninis eksperimentas:

1. Algoritmų atsparumas medžiagų kiekio pasikeitimams.
2. Algoritmų atsparumas triukšmui.
3. Algoritmų gebėjimas atskirti medžiagas priklausomai nuo hiper-spektrinio vaizdo dydžio.

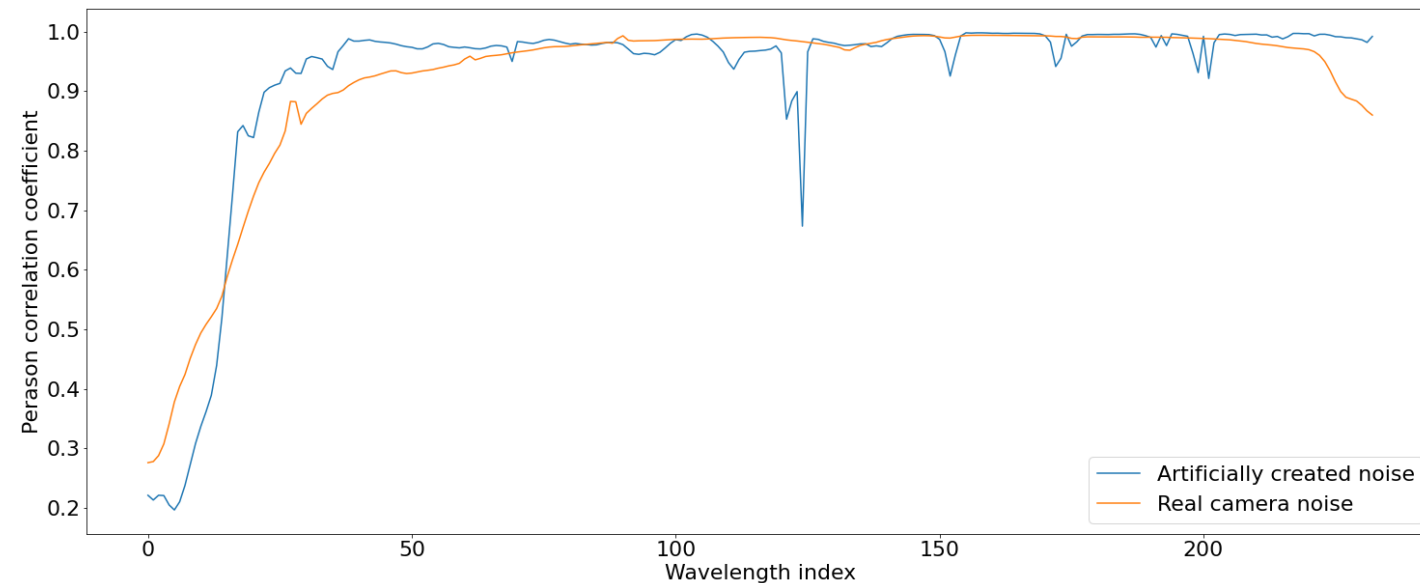
Cheminės sudėties nustatymo eksperimentas



“2018 IEEE GRSS data fusion” duomenų rinkinio fragmentas naudotas, kaip hiperspektrinių vaizdų generavimo pagrindas.

Algoritmų triukšmo atsparumo eksperimentas

- Sugeneruotas dirbtinis Gauso triukšmas skirtingais SNR kiekiais.
- Panaudoti realios kameros triukšmo parametrai.
- Skaičiuojama Pearson koreliacija tarp gretimų ruožų triukšmo kiekiui nustatyti.
- Atliekama optimizacija dirbtiniam triukšmui generuoti.



Algoritmai naudoti eksperimente

Retų duomenų regresija:

- Sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) [J. M. Bioucas-Dias and M. A. T. Figueiredo, 2010]
- SUnSAL and total variation (SUnSAL_TV) [J. M. Bioucas-Dias and M. A. T. Figueiredo, 2012]
- Spectral–spatial weighted sparse unmixing (S2WSU) [S. Zhang et al. 2018]
- Augmented linear mixing model (ALMM) [D. Hong et al., 2019]

Algoritmai naudoti eksperimente

Neneigiamų matricų faktorizavimas:

- Spatial group sparsity regularized NMF (SGSNMF) [X. Wang et al., 2017]
- Coupled nonnegative matrix factorization (CNMF) [N. Yokoya et al., 2012]
- Robust collaborative nonnegative matrix factorization (R-CoNMF) [J. Li et al., 2016]
- Total variation regularized reweighted sparse NMF (RSNMF) [W. He et al., 2017]

Autoenkoderių tinklai:

- Autoencoder network with adaptive abundance smoothing (AAS) [Z. Hua et al., 2020]

Išvados

- Tarp tirtų algoritmų nebuvo sukurta vieninga metodika šių algoritmų testavimui.
- SGSNMF algoritmas gavo prasčiausius rezultatus.
- SUnSAL algoritmas beveik visų atliktų eksperimentų metu gaudavo geriausius rezultatus.

Kito pusmečio planas

1. Kurso "Netiesiniai statistikos modeliai masinių duomenų analizėje" laikymas
2. Etaloninių eksperimentų ir algoritmų analizės rezultatų publikavimas ISI žurnale.
3. Hiper-spektrinių duomenų rinkinio sudarymas, tirtų algoritmų patikrinimas su sukurtu duomenų rinkiniu, rezultatų surinkimas į straipsnį.



**Vilnius
universitetas**

Kontaktai

Akademijos g. 4

LT-08663 Vilnius

+370 6 256 79560

vytautas.paura@mif.stud.vu.lt