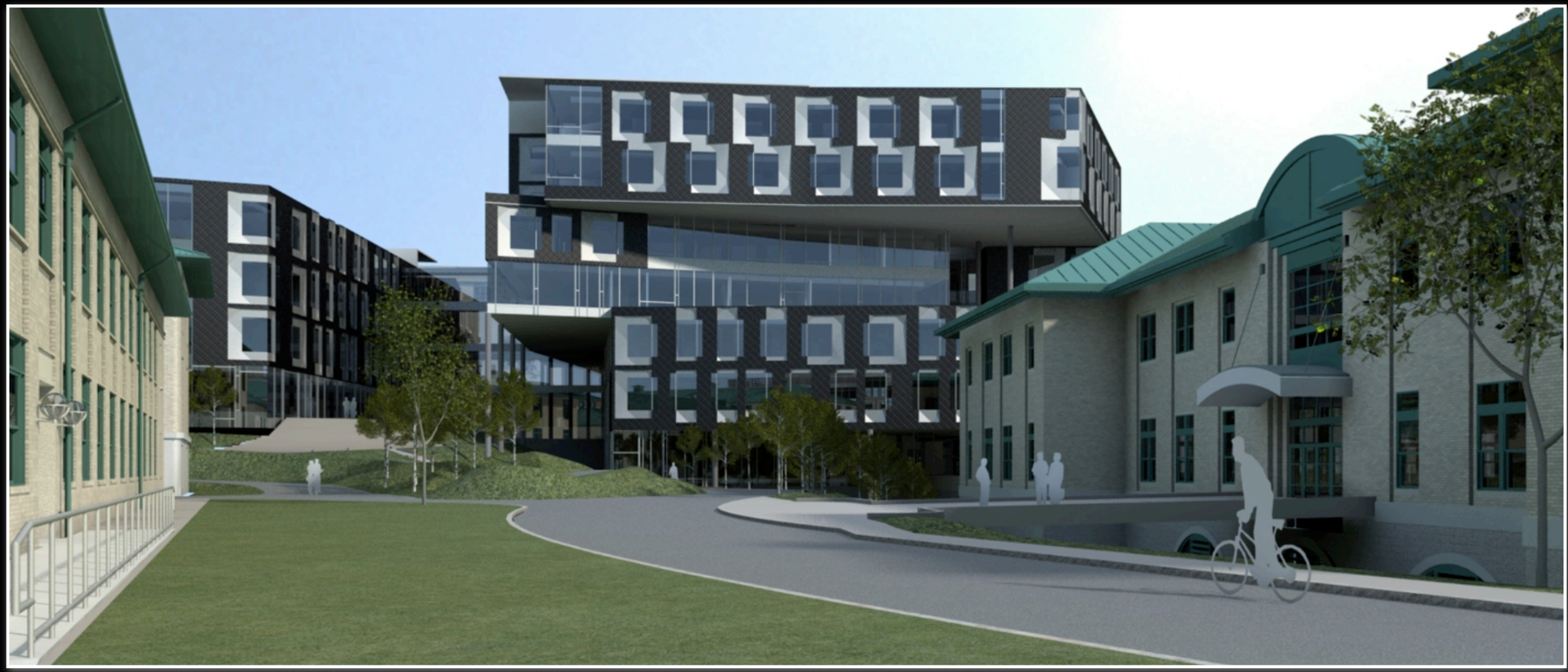
Photo Clip Art

Jean-François Lalonde, Derek Hoiem, Alexei A. Efros
Carnegie Mellon University

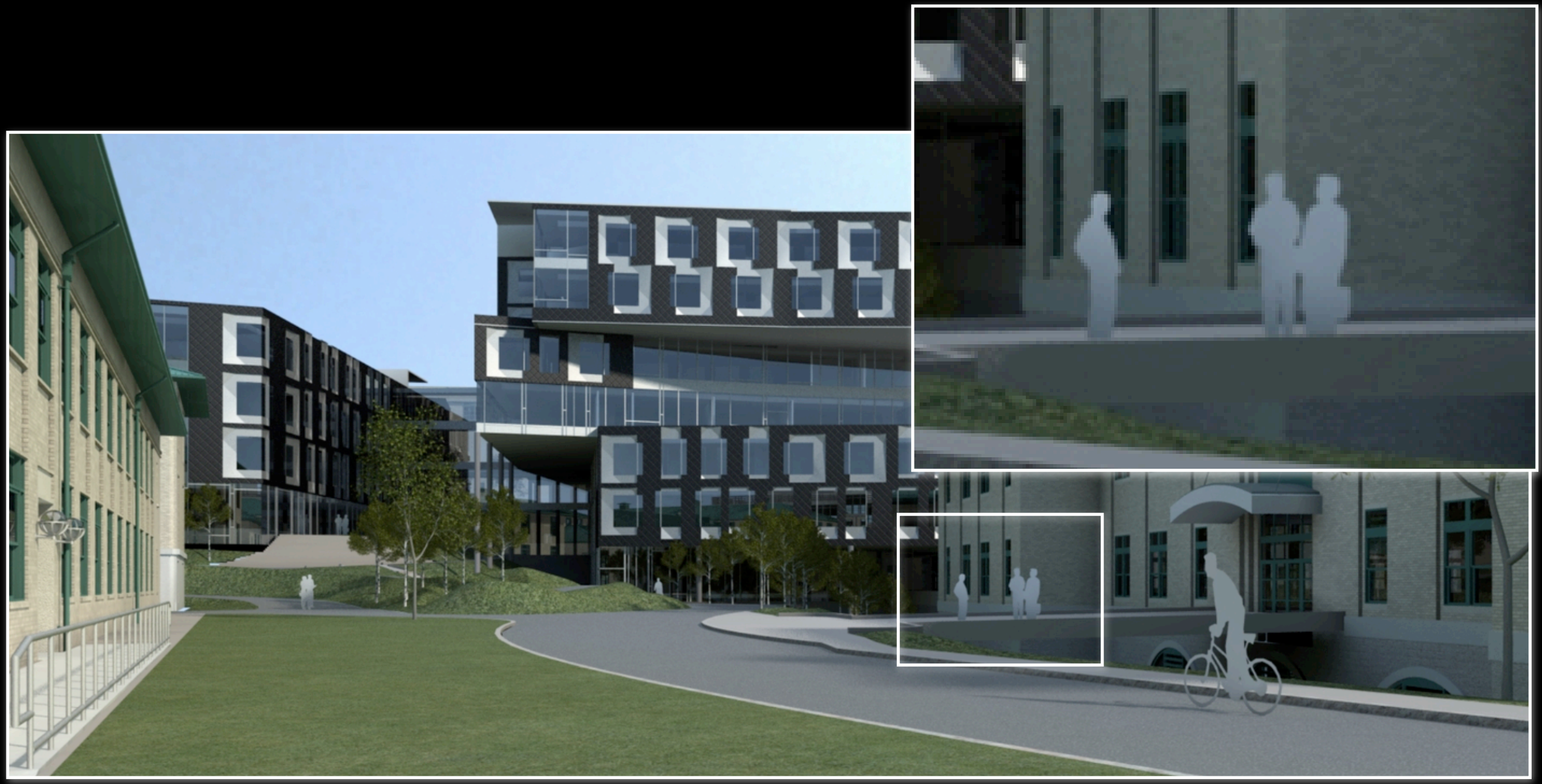
Carsten Rother, John Winn and Antonio Criminisi
Microsoft Research Cambridge

Gates center at CMU, circa 2009



Lalonde et al., SIGGRAPH 2007

Gates center at CMU, circa 2009



Lalonde et al., SIGGRAPH 2007

Gates center at CMU, circa 2009

Lalonde et al., SIGGRAPH 2007

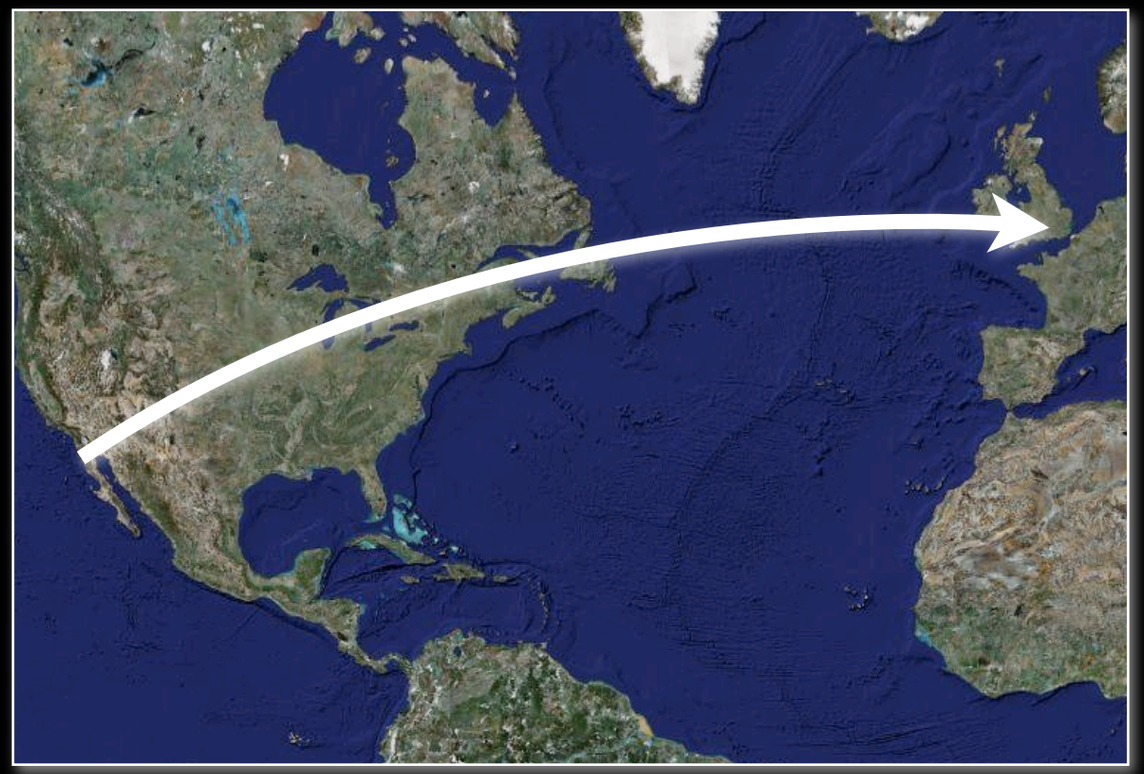
Inserting objects into images

Lalonde et al., SIGGRAPH 2007

Inserting objects into images - the traditional approach [Debevec ‘98]

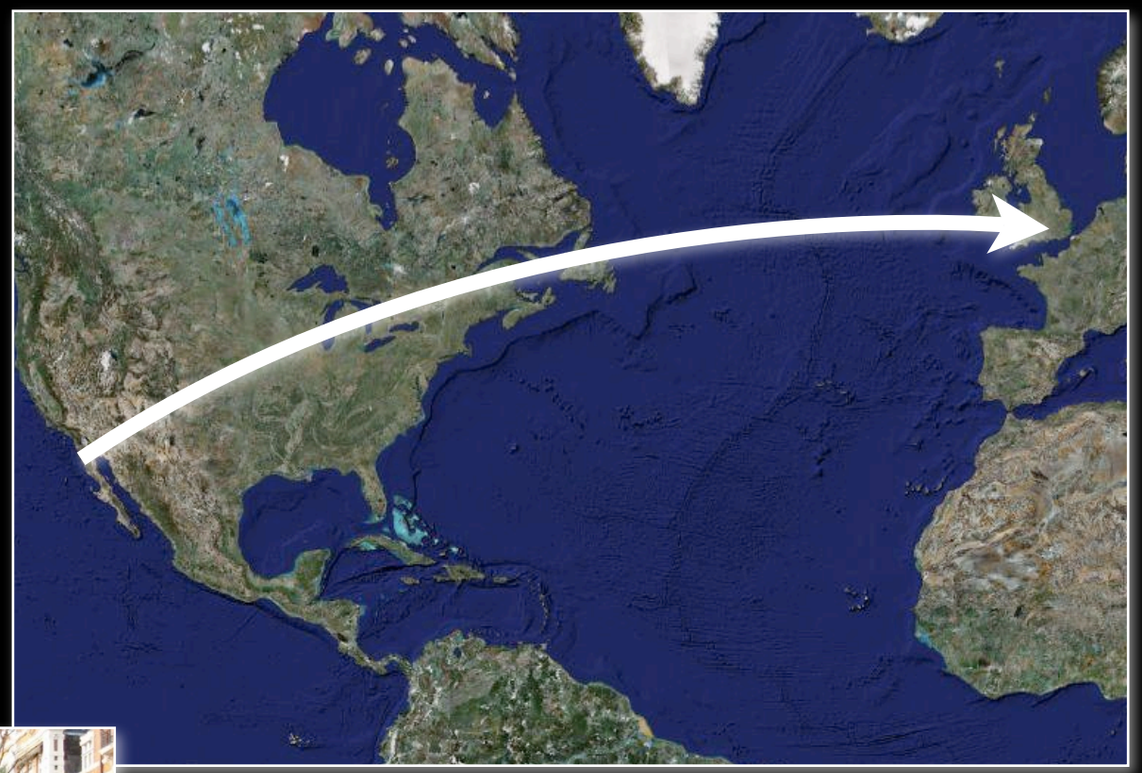
Lalonde et al., SIGGRAPH 2007

Inserting objects into images - the traditional approach [Debevec ‘98]



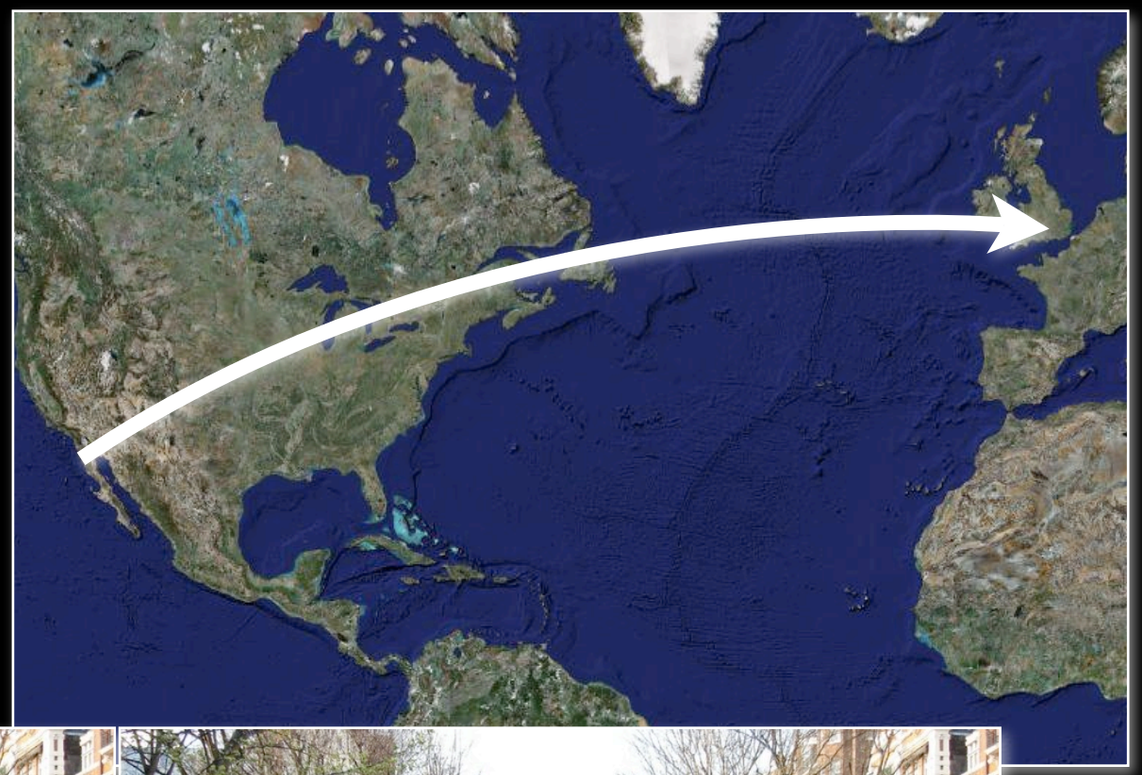
Lalonde et al., SIGGRAPH 2007

Inserting objects into images - the traditional approach [Debevec ‘98]



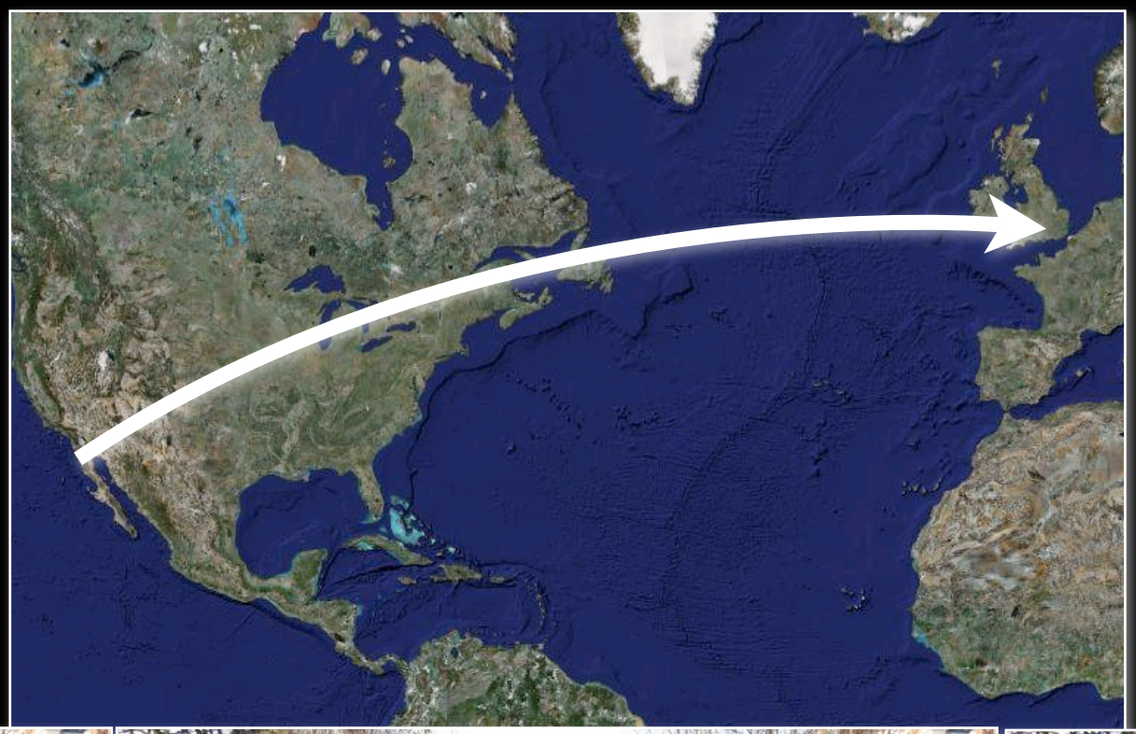
Lalonde et al., SIGGRAPH 2007

Inserting objects into images - the traditional approach [Debevec ‘98]



Lalonde et al., SIGGRAPH 2007

Inserting objects into images - the traditional approach [Debevec ‘98]



Lalonde et al., SIGGRAPH 2007

Inserting objects in images

[Debevec ‘98]

Lalonde et al., SIGGRAPH 2007

Inserting objects in images

[Debevec ‘98]

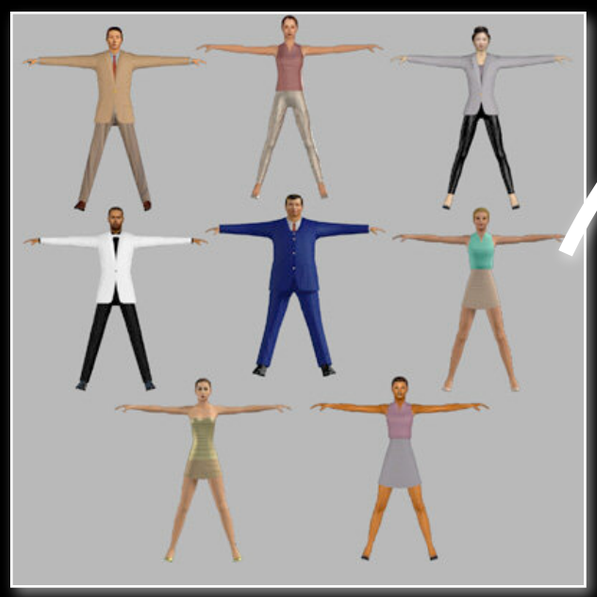
Highly detailed geometry
Highly detailed materials
Very expensive

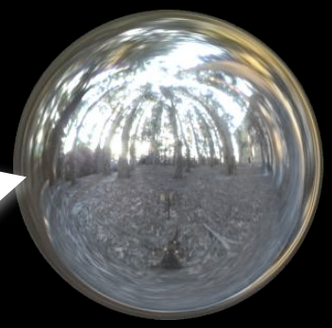
Lalonde et al., SIGGRAPH 2007

Inserting objects in images

[Debevec ‘98]

Highly detailed geometry
Highly detailed materials
Very expensive



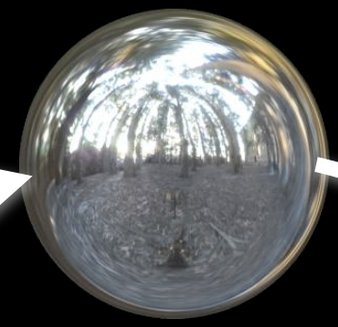


Lalonde et al., SIGGRAPH 2007

Inserting objects in images

[Debevec ‘98]

Highly detailed geometry
Highly detailed materials
Very expensive



- Realistic renderings
- Expensive and impractical

Lalonde et al., SIGGRAPH 2007

Alternative: Clip art

Lalonde et al., SIGGRAPH 2007

Alternative: Clip art

- Easy, intuitive, cheap
- Not realistic

Lalonde et al., SIGGRAPH 2007

Goal

Photo-realistic ⟷ **Cartoon**

Image-based rendering

Clip Art

Expensive and impractical ⟷ **Cheap and intuitive**

Lalonde et al., SIGGRAPH 2007

Goal

Photo-realistic **Cartoon**

Image-based rendering

Photo Clip Art

?

Clip Art

Expensive and impractical **Cheap and intuitive**

Lalonde et al., SIGGRAPH 2007

Challenges

- Insert **THIS** object: impossible!

Lalonde et al., SIGGRAPH 2007

Challenges

- Insert **THIS** object: impossible!

Lalonde et al., SIGGRAPH 2007

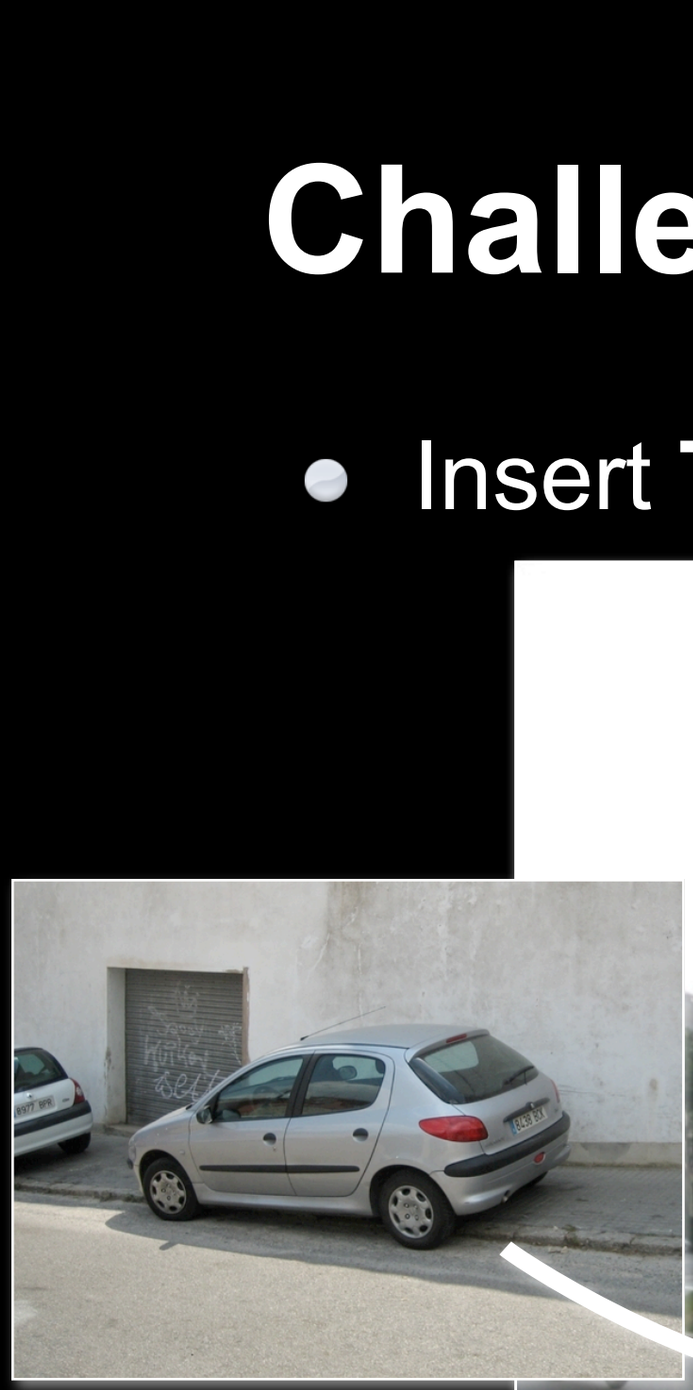
Challenges

- Insert **THIS** object: impossible!

Lalonde et al., SIGGRAPH 2007

Challenges

- Insert **THIS** object: impossible!



- object orientation
- scene illumination

Lalonde et al., SIGGRAPH 2007

The use of data

- Insert **SOME** object: much easier!

Lalonde et al., SIGGRAPH 2007

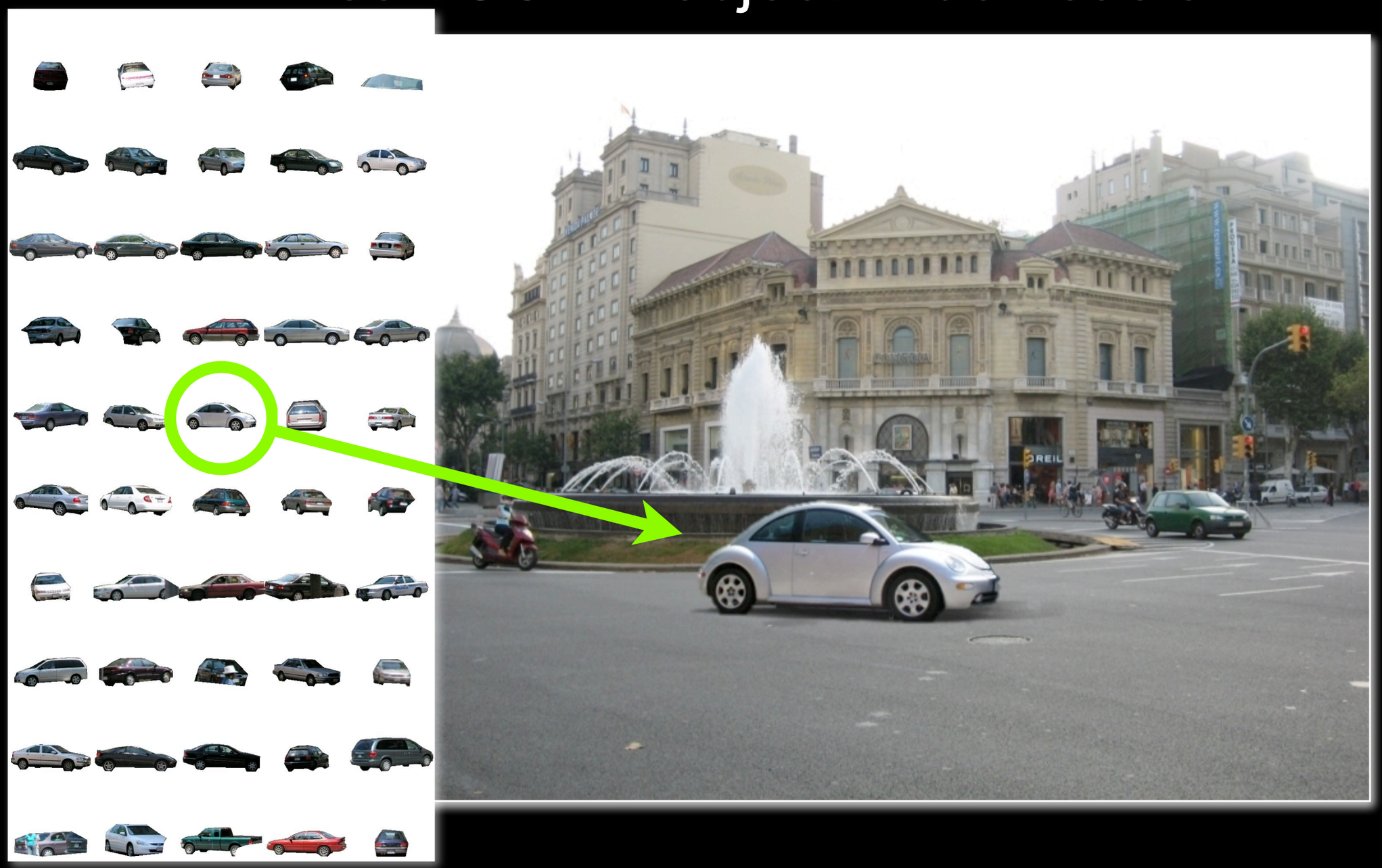
The use of data

- Insert **SOME** object: much easier!

Lalonde et al., SIGGRAPH 2007

The use of data

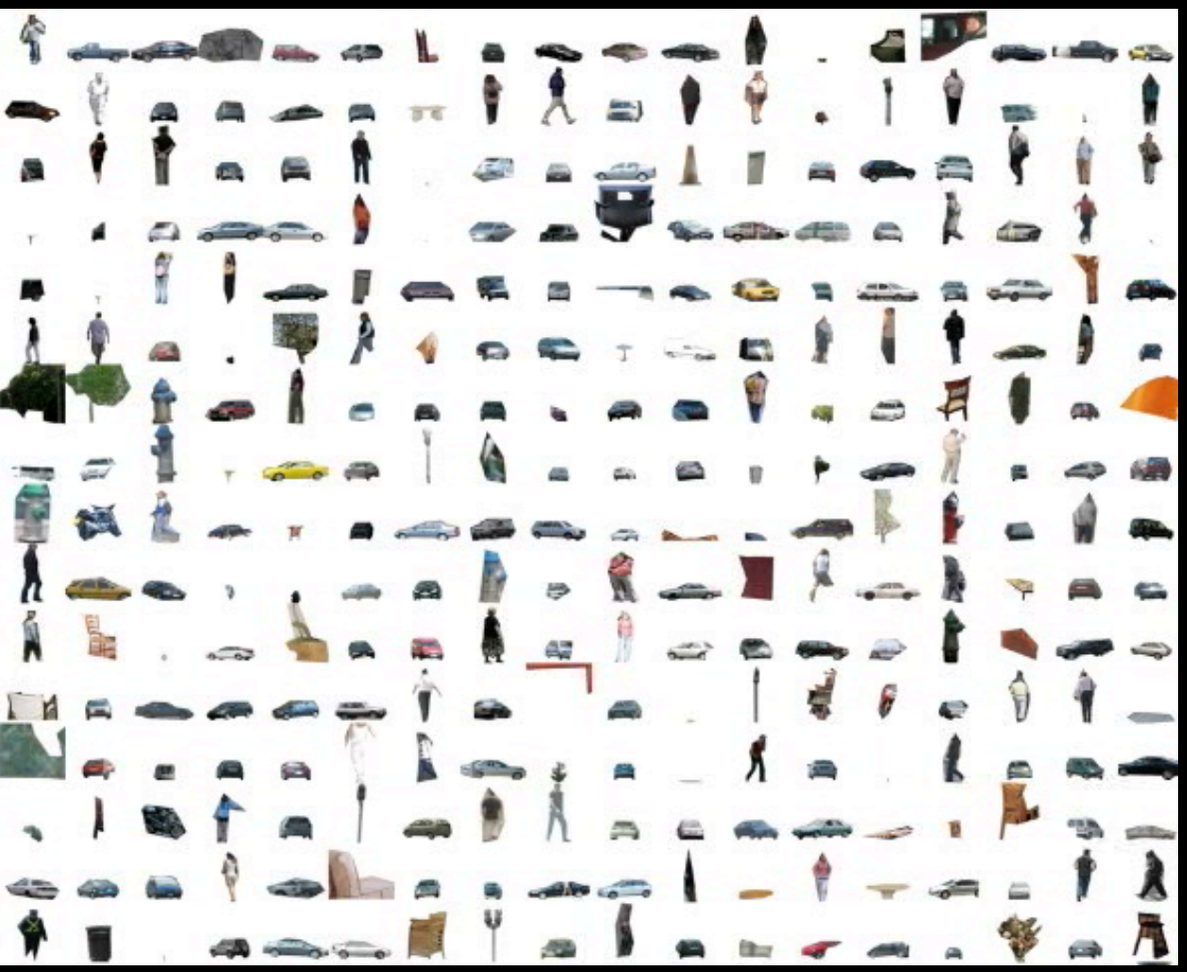
- Insert **SOME** object: much easier!



Lalonde et al., SIGGRAPH 2007

The Google model

Database



Lalonde et al., SIGGRAPH 2007

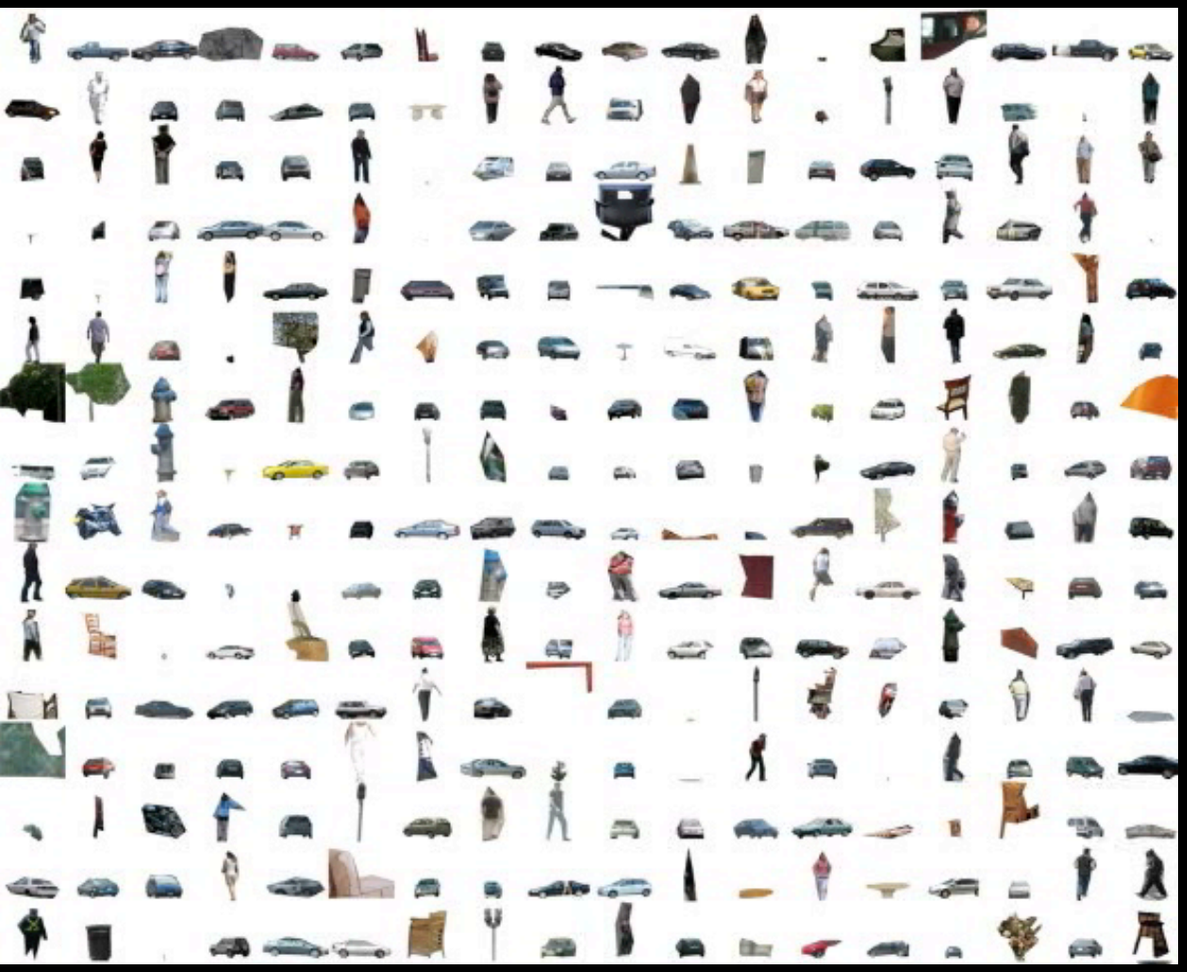
The Google model

Database **Query**

Lalonde et al., SIGGRAPH 2007

The Google model

Database



Query

Sort the objects

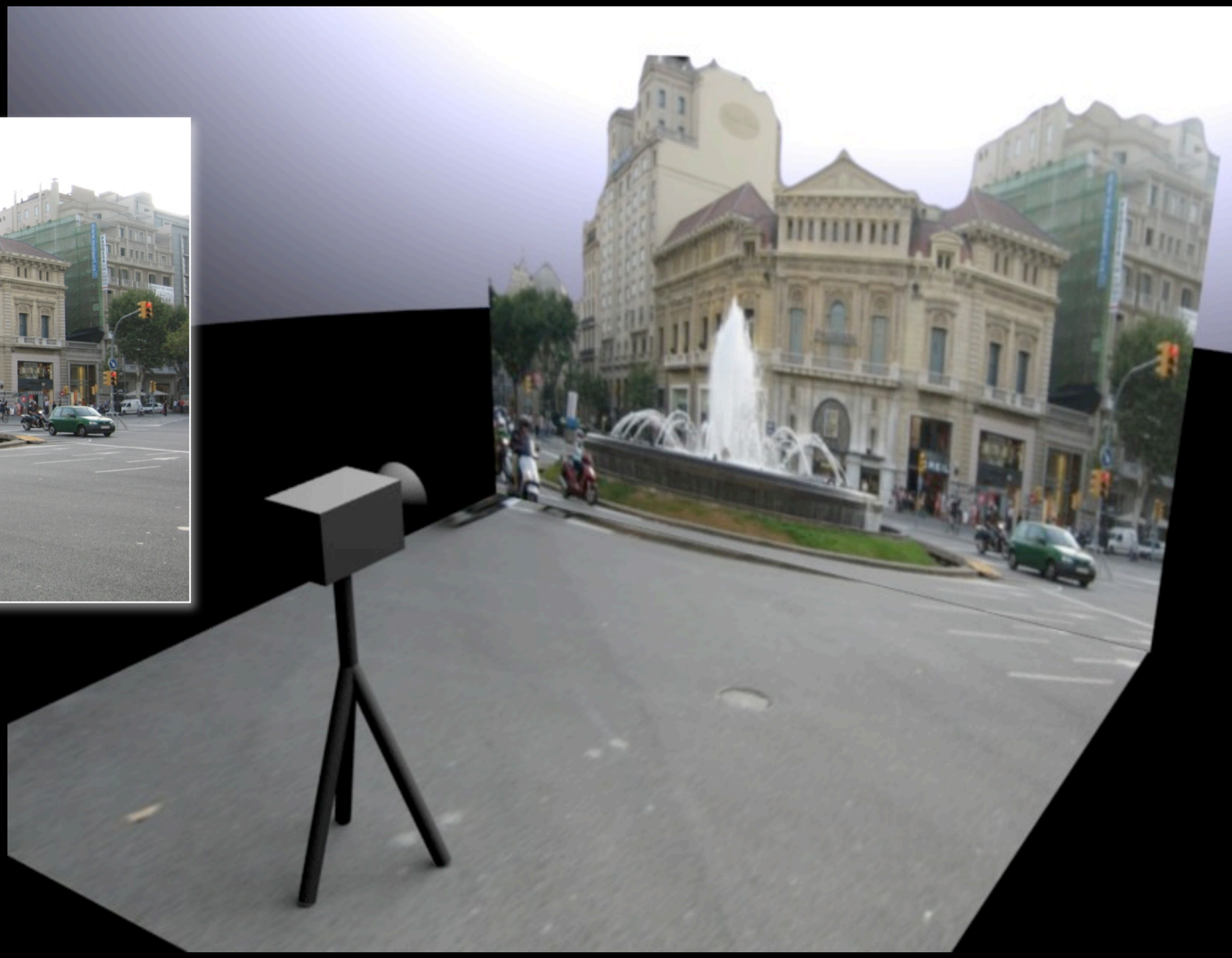
Results

Lalonde et al., SIGGRAPH 2007

2-D image vs 3-D scene

Lalonde et al., SIGGRAPH 2007

2-D image vs 3-D scene



Lalonde et al., SIGGRAPH 2007

Prior work

3-D from single image

Tour into the picture
[Horry *et al.*, ’97, Criminisi *et al.*, ‘00]

Image-based Modeling
and Photo Editing
[Oh *et al.*, ‘01]

Automatic Photo Pop-up
[Hoiem *et al.*, ‘05]

Lalonde et al., SIGGRAPH 2007

Prior work

3-D from single image

Tour into the picture
[Horry *et al.*, ’97, Criminisi *et al.*, ‘00]

Image-based Modeling
and Photo Editing
[Oh *et al.*, ‘01]

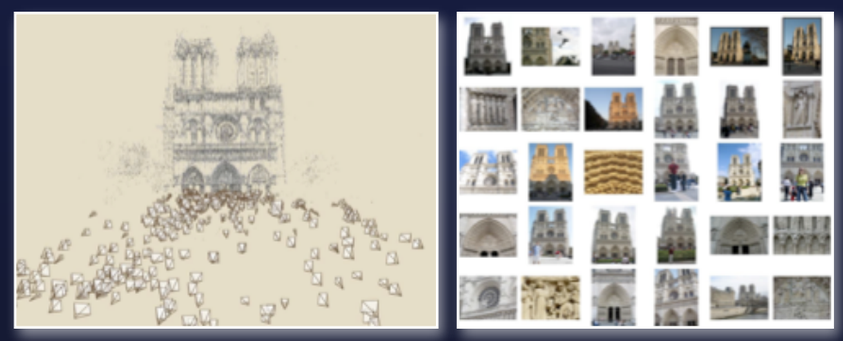
Automatic Photo Pop-up
[Hoiem *et al.*, ‘05]

Data-driven image synthesis

Content based Image Synthesis
[Diakopoulos *et al.*, ‘04]

Semantic Photo Synthesis
[Johnson *et al.*, ‘06]

Photo Tourism
[Snavely *et al.*, ’06]



Outline

Phase I: Database annotation

Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Outline

Phase I: Database annotation

Name: person
Subgroup: person, standing
Height: 1.5m
Local context: in shadow
Illumination: sunny, bright day, no cloud
Segmentation quality: excellent, >40 points
Upsampling blur: low

Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Outline

Phase I: Database annotation

Name: person
Subgroup: person, standing
Height: 1.5m
Local context: in shadow
Illumination: sunny, bright day, no cloud
Segmentation quality: excellent, >40 points
Upsampling blur: low

Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Outline

Phase I: Database annotation

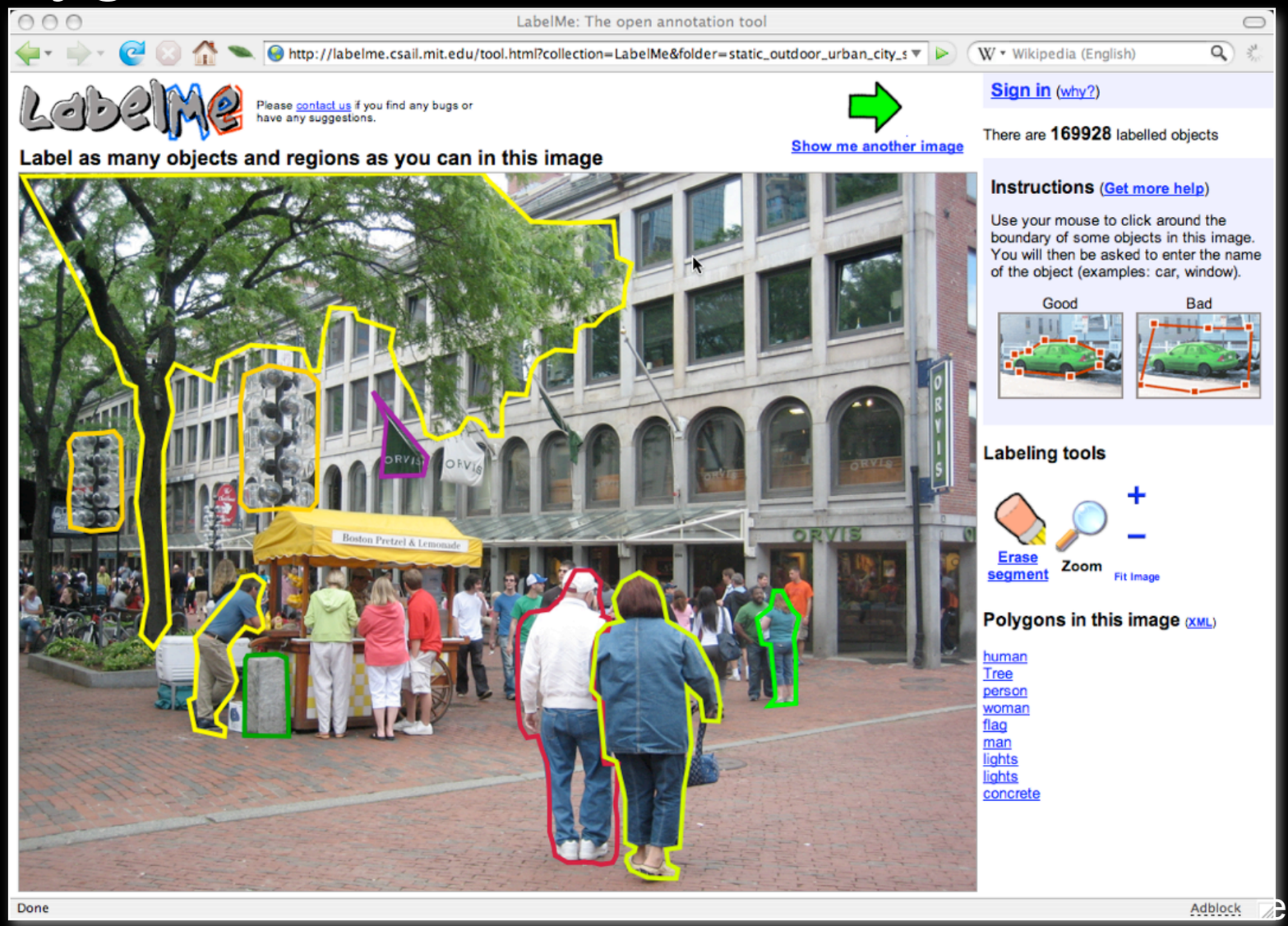
Name: person
Subgroup: person, standing
Height: 1.5m
Local context: in shadow
Illumination: sunny, bright day, no cloud
Segmentation quality: excellent, >40 points
Upsampling blur: low

Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Data source: LabelMe [Russell *et al.*, 2005]

- Online (http://labelme.csail.mit.edu), user-contributed
- 170,000 objects in 40,000 images
- Polygons and names

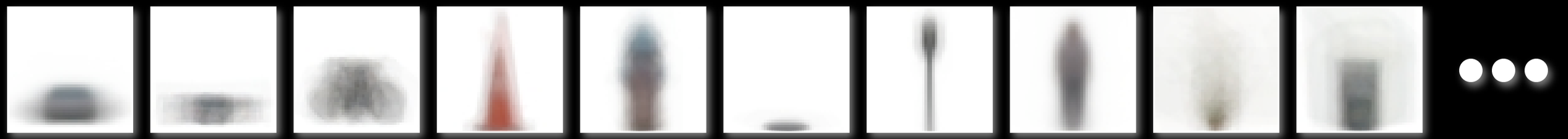


Data organization

Lalonde et al., SIGGRAPH 2007

Data organization

Top-level categories (chosen manually, 16 total)

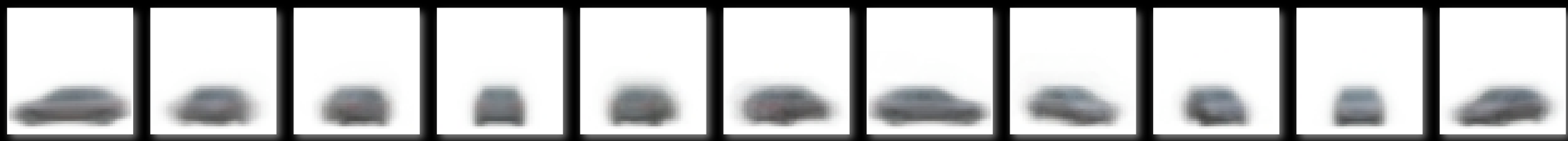


Lalonde et al., SIGGRAPH 2007

Data organization

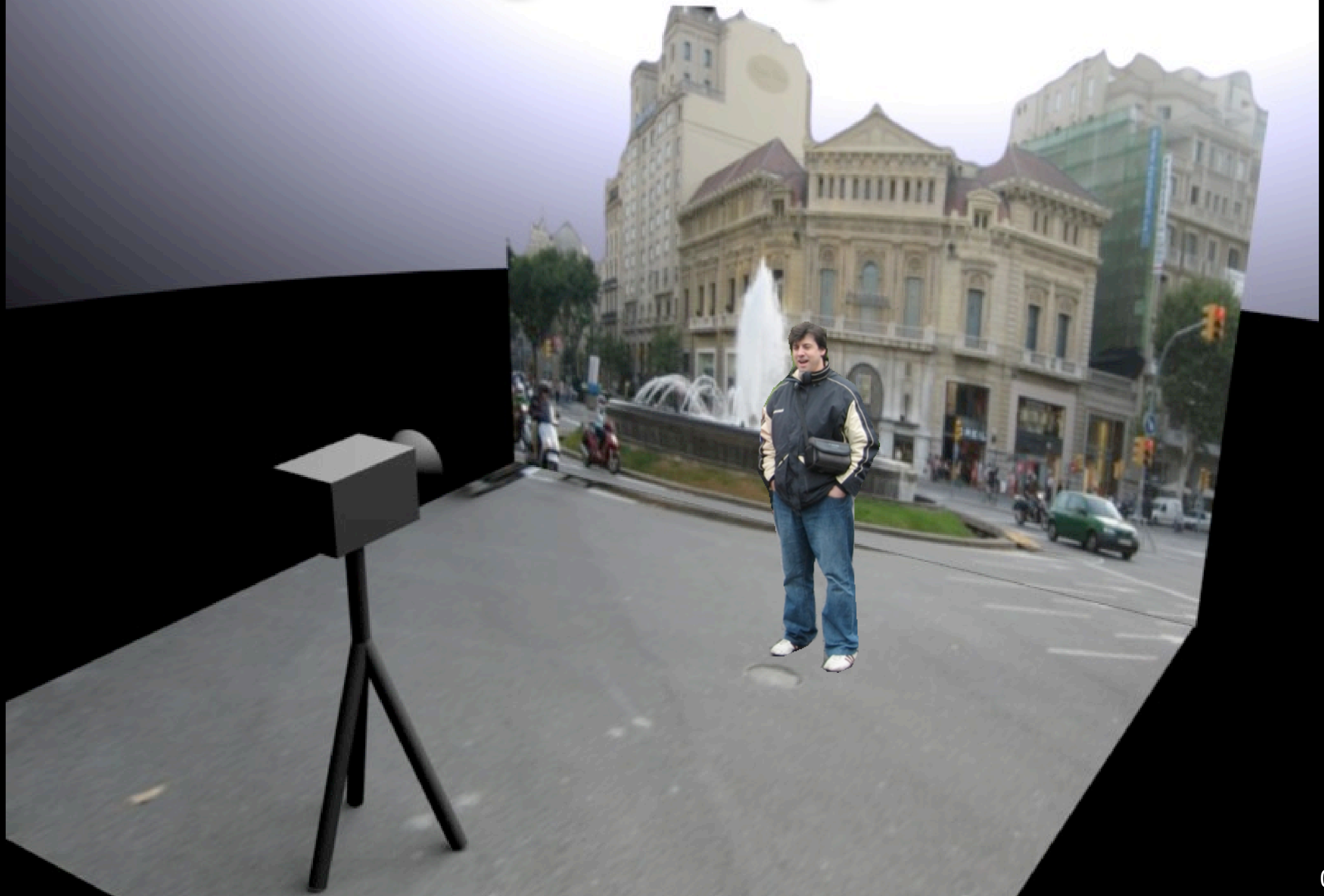
Top-level categories (chosen manually, 16 total)

Second-level categories (from annotations or clustering)

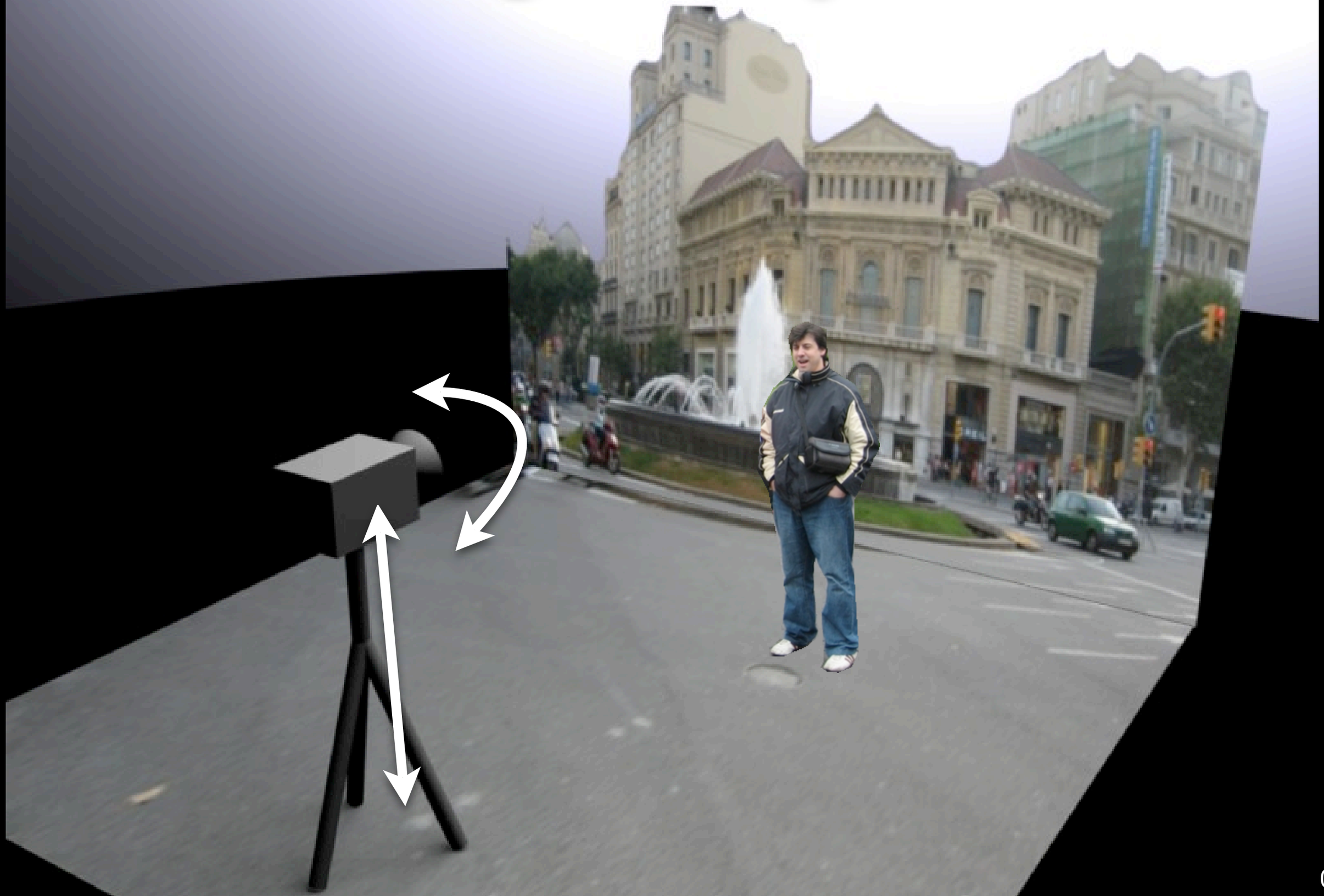


Lalonde et al., SIGGRAPH 2007

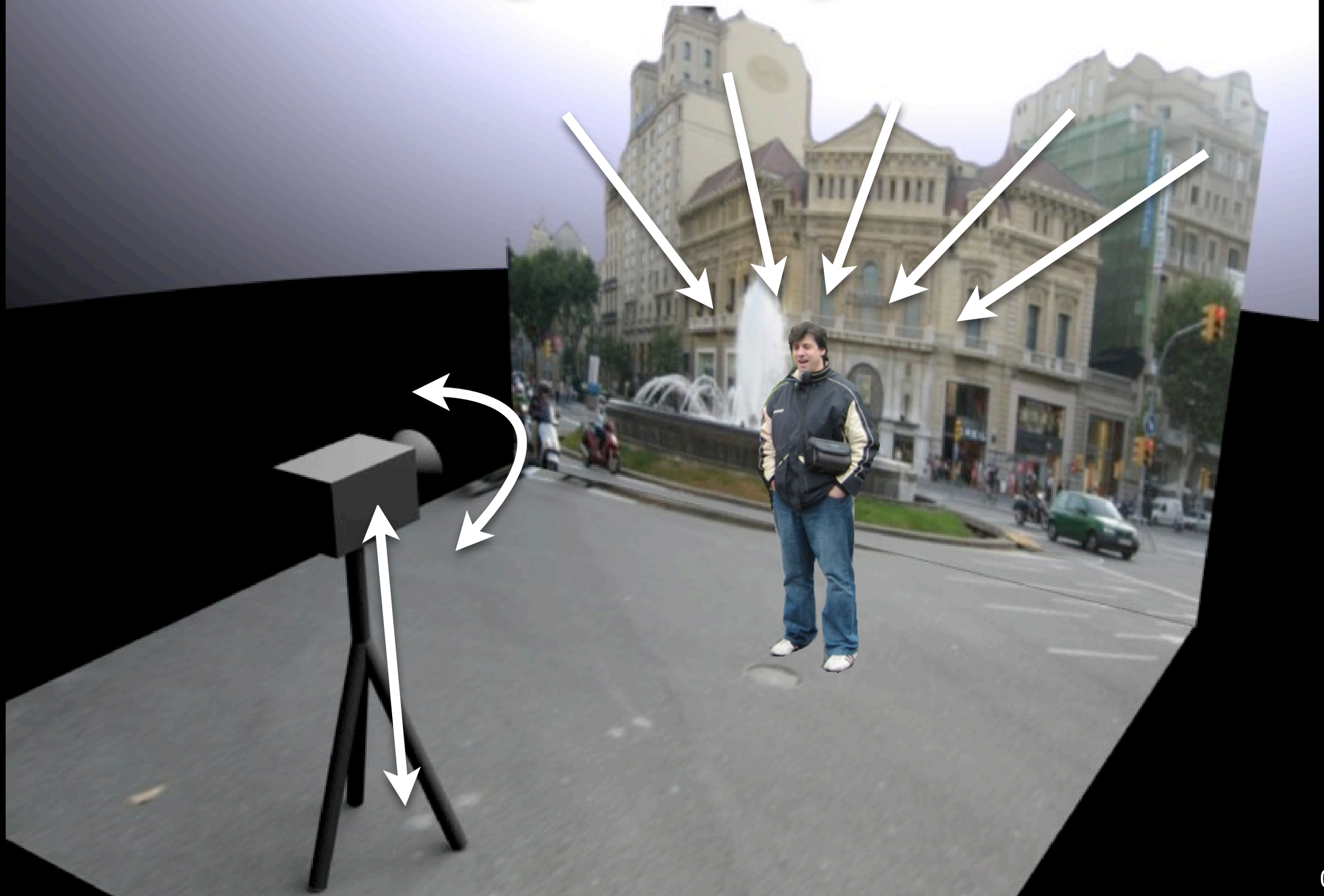
Annotating the objects



Annotating the objects



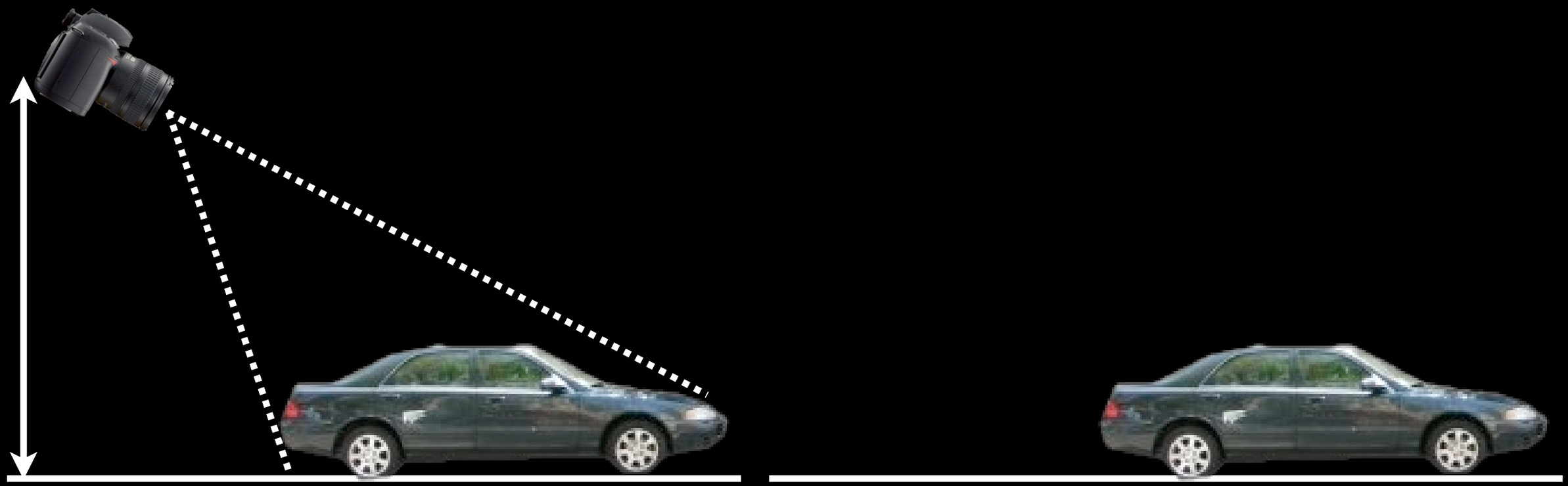
Annotating the objects



Camera parameters

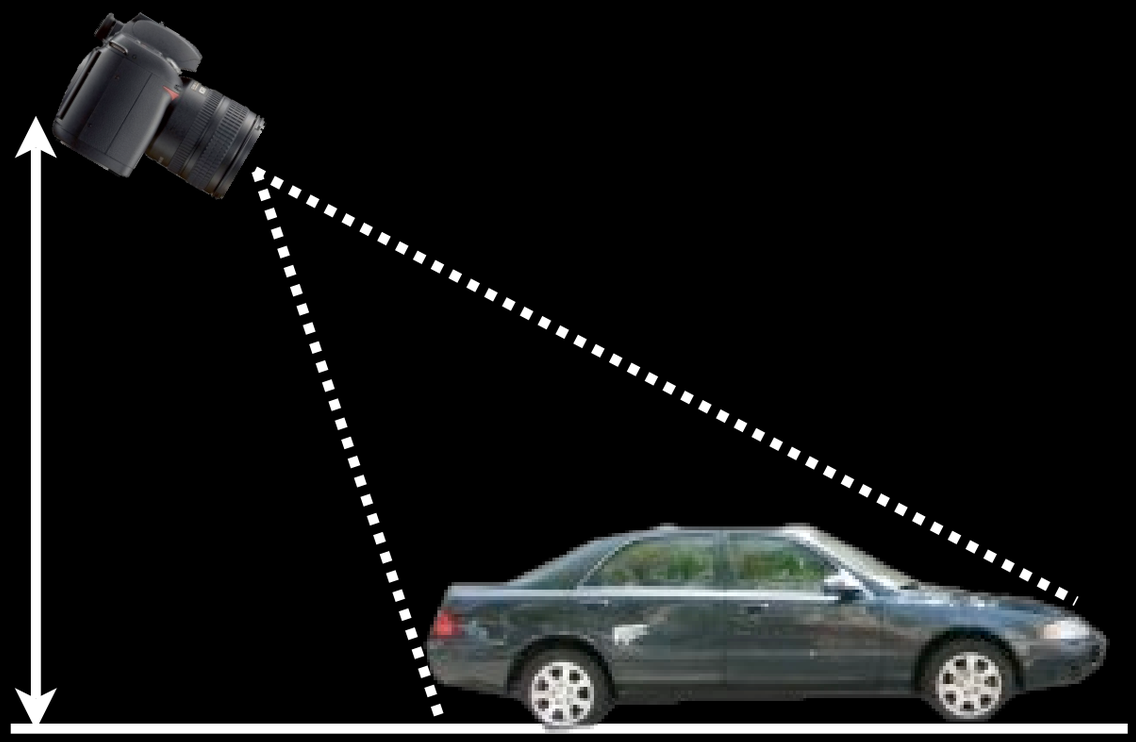
Lalonde et al., SIGGRAPH 2007

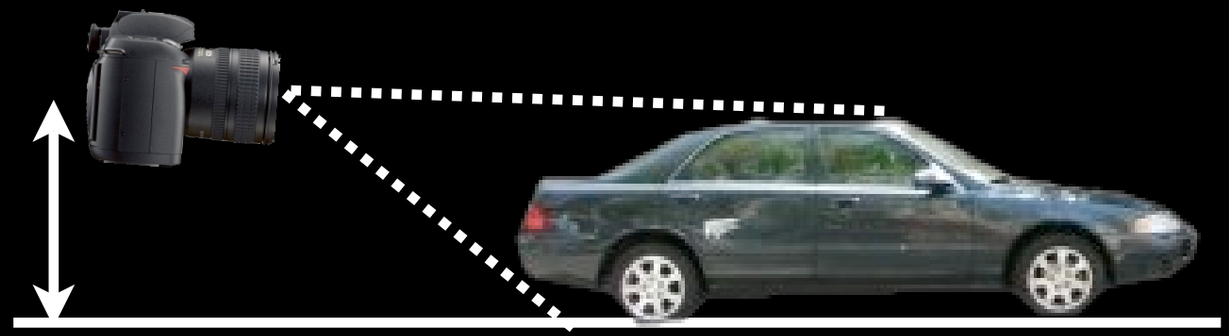
Camera parameters



Lalonde et al., SIGGRAPH 2007

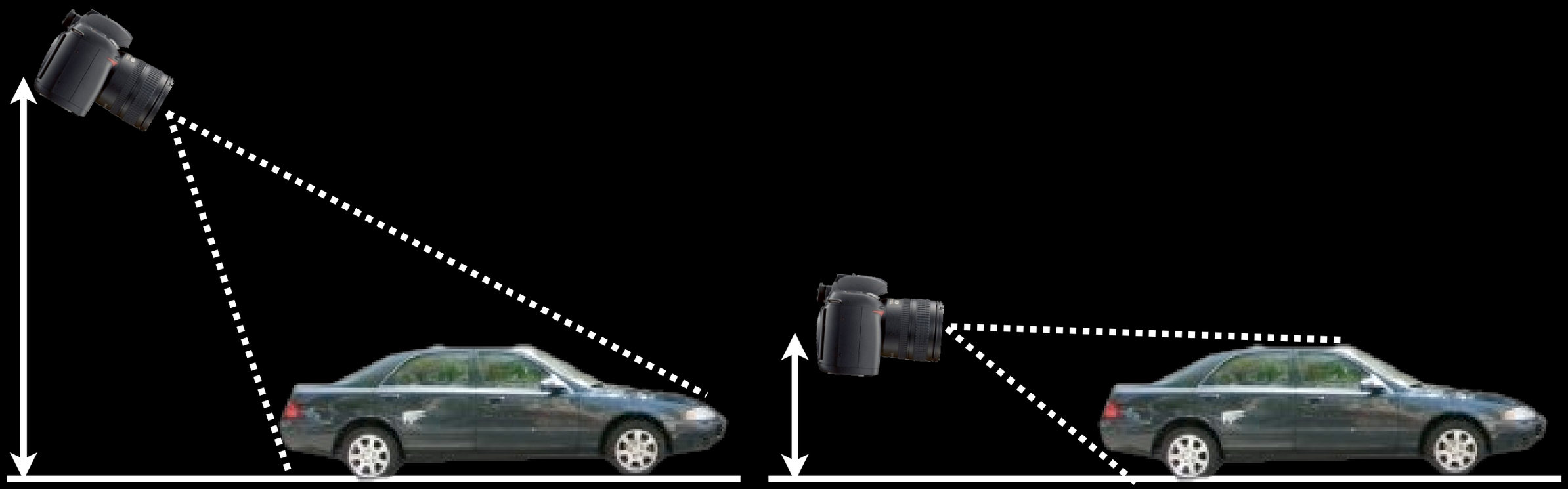
Camera parameters





Lalonde et al., SIGGRAPH 2007

Camera parameters



- Assume
  - flat ground plane
  - all objects on ground
  - camera roll is negligible (consider pitch only)
- Camera parameters: height and orientation

Lalonde et al., SIGGRAPH 2007

Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

Lalonde et al., SIGGRAPH 2007

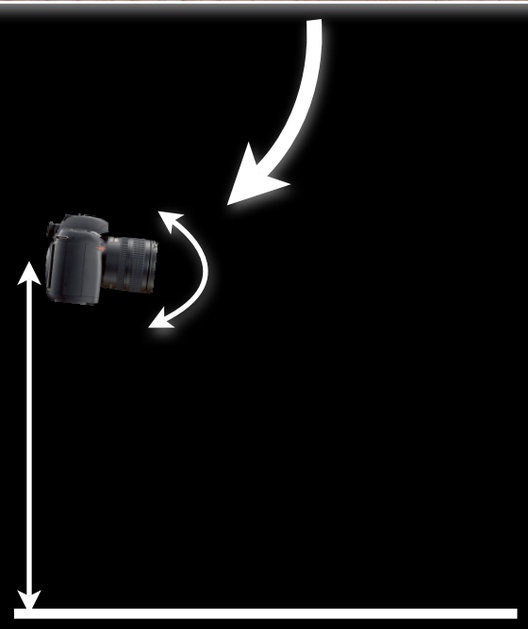
Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

Lalonde et al., SIGGRAPH 2007

Camera parameters

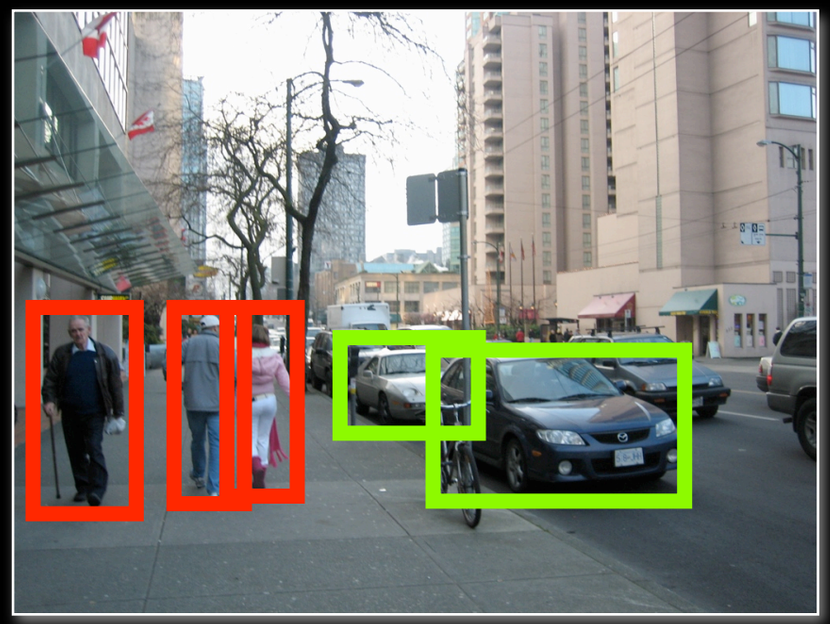
Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

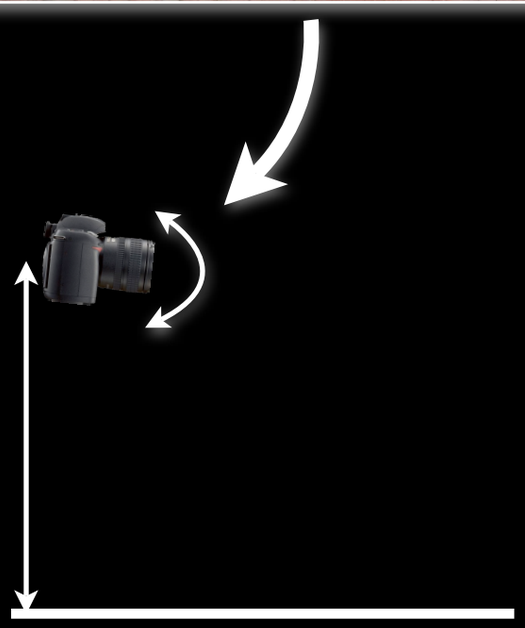


Lalonde et al., SIGGRAPH 2007

Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

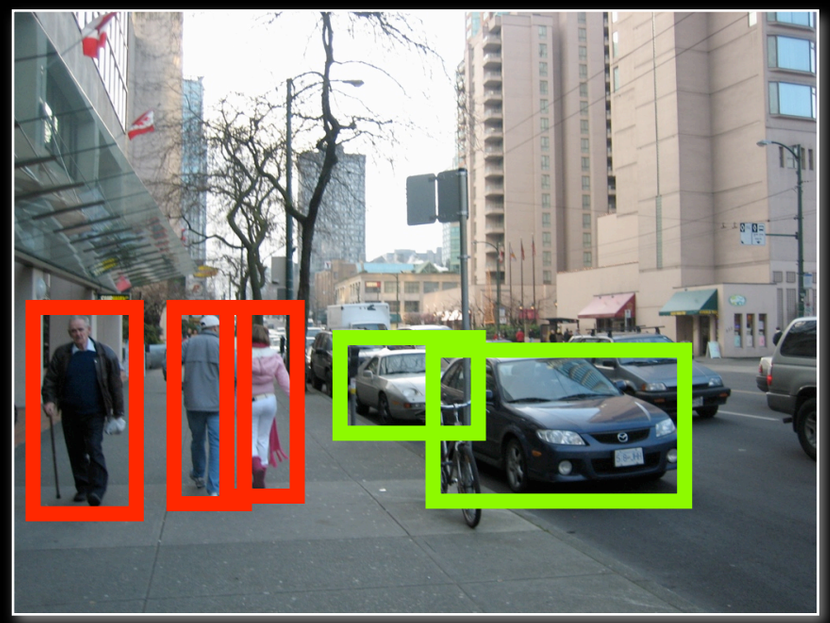


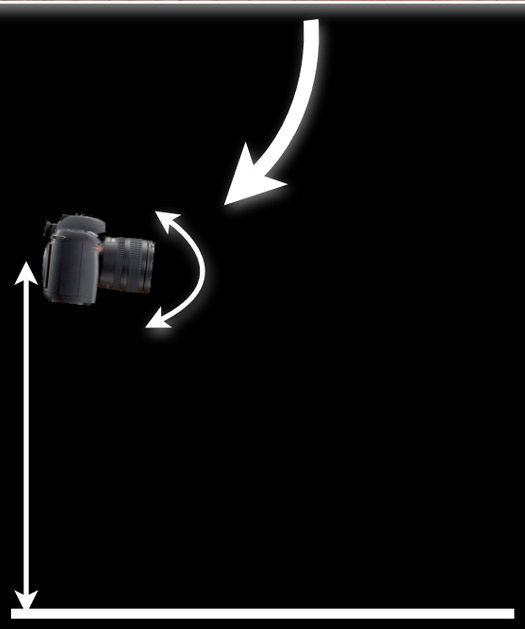


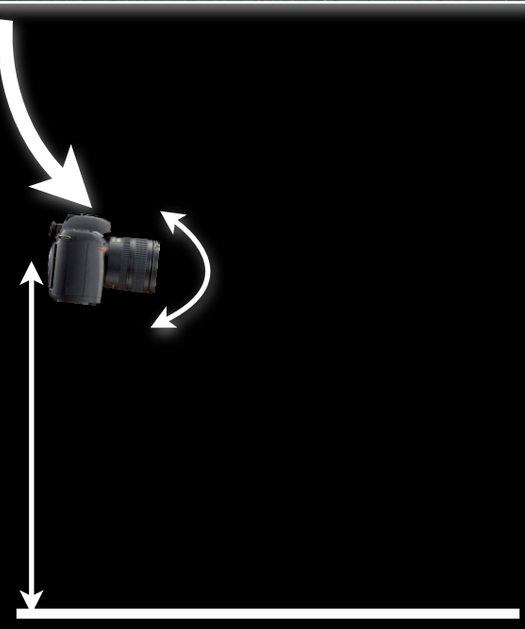
Lalonde et al., SIGGRAPH 2007

Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)





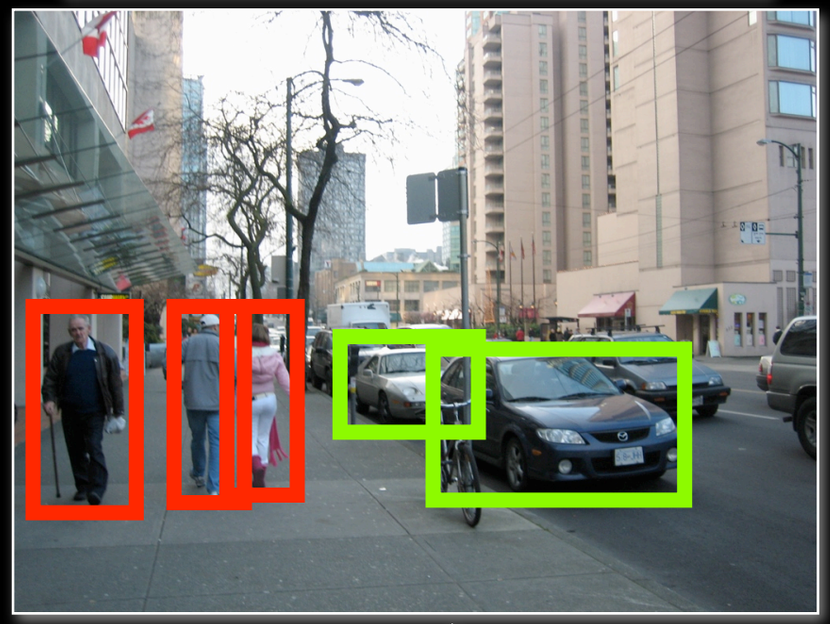


Lalonde et al., SIGGRAPH 2007

Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

Car height distribution
1.5 +/- 0.19 m
(automatically learned)

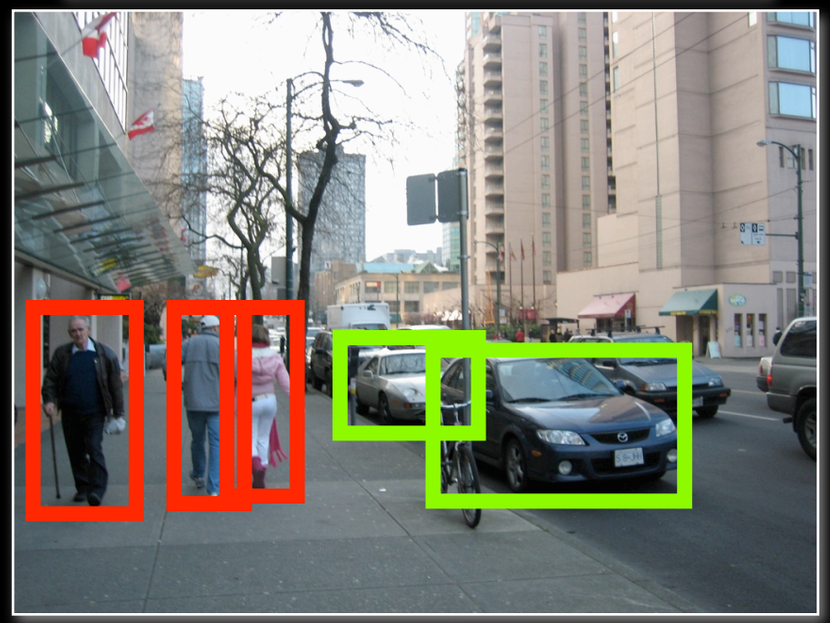


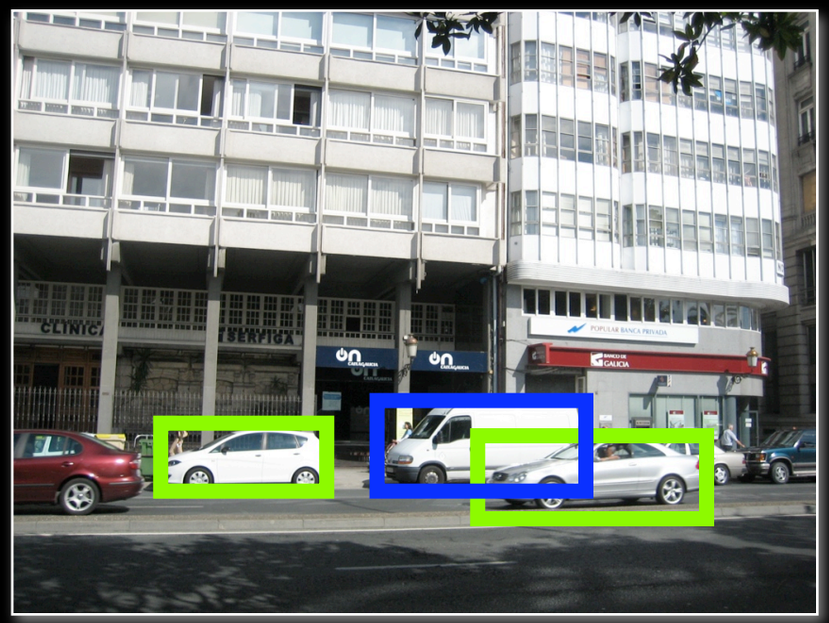
Lalonde et al., SIGGRAPH 2007

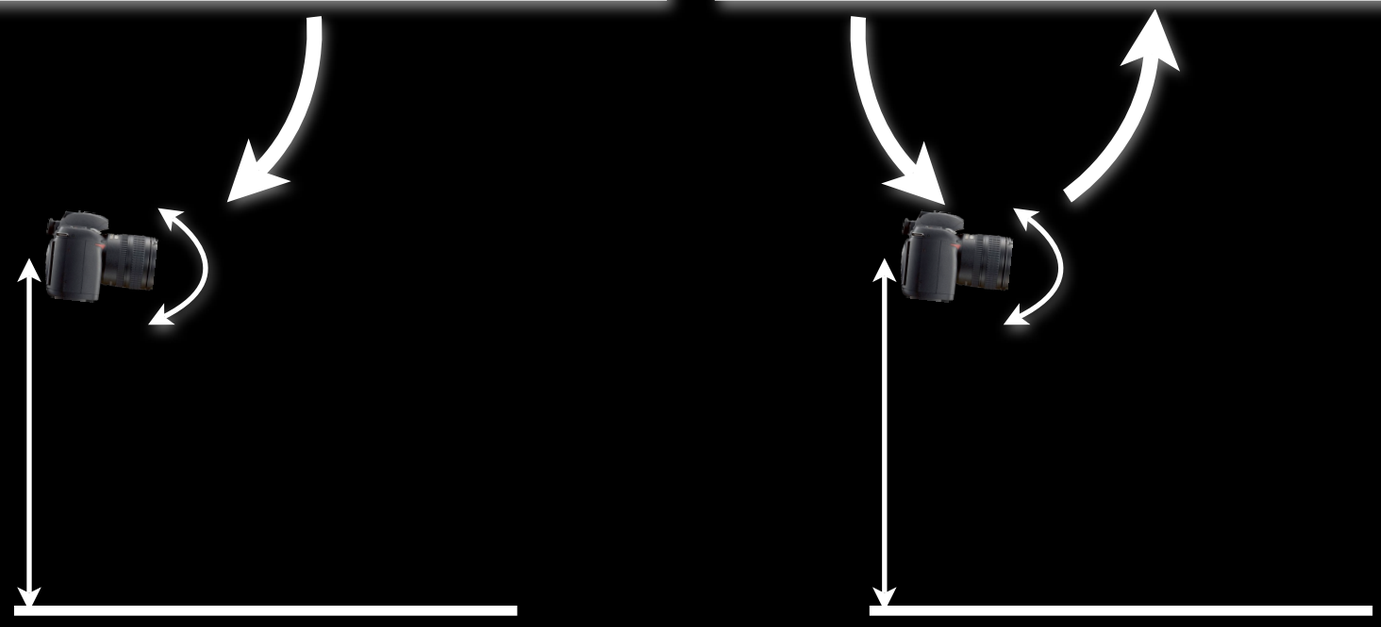
Camera parameters

Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

Car height distribution
1.5 +/- 0.19 m
(automatically learned)





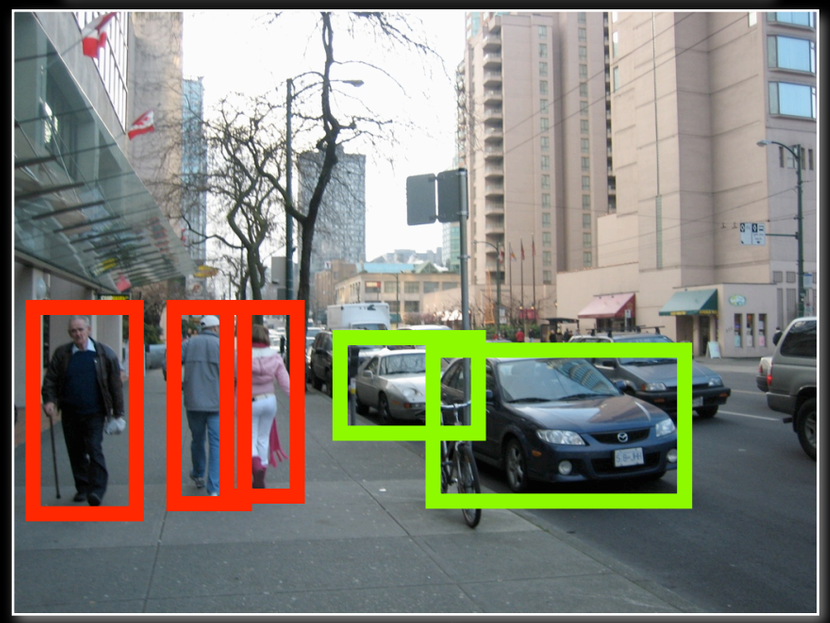


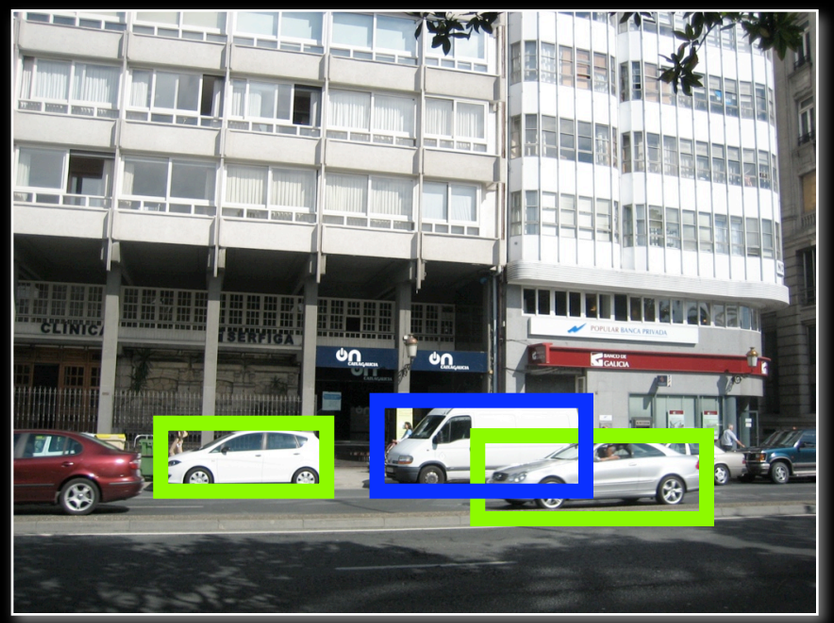
Lalonde et al., SIGGRAPH 2007

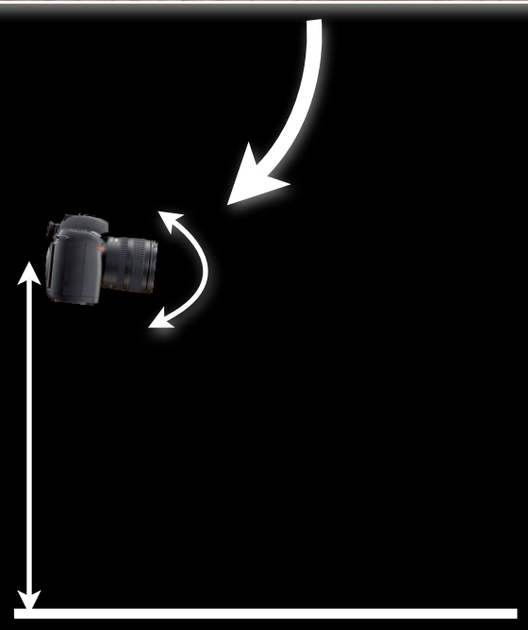
Camera parameters

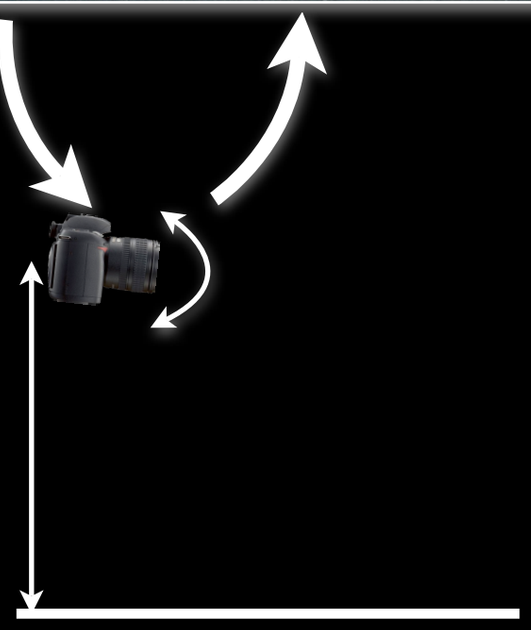
Human height distribution
1.7 +/- 0.085 m
(National Center for Health Statistics)

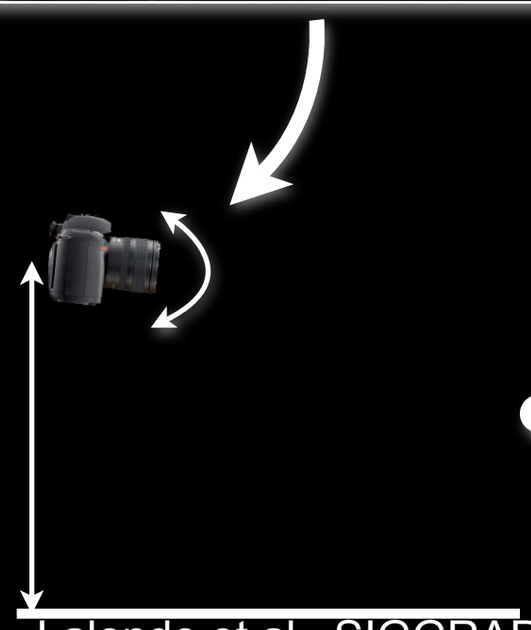
Car height distribution
1.5 +/- 0.19 m
(automatically learned)











Lalonde et al., SIGGRAPH 2007

Object heights

Database image

Lalonde et al., SIGGRAPH 2007

Object heights

Database image

Pixel heights

Lalonde et al., SIGGRAPH 2007

Object heights

Database image

Pixel heights

Real heights

Lalonde et al., SIGGRAPH 2007

Estimated object heights

Object	Estimated average height (m)
Car	1.51
Man	1.80
Woman	1.67
Parking meter	1.36
Fire hydrant	0.87

Lalonde et al., SIGGRAPH 2007

Geometry is not enough

Lalonde et al., SIGGRAPH 2007

Geometry is not enough

Lalonde et al., SIGGRAPH 2007

Geometry is not enough

Lalonde et al., SIGGRAPH 2007

Geometry is not enough

Lalonde et al., SIGGRAPH 2007

Illumination context

- Exact environment map is impossible
  - Approximations [Khan *et al.*, ‘06]

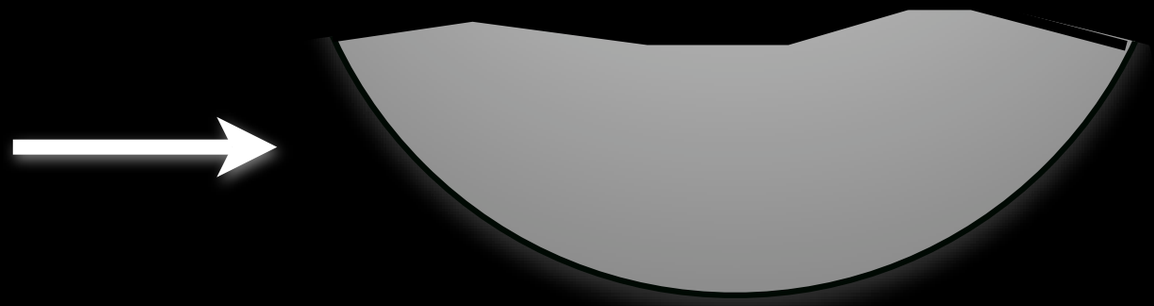
Database image

Illumination context

- Exact environment map is impossible
  - Approximations [Khan *et al.*, ‘06]

Database image

Environment map rough approximation

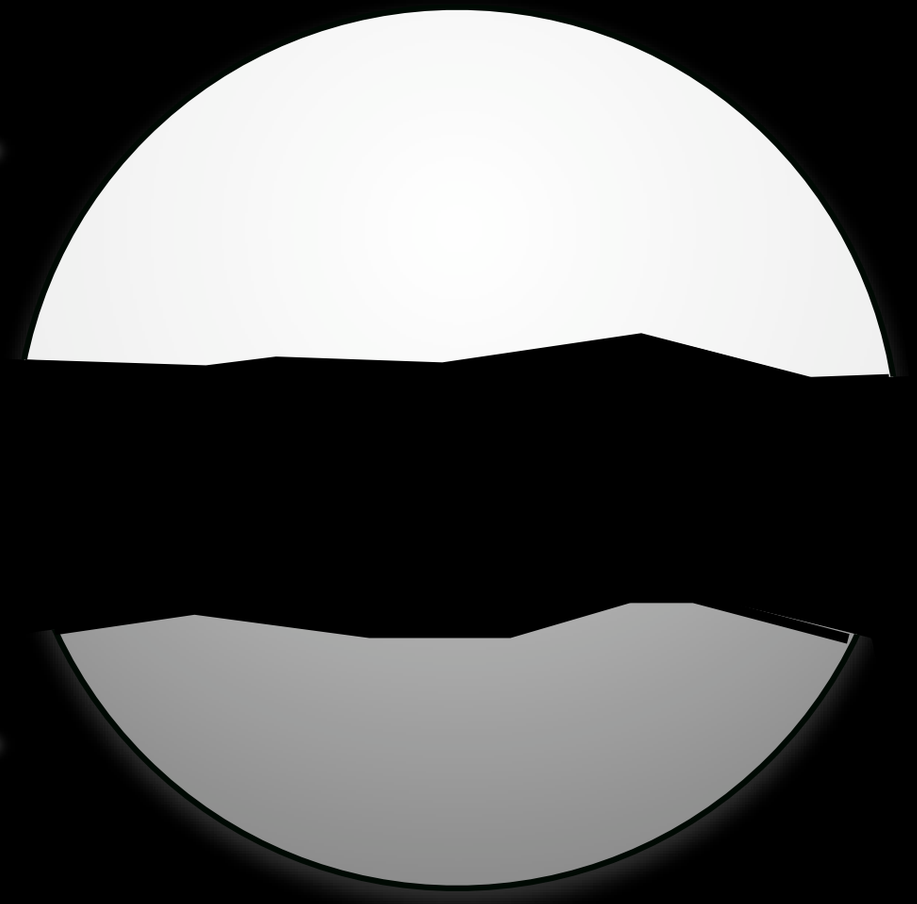


Illumination context

- Exact environment map is impossible
  - Approximations [Khan *et al.*, '06]

Database image

Environment map rough approximation

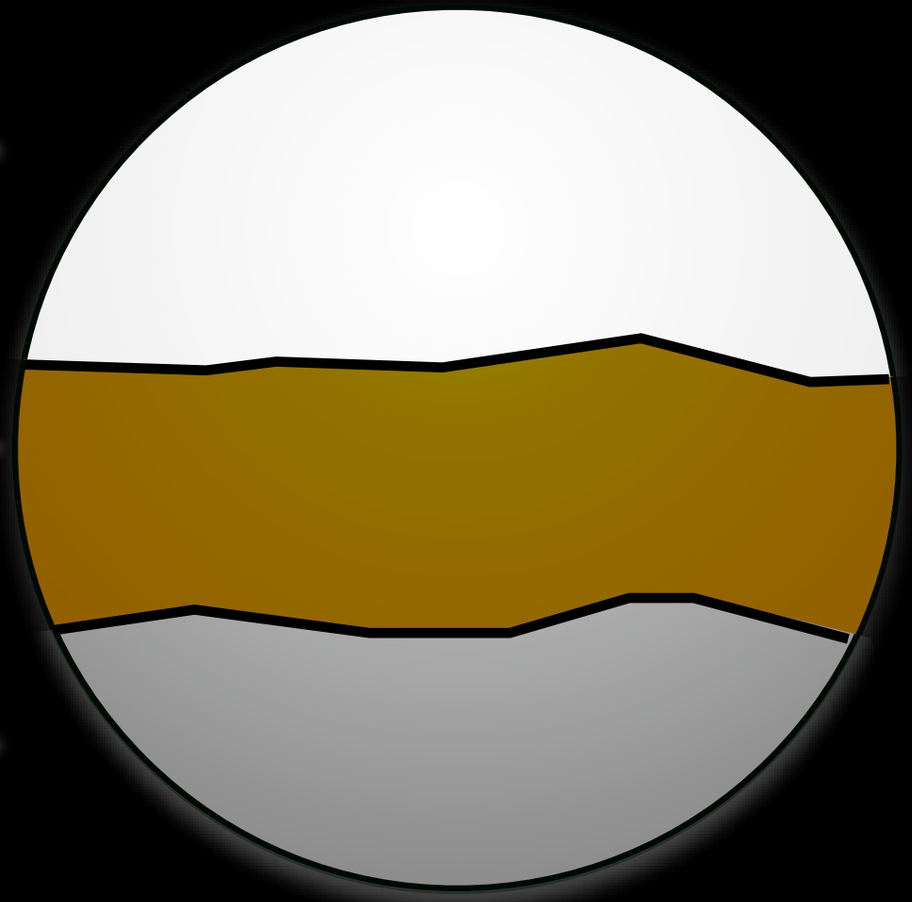


Illumination context

- Exact environment map is impossible
  - Approximations [Khan *et al.*, '06]

Database image

Environment map rough approximation



Lalonde et al., SIGGRAPH 2007

Illumination context

Database image

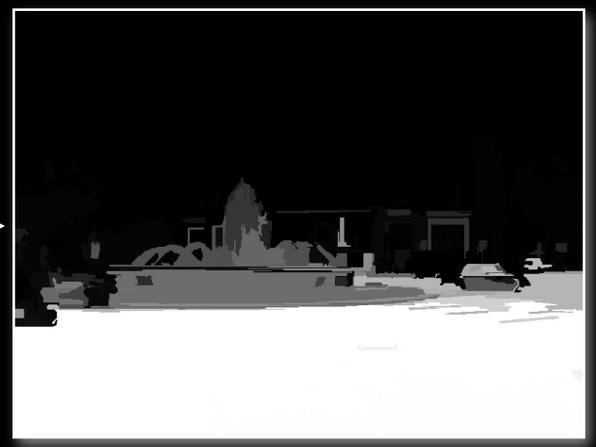
Automatic Photo Popup
Hoiem *et al.*, SIGGRAPH ‘05

Illumination context

Database image

Automatic Photo Popup
Hoiem *et al.*, SIGGRAPH ‘05

P(pixel|class)



Lalonde et al., SIGGRAPH 2007

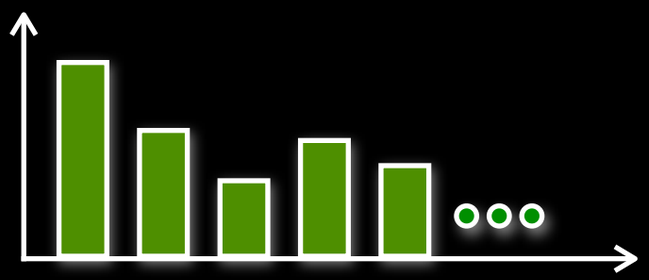
Illumination context

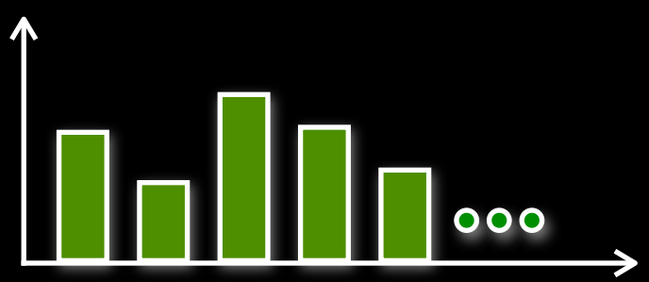
Database image

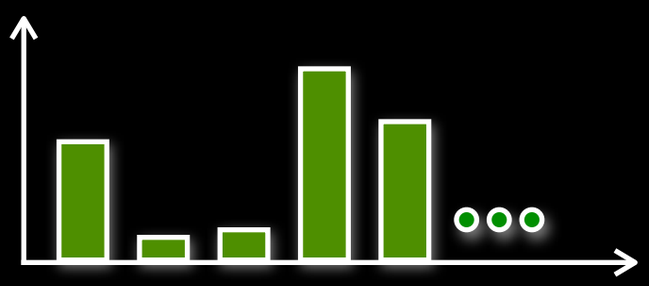
Automatic Photo Popup
Hoiem *et al.*, SIGGRAPH ‘05

P(pixel|class)

CIE *L*a*b** histograms







Lalonde et al., SIGGRAPH 2007

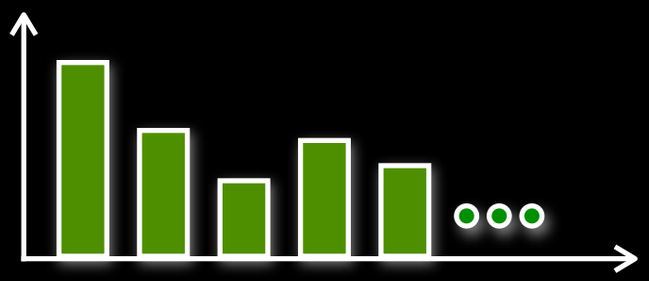
Illumination context

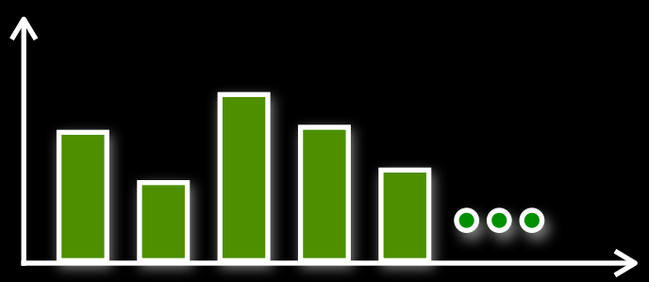
Database image

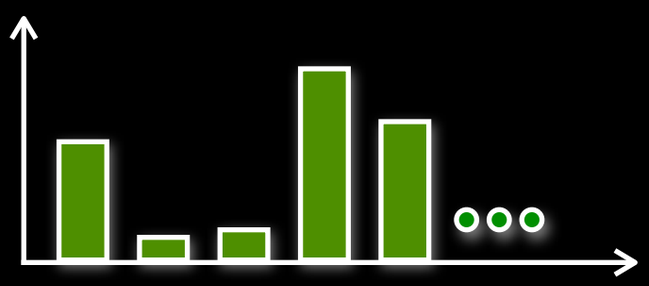
Automatic Photo Popup
Hoiem *et al.*, SIGGRAPH ‘05

P(pixel|class)

CIE *L*a*b** histograms







- Low Dynamic Range
  - matching vs. relighting

Lalonde et al., SIGGRAPH 2007

Illumination nearest-neighbors

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: local context

Lalonde et al., SIGGRAPH 2007

Other criteria: segmentation

- LabelMe contributers not always reliable
- Segmentation quality

Lalonde et al., SIGGRAPH 2007

Other criteria: segmentation

- LabelMe contributers not always reliable
- Segmentation quality

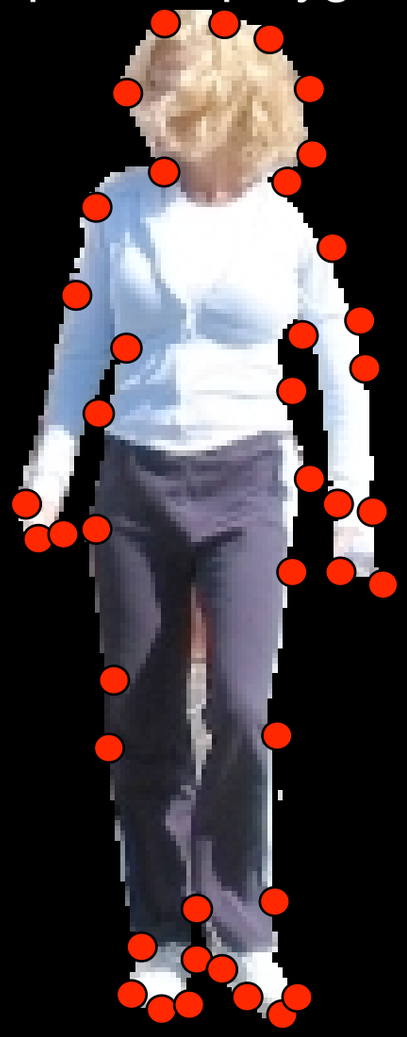
38 points / polygon

Lalonde et al., SIGGRAPH 2007

Other criteria: segmentation

- LabelMe contributers not always reliable
- Segmentation quality

38 points / polygon



4 points / polygon

Lalonde et al., SIGGRAPH 2007

Other criteria: blur

- Resolution: avoid up-sampling

x3 up-sampling

Lalonde et al., SIGGRAPH 2007

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label

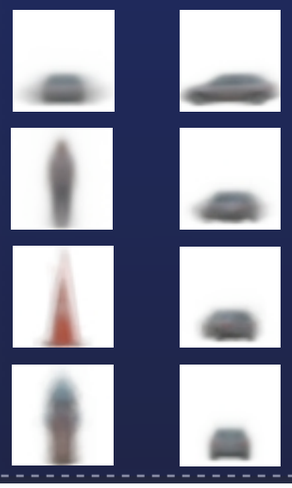
Phase II: Object insertion

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label Cluster



Phase II: Object insertion

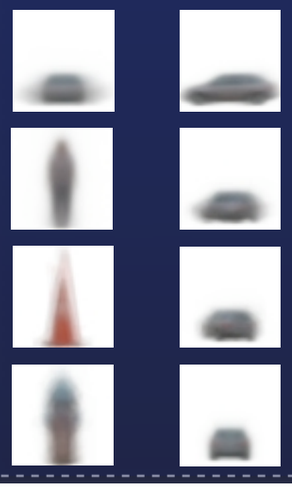
Lalonde et al., SIGGRAPH 2007

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label Cluster 3-D height



Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label

Cluster

3-D height

Camera

Phase II: Object insertion

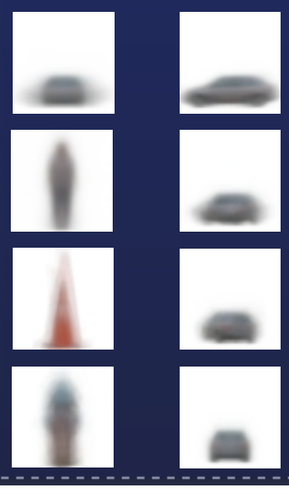
Lalonde et al., SIGGRAPH 2007

Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label Cluster 3-D height Camera Illumination context



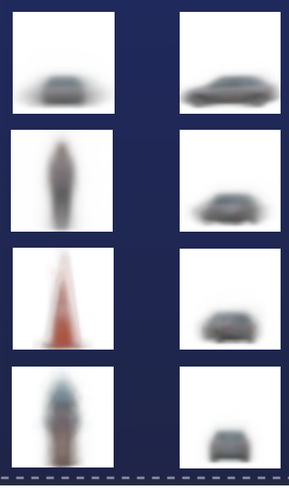
Phase II: Object insertion

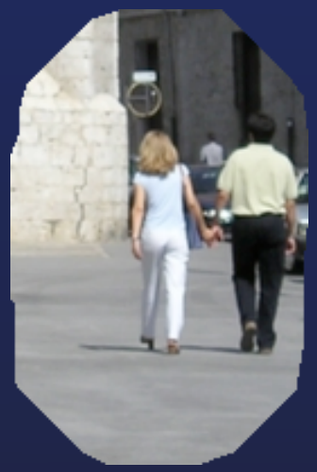
Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label | Cluster | 3-D height | Camera | Illumination context | Local context





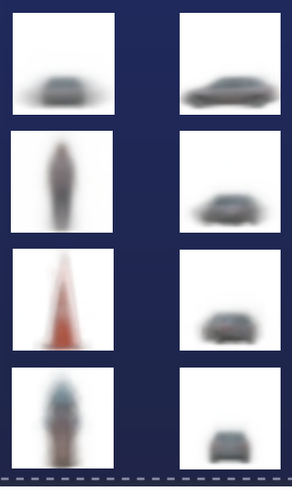
Phase II: Object insertion

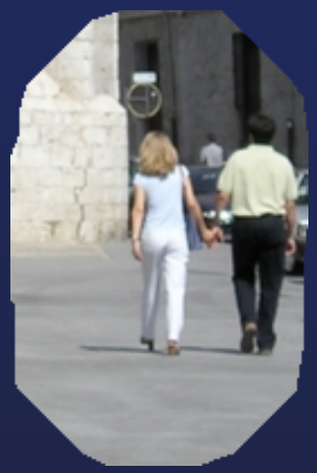
Recap

Phase I: Database annotation

Object properties (used for sorting the database)

Label | Cluster | 3-D height | Camera | Illumination context | Local context | Segmentation





Phase II: Object insertion

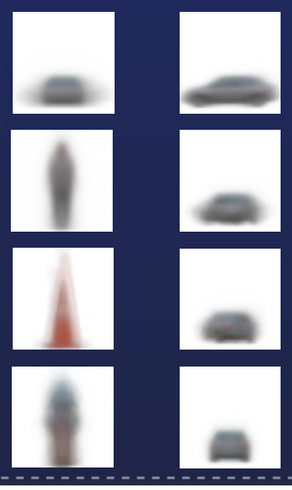
Lalonde et al., SIGGRAPH 2007

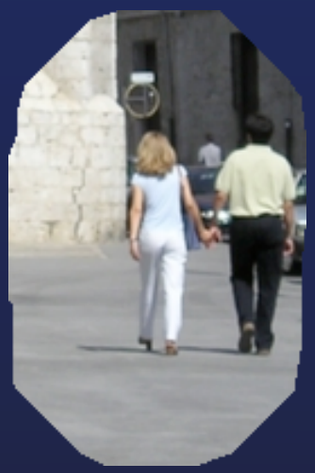
Recap

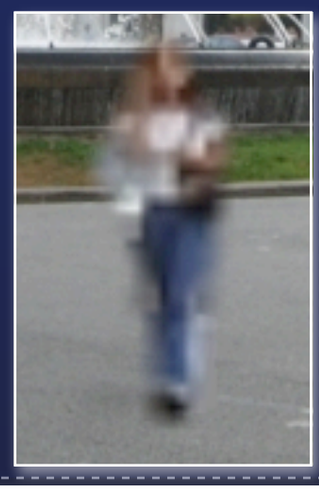
Phase I: Database annotation

Object properties (used for sorting the database)

Label | Cluster | 3-D height | Camera | Illumination context | Local context | Segmentation | Blur







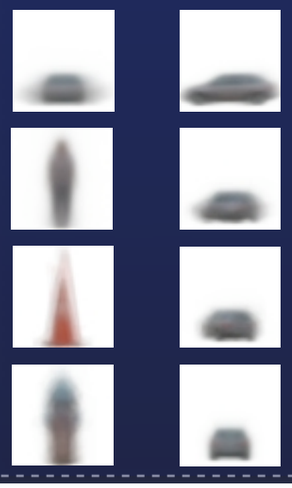
Phase II: Object insertion

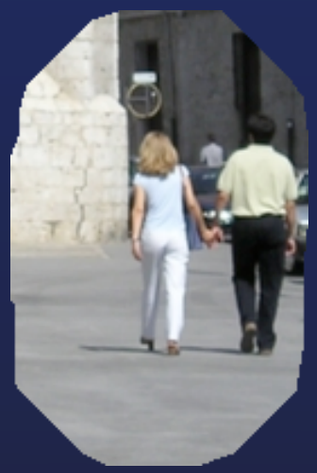
Recap

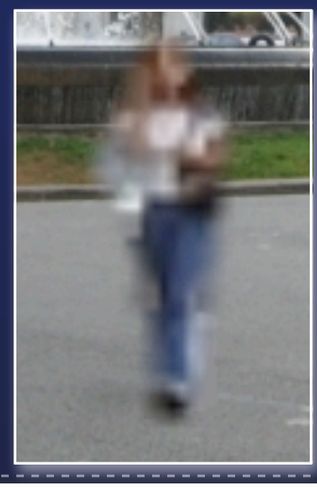
Phase I: Database annotation

Object properties (used for sorting the database)

Label | Cluster | 3-D height | Camera | Illumination context | Local context | Segmentation | Blur







Phase II: Object insertion

Lalonde et al., SIGGRAPH 2007

Let’s insert an object!

- Poor user-provided segmentations
- Noticeable seams

al., SIGGRAPH 2007

Let’s insert an object!

- Poor user-provided segmentations
- Noticeable seams

al., SIGGRAPH 2007

Segmentation & blending

Blending

Spline blending
[Burt and Adelson, ‘83]

Poisson image editing
[Pérez et al., ’03]

Lalonde et al., SIGGRAPH 2007

Segmentation & blending

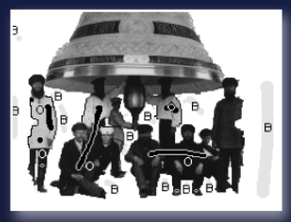
Blending

Spline blending
[Burt and Adelson, ‘83]

Poisson image editing
[Pérez et al., ’03]

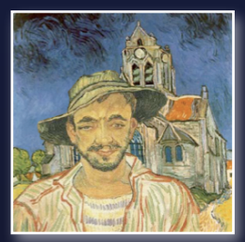
Cutting

Graph cut segmentation
[Boykov and Jolly, ’01]



GrabCut
[Rother et al., ’04]

Lazy Snapping
[Li et al., ’04]



Lalonde et al., SIGGRAPH 2007

Segmentation & blending

Blending

Spline blending
[Burt and Adelson, ‘83]

Poisson image editing
[Pérez et al., ’03]

Cutting

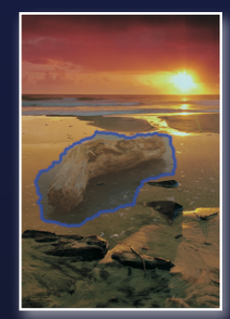
Graph cut segmentation
[Boykov and Jolly, ’01]

GrabCut
[Rother et al., ’04]

Lazy Snapping
[Li et al., ’04]

Both!

Drag-and-Drop Pasting
[Jia et al., ’06]



Lalonde et al., SIGGRAPH 2007

Segmentation & blending

Blending

Spline blending
[Burt and Adelson, ‘83]

Poisson image editing
[Pérez et al., ’03]

Cutting

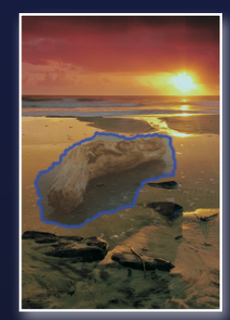
Graph cut segmentation
[Boykov and Jolly, ’01]

GrabCut
[Rother et al., ’04]

Lazy Snapping
[Li et al., ’04]

Both!

Drag-and-Drop Pasting
[Jia et al., ’06]



- Need to be automatic
- More challenging images

Lalonde et al., SIGGRAPH 2007

Segmentation refinement

- Graph Cut [Boykov and Jolly, 2001]
- Flux shape prior [Kolmogorov and Boykov, 2005]
  - Overcomes the shrinking bias
  - Does not over-smooth

Database image

Lalonde et al., SIGGRAPH 2007

Segmentation refinement

- Graph Cut [Boykov and Jolly, 2001]
- Flux shape prior [Kolmogorov and Boykov, 2005]
  - Overcomes the shrinking bias
  - Does not over-smooth

Database image

Band restriction for GC

Lalonde et al., SIGGRAPH 2007

Segmentation refinement

- Graph Cut [Boykov and Jolly, 2001]
- Flux shape prior [Kolmogorov and Boykov, 2005]
  - Overcomes the shrinking bias
  - Does not over-smooth

Database image

Band restriction for GC

GC segmentation

Lalonde et al., SIGGRAPH 2007

Segmentation refinement

- Graph Cut [Boykov and Jolly, 2001]
- Flux shape prior [Kolmogorov and Boykov, 2005]
  - Overcomes the shrinking bias
  - Does not over-smooth

Database image

Band restriction for GC

GC segmentation

With flux

Lalonde et al., SIGGRAPH 2007

Context-sensitive blending

- Blending: additional term to prevent discoloration
- Mask: similarity between fg & bg statistics

Input image

Lalonde et al., SIGGRAPH 2007

Context-sensitive blending

- Blending: additional term to prevent discoloration
- Mask: similarity between fg & bg statistics

Input image

Poisson editing [Perez *et al.*, ‘03]

Lalonde et al., SIGGRAPH 2007

Context-sensitive blending

- Blending: additional term to prevent discoloration
- Mask: similarity between fg & bg statistics

Input image

Poisson editing
[Perez *et al.*, ‘03]

Compositing result

Lalonde et al., SIGGRAPH 2007

Shadows are important [Kersten *et al.*, ‘96]

Shadows are important [Kersten *et al.*, ‘96]

Shadows are important [Kersten *et al.*, ‘96]

- Not so sensitive to shadow **direction** [Cavanagh, 2005]

Shadow transfer

Lalonde et al., SIGGRAPH 2007

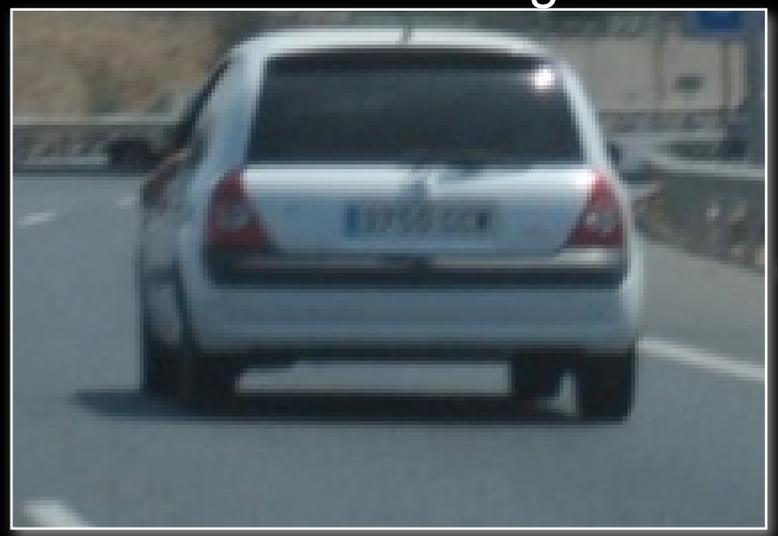
Shadow transfer

Database image

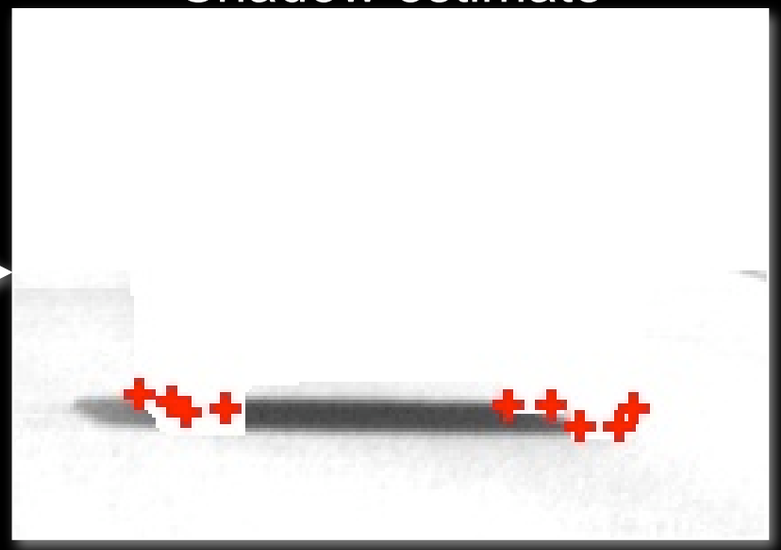
Lalonde et al., SIGGRAPH 2007

Shadow transfer

Database image



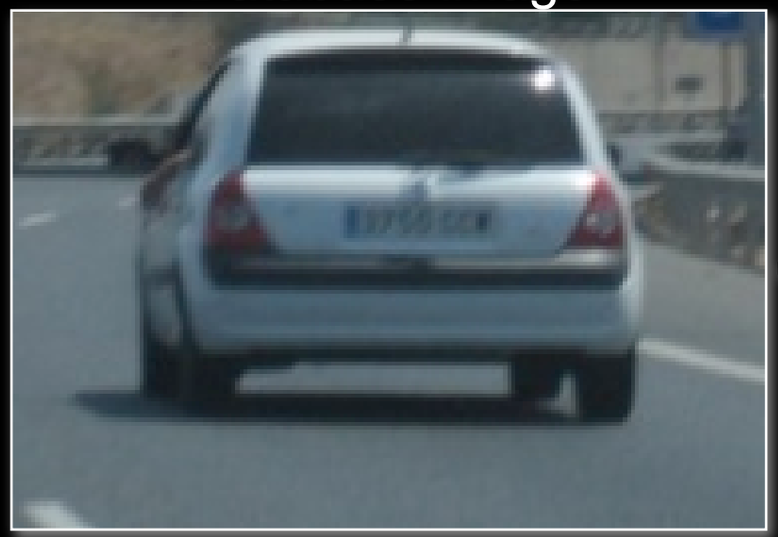
Shadow estimate



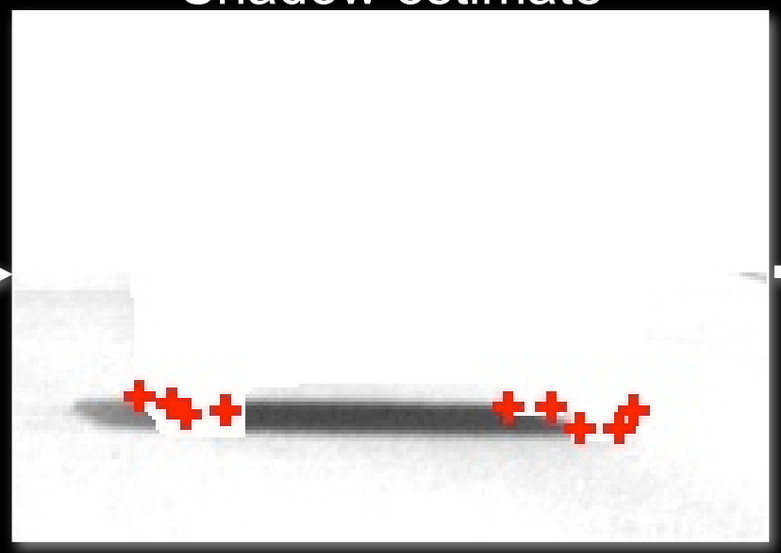
Lalonde et al., SIGGRAPH 2007

Shadow transfer

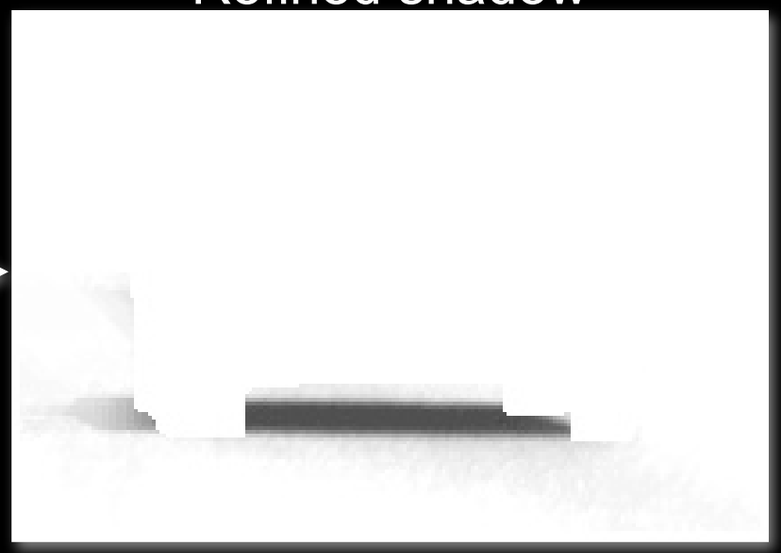
Database image



Shadow estimate



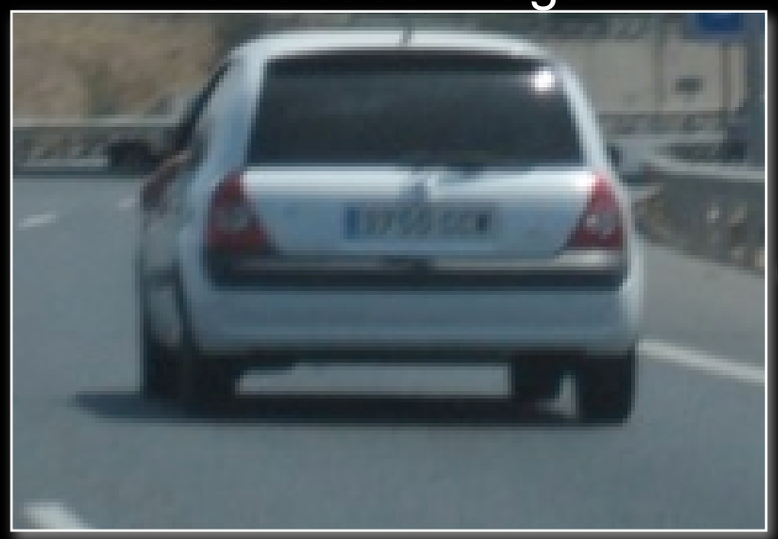
Refined shadow



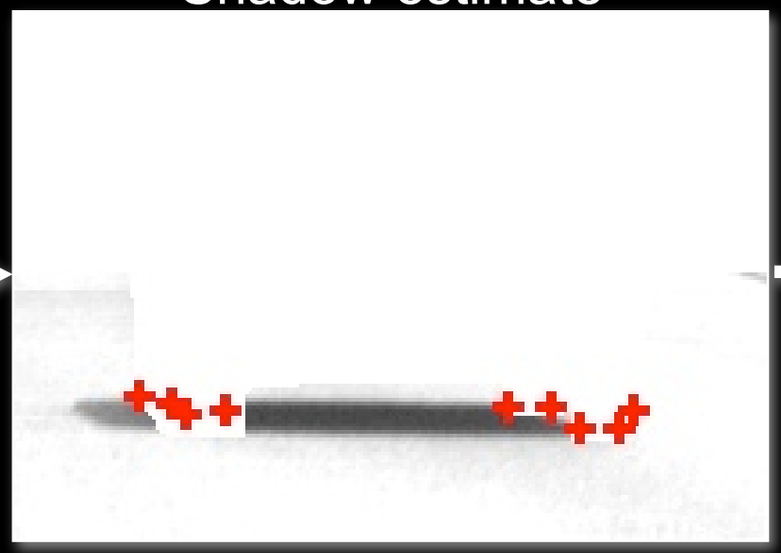
Lalonde et al., SIGGRAPH 2007

Shadow transfer

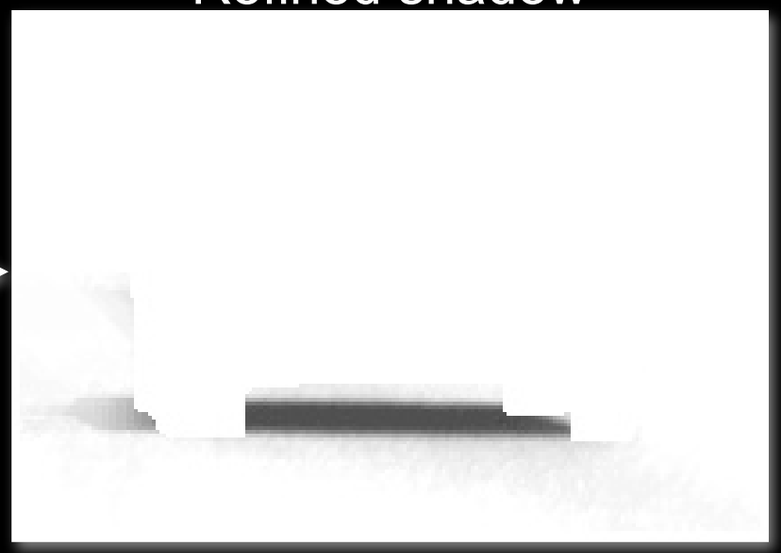
Database image



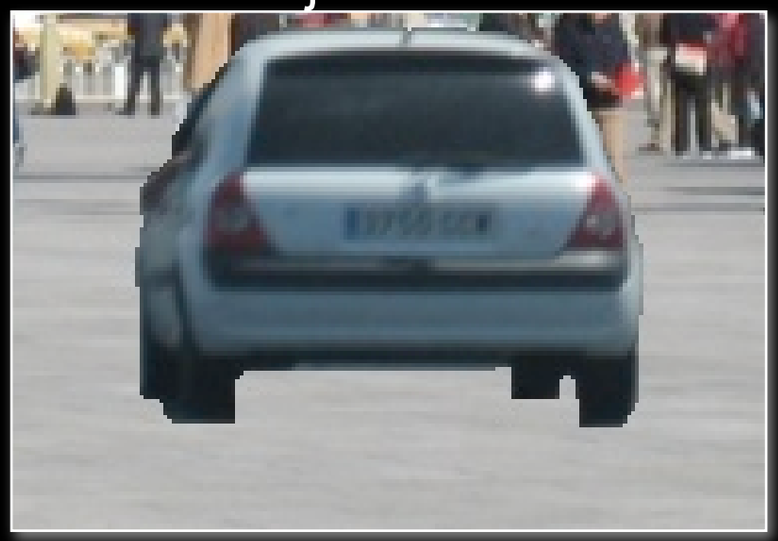
Shadow estimate



Refined shadow



Object alone

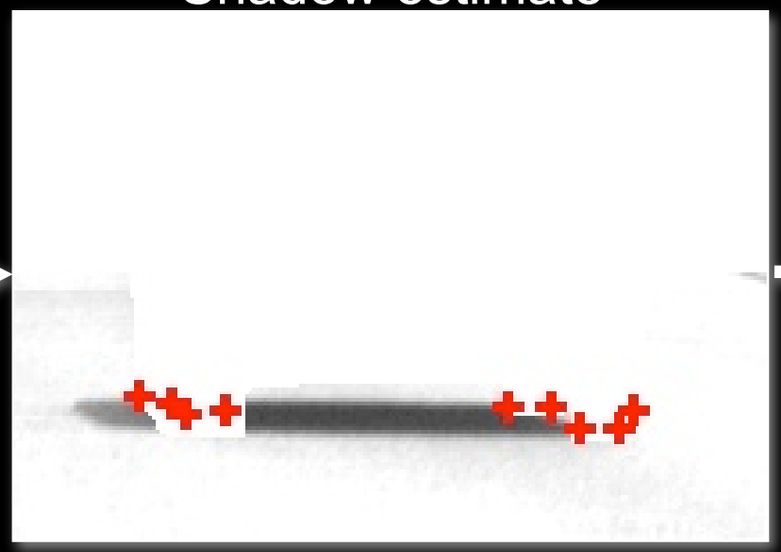


Lalonde et al., SIGGRAPH 2007

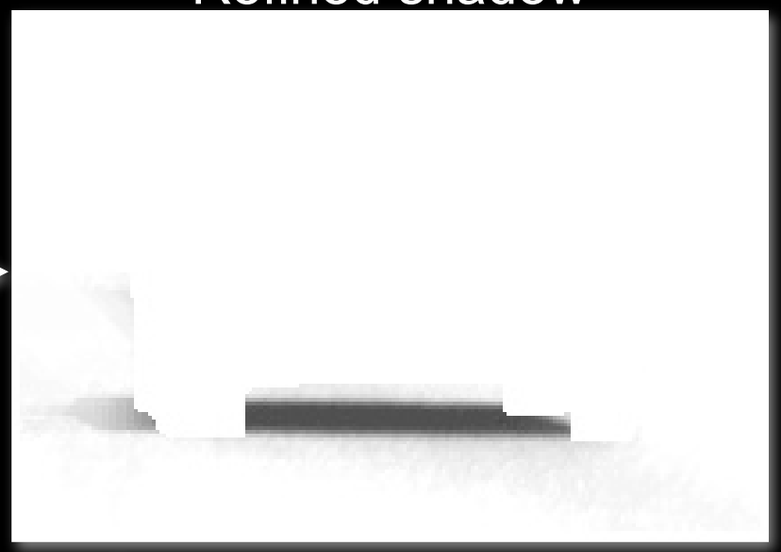
Shadow transfer

Database image

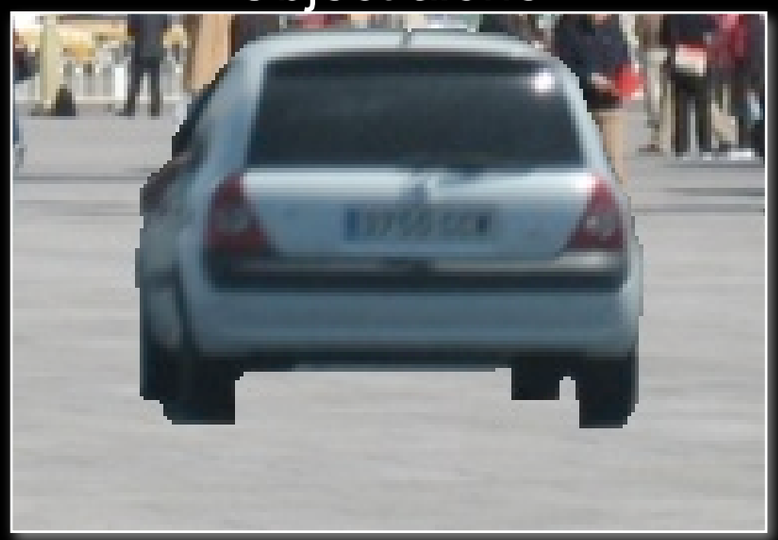
Shadow estimate



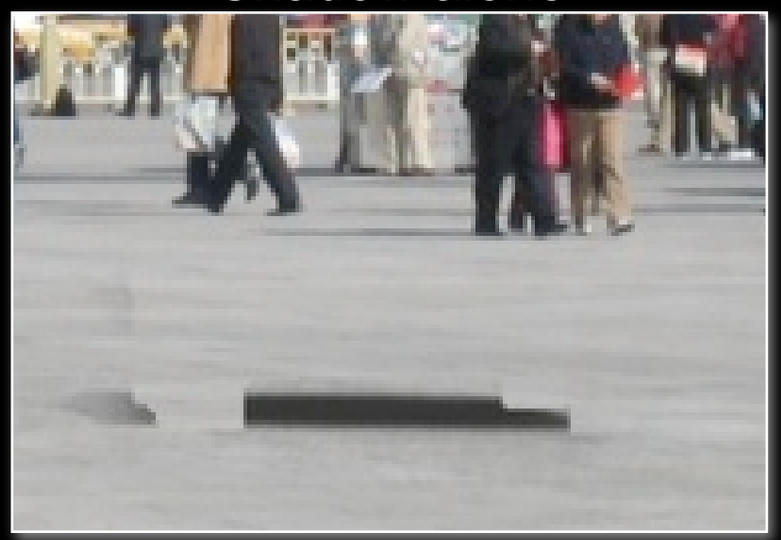
Refined shadow



Object alone



Shadow alone

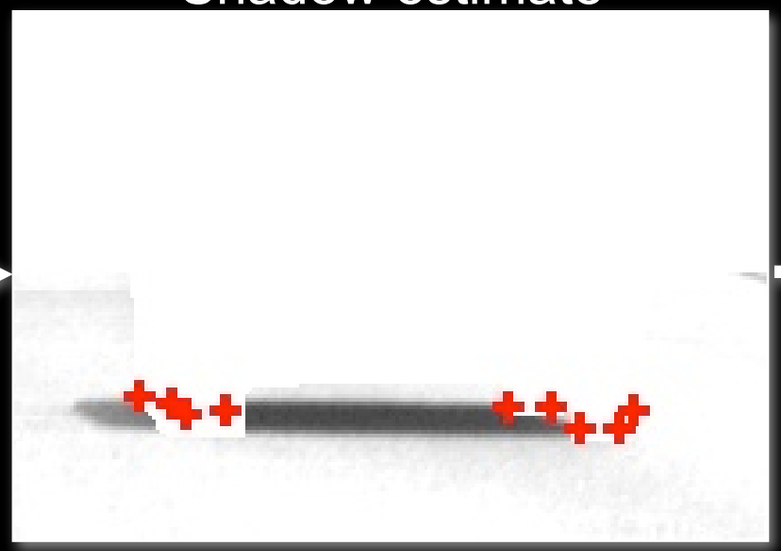


Lalonde et al., SIGGRAPH 2007

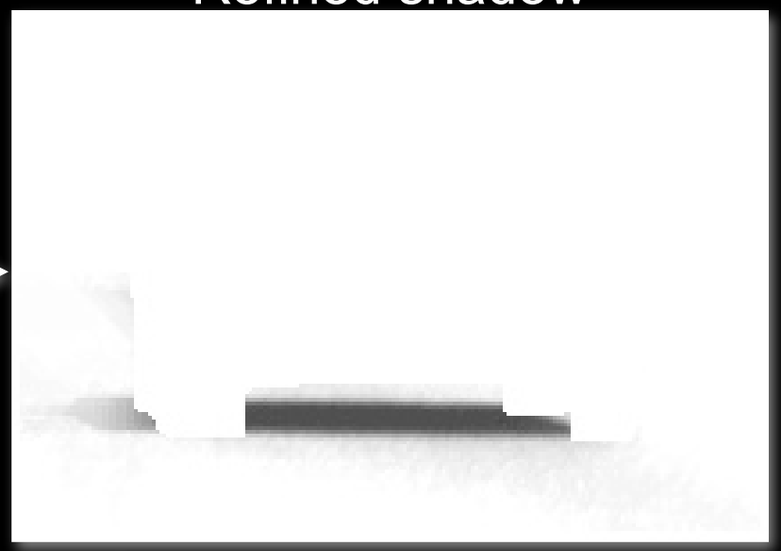
Shadow transfer

Database image

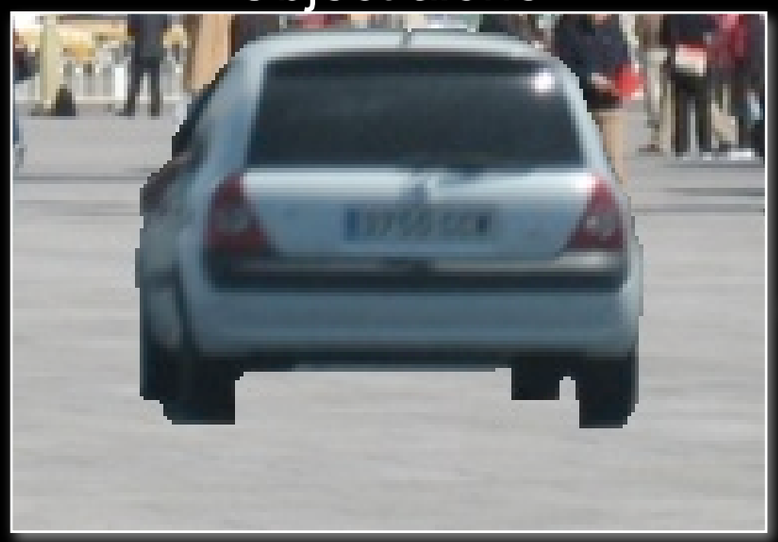
Shadow estimate



Refined shadow

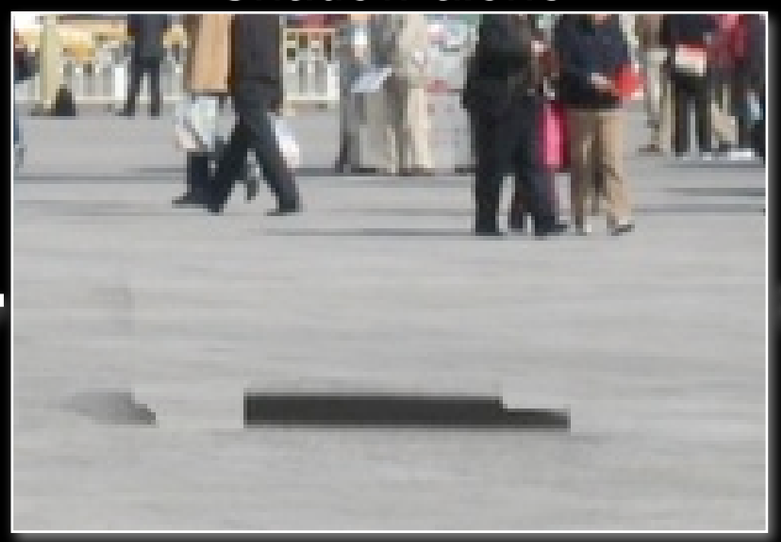


Object alone



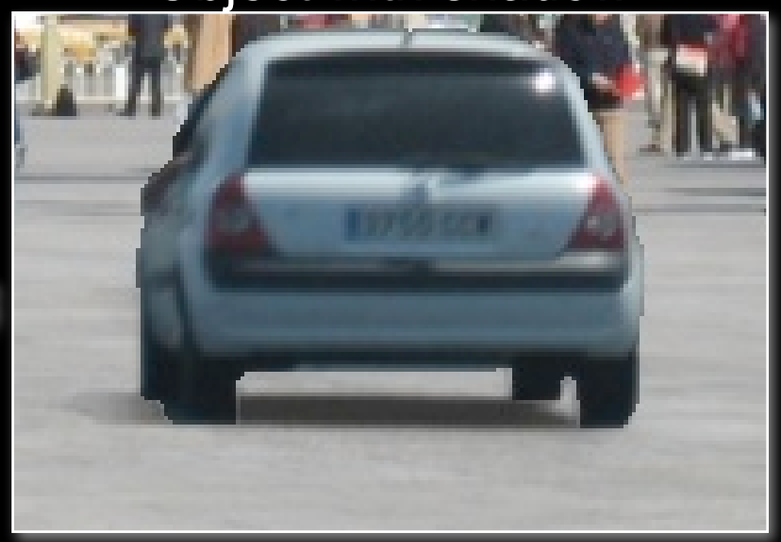
+

Shadow alone



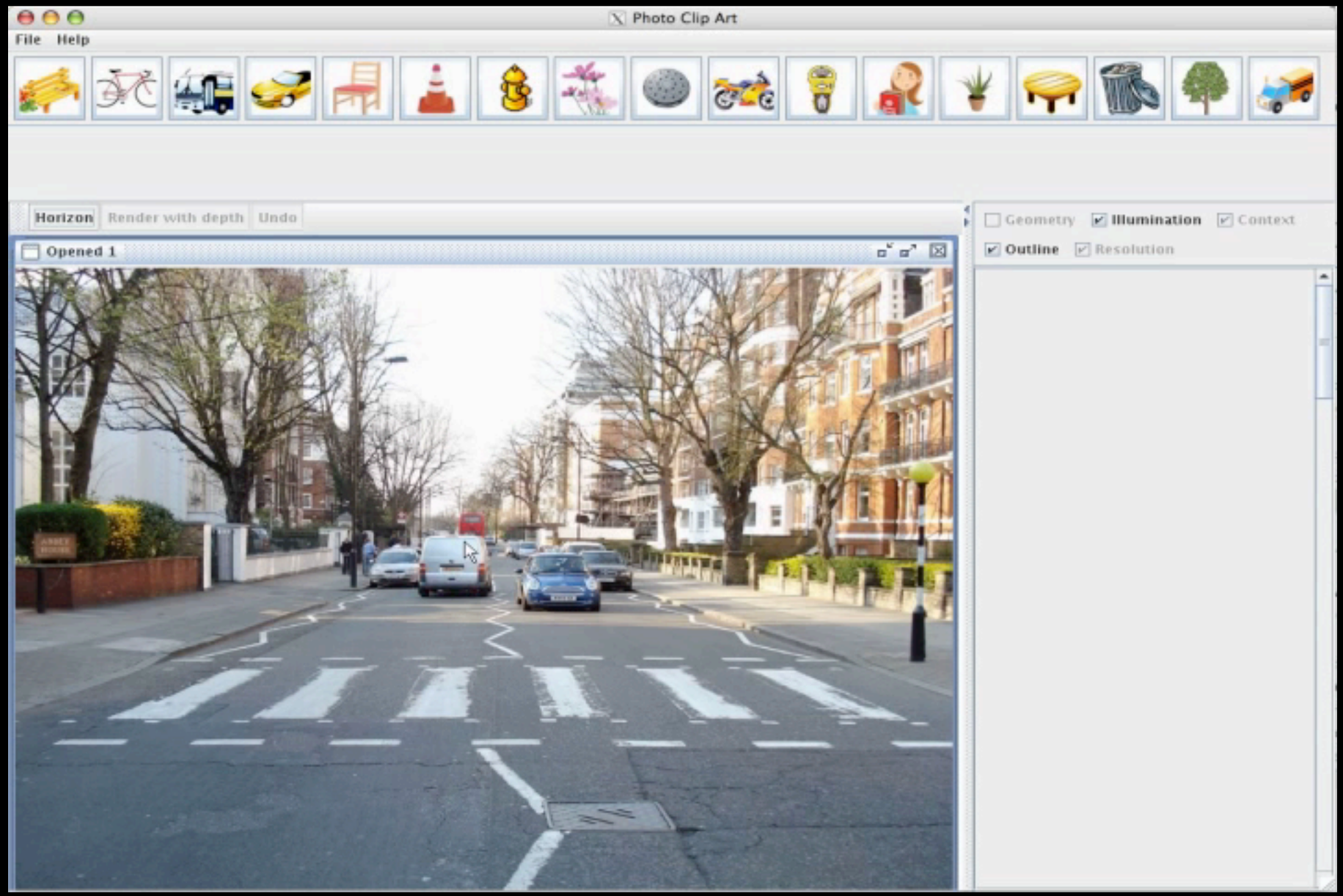
=

Object with shadow



Lalonde et al., SIGGRAPH 2007

User interface



Lalonde et al., SIGGRAPH 2007

Street accident

Lalonde et al., SIGGRAPH 2007

Street accident

Lalonde et al., SIGGRAPH 2007

Bridge

Bridge

Painting

Lalonde et al., SIGGRAPH 2007

Painting

Lalonde et al., SIGGRAPH 2007

Alley

Lalonde et al., SIGGRAPH 2007

Alley

Lalonde et al., SIGGRAPH 2007

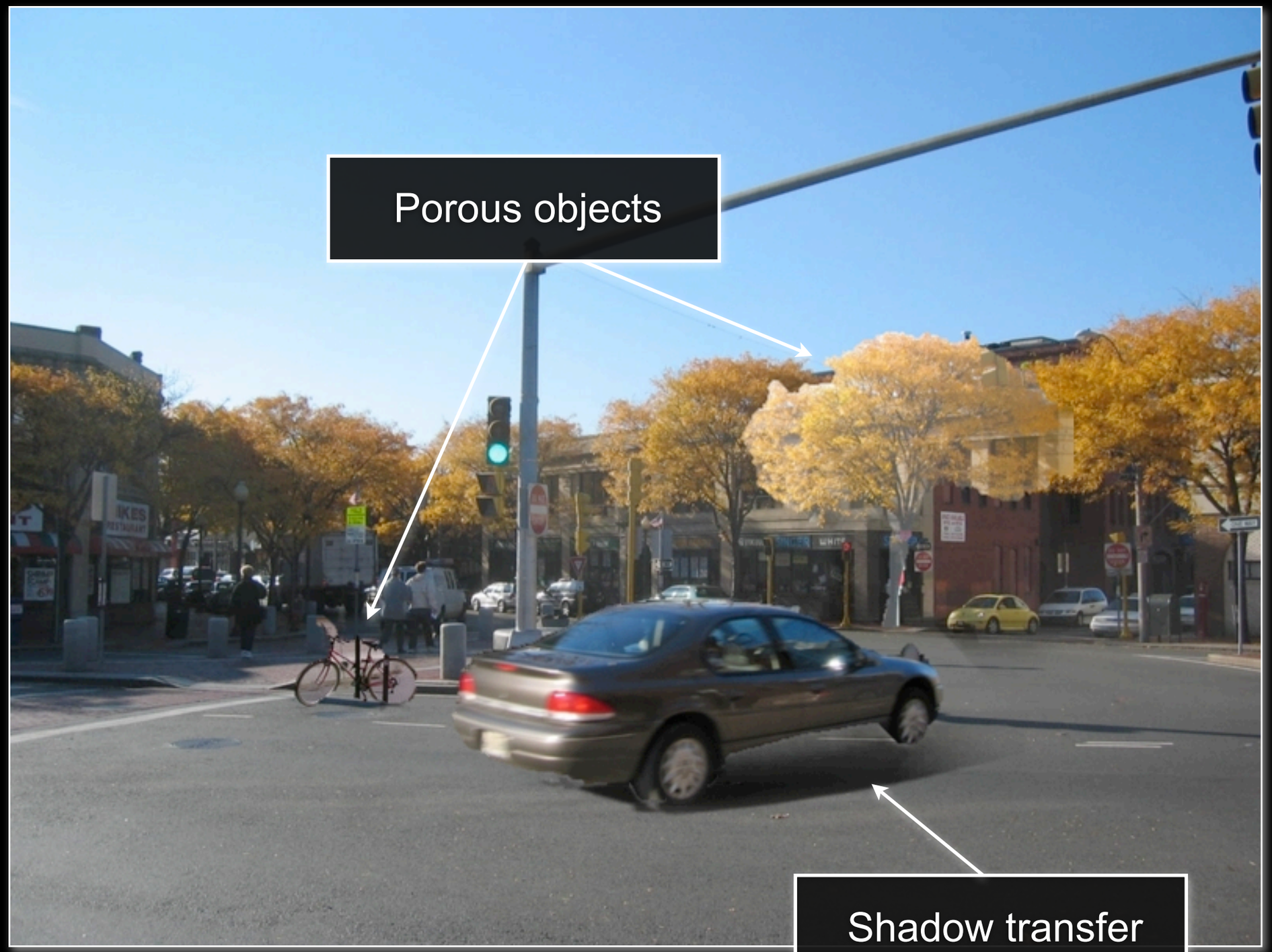
Failure cases

Lalonde et al., SIGGRAPH 2007

Failure cases

Lalonde et al., SIGGRAPH 2007

Failure cases

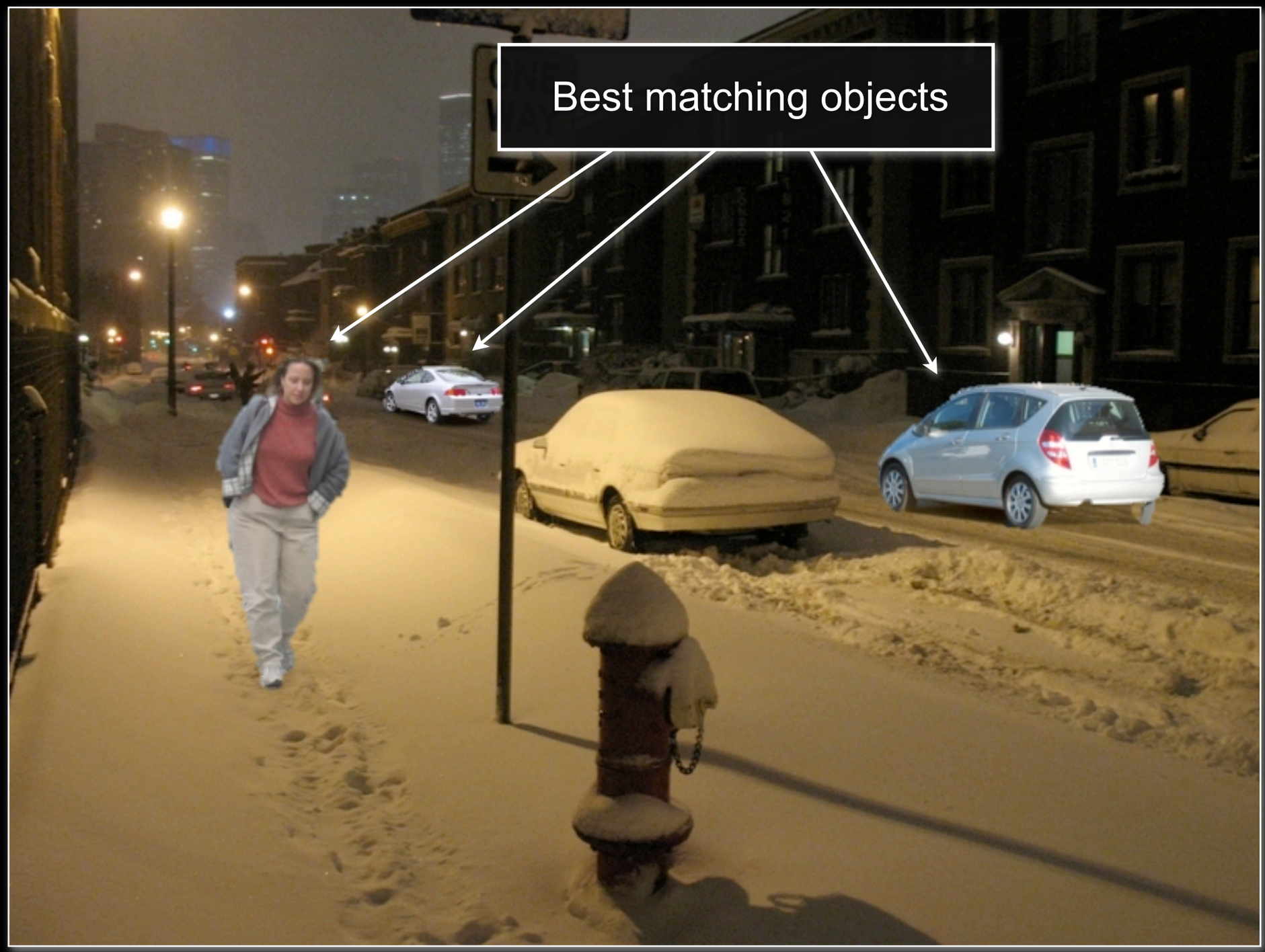


Lalonde et al., SIGGRAPH 2007

Failure cases

Lalonde et al., SIGGRAPH 2007

Failure cases



Lalonde et al., SIGGRAPH 2007

Conclusion

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven
  - Thinking in 3-D

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven

  - Thinking in 3-D

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven
  - Thinking in 3-D

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven
  - Thinking in 3-D

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven

  - Thinking in 3-D

- Plenty of challenges remaining

Lalonde et al., SIGGRAPH 2007

Conclusion

- Photo Clip Art: intuitive and photo-realistic
  - Data-driven

  - Thinking in 3-D

- Plenty of challenges remaining
- Many of them will diminish: Add more data!

Lalonde et al., SIGGRAPH 2007

Thank you!