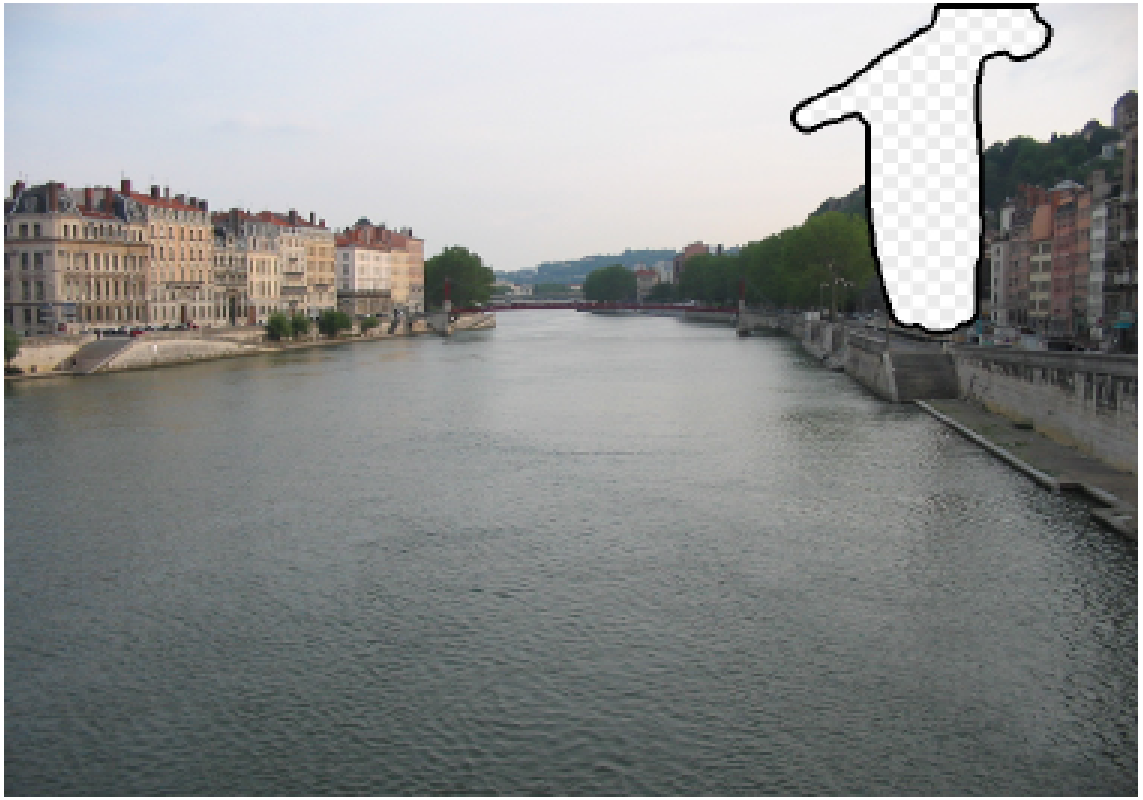
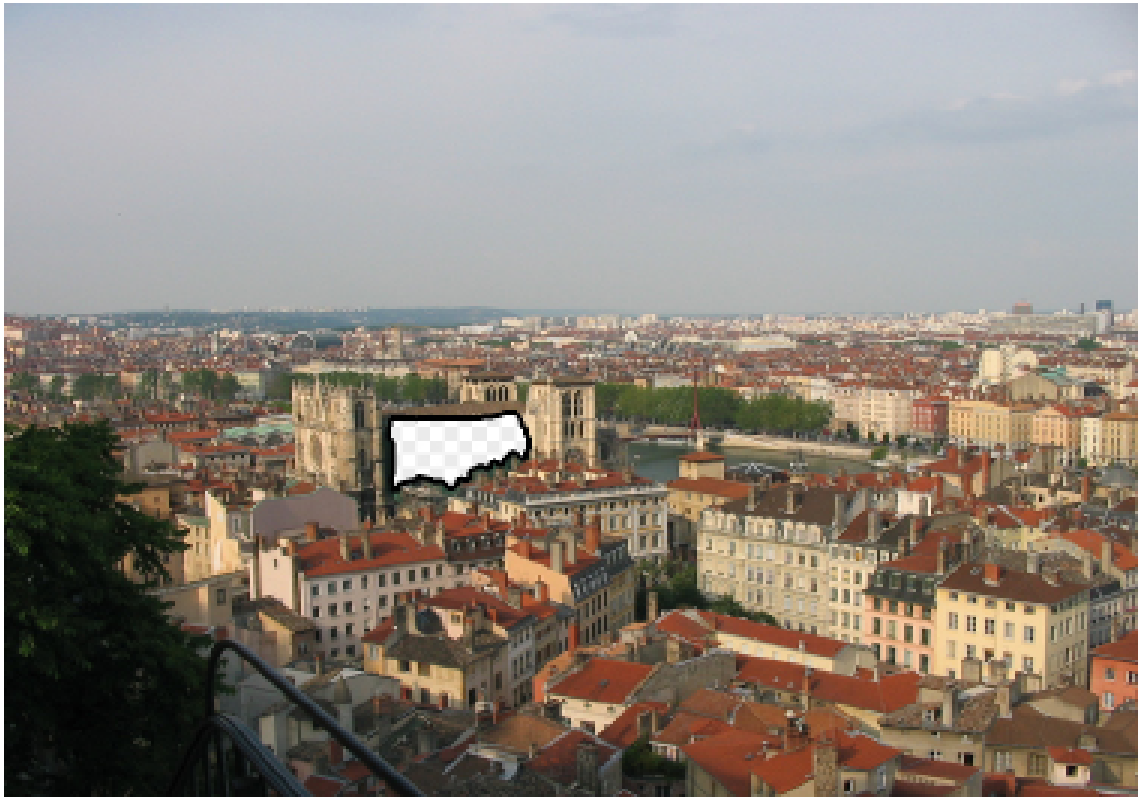
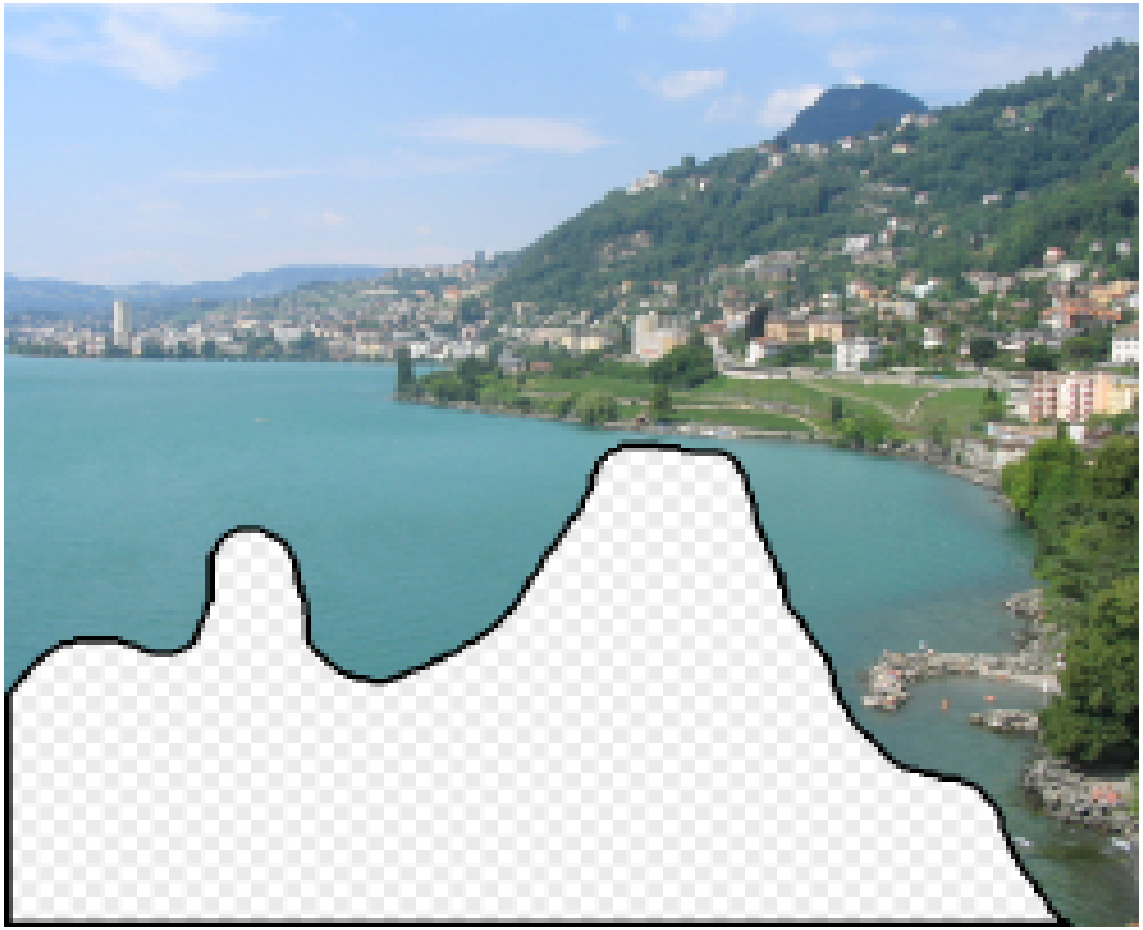
Scene Completion Using Millions of Photographs, Hays and Efros, International Conference on Computer Graphics and Interactive Techniques, 2007

Rich Turner (turner@gatsby.ucl.ac.uk)

Gatsby Unit, 23/08/2007







Algorithm

Database: 2×10^6 images (400Gb) from Flickr

1. **Find subset** of images depicting semantically similar scenes

2. **Find patches** in subset that match the context surrounding the missing region

3. **Blend** in the most similar patches

Write down a suitable cost-function at each stage...

1. Finding the most similar subset

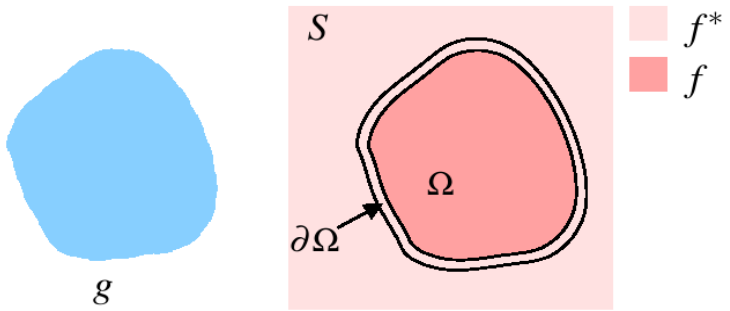
- Purpose:
  - Ensures filled in regions have sensible content.
  - Reduces the amount of data (to 200 images) for subsequent stages

- Measure similarity using:
  - colour histogram
  - a low dimensional representation of the images (GIST scene descriptors)

2. Finding matching patches

- Define some region **around the missing area**
- Find patches in subset that are a "closest" match to this region
- Distance = pixel-wise error + texture dissimilarity
- Search across translations and scales (small translations cost less)

3. Blending: Poisson Interpolation



Fill in the missing area Ω such that

- boundary conditions matched $f|_{d\Omega} = f^*|_{d\Omega}$
- close to the guide (g) as possible: $\min_f \int_\Omega |\nabla f - \nabla g|^2$
- Solved by the Poisson Eqn: $\nabla^2 f = \nabla^2 g$

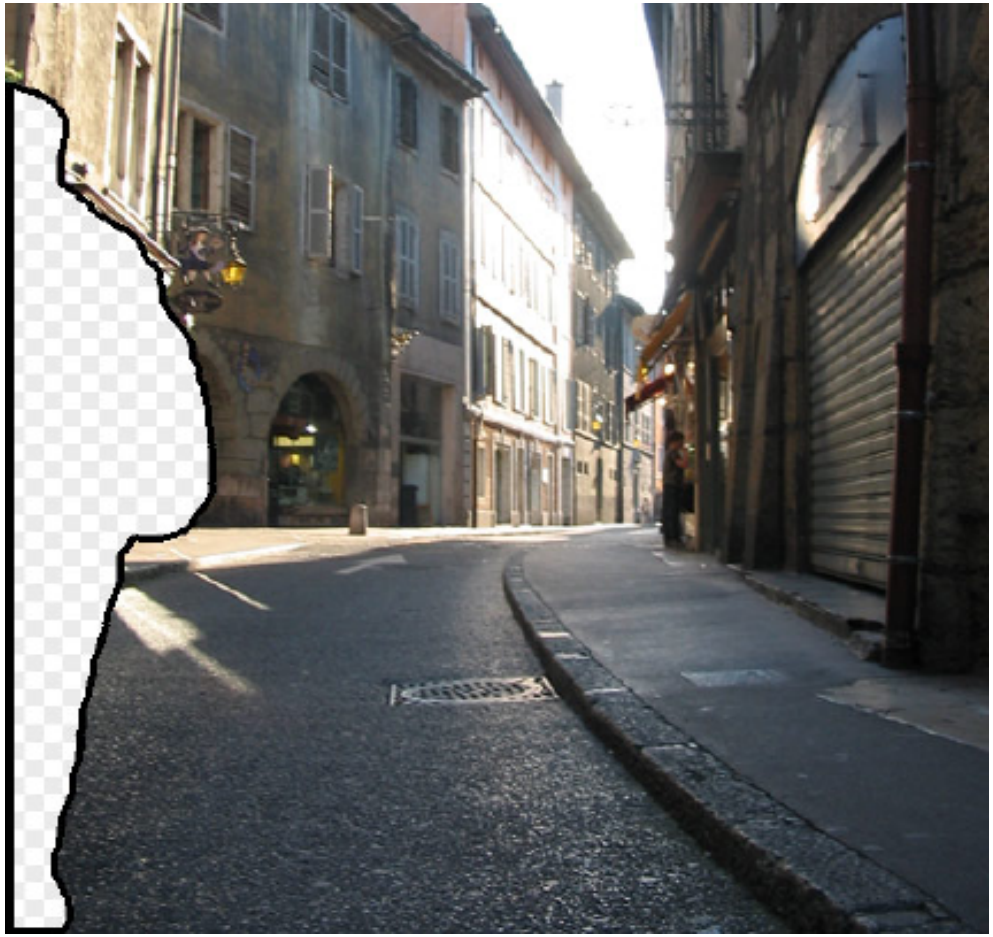
Additional Trick - Seam Finding

Find the optimal boundary ("seam finding") for blending

- minimise gradient differences round the boundary
- encourage the seam to be tight to the missing area
- optimise energy using graph-cut algorithm.

Hardware: Cluster of 15 machines, full algorithm takes $\approx$ 10mins.

Blending failure: textures not matched

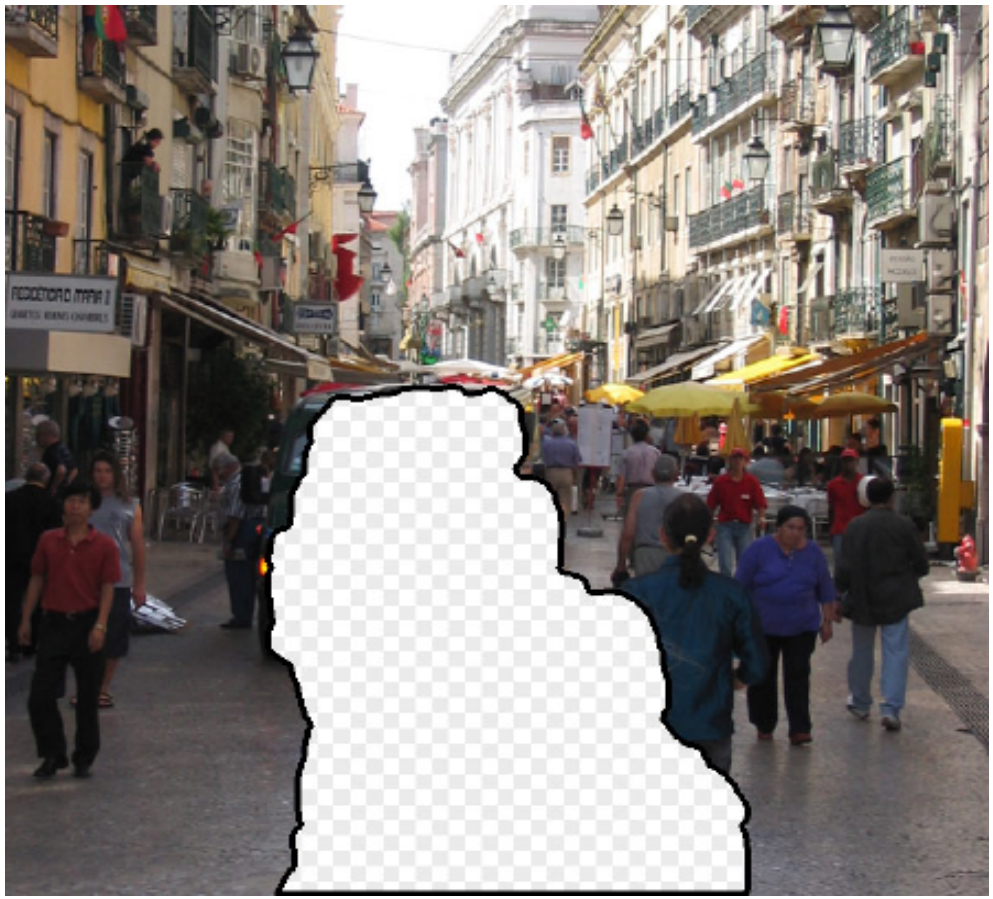


Blending failure: textures not matched

Scene matching failure

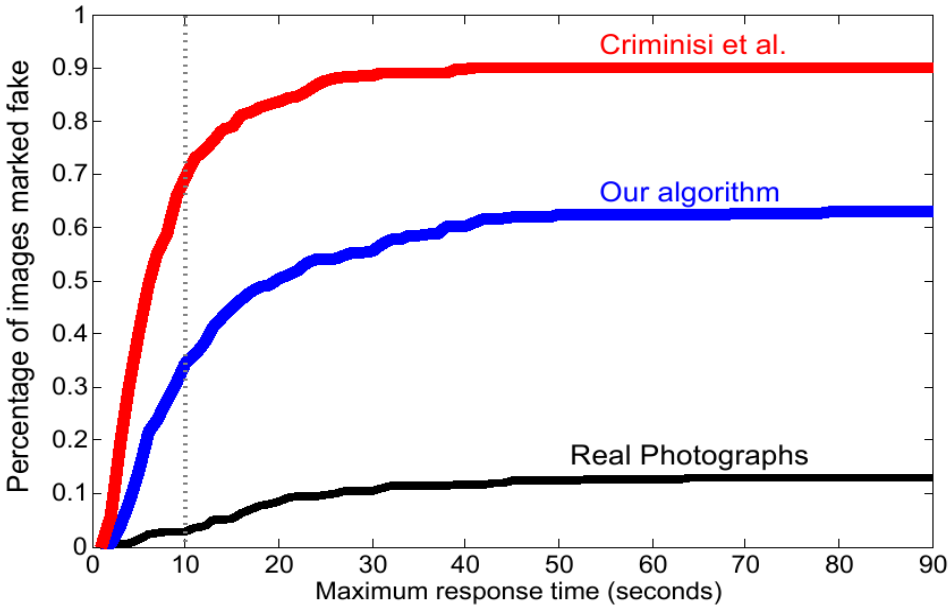
Scene matching failure

High level semantics violated



High level semantics violated

Psychophysical Benchmark



Summary

- If you've got millions of images, there are a number that share similarly shaped areas with similar semantic content.

- Moving toward the realm where you have
  - enough data to stop being Bayesian
  - too much data to be Bayesian