

Information Sciences Institute
Agent of Innovation: from visionary to viable

Geospatial Data Integration for High-Level Fusion

Craig Knoblock
knoblock@usc.edu

**University of Southern California
and
Geosemble Technologies**

High Levels of Information Needed for High-level Fusion

- **Information fusion has focused on low-level sensor fusion**
 - That is what the data can support
- **Yet analysts spend most of their time looking for information**
 - The majority of that on open source data
- **To do higher levels of fusion, requires higher levels of data**
 - News reports/intelligence articles about what is happening in an area
 - Identities of the businesses that are operating in a building
 - Explanations of changes seen in an image
 - Photos or videos of a location
 - ...

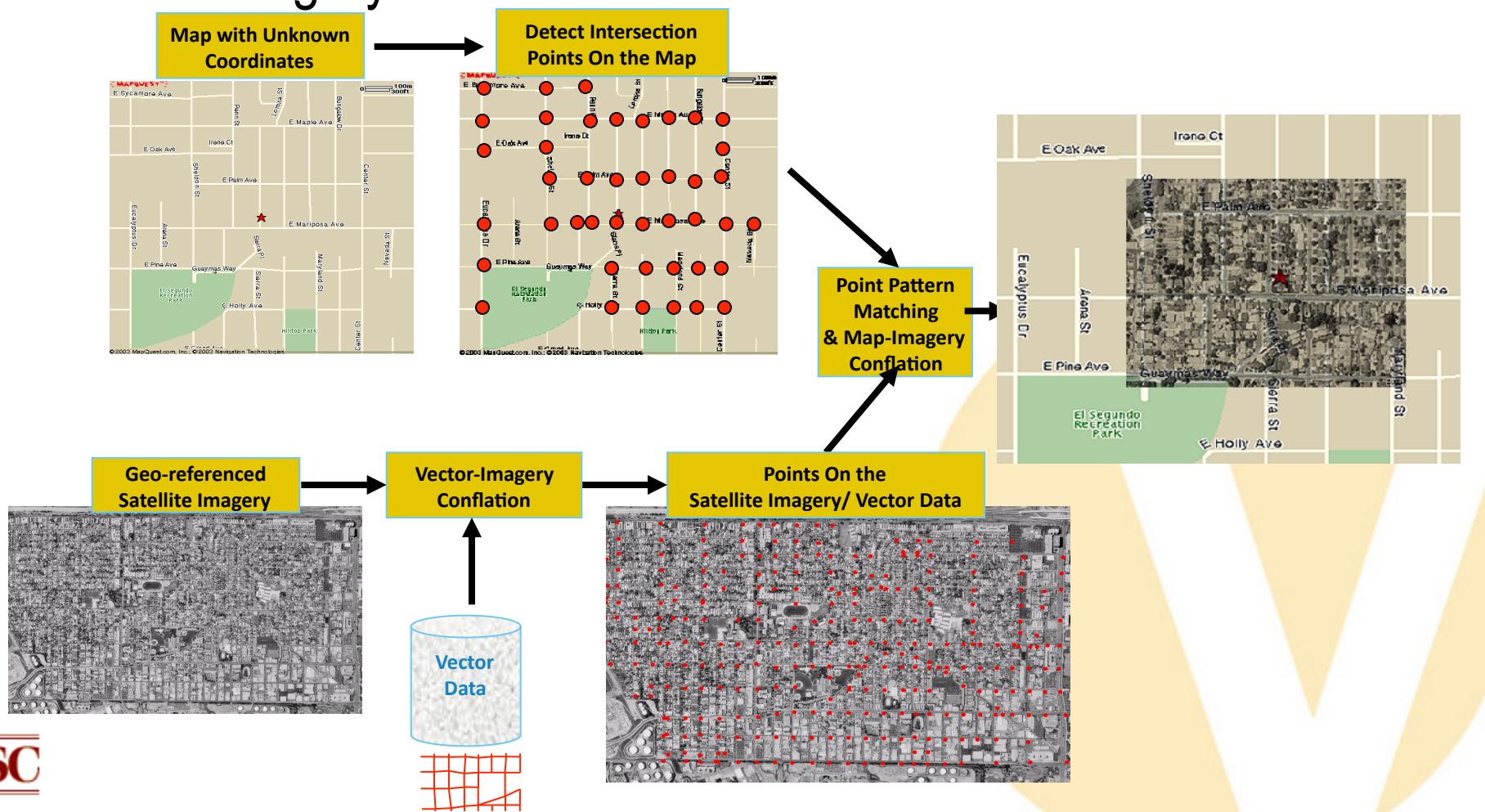


Talk Outline

- **Fusing online maps with imagery**
- **Linking news articles to imagery**
- **Exploiting telephone books to identify buildings in imagery**

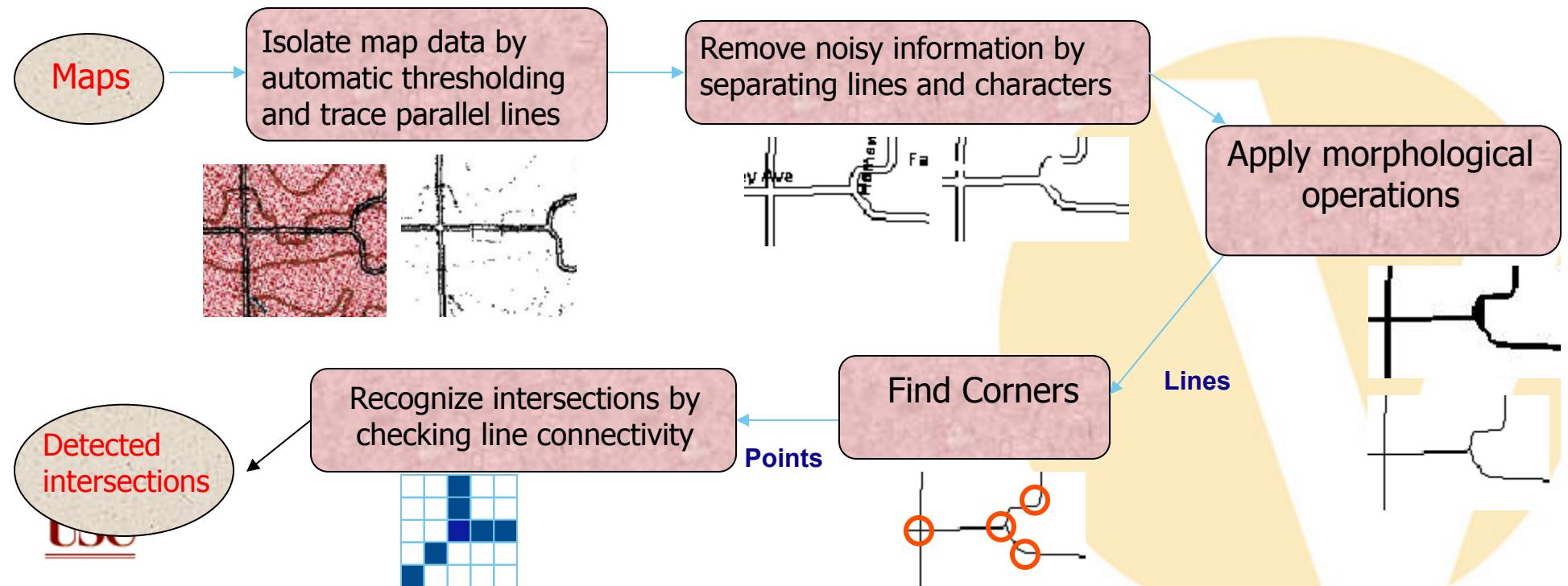
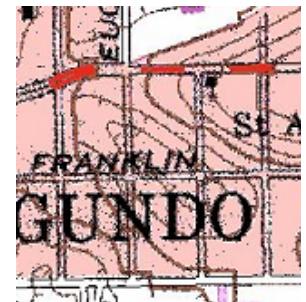
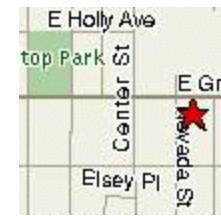
Fusing Maps & Imagery

- Exploit layout of intersections to accurately fuse maps with imagery



Finding Intersection Points on Maps

- Challenge is to identify intersection points automatically and accurately
 - Varying thickness of lines
 - Single-line map v.s. double-line map
 - Noisy information: symbols and alphanumeric characters

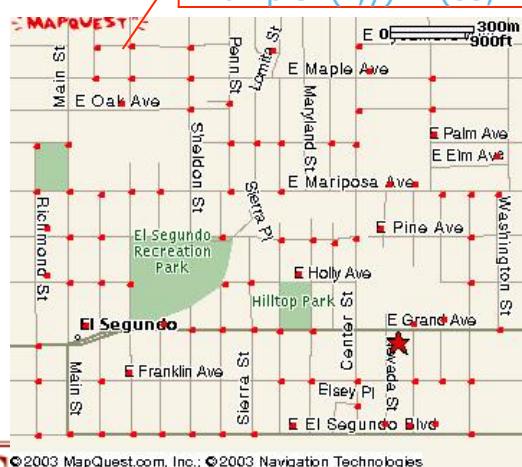


Point Pattern Matching

- Distribution of intersections is used to determine the relationship between a map and an image
- Find the mapping between these points to get a set of control point pairs
 - Find the transformation T between the layout (with relative distances) of the two point sets

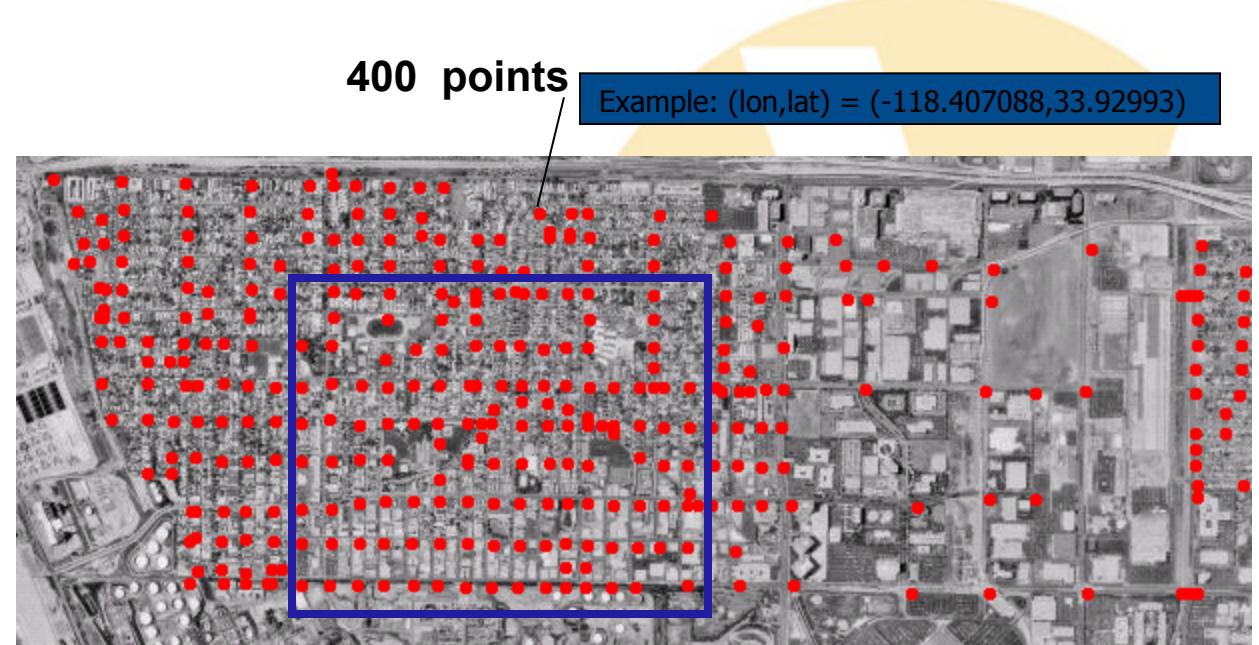
80 points

Example: (x,y) = (83,22)



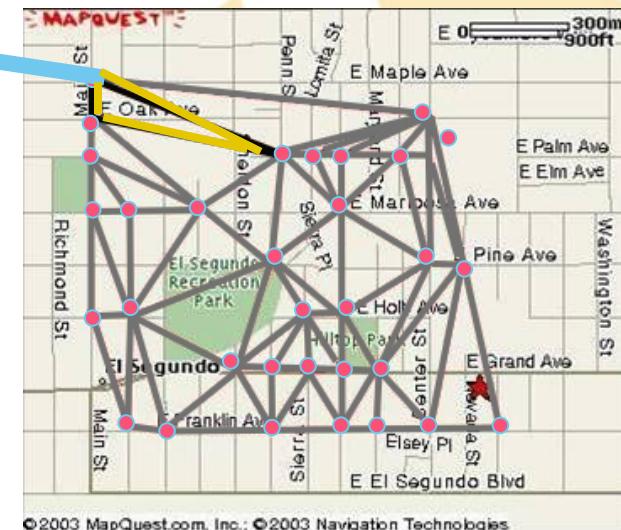
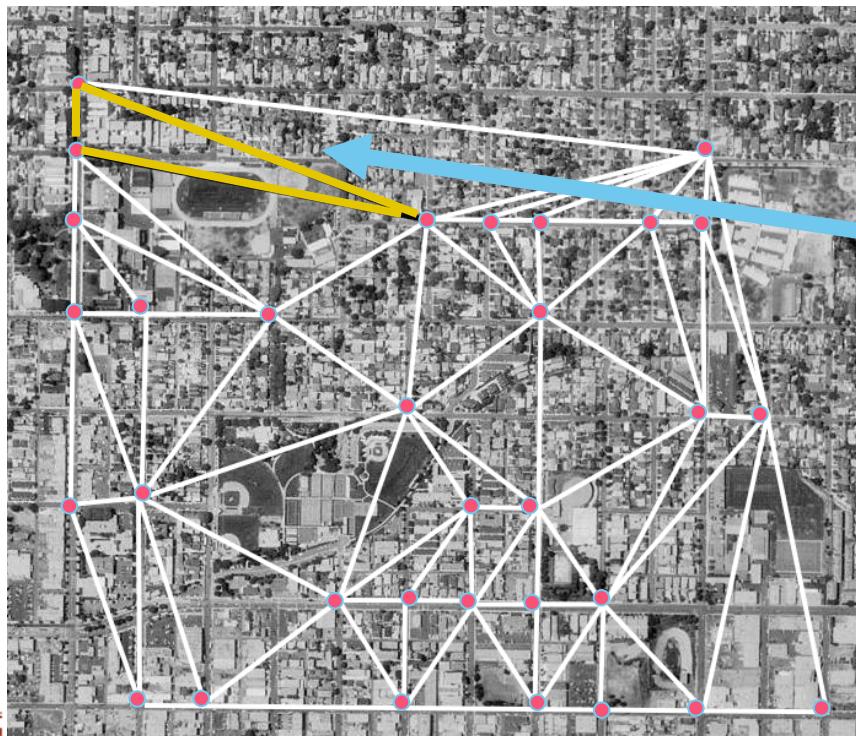
400 points

Example: (lon,lat) = (-118.407088,33.92993)



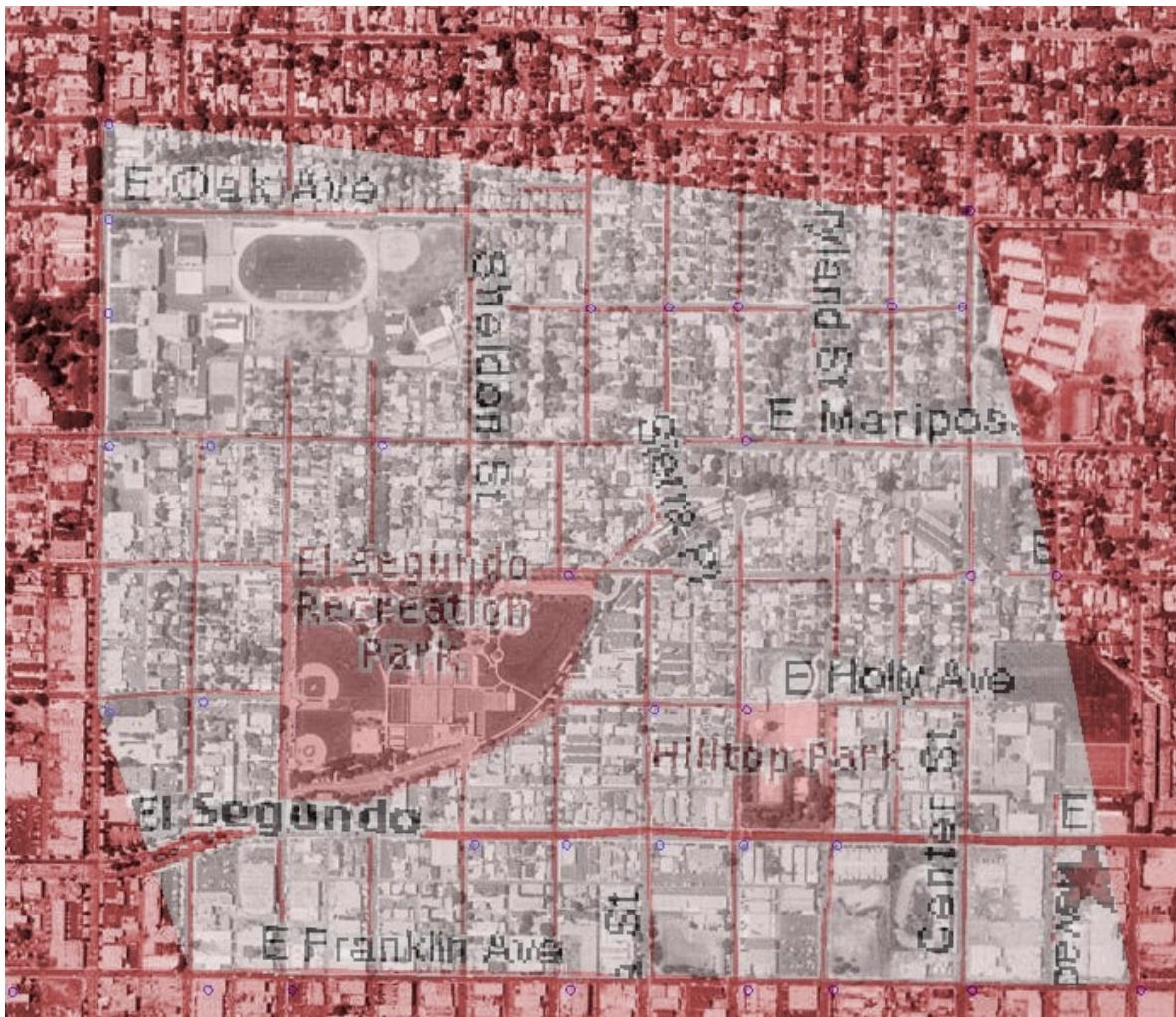
Aligning Maps and Imagery

- Using matched point pattern to align maps with imagery by rubber-sheeting



Aligning Maps and Imagery

- Using matched point pattern to align maps with imagery by rubber-sheeting



Talk Outline

- **Fusing online maps with imagery**
- **Linking news articles to imagery**
- **Exploiting telephone books to identify buildings in imagery**

Linking News to Imagery

- **Visualize various media in a geospatial context to better understand the context of an area**
 - Possible media includes databases, text documents, maps, audio recordings, video, etc
 - E.g., reports of illegal border crossing incidents, news articles from local papers, television news reports, etc.

The image shows a dual-pane application. On the left is a geospatial viewer with a satellite map of a industrial area. A callout bubble highlights a building labeled "Aircraft and Turbine Support (Company) 2701 N SHERIDAN RD, TULSA, OK". On the right is a web browser displaying a news article from the "Journal Record" about the same company's expansion plans. The browser window includes standard navigation buttons and a status bar indicating the URL and page title.

Approach

- **Link the online documents to an image by identifying the geographic features shown in the image, and then searching text sources to find documents relevant to the specific features**
 - The traditional approach is to process each possible document and associate the geographic references
 - *It requires preprocessing the documents and does not support fine-grained linking*
 - Our technology starts with an area, rather than a document
 - *It scales well with high accuracy and the linking to online documents can be done on-the-fly*



Traditional approach

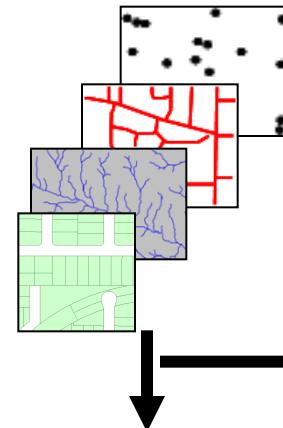


Our unique approach

Building Geospatial Knowledgebase

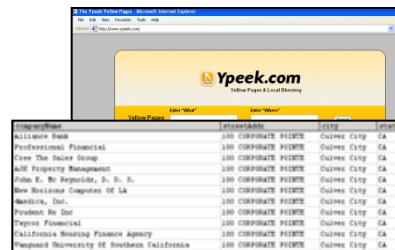
Available geospatial vector sources

- NGA gazetteers
- USGS gazetteers
- Parcel data
- US Census Tiger Files (e.g., road networks, hydrographic data)



Other sources relevant to geographic info.

- Phone books
- City business list



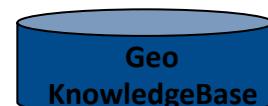
Parcel data

Mapping data to buildings

Businesses with geographic info.

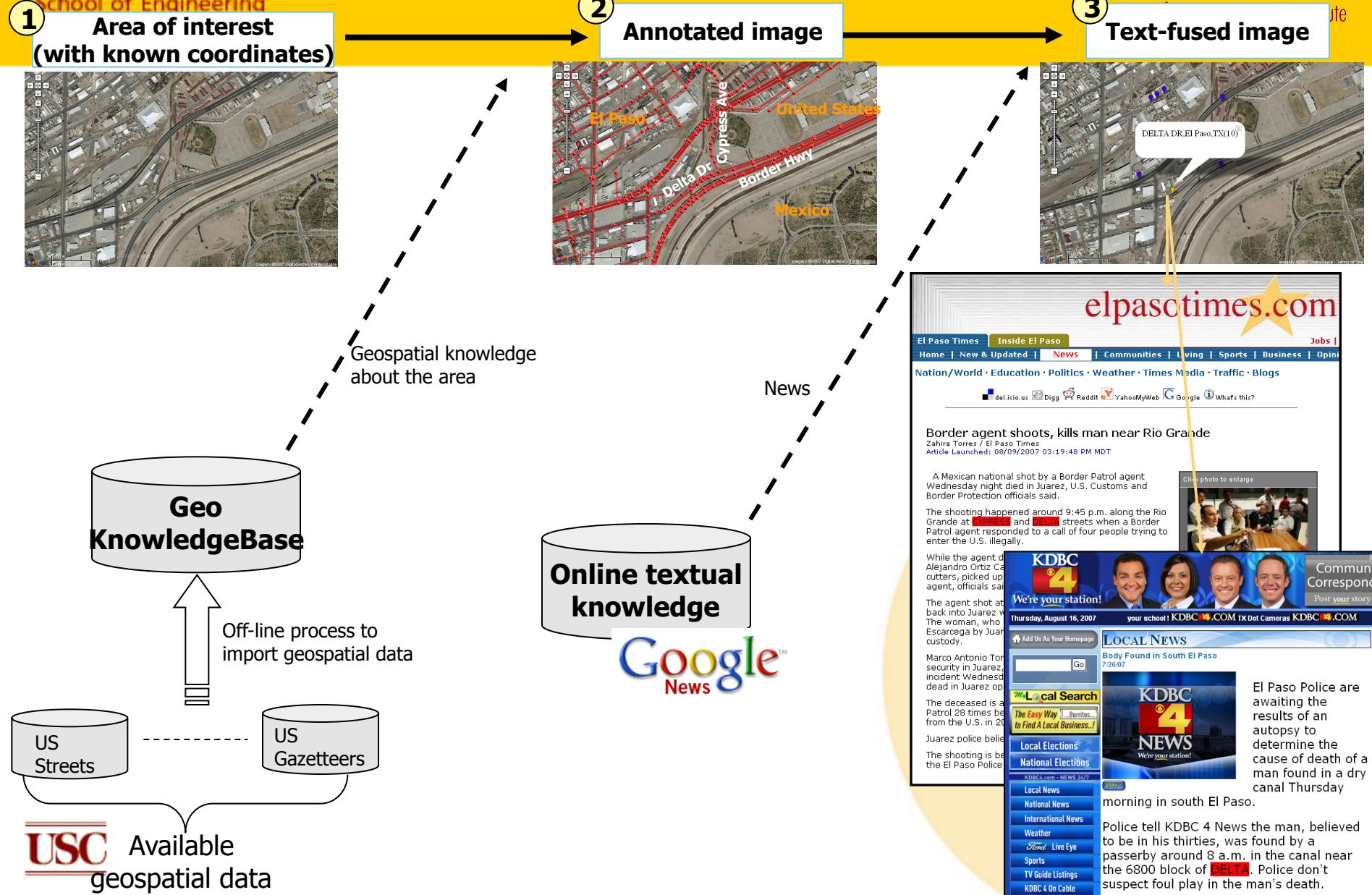
Integrating vector sources

Geospatial features



Name	Type	Coordinate
USC	Point	(34.3, -118.15)
Vine St.	Polygon	(34.21, -118.13) (34.22, -118.15) (34.25, -118.2)
...

Text-to-Image Linking



System GUI

Talk Outline

- **Fusing online maps with imagery**
- **Linking news articles to imagery**
- **Exploiting telephone books to identify buildings in imagery**

Building Identification (BID) Problem

Traditional Sources



Non-traditional Sources



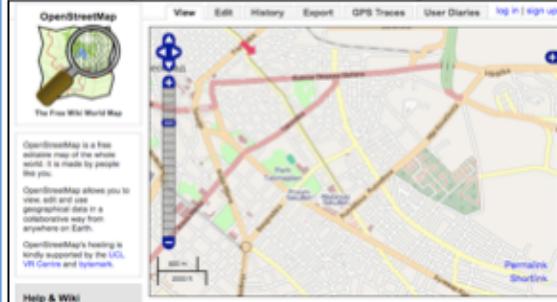
After

INHALT

BID Problem as a CSP

Vector Data

1. OpenStreetMap



- 2. KML, shapefiles etc
- 3. Manually Created

Information from Phone Book Data



Yellow and White Pages

Information from Image Or Map



CSP Model

CSP Solver

Street and address assignment for each building in an image

[Michalowski & Knoblock 2005]

- **Definition of a CSP**

- Given $P = (V, D, C)$

- V is a set of variables, $V = \{V1, V2, \dots, Vn\}$
 - D is a set of variable domains (domain values)
 - $D = \{DV1, DV2, \dots, DVn\}$
 - C is a set of constraints, $C = \{C1, C2, \dots, Cl\}$
 - $CVa, Vb, \dots, Vi = \{(x, y, \dots, z)\} \subseteq DVa \times DVb \times \dots \times DVi$

- Query: can we find a value for each variable such that all constraints are satisfied?

- **Useful for modeling & solving combinatorial problems**

Phone Book Linked to Streets

Street Name: Uskocka

White Pages Data		Yellow Pages Data	
Phone Number:	011 2631495	Phone Number:	011 202 9292
Address:	USKOČKA 7, BEOGRAD	Address:	Beograd, Vračar, Uskočka 6
Name:	ALABANDA, VLADETA	Business Name:	Državna Lutrija Srbije
Phone Number:	011 2621250	Category:	igre na sreću
Address:	USKOČKA 7/002, BEOGRAD	Phone Number:	011 627 605
Name:	ALABANDA, VLADETA	Address:	Beograd, Uskočka 4
Phone Number:	011 2620523	Business Name:	Fun casino
Address:	USKOČKA 7/4, BEOGRAD	Category:	kazina
		WebPage:	

STEP 1: SELECT AN AREA
Select an area on map using mouse

STEP 2: IDENTIFY STREETS

- Draw Street Manually ADD MORE
- Label Streets Manually
- Import KML File
- View Maps Online
- Load OpenStreetMaps Data

STEP 3: IDENTIFY BUILDING LOCATIONS

- Draw Points DONE
- Draw Polygon ADD MORE
- Import KML File
- Process LIDAR Data
- Import WikiMapia Layer

STEP 4: IDENTIFY BUILDING ADDRESSES

- Load Phone Book Data
- Predict Intersection Buildings
- Exploit Constraints
- Exploit Map Sources

OTHER OPTIONS

EXPORT KMZ
SAVE CURRENT SESSION DATA
LOAD PREVIOUS SESSION'S DATA

Constraint Reasoning to Link Data to Buildings

Map Satellite Hybrid

1. 7 Cara Lazara
2. 9 Cara Lazara
3. 3 Nikole Spasica

2 Uskocka
4 Uskocka

Google Imagery © 2009 DigitalGlobe, GeoEye Terms of Use

STEP 1: SELECT AN AREA
Select an area on map using mouse

STEP 2: IDENTIFY STREETS

Draw Street Manually [ADD MORE](#)
 Label Streets Manually
 Import KML File
 View Maps Online
 Load OpenStreetMaps Data

STEP 3: IDENTIFY BUILDING LOCATIONS

Draw Points [DONE](#)
 Draw Polygon [ADD MORE](#)
 Import KML File
 Process LIDAR Data
 Import WikiMapia Layer

STEP 4: IDENTIFY BUILDING ADDRESSES

Load Phone Book Data
 Predict Intersection Buildings
 Exploit Constraints
 Exploit Map Sources

OTHER OPTIONS

[EXPORT KML'S](#)
[SAVE CURRENT SESSION DATA](#)
[LOAD PREVIOUS SESSION'S DATA](#)

Exploit Maps for Disambiguation

The image consists of three main parts. On the left is a satellite map of a city street with several green location markers and two yellow circles highlighting specific points. A callout box contains a list of addresses and a link to a website, with a green arrow pointing to a detailed map on the right. The detailed map shows a specific intersection with a blue circle around it, labeled "SAT". Below the maps is a user interface with sections for "ADDRESSES" and "OTHER OPTIONS".

Address: 3 Nikole Spasica
Website: www.satoz.net

Map:

Map details:
Address: 3 Nikole Spasica
Website: www.satoz.net

ADDRESSES

- Load Phone Book Data
- Predict Intersection Buildings
- Exploit Constraints
- Exploit Map Sources

OTHER OPTIONS

EXPORT KMLs
SAVE CURRENT SESSION DATA
LOAD PREVIOUS SESSION'S DATA

Results Propagated to Further Reduce Ambiguity

Map Satellite Hybrid

10 Cara Lazara

3 Nikole Spasica

Google

Copyright ©2009 DigitalGlobe. Google, Terms of Use

STEP 1: SELECT AN AREA

Select an area on map using mouse

STEP 2: IDENTIFY STREETS

Draw Street Manually [ADD MORE](#)

Label Streets Manually

Import KML File

View Maps Online

Load OpenStreetMaps Data

STEP 3: IDENTIFY BUILDING LOCATIONS

Draw Points [DONE](#) [ADD MORE](#)

Draw Polygon

Import KML File

Process LIDAR Data

Import WikiMapia Layer

STEP 4: IDENTIFY BUILDING ADDRESSES

Load Phone Book Data

Predict Intersection Buildings

Exploit Constraints

Exploit Map Sources

OTHER OPTIONS

[EXPORT KML](#)

[SAVE CURRENT SESSION DATA](#)

[LOAD PREVIOUS SESSION'S DATA](#)

Conclusions

- **Open source data is a critical source of information for analysis and decision making**
- **I described how we can fuse three types of information for better understanding of a given region:**
 - Raster maps
 - New articles
 - Telephone books
 - All of these data sources are readily available for most of the world
- **There are many other types of data (e.g., photos, videos, text documents) and many ways this information can be integrated**
- **All of these sources are key to making progress in high-level information fusion**