BAE SYSTEMS

Intelligent Photogrammetry for Digital Elevation Model Production

Technical Brief

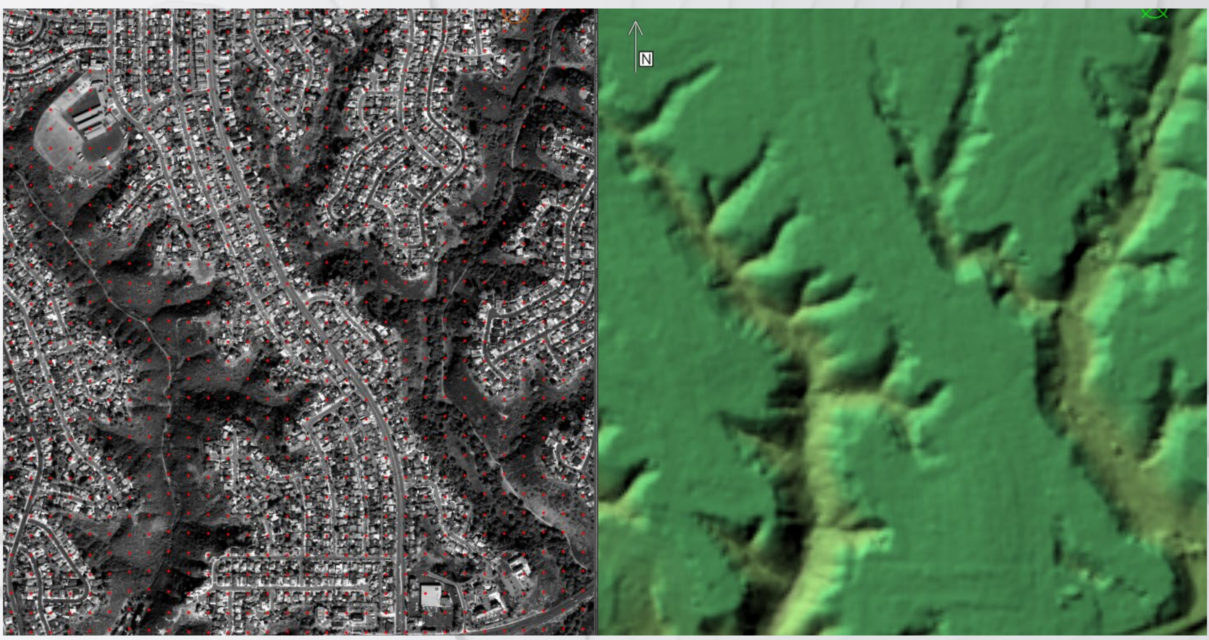


Image © 2021 Maxar Technologies.

GXP®
Geospatial eXploitation Products™

Geospatial solutions
to ensure a safer world.

1. Executive summary

Digital Elevation Model (DEM) production is one of the most time consuming tasks in digital photogrammetry. By applying machine learning to Digital Photogrammetry, BAE Systems' Intelligent Photogrammetry can significantly reduce the cost of DEM production from Digital Surface Models (DSM), which are generated from satellite images, aerial images, Unmanned Aerial Vehicle (UAV) images, sparse LiDAR 3-D point clouds, and dense LiDAR 3-D point clouds. There are various types of DSM, each containing different post spacing and accuracy. The following sets of 3-D models have been trained based on the different DSM types:

1. 3DLargeBuildingModel
2. 3DBuildingModel
3. 3DHouseModel
4. 3DTreeModel
5. 3DGroundPointModel

The first four models detect above ground 3-D objects and then remove them from DSM to generate DEM. The last model classifies 3-D points into thirteen categories, which are then used to generate DEM in difficult terrain such as dense forestry areas, where the ground is mostly unseen.

The main cost of DEM production using DSM generated from satellite images in difficult terrain is the transformation from DSM to DEM. Traditional handcrafted bare earth algorithms for DSM to DEM transformation cannot deal with many different cases for general purpose application and big data. Intelligent Photogrammetry can handle different cases by adding training samples. For this case study, the city of San Diego was used to generate DEM from Intelligent Photogrammetry to achieve Root Mean Square Error (RMSE) of 0.95 meters from stereo satellite images. This case study indicates that Intelligent Photogrammetry can reduce the DEM production cost by more than 50%.

The most time consuming component of DEM production is dense forestry areas, and in this case study, the forestry height is up to 19 meters causing the ground to be nearly invisible. This issue was resolved with 3DGroundPointModel based on machine learning, achieving RMSE 2.40 meter and meeting the desired DEM accuracy requirement of 2 to 3 meters using stereo satellite images.

DEM production from UAV images using our Intelligent Photogrammetry can achieve state-of-the-art accuracy. The Intelligent Photogrammetry can identify errors in DSM generated from UAV images and correct them; therefore, providing a very competitive DEM generation capability for UAV images.

2. Intelligent Photogrammetry

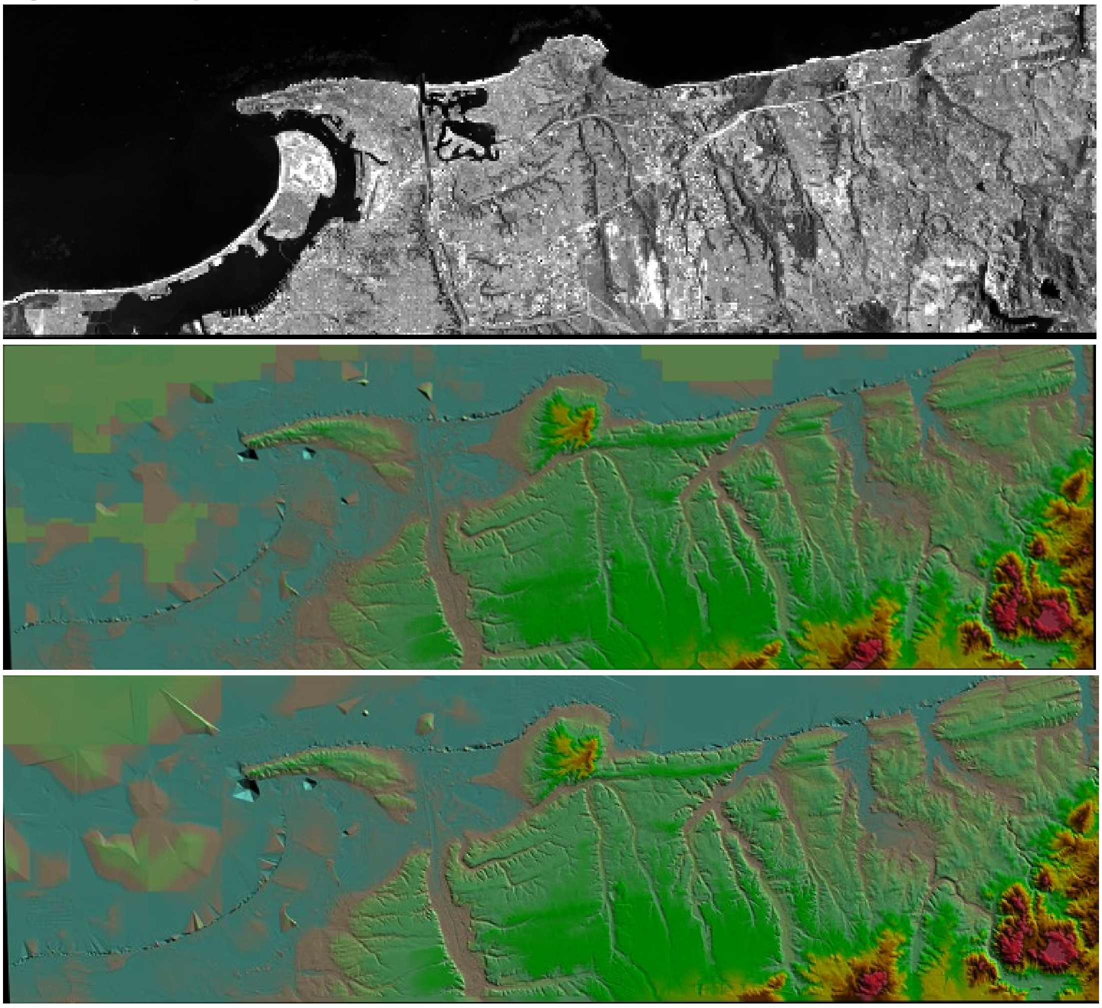
Deep learning Convolution Neuron Network (CNN) is applied to advance the five major automation tasks in Digital Photogrammetry:

1. Tie point extraction and matching
2. DSM generation
3. DEM generation
4. Automatic Feature Extraction (AFE)
5. Image segmentation

In the last six years, we have developed DeepObject™, which uses deep learning to detect objects from imagery and 3-D point clouds (Zhang, 2016; Zhang 2017a; Zhang 2017b; Zhang & Hammoud, 2019). One of the innovations in DeepObject is that the positional accuracy is sub-pixel; meaning, the positional accuracy allows detected objects such as corners of manufactured structures to be used as tie points for triangulation. Removing the detected vehicles, houses, buildings, vegetation, trees, and forestry from the DSM allows for DEM generation. SOCET GXP® AFE works accurately when the input is LiDAR 3-D point clouds. DeepObject is a natural extension to AFE and can detect 2-D objects from imagery with close to human-level accuracy. The Intelligent Photogrammetry uses a 3DGroundPointModel to classify 3-D points into thirteen categories, which are then used to generate DEM in dense forestry areas. Applying deep learning to photogrammetry transforms Digital Photogrammetry into Intelligent Photogrammetry.

Applying the five DeepObject models sequentially and to a DSM of 1,141,346,234 posts, 1 meter post spacing, covering the city of San Diego (Figure 1), an RMSE of 0.95 meter (Table 1) has been achieved.

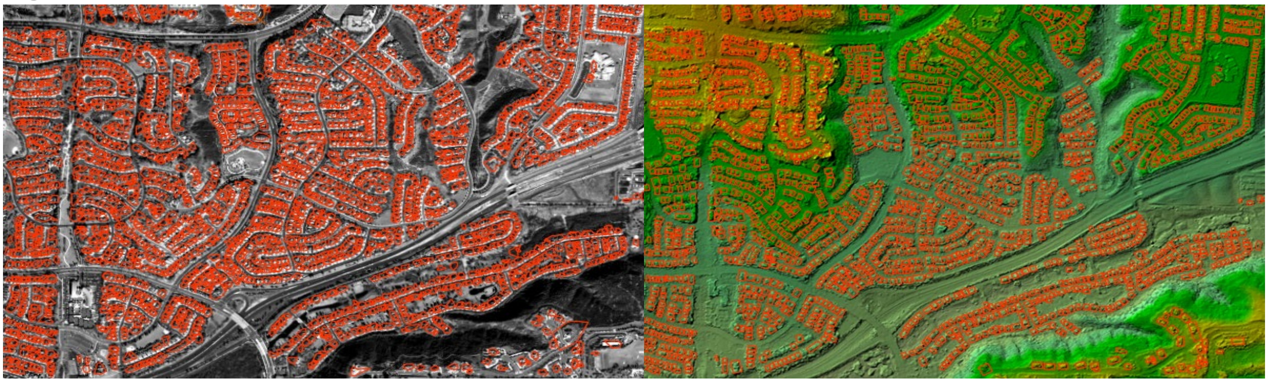
Figure 1. Study area for DSM and DEM.



Terrain Shaded Relief (TSR) image of DSM extracted using SOCET GXP ASM with stereo satellite images (WorldView 50cm images). The top image is the left image of the stereo pairs. The middle image is the TSR of DSM with 1 meter post spacing. The bottom image is the TSR of DEM, which is generated by our Intelligent Photogrammetry. The total area is over 1,000 km^2 and the land area is over 750km^2. This is a difficult area for DEM production because the terrain is not flat and there are lots of houses, trees, and buildings. There are also a few densely forested areas in this terrain. Image © 2021 Maxar Technologies.

An accurate 3-D model is needed to remove 3-D objects, such as a house, from DSM. The 3DHouseModel has achieved 99% precision and 94% recall with only 2,980 positive training samples and 1,827 negative training samples (Figure 2). All training samples are collected from DSM of 1 meter post spacing.

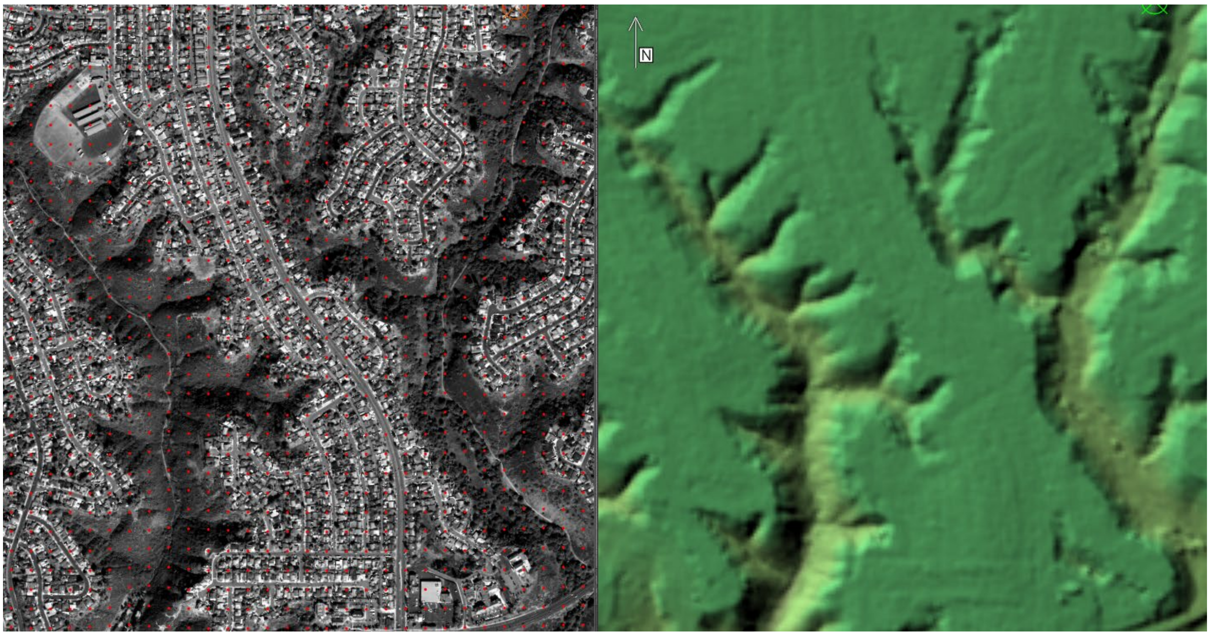
Figure 2. 3-D houses detected from DSM.



There are 2,227 two-story houses in this area (red polygons). The left image is one of the stereo pairs with Ground Sample Distance (GSD) of 0.5 meter. The right image is the TSR of DSM generated from a stereo pair of WorldView-2 images. The number of true positive is 2,105. The number of false positive is 16. The number of false negative is 122. We removed those detected houses from DSM to generate DEM automatically. Image © 2021 Maxar Technologies.

We compute the elevation differences of 1,007 manually edited ground posts against the elevation from the automatically generated DEM to evaluate the accuracy (Figure 3). The 1,007 manually edited ground posts have a post spacing of 50 meters, whereas the RMSE of the DEMO is 0.95 meters. At every 50 meter location, we manually edited a post to the ground. Using 3-D stereo view, posts such as tops of houses, trees, or buildings, were manually edited to the ground based on nearest ground posts and human intelligence. Because this is a very time consuming process, automatic DEM generation is the top priority of our Intelligent Photogrammetry.

Figure 3. Accuracy of DEM.



The left image has 1,007 manually edited ground posts (red dots) with a spacing of 50 meters. The right image is the TSR of DEM. The RMSE computed from the elevation differences of these 1,007 ground posts vs. the DEM is 0.95 meters. Image © 2021 Maxar Technologies.

To validate Intelligent Photogrammetry based DEM generation vs. other DEM generation algorithms, we used the same DSM, the same 1,007 ground posts, and generated an accuracy comparison table (Table 1). We evaluated a handcrafted, third-party DEM generation software which has an RMSE of 2.55 meters. SOCET GXP NGATE and ASM can automatically generate DEM from DSM with a DEM accuracy of 1.86 meters. Intelligent Photogrammetry based DEM generation is much more accurate than the third-party software (Table 1).

Table 1. Accuracy comparison of automatic DEM generation from stereo satellite images.

	Third-party DEM	NGATE/ASM DEM	Intelligent Photogrammetry DEM
RMSE (meter)	2.55	1.86	0.95

Intelligent Photogrammetry DEM can significantly reduce the DEM production cost from satellite stereo images, aerial stereo images, UAV stereo images, sparse LiDAR 3-D point clouds, and dense LiDAR 3-D point clouds. DEM generation from stereo satellite images is more difficult than the other sources.

3. 3DGroundPoint Model

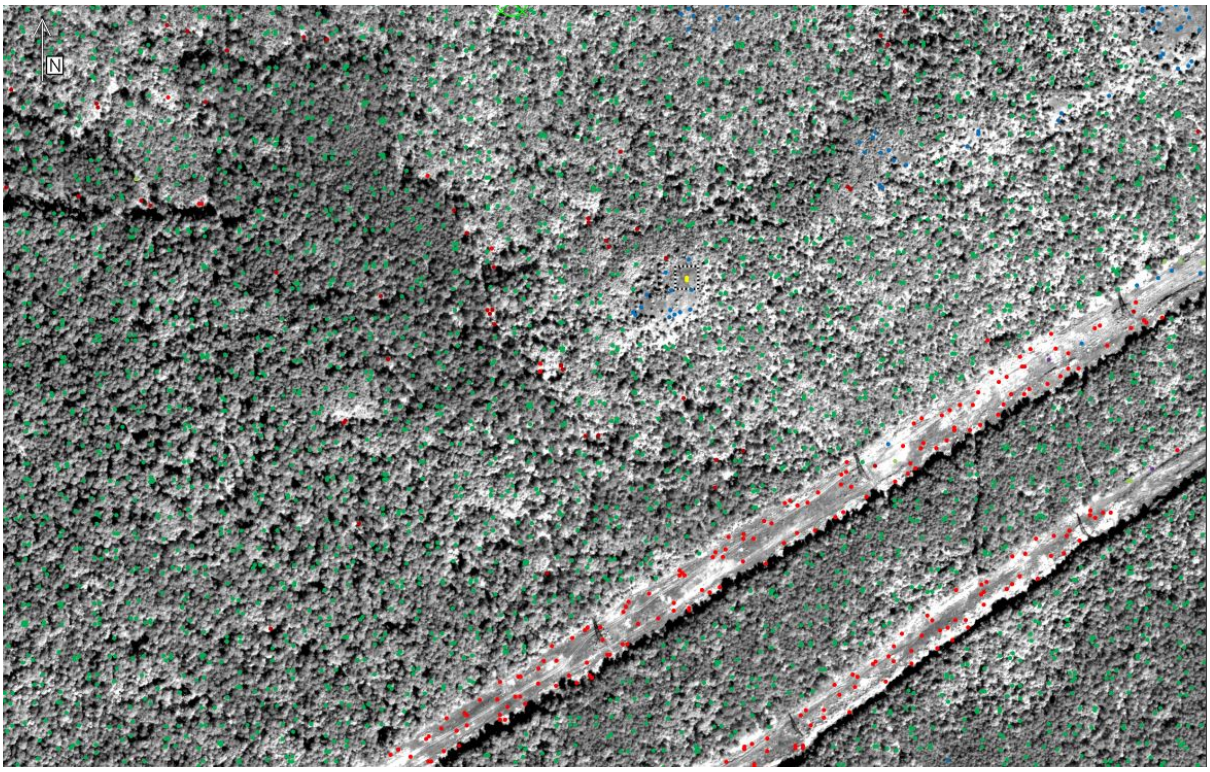
The main application of 3DGroundPointModel is to automatically achieve accurate DEM generation in dense forestry areas. 3-D object detection models (3DLargeBuildingModel, 3DBuildingModel, 3DHouseModel, 3DTreeModel) can only detect 3-D objects with specific shapes and dimensions. Dense forestry areas do not have certain shapes and dimensions causing the need for a classification model: 3DGroundPointModel. In this dense forestry area, 3-D DSM posts and 3-D point clouds are classified into thirteen categories; eight of the categories are displayed with different colors (Figure 4). We use a Triangulated Irregular Network (TIN) to model a forestry height map for each forestry post based on the classification. For example, posts of categories Ground_Low, Ground_High, Ground_Edge, Ground_Ridge, Ground_Valley, and Ground_Slope have a forestry height of 0. For each Ground_Low or Ground_Edge post, we find its closest Aboveground_Forestry post and use their difference in Z to determine the forestry height.

The following equation was used to compute the DEM:

$$DEM = DSM - Forestry\_Height\_TIN$$

Where Forestry_Height_TIN is determined based on the classification of 3-D points/posts from the 3DGroundPointModel.

Figure 4. 3DGroundPointModel.

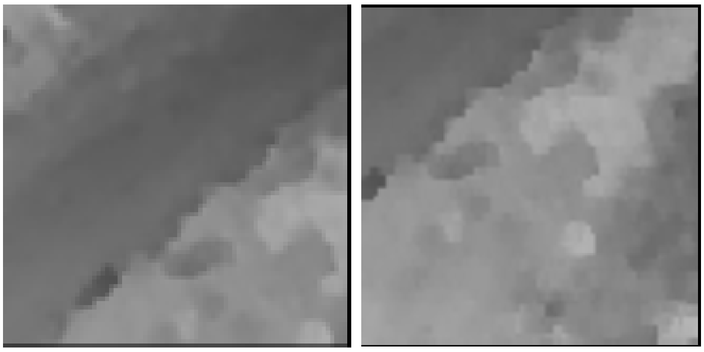


Ground_Low: ● Ground_High: ● Ground_Edge: ● Ground_Ridge: ● Ground_Valley: ● Ground_Slope: ● Aboveground_Forestry: ● Aboveground_Tree: ●

Most of the points in this dense forestry area are above ground, indicating that the 3DGroundPointModel accurately classified 3-D points in a DSM. Image © 2021 Maxar Technologies.

The training samples are collected from DSM, not images (Figure 5). The middle point of the left image is a training sample of Ground_Edge, which is on the ground and at the edge of a forestry area. The middle point of the right image is a training sample of Aboveground_Forestry. We have developed algorithms to transfer a patch of DSM into an image chip (Zhang, 2017a).

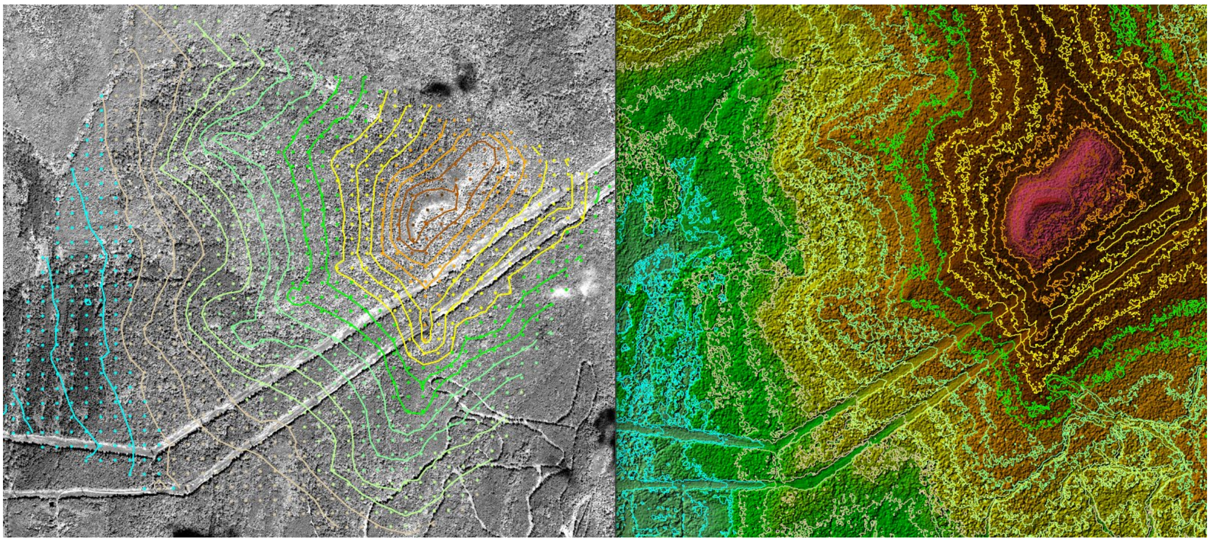
Figure 5. Training samples for 3DGroundPointModel.



Left image chip is a training sample of Ground_Edge and right image chip is a training sample of Aboveground_Forestry.

To evaluate DEM accuracy generated by the 3DGroundPointModel, we have manually measured 1,067 ground truth posts with a post spacing of 50 meters (Figure 6). The forestry is so dense that most of the ground is not visible from overhead.

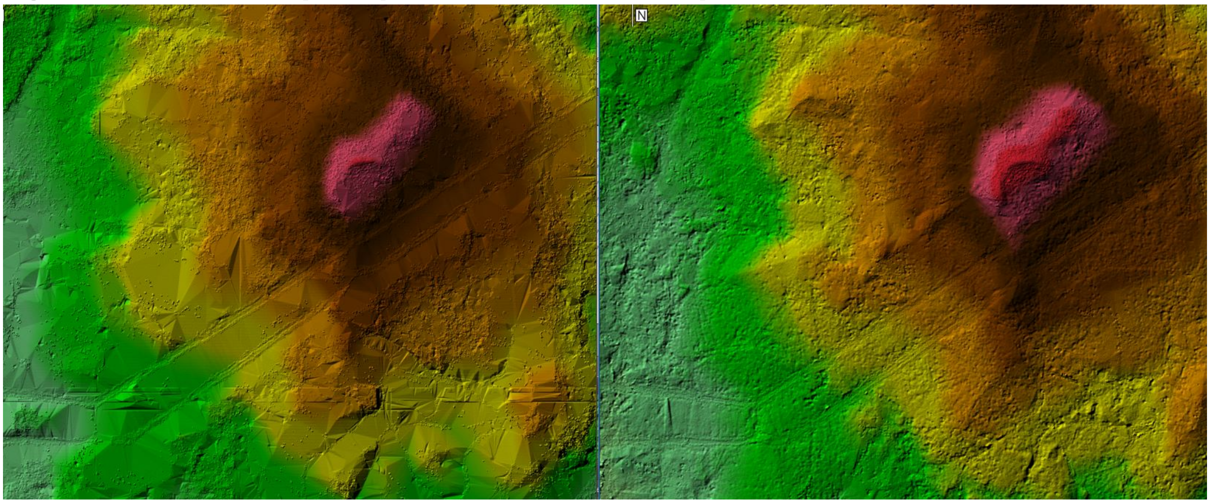
Figure 6. DEM in a dense forestry area.



Left image shows a dense forestry area with forestry height up to 19 meters. There are 1,067 manually edited ground posts with 50 meters post spacing as shown in the left image. The contours are of 10 meter interval. The right TSR image shows the DSM generated by SOCET GXP ASM with contours of 10 meter interval. Image © 2021 Maxar Technologies.

The difference in DEM accuracy between Digital Photogrammetry and Intelligent Photogrammetry in dense forestry areas shows great improvement (Figure 7). The left TSR image is Digital Photogrammetry showing DEM generated by NGATE/ASM. The right TSR image is Intelligent Photogrammetry showing DEM generated by the 3DGroundPointModel. DEM from NGATE/ASM has an RMSE of 4.83 meters, while DEM from 3DGroundPointModel has RMSE of 2.40 meters, which meets the desired accuracy requirement of 2 to 3 meters (Table 2).

Figure 7. DEM accuracy comparison in dense forestry areas.



The left TSR image shows the DEM generated by ASM. The right TSR image shows the DEM generated by the 3DGroundPointModel. As shown in Table 2, the DEM accuracy from 3DGroundPointModel is much more accurate (2.40 meters vs. 4.83 meters). Desired requirement of DEM accuracy in dense forestry area is 2 to 3 meters.

Table 2. DEM accuracy in dense forestry area from satellite images.

	NGATE/ASM DEM	Intelligent Photogrammetry DEM
RMSE (meter)	4.83	2.40
Standard deviation (meter)	3.65	2.23
Bias (meter)	-3.17	-0.70

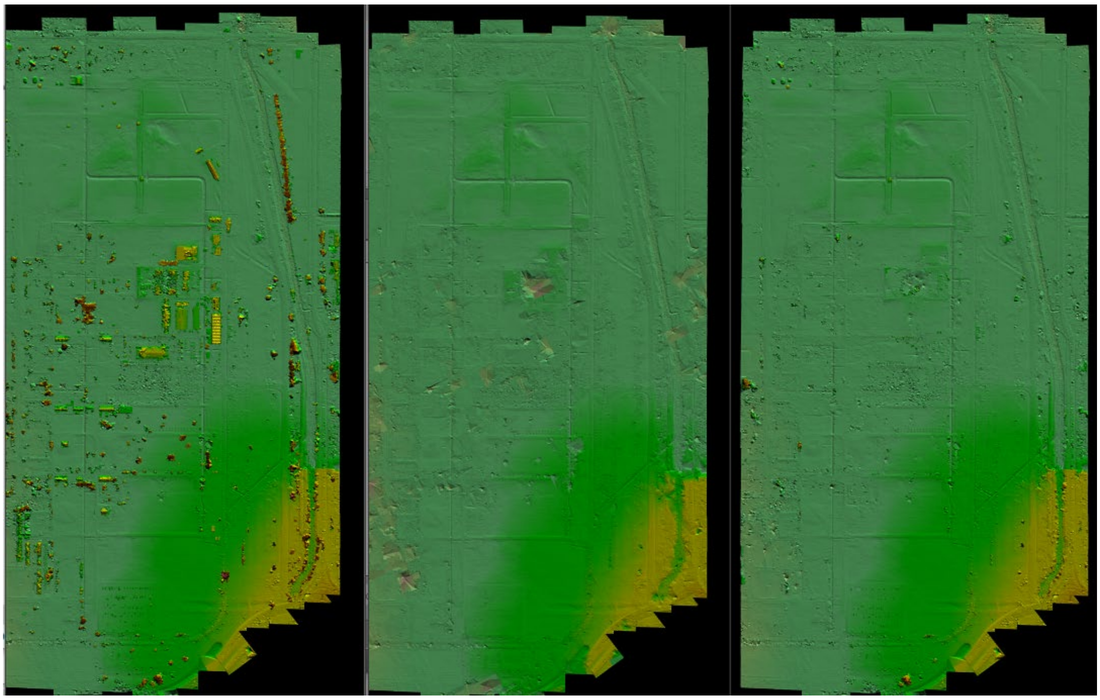
Intelligent Photogrammetry DEM can significantly reduce the DEM production cost from satellite stereo images, aerial stereo images, UAV stereo images, sparse LiDAR 3-D point clouds, and dense LiDAR 3-D point clouds. DEM generation from stereo satellite images is more difficult than the other sources. The DEM accuracy requirement in dense forestry areas is 2 to 3 meters according to a key customer. Our 3DGroundPointModel has achieved the required accuracy in this case study.

The 3DGroundPointModel is the equivalent of SOCET GXP terrain editing algorithm "*Dense_Trees/Buildings_with_Polygon_on_Ground*", which is the most efficient terrain editing algorithm in SOCET GXP software. This algorithm needs a human editor to manually define the boundary of a dense forestry area and identify ground points within the dense forestry area, whereas the 3DGroundPointModel does not require a human editor at all.

4. DEM production from UAV images

Intelligent Photogrammetry can achieve state-of-the-art DEM production accuracy from UAV images. In many uses cases, DEM is needed, but in this case study, SOCET GXP ASM generated a DSM from 660 UAV stereo images of GSD 0.025 meters. The DSM has a post spacing of 0.05 meters with 680 million posts (Figure 8). There are buildings, houses, and trees in this relatively flat area and the DSM accuracy is not very high. The middle image shows several errors in the ASM DEM, but the right image shows that our Intelligent Photogrammetry DEM corrects these errors.

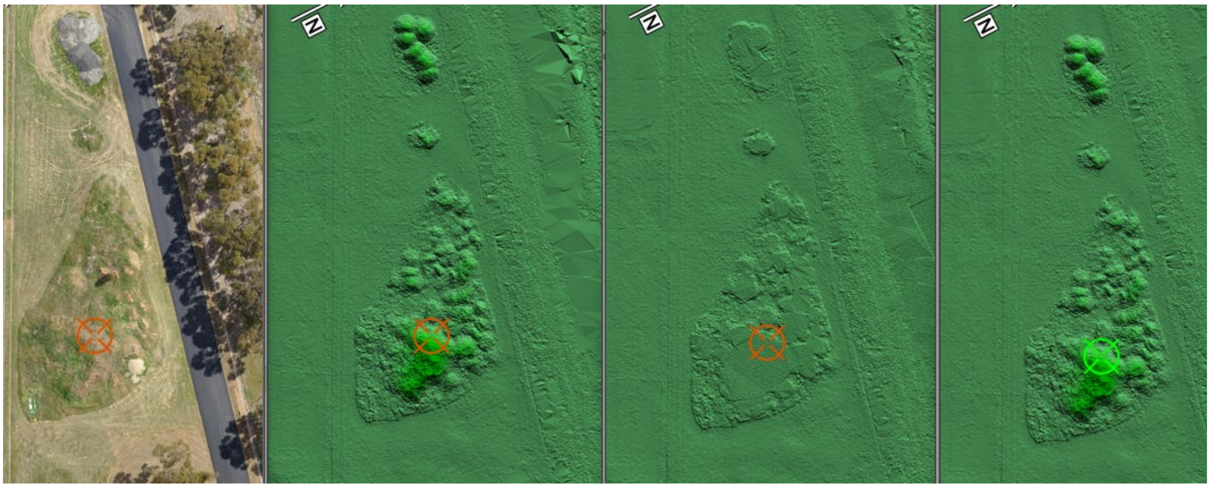
Figure 8. DEM from UAV images.



The left TSR image is the DSM. The accuracy of the DSM is not very high. The middle TSR image is the DEM generated by ASM. The right TSR image is the DEM generated by our Intelligent Photogrammetry.

Digital Photogrammetry uses handcrafted bare earth algorithms to generate DEM from DSM (Figure 9). The first ortho image shows a pile of dirt at the cursor location (ground). The second TSR image is the DSM and the pile of dirt is at the cursor location. The third TSR image is the DEM from Digital Photogrammetry, which uses handcrafted algorithms to generate DEM. The fourth TSR image is the DEM generated by our Intelligent Photogrammetry, which is based on machine learning. While Digital Photogrammetry wrongly removed the pile of dirt as above ground structures, our Intelligent Photogrammetry keeps the pile of dirt as ground. Our Intelligent Photogrammetry can be trained to know what a pile of dirt looks like by collecting a training sample on it. The Digital Photogrammetry bare earth algorithms cannot be trained to recognize a pile of dirt.

Figure 9. Intelligent Photogrammetry vs. Digital Photogrammetry.



Intelligent Photogrammetry can be trained to recognize a pile of dirt in terms of training samples, while Digital Photogrammetry bare earth algorithms (software) cannot be trained to recognize a pile of dirt. Piles of dirt were wrongly removed from DSM as shown in the third TSR image. The same piles of dirt were kept correctly as bare ground in the fourth TSR image by our Intelligent Photogrammetry.

This case study indicates that our machine learning based DEM production has much higher accuracy than handcrafted bare earth algorithms (Figure 10). The left TSR image is the DSM. The middle TSR image is the DEM generated by SOCET GXP ASM. The right TSR image is the DEM automatically generated by our Intelligent Photogrammetry using machine learning. In the lower left-hand corner, the area is already bare ground; that is, the DSM and DEM should be the same. There are two problems from ASM DEM as shown in the middle TSR image: the gridding effect and mistakenly lowered elevation of bare ground to below ground level.

Figure 10. Machine learning based DEM production from UAV images.

Left TSR image is DSM. Middle TSR image is DEM generated by handcrafted bare earth algorithm. Right TSR image is DEM from Intelligent Photogrammetry. This dataset is from a SOCET GXP customer.

5. Conclusion

The future of photogrammetry is Intelligent Photogrammetry, which uses machine learning to advance automation tasks in digital photogrammetry. Photogrammetry and 3-D mapping deal with images of the Earth's surface. The Earth's surface is complex and the images of it have orders of magnitude more complexities and variations. Traditional handcrafted automation algorithms cannot deal with so many variations and differences due to the dimensionality curse of pixels (Goodfellow, Bengio, and Courville, 2016). Prior to deep learning, we could not develop commercial grade AFE algorithms from images because of this dimensionality curse.

This case study indicates that DEM from Intelligent Photogrammetry can achieve RMSE of 0.95 meters in difficult terrain from stereo satellite image and may reduce DEM production cost by more than 50%. With more quality training samples, the DEM accuracy should continue to be higher because Intelligent Photogrammetry DEM is based on machine learning.

This case study indicates that the 3DGroundPointModel can automatically generate DEM in dense forestry areas and drastically reduce DEM production cost. Our 3DGroundPointModel classifies DSM 3-D posts and 3-D point clouds into thirteen categories, which are then used to compute forestry height. By subtracting the forestry height from DSM, we generated DEM in dense forestry areas, where the ground is unseen. This case study shows that Intelligent Photogrammetry reduces DEM error by 50% and meets required DEM accuracy in dense forestry areas.

6. Bibliography

Goodfellow, I., Y. Bengio and A. Courville, 2016. Deep Learning, The MIT Press, Cambridge, MA and London, UK, 775 pp.

Ren, S., K. He, R. Grishick, and J. Sun. 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks: arXiv:1506.01497.

Zhang, B., 2016. Big geospatial data + deep learning + high performance computing = geospatial intelligence, presentation S6260, 2016 GPU Technology Conference, San Jose, CA, 4-7 April. [http://on-demand.gputechconf.com/gtc/2016/presentation/s6260-bingcai-zhang-big-geospatial-data-deep-learning-high-performance-computing.pdf]

Zhang, B., 2017a. 3D DeepObject for precision 3D mapping, featured presentation S7149, 2017 GPU Technology Conference, San Jose, CA, 8-11 May. [https://on-demand.gputechconf.com/gtc/2017/presentation/s7149-bingcai_zhang-deepobject-for-precision-printing.pdf]

Zhang, B. 2017b. Rotation variant object detection in deep learning, Nov. 8. US Patent #10347720 [https://patents.google.com/patent/US10346720B2/en]

Zhang, B. and R. Hammoud, 2019. Fast and accurate target detection in overhead imagery using double convolution neural networks, invited presentation, SPIE Defence + Commercial Sensing, Baltimore Convention Center, Baltimore, Maryland, 14-18 April. [https://www.spiedigitallibrary.org/conference-proceedings-of-spie/10988/109880O/Fast-and-accurate-target-detection-in-overhead-imagery-using-double/10.1117/12.2519795.full?SSO=1]

Zhang, B. 2020. Intelligent photogrammetry for terrain production. BAE Systems technical report approved for public release.

BAE Systems, Inc. is the U.S. subsidiary of BAE Systems plc, an international defense, security and aerospace company which delivers a full range of products and services for air, land, and naval forces, as well as advanced electronics, security, information technology solutions, and customer support services.

Enabling development of the most advanced geospatial intelligence, BAE Systems GXP™ software enables rapid discovery, exploitation, and dissemination of mission-critical geospatial data. From key military, security, and incident response operations, to a variety of commercial development and research initiatives, GXP provides a comprehensive suite of solutions to inform effective decision-making and ensure a safer world.