

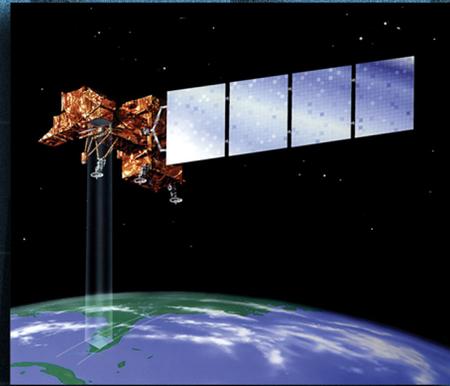
# PE&RS

August 2021

Volume 87, Number 8

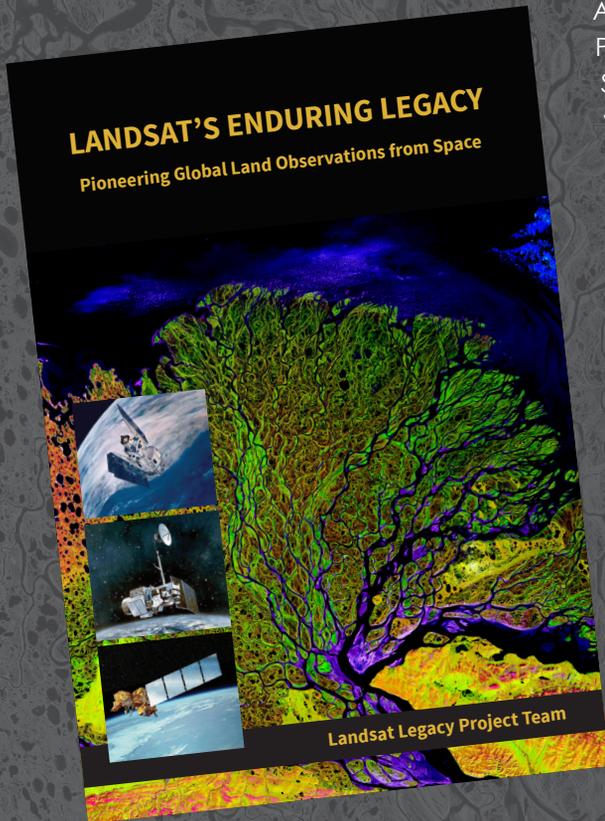
*The official journal for imaging and geospatial information science and technology*

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING



# LANDSAT'S ENDURING LEGACY

## PIONEERING GLOBAL LAND OBSERVATIONS FROM SPACE



After more than 15 years of research and writing, the Landsat Legacy Project Team is about to publish, in collaboration with the American Society for Photogrammetry and Remote Sensing (ASPRS), a seminal work on the nearly half-century of monitoring the Earth's lands with Landsat. Born of technologies that evolved from the Second World War, Landsat not only pioneered global land monitoring but in the process drove innovation in digital imaging technologies and encouraged development of global imagery archives. Access to this imagery led to early breakthroughs in natural resources assessments, particularly for agriculture, forestry, and geology. The technical Landsat remote sensing revolution was not simple or straightforward. Early conflicts between civilian and defense satellite remote sensing users gave way to disagreements over whether the Landsat system should be a public service or a private enterprise. The failed attempts to privatize Landsat nearly led to its demise. Only the combined engagement of civilian and defense organizations ultimately saved this pioneer satellite land monitoring program. With the emergence of 21st century Earth system science research, the full value of the Landsat concept and its continuous 45-year global archive has been recognized and embraced. Discussion of Landsat's future continues but its heritage will not be forgotten.

The pioneering satellite system's vital history is captured in this notable volume on Landsat's Enduring Legacy.

### Landsat Legacy Project Team

Samuel N. Goward  
Darrel L. Williams  
Terry Arvidson  
Laura E. P. Rocchio  
James R. Irons  
Carol A. Russell  
Shaida S. Johnston

### Landsat's Enduring Legacy

Hardback, 2017, ISBN 1-57083-101-7  
Student \$60\*  
Member \$80\*  
Non-member \$100\*

\* Plus shipping

## SAVE 40%

VISIT THE ASPRS BOOKSTORE at [ASPRS.ORG](http://ASPRS.ORG)

## ANNOUNCEMENTS

**Teledyne Optech**, a Teledyne Technologies company and global leader in advanced lidar sensors has delivered its next generation bathymetric lidar CZMIL SuperNova to leading Norwegian mapping firm Terratec AS. Terratec is the first private company to acquire the CZMIL SuperNova and the only company in the EU to engage in work around environmental change and coastal zone mapping using the CZMIL SuperNova's powerful depth penetration and advanced capabilities.

Teledyne Optech's new CZMIL SuperNova has the best bathymetric depth performance, the highest green laser point density in its class and greatly improved range precision. CZMIL SuperNova also introduces SmartSpacing technology for even and efficient point spacing across the swath, onboard processing capability for reduced post-processing and three configurable modes for maximizing performance in different marine environments.

As a complete geospatial solution, CZMIL SuperNova has integrated industry-leading software from Teledyne CARIS, to provide seamless integration for processing and deliverables. The complete CARIS Ping-to-Chart™ workflow allows for a myriad of deliverables that go far beyond the mere point cloud. Moreover, leveraging advanced AI techniques for automated land/water discrimination and noise classification, the CZMIL SuperNova bathymetric solution effectively sets a new standard in processing workflow efficiency through automation without compromising quality.

CEO of Terratec, Oivind O. Aase stated: "Terratec is excited to be the first customer worldwide to take advantage of the new CZMIL SuperNova sensor. We have been following the development in airborne bathymetry for a long time and for us this has been a maturation process. The combination of the SuperNova hardware and CARIS processing solution represents a great step forward for airborne bathymetry. Our pilot clients are very satisfied with the results we have presented thus far. We look forward to continued collaboration with Teledyne Optech in further development of the AI algorithms to release the great potential we see in this software solution."

For more information, visit [www.teledyneimaging.com](http://www.teledyneimaging.com)

**UP42 and SI Imaging Services (SIIS)** of Daejeon, South Korea, have signed an agreement to make imagery from the KOMPSAT satellites available on the UP42 marketplace and developer platform. The deal includes high-resolution optical

imagery from KOMPSAT-3 and -3A, and Synthetic Aperture Radar (SAR) data from KOMPSAT-5.

KOMPSAT imagery is a valuable addition to the more than 50 geospatial data sets now available on the UP42 marketplace, including satellite imagery from five international organizations. UP42 users will find imagery from the Korean constellation complements other data products by offering diverse spatial and spectral capabilities, broad dynamic ranges, afternoon acquisition times, extensive archives, and attractive price points.

"The addition of the KOMPSAT 3, 3A, and 5 broadens our current offering of high-resolution satellite and radar data. This high-quality constellation is invaluable for unlocking new use cases for our users globally through higher combined revisit and ranging acquisition modes," said UP42 CEO Sean Wiid.

Wiid predicts one of the most important impacts on new use cases for UP42 users will come in the area of intraday monitoring — with the aid of KOMPSAT's afternoon collection times. The addition of KOMPSAT data will also provide existing UP42 users with more options to augment infrastructure monitoring and vegetation management use cases.

"The UP42 marketplace enabled us to reach out to potential global customers. On the UP42 platform, customers will have easy access to high quality VVHR KOMPSAT EO and SAR constellation imagery up to 40cm resolution and gain valuable information through analysis," said Moon-Gyu Kim, CEO of SI Imaging Services.

In its initial offering from KOMPSAT, UP42 will make tasking and archive ordering available from KOMPSAT-3, KOMSAT-3A, and KOMPSAT-5 satellites.

The UP42 platform offers extensive Earth observation data sets and more than 70 analytics tools – along with cloud computing power – to create custom geospatial solutions easily and inexpensively. Users purchase just the data needed to cover their area of interest and then leverage off-the-shelf processing capabilities to analyze the data sets without investment in their own computing infrastructure. And for users who only want to purchase imagery, one simple API on the UP42 marketplace enables them to search multiple data providers and integrate selected imagery directly into their geospatial workflows.

Tasked and archive KOMPSAT imagery may be ordered now, along with other satellite and aerial data sets, by visiting <https://up42.com/order-data>.

## TECHNOLOGY

**Teledyne Optech** and **Teledyne CARIS**, both Teledyne Technologies companies and global leaders in advanced lidar sensors and marine mapping software, announce their next generation bathymetric lidar, the CZMIL SuperNova.

The CZMIL SuperNova boasts the best depth performance and the highest green laser point density in its class. Introducing

SmartSpacing technology for even and efficient point spacing, real-time processing capability for reduced post-processing time and configurable modes for maximizing performance in different water environments, the SuperNova provides a wide range of inputs for climate change modelling and is ideal for inland water environments, base mapping for coastal zones and shoreline.

To complete the solution, Teledyne CARIS has integrated its BASE Editor software for seamless data processing capacity. Leveraging AI techniques for land/water discrimination and noise classification the CZMIL SuperNova bathymetric solution effectively delivers on marketplace demands for efficiencies in the processing workflow.

CZMIL SuperNova Product Manager Jennifer Aitken explained: “CZMIL SuperNova builds on over 30 years of experience with airborne green laser bathymetry. The ability to configure the system on-the-fly and optimize data collection for local conditions has significant benefit. We are getting plenty of signal at 600 meters altitude and even higher, so our coverage area is increasing without sacrificing depth penetration.”

For more information visit <https://www.teledyneoptech.com/en/products/airborne-survey/czmil-supernova/>.

**Golden Software**, a developer of affordable software for visualizing and analyzing diverse data sets, has announced significant upgrades to its scientific graphing package, Grapher™. The new version released in June 2021 is faster and offers a broader, more consistent selection of features across all plot types.

“Grapher is now easier and more intuitive to use with nu-

merous functions and capabilities standardized throughout the application,” said Justine Carstairs, Grapher Product Manager. “Users will also find the new version provides improved display features for better communication of information.”

Grapher is used extensively by scientists and engineers in oil & gas operations, environmental consulting, climate research, mineral exploration, and academic pursuits. The package offers deeper insights into diverse data sets, including chemical, physical, geologic and geospatial data, through 80 different 2D and 3D plotting types.

Golden Software’s development team is already working on new features, functionality and enhancements for the next version of Grapher that will be released in late 2021. Users with active maintenance can take these new capabilities for a test drive by downloading Grapher Beta. They are invited to send comments and/or questions directly to the Golden Software development team as they use the Beta version in their day-to-day workflows.

“In the Beta version, we are continuing to build on our progress making Grapher more stable and easier to use,” said Carstairs.

For additional information, visit [www.GoldenSoftware.com](http://www.GoldenSoftware.com).

## CALENDAR

- 16-20 August, **URISA GIS Leadership Academy**. For more information, visit [www.urisa.org/education-events/urisa-gis-leadership-academy/](http://www.urisa.org/education-events/urisa-gis-leadership-academy/).
- 19 August, **2021 Florida Region Summer SAR Workshop**. For more information, visit <https://www.asprs.org/event/2021-florida-region-summer-sar-workshop>.
- 3-6 October, **GIS-Pro 2021**, Baltimore, Maryland. For more information, visit [www.urisa.org/gis-pro](http://www.urisa.org/gis-pro).
- 8-12 November, **URISA GIS Leadership Academy**, St. Petersburg, Florida. For more information, visit [www.urisa.org/education-events/urisa-gis-leadership-academy/](http://www.urisa.org/education-events/urisa-gis-leadership-academy/).
- 14-18 December, **30<sup>th</sup> International Cartographic Conference & International Cartographic Exhibition**, Florence, Italy. For more information, visit <https://icaci.org/icc2021>.
- 6-8 February 2022, **Geo Week 2022**, Denver, Colorado. For more information, visit [www.geo-week.com/](http://www.geo-week.com/).



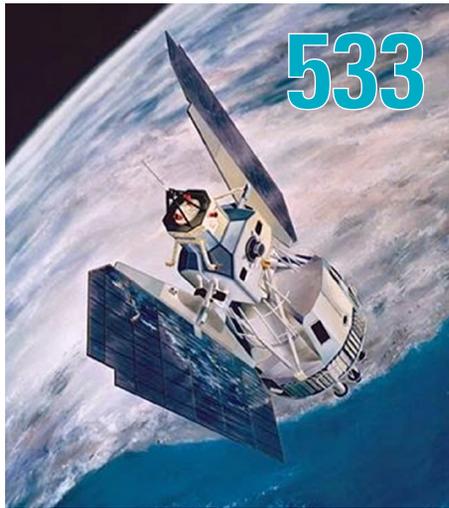
**Too young to drive the car? Perhaps!  
But not too young to be curious about geospatial sciences.**

The ASPRS Foundation was established to advance the understanding and use of spatial data for the betterment of humankind. The Foundation provides grants, scholarships, loans and other forms of aid to individuals or organizations pursuing knowledge of imaging and geospatial information science and technology, and their applications across the scientific, governmental, and commercial sectors.

**Support the Foundation, because when he is ready so will we.**

**[asprsfoundation.org/donate](http://asprsfoundation.org/donate)**





## Semi-Centennial of Landsat Observations & Pending Landsat 9 Launch

By Samuel N. Goward, Jeffrey G. Masek, Thomas R. Loveland, John L. Dwyer, Darrel L. Williams, Terry Arvidson, Laura E.P. Rocchio, and James R. Irons



## GIS Tips & Tricks

By Brittany Capra and Al Karlin, Ph.D., CMS-L, GISP

## COLUMNS

- 541** GIS Tips & Tricks—You don't have to accept the defaults in GlobalMapper
- 545** Book Review—*Remote Sensing Time Series Image Processing. First Edition*
- 547** Grids and Datums Update  
This month we look at the Republic of Yemen

## ANNOUNCEMENTS

- 540** Call for Submissions—*PE&RS* Special Issue on Remote Sensing Monitoring for Urban Environment
- 540** New ASPRS Members  
Join us in welcoming our newest members to ASPRS.
- 543** Certifications
- 549** Headquarters News

## DEPARTMENTS

- 529** Industry News
- 530** Calendar
- 544** ASPRS Sustaining Members
- 550** Who's Who in ASPRS
- 566** In-Press *PE&RS* Articles

### 551 Spinor-Based Attitude Determination with Star Sensor Considering Depth

Qinghong Sheng, Rui Ren, Weilan Xu, Hui Xiao, Bo Wang, and Ran Hong

A star sensor is a high-precision satellite attitude measurement device. Since its observation information has only two-dimensional direction vectors, when a star sensor is used for attitude determination the dimension of the observation information is less than the number of attitude angles determined, so mainstream algorithms usually only guarantee the accuracy of the pitch angle and the roll angle. In view of the lack of depth information in the observation's imaging geometric condition, this article proposes a spinor-based attitude determination model, which describes a straight line passing through two stars with the spinor and maps the depth information of the straight line with the pitch, to establish an imaging geometry model of the spinor coplanar condition.

### 557 Digital Building-Height Preparation from Satellite Stereo Images

Prakash P. S. and Bharath H. Aithal

Buildings are considered prominent objects for understanding the pattern of growth in an urban setting. Remote sensing technology plays a vital role in facilitating data generation pertaining to various urban applications. Digital surface models represent the elevation of the earth surface features, and can be obtained from stereo images, radar, laser scanning, and so on. Photogrammetric techniques applied to optical stereo satellite images are economical and fast ways to generate height information of buildings. In this article, a quantitative and qualitative analysis of digital surface models generated from Cartosat-1 stereo images is compared with openly available data.

### 567 Enhanced Lunar Topographic Mapping Using Multiple Stereo Images Taken by Yutu-2 Rover with Changing Illumination Conditions

Wenhui Wan, Jia Wang, Kaichang Di, Jian Li, Zhaoqin Liu, Peng Man, Yexin Wang, Tianyi Yu, Chuankai Liu, and Lichun Li

In a planetary-rover exploration mission, stereovision-based 3D reconstruction has been widely applied to topographic mapping of the planetary surface using stereo cameras onboard the rover. In this article, we propose an enhanced topographic mapping method based on multiple stereo images taken at the same rover location with changing illumination conditions.

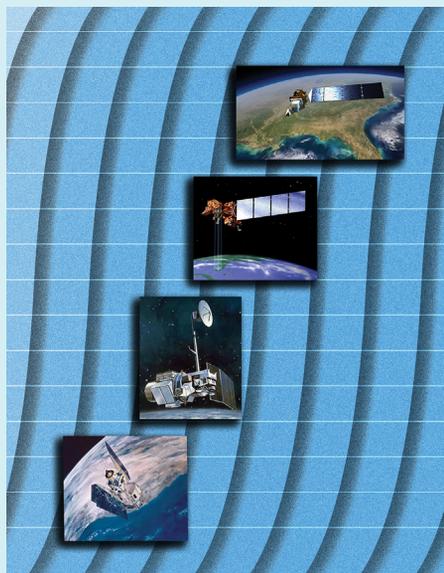
### 577 Unsupervised Representation High-Resolution Remote Sensing Image Scene Classification via Contrastive Learning Convolutional Neural Network

Fengpeng Li, Jiabao Li, Wei Han, Ruyi Feng, and Lizhe Wang

Inspired by the outstanding achievement of deep learning, supervised deep learning representation methods for high-spatial-resolution remote sensing image scene classification obtained state-of-the-art performance. However, supervised deep learning representation methods need a considerable amount of labeled data to capture class-specific features, limiting the application of deep learning-based methods while there are a few labeled training samples. An unsupervised deep learning representation, high-resolution remote sensing image scene classification method is proposed in this article to address this issue.

**See the Cover Description on Page 532**

# COVER DESCRIPTION



## Evolution of Landsat satellites

Left to right:

Artistic rendition of Landsats 1 (ERTS), 2 and 3.

Artistic rendition of Landsats 4 and 5.

Artistic rendition of Landsat 7

Artistic rendition of Landsat 8

On July 23, 1972, the Earth Resources Technology Satellite (ERTS-1) was launched into space onboard a Delta 900 rocket from Vandenberg Air Force Base, California. ERTS-1 was the first Earth-observing satellite launched to monitor and study our planet's landmasses. Also known as ERTS-A, the satellite was renamed Landsat 1 in 1975 and collected data until January 1978.

Landsat 2 was launched onboard a Delta 2910 rocket from Vandenberg Air Force Base, California on January 22, 1975. Originally named ERTS-B (Earth Resource Technology Satellite B), the spacecraft was renamed Landsat 2 before launch. On February 25, 1982 after seven years of service, Landsat 2 was removed from operations due to yaw control problems; it was officially decommissioned on July 27, 1983.

Landsat 3 (originally named Landsat C) was launched into space onboard a Delta 2910 rocket from Vandenberg Air Force Base, California on March 5, 1978. The objective of Landsat 3 was to extend the period of space-acquired Earth imagery, started by Landsat 1 and Landsat 2. The satellite was placed in standby mode on March 31, 1983 and decommissioned on September 7, 1983.

Landsat 4 was launched from Vandenberg Air Force Base in California on July 16, 1982 on a Delta 3920 rocket. With an updated design than the previous three missions, the satellite carried the Multispectral Scanner (MSS) as well as the new Thematic Mapper (TM) instruments. The sensors onboard the satellite collected data until late 1993, and the satellite was decommissioned on June 15, 2001.

Landsat 5 was launched from Vandenberg Air Force Base in California on March 1, 1984, and like Landsat 4, carried the Multispectral Scanner (MSS) and the Thematic Mapper (TM) instruments. Landsat 5 delivered Earth imaging data nearly 29 years - and set a Guinness World Record For 'Longest Operating Earth Observation Satellite', before being decommissioned on June 5, 2013.

Landsat 6 was launched on October 5, 1993 on a Titan II rocket from Vandenberg Air Force Base, California, but did not achieve orbit. The satellite carried the Enhanced Thematic Mapper, an improved version of the instruments on Landsat 4 and Landsat 5, and included a 15-meter panchromatic band.

Landsat 7 was launched from Vandenberg Air Force Base in California on April 15, 1999 on a Delta II rocket. The satellite carries the Enhanced Thematic Mapper (ETM+) sensor. Since June 2003, the sensor has acquired and delivered data with data gaps caused by the Scan Line Corrector (SLC) failure.

Landsat 8 (formally the Landsat Data Continuity Mission, LDCM) was launched on an Atlas-V rocket from Vandenberg Air Force Base, California on February 11, 2013. Landsat 8 is the most recently launched Landsat satellite and carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments.

Satellite description text via <https://www.usgs.gov/core-science-systems/nli/landsat>.



## PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING

JOURNAL STAFF

Publisher ASPRS

Editor-In-Chief Alper Yilmaz

Assistant Director — Publications Rae Kelley

Electronic Publications Manager/Graphic

Artist Matthew Austin

*Photogrammetric Engineering & Remote Sensing* is the official journal of the American Society for Photogrammetry and Remote Sensing. It is devoted to the exchange of ideas and information about the applications of photogrammetry, remote sensing, and geographic information systems. The technical activities of the Society are conducted through the following Technical Divisions: Geographic Information Systems, Photogrammetric Applications, Lidar, Primary Data Acquisition, Professional Practice, and Remote Sensing Applications. Additional information on the functioning of the Technical Divisions and the Society can be found in the Yearbook issue of *PE&RS*.

Correspondence relating to all business and editorial matters pertaining to this and other Society publications should be directed to the American Society for Photogrammetry and Remote Sensing, 425 Barlow Place, Suite 210, Bethesda, Maryland 20814-2144, including inquiries, memberships, subscriptions, changes in address, manuscripts for publication, advertising, back issues, and publications. The telephone number of the Society Headquarters is 301-493-0290; the fax number is 225-408-4422; web address is [www.asprs.org](http://www.asprs.org).

**PE&RS.** *PE&RS* (ISSN0099-1112) is published monthly by the American Society for Photogrammetry and Remote Sensing, 425 Barlow Place, Suite 210, Bethesda, Maryland 20814-2144. Periodicals postage paid at Bethesda, Maryland and at additional mailing offices.

**SUBSCRIPTION.** For the 2021 subscription year, ASPRS is offering two options to our *PE&RS* subscribers — an e-Subscription and the print edition. E-subscribers can plus-up their subscriptions with printed copies for a small additional charge. Print and Electronic subscriptions are on a calendar-year basis that runs from January through December. We recommend that customers who choose both e-Subscription and print (e-Subscription + Print) renew on a calendar-year basis.

The rate of the e-Subscription (digital) Site License Only for USA and Non-USA is \$1000.00 USD. e-Subscription (digital) Site License Only for Canada\* is \$1049.00 USD. e-Subscription (digital) Plus Print for USA is \$1365.00 USD. e-Subscription (digital) Plus Print for Canada\* is \$1424.00 USD. e-Subscription (digital) Plus Print for Non-USA is \$1395.00 USD. Printed-Subscription Only for USA is \$1065.00 USD. Printed-Subscription Only for Canada\* is \$1124.00 USD. Printed-Subscription Only for Non-USA is \$1195.00 USD. \*Note: e-Subscription/Printed-Subscription Only/e-Subscription Plus Print for Canada include 5% of the total amount for Canada's Goods and Services Tax (GST #135123065). **PLEASE NOTE: All Subscription Agencies receive a 20.00 USD discount.**

**POSTMASTER.** Send address changes to *PE&RS*, ASPRS Headquarters, 425 Barlow Place, Suite 210, Bethesda, Maryland 20814-2144. CDN CPM #/40020812)

**MEMBERSHIP.** Membership is open to any person actively engaged in the practice of photogrammetry, photointerpretation, remote sensing and geographic information systems; or who by means of education or profession is interested in the application or development of these arts and sciences. Membership is for one year, with renewal based on the anniversary date of the month joined. Membership Dues include a 12-month electronic subscription to *PE&RS*. Or you can receive the print copy of *PE&RS* Journal which is available to all member types for an additional fee of \$60.00 USD for shipping USA, \$65.00 USD for Canada, or \$75.00 USD for international shipping. Dues for ASPRS Members outside of the U.S. will now be the same as for members residing in the U.S. Annual dues for Regular members (Active Member) is \$150.00 USD; for Student members \$50.00 USD for USA and Canada \$53.00 USD; \$60.00 USD for Other Foreign members. A tax of 5% for Canada's Goods and Service Tax (GST #135123065) is applied to all members residing in Canada.

**COPYRIGHT 2021.** Copyright by the American Society for Photogrammetry and Remote Sensing. Reproduction of this issue or any part thereof (except short quotations for use in preparing technical and scientific papers) may be made only after obtaining the specific approval of the Managing Editor. The Society is not responsible for any statements made or opinions expressed in technical papers, advertisements, or other portions of this publication. Printed in the United States of America.

# SEMI-CENTENNIAL OF LANDSAT OBSERVATIONS & PENDING LANDSAT 9 LAUNCH

**Samuel N. Goward**

*Professor Emeritus,  
Department of Geographical Sciences,  
University of Maryland*

**Jeffrey G. Masek**

*Landsat Project Scientist,  
NASA Goddard Space Flight Center*

**Thomas R. Loveland**

*Retired Research Scientist,  
USGS Earth Resources Observation  
and Science (EROS) Center*

**John L. Dwyer**

*Research Scientist,  
USGS Earth Resources Observation  
and Science (EROS) Center*

**Darrel L. Williams**

*Chief Scientist,  
Global Science & Technology Inc.*

**Terry Arvidson**

*Retired Systems Engineer,  
Lockheed Martin Corp.*

**Laura E.P. Rocchio**

*Science Writer,  
Science Systems and Applications, Inc.*

**James R. Irons**

*Director, Earth Sciences Division,  
NASA Goddard Space Flight Center*

The first Landsat was placed in orbit on 23 July 1972, followed by a series of missions that have provided nearly continuous, two-satellite 8-day repeat image coverage of the Earth's land areas for the last half-century. These observations have substantially enhanced our understanding of the Earth's terrestrial dynamics, both as a major element of the Earth's physical system, the primary home of humans, and the major source of resources that support them. The history of Landsat is complex, reflective of the human systems that sustain it. Despite the conflicted perspectives surrounding the continuation of the program, Landsat has survived based on worldwide recognition of its critical contributions to understanding land dynamics, management of natural resources and Earth system science. Launch of Landsat 9 is anticipated in Fall 2021, and current planning for the next generation, Landsat Next is well underway. The community of Landsat data users is looking forward to another 50 years of the Landsat program.

## Introduction

Many technological advances emerged from the World War II, including rockets, electronic computing, and non-photographic sensors. By the mid-1950s these technologies set the stage for satellite monitoring of the Earth. First employed for intelligence gathering, some researchers also saw great potential for civilian applications, which until that time had primarily been serviced by aerial photography. Experimentation with airborne multispectral sensors suggested great potential to monitor land resources from space (Figure 1). Initial funding from US military agencies supported university and federal investigators in exploring applications in agriculture, forestry, hydrology, geology, and geography. University of Michigan, University of California, Berkley, and Purdue University were early leaders.

In the mid-1960s as NASA looked beyond the Apollo program, they became increasingly interested in Earth observation missions that would exploit technologies developed for Apollo. However, researchers from US Geological Survey (USGS), US Department of Agriculture (USDA), and the Army Corp of Engineers became more interested in free-flying land-monitoring missions. Early disagreements between NASA and USGS led Stewart Udall, the then Department of the Interior (DOI) Secretary—with the encouragement of USGS Director William Pecora, to announce in 1966 that DOI would develop and launch an Earth Resources Observation Satellite (EROS). After considerable Washington discussions, NASA was given the task to develop this satellite, called the Earth Resources Technology Satellite (ERTS) and later renamed Landsat.

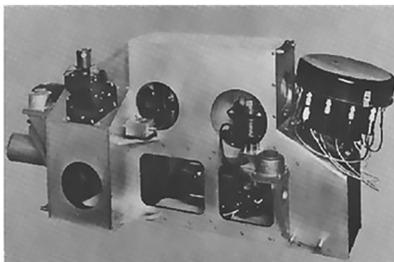
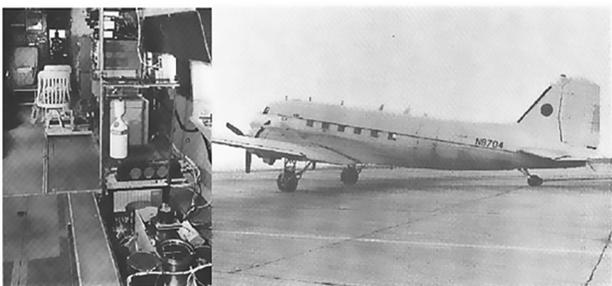


Figure 1. The M7 multispectral scanner (top; inspection plates removed) was mounted into the fuselage (bottom left) and flown aboard the ERIM C-47 aircraft (bottom right). Note the old operator chairs (bottom left). Multispectral images taken by the M7 were explored by researchers at the University of Michigan and Purdue University as an advanced means to monitor Earth resources. The M7 was a key inspiration for the MSS instruments flown on early Landsat satellites. (Photo credits: ERIM).



Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 533–539.

0099-1112/21/533–539

© 2021 American Society for Photogrammetry  
and Remote Sensing

doi: 10.14358/PERS.87.8.533

The primary source for this narrative is *Landsat's Enduring Legacy* published by the American Society for Photogrammetry and Remote Sensing (Goward *et al.*, 2017). Updates for Landsat 8 and the Landsat Archive can be found in the special issue of *Remote Sensing of Environment* (Loveland & Irons, 2016). Recent activities that have explored the future of the Landsat program have been detailed in the recent Congressional Research Service report *Landsat 9 and the Future of the Sustainable Land Imaging Program* (Normand, 2020) and RSE journal articles (Masek *et al.*, 2020).

## Early Landsat (L1-L3) and Agriculture

By 1967, NASA Goddard Space Flight Center (GSFC) had initiated the ERTS program with the goal of launching the first satellite by July 1972 (Figure 2). Potential users' first preference was to fly film-based camera systems that replicated aerial photography. However, the US intelligence community, already flying such systems, blocked their use for civilian applications. This left television technologies, such as the Return Beam Vidicon (RBV), used to find Apollo moon landing sites, or more novel multispectral scanners. Ultimately, NASA selected to fly both an RBV, developed by Radio Corporation of America (RCA), as the primary imaging instrument and the Multispectral Scanner System (MSS), developed by Hughes Aircraft, as a technical experiment. ERTS-1 was successfully launched on 23 July 1972. The remarkable quality of data returned from the digital MSS instrument astonished many. When an electrical short in the RBV nearly ended the mission shortly after launch, NASA made the decision to retire the RBV and make the MSS the primary imaging system (Figure 3).

Over the next three years, NASA funded an ERTS principal investigator (PI) program, the first such NASA program for the Earth science community. Over 300 PIs were selected to participate in this program. Meanwhile, the USGS initiated an ERTS archival system in Sioux Falls, South Dakota, the Earth Resources Observation System (EROS). In 1975, the ERTS program was renamed Landsat; Landsat 2 was launched, and NASA began studies to develop an advanced multispectral scanner system—the Thematic Mapper (TM). In 1978, Landsat 3 was placed into orbit.

The same year that Landsat 1 launched, the Union of Soviet Socialist Republics (USSR) concealed their major wheat crop failure and managed to purchase large portions of the US surplus wheat at low cost. When Congress heard of this intelligence lapse, they directed the USDA and NASA to engage the Landsat capabilities to monitor global wheat production; this became the multiagency Large Area Crop Inventory Experiment (LACIE) program. For the next decade, agriculture was the primary operational focus for the Landsat program.

## Thematic Mapper and Commercial Era (L4 & L5)

Landsat 4 launched on 16 July 1982 carrying a four-band MSS and a second-generation sensor, the Thematic Mapper (TM) (Figure 4). A sensor similar to the TM had been considered for Landsat 1, but nearly a decade of technological advances was needed to realize a flight-ready version. The TM had a 30m spatial resolution and seven spectral bands that expanded into blue, shortwave infrared, and thermal infrared wavelengths.

Communication and power issues plagued Landsat 4, which limited TM collection for the life of the mission. The identified issues were rectified before the Landsat 5 launch. These fixes, plus a large fuel tank installed for an unrealized Space Shuttle capture-and-repair capability, enabled Landsat 5 to collect data for 28.8 years, carrying the mission well past the 1993 Landsat 6 loss (Figure 5).



Figure 2. Artistic rendition of Landsats 1 (ERTS), 2 and 3. The early Landsats used the same satellite bus developed for the NASA Nimbus weather satellites, adapted for the new Landsat instruments. Only two years after the initial meetings with the selected Landsat contractors, Landsat 1 was ready for launch. (Image credit: NASA)

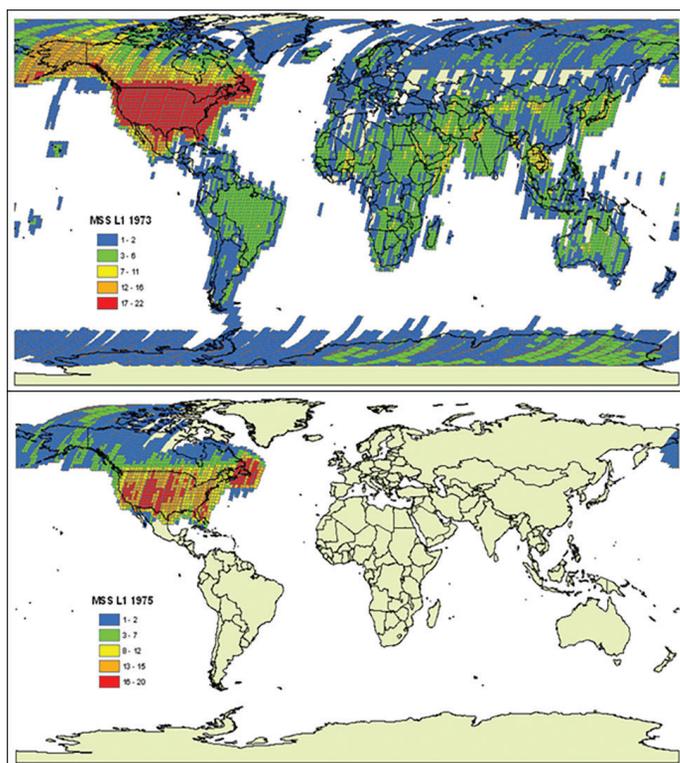


Figure 3. Landsat 1 (ERTS) coverage for 1973 (top) and 1975 (bottom). Onboard wide-band videotape recorders recorded image observations but generally failed after about 2.5 years. After this, only direct line-of-sight transmission to ground stations was possible, leading to only North American coverage via direct downlink in later years of the missions. (Image credit: USGS)



Figure 4. Artistic rendition of Landsats 4 and 5. The satellites had no onboard recorders; the large antenna on top of the satellite transmitted data from Landsat to TDRSS for global downlink. The first TDRS was placed over the Atlantic in 1983. The second TDRS was lost in the 1986 Challenger launch, so there was no ability for full global coverage via TDRSS until 1989. Additionally, Landsat’s commercialization and the higher priority of Space Shuttle operations substantially reduced global coverage via TDRSS. Fortunately, Landsat data captured by International Ground Stations during this time have now been added to the US Landsat archive. (Image credit: NASA)

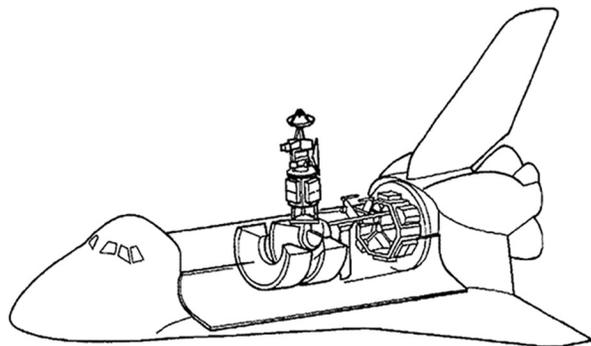


Figure 5. The proposed Shuttle Landsat retrieval system. To accomplish this feat, the Landsat satellite would need to be de-orbited to a possible Shuttle altitude and thus Landsats 4 and 5 were fitted with large fuel tanks. After the Challenger loss and the cancellation of Vandenberg Shuttle launch facilities, this retrieval was never attempted. However, the expanded fuel tank did permit Landsat 5 to continue operations for nearly 30 years, filling the gap in coverage that would likely have occurred after the loss of Landsat 6 during launch in 1993. (Image credit: NASA/Hughes SBRC)

Unlike weather satellites that were managed by the Federal government as a public good, Landsat followed the path of communication satellites to be spun-off to private industry for commercialization. The National Oceanographic and Atmospheric Administration (NOAA) had the unpopular task of recovering all Landsat operational costs via data sales and finding a commercial operator. EOSAT—the sole remaining bidder in a deeply flawed commercialization gambit—took control of the Landsat satellites in 1985. Despite lowering NOAA-set data prices, sales never took off. With no on-board recorders, Landsats 4 and 5 international acquisitions were dependent on either direct downlink to International Ground Stations or high-cost NASA tracking and data relay satellite system (TDRSS) transmissions, to the detriment of a global archive (Figure 6).

Despite the then-groundbreaking 30m resolution, data costs caused all but the best-funded scientists to migrate to NOAA’s no-cost Advanced Very High Resolution Radiometer (AVHRR) data with more frequent observation repeat cycles but much coarser spatial resolution. With these AVHRR data, researchers refocused their collective energy on the development of robust methodologies for global vegetation analyses.

Fortunately, the concept of the MSS Basic Data Set during this period cemented the idea of long-term Landsat data preservation and, when coupled with a growing awareness of global change ramifications and satellites’ monitoring role, helped propel Landsat data from a commodity back towards a public good.

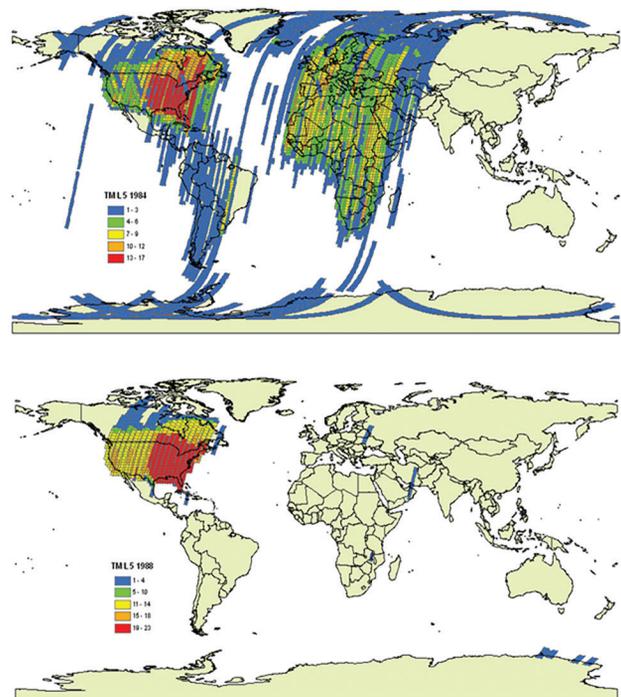


Figure 6. US Landsat 5 coverage for 1984 and 1988. Half global coverage was the result of losing the western TDRSS satellite during Challenger launch. By 1988 the satellite’s primary link to the TDRSS had failed and most acquisitions were being made by Landsat 4. EOSAT continued acquisitions over the US via ground-based antenna at NASA Goddard and exchange agreement with western Canadian station. (Image credit: USGS)

## Studying Global Change with Landsats 5 & 7

NASA's response to the 1990 Global Change Research Act was to focus its Earth Sciences division on the Mission to Planet Earth (MTPE) program and the development of the Earth Observation System (EOS). Landsat was integrated into MTPE as an EOS observatory. As such, MTPE supported The Landsat-based Pathfinder tropical deforestation and land cover change studies, investigation of follow-on Landsat technologies, and the USGS EROS rescue of deteriorating Landsat archive media. Later MTPE and USGS supported reprocessing and geodetic updating of global 1975, 1990, and 2000 Landsat-based EarthSat GeoCover data sets as well as adding 2005 and 2010 data. all of which were made freely available to users. With the at-launch loss Landsat 6 satellite in 1993, a primary goal of the MTPE Landsat program was to quickly develop and launch Landsat 7, no later than 1998 (Figure 7). The L7 design

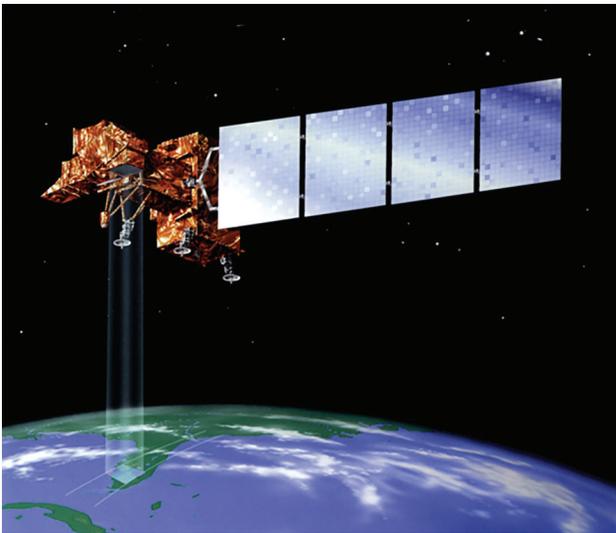


Figure 7. Artistic rendition of Landsat 7 imaging Florida. The three X-band antennae on the bottom of the satellite enabled transmission of image data to multiple ground stations at the same time. An onboard solid-state recorder enabled data collected outside of US ground station acquisition circles to be stored for later downlink and processing into the US archive. An automated Long-Term Acquisition Plan ensured global acquisition coverage priorities. (Image credit: NASA/USGS)

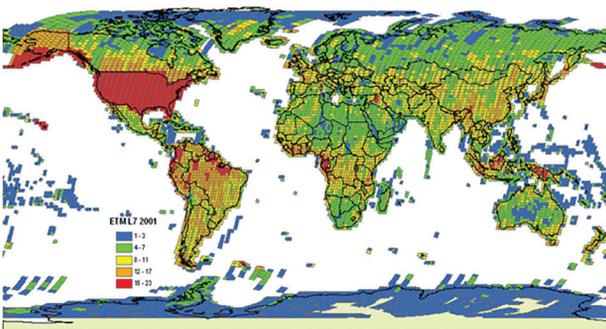


Figure 8. 2001 Landsat 7 global coverage. Landsat 7 was the first of the series to truly approach global coverage, including islands and reefs as well as the polar regions. This was thanks to technology improvements onboard and in the processing chain, as well as scheduling guidance from the Long-Term Acquisition Plan (LTAP). (Image credit: USGS)

included several satellite advances including a spacious solid-state recorder, the Enhanced Thematic Mapper Plus (ETM+) sensor with multiple calibration options, and increased downlink capacity. In addition, ground system included a high-speed large-capacity data processing, a dedicated Image Assessment System, and the Long-Term Acquisition Plan (LTAP) that implemented systematic global acquisition of data (Figure 8). Further, In 1996 MTPE funded formation of the first-ever Landsat Science Team (LST) in recognition of Landsat's new EOS role. The LST aided refinement of L7 operations as well as exploring specific PI-driven science goals. Unfortunately, the LST was abandoned in 2001 in anticipation of a second attempt to commercialize Landsat with Landsat 8. Throughout the operational life of Landsat 7 (1999-present), improvements have been made in acquisition rates (250 scenes/day to near 500), data accuracy, the routine provision of higher-level data products, and, no cost data access in December 2008. After return of Landsat 5 to government management in 2001, it was incorporated into the LTAP, thus maximizing the archive contents.

## Earth System Science with Landsats 8 & 9

The Landsat 8 era began with uncertainty and evolved towards a stable, sustainable strategy for follow-on Landsat missions (Figure 9). Following the launch of Landsat 7, NASA sequentially initiated and abandoned two implementation strategies for a follow-on mission before receiving direction to build a free-flyer satellite that was successfully launched as the Landsat Data Continuity Mission (LDCM) on 11 February 2013. The 1992 Land Remote Sensing Policy Act expressed a preference for a private sector follow-on to Landsat 7. In accordance with that preference, NASA first attempted to implement the follow-on mission as a data buy from a privately-owned and operated satellite and called the initiative the LDCM. NASA issued a Request for Proposals (RFP) and then declined to accept any of the tendered proposals in late 2003, deeming the proposed levels of public risk too high. The Office of Science and Technology Policy (OSTP) then in 2004 directed NASA to incorporate a Landsat



Figure 9. Artistic rendition of Landsat 8 in orbit over SE United States. Landsats 8 and 9 use the same satellite bus—a fourth-generation Landsat bus design. Improvements in communication bandwidth, science instrument design, and solid-state recorder capacity, and an updated LTAP produce coverage that substantially exceeds Landsat 7 coverage. The collection of nearly all possible images for much of the globe greatly enhances the ability to stitch together a less cloudy data base of imagery. (Image credit: NASA/USGS)

sensor into the ill-fated National Polar-Orbiting Operational Environmental Satellite System (NPOESS). With NPOESS encountering difficulties, OSTP revised its guidance in December 2005 and directed NASA to build a free-flyer follow-on to Landsat 7.

NASA retained the LDCM moniker and managed the development and launch of a satellite observatory consisting of a spacecraft bus procured from Orbital Sciences Corporation to carry a two-instrument payload, the Operational Land Imager (OLI) procured from Ball Aerospace and the Thermal InfraRed Sensor (TIRS) built by NASA Goddard. These two sensors were the first pushbroom imaging radiometers flown by a Landsat mission and offered improved radiometric performance relative to the whiskbroom sensors flown aboard prior Landsat satellites. LDCM was turned over to USGS for operation after in-orbit check out and USGS renamed the observatory Landsat 8.

Landsat 8 operates in a rapidly advancing Earth observation infrastructure. Where Landsat satellites were once the only civilian space assets designed for land observations, Landsat 8 now shares low Earth orbit with a proliferation of international and commercial satellites observing all components of the Earth system — the atmosphere, oceans, cryosphere, and continents. These global observations propel forward Earth system science where the Earth is studied as an integrated system of processes that encompass all the components.

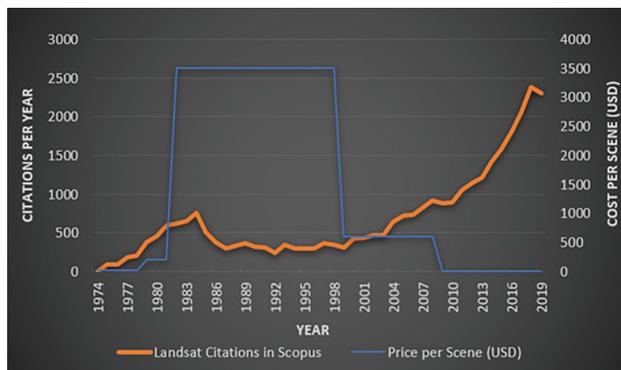


Figure 10. Number of Landsat-related citations since 1974. Citations greatly increased after the open data policy decision in December 2008. As of December 2020 over 100 million Landsat scenes have been downloaded from archive. For more details see Landsat Project Statistics ([usgs.gov](https://usgs.gov)) (Image credit: USGS)

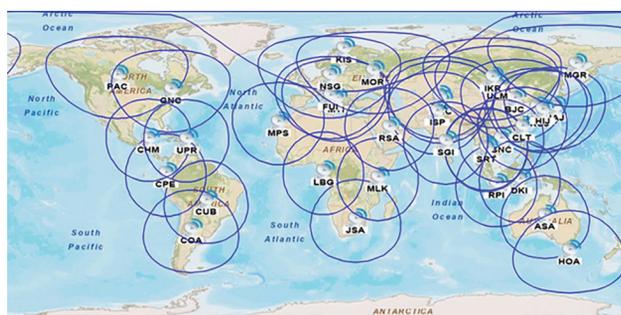


Figure 11. Locations of Landsat international ground stations from 1972 to current. The number of ground stations active at any given time was quite variable. For details concerning specific ground stations see <https://landsat.usgs.gov/historical-international-ground-stations> (Image credit USGS)

The long delay between the Landsat 7 and Landsat 8 launches led NASA and USGS to conceive a Sustainable Land Imaging (SLI) program to define and plan for future missions. The first SLI mission architecture study resulted in a decision to build Landsat 9 as a near-identical copy of the Landsat 8 observatory with a launch planned for Fall 2021. Landsat 9 will contribute to Earth system science by continuing to build upon the now five-decade-long record of Landsat observations.

### Advancing the Landsat Archive

Two decisions, made in the early days of the LDCM development, had major impacts on the use of Landsat. First, the USGS and NASA ratified a change to the Landsat Data Policy that resulted in the provision of non-discriminatory access to free orthorectified Landsat data. Second, USGS and NASA reconvened a Landsat Science Team (LST) in 2006 and every 5 years since that time.

The 2008 opening of the archive transformed Landsat data use (Figure 10). Landsat time series analyses for monitoring landscape change are now commonplace. With LDCM largely defined and underway, the LST tackled the role of community needs and defining science-based input that improve uses of the Landsat archive.

The LST focused on ideas to expand the historical Landsat archive, resulting in the Landsat Global Archive Consolidation initiative that started repatriation of all Landsat 1–7 data held by International Ground Stations (IGs) (Figure 11). With full cooperation from the IGs, over 5.41 million additional, unique historical scenes are now available to Landsat users worldwide (Figure 12).

Landsat 7 and 8 acquisition strategies pushed the technical capabilities of both missions doubling the daily growth of the archive. Landsat 7 now collects over 500 scenes per day and Landsat 8 collects up to 740 scenes per day. In addition, the SLI long-term commitment that includes the planned 2021 launch of Landsat 9 will expand the aggressive global acquisition strategy.

The USGS, with LST input, also focused on ideas for improving the usability of the archive. Community R&D led to major improvements in the geometric correction and geo-location of all Landsat data. Instrument cross calibration and measurement uncertainties were defined and applied to improve radiometric consistency and comparability for all 9 million Landsat 1-8 archive scenes. The USGS implemented and applied these advances using a collection management strategy that standardized and controlled the characteristics of Level 1 products and added derived Level 2 geophysical products and pixel-based metadata. The Landsat program is moving toward providing “analysis-ready” virtual data cubes that will further improve the efficiency and consistency of time series studies. The entire archive has recently been moved to a cloud environment that will ensure more timely reprocessing and more robust access to the archive.

### Landsat Next

Recognizing that Landsat observations continue to provide unique and critical information for science, the US Government has made a long-term commitment to continue

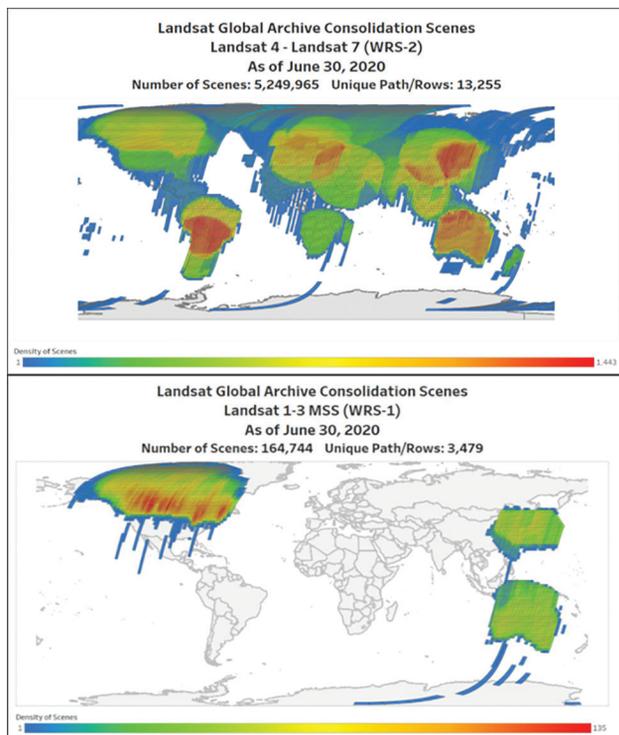


Figure 12. International station contributions to global archive consolidation. As of June 2020, the newly incorporated international scenes were nearly double the number of the previous holdings from Landsat 1–7. The unique path-rows are locations that had not been previously imaged for the US archive. As of September 2020, the LGAC scenes more than doubled the previous Landsat 1-7 holdings. (Image credit: USGS)

these measurements through the Sustainable Land Imaging program. Not surprisingly, both technology and user expectations have evolved since the Landsat 8/9 systems were scoped in the early 2000s. User surveys by USGS have revealed a strong preference for improved spatial resolution, more frequent observations, and additional spectral channels to support new applications such as water quality and cryospheric studies. In part, these findings reflect the influence of the operational European Sentinel-2 satellite series, which has acquired Landsat-like images with finer spatial resolution and greater frequency than the current Landsat program. At the same time, reports from the National Research Council and USGS Landsat Advisory Group have urged NASA and USGS to pursue more innovative ways of implementing Landsat missions, including exploration of commercial and international partnerships.

NASA and USGS have recently advanced the Landsat Next concept centered on collection of a “superspectral” data set encompassing some 25 spectral bands, with a spatial resolution of 10-20m (for the reflective bands) or 60m (for the thermal infrared). This mission, to be launched as early as the late-2020s, would significantly advance Landsat observational capabilities and provide a robust complement both to other international imaging systems, as well as to upcoming hyperspectral missions such as the NASA Surface Biology and Geology (SBG) mission. While the exact architecture of the Landsat Next mission will not be finalized until late 2021, non-traditional approaches, including acquiring the data set via a constellation of smaller satellites, are being explored.

### Summary

The Landsat program over the last half-century has created an unparalleled observation record of the state and dynamics of Earth’s land conditions. The major signals captured in these measurements document how life-sustaining biological activity varies in space and time. We have learned how to separate atmospheric and land surface signals to provide highly accurate estimates of this activity, which provides important insights into Earth system processes and humans’ role in modulating and managing this activity. The US has preserved this critical historical record of land processes that has already produced major new insights into how Earth sustains life on this planet. Continuation of this record into the future and enhancement of these measures is of vital importance today and for future generations.

### References

Goward, S. N., Williams, D. L., Irons, J. R., Arvidson, T., Rocchio, L., Russell, C., & Johnston, S. (2017). *Landsat’s Enduring Legacy: Pioneering Global Land Observations from Space*. Bethesda, Maryland: American Society of Photogrammetry and Remote Sensing.

Loveland, T. R., & Irons, J. R. (2016). Landsat 8 Science Results Special Issue. *Remote Sensing of Environment*, 185, 1-284.

Masek, J. G., Wulder, M.A., Markham, B.L., McCorkel, J., Crawford, C.J., Storey, J., and Jenstrom, D.T. (2020). Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, 248.

Normand, A. E. (2020). *Landsat 9 and the Future of the Sustainable Land Imaging Program* (R46560). Retrieved from Washington, DC: <https://crsreports.congress.gov/R46560>.

# Call for Submissions

## Remote Sensing Monitoring for Urban Environment

*Photogrammetric Engineering and Remote Sensing (PE&RS)* is seeking submissions for a special issue on Remote Sensing Monitoring for Urban Environment.

Urban remote sensing provides images with multiple spatio-temporal-spectral attributes, which can provide qualitative, quantitative, dynamic and comprehensive information and support for urban environmental monitoring and evaluation, and serve urban planning and management, ecological environment protection. In recent decades, global urban areas have been rapidly expanding, especially in developing countries. Rapid urbanization, along with manufacturing industries and large number of vehicles has resulted in serious environmental problems, called “urban diseases”, including increased vulnerability to natural hazards, natural vegetation cover decline and arable land loss, urban heat islands, air pollution, hydrological circle alteration and biotic homogenization. Urban ecosystems are strongly influenced by anthropogenic activities. Considering this, this special issue of PE&RS is aimed at reporting novel studies on exploiting remote sensing big data to monitor and improve urban environment, and showing the potential of remote sensing in developing sustainable cities, including but not limited to:

- Urban thermal-environment remote sensing
- Remote sensing image acquisition and processing for urban environment
- Remote sensing dynamic monitoring of urban expansion
- Remote sensing change detection of urbanization
- Remote sensing retrieval of urban ecological environment
- Remote sensing evaluation of urban human settlements
- Urban sustainability indicators and assessment
- Urban environmental monitoring

Papers must be original contributions, not previously published or submitted to other journals. Submissions based on previous published or submitted conference papers may be considered provided they are considerably improved and

extended. Papers must follow the instructions for authors at <http://asprs-pers.edmgr.com/>.

### Important Dates (Tentative)

- March 1, 2021—Submission system opening
- September 31, 2021—Submission system closing
- Planned publication date is December 2021
- Submit your manuscript to <http://asprs-pers.edmgr.com/> by September 31, 2021.

### Guest Editors

#### Zhenfeng Shao, *Wuhan University, China*

Prof. Zhenfeng Shao, Professor at the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, China. His research interests include urban remote sensing. He is now an associate editor of Email: [shaozhenfeng@whu.edu.cn](mailto:shaozhenfeng@whu.edu.cn).

#### Orhan Altan, *Istanbul Technical University, Turkey*

Prof. Orhan ALTAN, Professor at the Department of Geomatics Engineering, Istanbul Technical University, Turkey. He is Past President and Honorary Member of ISPRS, Honorary Member of Science Academy. He has published more than 200 scientific papers in scientific journals and conferences, and editor or co-editor of more than 20 international books. Email: [oaltan@itu.edu.tr](mailto:oaltan@itu.edu.tr).

#### J.L.van Genderen, *University of Twente, Netherlands*

Professor J.L.van Genderen, Professor at the Department of Earth Observation Science, Faculty of Geo-information Science and Earth Observation(ITC),University of Twente, Netherlands, he is an Associate Editor of the international journal GEO Spatial Information Science. His research interests include urban remote sensing and new technologies for urban management. Email: [Genderen@alumni.itc.nl](mailto:Genderen@alumni.itc.nl).

---

## NEW ASPRS MEMBERS

ASPRS would like to welcome the following new members!

Dhananjay Agrawal  
Michael S. Anders  
James Armstrong  
Chuck Boyer  
Jonathan Byham  
Alan Chace, PLS  
Peter G. Chirico  
Martha S. Compton, CMS  
Jean-Francois Dionne, GISP  
Dustin Frey  
Kristopher Gallagher

Isidro Xavier Garza  
Chris Ghormley  
John Joseph Hubler  
Jared Janacek  
Jiyeon Kim  
Derege Kobre  
Chu Kei Kwok  
Kyle Lloyd  
Kevin F. May, CMS  
Jeffrey Curtis Miller  
Hector Ricardo Padilla, GISP

John Lyman Reynolds  
Jenny Lynn Roberts-Axelrod  
Nathan Rowley  
Melissa J. (Tolene) Rura-Porterfield, Ph.D.  
Leilani Scott Young  
David Shean  
Will Siepmann  
Brad Smith  
Taylor Sparrow  
Emma Thomas  
Scott Porter Walls

FOR MORE INFORMATION ON ASPRS MEMBERSHIP, VISIT [HTTP://WWW.ASPRS.ORG/JOIN-NOW](http://www.asprs.org/join-now)

## You don't have to accept the defaults in GlobalMapper

The first thing that I tell my beginning GIS students is, “Never accept the defaults”. GIS software programs like ArcMap, QGIS, and GlobalMapper randomly choose colors, line weights, etc. when a new layer is added to a map. While the software’s defaults are easy enough to modify for vector layers, changing the default color ramps for raster layers usually requires multiple steps.

Rasters are most valuable at helping us visualize things that vary continually through space, like temperature, rain-fall, soil moisture, and elevation. These types of variables frequently are associated with “standard” color ramps. For example, we associate areas of high temperatures with reds and oranges, while lower temperatures are usually depicted in blues and greens. Likewise, we generally associate lower elevations with blues and greens, while higher elevations are symbolized with browns and whites as in the “standard” US Geological Survey topographic maps. So, when a raster is loaded into mapping software, the program tries to assign the most likely color ramp. For example, when an elevation raster is loaded into GlobalMapper, the color ramp will generally default to a blue (= lower elevations) to reds (=higher elevations). In ArcGIS, either Desktop or Pro, the ramp defaults to a gray-scale for all non-image rasters.

One workflow used at Dewberry to review lidar surfaces for poor ground filtering, noise, and other anomalies, is to construct an elevation raster for manual inspection. In Brittany’s work, as assistant quality control manager, she inspects these lidar-derived rasters using GlobalMapper. The “default” color ramp works well when there is considerable elevation variation in the surface, but when inspecting relatively flat areas, especially with low-lying vegetation, the slight differences in color are easily overlooked. So, here is her tip for constructing a custom color ramp to help visualize small differences in elevation.

In Figure 1, an elevation raster was loaded into GlobalMapper. The software displayed the raster in the standard blue to red elevation color ramp, but because the area of interest shown in the figure is generally so flat (this is a portion of a Digital Elevation Model in Florida), there is not much differentiation in color over almost a square mile of this elevation raster. So, to understand this area better, she needed to adjust the color ramp (Remember: Never accept the defaults.)

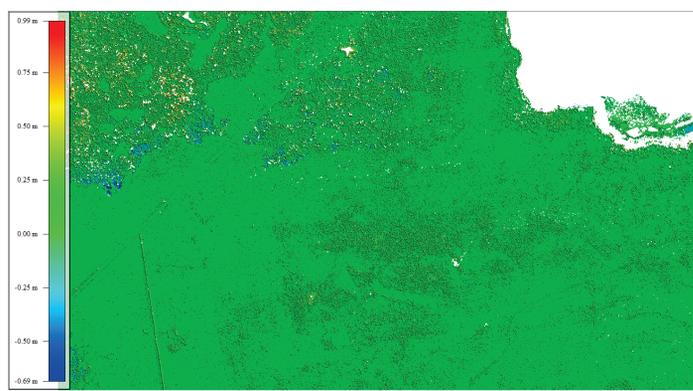


Figure 1. GlobalMapper default color ramp for a Digital Elevation Model in Florida.

As with most GIS software, there are multiple ways to change the symbology/color ramp and other display parameters (ex., the units, transparency, etc.), so while this is not the only way, here is a simple, six-step, workflow in GlobalMapper:

**Step 1 —**  
Right-clicking on the elevation legend as in Figure 2 starts the dialog box that allows you to change the units displaying the data in the two lower lines (Metric Units or Statute Units), and by left-clicking “Elevation Legend Options...” starts the “Configuration – Elevation Legend” (Figure 3) dialog.

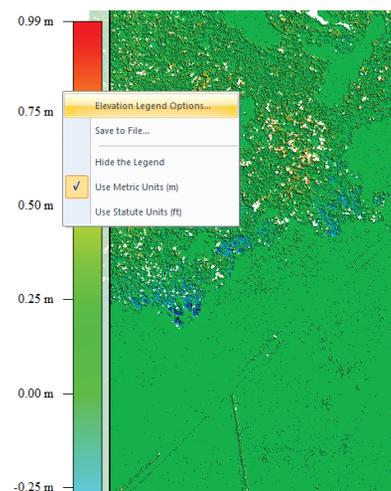


Figure 2. “Elevation Legend Options...” dialog box. (Note: this is one place where you can change between metric and imperial units.)

**Step 2** — Once the “Configure - Elevation Legend” dialog opens, use the options on the white portion of the dialog box to select “Display Options” and then the radio button to manipulate the “Elevation Legend” (gray side). There are many style changes that can be made! In this example, choose “Shader Options” under “Point Styles”. (Note: the top line of gray side of this dialog box provides another opportunity to change the legend units.)

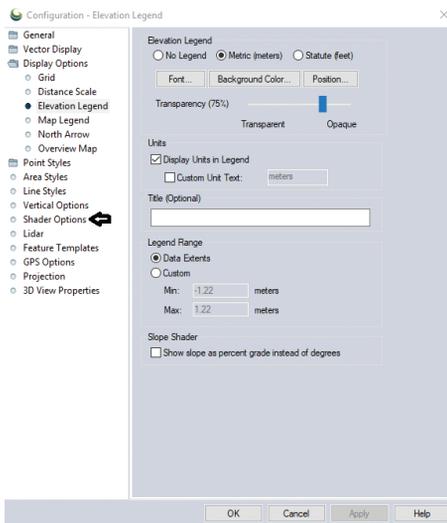


Figure 3. The “Configuration – Elevation Legend” dialog showing where to click to customize the color ramp.

**Step 3** — Choosing <new> at the bottom of the dialog box (Figure 4) will start the “Custom Shader” dialog. Enter a name for the new custom shader. In this case, we entered “My-CustomShader”. This shader file will be saved for future use.

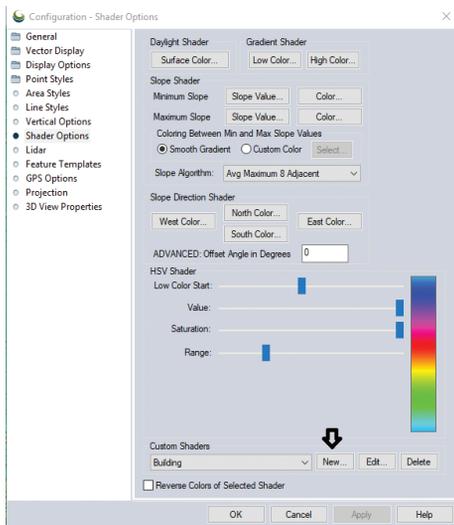


Figure 4. The “Configuration – Custom Shader” dialog box.

**Step 4** — From the “Custom Shader” dialog, the new filename will autopopulate (1), then select <Initialize from Other Shader> (2) and then use the dropdown to select “Global Shader” (3) as in Fig 5. Press <OK> to complete this step.

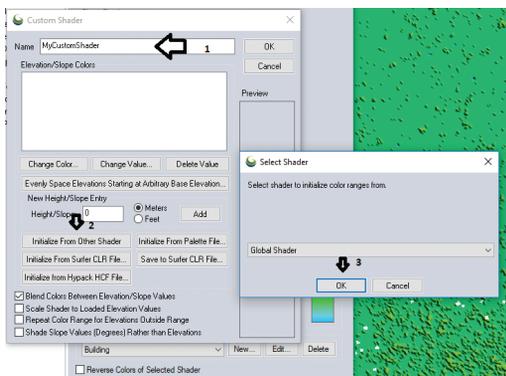


Figure 5. Dialog to construct your Custom Shader.

**Step 5** — The custom shader will load the Global Shader color ramp which you will customize by selecting each color, then the <Change Color> button and select a color from the color palette, and then the <Change Value> button and enter the elevation value to associate with the chosen color. The Custom Shader defaults as Figure 6, so select and change each color and value until the shader appears as Figure 7. You will need to add values by entering the value (in the Height/Slope box) and then pressing the <Add> button to select a color.

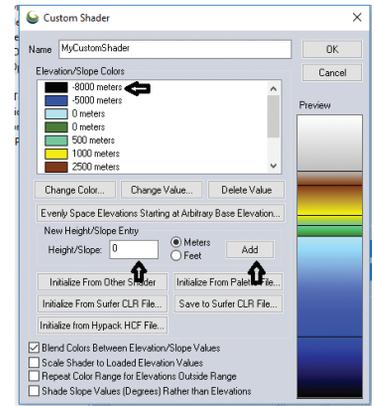


Figure 6. Starting configuration for Custom Shader.

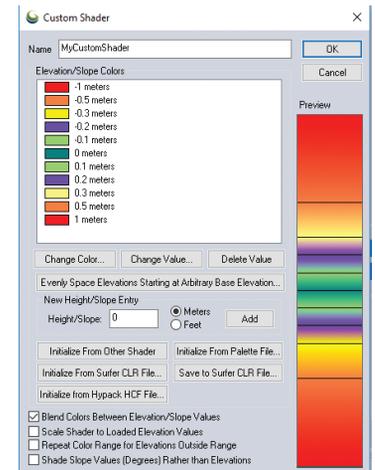


Figure 7. Ending configuration for Custom Shader.

Back on the Configuration – Shader Options dialog, press <Apply> and <OK> to complete this step.

**Step 6** — To apply the new “MyCustomShader”, right-click on the legend to return to the “Configuration – Elevation Legend” dialog, select the “Shader Options” (Figure 8.) and use the dropdown selector to find “MyCustomShader”. Press <Apply> and <OK> to complete this step.

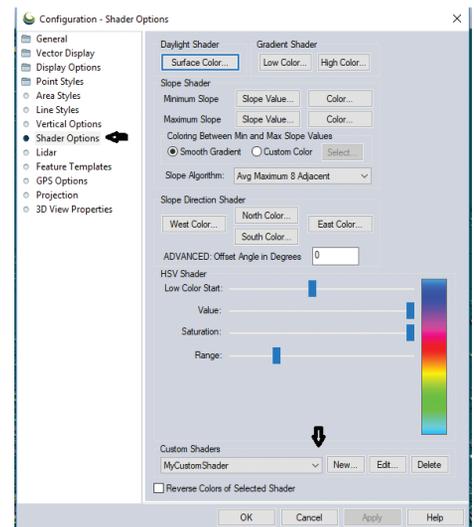


Figure 8. Selecting the custom display ramp on the Shader Options menu.

Finally, you may need to set the Minimum and Maximum Elevation Ranges for display on the Display Options menu (Figure 9).

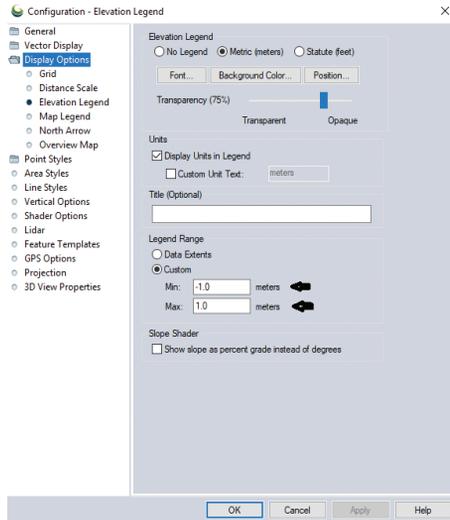


Figure 9. Setting Minimum and Maximum Elevation Ranges.

And... Voila, the vegetation (yellow) now stands out from the ground (blues)...



Figure 10. Custom color ramp applied to the Digital Elevation Model in Figure 1.

Send your questions, comments, and tips to [GISTT@asprs.org](mailto:GISTT@asprs.org).

Brittany Capra and Al Karlin, Ph.D., CMS-L, GISP are with Dewberry's Geospatial and Technology Services group in Tampa, FL. Brittany is an assistant quality manager working with landscape-level lidar projects. As a senior geospatial scientist, Al works with all aspects of lidar, remote sensing, photogrammetry, and GIS-related projects.

## STAND OUT FROM THE REST

### EARN ASPRS CERTIFICATION

ASPRS congratulates these recently Certified and Re-certified individuals:

#### CERTIFIED PHOTOGRAMMETRIST

**Volkan Akbay, Certification # CP1658**

effective March 12 2021, expires March 12 2026

#### RECERTIFIED PHOTOGRAMMETRIST

**Trevi Gigliotti, Certification # R1492CP**

effective May 31 2021, expires May 31 2026

**Michael Loose, Certification # R1199CP**

effective May 8 2021, expires May 8 2025

**Timothy McMillan, Certification # R1438CP**

effective Jan 25 2021, expires Jan 25 2026

**Christine Wasicek, Certification # R1296CP**

effective April 6 2021, expires April 6 2026

**David Holm, Certification # R1301CP**

effective July 27 2021, expires July 27 2026

**Michael Brian Wernau, Certification # R1474CP**

effective Dec 15 2020, expires Dec 15 2025

#### RECERTIFIED MAPPING SCIENTIST – REMOTE SENSING

**Steven Lennartz, Certification # R185RS**

effective Dec 6 2020, expires Dec 6 2025

#### CERTIFIED MAPPING TECHNOLOGIST GIS/LIS

**Jeffrey Williams, Certification # GST307**

effective April 6 2021, expires April 6 2024

#### RECERTIFIED MAPPING SCIENTISTS LIDAR

**Al Karlin, Certification # R007L**

effective April 19 2021, expires April 19 2026

#### CERTIFIED MAPPING SCIENTISTS LIDAR

**Aaron Hofstedt, Certification # L066**

effective April 15 2021, expires April 15 2026

ASPRS Certification validates your professional practice and experience. It differentiates you from others in the profession. For more information on the ASPRS Certification program: contact [certification@asprs.org](mailto:certification@asprs.org), visit <https://www.asprs.org/general/asprs-certification-program.html>.



# SUSTAINING MEMBERS

## ACI USA Inc.

Weston, Florida  
<https://acicorporation.com/>  
 Member Since: 2/2018

## Aerial Services, Inc.

Cedar Falls, Iowa  
[www.AerialServicesInc.com](http://www.AerialServicesInc.com)  
 Member Since: 5/2001

## Ayres Associates

Madison, Wisconsin  
[www.AyresAssociates.com](http://www.AyresAssociates.com)  
 Member Since: 1/1953

## Dewberry

Fairfax, Virginia  
[www.dewberry.com](http://www.dewberry.com)  
 Member Since: 1/1985

## Environmental Research Incorporated

Linden, Virginia  
[www.eri.us.com](http://www.eri.us.com)  
 Member Since: 8/2008

## Esri

Redlands, California  
[www.esri.com](http://www.esri.com)  
 Member Since: 1/1987

## F.R. Aleman & Associates, Inc.

Miami, Florida  
<https://fr-aleman.com>  
 Member Since: 7/2020

## GeoCue Group

Madison, Alabama  
<http://www.geocue.com>  
 Member Since: 10/2003

## Geographic Imperatives LLC

Centennial, Colorado  
<http://geographicimperativesllc.com>  
 Member Since: 9/2021

## GeoWing Mapping, Inc.

Richmond, California  
[www.geowingmapping.com](http://www.geowingmapping.com)  
 Member Since: 12/2016

## GPI Geospatial Inc.

*formerly Aerial Cartographics of America, Inc. (ACA)*  
 Orlando, Florida  
[www.aca-net.com](http://www.aca-net.com)  
 Member Since: 10/1994

## Green Grid Inc.

San Ramon, California  
[www.greengridinc.com](http://www.greengridinc.com)  
 Member Since: 1/2020

## Half Associates, Inc.

Richardson, Texas  
[www.half.com](http://www.half.com)  
 Member Since: 8/2021

## Keystone Aerial Surveys, Inc.

Philadelphia, Pennsylvania  
[www.kasurveys.com](http://www.kasurveys.com)  
 Member Since: 1/1985

## Kucera International

Willoughby, Ohio  
[www.kucerainternational.com](http://www.kucerainternational.com)  
 Member Since: 1/1992

## L3Harris Corporation

Broomfield, Colorado  
[www.harris.com](http://www.harris.com)  
 Member Since: 6/2008

## NV5 Geospatial

Sheboygan Falls, Wisconsin  
[www.quantumspatial.com](http://www.quantumspatial.com)  
 Member Since: 1/1974

## Pickett and Associates, Inc.

Bartow, Florida  
[www.pickettusa.com](http://www.pickettusa.com)  
 Member Since: 4/2007

## Robinson Aerial Surveys, Inc.(RAS)

Hackettstown, New Jersey  
[www.robinsonaerial.com](http://www.robinsonaerial.com)  
 Member Since: 1/1954

## Scorpius Imagery Inc.

Newark, Delaware  
[aerial@scorpiusimagery.com](mailto:aerial@scorpiusimagery.com)  
 Member Since: 6/2021

## Sanborn Map Company

Colorado Springs, Colorado  
[www.sanborn.com](http://www.sanborn.com)  
 Member Since: 10/1984

## Surdex Corporation

Chesterfield, Missouri  
[www.surdex.com](http://www.surdex.com)  
 Member Since: 12/2011

## Surveying And Mapping, LLC (SAM)

Austin, Texas  
[www.sam.biz](http://www.sam.biz)  
 Member Since: 12/2005

## T3 Global Strategies, Inc.

Bridgeville, Pennsylvania  
<https://t3gs.com/>  
 Member Since: 6/2020

## Terra Remote Sensing (USA) Inc.

Bellevue, Washington  
[www.terraremote.com](http://www.terraremote.com)  
 Member Since: 11/2016

## Towill, Inc.

San Francisco, California  
[www.towill.com](http://www.towill.com)  
 Member Since: 1/1952

## Wingtra

Zurich, Switzerland  
<https://wingtra.com/>  
 Member Since: 6/2020

## Woolpert LLP

Dayton, Ohio  
[www.woolpert.com](http://www.woolpert.com)  
 Member Since: 1/1985

# SUSTAINING MEMBER BENEFITS

## Membership

- ✓ Provides a means for dissemination of new information
- ✓ Encourages an exchange of ideas and communication
- ✓ Offers prime exposure for companies

## Benefits of an ASPRS Membership

- Complimentary and discounted Employee Membership\*
- E-mail blast to full ASPRS membership\*
- Professional Certification Application fee discount for any employee
- Member price for ASPRS publications
- Discount on group registration to ASPRS virtual conferences
- Sustaining Member company listing in ASPRS directory/website
- Hot link to company website from Sustaining Member company listing page on ASPRS website
- Press Release Priority Listing in *PE&RS* Industry News
- Priority publishing of Highlight Articles in *PE&RS* plus, 20% discount off cover fee
- Discount on *PE&RS* advertising
- Exhibit discounts at ASPRS sponsored conferences (exception ASPRS/ILMF)
- Free training webinar registrations per year\*
- Discount on additional training webinar registrations for employees
- Discount for each new SMC member brought on board (Discount for first year only)

\*quantity depends on membership level

**Are you...**

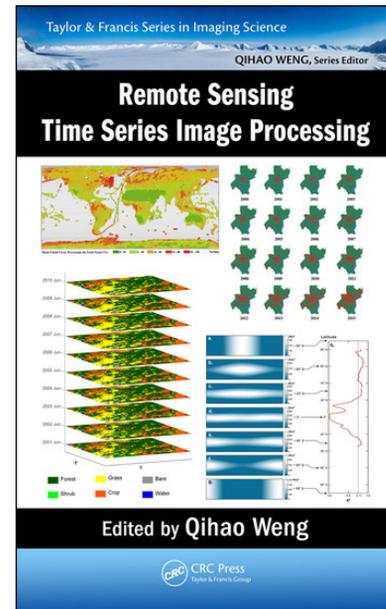
- » Engaged in Earth observation science?
- » Working on multitemporal change detection?
- » Mining data from disparate imagery archives and libraries?
- » Interested in image time series fundamentals and techniques?
- » Concerned with repairing or mitigating image artifacts and defects?
- » Reconstructing parts of images affected by contaminated pixels?
- » Seeking to create spatiotemporal data fusion products?

**If you answered “yes” to any of these questions, then this book is for you!**

In terms of current interest, the expression “change detection” signifies one of the premier applications of remote sensing. This book makes a momentous and substantial contribution showcasing the fundamentals of select processing techniques used with imagery time series at both coarse and fine spatiotemporal resolutions. In doing so, it also provides substantive examples of real-world applications for some of the algorithms. Often, as exemplified in this book, the existence of disparate sensor data at varying spectral, temporal, and spatial resolutions results in the creation of synthetic or fusion images and simulated time series.

This work sports a highly distinguished and recognized editor (Dr. Qihao Weng), and a list of twenty-five contributing authors and co-authors, responsible for the book’s ten chapters. On the Preface, a gem in and of itself, Dr. Weng outlines his vision for the direction in which remote sensing and imaging science are headed and provides a synopsis of the book. Written with impressive clarity, using precise and rich language, this must-read section provides an excellent opening. It paints such a clear and illustrative picture of the state-of-the-practice that it could well be used to guide the development of a course on the topic.

The book is laid out in three parts, the first part of which deals with the fundamentals of time series image and data generation and it contains three chapters: **Chapter 1**—*Cloud and Cloud Shadow Detection for Landsat Images: The Fundamental Basis for Analyzing Landsat Time Series*; **Chapter 2**—*An Automatic System for Reconstructing High-Quality Seasonal Landsat Time Series*; and **Chapter 3**—*Spatiotemporal Data Fusion to Generate Synthetic High Spatial and Temporal Resolution Satellite Images*. Chapter 1 starts out the leading thread for this section, with a categorization of algorithms to deal with specific image contamina-



## Remote Sensing Time Series Image Processing. First Edition

Qihao Weng (Editor)

263 pages; 24 color and 64 black and white illustrations. 2018. CRC Press. ISBN-13: 978-0367571795, ISBN-10: 0-36757179X. Hardback, paperback, and eBook.

Demetrio P. Zourarakis, PhD, GISP, CMS (GIS, RS, Lidar), Adjunct Assistant Professor, Dept. of Plant and Soil Sciences, University of Kentucky

tion such as clouds and cloud shadows. Chapter 2 focuses on an automated method for the reconstruction of contaminated pixels (i.e., due to clouds, cloud shadows, and SLC gaps). Chapter 3 delves into spatiotemporal data fusion for working with coarse spatial resolution/fine temporal resolution time series with fine spatial resolution/coarse temporal resolution data. Among the several different techniques contrasted, two recent algorithms are introduced in detail and evaluated, namely the NDVI Linear Mixing Growth Model (NDVI-LMGM) and Flexible Spatiotemporal DATA Fusion method (FSDAF).

The second part of the book examines topics related to feature development and information extraction; **Chapter 4**—*Phenological Inference from Times Series Remote Sensing*

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 545–546.

0099-1112/21/545–546

© 2021 American Society for Photogrammetry  
and Remote Sensing

doi: 10.14358/PERS.87.8.545

*Data*, **Chapter 5**—*Time Series Analysis of Moderate Resolution Land Surface Temperatures*, and **Chapter 6**—*Impervious Surface Estimation by Integrated Use of Landsat and MODIS Time Series in Wuhan, China*. Chapter 4 focuses on phenology, specifically on spectral-temporal trajectories of greenness indicators, for either single-year or multi-season, at a local or regional level. Examples of greenness analysis provided include crops in agroecosystems, forestry, and flooded landscapes. Chapter 5 exemplifies work with land surface temperature (LST) by using the data from collection-5 and collection-6 daily Aqua and Terra MODIS LST time series, with an application to surface urban heat islands (SUHI). Finally, chapter 6 shows an application of the Biophysical Composition Index (BCI) derived from a pre-processed Landsat data as a basis for mapping the annual evolution of impervious surfaces.

Finally, the third, longest and final part covers image applications of time series and contains four chapters. **Chapter 7**—*Mapping Land Cover Trajectories Using Monthly MODIS Time Series from 2001 to 2010*, **Chapter 8**—*Creating a Robust Reference Dataset for Large Area Time Series Disturbance Classification*, **Chapter 9**—*A General Workflow for Mapping Forest Disturbance History Using Pixel Based Time Series Analysis*, and **Chapter 10**—*Monitoring Annual Vegetated Land Loss to Urbanization with Landsat Archive: A Case Study in Shanghai, China*. Chapter 7 tackles the tough issues of what constitutes true land cover change and what the correct temporal interval for its determination should be, so as not to miss it – e.g., inter-annual versus intra-annual. The existence and availability of “hyper-temporal” datasets does not facilitate the task. The authors make use of the concept of “adaptive time series” to arrive at a temporal segmentation of a continuous 10-year time series. The land cover trajectories were reconstructed at 32-day intervals where most of the land cover changes that are missed or observed later than actual times in annual mapping practice can be captured.

The development of a robust reference dataset as a critical pre-requisite for meaningful spatiotemporal disturbance analysis is shown in Chapter 8. In it, trained interpreters create such dataset by using Google Earth imagery, the pixel’s trajectory as extracted from annual composites, local expert knowledge, and regional spatial datasets. To demonstrate the utility of the reference dataset, a subset is then used in a machine learning environment to produce classified disturbance maps over a 28-year period according to agent and severity categories. Chapter 9 illustrates the challenges, lessons learned and implications for future studies resulting

from quantifying temporal changes in forested lands represented by a wall-to-wall forest disturbance mapping effort at a national scale. The main factors analyzed included: automating image selection, compositing to overcome cloud-contaminated observations, achieving consistency across diverse landscapes, handling the large data volumes, and identifying and addressing quality assessment issues while under production, among others. Finally, Chapter 10 explores a workflow to quantify urban expansion at the expense of agricultural/vegetated land. Annual vegetated land conversion to urbanization information is extracted by using dense Landsat time series image stacks, creating 11 continuous years’ worth of Normalized Difference Vegetation Index (NDVI) mosaics from cloud/shadow free images. Computations were carried out on the Google Earth Engine (GEE) cloud-computing platform for large area applications.

This being my first review of a book in digital format, having access to it via a secure site and choosing a device with a screen of adequate size to accommodate the book at a whole-page scale readability were important factors. Now, something about the technology needed to access the book, depending on which version you purchase. This edition is available as paperback, hardback or digitally as a *VitalSource eBook* (<https://www.routledge.com/Remote-Sensing-Time-Series-Image-Processing/Weng/pl/book/9781138054592>). At the time of this writing, the eBook format can be either purchased at \$49.45 or rented for six months (\$27.48) or twelve months (\$32.97). If purchased or rented as a *Routledge & CRC Press eBook*, it can be accessed ubiquitously through the free *VitalSource Bookshelf® app*(lication), onto online/offline desktop, laptop or mobile/*eReader* devices. Depending on how you access the *eBook* (via *Bookshelf Online* in your browser or via the *Bookshelf app* on your PC or mobile device), this *VitalSource eBook* can be in a “re-flowable” **EPUB** format which allows you to resize text to suit you and enables other accessibility features. Where the content of the *eBook* requires a specific layout, or contains mathematical or other special characters, the eBook will be available in **PDF (PBK)** format, which cannot be re-flowed. For both formats the functionality available will depend on how you access the **eBook**.

In essence, the existence of a digital version of this highly recommended book, accessible anywhere/anytime reinforces the feasibility of its use in courses delivered virtually – a modality in vogue these days and which will likely remain an option for teaching and learning in the foreseeable future.



# & GRIDS & DATUMS

## REPUBLIC OF YEMEN

BY Clifford J. Mugnier, CP, CMS, FASPRS

The Grids & Datums column has completed an exploration of every country on the Earth. For those who did not get to enjoy this world tour the first time, *PE&RS* is reprinting prior articles from the column. This month's article on the Republic of Yemen was originally printed in 2003 but contains updates to their coordinate system since then.

**A**round 1000 *BC* the region of present Yemen was ruled by three successive civilizations: the Minaean, the Sabaean, and the Himyarite. These three kingdoms all depended for their wealth on the spice trade, consisting mainly of frankincense and myrrh. By the 11<sup>th</sup> century *BC*, land routes were greatly improved throughout Arabia by using the camel as the beast of burden. Frankincense was carried from its production center at Qana (now Bir'Ali) to Gaza. The chief incense traders were the Minaeans, who established their capital at Karna (now Sadah), before the Sabaean era in 950 *BC*, which lasted for about 14 centuries. The region was invaded by the Romans in the first century *AD*; by the 6<sup>th</sup> century *AD*, it was conquered first by the Ethiopians (*PE&RS*, March 2003), and then by the Persians. The region converted to Islam in the 7<sup>th</sup> century. North Yemen became independent of the Ottoman Empire (Turkey) in 1918. The British had set up a protectorate area around the southern port of Aden during the 19<sup>th</sup> century, but withdrew in 1967 from what had become South Yemen. After two decades of hostilities, the two countries were formally unified as the Republic of Yemen in 1990.

Yemen is bordered by Saudi Arabia on the north (1,458 km), Oman on the east (288 km), the Arabian Sea and Gulf of Aden to the south, and the Red Sea to the west (1906 km). Slightly larger than twice the size of Wyoming, the country is comprised of a narrow coastal plain backed by flat-topped hills and rugged mountains with dissected upland desert plains in the center that slope into the desert interior. The lowest point is the Arabian Sea and the highest point is Jabal an Nabi Shu'ayb at 3,760 m.



Original classical triangulation of Western Aden was done by the British Survey of India in the early 20<sup>th</sup> century. The first large-scale map series, based on ground surveys, was published in 1917. Later updated, the Aden GSGS 3879 series at 1:126,720 scale was published as a polychrome series in 1930. Carl Rathjens and Herman von Wissmann published the *Karte des Reisegebiets in Jemen* (Map of the Region Traversed in Yemen) in 1934. There was no grid on the map at 1:100,000 scale, and it covered a limited area of Yemen between Şan'ā' and the coast. The three-part poly-chrome map was based on road surveys, supplemented with diaries, sketches, and maps from other travelers. Relief was represented by form lines or hachures.

In 1925, the British Survey of India established the Aden Zone Lambert Conical Orthomorphic grid where the latitude of origin  $\varphi_0 = 15^\circ$  N, central meridian  $\lambda_0 = 45^\circ$  E, scale factor at the latitude of origin  $m_0 = 0.999365678$ , False Easting =

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 547–548.

0099-1112/21/547–548

© 2021 American Society for Photogrammetry  
and Remote Sensing

doi: 10.14358/PERS.87.8.547

1,500 km, and False Northing = 1,000 km. This typical “British Grid” had its limits of zone listed as “North: 150,000 meter Northing grid line of the Mecca-Muscat Zone (Lambert Conical Orthomorphic projection), East: Meridian of 60°E, South: From meridian of 60°E, southwest along loxodrome defined by the points 14° 30′ N, 56° E, and 11° 30′ N, 44° E., West: Loxodrome from 44° E, 11° 30′ N toward 19° N, 39° E, to intersection of 150,000 meter Northing grid line of Mecca-Muscat Zone with the loxodrome.” The ellipsoid of reference is the Clarke 1880 where  $a = 6,378,249.145$  m and  $1/f = 293.4663077$ . Curiously, no datum was listed by the British General Staff, Geographical Section (GSGS). However, reading the Survey of India Triangulation Dossiers reveals that all British chains in the region were based on the original Nahrwan Datum of Iraq where  $\Phi_0 = 33^\circ 19' 10.87''$  N,  $\Lambda_0 = 44^\circ 43' 25.54''$  E, and the orientation is based on the azimuth from South End Base of Nahrwan (“1M”) to “2M” as  $\alpha_0 = 169^\circ 04' 08.2''$  from south. In my opinion, the original classical datum of Yemen (“Aden Datum of 1925”) is actually the British-observed Nahrwan Datum. A test point provided by the U.S. Army Corps of Engineers, Lake Survey Unit in 1943 for Aden Zone is  $\varphi = 13^\circ 53' 46.728''$  N,  $\lambda = 37^\circ 37' 19.732''$  E,  $X = 703,075.269$  m,  $Y = 891,245.290$  m,  $\theta = -1^\circ 54' 34.3032$ , and  $m_\varphi = 0.91412282$ .

Numerous map series of Aden and Şan’ā’ were produced by the British and U.S. militaries in the period between 1950 and 1961. Scales varied from 1:10,000 to 1:100,000. Other datums reported to exist include Kamaran (Island) Datum of 1926-1927, Ras Karma (Island) Datum, Socotra (Island) Datum of 1957, and Socotra (Island) Datum of 1964-1965. All of these island “astro” datums are presumably referenced to the Clarke 1880 ellipsoid.

A curiosity in large-scale map projections was developed by the U.S. Geological Survey when a mapping project was completed of Şan’ā’ in the 1960s. The Şan’ā’ Azimuthal Equidistant Grid System was defined by the latitude of origin  $\varphi_0 = 15^\circ 37' 2''$  N, central meridian  $\lambda_0 = 42^\circ 59' 32.25''$  E, scale factor at the latitude of origin  $m_0 = 1.0$ , False Easting = 40 km, and False Northing = 20 km. The International ellipsoid was chosen for the grid, and the datum was presumably ersatz.

Yemen and Oman established their boundary in an 01 October 1992 agreement consisting of straight-line segments connecting eight turning points defined on the WGS84 Datum. The eighth point is a tripoint with Saudi Arabia. On 18 April 2001, Yemen signed a cooperation agreement based on

the 1999 International Tribunal resolution of the Permanent Court of Arbitration, the Hague, the Netherlands. In 2000, Yemen and Saudi Arabia agreed to a delimitation of their common border.

Large-scale local topographic maps of Yemen are controlled by the government. However, complete topographic coverage of the country is available from commercial map sellers worldwide in the form of Russian military mapping at the scales of 1:200,000 and 1:100,000. Considering the availability of these recent up-to-date Russian maps and the phenomenal accuracy achieved nowadays with shirt-pocket-sized consumer-grade GPS receivers, such secrecy of the Yemeni government is merely restricting the economic development of their own nation.

## Yemen Update

In January, 2006, a **TOTAL** (French Oil Company) document was released by the Yemen LNG Co., Ltd. About the “Geodetic Systems on Pipe Line project.” The new “Official Yemen Horizontal Datum is listed as Sana’a – IGN reference marker, calculated on ITRF91 (epoch 92.5) YGN96 (WGS84). The document lists the “Old South Yemen (OSY) Datum referenced to the Krassovsky 1940 ellipsoid where  $a = 6,378.245.00$  m and  $1/f = 298.3$  with a transformation to YGN96 (WGS84) Datum as:  $\Delta X = -76$ m,  $\Delta Y = -138$ m,  $\Delta Z = +67$ m, “Datum shift provided by the French *Institute Géographique National* in October 1996. An example Transformation Test Point is offered for Geodetic marker BH-01 Datum OSY:  $\varphi = 13^\circ 59' 06.145''$  N,  $\lambda = 48^\circ 10' 41.197''$  E,  $h = 23.89$  m., and Datum YNGN96:  $\varphi = 13^\circ 59' 09.5087''$  N,  $\lambda = 48^\circ 10' 40.0179''$  E,  $h = -0.730$  m.,  $H = 18.220$  m. “Geoid model: a local geoid model computed by the French IGN in 1996 is used to convert height above WGS84 spheroid into normal heights above MSL. (at *Al Hudaydah (MSL) – Ed.*). The geoid model software is available at **TOTAL** office.”

The contents of this column reflect the views of the author, who is responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the American Society for Photogrammetry and Remote Sensing and/or the Louisiana State University Center for Geoinformatics (C<sup>4</sup>G).

This column was previously published in *PE&RS*.

## ASPRS WORKSHOP SERIES

### It's not too late to earn Professional Development Hours

Miss one of our Live Online Workshops? Not to worry! You can purchase the workshops now and watch when you are ready! Check out the workshops offered by visiting:

<https://asprs.prolearn.io/catalog>

## JOURNAL STAFF

### Editor-In-Chief

Alper Yilmaz, Ph.D., PERSeditor@asprs.org

### Associate Editors

Rongjun Qin, Ph.D., qin.324@osu.edu

Michael Yang, Ph.D., michael.yang@utwente.nl

Petra Helmholz, Ph.D., Petra.Helmholz@curtin.edu.au

Bo Wu, Ph.D., bo.wu@polyu.edu.hk

Clement Mallet, Ph.D., clemallet@gmail.com

Jose M. Pena, Ph.D., jmpena@ica.csic.es

Prasad Thenkabil, Ph.D., pthenkabil@usgs.gov

Ruisheng Wang, Ph.D., ruishengwang@ucalgary.ca

Desheng Liu, Ph.D., liu.738@osu.edu

Valérie Gouet-Brunet, Ph.D., valerie.gouet@ign.fr

Dorota Iwaszczuk, Ph.D., dorota.iwaszczuk@tum.de

Qunming Wang, Ph.D., wqm11111@126.com

Filiz Sunar, Ph.D., fsunar@itu.edu.tr

Norbert Pfeifer, np@ipf.tuwien.ac.at

Jan Dirk Wegner, jan.wegner@geod.baug.ethz.ch

Hongyan Zhang, zhanghongyan@whu.edu.cn

Dongdong Wang, Ph.D., ddwang@umd.edu

Zhenfeng Shao, Ph.D., shaozhenfeng@whu.edu.cn

Ribana Roscher, Ph.D., ribana.roscher@uni-bonn.de

Sidike Paheding, Ph.D., spahedin@mtu.edu

### Contributing Editors

#### Highlight Editor

Jie Shan, Ph.D., jshan@ecn.purdue.edu

#### Feature Articles

Michael Joos, CP, GISP, featureeditor@asprs.org

#### Grids & Datums Column

Clifford J. Mugnier, C.P., C.M.S., cjmce@lsu.edu

#### Book Reviews

Sagar Deshpande, Ph.D., bookreview@asprs.org

#### Mapping Matters Column

Qassim Abdullah, Ph.D., Mapping\_Matters@asprs.org

#### Sector Insight

Lucia Lovison-Golob, Ph.D., lucia.lovison@sat-drones.com

Bob Ryerson, Ph.D., FASPRS, bryerson@kimgeomatics.com

#### GIS Tips & Tricks

Alvan Karlin, Ph.D., CMS-L, GISP akarlin@Dewberry.com

### ASPRS Staff

#### Assistant Director — Publications

Rae Kelley, rkelly@asprs.org

#### Electronic Publications Manager/Graphic Artist

Matthew Austin, maustin@asprs.org

#### Advertising Sales Representative

Bill Spilman, bill@innovativemediasolutions.com

## JAMES “JESSE” WINCH RETIRES FROM ASPRS

Jesse Winch joined ASPRS in October 1998, not long after Jim Plasker was hired as Executive Director. Jesse was hired for the position of Program Manager focusing on Certification and the Awards and Scholarship Programs. He also became the liaison with the Board of Directors, the Division, Region, and Committee officers and, eventually, Office Manager. Jesse worked closely with Jim and Kim on a great many projects including the retiring of the mortgage on the headquarters site and the restoration of the ASPRS Foundation in 2005. Jim subsequently appointed him to be Assistant Executive Director of the Foundation which allowed him to work with and develop lasting friendships with members on the Board of Trustees.

ASPRS went through a couple of rough years after Jim Plasker retired. Jesse was asked to serve as Acting Executive Director by then ASPRS President Becky Morton. Jesse, with guidance from Karen Schuckman and the Board of Directors, was able to steer ASPRS in the direction of becoming a stable and financially solvent organization.

“I’ll never forget Jesse’s love for good coffee and his concerts with the world-renowned Photogrammetric Ramblers,” says Rae Kelley, Director of Publications.

Jesse says, “I’m looking forward to retirement as I’m still very active in the Irish music world.” A couple of years ago he was inducted into the Hall of Fame for the Mid-Atlantic Region of Comhaltas Ceoltoiri Éireann, the international Irish musicians’ organization based in Dublin, Ireland.

## ASPRS PARTNERS WITH PROMETRIC TO PROVIDE ALL FUTURE ASPRS CERTIFICATION EXAMS

### This new partnership:

- Removes the need for candidates to coordinate proctors
- Allows for examinations to be offered in a COVID-19-safe environment
- Provides immediate test results to the candidate
- Allows for “remote proctoring” for candidates who do not have a Prometric testing center near their locations, to include international candidates
- For more information regarding the ASPRS Certification Program, visit <https://www.asprs.org/certification>.

A growing number of scientific and technical disciplines depend on photogrammetry and the mapping sciences for reliable measurements and information. It is in the interest of those who provide photogrammetric and mapping sciences services, as well as the user of these services, that such information and data be accurate and dependable. The ASPRS Certification Program has as its purpose the establishment and maintenance of high standards of ethical conduct and professional practice among photogrammetrists, mapping scientists, technologists, and interns.

### Who at ASPRS Do I Contact for help with...

#### Membership/PE&RS Subscription

Shalini Gogawale — office@asprs.org

#### Advertising/Exhibit Sales

Bill Spilman —  
bill@innovativemediasolutions.com

#### PE&RS Editor-in-Chief

Alper Yilmaz — PERSeditor@asprs.org

#### Calendar — calendar@asprs.org

ASPRS Bookstore — office@asprs.org

ASPRS Foundation — foundation@asprs.org

425 Barlow Place, Suite 210, Bethesda, MD 20814  
301-493-0290, 225-408-4422 (fax), [www.asprs.org](http://www.asprs.org)

# WHO'S WHO IN ASPRS

## BOARD OF DIRECTORS

### BOARD OFFICERS

#### President

Jason M. Stoker, Ph.D,  
U.S. Geological Survey

#### President-Elect

Christopher Parrish, Ph.D  
Oregon State University

#### Vice President

Lorraine B. Amenda, PLS, CP  
Towill, Inc.

#### Past President

Jeff Lovin  
Woolpert

#### Treasurer

Stewart Walker, Ph.D.  
photogrammetry4u

#### Secretary

Harold Rempel  
ESP Associates, Inc.

---

### BOARD MEMBERS

#### Sustaining Members Council

Chair: Joe Cantz  
<https://www.asprs.org/sustaining-members-council.html>

#### Technical Division Directors Council

Chair: Bandana Kar, Ph.D  
Vice Chair: Denise G. Theunissen

#### Committee Chairs Council

Chair: Mike Zoltek

#### Early-Career Professionals Council

Chair: Melissa Martin

#### Region Officers Council

Chair: Demetrio Zourarakis  
Vice Chair: Jason Krueger

#### Student Advisory Council

Chair: Evan Vega  
Vice Chair: Lauren McKinney  
<http://www.asprs.org/Students/Student-Advisory-Council.html>

#### Geographic Information Systems Division

Director: Denise Theunissen  
Assistant Director: Briana Williams  
[www.asprs.org/Divisions/GIS-Division.html](http://www.asprs.org/Divisions/GIS-Division.html)

#### Lidar Division

Director: Josh Nimetz, CMS  
Assistant Director: Ajit Sampath  
[www.asprs.org/Divisions/Lidar-Division.html](http://www.asprs.org/Divisions/Lidar-Division.html)

#### Photogrammetric Applications Division

Director: Kurt Rogers  
Assistant Director: Benjamin Wilkinson  
[www.asprs.org/Divisions/Photogrammetric-Applications-Division.html](http://www.asprs.org/Divisions/Photogrammetric-Applications-Division.html)

#### Primary Data Acquisition Division

Director: Greg Stensaas  
Assistant Director: Srini Dharmapuri  
[www.asprs.org/Divisions/Primary-Data-Aquisition-Division.html](http://www.asprs.org/Divisions/Primary-Data-Aquisition-Division.html)

#### Professional Practice Division

Director: Bill Swope  
Assistant Director: Hope Morgan  
[www.asprs.org/Divisions/Professional-Practice-Division.html](http://www.asprs.org/Divisions/Professional-Practice-Division.html)

#### Remote Sensing Applications Division

Director: Amr Abd-Ehrahman  
Assistant Director: Tao Liu  
[www.asprs.org/Divisions/Remote-Sensing-Applications-Division.html](http://www.asprs.org/Divisions/Remote-Sensing-Applications-Division.html)

#### Unmanned Autonomous Systems (UAS)

Director: Dan Hubert  
Assistant Director: Jacob Lopez  
<https://www.asprs.org/divisions-committees/uas-division>

---

### REGION PRESIDENTS

#### Alaska Region

#### Cascadia Region

Robert Hariston-Porter

#### Eastern Great Lakes Region

Michael Joos, CP, GISP  
<http://egl.asprs.org>

#### Florida Region

Xan Fredericks  
<http://florida.asprs.org>

#### Heartland Region

Whit Lynn  
<http://heartland.asprs.org/>

#### Intermountain Region

Robert T. Pack  
<https://asprsintermountain.weebly.com>

#### Mid-South Region

David Hughes  
<https://www.asprs.org/all-regions/mid-south.html>

#### Northeast Region

#### North Atlantic Region

<http://natlantic.asprs.org>

#### Pacific Southwest Region

John Erickson, PLS, CP  
<https://pswasprs.org>

#### Potomac Region

Dave Lasko  
<http://www.asprspotomac.org>

#### Rocky Mountain Region

<http://www.asprs-rmr.org/>

#### Western Great Lakes Region

Adam Smith  
<http://wgl.asprs.org/>

# Spinor-Based Attitude Determination with Star Sensor Considering Depth

Qinghong Sheng, Rui Ren, Weilan Xu, Hui Xiao, Bo Wang, and Ran Hong

## Abstract

A star sensor is a high-precision satellite attitude measurement device. Since its observation information has only two-dimensional direction vectors, when a star sensor is used for attitude determination the dimension of the observation information is less than the number of attitude angles determined, so mainstream algorithms usually only guarantee the accuracy of the pitch angle and the roll angle. In view of the lack of depth information in the observation's imaging geometric condition, this article proposes a spinor-based attitude determination model, which describes a straight line passing through two stars with the spinor and maps the depth information of the straight line with the pitch, to establish an imaging geometry model of the spinor coplanar condition. Experiments show that the yaw-angle attitude accuracy of the method is an order of magnitude better than that of mainstream algorithms, and the accuracy of the three attitude angles reaches the arc-second level.

## Introduction

Because of their advantages of strong anti-interference ability, high measurement accuracy, and autonomous attitude-sensor navigation, star sensors are widely used in satellite attitude measurement. The principle of the star-sensor-attitude determination algorithm is to solve the Wahba problem directly or indirectly (Habib *et al.* 2007; Xie *et al.* 2017; Guan *et al.* 2019; Meygret *et al.* 2019). Determination algorithms mainly include minimum least-squares (MLS; Xie 2009), Tri-Axial Attitude Determination (TRIAD), the recursive Euler-q algorithm, QUEST, Fast Optimal Attitude Matrix, and their improved versions (Gai *et al.* 1985; Shuster *et al.* 1991; Guan *et al.* 2019). The TRIAD algorithm constructs a new orthogonal coordinate system from two nonparallel star-direction vectors and calculates the rotation matrix (Liebe 2002). It uses quaternions to describe the attitude and constructs a K matrix (Shuster 2006). The QUEST algorithm performs batch processing in a single-point manner (Keat 1977), using Rodrigues parameters to represent quaternions; compared with the q-method algorithm, the QUEST algorithm simplifies the calculation of solving eigen values. Many scholars have expanded the QUEST algorithm into a Kalman-filter algorithm (Shuster and Oh 1981; Zhang *et al.* 2017; van Wyk *et al.* 2020) based on actual needs and combined with prior information, and made predictions based on the law of attitude change along with time to reduce systematic errors. The Fast Optimal Attitude Matrix method indirectly uses singular value decomposition to solve the attitude distribution matrix, yielding good robustness (Markley 1988, 1993).

These methods for determining the attitude of a star sensor are characterized by solving the optimal direction cosine

matrix with a determinant of +1 according to the collinear condition in the imaging geometry to minimize the cost function. However, since the closest star to the earth is about 4.2 light-years away, all stars observed by a star sensor can be considered to be approximately in the same two-dimensional “plane” space at infinity. The observation data acquired lack depth-control elements, and even if the number of star-direction vectors obtained is greater than the number of unknowns, when only the “plane” control information is included, the amount of information necessary to determine the direction of the optical axis is insufficient. So in solving attitude, there is a serious multicollinearity of normal equations, and a small change in the observed value of a star's image point coordinates causes instability in the solution.

In recent years, the field of robot precise kinematics modeling has used spinor theory to describe the attitude and position of a rigid body in space; the practice of calibrating robot systems' end geometric parameter error and tracking the three-dimensional trajectory has achieved high precision (Davidson and Hunt 2004). A spinor is a geometric quantity with five independent parameters, and the physical meaning of pitch is the linear distance that the rigid body moving in three-dimensional space translates along the axis; the movement of the rigid body makes the end rod become the axis, thereby drawing a three-dimensional ruled surface. In the observation of star geometric imaging, the rotation of a straight line passing through a binary star in space is expressed by movement around the imaging visual axis; its translation will then inevitably occur along the visual axis, and the translation distance is the pitch. By means of the pitch, each image spinor can map the depth information (along the visual axis) of the star in the two-dimensional “plane” at infinity, and the star control conditions are upgraded from the two-dimensional plane to the three-dimensional solid, thereby forming a strongly constrained star geometric imaging condition.

In order to improve the reliability of attitude sensors, similar to the method of spinors in the state-characteristics analysis of rigid-body motion, this article uses spinors to increase the control information dimension. In the second part of the article, an imaging geometric model based on spinors is established to analyze the complex collinearity of the normal equation matrix and the morbidity of the observation data system in the attitude determination process. The third part uses simulation data to verify the effectiveness of this method at overcoming the morbidity of observation data.

## Materials and Methods

The spinor-based attitude determination (SPIAD) method proposed in this article is mainly divided into two parts: using a spinor to represent the connection of two stars, and the

Qinghong Sheng, Rui Ren, Weilan Xu, Bo Wang, and Ran Hong are with the Nanjing University of Aeronautics and Astronautics, Nanjing, China.

Hui Xiao is with Nanjing Xiaozhuang University, Nanjing, China.

Contributed by Ruisheng Wang, August 25, 2020 (sent for review November 13, 2021; reviewed by Yongjun Zhang, Perpetual Akwensi)

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 551–556.  
0099-1112/21/551–556

© 2021 American Society for Photogrammetry  
and Remote Sensing  
doi: 10.14358/PERS.87.8.551

pitch to represent the depth information of this connection, to establish the coplanar spinor imaging geometric model; and deriving the error equation and the normal equation of the method to solve the satellite attitude.

### General Concept of a Spinor Based on the Imaging Geometric Model

#### Spinor Representation

As shown in Figure 1, the spinor  $S$  is a geometry composed of two vectors:

$$S = \begin{pmatrix} \mathbf{o} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{o} \\ \mathbf{r} \times \mathbf{o} + h\mathbf{o} \end{pmatrix} \quad (1)$$

The unit vector  $\mathbf{o}$  represents the direction vector of the spinor axis,  $\mathbf{r}$  is the position vector of the axis,  $h$  is the pitch, and  $h = \mathbf{o} \cdot \mathbf{p}$ —that is, the pitch is the projection of the subpart  $\mathbf{p}$  on the main part  $\mathbf{o}$ . If the pitch is zero, the spinor is a line vector and  $\mathbf{r}$  is perpendicular to  $\mathbf{o}$ . There are five independent parameters of a spinor: four describing position and attitude, and the pitch. The physical meaning of a spinor is that when a rigid body moves, the linear distance that the rigid body translates along the axis is independent of the origin and invariant. This means that the size of the pitch is only related to the amount of the spinor itself. When the coordinate system moves, the pitch does not change; when a straight line makes a helical motion, if it simultaneously translates along the axis it can draw a spiral line. According to the physical meaning of pitch, the height of the cylinder where the spiral line is located is the pitch. In other words, the amount of a spinor with pitch can increase the dimension of the information, and the added dimension parameter is a distance on the direction vector where the axis is located.

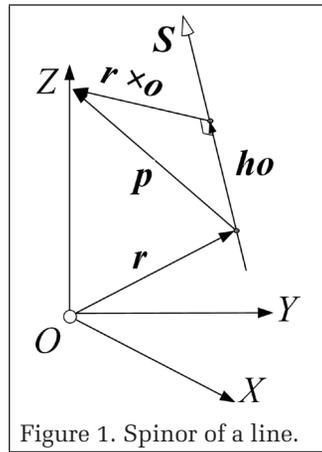


Figure 1. Spinor of a line.

Suppose that the direction vectors of the axes of the two spinors  $S_1$  and  $S_2$  are respectively  $\mathbf{o}_1$  and  $\mathbf{o}_2$ , the position vectors are  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , and the pitches are  $h_1$  and  $h_2$ ; then the spatial position relationship between  $S_1$  and  $S_2$  is the reciprocal product:

$$S_1 \circ S_2 = (\mathbf{r}_2 \times \mathbf{o}_2 + h_2\mathbf{o}_2) \cdot (\mathbf{r}_1 \times \mathbf{o}_1 + h_1\mathbf{o}_1) \quad (2)$$

The reciprocal product is related only to the distance between axes and the angle between the projection and the pitch of the two spinors. It is independent of the origin of the coordinate system and invariant. When the two spinors satisfy the condition of coplanarity, the reciprocal product is zero.

#### Coplanar Equation of Spinors with Pitch

The celestial coordinate system uses the equatorial plane as the basic plane, and the image space coordinate system  $O$ - $XYZ$  takes the photography center  $O$  as the original point. Stars  $A$  and  $B$  in the celestial sphere correspond to image points  $a$  and  $b$  in the image space coordinate system.

In the celestial coordinate system, suppose the right ascension and declination of stars  $A$  and  $B$  are  $(\alpha_1, \beta_1)$  and  $(\alpha_2, \beta_2)$ , and the spinor formed by the straight lines passing through stars  $A$  and  $B$  is  $S_0$ ; then the attitude vector  $\mathbf{o}_0$  and position vector  $\mathbf{r}_0$  of the spinor  $S_0$  are

$$\begin{aligned} \mathbf{o}_0 &= (\lambda_2 - \lambda_1 \quad \mu_2 - \mu_1 \quad \nu_2 - \nu_1) \\ \mathbf{r}_0 &= (\lambda_1 \quad \mu_1 \quad \nu_1) \end{aligned} \quad (3)$$

where  $\lambda$ ,  $\mu$ , and  $\nu$  are the projections of the right ascension.

Suppose the focal length of the sensor is  $f$ , the image space coordinates of image points  $a$  and  $b$  are  $(x_1, y_1, f)$  and  $(x_2, y_2, f)$ , and the connection with the photography center  $O$  is the imaging light, represented by spinors  $S_1$  and  $S_2$ . Since the imaging rays  $Oa$  and  $Ob$  pass through the origin  $O$  of the coordinate system, the position vectors of  $S_1$  and  $S_2$  are zero and the direction vectors of the axes of  $S_1$  and  $S_2$  are expressed by  $\mathbf{o}_1$  and  $\mathbf{o}_2$ :

$$\begin{aligned} \mathbf{o}_1 &= \frac{1}{\sqrt{x_1^2 + y_1^2 + f^2}} (x_1 \quad y_1 \quad f) \\ \mathbf{o}_2 &= \frac{1}{\sqrt{x_2^2 + y_2^2 + f^2}} (x_2 \quad y_2 \quad f) \end{aligned} \quad (4)$$

When the straight lines  $AB$ ,  $Oa$ , and  $Ob$  lie on the same plane, the reciprocal product  $F_1$  of  $S_0$  and  $S_1$  and the reciprocal product  $F_2$  of  $S_0$  and  $S_2$  are both zero. Therefore, the coordinate transformation relationship between the celestial coordinate system and the image space coordinate system can be established.

However, the deviation of the initial direction of the visual axis result in nonzero reciprocal products  $F_1$  and  $F_2$ , so the visual axis needs to be rotated to restore the correct direction of the imaging rays  $Oa$  and  $Ob$ . These imaging rays are constantly shifting  $h$  in the direction of the optical axis and continuously rotating around the optical axis by a quaternion  $q$  to achieve corrected image point coordinate observations. In other words, the rotation matrix  $\mathbf{R}$  composed of quaternion  $q$  is used to perform coordinate transformation on the spinors  $S_1$  and  $S_2$  of the imaging light, enabling the rotated imaging rays  $Oa$  and  $Ob$  to be coplanar with the straight line  $AB$  at the same time, as shown in Figure 2. By multiplying the rotation matrix  $\mathbf{R}$  with the direction vectors of the  $S_1$  and  $S_2$  axes, the spinor coplanar conditional equation for determining the attitude can be established:

$$\begin{cases} F_1 = \mathbf{R}\mathbf{o}_1 \cdot (\mathbf{r}_0 \times \mathbf{o}_0 + h_0\mathbf{o}_0) = 0 \\ F_2 = \mathbf{R}\mathbf{o}_2 \cdot (\mathbf{r}_0 \times \mathbf{o}_0 + h_0\mathbf{o}_0) = 0 \end{cases} \quad (5)$$

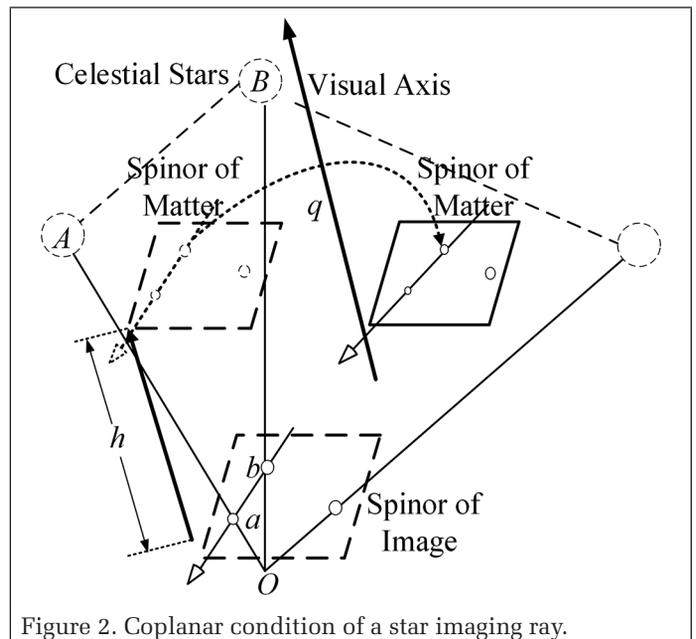


Figure 2. Coplanar condition of a star imaging ray.

## Mathematical Model of Spinor-Based Attitude Determination

Suppose that the three attitude angles are  $\varphi$ ,  $\omega$ , and  $\kappa$ , then the rotation matrix  $\mathbf{R}$  is

$$\mathbf{R} = \begin{bmatrix} \cos\varphi \cos\kappa + \sin\varphi \sin\omega \sin\kappa & -\cos\varphi \sin\kappa + \sin\varphi \cos\kappa & \sin\varphi \sin\omega \\ \cos\omega \sin\kappa & \cos\omega \cos\kappa & \cos\omega \cos\kappa \\ \sin\varphi \cos\kappa - \cos\varphi \sin\omega \sin\kappa & -\sin\varphi \sin\kappa - \cos\varphi \sin\omega \cos\kappa & -\cos\varphi \cos\omega \end{bmatrix} \quad (6)$$

Substituting equation 6 into equation 5 gives

$$\text{see below} \quad (7)$$

Equation 7 is linearized to establish the error equation:

$$\begin{aligned} F_1 &= F_{10} + a_{11}d\varphi + a_{12}d\omega + a_{13}d\kappa + a_{14}dh \\ F_2 &= F_{20} + a_{21}d\varphi + a_{22}d\omega + a_{23}d\kappa + a_{24}dh \end{aligned} \quad (8)$$

where  $F_{10}$  and  $F_{20}$  are, respectively, the residual errors of the corresponding observations, and  $d\varphi$ ,  $d\omega$ ,  $d\kappa$ , and  $dh$  are the corrections of the unknowns.

Assuming that  $n$  straight lines passing through two stars are observed,  $2n$  error equations can be listed:

$$(\mathbf{A}^T \mathbf{A})\mathbf{X} = \mathbf{A}^T \mathbf{F} \quad (9)$$

where  $\mathbf{A}$  is the coefficient matrix of unknowns and  $\mathbf{X}$  is the unknowns matrix composed of attitude parameters and pitch. If the normal equation is  $\mathbf{N}\mathbf{X} = \mathbf{W}$ , where  $\mathbf{N} = \mathbf{A}^T \mathbf{A}$  and  $\mathbf{W} = \mathbf{A}^T \mathbf{F}$ , then assuming that there is a slight perturbation  $\Delta\mathbf{N}$  and  $\Delta\mathbf{W}$ , which causes the estimator  $\mathbf{X}$  to deviate by  $\Delta\mathbf{X}$ , the normal equation after the disturbance can be expressed as

$$(\mathbf{N} + \Delta\mathbf{N})(\mathbf{X} + \Delta\mathbf{X}) = (\mathbf{W} + \Delta\mathbf{W}) \quad (10)$$

which means

$$\frac{\|\Delta\mathbf{X}\|}{\|\mathbf{X} + \Delta\mathbf{X}\|} \leq \|\mathbf{N}\| \|\mathbf{N}^{-1}\| \frac{\|\Delta\mathbf{N}\|}{\|\mathbf{N}\|} \quad (11)$$

If the solution process is accurately carried out, then whether the relative error of the solution of the equation system caused by the error of the observation data is controllable depends on  $\|\mathbf{N}\| \|\mathbf{N}^{-1}\|$ . If its value is large, a small disturbance in the observation data  $\mathbf{N}$  and  $\mathbf{W}$  can cause large changes in the solution and have a greater negative impact. The traditional attitude determination method for a star sensor projects the two-dimensional star direction-vector information on the two-dimensional "plane" space at infinity onto three-dimensional coordinate axes. On the one hand, the elements of the three-dimensional direction information are not independent of each other; and on the other hand, due to the lack of depth-control elements in the detected observation data, these star observation data cannot fully reflect the role of the three attitude parameters in the attitude determination system, especially the yaw angle rotating around the Z

direction, and even observing multiple data points does not bring enough control information. Therefore, the observation matrix cannot obtain reliable results from the observation data, resulting in an approximate linear correlation between the column vectors of the observation matrix. And because the approximate complex collinearity of the observation matrix is directly transmitted to the normal equation matrix through the power, it leads to severe complex collinearity of the normal equation matrix. Small changes in the star's image point coordinate observations will eventually cause ill-posed attitude systems.

In the SPIAD method proposed in this article, the star information in the observation data is the spinor S0 passing through two stars, and the pitch in S0 will be translated along the direction of the visual axis; in other words, the pitch is added to the control information. At this moment, the pitch maps the information of the star's depth direction. On the one hand, the process of solving the unknown is controlled by higher-dimensional information, and the correlation coefficient between the three attitude angles is reduced. On the other hand, the star control conditions are improved from two dimensions to three, which reduces the degree of the complex collinearity of the data, weakens the effect of data morbidity, and forms strongly constrained star geometric imaging conditions, thereby reducing the impact of small disturbances on attitude accuracy.

## Results

### Data Used and Parameter Settings

To evaluate the performance of SPIAD, we selected two different data sets: simulated and active pixel sensor images. Table 1 shows the parameters of the two data sets. The stars located on the surrounding edges of the image are selected, and the celestial sphere area captured by the attitude sensor is determined according to the right ascension and declination of these stars in the celestial coordinate system. In the star catalogue, all the stars in this area are selected, and the collinearity equation is used to project the celestial coordinates of the extracted stars onto the star map to obtain the image space coordinates of the stars.

### Experimental Results and Analysis

#### The Best Use Condition of the SPIAD Algorithm

This section discusses the influence of the number and distribution of stars on the accuracy of attitude determination using the simulated data and the SPIAD method.

The attitude determination results for different numbers of stars are shown in Figure 3. With increasing numbers of stars, the attitude determination accuracy of the SPIAD method generally shows an upward trend. The accuracy of the yaw angle is always lower than that of the other two angles; when the number of stars reaches 4, the attitude accuracy on the three direction axes tends to be stable: for the roll and pitch angles it reaches 1 arcsec and for the yaw angle <2 arcsec. At this moment, the accuracy of the three attitude angles is maintained at an order of magnitude.

Table 1. Technical parameters of the images in the experiments.

Data Set	Star Catalogue	Focus (mm)	Height (pixels)	Width (pixels)	Pixel Size (mm)
Simulated images	HIP (J2000)	43.3	1024	1024	0.015
Active pixel sensor images	HIP (J2000)	43.3	1024	1024	0.015

### Equation 7

$$\begin{aligned} F_1 &= l_1(l_0 + q_0 h)(\cos\varphi \cos\kappa + \sin\varphi \sin\omega \sin\kappa - \cos\varphi \sin\kappa + \sin\varphi \cos\kappa + \sin\varphi \sin\omega) + m_1(m_0 + s_0 h)(\cos\omega \sin\kappa + \cos\omega \cos\kappa + \sin\omega) + n_1(n_0 + t_0 h)(\sin\varphi \cos\kappa - \cos\varphi \sin\omega \sin\kappa - \sin\varphi \sin\kappa - \cos\varphi \sin\omega \cos\kappa - \cos\varphi \cos\omega) \\ F_2 &= l_2(l_0 + q_0 h)(\cos\varphi \cos\kappa + \sin\varphi \sin\omega \sin\kappa - \cos\varphi \sin\kappa + \sin\varphi \cos\kappa + \sin\varphi \sin\omega) + m_2(m_0 + s_0 h)(\cos\omega \sin\kappa + \cos\omega \cos\kappa + \sin\omega) + n_2(n_0 + t_0 h)(\sin\varphi \cos\kappa - \cos\varphi \sin\omega \sin\kappa - \sin\varphi \sin\kappa - \cos\varphi \sin\omega \cos\kappa - \cos\varphi \cos\omega) \end{aligned}$$

When the number of stars is 4, five distribution scenarios are set, as shown in Figure 4; the accuracy results are shown in Figure 5. In Figure 4a–d, along with the distribution of star points becoming wider and wider, the attitude accuracy becomes higher and higher. In the star point distributions of Figure 4b and 4c, the error of attitude accuracy is large, and there is at least one attitude angle whose accuracy is  $>10$  arcsec; in the star point distribution of Figure 4c, the attitude accuracy is slightly higher, but the accuracy of the three attitude angles is still within 10 arc; and in the star point distribution of

Figure 4d, the attitude accuracy is the highest, and the error of each of the three attitude angles reaches 1–2 arcsec.

#### Comparison Experiments

This section discusses the experimental results of the SPIAD algorithm, the MLS algorithm, and other mainstream algorithms when the number of stars is 4 and the stars are far from the visual axis and evenly distributed. The star pixel coordinates are randomly increased by an error of 0.01–0.8 pixel. The accuracy results of the SPIAD and MLS algorithms are compared in Figure 6.

It can be seen from Figure 6 that the two algorithms exhibit less difference in the accuracy of the pitch and roll angles,

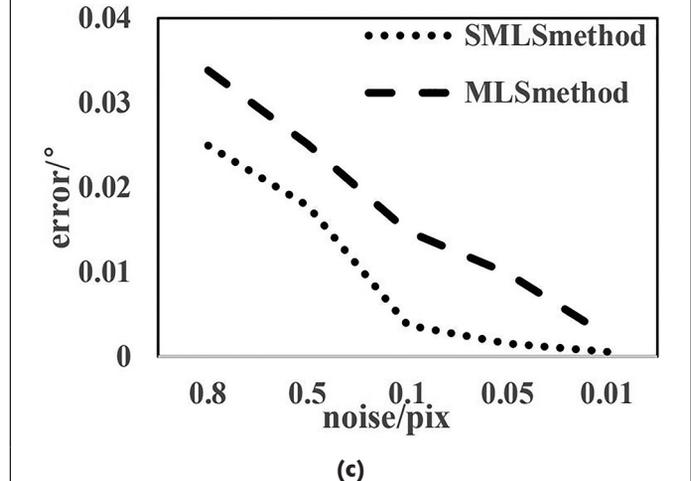
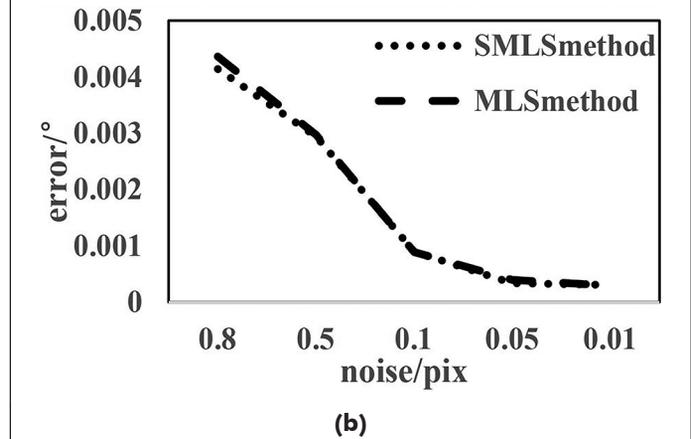
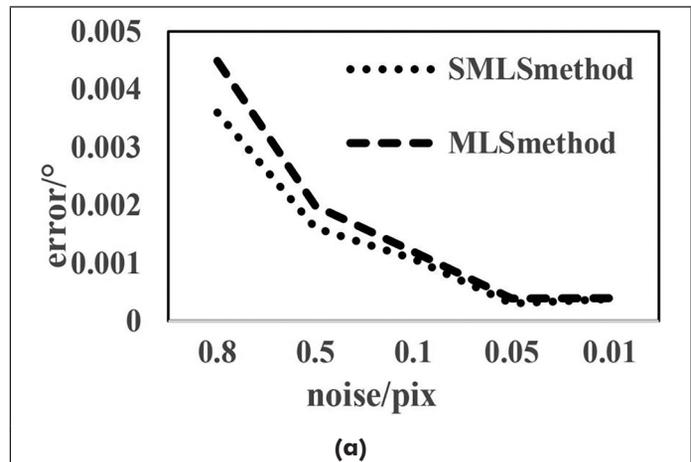
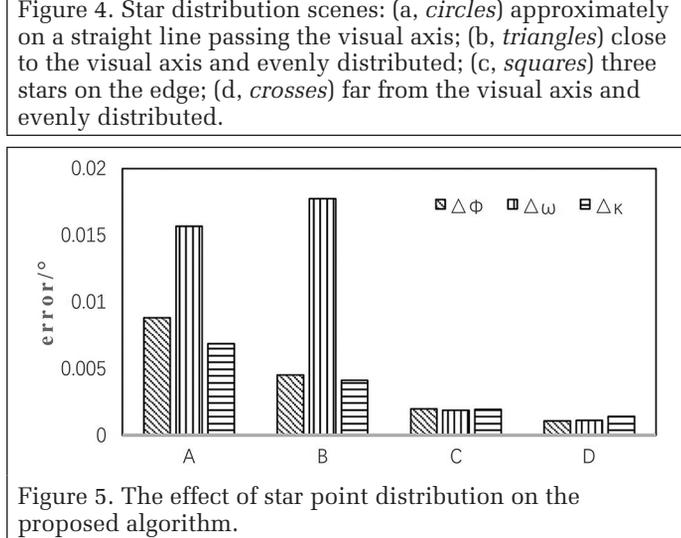
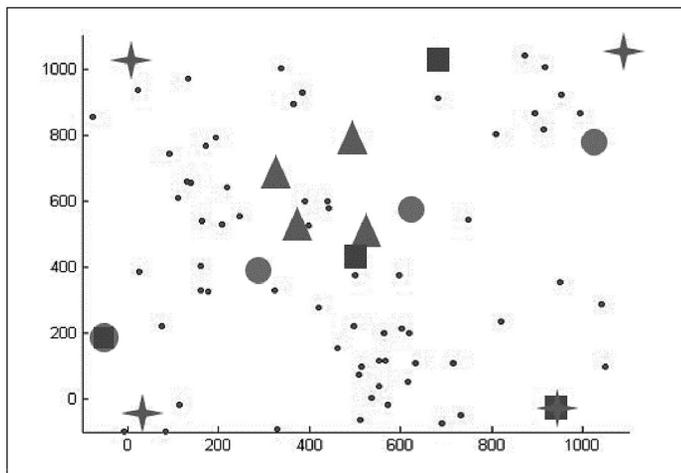
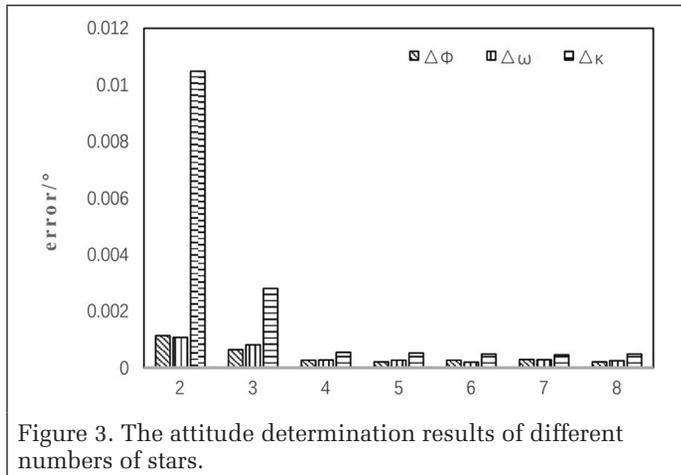


Figure 6. Error statistics for (a) pitch angle  $\phi$ , (b) roll angle  $\omega$ , and (c) yaw angle  $\kappa$ .

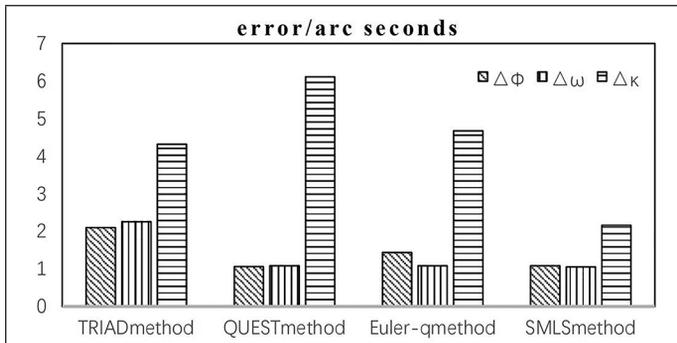


Figure 7. Error statistics.

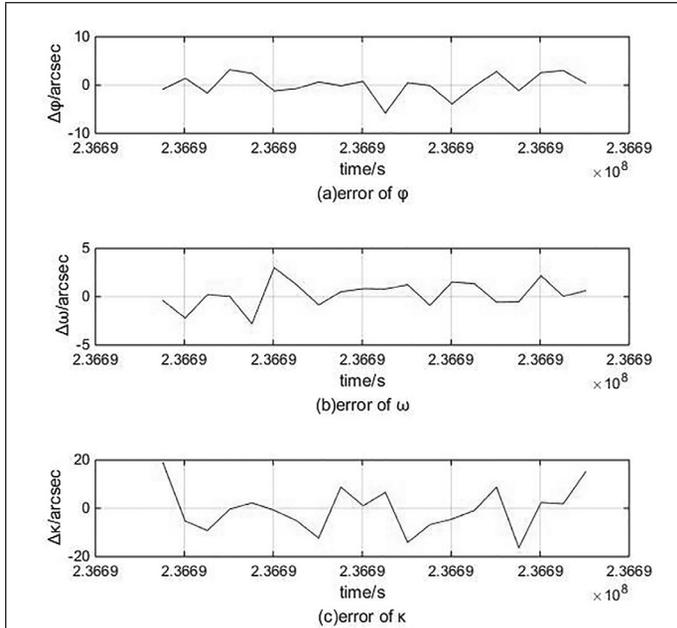


Figure 8. Error of attitude determination with the minimum least-squares algorithm.

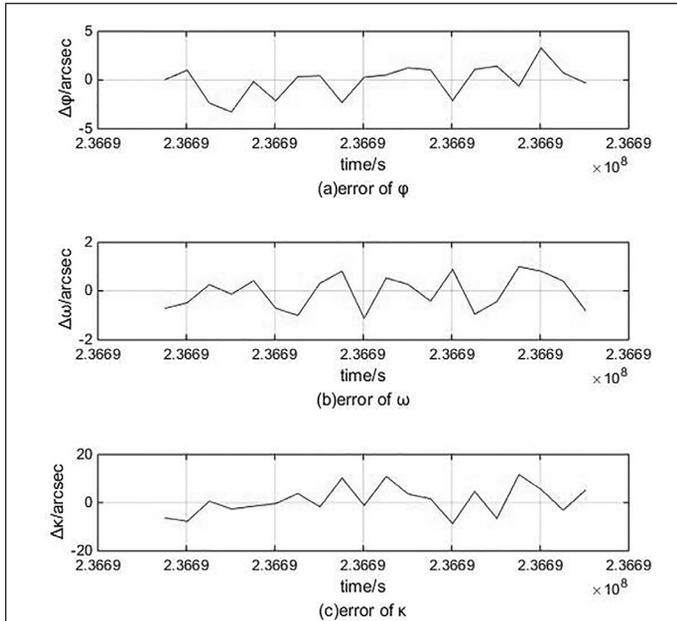


Figure 9. Error of attitude determination with the proposed algorithm.

and greater difference in the accuracy of the yaw angle; when the noise is within 0.1 pixel, the accuracy is stable—5 arcsec for the MLS algorithm and 2 arcsec for the SPIAD algorithm. It is obvious that the accuracy of the SPIAD algorithm is better than that of the MLS algorithm. Table 2 further shows the correlation coefficients and condition numbers of each attitude angle of the MLS and SPIAD algorithms, with lower results for the SPIAD algorithm.

Figure 7 shows the accuracy results of the SPIAD algorithm and mainstream algorithms under the same conditions. As can be seen, the accuracy of the roll and pitch angles with the TRIAD algorithm is 2–3 arcsec, and that of the yaw angle is 5 arcsec. The accuracy of the roll and pitch angles with the Euler-q algorithm is 1–2 arcsec, and that of the yaw angle is basically consistent with the results of the TRIAD algorithm. The accuracy of the roll and pitch angles with the QUEST algorithm is high, reaching 1 arcsec, but that of the yaw angle exceeds 6 arcsec. The pitch and roll angles with the SPIAD algorithm are similar to those with the QUEST algorithm, but the yaw-angle accuracy is much higher than with the QUEST algorithm, close to 2 arcsec. The SPIAD algorithm is more balanced in terms of the three attitudes, with all three attitude angles within 2.1 arcsec, giving an average two- to threefold improvement over these mainstream algorithms.

#### Active Pixel Sensor Data Experiments

The real experiments are discussed in this section. Table 3 shows the comparison experiment results for the attitude angle calculated in the first frame of the star map and real-time stellar attitude. For the 20 star images from star time 236689009.5265, MLS and SPIAD were used to determine the attitude, which was compared with the Euler angle transmitted from the star. The experimental results are shown in Figures 8 and 9. Comparisons of the maximum error and the root-mean-square error of each attitude parameter are shown in Tables 4 and 5.

Table 2. Correlations between orientation parameters.

Adjustment Method	Item	$\phi$	$\omega$	$\kappa$	Condition Number
Minimum least squares	$\phi$	1			12.49
	$\omega$	-0.76	1		
	$\kappa$	0.74	-0.89	1	
Spinor-based attitude determination	$\phi$	1			0.1678
	$\omega$	0.44	1		
	$\kappa$	-0.34	0.26	1	

Table 3. Attitude determination results for the first star map.

Algorithm	$\phi$ (°)	$\omega$ (°)	$\kappa$ (°)
Real attitude	65.457 060 55	57.715 672 30	89.265 835 98
Minimum least squares	65.456 124 27	57.716 684 67	89.270 932 50
Spinor-based attitude determination	65.457 491 28	57.715 534 86	89.264 811 41

Table 4. Comparison of maximum error.

Algorithm	$\Delta\phi$ (")	$\Delta\omega$ (")	$\Delta\kappa$ (")
Minimum least squares	3.1694	3.0019	18.8125
Spinor-based attitude determination	3.2899	1.0011	11.6337

Table 5. Comparison of root-mean-square error.

Algorithm	$\Delta\phi$ (")	$\Delta\omega$ (")	$\Delta\kappa$ (")
Minimum least squares	2.2186	1.3802	8.9944
Spinor-based attitude determination	1.5822	0.6861	5.9523

It can be seen that both the MLS and SPIAD algorithms can successfully achieve attitude determination, and both achieve high accuracy. The errors of the pitch and roll angles are smaller than that of the yaw angle. In the MLS algorithm, the pitch-angle accuracy reaches 2–3 arcsec, and the roll-angle accuracy reaches 1–3 arcsec; in the SPIAD algorithm, the pitch-angle accuracy reaches 1–3 arcsec, and the roll-angle accuracy can sometimes reach the sub-arc-second level. However, the attitude determination error of the yaw angle obtained by the two algorithms is relatively large, with a maximum above 10 arcsec, and the root-mean-square error is above 5 arcsec. The error of the SPIAD algorithm is generally smaller than that of the MLS algorithm.

## Discussion

When the star distribution field is four points located on the same line, the accuracy of the attitude angle is the worst. This is due to the strong correlation of the distribution of control points, leading to a large condition number for the coefficient matrix after matrix inversion, with the equation having serious ill-posed problems. As shown in Table 2, the correlation coefficients of  $\omega$  and  $\kappa$  and of  $\omega$  and  $\kappa$  are close to 1, indicating that a high correlation exists between the orientation elements in the MLS model. In the SPIAD algorithm, the pitch is the linear distance that the rigid body moves horizontally along the axis when it moves in a spiral. Therefore, the three attitude angles are controlled by higher-dimensional information, reducing the correlation coefficients between the attitude angles and the degree of complex collinearity of the data in the least-squares solution, making the attitude determination process less affected by disturbance of the control information. In particular, since the pitch is located in the pointing direction of the optical axis of the star camera, it is equivalent to generating new control information in the pointing direction of the optical axis, thus improving the accuracy of the yaw angle.

## Conclusion

This article presents a method for determining the attitude of a star sensor based on spinors. Based on the coplanar condition of spinors, a new imaging geometric model based on the attitude sensor is established. The results show that the algorithm in this article, with four stars far from the visual axis uniformly distributed, has the highest degree of precision. The pitch increases the depth information of the observation data and reduces the correlation between attitude parameters, making the ill-posed problem of the attitude determination system weaker. The accuracy of the yaw angle shows a significant improvement effect, and those of the pitch and roll angles are also maintained at the arc-second level, with the accuracy of the three attitude angles reaching 1–2.1 arcsec, and the overall accuracy is better than those of mainstream star-sensitive attitude determination algorithms. The next research plan is to analyze the stability of the algorithm in different directions and different visual fields and carry out dynamic attitude determination research.

## Acknowledgments

This research was funded in part by the National Natural Science Foundation of China (grant 41471381) and the Natural Science Foundation of Jiangsu Province of China (grants BK20171410 and BK20170782). Heartfelt thanks are also given for the comments and contributions of anonymous reviewers and members of the editorial team.

## References

- Davidson, J. K. and K. H. Hunt. 2004. Robots and screw theory: Applications of kinematics and statics to robotics. *Journal of Mechanical Design* 126(4):763–764.
- Gai, E., K. Daly, J. Harrison and L. Lemos. 1985. Star-sensor-based satellite attitude/attitude rate estimator. *Journal of Guidance, Control, and Dynamics* 8(5):560–565.
- Guan, Z., Y. Jiang, J. Wang and G. Zhang. 2019. Star-based calibration of the installation between the camera and star sensor of the LuoJia 1-01 satellite. *Remote Sensing* 11(18):2081.
- Habib, A., S. W. Shin, K. Kim, C. Kim, K.-I. Bang, E.-M. Kim and D.-C. Lee. 2007. Comprehensive analysis of sensor modeling alternatives for high resolution imaging satellites. *Photogrammetric Engineering and Remote Sensing* 73(11):1241–1251.
- Keat, J. E. 1977. *Analysis of Least-Squares Attitude Determination Routine DOAOP. Technical Report CSC/TM-77/6034*. Silver Spring, MD: Computer Sciences Corporation.
- Liebe, C. C. 2002. Accuracy performance of star trackers—A tutorial. *IEEE Transactions on Aerospace and Electronic Systems* 38(2):587–599.
- Markley, F. L. 1988. Attitude determination using vector observation and the singular value decomposition. *The Journal of the Astronautical Sciences* 36(3):245–258.
- Markley, F. L. 1993. Attitude determination using vector observations: A fast optimal matrix algorithm. *The Journal of the Astronautical Sciences* 41(2):261–280.
- Meygret, A., G. Blanchet, C. Latry, A. Kelbert and L. Gross-Colzy. 2019. On-orbit star-based calibration and modulation transfer function measurements for PLEIADES high-resolution optical sensors. *IEEE Transactions on Geoscience and Remote Sensing* 57(8):5525–5534.
- Shuster, M. D. 2006. The quest for better attitudes. *The Journal of the Astronautical Sciences* 54(3–4):657–683.
- Shuster, M. D. and S. D. Oh. 1981. Three-axis attitude determination from vector observations. *Journal of Guidance and Control* 4(1):70–77.
- Shuster, M. D., D. S. Pitone and G. J. Bierman. 1991. Batch estimation of spacecraft sensor alignments, I. Relative alignment estimation. *The Journal of the Astronautical Sciences* 39(4):519–546.
- van Wyk, F., Y. Wang, A. Khojandi and N. Masoud. 2020. Real-time sensor anomaly detection and identification in automated vehicles. *IEEE Transactions on Intelligent Transportation Systems* 21(3):1264–1276.
- Xie J.-F. 2009. *The Critical Technology of Data Processing of Satellite Attitude Determination Based on Star Sensor*. Ph.D. dissertation, Wuhan University.
- Xie J. F., X. M. Tang, F. Mo and Z. L. Liu. 2017. Domestic stellar image post-processing project design and system implementation for ZY-3 satellite. *Geomatics and Information Science of Wuhan University* 42(4):434–440.
- Zhang H., Y. Niu, J. Lu, C. Zhang and Y. Yang. 2017. On-orbit calibration for star sensors without priori information. *Optics Express* 25(15):18393–18409.

# Digital Building-Height Preparation from Satellite Stereo Images

Prakash P. S. and Bharath H. Aithal

## Abstract

Buildings are considered prominent objects for understanding the pattern of growth in an urban setting. Remote sensing technology plays a vital role in facilitating data generation pertaining to various urban applications. Digital surface models represent the elevation of the earth surface features, and can be obtained from stereo images, radar, laser scanning, and so on. Photogrammetric techniques applied to optical stereo satellite images are economical and fast ways to generate height information of buildings. In this work, a quantitative and qualitative analysis of digital surface models generated from Cartosat-1 stereo images is compared with openly available data. The study finds that it is possible to acquire about 50 percent of building heights with acceptable error limits. The experimental results indicate that the quality of height information is suitable for applications to assess urban development at a macro scale, but not for individual building-level modeling.

## Introduction

Cities across India have experienced unprecedented growth over the past three decades, posing numerous challenges to the ecosystem (Bharath *et al.* 2018). Bangalore has grown threefold in terms of its geographical area since the year 1990 (Ramachandra *et al.* 2013). In terms of land use change, the urban area of the city has increased from 8% to 60% over the past 40 years (Bharath *et al.* 2018). Buildings are prominent objects in the rapid expansion of paved surfaces in the urban environment. Information about building footprints and heights has several applications, including measuring urban growth (Taubenböck *et al.* 2017), evaluating rooftop solar potential (Kumar 2018), assessing disaster and hazard risk (Biljecki *et al.* 2017), planning telecommunication tower locations (Saikhom *et al.* 2016), and planning utilities and

detecting change (Qin *et al.* 2016). 3D building data sets are generated using multiple techniques such as laser (Wang 2013) and radar scanning (Qiu *et al.* 2018), stereo satellite images (Sharma *et al.* 2016), and aerial surveys (Suveg and Vosselman 2004). Optical stereo images have an advantage over other remote sensing methods such as laser and radar in terms of cost and suitability, although quality is a matter of concern.

Globally available free surface models such as SRTM, ASTER, ALOS PRISM, Cartosat-30, and ALOS PALSAR offer the possibility to obtain data generated in several ways (Misra *et al.* 2018). Publicly available surface models are useful for generating global-scale data, but at a finer scale their resolution appears limited. Earlier research studies indicate that deriving building height accurately from openly available data sets is challenging (Yap *et al.* 2019; Ren *et al.* 2020). The major limitation of these open data sets is that they are available for only a certain time period, and continuous temporal coverage is unavailable. However, deriving building heights from openly available digital elevation models (DEM) are less explored. Commercially available satellite images are vital in the preparation of digital building heights, because of their temporal coverage capability, higher resolution, and automated production pipeline. Satellite images captured with a stereo camera (Sharma *et al.* 2016) and multi-date images (Ghuffar 2018) have been to generate depth information. The details of generating a digital surface model (DSM) from each of these sensors are reviewed in work by Deilami and Hashim (2011). Several earlier studies evaluated the accuracy parameters of DSMs generated from satellite images, showing the effect of terrain (Bhardwaj *et al.* 2013), choice of algorithms (Eckert and Hollands 2010), number of ground control points (GCPs), and their distribution on the resultant quality of the surface generated. Among commercially available data sets, *Cartosat-1* offers low-cost stereo satellite images compared to other contemporaries, as shown in Table 1.

Table 1. Prominent satellite sensors with stereo imaging capabilities

Satellite Sensor	Resolution		Spectral Bands	Swath (km)	Availability Date	Approx. Cost/km <sup>2</sup>
	Pan	Multispectral				
<i>SPOT 6/7</i>	50 cm	50 cm	B, G, R, NIR	60	2014	\$8.15
<i>GeoEye-1</i>	46 cm	1.84 m	B, G, R, NIR	15	2008	\$55
<i>KOMPSAT-3</i>	40 cm	1.6 m	B, G, R, NIR	15	2012	\$48
<i>Pléiades-1B</i>	50 cm	2.0 m	B, G, R, NIR	20	2012	\$25
<i>Pléiades-1A</i>	50 cm	2.0 m	B, G, R, NIR	20	2011	\$25
<i>WorldView-2</i>	46 cm	1.84 m	B, G, R, NIR + 4	16	2009	\$35
<i>WorldView-3</i>	31 cm	1.24 m	B, G, R, NIR + 4	13	2014	\$35
<i>Ikonos</i>	82 cm	3.28 m	B, G, R, NIR	11	1999–2006	\$20
<i>Cartosat-1</i>	2.5 m	—	Panchromatic	27	2005	\$0.12
<i>ALOS</i>	2.5 m	10 m	B, G, R, NIR	35	2006	\$0.37

B = blue; G = green; NIR = near-infrared; R = red.

Prakash P. S. and Bharath H. Aithal are with the Ranbir and Chitra Gupta School of Infrastructure Design and Management, Indian Institute of Technology Kharagpur, West Bengal, India (psprakashgis@gmail.com, bharath@infra.iitkgp.ac.in).

Contributed by Valerie Gouet-Brunet, June 5, 2020 (sent for review January 5, 2021; reviewed by Linfu Xie, Li Yuan, Laurent Caraffa)

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 557–565.  
0099-1112/21/557–565

© 2021 American Society for Photogrammetry  
and Remote Sensing  
doi: 10.14358/PERS.87.8.557

In this work, building heights are obtained from the DSM of *Cartosat-1* stereo images using photogrammetric procedures. The mapping capability of the automatically generated surface model is analyzed to derive building heights. Various methods of generating a digital terrain model (DTM) are tested for their suitability in estimating building heights. An assessment of various open-source DSMs is performed to derive building heights. The accuracy of the DSM generated is evaluated using accurate GCPs and compared against freely available elevation models. The building-height estimates from the models are validated with the buildings' ground-measured heights.

This article is arranged as follows. In the next section, previous studies and their results are evaluated. The method adopted for the present work is described in the following section, including explanation of technical aspects related to DSM and DTM generation. Then come results and discussions, followed by the conclusions of the study.

## Related Works

The stereoscopic parallaxes in stereo models caused by building roofs and bases are used to measure the heights of buildings manually from satellite images (Zeng 2014). In automated techniques, shadows are used to compute the height of buildings from monocular images (Shettigara and Sumerling 1998). This method cannot overcome the quality issues related to detected shadows, as it is degraded by tall adjacent buildings and the spectral homogeneity of shadows and water (Raju *et al.* 2014). The DSMs derived from stereo satellite images are used for several applications, including estimating the heights of buildings (Misra *et al.* 2018) and trees (Immitzer *et al.* 2016). Computing the height of above-ground objects such as buildings or trees requires the generation of a bare-earth model or DTM from the DSM. The DSM represents the elevation of the earth surface, encompassing all above-ground objects, whereas the DTM represents the elevation of the bare earth surface after above-ground features are removed. Arefi *et al.* (2011) developed an iterative approach to producing a DTM from *Cartosat-1* stereo images. The DSM is filtered (Sithole and Vosselman 2004; Zhang *et al.* 2016) to obtain the nearest possible height estimations. Ranagalage *et al.* (2018) adopted a generalization technique by creating elevation grids throughout the area. Özcan *et al.* (2018) removed nonground objects using segmentation and later applied a region-growing algorithm to generate a DTM. Other work that has produced promising height estimates (Wurm *et al.* 2014), does so from a normalized DSM (nDSM) through resampling. A slope-based filtering technique was adopted by Misra *et al.* (2018) to prepare a DTM from an existing DSM to calculate digital building heights. Handayani *et al.* (2018) and Ranagalage *et al.* (2018) assume that buildings are above a plane surface having constant elevation. Alternatively, Zeng (2014) used zonal statistics, building polygons, and an nDSM to arrive at best approximate building heights.

## Method

The current work encompasses generating a DSM, deriving the DTM, evaluating the DSM in the estimation of building heights, and comparing the building heights generated with an openly available DEM. The overall method adopted to estimate building heights from stereo satellite images is shown in Figure 1. The following subsections aim to describe details of our method.

### Data Used

*Cartosat-1* stereo satellite images with a spatial resolution of 2.5 m were used to generate the DSM, which we compared with an openly available DEM from multiple sensors as given in Table 2. The accuracy of the stereo model is improved using 42 accurate GCPs, are collected across the study

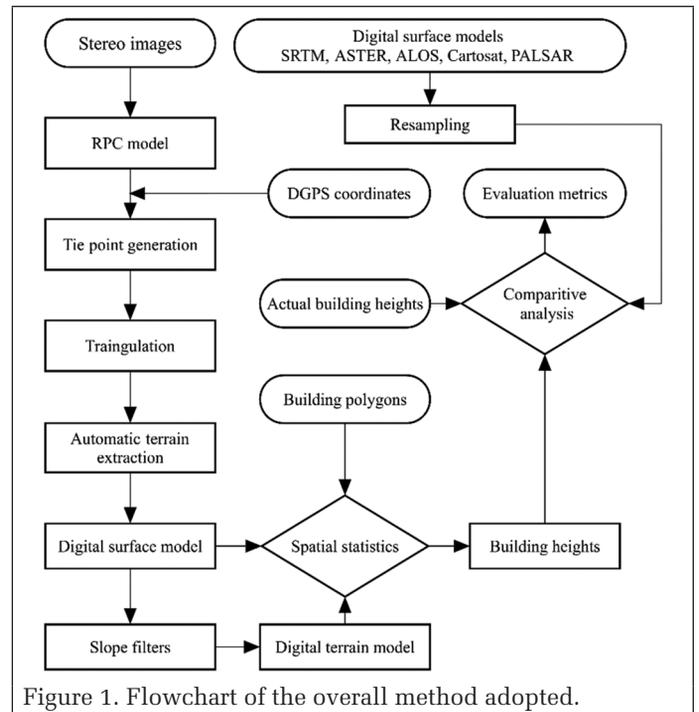


Figure 1. Flowchart of the overall method adopted.

Table 2. Spatial data sets used in the study.

Sensor	Spatial Resolution (m)	Type	Capture
ALOS PRISM	30	DEM	April 2015
ALOS PALSAR	12.5	DEM	August 2015
Cartosat-1	30	DEM	April 2015
ASTER	30	DEM	2000–2009
SRTM	30	DEM	September 2014
Cartosat-1	2.5	Stereo images	March 2018

DEM = digital elevation model.

area. For evaluating building-height estimation from the model output, we used 190 various kinds of building heights measured manually. The necessary building polygons were digitized manually.

### DSM Generation

The *Cartosat-1* data sets acquired consist of stereo images and a corresponding rational polynomial coefficient (RPC) file (Hu and Tao 2002; Sharma *et al.* 2015) for each image. The RPC files are used to describe the mathematical relationship of images—that is, normalized pixel coordinates ( $I_N$  and  $S_N$ ) and normalized ground coordinates ( $\varphi_N$ ,  $\lambda_N$ , and  $h_N$ ) at the time of image acquisition. Here  $h_N$  is the normalized elevation above ellipsoid, and the pixel coordinates (usually with respect to the World Geodetic System) are given by

$$I_N = \frac{\text{Num}_I(\varphi_N, \lambda_N, h_N)}{\text{Den}_I(\varphi_N, \lambda_N, h_N)} \quad (1)$$

$$S_N = \frac{\text{Num}_S(\varphi_N, \lambda_N, h_N)}{\text{Den}_S(\varphi_N, \lambda_N, h_N)} \quad (2)$$

where  $\text{Num}_I$ ,  $\text{Den}_I$ ,  $\text{Num}_S$ , and  $\text{Den}_S$  are third-order polynomials instead of sensor and satellite parameters. Initially, the stereo pair of images is oriented using the RPC files, and then accurate GCPs are used to refine the model and improve its accuracy (Rao *et al.* 2006). The stereo pair of images is subjected to interior orientation using the RPC files (Giribabu *et al.* 2013). Later, automatic tie points are generated, and

triangulation helps produce a more accurate model (Saha 2014). The enhanced automatic terrain extraction algorithm (Tsanis *et al.* 2014) is followed for DSM generation, resulting in the DSM shown in Figure 2. The output cell size is 5 m, twice the original image resolution (Gianinetta 2009). Initially, quasi-epipolar images are produced to limit the search range for the stereo matching to a single dimension. The algorithm used for object reconstruction (Hirschmuller 2008) minimizes the global energy function  $E$  for all disparities  $D$ :

$$E(D) = \sum_p \left( C(p, D_p) + \sum_{q \in N_p} P_1 T[|D_p - D_q| = 1] + T[P_2 T[|D_p - D_q| > 1]] \right) \quad (3)$$

Here  $C$  is a function that defines the matching cost for each pixel location  $p$  in the first image;  $N_p$  is the neighboring pixels of pixel  $p$ ; a logical function  $T$  is equal to 0 if the argument is false, and 1 otherwise;  $D$  is the disparity map computed by selecting disparities with minimum aggregated cost;  $P_1$  is the penalty

value for a disparity jump of one pixel, in cases of slant surfaces; and  $P_2$  is the penalty value for a disparity jump of more than one pixel, in situations where surfaces are discontinuous.

### DSM Quality Assessment

The vertical accuracy of the DSM generated from stereo images using automated terrain extraction was assessed using accurate GCPs, and a comparative analysis of the vertical accuracies of freely available DSMs was performed. The 42 GCPs used are shown in Figure 3. The control points are spread across the study area evenly and are obtained from a survey based on differential GPS (DGPS) carried out in the study area. By performing overlay analysis, we obtained the elevation corresponding to each point from the DSM. Then we calculated the root-mean-square error (RMSE) separately for each DSM height using the equation

$$RMSE = \sqrt{\frac{\sum (Z_{DGPS} - Z_{DSM})^2}{N}} \quad (4)$$

where  $Z_{DGPS}$  corresponds to the elevation of points measured using DGPS and  $Z_{DSM}$  is the corresponding elevation obtained from the DSM. Elevation profiles of the DSMs are generated for latitudes and longitudes 5 km apart, as shown in Figure 4.

### DTM Preparation

Generation of an elevation model excluding above-earth objects is a challenging task because of the complex nature of the urban environment. Since urban areas contain varied land use types spread haphazardly, it is difficult to generalize a single method to derive a terrain model. Earlier work demonstrated automated methods such as segmentation (Özcan *et al.* 2018), grid-based methods (Handayani *et al.* 2018; Ranagalage *et al.* 2018), multi-directional slope (MSD) filtering (Perko *et al.* 2015), and deep learning (Gevaert *et al.* 2018). Depending on the terrain complexity and the type of data used of surface generation, a suitable method needs to be developed. But it is unclear which DTM generation method will produce the best results for large areas and in undulating terrain. Hence, the idea of this study is to evaluate various methods and show the quality of their results. We tested five methods.

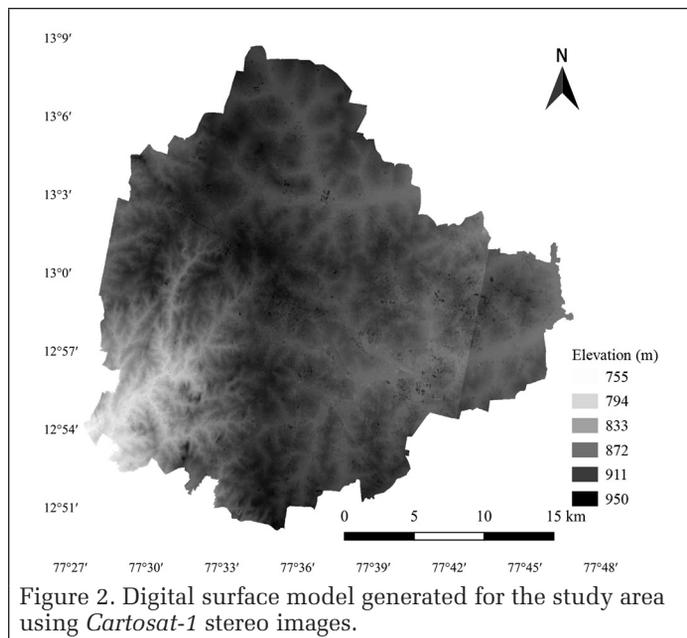


Figure 2. Digital surface model generated for the study area using *Cartosat-1* stereo images.

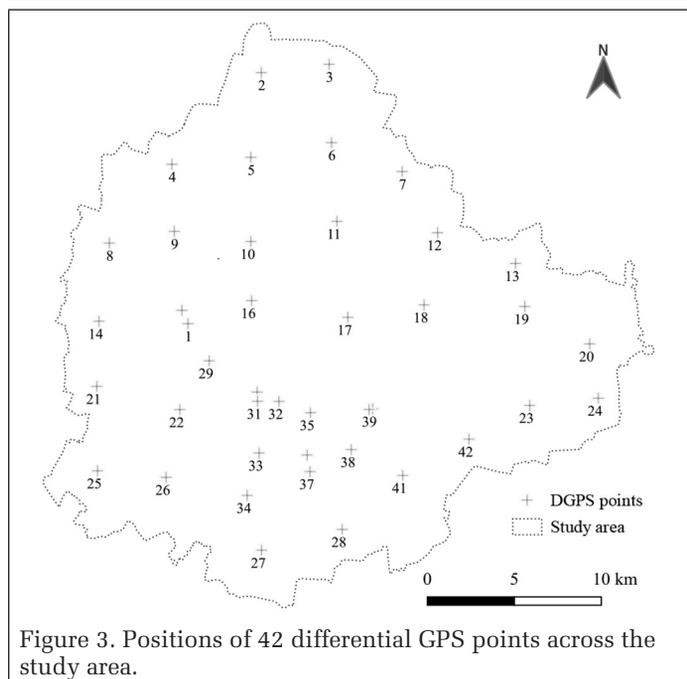


Figure 3. Positions of 42 differential GPS points across the study area.

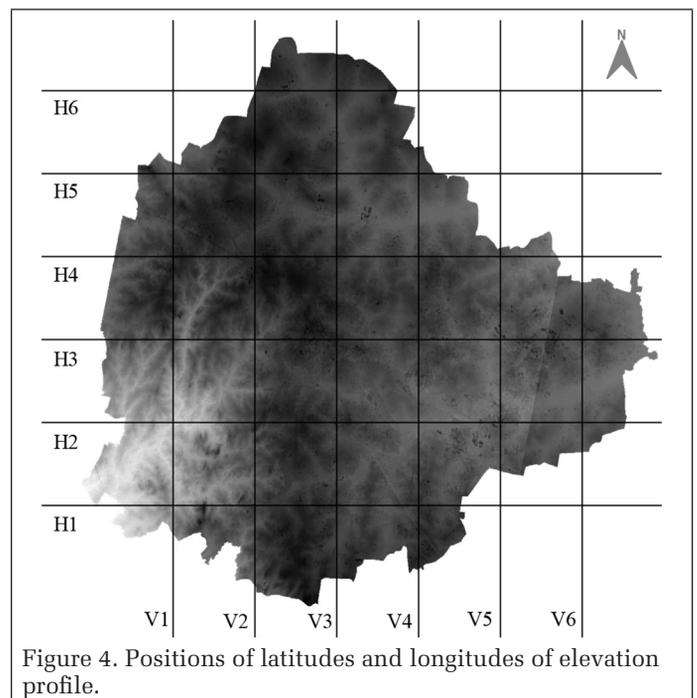


Figure 4. Positions of latitudes and longitudes of elevation profile.

### Multi-Directional Slope

The MSD method of filtering has achieved good results with a high-resolution DSM in undulating terrain (Mousa *et al.* 2017). Researchers have also demonstrated it for deriving a terrain model with a low-resolution DSM (Misra *et al.* 2018). The MSD technique requires fixing parameters such as height threshold, slope threshold, extent of the scan-line filter, and Gaussian smoothing kernel size. In this method, each pixel is classified as a terrain pixel or not by checking the height in five directions. If the height difference is more than the given threshold, the pixel is classified as nonground. This technique produces a terrain model that contains holes corresponding to nonground objects. After that, holes are filled with the nearest elevation values using linear interpolation techniques. The algorithm of MDS is explained using a flowchart given in Figure 5, where  $S_i$  is the slope threshold and  $Z_i$  is the height threshold given as input parameters.

### Grid-Based Method

This method involves two steps. In the first, the minimum elevation value is identified in a grid of fixed size (such as 50, 70, 100, or 120 m) throughout the study area by performing an overlay operation of the DSM and grid polygons. Then the minimum elevation value is assigned to each pixel that falls under that grid. The drawback of this method is that it creates discontinuity at the edges of the grid, causing wrong height estimates at those places. Our experiments showed that grid of 100 m produced the best results.

### Interpolation

In this technique, the lowest elevation value is initially identified in a grid and mapped to the center of the grid. The minimum value is then interpolated to the entire area using the inverse distance weighting method. This process is carried out throughout the study area to fill the elevation values. We tested various grid spacing values—such as 50, 70, 90, and 120 m—and selected 100 m for the best results.

### Slope-Based Filter

This technique was developed to derive a DTM from laser point-cloud data (Asal 2019; Vosselman 2000). It is similar to morphological erosion in grayscale imagery. Once the filter runs through the image, it creates a bare-earth surface by removing all above-ground objects such as buildings and trees. The mathematical representation of the technique is given by Vosselman (2000) in detail. The holes then need to be filled with elevation values using linear interpolation.

### Road Buffers

This is a simple technique in which the elevation values corresponding to roads in the study area are extracted and then extrapolated to the remaining places. The concept behind this idea is that roads represent zero heights throughout the scene, and all building heights are measured from the level of the road. The road layer is generated for the entire study area, with a buffer of 2.5 m on both sides of the road's center line. Through overlay analysis, all elevation values under the road buffer are clipped and saved as a separate raster layer. Then using linear interpolation, the area belonging to nonroad pixels is filled with elevation values.

### Building-Height Estimation

Two elevation models—a DTM and a DSM—are essential in estimating building height, in addition to building polygons. The building polygons that are needed to calculate height parameters are created using manual digitization. The overlay analysis is used to estimate the height of buildings as described later.

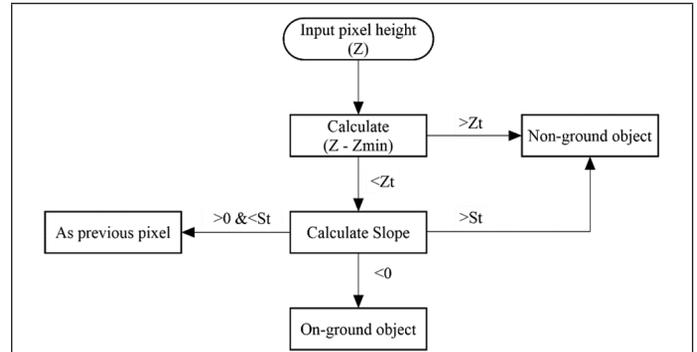


Figure 5. Flowchart showing the algorithm used to derive the digital terrain model.

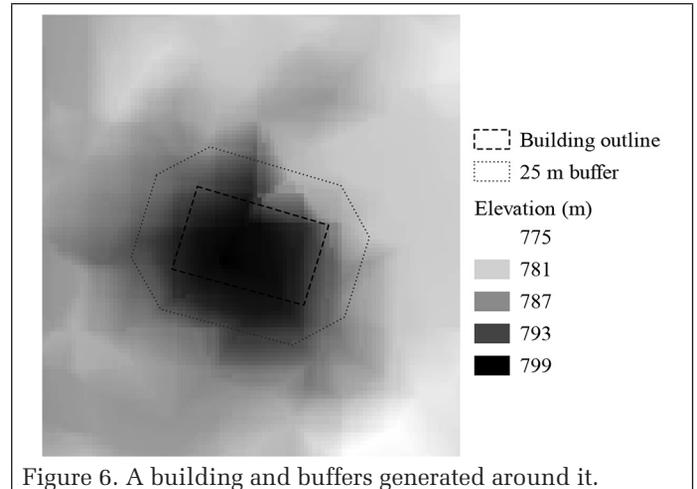


Figure 6. A building and buffers generated around it.

### DTM Method

The arithmetic subtraction of the DTM from the DSM produces a height map called the normalized DSM:

$$nDSM = DSM - DTM \quad (5)$$

The nDSM contains the height of above-ground objects such as trees and buildings. Using overlay analysis for each building polygon, the height corresponding to the maximum value is obtained. The height of each building is obtained from the nDSM maximum elevation value within the building polygon:

$$\text{Height of a building}_{(\text{polygon})} = \max(nDSM_{(\text{polygon})}) \quad (6)$$

### Building Buffers

Proposed by Sharma *et al.* (2016), a buffer area around each building is generated at a fixed distance of 25 m. Then from the DSM a pixel corresponding to the minimum value in that buffer zone is considered as the ground elevation. The maximum pixel value within a building polygon is considered the roof elevation. The difference between the two elevation values is considered the estimated building height. A pictorial representation of a building, buffer area, and DSM overlapped on imagery is shown in Figure 6.

## Results and Discussion

This work encompasses the generation of multiple products such as a DSM, DTM, and building heights. Each of these

Table 3. Accuracy metrics derived from 42 ground control points in the study area.

Model	Cartosat-30	ASTER	SRTM	ALOS PALSAR	ALOS PRISM	Cartosat 2.5 (Ours)
Root-mean-square error	5.26	7.24	4.39	4.57	4.67	9.85

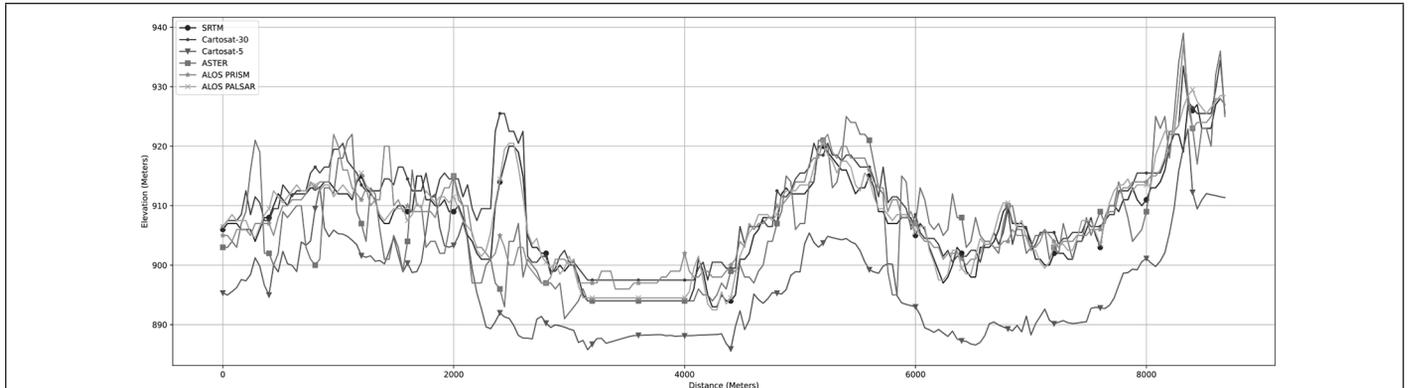


Figure 7. Elevation profile along section H2.

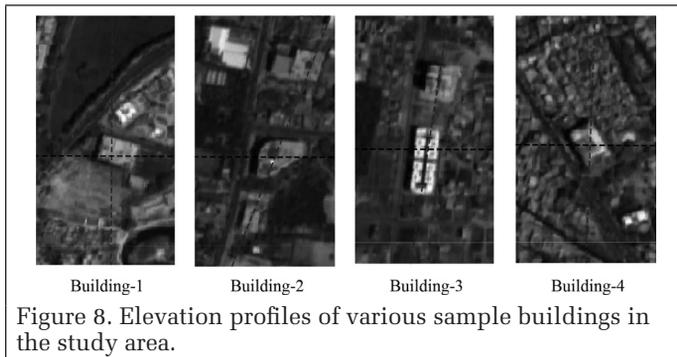


Figure 8. Elevation profiles of various sample buildings in the study area.

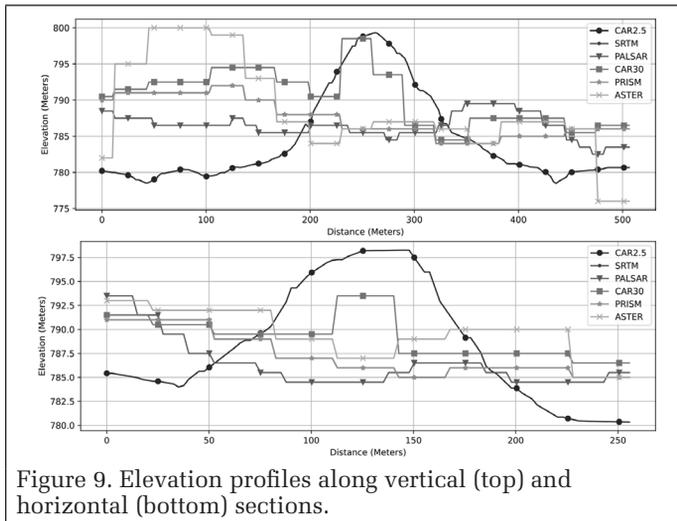


Figure 9. Elevation profiles along vertical (top) and horizontal (bottom) sections.

outputs is checked for quality using accuracy parameters. The quality of a DSM is assessed using accurate GCPs and drawing elevation profiles. The quality of a DTM is assessed using elevations of sample points and drawing elevation profiles at several places in the study area. The height estimates are tested against manually measured building heights.

### DSM Quality Assessment

RMSE values were calculated for 42 accurate GCPs measured across the study area using DGPS. Table 3 gives the RMSE values obtained for elevations of DEMs from various sources. The elevations obtained from SRTM and ALOS PRISM were found to be more accurate than those from other sources. The ASTER data have the maximum error among all openly available DEMs. The DEM generated from *Cartosat-1* stereo gave an RMSE value of 9.85, which is the highest among all other surface models. The elevations of Cartosat-30 and ALOS PALSAR are provided with respect to ellipsoidal height, unlike other DEMs, which provide geoidal height. In order to convert from

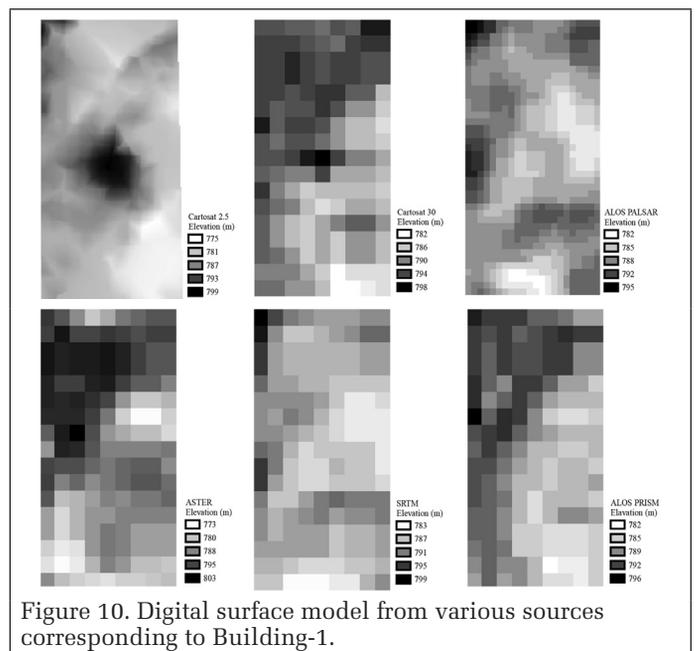


Figure 10. Digital surface model from various sources corresponding to Building-1.

ellipsoidal to geoidal elevation, a constant value of 86.5 m is used as an addition factor throughout the study area, which is the average value obtained from DGPS readings.

The openly available DEMs provide elevation at every 30 m as a pixel value. On the other hand, our model produced elevation at every 5 m as pixel value. This helps to generate heights of more ground features than openly available DEMs. Table 3 shows the comparison of RMSE values (geoidal height) of various DEMs and our model. The higher accuracies of open DEMs can be attributed to generalization and the radar method used for them. It is important to note that our model has 36 height values for each pixel of an openly available DEM.

Further, vertical profiles of DEMs were drawn along cross sections separated by 5 km, as shown in Figure 4. Figure 7 shows one such elevation profile along horizontal section H6. It gives the idea of land surface features such as a water body and certain hilly areas. It is evident from multiple elevation profiles that all DEMs follow the elevation profile of the land surface in a similar manner, though they do not capture the height variation of buildings suitably.

In order to assess the capability of a DSM to capture various kinds of buildings, elevation profiles are drawn at several places, as shown in Figures 8 and 9. From the elevation profiles, it is evident that openly available DEMs do not capture building geometry. On the other hand, the DSM generated using *Cartosat-1* stereo (shown in pink) gives significant elevation variation of the building and surrounding terrain.

Table 4. Difference in mean values of digital surface and terrain models under 190 buildings.

Method	Road Buffer	Multi-Directional Slope	Slope Filter	Grid 100 m	Inverse Distance Weighting
Mean difference	1.65	3.73	3.42	7.57	8.14

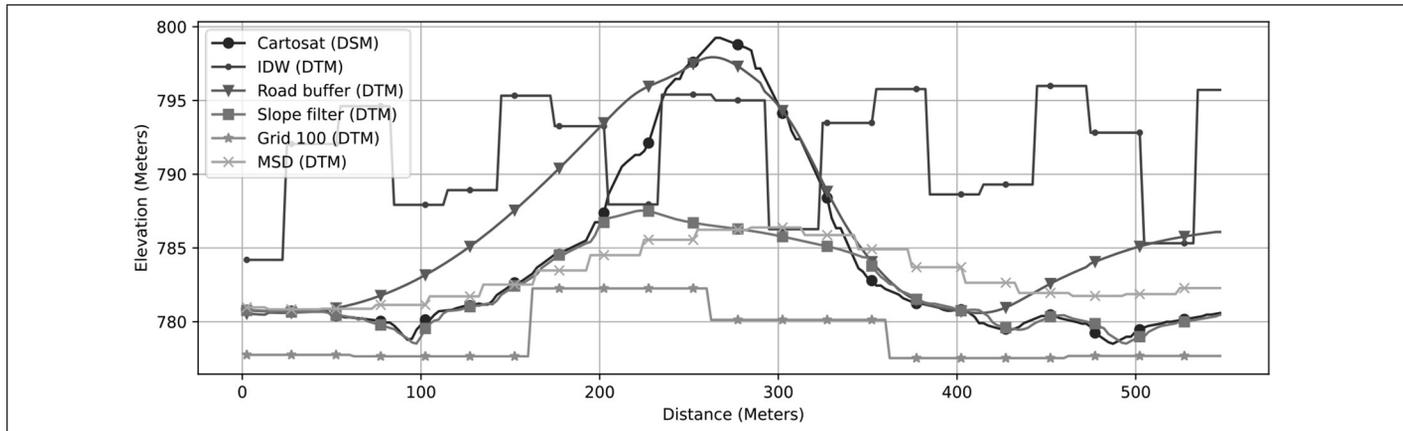


Figure 11. Elevation profiles of generated digital terrain model corresponding to Building-1 (vertical section).

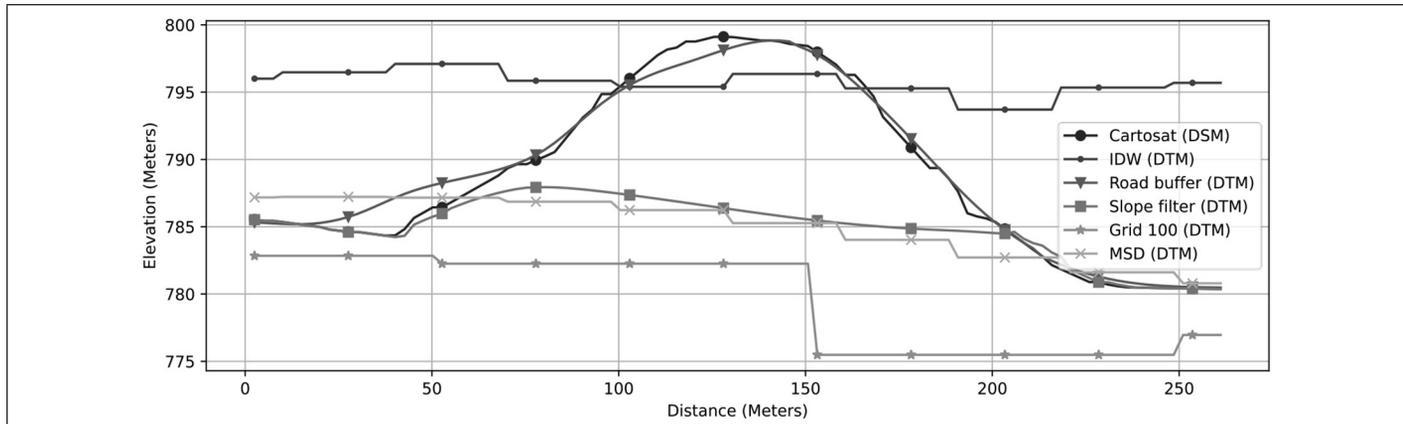


Figure 12. Elevation profiles of digital terrain model corresponding to Building-1 (horizontal section).

This can be clearly seen from the DSM in Figure 10. Other elevation models with a pixel size of 30 m do not capture the differences in elevation between building and ground surface. Even ALOS PALSAR, which has a spatial resolution of 12.5 m, fails to represent buildings in the scenes. The actual height of Building-1 is 20 m, according to field measurement. Observation of elevation profiles shows that an elevation variation from 780 to 800 (green line) obtained from *Cartosat-1* stereo images. From the elevation profiles, one could see that the corners of the building are not represented correctly.

#### DTM Quality Assessment

The DTMs obtained from the DSMs generated using *Cartosat-1* with multiple methods were compared. The mean differences of pixel values under 190 building polygons obtained for the various methods are given in Table 4. The elevation profiles of the DTMs produced using various methods are plotted and compared. From Figures 11 and 12, it is clear that elevation sampled on a 100-m grid tends to give maximum height estimations of the building features. Also, observation shows that MSD tends to produce a smoother DTM than the other methods, as shown in Figure 13. The DTM generated using the road-buffer method shows a profile similar to the DSM, which produces the least accurate building-height estimations.

#### Building-Height Estimation

Building-height estimation is performed using two methods: subtracting the DTM from DSM and using polygonal buffers around the buildings. In order to understand the estimation capability of the model output, the measured heights of 190 buildings were compared against estimated height. The details

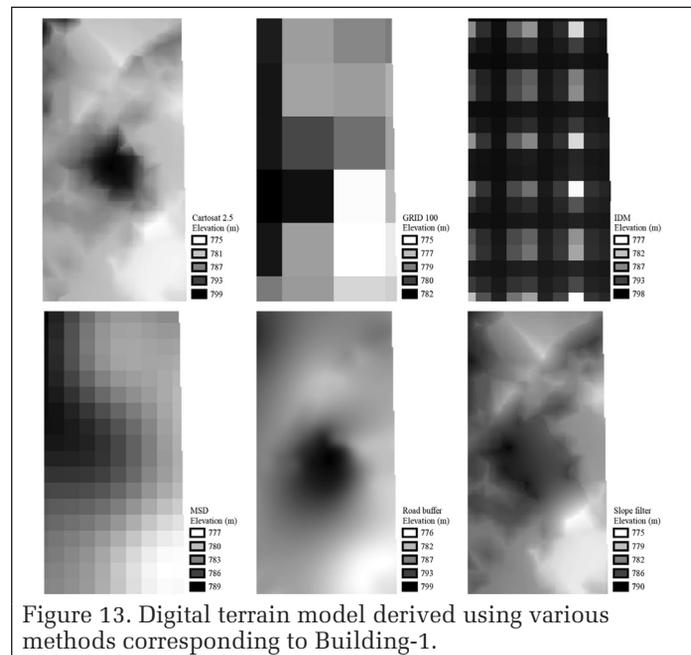


Figure 13. Digital terrain model derived using various methods corresponding to Building-1.

of building heights are given in Appendix A. Analysis of the building height derived from the model shows that about 50% of building heights were estimated within the stipulated error of  $\pm 10$  m for the current sensor capabilities. Among various

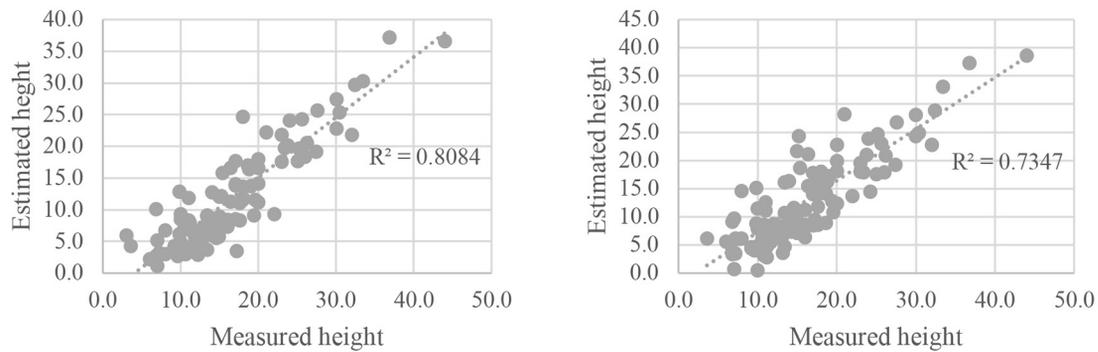


Figure 14. Relationship between estimated and measured building height using building buffers (left) and grid-based method (right).

DTM-generation methods for height estimation, the grid-based method proved to be most efficient, followed by interpolated surface. The heights generated from the building-buffer method proved to be the most suitable and time-efficient for a large geographical area. Figure 14 shows the relationship between measured and estimated height using grid-based and building-buffer techniques. It indicates that the model estimates the heights of buildings reasonably well in the range of 10 to 40 m. The higher height estimates can be attributed principally to two aspects. First, the height of the buildings is measured to the level of roof and plinth, as physical access to the top of the buildings is restricted, but DSM generation is based on the highest objects prominently visible on the top of the roof. Second, both methods use maximum and minimum pixel values from the DSM to calculate the height, which induces a certain amount of higher height estimation.

## Conclusions

This work focuses on procedural aspects estimating building heights from optical satellite stereo images. Initially, various imaging sensors that are commercially available were reviewed, and *Cartosat-1* stereo images were chosen for experimentation. As per previous research findings, the photogrammetric procedure was used instead of measuring shadow lengths to derive building heights. The automated photogrammetric techniques generate a DSM as base data for height estimation. In addition to this, openly available DEMs such as ASTER, SRTM, ALOS PRISM, and ALOS PALSAR (Appendix E), were compared with the generated DSM. Among them, SRTM showed the minimum deviation from the actual ground elevation. Two distinct approaches were used to estimate the heights: nDSM and generating buffers around each building. Various approaches were tested to derive a DTM, and grid-based generalization of height produced the closest height estimations. A slope-based filter generated a smoother surface, but its ability to estimate height was less than the grid-based method. The building-buffer method produced better estimates than all other nDSM-based techniques, and is computationally more straightforward and faster, since DTM generation is not necessary. Among openly available DEMs, ALOS PRISM showed better representation of buildings in the study area, followed by Cartosat-30.

The DSM generated from *Cartosat-1* stereo images is capable of estimating the heights of about 50% of buildings in the study area within the acceptable error limit of  $\pm 10$  m. Careful observation of the DSM shows that buildings that have large roof area with bright colors tend to give better representation. High-rise and small buildings are not captured appropriately by the model. Also, DSMs cannot capture the shape of buildings at the current resolution of 2.5 m. The experiment shows that the current sensor capability is suitable for macro-scale urban applications but not for individual assessment of buildings. In future, a similar analysis of height estimates can be performed on higher-resolution optical stereo images, to cover most

buildings present in the study area. Surface models generated from other methods, such as satellite microwave scanning, can be compared with optimal imaging methods. Further studies may include analysis of a large number of building heights measured from various parts of cities as well. The capability of sensors to capture surface features such as trees and building needs to be tested for flat as well as hilly terrain.

## Acknowledgments

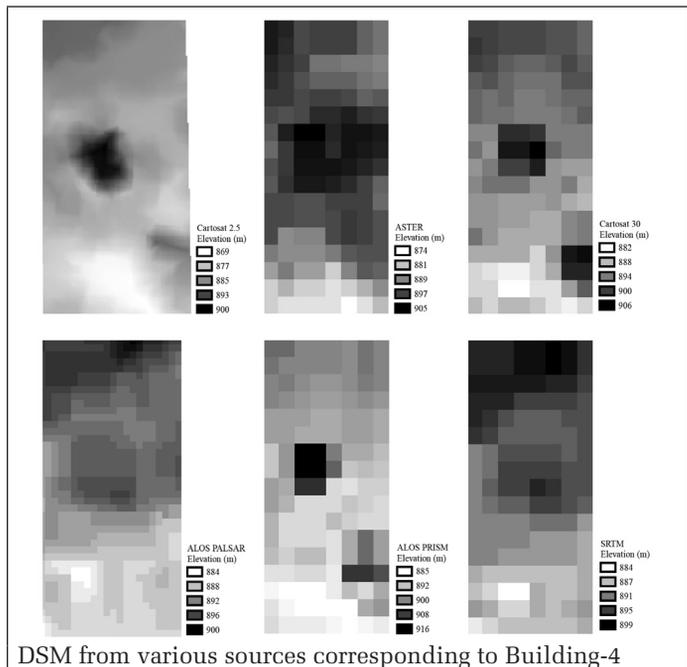
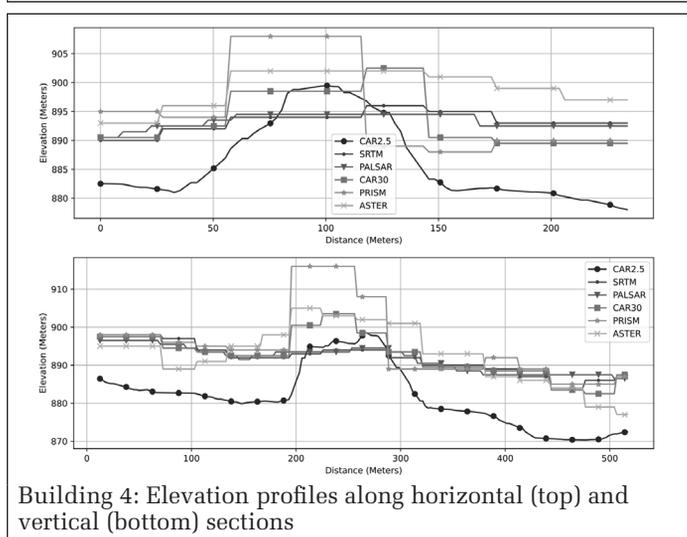
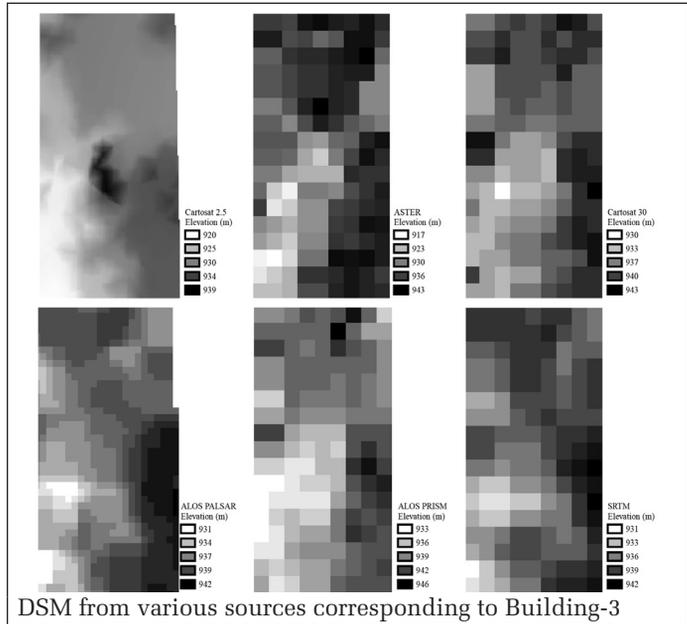
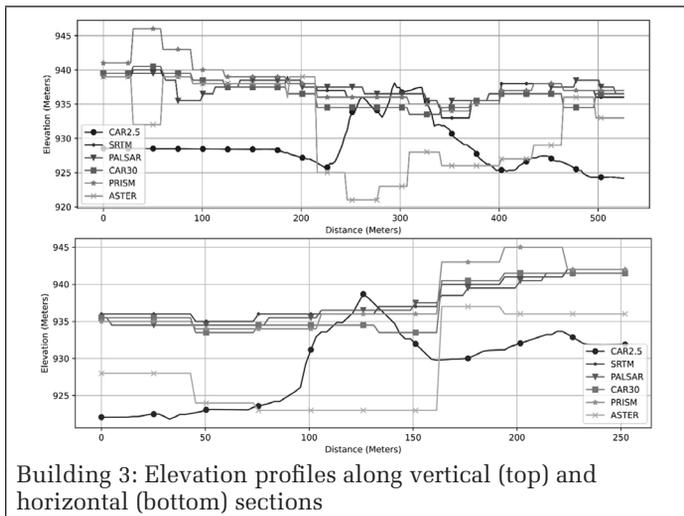
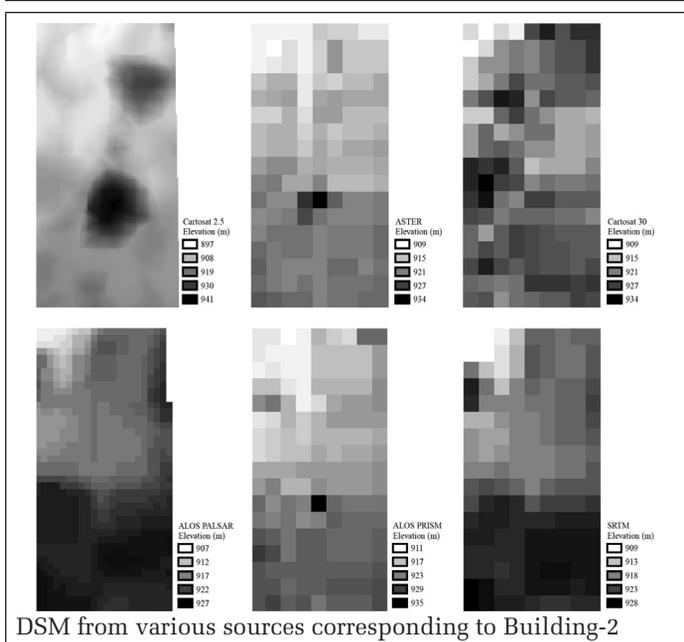
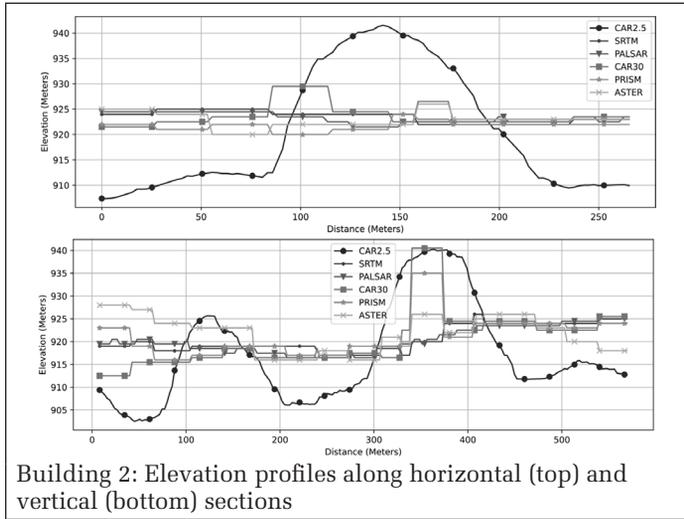
We are thankful to the Ranbir and Chitra Gupta School of Infrastructure Design and Management, NRDMS, Ministry of Science and Technology, Government of India, for financial and infrastructural assistance.

## References

- Arefi, H., P. d'Angelo, H. Mayer and P. Reinartz. 2011. Iterative approach for efficient digital terrain model production from CARTOSAT-1 stereo images. *Journal of Applied Remote Sensing* 5(1):053527.
- Asal, F.F.F. 2019. Comparative analysis of the digital terrain models extracted from airborne LiDAR point clouds using different filtering approaches in residential landscapes. *Advances in Remote Sensing* 8(2):93319.
- Bharath, H. A., M. C. Chandan, S. Vinay and T. V. Ramachandra. 2018a. Modelling urban dynamics in rapidly urbanizing Indian cities. *The Egyptian Journal of Remote Sensing and Space Science* 21(3):201–210.
- Bharath, H. A., S. Vinay, M. C. Chandan, B. A. Gouriand T. V. Ramachandra. 2018b. Green to gray: Silicon Valley of India. *Journal of Environmental Management* 206:1287–1295.
- Bhardwaj, A., R. S. Chaterjee and K. Jain. 2013. Evaluation of *Cartosat-1* satellite triangulation & DSMs in varied terrain conditions. *Scientific Research* 1(2):19–24.
- Biljecki, F., H. Ledoux and J. Stoter. 2017. Generating 3D city models without elevation data. *Computers, Environment and Urban Systems* 64:1–18.
- Deilami, K. and M. Hashim. 2011. Very high-resolution optical satellites for DEM generation: A review. *European Journal of Scientific Research* 49(4):542–554.
- Eckert, S. and T. Hollands. 2010. Comparison of automatic DSM generation modules by processing IKONOS stereo data of an urban area. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 3(2):162–167.
- Gevaert, C. M., C. Persello, F. Nex and G. Vosselman. 2018. A deep learning approach to DTM extraction from imagery using rule-based training labels. *ISPRS Journal of Photogrammetry and Remote Sensing* 142:106–123.
- Ghuffar, S. 2018. DEM generation from multi satellite PlanetScope imagery. *Remote Sensing* 10(9):1462.
- Gianinetto, M. 2009. Evaluation of *Cartosat-1* multi-scale digital surface modelling over France. *Sensors* 9(5):3269–3288.

- Giribabu, D., P. Kumar, J. Mathew, K. P. Sharma and Y.V.N. Krishna Murthy. 2013. DEM generation using *Cartosat-1* stereo data: issues and complexities in Himalayan terrain. *European Journal of Remote Sensing* 46(1):431–443.
- Handayani, H. H., Y. Murayama, M. Ranagalage, F. Liu and D.M.S.L.B. Dissanayake. 2018. Geospatial analysis of horizontal and vertical urban expansion using multi-spatial resolution data: A case study of Surabaya, Indonesia. *Remote Sensing* 10(10):1599.
- Hirschmuller, H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(2):328–341.
- Hu, Y. and C. V. Tao. 2002. Updating solutions of the rational function model using additional control information. *Photogrammetric Engineering and Remote Sensing* 68(7):715–723.
- Immitzer, M., C. Stepper, S. Böck, C. Straub and C. Atzberger. 2016. Use of WorldView-2 stereo imagery and National Forest Inventory data for wall-to-wall mapping of growing stock. *Forest Ecology and Management* 359:232–246.
- Kumar, A. 2018. Solar Potential Analysis of Rooftops Using Satellite Imagery. <<https://arxiv.org/abs/1812.11606>> Accessed DD Month YYYY.
- Misra, P., R. Avtar and W. Takeuchi. 2018. Comparison of digital building height models extracted from AW3D, TanDEM-X, ASTER, and SRTM digital surface models over Yangon City. *Remote Sensing* 10(12):2008.
- Mousa, Y.A.-K., P. Helmholz and D. Belton. 2017. New DTM extraction approach from airborne images derived DSM. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42(1/W1):75–82.
- Özcan, A. H., C. Ünsalan and P. Reinartz. 2018. Ground filtering and DTM generation from DSM data using probabilistic voting and segmentation. *International Journal of Remote Sensing* 39(9):2860–2883.
- Perko, R., H. Raggam, K. H. Gutjahr and M. Schardt. 2015. Advanced DTM generation from very high resolution satellite stereo images. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II(3/W4):165–172.
- Qin, R., J. Tian and P. Reinartz. 2016. 3D change detection—approaches and applications. *ISPRS Journal of Photogrammetry and Remote Sensing* 122:41–56.
- Qiu, C., M. Schmitt and X. X. Zhu. 2018. Towards automatic SAR-optical stereogrammetry over urban areas using very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing* 138:218–231.
- Raju, P.L.N., H. Chaudhary and A. K. Jha. 2014. Shadow analysis technique for extraction of building height using high resolution satellite single image and accuracy assessment. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL(8):1185–1192.
- Ramachandra, T. V., B. H. Aithal and S. Vinay. 2013. Land use land cover dynamics in a rapidly urbanizing landscape. *SCIT Journal* 13:1–12.
- Ranagalage, M., R. C. Estoque, H. H. Handayani, X. Zhang, T. Morimoto, T. Tadono and Y. Murayama. 2018. Relation between urban volume and land surface temperature: A comparative study of planned and traditional cities in Japan. *Sustainability* 10(7):2366.
- Rao, B. S., A.S.R.K.V. Murali Mohan, K. Kalyanaraman and K. Radhakrishnan. 2006. Evaluation of Cartosat-I stereo data of Rome. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* NNN(NN):1026–1029.
- Ren, C., M. Cai, X. Li, Y. Shi and L. See. 2020. Developing a rapid method for 3-dimensional urban morphology extraction using open-source data. *Sustainable Cities and Society* 53:101962.
- Saha, K. 2014. DSM extraction and evaluation from *Cartosat-1* stereo data for Bhopal city, Madhya Pradesh. *International Journal of Scientific and Research Publications* 4(5):56–60.
- Saikhom, V., D. Chutia, P. S. Sing, P.L.N. Raju and S. Sudhakar. 2016. A novel geospatial approach for identifying optimal sites for setting-up of mobile telecom towers strategically. *Journal of Geomatics* 10(2):192–198.
- Sharma, S. A., R. Agrawal and P. Jayaprasad. 2016. Development of “3D city models” using IRS satellite data. *Journal of the Indian Society of Remote Sensing* 44(2):187–196.
- Shettigara, V. K. and G. M. Sumerling. 1998. Height determination of extended objects using shadows in SPOT images. *Photogrammetric Engineering and Remote Sensing* 64(1):35–44.
- Sithole, G. and G. Vosselman. 2004. Experimental comparison of filter algorithms for bare-Earth extraction from airborne laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 59(1–2):85–101.
- Suveg, I. and G. Vosselman. 2004. Reconstruction of 3D building models from aerial images and maps. *ISPRS Journal of Photogrammetry and Remote Sensing* 58(3–4):202–224.
- Taubenböck, H., I. Standfuß, M. Wurm, A. Krehl and S. Siedentop. 2017. Measuring morphological polycentricity—a comparative analysis of urban mass concentrations using remote sensing data. *Computers, Environment and Urban Systems* 64:42–56.
- Tsanis, I. K., K. D. Seiradakis, I. N. Daliakopoulos, M. G. Grillakis and A. G. Koutroulis. 2014. Assessment of GeoEye-1 stereo-pair-generated DEM in flood mapping of an ungauged basin. *Journal of Hydroinformatics* 16(1):1–18.
- Vosselman, G. 2000. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing* 33(B3/2, part 3):935–942.
- Wang, R. 2013. 3D building modeling using images and LiDAR: A review. *International Journal of Image and Data Fusion* 4(4):273–292.
- Wurm, M., P. d’Angelo, P. Reinartz and H. Taubenböck. 2014. Investigating the applicability of *Cartosat-1* DEMs and topographic maps to localize large-area urban mass concentrations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7(10):4138–4152.
- Yap, L., L. H. Kandé, R. Nouayou, J. Kamguia, N. A. Ngouh and M. B. Makuate. 2019. Vertical accuracy evaluation of freely available latest high-resolution (30 m) global digital elevation models over Cameroon (Central Africa) with GPS/leveling ground control points. *International Journal of Digital Earth* 12(5):500–524.
- Zeng, C. 2014. Automated Building Information Extraction and Evaluation from High-Resolution Remotely Sensed Data. Ph.D. dissertation, *University of Western Ontario*.
- Zhang, Y., Y. Zhang, Y. Zhang and X. Li. 2016. Automatic extraction of DTM from low resolution DSM by two-steps semi-global filtering. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 3(3):249–255.

# Appendix A: Elevation profiles at various sample buildings in the study area



---

## IN-PRESS ARTICLES

- Longjie Ye, Wen Xiao, Yehua Sheng, Dong Su, Pengbo Wang, Shan Zhang, Na Zhao, Hui Chen, and Ka Zhang. Gaussian mixture model of ground filtering based on hierarchical curvature constraints for airborne LiDAR point clouds.
- Jonathan B. Thayne, Aaron M. Paque, and Megan C. Maher. Detecting Geo-Positional Bias in Imagery Collected using Small UASs.
- Chengming Ye, Hongfu Li, Ruilong Wei, Lixuan Wang, Tianbo Sui, Wensen Bai, and Saied Pirasteh. Twice Adaptive Intensity Threshold Method for Uneven LiDAR Data to Extract Road Markings.
- Majid Rahimzadegan, Arash Davari, and Ali Sayyadi. Estimating regional soil moisture with synergistic use of AMSR2 and MODIS images.
- Guoqing Zhou, Man Yuan, Xiaozhu Li, Hongjun Sha, Jiasheng Xu, Bo Song, and Feng Wang. Optimal regularization method based on L-curve for solution of the RFM parameters.
- Lisa M. LaForest, Tian Zhou, Seyyed Meghdad Hasheminasab, and Ayman Habib. System Calibration Including Time Delay Estimation for GNSS/INS-Assisted Push-Broom Scanners Onboard UAV Platforms.
- He Yanjun, Xu Jun, Li Jiansong, Peng Hao, Wu Bin. Information Extraction from High-resolution Remote Sensing Images Based on Multi-scale Segmentation and Case-based Reasoning.
- Radhika Ravi, Ayman Habib. Least Squares Adjustment with a Rank-deficient Weight Matrix and its Applicability towards Image/LiDAR Data Processing.
- Saket Gowravaram, Haiyang Chao, Andrew Molthan, Tiebiao Zhao, Pengzhi Tian, Harold Flanagan, Lori Schultz, and Jordan Bell. Spectral Reflectance Estimation of UAS Multispectral Imagery Using Satellite Cross-Calibration Method.
- Toshihiro Sakamoto. Early Classification Method for US Corn and Soybean by Incorporating MODIS-Estimated Phenological Data and Historical Classification Maps in Random-Forest Regression Algorithm.
- Mehdi Khoshboresh-Masouleh and Reza Shah-Hosseini. A Deep Multi-Modal Learning Method and a New RGB-Depth Data Set for Building Roof Extraction.
- Weining Zhu, Zeliang Zhang, Zaiqiao Yang, Shuna Pang, Jiang Chen, Qian Cheng. Spectral Probability Distribution of Closed Connected Water and Remote Sensing Statistical Inference for Yellow Substance.
- Abdullah Kayı, Bülent Bayram, and Dursun Zafer Şeker. The Analysis on the Annual Change of Digital Aerial Camera's IMUs Bore-sight Misalignment.
- Huan Xie, Binbin Li, Shijie Liu, Xiaohua Tong, Hong Tang, Xu Wang. A Method of Extracting High-accuracy Elevation Control Points from ICESat-2 Altimetry Data.
- Forrest Corcoran, Christopher Parrish. Diffuse Attenuation Coefficient (Kd) from ICESat-2 ATLAS Spaceborne Lidar Using Random Forest Regression.
- Samia Boukir, Li Guo, Nesrine Chehata. Improving remote sensing multiple classification by data and ensemble selection.
- Hari Shankar, Arijit Roy, Prakash Chauhan. Persistent Scatterer Interferometry for Pettimudi (India) Landslide Monitoring using Sentinel-1A Images.
- Ying Dong, Bo Yu, Fang Chen, Lei Wang, Ning Wang, Aqiang Yang. MSegnet, a practical network for building detection from high spatial resolution images.
- Lei Yan, Feizhou Zhang, Xufang Liu, Yun Xiang, Zihan Zhang, Siyuan Liu. Estimation of rock characteristics based on polarization spectra: surface roughness, composition and density.

# Enhanced Lunar Topographic Mapping Using Multiple Stereo Images Taken by Yutu-2 Rover with Changing Illumination Conditions

Wenhui Wan, Jia Wang, Kaichang Di, Jian Li, Zhaoqin Liu, Peng Man, Yexin Wang, Tianyi Yu, Chuankai Liu, and Lichun Li

## Abstract

*In a planetary-rover exploration mission, stereovision-based 3D reconstruction has been widely applied to topographic mapping of the planetary surface using stereo cameras onboard the rover. In this study, we propose an enhanced topographic mapping method based on multiple stereo images taken at the same rover location with changing illumination conditions. Key steps of the method include dense matching of stereo images, 3D point-cloud generation, point-cloud co-registration, and fusion. The final point cloud has more complete coverage and more details of the terrain than that conventionally generated from a single stereo pair. The effectiveness of the proposed method is verified by experiments using the Yutu-2 rover, in which two data sets were acquired by the navigation cameras at two locations and under changing illumination conditions. This method, which does not involve complex operations, has great potential for application in planetary-rover and lander missions.*

## Introduction

In a planetary-rover exploration mission, topographic mapping plays a crucial role to support science and engineering operations. It provides fundamental terrain information for topographic analysis, obstacle detection, path planning, rover navigation, and so on (Li *et al.* 2005; Alexander *et al.* 2006; Olson *et al.* 2007; Di *et al.* 2020). Thus, the performance of topographic mapping largely determines the safety of the rover's travel and its effectiveness in fulfilling scientific exploration tasks.

Due to their low power consumption, rich information acquisition, and reliable performance, stereo cameras have been used as the primary payload for topographic mapping in rover missions. The captured stereo images are usually transmitted from the rover to the ground-control station and inputted to the data-processing system for image preprocessing, 3D reconstruction, visual localization, and further scientific analysis. This off-line processing methodology has been successfully applied in Mars and lunar rover missions, such as the Mars Exploration Rover missions and Mars Science Laboratory mission (Li *et al.* 2005; Alexander *et al.* 2006; Matthies *et al.* 2007; Di *et al.* 2008; Stein *et al.* 2019) and Chang'e-3 and Chang'e-4 rover missions (Z. Liu *et al.* 2015, 2020; Di *et al.* 2020). During such a rover mission, topographic mapping is performed routinely with stereo images captured at the

waypoints along the rover traverse. The topographic mapping products are used not only in obstacle detection and path planning but also in measurement and designation of science targets for detailed in situ exploration.

The rover topographic mapping techniques used in previous missions are basically based on stereovision methods (Matthies *et al.* 2007). Image matching is the core technique in stereovision and has been extensively researched in the fields of photogrammetry and computer vision. Image-matching methods can be broadly classified as sparse or dense matching. Sparse matching methods match distinguished features (e.g., interest points; Gruen 1985; Förstner and Gülch 1987; Harris and Stephens 1988) to provide high-precision but sparse correspondence between the stereo images. Dense matching methods provide a pixel-by-pixel disparity map for dense 3D point-cloud generation (Scharstein and Szeliski 2002; Zhu *et al.* 2007), and have been applied in Mars and lunar rover exploration missions successfully (Alexander *et al.* 2006; Matthies *et al.* 2007; Di *et al.* 2008; Z. Liu *et al.* 2015, 2020). Recently, Semi-Global Matching and related optimized methods (Hirschmuller 2008; Michael *et al.* 2013; Besse *et al.* 2014) have become popular the vision-based 3D mapping applications because of their improved performance. Without requiring a manually designed model for image-feature extraction, deep learning-based dense matching methods have been proposed, using a large amount of proper training data (Luo *et al.* 2016; J.-R. Chang and Chen 2018). To overcome the limitation of the hard baseline of the stereo camera onboard a planetary rover, wide-baseline methods with image acquisition at two or more waypoints were developed for topographic mapping with higher accuracy for farther ranges (Olson *et al.* 2003; Di and Peng 2011; Wu *et al.* 2011).

Rover imagery of the lunar surface is often of poor texture, caused by the featureless surface itself or the severe influence of changes in illumination conditions due to the lack of atmosphere. This poses great challenges to the precision and completeness of topographic mapping products from stereo pairs of images. Recently, some illumination-invariant matching methods (Xu *et al.* 2016; Y.-J. Chang and Ho 2017) have been developed for urban or indoor scenes. These methods may not be directly applicable in matching of lunar rover images with different illumination conditions, though. Shading information associated with different illumination conditions has been exploited to produce lunar topographic products from orbital images using a photometric stereo technique (W. Liu *et al.* 2018; W. Liu and Wu 2020). The model and techniques are

Wenhui Wan, Kaichang Di, Zhaoqin Liu, Peng Man, and Yexin Wang are with the State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China (dikc@radi.ac.cn).

Jia Wang, Jian Li, Tianyi Yu, Chuankai Liu, and Lichun Li are with Beijing Aerospace Control Center, Beijing, China.

Contributed by Norbert Pfeifer, August 7, 2020 (sent for review January 21, 2021; reviewed by Wai Chung Liu, Cristina Re).

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 567–576.  
0099-1112/21/567–576

© 2021 American Society for Photogrammetry  
and Remote Sensing  
doi: 10.14358/PERS.87.8.567

also not directly applicable to lunar rover images, which are close-range perspective views.

During surface operations in a lunar rover mission—e.g., the *Chang'e-4* mission—the rover often stops and stays at the same location for quite a long time, for in situ exploration of the science targets or to hibernate in the extremely cold lunar night. If the rover took stereo images at the same location but at different times with different illumination conditions, would the multiple stereo images be helpful in producing better topographic products than a single stereo pair? This is the question that inspired us to begin this research.

This article proposes a novel and easy-to-implement method of rover-based lunar topographic mapping using multiple stereo images taken at the same location but with changing illumination conditions. The effectiveness of the proposed method is verified using stereo images acquired by the Yutu-2 rover. The enhanced lunar topographic mapping results, which have higher precision and more details, can be used in missions to support rover exploration planning and decision making. The proposed method is particularly useful when we want to get a higher-quality 3D model of the exploration area and science target but the rover cannot get closer to the target because of traversing constraints. It is also directly applicable to a lunar sample-return mission—e.g., the planned *Chang'e-5* mission—in which the stereo cameras are fixed on the lander, to provide enhanced mapping of the sampling areas.

The rest of the article is organized as follows. The next section elaborates the methodology and describes the unique image data sets acquired by the Yutu-2 rover for verification of the proposed topographic mapping method. In the section after that, the experimental results and detailed performance analysis based on the rover images are reported. Conclusions and future work are discussed in the last section.

## Methodology and Data

The main idea on which the method is proposed is that multiple stereo images taken at the same location and the same attitude but with different illumination conditions (i.e., different solar azimuth and elevation angles) can provide more information about the terrain than any single stereo pair in the data set. As we know, it is usually not possible to reconstruct the 3D terrain in shadow areas from one stereo pair through image matching. Multiple stereo images with different illumination conditions can solve this problem because shadow areas in one stereo pair may become illuminated in other stereo pairs. For a terrain patch not in shadow area, i.e., visible in all stereo images, the image-matching results of the same image pixel in the left images will be slightly different through subpixel matching (e.g., the least-squares matching method), because of the microstructural changes in visual appearance of the same patch in different stereo pairs. Consequently, the 3D coordinates of the same point in different stereo pairs will be different. Taking these 3D points as the repeated observations of the same physical point and merging them based on the measurement uncertainties will result in a 3D point with higher precision. Our proposed method aims to take advantage of multiple stereo images with different illumination conditions to enhance the capability of rover-based lunar topographic mapping.

The proposed lunar topographic mapping method, using stereo images captured in different illumination conditions, consists of three major steps: dense 3D point-cloud generation from each stereo pair, co-registration of multiple 3D point clouds, and fusion of surface point clouds. As shown in Figure 1, each pair of stereo images is inputted individually for dense matching and generation of 3D point clouds. Then, co-registration of multiple point clouds is performed through a global Iterative Closest Point (ICP) algorithm under the principle of 2D and 3D weighted-distance minimization.

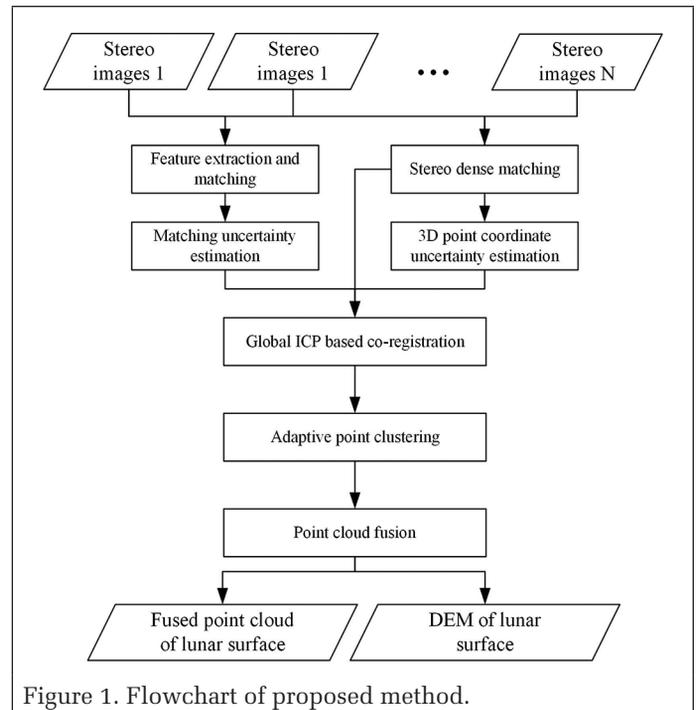


Figure 1. Flowchart of proposed method.

Theoretically, if the positions and attitudes of the multiple stereo pairs are exactly the same, this co-registration step can be skipped. But in practice, co-registration is necessary due to slight inconsistencies among the actual orientations of the stereo images caused by the control and measurement uncertainties of rover mast drives. Later, the co-registered 3D point clouds are fused into a merged surface point set with adaptive point clustering based on position consistency. Finally, the topographic products are provided in two forms: the fused point-cloud data, and the digital elevation model (DEM) generated by regular grid interpolation. The details of the proposed method follow.

### Multiple Stereo-Image Acquisition by Yutu-2 Rover at the Same Location

China's *Chang'e-4* probe successfully landed on the far side of the moon in the Von Kármán crater on 3 January 2019, and the rover Yutu-2 was released from the lander and started surface exploration on the same day (Di *et al.* 2019). The Yutu-2 rover carries three pairs of stereo cameras: the navigation camera (Navcam), the hazard-avoidance camera (Hazcam), and the panoramic camera (Pancam; Z. Liu *et al.* 2020). The Navcam and Hazcam are engineering sensors designed for navigation and obstacle avoidance, and have been used to acquire images at every waypoint. The Pancam is one of the science payloads and has been used at certain waypoints for long-term planning of scientific exploration. The Hazcam is fixed on the front of the rover with a downward angle of 30°. The Navcam and Pancam are mounted on the same camera bar atop the rover mast and can acquire stereo images pointing to different azimuth and elevation angles by rotations through the yaw and pitch drives. Controlled by the deployment drive, the camera mast is deployed to vertical position when working in a lunar day and stowed to horizontal position when hibernating in the lunar night. Before the *Chang'e-4* probe was launched, the interior orientation parameters and distortion-model parameters of the cameras were calibrated precisely in the calibration field containing evenly distributed 3D control points. The installation parameters of each camera and mast joints related to the rover body were also calibrated by cube prism-based measurements. The exterior orientation parameters of Navcam and Pancam images related to the rover body coordinate system can be calculated with the parameters of mast installation and yaw/pitch rotations.

In this study, in order to verify the proposed enhanced topographic mapping method, we designed the data-acquisition strategy and implemented it on the eighth, ninth, and 10th lunar day at the ground-control station, Beijing Aerospace Control Center. Specifically, three pairs of Navcam images were captured at waypoint LE01002 in the morning of the 10th lunar day. The captured stereo images and their detail views are shown in Figure 2.

The second data set contains three pairs of Navcam stereo images captured at waypoint LE00803 in the afternoon of the eighth lunar day and one pair in the morning of the ninth lunar day. The images are shown in Figure 3. The solar azimuth angles and elevation angles of the two image data sets in the local-site coordinate system are listed in Table 1.

As can be seen in the images of the two data sets, the areas of the shadowed image patches became larger with the decrease of the solar elevation angle. We note also that image pair 2d was captured with the opposite solar azimuth angle from image pairs 2a to 2c. Therefore, the shadow areas in image pairs 2a to 2c were imaged effectively through image pair

2d, which should be particularly helpful in reconstructing the lunar surface more completely.

### Dense Matching of Stereo Images and 3D Point-Cloud Generation

In our proposed method, dense matching is adopted for each pair of stereo images captured by the rover in order to obtain the 3D point cloud. Before the dense matching, it is necessary to perform image enhancement to improve the image quality, which is helpful for getting more matched points. We use methods of grayscale stretch and histogram specification to enlarge the image contrast of a single image and decrease the histogram differences between left and right images. Then the pinhole-based distortion model and calibrated relative exterior orientation parameters of the Navcam are adopted for epipolar rectification (Barnard and Fischler 1982).

The pinhole-based distortion correction model for Navcam imagery is

$$\begin{cases} x'_d = x'_s + x'_s(k_1r^2 + k_2r^4 + k_3r^6 + \alpha) + \beta y'_s + p_1(r^2 + 2x'_s y'_s) + 2p_2 x'_s y'_s \\ y'_d = y'_s + y'_s(k_1r^2 + k_2r^4 + k_3r^6) + 2p_1 x'_s y'_s + p_2(r^2 + 2y'^2_s) \end{cases} \quad (1)$$

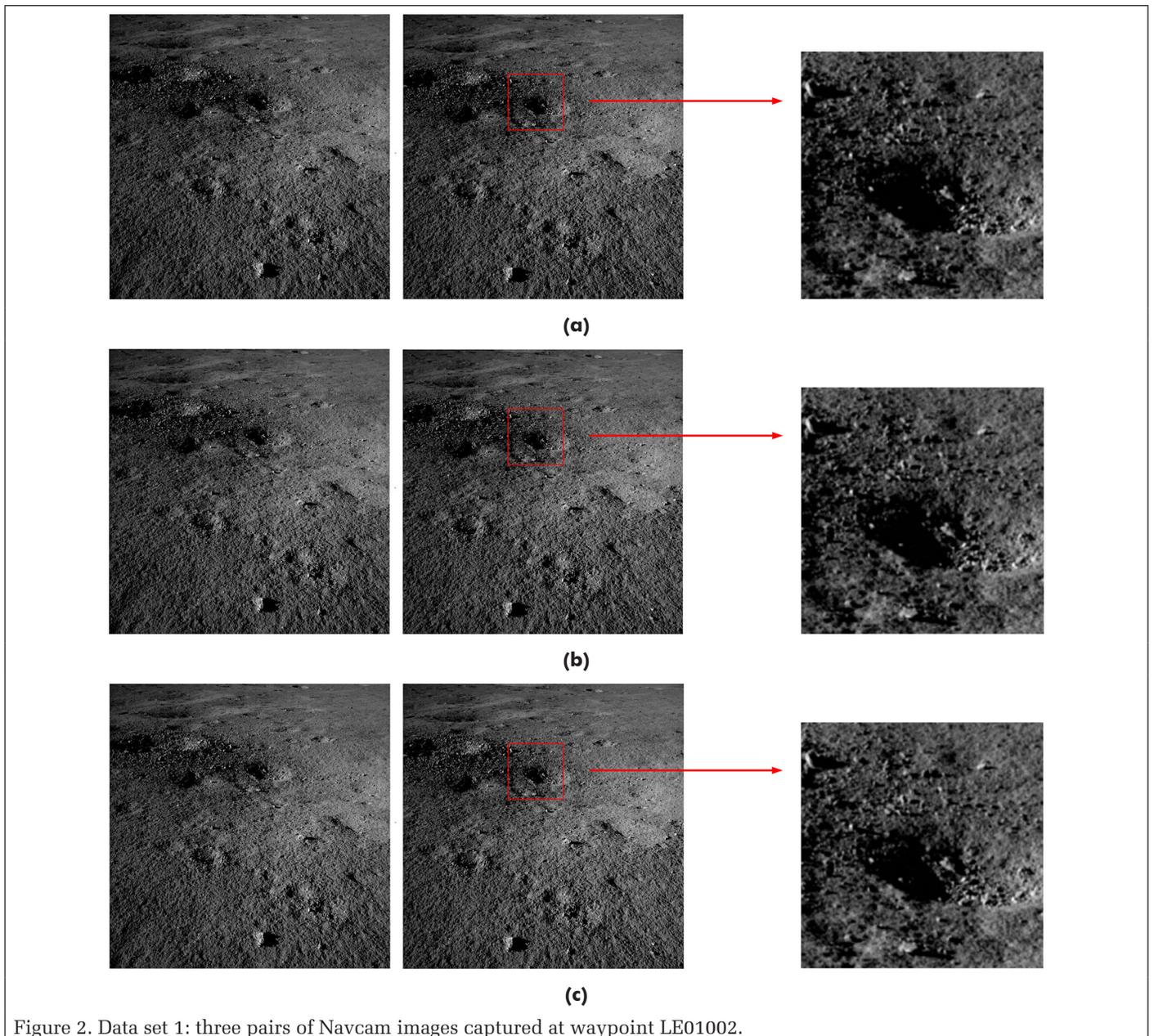


Figure 2. Data set 1: three pairs of Navcam images captured at waypoint LE01002.

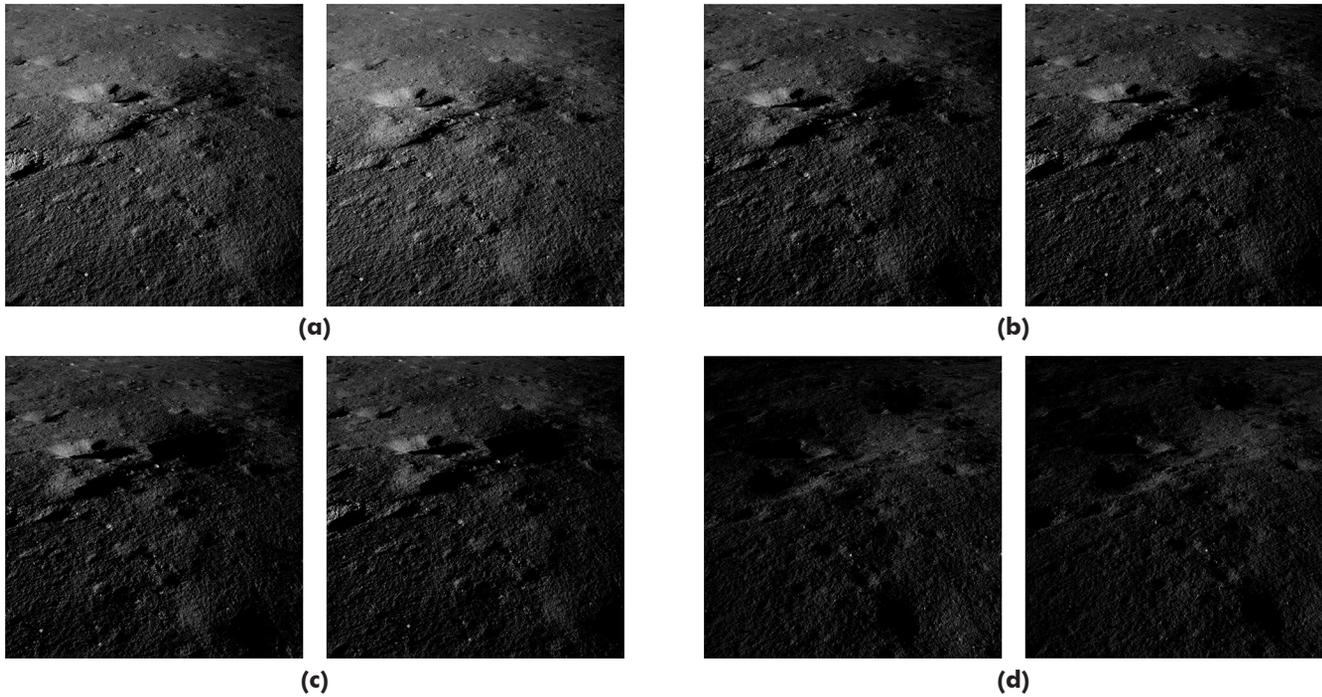


Figure 3. Data set 2: four pairs of Navcam images captured at waypoint LE00803.

where  $(x'_s, y'_s)$  and  $(x'_d, y'_d)$  are the coordinates of the original point and the corrected point in the principal point-centered image-plane coordinate system in millimeters;  $k_1, k_2, k_3$  are the radial distortion coefficients;  $p_1, p_2$  are the tangential distortion coefficients;  $r^2$  is a coefficient calculated by  $r^2 = x'_s + y'_s{}^2$ ;  $\alpha$  is the non-square scale factor of the charge-coupled device;  $\beta$  is the non-orthogonal distortion coefficient of the charge-coupled device, and the point coordinates  $(x', y')$  in the image-plane coordinate system in millimeters are calculated by

$$\begin{cases} x' = (x - x_0) \cdot p_s \\ y' = (y - y_0) \cdot p_s \end{cases} \quad (2)$$

where  $(x, y)$  are the image coordinates in pixels, with the origin centered at the top left corner of the image;  $(x_0, y_0)$  are the principal-point coordinates in pixels; and  $p_s$  is the pixel size in millimeters.

Semi-Global Matching (Hirschmuller 2008) is applied to the generation of the disparity map by scanning the image pixel with the minimum in energy function along the x-axis of the epipolar rectified images. The energy function is represented as

$$E(D) = \sum_{\mathbf{p}} \left( C(\mathbf{p}, D_{\mathbf{p}}) + \sum_{\mathbf{q} \in N_p} P_1 T[|D_{\mathbf{p}} - D_{\mathbf{q}}| = 1] + \sum_{\mathbf{q} \in N_p} P_2 T[|D_{\mathbf{p}} - D_{\mathbf{q}}| > 1] \right) \quad (3)$$

where  $E(D)$  is the sum of all pixel matching costs for the disparities of image  $D$ ;  $C(\mathbf{p}, D_{\mathbf{p}})$  is the matching cost of pixel  $\mathbf{p}$  with disparity  $D_{\mathbf{p}}$ ;  $\mathbf{q}$  is the pixel set in neighborhood  $N_p$  of  $\mathbf{p}$ ; and  $P_1$  and  $P_2$  are the penalty coefficients for the pixels with disparity changes in  $\mathbf{q}$  of, respectively, one pixel and more than one pixel. Considering that the grayscale distributions of the same physical point in the left and right images are still slightly different, due to different observation angles and different responses of the left and right cameras, the ADCensus algorithm is used for calculating the matching cost (Mei *et al.* 2011). It is expressed as

Table 1. Solar azimuth angles and elevation angles of the two stereo-image data sets.

Stereo Image No.	Solar Elevation Angle (°)	Solar Azimuth Angle (°)	Acquisition Time (UTC+8:00)
1a	18.3	67.9	2019-09-24 09:36:02
1b	19.2	66.9	2019-09-24 11:53:28
1c	20.0	65.9	2019-09-24 14:19:27
2a	16.2	-71.8	2019-08-06 17:45:04
2b	11.4	-77.8	2019-08-06 20:07:27
2c	9.9	-78.7	2019-08-07 09:54:51
2d	7.8	80.4	2019-08-24 14:19:22

$$C_{\text{ADCensus}}(\mathbf{p}, D_{\mathbf{p}}) = \rho(C_{\text{Census}}(\mathbf{p}, D_{\mathbf{p}}), \lambda_{\text{Census}}) + \rho(C_{\text{AD}}(\mathbf{p}, D_{\mathbf{p}}), \lambda_{\text{AD}}) \quad (4)$$

where  $C_{\text{Census}}$  and  $C_{\text{AD}}$  are the cost values from the census transform and absolute differences, respectively;  $\lambda$  is the pre-set coefficient of cost estimation; and  $\rho(c, \lambda)$  is a function on variable  $c$  represented as

$$\rho(c, \lambda) = 1 - \exp\left(-\frac{c}{\lambda}\right) \quad (5)$$

The census cost  $C_{\text{Census}}$  is computed as the Hamming distance of the two bit strings, which are formed from sequential judgments whether the grayscale value of the centered pixel is higher or lower than that of each neighboring pixel. Then the disparity map is obtained by means of minimizing  $E(D)$  based on dynamic programming. Note that the disparity map only contains the disparity values in the horizontal direction of the image, because of epipolar rectification. Finally, least-squares matching (Gruen 1985) is applied to refine the disparity map to subpixel values. After dense matching, the 3D coordinates of matched points are calculated by space intersection with the exterior orientation parameters of the stereo images. As a result, 3D point-cloud data are generated for each stereo pair of images.

### Co-registration of Multiple Point Clouds

Due to the control and measurement uncertainties of the rover's mast drives that deploy the mast or rotate the camera bar, inconsistencies may exist among the 3D point clouds generated from the stereo Navcam images, even if they were acquired at the same waypoint with the same nominal pointing directions. Thus, point-cloud co-registration is performed to obtain the relative parameters of rotation and translation. ICP is a useful method for point-cloud registration by minimizing the mean square distance of two point clouds without feature point matching (Besl and McKay 1992). However, because of flat terrain in some exploration areas, the registration solution may fall into a local minimum. Considering that the images captured under changing illumination conditions still have a certain grayscale similarity, 2D feature points are extracted and matched among different stereo images, so as to constrain the global ICP registration.

To obtain the matched points robustly, the scale-invariant feature transform (Lowe 2004) is used to extract and match the feature points in the left images of stereo pairs. Then the corresponding points in the right images are matched by the least-square matching method. Outlier detection is performed to eliminate incorrectly matched points. Finally, the left image is divided into evenly spaced grids, and the point in each grid with the highest correlation coefficient is chosen to get evenly distributed matched points within and among stereo pairs. Afterwards, the geometric constraints between stereo pairs, constructed by the feature matched points, will be incorporated into the global ICP solution.

In our global ICP solution, we also consider the uncertainties of the 3D points by calculating weighted 3D distances. According to the principle of photogrammetry (Di and Li 2007), the 3D coordinate uncertainties of object point  $P$  can be estimated by

$$\begin{cases} \sigma_x = \sqrt{\left(\frac{Z_p^2}{Bf}\right)^2 \left(\frac{x}{f}\right)^2 \sigma_d^2 + \left(\frac{Z_p}{f}\right)^2 \sigma_x^2} \\ \sigma_y = \sqrt{\left(\frac{Z_p^2}{Bf}\right)^2 \left(\frac{y}{f}\right)^2 \sigma_d^2 + \left(\frac{Z_p}{f}\right)^2 \sigma_y^2} \\ \sigma_z = \frac{Z_p}{Bf} \sigma_d \end{cases} \quad (6)$$

where  $(\sigma_x, \sigma_y, \sigma_z)$  are the 3D errors (standard deviations) of point  $P$ ;  $B$  and  $f$  are the baseline and focal length, respectively, of the stereo cameras;  $Z_p$  is the z-axis (perpendicular to the image plane) coordinate of  $P$  in the camera reference coordinate system;  $(x, y)$  are the 2D coordinates of the corresponding image point  $p$ ;  $(\sigma_x, \sigma_y)$  are the 2D measurement errors of  $p$ ; and  $\sigma_d$  is the error of horizontal disparity of  $p$ , represented by the stereo-matching error of  $p$  and calculated as (Zhang and Zhang 1997)

$$\sigma_d = \sqrt{\frac{2(1-\rho^2)}{N} \cdot \left(\frac{\sigma_g^2}{\sigma_g'^2}\right)} \quad (7)$$

where  $\rho$  is the correlation coefficient of the matched points,  $N$  is the number of pixels in the matching window,  $\sigma_g$  is the grayscale variance in the matching window, and  $\sigma_g'$  is the variance of first-order grayscale difference. By combining Equations 6 and 7, the 3D coordinate uncertainties of the point cloud are estimated for the following co-registration.

In our global ICP solution, the optimization principle is built on minimized 2D and 3D weighted distance, which is expressed as

$$L(R, T) = a \cdot D_{2D}(R, T) + b \cdot D_{3D}(R, T) \quad (8)$$

where  $R$  and  $T$  are the rotation and translation parameters between the reference point cloud  $P_R$  and the point cloud to be registered  $P_C$ ;  $a$  and  $b$  are the preset weight coefficients;  $D_{2D}(R, T)$  is the 2D mean weighted distance of the matched feature points, expressed as

$$D_{2D}(R, T) = \frac{1}{M} \sum_{i=1}^M m_i^R - F_p(m_i^C, R, T) \quad (9)$$

and  $D_{3D}(R, T)$  is the mean weighted 3D distance, which is expressed as

$$D_{3D}(R, T) = \frac{1}{N} \sum_{i=1}^N \omega_i [p_i^R - (R \cdot p_i^C - T)] \quad (10)$$

where  $M$  is the number of feature matched points;  $m_i^R$  and  $m_i^C$  are the  $i^{\text{th}}$  matched points in the left images of the reference stereo frame  $I_R$  and the co-registered frame  $I_C$ ;  $F_p$  is the projection function that transforms the coordinates in  $I_C$  to in  $I_R$ ;  $N$  is the pair number of the closest points between  $P_R$  and  $P_C$ ; and  $\omega_i$  is the distance weight between the corresponding 3D point  $p_i^R$  and  $p_i^C$ , which is

$$\omega_i = \frac{2}{\sigma_i^R + \sigma_i^P} \quad (11)$$

where  $\sigma_i^R$  and  $\sigma_i^P$  are the errors of points  $p_i^R$  and  $p_i^C$  obtained by Equations 6 and 7. By minimizing Equation 8, the rotation and translation parameters are solved to complete the co-registration.

### Fusion of Multiple Point Clouds

The co-registered point clouds are considered as the multiple observations of the lunar surface in different illumination conditions. The point clouds generated from the stereo images are then fused to obtain a density point cloud to better represent the lunar surface. In the data-fusion process, each point from all of the point clouds of single stereo pairs is inputted to the surface point set along with the point uncertainty. Point clustering is performed for all points one by one. As a result, each cluster contains a seed point having the lowest uncertainty and other points associated to the seed point as repeated observations. Then by merging the points in each cluster to one point using the weighted average, the final fused surface point set is obtained for enhanced topographic mapping. The workflow of multiple point-cloud fusion is shown in Figure 4.

Since the stereo matched image points are the original source of the 3D points, the estimation of position consistency is performed in the image space. All 3D points of the generated point clouds are projected to the left image space of the stereo pair which has the most matched points, to obtain the corresponding 2D coordinates. All the points are sorted by the 2D matched uncertainties in increasing order to obtain the ordered point set, which is defined as

$$\mathbf{P}_0 = \{p_i(i, l, X_{2D}, X_{3D}, \sigma_{2D}, \sigma_{3D}, C_{ID}) \mid i \in (1, m), l \in (1, n)\} \quad (12)$$

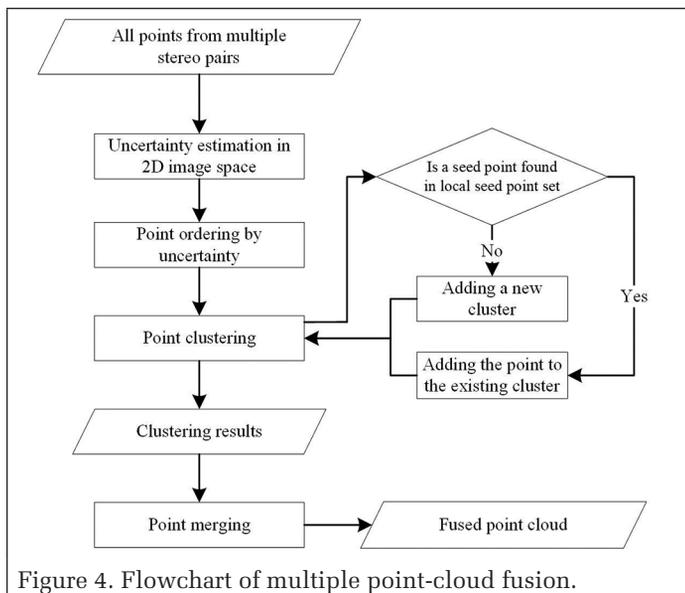


Figure 4. Flowchart of multiple point-cloud fusion.

where  $i$  is the point index in the point set;  $l$  is the point-cloud index;  $X_{2D}$  and  $X_{3D}$  are the 2D and 3D coordinates, respectively;  $\sigma_{2D}$  and  $\sigma_{3D}$  are the 2D and 3D uncertainties of the point calculated by Equations 7 and 6;  $m$  is the total number of points;  $n$  is the number of point clouds; and  $C_{ID}$  is the cluster index, which is set to  $i$  by default. Then each point  $p_i$  in  $\mathbf{P}_0$  is moved by the ascending order of the uncertainties to the initially empty surface point set  $\mathbf{P}_s$  by the uncertainty clustering. In the process of point clustering,  $p_i$ 's corresponding local seed point set from  $\mathbf{P}_s$ , which is defined as  $\mathbf{P}_N = \{p_j | j \in (1, k)\}$ , is obtained by searching in the neighboring area centered at  $X_{2D}^i$  of  $\mathbf{P}_s$  with the radius of  $\sigma_{2D}^i$ . If  $p_i$  and  $p_j$  have the minimal weighted distance  $L$ , which is calculated by

$$L = \sigma_{2D}^i + X_{2D}^i - X_{2D}^j \quad (13)$$

and  $l_i \neq l_j$ ,  $p_i$  is added to  $\mathbf{P}_s$  by setting  $C_{ID}^i$  to  $C_{ID}^j$  as a new member of the cluster containing  $p_j$ . Otherwise,  $p_i$  is used to construct a new cluster as the seed point. Due to the small search radius  $\sigma_{2D}^i$  (typically 0.2 pixel), the points in one cluster are close to each other, which ensures that the clustered points are repeated observations of the same physical point from multiple stereo pairs. After each point is clustered in  $\mathbf{P}_0$  iteratively, the optimally clustered point set  $\mathbf{P}_s$  is obtained. Then each point cluster containing more than one point is merged to an optimally estimated point  $p_r$ . The 3D coordinates  $X_{3D}^j$  of the merged point can be obtained by

$$X_{3D}^f = \sum_{h=1}^N w_h \cdot X_{3D}^h \quad (14)$$

where  $X_{3D}^h$  is the 3D coordinates of cluster point;  $N$  is the point number of the cluster; and  $w_h$  is the weight coefficient, calculated by

$$w_h = \frac{1}{\sigma_{3D}^h} / \sum_{c=1}^N \left( \frac{1}{\sigma_{3D}^c} \right) \quad (15)$$

According to measurement-error theory, the weighted average of the repeated 3D point measurements will have a higher precision than a single measurement. By processing all clusters and adding all the new points, the fused surface point set is obtained. This point cloud is optimal in that it has a higher

density and higher precision than the point cloud generated from a single stereo pair.

## Experimental Results

### Topographic Mapping Results

The experiments were performed using the two data sets already described. First, each image pair was used to generate a point cloud. Then the fused point sets of the lunar surface were obtained by applying the methods of point-cloud co-registration and fusion. Figure 5 shows the generated point clouds of the first data set and the detailed view of the point cloud corresponding to the image areas shown in Figure 2, with Figure 5d and 5e showing the fused surface point sets that have denser points than the single point cloud produced by one stereo pair of images. Figure 6 shows the reconstruction results of the second data set. The point numbers of the single stereo point clouds and the fused clouds can be found in Table 2.

From Figures 5 and 6 and Table 2, we can observe that in both data sets, the fused clouds have many more points than the point clouds generated from single stereo images, providing enhanced topographic mapping capability. Note that image pair 2d was captured with the opposite solar azimuth angle to the other three stereo pairs in data set 2, and thus the shadow areas in those pairs were imaged effectively through image pair 2d. This is particularly helpful for reconstructing the lunar surface with more complete coverage, as shown in Figure 6e.

### DEM Evaluation

To further evaluate the performance of the proposed method, we applied statistical analysis of the DEMs generated in the experiments. For convenience of comparison, two test areas were selected in each data set, at close and far range from

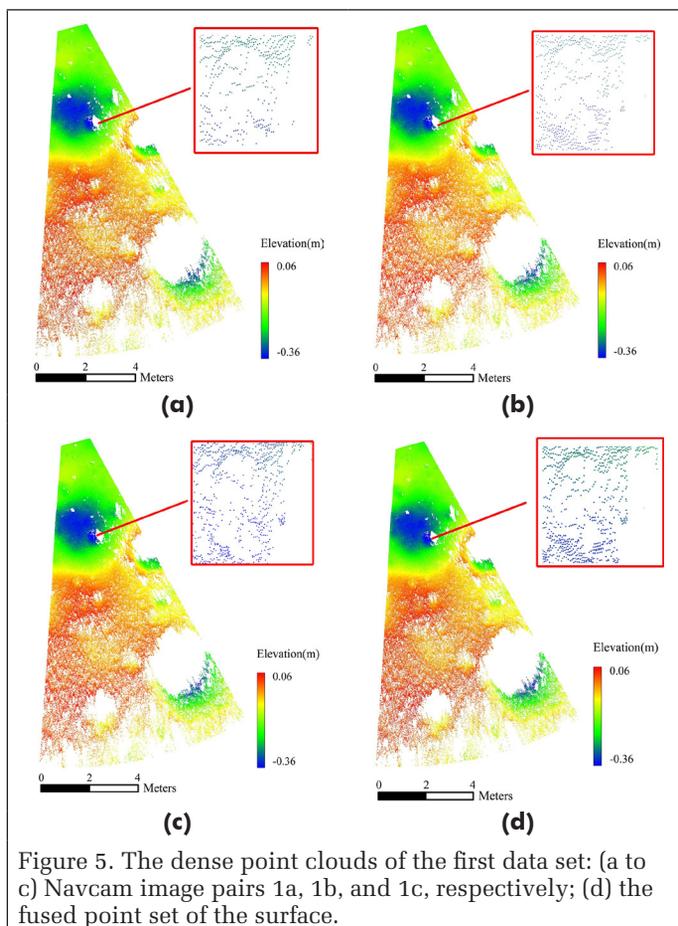


Figure 5. The dense point clouds of the first data set: (a to c) Navcam image pairs 1a, 1b, and 1c, respectively; (d) the fused point set of the surface.

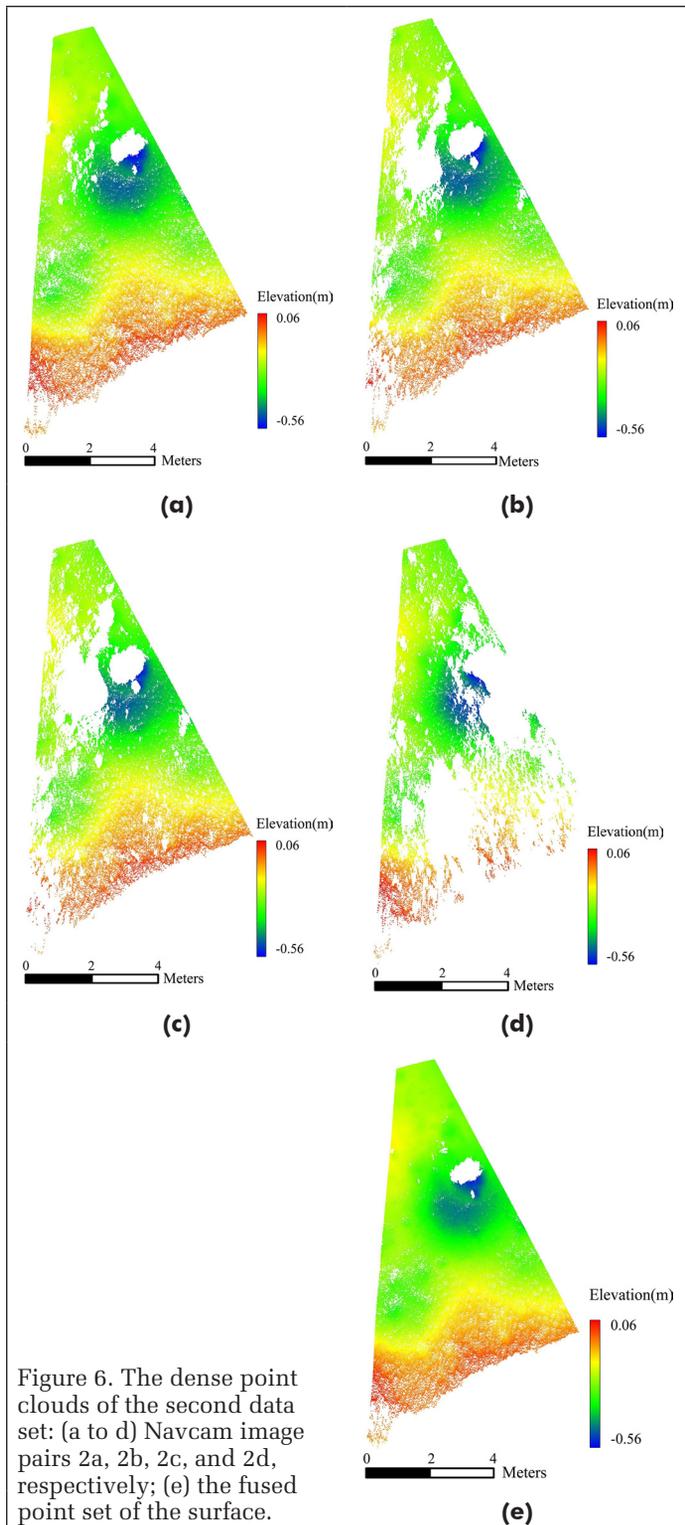


Figure 6. The dense point clouds of the second data set: (a to d) Navcam image pairs 2a, 2b, 2c, and 2d, respectively; (e) the fused point set of the surface.

Table 2. Point numbers of the point clouds.

Data Set	Point Cloud	Number of Points
1	1a	616 428
	1b	629 842
	1c	644 015
	Fused result	824 815
2	2a	634 784
	2b	401 470
	2c	331 660
	2d	266 062
	Fused result	910 187

the rover. Regular grid DEMs, interpolated with the points in selected areas, were used to perform the quality analysis using the statistical measures of sharpness and mean gradient. In evaluating the first data set, two areas A and B were selected to produce the DEMs from the fused point set. For comparative analysis, the points in the corresponding areas from the point cloud of image pair 1c—which had the most points among image pairs of the data set—were also used to generate single-pair DEMs. The two selected areas and the DEMs are shown in Figure 7. In order to preserve the details of the reconstruction results, the grid spacings of the DEMs generated in Areas A and B were set to 1 and 2 mm, respectively, to facilitate detailed comparison. To investigate the details of the topographic mapping results, we obtained the profiles of the DEMs along Profile A in Figure 7b and Profile B in Figure 7d. Figure 8 shows the red elevation profiles of the DEMs from the fused points and blue single-pair DEM profiles.

Similarly, we selected Areas C and D from the results of the second data set for DEM-based evaluation. The point cloud of image pair 2a was chosen as the data source of the single-pair DEM. Figure 9 shows the DEMs in Areas C and D. Figure 10 shows Profiles C and D. As shown in Figures 8 and 10, more details can be seen in the DEMs generated from fused point clouds.

Statistical measures of mean gradient and sharpness, Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE), and Natural Image Quality Evaluator (NIQE) were used to evaluate the produced DEMs. The sharpness depicts the mean local terrain relief, which is calculated as the distance-weighted average of height differences of a grid to its eight neighboring grids in the DEM (Wang *et al.* 2004). BRISQUE and NIQE (Mittal *et al.* 2012, 2013) were originally used to evaluate image quality; DEMs with more details and less noise will achieve better results (lower scores). As shown in Table 3, compared with the best results of a single image pair, the DEMs from fused points have better quality, demonstrating the effectiveness of the proposed method in lunar topographic mapping with higher quality and more details.

## Conclusion

In this study, a novel method for enhanced lunar topographic mapping is proposed using multiple rover stereo images acquired at one location with changing illumination conditions. The point clouds generated from each pair of stereo images were fused into a global optimized lunar surface point set. To eliminate errors in the measurements of rover mast drives, point-cloud registration was performed using the global ICP method with constraints from 2D matched features to enhance the consistency among point clouds. Each point was inputted for clustering based on the position consistency, to find repeated surface observations. The 3D coordinates were estimated by all 3D points in the cluster, with the position uncertainties. By processing all clusters of repeated observations and adding the new observations, a fused surface point cloud with larger spatial coverage (higher point density) and lower positioning uncertainty (higher precision) can be obtained.

To validate the effectiveness of the proposed method, we acquired two image data sets using the Yutu-2 rover's Navcam at waypoints LE01002 and LE00803 under changing illumination conditions. Experimental results show that the topographic mapping products (point clouds and DEMs) generated by the proposed method have higher quality than those conventionally derived from single stereo images in terms of revealing more details of the lunar surface. Furthermore, the spatial coverage of the products is improved by elimination of shadow areas with the help of changing illuminations. Without involving complex engineering operations, the proposed method provides a novel solution to enhanced

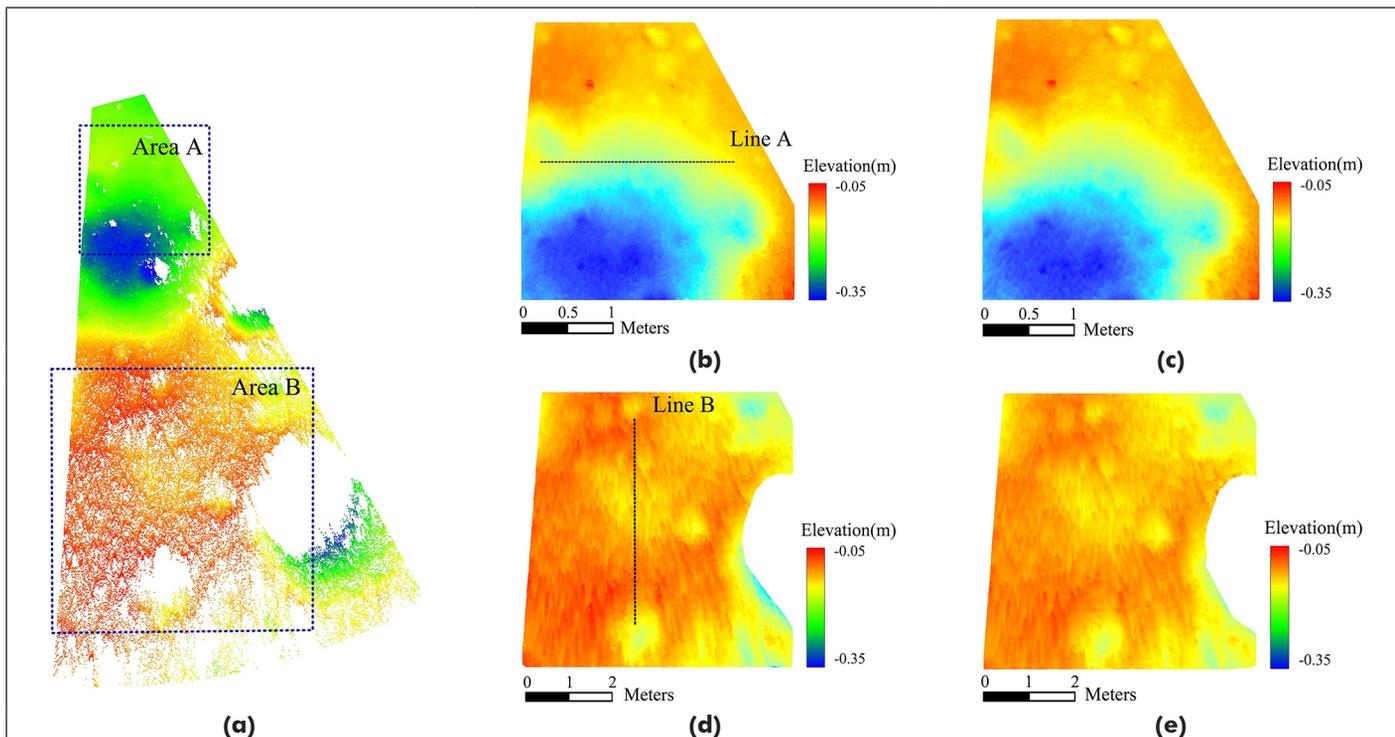


Figure 7. The selected areas and generated digital elevation models of the first data set: (a) test Areas A and B; (b, d) digital elevation models generated from the fused point sets; (c, e) digital elevation models from the point cloud of image pair 1c.

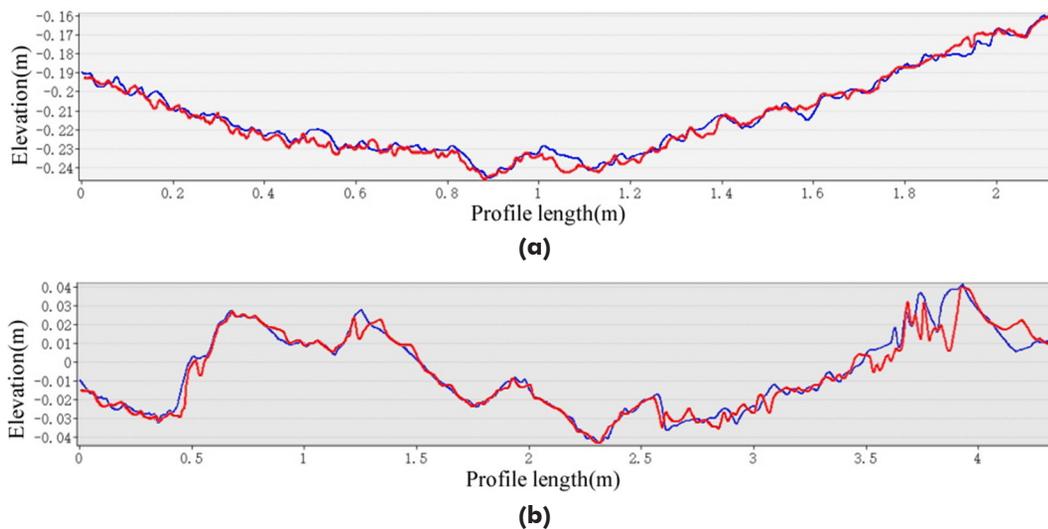


Figure 8. Elevation Profiles A (a) and B (b). Red lines are from fused digital elevation models and blue lines are from single-pair digital elevation models.

Table 3. Quality evaluation results of the digital elevation models.

Area	Source	Mean Gradient	Sharpness	BRISQUE	NIQE
A	Fused points	0.28	0.25	46.51	11.29
	Single pair	0.23	0.19	46.82	11.66
	Points of brute-force merging	0.26	0.21	47.46	11.46
B	Fused points	0.23	0.20	48.49	8.86
	Single pair	0.19	0.17	51.51	10.22
	Points of brute-force merging	0.20	0.18	49.02	8.94
C	Fused points	0.15	0.26	45.75	10.87
	Single pair	0.11	0.19	47.62	10.94
	Points of brute-force merging	0.12	0.21	47.78	10.90
D	Fused points	0.27	0.24	48.23	10.43
	Single pair	0.19	0.17	48.33	10.47
	Points of brute-force merging	0.22	0.19	48.40	10.43

BRISQUE = Blind/Referenceless Image Spatial Quality Evaluator; NIQE = Natural Image Quality Evaluator.

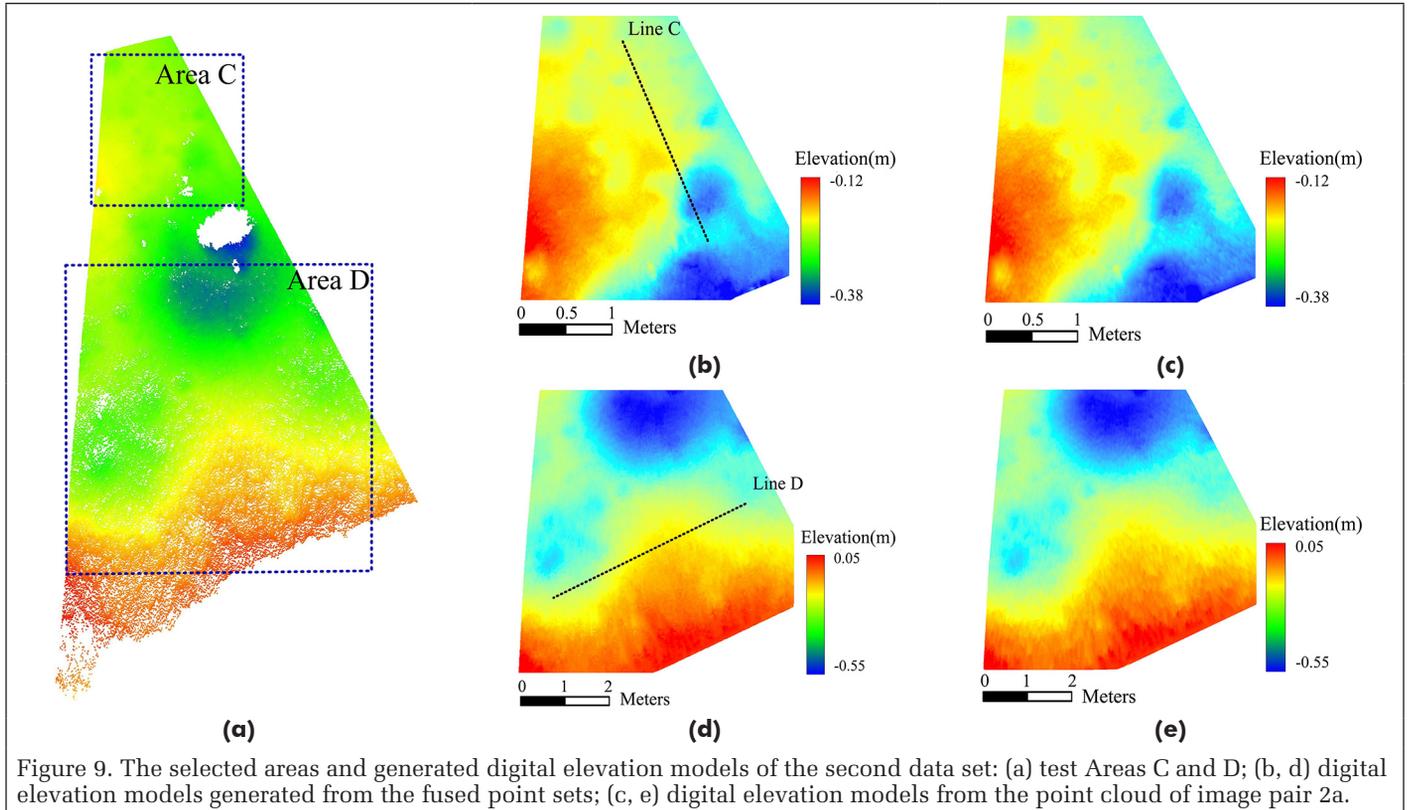


Figure 9. The selected areas and generated digital elevation models of the second data set: (a) test Areas C and D; (b, d) digital elevation models generated from the fused point sets; (c, e) digital elevation models from the point cloud of image pair 2a.

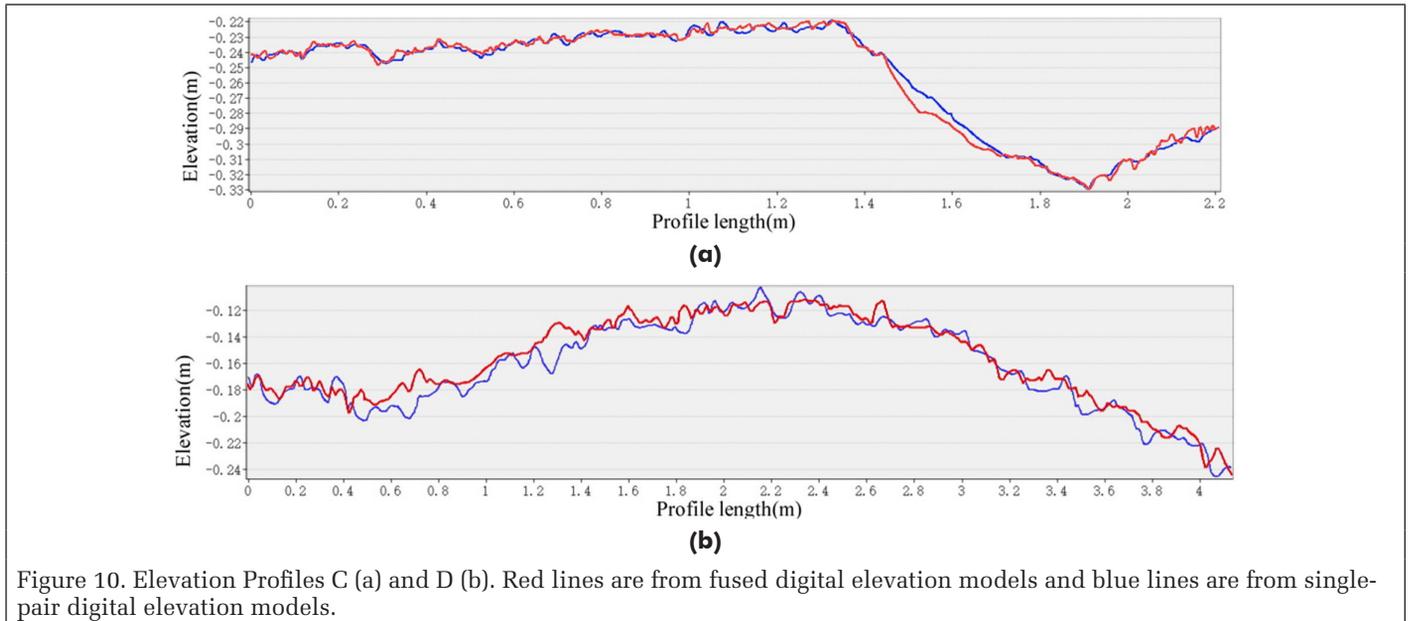


Figure 10. Elevation Profiles C (a) and D (b). Red lines are from fused digital elevation models and blue lines are from single-pair digital elevation models.

lunar topographic mapping, which has great potential to be applied in planetary rover and lander missions. In future work, shape-from-shading techniques could be combined with the proposed stereo mapping method to further enhance the topographic mapping capability using these kinds of image data sets. Furthermore, photometric properties of the surface materials could be studied based on these unique data sets with changing illumination conditions, after appropriate photometric calibration.

### Acknowledgments

The *Chang'e-4* mission was carried out by the Chinese Lunar Exploration Program. This work was supported by the National Key Research and Development Program of China

(grant 2018YFB1305004) and the National Natural Science Foundation of China (grants 41941003 and 41771488).

### References

- Alexander, D. A., R. G. Deen, P. M. Andres, P. Zamani, H. B. Mortensen, A. C. Chen, M. K. Cayan, J. R. Hall, V. S. Klochko, O. Pariser, C. L. Stanley, C. K. Thompson and G. M. Yagi. 2006. Processing of Mars Exploration Rover imagery for science and operations planning. *Journal of Geophysical Research: Planets* 111(E2):E02S02.
- Barnard, S. T. and M. A. Fischler. 1982. Computational stereo. *ACM Computing Surveys* 14(4):553–572.
- Besl, P. J. and N. D. McKay. 1992. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14(2):239–256.

- Besse, F., C. Rother, A. Fitzgibbon and J. Kautz. 2014. PMBP: PatchMatch belief propagation for correspondence field estimation. *International Journal of Computer Vision* 110(1):2–13.
- Chang, J.-R. and Y.-A. Chen. 2018. Pyramid stereo matching network. Pages 5410–5418 in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, held in City, St./Country, DD–DD Month YYYY. Edited by J. Editor. City, St./Country: Publisher.
- Chang, Y.-J. and Y.-S. Ho. 2017. Pixel-based adaptive normalized cross correlation for illumination invariant stereo matching. Pages 124–129 in *Proceedings of IS&T International Symposium on Electronic Imaging, Stereoscopic Displays and Applications, XXVIII*.
- Di, K. and R. Li. 2007. Topographic mapping capability analysis of Mars Exploration Rover 2003 mission imagery. Pages 28–31 in *Proceedings of the 5th International Symposium on Mobile Mapping Technology (MMT 2007)*.
- Di, K., Z. Liu, B. Liu, W. Wan, M. Peng, Y. Wang, S. Gou, Z. Yue, X. Xin, M. Jia and S. Niu. 2019. Chang'e-4 lander localization based on multi-source data. *Journal of Remote Sensing* 23(1):177–184.
- Di, K., Z. Liu, W. Wan, M. Peng, B. Liu, Y. Wang, S. Gou and Z. Yue. 2020. Geospatial technologies for Chang'e-3 and Chang'e-4 lunar rover missions. *Geo-spatial Information Science* 23(1):87–97.
- Di, K. and M. Peng. 2011. Wide baseline mapping for Mars rovers. *Photogrammetric Engineering and Remote Sensing* 77(6):609–618.
- Di, K., F. Xu, J. Wang, S. Agarwal, E. Brodyagina, R. Li and L. Matthies. 2008. Photogrammetric processing of rover imagery of the 2003 Mars Exploration Rover mission. *ISPRS Journal of Photogrammetry and Remote Sensing* 63(2):181–201.
- Förstner, W. and E. A. Gülch. 1987. A fast operator for detection and precise location of distinct points, corners and centres of circular features. Pages PP–PP in *Proceedings of the ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*.
- Gruen, A. W. 1985. Adaptive least squares correlation: A powerful image matching technique. *South African Journal of Photogrammetry, Remote Sensing and Cartography* 14(3):175–187.
- Harris, C. and M. Stephens. 1988. A combined corner and edge detector. Pages PP–PP in *Proceedings of the 4th Alvey Vision Conference*.
- Hirschmuller, H. 2008. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30(2):328–341.
- Li, R., S. W. Squyres, R. E. Arvidson, B. A. Archinal, J. Bell, Y. Cheng, L. Crumpler, D. J. Des Marais, K. Di, T. A. Ely, M. Golombek, E. Graat, J. Grant, J. Guinn, A. Johnson, R. Greeley, R. L. Kirk, M. Maimone, L. H. Matthies, M. Malin, T. Parker, M. Sims, L. A. Soderblom, S. Thompson, J. Wang, P. Whelley and F. Xu. 2005. Initial results of rover localization and topographic mapping for the 2003 Mars Exploration Rover mission. *Photogrammetric Engineering and Remote Sensing* 71(10):1129–1142.
- Liu, W. C. and B. Wu. 2020. An integrated photogrammetric and photoclinometric approach for illumination-invariant pixel-resolution 3D mapping of the lunar surface. *ISPRS Journal of Photogrammetry and Remote Sensing* 159:153–168.
- Liu, W. C., B. Wu and C. Wohler. 2018. Effects of illumination differences on photometric stereo shape-and-albedo-from-shading for precision lunar surface reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing* 136:58–72.
- Liu, Z. Q., K. C. Di, M. Peng, W. H. Wan, B. Liu, L. C. Li, T. Y. Yu, B. F. Wang, J. L. Zhou and H. M. Chen. 2015. High precision landing site mapping and rover localization for Chang'e-3 mission. *Science China Physics, Mechanics & Astronomy* 58(1):1–11.
- Liu, Z., K. Di, J. Li, J. Xie, X. Cui, L. Xi, W. Wan, M. Peng, B. Liu, Y. Wang, S. Gou, Z. Yue, T. Yu, L. Li, J. Wang, C. Liu, X. Xin, M. Jia, Z. Bo, J. Liu, R. Wang, S. Niu, K. Zhang, Y. You, B. Liu and J. Liu. 2020. Landing site topographic mapping and rover localization for Chang'e-4 mission. *Science China Information Sciences* 63(4):140901.
- Lowe, D. G. 2004. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2):91–110.
- Luo, W., A. G. Schwing and R. Urtasun. 2016. Efficient deep learning for stereo matching. Pages 5695–5703 in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, held in Las Vegas, Nev. 27–30 June 2016. <https://doi.org/10.1109/CVPR.2016.614>.
- Matthies, L., M. Maimone, A. Johnson, Y. Cheng, R. Willson, C. Villalpando, S. Goldberg, A. Huertas, A. Stein and A. Angelova. 2007. Computer vision on Mars. *International Journal of Computer Vision* 75(1):67–92.
- Mei, X., X. Sun, M. Zhou, S. Jiao, H. Wang and X. Zhang. 2011. On building an accurate stereo matching system on graphics hardware. Pages PP–PP in 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), held in Barcelona, Spain 6–13 November 2011. <https://doi.org/10.1109/ICCVW.2011.6130280>.
- Michael, M., J. Salmen, J. Stallkamp and M. Schlipf. 2013. Real-time stereo vision: Optimizing Semi-Global Matching. Pages 1197–1202 in 2013 IEEE Intelligent Vehicles Symposium (IV), held in Gold Coast, QLD, Australia 23–26 June 2013. <https://doi.org/10.1109/IVS.2013.6629629>.
- Mittal, A., A. K. Moorthy and A. C. Bovik. 2012. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing* 21(12):4695–4708.
- Mittal, A., R. Soundararajan and A. C. Bovik. 2013. Making a “completely blind” image quality analyzer. *IEEE Signal Processing Letters* 20(3):209–212.
- Olson, C. F., H. Abi-Rached, M. Ye and J. P. Hendrich. 2003. Wide-baseline stereo vision for Mars rovers. Pages 1302–1307 in *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, held in Las Vegas, Nev. 27–31 October 2003. <https://doi.org/10.1109/IROS.2003.1248825>.
- Olson, C. F., L. H. Matthies, J. R. Wright, R. Li and K. Di. 2007. Visual terrain mapping for Mars exploration. *Computer Vision and Image Understanding* 105(1):73–85.
- Scharstein, D. and R. Szeliski. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47(1–3):7–42.
- Stein, T. C., R. E. Arvidson and F. Zhou. 2019. PDS analyst's notebook for MSL and MER. Page PP in *Ninth International Conference on Mars*.
- Wang, H., W. Zhong, J. Wang and D. Xia. 2004. Research of measurement for digital image definition. *Journal of Image and Graphics* 9(7):581–586.
- Wu, B., Y. Zhang and Q. Zhu. 2011. A triangulation-based hierarchical image matching method for wide-baseline images. *Photogrammetric Engineering and Remote Sensing* 77(7):695–708.
- Xu, J., Q. Yang, J. Tang and Z. Feng. 2016. Linear time illumination invariant stereo matching. *International Journal of Computer Vision* 119(2):179–193.
- Zhang, Z. and J. Zhang. 1997. *Digital Photogrammetry*. Wuhan, China: Press of the Wuhan Technical University of Surveying and Mapping.
- Zhu, Q., B. Wu and Y. Tian. 2007. Propagation strategies for stereo image matching based on the dynamic triangle constraint. *ISPRS Journal of Photogrammetry and Remote Sensing* 62(4):295–308.

# Unsupervised Representation High-Resolution Remote Sensing Image Scene Classification via Contrastive Learning Convolutional Neural Network

Fengpeng Li, Jiabao Li, Wei Han, Ruyi Feng, and Lizhe Wang

## Abstract

*Inspired by the outstanding achievement of deep learning, supervised deep learning representation methods for high-spatial-resolution remote sensing image scene classification obtained state-of-the-art performance. However, supervised deep learning representation methods need a considerable amount of labeled data to capture class-specific features, limiting the application of deep learning-based methods while there are a few labeled training samples. An unsupervised deep learning representation, high-resolution remote sensing image scene classification method is proposed in this work to address this issue. The proposed method, called contrastive learning, narrows the distance between positive views: color channels belonging to the same images widens the gaps between negative view pairs consisting of color channels from different images to obtain class-specific data representations of the input data without any supervised information. The classifier uses extracted features by the convolutional neural network (CNN)-based feature extractor with labeled information of training data to set space of each category and then, using linear regression, makes predictions in the testing procedure. Comparing with existing unsupervised deep learning representation high-resolution remote sensing image scene classification methods, contrastive learning CNN achieves state-of-the-art performance on three different scale benchmark data sets: small scale RSSCN7 data set, midscale aerial image data set, and large-scale NWPU-RESISC45 data set.*

## Introduction

Due to the rapid development of remote sensing satellite technology, there are many available high-resolution remote sensing (HRRS) images acquired every day. Because of the abundant texture information and wide ground coverage, HRRS images have been widely applied in varied Earth observation tasks, such as land cover or land use classification (Milani *et al.* 2018; Cheng and Kustas 2019; Zhang *et al.* 2015; Gómez *et al.* 2016; Chen *et al.* 2017; Myint *et al.* 2011), change detection of land cover types (Liu *et al.* 2019; Saha *et al.* 2019; Lv *et al.* 2018), environment monitoring (Zhu *et al.* 2020; Song *et al.* 2019; Wang *et al.* 2019), and ground object detection or ground target tracking (Cheng *et al.* 2016; Cheng *et al.* 2014; Yu *et al.* 2018; Du *et al.* 2019; Zhang *et al.* 2019). To satisfy the requirement of practical Earth observation tasks,

understanding or analyzing HRRS images is a prerequisite. Furthermore, scene classification, which labels HRRS images based on semantic categories, is one of the fundamental research topics in understanding or analyzing HRRS images. Owing to the essential role of HRRS image scene classification playing in practical tasks, many approaches have been proposed to handle the challenging task (Li *et al.* 2017; Cheng *et al.* 2017; Lu *et al.*; 2017; Cheng *et al.* 2013; Han *et al.* 2018; Yang *et al.* 2018; Xia *et al.* 2017; Cheng *et al.* 2018; He *et al.* 2018; Wang *et al.* 2018; Liu *et al.* 2018; Yuan *et al.* 2018; Liu *et al.* 2017; Li *et al.* 2016; Zhao *et al.* 2015; Minetto *et al.* 2019). Although these methods have obtained apparent improvement, developing efficient and robust HRRS image scene classification approaches is an urgent and essential task owing to the large intraclass discrepancy and high interclass similarity (Minetto *et al.* 2019).

Current HRRS image scene classification methods are categorized into three classes: handcrafted feature-based methods, traditional unsupervised representation methods, and supervised deep learning representation methods (Cheng *et al.* 2017). Handcrafted, feature-based methods need engineering work by humans to design features, which is a time-consuming task. Moreover, a single handcrafted feature cannot adequately represent complex HRRS images, which results in loss of information (Wang *et al.* 2018). Traditional unsupervised representation methods, such as autoencoder, have poor performance on HRRS image scene classification, because the encoded representations by these methods only compress the dimension of input HRRS images and do not necessarily have strong category differentiation (Minetto *et al.* 2019). Thanks to the substantial capacity representation of deep learning, supervised deep learning representation HRRS image scene classification approaches have made excellent progress (Li *et al.* 2020).

However, traditional supervised deep learning representation methods need a massive amount of labeled information to avoid overfitting or underfitting issue (Zheng *et al.* 2019). Manually labeling data is a time-consuming task, and ImageNet (Krizhevsky *et al.* 2012), the largest data set in computer vision task, took several years to be created. For the HRRS image scene classification, the scales of data sets are limited. For example, the Northwestern Polytechnical University-Remote Sensing Image Scene Classification 45 (NWPU-RESISC45) data set is the most significant benchmark data set (Cheng *et al.* 2017) in the HRRS image scene classification task. There are

School of Computer Science and Hubei Key Laboratory of Intelligent Geo-Information Processing, China University of Geosciences, Wuhan 430078, P. R. China (lizhe.wang@foxmail.com).

Contributed by EVasit Sagan, July 23, 2020 (sent for review January 3, 2021; reviewed by Amr Abd-Elrahman, Mason Maimaitijiang).

Photogrammetric Engineering & Remote Sensing  
Vol. 87, No. 8, August 2021, pp. 577–591.  
0099-1112/21/577–591

© 2021 American Society for Photogrammetry  
and Remote Sensing  
doi: 10.14358/PERS.87.8.577

only 700 labeled images for each category, far less than the number of each class in ImageNet. The demand for massive labeled data limits supervised deep learning representation HRRS image scene classification methods. Hence, developing deep learning representation HRRS image scene classification methods with a few labeled samples or without supervised information becomes an urgent task.

To face the challenge, some unsupervised deep learning representation HRRS image scene classification approaches have been proposed in the past several years, and they can be divided into autoencoder-based methods (Romero *et al.* 2015; Tao *et al.* 2017) and generative model-based methods (Lin *et al.* 2017; Yu *et al.* 2020). Autoencoder-based methods encode HRRS images as low dimension vectors or matrices and then decode these vectors or matrices to generate HRRS images that need to be as close as possible to the original HRRS images. The vectors or matrices generated by the encoder procedure are used as representations of HRRS images. They replace authentic HRRS images and are input to a classifier to be used for classification operations. These methods only compress the dimension of data and cannot guarantee that obtained presentations by autoencoder are discriminative enough (Tian *et al.* 2019). Hence, these unsupervised representation methods for HRRS image scene classification usually do not have satisfying performance. The other typical methods are based on generative adversarial networks (GANs), which consist of a generator and a discriminator. The generator decodes samplings of random distribution as images and forces these images close to existing images, while the discriminator estimates whether the generated images are real enough comparing HRRS images used for training. GANs-based unsupervised representation methods use the encoded vectors or matrices by discriminator from HRRS images as representations to classify HRRS images. And the GANs-based unsupervised representation HRRS image

scene classification methods have better performance than previous unsupervised representation methods, such as bag-of-the-visual-words (BoVW) (Zhu *et al.* 2016), sparse coding (Olshausen and Field 1997; Sheng *et al.* 2012), and autoencoder (Saha *et al.* 2019; Masci *et al.* 2011; Fan *et al.* 2014; Bo *et al.* 2017; Cheng *et al.* 2015). However, training GANs is a difficult task, especially when there are a few training samples (Ma *et al.* 2019), which limits the application of the GAN-based approaches on small scale HRRS image data sets.

To overcome the disadvantages of existing unsupervised deep learning representation methods, contrastive learning-convolutional neural network (CL-CNN) is proposed in this study, which is motivated by the human visual system, and some existing work (Fu *et al.* 2016; Liu and Abd-Elrahman 2018). It is universally acknowledged that humans can recognize an entity from different viewpoints through view-invariant factors. However, acquiring HRRS images with different viewpoints for a scene is difficult in the scene classification task. Therefore, some alternative solutions are needed. In the proposed method, image channels that are simple to obtain are used as different viewpoints. As shown in Figure 1, this image presents a sample scenario. In this example, HRRS images with YDbDr color space (where Y is the luminance, and Db and Dr are the chrominance components) have three channels, Y channel and a combination of Db and Dr that are regarded as two views. CL-CNN aims to capture invariant factors shared by the two views, and the two views belonging to the same HRRS image measured by Euclidean distance or other metrics are mapped on nearby points, while the views from other HRRS images are mapped on far apart points. By these operations, CL-CNN can generate a unique representation for each HRRS image without any supervised information.

Our contributions to this paper can be summarized as follows:

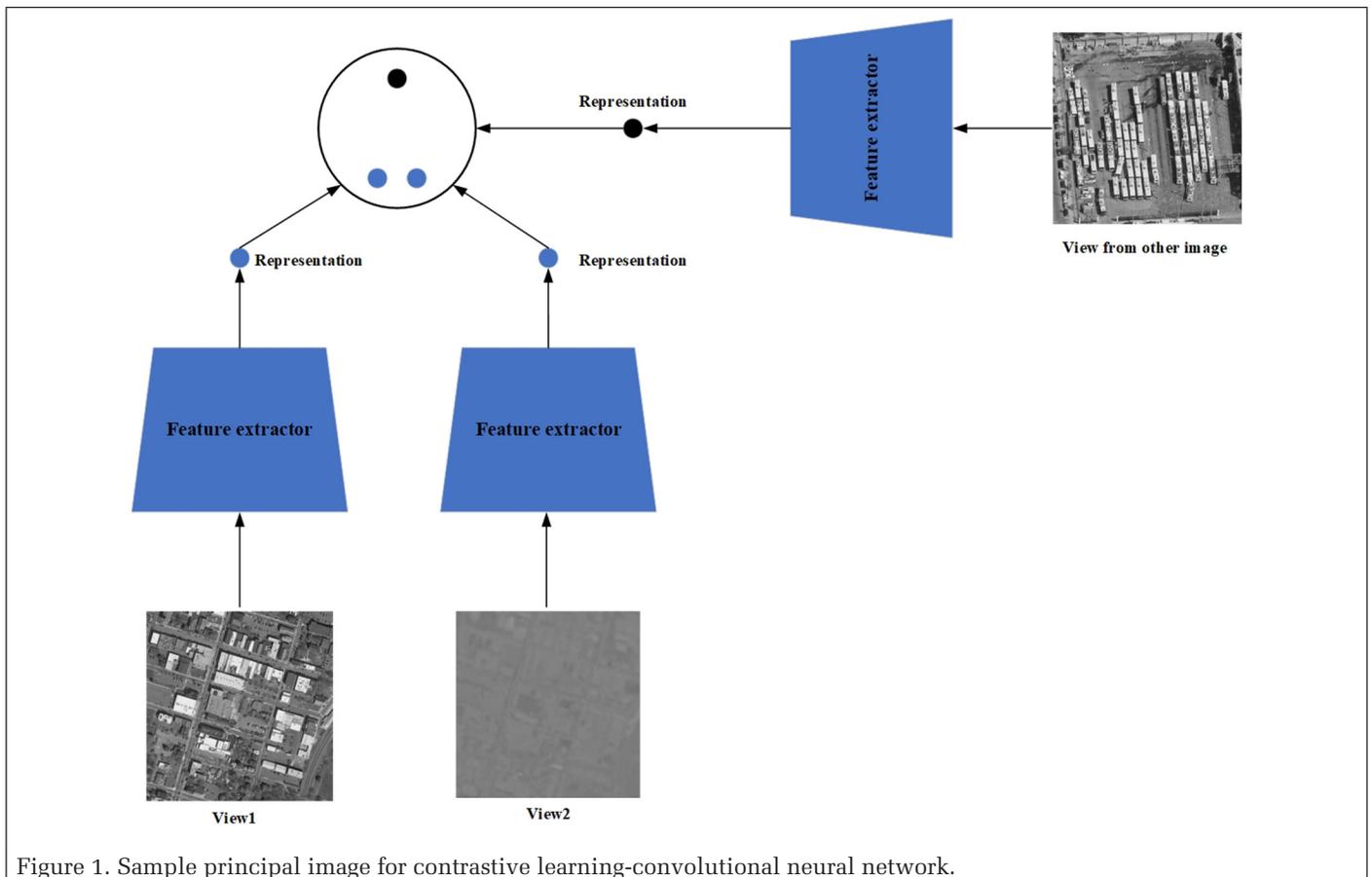


Figure 1. Sample principal image for contrastive learning-convolutional neural network.

- a. A simple unsupervised deep learning representation HRRS image scene classification method, CL-CNN is proposed. In the CL-CNN, views, which are color channels from the same HRRS images, are mapped into the same latent spaces, while views representing color channels from different HRRS images are expected to stay away from each other.
- b. The best parameter settings of CL-CNN are discussed, and the influence of color space to obtain multi-views is presented by experiments.
- c. Several groups of experiments are performed to evaluate the CL-CNN's effectiveness on HRRS image scene classification.

There are four sections in the rest of the paper. In the section "Related Work and Background", related work and some background knowledge are introduced. The section "Proposed Method" explains the details of CL-CNN, and some basic principles are also presented in this section. Some details, such as settings of parameters for training, benchmark data sets, and evaluation standards, are given. Moreover, experiment results are also presented and explained in this section. The conclusion is drawn in the last section.

## Related Work and Background

In the past several decades, the HRRS image scene classification attracts tremendous attention because of its actual role-playing in Earth observation tasks. According to the adopted types of features, existing HRRS image scene classification methods can be summarized as four types: handcrafted feature-based methods, traditional unsupervised representation methods, supervised deep learning representation methods, and unsupervised deep learning representation methods (Cheng *et al.* 2017). They are briefly reviewed in the following four subsections of this section.

### Handcrafted Feature-Based HRRS Image Scene Classification Methods

Handcrafted features, such as texture, spectral, color, and structure information are used in remote sensing image classification (Zhang *et al.* 2015), and they are expected to represent a complex image with class-specific and simple features. Hence, some broadly used handcrafted features, such as color histogram (Dala and Triggs 2005), local binary pattern (Ojala *et al.* 2002), texture descriptors (Bhagavathy and Manjunath 2006), scale-invariant feature transform (SIFT) (Lowe 2004), and histogram of oriented gradients (Dala and Triggs 2005) have been applied in HRRS image scene classification. However, adopting a single type of handcrafted feature representing a remote sensing image is impossible, and this representation operation results in loss of information (Wang *et al.* 2018).

To overcome the disadvantages of single type handcrafted feature for representation, approaches combining multiple handcrafted features are proposed (Risojević and Babić 2012; Zhu *et al.* 2016; Zou *et al.* 2016). One of them is BoVW. Zhu *et al.* (2016) introduced BoVW into HRRS image scene classification, which fuses the local dense SIFT features and global shape-based features. Although these handcrafted feature combination methods partly reduce the loss of data and improve the representation capacity of the handcrafted features, handcrafted features need to be adjusted according to the characteristics of the data, which is time-consuming and prevents the application of these methods in practical tasks.

### Traditional Unsupervised Representation HRRS Image Scene Classification Methods

To obtain features suitable for different types of HRRS images and reduce information loss, traditional unsupervised representation methods are proposed. Unsupervised features, such as principal component analysis; (Wold *et al.* 1987), spatial pyramid matching kernel (SPM; Lazebnik *et al.* 2006),

locality-constrained linear coding (LLC; Wangt *et al.* 2010), k-means clustering (Hartigan and Wong 1979), sparse coding (Olshausen and Field 1997; Sheng *et al.* 2012), and autoencoder (Saha *et al.* 2019; Masci *et al.* 2011; Fan *et al.* 2014; Bo *et al.* 2017; Cheng *et al.* 2015) aim to learn a function to transform original HRRS images into feature spaces which can distinguish HRRS images belonging to different categories without labeled information. Features extracted by these approaches are shallow features. SPM (Lazebnik2006) aims to obtain the representation of the original image by a spatial pyramid coding of a local descriptor, while BoVW tries to obtain the global features of the original image. Hence, BoVM and SPM (BoVM+SPM) are combined to obtain the representation of HRRS image to achieve better performance on scene classification task (Cheng *et al.* 2017). LLC (Wang *et al.* 2010) is a typical sparse coding-based method for HRRS image scene classification, which adopts building up local-coordinate system by coding local regions. And traditional unsupervised representation is not discriminative enough due to lack of labeled information, though unsupervised features do not need human engineering work and rely on supervised information.

### Supervised Deep Learning Representation HRRS Image Scene Classification Methods

In recent years, deep learning methods, especially CNNs, have made significant achievements. Supervised deep learning representation methods have been widely applied in remote sensing image processing tasks. Almost all of these methods use CNNs as a network backbone to extract features to represent input data (Yu *et al.* 2020). Furthermore, these models are trained with labeled data through backpropagation, which has achieved excellent HRRS image scene classification performance.

Existing deep learning can be mainly categorized into two types. At the beginning of the deep learning rising state, CNNs were regarded as approaches to extract features (Cheng *et al.* 2016), which then were handled by classifiers, such as SVM, K-nearest neighbor, and logic regression (LR). These methods need to train CNNs and classifiers separately, which increases processing time. The other types are end-to-end CNN-based models (Zhang *et al.* 2019), which train feature extractors-based on CNNs and classifiers concurrently. Moreover, these types of models are current mainstream for HRRS image scene classification. Owing to the limited numbers of labeled HRRS images, these CNNs-based methods need the fine-tuning or pretrain policy to introduce prior knowledge to avoid underfitting or overfitting issues instead of training models from scratch (Wang *et al.* 2018).

However, supervised deep learning representation methods rely heavily on the number of labeled training data. For supervised deep learning representation HRRS image scene classification, labeled data are scarce and challenging to obtain, limiting its application in the HRRS image scene classification task (Yu *et al.* 2020).

### Unsupervised Deep Learning Representation HRRS Image Scene Classification Methods

Due to the over-reliance on labeled training samples of supervised deep learning representation methods, developing a deep learning feature extractor without supervised information becomes an urgent task. In the past few years, various approaches were introduced into HRRS image scene classification to reduce the reliance on supervised information. In the early time, one of the typical unsupervised deep learning representation methods is stacked autoencoder; (Vincent *et al.* 2010), which consists of multiple layers of sparse autoencoders. This method has an apparent disadvantage in that the input data is transformed into a one-dimensional vector, which destroys spatial information of HRRS images.

CNN-based methods can retain spatial information in the feature extraction procedure. Hence, Romero *et al.* (2015) proposed a non-end-to-end CNN-based model, a greedily trained layer-by-layer. Comparing with non-end-to-end models, end-to-end models are easy to use in practical tasks. Therefore, Tao *et al.* (2017) adopted convolutional layers to replace traditional sparse autoencoder layers and introduced a CNN-based stacked autoencoder for HRRS image scene classification. Recently, motivated by the mechanism of GANs, multiple-layer feature-matching GANs (Lin *et al.* 2017) and attention GANs (Yu *et al.* 2020) were put forward successively for HRRS image scene classification. However, training GANs with limited samples is a difficult task. To overcome the disadvantages of the GANs-based methods, we proposed the CL-CNN in this study, which is easy to train and can obtain discriminative representation with limited training samples without supervised information.

## Proposed Method

To learn discriminative presentations of HRRS images by capturing information shared by different view-points without supervised information, each HRRS image's color channels need to be split and regrouped into views. For red, green, blue (RGB) color space, each channel reflects a piece of specific information and has little shared information. Therefore, other color spaces need to be selected to meet the requirements of CL-CNN. One of them is YDbDr color space. Y channel of YDbDr represents luminance, while Db, Dr are chrominance. Y and a combination of Db and Dr have much more shared information than RGB color space, which is evaluated by experiments in the section "Experiment". To facilitate the introduction of the proposed method, YDbDr is used as an example of color space in this section. Some background information about contrastive learning is introduced firstly. Contrastive learning is explained with equations and figures. Some symbols used in this section need to be described. For a collection  $C$  with  $M$  views, it can be denoted as  $C = \{V_1, V_2, \dots, V_M\}$ .  $v_i$  represents random variable representing for each view  $V_i$  in collection  $C$  following  $v_i \sim P(V_i)$ .

## Background Knowledge

To better introduce contrastive learning, a relevant topic, predictive learning, is presented in this part.  $V_1$  and  $V_2$  are supposed to be two views of a data set. In this situation,  $V_1$  represents the luminance of input HRRS images,  $V_2$  denotes chrominance of them. A deep nonlinear transformation from  $v_1$  to  $v_2$  by latent variables  $z$  is defined in this place as predictive learning. As shown in Equation 1 and Equation 2, where  $f$  represents encoder network,  $g$  denotes decoder network, and  $\hat{v}_2$  is prediction for  $v_2$ , these two equations express the procedure of the defined predictive learning. The parameters of model  $f$  and model  $g$  are trained by using an objective function to make  $\hat{v}_2$  similar to  $v_2$  as much as possible. Frequently-used objective functions are L1 and L2 norm. It is noticed that there is an assumption that each pixel in  $v_1$  and  $v_2$  is independent from each other (Higgins *et al.* 2017), which results in poor performance of represent complex images (Locatello *et al.* 2019). Predictive learning has been used in some representation methods, such as variational autoencoder (Kingma and Welling *et al.* 2014) and colorization (Cheng *et al.* 2015; Larsson *et al.* 2016; Poterek *et al.* 2020).

$$z = f(v_1) \quad (1)$$

$$\hat{v}_2 = g(z) \quad (2)$$

Different from predictive learning, in contrastive learning, as shown in Equation 3 and Equation 4, two view  $v_1$  and  $v_2$  of HRRS images need to be mapped into latent space as  $z_1$  and  $z_2$ . Furthermore, the latent representation of the two views needs

to be as close as possible to each other. Although contrastive learning and predictive learning aim to find the relevance of the two views, the former only focuses on the relation of two views in a high dimensional space. At the same time, the latter pays attention to how to transform  $v_1$  to  $v_2$ . Moreover, there is no requirement that each factor in  $z_1$  and  $z_2$  is independent of each other. And  $z_1$  and  $z_2$  need to be closed to or far from each other, which is the essential difference between contrastive learning and predictive learning.

$$z_1 = f_1(v_1) \quad (3)$$

$$z_2 = f_2(v_2) \quad (4)$$

## Contrastive Learning

The motivation of contrastive learning aims to learn discriminative embeddings that can distinguish each HRRS image. There are two views,  $V_1$  and  $V_2$ , for an HRRS image data set, which consists of a collection of samples  $\{v_1^i, v_2^i\}_{i=1}^N$ , where  $N$  represents the number of images in the data set. The two views are regrouped in the training procedure. For views from joint distribution  $x \sim (v_1, v_2)$  or  $x = (v_1^i, v_2^j)$ , which represents views from the same HRRS images, they are named positive views, while negative views groups are selected from product of marginals  $y \sim p(v_1)p(v_2)$  or  $y = (v_1^i, v_2^j)$ , where  $i$  is not equal to  $j$ .

A feature extractor,  $h_\theta$  consisting of two branches CNNs, is trained to make positive view pairs obtain a high value and make a low value for negative view pairs. Negative pairs and positive pairs are collected in a set  $S = \{x, y_1, y_2, \dots, y_k\}$ ,  $k$  presents the number of negative pairs for a positive pair  $x$ . And  $k$  needs to be larger than one and smaller than  $N$ . HRRS images are transformed into YDbDr color space in the training procedure, where Y channel is  $V_1$ , and Db and Dr channels compose  $V_2$ , and then establishing  $S$  for transformed HRRS images are invested into  $h_\theta$  for training. The objective function for training can be represented by Equation 5:

$$I_{\text{contrastive}} = -E_S \left\{ \log \frac{h_\theta(x)}{h_\theta(x) + \sum_{i=1}^k h_\theta(y_i)} \right\} \quad (5)$$

According to the definition of  $S$ , the objective function can be rewritten as Equation 6.

$$I_{\text{contrastive}} = -E_{\{x, y_1, y_2, \dots, y_k\}} \left\{ \log \frac{h_\theta(\{v_1^i, v_2^j\})}{\sum_{j=1}^{k+1} h_\theta(v_1^i, v_2^j)} \right\} \quad (6)$$

As described in the paper,  $h_\theta$  is a feature extractor consisting of two branches,  $h_{\theta_1}$  and  $h_{\theta_2}$ . Input channels for HRRS images for  $h_{\theta_1}$  and  $h_{\theta_2}$  are different, where  $\theta_1$  and  $\theta_2$  represent parameters for the two branches. For example, when original HRRS image transformed into YDbDr color space,  $V_1$  has a single channel for each input image, and there are two channels for  $V_2$ .

The latent representations for the two branches can be obtained by  $z_1 = h_{\theta_1}(v_1)$  and  $z_2 = h_{\theta_2}(v_2)$ . In the training procedure, the score value between the two branches is computed by cosine distance, as displayed in Equation 7.

$$h_\theta = \exp \left( \frac{h_{\theta_1}(v_1) \cdot h_{\theta_2}(v_2)}{\|h_{\theta_1}(v_1)\| \|h_{\theta_2}(v_2)\|} \right) \quad (7)$$

In the Equation 6, only one case that  $V_1$  selected as positive views and  $V_2$  representing negative views is taken into consideration. In the practical situation,  $V_1$  can be negative views from other HRRS images, and  $V_2$  are positive views. In this situation, negative view pairs can be denoted as  $y \sim p(v_1)p(v_2)$

or  $y=(v_1^i, v_2^j)$ , where  $i$  is not equal to  $j$ . For Equation 6, it can be rewritten as Equation 8.

$$J_{\text{contrastive}}^{V_1, V_2} = -E_{\{x, y_1, y_2, \dots, y_k\}} \left\{ \log \frac{h_{\theta}(\{v_1^i, v_2^j\})}{\sum_{j=1}^{k+1} h_{\theta}(v_1^i, v_2^j)} \right\} \quad (8)$$

Hence, these two cases need to be added to compose contrastive learning. And the whole loss function for the CL-CNN can be written as Equation 9.

$$J_{\text{contrastive}} = J_{\text{contrastive}}^{V_1, V_2} + J_{\text{contrastive}}^{V_2, V_1} \quad (9)$$

In the above description, only one situation with two views is considered. Expanding the method to more than two views is feasible. However, obtaining more than two views for HRRS images is quite tricky. In this section, YDbDr is introduced as an example. Other YUV color spaces also can be used to obtain multi-views of HRRS images. As shown in Figure 2, the Y channel of YDbDr or YUV possesses great texture and spatial information. In contrast, the Db and Dr or U and V channels only reflect chrominance, and each channel of them has limited information. Considering the practical situation of HRRS images and the characteristics of color spaces, only the case with two views are discussed. Furthermore, the effectiveness of different  $k$  and color space should be presented by experiments on HRRS image data sets.

### Model Settings

To better present the performance of CL-CNN, a CNN model needs to be selected as a feature extractor. In the past years, there are a great deal of CNN models proposed, such as AlexNet (Krizhevsky *et al.* 2012), VGGNet (Simonyan and Zisserman 2015), GoogleNet (Szegedy *et al.* 2015), and ResNet (He *et al.* 2016). Considering the excellent performance of ResNet on supervised deep learning HRRS image scene classification (Minetto *et al.* 2019), ResNet-50 is used in CL-CNN as a feature extractor. Its structure is displayed in Figure 3. Each branch of  $h_{\theta}$  is a completed ResNet-50. Furthermore, the number of input channels of a branch1 is 1, while the number of input channels for the other branch is 2. As shown in Figure 4, it shows the detailed structure of CL-CNN. For an HRRS image transformed into YDbDr color space  $X=\{\text{Channel}_Y, \text{Channel}_{Db}, \text{Channel}_{Dr}\}$  with a size  $224 \times 224$ , the feature extraction procedure can be present in Equations 10 and 11.

$$z_1 = f_{\theta_1}(\text{Channel}_Y) \quad \text{Channel}_Y \in 1 \times 224 \times 224, z_1 \in 128 \times 1 \times 1 \quad (10)$$

$$z_2 = f_{\theta_2}(\text{Channel}_{Db}, \text{Channel}_{Dr}) \quad \text{Channel}_{Db}, \text{Channel}_{Dr} \in 1 \times 224 \times 224, z_2 \in 128 \times 1 \times 1 \quad (11)$$

To make the training procedure efficient, a memory bank motivated by (He *et al.* 2019) is applied in CL-CNN. In the training procedure, all latent representation for HRRS images is saved in the memory bank. An opposing pair set  $S$  can be composed by selecting latent representation from the memory bank. The mechanism can weaken the negative impact caused by limited graphics processing unit (GPU) computing storage.

### Experiment

At the beginning of this section, some evaluation metrics for HRRS image scene classification are introduced. Then, some parameter settings for in training procedure for CL-CNN are explained in detail. Three benchmark data sets for HRRS image scene classification used in this paper are presented to display experiments better. In the last part of the section, we display all experimental results and explain them to avoid misunderstanding.

#### Evaluation Metrics

Confusion matrix (CM), overall accuracy (OA), and average accuracy (AA) are frequently used in HRRS image scene classification approaches evaluation as follows:

- Confusion matrix (CM): CM has been widely applied in supervised learning evaluation. The row labels represent the true categories for input data in the CM matrix, while the column labels are predictions. Moreover, CM can be shown visually, which makes it quite useful for analyzing experimental results.
- Overall accuracy (OA): OA is the proportion of correctly predicted data for all data, which is obtained by the number of correctly predict images dividing the total number of testing images.
- Average accuracy (AA): AA is the average prediction accuracy for each category. If the number of testing images for each category is the same, AA is equal to OA.

#### Parameter Settings

We perform all experiments with a code based on Pytorch. This paper's experiments are carried out on the Tianhe-2 V100 GPU distributed system, where there are four V100 GPUs with 16 GB random access memory (RAM) on each computation node. The learning rate is set as  $lr_1 = 0.1$  with the stochastic gradient descent optimizer for the feature extractor. We adopted a variable learning rate for the training feature extractor, which is updated, followed as  $lr_1 = lr_1 * 0.9$  for each 100 training steps. Moreover, the LR classifier training procedure uses a fixed learning rate at  $lr_2 = 0.1$ . We set the batch size to 128 for images with a size of  $224 \times 224$  with 3000 training epochs for training unsupervised feature extractor and 40 training echoes for LR classifier. The number of negative views,  $k$  is discussed by experiments on RSSCN7 to find the parameter settings rules. To obtain the best transformation color space, we compare CL-CNN performance with different color spaces on RSSCN7

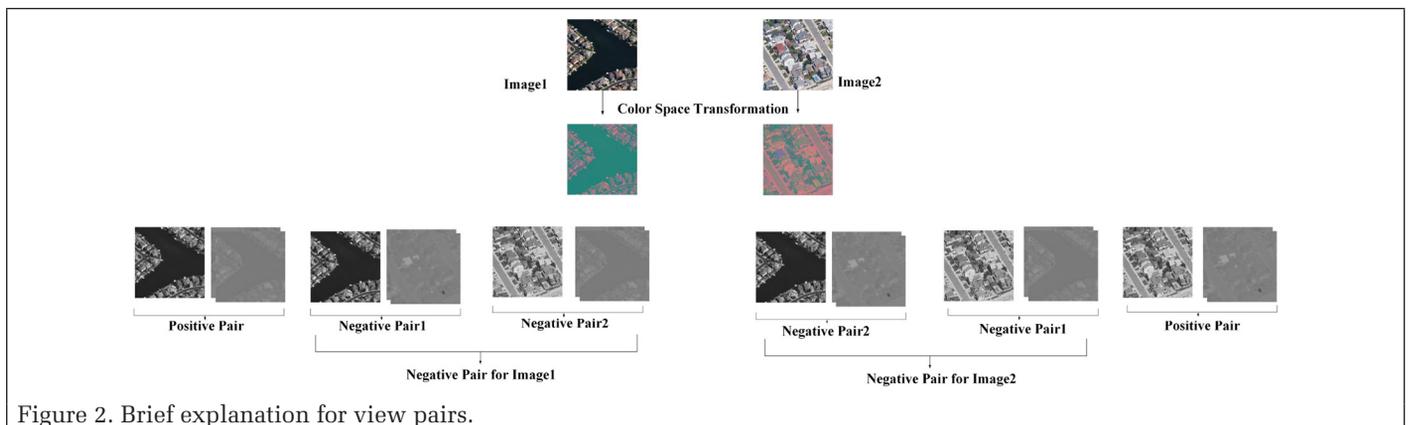
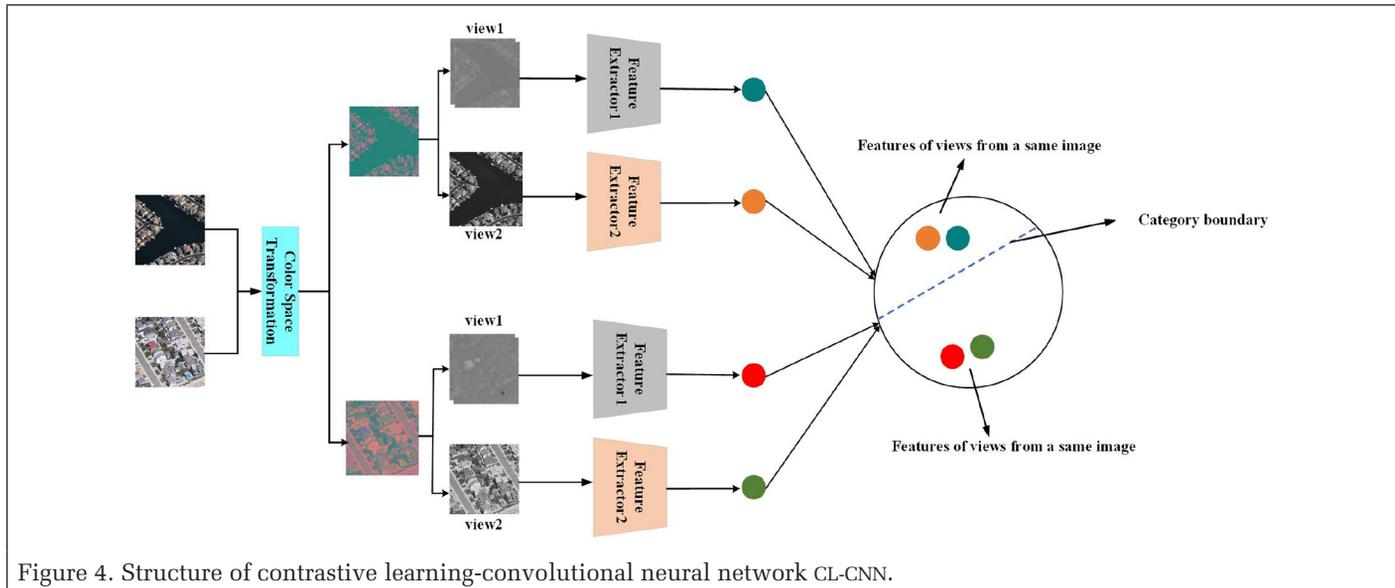
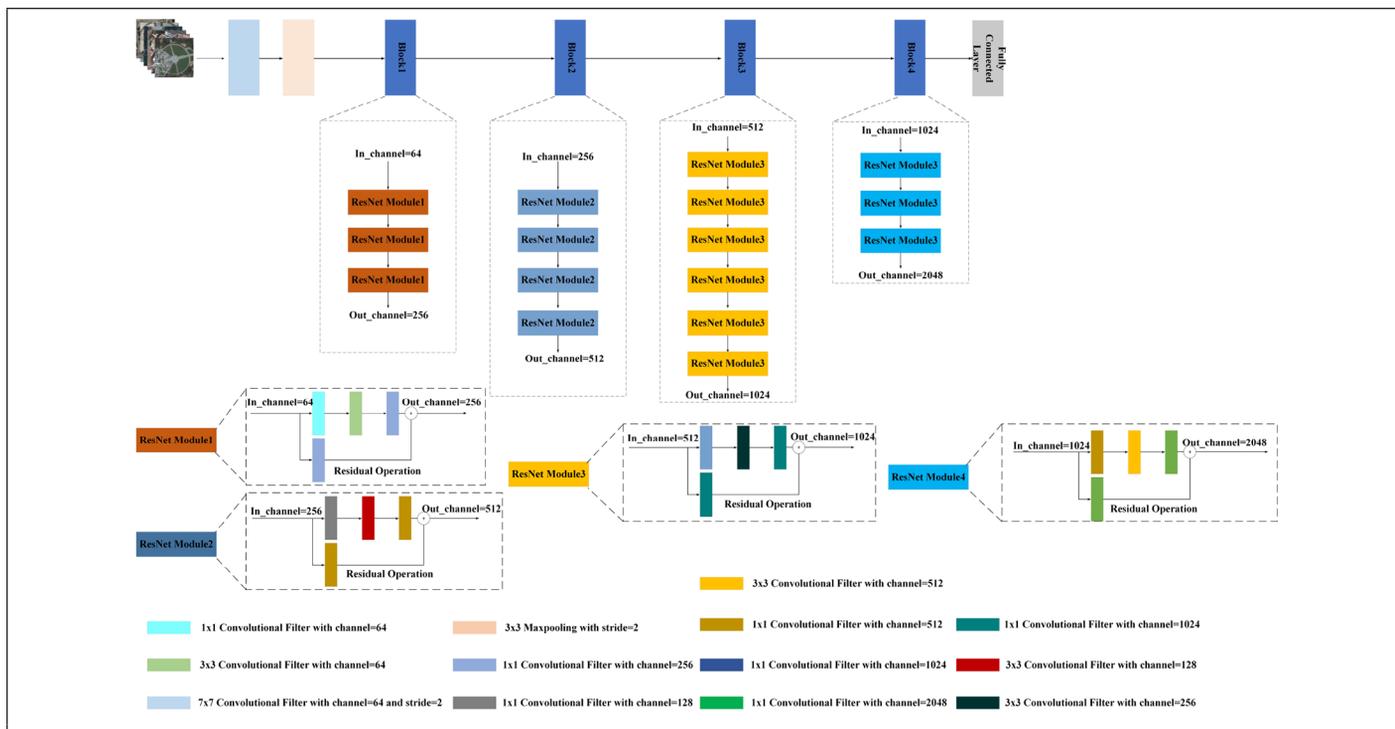


Figure 2. Brief explanation for view pairs.



with 20% data for training. To compare with current existing unsupervised representation HRRS image scene classification methods, we use the same training ratios with compared methods for each used data set in our experiment.

**Brief Introduction of Data Sets**

To evaluate the effectiveness and robustness of CL-CNN, we use the RSSCN7, the aerial image data set, and the NWPU-RESISC45 (Cheng *et al.* 2017) data set to perform experiments. The three used benchmark data sets in this paper have different data scales. RSSCN7 is a small data set, where there are only seven categories. There are 30 classes in aerial image data set, it is a middle-level data set, and the number of each category of this data set is unbalanced, making it challenging. Furthermore, NWPU-RESISC45 has 45 categories, which is the most significant data set for HRRS image scene classification. By the three data sets, CL-CNN can be evaluated for the ability to handle different scale data sets. Moreover, super parameter settings

obtained over RSSCN7 are applied in the other two data sets to present the robustness of CL-CNN.

1. **RSSCN7 Data Set (RSSCN7):** RSSCN7 data set is a small scale HRRS image scene classification (Zou *et al.* 2015). There are only seven categories in the data set, and each category has 400 images with a size of 400×400 pixels. All images of this data set are selected from Google Earth, which has been pretreated. To compare with existing unsupervised deep learning representation methods for HRRS image scene classification, we perform two group experiments with 20% and 50% samples for training. The rest of the data in this data set is used for testing on (Yu *et al.* 2020).
2. **Aerial Image Data Set (AID):** Aerial image data set (AID) (Xia *et al.* 2017) is a medium scale land-use data set for HRRS image scene classification where there are 30 categories. The AID data set has an unbalanced category of data, which is varied from 220 to 420 images. There are a total of 10 000

images with a size of 600×600. Furthermore, all of them are labeled by remote sensing image interpretation specialists. We perform two group experiments with 20% and 50% data for training separately, and the rest of the images are test data, according to the existing models in (Yu *et al.* 2020).

3. **NWPU-RESIS45 Data Set (NWPU-45):** NWPU-RESIS45 (Cheng *et al.* 2017) is a challenging data set for HRRS scene classification. There are 45 categories in the data set, where each category has 700 images with a size of 256×256. The spatial resolution of these images is range from 0.2 m to 0.3 m. According to Minetto *et al.* (2019), the NWPU-45 data set is the largest data set for HRRS image scene classification. In the data set, images belonging to the same category usually have a large discrepancy. In contrast, some images categorized as different labels are quite similar to each other, which means achieving satisfying performance on the data set is extremely difficult. According to Yu *et al.* (2020), we perform two groups of experiments, which use 10% and 20% images for each category as training data and the performance of models are tested on the rest of the data.

Data partition ways of training data sets and testing data sets have a significant impact on testing accuracy. To avoid the occasionality of experimental results and display the robustness of CL-CNN, we calculate the mean and standard deviation for each group of experiments performed by CL-CNN. We also use horizontal flipping, vertical flipping to augment the training data. It is noticed that the training feature extractor of CL-CNN does not need any labeled information. In the testing procedure, labeled information of training data is only used to train the LR classifier. The parameters of the feature extractor are frozen and cannot be updated.

### Experiment Results and Explanation

To better present the improvement and remarkable performance of CL-CNN, we use three parts to display all experiment results. In the first part, we discuss the influence of numbers of negative pairs and different color spaces on experiment results. The second part shows the comparisons between existing unsupervised representation HRRS image scene classification methods and explains experiment results in detail. To display CL-CNN improvement, we also give comparisons between CL-CNN with supervised deep learning representation HRRS image scene classification methods.

*Discussion About Influence of Different k Values and Different Color Spaces*  
To find the influence of different  $k$  for CL-CNN, we perform seven groups of experiments with  $k = 1, 200, 400, 600, 800, 1000$  over RSSCN7 data set with YDbDr color space. In these experiments, 20% of data are used as training samples, and  $N$  is 560 in this situation. Although  $k$  needs to be smaller than  $N$  and larger than 0, we want to present results if  $k$  is out of range. As shown in Figure 5, with an increase of  $k$  value, CL-CNN needs more training epochs to make its parameter convergence. Furthermore, when  $k$  is out of range, CL-CNN can still be trained. As shown in Table 1, there is a positive correlation between  $k$  and testing accuracy, while  $k$  is larger than 0 and smaller than  $N$ . Moreover, when  $k$  is larger than  $N$ , the growth of testing accuracy is stagnant, which means that the time-consuming training CL-CNN with a larger  $k$  does not have any return. Hence, when  $k$  is equal to  $N - 1$ , CL-CNN spends less time obtaining the best performance based on these experiments. In the following experiments, we set  $k$  as  $N - 1$ .

In the paper's previous description, we use YDbDr color space as an example to introduce CL-CNN. The influence of different color spaces used to obtain multi-views needs to be discussed. Therefore, we perform experiments over RSSCN7 with 20% data for training and calculate the average performance of CL-CNN with different color spaces. We use six kinds of color spaces in the experiment of discussing the influence of different color spaces for representation results, and they are

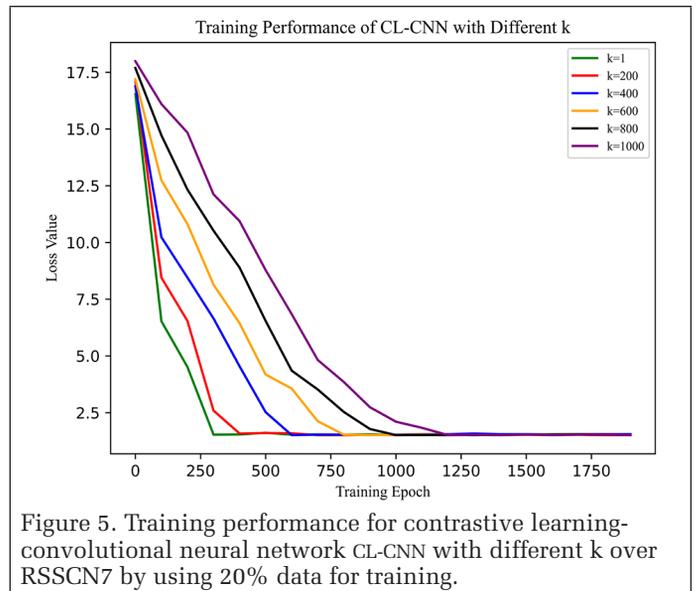


Figure 5. Training performance for contrastive learning-convolutional neural network CL-CNN with different  $k$  over RSSCN7 by using 20% data for training.

Table 1. Testing accuracy of contrastive learning-convolutional neural network over RSSCN7 with different values of  $k$ .

$k$ Value	OA (20%)
1	68.09 ± 0.66
200	73.28 ± 0.15
400	82.77 ± 0.54
559	83.02 ± 0.82
600	83.02 ± 0.82
800	83.02 ± 0.82
1000	83.02 ± 0.82

OA = overall accuracy.

Table 2. Testing accuracy of contrastive learning-convolutional neural network over RSSCN7 with different color spaces.

Color Space	OA (20%)
RGB	68.09 ± 0.66
YUV	73.28 ± 0.15
Lab	82.77 ± 0.54
YPbPr	83.02 ± 0.82
YDbDr	83.02 ± 0.82
YCbCr	83.02 ± 0.82

OA = overall accuracy;  
RGB = red, green, blue.

RGB color space, YUV color space, lab color space, YDbDr color space, YPbPr color space, and YCbCr color space. For RGB color space, R channels of the input images are used as a view, and the combination of the rest two channels is the other view. Moreover, experiments about the rest of the color spaces use two views: The Y channel and the remaining two channels, respectively. As shown in Table 2, original HRRS images achieve the worst performance, which means invariant-view information between different channels of RGB color space is low. For the other color spaces, HRRS images with YCbCr color space obtains the best average performance. Actually, except RGB color space, each channel of the other color spaces has an abundant amount of the same feature information with each other, while each channel of RGB color space is entirely independent. That is why CL-CNN with RGB color space has the worst performance on the RSCCN7 data set. Although YCbCr color space can be only the best selection for CL-CNN over the RSCCN7 data set, we apply this rule in experiments to present CL-CNN other's robustness data sets to compare with existing methods.

### Comparisons of Unsupervised Deep Learning Representation HRRS Image Scene Classification Methods over Selected Benchmark Data Sets

1. The RSCCN7 is a small-scale data set. Moreover, training a supervised or unsupervised deep learning representation method with limited data is a challenging task. To compare with existing methods, we use 20% and 50% data of RSCCN7 to train CL-CNN respectively and evaluate the performance by LR classifier. As shown in Table 3, CL-CNN obtains 2.52% and 3.79% improvement in the two groups of experiments over RSCCN7 comparing with the state-of-the-art

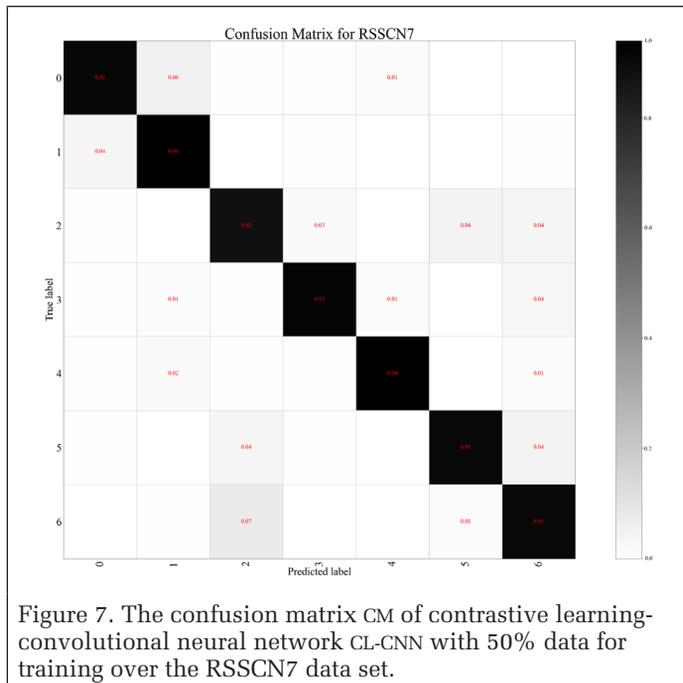
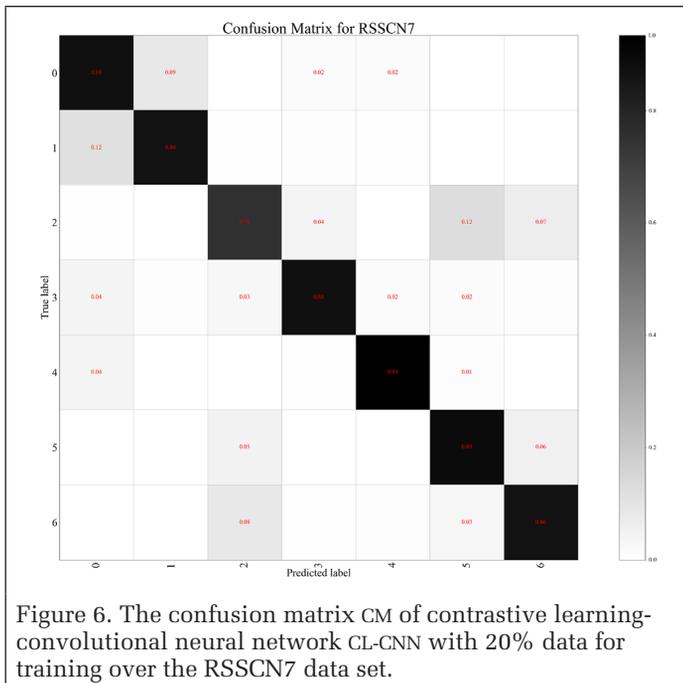


Figure 6. The confusion matrix CM of contrastive learning-convolutional neural network CL-CNN with 20% data for training over the RSSCN7 data set.

Figure 7. The confusion matrix CM of contrastive learning-convolutional neural network CL-CNN with 50% data for training over the RSSCN7 data set.

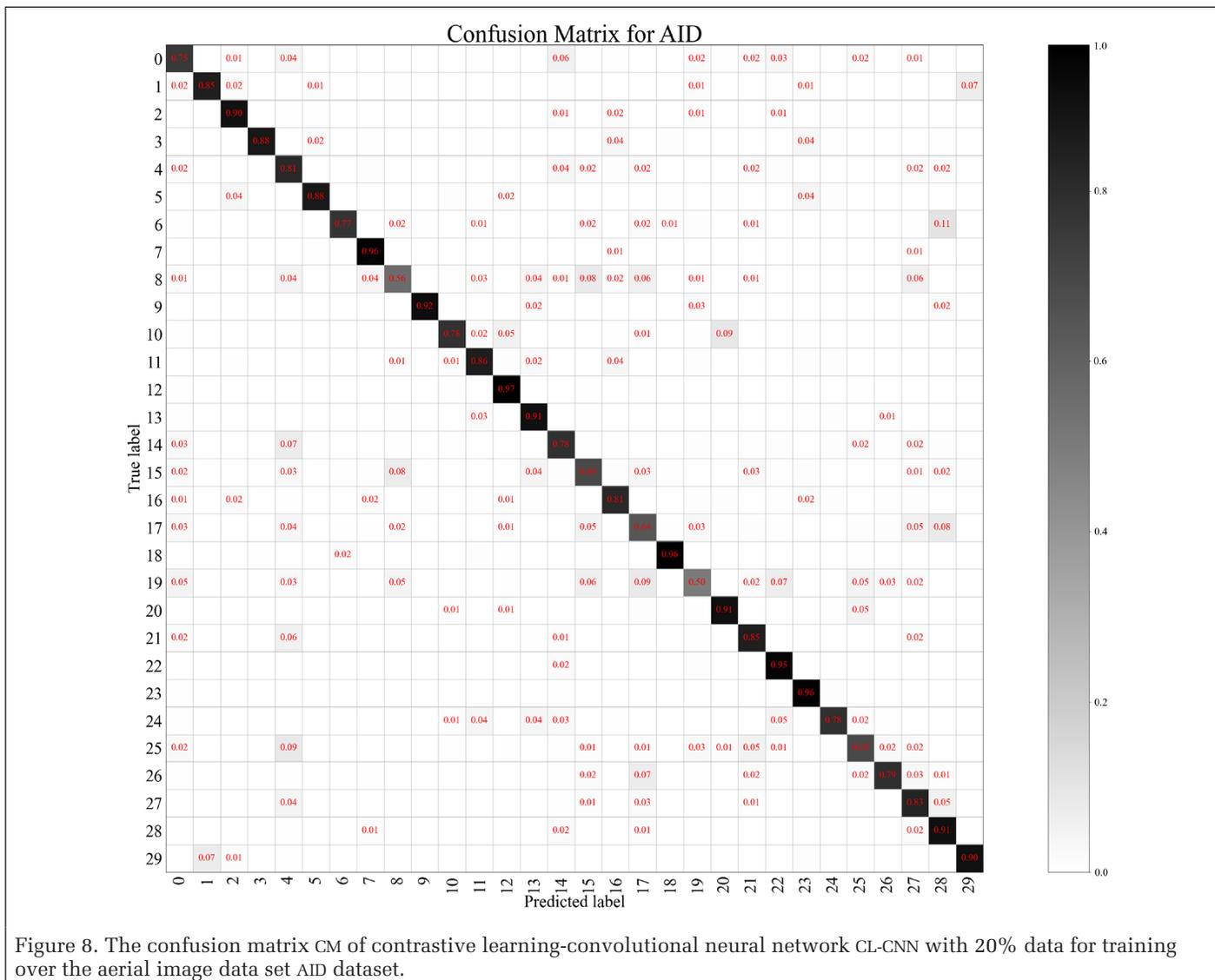


Figure 8. The confusion matrix CM of contrastive learning-convolutional neural network CL-CNN with 20% data for training over the aerial image data set AID dataset.

unsupervised representation method for HRRS image scene classification. Furthermore, in Figures 6 and 7, where the numbers in the axis from 0 to 6 represent grass, field, industry, river lake, forest, resident, and parking, we present the classification results of CL-CNN over the RSSCN7 data set by visualizing CMs. The performance over the RSSCN7 indicates that CL-CNN can generate class-specific representation, even though there are a few training samples.

- The AID is a medium scale data set for HRRS image scene classification, and AID has unbalanced samples for each category, which increases the difficulty of unsupervised representation. To compare with existing unsupervised representation methods and present the excellent ability of unsupervised representation in unbalanced categories, we use 20% and 50% images of the AID data set as training data to perform two group experiments. As shown in Table 4, which displays performance comparisons over the AID data set between CL-CNN and existing unsupervised representation methods, CL-CNN obtains 2.94% and 4.53% improvement of classification accuracy comparing with attention GANs in these two experiments. The improvement achieved by CL-CNN shows the ability of the proposed method to capture discriminative features with samples from imbalanced classes.

Figures 8 and 9, where numbers from 0 to 29 are corresponding to the airport, bare land, farmland, forest,

Table 3. The performance comparison of models on the RSSCN7 data set.

Model	OA (20%)	OA (50%)
LLC (Xia <i>et al.</i> 2017)	73.29 ± 0.63	77.11 ± 1.29
BoVW (Xia <i>et al.</i> 2017)	76.91 ± 0.59	81.28 ± 1.19
MARTA GANs (Lin <i>et al.</i> 2017)	79.86 ± 0.93	84.49 ± 0.62
Attention GANs (Yu <i>et al.</i> 2020)	83.47 ± 0.63	87.32 ± 0.54
CL-CNN	85.99 ± 0.47	91.11 ± 0.18

OA = overall accuracy; LLC = locality-constrained linear coding; BoVW = bag-of-visual-words; GAN = generative adversarial networks; CL-CNN = contrastive learning-convolutional neural network.

Table 4. The performance comparison of models on the aerial image data set.

Model	OA (20%)	OA (50%)
LLC (Xia <i>et al.</i> 2017)	58.06 ± 0.50	63.24 ± 0.44
BoVW (Xia <i>et al.</i> 2017)	62.49 ± 0.53	68.37 ± 0.40
MARTA GANs (Lin <i>et al.</i> 2017)	75.39 ± 0.49	81.57 ± 0.33
Attention GANs (Yu <i>et al.</i> 2020)	78.95 ± 0.49	84.52 ± 0.18
CL-CNN	81.89 ± 0.74	89.05 ± 0.23

OA = overall accuracy; LLC = locality-constrained linear coding; BoVW = bag-of-visual-words; GAN = generative adversarial networks; CL-CNN = contrastive learning-convolutional neural network.

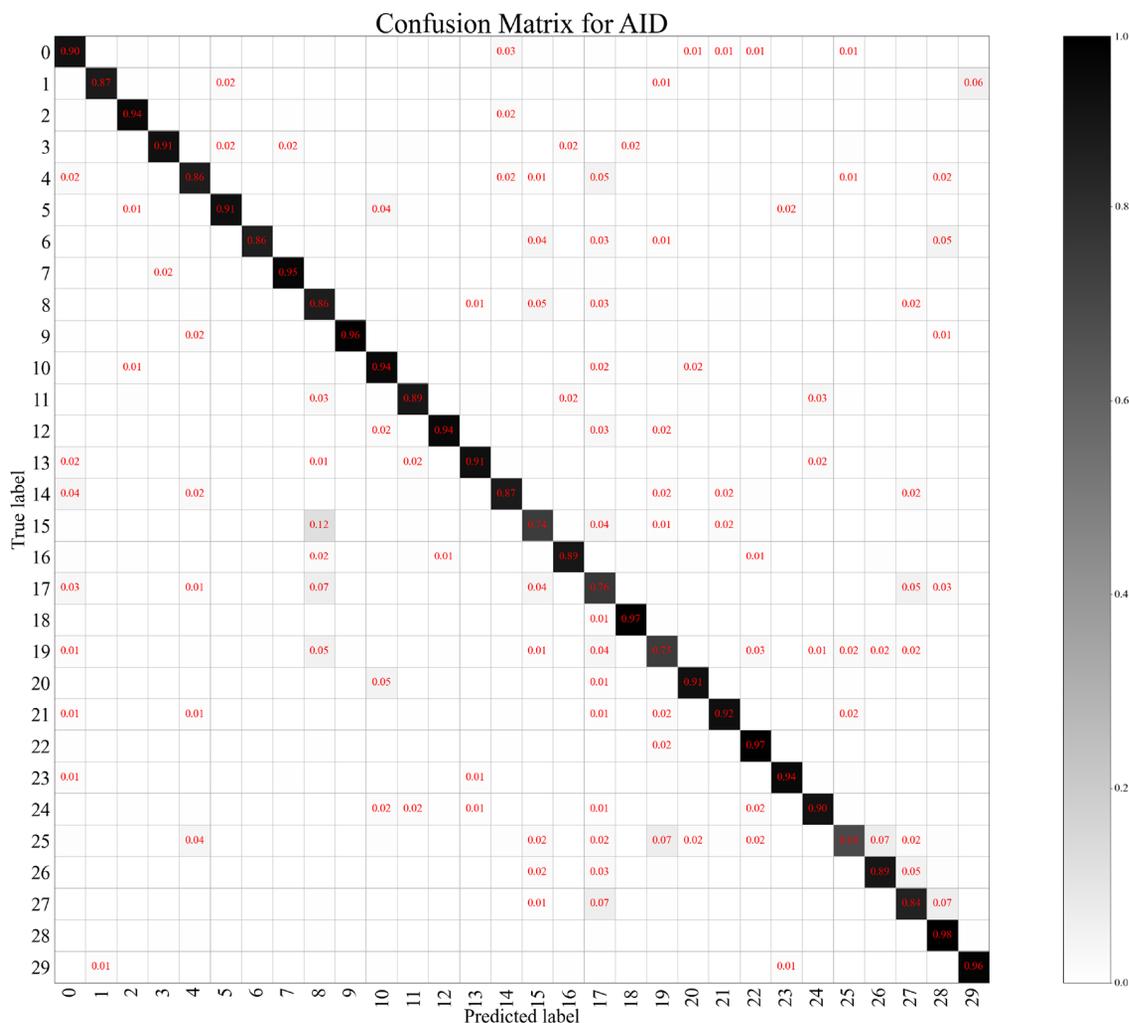


Figure 9. The confusion matrix CM of contrastive learning-convolutional neural network CL-CNN with 50% data for training over the aerial image data set AID dataset.

industrial, meadow, medium residential, mountain, park, parking, playground, pond, baseball field, port, railway station, resort, river, school, sparse residential, square, stadium, storage tanks, viaduct, beach, bridge, center, church, commercial, dense residential, and desert, respectively, show the classification cases of all categories. We find that the classification result for the center is not satisfying in the two groups of experiments. To better explain the issue, we select some samples from this category in the AID data set. As shown in Figure 10, these HRRS images are typical samples of the center, where the center buildings' outline is round. In contrast, misclassification HRRS images shown in Figure 11 have multiple kinds of outlines and are significantly different from typical samples in this category. That is the reason why the classification accuracy of this category is not significant enough.

3. A achieving satisfying performance over the NWPU-45 data set is a challenging task. However, this data set can present the ability of the model, capturing class-specific representation. To compare with existing methods, we perform two groups of experiments using 10% and 20% data for training. Furthermore, these training ratios mean there are only quite limited numbers of HRRS images used for training, which deepen the difficulty in obtaining excellent performance over NWPU-45. As shown in Table 5, although the NWPU-45 data set is challenging enough, and only a few HRRS images are used for training, CL-CNN still achieves significant improvement in the two groups of experiments. Comparing with the state-of-the-art model, CL-CNN obtains 2.85% and 6.01% improvement, which is excellent progress and proves the excellent performance of CL-CNN over complex data sets with limited training samples. Although CL-CNN makes excellent progress on scene classification over the NWPU-45 data set, CL-CNN classification performance over some categories in NWPU-45 is not satisfying. As shown in Figures 12 and 13, where numbers from 0 to 44 denote airplane, airport, commercial area, dense residential, desert, forest, freeway, golf course, ground track field, harbor, industrial area, intersection, baseball diamond, island, lake, meadow, medium residential, mobile home park, mountain, overpass, palace,

parking lot, railway, basketball court, railway station, rectangular farmland, river, roundabout, runway, sea ice, ship, snow-berg, sparse residential, stadium, beach, storage tank, tennis court, terrace, thermal power station, wetland, bridge, chaparral, church, circular farmland, and cloud; classification results of palace and church are not remarkable enough. As shown in Figures 14–17, images in these two categories have a sizeable intraclass discrepancy, which means data distributions of these two categories are difficult to be fitted with limited training samples.

*Comparisons Between CL-CNN with Supervised Deep Learning Representation HRRS Image Scene Classification Methods over Selected Benchmark Data Sets*

Although CL-CNN comparisons with current existing unsupervised deep learning representation HRRS image scene classification present the effectiveness and improvement of CL-CNN, they cannot display the great potential in practical tasks comparing supervised deep learning representation methods. Moreover, we still need experiments to compare CL-CNN with supervised deep learning representation methods to indicate the possibility of replacing supervised deep learning representation methods in certain circumstances where there are limited labeled samples, and a CNN feature extractor still needs to be trained.

We select three widely used supervised deep learning representation HRRS image scene classification methods, VG-Net-16, GoogleNet, and ResNet-50 to compare with CL-CNN.

Table 5. The performance comparison of models on the NWPU-45 data set.

Model	OA (20%)	OA (50%)
LLC (Cheng <i>et al.</i> 2017)	38.81 ± 0.23	40.03 ± 0.34
BoVW +SPM (Cheng <i>et al.</i> 2017)	27.83 ± 0.61	32.96 ± 0.47
BoVW (Cheng <i>et al.</i> 2017)	41.72 ± 0.21	44.97 ± 0.28
MARTA GANs (Lin <i>et al.</i> 2017)	68.63 ± 0.22	75.03 ± 0.28
Attention GANs (Yu <i>et al.</i> 2020)	71.21 ± 0.21	75.03 ± 0.28
CL-CNN	74.06 ± 0.14	81.04 ± 0.45

OA = overall accuracy; LLC = locality-constrained linear coding; BoVW = bag-of-visual-words; CL-CNN = contrastive learning-convolutional neural network.



Figure 10. Typical samples belonging to center in the aerial image data set AID dataset.



Figure 11. Misclassification high-resolution remote sensing HRRS images belonging to center in the aerial image data set AID dataset.

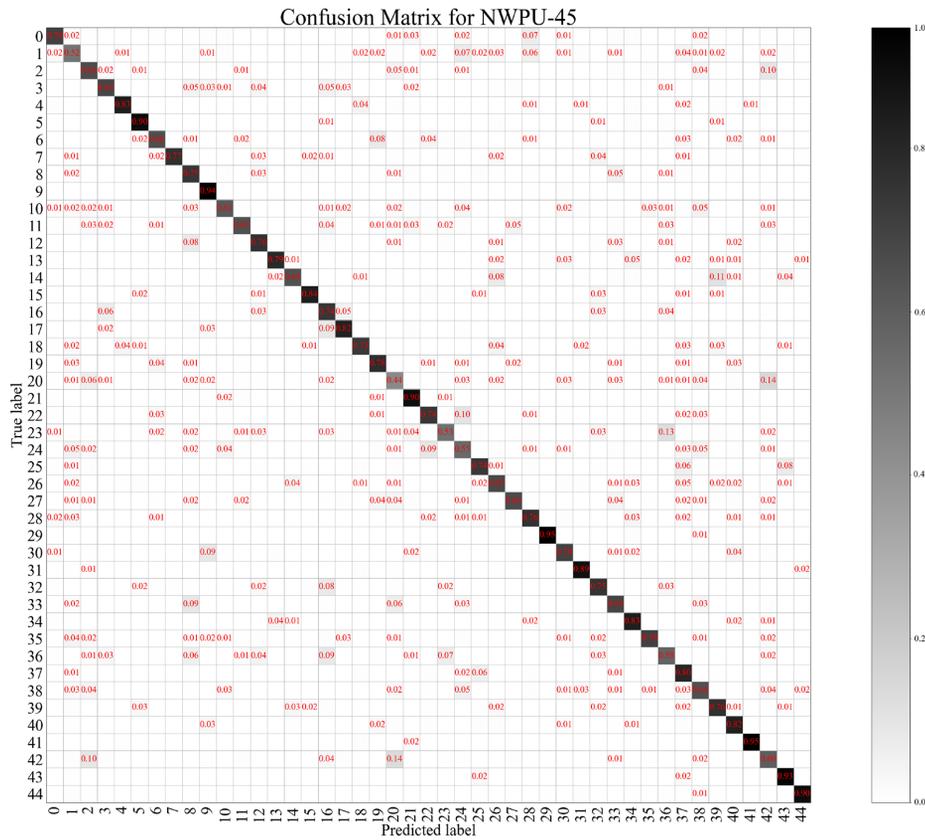


Figure 12. The confusion matrix of contrastive learning-convolutional neural network CL-CNN with 10% data for training over the NWPU-45 data set.

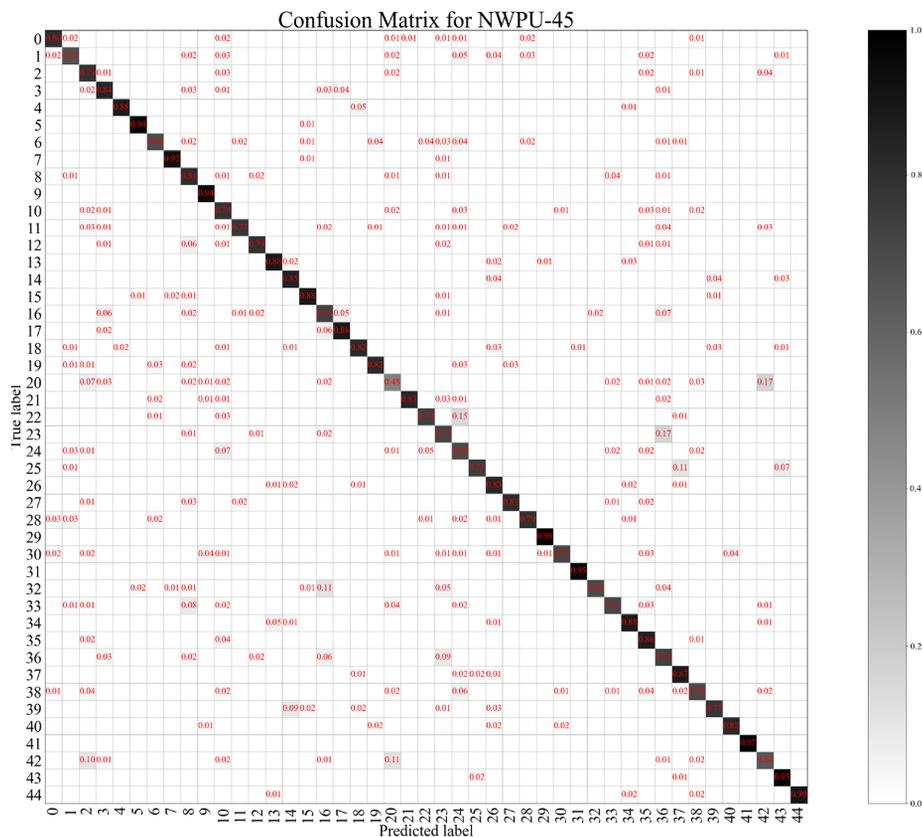


Figure 13. The confusion matrix of contrastive learning-convolutional neural network CL-CNN with 20% data for training over the NWPU-45 data set.

These three selected models are trained from scratch with the same ratios of labeled samples compared with CL-CNN. To make the performance of the selected three methods

more persuasive, we directly cite experimental results of the RSSCN7 data set and the AID data set from Cheng et al (2018) and experimental results of the NWPU-45 data set from Xia



Figure 14. Typical samples belonging to palace in the NWPU-45 data set.



Figure 15. Misclassification high-resolution remote sensing HRRS images belonging to palace in the NWPU-45 data set.



Figure 16. Typical samples belonging to church in the NWPU-45 data set.



Figure 17. Misclassification high-resolution remote sensing HRRS images belonging to church in the NWPU-45 data set.

Table 6. The performance comparisons between contrastive learning-convolutional neural network and supervised deep learning representation methods.

Model	Data Set					
	RSSCN7 OA (20%)	RSSCN7 OA (50%)	AID OA (20%)	AID OA (50%)	NWPU-45 OA(10%)	NWPU-45 OA (20%)
VGGNet-16	83.98 ± 0.87	87.18 ± 0.94	86.59 ± 0.29	89.64 ± 0.36	76.47 ± 0.18	79.79 ± 0.15
GoogleNet	82.55 ± 1.11	85.84 ± 0.94	83.44 ± 0.40	86.39 ± 0.55	76.19 ± 0.38	78.48 ± 0.26
ResNet-50	82.54 ± 0.21	89.20 ± 0.13	86.53 ± 0.12	90.22 ± 0.34	83.88 ± 0.26	85.19 ± 0.71
CL-CNN	85.99 ± 0.47	91.11 ± 0.18	81.89 ± 0.74	89.05 ± 0.23	74.06 ± 0.14	81.11 ± 0.45

OA = overall accuracy; AID = aerial image data set; CL-CNN = contrastive learning-convolutional neural network.

*et al.* (2017). However, there are only experimental results of ResNet-50 initiated with ImageNet weight. We obtain the performance of ResNet-50 on the three selected benchmark data sets by several experiments. The code of ResNet-50 is based on Pytorch. We use the Adam optimizer with  $lr = 1 \times 10^4$ , which is updated in every five training epoch. Furthermore, augmentation operation, including flipping vertical and flipping horizontal, are used in the training procedure.

As shown in Table 6, the CL-CNN has better performance than the other three supervised deep learning representation methods on the RSSCN7 data set. Although these three supervised representation methods are trained with labeled information, they do not consider each samples interclass and intraclass relations. And the category settings of the RSSCN7 data set is simple, which is the other reason why CL-CNN obtains excellent performance. It still presents that training a deep learning model on a small-scale data set with labeled samples cannot guarantee that extracted features are discriminative enough, especially when prior knowledge is not applied to initial the used deep learning model weights. The performance of CL-CNN over RSSCN7 shows that CL-CNN achieves much more discriminative data representation of HRRS images in the RSSCN7 data set comparing with supervised deep learning data representation methods.

However, the performance of the CL-CNN over the AID data set and the NWPU-45 data set is not competitive. When the ratios of training samples are low, the CL-CNN performance is not remarkable enough. These experimental results are only better than GoogleNet on the AID data set and the NWPU-45 data set. That displays that the representation capacity of CL-CNN is not strong enough. Although the performance of CL-CNN with high ratios of training samples has made a significant improvement, it is not remarkable comparing with supervised ResNet-50.

We find that the CL-CNN can replace supervised deep learning methods over the RSSCN7 data set by these comparisons. It can obtain better performance than supervised representation methods in the data set. Moreover, in large data sets, such as the AID data set and the NWPU-45 data set, the CL-CNN can achieve classification accuracy close to supervised representation methods with higher training sample ratios. However, the CL-CNN still needs to be improved to handle complex data sets with lower training sample ratios.

## Conclusion

In this paper, we introduce an unsupervised representation method, CL-CNN, for HRRS image scene classification. We introduce the current research status of unsupervised representation for HRRS image scene classification. Motivated by the human visual system, we propose CL-CNN. And then, we explain the details of CL-CNN. In CL-CNN, we regard color channels of HRRS images as positive views and color channels from different HRRS images as negative views. CL-CNN obtains class-specific representation by shrinking the distance between positive view pairs and enlarging the distance between negative view pairs without any labeled information during the training feature extractor procedure. To better present CL-CNN characteristics, we discuss the influence of different  $k$  denoting the number of negative view pairs values for representation results by experiments over the RSSCN7 data set. Furthermore, we present the difference resulted from several kinds of color spaces. The experimental results over RSSCN7, AID, NWPU-45, with  $k = N - 1$  and YCbCr color space, indicate the excellent ability of CL-CNN capturing class-specific representation without labeled information. In the future, we will improve the performance of CL-CNN further and develop a completely unsupervised deep learning representation HRRS image scene classification method.

## Acknowledgments

The authors would like to thank Prof. Qian Wang, Prof. Gong Cheng, and Prof. Guisong Xia for sharing the RSSCN7 data set, NWPU-45 data set, and AID data sets. We would also like to thank editors, associated editors, and anonymous reviewers for their insightful suggestions and comments, which significantly improve the paper. This work is supported in part by the National Natural Science Foundation of China (No. U1711266, No.41701429).

## References

- Bhagavathy, S. and B. Manjunath. 2006. Modeling and detection of geospatial objects using texture motifs. *IEEE Transactions on Geoscience and Remote Sensing* 44(12):3706–3715.
- Bo, D., X. Wei, W. Jia, Z. Lefei, Z. Liangpei and T. Dacheng. 2017. Stacked convolutional denoising auto-encoders for feature representation. *IEEE Transactions on Cybernetics* 47:1017–1027.
- Chen, F., K. Wang, T. Van de Voorde and T. F. Tang. 2017. Mapping urban land cover from high spatial resolution hyperspectral data: An approach based on simultaneously unmixing similar pixels with jointly sparse spectral mixture analysis. *Remote Sensing of Environment* 196:324–342.
- Cheng, G., L. Guo, T. Zhao, J. Han, H. Li and J. Fang. 2013. Automatic landslide detection from remote-sensing imagery using a scene classification method based on BoVW and pLSA. *International Journal of Remote Sensing* 34(1):45–59.
- Cheng, G., J. Han, P. Zhou and L. Guo. 2014. Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing* 98:119–132.
- Cheng, J. and W. P. Kustas. 2019. Using very high resolution thermal infrared imagery for more accurate determination of the impact of land cover differences on evapotranspiration in an irrigated agricultural area. *Remote Sensing* 11(6):613–625.
- Cheng, G., Z. Li, X. Yao, L. Guo and Z. Wei. 2017. Remote sensing image scene classification using bag of convolutional features. *IEEE Geoscience and Remote Sensing Letters* 14(10):1735–1739.
- Cheng, G., C. Yang, X. Yao, L. Guo and J. Han. 2018. When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs. *IEEE Transactions on Geoscience and Remote Sensing* 56(5):2811–2821.
- Cheng, Z., Q. Yang and B. 2015a. Sheng. 2015a. Deep colorization. Pages 415–423 in *Proceedings Computer Vision and Pattern Recognition*, held in Boston, Mass., June 2015.
- Cheng, G., P. Zhou, J. Han, L. Guo and J. Han. 2015b. Auto-encoder-based shared mid-level visual dictionary learning for scene classification using very high-resolution remote sensing images. *International Journal of Computer Vision* 9(5):639–47.
- Cheng, G., P. Zhou and J. Han. 2016. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 54(12):7405–7415.
- Dalal, N. and B. Triggs. 2005. Histograms of oriented gradients for human detection. Pages 886–93 in *Proceedings Computer Vision and Pattern Recognition*, held in San Diego, Calif., June 2005.
- Du, B., S. Cai and C. Wu. 2019. Object tracking in satellite videos based on a multi frame optical flow tracker. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 12(8):3043–3055.
- Fan, Z., D. Bo and Z. Liangpei. 2014. Saliency-guided unsupervised feature learning for scene classification. *IEEE Transactions on Geoscience and Remote Sensing* 53(4):2175–84.
- Fu, J., H. Zheng and T. Mei. 2016. Meina and shiguang shan and xilin chen. Pages 4847–4855 in *Proceedings Computer Vision and Pattern Recognition*, held in Las Vegas, Nev., June 2016.
- Gómez, C., J. C. White and M. A. Wulder. 2016. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 116:55–72.

- Han, W., R. Feng, L. Wang and Y. Cheng. 2018. A semi-supervised generative framework with deep learning features for high-resolution remote sensing image scene classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 145:23–43.
- Hartigan, J. A. and M. A. Wong. 1979. A k-means clustering algorithm. *Journal of the Royal Statistical Society, Series C (Applied Statistics)* 28(1):100–108.
- He, K., X. Zhang, S. Ren and J. Sun. 2016. Deep residual learning for image recognition. Pages 770–778 in *Proceedings Computer Vision and Pattern Recognition*, Las Vegas, Nev., June 2016.
- He, K., H. Fan, Y. Wu, S. Xie and R. Girshick. 2019. Momentum contrast for unsupervised visual representation learning. Pages 9726–9735 in *Proceedings Computer Vision and Pattern Recognition*, Long Beach, Calif., June 2019.
- He, N., L. Fang, S. Li, A. Plaza and J. Plaza. 2018. Remote sensing scene classification using multilayer stacked covariance pooling. *IEEE Transactions on Geoscience and Remote Sensing* 56(12):6899–6910.
- Higgins, I., L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed and A. Lerchner. 2017. beta-vae: Learning basic visual concepts with a constrained variational framework. In *Proceedings International Conference on Learning Representations*, held in Toulon, France, April 2017. <https://openreview.net/forum?id=Sy2fzU9gl>.
- Kingma, D. P. and M. Welling. Auto-encoding variational bayes. 20145. In *Proceedings International Conference on Learning Representations*, held in Banff, Alberta, Canada, April 2014. <http://arxiv.org/abs/1312.6114>.
- Krizhevsky, A., I. Sutskever and G. E. Hinton. 2012. ImageNet classification with deep convolutional neural networks. Pages 84–90 in *Proceedings Advances in Neural Information Processing Systems*, Lake Tahoe, Nev., December 2012.
- Larsson, G., M. Maire and G. Shakhnarovich. 2016. Learning representations for automatic colorization. Pages 577–593 in *Proceedings European Conference on Computer Vision*, Amsterdam, Netherlands, October 2016.
- Lazebnik, S., C. Schmid and J. A. Ponce. 2006. CNN-Based method for infant cry detection and recognition. Pages 2169–2178 in *Proceedings Computer Vision and Pattern Recognition*, New York, N.Y., June 2006.
- Li, A., Z. Lu, L. Wang, T. Xiang and J. Wen. 2017. Zero-shot scene classification for high spatial resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 55(7):4157–4167.
- Li, F., R. Feng, W. Han and L. Wang. 2020. High-resolution remote sensing image scene classification via key filter bank based on convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing* 58(11):8077–8092.
- Li, Y., C. Tao, Y. Tan, K. Shang and J. Tian. 2016. Unsupervised multilayer feature learning for satellite image scene classification. *IEEE Geoscience and Remote Sensing Letters* 13(2):157–161.
- Lin, D., K. Fu, Y. Wang, G. Xu and X. Sun. 2017. MARTA GANs: Unsupervised representation learning for remote sensing image classification. *IEEE Geoscience and Remote Sensing Letters* 14(11):2092–2096.
- Liu, T. and A. Abd-Elrahman. 2018. An object-based image analysis method for enhancing classification of land covers using fully convolutional networks and multi-view images of small unmanned aerial system. *Remote Sensing* 10(3):457–481.
- Liu, G., Y. Gousseau and F. Tupin. 2019. A contrario comparison of local descriptors for change detection in very high spatial resolution satellite images of urban areas. *IEEE Transactions on Geoscience and Remote Sensing* 57(6):3904–3918.
- Liu, Y. and C. Huang. 2017. Scene classification via triplet networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(1):220–237.
- Liu, Y., C. Y. Suen, Y. Liu and L. Ding. 2018. Scene classification using hierarchical Wasserstein CNN. *IEEE Transactions on Geoscience and Remote Sensing* 57(5):2494–2509.
- Locatello, F., S. Bauer, M. Lucic, G. Rätsch, S. Gelly, B. Schölkopf and O. Bachem. 2019. Challenging common assumptions in the unsupervised learning of disentangled representations. In *Proceedings 36th International Conference on Machine Learning*, New Orleans, La., May 2019. <https://openreview.net/forum?id=Byg6VhUp8V>.
- Lowe, D. G. 2004. Distinctive image features from scale-invariant key points. *International Journal of Computer Vision* 60(2):91–110.
- Lu, X., X. Zheng and Y. Yuan. 2017. Remote sensing scene classification by unsupervised representation learning. *IEEE Transactions on Geoscience and Remote Sensing* 55(9):5148–5157.
- Lv, P., Y. Zhong, J. Zhao and L. Zhang. 2018. Unsupervised change detection based on hybrid conditional random field model for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing* 56(7):4002–4015.
- Ma, D., P. Tang and L. Zhao. 2019. Siftinggan: Generating and sifting labeled samples to improve the remote sensing image scene classification baseline in vitro. *IEEE Geoscience and Remote Sensing Letters* 16(7):1046–1050.
- Masci, J., U. Meier, D. Cire an and J. Schmidhuber. 2011. Stacked convolutional auto-encoders for hierarchical feature extraction. Pages 52–59 in *Proceedings International Conference on Artificial Neural Networks*, Espoo, Finland, June 2011.
- Milani, G., M. Volpi, D. Tonolla, M. Doering, C. Robinson, M. Kneubühler and M. Schaepman. 2018. Robust quantification of riverine land cover dynamics by high-resolution remote sensing. *Remote Sensing of Environment* 217:491–505.
- Minetto, R., M. P. Segundo and S. Sarkar. 2019. Hydra: An ensemble of convolutional neural networks for geospatial land classification. *IEEE Transactions on Geoscience and Remote Sensing* 57(9):6530–6541.
- Myint, S. W., P. Gober, A. Brazel, S. Grossman-Clarke and Q. Weng. 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment* 115(5):1145–1161.
- Ojala, T., M. Pietikainen and T. Maenpaa. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7):971–987.
- Olshausen, B. A. and D. J. Field. 1997. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research* 37(23):3311–3325.
- Poterek, Q., P. Herrault, G. Skupinski and D. Sheeren. 2020. Deep learning for automatic colorization of legacy grayscale aerial photographs. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 13(16):2899–2915.
- Risojević, V. and Z. Babić. 2012. Fusion of global and local descriptors for remote sensing image classification. *IEEE Geoscience and Remote Sensing Letters* 10(4):836–840.
- Romero, A., C. Gatta and G. Camps-Valls. 2015. Unsupervised deep feature extraction for remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing* 54(3):1349–1362.
- Saha, S., F. Bovolo and L. Bruzzone. 2019. Unsupervised deep change vector analysis for multiple-change detection in VHR images. *IEEE Transactions on Geoscience and Remote Sensing* 57(6):3677–3693.
- Sheng, G., W. Yang, T. Xu and H. Sun. 2012. High-resolution satellite scene classification using a sparse coding based multiple feature combination. *International Journal of Remote Sensing* 33(8):2395–2412.
- Simonyan, K. and A. Zisserman. 2015. Very deep convolutional networks for large-scale image recognition. Pages 1–9 in *Proceedings International Conference on Learning Representations*, San Diego, Calif., May 2015.
- Song, J., X. Tong, L. Wang, C. Zhao and A. V. Prishchepov. 2019. Monitoring finer-scale population density in urban functional zones: A remote sensing data fusion approach. *Landscape and Urban Planning* 190:103580. <https://doi.org/10.1016/j.landurbplan.2019.05.011>.

- Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich. 2015. Going deeper with convolutions. Pages 1–9 in *Proceedings Conference on Computer Vision and Pattern Recognition*, Boston, Mass., June 2015.
- Tao, Y., M. Xu, F. Zhang, B. Du and L. Zhang. 2017. Unsupervised-restricted deconvolutional neural network for very high resolution remote-sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing* 55(12):6805–6823.
- Tian, Y., D. Krishnan and P. Isola. 2019. Contrastive multiview coding. arXiv preprint arXiv:190605849.
- Vincent, P., H. Larochelle, I. Lajoie, Y. Bengio and P. A. Manzagol. 2010. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research* 11(Dec):3371–3408.
- Wang, S., M. Garcia, P. Bauer-Gottwein, J. Jakobsen, P. J. Zarco-Tejada, F. Bandini, F. Bandini, V. Sobejano Paz, and A. Ibrom. 2019. High spatial resolution monitoring land surface energy, water and co2 fluxes from an unmanned aerial system. *Remote Sensing of Environment* 229:14–31.
- Wang, Q., S. Liu, J. Chanussot and X. Li. 2018a. Scene classification with recurrent attention of VHR remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing* 57(2):1155–1167.
- Wang, J., W. Liu, L. Ma, H. Chen and L. Chen. 2018b. IORN: An effective remote sensing image scene classification framework. *IEEE Geoscience and Remote Sensing Letters* 15(11):1695–1699.
- Wang, J., J. Yang, K. Yu, F. Lv, T. S. Huang and Y. Gong. 2010. Locality-constrained linear coding for image classification. Pages 3360–3367 in *Proceedings Conference on Computer Vision and Pattern Recognition*, San Francisco, Calif., June 2010.
- Wold, S., K. Esbensen and P. Geladi. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2(1–3):37–52.
- Xia, G. S., J. Hu, F. Hu, B. Shi, X. Bai, Y. Zhong, L. Zhang, and X. Lu. 2017. AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions Geoscience and Remote Sensing* 55(7):3965–3981.
- Yang, N., H. Tang, H. Sun and X. Yang. 2018. Drop band: A simple and effective method for promoting the scene classification accuracy of convolutional neural networks for VHR remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters* 15(2):257–261.
- Yu, Y., X. Li and F. Liu. 2020. Attention GANs: Unsupervised deep feature learning for aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing* 58(1):519–531.
- Yu, B., L. Yang and F. Chen. 2018. Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11(9):3252–3261.
- Yuan, Y., J. Fang, X. Lu and Y. Feng. 2018. Remote sensing image scene classification using rearranged local features. *IEEE Transactions on Geoscience and Remote Sensing* 57(3):1779–1792.
- Zhang, F., B. Du and L. Zhang. 2015. Scene classification via a gradient boosting random convolutional network framework. *IEEE Transactions on Geoscience and Remote Sensing* 54(3):1793–1802.
- Zhang, Y., Y. Yuan, Y. Feng and X. Lu. 2019. Hierarchical and robust convolutional neural network for very high-resolution remote sensing object detection. *IEEE Transactions on Geoscience and Remote Sensing* 57(8):5535–5548.
- Zhang, H., J. Zhang and F. Xu. Land use and land cover classification base on image saliency map cooperated coding. 2015. Pages 2616–2620 in *Proceedings International Conference on Image Processing*, Quebec City, Quebec, Canada, September 2015.
- Zhao, B., Y. Zhong, G. S. Xia and L. Zhang. 2015. Dirichlet-derived multiple topic scene classification model for high spatial resolution remote sensing imagery. *IEEE Transactions on Geoscience and Remote Sensing* 54(4):2108–2123.
- Zheng, X., Y. Yuan and X. Lu. 2019. A deep scene representation for aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing* 57(7):4799–4809.
- Zhu, Z., J. Zhang, A. Yang, A. H. Aljaddani, W. B. Cohen, S. Qiu and C. Zhou. 2020. Continuous monitoring of land disturbance based on Landsat time series. *Remote Sensing of Environment* 238:1111–1116.
- Zhu, Q., Y. Zhong, B. Zhao, G. S. Xia and L. Zhang. 2016. Bag-of-visual-words scene classifier with local and global features for high spatial resolution remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters* 13(6):747–751.
- Zou, J., W. Li, C. Chen and Q. Du. 2016. Scene classification using local and global features with collaborative representation fusion. *Information Sciences* 348:209–226.
- Zou, Q., L. Ni, T. Zhang and Q. Wang. 2015. Deep learning-based feature selection for remote sensing scene classification. *IEEE Geoscience and Remote Sensing Letters* 12(11):2321–2325.

# Digital Elevation Model Technologies and Applications The DEM Users Manual, 3<sup>rd</sup> Edition

Edited by David F. Maune, PhD, CP  
and Amar Nayegandhi, CP, CMS

To order, visit  
<https://www.asprs.org/dem>

The 3rd edition of the DEM Users Manual includes 15 chapters and three appendices. References in the eBook version are hyperlinked. Chapter and appendix titles include:

1. Introduction to DEMs  
*David F. Maune, Hans Karl Heidemann,  
Stephen M. Kopp, and Clayton A. Crawford*
  2. Vertical Datums  
*Dru Smith*
  3. Standards, Guidelines & Specifications  
*David F. Maune*
  4. The National Elevation Dataset (NED)  
*Dean B. Gesch, Gayla A. Evans,  
Michael J. Oimoen, and Samantha T. Arundel*
  5. The 3D Elevation Program (3DEP)  
*Jason M. Stoker, Vicki Lukas, Allyson L. Jason,  
Diane F. Eldridge, and Larry J. Sugarbaker*
  6. Photogrammetry  
*J. Chris McGlone and Scott Arko*
  7. IfSAR  
*Scott Hensley and Lorraine Tighe*
  8. Airborne Topographic Lidar  
*Amar Nayegandhi and Joshua Nimetz*
  9. Lidar Data Processing  
*Joshua M. Novac*
  10. Airborne Lidar Bathymetry  
*Jennifer Wozencraft and Amar Nayegandhi*
  11. Sonar  
*Guy T. Noll and Douglas Lockhart*
  12. Enabling Technologies  
*Bruno M. Scherzinger, Joseph J. Hutton,  
and Mohamed M.R. Mostafa*
  13. DEM User Applications  
*David F. Maune*
  14. DEM User Requirements & Benefits  
*David F. Maune*
  15. Quality Assessment of Elevation Data  
*Jennifer Novac*
- Appendix A. Acronyms  
Appendix B. Definitions  
Appendix C. Sample Datasets

This book is your guide to 3D elevation technologies, products and applications. It will guide you through the inception and implementation of the U.S. Geological Survey's (USGS) 3D Elevation Program (3DEP) to provide not just bare earth DEMs, but a full suite of 3D elevation products using Quality Levels (QLs) that are standardized and consistent across the U.S. and territories. The 3DEP is based on the National Enhanced Elevation Assessment (NEEA) which evaluated 602 different mission-critical requirements for and benefits from enhanced elevation data of various QLs for 34 Federal agencies, all 50 states (with local and Tribal input), and 13 non-governmental organizations.

The NEEA documented the highest Return on Investment from QL2 lidar for the conterminous states, Hawaii and U.S. territories, and QL5 IfSAR for Alaska.

Chapters 3, 5, 8, 9, 13, 14, and 15 are "must-read" chapters for users and providers of topographic lidar data. Chapter 8 addresses linear mode, single photon and Geiger mode lidar technologies, and Chapter 10 addresses the latest in topobathymetric lidar. The remaining chapters are either relevant to all DEM technologies or address alternative technologies including photogrammetry, IfSAR, and sonar.

As demonstrated by the figures selected for the front cover of this manual, readers will recognize the editors' vision for the future – a 3D Nation that seamlessly merges topographic and bathymetric data from the tops of the mountains, beneath rivers and lakes, to the depths of the sea.

Co-Editors

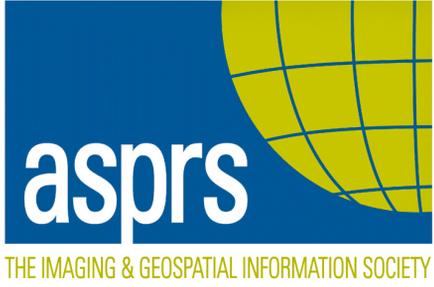
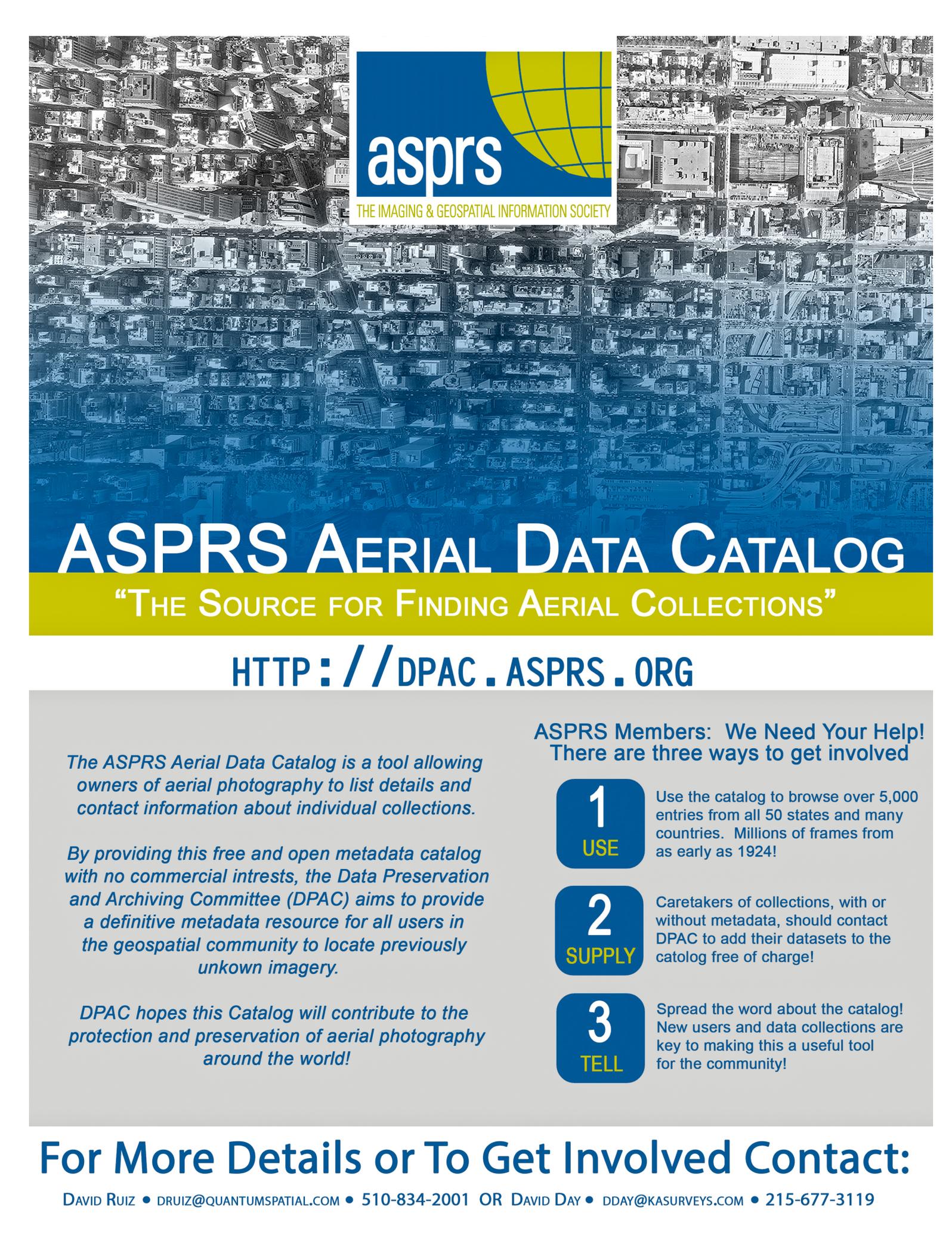
David F. Maune, PhD, CP and  
Amar Nayegandhi, CP, CMS

---

## PRICING

Student (must submit copy of Student ID)	\$50 +S&H
ASPRS Member	\$80 +S&H
Non-member	\$100 +S&H
E-Book (only available in the Amazon Kindle store)	\$85

---



asprs

THE IMAGING & GEOSPATIAL INFORMATION SOCIETY

# ASPRS AERIAL DATA CATALOG

“THE SOURCE FOR FINDING AERIAL COLLECTIONS”

[HTTP://DPAC.ASPRS.ORG](http://dpac.asprs.org)

*The ASPRS Aerial Data Catalog is a tool allowing owners of aerial photography to list details and contact information about individual collections.*

*By providing this free and open metadata catalog with no commercial interests, the Data Preservation and Archiving Committee (DPAC) aims to provide a definitive metadata resource for all users in the geospatial community to locate previously unknown imagery.*

*DPAC hopes this Catalog will contribute to the protection and preservation of aerial photography around the world!*

**ASPRS Members: We Need Your Help!**  
There are three ways to get involved

1

USE

Use the catalog to browse over 5,000 entries from all 50 states and many countries. Millions of frames from as early as 1924!

2

SUPPLY

Caretakers of collections, with or without metadata, should contact DPAC to add their datasets to the catalog free of charge!

3

TELL

Spread the word about the catalog! New users and data collections are key to making this a useful tool for the community!

## For More Details or To Get Involved Contact:

DAVID RUIZ • [DRUIZ@QUANTUMSPATIAL.COM](mailto:DRUIZ@QUANTUMSPATIAL.COM) • 510-834-2001 OR DAVID DAY • [DDAY@KASURVEYS.COM](mailto:DDAY@KASURVEYS.COM) • 215-677-3119

LEARN  
DO  
GIVE  
BELONG

**ASPRS Offers**

- » Cutting-edge conference programs
- » Professional development workshops
- » Accredited professional certifications
- » Scholarships and awards
- » Career advancing mentoring programs
- » *PE&RS*, the scientific journal of ASPRS

[asprs.org](http://asprs.org)

ASPRS