Image-Based Buildings and Facades

Peter Wonka
Associate Professor of Computer Science
Arizona State University

Daniel G. Aliaga
Associate Professor of Computer Science
Purdue University

Course: Modeling 3D Urban Spaces using Procedural and Simulation-based Techniques

Challenge

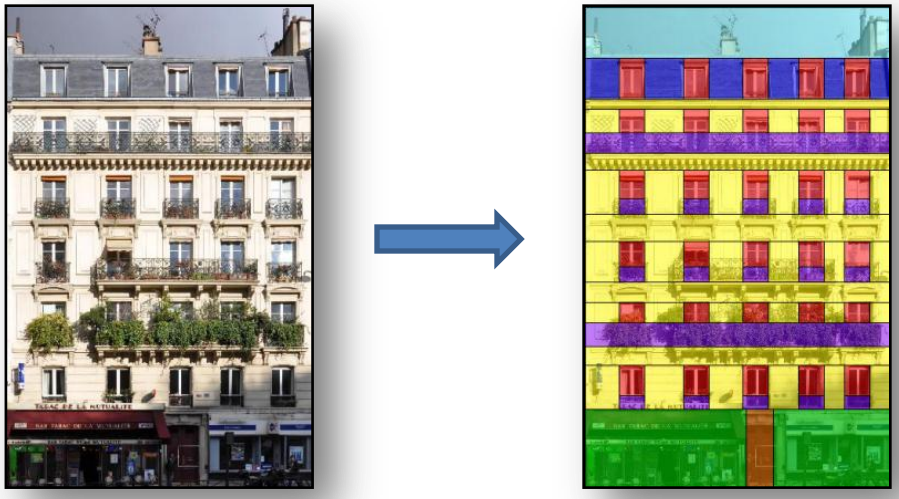
- Generate facades/building models starting with images (and maybe some meta-data)
- Common strategy is to assume specific styles
- Methods loosely divided into:
  - Non-interactive, and
  - Interactive

Non-Interactive Methods

- Work comes from graphics/vision community
- Facades:
  - Parsing Image Facades with Reinforcement Learning, O. Teboul, I. Kokkinos, P. Katsourakis, L . Simon, N. Paragios, CVPR 2011
- Buildings:
  - 2.5D Dual Contouring: A Robust Approach to Creating Building Models from Aerial LiDAR Point Clouds, Zhou and Neumann, ECCV 2010
  - Building Reconstruction using Manhattan-World Grammars, C. Vanegas, D. Aliaga, B. Benes, CVPR 2010
  - Style Grammars for Interactive Visualization of Architecture, D. Aliaga, P. Rosen, D. Bekins, TVCG 2007 (*towards interactive…*)

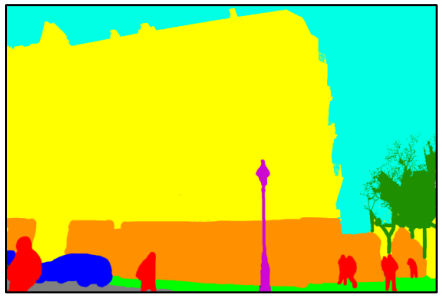
Image-Based Buildings and Facades

- **Parsing Image Facades with Reinforcement Learning**
  - Teboul, Kokkinos, Katsourakis, Simon, Paragios
- CVPR 2011



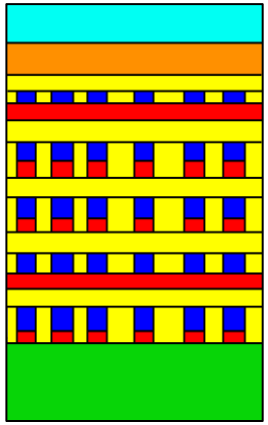
Overall Goal I: Semantic Segmentation of Urban Scenes

- Image Parsing



building
pedestrian
street lamp
pavement
road
cars
trees
sky
shops

- Cropped/Rectified Building Images: ‘Facade Parsing’



wall
balconnies
shop
windows
sky
roof

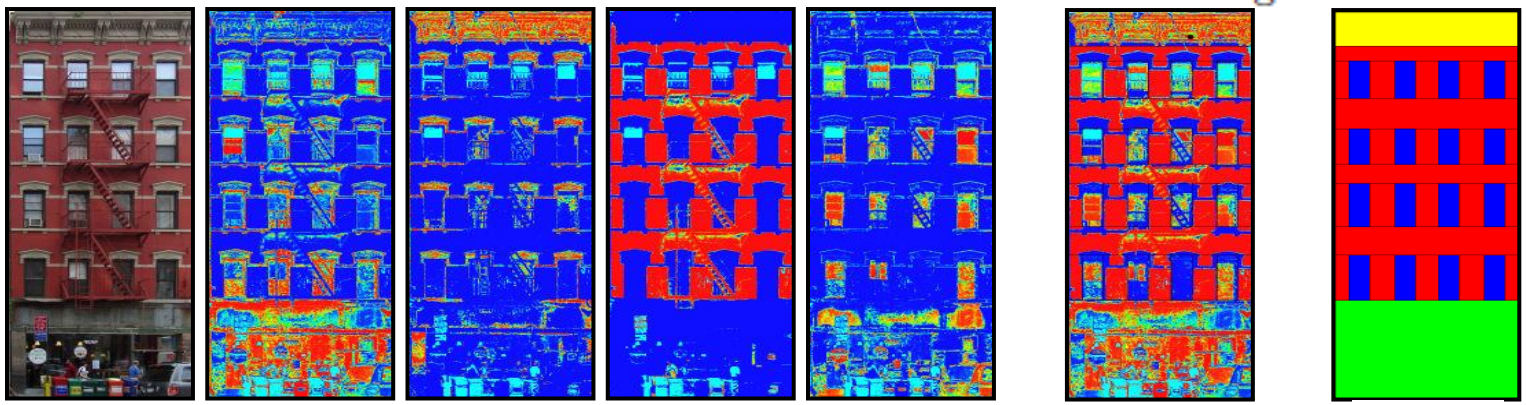
Overall Goal II: Image-based Procedural 3D Models

- Based on 2D parsing + simple extrude and insertion rules turn 2D to 3D…

Problem Statement

- Input: image $I(x, y)$
- Output: labelling $l(x, y)$
- Pixel-level classification function: $m(c, x, y) = p(c|I(x, y))$
- Objective: $C(l) = \sum_{x,y} m(l(x, y), x, y)$

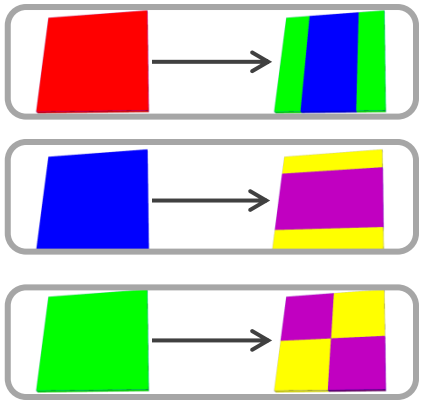
- Wanted: $l^* = \arg \max_{l:\text{building}} C(l)$

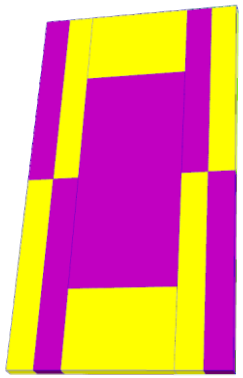


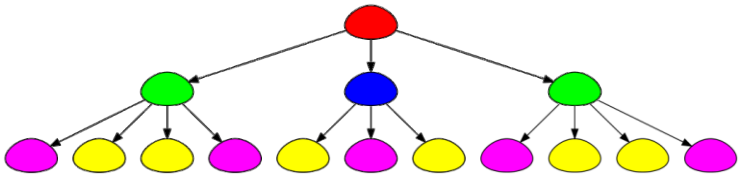
original $m(shop)$ $m(roof)$ $m(wall)$ $m(window)$ $m(l(x, y), x, y)$ $l(x, y)$

Shape Grammars: Recursive Derivation of Labelling

- Top level: axiom
- Recursive application of shape operators
  - Partition domain and assign label to each part
- Terminals: semantic labels (e.g., window, door etc)

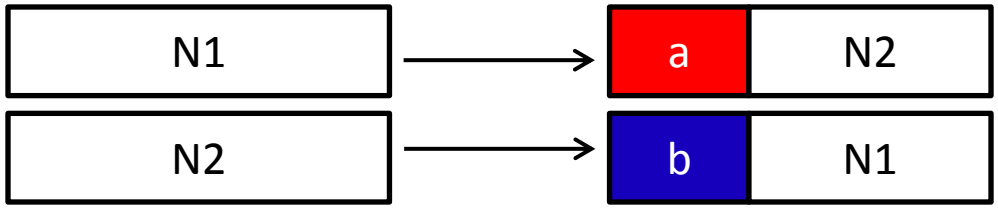




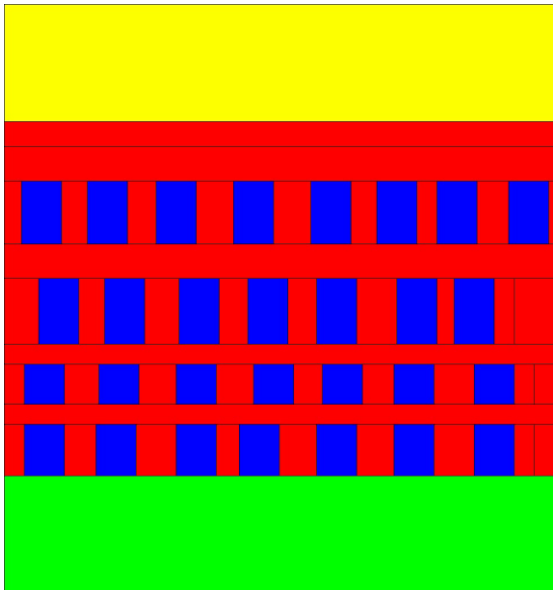


Binary Split Grammars

– Binary:

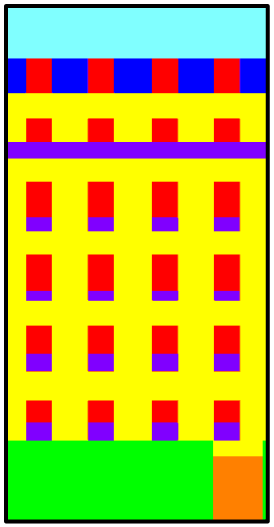


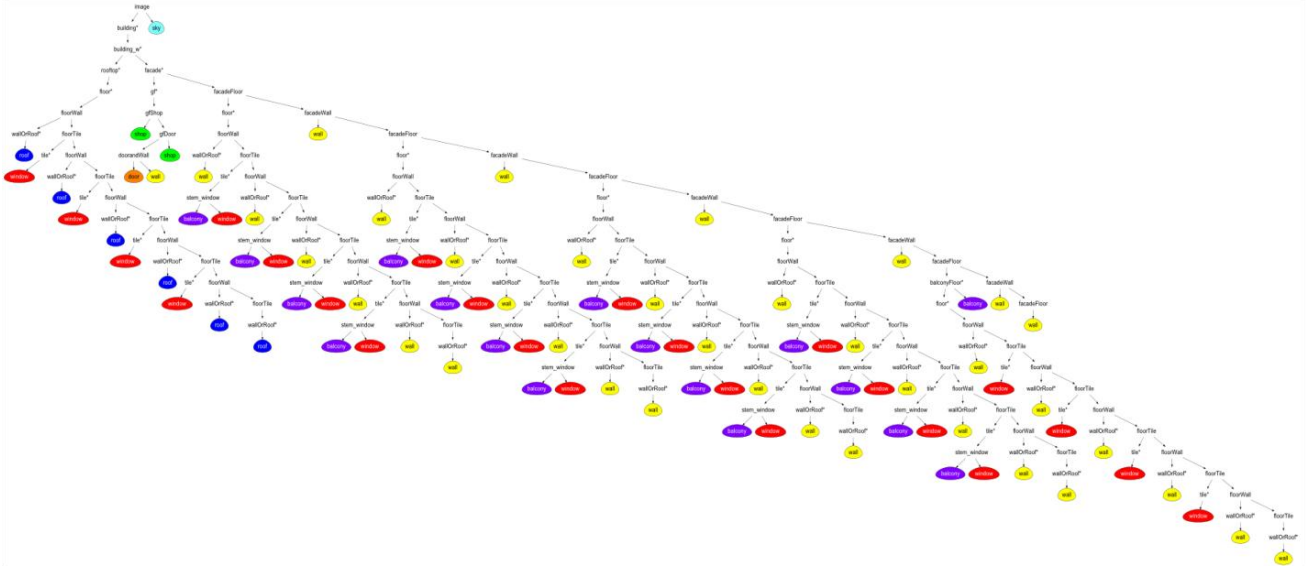
– Split: one dimension at a time



Challenges

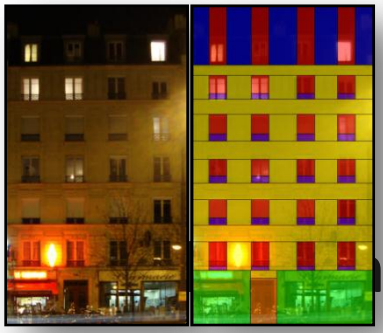
- Joint optimization: topology + geometry
- Enforce the result to be in the language of the grammar: $C \in L(G)$
- High and unknown dimensionality: $card\big(L(G)\big)$ up to 1 gogol! (10^{100})



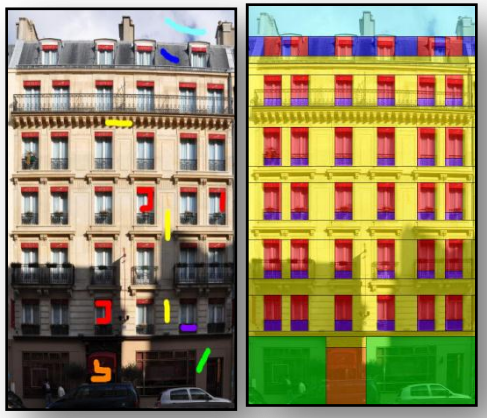


Example Results

- Natural Occlusions



- Cast Shadows



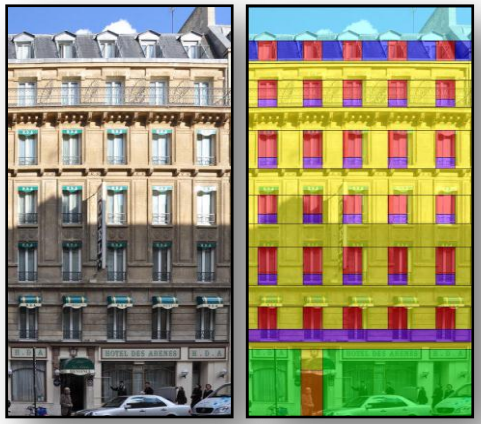
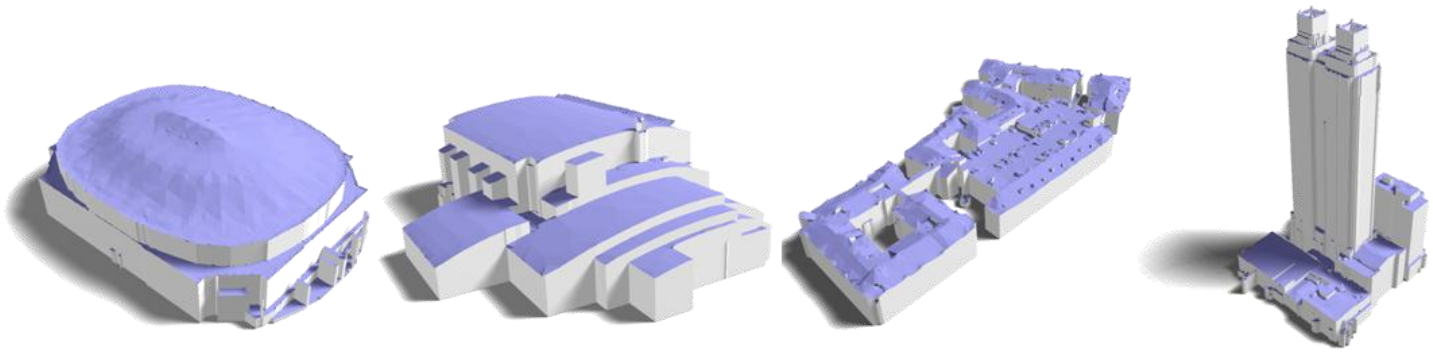


Image-Based Buildings and Facades

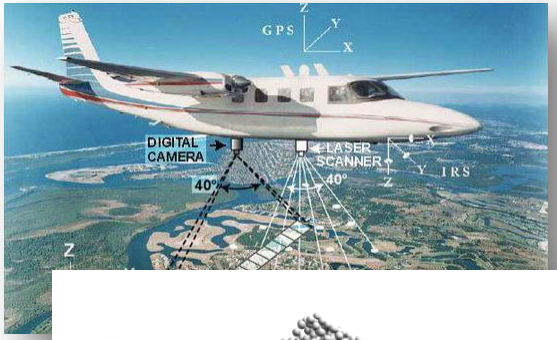
- **2.5D Dual Contouring: A Robust Approach to Creating Building Models from Aerial LiDAR Point Clouds**
  - Zhou, Neumann
- ECCV 2010

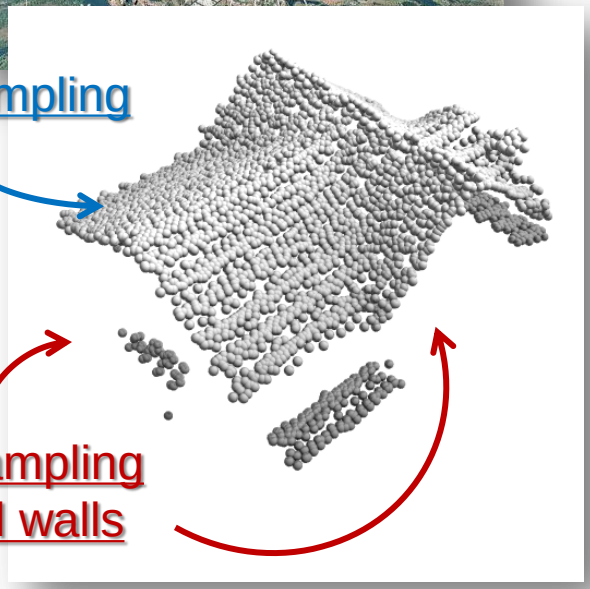


USC
UNIVERSITY
OF SOUTHERN
CALIFORNIA

2.5D Dual Contouring :

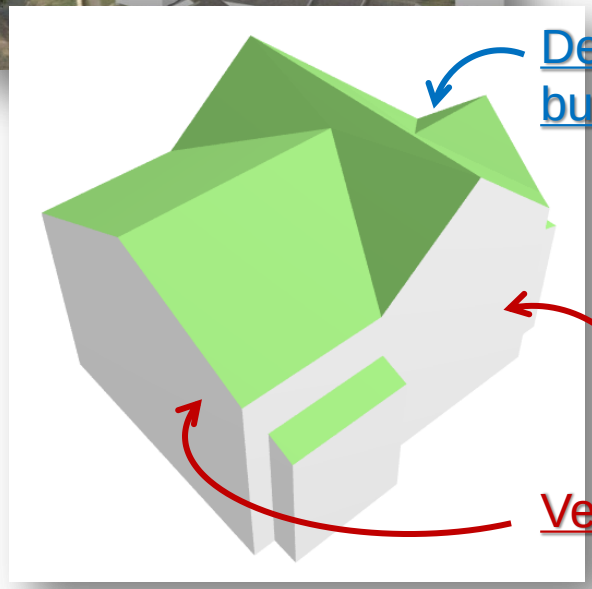
A Robust Approach to Creating Building Models from Aerial LiDAR Point Clouds





Dense sampling on roofs

Sparse sampling on vertical walls



Detailed building roofs

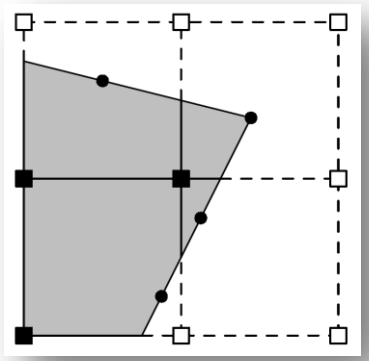
Vertical walls

2.5D Dual Contouring:

A Robust Approach to Creating Building Models from Aerial LiDAR Point Clouds

Classic dual contouring

2D

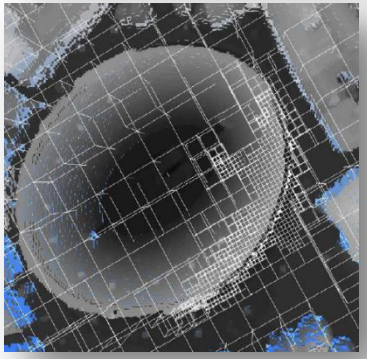


[Ju et al '02]

Create one 2D point per 2D quadtree cell

Cons: no 3D output

3D

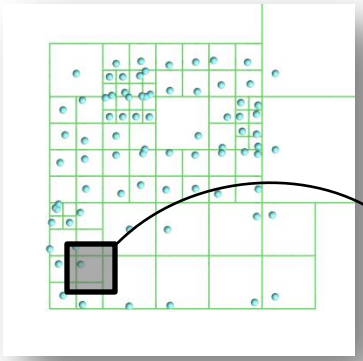


[Ju et al '02]
[Fiocco et al '05]

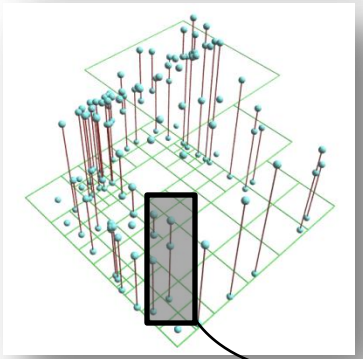
Create one 3D point per 3D octree cell

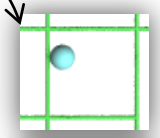
Cons: no 2.5D characteristic

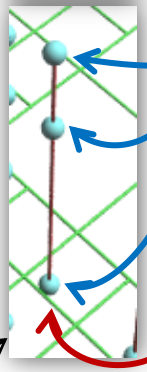
2.5D dual contouring



- Use 2D quadtree as supporting data structure
- Create one ***hyper-point*** per cell



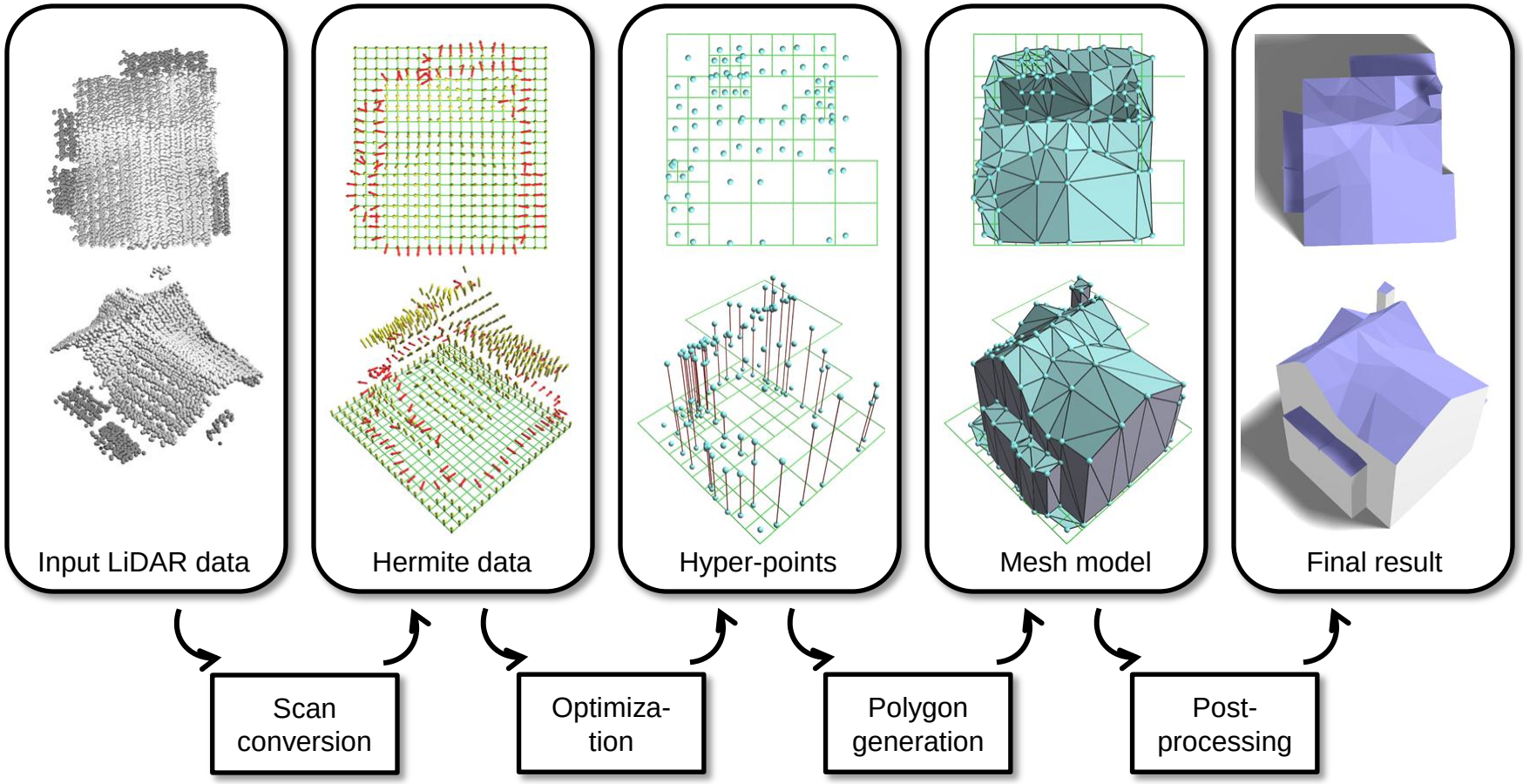




Different z values

Same x-y coordinates

Pipeline


Input LiDAR data
Hermite data
Hyper-points
Mesh model
Final result
Scan conversion
Optimiza-tion
Polygon generation
Post-processing

USC
UNIVERSITY
OF SOUTHERN
CALIFORNIA

Video

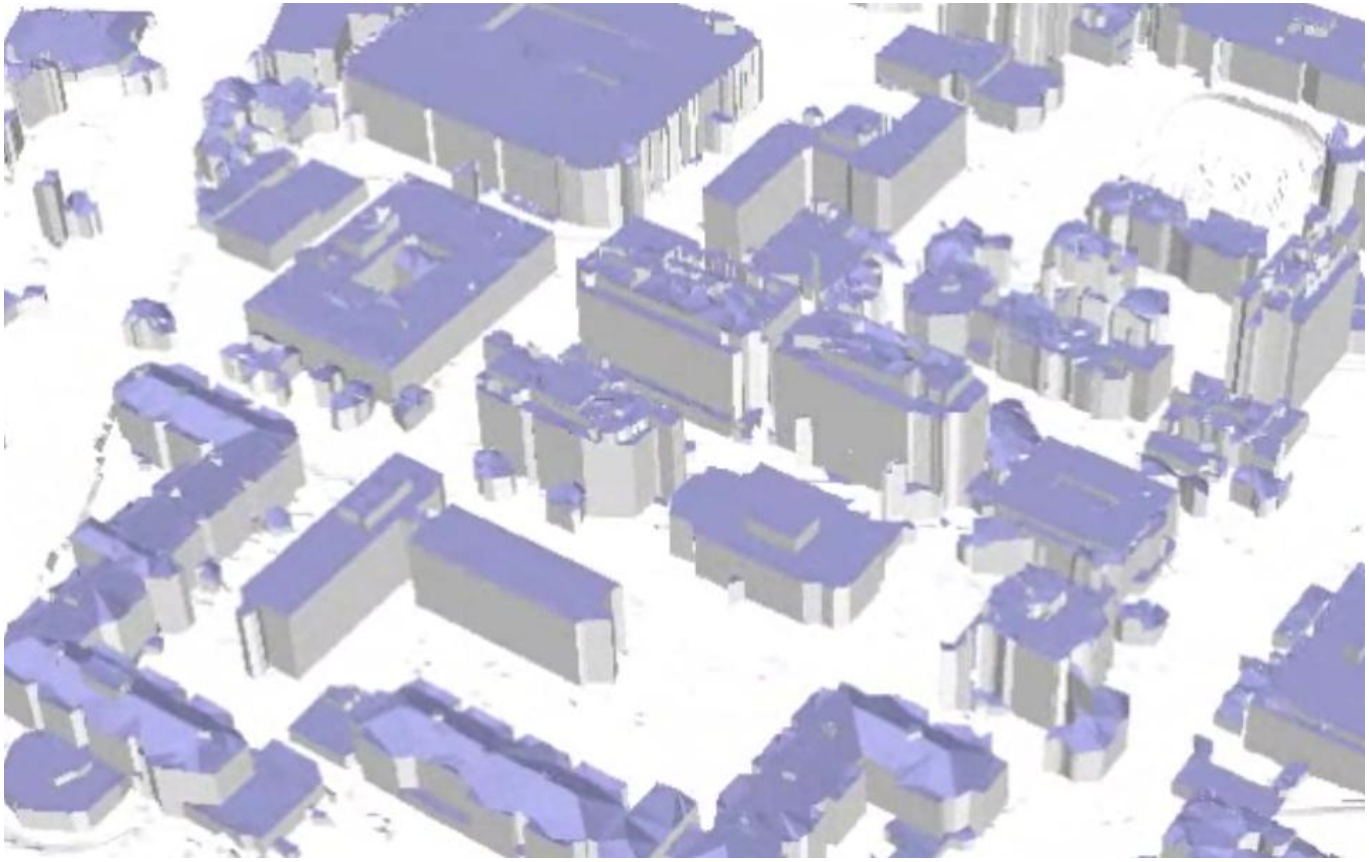
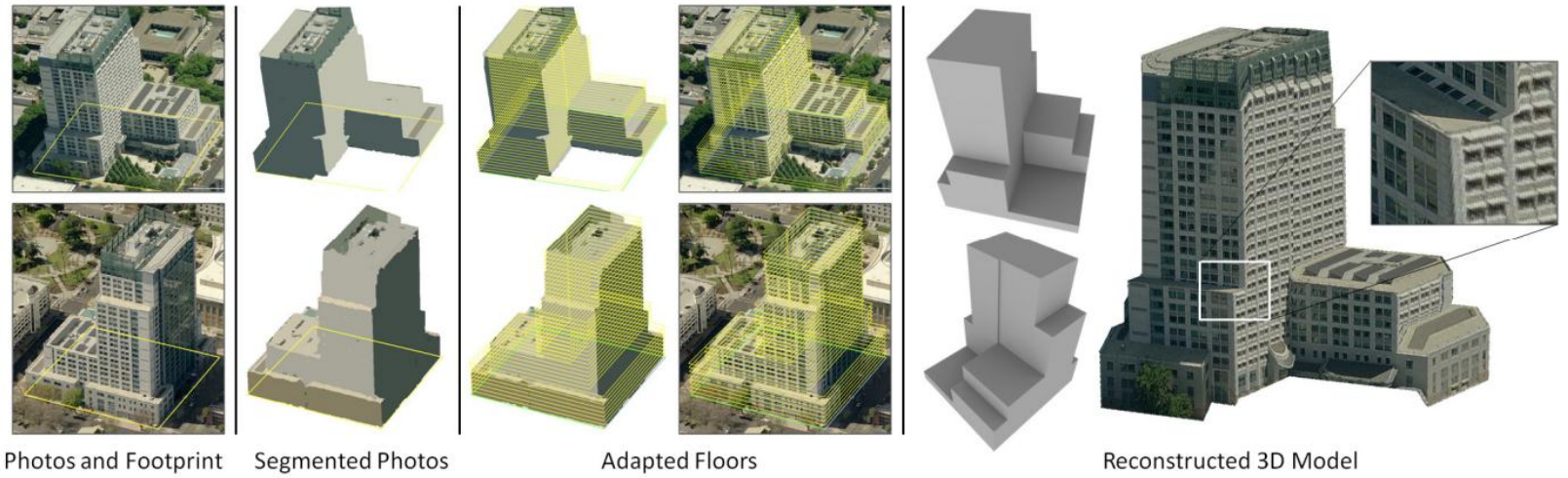


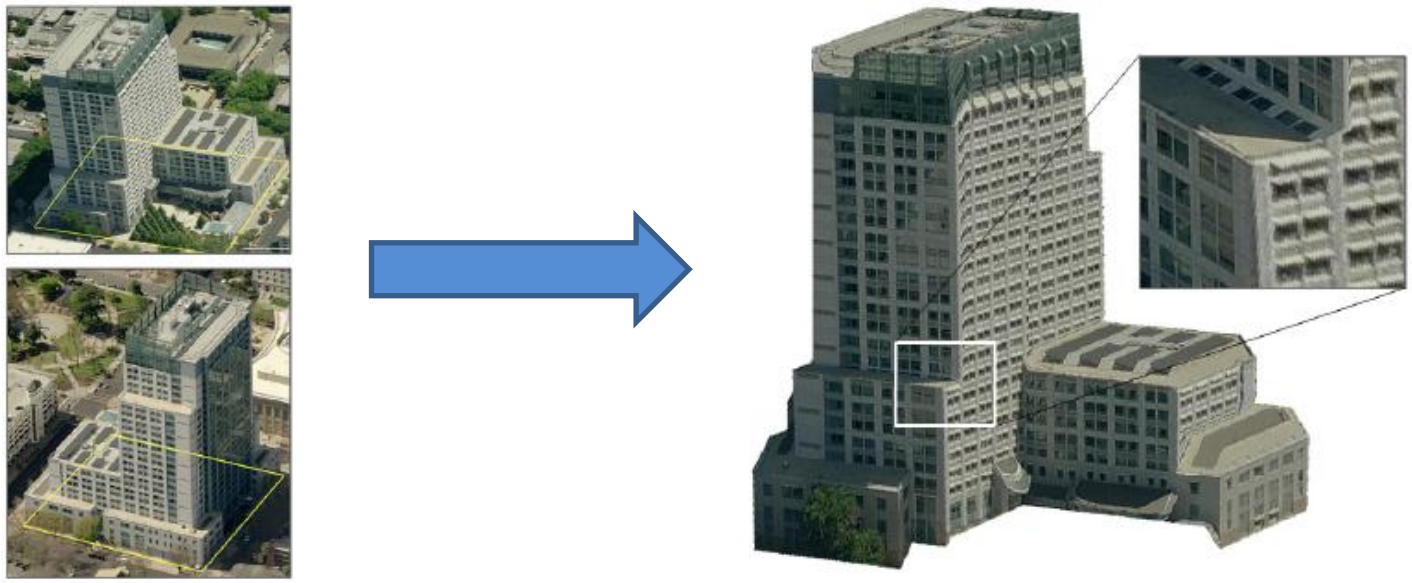
Image-Based Buildings and Facades

- **Building Reconstruction using Manhattan-World Grammars**
  - Vanegas, Aliaga, Benes
- CVPR 2010



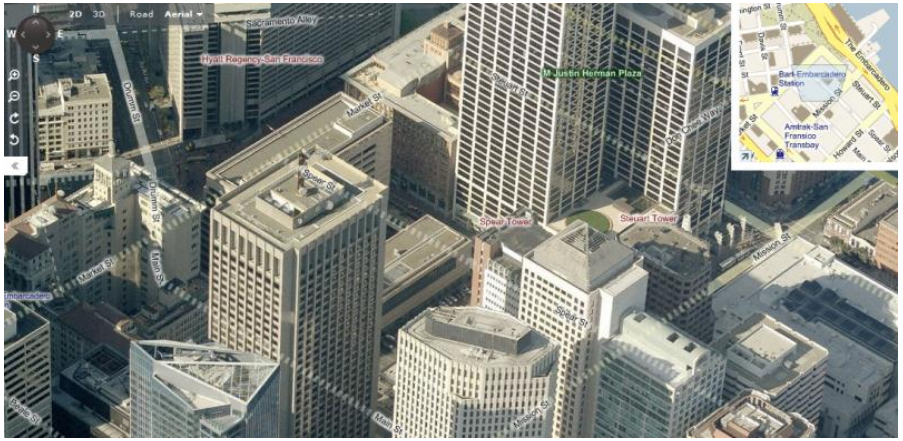
Manhattan World Buildings

- An example of automatic inverse procedural modeling of Manhattan-like buildings



Manhattan World Buildings

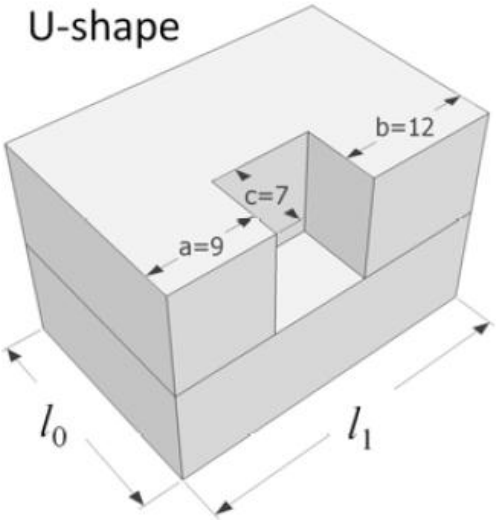
- Images of urban areas and street networks all around the world are readily available (e.g., Bing Maps, Google Maps…)



Manhattan World Observation

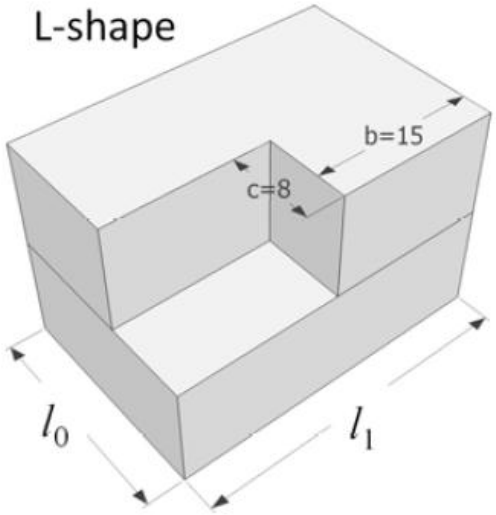
- Manhattan World (MW) transitions:

U-shape



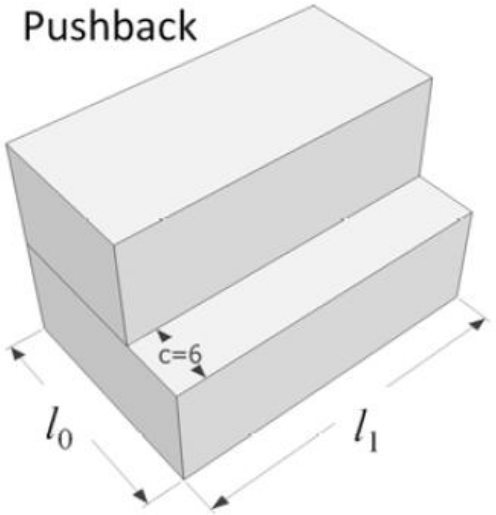
$f(l_0) - f(a) - f(c) + f(l_1\text{-}a\text{-}b) + f(c) - f(b) - f(l_0) - f(l_1)$

L-shape



$f(l_0\text{-}c) - f(l_1\text{-}b) + f(c) - f(b) - f(l_0) - f(l_1)$

Pushback



$f(l_0\text{-}c) - f(l_1) - f(l_0\text{-}c) - f(l_1)$

Manhattan World Observation

- These three transitions can be represented with just one rule that we call **generalized rewriting rule** (GRR):

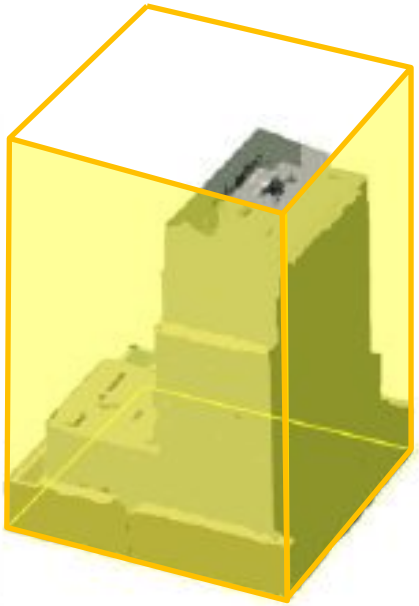
$$f(l) \rightarrow f(a) - f(c) + f(l - a - b) + f(c) - f(b)$$

Our Approach

- Generate initial 3D building envelope
  - Extrude bounding box of the building footprint extracted from GIS data

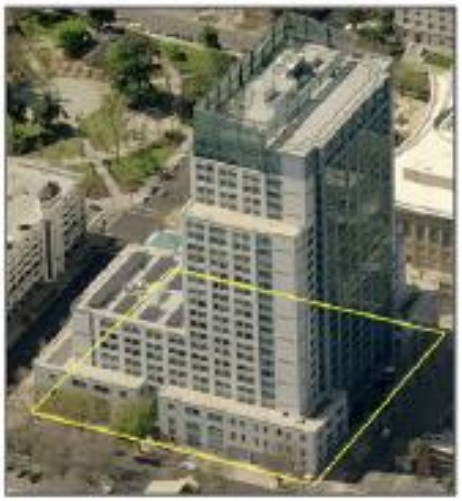
Our Approach

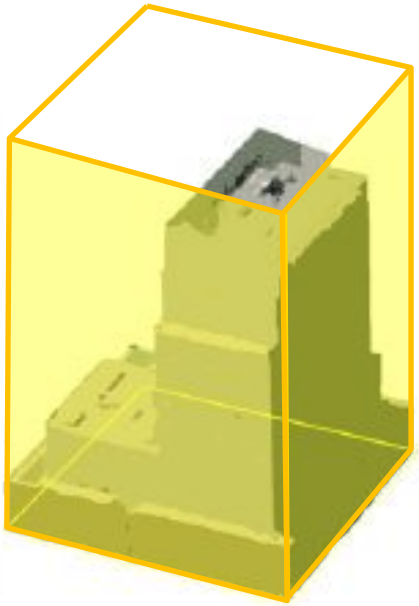
- Generate initial 3D building envelope
  - Extrude bounding box of the building footprint extracted from GIS data

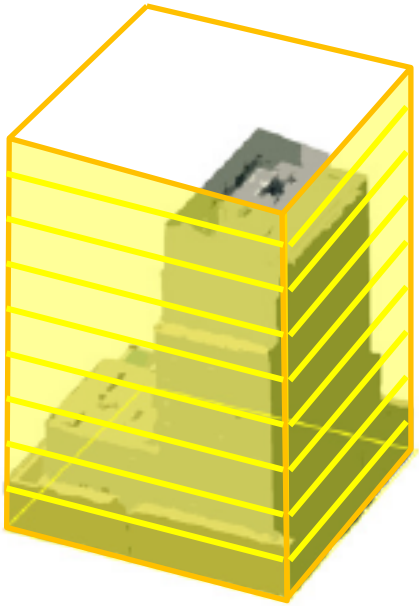


Our Approach

- Generate initial 3D building envelope
- Divide initial model into a sequence of floors

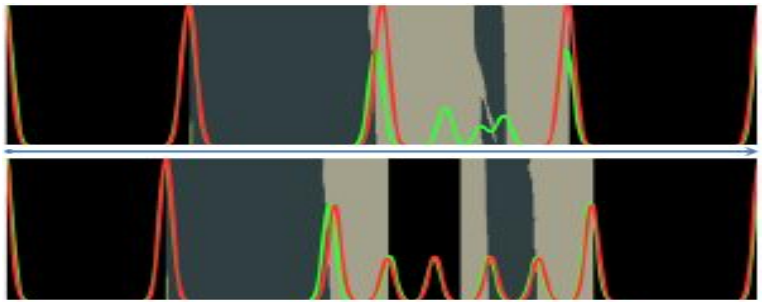






Our Approach

- Generate initial 3D building envelope
- Divide initial model into a sequence of floors
- For each floor, automatically modify geometry to match that observed in the images
  - Assume facades have different relative intensities
  - Use the generalized rewriting rule



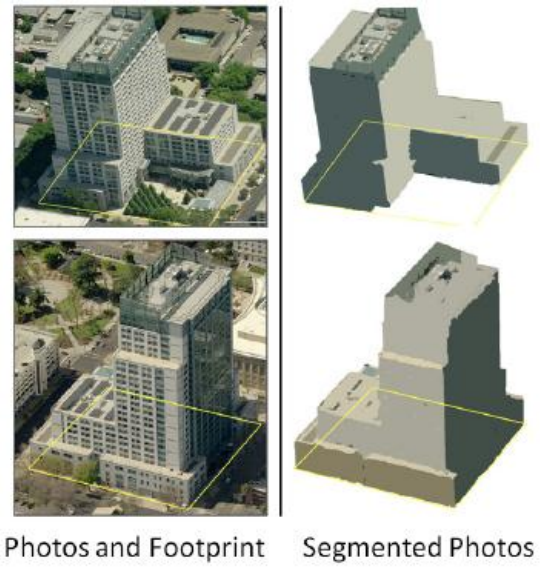
Our Approach

- Example result

Photos and Footprint

Our Approach

- Example result

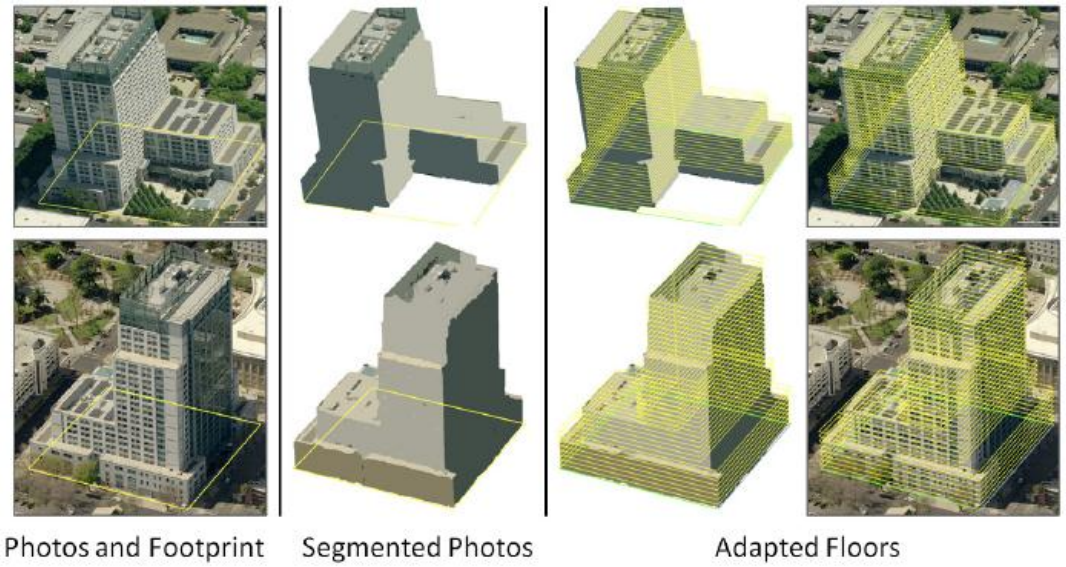


Photos and Footprint

Segmented Photos

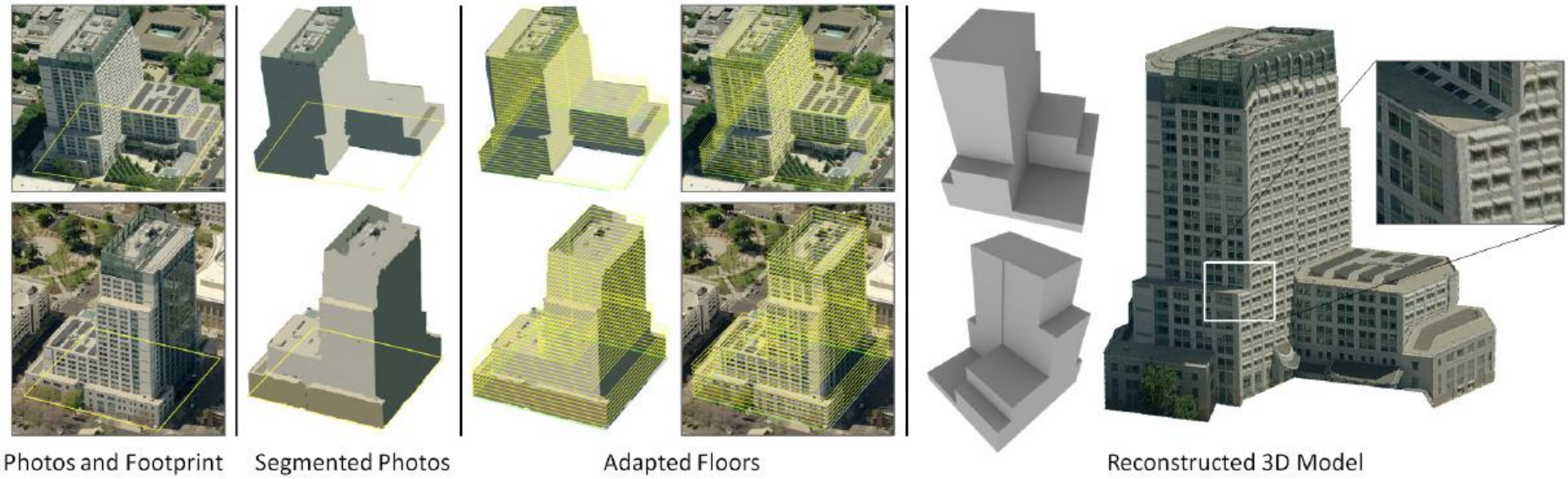
Our Approach

- Example result



Our Approach

- Example result



Manhattan World Buildings

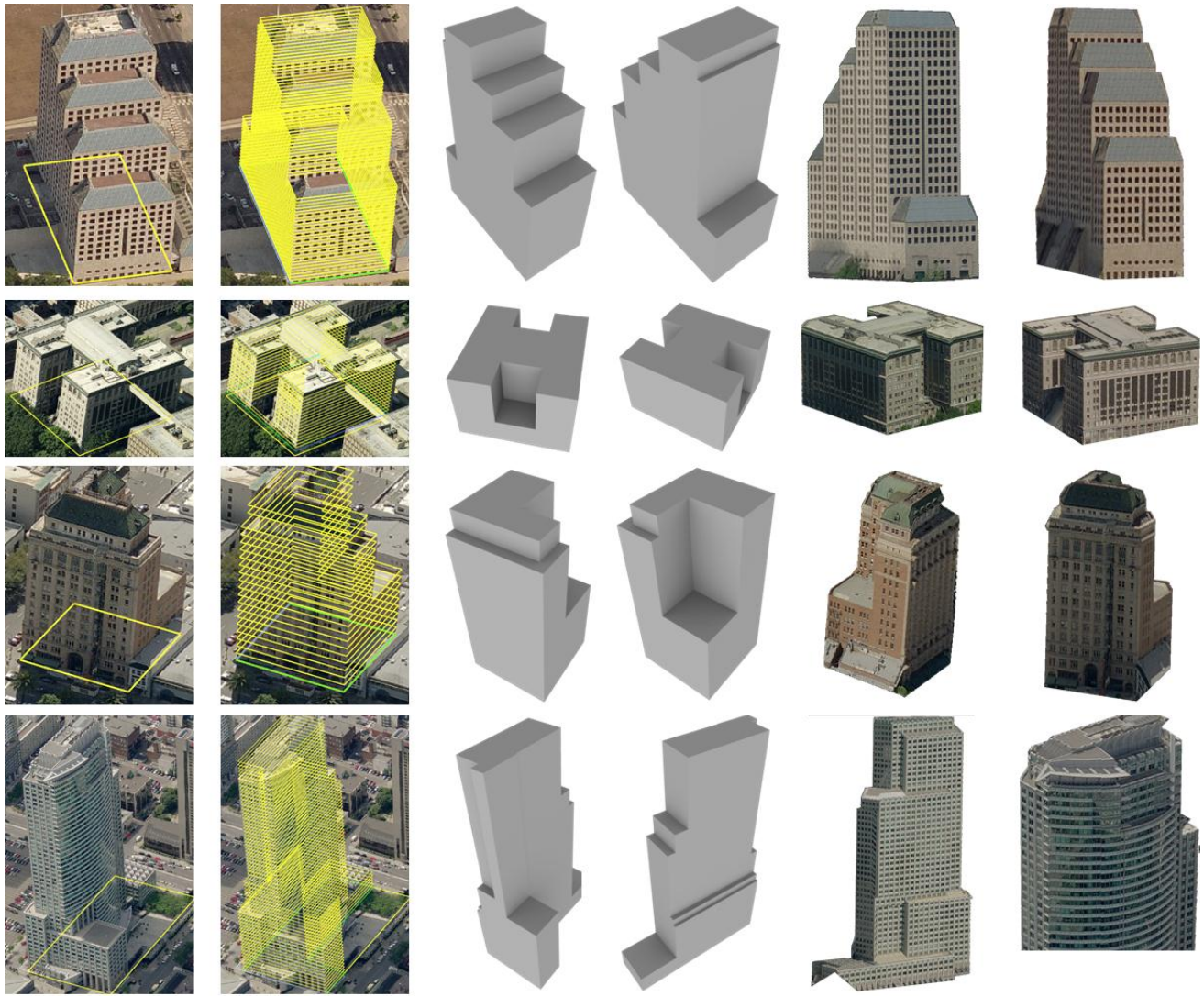


Image-based Buildings and Facades

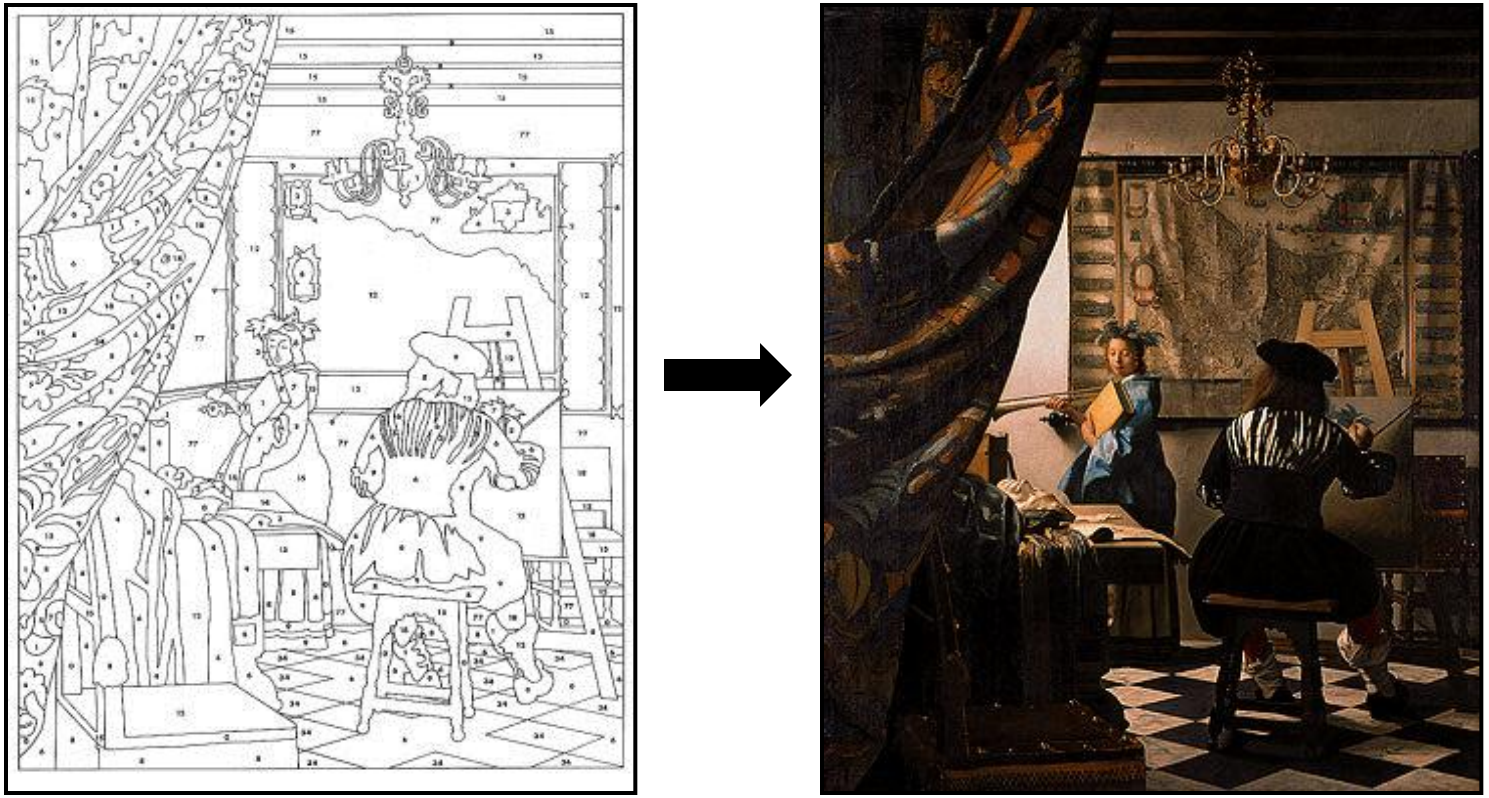
- **Style Grammars for Interactive Visualization of Architecture**

  - Aliaga, Rosen, Bekins

- IEEE TVCG 2007

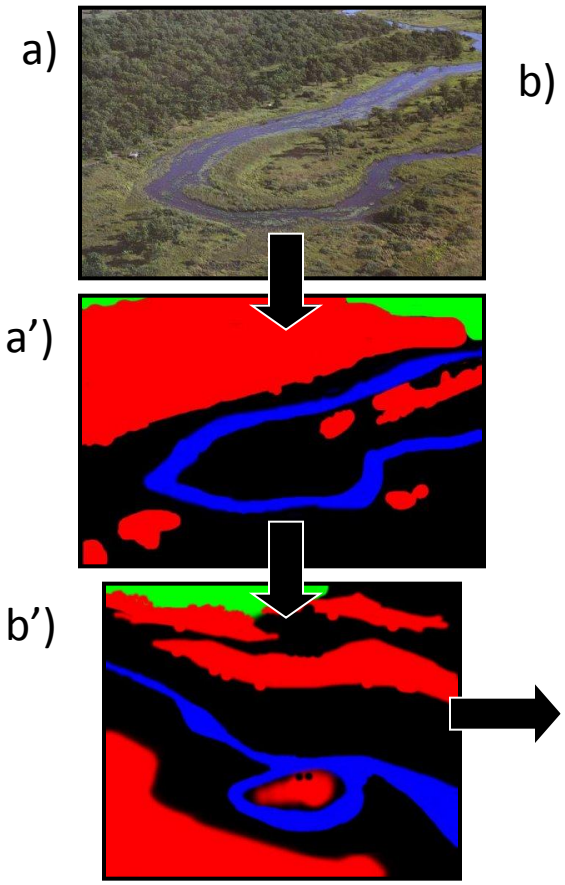
Inspiration: Paint-by-Number

Create detailed paintings by filling in numbered regions corresponding to paint colors.



Inspiration: Texture-by-Number

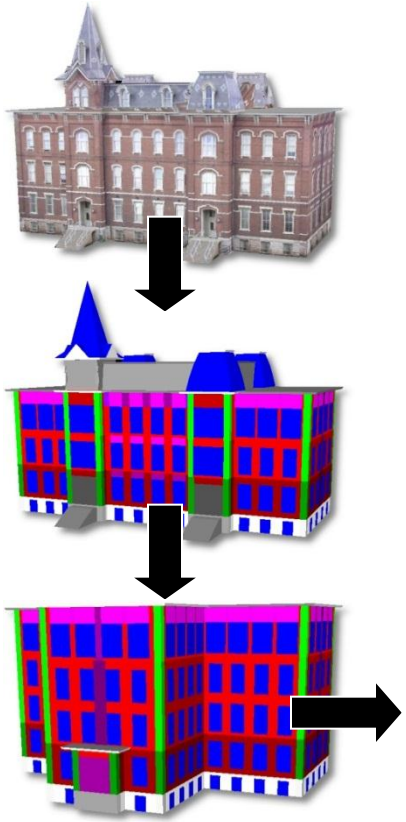
Create realistic images using a source photograph and color coding, along with a target color coding.



novel landscape [Hertzmann01]

This work: Build-by-Number

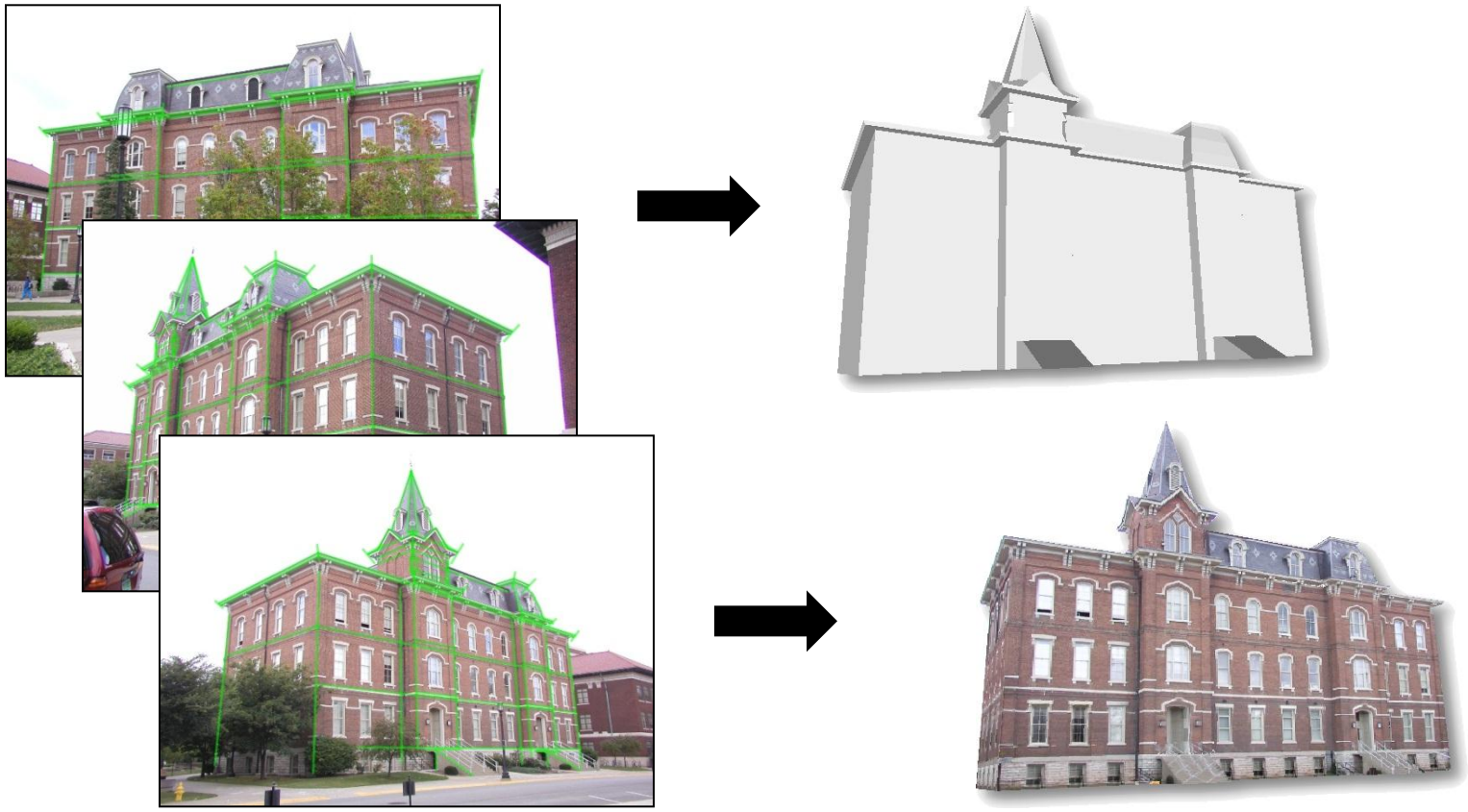
Create realistic architecture using a source image set and model, along with a target model.



novel building

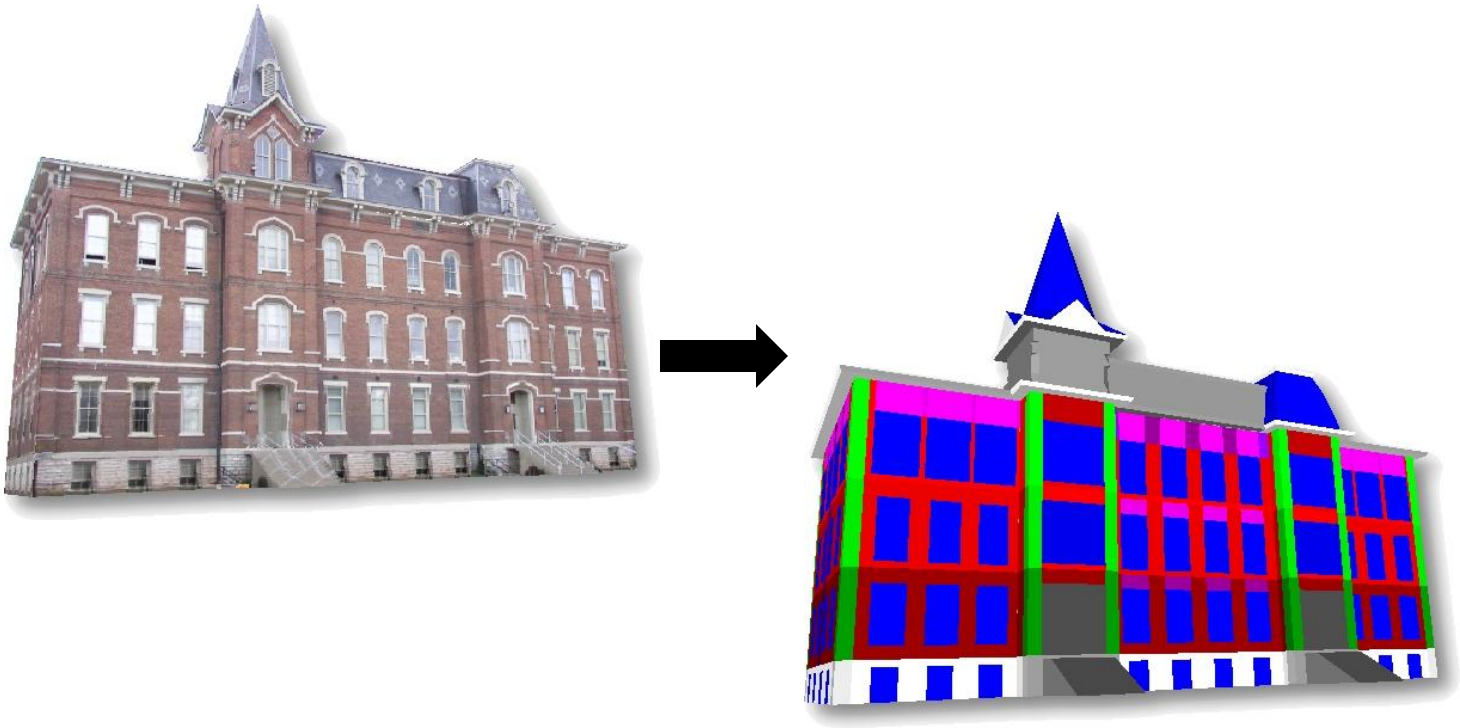
System Overview

1. A geometric model is recovered from a sparse image set.



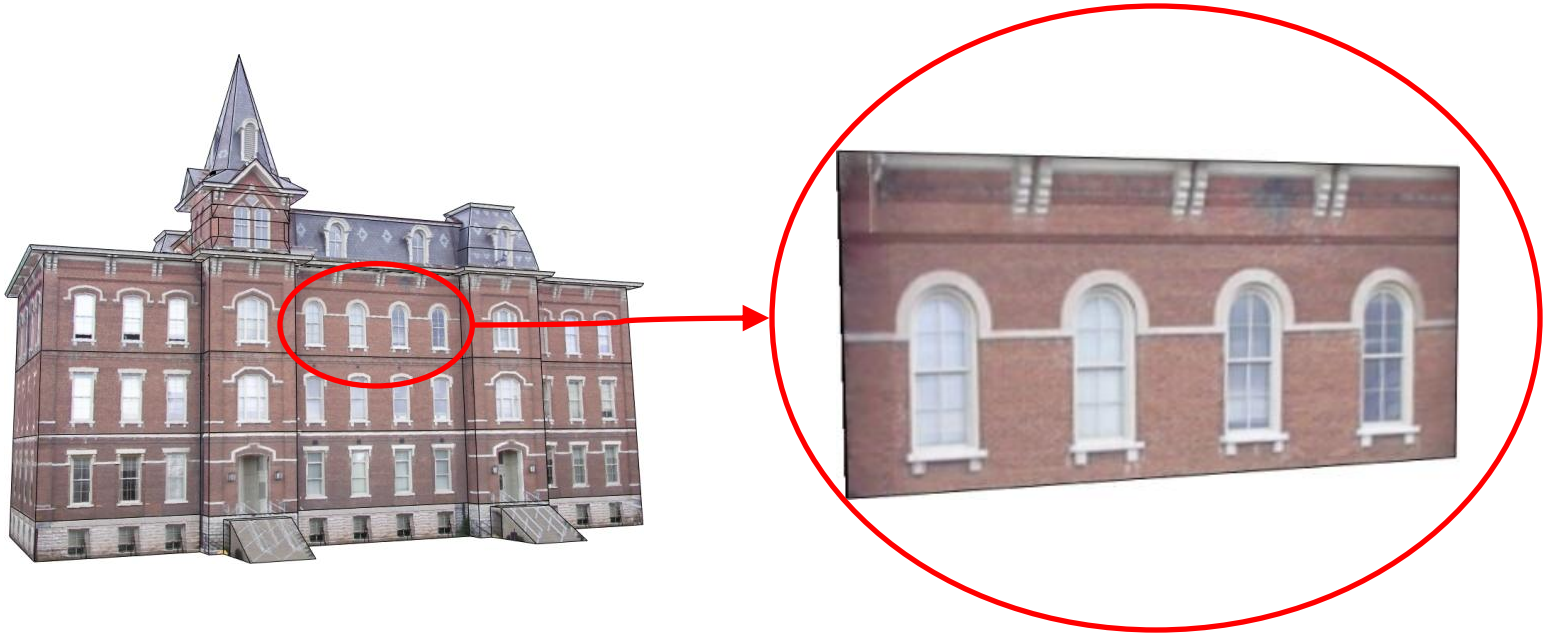
System Overview

2. The model is subdivided into feature regions such as brick, windows, and doors. Identical or similar features are grouped together.



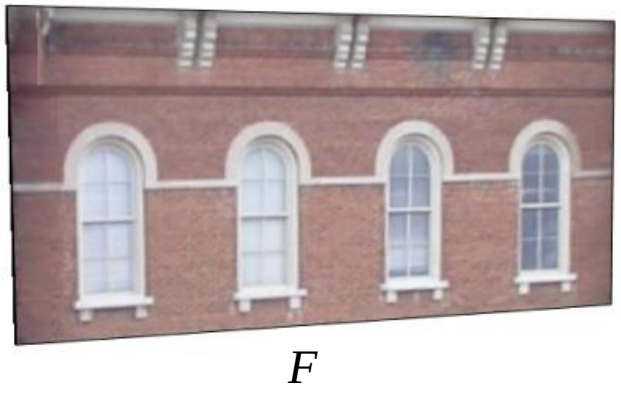
System Overview

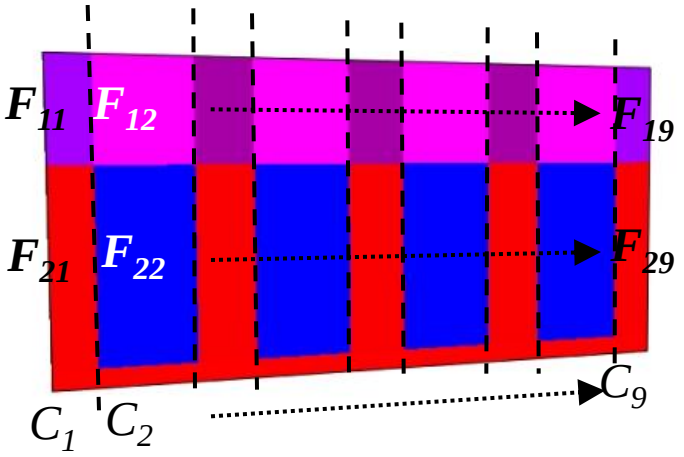
3. A **face schema,** comprising a symbolic growth rule and geometric properties of a single face, is derived and used



System Overview

3. A **face schema,** comprising a symbolic growth rule and geometric properties of a single face, is derived and used

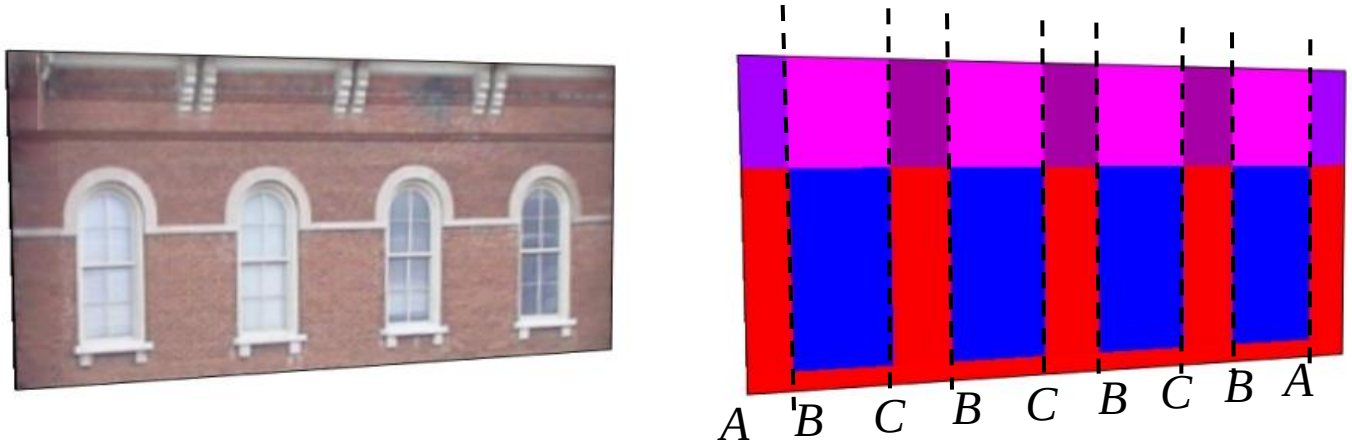




F can be represented symbolically as $F = C_1C_2...C_n$, where $C_j = (F_{1j}F_{2j}...F_{mj})$ is a column of subfaces.

System Overview

3. A **face schema,** comprising a symbolic growth rule and geometric properties of a single face, is derived and used

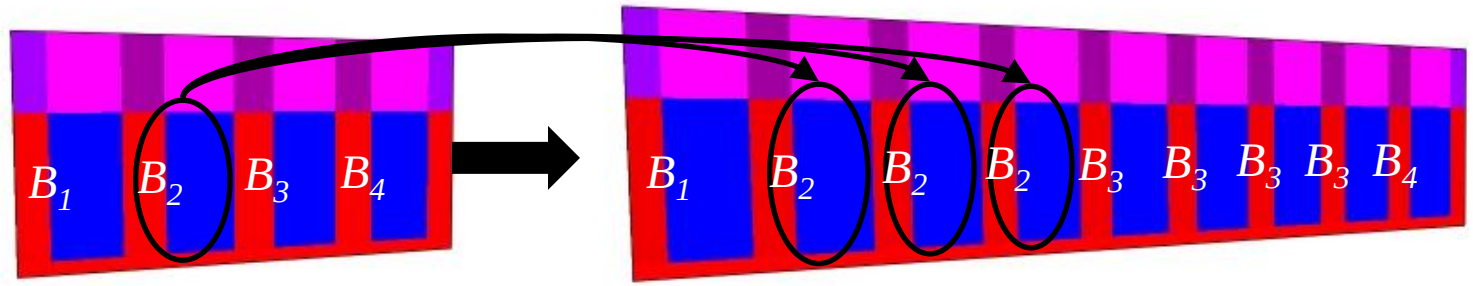


This face is described by the string $F = ABCBCBCBA$, with possible growth rule

$$F \rightarrow A(BC)^*BA$$

System Overview

3. A **face schema,** comprising a symbolic growth rule and geometric properties of a single face, is derived and used



$F = ABCBCBCBA$

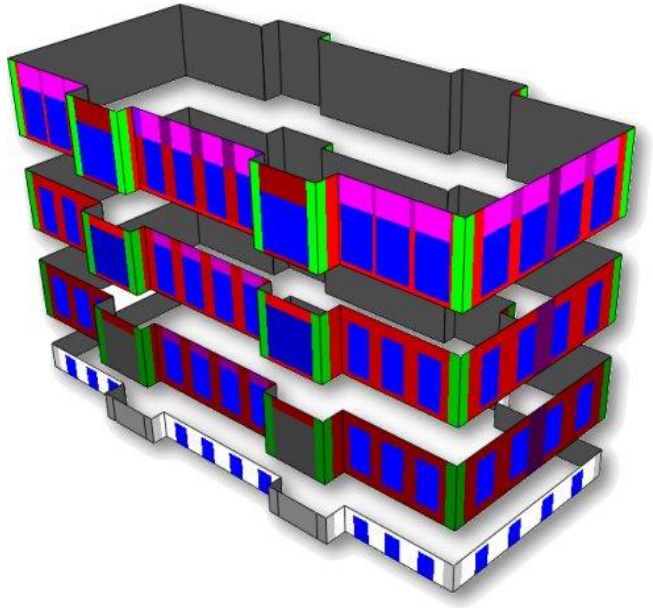
$F \rightarrow A(BC)^*BA$

$F' = ABCBCBCBCBCBCBCBA$

System Overview

4. Similarly, a **floor schema** is derived and used as well

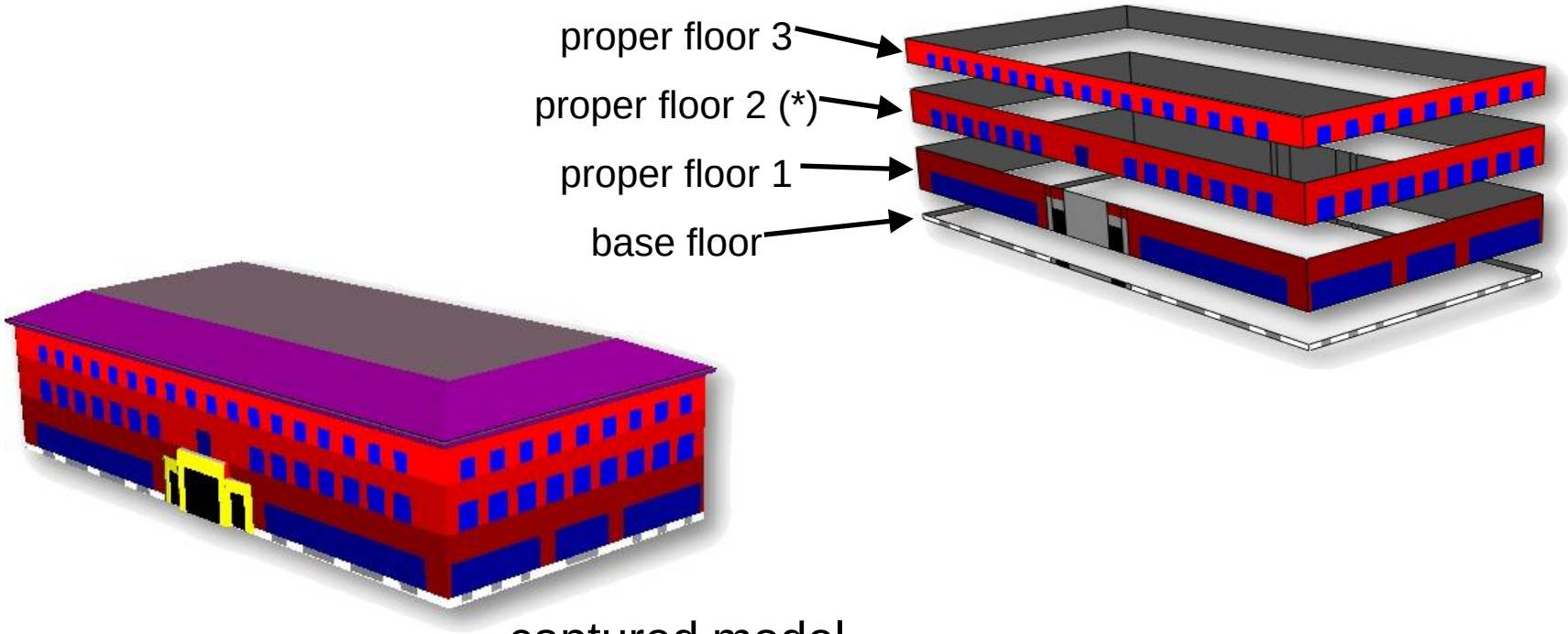
captured model



model floors

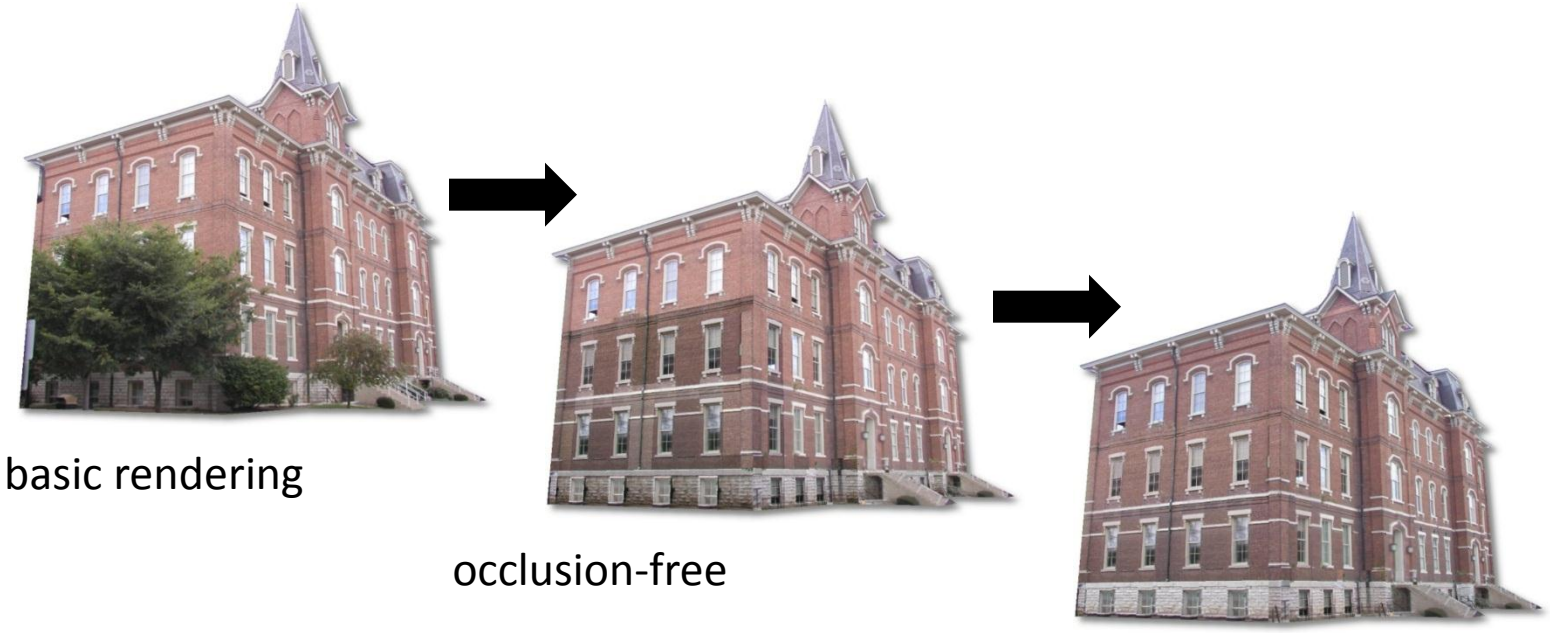
System Overview

5. Similarly, a **model schema** is derived and used as well



System Overview

6. Features from the original image set are textured onto the new model with occlusions removed and shading equalized.



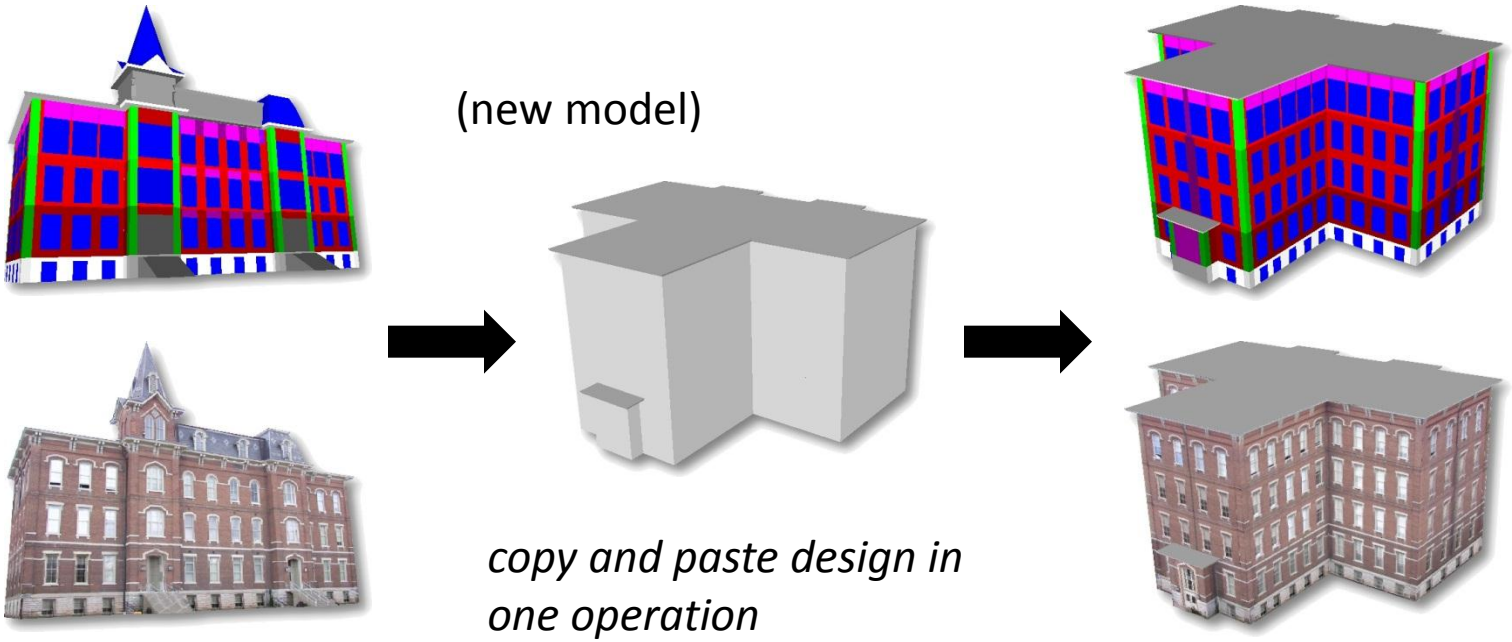
basic rendering

occlusion-free

color equalized

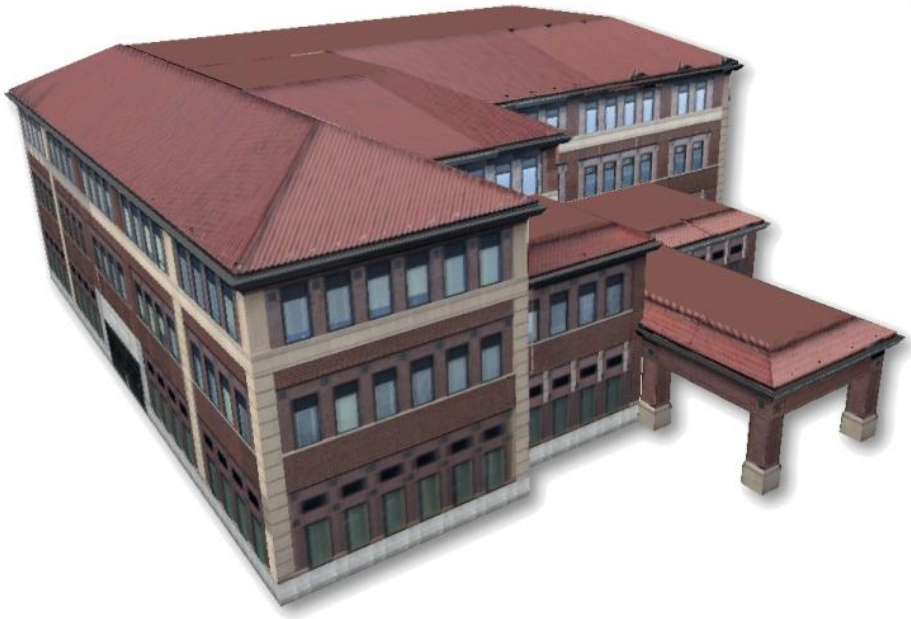
System Overview

7. A new model is created and automatically subdivided and colored using rules derived from the captured model's subdivision scheme.

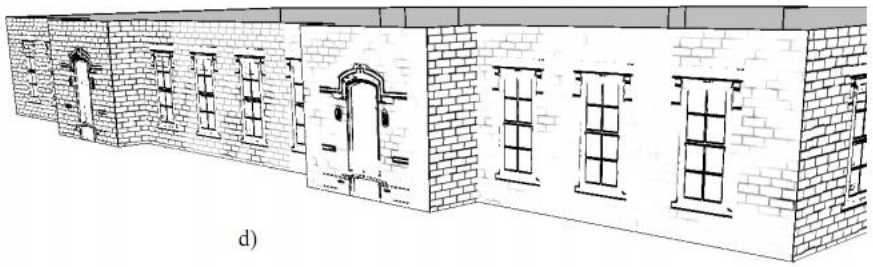


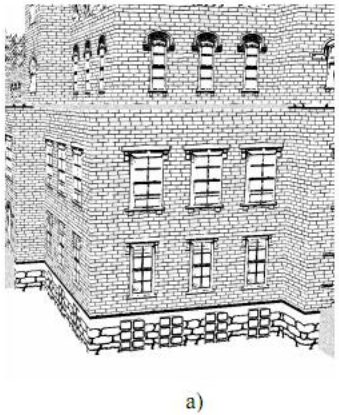
Video

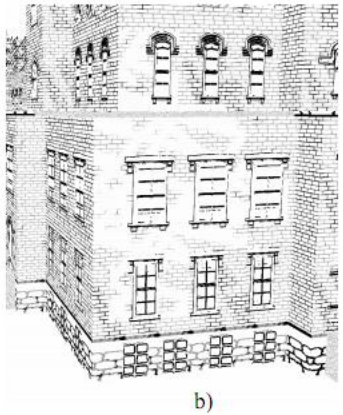
Examples

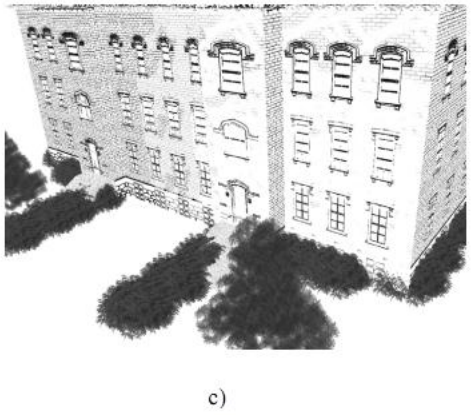


Examples









Examples