Paper Outline

Introduction

1. Characterization of Inland Water Bodies via Remote Sensing is Challenging
   - complex spectral signatures
   - limited spatial and temporal resolution (long satellite overpasses)
   - nevertheless remote sensing has significant potential
   - Harmful Algal Bloom (HAB) classification and mapping
   - Oil Spill mapping
2. Hyperspectral Imaging addresses spectral limitations of traditional remote sensing platforms
3. Supervised methods
   - Inversion of optically-active parameters
   - Key limitation: availability of in-situ reference data
     - Cite Aurin et al. and Ross et al. here
   - When ground truth data are limited, classification is an alternative approach, e.g. "bad" to "good" water quality
4. Unsupervised methods
   - Unsupervised classification & ML Techniques
   - Endmember Extraction and Spectral Unmixing
5. Robot Teams and HSI equipped Drones
6. This paper
   - Extend capabilities of robot team by showcasing capabilities for unsupervised classification
   - Utilize GTM as Bayesian extension of SOM to enable unsupervised classification
   - GTM also enables extraction of spectral features associated with each node

- Two case studies
  1. GTM trained on water-only HSI data for lake
     - Identify interesting regions for further examination by robot team
     - Investigate spectral variability of water
  2. GTM on dataset combining grass and water pixels with a simulated pollution event using Rhodamine tracer dye
     - Use extracted GTM "endmembers" together with NS3 to map abundance
     - Identify abundance of algae
     - Map evolution of Rhodamine dye plume

Materials and Methods

1. Data Collection

   - Describe collection at site in Montague, North Texas

2. Autonomous Robot Team

   - Describe drone
     - Alta X Quadcopter
     - Resonon Pika XC2 HSI
     - Intel Nuc compute
   - describe georectification procedure

3. Pre-Processing
4. Generative Topographic Mapping

- Original SOM paper (Kohonen 1990)
- Heskes reinterpretation of SOM algorithm as stochastic gradient descent on an energy suface i.e. justification with a cost function (Heskes 1999)
- GTM paper (Bishop, Svensén, and Williams 1998b)

  - Describe algorithm
  - Describe parameters

- Bayesian Information Criterion and Akaike Information Criterion for Hyperparameter Optimization

  - A full Bayesian treatment would involve marginalization over all possible parameter values but this is much more computational efficient.

5. Abundance Mapping with the Normalized Spectral Similarity Score

   - Describe different metrics for comparison of spectra
     - Trade-off between metrics that distinguish changes in overall intensity and changes in peak location, i.e. hue

Results

1. Two Case Studies
2. GTM Fit on Full Dataset
3. Hyperparameter Optimization with BIC
4. Spectral Signature Identification
5. Water Class Map (Water Only GTM)
6. Algae Identification
7. Dye Plume Identification

Discussion

1. GTM Applications in our Paper
2. Comparison to existing approaches
    - Unsupervised Classification
    - Endmember Extraction & Spectral Unmixing
    - Applications for either of these with Drones?
3. Use of GTM and Unsupervised methods for Robot Team
    - As a feature transformer/preprocessor for supervised methods (as opposed to, say, PCA or k-NN)
    - identification of "interesting" regions for intelligent deployment of robotic boat
        - don't continue collecting data in region that is similar
        - provision boat to maximize data collection across distribution of GTM classes
4. Extensions and future work
    - Incorporation of GTM extensions for efficient online-learning
        - Currently GTM classes can quickly be applied (in real time) once they are trained to generate abundance maps e.g. chlorophyll, algae, etc…
        - Key limitation of GTM is similar to SOM and k-NN: you need to compute distances between samples and latent nodes. This can scale poorly for very large datasets
        - Bishop suggests a way to augment GTM training algorithm to enable online training of GTM. This could be deployed on-board the UAV to perform classification and identify spectral signatures in near-real time.
        - Extensions of GTM (Bishop, Svensén, and Williams 1998a)
    - Utilize the GTM with in-situ data collection to identify spectral signatures of specific algal species

Conclusions

NOTES AND CITATIONS

1. Applications of remote sensing to water quality
    - machine learning (XGBoost) used with paired remote sensing imagery and in-situ data to classify imagery into 5 categories including 3 for harmful algal blooms (Ghatkar, Singh, and Shanmugam 2019)
    - Spectral signatures of classes are complex and often overlapping (Thenkabail, Lyon, and Huete 2018)
    - Remote sensing used for oil spill analysis: extent and thickness mapping (Kokaly et al. 2013; Leifer et al. 2012)
    - Sun glitter remains a key challenge for sensing in the visible portion of the spectrum but multi and hyperspectral imagers have been used for oil spill identification and to identify their impacts on vegetation stress and mortality (Fingas and Brown 2014; Khan et al. 2018)
2. Hyperspectral Imagery
    - applications include food quality & safety, medical diagnoses, precision agriculture, and forensic document examination (Khan et al. 2018)
    - Hyperspectral data were used to distinguish oils by type, e.g. crude, diesel, gasoline, and palm (Yang et al. 2020)
3. Supervised regression and classification for water quality
    - common approach is inversion of optically-active water quality parameters such as chlorophyll-a, blue-green algae, turbidity, and temperature (Ritchie, Zimba, and Everitt 2003)
    - Combining spectral indices such as the NDVI together with machine learning is a popular approach (Thenkabail, Lyon, and Huete 2018; Sagan et al. 2020; Lu et al. 2021)
    - Polynomial regression models for chlorophyll-a, turbidity (D. Zhang, Zeng, and He 2022)
    - Key limitation is collection of sufficient quantity of in-situ reference data
    - Ross et al. created a comprehensive dataset with over 600,000 water quality records matching optically active water quality parameters with associated satellite imagery from Landsat 5,7,8. To achieve this quantity of data, they needed records spanning 1984 to 2019 (Ross et al. 2019)
    - Aurin et al. took a similar route combining 30 years of remote sensing imagery with in-situ data from over 500 field compaigns for CDOM, organic carbon, etc… (Aurin, Mannino, and Lary 2018).

- When ground truth data aren't available in sufficient quantity, classification into water quality categories is another approach (ground truth can be easier to obtain by expert analysis of scene) (Koponen et al. 2002)
- When no ground-truth data are available, unsupervised classification can still help partition imagery into groups or clusters.
- Many data-driven, ML methods have been employed for the task including various matrix factorizations, k-nearest neighbors, fuzzy c-means, density estimation methods, etc. (L. Zhang et al. 2019)
- Additionally unsupervised approaches can be used to perform nonlinear dimensionality reduction & pre-processing for supervised approaches.
- SOM used for remote sensing imagery classification (Wan and Fraser 2000)
- SOM used for clustering and data compression of HSI cube-sat (Danielsen, Johansen, and Garrett 2021)
- SOM used for land-use and land-cover change analysis (Penfound and Vaz 2021)

4. Endmember Extraction and Spectral Unmixing

   - This is a related problem where the goal is to identify unique spectral signatures which combine (linearly or non-linearly) to produce the measured signal.
   - Having identified the set of endmembers, we then seek to determine their relative abundance in each pixel
   - Many statistical approaches for extracting endmembers (Berman et al. 2004)
   - Sparse PCA used for endmember extraction (Yousefi et al. 2016)
   - Popular spectral unmixing approach is to treat "pure" endmembers as vertices of a simplex [Plaza et al. (2012); nascimento2005vertex]
   - Convolutional Neural Networks with Autoencoder architectures are a popular ML approach which identify endmembers and perform non-linear unmixing (Palsson, Ulfarsson, and Sveinsson 2020; Su et al. 2017, 2019; Borsoi, Imbiriba, and Bermudez 2019)
   - Self Organizing Map is another approach which has been used for endmember extraction together with neural networks for abundance mapping (unmixing) (Cantero et al. 2004)
   - SOM has bee nused for identifying synoptic-scale patterns in wind and sea surface temperature data (Richardson, Risien, and Shillington 2003)
   - SOM used for HSI feature extraction (Ceylan and Kaya 2021)
   - drawback of autoencoder and other statistical approach is lack of a (topological) relationship between the classes/endmembers. Is endmember 1 closer to endmember 2 or endmember 10? The SOM addresses this. The GTM is even better…

5. Abundance Mapping

   - Many different spectral similarity functions exist for comparing spectra. Each have trade-offs between ability to distinguish differences in intensity versus differences in peak location (hue) (Deborah, Richard, and Hardeberg 2015)

- spectral angle mapper is popular similarity function used in endmember extraction. The SAM is nice because it is more sensitive to shape than scale (Jiang, Werff, and Meer 2020)
- spectral correlation mapper introduced to as a statistical alternative to SAM based on covariance instead of spectral angle (De Carvalho and Meneses 2000)
- Normalized spectral similarity score developed by Nidamanuri et al. combine the mean-squared error with spectral angle for a happy medium (Nidamanuri and Zbell 2011)

6. Drone-based HSI

- Near-earth HSI addresses spatial, spectral, and temporal limitations of satellite and airborne platforms. UAV-based HSI enable fine-scale mapping (Banerjee, Raval, and Cullen 2020)
- Drones can be equipped with HSI and compute to enable rapid generation of spectral indices like the NDVI for applications such as precision agriculture (Horstrand et al. 2019)
- UAV-based HSI can be georectified to centimeter-scales without need for ground control points by using on-board GPS and IMU (Arroyo-Mora et al. 2019)
- UAV-based HSI for turbidity estimation (Vogt and Vogt 2016)

Arroyo-Mora, J Pablo, Margaret Kalacska, Deep Inamdar, Raymond Soffer, Oliver Lucanus, Janine Gorman, Tomas Naprstek, et al. 2019. "Implementation of a UAV–Hyperspectral Pushbroom Imager for Ecological Monitoring." *Drones* 3 (1): 12.

Aurin, Dirk, Antonio Mannino, and David J Lary. 2018. "Remote Sensing of CDOM, CDOM Spectral Slope, and Dissolved Organic Carbon in the Global Ocean." *Applied Sciences* 8 (12): 2687.

Banerjee, Bikram Pratap, Simit Raval, and PJ Cullen. 2020. "UAV-Hyperspectral Imaging of Spectrally Complex Environments." *International Journal of Remote Sensing* 41 (11): 4136–59.

Berman, Mark, Harri Kiiveri, Ryan Lagerstrom, Andreas Ernst, Rob Dunne, and Jonathan F Huntington. 2004. "ICE: A Statistical Approach to Identifying Endmembers in Hyperspectral Images." *IEEE Transactions on Geoscience and Remote Sensing* 42 (10): 2085–95.

Bishop, Christopher M, Markus Svensén, and Christopher KI Williams. 1998a. "Developments of the Generative Topographic Mapping." *Neurocomputing* 21 (1-3): 203–24.

———. 1998b. "GTM: The Generative Topographic Mapping." *Neural Computation* 10 (1): 215–34.

Borsoi, Ricardo Augusto, Tales Imbiriba, and José Carlos Moreira Bermudez. 2019. "Deep Generative Endmember Modeling: An Application to Unsupervised Spectral Unmixing." *IEEE Transactions on Computational Imaging* 6: 374–84.

Cantero, MC, RM Perez, Pablo J Martinez, PL Aguilar, Javier Plaza, and Antonio Plaza. 2004. "Analysis of the Behavior of a Neural Network Model in the Identification and Quantification of Hyperspectral Signatures Applied to the Determination of Water Quality." In *Chemical and Biological Standoff Detection II*, 5584:174–85. SPIE.

Ceylan, Oğuzhan, and Gülsen Taskin Kaya. 2021. "Feature Selection Using Self Organizing Map Oriented Evolutionary Approach." *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 4003–6. https://api.semanticscholar.org/CorpusID:238750026.

Danielsen, Aksel S, Tor Arne Johansen, and Joseph L Garrett. 2021. "Self-Organizing Maps for Clustering Hyperspectral Images on-Board a Cubesat." *Remote Sensing* 13 (20): 4174.

De Carvalho, O Abilio, and Paulo Roberto Meneses. 2000. "Spectral Correlation Mapper (SCM): An Improvement on the Spectral Angle Mapper (SAM)." In *Summaries of the 9th JPL Airborne Earth Science Workshop, JPL Publication 00-18*, 9:2. JPL publication Pasadena, CA, USA.

Deborah, Hilda, Noël Richard, and Jon Yngve Hardeberg. 2015. "A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (6): 3224–34.

Fingas, Merv, and Carl Brown. 2014. "Review of Oil Spill Remote Sensing." *Marine Pollution Bulletin* 83 (1): 9–23.

Ghatkar, Jayesh Ganpat, Rakesh Kumar Singh, and Palanisamy Shanmugam. 2019. "Classification of Algal Bloom Species from Remote Sensing Data Using an Extreme Gradient Boosted Decision Tree Model." *International Journal of Remote Sensing* 40 (24): 9412–38.

Heskes, Tom. 1999. "Energy Functions for Self-Organizing Maps." In *Kohonen Maps*, 303–15. Elsevier.

Horstrand, Pablo, Raúl Guerra, Aythami Rodríguez, María Díaz, Sebastián López, and José Fco López. 2019. "A UAV Platform Based on a Hyperspectral Sensor for Image Capturing and on-Board Processing." *IEEE Access* 7: 66919–38.

Jiang, Tingxuan, Harald van der Werff, and Freek van der Meer. 2020. "Classification Endmember Selection with Multi-Temporal Hyperspectral Data." *Remote Sensing* 12 (10): 1575.

Khan, Muhammad Jaleed, Hamid Saeed Khan, Adeel Yousaf, Khurram Khurshid, and Asad Abbas. 2018. "Modern Trends in Hyperspectral Image Analysis: A Review." *Ieee Access* 6: 14118–29.

Kohonen, Teuvo. 1990. "The Self-Organizing Map." *Proceedings of the IEEE* 78 (9): 1464–80.

Kokaly, Raymond F, Brady R Couvillion, JoAnn M Holloway, Dar A Roberts, Susan L Ustin, Seth H Peterson, Shruti Khanna, and Sarai C Piazza. 2013. "Spectroscopic Remote Sensing of the Distribution and Persistence of Oil from the Deepwater Horizon Spill in Barataria Bay Marshes." *Remote Sensing of Environment* 129: 210–30.

Koponen, Sampsa, Jouni Pulliainen, Kari Kallio, and Martti Hallikainen. 2002. "Lake Water Quality Classification with Airborne Hyperspectral Spectrometer and Simulated MERIS Data." *Remote Sensing of Environment* 79 (1): 51–59.

Leifer, Ira, William J Lehr, Debra Simecek-Beatty, Eliza Bradley, Roger Clark, Philip Dennison, Yongxiang Hu, et al. 2012. "State of the Art Satellite and Airborne Marine Oil Spill Remote Sensing: Application to the BP Deepwater Horizon Oil Spill." *Remote Sensing of Environment* 124: 185–209.

Lu, Qikai, Wei Si, Lifei Wei, Zhongqiang Li, Zhihong Xia, Song Ye, and Yu Xia. 2021. "Retrieval of Water Quality from UAV-Borne Hyperspectral Imagery: A Comparative Study of Machine Learning Algorithms." *Remote Sensing* 13 (19): 3928.

Nidamanuri, Rama Rao, and Bernd Zbell. 2011. "Normalized Spectral Similarity Score (NS^3) as an Efficient Spectral Library Searching Method for Hyperspectral Image Classification." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4: 226–40. https://api.semanticscholar.org/CorpusID:32567143.

Palsson, Burkni, Magnus O Ulfarsson, and Johannes R Sveinsson. 2020. "Convolutional Autoencoder for Spectral–Spatial Hyperspectral Unmixing." *IEEE Transactions on Geoscience and Remote Sensing* 59 (1): 535–49.

Penfound, Elissa, and Eric Vaz. 2021. "Analysis of Wetland Landcover Change in Great Lakes Urban Areas Using Self-Organizing Maps." *Remote Sensing* 13 (24): 4960.

Plaza, Javier, Eligius MT Hendrix, Inmaculada García, Gabriel Martín, and Antonio Plaza. 2012. "On Endmember Identification in Hyperspectral Images Without Pure Pixels: A Comparison of Algorithms." *Journal of Mathematical Imaging and Vision* 42: 163–75.

Richardson, Anthony J, C Risien, and Frank Alan Shillington. 2003. "Using Self-Organizing Maps to Identify Patterns in Satellite Imagery." *Progress in Oceanography* 59 (2-3): 223–39.

Ritchie, Jerry C, Paul V Zimba, and James H Everitt. 2003. "Remote Sensing Techniques to Assess Water Quality." *Photogrammetric Engineering & Remote Sensing* 69 (6): 695–704.

Ross, Matthew RV, Simon N Topp, Alison P Appling, Xiao Yang, Catherine Kuhn, David Butman, Marc Simard, and Tamlin M Pavelsky. 2019. "AquaSat: A Data Set to Enable Remote Sensing of Water Quality for Inland Waters." *Water Resources Research* 55 (11): 10012–25.

Sagan, Vasit, Kyle T Peterson, Maitiniyazi Maimaitijiang, Paheding Sidike, John Sloan, Benjamin A Greeling, Samar Maalouf, and Craig Adams. 2020. "Monitoring Inland Water Quality Using Remote Sensing: Potential and Limitations of Spectral Indices, Bio-Optical Simulations, Machine Learning, and Cloud Computing." *Earth-Science Reviews* 205: 103187.

Su, Yuanchao, Jun Li, Antonio Plaza, Andrea Marinoni, Paolo Gamba, and Somdatta Chakravortty. 2019. "DAEN: Deep Autoencoder Networks for Hyperspectral Unmixing." *IEEE Transactions on Geoscience and Remote Sensing* 57 (7): 4309–21.

Su, Yuanchao, Andrea Marinoni, Jun Li, Antonio Plaza, and Paolo Gamba. 2017. "Nonnegative Sparse Autoencoder for Robust Endmember Extraction from Remotely Sensed Hyperspectral Images." In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 205–8. https://doi.org/10.1109/IGARSS.2017.8126930.

Thenkabail, Prasad S, John G Lyon, and Alfredo Huete. 2018. *Hyperspectral Indices and Image Classifications for Agriculture and Vegetation*. CRC press.

Vogt, Michael C, and Mark E Vogt. 2016. "Near-Remote Sensing of Water Turbidity Using Small Unmanned Aircraft Systems." *Environmental Practice* 18 (1): 18–31.

Wan, Weijian, and Donald Fraser. 2000. "A Multiple Self-Organizing Map Scheme for Remote Sensing Classification." In *International Workshop on Multiple Classifier Systems*, 300–309. Springer.

Yang, Junfang, Jianhua Wan, Yi Ma, Jie Zhang, and Yabin Hu. 2020. "Characterization

Analysis and Identification of Common Marine Oil Spill Types Using Hyperspectral Remote Sensing.” *International Journal of Remote Sensing* 41 (18): 7163–85.

Yousefi, Bardia, Saeed Sojasi, Clemente Ibarra Castanedo, Georges Beaudoin, François Huot, Xavier PV Maldague, Martin Chamberland, and Erik Lalonde. 2016. “Mineral Identification in Hyperspectral Imaging Using Sparse-PCA.” In *Thermosense: Thermal Infrared Applications XXXVIII*, 9861:312–22. SPIE.

Zhang, Dingyu, Siyu Zeng, and Weiqi He. 2022. “Selection and Quantification of Best Water Quality Indicators Using UAV-Mounted Hyperspectral Data: A Case Focusing on a Local River Network in Suzhou City, China.” *Sustainability* 14 (23): 16226.

Zhang, Lefei, Liangpei Zhang, Bo Du, Jane You, and Dacheng Tao. 2019. “Hyperspectral Image Unsupervised Classification by Robust Manifold Matrix Factorization.” *Information Sciences* 485: 154–69.