R&D FUNDING PROGRAM

The National Reconnaissance Office Director’s Innovation Initiative (DII) program funds cutting-edge scientific research in a high-risk, high-payoff environment to discover innovative concepts and creative ideas that transform overhead intelligence capabilities and systems for future national security intelligence needs. The program seeks the brightest minds and breakthrough technologies from industry, academia, national laboratories, and U.S. government agencies.

Visit the website for Broad Agency Announcement and Government Sources Sought Announcement requirements.

703.808.2769

www.nro.gov/Business-Innovation-Opportunities
ANNOUNCEMENTS

TCarta Marine, a global provider of marine geospatial products, has supplied the National Oceanic and Atmospheric Administration (NOAA) with satellite-derived bathymetry (SDB) validated by green laser altimeter data from the NASA ICESat-2 satellite for two U.S. shallow-water coastal areas.

TCarta was the subcontractor on the prime contract awarded to Woolpert, an international geospatial firm headquartered in Dayton, Ohio.

The NOAA pilot focused on two shallow-water regions that were 3,000 square kilometers in total area - one in the Green Bay area of western Lake Michigan and the other around Cape Cod and Nantucket Sound. Both areas experience natural forces that alter the underwater terrain faster than traditional bathymetric surveys can be completed.

SDB is being considered as a fast and inexpensive alternative for such coastal zones. A primary advantage of SDB is that orbiting satellites can be tasked to collect up-to-date imagery, with bathymetric measurements derived to create end products in a matter of days.

TCarta created products in the two project pilot areas to enable NOAA to test the use of SDB in shallow coastal zones. This provided updated bathymetry and infilled data gaps from traditional bathymetric measurement technologies, such as airborne lidar or marine sonar. The pilot was conducted jointly by NOAA’s Office of Coast Survey and National Geodetic Survey’s Remote Sensing Division. The SDB data sets measured the seafloor to a depth of 20-25 meters, with validation using ICESat-2.

“The integration of ICESat2 bathymetry with satellite imagery derived bathymetry is a powerful combination to increase accuracy,” said RDML Shepard Smith, the recently retired Director of NOAA’s Office of Coast Survey. “It holds the potential to dramatically improve our ability to monitor change in dynamic coastal waters and allow us to map remote areas for the first time.”

For more information, visit www.tcarta.com.

The U.S. National Science Foundation and NASA have signed a memorandum of understanding establishing the framework for collaborative efforts to broaden participation in engineering.

The collaboration will involve NASA’s Minority University Research and Education program, which engages underrepresented populations through a wide variety of initiatives, and NSF’s Broadening Participation in Engineering and NSF INCLUDES programs. NSF INCLUDES supports national infrastructure for collaborations that broaden participation in STEM fields for historically underrepresented groups. NSF’s Broadening Participation in Engineering program supports research to develop a diverse, inclusive and well-prepared engineering workforce.

The goal with this new agreement is to leverage NASA and NSF programs to build coalitions of public and private organizations who use evidence-based concepts for broadening participation of underrepresented groups in engineering,” said NSF Assistant Director for Education and Human Resources Karen Marrongelle.

Under the new agreement, NSF and NASA intend to collaborate on a common agenda and joint review of proposals; the agreement also provides more flexibility to support research, education, and workforce development proposals of mutual interest to advance diversity, equity and inclusion in engineering.

The new partnership aims to broaden participation in engineering by expanding opportunities for institutions and organizations to engage students and researchers through NSF and NASA programs. Activities may include educational experiences for students from kindergarten through college, professional development of educators, new course and curriculum development, and workforce inclusion research.

For information about NSF’s broadening participation efforts and agency programs, visit nsf.gov. For information on NASA’s programs, visit nasa.gov.

TECHNOLOGY

GeoCue Group Inc. has updated its RIEGL-based True View® 3D Imaging Systems (3DIS®) with the launch of the new True View 635/640 3DIS. GeoCue’s True View 3DIS family brings a unique combination of fused lidar sensors and photogrammetric cameras in a fully calibrated platform, allowing direct generation of high accuracy, colorized lidar data in less time than is required to fly the mission.

GeoCue’s first generation of RIEGL miniVUX-2UAV-based 3DIS integrated systems, the True View 615 and 620, introduced compact, survey-grade 3DIS providing high network accuracy and precision (low noise). Combined with GeoCue’s industry-leading integrated data processing software suite, True View EVO, all GeoCue 3DIS include the full post-processing software workflow, including direct integration with Applanix POSPac.

Unlike more common lidar/camera UAS software suites that have separate workflows for point classification or point colorization, the integrated EVO workflow generates fully geo-referenced, point-traced, colorized point clouds in less time than it takes to fly the mission. As an added bonus, True View EVO supports the direct creation of many standard project deliverables including ground classified point clouds, surface models, contours, Digital Elevation Models (DEMs), volumetric analysis, wire extraction and similar products, without the need for additional third-party software.

Staying in step with RIEGL’s latest products, the True View 635/640 are equipped with the recently released RIEGL miniVUX-3UAV laser scanner. The miniVUX-3UAV, a 360° rotating mirror scanner, increases the scanner frequency to 300 kHz and offers a unique mode where the 200,000 pulse per second scan rate is focused in a 120° cross-track field of view, providing significantly increased point densities in aerial map-
Phase One, a developer of digital imaging technologies, has introduced its next-generation Phase One Aerial Systems (PAS). Among its many enhancements, PAS is now supported by the new iX Controller series which incorporates an Applanix GNSS-Inertial unit, making airborne mapping more efficient with the possible use of direct georeferencing.

Designed for increased productivity in shorter flight time, the PAS 280MP is setting new standards in RGB large-format aerial imaging in terms of image quality and return on investment. The high image capture rate of 2fps and 20,000-pixel swath combined with advanced blur control motion compensation technology ensures sharp image collection at high flight speeds. The system’s high dynamic range and improved light sensitivity mean more flight hours per day and year.

In addition, the PAS 280MP is light and compact, allowing it to be easily installed in a wide range of aircraft. The large-format aerial system offers an optional 4-Band configuration for simultaneous RGB color and near-infrared (NIR) image collection used in agriculture, forestry, and vegetation mapping applications. The PAS 150MP provides the same accuracy and versatility at 150MP resolution.

“We are proud to introduce our next-generation PAS with the addition of AP+, the latest GNSS technology from Applanix, enabling direct georeferencing without the use of ground control points — a major cost saver. The complete integration with the iX Controller, enables our geospatial clients to complete their aerial mapping projects faster and more efficiently,” said Dov Kalinski, VP of Geospatial Business.

The PAS 280MP and 150MP are delivered fully integrated with the new iX Controller MK5, serving as a computerized command center onboard the aircraft, supporting a variety of accuracy levels.

For information on next-generation Phase One Aerial Systems, visit: https://geospatial.phaseone.com/aerial-solutions/.

IVION Core offers both a refreshed look together with new features and improvements.

The enhancements in NavVis IVION Core are expressly designed to support laser scanning service providers, surveyors, and AEC Companies. These include multi-site functionality, updated user management, and site coordinate systems for survey-grade geo-registration of data.

For more information, visit www.navvis.com/ivion/core.

---

EVENT

The Earth Archive, an emerging conservation initiative that is transforming the fight against the climate crisis, announced today it will host its inaugural Earth Archive Virtual Congress on June 15-16, 2021. The event will present a revolutionary campaign to scan the entire Amazon Basin and bring together stakeholders interested in joining the project.

By scanning the planet’s land surface with very high-resolution lidar, the Earth Archive will create a true three-dimensional twin of our world — an open source, digital record of the Earth that will reflect the landscape exactly as it was at the time of scanning. With this endeavor, the Earth Archive is positioned to provide geospatial data that will serve as the legacy baseline for understanding and conserving our world.

The Earth Archive Virtual Congress will connect participants at all levels of society interested in partaking in the Amazon campaign, including researchers, students, NGOs, indigenous groups, government officials and corporations. Archeologists, anthropologists and Earth Archive researchers and stakeholders will team with geospatial, mapping and GIS entities and organizations to present solutions and findings.

Day One of the two-day summit will focus on understanding the current state of Amazon mapping, along with outlining the contributions that Earth Archive data can make in improving current understandings of Amazonian socio-natural systems. Day Two will focus on the broader contributions that an open-source Amazon scan can make to an array of sciences, policy making and other related fields.

Visit www.theearthearchivecongress.com/congress for the full schedule, list of speakers and link to register.

---

CALENDAR

- 16-20 August, URISA GIS Leadership Academy. For more information, visit www.urisa.org/education-events/urisa-gis-leadership-academy/.
- 3-6 October, GIS-Pro 2021, Baltimore, Maryland. For more information, visit www.urisa.org/gis-pro.
405 Spatial Resolution Enhancement for Large-Scale Land Cover Mapping via Weakly Supervised Deep Learning
Qiutong Yu, Wei Liu, Wesley Nunes Gonçalves, José Marcato Junior, and Jonathan Li
Multispectral satellite imagery is the primary data source for monitoring land cover change and characterizing land cover globally. However, the consistency of land cover monitoring is limited by the spatial and temporal resolutions of the acquired satellite images. The public availability of daily high-resolution images is still scarce. This article aims to fill this gap by proposing a novel spatiotemporal fusion method to enhance daily low spatial resolution land cover mapping using a weakly supervised deep convolutional neural network.

413 Ecological Functions and Human Activity Interference Evaluation in Ecological Protection Redline for Urban Environment
Guoming Li, Binbin He, Liang Li, Miao Yang, Qiongyi Huang, and Zihan Guo
The influences of ecological protection redlines on urban sustainable development can significantly vary from place to place with distinct ecological functions and human activity interference. However, recent research lacks the evaluation of ecological functions and human activity interference in ecological protection redline areas near urban areas. This article presents an evaluation method consisting of two evaluation criteria systems: an ecological function evaluation criteria system based on conserving water resources and soil and maintaining biodiversity and a human activity interference evaluation criteria system based on the human activity impact index, population and road network density, and the intensity of tourism activity and livelihood sewage discharge.

421 A High-Resolution Satellite DEM Filtering Method Assisted with Building Segmentation
Yihui Li, Fang Gao, Wentao Li, Peng Zhang, Yuan An, Xing Zhong, Yuwei Zhai, and Yongjian Yang
Digital elevation model (DEM) filtering is critical in DEM production, and large-area meter-level resolution DEM is mainly generated from high-resolution satellite images. However, the current DEM filtering methods are mostly aimed at laser scanning data and tend to excessively remove ground points when processing a satellite digital surface model (DSM). To accurately filter out buildings and preserve terrain, we propose a DEM filtering algorithm using building segmentation results of orthophoto.

431 A Combined Unmixing Framework for Impervious Surface Mapping on Medium-Resolution Images with Visible Shadows
Hui Luo and Nan Chen
Spectral unmixing methods with medium-resolution remote sensing images have become the main approach to mapping urban impervious-surface information. However, as more tall buildings appear, numerous visible shadows exist in medium-resolution images; these have usually been ignored in previous research, but they seriously affect accuracy. To solve this problem, we propose a combined unmixing framework to extract impervious surface in nonshadow and shadow areas, using linear and nonlinear unmixing models, respectively.

445 An Incremental Isomap Method for Hyperspectral Dimensionality Reduction and Classification
Yi Ma, Zechong Zheng, Yutang Ma, Mingcang Zhu, Ran Huang, Xueye Chen, Qingjun Peng, Yong He, Yufeng Lu, Guoqing Zhou, Zhigang Liu, and Mujie Li
Many manifold learning algorithms conduct an eigen vector analysis on a data-similarity matrix with a size of N×N, where N is the number of data points. Thus, the memory complexity of the analysis is no less than O(N^2). This article presents an incremental manifold learning approach to handle large hyperspectral data sets for land use identification.
Tides, currents, and human activity combine to produce art-like patterns in the waters around China’s Leizhou Peninsula. The patterns are visible in these images, acquired by the Operational Land Imager (OLI) on Landsat 8 on January 1, 2021, Norman Kuring of NASA’s Ocean Biology group applied color-filtering techniques to draw out the fine details in the water, but the features are real.

The cover image shows the waters around the entire Leizhou Peninsula, a landform that juts out into the South China Sea from the southernmost portion of mainland China. The Gulf of Tonkin (Beibu Gulf) lies to the west, and the Gionghou Strait (Hainan Strait) passes to the south.

Numerous bays, harbors, and estuaries line more than 1500 kilometers of coastline along the peninsula. Two on the peninsula’s western side—Anpu Harbor and Liusha Bay—are visible in the detailed images to the right.

In both waterbodies, the geometric patterns of aquaculture stand out. Liusha Bay has historically been known for oyster farms that produce high-end “Nanzhu” pearls. In recent decades, excess sediment runoff, competition for space, and a series of natural disasters have threatened the industry, but news reports indicate that production of these pearls is recovering. The harvest of other types of shellfish and finfish is also common here.

According to Xiaochuan (Sean) Ma, a scientist at the Institute of Oceanology, Chinese Academy of Sciences, the color of the water might be the result of suspended sediment and a high concentration of phytoplankton—a result of the region’s ecological conditions and eutrophication of its waters. Ma noted that phytoplankton can be abundant around the equipment deployed by fishermen.

One fishing method traditional in some parts of China involves pairs of poles anchored in the seabed, with a net spanning the tens of meters between them. Ma speculated that the parallel lines off Liusha Bay could be from the water current passing by the poles of this type of fishing arrangement.

In contrast to the geometric patterns made by people, the paintbrush-like appearance of the water is a work of nature. Currents and tides move and mix the colorful water. “It is impressive how much power the longshore transport is showing,” said Peter Clift, a scientist at Louisiana State University.

Waters east of the peninsula also display complex patterns. In Leizhou Bay, sediment-laden river water contributes to the colorful swirls. Sediment from the Pearl River Delta, about 400 kilometers (250 miles) to the northeast, might also contribute.

NASA images by Norman Kuring/NASA’s Ocean Color Web, using Landsat data from the U.S. Geological Survey. Story by Kathryn Hansen with image interpretation by Xiaochuan Ma/Institute of Oceanology, Chinese Academy of Sciences, and Peter Clift/ Louisiana State University.

For more information, visit https://landsat.visibleearth.nasa.gov/view.php?id=148060.
Hello, my name is Jason Stoker and I am the incoming President of the American Society for Photogrammetry and Remote Sensing for 2021. I want to express how grateful I am to be able to serve this organization in this capacity. I am truly humbled to think that my name will be included on the same list as so many other amazing professionals who have led this Society before me. I promise you I will work hard to continue to build on the successes my predecessors have laid before me.

This last year has been a difficult one for all of us. As you can see, I am giving my address to you all virtually from my home in Fort Collins, Colorado. This is not how I expected to start my term as President when I first became Vice President, but these are the times we are now forced to adapt to. And like all of us, the Society has had to adapt and adjust to the current situation. I am happy to report that thanks to the hard work by so many in ASPRS, we are not only surviving but adapting, perhaps even evolving, to meet the moment we are faced with. I am very proud of the resilience this organization has shown over this past year.

As I start this journey as ASPRS president for the next year, I have begun to think of ASPRS like a community garden- a garden that was started almost 90 years ago. True pioneers in mapping and photogrammetry broke the ground, prepared the soil, and planted the seeds of a Professional Society- a Society that in all honesty looks vastly different than the one they started, but still carries many of the traditions and legacies forward. Over the years this garden has expanded and contracted, brought in new varieties of ideas and themes and people, and has had bouts of both health and struggle. In learning about these early pioneers I often wonder what leaders such as Colonel Birdseye would think about lidar, or mapping from a drone- something that people of that time probably could have imagined as science fiction, but not as something that could be performed by anyone willing to invest in the equipment, software and methods. I have spent some time studying the works of people who have previously tended this garden to prepare me for my tasks ahead. For example, there have been 18 Presidents of ASPRS before me that also were USGS employees, including our first President- a legacy I do not take lightly. One of my initiatives for the next year is to make the history of the Society more accessible via the web and other documents, because I feel the best way for us to tend this garden is by knowing all of the things that have previously gone in to it, the good and the bad. I encourage any of our emeritus members with stories and history to share to contact us, so we can capture memories of the Society that may be lost forever otherwise. I look forward to some of the stories we will hear this year in our virtual annual conference.

And as any gardener knows, diversity is critical to both current and future success. A monoculture never survives very long before it seems that you put more into it than you get out of it. Diversity of viewpoints, platforms, focus areas and techniques are critical to ensuring that the Society adapts and responds to the winds of change. Adaptation can even lead to evolution over time. Diversity of our members is so critical to ensuring that we are truly representing the larger population of both the United States and the world. We won’t solve the problems of our planet if we come from a narrow, constrained viewpoint. In this garden called ASPRS we must both foster new growth as well as prune the things that are limiting future growth. The Society has begun to take steps to address the issues with diversity of our membership by forming a Task Force and providing recommendations to the Board for moving forward. I know that I am obviously not the poster child for diversity, but I do promise to push any initiatives the Task Force recommends toward making the society more diverse, more representative, more inclusive, and as respectful as possible. Any suggestions for improvement are always welcome!

And as with any garden, success isn’t due to only how we tend to it- things outside of our control can affect our growth and productivity. The pandemic could have really have negatively impacted us, as one of our revenue streams has always been face-to-face conferences and meetings that allow us to network, share ideas, create partnerships, build friendships and provide a bolt of energy a couple times of year. Last year we were forced to pivot quickly and move our annual conference to a virtual one. And this effort was a great success- one that has opened our eyes to how we may be able to thrive in a future where virtual collaborations and education are more commonplace. There really is no replacement for the amount of energy and inspiration that is generated from being surrounded by such amazing people in person; however we have seen that by diversifying the ways we can connect with one another we can still advance ASPRS’ mission and even generate revenue. I hope to continue the expansion of virtual education and information sharing, as well as investigating options for revenue building and cost savings by having more of a virtual presence. While I am excited about this year’s annual virtual conference, I am definitely already looking forward to next year’s Geo Week, where the AEC Next Technology Expo & Conference, International Lidar Mapping Forum, and SPAR 3D Expo & Conference, are coming together in 2022 with partner events USIBD Annual Symposium and our own ASPRS Annual Conference.
My journey with ASPRS began over 20 years ago with Dr. Roger Hoffer, one of our Past Presidents and my M.S. Advisor- without whom I may have never become part of ASPRS. Thank you, Dr. Hoffer! I also want to expressly thank my immediate predecessors Jeff Lovin and Tommy Jordan, who have taken me under their wing and helped guide me. Finally, I want to personally thank Bobbi Lenczowski, Lorraine Amend, and Karen Shuckman, who have really helped me better understand how ASPRS operates. To our managing director, Karen Schuckman especially- thank you so much for all you do for us. Your love and passion for this society is contagious. I want to thank our board for their hard work and dedication, especially over this difficult year. Also, our Regions, Student Chapter and Early Career Professionals have been dedicated and inspiring. Thanks to our staff, including Matt Austin, Rae Kelley, and P&N, our association management firm, as well as Jesse Winch our executive director. As many of you are aware Jesse is retiring as Executive Director this year. Jesse, thank you so much for leadership and guidance over the years. You are leaving ASPRS in a very strong position, and I hope you can enjoy retirement knowing that you left us in a very healthy place. I know I am missing thanking many folks here, but I promise I will make it up to you over the next year.

I often wonder what our earliest members would think of this garden called ASPRS today. The Society is bearing fruit now that they could not have imagined would grow in their garden when they started it. Similarly, we may not see the benefits of actions we take or decisions we make now for years to come. One thing I do know though is that a garden does not thrive without people tending to it with knowledge, care and love. And this is truly a community garden- some just come and harvest from it, and that is wonderful. And as those of us who do devote time and energy tending to it are almost all volunteers, I hope you all will join me this year and for many years to come in tending and growing this garden in helping the Society thrive, grow, adapt, and evolve as we respond to changing times. Without your help we would not be where we are today, and without your continued contributions we cannot flourish and truly meet the moment. I look forward to seeing you all in person next year!

Jason Stoker, Ph.D., U.S. Geological Survey, 2021 ASPRS President

ASPRS BOARD OFFICERS AND MEMBERS

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. Amend, 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 – 2023
Chair: Joe Cantz
https://www.asprs.org/sustaining-members-council.html

Committee Chairs Council – 2022
Chair: Mike Zoltek
Early-Career Professionals Council – 2022
Chair: Melissa Martin
Region Officers Council – 2023
Chair: Demetrio Zourarakis
Student Advisory Council – 2022
Chair: Evan Vega
Vice Chair: Lauren McKinney
http://www.asprs.org/Students/Student-Advisory-Council.html

Technical Division Directors Council – 2022
Chair: Bandana Kar, Ph.D
Vice Chair: Denise G. Theunissen

TECHNICAL DIVISION OFFICERS
Geographic Information Systems Division – 2023
Director: Denise Theunissen
Assistant Director: Briana Williams
www.asprs.org/Divisions/GIS-Division.html

Lidar Division – 2022
Director: Josh Nimetz, CMS
Assistant Director: Ajit Sampath
www.asprs.org/Divisions/Lidar-Division.html

Photogrammetric Applications Division – 2022
Director: Kurt Rogers
Assistant Director: Benjamin Wilkinson
www.asprs.org/Divisions/Photogrammetric-Applications-Division.html

Primary Data Acquisition Division – 2023
Director: Greg Stensaas
Assistant Director: Srinivasa Dharmapuri
www.asprs.org/Divisions/Primary-Data-Acquisition-Division.html

Professional Practice Division – 2024
Director: Bill Swope
Assistant Director: Hope Morgan
www.asprs.org/Divisions/Professional-Practice-Division.html

Remote Sensing Applications Division – 2024
Director: Amr Abd-Ehrahman
Assistant Director: Tao Liu

Unmanned Autonomous Systems (UAS) – 2024
Director: Dan Hubert
Assistant Director: Jacob Lopez
https://www.asprs.org/divisions-committees/uas-division
The Executive Director’s Report was given during the 2021 ASPRS Annual Conference. The virtual conference was held during the week of March 29 to April 2, 2021.

2021 is my 23rd year with ASPRS and this will be my last conference since I plan to retire at the end of June.

I started working for ASPRS in October of 1998, not long after Jim Plasker was hired as Executive Director. The late Kim Tilley interviewed me and recommended me for the position of Program Manager focusing on Certification and Awards to complete the administrative staff. Jim, whose knowledge and leadership was inspirational, also assigned me to be liaison with the Board of Directors, the Division, Region, and Committee officers and, eventually, Office Manager. I worked closely with Jim and Kim on a great many projects including the retiring of our mortgage on the headquarters site and the restoration of the ASPRS Foundation in 2005. Jim subsequently appointed me to be Assistant Executive Director of the Foundation and that gave me the honor of working with the Board of Trustees over the years.

After Jim retired in 2013, we saw a couple of rough years after which I was asked by then President Becky Morton to serve as Acting Executive Director. At that point, I was intensely interested in helping to steer ASPRS in the direction of a stable and financially solvent organization, so I readily agreed to do all that I could to accomplish that.

If it were not for Karen Schuckman, newly installed Managing Director, none of the great successes we all have witnessed would have happened. As most of you know, Karen is an indefatigable whirlwind who never sleeps. While working with our Treasurer, Stewart Walker, she astutely managed our finances, efficiently organized our conferences including last year when we had to shift to a virtual platform, started our weekly zoom meetings which revolutionized our intra-society communication and so many other projects and initiatives.

I am also leaving with fond memories of working for over 20 years with Al Stevens

**The ASPRS Headquarters Staff**
- Matthew Austin, Graphic Designer, Digital Publications Manager
- Rae Kelley, Assistant Director, Publications
- Jesse Winch, Acting Executive Director
- Karen Schuckman, Managing Director

**Headquarters/Condo**
To date, I’ve shown the renovated space to the President of the Condo Homeowners Association that surrounds our physical location. She plans to inquire among their members if anyone is interested in leasing or buying the space.

I continue to work with our agent from Donohue for potential general listing purposes.

The Board still needs to decide the direction we want to go. I would recommend that we find another short-term tenant for a couple of years (like 1-3) that would give us time to clear out the other side and organize the photos, paper files, bound journals, ASPRS pubs, historic awards, etc.

**AWARDS**
At the recent Foundation meeting, Board of Trustees President Cliff Greve announced the launch of a new, fully endowed award focusing on government service.

**ANNUAL ELECTION**
This year we saw the highest percentage of members voting than any year in recent memory. Managing our list of Emeritus members and encouraging them to participate was an important change in the numbers. Thanks to Karen Schuckman for that. Also, frequent reminders to vote by P&N was helpful and sending ballots to new members who joined after the initial ballot was sent also contributed to the increased response.

**Membership**
Below are the current membership numbers for the year to date.

Some highlights include a significant decrease in the number of Emeritus members from last year due to a major investigation and clean-up of the records. Our total membership numbers are now more accurate as a result.

<table>
<thead>
<tr>
<th>Membership Type</th>
<th>2021</th>
<th>January</th>
<th>February</th>
<th>March</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional</td>
<td>824</td>
<td>797</td>
<td>869</td>
<td></td>
</tr>
<tr>
<td>Student</td>
<td>205</td>
<td>202</td>
<td>282</td>
<td></td>
</tr>
<tr>
<td>Unemployed Professional</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Executive</td>
<td>11</td>
<td>11</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Retired Professional</td>
<td>16</td>
<td>15</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Rising Star</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Startup Alley</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Sustaining</td>
<td>25</td>
<td>24</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Sustaining Member Contacts</td>
<td>71</td>
<td>70</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Emeritus Members</td>
<td>207</td>
<td>209</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1373</td>
<td>1342</td>
<td>1500</td>
<td></td>
</tr>
</tbody>
</table>

**Publications**

**Book Sales — Best Sellers List**
- Airborne Topographic Lidar Manual
- Landsat’s Enduring Legacy: Pioneering Global Land Observations from Space
- Manual of Photogrammetry - Sixth Edition
who was Chair of the Awards Committee until he retired last year, and Mike Renslow, Chair of the Evaluation for Certification Committee, who retired two years ago. We became good friends over the years, and I could not have asked for two more dedicated, professional colleagues to make my job an easy one. I also had the opportunity to meet and work with the late Sky Chamard, or Father Chamard, as Jim called him, founder of our certification program, and I made the acquaintance of a couple of the folks we now have memorial scholarships for – Abe Anson and John O. Behrens, to name two and some of the family members of others – Jack Liang, Ta Liang’s son and Bob and Lola Osborn, parents of Ken Osborn. And I have to add that the Awards and Certification Programs, respectively, are now in the very capable hands of Lindi Quackenbush and Mike Zoltek.

I got to play music with the renowned Photogrammetric Ramblers – Tommy Jordan, guitarist extraordinaire - and really a multi-instrumentalist – and Chris McGlone on the 5-string banjo. We brought the house down, so to speak, at the Student Advisory Council reception in Louisville in 2014! A couple of years later, we had a reunion and sing-along in Baltimore where we were joined by Priscilla Weeks and Karen on vocals and Cliff and Judy Greve on guitars.

It has also been a pleasure to work with such a stable and productive staff over the years and you are lucky to still have Rae Kelley and Matthew Austin at headquarters helping the good folks at P&N to keep things moving.

I have to say that I’m looking forward to retirement – at 78 –as I’m still very active in the Irish music world. A couple of years ago, I 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. I have been playing music with my younger kids and we’re working on a recording project which we will hopefully finish by the summer.

But, I will definitely miss ASPRS and all the dedicated, insightful, and hard-working member-volunteers who really make ASPRS the dynamic association that it is.

I thank you all for the good wishes that I have received and depart knowing ASPRS is on a great path to continued success.

Call if you need anything.

Jesse Winch, Acting Executive Director

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
When Out-of-the-Box Toolbars are Not Enough, Make Your Own

Every-day Esri ArcMap users, often involved in production-type GIS work, find themselves focusing on routine and systematic workflows throughout the life cycle of a project. As an every-day user, we can directly relate. One task will often take weeks or months of focus and only requires the same few ArcMap tools which may be scattered among multiple toolbars. For example, when “QCing” an auto-collected waterbody shapefile, you may routinely use tools from the Standard toolbar along with the Editor toolbar, Advanced Editing, Tools, Creative Features, LAS Dataset, Effects, and the Drawing toolbar. A working map document (MXD) starts to feel cluttered with so many docked toolbars, and the end-user wastes valuable time jumping between all these active toolbars.

To combat the fatigue that comes with repeatedly switching between the same few tools, consolidating multiple, diverse tools into one toolbar can provide a welcomed change to the workflow. Before starting a task that you know will be repetitive, just create a custom toolbar that can be revisited and switch between tools without hunting across the entire suite of active toolbars and drop-down menus.

Below is an example we created called “My Great Toolbar” with tools from nine (9) different toolbars! From left to right, tools are for panning up, left, right, down, and revisit a previous extent in the data frame (Pan/Zoom toolbar); cut a vertical cross section from an LAS Dataset (LAS Dataset toolbar); swipe between layers (Effects toolbar); annotate with rectangular polygons (Draw toolbar); select editable features and reshape editable features (Editor toolbar); Explode Multipart Features (Advanced Editing toolbar); delete selected features and undo edits (Standard toolbar); save edits (Editor toolbar); and Identify features (Tools toolbar).

To create your own custom toolbar in ArcMap, navigate to Customize > Customize Mode (Step 1 below). Within the Toolbars tab of the Customize window, select New..., then name your new toolbar (Step 2 below). The new toolbar will appear as a small, empty toolbar near the upper left of your data frame (shown below) or docked with your other toolbars in the toolbar dock.

continued on page 396
EarthArchive

The Earth Archive Virtual Congress

June 15th - 16th

An unprecedented scientific effort to create a digital twin of the entire surface of the Earth and everything on it before it’s too late.

Featuring Presentations From:

Chris Fisher
Founder and Director
The Earth Archive

Douglas Preston
Best Selling Author

Thomas Lovejoy
University Professor of Environmental Science & Policy

Keith Masback
CEO
USGIF 2008-2018

Barbara Ryan
Executive Director
WGIC

Supporting Organizations

COMPLIMENTARY REGISTRATION

www.EarthArchiveCongress.com
**QUESTION:**

**Question 1:** If an accurate digital elevation model (DEM) is already available, can frames collected without orientation angle information be accurately triangulated and processed at a lower overlap (20-30%) than the normal Structure from Motion (SfM) program recommendations of 75% overlap or more? Please consider the systems to be metric and internally calibrated with a post-processed, kinematic and “fixed” Global Navigation Satellite System (GNSS) solution.

**Question 2:** Is there a significant difference in determining the Exterior Orientation (EO) parameters from SfM-based programs as compared to traditional approaches? Does having an accurate internal calibration when beginning processing make much difference?

Nathan Eick, Flight Operations Manager, Aerial Services, Inc. (ASI)

“Experience has shown us that measured sensor orientation angles are not critical to the photogrammetric process if the acquired imagery is processed through a triangulation routine”

Dr. Abdullah, Question 1: Experience has shown us that measured sensor orientation angles are not critical to the photogrammetric process if the acquired imagery is processed through a triangulation routine. An inertial measurement unit (IMU) is used to measure sensor orientation. These measured orientation angles—whether roll, pitch, heading or omega, phi, kappa—are crucial to processing the imagery, if the intent is to bypass the imagery triangulation process by utilizing the direct orientation approach. Having said that, it is worth mentioning that although the use of IMU is not necessary, having measured sensor orientation angles can benefit the processing software. If used correctly with traditional photogrammetric or SfM-based software, these angles provide value during the initial estimation of tie, pass or key point selection. Now, you also asked about whether the excessive imagery overlap is necessary for processing the imagery if you have an accurate DEM to use for the orthorectification process. Increased overlap is beneficial to the two main imagery processing steps: the triangulation of the imagery to determine the accurate geometry of the imagery block, i.e. network; and, most importantly, to the production of clean and reliable point clouds. If you are not concerned about the quality of the point cloud, then you can decrease the amount of overlap between the imagery given the figures you provided, assuming you meant this overlap ratio is between imagery taken from adjacent flight lines—or what we traditionally call “sidelap.” You do not need to decrease the overlap ratio between imagery taken along a flight line, also known as “forwardlap,” as it does not save on the acquisition time and therefore does not affect airplane fuel efficiency. Increasing the amount of forwardlap results in more images but, considering the digital workflow everyone is using today in processing imagery, this added number of images has very little impact on the manual labor budget of a project and it only increases computer processing time. The traditional photogrammetric process uses 60% forwardlap and 30% sidelap. Based on my experience, this reduced overlap also works in SfM-based software if the intent is not to produce refined point clouds. A forwardlap ratio of 60% or more is always needed if the imagery will be used for stereo map compilation. Therefore, it is a good planning strategy to assume that the imagery may eventually be used to set up stereo pairs for map compilation and not just for the production of orthos, since you never know when the need for stereo compilation may arise.

“SfM-based software provides more flexibility in dealing with imagery that does not have absolute georeferencing information to start with, as compared to the traditional photogrammetric approach where you need at least an approximate camera position and orientation.”

**Photogrammetric Engineering & Remote Sensing**

Vol. 87, No. 6, June 2021, pp. 395–396. 0099-1112/21/395–396

© 2021 American Society for Photogrammetry and Remote Sensing
doi: 10.14358/PERS.87.6.395
“all software needs to accurately determine the orientations and positions of images relative to each other. Once that is done, the collinearity condition is used to intersect rays to generate point clouds”

Dr. Abdullah, Question 2: The thing we need to understand here is that SfM is a matching technique that is more efficient than the typical auto-correlation we used in photogrammetry. Whether traditional photogrammetric or SfM-based, all software needs to establish image geometry at the instant of exposure. In other words, all software needs to accurately determine the orientations and positions of images relative to each other. Once that is done, the collinearity condition is used to intersect rays to generate point clouds. Besides its efficient image matching technique, SfM-based software provides more flexibility in dealing with imagery that does not have absolute georeferencing information to start with, as compared to the traditional photogrammetric approach where you need at least an approximate camera position and orientation. Having accurate internal camera parameters are beneficial, but not absolutely required, for both approaches—especially if you are going to process the data through the standard workflow. Having a known geometry of internal camera parameters helps the software in

“Even if you have a metric photogrammetric camera that provides an accurate calibration report, you will find the software alters the calibrated camera parameters slightly to give you the best product from the imagery.”

practice to provide software with accurate internal camera parameters, but it is not absolutely necessary if the software has the capability to refine these parameters through a process that is referred to in traditional photogrammetric practices as “camera self-calibration.”

**Dr. Abdullah is Vice President and Chief Scientist at Woolpert, Inc. He is also adjunct professor at Penn State and the University of Maryland Baltimore County. Dr. Abdullah is ASPRS fellow and the recipient of the ASPRS Life Time Achievement Award and the Fairchild Photogrammetric Award.**

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, Woolpert, Inc., NOAA Hydrographic Services Review Panel (HSRP), Penn State, and/or the University of Maryland Baltimore County.

GIS Tips & Tricks, continued from page 393

Step 3 - In the Customize window, select the Commands tab. Search for different tools using the ‘Show commands containing:’ search box or by navigating through the Categories list. Then, drag-and-drop your favorite tools into the new, blank toolbar.

When the “Customize” mode is active, you can add new tools to a toolbar or rearrange the icons by simply dragging them to a new position. To delete an icon (tool) just drag it away from the toolbar and release it. Of course, the icon is still available to be chosen from the Commands menu and your new toolbar will be saved with your map document. It is really that easy to make your own custom toolbars.

Toolbar customizations are unlimited in diversity. Please feel free to share yours with us. Send your questions, comments, and tips to GISTT@ASPRS.org.

Kristopher Gallagher is a Senior Lidar Analyst focused on the production of digital mapping products derived from topographic lidar and utility surveys. As senior geospatial scientist, Al Karlin works with all aspects of Lidar, remote sensing, photogrammetry, and GIS-related projects.
GIS Tutorial for Crime Analysis, Second Edition, is based on Esri’s ArcGIS Desktop software and is a resource intended to familiarize anyone using or intending to use GIS for crime analysis through hands on software-based learning. The tutorials are geared toward use in contemporary police organizations with direct crime mapping applications. This resource provides immediate opportunity for analysis throughout the tutorials by replacing example data with existing department organization data.

The book is organized with well-defined chapter topics which progress and build upon previously learned concepts from beginning to end. This organizational structure makes it well-suited for the self-directed learner, and a resource for those already familiar with GIS who are interested in a specific application. It is also organized for class use with introduction to important GIS mapping concepts, tutorials, and assignments at the end of chapters.

The book consists of nine chapters and has a total of 348 pages. The first chapter serves as an introduction to how GIS can be utilized in police work. The second chapter is a tutorial-based introduction to ArcGIS Desktop software and the user interface. The subsequent chapters (3 through 9) are all tutorial-based with a focus on ArcGIS and ArcMap. The tutorials are well structured with easy to follow, step by step instructions for completing chapter projects. Full color images of the software interface provide clear examples and instruction for how to complete each task. After the software tutorials, assignments are presented for gaining additional familiarity with both the concepts presented and the software. The proceeding tutorials become an excellent reference when completing these assignments. Chapter 9 concludes the book with hot spot modeling and retrospective prediction analysis. Alternative scopes are also provided in this chapter for more customized application and instructor course tailoring.

GIS Tutorial for Crime Analysis, Second Edition, is a useful resource for individuals and departments looking to gain understanding in environmental criminology. It offers foundational information geared toward those with minimal to no GIS experience, and the progressive learning structure and assignment suggestions also makes it well suited for instructors. Utilizing crime mapping tools as described in this book, modern policing can see benefit through greater understanding of the spatial distribution of crimes and proximity analysis.

GIS Tutorial for Crime Analysis (2nd Ed.)
Wilpen L. Gorr, Kristen S. Kurland, Zan M. Dodson
Reviewed by Toby M. Terpstra, Instructor, Society of Automotive Engineers (SAE) and Principal Forensic Animator, Kineticorp LLC., Greenwood Village, Colorado.
GREETINGS FROM THE ASPRS STUDENT ADVISORY COUNCIL (SAC)!

As we network within our ASPRS Chapter community, we are taking this month to highlight student and K-14 educational resources on the ASPRS Website.

### ASPRS Community Forums

ASPRS hosts over 20 community forums related to events and activities, professional opportunities, volunteer opportunities, technology, articles, and more. Each forum is managed and please be mindful that posted comments should be appropriate and respectful. You must log in to the ASPRS Website. A few forums are listed below.

<table>
<thead>
<tr>
<th>Forum</th>
<th>Committee/Division</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASPRS Jobs</td>
<td>Education and Professional Development Committee</td>
</tr>
<tr>
<td>ASPRS Community</td>
<td>Technical Division — Primary Data Acquisition</td>
</tr>
<tr>
<td>Florida Region Forum</td>
<td>Technical Division — Geographic Information Systems</td>
</tr>
<tr>
<td>Standards Committee Community</td>
<td>Technical Division — Lidar</td>
</tr>
</tbody>
</table>

### Education, Career Information & K-14 Outreach

<table>
<thead>
<tr>
<th>Resource</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>GeoBytes</td>
<td><a href="https://www.asprs.org/geobytes.html">https://www.asprs.org/geobytes.html</a></td>
</tr>
<tr>
<td>GeoBits for Kids</td>
<td><a href="https://www.asprs.org/geobits.html">https://www.asprs.org/geobits.html</a></td>
</tr>
<tr>
<td>On-Demand Workshops</td>
<td><a href="https://asprs.prolearn.io/home">https://asprs.prolearn.io/home</a></td>
</tr>
</tbody>
</table>

### Events, Seminars & Social Media

<table>
<thead>
<tr>
<th>Event</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>Events &amp; Seminars</td>
<td><a href="https://www.asprs.org/ASPRS-Events.html">https://www.asprs.org/ASPRS-Events.html</a></td>
</tr>
<tr>
<td>Facebook</td>
<td><a href="https://www.facebook.com/SacAsprs">https://www.facebook.com/SacAsprs</a></td>
</tr>
<tr>
<td>LinkedIn</td>
<td><a href="https://www.linkedin.com/in/asprs-sac-2310131b1">https://www.linkedin.com/in/asprs-sac-2310131b1</a></td>
</tr>
<tr>
<td>Twitter</td>
<td>@SacAsprs</td>
</tr>
</tbody>
</table>
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 Kenya was originally printed in 2003 but contains updates to their coordinate system since then.

The coast of Kenya was long dominated by Arabs and was seized in the 16th century by the Portuguese. The Europeans were expelled by the Omani; the coast then came under the rule of the Sultan of Zanzibar, and later leased in 1867 to the British East Africa Association. The British extended their holdings into the interior and fixed an initial southern boundary with the German East Africa Company in 1886. Kenya is bordered on the north by Ethiopia (PE&RS, March 2003) (861 km), on the east by Somalia (682 km), on the southeast by the Indian Ocean (536 km), on the south by Tanzania (769 km), on the west by Uganda (933 km), and on the northwest by Sudan (232 km). The lowest point in Kenya is the Indian Ocean, and the highest point is Mount Kenya at 5,199 m. Comprised of the Nairobi Area and seven provinces — Central, Coast, Eastern, North Eastern, Nyanza, Rift Valley, and Western —Kenya is slightly larger than twice the size of Nevada. The former Colony of British East Africa gained independence on 12 December 1963.

Thanks to Morgan W. Davis, “The history of surveying in East Africa begins with the domination of those lands by European powers in the late 1800s. The British entered into a number of agreements defining spheres of influence in 1886, 1890, 1891, and 1894. The present day boundaries between Kenya and its neighbors are a result of legal descriptions hashed out in negotiations and subsequent triangulation and boundary surveys. Negotiations between the United Kingdom and Germany in 1886 and 1890 established spheres of influence north and south respectively, of a line beginning at the Indian Ocean near Vanga and extending to the eastern shore of Lake Victoria: ‘The line of demarcation starts from the mouth of the River Wanga or Umbe (Umba), runs direct to Lake Jipe, passes thence along the eastern side and round the northern side of the lake and crosses the Lumi River ...’

“After which it passes midway between the territories of Taveita and Chaga, skirts the northern base of the Kilimanjaro range, and thence is drawn direct to the point on the eastern side of Lake Victoria Nyaza (Lake Victoria) which is intersected by the 1st degree of south latitude.”

The line between the Indian Ocean and Lake Jipe was surveyed by plane table. Most of the mapping in East Africa between 1890 and 1910 was a result of the boundary commissions. Basic topographic mapping of varying quality was accomplished along the narrow zones of the surveyed boundaries. There was little opportunity to extend mapping to the
interiors of the colonies. An important outcome of this early phase was the consolidation of the War Office as the authority on boundary surveys and maps in Africa. Both the Foreign Office and the Colonial Office relied heavily on the expertise of the War Office on technical matters related to surveying and mapping, as well as for help in wording legal descriptions in negotiations. The Topographic Section of the General Staff of the War Office played a crucial role, as well in the policies and activities of survey departments in the colonies.

The colonial Survey Committee was created in an attempt to organize the mapping effort in the British East African colonies. The first meeting was held on 14 August 1905. They recommended that there be two survey departments, standardized topographic map scales at 1:62,500, 1:125,000, 1:250,000, and 1:1,000,000. In a 1907 meeting, they adopted the Clarke 1858 ellipsoid for Africa. They decided on the spelling of place names on maps. The Committee continued to be an important governing body up to the World War II years. Major E. H. Ellis was appointed Inspecting Officer to the departments in the Uganda and East Africa Protectorates (Kenya) in order to help expedite work. He submitted a comprehensive report in February 1907 in which he noted that a topographic section had not been constituted. He insisted that a full section of 2 officers and 6-8 surveyors be formed. He recommended map sheets covering 45° longitude and 30° latitude or 30° x 30°, at 1:125,000 scale for developed areas, and 1½° longitude by 1° latitude in undeveloped areas, utilizing the rectangular polyconic projection. *(This was the same specification utilized during the same era by the British Survey of India. – CJM)*

In late 1908 one officer, three NCOs and a civilian were assembled to begin fieldwork on 1:125,000 sheets for Kijabe and Nairobi, and a special 1:62,500 sheet for Nairobi. Mapping continued until the outbreak of WWI. The Africa Series GSGS 1764 in 33 sheets at 1:250,000 scale covered both Uganda and the East Africa Protectorates. The maps were published in monochrome, principally between 1905 and 1907. These were provisional sheets with a paucity of detail. The maps were later also published in color. Each sheet covered 1½° longitude and 1° latitude with a graticule spacing of 30'. They were reprinted in 1939-1941 during the East African Campaign. It was not until 1953 and thereafter that Series GSGS 1764 was replaced at the same scale by Series GSGS 4801 and subsequently Series Y503.

After WWI, the War Office was no longer available to do work in the African colonies. German East Africa had been assigned to Great Britain as a mandate from the League of Nations in 1919 and was renamed Tanganyika. The Arc of the 30th meridian was proposed as the foundation of triangulation in the East African colonies. Observations on a portion of the arc in western Uganda had been taken prior to 1914, and the triangulation net in Uganda was tied to it. Surveying on the arc had been done in northern Rhodesia, and it was felt that it was important to close the gap in the arc in Tanganyika.

Martin Hotine surveyed the arc of the 30th meridian in Tanganyika between 4½° and 9° South during the years 1931-1933. Depletion of funds in late 1933 left a gap in the arc between 1½° and 4½° South. From July 1936 to August 1937, a survey was conducted wholly within Tanganyika to fill the gap, consisting of observation angles and some azimuths. Uganda had withdrawn from the project due to fears that if their portion of the arc was connected to South Africa, they would be forced to recompute their already completed surveys on a new projection and grid system.

This leads to a major theme of discussion during the years between the two great wars – that of a common datum and projection for all of British Africa. Debate raged over this topic until the exigencies of war during the Second World War permitted the British military to force a solution. In a memorandum circulated in 1926, it was assumed that a common datum could be chosen, utilizing a meridional orthomorphic projection from Khartoum to Cape Town. The Clarke 1880 ellipsoid was suggested. During the second Conference of Empire Survey Officers (1931), it was assumed that all colonial governments would adopt the Transverse Mercator projection because it was already accepted by Egypt, South Africa and two of the West African territories. The width of the zones could not be agreed upon. Kenya saw little prospect in adopting the proposal because its cadastral work was computed on Clarke 1858 and the Cassini projection. Extension and re-computation of its triangulation was more urgent than conversion of its completed surveys to a new datum. In January 1934, GSGS proposed a coordinated projection and grid zone embracing South Africa, South Rhodesia, Sudan, Egypt, and the Central and East African territories. They recommended the Clarke 1880 ellipsoid and the Transverse Mercator projection on a 6-degree grid. The same parameters were recommended in a meeting of a sub-committee of the Colonial Survey Committee on 3 October 1935. Brigadier M. N. MacLeod insisted on the adoption of the meter as the map unit. Each time a new recommendation would be put forward for a common set of map parameters, one or more colonial governing bodies would shoot it down for various reasons.

Lord Hailey wrote *An African Survey* (1938) after his tour of Africa in 1935. His views were taken up by the Colonial Survey Committee in 1939, at which time they once again recommended a 6-degree grid and the adoption of the meter. Whittingdale replied that a 2-degree system was more appropriate for topographic mapping and military surveys. Huntley showed the military advantages of the 2-degree grid (artillery), and that it was inconvenient for cadastral surveyors to apply the corrections that a 6-degree grid would necessitate. *(The same reason for practicality continues to this day for NOT using the UTM grid for civil GIS and surveying applications. – CJM)* South Africa totally opposed the change from 2-degree to 6-degree zones. There was general agreement on the adoption of the meter on map grids.

A policy for military mapping was defined in July 1940, which utilized the Clarke 1880 ellipsoid and the Transverse Mercator projection with 5-degree zones. The central meridians were placed at 32° 30' E, 37° 30' E, and 42° 30' E. A scale
factor reduction of 0.05% was introduced to provide correct scale on two parallel meridians approximately 1° 49’ on either side of the central meridian. The scale error at the central meridian was about 1:2,000, and it was about 1:2,200 at the edges of the grid zones. The grid was originally in yards, but was later changed to meters. This became known as the East African War System, and it was eventually applied to an area bounded by 19° N, 15° E, 12° S and the Indian Ocean. The Directorate of Colonial Surveys was born on 1 March 1946, with Brigadier Martine Hotine as it first (and only) Director. An allowance of £2 million was approved for this centralized organization of geodetic and topographic surveys. For the first time in the eastern African colonies, two problems, which had plagued the survey effort from the earliest days, were addressed: lack of funds and the lack of a centralized organizing body. In 1947, fieldwork for basic topographic mapping was commenced. The first 1:50,000 scale sheets of Series Y731 were produced for the Kenya Ethiopia Boundary Commission. At least 470 sheets were produced, virtually all of which were contoured, and 64 sheets along the Ethiopian border.

The 1:100,000 scale sheets of Series Y633 were produced between 1958 and 1968, mainly by the Survey of Kenya and the Directorate of Military Survey. A general map series at 1:250,000 (Y503) has been derived from the 1:50,000 and 1:100,000 scale sheets. The Survey of Kenya produced 42 of the 50 sheets needed to cover the country. Kenya is covered by 7 sheets of the 1:1,000,000 International Map of the World. The most commonly used geodetic parameters for maps produced by the Kenyan authorities are: Arc Datum 1960 referenced to the Clarke 1880 (modified) ellipsoid, Transverse Mercator projection with coordinates on the UTM grid.

In the 1970s, first-order EDM traverses were run between stations adjusted on the Arc 1960 Datum and Clifford triangulation. In 1972-1973 the Survey of Kenya, in conjunction with the U.S. Defense Mapping Agency and the Directorate of Military Survey of the U.K. made the first experimental Doppler satellite survey in Kenya. Recently the Kenya Institute of Surveying and Mapping (KISM) took GPS observations on existing control points. A. S. Lwangasi of the University of Nairobi reported the results of a datum transformation carried out on 25 control points from Arc 1960 Datum to WGS 84 Datum: \( \Delta X = -179.1 \text{ m} \pm 0.7 \text{ m}, \Delta Y = -44.7 \text{ m} \pm 0.7 \text{ m}, \Delta Z = -302.6 \text{ m} \pm 2.2 \text{ m}. \)

In a letter dated 20th July 1989, J. R. R. Aganyo wrote for the Director of Surveys of the Survey of Kenya that the old Cassini-Soldner used in Kenya has the following parameters: grid name: Cassini-Soldner; years used: since introduction of cadastral surveys; central meridians: 33°, 35°, 37°, 39° East; unit: English Foot where 1 foot = 0.30480 International metres, exactly; ellipsoid: Clarke 1858 where \( a = 20,926,348 \text{ English feet and} \frac{1}{f} = 294.26. \)

In January 2000, Russell Fox of the U.K. Ordnance Survey sent a memo to me that was written by the famous H. F. Rainsford on 28 September 1961 with the (then) Directorate of Overseas Surveys:

“Since there appears to be some confusion of thought about the ‘origin’ of the trigonometric data lists produced by this Directorate, the purpose of this paper is to clarify the position so far as possible.

Up to the present date, all trig. Lists have included in the preamble the words – ‘New 1950 Arc Datum.’ This denotes that the results in the list are based on the Arc of the 30th meridian, which was computed by the D.O.S., from South Rhodesia to Uganda, in the 1950 (circa). The values of the stations accepted as a starting point in South Rhodesia had been computed continuously from South Africa. These Arc results have been held fixed since 1950, and it is hoped that they will remain so for as long as possible in the future, since they are used not only by the D.O.S., but also by the Congo and Portuguese Africa, and they provide a uniform system from the Cape to the Equator.

The South African datum is an arbitrary one, as at no station were the Astronomic and Geodetic latitude, longitude and Azimuth made coincident. On the Arc itself the (A-G) values vary (sometimes quite abruptly) between:

- latitude \( +20^\circ \text{ and} -30^\circ \)
- longitude \( +12^\circ \text{ and} -10^\circ \)
- azimuth \( +15^\circ \text{ and} -08^\circ \)

The only astronomic elements that have been held fixed on the Arc adjustment are – in South Africa one latitude, longitude and azimuth (but each at a different station) and an astronomic azimuth at Kichareere in Uganda, just south of the Equator.

The Year 1950 was used in the title as a convenient epoch mainly to distinguish from previous systems such as the ‘1935 Arc Datum.’ (original emphasis in color)

Tanganyika was the first East African territory in which geodetic trig. control was computed based on the Arc and used for control of topographic surveys. It was known that some of this trig. was not up to primary standards, but it was the only work available and it was hoped that recomputation based on the Arc would produce results of sufficient accuracy for the purpose required.

Since Laplace Azimuths had not been available for the Arc computation nor in Tanganyika, the Tanganyika trig. was computed without holding fixed any azimuths, which were, in any case, of doubtful value. When the trig. computation reached Malindi in Kenya from the Arc it was found that the (A-G) azimuth was approximately 20°.

It was then decided that a new approach was necessary. Put in new primary circuits based on the Arc, and observe frequent astronomical stations and tellurometer lengths, much closer together than the old measured bases. The trig. circuits (were) to be adjusted to the fixed (or nearly fixed) scale and azimuth checks. This policy has been carried out and results have already been circulated for:

The Lake Circuit
Uganda Primary
Kenya Primary
Dear Mr Mugnier,

I found your name when I came across your 2004 article on Iceland’s Grids and Datums. I’ve been researching the British invasion and subsequent occupation of Iceland in May 1940, more as a hobby, although with retirement approaching, this might turn into something more substantive.

Having come across your article, and noting that in the case of Iceland there were two datums, I wondered what mapping was available to the British in the lead up to the invasion. My expectation is that the British would have access to the nautical charts required for the maritime aspects of the operation, and probably, street maps for Reykjavík and Akureyri.

Whilst these would have met the needs for planning and meeting immediate objectives, as the troops moved into the hinterland there would have been need for something more detailed. Unfortunately, the plans of the initial operation are only in hard copy and held at the public records office in Kew, London. Unfortunately, due to the COVID lockdowns, this office is closed. Moreover, online research has only shown a 1:500,000 topographical (Landslagsuppdrattur) map produced in 1928 by a Samuel Eggertsson an Icelandic surveyor (1864 - 1949). Eggertsson’s work would at least suggest some mapping was available, but once again online searches has produced scant results.

So my question is, through your research are you aware of what mapping coverage was available pre May 1940? Also, if mapping was available are you aware of any German involvement in surveying the country or would this have been solely an Icelandic/Danish effort?

Thank you in advance for any information or insight you can provide

~ J M Lee BEM, CSMP, CBCI. MSyI
Doran Risk Consulting Ltd

As of 20 December 1946, a cartometric analysis of Islandic source material by William W. Baird of Army Map Service stated:

SUBJECT: Geodetic investigation of source material for AMS 1:50,000 series of Iceland (40045-3, 4, 4)

1. Map series investigated:
   A. S30-DGS-50
   B. S30-GSGS-50 (Grid Shown)
   C. 3-30-37003-100
   D. S30-GI-100
   E. S30-GSGS-100
   F. S30-AMS-100

2. 1:50,000
   The two sets of 1:50,000’s have only partial coverage of Iceland. The 50’s are quarters of the 100’s - detail, features and elevations agree. Both sets of 50’s carry a projection based on Copenhagen. To convert to Greenwich a correction of 12° 34’ 40.35” must be applied to the meridians.
   A. S30-DGS-50
   The mean differences (control-scaled) were Longitude 1.6” and Latitude 1.0”. Maximum differences were Longitude from +12.0” to −2.2” and Latitude from +2.3” to −0.7”.
   B. S30-GSGS-50
   The mean differences (control-scaled) were Latitude 1.4” and Longitude 1.4”. Maximum differences were Latitude from −2.8” to +0.3” and Longitude from −3.3” to +0.3”.
   At the scale of 1:50,000, one second of Latitude equals .024 inches and one second of Longitude equals .010 inches.

3. 1:100,000
   The 1:100,000 series has almost complete coverage of Iceland and on most of the sheets the projection is based on Greenwich, but a few sheets carry meridians based on Copenhagen. To convert to Greenwich, use the same correction as was given for the 50’s.
   A. 3-30-37003-100 and S30-GI-100 are original Danish map work. S30-GSGS-100 is a British G.S.G.S. War Office series and S30-AMS-100 is an A.M.S. redraft of the G.S.G.S. series. All four series seem equally reliable for most areas, but the original Danish maps have a few spots which do not agree with the control values by several seconds, mostly in the vicinity of Akureyri.
   B. About fifty points were scaled from the various 1:100,000 series and the mean differences (control-scaled) were Longitude 3.7” and Latitude 1.5”.
All these results have been headed, as before, ‘New 1950 Arc Datum’, because the fundamental datum, which is the Arc, has not been changed. Whenever the new coordinates differ from the previous, this is due to a re-computation (including new observations) of part of the trig. system.

To avoid any further misunderstanding in the future it is proposed to change the heading of trig. lists now to ‘New 1960 Arc Datum.’ Most of the Tanganyika main trig. has still to be recomputed and a letter will be sent to each territory indicating the particular trig. chains which have already been recomputed and circulated under the 1950 heading.

The Figure of the Earth used is the Modified Clarke 1880, for which \( a = 6378249.145 \) and \( r = 293.465 \) in International Metres. The geodetic tables used are Latitude Functions Clarke 1880 Spheroid, Army Map Service, but most D.O.S. computations are now done on the Electronic Computer, which computes its own geodetic factors ‘ab initio’. Coordinates are also produced on the U.T.M. projection.”

Thanks go to Washington Abuto wherein his letter of 24 November 1997 for the Director of Surveys of the Survey of Kenya enclosed a paper detailing much of Kenya’s history of Grids and Datums. That paper, authored by Mahinda, served as the basis of much of the specific geodetic history quoted in Davis’ graduate-level term paper of 1999.

Kenya Update

AFREF Newsletter No. 5, (2008) reports a Molodensky-Badekas 7-parameter transformation solution for Kenya based on twenty common points between WGS84 and Arc 1960 Datum. Although the parameters are listed, no coordinates are listed for the Arc60 Datum origin point necessary for a Molodensky-Badekas model so the model is likely Bursa-Wolfe, instead. Furthermore, no guidance is provided regarding which direction the parameters are intended to be used, nor are any test points provided. https://www.rcmrd.org/newsletters

AFREF Newsletter No. 11, (2010) reports on the KENREF (Kenya Reference Frame) proposal that will include CORS at 21 primary stations and eventually 71 stations at secondary locations. https://www.rcmrd.org/newsletters


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 (C4G).

This column was previously published in PE&RS.

Ad Index

National Reconnaissance Office | www.nro.gov/Business-Innovation-Opportunities/ | Cover 2
Earth Archive | www.EarthArchiveCongress.com | 394
NEW ASPRS MEMBERS

ASPRS would like to welcome the following new members!

Alaska
Bryan O’Malley, CP

At Large
Erik Birzulis

Casadia
Malcolm Baker
Matthew T. Kumpula, PLS
Natalie Martinus
Sean Pickner

Eastern Great Lakes
Mark Alan Ober, PLS

Florida
Eric J. Andelin, CP
Todd H. Bates, PLS

Gulf South
Hugh Bartlett
Wallace Scott Dunham, CP, CMS
Henry B. O’Steen, II
Russell White

Heartland
Aparajithan Sampath, PhD

Mid South
Nancy E. Graham
Ryan Griffin, CMS
James Maxwell Norton
Rory Smith
Michele Marie Thornton
Rick Yeatts

Northeastern
Christina Belton
Andrew Stott

Pacific Southwest
Saif AATI
Marian Cosmin Bolata
Eduardo Contreras
Flora Maria Delgado
Albert Hong
Dar A. Roberts
Steve Schuetz
Sheng Tan
Dan Taranik
Michael Tartaglia

Potomac
Joseph Jones
William McAuliffe

Rocky Mountain
Claire M. Hay, PhD
Kyle Knapp
Robert Carleton White, GISP

Western Great Lakes
Mr. Paul Crocker
Ben Leonard
Luke Peterson
Stephanie Lynn Rockwood

FOR MORE INFORMATION ON ASPRS MEMBERSHIP,
VISIT HTTP://WWW.ASPRS.ORG/JOIN-NOW

Your Path To Success In The Geospatial Community

Who at ASPRS Do I Contact for help with…

Membership/PE&RS Subscription
Shalini Gogawale — office@asprs.org

PE&RS Editor-in-Chief
Alper Yilmaz — PERSeditor@asprs.org

Advertising/Exhibit Sales
Bill Spilman —
bill@innovativemediasolutions.com

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
Spatial Resolution Enhancement for Large-Scale Land Cover Mapping via Weakly Supervised Deep Learning

Qiutong Yu, Wei Liu, Wesley Nunes Gonçalves, José Marcato Junior, and Jonathan Li

Abstract
Multispectral satellite imagery is the primary data source for monitoring land cover change and characterizing land cover globally. However, the consistency of land cover monitoring is limited by the spatial and temporal resolutions of the acquired satellite images. The public availability of daily high-resolution images is still scarce. This paper aims to fill this gap by proposing a novel spatiotemporal fusion method to enhance daily low spatial resolution land cover mapping using a weakly supervised deep convolutional neural network. We merge Sentinel images and moderate resolution imaging spectroradiometer (MODIS)-derived thematic land cover maps under the application background of massive remote sensing data and the large spatial resolution gaps between MODIS data and Sentinel images. The neural network training was conducted on the public data set SEN12MS, while the validation and testing used ground truth data from the 2020 IEEE Geoscience and Remote Sensing Society data fusion contest. The proposed data fusion method shows that the synthesized land cover map has significantly higher spatial resolution than the corresponding MODIS-derived land cover map. The ensemble approach can be implemented for generating high-resolution time series of satellite images by fusing fine images from Sentinel-1 and -2 and daily coarse images from MODIS.

Introduction
Remotely sensed satellite imagery is the primary data source for monitoring land cover change and characterizing land cover on a global scale (Song et al. 2017). Satellite images with daily coverage and fine spatial resolution are highly desired for Earth observation and related environmental applications (Sun and Zhang 2019). However, the consistency of daily land cover monitoring is often constrained by the spatial and temporal resolutions of the acquired satellite images freely available. For instance, Landsat satellites capture images with a moderate spatial resolution of 30 meters but with a long revisit period of 16 days. On the contrary, the moderate resolution imaging spectroradiometer (MODIS) can provide images daily, with coarser spatial resolutions of 250 m, 500 m, and 1 km. Hence, it is important to understand how to jointly leverage complementary data sources efficiently to conduct land cover classification. To have up-to-date land cover monitoring with a fine spatial scale, increasing the spatial resolution of coarse satellite imagery represents a continued advancement in remote sensing research. The availability of the MODIS data set has driven global-scale land cover mapping at coarse resolution. Previous works have conducted spatiotemporal fusion to blend MODIS and Landsat data to obtain improved classification results with a higher spatial resolution of 30 m (Gevaert and García-Haro 2015; Wang et al. 2015; Chen et al. 2017). Sentinel-1 and Sentinel-2 today can provide higher temporal resolution (three to five days) and higher spatial resolution (10 to 20 m) than Landsat satellites. However, these images are frequently unavailable for land cover mapping due to the presence of clouds. Hence, it is necessary to develop a feasible method to integrate remote sensing data from different sensors and time phases to acquire geospatial data with high spatial and temporal resolutions.

Recently, deep learning frameworks have enhanced the classification performance by automatic extraction of in-depth features. Therefore, deep learning-based land cover classification has become a current hotspot in the remote sensing research community. One of the significant advantages of using deep learning algorithms is that it is a learning-based method, which automatically learns an end-to-end mapping between coarse resolution images and fine resolution images. Previous research indicates that semantic segmentation classification with deep learning methods at the pixel level is promising in land cover mapping (Huang et al. 2018; Kemker et al. 2018).

To the best of our knowledge, no deep learning-based model has yet been introduced to conduct spatiotemporal fusion to blend MODIS data and Sentinel satellite images. The novelty is emphasized by the proposal of a weakly supervised approach. With the aim of providing enhanced land cover mapping through the fusion of multi-source satellite data, this paper extends one of the current state-of-the-art semantic segmentation networks, DeepLabV3+ (Chen et al. 2018), and then employ it to enhance the spatial resolution of MODIS-derived land cover maps, by integrating the maps (with an original spatial resolution of 500 m), synthetic-aperture radar (SAR) images derived from Sentinel-1, and multispectral images derived from Sentinel-2. The outputs of the model are high-resolution (10 m) land cover thematic maps. Technically, this is a task of supervised semantic segmentation of the Sentinel images since the MODIS maps are utilized as the target ground truth labels, and the model assigns one of the label classes to each pixel in the Sentinel images. However, due to the coarse resolution of MODIS maps, the Sentinel images only contain partial observations of the target ground truth labels, which makes the task become a weakly supervised semantic segmentation. To deal with weakly annotated ground truth labels, an

© 2021 American Society for Photogrammetry and Remote Sensing
doi: 10.14358/PERS.87.6.405

PHOTOGRAFMETRIC ENGINEERING & REMOTE SENSING
Vol. 87, No. 6, June 2021, pp. 405–412.
0099-1112/21/405–412
additional module was embedded in the model, and it automatically updates the coarse labels based on the intermediate predictions on the training sets. The contributions of this paper can be summarized as follows.

First, a deep learning-based method was developed for an effective fusion of the MODIS and Sentinel data. Second, a deep learning semantic segmentation network, DeepLabV3+, was comprehensively evaluated given that the ground truth labels are noisy and unreliable. Third, more challenging land cover types can be classified using the proposed method.

The rest of this paper is organized as follows. Section “Related Works” reviews some previous studies regarding spatiotemporal fusion and the use of DeepLabV3+ on remote sensing images. Section “Method” introduces the technical information about DeepLabV3+ and the modifications added to its original architecture. The section “Experiments and Discussion” describes the experimental setup, including the implementation detail of our proposed model and the data set used in this study. Finally, the last section concludes the paper with remarks and expectations.

Related Works

Spatiotemporal Fusion of Remote Sensing Images

In the field of remote sensing, many key application domains stand to benefit from data fusion techniques. For example, the increase of spatial resolution contributes to land cover classification and ground object identification. Recently, many remote sensed data fusion methods have been proposed to deal with the specific problems that arise from the trade-off between spatial resolution and temporal frequency. In general, they can be categorized into image pair-based and spatial unmixing-based methods (Ghamisi et al. 2019).

The image pair-based method utilizes the relationship between the available coarse/fine image pairs to guide the prediction of fine images from coarse images on other days. Image pair-based methods can be further classified into a filter-based algorithm and learning-based algorithm (Song et al. 2018).

Among existing image pair-based spatiotemporal data fusion algorithms, the spatial and temporal adaptive reflectance fusion model (STARFM) (Gao et al. 2006) was the first model developed. It has been widely applied for fusing Landsat and MODIS to monitor environmental changes (Chen et al. 2015; Gevaert and García-Haro 2015). It uses one known pair of Landsat and MODIS images and one MODIS image at the prediction date. STARFM assumes that for a pure coarse pixel where only one land cover type exists, the changes in fine pixels within that coarse pixel can be implied directly by the coarse pixel changes. For heterogeneous coarse pixels with two or more land cover types, a weighted function is used for prediction, which assigns higher weights to the neighboring fine pixels where they are physically closer and spectrally similar to the coarse pixels (Gao et al. 2006).

Since STARFM assumes that the temporal changes of all land cover classes within a coarse pixel are consistent, it is thereby suitable for homogeneous landscapes (Ghamisi et al. 2019). However, it is sensitive to high heterogeneity and abrupt land cover changes (Sun and Zhang 2019). Subsequently, several algorithms have been developed to improve the accuracy of STARFM. For instance, Hilker et al. (2009) proposed the spatial-temporal adaptive algorithm for mapping reflectance change, designed to detect reflectance changes using Tasseled Cap transformations of both Landsat and MODIS data. Zhu et al. (2010) developed enhanced spatial and temporal adaptive reflectance fusion model to deal with heterogeneous landscapes specifically. It requires two coarse/fine image pairs to estimate the temporal change rate of each land cover class separately and assumes the change rates to be consistent. To summarize, these methods are different mainly in modeling the relationship between the paired pixels. These methods, including STARFM, can be considered filter-based methods because each pixel is predicted from a filtering model, a weighted sum of spectrally similar neighboring pixels from the input images (Song et al. 2018).

Recently, some learning-based spatiotemporal fusion algorithms have been proposed, such as support vector machine (Wang et al. 2018), Hopfield neural networks (Fung et al. 2019), and deep convolutional neural network (Song et al. 2018). These models directly take image pairs as inputs and automatically learn the relationship between coarse/fine image pairs. The results indicate that learning-based algorithms are more robust than the traditional spatiotemporal fusion algorithm (Sun and Zhang 2019). However, it usually requires abundant data for training the mapping relationship between fine and coarse satellite images.

The spatial unmixing-based methods are applied to compute the endmember (i.e., label) of coarse pixels and estimate the fined pixels using weighted endmembers (Zurita-Milla et al. 2008). According to Gevaert and García-Haro (2015), there are four steps in a spatial unmixing-based fusion model: (1) clustering the high-resolution data set to define the endmembers, (2) calculating the fractions of each endmember within each coarse spatial resolution pixel, (3) unmixing the medium-resolution pixel, and (4) assigning reflectance spectra to the high-resolution pixels. The unmixing can be applied using only one land cover thematic map with a fine spatial resolution (i.e., prior classification results). The thematic map can be produced by interpreting the available fine spatial resolution data (e.g., land use database). For example, Zurita-Milla et al. (2008) produced a 30 m Landsat-like time series by integrating one 30 m thematic map obtained by the classification of an available Landsat image and 300 m medium resolution imaging spectrometer time series. Furthermore, recent research illustrates that image pair-based and spatial unmixing-based methods can be combined (Zhu et al. 2016; Xie et al. 2016). Gevaert and García-Haro (2015) combined the advantages of STARFM and unmixing-based algorithms to propose a novel spatial and temporal reflectance unmixing model, which directly estimates the land cover changes between two coarse images.

In summary, traditional spatiotemporal fusion methods are mostly based on fusing each fine-coarse image pair in a pixel-wise process, which is not suitable for large-scale remote sensing data sets as the prediction is very time-consuming. In recent years, a variety of deep learning networks and large-scale remote sensing data sets have been published. The potential of deep learning-based spatiotemporal fusion methods needs to be further investigated, and novel methods should be proposed, mainly based on weak supervision.

Semantic Segmentation of Remote Sensing Image Using DeepLabV3+

Several studies have been carried out on using DeepLabV3+ for land cover classification tasks for aerial images. In a comparison study by Pashaei et al. (2020), the authors evaluated the performances of multiple semantic segmentation architectures on unmanned aircraft vehicle images for efficient land cover mapping. The experimental results demonstrate that DeepLabV3+ has a great potential for accurate land cover prediction tasks on a limited labeled image. On the other hand, some researchers extend the original DeepLabV3+ network to be more applicable for land observation images. For instance, Chen et al. (2019) proposed an improved network-based on DeepLabV3+ for semantic segmentation of high-resolution remote sensing images. The authors adopt dilated convolution by adding an augmented atrous spatial pyramid pool layer and a fully connected fusion path layer. As a result, dilated convolution enlarges the receptive field of feature points while the feature map resolution remains unchanged.
In addition to general land cover classification, DeepLabV3+ has been utilized for specific land cover mapping applications such as agricultural mapping (Du et al. 2019) and vegetation mapping (Ayhan and Kwan 2020). Here, we proposed its usage in a weakly supervised deep learning-based data fusion method.

**Method**

The workflow of the proposed approach to weakly supervised deep learning-based data fusion is shown in Figure 1. Details about each stage are presented in the next subsections.

**Semantic Segmentation**

The basic framework of our data-fusion model is the semantic segmentation network developed by Chen et al. (2018), namely DeepLabV3+. It is the latest version of DeepLab semantic segmentation architecture, which utilizes an atrous spatial pyramid pooling (ASPP) module. It extends the previous version (DeepLabV3) by adding a decoder module to refine the segmentation results, especially along object boundaries (Chen et al. 2018). The framework achieves a state-of-the-art mean intersection-over-union of 89% on the PASCAL VOC 2012 test. The ASPP mechanism improves the segmentation performance by exploiting the multi-scale contextual information of the features. The encoder part of the network structure enables DeepLabV3+ to reduce the feature maps and capture semantic information, while the decoder part recovers the spatial information.

**Preprocessing of Sentinel-1 SAR Images**

The presence of speckle noise in the Sentinel-1 SAR images makes the interpretation of the contents difficult, thereby degrading the quality of the image. Therefore, an efficient speckle noise removal technique needs to be applied to the Sentinel-1 SAR images. In this study, SAR images are processed by the Enhanced Lee Filter (Lee 1981) to deal with the common problem of noisy edge boundaries. The filter algorithm operates by using edge directed windows. The local mean and local variance are computed using only the pixels in the edge directed window. After the speckle filtering, the images are enhanced by 2% of the linear stretch. The lowest and the highest 2% values are set to 0 and 255, respectively. Values in between are distributed from 0 to 255. As shown in Figure 2, the noise in the high contrast areas is effectively removed, and the edges are enhanced.

![Figure 1. The workflow of the proposed approach.](image1)

![Figure 2. Example of raw Sentinel-1 SAR image (left) and the processed counterpart (right).](image2)
Data Augmentation

Several augmentation techniques have been added to the data-loader module of the model network to improve the performance by enlarging the training data set. These include geometric transformations (e.g., flip, rotation, warp) and linear transformations (e.g., 2%-98% contrast stretch). All geometric transformations are randomly selected and applied to images, each with a probability of 0.5. The linear stretch is assumed to be useful as applying to images with low contrast (e.g., image taken during nighttime).

Label Refinement

In essence, the major task of this study is semantic segmentation on weakly-supervised training, in which the annotation (i.e., MODIS labels) is noisy and unreliable. To further improve the performance of the model, additional strategies were adopted to deal with noisy labels specifically. In the SEN12MS data set, images of each scene were selected and cropped to be relatively homogenous. Noises (or incorrect labels) normally exist at the edges of land cover parcels. For example, shorelines are not clearly shown on the MODIS maps. For that matter, an additional module was added to the model, which updates the labels every five epochs (an epoch refers to one cycle through the full training data set). Hence, only for the first five epochs, the model is trained on original MODIS labels. After the fifth epoch, the model outputs the intermediate predictions on all training samples and then obtains the updated labels by comparing the intermediate predictions with the original MODIS labels and the predictions are used for the next five epochs. Figure 3 shows the label refinement steps.

Implementation Details

Our model was implemented on PyTorch and worked on one graphics processing unit (NVIDIA 2070-super). The weights of a pretrained model on the ImageNet data set are used for the initialization of our model. It is worth mentioning that the number of land covers in the training data set is different from the number of classes in the ImageNet data set, so the logit weights in the pretrained model are excluded. In this work, several modifications were made to the original DeepLabV3+ network. Preprocessing of Sentinel-1 SAR images and data augmentation were added to the data-loader module of the network, and the structure of the network was altered to update the label during the training process. In addition, the original DeepLabV3+ is used as the baseline model to compare with our model. Both models were trained for 50 epochs, and the average time per epoch is around one hour. The training parameters of our model and the baseline are presented in Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretrained on ImageNet</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>SAR image preprocessing</td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Data augmentation</td>
<td>True</td>
<td>True</td>
</tr>
<tr>
<td>Label refinement</td>
<td>False</td>
<td>True</td>
</tr>
<tr>
<td>Backbone network</td>
<td>ResNet-101</td>
<td></td>
</tr>
<tr>
<td>Momentum</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Initial learning rate</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Number of epochs</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>Batch size</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Output stride</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Weight decay</td>
<td>0.00005</td>
<td></td>
</tr>
</tbody>
</table>

Experiments and Discussion

Data Set

The model is trained on a public satellite imagery data set, SEN12MS, which was published by Schmitt et al. (2019). This data set contains globally distributed scenes, covering inhabited continents during all meteorological seasons. SEN12MS includes 180,662 triplets of Sentinel land cover maps (see Figure 4), dual-polarized (VV and VH) SAR Sentinel-1 image patches, and Sentinel-2 multispectral image patches. Each image is cropped to a size of 256 × 256 pixels. While all data are oversampled to be at a ground sample distance of 10 m, the Sentinel images have a native resolution of about 10 to 60 m per pixel, and the MODIS-derived land cover has a native resolution of 500 m per pixel.

The Sentinel-1 SAR images were provided in the original form with no preprocessing (e.g., speckle filtering). For the Sentinel-2 multispectral images, a sophisticated mosaicking workflow was implemented to avoid the impacts of cloud-covered images. On the other hand, the MODIS land cover maps were created based on calibrated MODIS reflectance data in...
The raw reflectance data was classified following the International Geosphere-Biosphere Programme (IGBP) classification scheme (Loveland and Belward 1997) and land cover classification system (LCCS) scheme (Di Gregorio 2005). Moreover, sophisticated postprocessing is carried out for class-specific refinement, which integrates prior knowledge, auxiliary information, and temporal regularization based on a Markov random field (Schmitt et al. 2019). For different classification schemes, the provided MODIS maps have overall accuracies of approximately 67% under IGBP, 74% under LCCS land cover, and 81% under LCCS land use (Sulla-Menashe et al. 2019). For this study, a simplified version of IGBP was chosen to be the classification scheme. It means that the coarse label used in this study can only correctly annotate at most 67% of the image pixels.

The data set of the 2020 IEEE Geoscience and Remote Sensing Society Data Fusion Contest (DFC2020) was used to validate and test the performance of our deep learning spatiotemporal fusion model. The DFC2020 data set contains scenes with undisclosed geolocation and not contained in the SEN12MS data set, with semimanually derived high resolution (10 m) land cover maps as the ground truth labels. In addition to the high-resolution ground truth labels, the validation and testing images are provided in the same triplet format as the training data set (i.e., corresponding Sentinel-1, Sentinel-2, and MODIS labels). The validation set contains 986 quadruplets, and the testing set has 5128 quadruplets (see Figure 5).

### Classification Scheme and Evaluation Metric

A simplified version of the IGBP classification scheme is used for this project. As shown in Table 2, the original IGBP scheme has 17 classes in total. The simplified scheme has 10 classes.

The fusion results were evaluated using the classification accuracy as the quantitative indicator. The higher the accuracy is, the training model has a stronger ability to classify land cover features. The accuracy is defined by

\[
\text{Accuracy} = \frac{1}{m} \sum_{i=1}^{m} \delta (f_i = y_i),
\]

where \( m \) is the number of samples, \( f_i \) and \( y_i \) are the true and predicted pixel label values, and \( \delta \) is the Iverson bracket operator, which evaluates to 1 when the labels match, and to 0 when labels mismatch.

### Quantitative Results

It took about 25 minutes for the trained model to predict the 5128 images of the testing set. The results on the validation set and the testing set for the baseline and the proposed models are shown in Tables 3 and 4, respectively. The overall performances were assessed using average class accuracy (AA), which indicates the mean of the accuracies of all land cover classes in the simplified IGBP scheme. It is worth mentioning that the validation and testing data set does not include savanna (Class 3) and snow/ice (Class 8).

### Table 2. Original and simplified IGBP land cover classification schemes.

<table>
<thead>
<tr>
<th>Simplified Class No.</th>
<th>Simplified Class Name</th>
<th>IGBP Class Name</th>
<th>IGBP Class No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Forest</td>
<td>Evergreen Needleleaf Forest</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Deciduous Needleleaf Forest</td>
<td>Deciduous Needleleaf Forest</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>Mixed Forest</td>
<td>Mixed Forest</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>Shrubland</td>
<td>Closed Shrublands</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>Savanna</td>
<td>Woody Savannas</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>Grassland</td>
<td>Grasslands</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>Wetlands</td>
<td>Permanent Wetlands</td>
<td>11</td>
</tr>
<tr>
<td>8</td>
<td>Croplands</td>
<td>Croplands</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>Urban/Built-Up</td>
<td>Urban/Built-Up</td>
<td>13</td>
</tr>
<tr>
<td>10</td>
<td>Snow/Ice</td>
<td>Permanent Snow and Ice</td>
<td>15</td>
</tr>
<tr>
<td>11</td>
<td>Barren</td>
<td>Barren</td>
<td>16</td>
</tr>
<tr>
<td>12</td>
<td>Water</td>
<td>Water Bodies</td>
<td>17</td>
</tr>
</tbody>
</table>

**Figure 4.** An example of SEN12MS triplets.

**Figure 5.** An example of DFC2020 quadruplets.
Table 5. Normalized confusion matrix for our model on DFC2020 testing set.

<table>
<thead>
<tr>
<th>Class</th>
<th>Forest</th>
<th>Shrubland</th>
<th>Grassland</th>
<th>Wetland</th>
<th>Cropland</th>
<th>Urban/built-up</th>
<th>Barren</th>
<th>Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Forest</td>
<td>0.75</td>
<td>0.11</td>
<td>0.07</td>
<td>0.01</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Shrubland</td>
<td>0.23</td>
<td>0.14</td>
<td>0.45</td>
<td>0.02</td>
<td>0.14</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Grassland</td>
<td>0.12</td>
<td>0.15</td>
<td>0.47</td>
<td>0.01</td>
<td>0.23</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Wetlands</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
<td>0.28</td>
<td>0.14</td>
<td>0.00</td>
<td>0.00</td>
<td>0.47</td>
</tr>
<tr>
<td>Croplands</td>
<td>0.01</td>
<td>0.07</td>
<td>0.13</td>
<td>0.10</td>
<td>0.64</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
</tr>
<tr>
<td>Urban/built-up</td>
<td>0.06</td>
<td>0.02</td>
<td>0.04</td>
<td>0.03</td>
<td>0.78</td>
<td>0.01</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Barren</td>
<td>0.07</td>
<td>0.14</td>
<td>0.05</td>
<td>0.00</td>
<td>0.62</td>
<td>0.01</td>
<td>0.00</td>
<td>0.94</td>
</tr>
<tr>
<td>Water</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.05</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Qualitative Comparison

In addition to the accuracy evaluations, the visualization of the predicted maps was also presented for a qualitative overview of the spatial resolution enhancement of the land cover mapping. Enhanced land cover maps obtained by our model are shown in Figures 6a and 6b to demonstrate how the model performs on predicting different land covers. Each example includes the input Sentinel-2 multispectral image, the input Sentinel-1 SAR image, the original MODIS label/map, the enhanced map from the prediction of our model, and the DFC2020 ground truth label/map. As shown in Figure 6, the detection of shorelines and beaches are well recognized on the enhanced land cover map, with smoothed boundaries between land cover parcels.

Figure 6b shows that the model successfully reduced the impact of the misclassified label of grassland on the corresponding MODIS land cover map. Moreover, by visually analyzing the input image and the DFC ground-truth label, we can find that the DFC map underestimates the area of urban/built-up in this image. In contrast, the enhanced map correctly detects the presence of buildings. It indicates that even the ground truth label could still contain minor misclassifications. Additionally, Figure 6c shows that the model poorly identifies narrow rivers or small ponds despite the significant spectral differences. Both Figure 6c and 6d show that our model tends to misclassify cropland, wetland, and grassland.

In summary, the proposed model tends to be biased toward high represented classes such as forest, grassland, and urban. This is probably related to the fact that those classes exhibit more general textural and spectral characteristics, confusing the model prediction. In any case, our model presented results superior to the baseline with a significant margin.

Conclusion

In this paper, a weakly supervised deep learning-based approach was proposed for the fusion of satellite data at high spatial resolution (Sentinel-1 and Sentinel-2) with satellite-derived land cover maps at high temporal resolution (MODIS) to perform the enhanced land cover mapping. Considering the large spatial resolution gap between Sentinel and MODIS images, the fusion was conducted through a weakly supervised semantic segmentation. We modified the original DeepLabV3+ segmentation architecture by adding a label-update module to update the coarse label throughout the training automatically.
The experiment results have validated the effectiveness and potential of deep learning-based semantic segmentation architecture in the fusion of multi-source satellite data, improving land cover mapping.

References


Ecological Functions and Human Activity Interference Evaluation in Ecological Protection Redline for Urban Environment

Guoming Li, Binbin He, Liang Li, Miao Yang, Qiongyi Huang, and Zihan Guo

Abstract
The influences of ecological protection redlines on urban sustainable development can significantly vary from place to place with distinct ecological functions and human activity interference. However, recent research lacks the evaluation of ecological functions and human activity interference in ecological protection redline areas near urban areas. This article presents an evaluation method consisting of two evaluation criteria systems: an ecological function evaluation criteria system based on conserving water resources and soil and maintaining biodiversity and a human activity interference evaluation criteria system based on the human activity impact index, population and road network density, and the intensity of tourism activity and livelihood sewage discharge. This evaluation method was verified in Zoige County, China, by using remote sensing data to evaluate ecological functions and human activity interference in the ecological redline area for an urban environment. This evaluation method is one of the preliminary studies of urban sustainable development planning, and the result could provide a basis for the formulation of urban sustainable development strategy.

Introduction
The ecological protection redline is the lifeline of China’s national ecological security. It is a national strategy and also an important measure to promote ecological civilization, to delineate ecological protection redline, and to strictly abide by it. Ecological assessment and analysis are important topics for the urban environment. The economies and societies of megacities have developed rapidly, but they are also under pressure from increasing population, resources, and ecological threats (Li et al. 2019; Shao et al. 2019). New urban areas come mainly from cultivated land and ecological land (Zhang et al. 2019a). The Chinese government started an important project monitoring geographical conditions in 2012 that aims at fully revealing the spatial pattern of natural resources and economic and social development (Ding et al. 2016). The redline is designated to protect cultivated land from urban construction (Shao et al. 2020d). The ecological protection redline (hereinafter referred to as the “redline”) area refers to an area within the scope of ecological space providing special and important ecological functions that must be compulsorily and strictly protected. It is the baseline and lifeline for guaranteeing and maintaining national ecological security. Those areas include areas with important ecological functions, such as water resource conservation, biodiversity maintenance, soil conservation, and sand fixation, and areas that are ecologically sensitive and vulnerable, such as water loss, soil erosion, land and rock desertification, and salinization areas (Deng et al. 2018; Shao et al. 2020a, 2020b).

Redline delineation has almost been completed in China, and the redline supervision platform has also started construction under a national unified arrangement. Related research carried out abroad and at home has focused mainly on how to delineate the redline (Hu et al. 2018; Kong et al. 2019; Liang et al. 2020; Shao et al. 2020c), how to optimize and adjust the redline delineation (Shao et al. 2019; Zhang et al. 2019; Xu et al. 2020), and how to assess the ecological value provided by the redline area (Hou et al. 2018; Shao and Cai 2018; Xu et al. 2018; Deng et al. 2020; Xu et al. 2020; Zeng et al. 2020). However, how to scientifically evaluate the ecological status and human activity interference has become a hot spot of general concern and has also been greatly needed in the comprehensive management of the redline. Wu et al. (2020) carried out a quantitative analysis of the risk of interference and degradation on habitat receptors within the redline area caused by several risk sources, such as farmland, cities, towns, mining areas, and main traffic lines. Wang et al. (2020) constructed a remote sensing ecological index based on several indicators, such as greenness, humidity, heat, and dryness, and then applied this index to evaluate the quality of ecological environment in the redline area. Yang et al. (2020) studied the evaluation method of ecosystem health in the redline area based on land use and landscape pattern. Niu et al. (2017) used remote sensing images to monitor and evaluate three indicators of the redline area: the composition and change of the natural ecosystem, the change of the average patch area of the ecological landscape, and the Normalized Difference Vegetation Index (NDVI). Research on the evaluation of ecological functions and human activity interference in the redline area remains sparse if nonexistent. Moreover,
the redline area can have a more direct impact on the sustainable development of near-urban areas, and therefore it is more important to establish the ecological function evaluation criteria system in the redline area near urban areas. The research objectives of this article are to fill the research gap mentioned above and propose a comprehensive evaluation method in the redline area near urban areas, taking two aspects (ecological function and human activity interference) into account that could provide background support and a scientific basis for ecological protection and restoration there.

Establishment of the Evaluation Criteria System

The main principles of selecting evaluation criteria are easy to express and understand, convenient to collect data, and able to support the establishment of a sound analysis method because the evaluation criteria need to not only represent the protection status but also support the administrative management of the redline area. According to requirements of the “Green Shield 2017” Special Action Plan of Supervision and Inspection in National Nature Reserve, combined with the author’s practical work experience in the past five years, the establishment of the evaluation criteria system in this article follows these principles: the criteria system should be highly integrated to systematically reflect the main characteristics of the regional ecosystem and the relationship between various subsystems, and the evaluation criteria should be sensitive to the changes of ecosystem and related human activities and be able to take both the integrity and the systematism of the redline into account to not only respond to the requirements for the protection of mountains, rivers, forests, farmlands, lakes, and grasslands within the redline area but also highlight the characteristics of the redline area. On the one hand, redline areas consist mainly of areas with important ecological functions, such as water resource conservation, soil conservation, biodiversity maintenance, and so on, so the ecological function evaluation criteria system has been constructed from these three aspects mentioned above. On the other hand, the main threats to the redline area include deforestation, shrinking of wetlands, degradation of grasslands, environmental pollution, and so on, which are caused by both various development activities and livelihood project constructions, such as livelihood production of original residents, tourism activities, and road construction. For this reason, the construction of the human activity interference evaluation criteria system has taken these factors into account: the intensity of livelihood sewage discharge, the intensity of tourism activities, the intensity of traffic facility construction, the degree of destruction of the ecological landscape, and the intensity of pollutant emissions. The evaluation criteria system is discussed next.

Method of Evaluation

Evaluation of Ecological Functions

Evaluation of Water Resource Conservation Function

Total water resource conservation. In this article, the precipitation storage method has been selected. The total water resource conservation of different ecosystems can be measured by their impoundment effect. The calculation method proposed by Ouyang et al. (2016) has been applied here. The formulas are as follows:

\[ Q = A \times J \times R, \]  
\[ J = J_0 \times K, \]  
\[ R = R_0 - R_v, \]

where \( Q (\text{mm/(hm}^2\cdot\text{a}^{-1})) \) is the increment of water conservation in ecosystems, such as forests, grasslands, wetlands, arable lands, and deserts, compared with bare soil; \( A (\text{hm}^2) \) is the area of the ecosystem; \( J (\text{mm}) \) is the annual average runoff generation in the calculation area; \( J_0 (\text{mm}) \) is the total annual average rainfall in the calculation area; \( K \) is the proportion of runoff generation to total rainfall in the calculation area; \( R \) is the benefit coefficient of runoff reduction in the ecosystem compared with bare soil (or cutover land); \( R_0 \) is the rainfall runoff rate in bare land under the runoff generation condition; and \( R_v \) is the rainfall runoff rate in the ecosystem under the runoff generation condition.

Change rate of water resource conservation capacity. Change rate of water resource conservation capacity is tentatively calculated every five years in accordance with the frequency of redline adjustment. The rate can be determined by the change in the ratio of total water resource conservation to the area of the redline area (Ouyang et al. 2016).

Evaluation of Soil Conservation Function

Total soil conservation. Total soil conservation in ecosystem is the difference between potential soil erosion and actual soil erosion. In this article, the universal soil loss equation is used to calculate total soil conservation, and the formulas are as follows (Ouyang et al. 2016):

\[ SC = SE_p - SE_a, \]
\[ SE = R \times K \times LS, \] (5)
\[ SE_i = R \times K \times LS \times C, \] (6)

where \( SC \) (t×hm\(^{-2}\)×a\(^{-1}\)) is the total soil conservation, \( SE_i \) (t×hm\(^{-2}\)×a\(^{-1}\)) is the potential soil erosion, \( SE \) (t×hm\(^{-2}\)×a\(^{-1}\)) is the actual soil erosion, \( R \) (MJ×mm×hm\(^{-2}\)×h\(^{-1}\)×a\(^{-1}\)) is the rainfall erosivity factor, \( K \) (t×hm\(^{-2}\)×h/×MJ\(^{-1}\)×mm\(^{-1}\)) is the soil erodibility factor, \( LS \) is the topographic factor, and \( C \) is the vegetation coverage factor.

**Soil conservation capacity.** Soil conservation capacity refers to the amount of soil conservation in the unit redline area, which can be determined by the ratio of total soil conservation to the area of the redline area (Xu et al. 2016).

**Change rate of soil conservation capacity.** Change rate of soil conservation capacity is tentatively calculated every five years in accordance with the frequency of redline adjustment. The rate can be determined by the ratio of the difference between the capacity of the calculation year and that of the base year to the capacity of the base year (Ouyang et al. 2017; Kong et al. 2019).

**Evaluation of Biodiversity Maintenance Function**

**Biodiversity maintenance capacity.** Biodiversity maintenance capacity in the ecosystem can be evaluated from two aspects: habitat quality and biodiversity (Niu et al. 2017; Kong et al. 2019). The capacity can be represented by the sum of the habitat quality index and the biodiversity index. The formulas for calculating these two indices are as follows:

\[ Q_{xj} = H_j \left( 1 - \frac{D_j^2}{D_{xj}^2} + k^2 \right) \] (7)

where \( BI \) is the biodiversity index, \( Q_{xj} \) is the habitat quality index of grid \( x \) in land use (land cover) type \( j \), \( D_{xj} \) is the threat of grid \( x \) in land use (land cover) or habitat quality type \( j \) to habitat quality, \( k \) is the half-saturation constant, and \( H_j \) is the habitat quality adaptability of land use (land cover) type \( j \) and.

**Change rate of biodiversity maintenance capacity.** Change rate of biodiversity maintenance capacity is tentatively calculated every five years in accordance with the frequency of redline adjustment. The rate can be determined by the ratio of the difference between the capacity of the calculation year and that of the base year to the capacity of the base year (Xu et al. 2016).

**Evaluation Criteria System of Human-Activity Interference**

**Human Activity Impact Index**

According to the monitoring results of human activities, the author has calculated the human activity impact index \( RHI \) and then evaluated the human activity impact in the redline area based on the classification results. The evaluation formula is as follows:

\[ RHI = \frac{a_1b_1x_1 + a_2b_2x_2 + ... + a_nb_nx_n}{x}, \] (9)

where \( x_i \) is the area of a certain human activity, \( a_i, b_i \) is the weight of a certain human activity, \( a_i \) is determined by different functions in redline areas where certain human activity is located, \( b_i \) is determined by the different impacts of various types of human activities in the redline area, and \( x \) is the total area of the redline area.

Based on the different impacts of various human activities in the redline area, the expert scoring method has been applied in this article, selecting five experts in each of the three fields (land surveying and mapping, ecological and environmental protection, and social economy) to determine the weight of each index.

**Population Density**

Population density refers to the ratio of the number of original residents in the redline area to the total area of the redline area. The calculation formula is as follows:

\[ PD = \frac{NOR}{RA} \times 100\%. \] (10)

where \( PD \) is population density, \( NOR \) is the number of original residents in the redline area, and \( RA \) is the area of the redline area.

**Table 1. Weight of different human activity impacts.**

<table>
<thead>
<tr>
<th>First Class</th>
<th>Second Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Code</td>
<td>Criteria</td>
</tr>
<tr>
<td>1</td>
<td>Cultivated</td>
</tr>
<tr>
<td>2</td>
<td>Roads</td>
</tr>
<tr>
<td>3</td>
<td>Roads</td>
</tr>
<tr>
<td>4</td>
<td>Structures</td>
</tr>
<tr>
<td>5</td>
<td>Artificial</td>
</tr>
<tr>
<td></td>
<td>piles</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING
Road Network Density
Road network density refers to the ratio of the total length of railway and high-grade roads in the redline area to the area of the redline area. The calculation formula is as follows:

\[ RND = \frac{TL}{RA} \times 100\%, \tag{11} \]

where \( RND \) is road network density, \( TL \) is the total length of railway and high-grade roads in the redline area, and \( RA \) is the area of the redline area.

Tourism Activity Intensity
Tourism activity intensity refers to the ratio of the total annual tourists in the redline area with scenic tickets to the area of the redline area. The calculation formula is as follows:

\[ ITA = \frac{\sum_{j=1}^{m} NT_j}{RA} \times 100\%, \tag{12} \]

where \( ITA \) is the intensity of tourism activity, \( NT \) is the number of tourists, \( RA \) is the area of the redline area, and \( j \) is the scenic spots with tickets.

Livelihood Sewage Discharge Intensity
Livelihood sewage discharge intensity can be expressed by the ratio of the amount of livelihood sewage discharge and garbage discharge from original residents and tourism activities in the redline area to the area of the redline area. The calculation formula is as follows:

\[ EPL = \frac{PD}{RA} \times 100\%, \tag{13} \]

where \( EPL \) is the livelihood sewage discharge intensity, \( PD \) is the amount of sewage and garbage, and \( RA \) is the area of the redline area.

Comprehensive Evaluation
In this article, the weighted method has been adopted for the comprehensive result of the evaluation criteria system. The final comprehensive result for judging sustainable development can be obtained by weighing the weights of different criteria.

The author obtained subjective weight \( w \) for each criteria through the expert scoring method, based on different impacts of criteria on the environment, in order to indicate its relative importance in the whole criteria system:

\[ w = (W_1, W_2, W_3, ..., W_m)^T, \tag{14} \]

\[ \sum_{i=1}^{m} W_i = 1, W_i \geq 0 (i = 1, 2, ..., m). \tag{15} \]

Human activity intensity has been selected for quantitative expression to synthesize the impacts of each criterion. The synthesis method is a criteria weighting method through which human activity intensity \( F_i \) can be determined by summing the weighted average mean of each criterion above. The dimensionless calculation formula is as follows:

\[ F_i = \sum_{i=1}^{m} W_i f_i \tag{16} \]

Human activity intensity has then been classified into five levels according to the natural discontinuity method to make further comparisons.

---

Data Sources and Overview of the Study Area

Data Sources
The remote sensing data for monitoring the redline area were collected mainly from surveying and mapping departments, the environmental protection department, the forestry department, the meteorology department, and statistical yearbooks. High-resolution satellite series, resource satellite series, Beijing satellite series and other high-resolution remote sensing data (full color, multispectral), administrative divisions, road distribution, land cover, and other data were collected from the land surveying and mapping department and are used mainly for extracting human activities from UAV images and calculating criteria related to ecological functions (Zhang et al. 2020). Temperature, precipitation, surface runoff, evapotranspiration, and other data from the meteorological department were used mainly for calculating criteria related to ecological functions. Population, number of tourists, amount of livelihood sewage discharged, soil texture, soil organic matter content, vegetation type, and number of species from local statistical yearbooks were used mainly for calculating criteria related to human activity interferences. Aboveground biomass in the suburbs can be calculated from combined LiDAR and Landsat-8 data (Zhang et al. 2019b), and urban vegetation biomass can be calculated in combination with LiDAR and high-resolution remote sensing images (Zhang and Shao 2020).

Overview of the Study Area
Zoige County is located on the eastern edge of the Qinghai-Tibet Plateau and north of the Aba Tibetan and Qiang Autonomous Prefecture of Sichuan Province, China. Its geographical coordinates are between 102°08′E to 103°39′E and 32°56′N to 34°19′N. The watershed of the Yellow River and the Yangtze River divides Zoige County into two parts, with a total land area of 10 436.58 km². The total area of ecological protection redline area is 5107.07 km², accounting for 50.06% of the total area of Zoige County.

---

Figure 2. Location of the study area.
Results and Discussion

Ecological Functions Evaluation

Water resource conservation can be greatly influenced by topography and precipitation. The terrain of most of the redline area in Zoige County is relatively gentle, where there exist abundant rainfall, high vegetation coverage, and low evaporation. It is the main water resource conservation area in the upper reaches of the Yellow River, and the water resource conservation function is at a high level. The terrain in the west of the redline area is undulating, where water storage capacity is poor and water content is low, resulting in a relatively weak soil conservation function. However, the biomass and soil conservation function in the east of the redline area are higher and better because of the grasslands and wetlands there. In the north and the east of the redline area with low altitude, due to the rich ecological types and high vegetation coverage, the biodiversity maintenance function is strong. The main ecological types in high-altitude areas are glacier and snow, and therefore the assessment value of the habitat quality in these areas is relatively low. Furthermore, overgrazing in the western area (mostly grasslands) has reduced the vegetation coverage and intensified desertification, resulting in relatively poor habitat quality in this area.

Human Activity Interference Evaluation

Most areas in the Zoige County redline area are high-altitude virgin forests and nature reserves. Therefore, compared to areas outside the redline area, there are fewer roads or permanent residents. However, the large number of tourist attractions and the high intensity of tourism activities in Zoige County have left a certain influence on the ecological environment in the redline area. Based on the calculation of all criteria, human activity interference is slight. The human activity areas in the redline area can be divided into two parts. One is located in the nature reserves with some tourism activities, which have had influences on the ecological environment in the redline area. Therefore, illegal human activities in the redline area should be cleaned up and rectified in strict accordance with the “Regulations on the Management of Nature Reserves,” while legitimate human activities should be reasonably controlled under the condition that urbanization development in these areas is prohibited. The other one is located in the scientific evaluation areas, classified mainly as cultivated land, roads, and residential sites. Ecological protection and ecological construction need to be strengthened in these areas because agricultural construction has developed rapidly and there exists a greater risk of geological disasters. Furthermore, in order to promote the sustainable development of surrounding cities and towns, destruction of the ecological environment and large-scale development and construction should be strictly prohibited, while the implementation of ecological restoration projects needs to be focused on, and this could help to further strengthen ecological restoration and protection.

Conclusions

The ecological protection redline can be divided into three lines: the baseline of ecological function maintenance, the bottom line of environmental quality and safety, and the limit...
line of natural resources. However, at present, there are no relevant technical methods for the evaluation of both ecological function and human activity interference in the ecological protection redline area. This article has presented an evaluation method with two criteria systems. One is the ecological function evaluation criteria system based on the water resource conservation function, the soil conservation function, and biodiversity maintenance function criteria. The other one is human activity interference evaluation criteria system based on the intensity of socioeconomic activity of the original residents, the intensity of tourism development, the intensity of infrastructure and livelihood construction, and the impact of environmental pollution. This evaluation method could provide background support and a scientific basis for ecological protection and restoration in the redline area.

Acknowledgments
This study was supported by the Science and Technology Program of Sichuan (2018SZ0314, 2020YFS0453, 2020YFS0058).

References

Figure 6. Comprehensive evaluation of human activity interferences.


IN-PRESS ARTICLES

Xueyan Li, Yu Hou, Ruifeng Zhai, Junfeng Song, Xuehan Ma, Shuzhao Hou, and Shuxu Guo. Three-dimensional reconstruction of single input image based on point cloud.

Rongjun Qin, Xiao Ling, Ph.D., Xu Huang. A Unified Framework of Bundle Adjustment and Feature Matching for High-resolution Satellite Images.

Zezhong Zheng, Mujie Li, Mingcang Zhu, Yue He, Jun Xia, Xueye Chen, Qingjun Peng, Yong He, Xiang Zhang, and Pengshan Li. The spatio-temporal evolution of urban impervious surface for Chengdu, China.

Lei Zhang, Hongchao Liu, Xiaosong Li, and Xinyu Qian. Optimizing the segmentation of a high-resolution image by using a local scale parameter.


Qinghong Sheng, Rui Ren, Weilan Xu, Hui Xiao, Bo Wang, and Ran Hong. Spinor-based Attitude Determination with Star Sensor Considering Depth.

Bharath H Aitha and Prakash PS. Digital building height preparation from satellite stereo images.

Jiabao Li, M.D, Wei Han, Ruyi Feng, Lizhe Wang, and Fengpeng Li. Unsupervised representation high-resolution remote sensing image scene classification via contrastive learning convolutional neural network.


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.


Qinghong Sheng, Rui Ren, Weilan Xu, Hui Xiao, Bo Wang, and Ran Hong. Spinor-based Attitude Determination with Star Sensor Considering Depth.

Bharath H Aitha and Prakash PS. Digital building height preparation from satellite stereo images.

Jiabao Li, M.D, Wei Han, Ruyi Feng, Lizhe Wang, and Fengpeng Li. Unsupervised representation high-resolution remote sensing image scene classification via contrastive learning convolutional neural network.


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.


A High-Resolution Satellite DEM Filtering Method Assisted with Building Segmentation

Yihui Li, Fang Gao, Wentao Li, Peng Zhang, Yuan An, Xing Zhong, Yuwei Zhai and Yongjian Yang

Abstract
Digital elevation model (DEM) filtering is critical in DEM production, and large-area meter-level resolution DEM is mainly generated from high-resolution satellite images. However, the current DEM filtering methods are mostly aimed at laser scanning data and tend to excessively remove ground points when processing a satellite digital surface model (DSM). To accurately filter out buildings and preserve terrain, we propose a DEM filtering algorithm using building segmentation results of orthophoto. Based on morphological filtering, our method estimates the probability of being a built-up area or mountains for DSM, and according to this probability the filtering parameters are adaptively adjusted. For robustness, our method performs the above filtering operation on DSM through a sliding-window approach, and finally the nonground points are determined by the votes of multiple filtering. Experiments against six representative data sets have shown that our method achieved superior performance than classical algorithms and commercial software.

Introduction
Digital elevation model (DEM) is an important type of geo-spatial information that serves as the basic data for geologic modeling and analysis and is widely used in hydrology, geologic disaster prevention, land-use planning, civil engineering, etc. Large-area DEM with meter-level grid size is mainly produced from high-resolution satellite images. The traditional way of generating DEM is manually filtering out the nonground points like vegetation, buildings, and other man-made features in a digital surface model (DSM), such as the WorldDEM data set (Riegler et al. 2015). Obviously, this method is labor-intensive and time-consuming, and the result of it is limited by the operator’s experience and proficiency. With continuous progress in automatic DEM filtering, some software like PCI Geomatica (PCI Geomatics 2020) provides the functionality of converting DSM to DEM with little manual intervention. Currently, most DEM filtering methods are targeted to laser scanning point cloud. They share a basic assumption: the height differences between nonground points and nearby ground points are abrupt, and the elevation changes of undulating terrain are smooth. Based on this, the algorithms distinguish ground or off-ground points on DSM. These algorithms can be classified into several major categories, including slope-based algorithm, surface-based algorithm, morphological algorithm, clustering/segmentation-based algorithm, and machine-learning-based algorithm.

The slope-based filtering method was first proposed by Vosselman (2000), and it assumes the slope value at the junction between nonground points and ground points is larger than that of ground points. Therefore, it sets a slope threshold to filter out the nonground points. However, a single fixed-slope threshold filter can lead to large errors in mountainous areas, so Junichi (2012) presented a method that adaptively adjusts the slope threshold based on the local rough terrain. This type of method is easy to implement, but its performance relies on the threshold selection.

The morphological filtering method applied the opening operation to DEM filtering to eliminate bulges on DSM (Kilian et al. 1996). It compares the elevation of a DSM point before and after the opening operation, and if the elevation difference is larger than the preset threshold, the point will be classified as ground. The key parameter of this algorithm is the sliding window size, because if it is too small, some large buildings might not be eroded; if the size is too large, some part of the terrain bump will be classified as nonground and removed. Considering this disadvantage, Zhang et al. (2003) proposed a progressive morphological filter. It adopts multiple sizes of sliding windows with different elevation change thresholds to perform the opening operation, respectively. These results are weighted by sliding window size and interpolated to generate final DEM. We found the morphological method is suitable for parallelization and therefore we chose it as the basic filtering approach in our algorithm.

Generally, the surface-based filtering method iteratively uses ground seed points to fit the ground surface and extracts new ground points from those close to the surface. Proposed by Axelsson (2000), progressive triangular irregular network (TIN) densification (PTD) method uses TIN to fit the terrain and stepwisely extracts ground points to densify TIN. PTD achieved the best performance in comparative experiments on laser scanning data carried out by Sithole and Vosselman (2004). Many other surface-fitting methods with good filtered results are presented. Mongus and Zalik (2012), Chen et al. (2013), and Hu et al. (2014) employed thin plate spline to precisely fit the ground surface. Zhang et al. (2016) presented a cloth simulation filtering (CSF) method that fits the ground surface through simulating a piece of cloth with gravity and rigidness dropping down to the upside-down point cloud. However, the surface-based method is not adopted in our method because it is time-consuming and we are temporarily unable to accelerate it. This method can replace the morphological method as the basic filtering approach in our algorithm for further research.

Fang Gao and Yongjian Yang are with College of Computer Science and Technology, Jilin University, No.2699 Qianjin Rd, Changchun, Jilin, China (gaofang@163.com).

Yihui Li, Fang Gao, Wentao Li, Peng Zhang, Yuan An, Xing Zhong, and Yuwei Zhai are with Chang Guang Satellite Technology Co., Ltd., No. 1299 Mingxi Rd., Changchun, Jilin, China.

Xing Zhong is also with State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, No. 2 Linggong Rd., Dalian, Liaoning, China.

Contributed by Prasad S. Thenkabail, November 11, 2020. (sent for review December 16, 2020; reviewed by Iiya Aneece, Muruli Krishna Gama.)

Photogrammetric Engineering & Remote Sensing
Vol. 87, No. 6, June 2021, pp. 421–430.
0099-1112/21/421–430
© 2021 American Society for Photogrammetry and Remote Sensing
doi: 10.14358/PERS.87.6.421
The clustering/segmentation-based filtering method such as Filin (2012) and Zhang and Lin (2013) first segments the point cloud into clusters according to certain rules, and then performs filtering with a cluster as the basic unit. Jahromi et al. (2011) and Hu and Yuan (2016) applied machine learning to DEM filtering and trained neural networks with manually labeled point cloud samples. These two types of methods are still in the experimental research stage and will not be discussed here in any detail.

The process of generating DSM point cloud from high-resolution optical satellite images is largely different from laser scanning (Hanley et al. 2002; Grodecki and Dial 2003; Crespi 2012). The former DSM is computed from the pixel-wise dense matching of near-vertical viewing epipolar stereopairs and the latter is directly measured by laser. Some of the key differences between two types of DSM are shown in Figure 1: DSM from satellite images generally has (1) a lower resolution than laser scanning; (2) buildings facades inaccurately slantwise connected to the ground caused by limited view angles of images; and (3) larger than meter-sized ups and downs even on flat ground surface due to error. Because of these differences, we are facing a dilemma when using current DEM filtering algorithms to process satellite DSM: if the filtering parameters are set to “strict” (remove buildings as completely as possible), the terrain bumps like mountain tops and ridges will be filtered out; if the filtering parameters are set to “relaxed” (preserve terrain as much as possible), there will be residual building points. In order to solve this problem caused by fixed parameters, based on classical morphological filtering, we propose a new algorithm which can adaptively adjust parameters according to whether this area is built-up or mountainous. To fully exploit satellite images, we perform building segmentation on the corresponding orthophoto of DSM and employ the segmentation result as additional information to assist in locating buildings.

Current building segmentation methods (Huang et al. 2016; Audebert et al. 2017; Badrinarayanan et al. 2017; Peng et al. 2019) are mostly semantic segmentation. However, there is always conglutination between buildings in the result and it has an adverse effect on statistics and analysis. Hence, we adopt the instance segmentation method whose network structure looks like a combination of object detection and semantic segmentation. Kaiming et al. (2017) presented Mask region-based convolutional neural networks (R-CNN) based on Faster R-CNN (Ren et al. 2017). Mask R-CNN is a representative two-stage method that first locates candidate regions-of-interest (ROIs) and then classifies and segments those ROIs in the second stage. Two-stage methods depend heavily on feature localization to produce masks, which is inherently sequential and difficult to accelerate. One-stage methods like fully convolutional instance-aware semantic (FCIS) segmentation (Li et al. 2017) and you only look at coefficients (YOLACT) (Bolya et al. 2019) can perform these steps in parallel. FCIS requires significant amounts of postprocessing after localization, which severely limits its speed. In contrast, the assembly step of YOLACT is more lightweight and can be implemented with graphics processing unit acceleration. For real-time demands, we adopt YOLACT to detect and locate buildings in orthophoto.

To the authors’ best knowledge, the contribution of this work to the community includes: (1) a DEM filtering algorithm specially for DSM generated from high-resolution satellite images, which achieves good performance in both removing buildings and preserving terrain compared with current major methods; (2) we demonstrate the effectiveness of using orthophoto building segmentation result to assist in DEM filtering, offering a practical strategy from engineering perspective.

**Experimental Data**

Since there are few publicly available data sets about DSM, DEM ground truth, and orthophoto generated from high-resolution satellite images, we made experimental datasets for six areas by ourselves, using Jilin-1 optical satellite images. The DSM and orthophoto are produced through PCI Geomatica software and the DEM ground truth is generated by manually filtering out the buildings and man-made features. For the resolution of DSM, DEM and orthophoto is 2 m and the ground surface is almost invisible in forest areas; we did not remove the densely distributed trees in the DEM ground truth, otherwise that part of terrain will be missing and its interpolated elevation will be unreliable. Each data set covers an area of 15–25 km². As shown in Figure 2, landforms and surface objects of these six areas are representative, including mountains, bridges, forests, desert, rural areas with small buildings, metropolis business district with skyscrapers, etc.

**Methodology**

The main idea of the proposed method is adaptively using different filtering parameters to process DSM of built-up areas and

---

**Figure 1.** Examples of laser scanning DSM point cloud (upper-left) and satellite DSM point cloud (upper-right). Below are orthophoto, detail views of ground surface (yellow rectangle), and buildings (red rectangle) of satellite DSM, for which it is difficult to filter out the nonground points by current laser-orientated methods and assess accuracy.
mountainous areas. To determine whether each part of DSM is built-up or mountainous, we perform building segmentation on orthophoto and compute the slope of DSM and rough DEM. The result may not be reliable enough if we simply divide DSM into regular grids and process them respectively because some parts of mountains may be located in the same grid with buildings, and they will be filtered under the same parameters. We also cannot find feasible methods to estimate the boundary of built-up and mountainous areas. Therefore, we adopt a sliding-window approach to perform filtering operation multiple times for a single DSM point. The votes of multiple filtering results determine whether this point is filtered out.

The proposed method includes several steps summarized in Figure 3. After the generation of DSM and orthophoto, we respectively carried out building segmentation on orthophoto and calculating slope at different scales. Before calculating the slope at a large scale, we perform an operation named “initial filtering” to completely remove the nonground points in order to obtain rough DEM. Then, in the building segmentation result, slope images of large and small scales are used to compute the “building factor” for sliding-window filtering. This factor adaptively sets filtering parameters “strict” or “relaxed” in the current sliding window according to the probability of being a built-up area. For efficiency, we retrieve ground points from the nonground points classified by initial filtering and accept its ground points are correctly classified. Because we perform filtering in a sliding-window way, each point will be processed by the morphological filter repeatedly (Zhang et al. 2003) with adaptive parameters for multiple times. So, for a single DSM point, we count the times of being classified as ground or nonground like counting votes and this point will finally be classified as the one that has more votes.

Building Segmentation on Orthophoto

In the proposed method, we perform building segmentation by YOLACT for its high efficiency. The training data set is also labeled manually based on Jilin-1 optical satellite images. Because the appearance and shapes of buildings are of great complexity, the accuracy of building segmentation results is limited. As shown in Figure 4a, there are some false (lower right) and inaccurate (lower left) segmented cases and noises in the result. To make the segmented results more reliable, we reject those with areas that are too large or too small. The thresholds are respectively set to 35 000 pixels and 40 pixels for our data based on experience. For example, some segmented results with red outlines shown in Figure 4b will be eliminated based on the above-mentioned thresholds.

Figure 2. Orthophotos of six experimental sites. From (a) to (f) are data set 1–6.

Figure 3. The flowchart of the proposed method.
Relative Slope of DSM for Building Detection

Apparently, it is difficult to differentiate between some buildings and man-made ground surface in orthophoto without three-dimensional geometric information. Therefore, we compute the slope of DSM to assist in locating buildings. The slope of a ground surface point is the angle between the horizontal plane and the tangent plane of this point. The unit of slope is percentage (%) or degree (°). In raster DEM, slope is computed by a 3 × 3 window and there are several algorithms as Zhou and Liu (2004) stated. We employ the third-order finite difference weighted by reciprocal of squared distance algorithm presented by Horn (1981):

\[
\begin{align*}
S_x &= \arctan\left(\frac{f_x}{f_y}\right) \\
S &= \text{grid cell size, } S = \text{slope}
\end{align*}
\]

According to slope-based filtering algorithm, the slope at building walls is much steeper than the slope on ground surface because the change in elevation of the former is more abrupt than the latter. Therefore, we can detect the edges of buildings by finding pixels of DSM with large slope value. However, in mountainous areas, some slopes may be as steep as walls of buildings and it will cause false detection. Considering that mountainous area always has a wider elevation range than the built-up area, we can enhance the slope value in the built-up area and reduce the effect of mountains by transforming their elevation range to the same level, namely, computing relative slope value instead of absolute. For example, in the proposed method, we split the DSM raster into 100 m × 100 m parts and transform elevations of each part to the same grayscale range, like [0, 255]. The minimum elevation of this part is assigned 0 and the maximum is assigned 255. In this way, the slope value in mountainous area will relatively shrink and edges of buildings will be more significant (Figure 5). A threshold for extracting edges of buildings is given 110 from our experience.

Slope of Down-Sampled Rough DEM for Mountain Detection

Another major concern of our method is to preserve the point cloud of ground surface, especially in mountainous areas because the terrain there is full of big ups and downs and tends to be eroded in filtering. It will cause larger errors than flat ground when interpolating the missing part for DEM raster. Our method determines whether this part of DSM is mountainous by the slope of rough DEM. First, a filtering operation named “Initial Filtering” in the flowchart (Figure 3) is performed to obtain rough DEM. Morphological filtering method is adopted for initial filtering with parameters set to “strict” to remove buildings as completely as possible, although the summits and ridges of mountains will be eroded. For our DSM with 2 m resolution, we set the “max window size” as 100 and the “max distance” as 8 when using the morphological filter in Point Cloud Library (PCL) (Rusu and Cousins 2011; Zhang et al. 2003). Then the result is interpolated to generate rough DEM.

Considering there may still be some remaining points of large buildings, we down-sample the rough DEM to 50 m grids to reduce the effect of them. Finally, this down-sampled DEM is used for calculating the slope (Figure 6). In the proposed method, the slope threshold is given 16° based on experiments. That means in the current sliding window for filtering (Figure 3), if there is more than one pixel whose slope value is larger than the threshold, this part of DSM will be viewed as a mountainous area.

Figure 4. Building segmentation result of data set 1. There are some inaccurately and incorrectly segmented parts (in red rectangle in (a)) with corresponding orthophotos. Some segmented cases exceeding the thresholds with areas that are too large and too small are shown in (b) with red outlines. They will be eliminated in the final segmented image.

Figure 5. The relative slope of DSM of data set 1.

Figure 6. The slope of down-sampled rough DEM of data set 1.
Ground and Nonground Points
In DEM extraction, the ground points of DSM point cloud are the measurements located on the bare earth. In contrast, the nonground points are on the surfaces of nonground objects, mainly including buildings and vegetation, etc. However, unlike laser scanning, it is almost impossible for optical satellites to observe the ground surface and estimate the height of trees through foliage. So in this paper, the points of dense vegetation covering a large area like forests are considered as terrain and will not be filtered out. For example, if the forest-covered part on a mountaintop of DSM is removed, the terrain there can only be interpolated from surrounding points, probably causing larger errors than the effects of tree height.

Sliding-Window Filtering with Adaptive Parameters
As stated above, it is unreliable to determine the boundary between built-up areas and mountainous areas, so we are unable to process point cloud of buildings and mountains with different parameters separately. To reduce the error caused by this, we perform the filtering operation in a sliding-window way. That is, only filtering the point cloud inside the current sliding window. The distance between two adjacent sliding windows is set small in order to repeatedly process a specific point many times. The multiple results of being classified as ground or nonground can be more accurate and robust.

When the sliding window is moving, the filtering parameters are adjusted adaptively according to the building segmentation result and the slope of rough DEM and DSM as mentioned above. After we obtain the building segmentation result \( seg(x, y) \) from orthophoto, a simple discriminant function \( \alpha \) is given to determine if a part of \( seg(x, y) \) is a built-up area:

\[
\alpha \left( seg_{x}, b_{y} \right) = \begin{cases} 
1 & \text{if } b_{x} \geq b_{y} \\
0 & \text{if } b_{y} < b_{y}
\end{cases}
\]  

(3)

where \( seg_{x} \) represents the part of \( seg(x, y) \) inside the sliding window \( P \) and \( b_{y} \) represents the number of detected buildings in \( g_{y} \). We first find closed contours in \( g_{y} \) and then eliminate those with areas that are too large or too small (like Figure 4), which are usually wrong detection or very small buildings. \( b_{y} \) is the threshold for the number of buildings, and based on it the function \( \alpha \) returns 0 or 1 to describe whether this part is a built-up area. In our method, under the condition that the resolution of DSM and orthophoto is 2 m and the size of sliding window is 400 m, the threshold \( b_{y} \) is given 5. This threshold may be also influenced by the quality of orthophoto, the performance of building segmentation method, and the training data set. So, the discriminant function and threshold can be optimized in further studies.

Similarly, for the relative slope of rasterized DSM dsmslp\((x, y)\), we propose function \( \beta \) to check whether this part of DSM is built-up:

\[
\beta \left( dsmslp_{x}, e, r_{y} \right) = \begin{cases} 
1 & \text{if } r_{y} \geq r_{y} \\
0 & \text{if } r_{y} < r_{y}
\end{cases}
\]  

(4)

where \( dsmslp_{x} \) means the \( dsmslp(x, y) \) inside the sliding window \( P \) and \( e \) represents the relative building extraction threshold, and if the relative slope value of a pixel is larger than \( e \), it will be labeled as “building-wall” pixel. \( r_{y} \) represents the number of “building-wall” pixels and \( r_{y} \) is the threshold for \( r_{y} \). If \( r_{y} \) exceeds this threshold, \( \beta \) function will return 1 and \( dsmslp_{x} \) will be considered a built-up area. To reduce the effect of noise, we set \( e \) at a slightly large value, 110, and set \( r_{y} \) as 20 based on the characteristics of our data and experiments.

The proposed method performs the “initial filtering” operation with “strict” parameters to generate rough DEM without any nonground points left. Then we find the mountainous area based on the slope of down-sampled DEM dsmslp\((x, y)\). A discriminant function \( \gamma \) is given:

\[
\gamma \left( dsmslp_{x}, s_{x} \right) = \begin{cases} 
1 & \text{if } s_{x} \geq s_{x} \\
0 & \text{if } s_{x} < s_{x}
\end{cases}
\]  

(5)

where \( dsmslp_{x} \) is dsmslp\((x, y)\) inside the sliding window \( P \) and \( s_{x} \) represents the maximum slope value of dsmslp\((x, y)\). \( s_{x} \) is the threshold for \( s_{x} \). If \( s_{x} \) is larger than \( s_{x} \), function \( \gamma \) will return 1 and the corresponding part of DSM is likely to be a mountainous area.

We perform morphological filtering on DSM inside the sliding window, with the filtering parameters dynamically adjusted. According to Zhang et al. (2003) and the resolution and precision of our experimental data, we set the cell size as 2 m and the initial elevation difference threshold \( h_{0} \) as 3 m, using the default value of PCL for the initial window size (or the base of an exponential function) and the terrain slope. In the current sliding window \( P \), the maximum window size \( W_{\text{max}} \) and the maximum elevation difference threshold \( h_{\text{max}} \) are given by:

\[
W_{\text{max}} = W_{0} + \varphi_{1}W_{1}
\]

\[
h_{\text{max}} = h_{0} + \varphi_{1}h_{1}
\]

(6)

(7)

where \( W_{0}, W_{1}, \) and \( h_{1} \) are respectively set as 20, 60, and 7 in our experiments. \( \varphi_{1} \) is the “building factor” calculated by the following formula:

\[
\varphi_{1} = 0.5 \alpha + 0.5 \beta + 0.25(1 - \alpha)(1 - \beta) \quad 0 \leq \varphi_{1} \leq 1
\]

(8)

where \( \alpha, \beta, \gamma \) are the discriminant functions mentioned above. Finally, we count the number of “votes” of multiple filtering. For efficiency, the ground points classified by “initial filtering” are accepted, as all nonground points are removed. Therefore, we retrieve the other ground points from the removed points by “initial filtering”. The removed point cloud is set as inited\((x, y)\) and the preserved point cloud of the current sliding window after the morphological filtering operation with adaptive parameters as g\((x, y)\). For one point \((x', y')\) in inited\((x, y)\), we assume the total number of sliding windows that involves \((x', y')\) is \( n \), and the results of them are g\((x, y)\), \( 1 \leq n \). A function \( \delta \) is used to check whether \((x', y')\) belongs to g\((x, y)\):

\[
\delta \left( g_{x'}, (x', y') \right) = \begin{cases} 
0 & \text{if } (x', y') \in g_{x'}(x, y) \\
1 & \text{if } (x', y') \notin g_{x'}(x, y)
\end{cases}
\]  

(9)

The number of votes \( V(x', y') \) for \((x', y')\) is calculated by the following formula:

\[
V(x', y') = \sum_{i=1}^{n} \delta \left( g_{x'}(x', y'), (x', y') \right) / n \quad 0 \leq V(x', y') \leq 1
\]

(10)

In our experiments, we set the threshold for \( V \) as 0.3. If \( V \) is larger than this threshold, \((x', y')\) will be labeled as a nonground point.

Experiments and Discussion
Accuracy Assessment Measure
To quantitatively validate the accuracy of our DEM filtering method, we employ Type I errors, Type II errors, total errors, and Kappa coefficient according to Congalton (1991) and Sithole and Vosselman (2004). Type I errors (i.e., omission errors) represent the rejection of ground points and Type II errors (i.e., commission errors) represent the acceptance of
nonground points as ground surface. Total errors describe the overall accuracy of this algorithm, and the Kappa coefficient is a robust measure of overall agreement of reference (ground truth) and filtered result.

Results
In our experiments, we compare our algorithm with several major methods, including progressive morphological filtering, cloth simulation filtering, and PCI Geomatica software. Both “relaxed” and “strict” parameters are respectively employed for progressive morphological filtering (Mor1, Mor2) and cloth simulation filtering (CSF1, CSF2). A set of “moderate” parameters is employed in PCI Geomatica, and namely it is a compromise between removing buildings and preserving terrain. The ground truth data are produced manually as mentioned above. The comparison of errors and Kappa coefficient for various methods on data set 1–6 is shown in Figure 7. As we can see, our method has larger Type I errors and smaller Type II errors than other methods, which indicates our method can remove nonground points more completely, while some ground points are also incorrectly removed.

The methods with “relaxed” parameters like CSF1 and Mor1 have small Type I errors but large Type II errors because they tend to preserve more ground points together with some nonground points. For some data sets, PCI, CSF2, and Mor2 achieve quite good performance in Type II errors, total errors,

Table 1. Calculation equations of errors and Kappa coefficient. In the table, $P_0$ is the accuracy of observed agreement and $P_c$ is the estimate of chance agreement.

<table>
<thead>
<tr>
<th>Classified Data</th>
<th>Ground Points</th>
<th>Nonground Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference data</td>
<td>$a$</td>
<td>$b$</td>
</tr>
<tr>
<td>$e$</td>
<td>$c$</td>
<td>$d$</td>
</tr>
<tr>
<td>Number of points</td>
<td>$e = a + b + c + d$</td>
<td></td>
</tr>
<tr>
<td>Type I errors</td>
<td>$b/(a + b)$</td>
<td></td>
</tr>
<tr>
<td>Type II errors</td>
<td>$c/(c + d)$</td>
<td></td>
</tr>
<tr>
<td>$P_o$</td>
<td>$(b + c)/e$</td>
<td></td>
</tr>
<tr>
<td>$P_c$</td>
<td>$[(a + b)(a + c) + (c + d)(b + d)]/e^2$</td>
<td></td>
</tr>
<tr>
<td>Kappa coefficient</td>
<td>$(P_o - P_c)/(1 - P_c)$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 7. Errors and Kappa coefficient for various methods on data set 1–6. From the perspective of quantitative assessment, our method achieves slightly better performance in accuracy and stability, with more ground points removed (larger Type I errors) and less nonground points preserved (smaller Type II errors).
and Kappa coefficient, but their results are poor on other data sets. On the whole, the total errors and Kappa coefficient of our method are slightly better than others. From the perspective of quantitative assessment, the superiority of our method seems very limited. However, when qualitatively assessing the filtered results and interpolated DEM, the advantage of our method is obvious. As shown in Figure 8, Figure 9, and Figure 10, our method can get smooth ground surface of DEM in built-up areas and preserve the topographic feature in mountainous areas at the same time.

The results above are due to the drawback of Type I errors, for it assumes the points on a relatively flat ground surface are of the same importance with those on mountains. In fact, the points of areas with drastic topographic relief are more significant in preserving terrain. Because if some points in a flat area are filtered out by mistake, the missing part can be
Figure 9. Filtered results and detail views of DEM for various methods on data set 2.
interpolated with little error based on the neighboring points when generating DEM. For the points in mountainous areas, the interpolation may be inaccurate if they are missing, since the height difference between two close-range points can be very large on mountains. Therefore, although the Type I errors of our method are large, the terrain of mountains is preserved quite well (Figure 8, Figure 9, and Figure 10 data set 4, 6). For areas like the metropolis business district shown in Figure 10 data set 5, some closely spaced skyscrapers are merged together in DSM, causing it to be difficult to distinguish them from hills geometrically. Based on the building segmentation on orthophoto, they can be recognized and removed.

Conclusions
In this paper, we developed a novel and automated DEM filtering method for high-resolution optical satellite DSM. Our method fully utilizes the building segmentation results of the corresponding orthophoto derived from original images and DSM. Based on building segmentation results by YOLACT and the slope of DSM and rough DEM, the proposed method first estimates the probability of being a built-up area or mountains for DSM inside the current sliding window, and then performs morphological filtering with adaptive parameters. The final filtered results are determined by the votes of multiple filtering. For a DSM point, if the number of its “non-ground” votes exceeds a certain threshold, this point will finally be filtered out. Both quantitative and qualitative evaluation of six data sets shows that this method can filter out buildings to get smooth ground surface and preserve terrain in mountainous areas at the same time. Compared with other popular methods, our algorithm achieves smaller Type II errors, competitive total errors, and Kappa coefficients with slightly larger Type I errors.

Despite the superior performance, we are aware that some limitations still exist. We take no account of the effects of forests and view them as part of terrain, causing errors in generating DEM. Moreover, the parameters and thresholds of the proposed method depend on the characteristics of satellite images, training samples for building segmentation and experience. For further research, we plan to verify and improve the accuracy and adaptability of this method through experiments on various satellite data.

Acknowledgments
This research was supported by grants from Key Scientific and Technological Research and Development Projects of Jilin (20200401093GX), National Key Research and Development Program (No. 2019YFE0127000) and National Key Research and Development Program (No. 2020YFA0714104).
References


A Combined Unmixing Framework for Impervious Surface Mapping on Medium-Resolution Images with Visible Shadows

Hui Luo and Nan Chen

Abstract
Spectral unmixing methods with medium-resolution remote sensing images have become the main approach to mapping urban impervious-surface information. However, as more tall buildings appear, numerous visible shadows exist in medium-resolution images; these have usually been ignored in previous research, but they seriously affect accuracy. To solve this problem, we propose a combined unmixing framework to extract impervious surface in nonshadow and shadow areas, using linear and nonlinear unmixing models, respectively. First shadow is separated from nonshadow. Then a nonlinear unmixing method is selected to map impervious surface in shadow, which is more suitable to the complex imaging environment in shadow, and a classic linear unmixing model in nonshadow. Through experimental tests, the proposed combined unmixing framework is shown to effectively reduce error in two study areas compared with classical unmixing methods.

Introduction
Impervious surface is composed of objects that prevent water from penetrating into the soil, most of which are human-made objects such as buildings, roads, squares, and so on (Arnold and Gibbons 1996; Slonecker, Jennings and Garofalo 2001; Q. Weng 2012; Gong et al. 2020). It is not only used to reflect the level of urbanization in a city, but also has a significant impact on the urban environment (Arnold and Gibbons 1996; C. Wu and Murray 2003), affecting things such as evaluation of hydrological effects (Shao et al. 2019), non-point-source pollution of water (Conway 2007; Kaspersen et al. 2017), the urban heat island effect (Yuan and Bauer 2007; Deng and Wu 2013; Yang et al. 2020), and urban ecological diversity (Q. Weng 2020). Hence, it is meaningful to extract impervious-surface information accurately.

Remote sensing technology has the advantages of providing all-weather, large-area, repeatable observations, which have been widely used in many scientific studies and applications (Liu et al. 2019; Jiang, Jiang and Jiang 2020; Shao et al. 2020). Extraction of impervious-surface areas at different scales can be implemented through remote sensing data (Gallo and Xian 2016; Liu et al. 2019; Wang and Li 2019; Shao et al. 2020). Medium- and high-resolution images are often used as the main data source for extracting impervious surface in urban areas. Generally, impervious surface can present richer geometric features in higher resolution remote sensing images, such as rectangular-like roofs (Wang and Li 2019). However, the acquisition cost of high-resolution remote sensing images is relatively high, and the image data are complex and difficult to process (Sawaya et al. 2003). In contrast, medium-resolution images have become a suitable data source for monitoring the distribution of urban impervious surface at different scales due to their low cost, wide observation range, and multiple acquisition channels (Q. Weng 2012; Shao and Liu 2014; L. Zhang, Weng and Shao 2017). However, due to the resolution limitation of medium-resolution images and the complexity of urban features, there are a large number of mixed pixels in the images, which brings challenges to the extraction of impervious surface (Cracknell 1998; Demarchi et al. 2012; Weng 2012). For this situation, the traditional per-pixel classification method has difficulty obtaining accurate proportions of impervious surface, since it just classifies each pixel into a specific category label (Phinn et al. 2002). A large number of spectral unmixing methods have been proposed to solve this problem.

Spectral unmixing is a technique that can decompose mixed pixels into corresponding pure components and obtain the proportion of each component in the mixed pixel (Keshava and Mustard 2002; H. Li et al. 2021), so as to obtain more accurate quantitative results for land cover. In the past decades, a variety of unmixing methods for impervious surface have been proposed (Yu et al. 2017), which can be divided into two categories based on different assumptions: linear and nonlinear (Petrou 2015). Linear unmixing methods are mainly based on the basic assumption that the spectrum of each component is linearly weighted to form the spectrum of the mixed pixel. Linear spectral mixture analysis (LSMA) is one of the most classic linear unmixing methods, widely used in subpixel mapping of urban impervious surface (C. Wu and Murray 2003; Li, Moran and Hetrick 2011; Q. Weng 2012; Lu et al. 2014). Subsequent research on linear unmixing methods has extended LSMA, such as normalized LSMA (C. Wu 2004), hierarchical SMA (Sun et al. 2017), geographically weighted SMA (Deng and Wu 2013), and multiple-end-member SMA (MESMA; F. Weng and Pu 2013). In contrast, nonlinear unmixing methods consider that the formation of a mixture is a nonlinear process, usually accompanied by the occurrence of multi-point scattering (Ray and Murray 1996; Gilabert, García-Haro and Meliá 2000; Keshava and Mustard 2002). Some studies have shown that the unmixing accuracy of nonlinear methods such as artificial neural networks (ANNs), fuzzy classification, and support vector machines (SVMs) is better than that of linear methods in some complex scenes (J. Zhang and Foody 2001; B. Wu, Zhang and Li 2006; Q. Weng and Hu 2008), such as compact soil composed of grains of...
sand of different composition, canopies of different heights, and shaded areas (Keshava and Mustard 2002; Somers, Tits and Coppin 2014; Yu et al. 2017).

When using these linear and nonlinear methods to unmix impervious surface, it is necessary to determine a model that can describe the urban composition. The V-I-S model (Ridd 1995) is composed of three typical land covers—vegetation, impervious surface, and soil—and is widely used in the extraction of urban impervious surface (Phinn et al. 2002; Fan, Deng and Zhu 2013; F. Xu et al. 2019). Due to the large variation in the materials that make up impervious surface, it will show different brightnesses in images. Thus the four-end-member VHLS model (vegetation, high-albedo impervious surface, low-albedo impervious surface, and soil) was proposed, extending impervious surface into high- and low-albedo components (C. Wu and Murray 2003) and further reducing its unmixing error. The VHLS model is often used for unmixing impervious surface in medium-resolution images, such as the Landsat series of images (Hu and Weng 2009; C. Wu 2004; F. Weng and Pu 2013; H. Li et al. 2019; Zhao et al. 2020).

Although these methods have been widely used in urban impervious-surface unmixing, there are still some key issues that affect accuracy (Wang and Li 2019). For example, bright impervious surface is easily confused with bare soil, and water bodies and shadow are usually misclassified as dark impervious surface, leading to some errors in estimation (C. Wu and Murray 2003; Lu and Weng 2006; Lu et al. 2011; L. Li, Lu and Kuang 2016). Among these problems, the confusion between shadow and dark impervious surface has rarely been fully taken into consideration (Dennison and Roberts 2003; C. Wu 2004; Hu and Weng 2009). There are two common strategies to solve this problem in medium-resolution images. First, ignoring the shadow problem in the study site generally is based on the assumption that the ratio of shadow area to study site is relatively low, so that the shadow can be ignored in unmixing the impervious surface (C. Wu and Murray 2003). Second, some studies regard shadow as an independent end member in impervious-surface unmixing (Jie and Liu 2009; F. Weng and Pu 2013). Generally, these strategies still cause some problems. Ignoring shadow and unmixing the image directly will cause the shadow to be classified as dark impervious surface, resulting in an overestimation in the final result. Unmixing the shadow as an end member alone may lose impervious surface under shadow, because the spectral characteristics of impervious surface under shadow areas are different from those under nonshadow areas, resulting in underestimation in the final result (Yang and Li 2015; Yang and He 2017). With the rapid increase of urbanization, a large amount of visible shadow can also be clearly seen in some medium-resolution images. And it is meaningful to map the impervious surface in shaded areas, even though shadows bring challenges to estimating real land cover types in these scenes (C. Wu 2004; Lu et al. 2011; L. Li et al. 2016; Lin et al. 2019).

Therefore, in this article we propose an impervious-surface unmixing framework considering impervious surface under shadow during the unmixing process for medium-resolution images. We use the idea of separate processing of shadow and nonshadow areas (P. Li et al. 2011; T. Zhang and Huang 2018). First, the entire image was divided into nonshadow and shadow areas, and different end-member models were selected for these two different scenarios (Yang and He 2017). Then linear unmixing was applied in nonshadow areas and nonlinear unmixing in shadow areas. The final impervious-surface distribution result was obtained by combining the results for nonshadow and shadow areas. Experiments showed that the combination of linear unmixing in nonshadow areas and nonlinear unmixing method in shadow areas can not only avoid confusion between shadow and dark impervious surface but also effectively extract the amount of impervious surface under shadow, thereby improving the overall accuracy of impervious-surface unmixing.

**Study Area and Data Sets**

In order to test the performance of the combined framework for unmixing impervious surface with shadows, two typical cities—Wuhan and Hangzhou—were selected as the study sites. They are the capital cities, respectively, of Hubei and Zhejiang provinces in China. Wuhan and Hangzhou are important cities in central and eastern China, with relatively high population density and urbanization levels and relatively prosperous economies. The Wuhan research area comes from the city’s Jiangxia District, where urbanization has been vigorously promoted in recent years. The Hangzhou research area in in the city’s Xihu District, one of its central urban areas, with dense buildings. There are rich shadow scenes in these two study areas, which are ideal for unmixing urban impervious surface considering shadow. Figure 1 shows the location of study sites used.

![Figure 1. The study sites: (a) Wuhan and (b) Hangzhou.](image-url)
by ENVI software. The true-color sample Landsat 8 OLI data for the study sites, at 15 m resolution, are shown in Figure 2a–b, at 400×400 and 240×240 pixels, respectively. Two areas (red boxes A and B in Figure 2a–b) were selected as representative shadow scenes to qualitatively compare the results of different unmixing methods. It should be noted that before the 15-m Landsat 8 OLI images were used to unmix impervious surface, the water pixels in the image had to be masked based on the modified normalized difference water index (H.-Q. Xu 2005).

For each study area, a high-resolution remote sensing image acquired at close to the same time as the Landsat 8 OLI image was selected as a reference image for assessing the abundance of impervious surface. The reference image used for the Wuhan area was a Gaofen-2 image at 1 m spatial resolution after fusion with its panchromatic band (Figure 2c), 6000×6000 pixels, covering the same area as Figure 2a. The reference image used for the Hangzhou area was a WorldView-2 image that also underwent fusion with its panchromatic band, but had a higher resolution of 0.5 m (Figure 2d). It was 7200×7200 pixels, covering the same area as Figure 2b. The land cover classification results from the high-resolution images are sufficient as the ground truth for the unmixing results acquired from the 15-m Landsat 8 OLI images; however, classification under shadow of high-resolution images has difficulty obtaining a convincing result with very high accuracy. Therefore, we acquired all types of impervious surface in the reference high-resolution images of these two cities by visual interpretation. In this interpretation step, the impervious-surface area under shadow was selected with auxiliary data from Google Earth. Finally, the whole impervious-surface binary map with high precision was obtained. Before the impervious-surface of the high-resolution images was acquired, it was necessary to geo-register them with their corresponding Landsat 8 OLI images to ensure the accuracy of the impervious-surface evaluation. In this article, we used the automatic registration module in ENVI to complete the image registration work. Since resampling during registration will change the spectral value of the image, we used the Landsat 8 image as the reference image, the high-resolution image as the image to be registered, and cross correlation as the matching algorithm. The final geo-registration error of the two study areas is less than one pixel of the reference image.

**Methodology**

A combined strategy was proposed to extract the impervious-surface abundance, consisting of three steps (detailed in Figure 3). First, the 15-m Landsat 8 OLI images were divided into two different scenes by shadow detection: nonshadow and shadow areas. Second, considering that impervious surface presents...
different spectral characteristics under shadow and in non-shadow, the impervious surface in the nonshadow area was extracted using the commonly used VHLS model (C. Wu and Murray 2003). Because land covers in shadow are not easy to distinguish, we did not use the general four-end-member model to unmix the impervious surface in shadow, but instead used the two-end-member IP (impervious surface, pervious surface) model to describe the mixed pixels in the shadow area. Two representative linear unmixing methods—LSMA (Zhao et al. 2020) and MESMA (F. Weng and Pu 2013)—were selected for unmixing nonshadow area, and the classical nonlinear methods of SVM (Brown, Gunn and Lewis 1999) and ANN (Hu and Weng 2009) were used for impervious surface in shadow. Finally, we combined the results for the nonshadow and shadow areas to form a complete impervious-surface abundance map of the entire image. The four types of unmixing results were also assessed without regard to shadow. In the following sections, each step of the method will be introduced in detail.

We combined high-resolution images and Google Earth images to help in selecting pure pixels directly from 15-m Landsat 8 OLI images. The Landsat 8 OLI images were divided into shadow and nonshadow areas, so the end-member models used in the two scenes were different. Similarly, due to the different size and distribution of land covers in the two study areas, the number of selected samples was different. There were four types of end member in nonshadow areas: vegetation (V), high-albedo impervious surface (H), low-albedo impervious surface (L), and soil (S). In the shadow area, two types were selected: impervious surface (I) and pervious surface (P). The distribution of all candidate end-member samples in the nonshadow and shadow areas is shown in Figure 4. To highlight the distribution, we added a less-transparent layer onto the original image. Because the size of study sites in Wuhan and Hangzhou are different, the sample numbers selected for each class are also different. In the Wuhan research area, we used 100 end members of each type; in the Hangzhou research area, the number was 20. LSMA directly uses the average end-member spectrum of all samples to unmix, whereas MESMA uses the combination of spectra in the spectral library to iteratively unmix and select the one with the smallest unmixing error. If there are many combinations, MESMA is time-consuming. In order to reduce the time spent in iteration and ensure the diversity of the spectra in the spectrum library, we used stratified random sampling to select five spectra from each end-member type to establish the MESMA unmixing spectrum library. With the

---

**Figure 3.** Flowchart of the proposed combined strategy, with three steps: (1) shadow detection; (2) impervious-surface unmixing test in shadow area with nonlinear method and nonshadow area with linear method; (3) combination of unmixing results for nonshadow and shadow areas.

**Figure 4.** Distribution map of candidate end-member samples in shadow and nonshadow areas: (a) Wuhan; (b) Hangzhou.
nonlinear unmixing methods (SVM, ANN), all samples were used for training and then unmixing.

### Shadow Detection
The 15-m Landsat 8 OLI images were divided into nonshadow and shadow areas by shadow detection. Property-based shadow detection is widely used in urban areas (Tsai 2006; Chung, Lin and Huang 2009) because it makes full use of the spectral characteristics of the image without the need for auxiliary data. It mainly implements the idea of enhancing the shadow characteristics first and then using threshold segmentation to extract them. We used the shadow detector index (SDI; Mostafa and Abdelhafiz 2017), a type of property-based shadow detection operator, to perform shadow extraction on the 15-m Landsat 8 OLI images:

\[
SDI = \frac{(1 - PC_i) + 1}{((G - B) \times R) + 1}
\]

where \( R \), \( G \), and \( B \) are the respective normalized values of the red, green, and blue bands and \( PC_i \) is the normalized value of the first principal component.

### Linear Mixing Model
The linear mixing model (LMM) is the most commonly used model for unmixing the mixed pixels of medium-resolution multispectral images. It follows the assumption that each photon received by the sensor interacted only with a land cover before it reached the sensor, and it therefore takes the reflectance spectrum of mixed pixels as a linear weighted combination of pure-pixel spectral signatures; these pure pixels are also called end members (Keshava and Mustard 2002). The weight is expressed by the proportion of the area occupied by each end member in the mixed pixel. The LMM can be expressed as

\[
r = Mf + e
\]

where \( r \) is a \( k \times 1 \) vector of mixed reflectance spectra, \( k \) is the number of image bands, \( M \) is a \( k \times n \) matrix whose columns represent the respective reflectance signal of each end member, \( n \) is the number of end members, \( f \) is an \( n \times 1 \) vector of the proportions of each end member in the mixed pixel, and \( e \) is a \( k \times 1 \) error vector consisting of sensor and model system error. Equation 2 is usually calculated by least-square estimation. To guarantee the physical significance of the end-member abundance, some constraints should be considered. Commonly used ones include the sum-to-one constraint and the non-negative constraint:

\[
\sum_{i=1}^{n} f_i = 1 \quad \text{and} \quad f_i \geq 0
\]

where \( f_i \) is the proportion of end member \( i \) and \( n \) is the number of end-member categories. Previous research has shown that adding these constraints to the LMM solution can effectively improve unmixing accuracy (Heinz and Chang 2001). In this case, the model will be solved by a method called a fully constrained least-squares (FCLS) solution. LSMA with FCLS was chosen as one of the linear unmixing approaches in this paper.

LSMA generally uses a fixed end-member matrix on the entire image. However, in complex urban scenes with different illumination conditions, various land covers usually show different degrees of spectral variation, called intraclass and interclass variability. When a fixed end-member matrix is used in complex scenarios, this spectral variation leads to imprecise abundance results. To overcome LSMA’s shortcomings with regard to spectral variability, MESMA was proposed (Roberts et al. 1998). MESMA is one of the widely used methods for mitigating intraclass spectral variability. It allows for the type and number of end members to be changed while unmixing mixed pixels, and selects the final result according to given optimal criteria. MESMA is an iterative optimization and unmixing method, and it usually uses LSMA for its basic unmixing model. The concept of MESMA is to evaluate the calculation results of all LSMA models based on error statistics, then select the best-fitting LSMA model and use its unmixing results as the ultimate unmixing results. Therefore, in addition to LSMA with FCLS, we used MESMA as our second LMM. To evaluate the optimal result with MESMA, we used the root-mean-square of the model error (RMS errors; Dennison and Roberts 2003). For comparison with the simple SML model, we chose LSMA with FCLS as the basic unmixing model for MESMA and selected the smallest RMS errors value of each pixel as its final result.

### Nonlinear Mixing Model
In contrast to the LMM, a nonlinear mixing model considers a large amount of real nonlinear mixing phenomena in the image. Especially when the scattering of photons interacting with multiple land covers cannot be ignored, or there are more shadows in the image, the mixture will show more nonlinearity (Keshava and Mustard 2002). A nonlinear mixing model is then more consistent with the real sensor imaging situation, and provides more accurate results than an LMM. A variety of nonlinear mixing models exist—including geometrical optics, probability-based, neural-network, and bilinear models—which can generally be divided into two categories: physical-based and data-driven models (Heylen, Parente and Gader 2014). Physical-based models attempt to physically model spectral mixing processes with multiple scattering phenomena, such as introducing a cross-multiplication term for the end members or a polynomial term in a bilinear model to represent multiple scattering (Altman et al. 2012). Although they have been successfully applied in some studies, each of these models has its own limitations. The biggest limitation is that they contain more parameters, so that often they require many image bands to solve. Therefore, physical-based nonlinear mixing model are widely used with hyperspectral rather than multispectral images.

Data-driven models, on the other hand, are used for building a fixed expression with parameters but to train a nonlinear mixing model based on the data or directly unmix images based on training data. Therefore, they count as supervised machine-learning techniques. Another advantage of this type model is that they have been deemed to have some potential for mitigating end-member variability when sufficient representative training data are given. Considering that the unmixing data source is Landsat 8 multispectral images, we selected a data-driven model as the nonlinear mixing model to test, and chose two supervised algorithms—SVM and ANN—for unmixing. Specifically, in the SVM model we used the multi-class posterior probability of the trained SVM classifier to estimate the abundances of various end members of the mixed pixels. Generally, the SVM model is often used for classification (X. Li et al. 2015). However, in some studies, mixed-pixel decomposition has been proven to be a special case based on posterior probability classification (Kolaczyk 2003). Therefore, we can consider the posterior probability value of the SVM classification model as the abundance value of each end member in the mixed pixel. As with the SVM model for
classification, we chose the corresponding end members as training data and specified the corresponding category (four in nonshadow areas and two categories in shadow). Because the direct output value of the SVM model has not been calibrated, it is necessary to perform a regression operation on its output value to calculate the posterior probability value corresponding to each class:

\[ p(C \mid x) = \frac{1}{1 + e^{-(A \cdot x + B)}} \]  

where \( p(C \mid x) \) is the posterior probability of category \( C \) calculated based on the training data \( x \), \( g(x) \) is the output value of the SVM model, and \( A \) and \( B \) are regression parameters estimated based on \( x \). The LIBSVM library (Chang and Lin 2011) was chosen to solve the posterior probability of the SVM multi-class classification model.

The ANN model often shows good results for some complex nonlinear tasks by simulating the neuronal mechanism of the brain (Han et al. 2018). Many previous studies suggest that ANN models can be used in image-unmixing tasks, especially for complex scenes (Hu and Weng 2009; Ahmed et al. 2017). We chose the classic three-layer perceptron architecture to form an ANN model with multi-layer perceptron (Q. Weng and Hu 2008). This model is one of the most commonly used neural-network architectures, and has been proven to have good performance in image unmixing. The three-layer perceptron model is composed of an input layer, a hidden layer, and an output layer, each containing several neurons. When it is applied to image unmixing, the number of neurons in the input layer is equal to the number of image bands, and the number in the output layer is equal to the number of end members. In the hidden layer, the number of neurons is usually determined subjectively by expert experience. Referring to the literature (Chettri, Cromp and Birmingham 1992; Yu et al. 2017) and comprehensively considering the training cost and degree of nonlinearity of the ANN model, we used \((2s + 1)\) neurons in our hidden layer, where \( s \) is the number of image bands. For each neuron in this layer, the output of the neuron in the previous layer is multiplied by the weights on the connected edges and summed as the input of the neuron in this layer. Then an activation function was utilized to generate the output of this layer, namely the new input of the neuron connected to the next layer:

\[ y_q = F \sum_{j=0}^{m} w_{pq} x_p \]  

where \( y_q \) is the output of the neuron \( q \) in the current layer, \( F \) is the activation function of \( q \), \( w_{pq} \) is the weight connecting neuron \( p \) of the previous layer and \( q \), \( x_p \) is the output of \( p \), and \( m \) is the number of neurons in the previous layer.

In order to introduce nonlinearity into the ANN model, the activation function needs to be set to a nonlinear form. We chose the sigmoid function as the activation function for the hidden and output layers; the value of the output layer ranges from 0 to 1. In order to satisfy the sum-to-one constraint, the result of output layer must be normalized to represent the abundance value of each end member.

Accuracy Assessment

To evaluate the accuracy of different methods for unmixing impervious surface, it is necessary to calculate and compare some statistical indicators of the estimated and reference impervious surface. We used three common quantitative indicators for evaluation, namely the root-mean-square error (RMSE; different from RMSE errors), the correlation coefficient \( R \), and the systematic error (SE); detailed calculation formulas are given by Sun et al. (2017). RMSE and SE quantitatively evaluate the accuracy of the method through error statistics, whereas \( R \) is used to describe the linear relationship between the estimated and reference abundance of impervious surface. Specifically, RMSE is a common indicator used to measure the difference between the estimated and reference values of impervious surface. In comparing the accuracy of different unmixing methods, the method with the lower RMSE has better performance. SE reflects whether the overall trend of estimation is over or under that calculated by the deviation between the estimated and reference impervious surface; the smaller the absolute value of SE, the better the overall performance of the method. \( R \) is used to evaluate the performance of different methods by describing the linear relationship between the estimated and reference values; the higher the \( R \), the better the linear relationship.

Later, we will separately evaluate the accuracy of the unmixing results for shadow and nonshadow areas to explore the performance of four unmixing methods (LSMA, MESMA, SVM, ANN), using only RMSE as a representative statistical indicator. We used block-by-block statistics to calculate RMSE results for the shadow and nonshadow areas of the Landsat 8 OLI image independently—that is, one pixel corresponds to 15×15 pixels on the Gaofen-2 image (Wuhan) or 30×30 pixels on the WorldView-2 image (Hangzhou). Since the combined impervious-surface unmixing results of the nonshadow and shadow areas will be merged to form a complete impervious-surface abundance map of the entire image, the results of the four methods should be compared without considering the influence of shadow. Therefore, we will use a larger block-statistics mode to count the impervious-surface results of the four types of combined method and the four unmixing methods without considering shadow, to further eliminate possible statistical errors caused by registration. A 3×3-pixel block is selected as the statistical unit in the 15-m Landsat 8 OLI images, and the high-resolution reference value corresponding to that unit is calculated according to the resolution ratio. Three statistical indicators (RMSE, \( R \), SE) were computed to assist in the comprehensive comparative analysis.

Results

Shadow Detection

The first step of the combined framework proposed in this article is to extract the shadow pixels in the 15-m Landsat 8 OLI images, dividing them into two different scenes—non-shadow and shadow—for the impervious-surface unmixing task. Based on the SDI operator, the shadow information in the Landsat 8 image was enhanced and the shadow pixels extracted by setting an appropriate threshold determined by the classic Otsu algorithm (Otsu 2007). There are many shadows visible in the Wuhan and Hangzhou study sites (Figure 2a–b). After shadow detection, the final shadow binary maps of the two research areas are shown in Figure 5a–b; most shadows have been detected. In the Wuhan study area, the proportion of shadow is about 5.6% of the entire image. Except for the shadow caused by dense buildings, the distribution is somewhat scattered. The proportion of shadow in Hangzhou reaches 13.1%, and since the study area contains a large number of tall and dense buildings, the distribution of shadow presents mainly large and dense features.

To check the accuracy of shadow detection, the criterion of the two-classification task was adopted as reference. Based on the idea of stratified random sampling, 200 pixels each were selected as the test sample in the nonshadow and shadow areas; then the overall accuracy and \( k \) statistic were used for
accuracy evaluation. The overall accuracy for Wuhan was 97%, $\kappa = 0.95$, and the overall accuracy for Hangzhou was 98%, $\kappa = 0.97$.

**Unmixing of Impervious Surface Under Nonshadow and Shadow Areas Separately**

Through shadow detection, the Landsat 8 image was divided into shadow and nonshadow areas, with different spectral characteristics for impervious surface. Therefore, different end-member models and data sources were required to unmix the impervious surface. Specifically, the VHLS end-member model was used in nonshadow areas; its impervious-surface abundance is the sum of the H and L end-member abundances. In shadow areas, a simple IP end-member model was used; the abundance of the I end member is the impervious-surface abundance. Two representative unmixing methods for linear (LSMA, MESMA) and nonlinear (ANN, SVM) models were selected and carried out in the nonshadow and shadow areas, respectively. In addition, two linear/nonlinear mixing models were used in the nonshadow/shadow areas to complete the comparative experiments. Finally, through block-by-block statistics, the RMSE value was used as a comparison index among the four methods for unmixing impervious surface in nonshadow and shadow areas. It should be emphasized that the two nonlinear unmixing methods (SVM and ANN) are based on data training, and the results of these models are related to the training environment and data quality, so they will have a certain volatility. Therefore, when using these two methods, we repeated experiment 10 times independently and used the average RMSE result.

Table 1 lists the RMSE results for the four unmixing methods in the nonshadow and shadow areas in Wuhan. In nonshadow, MESMA achieves the best result, with an RMSE value of 0.3635, and SVM the worst, with an RMSE value of 0.4542. In shadow, however, LSMA reaches an RMSE value as high as 0.5260, and MESMA reaches 0.4589. The accuracy of both of these linear methods in shadow is inferior to those of the nonlinear unmixing methods: ANN obtains the best result in shadow, with an RMSE of 0.3846, and SVM achieves an RMSE of 0.4292. The statistical results for the Hangzhou study area are shown in Table 2. In nonshadow areas, MESMA achieves the best result, with an RMSE value of 0.3592. The RMSE of LSMA is 0.3900, and that of SVM is 0.3708 (slightly better). ANN has the worst result, with an RMSE of 0.4074. Different from the Wuhan study area, the optimal unmixing method for impervious surface in shadow areas is SVM, with an RMSE of 0.3686. ANN has an RMSE of 0.3821, which is also better than the two linear methods: LSMA with an RMSE of 0.4292 and MESMA with an RMSE of 0.3971.

Table 1. Block-by-block root-mean-square error for nonshadow and shadow in Wuhan (1×1-pixel unit).a

<table>
<thead>
<tr>
<th>Unmixing Method</th>
<th>Nonshadow Area</th>
<th>Shadow Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMA</td>
<td>0.3740</td>
<td>0.5260</td>
</tr>
<tr>
<td>MESMA</td>
<td>0.3635</td>
<td>0.4589</td>
</tr>
<tr>
<td>SVM (average)</td>
<td>0.4542</td>
<td>0.4201</td>
</tr>
<tr>
<td>ANN (average)</td>
<td>0.3709</td>
<td><strong>0.3846</strong></td>
</tr>
</tbody>
</table>

ANN = artificial neural network; LSMA = linear spectral mixture analysis; MESMA = multiple-end-member spectral mixture analysis; SVM = support vector machine.

aBoldface indicates the best result in each scene.

Table 2. Block-by-block root-mean-square error for nonshadow and shadow in Hangzhou (1×1-pixel unit).a

<table>
<thead>
<tr>
<th>Unmixing Method</th>
<th>Nonshadow Area</th>
<th>Shadow Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMA</td>
<td>0.3900</td>
<td>0.4292</td>
</tr>
<tr>
<td>MESMA</td>
<td><strong>0.3592</strong></td>
<td>0.3971</td>
</tr>
<tr>
<td>SVM (average)</td>
<td>0.3708</td>
<td><strong>0.3686</strong></td>
</tr>
<tr>
<td>ANN (average)</td>
<td>0.4074</td>
<td>0.3821</td>
</tr>
</tbody>
</table>

ANN = artificial neural network; LSMA = linear spectral mixture analysis; MESMA = multiple-end-member spectral mixture analysis; SVM = support vector machine.

aBoldface indicates the best result in each scene.

From these accuracy results, we can conduct some preliminary analysis. In nonshadow areas, MESMA achieves the best results among the four methods, while the worst results both come from nonlinear methods (SVM for Wuhan and ANN for Hangzhou). Moreover, an obvious trend can be seen from the results in shadow: the results of the two nonlinear methods are always better than those of the linear methods. This is probably because the end members in shadow areas have a more nonlinear form, making the situation difficult for linear unmixing methods to simulate and leading to their poor results. For instance, the classic unmixing method LSMA, based on the linear assumption, always achieves the worst results.
in shadow in the two study areas. In addition, the accuracy of MESMA is always better than that of LSMA, whether in shadow or not. Since MESMA alleviates spectral variability through iteration operation, it can achieve better results than LSMA, with its fixed-end-member model. As for comparing SVM and ANN, since their accuracy has a greater relationship with the sample and training environment, it may be difficult to tell which method is better.

### Unmixing Impervious Surface with a Combined Method

We combined the unmixing results for the nonshadow and shadow areas into the impervious-surface abundance of the entire image. There are four possible combinations of results: LSMA + SVM, LSMA + ANN, MESMA + SVM, and MESMA + ANN, which are shown in Tables 3 and 4, respectively, for Wuhan and Hangzhou. The four individual methods were also used directly for the entire image of the study site without consideration of unmixing the shadow and nonshadow scenes separately; these are the comparative experiments displayed in Tables 3 and 4, using the VHLs end-member model and corresponding samples to directly unmix the impervious surface of the entire image. After comprehensive evaluation of three indicators for the Wuhan area, according to Tables 1 and 3, the optimal combined method is MESMA in nonshadow and ANN in shadow. The RMSE, R, and SE of the proposed combined method are 0.2477, 0.6664, and 0.0784, respectively. Among the four single methods conducted without separating shadow and nonshadow, MESMA performs the best in both two study sites (Tables 3 and 4), which is consistent with the results in nonshadow areas (Tables 1 and 2).

Table 3. Accuracy results of proposed combined method and comparative experiments in Wuhan (3x3-pixel unit).a

<table>
<thead>
<tr>
<th>Unmixing Method</th>
<th>RMSE</th>
<th>R</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMA</td>
<td>0.2860</td>
<td>0.5890</td>
<td>0.0838</td>
</tr>
<tr>
<td>MESMA</td>
<td>0.2511</td>
<td>0.6567</td>
<td>0.0842</td>
</tr>
<tr>
<td>SVM (average)</td>
<td>0.3389</td>
<td>0.5523</td>
<td>0.2164</td>
</tr>
<tr>
<td>ANN (average)</td>
<td>0.2685</td>
<td>0.6125</td>
<td>0.0876</td>
</tr>
<tr>
<td>LSMA + SVM (average)</td>
<td>0.2727</td>
<td>0.6069</td>
<td>0.0947</td>
</tr>
<tr>
<td>LSMA + ANN (average)</td>
<td>0.2712</td>
<td>0.6107</td>
<td>0.0927</td>
</tr>
<tr>
<td>MESMA + SVM (average)</td>
<td>0.2485</td>
<td>0.6641</td>
<td>0.0803</td>
</tr>
<tr>
<td>MESMA + ANN (average)</td>
<td>0.2477</td>
<td>0.6664</td>
<td>0.0784</td>
</tr>
</tbody>
</table>

| ANN = artificial neural network; LSMA = linear spectral mixture analysis; MESMA = multiple-end-member spectral mixture analysis; R = correlation coefficient; RMSE = root-mean-square error; SE = systematic error; SVM = support vector machine. |

| Table 4. Accuracy results of proposed combined method and comparative experiments in Hangzhou (3x3-pixel unit).a |

<table>
<thead>
<tr>
<th>Unmixing Method</th>
<th>RMSE</th>
<th>R</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSMA</td>
<td>0.3325</td>
<td>0.5637</td>
<td>0.2430</td>
</tr>
<tr>
<td>MESMA</td>
<td>0.2928</td>
<td>0.5914</td>
<td>0.1896</td>
</tr>
<tr>
<td>SVM (average)</td>
<td>0.3023</td>
<td>0.5977</td>
<td>0.2075</td>
</tr>
<tr>
<td>ANN (average)</td>
<td>0.3226</td>
<td>0.5752</td>
<td>0.2206</td>
</tr>
<tr>
<td>LSMA + SVM (average)</td>
<td>0.2952</td>
<td>0.6004</td>
<td>0.1988</td>
</tr>
<tr>
<td>LSMA + ANN (average)</td>
<td>0.2962</td>
<td>0.6061</td>
<td>0.2013</td>
</tr>
<tr>
<td>MESMA + SVM (average)</td>
<td>0.2584</td>
<td>0.6364</td>
<td>0.1498</td>
</tr>
<tr>
<td>MESMA + ANN (average)</td>
<td>0.2598</td>
<td>0.6391</td>
<td>0.1523</td>
</tr>
</tbody>
</table>

ANN = artificial neural network; LSMA = linear spectral mixture analysis; MESMA = multiple-end-member spectral mixture analysis; R = correlation coefficient; RMSE = root-mean-square error; SE = systematic error; SVM = support vector machine.

*Boldface indicates the best result in each scene.

According to Table 4, the optimal combined method for the Hangzhou area is MESMA in nonshadow and SVM in shadow, with an RMSE, R, and SE of 0.2584, 0.6364, and 0.1498. This combination in the Hangzhou image performs the best among all eight combinations across both study sites. Among the four single methods for the entire image, MESMA again performs the best, but LSMA has the worst performance in this study—perhaps because the Hangzhou study area has more shadow than the Wuhan study area, so that LSMA’s linear assumption leads to poor performance. As for the two nonlinear methods, different from in the Wuhan study area, SVM performs better than ANN, which may be due to the fact that the training data for these study areas are different.

The unmixing results for nonshadow and shadow areas were merged, and through experimental comparison of the two study areas, we found that the results of the combined method are generally better than those for either of the individual methods involved in the combined method. For example, the combination of LSMA in nonshadow and SVM in shadow outperforms both individual LSMA and SVM conducted for the entire image without subscene unmixing (division into shadow and nonshadow scenes). Comparing the two study areas, the optimal unmixing method for nonshadow is the same in both, whereas the optimal methods for shadow are different: MESMA and ANN for Wuhan and MESMA and SVM for Hangzhou. In both study sites, this optimal combined method outperforms MESMA implemented for the entire image. The difference in final accuracy between the optimal combined method and MESMA alone is caused by the different strategies used in the shadow area. MESMA alone implemented for the entire image does not consider the unmixing of impervious surface in shadow separately, so it is likely to incorrectly estimate the final impervious surface. In contrast, the combined method uses a different end-member model (IP) and samples for impervious surface in shadow separately. In addition, there is a greater possibility of multi-point scattering in shadow; making it perhaps more effective to choose a nonlinear unmixing model for that scene (see Tables 3 and 4).

This combination strategy alleviates the confusion between shadow and impervious surface, and makes full use of the advantages of nonlinear unmixing methods in shadow areas. When unmixing impervious surface with visible shadows, using nonlinear methods in the shadow areas can effectively extract the impervious-surface information, thereby improving the overall accuracy of unmixing. Based on Tables 3 and 4, the RMSE difference between the optimal combined method (involving MESMA) and MESMA alone for the entire image is smaller for Wuhan than for Hangzhou; this may be because more improvement can be obtained with the combined method when there are more shadows in the image.

All of these results and analyses indicate that it is reasonable and effective to separately unmix impervious surface with a linear model in nonshadow areas and a nonlinear model in shadow areas. The final results combining impervious-surface abundance acquired in shadow and nonshadow areas separately show higher accuracy than all of the individual methods involved when used without considering shadow.

The four combinations of unmixing methods and four comparative experiments in both study sites are shown in Figures 6 and 7, respectively. Since the nonlinear methods SVM and ANN are data driven, the results have certain fluctuations with the training situation; so all the results involving SVM and ANN from Figure 6 to Figure 9 are given as the optimal results of 10 independent experiments. Figures 6i and 7i show the binary visual-interpretation results for impervious surface in the high-resolution images for Wuhan and Hangzhou, which can be used as the ground truth for qualitatively and quantitatively judging the unmixing methods. According to Tables 3 and 4, MESMA is the best of the individual methods not considering subscene unmixing. The impervious-surface abundance
Figure 6. The abundance of impervious surface for each unmixing method in Wuhan: (a) linear spectral mixture analysis (LSMA) for the entire image; (b) multiple-end-member spectral mixture analysis (MESMA) for the entire image; (c) support vector machine (SVM) for the entire image; (d) artificial neural network (ANN) for the entire image; (e) LSMA in nonshadow and SVM in shadow; (f) LSMA in nonshadow and ANN in shadow; (g) MESMA in nonshadow and SVM in shadow; (h) MESMA in nonshadow and ANN in shadow; (i) ground truth.

Figure 7. The abundance of impervious surface for each unmixing method in Hangzhou: (a) linear spectral mixture analysis (LSMA) for the entire image; (b) multiple-end-member spectral mixture analysis (MESMA) for the entire image; (c) support vector machine (SVM) for the entire image; (d) artificial neural network (ANN) for the entire image; (e) LSMA in nonshadow and SVM in shadow; (f) LSMA in nonshadow and ANN in shadow; (g) MESMA in nonshadow and SVM in shadow; (h) MESMA in nonshadow and ANN in shadow; (i) ground truth.

Figure 8. Enlarged view of unmixing impervious surface of red rectangle A in Wuhan (Figure 2a): (a) multiple-end-member spectral mixture analysis for the entire image; (b) multiple-end-member spectral mixture analysis in nonshadow and artificial neural network in shadow; (c) ground truth.

Figure 9. Enlarged view of unmixing impervious surface of red rectangle B in Hangzhou (Figure 2b): (a) multiple-end-member spectral mixture analysis for the entire image; (b) multiple-end-member spectral mixture analysis in nonshadow and support vector machine in shadow; (c) ground truth.
from MESMA (Figures 8a and 9a) is obviously higher than the ground truth (Figures 8c and 9c), showing that the performance of MESMA is weaker when more shadows exist. This is probably because when the shadow is not considered separately, it increases the probability of confusion between shadow and low-albedo impervious surface, leading to a higher result for impervious surface. In contrast, the proposed combined framework can alleviate this situation, especially in areas with more shadows, such as Figure 9b showing that the combined method can reduce the overestimation effect of shadow on the unmixing of impervious surface.

**Discussion**

**Variability Analysis of Nonlinear Unmixing Methods in Shadow Areas**

Although the experimental results show that nonlinear unmixing in shadow areas is far better than linear unmixing, the accuracy of the two nonlinear methods (SVM, ANN) is greatly affected by the samples and training environment, and therefore their performance will inevitably show a certain degree of volatility. In the previous experiments, the final results of the two nonlinear methods were obtained by averaging 10 independent experiments. In order to further explore the trend of the effects of these two nonlinear methods in shadow areas, we performed another 100 nonlinear unmixing tests in shadow for both Wuhan and Hangzhou, and used the RMSE value as the accuracy evaluation index to draw the fluctuation graph. The experimental results are shown in Figure 10a–b; the black line is ANN and the magenta line is SVM. The two linear methods, LSMA and MESMA, serve as the comparison; since linear unmixing results are constant, they appear as two straight lines in the figure (red, LSMA; blue, MESMA).

The results reflected in Figure 10 are the same as the previous conclusions—that is, for the two nonlinear unmixing methods of SVM and ANN, their performance in shadow areas is basically better than that of the two linear methods. However, the results also fluctuate with the increasing number of experiments. For example, the accuracy of ANN will occasionally show poor results, even worse than LSMA. This may be related to overfitting of the model during the training process, because the performance of the data-based nonlinear method is directly affected by the training of the corresponding model. In addition, we found that the two study areas have some different results. In the Wuhan research area, averaging 100 experiments each for SVM and ANN, the RMSE for ANN is 0.4246, whereas for SVM it is 0.4256. Moreover, it is obvious from Figure 10 that ANN can get the optimal result below 0.4 many times, and SVM shows relatively stable unmixing ability. But in Hangzhou, the average RMSE for SVM is 0.3600, which is better than ANN (RMSE = 0.3773). This may be due to the small number of samples in Hangzhou, preventing ANN from effectively exerting its nonlinear unmixing ability and making the results of SVM more volatile compared with the Wuhan research area and its large number of samples. Finally, the question of how to choose between these two nonlinear methods can be considered based on the number of samples and on stability requirements. If the number of samples is sufficient, ANN will have a greater chance of achieving the best result at unmixing impervious surface in shadow. When the number of samples is limited, SVM, with relatively stable performance in shadow areas, can be given priority.

**Comparison of Linear and Nonlinear Unmixing Methods in Two Scenes**

Linear and nonlinear unmixing methods have different assumptions about the formation of mixed pixels. Linear methods assume that the spectral value of a mixed pixel is linearly weighted by each end member, and the abundance information is generally obtained based on the restricted least-squares solution. Models of this kind are simple and interpretable, and therefore have been widely used. However, more and more studies have found that the unmixing effect of linear models in some complex scenes is not ideal. Instead, some nonlinear methods that take into account nonlinear factors perform better. These nonlinear methods include data-driven models and physical-based models simulating multi-point scattering. By explicitly or implicitly introducing nonlinear factors into the model, these nonlinear methods can be used to simulate the formation of mixed pixels more accurately, sometimes achieving better results than traditional linear methods—especially in scenes with strong nonlinearity. For example, the two nonlinear methods selected in this article—SVM and ANN, which are data driven and have strong nonlinear unmixing ability—are always better than the linear methods selected (LSMA and MESMA) in shadow areas with complex nonlinear components.

Since shadow will have a certain impact on the unmixing of impervious surface, this article divides the task of that unmixing into subtasks in two different scenes: nonshadow and shadow areas. Linear and nonlinear unmixing, respectively, were carried out in these two scenes. Judging from the results in shadow areas, the accuracy of the data-based nonlinear unmixing methods is basically much higher than that of the linear methods in shadow areas with obvious nonlinearity.
In nonshadow areas, although MESMA achieved the optimal unmixing result, the result of the linear method LSMA is not always better than those of the two nonlinear methods. LSMA is based on fixed end members, which will lead to certain errors. In other words, if the nonlinear model is properly trained in nonshadow area, the two nonlinear methods may achieve better results than LSMA, as shown in Tables 1 (RMSE$_{ANN} = 0.3709$ versus RMSE$_{LSMA} = 0.3740$) and 2 (RMSE$_{ANN} = 0.3708$ versus RMSE$_{LSMA} = 0.3900$). For the special nonlinear scene of shadow, we can use a nonlinear method based on data training, which can greatly improve the unmixing accuracy in shadow. At the same time, considering that nonlinear methods require higher training costs and may yield a certain volatility in results, we can choose MESMA, a linear method with good effect in nonshadow areas, to unmix impervious surface in those areas. The combined framework proposed in this article can maximize the respective advantages of linear and nonlinear unmixing methods.

**Optimization Directions of the Combined Framework**

In addition to the various advantages already mentioned, the combined framework of this article still has some directions worthy of improvement. First of all, it divides the image into two different scenes (shadow and nonshadow) through shadow extraction. Since the end-member models and data sets of the two regions are different, the correct classification of each pixel image will affect the subsequent unmixing. In this article the S3D operator was selected to enhance and extract shadows, but due to the confusion of shadow with other dark objects and errors caused by threshold segmentation, the final shadow information extracted by this approach has some errors relative to the ideal situation. Therefore, one of the optimization directions is to use more accurate shadow-extraction algorithms to eliminate related errors and improve the subsequent unmixing accuracy.

Another direction that can be optimized is increasing the number of samples and end-member categories in shadow areas. Due to the limitations of the actual situation in the study areas, only a small number of samples and a simple two-end-member model were selected for impervious-surface unmixing. The two-end-member model chosen cannot fully describe the shadow scene, and there may be a certain error when unmixing impervious surface in shadow areas; a model with more end members may better describe mixed pixels and reduce unmixing error. From the results in Tables 3 and 4, the use of a data-driven nonlinear method in the shadow area can effectively improve the unmixing accuracy of impervious surface in shadow. However, the results of these nonlinear methods have certain volatility, as shown in Figure 10, and are greatly affected by the sample and training environment. By increasing the number of samples in the shadow area, the model can be fully trained, thereby improving the impervious-surface unmixing ability of these methods. In the future, we can select an appropriate end-member model and sufficient samples to apply to these nonlinear methods and improve the overall accuracy of unmixing impervious surface. Finally, exploring a more automated processing flow of the combined framework could make it more efficient, which is also a direction of optimization.

**Conclusion**

This article proposed a new combined framework to draw the impervious-surface abundance map of a 15-m-spatial-resolution Landsat 8 OLI image with visible shadows. First the image was divided into nonshadow and shadow areas using shadow detection, and then linear models in nonshadow areas and nonlinear models in shadow areas were used to unmix the impervious surface separately. The results of the two scenes were combined into the impervious-surface abundance map of the entire image. The unmixing accuracy of the four combinations of methods was assessed, and compared with the four methods individually for the entire image (ignoring shadow).

We can draw some conclusions from the experimental results and analysis. First of all, the results of separately processing shadow and nonshadow areas basically shows better performance than those obtained by the same individual processes across the entire image. This indicates that shadows should not be ignored in impervious-surface extraction, since they may occupy lots of areas in urban regions. Second, the proposed optimal combined framework applying a linear model (MESMA) in nonshadow areas and a nonlinear model (SVM or ANN) in shadow areas obtained the best accuracy. The experiments indicate the effectiveness of nonlinear models in extracting impervious surface under shadow coverage. In the future, our research will further improve the unmixing accuracy and automation of the combined method according to the optimization direction.

**Acknowledgments**

This research was funded by the National Natural Science Foundation of China (grants 41801285 and U1711267) and Fundamental Research Funds for National University, China University of Geosciences (Wuhan) (grant G1323520032). We thank Geospatial Data Cloud (http://www.gscloud.cn/) for providing the Landsat 8 OLI remote sensing images. We also thank Shubin Deng, Qijin Chen, Huijian Du, and Enhui Xu—the authors of the book ENVI Remote Sensing Image Processing Methods, published by Higher Education Press in 2014—for providing high-resolution remote sensing image data from the CD-ROM for that book.

**References**


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

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.

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.

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.
An Incremental Isomap Method for Hyperspectral Dimensionality Reduction and Classification

Yi Ma, Zezhong Zheng, Yutang Ma, Mingcang Zhu, Ran Huang, Xueye Chen, Qingjun Peng, Yong He, Yufeng Lu, Guoqing Zhou, Zhigang Liu, and Mujie Li

Abstract
Many manifold learning algorithms conduct an eigenvector analysis on a data-similarity matrix with a size of $N \times N$, where $N$ is the number of data points. Thus, the memory complexity of the analysis is no less than $O(N^2)$. We present in this article an incremental manifold learning approach to handle large hyperspectral data sets for land use identification. In our method, the number of dimensions for the high-dimensional hyperspectral-image data set is obtained with the training data set. A local curvature variation algorithm is utilized to sample a subset of data points as landmarks. Then a manifold skeleton is identified based on the landmarks. Our method is validated on three AVIRIS hyperspectral data sets, outperforming the comparison algorithms with a $k$–nearest-neighbor classifier and achieving the second best performance with support vector machine.

Introduction
Hyperspectral images (HSIs) consist of hundreds of channels that offer a large amount of information for remote sensing applications (Bachmann et al. 2005; Shao and Zhang 2014; Shao et al. 2014, 2015; X. Sun et al. 2015). However, these channels are often highly correlated and contain redundant information, posing significant computational challenges (Bioucas-Dias et al. 2013; Shao and Zhang 2014; Shao et al. 2014; Paolletti et al. 2018). Thus, exploration of dimensionality-reduction methods for HSIs is highly desirable. Many manifold learning approaches have been explored in the field of HSI applications, where the HSI data points are presumed to be in a manifold in the original high-dimensional space (W. Li et al. 2012; W. Sun et al. 2014; Lunga et al. 2014; Veganzones et al. 2016). Manifold learning algorithms aim at finding the low-dimensional structure for subsequent applications such as land use classification and object detection. The main challenge of manifold learning methods for HSI data sets is the storage and computational complexities associated with the eigen decomposition process of the $N \times N$ similarity matrix derived from the data sets, where $N$ is the number of data points (Bioucas-Dias and Nascimento 2008; C. Li et al. 2012; Álvarez-Cortés et al. 2018). This step usually becomes a computational bottleneck when $N$ is sufficiently large.

Manifold learning methods can be divided into linear and nonlinear types. Traditional linear approaches, such as principal component analysis (PCA) and probabilistic principal component analysis (PPCA), sometimes fail to reveal the intrinsic data structures in real-world high-dimensional data sets (Tang et al. 2014; Xia et al. 2014). Nonlinear manifold algorithms can be either local techniques—such as local tangent space alignment (LTSA; Ma et al. 2010a, 2010b), locally linear embedding (LLE; Roweis and Saul 2000), and t-distributed stochastic neighbor embedding (t-SNE; van der Maaten and Hinton 2008)—or global methods such as isometric feature mapping (Isomap; Tang et al. 2014). These manifold methods show more advantages in dimensionality reduction and achieve promising results in a variety of applications including feature extraction (e.g., Y. Tang et al. 2014; Xia et al. 2014), object detection (e.g., Nasrabadi 2014; L. Zhang et al. 2014; Zou and Shi 2016), and land use classification (e.g., X. Wang et al. 2005; Chen et al. 2006; Ma et al. 2010a, 2010b; Ding et al. 2013; Huang et al. 2014; W. Sun et al. 2014; Ma et al. 2015).

For high-dimensional HSIs, Y. Tang et al. (2014) have presented a manifold algorithm based on sparse representations of HSIs for land use classification, and the local geometric property is embedded into the land use classification system with a local manifold representation approach. Van der
Maaten and Hinton (2008) have introduced a t-SNE algorithm for dimension reduction of multiple features under a probability-preserving projection framework for subsequent land use classification. Chen et al. (2006) have proposed improved nonlinear manifold learning based on intelligent landmark selection for land use classification. Ma et al. (2015) have developed a semi-supervised HSI classification algorithm based on the local manifold learning approach with an improved k–nearest-neighbor classifier. Bachmann et al. (2009) have introduced a data-driven algorithm to describe the nonlinear structure of HSI with manifold coordinate representations.

L. Zhang et al. (2013) have developed a tensor discriminative locality alignment with HSIs for spectral/spatial feature extraction and land use classification. Du et al. (2012) have introduced a discriminative manifold learning method for HSI dimensionality reduction and classification.

All these algorithms have achieved very promising results. Though many consider the complexity bottleneck associated with nonlinear manifold learning, it still remains a challenge for large HSI data sets (Chen et al. 2006; van der Maaten and Hinton 2008). In this article, we present an incremental manifold learning method of dimensionality reduction for large HSI data sets. We first identify the number of dimensions with the training data set. Then the Isomap algorithm (Tang et al. 2014) is used to learn a manifold skeleton for a representative subset of the large data set. The remaining data points are then inserted into the skeleton through local linear embedding. Isomap is susceptible to producing shortcuts when computing geodesic distances for data pairs if noise data points exist between the separated portions of a manifold. The geodesic distance for a data pair is calculated by searching and propagating through local nearest neighbors. In fact, it is difficult to determine the number of neighbors that should be searched. In this study, we propose to use a L1 norm-based optimization strategy to automatically select nearest neighbors for computing geodesic distance; these nearest neighbors are restricted in a local linear patch to achieve adaptive graph construction. Noise data points usually exist off the local linear patch, so that the potential shortcuts are eliminated. The original raw data set and the dimension-reduced HSI data set are then classified with k–nearest-neighbor (kNN) and support vector machine (SVM) classifiers. SVM usually shows good performance for classifying high-dimensional data (Camps-Valls and Bruzzone 2005; Bo et al. 2016).

The main contribution of our work is that we introduce L1-norm optimization into the Isomap algorithm for HSI data sets to resolve the shortcut problem in the construction of the neighborhood graphs. In our approach, neighborhoods for each data point are selected adaptively and confined to the local linear patch. The rest of the article proceeds as follows: If the next section presents the proposed incremental manifold learning method. After that, the experimental setup is described. Then the experimental results on three AVIRIS data sets are presented and discussed, and we conclude with a summary and discussion of future work.

**Methods**

Our proposed method includes the following major steps. First, the radiation values of the HSI data set are normalized. Second, landmarks are sampled from the training data set to represent the low-dimensional manifold hidden in the original data, and the Isomap algorithm is utilized to obtain the manifold skeleton of the HSI data set. Third, the remaining training and testing data set is inserted into the skeleton using the LLE algorithm. Finally, the dimension-reduced HSI data set is classified for evaluation. The diagram of our system is given in Figure 1.

**Normalization**

The radiation values of the HSI data set may differ from channel to channel, and the absolute values of pixels in different channels may vary by a few hundred. As a pretreatment, normalization is carried out to ensure that all pixel values range from 0 to 1.

**Sampling**

Landmarks from the training data set for dimensionality reduction are sampled to describe the low-dimensional manifold hidden in the original data. In this article, landmark data points are selected according to the local curvature variation approach (Tran et al. 2015). Data points with larger curvature values have more opportunities to be sampled. Figure 2 depicts the fundamental principle of the approach. It is obvious that more data points should be sampled approaching area B than approaching area A to effectively represent the structure of the data.

The local curvature \( c_i \) of each point \( x_i \) is found among its \( k_{\text{curv}} \) neighbors as

\[
c_i \approx \frac{1}{k_{\text{curv}} - 1} \sum_{l=2}^{k_{\text{curv}}} \arccos \left( \frac{\sigma_{\text{min}}(Q_i^lQ_i^l)}{\|Q_i^l\|} \right)
\]

where \( \sigma_{\text{min}}(\cdot) \) denotes the smallest singular value, \( Q_i \) represents an orthonormal basis of the tangent space in the \( k_{\text{curv}} \) neighbors, and \( \theta_i \) denotes the tangent space projection \( \theta_i = Q^l(x_i - X) \), where \( X \) describes the average value of neighbor points (Z. Zhang et al. 2012).

---

**Figure 1. Diagram of our proposed system.**

**Figure 2. Local curvature variation.**
A probability density function \( p(x) \) is derived for importance sampling as

\[
p(x_i) = c_i / \sum_{j=1}^{n} c_j
\]

(2)

For each point \( x_i \), its probability value is proportional to its curvature value \( c_i \).

**Manifold Skeleton Learning with Isomap**

The Isomap algorithm is utilized to obtain the manifold skeleton (Tran et al. 2013, 2015). It can be summarized as finding a map \( f: X \rightarrow Y \), where \( X \) is the original high-dimensional space and \( Y \) is the lower-dimensional manifold. The approach can be divided into three steps. First, a neighborhood graph \( G \) can be constructed based on the \( k \) nearest neighbors of each data point \( x_i \) in \( X \). Second, a similarity graph \( D_s = d_s(i,j) \) can also be obtained, where \( d_s(i,j) \) is the geodesic distance between data points \( x_i \) and \( x_j \) in the neighborhood graph \( G \). The geodesic distance is calculated with Dijkstra’s algorithm (Dijkstra 1959; J. Yu and Kim 2016). Third, the dimension of the raw data set can be reduced by minimizing

\[
E = \| \tau(D_s) - \tau(D_r) \|_2
\]

(3)

where \( D_r \) represents a Euclidean-distances matrix \( (d_Y(i,j) = \| y_i - y_j \|) \). \( \| M \|_2 \) describes the L2 matrix norm \( \sqrt{\sum_{i,j} M_{ij}^2} \), and \( \tau(\cdot) \) denotes the second-order variation of the geodesic distance that converts geodesic distance into an inner product:

\[
\tau = -1/2H^TSH
\]

(4)

where

\[
S(i,j) = (d(i,j))^2
\]

(5)

is a matrix of squared distances,

\[
H(i,j) = \delta(i,j) - \frac{1}{n}
\]

(6)

\( H \) denotes a centering matrix, and \( \delta \) represents the identity matrix. The value of \( E \) can be minimized by representing \( y_i \) by the first \( d' \) eigen vectors of the matrix \( \tau(D_r) \). With each successive eigen vector, the residual variance of the data set is reduced. The residual variance is minimized if the number of eigen vectors reaches the intrinsic dimension of the manifold.

In our improved Isomap (Lr-Isomap) algorithm, \( X \) is a high-dimensional data set that needs to be reduced, and \( x_i \) is a point belonging to \( X \). The data point \( x_i \) can be reconstructed using its nearest neighbors \( X(\bar{i}) \) by solving

\[
\min \frac{1}{2} \left( X(\bar{i}) - x_i \right)^T W x_i + 2 \| W \|_F^2
\]

(7)

where the first term presents the reconstruction error and the second term denotes the Lr-norm penalty to ensure a sparse solution (Boutsidis et al. 2015; Oguslu et al. 2015; W. Sun et al. 2017). The neighborhood is confined to the same \( k \) nearest neighbors as the original Isomap implementation for a fair comparison.

**Local Linear Embedding**

Once the landmarks are defined in a lower-dimensional space, the remaining training data set can be inserted into the dimension-reduced landmark \( t \) data set with LLE. For each data point \( x_i \), its \( k \) nearest landmarks \( X(k) \) in the original high-dimensional space can be used to reconstruct \( x_i \) by minimizing

\[
E(W) = \sum_{i} \| x_i - \sum_{k} W_{ik} x_k \|_r^2
\]

(8)

where \( \Sigma W_{ik} = 1 \). To embed \( x_i \) into the lower-dimensional manifold, it has to be reconstructed in the lower-dimensional space as \( Z_i \), according to the aforementioned weights \( W_{ik} \):

\[
Z_i = \sum_k W_{ik} Z_k
\]

(9)

where \( Z_i \) denotes the low-dimensional representation of the landmarks \( X_i \). The LLE approach is described in Figure 3, where the white dot denotes a data point to be embedded into the manifold skeleton and the black dots represent the landmarks of the skeleton.

**Classification**

Both the original raw data set and the dimension-reduced data set are classified with the KNN and SVM classifiers for comparison. KNN and SVM are widely used classification algorithms (Gualtieri and Cromp 1999; Xue et al. 2014; Tan et al. 2015; Wu et al. 2016; H. Yu et al. 2017). In KNN, we first identify \( k \) nearest neighbors in the training data set for a given testing data point, then assign the majority of labels among the \( k \) neighbors as the predicted labels for the testing data points. In SVM, the decision boundaries between different classes are optimized based on the training data set, and the trained model is then used to classify the testing data set (Benabdeslem and Bennani 2006; Yang et al. 2009; Xiao et al. 2014).

**Experimental Setup**

**Data Sets**

Our proposed framework is evaluated on three AVIRIS data set collected by a 224-channel AVIRIS sensor with a spectral coverage of 0.4–2.5 μm (Luo et al. 2017; W. Sun et al. 2017):

- Salinas-A: The Salinas-A scene was acquired over Salinas Valley, California, in 1998. The scene is composed of 80×83 pixels, and the spatial resolution is 3.7×3.7 m/pixel. Twenty water-absorption channels were discarded: 108–112, 154–167, and 224. The Salinas-A ground truth consists of seven types of land use: unlabeled points, broccoli_green_weeds_1, corn_senesced_green_weeds, lettuce_romaine_4wk, lettuce_romaine_5wk, lettuce_romaine_6wk, and lettuce_romaine_7wk.

![Figure 3. Local linear embedding.](image-url)
• Indian Pines: The Indian Pines scene was captured over the Indian Pines test site in northwestern Indiana in 1992. The scene is composed of 145×145 pixels, with a spatial resolution of 20×20 m/pixel. Similar to Salinas-A, 20 water-absorption channels were removed. It consists of 16 classes, including vegetation and forests. Seven classes and 4584 pixels were randomly chosen for experiments.

• Salinas-All: The Salinas-All scene was also gathered over Salinas Valley, California, in 1998. It covers a larger area than Salinas-A—i.e., Salinas-A is a subscene of Salinas-All. The size of Salinas-All is 512×217 pixels, and the spatial resolution is also 3.7×3.7 m/pixel. After preprocessing, the Salinas-All data set is constituted of 204 spectral reflectance channels and 16 classes. We randomly selected 5000 data points, belonging to seven classes, for the third experiment. The sample channels and ground truths of these data sets are shown in Figure 4.

Experiments

Using more data for training will reduce the size of the testing data set and may give optimistic performance estimation. After trial and error, we found that equally–equally sized partitions can give us stabilized estimations for the parameters involved in the proposed method—the number of target dimensions to keep, the number of landmarks to be selected, and so on. With the training data set, we first identify the number of dimensions for the HSI data set. Second, landmark selection is performed with the local curvature variation algorithm. Third, we utilize the $L_1$-Isomap to learn a manifold skeleton for the selected landmarks. Then we insert the remaining data points, including the training and testing data sets, into the manifold skeleton with the LLE algorithm (Zheng et al. 2016a, 2016b, 2016c).

Six manifold learning algorithms are also implemented for comparison: four nonlinear methods and two linear. The four nonlinear approaches are LLE, Isomap, and t-SNE. The two linear algorithms are PCA and PPCA. PCA is usually chosen as the baseline method, since it is widely used in practice.

Once the manifolds are obtained for the data sets, we apply the two classifiers, KNN and SVM. The more data employed for classifier training, the better the performance of the classifier for unseen data. However, more training data will boost different classifiers to a level that may diminish the performance differences among the classifiers. After trial and error, we found that using 100 samples from the training data set gives a good trade-off; the classifiers trained with 100 data samples were then applied to the testing data set for performance evaluation.

Performance Metrics

Overall accuracy (OA), user accuracy (UA), producer accuracy (PA), and $\kappa$ coefficient (KC) are used as performance metrics in this study (Sim and Wright 2005; McHugh 2012; Zheng et al. 2013):

$$OA = \frac{\text{Number of correctly classified pixels}}{\text{Number of overall pixels}} \times 100\% \quad (10)$$

$$UA = \frac{\text{Number of correctly classified pixels for a specific class}}{\text{Number of pixels classified as this class}} \times 100\% \quad (11)$$

$$PA = \frac{\text{Number of correctly classified pixels for a specific class}}{\text{Number of reference pixels classified for this class}} \times 100\% \quad (12)$$

$$KC = \frac{P_D - P_C}{1 - P_C} \quad (13)$$

where KC ranges between 0 and 1, PO denotes the proportion of observed agreements, and PC indicates the proportion of agreements expected by chance.

Results

Results on Salinas-A

First, experiments were performed to decide the number of dimensions that the Salinas-A data set should be reduced to. The ultimate goal of our method is to preserve the intrinsic structure of the HSI data set and reduce the redundancy. The number of dimensions for the data set should be as low as possible to achieve satisfactory classification accuracy. The Salinas-A training data set was reduced to different numbers of dimensions from 1 to 199, and the OA results are shown in Figure 5. Based on Figure 5, the optimal number of dimensions for the reduced data set is 10.

Second, the relationship between the number of landmarks and OA performance is investigated. The number of landmarks varies from 100 to 4000 with a step size of 100, and the results are depicted in Figure 6. The CPU time needed for different landmarks is shown in Figure 7. According to Figures 6 and 7, the optimal number of landmarks is 1000. In
addition, the computational load increases as \( m^2 \), where \( m \) is the number of landmarks.

Based on these experiments, we set the number of dimensions to keep as 10 and the number of landmarks as 1000 for this data set. Then we applied the proposed method to reduce the dimensionality of the data set to 10. Finally, we trained a KNN classifier and an SVM classifier on 100 dimension-reduced samples from the training data set and tested them on the dimension-reduced testing data set. The results from our proposed method and the other seven algorithms compared are described in Tables 1 and 2. The computational times of these methods are given in Table 3. Although some classes with other approaches may have higher UA, our proposed method achieves the best OA and KC with KNN (Table 1). With SVM, our approach is comparable to PPCA, with the cost of a threefold penalty in computational complexity (Table 3). Both \( L_1 \)-Isomap and PPCA are slightly inferior to \( t \)-SNE (Table 2), but our method requires less computational time (Table 3).

**Results on Indian Pines**

Similarly, experiments were conducted to find the number of dimensions in the Indian Pines data set to keep. The training set was also reduced to different numbers of dimensions from 1 to 199, as before. The number of dimensions for the reduced data set was chosen as 9 based on the results shown in Figure 8.

The relationship between the number of landmarks and OA performance is shown in Figure 9, where the number of landmarks changes from 100 to 3500 with a step size of 100. The CPU times required for different numbers of landmarks are presented in Figure 10. Based on these results, the number of landmarks was chosen as 1000. Moreover, the computational load again increases as \( m^2 \), where \( m \) denotes the number of landmarks.

### Table 1. Classification accuracies of different methods in the Salinas-A data set with the \( k \)-nearest-neighbor classifier.

<table>
<thead>
<tr>
<th>Method</th>
<th>U</th>
<th>B</th>
<th>C</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
<th>OA</th>
<th>KC*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw (original data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>72.83</td>
<td>0.67</td>
</tr>
<tr>
<td>PA</td>
<td>48.41</td>
<td>79.59</td>
<td>89.67</td>
<td>74.89</td>
<td>79.99</td>
<td>72.34</td>
<td>73.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>40.63</td>
<td>85.28</td>
<td>90.45</td>
<td>79.48</td>
<td>85.52</td>
<td>71.22</td>
<td>81.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>47.93</td>
<td>77.82</td>
<td>88.40</td>
<td>77.16</td>
<td>78.17</td>
<td>63.11</td>
<td>72.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>43.34</td>
<td>88.52</td>
<td>87.66</td>
<td>76.65</td>
<td>85.79</td>
<td>63.92</td>
<td>68.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70.80</td>
<td>0.65</td>
</tr>
<tr>
<td>PPCA</td>
<td>40.83</td>
<td>60.77</td>
<td>76.85</td>
<td>52.05</td>
<td>69.84</td>
<td>59.66</td>
<td>62.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>31.02</td>
<td>40.68</td>
<td>76.31</td>
<td>56.60</td>
<td>81.66</td>
<td>69.52</td>
<td>71.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>33.52</td>
<td>60.79</td>
<td>62.60</td>
<td>48.20</td>
<td>63.84</td>
<td>50.64</td>
<td>50.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LTSA</td>
<td>29.71</td>
<td>58.41</td>
<td>55.15</td>
<td>52.97</td>
<td>72.24</td>
<td>53.18</td>
<td>49.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>39.16</td>
<td>72.63</td>
<td>65.14</td>
<td>55.73</td>
<td>69.63</td>
<td>54.08</td>
<td>70.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>34.83</td>
<td>82.33</td>
<td>62.34</td>
<td>63.19</td>
<td>79.21</td>
<td>51.43</td>
<td>63.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLE</td>
<td>43.96</td>
<td>63.78</td>
<td>77.59</td>
<td>67.60</td>
<td>81.28</td>
<td>64.54</td>
<td>70.31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isomap</td>
<td>40.98</td>
<td>79.96</td>
<td>72.33</td>
<td>66.78</td>
<td>87.45</td>
<td>71.30</td>
<td>66.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>47.73</td>
<td>72.08</td>
<td>81.85</td>
<td>70.94</td>
<td>77.22</td>
<td>69.56</td>
<td>67.26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>U</td>
<td>42.07</td>
<td>77.98</td>
<td>78.83</td>
<td>61.62</td>
<td>85.15</td>
<td>71.82</td>
<td>74.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-SNE</td>
<td>64.31</td>
<td>82.30</td>
<td>89.69</td>
<td>83.28</td>
<td>90.45</td>
<td>78.41</td>
<td>81.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>61.01</td>
<td>88.36</td>
<td>91.86</td>
<td>90.41</td>
<td>90.21</td>
<td>77.36</td>
<td>76.63</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*All values given as percentages except \( \kappa \) coefficients.

B = broccoli_green_weeds_1; C = corn_senesced_green_weeds; \( \kappa \) = \( \kappa \) coefficient; L4 = lettuce_romaine_4wk; L5 = lettuce_romaine_5wk; L6 = lettuce_romaine_6wk; L7 = lettuce_romaine_7wk; LLE = locally linear embedding; LTSA = local tangent space alignment; OA = overall accuracy; PA = producer accuracy; PCA = principal component analysis; PPCA = probabilistic principal component analysis; \( t \)-SNE = \( t \)-distributed stochastic neighbor embedding; U = unlabeled points; UA = user accuracy.
Once the parameters were chosen, we repeated the same experiments as with the Salinas-A data set. Results are shown in Tables 4 and 5, and the corresponding computational times are presented in Table 6. Again, although some classes with some methods may have a bit higher PA or UA, our proposed method still obtains the best OA and KC with KNN. The t-SNE method achieves the best OA and KC with SVM, but our proposed approach obtains the second best, with less computational time.

Results on Salinas-All

Experiments were carried out to identify the number of dimensions that the Salinas-All data set should be reduced to. The training data set was reduced to different numbers of dimensions from 1 to 199, with OA results described in Figure 11. Based on Figure 11, the number of dimensions for the reduced data set was set to 10.

The relationship between the number of landmarks and OA performance was investigated, and the results are shown in Figure 12. The CPU times for different landmarks are shown in Figure 13. Based on these results, the number of landmarks

### Table 2. Classification accuracies of different methods in the Salinas-A data set with the support vector machine classifier.

<table>
<thead>
<tr>
<th>Method</th>
<th>Class</th>
<th>Raw (original data)</th>
<th>PCA</th>
<th>PPCA</th>
<th>LTSA</th>
<th>LLE</th>
<th>t-SNE</th>
<th>L₁-Isomap</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>B</td>
<td>C</td>
<td>L4</td>
<td>L5</td>
<td>L6</td>
<td>L7</td>
</tr>
<tr>
<td>Raw (original data)</td>
<td></td>
<td>72.28</td>
<td>54.65</td>
<td>91.37</td>
<td>80.12</td>
<td>83.37</td>
<td>71.88</td>
<td>78.51</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>87.91</td>
<td>46.07</td>
<td>83.56</td>
<td>82.02</td>
<td>90.43</td>
<td>85.36</td>
<td>78.51</td>
</tr>
<tr>
<td></td>
<td>UA</td>
<td>84.83</td>
<td>90.42</td>
<td>90.81</td>
<td>87.09</td>
<td>98.61</td>
<td>97.72</td>
<td>92.91</td>
</tr>
<tr>
<td>PCA</td>
<td></td>
<td>83.09</td>
<td>85.72</td>
<td>93.06</td>
<td>97.43</td>
<td>98.01</td>
<td>96.89</td>
<td>99.51</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>98.00</td>
<td>95.81</td>
<td>89.33</td>
<td>93.74</td>
<td>98.07</td>
<td>98.68</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>UA</td>
<td>93.12</td>
<td>91.97</td>
<td>93.84</td>
<td>91.90</td>
<td>99.77</td>
<td>100.00</td>
<td>98.76</td>
</tr>
<tr>
<td>PPCA</td>
<td></td>
<td>88.75</td>
<td>81.05</td>
<td>75.42</td>
<td>79.81</td>
<td>91.33</td>
<td>95.13</td>
<td>98.94</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>71.27</td>
<td>87.33</td>
<td>84.64</td>
<td>55.19</td>
<td>92.68</td>
<td>83.04</td>
<td>84.89</td>
</tr>
<tr>
<td></td>
<td>UA</td>
<td>94.77</td>
<td>89.27</td>
<td>89.05</td>
<td>92.83</td>
<td>98.01</td>
<td>96.70</td>
<td>99.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>99.33</td>
<td>92.74</td>
<td>88.67</td>
<td>83.32</td>
<td>96.58</td>
<td>94.64</td>
<td>95.50</td>
</tr>
<tr>
<td>LTSA</td>
<td></td>
<td>95.61</td>
<td>92.77</td>
<td>88.69</td>
<td>93.93</td>
<td>98.40</td>
<td>96.04</td>
<td>99.04</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>99.56</td>
<td>89.40</td>
<td>91.37</td>
<td>92.38</td>
<td>99.15</td>
<td>97.35</td>
<td>95.01</td>
</tr>
<tr>
<td></td>
<td>UA</td>
<td>84.49</td>
<td>96.17</td>
<td>98.95</td>
<td>99.35</td>
<td>98.28</td>
<td>98.76</td>
<td>99.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td>99.66</td>
<td>95.22</td>
<td>94.85</td>
<td>95.03</td>
<td>99.89</td>
<td>99.51</td>
<td>98.94</td>
</tr>
<tr>
<td>LLE</td>
<td></td>
<td>89.38</td>
<td>90.44</td>
<td>95.78</td>
<td>97.34</td>
<td>98.81</td>
<td>99.38</td>
<td>99.66</td>
</tr>
<tr>
<td></td>
<td>PA</td>
<td>99.44</td>
<td>94.21</td>
<td>91.70</td>
<td>94.76</td>
<td>98.56</td>
<td>97.99</td>
<td>94.62</td>
</tr>
</tbody>
</table>

B = broccoli_green_weeds_1; C = corn_senesced_green_weeds; KC = κ coefficient; L4 = lettuce_romaine_4wk; L5 = lettuce_romaine_5wk; L6 = lettuce_romaine_6wk; L7 = lettuce_romaine_7wk; LLE = locally linear embedding; LTSA = local tangent space alignment; OA = overall accuracy; PA = producer accuracy; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding; U = unlabeled points; UA = user accuracy.

*All values given as percentages except κ coefficients.

### Table 3. Computational time (s) of different methods in the Salinas-A data set.

<table>
<thead>
<tr>
<th>Method</th>
<th>PCA</th>
<th>PPCA</th>
<th>LTSA</th>
<th>LLE</th>
<th>Isomap</th>
<th>t-SNE</th>
<th>L₁-Isomap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>9.9</td>
<td>9.5</td>
<td>9.8</td>
<td>5.9</td>
<td>13.1</td>
<td>61.0</td>
<td>27.9</td>
</tr>
</tbody>
</table>

LLE = locally linear embedding; LTSA = local tangent space alignment; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding.

Figure 8. Changes of overall accuracy with target dimensions in Indian Pines.

Figure 9. Changes of overall accuracy with number of landmarks in Indian Pines.

Figure 10. CPU time needed for different numbers of landmarks in Indian Pines.
Table 4. Classification accuracies of different methods in the Indian Pines data set with the support vector machine classifier.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw (original data)</td>
<td>72.16</td>
<td>73.00</td>
<td>91.64</td>
<td>98.72</td>
<td>64.33</td>
<td>96.51</td>
<td>93.01</td>
<td>83.43</td>
<td>0.80</td>
</tr>
<tr>
<td>PA</td>
<td>76.63</td>
<td>77.18</td>
<td>94.83</td>
<td>99.72</td>
<td>56.65</td>
<td>91.18</td>
<td>88.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td>68.77</td>
<td>70.83</td>
<td>91.58</td>
<td>98.91</td>
<td>55.20</td>
<td>86.70</td>
<td>82.55</td>
<td>89.11</td>
<td>0.74</td>
</tr>
<tr>
<td>PCA</td>
<td>66.39</td>
<td>55.88</td>
<td>89.93</td>
<td>97.68</td>
<td>54.61</td>
<td>86.48</td>
<td>91.51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPCA</td>
<td>34.89</td>
<td>25.98</td>
<td>26.51</td>
<td>99.19</td>
<td>31.59</td>
<td>18.98</td>
<td>45.66</td>
<td>42.19</td>
<td>0.29</td>
</tr>
<tr>
<td>PA</td>
<td>34.48</td>
<td>24.69</td>
<td>29.37</td>
<td>99.25</td>
<td>28.36</td>
<td>16.40</td>
<td>46.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td>51.59</td>
<td>20.59</td>
<td>44.99</td>
<td>71.67</td>
<td>44.68</td>
<td>20.03</td>
<td>57.81</td>
<td>59.74</td>
<td>0.38</td>
</tr>
<tr>
<td>LLE</td>
<td>52.80</td>
<td>20.82</td>
<td>45.05</td>
<td>70.38</td>
<td>37.41</td>
<td>17.14</td>
<td>61.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>67.43</td>
<td>54.97</td>
<td>92.42</td>
<td>82.28</td>
<td>51.48</td>
<td>95.21</td>
<td>83.98</td>
<td>82.50</td>
<td>0.70</td>
</tr>
<tr>
<td>UA</td>
<td>69.56</td>
<td>45.30</td>
<td>91.84</td>
<td>87.92</td>
<td>47.59</td>
<td>94.50</td>
<td>86.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isomap</td>
<td>71.71</td>
<td>61.81</td>
<td>92.13</td>
<td>88.66</td>
<td>58.15</td>
<td>94.48</td>
<td>87.26</td>
<td>84.79</td>
<td>0.75</td>
</tr>
<tr>
<td>PA</td>
<td>71.94</td>
<td>50.97</td>
<td>93.04</td>
<td>96.46</td>
<td>57.08</td>
<td>93.16</td>
<td>89.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-SNE</td>
<td>69.93</td>
<td>63.81</td>
<td>87.99</td>
<td>99.11</td>
<td>63.54</td>
<td>93.28</td>
<td>87.14</td>
<td>87.26</td>
<td>0.76</td>
</tr>
<tr>
<td>PA</td>
<td>78.37</td>
<td>56.04</td>
<td>93.71</td>
<td>99.79</td>
<td>53.17</td>
<td>85.31</td>
<td>87.71</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L_t-SNE</td>
<td>82.47</td>
<td>86.68</td>
<td>96.15</td>
<td>99.27</td>
<td>75.01</td>
<td>95.47</td>
<td>95.91</td>
<td>90.45</td>
<td>0.88</td>
</tr>
<tr>
<td>PA</td>
<td>82.30</td>
<td>86.04</td>
<td>97.35</td>
<td>96.85</td>
<td>73.50</td>
<td>97.26</td>
<td>94.97</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KC = κ coefficient; LLE = locally linear embedding; LTSA = local tangent space alignment; OA = overall accuracy; PA = producer accuracy; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding; UA = user accuracy.

*All values given as percentages except κ coefficients.

Table 5. Classification accuracies of different methods in the Indian Pines data set with the support vector machine classifier.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw (original data)</td>
<td>87.97</td>
<td>96.91</td>
<td>97.44</td>
<td>99.65</td>
<td>63.18</td>
<td>99.04</td>
<td>98.23</td>
<td>89.87</td>
<td>0.88</td>
</tr>
<tr>
<td>PA</td>
<td>80.91</td>
<td>78.94</td>
<td>96.04</td>
<td>97.88</td>
<td>84.37</td>
<td>90.73</td>
<td>95.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA</td>
<td>94.00</td>
<td>97.63</td>
<td>95.77</td>
<td>99.94</td>
<td>87.91</td>
<td>94.98</td>
<td>99.47</td>
<td>95.60</td>
<td>0.95</td>
</tr>
<tr>
<td>PA</td>
<td>94.55</td>
<td>87.29</td>
<td>96.68</td>
<td>97.84</td>
<td>94.92</td>
<td>93.05</td>
<td>98.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td>96.04</td>
<td>91.74</td>
<td>97.79</td>
<td>99.87</td>
<td>92.69</td>
<td>77.79</td>
<td>99.92</td>
<td>95.28</td>
<td>0.94</td>
</tr>
<tr>
<td>LTSA</td>
<td>92.74</td>
<td>92.90</td>
<td>98.81</td>
<td>99.76</td>
<td>88.80</td>
<td>89.69</td>
<td>98.09</td>
<td>90.25</td>
<td>0.88</td>
</tr>
<tr>
<td>PA</td>
<td>97.31</td>
<td>99.39</td>
<td>94.97</td>
<td>86.15</td>
<td>73.32</td>
<td>35.65</td>
<td>98.42</td>
<td>96.04</td>
<td>0.95</td>
</tr>
<tr>
<td>UA</td>
<td>91.70</td>
<td>94.10</td>
<td>96.32</td>
<td>86.35</td>
<td>81.52</td>
<td>43.19</td>
<td>97.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isomap</td>
<td>97.65</td>
<td>92.59</td>
<td>96.72</td>
<td>97.29</td>
<td>90.48</td>
<td>98.07</td>
<td>98.61</td>
<td>96.04</td>
<td>0.95</td>
</tr>
<tr>
<td>PA</td>
<td>95.38</td>
<td>88.35</td>
<td>94.84</td>
<td>95.88</td>
<td>97.06</td>
<td>97.16</td>
<td>99.54</td>
<td>96.71</td>
<td>0.96</td>
</tr>
<tr>
<td>UA</td>
<td>96.31</td>
<td>96.49</td>
<td>95.43</td>
<td>95.61</td>
<td>93.88</td>
<td>97.86</td>
<td>99.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>t-SNE</td>
<td>97.97</td>
<td>88.26</td>
<td>96.41</td>
<td>97.86</td>
<td>95.49</td>
<td>95.65</td>
<td>99.61</td>
<td>98.51</td>
<td>0.98</td>
</tr>
<tr>
<td>PA</td>
<td>97.92</td>
<td>95.56</td>
<td>98.50</td>
<td>99.51</td>
<td>99.56</td>
<td>94.51</td>
<td>99.91</td>
<td>98.10</td>
<td>0.98</td>
</tr>
<tr>
<td>UA</td>
<td>99.68</td>
<td>94.77</td>
<td>97.09</td>
<td>99.81</td>
<td>99.26</td>
<td>99.10</td>
<td>99.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L_t-SNE</td>
<td>95.95</td>
<td>98.34</td>
<td>98.70</td>
<td>98.64</td>
<td>93.36</td>
<td>98.70</td>
<td>99.94</td>
<td>97.56</td>
<td>0.97</td>
</tr>
<tr>
<td>PA</td>
<td>97.46</td>
<td>94.67</td>
<td>98.01</td>
<td>95.28</td>
<td>97.11</td>
<td>97.45</td>
<td>99.55</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KC = κ coefficient; LLE = locally linear embedding; LTSA = local tangent space alignment; OA = overall accuracy; PA = producer accuracy; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding; UA = user accuracy.

*All values given as percentages except κ coefficients.

was selected as 1000, because more landmarks could not improve OA obviously but required more computational time.

The results of the proposed method and the other seven algorithms with the chosen parameters are displayed in Tables 7 and 8, and the corresponding computational times are shown in Table 9. Although some classes with some approaches may have higher PA or UA, the proposed approach

Table 6. Computational time (s) of different methods in the Indian Pines data set.

<table>
<thead>
<tr>
<th>Method</th>
<th>PCA</th>
<th>PPCA</th>
<th>LTSA</th>
<th>LLE</th>
<th>Isomap</th>
<th>t-SNE</th>
<th>L_t-SNE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>7.2</td>
<td>7.3</td>
<td>6.8</td>
<td>6.3</td>
<td>15.2</td>
<td>58.7</td>
<td>24.5</td>
</tr>
</tbody>
</table>

LLE = locally linear embedding; LTSA = local tangent space alignment; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding.
achieves the best and second-best OA and KC with KNN and SVM, respectively. The t-SNE algorithm with SVM again obtains the best OA and KC on this data set, but at the cost of more computational time.

**Discussion**

The experimental results on Salinas-A with the KNN classifier show that our proposed L1-Isomap method yields the best results, with the highest OA (80.37%) and KC (0.76) among all eight approaches, as shown in Table 1. The performance is even better than that of the raw data set (72.83% and 0.67, respectively). It is worth noting that the dimensionality of the reduced data set is only 10, compared to 204 for the raw data set. PCA obtains the third-best results, with an OA of 70.80% and a KC of 0.65. The accuracy of Isomap is comparable to t-SNE, while PPCA, LTSA, and LLE all perform much worse. The experimental results with the SVM classifier show much better performance than those achieved with KNN, indicating that SVM has a much better generalization capability on small data sets than KNN. The proposed approach ranks third, but is comparable to the top two (t-SNE and PPCA).

Similar trends can be observed on the Indian Pines data set. With KNN, the proposed method obtains an OA of 90.45% and a KC of 0.88, which is the best among the eight approaches (Table 4). The performance margin of improvement of our

![Figure 11. Changes of overall accuracy with target dimensions in Salinas-All.](image1)

![Figure 12. Changes of overall accuracy with number of landmarks in Salinas-All.](image2)

![Figure 13. CPU time needed for different numbers of landmarks in Salinas-All.](image3)

<table>
<thead>
<tr>
<th>Method</th>
<th>Broccoli</th>
<th>Fallow</th>
<th>Stubble</th>
<th>Celery</th>
<th>Corn</th>
<th>Lettuce</th>
<th>Vineyard</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OA</td>
<td>KC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw (original data)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>99.13</td>
<td>94.52</td>
<td>94.81</td>
<td>97.71</td>
<td>92.64</td>
<td>82.97</td>
<td>95.49</td>
</tr>
<tr>
<td>UA</td>
<td>99.21</td>
<td>91.11</td>
<td>99.68</td>
<td>98.61</td>
<td>88.33</td>
<td>90.09</td>
<td>88.37</td>
</tr>
<tr>
<td>PCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>94.10</td>
<td>83.83</td>
<td>98.86</td>
<td>97.86</td>
<td>88.72</td>
<td>70.09</td>
<td>94.15</td>
</tr>
<tr>
<td>UA</td>
<td>98.11</td>
<td>81.81</td>
<td>99.09</td>
<td>99.14</td>
<td>85.52</td>
<td>78.33</td>
<td>86.61</td>
</tr>
<tr>
<td>PPCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>56.46</td>
<td>86.74</td>
<td>87.97</td>
<td>69.73</td>
<td>56.07</td>
<td>50.31</td>
<td>60.72</td>
</tr>
<tr>
<td>UA</td>
<td>58.16</td>
<td>81.67</td>
<td>87.42</td>
<td>63.80</td>
<td>55.08</td>
<td>50.00</td>
<td>69.27</td>
</tr>
<tr>
<td>LTSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>73.92</td>
<td>46.67</td>
<td>86.91</td>
<td>76.47</td>
<td>52.44</td>
<td>23.92</td>
<td>59.38</td>
</tr>
<tr>
<td>UA</td>
<td>78.20</td>
<td>45.87</td>
<td>86.11</td>
<td>73.50</td>
<td>51.40</td>
<td>27.91</td>
<td>51.39</td>
</tr>
<tr>
<td>LLE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>72.07</td>
<td>52.13</td>
<td>61.68</td>
<td>66.90</td>
<td>50.95</td>
<td>39.82</td>
<td>59.78</td>
</tr>
<tr>
<td>UA</td>
<td>74.32</td>
<td>56.80</td>
<td>63.07</td>
<td>68.40</td>
<td>56.04</td>
<td>33.30</td>
<td>66.90</td>
</tr>
<tr>
<td>Isomap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>69.86</td>
<td>58.92</td>
<td>63.60</td>
<td>75.81</td>
<td>61.43</td>
<td>35.33</td>
<td>71.47</td>
</tr>
<tr>
<td>UA</td>
<td>72.66</td>
<td>64.78</td>
<td>68.47</td>
<td>82.14</td>
<td>59.92</td>
<td>31.47</td>
<td>69.89</td>
</tr>
<tr>
<td>t-SNE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>99.80</td>
<td>91.23</td>
<td>97.86</td>
<td>99.01</td>
<td>91.68</td>
<td>87.60</td>
<td>93.74</td>
</tr>
<tr>
<td>UA</td>
<td>99.61</td>
<td>93.41</td>
<td>99.37</td>
<td>98.81</td>
<td>90.51</td>
<td>88.44</td>
<td>80.01</td>
</tr>
<tr>
<td>L1-Isomap</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA</td>
<td>99.70</td>
<td>96.27</td>
<td>99.20</td>
<td>99.37</td>
<td>95.08</td>
<td>89.62</td>
<td>95.22</td>
</tr>
<tr>
<td>UA</td>
<td>99.46</td>
<td>93.40</td>
<td>99.25</td>
<td>98.11</td>
<td>95.07</td>
<td>94.44</td>
<td>96.87</td>
</tr>
</tbody>
</table>

KC = κ coefficient; LLE = locally linear embedding; LTSA = local tangent space alignment; OA = overall accuracy; PA = producer accuracy; PCA = principal component analysis; PPCA = probabilistic principal component analysis; t-SNE = t-distributed stochastic neighbor embedding; UA = user accuracy.

*aAll values given as percentages except κ coefficients.
The experimental results with the KNN classifier again show that our method is the best (OA = 97.08%, KC = 0.96) among the eight approaches, even compared to t-SNE (OA = 95.26%, KC = 0.94). The other four manifold learning algorithms do not perform well on this data set (Table 7). In comparison, the experimental results with the SVM classifier illustrate that all eight algorithms perform well, including on the raw data set, and our proposed method ranks second (Table 8).

In summary, with the KNN classifier our proposed method achieves the best results on all three data sets for HSI dimensionality reduction and classification. With the SVM classifier, our approach obtains the second-best results, on average, for the three data sets. We also find that PCA is a robust linear manifold learning method. Nonlinear manifold learning algorithms such as Isomap perform dimensionality reduction in a similar manner as PCA, with the data-pair similarity measure replaced by the geodesic distance rather than the Euclidean distance. The geodesic distance is computed by searching through a k nearest neighborhood on the manifold and sometimes is easily confused by noise, which could degrade the performance of the algorithm. In our experiments, Isomap did not perform well on all three data sets with KNN. With the proposed L₀-norm-based optimization technique, the performance of Isomap on real data sets is improved, and it is the best in our experiments with KNN.

In the SVM optimization, regularization strategies have been involved to improve its generalization capability on unseen data. Using only 100 samples for SVM training, the trained models perform very well on the testing data sets as compared to the KNN classifier, where no regularization scheme is used.

The t-SNE method, as one of the state-of-the-art manifold learning algorithms, performs best with the SVM classifier. However, the performance boost may be due to SVM’s generalization capability. For example, t-SNE with KNN is inferior to our proposed approach for all three data sets in our experiments. In addition, t-SNE requires more computation resources, and this burden will become prohibitive for large data sets, because it is not an incremental manifold learning algorithm. Therefore, it may not be a suitable solution for large-scale remote sensing images.

The experimental results show that our proposed approach outperforms the other seven methods, including t-SNE, on all three data sets with the KNN classifier and yields the second-best results, on average, with the SVM classifier. But we note that the computational complexity of our proposed L₀-Isomap method ranks the second-highest among the eight algorithms, which should be taken into account in practice.
Conclusions
In this article, we develop an incremental Isomap approach for HSI dimensionality reduction and classification. The method is an Isomap-based incremental manifold learning approach that is suitable for handling large data sets. We validate the proposed method on three AVIRIS data sets for dimensionality reduction and classification and compare the results with those obtained by four nonlinear methods—t-SNE, Isomap, LLE, and LTSA—and two linear methods: PCA and PPCA. The experimental results show that our proposed method outperforms the other seven methods on all three data sets with the KNN classifier. It obtains the second-best results, on average, for the three data sets with the SVM classifier, only slightly worse than t-SNE. In addition, the results achieved by our method are much better than those obtained by using the raw data sets. Therefore, our proposed method is robust and efficient for HSI dimensionality reduction and classification. But we note that the computational complexity of our proposed L,-Isomap method ranks the second-highest among the eight algorithms, which should be taken into account in practice.

Our future work includes testing more HSI data sets and classifiers to demonstrate the robustness of our proposed HSI dimensionality-reduction method and implementing other new algorithms, such as a deep convolutional neural network, for comparison with the proposed method for HSI data classification.

Acknowledgments
This research is supported in part by the Fund of Research and Application of Constructing Satellite Remote Sensing Technology Power Application Comprehensive Test Site and Environment Wide Area Intelligent Monitoring (YNKJM20191246), the Open Fund of Key Laboratory of Urban Land Resources Monitoring and Simulation, Ministry of Natural Resources (grantsKF-2019-04-069 andKF-2019-04-074), the Open Fund of State Key Laboratory of Remote Sensing Science (grant OFSLRSS201819), the Science Research Program of Natural Resources Department of Sichuan Province (grants KJ-2020-3 and KJ-2021-8), the Key Technologies Research and Development Program of Sichuan Province (grant 2020YSFY 0041), the Open Fund of State Key Laboratory of Water Resources and Hydropower Engineering Science (grants 2014SWG04 and 2018SWG02), and the Guangxi Key Laboratory for Spatial Information and Geomatics (grants 17-259-16-14 and 19-050-11-06).

References
SUSTAINING MEMBERS

Aerial Services, Inc.
Cedar Falls, Iowa
www.AerialServicesInc.com
Member Since: 5/2001

Ayres Associates
Madison, Wisconsin
www.AyresAssociates.com
Member Since: 1/1953

Cardinal Systems, LLC
Flagler Beach, Florida
www.cardinalsystems.net
Member Since: 1/2001

Dewberry
Fairfax, Virginia
www.dewberry.com
Member Since: 1/1985

Esri
Redlands, California
www.esri.com
Member Since: 1/1987

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
Member Since: 12/2016

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

Green Grid Inc.
San Ramon, California
www.greengridinc.com
Member Since: 1/2020

Half Associates, Inc.
Richardson, Texas
www.half.com
Member Since: 8/2021

Keystone Aerial Surveys, Inc.
Philadelphia, Pennsylvania
www.kasurveys.com
Member Since: 1/1985

Kucera International
Willoughby, Ohio
www.kucerainternational.com
Member Since: 1/1992

L3Harris Corporation
Broomfield, Colorado
www.harris.com
Member Since: 6/2008

NV5 Geospatial
Sheboygan Falls, Wisconsin
www.quantumspatial.com
Member Since: 1/1974

Pickett and Associates, Inc.
Bartow, Florida
www.pickettsusa.com
Member Since: 4/2007

Sanborn Map Company
Colorado Springs, Colorado
www.sanborn.com
Member Since: 10/1984

Surdex Corporation
Chesterfield, Missouri
www.surdex.com
Member Since: 12/2011

Surveying And Mapping, LLC (SAM)
Austin, Texas
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
Member Since: 11/2016

Towill, Inc.
San Francisco, California
www.towill.com
Member Since: 1/1952

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

Woolpert LLP
Dayton, Ohio
www.woolpert.com
Member Since: 1/1985

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 per year*
- Professional Certification Application fee discount for any employee
- Discounts on PE&RS Classified Ad
- 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, one per year
- Discount on PE&RS full-page advertisement
- 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

456 June 2021
ASPRS MEMBER EXCLUSIVE 20% DISCOUNT ON ALL WHITTLIES PUBLISHING BOOKS USING CODE WPASPRS2

WHITTLES PUBLISHING’S STABLE OF CLASSIC GEOMATICS BOOKS INCLUDES THE THREE WINNERS OF THE PRESTIGIOUS KARL KRAUS MEDAL AWARDED BY ISPRS

978-1870325-73-8 (Available as a CD)

978-1904445-87-6

ISBN? Website only has 2nd edition

UNMANNED VEHICLE SYSTEMS FOR GEOMATICS TOWARDS ROBOTIC MAPPING

978-184995-390-0

BROWSE OUR WEBSITE TO SEE OUR FULL RANGE OF ACCLAIMED GEOMATICS BOOKS

WWW.WHITTLIESPUBLISHING.COM
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