Kun Liu

Summary

I am currently a data scientist at MediaGamma Ltd. I design and build automatic solutions to address large-scale machine learning problems in computational advertising. From 2013 to 2016, I was a research associate at University College London (UCL). I worked on a geospatial big data project called IQmulus which provided a platform to process massive amounts of geospatial data and serve useful knowledge by processing algorithms developed by the project. From 2010 to 2013, I was a Ph.D. researcher at the ALICE team of INRIA (French Institute for Research in Computer Science and Automation). I carried out extensive research which proposed novel methods to process 3D data acquired by multiple cameras. I received my doctorate in computer science from Institut National Polytechnique de Lorraine. Please see my LinkedIn profile `https://www.linkedin.com/in/kwunlyou` for more information.

Specialities

Large-scale Machine Learning, Optimization, Numerical Analysis, Geometry Processing,

Skills & Expertise

Programming	Python, C++, Scala
Libraries	Apache Spark, Scikit-learn, SuiteSparse, OpenGL, QT
Operating Systems	Linux, macOS, Windows
Others	Amazon AWS, Git, Jenkins, Docker

Experience

2016 - Present **Data Scientist**, *MediaGamma Ltd*, United Kingdom.
As a data scientist, I am developing applications for programmatic advertising by applying cutting-edge machine learning technologies. The role is to design and build automatic solutions which address large-scale machine learning problems in computational advertising.

2013 - 2016 **Research associate**, *University College London*, United Kingdom.
I am a contributor to the project IQmulus (`http://www.iqmulus.eu/`) funded by European Commission. This project aims to create a high-volume fusion and analysis platform for large geospatial datasets. I am developing and implementing methods to process large LiDAR datasets using the state-of-the-art cluster computing engine Apache Spark.

2010 - 2013 **PhD researcher**, *INRIA-Nancy Grand-Est*, France.
I was involved in the project Physigrafix which combines acquisition techniques, geometry, and mechanics to develop simulation models that accurately reproduces realistic deformation. My work aims to develop specialized methods suitable for processing point data originating from multi-view reconstruction.

Education

2010 - 2013 **PhD candidate**, Computer Science, Institut National Polytechnique de Lorraine.

2008 - 2010 **M.S.**, Applied Mathematics, Zhejiang University.
Excellent Graduation Thesis in Zhejiang University

2004 - 2008 **B.S**, Mathematics, Zhejiang University.
Excellent Graduation Thesis in Zhejiang University
Overall GPA: 3.82/4.0 (87.22/100)
GPA of the last two years: 3.86/4.0 (89.75/100)

(+44) 79 27 59 71 77 • zju.liukun@gmail.com • kun-liu.com

Languages

English	Professional working proficiency
Mandarin Chinese	Native

Honors & Awards

2005	Excellent Undergraduate of the year
2005	Scholarship of State-level Training Base
2007	Scholarship of State-level Training Base
2009	Excellent Graduate Scholarship of the year

Publications

Change Detection of Mobile LiDAR Data using Cloud Computing
Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 2016
Authors: **Kun Liu**, Jan Boehm, Christian Alis
Change detection has long been a challenging problem although a lot of research has been conducted in different fields such as remote sensing and photogrammetry, computer vision, and robotics. In this paper, we blend voxel grid and Apache Spark together to propose an efficient method to address the problem in the context of big data. Voxel grid is a regular geometry representation consisting of the voxels with the same size, which fairly suites parallel computation. Apache Spark is a popular distributed parallel computing platform which allows fault tolerance and memory cache. These features can significantly enhance the performance of Apache Spark and results in an efficient and robust implementation. In our experiments, both synthetic and real point cloud data are employed to demonstrate the quality of our method.

Classification of Big Point Cloud Data using Cloud Computing
Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-3/W3, 553-557, 2015
Authors: **Kun Liu**, Jan Boehm
Point cloud data plays a significant role in various geospatial applications as it conveys plentiful information which can be used for different types of analysis. Semantic analysis, which is an important one of them, aims to label points as different categories. In machine learning, the problem is called classification. In addition, processing point data is becoming more and more challenging due to the growing data volume. In this paper, we address point data classification in a big data context. The popular cluster computing framework Apache Spark is used through the experiments and the promising results suggests a great potential of Apache Spark for large-scale point data processing.

Sphere Packing Aided Surface Reconstruction for Multi-view Data
10th International Symposium on Visual Computing (ISVC), 2014
Authors: **Kun Liu**, Patricio A. Galindo, Rhaleb Zayer
Surface reconstruction has long been targeted at scan data. With the rise of multi-view acquisition, existing surface reconstruction techniques often turn out to be ill adapted to the highly irregular sampling and multilayered aspect of such data. In this paper, a novel surface reconstruction technique is developed to address these new challenges by means of an advancing front guided by a sphere packing methodology. The method is fairly simple and can efficiently triangulate point clouds into high quality meshes. The substantiated experimental results demonstrate the robustness and the generality of the proposed method.

(+44) 79 27 59 71 77 • zju.liukun@gmail.com • kun-liu.com

A New Framework For Interactive Segmentation of Point Clouds

Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XL-5, 357-362, 2014

Authors: **Kun Liu**, Jan Boehm

Point cloud segmentation is a fundamental problem in point processing. Segmenting a point cloud fully automatically is very challenging due to the property of point cloud as well as different requirements of distinct users. In this paper, an interactive segmentation method for point clouds is proposed. Only two strokes need to be drawn intuitively to indicate the target object and the background respectively. The draw strokes are sparse and don't necessarily cover the whole object. Given the strokes, a weighted graph is built and the segmentation is formulated as a minimization problem. The problem is solved efficiently by using the Max Flow Min Cut algorithm. In the experiments, the mobile mapping data of a city area is utilized. The resulting segmentations demonstrate the efficiency of the method that can be potentially applied for general point clouds.

Bundle Adjustment Constrained Smoothing For Multi-View Point Cloud Data

8th International Symposium on Visual Computing (ISVC), 2012

Authors: **Kun Liu**, Rhaleb Zayer

Direct use of denoising and mesh reconstruction algorithms on point clouds originating from multi-view images is often oblivious to the reprojection error. This can be a severe limitation in applications which require accurate point tracking, e.g., metrology. In this paper, we propose a method for improving the quality of such data without forfeiting the original matches. We formulate the problem as a robust smoothness cost function constrained by a bounded reprojection error. The arising optimization problem is addressed as a sequence of unconstrained optimization problems by virtue of the barrier method. Substantiated experiments on synthetic and acquired data compare our approach to alternative techniques.

Paint Mesh Cutting

Computer Graphic Forum (Proceedings of Eurographics 2011)

Authors: Lubin Fan, Ligang Liu, **Kun Liu**

We present a novel progressive painting-based mesh cut out tool, called Paint Mesh Cutting, for interactive mesh segmentation. Different from the previous user interfaces, the user only needs to draw a single stroke on the foreground region and then obtains the desired cutting part at an interactive rate. Moreover, the user progressively paints the region of interest using a brush and has the instant feedback on cutting results as he/she drags the mouse. This is achieved by efficient local graph-cut based optimizations based on the Gaussian mixture models (GMM) on the shape diameter function (SDF) metric of the shape. We demonstrate a number of various examples to illustrate the flexibility and applicability of our system and present a user study that supports the advantages of our user interface.