Forest structure from terrestrial laser scanning – in support of remote sensing calibration/validation and operational inventory

David Kelbe

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation

This Dissertation is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
Forest structure from terrestrial laser scanning – in support of remote sensing calibration/validation and operational inventory

by

David Kelbe

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Chester F. Carlson Center for Imaging Science College of Science Rochester Institute of Technology

July 16, 2015

Signature of the Author ________________________________

Accepted by ____________________________________________

Coordinator, Ph.D. Degree Program Date
The Ph.D. Degree Dissertation of David Kelbe has been examined and approved by the dissertation committee as satisfactory for the dissertation required for the Ph.D. degree in Imaging Science.

Dr. Jan van Aardt, Dissertation Advisor

Dr. David Ross, External Chair

Dr. David Messinger

Dr. Robert Kremens

Date
DISSERTATION RELEASE PERMISSION
ROCHESTER INSTITUTE OF TECHNOLOGY
COLLEGE OF SCIENCE
CHESTER F. CARLSON CENTER FOR IMAGING SCIENCE

Title of Dissertation:
Forest structure from terrestrial laser scanning – in support of remote sensing calibration/validation and operational inventory

I, David Kelbe, hereby grant permission to Wallace Memorial Library of R.I.T. to reproduce my thesis in whole or in part. Any reproduction will not be for commercial use or profit.

Signature

Date
To Dad, who inspired me in all things imaging, and life.
Disclaimer

This material is based upon work supported by the National Science Foundation and National Science Foundation Graduate Research Fellowship under Grant No. DGE-1102937. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.
Forest structure from terrestrial laser scanning – in support of remote sensing calibration/validation and operational inventory

by

David Kelbe

Submitted to the Chester F. Carlson Center for Imaging Science in partial fulfillment of the requirements for the Doctor of Philosophy Degree at the Rochester Institute of Technology

Abstract

Forests are an important part of the natural ecosystem, providing resources such as timber and fuel, performing services such as energy exchange and carbon storage, and presenting risks, such as fire damage and invasive species impacts. Improved characterization of forest structural attributes is desirable, as it could improve our understanding and management of these natural resources.

However, the traditional, systematic collection of forest information – dubbed “forest inventory” – is time-consuming, expensive, and coarse when compared to novel 3-D measurement technologies. Remote sensing estimates, on the other hand, provide synoptic coverage, but often fail to capture the fine-scale structural variation of the forest environment. Terrestrial laser scanning (TLS) has demonstrated a potential to address these limitations, but its operational use has remained limited due to unsatisfactory performance characteristics vs. budgetary constraints of many end-users.

To address this gap, my dissertation advanced affordable mobile laser scanning capabilities for operational forest structure assessment. We developed geometric reconstruction of forest structure from rapid-scan, low-resolution point cloud data, providing for automatic extraction of standard forest inventory metrics. To augment these results over larger areas, we designed a view-invariant feature descriptor to enable marker-free registration of TLS data pairs, without knowledge of the initial sensor pose. Finally, a graph-theory framework was integrated to perform multi-view registration between a network of disconnected scans, which provided improved assessment of forest inventory variables.

This work addresses a major limitation related to the inability of TLS to assess forest structure at an operational scale, and may facilitate improved understanding of the phenomenology of airborne sensing systems, by providing fine-scale reference data with which to interpret the active or passive electromagnetic radiation interactions with forest structure. Outputs are being utilized to provide antecedent science data for NASA’s HyspIRI mission and to support the National Ecological Observatory Network’s (NEON) long-term environmental monitoring initiatives.
Acknowledgements

An African proverb reads,

“If you want to go fast, go alone. If you want to go far, go together.”

I am so thankful for all who have come alongside me during this journey. Jan, you’ve taught me so much more than what is contained in this dissertation. I count it a privilege to have “grown up” as a scientist under your leadership and mentorship.

Dr. Messinger, you’ve always been so supportive of the students in the DIRS lab. You believed in me long before I believed in myself, and for that I am grateful. Dr. Ross, I really admire the effort you put into developing relationships with your students, even in undergraduate calculus, where the majority of students were just “passing through” for the semester. Krem, thank you for being approachable, humble, and down-to-earth. There is so much more I wish to learn from you.

Thank you, all, for investing in me, both professionally and personally.

Sue, I credit so many of my opportunities at RIT to your optimistic, make-it-happen, attitude. Thank you for working magic on my undergraduate course plan, and enabling a host of incredible opportunities. We could all learn from your infectious curiosity for continued learning. Thank you always being ready to stop what you are doing, and listen. Cindy, where would we be without you. Every office needs someone as affectionate– and firm– as you.

Carl, you are our defacto sounding board for all questions related to career and professional plans. I hope you know how invaluable your perspective has been in helping me, and my classmates, get on track and sort out the complex road ahead. Your class made me fall in love with remote sensing – thank you. You are an inspiration both professionally and personally. Jeff, the way you think, communicate, and teach has always captivated me. Thank you for all you have done by example, to encourage me along the same path. Roger, thank you for coming beside me and making my childhood dreams of “Indiana Jones” come true. I am grateful for our friendship.

Paul, the unwritten “coauthor” of my dissertation, thank you for the technical help, for bouncing ideas around, and for always - always - being willing and eager to help a friend out. Your selflessness is an inspiration to many. Now get to work on your own thesis!

Monica, thank you for keeping me company in the front and center seats of every classroom, for your willingness to lend a hand, and for long nights studying in the Reading Room.

To Wei, Madhurima, Colin, Amanda, Katie - thank you for making endless days in a windowless office a bit more enjoyable. To Dong, Ming, Bo Ding, thanks for bringing diversity and joy to grad school. Claudia, thank you for teaching me to enjoy coffee breaks the Italian way.

I am grateful to Jason Faulring for his development of the TLS system, and to the Chester F. Carlson Center for Imaging Science for their continued support. It is truly a special place.

Finally, I owe a heartfelt acknowledgement to my collaborators, without whom this work would not have been possible: including those at the National Ecological Observatory Network (NEON), the University of Michigan Biological Station (UMBS), Virginia Commonwealth University (VCU), Canterbury University, New Zealand, and Chiba University, Japan.
Contents

Foreword ............................................. i
Declaration ............................................. i
Approval ............................................. ii
Abstract ............................................. vii
Acknowledgements ..................................... viii
Table of Contents .................................... xii
List of Figures ......................................... xiv
List of Tables ........................................... xv
Glossaries ............................................. xv
  List of Acronyms ..................................... xv
  List of Latin Symbols ................................ xviii
  List of Greek Symbols ................................ xx

1 Introduction ........................................ 1
  1.1 Context ........................................... 1
  1.2 Objectives ....................................... 2
  1.3 Dissertation Layout ............................... 3
    1.3.1 Chapter 2: Background ....................... 3
    1.3.2 Chapter 3: Measuring stem attributes (Objective 1) ................... 4
    1.3.3 Chapter 4: Pairwise marker-free registration (Objective 2) ........... 5
    1.3.4 Chapter 5: Graph-based multi-view registration (Objective 3) ........ 6
    1.3.5 Chapter 6: Conclusions, impact, and outlook ....................... 6
  1.4 Novel Contributions ................................ 7
  1.5 Related Publications ............................. 10
    1.5.1 Refereed Journal Articles .................... 10
    1.5.2 Conference Proceedings ....................... 11
    1.5.3 Conference Presentations (No proceedings) .................. 13
    1.5.4 Conference Posters .......................... 14

2 Background .......................................... 15
  2.1 Forests .......................................... 15
  2.2 Terrestrial Laser Scanning (TLS) .................. 17
  2.3 Airborne laser scanning (ALS) .................... 18
  2.4 Measuring Stem Attributes using TLS ............... 20
  2.5 Pairwise Registration of TLS data ................ 26
  2.6 Multi-view Registration of TLS data ............... 33
## Measuring stem attributes

3.1 Foreword .................................................. 41
3.2 Abstract .................................................. 42
3.3 Introduction .............................................. 43
3.4 Methods .................................................... 44
   3.4.1 Terrestrial laser scanning (TLS) system ............ 44
   3.4.2 Study Area ........................................... 48
   3.4.3 Experimental Design ................................. 49
   3.4.4 Presuppositions ..................................... 49
   3.4.5 Algorithm ............................................ 52
   3.4.6 Validation ............................................ 58
3.5 Results .................................................... 58
   3.5.1 Quantitative Stem Models ............................ 58
   3.5.2 Classification Accuracy .............................. 60
   3.5.3 Forest Inventory Parameters ......................... 63
3.6 Discussion ............................................... 67
   3.6.1 Quantitative Stem Models ............................ 67
   3.6.2 Classification Accuracy .............................. 68
   3.6.3 Forest Inventory Parameters ......................... 69
3.7 Conclusions .............................................. 71

## Pairwise marker-free registration of TLS data

4.1 Foreword .................................................. 73
4.2 Abstract .................................................. 74
4.3 Introduction .............................................. 75
4.4 Methods .................................................... 80
   4.4.1 Background ........................................... 80
   4.4.2 Study Area ............................................ 82
   4.4.3 Algorithm Problem Statement ......................... 85
   4.4.4 Algorithm Input ..................................... 87
   4.4.5 Pairwise Registration Parameter Fitting ............ 89
   4.4.6 Pairwise Error Assessment ............................ 97
   4.4.7 Algorithm Performance Analyses ...................... 101
4.5 Results .................................................... 102
   4.5.1 Point cloud .......................................... 102
   4.5.2 Transformation Parameters ............................ 104
   4.5.3 Tree Locations ....................................... 104
   4.5.4 Performance Analyses ................................ 109
List of Figures

1.1 Dissertation concept map ................................. 7
2.1 Range measurement principles of TLS ................................. 18
3.1 Low-cost terrestrial lidar system ................................. 45
3.2 Parametrization of a tapered cylinder ................................. 51
3.3 Stem modeling algorithm flowchart ................................. 52
3.4 Stem modeling conceptual diagram ................................. 54
3.5 Quantitative 3-D reconstruction of stem structure ................................. 61
3.6 Classification maps ................................. 62
3.7 Classification accuracy plots ................................. 64
3.8 Range measurement results ................................. 65
3.9 DBH measurement results ................................. 65
3.10 Plot-level parameters results (corrected for occlusion) ................................. 66
4.1 Limitations of point-based registration for TLS data in forests ................................. 81
4.2 Advantages of object-based registration for TLS data in forests ................................. 82
4.3 Example structural diversity of study site ................................. 83
4.4 Algorithm flowchart including confidence metrics ................................. 87
4.5 Pairwise registration flowchart ................................. 90
4.6 Tie points in the local coordinate system (LCS) ................................. 91
4.7 Tie points in the world coordinate system (WCS) ................................. 92
4.8 Geometric constraints on voting process ................................. 95
4.9 Circular path construction for embedded confidence metrics ................................. 99
4.10 Pairwise point cloud registration example ................................. 103
4.11 Receiver operator characteristic (ROC) curves for embedded confidence metric ................................. 106
4.12 Reported registration error vs. range ................................. 107
4.13 Percent of scans detected vs. range ................................. 109
4.14 Impact of input tie point error on registration results ................................. 110
4.15 Sensitivity of tie point diameters to registration results ................................. 111
4.16 Sensitivity of registration results to number of matching tie points ................................. 112
5.1 Example structural diversity of study site ................................. 128
5.2 Multi-view registration flowchart ................................. 131
5.3 Pairwise registration schematic showing output edges ................................. 133
5.4 Conceptual illustration of graph theory ................................. 134
5.5 Model graph encodes topological relationships between scan pairs ................................. 136
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6</td>
<td>Minimal spanning forests</td>
<td>138</td>
</tr>
<tr>
<td>5.7</td>
<td>Pose conflict due to competing paths through the graph</td>
<td>141</td>
</tr>
<tr>
<td>5.8</td>
<td>Output absolute poses</td>
<td>142</td>
</tr>
<tr>
<td>5.9</td>
<td>Circular self-closure framework provides embedded confidence</td>
<td>144</td>
</tr>
<tr>
<td>5.10</td>
<td>Multi-view registration results</td>
<td>148</td>
</tr>
<tr>
<td>5.11</td>
<td>Predictive power of reported RMSE metric</td>
<td>151</td>
</tr>
<tr>
<td>5.12</td>
<td>Predictive power of pairwise RMSE metric</td>
<td>153</td>
</tr>
<tr>
<td>5.13</td>
<td>Impact of input tie point error on multi-view registration results</td>
<td>154</td>
</tr>
<tr>
<td>5.14</td>
<td>Improvement of multi-view RMSE vs. pairwise RMSE</td>
<td>156</td>
</tr>
<tr>
<td>6.1</td>
<td>Exploratory LAI assessment</td>
<td>168</td>
</tr>
<tr>
<td>6.2</td>
<td>Theoretical trends of forest structure vs. succession</td>
<td>170</td>
</tr>
<tr>
<td>6.3</td>
<td>Forest chronosequence</td>
<td>171</td>
</tr>
<tr>
<td>6.4</td>
<td>Virtual scene generation using TLS-derived outputs</td>
<td>173</td>
</tr>
<tr>
<td>6.5</td>
<td>Hyperspectral Infrared Imager (HyspIRI) radiance contributions are modified by large-footprint point spread function (PSF)</td>
<td>174</td>
</tr>
<tr>
<td>6.6</td>
<td>Sub-pixel structural impact on spectroscopy</td>
<td>176</td>
</tr>
<tr>
<td>6.7</td>
<td>Potential for calibration/validation of remote sensing</td>
<td>178</td>
</tr>
<tr>
<td>6.8</td>
<td>Airborne laser scanning (ALS)-TLS fusion for crown parameter estimation</td>
<td>180</td>
</tr>
</tbody>
</table>
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Lidar systems specifications</td>
<td>47</td>
</tr>
<tr>
<td>3.2</td>
<td>Validation plot characteristics</td>
<td>49</td>
</tr>
<tr>
<td>3.3</td>
<td>Parameter values used in this study</td>
<td>52</td>
</tr>
<tr>
<td>3.4</td>
<td>Per-plot classification accuracies</td>
<td>62</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of ground validation plots in Harvard Forest, MA, USA.</td>
<td>84</td>
</tr>
<tr>
<td>4.2</td>
<td>Reported rigid transformation parameter estimates and errors for a registration pair from the site shown in Figure 4.10.</td>
<td>104</td>
</tr>
<tr>
<td>4.3</td>
<td>Mean upper-bound transformation parameter errors for all transformation pairs.</td>
<td>104</td>
</tr>
<tr>
<td>5.1</td>
<td>Quantitative plot characteristics</td>
<td>128</td>
</tr>
<tr>
<td>5.2</td>
<td>Quantitative multi-view registration results</td>
<td>150</td>
</tr>
</tbody>
</table>
### List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-D</td>
<td>two-dimensional. 22</td>
</tr>
<tr>
<td>3-D</td>
<td>three-dimensional. 1, 19, 20, 22, 24, 34, 97, 112, 130, 139, 148, 165, 168, 184, 186</td>
</tr>
<tr>
<td>ALS</td>
<td>airborne laser scanning. xiii, 9, 10, 20–22, 33, 35, 81, 125, 168, 179–182, 184</td>
</tr>
<tr>
<td>AOP</td>
<td>Airborne Observation Platform. 179</td>
</tr>
<tr>
<td>BA</td>
<td>basal area. 77</td>
</tr>
<tr>
<td>C</td>
<td>carbon. x, 7, 9, 170, 171, 173</td>
</tr>
<tr>
<td>CGI</td>
<td>computer-generated imagery. 173</td>
</tr>
<tr>
<td>COTS</td>
<td>commercial off-the-shelf. 79</td>
</tr>
<tr>
<td>DBH</td>
<td>diameter at breast height. xix, 4, 21–24, 26, 44, 89, 90, 92–94, 96, 98, 103, 104, 113, 114, 117, 118, 120, 146, 164, 165, 183, 186</td>
</tr>
<tr>
<td>DEM</td>
<td>digital elevation model. 33, 90</td>
</tr>
<tr>
<td>DIRSIG</td>
<td>Digital Image and Remote Sensing Image Generation. 9, 173–175, 178, 179</td>
</tr>
<tr>
<td>DOF</td>
<td>degrees of freedom. 165</td>
</tr>
<tr>
<td>DWEL</td>
<td>Dual Wavelength Echidna® Lidar. 186</td>
</tr>
<tr>
<td>EM</td>
<td>electro-magnetic. 179</td>
</tr>
<tr>
<td>GPA</td>
<td>generalized procrustes analysis. 39, 128</td>
</tr>
<tr>
<td>GPS</td>
<td>global positioning system. 28, 77</td>
</tr>
<tr>
<td>GSD</td>
<td>ground sample distance. 9, 176</td>
</tr>
<tr>
<td>GUI</td>
<td>graphical user interface. 174</td>
</tr>
<tr>
<td>HyspIRI</td>
<td>Hyperspectral Infrared Imager. xiii, 9, 10, 174, 176–178</td>
</tr>
<tr>
<td>ICP</td>
<td>iterative closest point. 32–34, 36, 39, 80–82, 126, 128</td>
</tr>
<tr>
<td>IDW</td>
<td>inverse distance weighting. 135, 139, 143</td>
</tr>
<tr>
<td>IFOV</td>
<td>instantaneous field of view. 176, 177</td>
</tr>
<tr>
<td>IMU</td>
<td>inertial measurement unit. 28, 77</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>LAI</td>
<td>leaf area index. 10, 21, 86, 133, 167–171, 179, 181</td>
</tr>
<tr>
<td>LCS</td>
<td>local coordinate system. 87, 93, 96, 102, 135, 146–148</td>
</tr>
<tr>
<td>LEO</td>
<td>low earth orbit. 176</td>
</tr>
<tr>
<td>lidar</td>
<td>light detection and ranging. 19, 130, 174</td>
</tr>
<tr>
<td>MCE</td>
<td>maximum correspondence error. 39</td>
</tr>
<tr>
<td>MSS</td>
<td>minimum sample set. 91–93, 115, 116, 118–120</td>
</tr>
<tr>
<td>MST</td>
<td>minimal spanning tree. 139, 149</td>
</tr>
<tr>
<td>NPP</td>
<td>net primary production. 9, 170–172</td>
</tr>
<tr>
<td>PAR</td>
<td>photosynthetic active radiation. 167</td>
</tr>
<tr>
<td>PSF</td>
<td>point spread function. xiii, 176, 177, 186</td>
</tr>
<tr>
<td>$R^2$</td>
<td>coefficient of determination. 22–24, 26, 31, 79, 164, 170</td>
</tr>
<tr>
<td>RAM</td>
<td>random-access memory. 178</td>
</tr>
<tr>
<td>RANSAC</td>
<td>RAndom SAmple Consensus. xviii, 8, 25, 91, 92, 97, 98, 115, 116, 118–120, 165</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristic. 2, 77, 81, 103, 106–108, 148, 154, 160, 161</td>
</tr>
<tr>
<td>RSR</td>
<td>reduced simple ratio. 181</td>
</tr>
<tr>
<td>SAR</td>
<td>synthetic aperture radar. 174</td>
</tr>
<tr>
<td>SAS</td>
<td>Statistical Analysis Software. 168</td>
</tr>
<tr>
<td>SSE</td>
<td>sum-of-squared error. 82</td>
</tr>
<tr>
<td>SVD</td>
<td>singular value decomposition. 91, 92, 98, 118, 120, 134, 136, 149, 165</td>
</tr>
<tr>
<td>TIR</td>
<td>thermal infrared. 176</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>UN-REDD</td>
<td>United Nations Reduced Emission from Deforestation and Forest Degradation. 180, 186</td>
</tr>
<tr>
<td>VSWIR</td>
<td>visible to short wave infrared. 176</td>
</tr>
<tr>
<td>WCS</td>
<td>world coordinate system. xx, 38, 39, 94, 127, 128, 134, 135, 142, 144, 146, 147</td>
</tr>
<tr>
<td>wlidar</td>
<td>waveform light detection and ranging. 10, 19, 21, 178, 179</td>
</tr>
</tbody>
</table>
## List of Latin Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>a weighted adjacency matrix. 139</td>
</tr>
<tr>
<td>B</td>
<td>background. 90</td>
</tr>
<tr>
<td>D</td>
<td>directed graph. 145</td>
</tr>
<tr>
<td>d</td>
<td>A triplet of $d$ values. xx, 94, 95</td>
</tr>
<tr>
<td>$d$</td>
<td>tree DBH. xviii, 88–90, 94, 96, 98</td>
</tr>
<tr>
<td>E</td>
<td>a graph edge. 138, 140, 141</td>
</tr>
<tr>
<td>e</td>
<td>Euclidian distance between two points. 92, 98</td>
</tr>
<tr>
<td>$\mathcal{F}$</td>
<td>spanning forest. 139–142, 149</td>
</tr>
<tr>
<td>$\mathcal{G}$</td>
<td>a graph. 135, 137–140</td>
</tr>
<tr>
<td>k</td>
<td>Binomial coefficient index $k$, i.e., “n choose k”. 91, 94, 95</td>
</tr>
<tr>
<td>$k$</td>
<td>path length. 141</td>
</tr>
<tr>
<td>LCS</td>
<td>local coordinate system. 87, 92, 98, 135, 142</td>
</tr>
<tr>
<td>$\mathcal{L}'$</td>
<td>Likelihood of a pair’s correspondence based on intrinsic geometric similarity. 92, 96–98</td>
</tr>
<tr>
<td>l</td>
<td>tree location at ground (i.e., intersection of tree with DEM). xviii, 88–90, 95, 98, 100–102</td>
</tr>
<tr>
<td>$l$</td>
<td>A triplet of $l$ values. xx, 95, 97, 99, 100</td>
</tr>
<tr>
<td>M</td>
<td>number of scans. 86, 102, 135, 137–144, 146, 149</td>
</tr>
<tr>
<td>$N_C$</td>
<td>Number of combinations. 92, 96</td>
</tr>
<tr>
<td>$N_P$</td>
<td>Number of triplet pairs between $i$ and $j$. 96–98</td>
</tr>
<tr>
<td>$N_R$</td>
<td>Number of RAndom SAmple Consensus (RANSAC) iterations. 98</td>
</tr>
<tr>
<td>$N$</td>
<td>number of trees detected. xx, 89, 91–96, 98, 101</td>
</tr>
<tr>
<td>$O$</td>
<td>Computational complexity. 91</td>
</tr>
</tbody>
</table>
\( P \) a path. 140, 141
\( \mathcal{P} \) point cloud. 82–84, 87, 90, 130, 135, 144
\( \mathbf{R} \) rotation matrix. 88, 89, 92, 97, 98, 100–102, 135–138, 141, 142, 145
\( r \) range. 82, 130
\( \mathcal{S} \) sensor. 82–84, 87–89, 93, 94, 134–140, 142–146
\( \hat{\mathcal{S}} \) modeled object surface. 90
\( \mathcal{S} \) object surface. 82, 83, 90
\( \mathcal{T}_i \) set of tree model parameters for the \( i \)-th sensor. 88
\( \mathcal{T} \) facetized geometric tree model. 87, 89, 92–94, 98–101, 135, 145, 147, 148
\( t_d \) Relative diameter at breast height (DBH) difference. 96, 118
\( \mathbf{t} \) translation vector. 88, 89, 92, 97, 98, 100–102, 135–138, 141, 142, 145
\( t_x \) translation in x. 88, 89, 98, 99, 102, 106, 136, 143
\( t_y \) translation in y. 88, 89, 98, 99, 102, 106, 136, 143
\( t_z \) translation in z. 88, 98, 99, 102, 106, 136, 143
\( V \) a graph vertex. 138, 140, 141
WCS world coordinate system. 142
\( W \) Weight of a graph edge. 139
\( e \) error. 88, 89, 102, 135–138, 140–142, 146–148, 152, 154, 155, 157, 158
\( W \) weight of a graph edge. 138
\( \mathbf{x} \) a point in 3-D. 82, 130
List of Greek Symbols

\( \mathbf{D} \) set of \( d_{c=1\ldots N} \). 92, 94, 96
\( \Lambda \) set of \( \lambda_{c=1\ldots N} \). 92, 95, 96

\( \theta \) scan mirror angle. 82, 130
\( \theta_x \) rotation angle about x axis. 88, 98, 99, 102, 106, 136, 143, 146
\( \theta_y \) rotation angle about y axis. 88, 98, 99, 102, 106, 136, 143, 146
\( \theta_z \) rotation angle about z axis. 88, 98, 99, 102, 106, 136, 143, 146

\( \mathbf{L} \) set of \( l_{c=1\ldots N} \). 92, 95, 96
\( \lambda \) eigenvalue. xx, 95–97
\( \lambda \) A triplet of \( \lambda \) values. xx, 95

\( \mu \) mean. 30, 79

\( \sigma \) standard deviation. 30, 39, 79, 104, 113, 114

\( \hat{\tau} \) Estimated effective translation (i.e., position relative to world coordinate system (WCS)). 140–142, 144

\( \phi \) rotation stage position. 82, 130

\( \Psi \) Estimated effective error. 140–144, 146–148, 151–158

\( \hat{\Omega} \) Estimated effective rotation (i.e., orientation relative to WCS). 140–142, 144
Chapter 1

Introduction

1.1 Context

Forests are an important part of the natural ecosystem, providing resources such as timber and fuel, performing services such as energy exchange and carbon storage, and presenting risks, such as fire damage and invasive species impacts. Improved characterization of forest structural attributes is desirable, as it could improve our understanding and management of these natural resources.

Traditionally, the systematic collection of forest information related to stem volume and biomass—dubbed “forest inventory”—is achieved via relatively crude, readily-measured explanatory variables, such as tree height and stem diameter. Such field inventories are time-consuming, expensive, and coarse when compared to novel three-dimensional (3-D) measurement technologies. Remote sensing estimates, on the other hand, provide synoptic coverage, but often fail to capture the fine-scale structural variation of the forest environment. Terrestrial laser scanning (TLS) has demonstrated a potential to address these limitations, while offering opportunity to support remote sensing efforts by providing spatially-explicit ground-truth data for calibration/validation in forest environments. An additional benefit is the potential to extract realistic 3-D forest models, for use in simulation and visualization studies. However, despite this poten-
tial, operational use has remained limited due to unsatisfactory performance characteristics vs. budgetary constraints of many end-users. To address this limitation, this dissertation advanced operational, TLS-based forest structure assessment capabilities, in terms of three main objectives.

1.2 Objectives

This dissertation addresses three main objectives related to improving forest structure assessment capabilities from TLS. The objectives, and corresponding sub-objectives, are as follows:

1. Assess the feasibility of a low-cost, low resolution (spatial impulse and angular sampling) TLS for automatic forest stem inventory

   1.1. Quantify the error of stem detection in terms of per-pulse classification accuracy
   1.2. Determine the measurement accuracy and precision of visible tree stem structural inventory parameters
   1.3. Evaluate the impact of occlusion in terms of limiting plot-level inventory assessment

2. Determine the error associated with automatic, blind, marker-free registration of TLS data pairs in forest environments

   2.1. Quantify the root mean square error (RMSE) of the proposed marker-free registration approach
   2.2. Assess the validity of the embedded confidence metric using receiver operator characteristic (ROC) curves
   2.3. Inform optimal sample spacing considerations for TLS data collection in New England or similar forests

3. Determine the error associated with multi-view, marker-free registration of TLS data in forest environments

   3.1. Evaluate the performance of the proposed multi-view registration algorithm
   3.2. Assess the validity of the embedded multi-view registration confidence metric
3.3. Demonstrate the improvement in plot-level forest parameter estimation afforded by multiple scans

1.3 Dissertation Layout

This dissertation contains six chapters and an appendix. The Introduction is the first chapter.

1.3.1 Chapter 2: Background

Chapter 2 provides a comprehensive background on the context for this work. Forests are an important part of the natural ecosystem. Traditionally, the systematic collection of forest information, related to stem volume and biomass – dubbed “forest inventory” – is achieved via relatively crude, readily-measured explanatory variables, such as tree height and stem diameter. Remote sensing estimates, on the other hand, provide synoptic coverage, but often fail to capture the fine-scale structural variation of the forest environment. This chapter describes the motivation for forest structure assessment using TLS, and highlights the potential for TLS to address the limitations of traditional forest inventory, while also supporting remote sensing efforts through calibration and validation with spatially-explicit ground truth data. A state-of-the-art in forest structure assessment using TLS is presented, from which four limitations are identified. The four limitations reflect an overall gap in the operational capability of TLS to satisfy performance criteria against the constraints (practicality, efficiency, and cost) of many end-users, and are summarized as follows (Lovell et al., 2011; Bi and Wang, 2010; Pueschel, 2013). First, we identified the need for quantitative stem measurement from low-cost, low-resolution, single-scan TLS data. Second, we identified a need to develop blind, marker-free registration approaches for scan pairs in the forest environment. A view-invariant feature metric is developed, which enables efficient data acquisition without artificial targets or tie points. Third, we identified the motivation for multi-view registration of forest point cloud data, which exploits multiple pairwise connections in order to bring occluded scans into global alignment. Finally, we identified the need for canopy structural assessment from low-resolution TLS. These limitations are the foundation of four objectives,
which are addressed in sequence in the subsequent chapters.

1.3.2 Chapter 3: Measuring stem attributes (Objective 1)

From a comprehensive review of the literature (chapter 2), we found that TLS has demonstrated a potential to address the limitations of traditional forest inventory. However, despite significant research focus over the past decade, its application to operational inventory has been limited. This is due in part to the high cost (Pueschel et al., 2013; Clawges et al., 2007) of systems primarily designed for engineering, architecture, and forensics (Danson et al., 2007; Henning and Radtke, 2006b; Yang et al., 2013). A mobile laser scanning system, recently-developed by Rochester Institute of Technology (RIT), provides the cost-effective hardware necessary to rectify this knowledge gap, but leaves uncertainty as to the validity of existing measurement techniques for a low-cost system, which is limited by angular sampling resolution, registration, and laser beam divergence.

Therefore, we identified a first objective as follows:

1. Assess the feasibility of a low-cost, low resolution (spatial impulse and angular sampling) TLS for automatic forest stem inventory.

Chapter 3 describes in detail the study, methodology, and results related to objective 1. Outputs include an original, robust methodology to model visible tree stem structure based on intuitive 2-D projections, from which traditional inventory metrics (tree location, DBH, basal area, stem density) readily can be extracted. Additional structural information is encapsulated in the face-tized 3-D tree stem models, which have implications for parameterization of models for virtual scene generation, calibration/validation, etc. Although good recovery of visible stem structure is achieved, a limitation remains in terms of plot-level forest assessment, due to the occlusion of objects from the sensor field of view, this will be addressed in the subsequent chapter. This research fills a vital gap in our ability to assess sub-canopy forest structure, as dramatic changes in land use and other human activities necessitate improved monitoring and assessment of the biosphere (National Ecological Observatory Network, 2014)
1.3.3 Chapter 4: Pairwise marker-free registration (Objective 2)

Results and conclusions from Chapter 3 identified occlusion - not resolution or algorithm limitations - as the primary challenge to accurate plot-level forest inventory using a low-cost TLS. This application confirms the need for point cloud registration; however, from a review of the literature (chapter 2), the need for registration of TLS data is much more pervasive, with application to robotics/mobile perception (Forsman and Halme, 2005), mapping, and the majority of techniques for canopy structure assessment. In particular, we identified a critical knowledge gap in the ability to perform automatic, blind, marker-free registration of TLS data pairs in forest environments. Such an approach is necessary in order to support operational, efficient, data acquisition.

Therefore, we identified a second objective as follows:

2. Determine the error associated with automatic, blind, marker-free registration of TLS data pairs in forest environments.

Chapter 4 describes in detail the study, methodology, and results related to objective 2 and its corresponding sub-objectives. Outputs include an original, robust methodology to register TLS data using view-invariant tie points derived from the stem-terrain intersection points, i.e., the geometric primitives derived in Chapter 3. Moreover, geometric properties are exploited to constrain the search space and improve computational efficiency, enabling automatic registration without knowledge of initial sensor pose. We also present an innovative approach for providing an embedded error metric, by exploiting circular self-closure through disjoint tie point sets. This research fills a vital knowledge gap in the efficient, pairwise registration of forest point cloud data, and has implications for informing optimal sample spacing for TLS data collection, improving the plot-level assessment of inventory (Chapter 3) and supporting the assessment of canopy structure. However, global inconsistencies may exist despite local matches, and some scans may be occluded from view of the reference scan, requiring global registration.


1.3.4 Chapter 5: Graph-based multi-view registration (Objective 3)

Pairwise registration, including the approach developed in Chapter 4, can provide positive registration results between scans, which share corresponding tie points, but has several remaining limitations. For example, occlusion or view disparities may reduce the number of scans that can be successfully linked to a single reference node, thus limiting the geographic extent. Moreover, pairwise registration results may be globally inconsistent, despite purported consistency at the local level, i.e., between pairs. As a result, multi-view registration is needed to perform global registration of the network of pairwise correspondences.

Therefore, we identified a third objective as follows:

3. Determine the error associated with multi-view, marker-free registration of TLS data in forest environments.

Chapter 5 describes in detail the study, methodology, and results related to objective 3. Outputs include an original, robust methodology, which performs multi-view registration of TLS data using a graph theory approach. Pairwise registration connections from chapter 4 are used to initialize the edges of a graphical framework. We define edge weights from the pairwise embedded confidence metric of chapter 4, which has the potential to simplify the registration process, while improving the resistance to noise (Huber and Hebert, 2003). We compare the trade space of both sequential and simultaneous registration paradigms, and develop a hybrid approach, which maintains the advantages of each. Finally, we demonstrate the improvement of plot-level forest parameter estimation afforded by multi-view registration.

1.3.5 Chapter 6: Conclusions, impact, and outlook

The cumulative result of these chapters is the generation of spatially-explicit, georeferenced, forest structure products. Chapter 6 describes the impact of this work, and identifies the impact across various domains: We highlight the potential for establishing structure-function relationships from TLS data towards understanding forest growth and production in response to disturbance and management. We demonstrate how TLS structural outputs from this dissertation can be used to
1.4. NOVEL CONTRIBUTIONS

parameterize virtual scene generation, and share some examples of where these virtual scenes are being utilized to (i) provide insight on the phenomenology of remote sensing systems, and (ii) provide antecedent science data for pre-launch remote sensing missions. We illustrate the potential for TLS data to support remote sensing calibration/validation by providing synergistic structural ground truth. Finally, we propose suggestions for future investigations based on the algorithm and system-level limitations identified in this dissertation. This flow of information is encapsulated in concept map of Figure 1.1.

Figure 1.1: This dissertation is arranged sequentially according to four identified limitations of forest structure assessment from TLS. Each chapter lays the foundation for the next, by utilizing the outputs (left hand arrows) as input in the subsequent chapter. Moreover, each subsequent chapter develops additional computational tools, which address the identified limitations of the previous chapter, i.e., right hand itemization (−).

1.4 Novel Contributions

Chapter 3: Measuring stem attributes (Objective 1)

• Development of a robust technique for tree geometric reconstruction from low-resolution, single-scan point cloud data, thus enabling measurement of tree stems, which subtended at
least 15 mrad – the angular beam width of our system.

- Provision of computational tools, which demonstrate and support the utility of low-resolution, low-cost TLS instruments for operational forest inventories.

**Chapter 4: Pairwise marker-free registration (Objective 2)**

- Development of a view-invariant feature descriptor, i.e., stem-terrain intersection points, for enabling efficient, marker-free registration of point cloud data in forest environments.
- Integration of view-invariant geometric properties, i.e., feature-triplet eigenvalues, which constrain the search space and thus enable blind registration without initial pose estimates.
- Development of a framework, which exploits RANSAC to reduce error and provide output registration results, which are precision-limited by the noise of the input tie points.
- Design and integration of an embedded error metric, which provides an upper-bound error metric associated with each registration, by exploiting circular-self-closure between disjoint tie point sets.
- Establishment of optimal sample-spacing considerations for TLS data collection in forest environments.

**Chapter 5: Graph-based multi-view registration (Objective 3)**

- Development of a graph-based framework for multi-view registration, which exploits pose conflict as redundant information, in order to improve the precision of output registration parameters.
- Integration of embedded pairwise error metrics associated with each “edge” to simplify the graph-based optimization framework, while improving resistance to noise.
- Development of a method to merge modeled geometry outputs at the object-level, thus providing improved plot-level inventory estimation, and enabling virtual scene generation.
• Cumulatively, chapters 3-5 provide an end-to-end framework for plot-level forest structural parameter estimation from low-resolution TLS. Computational tools successively build on each other to provide spatially-explicit, georeferenced, forest structure products, which have an number of impacts across domains, as outlined in Chapter 7: Conclusions, impact, and outlook.

Conclusions, impact, and outlook

• Foundation for deriving TLS canopy structure metrics, which will be linked to forest function by collaborators, e.g., net primary production (NPP), in an effort to understand how and why C fluxes change at timescales relevant to ecological succession (MSc Biology Thesis, Cynthia M. Scheuermann, VCU).

• Provision of forest structural inputs to automatically parameterize realistic, virtual scene generation based on actual TLS-measured tree geometry.

• Virtual forest scenes are being utilized in the Digital Image and Remote Sensing Image Generation (DIRSIG) model to provide insight on the phenomenology of airborne sensing systems, i.e., by providing fine-scale reference data with which to interpret the active or passive electromagnetic radiation interactions with forest structure.

• Structural outputs are being utilized in the simulation environment to provide antecedent science data to the National Aeronautics and Space Administration (NASA)’s HyspIRI mission, specifically in terms of the effect of structural variation on spectroscopy for 30-60 m ground sample distance (GSD).

• The end-to-end output of spatially-explicit forest structure products provides an opportunity to link the fidelity of TLS reference data with the synoptic perspective of ALS, which may support remote sensing calibration/validation, specifically the NEON’s environmental monitoring initiatives.

TLS instrumentation

• Pioneered the exploitation of RIT’s TLS system since its inception in 2010.
• Extensive use and field testing of the TLS system led to the identification and support of significant system design improvements.

**Data collections**

• Hemlock Forest, NY: Initial testing of the TLS system for operational inventory.

• Harvard Forest, MA: Participated in a collaborative effort with NEON and UMBS to collect TLS and leaf area index (LAI) data in support of NEON’s long-term environmental monitoring initiatives.

• Hagley Park, Christchurch, New Zealand: Provided TLS data support on how to link ALS to TLS for urban forests.

• San Joaquin Experimental Reserve and Soaproot Saddle, CA: Contributed to the design and execution of a field campaign to collect TLS, LAI, and herbaceous biomass data in conjunction with overflights of both NEON spectrometer and waveform light detection and ranging (wlidar) data, and NASA’s HyspIRI’s test flights.

• University of Michigan Biological Station, MI: Led the design and execution of a field campaign to assess forest canopy structure across a forest disturbance chronosequence, representing a 200-year range of forest regrowth following “cut and burn” disturbance.

### 1.5 Related Publications

Portions of this dissertation, as well as preliminary and/or closely related work, have been published in the following outlets:

#### 1.5.1 Refereed Journal Articles

   
   DOI: [10.1109/JSTARS.2015.2416001](https://doi.org/10.1109/JSTARS.2015.2416001). In press.


### 1.5.2 Conference Proceedings


### 1.5.3 Conference Presentations (No proceedings)


1.5.4 Conference Posters


Chapter 2

Background

2.1 Forests

Changes in the biosphere, in response to human and other impacts, necessitate increased monitoring and assessment of natural systems. Forests, in particular, occupy a prominent role in natural resource management, policy, and economics. Management of forests is vital to the economy, providing resources such as timber and fuel (Klemperer, 1996), and to environmental health, through regulation of services such as animal habitat provision (Lindenmayer and Franklin, 2002) and emissions reductions (Gibbs et al., 2007), and through mitigation of risk due to erosion (Booth et al., 2002), forest fires (Chandler et al., 1983), and invasive species (Pimentel et al., 2005).

One key to effective management is information about the state and dynamics of forest resources. This is obtained through a systematic collection of forest attribute data (Tansey et al., 2009) referred to as “forest inventory” (Hamilton, 1975). Conventional inventory techniques aim to provide coarse, readily-available proxies for understanding higher-order system information, such as biomass, timber value, and carbon sequestration. Typically, statistical sampling techniques are used to allocate a series of plot-level (fixed or variable) measurements and then extrapolate these measurements to larger areas (Kangas and Maltamo, 2006). These plot-level measurements may include tree species, tree height, crown width, canopy thickness, leaf area index, stem den-
2.1. FORESTS

Sity, basal area, and DBH (diameter at breast height; measured 1.3 m above ground), and are acquired using a range of tools including tape measures, hypsometers, and rangefinders (Kangas and Maltamo, 2006; Liang et al., 2012).

Decades of research have made traditional forest inventory the underpinning of forest study, management, and policy (Tansey et al., 2009). Yet, in the context of both dramatic changes to the environment and rapid technological innovation, traditional mensuration techniques do not always provide a desired level of structural fidelity that is needed to understand forest functioning. For example, manual measurement fails to adequately capture the fine-scale structural variability of forests or information about explicit stem, branch, and canopy structure (Henning and Radtke, 2006b). Furthermore, it is susceptible to subjective errors, which sometimes makes reproducibility sometimes challenging (Hopkinson et al., 2004; Bréda, 2003). Increased characterization of forest structural attributes is desirable, as it could improve our understanding and response to a diverse range of processes, such as photosynthesis and respiration (Hilker et al., 2012c), leaf acclimation (Alton and North, 2007), branch decay times (Raumonen et al., 2011), carbon cycle estimations (Zhao et al., 2011), and fire propagation (Loudermilk et al., 2009).

Recent technological advancements have demonstrated the capacity of laser scanning to rapidly record detailed structural information, both on the ground and - for large-area operations - from air and space (Bachman, 1979). Airborne (ALS) and terrestrial laser scanning (TLS) are active sensing systems, which measure geometric characteristics, as opposed to reflectance or other radiation signatures obtained by passive sensing systems. Thus, they provide an important link between vegetation structural and material properties, and the subsequent ecological features of interest (Zhou et al., 2014). Airborne laser scanning (ALS) has matured to operational use over the past decade for large-scale forest structure assessment (e.g., Wehr and Lohr, 1999; Nelson et al., 1988; Lefsky et al., 2002; Næsset, 2007); the reader is referred to Hyyppä et al. (2008) for a detailed review. However, airborne analyses rely on ground-truth information (e.g., inventory) for calibrating and validating landscape models (Liang et al., 2012; Yu et al., 2010). As such, they too are limited by the fidelity - the structural resolution - of ground-reference data provided from traditional forest inventory. Terrestrial laser scanning (TLS), on the other hand, is well-poised
to address both the limitations in forest inventory (Hopkinson et al., 2004; Maas et al., 2008) and
the calibration needs of airborne forest sensing, including ALS (Hilker et al., 2012a; Jupp, 2011;
Lindberg et al., 2012; Liang et al., 2012).

2.2 Terrestrial Laser Scanning (TLS)

TLS is an active sensing technique, which utilizes a laser-based ranging sensor mounted on a
ground-based platform. A pulsed laser, typically with a wavelength in the infrared (905 or 1064
nm) is deflected by a rotating mirror assembly to rapidly interrogate a scene in a “fan” pattern
(Figure 2.1). This sampling in zenith angle is coupled to platform movement, either by translation
(e.g., for vehicular or airborne platforms), or azimuthal rotation (e.g., from a tripod-mounted
rotation stage) to build up a description of the surrounding structure. Range is recorded by either
phase differences or time-of-flight. For time-of-flight lidar, as used in this dissertation, the emitted
pulse interactions with object structure cause a deformation in the temporal profile of the backscat-
tered energy. The backscattered energy profile, or “waveform”, is effectively a convolution of the
temporal laser beam impulse response (typically a Gaussian), and the target(s). Internal digitiza-
tion routines then sample the backscattered energy profile and estimate the temporal locations
of target interaction. For each target, range, \( r \), is then computed based on the return-trip travel
\( t \), and the speed of light, \( c \) as in Equation 2.1. Target range is coupled to encoded angular
information, providing explicit measurement of three-dimensional (3-D) location.

\[
r = \frac{c \cdot t}{2}
\]

Waveform light detection and ranging (wlidar) is capable of digitizing the full backscattered
energy profile, whereas discrete light detection and ranging (lidar), as used in this study, detects the
peaks of the energy profile and records just the locations of these object interceptions. Resolution is
defined in terms of the temporal and spatial profile - the “impulse response” - of the emitted laser
pulse. Importantly, the spatial profile is range-dependent based on the divergence of the laser
beam, and therefore is often expressed simply in terms of the angular beam divergence. Likewise,
sampling distance is expressed in terms of the angular step-width of the mirror assembly or rotation stage. These concepts are illustrated in Figure 2.1.

Figure 2.1: TLS measures 3-D object location based on the return-trip travel time of an emitted laser pulse. The emitted laser pulse has a temporal profile, as illustrated by the Gaussian impulse response, and a range-dependent spatial profile, as illustrated by the diverging beam-width. Interaction of the laser pulse with object structure causes a temporal deformation in the backscattered energy profile, which is digitized and converted to range. This, in conjunction with angular information, provides for precise measurement of 3-D location.

2.3 Airborne laser scanning (ALS)

Like TLS, airborne laser scanning (ALS) is an active sensing technique, which records range data based on an emitted laser pulse. ALS provides systematic, wide-area coverage, but faces several limitations. A first limitation of ALS is the reduced ability to measure within-canopy structure. Discrete-return ALS, for example, records only the first and last, or perhaps, a few, e.g., up to 5,
backscattered returns from each emitted laser pulse. As a result, limited information from the inner canopy is obtained (Lovell et al., 2003; Chasmer et al., 2004). TLS systems, on the other hand, provide hemispherical scanning from a ground-based platform, and thus sample different parts of the forest structure, but often under-sample the upper canopy. Chasmer et al. (2004) examined the voxel column percentile distributions of point returns for both ALS and TLS and demonstrated that a higher percentage of laser pulses intercept the top of the canopy for ALS, with limited returns within the canopy and understory. Likewise, TLS exhibited a higher number of returns from the understory, stems, and lower-canopy, but had fewer returns in the upper canopy due to occlusion. Substantial gaps in measurement can be reconciled by augmenting data from both modalities, with implications for improving estimates of leaf area index (LAI), clumping, and canopy closure, e.g., for radiation modeling (Chasmer et al., 2004).

A second limitation concerns the inability of ALS to measure biophysically relevant parameters of interest. Due to the predominantly nadir perspective and signal density and attenuation, measurement of stem structure, including diameter at breast height (DBH), volume, etc., is difficult (Lovell et al., 2003). These stem variables, however, are important for assessing biomass storage, merchantable timber volume, etc. Moreover, with limited incidence angles, and finite footprint sizes on the order of 0.1–0.5 m, ALS may be unable to detect small canopy gaps (Henning and Radtke, 2006b; Parker et al., 2004; Tickle et al., 2006), which are important in understanding the distribution of foliage elements within a canopy. As a result, canopy cover is often overestimated using ALS (Lovell et al., 2003).

A third limitation is related to the complexity of emerging small-footprint wlidar, which digitizes the entire backscattered pulse, rather than recording just a few discrete return locations. Despite the potential of wlidar to make addition within-canopy measurements, there has so far been limited use of the technology in forest inventories, owing to uncertainty in appropriate physical interpretation of the complex, backscattered waveform. TLS provides fine-scale structural reference data, and may be useful to better understand the complex interactions of an emitted laser pulse with object structure.

Finally, ALS-based inventories rely on biophysical reference data obtained from sample forest
2.4. MEASURING STEM ATTRIBUTES USING TLS

Availing upon advances in 3-D surveying hardware (Lichti et al., 2000; Lichti et al., 2002), pilot studies focused on the identification and measurement of tree stems, by two-dimensional (2-D) circle fitting of a slice of points at 1.3 m above ground. A height of 1.3 m is typically chosen as it allows convenient extension to the measurement of DBH. Simonse et al. (2003) reported DBH estimation within ±5.8 cm, with errors attributed to branches, which prevented unimpeded measurement of the main stem. To overcome this, the authors suggested future studies could utilize diameter measurements at variable heights above ground in order to reduce the effect of anomalous point samples. This approach was taken by Hopkinson et al. (2004), wherein a geometric cylinder fit was applied to points between 1.25 m and 1.75 m above ground, after manual detection of tree location. DBH estimation was achieved with coefficient of determination ($R^2$)=0.85. Moreover, by incorporating TLS-derived height, stem volume was estimated to within 7% of manual field measurements. However, the authors cautioned that in complex forests, i.e., other than single-tier plantations used in this study, substantial manual measurements or sophisticated algorithms would be required.

Subsequent studies increased the level of automation (e.g., Liang et al., 2012) and application to a range of forest types, for example stands with greater stem density (Liang et al., 2012; Watt and Donoghue, 2005; Brolly and Király, 2009), terrain variation (Maas et al., 2008) and species...
2.4. MEASURING STEM ATTRIBUTES USING TLS

heterogeneity (Moskal and Zheng, 2012). The reliability of forest parameter estimation in forest stands with greater stem density is of particular importance, because occlusion of the laser beam reduces the quality of information that can be obtained (Watt and Donoghue, 2005). To investigate these effects, Watt and Donoghue (2005), scanned densely stocked plantation forest plots, with stem densities up to 2800 stems·ha⁻¹. DBH retrieval for a small sample of 10 stems was achieved with $R^2$=0.92, with errors attributed to partial occlusion. Likewise, Brolly and Király (2009) estimated DBH from high-resolution (0.9 mrad) data collected in an unmanaged forest plot, with reported root mean square error (RMSE)'s of 4.2 cm and 7.0 cm obtained for geometric circle and cylinder fitting, respectively. A review of DBH estimation was provided by Maas et al. (2008), which compared DBH estimation for a combination of instrument types, forest plot characteristics, and data collection modes. DBH estimation was reported to within 1.8 cm for multiple-scan, high-resolution, 0.25 mrad data, with suggestions for future work directed at improving the technique for estimating DBH using various “slices” above ground.

Methods for DBH measurement also were applied at various heights above ground in an effort to model stem taper and form. The majority of studies used multi-scan data; the basic approach is to isolate point subsets at various slices above ground and then fit circles to the projected 2-D data (Aschoff et al., 2004; Henning and Radtke, 2006a; Király and Brolly, 2008; Olofsson et al., 2014), or cylinders to the 3-D data (Thies et al., 2004; Pfeifer and Winterhalder, 2004; Liang et al., 2014). Diameter estimates generally decreased in precision vertically along the stem, due to reduced point density, projected area effects on the predominantly vertical tree stems, and the increased occlusion due to canopy at successive heights above ground (Henning and Radtke, 2006a). Henning and Radtke (2006b), for example, achieved stem diameter errors of 1 cm below the live crown, but less accurate estimates (2 cm) at heights up to 13 m, when compared to manual measurements obtained from felled trees.

Numerous variants of the traditional circle and cylinder-fitting approaches have been proposed; for example Pfeifer and Winterhalder (2004) fit free-form curves to account for stem cross-sections that were not completely circular, with diameter errors of $\pm 1 – 2$ cm attributed to both surface roughness and point noise. Forsman and Halme (2005) projected points onto a plane
orthogonal to growth to better model stems, which were not perfectly vertical. Precision was evaluated by comparing the derived circle radius for a single cylindrical “pillar” used to simulate a tree stem indoors. Larger errors (up to 2.7 cm) were attributed to the relatively coarse, 4.2 mrad beam diverge (cf. the 15 mrad beam divergence of the SICK system used in this study (SICK, 2009).

In an effort to increase efficiency of data acquisition, other studies focused on single-scan data, despite the reduced accuracy and precision (Maas et al., 2008). The majority of techniques have included circle fitting (Bienert et al., 2007; Lindberg et al., 2012), cylinder fitting (Moskal and Zheng, 2012; Liang et al., 2012) and line detection of stem edges in the 2-D Andrieu range images (van Leeuwen et al., 2013). For example, Bienert et al. (2007) reported a standard deviation of 2.48 cm for diameter measurements obtained along 10 cm intervals for 22 Sitka spruce trees after comparison to harvester data, with a slight underestimate (mean = -0.64 m) and a maximum error of 19.6 cm. Due to the challenge of obtaining truth measurement of 3-D stem structure using conventional techniques, many studies presented higher-fidelity models with quantitative validation limited to lower-level parameters such as DBH or detection accuracy. For example, Lindberg et al. (2012) compared diameter estimations at 1.3 m obtained from 0.17 mrad data to manual DBH measurements, achieving 3.8 cm RMSE. Alternatively, Moskal and Zheng (2012) performed geometric cylinder fitting for a horizontal slice of points centered at 1.3 above ground ($R^2 = 0.91$, RMSE = 9.2 cm). Liang et al. (2012) classified the point cloud (e.g., stem, branch, etc.) using local point distribution metrics, and then fit successive cylinders along the main stem. This allowed 73% tree detection accuracy for plots of up to 1022 stems·ha$^{-1}$, though quantitative stem measurement, e.g., DBH, was not assessed.

Several less-conventional techniques were presented for DBH estimation from single-scan data. For example the crescent moon method was applied by Király and Brolly (2008), with reported DBH estimation errors of ±2 cm for 50% of the samples. Finally, radiometric information was utilized by Lovell et al. (2011), who computed the angular span from an intensity transect. The latter approach (Lovell et al., 2011) was developed for a moderately low-resolution system (5 mrad beam divergence), and reported DBH errors between 4.3–9.1 cm for 19 trees whose angular
span was at least twice the system beam divergence. An analysis of the effect of single scan mode on the retrieval of stem form was presented by Pueschel et al. (2013). Results consistently showed that lower RMSE’s (0.66 – 1.21 cm) can be achieved for merged scan data, as opposed to RMSE’s of 1.39–2.43 cm for single scans when using standard least-squares circle-fit algorithms.

Extension to forest stands with high stem density proved more difficult, requiring novel techniques such as RAndom SAmple Consensus (RANSAC) for noise resistance (Olofsson et al., 2014) or extrapolation based on taper relationships (Tansey et al., 2009) to estimate upper-stem diameters, which were occluded by the canopy. Despite these efforts, occlusion of upper-stems remained a challenge, e.g., with Tansey et al. (2009) reporting unsuccessful stem volume estimation. These measurements were used to create precise facetized models (e.g., Aschoff et al., 2004), and to assess a number of tree variables such as stem form (Thies et al., 2004), stem volume (Liang et al., 2014), stem ovality (Pfeifer and Winterhalder, 2004), diameter profiles (Henning and Radtke, 2006a; Bienert et al., 2007), and height (Olofsson et al., 2014). This has had numerous applications in forest inventory, e.g., providing timber value estimates to within 7% (Murphy, 2008). Moreover, these techniques facilitated retrieval of plot-level attributes, such as basal area and stem density (e.g., Yao et al., 2011; Strahler et al., 2008; Lovell et al., 2011; Tansey et al., 2009). The reader is referred to van Leeuwen et al. (2011) for a review of relevant applications in forestry.

For applications where reconstruction detail is prioritized over sample size and collection efficiency, a number of studies have also demonstrated the ability of TLS to recover precise descriptions of tree architecture. Early approaches demonstrated the potential to perform skeletonization of branching structure based on morphology (Gorte and Winterhalder, 2004; Gorte and Pfeifer, 2004), but provided no quantitative results. In general, explicit validation of tree topology has been difficult, with many researchers resorting to validation via lower-dimensional variables, distributional metrics obtained from simulated data, or by limiting reconstruction to small saplings. For example, Delagrange and Rochon (2011) explored a clustering approach for reconstructing a sapling. Parameters such as branch length, and height of insertion into the main stem, were compared to manual measurements with mean absolute errors of 15 cm and 3 cm, respectively.
For larger trees, Schilling et al. (2012) repetitively applied the depth first search algorithm to generate tree branch topology. Quantitative validation of the results were limited to lower-level parameters, e.g., average deviations in DBH (2.1 cm), tree height (1.75 m), and crown base height (1.29 m). Other graph-theory approaches were explored by (Bucksch and Lindenbergh, 2008; Bucksch et al., 2010). Côté et al. (2011) and Côté et al. (2013) extended the skeletonization approach of Bucksch et al. (2010) by using the point cloud as attractors in a space colonization model using L-systems grammar (Prusinkiewicz and Lindenmayer, 1996), and then added foliage based on light availability models; this provided accurate modeling of distributional branch characteristics (diameter: \( \text{RMSE} = 21\% \), length: \( \text{RMSE} = 20\% \), and insertion angle: \( \text{RMSE} = 25\% \)).

Raumonen et al. (2013) demonstrated a fully-automatic technique for comprehensive tree description from high resolution (0.22 mrad) TLS data. The point cloud was filtered, split into cover sets, segmented, and then reconstructed as a series of cylinders. Artificial tree models were used to generate synthetic point clouds, from which reconstruction results could be assessed relative to “truth” tree architecture, and without the effects of wind, etc. Parameters such as total branch/stem volume, branch length distributions, and stem taper were modeled, but without truth data for comparison. Additionally, manual caliper measurements were obtained for a sample of small branches, with diameter accuracy achieved commensurate to the TLS system resolution.

Other researchers explored a voxel analysis approach (Vonderach et al., 2012; Hosoi et al., 2013; Lefsky and McHale, 2008). Vonderach et al. (2012) reported total woody volume estimates within -5.1%–14.3%, as compared to control values obtained by harvesting and weighing nine deciduous trees. Direct, manual measurement of the TLS point cloud was used to provide truth volume measurements by Hosoi et al. (2013), with reported estimation errors of 0.5% and 34% for the main stem and cumulative branches, respectively. Lefsky and McHale (2008) reported distributional statistics on branch volume and length, in addition to the fit between modeled and measured stem diameters \((R^2=0.98)\).

Finally, limited by the complexity of the point cloud, other techniques relied on manual digitization (Eysn et al., 2013), semi-automatic retro-engineering software (Dassot et al., 2012; Delagrange et al., 2014). Eysn et al. (2013) compared modeled stem diameters at 1 m intervals
above ground to manual measurements obtained from the point cloud, with standard deviation of residuals between 1–2 cm. Dassot et al. (2012) measured woody volume using destructive sampling, and found that TLS-derived estimates were within ±10% for the main stem volume and ±30% for the cumulative branch volumes. Similarly, PypeTree, a visual modeling environment developed by Delagrange et al. (2014) utilized semi-supervised adjustment tools to address point cloud inaccuracies and improve reconstruction. Though small branches (length < 3.5 cm) were difficult to detect, errors in cumulative skeleton length were as low as 1.8% when compared to destructive sampling of two small saplings. These outputs may inform gas exchange models (Bienert et al., 2010), improve carbon-cycle estimations (Raumonen et al., 2013; Vonderach et al., 2012), or via linkages to wood fiber attributes, aid optimization of resource management (Côté et al., 2013).

The rapid increase in point cloud algorithms over recent years demonstrates the unique ability of TLS for extracting forest attributes. However, existing algorithms have so far been unable to extend the applicability of forest mensuration using TLS to more operationally practical, low-cost instrumentation. It remains to be seen whether recent technological innovations can be harnessed towards addressing the limitations in detail and scope of operational forest inventory (Olofsson et al., 2014), without exceeding the budgets of traditional methods (Mackrory and Daniels, 1995). To gain traction in an operational capacity, instrument and operation costs need to be reduced to meet the traditional inventory budget (Pueschel et al., 2013; Bi and Wang, 2010). We have developed a new TLS system, assembled from readily-available instrumentation and components, which offers an affordable (US$10k) and rapid (40 second scan time; 5 second set-up time, mobile platform) lidar solution for detailed forest structure assessment. This hardware development has the potential to address the existing knowledge gaps, but leaves uncertainty as to the validity of existing measurement techniques for a low-cost system, which is subject to several critical limitations. These limitations include a coarse beam divergence (15 mrad) and sampling resolution (4.13 mrad) that is two orders of magnitude coarser than previously used scanners. Furthermore, no marker-free registration software is provided by the manufacturer, limiting analysis to single scans. As a result of this knowledge gap, we identified a first objective
as follows:

1. Assess the feasibility of a low-cost, low resolution (spatial impulse and angular sampling) TLS for automatic forest stem inventory.

2.5 Pairwise Registration of TLS data

Registration of point cloud data is an important precursor to data analysis using TLS (Grant et al., 2012). The majority of forest inventory studies utilize multiple, co-registered scans in order to avoid the critical limitations of data obscuration due to laser occlusion. The need for registration of TLS data is much more pervasive than inventory, however, with application to robotics/mobile perception (Forsman and Halme, 2005), mapping, and the majority of techniques for canopy structure assessment.

Registration is the process of aligning data into a common coordinate system. We can define two types of registration, (i) relative, i.e., the combination of data from multiple scanner positions into a single scanner’s coordinate system, and (ii) absolute, i.e., the referencing of all data to an absolute global coordinate system (Bi and Wang, 2010). We will focus on relative registration techniques in this literature review. Registration is performed by estimating the three translation and three rotation parameters between two coordinate systems and then modifying the data’s spatial coordinates based on these transformation parameters (Hilker et al., 2012b).

One technique for registration is to measure the position and orientation parameters of each TLS system. A differential global positioning system (GPS) or total station is used to survey the precise location of each scanner location, and an inclinometer or inertial measurement unit (IMU) is used to measure the instrument’s orientation. Note that some TLS instruments have a motorized head, which automatically levels the z-axis, requiring only azimuthal correction (Hilker et al., 2012b). In a study by Van der Zande et al. (2006) a SICK sensor, similar to the one in this dissertation, was stepped laterally along a translation stage at several known positions adjacent to artificial tree. Data registration was then performed based on the precise knowledge of the scanner pose at each measurement location. Despite the controlled set-up, registration proved to
be the most difficult obstacle of this study, requiring manual corrections to compensate for small errors in the initial measurement. In addition to these challenges, manual measurement is labor-intensive and time-consuming (Van der Zande et al., 2006), and requires high-precision surveying equipment, which may not be operationally tenable to foresters (Hilker et al., 2012b). Based on these limitations, manual measurement based on surveying instrument location is arguably unsuitable for registration.

Note that in general, quantitative analyses of the precision of registration has been limited in the literature. This is because of two primary reasons. First, registration is often performed as a precursor to subsequent data analysis, and therefore quantitative validation is not the primary research focus. Second, the generation of truth data necessary to assess registration is difficult to obtain. Measuring sensor pose directly, e.g., as above (Van der Zande et al., 2006) is prone to errors, and therefore is not tenable as means to provide truth information for more automated approaches. As a result, most studies have reported a measure of the deviation between tie points after registration (cf. “tension” metric; (Cifuentes et al., 2014)), which provides a first-order estimate of the error, but is not supported by the rigor of explicit, true reference information. Given these challenges, quantitative results are provided in the following background where available, and as such, gaps in quantitative assessment may remain.

In contrast to the manual measurement approach of Van der Zande et al. (2006), other researchers have performed manual alignment of the point clouds after data collection, based on visual inspection. Yang et al. (2013), for example, adjusted the rotation and translation matrices of multiple scans in Pointools View Pro software, to align features such as trunk shapes, terrain patterns, and crown characteristics. However, some authors have cautioned that the identification of specific plant elements within a point cloud can be time consuming and subjective (Henning and Radtke, 2008). A positional accuracy of $\pm$ 20 cm was estimated based on the researchers’ experience, though no rigorous validation was performed. Despite the simplicity, the researchers’ cited some prohibitive disadvantages of this approach. Accuracy is dependent upon the subjective clarity of these features to the interpreter, and the process is labor-intensive and time-consuming; as a result it has not been advised for future work (Van der Zande et al., 2006).
A third class of registration algorithms involve manually placing artificial targets (or “markers”) in the scene, which serve as precise, unambiguous tie points (Yang et al., 2013; Van der Zande et al., 2006; Hilker et al., 2012b; Cifuentes et al., 2014). Commonly, retroreflective spheres mounted on poles are used (Hilker et al., 2012b), although reflective tape (Henning and Radtke, 2006a), plain A4 paper (Aschof et al., 2004), and checkerboards (FARO, 2012) have also been used. Targets may then be detected either automatically with some commercial software packages (e.g., FARO, 2012; RIEGL, 2005; Leica, 2014), or manually (Hilker et al., 2012b). Note that even with the automatic detection capabilities of commercial software, manual detection may still be required for missed targets (FARO, 2012). Given that the majority of available systems are commercial scanners with corresponding software packages (Bi and Wang, 2010), these methods have been widely used in forest assessment studies.

First, we review studies which utilized marker-based registration via commercial software packages. For example, Zheng and Moskal (2012) registered high-resolution (0.16 mrad sampling) TLS data using Leica Cyclone software with eight in-field reference targets, achieving a mean absolute error of 3.4 cm. Another study used FARO SCENE software (FARO, 2012) to register TLS data from nine positions in a 20 × 20 m plot (Cifuentes et al., 2014), but the reported accuracy “tension” metric precluded direct comparison to other studies. Pueschel (2013) used a set of FARO targets (four spheres and one planar target) and performed manual registration with very low registration errors ($\mu = 0.3 \text{ mm}$, $\sigma = 1-2 \text{ mm}$) for high resolution data (angular step width = 0.6 mrad).

Despite the capability of achieving high accuracy, the practicality and scope of these methods are limited by several key restrictions. First, automatic detection of targets is not necessarily trivial, and may require additional user interaction for identification of undetected markers (Pueschel, 2013; FARO, 2012). Second, target size and resolution parameter settings enforce a limit on the maximum distance between the sensor and target, in order for the target to be clearly identified (Pueschel, 2013; FARO, 2012; Henning and Radtke, 2006b; Aschoff et al., 2004). Finally, and perhaps most relevant, is the issue of cost: Commercial registration packages may be outside the budget constraints of many end users (Hilker et al., 2012b).
This budgetary restriction is increasingly salient as a growing number of affordable, small sensor providers (Hilker et al., 2012b) lack registration expertise (Bi and Wang, 2010) and provide no such software with their systems (cf. SICK, 2009). In response, some researchers explored means to automatically detect retroreflective spherical targets. Bienert and Maas (2009) performed automatic sphere detection in addition to manual measurement using the FARO SCENE software (FARO, 2012), and compared distances between tie points in both scans in order to assign tie point correspondence. Hilker et al. (2012b) developed a simple approach for registration, based on the use of artificial targets, which avoided the need for commercial software packages (Hilker et al., 2012b). The proposed method used 0.2 m diameter reflective polystyrene spheres mounted on wooden pegs; their location was first identified in the image based on manual field-measurement of the location and bearing, and then tie points were extracted for each sphere by averaging the positional coordinates of high intensity returns. Note that averaging the location of points sampled on a 0.2 m hemisphere will give an ambiguous center coordinate (tie point), depending on sensor perspective, which we consider a major contributor to the higher reported tie point correspondence errors (RMSE = 0.04-0.7 m and $R^2 = 0.70-0.99$). One disadvantage of this approach is its reliance on accurate reflectance information for automatic thresholding of points on the sphere, and limitations of practicality related to the use of manual targets.

Transportation and placement of supporting bases and reflective targets is cumbersome and tedious (Calders et al., 2014; Bienert et al., 2006; Pueschel et al., 2013; Cifuentes et al., 2014), and requires additional personnel, equipment, and time (Calders et al., 2014). For example, while the scan time duration for the RIEGL VZ-400 is 1.5 minutes, Calders et al. (2014) reported that setting up registration targets and collecting just five scans (center and plot corners) would take 2-4 hours, in comparison. In addition to time requirements, it may also be difficult to place targets in positions that can be seen from multiple viewpoints (Henning and Radtke, 2006b). As a result, some studies have raised concerns of the practicality of using artificial targets in forested environments (Aschoff et al., 2004) This is because forest mensuration necessitates maximizing the sample size and performance ability, while minimizing the required time, personnel, and cost (Lovell et al., 2011; Bi and Wang, 2010; Pueschel et al., 2013). The use of artificial targets
severely reduces mobility and efficiency, and represents a major limitation to the utility of TLS for operational objectives, especially in terms of the number of samples that can be measured (cf. Pueschel et al., 2013, Tables 4 and 5). Marker-free registration techniques are needed to reconcile these current limitations (Pueschel, 2013).

The most common technique for marker-free data registration is based on the iterative closest point (ICP) algorithm, which minimizes the Euclidean distance between points in regions of overlap, based on an initial estimation of sensor pose (Bi and Wang, 2010). First introduced by Chen and Medioni (1991) and Besl and McKay (1992), there is continued work to improve this method for registration, which suffers from drawbacks such as convergence to local minima and the requirement of estimating an initial transformation (which is not easy, and prevents full automation) (Bi and Wang, 2010; Henning and Radtke, 2008). Subsequent modifications have sought to automatically determine appropriate initial alignments (Chetverikov et al., 2005; Kim et al., 2004) and to avoid convergence to a local minima (Luck et al., 2000; Dalley and Flynn, 2002; Gelfand et al., 2005). Relevant information on the quantitative accuracy/precision afforded by these methods for forest point cloud data is reserved for subsequent discussion in the context of several applicational studies.

Several reviews of marker-free point cloud data registration approaches can be found in the literature (e.g., Williams et al., 1999; Mian et al., 2005; Salvi et al., 2007). However, these marker-free registration techniques have been primarily developed for robotics, computer vision, and engineering applications, and are ill-suited to the forest environment (Henning and Radtke, 2006b; Henning and Radtke, 2008; Bi and Wang, 2010). This is due to the complex, irregular shape of natural elements (Cifuentes et al., 2014), and the sparsity of forest structure, such that the visible sampled surfaces are vastly different even with small perspective shifts. In contrast, most studies applying ICP in the literature have been based on data with significant overlap (> 50%), only one surface of interest, and relatively high point density (Henning and Radtke, 2008).

Nevertheless, due to availability, convenience, and a lack of other options, these techniques have been occasionally applied in forest environments, but with unsatisfactory results. Hopkinson et al. (2004) used the IMAlign module of the Polyworks software suite (InnovMetric Software,
2.5. PAIRWISE REGISTRATION OF TLS DATA

2007), which computes the best fit transformation parameters using an iterative analysis of point cloud residuals in the region of scan overlap after initial coarse registration. As this registration step was simply a precursor to stem modeling, quantitative assessment of the results were not given. And Calders et al. (2014) used the Multi Station Adjustment algorithm from RIEGL RiSCAN PRO software, which employs a similar ICP technique for combining TLS and ALS data. Registration precision was assessed in terms of the standard deviation between the ALS and TLS-derived digital elevation model (DEM) (5–7 cm). It was reported that these techniques faced difficulties in the forest environment because of the lack of unique features (especially for homogeneous plantations) and difficulty of point correspondence.

As a result of these limitations of point correspondence and complex natural features, there is a need for marker-free registration approaches, which are robust in the forest environment. Such approaches could maximize collection efficiency and mobility in-field, while still providing the benefits of multiple-scan information. A hybrid marker-free approach was presented by Bienert and Maas (2009). A single spherical target was used in conjunction with detected stem objects in order to extract planes and ultimately tie points. This provided coarse registration, though detailed quantitative results were not presented. Another technique was presented by Henning and Radtke (2006a) and summarized in Henning and Radtke (2006b) and overcame the issue of point correspondence by making some geometric assumptions about the tree objects prior to applying the ICP algorithm. Coarse alignment was first performed based on either reflective tape affixed around tree stems at breast height (Henning and Radtke, 2006a) or measurement of the scanner location and orientation (Henning and Radtke, 2006b). Then, to avoid the problem of different scanner positions sampling different sides of an object surface (e.g., of the tree stem), the authors assumed that tree boles were approximately circular in cross-section and extracted the centers of the tree stem at various heights above ground. Pairwise registration was then performed using these (perspective-independent) data with the ICP algorithm. These constraints reduced the error in the x-y direction (maximum reported x-y registration errors of 2.1 cm); ground surfaces were also incorporated to reduce error in the z direction. Though this unique approach demonstrated the potential for utilizing point sets, which were invariant to view or perspective
differences, its scope was limited, as high resolution (1.05 mrad angular sampling) data were used, vegetation was chemically controlled to reduce occlusion, and registration focused on just three clearly visible trees within 5 m of four scanner positions.

The authors later extended this approach by extrapolating stem-center tie points at various heights above ground, and incorporating tie points from the terrain, in order to more widely distribute tie points into the volume (Henning and Radtke, 2008). A two-stage procedure first modified orientation about the x and y axes, and shifts in the z axis, based on terrain tie points. A second stage then accounted for rotation about the z axis and shifts along the x and y axes using stem tie points. Alignment error, reported as the mean Euclidean errors between corresponding tie points, was as low as 0.16 cm after the second stage, although removal of poorly matched tie points may have biased results. Despite the novelty of deriving explicit tie points from natural objects, this approach required an initial coarse registration, based on field-recorded positions and scanner orientations. As such, this technique is limited by many of the same challenges faced by the ICP algorithm. Initial orientation errors were assumed to be no more ±3°, which may limit the potential for rapid, operational data acquisition.

Despite the advances in feature estimation from natural surfaces introduced by Henning and Radtke (2008), registration remains one of the most pressing challenges for 3-D data processing (Kang et al., 2009). A blind approach, which is invariant to initial sensor pose, would greatly increase the efficiency of TLS data acquisition. A critical knowledge gap, therefore, is the ability to perform automatic, blind, marker-free registration of TLS data in forest environments, which is necessary to maintain data acquisition efficiency and support rapid operational forest structure assessment (Pueschel, 2013).

Consequently, we identified a second objective as follows:

2. Determine the error associated with automatic, blind, marker-free registration of TLS data pairs in forest environments
2.6 Multi-view Registration of TLS data

TLS has emerged as an effective tool for rapid and comprehensive measurement of object structure. Although initially developed for applications in the built environment, increasing potential has been shown in the assessment of forest structure. A persistent challenge, however, concerns the registration of data collected from multiple scanner locations into a single common coordinate system (Kang et al., 2009; Pingi et al., 2005; Theiler and Schindler, 2012; Stamos and Leordeanu, 2003). Relative registration is performed by estimating the three translation and three rotation parameters between two coordinate systems and then modifying the data’s spatial coordinates accordingly (Hilker et al., 2012b). This registration is often a necessary preprocessing step in order to reduce occluded areas (Eo et al., 2012; Sharp et al., 2004; Salvi et al., 2007; Weinmann et al., 2011) and compensate for decreased range-dependent point density and resolution (Henning and Radtke, 2008). Moreover, it allows for multitemporal analyses or permanent monitoring (Henning and Radtke, 2008), and has applications to virtual reality (Stamos and Leordeanu, 2003), urban planning (Stamos and Leordeanu, 2003), and other domains. In contrast to the previous section, which examined approaches for pairwise registration, this section focuses on multi-view registration—the alignment of multiple scans into a common coordinate system.

Of particular interest is the registration of TLS data in forest environments. This is necessary for extraction of dendrometric parameters (Bucksch and Khoshelham, 2013; Zhou et al., 2014), canopy assessment (Henning and Radtke, 2008), and plot-level inventory (Kelbe et al., 2015b). Ultimately, multisensor registration between ALS and TLS data could provide synergistic structural ground truth data to support calibration/validation of large-scale, airborne sensing models (Henning and Radtke, 2008). Despite this need, the majority of registrations algorithms in the literature have limited use in forest environments due to factors such as occlusion, spatial variability, and movement, e.g., due to wind, (Henning and Radtke, 2008). This is confounded by system contributors, such as the range-dependent point density and discrete sampling nature of laser scanning technology (Barnea and Filin, 2008). As a result, small sensor-displacements may yield drastic changes in scene content (Forsman and Halme, 2005), which challenge the establishment of reliable point or feature correspondence (Zhou et al., 2014).
Traditionally, and as discussed previously, this challenge is overcome by placing manual targets in the scene, which serve as control points for marker-based registration (Theiler and Schindler, 2012; Hilker et al., 2012b; Bucksch and Fleck, 2011). However, the use of targets is time-consuming, tedious and costly (Zhou et al., 2014). As a result, marker-free techniques are preferred, in order to improve field-scanning efficiency (Zhou et al., 2014) and make TLS cost-competitive relative to traditional forest inventory techniques (Ducey and Astrup, 2013). Unfortunately, the majority of existing marker-free techniques utilize iterative point matching (e.g., ICP) or surface matching (Huber and Hebert, 2003), both of which are successful only for engineered surfaces (Henning and Radtke, 2008).

Recent automatic, marker-free registration approaches, such as (Kelbe et al., 2015a) and (Henning and Radtke, 2008), offer the potential to rectify this disparity and improve the operational capabilities of TLS in forest environments. Previously, we developed a marker-free registration approach, which extracted view-invariant tie points derived from the modeled tree and terrain geometry (Kelbe et al., 2015a). Pairs of scans were then registered with an efficient voting method based on geometric constraints. Finally, an embedded upper-bound error metric was provided with each output transformation parameter set by exploiting circular self-closure along a composite transformation between disjoint tie point subsets. This provided a robust solution for pairwise registration of forest terrestrial laser scanner data pairs. However, multiple (i.e., > 2) scans are often collected in a forest plot, requiring multi-view registration (Huber and Hebert, 2003). Multi-view registration offers the potential to identify and remove locally consistent, but globally incorrect matches (Huber and Hebert, 2003), and to bring into alignment disconnected scans through a connected sequence. However, due to the large nonlinear search space and the volume of input TLS data involved, multi-view registration is considered more challenging than pairwise registration (Stamos and Leordeanu, 2003). This research extends the pairwise framework of (Kelbe et al., 2015a) to perform multi-view registration.

Little attention in the literature has been given to multi-view registration of forest TLS data. For pairwise registration of single trees, several authors have presented a local registration approach based on alignment of tree skeletons (Zhou et al., 2014; Bucksch and Khoshelham, 2013). These
techniques extracted skeletons from the point clouds and then applied a minimization between
the point cloud and skeleton which allowed for local variation in transformation parameters at
the branch level. Bucksch and Khoshelham (2013) reported an average registration error of 5 mm,
although this was based on a coarse initial alignment of 15 mm. Zhou et al. (2014) extended this
approach so that a coarse initial alignment was not required, however the authors did not provide
quantitative validation of the results. For registration of plot-level TLS data, e.g., including
multiple trees, Henning and Radtke (2008) presented a two-stage approach based on tie points
derived from the terrain and modeled stem axes. Mean Euclidean errors between tie points were
as low (0.16 cm), however were dependent upon initial coarse pose estimates, e.g., orientation
errors of ≤ 3°.

Given the limited background on multi-view registration of forest terrestrial laser scanner data,
a review of existing approaches in other domains and sensing modalities provides additional con-
text on the state-of-the-art. Multi-view registration techniques are classified as either sequential,
simultaneous, or hybrid. Sequential alignment iteratively registers subsequent pairs of data from
an ordered sequence (e.g., A to B, B to C, C to D). Although this has inherent applications to se-
quential video frames or linear sampling protocols, it is subject to propagation and magnification
of errors throughout the sequence (Henning and Radtke, 2008; Pingi et al., 2005; Kim and Hong,
2003; Kang et al., 2000). As a result, simultaneous registration is considered optimal (Bergevin et
al., 1996; Blais and Levine, 1995; Jokinen and Haggrén, 1998). Simultaneous, or global registration
(Kang et al., 2000; Pingi et al., 2005) utilizes poses estimates between all pairs of scans to minimize
the accumulated transformation errors by distributing them throughout the rigid network (Pulli,
1999). Moreover, because the overlap area between all scans is used (as opposed to just a pair),
there is a greater potential to identify and utilize tie points that are dispersed throughout the
volume, thus improving registration results (Henning and Radtke, 2008). The final set class of
techniques, hybrid approaches, incorporate both sequential and simultaneous elements.

Graphical frameworks, which encode connectivity between overlapping views, have been
widely used for multi-view alignment. Typically, a node represents a single input view, sensor,
image frame, or point cloud (Kang et al., 2000; Huber and Hebert, 2003). Likewise, an edge repre-
sents a connection between nodes, as determined from pairwise registration. A video sequence, for example, would be represented as a predominantly linear graph due to the temporal adjacency of neighboring frames (Kang et al., 2000). Associated with each edge is a relative pose (Huber and Hebert, 2003). Typically, a reference node is chosen as the world coordinate system (WCS). The absolute pose between two pairwise-disconnected views can then be determined by composing the relative poses associated with each edge along a path connecting the scan to the reference WCS (Huber and Hebert, 2003).

A potential solution to multi-view registration exists when the graph is connected (i.e., a composite transformation path exists between each node and the reference) (Huber and Hebert, 2003). A minimal solution is defined by a spanning tree, which is a connected graph with no cycles. Additional edges introduce cycles in the graph, which may result in pose conflict due to the composition of pairwise transforms along different paths between views (Huber and Hebert, 2003). These pose inconsistencies are caused by small errors in pairwise pose estimates, which are accumulated through a graph path.

The network of pairwise correspondences are encoded in an adjacency matrix. An unweighted adjacency matrix assigns each \((i, j)\) element as “true” or “false” based on the existence of an edge connecting nodes \(i\) and \(j\). To the contrary, a weighted adjacency matrix provides a value or weight associated with each edge, i.e., based on image correlation (Kang et al., 2000), geometric distance (Kang et al., 2000), spatial overlap (Pingi et al., 2005), tie point registration error (Bendels et al., 2004), or the number of corresponding feature pairs within fixed thresholds (Stamos and Leordeanu, 2003). Graphs are typically undirected, i.e., no directional information is encoded between edges. Many previous studies have demonstrated the utility of graphs to perform sequential, simultaneous, and hybrid multi-view alignment.

Sequential alignment avoids the issue of pose conflict by finding the minimal spanning tree that connects all nodes to the reference node, using an assessment of “minimum path length” performed on the adjacency matrix. Because the spanning tree is acyclic, the sub-graph is guaranteed to be pose-consistent (Huber and Hebert, 2003). For example, Stamos and Leordeanu (2003) defined pairwise edge weights according to the number of corresponding line pairs detected in
urban point clouds, and then used Dijkstra’s algorithm (Dijkstra, 1959) to perform sequential alignment of each node into the WCS, based on a weighting of correspondence pairs. The authors computed registration errors as the average distance between matched planes, with reported errors between 0.1–5.6 cm. A sequential shortest path technique has also been applied in urban image mosaicing domains (Kim and Hong, 2003; Kim and Hong, 2006; Kang et al., 2000; Bendels et al., 2004). This sequential alignment strategy utilizes all possible pairwise connections, while rejecting weak fits, but does not solve the correspondence problem simultaneously (Stamos and Leordeanu, 2003), and does not exploit the redundant information provided by multiple edges in order to reduce registration error.

Graphical methods have also been applied to simultaneous multi-view registration, often by linear optimization of pose parameters to minimize registration error. For example, Huber and Hebert (2003) performed multi-view point cloud registration of small (40 cm) manmade objects by building a subgraph containing only correct pairwise matches. Global consistency was used to eliminate bad pairwise matches, with absolute poses adjusted to minimize surface overlap distance. No ground truth poses were available for quantitative validation, therefore synthetic data was generated with added Gaussian noise (σ=1 mm). The reported maximum correspondence error (MCE) was <1.2 mm. Eo et al. (2012) applied generalized procrustes analysis (GPA) to simultaneously adjust registration and found favorable results compared to sequential registration with ICP in urban point cloud data. The proposed method was validated by comparing transformation parameters to those obtained using the commercial PolyWorks software InnovMetric Software, 2007. Translation differences between techniques were < 3 cm, with both methods reporting an RMSE of 8 cm.

A third class of “hybrid” registration algorithms includes elements of both sequential and simultaneous registration. In order to reduce propagation of alignment errors and exploit multiple cycles through the graph, additional steps may be added to sequential registration, such as global averaging (Sharp et al., 2004) or cycle detection (Kang et al., 2009; Borrmann et al., 2008), which distribute errors across the path sequence. For example, Kang et al. (2009) collected data of an urban environment in a ring scheme, providing circular self-closure in order to redistribute errors. A high
section FARO scanner was used (angular resolution of 0.6–0.8 mrad), with error evaluated by comparing the RMSE between select corresponding tie points. Pairwise RMSE between tie points were reported to be < 5 mm, despite much larger closure errors through the circular network of 20 scans (1.6–6 m prior to error redistribution). After redistribution errors throughout the network, RMSE’s were ≈ 3.5 cm. For range data collected from autonomous robots, Borrmann et al. (2008) performed loop detection in order detect closed edge cycles in the sequential graph. Validation was performed by adding random errors to the initial pose estimates obtained from marker-based registration (assumed to be truth), and then computing the error of output parameters relative to the true pose estimates. This allowed even redistribution of errors between scans, with average positional errors of 5 cm. For arbitrary sampling protocols, Sharp et al. (2004) decomposed the model graph into a series of range-image cycles, so that nonlinear optimization could be performed over each basis cycle in closed form. Cycles were then reintegrated into a global model using an averaging technique. This afforded the advantages of sequential registration, while adding a secondary “global” pose adjustment to minimize error. Quantitative results were not presented.

Limited attention has been paid to multi-view registration in the forest environment, despite the expressed need (Henning and Radtke, 2008). Registration techniques that leverage global consistency to remove erroneous local matches and reduce propagation errors, i.e., either simultaneous or hybrid approaches, are preferred for multi-view registration, although there is still no consensus as to the best approach (Sharp et al., 2004). This research presents a hybrid multi-view registration approach for blind, marker-free registration of forest terrestrial laser scanner data, and compares it against standard sequential and simultaneous registration approaches. This work builds on Kelbe et al. (2015a) by providing an automatic, blind, marker-free, multi-view registration of a network of TLS scans collected at arbitrary locations within a forest plot. A primary contribution is the integration of embedded error metrics associated with each pairwise transformation (Kelbe et al., 2015a), which simplifies the global alignment problem, while adding resistance to noise (Huber and Hebert, 2003). Previously, a variety of heuristics have been used to assess the quality of a pairwise “edge”, including image correlation (Kang et al., 2000), geometric distance (Kang et al., 2000), spatial overlap (Pingi et al., 2005), tie point registration error (Bendels...
et al., 2004), or the number of corresponding feature pairs within fixed thresholds (Stamos and Leordeanu, 2003). However, none have provided adequate information on the precision of a pairwise transformation. We extend the pairwise error metric of Kelbe et al. (2015a) to provide an embedded multi-path confidence metric associated with each transformation model.

A critical knowledge gap, therefore, is the ability to perform automatic, blind, marker-free registration of TLS data in forest environments, which is necessary to maintain data acquisition efficiency and support rapid operational forest structure assessment (Pueschel, 2013).

Consequently, we identified a third objective as follows:

3. Determine the error associated with multi-view, marker-free registration of TLS data in forest environments

Specific objectives are to (i) validate the proposed embedded error metric for multi-edge paths through a graphical network, (ii) assess the performance of the proposed, hybrid multi-view registration technique for TLS data collected in New England forests, and (iii) demonstrate the improvement of plot-level inventory assessment compared to single-scan data collection.
Chapter 3

Measuring stem attributes

3.1 Foreword

From a comprehensive review of the literature (chapter 2), we found that TLS has demonstrated a potential to alleviate the limitations of traditional forest inventory in terms of efficiency, repeatability, and the fidelity of structural information that can be recorded. However, despite significant research focus over the past decade, its application to operational inventory has been limited. This is due in part to the high cost (Pueschel et al., 2013; Clawges et al., 2007) of systems primarily designed for engineering, architecture, and forensics (Danson et al., 2007; Henning and Radtke, 2006b; Yang et al., 2013). A recently-developed, mobile laser scanning system provides the cost-effective hardware necessary to rectify this knowledge gap, but leaves uncertainty as to the validity of existing measurement techniques for a low-cost system, which is in turn limited by angular sampling resolution, registration, and laser beam divergence.

Therefore, we identified a first objective as follows:

1. Assess the feasibility of a low-cost, low resolution (spatial impulse and angular sampling) TLS for automatic forest stem inventory.

This chapter describes in detail the study, methodology, and results related to objective 1. Outputs include an original, robust methodology to model visible tree stem structure based on
3.2 Abstract

Despite the active research, terrestrial laser scanning (TLS) has remained underutilized for forest structure assessment due to reliance of processing algorithms on high-resolution data, that may be costly and time-consuming to collect. Operational inventories, however, necessitate maximizing sample size while minimizing time and cost. The objective of this study was to assess the performance of a novel technique that enables stem reconstruction from low-resolution, single-scan TLS data in an effort to satisfy performance criteria against operational acquisition constraints. Instead of utilizing the curvature of the tree stem, e.g., by circle or cylinder fitting, we take advantage of the sensor-object geometry and reduce the dimensionality of the modeling to a series of 1-D line fits. This allowed robust recovery of tree stem structure in a range of New England forest types, for tree stems, which subtended at least an angular width of 15 mrad—the beam divergence of our system. Assessment was performed by projecting the 3-D data onto a 2-D images and evaluating the per-point classification accuracies using manually-digitized truth maps. Manual forest inventory measurements were also collected for each 20×20 m plot and compared to measurements derived automatically. Good retrievals of stem location ($R^2 = 0.996$, RMSE = 0.17 m) and diameter at breast height ($R^2 = 0.80$, RMSE = 0.06 m) were achieved. This study demonstrates that low-resolution sensors may be effective in providing data for operational...
3.3. INTRODUCTION

Key to effective management of forest natural resources is information obtained through a systematic collection of forest attribute data (Tansey et al., 2009) referred to as “forest inventory” (Hamilton, 1975). Decades of research have made traditional forest inventory the underpinning of forest study, management, and policy (Tansey et al., 2009). Yet, in the context of both dramatic changes to the environment and rapid technological innovation, traditional mensuration techniques do not always provide a desired level of structural fidelity. Increased characterization of forest structural attributes is desirable, as it could improve our understanding, management, and regulation of resources such as timber and fuel (Klemperer, 1996), services such as animal habitat provision (Lindenmayer and Franklin, 2002) and carbon cycling (Gibbs et al., 2007), and risks such as erosion (Booth et al., 2002), forest fires (Chandler et al., 1983), and invasive species (Pimentel et al., 2005).

Recent technological advancements have demonstrated the capacity of laser scanning to rapidly record detailed structural information both on the ground and - for large-area operations - from air and space (Bachman, 1979). Range is computed from either phase differences or return travel time for an emitted laser pulse. Emitted laser pulses are coupled to platform movement and deflected by a rotating mirror to rapidly interrogate the scene. For each pulse, interaction with object structure causes a deformation in the temporal profile of the backscattered pulse, providing a precise and measurable record of object location.

Airborne laser scanning (ALS) has matured to operational use over the past decade for large-scale forest structure assessment, e.g., (Wehr and Lohr, 1999; Lefsky et al., 2002); the reader is referred to Maltamo et al. (2014) for a detailed review. Typically, ground-based field measurements are collected within sample plots and then used to develop empirical calibrations between observed ALS data and biophysical parameters at the plot or stand level (Hauglin et al., 2014). In recent years, ALS systems with increasing point density have allowed extraction of tree-level structural attributes, including DBH (Bucksch et al., 2014), however limitations remain in terms
of the minimum footprint size and the range of possible incidence angles achievable with an airborne platform (Henning and Radtke, 2006b; Lovell et al., 2011). As a result, a high percentage of laser pulses are intercepted within the upper canopy (Hilker et al., 2010), making precise, direct measurement of sub-canopy tree structure difficult (Hosoi and Omasa, 2006; Hilker et al., 2012a).

Terrestrial laser scanning (TLS), on the other hand, operates from a ground platform and can directly measure sub-canopy tree attributes, with improvements in structural detail compared to manual forest inventory measurements (Liang et al., 2012). This offers the potential to both support ground-based forest inventory (Hopkinson et al., 2004; Maas et al., 2008) and complement the aforementioned limitations of ALS (Lovell et al., 2011; Hilker et al., 2012a). For example, biophysical reference data obtained using TLS could be used to calibrate ALS models for extrapolation over larger scales (Hauglin et al., 2014; Liang et al., 2012; Lindberg et al., 2012).

Pilot studies using TLS for forest structure assessment focused on the identification and measurement of tree stems, including location, height and diameter at breast height (DBH) using the Hough transform (Simonse et al., 2003) and geometric cylinder fitting (Hopkinson et al., 2004). Subsequent studies extended the level of automation and applicability of TLS to a range of forest types. Study areas included stands with greater stem density (Watt and Donoghue, 2005; Brolly and Király, 2009), sub-canopy complexity (Tansey et al., 2009), terrain variation (Maas et al., 2008), and species heterogeneity (Moskal and Zheng, 2012). Motivated by constraints of data collection efficiency, additional studies measured DBH from single-scan data, for example using the crescent moon method (Király and Brolly, 2008) or by computing the angular span from an intensity transect (Lovell et al., 2011). These techniques facilitated retrieval of plot-level forest structural attributes, such as basal area and stem density, e.g., (Tansey et al., 2009; Yao et al., 2011).

### 3.4 Methods

#### 3.4.1 TLS system

We used a new, low-cost terrestrial lidar system integrated from commercial-off-the-shelf (COTS) components by Rochester Institute of Technology. The system is based on a design first imple-
Figure 3.1: Low-cost terrestrial lidar system integrated from commercial-off-the-shelf (COTS) components. A SICK LMS-151 laser scanner, (a), is mounted to a rotation stage, (b), and tethered to a data logger, (c), and battery, (d), which are mounted on a backpack and worn by the operator. Instrument control is achieved via a wireless mobile application (e).
mented by a team at the Katholieke Universiteit Leuven, Belgium (Van der Zande et al., 2006). It was designed to overcome the limitations of high cost, low mobility, and long scan time, which have so far precluded operational forest structure assessment using terrestrial lidar. The emergence of this and other low-cost, low-resolution sensors (Hilker et al., 2012b) has accelerated the need to develop structural modeling algorithms, which are not constrained by high-resolution point cloud data. Unlike many commercial scanners that provide for high-density point cloud data, lower-cost systems are geared towards the efficient and fast sampling of structural data, often at a much lower resolution. This system has a minimum angular step-width of 4.36 mrad, and a beam divergence of 15 mrad, both approximately two orders of magnitude coarser than comparable instrumentation (See Table 3.1). These limitations provide an opportunity to address the knowledge gap in terms of structural algorithms, which are robust to low-resolution data:

The basic components of our TLS system can be seen in Figure 3.1. The sensor head is a SICK LMS-151 laser scanner, which is compact, lightweight (1.1 kg), and weather-resistant (SICK, 2009). A 905 nm laser is pulsed at 27 kHz with range measurement recorded based on time-of-flight. The laser pulse is deflected by a rotating mirror to sample a 270° arc, swept out in elevation angle. This sensor head is coupled to an azimuthal rotation stage, which provides coverage of the full hemisphere above the instrument and a portion of the hemisphere below (270° V x 360° H coverage). Up to two returns per outgoing pulse are digitized. The instrument is tethered to a data logger and battery, which are mounted on a backpack and worn by the operator. Sensor control is achieved via a wireless mobile application.
Table 3.1: Specifications of selected terrestrial lidar instrumentation (adapted from Dassot et al., 2011). Sensor resolution is defined by both the intrinsic laser beam divergence and the angular sampling step-width.

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Model</th>
<th>Availability</th>
<th>Range finder</th>
<th>Wavelength (nm)</th>
<th>Measurement range (m)</th>
<th>Range accuracy (mm)</th>
<th>Spot size at exit (mm)</th>
<th>Beam divergence (mrad)</th>
<th>Min angular step-width (mrad)</th>
<th>V×H field of view (°)</th>
<th>Max pulse frequency (Hz)</th>
<th>Weight (kg)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>RIEGL</td>
<td>LMS-Z420i</td>
<td>commercial</td>
<td>time-of-flight</td>
<td>1,550</td>
<td>2-1,000</td>
<td>10 (at 50 m)</td>
<td>8</td>
<td>0.25</td>
<td>0.07</td>
<td>80 × 360</td>
<td>11,000</td>
<td>18</td>
<td>RIEGL (2010)</td>
</tr>
<tr>
<td>Optech</td>
<td>Ilris-3D</td>
<td>commercial</td>
<td>time-of-flight</td>
<td>1,535</td>
<td>3-1,500</td>
<td>7 (at 50 m)</td>
<td>14</td>
<td>0.15</td>
<td>0.02</td>
<td>40 × 40</td>
<td>10,000</td>
<td>20</td>
<td>Optech (2009)</td>
</tr>
<tr>
<td>Z+F</td>
<td>Imager 5003</td>
<td>commercial</td>
<td>phase shift</td>
<td>780</td>
<td>1-53.5</td>
<td>3 (at 25 m)</td>
<td>3</td>
<td>0.22</td>
<td>0.03</td>
<td>310 × 360</td>
<td>500,000</td>
<td>16</td>
<td>Zoller + Fröhlich (2013)</td>
</tr>
<tr>
<td>Faro</td>
<td>Focus 3D 120</td>
<td>commercial</td>
<td>phase shift</td>
<td>905</td>
<td>0.6-153</td>
<td>2 (at 25 m)</td>
<td>3.8</td>
<td>0.16</td>
<td>0.16</td>
<td>305 × 360</td>
<td>976,000</td>
<td>5</td>
<td>FARO (2011)</td>
</tr>
<tr>
<td>SICK</td>
<td>LMS-151</td>
<td>experimental</td>
<td>time-of-flight</td>
<td>905</td>
<td>0.5-50</td>
<td>±30</td>
<td>8</td>
<td>15.0</td>
<td>4.36</td>
<td>270 × 360</td>
<td>27,000</td>
<td>5</td>
<td>SICK (2009)</td>
</tr>
</tbody>
</table>
3.4.2 Study Area

In order to assess the robustness of our stem reconstruction method across a range of New England forest types, our study area follows the ground validation work of the National Ecological Observatory Network (NEON) (Kampe et al., 2010). NEON is a continental-scale ecological monitoring platform designed to monitor changes in the biosphere in response to human impact. NEON has divided the U.S. into a series of distinct ecological domains in order to characterize the ecological diversity at the continent scale. The study area for this work corresponds NEON’s core site for the Northeastern ecological domain, including both Harvard Forest and Quabbin Reserve, Massachusetts, USA (bounded by 42.428° N, 72.284° W and 42.558° N, 72.170° W; WGS1984). Harvard Forest is a 1,200 ha reserve with a long history of ecological research and management. Quabbin Reserve is a 23,000 ha public forest and provides additional diversity, e.g., disturbance regimes.

NEON has set up twenty $20 \times 20$ m ground validation plots representing a diverse range of Northeastern USA forest structure. Plots were selected using a stratified random sampling scheme and evaluated in person to ensure forest structure variability. We investigated eleven of NEON’s twenty $20 \times 20$ m plots. Individual plot-level characteristics are shown in Table 5.1, and include a range of young, mature, single-tiered, multi-tiered, sparse, dense, deciduous, coniferous, and mixed forest types. Basal area (BA) ranged from 3.21 m$^2$ · ha$^{-1}$ to 73.73 m$^2$ · ha$^{-1}$. Stem densities were recorded separately for stems of DBH ≥ 10 cm and for stems of DBH < 10 cm, and ranged from sparse and mature (700 stems/ha, all of which were ≥ 10 cm DBH) to dense and young (2475 stems · ha$^{-1}$: 625 with DBH ≥ 10 cm and 1850 with DBH < 10 cm). Ground vegetation parameters (mean height ($\bar{z}$) and percent cover ($p$)) were also recorded, with plot attributes ranging from bare ground to full coverage of 1.1 m tall vegetation. A few sample lidar images of the plots are shown in Figure 3.6 to highlight the gradient of forest structure. The structural variability represented by our study area is unique among previous research and provides a diverse data set from which to evaluate the robustness of stem reconstruction across a range of forest types.
Table 3.2: Summary of ground validation plots in Harvard Forest, MA, USA.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Dominant Species (Secondary Species)</th>
<th>BA m² · ha⁻¹</th>
<th>Stems ha⁻¹</th>
<th>Ground veg.</th>
<th>Ground veg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>05</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>73.73</td>
<td>775</td>
<td>250</td>
<td>0.5</td>
</tr>
<tr>
<td>06</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>18.43</td>
<td>625</td>
<td>1850</td>
<td>0.9</td>
</tr>
<tr>
<td>08</td>
<td><em>Pinus strobus</em> - <em>Tsuga canadensis</em></td>
<td>55.70</td>
<td>600</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td><em>Pinus strobus</em></td>
<td>53.84</td>
<td>950</td>
<td>0</td>
<td>0.5</td>
</tr>
<tr>
<td>12</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>37.91</td>
<td>800</td>
<td>150</td>
<td>0.6</td>
</tr>
<tr>
<td>13</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>51.67</td>
<td>425</td>
<td>675</td>
<td>0.9</td>
</tr>
<tr>
<td>15</td>
<td><em>Tsuga canadensis</em> - <em>Betula alleghaniensis</em></td>
<td>51.56</td>
<td>700</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td><em>Tsuga canadensis</em> - <em>Betula alleghaniensis</em></td>
<td>70.33</td>
<td>725</td>
<td>250</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>Upland successional shrubland</td>
<td>3.21</td>
<td>225</td>
<td>425</td>
<td>1.1</td>
</tr>
<tr>
<td>19</td>
<td>Variant: <em>Tsuga canadensis</em> with mixed</td>
<td>66.38</td>
<td>950</td>
<td>125</td>
<td>0.6</td>
</tr>
<tr>
<td>31</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>40.96</td>
<td>525</td>
<td>125</td>
<td>0.7</td>
</tr>
</tbody>
</table>

3.4.3 Experimental Design

Our experimental design was to evaluate the performance of stem reconstruction using low-resolution single-scan TLS data collected across a range of forest types. Plots were first marked based on a center GPS coordinate, compass, and tape measure. Lidar data acquisition was designed to mimic an operational inventory survey and included single scan measurements at the center of each 20×20 m plot. Reconstructed stem models were validated against manually digitized truth maps based on the TLS point cloud images. Moreover, to evaluate the potential for forest inventory, manual measurement of forest parameters were recorded for comparison to estimates derived from the lidar data. Data were collected during August 2012 leaf-on conditions. Data collection during leaf-off conditions is preferable for measurement of woody structure, in that it reduces the effects of occlusion (Dassot et al., 2012; Raumonen et al., 2013; Côté et al., 2009). However, our objective was to evaluate the robustness of our algorithm given varying degrees of sub-canopy occlusion, with an eye towards supporting operational forest structure assessment. Since surveys are often done during growing season, leaf-on data collection allowed us to better evaluate the feasibility of our algorithm under these conditions.

3.4.4 Presuppositions

A natural scene contains a variety of complex object structures. Examples may include tree stems, branches, foliage, herbaceous vegetation, and soil. For the purpose of our stem detection algo-
rithm, let the tree stems be designated $S$ and the remaining structure designated as background, $B$. For each scan mirror angle, $(\theta)$, and rotation stage position, $(\phi)$, a laser pulse is emitted into the scene. The mirror scans over a zenith range of $270^\circ$; this is coupled to platform rotation of $180^\circ$ to fully sample the upper hemisphere and a significant portion of the lower hemisphere below the instrument. Moreover, the laser pulse width diverges with range $(r)$, in order to fully sample the portion of the sphere within the scanner field of view. The beam diameter, $d(r)$, can be expressed as

$$d(r) = d_{\text{exit}} + \left(2 \cdot r \cdot \tan \left(\frac{\Theta_d}{2}\right)\right) \quad \text{(3.1)}$$

where $d_{\text{exit}}$ is the beam diameter at the exit optic and $\Theta_d$ is the beam divergence with units [radians]. The distance between adjacent samples, $\Delta_s$, is also a function of range, defined as

$$\Delta_s(r) = 2 \cdot r \cdot \tan \left(\frac{\Theta_s}{2}\right) \quad \text{(3.2)}$$

where $\Theta_s$ is the angular step-width between consecutive scan lines, expressed in radians.

For each scan mirror angle, $(\theta)$, and rotation stage position, $(\phi)$, the return trip travel time of a laser pulse is digitized and converted to range based on the speed of light. This gives an unambiguous triplet, $(\theta, \phi, r)$, for each digitized pulse. A point cloud, $P$, is the aggregate of all digitized range measurements. Upon conversion from spherical to cartesian coordinates, we can define the point cloud $P = \{x_1, x_2, \ldots, x_n\}$ where $x_i \in \mathbb{R}^3$ is the $x, y, z$ position for the $i^{th}$ point in $P$.

We assume that the point cloud, $P$, is a finite set of samples of an object surface, $S$, and background surface, $B$. The task is to detect and reconstruct the underlying object surface, $\hat{S}$, from the unclassified sample measurements, $\{x_1, x_2, \ldots, x_n\}$. Mathematically, we wish to recover the underlying object structure, $\hat{S} = F(P)$.

This surface reconstruction is an example of an inverse problem, which for this case is ill-posed due to the absence of a unique solution (Åkerblom, 2012; Hubbard, 2006). Because of the limited view angles possible from terrestrial laser scanning, there will always be measurement gaps in the point cloud data due to occlusion of the emitted laser pulse by opaque objects. This is true
for multiple, co-registered scans, where a terrestrial platform precludes adequate sampling of the upper canopy, but even more so for single scans, which can only sample the surface visible to the instrument at one location.

This ill-posed problem can be constrained by imposing geometrical assumptions on the object surface. We assume that tree stems can be approximated as a series of contiguous conical frustums (herein referred to as “tapered cylinders”). A tapered cylinder segment, \( C \), is parametrized according to Fig. 3.2 with lower and upper axis locations \( l_0 = \{x, y, z\} \) and \( l_1 = \{x, y, z\} \), respectively, and radii \( r_0 \) and \( r_1 \). From these parameters, a lean angle \( \omega \) can be computed with respect to the z-axis (vertical). Additionally, the taper \( \tau \) can be computed based on the change in radius with height. We make several assumptions about the tapered cylinder model’s maximum and minimum parameter values. Additionally, a coverage parameter, \( c \), is defined to provide an estimate of the quality of fit. The coverage defines the percentage of the tree stem surface, which is sampled by the sensor. These parameters are reported in Table 3.3.

Figure 3.2: A tapered cylinder segment, \( C \), is parametrized with lower and upper axis locations \( (l_0 \text{ and } l_1) \), and radii \( (r_0 \text{ and } r_1) \). Additionally, these parameters define a lean angle \( (\omega) \) and taper angle \( (\tau) \).

The underlying terrain topography is modeled as a precursory requirement to stem modeling. Previous studies have either assumed the terrain is a simple flat plane, e.g., (Yao et al., 2011) or have modeled the terrain without quantitative validation. Similar methods have been employed in literature with assumed accuracy, e.g., (Thies et al., 2004); therefore quantitative validation is not performed in this paper. Nevertheless, errors in terrain modeling, primarily due to vegetation, which interferes with the sensing of the true ground surface, may cause an underestimation of the
Table 3.3: Parameter values used in this study

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radius</td>
<td>$r$</td>
<td>0.1 m – 1.0 m</td>
</tr>
<tr>
<td>Taper</td>
<td>$\tau$</td>
<td>$\leq 5^\circ$</td>
</tr>
<tr>
<td>Lean</td>
<td>$\omega$</td>
<td>$\leq 25^\circ$</td>
</tr>
<tr>
<td>Coverage</td>
<td>$c$</td>
<td>$\geq 50%$</td>
</tr>
<tr>
<td>Voxel size</td>
<td>$s$</td>
<td>2 m</td>
</tr>
</tbody>
</table>

measured tree diameter due to the effects of stem taper.

3.4.5 Algorithm

The stem reconstruction algorithm is divided into several stages as shown in Figure 3.3. First, the point cloud is subdivided into disjoint voxels. Second, each voxel is assessed to determine the parameters of candidate tapered cylinder segments (i.e., portions of a tree stem) within that voxel. Third, the set of all segments are synthesized into complete tree stem models via outlier filtering and gap filling. Finally, forest parameters are extracted from the quantitative stem structure. Each of these stages is discussed in detail below.

Figure 3.3: Algorithm flowchart identifying four main stages. Corresponding figures are referenced in parentheses.

Stage one subdivides the point cloud into disjoint voxels. We assume that tree stems are made
up of a series of contiguous tapered cylinders; this step enables the fitting of such models on a segment-by-segment basis. Mathematically, let a voxel be

\[ V_i = \{ x_1, x_2, \ldots, x_n \} \in \mathcal{P}(x_i < x \leq x_{i+1}) \cap \\
(y_i < y \leq y_{i+1}) \cap (z_i < z \leq z_{i+1}), \tag{3.3} \]

as shown in Figure 3.4a. Voxels have a side length, \( s \), such that \( s = x_{i+1} - x_i \) for all \( i \). A tradeoff between model resolution and processing time guides the choice of an appropriate voxel size. For this study, \( s = 2 \) m.

Stage two parameterizes candidate tree stem segments within each voxel. The majority of previous methods determine the diameter, and optionally direction, of a stem segment based on the curvature of points incident on the stem’s surface, e.g., (e.g., by cylinder fitting, circle fitting, or the Hough transform, see Table 1 in Pueschel et al. (2013)). Low-resolution instruments, however, may be unable to resolve the stem curvature with sufficient precision for stem diameter measurement, especially for trees with a diameter on the order of the laser beam size (See Figure 3.5c). To overcome this challenge we take advantage of the known sensor-object geometry and reduce the dimensionality of the fitting problem. Instead of fitting 3-D tapered cylinders or 2-D circles, we fit 1-D lines to the edge points of tree stems as viewed from the sensor. This reduces the degrees of freedom of the modeling problem while improving robustness to noise. A description of the approach follows in detail. For each voxel \( V_i \), the following steps are performed in the stem segment modeling stage.
Figure 3.4: Stem modeling is performed by iteratively dividing the point cloud into disjoint voxels (a) and fitting a tapered cylinder to the point samples (b-f). For each voxel, the estimated lean angle and direction is determined (b), from which the points are projected onto a new basis via Gram-Schmidt Orthogonalization (c). This allows us to examine 2-D projections of points, e.g., as from “top-down,” “face-on,” and “side” views. Samples of the background, B are removed using density estimation (d-left) and tree stem segments are isolated using alpha shapes (d-right). Points on the front, left, and right edges as viewed by the sensor are then isolated using alpha shapes (e). Finally, RANSAC is used to fit lines to these point sets (f), from which parameters of a tapered cylinder may be directly computed.
3.4. METHODS

1. Determine the estimated lean angle and direction \( \hat{\omega} \) by finding the vector, which maximizes the overlap of points along the stem surface (see Figure 3.4b).

   (a) Assume the true lean angle \( \omega \) can vary from the +z-axis according to \( 0 \leq \omega \leq \omega_{\text{max}} \). In this study, \( \omega_{\text{max}} = 25^\circ \), see Table 3.3 for list of parameters.

   (b) Systematically adjust \( \hat{\omega} \) within \( 0 \leq \hat{\omega} \leq \omega_{\text{max}} \).

   (c) For each \( \hat{\omega} \), compute the discrete point density after projection onto the subspace, \( S_{\perp} \), orthogonal to the vector described by \( \hat{\omega} \). The bin size is equal to the system sampling distance, \( \Delta_s(r) \).

   (d) Choose \( \hat{\omega} \) as the vector giving the maximum discrete point density.

2. Reproject the point samples in \( \mathcal{V}_i \) onto a new basis set using Gram-Schmidt orthogonalization, in which 2-D projections correspond to “top-down/face-on/side” views (Figure 3.4c).

   (a) Let \( \{u, v, w\} \) be a new orthonormal basis set such that \( w \) is the unit vector in the direction of the lean, \( \hat{\omega} \), \( v \) is the unit vector from the sensor to the object, and \( u \) is the unit vector mutually orthogonal to \( (w, v) \).

   (b) Then, \( \mathcal{V}_{uv} \) is the projection of \( \mathcal{V}_i \) onto the \( uv \) subspace (i.e., top-down view), \( \mathcal{V}_{uw} \) is the projection of \( \mathcal{V}_i \) onto the \( uw \) subspace (i.e., face-on view), and \( \mathcal{V}_{vw} \) is the projection of \( \mathcal{V}_i \) onto the \( vw \) subspace (i.e., side view).

3. Remove points due to branches, digitization noise, and other samples of the background, \( \mathcal{B} \).

   To do this, we assume that the stem direction is aligned with \( w \) and remove points, which are sparsely sampled along this vertical extent (Figure 3.4d-left).

   (a) Compute the discrete 2-D histogram \( N_{\text{hist}} \) of \( \mathcal{V}_{uw} \) with bin size equal to the system sampling distance, \( \Delta_s(r) \).

   (b) Calculate the expected maximum number of points \( N_{\text{exp}} \) for each bin based on \( \Delta_s(r) \) and the voxel side length, \( s \). (i.e., the number of points in each bin if a vertical surface is fully sampled).
(c) Compute the binary map, $I$, which satisfies $\left( N_{\text{hist}}/N_{\text{exp}} \right) \geq c_{\min}$, and dilate by the sampling distance, $\Delta_s(r)$, to account for edge effects (purple-shaded cells in Figure 3.4d-left). The coverage threshold $c_{\min}$ permits inclusion of stem surfaces, which are not fully sampled along the vertical extent due to occlusion, etc. For this study, $c_{\min} = 0.5$ (see Table 3.3 for list of parameters). Points that satisfy this condition are shaded green in Figure 3.4d-left. Remaining points (red) are removed.

4. Isolate points belonging to the most prominent tree stem, as defined by point cardinality (Figure 3.4d-right), and pass to step 6.

   (a) Compute alpha shapes (Edelsbrunner and Mücke, 1994) in the $V_{uv}$ projection, with probe radius equal to the system sampling distance, $\Delta_s(r)$.

   (b) Points contained in the alpha shape with the largest cardinality are passed to step 6. The remaining points will be reexamined in subsequent iterations, see step 9.

5. Find front, left, and right edge points from the sensor point of view (shaded red in Figure 3.4e) in preparation for line-fitting.

   (a) Compute alpha shapes in the “face-on”, $V_{vw}$, and “side”, $V_{uw}$, projections with probe radius equal to the system sampling distance, $\Delta_s(r)$.

   (b) Edge ownership is determined based on the angle of the line segment between successive point pairs.

6. Estimate parameters of a 2-D line for each set of boundary points (Figure 3.4f) using RANSAC (Fischler and Bolles, 1981). The inlier threshold is equal to one-half the system sampling distance, $0.5 \cdot \Delta_s(r)$, and the number of iterations is set to 500.

   (a) Fit $v = m_1 \cdot w + b_1$ from $\{v_1, w_1\}$ corresponding to front edge of tree.

   (b) Fit $u = m_2 \cdot w + b_2$ from $\{u_2, w_2\}$ corresponding to left edge of tree.

   (c) Fit $u = m_3 \cdot w + b_3$ from $\{u_3, w_3\}$ corresponding to right edge of tree.
3.4. METHODS

7. Compute tapered cylinder parameters in \( \{u, v, w\} \) space (see Figures 3.2 and 3.4f).

\[
r_0 = \frac{|(m_2 \cdot w_{\text{min}} + b_2) - (m_3 \cdot w_{\text{min}} + b_3)|}{2} \quad (3.4)
\]
\[
r_1 = \frac{|(m_2 \cdot w_{\text{max}} + b_2) - (m_3 \cdot w_{\text{max}} + b_3)|}{2} \quad (3.5)
\]
\[
l_0(u) = \frac{|(m_2 \cdot w_{\text{min}} + b_2) + (m_3 \cdot w_{\text{min}} + b_3)|}{2} \quad (3.6)
\]
\[
l_0(v) = m_1 \cdot w_{\text{min}} + b_1 + r_0 \quad (3.7)
\]
\[
l_0(w) = w_{\text{min}} \quad (3.8)
\]
\[
l_1(u) = \frac{|(m_2 \cdot w_{\text{max}} + b_2) + (m_3 \cdot w_{\text{max}} + b_3)|}{2} \quad (3.9)
\]
\[
l_1(v) = m_1 \cdot w_{\text{max}} + b_1 + r_1 \quad (3.10)
\]
\[
l_1(w) = w_{\text{max}} \quad (3.11)
\]
\[
\tau = \arctan \left( \frac{r_0 - r_1}{w_{\text{max}} - w_{\text{min}}} \right) \quad (3.12)
\]
\[
\omega = \hat{\omega} \text{ from step 1} \quad (3.13)
\]

These parameters are then reprojected from \( \{u, v, w\} \) space back into \( \{x, y, z\} \) space.

8. Expand voxel to \( V_i^* \) to include all points of distance \( \leq r_{\text{max}} \) from the modeled tree axis and repeat steps 3-6 with \( V_i = V_i^* \) and \( \hat{\omega} = \omega \) to ensure that a voxel contains the full stem.

9. Repeat steps 3-8 until a stopping criterion is reached. This addresses additional stems that are also in the voxel. When a stopping criterion is reached, proceed to the next voxel and reinitiate the modeling process.

The result of stage two is a series of candidate tree stem segments. In stage three, individual segments are synthesized into single tree stem models using outlier filtering and gap filling techniques. Outlier segments are filtered and removed if their parameters (radius, taper, lean, coverage) exceed the user-specified parameters, as tabulated in Table 3.3. Remaining segments may exhibit gaps due to partial occlusion of the stem surface. These gaps are filled by interpolating between neighboring segments of the same tree, and extended until intersection with the DEM is reached.
Finally in stage four, structural forest parameters are extracted directly from the models. We extract tree-level parameters including stem location and DBH, and plot-level parameters including basal area and stem density. Average computation time is 5 minutes on an Apple Macbook Pro with a 2.6 GHz processor.

3.4.6 Validation

Assessment of the 3-D geometric models was performed by projecting the 3-D data onto a 2-D Andrieu image and evaluating the per-point classification accuracies using manually-digitized truth maps. Additionally, in order to assess the potential of forest inventory using single-scan, low-resolution point cloud data, we automatically extracted forest parameters from the 3-D models and compared them to measurements made using conventional techniques (e.g., tape measure). We evaluated the ability to automatically (i) locate tree stems, measure stem diameter at breast height, (iii) count stems, and (iv) estimate basal area, using $R^2$ and root mean square error (RMSE) metrics.

3.5 Results

3.5.1 Quantitative Stem Models

The output of our algorithm is a quantitative 3-D reconstruction of the sub-canopy stem structure. These 3-D reconstructions are facetized geometric surfaces, which capture the underlying stem structural information (Figure 3.5), and could be used to derive detailed parameters such as taper, sweep, lean, merchantable volume, etc.

However, comprehensive validation of detailed 3-D structural information is difficult without destructive sampling (Henning and Radtke, 2006b; Raumonen et al., 2013; Garber and Maguire, 2003; Clark et al., 2000; Strahler et al., 2008). Many previous TLS studies have thus resigned to presenting higher fidelity 3-D models without rigorous validation, e.g., (Henning and Radtke, 2006a; Zhao et al., 2011). In order to provide quantitative validation of the 3-D stem models, we projected the 3-D data onto a 2-D Andrieu projection and evaluated stem detection accuracies.
3.5. RESULTS

based on manually digitized truth maps. Figure 3.6 shows Andrieu images of four plots chosen to represent the diversity in forest type and occlusion. An Andrieu projection (Andrieu et al., 1994) is an equal-angle map projection, which presents each point in the point cloud as a pixel in the image (no interpolation). Rows of the image correspond to samples in the elevation angle provided by mirror rotation, and columns of the image correspond to samples in azimuth angle provided by the rotation stage. The brightness of each pixel is the histogram-equalized intensity of the returned pulse. Note that the distortion increases at the top of the images (the upper hemisphere is oversampled). Furthermore, the black areas at the bottom of the image correspond to shadows or areas that are obscured by the instrument rotation stage and cables. Each image is 1440 × 540 pixels at full resolution, corresponding to the 360° × 135° field of view sampled at 1/4° spacing.

Visual assessment shows the algorithm is robust across a range of plot types and quality. Figure 3.6a is a single-tiered, low-density (525 stems · ha⁻¹) deciduous woodland with limited occlusion due to sub-canopy. Faithful reconstruction of the tree stem profile (incorporating diameter, taper, sweep, lean) was achieved. Some tree stem diameters were incorrectly modeled due to the partial obstruction by another tree. Reduced precision is achieved at large elevation angles. Figure 3.6b is a young coniferous plot with higher density and moderate sub-canopy branching. Our approach achieved good recovery of tree stem structure, even when partially or fully occluded. Some stems are incorrectly modeled due to the interference and complexity of the sub-canopy forest structure. Figure 3.6c is a young deciduous plot with significant occlusion due to foliage and vegetation. Good recovery was achieved for those stems that are visible to the sensor. However, it is apparent that significant occlusion will make plot-level retrieval of stem densities and basal area difficult. Finally, Figure 3.6d is a successional shrubland with 225 stems · ha⁻¹ > 10 cm and 425 stems · ha⁻¹ ≤ 10 cm. The algorithm missed a stem in the right quadrant, but was robust against errors of commission.
3.5.2 Classification Accuracy

We evaluated the accuracy of the tree stem models by examining the per-point classification accuracy for tree stem detection. For each plot, truth ROI’s (regions of interest) were hand-digitized based on the Andrieu projection intensity images. A separate ROI was digitized for each visible tree stem. Stems were digitized only where visible and regions were not interpolated between occlusion gaps, nor extended until intersection with the DEM. Additionally, the position and DBH were computed for each tree stem ROI to provide truth measurements, which were compensated to only include stems, which were visible to the sensor. This is used for validation in a later section.

Figure 3.6 shows the classification maps for the plots exemplified in Figure 3.6. Green pixels denote correct detection. Red pixels denote incorrect detection, i.e., errors of commission. Blue pixels denote incorrect rejection, i.e., errors of omission, and black pixels denote correct rejections. Pixel-based classification of the Andrieu images has a natural connection to the point data itself: Each pixel represents a single laser pulse, therefore, pixel-based classification can also be interpreted as the percentage of laser pulses, which are correctly or incorrectly identified as tree stems. However, it is evident from Figure 3.6 that there are limitations to this approach. Truth maps are based on subjective criteria of the operator. Specifically, it is not a trivial decision to determine where a tree stem transitions from “visible” to “not visible” if there are branches or foliage, which partially occlude its surface. This is further compounded by the fact that truth digitization was limited to recording the visible stem surface (no interpolation or extension), whereas the derived models do include interpolation between gaps and extension from the base to the terrain model. As a result, some plots (e.g., Figure 3.6a) consistently missed pixels above the detected vertical stem extent, while others (e.g., Figure 3.6b) exhibit consistent errors of commission. In the former case, the manual stem digitization is too liberal, in the latter, it is too conservative. Furthermore, due to errors in the projection of 3-D stem models onto the 2-D Andrieu images, the lidar-derived ROI’s may not be perfect (e.g., notice the intersection of stems with the DEM). Nevertheless, a literature review suggests that detailed tree stem models are historically difficult to quantitatively assess. This approach, though imperfect, is a viable metric for validation of parametric tree stem modeling.
3.5. RESULTS
despite the inflation of reported error due to subjective difficulties in manual digitization.

Figure 3.5: The output of our algorithm is a quantitative 3-D reconstruction of the sub-canopy stem structure and terrain (a). Stems are modeled as a series of contiguous tapered cylinders (see detail; (b)). Each segment is a facetized polygonal mesh providing precise and measurable structural information, e.g., volume. Facets are colored for visualization. Stems intersect the terrain and extend vertically upward until the underlying stem structure is insufficiently sampled. Reconstruction is achieved for stems, which subtended at least an angular width of 15 mrad—the beam divergence of our system (c). Note that the spatial beam width is oversampled such that 3.44 points correspond to one beam width.

Classification accuracies are provided in Table 3.4 on a per-plot basis. Errors of commission were consistently below 5%. Correct detection rates were much more variable. Approximately 80% correct detection rates were achieved for plots 13,15,16, and 31. These corresponded to plots with large basal area, yet few stems per hectare. Poor correct detection rates (20%) were achieved for plots 6 and 17. Plot 6 is a young mixed forest with 2475 stems · ha$^{-1}$, 1850 of which were of DBH < 10 cm. Stem diameters are small, with a plot basal area of only 18.43 m$^2$ · ha$^{-1}$. Plot 17 is an upland successional scrubland, shown in Figure 3.6d. Error is attributed to an overly liberal digitization of visible tree stem structure (e.g., within the bush in the left quadrant). Across
3.5. RESULTS

(a) Plot A3  
(b) Plot A10  
(c) Plot A13  
(d) Plot A17

Figure 3.6: Classification maps showing correct detections (green), errors of omission (blue), errors of commission (red), and correct rejections (black) based on manually-digitized truth data for the sample plots in Figure 3.6. Note the incorrectly classified errors of commission and omission. This underlies the difficulty of objective truth digitization.

all sites, percentage-based errors of omission were inflated due to the relatively few number of pixels image-wide, which are truly classified as stems (i.e., a few pixels error contributes to a large percentage-wise error). Likewise, errors of commission were arguably deflated because the majority of pixels in the image are not stems, and thus many erroneous pixels contributes to a small percentage-wise error. Moderate per-pixel classification rates (60% correct detection) were achieved for plots 5, 8, 10, and 19. These plots are dominated by coniferous trees whose sub-canopy tree structure increase the complexity of modeling. Plot 12 (BA = 27.91 m²·ha⁻¹ with 800 stems ≥ 10 cm) achieved just a 40% correct detection rate; this is due to a single large tree at close range, which was missed because of partial occlusion.

Table 3.4: Per-plot classification accuracies.

<table>
<thead>
<tr>
<th>Plot #</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>13</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>19</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corr. Detect [%]</td>
<td>56</td>
<td>8</td>
<td>67</td>
<td>60</td>
<td>41</td>
<td>83</td>
<td>79</td>
<td>76</td>
<td>19</td>
<td>58</td>
<td>80</td>
</tr>
<tr>
<td>Corr. Reject [%]</td>
<td>100</td>
<td>99</td>
<td>98</td>
<td>97</td>
<td>98</td>
<td>98</td>
<td>98</td>
<td>97</td>
<td>100</td>
<td>98</td>
<td>98</td>
</tr>
<tr>
<td>Incorr. Detect [%]</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Incorr. Reject [%]</td>
<td>43</td>
<td>92</td>
<td>32</td>
<td>40</td>
<td>59</td>
<td>17</td>
<td>19</td>
<td>22</td>
<td>81</td>
<td>42</td>
<td>19</td>
</tr>
</tbody>
</table>
Classification accuracies were aggregated over all plots and plotted as a function of xy-range in Figure 3.7a. Poor detection (30%) was achieved between 0 – 2 m, after which detection rates peaked between 2 – 6 m at 80% and then gently decreased to 75% at 14 m. There was a slight increase in errors of commission and omission as range increased. Likewise, classification accuracies were aggregated over all plots and plotted as a function of elevation angle in Figure 3.7b. Correct detection rates peaked near an elevation angle of 0°. As elevation angles deviated from 0° up to ±30°, detection rates deteriorated. Past this point, excessive noise due to insufficient sample sizes precluded meaningful analysis.

### 3.5.3 Forest Inventory Parameters

Further analysis evaluated the capability to automatically retrieve typical forest inventory parameters, individual tree-level stem location and DBH, and plot-level stem density and basal area. The positional (range) accuracy of stem detection was evaluated by comparing the range measurements to corresponding tree centers (Figure 3.8). Stem DBH was evaluated against DBH measured in-field with good correlation ($R^2 = 0.80$, RMSE = 0.06 m, see Figure 3.9). DBH retrieval was evaluated only for visible stems: stems that are partially or fully occluded were flagged and excluded.

We also evaluated the ability to recover stem densities at the plot-level using a single scan placed at the center of the 20 × 20 m plot. This is an inherently difficult problem, due to the occlusion of stems by other object structure. Therefore, we performed two analyses. The first, shown in Figure 3.10a, assessed the ability to estimate the number of visible stems (as digitized from the intensity images). The second, shown in Figure 3.10b, assessed the ability to estimate the number of all stems, regardless of whether they were visible to the sensor or not (based on conventional inventory). These measurements were then modulated by the plot area to estimate plot-level stem density. Good retrieval of visible stem density was achieved ($R^2 = 0.86$, RMSE = 79.38 stems·ha$^{-1}$); however, retrieval of stem density at the plot-level (including both visible and occluded trees) was less successful ($R^2 = 0.19$, RMSE = 188.08 stems·ha$^{-1}$), as was expected.

We also evaluated the ability to estimate 20×20 m plot-level basal area, which can be directly
Figure 3.7: Classification accuracies as a function of (a) range and (b) elevation angle. (a) Poor detection was achieved between 0-2 m. This was attributed to partial occlusion of the lower stem by the instrument rotation stage, increased sensor noise, and the error associated in our model due to high within-voxel variation in point density. (b) Peak detection rates were achieved at an elevation angle of 0°, after which classification accuracy decreased due to the reduced effective resolution of the beam relative to the stem diameter, and the increasing presence of ground vegetation or canopy structure at other angles cf. (Henning and Radtke, 2006a).
3.5. RESULTS

Figure 3.8: Tree stem locations were automatically detected with high positional (range) accuracy when compared to inventory measurements made using a tape measure. We achieved an $R^2 = 0.996$ and RMSE = 0.17 m.

Figure 3.9: Stem DBH were automatically detected and measured with high accuracy when compared to inventory measurements made using a tape measure. We achieved an $R^2 = 0.80$ and a RMSE, RMSE = 0.06 m.
Figure 3.10: (a) Stem density was recovered with $R^2 = 0.86$, RMSE = 79.38 stems·ha$^{-1}$ for all visible stems. (b) Retrieval of stem density compared to field inventory measurements (including all trees, visible or not) was less successful ($R^2 = 0.19$, RMSE = 188.08 stems·ha$^{-1}$), as was expected.
3.6. DISCUSSION

computed from the reconstructed stem models by summing all detected tree diameters:

\[ BA = \sum_{plot} \pi \left( \frac{DBH_i}{2} \right)^2 \cdot \frac{1}{A_{plot}} \]  

(3.14)

Results were similar to plot-level stem density estimation, with poor retrieval of the plot-level basal area due to occlusion \( (R^2 = 0.21, \text{RMSE} = 16.23 \text{ m}^2 \cdot \text{ha}^{-1}) \), but good retrieval \( (R^2 = 0.82, \text{RMSE} = 7.43 \text{ m}^2 \cdot \text{ha}^{-1}) \) when adjusted to include only stems that were visible to the sensor.

3.6 Discussion

Laser scanning has potential to collect a significantly larger range of forest structural attributes than traditional inventory techniques. However, the reliance of existing algorithms on high-resolution point cloud data are considered a major constraint for operationalization in forestry practice. Commercial TLS systems, frequently used in literature, have a market that is predominantly outside forestry, and includes fields in engineering, architecture, archeology and forensics, where negligence of registering fine details comes at high cost (Lovell et al., 2011; Bi and Wang, 2010). Traditional forest inventories, however, aim to characterize plot or stand level structural attributes for statistical interpretation, and opens opportunity for lower-resolution, lower-cost laser scanners. We developed a robust technique for reconstruction of tree stem models and subsequent retrieval of inventory parameters from low-resolution, single-scan laser scanner data collected by a low-cost TLS instrument designed to accommodate this niche application.

3.6.1 Quantitative Stem Models

The effects of system parameters on mensuration were complicated by their interdependence on stand and plot characteristics (Pueschel et al., 2013) as seen in Figure 3.6. For example, tree species composition, age, and density affect the degree of occlusion due to branch, foliage, and ground vegetation structure, which is a primary challenge for structural assessment using TLS. In some regions of the upper canopy, there are large voids in the measured point cloud, hence the stem models, correspondingly, are incomplete in these upper regions and we were unable to
accurately estimate tree height and stem volume. At lower heights, effects of occlusion play a role too and the ability to reconstruct stem models depends on the visibility of stem “edges”. For example, branches or foliage traversing laterally across the stem surface (see leftmost dominant tree in Figure 3.6b) often have a limited effect on the accuracy of the derived stem models (due to interpolation), while partial obscuration of two or more stems is likely to result in underestimates of stem diameters (see the right quadrant of Figure 3.6a where only the left-most edges of distant co-azimuthal stems are visible). These cases often resulted in exaggerated taper, and sometimes in rejections when the derived taper exceeded preset assumptions.

3.6.2 Classification Accuracy

We assessed per-point classification accuracy in order to supplement available inventory measurements and to validate the algorithm along the stem height and range sequence. Errors in point classification present a worse-case assessment and are exacerbated as a result of digitization errors, and ambiguity in cases where multiple objects overlap leading to a single return.

Highest classification accuracies were achieved for plots with large basal area and correspondingly a limited sub-canopy, owing to the lower noise levels in the data as well as improved accuracy in the manual digitization process. Analysis of classification accuracy by range revealed low detection rates for trees within 0-2 m. This was unexpected and primarily observed for plots 6, 8, 10, 12, and 19 and is likely caused by the variable point density at near range that has severe implications for both the lean angle estimation (which in turn affects the orthogonalization and projection onto orthogonal subspaces), noise removal, and tapered cylinder parameterization using alpha shapes. Future work will investigate a hierarchical approach to improve consideration of local point density variations while maintaining computational efficiency.

Evaluating classification accuracy as a function of elevation angle revealed that detection rates peaked for small elevation angles but decreased thereafter until about ±30°, where noise precluded meaningful analysis. This trend was due in part to the difficulty of subjective ROI digitization in regions where upper canopy (large elevation angles) or ground vegetation (low elevation angles) partially occluded the stem. In addition, measurement precision at higher elevation angles is
impacted by a reduced point density due to the range-dependent resolution, projected area effects on the predominantly vertical tree stems, and the increased occlusion at successive heights above ground (Henning and Radtke, 2006a).  

### 3.6.3 Forest Inventory Parameters

We also assessed tree-level and plot-level forest structure variables. Stem location (range) was evaluated with a high $R^2$ of 0.996, but a relatively large RMSE of 0.17 m. This error exceeds sensor specifications of ±3 cm bias and ±1.2 cm error (1σ) (SICK, 2009), suggesting inaccuracies in measurement using the tape measure and compass, due to aspects such as path obstructions, terrain effects, and read out (Henning and Radtke, 2006b). Stem diameter was evaluated with an $R^2 = 0.80$ and an RMSE of 0.06 m. Error sources in diameter retrieval included the potential mis-registration between lidar-derived stems and their field-measured counterparts, the partial shadowing between tree stems, and errors in estimating terrain elevation, which impacted DBH retrieval. A final source of error is due to the coarse beam divergence and corresponding probability of sampling stem edges when the beam center is outside the stem surface area, and this resulted in a 4 cm positive bias. Absolute diameter errors appeared to increase after 30 cm DBH. This may be due to larger stems having less regular shapes that are poorly estimated with a geometric fit.

While successful reconstruction of visible stem structure was achieved, occlusion significantly decreased the ability to measure plot-level attributes via single scans, cf. (Watt and Donoghue, 2005; Pueschel, 2013). Good results were obtained for plot-level basal area ($R^2 = 0.82$, RMSE = 7.43 m$^2$·ha$^{-1}$) and stem density ($R^2 = 0.86$, RMSE = 79.38 stems·ha$^{-1}$) when truth measurements were adjusted to include only stems that were visible to the sensor. However, poor retrieval of plot-level basal area ($R^2 = 0.21$, RMSE = 16.23 m$^2$·ha$^{-1}$) and stem density ($R^2 = 0.19$, RMSE = 188.08 stems·ha$^{-1}$) was achieved when all stems were included. Occlusion is the principal challenge in ground-based laser scanning acquisition as evidenced in this and other studies. Several techniques have sought to address this “inherent limitation” (Pueschel et al., 2013), making use of either statistics or multiple scans. Statistical approaches adjust plot-level estimates based on assumptions of occlusion due to tree stems, e.g., (Lovell et al., 2011; Ducey and Astrup, 2013),
but are not appropriate for forests with dense sub-canopy structure, such as those used in this study. To the contrary, the good performance of tree-level reconstruction shown here suggest that co-registration, not improved scanner resolution, is sufficient to address this limitation. Existing multiple-scan registration approaches, e.g., (Henning and Radtke, 2008; Eysn et al., 2013), may be effective in combining information at either the data or feature level, while retaining the efficiency of single-scan data acquisition (Pueschel, 2013). Using this mobile TLS system we were able to scan a $20 \times 20$ m plot with 25 scans in 30 minutes.

Results were assessed against system resolution parameters (sampling step-width and beam divergence) of comparable studies in the literature. We found that DBH retrieval was achieved commensurate to the sampling step-width. A review of DBH retrieval accuracies by Pueschel et al. (2013) reported RMSE's between 2-8 cm for instruments operating between 0.5-4 mrad, respectively. This study used a sampling step-width of 4.36 mrad and achieved a comparable RMSE of 0.06 m. In the context of beam divergence, however, stem diameter extraction was possible at much larger beam sizes than previously published limits. Pueschel (2013) assessed the influence of scan parameters on tree metric extraction using the popular Hough transform, and found that beam divergence and range together enforced a limit (i.e., in terms of beam size) in the ability to retrieve stem diameter via geometric circle fitting. They found that the maximum beam size from which DBH could be successfully extracted was 4 cm. To the contrary, we achieved successful stem measurement with a beam size of up to 25 cm, or more succinctly, for tree stems, which subtended at least an angular width of 15 mrad—the beam divergence of our system. Robust measurement was achieved despite data limitations by employing a unique approach to modeling using intuitive face-on/side-view projections, which overcame many of the challenges associated with conventional methods. We concluded that lower-resolution sensors, such as the one used in this study, may be effective in providing data for forest inventories constrained by sample size, time, and cost. Because acquisition time is proportional to the square of the sampling resolution, the feasibility of lower-resolution systems has implications towards improving the efficiency of data acquisition and thus the collection of adequate sample sizes for statistical analysis. Moreover, because lower-resolution sensors may be less expensive, these algorithms make possible the use
3.7. CONCLUSIONS

and integration of lower-cost, off-the-shelf laser scanners, such as the one used in this study. This directly addresses the needs of practical forestry studies, which require the collection of a large number of samples in order to account for spatial variance, while concurrently minimizing time and cost (Lovell et al., 2011).

3.7 Conclusions

This study assessed the performance of a novel tree stem reconstruction algorithm developed to enable the use of low-resolution, single-scan terrestrial laser scanner data. Instead of utilizing the curvature of the tree stem, e.g., by circle or cylinder fitting, we take advantage of the sensor-object geometry and reduce the dimensionality of the modeling to a series of 1-D line fits. This allowed robust recovery of tree stem structure in a range of New England forest types, for tree stems, which subtended at least an angular width of 15 mrad—the beam divergence of our system. From these facetized geometric models, standard forest inventory parameters were extracted and compared to measurements made using conventional techniques. Unbiased retrieval of tree location and diameter at breast height (DBH) was achieved within a 20×20 m plot for stems of diameter ≥10 cm. Plot-level estimates of stem density and basal area were limited by occlusion owing to a single scan location; future work will use existing registration approaches to overcome this inherent limitation. These results demonstrated the feasibility of lower-resolution sensors in providing data for operational forest inventories constrained by sample size, time, and cost. Such developments may facilitate an improved ability to study the diverse relationships occupying the structure-function interface (e.g., physiology, animal habitat, carbon sequestration, forest fire risks, and timber value). Additionally, since TLS provides a complementary and commensurate data set to ALS, improved sub-canopy structural assessment may further advance the exploitation of ALS data in forested environments.
Chapter 4

Pairwise marker-free registration of TLS data

4.1 Foreword

Results and conclusions from Chapter 3 identified occlusion - not resolution or algorithm limitations - as the primary challenge to accurate plot-level forest inventory using a low-cost terrestrial laser scanning (TLS). This application confirms the need for point cloud registration; however, from a review of the literature (chapter 2), the need for registration of TLS data is much more pervasive, with application to robotics/mobile perception (Forsman and Halme, 2005), mapping, and the majority of techniques for canopy structure assessment. In particular, we identified a critical knowledge gap in the ability to perform automatic, blind, marker-free registration of TLS data pairs in forest environments. Such an approach is necessary in order to support operational, efficient, data acquisition.

Therefore, we identified a second objective as follows:

2. Determine the error associated with automatic, blind, marker-free registration of TLS data pairs in forest environments.
Chapter 4 describes in detail the study, methodology, and results related to objective 2 and its corresponding sub-objectives. Outputs include an original, robust methodology to register TLS data using view-invariant tie points derived from the stem-terrain intersection points, i.e., the geometric primitives derived in Chapter 3. Moreover, geometric properties are exploited to constrain the search space and improve computational efficiency, enabling automatic registration without knowledge of initial sensor pose. We also present an innovative approach for providing an embedded error metric, by exploiting circular self-closure through disjoint tie point sets. This research fills a vital knowledge gap in the efficient, pairwise registration of forest point cloud data, and has implications for informing optimal sample spacing for TLS data collection, improving the plot-level assessment of inventory (Chapter 3) and supporting the assessment of canopy structure.

This chapter was submitted for peer-review as a manuscript entitled, *Marker-free registration of forest terrestrial laser scanner data pairs with embedded confidence metrics*. As such, certain sections may be repetitive with the background section (chapter 2).

### 4.2 Abstract

TLS has emerged as an effective tool for rapid, comprehensive measurement of object structure. Registration of TLS data is an important prerequisite to overcome the limitations of occlusion. However, due to the high dissimilarity of point cloud data collected from disparate viewpoints in the forest environment, adequate marker-free registration approaches have not been developed. The majority of studies instead rely on the utilization of artificial tie points (e.g., reflective tooling balls) placed within a scene to aid in coordinate transformation. We present a technique for generating view-invariant feature descriptors that are intrinsic to the point cloud data, and thus enable blind, marker-free registration in forest environments. To overcome the limitation of initial pose estimation, we employ a voting method to blindly determine the optimal pairwise transformation parameters, without an a priori estimate of the initial sensor pose. To provide embedded error metrics, we developed a set theory framework in which a circular transformation path is traversed between disjoint tie point subsets. This provides an upper estimate of the root mean square error (RMSE) confidence associated with each pairwise transformation. Output
RMSE errors are commensurate with the RMSE of input tie points locations. Thus, while the mean output RMSE=16.3 cm, improved results could be achieved with a more precise laser scanning system. This study (i) quantifies the RMSE of the proposed marker-free registration approach, (ii) assesses the validity of embedded confidence metrics using receiver operator characteristic (ROC) curves, and (iii) informs optimal sample spacing considerations for TLS data collection in New England forests. While the implications for rapid, accurate, and precise forest inventory are obvious, the conceptual framework outlined here could potentially be extended to built environments.

4.3 Introduction

Spatial registration is the process of aligning data into a common coordinate system. Combining data from multiple, co-registered laser scans is a common pre-processing step in order to avoid the critical limitations of data obscuration due to laser occlusion. This has applications to robotics/mobile perception (Forsman and Halme, 2005), mapping (Nüchter, 2009), and other domains, but is especially important for forest inventory studies, which estimate forest structural variables such as basal area (BA) or stem density over a broad sample area. Though single-scan terrestrial laser scanning (TLS) acquisition is effective at characterizing visible tree structure, multiple scan information is often necessary in order to assess plot-level variables (Kelbe et al., 2015b). We can define two types of registration, (i) relative, i.e., the combination of data from multiple scanner positions into a single scanner’s coordinate system, and (ii) absolute, i.e., the referencing of data to an absolute global coordinate system (Bi and Wang, 2010). This paper describes a technique for relative registration. Rregistration is performed by estimating the three translation and three rotation parameters between two coordinate systems and then modifying the data’s spatial coordinates according to these parameters (Hilker et al., 2012b).

Direct registration approaches manually measure the position and orientation parameters of each TLS system. A differential global positioning system (GPS) or total station is used to survey the precise location of each scanner location, and an inclinometer or inertial measurement unit (IMU) is used to measure the instrument’s orientation. Note that some TLS instruments
have a motorized head, which automatically levels the z-axis, thus requiring only azimuthal correction (Hilker et al., 2012b). In a study by Van der Zande et al. (2006), a SICK sensor, similar to the one in this study, was stepped laterally along a translation stage at several known positions adjacent to an artificial tree. Data registration was then performed based on the precise knowledge of the scanner pose at each measurement location. Despite the controlled set-up, registration proved to be the most difficult obstacle, requiring labor-intensive and time-consuming manual corrections. Moreover these techniques require high-precision surveying equipment that may not be operationally tenable to foresters (Hilker et al., 2012b).

An alternative technique is to manually align the data after collection, based on visual inspection in computer software. Yang et al. (2013) adjusted the rotation and translation matrices of multiple scans in Pointools View Pro software to align features such as trunk shapes, terrain patterns, and crown characteristics. An estimated positional accuracy of \( \pm 20 \) cm was achieved. However, accuracy is dependent upon the subjective clarity of these features to the interpreter, and the process is labor-intensive and time-consuming; as a result it has not been advised for future implementation (Van der Zande et al., 2006).

A third class of registration algorithms involve manually placing artificial targets (or “markers”) in the scene, which serve as precise and unambiguous tie points (Yang et al., 2013; Van der Zande et al., 2006; Hilker et al., 2012b; Cifuentes et al., 2014). Commonly, retroreflective spheres mounted on poles are used (Hilker et al., 2012b), although reflective tape (Henning and Radtke, 2006a), plain A4 paper (Ascho et al., 2004), and checkerboards (FARO, 2012) have also been used. Targets may then be detected either manually (Hilker et al., 2012b), or automatically with commercial software packages, e.g., (FARO, 2012; RIEGL, 2005; Leica, 2014). Given that the majority of available systems are commercial scanners with corresponding software packages (Bi and Wang, 2010), these methods have been widely used in forest inventory studies. For example, Zheng and Moskal (2012) registered high-resolution (0.16 mrad sampling) TLS data using Leica Cyclone software with eight in-field reference targets, achieving a mean absolute error of 3.4 cm. Another study used FARO SCENE software (FARO, 2012) to register TLS data from nine positions in a 20 × 20 m plot (Cifuentes et al., 2014), but the reported accuracy “tension” metric precluded direct
Pueschel (2013) used a set of FARO targets (four spheres and one planar target) and performed manual registration with very low registration errors ($\mu = 0.3$ mm, $\sigma = 1-2$ mm) for high resolution data (angular step width = 0.6 mrad).

Despite the capability of achieving high accuracy, the practicality and scope of these methods are limited by several key restrictions. First, automatic detection of targets is not trivial, and may require additional user interaction for identification of undetected markers (Pueschel, 2013; FARO, 2012). Second, target size and instrument resolution enforce a limit on the maximum distance between the sensor and target, in order for successful detection (Pueschel, 2013; FARO, 2012; Henning and Radtke, 2006b; Aschoff et al., 2004). Finally, and perhaps most relevant, is the issue of cost: Commercial registration packages may be outside the budget constraints of many end users (Hilker et al., 2012b). This budgetary restriction is increasingly salient as a growing number of affordable, small sensor providers (Hilker et al., 2012b) lack registration expertise (Bi and Wang, 2010) and provide no such software with their systems, cf. (SICK, 2009).

In support of these commercial off-the-shelf (COTS) systems, Hilker et al. (2012b) developed a simple approach for registration, based on the use of artificial targets, which avoided the need for commercial software packages (Hilker et al., 2012b). The proposed method used 0.2 m reflective polystyrene spheres mounted on wooden pegs; their location was first identified in the image based on manual field-measurement of the location and bearing, and then tie points were extracted for each sphere by averaging the positional coordinates of high intensity returns within that identified region. Note that averaging the location of points sampled on a 0.2 m hemisphere will give an ambiguous center coordinate (tie point), depending on sensor perspective, which may have contributed to the higher reported correspondence errors (root mean square error (RMSE) = 0.04-0.7 m and coefficient of determination ($R^2$) = 0.70-0.99 between tie points).

Disadvantages of this approach are its reliance on accurate reflectance information for automatic thresholding of points on the sphere, and limitations of practicality related to the use of manual targets. Transportation and placement of supporting bases and reflective targets is cumbersome and tedious (Calders et al., 2014; Bienert et al., 2006; Pueschel et al., 2013; Cifuentes et al., 2014), and requires additional personnel, equipment, and time (Calders et al., 2014). Furthermore,
it may be difficult to place targets in positions that can be seen from multiple viewpoints (Henning and Radtke, 2006b). As a result, some studies have raised concerns of the practicality of using artificial targets in forested environments (Aschoff et al., 2004). This is because forest mensuration necessitates maximizing the sample size and performance ability, while minimizing the required time, personnel, and cost (Lovell et al., 2011; Bi and Wang, 2010; Pueschel et al., 2013). The use of artificial targets severely reduces mobility and efficiency, and represents a major limitation to the utility of TLS for operational objectives, especially in terms of the number of samples that can be measured, cf. Tables 4 and 5 from (Pueschel et al., 2013). In this context, marker-free registration techniques are needed to address these current limitations (Pueschel, 2013).

The most common technique for marker-free data registration is based on the iterative closest point (ICP) algorithm, which minimizes the Euclidean distance between points in regions of overlap, based on an initial estimation of sensor pose (Bi and Wang, 2010). First introduced by Chen and Medioni (1991) and Besl and McKay (1992), there has been continued work to improve this method for registration, which suffers from drawbacks, such as convergence to local minima and the requirement of estimating an initial sensor pose (which is not easy, and prevents full automation) (Bi and Wang, 2010). Improvements to this framework have been developed, for example, by Gelfand et al. (2005) who computed descriptors for each point based on local geometry, in order to eliminate the need for initial pose estimation.

However, despite the prevalence of these methods in computer vision and robotics, existing marker-free registration techniques have been primarily developed for industrial and engineering applications, and are ill-suited to the forest environment (Henning and Radtke, 2006b; Bi and Wang, 2010). This is due to the complex, irregular shape of natural elements (Cifuentes et al., 2014), and the sparsity of forest structure, such that small shifts in perspective record vastly different sampled surfaces. Nevertheless, due to availability, convenience, and a lack of alternatives, these techniques have been occasionally applied in forest environments, but with unsatisfactory results. Hopkinson et al. (2004) used the IMAlign module of the Polyworks software suite (Innov-Metric Software, 2007), which computes the best fit transformation parameters using an iterative analysis of point cloud residuals in the region of scan overlap after initial coarse registration. As
this registration step was simply a precursor to stem modeling, quantitative assessment of the results were not given. And Calders et al. (2014) used the Multi Station Adjustment algorithm from RIEGL RiSCAN PRO software, which employs a similar ICP technique for combining TLS and airborne laser scanning (ALS) data. The authors found that marker-free registration techniques were inadequate in the forest environment, because of the difficulty of establishing point correspondence between data obtained from various sensor perspectives.

One marker-free technique was presented by Henning and Radtke (2006a) and summarized in Henning and Radtke (2006b), which overcame the issue of point correspondence by making some geometric assumptions on the tree objects. Coarse alignment was first performed based on either reflective tape affixed around tree stems at breast height (Henning and Radtke, 2006a) or measurement of the scanner location and orientation (Henning and Radtke, 2006b). Then, to avoid the problem of different scanner positions sampling different sides of an object surface (e.g., of the tree stem), the authors assumed that tree boles were approximately circular in cross-section and extracted the centers of the tree stem at various heights above ground. Pairwise registration was then performed using these (perspective-independent) data with the ICP algorithm. These constraints reduced the error in the x-y direction; ground surfaces were also incorporated to reduce error in the z direction. Though this approach demonstrated the potential for utilizing point sets that were invariant to view or perspective differences, its scope was limited, aligning just three clearly visible trees within 5 m of four scanner positions. Moreover, an initial pose estimate was required, limiting the utility for rapid, operational inventory.

In light of this gap, there is a compelling need for developing a rapid, marker-free registration which is robust in the forest environment. Such an approach could maximize collection efficiency and mobility in-field, while still providing the benefits of multiple-scan registration for rapid operational forest structure assessment (Pueschel, 2013). Consequently, the objectives of this paper are to (i) quantify the RMSE error of a proposed marker-free registration approach, (ii) assess the validity of embedded confidence metrics using receiver operator characteristic (ROC) curves, and (iii) identify an optimal sample spacing for TLS data collection in New England forests.
4.4 Methods

4.4.1 Background

For a typical stationary TLS system, a pulsed laser beam is rapidly emitted into the scene in a fan pattern based on the deflection by a rotating mirror. This scanning in elevation angle is coupled to azimuthal platform rotation to sample nearly the full sphere, except for a small occlusion cone below the instrument. For each scan mirror angle, $\theta$, and rotation stage position, $\phi$, the return trip travel time of a laser pulse is digitized and converted to range, $r$, based on the speed of light. This gives an unambiguous triplet ($\theta$, $\phi$, $r$) for each digitized pulse. A point cloud, $\mathcal{P}$, is the aggregate of all digitized range measurements. Upon conversion from spherical to cartesian coordinates, we can define the point cloud $\mathcal{P} = \{x_1, x_2, \ldots, x_n\}$, where $x \in \mathbb{R}^3$ is the $x, y, z$ position for the $i^{th}$ point in $\mathcal{P}$.

The challenge with point-based registration, and the solution afforded by object-based registration, are illustrated in Figures 4.1 and 4.2, respectively. In Figure 4.1, two scans are taken from opposite sides of a tree stem. Sensor 1 ($\mathcal{S}_1$) samples the visible (left) side of that object surface, ($\mathcal{S}$), but cannot record any measurements of the right side. Sensor 2 ($\mathcal{S}_2$) samples the right side of the surface, but cannot record any measurements of the left side. Point-based registration methods, e.g., iterative closest point (ICP), which iteratively adjust parameters of a rigid transformation while minimizing the sum-squared error (SSE) between corresponding points of $\mathcal{P}_i$ and $\mathcal{P}_j$, maps the new sensor 1 location, $\hat{\mathcal{S}}_1$, to $\mathcal{S}_2$ and, likewise, map $\hat{\mathcal{P}}_1$ to incorrectly overlap $\mathcal{P}_2$. 
4.4. METHODS

Figure 4.1: Point-based registration techniques generally fail for TLS scans from forested environments. This is because (a) multiple scanners (e.g., \( S_1 \) and \( S_2 \)) sample different parts of \( S \) (e.g., tree stem). There is insufficient overlap of \( P_1 \) and \( P_2 \), therefore (b) \( \hat{P}_1 \) is mapped onto \( P_2 \).

The solution we present determines view-invariant tie-points based on the underlying objects themselves, and then employs a registration algorithm developed upon that framework. This is shown in Figure 4.2. Sensors \( S_1 \) and \( S_2 \) still sample opposite surfaces of \( S \). However, an estimate of the underlying tree stem surface is first modeled from each set of points, and then view-invariant tie points are extracted as the intersection point of the tree stem axis and the terrain, based on assumptions of radial symmetry. The result is that \( \hat{S}_1 \) is mapped to near its true location \( S_1 \), and likewise, the points \( \hat{P}_1 \) are mapped to near their true location \( P_1 \) (Figure 4.2b). The algorithm is expounded in full detail below.
4.4. METHODS

Figure 4.2: Object-based registration can overcome the challenges caused by limited point overlap. (a) Instead of registering the points, $P_1$ and $P_2$, tie points are extracted from the modeled surfaces, and these view-invariant tie points are used to perform the registration. (b) The result is that $\hat{S}_1$ is mapped to near its true location, $S_1$, and the points, $\hat{P}_1$, are mapped to near their true location $P_1$.

4.4.2 Study Area

To assess the feasibility of point cloud registration in the forest environment, we assessed the error registration for 11 plots spanning a diverse range of structural complexity. The study area for this work corresponds to the National Ecological Observatory Network (NEON)'s (Kampe et al., 2010) core ecological site for the Northeastern ecological domain, including both Harvard Forest and Quabbin Reserve, Massachusetts, USA (bounded by 42.428° N, 72.284° W and 42.558° N, 72.170° W; WGS1984). Harvard Forest is a 1,200 ha reserve with a long history of ecological research and management. Quabbin Reserve is a 23,000 ha public forest and provides additional diversity, e.g., disturbance regimes.
Eleven 20 × 20 m NEON plots were selected in this region, representing a diverse range of Northeastern USA forest structure. Plots were selected using a stratified random sampling scheme and evaluated in person to ensure forest structure variability. Individual plot-level characteristics are shown in Table 5.1. Plots include a range of young, mature, single-tiered, multi-tiered, sparse, dense, deciduous, coniferous, and mixed forest types. Basal area (BA) ranged from 3.21 m² · ha⁻¹ to 73.73 m² · ha⁻¹. Stem densities were recorded separately for stems of DBH ≥ 10 cm and for stems of DBH < 10 cm, and ranged from sparse and mature (700 stems · ha⁻¹, all of which were ≥ 10 cm DBH) to dense and young (2475 stems · ha⁻¹: 625 with DBH ≥ 10 cm and 1850 with DBH < 10 cm). Ground vegetation characteristics (mean height (\( \bar{z} \)) and percent cover (\( p \))) were also recorded. Plots exhibited a range of sub-canopy characteristics, from bare ground to full coverage of 1.1 m tall ground vegetation. Lidar images of a sample of the plots are shown in Figure 5.1. The structural variability represented by our study area is unique among previous research and provides a diverse data set from which to evaluate the ability of TLS for operational forest inventory.

Figure 4.3: Examples sites represent a diverse range of Northeastern USA forest structure, ground vegetation, and terrain characteristics (See Table 5.1).
For each plot, 25 scans were collected in a nominal grid pattern with 5 m spacing. Plots were first laid out based on a center GPS coordinate, compass, and tape measure. This regular sampling method was maintained to ensure consistent, objective data coverage. Knowledge of this initial pose information and the regular pattern of data collection, however, was not used a priori in the development of this algorithm. In other words, the pairwise registration technique developed in this paper is blind, or independent of initial scanner pose. We verify this independence with analyses outlined in subsection 4.4.7.

Our experimental design was to perform blind, marker-free registration using the proposed technique between all possible scan pairs in each of the $M = 25$ scans for each 20 x 20 m plot. Thus, there are $\binom{M-25}{2} = 300$ possible pairwise transformations between scans collected for each plot, and $\binom{M-25}{2} \times 11$ plots = 3300 pairwise transformations results collated in this paper.

Data were collected during August 2012 leaf-on conditions. Data collection during leaf-off conditions is preferable for measurement of woody structure, in that it reduces the effects of occlusion (Dassot et al., 2012; Raumonen et al., 2013; Côté et al., 2009). However, our objective was to evaluate TLS as an operational forest inventory tool. Since some parameters, e.g., leaf area index (LAI), are relevant only during the growing season, leaf-on data collection allowed us to better evaluate the feasibility of this objective.

The instrument used in this study was a low-cost mobile terrestrial lidar system integrated from commercial-off-the-shelf (COTS) components by Rochester Institute of Technology. The system is

### Table 4.1: Summary of ground validation plots in Harvard Forest, MA, USA.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Easting</th>
<th>Northing</th>
<th>Dominant Species (Secondary Species)</th>
<th>BA $\text{m}^2 \cdot \text{ha}^{-1}$</th>
<th>Stems $\text{ha}^{-1}$ $\geq 10$cm</th>
<th>Stems $\text{ha}^{-1}$ $&lt; 10$cm</th>
<th>$\bar{z}$ [m]</th>
<th>$p$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>05</td>
<td>-72.96080</td>
<td>47.089083</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>73.73</td>
<td>775</td>
<td>250</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>06</td>
<td>-73.11566</td>
<td>47.126714</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>18.43</td>
<td>625</td>
<td>1850</td>
<td>0.9</td>
<td>45</td>
</tr>
<tr>
<td>08</td>
<td>-72.81538</td>
<td>47.050078</td>
<td><em>Pinus strobus</em> - <em>Tsuga canadensis</em></td>
<td>55.70</td>
<td>600</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>-73.32665</td>
<td>47.057374</td>
<td><em>Pinus strobus</em></td>
<td>53.84</td>
<td>950</td>
<td>0</td>
<td>0.5</td>
<td>40</td>
</tr>
<tr>
<td>12</td>
<td>-73.07741</td>
<td>47.134911</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>37.91</td>
<td>800</td>
<td>150</td>
<td>0.6</td>
<td>30</td>
</tr>
<tr>
<td>13</td>
<td>-73.21947</td>
<td>47.131693</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>51.67</td>
<td>425</td>
<td>675</td>
<td>0.9</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>-72.85960</td>
<td>47.051639</td>
<td><em>Tsuga canadensis</em> - <em>Betula alleghaniensis</em></td>
<td>51.56</td>
<td>700</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>-73.28461</td>
<td>47.039323</td>
<td><em>Tsuga canadensis</em> - <em>Betula alleghaniensis</em></td>
<td>70.33</td>
<td>725</td>
<td>250</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>-72.64715</td>
<td>47.036012</td>
<td>Upland successional shrubland</td>
<td>3.21</td>
<td>225</td>
<td>425</td>
<td>1.1</td>
<td>100</td>
</tr>
<tr>
<td>19</td>
<td>-72.95442</td>
<td>47.100031</td>
<td>Variant: <em>Tsuga canadensis</em> with mixed</td>
<td>66.38</td>
<td>950</td>
<td>125</td>
<td>0.6</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>-72.71459</td>
<td>47.042133</td>
<td><em>Quercus rubra</em> - (Acer saccharum)</td>
<td>40.96</td>
<td>525</td>
<td>125</td>
<td>0.7</td>
<td>70</td>
</tr>
</tbody>
</table>
based on a design first implemented by a team at the Katholieke Universiteit Leuven, Belgium (Van der Zande et al., 2006). A SICK LMS-151 laser scanner (SICK, 2009) pulses a 905 nm laser at 27 kHz, with range measurement recorded based on time-of-flight. The emitted laser pulse is deflected by a rotating mirror and coupled to azimuthal platform rotation in order to sample full hemisphere above the instrument and a portion of the hemisphere below (270° V x 360° H coverage). The instrument is tethered to a data logger and battery, which are mounted on a backpack and worn by the operator. Sensor control is achieved via a wireless mobile application. This instrument was designed to overcome the limitations of high cost, low mobility, and long scan time, which have so far precluded operational forest structure assessment using terrestrial lidar. However unlike many commercial scanners that provide for high-density point cloud data, this system has a minimum angular step-width of 4.36 mrad, and a beam divergence of 15 mrad, both approximately two orders of magnitude coarser than comparable instrumentation (see Table 1 from (Kelbe et al., 2015b)). These limitations provide an opportunity to develop registration algorithms, which are robust to low-resolution data, while focusing on rapid, operational inventories. Our instrument’s mobile platform allowed a single operator to traverse between plots without disassembly. The total time to characterize a 20×20 m plot with 25 scans was approximately 30 minutes for a single operator. This includes 30 seconds per scan duration and an additional 15 seconds required to move between scan locations.

4.4.3 Algorithm Problem Statement

Consider two scans from sensors $S_i$ and $S_j$. Each scan records points in a local coordinate system (LCS$_i$, LCS$_j$), i.e., a coordinate system with its origin and axes defined by the pose of instrument. The task is to determine the pairwise rotation and translation parameters, which transform LCS$_j$ to LCS$_i$, without a priori knowledge of the initial scan configurations (position, orientation).

The algorithm flowchart is shown in Figure 5.2. The input data to this algorithm are the stem maps, $T_i$ and $T_j$, derived from point clouds $P_i$ and $P_j$. Stem maps are a common output for TLS algorithms in forest inventory, making this algorithm readily accessible to existing data and lidar systems. However, in the case where stem maps are unavailable from single scans, they may be
extracted using the techniques described in (Kelbe et al., 2015b). We defined the stem map, $\mathcal{T}_u$, as the collection of all tree diameters, $d$, and locations, $l$, for a scan $i$, with the added condition that in order to constrain the $z$ dimension, the tree locations, $l$, are $\in \mathbb{R}^3$. The algorithm proceeds in two stages, (i) parameter fitting and (ii) error estimation. In the parameter fitting stage (first row of Figure 5.2), pairwise registration (see flowchart Fig. 4.5 for details) is performed to determine the rigid transformation parameters $(R_i^j = \{\theta_x, \theta_y, \theta_z\}$, and $t_i^j = \{t_x, t_y, t_z\}$) which map scan $j$ into the coordinate system of scan $i$. The error estimation stage (bottom row of Figure 5.2) utilizes the same pairwise registration framework, but with disjoint tie point sets $A$ and $B$ in order to derive a pair of independent transformations, one from $i^B \rightarrow j^B$, and another from $j^A \rightarrow i^A$. These “forward” and “reverse” transformation parameters are then applied in sequence to determine the effective transformation parameters for a circular path from disjoint sets ($i^B \rightarrow j^B; j^A \rightarrow i^A$), i.e., back to the original coordinate system. This provides an upper estimate of the error ($e_{\theta_x}, e_{\theta_y}, \ldots e_{\theta_z}$) associated with each transformation parameter, as estimated from stage one. Moreover, we also compute an RMSE error between corresponding tie points for the circular path (since their truth location is known) in order to provide a more intuitive, readily-comparable, upper bound error metric associated with the primary registration stage. The outputs of the algorithm are the rotation and translation parameters, which transform scan $\mathcal{S}_j$ into $\mathcal{S}_i$, along with confidence metrics of the error for each transformation. Notes on the input data are discussed in section 4.4.4, while details of the primary stage (rigid transformation parameters) and secondary stage (confidence metrics) are examined in sections 4.4.5 and 4.4.6, respectively.
Figure 4.4: The algorithm proceeds in two parallel stages. The first (top row) determines the estimated rigid transformation parameters based on input stem maps (tie points) (see section 4.4.6). The second (bottom row) provides upper-bounds confidence metrics, associated with these estimated rigid transformation parameters, by construction of a circular path wherein the truth tie point locations are explicitly known (see section 4.4.5).

### 4.4.4 Algorithm Input

The input to this algorithm are two view-invariant tie-point sets, $\mathcal{T}_i$ and $\mathcal{T}_j$. The $i^{th}$ tie point set, $\mathcal{T}_i$, is the set of $N_i$ stem-terrain intersection points, $l$, and the $N_i$ diameter at breast height (DBH) values, $d$, where there are $N_i$ trees detected in the $i^{th}$ scan. More specifically, $\mathcal{T}_{i,j}$ is the $i^{th}$ tie point for scan $\mathcal{S}_i$.

$$
\mathcal{T}_i = \left\{ l_i = l_1 \ldots l_{N_i}, \in \mathbb{R}^3 \right\}
$$

Conceptually, this is analogous to the traditional “stem map,” making the algorithm readily accessible to existing data and lidar systems. Any range-measurement system (e.g., stereo photography, time-of-flight camera, high-resolution TLS, low-resolution TLS) capable of estimating the 3-D position of stem locations ($l \in \mathbb{R}^3$) and diameters can apply the algorithm developed here for coordinate transformation. In fact, we will show in subsection 4.4.7 that the error of the proposed approach is roughly equivalent to the error associated with the initial tie points. Therefore, while the results shown here are for a low-resolution, cost-effective TLS system, improved RMSE results would be achieved using a scanner with greater measurement precision.

Note that $l \in \mathbb{R}^3$, i.e., a $z$ value is required for each stem location in order to constrain the
transformation in 3-D. Where such data are not already available, \( T_i \) can be derived using the single-scan tree stem modeling approach outlined in (Kelbe et al., 2015b). In Kelbe et al. (2015b), we assumed that \( P \) is a finite set of samples of both the object surface, \( (S, \text{e.g., tree stems, terrain}) \) and background \( (B, \text{e.g., branches, foliage, noise}) \). The underlying object surface was then reconstructed, i.e., \( \hat{S} = F(P) \), by modeling tree stems as a series of contiguous conical frustrums and by modeling the digital elevation model (DEM) as a minimum-z Delaunay triangulation. \( l \) is computed as the intersection of the tree stem axis and the corresponding DEM facet. \( d \) is computed from trigonometry based on the conical frustum’s diameter and taper parameters for an elevation of 1.3 m above ground (diameter at breast height; DBH).

We only used tie points at ground level. Justifications for this decision are as follows: It is widely understood that a mathematical transformation performs best (i.e., minimizes error) at or near the locations of the tie points. Additional tie point features (derived from the extended set of cylinder parameters) could thus improve the precision of the transformation, especially in regions that are not well-represented by the reduced set of tie points. In other words, the full cylinder models extend well into the \( +z \) direction, while the extracted tie point set are confined primarily to the \( \{x, y, z \approx 0 \text{ m}\} \) subspace. A transformation derived from tie points confined to this subspace will have low point error within this subspace, but higher error in the upper canopy regions. Nevertheless, the prevailing techniques for registration of TLS data (using artificial targets placed between 0–2 m above ground) are also constrained by this limitation, and the technique has been deemed sufficiently precise for most applications. A further justification is accessibility: tree locations and DBH’s, i.e., stem maps, are a common output of TLS algorithms for forest inventory. In contrast, fully facetized tree stem models are not a common output, making a registration technique dependent on these inputs less accessible for the user. Finally, the complexity of this problem increases dramatically for large numbers of tie points. We hypothesized that marker-free registration, using a reduced parameter set, will yield sufficiently precise transformation parameters, and our results confirmed this hypothesis.
4.4.5 Pairwise Registration Parameter Fitting

The pairwise registration approach is best understood in the context of some preliminary background on (i) available tools and (ii) the remaining computational challenge. Two tools of interest are the rigid transformation model based on singular value decomposition (SVD) and the RAn-dom SAmple Consensus (RANSAC) paradigm. The rigid transformation model, based on SVD, deterministically solves for the 3-D rigid transformation parameters, which minimize the least-squared error between a pair of corresponding point sets with cardinality $\geq 3$ (Besl and McKay, 1992).

The crux of this approach is then to determine accurate correspondences between two point sets. One approach is to randomly sample many possible correspondence sets and determine the 3-D rigid transformation model, which provides the best fit, cf. the RANSAC paradigm. RANSAC is an iterative voting method, which estimates the parameters of a mathematical model from data containing outliers. In each iteration, a minimum sample set (MSS) of points are chosen from the data and used to instantiate a candidate model. The fit (e.g., number of inliers) is computed based on the candidate model and compared to the fit obtained from the current best model. The best model is updated if the fit is improved, with the result being that an estimate of the true model reliably can be obtained from data containing large amounts of noise.

Despite the potential of this method, the computational complexity, $O$, of a random search is large for the scenario posed in this paper, i.e., testing all possible $k$-permutations of $N_i$ (the number of trees detected in scan $i$) against all possible $k$-permutations of $N_j$.

$$O\left(\frac{N_i!}{(N_i - k)!} \cdot \frac{N_j!}{(N_j - k)!}\right)$$

(4.2)

To overcome this prohibitive limitation, we take advantage of readily available a priori knowledge to sort the list of correspondence sets by geometric similarity and then evaluate them in sequence using an iterative RANSAC framework. The details of pairwise registration are outlined in the flowchart in Figure 4.5, with each step described in detail in the following descriptive sequence.
Generate set of tie point triplets for $i$ and $j$.
(cf. $\binom{N}{3}$ triangles in Figure 4.6a, and $\binom{N}{3}$ triangles in Figure 4.6b).

Populate triplet sets with tie point feature information
DBH values $D_i, D_j$  
Spatial location $L_i, L_j$  
Eigenvalues $\Lambda_i, \Lambda_j$  
*Values in each $c^{th}$ and $c^{th}$ triplet are ordered by decreasing radius.

Remove highly collinear triplets
Reduced sets are: $\tilde{D}$, $\tilde{L}$, and $\tilde{\Lambda}$
There are $N_{C_i} < \binom{N_i}{3}$ triplets for set $i$, and $N_{C_j} < \binom{N_j}{3}$ triplets for set $j$

Evaluate correspondence between triplet pairs from $i$ and $j$.
There are $N_{C_i} \times N_{C_j}$ possible pairs.
1. Remove pairs with dissimilar radii.
2. Sort pairs by intrinsic geometric similarity, or “likelihood”, $\mathcal{L}$.
(See Figure 4.8).

Determine 3-D rigid transformation parameters (cf. RANSAC)
Iteratively,
1. Select minimum sample set (MSS), i.e., a pair of triplet tie points
2. Fit model using SVD.
3. Transform full set $T_j$ into LCS$_i$.
4. Determine the number of inliers (trees whose locations match within $\epsilon_{\text{min}}$)
   Save best transformation parameters, $R$, $t$.

Figure 4.5: Flowchart for the pairwise transformation from scan $j \rightarrow i$. 
4.4. METHODS

4.4.5.1 Generate a set of tie point triplets

The MSS for a rigid transformation in 3-D is three, i.e., a “triplet” of tie points. Therefore, we determine all possible combinations of the $N_i$ tie points, taken $k = 3$ at a time for scan $i$, and all possible combinations of the $N_j$ tie points, taken $k = 3$ at a time for scan $j$. There are $\binom{N_i}{3}$ triplets for scan $i$ and $\binom{N_j}{3}$ triplets for scan $j$. Triplets are visualized in Figure 4.6 as the set of all possible (red) and (blue) triangles created by drawing edges between the trees (tie points) in each stem map. Recall that each shaded circle represents a stem detected at that location, with a DBH corresponding to the diameter of that circle. In this example, there are $N_i = 12$ trees detected in scan $i$, and $N_j = 13$ trees detected in scan $j$. The larger black circles delineate the nominal range of $S_i$ and $S_j$, so that only tie points that are shared between scans $i$ and $j$ (i.e., purple) can be used for successful transformation.

![Figure 4.6: Tie point sets (a) $T_i$ (red) and (b) $T_j$ (blue), as plotted in their local coordinate system (LCS)'s, where a 90 degree rotation has been applied around the z-axis (yaw). The darker, shaded purple edges identify shared tie points. The objective of registration is to find the transformation which maps $T_j$ into the coordinate system of $T_i$ (see Figure 4.7).](image)

Shared points can be visualized by plotting the tie points in the same coordinate system, i.e., as they exist in the world. This is shown in Figure 4.7, where scan $j$ was positioned with an $x, y$ translation offset and $90^\circ$ rotation offset around the z-axis (yaw) relative to scan $i$. 
4.4. METHODS

Figure 4.7: Tie point sets $T_i$ (red) and $T_j$ (blue), as plotted in the world coordinate system (WCS) defined by set $i$. The larger black circles delineate the nominal range of $S_i$ and $S_j$, so that only tie points that are shared (purple) between scans $i$ and $j$ can be used for successful transformation between pairs.

4.4.5.2 Populate triplets with feature information

Triplet sets for each scan are then populated with feature information, including their DBH and location values. We use the stem axis-terrain intersection point as an unambiguous localization of a stem’s position.

Let $D_i$ be the set of $\binom{N_i}{3}$ triplets of DBH values, $d$, for scan $i$.

Let $D_i = \{d_{c_i=1-\binom{N_i}{3}}\}$ \hfill (4.3)

$d_{c_i} = \begin{bmatrix} d_{c_i,1} & d_{c_i,2} & d_{c_i,3} \end{bmatrix}^T$ \hfill (4.4)

$k$-combinations of $N_i$ are sufficient, rather than the $(k!)^2$ times more computationally expensive set of permutations, by sorting the triplet of DBH values such that $d_{c_i,1} \leq d_{c_i,2} \leq d_{c_i,3}$. This forces the correspondence between tie points that have the most likely similar radii.
Likewise, let $L_i$ be the set of $\binom{N_i}{3}$ triplets of tie-point location values, $l$,

$$L_i = \left\{ l_i = \binom{N_i}{3} \right\}$$  \hspace{1cm} (4.5)

where $l$ is a set of three tie points, $l$, with each tie point defined as a stem-terrain intersection point.

$$l_i = \begin{bmatrix} l_{i,1} & l_{i,2} & l_{i,3} \end{bmatrix}^T$$  \hspace{1cm} (4.6)

The ordering of $d_i$ is applied to $l_i$ for consistency.

An additional feature is then computed to characterize the triplet’s intrinsic geometry. A triplet’s intrinsic geometry can be assessed by eigenvalue analysis of the covariance matrix of the set of three location points. For the $c_i$th triplet, $\lambda_{c_i,1}$ describes the variance in the principal direction (described by the first eigenvector), while $\lambda_{c_i,2}$ describes the variance in the direction orthogonal to the first eigenvector. While the mean describes the location of a set of points in space, and the eigenvectors describe their orientation, the eigenvalues describe the intrinsic geometry of that set, regardless of its extrinsic position and orientation. This information is used in subsequent steps to reduce the computational complexity by only searching tie point sets with similar geometry.

Mathematically, from $L_i$ we can compute the set $\Lambda_i$ of $\binom{N_i}{3}$ eigenvalue triplets, $\Lambda$.

$$\Lambda_i = \left\{ \lambda_i = \binom{N_i}{3} \right\}$$  \hspace{1cm} (4.7)

where

$$\lambda_i = \begin{bmatrix} \lambda_{c_i,1} & \lambda_{c_i,2} & \lambda_{c_i,3} \end{bmatrix}^T$$  \hspace{1cm} (4.8)

Note that $\lambda_{i,3} = 0$, because $k = 3$ points lie in a plane.
4.4.5.3 Remove highly collinear triplets

After populating these feature sets, we remove triplet sets, which are highly collinear, because collinear point sets do not adequately constrain the third degree of freedom for a rigid transformation and may introduce error in the output parameters. We subset $D_i, L_i, \Lambda_i$ by only keeping combinations $c_i$, such that the following collinearity condition is met:

$$\frac{\lambda_{c_i,2}}{\lambda_{c_i,1} + \lambda_{c_i,2}} \geq \lambda_{\text{coll}}$$

(4.9)

The reduced sets, $\tilde{D}_i, \tilde{L}_i, \tilde{\Lambda}_i$ contains $N_C < \binom{N_i}{3}$ candidate tie point sets.

4.4.5.4 Evaluate correspondence between triplet pairs

The next step is to sort potential correspondences between pairs of feature triplets for a given scan $i$ and $j$. First, the reduced set of tie-point triplets are generated for a scan pair $i$ and $j$, giving a pair of feature vectors, including the ordered DBH values ($\tilde{D}_i, \tilde{D}_j$), stem-terrain intersection points ($\tilde{L}_i, \tilde{L}_j$), and their intrinsic geometry ($\tilde{\Lambda}_i, \tilde{\Lambda}_j$), with $N_{C_i}$ triplets in set $i$ and $N_{C_j}$ triplets in set $j$. There are $N_{C_i} \times N_{C_j}$ possible pairs.

We remove pairs that have dissimilar diameter values for any of the three corresponding DBH values. Because the DBH values in each triplet were previously sorted, we can compare diameter values on an element-by-element bases and remove pairs, which do not satisfy the following requirement, with $t_d = 20\%$.

$$\frac{d_{c_i,l} - d_{c_j,l}}{(d_{c_i,l} + d_{c_j,l})/2} < t_d \quad \text{for } l = 1, 2, 3$$

(4.10)

The remaining $N_P < N_{C_i} \times N_{C_j}$ triplet pairs are then sorted according to their geometric similarity to determine the most likely correspondence pairs. This rotation- and translation-invariant description of geometry allows us to evaluate the geometric similarity of point sets measured in independent LCS’$s$. We compare the eigenvalues between each triplet pair to assess the intrinsic geometric similarity of a tie-point feature pair. We define the geometric similarity, or “likelihood”, $\mathcal{L}$, of correspondence between a pair of tie-point triplets for sets $i$ and $j$ as
\[ L_p = (\lambda_{c,1} - \lambda_{c,1})^2 + (\lambda_{c,2} - \lambda_{c,2})^2 \]  

(4.11)

where \( L_p \) is evaluated for all \( N_p \) triplet pairs. All possible triplet pairs are then sorted according to \( L_p \) (Figure 7).

4.4.5.5 Determine rigid transformation parameters (RANSAC)

An iterative process then evaluates each pair in order of decreasing likelihood (Figure 4.8) by computing the parameters of a rigid transformation model and determining the pair, which gives the best “fit”. A pair of triplet location values from set \( i \) and \( j \) are selected, \((l_i, l_j)\), based on their likelihood, \( L \). Parameters of a three-dimensional (3-D) rigid transformation are then computed, which minimize the least-squares error between the three corresponding points in the direction.

Figure 4.8: Pairs of tie-point triplets are examined iteratively in order of decreasing likelihood, \( L \).
4.4. METHODS

Using singular value decomposition (SVD) (Besl and McKay, 1992). There are six parameters in the model, i.e., three Euler angles \((\theta_x, \theta_y, \theta_z)\) and three translation parameters \((t_x, t_y, t_z)\). From these, a rotation matrix, \(R_{ij}\), and translation vector, \(t_{ij} = [t_x \ t_y \ t_z]^T\) can be constructed. The full tie-point set \(I_j\) (containing all \(N_j\) tie points), is then transformed into \(LCS_i\) using the rotation and translation parameters.

\[
I_j = R_{ij}I_j + t_{ij} \tag{4.12}
\]

where we have added the subscript to designate source scan, \(j\), and the superscript to designate the target scan, \(i\), i.e., the transformation pair \((R_{ij}, t_{ij})\) corresponds to a transformation from \((j \rightarrow i)\).

The transformed locations are updated in the new tie point set, \(T_{ij}\). The DBH values remain unchanged.

\[
T_{ij} = \begin{cases} 
I_j = I_{i1} \ldots I_{it} \ldots I_{iN_j} \in \mathbb{R}^3 \\
T_{ij} = d_{i1} \ldots d_{it} \ldots d_{iN_j}
\end{cases}
\]

Tie point sets \(T_{ij}\) and \(T_i\) are then compared to determine the number of matching tie points, based on Euclidean distance. A match is successful between any tree \(I_i\) and \(I_j\) if the Euclidean distance between them is less than a predefined threshold, \(e_{min} = 0.4\) m.

\[
\left[ l_i(x) - l_j(x) \right]^2 + \left[ l_i(y) - l_j(y) \right]^2 + \left[ l_i(z) - l_j(z) \right]^2 < e_{min} \tag{4.13}
\]

The number of matches is tallied, and if the quantity is larger than the current “best” number of matches, the model parameters are updated as the current “best” model. This process is repeated for each triplet pair in descending order of likelihood, \(L\). The algorithm terminates after evaluating all \(N_P\) pairs, or after a predefined number of RANSAC iterations, \(N_R\). Because the pairs are sorted by likelihood, we can have a low \(N_R\) value (for computational speed), while maintaining a high probability of finding the best model. As a final step, the model parameters are recomputed based on the full inlier set of corresponding point sets. This improves the precision of model parameters by taking into account a more distributed network of corresponding points.
The output of this parameter fitting phase (See Figure 5.2, first row) are estimates of the six transformation parameters, \((t_x, t_y, t_z, \theta_x, \theta_y, \theta_z)\), which transform point cloud data \(j\) into \(i\).

### 4.4.6 Pairwise Error Assessment

While section 4.4.5 provides an estimate of the six transformation parameters, \((t_x, t_y, t_z, \theta_x, \theta_y, \theta_z)\), section 4.4.6 provides an estimate of the error associated with each of these output parameters. This embedded confidence metric is produced for each output transformation model by performing the following steps: Let \(T^A_i\) and \(T^B_i\) be disjoint sets of \(T_i\), such that

\[
T^A_i \cap T^B_i = \emptyset \tag{4.14}
\]

\[
T^A_i \cup T^B_i = \{T_i\} \tag{4.15}
\]

Although ownership in each set could be determined randomly, we found that this approach reduced the frequency of correct linkages for the “return” trip, i.e., from \(j \rightarrow i\). This is because an excess of corresponding points may be naively assigned to set \(A\), leaving set \(B\) with an insufficient number of corresponding tie points necessary for round-trip validation. Instead, we instantiate set \(A\) as the triplet of tie points, which provided the optimal transformation model in the first phase.

\[
T^A_i = l_{i,\text{best}} \tag{4.16}
\]

\[
T^A_j = l_{j,\text{best}} \tag{4.17}
\]

Consequently, set \(B\) contains the remaining tie points for scans \(i\), and \(j\):
This maximizes the opportunity for determining two disjoint transformation paths, while maintaining set independence necessary for validation. The formulation from section 4.4.5 can then be repeated for each disjoint set in order to determine a pair of transformation parameters associated with the circular path from $i \to j \to i$. Namely, for each transformation pair $(i, j)$ we determine the rotation, $R_{jB}^i$, and translation, $t_{jB}^i$, associated with the forward path from $i \to j$ via the disjoint set $B$ (Figure 4.9a).

\[ T_j^B = l_j - 1_{c_{j, \text{best}}} \]  
(4.18) \[ T_i^B = l_i - 1_{c_{i, \text{best}}} \]  
(4.19)

Likewise, we determine the rotation, $R_{iA}^j$, and translation, $t_{iA}^j$, associated with the reverse path from $j \to i$ via the disjoint set $A$ (Figure 4.9b).

\[ l_j^{Bj} = R_{jB}^i l_i^{Bj} + t_{jB}^i \]  
(4.20) \[ l_i^{Aj} = R_{iA}^j l_j^{Aj} + t_{iA}^j \]  
(4.21)
4.4. METHODS

Figure 4.9: (a) Disjoint sets $\mathcal{T}_i^B$ (red) and $\mathcal{T}_j^B$ (blue) provide a “forward” path from $i \rightarrow j$. (b) Disjoint sets $\mathcal{T}_i^A$ (red) and $\mathcal{T}_j^A$ (blue) provide a “reverse” path from $j \rightarrow i$. Shared tie points are colored purple. Note the vastly fewer number of connections between disjoint scans, because $((N_i^2)/(N_j^2))^{\frac{1}{2}} \approx 0.01$ for large $N_i$.

We can then travel a circular path using the disjoint point sets $(\mathcal{T}_i^B \rightarrow \mathcal{T}_j^B; \mathcal{T}_j^A \rightarrow \mathcal{T}_i^A)$ and apply the transformation parameters associated with each path in sequence, i.e.,

$$\begin{align*}
    t_{BA} &= R_{jA} \left( R_{iB} t_i^B + t_{iB} \right) + t_{jA}^A \\
    &= R_{BA} t_i^B + t_{BA}^A
\end{align*}$$

(4.22)

(4.23)

with the effective rotation $R_{BA}$ and translation $t_{BA}$ of the circular path computed as follows:

$$R_{BA} = R^A R^B$$

(4.24)

$$t_{BA} = R^A t^B + t^A$$

(4.25)

(4.26)
By construction, \( \mathbf{R}_{\text{true}} = \mathbb{I} \) and \( \mathbf{t}_{\text{true}} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \). The extent to which the observed transformation parameters for a circular path \((\mathbf{R}^A, \mathbf{t}^A)\) deviate from truth (0) provides an explicit upper-bound estimate of the precision of output model parameters. We define an upper estimate of the error in a transformation parameter, \( p \), as

\[
e_p = |p^{AB}|, \text{ for } p = t_x, t_y, t_z, \theta_x, \theta_y, \theta_z
\]  

(4.27)

This provides an embedded upper-bound error metric associated with each output transformation parameter. Note that the composition and decomposition of the rotation matrix, \( \mathbf{R} \), into component Euler angles \((\theta_x, \theta_y, \theta_z)\), is applied consistently in order to allow direct comparison between true and observed Euler angles.

In order to provide a more intuitive, readily comparable error metric, we also applied the same framework to compute an RMSE of the tie points. From Equation 4.23, note the superscript and subscript of the result, \( \hat{t}^{Ai} \), are identical, i.e., the tie points are mapped back into the original coordinate system \( LCS_i \). Thus by construction, there is element-by-element correspondence between \( l \) and \( \hat{l}^{Ai} \). We compute the RMSE for a transformation pair as

\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{y} (\hat{t}^{Ai} - \hat{l})^2}
\]  

(4.28)

These are both upper estimates of the error, because only half of the tie point set was available for each transformation path (i.e., the sets were disjoint). Moreover, errors associated with the circular path are compounded (doubled) compared to the traverse of a single path \( j \rightarrow i \), used to compute the best model in section 4.4.5. Moreover, fewer tie points are used, which introduces greater error (see Figure 4.16).

These error assessments are automatically produced with every output transformation to provide an embedded confidence metric for the end user. We validated this error metric by manually labeling truth data for \((\binom{M=25}{2} = 300 \text{ pairs per site}) \times 11 \text{ sites} \times 2 = 6,600 \text{ pairs}, where the
factor 2 is added because both the forward and reverse transformation paths are independently validated. Some sites had a reduced number of scans due to feasibility limitations, thus reducing the total number of pairs = 5,585. These pairwise registration results were then collated to generate a histogram of RMSE error for various labeled classes, and corresponding ROC curves.

4.4.7 Algorithm Performance Analyses

We evaluated the performance of the algorithm in terms of its sensitivity to (i) rotation/translation offset, (ii) RMSE in the input tie-point locations, (iii) error in the tree DBH measurements, and (iv) reducing the number of matching tie points. To capture the variability of the New England study sites’ forest structure, these sensitivity analyses were performed for tie point data derived from each scan collected during the field campaign, i.e., for each of the (nominally) 25 scans per 22 sites, yielding a total of 550 unique input tie point sets.

To assess the sensitivity to rotation/translation, we designed an experiment where a truth tie point set was rotated and translated by known rotation and translation parameters. Pairwise registration was then performed, as outlined in 4.4.5, to estimate the rotation and translation parameters, which map the transformed tie point set back to its initial coordinate system. Since, by construction, the truth parameters are known exactly, we can determine the error in truth parameters (or more intuitively, the RMSE between corresponding tie points), to assess the sensitivity of pairwise registration to rotation and translation offsets. By design i.e., iteratively searching possible combinations of triplet tie points, the proposed algorithm should be invariant to this modulation of sensor pose.

The second experiment assessed the sensitivity of the registration algorithm to errors in the input tie point tree locations, $\text{RMSE}_{\text{in}}$. Error in the input tie point locations is due to external processing algorithms, which are unable to precisely localize the stem-terrain intersection point. A perfect registration algorithm would be $\text{RMSE}_{\text{in}}$-limited, with any additional RMSE (implicit subscript “out”) due to the registration algorithm. We hypothesized that the RMSE of registration would be roughly equivalent to $\text{RMSE}_{\text{in}}$ of the input tie point locations, up to a limit where noise approached the between-tree distance, and errors largely increased. To evaluate this hypothesis,
we rotated and translated the truth tie point locations and then added noise from a Gaussian distribution with mean zero and successively larger $\sigma$ to the truth tie point locations. RMSE$_{in}$ was computed from the noise-added input tie point set. We then performed registration between the truth tie point set, and the noise-added tie point set, and calculated the RMSE of the result.

A third experiment assessed the sensitivity of the registration algorithm to error in the input tie point DBH values. This error is also due to external processing algorithms, and reflects the sensitivity of the proposed algorithm to noise in the input data. We rotated and translated the truth tie point locations and then added noise from a Gaussian distribution with mean zero and successively larger $\sigma$ to the truth tie point DBH values. Noise was defined as a percentage of the true DBH value.

The final experiment assessed the number of matching tie points that are required between scans. For an initial tie point set, we progressively subset the tie point set to fewer and fewer tie points, until the minimum required for a rigid transformation (3). We then performed pairwise registration between the subset tie point set and the full set, and calculated the RMSE of the result. We hypothesized that the error would be slightly larger for small subsets due to the losses in precision when re-estimating the model without a full set, where errors are distributed randomly throughout many samples and are effectively minimized.

4.5 Results

4.5.1 Point cloud

An example of a registered point cloud is shown for a selected site in Figure 4.10. The scans were collected 13.9 m apart (note that the center occlusion cones indicate nominal sensor position) and were registered without a priori knowledge of scanner pose. The output transformation parameters and per-parameter errors for this site are reported in Table 4.2. Embedded upper-bound confidence metrics report an RMSE = 18 cm. The advantage of multi-scan data collection is readily apparent. Note how the sparse sampling of the center, v-shaped stems by sensor $j$ is augmented by data from $i$, which has a higher effective resolution at closer range.
Figure 4.10: (a) Point cloud $j$ (red) is mapped to point cloud $i$ (blue) using the proposed marker-free approach with a reported $RMSE = 18$ cm. Between-scanner distance was $\approx 14$ m, and no a priori estimate of pose was used in the transformation. (b) Detail inset shows measurements samples on opposing sides of tree stem structure due to the differing sensor positions. Note how the sparse sampling of the center, v-shaped stems by sensor $j$ is augmented by registered data from $i$. Error estimates for this site are listed in Table 4.2.
Table 4.2: Reported rigid transformation parameter estimates and errors for a registration pair from the site shown in Figure 4.10.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Upper-bound error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_x$</td>
<td>8.95 m</td>
<td>0.8 cm</td>
</tr>
<tr>
<td>$t_y$</td>
<td>-10.60 m</td>
<td>11.1 cm</td>
</tr>
<tr>
<td>$t_z$</td>
<td>-1.06 m</td>
<td>6.3 cm</td>
</tr>
<tr>
<td>$\theta_x$</td>
<td>1.11°</td>
<td>0.43°</td>
</tr>
<tr>
<td>$\theta_y$</td>
<td>0.51°</td>
<td>0.15°</td>
</tr>
<tr>
<td>$\theta_z$</td>
<td>-4.27°</td>
<td>0.65°</td>
</tr>
</tbody>
</table>

4.5.2 Transformation Parameters

We collated the error in transformation parameters for all scan pairs, identified manually as true matches, and reported the mean values in Table 4.3. The mean error in $z$–translation was slightly higher than that for $x$ and $y$ (12.4 cm vs. 7.8 cm and 7.2 cm, respectively). The mean errors in Euler angles were below 1°.

Table 4.3: Mean upper-bound transformation parameter errors for all transformation pairs.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean absolute error, $e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_x$</td>
<td>7.8 cm</td>
</tr>
<tr>
<td>$t_y$</td>
<td>7.2 cm</td>
</tr>
<tr>
<td>$t_z$</td>
<td>12.4 cm</td>
</tr>
<tr>
<td>$\theta_x$</td>
<td>0.934°</td>
</tr>
<tr>
<td>$\theta_y$</td>
<td>0.480°</td>
</tr>
<tr>
<td>$\theta_z$</td>
<td>0.188°</td>
</tr>
</tbody>
</table>

While the parameter errors provide insight into the transformation model, the RMSE provides a more convenient metric for comparison. The mean RMSE for true matches was 16.3 cm. The subsequent results report error in terms of RMSE.

4.5.3 Tree Locations

A histogram of RMSE values and corresponding ROC curves are plotted in Figure 4.11. Each of the 5,585 registration pairs were manually classified as either “true” (T) or “false” (F). Since an RMSE is calculated based on a combined traversal through forward and reverse disjoint sets,
we have the following pairs of truth labels: TT (true forward path, true reverse path), FF (false forward path, false reverse path), TF (true forward path, false reverse path), and FT (false forward path, true forward path). Note that any of the circular paths with a false path will have high error. Thus, we combine FF, TF, and FT as a new class, F, yielding the familiar two-class problem.

From Figure 4.11a we observe that there are a high number of TT paths that have a computed RMSE error $< 50$ cm. The mean RMSE value for TT paths is 16.32 cm. There is significant class separation between this distribution and the distribution of F classes, which exhibits a nominal Gaussian distribution with mean 20 m and standard deviation 5 m. Of the three F classes (FF, TF, FT), we note that the majority of high-error paths were due to incorrect matches in both the forward and reverse paths (FF). Fewer pairs had a correct forward path, but an incorrect reverse path (TF). Still fewer had an incorrect forward path, but a correct reverse path (FT).

The two-class labeling of truth data yields the ROC curves for quantifying the robustness of the embedded RMSE confidence metric. Figure 4.11b shows the correct detection rate vs. incorrect detection rate. The correct detection rate is nominally 1 across the range of incorrect detections, i.e., the RMSE provides a robust metric for class separation. Figures 4.11c,d expound on this by coupling the RMSE to detection rates. Figure 4.11c reveals that unit correct detection is possible with a threshold greater than RMSE $= 3$ m. Likewise, in Figure 4.11d, incorrect detection rates are very low for threshold RMSE $< 3$ m. These results confirm the validity of the proposed embedded confidence metric.
Figure 4.11: (a) A histogram of RMSE values classified according to manual truth data to identify each path of a circular transformation pair as either true (T) or false (F). (b) This classification yields the ROC curve, with (c) near perfect correct detection and (d) expected incorrect detection rates as a function of RMSE threshold. These graphs confirm the validity of proposed embedded confidence metrics.

Figure 4.12 plots RMSE as a function of distance between sensors for all identified registration pairs with a valid error (i.e., TT). As expected, we observe a slight trend of increasing RMSE vs. distance.
4.5. RESULTS

Figure 4.12: Box plot showing the middle 50% quantiles for collated RMSE values vs. distance between sensors. RMSE increased only slightly with respect to between-sensor range.

Perhaps more salient than RMSE is the analysis of pairwise transformation detection rates vs. distance. As the sensors are further separated in range, there is a reduced probability that they sample a sufficient number of corresponding trees, which are necessary to generate matching tie points for registration. As a result, “detection rate”, or the percentage of successful pairwise transformations, is a more meaningful analysis to inform recommendations on sample spacing. Figure 4.13 plots the detection rate vs. their between-sensor range. These were plotted separately for both forward and reverse paths in order to provide insight into the circular path construction. The asterisk marker identifies contributions from forward paths. The circle marker identified
contributions from reverse paths. There is a reduced area of overlap (dashed line) due to the input point cloud data being cropped to exclude points greater than 16 m in range from the sensor. To account for this, we normalized the raw data (gray lines) by the range-dependent overlap area in order to produce a normalized detection percentage (black lines) vs. range for both forward (asterisk) and reverse (circle) paths. We observe that forward paths maintain high detection rates (85%) until rapid fall-off after 15 m. Reverse paths, which operate from a reduced set of tie points, exhibit a more linear reduction in detection rate vs. range.
4.5. RESULTS

Figure 4.13: Pairwise transformation detection percentages detailing the number of scanner pairs successfully linked (according to truth labeling) vs. their between-sensor range. Raw results (gray lines) are normalized by sensor overlap area (black lines). Separate plots are provided for both forward (asterisk marker) and reverse (circle marker) paths in order to provide insight into the circular path construction. The range-dependent sensor overlap area, normalized to 100%, is plotted in dotted gray.

4.5.4 Performance Analyses

The first analysis assessed the invariance of the proposed algorithm to sensor pose. Results showed that an \( \text{RMSE} = 0 \) was achieved, regardless of translation and rotation offset between two tie point sets, as expected. On the other hand, we hypothesized that RMSE would be largely driven by noise in the input data sets (stem-terrain intersection points). External algorithms may be unable
to precisely localize these 3-D locations, and we postulated that this source data noise will have a major impact on the output transformation parameters. Figure 4.14 plots the output RMSE vs. the input RMSE$_{in}$ of the tie point locations for all manually-identified $TT$ transformation pairs. We observe that output RMSE is consistently lower than RMSE$_{in}$, except for RMSE$_{in}$ $> 0.45$ m, at which point the variance of RMSE increases, along with a progressively larger bias towards high RMSE for associated increases in RMSE$_{in}$.

![Figure 4.14: Collated RMSE values vs. the RMSE$_{in}$ added to the input tie point set reveal that registration results are largely dependent upon the quality of input tie point sets. This extends the results to other study areas/TLS sensors, and suggests that higher precision could be achieved for TLS instruments with improved precision.](image-url)

Figure 4.15 shows a box plot of RMSE vs. percent error (drawn from a normal distribution...
with \( \mu = 0, \sigma \) added to the input tie point DBH values. Center-50% quantile RMSE errors are small (< 10 cm) for \( \sigma < 360\% \). For \( \sigma > 120\% \), there is a distinctly greater spread of high-RMSE outliers, though this is not a concern because embedded confidence metrics, in conjunction with an appropriate threshold determined from Figure 4.11, will flag these as bad transformation pairs.

![Figure 4.15: Collated RMSE values vs. percentage \( \sigma \) added to the input tie point DBH values reveal low sensitivity of the algorithm to noise in DBH, up to \( \sigma > 120\% \), after which center-50% quantiles remain low, with increasing high-RMSE outliers.](image)

The final analysis added noise and a random transformation to a reference tie point set and then progressively subset it to reduce the number of matching tie points in the second tie point set. Figure 4.16 plots the RMSE as a function of the number of subset tie points in the second set.
Note that when no noise was added to the tie point locations, zero error was achieved regardless of the number of tie points, as expected from previous results. Therefore, an $\text{RMSE}_{\text{in}}$ floor was added to the input tie points (dotted line), which serves as a baseline for evaluation. We observe that the RMSE is fairly consistent for subsets of 13-60 matching tie points. As the number of tie points is progressively reduced (from 13 to the MSS of 3), the error increases. This is because RANSAC recomputes the model parameters with the full inlier set in order to improve fit. For small numbers of tie points, this secondary step is less effective at minimizing RMSE error. An equation was fit to the data points, resulting in $\text{RMSE} = \frac{0.56}{N_{\text{sub}}} + 0.31$.

![Figure 4.16: Input tie point sets were progressively subset to reduce the number of matching tie points. Consistent RMSE on the order of $\text{RMSE}_{\text{in}}$ is observed for 13-60 corresponding tie points, with an increase in RMSE for few corresponding tie points. Although the model is still able to determine accurate registration parameters with at least the MSS, precision is lost due to RANSAC’s inability to minimize error through re-estimation of the full set.](image-url)
4.6 Discussion

Terrestrial laser scanning offers a compelling potential for comprehensive measurement of forest structure, especially when multiple scans are combined to overcome limitations of laser beam occlusion. Efficient registration of multiple-scan information in the forest environment, however, has remained a challenge for operational inventories, especially for the growing number of affordable laser scanner sensors that are not supported by current registration packages (Hilker et al., 2012b). We addressed this knowledge gap by developing a robust, automatic registration approach that is invariant to differential sensor pose and that does not require external markers.

Moreover, an inherent limitation of existing marker-free registration approaches is the lack of output error metrics associated with each transformation pair. We developed an error assessment framework, using set theory, to produce an upper estimate of the six transformation parameter errors associated with each registration. Note that this framework for embedded confidence metrics is extensible to any view-invariant tie point feature set, and thus may have application outside the domain of forest terrestrial laser scanning. Output transformation parameter errors were produced for each registration pair, in order to provide explicit insight into the rigid transformation model. Mean reported absolute errors of translation were greatest for the $z$ component (Table 4.3), which is expected due to the challenge of sampling the terrain at forest plots with significant ground vegetation. As a result, for trees farther from the sensor, the increased beam size, as a function of system beam divergence, and shallow incidence angle make it more difficult to correctly localize the $z$-component of the stem-terrain intersection point. Mean reported absolute errors of rotation were all under 1°, with $x$, $y$, $z$ components having progressively smaller error. This may be due to the order of decomposition of the rotation matrix. We also computed an RMSE between tie-points to provide a simple, readily-comparable metric of error. Subsequent discussion utilizes this error metric for analysis.

RMSE was evaluated vs. (i) sensor pose offset, (ii) RMSE$_{in}$ of the input tie point locations (Figure 4.14), (iii) percent error added to the input tie point DBH values (Figure 4.15), and (iv) number tie points (Figure 4.16). Zero RMSE values were observed regardless of differential sensor pose. In other words, the location and orientation offset between two laser scanner measurements
did not affect RMSE. This was as expected, given the view-invariant description of tie points, i.e., the stem-terrain intersection point is computed as approximately the same position in space regardless of from which side the tree is viewed.

We identified several challenges where this approach may fail. First, homogeneous forest stands, e.g., plantations, may have insufficient geometric dissimilarity necessary to provide an unambiguous coordinate transform. Initial pose estimates may be necessary in this case to constrain the transformation. Second, there are two sources of error in the tie points, which introduce noise in the source data. The assumption of radial symmetry may fail for tree boles that are not cylindrical, introducing error in the (x,y-position) of the measured tie points. Moreover, for forested areas with significant ground vegetation, inabilities in sampling the obscured terrain could introduce errors in the (z-position) of the stem-terrain intersection point. Mean absolute errors were largest for the z component (Table 4.3), which corroborated the hypothesis that terrain detection was the primary challenge for forest sites with ground vegetation. A third analysis assessed the sensitivity of the algorithm to error in the DBH values. RMSE was found to be largely insensitive to error of the input tie point DBH values (Figure 4.15). While the algorithm removes tie points whose radii are \(< t_d = 20\% \) (Equation 4.10), Figure 4.15 demonstrates a resistance to diameter values of much larger deviation. This is explained by the capability of a voting method such as RANSAC to find an optimal MSS from among a large number of samples containing noise.

As expected, we showed that it is the noise in the input tie point locations, which drives the RMSE metric of the output registration results. In Figure 4.14, we added noise to an input tie point set, and then rotated/transformed those points by a random amount. The modified tie point set was then registered to the initial tie point set, in order to assess the RMSE error as a function of RMSE\(_{in}\). Output RMSE was consistently lower than RMSE\(_{in}\) for RMSE\(_{in}\) < 45 cm. This shows that the proposed registration approach is RMSE\(_{in}\)-limited, i.e., the output error is lower than the error of the input source data. This promising result demonstrates the utility of RANSAC and the deterministic SVD model to determine, in the presence of noise, the rigid transformation parameters that minimize the least-squares error. This is especially encouraging given that the RMSE metric represents an upper estimate of the error associated with the pairwise transformation,
4.6. DISCUSSION

owing to the requirement of a round-trip traversal in order to generate truth data. As a result reported errors are compounded (doubled). Moreover, the second path has less precision than the first, because the three best tie points have been removed from the model by necessity of ensuring independence (i.e., disjoint sets). The true RMSE errors for a single forward path can be expected to be on the order of half the reported RMSE metric. For $\text{RMSE}_\text{in} > 45$ cm, the variance of output RMSE increases, along with an increasing bias towards high RMSE values. However, real data are unlikely to have such high $\text{RMSE}_\text{in}$ values. This effectively constrains the results of this algorithm to the $\text{RMSE}_\text{in}$-limited regime.

Moreover, this linear relationship suggests that a higher-precision instrument (with lower $\text{RMSE}_\text{in}$) could achieve output RMSE on the order of the instrument precision. This extends conclusions on the algorithm precision beyond the specific sensor and sample areas used in this study. In this study, input data were derived from (Kelbe et al., 2015b), which used a low-resolution (15 mrad beam divergence) sensor and reported an $\text{RMSE}_\text{in} = 16.53$ cm. Mean output RMSE for all true matches was 16.3 cm, following the expected output RMSE vs. $\text{RMSE}_\text{in}$ analysis in Figure 4.14. Likewise, a higher-resolution sensor could provide more precise localization of tie-point locations, with RMSE of registration results expected to follow commensurate to the $\text{RMSE}_\text{in}$ of tie points.

The final analysis progressively subset an input tie point set after adding noise and a rotation/translation offset. The reduced subset of tie points was then registered to the full set, and the RMSE was plotted vs. number of subset tie points. Results confirmed our hypothesis that the algorithm is fairly invariant to the number of corresponding tie points (Figure 4.16). RANSAC, by design, needs just a MSS in order to compute the rigid transformation model. Therefore, RMSE values are $\text{RMSE}_\text{in}$-limited, except for a slight increase in error for low ($< 13$) tie points. This is because RANSAC’s secondary stage, which re-estimates the model using the full inlier set, is less effective at minimizing global RMSE error with fewer available tie points.

In an effort to inform optimal sampling in New England forest environments, we collated RMSE errors for 5,585 registration pairs and plotted RMSE vs. between-sensor distance (Figure 4.12). Distance between scans ranged from 5 - 28 m. RMSE increased only slightly with respect
to range, with mean RMSE values ranging from 16 cm (0-5 m between-sensor range) to 52 cm (20-25 m between-sensor range). If RMSE was the only criterion, this relationship gives little motivation for reducing the sensor spacing to maximize registration precision. However, we also expected that the percent detection between scans would decrease as a function of range. We plotted the percentage of scans that were successfully linked to each other vs. their between-sensor range (Figure 4.13), and found this to be the driving factor affecting sample spacing. The forward path maintained high detection percentages (85%, after correction for decreased area of overlap) until rapid falloff at 15 m range. The reverse path detection percentages, however, more linearly decreased as a function of range, dropping from 88% (0 m) to 0% (25 m). The implications are that the requirement of a disjoint return transformation path, which is necessary in order to compute the RMSE confidence metric, greatly inhibited the number of scans that could be correctly linked with an associated confidence metric. In other words, there may be cases where an output transformation path is deemed incorrect due high RMSE, but the forward path was able to estimate correct registration parameters. Thus, in cases where the algorithm is unable to produce a transformation model with low RMSE error, registration parameters from the forward path may contain an appropriate transformation model. This offers an opportunity to increase detection rates in cases where there are insufficient matching tie points to compute an embedded confidence metric.

4.7 Conclusions

This study quantified the RMSE of a proposed blind, view-invariant, marker-free registration approach for terrestrial laser scanner data in forest environments. An embedded confidence metric was developed using set theory to provide an upper estimate of the error associated with each transformation pair, and was validated using manual truth classification and receiver operator curves (ROC’s). Rigorous analyses showed that (i) the algorithm is invariant (blind) to initial sensor pose, insensitive to error in DBH values, and possible with at least 3 (the MSS) corresponding tie points between scans. We collated transformation results for 5,585 registration pairs in the New England forest environment, and found that while RMSE increased slightly
with range between scanner locations, there was a much more prominent effect on the percentage of scan pairs that could be successfully linked, due to occlusion and a lack of corresponding objects within the scanners’ fields of view. This informed considerations for optimal sample spacing for TLS data collection in New England or similar forests. Finally, we demonstrated that the registration algorithm is RMSE\textsubscript{in}-limited, which extends results to other sensors and study areas. Owing to the minimization of least-squares error by RANSAC and SVD, output RMSE of registration can be expected to be lower than the input error of the source data. This work provides an accessible and fully automatic approach for registering terrestrial laser scanner data without artificial targets, thus enabling rapid structural assessment for domains of forest inventory, airborne calibration/validation, and computer vision.
4.7. CONCLUSIONS
Chapter 5

Graphical marker-free registration of TLS data

5.1 Foreword

Pairwise registration, including the approach developed in Chapter 4, can provide positive registration results between scans, which share corresponding tie points, but has several remaining limitations. For example, occlusion or view disparities may reduce the number of scans that can be successfully linked to a single reference node, thus limiting the geographic extent. Moreover, pairwise registration results may be globally inconsistent, despite purported consistency at the local level, i.e., between pairs. As a result, multi-view registration is needed to perform global registration of the network of pairwise correspondences.

Therefore, we identified a third objective as follows:

3. Determine the error associated with multi-view, marker-free registration of terrestrial laser scanning (TLS) data in forest environments.

Chapter 5 describes in detail the study, methodology, and results related to objective 3. Outputs include an original, robust methodology, which performs multi-view registration of TLS data using
a graph theory approach. Pairwise registration connections from chapter 4 are used to initialize the edges of a graphical framework. We define edge weights from the pairwise embedded confidence metric of chapter 4, which has the potential to simplify the registration process, while improving the resistance to noise (Huber and Hebert, 2003). We compare the trade space of both sequential and simultaneous registration paradigms, and develop a hybrid approach, which maintains the advantages of each.

5.2 Abstract

TLS has demonstrated increasing potential for rapid, comprehensive measurement of forest structure, especially when multiple scans are spatially registered in order to reduce the limitations of occlusion. Although marker-based registration techniques (based on retro-reflective spherical targets) are commonly used in practice, a blind, marker-free approach is preferable, insofar as it supports rapid, operational data acquisition. To support these efforts, we extend the pairwise registration approach of Kelbe et al. (2015a), and develop a graphical framework to perform blind, marker-free, global registration of multiple point cloud data sets. Pairwise pose estimates are weighted based on their estimated error, in order to overcome pose conflict while exploiting redundant information and improving precision. The proposed approach was tested for eight diverse New England forest sites, with 25 scans collected at each site. Quantitative assessment was provided via a novel embedded confidence metric, with mean estimated RMSE of 7.2 cm and 89% of scans connected to the reference node. This study (i) assesses the validity of the embedded multi-view registration confidence metric (ii) evaluates the performance of the proposed registration algorithm, and (iii) demonstrates the improvement in plot-level forest parameter estimation afforded by multiple scans.

5.3 Introduction

Recent technological advancements have demonstrated the capacity of laser scanning to rapidly record detailed structural information, both on the ground and - for large-area operations - from air
and space (Bachman, 1979). Airborne (ALS) and terrestrial laser scanning (TLS) are active sensing systems, which measure geometric characteristics, as opposed to reflectance or other radiation signatures obtained by passive sensing systems. Thus, they provide an important link between vegetation structural and material properties, and the subsequent ecological features of interest (Zhou et al., 2014). Airborne laser scanning (ALS) has matured to operational use over the past decade for large-scale forest structure assessment (e.g., Wehr and Lohr, 1999; Nelson et al., 1988; Lefsky et al., 2002; Næsset, 2007); the reader is referred to Hyyppä et al. (2008) for a detailed review. However, airborne analyses rely on ground-truth information (e.g., inventory) for calibrating and validating landscape models (Liang et al., 2012; Yu et al., 2010). As such, they too are limited by the fidelity - the structural resolution - of ground-reference data provided from traditional forest inventory. Terrestrial laser scanning (TLS), on the other hand, is well-poised to address both the limitations in forest inventory (Hopkinson et al., 2004; Maas et al., 2008) and the calibration needs of airborne forest sensing, including ALS (Hilker et al., 2012a; Jupp, 2011; Lindberg et al., 2012; Liang et al., 2012).

TLS operates from a ground platform, and has emerged as an effective tool for rapid and comprehensive measurement of object structure. Early studies using TLS demonstrated the potential for TLS to provide automatic forest inventory of tree stems, including their location (Liang et al., 2012), height (Olofsson et al., 2014; Hopkinson et al., 2004) and diameter (Lovell et al., 2011). Other authors have demonstrated the extraction of plot-level variables, such as stem density and basal area (Tansey et al., 2009), and biomass (Yao et al., 2011). More recently, some researchers have focused on high-fidelity geometric reconstruction of tree architecture from TLS data, which may provide more detailed modeling of individual trees and canopies (Raumonen et al., 2013).

A persistent challenge for structural assessment from TLS, however, concerns the occlusion of the emitted laser beam (Eo et al., 2012; Sharp et al., 2004; Salvi et al., 2007; Weinmann et al., 2011). This is especially important in the forest environment, where the line of sight from a single view is typically short, due to occlusion of the laser beam by forest elements. To address this, data collected from multiple scanner locations are typically registered into a single common
coordinate system (Kang et al., 2009; Pingi et al., 2005; Theiler and Schindler, 2012; Stamos and Leordeanu, 2003). Relative registration is performed by estimating the three translation and three rotation parameters between two coordinate systems and modifying the data’s spatial coordinates accordingly (Hilker et al., 2012b). This registration is often a necessary preprocessing step for various TLS-based forest structure studies, including the extraction of dendrometric parameters (Bucksch and Khoshelham, 2013; Zhou et al., 2014), canopy assessment (Henning and Radtke, 2008), plot-level inventory (Kelbe et al., 2015b), and multitemporal forest monitoring (Henning and Radtke, 2008). Ultimately, multisensor registration between ALS and TLS data could provide synergistic structural ground truth data to support calibration/validation of large-scale, airborne sensing models (Henning and Radtke, 2008).

Traditionally, registration is commonly performed by placing manual targets in the scene, which serve as control points for marker-based registration (Boehm and Becker, 2007; Theiler and Schindler, 2012; Barnea and Filin, 2008). However, the placement of artificial targets is time-consuming, tedious, and hence costly. As a result, marker-free techniques are preferred, in order to improve field-scanning efficiency (Zhou et al., 2014) and make TLS cost-competitive relative to traditional forest inventory techniques (Ducey and Astrup, 2013). Unfortunately, the majority of existing marker-free techniques utilize iterative point matching (e.g., iterative closest point (ICP)) or surface matching (Huber and Hebert, 2003), both of which are successful only for engineered surfaces (Henning and Radtke, 2008). Thus, there limited use in forest environments due to factors such as occlusion, spatial variability, and movement, e.g., due to wind (Henning and Radtke, 2008). This is confounded by system contributors, such as the range-dependent point density and discrete sampling nature of laser scanning technology (Barnea and Filin, 2008). As a result, small sensor-displacements may yield drastic changes in scene content (Forsman and Halme, 2005), which challenge the establishment of reliable point or feature correspondences (Zhou et al., 2014).

Recent automatic, marker-free registration approaches, such as (Kelbe et al., 2015a) and (Henning and Radtke, 2008), offer the potential to rectify this disparity and improve the operational capabilities of TLS in forest environments. Multi-view registration offers the potential to identify
and remove locally consistent, but globally incorrect matches (Huber and Hebert, 2003), and to bring into alignment disconnected scans through a connected sequence. However, due to the large nonlinear search space and the volume of input TLS data, multi-view registration imposes significant new challenges (Stamos and Leordeanu, 2003).

Although there is limited background on multi-view registration of forest terrestrial laser scanner data, a review of existing approaches in other domains and sensing modalities will provide useful context on the state-of-the-art. Multi-view registration techniques are classified as either sequential, simultaneous, or hybrid. Sequential alignment iteratively registers subsequent pairs of data from an ordered sequence (e.g., A to B, B to C, C to D). Although this has inherent applications to sequential video frames or linear sampling protocols, it is subject to propagation and magnification of errors throughout the sequence (Henning and Radtke, 2008; Pingi et al., 2005; Kim and Hong, 2003; Kang et al., 2000). As a result, simultaneous registration is considered optimal (Bergevin et al., 1996; Blais and Levine, 1995; Jokinen and Haggrén, 1998). Simultaneous, or global registration (Kang et al., 2000; Pingi et al., 2005) utilizes pose estimates between all pairs of scans to minimize the accumulated transformation errors by distributing them throughout the rigid network (Pulli, 1999). Moreover, because the overlap area between all scans is used (as opposed to just a pair), there is a greater potential to identify and utilize tie points that are dispersed throughout the volume, thus improving registration results (Henning and Radtke, 2008). The final class of techniques, hybrid approaches, incorporate both sequential and simultaneous elements.

Graph-theory frameworks, which encode connectivity between overlapping views, have been widely used for multi-view alignment. Typically, a node represents a single input view, sensor, image frame, or point cloud (Kang et al., 2000; Huber and Hebert, 2003). Likewise, an edge represents a connection between nodes, as determined from pairwise registration. A video sequence, for example, would be represented as a predominantly linear graph due to the temporal adjacency of neighboring frames (Kang et al., 2000). Associated with each edge is a relative pose (Huber and Hebert, 2003). Typically, a reference node is chosen as the world coordinate system (WCS). The absolute pose between two pairwise-disconnected views then can be determined by composing the relative poses associated with each edge along a path connecting the scan to the
A potential solution to multi-view registration exists when the graph is connected (i.e., a composite transformation path exists between each node and the reference) (Huber and Hebert, 2003). A minimal solution is defined by a spanning tree, which is a connected graph with no cycles. Additional edges introduce cycles in the graph, which may result in pose conflict due to the composition of pairwise transforms along different paths between views (Huber and Hebert, 2003). These pose inconsistencies are caused by small errors in pairwise pose estimates, which are accumulated through a graph path.

The network of pairwise correspondences are encoded in an adjacency matrix. An unweighted adjacency matrix assigns each \((i, j)\) element a value of “true” or “false”, based on the existence of an edge connecting nodes \(i\) and \(j\). To the contrary, a weighted adjacency matrix provides a value or weight associated with each edge, i.e., based on image correlation (Kang et al., 2000), geometric distance (Kang et al., 2000), spatial overlap (Pingi et al., 2005), tie point registration error (Bendels et al., 2004), or the number of corresponding feature pairs within fixed thresholds (Stamos and Leordeanu, 2003). Graphs are typically undirected, i.e., no directional information is encoded between edges. Many previous studies have demonstrated the utility of graphs to perform sequential, simultaneous, and hybrid multi-view alignment.

Sequential alignment avoids the issue of pose conflict by finding the minimal spanning tree that connects all nodes to the reference node, using an assessment of “minimum path length”, performed on the adjacency matrix. Because the spanning tree is acyclic, the sub-graph is guaranteed to be pose-consistent (Huber and Hebert, 2003). For example, Stamos and Leordeanu (2003) defined pairwise edge weights according to the number of corresponding line pairs detected in urban point clouds, and then used Dijkstra’s algorithm (Dijkstra, 1959) to perform sequential alignment of each node into the WCS, based on a weighting of correspondence pairs. A sequential shortest path technique has also been applied in urban image mosaicing domains (Kim and Hong, 2003; Kim and Hong, 2006; Kang et al., 2000; Bendels et al., 2004). This sequential alignment strategy utilizes all possible pairwise connections, while rejecting weak fits, but does not solve the correspondence problem simultaneously (Stamos and Leordeanu, 2003), and does not exploit the
redundant information provided by multiple edges in order to reduce registration error.

Graphical methods also have been applied to simultaneous multi-view registration, often by linear optimization of pose parameters to minimize registration error. For example, Huber and Hebert (2003) performed multi-view point cloud registration of manmade objects by building a subgraph containing only correct pairwise matches. Global consistency was used to eliminate bad pairwise matches, with absolute poses adjusted to minimize distance between manmade surfaces. Eo et al. (2012) applied generalized procrustes analysis (GPA) to simultaneously adjust registration and found favorable results compared to sequential registration with ICP in urban point cloud data.

A third class of “hybrid” registration algorithms includes elements of both sequential and simultaneous registration. In order to reduce propagation of alignment errors and exploit multiple cycles through the graph, additional steps may be added to sequential registration, such as global averaging (Sharp et al., 2004) or cycle detection (Kang et al., 2009; Borrmann et al., 2008), which distribute errors across the path sequence. For example, Kang et al. (2009) collected data providing circular self-closure (i.e., A to B, B to C, C to A) in order to redistribute errors through the “cycle”. For range data collected from autonomous robots, Borrmann et al. (2008) performed loop detection in order to detect closed edge cycles in the sequential graph. For arbitrary sampling protocols, Sharp et al. (2004) decomposed the graph into a series of closed cycles, so that nonlinear optimization could be performed over each basis cycle in closed form. Cycles were then reintegrated into a global model using an averaging technique. This afforded the advantages of sequential registration, while adding a secondary “global” pose adjustment to minimize error.

Registration techniques that leverage global consistency to remove erroneous local matches and reduce propagation errors, i.e., either simultaneous or hybrid approaches, are preferred for multi-view registration, although there is still no consensus as to the best approach (Sharp et al., 2004). This research presents a hybrid multi-view registration approach for blind, marker-free registration of forest terrestrial laser scanner data, and compares it against standard sequential and simultaneous registration approaches. This work builds on Kelbe et al. (2015a) by providing an automatic, blind, marker-free, multi-view registration of a network of TLS scans collected at
arbitrary locations within a forest plot. A primary contribution is the integration of embedded error metrics associated with each pairwise transformation (Kelbe et al., 2015a), which simplifies the global alignment problem, while adding resistance to noise (Huber and Hebert, 2003). Previously, a variety of heuristics have been used to assess the quality of a pairwise “edge”, including image correlation (Kang et al., 2000), geometric distance (Kang et al., 2000), spatial overlap (Pingi et al., 2005), tie point registration error (Bendels et al., 2004), or the number of corresponding feature pairs within fixed thresholds (Stamos and Leordeanu, 2003). However, none have provided adequate information on the precision of a pairwise transformation. We extend the pairwise error metric of Kelbe et al. (2015a) to provide an embedded multi-path confidence metric associated with each transformation model. Specific objectives are to (i) validate the proposed embedded error metric for multi-edge paths through a graphical network, (ii) assess the performance of the proposed, hybrid multi-view registration technique for TLS data collected in New England forests, and (iii) demonstrate the improvement of plot-level inventory assessment compared to single-scan data collection.

5.4 Methods

5.4.1 Background

For a typical stationary TLS system, a pulsed laser beam is rapidly emitted into the scene in a radial pattern based on the deflection by a rotating mirror. This scanning in elevation angle is coupled to azimuthal platform rotation to sample nearly the full sphere, except for a small occlusion cone below the instrument. For each scan mirror angle, $\theta$, and rotation stage position, $\phi$, the return trip travel time of a laser pulse is digitized and converted to range, $r$, based on the speed of light. This gives an unambiguous triplet ($\theta$, $\phi$, $r$) for each digitized pulse. A three-dimensional (3-D) point cloud, $P$, is the aggregate of all digitized range measurements. The reader is referred to Bachman (1979) for additional description of laser ranging or more commonly, light detection and ranging (lidar) principles. Upon conversion from spherical to cartesian coordinates, we define the point cloud $P = \{x_1, \ldots, x_i, \ldots, x_n\}$ where $x \in \mathbb{R}^3$ is the $x, y, z$ position for the $i^{th}$ point in $P$. 
5.4. METHODS

5.4.2 Study Area

To assess the feasibility of multi-view point cloud registration in the forest environment, we assessed the error registration for eight forest plots spanning a diverse range of structural complexity. The study area for this work corresponded to the National Ecological Observatory Network (Kampe et al, 2010) (NEON) core ecological site for the Northeastern ecological domain, including both Harvard Forest and Quabbin Reserve, Massachusetts, USA (bounded by 42.428° N, 72.284° W and 42.558° N, 72.170° W; WGS1984). Harvard Forest is a 1,200 ha reserve with a long history of ecological research and management. Quabbin Reserve is a 23,000 ha public forest and provides additional diversity, via various disturbance regimes (Winkler et al., 2010).

Eight 20 × 20 m plots were selected in this region, representing a diverse range of Northeastern USA forest structure, and were selected using a stratified random sampling scheme and evaluated in person to ensure forest structure variability. Individual plot-level characteristics are shown in Table 5.1. Plots include a range of age, densities, structural complexities, and species compositions. Basal area (BA) ranged from 40.96 m$^2$ · ha$^{-1}$ to 66.38 m$^2$ · ha$^{-1}$. Stem densities were recorded separately for stems of DBH ≥ 10 cm and for stems of DBH < 10 cm, and ranged from sparse and mature (700 stems/ha, all of which were ≥ 10 cm DBH) to dense and young (1300 stems·ha$^{-1}$; 84% with DBH ≥ 10 cm and 16% with DBH < 10 cm). Ground vegetation characteristics varied from bare ground to 70% coverage of 0.7 m tall ground vegetation. Lidar images of a sample of the plots are shown in Figure 5.1. The structural variability represented by our study area is unique among previous research and provides a diverse data set from which to evaluate the ability of TLS for operational forest inventory.
Figure 5.1: Equal-angle projection images of lidar intensity values for several example forest sites. Example sites represent a diverse range of Northeastern USA forest characteristics, including (a) terrain variation, (b) sub-canopy branches and foliage, and (c) ground vegetation.

Table 5.1: Summary of ground validation plots in Harvard Forest, MA, USA.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Dominant Species (Secondary Species)</th>
<th>BA m² ha⁻¹</th>
<th>Stems ha⁻¹</th>
<th>Ground veg. z [m]</th>
<th>p [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td><em>Acer saccharum</em> - <em>Pinus strobus</em> / <em>Acer pensylvanicum</em></td>
<td>47.83</td>
<td>1075 200 0.5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>04</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>41.19</td>
<td>1100 175 0.5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>07</td>
<td><em>Pinus strobus</em> - <em>Quercus</em> (alba, rubra, velutina)</td>
<td>42.19</td>
<td>1100 200 0.5</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>08</td>
<td><em>Pinus strobus</em> - <em>Tsuga canadensis</em></td>
<td>55.70</td>
<td>600 100 0.5</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td><em>Pinus strobus</em></td>
<td>53.84</td>
<td>950 0 0.5</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td><em>Tsuga canadensis</em> - <em>Betula alleghaniensis</em></td>
<td>51.56</td>
<td>700 0 0 0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Variant: <em>Tsuga canadensis</em> with mixed</td>
<td>66.38</td>
<td>950 125 0.6</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td><em>Quercus rubra</em> - (<em>Acer saccharum</em>)</td>
<td>40.96</td>
<td>525 125 0.7</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>

For each plot, 25 scans were collected in a nominal grid pattern with 5 m spacing between scans (5 × 5 scans). Plots were first laid out based on a center GPS coordinate, compass, and tape measure. This regular sampling method was maintained to ensure consistent, objective data.
coverage. Knowledge of this initial pose information and the regular pattern of data collection, however, was not used a priori in the development of this algorithm, so that the multi-view registration technique developed in this paper is blind, independent of initial scanner pose.

Our experimental design was geared to evaluate multi-view registration by aligning each of the 25 scans collected within a forest plot into a common coordinate system. Data were collected during August 2012 leaf-on conditions. Data collection during leaf-off conditions is preferable for measurement of woody structure, in that it reduces the effects of occlusion due to foliage (Dassot et al., 2012; Raumonen et al., 2013; Côté et al., 2009). However, our objective was to evaluate TLS as an operational forest inventory tool. Since some parameters, e.g., leaf area index (LAI), are relevant only during the growing season, leaf-on data collection allowed us to better evaluate the feasibility of the approach in context of typical forestry and ecological needs.

The instrument used in this study was the “Canopy Biomass Lidar (CBL)”, a low-cost mobile terrestrial lidar system integrated from commercial-off-the-shelf (COTS) components by Rochester Institute of Technology and later upgraded by the University of Massachusetts Boston. The system is based on a design first implemented by a team at the Katholieke Universiteit Leuven, Belgium (Van der Zande et al., 2006). A SICK LMS-151 laser scanner (SICK, 2009) pulses a 905 nm laser at 27 kHz, with range measurement recorded based on time-of-flight. The emitted laser pulse is deflected by a rotating mirror and coupled to azimuthal platform rotation in order to sample the full hemisphere above the instrument and a portion of the hemisphere below (270° V x 360° H coverage). The instrument is tethered to a data logger and battery, which are mounted on a backpack and worn by the operator. Sensor control is achieved via a wireless mobile application. This instrument was designed to overcome the limitations of high cost, low mobility, and long scan time, which have so far precluded operational forest structure assessment using terrestrial lidar. However, unlike many commercial scanners that provide for high-density point cloud data, this system has a minimum angular step-width of 4.36 mrad, and a beam divergence of 15 mrad, both approximately two orders of magnitude coarser than commerical TLS instrumentation typically used in forest inventory assessments (see Table 1 from (Kelbe et al., 2015b)). These limitations provide an opportunity to develop registration algorithms that are robust to low-resolution data,
while focusing on rapid, operational inventories. Our instrument’s mobile platform allowed a single operator to traverse between plots without disassembly. The total time to characterize a 20×20 m plot with 25 scans was approximately 30 minutes for a single operator. This includes 30 seconds per scan duration and an additional 15 seconds required to move between scan locations.

5.4.3 Methods Problem Statement

The task of multi-view registration is to bring a set of sensors, $\mathcal{S}$, into spatial alignment. We make no assumptions on the a priori sensor locations, nor the overlap between sensor views. Moreover, the list of sensor views is unordered, meaning that consecutive views do not imply spatial adjacency.

We follow the precedent from previous research and separate the multi-view registration process into a two stages, as shown in Figure 5.2. The first stage (pairwise registration) takes as input point cloud data for pairs of sensor views, and computes the relative transformation between sensor pairs. Details of the pairwise registration algorithm can be found in (Kelbe et al., 2015a), which outlined a blind, marker-free registration of forest terrestrial laser scanner data pairs using view-invariant tie points. Briefly, input point clouds were first reconstructed to model individual tree stems from single scans (Kelbe et al., 2015b), from which view-invariant tie points were extracted as the intersection of the tree stem axis and terrain. Geometric similarity was then employed to constrain an iterative voting method, which determined the three rotation and three translation parameters for a rigid transformation using singular value decomposition (SVD). An embedded upper-bound root mean square error (RMSE) metric was provided with each transformation output, by traversing a closed loop along a forward and reverse path through disjoint tie point sets. This provided explicit assessment of the error associated with each pairwise transformation. This information has the potential to simplify the multi-view registration, while improving resistance to noise (Huber and Hebert, 2003).

The second stage (hybrid graphical registration) takes as input the set of relative pairwise transformations between all scan pairs, and their associated errors, and determines the absolute transformations, which align each scan, $j$, into a world coordinate system (WCS), where the WCS
is defined as the \( w \)th local coordinate system, \( w \). Details of the multi-view registration are provided below.

---

**5.4.4 Input**

Consider \( M \) scans from sensors \( S_1 \ldots S_i \ldots S_M \). Each scan records points in a local coordinate system, LCS\(_i\), i.e., a coordinate system with its origin and axes defined by the pose of instrument \( i \). The collection of points from the \( i \)th sensor is the \( i \)th point cloud, \( P_i \). View-invariant tie points, \( T_i \), are extracted from each \( P_i \) as the unambiguous 3-D locations of the stem-terrain intersection points (Kelbe et al., 2015a), obtained from modeled geometry (Kelbe et al., 2015b).

---

**5.4.5 Pairwise Registration**

After detecting view-invariant tie points, pairs of scans were aligned using an efficient voting method based on geometric constraints (Kelbe et al., 2015a). With \( M = 25 \) scans within each plot, there are \( \binom{M-25}{2} = 300 \) possible pairwise combinations. For each pair \( (i = 1 \ldots M, j = 1 \ldots M) \) the rigid transformation parameters, \( R_j^i = [\theta_x, \theta_y, \theta_z] \), and \( t_j^i = [t_x, t_y, t_z] \) are determined, which map scan \( j \) into scan \( i \) in a least squares sense using singular value decomposition (SVD) (Kelbe et al., 2015b; Kelbe et al., 2015a). In some cases, however, there is no view overlap between \( S_i \) and \( S_j \).
due to range, occlusion, etc. Or, correspondences between tie points are incorrect, resulting in an erroneous output transformation. It therefore is desirable to have an embedded confidence metric, which can simplify the removal of incorrect matches.

Various metrics have been identified to assess the quality of an output registration, including image correlation (Kang et al., 2000), geometric distance (Kang et al., 2000), spatial overlap (Pingi et al., 2005), surface consistency (Huber and Hebert, 2003), registration error between feature points (Bendels et al., 2004), and the number of corresponding feature pairs (Stamos and Leordeanu, 2003). Huber and Hebert (2003) noted that pairwise registration error metrics, such as those above, may not provide satisfactory performance criteria. This complicates the registration process, because falsely matched feature points can result in globally inconsistent transformations (Huber and Hebert, 2003). Absolute knowledge of the pairwise registration accuracy, therefore, could greatly simplify the multi-view registration, while improving resistance to noise (Huber and Hebert, 2003).

To address this, Kelbe et al. (2015a) developed a circular self-closure framework, where a transformation path is traversed forward and backward between disjoint tie point subsets for a scan pair. The transformations associated with each path were applied in sequence, such that by construction, \( R_{true} = I \) (the identity matrix) and \( t_{true} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \). For each pairwise transformation, the researchers then transformed the initial tie point set through this circular path, bringing the tie points back to their original coordinate system. A perfect transformation maps each point back to its original location. Errors in the transformation introduce deviations, which can be computed on a per-element basis. This reported RMSE metric, \( e \), provided an embedded upper bound estimate of the precision of each pairwise transformation. In this research, we extended this construction to a graphical framework in order to simplify the multi-view registration.
5.4. METHODS

Figure 5.3: Pairwise registration determined the rigid transformation that aligned sensor $S_j$ into $S_i$ for all $i, j$. Associated with each registration are the output transformation parameters, $R_{ij}^j$ and $t_{ij}^j$, and an embedded upper-bound error metric, $e_{ij}^j$. Note that some paths are missing due to insufficient corresponding tie points for a rigid transformation, and some paths may have high error.

5.4.6 Graph Initialization

From the network of pairwise correspondences in Figure 5.3, a model graph, $G$, was constructed. In the computational sciences and mathematics domain, a “graph” is a structural framework, which encodes connectivity and relationships between different objects. The concept of a graph can be understood with the following example (Figure 5.4). Consider three people sitting in their cubicles at work. John and Bill are in cubicles next door to each other, and Mike is across the hall. Where John is sitting, he can look back and see Mike in the cubicle across the hall. But Bill, who is in the cubicle next to John, is occluded from John’s view by the dividing partition. John, Mike, and Bill can only communicate visually, requiring a line of sight (an “edge”, or connection) between two communicative parties. Although John can’t see Bill directly, John can see Mike, who in turn can see Bill. John can ask Mike for information about Bill; for example, his shirt color, which way
he is facing, or his location. In this way, he can obtain information about Bill indirectly through Mike, who serves as a “link” in the communication chain between John and Mike.

![Figure 5.4](image)

Figure 5.4: The concept of a graph can be illustrated by three people sitting in their cubicles at work, in a hypothetical scenario where only nonverbal communication is possible, requiring a line of sight between two parties. Bill is occluded from view of John, preventing their communication. However, John can obtain information about Bill via Mike, who serves as a “link” in the communication chain.

When there are multiple “nodes”, or people in the chain, this framework is reminiscent of the childhood game, “telephone”, where a secret message is whispered from one person to the next. The amusing point of this game is that small communication errors between pairs are compounded
over the transmission sequence, resulting in a wildly ridiculous final message. However, we can
generalize the concept even further. The childhood “telephone” game represents what is called
a sequential graph, which is a simplified framework, where a message travels linearly, without
repetition, from A to B to C to D, etc. A more complicated game can be envisioned, which
allows for messages to be communicated arbitrarily between any two pairs of people. In this
case, different communication errors are accumulated depending on the path sequence through
which the message was communicated. This is illustrative of the concept of pose conflict, wherein
competing messages are received by each person, or node, in the graph. A central challenge of
graph theory is to rectify the many competing messages, in order to determine the best estimate
of the true message. In the context of TLS data registration, we wish to rectify the competing
position and orientation estimates obtained from the network of pairwise correspondences.

These concepts can be formulated mathematically, as follows. A weighted, undirected graph,
$G$, is defined as the collection of vertices, $V$, edges, $E$, and weights, $W$. This is visualized in Figure
5.5.

$$G = \{V, E, W\} \quad (5.1)$$

Let the vertex set, $V$, be the collection of scans, $\mathcal{S}_{i=1...M}$, such that each scanner view, $\mathcal{S}_i$, represents
a vertex in the graph (Equation 5.2). Pairwise connections are represented as edges, $E$, which
contain the rigid transformation parameters, i.e., the rotation, $R$, and translation, $t$, associated
with each edge (Equation 5.3), and the weights, $W$, of each edge determined from the embedded
confidence metric outlined in Section 5.4.5 (Equation 5.4).

$$V = \{\mathcal{S}_{i=1...M}\} \quad (5.2)$$

$$E_{i,j} = \{R_{ij}, t_{ij}\} \quad (5.3)$$

$$W_{i,j} = \epsilon_{ij} \quad (5.4)$$
Recall that the minimal solution is defined by a spanning tree, which is a connected graph with no cycles. Additional edges introduce cycles in the graph, which may result in pose conflict due to the composition of pairwise transforms along different paths between views (Huber and Hebert, 2003). These pose inconsistencies are caused by small errors in pairwise registration estimates. As a result, the aggregate network of multiple pairwise correspondences may enforce/contradict each other in a global, rigid 3-D transformation model.

5.4.7 Dijkstra Spanning Tree

The first step of the proposed multi-view registration approach is drawn from sequential registration techniques, which commonly use Dijkstra’s algorithm (Stamos and Leordeanu, 2003; Kim and Hong, 2006) or other shortest-path techniques (Bendels et al., 2004; Kim and Hong, 2003; Kang et al., 2000). This is used to compute the minimal spanning tree (MST) between a reference node, $w$, and all other nodes in a graph, based on the weights associated with each edge (Dijkstra, 1959). Absolute poses can then be determined by composing transformations associated with each path edge.

Although this sequential registration approach is straightforward and resistant to noise (Huber
and Hebert, 2003), it does not take advantage of redundancy within the graph in order to optimize fit (Stamos and Leordeanu, 2003). Therefore, we implemented a hybrid registration approach, which includes additional edges in the global transformation network. This was done as follows: Instead of determining a single spanning tree, all \( M \) scanning trees were computed (Section 5.4.7), and then aligned into a common coordinate system (Section 5.4.8). The redundant information encoded in the model was then exploited to determine consensus-based, absolute poses for each node, using inverse distance weighting (IDW) (Section 5.4.9). This effectively reduces the effect of erroneous transformation parameters by averaging.

Mathematically, from the graph, \( G \), we generated a weighted adjacency matrix, \( A \), with weights \( A(i, j) = W_{i,j} \). The shortest path from each node \( S_j \) to \( S_w \) was computed from \( A \) using Dijkstra’s algorithm (Dijkstra, 1959). This resulted in \( M \) minimal spanning forests, \( F_1 \ldots F_M \), as is illustrated in Figure 5.6. Note that in data science, the name “spanning forest” is used when allowing for a disjoint graph.
A minimal spanning forest is simply a subgraph of $G$, which contains a list of the optimal paths that link each node to the reference node. Formally, we define the spanning forest, $F$, as the set of vertices, $V$, edges, $E$, and effective aggregate path weights, $\Psi$, where the edge set includes the set of paths, $P$, effective rotations, $\hat{\Omega}$, and effective translations $\hat{\tau}$:
\[ \mathcal{F} = \{V, E, \Psi\} \quad (5.5) \]
\[ E = \{P, \hat{\Omega}, \hat{\tau}\} \quad (5.6) \]

Note that there are \( M \) spanning forests for each possible reference node, i.e., \( \mathcal{F}_i \) is the subgraph associated with reference node \( i \). Moreover, we follow the previous convention and let the subscript of the effective parameters designate the source scan, and likewise, let the superscript designate the target scan. Thus, the effective rotation, \( \hat{\Omega}_j^i \), and effective translation, \( \hat{\tau}_j^i \), are the transformation parameters that map \( j \) \( \rightarrow \) \( i \) via the appropriate path, \( P_j^i \). Let the path, \( P_j^i \), contain the ordered list of nodes that are traversed from node \( j \) to reference node \( i \), as determined from Dijkstra’s algorithm. We computed the estimated effective rotation matrix, \( \hat{\Omega}_j^i \), and the estimated translation vector, \( \hat{\tau}_j^i \), by composing the pairwise transformations in sequence along this path. In other words, for a path of length, \( k \), with \( P(1) = j \) and \( P(k) = i \),

\[ \hat{\Omega}_j^i = R_{P(k-1)}^{P(k)} \cdots R_{P(2)}^{P(3)} \cdot R_{P(1)}^{P(2)} \cdot \hat{\theta}_{P(1)} \quad (5.7) \]
\[ \hat{\tau}_j^i = t_{P(k-1)}^{P(k)} + (R_{P(k-1)}^{P(k)} \cdot t_{P(k-1)}^{P(k-2)}) + \cdots \quad (5.8) \]

Moreover, we extended the pairwise RMSE metric, \( \epsilon \), in order to compute an estimate of the aggregate path RMSE, \( \Psi \), which was then encapsulated in the graphical model as the edge weight. The aggregate path RMSE, \( \Psi \), was defined as the sum of RMSE values associated with each pairwise path segment.

\[ \Psi_j^i = \epsilon_{P(k-1)}^{P(k)} + \cdots + \epsilon_{P(2)}^{P(3)} + \epsilon_{P(1)}^{P(2)} \quad (5.9) \]

Additional relationships were explored to compute aggregate the RMSE, including the product, maximum value, and geometric mean of \( \epsilon_j^i \); however, the additive model (Equation 5.9) was found to be the best predictor of the true RMSE.
The embedded RMSE metric from Kelbe et al. (2015a) provided a robust, first-order approximation of the error associated with each path. This metric improved the ability to find the best spanning tree and represents a novel improvement over existing approaches. However, in some cases errors may remain, for example, due to symmetries in the data, e.g., regularly-spaced forest plantations (Huber and Hebert, 2003), that complicate the unique identification of feature points. As a result, locally consistent but globally incorrect matches may introduce conflict in the global transformation model.

### 5.4.8 Align Spanning Trees

To address these potential inconsistent matches, we exploited the mutual information, i.e., conflict, encoded in the network of pairwise poses in order to reduce the global registration error. This was done by aligning all $M$ spanning forests, $F_i$, into a common WCS. For numerical reasons, we define the WCS as the LCS$^w$ that minimizes the sum of the edge weights, e.g., estimated RMSE values, from $F_w$ to all other nodes (Bendels et al., 2004).

$$w = \arg \min _i \sum _{j=1} ^{M} \Psi _{ij}$$  \hspace{1cm} (5.10)

To align all $M^2$ effective pose parameters ($M$ nodes with $M$ reference nodes) into the WCS, the effective rotations, $\hat{\Omega}$, and translations, $\hat{t}$, were transformed into the WCS by applying the pairwise transformations, $R$ and $t$, associated with the edge from $i \rightarrow w$.

$$\hat{\Omega} _{iw} = R _{iw} \cdot \hat{\Omega} _{ij}$$  \hspace{1cm} (5.11)

$$\hat{t} _{iw} = R _{iw} \cdot \hat{t} _{ij} + R _{iw} \cdot t _{iw}$$  \hspace{1cm} (5.12)

Note that the superscript, $iw$, indicates that the scanner pose was transformed into the target coordinate system, $w$, via reference node, $i$. Likewise, the pairwise RMSE, $e _{ij}$, was added to the effective RMSE, $\Psi$, in order to determine the effective aggregate RMSE from $j \rightarrow i \rightarrow w$.

$$\Psi _{ij} ^{iw} = e _{ij} ^{iw} + \Psi _{ij}$$  \hspace{1cm} (5.13)
This alignment is visualized in Figure 5.7, where the reference node is shaded dark red. The $\mathcal{M}^2$ effective transformations are represented by black edges, with the underlying pairwise edges included in light gray. From this figure, the advantage of exploiting pose inconsistency in the graphical model becomes apparent. Different paths throughout the graph result in deviations in the reported effective pose estimates. In the next section, we determined the estimated effective poses for each node, by averaging the reported transformation parameters weighted by the inverse of the associated edge weight, $\Psi_{ij}^{iw}$ (Section 5.4.9).

![Diagram of network with nodes and edges](attachment:image.png)

Figure 5.7: Different paths throughout the network introduce deviations in the estimated output pose. We exploit this redundant information by averaging conflicting pose estimates weighted by the inverse aggregate RMSE, $\Psi$ (Section 5.4.9).

### 5.4.9 Compute Effective Pose

From Figure 5.7, we observed that traversing the graph along different edges introduced deviations in the estimated effective pose. We exploited this redundant information by averaging the $\mathcal{M}$ pose estimates for each node, in order to increase pose estimate precision. An inverse distance weighting (IDW) interpolation was used to perform a weighted average on the pose parameters, such that pose estimates with higher estimated error, i.e., $\Psi$, had less impact on the output model. For each of the six rotation parameters, $p = \{\theta_x, \theta_y, \theta_z, t_x, t_y, t_z\}$, the weighted parameter estimate,
\[ \hat{p}_i = \frac{\sum_{j=1}^{M} f(\Psi_j)p_i}{\sum_{j=1}^{M} f(\Psi_j)} \] (5.14)

An exponential weighting function of the form, \( f(x) = e^{-5x} \), rather than \( f(x) = 1/x \) was used to reduce the impact of high-RMSE paths. The exponential factor, 5, was chosen such that an RMSE of 1 m has < 1% impact on the weighted average. The result is the set of M effective rotations, \( \hat{\Omega}_j^w \), and translations, \( \hat{\tau}_j^w \), which map scanner \( j \rightarrow w \) (Figure 5.8). Using \( \hat{\Omega} \) and \( \hat{\tau} \), we can then transform the point cloud, \( P_j \), into the WCS,

\[ \hat{P}_w = \hat{\Omega}_j^w \cdot P_j + \hat{\tau}_j^w \] (5.15)

Figure 5.8: Output absolute poses are estimated by weighted averaging. An embedded aggregate path RMSE, \( \Psi \), provides an upper-bound estimate of the confidence associated with each registration.

### 5.4.10 Compute Embedded Error

We extended the circular self-closure framework presented in Kelbe et al. (2015a) to determine an estimate of the error associated with each global transformation, from \( j \rightarrow w \). The aggregate path
error was described briefly in section 5.4.7 (Equation 5.9). Here we elaborate on the circular path construction in the context of multi-sensor registration.

For each ordered pair of nodes \( \mathcal{S}_i, \mathcal{S}_j \), the tie points, \( \mathcal{T}_i \), are split into two disjoint sets, \( \mathcal{T}_i^A \) and \( \mathcal{T}_i^B \). From these disjoint tie point sets, we determined the pairwise transformation parameters from \( j \to i \) through disjoint set \( A \), and the corresponding parameters associated with the reverse path, \( i \to j \), through disjoint set \( B \). This is visualized in Figure 5.9, where the nodes are symbolically split into two disjoint graphs, \( D^A \) and \( D^B \) (shaded red and blue). Moreover, there are two unique edges associated with the forward (red) and reverse (blue path) from each node \( i \) to node \( j \). To compute an estimate of the per-edge error for each pair of nodes \((i, j)\), we traversed a circular path from \( i \to j \to i \), where the transformation parameters associated with the path from \( i \to j \) are taken from \( D^B \) and the transformation parameters associated with the return path from \( j \to i \) are taken from \( D^A \). The functional relationships associated with the circular path were then applied in sequence, where we know by construction that \( R_{true} = I \) and \( t_{true} = [0 \ 0 \ 0]^T \). An upper estimate of the confidence associated with a single pairwise edge was determined by computing the RMSE between the tie points from \( i \) before and after transformation through the circular path, i.e., back to the source coordinate system (See Equation 28 from (Kelbe et al., 2015a)). This framework was extended to multi-view registration by summation along the path. In other words, the aggregate RMSE for a path consisting of multiple edges, e.g., \( \mathcal{S}_6 \to \mathcal{S}_4 \to \mathcal{S}_3 \) in Figure 5.9, was computed by summation of the RMSE values for each pairwise segment within the path (See Equation 5.9).
Figure 5.9: A circular self-closure framework between disjoint sets (shown symbolically as red and blue) provides an embedded error metric. The pairwise error metric, $e^j_i$, is obtained by composing the transformations associated with an edge’s forward (red) and reverse (blue) path, and computing the RMSE between the input tie points and transformed tie points through the circular path. Aggregate path errors were determined as the summation of the errors associated with each path edge, e.g., $\Psi^3_6 = e^4_6 + e^3_4$.

5.4.11 Verification Analyses

To assess the validity of the proposed confidence metric (objective i) and to quantify the performance of the proposed hybrid registration approach (objective ii), several experiments were designed where truth data were inherently known. For all experiments, truth tie points were generated in the WCS from a uniform distribution with limits specified by nominal plot boundaries, e.g., a $60 \times 60 \times 2$ m volume with origin at (0,0,0) and orientation defined by the Cartesian axes. A constraint was added such that the minimum allowable distance between tie points was 2 m. Tie point locations were attributed a truth diameter at breast height (DBH) value drawn from the uniform distribution, $U(10 \text{ cm}, 50 \text{ cm})$.

$M=25$ “sensors” were then placed within the scene to record a sample of the full set of true tie points in each scanner’s local coordinate system (LCS). Scanners were placed in a nominal grid pattern with 10 m spacing within the center $40 \times 40$ m of the volume. A uniform random deviation from the interval $(x \pm 3, y \pm 3, z \pm 1 \text{ m})$ was added to the nominal sensor position, and a uniform random deviation from $(\theta_x \pm 20^\circ, \theta_y \pm 20^\circ, \theta_z \pm 180^\circ)$ was added to the nominal orientation. This
generated random scanner poses representative of a nominal sampling protocol: representatively spaced throughout the volume, with the scanner head roughly level, and with arbitrary azimuthal orientation.

To generate tie points as they would be observed by each sensor, i.e., in the LCS, each sensor recorded a subset of the full tie point set, constrained by a maximum range drawn from the uniform distribution $U(15 \text{ m}, 30 \text{ m})$ and randomly subset to $U(20, 30)$ points. Detected tie points in the WCS were then transformed into the LCS, for each sensor $i$ using the $i^{th}$ true sensor pose. Let the true tie points in each LCS be $\tilde{T}_i$. Noise was then added to the true tie point locations, $\tilde{T}_i$, from a zero-mean normal distribution, $N(0, \sigma)$ to simulate the error in detecting the true tie point location, which may be due to factors such as terrain obscuration, non-symmetric stem cross-sections, sensor noise, or occlusion (Kelbe et al., 2015a). Let the observed tie points (i.e, with noise added) in each LCS be $T_i$, where $T_i = \tilde{T}_i + N(0, \sigma)$. 100 trials were performed with standard deviation, $\sigma$, ranging from $10^{-5}$ m to 0.5 m. A random error was also added to the tie point diameters, drawn from a uniform distribution ranging from 0 - 50% diameter difference. This construction allowed us to perform registration on the noise-added tie point data, while retaining absolute truth information for subsequent evaluation.

The first experiment assessed the validity of the proposed confidence metric for each graphical registration technique. For each $i^{th}$ sensor pose, the observed tie point set, $T_i$, was transformed into the WCS, $T_i^\text{w}$, using the estimated rigid transformation parameters. Note that we continue the notation designating source pose with a subscript, and target pose with a superscript. The associated reported RMSE, $\Psi_i^\text{w}$, was computed between $T_i^\text{w}$ and $T_w$. Likewise, for each $i^{th}$ sensor pose, the same rigid transformation parameters were used to transform the true tie point set, $\tilde{T}_i$ (no noise added) into LCS$_w$. The true RMSE, $\tilde{\Psi}_i^\text{w}$, was calculated between $\tilde{T}_i^\text{w}$ and $\tilde{T}_w$. We then plotted $\Psi, \tilde{\Psi}_i$ for all $i$ (Figure 5.11) to evaluate whether $\Psi$ was a good predictor of $\tilde{\Psi}$.

The proposed graph-based confidence metric, $\Psi$, is an aggregate sum of the pairwise confidence metrics, $e$, associated with an edge. To uncouple these additive effects and provide additional insight into the underlying mechanisms of the graph-based confidence metric, we performed a second experiment, which assessed the validity of the pairwise confidence metric in the
same manner as above. Although Kelbe et al. (2015a) demonstrated the usability of this metric to first order using receiver operator characteristic (ROC) curves, a study on the precision of this metric provides additional insight. For each pairwise combination of source pose, $i$, and target pose, $j$, the observed tie point set $T_j$ was transformed into $LCS_i$ via the estimated rigid transformation parameters, yielding $T^i_j$. The associated reported RMSE, $e^i_j$, was computed between $T^i_j$ and $T_j$. Likewise, for each $i^{th}$ sensor pose, the same rigid transformation parameters were used to transform the true tie point set, $\tilde{T}_j$ (no noised added) into $LCS_i$, yielding $\tilde{T}^i_j$. The true $\tilde{e}^i_j$ was calculated between $\tilde{T}^i_j$ and $\tilde{T}_j$. We then plotted $e$ vs. $\tilde{e}$ for all $i, j$ combinations (Figure 5.12), to evaluate whether $e$ was a good predictor of $\tilde{e}$.

Next, we quantified the performance of the graph-based registration technique (objective ii), and compared it to a standard sequential graphical approach and a standard simultaneous graphical approach. The third experiment assessed the impact of the true RMSE on the input RMSE ($RMSE_{in}$) added to the tie points. Error in precisely localizing the 3-D points, e.g., due to ground vegetation, branches, or tree stems that are not radially symmetric, may introduce deviations in the input tie point sets. These deviations enforce a limit in the output precision of registration, because a perfect transformation between scan pairs will still yield an RMSE equivalent to the RMSE inherent in the tie points. We hypothesized that the true RMSE would be limited primarily by this input deviation, $RMSE_{in}$. To assess this, 100 trials were performed with zero-mean normal random noise added to the tie point locations, and with a $\sigma$ logarithmically increased from $10^{-5}$ m to 0.5 m. For each trial, the $RMSE_{in}$ was calculated from the observed tie point locations. The true error, $\Psi$, was then determined and plotted vs. $RMSE_{in}$ (Figure 5.13).

The final experiment assessed the performance of the graphical registration compared to pairwise registration. While the graphical registration is superior simply in its ability to link disconnected scans to a reference node via connected paths, it would be advantageous if the redundant information encoded within a graphical network could be used to improve the output RMSE of associated paths as well. To assess this, the true error of a graphical path, $\Psi^w_i$, was plotted vs. the corresponding true error of the pairwise path, $\tilde{e}^w_i$ (Figure 5.14).

For all analyses, we compared the proposed hybrid approach against a standard sequential
and simultaneous registration approach. The sequential registration approach was implemented by determining the single MST using Dijkstra’s algorithm, with edge weights determined from the embedded pairwise confidence metric. The simultaneous registration technique implemented in this study used SVD to align all $M$ spanning forests, $\mathcal{F}_i$, following the geometry-constrained voting approach from Kelbe et al. (2015a). These additional algorithms provided an opportunity to assess performance of the proposed hybrid approach in comparison to other techniques (objective ii).

5.5 Results

5.5.1 Point Cloud

Output registered point clouds are shown in Figure 5.10 for the example plots from Figure 5.1. For visualization, a 40 m $\times$ 5 m transect was cropped from the entire aggregate point cloud. Different scans are distinguished by color. These example plots were chosen to reflect a diversity of structural characteristics, including terrain variation (Figure 5.10a), sub-canopy branching and foliage (Figure 5.10b), and ground vegetation (Figure 5.10c). Insets provide detailed views of fine-scale tree structure.
Figure 5.10: (a) Registered point clouds correspond to the example plots from Figure 5.1, and reflect a diversity of structural characteristics, including (a) terrain variation, (b) sub-canopy branching and foliage, and (c) ground vegetation. Quantitative registration metrics are included in Table 5.2, where we report 92% – 100% of scans successfully connected, with a reported RMSE $= 28 – 42$ cm. Based on the analysis from Figure 5.11, true RMSE values are expected to be $5.4 \times$ lower, i.e., 5-7 cm.
In Table 5.2, quantitative results from the proposed graphical approach are collated and compared to the pairwise registration approach from Kelbe et al. (2015a). The reported graphical RMSE, $\Psi$ [cm], was consistently < 50 cm. However, the reported graphical RMSE metric, $\Psi$, was found to be an overestimate of the true RMSE, $\Psi$ (Figure 5.11), with a multiplicative bias of 5.40. As a result, actual RMSE values for pairwise and graphical registration are expected to be $3.19 \times$ and $5.40 \times$ lower, respectively, than the reported values. The estimated true RMSE was included for the average of all plots in the final row of Table 5.2, by adjusting for the multiplicative bias determined from Equations 5.17 and 5.19. Average estimated graphical registration error was 7.24 cm after compensation for the multiplicative bias. The graphical approach achieved registration on the order of, or better than the corresponding pairwise registration proposed by Kelbe et al. (2015a) (cf. Figure 5.14).

The primary advantage of the graphical approach is the ability to bring additional scans into alignment, which are otherwise not connected to the reference node via a direct path (see bolded results in Table 5.2). This is especially apparent for plots with significant sub-canopy foliage and branching, which occlude some sensor positions from view of a single reference node.

For example, plot 15 (see Figures 5.1b and 5.10b) was a mixture of Tsuga canadensis and Betula alleghaniensis, with significant occlusion in the corresponding point clouds. Only $13/25 = 52\%$ of scans were linked to the reference node using pairwise registration. The proposed graphical approach, however, resulted in 92% of scans being connected. Likewise, for plot 7, we achieved an improvement from 35% to 100% connectivity.
Table 5.2: Quantitative results from the proposed graphical approach were collated and compared to the pairwise registration approach from Kelbe et al. (2015a). The primary advantage of multi-view registration is the ability to bring additional scans into alignment (bold percentages) by linking disconnected scans to the reference node via an indirect path. Est* is the estimated true RMSE after adjusting for the multiplicative bias, determined from Equations 5.17 and 5.19.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Pairwise RMSE, ε [cm]</th>
<th>% Connected</th>
<th>Graphical Approach RMSE, Ψ [cm]</th>
<th>% Connected</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.25 20/25 = 80%</td>
<td></td>
<td>42.87 25/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>7.39 2/25 = 8%</td>
<td></td>
<td>43.44 21/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>18.14 9/25 = 36%</td>
<td></td>
<td>46.72 25/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>21.82 21/25 = 84%</td>
<td></td>
<td>33.92 25/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>25.16 15/25 = 60%</td>
<td></td>
<td>39.40 25/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>22.09 13/25 = 52%</td>
<td></td>
<td>42.90 23/25 = 92%</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>8.45 2/25 = 8%</td>
<td></td>
<td>31.49 9/25 = 36%</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>16.61 19/25 = 76%</td>
<td></td>
<td>28.86 25/25 = 100%</td>
<td></td>
</tr>
<tr>
<td>Ave.</td>
<td>17.47 51%</td>
<td></td>
<td>39.14 89%</td>
<td></td>
</tr>
<tr>
<td>Est*</td>
<td>5.48 51%</td>
<td></td>
<td>7.24 76%</td>
<td></td>
</tr>
</tbody>
</table>

Confidence Metric

Figure 5.11 provides insight on the validity of the proposed confidence metric. Experiment #1 was performed for each of the proposed graphical techniques, and the reported RMSE, Ψ, was plotted vs. the true RMSE, Ψ. The analysis was also repeated for the proposed hybrid (hyb), sequential (seq), and simultaneous (sim) techniques for comparison. The reported RMSE vs. true RMSE exhibited heteroscedasticity, violating the assumptions of equal variance across the predictor values, that are required for linear regression. To address this, we utilized a log-log analysis in this and subsequent plots, which is better suited to describe the relationship of underlying data.
5.5. RESULTS

Figure 5.11: The reported RMSE metric is plotted vs. true RMSE (obtained via simulation analyses), revealing that, as expected, the reported RMSE metric is an overestimate of the true RMSE of registration. Note the log-log axes, therefore the intercept should be interpreted as a multiplicative bias. The proposed hybrid approach has the largest multiplicative bias.

The embedded confidence metrics for all three graphical methods provide a linear estimate of the true RMSE in log-log space. Note the interpretation of log-log regression coefficients is different than standard level-level regression. Here, the linear plot \( \ln \Psi = \beta_1 \cdot \ln(\hat{\Psi}) + \beta_0 \) corresponds to an exponential equation of the form \( \Psi = a \cdot \ln(\hat{\Psi})^k \), with \( \beta_1 = k \) referred to as the partial elasticity, and
$\beta_0 = \ln a$ as the multiplicative bias. The reported fits, transformed into linear space, are as follows:

\[
\Psi_{seq.} = 3.60 \cdot (\bar{\Psi})^{0.96}
\]

\[
\Psi_{hyb.} = 5.40 \cdot (\bar{\Psi})^{0.87}
\]

\[
\Psi_{sim.} = 4.94 \cdot (\bar{\Psi})^{0.93}
\]

Experiment #2 collated the reported pairwise RMSE, $\bar{e}$, vs. the true pairwise RMSE, $e$, for tie points with added positional deviations drawn from a normal distribution with a range of input variances. Kelbe et al. (2015a) demonstrated the usability of this metric to first order using ROC curves, and suggested that the pairwise embedded confidence metric represents an upper bound to the true error (i.e., a bias exists). The precision of the error, however, has not yet been assessed. Figure 5.12 provides information on the precision of the pairwise confidence metric. Again, an unequal variance of the predicted value was observed with respect to the predictor, yielding a log-log plot and subsequent regression. The reported fit, transformed into linear space, is as follows:

\[
e = 3.19 \cdot (\bar{e})^{0.95}
\]
5.5. RESULTS

Figure 5.12: The total path RMSE, $\Psi$, is an aggregate of the pairwise RMSE, $e$, associated with each edge. To uncouple this effect, we plotted the reported pairwise RMSE vs. the true pairwise RMSE, and found that the multiplicative bias from Figure 5.11 is inherited from a similar multiplicative bias in the pairwise RMSE estimate.

5.5.2 Graph-based Registration

In an effort to quantify the performance of the proposed graph-based registration approaches (objective ii), experiments #3 – 5 assessed various performance characteristics of the proposed graph-based registration techniques. Experiment #3 collated the true RMSE values, $\Psi$, vs. the corresponding $\text{RMSE}_{\text{in}}$, associated with the input tie point sets. The results are plotted in Figure 5.13.
Figure 5.13: Assessment of the true RMSE vs. the input RMSE of the tie points. Deviations of the input tie points introduce an inherent theoretical limit in the RMSE after registration, and we see that output results are primarily limited by this $\text{RMSE}_{\text{in}}$, although the proposed approach offers a 10% improvement (see regression fits).

The reported fits, transformed out of log-log space, are as follows:

\[
\Psi_{\text{seq}} = 1.00 \cdot (\text{RMSE}_{\text{in}})^{1.00} \tag{5.20}
\]
\[
\Psi_{\text{hyb}} = 0.90 \cdot (\text{RMSE}_{\text{in}})^{1.01} \tag{5.21}
\]
\[
\Psi_{\text{sim}} = 1.06 \cdot (\text{RMSE}_{\text{in}})^{1.02} \tag{5.22}
\]

The RMSE of the proposed graph-based technique was commensurate to the $\text{RMSE}_{\text{in}}$ of the input...
tie points. A significance test was performed on the parameters to determine if they are statistically
different from 1 ($\alpha = 0.10$). For several cases we failed to reject the null hypothesis, yielding the
simplified equations as follows:

\[
\begin{align*}
\Psi_{\text{seq}} &= \text{RMSE}_{\text{in}} \\
\Psi_{\text{hyb}} &= 0.90 \cdot (\text{RMSE}_{\text{in}}) \\
\Psi_{\text{sim}} &= (\text{RMSE}_{\text{in}})^{1.02}
\end{align*}
\]

Experiment #4 sought to answer the question, “does the graphical registration offer an improvement in output RMSE when compared to the RMSE of pairwise registration?” Ideally, the redundant information encoded in the graphical network could provide opportunity to reduce overall error by effectively averaging conflicting pairwise transformations. However, there is also the challenge that bringing additional scans into global alignment introduces additional rigid constraints, which may increase overall error, compared to a smaller quantity of tie point sets. This may increase the overall error when compared to a subset of tie point sets. The true RMSE for the proposed graph-based registration approach, $\Psi$, was plotted vs. the true RMSE obtained from pairwise registration, $\hat{\varepsilon}$, in Figure 5.14.
Figure 5.14: Assessment of true RMSE error obtained via graphical methods vs. the true RMSE error obtained by pairwise registration. Fractional multiplicative biases revealed that, as expected, the redundant information encoded in the graphical network provided opportunity to reduce overall error by averaging.

The reported fits, transformed into linear space, are as follows:

\[
\begin{align*}
\Psi_{seq} &= 0.65 \cdot (\bar{e})^{0.98} \\
\Psi_{hyb} &= 0.58 \cdot (\bar{e})^{0.98} \\
\Psi_{sim} &= 0.59 \cdot (\bar{e})^{0.97}
\end{align*}
\]

The significance of the regression parameters were tested, and it was confirmed that the exponents
in Equations 5.26-5.28 are statistically different from 1.

5.6 Discussion

The proposed graph-based registration achieved positive results for forest plots with a range of structural characteristics, including terrain topology (Figure 5.10a), sub-canopy occlusion (Figure 5.10b), and relatively open forest structure (Figure 5.10c). Figure 5.10a corresponds to site 10 (see Table 5.1; Figure 5.1a), which is dominated by *Pinus strobus* and has an elevation differential of 5.2 m across the scanned plot area of 20 m × 20 m = 400 m² (nominally, a 10% grade). Visually, all scans are aligned roughly to the appropriate position and orientation, although some errors are noticeable in the lower left section of terrain (see vertical offset between cyan and blue terrain points). However, these points are outside the plot area of interest (center 20 × 20 m). It is expected, and even desired, that errors are greater outside the nominal volume of interest, i.e., where tie points were located, in order to improve fit within the region of interest. Figure 5.10b corresponds to site 15 (see Table 5.1; Figure 5.1b), which is a mixed forest of both *Tsuga canadensis* and *Betula alleghaniensis*. Scans from this plot exhibited significant occlusion due to both sub-canopy branches and foliage. The nominal scanner range due to occlusion was ≈ 10 m, as evidenced by the sharp and consistent falloff in scanner returns above 10 m in the canopy. As a consequence, pairwise registration was only able to link 52% of the scans to the reference coordinate system. Graph-based registration is especially important in such a case of significant occlusion. By utilizing connections between multiple nodes, the proposed graph-based registration technique was able to bring 92% of the scans into global alignment. Figure 5.10c corresponds to site 31 (see Table 5.1; Figure 5.1c), which is a relatively open, mixed deciduous forest dominated by *Quercus rubra* and *Acer saccharum*. 70% coverage of 0.7 m ground vegetation, however, provides a challenge in accurately detecting the underlying terrain, which is necessary to localize the $z$ component of the tie points. We observe good registration results, however, as evidenced by the alignment of trunk, branch, and foliage clumping features (see detail inset).

A critical limitation of existing marker-free registration approaches is the lack of an associated confidence metric provided with each registration output. The graph-based confidence metric
proposed here is aggregated from the pairwise confidence metrics associated with each edge, which was first proposed in Kelbe et al. (2015a). The authors validated this metric to first order using ROC curves, and in this research, we augmented this with an experiment where true RMSE values were explicitly known. This provided opportunity to assess the validity of the proposed graph-based confidence metric (objective i).

From the second experiment, the reported pairwise RMSE was found to be linearly related to the true RMSE in log-log space with slope $\beta_1$ and intercept $\beta_0$. Transformed to level-level space, this represents an exponential model (cf. Equation 5.19) with an exponential factor, $k = \beta_1$, referred to as the partial elasticity, and a multiplicative bias, $a = 10^{\beta_0}$. Though the exponential factor in Equation 5.19 (i.e., the fitted slope in 5.12) is statistically different from 1, to first order it is close to 1. The parameter of interest is the intercept, which suggests that the pairwise confidence metric overestimates the true RMSE by a small multiplicative bias of 3.19. Note that this overestimation introduces an error that is several orders of magnitude smaller than the range of RMSE values between correct and incorrect transformations. For example, an incorrect transformation yields an RMSE on the order of $10^1$, see Figure 11 from Kelbe et al. (2015a). Thus, the error metric is able to distinguish between correct and incorrect transformations, but has less power in precisely predicting true RMSE. We suggest that this limitation is due to the random error of the tie points and the circular path construction for pairwise error estimation. With relatively small numbers of tie points, the random errors associated with these tie points influence the output rigid transformation parameters, introducing deviations to the nominal output transformation parameters associated with the forward and reverse paths. These deviations may add either constructively or destructively in the circular path construction. As a result, noise is added to the estimated RMSE values, reducing the precision of the pairwise RMSE estimate.

This underlying source of noise is aggregated in the construction of the proposed multi-sensor RMSE metric, as observed in the first experiment. As a result, Equations 5.16 - 5.18 (i.e., the fitted intercepts in 5.11) also exhibit a multiplicative bias, which is slightly larger than Equation 5.19. The embedded confidence metric associated with the sequential technique has the smallest multiplicative bias (3.60) followed by the simultaneous approach (4.94), and finally the proposed
hybrid approach (5.40). The implications are as follows: The reported graph-based RMSE retains its predictive power in differentiating between correct and incorrect transformations (cf. ROC curves in Figure 11 from Kelbe et al. (2015a)). The reported RMSE (both pairwise and graph-based), however, has limited predictive ability in quantifying the precise amount of error associated with paths of comparable error.

Figures 5.13 - 5.14 supplement the performance analysis of the proposed graph-based registration technique (objective ii). Recall that the reported RMSE metrics were found to have a multiplicative bias. To avoid this bias in our evaluation of the proposed approach, we used the true RMSE values obtained from simulated data, as the response variable in subsequent analyses. From the third experiment, we found that the true RMSE of registration was achieved commensurate with the input RMSE of the tie points (Figure 5.13). In other words, the input tie points have deviations, $RMSE_{in}$, about their true location due to an inability to precisely localize the stem-terrain intersection points. After simplifying the fitted models using significance testing, we found that the proposed hybrid method achieved the best performance, with true RMSE on average 10% lower than the deviation of the input tie points, $RMSE_{in}$ (Equation 5.23). This positive outcome suggests that the proposed hybrid technique is able to produce a registration result that is superior to the limit imposed by the deviations of the input tie point sets. This improvement is achieved by the exploitation of redundant information encoded in the graphical model. Recall that cycles in the graphical network introduce pose conflict (Huber and Hebert, 2003), which can be exploited to reduce error by averaging. As expected, we find that the proposed hybrid approach is able to achieve registration errors lower than the error of the input tie point sets, because up to 25 registration parameters are averaged for each node, thus reducing the impact of random deviations associated with individual paths.

Note that the proposed hybrid registration approach can be applied to any arbitrary set of tie points, including those extracted from externally-placed markers. Thus, this algorithm is extensible beyond the applications provided in the study. The framework presented here could be utilized in other domains to perform multi-view registration with embedded confidence metrics.
5.7 Conclusions

This study assessed the performance of a marker-free, multi-view registration approach, which achieved automatic, blind, global alignment of TLS point cloud data from multiple scans in forest environments. We extended the embedded RMSE metric associated with each pairwise correspondence (Kelbe et al., 2015a), in order to build a weighted graphical network. Estimated errors associated with different paths through the graph were then exploited to weight competing pose estimates and thus improve the precision of the output transformation parameters. As a result, we were able to achieve output RMSE registration results on average 10% better than the limit imposed by the RMSE of the input tie points, representing an improvement over standard sequential and simultaneous registration approaches. Additionally, we developed a circular self-closure framework, which provided an embedded estimate of the aggregate error associated with each transformation path. Quantitative analyses found this reported RMSE metric to exhibit heteroscedasticity, and have a small multiplicative bias compared to the true RMSE. As a result, we concluded that it is useful for identifying and rejecting poor transformations, but has limited predictive ability in precise error quantification. Finally, we demonstrated the potential to improve plot-level inventory assessment via global registration of multiple scans. This work offers an approach for multi-view registration of terrestrial laser scanner data without artificial targets, and opens the door to rapid plot-level and even stand-level structural assessment in domains such as forest inventory, airborne calibration/validation, and computer vision.
Chapter 6

Conclusions, Impact, and Outlook

6.1 Conclusions

Over the past decade, terrestrial laser scanning (TLS) has demonstrated a potential to address the limitations of conventional forest structure assessment, including the accuracy, reproducibility, detail, and fidelity of mensuration that can be achieved. Yet despite this promising outlook, operational use of TLS for forest structure assessment has been limited, due in part to the high cost of system hardware primarily designed for engineering, architecture, and forensics (Pueschel et al., 2013) and the lack of adequate, automated processing methods to facilitate rapid, economical data collection and processing. If TLS is to take hold as an effective tool for operational forest structure assessment, performance criteria must be satisfied against the budget and time constraints of end users (Mackrory and Daniels, 1995), and efficient, automated processing methods must be developed in order to facilitate economical data collection and processing procedures.

As a contribution to rectify this gap, this dissertation presented a set of robust forest structure assessment tools to support a recently-developed, low-resolution, low-cost, mobile terrestrial laser scanning system. Our contributions focused on three identified knowledge gaps in the literature, each contributing to the overall theme of reconciling performance criteria with operational limitations of TLS. The three objectives were as follows: (1) stem reconstruction from low-resolution
single-scan point cloud data, (2) blind, marker-free point cloud registration between scan pairs using view-invariant features, and (3) multi-view forest point cloud registration using graph theory. The combined work, we believe, addresses a major limitation in the feasibility of TLS for structural assessment of forest environments.

6.1.1 Single-scan stem reconstruction

The first objective (Chapter 3) assessed the performance of a novel tree stem reconstruction algorithm, which was developed to enable the use of low-resolution, single-scan terrestrial laser scanner data. Traditional approaches, which rely on the measured curvature of the tree stem, e.g., by circle or cylinder fitting, have proven to be unsuccessful for low-resolution instruments, such as the one used in this dissertation. Instead, we take advantage of the sensor-object geometry and reduce the dimensionality of the modeling to a series of 1-D line fits. This allowed for robust recovery of tree stem structure in a range of New England forest types, for tree stems that subtended at least an angular width of 15 mrad—the beam divergence of our system. From these facetized geometric models, standard forest inventory parameters were extracted and manually compared to measurements made using conventional techniques, e.g., a measuring tape. Unbiased retrieval of tree location (coefficient of determination ($R^2 = 0.996$, root mean square error (RMSE) = 0.17 m) and diameter at breast height (DBH) ($R^2 = 0.80$, RMSE = 0.06 m) was achieved within a 20×20 m plot for stems of diameter ≥10 cm, which were visible to the sensor. Plot-level estimates of stem density ($R^2 = 0.19$, RMSE = 188.08 stems·ha⁻¹) and basal area ($R^2 = 0.21$, RMSE = 16.23 m²·ha⁻¹), however, were limited by occlusion owing to a single scan location, as expected.

These results demonstrated the feasibility of lower-resolution sensors in providing data for operational forest inventories constrained by sample size, time, and cost. However, limitations due to occlusion confirmed the importance of point cloud registration in providing a plot-level inventory assessment. Therefore, the subsequent objective addressed the pairwise registration problem for forest point cloud data.
6.1.2 Marker-free registration of TLS data pairs

Registration of point cloud data for TLS in forest environments is notoriously time-consuming and labor-intensive, with the majority of studies relying on manually-placed external targets, to serve as control points in data registration. As a result, the registration process has represented a major bottleneck to the collection of sufficient sample sizes for operational inventory. Nevertheless, the integration of adjacent TLS measurements is a necessary step in order to compensate for the effects of occlusion, reduced point density, and range limitation, therefore providing adequate plot-level inventory estimates.

Therefore, the second objective quantified the RMSE of a blind, marker-free registration approach for terrestrial laser scanner data pairs in forest environments. A view-invariant feature metric was designed from the intersection of modeled tree stem geometry and terrain (Chapter 3). View-invariant geometric properties were incorporated to provide robust registration without knowledge of the initial sensor pose. An embedded confidence metric was developed, using set theory, to provide an upper estimate of the error associated with each transformation pair, and was validated using manual truth classification and receiver operator curves (ROC’s).

Analyses showed that (i) the algorithm is invariant (blind) to initial sensor pose, insensitive to error in DBH values, and possible with at least three corresponding tie points between scans, i.e., the degrees of freedom (DOF) necessary to constrain a three-dimensional (3-D) transform. We collated transformation results for 5,585 registration pairs in the New England forest environment, and found that while RMSE increased slightly with range between scanner locations, there was a much more prominent effect on the percentage of scan pairs that could be successfully linked. This was due to occlusion and a lack of corresponding objects within the scanners’ fields of view. Results informed considerations for optimal sample spacing for TLS data collection in New England or similar forests. Finally, we demonstrated that the registration algorithm is \( \text{RMSE}_{\text{in}} \)-limited, which extends results to other sensors and study areas. Owing to the minimization of least-squares error by RAndom SAmple Consensus (RANSAC) and singular value decomposition (SVD), the output RMSE of registration can be expected to be lower than the input error of the source data. This work provides an accessible and fully automatic approach for registering terrestrial laser scanner
data without artificial targets, thus enabling rapid structural assessment for domains of forest
inventory, airborne calibration/validation, and computer vision.

Several limitations were identified in the pairwise registration approach. A limited number
of shared tie points between scan pairs, in conjunction with periodic tree spacing, e.g., plantation
forests, may erroneously produce locally consistent registration results, despite global inconsis-
tency. Moreover, occlusion of emitted laser pulses limits the number of scans, and thus the
geographic extent, which can be linked to a central reference node via single, pairwise connec-
tions. To address these challenges, the subsequent chapter developed a multi-view registration
framework, which utilizes a global network of pairwise scan connections in order to link pairwise-
disconnected scans through a connected path.

6.1.3 Multi-view registration of TLS data

Chapter 5 assessed the performance of a marker-free, multi-view registration approach, which
achieved automatic, blind, global alignment of TLS point cloud data in forest environments. We
extended the embedded RMSE metric associated with each pairwise correspondence (Kelbe et al.,
2015a), in order to build a weighted graphical network. Estimated errors associated with different
paths through the graph were then exploited to weight competing pose estimates and thus improve
the precision of the output transformation parameters. As a result, we were able to achieve output
RMSE registration results on average 10% better than the limit imposed by the RMSE of the input
tie points, representing an improvement over standard sequential and simultaneous registration
approaches. Additionally, we developed a circular self-closure framework, which provided an
embedded estimate of the aggregate error associated with each transformation path.

Quantitative analyses found this reported RMSE metric to exhibit heteroscedasticity, and have
a small multiplicative bias compared to the true RMSE. As a result, we concluded that it is useful for
identifying and rejecting poor transformations, but has limited predictive ability in precise error
quantification. Finally, we demonstrated the potential to improve plot-level inventory assessment
via global registration of multiple scans. This work offers an approach for multi-view registration
of terrestrial laser scanner data without artificial targets, and enables alignment of point cloud
data over a broader geographic area. This has implications to approaches that provide improved structural estimates at scales relevant to plot-level inventory.

The combined work fills a vital gap in our ability to assess forest ecosystem structure, as dramatic changes in land use and other human activities necessitate improved monitoring and assessment of the biosphere (National Ecological Observatory Network, 2014). The outputs of this dissertation led to numerous applications and impacts across various domains. A few relevant areas of impact are highlighted in the following section.

6.2 Impact

6.2.1 Forest canopy assessment

The developments of this dissertation have several impacts across domains. First, additional TLS-derived structural outputs, at the expanded site-level (graph-based traversing of connected scans) could be explored towards understanding forest canopies. Canopy structure is defined as the spatial organization of the above-ground components of vegetation (Parker, 1995), and encompasses a description of the position, quantity, type, and connectivity of both the foliage and supporting woody components (Ross, 1981). Canopy structural information is both an influence on, and an indicator of, the state and dynamics of forest function (Ellsworth and Reich, 1993; Parker et al., 2004). Of particular interest is measuring the surface area of photosynthetic tissue (leaf area index (LAI)) within a forest canopy, as this determines the size of the plant-atmosphere interface (Weiss et al., 2004) and thus has both explicit biological (e.g., respiration potential) and physical (e.g., radiation interception) implications (Jonckheere et al., 2004). LAI is a dimensionless variable defined as one half the total leaf area per unit ground surface area (Lang et al., 1991; Chen and Black, 1992), with typical values between 3 and 19 for forests (Schulze, 1982).

The reader is referred to Jonckheere et al. (2004) for a review of traditional methodologies for ground-based measurement of LAI, and their various limitations. Importantly, commonly-used optical methods (which measure differential photosynthetic active radiation (PAR)) rely on external passive radiation, and thus require exacting sky conditions (e.g., cloud-free), which may
be impractical given the constraints of data acquisition (Strahler et al., 2008). Moreover, they provide simple spatial summaries of LAI with no opportunity to analyze its spatial variation (Jonckheere et al., 2004; Henning and Radtke, 2006b). To the contrary, TLS is an active sensing technique, and thus is not dependent on external radiation conditions. Moreover, explicit 3-D measurement offers an opportunity to provide enhanced structural information on the spatial distribution of leaf area (Danson et al., 2007), which may improve our understanding of forest canopies (Parker, 1995; Lefsky et al., 2002).

Although it falls outside the scope of this dissertation, we performed exploratory analyses in order to assess the potential for measuring LAI using low-cost, low-resolution TLS. Physically, this is a challenging problem for low-resolution instruments (Clawges et al., 2007), due to the effect of the beam diameter on LAI estimates (Wilson, 1963; Denison, 1997). Therefore, initial exploration focused on a statistical point cloud distributional approach, where site-level point distributional metrics, i.e., those metrics extracted from a connected set of scans, are correlated to forest biophysical variables (van Aardt et al., 2006). For example, accumulated point returns at various height bins may relate to laser penetration depth—and thus LAI—of forest canopies. Although there are obvious system-level biases for hemispherical TLS scanning systems, e.g., variable point density and beam-width across the volume of interest, point distributional approaches have enabled vertical canopy structure characterization from airborne laser scanning (ALS) instruments (van Aardt et al., 2006; Magnussen and Boudewyn, 1998; Means et al., 2000; Drake et al., 2002; Lefsky et al., 2002; Næsset, 2002), and may be offer comparable success for ground-based instruments.

Truth LAI data were collected coincident to TLS measurements for 19 sites at Harvard Forest, Massachusetts. At each site, 25 measurements were taken in a grid pattern (5 m spacing), and averaged to compute site-level LAI and lidar point distributions. Site-level distributional metrics were extracted from point clouds using Statistical Analysis Software (SAS) (SAS, Inc.) for subsequent regression analysis. Potential explanatory variables included standard distributional metrics of the z-values (mean, median, minimum, maximum, standard deviation, kurtosis, etc.), both for the entire point cloud, and for data binned from proportional height intervals (0-10% of the maximum, 10-20%, . . . 90-100% of the maximum). A stepwise linear regression approach was
taken to determine statistically significant predictor variables at the $\alpha = 5\%$ confidence level. We acknowledge the limitations of this approach, including the potential for overfitting a model (Butler, 2013; Lazer et al., 2014), especially when a large number of predictor variables are used relative to the number of samples. However, exploratory results demonstrated that some information related to LAI is embedded in the lidar point cloud distribution, with 66% of the variance (RMSE = 0.42; unitless LAI of $m^2/m^2$) explained by distributional metrics obtained from points in the upper height percentiles (80-100% of the maximum height); see Figure 6.1a. This has meaningful physical interpretation relating to the penetration depth of an emitted laser pulse and, by extension, the amount and distribution of foliage within the canopy. Residuals were approximately Gaussian and evenly distributed (Figure 6.1b). This is an aspect that we consider ideal for further exploration in future research, as discussed later.
6.2. IMPACT

Figure 6.1: A stepwise regression approach was used in an exploratory study to determine which lidar distribution metrics were significant predictors of LAI at the $\alpha = 5\%$ confidence level. Distributional metrics from the upper-height percentiles were significant predictors of LAI ($R^2=66\%$; RMSE = 0.42).

6.2.2 Forest carbon (C) cycling

A second example is the use of TLS-derived structural outputs towards understanding forest canopies. Specifically, structural outputs provided by TLS can be estimated as a proxy for net pri-
mary production (NPP) to understand the mechanisms controlling C cycling trends over decadal to century timescales. Aboveground NPP is the sum of annual wood and leaf production at the ecosystem scale, and is a primary indicator of C sequestration. This will constitute the MSc Biology thesis of collaborator, Cynthia M. Scheuermann, Virginia Commonwealth University (VCU).

Carbon sequestration by temperate forest ecosystems play an important role in regulating the climate system, currently removing 9% of annual global C emissions from fossil fuel combustion (Pan et al., 2011). Monitoring of forest C flux is important, providing predictive insight into long-term C cycling and climate impact. However, tower-based monitoring of C fluxes is well-understood at hourly to yearly timescales (Dragoni et al., 2011; Chen et al., 2009; Gough et al., 2008), but there is uncertainty as to how and why C fluxes change across decadal to century timescales, i.e., at scales relevant to ecological succession (Urbanski et al., 2007). This knowledge gap is becoming increasingly important, as a diverse patchwork of US East and Midwest forests – clear-cut and harvested during the mid-19\textsuperscript{th} and early 20\textsuperscript{th} centuries – are broadly reemerging due to changes in land-use across the region. These emerging forests are approaching successional stages, and present a challenge in predicting future C cycling trends.

Ms. Scheuermann proposed that decadal to century changes in NPP, a key indicator of C sequestration, may correspond with shifts in forest canopy structure, which alter the distribution and efficient use of limiting resources necessary for plant growth (Reich, 2012). For example, theoretical structural changes suggest that canopy gaps – which are present in very young forests – close during early succession, before reforming in mid-late stages of succession as dominant trees senesce (Figure 6.2). Other metrics, such as LAI, increase through early-succession, before reaching an asymptote at mid-late successional stages. In order to test this hypothesis, an experiment was designed in which forest structural outputs, obtained from this dissertation and closely related work, will be correlated against truth NPP data. This will establish a linkage between forest function, e.g., C cycling, growth and production, and forest structure, e.g., the spatial arrangement of leaves within a forest canopy.
6.2. IMPACT

In support of these goals, TLS data were collected for sample plots within a large-scale experimental forest chronosequence (Figure 6.3), representing a diverse range of forest structure at various stages of succession. Adjacent stands within this chronosequence have been systematically cut and burned over the past century, in an effort to support forest ecological research related to the emerging patchwork of forests in the US East and Midwest. Ms. Scheuermann will link TLS-derived structural outputs to true wood and leaf NPP data in order to identify the best structural correlates of NPP. Wood NPP data were obtained by annual dendrometer band measurements (Gough *et al.*, 2008), and leaf NPP data were obtained using leaf litter traps. A non-linear model will fit the time series (chronosequence) of structural parameters to stand age, in order to examine hypothesized structural trends over successional time. These efforts could
improve our understanding of carbon cycling trends over decadal to century time-scales, and thus will contribute to our improved ability to predict long-term climate impact due to emissions reductions from temperate forest ecosystems.

Figure 6.3: Controlled cut and burn management within adjacent stands of a forest chronosequence provides opportunity to understand C cycling trends in emerging mid-late successional forests.

### 6.2.3 Virtual scene generation

A second impact of this dissertation concerns the generation of virtual forest scenes. Virtual scenes are used to convey realistic scene content in applications such as computer-generated imagery (CGI), gaming, and animation. Of particular interest, however, is the generation of virtual scenes for scientific applications, e.g., for evaluating image system designs, validating image exploitation algorithms, and providing training data for image analysts. The Digital Image and Remote Sensing Image Generation (DIRSIG) model is one such synthetic image generation model, and has a 20-year development history at the Digital Imaging and Remote Sensing Laboratory (DIRS) at Rochester Institute of Technology (RIT). DIRSIG is a first-principles, physics-based ray tracing model capable of generating synthetic imagery for a range of modalities, including passive single-band, multispectral (Schott et al., 1999), or hyperspectral (lentilucci and Brown, 2003) imagery in
the visible through thermal infrared region of the electromagnetic spectrum, in addition to low-light (Ientilucci, 1998), light detection and ranging (lidar) (Brown et al., 2005), synthetic aperture radar (SAR) (Gartley et al., 2010), polarimetric (James R. Shell, 2005), and other modalities.

DIRSIG requires both a sensor model and a scene model. While the sensor model is relatively simple to create using an integrated graphical user interface (GUI) and basic system specifications, the generation of realistic scenes is much more time-consuming. In order to generate virtual scenes that contain the spatial and spectral complexity of real-world data, the scene model must be able to reproduce the first-principles, radiative mechanisms, which combine to produce the data observed by real-world imaging systems. Moreover, in order to link virtual data back to real-world signal outputs, it is desirable that virtual scenes represent, to first order, the structural scene complexity of a real-world study site imaged by a sensor platform. In the context of simulated imaging of forests, for example, synthetic trees ideally should have similar structural characteristics – size, canopy extent, and location – as a reference field plot in the real world. Currently, the generation of such scenes is a difficult task, requiring labor-intensive modifications of software-derived tree models, which poorly reflect the competition for resources within a closed forest canopy.

To address this challenge, Dr. Martin van Leeuwen (Postdoctoral Research Associate, RIT), has developed a technique for parameterizing virtual forest scenes based on the structural outputs provided by this dissertation. A Python script takes as input the unified (plot-level) geometric stem models reconstructed for each forest plot, and grows a “forest” of virtual trees using the Arboro tree generation software. The geometric characteristics of each tree stem (taper, radius, sweep, etc.) are parameterized by the measurements obtained from the TLS. Canopy extent is modulated to account for resource competition within a forest canopy by constraining the Arboro tree geometries to the volumetric Voronoi cells obtained from the TLS stem models. Future work could integrate canopy structural estimates, such as vertical foliage distribution, in order to spatially allocate the distribution of foliage based on measured, real-world data.

These efforts provide an opportunity to efficiently generate virtual scenes, which approximate the characteristics of actual forest plots in the real world. Note that at the scales of interest, e.g., 60 m pixels for the Hyperspectral Infrared Imager (HyspIRI) mission, the synthetic scene content
6.2. IMPACT

need not be an exact replica of the true spatial distribution of real-world branch and leaf structure. However, a first-order approximation is a step towards linking the simulation environment back to real-world earth observation. This has implications for synthetic image generation and testing, which will be described in the next section.

Figure 6.4: Output geometric stem models are used to reconstruct virtual forest scenes based on observed structural data. This has implications for synthetic image generation and testing.

6.2.4 System phenomenology studies

Realistic virtual scenes, such as those identified in the previous section, may offer opportunity to improve our understanding of system phenomenology, especially when combined with pre-launch test flights of planned remote sensing missions. For example, preliminary test flights could be used to predict data outcomes or challenges, or may provide opportunity for anticipatory algorithm development. This is especially informative when combined with the DIRSIG image simulation environment, in which a replicate sensor model “images” a corresponding virtual scene replicate. With this framework, system design parameters, e.g., outgoing lidar pulse width,
point spread function (PSF), etc., can be adjusted for the virtual instrument, in order to evaluate how those changes affect the output data. Or, the scene content can be purposefully modified, in order to isolate the effects of structural variation on observed, simulated imagery. With the support of the structural outputs from this dissertation, Mr. Wei Yao (Ph.D. candidate, Imaging Science, RIT) is pursuing these efforts in the context of the planned HyspIRI mission (Yao et al., 2015b; Yao et al., 2015a).

As part a National Aeronautics and Space Administration (NASA)’s decadal survey strategy, planning for the HyspIRI mission is currently underway towards addressing key science questions related to the world’s ecosystems. Preparatory work is ongoing to provide antecedent science data for the anticipated HyspIRI mission and associated science products. The HyspIRI mission includes a visible to short wave infrared (VSWIR) imaging spectrometer and a multispectral thermal infrared (TIR) instrument, mounted on a satellite in low earth orbit (LEO) (Roberts et al., 2012). Global coverage is provided with a relatively coarse, 60 m ground sample distance (GSD). This large ground-projected pixel size introduces some uncertainty in terms of how observed spectral radiance is affected by structural variations within the instrument footprint. This is because the system’s PSF weights spectral contributions unevenly across the instantaneous field of view (IFOV) of a pixel. Therefore, a tree at the periphery of a pixel, for example, may contribute differently to the observed radiance than a tree at the center (Figure 6.5).

![Figure 6.5: HyspIRI’s (and other similar sensors’) large-footprint PSF introduces uncertainty in terms of how within-pixel structural arrangements affect observed radiance.](image)

In order to investigate these science questions, Mr. Yao is utilizing outputs from this dis-
sertation in order generate virtual scenes, which have structural attributes modeled after the real-world study sites imaged during pre-launch HyspIRI test flights. For example, Figure 6.6 shows a virtual scene modeled after the remote sensing calibration plot for the National Ecological Observatory Network (NEON)'s Pacific Southwest domain (37°6′43.77″N, 119°44′11.85″W). Multiple simulated HyspIRI data sets were then generated by varying within-pixel scene variables, such as forest density, the position and distribution of trees, crown size, etc. Statistically significant differences among a series of narrow-band vegetation indices were used to assess the impact of sub-pixel vegetation structure on spectral response. Early results indicated that HyspIRI is sensitive to sub-pixel vegetation structural variation, even outside the IFOV of a single pixel, i.e., due to impact from the tails of a Gaussian-like PSF. These developments may inform future HyspIRI vegetation data products, e.g., by adapting calibration strategies to account for this sub-pixel variation of structure.
Figure 6.6: Structural outputs from this dissertation are being used to investigate the impact of sub-pixel structural variation on the assessment of vegetation structure via spectroscopy for the HyspIRI mission.

Moreover, Mr. Paul Romanczyk (Ph.D. candidate, Rochester Institute of Technology) is utilizing virtual scenes in the DIRSIG environment to improve understanding of waveform light detection and ranging (wliadar). Early work identified the level of geometric fidelity necessary for small-footprint wliadar simulations of virtual forest scenes (Romanczyk et al., 2013b). In other words, synthetic tree models may contain various levels of geometric fidelity, e.g., trunk, boughs, various levels of branches, twigs, etc. These geometric components enforce a tradeoff between model fidelity and random-access memory (RAM)/computation speed. It was found that several levels of geometric fidelity can be removed without statistically impacting the observed signals, which has implications to improving the practicality of forest simulation studies. An additional study examined the effect of positioning error on the repeatability of small-footprint wliadar signals.
6.2. IMPACT

(Romanczyk et al., 2013a), using virtual forest scenes. Additionally, related work by our research group has explored the utility of DIRSIG in understanding energy attenuation within the forest canopy in small-footprint wlidar signals (Cawse-Nicholson et al., 2013b), and the scalability of spectral LAI metrics (Cawse-Nicholson et al., 2013a).

6.2.5 Remote sensing calibration/validation

Finally, the developments of this thesis have implications for supporting the calibration and validation of remote sensing missions, by providing spatially-explicit structural reference data with which to understand the passive or active interactions of electro-magnetic (EM) energy with object structure. Traditional, ground-based forest structure assessment, e.g., inventory, is rapidly becoming outmoded by the increasingly sophisticated airborne and spaceborne imaging systems, which rely on fine-scale ground-level reference data for calibration and validation of large-scale models (Hilker et al., 2012a; Jupp, 2011; Lindberg et al., 2012; Liang et al., 2012; Yu et al., 2010). ALS has matured to operational use over the past decade for large-scale forest structure assessment (e.g., Wehr and Lohr, 1999; Nelson et al., 1988; Lefsky et al., 2002; Næsset, 2007), while a new sophistication of sensing systems, namely, small-footprint wlidar, remain underutilized due to the limited understanding of the interaction between a small laser footprint and forest structure. For an improved exploitation of the next generation of airborne/spaceborne sensing systems, TLS data could be used to link observed data back to true object structure at the fine scale. TLS provides the same structural assessment as ALS, but from a ground platform. This complementary perspective could provide a synergistic structural data product to link the observed airborne data back to the true forest structure at the fine scale (Hilker et al., 2012a; Jupp, 2011), i.e., thereby potentially improving calibration and validation of airborne sensing structural products. For example, Figure 6.7 shows ALS data provided by NEON’s Airborne Observation Platform (AOP) from Soaproot Saddle, CA. Coincident TLS data were collected for a number of ground validation plots within the study area. The below-canopy, hemispherical perspective of TLS (Figure 6.7, inset) provides additional information on sub-canopy vegetation distribution, etc., which is not measured by the airborne platform. These contributions reflect an overall goal of RIT to support
NEON’s long-term monitoring initiatives with high-fidelity structural ground truth (van Aardt et al., 2014).

![Image of road and TLS point cloud](image)

Figure 6.7: TLS provides a complementary structural data set to ALS offering potential to support calibration/validation of remote sensing missions. (a) Nadir image shows a road spanning from the northwest to southeast corner. Swath width is \( \approx 240 \) m. (b) TLS point cloud (inset) shows additional structural detail, including the road, sub-canopy vegetation, and individual tree stems.

TLS outputs have also demonstrated potential to provide ground truth for satellite remote sensing, with implications for forest monitoring in context of the United Nations Reduced Emission from Deforestation and Forest Degradation (UN-REDD) initiative. The UN-REDD program provides incentives to developing countries to reduce net greenhouse gas emissions through enhanced forest management. Typically, field validation of satellite data is provided by manual estimates of stem volume obtained from national inventory data. The limitations of traditional forest inventory, e.g., financial constraints on scope and scale, however, underscore the need for more objective, efficient forest monitoring using recent technological advances, such as TLS. In support of this, Dr. Akira Kato (Chiba University, Japan) explored the utility of TLS in providing ground-truth measurements for satellite remote sensing (Kato et al., 2013). LAI was extracted...
from TLS data and compared to estimates obtained from the Landsat-derived reduced simple ratio (RSR), with $R^2 = 0.79$. The previous discussion concerned calibration of remote sensing data using TLS. In contrast, other efforts have focused on the fusion of modalities to improve overall forest parameter estimates.

Ms. Claudia Paris (Ph.D. candidate, University of Trento, Italy) investigated a fusion approach to improve canopy volume assessment using both TLS and ALS (Paris et al., 2015). Canopy volume has important implications to fuel loading, habitat provision, etc. However, explicit measurement of canopy volume using conventional techniques is nontrivial, due to the complex structure of irregular, natural surfaces. Moreover, while laser scanning offers objective measurement, there are limitations in providing comprehensive coverage of a tree canopy from a single perspective. To address this, Ms. Paris coregistered data from both ALS and TLS for sites in San Joaquin Experimental Reserve, California, USA. In the fused point cloud, measurement of the upper canopy, provided by ALS, was augmented with samples of the below-canopy structure via data from a hemispherical scanning TLS placed around the location of a single tree. Crown parameters were then estimated by applying alpha shapes to the spatial extent of the fused point cloud (Figure 6.8).
6.3 Outlook / Future Work

Despite the opportunity for impact across a range of disciplines, some limitations of this research have emerged. Here we describe algorithmic and system-level limitations in the context of the overall objectives of this work. Addressing these limitations in future work could improve the overall outcomes and strengthen subsequent research.

6.3.1 Algorithmic Limitations

The stem reconstruction approach developed in chapter 3 provided robust measurement of the un-occluded, main tree stems, from low-resolution (15 mrad) TLS data. It should be noted that a recent research focus has emerged in the literature of modeling complete tree architecture— including boughs, branches, etc.— from high resolution (≈0.15 mrad) data. In contrast, our approach does not attempt to maximize geometric fidelity at the cost of operational efficiency. Rather, we
presented tools for rapid measurement of first-order stem parameters, relevant to forest inventory and management. In an effort to support these objectives, we impose geometric assumptions on the form of a single tree stem, which constrain the fidelity of the output models. Nevertheless, treatment of more complex tree structures, e.g., split stems, boughs, etc., could improve the fidelity of tree models obtained from TLS. Improved structural information could provide additional information relevant to forest management and harvesting. For example, information on the height-to-first branch could provide insight into merchantable timber volume, or timber quality assessment. Additional modeling fidelity would require less strict geometric assumptions, potentially leading to greater modeling errors due to noise. Therefore, future work could examine the impact of commission errors, related to structural elements, as a function of improved geometric fidelity.

The marker-free registration approach developed in chapters 3 and 5 provided robust alignment of TLS data, with significant improvements over the state-of-the-art. An approach balancing simplicity and rigor was taken to minimize error, while maintaining computational efficiency. However, some authors have emphasized the importance of extracting tie points, which are broadly distributed throughout the measurement volume. This strengthens the robustness of the rigid transformation by reducing the potential for degeneracy, e.g., coplanar tie points, and distributes errors more evenly throughout the volume of interest. Our approach utilizes tie points, whose distribution through the z-dimension was limited by the terrain. With an eye for forest inventory, we opted to minimize errors at the ground plane, where stem-level measurements, e.g., DBH, are typically recorded. However, we acknowledged that errors may be larger at other locations throughout the volume, e.g., upper-canopy regions. Future applications may require greater registration precision in the upper-canopy. This could be addressed in future work as follows: By modeling trees with improved geometric fidelity, additional view-invariant tie points are envisioned, e.g., as the intersection of first-order branches/boughs and stems. These tie points could be integrated into the registration algorithm with minimal effort, in order to distribute registration errors more evenly throughout the scan volume or instrument field-of-view.

A second limitation is that forests with certain structural characteristics will not produce
positive results. The algorithm was tested for TLS data collected in an open, woodland savanna (San Joaquin Experimental Range (SJER), Fresno, California) with poor results. This was attributed to occlusion of the stem surface of interest. As opposed to closed forest canopies, where resource competition discourages sub-canopy vegetation growth, open woodland savannas may have significant occlusion of the main stem, thus preventing the establishment of adequate tie points for registration. For exploitation of TLS data collected in open woodland areas such as SJER, researchers may have to rely on coarse alignment obtained from a regularly measured sampling pattern. Note that we expect good performance for registration in urban or park settings, where manual pruning provides for adequate sampling of the tree stem and terrain. Additionally, we expect our approach to fail for homogeneous forest plantations, where trees are planted evenly in a grid pattern, yielding a symmetric or non-unique pattern of tie points. This could be addressed in future work by incorporating additional tie points, e.g., based on branch structure, or by utilizing a regular sampling scheme to constrain the search space and avoid convergence to a local minimum.

Finally, the proposed registration approach enforces a fairly strict constraint that two disjoint transformation paths must exist between scan pairs. While just three corresponding tie points are sufficient for a single rigid transformation in 3-D, we require at least six tie points in order to compute an estimated transformation error via a circular path (forward/reverse) between point clouds. Thus, for data sets with only 3–5 corresponding tie points, the algorithm will be unable to report an output transformation, even if a single path between exists between scans. This limitation was observed in our experimental results, and becomes especially relevant for forest sites with significant occlusion and/or low stem densities. Future work could examine the opportunity to relax constraints, perhaps with additional user interaction, in order to improve the percentage of scans that can be linked to a reference scan, albeit without an embedded error.

### 6.3.2 System-level Limitations

A primary system-level limitation concerns the hemispherical scanning characteristics of many point-based sampling systems, such as TLS. As opposed to ALS, which provides roughly consistent data characteristics (point density, beam size, etc.) across the collection area, TLS data have
variable characteristics across the measurement range. For example, point density differs by three orders of magnitude across the zenith range, due to equal-angle sampling of both the rotation stage (azimuth direction) and mirror rotation (zenith direction). In an effort to reduce user-defined thresholds to system-dependent parameters, the developed stem modeling algorithm relies on estimating a range-dependent point density in order to reduce noise. However, this required an assumption of constant point density within a voxel. Although this assumption was satisfied for voxels at far-range, it was less valid for voxels at close-range, which subtended a greater range of zenith angles. A reduced classification accuracy at near-range was attributed to this system-level limitation of point density.

In contrast to point density, which is a function of the equi-angular sampling pattern, an additional specification of interest concerns the temporal and spatial beam-width of our system. Spatially, the 15 mrad beam divergence causes a linear reduction in effective resolution with range. Algorithms in this dissertation were developed to address these limitations, enabling reconstruction of tree stems which subtended at least one beam-width. Nevertheless, some remaining challenges were observed relating to the finite, i.e., on the order of 1 ns \( \approx 0.15 \text{ m} \), temporal pulse-width. Interaction between an emitted laser pulse and intercepted objects causes a temporal deformation in the backscattered pulse. For hard targets, which are fully subtended by the emitted pulse, the backscattered waveform has a similar Gaussian shape as the outgoing pulse, providing straightforward digitization of range. In contrast, natural surfaces (vegetation elements) may be smaller than the beam size (15.8 cm at 10 m; 45.8 cm at 30 m for the system used in this dissertation). As a result of partial interception, the backscattered energy profile is complex, and less suited to digitization of explicit returns. This often causes digitization of “phantom points” (Eysn et al., 2013), which are located between two targets (Parker et al., 2004). Similarly, intensity information is virtually ignored in forest TLS studies due to the untenable calibration between recorded intensity values and true object material properties. However, this may provide useful information relating to tree species, health, or even differentiation between bark and foliage.

These limitations relate to an overall challenge of using commercial system components,
which operate as a “black box” in the imaging chain. A record of the backscattered waveform, internal digitization routines, laser PSF’s, etc., are generally proprietary, and are not disclosed by sensor manufacturers. Knowledge of this information, however, could provide valuable insight towards improving the spatial and radiometric fidelity of TLS instruments. For example, spatio-temporal deconvolution of the waveform could improve the accuracy of spatial measurements and reduce the phantom points between targets. Moreover, knowledge of emitted laser pulse characteristics could support radiometric calibration of the recorded intensity (Wagner, 2010). Towards addressing these challenges, some researchers have focused on developing custom laser scanning systems (Strahler et al., 2008; Douglas et al., 2012), which provide for explicit knowledge of system parameters. The Echidna® (Strahler et al., 2008) and Dual Wavelength Echidna® Lidar (DWEL) (Douglas et al., 2012) provide for digitization of the entire backscattered waveform, in addition to explicit knowledge of system parameters. Although internal system parameters remain out of reach for the system used in this study, future work could utilize laboratory tools to (i) characterize measurable system parameters and (ii) quantify the effects of these parameters on structural measurement.

Finally, these system effects are exacerbated by occlusion, which is consistently cited as a primary limitation in forest structure assessment from TLS. While registration (chapters 4 - 5) addressed this limitation by incorporating additional measurements in the x-y plane, occlusion in the z direction remains a limitation of this research. For example, tree height is difficult to assess due to the reduced penetration of the laser beam into the forest canopy. This is especially challenging for systems with a large beam divergence, due to the reduced probability of unobstructed laser penetration through canopy gaps. Although not critical to extraction of volume– DBH is the primary explanatory variable– height is still an important variable for volume assessment, with implications to timber yield, carbon storage, UN-REDD, etc.

As a consequence of these limitations, forest structural assessment from TLS arguably is still incomplete from a forest manager’s perspective. In other words, TLS makes significant advances in some areas of forest measurement fidelity, e.g., 3-D stem sweep, but remains limited in other key information areas for forest inventory, e.g., height and species. This is confounded
by the gap between the development of technical tools, presented here, and the application to real-world forest management. In particular, the accuracy/precision of forest management using TLS is dependent upon numerous interrelated factors, including system-level limitations, data collection/setup parameters, and forest type. Analyses in this dissertation provided both qualitative and quantitative guidance on the impact of various factors independently, e.g., tree stem classification accuracy vs. forest type (Table 3.4) pairwise registration error vs. range (Figure 4.12), and multi-view registration accuracy vs. forest type (Table 5.2). However, future work could consolidate these analyses, in order to provide an authoritative “best practices” manual for optimal data collection parameters, given accuracy requirements and forest type. In conclusion, this research developed novel technical tools, which support forest inventory and remote sensing calibration/validation from an operational perspective. However, additional work is needed to understand the specific management implications relative to the algorithm and system-level limitations discussed above.
Bibliography


