Optimal Minimum Spanning Tree Algorithm
for Improving Coherence of Persistence Scatter-Interferometric
Synthetic Aperture Radar Technique

Youm Minkyo,[1] Kim Junkyeong,[1*] Park Sehwan,[1]
Lee Taeyoung,[2] and Jeong Myeong-Hun[2]

[1]Inspection for Infrastructure Laboratory (SIIL), Advanced Institute of Convergence Technology (AICT),
145 Gwanggyo-ro, Yeongtong-gu, Suwon-Si, Gyeonggi-do 16229, Korea
[2]Department of Civil Engineering, Chosun University, 309, Pilmun-daero, Dong-gu, Gwangju 61452, Korea

The persistence scatter Interferometric Synthetic Aperture Radar (PS-InSAR) analysis method is a technique that utilizes the persistent scatter in synthetic aperture radar images and image analysis is used via the interference of 25 or more slave images in a master image. The accuracy of the method is determined by the similarity between images, where the higher the coherence, the more accurate the image. Minimum spanning tree (MST) algorithms are used to find the optimal coherence by considering the temporal and spatial coherence of each image, instead of using the more common star graph, which interferes with the slave images in a single master image. When MST algorithms are employed considering the high coherence between images, pairs with high coherence can be connected; however, a higher number of images requires a higher processing speed. In this study, MST algorithms were therefore used by only considering the basic information in an image such as the spatial and temporal baselines, without the need for image processing. To verify the MST algorithms, a three-dimensional regression analysis was conducted, considering the correlation between the spatial and temporal baselines and the coherence. Furthermore, a novel MST algorithm was performed, which considered the weights derived from the preceding analysis. The results showed a high coherence of 98.5%, which was achieved rapidly, and a processing speed increase of approximately 120% over the results attained using the star graph. This study could be of great help in performing SAR image processing, which is part of remote sensing.

1. Introduction

One of the InSAR techniques, the persistence scatter Interferometric Synthetic Aperture Radar (PS-InSAR), was developed by Colesanti *et al.* at the POLIME Institute in Italy in the late 1990s. It is performed using a high-density persistent scatterer with the superior

---

reflectivity of the surface.[1–3] The average coherence depends on the pair of interference graphs that are adopted in carrying out the PS-InSAR technique. Three algorithms are used to construct the interference pairs: star graphs, baselines, and the minimum spanning tree (MST). In the existing PS-InSAR technique, a star graph is used to select the master image of the graph and create the link to the slave image. Star graphs ensure the temporal continuity of the deformation measurements and thus the possibility of phase unwrapping. However, a disadvantage is that the perpendicular baseline between the master image and the slave image might affect the temporal and spatial decorrelation. A small baseline subset (SBAS) algorithm was also developed to reduce geometric errors and temporal inconsistencies. However, the SBAS algorithm may not use all the available images, and there are concerns about the degradation of the coherence as the image graphs frequently disconnect.[4] Finally, an MST graph was developed and an optimal framework was designed to select the optimal cross-pair while considering both the spatial and temporal decorrelation.[5,6]

The PS-InSAR technique was implemented using star and MST graphs such as China Tianjin surface displacement monitoring[7] and InSAR analysis with low precision.[8] Examples in which PS-InSAR has been used include the use of MST algorithms for measurements such as surface deformation in northern Italy,[9] monitoring volcanoes,[10] ground deformation monitoring,[11] and building deformation monitoring.[12] In this study, the basic information provided in the images, such as the spatial or temporal baseline, was utilized. An MST graph was derived without image processing to obtain coherence.

2. Materials and Methods

2.1 PS-InSAR technique

The InSAR technique is an interfering radar technique that combines two or more SAR images to obtain three-dimensional information about a surface. It also has a high spatial resolution over a wide area range, so global monitoring is possible.[1] The PS-InSAR technique, one of the InSAR techniques, is used to select and analyze persistent scatter with high reflectivity in SAR images, and increases the coherence of the images when there is scatter that provides excellent reflectivity within one pixel of the SAR images, in a process known as persistent scatter. As the PS-InSAR technique works by matching at least 25 images, it can effectively eliminate various error components such as atmospheric effects and satellite orbit error. Also, digital elevation model (DEM) errors are deleted during the convergence of several SAR images, and no high-precision DEM adjustments are therefore required.[2,5,6,8]

2.2 Coherence

Coherence is an indicator of the quality of a SAR image. A fundamental step in calculating the coherence in the PS-InSAR process is to produce a reflective map of the master and slave images. The coherence is then calculated for each pixel using the following equation:

$$\left| \frac{E y_1 \times y_2}{\sqrt{E |y_1|^2} \times \sqrt{E |y_2|^2}} \right|, \, 0 \leq \gamma \leq 1 . \tag{1}$$

Coherence is affected by not only the geometric relationship between images but also the spatial and temporal baselines between images, which is known as decorrelation. This causes the coherence to degrade while also affecting the quality of the interferogram. Equation (2) indicates the element of decorrelation.[10,11]

$$\gamma_{total} = \gamma_{bl} \times \gamma_{dop} \times \gamma_{vol} \times \gamma_{thermal} \times \gamma_{temp} \times \gamma_{proc} \tag{2}$$

2.3   MST algorithm

The spanning tree shows the minimum possible number of connections in a graph. All nodes within a graph should be connected without containing cycles. Therefore, the spanning tree connects *n* nodes with (*n* − 1) lines. During the performance of the above algorithm, the MST describes a tree with the minimum weight sum of the edges used. In the MST algorithm, when the weight of each line differs, merely using the least line does not necessarily result in the minimum cost.[13,14] Therefore, the MST algorithm considers the weight of each line and selects the spanning tree with the minimum cost. The algorithm used to implement an MST algorithm is largely divided into the Kruskal or Prim MST algorithm.[15–17] The Kruskal MST algorithm is executed as shown in Fig. 1.

Line CD, the line with the smallest weight is selected first, followed by line AC, the line with the second smallest weight. The edge with the next smallest weight (edge AD) is not selected because it would form a cycle with the previously selected edges. Therefore, edge AB is selected, thus completing the scanning tree.

The processing of the Prim MST algorithm is as shown in Fig. 2. If starting from node A, the minimum cost line AC between lines AB and AC is selected. After selecting line CD for connection, line AD, which is the cycle of lines AD and BD, is removed, and line AB is selected.

Line	CD	AC	AD	AB	BD
Weight	2	3	4	5	6

Line	CD	AC	AD	AB	BD
Weight	2	3	4	5	6

Line	CD	AC	AD	AB	BD
Weight	2	3	4	5	6
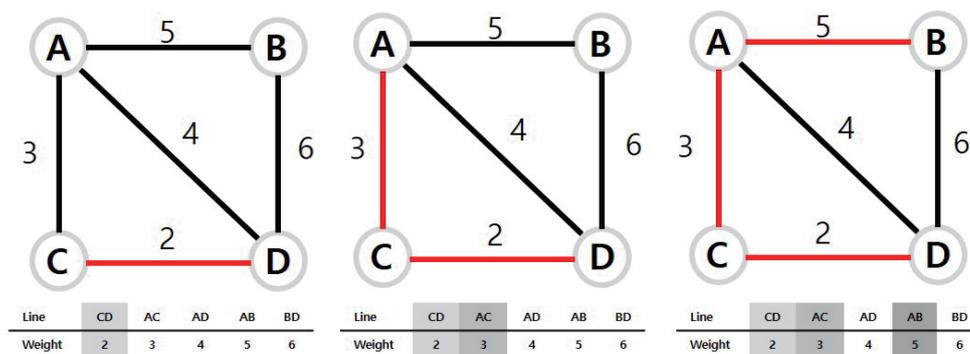
Fig. 1.   (Color online) Kruskal MST algorithm.
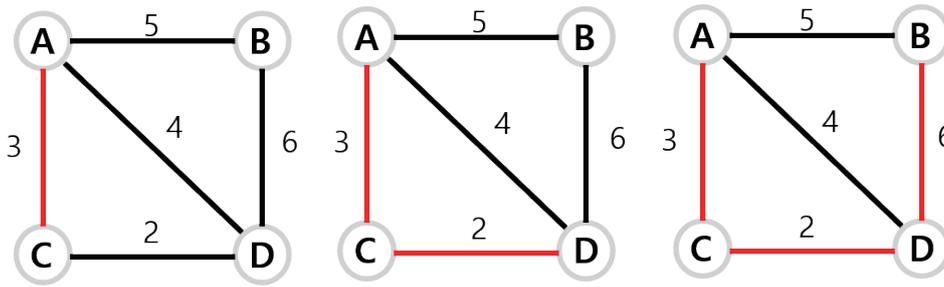
Fig. 2.   (Color online) Prim MST algorithm.

3.   Results and Discussion

3.1   Experimental overview

The study was conducted using 31 TerraSAR-X and TanDEM-X satellite images taken over a total period of 28 months. From the above images, 31 nodes were obtained and 465 lines were generated, as shown in Fig. 3.

3.2   Experimental method

In this study, we introduce an MST graph that uses new parameters to replace an MST graph that is generated on the basis of coherence. Our purpose is to reduce the time required for image processing and to ensure that the coherence is also maintained to at least a certain level. In terms of verification, since our priority is to shorten the processing time, we used the fast Prim algorithm as the MST algorithm. We compared the average coherences of four graphs, the contents of which are as follows:
(a) an MST graph based on the coherence between images,
(b) an MST graph based on the temporal and spatial coherence baselines (no weights),
(c) an MST graph based on the temporal and spatial coherence baselines (multilinear regression weights), and
(d) a star graph.

3.3   Multilinear regression method

The most valuable comparison group among the graphs to be analyzed in this study is the coherence-based MST graph. Coherence has the greatest influence on PS-InSAR accuracy, and the weight between the images is selected on the basis of this fact. In this study, the MST graph based on the temporal and spatial baselines was selected for experiments. In this process, the MST graph was derived using multilinear regression weights. Equation (3) shows the multilinear regression equation when there are two fixed variables.
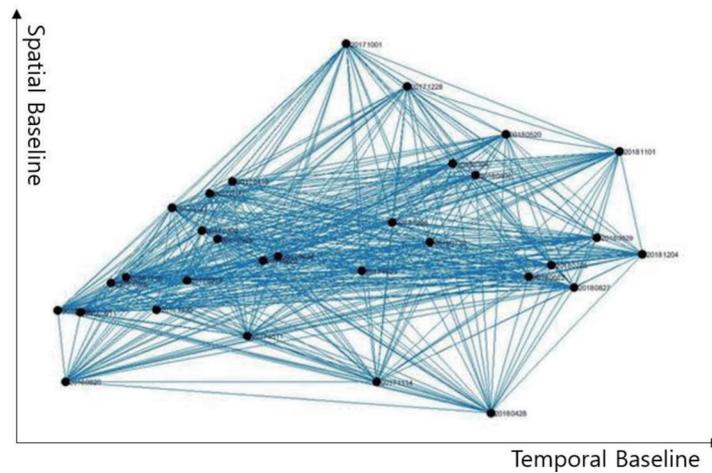
Fig. 3.   (Color online) All node connections.

$$Y = \beta_1 X_1 + \left(\beta_2 + \beta_{21} X_1 \right) X_2 + \beta_3 \tag{3}$$

The temporal and spatial baselines were selected as independent variables to perform the multilinear regression, and a graph was formed on the basis of the coherence between images. The interaction between the independent variables should be excluded in the multilinear regression process because there is an interaction between the coherence and the baselines. Therefore, the temporal and spatial baselines, which do not interact, are used as independent variables, and the coherence is used as a fixed variable. Assuming that there is no interaction between the independent variables, Eq. (4) can be expressed as

$$Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 , \tag{4}$$

where Y is the dependent variable, X is the independent variable, and both X and Y are the weights for each of the independent variables. Figure 4 illustrates the multilinear regression performed in this study. The weights derived from the above procedure are given in Eqs. (5) and (6).

$$0.583150Y = -\left(0.000386 X_1 + 0.000169 X_2 \right) \tag{5}$$

$$Y = -\left(6.619e^{-4} X_1 + 2.898e^{-4} X_2 \right) \tag{6}$$

As shown in Table 1, the coefficients of the temporal and spatial baselines are 0.6955 and 0.3045, respectively. The R-square value is 0.726 and the adjusted R-square value is 0.725 if there is a significant difference. If the difference between these coefficients is significant, the correlation between the two variables is high. The difference between the two variables derived from this study is 0.001, which means that these variables can be inferred to be almost completely independent.
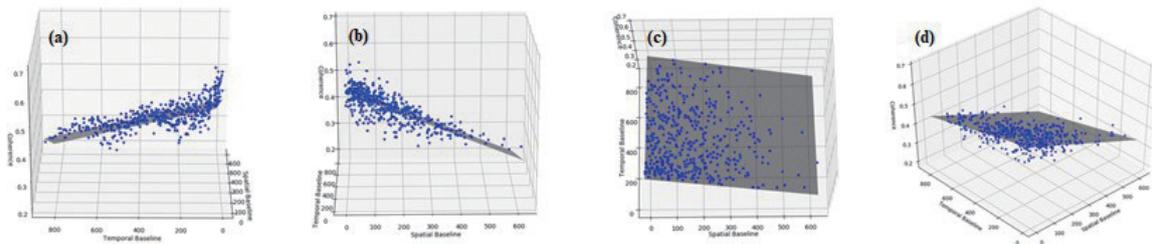
Fig. 4.   (Color online) Multilinear regression: (a) front image (temporal baseline), (b) front image (spatial baseline), (c) plane image, and (d) quarter image.

Table 1
Multilinear regression values.

	Coefficient	Adj-coefficient	Std. error
Coherence	0.5832	—	0.0004
Spatial baseline	3.86e−4	0.6955	7.88e−6
Temporal baseline	1.69e−4	0.3045	1.37e−5

4.   Results

In this study, 31 images were analyzed, and coherence was extracted through matching 465 images. A graph including all the images was generated, compared, and analyzed, and the contents are plotted in Fig. 5 and described in Table 2. As shown in Fig. 5, the coherence-based MST graph was selected as a comparative group, and a baseline-based MST graph with no weights, a baseline-based MST graph using multilinear regression weights, and a star graph, which is a standard graph used with PS-InSAR, were used for comparison and analyzed. The average coherence was determined to be 0.5677 (94.91%) for the graph with no weights and 0.5966 (98.53%) for the graph with multilinear regression weights. The coherence-based MST graph has the lowest standard deviation among the coherence-based MST graphs. Finally, the line concordance rates were 63.33 and 60%, which correspond to MLR weights and no weights, respectively, based on graph (a). The star graph confirmed that indicators such as the coherence mean and standard deviation are lower than those in the MST algorithms.

Figure 6 shows the mean, maximum, and minimum coherence images of the network graphs, with the corresponding values provided in Table 3. As indicated in Table 3, the max-min deviations in coherence were in the order (a)<(d)<(c)<(b). In particular, the data shows a similar deviation for the MST graph with no weights and the star graph. Also, the comparison of the MST graphs with the multilinear regression MST graphs indicates that the max-min deviation is lower, and thus the average is higher.

Figure 7 shows the average coherence distribution of each graph. As demonstrated in the figure, the coherence-based MST graph has a high average coherence because of its high overall density and small variance. The baseline-based MST graph (no weights) is less dense than the coherence-based MST graph and has a broader dispersion. The baseline-based MST graph (MLR weights) was found to be skewed towards larger densities with a narrow distribution and a high
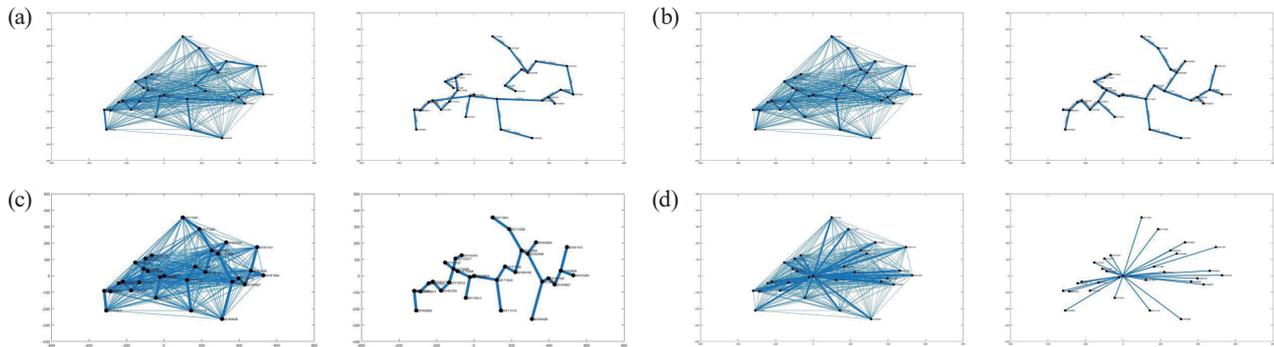
Fig. 5. (Color online) Connection graphs: (a) coherence-based MST, (b) baseline-based MST (no weights), (c) baseline-based MST (MLR weights), and (d) star graph.

Table 2
Experimental results.

	(a)	(b)	(c)	(d)
Coherence	0.6055	0.5747	0.5966	0.4988
Average (%)	100	94.91	98.53	82.38
Standard deviation	0.384	0.0554	0.0406	0.0575
Matching rate (%)	100	63.33	60	—
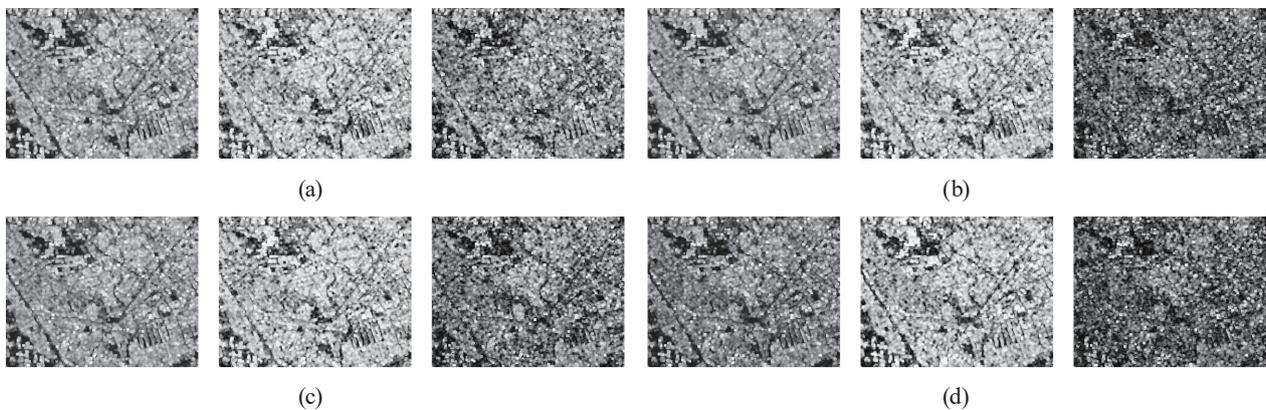
Fig. 6. Coherence image (average: left, maximum: center, minimum: right) (a) coherence-based MST, (b) baseline-based MST (no weights), (c) baseline-based MST (MLR weights), and (d) star graph.

Table 3
Experimental results.

	Coherence-based MST	Baseline-based MST (no weights)	Baseline-based MST (MLR weights)	Star graph
Avg	0.6055	0.5747	0.5966	0.4988
Max	0.6774	0.6774	0.6774	0.6549
Min	0.5514	0.4255	0.5185	0.4054
Max − Min	0.1260	0.2519	0.1589	0.2495
Processing time	2 h 15 min	21 min	28 min	8 min
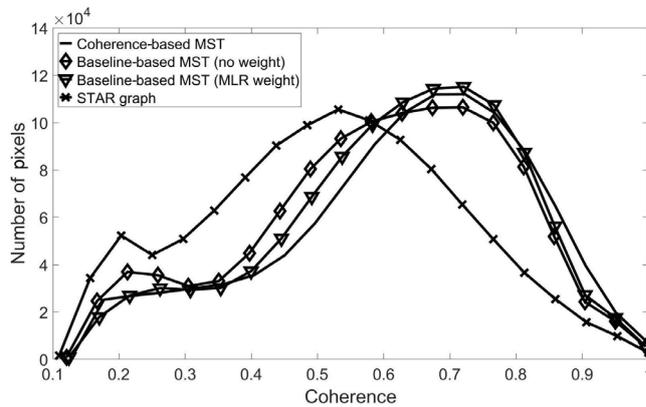
Fig. 7.   Average coherence graph.
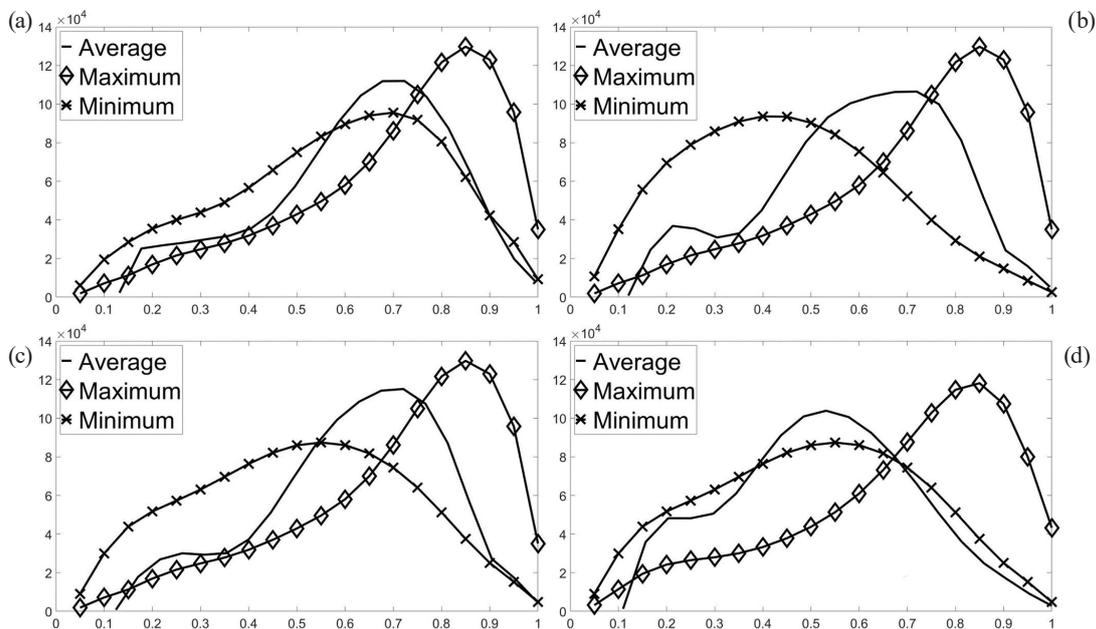
Fig. 8.   Maximum, minimum, and average coherence images: (a) coherence-based MST, (b) baseline-based MST (no weights), (c) baseline-based MST (MLR weights), and (d) star graph.

center value as compared with the baseline-based MST graph (no weight). Finally, for the star graph, the variance is equal to that of the baseline-based MST graph (no weights), whereas the center value of the graph shifts towards the lower side. Figure 8 shows the maximum, minimum, and average coherence values of each graph. As can be seen in the figure, it can be concluded that the maximum coherence values of the remaining graphs, except for the star graph, are the same, but the graphs differ in terms of the minimum coherence. In particular, it can be seen in graphs (b) and (c) that the center value of the least coherent graph increases through the MLR weights. Table 3 shows detailed information about the graphs in Fig. 8. The processing speed of the proposed method is 6.43 times higher than that of the coherence-based MST graph.

5.   Conclusions

We derived an MST graph using a new method to improve a coherence-based MST graph for determining the accuracy of PS-InSAR, which is accurate but requires a significant amount of time for image processing in coherence calculation. Multiple linear regression was performed using basic image information and the temporal and spatial baselines. The weight of each baseline was used to generate the MST graph. The results are as follows.

1. The MST algorithm using the first test group, with multilinear regression weights, showed a coherence level of 98.53% as compared with the coherence-based MST graph (100%), which is a comparative group. It can be processed faster than the existing coherence-based MST graphs, although the accuracy is slightly lower.

2. In the case of the baseline-based MST graph with no weights, which is the second group in the experiment, although the accuracy is lower than that of the MST algorithm using multilinear regression weights, the image processing can be performed at a higher rate because the weight determination step that uses multilinear regression is omitted,

3. Finally, in the process of obtaining the MST graph, the line concordance rate was found to be 63.33%, and no correlation between the coincidence rate and the coherence was found.

4. The MST graph produced using multilinear regression weights shows that the maximum coherence does not differ as compared with the MST graph without weights, but a significant difference has occurred in minimum coherence. This indicates that the multilinear regression does not increase the maximum coherence, but increases the minimum coherence, resulting in an increase in average coherence.

5. From the analysis of the average coherence of each graph, MST techniques that use multiple linear circuit weights show high average densities and narrow data distributions, meaning that they are evaluated as techniques with higher utilization values than other experimental groups.

6. Baseline-based MST graphs that use MLR weights show significantly higher processing speeds than conventional coherence-based MSTs. Therefore, it is possible to prove the high accuracy and processing speed of the method used in this study.

However, 465 cases were considered in this study, but image processing is required under other conditions. This is because the multiple linear regression coefficients derived from this study can be similar depending on the number of images, the number of lines, and the location of satellites. However, through this study, a method of increasing the speed of SAR image processing, a type of remote sensing, was established, and meaningful results were derived.

Acknowledgments

References

1   C. Colesanti, A. Ferretti, C. Prati, and F. Rocca: Eng. Geol. **68** (2003) 3. https://doi.org/10.1016/S0013-7952(02)00195-3
2   C. Colesanti, A. Ferretti, F. Novali, C. Prati, and F. Rocca: IEEE Trans. Geosci. Remote Sens. **41** (2003) 1685. https://doi.org/10.1109/TGRS.2003.813278
3   N. Cao, H. Lee, and H. C. Jung: IEEE Trans. Geosci. Remote Sens. **54** (2015) 1074. https://doi.org/10.1109/TGRS.2015.2473818
4   B. Hu, H.-S. Wang, Y.-L. Sun, J.-G. Hou, and J. Liang: Remote Sens. **6** (2014) 3648. https://doi.org/10.3390/rs6053648
5   Q. Luo, G. Zhou, and D. Perissin: Remote Sens. **9** (2017) 717. https://doi.org/ 10.3390/rs9070717
6   D. Perissin and T. Wang: IEEE Trans. Geosci. Remote Sens. **50** (2011) 271. https://doi.org/10.1109/TGRS.2011.2160644
7   L. San-Ming, S. Xin-Jian, Z. Wen-Wu, D. Kai-Fu, W. Wen-Ni, L. Hong-Bao, and L. Zhi-Guang: Chin. J. Geophys. Chin. Ed. **57** (2014) 3129. https://doi.org/10.6038/cjg20141004
8   L. Zhang, X. Ding, and Z. Lu: Int. J. Image Data Fusion **6** (2015) 289. https://doi.org/10.1080/19479832.2015.1068874
9   C. Meisina, F. Zucca, F. Conconi, F. Verri, D. Fossati, M. Ceriani, and J. Allievi: Quat. Int. **171** (2007) 90. https://doi.org/10.1016/j.quaint.2006.12.011
10 A. Bonforte, F. Guglielmino, M. Coltelli, A. Ferretti, and G. Puglisi: Geochem. Geophys. Geosyst. **12** (2011). https://doi.org/10.1029/2010GC003213
11 C. Meisina, F. Zucca, D. Fossati, M. Ceriani, and J. Allievi: Eng. Geol. **88** (2006) 240. https://doi.org/10.1016/j.enggeo.2006.09.010
12 A. Ciampalini, F. Bardi, S. Bianchini, W. Frodella, C. Del Ventisette, S. Moretti, and N. Casagli: Int. J. Appl. Earth Obs. Geoinf. **33** (2014) 166. https://doi.org/10.1016/j.jag.2014.05.011
13 G. Navarro and R. Paredes: Algorithmica **57** (2010) 585. https://doi.org/10.1007/s00453-010-9400-6
14 G. R. Raidl and B. A. Julstrom: IEEE Trans. Evolutionary Comput. **7** (2003) 225. https://doi.org/10.1109/TEVC.2002.807275
15 T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein: Introduction Algorithms **3** (2009) 631.
16 R. C. Prim: The Bell Syst. Tech. J. **36** (1957) 1389. https://doi.org/10.1002/j.1538-7305.1957.tb01515.x
17 J. B. Kruskal: Proc. American Mathematical Society **7** (1956) 48. https://doi.org/10.2307/2033241

About the Authors

Youm Minkyo received his M.S. and Ph.D. degrees from Sungkyunkwan University in 2013 and 2019, respectively. From 2013 to 2014, he was a researcher at the Korea Atomic Energy Research Institute, Korea. From 2014 to 2015, he was a researcher at the National Disaster Management Institute, Korea. Since 2020, he has been a senior researcher at the Advanced Institute of Convergence Technology. His research interests are in remote sensing, image processing, InSAR, and satellite image processing. (tomsmith850918@gmail.com)

Kim Junkyeong received his M.S. and Ph.D. degrees from Sungkyunkwan University in 2012 and 2018, respectively. From 2012 to 2013, he was a researcher at Kumkang Coryo Chemicals, Korea. Since 2019, he has been a senior researcher at the Advanced Institute of Convergence Technology. His research interests are in sensor-based infrastructure safety inspection. (junkyeong@snu.ac.kr)

Park Sehwan received his M.S. degree from Sungkyunkwan University, Korea, in 2020. Since 2020, he has been a researcher at the Advanced Institute of Convergence Technology. His research interests are in sensors, image processing, and computer vision. (sehwan0721@snu.ac.kr)

Lee Taeyoung is studying for a master's degree in the Department of Civil Engineering, Chosun University, Republic of Korea, from which he received his bachelor's degree in 2020. (taeyounglee1@naver.com)

Jeong Myeong-Hun is an assistant professor in the Department of Civil Engineering at Chosun University, Republic of Korea. His research interests are in spatial data science, dynamism, and uncertainty with visualization. Previously, he was a postdoctoral research associate at the CyberGIS Center for Advanced Digital and Spatial Studies in the National Center for Supercomputing Applications, The University of Illinois at Urbana-Champaign, USA, and he received a Ph.D. degree from the University of Melbourne, Australia. (mhjeong@myeonghun.org)