Finding common geographic features in old maps and photographs for reconstructing historical landscape – a data criticism approach for Silk Road ruins

Asanobu KITAMOTO, National Institute of Informatics
Yoko NISHIMURA, Toyo Bunko

http://dsr.nii.ac.jp/
Digital Silk Road Project

• Started in 2001.

• Digital Humanities: Collaborative work among informatics + humanities scholars.

• Databases and digital resources are publicly accessible on the Web.

http://dsr.nii.ac.jp/
Toyo Bunko Rare Books

- Digitization of 203 books, 59358 pages.
- Collection of relevant academic references.
- Manual input of captions and TOCs.
- OCR for full-text search (with errors).

http://dsr.nii.ac.jp/toyobunko/
Dealing with Many Types of Data

Text

Photograph

Map

Gazetteer
Platform for Linking Entities

• An **entity** (a geographic feature) may have different representations in **historical sources**.

• **[Question]** How to identify and link them?

• **[Answer]** Data criticism approach – critically “read” data to reconstruct historical facts.

• **[Problem]** Error, variation, and different conceptualization; such as location error, name variation, different definitions, …
Stein’s map “Innermost Asia” was registered and displayed on Google Earth satellite images.
Problem of “Missing” Ruins

- Oi-tam, ruined fort
- Bögan-tura
- Buluyuk (Shipang, Sassik-bulak, Kazma)
- Murtuk-ruins

- Yogan-tura
- Chikkan-köl
- Bedaulat’s town, Bēsh-kāwuk, Kosh-gumbaz
- Yutōgh
Error Distribution in Turfan Basin / White: Innermost Asia / Black: Serindia

• Some ruins were reported by 20\textsuperscript{th} expeditions, but are missing in recent survey reports.

2013/12/10 PNC 2013
Removing Location Errors

Based on error information of maps, our guess about the location of Murtuk Ruins is represented as

Estimated error: west-southwest 5.6km
Matching Entities

Stein’s map and satellite images for the same area. Each source reports different ruins due to different conceptualization.
Photographs as Evidence

伯西哈石窟(烏江不拉克仏塔)

Murtuk Ruins (M. B. I)

烏江不拉克烽火台

Murtuk Ruins (Ruined Shrine M. C. I)
Grünwedel Map (Gaochang)

How to use the previously “untrusted” map?
Inconsistencies in Gaochang Maps

Aurel Stein

Albert Grünwedel
Topological Interpretation

Where is ν and μ? We have multiple candidates...

Search for a road in north of γ and ο, and a road between the inner wall to the wall gate.

Hypothesize the location of ν and μ (to be verified later).

Maps designed for navigation purposes should preserve the topology, not to get lost when visiting there again.
Most Ruins are Identified

Linking expedition results to recent surveys may lead to new interpretations.
Linking Entities Across Sources

Textual criticism
Textual source A
Place name S

Geographic
And Visual
Relationship

Place name U

Data Source C, D, E...

Data criticism

Place name V

Textual source B
Place name T
Registration of Maps

Geometric Correction

Map 1

- All points are registered.
- Shapes are distorted.

Map 2

Single-Point Registration

Map 1

- Single point is registered (but no other points).
- Shapes are not distorted.

Register ‘B’
Qianlong Map (Beijing)
Massive Geometric Correction

Huge size = W 13 m x H 14 m
Many sheets = 203 sheets in total
Massive pixels = 29 billion pixels

• Control points + lines:
  We proposed a new geometric correction.
Discovery and Digital Survey

- **Discovery:** 5 sheets have misarrangements due to improper reconstruction in the past.
- **Digital Survey:** Place-names were checked to make a comprehensive gazetteer.
Mappinning (Map+Pinning)
http://dsr.nii.ac.jp/digital-maps/mappinning/
Stein Placename Database
http://dsr.nii.ac.jp/digital-maps/stein/place-names/

- Unique IDs for about 5500 ruins in Silk Road.
- Locations on multiple maps are compared and evaluated on the “mappinning” tool.
Kara-Khoja Placename Database
http://dsr.nii.ac.jp/digital-maps/kara-khoja/place-names/

Linked Data Concept

Resources are linked at unique IDs.
Different Conceptualization

Ruin A in X

Ruin C in Y

Ruin D in Y

Ruin E in Y

Ruin B in X
Different Conceptualization

• Ruins are named ...
  – after local placenames.
  – after geographic features (such as river).
  – after the type of entity (such as temple).

• Different “languages” are used for description and communication.

• **Semiotic problem**: Articulation of the world and how entities are named is arbitrary.
Data Criticism Approach

Data
- Numerical data
- Spatial data (map)
- Visual data (photo)

Criticism

Inquiry

Historical Fact
- Spatial / visual evidences for historical studies

Digital Tools: Computational Algorithms and Databases
Textual and Data Criticism

**Textual Criticism**
- Well developed.
- Qualitative.
- Human reading and interpretation.
- Textual media.

**Data Criticism**
- To be explored.
- Quantitative.
- Computational algorithm.
- Spatial and visual media or multimedia.
Historical GIS and Map Criticism

Historical GIS

Source

Criticism by human

Digital Tools (GIS)

Analysis of Digital Data

Map Criticism

Source

Analysis of Sources

Criticism by human and computers

Digital Tools
Our Contribution

1. **Data criticism approach**: How to “read” data for reconstructing facts, under error, variation, and different conceptualization?

2. **Methodological commons**: map registration, 3D visual interpretation, linked data, ...

3. **Platform for linking entities**: data editing and managing platform for linking resources and combining evidences across sources.
Digital Silk Road Project

• Website
  – http://dsr.nii.ac.jp/

• Toyo Bunko Digital Archive
  – http://dsr.nii.ac.jp/toyobunko/

• Silk Road Maps
  – http://dsr.nii.ac.jp/geography/

Contact:
Asanobu KITAMOTO : kitamoto@nii.ac.jp