Acoustic-to-Hyper-Spectral: Hyper-Spectral Image Construction from Frequency Spectrums through Simulated Annealing (Student Abstract)

Ruth-Emely Pierau[1,2], **Alaster Meehan**[2], **Hamid Rezatofighi**[1], **Peter J. Stuckey**[1]
[1] Department of Data Science and Artificial Intelligence, Monash University
[2]Future Fibre Technologies, 10 Hartnett Cl, Mulgrave VIC 3170, Australia
ruth-emely.pierau@monash.edu, ameehan@fftsecurity.com, hamid.rezatofighi@monash.edu, peter.stuckey@monash.edu

Introduction

Hyper-spectral images describe all images with more than three image channels that are derived from a frequency spectrum to obtain the spectrum for each image pixel. They can be constructed by hyper-spectral image cameras or from the frequency spectrum of various sensing applications such as distributed acoustic sensing (DAS) systems. In this paper, we use the frequency spectrum from DAS systems to construct hyper-spectral images for downstream computer vision tasks such as object detection. The aim is to construct an image that represents features of signals from events while repressing noise.

The frequency spectrums are collected from a commercially deployed DAS system from my industry partner, Future Fibre Technology, an AVA Risk Group company. DAS systems can be used to continuously monitor vibrations and acoustic signals over long distances in real-time with a high sensitivity (Wang et al. 2014; Gabai and Eyal 2016; Xu et al. 2022). They are often deployed for the monitoring and protection of commercial assets and infrastructure such as airports, power plants, and pipelines (Zhu et al. 2022; Chen et al. 2023; Wu et al. 2015). Each system generally consists of a controller unit and a fiber optical cable. The fiber transmits laser pulses and scatters the light. The reflection of the scattered light is sampled every 0.5 m. The pattern of the backscatter is changed by disturbances. Disturbances can either be types of events or noise such as wind and wild animals. Events are typically types of intrusions, structural changes, or defined actions and structures, e.g. fence climbing. The raw acoustic signal sensed by the fiber cable is transformed via Fast Fourier Transformation to a frequency spectrum. The frequency spectrum is then used to construct hyper-spectral images that represent distance in width and time in height (Sha et al. 2021; He and Liu 2021; Wang et al. 2020).

Mapping one frequency to one image channel generally leads to overlarge images due to the wide range of frequencies considered, i.e. we consider frequency ranges of 1000 Hz or 2500 Hz. We need to select frequencies that represent signal information and drop frequencies that contain noise. Frequencies can be binned together into frequency bands where each band forms one image channel. The current state-of-the-art approaches for DAS systems integrated with object detectors rely on frequency band construction either by expert knowledge or by trial-and-error (Ma et al. 2023; Liu et al. 2019; Tejedor et al. 2017).

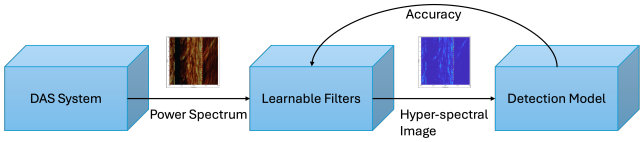

Figure 1: Schema of Learnable Filters integrated with a DAS system and an Object Detector

Our approach is based on simulated annealing to train filters to automatically select frequencies and bin them into frequency bands. The aim is to represent a variety of signal information and decrease noise. Signal representation is optimized for object detection on time-sequenced images with a set number of image channels. The object detection task consists of localizing and classifying events in the constructed hyper-spectral images. Events and noise often share at least some frequencies and vary between application types.

Hyper-Spectral Image Construction

We design a simulated annealing based approach that constructs hyper-spectral images from the frequency spectrums of a DAS system and iteratively improves them through the training of learnable filters. The learnable filters select frequency bands from the frequency spectrum, also power spectrum, that represent signal information, and drop frequencies that contain only noise. Our fully integrated model consists of a DAS system, learnable filters, and an object detector. Figure 1 shows a schema of our model. The object detector at the end of the pipeline is trained in conjunction with the learnable filters. The hyper-spectral images are specifically constructed for the downstream computer vision task. The object detector serves as both the measure and the selector by providing accuracy as a metric to the learnable filters. The filters are adjusted through simulated annealing relying on the provided metric.

Our learnable filters take as input a frequency spec-

trum from our DAS system. Before joint model training of the learnable filters and the object detector starts, hyper-spectral images are constructed via uniform filters, i.e. the frequency spectrum has been uniformly divided into n frequency bands. The frequency bands are formed by binning and averaging frequencies. Each band is mapped to one image channel. The object detection model is trained on the uniformly created hyper-spectral images for a set number of epochs and provides mean average precision (mAP), an accuracy metric, to the learnable filters. The filters are then adjusted through simulated annealing that uses mAP in its acceptance criterion.

We designed two actions for our simulated annealing model; each consists of two local moves: Split and drop, or split and merge. For split and drop, one channel is split equally into half, the another is dropped. For split and merge, one channel is split equally into half, the another is merged with its neighboring channel. Both the action and the two channels are randomly chosen. Once the action has been applied to the frequency spectrum, the hyper-spectral images are re-created. The object detector is further trained on the re-created hyper-spectral images. Before training resumes, the weights of the object detector are similarly adjusted by dropping the weight in the first layer associated with the dropped frequency band and replacing it with the weight of the frequency band that was either split or merged.

After training for a set number of epochs, the acceptance criterion is calculated based on the provided mAP. If model performance increases, the adaptation for the filters is accepted. If model performance decreases, the adjustment is discarded and the model reverts back to the previous filter configuration and model weights. Temperature is decreased. The model continues to adapt the filters for the hyper-spectral image construction and to train on the adapted hyper-spectral images until either the temperature decreases to zero or the total number of epochs for model training is reached. Alternatively, if the temperature is not set, the acceptance criterion of the simulated annealing must strictly improve. The filters are adjusted via hill climbing and the model is trained for the set number of epochs.

We integrated our simulated annealing based approach with yolov7 tiny (Wang, Bochkovskiy, and Liao 2023). We set an initial temperature of 100 and decrease by 4. We also consider a hill climb approach. The model is trained for 300 epochs in total. The initial 60 epochs are reserved for training on the uniformly constructed hyper-spectral images. Afterwards, the model is trained for 10 epochs after each filter adaptation before the acceptance criterion is calculated. We use COCO mAP evaluated on the training set as the fitness score (Lin et al. 2014). We also calculate the acceptance criterion before model training resumes. We evaluate 10 different random filter adaptations and select the best performing filter configuration. The dataset is from a commercially deployed DAS system installation and consists of 3442 images split into 80% training and validation set and 20% test set. We consider as baseline methods yolov7 trained on hyper-spectral images hand-crafted with expert knowledge and constructed with uniform filters.

Table 1 shows the results of our ablation study. Our simulated annealing approach significantly outperforms the expert knowledge based approach by a margin of 22.2% and the uniform filtering method by a margin of 18.3% in mAP Hill climb is outperformed by a smaller margin of 0.4% - 2.5% in mAP which seems to confirm that allowing performance to temporarily decrease can lead to a wider exploration of the search domain and eventually a better model performance. The recall, measuring how many of the objects are detected, shows similar results and corresponds to the conclusions drawn based on the accuracy.

Method	Variation	mAP	Recall
Hand-crafted	N/A	0.571	0.706
Uniform	N/A	0.610	0.744
Hill Climb	Seed 132	0.789	0.868
Hill Climb	Seed 768	0.768	0.853
Simulated Annealing	Seed 132	0.793	0.869

Table 1: Ablation Study

We also compare the same methods initiated on different seeds. The performance of the hill climb approach varies by 2.1%. Our model depends on randomness to search the solution space. Different random seeds can have a large impact on the filter configuration found and ultimately the model performance. One reason might be the limitations imposed by the actions. Once a frequency band is dropped, it cannot be re-introduced. Essentially, parts of the search domain are cut off with each accepted filter configuration. Large parts of the frequency spectrum are discarded without further consideration. Additionally, already formed frequency bands can only be split, merged, and dropped. Many potential combinations of frequencies are excluded by design. Each frequency band can only be split a set number of times before it becomes too small.

The limitations to the available actions was introduced to decrease model training time and purposefully guide hyper-spectral image construction similarly to previous approaches via hand-crafting by expert knowledge. Further work to explore more and different sets of actions is needed. Likewise, the model should become less reliant on randomness. Despite our current limitations, our model has generalized well to the unseen test set of our dataset and outperforms previous state-of-the-art methods.

Acknowledgments

This research was fully funded by the Australian Government through the Australian Research Council Industrial Transformation Training Centre in Optimisation Technologies, Integrated Methodologies, and Applications (OPTIMA), Project ID IC200100009, and by Future Fibre Technologies, an AVA Risk Group company.

References

Chen, S.; Han, J.; Sui, Q.; Zhu, K.; Lu, C.; and Li, Z. 2023. Advanced Signal Processing in Distributed Acoustic Sensors Based on Submarine Cables for Seismology Applications. *Journal of Lightwave Technology*.

Gabai, H.; and Eyal, A. 2016. On the sensitivity of distributed acoustic sensing. *Optics letters*, 41(24): 5648–5651.

He, Z.; and Liu, Q. 2021. Optical fiber distributed acoustic sensors: A review. *Journal of Lightwave Technology*, 39(12): 3671–3686.

Lin, T.-Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; and Zitnick, C. L. 2014. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13*, 740–755. Springer.

Liu, H.; Ma, J.; Xu, T.; Yan, W.; Ma, L.; and Zhang, X. 2019. Vehicle detection and classification using distributed fiber optic acoustic sensing. *IEEE Transactions on Vehicular Technology*, 69(2): 1363–1374.

Ma, B.; Gao, R.; Zhang, J.; and Zhu, X. 2023. A YOLOX-Based Automatic Monitoring Approach of Broken Wires in Prestressed Concrete Cylinder Pipe Using Fiber-Optic Distributed Acoustic Sensors. *Sensors*, 23(4): 2090.

Sha, Z.; Feng, H.; Rui, X.; and Zeng, Z. 2021. PIG Tracking utilizing fiber optic distributed vibration sensor and YOLO. *Journal of Lightwave Technology*, 39(13): 4535–4541.

Tejedor, J.; Macias-Guarasa, J.; Martins, H. F.; Pastor-Graells, J.; Corredera, P.; and Martin-Lopez, S. 2017. Machine learning methods for pipeline surveillance systems based on distributed acoustic sensing: A review. *Applied Sciences*, 7(8): 841.

Wang, C.-Y.; Bochkovskiy, A.; and Liao, H.-Y. M. 2023. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 7464–7475.

Wang, Z.; Lu, B.; Ye, Q.; and Cai, H. 2020. Recent progress in distributed fiber acoustic sensing with Φ-OTDR. *Sensors*, 20(22): 6594.

Wang, Z.; Zeng, J.; Li, J.; Peng, F.; Zhang, L.; Zhou, Y.; Wu, H.; and Rao, Y. 2014. 175km phase-sensitive OTDR with hybrid distributed amplification. In *23rd international conference on optical fibre sensors*, volume 9157, 1562–1565. SPIE.

Wu, H.; Sun, Z.; Qian, Y.; Zhang, T.; and Rao, Y. 2015. A hydrostatic leak test for water pipeline by using distributed optical fiber vibration sensing system. In *Fifth Asia-Pacific Optical Sensors Conference*, volume 9655, 568–571. SPIE.

Xu, W.; Yu, F.; Liu, S.; Xiao, D.; Hu, J.; Zhao, F.; Lin, W.; Wang, G.; Shen, X.; Wang, W.; et al. 2022. Real-time multi-class disturbance detection for Φ-OTDR based on YOLO algorithm. *Sensors*, 22(5): 1994.

Zhu, H.-H.; Liu, W.; Wang, T.; Su, J.-W.; and Shi, B. 2022. Distributed acoustic sensing for monitoring linear infrastructures: Current status and trends. *Sensors*, 22(19): 7550.