Master in
Computer Vision
Barcelona

Module: Optimization and Inference Techniques in CV
Project: Scene Completion
Teachers: Vadim Fedorov, J.F. Garamendi

Scene Completion Using Millions of Photographs

Aim

- To set-up a graphical model, and to do the inference using the UGM library.

Scene Completion Using Millions of Photographs

Proposed Algorithm

Scene Completion Using Millions of Photographs

James Hays Alexei A. Efros
Carnegie Mellon University

Original Image Input Scene Matches Output

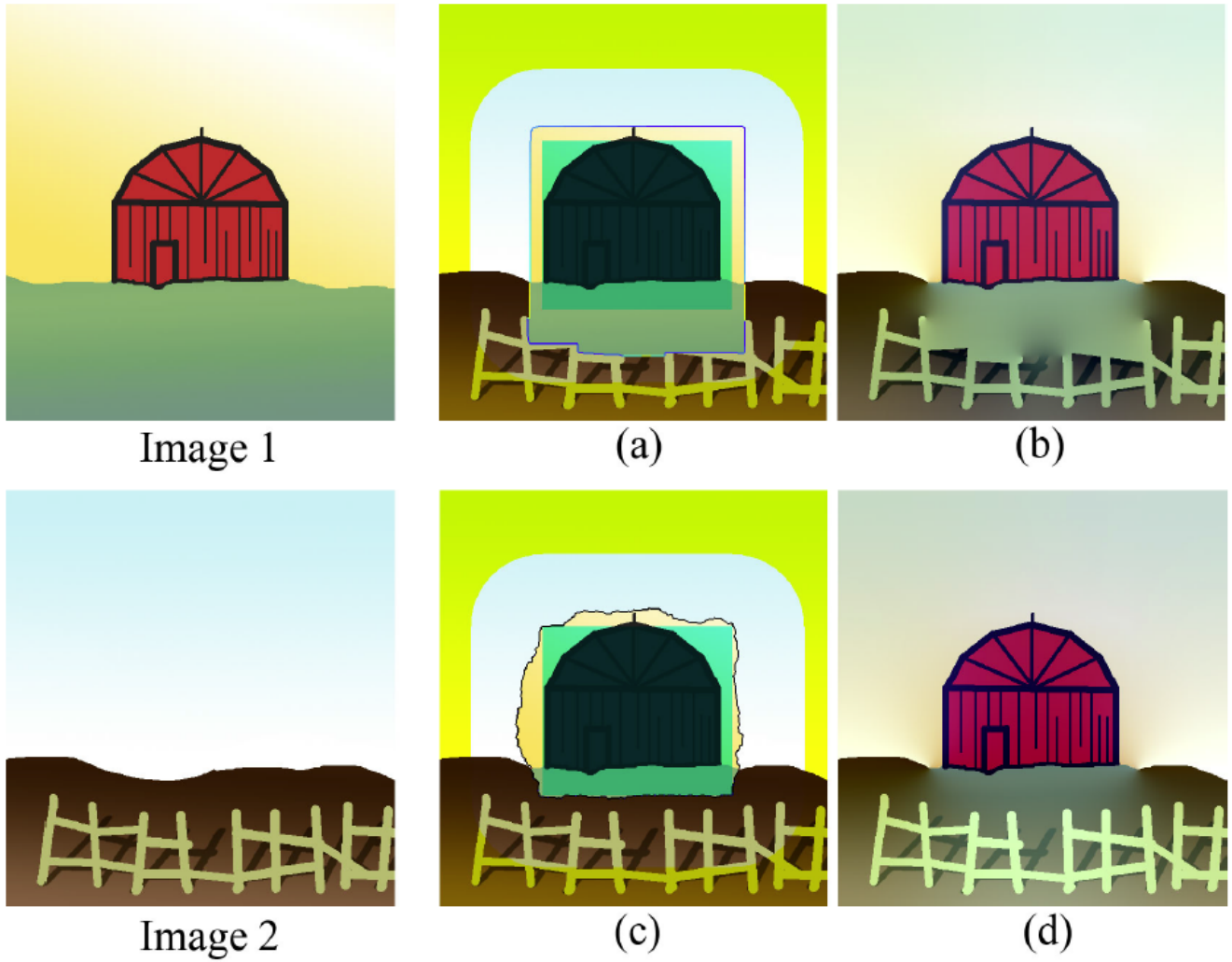
Figure 1: Given an input image with a missing region, we use matching scenes from a large collection of photographs to complete the image.

Given a target image to complete

- From a huge database, finding the most similar images to the target image using GIST descriptor.
- From every found image, find the most similar region to the hole of the target.
- **Authomatically fine tune of the editing mask. ← TO IMPLEMENT using graphical models.**
- Copy the region.
- Perform the Poisson Editing correction.
- Choose the best result (visual inspection).

Authomatically fine tune of the editing mask.

Described in "Scene Completion Using Millions of Photographs"



Authomatically fine tune of the editing mask.

Described in "Scene Completion Using Millions of Photographs"

$$C(L) = \sum_{p} C_d(p, L(p)) + \sum_{p,q} C_i(p, q, L(p), L(q)) \qquad (1)$$

Where $L(p)$ is the label

- $L(p)$= *patch*, Pixel from the source (it is the patch or hole).
- $L(p)$= *exist*, pixel from the incomplete image(target).

C_d are the unary potentials.

- if p is outside the area covered by the source, $C_d(p, patch) = \infty$ (different image sizes).
- if $p \in mask \rightarrow L(p) = patch$
  - $C_d(p, exist) = \infty$, $C_d(p, patch) = 0$, probabilities?
- if $p \notin mask \rightarrow L(p) = exist$
  - $C_d(p, exist) = 0$, $C_d(p, patch) = (k \cdot dist(p, mask))^3$, probabillities?

Authomatically fine tune of the editing mask.

$$C(L) = \sum_{p} C_d(p, L(p)) + \sum_{p,q} C_i(p, q, L(p), L(q)) \qquad (1)$$

Where $L(p)$ is the label
C_i are the binary potentials.

- if pixels p and q are not adjacent, the cost is zero.
- if $L(p) == L(q)$ The cost is zero
- $\nabla ssd(p, q)$ (see the paper)

Scene Completion Using Millions of Photographs

Considerations about the given code

- The images are preprocessed (using `computeGIST.m` function) and stored in the `descriptorsGIST.mat` file.
- Use the given poisson editing function (the data structure of the paremeters is slight different).

Scene Completion Using Millions of Photographs

Code to Complete

TO DO 1: How many number of states for the hidden variable?

TO DO 2: Complete the function `CreateGridUGMModel.m`
Hint: remember the guided exercise in class.

TO DO 3: Build the unary potentials (or factors) of each node based on the euclidean distance of the pixel to the boundary of the hole, see equation (2). If the node corresponds to a pixel inside the hole, then the probability of belonging to the mask is 1. Otherwise it should be 0.

TO DO 4: Build the edge potentials as $\nabla(diff(p, q))$.
Hint: the image is a color image, so every pixel is a vector function but diff(p,q) is a scalar function, so diff(p,q) should be a norm. Use L2-norm.

TO DO 5: Call inference algorithms to solve the MAP problem

Deliverable: Mandatory

Mandatory means if there any point that it is not done, then the weekly task will FAIL.

- Complete the given files.
  - `CreateGridUGMModel.m`
  - `fill_image.m`
- Try several values for the size of the area around the mask used to compare (`mask_extended`).

Maximum: 10 points. The evaluation will depend on the document and the code.

Deliverable: Optional

- Download your own dataset
  - In the material you could find how to build scripts for downloading images.
  - More information on the flickr API.
  - You should compute the GIST descriptors for the images. There are code on the material ('`computeGIST.m`').
  - **WARNING!!! Not everything on Internet is free**, take care about the image license. You can download the images with a specific license.

CODE

- You should to install the UGM library (Class exercise with Oriol and Joan)
- 'images' folder is a small data set with precomputed Gist descriptors saved on 'computeGist.m' file.

Some results

Some results

Some results

Some results

Some results

Some results

Some results

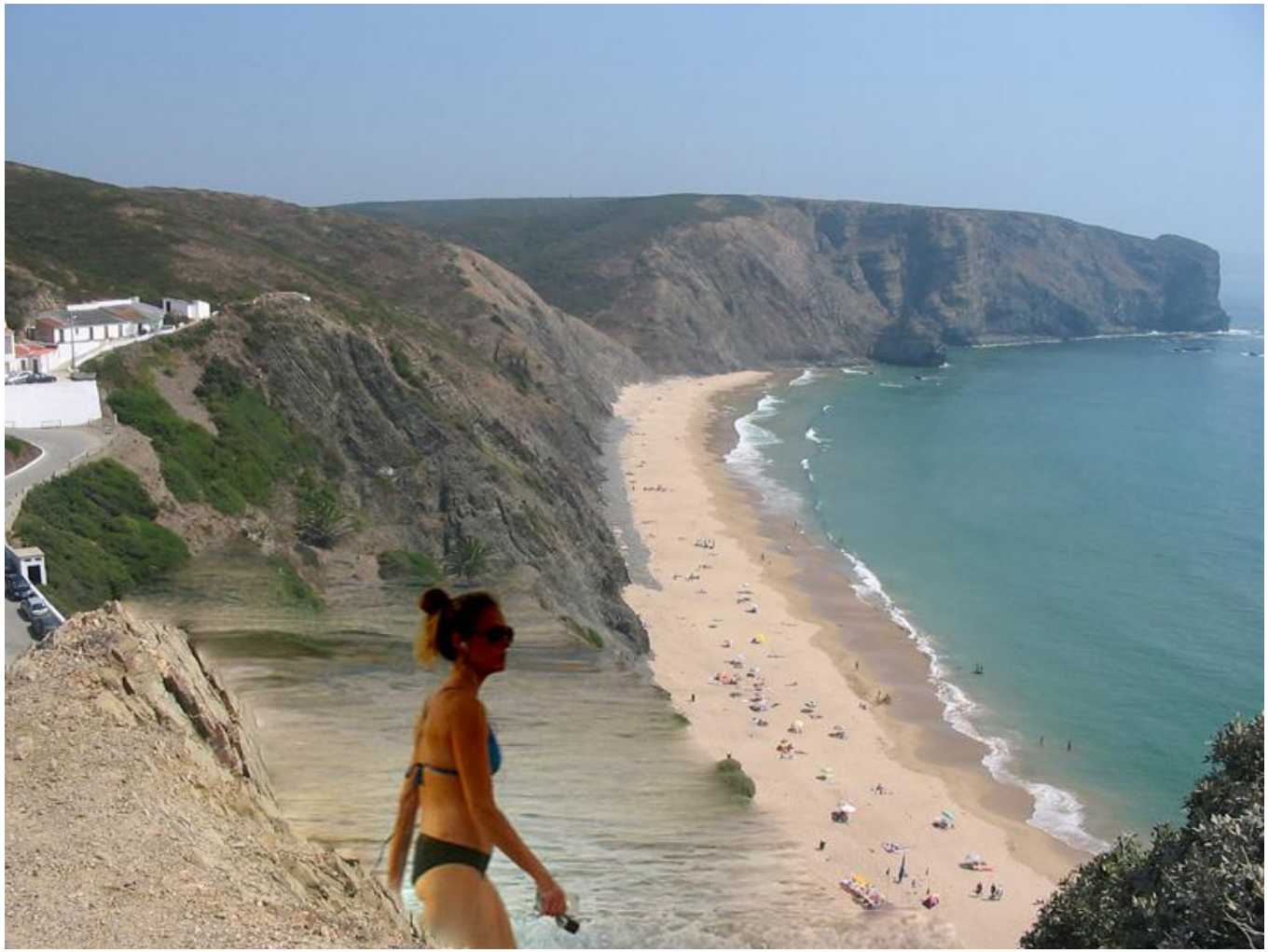
Some results

Some results

Some results

Some results

Some results



Some results

Final Presentation

Oral presentation of the project putting in relation all the stages of the project.

- 10 Minutes presentation: **NO MORE!!!!**
- English.
- For each group, I will choose who begins, who conitnues and who ends.
- Put in relation all the stages of the project.

Deliver of the presentation the same day at 12:00 hours.

Optional parts (all of them)

- If the optional is related with the work flow of the project, you should talk about it on the presentation.
- If not, Do not break the final presentation coherence, you should come to our office and tell us what did you do.