SLAM in Mining Environment

A proposal for an immediate control of sensor fusion

Tuan Khoi Nguyen

The University of Melbourne
Faculty of Engineering & IT

November 19, 2022

Background Context

Working in mining environment

Human have to operate in hazard environment of mining: fumes, debris, noises [1], etc.

The use of Unmanned Vehicles

Unmanned vehicles (UVs) can operate in place of human. Some example tasks:

- Navigation
- Exploration

Figure: *A vehicle approaching mine tunnel [2]*

Simultaneous Localization and Mapping (SLAM)

- Constructing a map of the surrounding environment, while trying to get self's location in that map.
- Have many different methods and specialized methods: Visual SLAM, LIDAR SLAM, EKF SLAM, etc.

SLAM applications in mining

Gain useful information for user, for example:

- Slope detection
- Obstacle detection and tracking
- Path planning

SLAM Challenges

- Mining environment dark & dusty [3, 4], not good for SLAM inputs.
- GPS may not work [3].
- Current industry computational resources has limited capacity [5].

Figure: *Object detection in a dark environment [6]*

Background Context

The SLAM problem in mining

Aim and Direction (Initial Research Question)

How can we utilise sensors to improve on our SLAM results, so that the UV can efficiently operate in visually degraded environments?

Literature Review

The Process

Aim

Look for potential ideas that can solve our research problem.

Keywords: 'low-light', 'foggy view', 'SLAM'

Questions of Interest

- What methodologies are in this paper?
- What environment condition is of interest for the method?
- What physical resources were required (sensors, cameras,. . .)?
- What factor has not been considered?

Device	Advantage	Problem
Unmanned Ground Vehicle (UGV)	Large load capacity, accurate close-up view, simpler control system	Small view scope, ground-only operation [7]
Unmanned Aerial Vehicle (UAV)	Move easily in complex terrains, larger observation scope from above for wide environment	Low load capacity, more complicated control system, limited close-up observation [7, 8]

Observation

- UGV is suitable for common mining task: close-up obstacle detection, load carrying.
- UAV's airborne advantages may not be utilized in mining environment, which have narrow spaces.

Therefore, UGV would be best-suited for our mining problem.

Device	Advantage	Problem
Microprocessor/ Micro-controller	Lightweight	Only capable of executing single code lines [9]
Miniature PC	Compact size, PC-standard computation power [9]	Still have memory limitation on advanced operations like deep learning
Multi-device architectures [10, 11]	Flexible computation capacity	Heavy load, costly

Observation

- Microprocessors and micro-controllers is not feasible for SLAM.
- Multi-device architectures might overload the UGV, given that it needs to carry task loads as well.

A mini-PC (such as Intel NUC 11 Pro) will fit with our SLAM problem. However, careful computation needs to be considered, given that its memory capacity is not unlimited [1].

[1] For Intel NUC 11 Pro, the capacity is 16 Gigabytes. Less costly alternatives have 4 to 8 Gigabytes

Sensor variations of SLAM

- Visual SLAM (Camera, LED): Take pictures from various angles and construct map from stitched images.
- LIDAR SLAM (2D LIDAR, 3D LIDAR): Uses LIDAR to scan the surrounding map.
- RGB-D SLAM (RGB-D camera): Uses RGB-D camera to retrieve estimation of map.

Additional sensors

- Internal Measurement Unit (IMU)
- Radar

Device	Advantage	Problem
Camera	Primary input for Visual SLAM [12]	Not working well under visual challenges
LED [4, 5]	No extra computation, efficient in low-light	Not efficient under diffusion (dust, fog)
2D LIDAR [5]	Input of LIDAR SLAM, efficient computation, work in low-light	Inefficient during movement, restricted to planar projection
3D LIDAR [13]	Input of LIDAR SLAM, 360^o cover, maximized depth capture [13]	Computationally expensive, sparse resolution [13]
RGB-D [3, 6]	Input of RGB-D SLAM, simple integration, high resolution data	Lack depth information [13]
IMU [4, 6]	Immune to all visual challenges	Inefficient if GPS not available, drifting
Radar [14, 15]	Robust to diffusion [14]	Low accuracy, hard to track movement, can be distorted in close-range [15]

THE UNIVERSITY OF MELBOURNE

Observation

- Very few papers considered diffused environments [14, 15], which is common in mining.
- One's advantage is able to cover other's disadvantage → Sensor fusion is beneficial, regardless!

Choosing the main SLAM input

- To pick a specific algorithm requires in-depth research and rigid experimentation procedure.
- But if no information is known, Visual SLAM can be tried first:
  - Does not store depth information → Most memory efficient.
  - Image input - supported by many code libraries → Easy to work with.

The next parts will visualize on visual SLAM images for simplicity. Same properties can apply to other input types.

Deep Learning

- Uses neural network to learn deep latent representation of data.
- Can be tailored to achieve state-of-the-art in specific tasks.
- 2 forms of learning:
  - **Supervised**: uses already available data to train model, determine action using prior knowledge.
  - **Unsupervised**: dynamically adjust model to adapt to changes within the environment.

- **Supervised**: Not able to get the ground truth before operation (exploration tasks)
- **Unsupervised**: Need to store large amount of information, repeatedly consuming computation processes

$\Longrightarrow$ Deep learning not suitable with our problem

GraphSLAM [16]

- Uses graph and probabilities to represent data.
- Make estimation based on likelihood optimization and Bayes Theorem.
- Robust to external disturbance, logic-based

- Extensive probabilistic models, may exceed memory if state space large. Also requires some pre-training [16].
- Bayes Theorem assume independence, may not hold well if sensor effects are related

$\Longrightarrow$ GraphSLAM not suitable with our problem

Literature Review

Fusion Method

EKF SLAM [3, 6, 14]

- Feedback model: use estimation correctness to adjust the model.
- Extended Kalman Filter: Weighted balance between model estimation and measurement.
- Computationally efficient & fast: Discard past information, approximated low-order model.

- Exactly matches our need of low computation cost and online SLAM learning
- Diversified system compatibility: Can combine multiple flows of sensors

$\Longrightarrow$ EKF SLAM is suitable with our problem

Problem Formulation

Summary & Refine

Summary

- UGV is suitable for its compatibility in mining tasks.
- Computation should not exceed the capacity of a standard mini-PC.
- Sensor combination is the best way to alleviate each's problems.
- EKF is best fit as sensor fusion method for online SLAM learning problem.

Refined Research Problem

"How can we utilise measurements of multiple sensors to improve on our SLAM results, so that the UGV can operate in visually degraded environments, with sufficient use of computation similar to a standard industrial PC?"

Semi-Discrete Environment

Context

EKF SLAM Common Structure

Consists of one high resolution - low frequency and one low resolution - high frequency estimator. The latter is used to update the former's measurement.

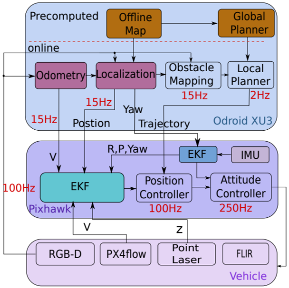

Figure: *Structure in pose estimation [3]*

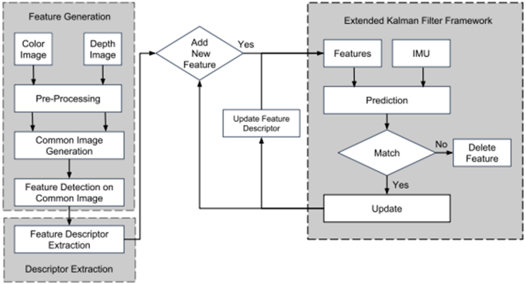

Figure: *Structure in odometry estimation [6]*

Low-frequency signal (LFS)

Usually the primary inputs of SLAM. Rich in information, but requires longer time to get:

- High-resolution camera
- LIDAR
- RGB-D camera

High-frequency signal (HFS)

Does not contain much information, but can be retrieved fast:

- IMU
- Low-resolution camera

Semi-Discrete Environment

Context

Observation

The high-frequency signal (HFS) can give extra information for the low-frequency signal (LFS).

Question

Can we provide the high-frequency signal (HFS) with something from low-frequency signal (LFS)?

Semi-Discrete Environment

Formulation

Idea

An immediate controller can give feedback to HFS from LFS. To do that, it can extract details that will not change rapidly over time t.

- L: Does not change over time.
- $\Delta l(t)$: Little change over time.
- $h(t)$: Changes rapidly over time.

Source: Images taken from Youtube

$$l(t) + h(t) = L + \Delta l(t) + h(t)$$

$$L + \Delta l(t) \approx L$$

Figure: *Formulation of an image sequence*

Semi-discrete Environment (SDE)

The environment $E(t)$ at time t consists of 2 parts: one that changes fast over time: $h(t)$, and another that has little change in the same amount of time, which can be approximated to no change at all: $l(t) \approx L$.

$$E = l(t) + h(t) \approx L + h(t)$$

Immediate Controller

An immediate controller will be able to extract L from LFS, and then combine it with $h(t)$ from HFS to get an estimation of the environment.

Semi-discrete Environment in SLAM

For our SLAM problem, an environment E have the following elements:

- L: locations of objects and ground truth map.
- $h(t)$: changes in object position relative to the UGV.

Visual SLAM image input as an environment

- L (static objects, surroundings): Color frequencies, relative distance between objects.
- $h(t)$ (relative position): Colors' spatial locations.

Example below: Right camera is the same as left camera, but shifted slightly to the right.

Figure: *Left stereo camera* [1]

Figure: *Right stereo camera* [1]

[1] Taken from Driving Stereo dataset [17]

SDE Extraction

Observation

Image inputs: color frequencies and their relative position does not change, but the overall spatial location changes.

$\Longrightarrow$ Needs a frequency extraction method

THE UNIVERSITY OF MELBOURNE

Method

- Decompose signal into linear combination of sinusoids.
- Only capture frequencies and scale, but don't know where they locate.

Short-Time Fourier Transform

- Segments the signal first, then perform Fourier Transform.
- Problem: each signal requires different segmentation, and auto-segmentation is a big problem of its own.

Wavelet Transform

Wavelet transform represents a signal as 2 informative parts: frequency appearances and their temporal locations [18] in the form of coefficients.

Transforms (decomposition) by convoluting signal to a given basis function [19]. Each Wavelet coefficient will represent the information amount at a specific frequency and scale.

Observation

Wavelet transform is suitable as an SDE extraction method, as it can retrieve both frequency and temporal information.

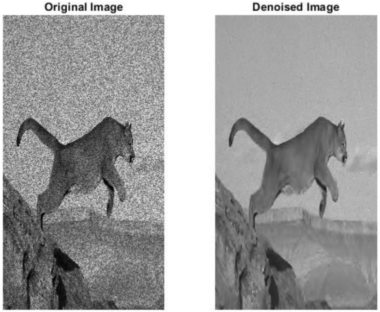

Figure: *A noisy image reconstructed with Discrete Wavelet Transform*[2]

[2] Source: MATLAB Wavelet Tutorial

THE UNIVERSITY OF MELBOURNE

Basis Function

Each basis function belongs to a wavelet family, which shares the support length and is uniquely defined by its number of vanishing points.

Parameters

- Support length (SL): Interval length of non-zero coefficients [20]. Higher values can detect more dense features in a signal.
- Number of vanishing points (NVP): 'Resolution' of wavelet representation. The larger value, the more complex rule fitted [21].
- Orthogonal: If decomposition and reconstruction have same NVP [19]. Discretized, but can maintain mean pixel values and in turn, maximize preservation of distinctive features.
- Biorthogonal: If decomposition and reconstruction have different NVP [19]. Does not maintain mean pixel values, but has linear phase that does interpolation, good for reconstructing details.

Common Wavelets

- Orthogonal, short SL: Haar, Daubechies, Symlet
- Biorthogonal: biorthogonal, reverse biorthogonal

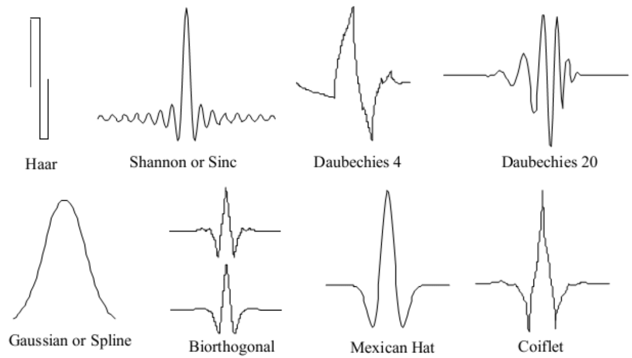

Figure: *Common wavelet examples [22]*

Wavelet Family

Can start with short SL, orthogonal wavelets (Daubechies, Symlets, etc.): all-rounded, good at preserving important information, simple.

Can switch to others when requirements are more specific.

Number of vanishing points

NVP is a parameter that have different optimal value for different signals
$\longrightarrow$ Can only tune on case-by-case basis

Wavelet Experiments

Purpose

Test the feasibility of different wavelets in terms of efficiency and computation.

Dataset

- DARPA SubT Challenge dataset [2]: images of UGV exploring tunnels, similar to mining environment.
- Driving Stereo dataset [17]: images of stereo camera pairs, demonstrating shift effects

Figure: *Samples of DARPA Dataset*

Figure: *Samples of Driving Stereo Dataset*

Procedure

1. Decomposition: Perform wavelet transform
2. Main process: Perform modification on Wavelet coefficients
3. Reconstruction: Perform inverse wavelet transform

Main process

- Denoise: Set a threshold, and set all smaller coefficient values to 0 $\longrightarrow$ Constant $O(1)$
- Compression: Set number k, set top k smallest coefficient values to 0 $\longrightarrow$ Linear $O(n)$

Wavelet Experiments

Wavelet Family Comparison (Decomposition and Reconstruction)

Result Analysis

Result does not vary between families.

Runtime includes computation overheading.

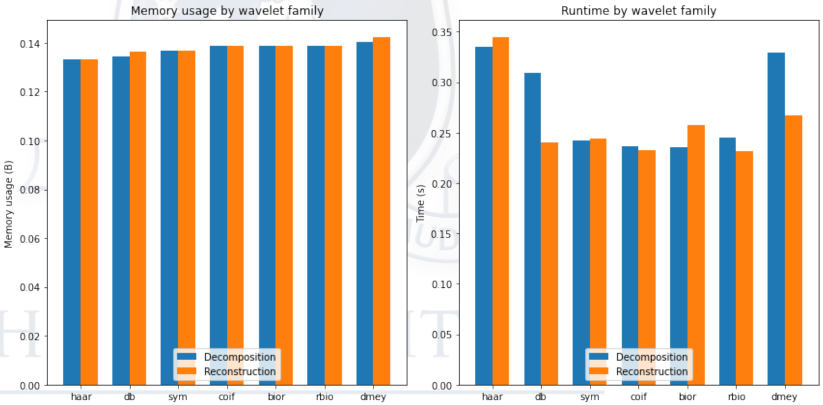

Figure: *Runtime & memory consumption for decomposition and reconsturction of different wavelet families*

Wavelet Experiments

Wavelet Family Comparison Results (Denoising)

Result Analysis

Orthogonal wavelets (first row) are strong in removing redundant features, having the best MSE

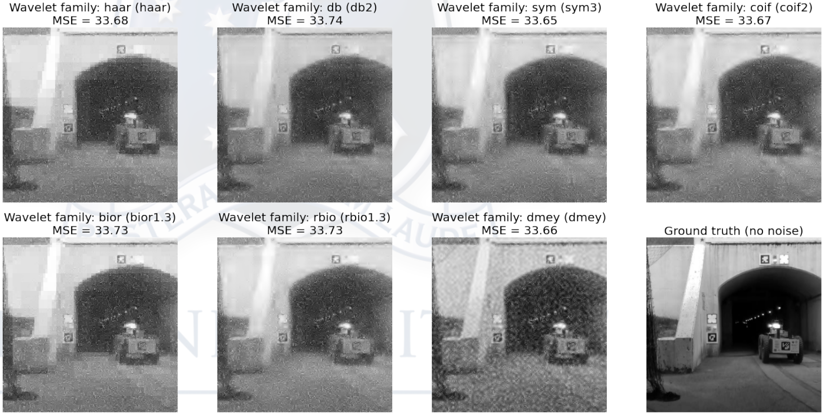

Figure: *Denoise of different wavelet families*

Wavelet Experiments

Wavelet Family Comparison Results (Image Compression)

Result Analysis

Orthogonal wavelets works well by removing redundancy.

Biorthogonal wavelets works well by interpolating features.

Symlets with 3 NVP (sym3) have significantly low MSE for both denoising and compression $\longrightarrow$ Can use as starting wavelet

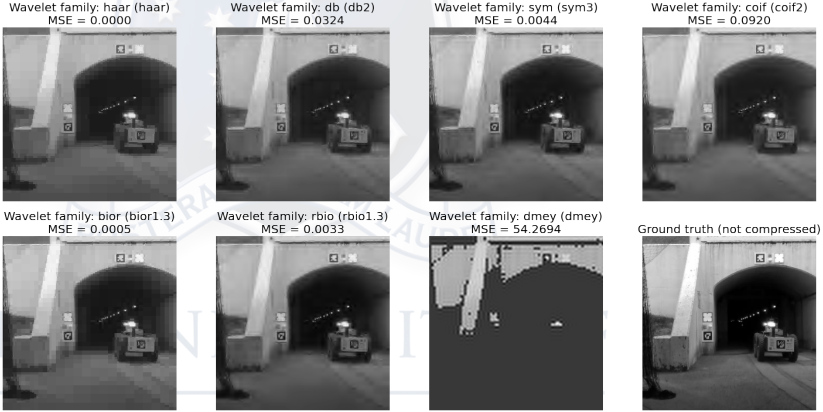

Figure: *Compression of different wavelet families*

Wavelet Experiments

Performance Results

Result Analysis

As expected, the runtime is dependent of main process's time complexity $\longrightarrow$ Need to make main process efficient as well.

The memory usage is negligible due to no local variables in main process $\longrightarrow$ Important to minimize number of variable in main process.

Ignoring computation overhead (approximately 0.2 seconds), these operations are feasible for high frequencies.

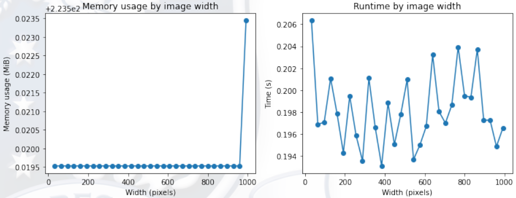

Figure: *Denoise of different image sizes*

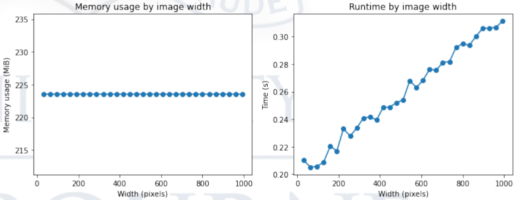

Figure: *Compression of different image sizes*

Summary

What have been investigated so far

Ideas

Overall, the following ideas have been made:

1. **Efficient computation & runtime is important** for our problem.
2. Different types of **sensors can be combined** to compensate each other.
3. **EKF SLAM model** will work as the fusion method for our SLAM problem, where its feedback structure will help with the computation limit.
4. Potential to add an **SDE extractor** to further save up computation, by only taking necessary information from low-resolution data.
5. Potential of using **frequency methods like Wavelet Transform** to **extract necessary information from low-resolution data**.

Summary

What have been investigated so far

Design choices

Overall, the following choices have been made for the system design:

1. **UGV** as main operation vehicle for high compatibility with mining tasks.
2. Processor uses **1 standard mini-PC** to balance between resource capacity and load.
3. **Visual SLAM** is a good start for lightweight information storage and well-supported libraries.
4. **Symlet 3 NVP** is a good start for wavelet choice as it is versatile and retain important information well.

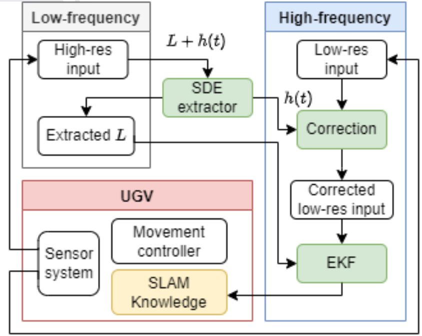

Figure: *The proposed system design*

Next Step

Possible directions from here

Choosing sensors

Rigorously choose a good combination of HFS-LFS inputs.

1. Some potential candidates: low resolution-high resolution camera pair, LIDAR and IMU, etc.
2. Requires complicated thought process, and needs well-laid experiments for approval.

Information Stitching

Develop mechanisim of combining HFS and LFS information efficiently.

1. Some examples: image stitching, panorama, LIDAR correction, etc.
2. Extensive development process, is a vast topic of its own.

References I

[1] Neal R. Haddaway, Steven J. Cooke, Pamela Lesser, Biljana Macura, Annika E. Nilsson, Jessica J. Taylor, and Kaisa Raito. Evidence of the impacts of metal mining and the effectiveness of mining mitigation measures on social–ecological systems in arctic and boreal regions: a systematic map protocol. *Environmental Evidence*, 8(1), February 2019.

[2] Chen Wang, Wenshan Wang, Yuheng Qiu, Yafei Hu, and Sebastian Scherer. Visual memorability for robotic interestingness via unsupervised online learning. In *European Conference on Computer Vision (ECCV)*, 2020.

[3] Zheng Fang, Shichao Yang, Sezal Jain, Geetesh Dubey, Silvio Maeta, Stephan Roth, Sebastian Scherer, Yu Zhang, and Stephen Nuske. *Robust Autonomous Flight in Constrained and Visually Degraded Environments*, pages 411–425. Springer International Publishing, Cham, 2016.

[4] Frank Mascarich, Shehryar Khattak, Christos Papachristos, and Kostas Alexis. A multi-modal mapping unit for autonomous exploration and mapping of underground tunnels. In *2018 IEEE Aerospace Conference*, pages 1–7, 2018.

[5] Shehryar Khattak, Christos Papachristos, and Kostas Alexis. Vision-depth landmarks and inertial fusion for navigation in degraded visual environments. In *ISVC*, 2018.

[6] Mikhail Sizintsev, Abhinav Rajvanshi, Han-Pang Chiu, Kevin Kaighn, Supun Samarasekera, and David P. Snyder. Multi-sensor fusion for motion estimation in visually-degraded environments. In *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 7–14, 2019.

[7] Jie Chen, Xing Zhang, Bin Xin, and Hao Fang. Coordination between unmanned aerial and ground vehicles: A taxonomy and optimization perspective. *IEEE Transactions on Cybernetics*, 46(4):959–972, 2016.

[8] Tam Willy Nguyen, Laurent Catoire, and Emanuele Garone. Control of a quadrotor and a ground vehicle manipulating an object. *Automatica*, 105:384–390, July 2019.

[9] Adam Stager, Herbert G. Tanner, and Erin Sparks. Design and construction of unmanned ground vehicles for sub-canopy plant phenotyping. In *Methods in Molecular Biology*, pages 191–211. Springer US, 2022.

[10] B.A. Draper, G. Kutlu, E.M. Riseman, and A.R. Hanson. ISR3: communication and data storage for an unmanned ground vehicle. In *Proceedings of 12th International Conference on Pattern Recognition*. IEEE Comput. Soc. Press.

[11] Tim Mueller-Sim, Merritt Jenkins, Justin Abel, and George Kantor. The robotanist: A ground-based agricultural robot for high-throughput crop phenotyping. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2017.

References II

[12] Michael Milford, Eleonora Vig, Walter J. Scheirer, and David Cox. Vision-based slam in changing outdoor environments. 2014.

[13] Chen Fu, Christoph Mertz, and John M. Dolan. Lidar and monocular camera fusion: On-road depth completion for autonomous driving. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pages 273–278, 2019.

[14] Ankith Manjunath, Ying Liu, Bernardo Henriques, and Armin Engstle. Radar based object detection and tracking for autonomous driving. In *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, pages 1–4, 2018.

[15] Ziyang Hong, Yvan Petillot, and Sen Wang. Radarslam: Radar based large-scale slam in all weathers. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5164–5170, 2020.

[16] Sebastian Thrun and Michael Montemerlo. The graph SLAM algorithm with applications to large-scale mapping of urban structures. *The International Journal of Robotics Research*, 25(5-6):403–429, May 2006.

[17] Guorun Yang, Xiao Song, Chaoqin Huang, Zhidong Deng, Jianping Shi, and Bolei Zhou. Drivingstereo: A large-scale dataset for stereo matching in autonomous driving scenarios. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[18] Dipalee Gupta and Siddhartha Choubey. Discrete wavelet transform for image processing. 2014.

[19] Ingrid Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, January 1992.

[20] Cristina Stolojescu, Ion Railean, Sorin Moga, and Alexandru Isar. Comparison of wavelet families with application to WiMAX traffic forecasting. In *2010 12th International Conference on Optimization of Electrical and Electronic Equipment*. IEEE, May 2010.

[21] Jonas Gomes and Luiz Velho. *From Fourier analysis to wavelets*. IMPA Monographs. Springer International Publishing, Cham, Switzerland, 1 edition, September 2015.

[22] Nasir Al-geelani, M. Afendi M. Piah, Ibrahim Saeh, Nordiana Othman, Fatin Muhamedin, and Nur Faizal Kasri. Identification of acoustic signals of internal electric discharges on glass insulator under variable applied voltage. 6:827–834, 04 2016.