Research on 3D Reconstruction Technology of Power Line Based on Image Semantic Segmentation

Xionggang Li¹, Feng Zhang¹, Guoqiang Li¹, Dongwei Xiang²* and Chengcheng Yang²

¹ Guangdong Power Grid Co., Ltd, Machine Operation Center, Guangdong Province, Guangzhou, China;
² Wuhan Huizhuohang Technology Co., Ltd, Hubei Province, Wuhan, China

*Corresponding author email: ssrs_xdw03265@163.com

Abstract. Because it was difficult to distinguish the characteristics of the power lines by the traditional methods of extracting the power lines, which led to the current situation of incomplete reconstruction and a large number of noise in the process of rebuilding the power lines only by the inclined photographing. In this paper, the power line information in the image is segmented pixel by pixel by introducing in-depth learning semantics segmenting neural network. The three-dimensional coordinates of the power line are calculated by the principle of multi-view three-dimensional reconstruction. Finally, the power line is fitted by the catenary equation to complete the incomplete power line reconstruction. The results show that the fitted power line model has high accuracy and meets the requirements of power related applications. Based on the traditional three-dimensional reconstruction, a new idea for power line reconstruction is proposed.

Keywords: Power line; Tilt photogrammetry; Deep learning; Semantic segmentation; Three-dimensional reconstruction; Catenary equation.

1. Introduction

With the development of society, power system has already entered thousands of households, industries and even the whole society. Power system plays an irreplaceable role in the economic development of contemporary society. That is why the development of power network is not only directly related to all aspects of national life, but also to the stable development of national economy, which can more reflect the comprehensive national strength of a country.

The elevated power transmission corridor was an indispensable part of the power network. It was mainly responsible for the transmission of power. At the same time, it was the connection between the power equipment, stations, and users, and played a very important role in the reliable operation of the power network. Therefore, it was urgent to enhance the power supply department's operation and maintenance of the transmission line and construct a digital transmission and maintenance system. With the development of the tilted test and measurement technology, it also provided a new idea for the digital transmission and maintenance system.

This paper presents a three-dimensional reconstruction method of overhead transmission corridor power line based on deep learning semantic segmentation. Firstly, the position of the power line in the image is segmented pixel by pixel through deep learning semantic segmentation. Secondly, the external orientation elements (i.e. the three-dimensional coordinates and pitch rotation angle of the photography center during camera shooting) and the camera internal parameters are obtained through spatial three-dimensional reconstruction to solve the three-dimensional coordinates of the power line.
Finally, the power line is fitted by catenary equation.

2. Key Technologies and Processes

2.1. Overall Technical Process
The whole technical process of this article is shown in Fig.1, mainly including feature point extracting and matching, increasing reconstruction, image power line division and matching.

![Image Power Line Segmentation](image.png)

Figure 1. Three-dimensional power line reconstruction process for overhead transmission corridors based on in-depth learning semantics segmentation.

2.2. Image Data Acquisition
There were three ways to collect the image data: the multiple wing drone, the fixed wing drone, and the helicopter with people. As shown in Fig.2, the advantage of the multiple wing drone was that it had low equipment cost, a variety of flight methods, and the disadvantage was that it had short lifespan and low operating efficiency; The fixed wing drone had the advantages of low equipment cost and long flight time, but the disadvantage was that it only flew in a single way; Some helicopters had the advantage of long-term flight, but the disadvantage was the poor quality of their images.

![Multi-Rotor UAV](image.png) ![Fixed-wing UAV](image.png) ![Helicopter](image.png)

Figure 2. Three methods of image data collection.

2.3. Feature Point Extracting and Matching
In this paper, feature point extraction is based on Scale-invariant feature transformation (SIFT) algorithm, which is robust to light, angle of view fine-tuning, noise and other changes, and has a high detection rate for partially occluded objects. On this basis, local descriptors are introduced to weaken the background image and improve the image recognition highlight rate. The Gaussian kernel function has been proved to be the only linear kernel function by Koenderink, Lindeberg.

2.4. Incremental Reconstruction
Incremental reconstruction starts with filtering the images and choosing the better ones for initial reconstruction. The initial selected images correspond to a wider photographic baseline with a higher heading and side-by-overlap ratio. Then the five-point algorithm for the complete pinhole model is used for relative orientation, and the beam adjustment method is used to optimize the minimum error non-linearly. Then the rest of the images are registered in the system, and the structure and camera parameters are continuously optimized until all the images are registered in the system. The beam adjustment optimization runs through the incremental reconstruction process, and the exterior orientation elements and camera internal parameters are finally solved.
2.5. Image Power Line Semantic Segmentation

Currently, semantic segmentation model algorithms mainly include FCN\textsuperscript{[7]}, SegNet\textsuperscript{[8]}, Deconvnet\textsuperscript{[9]}, U-net\textsuperscript{[10]}, DeepLab\textsuperscript{[11]}. RefineNet\textsuperscript{[12]}, PSPNet\textsuperscript{[13]}, GCN\textsuperscript{[14]}, DeepLabV3 ASPP\textsuperscript{[15]}, GAN\textsuperscript{[16]}, Mask R-CNN\textsuperscript{[17]}, etc. The FCN has poor robustness to image details and does not take into account the relationship between the pixel and the nearest pixel. SegNet has neighboring pixel information ignored when pooling feature maps of low-resolution images. Deconvnet cannot process the image details; U-net because each convolution will reduce two pixels, resulting in different detection results and the size of the detected object; DeepLab only detects one-eighth the size of the original input image due to network structure. RefineNet is unable to extend the network structure due to the decoder module. PSPNet has relatively strict requirements for detail handling; GCNs require a long time to detect due to the complexity of calculation due to the large number of structural parameters. DeepLabV3 ASPP can't capture a wide range of image information; GAN does not have a perfect mechanism of full and semi-supervision. Mask-RCNN introduces a new method of region feature clustering to improve the accuracy of pixel segmentation.

In this paper, based on the detectron2 framework platform, due to the particularity of power lines in the segmented image, rebuilding the power line model needs a high accuracy of pixel-by-pixel segmentation. At the same time, analyzing the deficiencies of the current semantic segmentation algorithm model, the Mask-RCNN algorithm model is finally selected, which is mainly an example segmentation to complete target detection, semantic segmentation, and so on. Many complex machine vision tasks such as scene segmentation.

In this paper, based on the detectron2 framework platform, due to the particularity of power lines in the segmented image, rebuilding the power line model needs a high accuracy of pixel-by-pixel segmentation. At the same time, analyzing the deficiencies of the current semantic segmentation algorithm model, the Mask-RCNN algorithm model is finally selected, which is mainly an example segmentation to complete target detection, semantic segmentation, and so on. Many complex machine vision tasks such as scene segmentation.

In this paper, based on the detectron2 framework platform, due to the particularity of power lines in the segmented image, rebuilding the power line model needs a high accuracy of pixel-by-pixel segmentation. At the same time, analyzing the deficiencies of the current semantic segmentation algorithm model, the Mask-RCNN algorithm model is finally selected, which is mainly an example segmentation to complete target detection, semantic segmentation, and so on. Many complex machine vision tasks such as scene segmentation.

\[
\frac{\partial L}{\partial x_i} = \sum_{m} \sum_{n} \left[ d(i, j^*(m, n)) < 1 \right] \left( 1 - \Delta h \right) \left( 1 - \Delta w \right) \frac{\partial L}{\partial y_{mn}}
\]

\( d \) is the distance between two points, \( \Delta h \) and \( \Delta w \) are the difference between \( x_i \) and \( j^*(m, n) \). It can be seen from the formula that each point whose aspect coordinate is less than 1 than \( j^*(m, n) \) must accept the gradient value returned by the corresponding \( y_{mn} \) value.

2.6. Fitting Power Lines

In power transmission system, power lines are suspended on the insulation of adjacent power towers. Overhead line suspension equation is usually introduced in engineering design to predict the shape of power lines between adjacent poles and towers, so as to simulate and calculate the rationality of Engineering design. This equation expresses the trend of overhead transmission lines with specific load and stress. Because the height of adjacent poles and towers is not always at the same elevation when the power network is built during actual construction, the elevation of suspensions on both sides of the power line is different. By considering the relevant factors, the suspension equation is as follows:

\[
y = \frac{h}{\sigma_0} \sqrt{\frac{2\delta_y}{(y-x)^2}} \frac{r(1-x)}{2\delta_y} - \frac{1}{\sqrt{1 + \left( \frac{h}{\sigma_0} \right)^2 \frac{2\delta_y}{(y-x)^2} \frac{r(1-x)}{2\delta_y}}} \]

\( \delta_y \) is the distance between two points, \( \Delta h \) and \( \Delta w \) are the difference between \( x_i \) and \( j^*(m, n) \). It can be seen from the formula that each point whose aspect coordinate is less than 1 than \( j^*(m, n) \) must accept the gradient value returned by the corresponding \( y_{mn} \) value.
3. Experiments and Result Analysis

3.1. Making In-depth Learning Datasets
Because image semantics segmentation is different from general target detection, when making a dataset, the object to be segmented is labeled with a few pixels on the image, usually using the image labeling tool Labelme, as shown in the figure below.

(a) Labeled image data (b) Mark the image Json information

Figure 3. Samples of training data.

3.2. Target Segmentation Results

3.2.1. Experimental environment

Table 1. Experimental variables.

| Variable          | Parameter                        |
|-------------------|----------------------------------|
| OS                | Ubuntu 16.04                     |
| CPU               | Intel(R) Core(TM) i7-7700 3.6GHZ |
| GPU               | NVIDIA GTX 1080                  |
| Python            | 3.5.8                            |
| Visual Studio     | Visual Studio Enterprise 2017    |
| cmake             | 2.8.7                            |
| cuda              | 10.1                             |
| cuDNN             | 7.6.5                            |
| OpenCV            | 3.4.10                           |

3.2.2. Segmentation results and accuracy. With 50 pre-training sample data, the result of power line segmentation is shown in Fig. 4 below. The detection data were collected from 100 images collected by fixed-wing and multi-rotator UAVs. The data were collected from different voltage levels of main and distribution networks in a region of Guangdong. The final detection results are shown in Table 2 below.
Figure 4. Effect diagram of image semantically segmenting power lines.

Table 2. Accuracy and segmentation time of semantically segmented images with different parameter configurations for Mask-RCNN networks.

| Parameter Configuration   | Average accuracy(mAP) | Time per slice |
|--------------------------|-----------------------|----------------|
| mask_rcnn R_50_c4_3x     | 85.0%                 | 0.0263s        |
| mask_rcnn R_50_DC5_3x    | 85.3%                 | 0.0258s        |
| mask_rcnn R_50_FPN_3x    | 86.5%                 | 0.0276s        |
| mask_rcnn R_101_DC5_3x   | 87.0%                 | 0.0363s        |

As can be seen from Table 2 above, the accuracy of segmentation using different Mask-RCNN parameter configurations is generally above 85%, which lays a solid foundation for subsequent three-dimensional power line reconstruction.

3.3. Power Line 3D Reconstruction

The external orientation elements and camera internal parameters are obtained by incremental reconstruction, and the power line is semantically segmented pixel by pixel. Then the three-dimensional coordinates of each pixel are calculated according to the principle of multi-view three-dimensional reconstruction. Finally, the complete power line is fitted according to the catenary equation, and the incomplete power line reconstruction is completed.

The experimental data is a distribution network line in Shaoguan City, Guangdong Province. The data collection uses a single-channel flight mode with a flight altitude of 100m relative to the ground, a flight speed of 7m/s and a gap of 60m. The reconstruction results of Pix4Dmapper and Xinjiang Smart Map using the three-dimensional modeling software on the market are shown in Fig. 5 below. In this paper, the effect of power line three-dimensional reconstruction by introducing in-depth learning semantics segmentation method is shown in Fig. 6 below.

Figure 5. Effects of other software modeling.
Figure 6. Effect diagram of the ideas in this paper.

The different colored balls in Fig.6 are the key points for fitting power lines based on the line information split by deep learning in this paper.

Figure 7. Effect diagram of single-strand traverse fitting.

Fig.7 above shows the effect of importing the fitted wires into Google Earth to show single-phase multiple-strand wires. As shown in the figure above, after deep learning semantics segmentation, single-phase multiple-strand power lines can be fitted completely according to the characteristics of the segmented points.

From the diagram, it can be seen that this paper uses in-depth learning semantics segmentation method to fit the incomplete problem of power line completion by using mathematical model alone in three-dimensional reconstruction, which provides a better data model for the later power related application analysis.

To verify the accuracy of reconstructed power lines, high-precision radar is used for data collection at the same interval. Table.3 shows the error statistics (expressed as horizontal direction and expressed as elevation equation) of the modeling results of two base distribution pole towers with a spacing of 60m in phase A and phase B traverses.

Table 3. Comparison of reconstructed power lines with measured points.

| phase | Measured traverse points | Interpolate traverse points | Offset |
|-------|--------------------------|-----------------------------|--------|
|       | X/m   | Z/m   | X/m   | Z/m   | dX/m  | dZ/m  |
| A     | -5495.066 | 58.456 | -5495.144 | 58.106 | 0.078  | 0.35  |
|       | -5495.875 | 58.185 | -5495.862 | 58.275 | -0.013 | -0.09 |
| B     | -5485.966 | 58.561 | -5485.973 | 58.111 | 0.007  | 0.45  |
|       | -5485.305 | 58.362 | -5485.311 | 58.312 | 0.006  | 0.05  |

It was known from table 3 that the maximum horizontal deviation was only 0.078m, and the maximum height deviation was only 0.45m. The deviation between the two directions met the requirements of electric power related applications. At the same time, it showed that the method of rebuilding the
electric power line through fitting the electric power line had a high precision, which provided a possibility for the following self-driving inspection of the transmission line.

4. Conclusion
In this paper, the oblique photogrammetry technology is used as the theoretical basis, through the introduction of in-depth learning Mask-RCNN neural network model to semantically segment power lines, in the modeling process, through the relevant mathematical models to fit the complete power lines, to solve the current image data occupies fewer pixels and cannot be completely reconstructed by oblique photogrammetry technology. The results show that the power line with incomplete three-dimensional reconstruction by fitting and completing has high accuracy of location spatial information, has certain theoretical significance and practical value, and lays a good foundation for subsequent auto-driving inspection along the line. However, the current in-depth learning semantics segmentation Mask-RCNN model is expected to further improve the efficiency and accuracy of segmentation. The next step will focus on improving the Mask-RCNN network to improve the accuracy and efficiency of segmentation.

Acknowledgements
This work is sponsored by the following funds: supported by the science and technology project of Guangdong Power Grid Co., Ltd. (Project No.: GDKJXM20184737)

References
[1] H Hirschmüller. Stereo Processing by Semiglobal Matching and Mutual Information[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 30.
[2] Ng P C, Henikoff S. SIFT: predicting amino acid changes that affect protein function[J]. Nucleic Acids Research, 2003, 31(13):3812-3814.
[3] Koenderink, J. J.. (1984). The structure of images. Biological Cybernetics, 50(5), 363-370.
[4] Lindeberg, Tony. Scale-space theory: A basic tool for analysing structures at different scales[J]. Journal of Applied Statistics, 1994, 21(2):225-270.
[5] Tian, Z., & BC Li. (2012). Optimize preview model parameters evaluation of ransac. Applied Mechanics & Materials, 687-691, 3984-3987.
[6] Shi, C., & Xie, D. L.. (2007). Design of an automatic system based on four-quadrant detector for beam-adjustment. Electronic Instrumentation Customer.
[7] Evan Shelhamer, Jonathan Long, Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation[M]. IEEE Computer Society, 2017.
[8] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017:1-1.
[9] Noh H, Hong S, Han B. [IEEE 2015 IEEE International Conference on Computer Vision (ICCV) - Santiago, Chile (2015.12.7-2015.12.13)] 2015 IEEE International Conference on Computer Vision (ICCV)- Learning Deconvolution Network for Semantic Segmentation[J]. 2015: 1520-1528.
[10] Eicken T V, Basu A, Buch V, et al. U-Net: a user-level network interface for parallel and distributed computing (includes URL)[J]. Acm Sigops Operating Systems Review, 1995, 29(5):40-53.
[11] Niu, Zijia, Jiang,. DeepLab-Based Spatial Feature Extraction for Hyperspectral Image Classification[J]. IEEE Geoscience and Remote Sensing Letters, 2019.
[12] Rajaram R N, Ohn-Bar E, Trivedi M M. RefineNet: Iterative refinement for accurate object localization[C]// IEEE International Conference on Intelligent Transportation Systems. IEEE, 2016.
[13] Zhao H, Shi J, Qi X, et al. Pyramid Scene Parsing Network[J]. 2016.
[14] Li J, Zhang T, Tian H, et al. SGCN: A Graph Sparsifier Based on Graph Convolutional Networks[J]. 2020.
[15] Chen L C, Papandreou G, Schroff F, et al. Rethinking Atrous Convolution for Semantic Image Segmentation[J]. 2017.

[16] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks[J]. Advances in Neural Information Processing Systems, 2014, 3:2672-2680.

[17] Kaiming H, Georgia G, Piotr D, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, PP:1-1.