24-091

Using Event-Based Imaging and Deep Learning to Generate 3D Surface Maps for Autonomous Roof Bolting

Rik Banerjee and Andrew J. Petruska
M3 Robotics Lab
Colorado School of Mines Golden CO, USA

ABSTRACT

This study explores implementing a machine learning-based system to generate a 3D surface repre- sentation of the roof and support straps in the mine. Event cameras have been chosen for their performance in high-dynamic-range lighting conditions and for their low latency. To enable automated drilling and bolting, 3D vision using event-based cameras has been developed. A ground-truth set is created using two, time-synced event cameras and a LiDAR camera. These sensors are used to construct a ground-truth dataset of corresponding event- camera images and surface maps from the LiDAR. The network is tested with stereo-pairs of event images and produces a depth image with ±5 mm RMS error on average across 1000 test images.

INTRODUCTION

This paper looks to introduce the use of neuro- morphic sensors called event cameras, to perform the 3D perception task [1] of surface representation in the harsh environmental and lighting conditions of underground mines. These cameras offer several advantages over traditional cameras such as low latency, high temporal resolution, and very high dynamic range [2]. This allows for the capture of information in an active mining environment as it is resistant to motion blur due to vibrating sensor platforms, shadows due to single-point source lighting, and dust clouds. However, this sensor is novel enough to not have an associated data set that is both large and diverse [3]. This is an additional hurdle to using commercial vision products as there simply is not enough labeled data to design and train a generalized machine learning model. This problem is usually solved by using simulated data [4]. Nevertheless, simulators frequently lack the representation of visual features commonly present in mining environments, as can be seen in Figure 1. Additionally, the shift from simulation to real- world conditions is recognized to be challenging [5]. The pipeline suggested in this paper serves as an alternative, offering a quick and cost-effective means of generating a labeled data set.

The 3D perception task requires that image points (x, y) in pixels be re-projected into world coordinates (X, Y, Z) in physical distance units. This can be achieved by incorporating depth by using either a depth sensor [6] or a stereo camera pair [7], during operation. To mitigate power and dust concerns associ- ated with active sensors, stereo-vision has been chosen to compute the depth of each pixel.

Stereo vision takes two images taken by cameras that have a known distance offset and outputs a dis- parity map. Disparity is the lateral perspective shift between a pair of corresponding pixels in the left and right images. Since the baseline distance (distance between the cameras) is known, the per-pixel disparity can be projected into depth using epipolar geometry [8] as in Equation 1.

$$d = \frac{b \times f}{z}$$

where,

d = the per-pixel disparity value,
b =the distance between the stereo cameras,
f = the focal length of the cameras used,
z = the depth of the pixel.

This work is supported by National Institute for Occupational Safety & Health | NIOSH/Project 75D30121C12206.

Figure 1. Shows an image of the roof with a bolted support strap, taken at the Edgar Mine in Idaho Springs, CO

Roof bolting is generally considered to be one of the most dangerous jobs in underground mines in the United States [9]–[12]. This is due to two factors - the operator is at risk of being injured by the bolting machine itself [13], and there is a risk of being a casu- alty of a roof fall [14]. Accidents, particularly affecting less-experienced operators, are prevalent. Miner safety and productivity could be greatly improved if the operator was not required to be in situ, identifying areas of the roof to drill, positioning, and operating the drill. Autonomous technologies in the mining industry offer numerous advantages by minimizing workers' exposure to dangerous conditions, increasing safety standards, lowering costs, and enhancing efficiency a[15].

There are several challenges to automating this process. Firstly, the semantic recognition of regions of the roof that are rock versus the support strap. It is critical for the autonomous system to be able to differentiate areas of the roof from straps that are already bolted to the roof. Secondly, the system needs to create a 3D surface representation of the scene during a drilling operation. This is to allow for the roof bolter to localize itself relative to the roof. Both of these tasks, semantic segmentation and 3D reconstruction, need to be per- formed in an active mining environment. This means that the system has to be robust to vibrational loads, dust clouds, and severe lighting conditions. These factors make it difficult to implement off-the-shelf computer vision products to solve these problems. Traditional stereo-vision products exist for stereo pairs of images. However, since this class of algorithms relies on matching corresponding pixels in the images taken by the left and right cameras using feature detectors [16], their performance degrades significantly when used with noisy event-based images.

The alternative is to train a convolutional neural network (CNN) with event-based stereo pairs, labeled with ground-truth depth data, to learn feature match- ing. This approach means that the system learns how to match corresponding points. The pyramid stereo matching network [17] utilizes a 3D CNN and region- scale features, rather than pixel-scale features, to in- corporate global context information. This is essential to predict a disparity map given a stereo pair of noisy, event-based images.

This paper presents a novel approach to solving the 3D reconstruction problem in harsh lighting and environmental conditions. Specifically, the contribu- tion is the pipeline for generating very large training data sets for event-based images during active mining operations.

This paper is laid out with a discussion of the planned approach, a description of the experimen- tal setup, a presentation of the results, a discussion grounding the results in the context of the project, and a conclusion.

METHODS

The goal is to predict depth images given time- synced accumulated event images. To do so, the dataset needs to be comprised of event images captured from a mine, and an associated depth ground-truth image.

Disparity is used as a proxy feature to ensure model modularity. Therefore, a disparity image is computed from the corresponding ground truth depth image us- ing the camera intrinsics of the sensor being used. This disparity is combined with the time-synced stereo- image pair and input to a 3D CNN architecture. To test, a time-synced event stereo pair of images is input to the network. The output is a disparity image, which can then be reprojected into depth, given the focal length of the stereo sensors.

A. Stereo-Disparity

The disparity ground-truth images are computed from the depth images taken by the L515. To do so the horizontal distance between the event cameras needs to be calculated. Once this is known, the disparity is calculated as in (1). Disparity is chosen as the quantity of interest as it is hardware agnostic. This means that depth can be inferred from any disparity given the focal length, as in 1. Therefore, the stereo-disparity training dataset is comprised of time-synced stereo- event pairs and a ground-truth disparity image. When testing the network, a stereo-pair of event images and a disparity image is predicted. This predicted disparity image is projected to depth using Equation 1.

B. Neural Network Architecture

The network used is the stacked hourglass architec- ture as first presented by Chang et al. in pyramid stereo matching network [17]. This particular architecture was chosen for the ability to extract global context information and infer features from noisy event-based images. The architecture is as shown in Figure 3.

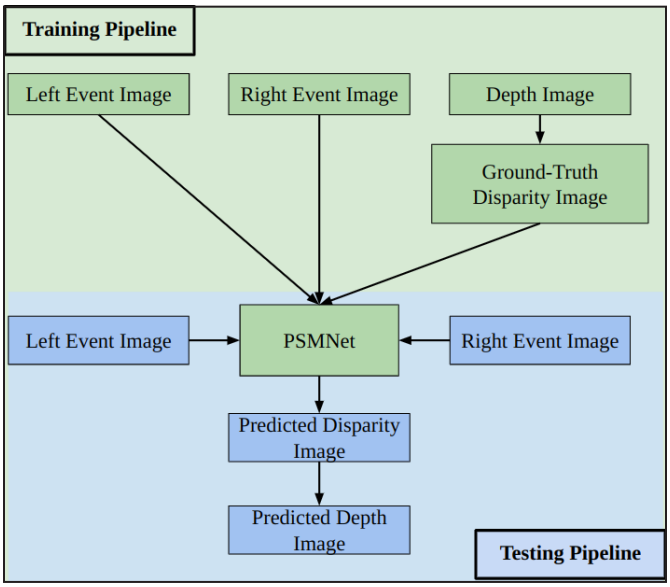

Figure 2. Shows the process for training the network to predict depth images

This work utilizes the stacked hourglass design. This allows for the 3D CNN to regularize cost volume [17].

C. Training the Deep Neural Network

The process described above is used to create a dataset of around 2400 stereo-pairs registered to the corresponding depth image. These are divided into batches of two, and trained for 26 epochs as seen in Figure 4 using a NVIDIA GeForce RTX 3070 graphical processing unit.

Several strategies were used so as to prevent overfit- ting [18]. Firstly an early-stopping strategy was used to stop the training process before validation variance started rising. Secondly, the training and validation set was shuffled at each epoch.

D. Extrinsic and Intrinsic Calibration

The sensor extrinsics are calibrated using computer vision techniques to compute the relative positions of the sensor axes. From this, we can get the offset distance between the two optical axes, the baseline, as well as the mapping between the Lidar and each of the event cameras [7]. OpenCV's library [19] is leveraged to compute the principal point and focal lengths of each of the three sensors. The focal length, f, is used for pose computation and mapping between disparity and depth as in 1.

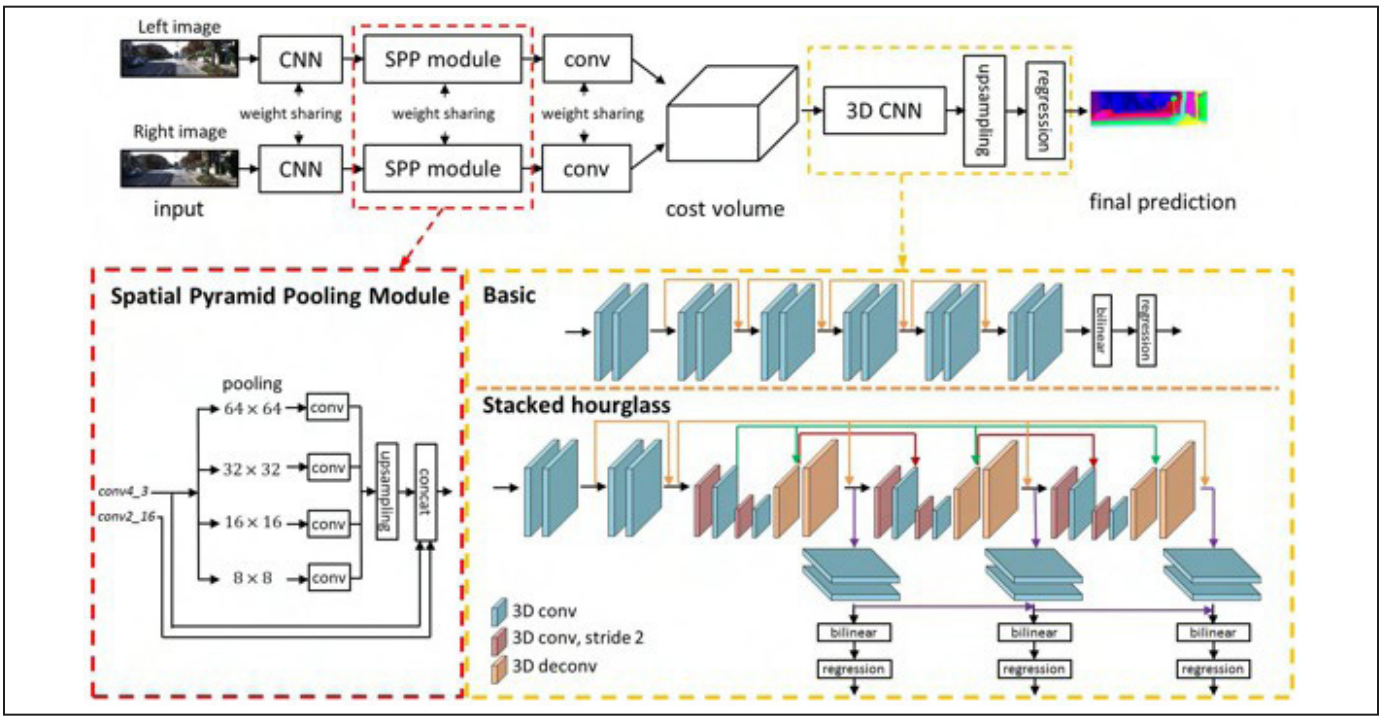

Figure 3. Shows the architecture of the Pyramid Stereo Matching Network [17]

E. Sensor Calibration Setup

The relative position of each sensor is computed with reference to the geometric center of the mobile rig using pose computation using a fiducial marker as in Figure 5.

Homogenous transforms can be chained together to map a depth image from the Lidar to be centered around each of the event camera axes. This is used as the ground-truth depth image, from the perspective of the event cameras, ${}^{E}\mathrm{I}$.

$$ {}^{E}\mathrm{I} = {}^{E}T_{\mathrm{L}}\ {}^{\mathrm{L}}\mathrm{I} \tag{2} $$

where ${}^{\mathrm{L}}\mathrm{I}$ is the depth image taken by the LiDAR, ${}^{L}\mathrm{I}$ is the ground-truth depth image from the perspective of each of the event cameras, and ${}^{E}T_{\mathrm{L}}$ is the transform between the depth camera and each of the event cameras found from extrinsic calibration.

Through this approach, the generation of a diverse, large training data set is automatic. Thus, any super- vised learning approach can be used as the constraint of the complication of labeling a data set is removed.

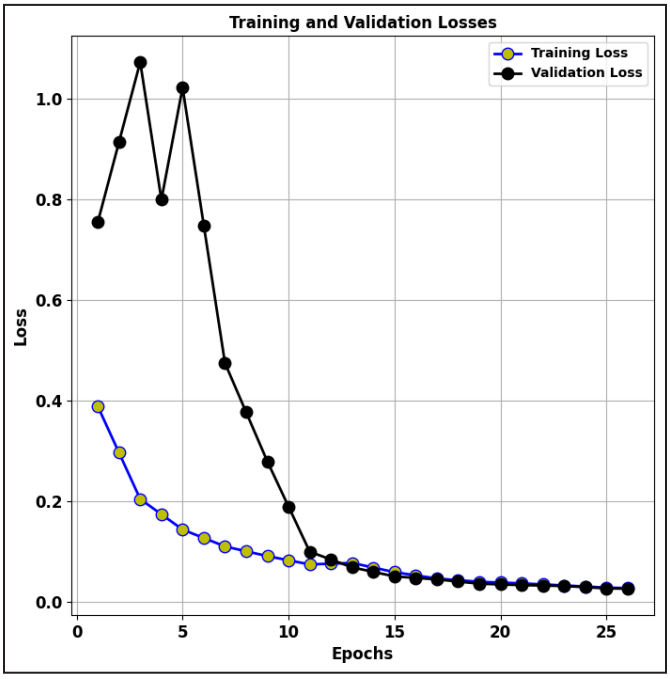

Figure 4. Shows the loss curves generated during the training of the network. An early stopping strategy was utilized so as not to over- train the network

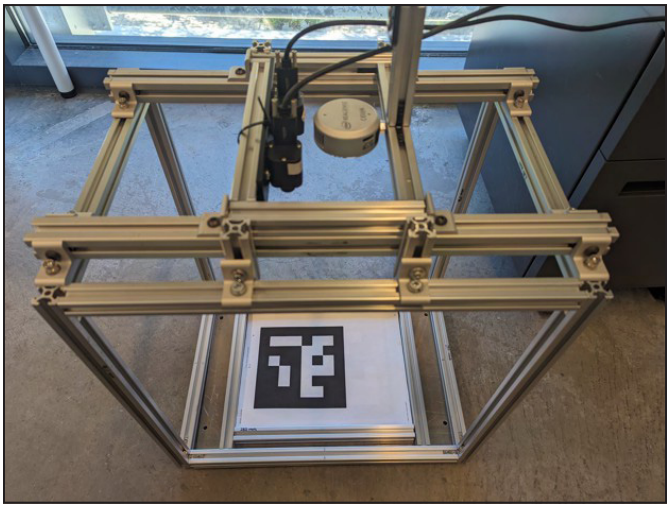

Figure 5. Shows the calibration rig used to extract the relative sensor poses. The fiducial marker can be swapped with a checkerboard to get the intrinsic matrix, K, for each of the sensors

DATA SET COLLECTION

The sensors used are two iniVation DAVIS (Dynamic and Active-pixel Vision Sensor) event cameras and one Intel RealSense L515 (Lidar-Camera). These are mounted rigidly as shown in Figure 7.

We evaluated the entire pipeline by collecting both training and testing data in the Edgar Mine in Idaho Springs, Colorado. The hypothesis is that disparity maps can be predicted at a higher speed and with higher fidelity than when computed with traditional stereo vision techniques, during an active drilling operation. To map the real-time disparities to depth, Equation (1) is used again. Using just the stereo setup, two-time synced images are input to the trained network as shown in Figure 8.

As can be seen, the input images have a significant amount of noise that would disrupt traditional stereo vision [20] that is based on feature matching. When the

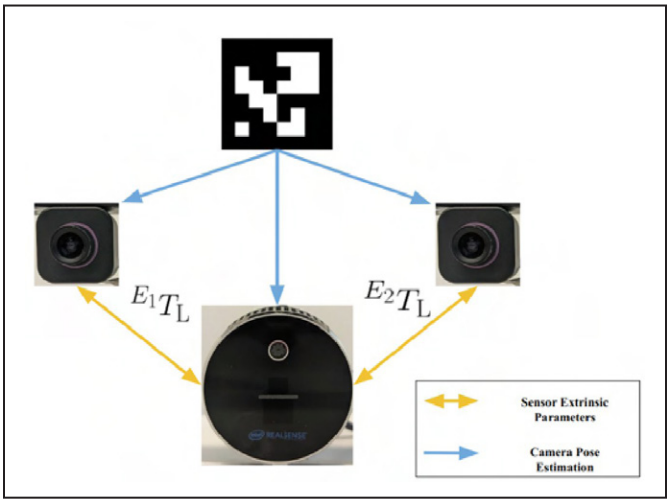

Figure 6. Shows the process used to locate the three sensors relative to each other

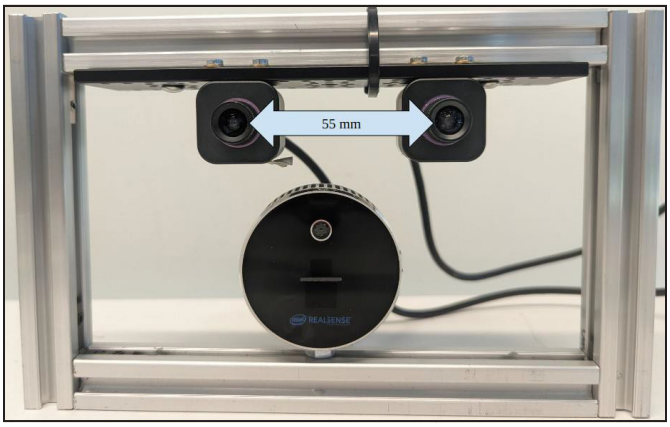

Figure 7. Shows the sensor rig used during mobile data collection operations in the mine. The stereo event cameras are mounted to the top of the rig, while the L515 lidar camera is mounted to the bottom rail

stereo-pair is input to the trained network, the output is a disparity map as shown in Figure 9.

A. Evaluation Metric

To evaluate these predictions, the Frobenius norm of the difference between the prediction and the ground truth image is used.

$$\| e \|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \left| d_{ij} - \widehat{d}_{ij} \right|^2 \right)^{1/2} \quad (3)$$

where d is the ground truth depth and $\widehat{d}$ predicted depth.

These prediction images can be used to re-project images taken by the event cameras into 3D world coordinates. This is used to compute a surface rep- resentation of the roof.

B. Data Sets Collected

Data was collected from four different areas of the Edgar Mine - two of these were reserved for unseen testing, with the LiDAR off to prevent unintentional triggering of the event cameras. These areas already have bolted roofs, with straps and meshes in all align- ments. The lighting conditions are all single-source, point lighting. Stereo pairs of accumulated event-based images are input to the trained network and a predicted depth image is the output. Ten % of the labeled data is reserved as a validation set. For this subset of the data, ground truth is available to evaluate the output of the network. Figure 9, from left to right, shows the right event image, the ground-truth disparity image, and the predicted disparity image.

Two models with the same architecture were instantiated and trained on two different training data sets. The first was a data set from one adit of the mine called the Miami. The distance to the roof was kept within a range of one meter to two meters. The focal length of the event cameras was adjusted appropriately to ac- count for this range and the cameras were intrinsically calibrated afterwards. The features in the data set were horizontal straps and fairly regular gneiss rock. The light source was a single-point headlamp. Ego motion was limited to handheld motion, mimicking a rigid mount to a drill rig.

The second data set was from a different adit called the Army. The distance to the roof was a wide range from one meter to ten meters. Therefore, the images from the data set are not necessarily in focus for the duration of the collection run. The data set is comprised of a variety of features - straps in various orientations, mesh, gneiss rock with discontinuities, and a desk-sized rolling toolbox. The light

Figure 8. Shows the time-synced accumulated event images as taken by the left event camera and right event camera

source was the same as the first set. Ego motion was a wide variety of speeds in each of the six degrees of freedom.

RESULTS

A. Training Set One

When projected to depth, this data set had $||e||_F$ = 5mm ± 1mm. This means that on average the Root Mean Squared Error(RMSE) between the prediction and the ground-truth is 5 mm.

B. Testing Set One

The figures below show predictions from the net- work. The input images are taken in a different area of the mine, i.e., unseen to the network prior to testing. The images are taken under different lighting conditions.

DISCUSSION

This work was able to generate virtually unlimited amounts of training data. Using the PSMNEt-Stereo network, event-based images were used to generate a 3D surface representation of the roof of the mine. Even though each mine can have unique lighting conditions, this pipeline can be used to collect data from the field quickly and cheaply. This data can be immediately added to the training set, without any pre-processing and a quick training regiment can be run.

As can be seen in Figure 9, the features detected by the tightly scoped network, trained on Training Set One, are predicted with high resolution and the right distance scale. The widely generalized model was able to get the right distance scale but was not able to learn the features to the same fidelity. This implies that the fidelity of the model is correlated with the resources devoted to the training of the model.

As Figures 9, 10, and 11 show, the event images can be used to produce a depth image of acceptable fidelity. The straps have holes to allow for bolting. The network can predict the depths of these through holes. Additionally, the outline of the strap is detected with the appropriate 20mm depth difference between the strap and the roof, where ground truth was available. This work shows that the LiDAR sensor perfor- mance can be replicated when the rock is smooth. However, the neural network exceeds the performance of the LiDAR in certain areas of the mine. Specifically, the support straps appear to absorb the wavelength generated [21] by the RealSense L515. The effect of the attenuated return signal is that straps in depth images appear to have the same depth. This introduces a significant measure of noise into the measurement. This is not

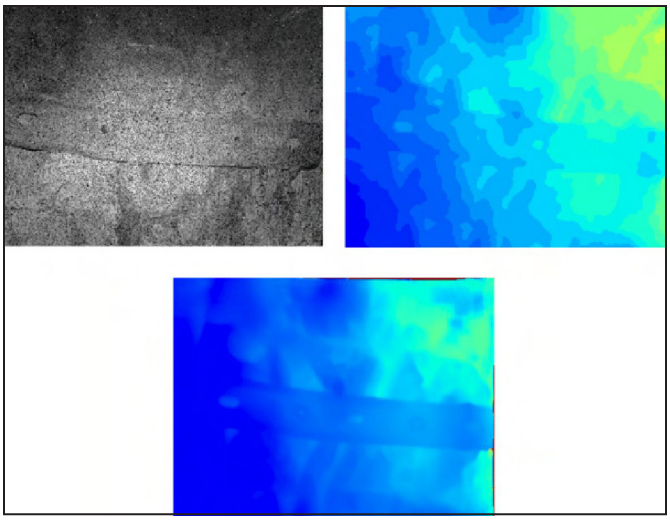

Figure 9. Top Row (Left to Right). Input event based image, LiDAR based depth image. Bottom Row. Predicted depth image. Shows the depth map predicted by the trained network. We can use the computed focal lengths and the baseline distance between the two event cameras to generate the real-time depth map from the perspective of each of the event cameras

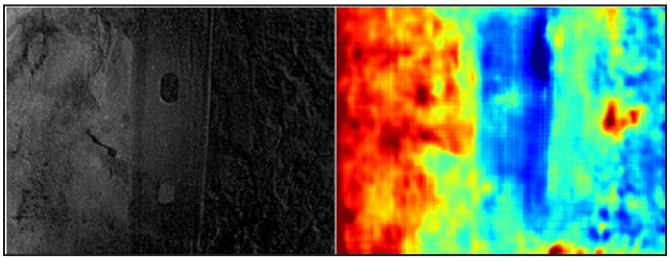

Figure 10. Top Row (Left to Right). Input event based image, Predicted depth image. Shows the prediction of the network from an input event image with the strap partially in a shadow. The 20mm difference in depth between the strap and the roof is captured

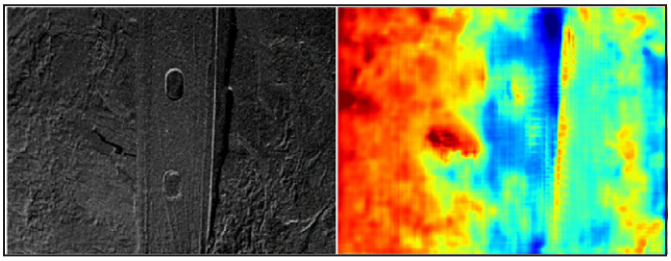

Figure 11. Top Row (Left to Right). Input event based image, Predicted depth image. Shows the prediction of the network from a well-lit input event image. The difference in depth between the holes in the strap and the strap itself is captured

an issue with the event cameras as they are passive sensors. However, it is an issue when trying to baseline the proposed architecture.

In addition to the advantages gained from learned feature matching, a noteworthy benefit is that the cameras do not necessitate stereo-calibration to obtain the essential matrix [22]. The essential matrix is used to encode the relative poses of each camera, aiding in solving the pixel correspondence problem. Tradition- ally, stereo calibration is a crucial step in the process of determining the intrinsic and extrinsic parameters of the two cameras. However, in the context of the learned stereo-vision approach, this calibration step becomes redundant. The method leverages the pixel correspon- dence learned during training, allowing the cameras to perform feature-matching without the need for explicit stereo-calibration. This not only simplifies the overall pipeline but also reduces the complexity associated with the setup and maintenance of stereo-calibrated cameras. By eliminating the stereo-calibration require- ment, the learned stereo-vision approach streamlines the implementation and deployment of multi-camera systems. This reduction in complexity contributes to the practical feasibility and efficiency of deploying vision systems in various applications, particularly in harsh environments like underground mines.

Another network with the same architecture was instantiated separately and trained on a tightly scoped scene with a horizontal strap and constant lighting conditions. This was done to separate the overhead costs of the training from the experimental parts of the pipeline. The output is shown in Figure 9. This performs better than the Lidar as with a higher number of training epochs, all the stochastic noise introduced by the environment is filtered out. This is promising moving forward, as the surface reconstruction problem can be solved using just two event cameras, allowing for resources for machine learning model training [23]. The sensor rig is expected to be mounted onto the drill rig. Therefore, this network can be specialized to a narrow field of view. Since the dataset is representative of the final use case, i.e. on the drill, the performance of this network is severely degraded in depths greater than four meters. However, since the labeling is cheap, the network can be generalized given the appropriate resources.

Looking ahead, this work can be extended to a simultaneous localization and mapping (SLAM) [24] module, offering real-time localization inputs to the roof-bolting machine. This integration presents a prac- tical solution to enhance the autonomy and efficiency of the roof-bolting process by providing accurate and up-to-date information about the machine's position within the mine environment.

Furthermore, the applicability of this approach extends beyond roof bolting, holding significant value for localization and mapping in the context of mobile robots operating in challenging environmental and optical conditions. Autonomous vehicles, for instance, encounter diminished LiDAR performance in adverse weather conditions such as fog and rain [25]. By lever- aging the robustness of event-based images and the learned stereo-vision approach, the SLAM module can potentially overcome these challenges and maintain reliable localization and mapping capabilities even in severe conditions.

This broader application underscores the versatility of the developed methodology and its potential impact on advancing autonomous systems in various domains. The adaptability to challenging environmental factors sets apart the event camera data pipeline as a promising solution for improving the reliability and performance of mobile robots and autonomous vehicles operating in real-world scenarios marked by adverse optical conditions.

In essence, this pipeline offers a scalable and ef- ficient means of continuously updating and improv- ing the performance of learned stereo-vision models through the integration of fresh, real-world data. The combination of quick data collection, minimal pre- processing, and quick training cycles enhances the adaptability and efficacy of the model in addressing the challenges posed by unique and evolving mining conditions.

CONCLUSION

This work has shown the capability to use event cameras and leverage their superior optical properties to create a 3-dimensional representation of the roof during an active mining operation. The contribution is the capability to produce virtually free, labeled data. This can be used in conjunction with any supervised approach. Moreover, a small sensor rig comprising just the two event cameras can be integrated onto a roof bolter. This work has shown that a real-time surface representation can be computed and used as inputs to planning algorithms. This is made easier due to the low power requirements and the superior optical properties of the event cameras themselves.

REFERENCES

[1] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide Scaramuzza. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(1):154–180, 2022.

[2] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea Censi, Stefan Leutenegger, Andrew J Davison, Jörg Conradt, Kostas Daniilidis, et al. Event-based vision: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 44(1):154–180, 2020.

[3] Alex Zihao Zhu, Dinesh Thakur, Tolga Özaslan, Bernd Pfrom- mer, Vijay Kumar, and Kostas Daniilidis. The multivehicle stereo event camera dataset: An event camera dataset for 3d perception. *IEEE Robotics and Automation Letters*, 3(3):2032– 2039, 2018.

[4] Shital Shah, Debadeepta Dey, Chris Lovett, and Ashish Kapoor. Airsim: High-fidelity visual and physical simulation for autonomous vehicles. In *Field and Service Robotics: Results of the 11th International Conference*, pages 621–635. Springer, 2018.

[5] Raghad Alghonaim and Edward Johns. Benchmarking domain randomisation for visual sim-to-real transfer. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12802–12808. IEEE, 2021.

[6] Fumio Itami and Takaharu Yamazaki. A simple calibration procedure for a 2d lidar with respect to a camera. *IEEE Sensors Journal*, 19(17):7553–7564, 2019.

[7] Thao Dang, C. Hoffmann, and C. Stiller. Self-calibration for active automotive stereo vision. In *2006 IEEE Intelligent Vehicles Symposium*, pages 364–369, 2006.

[8] Gang Xu and Zhengyou Zhang. *Epipolar geometry in stereo, motion and object recognition: a unified approach*, volume 6. Springer Science & Business Media, 1996.

[9] JJ Sammarco, A Podlesny, EN Rubinstein, and B Demich. An analysis of roof bolter fatalities and injuries in us mining. *Transactions of Society for Mining, Metallurgy, and Explo- ration, Inc*, 340(1):11, 2016.

[10] Mark Aldrich. Engineers attack the" no. one killer" in coal mining: The bureau of mines and the promotion of roof bolting, 1947–1969. *Technology and culture*, pages 80–118, 2016.

[11] Christopher Mark. The introduction of roof bolting to us underground coal mines (1948–1960): a cautionary tale. 2002.

[12] Ronald C Althouse, Michael J Klishis, and R Larry Grayson. Microanalysis of roof bolter injuries. *Applied occupational and environmental hygiene*, 12(12):851–857, 1997.

[13] Fred C Turin. Human factors analysis of roof bolting hazards in underground coal mines. 1995.

[14] Syd Peng. Roof bolting and underground roof falls. *Geohazard Mechanics*, 1(1):32–37, 2023.

[15] Zeshan Hyder, Keng Siau, and Fiona Nah. Artificial intel- ligence, machine learning, and autonomous technologies in mining industry. *Journal of Database Management (JDM)*, 30(2):67–79, 2019.

[16] Norhayati Mohd Suaib, Mohammad Hamiruce Marhaban, M Iqbal Saripan, and Siti Anom Ahmad. Performance eval- uation of feature detection and feature matching for stereo visual odometry using sift and surf. In *2014 IEEE Region 10 Symposium*, pages 200–203. IEEE, 2014.

[17] Jia-Ren Chang and Yong-Sheng Chen. Pyramid stereo match- ing network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5410–5418, 2018.

[18] Xue Ying. An overview of overfitting and its solutions. In *Journal of physics: Conference series*, volume 1168, page 022022. IOP Publishing, 2019.

[19] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

[20] Lazaros Nalpantidis and Antonios Gasteratos. Stereo vision for robotic applications in the presence of non-ideal lighting conditions. *Image and Vision Computing*, 28(6):940–951, 2010.

[21] Larissa T. Triess, David Peter, Christoph B. Rist, Markus Enzweiler, and J. Marius Zöllner. Cnn-based synthesis of realistic high-resolution lidar data. In *2019 IEEE Intelligent Vehicles Symposium (IV)*, pages 1512–1519, 2019.

[22] Kaveh Fathian, J. Pablo Ramirez-Paredes, Emily A. Doucette,.

[23] J. Willard Curtis, and Nicholas R. Gans. Quest: A quaternion- based approach for camera motion estimation from mini- mal feature points. *IEEE Robotics and Automation Letters*, 3(2):857–864, 2018.

[24] David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. Carbon emissions and large neural network training. *arXiv preprint arXiv:2104.10350*, 2021.

[25] A Alan B Pritsker. *Introduction to Simulation and SLAM II*. Halsted Press, 1984.

[26] Chen Zhang, Zefan Huang, Marcelo H. Ang, and Daniela Rus. Lidar degradation quantification for autonomous driving in rain. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3458–3464, 2021.

24-092

Utilizing Big Data Statistical Techniques in Python to Optimize Geometallurgy Workflow for Metallurgical Test Work Sample Selection

Muhammad Usman Siddiqui
Ausenco, Vancouver, BC

Kevin Erwin
Ausenco, Vancouver, BC

Rajiv Chandramohan
Ausenco, Vancouver, BC

Connor Meinke
Ausenco, Vancouver, BC

Shaihroz Khan
Ausenco, Toronto, ON

ABSTRACT

High-quality sample selection for metallurgical test work is essential to a geometallurgy study, but the large multi-dimensional dataset makes sample selection a daunting task, as classifying the dataset while respecting its heterogeneity is difficult. This paper presents a streamlined approach for sample selection, utilizing custom-built tools in Python to standardize the methodology, saving time and costs. This approach uses the cumulative sum method, principal component analysis, and k-means clustering method to elegantly cluster the data and select representative samples. A case study is used to demonstrate the effectiveness of the methodology by selecting 40 samples for flotation test work.

NOMENCLATURE

PCA—Principal component analysis
CuSum—Cumulative sum
PC—Principal component
WCSS—Within-cluster sum of square

INTRODUCTION

Geometallurgy is a compelling methodology in the mining industry to bridge the gap between metallurgy and geology. It aims to reduce technical risk and enhance the economic performance of a mineral processing plant by accounting for the variability in a deposit to strengthen investor confidence, facilitating robust revenue models, and developing scenarios with variable throughput rates to more accurately forecast cash flows for future mining operations. A geometallurgy study looks at the relationship between a deposit's geological characteristics, its variability, and its response to metallurgical processes. Selecting samples that capture the heterogeneity of the deposit for metallurgical test work is an essential component of a geometallurgy study.

High-quality sampling is vital across the entire mine value chain, as sampling errors are additive and generate monetary and intangible losses (Dominy et al., 2018). The goal is to select samples that accurately describe the deposit (Dominy et al., 2018). A geometallurgy database usually consists of sample id, mineral grade, lithology, alteration, and test work data if available. It is the basis for choosing representative samples for metallurgical characterization test work (competency, hardness, recoveries) which is then used in robust flowsheet development and equipment selection for optimum life of mine performance. However, these datasets can be large with multiple columns, making sample selection a daunting task during the analysis.

Michaux et al. (2020) present a framework to develop a geometallurgical program in which they discuss the methodology to cluster a geometallurgy database into similar clusters, and they present the cumulative sum (CuSum)